Spring 5-19-2018

Cyclic Urethane: A Versatile Handle for Synthetic Peptide Applications

Hader E. Elashal
hader.elashal@student.shu.edu

Cyclic Urethane: A Versatile Handle for Synthetic Peptide Applications

A dissertation submitted to Seton Hall University in partial fulfillment for the Doctor of Philosophy Degree

By:

Hader E. Elashal

May 2018

Department of Chemistry and Biochemistry

Seton Hall University

South Orange. NJ. 07079

USA

We certify that we have read this dissertation and that in our opinion it is adequate in scientific scope and quality as a dissertation for the degree of Doctor of Philosophy.

Dr. Monika Raj: _____ **Mentor (No Longer a SHU Faculty Member)**
Initialed by Dr. Cecilia Marzabadi, Co-Mentor, Chair

Dr. Cecilia Marzabadi: _____
Co-Mentor, Chair of Department

Dr. David Sabatino: _____
Member of Dissertation Committee

Dr. Cosimo Antonacci: _____
Member of Dissertation Committee

Seton Hall University

I dedicate this thesis to my parents, Eihab and Magda, my husband, Ahmed, and my daughter, Talia, for their love, support, and encouragement.

ABSTRACT

While peptides have impeccable potential due to their biological and therapeutic abilities, they are inherently limited in clinical practice due to their instability towards proteolysis and poor bioavailability. In the past few years peptidomimetics have emerged as a powerful class of compounds that overcomes the shortcomings associated with peptides. Peptidomimetics essentially mimic a natural peptide or protein in its chemical composition, structure and its ability to perform a biologically relevant role. In addition to increased stability and bioavailability, peptidomimetics can also have increased receptor binding affinity and selectivity, thereby making these mimics lead compounds in the field of drug design and discovery. Most recently, modifications to the peptide backbone have been introduced to increase the versatility of peptidomimetic applications. This thesis work is based on the unrestricted ability to modify amino acid side chain residues within biologically relevant peptides and form new classes of peptidomimetics for studying their structure-function properties.

Motivated by the idea of introducing a multipurpose moiety to increase the functionality of peptides, we sought to incorporate the cyclic urethane moiety otherwise known as 2-oxazolidinone and to explore the plethora of applications that accompany this incorporation. In the process of designing cyclic urethane containing peptides, we sought to peruse an approach where the cyclic urethane moiety would be derived from naturally occurring amino acids serine, threonine, glutamic acid, and/or cysteine.

Modification of serine's hydroxymethyl side chain to a 2- oxazolidinone moiety activates the peptide backbone chain and increases its susceptibility towards cleavage. Due to the versatility

of this moiety, 2-oxazolidinone has been used to explore various applications such as protease mimics, formation of peptide thioesters, and modified C-terminal peptides. When used as a protease mimic, formation of 2-oxazolidinone allows for site-selective cleavage of extremely unreactive peptide bonds using neutral aqueous conditions. This method exhibits broad substrate scope and selectively cleaves various bioactive peptides with post-translational modifications (e.g. N-acetylation and N-methylation) and mutations (D- and β-amino acids), which are unsuitable substrates for enzymes. Further application of this method has been demonstrated by the sequencing of cyclic peptides which is difficult to achieve by utilizing traditional methods such as Edman's degradation and MS/MS. Identifying the sequence of macrocyclic peptides is vital in exploring potential therapeutic candidates found in nature and/or created through split and pool techniques. Building on the susceptibility of 2-oxazolidinone to cleavage, this moiety was also utilized for the formation of peptide thioesters, which is of significance in native chemical ligation for the synthesis of large proteins. This approach allows the synthesis of peptide thioesters by using Fmoc SPPS, which is usually an incompatible method due to the nucleophilic secondary amine required for Fmoc removal. Moreover, 2-oxazolidinone was used for the synthesis of various C-terminally modified peptides which are otherwise unattainable without the use of specialized resins, handles, or linkers. By exploring the various applications of the cyclic urethane moiety, this thesis work has demonstrated the effectiveness and wide-spread applicability of cyclic urethane derived peptidomimetics in the field of synthetic organic chemistry.

AKNOWLEDGMENTS

"There are no secrets to success. It is the result of preparation, hard-work, and learning from failure." -Colin Powell

One does not stumble upon success; it is the result of hard-work and determination. For a person to maintain their drive for success, they must be surrounded by a strong support system, guidance, and encouragement. Many have contributed to the above-mentioned attributes, and for that I am forever in dept to them. No words can truly express my gratitude towards those who have helped me during my 4-year journey toward my degree.

I would like to recognize Dr. Monika Raj for accepting me as her first graduate student and for encouraging me to transfer into the Doctorate program. Under her guidance, I learned to work hard and persevere to accomplish my goals. She is very passionate about her research and invests much of her time towards the success of her students. The knowledge she imparted upon me has been a great help in my endeavors, and I believe my success is in part due to her support and mentorship. The skills I have learned from her will forever shape my professional life and career. Without her guidance, this fruitful dissertation would not have been possible.

I would also like to thank Dr. David Sabatino for all his help and guidance throughout my graduate school journey. While I was not officially a part of his research group, he always served as a second mentor to me, and he was never frugal with advice and encouragement. It was during my junior year as an undergraduate that I truly fell in love with the subject of Biochemistry while sitting in his class. He exhibits a true passion for teaching and instilling the love of science in others. My knowledge and foundation in Biochemistry is rooted in his teachings and remain with me until this day.

A special thank you is also in order for Dr. Cecilia Marzabadi and Dr. Cosimo Antonacci for their guidance and support during my final year of studies. I am grateful for their advice and suggestions. Dr. Marzabadi also played a role in developing my foundation in Organic Chemistry as a sophomore in my undergraduate studies. Without this foundation, I would not have been able to pursue a graduate degree that demands a strong organic chemistry skill set.

I would also like to thank the other members of my dissertation committee which include Dr. Wyatt Murphy and Dr. Yuri Kazakevich. Not only did they serve on my committee, but they have also imparted me with knowledge and wisdom as a student in their classes. Dr. Murphy has provided me with advice on multiple occasions, to which I am thankful. Meanwhile, Dr. Kazakevich has a direct hand in making sure my research proceeded smoothly by fixing the HPLC and LCMS whenever there was a technical issue.

One cannot mention the Department of Chemistry and Biochemistry without thanking Maureen. She is the core of the department; she is supportive and never fails to help students with any problems they are facing. In short, she is truly wonderful. I would also like to thank Dr. Snow, Dr. Kelty, Dr. Hanson, Fr. Gerry, Dr. Fadeev, and Dr. Gorun. They have provided me with knowledge and guidance either as my professors or as professors to classes to which I served as a TA.

Moreover, success is not possible without friends to share your experience, triumphs, and failures with. A journey is easier and more bearable with friends, and I know my journey would not have been the same without Tiauna Howard, Ryan Cohen, Neelam Lahankar, Zilma Muneeswaran, Lyssa Buisserth, and Yonnette Sim. We are all amino acids connected by a strong peptide bond, Dr. Raj's group. I am grateful for their support and friendship, and I wish them the

best of luck in what is yet to come. I know that they all have the strength to face anything that may come their way.

I am forever grateful for the love, support, and encouragement of my family. My parents, Eihab and Magda, raised me to be strong and determined. They always taught me to reach for the stars and that I can achieve all my dreams. The education and success of my siblings and I was always their priority, and they have made countless sacrifices towards that end. They are the best parents anyone can ask for, and I hope that I can continue to make them proud. I would not be where I am today without them. Next, I would like to thank Heidi, my sister. She is the embodiment of how a true sister should be. She always listens to me when I need to talk, encourages me when I lack motivation, and she is always here for me. Meanwhile, Basem, her twin and my younger brother, exhibits a loving and caring nature to which I am appreciative. I would like to thank my husband, Ahmed, for being by my side through the ups and downs of both my academic and personal journey. He accepts me as I am, through the good and bad, and that kind of support is at the core of my success. He always encourages me to push forward when I am ready to give up, and he always restores my confidence in my abilities when my faith waivers. I thank him for putting up with me and for his never-ending love and support. He is an amazing husband and an amazing father. Finally, I would like to thank the newest addition to our family, Talia, my daughter. She has given our lives meaning, joy, and beauty. I cannot imagine life without hearing the words "Mommy" everyday; these words give me strength to strive for success and to be the best role model that I can be.

TABLE OF CONTENTS

List of Supplemental Figures

ABBREVIATIONS AND SYMBOLS

µg	Microgram
µL	Microliter
Ala or A	Alanine
ANP	β-amino acid 3-amino-3-(2-nitrophenyl)propionic acid
Arg or R	Arginine
Asn or N	Asparagine
Asp or D	Aspartic Acid
Boc	*t*-butyloxycarbonyl Protecting Group
C	Carbon
Cbz	Benzyloxycarbonyl
CDI	1,1-Carbonyldiimidazole
CNBr	Cyanogen Bromide
CUT	Cyclic Urethane Technique
Cys or C	Cysteine
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DIAD	Diisopropyl azodicarboxylate
DIB	Diacetoxyiodobenzene
DIC	Diisopropylcarbodiimide
DIEA	N,N-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	N,N-Dimethylformamide
DMSO	Dimethyl sulfoxide
DSC	N,N-disuccinimidyl carbonate
DVB	Divinylbenzene
E. coli	Escherichia coli

Equiv.	Equivalents
ESI	Electrospray Ionization
FA	Formic Acid
Fmoc	9-fluorenylmethoxycarbonyl
Gln or Q	Glutamine
Glu or E	Glutamic Acid
Gly or G	Glycine
H or hr	Hour
H_2O	Water
HATU	*O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*'*N*'-tetramethyluronium hexafluorophosphate
HBr	Hydrobromic acid
HBTU	*O*-benzotriazol-1-yl-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate
HCTU	O-(6-Chlorobenzotriazol-1-yl)-N, N,N',N'-tetramethyluronium hexafluorophosphate
HF	Hydrogen fluoride or Hydrofluoric Acid
HFIP	1,1,1,3,3,3-hexafluoro-2-propanol
His or H	Histidine
HMBS	Heteronuclear Multiple Bond Correlation
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxy-benzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
Hse	Homoserine
HTS	High Throughput Screening
Ile or I	Isoleucine
IMS-MS	Ion Mobility Spectrometry- Mass Spectrometry
k_{cat}	Rate of Reaction
k_m	Substrate Concentration

LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
Leu or L	Leucine
Lys or K	Lysine
mCPBA	Meta-Chloroperoxybenzoic Acid
MDM2	Murine Double Minute 2 Protein
MeOH	Methanol
^{Me}*Oxd*	Methylated Oxazolidinone
Met or M	Methionine
mg	Milligram
Min	Minute
mL	Milliliter
mmol	Millimole
MPAA	4-mercaptophenylacetic acid
MS	Mass Spectrometer
MS/MS	Tandem Mass Spectrometry
N	Nitrogen
NCL	Native Chemical Ligation
NH_2	Amino group
nm	Nanometer Wavelength
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
^{o}	Degree
OAc	Acetoxy
OBOC	One Bead One Compound
OBTC	One Bead Two Compound
o-NBS-Cl	2-nitrobenzenesulfonyl chloride
Oxd	Oxazolidinone
Oxyma	Ethyl cyanohydroxyiminoacetete
p53	Tumor Protein

$Pd(PPh_3)_4$	Tetrakis(triphenylphosphine)palladium (0)
PEG	Polyethylene glycol
Pg	Protecting Group
pGlu	Pyro-Glutamyl
Phe or F	Phenylalanine
PhSNa	Sodium thiophenolate
Pro or P	Proline
PS	Polystyrene
PyBOP	Benzotriazol-1-yl-oxy tris(pyrrolidino)phosphonium hexafluorophosphate
Q-TOF	Quad Time of Flight
RP HPLC	Reverse Phase High Performance Liquid Chromatography
RT	Room Temperature
Ser or S	Serine
SPPS	Solid Phase Peptide Synthesis
TBTU	*O*-benzotriazol-1-yl-*N*,*N*,*N'*,*N'*-tetramethyluronium tetrafluoroborate
tBu	Tert-butyl group
TFA	Trifluoroacetic Acid
TG	TentaGel Resin
THF	Tetrahydrofuran
Thr or T	Threonine
Thz	Thiazolidinone
TIPS	Triisopropylsilane
tR	Retention Time
Trp or W	Tryptophan
Trt	Trityl Group
Tyr or Y	Tyrosine
UV-Vis	Ultraviolet-Visible Spectroscopy
Val or V	Valine
A	Alpha

B	Beta
γ	Gamma

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF SCHEMES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 1: GENERAL OVERVIEW OF PEPTIDE BIOCHEMISTRY AND METHODS FOR CHEMICAL SYNTHESIS AND MODIFICATION

1.1 GENERAL INTRODUCTION TO AMINO ACID AND PEPTIDE STRUCTURE

Amino acids play a fundamental role in nature as peptide building blocks and as intermediates in metabolic processes.[1] The twenty naturally occurring amino acids contain a wide array of chemical versatility.[2] The general structure of amino acids contains terminal amino and carboxyl groups which are bound by a central α-carbon. Attached to the α-carbon is a H atom and a side chain group designated as R, which serves as the basis of differentiation between one amino acid and another (Figure 1.1). This R group is what gives each amino acid its distinct identity and unique physical and chemical properties.

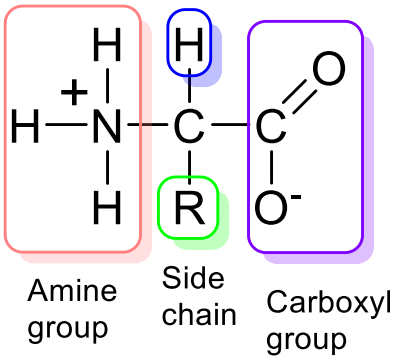


Figure 1.1 General amino acid ionic structure.

Based on the nature of the R group, amino acids can be classified into the following four categories: hydrophobic (nonpolar side chains), polar (uncharged residues with hydrophilic character), acidic (side chains with a negatively charged carboxyl group), and basic (side chains with a positively charged amino group). Moreover, due to the presence of four different functional groups attached

to the α-carbon, amino acids can exist in the form of two enantiomers, D(R) and L(S). While the L-form is predominant in nature, mutations can give rise to some D-amino acids as well. Amino acids also have a predominant ionic structure at physiological pH. The neutral zwitterion is formed by a balance in charges in between the N-/C-termini and it may also implicate a polar, charged amino acids side chain group (Figure 1.1). The latter can also affect the acid/base properties of amino acids which contain ionizable groups that can affect the overall/net charge of the amino acids in solution and as a function of changes in pH of the buffer system (Figure 1.2).

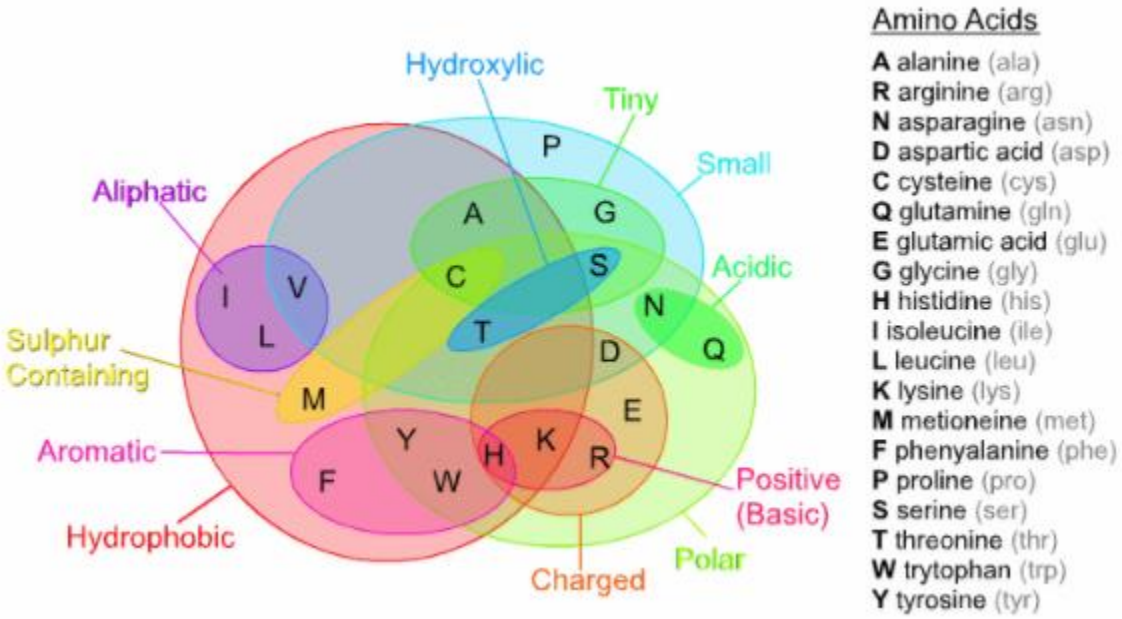
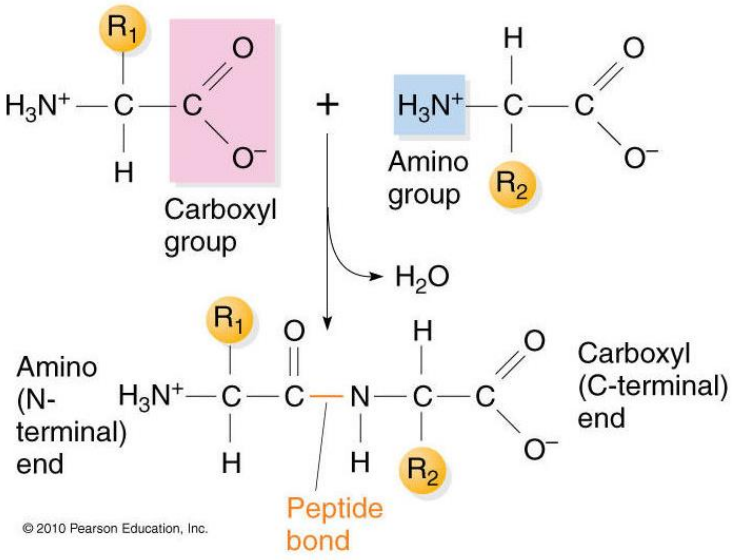


Figure 1.2 Venn diagram of amino acid properties. Figure adapted from reference (2).

Individual amino acids can be further linked together through peptide (amide) bonds to form a higher order structure, peptides. The formation of peptides takes place by condensation reactions in between the carboxylic group of one amino acid and the amino group of the adjacent amino acid thereby releasing a water molecule.[1] Peptides are comprised of amino acid linkages encompassing 50 or less residues; they are essential to the most basic biological processes, including signal

transduction and metabolism. For peptides to have the ability to perform a biologically important function, not only do they have to contain a certain amount of amino acids, the residues must also be incorporated in a specific order. The primary sequence of amino acids which make up the composition of a peptide is critical to its structure-function relationship. Small peptides ranging from two to dozens of amino acid residues can have a discernable effect on organism physiology.



Scheme 1.1 Peptide bond formation. Scheme adopted from reference (1).

Nature carries an extraordinary selection of biologically active peptides which participate in crucial roles in human function including hormones, neurotransmitters, and growth factors (Figure 1.3).[3] Furthermore, they serve as regulatory molecules in defense, stress, and reproductive processes.[4] With the discovery of over 7000 naturally occurring peptides, they are at the core of a living organism's daily operation.[3] Some of the most notable peptides include: carnosine **(1.1)**, a dipeptide found in muscle tissue; glutathione **(1.2)**, a tripeptide that serves as a scavenger for oxidizing agents which have been implicated in cancer; and enkephalins **(1.3)**, two pentapeptides that are naturally occurring pain relievers.[5] Even more vital to the human body are oxytocin **(1.4)**

and vasopressin **(1.5)**, two notable small peptide hormones with big effects.[5] Oxytocin is essential to child birth, where it aids in labor induction, contraction of the uterine muscle, and simulation of milk flow in a nursing mother. It also stimulates emotions of bonding and affection. Meanwhile, vasopressin maintains regular blood pressure level by acting as an antidiuretic in patients with low blood pressure. As demonstrated by the abovementioned peptides, naturally occurring peptides exhibit inherent properties which act as an inspiration for exploring their therapeutic profile and structure-function activity through chemical synthetic routes.

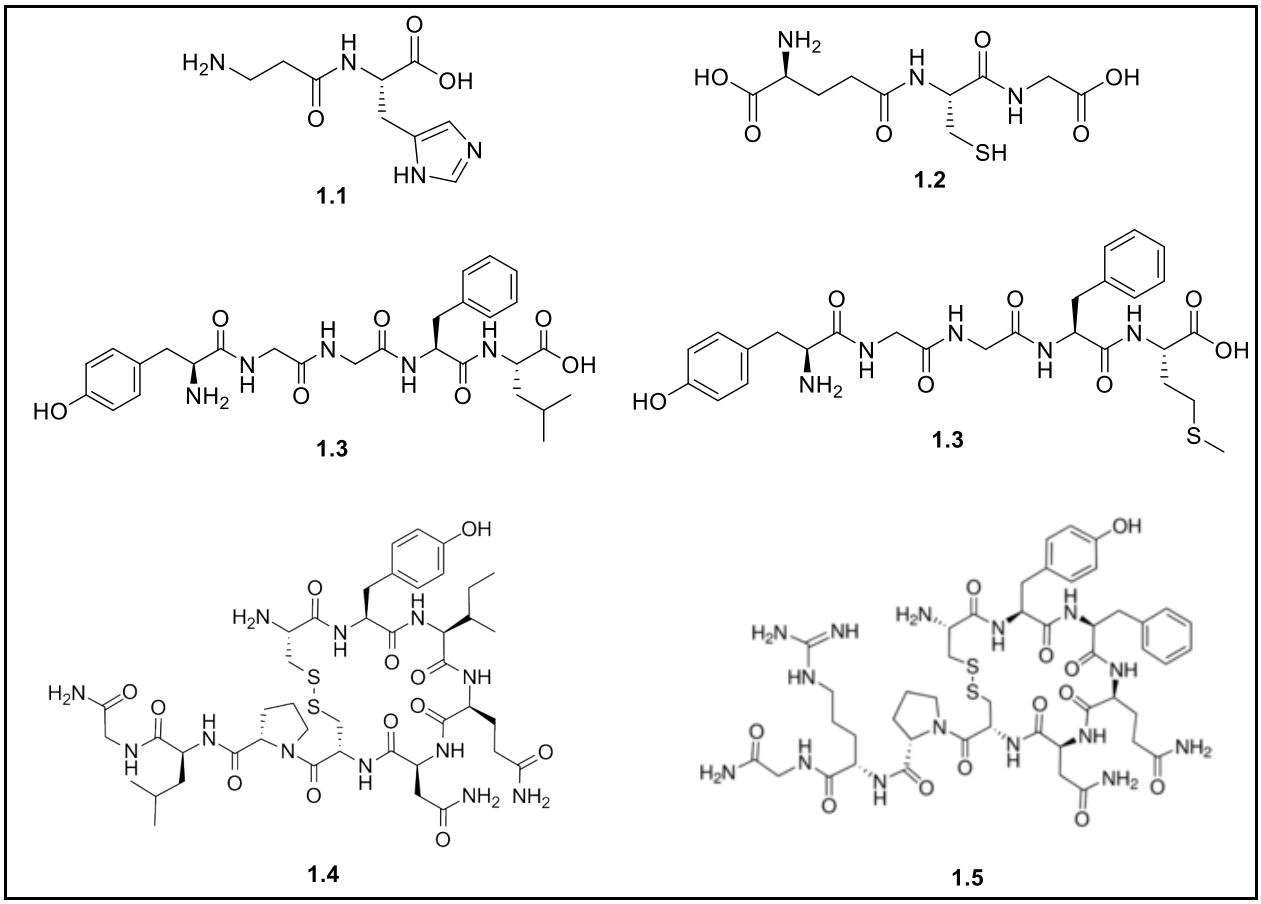


Figure 1.3 Bioactive peptides found in nature.

