DEVELOPMENT OF NEW APPROACHES FOR THE SYNTHESIS AND DECODING OF ONE-BEAD ONE-COMPOUND CYCLIC PEPTIDE LIBRARIES

Thèse

Xinxia Liang

Doctorat en sciences pharmaceutiques
Philosophiae doctor (Ph.D.)

Québec, Canada

DEVELOPMENT OF NEW APPROACHES FOR THE SYNTHESIS AND DECODING OF ONE-BEAD ONE-COMPOUND CYCLIC PEPTIDE LIBRARIES

Thèse

Xinxia Liang

Sous la direction de :

Éric Biron, directeur de recherche

RÉSUMÉ

La plupart des processus cellulaires et biologiques reposent, à un certain niveau, sur des interactions protéine-protéine (IPP). Leur manipulation avec des composés chimiques démontre un grand potentiel pour la découverte de nouveaux médicaments. Malgré la demande toujours croissante en molécules capables d'interrompre sélectivement des IPP, le développement d'inhibiteurs d'IPP est fortement limité par la grande taille de la surface d'interaction. En considérant la nature de cette surface, la capacité à mimer des structures secondaires de protéines est très importante pour lier une protéine et inhiber une IPP. Avec leurs grandes capacités peptidomimétiques et leurs propriétés pharmacologiques intéressantes, les peptides cycliques sont des prototypes moléculaires de choix pour découvrir des ligands de protéines et développer de nouveaux inhibiteurs d'IPP. Afin d'exploiter pleinement la grande diversité accessible avec les peptides cycliques, l'approche combinatoire «*one-bead-one-compound*» (OBOC) est l'approche la plus accessible et puissante. Cependant, l'utilisation des peptides cycliques dans les chimiothèques OBOC est limitée par les difficultés à séquencer les composés actifs après le criblage. Sans amine libre en N-terminal, la dégradation d'Edman et la spectrométrie de masse en tandem (MS/MS) ne peuvent pas être utilisées.

À cet égard, nous avons développé de nouvelles approches par ouverture de cycle pour préparer et décoder des chimiothèques OBOC de peptides cycliques. Notre stratégie était d'introduire un résidu sensible dans le macrocycle et comme ancrage pour permettre la linéarisation des peptides et leur largage des billes pour le séquençage par MS/MS. Tout d'abord, des résidus sensibles aux nucléophiles, aux ultraviolets ou au bromure de cyanogène ont été introduits dans un peptide cyclique et leurs rendements de clivage évalués. Ensuite, les résidus les plus prometteurs ont été utilisés dans la conception et le développement d'approches en tandem ouverture de cycle / clivage pour le décodage de chimiothèques OBOC de peptides cycliques. Dans la première approche, une méthionine a été introduite dans le macrocycle comme ancrage pour simultanément permettre l'ouverture du cycle et le clivage des billes par traitement au bromure de cyanogène. Dans la seconde approche, un résidu photosensible a été utilisé dans le macrocycle comme ancrage pour permettre l'ouverture du cycle et le clivage suite à une irradiation aux ultraviolets. Le

peptide linéaire généré par ces approches peut alors être efficacement séquencé par MS/MS. Enfin, une chimiothèque OBOC a été préparée et criblée la protéine HIV-1 Nef pour identifier des ligands sélectifs.

Le développement de ces méthodologies permttra l'utilisation de composés macrocycliques dans les chimiothèques OBOC et constitue une contribution importante en chimie médicinale pour la découverte de ligands de protéines et le développement d'inhibiteurs d'IPP.

ABSTRACT

A great number of cellular and biological processes depend, at some level, on protein-protein interactions (PPI). Their manipulation with chemical compounds has provided a great potential for the discovery of new drugs. Despite the increasing demand for molecules able to interrupt specific PPI, the development of small PPI inhibitors is beset by a number of challenges such as the large size of the interaction interface. Based on the interface's nature, the ability to mimic protein secondary structures is very important to bind a protein and inhibit PPI. With their interesting peptidomimetic abilities and pharmacological properties, cyclic peptides are very promising templates to discover protein ligands and development new PPI inhibitors. To fully exploit the great diversity accessible with cyclic peptides, the one-bead-one-compound (OBOC) combinatorial method is certainly the most accessible and powerful approach. Unfortunately, the use of cyclic peptides in OBOC libraries is limited by difficulties in sequencing hit compounds after the screening. Lacking a free N-terminal amine, Edman degradation cannot be used on cyclic peptides and complicated fragmentation patterns are obtained by tandem mass spectrometry (MS/MS).

In this regard we have designed and developed new convenient ring-opening approaches to prepare and decode OBOC cyclic peptide libraries. Our strategy was to introduce a cleavable residue in the macrocycle and as a linker to allow linearization of peptides and their release from the beads for sequencing by MS/MS. First, amino acid residues sensible to nucleophiles, ultraviolet irradiation or cyanogens bromide were introduced in a model cyclic peptide. Afterward, the most promising residues were used to design and develop tandem ring-opening/cleavage approaches to decode OBOC cyclic peptide libraries. In the first approach a methionine residue was introduced in the macrocycle and as a linker to allow a simultaneous ring-opening and cleavage from the beads upon treatment with cyanogens bromide. In the second approach, a photosensitive residue was used in the macrocycle and as a linker for a dual ring-opening/cleavage upon UV irradiation. The linear peptide generated by these approaches can be efficiently sequenced by tandem mass spectrometry. Finally, an OBOC library has been prepared and screened against the HIV-1 Nef protein to identify selective ligands.

The development of these methodologies will prompt the use of macrocyclic compounds in OBOC libraries and be an important contribution in medicinal chemistry for the discovery of protein ligands and the development of PPI inhibitors.

TABLE OF CONTENTS

RÉSUMÉ III
ABSTRACT V
TABLE OF CONTENTS VI
LIST OF TABLES IX
LIST OF FIGURES X
ABBREVIATION XV
ACKNOWLEDGMENT XX
Chapter 1 Introduction 1
1.1 Protein-Protein Interactions 2
1.1.1 Protein-Protein Interactions Interface and Hot Spots 3
1.1.2 Peptides as Protein-Protein Interactions Inhibitors 6
1.1.3 Cyclic Peptides 8
1.2 Combinatorial Chemistry Approaches to Prepare Peptide Libraries 9
1.2.1 Biological Libraries Method 13
1.2.2 Spatially Addressable Parallel Library Method 15
1.2.3 Synthetic Libraries Requiring Deconvolution 17
1.2.4 Synthetic Library Method Using Affinity Chromatography Selection 19
1.2.5 One-bead One-compound Library Method 19
1.3 Screening Strategies for One-Bead-One-Compound Libraries 21
1.3.1 Colorimetric Assays 23
1.3.2 Fluorescent Assays 24
1.3.3 Complex Object Parametric Analyzer and Sorter (COPAS) 25
1.3.4 Magnetic Beads 26
1.4 Strategies to Decode One-Bead-One Compound Cyclic Peptide Libraries 27
1.4.1 The Ladder Synthesis Approach 27
1.4.2 One-Bead-Two-Compound (OBTC) Method 29
1.4.3 Ring-Opening Approach 32
Chapter 2 Working Hypothesis and Objectives 39
2.1 Working Hypothesis 40
2.2 Objectives 41
2.3 Research Plan 42

Chapter 3 Evaluation of Different Cleavable Residues and Linkers for Ring-Opening/ Cleavage Approaches 43

3.1 Selection of the Ring-Opening Residues and Linkers 44

3.2 Model Cyclic Peptides Synthesis and Ring-Opening 45

3.3 Evaluation of the Ring-Opening and Cleavage Combinations 46

3.4 Conclusion 59

3.5 Experimental Section 60

Chapter 4 Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One Compound Libraries 62

Forward 63

Résumé 64

Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One-Compound Libraries 65

Supporting Information (SI) 78

Chapter 5 A Dual Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of One-Bead-One-Compound Cyclic Peptide Libraries 99

Forward 100

Résumé 101

A Dual Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of One-Bead-One-Compound Cyclic Peptide Libraries 102

Supporting Information (SI) 114

Chapter 6 Synthesis and Screening of One-Bead-One-Compound Combinatorial Peptide Libraries for the Development of HIV-1 Nef Protein Inhibitors. 143

6.1 Introduction 144

6.2 Results 145

6.2.1 Preparation of OBOC library 145

6.2.2 OBOC Library Screening and Sequence Determination 146

6.2.3 Binding Assay of Selected Peptides by Nanoporous Optical Interferometry 149

6.3 Conclusion 150

6.4 Experimental Section 151

6.4.1 Materials and Equipment 151

6.4.2 Peptides Synthesis 152

6.4.3 On-bead Library Screening 154

6.4.4 Binding Capacity Measurement 155

Chapter 7 Discussion and conclusion 156
7.1 General Discussion 157
7.2 Conclusion 161
REFERENCE 162

LIST OF TABLES

Chapter 1

Table 1. Milestones in early combinatorial chemistry[55] 10

Table 2. Combinatorial peptide library methods 12

Chapter 3

Table 1. Ring-opening/cleavage Approaches 45

Table 2. Calculated $[M+H]^+$ (Da) of model peptides at different stage 46

Chapter 6

Table 1. Sequenced peptides from positive beads 148

Table 2. Kinetic characterization of selected peptides by NPOI 150

LIST OF FIGURES

Chapter 1

Figure 1. Regulation of PPI involved in apoptosis signalling pathway.[5] 2

Figure 2. Surface of interaction between two proteins.[18] 3

Figure 3. Cellular functions of the pro-apoptotic p53 protein. 4

Figure 4. Interaction surface between a MDM2 binding domain and a peptidic fragment from p53. 5

Figure 5. Molecular inhibitors of the p53-MDM2 protein-protein interaction. 6

Figure 6. Structures of peptide-based marcocycles used as therapeutic agents. 7

Figure 7. Strategies to prepare cyclic peptides. 8

Figure 8. Synthetic approaches used in combinatorial chemistry: the parallel synthesis method (left) and the split-and-pool synthesis method (right). 11

Figure 9. Peptides displayed on M13 phage surface as fusion to the amino terminus of (A) pVIII or (B) pIII; (C) affinity-based selection procedure adapted in phage display technology.[75] 14

Figure 10. Techniques based on the spatially addressable parallel library strategy. (A) Multi-pin technology, (B) Tea bag method and (C) Spot synthesis technique. 15

Figure 11. Descriptive model of (A) iterative process and (B) positional scanning deconvolution approaches. In the case of positional scanning, the sequence is reconstructed at the end of the process, whereas in the iterative process the sequence is obtained step by step.[97] 18

Figure 12. The "*split-and-pool*" synthesis method to generate a one-bead-one-compound combinatorial library (A) a number of permutations for random peptide libraries (P, E, and T are building blocks in this case amino acids).(B) the generated diversity.[102] 20

Figure 13. On-bead screening of OBOC libraries and hit identification by MS/MS. 21

Figure 14. Protein-labelling strategies and sorting methods for the screening of OBOC libraries and selection of positive beads. 22

Figure 15. Colorimetric sorting with SA-AP. 23

Figure 16. Schematic of COPAS sorting. 25

Figure 17. Sorting and bead isolation with magnetic beads. 26

Figure 18. The Ladder Synthesis Approach.[136] 28

Figure 19. One-bead-two-compound approach on topological segregated bilayer beads... 30

Figure 20. Partial Edman degradation approach described by Chait et *al.* (h = homoserine lactone).[150] 31

Figure 21. Alkylthioaryl bridge ring-opening strategy described by Lee *et al.*[155] 32

Figure 22. The dual ring-opening/cleavage strategy described by Simpson and Kodadek.[127] 33

Figure 23. Mechanism involved in methionine ring-opening strategy..... 34

Figure 24. The dual ring-opening/cleavage strategy with reverse methionine linker described by Liang et *al*.[156] 34

Figure 25. Homocysteine system one-pot ring-opening/cleavage reaction by CNBr mediated.[157] 35

Figure 26. The cyclic dilactone approach to generate linear peptides.[159]..... 36

Figure 27. Dual ring-opening/cleavage of cyclic depsipeptides with ammonia described by Gurevich-Messina *et al.*[161]..... 37

Chapter 2

Figure 1. Design of a ring-opening approach to prepare and decode OBOC cyclic peptide libraries and its use in screening and ligands identification. 41

Chapter 3

Figure 1. Combination of methionine as ring-opening residue and ANP as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H]^+$ 895.50 Da molecular ion. 47

Figure 2. Combination of methionine as ring-opening residue and HMBA as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H+17]^+$ 912.52 Da molecular ion. 49

Figure 3. Combination of methionine as ring-opening residue at the N-terminal position and HMBA as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H]^+$ 895.50 Da molecular ion. 50

Figure 4. Combination of methionine as ring-opening residue at the N-terminal position and spacer sequence added to the HMBA linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the aminolyzed specie $[M+17+H]^+$ 1352.85 Da. (Spacer = LBBKG) ... 52

Figure 5. Combination of methionine as ring-opening residue at the N-terminal position and spacer sequence added to the ANP linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the molecular ion $[M+H]^+$ 1335.83 Da. (Spacer = LBBKG) 53

Figure 6. Combination of ANP as ring-opening residue at the N-terminal position and Met as linker with MALDI-TOF MS spectrum. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead. (Spacer = LBBK) 55

Figure 7. Combination of ANP as ring-opening residue at the N-terminal position and spacer sequence added to the HMBA linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the molecular ion $[M-H_2O+H]^+$ 1408.83 Da. (Spacer = LBBKG) 56

Figure 8. Use of ANP residue in the macrocyle and linker for dual ring-opening and cleavage with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead. C) MS/MS spectra of the dehydrated molecular ion $[M-H_2O+H]^+$ 1408.84 Da and D) MeOH adduct molecular ion $[M+14+H]^+$ 1440.84 Da. (Spacer = LBBKG) 58

Chapter 4

Figure 1. Incorporation of a Met residue in the macrocycle at different positions and the corresponding linear peptide generated after ring reopening and cleavage from the resin. . 69

Figure 2. Design and synthesis of cyclic peptides for the tandem ring-opening/cleavage approach. Reagents and conditions: (a) succinic anhydride, DIPEA, DMF; (b) H-Met-OFm, HATU, DIPEA, DMF; (c) 20% piperidine/DMF; (d) Fmoc-Lys-OAll, HATU, DIPEA, DMF; (e) standard Fmoc solid-phase peptide chemistry with HCTU; (f) $Pd(PPh_3)_4$, $PhSiH_3$, CH_2Cl_2; (g) PyBOP, HOBt, DIPEA, DMF; (h) TFA/H_2O/TIS (95:2.5:2.5); (i) CNBr, CH_3CN/AcOH/H_2O (5:4:1). 71

Figure 3. HPLC profiles (λ = 220 nm) and ESI-MS spectra of crude model peptides. (A) cyclo[AYKPFNMK(M-succinamide)] **5a** and (B) H_2N-KAYKPFNh* **6a** after tandem ring-opening/cleavage from TG resin. (h* = homoserine lactone) 72

Figure 4. MALDI MS and MS/MS spectra of peptides after tandem ring-opening/cleavage on a single bead. (A) H_2N-KAYKPFNh* **6a**, MS/MS for precursor ion *m/z* 950.59. (B) H_2N-KGYGSKh* released from a bead randomly selected from the OBOC cyclic heptapeptide library, MS/MS for precursor ion *m/z* 722.37. (h* = homoserine lactone)..... 74

Chapter 5

Scheme 1. Synthetic Route to Cyclic Peptides with ANP Residue and Ring-Opening/Cleavage Reaction……118

Figure 1. HPLC and ESI-MS profiles of crude products showing cyclization and ring-opening. (A) cyclo[ANP-LGYGKFE]-NH2 **2a**; and (B) 3-oxo-3-(2-nitrosophenyl)-propionyl-LGYGKFQ-NH2 **3a**. (C) Proposed structures for the dehydrated product **3a*** and its adduct **3a'** observed during ESI-MS analysis. 106

Figure 2. MALDI-TOF MS and MS/MS spectra obtained after dual ring opening/cleavage on a single bead of cyclic peptide **4a**. (A) MS of the crude product; (B) MS/MS of the dehydrated specie **3a*** (968.4 Da) $[\mathbf{3a}\text{-}18+H]^+$. 108

Figure 3. MS and MS/MS spectra of ANP*-HFSKGQLBBKG-NH_2 after dual ring opening/cleavage on a randomly selected bead from the OBOC cyclic peptide library. (A) MS; (B) MS/MS for precursor ion *m/z* 1299.5. (B = β-alanine). 109

Chapter 6

Figure 1. A) Structure of an HIV virion particle; **B)** Structure of the HIV-1 Nef protein; **C)** Sequential steps governing the downmodulation of MHC class I molecules and CD4. (Left) Nef accelerates the endocytosis of MHC I molecules (**a**). Nef activates PI3K (**b**). ARF6 becomes activated (**c**). Together with Nef, ARF6 mediates the internalization of MHC I molecules (**d**). The latter are retrieved to the TGN, where they remain trapped (**e**). (Right) The two steps of Nef-induced CD4 downmodulation. Nef connects the cytoplasmic tail of CD4 triggering rapid endocytosis of the CD4 receptor (**a**). In the early endosome, Nef interacts with the COPI coatomer, which targets CD4 for lysosomal degradation (**b**).[169]. 144

Figure 2. Solid-phase synthesis of the OBOC peptide library by the split & pool method. 146

Figure 3. A) First screening round with Dynabeads® M-280 Streptavidin; B) Second screening round with streptavidin-conjugated alkaline phosphatase and BCIP..... 147

Figure 4. A) MALDI-TOF MS and B) MS/MS spectrum of selected peptide NO.8 148

Figure 5. A) Artist's rendering of nano-porous biosensor principle; B) Schematic representation of interferogram of a typical nanoporous silicon biosensor experiment.[176] 149

ABBREVIATION

Chemicals and solvents

ACN	Acetonitrile
AcOH	Acetic acid
ANP	3-Amino-3-(2-nitrophenyl)propionic acid
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BSA	Bovine serum albumin
CNBr	Cyanogen bromide
DBU	1,8-Diazabicycloundec-7-ene
DCM	Dichloromethane
DMF	Dimethylformamide
EtOAc	Ethyl acetate
Fmoc	9-Fluorenylmethyloxycarbonyl
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HCTU	1-[Bis(dimethylamino)methylene]-5-chlorbenzotriazolium 3-oxid-hexafluorophosphate
HMBA	4-Hydroxymethylbenzoic acid
HOAt	1-Hydroxy-7-azabenzotriazole
MeOH	Methanol
$MgSO_4$	Magnesium sulfate
NMM	N-Methylmorpholine
PBS	Phosphate-buffered saline
Pip	Piperidine
PIC	Phenylisocyanate
PITC	Phenylisothiocyanate
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

SA-AP	Streptavidin-Alkaline Phosphatase
SDS	Sodium dodecyl sulfate
TBST	Tris-buffered saline and tween 20
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane

Units

℃	Degrees Celsius
cm	Centimeter
cm^{-1}	Reciprocal centimeter
g	Gram
h	Hour
mg	Milligram
min	Minute
mL	Milliliter
µL	Microliter
mmol	Millimole
mol	Mole
m/z	Mass-to-charge ratio
s	Second
nm	Nanometer
ppm	Part per million
µm	Micrometer
mm	Millimeter
Å	Angstrom
equiv	Equivalent
m/v	Mass to volume

v/v	Volume to volume

Analytical methods

^{1}H NMR	Proton nuclear magnetic resonance
^{13}C NMR	Carbon-13 nuclear magnetic resonance
COPAS	Complex object parametric analyzer and sorter
ESI	Electrospray ionization
LC-MS	Liquid chromatography mass spectrometry
MALDI-MS	Matrix assisted laser desorption ionization mass spectrometry
MS	Mass spectrometry
RP-HPLC	Reversed-phase high-performance liquid chromatography
UV	Ultraviolet

Others

ELISA	Enzyme linked immunosorbent assay
FP	Fluorescence polarization
HTS	High-throughput screening
HIV	Human immunodeficiency virus
IC_{50}	Half maximal inhibitory concentration
MDM2	Mouse double minute 2 homolog
MW	Molecular weight
Nef	Negative regulatory factor
NPOI	Nanoporous optical interferometry
OBOC	One bead one compound
OBTC	One bead two compound
PED	Partial Edman degradation
PID	Protein interaction domain

PPI	Protein-protein interaction
PTM	Post translational modification
RT	Room temperature
SPPS	Solid phase peptide synthesis
SPR	Surface plasmon resonance
TGN	Trans-Golgi network

Standard one- or three- letter codes are used for amino acids

It does not matter how slowly you go, so long as you do not stop.

Confucius 551 - 479 BC

ACKNOWLEDGMENT

It would not have been possible to achieve the degree of Doctor of Philosophy (Ph.D.) without unlimited support, assistance, and cooperation from numerous people in the process of my study. I would like to express to my deep and sincere appreciation to all of these people although I know it will have never been enough.

First and foremost, I would like to appreciate my supervisor, Dr. Éric Biron, for providing me the opportunity to pursue my Ph.D. degree in his lab and his constant intellectual guidance, patience, and encouragement throughout these years. It is that he guides me to the field of solid phase synthesis which is a powerful and great potential methodology for life-science research. His impressive knowledge, brilliant ideas, and sincere passion about science direct me to grow as a researcher, in which undoubtedly I will continuously benefit in my further career. With his humour and dynamism, I have been able to work in a pleasant environment.

I would like to express my gratitude to the programme director, Dr. Thérèse Di Paolo, for her understanding of my language barrier and managing my study processes. I wish to thank my committee members for valuable observations and thesis defense. I would also like to appreciate secretary department, Lysanne Tanguay who helped me a lot at beginning of my study, Sylvie Lacasse and Jessica Jean who organized my exams and defence.

I will also take this chance to thank all my great colleagues in CHU de Quebec Research Centre. My lab life would not be going well without their kind help and accompanying. Anick Girard taught me a lot and gave me great help for experiments when I joined in the lab. Marie-Pier Thibeault showed her clean and tidy labbook to make me have the good habit. Simon Vézina-Dawod gave suggestions for NMR analysis and we had lots of fun during working with his humour. François Bédard has been brought a lot of convenience with his skills in repairing various lab equipments. Marie-Claude Trottier, technical assistant, is responsible for the manipulation of NMR and LC-MS. I am really happy and lucky to have known and worked with all lab mates.

I would like to thank all my friends who shared happiness and sorrow with me in Quebec City. Without your accompanying, it will be hard and lonely of both my life and study.

They are Yi Dong, Lei He, Xiaoqiang Wang, Dan Xu, Ruixuan Wang, Xiaoye Sang, Hui Han, Huan Liang, Lucie Carolle Kenmogne, Guy Bertrand Djigoué, and Dossa Richi.

China Scholarship Council which is the most appreciated sponsor provided four-year financial support for my study in Université Laval. I also appreciate the Consortium Québécois sur la découverte du medicament (CQDM) supports me to finalize my thesis.

Finally, I will present my deepest acknowledgement from the bottom of my heart to my family. My parents, they are continuously and unselfishly inspiring me to pursue my dream with their infinite love. My boyfriend, Cang, is accompanying me and making my life to be wonderful. Thank you for your consistent comprehension and encouragements. Without my family support, I would not be where I am now today.

Chapter 1

Introduction

1.1 Protein-Protein Interactions

The cellular function of a vast majority of proteins is performed through physical interactions with other molecules, which are most frequently other proteins. Protein-protein interactions (PPI) occur when two or more proteins interact together to carry out their biological function. Essentially all of the known cellular and biological processes depend, at some level, on PPIs (Figure. 1).[1,2] These interactions play important roles in a wide range of diseases and infections. Therefore, the controlled interference of PPI with chemical compounds has provided tremendous potential for the discovery of novel molecular tools to improve our understanding of biochemical pathways as well as the development of new therapeutic agents.[3,4]

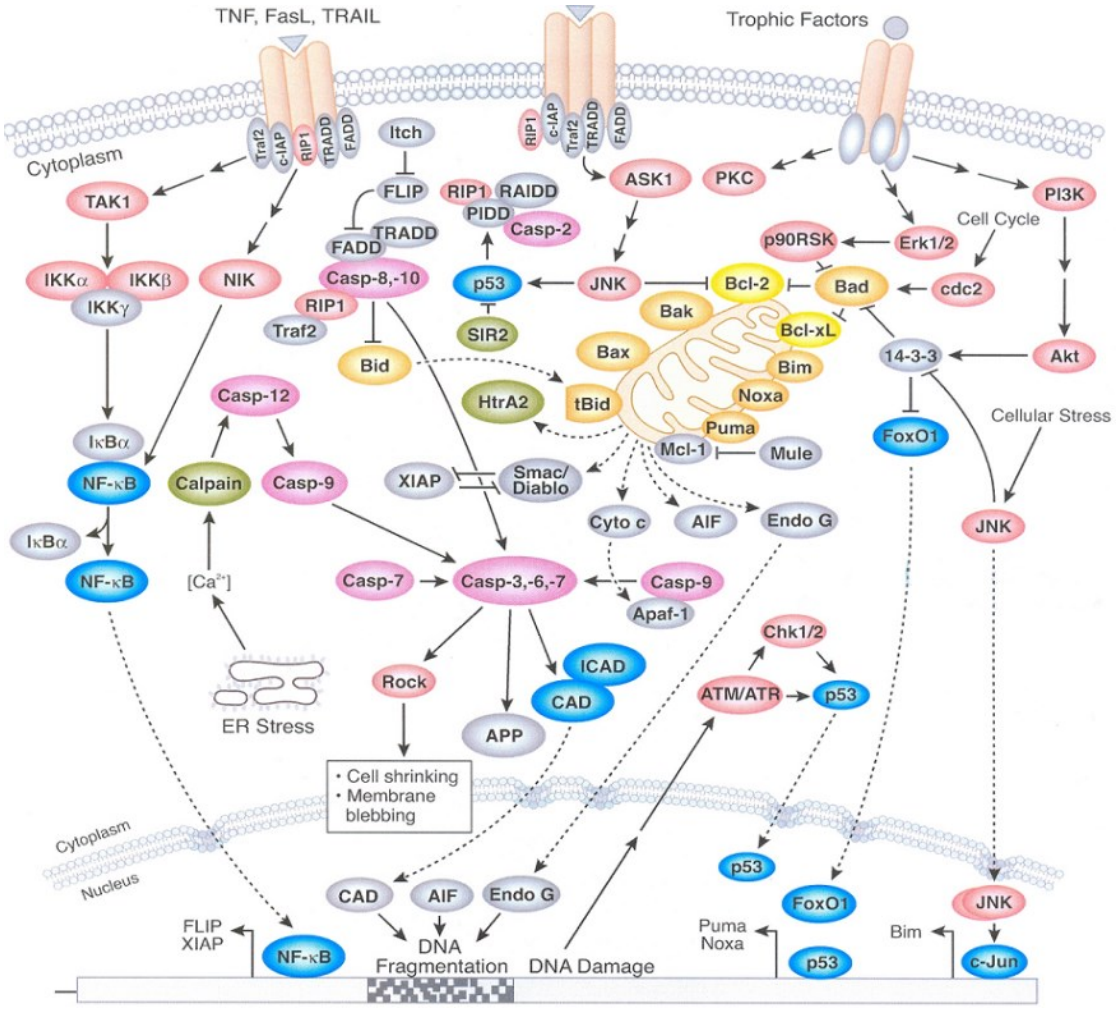


Figure 1. Regulation of PPI involved in apoptosis signalling pathway.[5]

1.1.1 Protein-Protein Interactions Interface and Hot Spots

There is a growing interest for PPI in biotech and pharmaceutical companies and the global market for PPI technology and therapeutics is expected to reach $50 billion in 2015.[6,7] As a result, the demand for molecules able to selectively bind protein interaction domains and interrupt specific PPI is rapidly increasing.[4,8,9] However, the development of small PPI inhibitors is beset by a number of challenges such as the large size of the surface area for specific recognition and the intracellular localization of most PPI (Figure. 2).[3,10–13] The interfaces involved in PPI are large (~ 1500 – 3000 $Å^2$)[14,15] compared to those involving small molecules with proteins (~ 300 – 1000 $Å^2$).[16,17] Therefore the size of the interfaces complicates the structure-activity relationship studies and impedes the discovery and development of effective small molecule inhibitors.

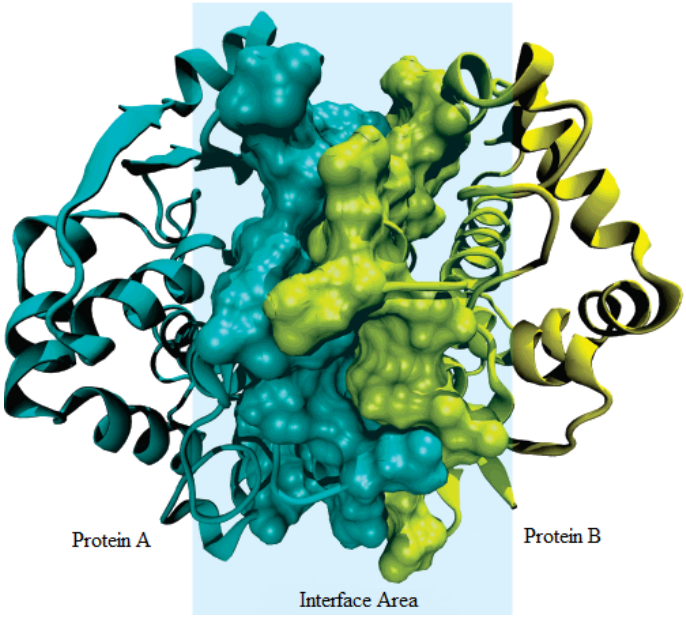


Figure 2. Surface of interaction between two proteins.[18]

Fortunately, different studies have showed that most interfaces involved in PPI contain compact and centralized regions that contribute to high affinity binding and are crucial for the interaction. These small high affinity regions, which are found on both sides of the PPI interface, are called “hot spots”. It has been demonstrated that it is possible to prevent the interaction between two protein partners by blocking a hot spot with a small molecule.[19] These results suggested that it was not necessary to occupy the entire binding surface to

disrupt the PPI and that small molecules may modulate PPI when binding to these "hot spots''.[3,20]

A good example of PPI inhibitors targeting interface hot spots is the development of MDM2-p53 interaction inhibitors. The p53 protein, also known as the p53 tumour suppressor, is a transcription factor composed of 393 amino acids playing a major role in cell cycle, apoptosis and DNA repair via the regulation of key genes expression.[21–24] The MDM2 protein (Mouse Double Minute 2 or HMD2, its equivalent in man) is an E3 ubiquitin ligase involved in p53 downregulation. Composed of 491 amino acids, MDM2 binds to the transcription activation domain of p53 to inhibit its role in healthy cells (Figure 3). Virtually all cancer cells exhibit a malfunction of p53 due to p53 gene mutation or overexpression of the protein MDM2.[25–27] Given the crucial role of the MDM2-p53 interaction in cancer cells survival, the development of inhibitors for this PPI was extremely promising for the treatment of many cancers. Therefore, the strategy was to develop a molecule capable of binding to MDM2 and prevent its interaction with the p53 protein.

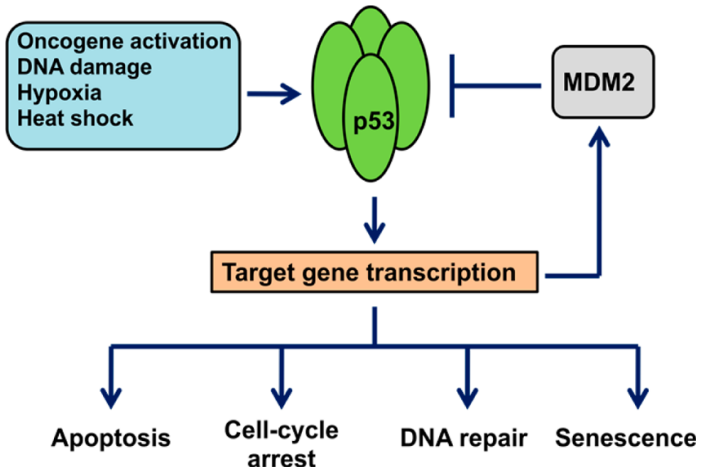


Figure 3. Cellular functions of the pro-apoptotic p53 protein.

Analysis of the MDM2 protein crystal structure with a short 15-residue peptide fragment from p53 transactivation domain (N-terminus) showed the presence of a small well-defined interaction surface. This p53 fragment was then used as a starting point for the development of several inhibitors for this PPI.[28–30] From this structure, researchers observed that side

chains from Phe19, Leu22, Trp23 and Leu26, which are contained in an α helix section of p53, interact with a large hydrophobic pocket in MDM2 (Figure 4).

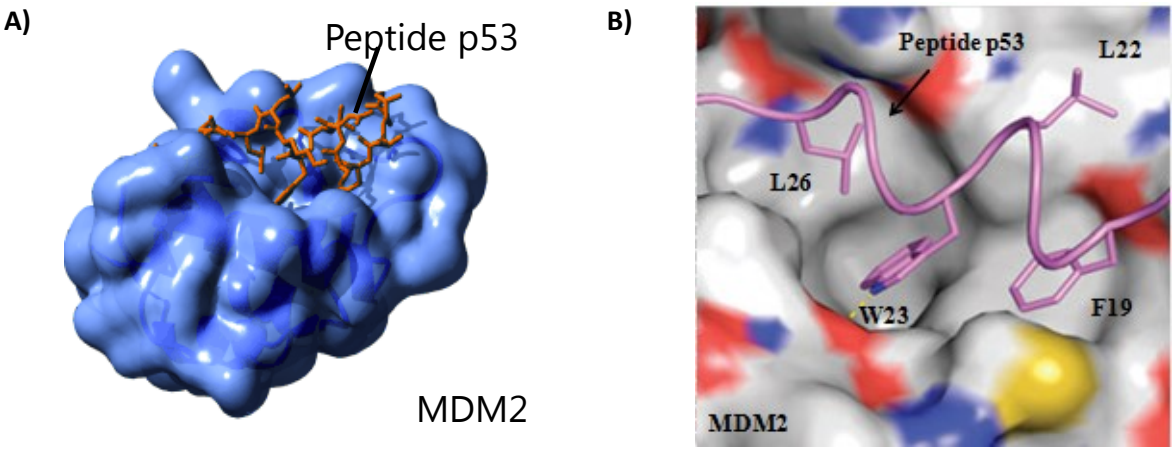


Figure 4. Interaction surface between a MDM2 binding domain and a peptidic fragment from p53.

These observations led to the development of several MDM2-p53 interaction inhibitors (Figure 5). In 2001, Moore and his collaborators discovered chlorofusine **1**, a naturally occurring cyclic peptide isolated during the screening of bacterial extracts.[31] Subsequently, Robinson's group has developed the macrocyclic peptide inhibitor **2** showing an IC_{50} of 140 nM. This macrocycle has a β-hairpin conformation in which the bioactive section is spatially a mimetic of Phe19 and Trp23 from p53 helical section.[32] In 2004, Vassilev's team have discovered the Nutlin, cis-imidazoline derivatives **3** identified by library screening.[33] These molecules were able to efficiently inhibit the formation of the p53-MDM2 complex with IC_{50} between 100 and 300 nM. A bromophenyl group is found in the Trp23 pocket, another bromophenyl group is found in the Leu22 pocket and the ethyl ether chain in the Phe19 pocket. Finally, in 2005 1,4-benzodiazepine-2,5-dione derivatives **4** have been reported as p53-MDM2 interaction inhibitors.[34] Compounds showing IC_{50} ranging from low μM to high nM have been identified by the combined screening of combinatorial chemical and virtual libraries.

Other examples for the successful rational development of small PPI inhibitors include interactions between XIAP and SMAC, HSP90 and GR, co-activators and estrogen alpha-receptor, Bcl-Xl and Bak, TNF-α and TNFRc1.[4,9,35–37]

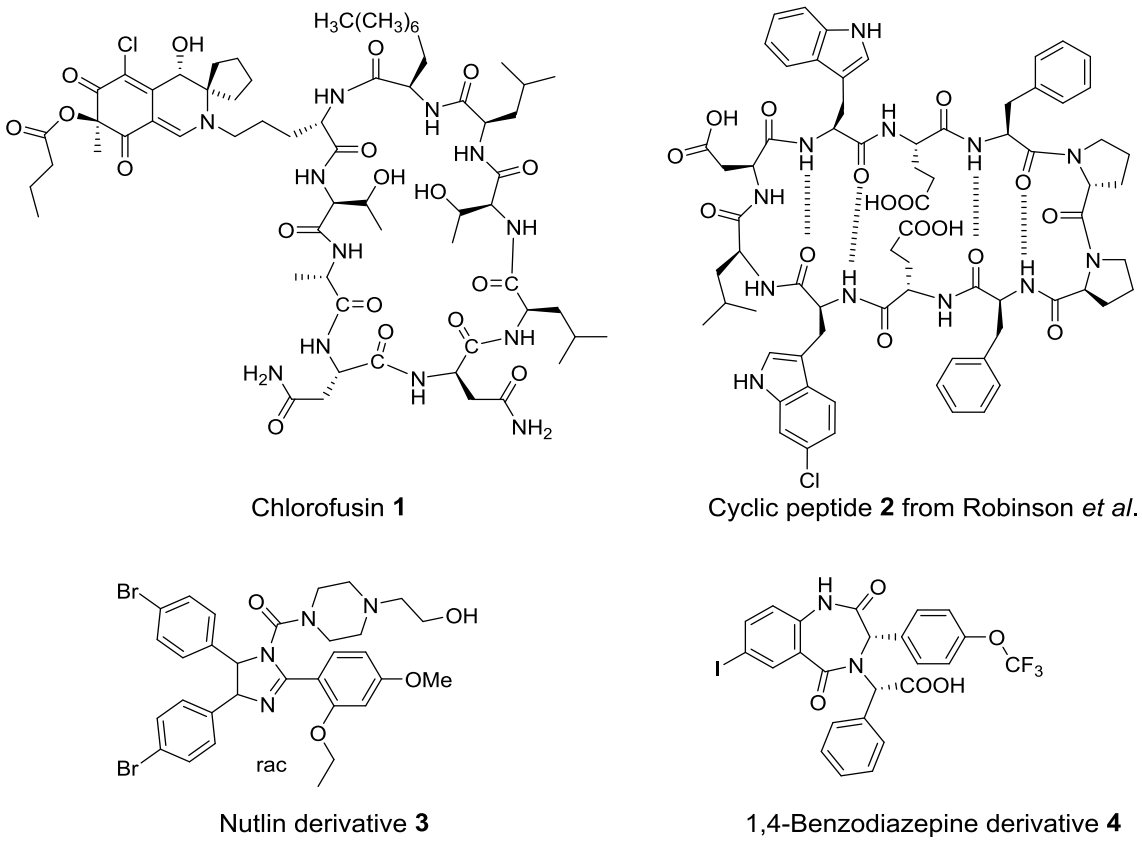


Figure 5. Molecular inhibitors of the p53-MDM2 protein-protein interaction.

1.1.2 Peptides as Protein-Protein Interaction Inhibitors

Considering the interaction surface nature, protein secondary structures play a major role in the molecular recognition and are essential for specific binding to protein interaction domains (PID). Therefore the ability to mimic protein secondary structures is very important in the design of PPI inhibitors and peptides represent a template of choice to mimic these structures and modulate PPI.[20] Peptides are a very interesting class of therapeutics since they show strong activity, high selectivity, low toxicity and few drug-drug interactions.[38,39] However, linear peptides show poor oral bioavailability and are easily degraded by proteases, limiting their use as therapeutic agents. To overcome these drawbacks, peptide cyclization is increasingly used. Compared to small organic molecules (MW < 500 Da), the size of most macrocycles (MW = 500~2000 Da) is large enough to compete with proteins for flat interaction surfaces and are small enough compared to antibody (MW > 50000 Da) to retain many drug properties. Peptide macrocyles are a rich source of biologically active compounds and are widely produced in nature by plants,

bacteria, fungi, marine invertebrates and many other organisms. A wide range of peptide-based macrocycles such as cyclosporine A (immunosuppressant), daptomycin (antibiotic), caspofungin (antifungal), eptifibatide (antiplatelet), octreotide (anticancer), and oxytocin (uterine contractions) are clinically used as therapeutic agents (Figure 6).[40]

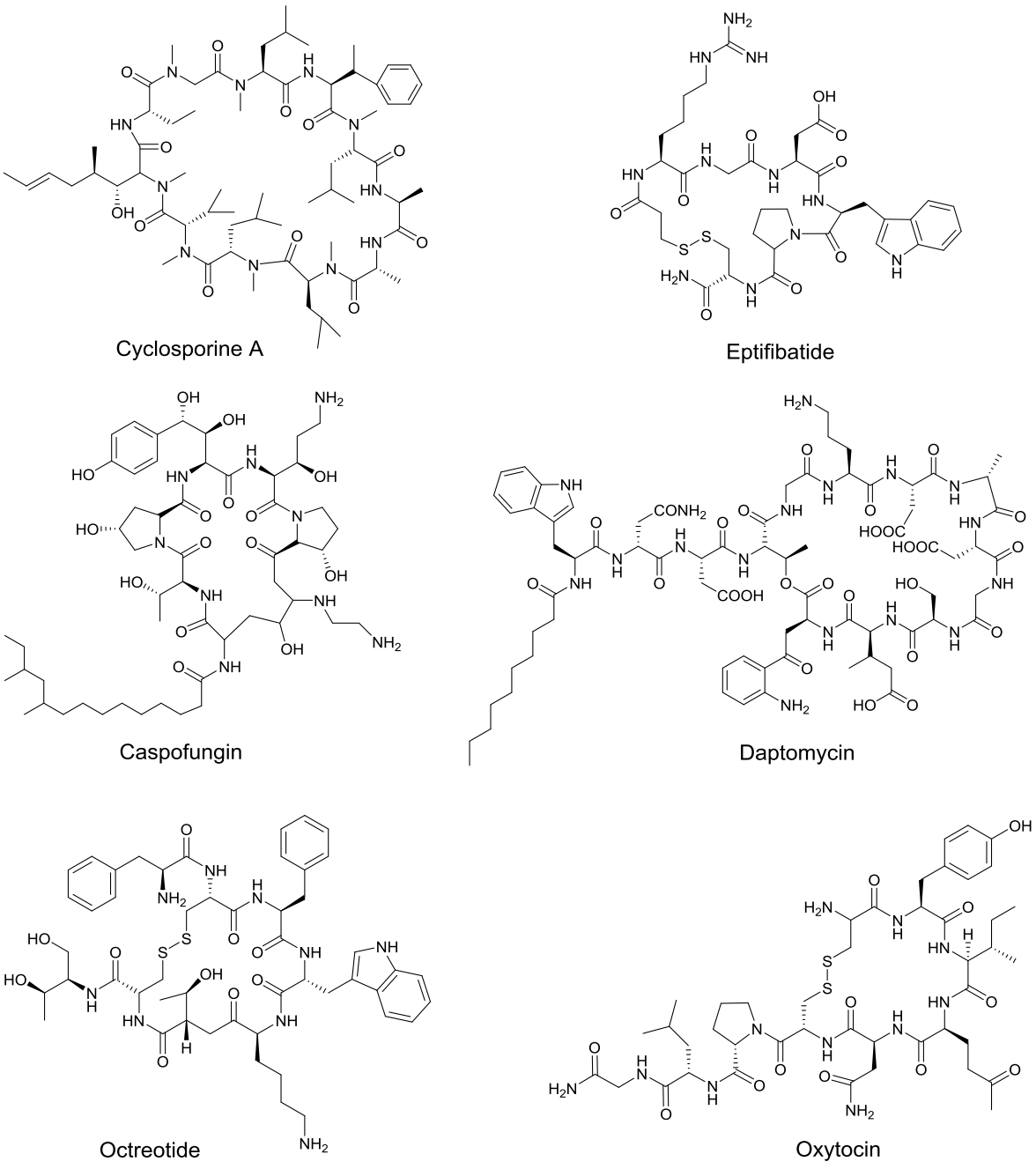


Figure 6. Structures of peptide-based marcocycles used as therapeutic agents.

1.1.3 Cyclic Peptides

With a wide spectrum of activity and a great therapeutic potential, cyclic peptides have gained a lot of interests in drug discovery. Cyclic peptides show many advantages compared to their linear counterparts. With increased structural rigidity, cyclic peptides show greater stability against endo- and exo-proteases[41,42] and the entropic advantages make them tighter-binding and may confer a higher binding specificity for a given macromolecular receptor.[43] Moreover, their conformational analyses in solution by 2D NMR are more precise and reliable;[42,44] and in some cases, they show increased cell permeability.[45]

Based on the type of chemical bonds found in the cycle backbone, two different classes of cyclic peptides have been established, namely homodetic and heterodetic cyclopeptides. For homodetic cyclic peptides, the macrocyle backbone is exclusively formed by amide bonds. In contrast, the backbone of heterodetic cyclic peptides may contain also one or many ester, ether, thioester, thioether or disulfide bonds in addition to the amide bonds. Cyclic peptides can be obtained from linear precursors by four different cyclization approaches: head-to-tail (C-terminus to N-terminus), head-to-side chain, side chain-to-tail, and side chain-to-side chain (Figure 7). In most conditions, the ring-closure reaction is performed by lactamization, lactonization or disulfide bridge formation. For the cyclization step, synthetic yields can vary greatly and will depend on various parameters such as reaction conditions, ring size, specific structure, and even amino acid sequence.

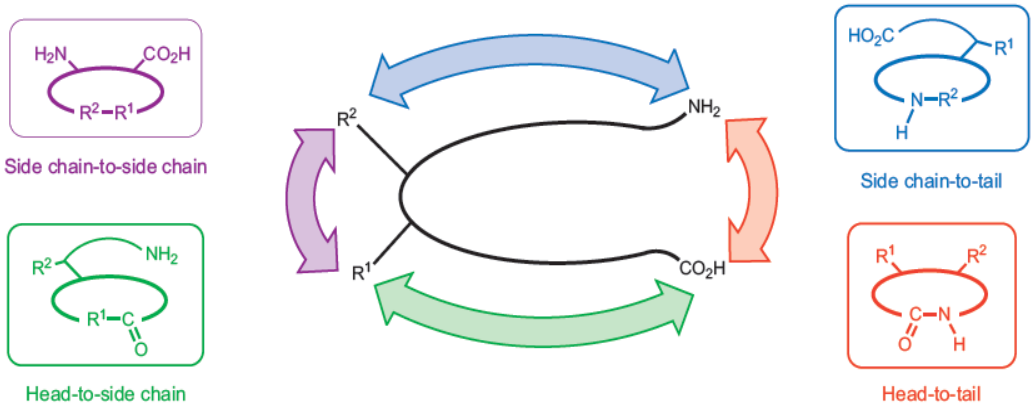


Figure 7. Strategies to prepare cyclic peptides.

Cyclic peptides are very convenient molecular probes in chemical biology and very useful tools in drug discovery and development. Most importantly, cyclic peptides are a class of privileged structures since they have been reported to bind to multiple, unrelated classes of receptor with high affinity.[46–48] The great degree of molecular complexity and diversity that can be accessed by simple changes in their linear sequence, makes the use of cyclic peptide-based scaffolds very appealing in drug design and discovery.[49] Upon cyclization, small changes, *e.g.* in α-carbon stereochemistry, ring size, or constraining residues, can have a dramatic effect on the backbone overall conformation. These conformational effects allow the presentation of side-chains with diverse relative orientations. Their abilities to mimic protein secondary structures such as β- and γ-turns, β-hairpins, β-sheets, α-helices and other helix types have been widely studied and demonstrated.[50–52]

Based on the characteristics described above, cyclic peptides represent a very attractive template to discover protein ligands and develop promising PPI inhibitors. In order to fully exploit the great conformational and functional diversity accessible with cyclic peptides, combinatorial chemistry is certainly the most powerful tool. A major asset of combinatorial cyclic peptide libraries over classic combinatorial libraries, where the scaffold is fixed, is the possibility to generate conformational diversity as well as functional diversity in order to encompass a very large chemical space (diversity).

1.2 Combinatorial Chemistry Approaches to Prepare Peptide Libraries

During the last three decades, the development of combinatorial methodologies has been greatly stimulated by the progress made in biological assays and their increasing testing capacities. To match the growing needs in high-throughput screening (HTS), the preparation of libraries containing a large number of chemical entities was required and combinatorial chemistry represented the most convenient and efficient approach to prepare such libraries. Allowing the simultaneous generation of a large number of synthetic compounds, combinatorial libraries have been used by many research groups to quickly test millions of compounds for their activity on a selected target in a short period of time. This approach is probably the most powerful approach that has occurred in pharmaceutical and life sciences to reduce the time and cost for generating high affinity protein ligands.

A combinatorial library is usually referred as a collection of 1×10^4 to 1×10^{10} different molecules generated by synthetic or biological approaches. The development of combinatorial peptide library technologies can be attributed to Bruce Merrifield who won a Nobel Price for the use of functionalized styrene-divinylbenzene beads as polymer supports in 1963 and his important contribution to the development of solid phase peptide synthesis.[53] This approach has faced a renascence in combinatorial chemistry in the middle 1980s with the synthesis of the first limited peptide library using multi-pin technology in 1984,[54] and has since become a powerful tool that not only facilitates the drug discovery process but also provides important information for the fundamental understanding of molecular recognition. These applications led to the development of a wide variety of combinatorial approaches to prepare peptide and chemical compound libraries (Table 1).

Table 1. Milestones in early combinatorial chemistry[55]

Year	Milestone
1984	Limited peptide library with the multi-pin technology
1985	Limited peptide library using tea-bag method
1986	Iterative approach on solid phase peptide library screening using the multi-pin synthesis
1986-90	Development of polynucleotide library methods
1988	Introduction of the split synthesis method on synthesizing a limited library of solution peptides
1990	Light directed parallel peptide synthesis of a library of 1024 peptides on chip
1990	Successful use of the filamentous phage to display a peptide library
1991	Introduction of the one-bead-one-compound method and successful application of this concept to a huge bead-bound peptide library
1991	Successful application of the interactive approach on a huge solution phase peptide library
1992	Synthesis of a limited benzodiazepine-based small molecule library
1992-93	Development of encoding methods for the one-bead-one-compound non-peptide library

Among the different synthetic approaches to prepare combinatorial libraries, the two most commonly used are the parallel synthesis method and the split-and-pool synthesis method. In parallel synthesis, like in a general organic synthesis, each reaction is performed in separate reaction conditions (Figure 8). The main difference is that many reactions are performed simultaneously in many reactors. To treat many reactions, easy treatment of each reaction is needed, and so solid-phase reaction is the more preferred method. The other option is to use expensive automatic synthesizer. If resins are not used as solid supports, it is possible to synthesize as a chip form by using photolithography of semi-conductor chips.

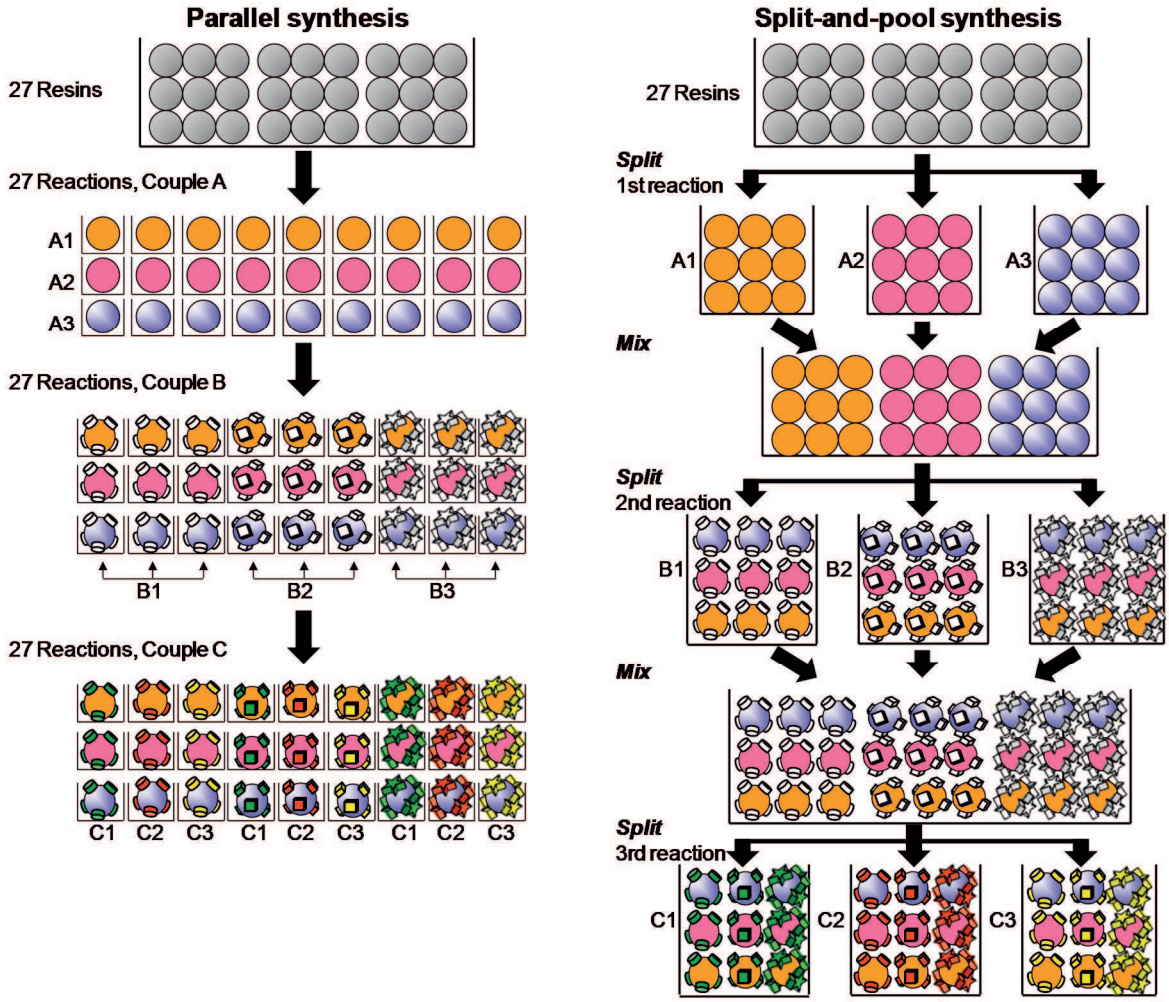


Figure 8. Synthetic approaches used in combinatorial chemistry: the parallel synthesis method (left) and the split-and-pool synthesis method (right).

The split-and pool synthesis method is only used with solid phase reactions (Figure 8). This method is most commonly used in the one-bead-one-compound combinatorial technology and will be described in details in section 1.2.5. The "split-and-pool synthesis" method was first described by Furka *et al.*[56] also developed by *Lam et al.*[57] and Houghten *et al.*.[58] The beads are split and placed into separate reaction vessels. Different building blocks are coupled on and within the beads. After each coupling step, the beads from each reaction vessel are combined and then randomly split again to the different reaction vessels for next coupling step. A library is prepared after repeating those processes several times. For example, if we try to make all the possible oligomers from monomers A, B and C, the total possible number of compounds is 27 ($3 \times 3 \times 3$) and only three coupling steps are required in the case of the split-and pool method. The number of total solid supports should be larger than 27 total compounds to be synthesized so that statistics of split-and-pool process is even. The parallel synthesis approach is more convenient for small size libraries while the split-and-pool approach is more efficient to prepare large size libraries.

Table 2. Combinatorial peptide library methods

Combinatorial peptide library methods	Key points of method
Biological libraries	Phage-display Plasmid Polysome Bacterial display Yeast-display
Spatially addressable parallel libraries	Multi-pin technology Tea bags and NanoKan technologies SPOTs-membrane method Light-directed peptide synthesis on chip
Synthetic libraries requiring deconvolution	Iterative approach Positional scanning Recursive deconvolution Orthogonal partition approach Dual recursive deconvolution
Select library by affinity column	Using affinity chromatography selection
One-bead one-compound libraries	Split-and-pool synthesis

Most combinatorial methods are composed of three main steps: (i) preparation of the library, (ii) screening of the library components, and (iii) determination of the chemical structures of active compounds.[59] The key feature of combinatorial chemistry is that a large range of analogs is synthesized using the same reaction conditions with different building blocks and in the same of different reaction vessels.[60] In this way, chemists can synthesize many hundreds to millions of compounds simultaneously instead of preparing only a few compounds one by one. To date, there are five general methods of preparing and screening combinatorial libraries: (i) the biological peptide library method (e.g., phage-display peptide library),[61–63] (ii) the spatially addressable parallel library method,[64,65] (iii) combinatorial library methods requiring deconvolution,[58,66] (iv) the affinity selection method[67] and (v) the OBOC combinatorial library method (Table 2).[57,59]

1.2.1 Biological Libraries Method

Polypeptide libraries can be obtained by molecular biology techniques with different biological entities, such as bacteria, phages, yeast and ribosomes. Among these biological peptide library methodologies, the phage-display technique based on phage exposition with peptides displayed by fusion with a viral coat protein is the most widely used.[68–70] Many filamentous bacteriophages are used for displaying peptides but M13 is the most commonly used because of its high capacity for replication and the ability to receive large DNA inserts into its genome.[71,72] Smith firstly developed the "fusion phage" method by inserting foreign DNA fragments into the encoding gene of the pIII protein to display short peptide library on the surface of filamentous M13-derived bacteriophage for mapping antibody epitopes.[73] This led to the expression of L-amino acid containing peptide on the virion surface, which did not affect virus infectivity. Bacteriophages are single-stranded DNA viruses that infect bacteria.[74] The major coat protein pVIII which is represented by about 2700 copies is the most abundant capsid component. At one end of the phage virion, there are 5 copies each of two minor coat proteins pIII and pVI for bacterial cell binding and the termination of phage particle assembly. Three to five copies of minor coat proteins pVII and pIX are at the other end of the phage for initiation and maintenance of phage assembly. The most commonly used coat proteins for peptide libraries displayed on M13 are pIII and pVIII (Figure 9). Higher affinity ligands are usually obtained by protein pIII.

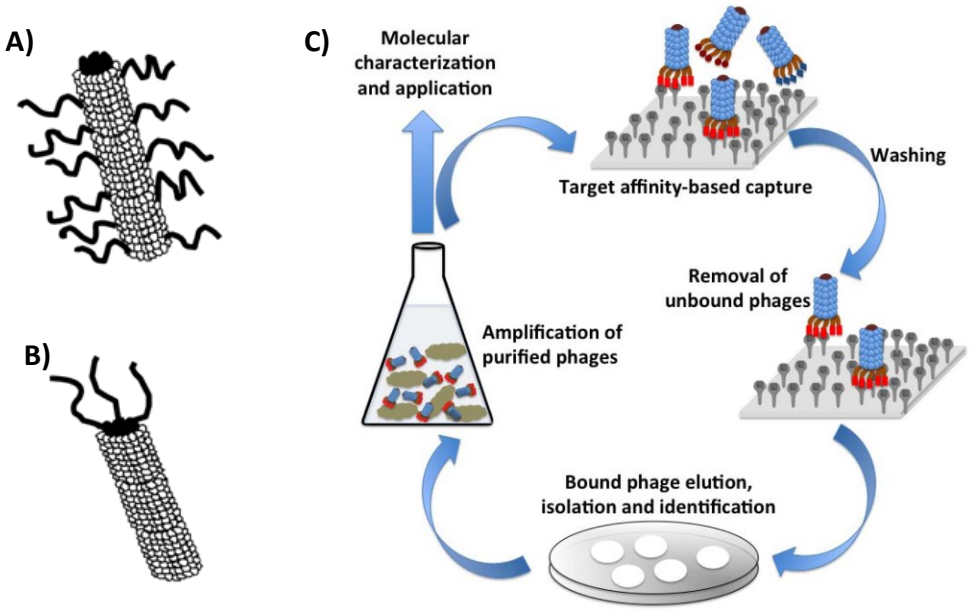


Figure 9. Peptides displayed on M13 phage surface as fusion to the amino terminus of (A) pVIII or (B) pIII; (C) affinity-based selection procedure adapted in phage display technology.[75]

In general, the extremely powerful biologic library method provides the following advantages. (i) It enables the researcher to routinely generate 10^8 to 10^9 different phages; (ii) a large number of different peptide sequences are easily produced on viral surface for screen. Longer peptides or proteins can be easily constructed without the limitation as in the case of the synthetic peptide library; (iii) it is convenient to generate the library by simply growing the microorganisms; (iv) random oligopeptides can be grafted on such tertiary folds by taking advantage of known protein folds, such as immunoglobulin fold, zinc-finger fold, or conotoxin fold. However, phage display for high-quality, large-diversity library requires specialized resources and skills. And several major disadvantages are present, such as (i) only the natural L-amino acid peptide libraries (20 natural proteinogenic amino acids) can be incorporated into these libraries, unnatural amino acids or other organic building blocks is not feasible; (ii) although simple disulfide cyclization is feasible, complicated bicyclic, compact scaffolding, branched structures, or molecules with special chemistry of cyclization are impossible; and (iii) screening assays of the biological libraries are generally limited to the binding assays (e.g., panning) and some functional assays such as protease substrate determination.

1.2.2 Spatially Addressable Parallel Library Method

Combinatorial libraries can also be generated by using another effective method, spatially addressable parallel library method. In this method, a collection of compounds is synthesized on a variety of carriers or in solution in a spatially addressable format. Depending on the library method used, screening can be performed either by a solid-phase binding assay or by a solution-phase assay. The chemical structure of positive compound can be inferred by the location, decoding is not needed. The reported techniques based on this strategy include multi-pin technology,[54] tea bag method,[76,77] SPOT synthesis,[78] and peptide microarray[64,79] (Figure 10). The main advantages of this method is structural determination of the library member is not required because these compounds is spatially addressable. In additional, D-amino acids, other unnatural amino acids or even other chemical moieties can be included in the library's structure. The major drawback of this method is that only limited number of compounds can be synthesized and therefore the library is very small.

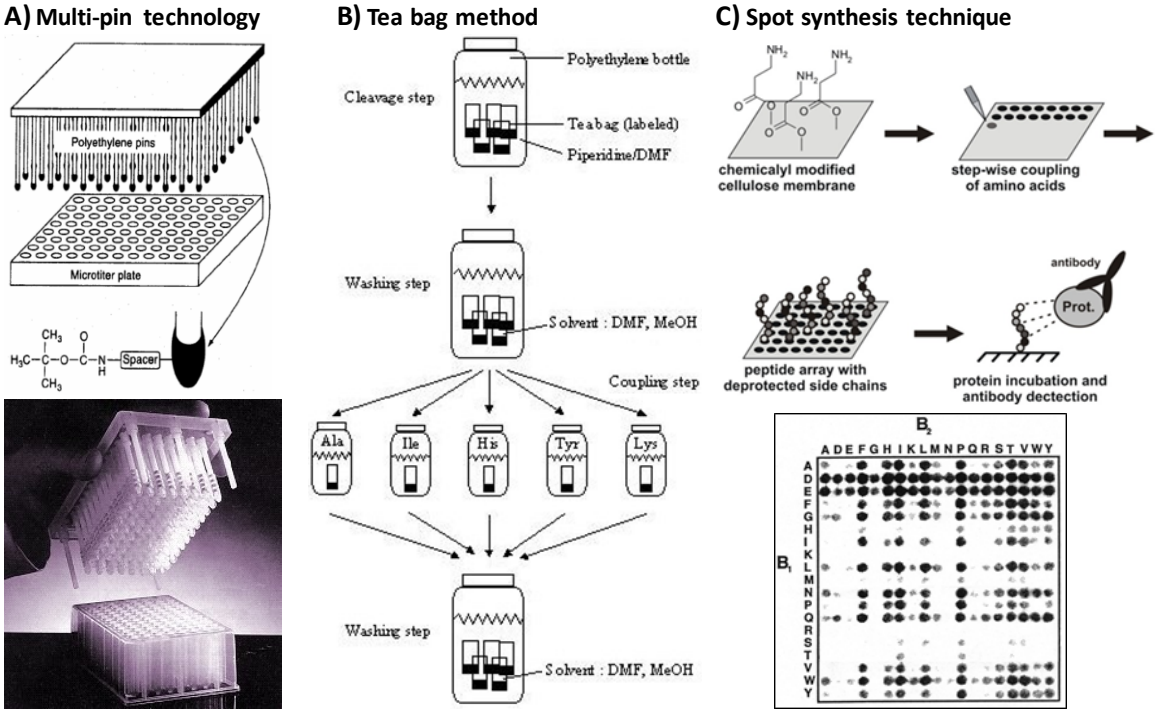


Figure 10. Techniques based on the spatially addressable parallel library strategy. (A) Multi-pin technology, (B) Tea bag method and (C) Spot synthesis technique.

Geysen et *al*.[54] initially introduced ***Multi-pin*** method in 1984 (Figure 10A). In this method, amino-functionalized polyacrylic-acid-grafted polyethylene rods which were called "pins" were inserted into an adapter fitting over a 96-well microtiter plate containing the reagent solutions. A unique compound is performed in each individual well of the microtiter plate. Biological screening was performed by an ELISA on the peptide pin. Latterly, peptides were able to release by cleavable linkers for solution assay in this pin method.[80–82] The chemical structure of any component is easily determined by its spatial location. Today, the multi-pin technology has been commercialized by Mimotopes (San Diego, CA, USA) as Pepsets. Peptide loading of each pin has been increased significantly by introducing the so-called "lantern" to fit into the tip of each pin.

A year later, Houghten introduced the "***Tea bag***" method for simultaneous multi-peptide synthesis (Figure 10B). [83] In this method, batches of resin (~25-100 mg) are sealed inside labeled, porous polypropylene packets (tea bags). Amino acid coupling reaction is performed in individual bag to generate activated monomers. After completion of each coupling, all tea-bags are collected into a large vessel and all common steps, such as resin washing and amino acid protecting group removing, are treated simultaneously. In 1995, a update method by using radio-frequency microchips (R_f tagging) for encoding combinatorial chemical libraries and a sample of resin beads for peptide synthesis in mini-baskets (e.g., NanoKan) was reported.[76,77] Each of these mini-baskets will be scanned with an electronic reader prior to or right after each coupling cycle during the "split-and-pool" synthesis.[56–58] Therefore, the construction of each compound can be identified by the recording history. At the end of the synthesis, single compounds are released from the beads which are inside of each mini-basket and placed in a 96-well plate to form a spatially addressable compound library. This approach takes advantage of the power of "split-and-pool" synthesis method. The main advantage of this method is that it offers considerable synthetic flexibility as one may use any resin beads in the synthesis.

Frank and coworkers synthesized peptides on a cellulose membrane or paper as the solid support instead of using polyethylene pins (Figure 10C).[84,85] In this **SPOT-synthesis** method, a circular spot is created by the volume of amino acids and coupling reagent. The size of the spot determines both the scale of reaction and the number of compounds. Different peptides are synthesized at different locations in a single sheet of cellulose paper.

The common steps of synthesis are treated by washing the whole membrane or surface with the respective reagents and solvents. The compounds synthesized can be evaluated using conventional high-throughput screening techniques while still attached to the membrane or in solution following membrane release. Cotton[86,87] (another form of cellulose) and polystyrene-grafted polyethylene film[88] segments have also been used as solid supports. SPOT synthesis is amenable to miniaturization and automation. The advantages of the method are that it is simple, cheap and provides sufficient quantities of peptides for various applications.[89]

Pioneered by the Affymax group[64] peptide **microarrays** are prepared by immobilizing many peptide molecules on the surface of a solid support in a small area in an addressable fashion. In 1991, Fodor *et al.* synthesized minute quantities of 1024 peptides on a single glass slide by using a photolithographic masking process in conjunction with light-directed peptide synthesis.[64] Each peptide spot occupies a 50 × 50 μm area. This accomplishment, in fact, predated DNA microarray. The immobilization can be achieved via *in situ* synthesis or chemical ligation through a covalent bond. A hydrophilic linker between the solid surface and the peptide usually is added to minimize steric hindrance caused by the solid support. The most commonly used solid support for microarray printing is a standard microscope glass slide but other solid supports include cellulose sheets[65] and polymer-based membranes.[90,91] Because of the limited quantity of peptide available, biological assay is restricted primarily to binding or functional assay on the slide. In addition, this method is not widely available since the requirement of complicated instrumentation.

1.2.3 Synthetic Libraries Requiring Deconvolution

In this method, mixtures of compounds are first synthesized and then screened for specific biological property. On the basis of the results from biologic assays, the structural determination of active compound may be able to be deduced without need of further structure elucidation. The term “deconvolution” has been used to describe the process whereby the active molecules in a library are identified, usually by the iterative testing of mixtures of compounds for a specific biologic property. The general deconvolution approaches includes the iterative approach,[94] the positional scanning approach,[93] the orthogonal partition approach,[94] the recursive deconvolution approach,[95] and the dual

recursive deconvolution approach.[96] There are following advantages in the method. A large number (10^6 to 10^8) of peptide library is able to be synthesized and screened for many existing biological assays. Decoding library is also not necessary because the structure of the active compounds can be deduced. However, in general, additional synthesis and testing, which are time-consuming, are needed to reach a final solution.

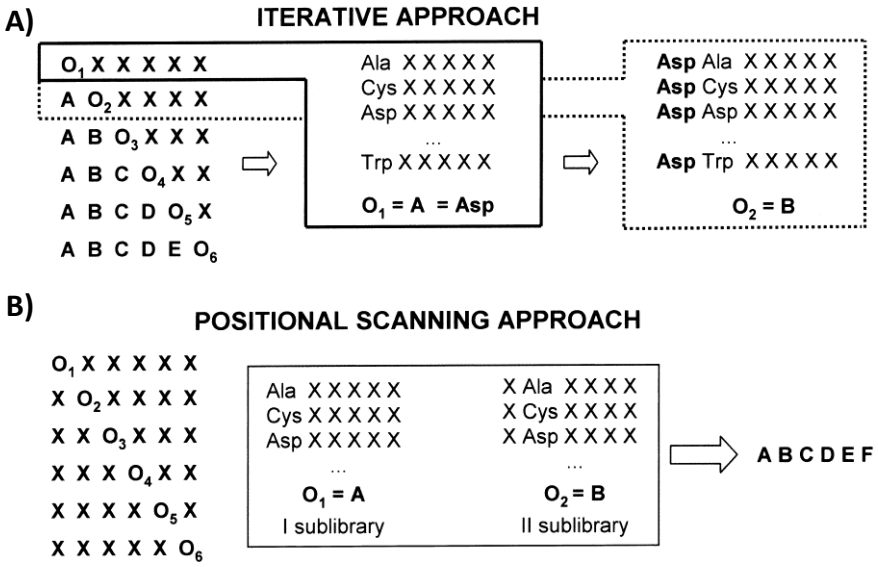


Figure 11. Descriptive model of (A) iterative process and (B) positional scanning deconvolution approaches. In the case of positional scanning, the sequence is reconstructed at the end of the process, whereas in the iterative process the sequence is obtained step by step.[97]

The two main deconvolution approaches are the iterative process[92] and the positional scanning (Figure 11).[93,98] The iterative approach was first reported by Geysen *et al.* in the multi-pin system. [92] Each position is chosen by one amino acid using a progressive selection way. Sublibraries are generated based on the result of the previous one; therefore, the structure does not need to be sequenced. A related approach, positional scanning, was later introduced by Dooley & Houghten.[93] In this method, sublibraries are synthesized with one of 20 amino acid fixed at one specific position while randomized in other positions. Then the sublibraries are tested for biologic activities. At the end of process, the sequence of active peptides can be deduced based on the biologic assay results. This method assumes that the contribution of each amino acid residue to the biologic activity is independent of

each other and it works very well for one existing predominant. If there are multiple binding motifs, scramble or uninterpretable results may be obtained.

1.2.4 Synthetic Library Method Using Affinity Chromatography Selection

In this method, the peptides, which are usually synthesized by "split-and-pool" method, are cleaved off the resins to form a solution phase peptide library.[56–58] The generated solution of peptide mixture is then loaded onto an affinity column with immobilized receptor to select the active molecules. After thorough washing, the bound peptides are eluted and structure are determined.[99] This method has been applied successfully for combinatorial oligodeoxynucleotide library in which the bound oligodeoxynucleotides are eluted and amplified. However, there are only few applications for peptide library because peptides cannot be amplified or cloned. Zuckermann *et al.* were able to apply the affinity selection method to retrieve three peptides from a very small peptide library (19 compounds) in solution.[100] Cantley *et al.* reported the successful application of the affinity selection method to identify peptide motifs for SH2 domains and kinase domains of protein tyrosine kinases.[67,101] The following points should be considered (i) nonspecific binding, if a huge peptide library is used or the affinity binding is not very high; (ii) uninterpretable sequencing data, if there is more than one predominant motif in the peptides mixture; (iii) To have enough material for structure determination may be another problem since the peptides cannot be amplified. Generally this method can be applied only to a relatively small peptide library (e.g. <10 000 peptides).

1.2.5 One-Bead One-Compound Library Method

The «one-bead one-compound approach» (OBOC) is a powerful approach that has been widely used to generate large peptide libraries. The concept was initially reported by Lam *et al*. where a library including 10^{13} compounds was prepared by "split-and-pool synthesis".[57] The described library contained a large number of single beads displaying many copies of a single compound and was named one-bead-one-compound library. In the OBOC method, the beads are split and placed into separate reaction vessels. Different building blocks are coupled on and within the beads. After each coupling step, the beads from each reaction vessel are combined and then randomly split again to the different reaction vessels for next coupling step. A library is prepared after repeating those process

several times. A simple example is illustrated in Figure 12 with the preparation of a tripeptide library by the "split-and-pool" methodology with proline (P), glutamic acid (E) and threonine (T) as building blocks by standard solid-phase peptide synthesis.[56,57] After each coupling step, beads from each reaction vessel are combined then split into the three vessels again for the next coupling step. After three such steps the 27 possible peptide sequences are all represented on separate beads (Figure 12). By using standard amino acid coupling procedure, an octapeptide library with 20 amino acids (including 20^8 possible copies) can be generated in 2-3 days. In this efficient method to identify ligands, thousands to millions of compounds are rapidly generated using the "split-and-pool" synthesis approach on solid support, in such a manner that each bead displays only a single compound entity.

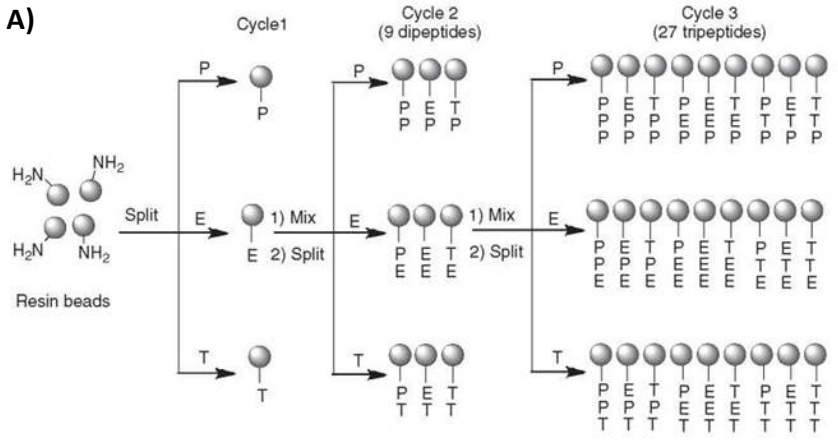


Figure 12. The "*split-and-pool*" synthesis method to generate a one-bead-one-compound combinatorial library (A) a number of permutations for random peptide libraries (P, E, and T are building blocks in this case amino acids).(B) the generated diversity.[102]

After its construction, high-throughput on-bead screening can be performed on the OBOC library to identify ligands, inhibitors or modulators for the molecular target of interest (Figure 13). OBOC libraries have been successfully used to discover linear peptidic ligands and modulators for various receptors,[103,104] enzymes,[53,105–108] transcription factors,[109] and other protein targets.[43,110–114] Because of its accessibility, flexibility and huge potential , the OBOC technology is particularly interesting to discover ligands for a protein of interest and develop new PPI inhibitors.

1.3 Screening Strategies for One-Bead-One-Compound Libraries

In general, the screening of OBOC libraries is performed on-bead. Most on-bead screening methods allow a large number of beads to be screened simultaneously against a target of interest and identify the positive beads containing the bioactive peptides. Therefore, the nature and behavior of the solid support become very important. The polymer beads used for the preparation and screening should show good swelling properties and be compatible with organic solvents for library synthesis and aqueous solution for biological assays. Among the different amphiphilic solid supports, PEG-based resins like ChemMatrix or PEG-grafted resins such as TentaGel® are the most commonly used. Compared to ChemMatrix which is composed of cross-linked PEG, TentaGel® have a polystyrene core grafted with poly(ethylene glycol) chains.[115,116] For TentaGel® resin, the loading capacity is typically around ~0.3 mmol/g and a wide variety of bead size can be used. The most commonly used are the 90 µm and the 130 µm which carry ~0.1 and 0.35 nmol compound per bead, respectively.[117] This amount is more than enough to identify the sequence of the peptide on the beads by tandem mass spectrometry (MS/MS). Moreover, the uniform size of TentaGel® beads makes the load capacity of beads consistent.

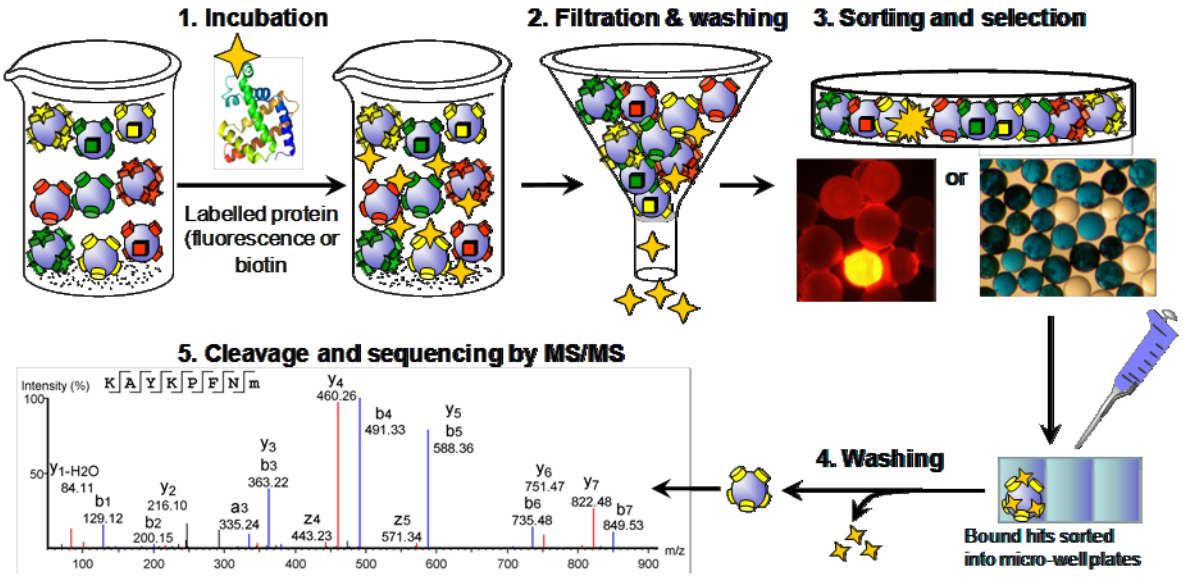


Figure 13. On-bead screening of OBOC libraries and hit identification by MS/MS.

Currently, the reported screening procedures for OBOC libraries generally involve the following steps: (1) incubation of the library with the target; (2) filtration and washing to remove unbound material; (3) sorting of the library and identification/selection of positive beads; (4) washing to remove bound material; (5) cleavage and sequence determination. Steps 1 to 4 can be repeated a few times to eliminate false positive. In the first step, tens of thousands to millions of compounds/beads are incubated with a target of interest that has been labelled. The choice of the label is extremely important in the screening process and will be based on the detection method used for the sorting in step 3. After incubation with the labelled target, beads interacting with the target are identified and isolated for compound structure determination.

Two main strategies have been used to label proteins of interest, direct protein-labelling and the use of antibodies (Figure 14). In the first strategy, the protein is directly labelled with a fluorescent derivative or with biotin. In the other strategy, a primary or secondary labelled antibody can be used. Whether direct protein labeling or antibody is used, the choice of on-bead detection technique is the same. The three most commonly used detection and sorting techniques for screening OBOC libraries are the colorimetric assays, fluorescence assays and the magnetic sorting.

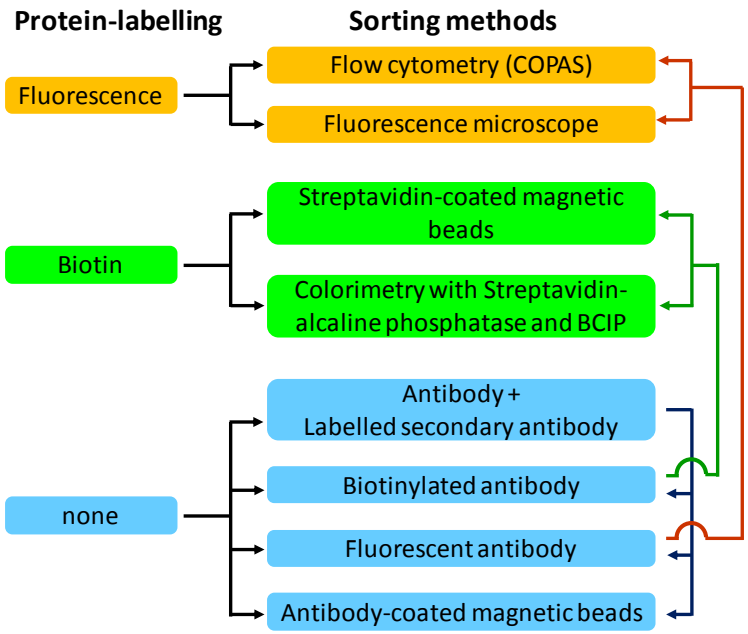


Figure 14. Protein-labelling strategies and sorting methods for the screening of OBOC libraries and selection of positive beads.

1.3.1 Colorimetric Assays

In colorimetric assay, an enzyme conjugate is used to transform a soluble colorless molecule into a colored precipitate. During the transformation the beads become colored and can be easily isolated when the library is visualized under a microscope. The target protein is usually labelled with a biotin moiety which is a well-known molecule that binds very tightly to streptavidin, a tetrameric protein with four biotin binding sites.[118] The binding of biotin to streptavidin, with a dissociation constant K_D on the order of 10^{-15} M, is one of the strongest known protein-ligand interactions. The small size of biotin means that biological activity of the protein will most likely be unaffected. The coupling of a biotin moiety to a protein or other molecule has been used in a wide variety of biotechnological applications and is called biotinylation.

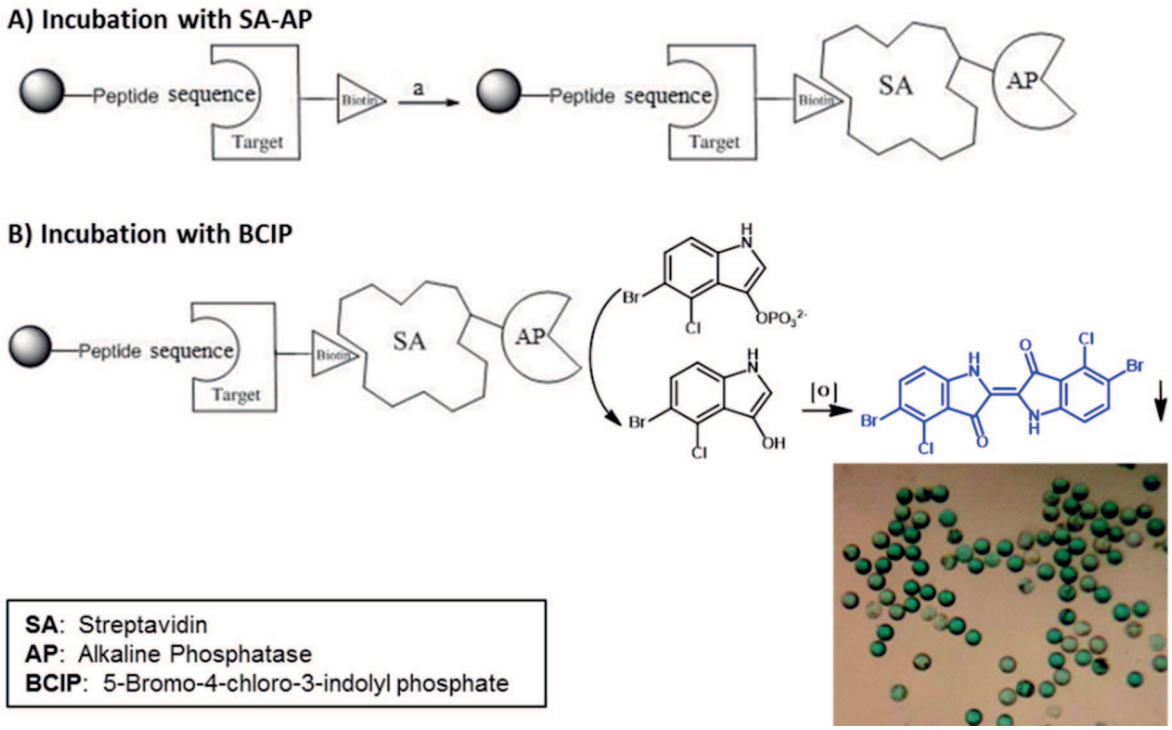


Figure 15. Colorimetric sorting with SA-AP.

To detect beads interacting with the target biotinylated protein, a streptavidin-alkaline phosphatase (SA-AP) conjugate is first added to the library that has been incubated with the target and washed (Figure 15). After some time to allow the binding of streptavidin to

biotin, 5-bromo-4-chloro-3-indolyl phosphate (BCIP) is added to the mixture to color the positive beads. The alkaline phosphatase conjugated to the streptavidin molecule hydrolyses the phosphate group of BCIP to generate the indole derivative, which is oxidized by oxygen to form an insoluble blue dye. After washing, the beads can be observed under a microscope and the positive colored ones picked up individually with a micropipette. When antibodies are used, the assay is called an on-bead ELISA (Enzyme-Linked Immunosorbent Assay). After the incubation step with the target protein and washings, the assay is performed with a primary antibody against the target protein followed by a secondary antibody conjugated to an alkaline phosphatase. The coloration process is the same as described above.

1.3.2 Fluorescent Assays

In the fluorescence assay, a fluorochrome is directly linked to the target protein or to a secondary antibody. Fluorescence is a very sensitive detection technique that involves the absorption of light or other electromagnetic radiation and emission of light at longer wavelength. The main advantage of this technique is that, in contrast to the colorimetric assay, fluorescence does not require post-incubation chemical reaction. After incubation and washing, the positive beads can be directly identify and isolated under a fluorescence microscope. Amongst the great number of fluorochrome commercially available, Fluorescein and Texas red are the most commonly used for OBOC library screening. They differ in optical properties, such as the intensity and spectral range of their excitation and emission wavelengths. The choice of the fluorochrome depends on the sorting methodology that will be used and the available filters. Special caution should be taken to avoid auto-fluorescence from the solid support. To reduce interference of beads auto-fluorescence, the protein can be labeled by dyes with emission maximum in the red that emitted in a region where the TentaGel beads fluorescence is less intense.[112]

However, after screening by colorimetry or fluorescence assays individual positive beads are picked manually using a pipette by naked eye under a microscope, which is a time-consuming and labour-intensive process. Therefore it was necessary to introduce more high-throughput sorting techniques for efficient selecting positive beads from a library containing millions of compound beads.

1.3.3 Complex Object Parametric Analyzer and Sorter (COPAS)

High-throughput screening which is defined by the number of compounds tested to be in the range of 10,000-100,000 per day is the key element for discovery new chemical ligands or drugs.[119] A flow cytometer for larger objects called a *Complex Object Parametric Analyzer and Sorter* (COPAS) has been used to sort the beads in a library (Figure 16). This instrument is capable of analyzing and sorting objects sized from 50-500 microns automatically.[120] Five different parameters are analyzed during sorting: object size, optical density and up to three spectrums of fluorescence (green, yellow, red). Compare to the traditional flow cytometer, the COPAS has been designed for large particles such as drosophila eggs and caenorhabditis elegans larvae and the pneumatic sorting mechanism does not damage the beads.[121] The sorting rate of the COPAS is 50,000 to 100,000 beads per hour on the basis of the chosen parameters, beads size, molecule density and fluorescence signals.[122]

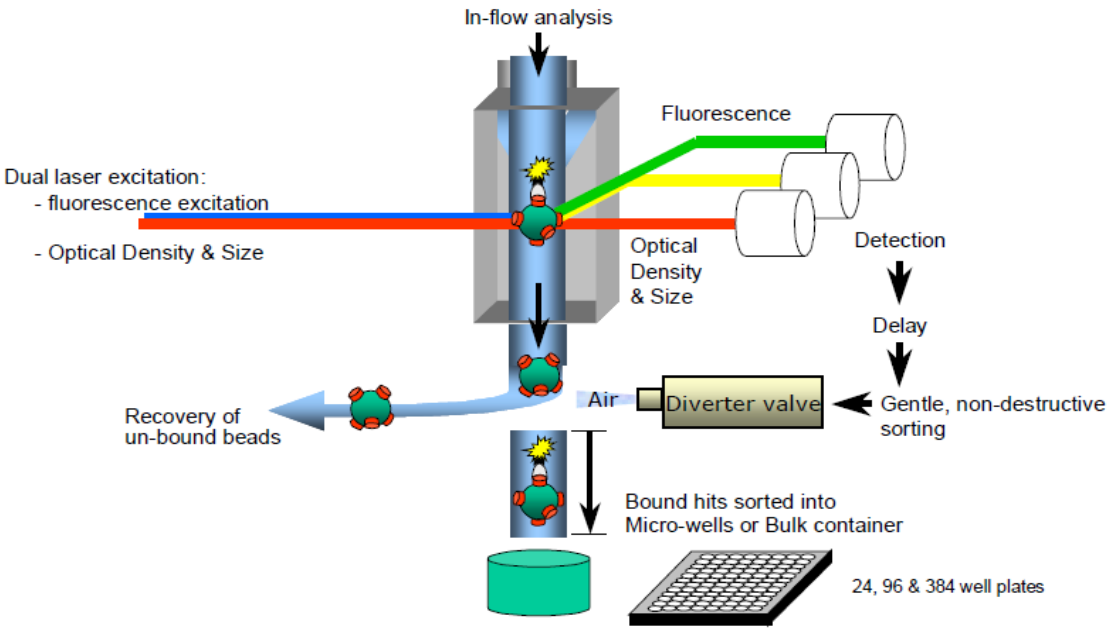


Figure 16. Schematic of COPAS sorting.

Analyzed peptide beads which have been incubated with the fluorescent labelled protein target can be sorted and dispensed into 24, 96 or 384 microwell plates.[123] Another major advantage of the COPAS is the statistics that are generated during the analysis and the possibility to choose aimed emission intensities for the sorting or to split the positive in

groups of different intensities. Reddy *et al.* used COPAS to do pre-screening to remove the most intensely auto-fluorescent beads and sort the remaining beads again to obtain few hits showing high K_D.[124]

1.3.4 Magnetic Beads

Magnetic beads have been widely used to isolate macromolecules and cells over the last decades. More recently, Astle *et al.* simplified the isolation process for positive beads by using magnetic beads.[125] In the described experiment, the target protein was linked to magnetic nanoparticles. After incubation with the library, the positive beads, which are bound with the magnetic target protein, are separated from the rest of the beads by using a magnet on the outside of the tube (Figure 17). This technique is very convenient to prune large OBOC libraries. Based on these results Pei group used commercially available magnetic beads bearing streptavidin named Dynabeads® to sort and isolate beads from libraries incubated with a biotinylated target protein.[126] The use of this assay method for solid-phase combinatorial screening is rapidly growing by several groups.[127–129] It offers the opportunity to use the same labelled protein for the first and subsequent screenings.

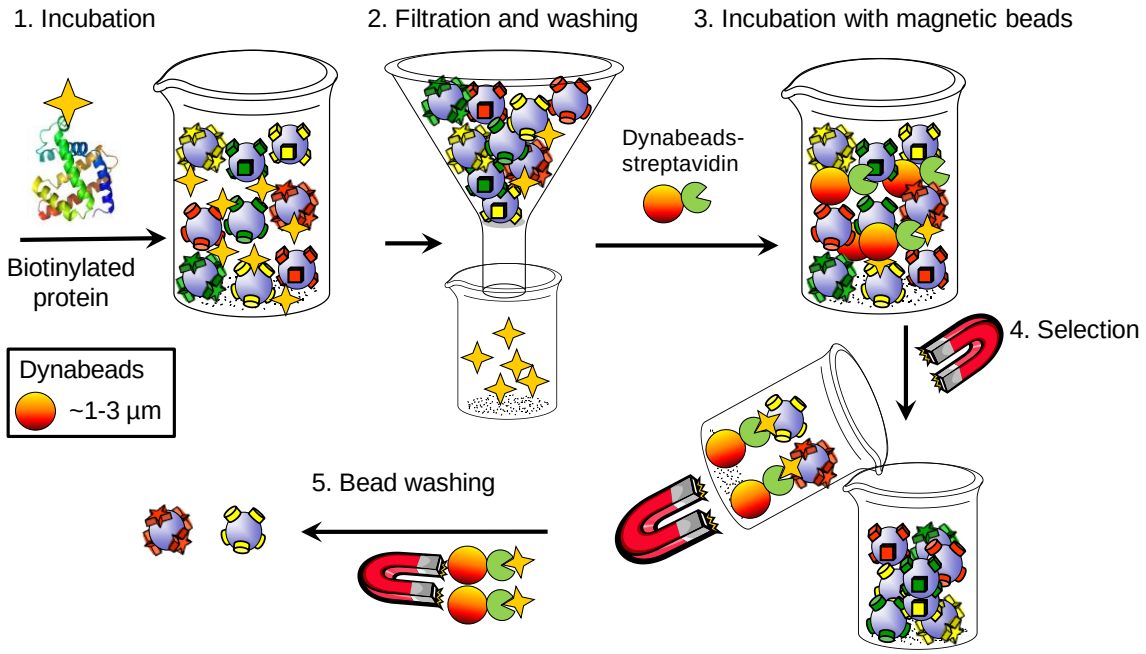


Figure 17. Sorting and bead isolation with magnetic beads.

1.4 Strategies to Decode One-Bead One-Compound Cyclic Peptide Libraries

After the screening process and selection of positive beads (hit), the identity of the exposed compound needs to be determined. This very important step is often the bottleneck in a screening process. In the case of peptide-based libraries, several high-throughput sequencing methods including Edman degradation,[102,130] encoding tags[131,132] and tandem mass spectrometry[133] have been used to identify peptide sequences. However, most of these methods have important drawbacks. For instance, Edman degradation is expensive and time-consuming, and can't be used for large numbers of peptides. In the other hand, encoding strategies require additional and compatible chemical steps in the library synthesis and the tag can potentially interfere during screening. Finally, the most accessible method seems to be the tandem mass spectrometry. The only disadvantage is the high price of the instrumentation that limits its utilization small laboratories.

With the exception of genetically encoded libraries, the use of cyclic peptides in OBOC libraries has been limited by difficulties in sequencing hit compounds after the screening. Lacking a free N-terminal amine, Edman degradation sequencing cannot be used on cyclic peptides and due to random ring-opening during ionization, complicated fragmentation patterns are obtained by tandem mass spectrometry (MS/MS) making spectral interpretation difficult or impossible.[134,135] For this reason, the use of cyclic peptides in OBOC libraries requires encoding. To overcome these drawbacks, different approaches to decode cyclic peptides in OBOC libraries have been developed. To date, there are three general strategies to decode OBOC cyclic peptide libraries: (i) the ladder synthesis approach, (ii) the one-bead-two-compound method and (iii) the ring-opening approach.

1.4.1 The Ladder Synthesis Approach

Peptide ladder synthesis method was initially introduced for linear peptides by Youngquist *et al*.[136] and Sepetov *et al*.[137] In their study, each polymer bead displays not only the full-length molecule but also small amounts of N-terminally capped truncated ladder members (Figure 18). During the coupling of each position in the randomized region, a small proportion (10%) of a capping agent, N-acetyl-alanine (Ac-Ala), was added to each

reaction vessel along with the individual amino acid (90%) (Figure 18A). This terminated the synthesis at a small percentage of the growing peptide chains as each monomer was added. These terminated peptides provide a record of the synthesis on each resin bead. At the end of the synthesis, each resin bead contains a full-length peptide as the major product as well as the corresponding family of termination products. After cleavage with CNBr, the released peptide mixture was analysed by MALDI MS. The amino acid identities can be established from the mass differences observed between adjacent members of the peptide ladder.

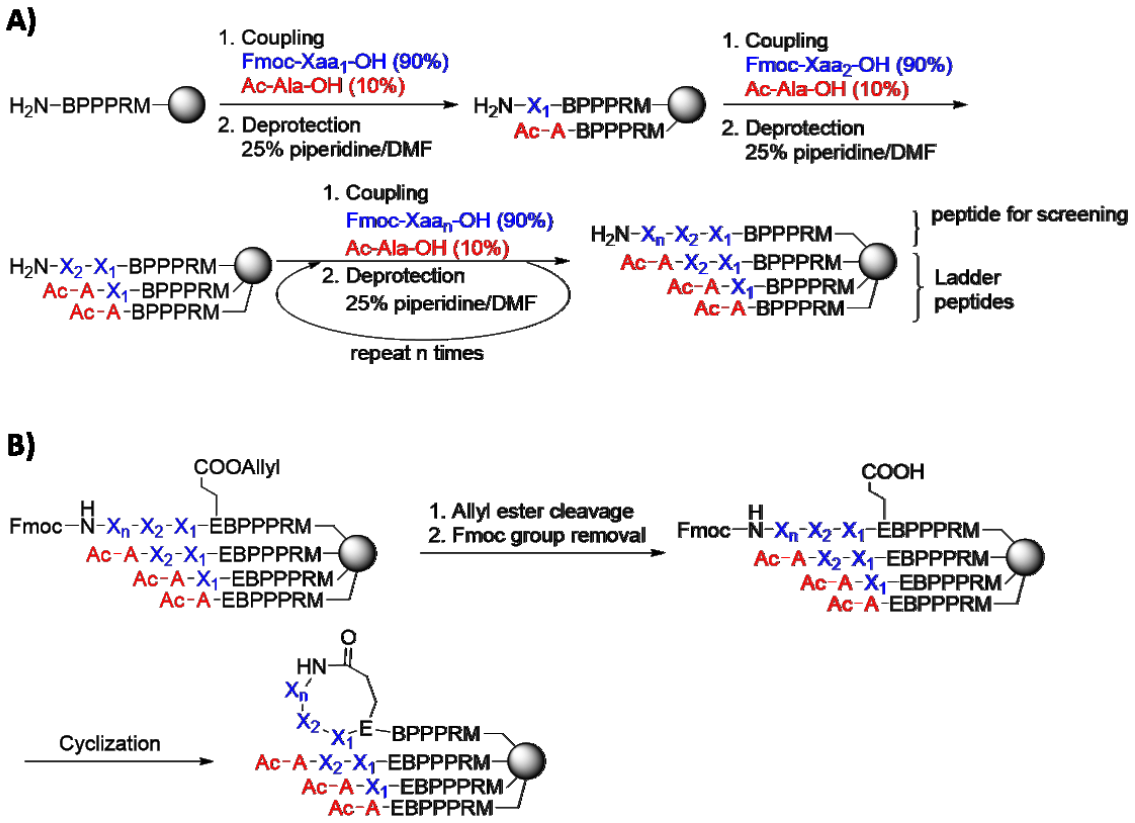


Figure 18. The Ladder Synthesis Approach.[136]

This efficient method was also developed for unnatural amino acids or other non-sequenceable building blocks libraries. Other different modified ladder synthesis approaches were introduced. Davies *et al.* used a small portion of methionine in amino acid as coupling reagents in each step to generate such a ladder by cyanogen bromide cleavage.[138] Another group used a different protecting group, such as Boc-AA instead of

Fmoc-AA, as capping reagent.[139] For cyclic peptides, the cyclization can be performed on the full-length peptide after selective side chain deprotection of a glutamic or aspartic acid residue (Figure 18B).

This approach has the disadvantage that the amount of chain termination is unpredictable since the building blocks have different relativities and the percentage of full-length peptides is significantly lower after multiple cycles.[140] Another drawback of this encoding approach is the fact that ladder peptides are displayed on the bead surface together with the full-length peptide, thereby causing interference by the coding tags during screening and yielding false positives due to interaction of the linear peptides with the target.[141] The ideal ladder-synthesis method for OBOC libraries would be to exclusively expose cyclic peptides on the surface of bead for screening and only allow the linear tags display in the bead interior for sequencing and compound identification.

1.4.2 One-Bead-Two-Compound (OBTC) Method

This problem has been elegantly overcome by Pei and coworkers with a one-bead-two-compound (OBTC) approach on topologically segregated bilayer beads, compatible with Edman degradation and mass spectrometry for the sequencing (Figure 19).[142] The bilayer-bead concept was initially introduced by Vágner and coworkers[143,144] with an enzymatic "shaving'' strategy using proteases to topologically segregate bead into double layers and later via different chemical approaches by different groups.[131,132,145–147] Topological bilayer segregation offers the opportunity to synthesize two compounds per bead, namely one exposed on the surface for screening (cyclic peptide) and the other inside (linear) as a tag for sequencing and compound identification. After the screening, sequence determination on the selected OBTC beads can be performed as described above, i.e. by Edman degradation or tandem MS. Our group has evaluated different parameters to improve and control the segregation step and showed that this step can be efficiently performed without water.[148] On the other hand, Kodadek group developed a different OBTC method for cyclic peptoid microarrays, that only cyclic peptoid containing a cysteine can be spotted onto a maleimide-activated microscope slide otherwise the linear encoding molecule without cysteine will not stay on the slide to interfere screening.[149]

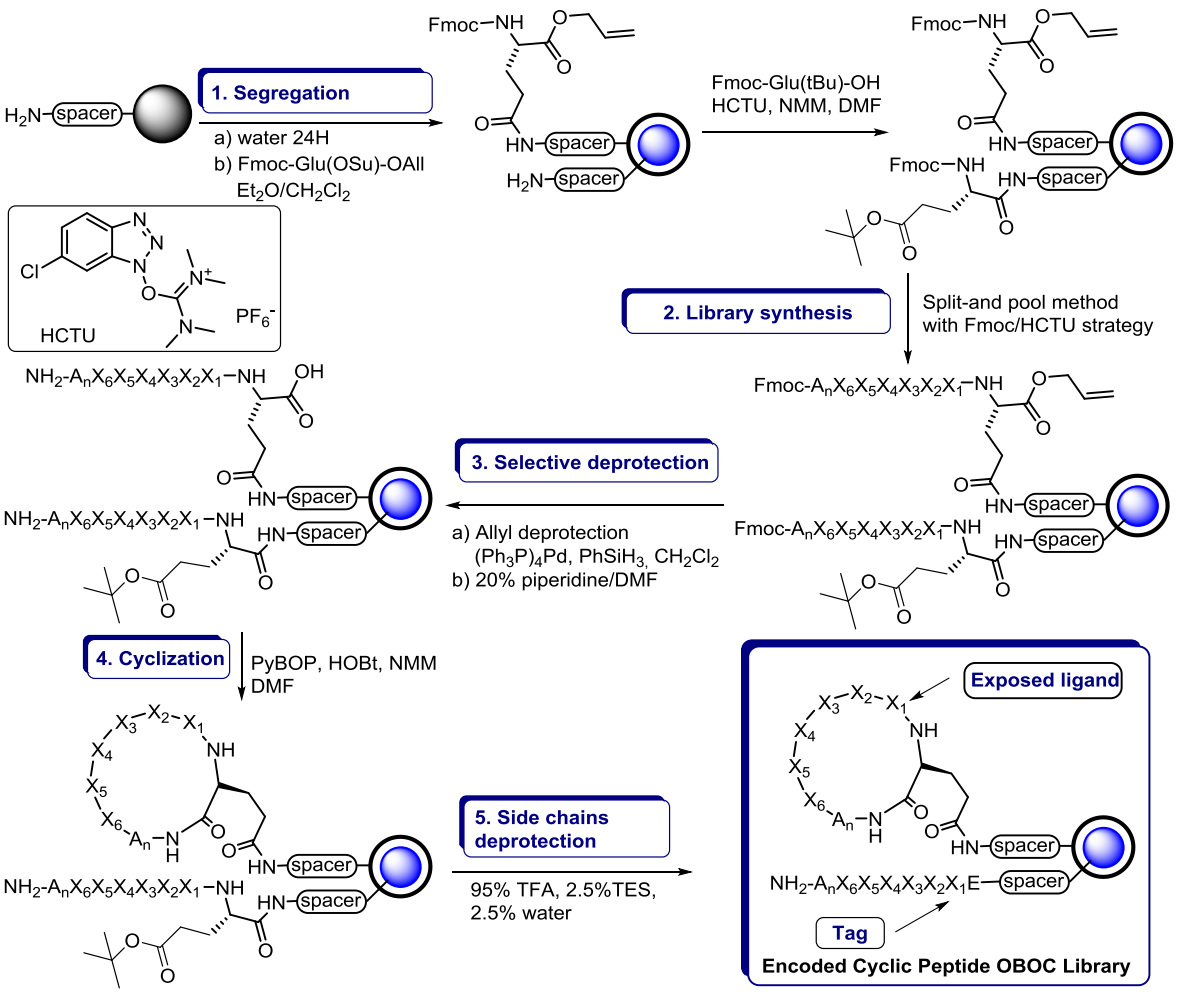


Figure 19. One-bead-two-compound approach on topological segregated bilayer beads.

Later, Chait *et al.* have developed a 'peptide ladder sequencing' method for linear peptides in solution phase.[150] This approach, called partial Edman degradation sequencing (PED), is based on the ladder synthesis with the exception that the ladder sequences are formed after the screening by partial degradation (Figure 20). Originally, the PED method used a mixture of phenylisothiocyanate (PITC) as Edman degradation reagent and phenylisocyanate (PIC) as capping reagent to treat full-length linear peptides followed by washing with TFA. The cycle was repeated many times to generate a peptide ladder. Subsequently the ladder members are released from the resin and analysed by MALDI MS to determine the sequence of cyclic peptides by calculation of the mass differences between ladder members. Lately other variations of PED was modified by other groups using

different capping reagent like Z-Osu or Fmoc-Osu (Figure 20).[140,151,152] Unfortunately, PED is very time consuming and involves the use of hazardous reagents, such as PITC.

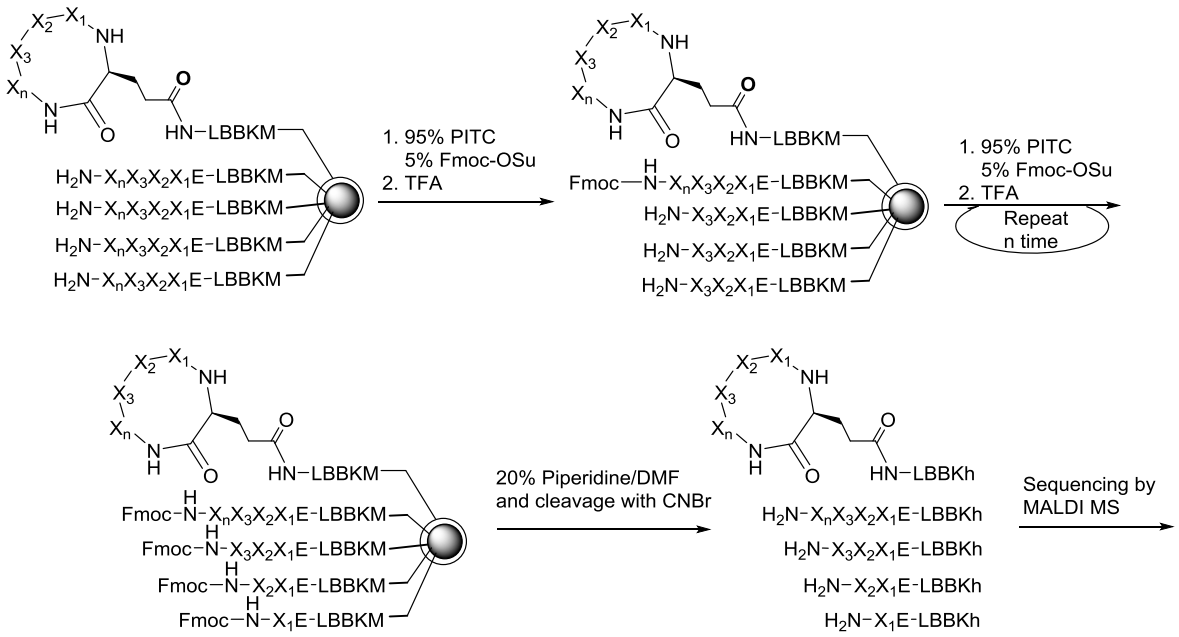


Figure 20. Partial Edman degradation approach described by Chait et *al.* (h = homoserine lactone).[150]

More recently, the friendlier linker 4-hydroxymethylbenzoic acid (HMBA), which allows to release peptides from resin by using vapour phase ammonium hydroxide, was used to avoid the use of hazardous cyanogen bromide to release the peptides from the resin.[153] In this method, a minor proportion (0.2 equivalent) of Fmoc-Ala-OH was added during the coupling of Fmoc-Asp(OPp)-OH as a branching point for cyclization to allow the simultaneous synthesis of a cyclic peptide and its linear counterpart for sequencing. Peptides were cleaved from each bead by treatment with ammonia and the sequence was deduced by MS/MS analysis of the linear peptides.

The ability of a molecule to access the interior of a bead depends on its size and the resin pore diameter.[144] Therefore, the control of the outer/inner layer ratio in the topologically segregated bilayer bead strategy is important to avoid contact between the target and the tag. It has been demonstrated that proteins under 23.8 kDa are capable of penetrating inside TentaGel beads of 90 μm diameter.[154] Thus precautions should be taken when small proteins are used in a screening. Another drawback is that, during library preparation, the

cyclization step cannot be monitored because of the presence of the linear peptide tag. This limitation and the potential interference of the encoding tag during screening have increased the interest in the development of a new encoding-free strategy to prepare OBOC cyclic peptide libraries.

1.4.3 Ring-Opening Approach

Lim's group reported an elegant ring-opening strategy on cyclic peptoids to eliminate the need for encoding (Figure 21).[155] The strategy involves the introduction of a cleavable alkylthioaryl bridge in the cycle to allow linearization of the molecule after the screening. The macrocycle "re-opening" reaction is performed in two steps. First the thioether is oxidized into a sulfone by m-chloroperoxybenzoic acid (mCPBA) then followed by nucleophilic displacement of the sulfone with a NaOH solution to generate a linear peptoid which can be sequenced by MS/MS. However, this approach was tested only on unfunctionalized side chains and the oxidation step could lead to side reactions with commonly used functionalized amino acids.

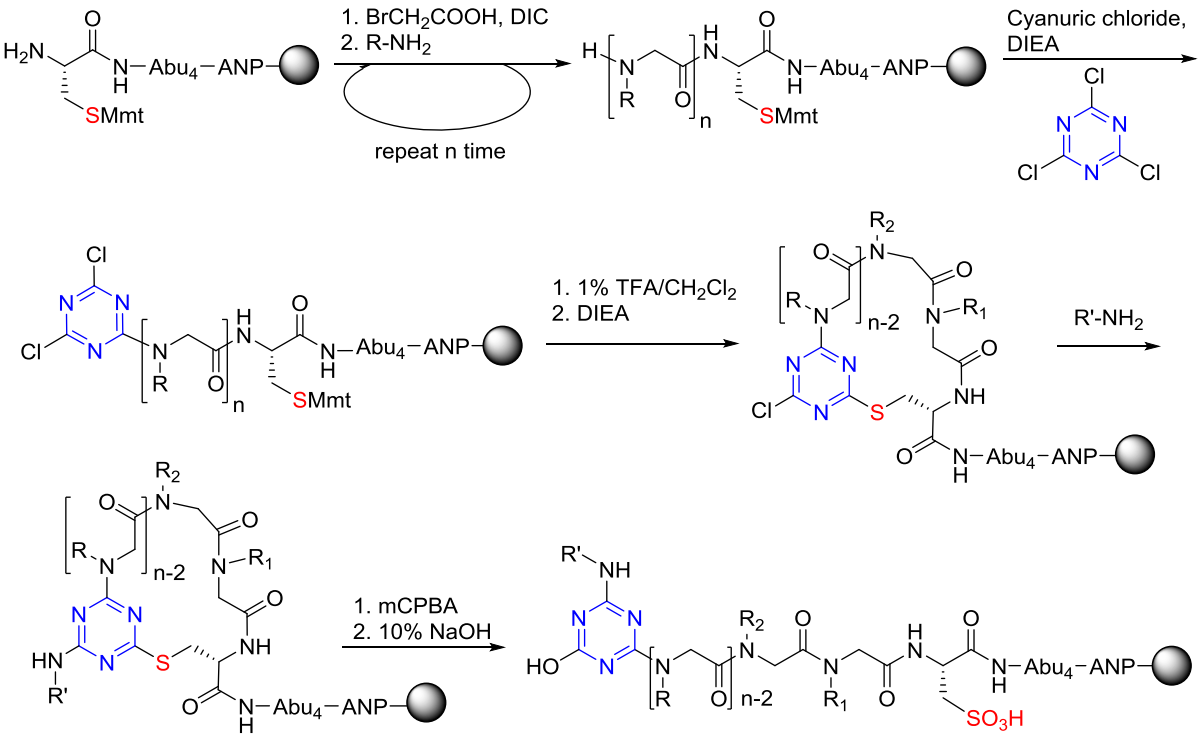


Figure 21. Alkylthioaryl bridge ring-opening strategy described by Lee *et al.*[155]

Using a similar strategy with cyclic peptoids, Kodadek group placed a methionine as a linker and within the cycle to allow a simultaneous linearization and compound release from the bead with cyanogen bromide (CNBr) (Figure 22).[127] It has been shown that CNBr can selectively cleave methionine C-terminal amide bonds to form a stable homoserine lactone (Figure 23). This reaction is well known in proteomics research and is commonly used in protein digestion experiments. Moreover, with the exception of cysteine, CNBr is compatible with every functionalized amino acid. This approach allows a dual ring-opening/cleavage with CNBr to generate linear peptides that can be sequenced by tandem MS/MS. Unfortunately, in this case poor cyclization yields (~70%) were observed and the technique stills need improvement. The fact that the reopening and cleavage can be performed in a single step is certainly a major advantage of this approach. Another drawback is the presence of two homoserine lactone residues in the linearized peptide. These residues can be cleaved or hydrolysed increasing the number of generated molecular ions and the complexity of the fragmentation patterns in MS/MS spectra.

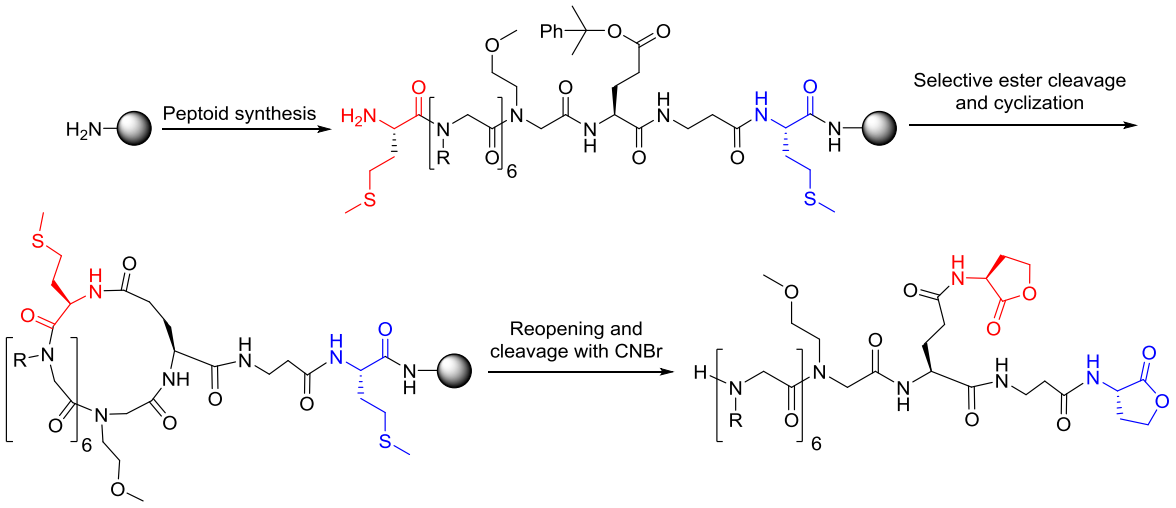


Figure 22. The dual ring-opening/cleavage strategy described by Simpson and Kodadek.[127]

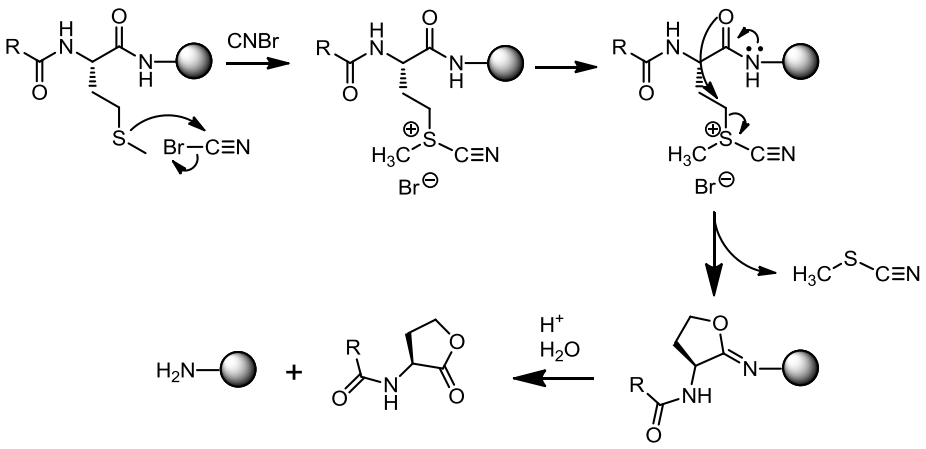


Figure 23. Mechanism involved in methionine ring-opening strategy.

To avoid the presence of two homoserine lactone residues in the linearized peptide, our group has designed an approach with a reverse methionine linker and a methionine at C-terminal as ring-opening point (Figure 24).[156] In this strategy, a Fmoc-Lys-OAll was anchored by its side chain to a reverse methionine linker chain to allow peptide cyclization and a methionine residue was introduced in the macrocycle for ring opeing. After treatment with CNBr, a linearized peptide bearing a single C-terminal homoserine lactone and a free N-terminal lysine was generated. The other homoserine lactone still remains attached to the resin. The design and development of this approach will be described in Chapter 4.

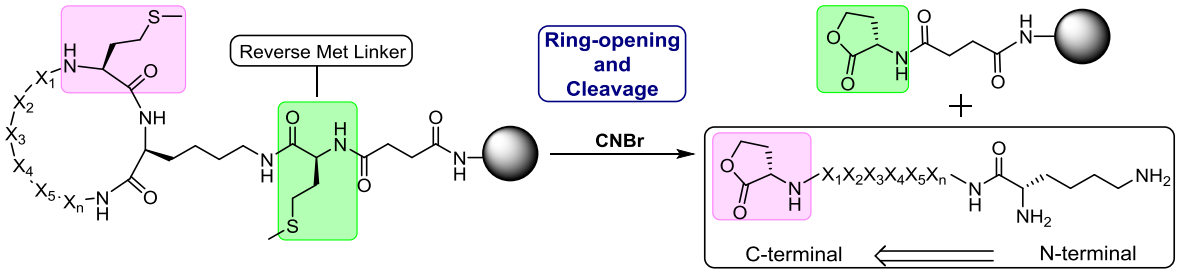


Figure 24. The dual ring-opening/cleavage strategy with reverse methionine linker described by Liang et *al*.[156]

More recently, based on the approaches with methionine described by our laboratory and Kodadek's group, Lim and coworkers used an homocysteine residue to prepare tail-to-side chain cyclic peptides and peptoids by thioether formation (Figure 25).[157] The strategy used this thioether bond to allow a simultaneous ring-opening and cleavage from the resin upon

treatment with CNBr. In this method, a trityl-protected homocysteine was firstly coupled on resin. After preparing peptide/peptoid sequences, the cyclization was performed by chloride displacement on chloroacetylated N-terminus with the deprotected homocysteine side chain. The expected linearized product was successfully detected after the treatment with a CNBr solution. Compared to the previous approach with cyanuric chloride developed by this group, this new method avoids the usage of the strong oxidizing agent mCPBA and the risks to oxidize side chains functional groups.

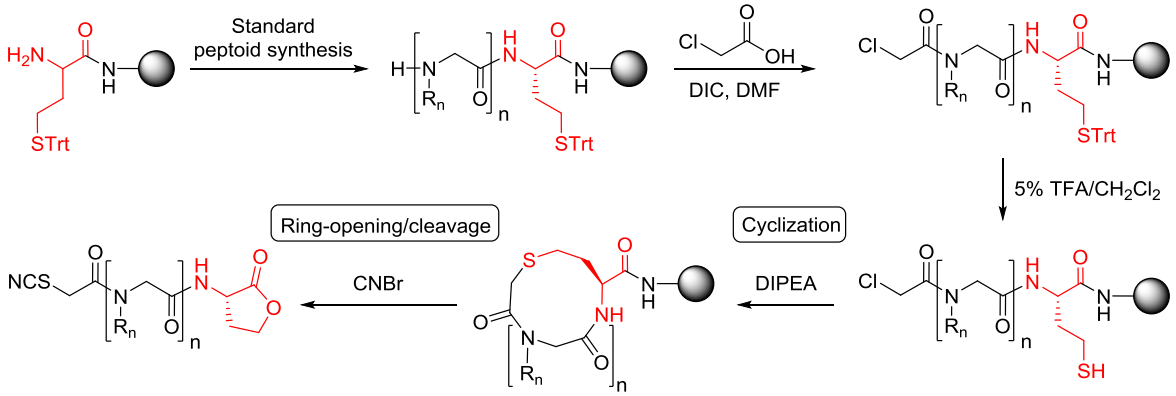


Figure 25. Homocysteine system one-pot ring-opening/cleavage reaction by CNBr mediated.[157]

Like strategies using methionine, inadvertent oxidation of the sulfur atom could happen during storage, manipulation or screening and, in addition to significantly decreasing the yields, prevent ring-opening of the macrocycle. To avoid this problem, it is strongly suggested to treat the cyclic peptide library with a reducing solution to reduce any oxidized methionine before the tandem ring-opening/cleavage reaction.[158]

The ring-opening strategy was also applied to cyclic depsipeptides. Menegatti *et al*. used lactic acid and the dipeptide ester *N*-Ac-Ser(OAla)-OH as linkers for dilactonization and to introduce two ester bonds in the macrocycle (Figure 26).[159] Upon alkaline treatment with a solution 0.1M NaOH in acetonitrile/water (8:2) during 20 min, the cyclic depsipeptides are linearized and released from the solid support to be sequenced by MS/MS. In this case, the ChemMatrix resin was used as the solid support because it contains exclusively primary ether bonds and is stable under nucleophilic conditions.[160] The identity of the generated peptides could be efficiently determined but they showed that purities over 80-85% could

hardly be achieved. The impurities, comprised of truncated and acetylated linear depsipeptides, are mainly due to the two esterification reactions, whose yields are lower than those of other coupling steps.

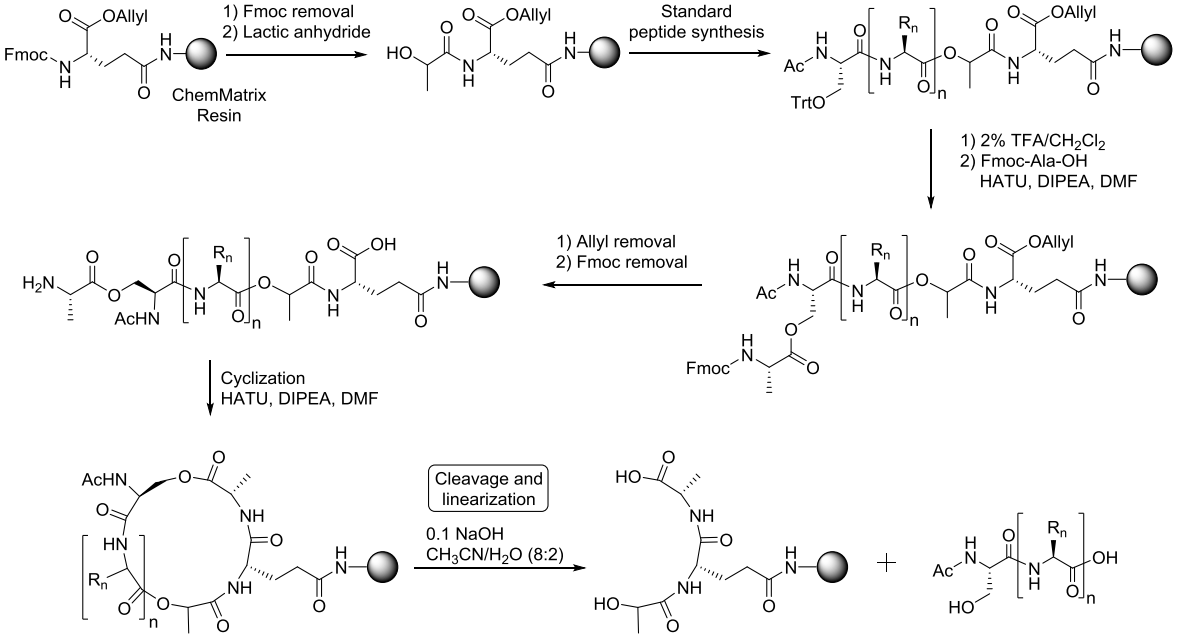


Figure 26. The cyclic dilactone approach to generate linear peptides.[159]

More recently, Gurevich-Messina *et al.* used the HMBA linker on ChemMatrix resin to prepare cyclic depsipeptides allowing simultaneous ring-opening and cleavage from the resin upon treatment with vapour phase ammonium hydroxide during 15 h (Figure 27).[161] Unlike other nucleophiles such as NaOH, the ammonia is easily removed by evaporation and hence does not interfere with the MS peptide analysis. After dual cleavage, the peptide sequences could be deduced from the MS/MS spectra. Incomplete cyclization was also observed for the tested depsipeptides. Overall this strategy avoids the use of hazardous reagents such as Pd or CNBr, and therefore, can be applied in a broad range of laboratories without special expertise in organic synthesis.

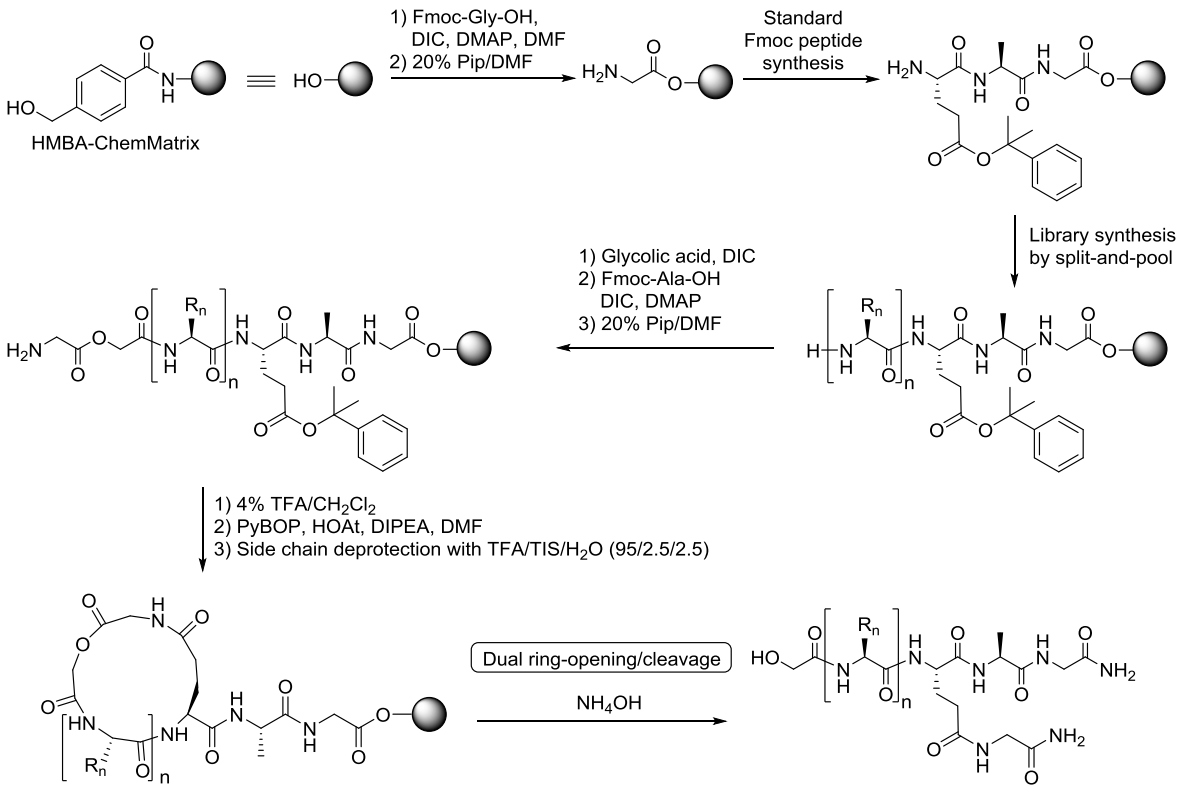


Figure 27. Dual ring-opening/cleavage of cyclic depsipeptides with ammonia described by Gurevich-Messina *et al.*[161]

Macrocyclic peptides represent excellent templates to discover protein ligands and to develop new PPI inhibitors. Unfortunately their use in OBOC libraries is limited by difficulties in sequencing hit compounds after the screening. Lacking a free N-terminal amine, Edman degradation sequencing cannot be used on cyclic peptides and due to random ring-opening during ionization, complicated fragmentation patterns are obtained by tandem mass spectrometry (MS/MS) making spectral interpretation impossible.

To overcome this major drawback, different strategies have been developed and include the ladder synthesis, encoding tags and ring-opening approaches. For the ladder synthesis and encoding methods, choosing a tag that would not interfere with biological assays is the key point. Although topologically segregated bilayer beads method reduce the risk of interferences, small size proteins often used in library screening can still penetrate inside the beads and interact with the tags. Until now, the most promising methods use a ring-opening approach. The different strategies that have been developed reduce the needs for

post-screening chemical modification and allow a fast sequence determination of cyclic peptides from combinatorial libraries by MS/MS. Indeed, in contrast to other encoding methods, the re-opening strategies display the same molecules in both the interior and the exterior of the beads, thus eliminating the risk that the interior molecules (tags) could interfere with screening by interacting with target proteins. This approach is very appealing and will be a starting point to develop a strategy adapted to our needs.

Chapter 2

Working Hypothesis and Objectives

2.1 Working Hypothesis

Based on the nature of the interactions interface that involves the recognition of protein secondary structures in protein interaction domains, we consider that cyclic peptides are very promising templates to discover protein ligands and development new PPI inhibitors. Cyclic peptides show very interesting peptidomimetic abilities and compared to their linear counterpart they exhibit tighter-binding for receptors, higher stability against proteolytic degradation and better cell permeability. Moreover, their synthesis is straightforward and the great degree of molecular complexity and diversity that can be accessed quickly and easily by simple changes in their linear sequence makes the use of cyclic peptides as scaffolds in drug design and discovery very appealing.

In order to fully exploit the great conformational and functional diversity accessible with cyclic peptides, combinatorial chemistry is certainly the most powerful approach. Among the different combinatorial library technologies, we are particularly interested in the OBOC method. This method is probably the most accessible and does not require specialized and expensive equipment. It also allows the generation of libraries containing tens of thousands to millions of compounds in a short period of time. Besides being affordable, the screening of OBOC libraries is very flexible and allows the use of different labelling agents and sorting approaches. This flexibility offers the opportunity to perform many biological tests to prune the library and eliminate false positives to accelerate the discovery of ligands for the target protein. Therefore, to cover a large molecular diversity, we were interested in the use of the OBOC combinatorial approach to prepare cyclic peptide libraries and screen them against proteins involved in PPI of interests.

When the project was initiated, the use of cyclic peptides in OBOC libraries was limited by difficulties in sequencing hit compounds after the screening since methods like Edman degradation and MS/MS analysis need a free N-terminal amine. Therefore new approaches allowing the decoding of OBOC macrocyle libraries were strongly needed. Based on the different methods developed to encode macrocycle and peptidomimetic libraries, we proposed to design and develop new convenient ring-opening approaches to prepare and decode OBOC cyclic peptides. Our strategy was to introduce cleavable residues in the macrocycle and as a linker to allow linearization of peptides and their release from the

beads for sequencing by MS/MS (Figure 1). The development of these methodologies will prompt the use of macrocyclic compounds in OBOC libraries and be a very important contribution to the fields of combinatorial chemistry and library screening to discover protein ligands and lead compounds.

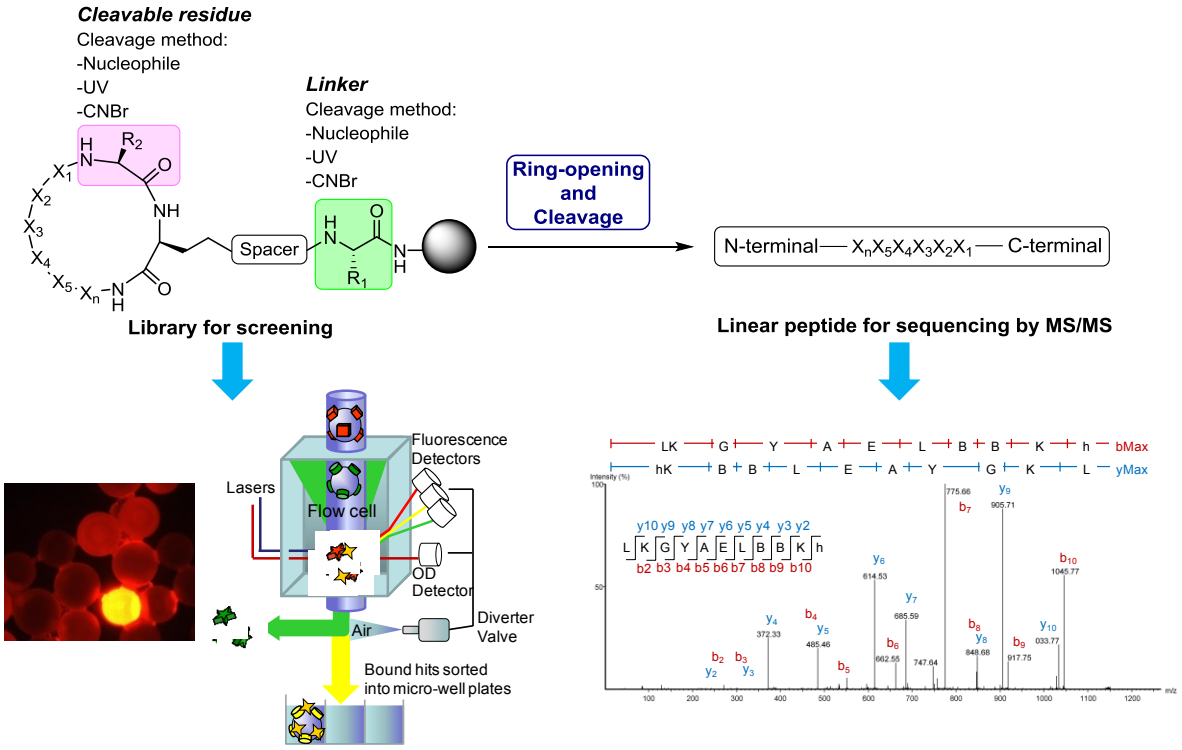


Figure 1. Design of a ring-opening approach to prepare and decode OBOC cyclic peptide libraries and its use in screening and ligands identification.

2.2 Objectives

The main objective of this project was to develop new high-throughput ring-opening approaches for the construction and decoding of combinatorial OBOC cyclic peptide libraries and to apply these methods in screenings against a protein of interest. To reach this objective, a multidisciplinary approach combining organic chemistry, solid-phase synthesis, photochemistry, mass spectrometry, peptide sequencing and biochemistry was used. The progress of the proposed research project was based on the following specific aims:

1) Evaluate and compare different cleavable residues in the ring-opening approach;
2) Design and develop dual ring-opening/cleavage strategies for the synthesis and sequence determination of OBOC cyclic peptide libraries;
3) Prepare a OBOC cyclic peptide library and perform on-bead screening against HIV-1 Nef protein;
4) Select hits and sequence positive beads to identify new ligands;
5) Validate the ligands for their ability to bind and inhibit target protein functions.