1.2 PEPTIDE SYNTHESIS OVERVIEW

During the last 50 years, there has been immense expansion and progress in the field of peptide and protein chemistry. Modern peptide chemistry includes many branches of study including synthesis, purification, analysis, structure and conformation investigation, and molecular modeling. Advances in biotechnology have facilitated the use of natural peptides as a starting point for novel synthetic peptide therapeutics with excellent specificity and efficacy which revolutionized peptide-based drug design and discovery. Since 2008, the peptide therapeutic market has exceeded the multi-billion-dollar level and over 400 peptides have been entered into clinical studies thus far.[6] Moreover, chemical synthesis has provided a practical method for allowing structural elucidation and synthetic manipulation of biologically relevant peptides that can only be isolated in small quantities. Synthetically derived peptides have also been essential in studying enzyme substrates, binding properties, and in NMR structural studies.

Generally, peptide synthesis can be described as the assembly of amino acid monomers stepwise through a continuous cycles of coupling reactions in between adjacent amino acids followed by the removal of a reversible protecting group leading to peptide elongation until the point of completion. The term peptide was first coined by Fischer and Fourneau who performed the first peptide synthesis at the turn of the 20th century.[7] Next, benzyloxycarbonyl (Cbz), the first reversible amino protecting group, was created which paved the way for further advances in the newly discovered field.[8] In 1953, Du Vigneaud was able to combine the earlier strategies and bring them to sophistication to construct a peptide successfully mimicking the properties of the hormone oxytocin.[9] However, the true breakthrough came with the introduction of dicyclohexylcarbodiimide (DCC), an unprecedented coupling reagent which refined the methodology of peptide bond formation and led to the milestone of solid phase peptide synthesis

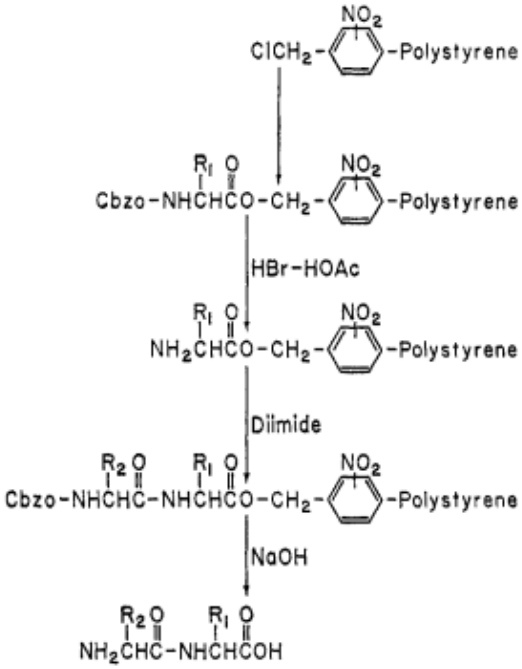
by Merrifield.[10] Building on previous discoveries in addition to his own extensive research, Merrifield was able to prefect the conditions for the synthesis of peptides on solid support and was awarded the Nobel Prize in Chemistry in 1984 for his pioneering contributions to the field of solid phase peptide synthesis.

While synthetic peptides can be made in solution or on a solid support, the solid support method is preferred due to the inherent limitations of solution phase synthesis. In solution, chemical synthesis gives rise to side products after each coupling step. Therefore, the intermediates and/or desired products must be purified, isolated, and characterized after each step in the deprotection/coupling cycle. Equally important, peptide insolubility in the solvents essential for synthesis proves problematic. Although attractive for large scale reactions due to quantitatively high yields, the demanding and laborious task of solution phase synthesis in addition to being ill-suited for synthesis of long peptides detract from its practical use in the advancement of peptide chemistry.

1.3 SOLID PHASE PEPTIDE SYNTHESIS

The discovery of solid phase peptide synthesis (SPPS) created a paradigm shift in the field of peptide synthesis by allowing the assembly of a peptide chain onto an insoluble and filterable porous support without the tedious task of purification and characterization after each reaction step. Pioneering the concept in the early 1960s, Merrifield proposed the use of a polystyrene-based solid support to assemble peptides in a stepwise fashion from the C to N terminus using amino terminal protected amino acids.[10] He synthesized a tetrapeptide by utilizing Cbz as an N^{α} protecting group, DCC as the coupling reagent, and hydrobromic acid (HBr) or saponification as the cleaving agent that releases the peptide from the solid support.[10]

Further advancement of the SPPS technique included the use of *t*-butyloxycarbonyl (Boc) as a terminal amino protecting group which would allow for the use of hydrogen fluoride (HF) as the acid medium for peptide release from the resin.[11,12] In its initial development stages, the methodology was solely based on acidolysis where side chain protecting groups and the linkage between the peptide and resin were relatively stable to moderate amounts of acid and labile in the presence of strong acidic conditions. Strong acids allow for the simultaneous deprotection of the amino acid side chain functionalities and liberation of the peptide from the resin support, thereby yielding an unprotected peptide to be purified. While the chemistry of deprotection and coupling was subject to variations, the standard procedure of SPPS allowed for the process to be automated, unlike solution-based synthesis.

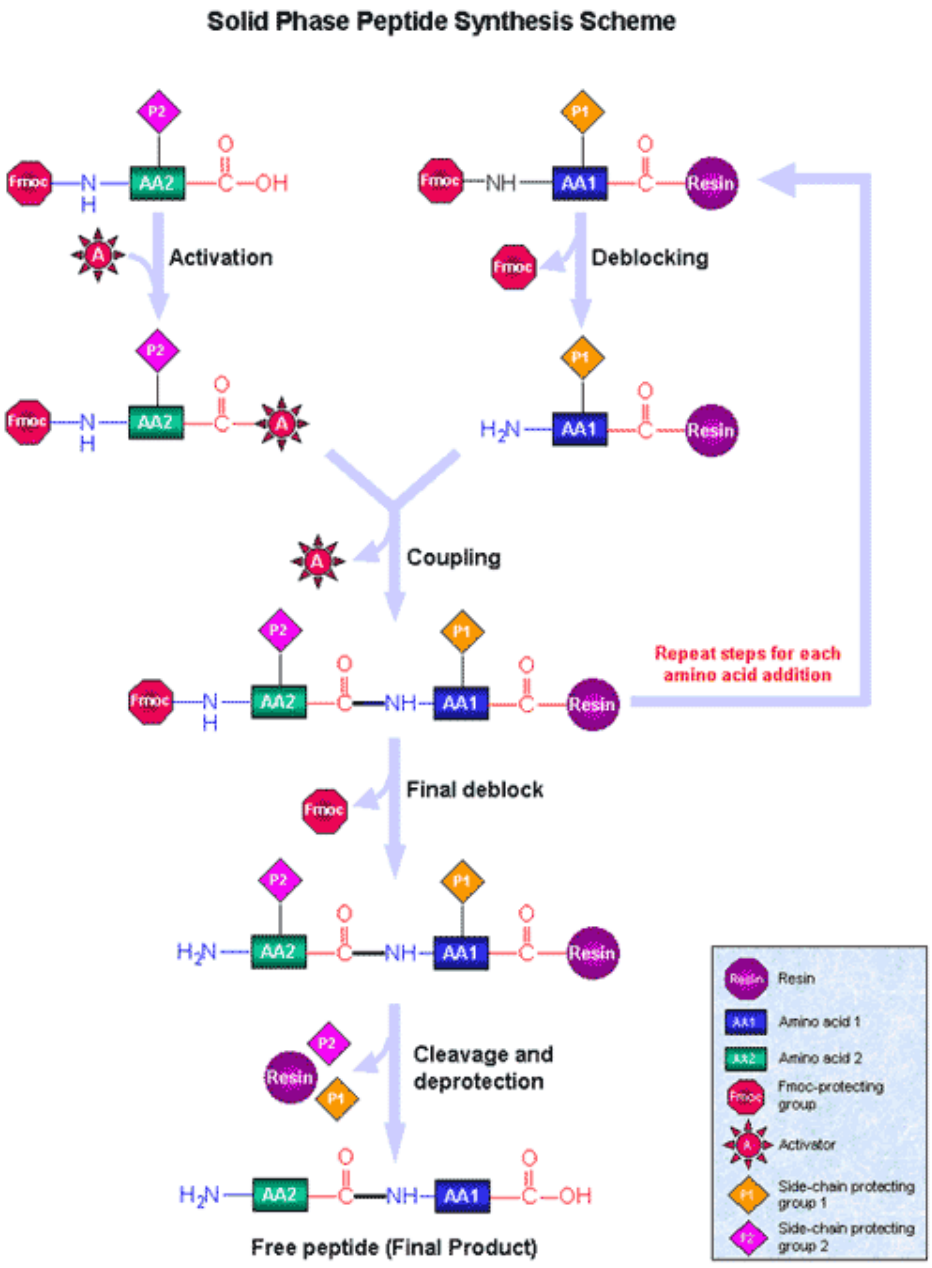


Scheme 1.2 SPPS procedure followed by Merrifield. Scheme adopted from reference (10) with permission.

In the period ranging from the 1960s to 1980s, progress had been made in the refinement of classical peptide synthesis methods to facilitate the synthesis of larger and biologically complex

peptides. The introduction of 9-fluorenylmethoxycarbonyl (Fmoc) by Carpino for N^{α} protection of amino acids offered a mild alternative to the acid-labile Boc-based synthesis strategies.[13,14] By utilizing Fmoc as a protecting group, SPPS shifted from a pure acidylosis methodology to an orthogonal technique which required a base for deprotection of the terminal Fmoc-protected amino group and an acid for the removal of side chain protecting group and peptide cleavage.[6] Advantages of utilizing Fmoc chemistry in comparison to the Boc protecting group include mild synthetic conditions, rapid deprotection of the N^{α} protecting group, and the possibility of monitoring coupling and deprotection through ultraviolet-visible spectroscopy (UV-Vis). Moreover, some peptides may be susceptible to acid-catalyzed side reactions which will interfere with the peptide efficacy and yields.

The advantages associated with SPPS outweigh the limitations that may be associated with the yields related to the synthesis of lengthy peptides. Generally, reagents that are low in molecular weight can be used in excess to yield a considerable quantity of product. Moreover, the easy filtration of excess reagents eliminates the need for intermediate purification that consumes ample amounts of time during solution phase synthesis. Equally important, the mild conditions promoted by Fmoc chemistry eliminated the possibility of any side reactions that may have risen from acid sensitive peptide components. Amino acid solubility is not an issue during the coupling cycle, while the solid support used to anchor the peptide is chemically compatible with the needed substrates, solvents, and reagents. The abovementioned benefits lend themselves to the successful expansion of SPPS into the billion-dollar industry it is today.



Scheme 1.3 General SPPS Protocol. Scheme adopted from: Sigma Aldrich

1.3.1 SOLID SUPPORT RESINS

Since the development of SPPS, various resins including those with linkers and spacers have been created to be adaptable to various synthetic conditions.[15] Since each type of resin has unique physical and chemical properties, the choice of solid support often dictates the success of

a peptide synthesized via SPPS. Currently, there are a plethora of commercially available solid supports that are compatible with peptide synthesis. Solid support specifications include stability, swelling and insolubility properties, and they must allow for filtration of excess reagents. Moreover, the resin must be unreactive towards all reagents and solvents that may be used during the synthesis process. Even more crucial to the successful chain elongation is extensive swelling to allow for reagent penetration into the resin pore and attachment of the first amino acid to the resin linker.

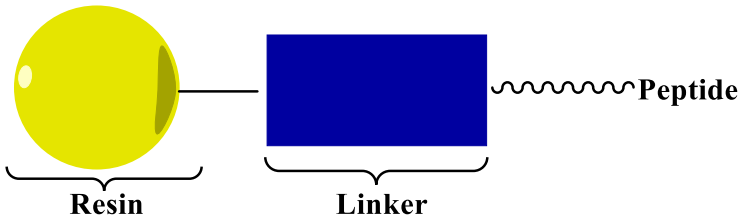


Figure 1.4 Solid support for peptide synthesis.

Resins are mainly classified into two main categories: polystyrene (PS) and polyethylene glycol (PEG). The hydrophobic PS resins can either be lightly or highly cross-linked, where the amount of divinylbenzene (DVB) can vary.[16] Lightly cross-linked PS resins tend to be gel-type and microporous, where space in between crosslinks contains solvent and solutions of polymer segments. Meanwhile, the more polar PEG resins are predominantly composed of ethylene glycol spacers. Moreover, they have no preference for swelling solvent since the PEG chain has good compatibility with most solvents. Other materials that have also been utilized as solid supports for SPPS including: silica, polymethacrylate, and controlled pore glass.[16] Another important resin property that influences reactivity and kinetics is size distribution. Most commonly, resin beads have a diameter of 100-200 mesh and 200-400 mesh which allow for effective swelling and coupling.[16]

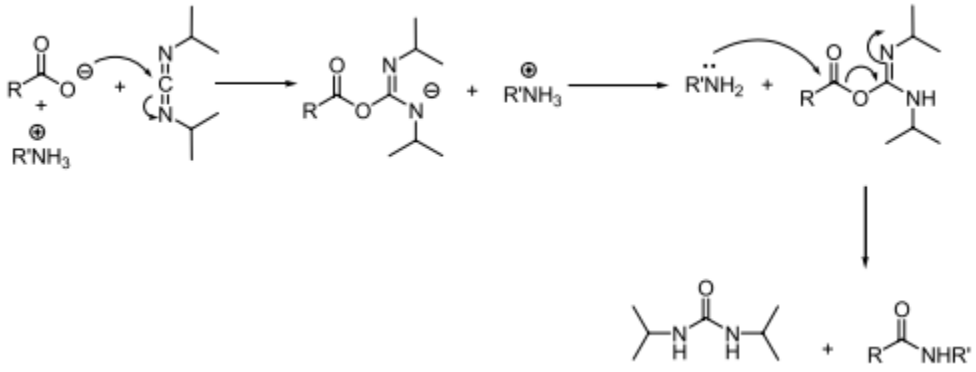
Since SPPS is traditionally executed in the C to N terminal direction, the linker present in the resin pore dictates the functional group that will be present at the C-terminus following cleavage of the peptide from the solid support since it serves as the point of connection between the first amino acid and resin bead.[17] The first resin ever used for SPPS was the non-polar cross-linked PS with a chloro-methyl group as a linker, which was later named the Merrifield resin. The first amino acid is loaded onto the resin by nucleophilic displacement of the chloride by the carboxylate group located on the amino acid. Upon HF-mediated cleavage of the peptide from the solid support, the carboxylic acid is regenerated at the C-terminus of the peptide. Presently, the Wang linker is more commonly used to yield a carboxylic acid at the C-terminus of a peptide; it was specifically designed for Fmoc-based synthesis, where the resulting peptide can be cleaved by TFA.[18] Another resin that can popularly yield a C-terminal carboxylic acid in addition to a peptide with protected side chains is 2-chlorotrityl chloride resin. In addition to generating a carboxylic acid at the C-terminus of a peptide, a carboxyamide is also another standard choice which is obtained by using the Rink linker.[19] There are a multitude of readily available resins and linkers which ease the synthetic burden and allow flexibility in the choice of C-terminal protecting group and loading condition pertaining to the first amino acid.

1.3.2 COUPLING REAGENTS

In the process of SPPS, the coupling step consists of the formation of an amide bond between two neighboring amino acid residues. The peptide bond is formed through activation of the carboxy component of one amino acid and the subsequent attack of the N^{α} group. While this reaction is thermodynamically favorable, the formation of a new amide bond has a high activation energy requirement. To aid in the coupling reaction initiation, coupling reagents serve as carboxy activators, often forming a more reactive C-terminal ester. In the past fifty years, many reagents

have been used to increase the efficiency of amino acid couplings and play a key role in generating considerably high yields and purities of the isolated, crude synthetic peptides.

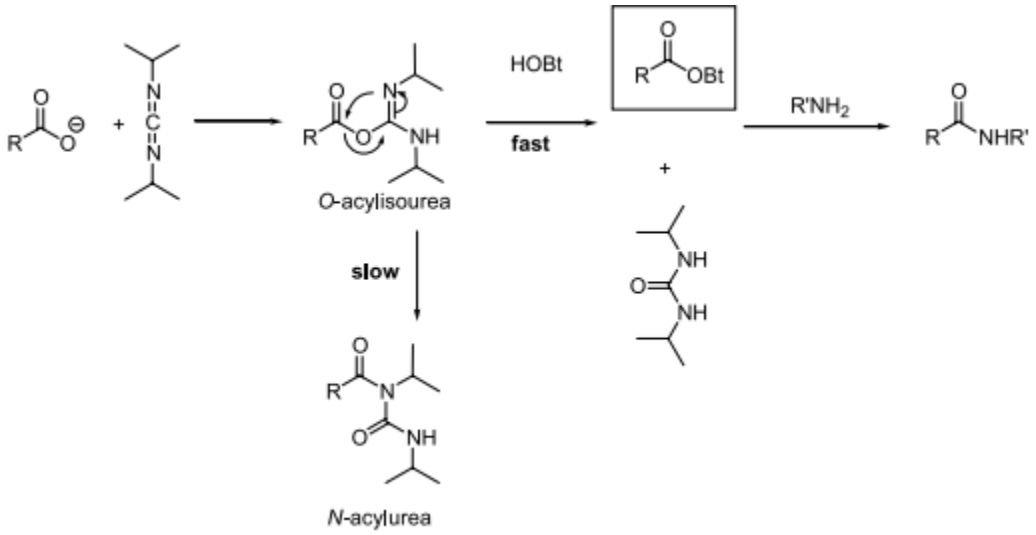
Carbodiimides such as dicyclohexylcarbodiimide (DCC) and diisopropylcarbodiimide (DIC) are considered to be classical *in situ* coupling reagents.[6,20,21] The activation step proceeds via nucleophilic attack of the carboxylate oxygen on the carbodiimide carbon. Upon formation of a reactive *O*-acylisourea species, a carboxylic ester with an activated leaving group, nucleophilic attack by the *N*-terminal amino group results in the peptide bond formation. While DCC and DIC proceed through similar mechanistic pathways, DIC is the reagent of choice due to its solubility in a majority of organic solvents, which also makes the removal of the unwanted urea by-product easier to accomplish via solvent filtration.



Scheme 1.4 One pot DIC coupling mechanism. Scheme adopted from reference (20) with permission.

While DIC executes coupling with high efficiency, it can be problematic. This method of activation has the potential to promote racemization and dehydration of asparagine (Asn) and glutamine (Gln) side chains. Using an additive in combination with DIC eliminateds the abovementioned limitations and accelerates carbodiimide-mediate couplings (Figure 1.5). Popular additives include 1-hydroxy-benzotriazole (HOBt) **[A]** and 1-hydroxy-7-aza-benzotiazole (HOAt) **[B]**, which are triazoles that attack the *O*-acylisourea to form an activated ester.[20] More recently,

ethyl cyanohydroxyiminoacetete (Oxyma) **[C]** has been developed as an alternative to HOAt and HOBt with comparable coupling efficiency.



Scheme 1.5 HOBt and DIC coupling mechanism. Scheme adopted from reference (18) with permission.

While carbodiimide reagents were the standard coupling reagents in peptide synthesis, newly developed activators have more commonly been put into practice due to their increased coupling kinetics. Examples of newly established coupling reagents in SPPS include: benzotriazol-1-yl-oxy tris(pyrrolidino)phosphonium hexafluorophosphate (PyBOP) **[D]**, *O*-benzotriazol-1-yl-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate (HBTU) **[E]**, *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*'*N*'-tetramethyluronium hexafluorophosphate (HATU) **[F]**, O-(6-Chlorobenzotriazol-1-yl)-N, N,N',*N*'-tetramethyluronium hexafluorophosphate (HCTU) **[G]**, and *O*-benzotriazol-1-yl-*N*,*N*,*N*',*N*'-tetramethyluronium tetrafluoroborate (TBTU) **[H]**.[6] The abovementioned coupling reagents also proceed via a similar mechanistic pathway to that of the carbodiimide reagents; the main difference in reagents classes is the initial rate of activation which lends itself to rapid coupling rates.