2.3 Research Plan

Our research plan was to select amino acid residues that can be cleaved in different conditions (e.g. residues cleaved with nucleophiles, UV or CNBr) and to introduce them in a model cyclic peptide. Their cleavage efficiency and selectivity will be evaluated and compared in terms of time, handling, high-throughput potential and purity of the generated linear peptides. The most promising residues will then be used in the design and development of tandem ring-opening approaches for OBOC cyclic peptide libraries. On the other hand, OBOC cyclic peptide libraries will be prepared and different on-bead screening methods will be tested to identify the best conditions for a screening against the HIV-1 Nef protein. Afterward, a peptide library will be screened against the HIV-1 Nef protein and the selected ligands will be individually synthesized to validate their binding to the Nef protein and determine their affinity.

In this thesis, the results obtained during the research project are divided into different chapters as follows:

- In **chapter 3**, we report the results obtained with the investigation on cleavable amino acid residues.
- In **chapters 4 and 5**, two dual ring-opening/cleavage approaches that have been designed and developed during the research project are presented.
- In **chapter 6**, the synthesis and screening of a OBOC library against HIV-1 Nef protein and the ligands identified are described.

Chapter 3

Evaluation of Different Cleavable Residues and Linkers for Ring-Opening/Cleavage Approaches

3.1 Selection of the Ring-Opening Residues and Linkers

The ring-opening approach has been developed to eliminate the need for encoding OBOC libraries and allow hit compounds identification after a screening. In this very promising approach, a cleavable residue is introduced in the macrocycle backbone to allow linearization of the molecule followed by cleavage from the beads. The generated linear peptide can then be sequenced by tandem mass spectrometry. The main objective of this chapter is to evaluate the efficiency and compatibility of different cleavable residues in the ring and/or as linkers for the ring-opening approach and peptide sequencing. Among the great number of cleavable amino acids and linkers available, we were particularly interested in methionine, β-amino acid 3-amino-3-(2-nitrophenyl)propionic acid (ANP) and hydroxymethylbenzoic acid (HMBA). The selection of these amino acids and linkers was based on their compatibility with most functionalized amino acid side chains and the different conditions required for the cleavage.

Methionine is one of two sulfur-containing proteinogenic amino acids and has been widely used as a cleavable residue in protein sequencing and identification.[162] Upon treatment with cyanogen bromide, the C-terminal peptide bond of the methionine (Met-Xaa) is cleaved to generate a C-terminal homoserine lactone and release a free amine N-terminal fragment (H-Xaa).[163] Therefore, as a cleavable residue or as a linker, methionine is compatible with standard solid-phase peptide synthesis, acid or base-sensitive protecting groups and its cleavage condition is also compatible with deprotected functionalized side chains. Another attractive approach is to use photosensitive linkers. Such photolabile residues are cleaved under UV irradiation rather than aggressive or corrosive reagents. *o*-Nitrobenzyl derivatives, such as ANP, are the most widely used.[66,164] The ANP residue is cleaved at wavelength of 365 nm, which is easy to obtain in chemical laboratory with standard UV lamps. Finally, peptides anchored to resin beads through HMBA linker can be released with rather strong nucleophiles, such as NaOH or ammonia vapor.[165] The benzyl ester formed between the peptide C-terminal carboxyl and the HMBA hydroxyl is stable to piperidine, which is required for compatibility with Fmoc-based solid-phase peptide synthesis. The HMBA linker can be efficiently cleaved by treatment with ammoniac gas and the released peptide washed out the beads with a suitable aqueous solvent.[153] This convenient procedure is very

economic and safer than CNBr handling and evaporation. While Met and ANP were chosen as ring-opening residues for our study, Met, ANP, and HMBA have been used as linkers. The efficiency of ring-opening and cleavage was first evaluated on a model peptide with or without a C-terminal spacer by combining the various cleavage residues and linkers (Table 1). The results obtained with the combination of two Met residues in OBOC cyclic peptide libraries will be presented in Chapter 4 and those obtained by dual photochemical ring-opening/cleavage with ANP will be discussed in Chapter 5.

Table 1. Ring-opening/cleavage Approaches

NO.	Model Peptides	Cleavable residue	Linker
1	*cyclo*[Leu-Gly-Tyr-Gly-Lys-Phe-**Met**-Glu]-NH_2	Met	ANP
2	*cyclo*[Leu-Gly-Tyr-Gly-Lys-Phe-**Met**-Glu]-NH_2	Met	HMBA
3	*cyclo*[**Met**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-NH_2	Met	HMBA
4	*cyclo*[**Met**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-Spacer-Gly-NH_2	Met	HMBA
5	*cyclo*[**Met**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-Spacer-Gly-NH_2	Met	ANP
6	*cyclo*[**ANP**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-Spacer-**Met**-NH_2	ANP	Met
7	*cyclo*[**ANP**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-Spacer-Gly-NH_2	ANP	HMBA
8	*cyclo*[**ANP**-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-Spacer-Gly-NH_2	ANP	ANP

*Spacer ≡ Leu-βAla-βAla-Lys

3.2 Model Cyclic Peptides Synthesis and Ring-Opening

Linear peptides were synthesized on TentaGel® S NH_2 or HMBA-TentaGel® S resins by standard Fmoc solid-phase peptide synthesis with HCTU as coupling reagent. After selective N-terminal Fmoc and side chain allyl ester removal, the peptides were cyclized directly on the solid support through a macrolactamization in a tail-to-side chain topology between the N-terminal amine and the carboxylic acid from a Glu residue side chain with HATU/HOAt as coupling reagent. Finally, the side chains protecting groups were removed with a TFA cocktail. To evaluate the efficiency of ring-opening and cleavage reactions, single beads were picked up and treated with suitable cleavage conditions to allow peptide

linearization and release from the beads. For ANP-containing beads, the cleavage was performed in MeOH under UV irradiation at 365 nm for 3 h. For Met-containing beads, an overnight treatment with CNBr solution was used for the ring-opening or cleavage reaction. Finally, for beads bearing HMBA as linker, the cleavage was conducted by gaseous aminolysis in a sealed desiccator with a beaker containing NH_4OH overnight. Linearized peptides were then immediately analyzed by MALDI-TOF/TOF MS. The theoretical mass by different approaches is shown in Table 2.

Table 2. Calculated $[M+H]^+$ (Da) of model peptides at different stage

NO.	Linear	Cyclic	Ring-opening	Dehydration	Adduct
1	943.46	925.45	895.46	----------	----------
2	943.46	925.45	895.46	----------	912.49
3	943.46	925.45	895.46	----------	912.49
4	1383.74	1365.73	1335.73	----------	1352.76
5	1383.74	1365.73	1335.73	----------	1251.71
6	1279.70	1453.74	1453.74	1435.73	----------
7	1252.70	1426.74	1426.74	1408.73	1425.76
8	1252.70	1426.74	1426.74	1408.73	1440.76

3.3 Evaluation of the Ring-Opening and Cleavage Combinations

In the first approach (Met/ANP), methionine was placed as ring-opening residue directly after the cyclization anchor Glu(OAll) and ANP as linker on TentalGel® S resin (Figure 1A). The bead was first treated with CNBr to yield a supported linear peptide that was thereafter released in solution upon UV irradiation. The produced linear peptides contained a C-terminal homoserine lactone and an N-terminal side chain anchored glutamine. MALDI-TOF MS analysis of the generated linear peptide showed the expected molecular ion at 895.50 Da that could be unambiguously sequenced both manually and by De Novo sequencing with the PEAKS software (Figures 1B and 1C). A peak corresponding to the oxidized cyclic peptide was also observed in the MS spectrum showing that the Met residue has been oxidized during the process.

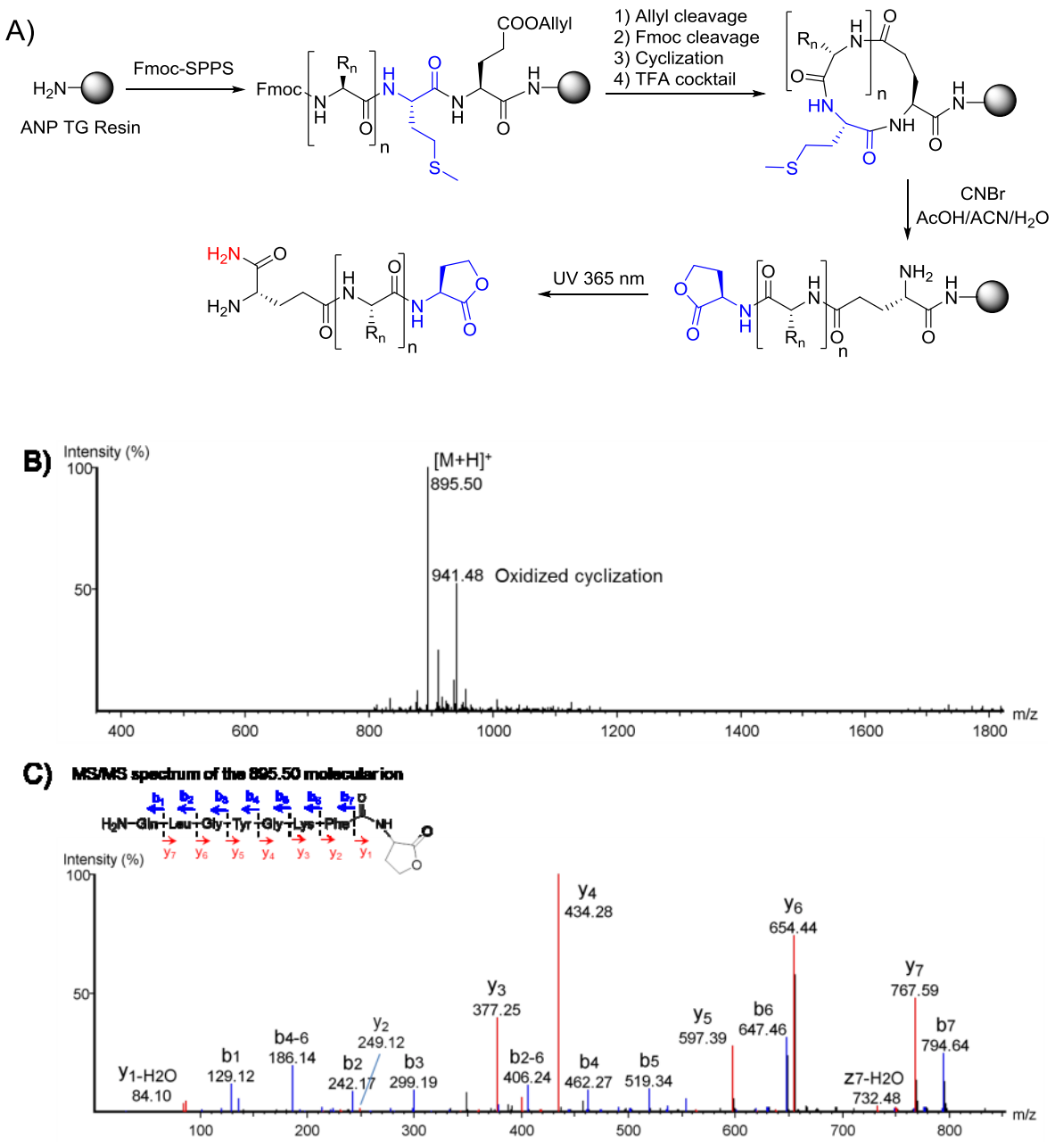
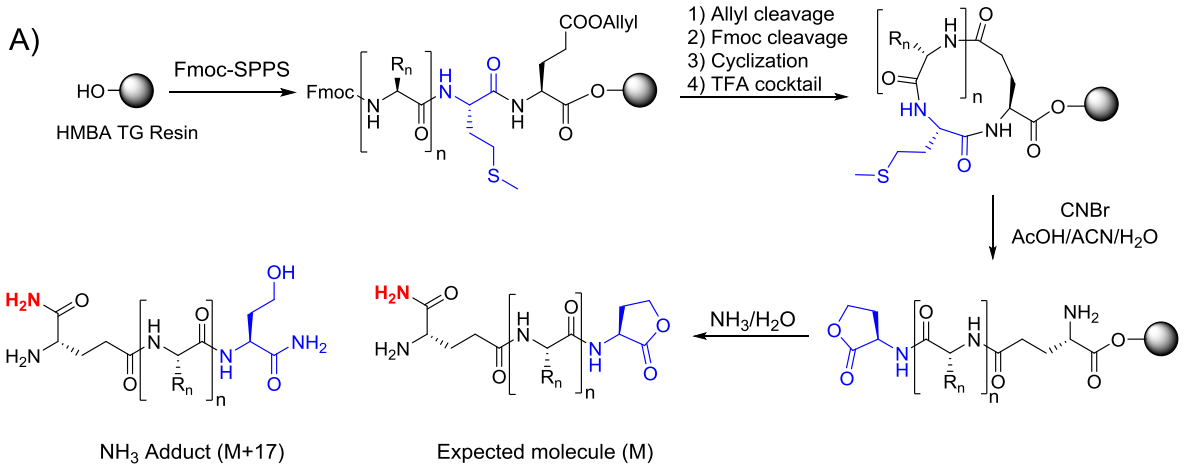
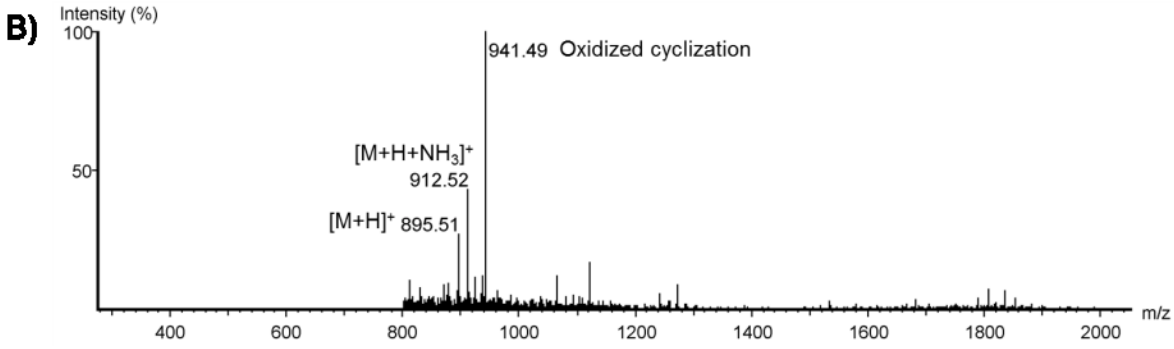


Figure 1. Combination of methionine as ring-opening residue and ANP as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H]^+$ 895.50 Da molecular ion.

In the second approach (Met/HMBA), methionine was placed as above in the peptide sequence and HMBA used as a linker (Figure 2A). The macrocycle was opened by treatment with CNBr and the bead cleaved with gaseous ammonia to yield a linear peptide containing a C-terminal homoserine lactone and an N-terminal side chain anchored glutamine as above. In this case, in addition to the expected molecular ion at 895.51 Da, a peptide bearing an aminolyzed (reopened) homoserine lactone at 912.52 Da (M+H+17) was observed in the MS spectrum. Moreover, the peak corresponding to the oxidized cyclic peptide was also observed (Figures 2B and 2C). The linearized peptide with the reopened homoserine lactone could be sequenced both manually and by De Novo sequencing with the PEAKS software.





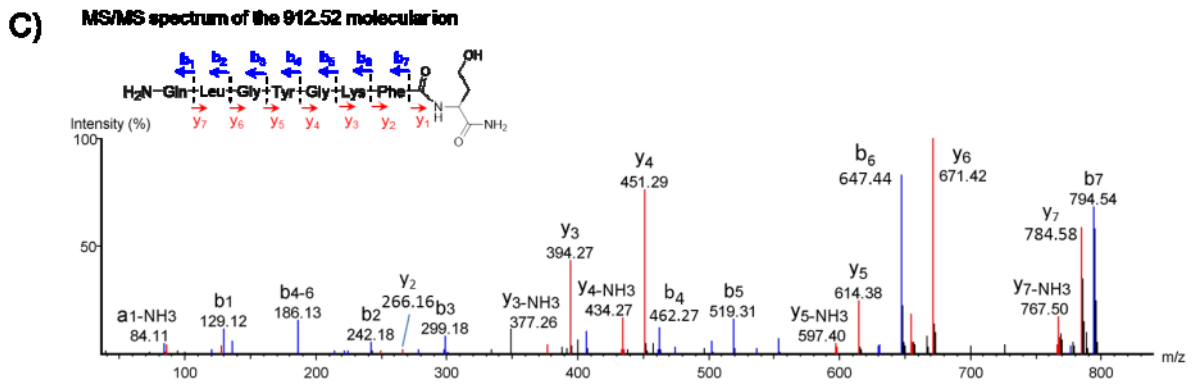
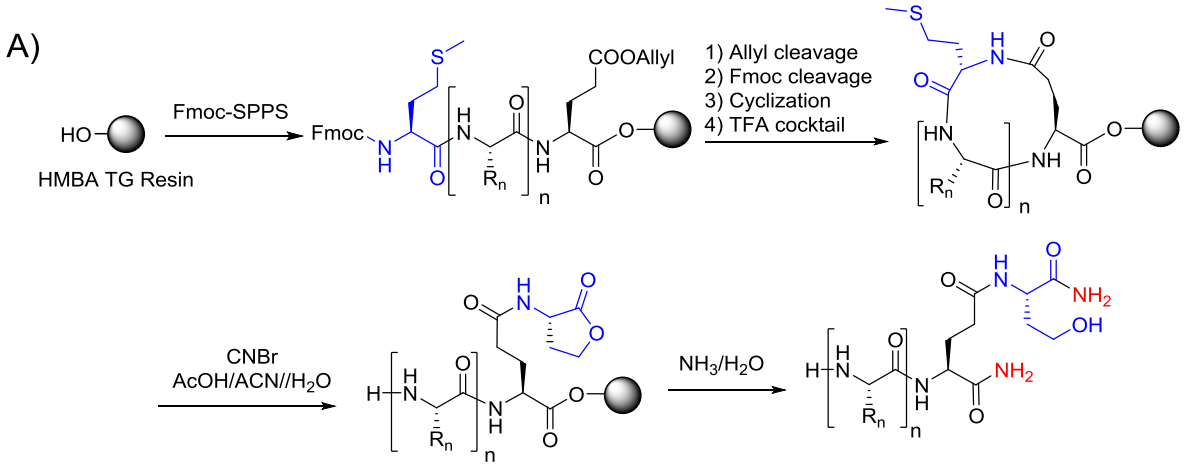


Figure 2. Combination of methionine as ring-opening residue and HMBA as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H+17]^+$ 912.52 Da molecular ion.

In the third approach (N-Met/HMBA), methionine was placed as ring-opening residue at the N-terminal position prior to cyclization and HMBA used as linker (Figure 3A). After ring-opening with CNBr and cleavage from the bead with gaseous ammonia, a linear peptide with a C-terminal glutamyl-homoserine lactone dipeptide was obtained. A homoserine residue from the aminolysis of the homoserine lactone with ammonia was also obtained in this approach. Here again, an oxidized molecular ion $M+NH_3+16$ was observed in the MALDI MS spectrum (Figure 3B). The linearized peptide bearing the homoserine lactone at 895.50 Da could be sequenced both manually and by De Novo sequencing with the PEAKS software.



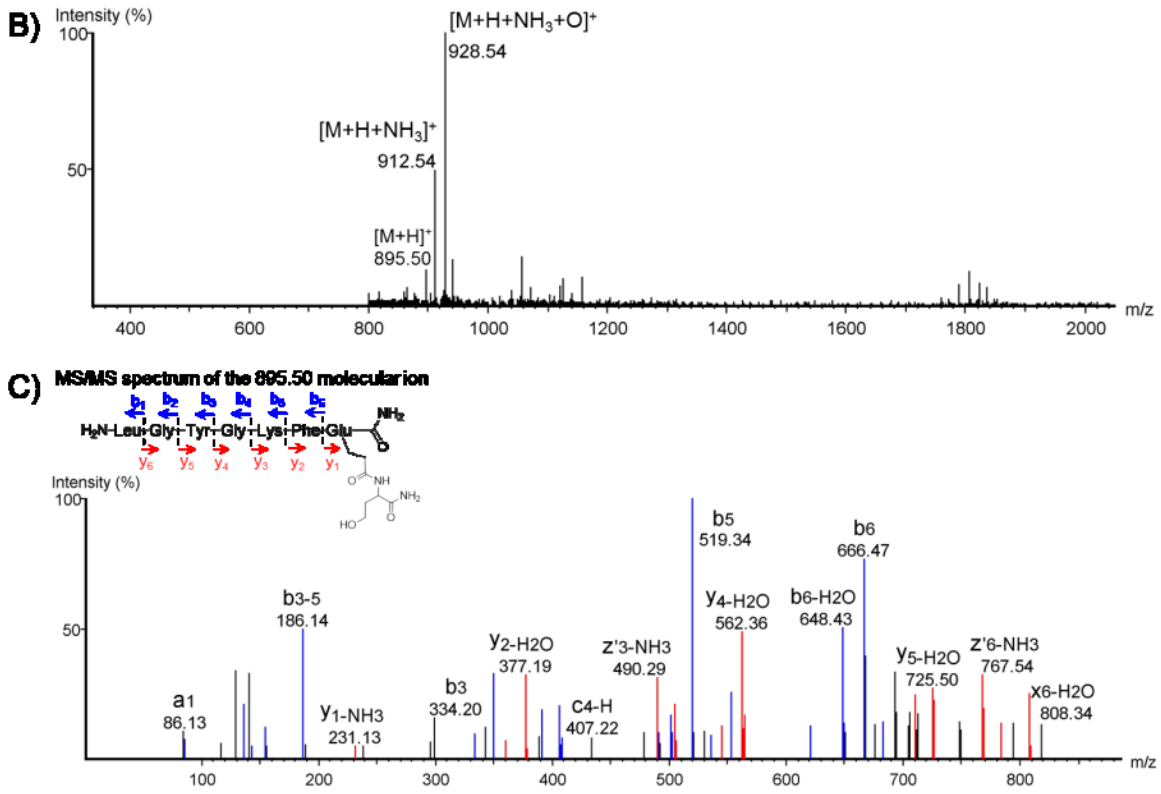
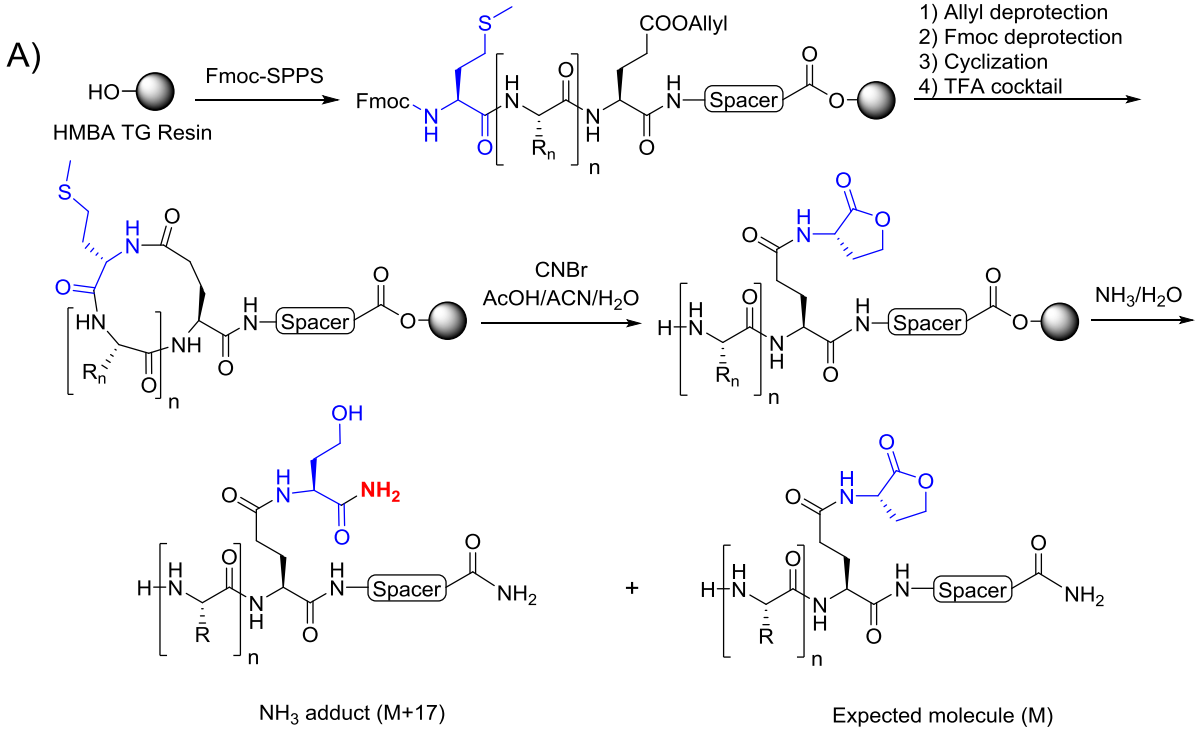


Figure 3. Combination of methionine as ring-opening residue at the N-terminal position and HMBA as linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the $[M+H]^+$ 895.50 Da molecular ion.

For the screening of OBOC libraries, it is highly recommended to insert a spacer between the bead and the peptide sequence. The spacer is generally used to keep the compound to be screened from the bead and facilitate its binding with the target macromolecule. In order to simulate real screening conditions, a spacer was added for the following approaches. In the fourth approach (N-Met/HMBA spacer), Met was placed at the N-terminal position as above and a Leu-βAla-βAla-Lys-Gly (LBBKG) spacer sequence added to the HMBA linker (Figure 4A). Following ring-opening with CNBr and cleavage from the bead with gaseous ammonia, a linear peptide containing a glutamic acid residue bearing a homoserine lactone on its side chain was obtained. As observed in the previous approach, the homoserine lactone was also partly aminolyzed during the cleavage step. MALDI MS analysis showed the presence of four different molecular species including the expected mass (M+H = 1335.83 Da), the aminolyzed peptide (M+H+17 = 1352.85 Da), the cyclic peptide (1365.82 Da) and the oxidized cyclic peptide (1381.81 Da) (Figure 4B). The molecular ion of the aminolyzed specie at 1352.85 Da could be sequenced both manually and by De Novo sequencing with the PEAKS software (Figure 4C).



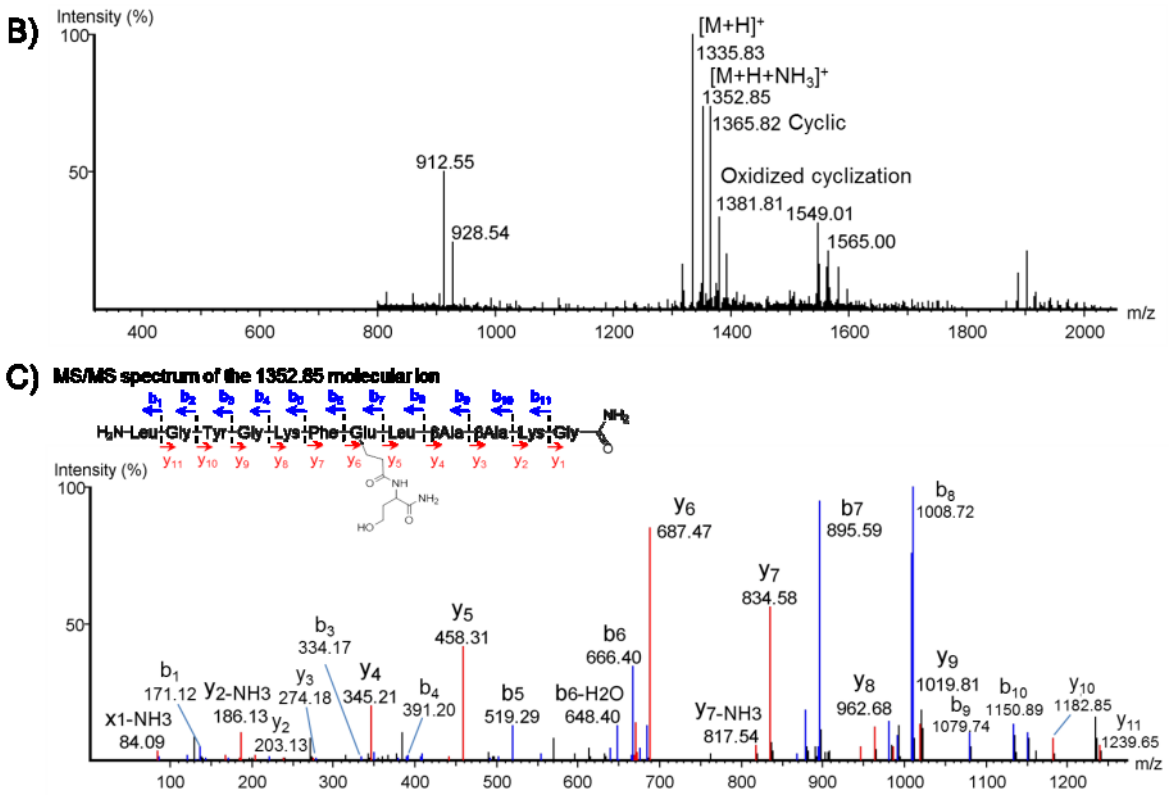


Figure 4. Combination of methionine as ring-opening residue at the N-terminal position and spacer sequence added to the HMBA linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the aminolyzed specie $[M+17+H]^+$ 1352.85 Da. (Spacer = LBBKG)

In the fifth approach (N-Met/ANP spacer), the HMBA linker was replaced by ANP (Figure 5A). After ring-opening with CNBr and cleavage from the bead upon UV irradiation, the generated linear peptide contains a homoserine lactone on the side chain of a glutamic acid residue in the middle of sequence which increases the sequencing difficulty. Here again the oxidized cyclic peptide was obtained in important proportion but the expected molecular ion at 1335.83 Da was observed in the MALDI MS spectrum (Figure 5B). In this case, the molecular ion could be sequenced both manually and by De Novo sequencing with the PEAKS software (Figure 5C).

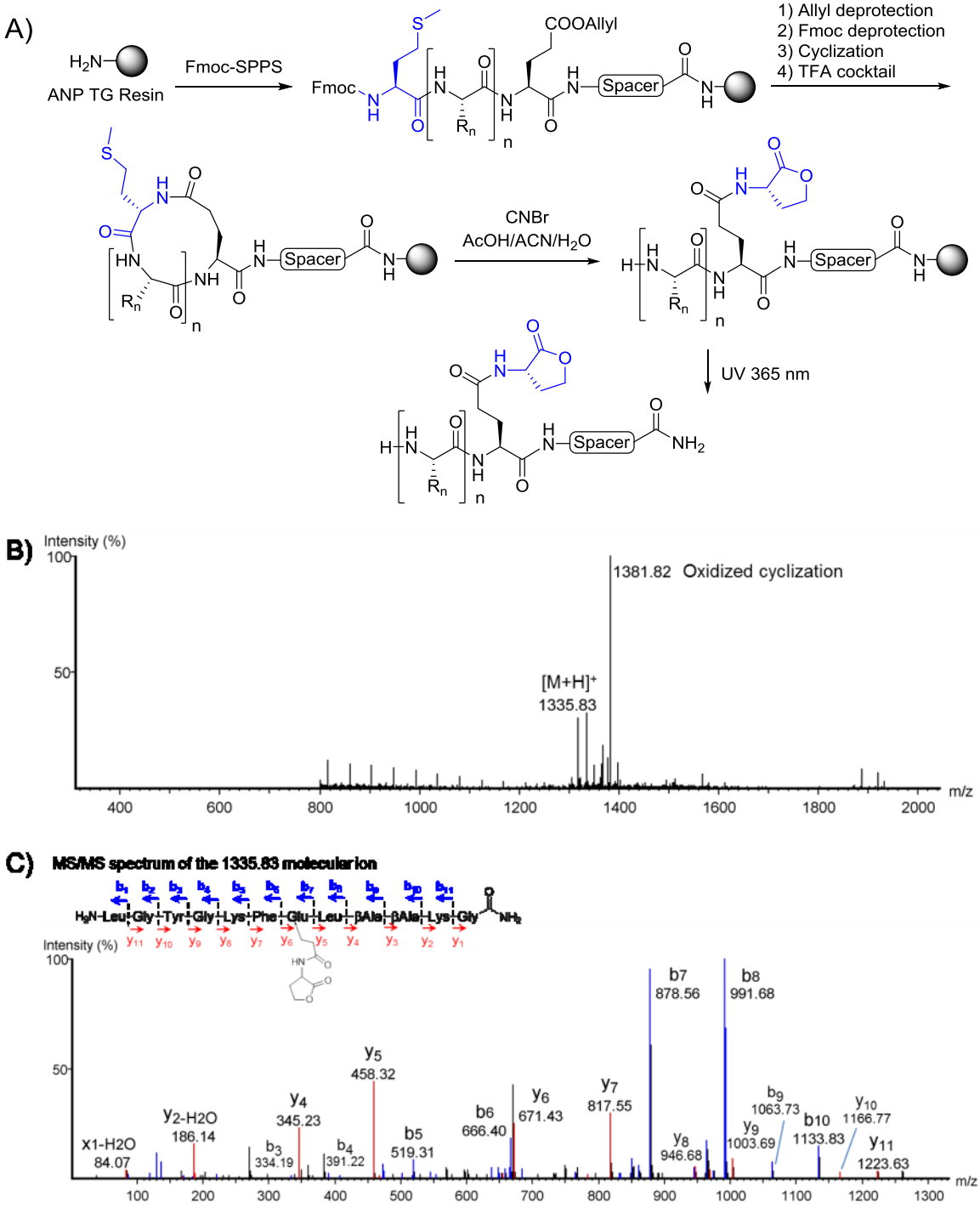
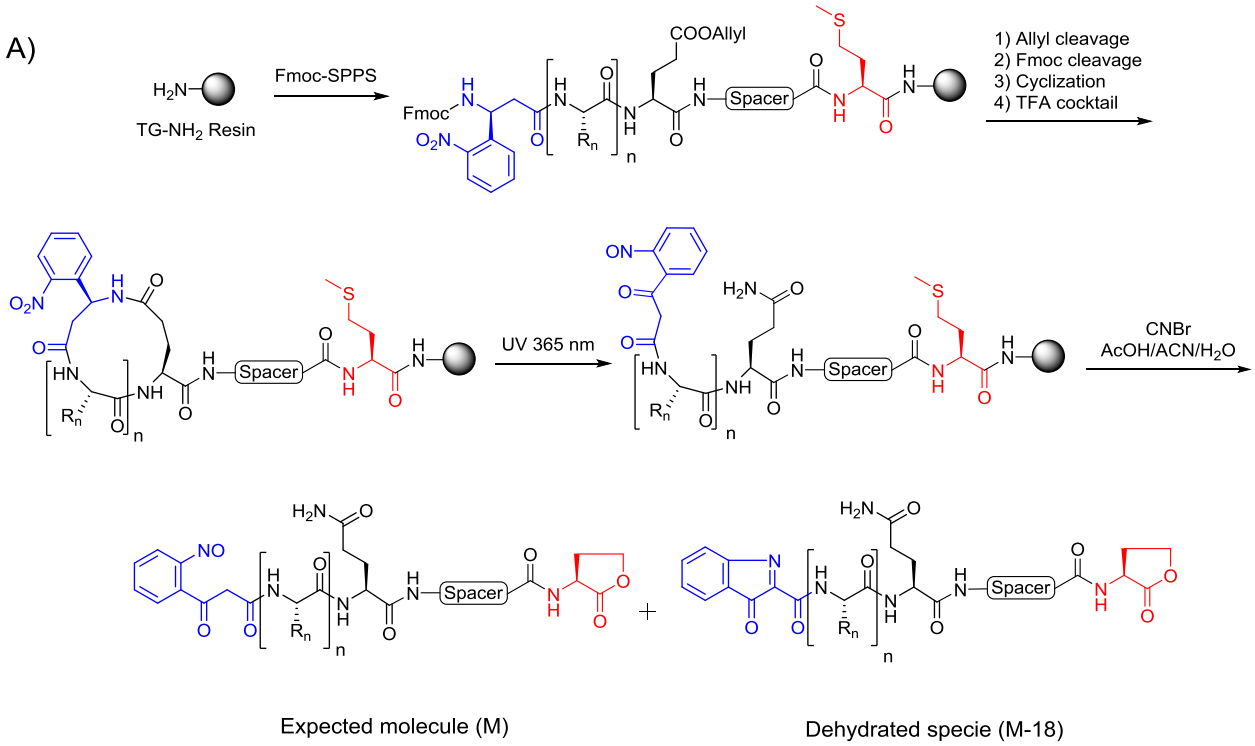


Figure 5. Combination of methionine as ring-opening residue at the N-terminal position and spacer sequence added to the ANP linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the molecular ion $[M+H]^+$ 1335.83 Da. (Spacer = LBBKG)

In the last three approaches, an ANP residue was placed at the N-terminal position prior to cyclization and different linker were used. In the sixth approach (N-ANP/Met), the peptide sequence was built on a methionine spacer (Figure 6A). The ring was first opened by UV irradiation and the linearized peptide cleaved with CNBr. Since the ring-opening was performed under UV irradiation 3 hours without any inert gas protecting, there is a high risk of oxidation of the Met linker. Once the Met is oxidized the cleavage step cannot be performed. As shown in the MALDI MS spectrum, the expected species of the model peptide at 1435.73 or 1453.74 Da were not observed suggesting that the Met linker has been oxidized during UV irradiation (Figure 6B). Therefore the peptides could not be sequenced by MS/MS. On the other hand, the species generated by this approach are all C- or N-terminally capped which significantly decrease the ionization intensity and considerably complicate MS/MS analyses.



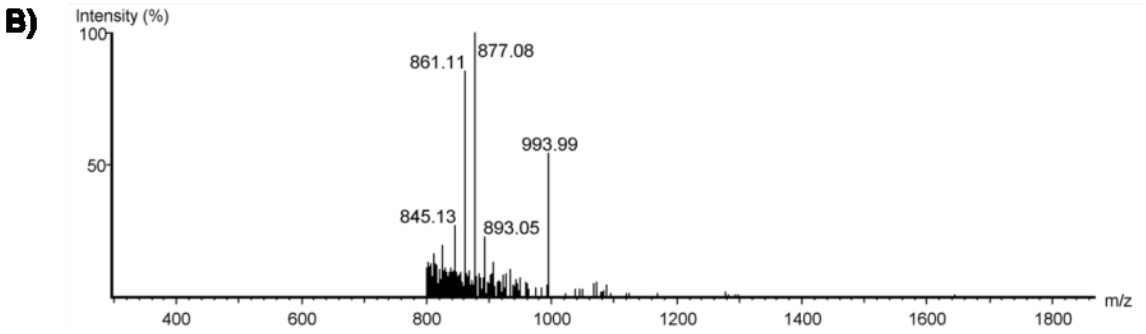


Figure 6. Combination of ANP as ring-opening residue at the N-terminal position and Met as linker with MALDI-TOF MS spectrum. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead. (Spacer = LBBK)

In the seventh approach (N-ANP/HMBA spacer), the Met linker was replaced by HMBA (Figure 7A). After ring-opening with UV irradiation, the linearized peptide was released from the bead by treatment with gaseous ammonia overnight. In this case the peak corresponding to the expected peptide containing an N-terminal nitroso compound was not observed (Figure 7B). Instead three peaks corresponding to dehydrated species at 1392.83 and 1408.83 Da and one adduct at 1423.84 Da were noticed in the MS spectrum. We found out that all these modifications were happening on the N-terminal ANP residue and caused by reactions of the nitroso followed by different rearrangements. These modifications will be discussed in more details in Chapter 5. Nevertheless the molecular ion at 1408.83 Da corresponding to the dehydrated specie could be sequenced both manually and by De Novo sequencing with the PEAKS software (Figure 7C).

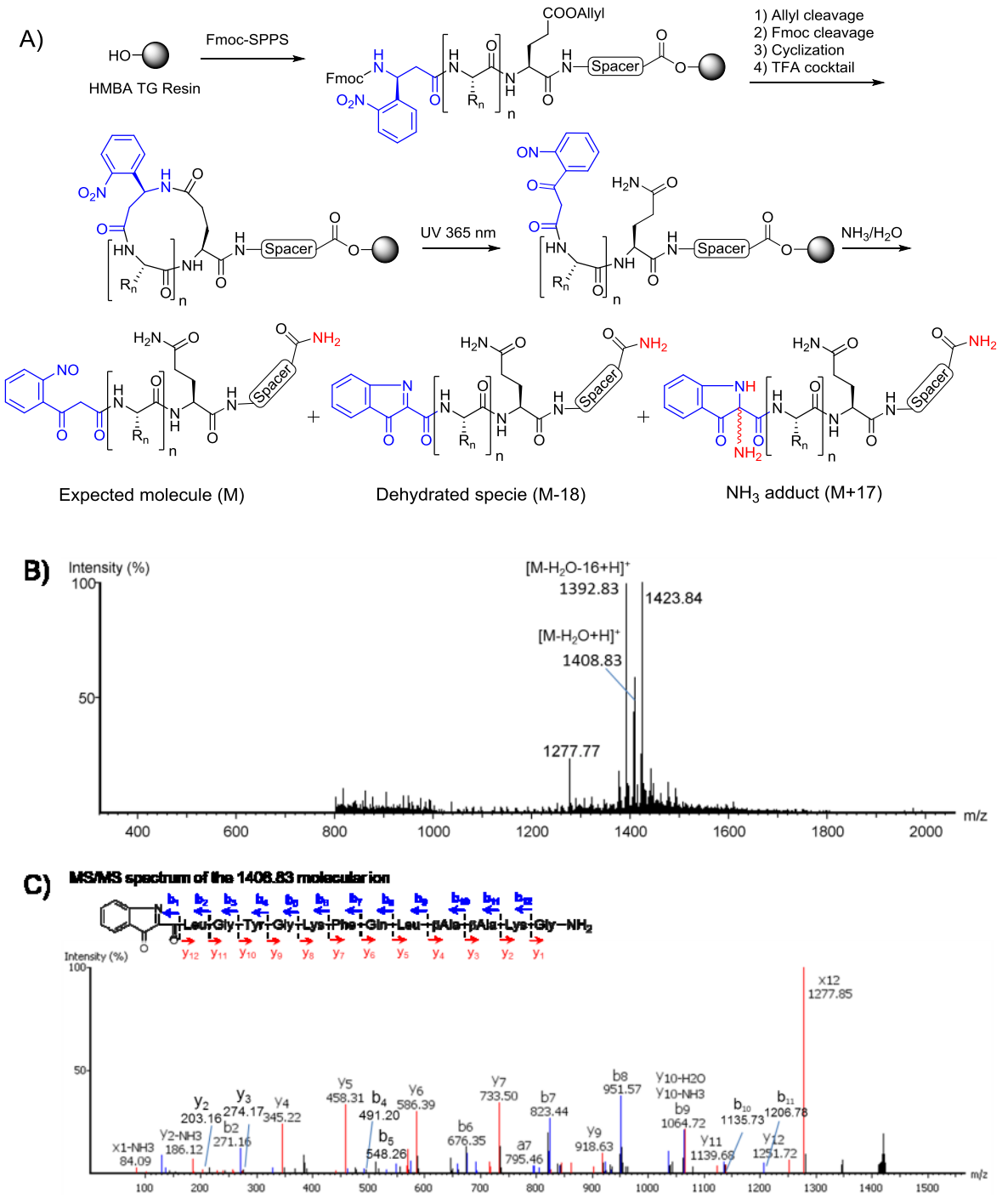
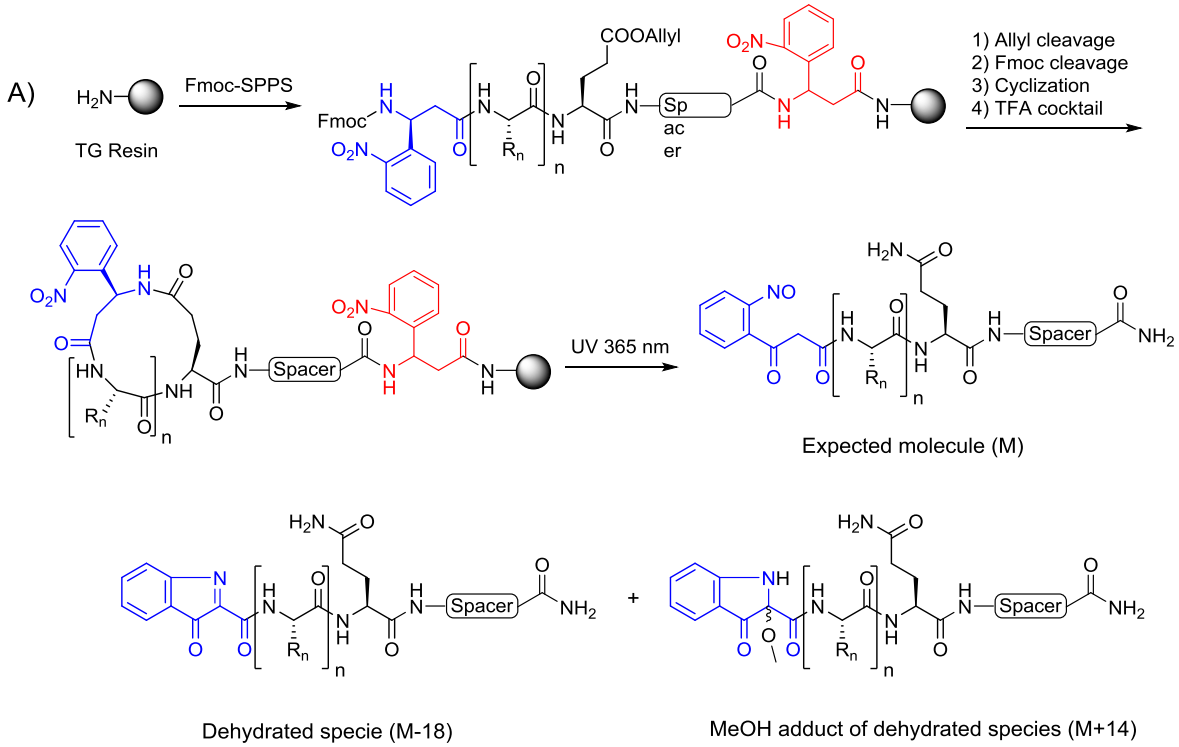


Figure 7. Combination of ANP as ring-opening residue at the N-terminal position and spacer sequence added to the HMBA linker with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead and C) MS/MS spectra of the molecular ion $[M\text{-}H_2O+H]^+$ 1408.83 Da. (Spacer = LBBKG)

In the last approach, the photolabile residue ANP within the macrocycle and as a linker to allow a simultaneous ring-opening and cleavage from the resin upon UV irradiation (Figure 8A). After ring-opening and cleavage by UV irradiation, MALDI TOF MS analysis of the released linear peptide showed two major products corresponding to the dehydrated specie ($M+H^+-18$) and to a MeOH adduct on the dehydrated specie ($M+H^+-18+MeOH$ or $M+H^++14$) (Figure 8B). Nevertheless the spectrum is quite clean and both molecular ions 1408.84 and 1440.84 Da could be sequenced by MS/MS (Figures 8C and 8D). Having two different peaks that can be used for sequencing can be very advantageous to identify compounds from OBOC libraries and confirm the sequence when both give the same results. Performed in a single step, this dual approach is very attractive and promising. Therefore, we have investigated the efficiency of this approach on OBOC cyclic peptide libraries and the results are presented in Chapter 5.



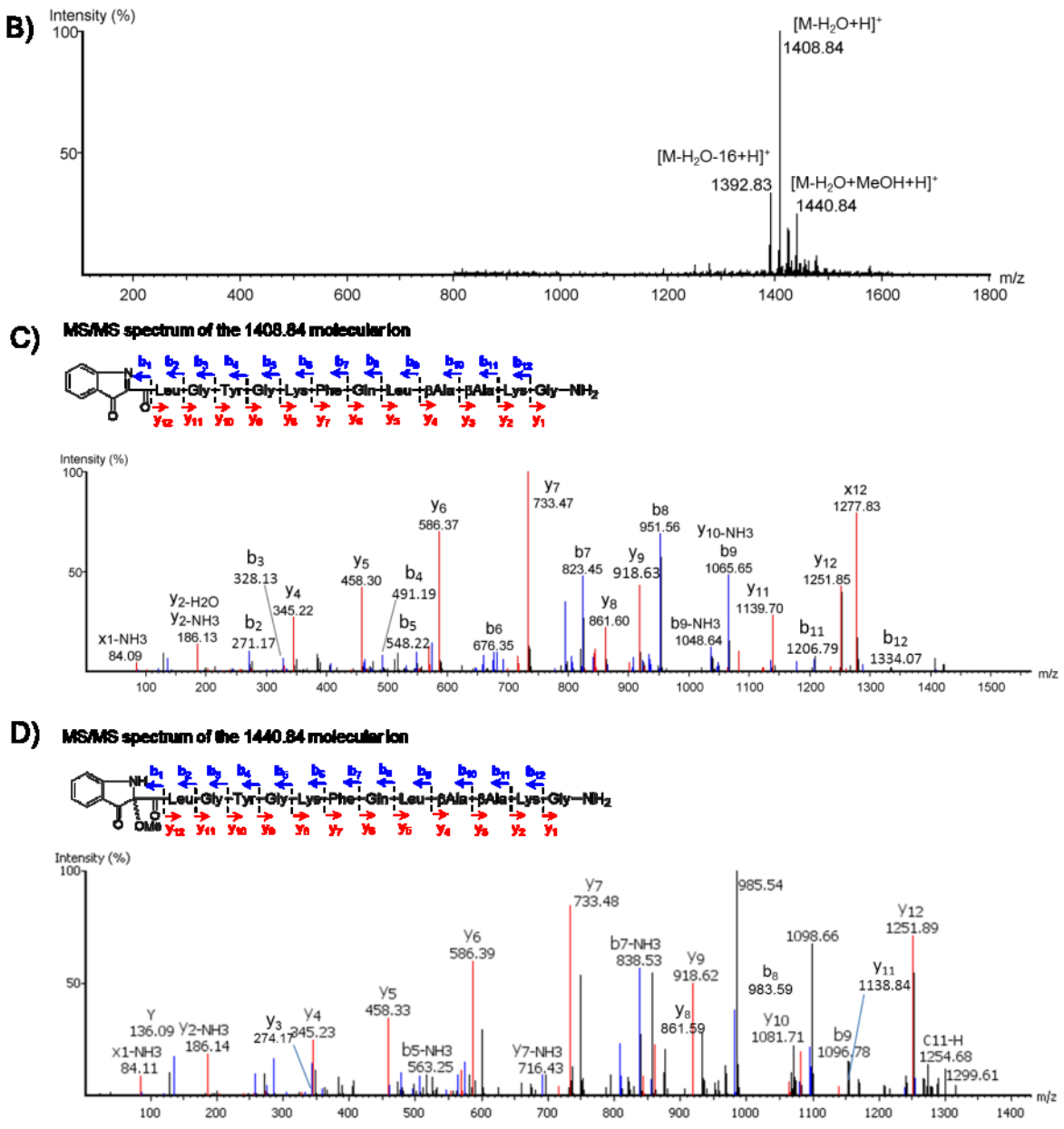


Figure 8. Use of ANP residue in the macrocyle and linker for dual ring-opening and cleavage with MALDI-TOF MS and MS/MS spectra. A) Synthesis of the cyclic peptides and ring-opening/cleavage. B) MALDI-TOF MS spectrum of model peptide obtained after ring-opening/cleavage on a single bead. C) MS/MS spectra of the dehydrated molecular ion $[M\text{-}H_2O+H]^+$ 1408.84 Da and D) MeOH adduct molecular ion $[M+14+H]^+$ 1440.84 Da. (Spacer = LBBKG)

3.4 Conclusion

Overall eight different combinations of Met, ANP and HMBA as ring-opening residue or linker have been investigated for their compatibility with other residues and efficiency in a ring-opening/cleavage strategy. Among the tested combinations, only the ANP/Met approach did not yield a sequenceable linear peptide. We also observed that the Met linker can be useful but oxidation of his side chain during the manipulation and storage is a major problem as it prevents efficient cleavage. Therefore it would be strongly recommended to store a library containing Met as a ring-opening residue or linker under inert and to perform screening and sequencing as fast as possible to avoid insufficient amount of peptide for MS analysis or absence of ring-opening. Otherwise it is possible to reduce oxidized methionine side chains with a dimethylsulfide solution prior to the ring-opening/cleavage reactions. An advantage of using Met as a ring-opening residue combined to another type of linker is that after treatment with CNBr, the reagent can be eliminated by simple filtration. As CNBr is very corrosive and its handling hazardous, avoiding its evaporation can significantly prolonged vacuum pump and material lifetime. The HMBA linker allows cleavage in mild conditions and is quite efficient. Works with this linker in OBOC libraries are currently underway in our laboratory. The ANP residue is very attractive as it is orthogonal to many protecting groups and linkers. In this study we observed that when ANP is used within the macrocycle, side reactions and subsequent rearrangements upon UV irradiation can complicate MS and MS/MS analysis. Nevertheless we were able to identify these derivatives and sequence some of them. In the case of the dual ring-opening/cleavage approach, it was very interesting to sequence two molecular ions from the same experiment and deduce the same sequence. These carbon copies could be very useful to confirm and increase the confidence level in peptide sequence identification.

In this study we showed that many ring-opening/cleavage approaches are feasible when Met, ANP and HMBA are combined with each others. One of the tested approaches was particularly attractive as it allows a simultaneous ring-opening of the macrocycle and cleavage from the resin in a single step. Because of this convenience, dual ring-opening/cleavage approaches were investigated more deeply and are presented in Chapters 4 and 5.

3.5 Experimental Section

Peptide Synthesis. Peptides were synthesized by standard Fmoc solid phase synthesis. Briefly, amino acid couplings were performed with a solution of Fmoc-Xaa-OH (3 equiv), HCTU (3 equiv) and NMM (6 equiv) in DMF for 20 min. The coupling step was repeated once and the resin washed with DMF (5 × 30 s). The Fmoc protecting group was removed by treating the resin twice with a solution of 20% piperidine in DMF (v/v) for 10 min followed by washing with DMF (5 × 30 s).

On resin peptide cyclization. The resin was swelled with CH_2Cl_2 and the allyl protecting group was removed with a solution of $Pd(PPh_3)_4$ (0.24 equiv) and $PhSiH_3$ (20 equiv) in CH_2Cl_2 for 2 × 30 min. The resin was washed with CH_2Cl_2 (5 × 30 s), 0.5% DIPEA (v/v) in DMF (2 × 2 min), 0.5% diethyl dithiocarbamate (v/v) in DMF (2 × 5 min) and DMF (5 × 30 s). The Fmoc group was then removed with 20% piperidine in DMF (2 × 10 min) and the resin washed with DMF (5 × 30 s). Peptide cyclization was performed on solid support in the presence of HATU (3 equiv), HOAt (3 equiv) and NMM (6 equiv) in DMF for 3 h. Cyclization was monitored by the chloranil test. After reaction completion, the resin was washed with DMF (5 × 30 s) and CH_2Cl_2 (5 × 30 s).

Side-chain deprotection on TentaGel resin. The resin was swelled with CH_2Cl_2 and side chain deprotection was performed with a mixture of TFA, water and triisopropylsilane (TIS) (95:2.5:2.5) for 3 h followed by washing with CH_2Cl_2 (5 × 30 s).

Ring-opening reaction and cleavage from TentaGel resin. Single beads were isolated randomly under microscope and placed in microcentrifuge tubes. The following manipulation is different depending on the different approaches. ***a) ANP*** Single beads were swelled into MeOH (200 μL). Then, a UV lamp was placed at 2 cm above microcentrifuge tubes and irradiation conducted at 365 nm for 3 h. MeOH was added every 1-2 h to avoid extensive evaporation and prevent the drying. If ANP was as ring-opening residue, send beads to cleavage step after filtration, washing and dryness. If ANP was as cleavage linker, the extra solvent was evaporated under vacuum and the peptides released from the beads were sent for MS/MS analysis. ***b) Met*** The bead was treated overnight with 20 μL of CNBr (40 mg/ mL) in $CH_3CN/AcOH/H_2O$ (5:4:1) in the dark. If Met was as ring-opening residue, send beads to cleavage step after filtration, washing and dryness. If Met was as cleavage

linker, the extra solvent was evaporated under vacuum and the peptides released from the beads were sent to MS/MS analysis. ***c) HMBA*** Single beads were placed in a sealed desiccator with a beaker containing NH_4OH overnight. Released peptides were eluted from bead by 10 μL MeOH. After drying the solvent, the beads were sent to MS/MS analysis.

MALDI-TOF MS Analysis. The peptides released from the bead were dissolved in 10 μL of 0.1% TFA in water. 1 μL of the peptide solution was mixed with 1 μL of 4-hydroxy-α-cyanocinnamic acid (5 mg/mL) in CH_3CN/0.1% TFA (1:1) and 1 μL of the mixture was spotted onto a MALDI sample plate for MS/MS analysis. The PEAKS Studio software was used for spectra analysis and DE NOVO sequencing.

Chapter 4

Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One Compound Libraries

Forward

Cyclic peptides are useful tools in chemical biology and drug discovery with a great therapeutic potential. The sequencing of hit compounds after the screening is the bottleneck in one-bead one-compound combinatorial chemistry library. In order to sequence simply and effectively, the objective of chapter 3 is to develop an approach for sequence determination after simultaneous macrocycle opening and release from solid support.

My contribution of this article was to perform cyclic peptides which including a methionine residue in the macrocycle and as a linker to introduce a tandem ring-opening/cleavage from resin. With Anick Girard's helping, a master student, I synthesized a small cyclic heptapeptide library to demonstrate the compatibility of the approach. I also did peptide sequences analysis with a de novo sequencing software. Finally, I drafted the preliminary manuscript and Éric Biron edited the final version. This article was published in the journal ACS Combinatorial Science (September 9th, 2013).

Résumé

L'utilisation de chimiothèques de peptides cycliques par la technique « one-bead-one-compound » se retrouve limitée par la difficulté de séquencer les composés positifs. Pour contourner ce problème, nous proposons une approche de réouverture de cycle et de clivage en tandem avec le bromure de cyanogène. En effet, en utilisant une méthionine dans le cycle et comme ancrage à la résine, il devient possible de cliver et de rouvrir le cycle simultanément, ce qui permet le séquençage des composés initialement cycliques. L'utilisation d'un ancrage inversé relativement à la séquence peptidique facilite le séquençage en évitant la présence de deux homosérine lactones C-terminales. Après l'analyse MALDI-TOF MS/MS, les peptides clivés à partir de chaque bille ont été séquencés. La compatibilité de notre méthode a été démontrée avec la plupart des acides aminés naturels et permet le séquençage de peptides cycliques obtenus par la technique « one-bead-one-compound », éliminant ainsi le besoin d'encodage.

ACS Combinatorial Science

Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One-Compound Libraries

Xinxia Liang, Anick Girard, and Eric Biron*

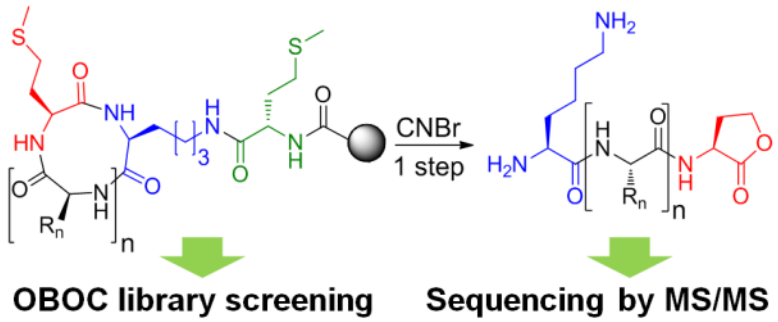
Faculty of Pharmacy, Université Laval, Pavillon Ferdinand-Vandry, Québec, Québec, G1V 0A6, Canada

Laboratory of Medicinal Chemistry, CHU de Québec Research Center (CHUL Section), 2705 Boulevard Laurier, Québec, Québec, G1V 4G2, Canada

*Corresponding author:
Prof. Éric Biron, Laboratory of Medicinal Chemistry, CHU de Québec Research Centre (CHUL Section), 2705 Boulevard Laurier, Québec, Québec, G1V 4G2, Canada, Phone: 1 418 524 2296; Fax: 1 418 654 2210, E-mail: Eric.Biron@pha.ulaval.ca

ABSTRACT

The use of cyclic peptides in one-bead-one-compound libraries is limited by difficulties in sequencing hit compounds. Lacking a free N-terminal amine, such peptides cannot be sequenced by the Edman degradation approach, and complex fragmentation patterns are obtained by tandem mass spectrometry. To overcome this problem, we designed an alternative approach introducing a methionine residue within the macrocycle and as a linker to allow simultaneous ring-opening and release from the resin upon treatment with cyanogen bromide. The methionine linker was inverted relative to the peptide chain to allow the synthesis of cyclic peptides anchored by a lysine side chain and to avoid the presence of two C-terminal homoserine lactones on the released linear peptides. After MALDI-TOF MS/MS analysis, the peptides released from a single bead were sequenced manually and with a de novo sequencing software. The strategy described herein is compatible with commonly used amino acids and allows sequencing of cyclic peptides in one-bead-one compound libraries, thus reducing the need for encoding.



KEYWORDS

Cyclic peptide libraries, one-bead-one-compound libraries, peptide macrocycles, ring-opening, peptide sequencing

Combinatorial chemistry has become a powerful tool for discovering potent and selective bioactive compounds for therapeutic and diagnostic applications, and is now an important component of the drug discovery process. Among the different combinatorial methodologies used to prepare and screen large peptide libraries, the one-bead-one-compound (OBOC) approach is one of the most accessible and economical to discover new hits against a target of interest.[1−4] The OBOC approach exploits the split-and-mix method to generate combinatorial libraries in which each bead displays many copies of a unique chemical entity.[5−7] OBOC libraries containing thousands to millions of different compounds can be readily synthesized and screened simultaneously on-bead against a target of interest. Such libraries have been successfully used to discover ligands and modulators for a wide variety of macromolecular targets.[2, 3, 8−12]

Cyclic peptides represent an important class of privileged structures, and they have gained a lot of interest in drug discovery.[13, 14] Compared to their linear counterparts, cyclic peptides show a greater stability against proteases and a higher selectivity because of their increased conformational rigidity.[14, 15] The great degree of molecular complexity and diversity that can be accessed by simple changes in their sequence has prompted the use of cyclic peptides in combinatorial libraries. However, the use of cyclic peptides in OBOC libraries has been limited by difficulties in sequencing hit compounds after screening. Since they lack a free N-terminal amine, such peptides cannot be sequenced via Edman degradation, and yield excessively complex fragmentation patterns with tandem mass spectrometry (MS/MS).[16] Structural elucidation of the selected compounds is a critical step and in this regard, different encoding strategies have been developed for OBOC libraries. The ladder synthesis, in which a small fraction of the peptides are N-terminally capped at each coupling cycle during peptide synthesis before the final cyclization, could be considered an efficient method.[17] However, the latter approach has the disadvantage of displaying the ladder peptides on the bead surface together with the cyclic peptide, thereby causing interference by the coding tags during screening.[8,18] To circumvent this problem, Pei and co-workers used a one-bead-two-compound (OBTC) approach relying on topologically segregated bilayer beads.[19] This strategy offers the opportunity to expose the cyclic peptide on the bead surface for screening while its linear counterpart for sequencing purposes is

found inside the bead.[4,12,20–22] The main disadvantage of this approach is that the cyclization step cannot be monitored because of the presence of the linear peptide tag.

More recently, a strategy based on the "reopening" of the macrocycle after the screening step has emerged to eliminate the need for encoding OBOC libraries.[23, 24] In this approach, a cleavable residue is introduced in the cycle backbone to allow linearization of the molecule under specific conditions and sequencing of the linear variant by MS/MS. Lim and coworkers reported a ring-opening approach involving the introduction of a cleavable alkylthioaryl bridge in peptoid macrocycles to allow linearization of the molecule by oxidation of the thioether, followed by nucleophilic displacement of the sulfone.[23] Using a similar strategy, Simpson and Kodadek inserted a methionyl residue within a peptoid macrocycle and as a linker to allow the simultaneous linearization and release of the compound from the bead upon treatment with CNBr.[24] A major asset of the ring-opening strategy is that, in contrast to OBTC and ladder synthesis methods, the same chemical entity is displayed inside and on the surface of the bead, eliminating the risk of interference by the coding tags during the screening. Moreover, the cyclization step can be easily monitored to ensure a complete conversion of the linear variant. On the basis of these elegant strategies with peptoids, we looked for an efficient and single step approach compatible with free amino acid side chains and that would allow simultaneous ring-opening of the cyclic peptides and cleavage from the resin. Herein, we report an alternative, straightforward version of the ring-opening approach for a fast and simple sequence determination of cyclic peptides from OBOC libraries.

The cleavable residue plays a critical role in the ring-opening approach and among the different linkers and cleavable residues that are readily available, we were particularly interested in methionine (Met). As any amino acid, Met can be used in standard solid-phase peptide synthesis and its impact on peptide conformation will be limited compared to extended aromatic residues and linkers. Moreover, Met is stable in the acidic, basic or reductive conditions commonly used to remove side chain protecting groups. Finally, the reaction conditions used to cleave Met residues are selective and compatible with other free amino acid side chains. Indeed, Met has been widely used as a linker in OBOC peptide

libraries and can be selectively cleaved upon treatment with CNBr to yield a C-terminal homoserine lactone.

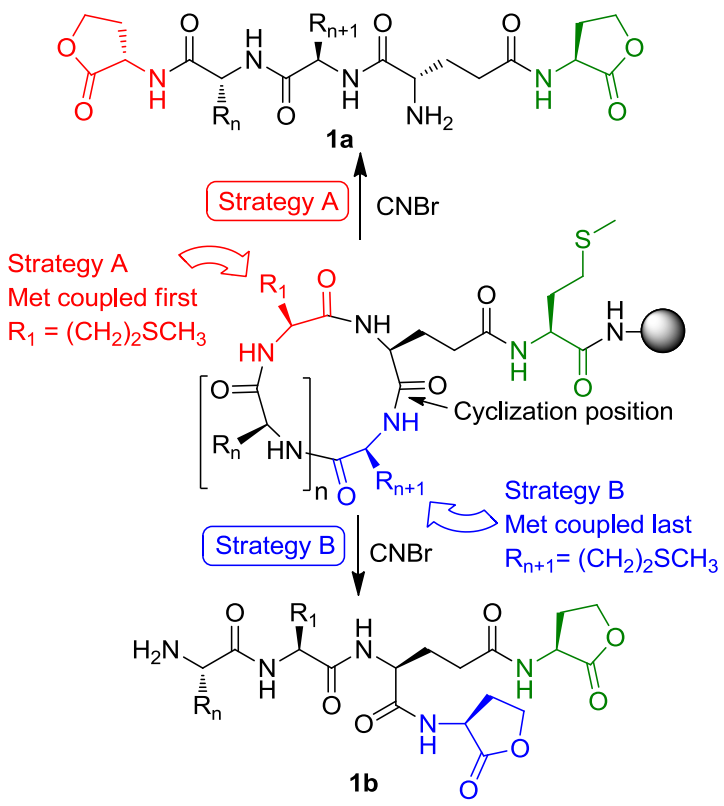


Figure 1. Incorporation of a Met residue in the macrocycle at different positions and the corresponding linear peptide generated after ring reopening and cleavage from the resin.

Our initial strategy to prepare OBOC cyclic peptides libraries was to use a solid supported Met bearing a side chain anchored Fmoc-Glu-OAll and to perform head-to-tail cyclization after peptide synthesis and selective deprotection of the carboxylic acid. In the ring-opening strategy, the Met residue can be introduced at two different positions in the macrocycle to allow its linearization. The Met residue can be incorporated either as the first residue following the side chain anchored Glu or as the last amino acid of the chain before cyclization (Figure 1). Depending on the Met position, the pattern of the linearized peptides thus yielded will be completely different. First, when Met is coupled directly to the Glu anchor before library synthesis, a linear peptide **1a** with two C-termini each bearing a homoserine lactone is obtained after cleavage with CNBr. Unfortunately, we observed that

linear peptides with such a pattern can entail problematic sequencing even when their amino acid composition is known. Next, when Met is coupled as the last residue during peptide elongation prior to cyclization, a linear peptide **1b** with a C-terminus bearing two homoserine lactones is generated. In this case, as observed by Simpson and Kodadek, [24] the cyclization efficiency was moderate and complete ring formation could not be obtained when Met was the N-terminal amino acid. The latter behaviour was not observed using the former strategy. Moreover, a linear peptide bearing two homoserine lactones was also generated.

To circumvent the problems associated with the presence of two homoserine lactone residues for the sequencing of the linearized peptide and the moderate cyclization efficiency obtained when Met is at the N-terminal position, an alternative approach was designed. Our strategy involved the inversion of the Met handle in order to eliminate one of the homoserine lactones in the linearized peptide combined with the introduction of the cleavable Met residue at the C-terminal position (Figure 2). Initially reported by Kappel and Barany for the synthesis of lysine-containing head-to-tail cyclic peptides,[25] inversion of the Met handle allows the release of a free amino group while the homoserine lactone remains attached to the resin. Therefore, the Fmoc-Glu-OAll anchor was replaced by Fmoc-Lys-OAll as a side chain anchored amino acid to allow on-resin peptide cyclization after standard solid-phase peptide synthesis and appropriate deprotection (Figure 2). The introduction of a Met residue directly next to the side chain anchored Lys in the cyclic peptide allows simultaneous ring-opening and peptide cleavage from the resin upon treatment with CNBr, yielding a linear peptide bearing a single C-terminal homoserine lactone and a free N-terminal lysine. The invariable residues at the C- and N-termini can be used as starting points in MS/MS spectrum analysis, thereby significantly helping the sequencing process. Moreover, because of its positive charge, the lysine residue facilitates ionization of the peptide during MALDI MS analyses.

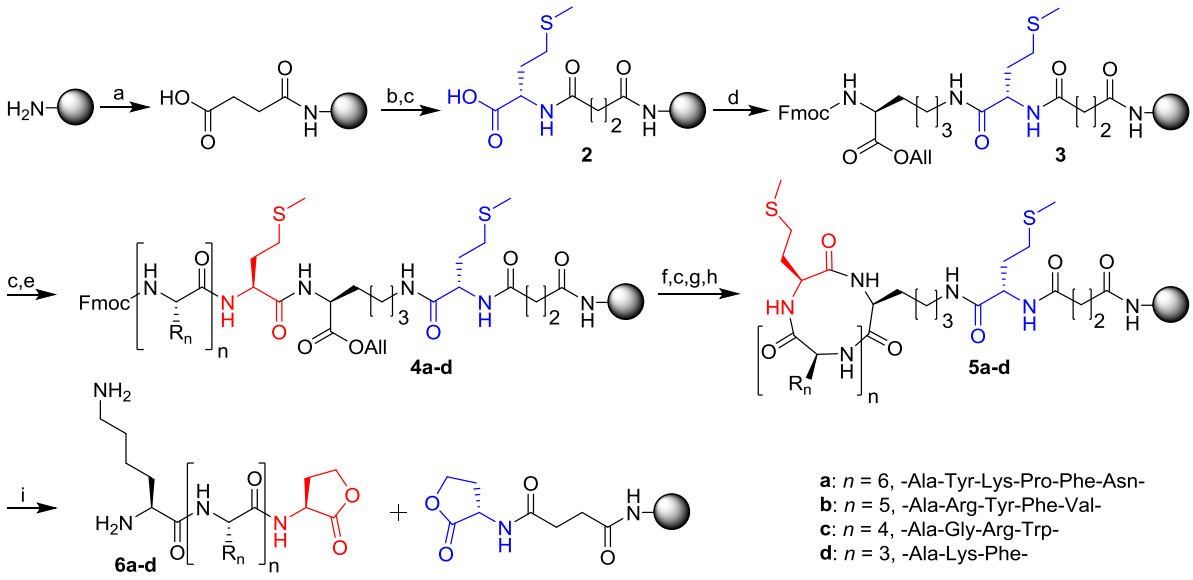


Figure 2. Design and synthesis of cyclic peptides for the tandem ring-opening/cleavage approach. Reagents and conditions: (a) succinic anhydride, DIPEA, DMF; (b) H-Met-OFm, HATU, DIPEA, DMF; (c) 20% piperidine/DMF; (d) Fmoc-Lys-OAll, HATU, DIPEA, DMF; (e) standard Fmoc solid-phase peptide chemistry with HCTU; (f) $Pd(PPh_3)_4$, $PhSiH_3$, CH_2Cl_2; (g) PyBOP, HOBt, DIPEA, DMF; (h) TFA/H_2O/TIS (95:2.5:2.5); (i) CNBr, CH_3CN/AcOH/H_2O (5:4:1).

To evaluate the efficiency of the proposed strategy, four model peptides of different size were prepared on Rink Amide AM resin and TentaGel S-NH_2 (TG) resin bearing an inverted Met handle **2** (Figure 2). The latter were prepared by treatment of the amino functionalized resins with succinic anhydride in the presence of DIPEA followed by coupling of H-Met-OFm with HATU on the pending carboxyl sites. After removal of the 9-fluorenylmethyl ester with piperidine in DMF, Fmoc-Lys-OAll was anchored via its side chain to the C_α-carboxyl of the resin-bound Met residue with HATU to yield resins **3**. The loadings were determined by quantification of the Fmoc group with the UV method and found to be 0.43 mmol/g and 0.22 mmol/g for the Rink Amide AM and TG resins, respectively.[26] After Fmoc group removal with piperidine in DMF, Fmoc-Met-OH was coupled to the free amino group and the peptides were assembled by standard Fmoc solid-phase synthesis with HCTU as coupling reagent. Afterward, the C-terminal allyl ester of the Lys anchor was selectively cleaved with $Pd(PPh_3)_4$ followed by removal of the N-

terminal Fmoc group. Head-to-tail peptide cyclization was performed with PyBOP, and progress of the reaction monitored using the chloranil test.[27] To confirm the presence of cyclic peptides, the Rink Amide AM resin was treated with a TFA cocktail to afford the fully deprotected cyclic peptides **5a−d** in solution. HPLC and MS analyses confirmed transformation of the linear peptides into their cyclic counterparts (Figure 3a and Supporting Information, Figure S5). Next, the tandem ring-opening/cleavage reaction was performed on the TG resin bearing the side chain deprotected cyclic peptides by treatment with a solution of CNBr in $CH_3CN/AcOH/H_2O$. The compounds thus released were analyzed by HPLC and MS, which showed that linear peptides **6a−d** have been generated, suggesting that the tandem reaction proceeded successfully (Figure 3b and Supporting Information, Figure S6).

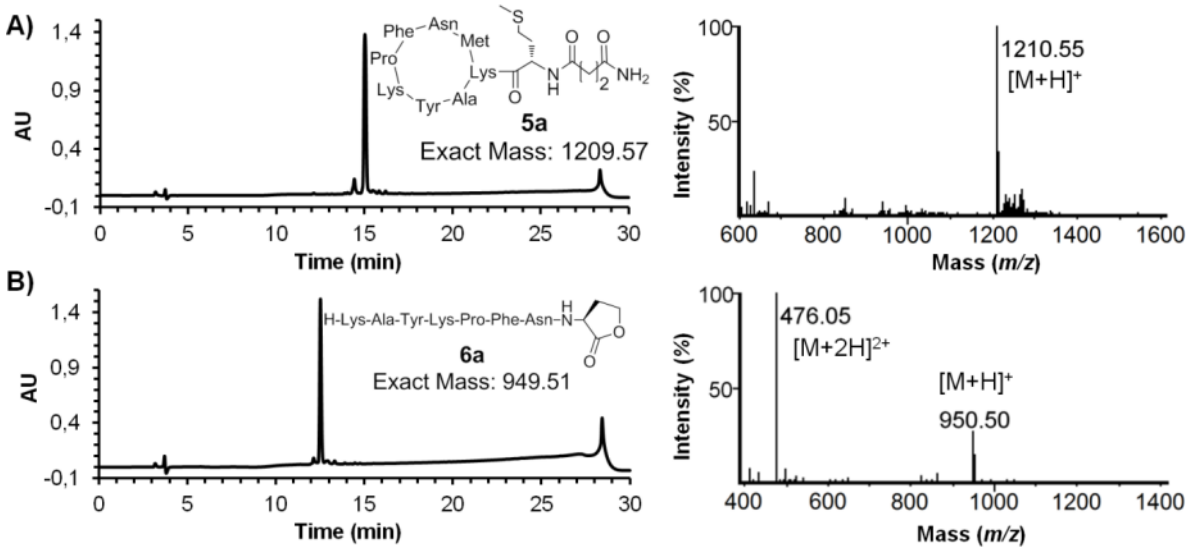


Figure 3. HPLC profiles (λ = 220 nm) and ESI-MS spectra of crude model peptides. (A) cyclo[AYKPFNMK(M-succinamide)] **5a** and (B) H_2N-KAYKPFNh* **6a** after tandem ring-opening/cleavage from TG resin. (h* = homoserine lactone)

To validate the efficiency of the approach in a one-bead-one-compound context, a single bead was picked up from the TG resins **5a−d** and treated with the CNBr solution. The crude products released from each single bead were immediately subjected to MALDI-TOF MS. The mass spectra of linearized peptides **6a−d** typically showed the expected molecular ion as the major peak and a sodium adduct at +22 Da (Figure 4 and Supporting Information,

Figure S7). In the case of peptide **6a**, a second unknown minor peak at −30 Da was observed. With peptide **6c**, a second minor peak at +16 Da showed oxidation of the Trp residue. MS/MS analysis of the molecular ions yielded high-quality spectra from which the linearized peptides **6a−d** could be sequenced manually and also by using de novo sequencing with the Peaks software (Figure 4a and Supporting Information, Figure S7).[28]

To demonstrate the compatibility of our strategy with OBOC libraries, a small cyclic heptapeptide library was prepared on 100 mg of TG resin bearing a side chain anchored Fmoc-Lys-OAll **3**. After coupling Fmoc-Met-OH to the resin, the library was prepared by split-and-pool synthesis using standard Fmoc/tBu solid-phase peptide chemistry. The next four positions within the peptide library are filled by a random combination of sixteen L-amino-acids. Met and Cys were excluded for reactivity reasons while Gln and Ile, which are isobaric with Lys and Leu, respectively, were not used for the sake of simplicity at the sequencing step. The penultimate position (i.e., relative to the Lys anchor found at the NH_2 end of the linearized heptapeptide) was randomly filled by either of three aliphatic L-amino acids with the least hindrance (namely, G, A, and L) to promote cyclization.[29] Following peptide cyclization and side chain deprotection as described above, five beads were randomly selected and individually treated with the CNBr solution in $CH_3CN/AcOH/H_2O$. The resulting crude peptides were analyzed by tandem MALDI-TOF MS. For each selected bead, the resulting MS spectrum showed the presence of peptides that could be unambiguously sequenced by MS/MS of the most important peak by manual analysis or de novo sequencing with the Peaks software (Figure 4b and Supporting Information, Figure S8). The procedure was performed on a freshly prepared library and no notable oxidation byproducts were observed. Nevertheless, inadvertent oxidation of the methionine residues may happen during storage, manipulation or screening and, in addition to significantly decreasing the yields, might prevent opening of the macrocycle. To avoid this problem, it is strongly suggested to treat the cyclic peptide library with a reducing solution to reduce any oxidized methionine residues prior to the tandem ring-opening/cleavage reaction.[12]

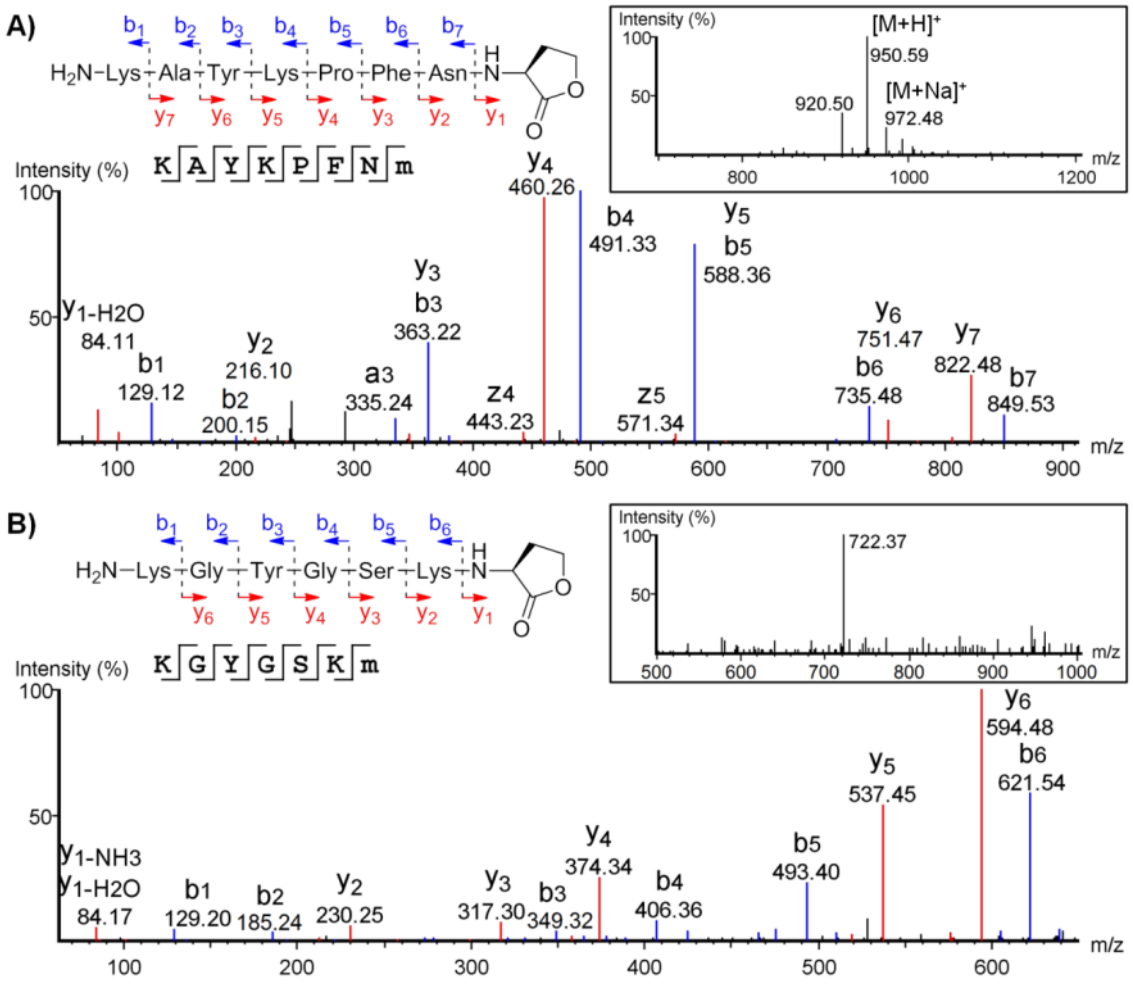


Figure 4. MALDI MS and MS/MS spectra of peptides after tandem ring-opening/cleavage on a single bead. (A) H_2N-KAYKPFNh* **6a**, MS/MS for precursor ion *m/z* 950.59. (B) H_2N-KGYGSKh* released from a bead randomly selected from the OBOC cyclic heptapeptide library, MS/MS for precursor ion *m/z* 722.37. (h* = homoserine lactone)

In conclusion, a straightforward and effective alternative approach was developed to allow sequence determination of cyclic peptides from OBOC combinatorial libraries. The results from this work demonstrate that the tandem ring-opening/cleavage strategy developed herein is compatible with commonly used amino acids and can be used on a single bead to release linear peptides that can be clearly and conclusively sequenced by MS/MS. The procedure described herein for the synthesis of the Met handle and the preparation of unencoded cyclic peptide libraries is simple and affordable for any peptide science or combinatorial chemistry laboratory.

ACKNOWLEDGMENTS

The authors are grateful to Isabelle Kelly from the Proteomic Platform of the Quebec Genomics Center at the CHU de Québec for MALDI-TOF/TOF experiments and de novo sequencing. X.L. thanks the China Scholarship Council for postgraduate scholarships. We are grateful to Dr. Richard Poulin for his help in the preparation of this manuscript.

REFERENCES

1. Lam, K. S.; Lebl, M.; Krchnak, V. "One-Bead-One-Compound" Combinatorial Library Method. *Chem. Rev.* **1997**, 97, 411−448.
2. Lam, K. S.; Liu, R.; Miyamoto, S.; Lehman, A. L.; Tuscano, J. M. Applications of One-Bead One-Compound Combinatorial Libraries and Chemical Microarrays in Signal Transduction Research. *Acc. Chem. Res.* **2003**, 36, 370−377.
3. Meldal, M. One-Bead Two-Compound Libraries for Detecting Chemical and Biochemical Conversions. *Curr. Opin. Chem. Biol.* **2004**, 8, 238−244.
4. Liu, R.; Wang, X.; Song, A.; Lam, K. S. Development and Applications of Topologically Segregated Bilayer Beads in One-Bead One-Compound Combinatorial Libraries. *QSAR Comb. Sci.* **2005**, 24, 1127−1140.
5. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature* **1991**, 354, 82−84.
6. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery. *Nature* **1991**, 354, 84−86.
7. Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. General Method for Rapid Synthesis of Multicomponent Peptide Mixtures. *Int. J. Peptide Protein Res.* 1991, 37, 487−493.
8. Aina, O. H.; Liu, R.; Sutcliffe, J. L.; Marik, J.; Pan, C. X.; Lam, K. S. From Combinatorial Chemistry to Cancer-Targeting Peptides. *Mol. Pharmaceutics* **2007**, 4, 631−651.
9. Martínez-Ceron, M. C.; Marani, M. M.; Taulés, M.; Etcheverrigaray, M.; Albericio, F.; Cascone, O.; Camperi, S. A. Affinity Chromatography Based on a Combinatorial Strategy for rErythropoietin Purification. *ACS Comb. Sci.* **2011**, 13, 251−258.

10. Kumaresan, P. R.; Wang, Y.; Saunders, M.; Maeda, Y.; Liu, R.; Wang, X.; Lam, K. S. Rapid Discovery of Death Ligands with One-Bead-Two-Compound Combinatorial Library Methods. *ACS Comb. Sci.* **2011**, 13, 259−264.

11. Alluri, P.; Liu, B.; Yu, P.; Xiao, X.; Kodadek, T. Isolation and Characterization of Coactivator-Binding Peptoids from a Combinatorial Library. *Mol. BioSyst.* **2006**, 2, 568−579.

12. Liu, T.; Qian, Z.; Xiao, Q.; Pei, D. High-Throughput Screening of One-Bead-One-Compound Libraries: Identification of Cyclic Peptidyl Inhibitors against Calcineurin/NFAT Interaction. *ACS Comb. Sci.* **2011**, 13, 536−547.

13. Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery An Underexploited Structural Class. *Nat. Rev. Drug Discovery* **2008**, 7, 608−624.

14. Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Med. Chem.* **2012**, 4, 1409−1438.

15. Adessi, C.; Soto, C. Converting a Peptide into a Drug: Strategies to Improve Stability and Bioavailability. *Curr. Med. Chem.* **2002**, 9, 963−978.

16. Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. Automated Mass Spectrometric Sequence Determination of Cyclic Peptide Library Members. *J. Comb. Chem.* **2003**, 5, 33−40.

17. Youngquist, R. S.; Fuentes, G. R.; Lacey, M. P.; Keough, T. Generation and Screening of Combinatorial Peptide Libraries Designed for Rapid Sequencing by Mass Spectrometry. *J. Am. Chem. Soc.* **1995**, 117, 3900−3906.

18. Shin, D. S.; Kim, D. H.; Chung, W. J.; Lee, Y. S. Combinatorial Solid Phase Peptide Synthesis and Bioassays. *J. Biochem. Mol. Biol.* **2005**, 38, 517−525.

19. Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. High-Throughput Sequence Determination of Cyclic Peptide Library Members by Partial Edman Degradation/Mass Spectrometry. *J. Am. Chem. Soc.* **2006**, 128, 13000−13009.

20. Liu, R.; Marik, J.; Lam, K. S. A Novel Peptide-Based Encoding System for "One-Bead One-Compound" Peptidomimetic and Small Molecule Combinatorial Libraries. *J. Am. Chem. Soc.* **2002**, 124, 7678−7680.

21. Meldal, M. The One-Bead-Two-Compound Assay for Solid Phase Screening of Combinatorial Libraries. *Biopolymers* **2002**, 66, 93−100.

22. Bédard, F.; Girard, A.; Biron, E. A Convenient Approach to Prepare Topologically Segregated Bilayer Beads for One-Bead Two- Compound Combinatorial Peptide Libraries. *Int. J. Pept. Res. Ther.* **2013**, 19, 13−23.

23. Lee, J. H.; Meyer, A. M.; Lim, H. S. A Simple Strategy for the Construction of Combinatorial Cyclic Peptoid Libraries. *Chem. Commun.* **2010**, 46, 8615−8617.

24. Simpson, L. S.; Kodadek, T. A Cleavable Scaffold Strategy for the Synthesis of One-Bead One-Compound Cyclic Peptoid Libraries That Can Be Sequenced by Tandem Mass Spectrometry. *Tet. Lett.* **2012**, 53, 2341−2344.

25. Kappel, J. S.; Barany, G. Methionine Anchoring Applied to the Solid-Phase Synthesis of Lysine-Containing "Head-to-Tail" Cyclic Peptides. Lett. Pept. Sci. **2003**, 10, 119−125.

26. Gude, M.; Ryf, J.; White, P. D. An Accurate Method for the Quantification of Fmoc-Derivatized Solid Phase Supports. *Lett. Pept. Sci.* **2002**, 9, 203−206.

27. Gaggini, F.; Porcheddu, A.; Reginato, G.; Rodriguez, M.; Taddei, M. Colorimetric Tools for Solid-Phase Organic Synthesis. *J. Comb. Chem.* **2004**, 6, 805−810.

28. Ma, B.; Lajoie, G. PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, 17, 2337−2342.

29. Thakkar, A.; Trinh, T. B.; Pei, D. Global analysis of peptide cyclization efficiency. *ACS Comb. Sci.* **2013**, 15, 120−129.

Supporting Information (SI)

A Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One-Compound Libraries

Xinxia Liang, Anick Girard and Eric Biron*

Faculty of Pharmacy, Université Laval, Pavillon Ferdinand-Vandry, Quebec (QC) G1V 0A6 Canada

Laboratory of Medicinal Chemistry, CHU de Québec Research Center (CHUL Section), 2705 Boulevard Laurier, Quebec (QC), G1V 4G2, Canada

eric.biron@pha.ulaval.ca

ELECTRONIC SUPPLEMENTARY INFORMATION

Materials and general methods 88

Synthesis of the Met handle 89

Peptide synthesis and preparation of the OBOC library 92

Figure S1-S4. Characterization of H-Met-OFm **1** and Fmoc-Lys-OAll **2** 97

Figure S5. HPLC and ESI-MS profiles of cyclic peptides. 100

Figure S6. HPLC and ESI-MS profiles of peptides after ring-opening 102

Figure S7. MALDI MS and MS/MS spectra of linear peptides obtained from tandem ring-opening/cleavage on a single TG bead. 104

Figure S8. MALDI MS and MS/MS spectra of the randomly selected beads from the cyclic peptide library after tandem ring-opening/cleavage. 106

Materials and equipment

All the chemical reagents and solvents from commercial sources were used without further purification. TentaGel S-NH_2 resin (130 μm, 0.29 mmol/g) was purchased from Rapp Polymere GmbH (Tübingen, Germany). Coupling reagents and amino acid derivatives were purchased from Matrix Innovation Inc. (Quebec, QC, Canada). Rink Amide AM resin (0.65 mmol/g) was purchased from ChemImpex (Wood Dale, IL, USA). All other reagents and solvents were purchased from Sigma-Aldrich (St. Louis, MO, USA). Reactions on solid support were performed in filter columns (2 and 10 mL) from Roland Vetter Laborbedarf OHG (Ammerbuch, Germany). Flash chromatography was performed on silica gel F60 (230-400 mesh) from Silicycle (Quebec, QC, Canada). 1H spectra were obtained in $CDCl_3$ or DMSO-d_6 as solvent and internal reference on a Bruker AVANCE 400 spectrometer (Billerica, MA, USA). RP-HPLC analyses were conducted on a Shimadzu Prominence instrument (Columbia, MD, USA) using a Phenomenex Gemini-NX column (4.6 mm × 250 mm, 5 μm C_{18}, 110Å, 1 mL/min) with a 20 min linear gradient from water (0.1% TFA) and CH3CN (0.1% TFA) (CH3CN 10-100%) and detection at 220 nm and 254 nm. Mass spectra were performed on a Shimadzu Prominence LCMS-2020 equipped with an ESI and APCI ion source. Absorbance analyses were performed on a Genesys 10uV UV-visible spectrophotometer from Thermo Scientific (Madison, WI, USA). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF) was performed on a AB SCIEX 4800 Plus MALDI-TOF/TOF instrument using alpha-cyano-4-hydroxycinnamic acid as matrix. The spectra were acquired using the 4000 Series Explorer Software (AbSciex, v 3.2.3). The PEAKS Studio software (Bioinformatics Solutions, v.5.3) was used for spectra analysis and DENOVO sequencing.

Fmoc quantification and loading determination

The procedure described by Gude and coworkers was used.[1] Briefly, approximately 25 mg of resin were weighed into a 25-mL volumetric flask and 5 mL of 2% 8-diazabicyclo[5.4.0]undec-7-ene (DBU) in DMF (v/v) were added. The suspension was stirred for 30 min and diluted with CH_3CN to 25 mL. Two mL of this solution were transferred to a 10-mL volumetric flask and diluted with CH_3CN to 10 mL. A reference solution was prepared in the same manner but without resin. To quantify the Fmoc group, 3

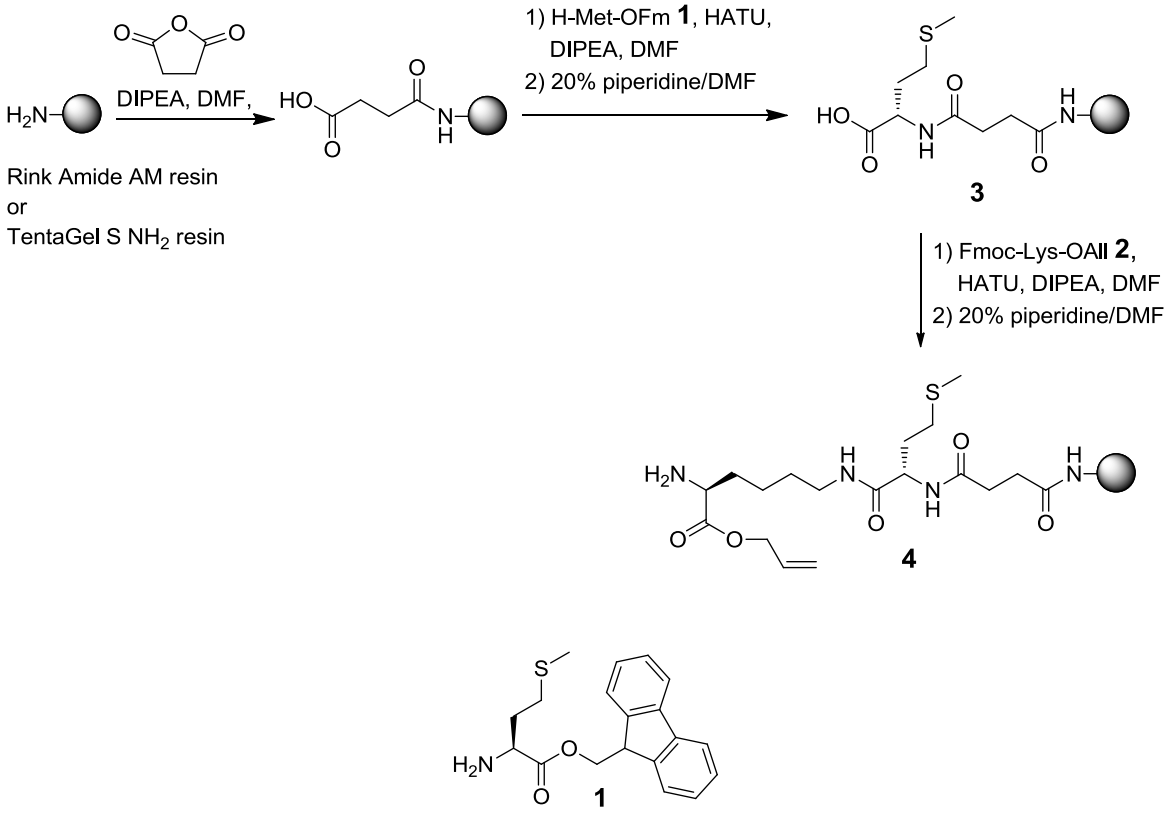
mL of solution were transferred in 1-cm cuvettes. The A_{304} of the sample solution was measured with a UV spectrophotometer and compared with the reference solution. The loading was calculated by inserting the A_{304} value into the following equation where A is the absorbance and mg is the mass of resin weighed:

$$Loading(mmol/g) = \left(A \times \frac{25}{mg} \right) \times 1.3117$$

Chloranil test

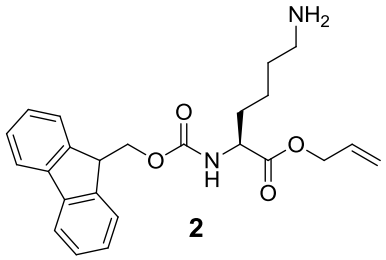
The presence of free amines on the resin was detected with the chloranil test.[2] Briefly, 2-3 drops of a solution of 2% acetaldehyde (v/v) in DMF and 2-3 drops of a solution of 2% 2,3,5,6-tetrachloro-1,4-benzoquinone (w/w) in DMF were added to few beads of the tested resin. The mixture was shaken at room temperature for 5 min. The presence of free amines is indicated by a green- or blue-colored solution. Negative samples (absence of free amines) register as yellow, amber, or brown.

Synthesis of the methionine handle.



***Synthesis of H-Met-OFm* (1).** The amino acid was prepared as previously described.[3] Briefly, Boc-Met-OH (1 g, 4.0 mmol) was dissolved in dry CH_2Cl_2 and N,N′-dicyclohexylcarbodiimide (990 mg, 4.8 mmol), 9-fluorenylmethanol (942 mg, 4.8 mmol) and 4-dimethylaminopyridine (DMAP) (97 mg, 0.8 mmol) were added in turn at 0°C. Stirring was continued for 20 min at 0°C and overnight at room temperature. After filtration, the reaction mixture was washed with a 0.1 N HCl solution, a saturated $NaHCO_3$ solution and water, and dried over $MgSO_4$. The solvent was almost entirely removed under reduced and the product precipitated with hexane. The white powder was washed 3 times with hexane and dried with a vacuum pump to afford Boc-Met-OFm as a white powder (1.26 g, 74%). MS (ESI) *m/z*: calcd for $C_{24}H_{29}NO_4SNa$ ($M+Na^+$) 450.17; observed 450.25; RP-HPLC t_R=23.9 min (10-100%).

The obtained Boc-Met-OFm was used without any further purification and dissolved in 30 mL of 4 N HCl in dioxane. The solution was stirred 1 h at room temperature and the solvent was removed under reduced pressure. The yellow solid was washed with diethyl ether (5 x 20 mL) by removing the supernatant by suction and dried under vacuum to afford H-Met-OFm as a white powder (1.0 g, 95%, 70% for two steps). ^{1}H NMR (400 MHz, CD_3OD) δ 1.66 (q, J = 7.1 Hz, 2H), 1.88 (s, 3H), 2.16-2.21 (m, 2H), 2.10-2.18 (m, 1H), 3.96 (t, J = 6.4 Hz, 1H), 4.30 (t, J = 4.5 Hz, 1H), 4.72 (dd, J_1 = 4.4 Hz, J_2 = 11.0 Hz, 1H), 5.07 (dd, J_1 = 4.6 Hz, J_2 = 11.0 Hz, 1H), 7.32-7.36 (m, 2H), 7.40 (t, J = 7.5 Hz, 2H), 7.63 (q, J = 7.5 Hz, 2H), 7.82 (d, J = 7.5 Hz, 2H); ^{13}C NMR (400 MHz, CD_3OD) δ 14.7, 29.7, 30.4, 48.2, 52.3, 67.9, 121.1, 125.6, 125.8, 128.3, 128.4, 129.0, 129.1, 142.7, 142.9, 144.4, 144.9, 170.5; MS (ESI) *m/z*: calcd for $C_{19}H_{22}NO_2S$ ($M+H^+$) 328.14; observed 328.20; RP-HPLC t_R=17.7 min (10-100%).



2

***Synthesis of Fmoc-Lys-OAll* (2).** Fmoc-Lys(Boc)-OH (1 g, 2.1 mmol), allyl alcohol (0.7 mL, 10.5 mmol), DMAP (14 mg, 0.1 mmol) were dissolved in 100 mL of CH_2Cl_2 and

stirred for 20 min under argon at 0°C. Then, 0.44 g of EDAC (2.3 mmol) was added and the reaction was stirred 5 min at 0°C and for 2 h under argon at room temperature. The reaction mixture was washed with a 0.1N HCl solution, a saturated $NaHCO_3$ solution and water, and dried over $MgSO_4$. The solvent was removed under reduced pressure and the product intensively dried with a vacuum pump to afford Fmoc-Lys(Boc)-OAll as a brown oil (0.98 g, 92%). MS (ESI) *m/z*: calcd for $C_{29}H_{37}N_2O_6$ ($M+H^+$) 509.26; observed 509.40; RP-HPLC t_R=23.8 min (10-100%).

The obtained Fmoc-Lys(Boc)-OAll was used without any further purification and dissolved in a solution of 50% TFA in CH_2Cl_2. The mixture was stirred for 30 min at room temperature. The solvent was removed under reduced pressure and the crude product purified by flash chromatography on silica gel (70% EtOAc in hexanes) to afford a colourless oil. The product was dissolved in a minimum of diethyl ether and precipitated by adding hexane. The supernatant was removed by suction and the product dried under vacuum to afford a white solid (0.68 g, 86%, 79% for two steps). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 1.35-1.43 (m, 2H), 1.48-1.59 (m, 2H), 1.65-1.73 (m, 1H), 1.84-1.92 (m, 1H), 3.05-3.19 (m, 2H), 4.23 (t, *J* = 7 Hz, 1H), 4.36-4.45 (m, 3H), 4.55-4.61 (m, 1H), 4.65 (d, *J* = 6 Hz, 2H), 5.20 (d, *J* = 10.4 Hz, 1H), 5.27 (d, *J* = 18.0 Hz, 1H), 5.87-5.96 (m, 1H), 7.32 (t, *J* = 7.4 Hz, 2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.61 (d, *J* = 7.4 Hz, 2H), 7.77 (d, *J* = 7.5 Hz, 2H); ^{13}C NMR (400 MHz, $CDCl_3$) δ 21.8, 26.3, 31.5, 40.3, 46.8, 53.7, 66.7, 67.8, 119.5, 120.1, 124.8, 127.1, 127.9, 130.8, 141.3, 143.3, 157.2, 172.3; MS (ESI) *m/z*: calcd for $C_{24}H_{29}N_2O_4$ ($M+H^+$) 409.21; observed 409.30; RP-HPLC t_R=18.4 min (10-100%).

***Reverse Met handle resin* (3).** TentaGel S NH_2 130 µm (1 g, 29 mmol) or Rink Amide resin (1 g, 65 mmol) was first swelled in 20 mL of DMF followed by addition of succinic anhydride (3 equiv) and NMM (6 equiv). The mixture was stirred overnight and the absence of free N-amine was confirmed by the chloranil test. Following resin filtration and washing with DMF (5 × 30 s), H-Met-OAll (5 equiv) was coupled to the resin in the presence of HATU (5 equiv) and collidine (10 equiv) in DMF (20 mL) for 6 h. The coupling step was repeated another time and the resin was washed with DMF (5 × 30 s) and CH_2Cl_2 (5 × 30 s). Next, the allyl protecting group was removed with a solution of $Pd(PPh_3)_4$ (0.4 equiv) and $PhSiH_3$ (12 equiv) in CH_2Cl_2 for 2 × 30 min. The resin was

washed with CH_2Cl_2 (5 × 30 s), 0.5% DIPEA (v/v) in DMF (2 × 2 min), 5% diethyl dithiocarbamate (v/v) in DMF (2 × 5 min), DMF (5 × 30 s), CH_2Cl_2 (5 × 30 s) and dried under vacuum overnight.

***Side chain anchored Fmoc-Lys-OAll on reverse Met handle resin* (4).** The reverse Met handle resin was swelled in DMF and Fmoc-Lys-OAll (5 equiv) coupled to the resin in the presence of HATU (5 equiv) and collidine (10 equiv) in DMF (20 mL) for 6 h. The coupling step was repeated another time and the resin was washed with DMF (5 × 30 s). Next, the resin loading was determined by the UV method. Loadings of 0.23 mmol/g and 0.57 mmol/g were observed for TG and Rink Amide resin, respectively. Finally, the Fmoc protecting group was removed by treating the resin twice with 20% piperidine in DMF (v/v) for 10 min. After washing with DMF (5 × 30 s) and CH_2Cl_2 (5 × 30 s), the resin was dried under vacuum overnight.

Peptide synthesis

Peptides were synthesized by standard Fmoc solid phase synthesis.[3] Briefly, amino acid couplings were performed with a solution of Fmoc-Xaa-OH (3 equiv), HCTU (3 equiv) and NMM (12 equiv) in DMF for 15 min. The coupling step was repeated once and the resin washed with DMF (5 × 30 s). The Fmoc protecting group was removed by treating the resin twice with a solution of 20% piperidine in DMF (v/v) for 10 min followed by washing with DMF (5 × 30 s).

On resin peptide cyclization. The resin was swelled with CH_2Cl_2 and the allyl protecting group was removed with a solution of $Pd(PPh_3)_4$ (0.4 equiv) and $PhSiH_3$ (12 equiv) in CH_2Cl_2 for 2 × 30 min. The resin was washed with CH_2Cl_2 (5 × 30 s), 0.5% DIPEA (v/v) in DMF (2 × 2 min), 5% diethyl dithiocarbamate (v/v) in DMF (2 × 5 min) and DMF (5 × 30 s). The Fmoc group was then removed with 20% piperidine in DMF (2 × 10 min) and the resin washed with DMF (5 × 30 s). Peptide cyclization was performed on solid support in the presence of PyBOP (3 equiv), HOBt (3 equiv) and DIPEA (6 equiv) in DMF for 3 h.[4] Cyclization was monitored by the chloranil test. After reaction completion, the resin was washed with DMF (5 × 30 s) and CH_2Cl_2 (5 × 30 s).

Cleavage from Rink Amide AM resin. The resin (10 mg) was swelled with CH_2Cl_2 followed by treatment with a cleavage solution composed of TFA/water/TIS (95:2.5:2.5) for 3 h. After filtration the resin was washed with CH_2Cl_2 (2 × 30 s) and the filtrates were evaporated to dryness. The resulting product was precipitated with cold diethyl ether and dried under vacuum.

Side-chain deprotection on TentaGel resin. The resin was swelled with CH_2Cl_2 and side chain deprotection was performed with a mixture of TFA, water and triisopropylsilane (TIS) (95:2.5:2.5) for 3 h followed by washing with CH_2Cl_2 (5 × 30 s).

Ring-opening reaction and cleavage from TentaGel resin. The resin (20 mg) was washed with CH_3CN (3 × 1 min) and treated overnight with a solution (500 µL) of CNBr (40 mg/mL) in $CH_3CN/AcOH/H_2O$ (5:4:1) in the dark. After filtration, the solvent was removed under reduced pressure to yield the free peptide.

Ring-opening/cleavage from a single bead and sequencing. The bead was transferred into a microcentrifuge tube and treated overnight with 20 µL of CNBr (40 mg/mL) in $CH_3CN/AcOH/H_2O$ (5:4:1) in the dark. The solvents were evaporated under vacuum to dryness and the peptides released from the bead were dissolved in 10 µL of 0.1% TFA in water. 1 µL of the peptide solution was mixed with 1 µL of 4-hydroxy-α-cyanocinnamic acid (5 mg/mL) in CH_3CN/0.1% TFA (1:1) and 1 µL of the mixture was spotted onto a MALDI sample plate for MS/MS analysis.

Cyclo[Ala-Tyr-Lys-Pro-Phe-Asn-Met-Lys(Met-Succinamide)] (**5a)**: 89% purity (determined by HPLC); RP-HPLC t_R= 15.06 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{56}H_{84}N_{13}O_{13}S_2$ ($M+H^+$) 1210.57; observed 1210.55.

Cyclo[Ala-Arg-Tyr-Phe-Val-Met-Lys(Met-Succinamide)] (**5b)**: 91% purity (determined by HPLC); RP-HPLC t_R= 15.79 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{52}H_{80}N_{13}O_{11}S_2$ ($M+H^+$) 1126.55; observed 1126.60.

Cyclo[Ala-Gly-Arg-Trp-Met-Lys(Met-Succinamide)] (**5c)**: 94% purity (determined by HPLC); RP-HPLC t_R= 14.41 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{42}H_{66}N_{13}O_9S_2$ ($M+H^+$) 960.45; observed 960.45.

Cyclo[Ala-Lys-Phe-Met-Lys(Met-Succinamide)] (**5d)**: 79% purity (determined by HPLC); RP-HPLC t_R= 14.10 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{38}H_{62}N_9O_8S_2$ ($M+H^+$) 836.42; observed 836.50.

H-Lys-Ala-Tyr-Lys-Pro-Phe-Asn-homoserine lactone (**6a)**: 95% purity (determined by HPLC); RP-HPLC t_R= 12.52 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{46}H_{68}N_{11}O_{11}$ ($M+H^+$) 950.51; observed 950.50.

H-Lys-Ala-Arg-Tyr-Phe-Val-homoserine lactone (**6b)**: 89% purity (determined by HPLC); RP-HPLC t_R= 13.26 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{42}H_{64}N_{11}O_9$ ($M+H^+$) 866.49; observed 866.50.

H-Lys-Ala-Gly-Arg-Trp-homoserine lactone (**6c)**: 93% purity (determined by HPLC); RP-HPLC t_R= 12.05 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{32}H_{50}N_{11}O_7$ ($M+H^+$) 700.39; observed 700.45.

H-Lys-Ala-Lys-Phe-homoserine lactone (**6d)**: 92% purity (determined by HPLC); RP-HPLC t_R= 10.19 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{28}H_{46}N_7O_6$ ($M+H^+$) 576.35; observed 576.45.

Synthesis of the random cyclic heptapeptide library

The OBOC combinatorial cyclic peptide library was prepared on 100 mg of 130 μm TG beads bearing Fmoc-Lys-OAll anchored by its side chain to the reverse Met linker. After removal of the Fmoc group with piperidine and coupling of Fmoc-Met-OH with HCTU as described above, the small library was synthesized by the split-and-pool method using Fmoc/HCTU solid phase peptide chemistry (see above). The next four position within the peptide library has a random combination of sixteen L-amino-acids (A, D, E, F, G, H, K, L, N, P, R, S, T, V, W and Y). The last position before cyclization has a random combination of three L-amino acids (G, A and L). Following amino terminus deprotection and allyl cleavage, peptide cyclization was carried out with PyBOP as described above. The side chain protecting groups were removed by treatment with a mixture of TFA and the beads washed extensively with CH_2Cl_2 (5 × 30 s), DMF (5 × 30 s) and MeOH (5 × 30 s).

Reference

1) Gude, M.; Ryf, J.; White, P.D. An accurate method for the quantification of Fmoc-derivatized solid phase supports. *Lett. Pept. Sci.* **2002**, *9*, 203-206.

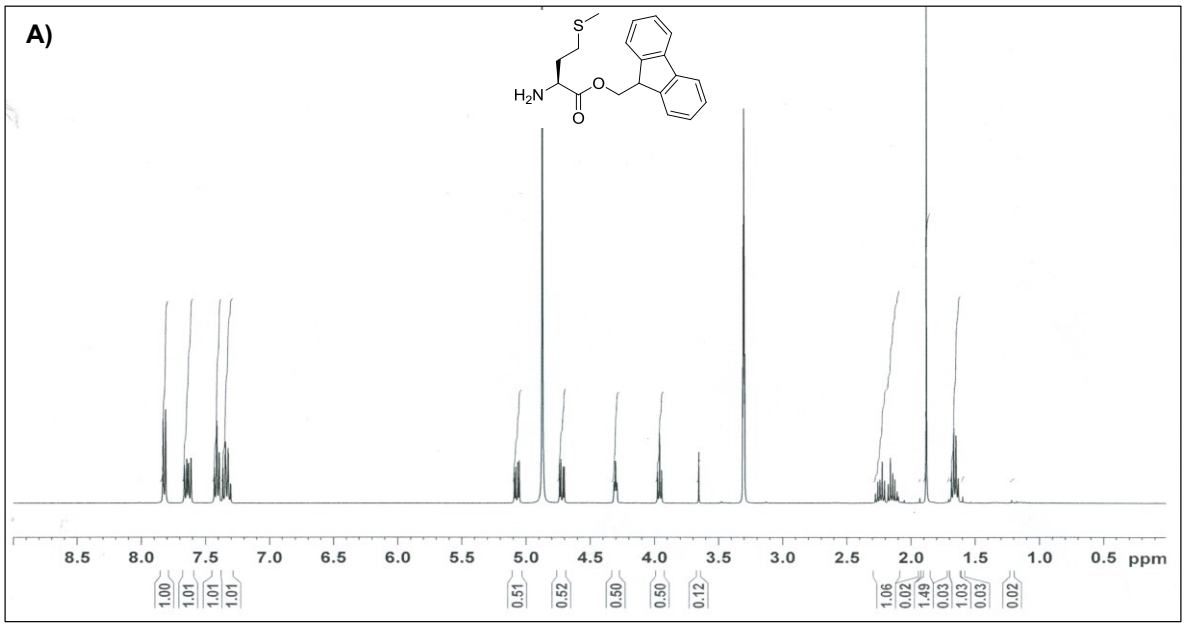
2) Gaggini, F.; Porcheddu, A.; Reginato, G.; Rodriguez, M.; Taddei, M. Colorimetric tools for solid-phase organic synthesis. *J. Comb. Chem.* **2004**, *6*, 805-810.

3) Kappel, J. S.; Barany, G. Methionine anchoring applied to the solid-phase synthesis of lysine-containing head-to-tail' cyclic peptides. *Lett. Pept. Sci.* **2003**, *10*, 119-125

4) Fields, G.B.; Noble, R.L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* **1990**, *35*, 161-214.

5) Thakkar, A.; Trinh T.B.; Pei, D. Global analysis of peptide cyclization efficiency. *ACS Comb. Sci.* **2013**, *15*, 120-129.

Figure S1. ^{1}H NMR spectra of A) H-Met-OFm **1** in CD_3OD and B) Fmoc-Lys-OAll **2** in $CDCl_3$



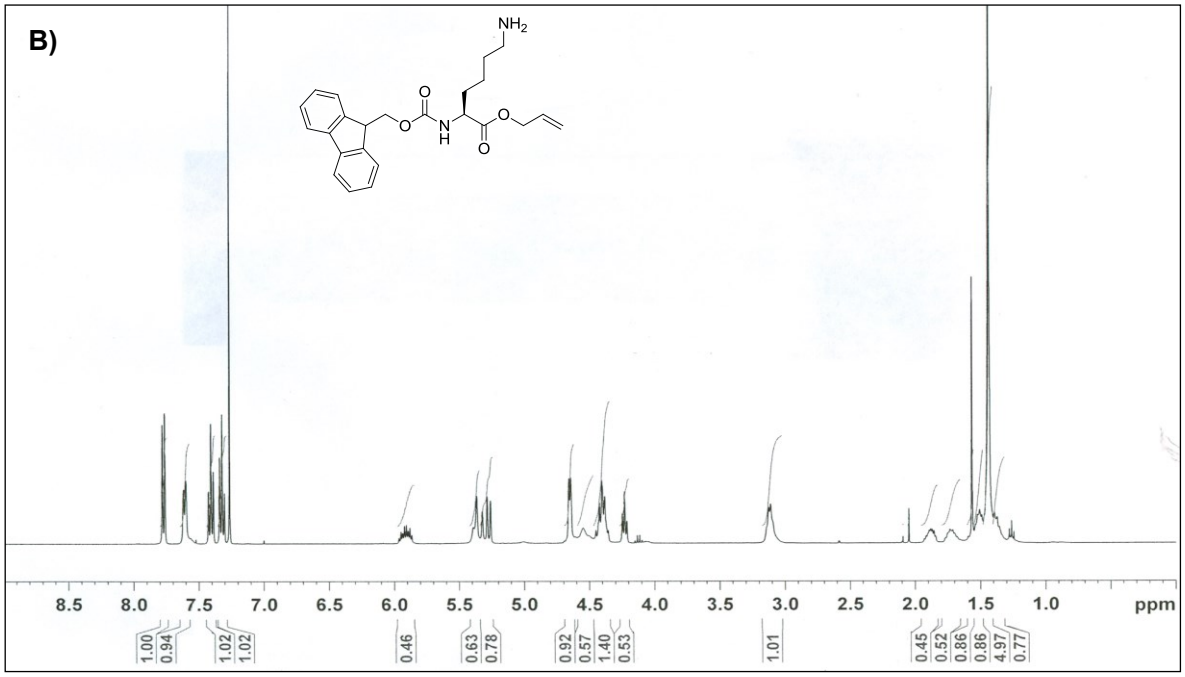
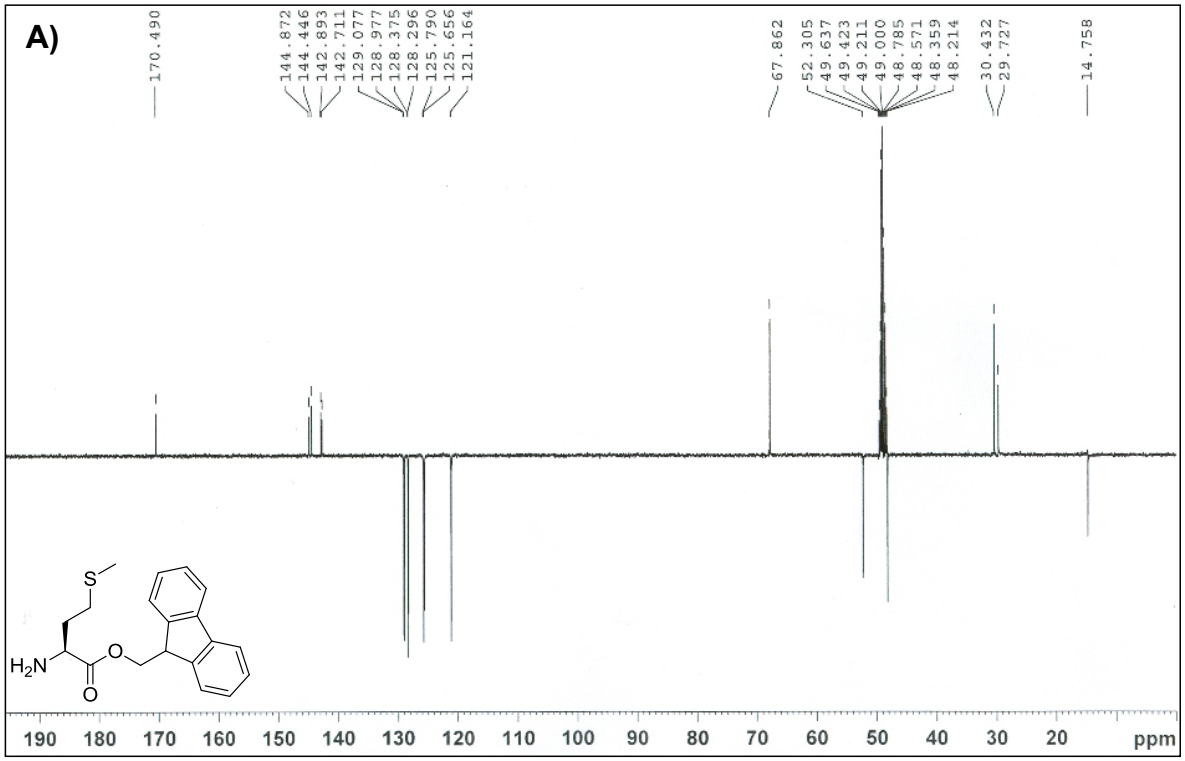


Figure S2. ^{13}C APT NMR spectra A) H-Met-OFm **1** in CD_3OD; B) Fmoc-Lys-OAll **2** in $CDCl_3$



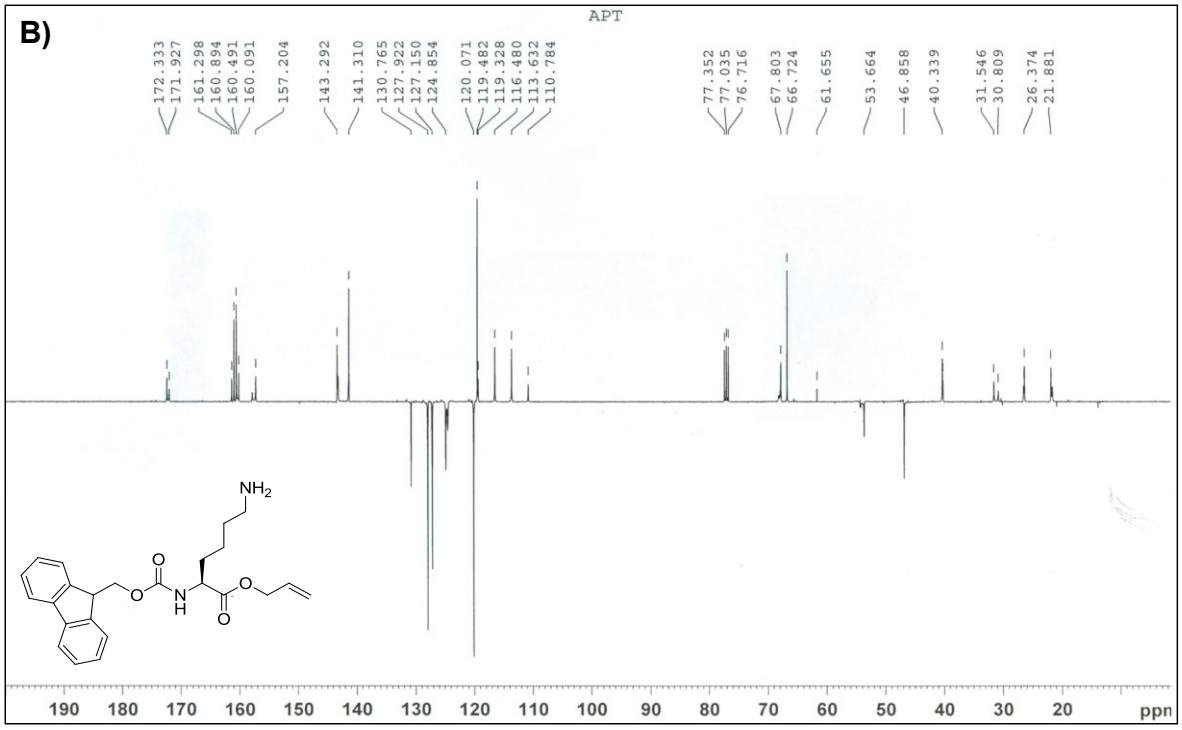


Figure S3. HPLC profiles (λ = 220 nm) of A) H-Met-OFm **1** and B) Fmoc-Lys-OAll **2**

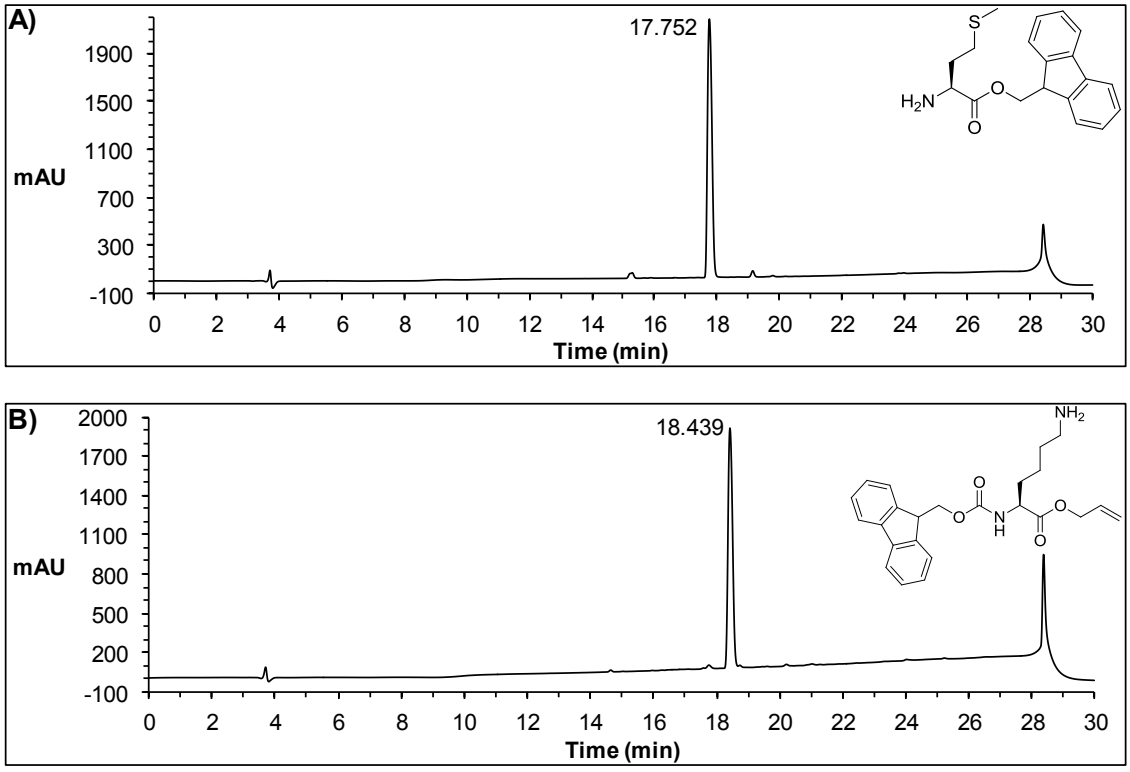


Figure S4. ESI-MS spectra of A) H-Met-OFm **1** and B) Fmoc-Lys-OAll **2**

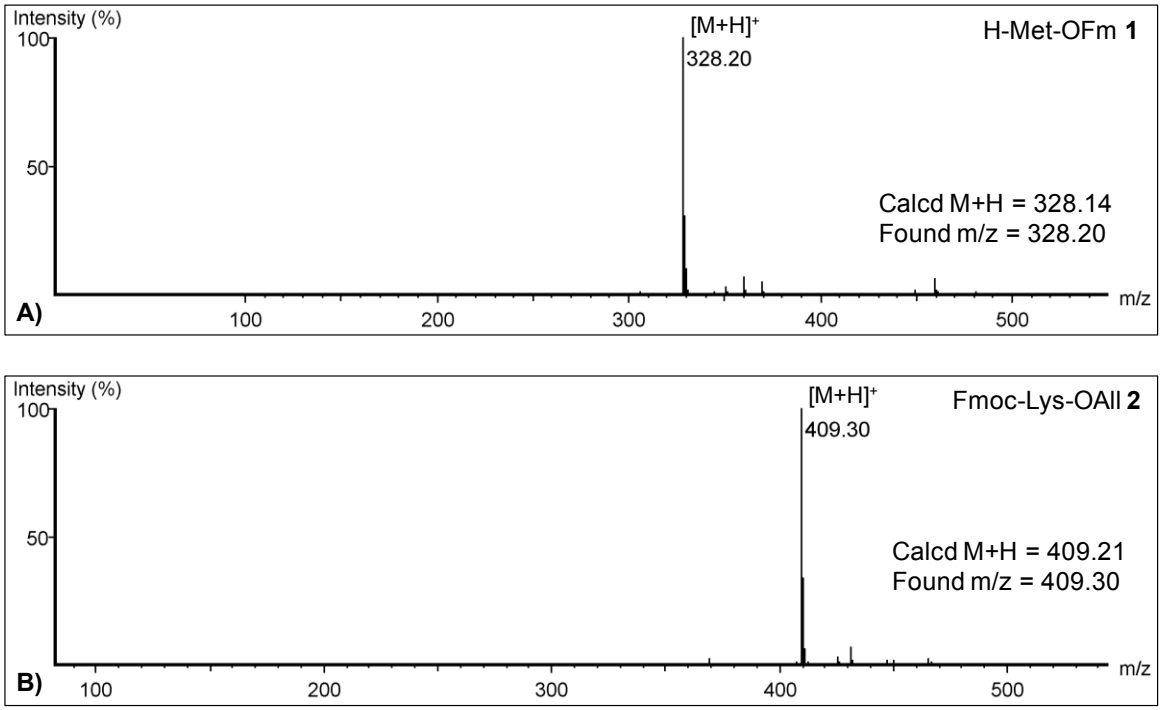


Figure S5. HPLC (λ = 220 nm) and ESI-MS profiles of cyclic peptides.

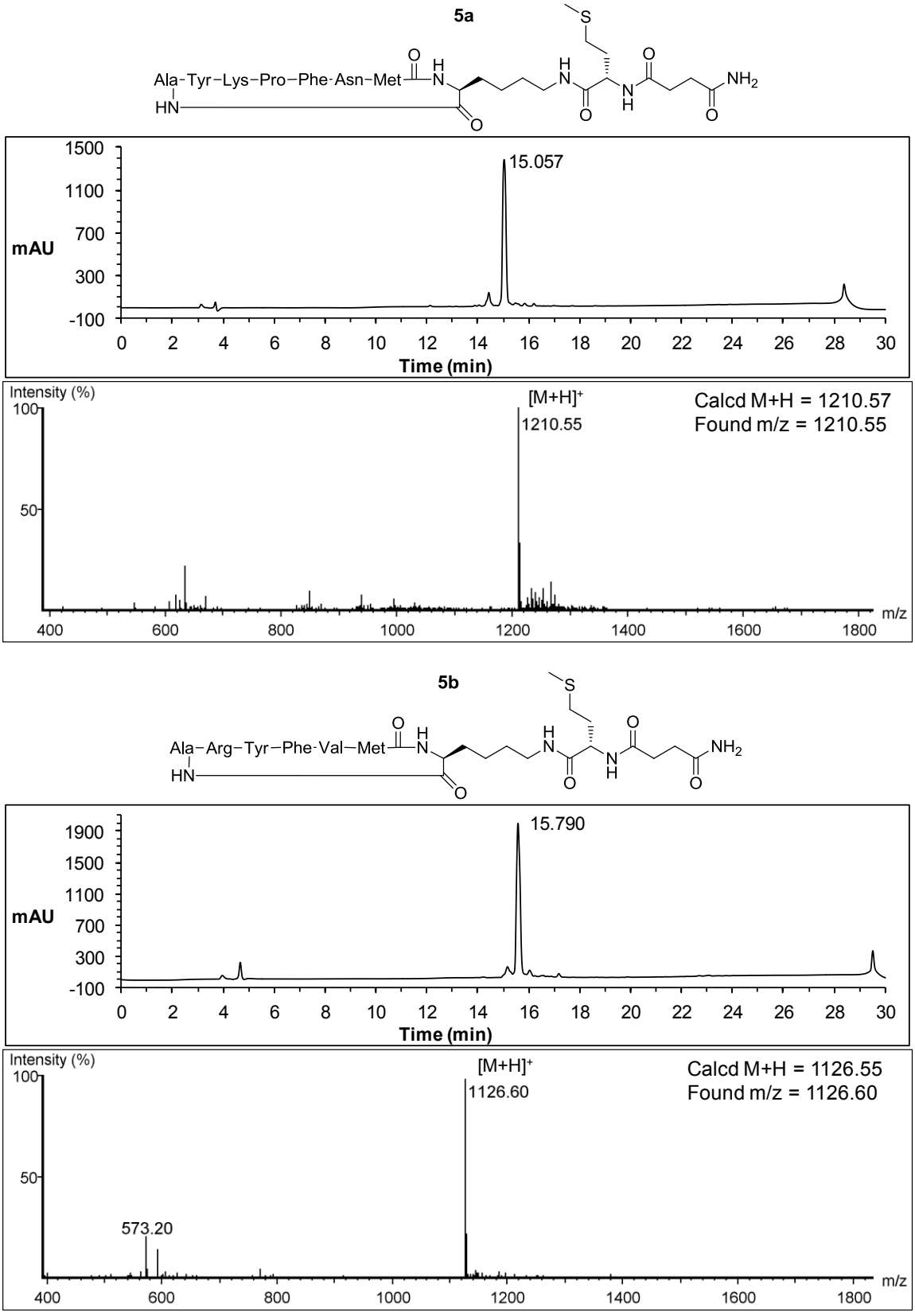
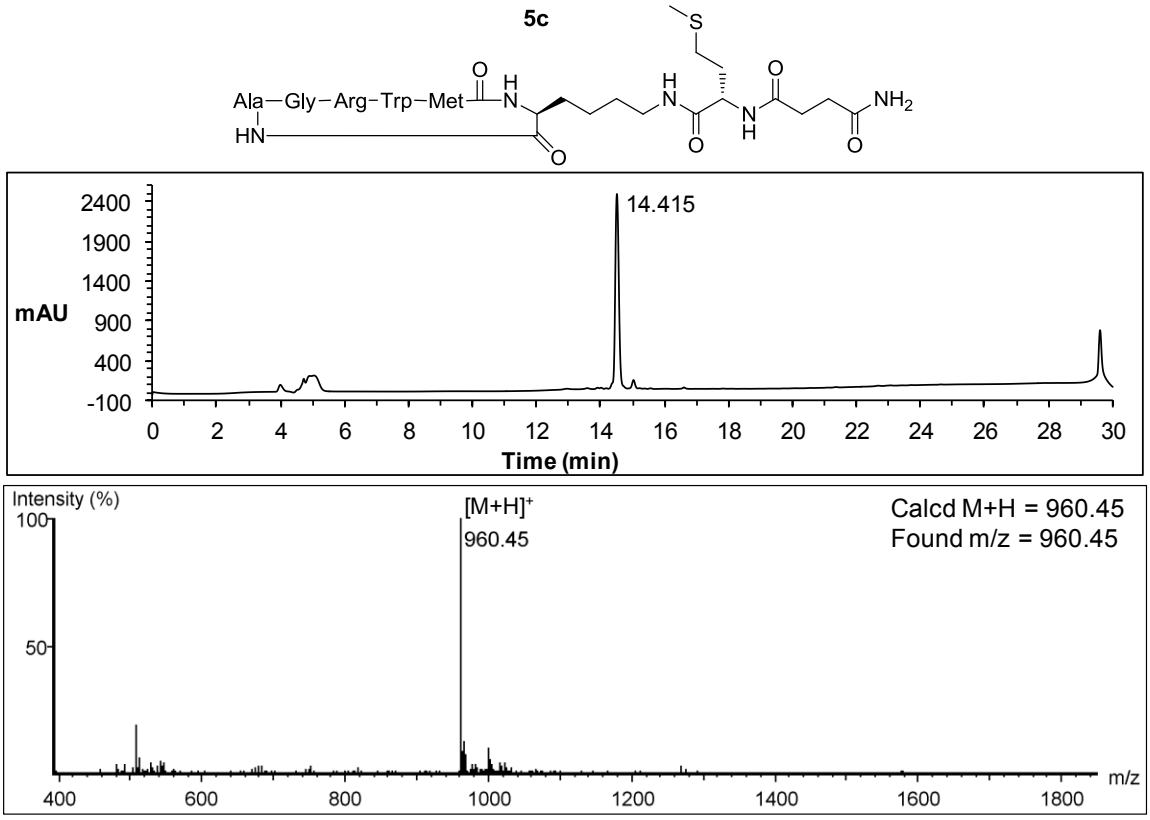


Figure S5. (Continued)



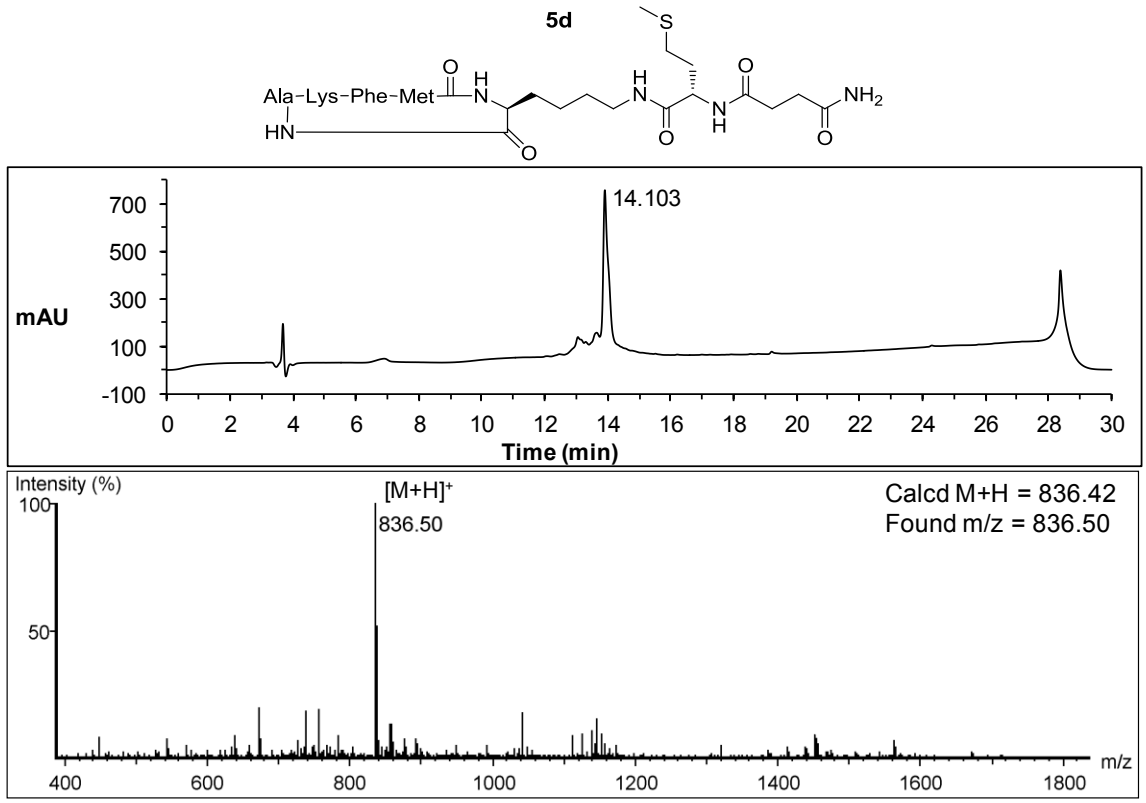
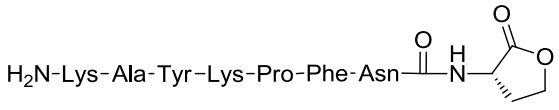
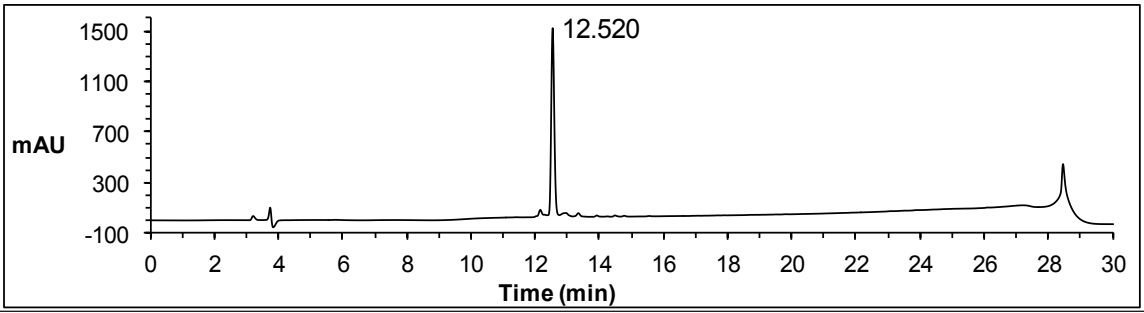
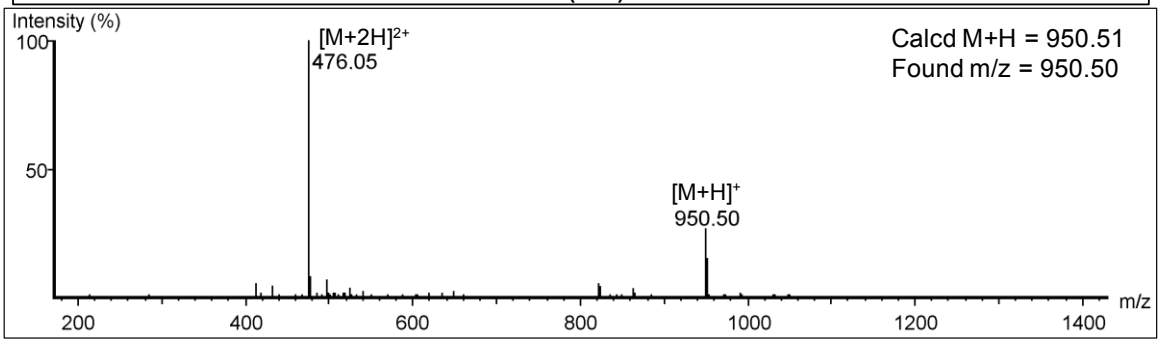


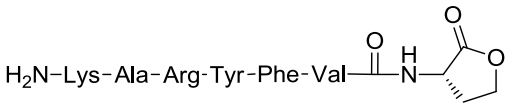
Figure S6. HPLC (λ = 220 nm) and ESI-MS profiles of linear peptides after ring-opening.