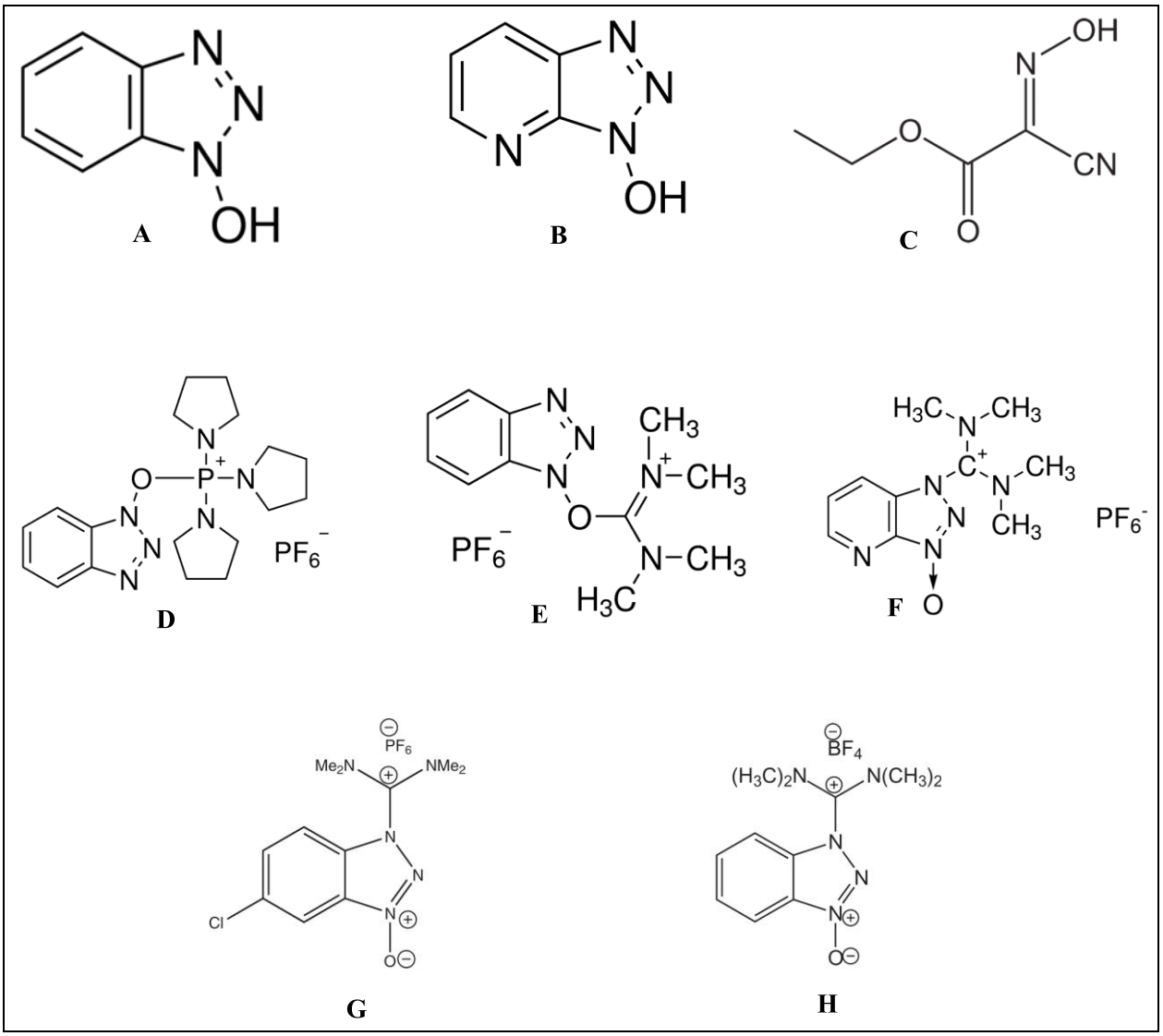


Figure 1.5 Structure of reagents used as additives during peptide coupling.

1.3.3 MODIFIED AMINO ACIDS

While peptides have impeccable potential due to their biological and therapeutic abilities, they are inherently limited due to their instability towards proteolysis and poor bioavailability.[22] In the past few years peptidomimetics have emerged as a powerful class of compounds which overcomes the shortcomings associated with peptides. Modified peptides (peptidomimetics) essentially mimic a natural peptide or protein in its chemical composition, structure and its ability

to perform a biologically relevant role. In addition to increased stability and bioavailability, peptidomimetics can also have increased receptor binding affinity and selectivity, thereby making these mimics leading compounds in the field of drug design and discovery.[23] By mapping structure-activity relationships, key residues that are responsible for a physiological effect can be identified while other non-essential amino acids can be manipulated to yield a more potent effect.

Incorporation of unnatural amino acids or modification of the peptide backbone chain provides peptidomimetics several advantages over naturally structured peptides. Unnatural amino acids can include *β*/*γ*-amino acids, which respectively contain two and three carbon atoms between the amino terminus and carboxyl terminus (Figure 1.6). In the case of *β*-amino acids, a total of 4 diastereomers are possible which expands the structural diversity in molecular design.[24] Moreover, D-amino acids can also be incorporated into a peptide to increase stability and induce a conformational change. By including several D-amino acid residues, the propensity for a peptide to form a secondary structure can be increased.[25] Most recently, modifications to the peptide backbone have been introduced to increase the versatility of peptidomimetic applications. Some notable examples include: glutamic acid (Glu) side chain cyclization to form a backbone pyro-glutamyl moiety (Figure 1.6),[26] which has been used in the synthesis of a solid support thioester precursor and enzyme mimetics; Hoffmann rearrangement of the Asn side chain to form a five-membered N-acylurea moiety,[27] which has been used as an enzyme mimic due to the susceptibility of C-N bond proceeding N-acylurea to hydrolysis; and modification of Methionine's (Met) side chain to a homoserine lactone moiety via treatment with cyanogen bromide (CNBr) which has long been used in the field of chemical cleavage as well.[28,29] Inspired by the versatility of applications relevant to the modification of amino acid side chains, this was designated as the topic of choice for the purpose of this thesis work.

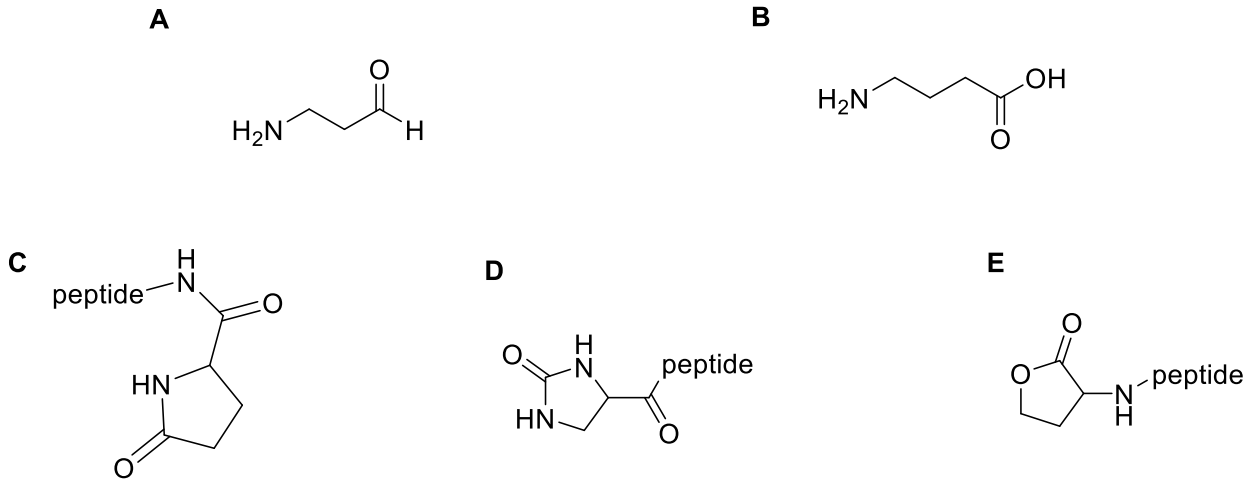


Figure 1.6 Modified amino acids. (A) β-amino acid (B) γ-amino acid (C) pyro-glutamyl moiety (D) N-acylurea (E) N-acylurea

1.4 CYCLIC URETHANE MOIETY

Amino acid modification to versatile moieties with broad applicability has gained considerable attention in the past few years. Motivated by the idea of introducing a multipurpose moiety to increase the functionality of peptides, we sought to incorporate the cyclic urethane moiety otherwise known as 2-oxazolidinone and explore the plethora of applications that accompany this incorporation (Figure 1.7).

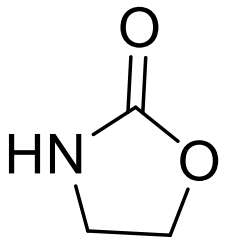


Figure 1.7 General structure of cyclic urethane.

Before its incorporation into peptides, cyclic urethanes gained wide-spread recognition as antimicrobial agents and chiral auxiliaries as popularized by Evans. As antimicrobial agents, they exhibit activity against multi-drug resistant gram-positive bacteria including *Staphylococcus*

aureus, *Staphylococcus epidermis*, *Streptococcus pneumoniae*, and *enterococci*.[30] Factors contributing to the successful application of 2-oxazolidinone in the medicinal chemistry field are related to its novel mode of action, binding selectivity, and its inhibition at the initiation phase of bacterial translation.[30] Owing to its unique chemical and biological properties, this moiety has promoted the development of derivatives that also exhibit potent activity.

Meanwhile, the use of chiral auxiliaries including oxazolidinones has been widely used in the synthesis of enantiomerically pure compounds since the early 1980s.[31]Auxiliaries such as oxazolidinone are substituted at the 4 and 5 positions to dictate controlled reactions. Through steric hinderance, the substituted positions direct the manner by which stereoselective transformations take place. Specifically, oxazolidinone chiral auxiliaries have been applied to aldol reactions, alkylation reactions, Diels-Alder reactions, and the synthesis of β-amino acids among many other synthetic transformations.[32-34] Requirements for the efficient use of a chiral auxiliary include a strong predisposition for a highly selective enolate reaction, a strong bias for enolate diastereofacial selection in new bond formation, and cleavage must occur in mild conditions that do not promote racemization of the desired product.[34] The abovementioned requirements are easily fulfilled through utilizing oxazolidinone as a chiral auxiliary.

Moreover, cyclic urethane has been incorporated into peptides to serve as a pseudo-proline surrogate due to its structural similarity to 5-membered proline (Pro) and its stability towards standard Fmoc synthesis, chemical ligation, and convergent synthesis conditions.[35] Owing to its structure, pseudo-prolines have the ability to introduce a "kink" into the peptide backbone which allows for disruptions of β-sheet structures which tend to be problematic during chain elongation due to intermolecular aggregation.[36,37] Incorporation of cyclic urethane into the peptide chain

increases peptide solvation, increases coupling kinetics, and induced a slight turn into the peptide backbone.

Due to their versatile applicability, many synthetic routes have been explored in the synthesis of oxazolidinones. Generally, this moiety is synthesized through the reaction of amino alcohols with several reagents including phosgene, dialkyl carbonates, urea, carbon dioxide, carbon monoxide, and oxygen.[38] Several notable methods have also been used for the synthesis process such as: the chiral resolution method, iodocyclocarbamation, The Heweh-Kauffman/Speranza-Peppel Method, and direct catalytic asymmetric amination of aldehydes.[30] Moreover, serine (Ser), threonine (Thr), and cysteine (Cys) have been used as building blocks for amino acid derived pseudo-prolines through the reaction of carbonyl containing reagents such as aldehydes and ketones.[36,37] Building upon this knowledge, we progressed to develop a rational design approach for the synthesis of cyclic urethane containing peptides.

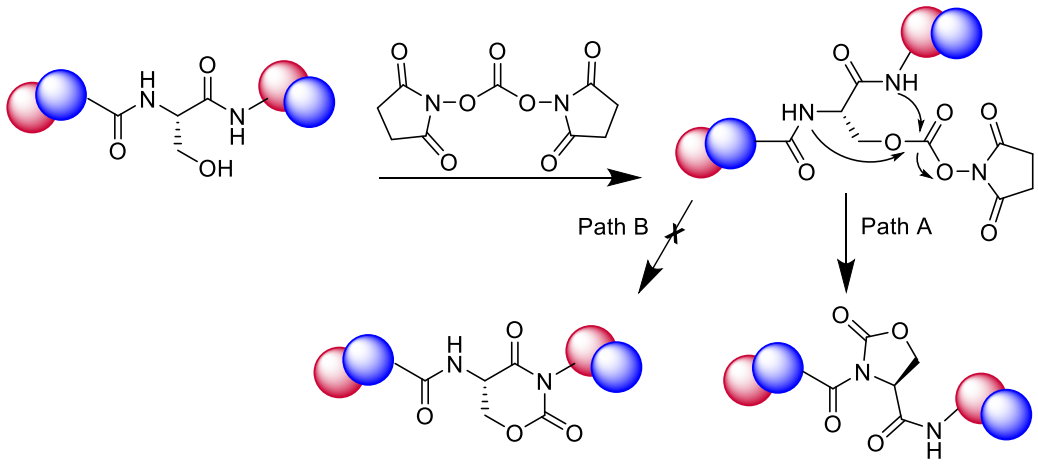
1.4.1 DESIGN OF CYCLIC URETHANE CONTAINING PEPTIDES

In the process of designing cyclic urethane containing peptides, we sought to peruse an approach where the cyclic urethane moiety would be derived from naturally occurring amino acids serine, threonine, and/or cysteine. Moreover, this approach would include modification of the amino acid residue(s) within the peptide sequence. Unique to our submonomer approach are a few factors: i) the cyclic urethane is to be derived from a natural amino acid, ii) the modification would take place within the peptide sequence rather than involving the prerequisite solution-phase synthesis of the cyclic urethane modified amino acid building block followed by incorporation into the peptide, and iii) modification to the desired moiety does not require a special protecting group or yield the dehydrated product as previously reported.[39]

We proceeded to rationalize the reaction mechanism. In this case, Ser was selected as the reactive amino acid residue. Based on Ser's nucleophilic side chain, we sought to activate the hydroxymethyl side chain to promote nucleophilic attack on a carbonyl containing compound to form an activated intermediate with a good leaving group (Scheme 1.6). Subsequently, intramolecular nucleophilic attack of the amide nitrogen of the peptide backbone on the intermediate could proceed to produce a five-membered cyclic urethane ring or a six-membered ring. Based on proximity of the amidic nitrogen, we hypothesized that the five-membered cyclic urethane would be the more kinetically favorable product, which would prove to be more resourceful in its variety of application.[40]

1.4.2 SYNTHESIS

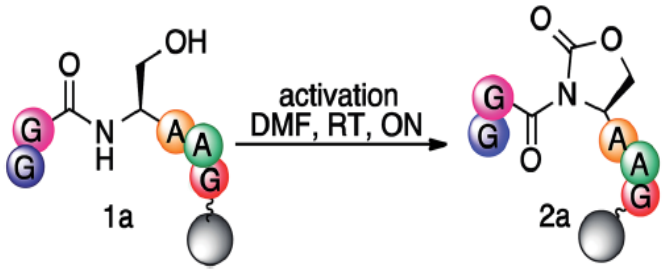
Once the rational resign was in place, we proceeded to execute the methodology experimentally. The focus of our design was to form the cyclic urethane moiety on a solid support post peptide synthesis. First, we screened carbonyl donating agents to determine the most effective reagent in the synthesis of cyclic urethane.[41]



Scheme 1.6 Serine modification mechanism.

For activation of serine's side chain, various electrophilic reagents such as 4-nitrophenyl chloroformate, 1,1-carbonyldiimidazole (CDI), and N, N-disuccinimidyl carbonate (DSC) were screened and quantitative conversion of Ac-GGSAAG to the cyclic urethane containing version was observed by HPLC. Further confirmation of the modification by molar mass was obtained by mass spectrometry. With a conversion percentage of approximately 99%, DSC was determined to be the reagent of choice for efficient conversion of Ser to cyclic urethane.

Table 1.1 Reagent screening for the backbone activation of serine[a]



Entry	Reagent	Additive	Conv.[b] (%)
1	4-Nitrophenyl chloroformate	—	0
2	4-Nitrophenyl chloroformate	DMAP	20
3	CDI	DMAP	60
4	DSC	DMAP	99

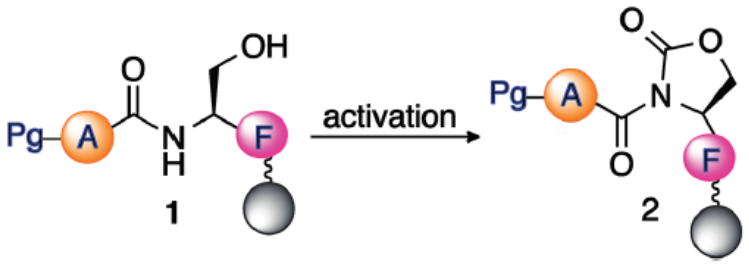
[a] Reaction conditions: to peptide Ac-GGSAAG **1a** on solid support (25 mg, 0.52 mm g^{-1}), a solution of 4-nitrophenylchloroformate/CDI/DSC (10 equiv.), DIEA (10 equiv.) and DMAP (catalytic amount) in DMF (3 mL) was added and the resin was left on a shaker overnight (ON) at room temperature (RT). [b] Conversion to Ac-GG*Oxd*AAG **2a** was calculated from the absorbance at 220 nm using HPLC.

Table adopted from reference (41) with permission.

Following reagent screening, formation of the cyclic urethane was further optimized on solid support by exploring other reaction conditions including solvent, time, temperature, and stoichiometry. A low conversion of 40% was observed in DCM (dichloromethane) due to poor solubility of DSC. When a mixture of DCM: DMF (dimethylformamide) (1:1) was screened as a solvent, the conversion increased by 20%. Meanwhile, an increased solubility in DMF led to a

conversion increase to 99%. Moreover, the role of N-terminal protecting groups was explored in the formation of cyclic urethane. Groups such as acetyl, tosyl, and Boc were investigated as to determine whether they would influence the efficiency of modification. Contrary to previously reported literature, the results indicated that formation of cyclic urethane is independent of the nature of the N-terminal protecting group.[39]

Table 1.2 Optimization of reaction conditions for cyclic urethane formation on solid support[a]



Entry	Pg	DSC (equiv.)	Solvent	Time (h)	Conv.[b] (%)
1	**1b** Fmoc	10	DCM	17	40
2	**1b** Fmoc	10	DCM : DMF	17	60
3	**1b Fmoc**	**10**	**DMF**	17	>99
4	**1b** Fmoc	5	DMF	17	70
5	**1b** Fmoc	10	DMF	3	30
6	**1b** Fmoc	10	DMF	7	90
7	**1c** Ac	10	DMF	17	>99
8	**1d** tosyl	10	DMF	17	>99
9	**1e** Boc	10	DMF	17	>99

[a] Reaction conditions: peptide **1** (25 mg, 0.87 mm g^{-1}) on solid support was reacted with DSC (5–10 equiv.), DIEA (5–10 equiv.), and a crystal of DMAP in different solvents (3 mL) at room temperature for 3–17 h. [b] Conversion to cyclic urethane moiety **2** was calculated from the absorbance at 220 nm using HPLC.

[a]Table adopted from reference (41) with permission.

Following optimization studies on solid support, varying amounts of DSC were explored on Fmoc-GASFAG post-cleavage from the solid support to determine the equivalents of reagent required for solution-phase modification.[40] Ten equivalents of DSC produced an 80% yield, while twenty equivalents yielded 90% modified peptide.

Table 1.3 Optimization of reaction conditions for cyclic urethane formation in solution[a]

Entry	Equivalents	(%) Yield[b]
1	5	75
2	10	80
3	20	90

[a]Reaction conditions: peptide (1 equiv.), DSC (5-20), DIEA (5-20 equiv.) and a crystal of DMAP in DMF were stirred overnight at room temperature. Purified by HPLC. [b]Isolated yield.

1.4.3 NMR CHARACTERIZATION

Upon optimization of the synthesis conditions, NMR characterization was performed to confirm the formation of the five-membered cyclic urethane ring rather than the six-membered ring on the sequence Fmoc-GS (Figure 1.8 and 1.9).[40]

The key differences in chemical shift that confirm the formation of the five-membered cyclic urethane ring include: 1) the methine carbon has a predicted ^{13}C shift of 55.5 ppm for the cyclic urethane ring and the observed chemical shift value was 55.4 ppm 2) HMBS correlation from both NH_2 protons to the neighboring carbonyl carbon indicate that the amino group is free and not involved in ring formation 3) the methine carbon would have a ^{13}C shift value of 51.6 ppm if the six-membered ring had formed.

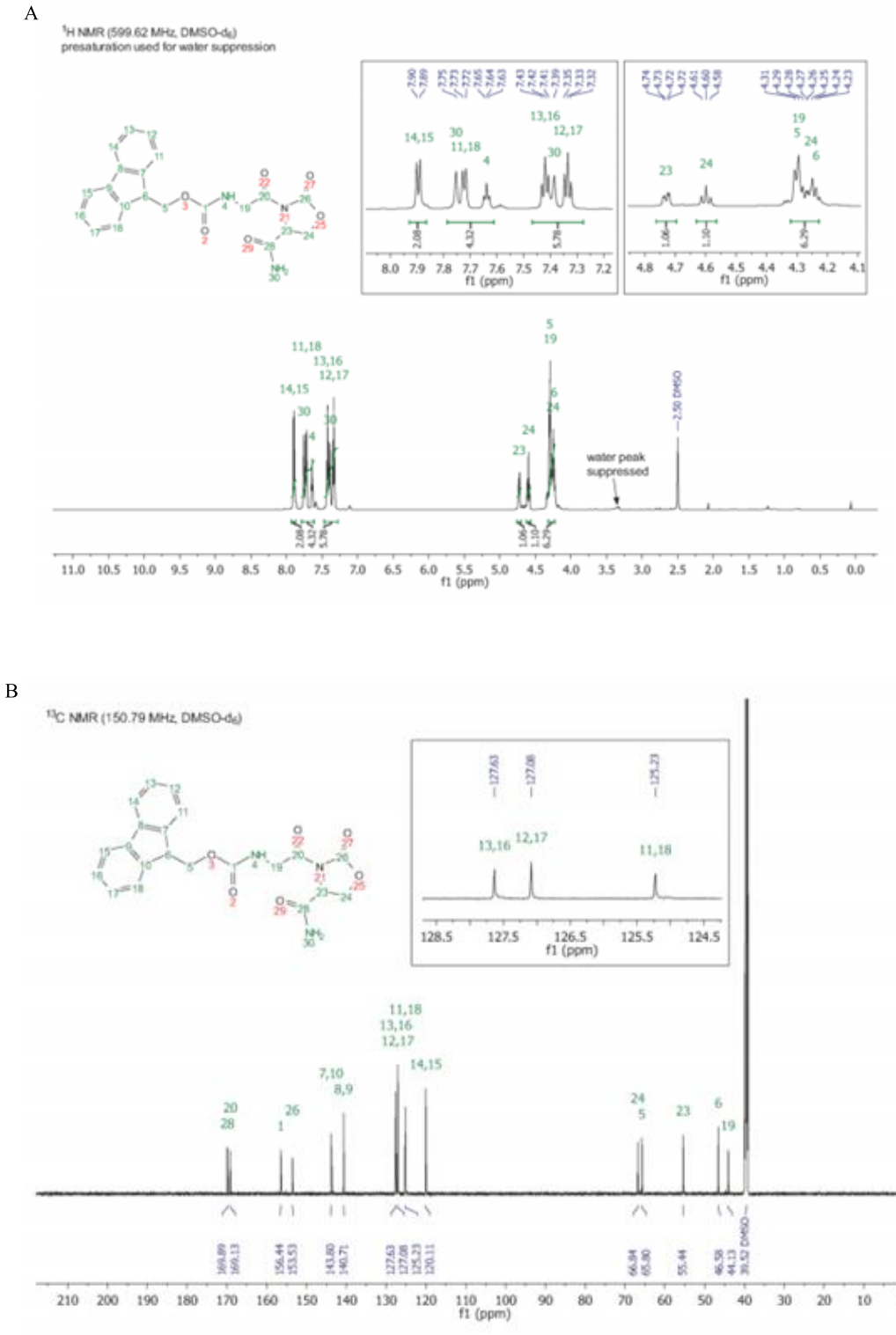


Figure 1.8 NMR spectra for modified Fmoc-GS. (A) ^{1}H NMR (B) ^{13}C NMR. Figure adopted from reference (40) with permission.