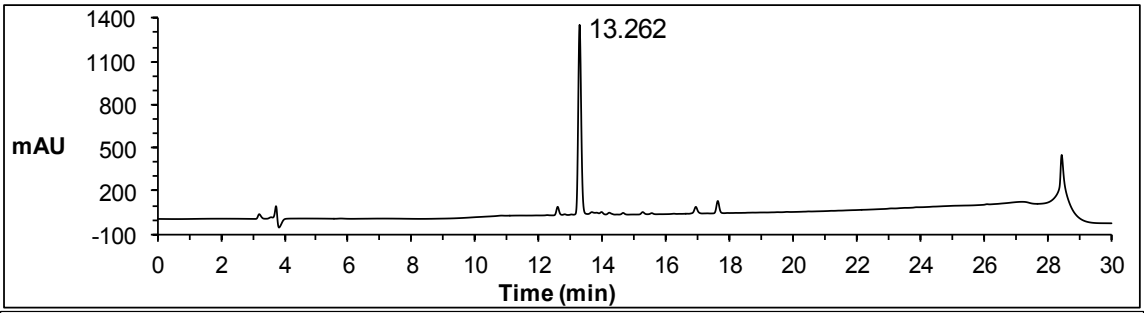
6a







6b



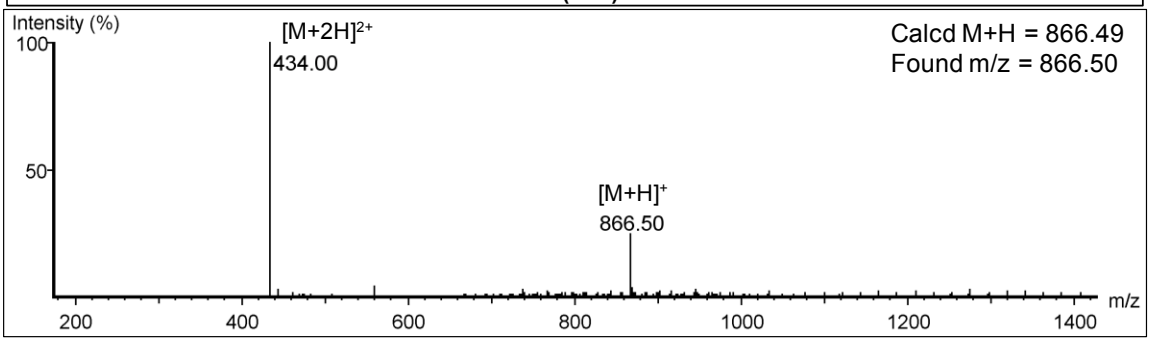
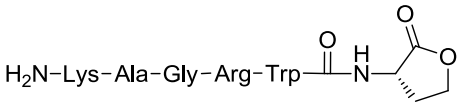
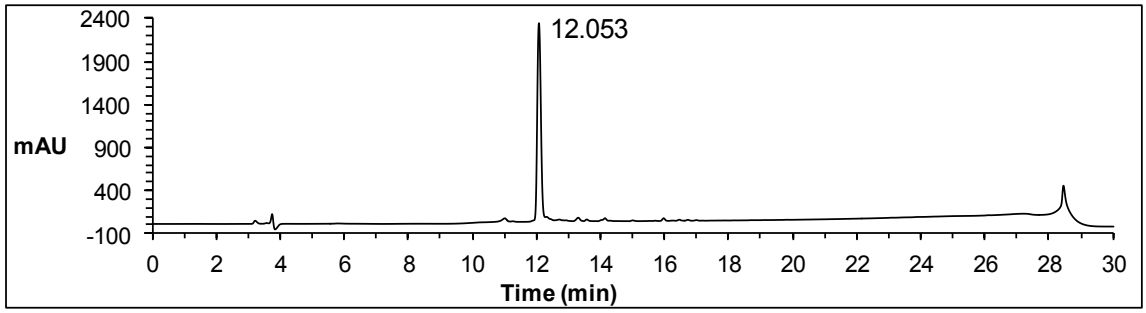
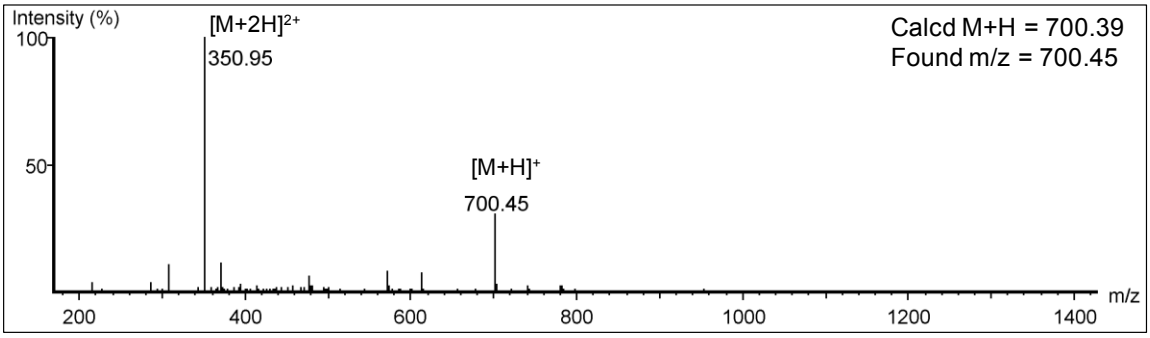


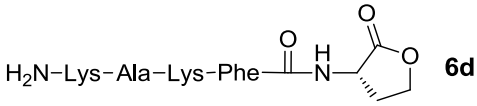
Figure S6. (Continued)



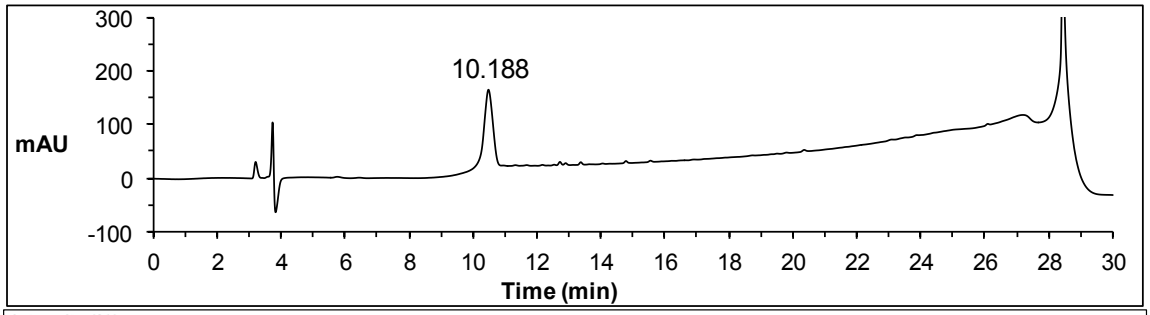
6c







6d



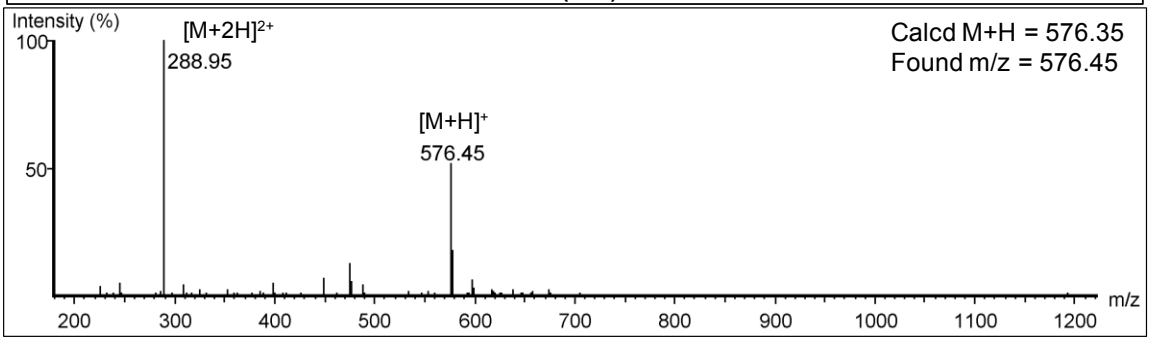
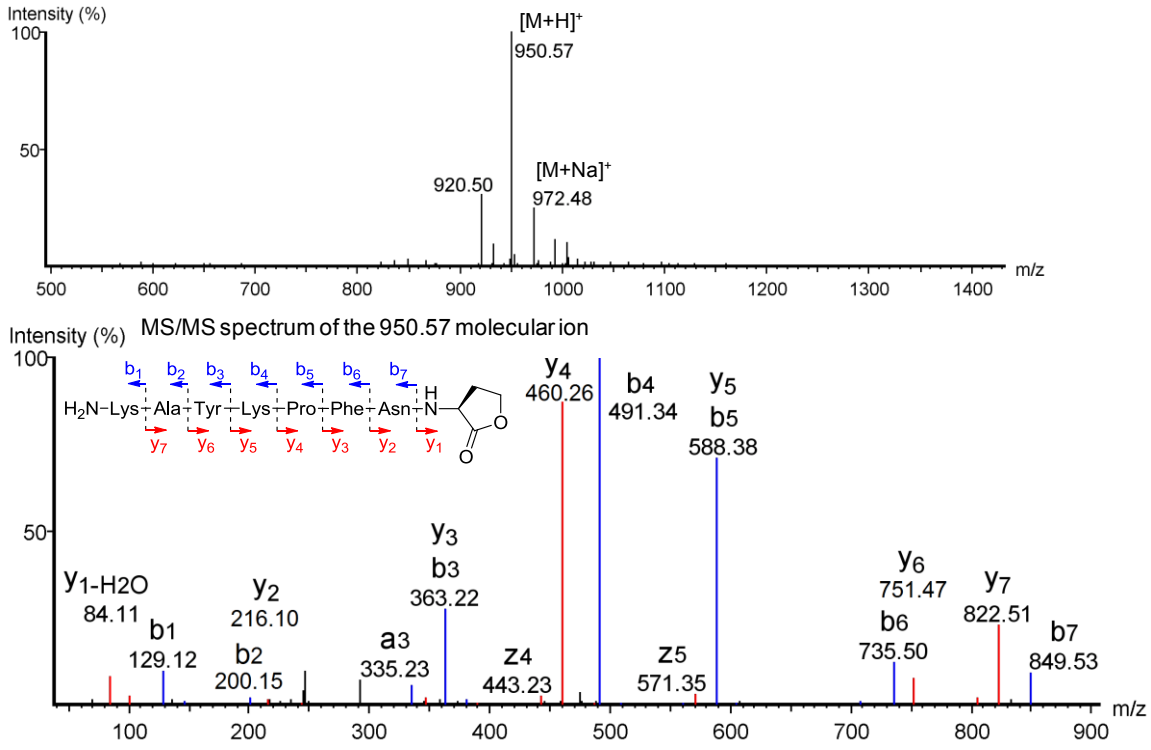


Figure S7. MALDI-TOF MS and MS/MS spectra of linear peptides obtained from tandem ring-opening/cleavage on a single TG bead

H-Lys-Ala-Tyr-Lys-Pro-Phe-Asn-homoserine lactone **6a**



H-Lys-Ala-Arg-Tyr-Phe-Val-homoserine lactone **6b**

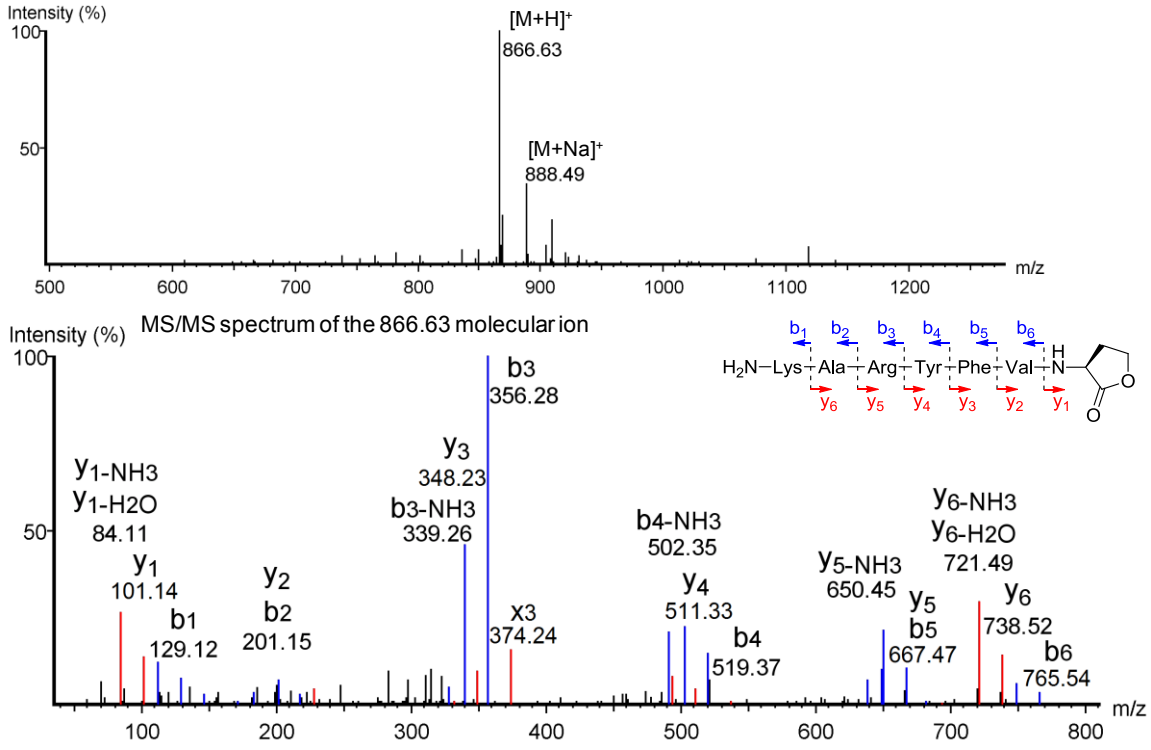
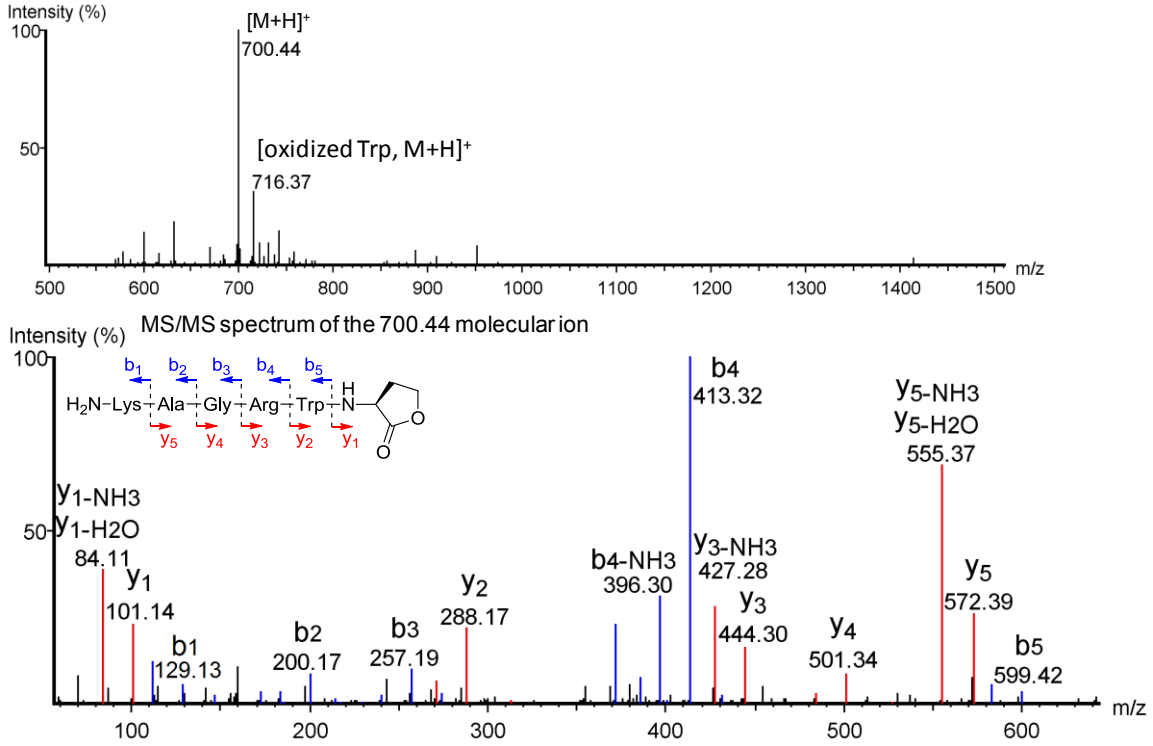


Figure S7. (Continued)

H-Lys-Ala-Gly-Arg-Trp-homoserine lactone **6c**



H-Lys-Ala-Lys-Phe-homoserine lactone **6d**

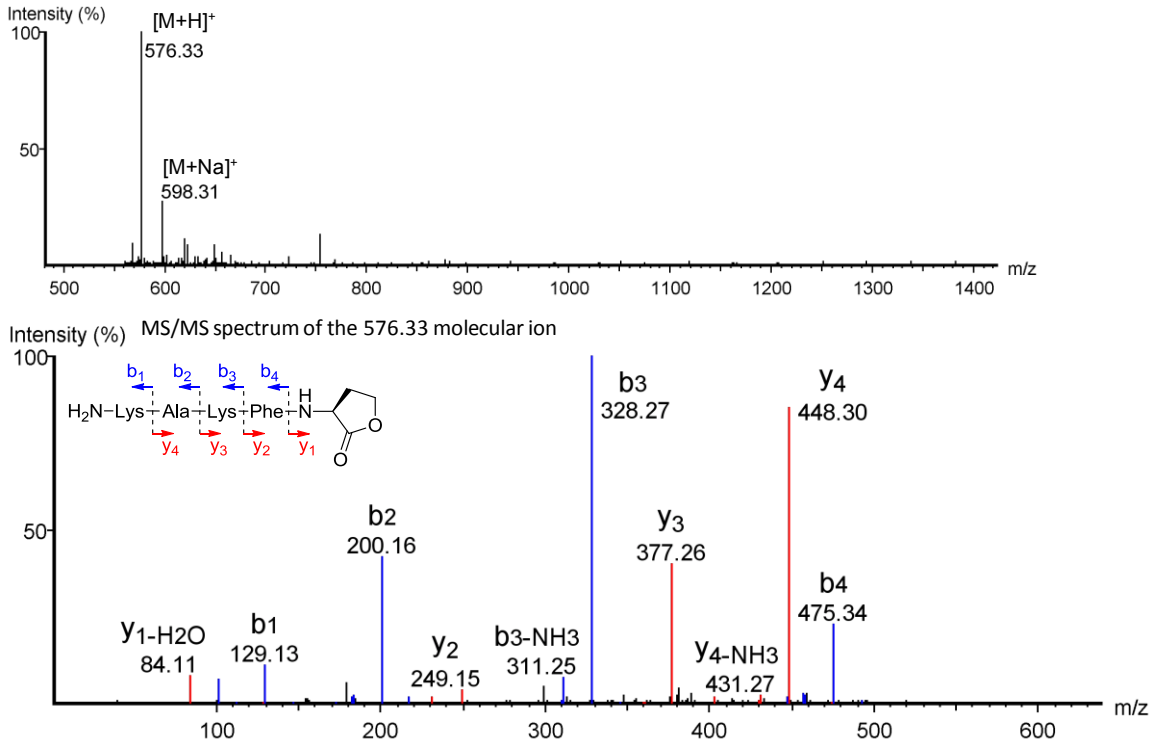
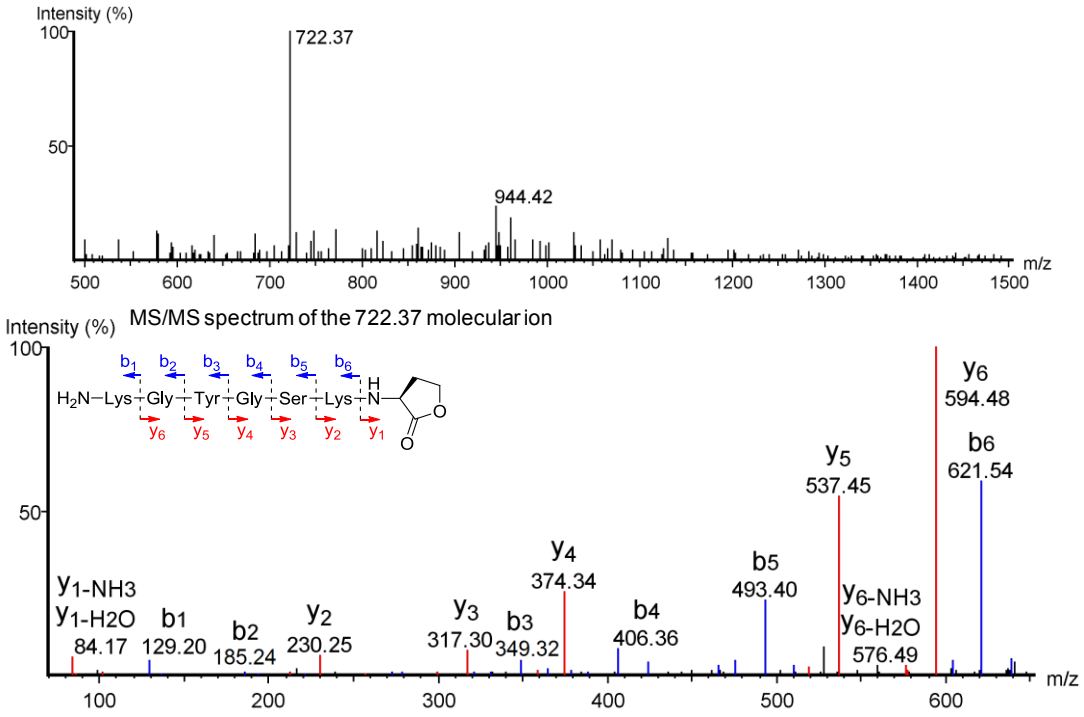


Figure S8. MALDI-TOF MS and MS/MS spectra of the randomly selected beads from the cyclic peptide library after tandem ring-opening/cleavage.

Bead #1



Bead #2

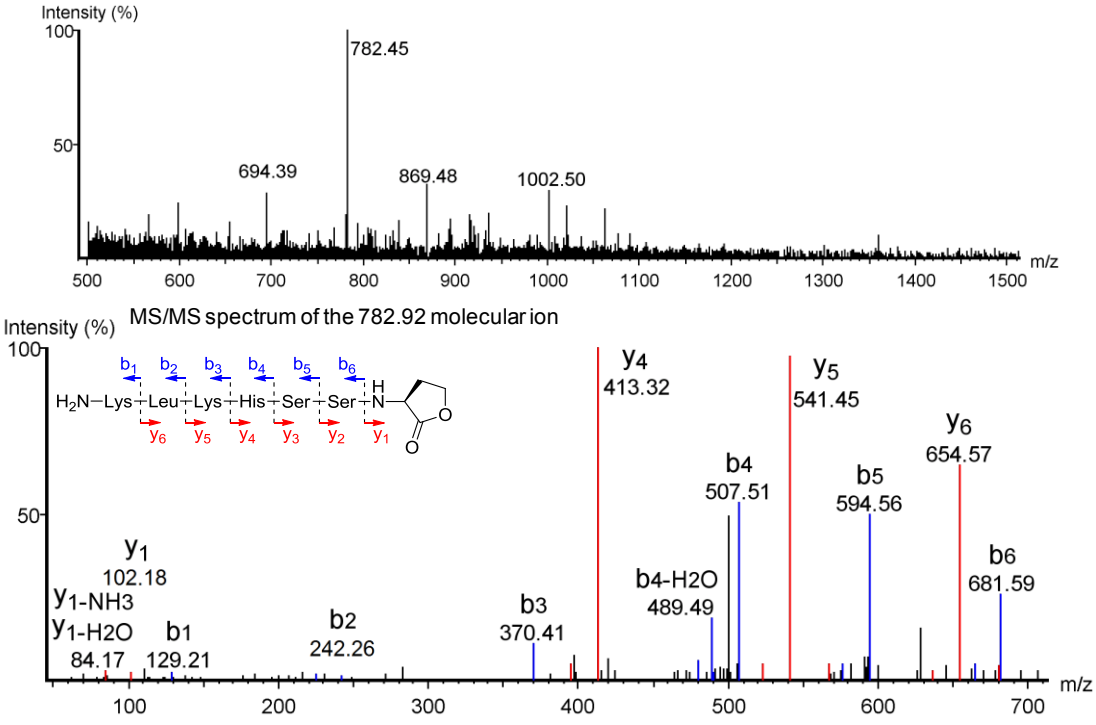
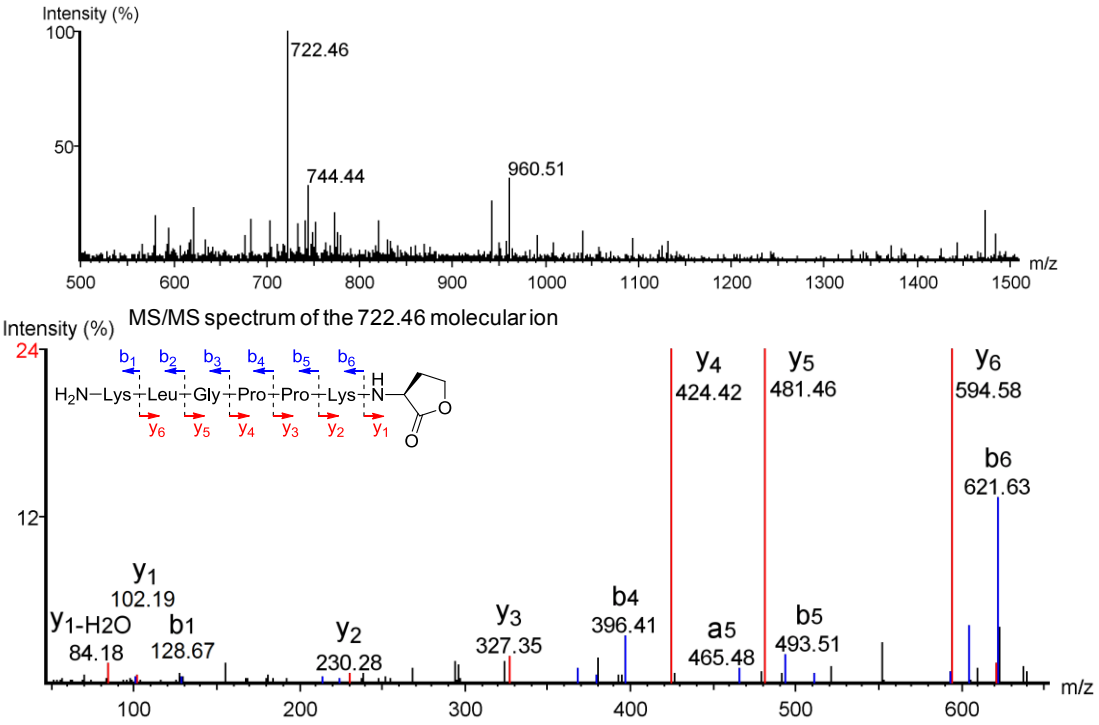


Figure S8. (Continued)

Bead #3



Bead #4

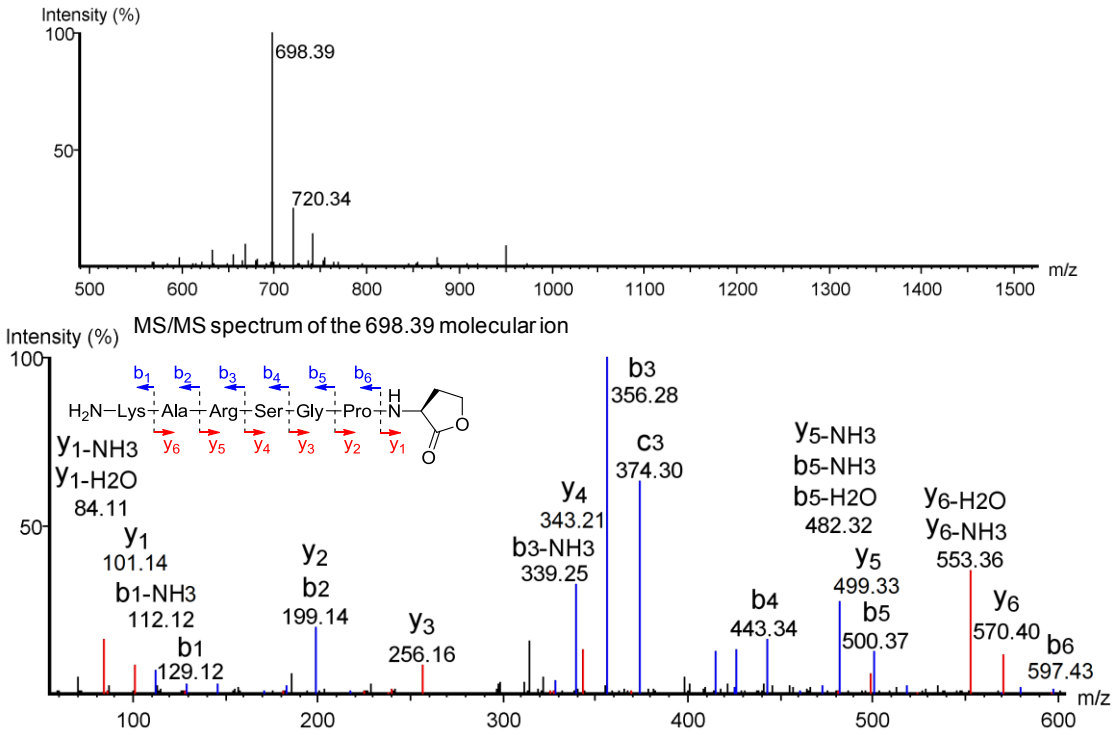
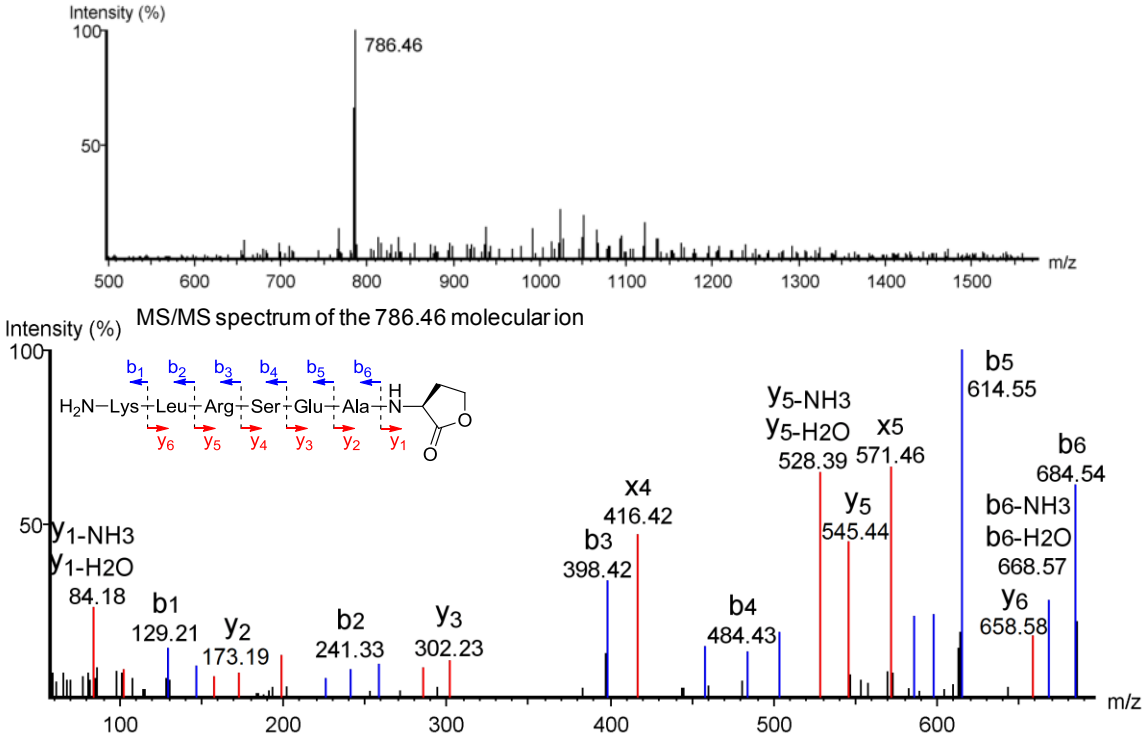


Figure S8. (Continued)

Bead #5



Chapter 5

One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptide Libraries

Forward

Photocleavable residue, 3-amino-3-(2-nitrophenyl)propionic acid (ANP), which is compatible with common acid and base-labile protecting groups is used as a linker in the solid-phase synthesis. Inspired by the methionine based method (Chapter 3), we designed a simultaneous ring-opening and cleavage approach using ANP as a linker and within the macrocycle from the resin upon UV irradiation. It is a chemical reagent free approach reducing the need for post-screening chemical modification for sequencing and allows a fast sequence determination of cyclic peptides from combinatorial libraries by MS/MS.

My contribution of this article was first to introduce two ANP residues in cyclic peptides. A part of the work, testing different cleavage conditions and the synthesis of a small cyclic peptide library, was done by Simon Vézina-Dawod and François Bédard. I also did sequencing determination of this small library using de novo sequencing. Finally, I drafted the preliminary manuscript and Éric Biron edited the final version. This article was published in the journal Org. Lett. (March 4th 2016).

Résumé

Les peptides cycliques sont des outils d'intérêt et d'une grande utilité en chimie biologique et médicinale pour étudier et moduler les interactions protéine-protéine. L'énorme diversité moléculaire qui est accessible avec les peptides cycliques a propulsé leur utilisation en chimie combinatoire.L'approche «one-bead-one-compound» (OBOC) est une méthode de criblage à haut débit très performante. Par contre, l'utilisation des peptides cycliques dans l'approche OBOC est limitée par les difficultés à séquencer les composés retenus après le criblage. En effet, l'absence d'amine libre en N-terminal rend la dégradation d'Edman impossible et le patron de fragmentation en spectrométrie de masse en tandem (MS/MS) trop complexe pour élucider la structure. Notre approche est d'incorporer un résidu photosensible au sein du macrocycle et comme ancrage afin de linéariser suite à la réouverture du cycle et de cliver les peptides du support simultanément. Les candidats linéarisés ainsi retenus peuvent être séquencés par MS/MS.

Org. Lett. 2016, **18**, 1174−1177

One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptide Libraries

Xinxia Liang, ‡[ab] Simon Vézina-Dawod, ‡[ab] François Bédard,[ab] Karine Porte[ab] and Eric Biron*[ab]

Faculty of Pharmacy, Université Laval, Pavillon Ferdinand-Vandry, Québec, Québec, G1V 0A6, Canada

Laboratory of Medicinal Chemistry, CHU de Québec Research Center (CHUL Section), 2705 Boulevard Laurier, Québec, Québec, G1V 4G2, Canada

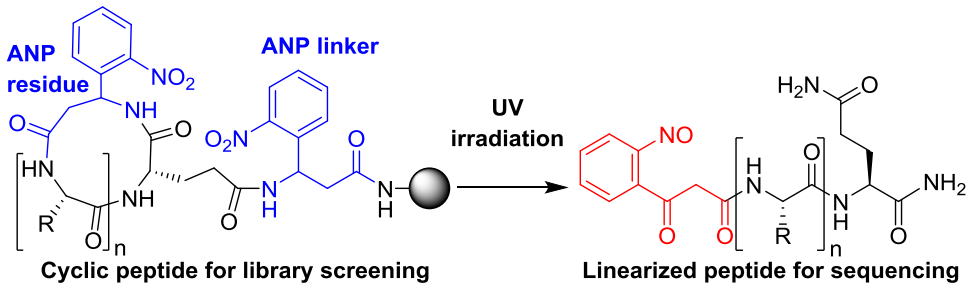
‡ X.L. and S.V.-D. contributed equally.

*Corresponding author:

Prof. Éric Biron, Laboratory of Medicinal Chemistry, CHU de Québec Research Centre (CHUL Section), 2705 Boulevard Laurier, Québec, Québec, G1V 4G2, Canada, Phone: 1 418 524 2296; Fax: 1 418 654 2210, E-mail: Eric.Biron@pha.ulaval.ca

ABSTRACT

A novel dual ring-opening/cleavage strategy to determine the sequence of cyclic peptides from one-bead one-compound libraries is described. The approach uses a photolabile residue within the macrocycle and as a linker to allow a simultaneous ring opening and cleavage from the beads upon UV irradiation and provide linearized molecules. Cyclic peptides of five to nine residues were synthesized and the generated linear peptides successfully sequenced by tandem mass spectrometry.

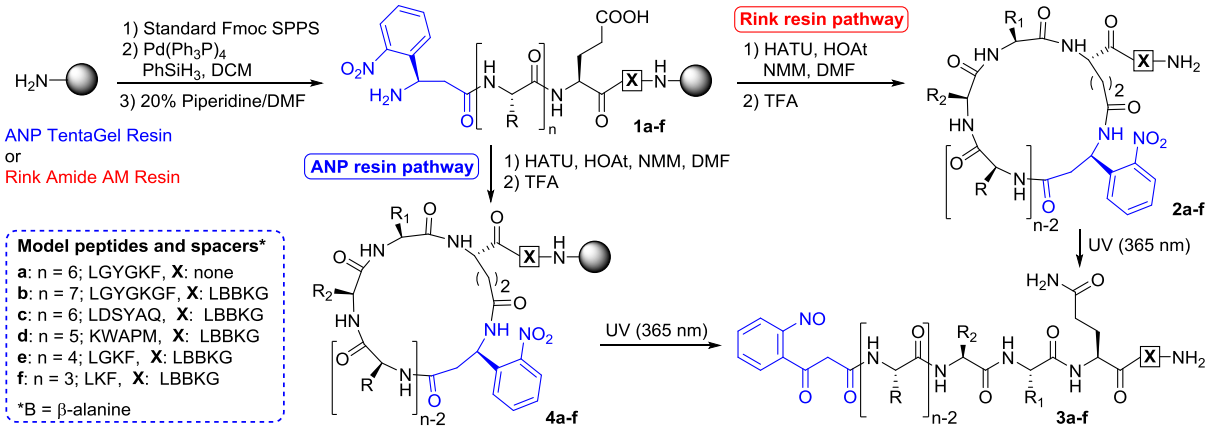


Peptide macrocycles are useful tools in chemical biology and medicinal chemistry to study and modulate protein functions.[1] With great potential as therapeutic agents, they have gained a great deal of interest in drug discovery.[2] Compared to their linear counterparts, cyclic peptides are more resistant to proteases, and their increased conformational rigidity lowers the entropic cost of binding, making them tighter binding to a given macromolecule.[3] The great degree of molecular diversity and complexity that can be accessed by simple changes in their sequence has prompted the use of cyclic peptides in combinatorial chemistry. The one bead, one compound (OBOC) approach, in which each bead carries many copies of a unique compound, has become a powerful tool in the drug discovery process.[4] However, the use of cyclic peptides in combinatorial OBOC libraries has been limited by difficulties in sequencing hit compounds after the screening. Lacking a free N-terminal amine, Edman degradation sequencing cannot be used on cyclic peptides, and complicated fragmentation patterns are obtained by tandem mass spectrometry (MS/MS).[5]

In this regard, a one bead, two compound approach on topologically segregated bilayer beads has been developed.[6] Initially introduced with an enzymatic shaving strategy[7] and later via chemical approaches, [8] topological bilayer segregation offers the opportunity to synthesize two compounds per bead, namely one which is exposed on the bead surface for screening (cyclic peptide) and the other found inside as a tag for sequencing and compound identification (linear peptide). More recently, Lim and co-workers reported a ring-opening strategy on cyclic peptoids to eliminate the need for encoding.[9] The approach involved the introduction of a cleavable alkylthioaryl bridge in the cycle to allow linearization of the molecule after the screening by thioether oxidation followed by nucleophilic displacement of the sulfone to generate a linear peptoid which can be sequenced by MS/MS. An attractive advantage of the ring-opening strategy over encoding methods is the absence of interference by the coding tag during screening since the same molecules are displayed inside and on the surface of the beads. Based on this strategy, we and other groups placed a methionine or a thioether bridge in cyclic peptides or peptoids and as a linker to allow a simultaneous linearization and compound release from the bead upon treatment with cyanogen bromide.[10] A similar approach was applied to cyclic depsipeptides where an aminolysis with NH_3 or hydrolysis with aqueous NaOH was used to cleave the ester bonds

and release the linear compound from the bead.[11] Most reported methods require aggressive chemical reagents or postscreening reactions that could lead to side-chain modifications. Based on these strategies, we were looking for an efficient, single-step, and reagent-free ring-opening approach that would be compatible with free amino acid side chains. Here, we report a novel and convenient approach for fast and simple sequence determination of cyclic peptides from OBOC libraries.

Our approach utilizes the photocleavable β-amino acid 3-amino-3-(2-nitrophenyl)propionic acid (ANP)[12] as a linker and within the macrocycle (Scheme 1). The strategy is to use UV irradiation to simultaneously convert cyclic peptides into their linear counterpart and release them from the beads. The generated linear peptides could then be sequenced by MS/MS. To evaluate the ring-opening upon UV irradiation, a first series of cyclic peptides of different sizes containing various functionalized amino acids and with or without a C-terminal spacer commonly used in OBOC libraries were synthesized on Rink Amide AM polystyrene resin. Scheme 1 illustrates the general procedure for the formation of cyclic peptides.



Scheme 1. Synthetic Route to Cyclic Peptides with ANP Residue and Ring-Opening/Cleavage Reaction

Linear peptides were synthesized on solid support by standard Fmoc chemistry. After removing the allyl and Fmoc protecting groups, the peptides **1a-f** were cyclized on resin with HATU. The chloranil test was used to qualitatively monitor the reaction.[13] After cyclization, the side chains protecting groups were removed and the compounds cleaved from the resin with a TFA cocktail. The released cyclic peptides **2a-f** were then analyzed by HPLC and electrospray ionization MS (ESI-MS) to confirm the absence of linear peptides (Figure 1A and Figure S3). Cyclic peptides **2a-f** were then subjected to UV-irradiation at 365 nm in MeOH for ring cleavage. HPLC analyses confirmed the transformation of the cyclic peptides into their linear counterpart **3a-f** (Figure 1B and Figure S4). While a mixture of different products was observed in most cases, a doublet shaped main peak with both peaks having the same mass was noticed for **3a**. This result led us to presume the formation of diastereomers during ring-opening. Moreover, even if highly pure products were not observed in the HPLC analyses for other model peptides, ESI-MS spectra showed the presence of a single major molecular specie (Figure 1B and Figure S4).

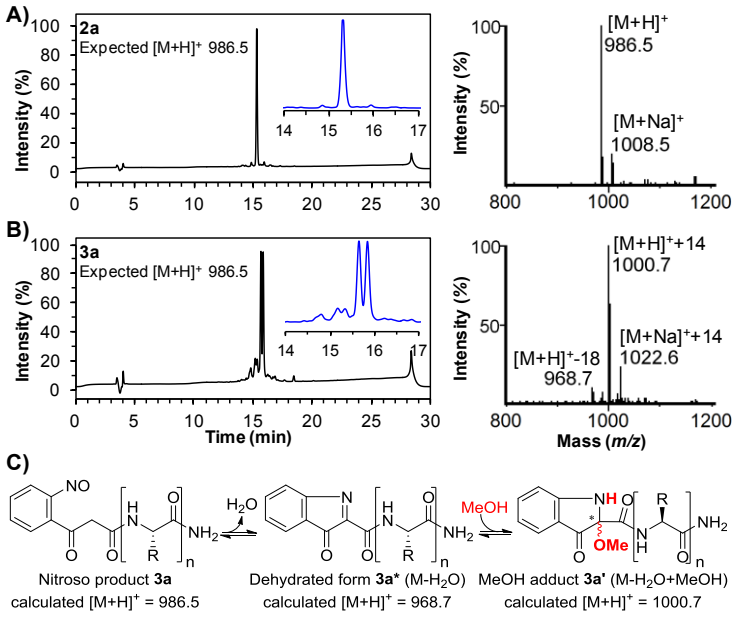


Figure 1. HPLC and ESI-MS profiles of crude products showing cyclization and ring-opening. (A) cyclo[ANP-LGYGKFE]-NH2 **2a**; and (B) 3-oxo-3-(2-nitrosophenyl)-propionyl-LGYGKFQ-NH2 **3a**. (C) Proposed structures for the dehydrated product **3a*** and its adduct **3a'** observed during ESI-MS analysis.

Since the cyclic and linearized peptides have exactly the same elemental composition, the use of MS to differentiate cyclic **2a-f** from linear compounds **3a-f** is theoretically limited. Surprisingly, after MS analysis of the generated linear peptides **3a-f**, the anticipated molecular ion was not obtained and the most important peak was observed at +14 Da (Figure 1B and Figure S4). This observation led us to suspect the formation of a C-terminal methyl ester as described by other groups[10c, 14] or the transformation of the expected N-terminal nitroso product into an indolin-3-one residue.[14] The formation of a methyl ester from the C-terminal amide would have generated a +15 Da adduct and could not explain the presence of diastereoisomers in the HPLC chromatograms. Further analysis of the photocleavage reaction mechanism of ortho-nitrophenyl derivatives led us to propose that the +14 Da peak is generated during UV irradiation from the addition of MeOH (32 Da) on a dehydrated intermediate **3*** ($[M-H_2O+H]^+$) of the nitroso derivative **3** (Figure 1C). The adduct product $M-H_2O+MeOH$ **3'** could be formed by an Ehrlich-Sachs-like reaction and as a result generate two epimers (Figure 1C and Figure S1).[15] This hypothesis was supported by the following results with peptide **3a**: 1) when methanol-d4 (36 Da) was used as solvent, the adduct peak was observed at +18 Da (Figure S5); 2) addition of $BuNH_2$ (73 Da) to the MeOH (5% v/v) during UV irradiation, yielded the adduct peak at +55 Da (Figure S5); 3) only the M-18 peak was observed in absence of MeOH during MALDI-TOF MS analysis and 4) all *b* ions (N-terminal fragments) were dehydrated (*b*-18) while the expected mass was observed for the *y* ions (C-terminal fragments) in the MS/MS spectrum (Figure 2B). Further investigations to understand this reaction are underway.

To our knowledge, besides the rearrangement observed with *o*-nitrotyrosine,[14] such rearrangements or modifications of the *o*-nitrosobenzyl moiety in ANP have not been previously reported. Since ANP is most commonly used as a linker and its *o*-nitrobenzyl analogs as protecting groups, little effort has been made to understand what happen to the *o*-nitrosobenzyl derivative after UV irradiation and characterize the products. For this reason, the post-cleavage modifications observed in this study are very attractive as they could be exploited to produce specific adducts to facilitate peptide sequencing. Moreover, a better understanding and control of post-cleavage rearrangements or modifications of the *o*-nitrosobenzyl derivative could lead to the development of interesting synthetic approaches to access polysubstituted heterocycles.

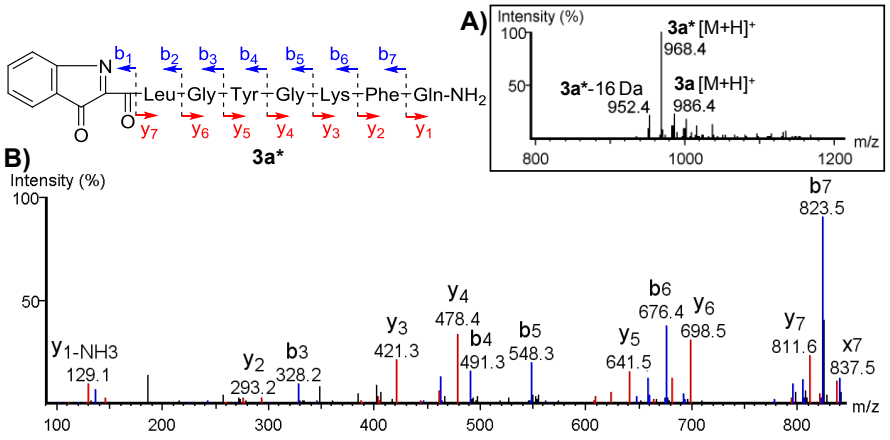


Figure 2. MALDI-TOF MS and MS/MS spectra obtained after dual ring opening/cleavage on a single bead of cyclic peptide **4a**. (A) MS of the crude product; (B) MS/MS of the dehydrated specie **3a*** (968.4 Da) [**3a**-18+H]$^+$.

To evaluate the efficiency of the simultaneous ring-opening/cleavage strategy, the model peptides **1a-f** were synthesized on TentaGel S NH_2 resin (130 µm) bearing the ANP as linker (Scheme 1). After cyclization and side chains deprotection, a small amount of resin **4a-f** was subjected to UV-irradiation and the released products analysed by HPLC and ESI-MS. The results confirmed the presence of linear peptides **3'a-f** and showed high degrees of similarity with HPLC profiles and ESI-MS spectra obtained from the Rink resin pathway. Next, a single bead was picked up from resins **4a-f** and exposed to UV-irradiation in MeOH. The crude products released from each single bead were immediately subjected to MALDI-TOF MS. In this case, the mass spectra showed very little MeOH adduct 3' ($M+H^+$+14) but the dehydrated product **3*** ($M-H_2O+H^+$) was observed as the major peak (Figure 2A and Figure S6). These results suggest that the adduct product is modified into the dehydrated specie when exposed to the laser under high vacuum during ionization in the MALDI instrument. With a wavelength of 355 nm, the ability of the MALDI's Nd:YAG laser to induce photochemical reaction has been shown[16] and used by some research groups for monitoring and structural elucidation of biomolecules.[17] More recently, Luyt's group used this "on-target" approach to sequence linear peptides on beads with a MALDI-TOF instrument equipped with a Nd:YAG laser.[18] In most of these cases, since the *o*-nitrobenzyl groups were used as linkers or protecting groups, only the peak corresponding to the desired molecule was analyzed and the residual photolabile group not characterized further.

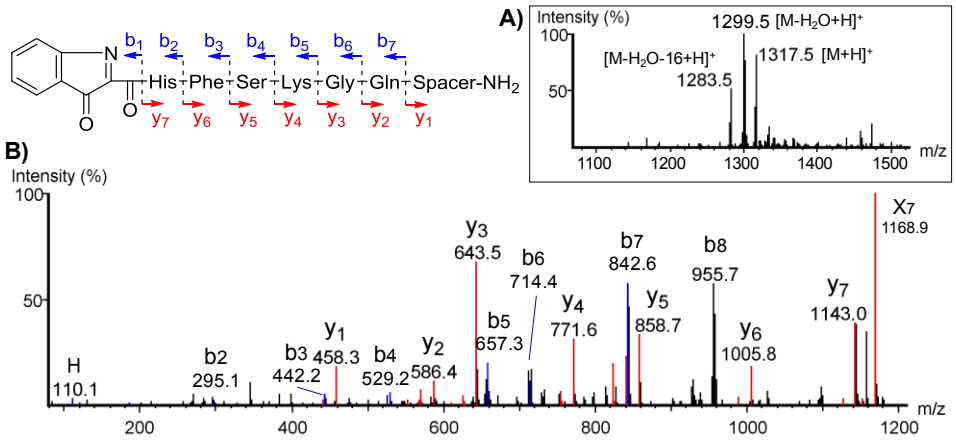


Figure 3. MS and MS/MS spectra of ANP*-HFSKGQLBBKG-NH_2 after dual ring opening/cleavage on a randomly selected bead from the OBOC cyclic peptide library. (A) MS; (B) MS/MS for precursor ion *m/z* 1299.5. (B = β-alanine).

In our study, the photosensitive residue remains attached to the peptide N-terminus and a common pattern containing a series of three peaks was observed on most MS spectra. Formed by a **3** $[M+H]^+$ peak, a ^{3*} [**3**-18+H]$^+$ peak and a **3***-16 Da peak (Figure 2A and Figure S6), this signature can be very useful to identify the most efficient molecular ion for sequencing unknown peptides from OBOC libraries. Unfortunately, we were not able to structurally define the **3***-16 peak but MS/MS spectra analyses of **3a-f** showed that this modification also happens on the ANP residue after UV irradiation and exposition to the Nd:YAG laser. Further studies are underway to characterize this modification. Compared to data obtained by ESI-MS after dual photocleavage of peptide **4d**, a significant level of Met oxidation was observed in the MALDI MS spectrum (Figure S6). With the characteristic isotopic distribution of the sulfur atom, the presence of oxidized species in the MALDI MS spectrum can be considered as a good indicator for the presence of a Met residue in the sequence.

MS/MS analysis of the dehydrated product molecular ions yielded high-quality spectra from which the linearized peptides **3*a-f** could be unambiguously sequenced manually and by using de novo sequencing with the PEAKS software (Figure 2B and Figure S6).19 In comparison, MS/MS spectra of most **3a-f** $[M+H]^+$ molecular ion could not be efficiently sequenced. The obtained complex fragmentation patterns for **3a-f** suggest that macrocyclic

structural isomers **2a-f** were remaining in the mixture. The presence of a C-terminal spacer in the peptides was very helpful in the sequencing process. With a fixed known mass, it allowed us to initiate sequencing of *y* ions from 458.35 Da and avoid often missing low mass *y* ions. On the other hand, since the identity of the five C-terminal amino acids is known, we were able to properly initiate sequencing of the *b* ions from the precursor molecular ions.

To demonstrate the compatibility of our strategy with OBOC libraries, a small cyclic heptapeptide library was prepared on 100 mg of ANP TG resin bearing the spacer Leu-βAla-βAla-Lys-Gly. The library was prepared by split-and-pool synthesis using standard Fmoc/tBu solid-phase peptide chemistry.[4b, 20] Following Glu(OAll), the next six positions within the peptide library were filled by a random combination of seventeen L-amino-acids (Cys, Gln and Ile were excluded). Finally, Fmoc-D-ANP-OH was added and the peptide cyclized after selective deprotection as described above. After removing the side chain protecting groups, 15 beads were randomly selected and individually irradiated at 365 nm in MeOH for 3 h. The resulting crude peptides were analyzed by MALDI-TOF MS. For each selected bead, the resulting MS spectrum showed the three peaks signature observed for the model peptides with the dehydrated specie as the most important peak (Figure 3A and Figure S7). The peptides could be unambiguously sequenced by MS/MS of the dehydrated specie by manual analysis and/or de novo sequencing with the PEAKS software (Figure 3B and Figure S7).[19] Some beads showed oxidized species in the MS spectrum suggesting the presence of Met in the sequence. In some case, the oxidized [M-18] peak was successfully sequenced by MS/MS when Met and Trp oxidation were included as post-translational modification (PTM) during de novo sequencing with the PEAKS software (Bead #3, Figure S7). These results demonstrate that the developed dual ring-opening/cleavage strategy is compatible with the side chains of commonly used amino acids and can be used on a single bead to release linear peptides that can be clearly and conclusively sequenced by MS/MS. In addition to not having to use aggressive chemical reagents for ring-opening and cleavage, another advantage of the developed approach over some other reported strategies is the presence of a spacer sequence on the released linearized peptide. This spacer facilitates MS/MS spectra analysis since its composition is known and constitutes a reliable starting point for sequencing. It can also be exploited to

increase the molecular mass and improve ionization of *y* ions by adding positively charged residues in the spacer. The procedure was performed on a freshly prepared library but it is strongly recommended to protect the library from light during synthesis and handling to prevent opening of the macrocycle or cleavage from the resin. It is also important to consider that hit compounds identified after OBOC library screening will contain an ANP residue within the macrocycle. Therefore, since this photolabile monomer can have an impact on the stability of hit macrocyles in bioassays, the ANP could be replaced by a photostable bioisostere such as 3-amino-3-(2-cyanophenyl)propionic acid (ACP) when hit compounds are resynthesized to avoid ring-opening during binding and biological assays.

In summary, we report the use of a photolabile residue within a macrocyclic peptide and as a linker to allow a one-pot ring-opening/cleavage reaction upon UV irradiation and provide linearized peptides that can be efficiently sequenced by MS/MS. While the ANP linker is well known in combinatorial chemistry, we demonstrated that the generated o-nitrosobenzyl product during UV irradiation undergoes a rearrangement leading to an indolin-3-one moiety and adducts that can be used for sequencing. Compatible with commonly used amino acids, the described approach avoids the use of harsh chemical reagents and post-screening reactions to produce a linearized molecule and allows a fast sequence determination of cyclic peptides from OBOC combinatorial libraries by MS/MS. Simple and affordable for any peptide science or combinatorial chemistry laboratory, the described dual ring-opening/cleavage strategy will be useful for the preparation and screening of OBOC macrocyclic peptide libraries.

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). The China Scholarship Council (XL) and NSERC of Canada (SVD, FB) and are acknowledged for under- and postgraduate scholarships. The authors are grateful to Isabelle Kelly of the Proteomics Core Facility of CHU de Québec Research Centre for MALDI-TOF/TOF analyses.

REFERENCES

1. (a) Thapa, P.; Espiritu, M.; Cabalteja, C.; Bingham, J.-P. *Int. J. Pept. Res. Ther.* **2014**, *20*, 545-551. (b) Liskamp, R. M. J.; Rijkers, D. T. S.; Kruijtzer, J. A. W.; Kemmink, J. *ChemBioChem* **2011**, *12*, 1626-1653. (c) Marsault, E.; Peterson, M. L. *J. Med. Chem.* **2011**, 54, 1961-2004. (d) Hamada, Y.; Shioiri, T. *Chem. Rev.* **2005**, *105*, 4441-4482.
2. (a) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. Nature Reviews. *Drug Discovery* **2008**, *7*, 608-624. (b) Mallinson, J.; Collins, I. *Future Med. Chem*. **2012**, *4*, 1409-1438. (c) Katsara, M.; Tselios, T.; Deraos, S.; Deraos, G.; Matsoukas, M. T.; Lazoura, E.; Matsoukas, J.; Apostolopoulos, V. *Curr. Med. Chem*. **2006**, *13*, 2221-2232.
3. (a) Tyndall, J. D. A.; Nall, T.; Fairlie, D. P. *Chem. Rev*. **2005**, *105*, 973-999. (b) Adessi, C.; Soto, C. *Curr. Med. Chem*. **2002**, *9*, 963-978. (c) Gilon, C.; Halle, D.; Chorev, M.; Selincer, Z.; Byk, G. *Biopolymers* **1991**, *31*, 745-750.
4. (a) Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411-448. (b) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84.
5. (a) Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. *J. Comb. Chem.* **2003**, *5*, 33-40. (b) Ngoka, L. M.; Gross, M. *J. Am. Soc. Mass. Spectrom.* **1999**, *10*, 732-746.
6. Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. *J. Am. Chem. Soc.* **2006**, *128*, 13000-13009.
7. Vágner, J.; Barany, G.; Lam, K. S.; Krchnák, V.; Sepetov, N. F.; Ostrem, J. A.; Strop, P.; Lebl, M. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8194-8199.
8. (a) Liu, R.; Marik, J.; Lam, K. S. *J. Am. Chem. Soc.* **2002**, *124*, 7678-7680. (b) Wang, X.; Peng, L.; Liu, R.; Gill, S. S.; Lam, K. S. *J. Comb. Chem.* **2005**, *7*, 197-209. (c) Liu, R.; Wang, X.; Song, A.; Bao, T.; Lam, K. S. *QSAR Comb. Sci.* **2005**, *24*, 1127-1140.
9. (a) Lee, J. H.; Meyer, A. M.; Lim, H.-S. *Chem. Commun.* **2010**, *46*, 8615-8617. (b) Lee, J. H.; Kim, H.-S.; Lim, H.-S. *Org. Lett.* **2011**, *13*, 5012-5015.
10. (a) Simpson, L. S.; Kodadek, T. *Tetrahedron Lett.* **2012**, *53*, 2341-2344. (b) Liang, X. X.; Girard, A.; Biron, E. *ACS Comb. Sci.* **2013**, *15*, 535-540. (c) Lee, K. J.; Lim, H.-S. *Org. Lett.* **2014**, *16*, 5710-5713.
11. (a) Gurevich-Messina, J. M.; Giudicessi, S. L.; Martínez-Ceron, M. C.; Acosta, G.; Erra-Balsells, R.; Cascone, O.; Albericio, F.; Camperi, S. A. *J. Pept. Sci.* **2015**, *21*, 40-45. (b) Menegatti, S.; Ward, K. L.; Naik, A. D.; Kish, W. S.; Blackburn, R. K.; Carbonell, R. G. *Anal. Chem.* **2013**, *85*, 9229-9237.
12. Brown, B.; Wagner, D.; Geysen, H. M. *Mol. Diversity* **1995**, *1*, 4-12.
13. Vojkovsky, T. *Pept. Res.* **1995**, *8*, 236-237.
14. Peters, F. B.; Brock, A.; Wang, J.; Schultz, P. G. *Chem. Biol*. **2009**, *16*, 148-152.
15. Ehrlich, P.; Sachs, F. Ber. Dtsch. *Chem. Ges*. **1899**, *32*, 2341-2346.

16. Fitzgerald, M. C.; Harris, K.; Shevlin, C. G.; Siuzdak, G. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 979-982.

17. (a) Guerlavais, T.; Meyer, A.; Debart, F.; Imbach, J. L.; Morvan, F.; Vasseur, J. J. *Anal. Bioanal. Chem.* **2002**, *374*, 57-63. (b) Nicolaou, K. C.; Watanabe, N.; Li, J.; Pastor, J.; Winssinger, N. *Angew. Chem. Int. Ed.* **1998**, *37*, 1559-1561. (c) St. Hilaire, P. M.; Lowary, T. L.; Meldal, M.; Bock, K. *J. Am. Chem. Soc.* **1998**, *120*, 13312-13320. (d) Aubagnac, J. L.; Enjalbal, C.; Subra, G.; Bray, A. M.; Combarieu, R.; Martinez, J. *J. Mass Spectrom*. **1998**, *33*, 1094-1103.

18. Amadei, G. A.; Cho, C.-F.; Lewis, J. D.; Luyt, L. G. *J. Mass Spectrom.* **2010**, *45*, 241-251.

19. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2337-2342.

20. (a) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86. (b) Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487-493.

.

Supporting Information (SI)

One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptides Libraries

Xinxia Liang, Simon Vézina-Dawod, François Bédard, Karine Porte and Eric Biron*

Faculty of Pharmacy, Université Laval, Pavillon Ferdinand-Vandry, Quebec (QC) G1V 0A6 Canada

Laboratory of Medicinal Chemistry, CHU de Québec Research Centre (CHUL Section), 2705 Boulevard Laurier, Quebec (QC), G1V 4G2, Canada

eric.biron@pha.ulaval.ca

ELECTRONIC SUPPLEMENTARY INFORMATION

Contents

Materials and equipment 130

Peptide synthesis 130

Figure S1: Proposed mechanism of photocleavage reaction and side product formation

Preparation of the OBOC library and references 136

Figure S2. HPLC profiles and ESI-MS spectra of linear precursors **1a-f** 137

Figure S3. HPLC profiles and ESI-MS spectra of cyclic peptides **2a-f** 140

Figure S4. HPLC profiles and ESI-MS spectra of linear peptides **3a-f** after ring-opening 143

Figure S5. HPLC profiles and ESI-MS spectra of linear peptides adducts after ring-opening 146

Figure S6. MALDI MS and MS/MS spectra of linear peptides **3a-f** obtained from dual ring-opening/cleavage on a single TG bead 147

Figure S7. MALDI MS and MS/MS spectra of the randomly selected beads from the cyclic peptide library after dual ring-opening/cleavage 150

Materials and equipment

TentaGel S NH_2 (TG) resin (130 μm, 0.29 mmol/g) was purchased from Rapp Polymer (Tübingen, Germany). Fmoc-protected amino acids and coupling reagents, 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) and 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) were purchased from Matrix Innovation (Quebec, QC, Canada). Rink Amide AM resin (0.65 mmol/g) and 3-N^{α}-Fmoc-amino-3-(2-nitrophenyl)propionic acid (ANP) were purchased from Chem-Impex International (Wood Dale, IL, USA). All other commercial reagents and solvents were purchased from commercial suppliers and used without further purification. Reactions on solid support were performed in polypropylene fritted syringes from Roland Vetter Labordedarf OHG (Ammerbuch, Germany). Photochemistry was performed with a Spectroline UV lamp (ENF-260C 115V 60Hz 0.2A). RP-LC/MS analyses were conducted on Shimadzu Prominence LCMS-2020 system equipped with an electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) probes using a Gemini-NX column (4.6 mm × 250 mm, 5 μm C_{18}, 110Å, 1 mL/min) with a 20 min linear gradient for **1a**, **2a** and **3a** or a Kinetex column (4.6 mm x 100 mm, 2.6 μm XB-C_{18}, 100 Å, 1.8 mL/min) with a 10.5 min gradient for **1b-f**, **2b-f** and **3b-f** from water (0.1% TFA) and CH_3CN (0.1% TFA) (CH_3CN 10-100%) and detection at 220 nm and 254 nm. Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry was performed on a AB SCIEX 4800 Plus MALDI-TOF/TOF instrument using alpha-cyano-4-hydroxycinnamic acid as matrix. The spectra were acquired using the 4000 Series Explorer Software (AbSciex, v 3.2.3). The PEAKS Studio software (Bioinformatics Solutions, v.7.0) was used for spectra analysis and DENOVO sequencing.