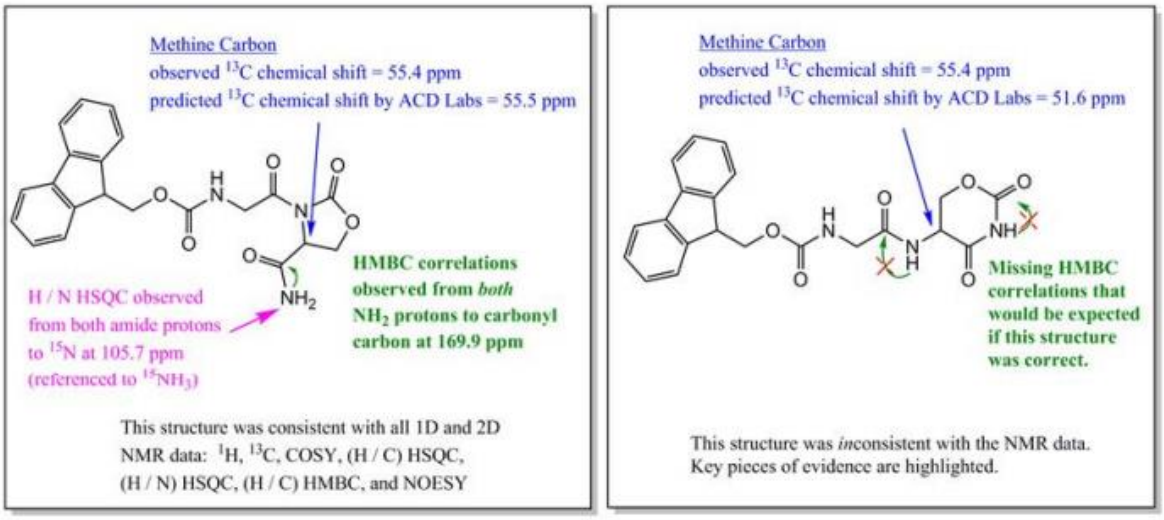


Figure 1.9 Comparison of NMR chemical shifts for cyclization of Ser to a five-membered ring versus a six-membered ring. Figure adopted from reference (40).

1.5 THESIS OBJECTIVES

The growing number of applications related to the modification of the amino acid side chains in a submonomer approach, *i.e.* directly within the peptide sequence, served as the framework for this thesis work.[42,43] Pursuing Ser as the model amino acid for the formation of the reactive cyclic urethane moiety results in a labile C-N bond which renders it susceptible to cleavage. Building on this principle, we sought to use the susceptibility of this moiety to cleavage and fulfill the following objectives: (1) use cyclic urethane as an enzyme mimic, thereby providing a new chemical cleavage methodology that is applicable to natural peptide sequences as well as mutated sequences that tend to be highly resistant to proteolysis, (2) use the cyclic urethane as a chemical cleavage method for cyclic peptides and uniquely structured lasso peptides which cannot be readily sequenced by MS/MS due to complex fragmentation patterns, (3) use the cyclic urethane moiety as a means for sequencing cyclic peptide libraries (4) use the cyclic urethane to make C-terminally modified peptides with various functional groups from the Rink amide resin and (5) use the cyclic

urethane to make the challenging peptide thioester precursor obtained from solid support through Fmoc synthesis which would lend to important applications in the native chemical ligation of synthetic peptides into chemically derived proteins.

1.6 REFERENCES

1) Russell, P. J. *iGenetics: A Molecular Approach Third Edition*, 3rd ed.; Pearson Education, 2010.

2) Esquivel, R.; Molina-Espíritu, M. In *Advances in Quantum Mechanics*; INTECH, 2013; pp 641–669.

3) Fosgerau, K.; Hoffmann, T. *Drug Discov. Today* **2015**, *20*, 122.

4) Uhlig, T.; Kyprianou, T.; Martinelli, F. G.; Oppici, C. A.; Heiligers, D.; Hills, D.; Calvo, X. R.; Verhaert, P. *EuPA Open Proteomics* **2014**, *4*, 58.

5) Campbell, M. K.; Farrell, S. O. *Biochemistry*, 7th ed.; Brooks/Cole, Cengage Learning, 2012.

6) Stawikowski, M.; Fields, G. B. *Curr. Protoc. Protein Sci.* **2002**, *26*, 1.

7) Fischer, E.; Fourneau, E. *Berichte der deutschen chemischen Gesellschaft* **1901**, *34*, 2868.

8) Bergmann, M.; Zervas, L. *Berichte der deutschen chemischen Gesellschaft* **1932**, *65*, 1192.

9) Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc.* **1953**, *75*, 4879.

10) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149.

11) Merrifield, R. B. *Recent Prog. Horm. Res.* **1967,** *23*, 451.

12) Sakakibara, S.; Shimonishi, Y.; Kishida, Y.; Okada, M.; Sugihara, H *Bull Chem. Soc. Jpn.* **1967**, *40*, 2164.

13) Carpino, L. A.; Han, G. Y. *J. Am. Chem. Soc.* **1970**, *92*, 5748.

14) Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1979**, *44*, 3739.

15) Fields, G. B.; Simon, M. I.; Abelson, J. N. *Methods in Enzymology*; Academic Press, 1997; Vol. 448.

16) Yan, B. *Comb. Chem. High Throughput Screen.* **1998**, *1*, 215.

17) Guillier, F.; Orain, D.; Bradley, M. *Chem. Rev.* **2000**, *100*, 3859.

18) Wang, S. *J. Am. Chem. Soc.* **1973**, *95* (4), 1328.

19) Rink, H. *Tetrahedron Lett.* **1987**, *28* (33), 3787.

20) Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827.

21) Rich DH, Singh J. The carbodiimide method. In: Gross E, Meienhofer J, editors. The Peptides. Vol. 1. Academic Press; New York: 1979. pp. 241–314

22) Vagner, J.; Qu, H.; Hruby, V. J. *Curr. Opin. Chem.* **2008**, *12*, 292.

23) Gokhale, A. S.; Satyanarayanajois, S. *Immunotherapy* **2014**, *6*, 755.

24) Checco, J. W.; Lee, E. F.; Evangelista, M.; Sleebs, N. J.; Rogers, K.; Pettikiriarachchi, A.; Kershaw, N. J.; Eddinger, G. A.; Belair, D. G.; Wilson, J. L.; Eller, C. H.; Raines, R. T.; Murphy, W. L.; Smith, B. J.; Gellman, S. H.; Fairlie, W. D. *J. Am. Chem. Soc.* **2015**, *137* (35), 11365.

25) White, C. J.; Yudin, A. K. *Nat. Chem*. **2011**, *3*, 509.

26) Tofteng, A. P.; Sørensen, K. K.; Conde-Frieboes, K. W.; Hoeg-Jensen, T.; Jensen, K. J. *Angew. Chemie - Int. Ed.* **2009**, *48* (40), 7411.

27) Tanabe, K.; Taniguchi, A.; Matsumoto, T.; Oisaki, K.; Sohma, Y.; Kannai, M. *Chem. Sci.* **2014**, *5*, 2747.

28) Gross, E. *Methods Enzymol*. **1967**, *11*, 238.

29) Simpson, L. S.; Kodadek, T. *Tetrahedron Lett.* **2012**, *53*, 2341.

30) Pandit, N.; Singla, R. K.; Shrivastava, B. *Int. J. Med. Chem.* **2012**, *2012*, 1.

31) Houben-Weyl. *Methods of Organic Chemistry*, 4th ed.; Thieme, 1995.

32) Evans, D. A.; Wu, L. D.; Wiener, J. J. M.; Johnson, J. S.; Ripin, D. H. B.; Tedrow, J. S. *J. Org. Chem.* **1999**, *64* (17), 6411.

33) Diaz, G.; De Freitas, M. A. A.; Ricci-Silva, M. E.; Diaz, M. A. N. *Molecules* **2014**, *19* (6), 7429.

34) Evans, D. A.; Helmchen, G. *Asymmetric Synth. Essentials* **2007**, 3.

35) Wöhr, T.; Mutter, M. *Tetrahedron Lett.* **1995**, *36*, 3847

36) Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218.

37) Dumy, P.; Keller, M.; Ryan, D. E.; Rohwedder, B.; Wöhr, T.; Mutter, M. *J. Am. Chem. Soc.* **1997**, *119*, 918.

38) Fujita, S. I.; Kanamaru, H.; Senboku, H.; Arai, M. *Int. J. Mol. Sci.* **2006**, *7*, 438.

39) De Marco, R.; Tolomello, A.; Campitello, M.; Rubini, P.; Gentilucci, L. *Org. Biomol. Chem.* **2012**. *10*. *2307.*

40) Elashal, H. E.; Raj, M. *Chem. Commun.* **2016**, *52*, 6304.

41) Elashal, H. E.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, *8*. 117.

42) Sabatino, D.; Proulx, C.; Klocek, S.; Bourguet, C. B.; Boeglin, D.; Ong, H.; Lubell, W. D. *Org. Lett.* **2009**, *11* (16), 3650.

43) Zuckermann, R. N.; Kerr, J. M.; Moosf, W. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **1992**, *114* (26), 10646.

CHAPTER 2: SITE-SELECTIVE CHEMICAL CLEAVAGE OF PEPTIDE BONDS

2.1 ABSTRACT

The cleavage of peptides and proteins is an essential tool used in bioanalytical, biotherapeutic, and bioengineering techniques. While enzymatic digestion is the preferred method of protein analysis, sequences which contain mutations and/or modifications deviate from nature's design for a native peptide/protein and tend to be resistant towards proteolysis. Therefore, there is a growing need for chemical cleavage methods which can analyze peptidomimetics that are unable to be recognized and processed by proteases. More recently, several methods have emerged which allow for site-selective cleavage of extremely unreactive bonds, including amides, thereby providing invaluable information regarding protein sequence, structure, and function. For controlled and selective cleavage of peptide bonds, chemical cleavage must selectively bind to one or more amino acid residues within a peptide chain and selectively cleave the peptide bond. Building on this principle, we have developed an approach that utilizes DSC to selectively modify Ser, Thr, Cys, and/or Glu, rendering them susceptible to peptide bond fragmentation. Subsequent incubation in buffer leads to the hydrolysis of the peptide bond at the N-terminal side of the modified amino acid. This methodology lends itself to peptide sequencing applications by exhibiting broad substrate scope in the cleavage of various bioactive peptides with post-translational modifications (e.g. N-acetylation and N-methylation) and mutations (D- and β- amino acids) which have been implicated in age related diseases.

2.2 CHAPTER OBJECTIVES

In this chapter, the development of an artificial chemical peptidase for site-selective cleavage of unreactive peptide bonds at serine, threonine, and cysteine residues is discussed. The site-selective cleavage of a peptide bond is a crucial procedure in protein sequencing and various new bioanalytical and bioengineering applications. Peptide bonds, however, are extremely stable towards hydrolysis, with a half-life of 500-1000 years at room temperature and pH 4-8. This limits the range of appropriate cleavage reagents, especially those that may trigger selective peptide bond hydrolysis at specified amino acid residues. Commonly used reagents include proteases, self-cleaving intein sequences, metals, and chemical agents, such as cyanogen bromide, albeit with certain restrictions associated with poor selectivity, efficiency and limited substrate scope, especially related to fragmenting modified peptides or peptidomimetics. In order to address these limitations, *this strategy provides a method that selectively modifies serine, threonine, and cysteine residues in any peptide chain and cleaves the peptide chain at the site of modification under neutral aqueous buffer conditions.* The site-selective cleavage parlays the reaction of an electrophilic reagent with the nucleophilic hydroxymethyl group of serine/threonine and/or the thiol of cysteine to activate the backbone amide chain. In this chapter, the rationale for site-selective modification of these selected residues in a peptide chain, the potential for selective-cleavage under mild conditions, and the effectiveness of the methodology for mutated and post-translationally modified peptides is described. This artificial chemical peptidase provides a new scissile position, Pro-Ser peptide bonds, which are typically resistant to cleavage by most proteolytic enzymes. Site-selective modification and cleavage of peptides is essential for ongoing efforts to determine the structure activity relationship of various biologically relevant peptides/proteins in addition to their modified analogs.

2.3 GRAPHICAL ABSTRACT

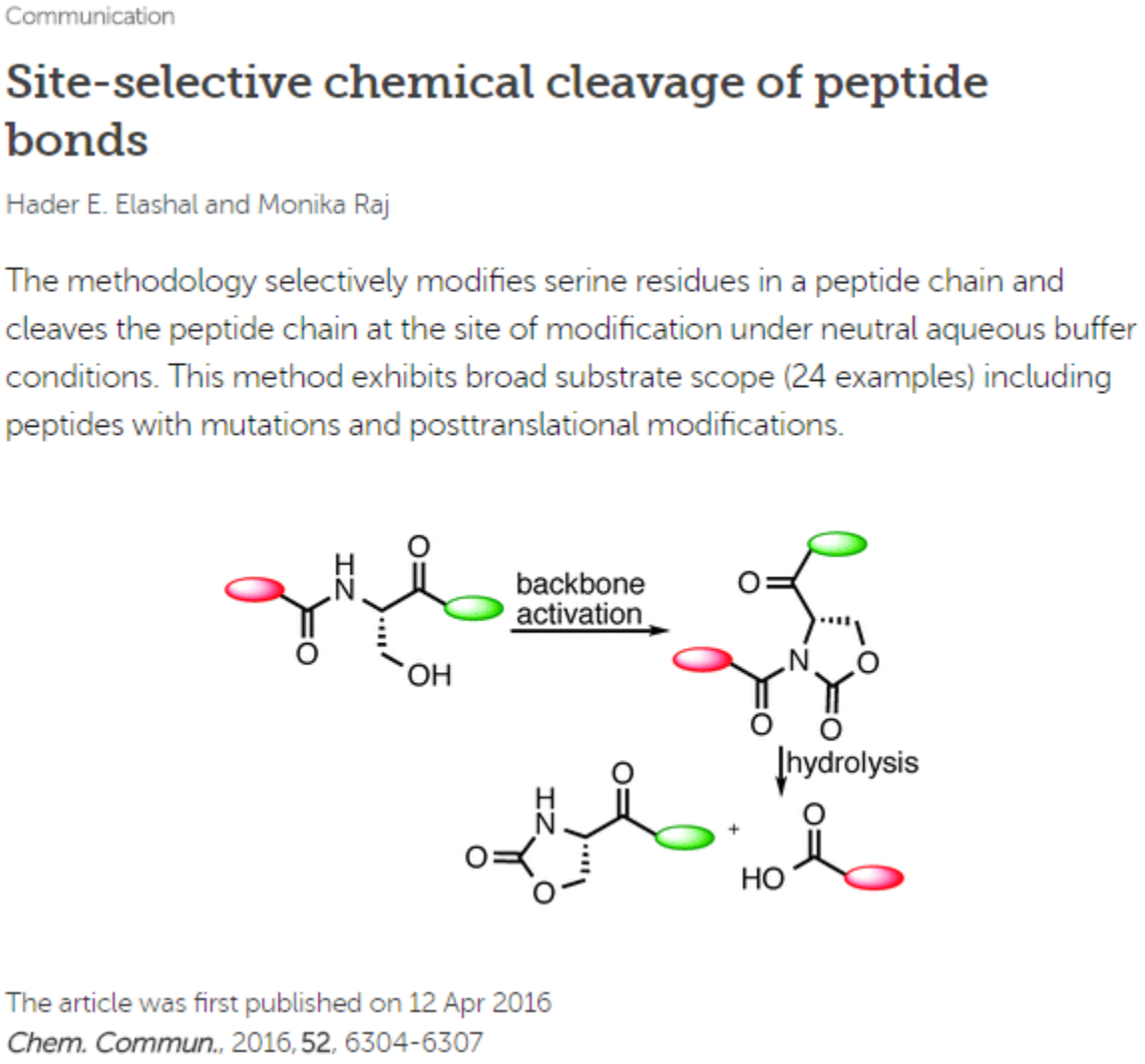


Figure 2.1 Graphical abstract for *Site-Selective Cleavage of Peptide Bonds*. Figure adopted with permission from: Elashal, H.; Raj, M. *Chem. Commun.* **2016,** *52*, 6304.[1]

2.4 INTRODUCTION

Site-selective cleavage of peptide bonds is an important reaction that has wide spread application in the vast field of proteomics, including sequencing, therapeutics, and the development of new bioanalytical techniques.[2,3] Enzymatic digestion is one of the most prominent methods of metabolizing proteins to their basic components in order to facilitate its sequence analysis.[4] Proteases that are commonly employed in protein hydrolysis include: trypsin which

cleaves at charged residues (Arg and Lys), chymotrypsin which cleaves at aromatic residues (Phe, Trp, and Tyr) in addition to Leu and Met, and pepsin which cleaves at residues with bulky side chains (Leu and Phe). While proteases cleave proteins with high specificity, precision, and fidelity, they have several disadvantages. Proteases have been known to contaminate the protein/peptide digestion and they require specific ranges of temperature and pH for optimal activity or enzyme denaturation can limit performance. Moreover, they lack efficacy when applied to diseased, post-translationally modified, and unnatural sequences, where their utility is greatly needed to facilitate the understanding of the onset of disease and the body's purpose for modifying proteins from their native state.

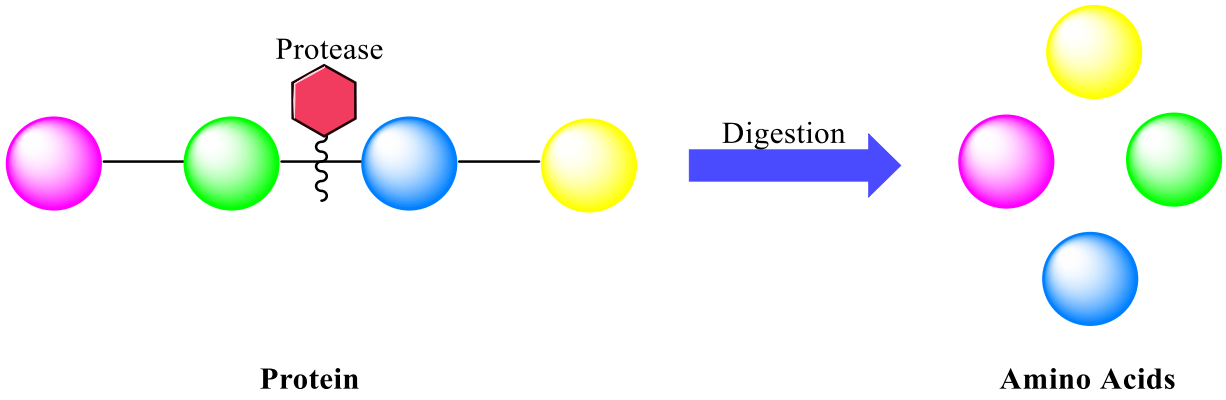
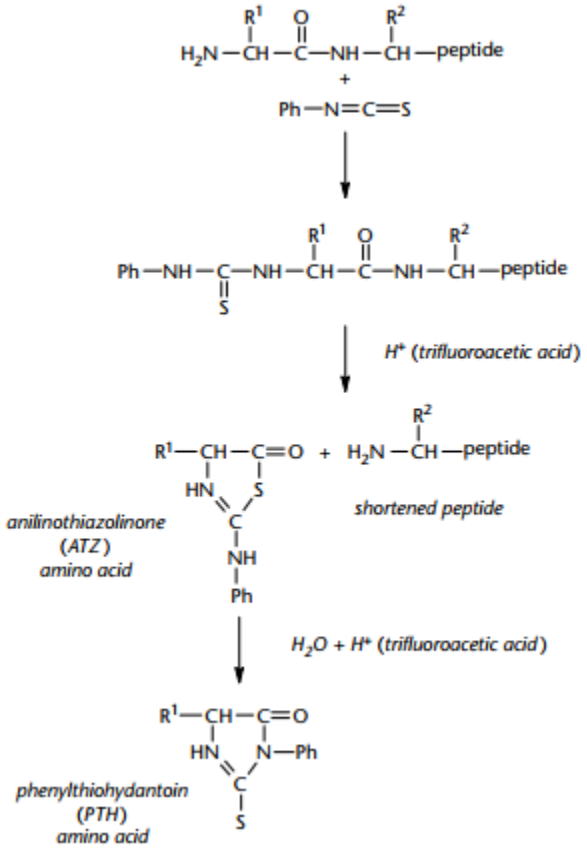


Figure 2.2 Protease digestion of a protein.

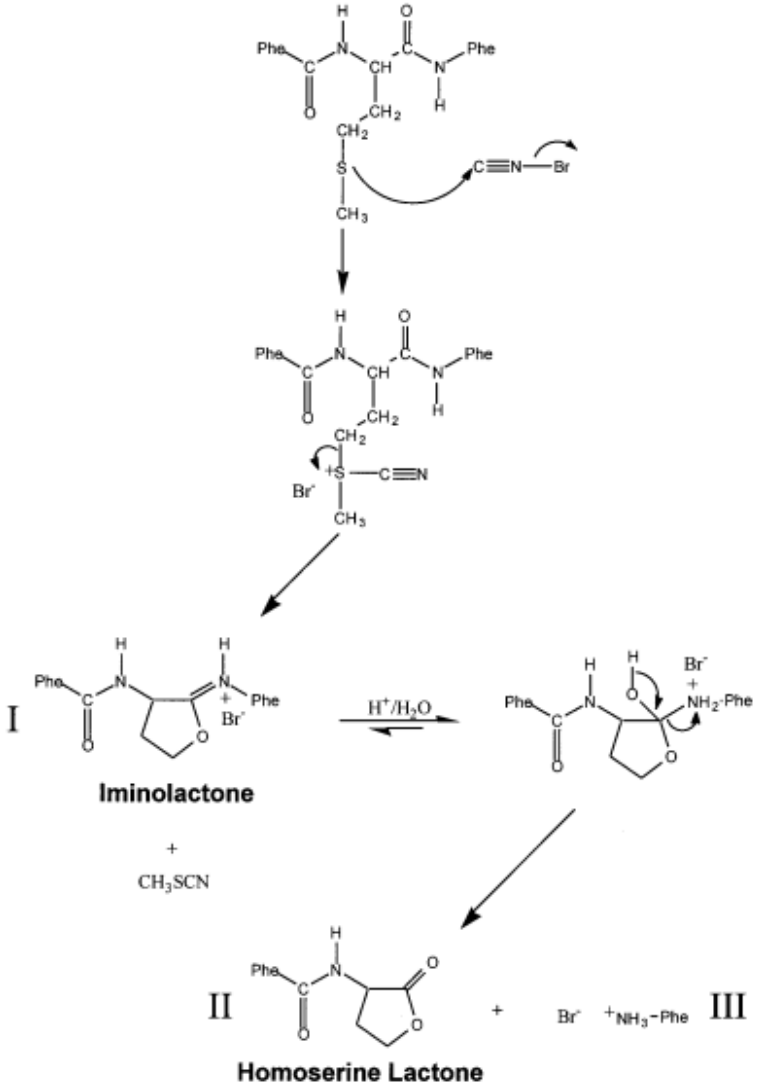
Owing to this limitation, efficient chemical cleavage methods have been developed as a way of offsetting the limitations associated with enzymes. The most commonly utilized chemical cleavage method is Edman degradation which gained wide spread popularity due to its broad applicability.[5] This technique utilizes phenyl isothiocyanate for N-terminal cleavage of peptides without disrupting other peptide bonds within the peptide sequence. The cleavage process proceeds under mild alkaline conditions; phenyl isothiocyanate is reacted with an uncharged N-terminal amino group to form a cyclized phenylthiocarbamyl derivatized moiety (Scheme 2.1).[6]

Upon the addition of acid, this N-terminal moiety is cleaved as a thiazolinone derivative which is extracted into organic solvent. Once treated with acid, the stable phenylthiohydantoin can be identified by utilizing chromatographic techniques or gel electrophoresis.[6] While this process can be executed repeatedly to identify the N-terminal residues in the protein chain, the efficiency of this methodology diminishes after 30 cycles of cleavage. Moreover, the process of determining each amino acid can be tedious and it lacks selectivity. Similar methods for N-terminal amino acid determination include Sanger's method, which has been widely applied in protein sequencing and ultimately resulted in the Nobel prize awarded to Frederick Sanger in 1958.[7,8]



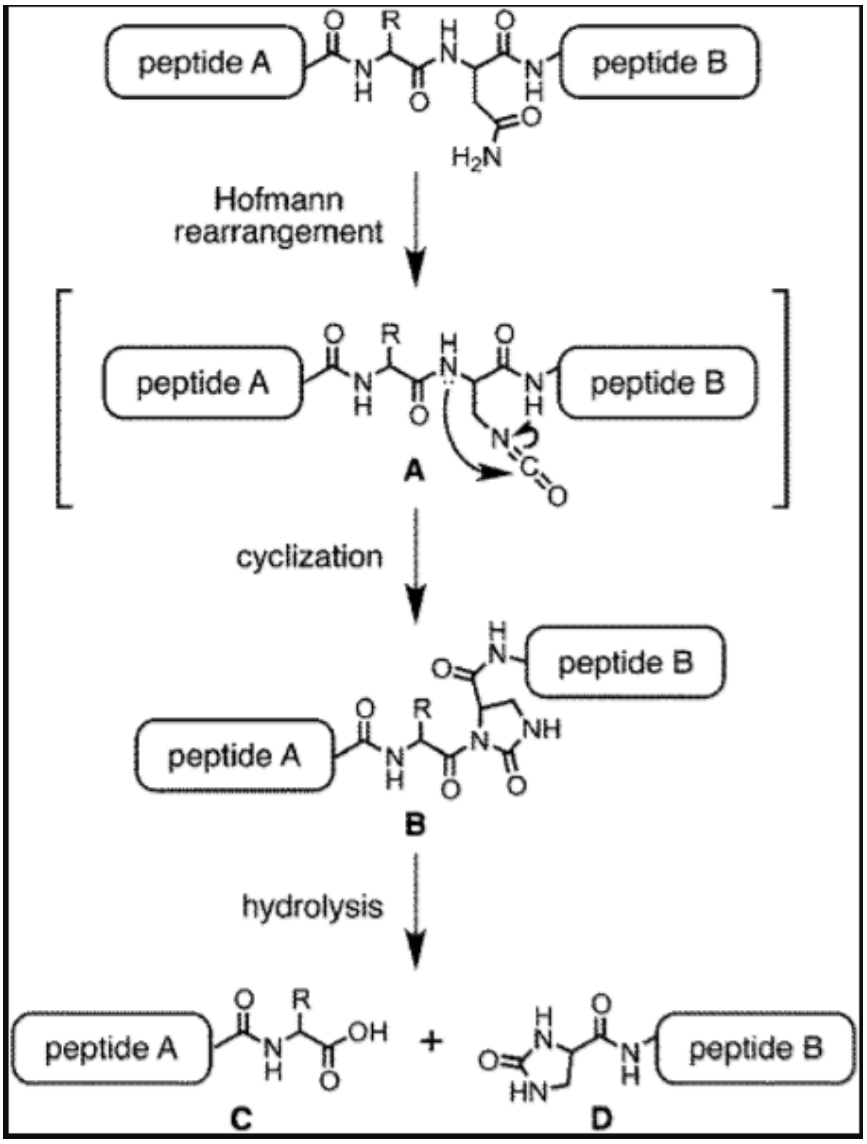
Scheme 2.1 Edman degradation. Scheme adopted from reference (6).

More recently, chemical cleavage methodologies have been developed based on amino acid side chain modification. These methodologies provide selectivity at the point of cleavage and they are uncomplicated in their execution, thereby facilitating the cleavage process and analysis of unknown sequences and/or proteolytically resistant sequences. Currently utilized selective cleavage methods include: cyanogen bromide for Met (Scheme 2.2),[9] 2-nitro-5-thio-cyanobenzoic acid,[10] and dehydroalanine for Cys,[11,12] and 2-iodosylbenzoic acid for Trp and Tyr.[13]



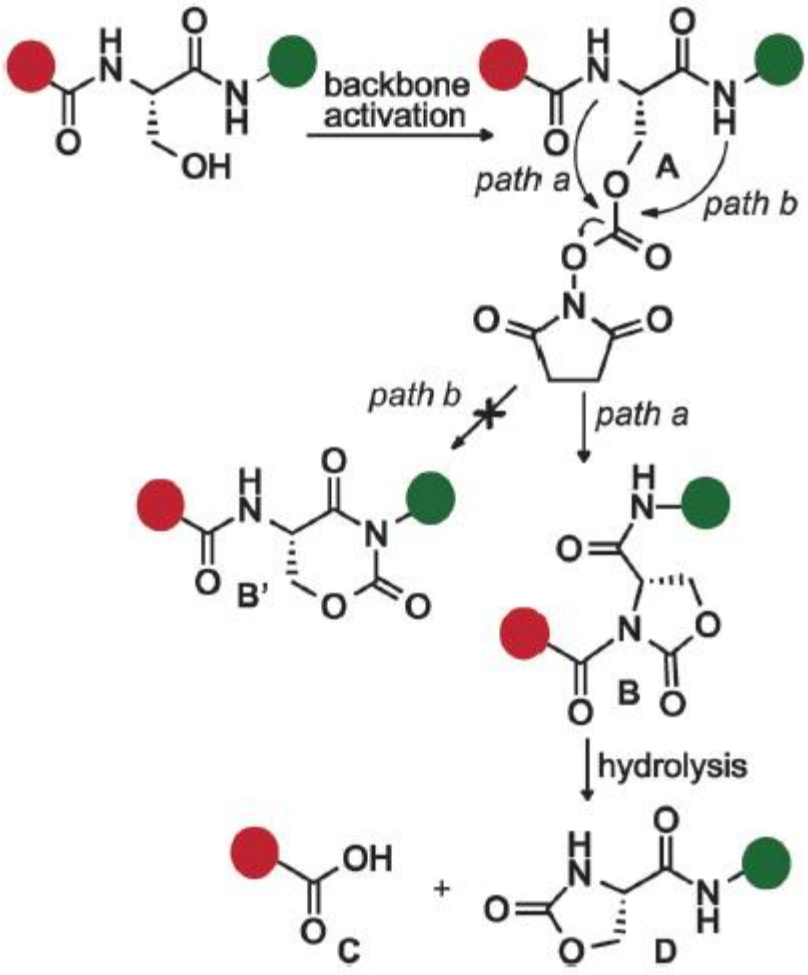
Scheme 2.2 Methionine mediated cleavage of peptide bonds. Scheme adopted with permission from: Kaiser, R.; Metzka, L. *Anal. Biochem.* **1999**, *266*, 1.[14]

While these techniques have proven complementary to proteases, they are limited by: inadequate selectivity, harsh conditions and low yields.[15] In the case of cyanogen bromide, the methodology uses toxic conditions and has a limited scope due to side reactions in the presence of methionyl-serine and methionyl-threonine.[14] Most recently, diacetoxyiodobenzene (DIB) was reported for asparagine-selective cleavage, yet oxidative modifications were observed at other reactive amino acid residues such as Cys and Met (Scheme 2.3).[16] Metals have been intensively studied for hydrolysis of peptide bonds, but their practical use in structure determination of proteins is still in its infancy.[17-24] Even with their associated drawbacks, chemical cleavage methods are a necessity in the analysis and characterization of proteolytically resistant proteins/peptides. Chemical methods aid in cleaving large proteins into smaller fragments which may be analyzed by complementary cleavage methods. They provide information regarding the presence of a particular amino acid in the protein/peptide sequence, its quantity based on the successful execution of cleavage, and they allow for cleavage even in the presence of unnatural residues such as D an β amino acids.



Scheme 2.3 Asparagine mediated cleavage of peptide bonds. Scheme adopted with permission from reference (16).

Owing to the need for more efficient chemical cleavage methodologies that are applicable to a broader variety of amino acids and peptide sequences, we sought to develop a chemical cleavage method that utilizes a cyclic urethane. Chemically modifying the amide backbone chain at Ser to include a cyclic urethane moiety renders the neighboring scissile peptide bond weak and susceptible to hydrolysis under neutral, aqueous conditions (Scheme 2.4).



Scheme 2.4 Rationale for the serine-selective modification and peptide cleavage in neutral aqueous solution. Scheme adopted from reference (1) with permission.

2.5 RESULTS AND DISCUSSION

To achieve the cyclic urethane modification at Ser, reaction conditions were optimized using varying amounts of DSC, 4-dimethylaminopyridine (DMAP), and N,N-diisopropylethylamine (DIEA) on a model Fmoc-protected hexapeptide Fmoc-Gly-Ala-Ser-Phe-Ala-Gly [retention time (Rt)=16.9 min)]. Previous attempts of utilizing DSC to form this cyclic urethane moiety on Fmoc/Boc peptides lead to the formation dehydroalanine.[25] Using our optimized conditions[1], Ser cyclization proceeded in 17 hrs as indicated by the sharp peak **2a** with a Rt=17.9 min, which was

analyzed by mass spectrometry (MS) (Figure 2.3). Following cyclization, the modified peptide was incubated in phosphate buffer with a pH of 7.5 for 31 hrs, and the reaction was monitored by HPLC. After the incubation period, the peak at 17.9 min disappeared and two new peaks appeared which was indicative of peptide cleavage into two fragments. MS analysis of the peaks at 8.9 min and 18.6 min proved these peaks to be the products resulting from cleavage at the N-terminal side of Ser. The N-terminal fragment **3a** eluted at 18.6 min, while the modified C-terminal fragment **4** eluted at 8.9 min.

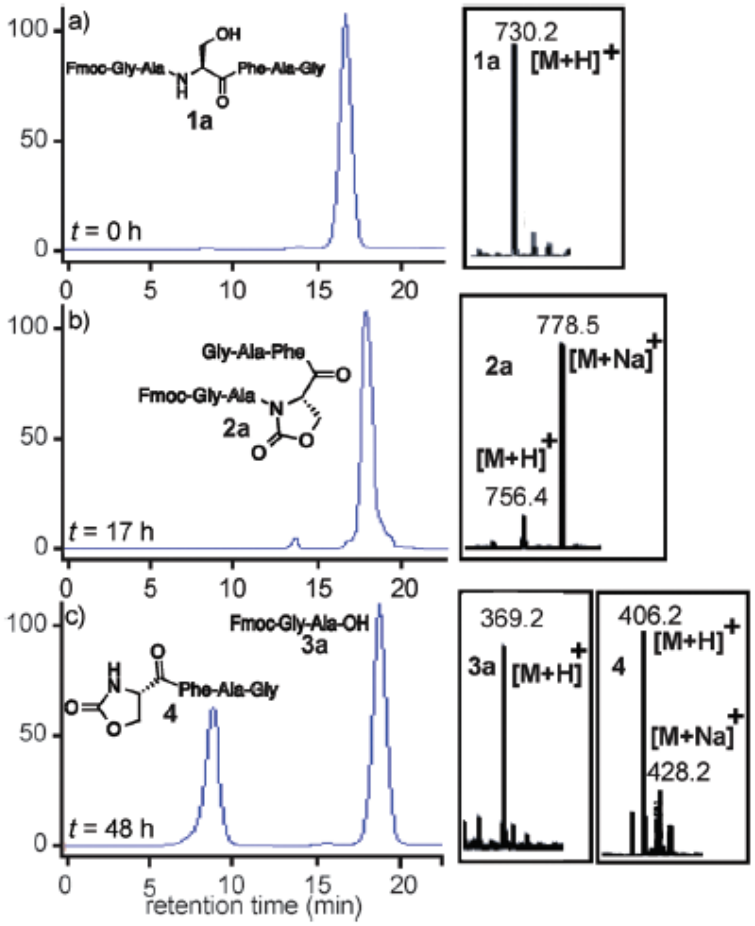


Figure 2.3 HPLC charts of the reaction mixture of 1a and DSC (10 equiv.), DIEA (10 equiv.), and a crystal of DMAP in DMF at (a) t = 0 h (top), (b) at t = 17 hrs (middle), and (c) after addition of 0.1 M phosphate buffer (pH 7.5) at t = 48 h(bottom) at 37 ºC. HPLC conditions: 0.1% FA (v/v) in water, 0.1% FA in CH_3CN, gradient: 0–80% in 25 min, 1.0 mL min_1 flow rate, detected at 254

nm. Insets show the LCMS spectrum corresponding to Rt at 16.9 min (1a, top), 17.9 min (2a, middle), 18.6 min (3a, bottom) and 8.9 min (4, bottom). Figure adopted from reference (1) with permission.

Following initial cyclization and cleavage of Ser, we sought to determine the effects of side chain functionality of the residue preceding Ser on the cyclization and cleavage efficiency of peptides. Moreover, we wanted to explore the effect of reaction conditions on unprotected side chains to determine if the reaction generated undesired side products. Using the sequence Fmoc-Gly-Xaa-Ser-Phe-Ala-Gly, various hexapeptides were synthesized to include different residues at the Xaa position (Table 2.1).

Table 2.1 Serine-selective amide bond cleavage of Fmoc-Gly-**Xaa**-Ser-Phe-Ala-Gly (1a-j)[a]

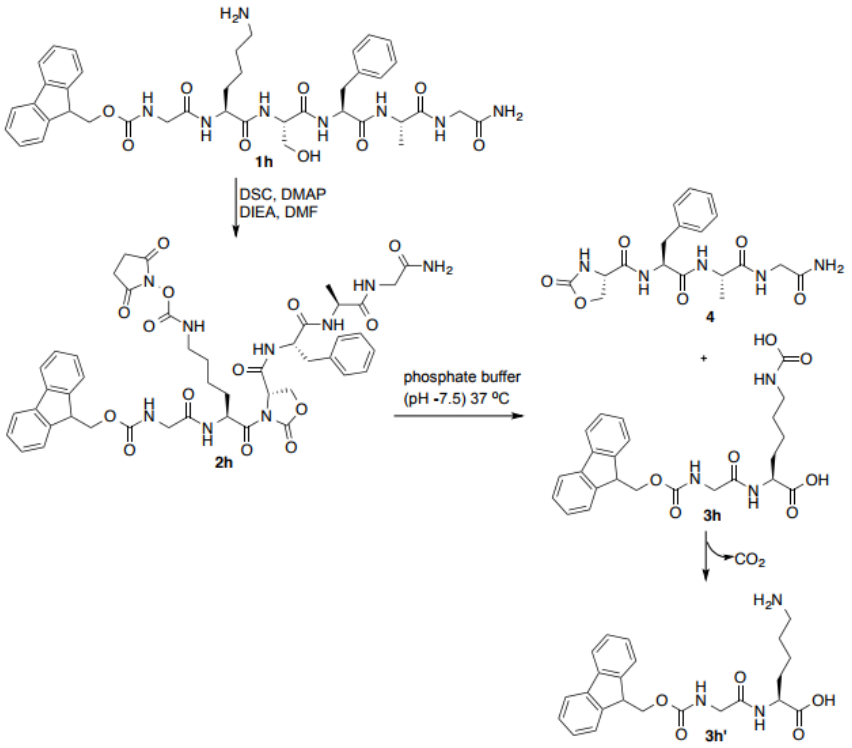
Entry	Substrate	Xaa	Conv.[b] (%)
1	**1a**	Ala	>99
2	**1b**	Gly	>99
3	**1c**	Met	98
4	**1d**	His	98
5	**1e**	Tyr	97
6	**1f**	Trp	97
7	**1g**[c]	Val	70
8	**1h**[d]	Lys	80
9	**1i**	Asp	95
10	**1j**	Pro	98

[a] Reaction conditions: peptide (1 equiv.) was reacted with DSC (5–20 equiv.), DIEA (5–20 equiv.) and a crystal of DMAP in DMF followed by cleavage with 0.1 M phosphate buffer (pH 7.5) at 37 °C for 48 h. [b] Conversion to N-terminal fragment, Fmoc-Gly-Xaa-OH (**3a–j**), was calculated from the absorbance at 254 nm using HPLC. [c] Cleavage with 0.1 M phosphate buffer (pH 7.8) at 37 °C for 4 days. [d] Acylation of the side chain of the lysine with DSC generated acylated lysine fragment **2h** and after buffer treatment, generated fragments **3h′** and **4**

Table 2.1 adopted from reference (1) with permission.

After incubation in buffer (pH=7.5) for 48 hrs, the C-terminal fragment Oxd-Phe-Ala-Gly **4** was obtained as a product post cleavage of all peptides with varying HPLC conversions ranging

from 70 to 99%. Peptides which included Ala **1a,** Gly **1b**, Met **1c**, His **1d**, Tyr **1e**, Trp **1f,** and Asp **1i** substituted at the Xaa position underwent conversion to the cleavage products after 48 hrs in a highly efficient manner (entries 1-6 and 9, Table 2.1). This conversion proceeds smoothly in contrast to other chemical cleavage methods which result in the over-oxidation of Tyr, Trp, and Met containing peptides. Thus, our method is tolerable of many side chain functional groups. Meanwhile, the peptide **1g** containing Valine adjacent to Ser gave moderate yield after four days in phosphate buffer (pH= 7.8) due to the presence of a bulky side chain (entry 7, Table). The peptide containing Xaa= Lys **1h** proceeded to modify in a unique manner due to the presence of a free amine in the Lys side chain (Scheme 2.5). The amine side chain also reacted with DSC; upon treatment with buffer, however, the desired Lys-Ser cleavage occurred, and an N-terminal carboxylate lysine derivative was generated **3h** in addition to Oxd-FAG-NH_2. Eventually, the N-terminal carboxylated lysine derivative underwent decarboxylation of the unstable carbamic acid to generate Fmoc-GK-OH (entry 8, Table 2.1) leaving the Lys residue unmodified. The final substitution included Proline **1j** in the position adjacent to Ser, and the cyclization and cleavage of the Pro-Ser bond proceeded with high efficiency. Typically, enzymatic digestion at a location neighboring to Pro is nearly unfeasible due to the cis amide bond rotation produced by Pro in the peptide backbone which makes the enzyme active site inaccessible.[26] However, in the case of utilizing our DSC-mediated chemical methodology, peptide-bond cleavage smoothly proceeded with a conversion of 98% which illustrated the efficiency of this technique regardless of the neighboring amino acid to Ser.



Scheme 2.5 Modification scheme of lysine containing peptide. Scheme adopted from reference (1) with permission.