Peptide synthesis

Peptides were synthesized by standard Fmoc solid phase synthesis.[1] Briefly, amino acid couplings were performed with a solution of Fmoc-Xaa-OH (3 equiv), HCTU (3 equiv) and NMM (6 equiv) in DMF for 30 min. The coupling step was repeated once and the resin washed with DMF (5 × 30 s). The Fmoc protecting group was removed by treating the resin twice with a solution of 20% piperidine in DMF (v/v) for 10 min followed by washing with DMF (5 × 30 s). Resins bearing ANP residues were protected from light.

Preparation of the TentaGel ANP resin. TentaGel S NH_2 130 µm (1 g, 0.29 mmol) was first swelled in 20 mL of DMF followed by addition of Fmoc-ANP-OH (3 equiv), HCTU (3 equiv) and NMM (6 equiv). The mixture was shaken mechanically for 45 min, the resin filtered and the coupling step repeated once. After filtration and washing the resin with DMF, the remaining amino groups were capped with a solution of Ac_2O/DIEA/DMF (7:2:91) for 15 min. The resin was filtered and washed thoroughly with DMF (5 × 30 s), DCM (5 × 30 s), MeOH (5 × 30 s) and dried *in vacuo*. The loading capacity was determined by the Fmoc dosage assay described by Gude et *al.* and estimated to 0.21 mmol/g.[2]

On resin peptide cyclization. The resin bearing the Fmoc-protected peptide was swelled with CH_2Cl_2 and the allyl protecting group was removed with a solution of $Pd(PPh_3)_4$ (0.25 equiv) and $PhSiH_3$ (20 equiv) in CH_2Cl_2 for 2 × 40 min. The resin was washed with CH_2Cl_2 (5 × 30 s), 0.5% DIPEA (v/v) in DMF (2 × 30 s), 0.5% diethyl dithiocarbamate (m/v) in DMF (2 × 30 s) and DMF (5 × 30 s). The Fmoc group was then removed with 20% piperidine in DMF (2 × 10 min) and the resin washed with DMF (5 × 30 s). Peptide cyclization was performed on solid support in the presence of HATU (3 equiv), HOAt (3 equiv) and NMM (6 equiv) in DMF for 3 h. Cyclization was monitored by the chloranil test.[3] After reaction completion, the resin was washed with DMF (5 × 30 s) and CH_2Cl_2 (5 × 30 s).

Cleavage from Rink Amide AM resin. The resin (25 mg) was treated with a cleavage solution composed of TFA, water and triisoproylsilane (TIS) (95:2.5:2.5) for 3 h followed by filtration and washing with CH_2Cl_2 (2 × 30 s). The filtrate was evaporated under reduced pressure and the resulting mixture precipitated with cold diethyl ether. The solid was washed twice with diethyl ether and dried under vacuum.

Ring-opening of cyclic peptides from Rink Amide AM resin. Crude cyclic peptide obtained from Rink Amide AM resin was dissolved in MeOH (1.5 mL). A UV lamp was then installed 2 cm over the sample and irradiation conducted at 365 nm for 3 h. The solvent was removed under reduced pressure and the resulting product was analyzed by HPLC and ESI-MS.

Side-chain deprotection on TentaGel ANP resin. The resin was swelled with CH_2Cl_2 and treated with a mixture of TFA/water/TIS (95:2.5:2.5) for 3 h. After side-chain deprotection, the resin was washed with CH_2Cl_2 (5 × 30 s), AcOH/CH_3CN/H_2O (3:4:3) (5 × 30 s), CH_3CN/H_2O (1:1) (5 × 30 s), MeOH (5 × 30 s) and dried under vacuum.

Ring-opening reaction and cleavage from a single bead of TentaGel ANP. Single beads were isolated one by one under microscope, placed in an open-top 96-well polypropylene plate and swelled into MeOH (200 µL). Then, a UV lamp was placed at 2 cm above the plate and irradiation conducted at 365 nm for 6 h. MeOH was added every 1-2 h to avoid extensive evaporation and prevent wells from drying. The solutions were transferred into 0.5 mL microcentrifuge tubes and the wells washed with MeOH (50 µL). The solvent was evaporated under vacuum to dryness and the peptides released from the bead were dissolved in 10 µL of 0.1% TFA in water. 1 µL of the peptide solution was mixed with 1 µL of 4-hydroxy-α-cyanocinnamic acid (5 mg/mL) in CH_3CN/0.1% TFA (1:1) and 1µL of the mixture was spotted onto a MALDI sample plate for MS/MS analysis.

H-D-ANP-Leu-Gly-Tyr-Gly-Lys-Phe-Glu-NH_2 (**1a**): 96 % purity (determined by HPLC); RP-HPLC t_R= 13.85 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{48}H_{65}N_{11}O_{13}$ ($M+H^+$) 1004.48; observed 1004.45.

H-D-ANP-Leu-Gly-Tyr-Gly-Lys-Gly-Phe-Glu-Leu-βAla-βAla-Lys-Gly-NH_2 (**1b**): 96 % purity (determined by HPLC); RP-HPLC t_R= 6.72 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{70}H_{104}N_{18}O_{19}$ ($M+H^+$) 1501.77; observed 1501.95.

H-D-ANP-Leu-Asp-Ser-Tyr-Ala-Gln-Glu-Leu-βAla-βAla-Lys-Gly-NH_2 (**1c**): 93 % purity (determined by HPLC); RP-HPLC t_R= 6.56 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{64}H_{97}N_{17}O_{22}$ ($M+H^+$) 1456.70; observed 1456.85.

H-D-ANP-Lys-Trp-Ala-Pro-Met-Glu-Leu-βAla-βAla-Lys-Gly-NH_2 (**11d**): 88 % purity (determined by HPLC); RP-HPLC t_R= 6.48 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{64}H_{97}N_{17}O_{16}S$ ($M+H^+$) 1392.70; observed 1392.85.

H-D-ANP-Leu-Gly-Lys-Phe-Glu-Leu-βAla-βAla-Lys-Gly-NH_2 (**1e**): 99 % purity (determined by HPLC); RP-HPLC t_R= 6.90 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{57}H_{89}N_{15}O_{15}$ ($M+H^+$) 1224.67; observed 1224.75.

H-D-ANP-Leu-Lys-Phe-Glu-Leu-βAla-βAla-Lys-Gly-NH_2 (**1f**): 99 % purity (determined by HPLC); RP-HPLC t_R= 6.59 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{55}H_{86}N_{14}O_{14}$ ($M+H^+$) 1167.64; observed 1167.75.

Cyclo[D-ANP-Leu-Gly-Tyr-Gly-Lys-Phe-Glu]-NH_2 (**2a**): 94 % purity (determined by HPLC); RP-HPLC t_R= 15.32 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{48}H_{64}N_{11}O_{12}$ ($M+H^+$) 986.47; observed 986.45.

Cyclo[D-ANP-Leu-Gly-Tyr-Gly-Lys-Gly-Phe-Glu]-Leu-βAla-βAla-Lys-Gly-NH_2 (**2b**): 91 % purity (determined by HPLC); RP-HPLC t_R= 7.12 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{50}H_{67}N_{12}O_{13}$ ($M+H^+$) 1483.76; observed 1483.90.

Cyclo[D-ANP-Leu-Asp-Ser-Tyr-Ala-Glu]-Leu-βAla-βAla-Lys-Gly-NH_2 (**2c**): 93 % purity (determined by HPLC); RP-HPLC t_R= 7.21 (10 - 100%); MS (ESI) (negative ionisation) *m/z*: calcd for $C_{44}H_{58}N_{11}O_{16}$ ($M-H^-$) 1439.56; observed 1439.85.

Cyclo[D-ANP-Lys-Trp-Ala-Pro-Met-Glu]-Leu-βAla-βAla-Lys-Gly-NH_2 (**2d**): 90 % purity (determined by HPLC); RP-HPLC t_R= 7.05 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{64}H_{95}N_{17}O_{15}S$ ($M+H^+$) 1374.69; observed 1374.80.

Cyclo[D-ANP-Leu-Gly-Lys-Phe-Glu]-Leu-βAla-βAla-Lys-Gly-NH_2 (**2e**): 90 % purity (determined by HPLC); RP-HPLC t_R= 7.41 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{57}H_{87}N_{15}O_{14}$ ($M+H^+$) 1206.66; observed 1206.90.

Cyclo[D-ANP-Leu-Lys-Phe-Glu]-Leu-βAla-βAla-Lys-Gly-NH_2 (**2f**): 60 % purity (determined by HPLC); RP-HPLC t_R= 7.38 (10 - 100%); MS (ESI) *m/z*: calcd for $C_{55}H_{84}N_{14}O_{13}$ ($M+H^+$) 1149.63; observed 1149.85.

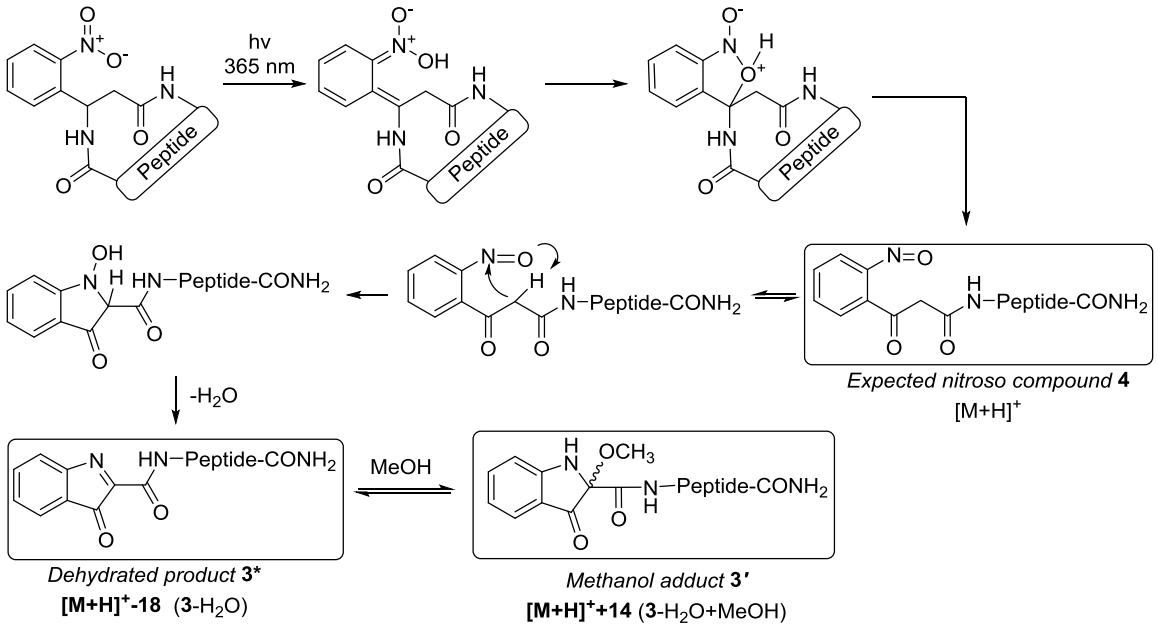
Table S1. Characterization data for linear peptides **3a-f** after ring-opening by UV irradiation

	3			**3***			**3′**		
	Calculated	*Observed* (Da)		*Calculated*	*Observed* (Da)		*Calculated*	*Observed* (Da)	
	$[M+H]^+$ (Da)	ESI	MALDI	$[M+H]^+$ (Da)	ESI	MALDI	$[M+H]^+$ (Da)	ESI	MALDI
a	$C_{48}H_{63}N_{11}O_{12}$ 986.47	986.45	986.45	$C_{48}H_{61}N_{11}O_{11}$ 968.46	968.45	968.45	$C_{49}H_{65}N_{11}O_{12}$ 1000.49	1000.45	----------
b	$C_{70}H_{102}N_{18}O_{18}$ 1483.76	----------	1483.61	$C_{70}H_{100}N_{18}O_{17}$ 1465.75	----------	1465.62	$C_{70}H_{104}N_{18}O_{18}$ 1497.77	1497.85	----------
c	$C_{64}H_{95}N_{17}O_{21}$ 1438.56	----------	----------	$C_{64}H_{93}N_{17}O_{20}$ 1420.68	----------	1420.54	$C_{65}H_{97}N_{17}O_{21}$ 1452.70	1452.75	1452.55
d	$C_{64}H_{95}N_{17}O_{15}S$ 1374.69	----------	----------	$C_{64}H_{93}N_{17}O_{14}S$ 1356.68	----------	1356.57	$C_{65}H_{97}N_{17}O_{15}S$ 1388.70	1388.75	1388.55
e	$C_{57}H_{87}N_{15}O_{14}$ 1206.66	1206.80	----------	$C_{57}H_{85}N_{15}O_{13}$ 1188.65	----------	1188.50	$C_{58}H_{89}N_{15}O_{14}$ 1220.67	1220.80	----------
f	$C_{55}H_{84}N_{14}O_{13}$ 1149.63	----------	----------	$C_{55}H_{82}N_{14}O_{12}$ 1131.62	----------	1131.49	$C_{56}H_{86}N_{14}O_{13}$ 1163.65	1163.85	----------

3* = **3**-H_2O

3′ = **3**-H_2O+MeOH

Figure S1. Proposed mechanism of photocleavage reaction and side product formation (Based on the mechanism reported for 2-nitrophenylalanine derivatives by Peters and al.).[4]



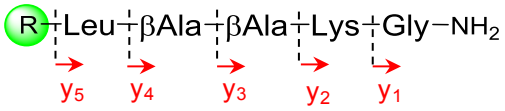
Synthesis of the random cyclic heptapeptide library

The OBOC combinatorial cyclic peptide library was prepared on 100 mg of 130 µm TentaGel ANP resin. After removal of the Fmoc group with piperidine, the small library was synthesized by the split-and-pool method using Fmoc/HCTU solid phase peptide chemistry (see above). The next five positions within the peptide library have a random combination of seventeen L-amino-acids (A, D, E, F, G, H, K, L, M, N, P, R, S, T, V, W and Y). Coupling for the last position was performed with Fmoc-ANP-OH. Following allyl ester cleavage and amino terminus deprotection, peptide cyclization was carried out with HATU as described above. The side chain protecting groups were removed by treatment with a mixture of TFA and the beads washed extensively with CH_2Cl_2 (5 × 30 s), AcOH/CH_3CN/H_2O (3:4:3) (5 × 30 s), CH_3CN/H_2O (1:1) (5 × 30 s) and MeOH (5 × 30 s).

Peptide sequencing by MS/MS

Following data acquisition, the PEAKS Studio software was used to analyze MS and MS/MS spectra and perform manual sequencing or automatic de novo sequencing with the DENOVO program in Peaks. When the C-terminal spacer Leu-βAla-βAla-Lys-Gly-NH_2 was present, the sequencing of y ions was initiated at 458.30 Da (y5 in Table S2). To avoid often missing low mass y ions, this ion at 458.30 Da was considered as y1 in the sequencing of peptides containing the C-terminal spacer. On the other hand, since the identity of the five C-terminal amino acids is known, sequencing of the b ions was initiated from the precursor molecular ions.

Table S2. Calculated mass for the *y* ions from the C-terminal spacer



	*y*1	*y*2	*y*3	*y*4	*y*5
y ions	75.05	203.15	274.18	345.22	458.30

References

1) Fields, G.B.; Noble, R.L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* **1990**, *35*, 161-214.

2) Gude, M.; Ryf, J.; White, P.D. An accurate method for the quantification of Fmoc-derivatized solid phase supports. *Lett. Pept. Sci.* **2002**, *9*, 203-206.

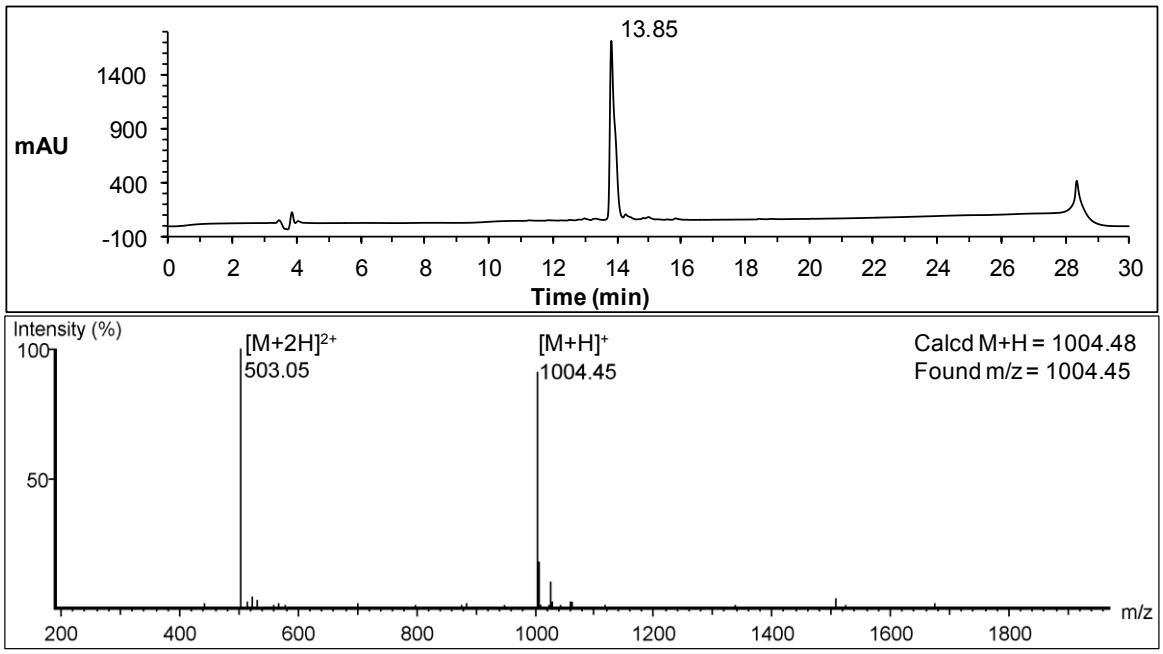
3) Gaggini, F.; Porcheddu, A.; Reginato, G.; Rodriguez, M.; Taddei, M. Colorimetric tools for solid-phase organic synthesis. *J. Comb. Chem.* **2004**, *6*, 805-810.

4) Peters, F.B.; Brock, A.; Wang, J.; Schultz P.G. Photocleavage of the polypeptide backbone by 2-nitrophenylalanine. *Chem. Biol.* **2009**, *16*, 148-152.

Figure S2. HPLC profiles (λ = 220 nm) and ESI-MS spectra of linear peptides **1a-f**.

H−ANP−Leu−Gly−Tyr−Gly−Lys−Phe−Glu−NH_2

1a



H−ANP−Leu−Gly−Tyr−Gly−Lys−Gly−Phe−Glu−Leu−βAla−βAla−Lys−Gly−NH_2

1b

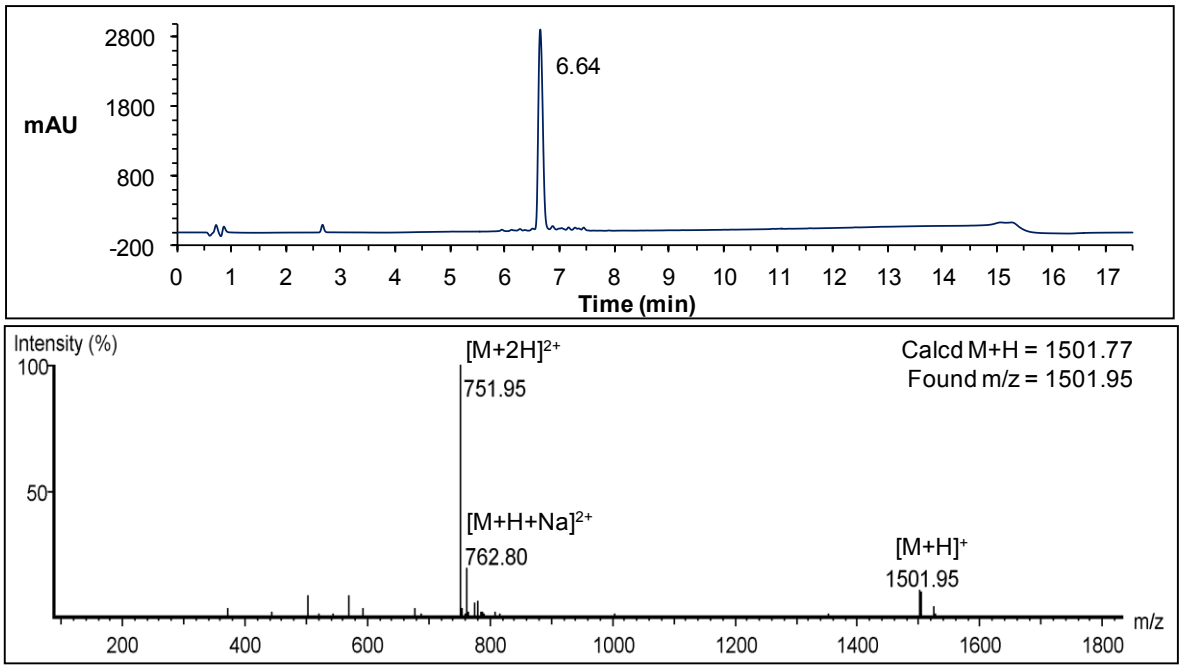
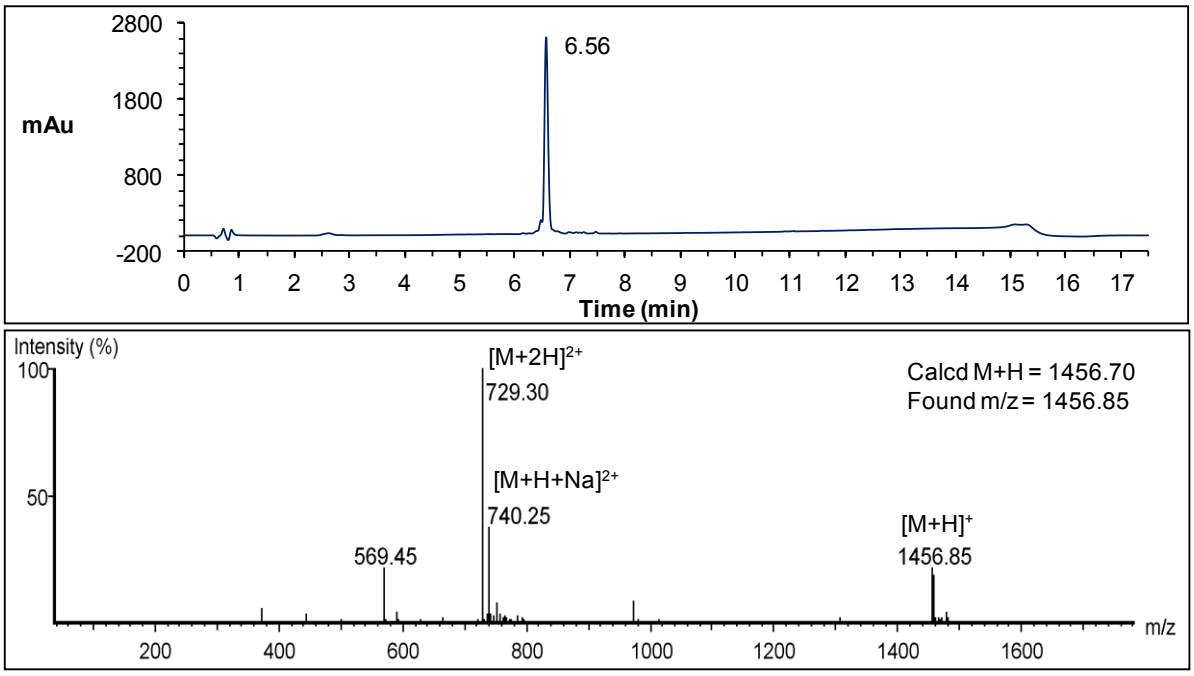


Figure S2. (Continued)

H−ANP−Leu−Asp−Ser−Tyr−Ala−Gln−Glu−Leu−βAla−βAla−Lys−Gly−NH_2

1c



H−ANP−Lys−Trp−Ala−Pro−Met−Glu−Leu−βAla−βAla−Lys−Gly−NH_2

1d

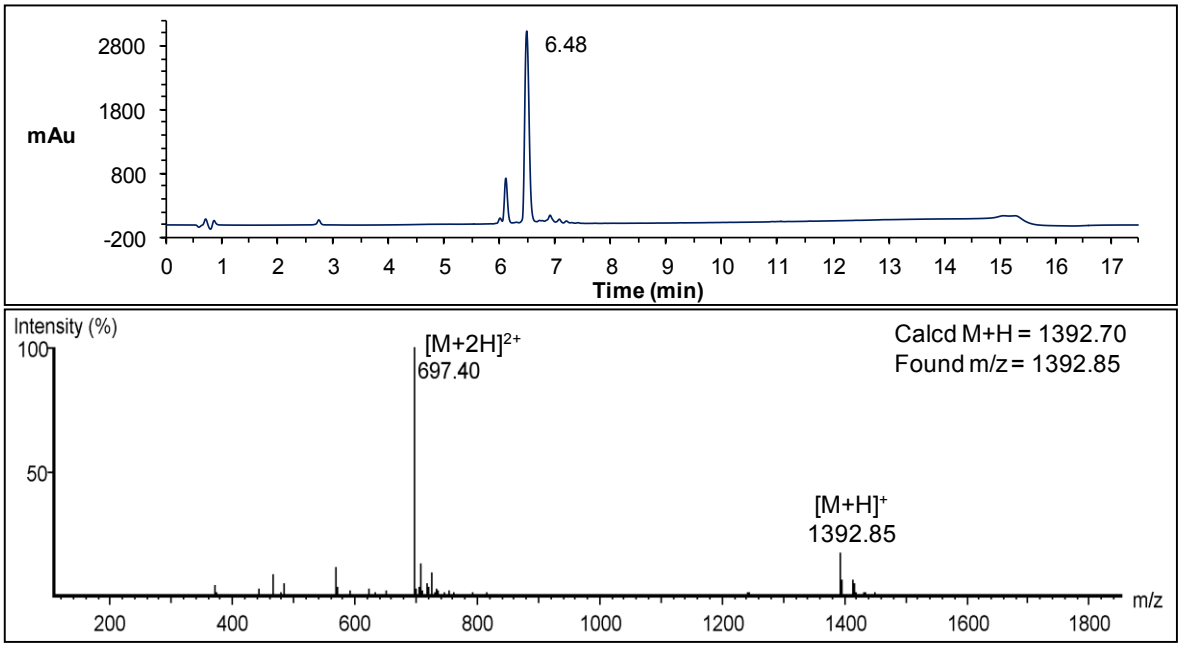
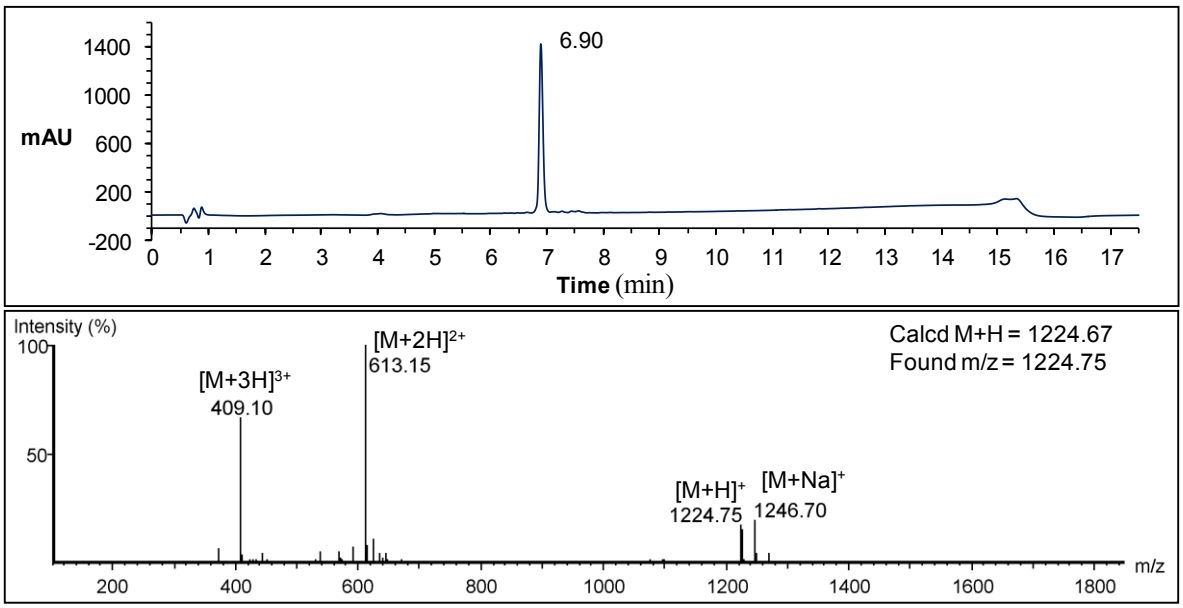


Figure S2. (Continued)

H−ANP−Leu−Gly−Lys−Phe−Glu−Leu−βAla−βAla−Lys−Gly−NH_2

1e



H−ANP−Leu−Lys−Phe−Glu−Leu−βAla−βAla−Lys−Gly−NH_2

1f

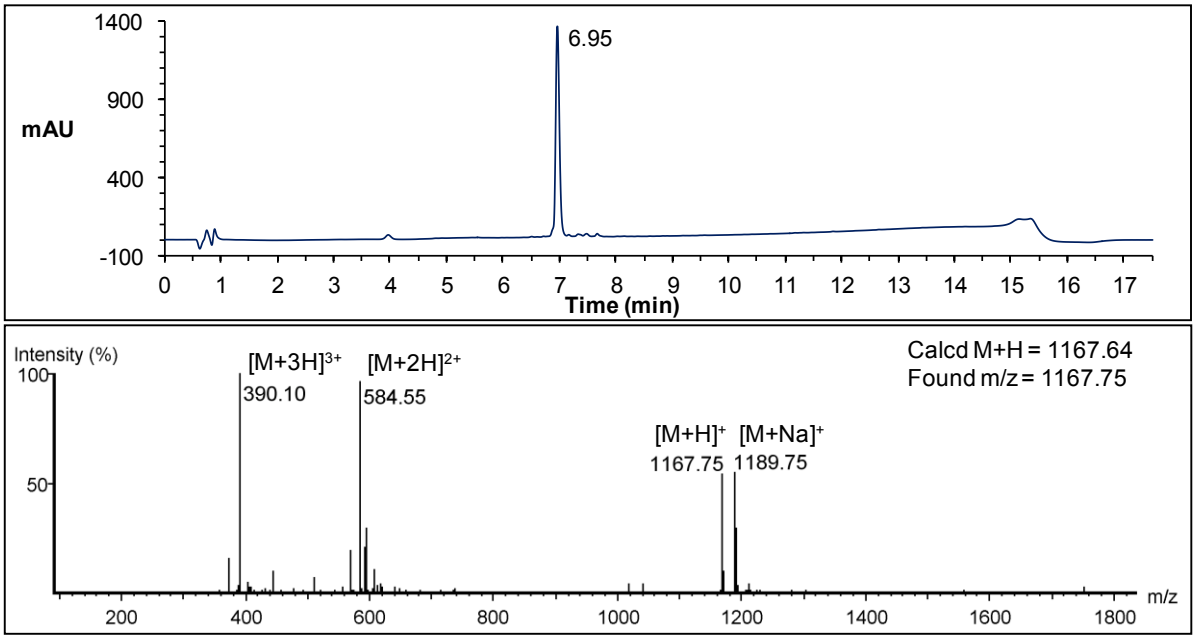
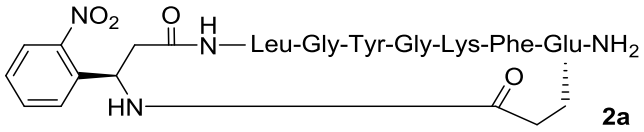
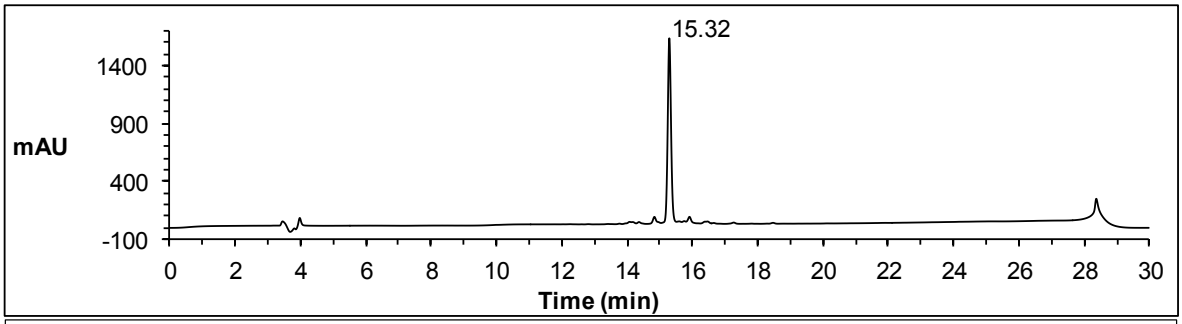
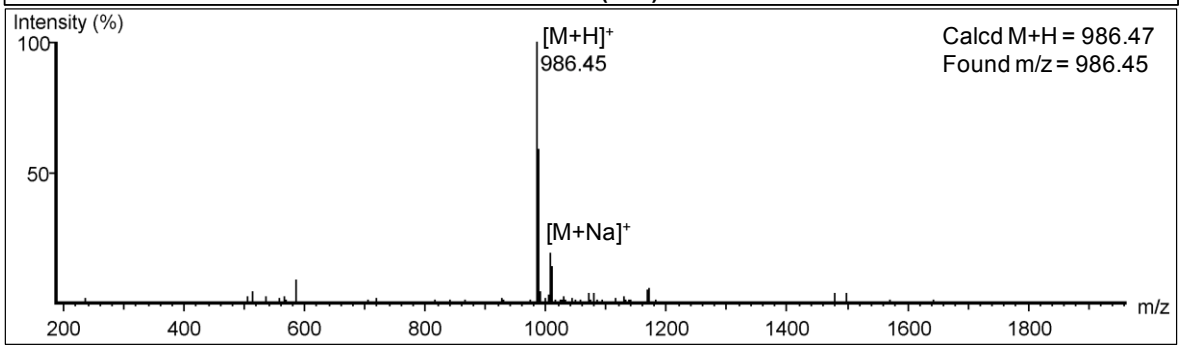


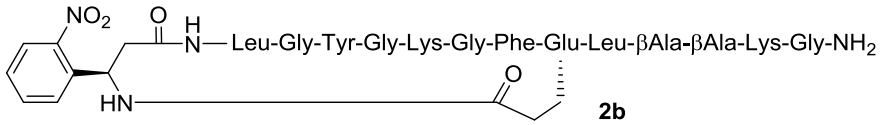
Figure S3. HPLC profiles (λ = 220 nm) and ESI-MS spectra of cyclic peptides **2a-f**.



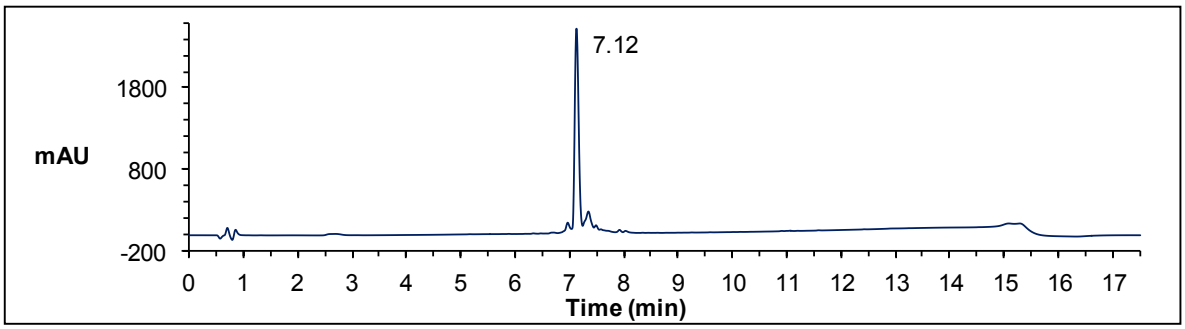
2a







2b



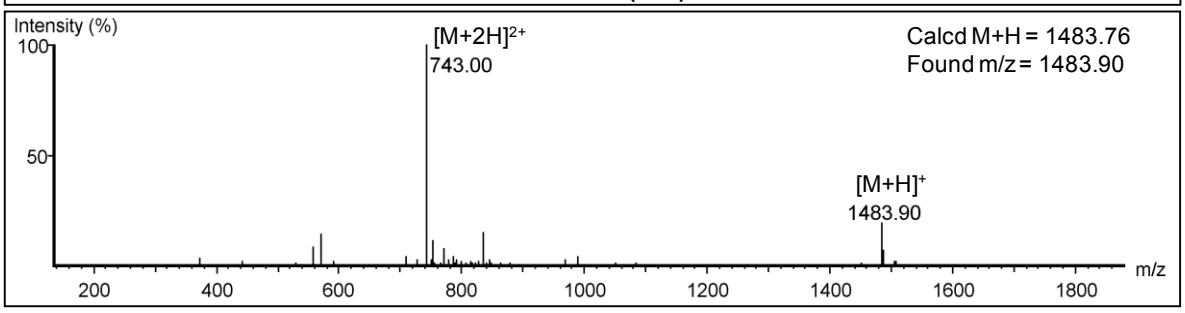
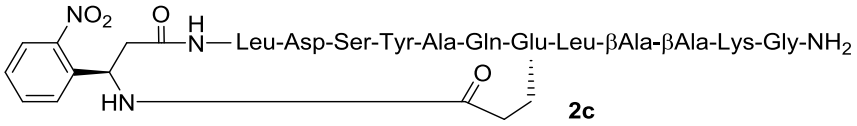
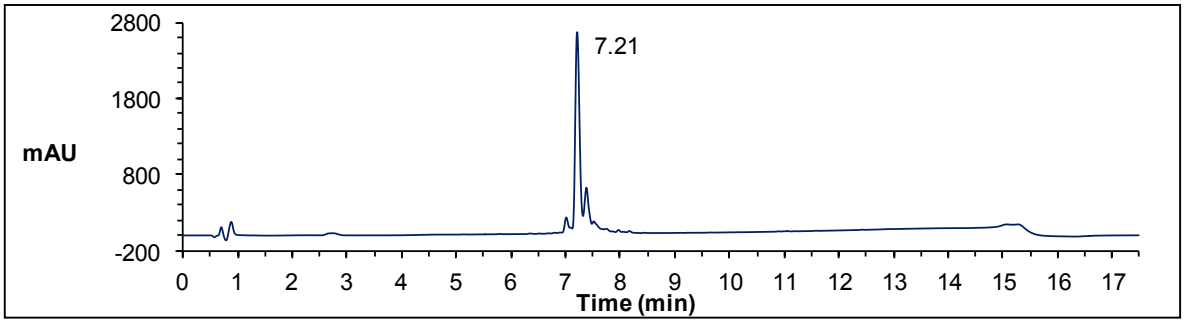
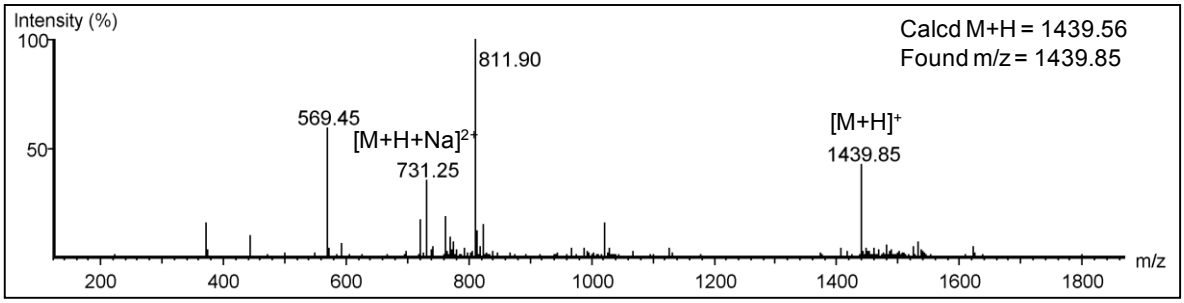


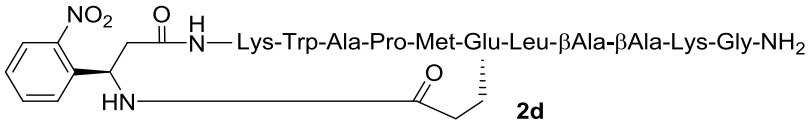
Figure S3. (Continued)



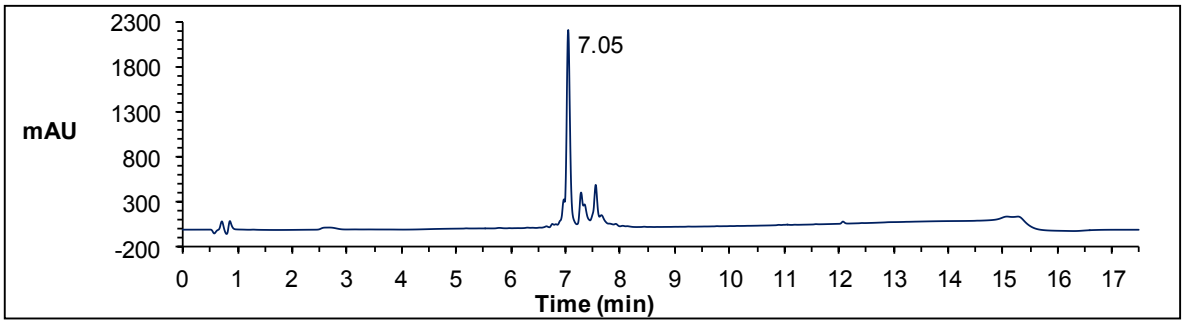
2c







2d



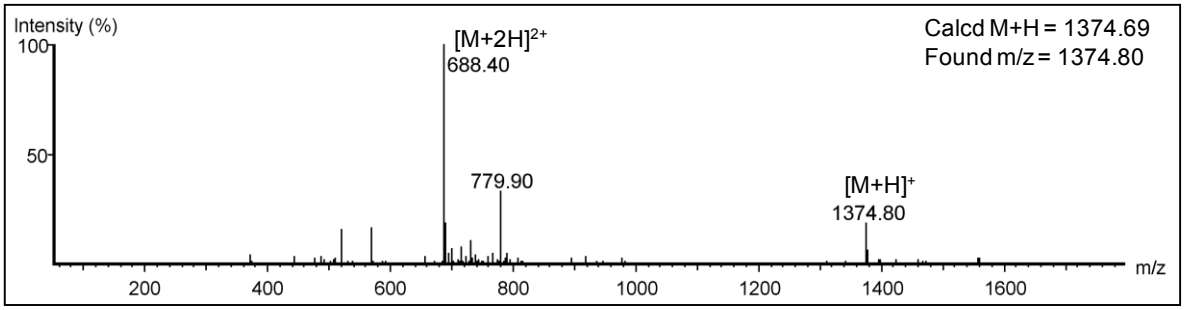
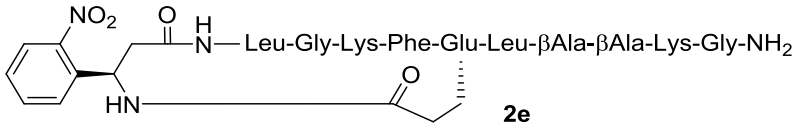
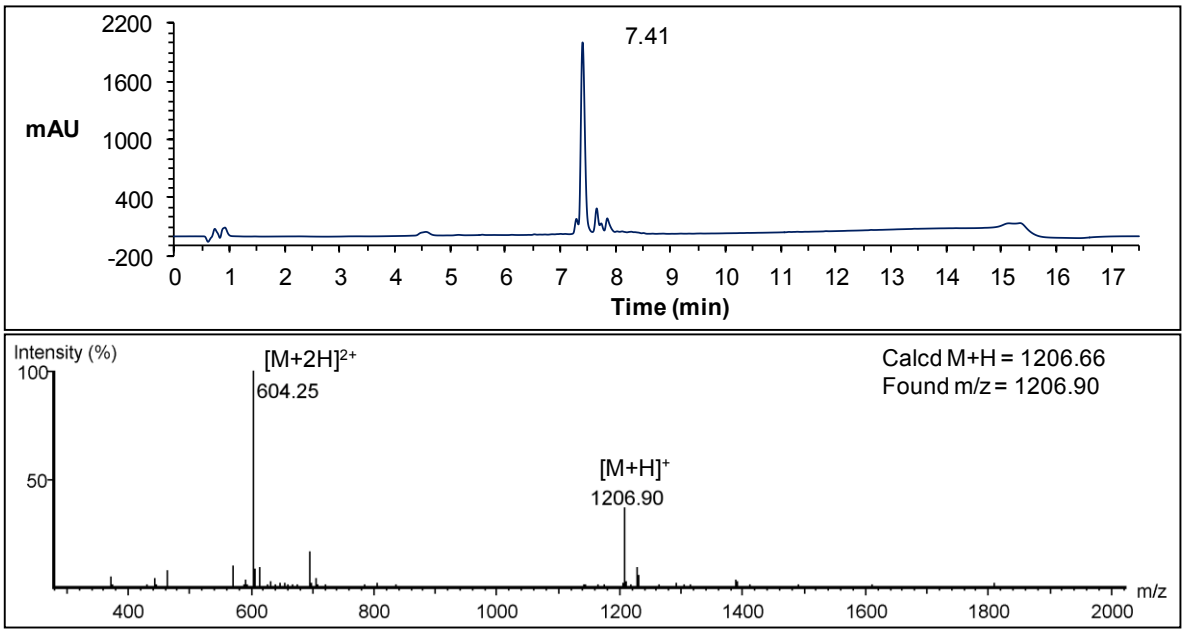
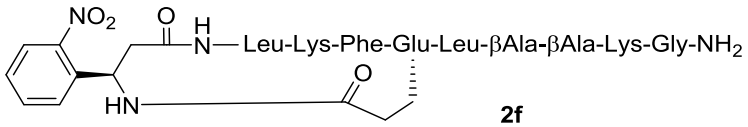


Figure S3. (Continued)



2e





2f

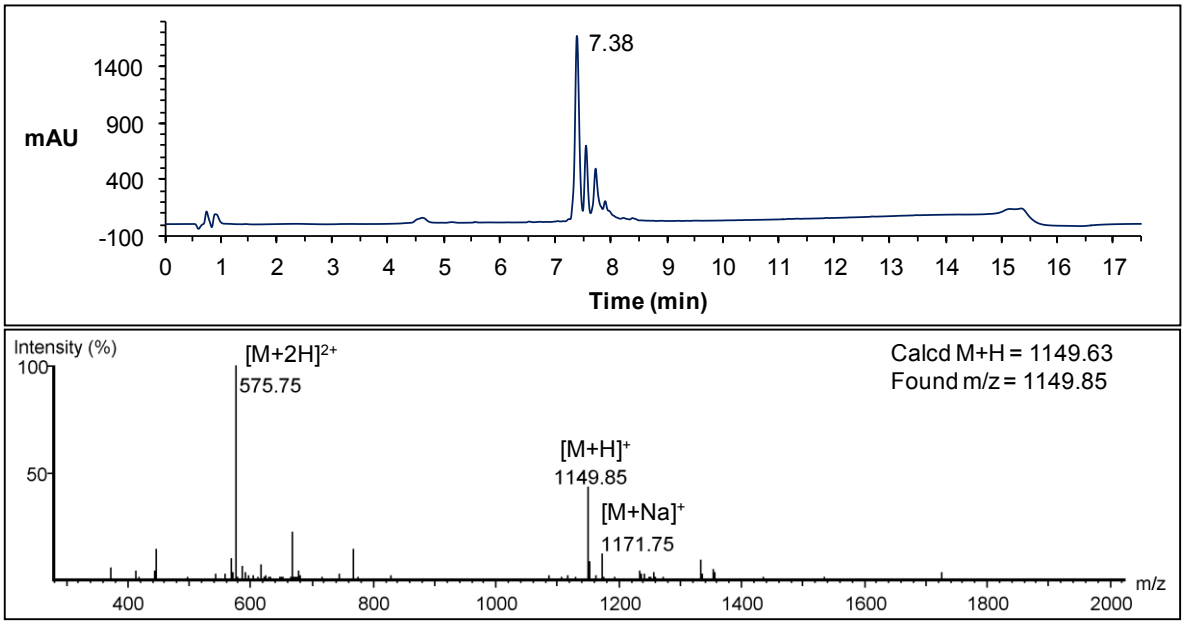
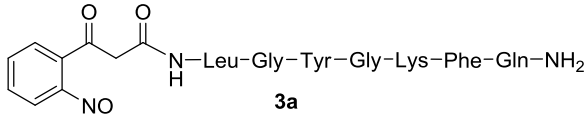
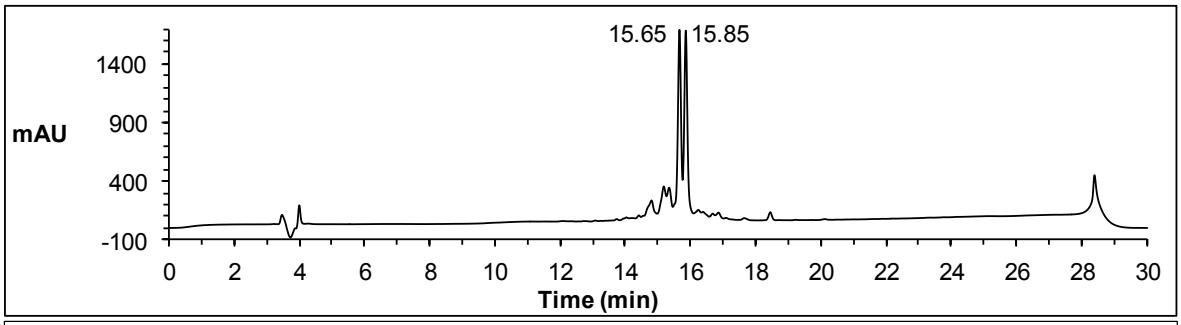
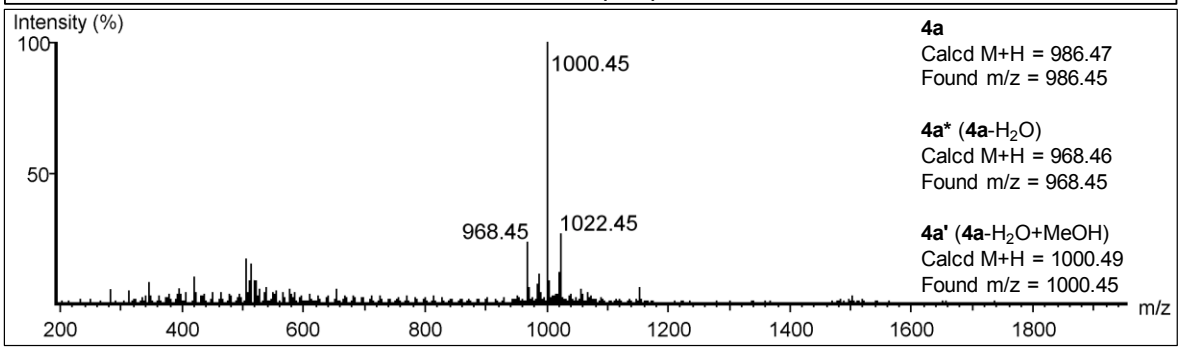


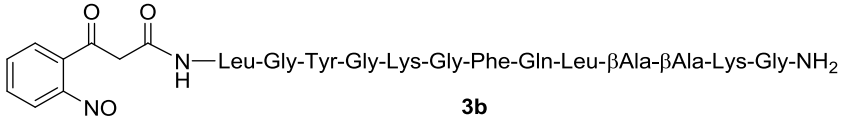
Figure S4. HPLC profiles (λ = 220 nm) and ESI-MS spectra of linear peptides **3a-f** after ring-opening.



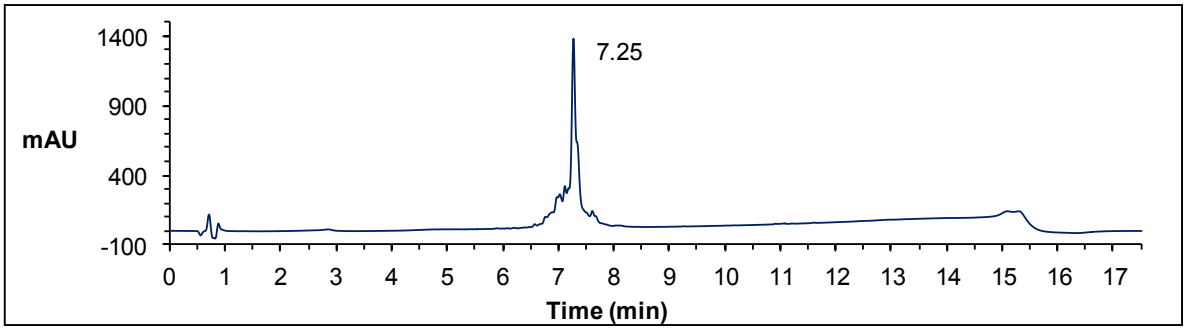
3a







3b



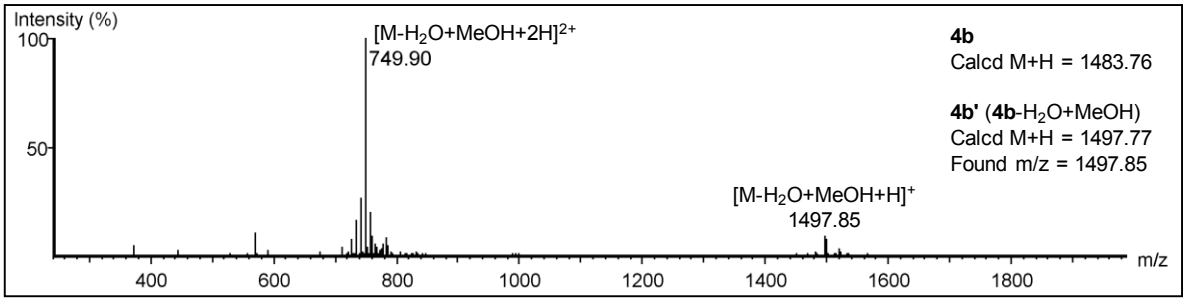
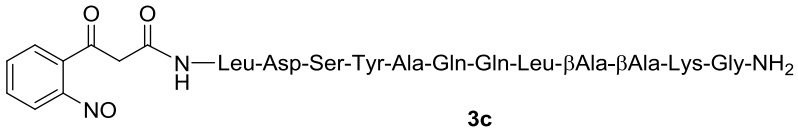
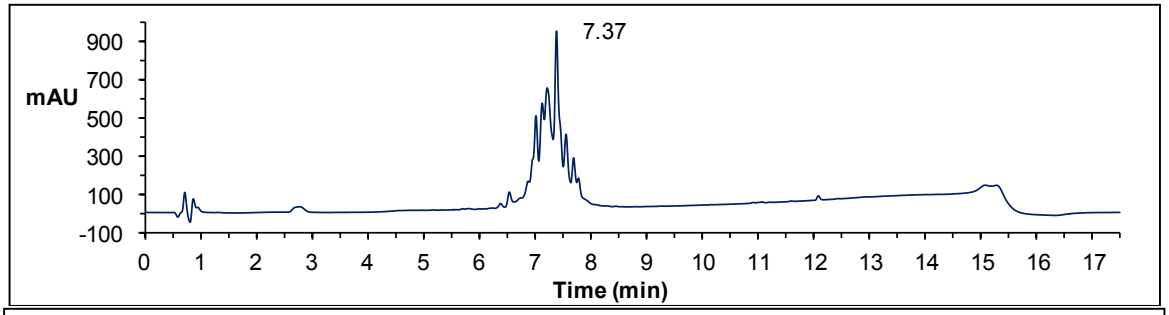
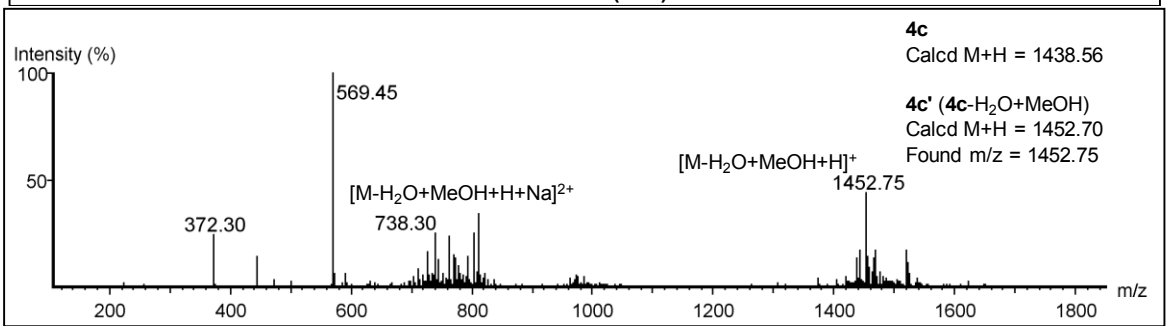


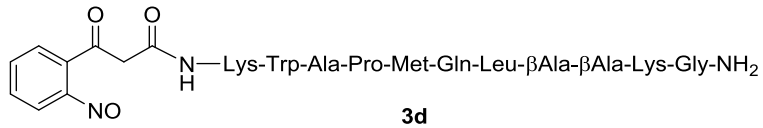
Figure S4. (Continued)



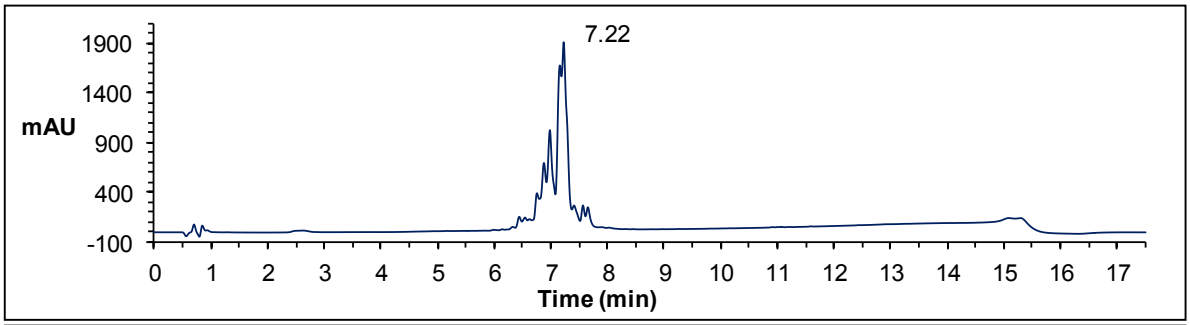
3c







3d



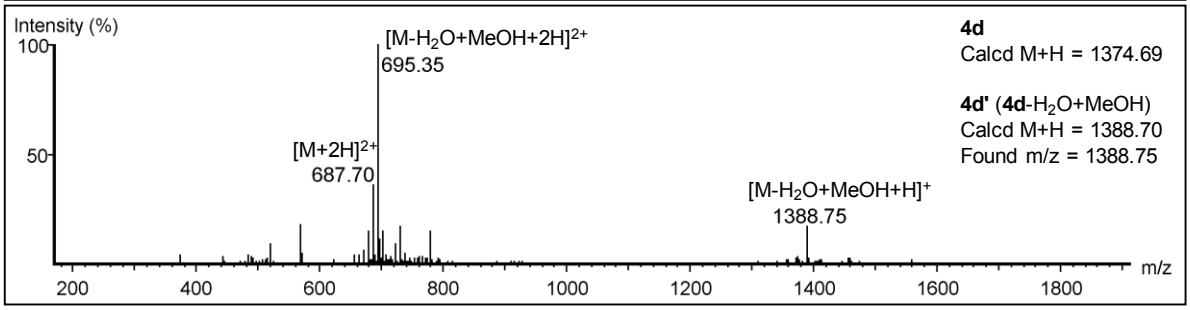
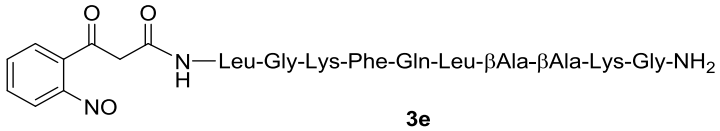
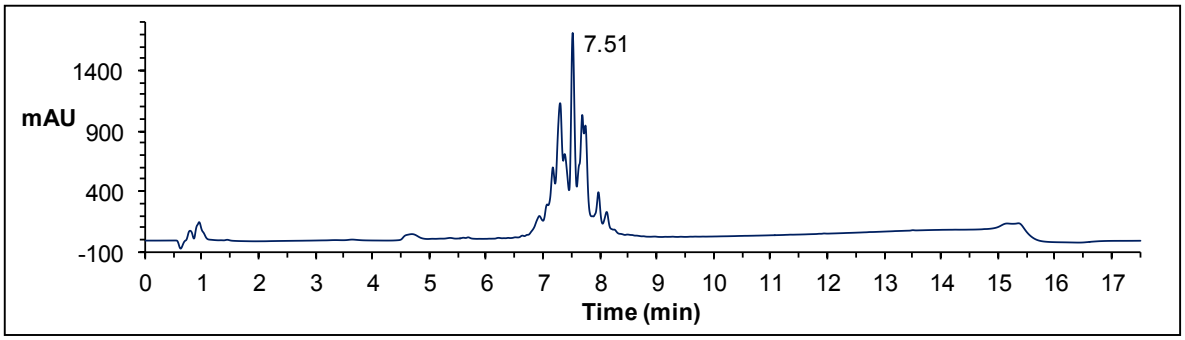
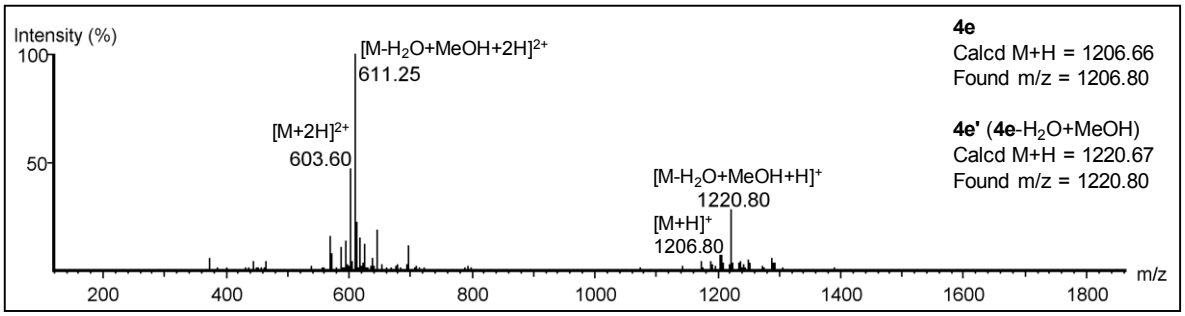


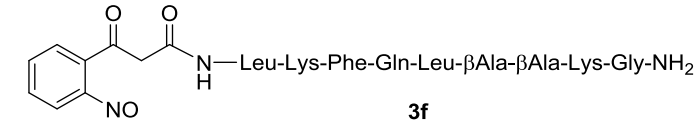
Figure S4. (Continued)



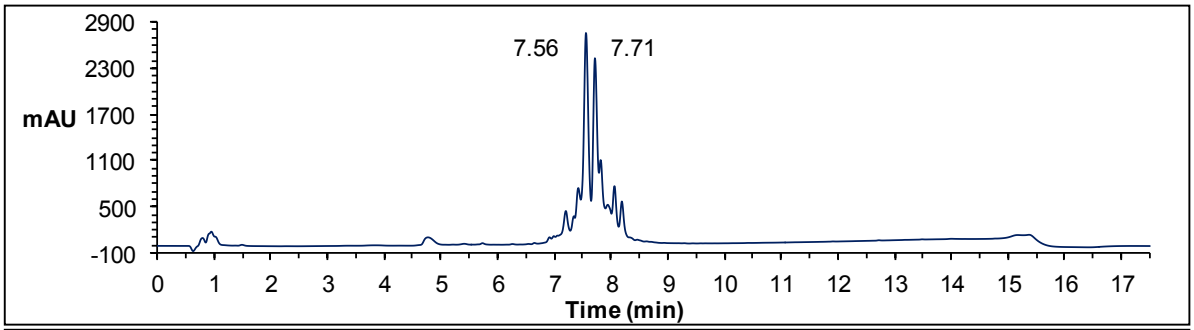
3e







3f



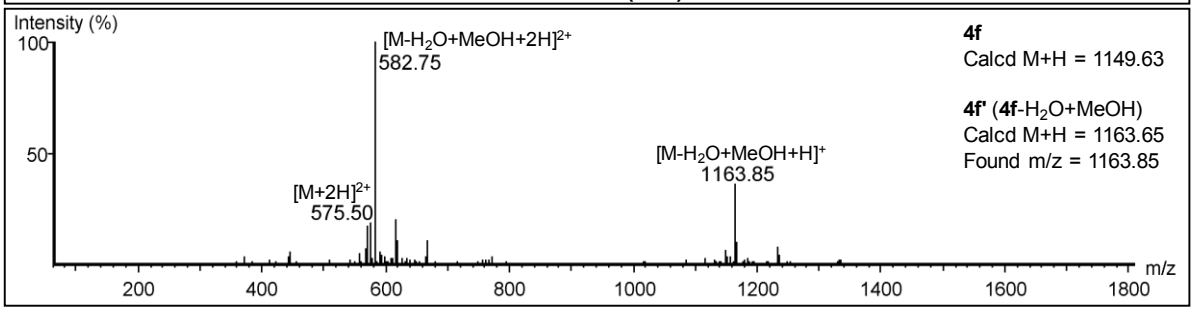


Figure S5. HPLC profiles (λ = 220 nm) and ESI-MS spectra of linear peptides after dual ring-opening/cleavage from TG resin **2a** in presence of **A)** CD_3OD or **B)** *n*-Butylamine.