Table 2.2 Substrate scope of site-selective cleavage of peptide bonds[a]

Entry	Substrate	Conv.[b] (%)
1	Fmoc-Gly-Ala-**Thr**-Phe-Arg-Phe-Gly-NH_2 (5)	78
2	Fmoc-Ala-**Ser**-Phe-Val-Gly-Ala-**Thr**-Phe-Arg-Phe-Gly-NH_2 (6)[c]	75
3	Fmoc-Ala-Val-Arg-**Ser**-Phe-**Ser**-Ala-Arg-Gly-Phe-Gly-NH_2 (7)[c]	75
4	Fmoc-Arg-Ala-Gly-Ala-**Ser**-Val-Arg-Phe-Ala-**Ser**-Phe-Gly-NH_2 (8)[c]	79
5	Fmoc-D-Val-D-Arg-D-Lys-D-Ala-D-**Ser**-D-Arg-D-Ala-D-Ala-NH_2 (9)	80
6	Fmoc-D-Val-D-Arg-D-Lys-D-Ala-L-**Ser**-D-Arg-D-Ala-D-Ala-NH_2 (10)	82
7	Fmoc-Cys-Gly-Arg-Arg-Ala-Cys-Gly-**Ser**-Phe-Ala-Gly-NH_2, disulfide bond (11)	72
8	Fmoc-Arg-Ala-**Glu**-Ala-Gly-**Ser**-Gly-Phe-NH_2 (12)[d]	70

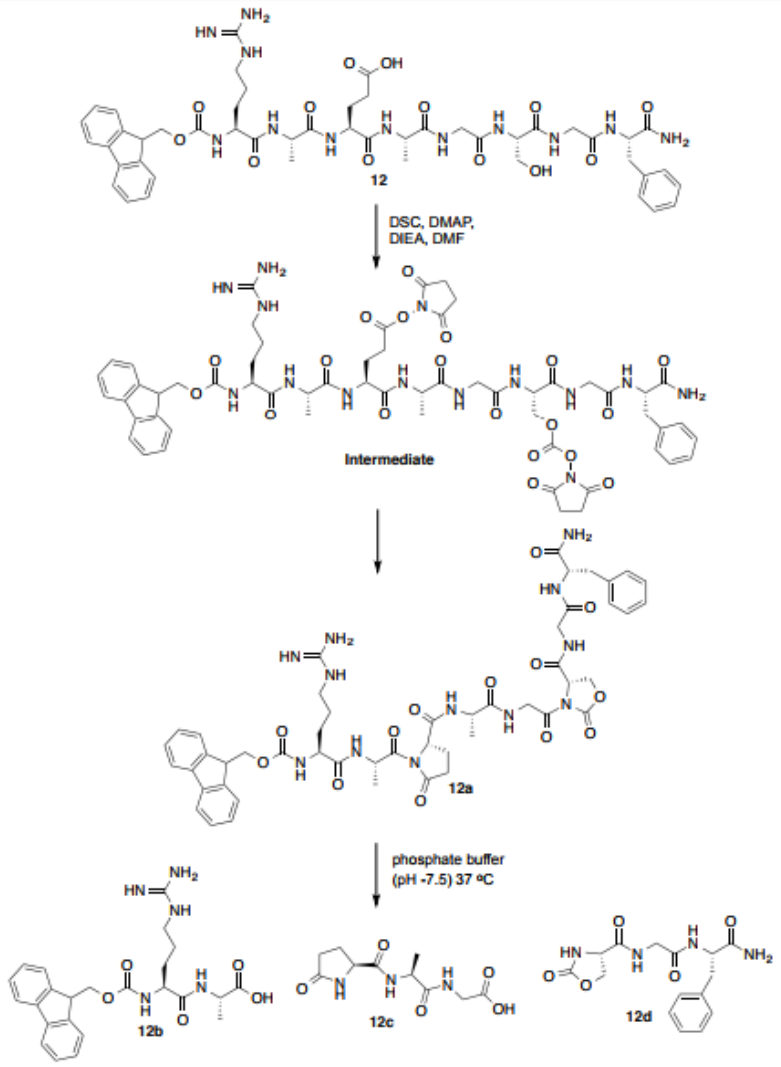
[a] Reaction conditions: peptide (1 equiv.) was reacted with DSC (20 equiv.), DIEA (20 equiv.), and a crystal of DMAP in DMF followed by cleavage with 0.1 M phosphate buffer (pH 7.5) at 37 °C for 48 h unless otherwise noted. [b] N-terminal fragment. [c] Three fragments were detected in the HPLC trace for cleavage at Ser and Thr at both the Ser residues. [d] Three fragments were detected in the HPLC trace for cleavage at both Ser and Glu.

Table 2.2 adopted from reference (1) with permission.

The substrate scope of the reaction was further evaluated (Table 2.2). Modification of Thr was also feasible as demonstrated by entries 1 and 2 (Table 2.2). After incubation with buffer, cleavage was observed at the N-terminal side of Ser and Thr, which is expected due to the similar

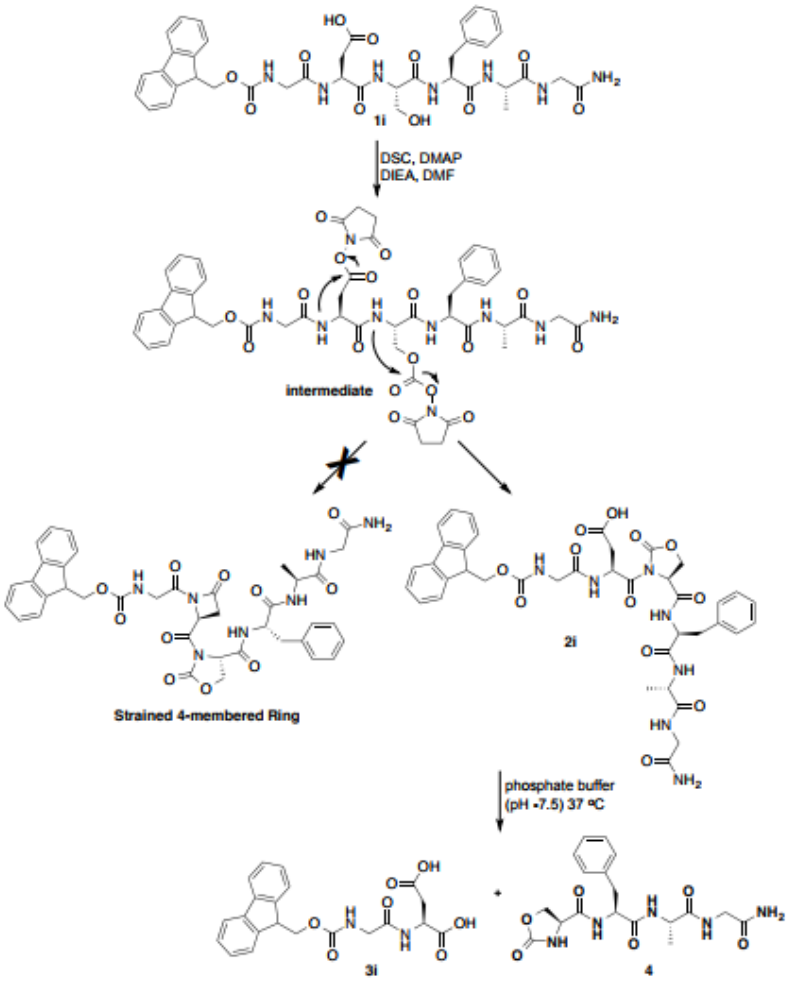
side chain functionality exhibited by both residues, thereby mimicking enzyme specificity towards similar substrates. However, cyclization of Threonine's side chain required a higher amount of DSC, which may be due to the presence of a methyl group neighboring to the hydroxyl group (entry 2, Table 2.2). Peptide Fmoc-GACFRFG-NH_2 also underwent modification and cleavage at the N-terminus of Cys and generated a thiazolidinone modified C-terminal fragment Thz-FRFG-NH_2 and the N-terminal fragment Fmoc-GA-OH. Peptides with longer lengths as well as multiple Ser residues gave comparable results (75-79%) illustrating the utility of this methodology of cleaving lengthy peptides into smaller fragments that are more easily analyzed (entries 3 and 4, Table 2.2). Meanwhile, peptides that contained D-amino acid residues which are generally resistant to enzymatic digestions due to their unnatural confirmation, were successfully cleaved under the mild reaction conditions (entries 5 and 6, Table 2.2). Additionally, peptide **11** comprising of an intramolecular disulfide bridge afforded the cleavage product without disrupting the disulfide bond (entry 7, Table 2.2). Therefore, this methodology allows for the disulfide mapping, which is an advantage over other chemical methodologies which tend to break the disulfide bond.[27]

While exploring the substrate scope in a Glu containing peptide **12**, surprisingly, cleavage occurred both at the N-terminal side of Glu and Ser to generate three fragments under the reaction conditions (entry 8, Table 2.2). The activation of Glu by DSC proceeds through the formation of a five-membered pyro-glutamyl imide moiety (pGlu) (Scheme 2.6).



Scheme 2.6 Modification scheme of glutamic acid to a pyro-glutamyl moiety. Scheme adopted from reference (1) with permission.

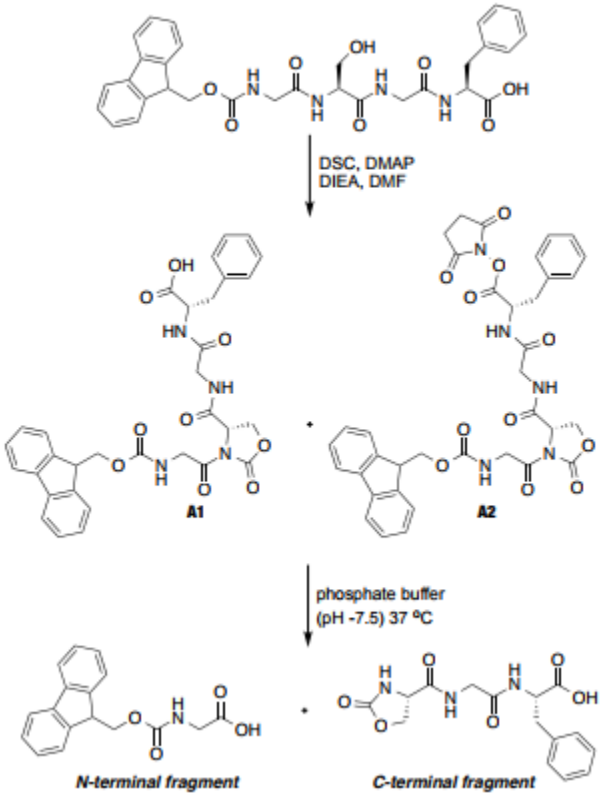
In the presence of Asp **1i**, while also containing a free carboxylic acid side chain, modification and cleavage is not observed due to side chain length (Scheme 2.7).



Scheme 2.7 Modification of Asp side chain vs. Glu side chain. Scheme adopted from reference (1) with permission.

Next, compatibility with a free main chain carboxylic group was explored by synthesizing Fmoc-GSGF-OH on Wang resin. Under the reaction conditions, cleavage was observed only at the N-terminus of Ser and generated an N-terminal Fmoc-Gly fragment and Oxd-Gly-Phe-OH as the C-terminal fragment (Scheme 2.8). This experiment in addition to the scope studied in Table 2.1 confirmed that most reactive side chains remain unmodified during the execution of these reaction conditions, contrary to other cleavage methods which lead to the oxidation of reactive amino acid side chains. Moreover, this method is applicable to peptide bond hydrolysis at Ser, Thr, reduced

Cys and Glu residues. The observed rates of reaction indicated that the Ser residue was most reactive to the DSC-mediated activation and hydrolysis conditions. Moreover, the multiple fragmentation patterns at the selected Ser, Thr, Cys and Glu residues can facilitate peptide sequencing of a broad range of naturally occurring peptides.



Scheme 2.8 DSC reaction scheme in the presence of a C-terminal carboxylic acid. Scheme adopted from reference (1) with permission.

Table 2.3 Site-selective cleavage of mutated and post-translationally modified peptides

Entry	Substrate	Conv.[a] (%)
1	Fmoc-Arg-Pro-Pro-Gly-Phe-**Ser**-Pro-Phe-Arg-NH_2 (**13**)	80
2	Fmoc-Leu-Asn-Asp-Arg-Leu-Ala-**Ser**-Tyr-Leu-NH_2 (**14**)	77
3	OAc-Ala-Val-Ala-Pro-Ala-Ala-**Ser**-Ile-Val-Ala-NH_2 (**15**)	73
4	OAc-Ala-Val-Ala-Pro-βAla-Ala-**Ser**-Ile-Val-AlaNH_2 (**16**)	75
5	Fmoc-Leu-Arg-Arg-Ala-**Ser**-(*N*-methyl)Leu-Gly-NH_2 (**17**)	79
6	Fmoc-Ile-βAla-Gly-Lys-Asn-Ser-Gly-Val-NH_2 (**18**)	70

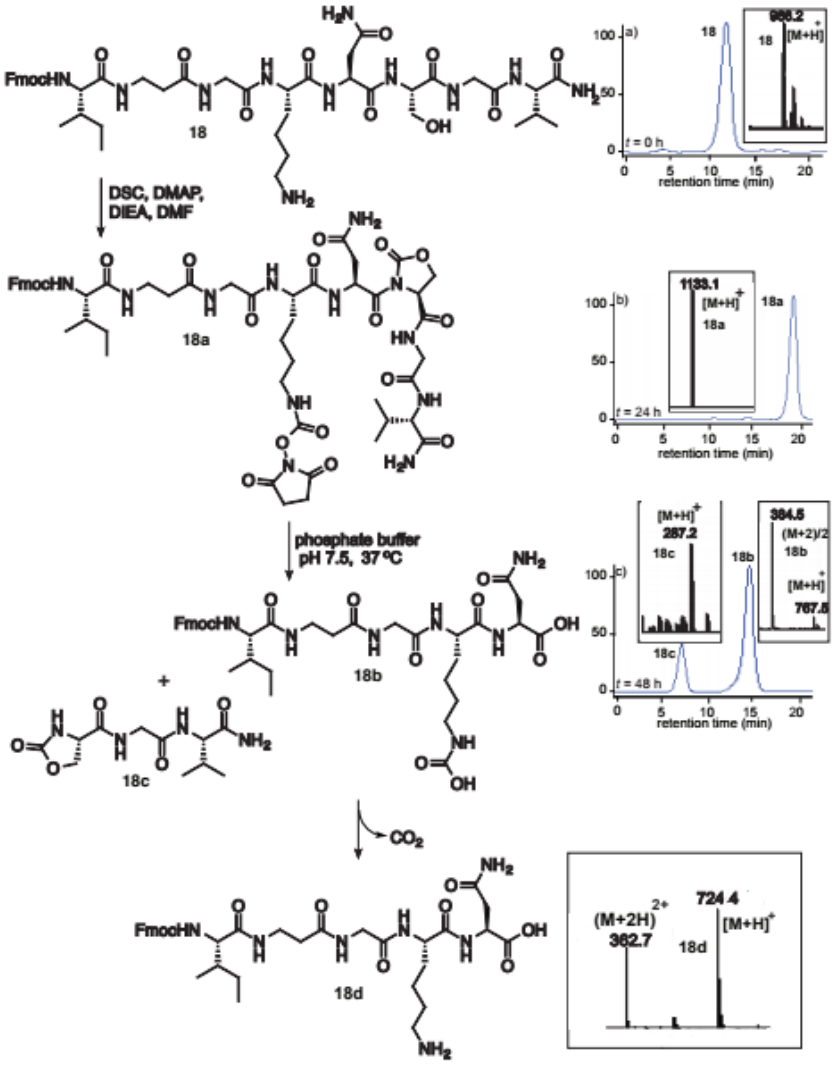
[a] N-terminal fragment.

Table 2.3 adopted from reference (1) with permission.

Subsequently, bioactive peptides, Bradykinin **13**[28] and type I hair keratin fragment **14**, were cleaved successfully with conversion rates ranging from 77-80% (entries 1 and 2, Table 2.3). Furthermore, we extended the scope of the reaction conditions to the scission of bioactive peptides with various post-translational modifications such as *N*-acetylated antimicrobial peptide AP00011 **15**, *N*-acetylated β amino acid containing antimicrobial peptide **16**, and *N*-methylated bioactive peptide **17**, which serves as a substrate for cAMP-dependent protein kinases[29] (entries 3-5, Table 2.3). The abovementioned modified peptides cleaved selectively at Ser with ease and high conversion. Chemical cleavage of these peptides proceeds independently of the presence of unnatural amino acid residues which is a great advantage over proteases which are unable to selectively recognize and cleave these modified peptides.

Moreover, this methodology was successfully applied on an amyloid β peptide, Fmoc-Ile-β Ala-Gly-Lys-Asn-Ser-Gly-Val-NH_2 **18**, which contained a mutated alanine β-amino acid residue, which is a notable mutation responsible for various age-related disease such as cataracts

and Alzheimer's (entry 6, Table 2.3) (Scheme 2.9).[30] Finally, the reaction was carried out on a fully unprotected peptide, SGISGPLS, a fragment of antimicrobial Bovine β-defensin 13. Treatment with DSC followed by hydrolysis generated fragments Oxd-GI, Oxd-GPL, and Oxd.



Scheme 2.9 Application of cleavage methodology on an amyloid β peptide fragment. Scheme adopted from reference (1) with permission.

CONCLUSIONS

In this study, a site-selective approach for the cleavage of peptides at the N-terminal of Ser, Thr, Cys, and Glu was developed by modification under metal free, mild conditions. The reactive side chains of other amino acids remained unmodified under the optimal reaction conditions,

which is an advantage over other chemical cleavage methods that yield modified fragments resulting from over-oxidation. Moreover, disulfide bonds are stable to the reaction conditions, thereby enabling the determination of disulfide positions within a peptide. The method also demonstrated broad substrate scope including cleavage of mutated and post-translationally modified peptides, which are unsuitable substrates for proteases. Additionally, a Pro-Ser bond was successfully cleaved, which is generally a difficult task for proteases due to the conformational restraints of Pro residues. Moreover, this method is also applicable to sequencing challenging peptides that have a high propensity for aggregation such as the amyloid peptides. The latter may be potentially applicable for determining the underlying mutations responsible for various age-related disorders. These results provided a firm basis for subsequent studies aimed at the development of artificial chemical proteases for the cleavage of target proteins.

2.6 EXPERIMENTAL SECTION

General: All commercial materials (Aldrich, Fluka, Nova) were used without further purification. All solvents were reagent grade or HPLC grade (Fisher). Anhydrous THF, diethyl ether, CH_2Cl_2, and DMF were obtained from a dry solvent system (passed through column of alumina) and used without further drying. All reactions were performed under air in round bottom flask. Yields refer to chromatographically pure compounds; % conversion were obtained by comparison of HPLC peak areas of products and starting material. HPLC was used to monitor reaction progress.

Materials: Fmoc-amino acids were obtained from Nova Biochem is under (EMD Millipore Corporation) (Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide resin was obtained from ChemPep Inc (Wellington, Florida). N,N,N′,*N*′-Tetramethyl-O (1Hbenzotriazol-uronium hexafluorophosphate (HBTU) was obtained from CreoSalus (Louisville,

Kentucky). N,*N*′-Disuccinimidyl carbonate (DSC) was obtained from Nova Biochem, under (EMD Millipore Corporation) (Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP): Merck KGaA (Darmstadt, Germany). N,*N*-Dimethylformamide (DMF): Macron Fine Chemicals (Center Valley, Pennsylvania). Dichloroethane (DCE), acetonitrile, N,*N*-Diisopropylethylamine (DIEA), N,*N*′-diisopropylcarbodiimide (DIC), were purchased from (EMD Millipore Corporation)(Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR 100 Matsonford Road Radnor, PA. Diethyl Ether: Sigma Aldrich (St. Louis, Missouri). Water was purified using a Millipore MilliQ water purification system.

NMR

This was done by Ryan Cohen at Merck facilities at Rayway, NJ. ^{1}H NMR spectra were recorded on a 600 MHz spectrometer and ^{13}C NMR spectra on a151 MHz, spectrometer at ambient temperature. All NMR chemical shifts (™) are referenced in ppm relative to residual solvent or internal tetramethylsilane. ^{1}H NMR chemical shifts referenced to residual DMSO-d_6 at 2.50 ppm, and ^{13}C NMR chemical shifts referenced to DMSO-d_6 at 39.52 ppm. ^{13}C NMR spectra are proton decoupled. NMR spectral data are reported as chemical shift (multiplicity, coupling constants (*J*), integration). Multiplicity is reported as follows: singlet (s), broad singlet (bs), doublet (d), doublet of doubles (dd), doublet of triplet (td), triplet (t) and multiplet (m). Coupling constant (*J*) in Hertz (Hz).

HPLC

Semi-Preparative HPLC:

Preparative HPLC chromatography (HPLC) was performed on Beckman Coulter equipped with System Gold 168 detector and 125P solvent module HPLC with a 10 mm C-18 reversed-phase column. All separations involved a mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1%

FA (v/v) in acetonitrile (solvent B). Semi-preparative HPLC method using a linear gradient of 0–80% acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 3.0 mL min^{-1}. The peptide absorbance was monitored at 220 nm (peptide bond) and 254 nm (aromatics) unless otherwise noted.

Analytical HPLC:

Analytical HPLC chromatography (HPLC) was performed on an Agilent 1200 series HPLC equipped with a 4.6x150 mm (5µm) C-18 reversed-phase column. All separations involved mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1% FA (v/v) in acetonitrile (solvent B). Peptide compositions were evaluated by analytical reverse phase HPLC using a gradient of 0.1% FA in acetonitrile versus 0.1% FA in water. Analytical HPLC method using a linear gradient of 0–80% 0.1% FA (v/v) acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 1.0 mL min^{-1}. The peptide was monitored by absorbance at 254 nm unless otherwise noted.

LCMS:

Mass spectrometry was performed using ultra high-performance liquid chromatography-mass spectrometry using the Agilent 1100 Series ESI LC MSD Spectrometer.

Fmoc Solid-Phase Peptide Synthesis:[31]

Peptides were synthesized manually on a 0.25 mmol scale using Rink amide resin. Fmoc–group was deprotected using 20% piperidine–DMF for 20 min to obtain a deprotected peptide-resin. Fmoc-protected amino acids (1.25 mmol) were sequentially coupled on the resin using a HBTU (1.25 mmol) and DIEA (1.25 mmol) for 2 h at room temperature. Peptides were synthesized using

standard protocols.[31] The peptide was cleaved from the resin using a cocktail of 95:2.5:2.5, trifluoroacetic acid (TFA):triisopropylsilane (TIPS):water (H_2O) for 2 h. The resin was removed by filtration and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was analyzed and purified by RP HPLC.

General procedure for cyclization: To a 5-mL round-bottom flask containing peptide (2-20mg (1 equiv.) in 0.5-2 mL dimethylformamide (DMF) was added a solution of DSC (5-20 equiv.), DIEA (5-20 equiv.) and crystal of DMAP in DMF (0.2-0.5 mL). The mixture was stirred at room temperature for 10 h. The reaction was concentrated under vacuum and resulting peptide was dissolved in 1:1 mixture of water and acetonitrile and purified by RP HPLC. The purified fractions were lyophilized to afford cyclized peptide as a white powder; yield: (60-80 %).

NMR Chemical Shifts for Fmoc-G-Oxd:

Analysis done by Ryan Cohen utilizing Merck facilities located in Rahway, NJ. ^{1}H NMR chemical shifts referenced to residual DMSO-d_6 at 2.50 ppm, and ^{13}C NMR chemical shifts referenced to DMSO-d_6 at 39.52 ppm. 1 H NMR (600 MHz, DMSO-d_6) δ 7.90 (d, J = 7.5 Hz, 2H), 7.75 (s, 1H), 7.72 (d, J = 7.4 Hz, 2H), 7.64 (t, J = 6.0 Hz, 1H), 7.42 (t, J = 7.5 Hz, 2H), 7.39 (s, 1H), 7.33 (t, J = 7.4 Hz, 2H), 4.73 (dd, J = 9.1, 3.1 Hz, 1H), 4.60 (t, J = 9.0 Hz, 1H), 4.36 – 4.15 (m, 6H).

^{13}C NMR (151 MHz, DMSO-d_6) δ 169.9, 169.1, 156.4, 153.5, 143.8, 140.7, 127.6, 127.1, 125.2, 120.1, 66.8, 65.8, 55.4, 46.6, 44.1.

2.7 REFERENCES

1. Elashal, Hader; Raj, M. *Chem. Commun.* **2016,** *52*, 6304.
2. Lundblad, R. L. Techniques in Protein Modification, CRC Press, Boca Raton, Florida, 1995.
3. Antonelli, G.; Turriziani, O. *Int J Antimicrob Agents* **2012**, *40* (2), 95.
4. Mótyán, J.; Tóth, F.; Tőzsér, J. *Biomolecules* **2013**, *3* (4), 923.
5. (a) Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1*, 80. (b) Edman, P.; Högfeldt, Erik; Sillén, Lars Gunnar; Kinell, Per-Olof. *Acta Chem. Scand*. **1950**, *4*, 283.
6. Smith, J. B. *ENCYCLOPEDIA OF LIFE SCIENCES*; Macmillan Publishers Ltd, Nature Publishing Group, 2001.
7. Sanger, F. *Biochem. J.* **1945**, *39*, 507.
8. Sanger, F. *Les Prix Nobel* **1958**, 134.
9. Gross, E. *Methods Enzymol.* **1967**, *11*, 238.
10. Degani, Y.; Patchornik, A. *Biochemistry.* **1974**, *13*, 1.
11. Holmes, T. J.; Lawton, R. G. *J. Am. Chem. Soc.* **1977**, *99*, 1984.
12. Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem Sci*. **2011**, *2*, 1666.
13. Mahoney, W. C.; Smith, P. K.; Hermodson, M. A. *Biochemistry*. **1981**, *20*, 443.
14. Kaiser, R.; Metzka, L. *Anal. Biochem.* **1999**, *266*, 1.
15. Walker, J. M. The Protein Protocols Handbook, Humana Press, Totowa, NJ, 2002.
16. Tanabe, K.; Taniguchi, A.; Matsumoto, T.; Oisaki, K.; Sohma, Y.; Kanai, M. *Chem. Sci.* **2014**, *5*, 2747.

17. Kaminskaia, N. V.; Johnson, T. W.; Kostic, N. M. *J. Am. Chem. Soc.***1999**, *121*, 8663.

18. Parac, T. N.; Ullmann, G. M.; Kostic, N.M. *J. Am. Chem. Soc.* **1999**, *121*, 3127.

19. Zhu, L.; Qin, L.; Parac, T. N.; Kostoc, N. M. *J. Am. Chem. Soc.* **1994**, *116*, 5218.

20. Hoyer, D.; Cho, H.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 3249.

21. Schepartz, A.; Cuenoud, B. *J. Am. Chem. Soc.* **1990**, *112*, 3247.

22. Kita, Y.; Nishii, Y.; Higuchi, T.; Mashima, K. *Angew. Chem. Int. Ed.* **2012**, *51*, 5723.

23. Milovic, N. M.; Kostic, N. M. *J. Am. Chem. Soc.* **2003**, *125*, 781.

24. Seki, Y.; Tanabe, K.; Sasaki, D.; Sohma, Y.; Oisaki, K.; Kanai, M. *Angew. Chem.* **2014**, *126*, 6619. (*Angew Chem, Int. Ed.* **2014**, *53*, 6501).

25. De Marco, R.; Tolomelli, A.; Campitello, M.; Rubini, P.; Gentilucci, L. *Org. Biomol. Chem.* **2012**, *10*, 2307.

26. Schacherl, M.; Pichlo, C.; Neundorf, I.; Baumann, U. *Structure.* **2015**, *23* (9), 1632.

27. Gorman, J. J.; Wallis, T. P.; Pitt, J. J. *Mass Spectrum Rev.* **2002**, *21*, 183.

28. Golias, C.; Charalabopoulos, A.; Stagikas, D.; Charalabopoulos, K.; Batistatou, A. *Hippokratia.* **2007**, *11*, 124.

29. Bramson, H. N.; Thomas, N. E.; Kaiser, E. T. *J. Biol. Chem.* **1985,** *260*, 15452.

30. Maeda, H.; Takata, T.; Fujii, N.; Sakaue, H.; Nirasawa, S.; Takahashi, S.; Sasaki, H.; Fujoo, N. *Anal. Chem.* **2015**, *87*, 561.

31. Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.

CHAPTER 3: CYLIC AND LASSO PEPTIDES: STRUCTURE DETERMINATION, ROTAXANE FORMATION, AND TOPOLOGY ANALYSIS

3.1 ABSTRACT

Natural products play a vital role in drug discovery and design as new peptide-based therapeutics continue to enter clinical trials. Bioactive peptides produced by non-ribosomal pathways such as N-methylated cyclic peptides and lasso peptides contain unique properties such as high receptor-binding activity, specificity towards their target, metabolic stability, and conformational rigidity. However, due to their distinctively complex topology, the ability to effectively determine their sequence and utilize the structure activity relationship in developing lead compounds remains limited. These limitations are primarily due to incompatibility with the majority of available sequencing methods including enzymatic digestion, chemical methods, and tandem mass spectroscopy. To overcome these limitations, we have developed a chemical methodology that can modify the backbone peptide chain at Ser to generate the cyclic urethane moiety, 2-oxazolidinone, thereby rendering the Ser N-terminal peptide bond susceptible to cleavage under hydrolytic conditions. This fragmentation method allowed for easy and rapid sequence determination by tandem mass spectroscopy. The methodology was applied to a broad substrate scope including the ability to modify peptides containing Thr, Cys, Glu, and unnatural amino acids such as homoserine which account for 60% of naturally occurring cyclic peptides. Utilizing this methodology, the cleavage and sequencing of stable N-methylated cyclic peptides, Polycarponin C and a Somatostatin analogue, has been accomplished. In addition, lasso peptides lariatin A, albusnodin, and benenodin-1 have also been structurally validated. Additionally, we created a peptide rotaxane by the selective cleavage of benenodin-1. This is the first chemical

methodology to be effectively applied for the cleavage of lasso peptides which is a major milestone in simplifying sequence determination of lasso peptides, studying the lasso topology, monitoring the effect of cleavage on sterically locked residues, and using lasso peptides as scaffolds for making rotaxane type interlocked machines.

3.2 CHAPTER OBJECTIVES

This study was a collaborative effort between our lab and the Link Lab at Princeton University, which specializes in lasso peptides. Chuhan Zong expressed and supplied the lasso peptides and helped in lasso data analysis post reaction. Meanwhile, cyclic peptide synthesis, purification, modification, and analysis in addition to lasso peptide modification was executed by the Raj group.

This chapter describes a simple and versatile chemical methodology that facilitates sequence determination of cyclic and lasso peptides, explores the lasso topology and the effect of cleavage on sterically locked residues, and uses lasso peptides as scaffolds for making rotaxane type interlocked machines. This methodology is the first chemical methodology to be used for the abovementioned applications which is a major milestone in understanding of the importance of residue location in maintaining topologically distinct structures and harnessing their unique activity. Based on the results described in Chapter 2, this chapter applies our strategy for the ring opening of cyclic peptides and/or cleavage of lasso peptides to yield simplified fragments for peptide sequencing. *The significance of this approach is based on a wide substrate scope and its compatibility with highly abundant amino acids which are found in over 60% of naturally occurring cyclic and lasso peptides*. Moreover, this method is chemo-selective and does not lead to side reactions or byproducts in the presence of unprotected amino acid side chains. We have developed a general method for the cleavage and sequencing of peptides with unique topologies.

As compared to other cleavage methodologies, this method is the first chemical method to be applied for the cleavage of N-methylated peptides and lasso peptides which provides a gateway to further probe the mechanics of how the lasso structure is threaded and the effectiveness of steric locks post cleavage. We have demonstrated the effectiveness of our method for the cleavage of cyclic peptides, N-methylated cyclic peptides, and lasso peptides.

3.3 GRAPHICAL ABSTRACT

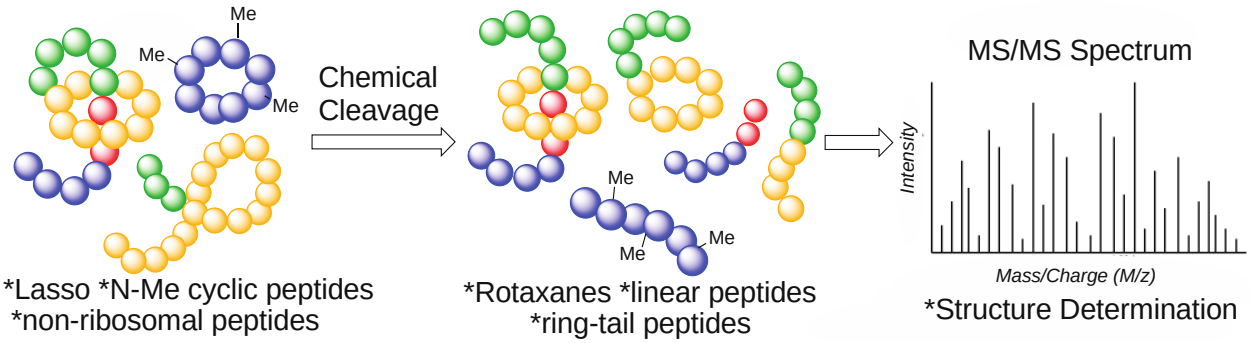


Figure 3.1 *Graphical abstract for Cyclic and Lasso Peptides: Sequence Determination, Topology Analysis, and Rotaxane Formation.* Figure adopted with permission from: Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A.J.; Raj, M. *Angew. Chem.* **2018**, DOI: 10.1002/anie.201801299.[1]

3.4 INTRODUCTION

The discovery of active natural products has been at the root of modern drug design and they continue to play a vital role as ground-breaking peptide based drugs make their way towards clinical trials.[2] Biologically active peptides produced by non-ribosomal pathways carry unique characteristics which allows them to have diverse functionalities.[3] Specifically, cyclic peptides are of considerable interest since they contain the following key properties: high receptor-binding activity, specificity towards their target, metabolic stability, conformational rigidity, and the potential to be cell permeable.[4] Unique peptides such as Cyclosporine A (an immunosuppressant),[5] peptides rich in N-methylation, microcin J25 (lasso peptide with anti- bacterial activity),[6] and

Kalata B1[7,8] (cyclotide with insecticidal activity) are all natural products with exceptional stability and bioactivity. However, the ability to immediately study and harness the structure activity relationship of these types of cyclic peptides in order to develop lead compounds is hindered due to the difficulty in determining the sequence of cyclic peptides.

While enzymatic cleavage is one of the most utilized methods for peptide sequencing, lasso peptides, cyclic peptides (especially those that are N –methylated), and other uniquely structured peptides are resistant to enzymatic cleavage, with slower kinetic rates by 20-300 folds as compared to their linear counterparts.[9] Mechanistically, the decrease in cleavage rates results from a slower rate of cleavage (k_{cat}) and a weaker interaction with proteases (K_m).[10] On the other hand, chemical methods for peptide bond cleavage can provide an efficient alternative to overcoming the abovementioned limitations. However, the existing methods for selective cleavage carry several limitations. Since cyclic peptides lack a free N-terminal amine, they cannot be sequenced via Edman degradation.[11] The most commonly used method for chemical cleavage of cyclic peptides utilizes CNBr for the selective modification of methionine;[12] however, methionine is the second least abundant amino acid in nature,[13] therefore its presence in a naturally isolated product is a rarity. Moreover, the cleavage is dependent on the neighboring amino acids; for example if methionine is followed by serine or threonine, side reactions convert methionine to homoserine without peptide bond cleavage and also leads to the modification of tyrosine's side chain.[14] Due to the lack of available chemical methods which can effectively modify abundant amino acid residues and afford cleavage of complex cyclic peptides, their feasibility for the sequencing of unique biologically active peptides remains limited.

While many advancements have been made in the field of sequencing technology, their practical utility in the sequencing of naturally available cyclic peptides remains problematic.[15]

Typically, tandem mass spectroscopy is the most rapid and efficient method to determine the sequence of a peptide. While this holds true for linear peptides, structure and sequence assignment becomes more complicated for cyclic peptides. Cyclic peptides undergo ring opening at various positions to generate a family of mass degenerate ions which further fragments to generate complex mixtures of various short fragments.[16-20] The interpretation of the various fragmentation patterns becomes difficult and impractical. Moreover, analysis of the MS/MS spectra is also hindered by the lack of defined N and C termini. Additionally, many biologically active peptides are synthesized through non-ribosomal pathways which increases the likelihood that unnatural amino acid residues may be present in the peptide which adds to the complexity of spectroscopic analysis. To this end, nuclear magnetic resonance spectroscopy (NMR) remains the only viable method for accurately predicting the structure and sequence of cyclic peptides, lasso peptides, and other peptides with distinct topologies. Furthermore, wrongful assignment of NOE contacts can lead to an inaccurate structure determination.[21] For example, a branched- cyclic peptide can be mistaken for a lasso structure and the opposite is also true.[21] Therefore, careful, lengthy, and tedious data interpretation are required in order to yield the correct sequence and structure assignment of peptide natural products through NMR.

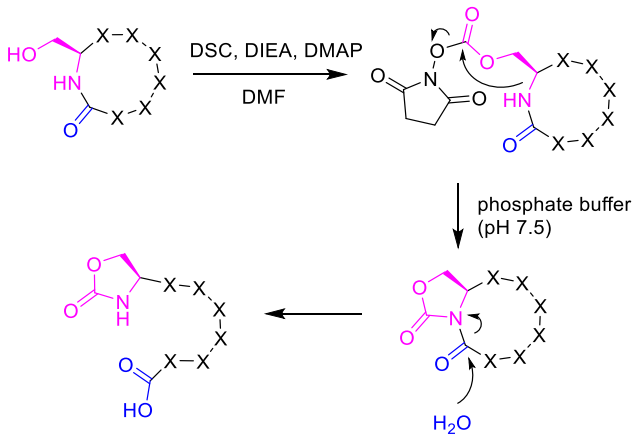
In this chapter, a simple and convenient approach for sequence determination of non-ribosomally produced highly stable lasso peptides and cyclic peptides with unnatural amino acid residues and N-methylation is reported. The methodology entails the site-selective amine backbone activation at Ser, Thr, Glu, and/or Cys. Upon activation, the amide bond at the N-terminal side of the residue(s) becomes susceptible to cleavage thereby converting the peptide into its linear counterpart which can easily be sequenced by MS/MS. The proposed strategy has a broad substrate scope and is compatible with some of the most highly abundant amino acids found in

nature.[13] To implement this approach, at least one Ser, Thr, Glu, or Cys should be present in the desired cyclic peptide, which is of high probability due to the presence of at least one of the abovementioned residues in approximately 60% of naturally occurring cyclic peptides and 66% of lasso peptides as determined by a survey sample screened through NORINE, [22] a cyclic peptide database, and a comprehensive list of all lasso peptides discovered to date.[21] Additionally, this methodology can also be utilized to modify various unnatural amino acid residues that are readily found in numerous bioactive cyclic peptides obtained from nature. Unnatural amino acids with a reactive side chain functionality such as a hydroxyl group or an amine and have the ability to form a five membered or six membered ring would be compatible with this strategy. For example, homoserine, beta-hydroxy-Phe (*β*-OH-Phe), *β*-OH-Asn, *β*-OH-Gln, *β*-OH-Asp, *β*-OH-Val, *β*-OH-Tyr, and beta homoserine (*β*-Hse) are amongst the unnatural amino acids which fulfill these criteria. Moreover, this method is chemo-selective and does not interfere with any other unprotected amino acid side chain; it does not generate byproducts or lead to side reactions. To the best of our knowledge, this is the only chemical method available that has a potential of sequencing more than 60% of lasso peptides and N-methylated cyclic peptides which is a significant advancement as compared to pre-existing techniques.

3.5 RESULTS AND DISCUSSION

Due to the difficulties associated with utilizing tandem mass spectroscopy for cyclic peptide analysis, we developed a strategy that converts these peptides to their linear counterparts. Upon linearization of the cyclic peptide, the previously mentioned obstacles related to the use of MS/MS for sequence determination become eliminated thereby simplifying the process of MS/MS analysis. The proposed strategy involves the conversion of Ser to an oxazolidinone moiety, thereby making the amide bond susceptible to cleavage at the Ser N-terminus under hydrolytic conditions.

The modification of Ser encompasses the reaction of its hydroxymethyl group with *N*,*N*-disuccinimidyl carbonate (DSC) to generate an intermediate (Scheme 3.1). Next, nucleophilic attack of the amide nitrogen of the peptide backbone produces the desired 2-oxazolidinone moiety. Upon formation of 2-oxazolidinone, the C-N bond is easily hydrolyzed under mild conditions to generate a ring-opened (linear) peptide. Earlier attempts at the formation of this moiety on Fmoc/Boc protected peptides led to the formation of dehydroalanine.[23] This methodology is also applicable to Thr and Cys which exhibit similar side chain functionalities as serine. Moreover, glutamic acid also serves as a substrate for this modification due to its nucleophilic side chain which has the ability to form a pyroglutamyl imide moiety. Our group previously reported the application of this moiety for the synthesis of thioesters and for the cleavage of linear peptides.[24,25]



Scheme 3.1 Macrocyclic Peptide Sequencing Strategy. Scheme adopted from reference (1) with permission.

Using previously determined reaction conditions,[25] we proceeded to sequence Polycarponin C, a mono-methylated cyclic peptide with anti-cancer activity.[26] First, we incubated the peptide with chymotrypsin[27] since it contains a phenylalanine to assess if enzymes can be utilized for opening of these N-methylated peptides. After 48 hours, no detectable cleavage of Polycarponin C was noted which confirmed the limitation of using enzymatic degradation in

sequencing of these peptides (Figure 3.2a). Next, direct MS/MS analysis was carried out in an effort to determine the peptide sequence without any modification (Figure 3.2b). However, the spectra generated complex fragmentation patterns that were difficult to assign. Moreover, several fragments corresponded to the same mass, which made accurate assignment a challenging task. We then proceeded to selectively modify the peptide backbone chain at the N-terminus of Thr by using DSC, which is otherwise very stable and has a half-life of 500-1000 years at pH 4-8.[28] Following the treatment with DSC, the peptide was hydrolyzed to its linear counterpart under mild aqueous buffer conditions (pH 7.4) at room temperature (Figure 3.2c). Noteworthy, the cleavage proceeded smoothly even in the presence of a Pro residue neighboring the site of cleavage which highlights a key advantage over enzymatic digestion which cannot proceed in the presence of a nearby Pro.[29] Upon analysis of the linear counterpart by MS/MS, the spectra clearly showed the ring opened peptide with a clear fragmentation pattern that is easier to decipher as compared to the MS spectra obtained by direct analysis of the cyclic peptide (Figure 3.2c).

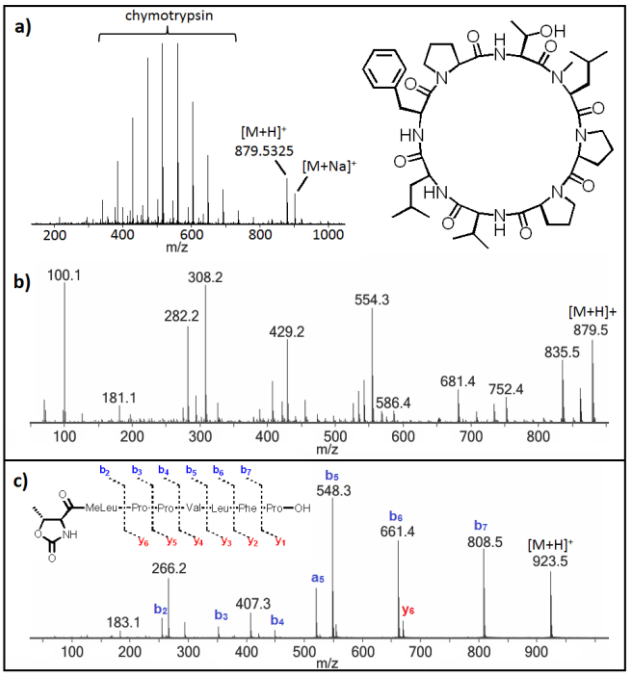


Figure 3.2 Sequencing polycarponin C. a) HRMS spectrum after 48 h incubation with chymotrypsin showing proteolytic resistance, b) complex MS/MS spectrum of polycarponin C, and c) simplified MS/MS spectrum of its ring-opened analogue. Figure adopted from reference (1) with permission.

Next, we sought to sequence a stable tri-N-methylated Somatostatin analogue.[30] After incubation with a mixture of trypsin and chymotrypsin[27] for 48 hours, the peptide remained intact, with no observable cleavage (Figure 3.3). Direct MS/MS analysis was also troublesome due to the complications noted above. Conversely, upon utilizing our method for sequencing, the peptide was opened to its linear counterpart following Thr activation and hydrolysis (Figure 3.3). By applying our methodology on these two cyclic peptides, we have demonstrated the efficiency of our strategy in the sequencing of bioactive cyclic peptides containing multiple methyl groups.