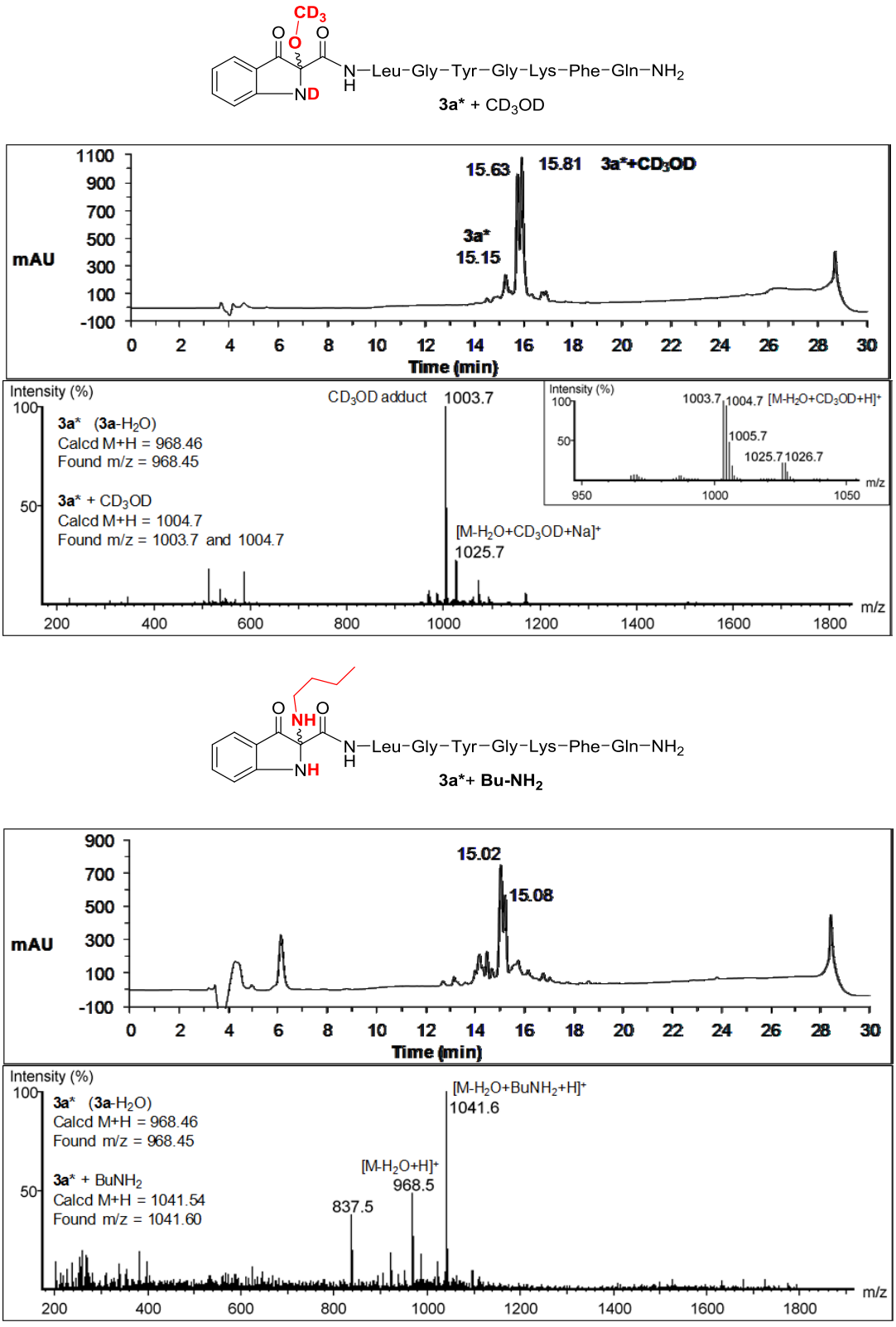
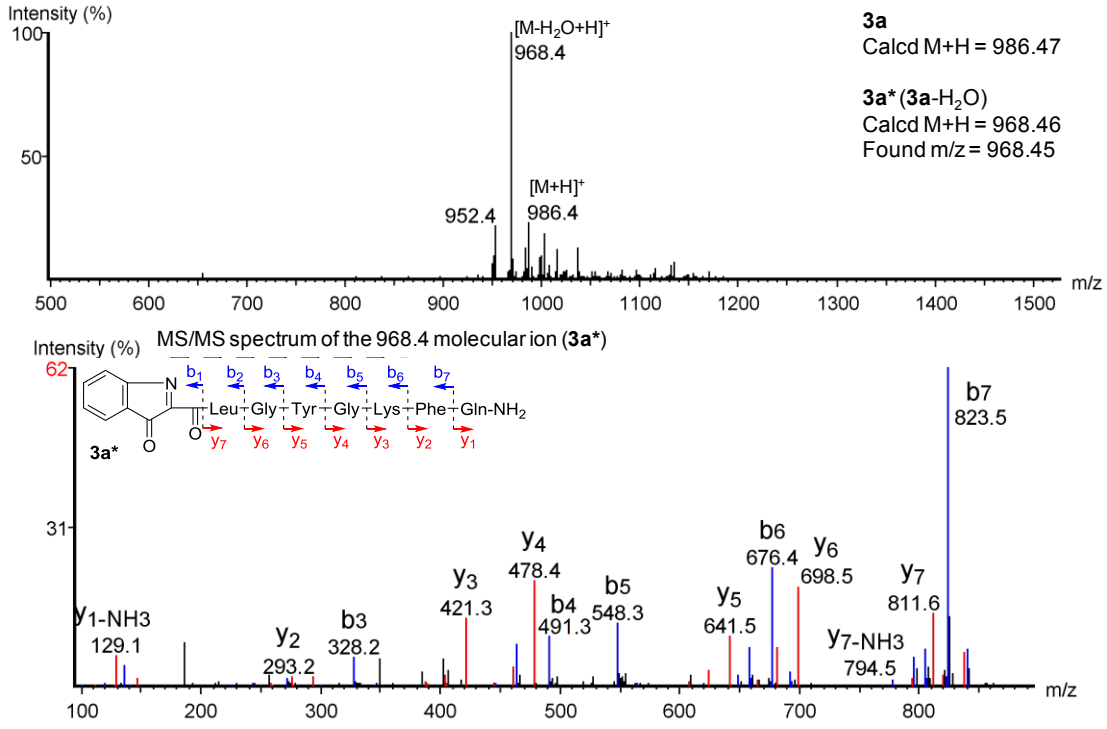


Figure S6. MALDI-TOF MS and MS/MS spectra of peptides obtained from tandem ring-opening/cleavage on a single TG bead.

ANP*-Leu-Gly-Tyr-Gly-Lys-Phe-Gln-NH_2 **3a**



ANP*-Leu-Gly-Tyr-Gly-Lys-Gly-Phe-Gln-Leu-βAla-βAla-Lys-Gly-NH_2 **3b**

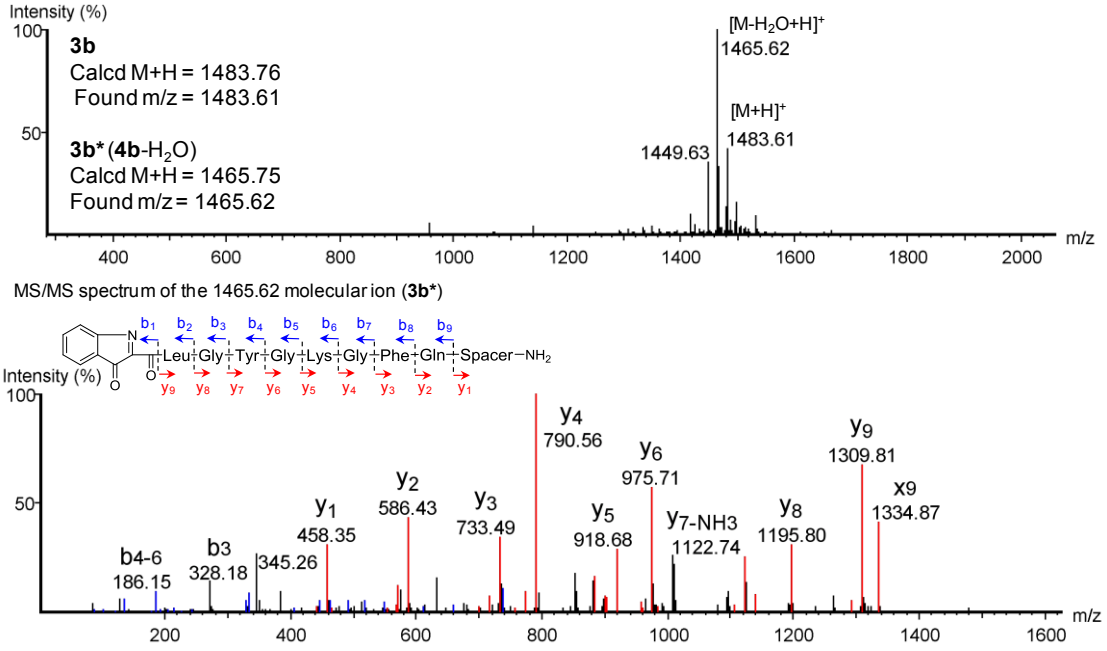
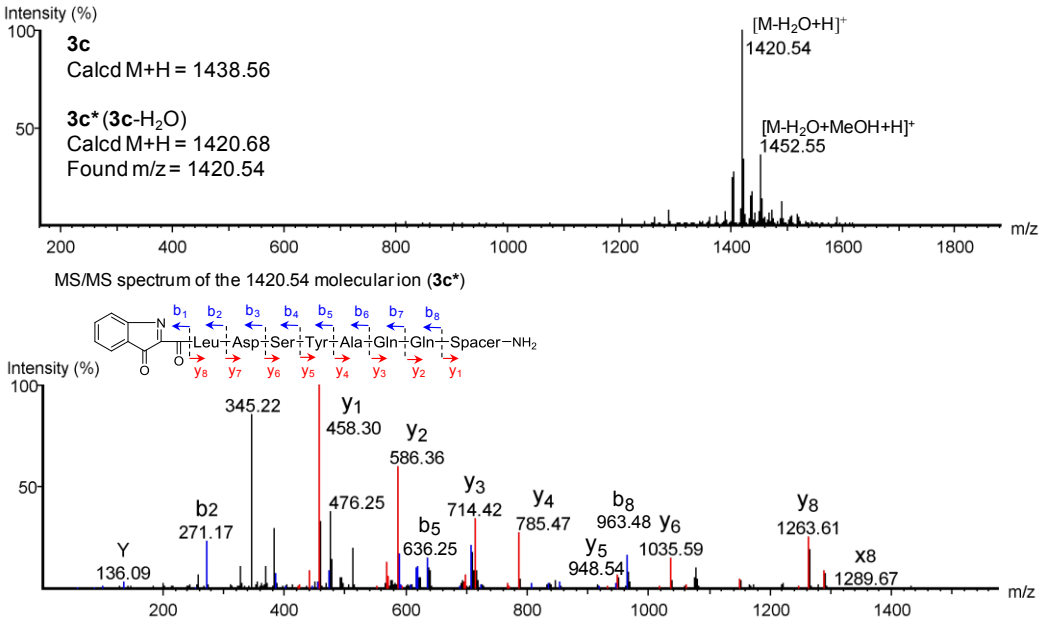


Figure S6. (Continued)

ANP*-Leu-Asp-Ser-Tyr-Ala-Gln- Gln-Leu-βAla-βAla-Lys-Gly-NH_2 3**c**



ANP*-Lys-Trp-Ala-Pro-Met-Gln-Leu-βAla-βAla-Lys-Gly-NH_2 **3d**

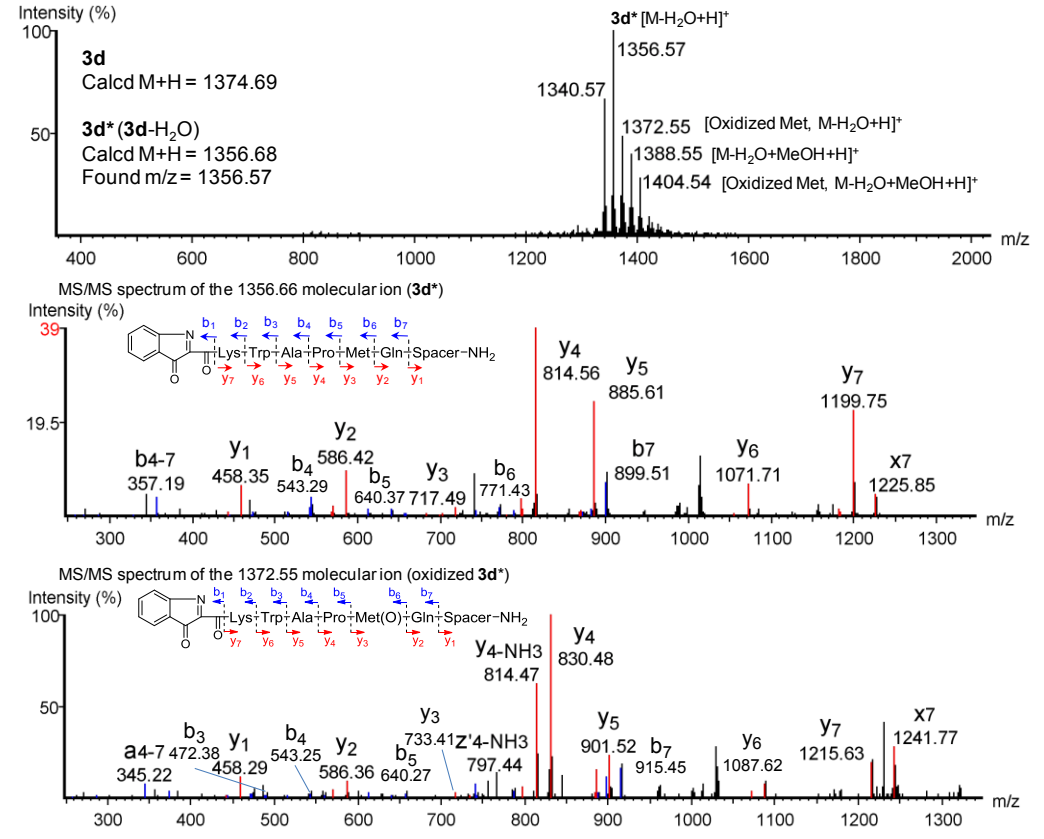
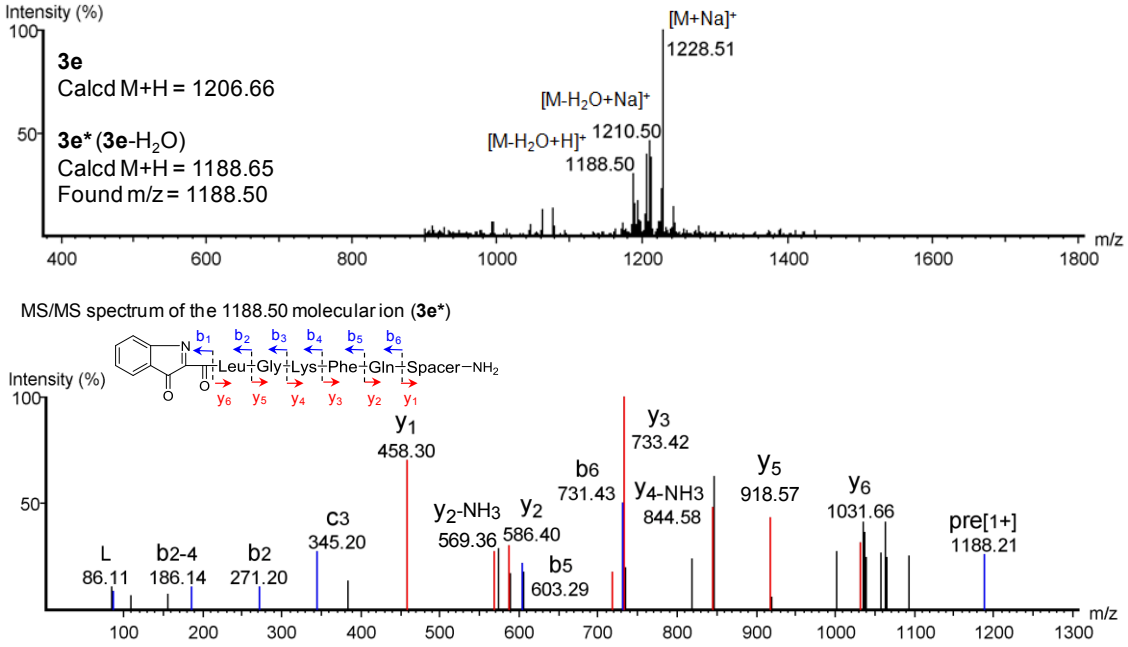


Figure S6. (Continued)

ANP*-Leu-Gly-Lys-Phe-Gln-Leu-βAla-βAla-Lys-Gly-NH_2 **3e**



ANP*-Leu-Lys-Phe-Gln-Leu-βAla-βAla-Lys-Gly-NH_2 **3f**

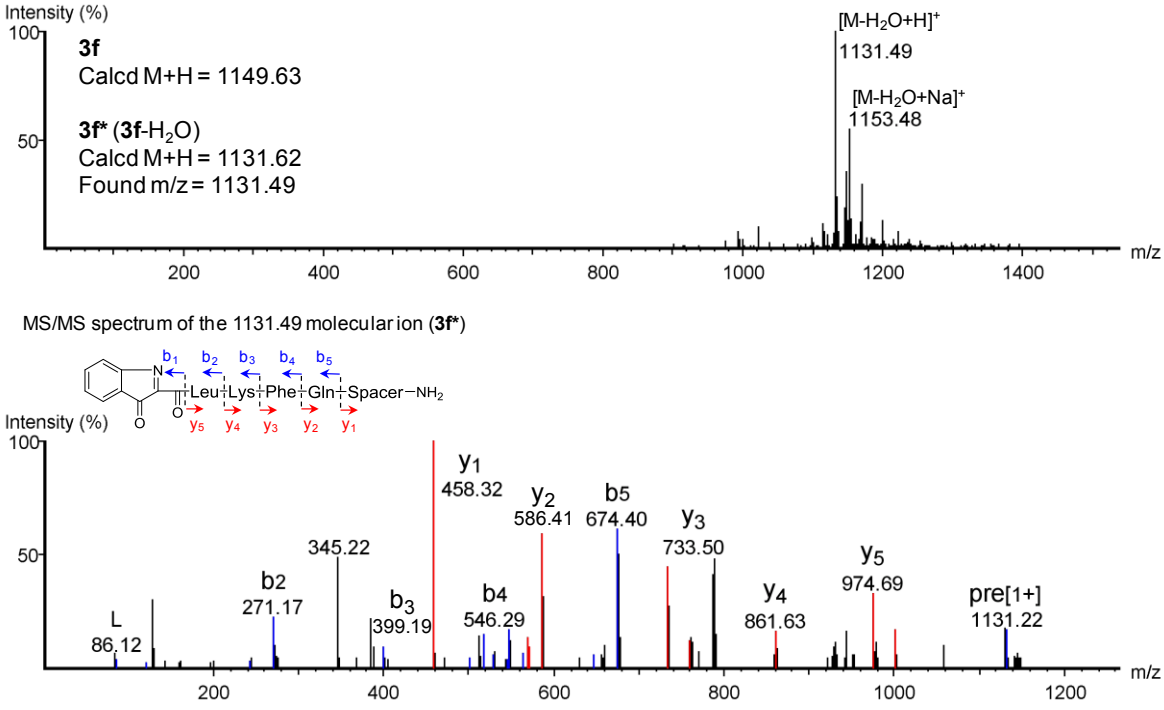
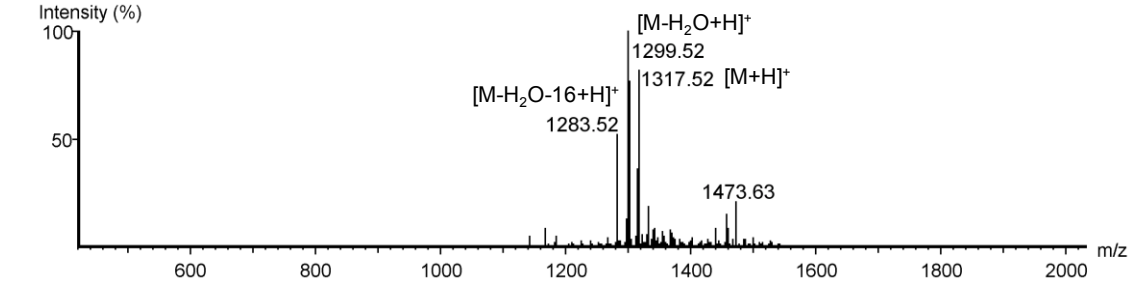
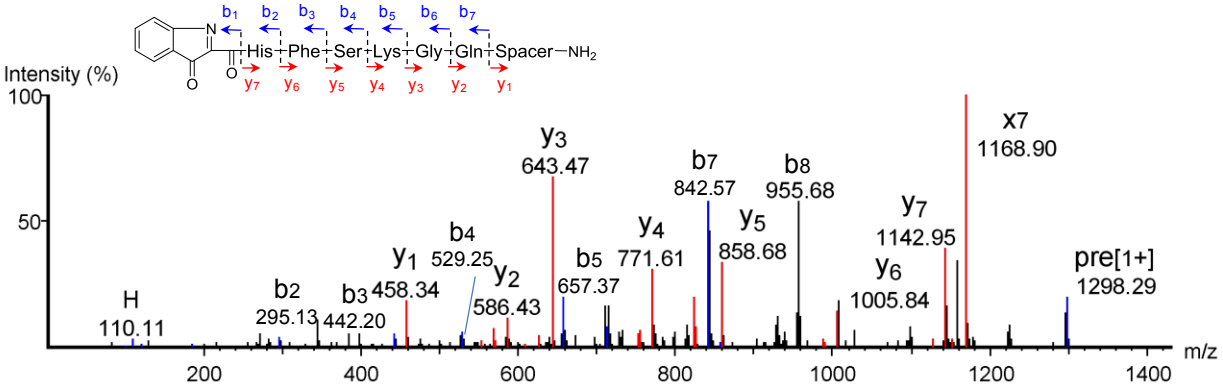


Figure S7. MALDI-TOF MS and MS/MS spectra of the randomly selected beads from the cyclic peptide library after tandem ring-opening/cleavage.

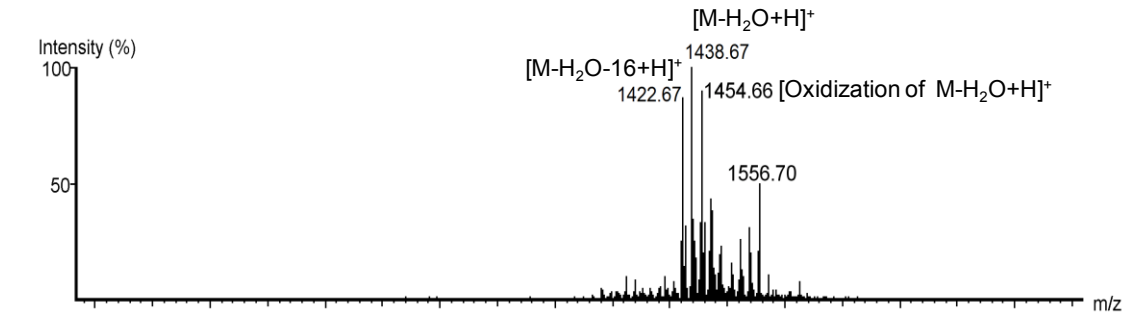
Bead #1



MS/MS spectrum of the 1299.52 molecular ion



Bead #2



MS/MS spectrum of the 1438.67 molecular ion

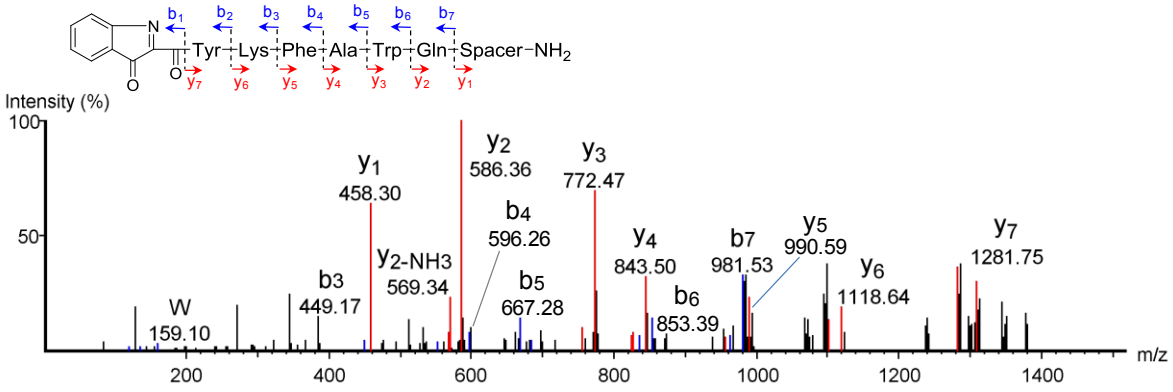
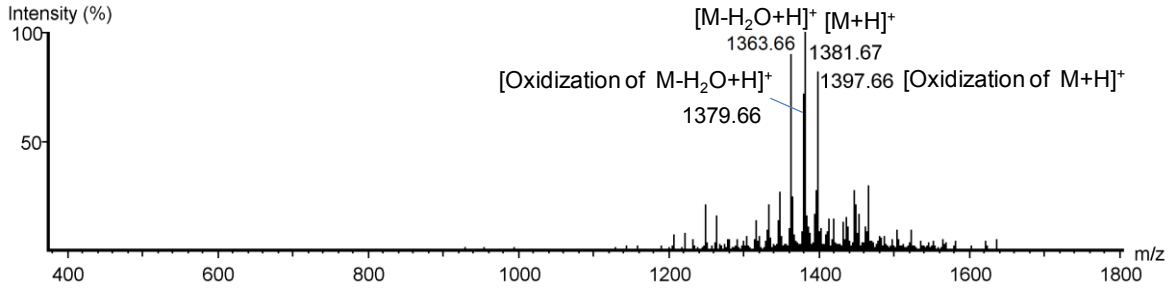
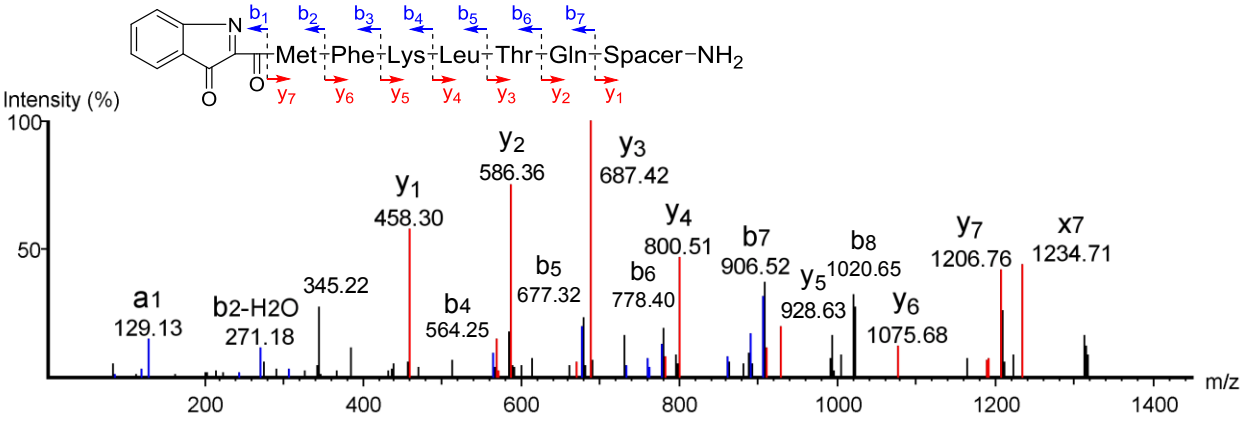


Figure S7. (Continued)

Bead #3



MS/MS spectrum of the 1363.66 molecular ion



MS/MS spectrum of the 1379.66 molecular ion

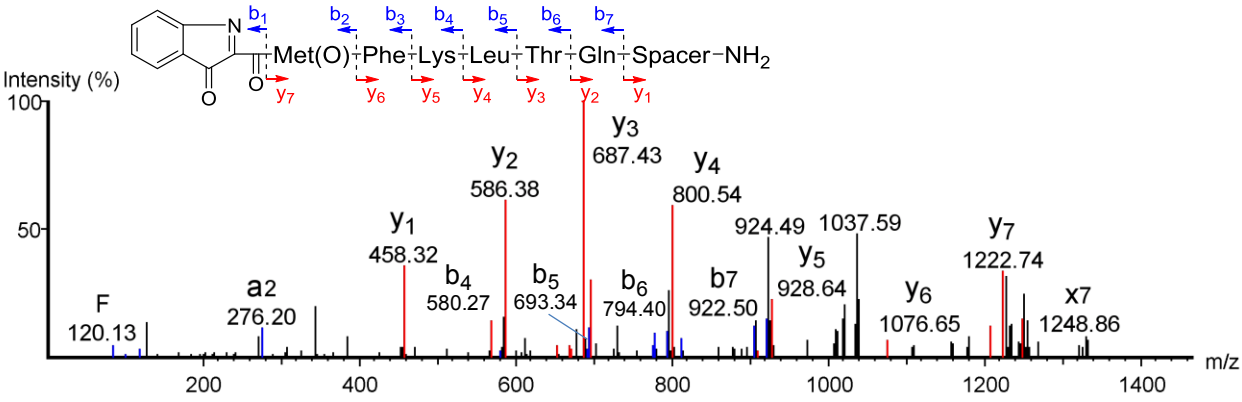
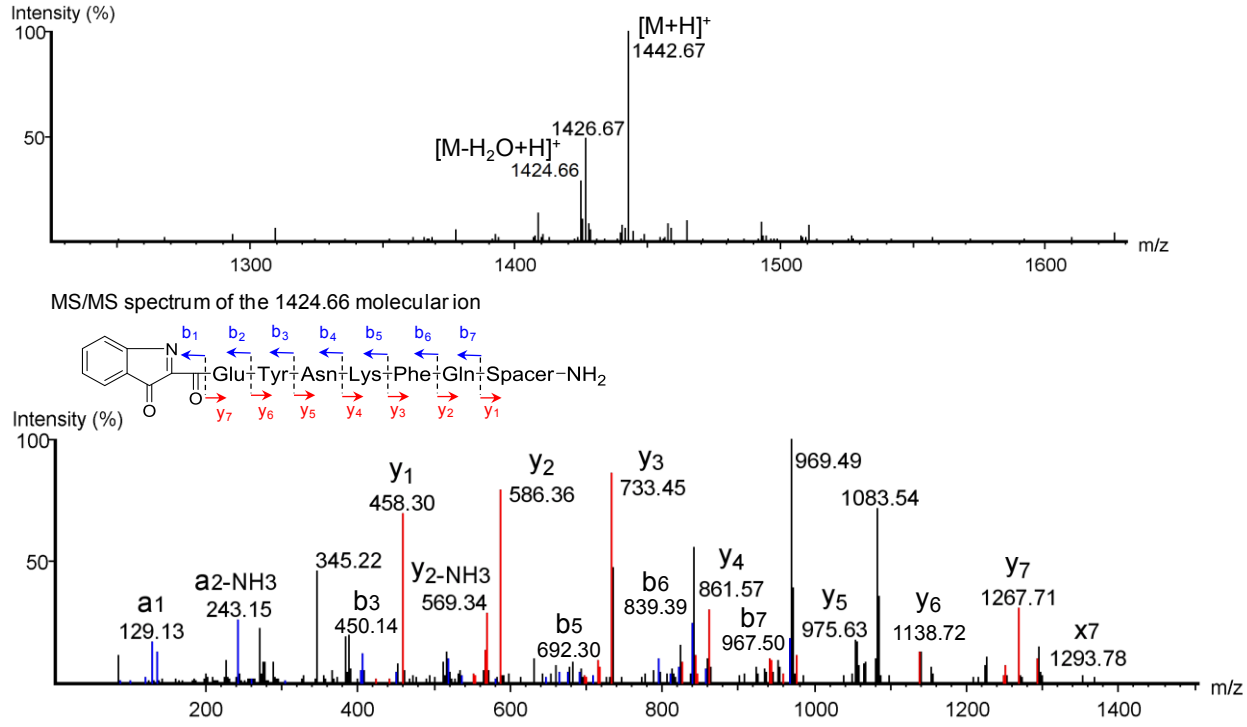


Figure S7. (Continued)

Bead #4



Bead #5

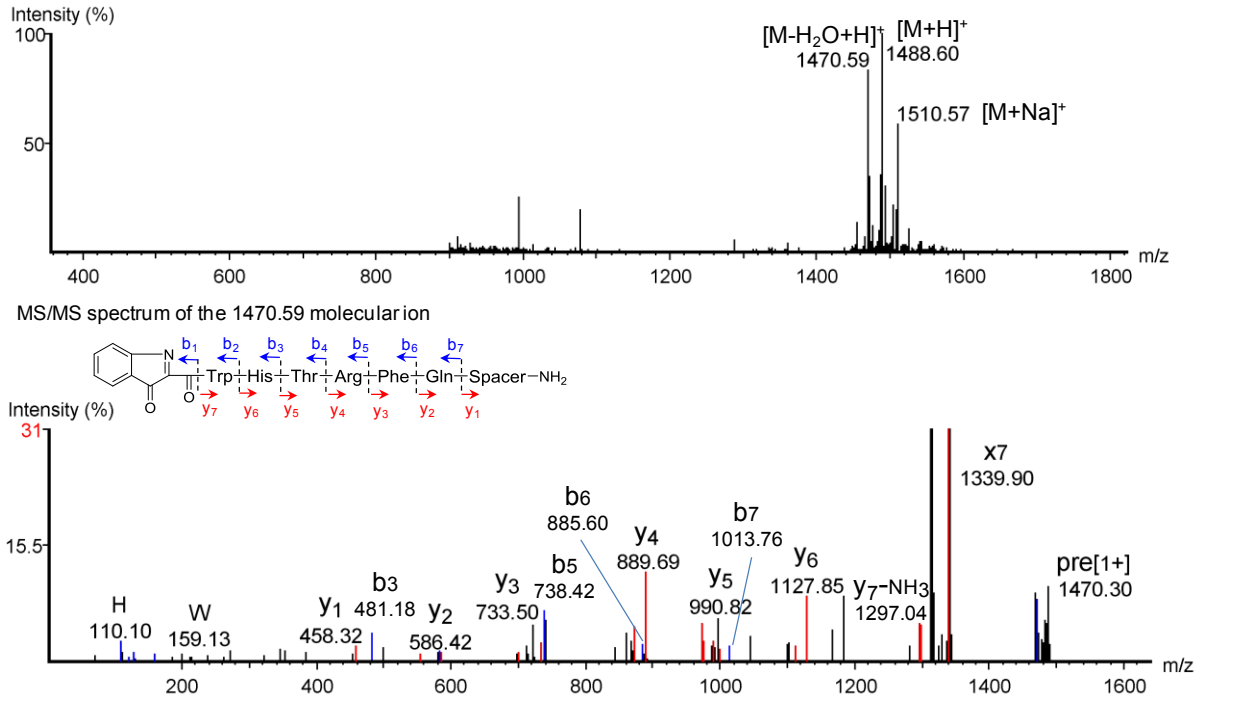
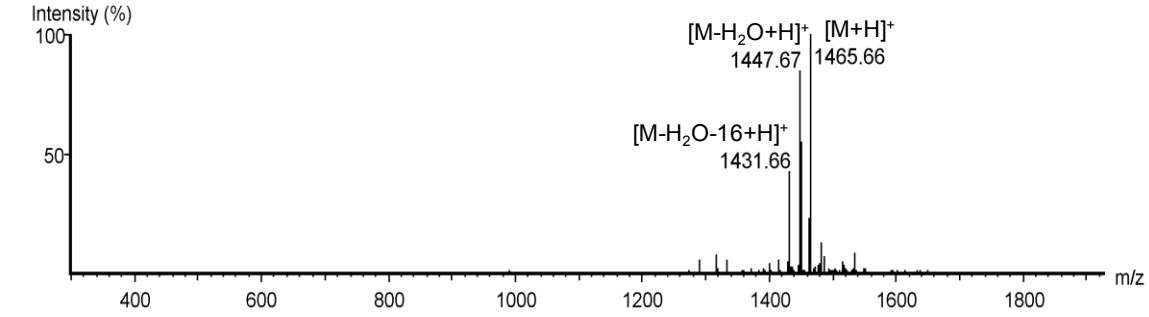
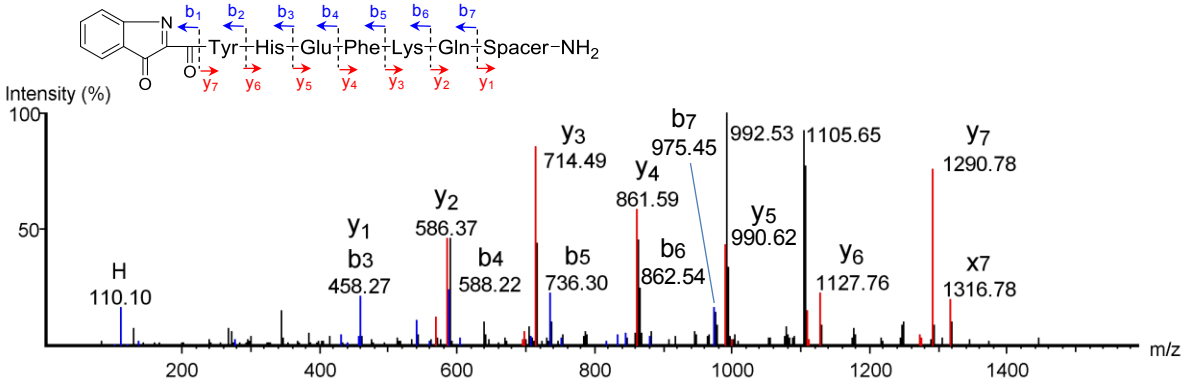


Figure S7. (Continued)

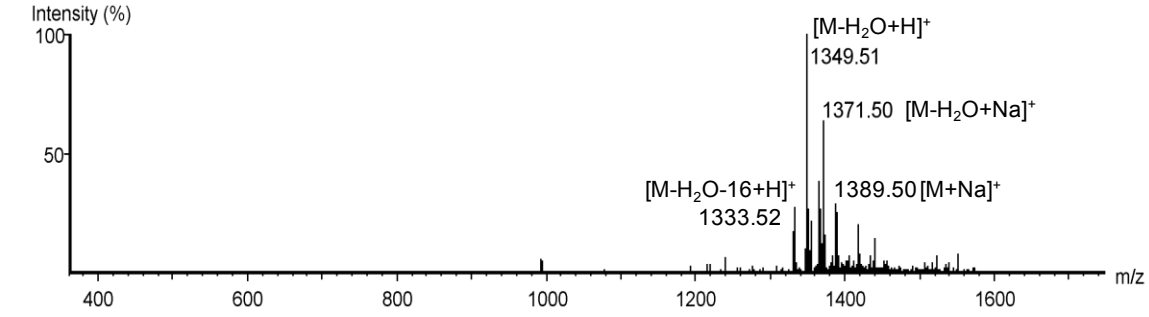
Bead #6



MS/MS spectrum of the 1447.67 molecular ion



Bead #7



MS/MS spectrum of the 1349.51 molecular ion

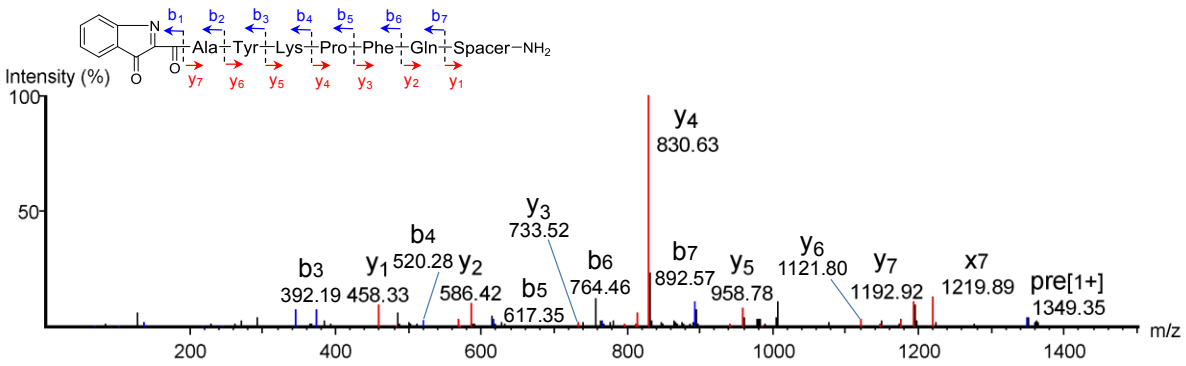
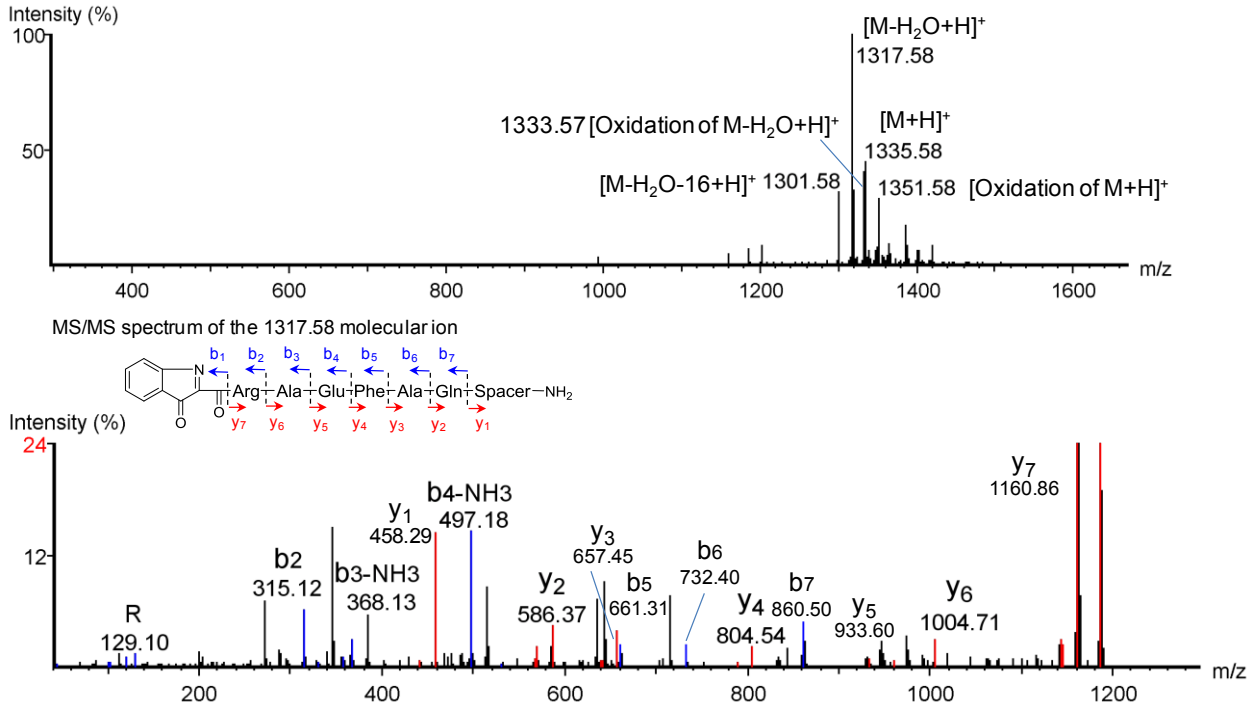


Figure S7. (Continued)

Bead #8



Bead #9

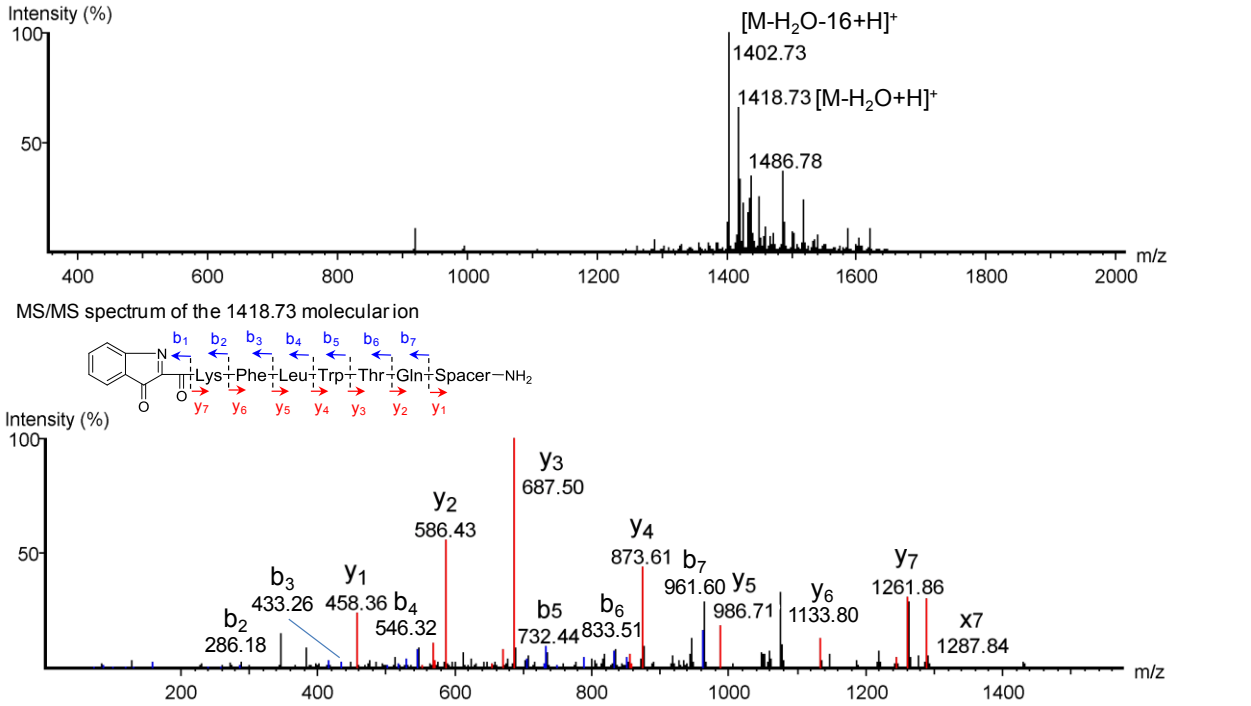
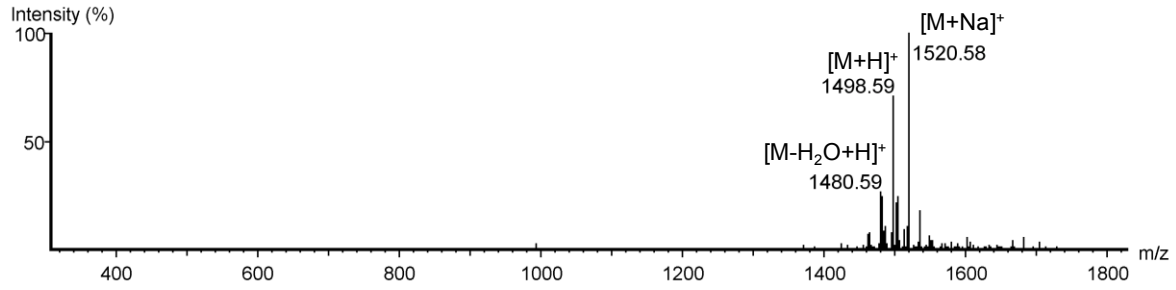
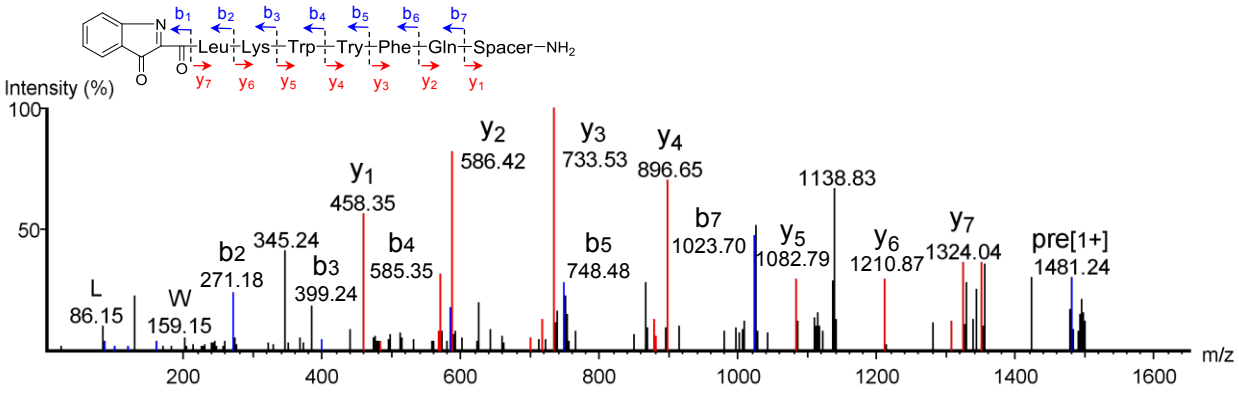


Figure S7. (Continued)

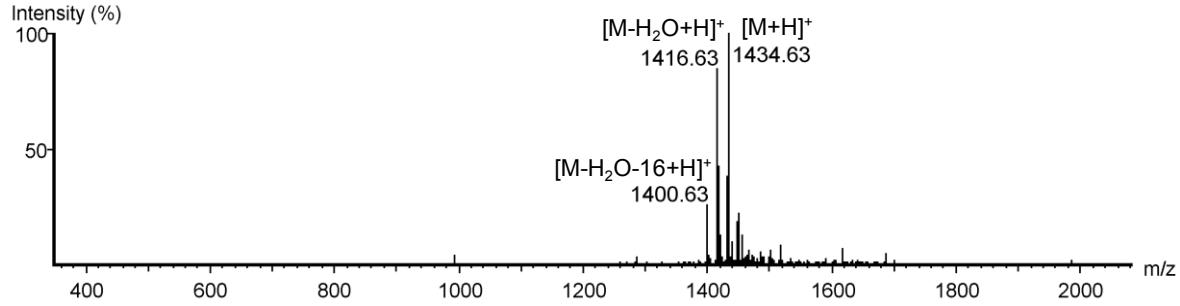
Bead #10



MS/MS spectrum of the 1480.59 molecular ion



Bead #11



MS/MS spectrum of the 1416.63 molecular ion

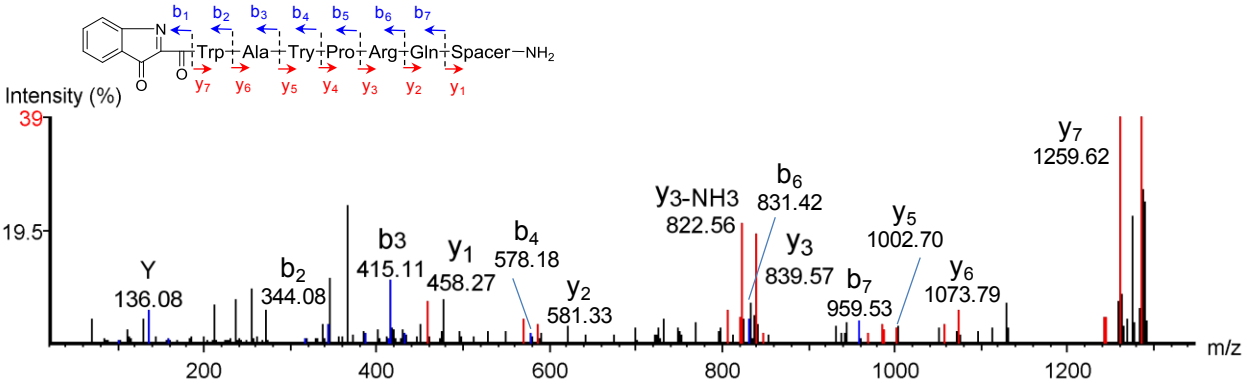
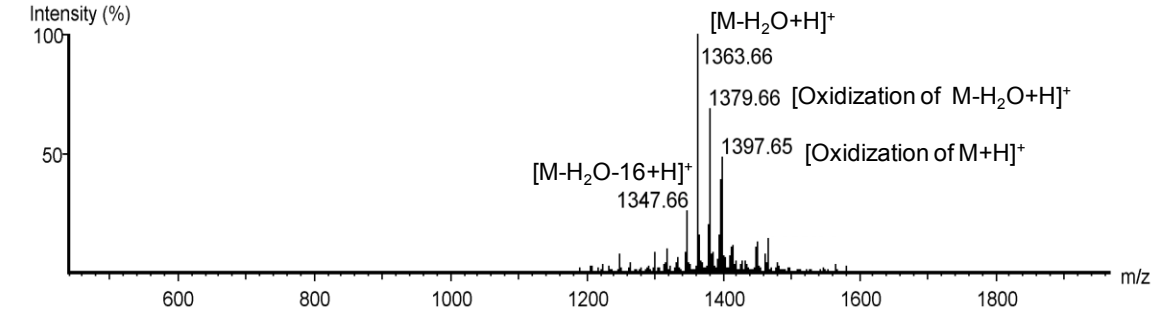
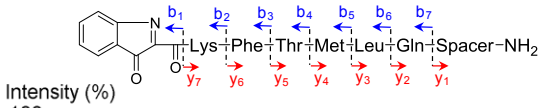


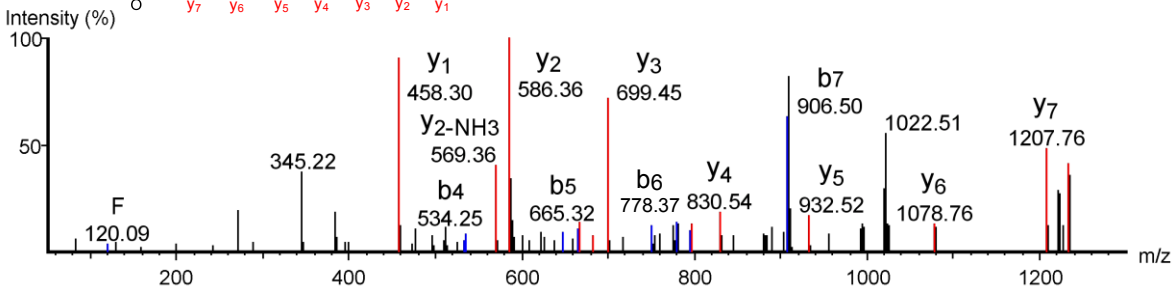
Figure S7. (Continued)

Bead #12

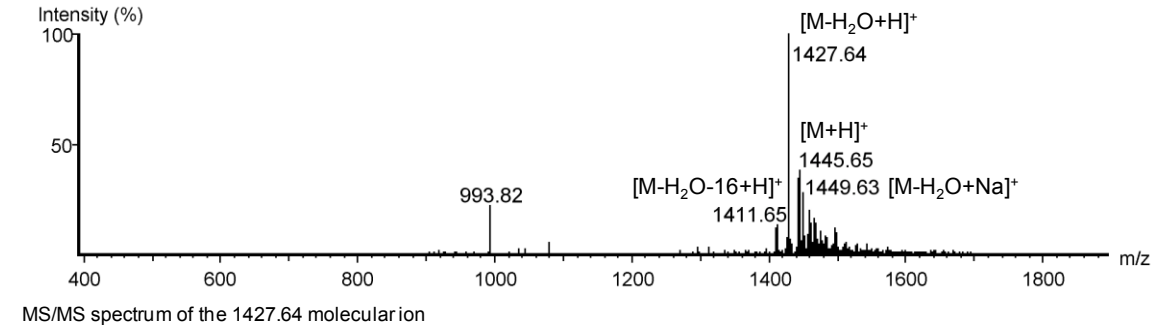


MS/MS spectrum of the 1363.66 molecular ion

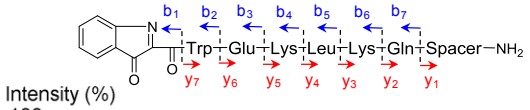




Bead #13



MS/MS spectrum of the 1427.64 molecular ion



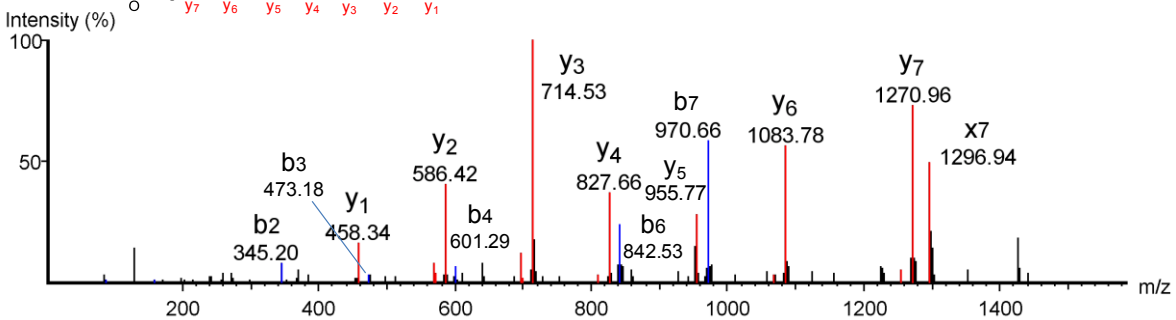
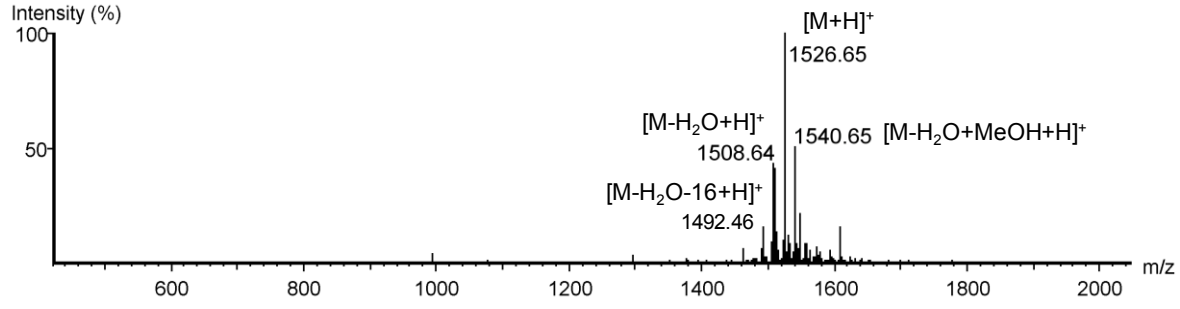
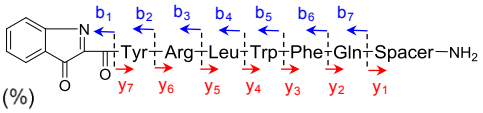


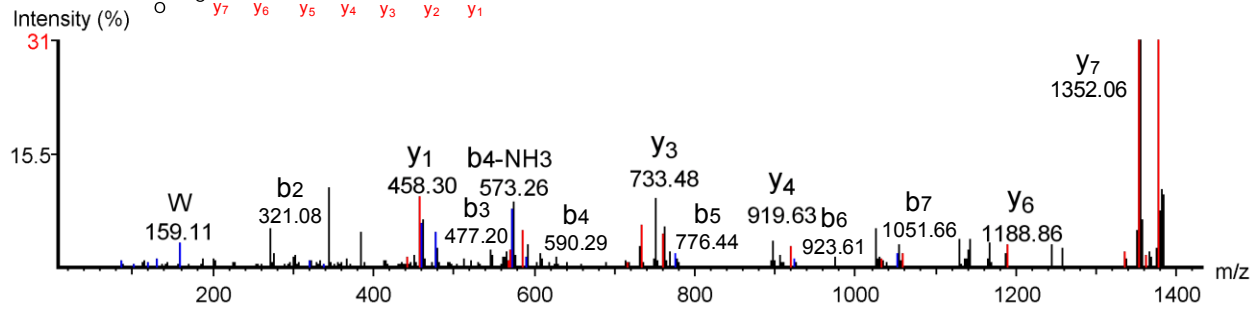
Figure S7. (Continued)

Bead #14

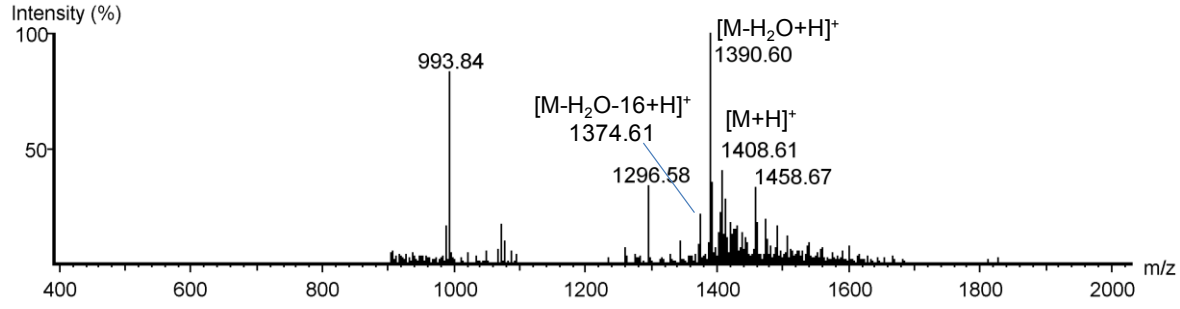


MS/MS spectrum of the 1508.64 molecular ion

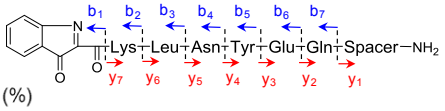


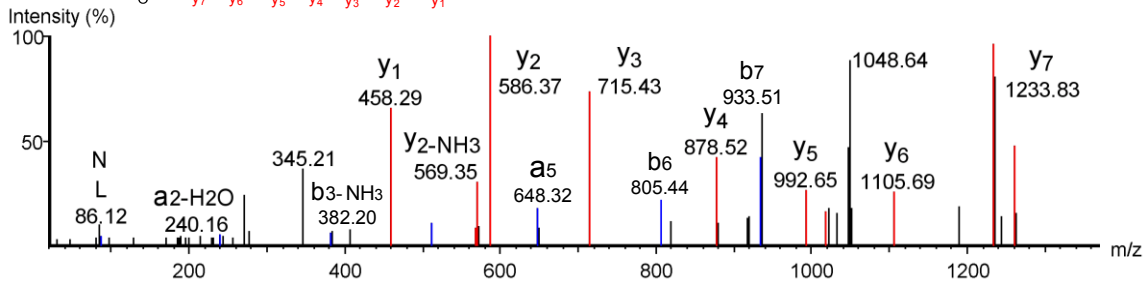


Bead #15



MS/MS spectrum of the 1390.60 molecular ion





Chapter 6

Synthesis and Screening of One-Bead-One-Compound Combinatorial Peptide Libraries for the Development of HIV-1 Nef Protein Inhibitors

6.1 Introduction

Nef (Negative Regulatory Factor), a small myristoylated protein of 200-215 amino acids (27-35 kDa), is an accessory protein of the Human Immunodeficiency Virus (HIV-1 and HIV-2) that promotes virus replication and pathogenesis in the infected host.[166,167] Devoid of enzymatic activity, Nef serves as an adaptor protein to divert host cell proteins to aberrant functions that amplify viral replication.[168] Acting via protein-protein interactions, Nef plays a critical role in HIV pathogenesis (Figure 1).[169]

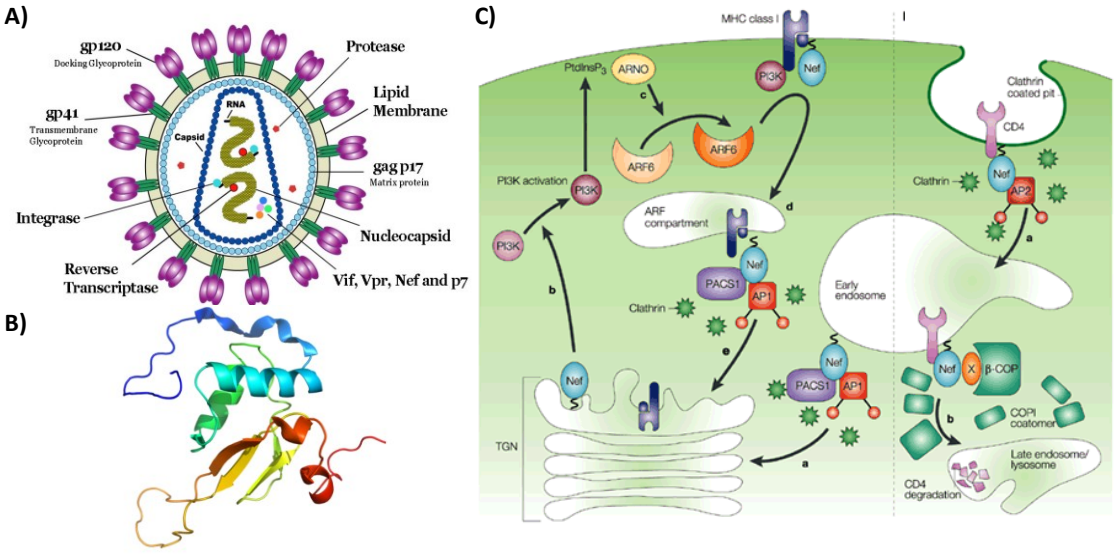


Figure 1. A) Structure of an HIV virion particle; **B)** Structure of the HIV-1 Nef protein; **C)** Sequential steps governing the downmodulation of expression of MHC class I molecules and CD4. (Left) Nef accelerates the endocytosis of MHC I molecules (**a**). Nef activates PI3K (**b**). ARF6 becomes activated (**c**). Together with Nef, ARF6 mediates the internalization of MHC I molecules (**d**). The latter are retrieved to the TGN, where they remain trapped (**e**). (Right) The two steps of Nef-induced CD4 downmodulation. Nef connects the cytoplasmic tail of CD4 triggering rapid endocytosis of the CD4 receptor (**a**). In the early endosome, Nef interacts with the COPI coatomer, which targets CD4 for lysosomal degradation (**b**).[169]

In the early stages of infection, Nef down-regulates CD4 for the fusion of the viral and cellular membranes and facilitate the entry of the virus into the host cell.[170] In order to escape the attack of host cell immune system, Nef must down-regulate major histocompatibility class I (MHC-I) from the cell surface to avoid recognition of cytotoxic T lymphocytes (CTLs).[171] In contrast to CD4, down-regulation of MHC-I happen when the concentration of Nef protein is high. Many signalling proteins are present in glycolipid-enriched microdomains where Nef is also found. The function of mediating cellular signalling by Nef enhances escape of the immune system and rises viral particle releasing.[172]

Involved in many cellular processes that are essential for virus pathogenesis and viral fitness, the HIV-1 Nef protein is a very promising therapeutic target to develop innovative anti-HIV agents. Despite this great potential, no effective ligand has been found to date. In order to identify ligands and inhibitors for Nef, we planned to prepare and screen a one-bead-one-compound (OBOC) peptide library. As described in the first chapter, the OBOC approach is one of the most powerful and convenient combinatorial methodologies to generated and screen large libraries.[56,57,59,102,173] This chapter describes the preparation and screening of a OBOC peptide library and the evaluation of hit compounds binding affinities for Nef.

6.2 Results

6.2.1 Preparation of OBOC library

In this work, an octapeptide linear library containing eight random residues bearing a spacer (LBBRM) was prepared on 400 mg of TentaGel® S NH_2 resin (~ 312,000 beads, 0.35 nmol peptides /bead) with a diameter of 130 μm (Figure 2). The TentaGel resin was chosen because of its excellent chemical and physical properties with respect to swelling and compatibility with organic and aqueous solvents. The LBBRM was used to provide a flexible spacer to facilitate protein binding and increase the molecular weight to facilitate MALDI TOF MS/MS analysis. Moreover, the Arg residue provides a fixed positive charge which boosts peptide ionization, reduces spectral noises and helps analyses of MS and

MS/MS spectra. Methionine was used as a linker and coupled as the first amino acid before the spacer. Each position was filled by a random combination of seventeen L-amino-acids (Met, Cys, and Ile were excluded) using the split-and-pool synthesis method. The theoretical diversity of the library is 17^8 (~6 975 757 441 different compounds). Finally, the peptide library was *N*-capped by N-terminal acetylation and the side chains deprotected with a TFA cocktail (Figure 2).

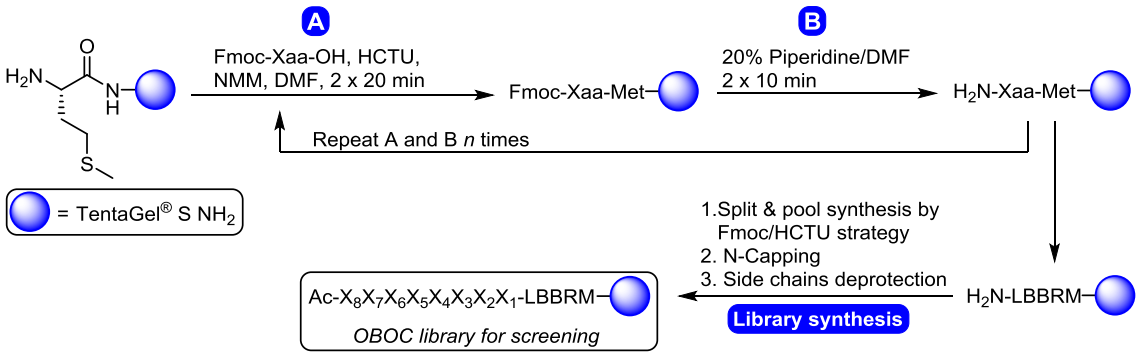


Figure 2. Solid-phase synthesis of the OBOC peptide library by the split & pool method.

6.2.2 OBOC library screening and sequence determination

In order to perform the screening, the Nef protein was first labeled with a biotin using a commercially available biotinylation kit. Then, the labelled proteins were subjected to library screening. For large libraries (>100 mg of resins), the identification and isolation of positive beads can be time-consuming and laborious by naked eye with the classical method. To facilitate and accelerate the screening process, the on-bead screening was performed by a two-step orthogonal screening involving magnetic bead sorting assay[125] and enzyme-linked colorimetric assay.[174] First, 100 mg of the peptide library was incubated with 500 nM biotinylated Nef for 6 h. Then streptavidin-conjugated magnetic beads (Dynabeads, ~2 µm) were added. After a short incubation to allow an efficient streptavidin-biotin interaction, a magnet was used to separate hits from the rest of the library beads (Figure 3A). The magnetic separation was performed several times to reduce false positive beads and a few thousands beads were selected. In this step, the population of peptide library beads was reduced largely. The resulting beads were subjected to a second screening round by incubating the beads with the streptavidin-alkaline phosphatase (SA-AP)

conjugate. Upon subsequent addition of the SA-AP substrate 5-bromo-4-chloro-3-indolyl phosphate (BCIP), the alkaline phosphatase hydrolyses the phosphate group of BCIP to generate the indole derivative, which is oxidized by oxygen to form an insoluble blue dye (Figure 3B). Approximately 100 colored beads were isolated under a microscope with a micropipette.

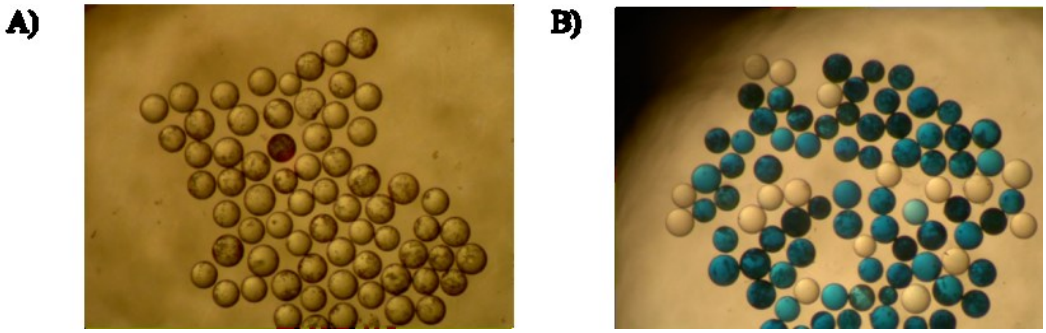
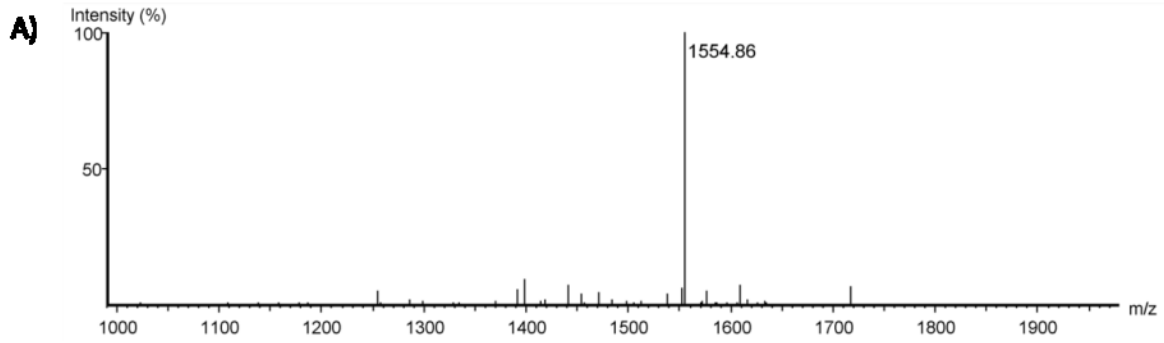


Figure 3. A) First screening round with Dynabeads® M-280 Streptavidin; B) Second screening round with streptavidin-conjugated alkaline phosphatase and BCIP.

After the screening rounds, the isolated beads were washed with 1% SDS solution at 95 °C for 5 min to remove any bound material. Finally, the peptides were individually released from the beads by treatment with a CNBr solution and analyzed by MALDI-TOF/TOF MS (example of MS and MS/MS spectra for bead #8, Figure 4). Their sequences were determined by using the *de novo* sequencing program in the Peaks software and 64 peptides were successfully sequenced (Table 1). The identified peptides were individually resynthesized and purified to be evaluated for their binding capacity to Nef by Nanoporous Optical Interferometry (NPOI).



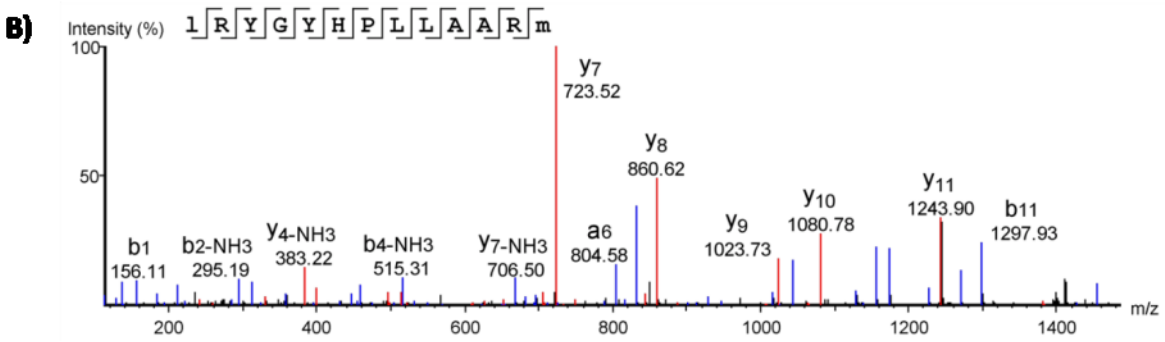


Figure 4. A) MALDI-TOF MS and B) MS/MS spectrum of selected peptide NO.8

Table 1. Sequenced peptides from positive beads

	Peptide		Peptide		Peptide
1	YRSEK/QFFH	**23**	NFWK/QAFYS	**45**	WSVDRVK/QS
2	FPSYVRHT	**24**	K/QYYK/QLPK/QL	**46**	AK/QTLFPWK/Q
3	FFK/QYTFFG	**25**	RLFPHTWL	**47**	TWAWDRFR
4	HAHSLRWH	**26**	RGTWAAHV	**48**	RNFWAVSD
5	NK/QK/QLWYVY	**27**	GSGYRFYE	**49**	LVFPSRFN
6	WSTTYNNR	**28**	YWK/QGPAAK/Q	**50**	TYLFFHTR
7	RNGWTYRS	**29**	DFRSVYK/QR	**51**	WAHNK/QYHN
8	LRYGYHPL	**30**	DHLSRWK/QH	**52**	RFLVHFK/QV
9	RDYK/QPPWW	**31**	WSPFK/QRRF	**53**	HK/QFAFRVK/Q
10	RGFFYPTL	**32**	WHANLYHY	**54**	RFRWFK/QYP
11	RYHPHVAY	**33**	NK/QRYEPYR	**55**	YAAK/QK/QFEW
12	LRATANLR	**34**	FK/QLHSLAR	**56**	RLYK/QAAK/QD
13	FK/QWHPTPA	**35**	LRYHSHK/QL	**57**	NGGRFK/QVY
14	FGVRYTK/QY	**36**	WK/QNSYRLK	**58**	RVFWLDSS
15	GYTYVRHK/Q	**37**	NVERNVWL	**59**	WRTWHVEG
16	RLAK/QHYWA	**38**	LNRTSFLG	**60**	SLK/QGLHAR
17	VHWFTYK/QE	**39**	DK/QYHKDHG	**61**	SRDWDRFR
18	LAHTTRYH	**40**	P FWWGGSF	**62**	RGERNVWL
19	HSGARFYR	**41**	RALEVWND	**63**	VNGVK/QHPS
20	RLHGFAHV	**42**	GRGVK/QFYP	**64**	LNRTSFLG
21	NFRK/QWGGE	**43**	GYFRPK/QHP		
22	VTGFFLNR	**44**	LRAAWAHN		

6.2.3 Binding assay of selected peptides by nanoporous optical interferometry

Nanoporous optical interferometry (NPOI) is a new interferometric technology for the measurement of protein receptor-ligand binding.[175] NPOI is a label-free binding assay using a defined nano-porous silicon substrate to measure the refractive index changes on the porous surface by white-light interferometry (Figure 5A).[176] The binding of ligands to a receptor surface leads to a change of refractive index which is read by the detector, resulting in a wavelength shift of the spectrum. Shifts are computed and translated into an optical path difference (OPD), which is proportional to the amount of surface bound protein and describes the kinetics of the reaction (Figure 5B).[176]

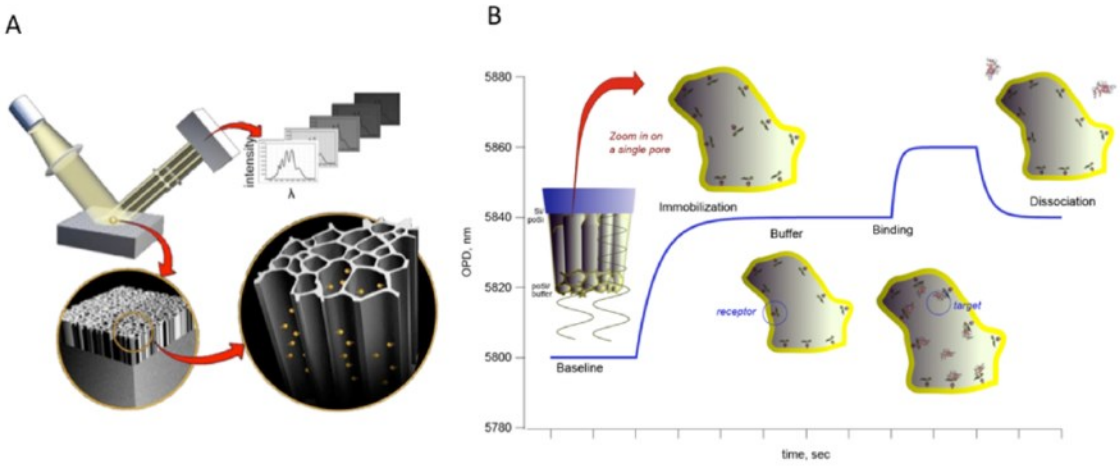


Figure 5. A) Artist's rendering of nano-porous biosensor principle; B) Schematic representation of interferogram of a typical nanoporous silicon biosensor experiment.[176]

The NPOI technology uses a 3D biosensor chip with a series of parallel pores with diameter between 10-150 nm and depth between 0.4-2.1 µm oriented on their surface.[177] With a three-dimensional surface area approximately 100 times more important than surface plasmon resonance (SPR), the sensitivity of NPOI is much higher than typical SPR.[177] The NPOI biochips nanoporous silicon substrate can be easily derivatized to attach ligands of interests by hydrophilic surface chemistries. Otherwise, NPOI can be easily performed in microtiter plate with high throughput or in flow cells with high data quality instead of a manual plate reader.

In our case, streptavidin chips which have streptavidin pre-immobilized and ready for immobilization of biotin-tagged molecules were used. A biotinylated Nef solution was passed in both sample and reference channels to finish immobilization.[178] The identified peptides which have been individually synthesized and purified were injected in the sample channel and ddH_2O was injected in the reference channel. Because of the high binding affinity of the streptavidin-biotin complex, the interaction of biotinylated Nef protein on the streptavidin chips is irreversible. Only the binding peptides were removed during the dissociation step. After regeneration step, the chip can be reused for next binding cycles. Dissociation constant (K_d) was obtained by Ski Report software (Table 2). NPOI reveals not only the binding constant but also the binding kinetics. The results obtained by NPOI showed that the selected peptides are poor to moderate ligands with the best K_d of 89.9 μM observed for peptide **33**.

Table 2. Kinetic characterization of selected peptides by NPOI

	Peptide	OPD	K_d(μM)		Peptide	OPD	K_d(μM)
33	NQRYEPYR	0.61	89.9	**16**	RLAQHYWA	1.21	150.8
56	RLYQAAQD	0.63	91.5	**17**	VHWFTYQE	0.61	164.6
8	LRYGYHPL	6.69	102.8	**51**	WAHNKYHN	0.65	165.2
3	FFKYTFFG	1.00	105.9	**55**	YAAQKFEW	1.11	194.5
59	WRTWHVEG	3.01	110.3	**54**	RFRWFKYP	6.30	205.4
52	RFLVHFKV	1.17	125.3	**9**	RDYKPPWW	8.49	209.6
47	TWAWDRFR	1.77	128.6	**31**	WSPFQRRF	1.57	235.4
23	NFWKAFYS	5.08	132.2	**62**	RGERNVWL	0.87	244.4
14	FGVRYTKY	1.05	139.3	**7**	RNGWTYRS	1.38	267.9
61	SRDWDRFR	1.11	143.1	**29**	DFRSVYQR	3.40	324.2

6.3 Conclusion

In summary, an OBOC octapeptide library has been prepared by the split-and-pool synthesis method. The library has been screened against the HIV-1 Nef protein and after 2 screening rounds, 64 hits have been isolated and their sequence successfully identified. The binding affinity of the selected peptides was determined by NPOI and the results showed that some selected peptides are poor to moderate ligands. These peptides will be used in

machine learning experiments to design new focused and biased peptide libraries. Nevertheless, this study was a good practice to identify pitfalls and weaknesses in the screening and validation process. Generally, the synthesis and screening steps were straightforward and efficient. The hit peptide sequencing by MS/MS was also efficient but the NPOI analyses showed important drawbacks. First the generated data were very variable and Nef leaching was also observed. To overcome this problem, another binding assay should be tested. For example, the ability of the selected peptides to inhibit specific PPI involving Nef could be determined in fluorescence polarization high-throughput assays with different Nef partners. Finally, the anti-HIV activity of the selected peptides coupled to a cell-penetrating peptide will be evaluated in a cellular assay.

6.4 Experimental Section

6.4.1 Materials and Equipment

TentaGel S NH_2 (TG) resin (130 μm, 0.29 mmol/g) was purchased from Rapp Polymer (Tübingen, Germany) and Rink Amide AM resin (0.65 mmol/g) from Chem-Impex International (Wood Dale, IL, USA). Fmoc-protected amino acids and coupling reagents, 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluoro phosphate (HCTU) were bought from Matrix Innovation Inc. (Quebec, QC, Canada). EZ-Link™Sulfo-NHS-Biotinylation Kit, Dynal® MPC™-1 magnet and Dynabeads® M-280 Streptavidin were acquired from ThermoFisher Scientific. Tris(hydroxymethyl) aminomethane (Tris) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) disodium salt were bought from Bio Basic Canada Inc. (Markham, Ontario, Canada). Alkaline phosphatase streptavidin (SA-AP) was bought from MP Biomedicals. Tween-20, bovine serum albumin (BAS) and sodium dodecyl sulphate (SDS) were bought from SIGMA-ALDRICH. Nef protein was obtained from the NIH AIDS program (catalog number 11478). Streptavidin functionalized nanoporous silicon biochips were from Silicon Kinetics (San Diego, CA, USA). All other commercial reagents and solvents were purchased from commercial suppliers and used without further purification. Reactions on solid support were performed in polypropylene fritted syringes from Roland Vetter Labordedarf OHG (Ammerbuch, Germany). RP-LC/MS analyses were conducted on Shimadzu Prominence LCMS-2020

system equipped with an electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) probes using a Gemini-NX column (4.6 mm × 250 mm, 5 μm C_{18}, 110Å, 1 mL/min) with a 20 min linear gradient from water (0.1% TFA) and CH_3CN (0.1% TFA) (CH_3CN 10-100%) and detection at 220 nm and 254 nm. Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry was performed on a AB SCIEX 4800 Plus MALDI-TOF/TOF instrument using alpha-cyano-4-hydroxycinnamic acid as matrix. The spectra were acquired using the 4000 Series Explorer Software (AbSciex, v 3.2.3). GeneMate Digital Dry Baths (Serial Number SB 1307382, BioExpress.) was used to heat beads suspended in 1% SDS solution for denaturation of Nef protein and any other materials bind on peptides. The PEAKS Studio software (Bioinformatics Solutions, v.7.0) was used for spectra analysis and DENOVO sequencing.

Buffer components: TBST buffer (50 mM Tris, 150 mM NaCl, 0.1% Tween 20, pH 7.4,), blocking buffer (3% BSA in TBST buffer), staining buffer (30 mM Tris, pH 8.5, 100 mM NaCl, 5 mM $MgCl_2$, 20 μM $ZnCl_2$), and PBS buffer (137 mM NaCl, 1.6 mM KCl, 1.5 mM KH_2PO_4, 6.4 mM Na_2HPO_4, pH 7.0).

6.4.2 Peptides synthesis

Peptides were synthesized by standard Fmoc solid-phase peptide synthesis. Briefly, amino acid couplings were performed with a solution of Fmoc-Xaa-OH (3 equiv), HCTU (3 equiv) and NMM (6 equiv) in DMF for 30 min. The coupling step was repeated once and the resin washed with DMF (5 × 30 s). The Fmoc protecting group was removed by treating the resin twice with a solution of 20% piperidine in DMF (v/v) for 10 min followed by washing with DMF (5 × 30 s).

Preparation of the peptide library. TentaGel S NH_2 130 μm (400 mg, 0.29 mmol, ~350 pmol/bead) was first swelled in 2 mL of DMF. A spacer Fmoc-Leu-βAla-βAla-Arg-Met-NH_2 was prepared on resin by standard Fmoc solid-phase peptide synthesis. The beads were suspended in a mixer DMF/DCM (1:3) then evenly split and placed into 17 separate reaction vessels. After washing with DMF (5 × 30 s), 17 different Fmoc-Xaa-OH (3 equiv) mixed with HCTU (3 equiv) and NMM (6 equiv) were independently added to a corresponding reaction vessels and the mixture agitated for 30 min. The coupling step was

repeated once and the resin washed with DMF (5 × 30 s). After each coupling step, the beads from each reaction vessel are combined to perform Fmoc deprotection with 20% piperidine in DMF for 2 × 10 min. Then the beads are randomly split again into the 17 different reaction vessels for the next coupling step. The library was prepared after repeating those process 8 times. After the last coupling/Fmoc-removal step, the N-terminal amine was acetylated with a solution of Ac_2O/DIEA/DMF (7:2:91) for 15 min. Finally, the resin was treated a solution of TFA, water and triisopropylsilane (TIS) (95:2.5:2.5) for 3 h for side chain deprotection followed by filtration and washing with DCM (5 × 30 s), MeOH (5 × 30 s) and dried *in vacuo*.

Cleavage from a single bead and sequencing. After removing the bound materials, positive beads were transferred in a microcentrifuge tube and treated overnight with 20 µL of CNBr (40 mg/mL) in CH_3CN/AcOH/H_2O (5:4:1) in the dark. The solvents were evaporated under vacuum to dryness and the peptides released from the bead were dissolved in 10 µL of 0.1% TFA in water. 1 µL of the peptide solution was mixed with 1 µL of 4-hydroxy-α-cyanocinnamic acid (5 mg/mL) in CH_3CN/0.1% TFA (1:1) and 1 µL of the mixture was spotted onto a MALDI sample plate for MS/MS analysis.

Synthesis of isolated peptides. All positive sequences were individually resynthesized without spacer on 100 mg of Rink Amide AM resin by standard Fmoc solid-phase peptide synthesis. After N-terminal acetylation with a solution of Ac_2O/DIEA/DMF (7:2:91) for 15 min, the peptide was cleaved from the resin and the side chains deprotected by treatment with a cleavage cocktail composed of TFA, water and triisoproylsilane (TIS) (95:2.5:2.5) for 3 h. The solvent was evaporated under reduced pressure and the resulting material precipitated with cold diethyl ether (3 × 30 mL). The solid was washed (2 × 30 mL) with cold diethyl ether and dried under vacuum.

HPLC purification of crude peptides. Crude peptides were purified by RP-HPLC on a Vydac 218MS C18 column (22.0 x 250 mm, 300 Å, 10 µm) using 0.1% TFA/H_2O (A) and 0.1% TFA/CH_3CN (B), with a linear gradient of 10% to 100% (B) for 20 min at 10 mL min^{-1} and UV detection at 220 nm and 254 nm. The collected fractions were lyophilized to afford the desired peptides as white powder. Peptides characterizations were done by RP-

HPLC and ESI-MS and all purities were over 95%. The purified peptides were sealed in amber vials and storage under -20 °C.

6.4.3 On-bead library screening

Two rounds of screening were performed for the library to minimize false positives due to nonspecific binding. Briefly, 100 mg of the resin was swelled in DMF (2 mL) and rotated in a syringe mounted with a fritted disc for 1 h at room temperature. Following the filtration, the resin was washed with DMF (3 × 2 mL), distilled H_2O (5 × 2 mL), TBST buffer (5 × 2 mL) and blocking buffer (5 × 2 mL). Afterwards the beads were blocked overnight at 4 °C with 2 mL of blocking buffer. The next day, the beads were washed with the blocking buffer (5 × 2 mL) and incubated with the blocking buffer containing a final concentration of 500 nM of the biotinylated Nef protein (~0.5 nmole of protein per mg of resin) for 6 h at 4 °C under gentle agitation. The beads were washed with TBST (5 × 2 mL) to remove any unbound protein. In another tube, 20μL of Dynabeads M-280 streptavidin (10 mg/mL) were washed with TBST to remove preservatives using Dynal® MPC™-1 magnet and resuspend in 100 μL TBST. Then the Dynabeads were poured in the syringe containing the library beads and rotated for 30 min at room temperature. The beads were transferred into a 15 mL Falcon tube and the tube placed in a Dynal® MPC™-1 magnet. The tube was swiveled slowly to make sure all the beads can come in close contact with the sides of the tube where the magnets are. Non-magnetized beads sank to the bottom. While the tube was touching the magnet to retain hits on the wall, extra liquid was poured and the non-magnetized collected in another Falcon tube. This procedure was repeated until no more beads appeared at the bottom of the tube. The same procedure was used once on nonmagnetized beads to make sure no positive hits were missing. The positive beads were washed with the blocking buffer (5 × 2 mL), transferred into a 2 ml syringe and incubated with SA-AP (final concentration of 10 μg/mL) for 10 min at 4 °C. Afterward, the beads were washed with the blocking buffer (5 × 2 mL) and the staining buffer (5 × 2 mL). Then 500 μL of the staining buffer containing 0.5 mg/mL of BCIP were added to the beads and intense turquoise color developed on positive beads in 15 min. The beads were poured into a petri dish and the turquoise colored beads picked up manually under a dissection

microscope with a micropipette and transferred into an eppendorf tube. Overall, 81 beads were selected after the two screening rounds.

Bound material removal. All hits were washed with distilled H_2O (5 × 1 ml) and suspended in 500 µL of 1% SDS solution for 5 min at 95 °C to remove any bound material. Afterward, the beads were extensively washed with AcOH/ACN/H_2O (3:4:3) (5 × 1 ml), AcOH/ACN (1:1) (5 × 1 mL) and distilled H_2O (5 × 1 mL) and individually isolated into microtubes for cleavage and MALDI-TOF MS/MS analysis.

6.4.4 Binding Capacity Measurement

Fresh SA-Chips were equilibrated by running 10 µL/min of PBS until stabilization was observed. Then 5 µg of biotinylated Nef protein diluted in PBS was immobilized using manufacturer's guidelines onto streptavidin SA-Chips.[178] All purified peptides were individually dissolved in ddH_2O at a final concentration of 539 µM. Afterward, 200 µL of peptide was injected to the sample channel and 200 µL of ddH_2O to the reference channel, the association step lasted 2000 s at 3 µL/min flow rate. On the other hand, the dissociation step was performed during 800 s at a 10µL/min flow rate. Finally, the chip was regenerated for next round with 20 mM glycine (pH=2.0) for 500 s at 5µL/min flow rate. All data obtained on a Silikon Kinetics SkiPro system were analyzed with the Ski Report software (version 1.1.64).

Chapter 7

Discussion and conclusion

7.1 General Discussion

Macrocyclic peptides have several attractive properties for the design and development of bioactive molecules and show great potential for the discovery of protein-protein interaction inhibitors. Combinatorial chemistry is one of the most powerful methods to optimally exploit the great molecular diversity that can be accessed with cyclic peptides. Among the different combinatorial methods available, the "one-bead-one-compound" approach is certainly the most accessible and economical. This attractive method is very straightforward and can be performed in any peptide science, combinatorial chemistry or chemical biology laboratory. However, the use of cyclic peptides in combinatorial OBOC libraries has been limited by difficulties in sequencing hit compounds after the screening. Lacking a free N-terminal amine, Edman degradation sequencing cannot be used on cyclic peptides and complicated fragmentation patterns are obtained by tandem mass spectrometry.

In order to avoid post-screening chemical modifications of the hit compounds, our objective was to develop more efficient decoding methods for OBOC cyclic peptide libraries. In a first approach described in chapter 3, we tested different combinations of ring-opening residues and linkers to evaluate their compatibility and efficiency to generate sequenceable linear peptides. Our goal was to identify the best residues for ring-opening residues and combinations and select a suitable approach to eliminate the encoding step in the preparation of OBOC cyclic peptide libraries. Methionine and ANP were tested as ring-opening residues in the different approaches and evaluated in combination with methionine, ANP or HMBA as linkers. Among the eight tested approaches, only the combination of ANP as ring-opening residue with methionine as linker did not generate a sequenceable linear peptide. The other approaches could be used for the decoding of OBOC cyclic peptide libraries and further studies are underway to evaluate their application potential. Among the approach tested, the dual ring-opening/cleavage approach was particularly attractive as it allows a single-step simultaneous linearization and cleavage from the bead. Therefore we continued our study with the development of new dual ring-opening/cleavage strategies with methionine and ANP.

In chapter 4, based on the macrocycle reopening approaches described by Lim *et al.* and Kodadek *et al.* with peptoids, we investigated the use of a methionine residue within the

macrocycle and as a linker. This strategy was very promising to allow reopening of the macrocyle and cleavage from the resin in a single step upon treatment with CNBr. The methionine used as linker had to be reversed to avoid the problems associated with the sequencing of linearized peptides bearing two C-terminal branches ending with a homoserine lactone. To adapt the strategy to an inverted methionine, the cyclic peptides were anchored via the side chain of a lysine residue. This modification allowed the release of a linear peptide with a single C-terminal homoserine lactone after treatment with CNBr. We also demonstrated that the linearized peptides obtained by this approach can be easily and efficiently sequenced by MALDI-TOF/TOF. The determination may be performed by manual analysis of the spectra MS/MS or by using sequencing software PEAKS. To demonstrate the compatibility of the strategy with OBOC libraries, a small cyclic heptapeptide library was prepared. The resulting crude peptides were analyzed by tandem MALDI-TOF MS. For each selected bead, the resulting MS spectrum showed the presence of peptides that could be unambiguously sequenced by MS/MS of the most important peak by manual analysis or de novo sequencing with the PEAKS software. One possible problem of this strategy is the oxidation of methionine residues. Inadvertent oxidation of the methionine residues may happen during storage, manipulation or screening and, in addition to significantly decreasing the yields, might prevent opening of the macrocycle. To avoid this problem, it is strongly suggested to treat the cyclic peptide library with a reducing solution to reduce any oxidized methionine residues prior to the tandem ring-opening/cleavage reaction.

After the development of the dual approach with methionine, we continued to look for a more efficient and less toxic approach for decoding OBOC macrocycle libraries. A chemical reagent free tandem ring-opening/cleavage approach was developed and presented in chapter 5. In this strategy, a photocleavable residue, 3-amino-3-(2-nitrophenyl) propionic acid (ANP), was introduced within the macrocycle and as a linker to allow simultaneous ring-opening and release from the resin upon treatment with UV irradiation. A series of cyclic peptides with different ring size and functionalized amino acids were tested. After dual ring-opening and cleavage, the MS spectra showed not only the expected $M+H^+$ molecular ion but also two other peaks corresponding to $M+H^+$-18 Da and $M+H^+$+14 Da. This observation led us to suspect that the reaction did not stop at the nitroso

product as expected. Further analysis of the photocleavage reaction mechanism of ortho-nitrophenyl derivatives, proposed that the $M+H^+$+14 Da ($M+H^+$-H_2O+MeOH) peak is generated during UV irradiation from the addition of MeOH (32 Da) on a dehydrated specie ($M+H^+$-18) of the nitroso derivative. To support this hypothesis, we used methanol-d_4 (36 Da) instead of MeOH （32 Da）as solvent and after UV irradiation the adduct peak was observed at +18 Da ($M+H^+$-18+36). In another test, $BuNH_2$ (73 Da) was added to the MeOH (5% v/v) during UV irradiation and the adduct peak was observed at +55 Da ($M+H^+$-18+73). Fortunately, all the MS/MS spectrum from the molecular ions $M+H^+$, $M+H^+$-18, $M+H^+$+14 can be sequenced manually or by using sequencing software PEAKS. These multiple MS/MS sequencing can be used as carbon copies of the same sequence and be very useful to identify the most efficient molecular ion for sequencing unknown peptides from OBOC libraries and significantly increase the confidence level during the sequencing process. Afterward, a 100 mg of ANP TG resin was used to prepare a small cyclic peptide library to prove the compatibility of our strategy with OBOC library. After the preparation of a small-sized cyclic peptide library, 15 beads were collected randomly and individually irradiated at UV 365 nm in MeOH. For each selected bead, the resulting MS spectrum showed the presence of peptides that could be unambiguously sequenced by MS/MS of the most important peak by manual analysis and/or de novo sequencing with the Peaks software. It should be noted that, since the ANP is photosensitive, it is strongly recommended to protect the library from light during synthesis and handling to prevent opening of the macrocycle or cleavage from the resin. Oxidation of tyrosine, tryptophane and methionine side chains was also observed. We suppose that oxidation of these side chains can be promoted by the UV irradiation. The oxidized species could also be successfully sequenced.

For the moment, we have developed two novel simultaneous ring-opening/cleavage methods for decoding OBOC macrocycle libraries. The greatest advantage of these methods is that the use of encoding tags is no longer necessary to decode and identify the compounds on the beads. As we know, the control of the outer/inner layer ratio in the topologically segregated bilayer bead strategy is important to avoid contact between the target and the tags. However, the ability of a molecule to access the interior of a bead depends on its size and the pores diameter. It has been demonstrated that proteins of 23.8

kDa are capable of penetrating inside TentaGel bead of 90 μm diameter. Compared to topological segregated bilayer beads method, the ring-opening strategy in OBOC libraries eliminates the risks of interference by the tag during screening since the same macrocyclic compound is found inside and outside the beads. Therefore the target protein may enter inside the beads without risks of interference.

Finally, in chapter 6, a OBOC combinatorial library containing linear octapeptides was synthesized and screened against the HIV-1 Nef protein for the discovery of new ligands. The main goal of this study was to practice the screening procedure to identify problems and potential pitfalls in the screening methods, library handling and MS analysis. Overall two screening rounds using two different methods were performed against biotinylated Nef protein. In the first round, the sorting was achieved with magnetic beads conjugated to streptavidin to significantly reduce the size of the library for the following screening method. In the second round, the sorting was performed with a colorimetric assay using a streptavidin-alkaline phosphatase (SA-AP) conjugate in presence of BCIP to eliminate false positive beads. After two screening rounds, 64 beads were selected and sequenced by MS/MS. Afterwards, the identified peptides were individually synthesized and purified for ligands validation by the nanoporous optical interferometry (NPOI) technology. Unfortunately, the results obtained with the NPOI showed that the selected peptides are poor ligands. These peptides are currently used in a machine learning approach to design new focused and biased libraries. Nevertheless, this study was an effective general repetition and showed that the screening of OBOC libraries does not afford black and white results but rather a large distribution of shades of grey. For this reason, the number of hits can significantly vary from one user to another as the selection is based on the cutoff of the person performing the experiment. Another important variant in the screening process is the blocking buffer that is used to increase or decrease the stringency of the experiment. Buffer components such as proteins (e.g. BSA or gelatin) and surfactants (e.g. Tween-20) can seriously influence specific and non-specific binding of the target protein to the beads. We observed that the number of hits can be considerably increased when lower stringent buffers are used during the screening. In a near future, we plan to screen OBOC cyclic peptide libraries against targets of interest using the tandem ring-opening/cleavage methods presented in chapter 4 and 5.

7.2 Conclusion

Cyclic peptides are very useful tools in chemical biology and medicinal chemistry to study and modulate proteins functions. With a great therapeutic potential, they have gained a lot of interest in drug discovery. To exploit the great molecular that can be accessed with cyclic peptides and allow their use in OBOC combinatorial libraries, we have investigated different ring-opening approaches and developed two novel dual ring-opening/cleavage strategies to prepare and decode OBOC cyclic peptide libraries. The developed strategies are very straightforward, reduce the need for post-screening chemical modification for sequencing and eliminate the risks of interference with the encoding tags. We have also showed that the developed approaches efficiently generate linear peptides and allow a fast sequence determination of selected peptides from combinatorial libraries by MS/MS.

A large amount of preliminary work and optimization have been made in this thesis. Some of the approaches presented in chapter three will be investigated in more details to develop new ring-opening strategies. On the other hand, strategies developed in chapter 4 and 5 will be used in a near future in screening programs against proteins of interest to discover new macrocyclic ligands. The developed ring-opening methods will also be tested for their compatibility and use with OBOC cyclic peptiod libraries. Finally the peptides identified in chapter 6 are currently used in a machine learning approach to design and prepare new libraries for additional screening rounds and to discover more efficient ligands against the Nef protein. Other binding assays such as fluorescence polarization are also going to be used to determine the affinity of the peptides for the Nef protein and test their ability to inhibit specific PPI with different Nef partners. Overall, the work presented in this thesis provides new strategies to use macrocycles in combinatorial libraries and their use will certainly help the discovery of new protein ligands and the development of new macrocycle-based therapeutic agents.

REFERENCE

1. Toogood, P. L. Inhibition of protein-protein association by small molecules: approaches and progress. *J. Med. Chem.* **45**, 1543–1558 (2002).
2. Albert, R. Scale-free networks in cell biology. *J. Cell Sci.* **118**, 4947–4957 (2005).
3. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3**, 301–317 (2004).
4. Wells, J. A & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001–1009 (2007).
5. Regulation of Apoptosis Overview. at <http://www.cellsignal.com/contents/science-pathway-research-apoptosis/regulation-of-apoptosis-signaling-pathway/pathways-apoptosis-regulation>
6. Nadamuni, S. Targeting Protein-Protein Interactions. *Gen. Eng. Biotech. News* **33**, 33 (2013).
7. Ivanov, A. A., Khuri, F. R. & Fu, H. Targeting protein – protein interactions as an anticancer strategy. *Trends Pharmacol. Sci.* **34**, 393–400 (2013).
8. Labbe, M., Laconde, G., Kuenemann, A., Villoutreix, B. O. & Sperandio, O. PPI-DB : a manually curated and interactive database of small non-peptide inhibitors of protein – protein interactions. *Drug Discov.Today.* **18**, 958–968 (2013).
9. Mullard, A. Protein – protein interaction inhibitors. *Nat. Rev. Drug Discov.* **11**, 173–175 (2012).
10. Cochran, A. G. Antagonists of Protein-Protein Interactions. *Chem. Biol.* **7**, 85-94 (2000).
11. Yin, H. & Hamilton, A. D. Strategies for targeting protein–protein interactions with synthetic agents angewandte. *Angew. Chem. Int. Ed. Engl.* **44**, 4130–4163 (2005).
12. Berg, T. Small-molecule inhibitors of protein-protein interactions. *Curr. Opin. Drug. Disc. Dev.* **11**, 666–674 (2008).
13. Zhao, L. & Chmielewski, J. Inhibiting protein – protein interactions using designed molecules. *Curr. Opin. Struct. Biol.* **15**, 9–12 (2005).
14. Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci.* **93**, 13–20 (1996).
15. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198 (1999).
16. Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., Salzberg, A. C., & Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* **25**, 71–75 (2007).

17. Smith, R. D., Hu, L., Falkner, J. A., Benson, M. L., Nerothin, J. P. & Carlson, H. A. Exploring protein-ligand recognition with Binding MOAD. *J. Mol. Graph. Model.* **24**, 414–425 (2006).

18. Keskin, O., Gursoy, A., Ma, B. & Nussinov, R. Principles of protein-protein interactions: what are the preferred ways for proteins to interact ? *Chem. Rev.* **108**, 1225–1244 (2008).

19. Bogan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1–9 (1998).

20. de Vega, M. J. P., Martín-Martínez, M. & González-Muñiz, R. Modulation of protein-protein interactions by stabilizing/mimicking protein secondary structure elements. *Curr. Top. Med. Chem.* **7**, 33–62 (2007).

21. Vazquez, A., Bond, E. E., Levine, A. J. & Bond, G. L. The genetics of the p53 pathway, apoptosis and cancer therapy. *Nat. Rev. Drug Discov.* **7**, 979–987 (2008).

22. Chène, P. Inhibiting the p53-MDM2 interaction: an important target for cancer therapy. *Nat. Rev. Cancer* **3**, 102–109 (2003).

23. Vousden, K. H. & Lane, D. P. p53 in health and disease. *Nat. Rev. Mol. Cell Biol.* **8**, 275–283 (2007).

24. Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* **408**, 307–310 (2000).

25. Soussi, T., Dehouche, K. & Béroud, C. p53 website and analysis of p53 gene mutations in human cancer : forging a link between epidemiology and carcinogenesis. *Hum. Mutat* **15**, 105–113 (2000).

26. Nigro, J. M., Baker, S. J., Preisinger, A. C., Jessup, J. M., Hosteller, R., Cleary, K., Signer, S. H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F. S., Weslon, A., Modali, R., Harris, C. C. & Vogelstein, B. Mutations in the p53 gene occur in diverse human tumour types. *Nature* **342**, 705–708 (1989).

27. Momand, J., Jung, D., Wilczynski, S. & Niland, J. The MDM2 gene amplification database. *Nucleic Acids Res.* **26**, 3453–3459 (1998).

28. Shair, M. D. A closer view of an oncoprotein-tumor suppressor interaction. *Chem. Biol.* **4**, 791–794 (1997).

29. Che, P. Inhibition of the p53-MDM2 interaction : targeting a protein-protein interface. *Mol. Cancer Res.* **2**, 20–28 (2004).

30. Buolamwini, J. K., Addo, J., Kamath, S., Patil, S. & Mason, D. Small molecule antagonists of the MDM2 oncoprotein as anticancer agents. *Curr. Cancer Drug Tar.* **5**, 57–68 (2005).

31. Duncan, S. J., Gru, S., Williams, D. H., Mcnicholas, C., Purewal, R., Hajek, M., Gerlitz, M., Martin, S., Wrigley, S. K., & Moore, M. Isolation and structure elucidation of chlorofusin , a novel p53-MDM2 antagonist from a fusarium sp . *J. Am. Chem. Soc.* **123**, 554–560 (2001).

32. Fasan, R., Dias, R. L. A.Moehle, K., Zerbe, O., Vrijbloed, J. W., Obrecht, D. & Robinson, J. A. Using a beta-hairpin to mimic an alpha-helix: cyclic peptidomimetic inhibitors of the p53-HDM2 protein-protein interaction. *Angew. Chem. Int. Ed. Engl* **43**, 2109–2112 (2004).

33. Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Klein, C., Fotouhi, N. & Liu, E.A. In vivo activation of the p53 pathway by small-Molecule antagonists of MDM2. *Science* **303**, 844–849 (2004).

34. Parks, D. J., Lafrance, L. V., Calvo, R. R., Milkiewicz, K. L., Gupta, V., Lattanze, J., Ramachandren, K., Carver, T. E., Petrella, E. C., Cummings, M. D., Maguire, D., Grasberger, B. L. & Lu, T. 1,4-Benzodiazepine-2,5-diones as small molecule antagonists of the HDM2–p53 interaction : discovery and SAR. *Bioorg. Med. Chem. Lett.* **15**, 765–770 (2005).

35. Didier, R. Rational design of protein – protein interaction inhibitors. *Medchemcomm* **6**, 51–60 (2014).

36. White, A. W., Westwell, A. D. & Brahemi, G. Protein-protein interactions as targets for small-molecule therapeutics in cancer. *Expert. Rev. Mol. Med.* **10**, 1–8 (2008).

37. Morelli, X. Chemical and structural lessons from recent successes in protein–protein interaction inhibition (2P2I). *Curr. Opin. Chem. Biol.* **15**, 475–481 (2011).

38. Huther, A. The emergence of peptides as therapeutic drugs for the inhibition of HIV-1. *Aids Rev.* **9**, 208–217 (2007).

39. Mason, J. M. Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. *Futur. Med. Chem.* **2**, 1813–1822 (2010).

40. Vlieghe, P., Lisowski, V., Martinez, J. & Khrestchatisky, M. Synthetic therapeutic peptides: science and market. *Drug Discov.Today.* **15**, 40–56 (2010).

41. Samanen, J., Ali, F., Romoff, T., Calvo, R., Sorenson, E., Vasko, J., Storer, B., Berry, D., Bennett, D. & Strohsacker, M. Development of a small RGD peptide fibrinogen receptor antagonist with potent antiaggregatory activity in vitro. *J. Med. Chem.* **34**, 3114–3125 (1991).

42. 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 analogu *Angew. Chem. Int. Ed.* **47**, 2595–2599 (2008).

43. Zhang, Y., Zhou, S., Wavreille, A.-S., DeWille, J. & Pei, D. Cyclic peptidyl inhibitors of Grb2 and tensin SH2 domains identified from combinatorial libraries. *J. Comb. Chem.* **10**, 247–255 (2008).

44. Weide, T., Modlinger, A., Kessler, H. & München, T. U. Spatial screening for the identification of the bioactive conformation of integrin ligands. *Top. Curr. Chem.* **272**, 1–50 (2007).

45. Chen, J., Warren, J. D., Wu, B., Chen, G., Wan, Q. & Danishefsky, S. J. A route to cyclic peptides and glycopeptides by native chemical ligation using in situ derived thioesters. *Tetrahedron Lett.* **47**, 1969–1972 (2006).

46. Souers, A. J. & Ellman, J. A. β-Turn mimetic library synthesis : scaffolds and applications. *Tetrahedron* **57**, 7431–7448 (2001).

47. Hanessian, S. *et al.* Design and synthesis of conformationally constrained amino acids as versatile scaffolds and peptide mimetics. *Tetrahedron* **53**, 12789–12854 (1997).

48. Stradley, S. J., Rizo, J., Bruch, M. D., Stroup, A. N. & Gierasch, L. M. Cyclic pentapeptides as models for reverse turns: determination of the equilibrium distribution between type I and type II conformations of Pro-Asn and Pro-Ala beta-turns. *Biopolymers* **29**, 263–287 (1990).

49. Koehn, F. E. & Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* **4**, 206–220 (2005).

50. Che, Y., Marshall, G. R., Che, Y. & Marshall, G. R. Priviledge scaffolds targeting reverse-turn and helix recognition. *Exp. Opin. Ther. Targets* **12**, 101–114 (2008).

51. Tapeinou, A., Matsoukas, M., Simal, C. & Tselios, T. Cyclic peptides on a merry-go-round; towards drug design. *Biopolym. (Peptide Sci.)* **104**, 453–461 (2015).

52. Liskamp, R. M. J., Rijkers, D. T. S., Kruijtzer, J. A. W. & Kemmink, J. Peptides and proteins as a continuing exciting source of inspiration for peptidomimetics. *ChemBioChem.* **12**, 1626–1653 (2011).

53. Merrifield, R. B. Solid phase peptide synthesis I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).

54. Geysen, H. M., Meloent, R. O. B. H. & Bartelingt, S. J. Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. U.S.A* **81**, 3998–4002 (1984).

55. Lam, K. S. Mini-review: application of combinatorial library methods in cancer research and drug discovery. *Anticancer drug Des.* **12**, 145–167 (1997).

56. Furka, A., Sebestyén, F., Asgedom, M. & Dibó, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493 (1991).

57. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354**, 82–84 (1991).

58. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature* **354**, 84–86 (1991).

59. Lam, K. S., Lebl, M. & Krchnák, V. The 'One-Bead-One-Compound' combinatorial library method. *Chem. Rev.* **97**, 411–448 (1997).

60. Patel, S., Giles, D., Chowdary, A., Patel, E. & Thomas, V. A Brief Introduction of combinatorial chemistry. *Res. J. Pharm., Biol. Chem. Sci.* **3**, 125–136 (2012).

61. Scott Jamie K., S. G. P. Searching for peptide ligands with an epitope library.pdf. *Science.* **249**, 386–390 (1990).

62. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. Peptides on phage : A vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. U.S.A* **87**, 6378–6382 (1990).

63. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. Random peptide libraries: a source of specific protein binding molecules. *Science.* **249**, 404–406 (1990).

64. Fodor, S., Read, J., Pirrung, M., Stryer, L., Lu, A. & Solas, D. Light-directed, spatially addressable parallel chemical synthesis. *Science.* **251**, 767–773 (1991).

65. Frank, R. Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* **48**, 9217–9232 (1992).

66. Bradley B. Brown, D. S. W. & H. M. G. A single-bead decode strategy using electrospray ionization mass spectrometry and a new photolabile linker : 3-Amino-3-(2-nitrophenyl) propionic acid. *Mol. Divers.* **1**, 4–12 (1995).

67. Zhou S., Kermit L. Carraway lll, Michael J. Eck, Stephen C. Harrison, Ricardo A. Feldman, Moosa Mohammadi, Joseph Schlessinger, Stevan R. Hubbard, Darrin P. Smith, Charis Eng, Maria J. Lorenzo, Bruce A. J. Ponder, & Lewis C. Cantley. Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536–539 (1995).

68. Trepel, M., Arap, W. & Pasqualini, R. In vivo phage display and vascular heterogeneity: implications for targeted medicine. *Curr Opin Chem Biol* **6**, 399–404 (2002).

69. Rodi, D. J., Makowski, L. & Kay, B. K. One from column A and two from column B: The benefits of phage display in molecular-recognition studies. *Curr. Opin. Chem. Biol.* **6**, 92–96 (2002).

70. Hoess, R. H. Protein design and phage display. *Chem. Rev.* **101,** 3205–3218 (2001).

71. Greenwood, J., Willis, A. E. & Perham, R. N. Multiple display of foreign peptides on a filamentous bacteriophage. *J. Mol. Biol.* **220**, 821–827 (1991).

72. Sternberg, N. & Hoess, R. H. Display of peptides and proteins on the surface of bacteriophage lambda. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1609–1613 (1995).

73. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).

74. Sidhu, S. S. Phage Display in Biotechnology and drug discovery. Group (Taylor & Ftancis/CRS: Boca Raton, FL, 2005).

75. Singh, A., Poshtiban, S. & Evoy, S. Recent advances in bacteriophage based biosensors for food-borne pathogen detection. *Sensors* **13**, 1763–1786 (2013).

76. Nicolao K. C. u, Xiao X.Y., Parandoosh Z., Senyei A. & Nova M. P. Radiofrequency encoded combinatorial chemistry. *Angew. Chem. Int. Ed.* **34**, 2289–2291 (1995).

77. Moran, E. J., Sarshar, S., Cargill, J. F., Shahbaz, M. M., Lio, A., Mjalli, A. M. M. & Armstrong, R. W. Radio frequency tag encoded combinatorial library method for the discovery of tripeptide-substituted cinnamic acid inhibitors of the protein tyrosine phosphatase PTPlB. *J. Am. Chem. Soc.* **117**, 10787–10788 (1995).

78. Frank, R. The SPOT-synthesis technique Synthetic peptide arrays on membrane supports—principles and applications. *J. Immunol. Methods* **267**, 13–26 (2002).

79. Lam, K. S. & Renil, M. From combinatorial chemistry to chemical microarray. *Curr. Opin. Chem. Biol.* **6**, 353–358 (2002).

80. Eden, W. Van, Meloen, R. H., Noordzu, A. & Embden, J. D. A. Van. Efficient mapping and characterization of a T cell epitope by the simultaneous synthesis of multiple peptides. *Eur J Immunol* **19**, 43–48 (1989).

81. Jung, G. & Beck-Sickinger, A. G. Multiple peptide synthesis methods and their applications. *Angew. Chem. Int. Ed.* **31**, 367–486 (1992).

82. Joe Maeji, N., Bray, A. M. & Mario Geysen, H. Multi-pin peptide synthesis strategy for T cell determinant analysis. *J. Immunol. Methods* **134**, 23–33 (1990).

83. Houghten, R. A. General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 5131–5 (1985).

84. Frank, R. Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* **48**, 9217–9232 (1992).

85. Frank, R. & Overwin, H. SPOT synthesis. Epitope analysis with arrays of synthetic peptides prepared on cellulose membranes. *Methods Mol. Biol.* **66**, 149–69 (1996).

86. Eichler, J., Beyermann, M. & Bienert, M. Application of cellulose paper as support material in simultaneous solid phase peptide synthesis. *Coll. Czech. Chem. Commun.* **54**, 1746–1752 (1989).

87. Eichler J, Bienert M & Stierandova A, L. M. Evaluation of cotton as a carrier for solid-phase peptide synthesis. *Pept. Res.* **4**, 296–307 (1991).

88. Wang Z., Laursen R. A. Multiple peptide synthesis on polypropylene membranes for rapid screening of bioactive peptides. *Pept. Res.* **5**, 275–280. (1992).

89. Reineke, U., Volkmer-Engert, R. & Schneider-Mergener, J. Applications of peptide arrays prepared by the spot-technology. *Curr. Opin. Biotechnol.* **12**, 59–64 (2001).

90. Stillman, B. a. & Tonkinson, J. L. FAST(TM) slides: A novel surface for microarrays. *Biotechniques* **29**, 630–635 (2000).

91. Benters, R., Niemeyer, C. M. & Wöhrle, D. Dendrimer-activated solid supports for nucleic acid and protein microarrays. *Chembiochem* **2**, 686–94 (2001).

92. Geysen, H. M., Rodda, S. J. & Mason, T. J. A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol. immunol.* **23**, 709–715 (1986).

93. Dooley, C. T. & Houghten, R. A. The use of positional scanning synthetic peptide combinatorial libraries for the rapid determination of opioid receptor ligands. *Life Sci.* **52**, 1509–1517 (1993).

94. Deprez, B. *et al.* Orthogonal Combinatorial Chemical Libraries. *J. Am. Chem. Soc.* **117**, 5405–5406 (1995).

95. Erb, E., Janda, K. D. & Brenner, S. Recursive deconvolution of combinatorial chemical libraries. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11422–11426 (1994).

96. Fenniri, H., Ding, L., Ribbe, A. E. & Zyrianov, Y. Barcoded resins: A new concept for polymer-supported combinatorial library self-deconvolution. *J. Am. Chem. Soc.* **123**, 8151–8152 (2001).

97. Falciani, C., Lozzi, L., Pini, A. & Bracci, L. Bioactive peptides from libraries. *Chem. Biol.* **12**, 417–426 (2005).

98. Pinilla C., Appel J.R., Blanc P & Houghten R.A. Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. *Bio Tech.* **13**, 901-905. (1992).

99. Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C. & Santi, D. V. Identification of highest-affinity ligands by affinity selection from equimolar peptide mixtures generated by robotic synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4505–4509 (1992).

100. Kerr, J. M., Banville, S. C. & Zuckermann, R. N. Encoded combinatorial peptide libraries containing non-natural amino acids. *J.Am.Chem.Soc.* **115**, 2529–2531 (1993).

101. Zhou, S. SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767–778 (1993).

102. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354**, 82–84 (1991).

103. Kumaresan, P. R., Wang, Y., Saunders, M., Maeda, Y., Liu, R., Wang, X. & Lam, K. S. Rapid discovery of death ligands with one-bead-two-compound combinatorial library methods. *ACS Comb. Sci.* **13**, 259–264 (2011).

104. Peng, L., Liu, R., Marik, J., Wang, X., Takada, Y. & Lam, K. S. Combinatorial chemistry identifies high-affinity peptidomimetics against alpha4beta1 integrin for in vivo tumor imaging. *Nat. Chem. Biol.* **2**, 381–389 (2006).

105. Christensen, C., Groth, T., Bruun Schiødt, C., Tækker Foged, N. & Meldal, M. Automated sorting of beads from a 'One-Bead-Two-Compounds' combinatorial library of metalloproteinase inhibitors. *QSAR Comb. Sci.* **22**, 737–744 (2003).

106. Liu, T., Liu, Y., Kao, H.-Y. & Pei, D. Membrane permeable cyclic peptidyl inhibitors against human peptidylprolyl isomerase Pin1. *J. Med. Chem.* **53**, 2494–24501 (2010).

107. Garaud, M. & Pei, D. Substrate profiling of protein tyrosine phosphatase PTP1B by screening a combinatorial peptide library. *J. Am. Chem. Soc.* **129**, 5366–5367 (2007).

108. Protein, B. *et al.* Solid-phase peptide library synthesis on HiCore resin for screening substrate specificity of Brk protein tyrosine kinase. *J. Comb. Chem.* **10**, 20–23 (2008).

109. Alluri, P., Liu, B., Yu, P., Xiao, X. & Kodadek, T. Isolation and characterization of coactivator-binding peptoids from a combinatorial library. *Mol. BioSyst.* **2**, 568–579 (2006).

110. Martínez-Ceron, M. C. *et al.* Affinity chromatography based on a combinatorial strategy for rerythropoietin purification. *ACS Comb. Sci.* **13**, 251–258 (2011).

111. Sweeney, M. C., Wavreille, A., Park, J., Butchar, J. P., Tridandapani, S. & Pei, D. Decoding Protein-Protein Interactions through Combinatorial Chemistry : *Biochem.* **44**, 14932–14947 (2005).

112. Alluri, P. G., Reddy, M. M., Bachhawat-Sikder, K., Olivos, H. J. & Kodadek, T. Isolation of protein ligands from large peptoid libraries. *J. Am. Chem. Soc.* **125**, 13995–14004 (2003).

113. Camperi, S. A., Iannucci, N. B., Albanesi, G. J., Oggero Eberhardt, M., Etcheverrigaray, M., Messeguer, A., Albericio, F. & Cascone, O. Monoclonal antibody purification by affinity chromatography with ligands derived from the screening of peptide combinatory libraries. *Biotechnol. Lett.* **25**, 1545–1548 (2003).

114. Joo, S. H. & Pei, D. Synthesis and screening of support-bound combinatorial peptide libraries with free C-termini: determination of the sequence specificity of PDZ domains. *Biochemistry* **47**, 3061–3072 (2008).

115. Quarrell, R., Claridge, T. D., Weaver, G. W. & Lowe, G. Structure and properties of TentaGel resin beads: implications for combinatorial library chemistry. *Mol. Divers.* **1,** 223–232 (1996).

116. Bayer, E. & Rapp, W. Graft copolymers of crosslinked polymers and polyoxyethylene, processes for their production, and their usage. *US Pat.* **4**,**908**,**405** (1990).

117. Pei, D. On-bead library screening made easier. *Chem. Biol.* **17**, 3–4 (2010).

118. Hendrickson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proc. Natl. Acad. Sci.* **86**, 2190–2194 (1989).

119. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* **9**, 580–588 (2009).

120. Biometrica, U. COPAS Imstruments for Large Particle Flow Cytometry.

121. Cho, C.-F., Behnam Azad, B., Luyt, L. G. & Lewis, J. D. High-throughput screening of one-bead-one-compound peptide libraries using intact cells. *ACS Comb. Sci.* **15**, 393–400 (2013).

122. Marani, M. M., Martínez Ceron, M. C., Giudicessi, S. L., Oliveira, E., Côté, S., Erra-Balsells, R., Albericio, F., Cascone, O. & Camperi, S. A. Screening of one-bead-one-peptide combinatorial library using red fluorescent dyes. Presence of positive and false positive beads. *J. Comb. Chem.* **11**, 146–150 (2009).

123. Donck, K .V., Bols, L., Bongaarts, R., Osta, P. V., Brouwer, J., Liskamp, R. M. J., Geysen, J., Perrault, D. & Kalutkiewicz, K. New technology automates sorting of large, bead-based combinatorial chemistry libraries. *18th Am. Pept. Symp.* **July 19-23,** (2003).

124. Reddy, M. M., Bachhawat-sikder, K. & Kodadek, T. Transformation of low-Affinity lead compounds into high-affinity protein capture agents. *Chem. & Biol.* **11**, 1127–1137 (2004).

125. Astle, J. M., Simpson, .L. S., Huang, Y., Reddy, M. M., Wilson, R., Connell, S., Wilson, J. & Kodadek, T. Seamless bead to microarray screening: rapid identification of the highest affinity protein ligands from large combinatorial libraries. *Chem. & Biol.* **17**, 38–45 (2010).

126. Liu, T., Qian, Z., Xiao, Q. & Pei, D. High-throughput screening of one-bead-one-compound libraries: identification of cyclic peptidyl inhibitors against calcineurin/NFAT interaction. *ACS Combi. Sci.* **13**, 537–546 (2011).

127. Simpson, L. S. & Kodadek, T. A cleavable scaffold strategy for the synthesis of one-bead one-compound cyclic peptoid libraries that can be sequenced by tandem mass spectrometry. *Tetrahedron Lett.* **53**, 2341–2344 (2012).

128. Upadhyaya, P., Qian, Z., Habir, N. A. A. & Pei, D. Direct Ras inhibitors identified from a structurally rigidified bicyclic peptide library. *Tetrahedron* **70**, 7714–7720 (2014).

129. Doran, T. M., Gao, Y., Simanski, S., McEnaney, P. & Kodadek, T. High affinity binding of conformationally constrained synthetic oligomers to an antigen-specific antibody: Discovery of a diagnostically useful synthetic ligand for murine Type 1 diabetes autoantibodies. *Bioorg. Med. Chem. Lett.* **25**, 4910–4917 (2015).

130. Wu, J., Ma, Q. N. & Lam, K. S. Identifying substrate motifs of protein kinases by a random library approach. *Biochem.* **33**, 14825–14833 (1994).

131. Song, A., Zhang, J., Lebrilla, C. B. & Lam, K. S. A novel and rapid encoding method based on mass spectrometry for 'one-bead-one-compound' small molecule combinatorial libraries. *J. Am. Chem. Soc.* **125**, 6180–6188 (2003).

132. Wang, X., Zhang, J., Song, A., Lebrilla, C. B. & Lam, K. S. Encoding method for OBOC small molecule libraries using a biphasic approach for ladder-synthesis of coding tags. *J. Am. Chem. Soc.* **126**, 5740–5749 (2004).

133. Biederman, K. J., Lee, H., Haney, C. a, Kaczmarek, M. & Buettner, J. a. Combinatorial peptide on-resin analysis: optimization of static nanoelectrospray ionization technique for sequence determination. *J. Pept. Res.* **53**, 234–243 (1999).

134. Ngoka, L. C. & Gross, M. L. Multistep tandem mass spectrometry for sequencing cyclic peptides in an ion-trap mass spectrometer. *J. Am. Soc. Mass. Spectrom.* **10**, 732–746 (1999).

135. Redman, J. E., Wilcoxen, K. M. & Ghadiri, M. R. Automated mass spectrometric sequence determination of cyclic peptide library members. *J. Comb. Chem.* **5**, 33–40 (2003).

136. Youngquist, R. S., Fuentes, G. R., Lacey, M. P. & Keough, T. Generation and screening of combinatorial peptide libraries designed for rapid sequencing by mass spectrometry. *J. Am. Chem. Soc.* **117**, 3900–3906 (1995).

137. Sepetov, N., Issakova, O., Krchnak, V. & Lebl, M. Peptide sequencing using mass spectrometry. *US Pat.* **5**,**470**,**753**, (1995).

138. Davies, M. & Bradley, M. Inverted peptides-single bead analysis by methionine scanning and mass spectrometry. *Tetrahedron Lett.* **38**, 8565–8568 (1997).

139. Sthilaire, P. M., Lowary, T. L., Meldal, M. & Bock, K. Oligosaccharide mimetics obtained by novel, rapid screening of carboxylic acid encoded glycopeptide libraries. *J. Am. Chem. Soc.* **120**, 13312–13320 (1998).

140. Wang, P., Arabaci, G. & Pei, D. Rapid sequencing of library-derived peptides by partial edman degradation and mass spectrometry. *J. Comb. Chem.* **3**, 251–254 (2001).

141. Shin, D. S., Kim, D. H., Chung, W. J. & Lee, Y. S. Combinatorial solid-phase peptide synthesis and bioassays. *ChemInform* **37**, 517–525 (2006).

142. Joo, S. H., Xiao, Q., Ling, Y., Gopishetty, B. & Pei, D. High-throughput sequence determination of cyclic peptide library members by partial Edman degradation/mass spectrometry. *J. Am. Chem. Soc.* **128**, 13000–13009 (2006).

143. Lebl, M., Lam, K.S., Salmon, S.E., Krchnak, V., Sepetov, N., & Kocis, P. Topologically segregated, encoded solid phase libraries. *US Pat.* **5**,**840**,**48**, (1998).

144. Vágner, J., Barany, G., Lam, K. S., Krchnák, V., Sepetov, N. F. Ostrem, J. A., Strop, P. & Lebl, M. Enzyme-mediated spatial segregation on individual polymeric support beads: application to generation and screening of encoded combinatorial libraries. *Proc. Natl. Acad. Sci.* **93**, 8194–8199 (1996).

145. Liu, R., Marik, J. & Lam, K. S. A novel peptide-based encoding system for 'one-bead one-compound' peptidomimetic and small molecule combinatorial libraries. *J. Am. Chem. Soc.* **124**, 7678–7680 (2002).

146. Liu, R., Wang, X., Song, A., Bao, T. & Lam, K. S. Development and applications of topologically segregated bilayer beads in one-beone-compound combinatorial libraries. *QSAR Comb. Sci.* **24**, 1127–1140 (2005).

147. Wang, X., Peng, L., Liu, R., Gill, S. S. & Lam, K. S. Partial alloc-deprotection approach for ladder synthesis of 'one-bead one-compound' combinatorial libraries. *J. Comb. Chem.* **7**, 197–209 (2005).

148. Bédard, F., Girard, A. & Biron, É. A convenient approach to prepare topologically segregated bilayer beads for one-bead two-compound combinatorial peptide libraries. *Int. J. Pept. Res. Ther.* **19**, 13–23 (2013).

149. Kwon, Y.-U. & Kodadek, T. Encoded combinatorial libraries for the construction of cyclic peptoid microarrays. *Chem. Commun.* **44***,* 5704–5706 (2008).

150. Chait, B. T., Wang, R., Beavis, R. C. & Kent, S. B. Protein ladder sequencing. *Science* **262**, 89–92 (1993).

151. Sweeney, M. C. & Pei, D. An improved method for rapid sequencing of support-bound peptides by partial edman degradation and mass spectrometry. *J. Comb. Chem.* **5**, 218–222 (2003).

152. Thakkar, A., Wavreille, A. S. & Pei, D. Traceless capping agent for peptide sequencing by partial Edman degradation and mass spectrometry. *Anal. Chem.* **78**, 5935–5939 (2006).

153. Giudicessi, S. L. *et al.* Friendly strategy to prepare encoded one bead-one compound cyclic peptide library. *ACS Comb. Sci.* **15**, 525–529 (2013).

154. Quarrell, R., Claridge, T. D., Weaver, G. W. & Lowe, G. Structure and properties of TentaGel resin beads: implications for combinatorial library chemistry. *Mol. Divers.* **1**, 223–232 (1996).

155. Lee, J. H., Meyer, A. M. & Lim, H.-S. A simple strategy for the construction of combinatorial cyclic peptoid libraries. *Chem. Commun.* **46**, 8615–8617 (2010).

156. Liang, X., Girard, A. & Biron, E. Practical ring-opening strategy for the sequence determination of cyclic peptides from one-bead-one-compound libraries. *ACS Comb. Sci.* **15**, 535–540 (2013).

157. Lee, K. J. & Lim, H. S. Facile method to sequence cyclic peptides/peptoids via one-pot ring-opening/cleavage reaction. *Org. Lett.* **16**, 5710–3 (2014).

158. Liu, T., Qian, Z., Xiao, Q. & Pei, D. High-throughput screening of one-bead-one-compound libraries: Identification of cyclic peptidyl inhibitors against calcineurin/NFAT interaction. *ACS Comb. Sci.* **13**, 537–546 (2011).

159. Menegatti, S. *et al.* Reversible cyclic peptide libraries for the discovery of affinity ligands. *Anal. Chem.* **85**, 9229–9237 (2013).

160. Gurevich-Messina, J. M., Giudicessi, S. L., Martínez-Ceron, M. C, Acosta, G., Erra-Balsells, R., Cascone, O., Albericio, F. & Camperi, S.A. An efficient strategy for the preparation of one-bead-one-peptide libraries on a new biocompatible solid support. *Tetrahedron Lett.* **46**, 1561–1564 (2005).

161. Gurevich-Messina, J. M. *et al.* A simple protocol for combinatorial cyclic depsipeptide libraries sequencing by matrix-assisted laser desorption/ionisation mass spectrometry. *J. Pept. Sci.* **21**, 40–5 (2015).

162. Gross, E. & Witkop, B. Selective cleavage of the methionyl peptide bonds in ribonuclease with cyanogen bromide 1. *J. Am. Chem. Soc.* **83**, 1510–1511 (1961).

163. Kaiser, R. & Metzka, L. Enhancement of cyanogen bromide cleavage yields for methionyl-serine and methionyl-threonine peptide bonds. *Anal. Biochem.* **266**, 1–8 (1999).

164. Stemson, S. M. & Schreiber, S. L. An acid- and base- stable o-nitrobenzyl photolabile linker for solid phase organic synthesis. *Tetrahedron* **39**, 7451–7454 (1998).

165. Atherton, E., Logan, C. J. & Sheppard, R. C. Peptide synthesis. Part 2. Procedures for solid-phase synthesis using N^{α}-fluorenylmethoxycarbonylamino-acids on polyamide supports. Synthesis of substance P and of acyl carrier protein 65-74 decapeptide. *J. Chem. Soc. Perkin Trans.* **1**, 538-546 (1981).

166. Ariën, K. K. & Verhasselt, B. HIV Nef: role in pathogenesis and viral fitness. *Curr. HIV Res.* **6**, 200–208 (2008).

167. Abraham, L. & Fackler, O. T. HIV-1 Nef: a multifaceted modulator of T cell receptor signaling. *Cell Commun. Signal.* **10**, 39 (2012).

168. Landi, A., Iannucci, V., Van Nuffel, A., Meuwissen, P. & Verhasselt, B. One protein to rule them all: modulation of cell surface receptors and molecules by HIV Nef. *Curr. HIV Res.* **9**, 496–504 (2011).

169. Peterlin, B. M. & Trono, D. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat. Rev. Immunol.* **3**, 97–107 (2003).

170. Greenway, A. L., Holloway, G., McPhee, D. A., Ellis, P., Cornall, A. & Lidman, M. HIV-1 Nef control of cell signalling molecules: multiple strategies to promote virus replication. *J. Biosci.* **28**, 323–35 (2003).

171. Schwartz, O., Maréchal, V., Gall, S. Le, Lemonnier, F. & Heard, J.-M. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV–1 Nef protein. *Nat. Med.* **2**, 338–342 (1996).

172. Das, S. R. & Jameel, S. Biology of the HIV Nef protein. *Indian J Med Res.* **121(4)**, 315–332 (2005).

173. Miyamoto, S., Lehman, A. L. & Tuscano, J. M. Applications of one-bead libraries and chemical microarrays in Signal. *Acc. Chem.Res.* **36**, 370–377 (2003).

174. Lam, K. S. & Leblt, M. Streptavidin and avidin recognize peptide ligands with different motifs. *Immunomethods* **15**, 11–15 (1992).

175. Kinetics, S. Nanopore Optical Interferometry. at <http://www.siliconkinetics.com/pdf/SKi_Nanopore_Optical_Interferometry_AN2.pdf>

176. Latterich, M. & Corbeil, J. Label-free detection of biomolecular interactions in real time with a nano-porous silicon-based detection method. *Proteome Sci.* **6**, 31 (2008).

177. Kinetics, S. Benchmarking Nanopore Optical Interferometry (NPOI) against SPR. 1–4 at <http://www.siliconkinetics.com/pdf/SKi_SPR_Benchmarking_AN7.pdf>

178. Kinetics, S. & Diego, S. SKi Pro TM X-10 Instrument Manual. http://www.siliconkinetics.com/index.php (2010).