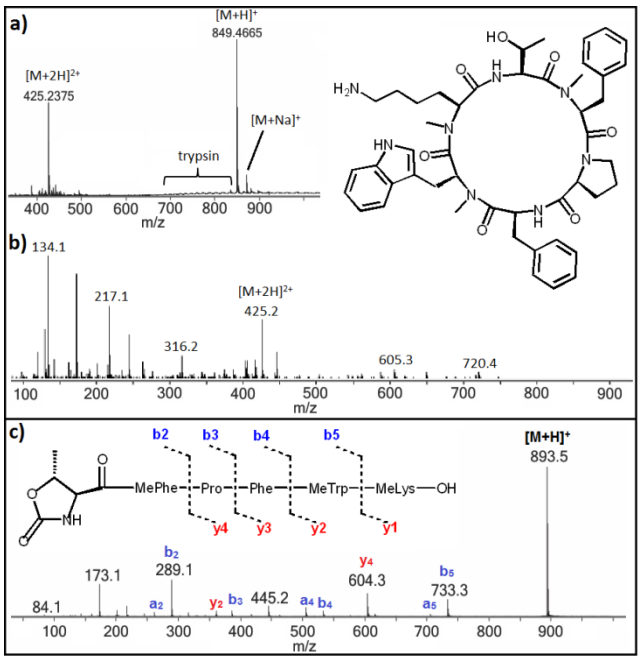
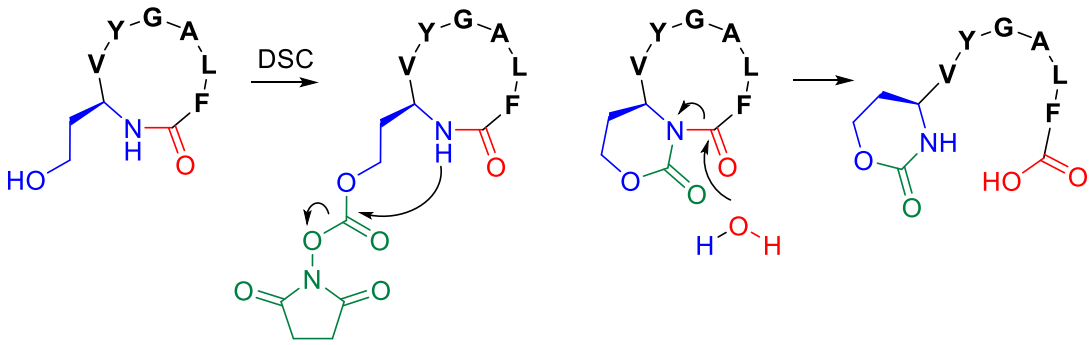


Figure 3.3 Sequencing data for a somatostatin analog. A) HRMS spectrum after protease incubation showing no detectable cleavage, b) MS/MS spectrum of cyclized peptide, and c) MS/MS spectrum of ring-opened peptide. Figure adopted from reference (1) with permission.

To demonstrate compatibility of our method with unnatural amino acids which are abundant in cyclic peptides isolated from nature, we synthesized one such model cyclic peptide

Gly-Tyr-Val-Hse-Phe-Leu-Ala containing homoserine (Hse). Due to hydroxyl side chain of homoserine and the side chain proximity to the amidic nitrogen, we hypothesized that homoserine can effectively react with DSC to form a six-membered oxazinanone ring making the homoserine N-terminal amide bond susceptible to hydrolysis (Figure 3.4). As predicted, homoserine in a cyclic peptide underwent modification upon treatment with DSC to a 6-membered ring followed by hydrolysis to its linear counterpart under neutral buffer conditions.



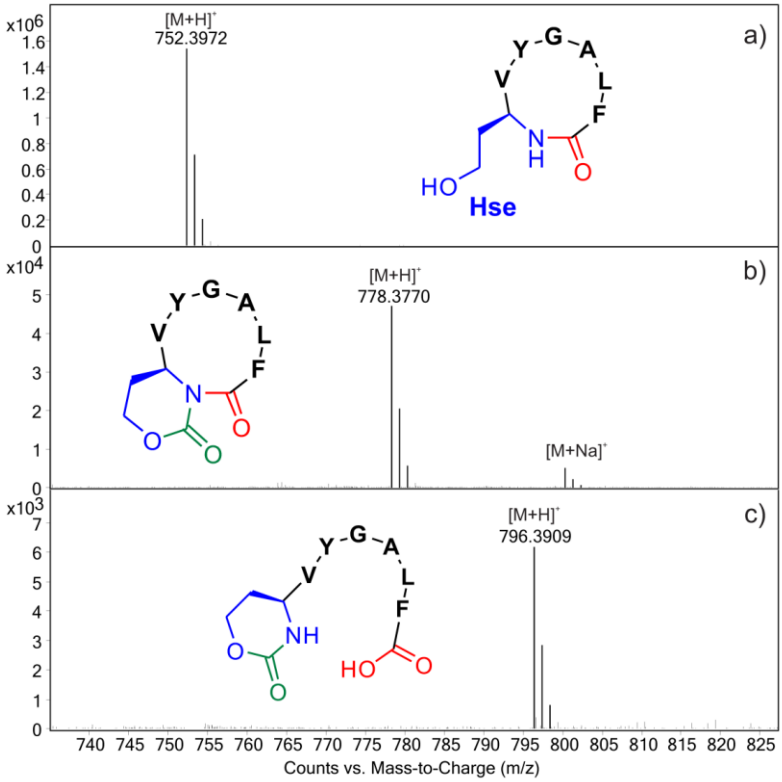


Figure 3.4 Cleavage data for Homoserine (Hse)-containing macrocyclic peptide. HRMS spectra of (a) *cyc*-Gly-Tyr-Val-Hse-Phe-Leu-Ala, (b) activated six-membered oxazinanone moiety, and (c) hydrolytically-cleaved linear peptide, which is amenable for MS/MS sequencing. Figure adopted from reference (1) with permission.

We have also applied this approach for sequencing a model cyclic peptide with serine residue *cyc*(Ser-Phe-Arg-Tyr-Ala-Glu) by selective modification of serine to 5-membered 2-oxazolidinone moiety followed by the cleavage of the backbone chain at the N-terminus of serine under hydrolytic conditions (Figure 3.5). The resulting linear peptide was then sequenced by tandem mass spectrometry. By displaying applicability to homoserine in addition to Ser, Cys, Thr, and Glu, we can effectively apply our method to sequence over 60% of naturally occurring cyclic peptides.

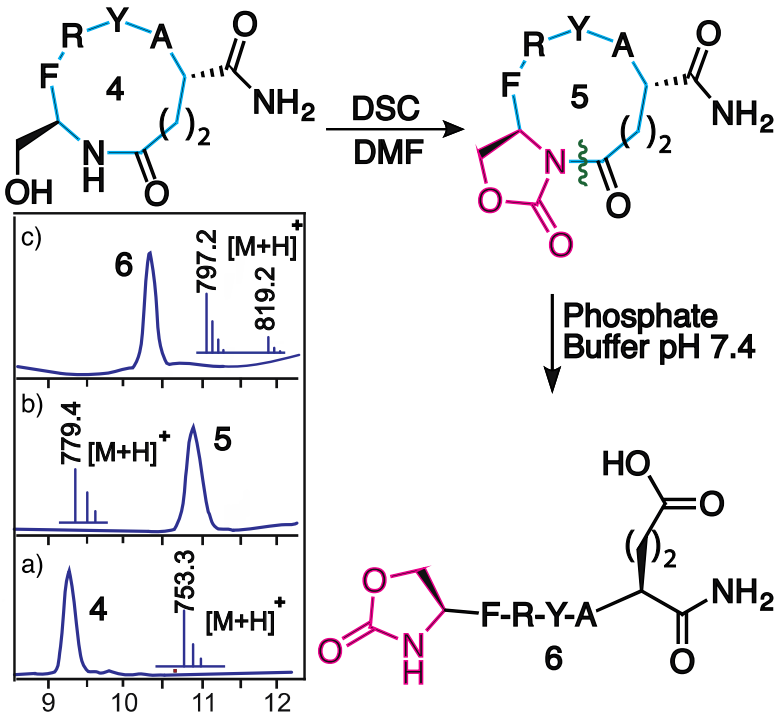


Figure 3.5 Serine activation to cyclic urethane moiety. a) **4:** Rt = 9.3 min, b) **5**: Rt = 10.8 min, c)

6: Rt = 10.3 min. Rt = retention time. Inset shows MS of compounds **4**, **5** and **6**. Figure adopted from reference (1) with permission.

To further assess the scope of our methodology, we explored the applicability of our strategy with lasso peptides which are known for their wide range of bioactivity including antimicrobial, enzyme inhibitory, and receptor antagonistic in addition to their unique topology which contains a ring, loop, and a tail component.[21,31] Sequencing of this distinctive class of peptides is problematic due to the unique lasso topology which is highly stable toward enzymatic cleavage by common digestive enzymes. On the other hand, isopeptidases have been found in organisms that express lasso peptides. These isopeptidases have the ability to cleave only the isopeptide bond present in the ring to yield a ring opened linear peptide.[32-35] However, each isopeptidase is specific only to the lasso peptide isolated from the same prokaryote.[32-35] Therefore, there is a need for chemical cleavage methods since there is no general method that can further assist in the sequence determination of the majority of lasso peptides.

To test the stability of lasso peptides towards enzymes before proceeding with cleavage through chemical modifications, we chose lariatin A and albusnodin. Lariatin A is an antimicrobial lasso peptide which contains an Arg residue in the ring region,[36] while albusnodin, another lasso peptide, contains two consecutive Arg residues in the loop region.[37] Since Arg is a substrate for trypsin, we treated both peptides with trypsin to investigate if cleavage would occur. After incubating the peptides with trypsin overnight, no observable cleavage was detected, thereby confirming the resistance of lasso peptides to proteolytic degradation. As further proof, we also incubated lariatin A with a mixture of trypsin and chymotrypsin since it also contains Trp, a substrate for chymotrypsin. After 24 hr, lariatin A remained completely intact. In fact, the lasso spectra after treatment with enzymes looks more complicated because these enzymes cleave themselves and generate multiple small fragments.

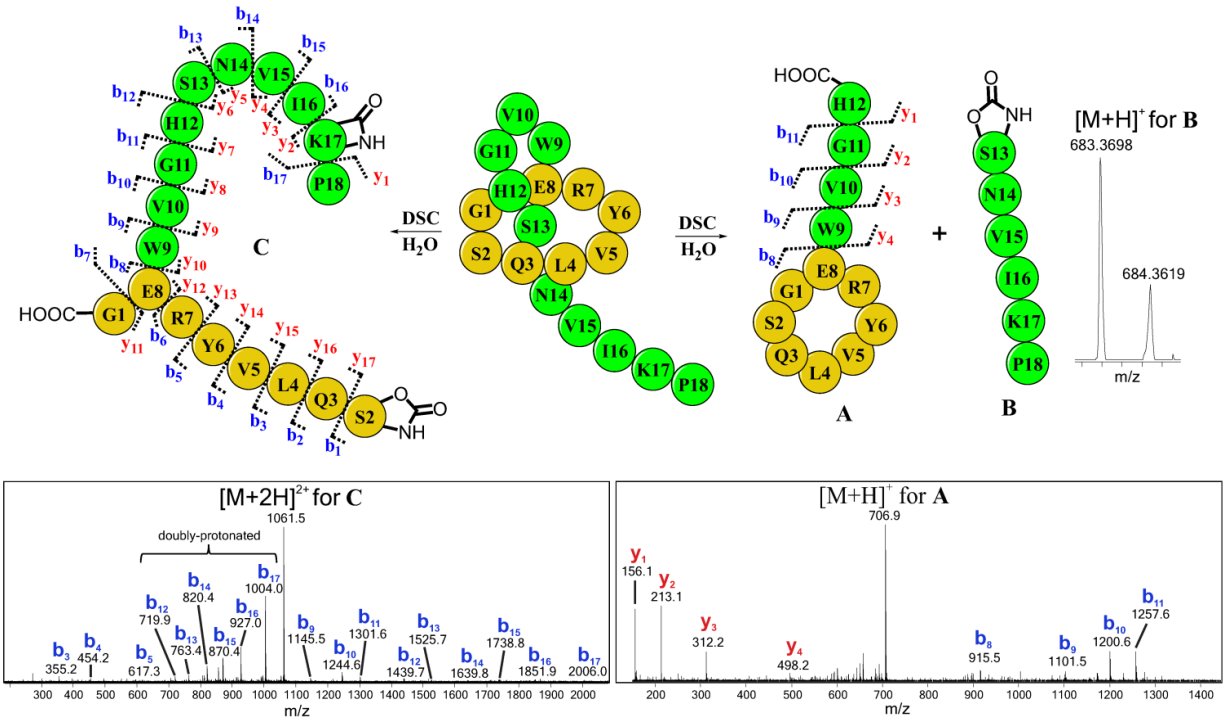


Figure 3.6 Sequencing data for lasso peptide, lariatin A. MS/MS spectra of the fully ring-opened peptide, **C**, and cleavage product, **A**, from reaction at Ser-13, which is present at the steric lock region. Figure adopted from reference (1) with permission.

We then proceeded to utilize our methodology for the cleavage of lariatin A and albusnodin since both peptides contain Ser within their sequence. Since lariatin A contains two Ser, one in the ring and the other in the loop region, we expected that treatment with DSC overnight will modify only the Ser which is present in the ring because the serine present in the loop region is threaded inside the ring making it less exposed. To our surprise, the major product was modification of Ser-13, and after hydrolysis, this generated a branched, cyclized peptide (**A**) and a short, linear peptide (**B**). A small amount of the expected branched, ring-opened product (**C**) was also observed; however, unexpectedly Lys17 was converted into a cycloocta-urea moiety (Figure 3.6). Previously on short model peptides, modification of the Lys side chain amino group was observed after treatment with DSC to the corresponding carbamic acid that subsequently would regenerate Lys under the hydrolytic conditions.[25a] However, in this case, Lariatin A's molecular topology may

facilitate intramolecular ring formation. Meanwhile, albusnodin has one Ser located in the ring region of the peptide which was expected to afford a linear peptide upon hydrolysis.[25b] Treatment with DSC and subsequent hydrolysis efficiently yielded the conversion of albusnodin to a ring open peptide. Although albusnodin has two Arg residues, no lasso peptide cleavage was observed upon treatment with carboxypeptidase and trypsin which emphasizes the high stability of lasso peptides towards enzymes even in the presence of two enzymatic cleavage sites.

Notably, efficiency of the reaction conditions decreases when applied to lasso peptides due to variations in structure such as loop size, ring size, steric locks, and residues threaded through the ring making them difficult to access. However, even with lower conversion rates, the cleaved product is enough to conduct MS/MS analysis thereby providing insight into lasso sequencing and structure determination.

In addition to exploring our strategy with the above-mentioned lasso peptides, we also studied reactivity with benenodin-1, another lasso peptide. Benenodin-1 is unique to our study since it contains three possible cleavage sites. It contains Ser, Thr, and Glu in the loop region of the peptide.[32] It was interesting to monitor if modification and cleavage of Glu would occur since it serves as a steric lock.[31] Interestingly, we observed the modification and cleavage of only Ser in benenodin-1 which resulted in the formation of a rotaxane structure (Figure 3.7). Rotaxanes have an interlocked architecture where a dumbbell shaped molecule is threaded through a macrocyclic ring which gives them immense potential as molecular machines.[38-40] In the case of benenodin-1, Glu and Gln serve as steric locks which prevent the un-threading of the linear peptide obtained after cleavage at the N-terminus of Ser thereby disconnecting the loop from the ring portion. This result showed that our methodology has the potential to utilize lasso peptides as scaffolds for the

formation of rotaxanes and various other interlocked peptide based molecular machines which are challenging to make through the available synthetic methods.[41]

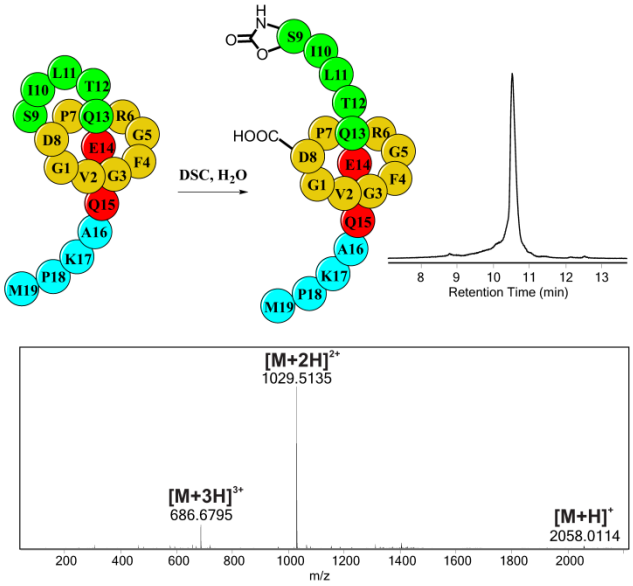
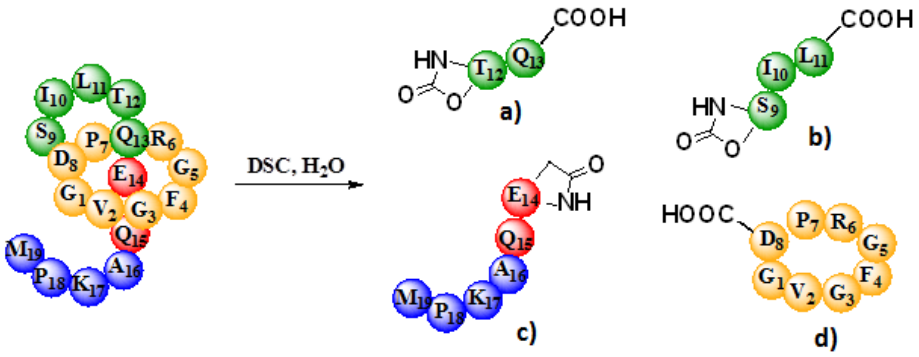


Figure 3.7. LC-MS of a rotaxane obtained from the modification/cleavage at Ser-9 of benendonin-1 ΔC5. Figure adopted from reference (1) with permission.

Other benenodin-1 ΔC5 cleavage products were also observed from Ser, Thr, and Glu residues that facilitated sequencing (Scheme 3.2). In these cases, linear peptides were generated from the loop region. Cleavage at or near the steric lock unthreads the loop from the ring, while distal site cleavage generates the rotaxane structure through conservation of the threaded architecture.



Scheme 3.2 Benenodin-1 ΔC5 Cleavage Products. Figure adopted from reference (1) with permission.

Overall, our approach simplifies the complicated lasso topology to simple fragments such as linear peptides and/or ring-tail peptides which can then be easily sequenced by further enzymatic digestion and/or tandem mass spectroscopy. Importantly, this method also has the potential to make peptide rotaxanes from lasso scaffolds.

3.6 CONCLUSIONS

We have developed a general method for the cleavage and sequencing of peptides with unique topologies. We have demonstrated the effectiveness of our method for the cleavage of cyclic peptides, N-methylated cyclic peptides, and lasso peptides. As compared to other cleavage methodologies, this method is the first chemical method to be applied for the cleavage of N-methylated peptides and lasso peptides which provides a gateway to further probe the mechanics of how the lasso structure is threaded and the effectiveness of steric locks post cleavage. Moreover, this methodology has a broad substrate scope due to its reactivity with highly abundant residues including Ser, Thr, Glu, and Cys as well as it compatibility with several unnatural amino acid residues such as homoserine and other beta-hydroxy amino acids. Overall, this methodology provides a means with the potential to sequence over 60% of naturally occurring cyclic peptides and 66% of all lasso peptides. Adding to its attractiveness, this methodology is selective to the abovementioned residues; it does not react with other unprotected amino acid side chains, generate byproducts, or lead to any side reactions. By utilizing this methodology, we are able to sequence peptides with unique structures, study the effect of loop size and residue location on lasso peptide cleavage, make interlocked structures like rotaxanes from lasso scaffolds, and provide a platform for discovering newly isolated biologically active peptides.

3.7 EXPERIMENTAL SECTION

General: All commercial materials (Sigma-Aldrich, Fluka, and Novabiochem) were used without further purification. All solvents were reagent or HPLC (Fisher) grade. Diethyl ether, dichloromethane, and *N*,*N*-dimethylformamide were obtained from a dry solvent system (passed through an alumina column) and used without further drying. All reactions were performed under air in glass vials. HPLC was used to monitor reaction progress.

Materials: Fmoc-amino acids were obtained from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide, Wang, and 2-chlorotrityl (2-ClTrt) resins were obtained from ChemPep Inc. (Wellington, Florida). *N*,*N*,*N*′,*N*′-Tetramethyl-*O*-(1H-benzotriazol-1-yl)uronium hexafluorophosphate **(**HBTU) was obtained from CreoSalus (Louisville, Kentucky). *N*,*N*′-Disuccinimidyl carbonate (DSC) was obtained from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP) was obtained from Merck KGaA (Darmstadt, Germany). *N*,*N*-Dimethylformamide (DMF) was obtained from Macron Fine Chemicals (Center Valley, Pennsylvania). Acetonitrile, *N*,*N*- diisopropylethylamine (DIEA), triisopropylsilane (TIPS), and *N*,*N*'-diisopropylcarbodiimide (DIC) were purchased from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR (100 Matsonford Road Radnor, Pennsylvania). Diethyl ether, dichloromethane (DCM), methanol (MeOH), phenylsilane, 1-hydroxy-7-azabenzotriazole (HOAt), tetrakis(triphenylphosphine)palladium(0) [$Pd(PPh_3)_4$], 2-nitrobenzenesulfonyl chloride (*o*-NBS-Cl), 2,4,6-trimethylpyridine, triphenylphosphine, tetrahydrofuran (THF), diisopropyl azodicarboxylate (DIAD), mercaptoethanol, 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), acetonitrile, toluene, dimethyldichlorosilane, and

1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) were obtained from Sigma-Aldrich (St. Louis, Missouri). Water was purified using a Millipore Milli-Q water purification system. Lariatin A was purchased from Abcam (Cambridge, Massachusetts).

LC-MS:

A Portion of analysis was done at Seton Hall University; the other portion was done at Merck facilities at Rahway, NJ by Ryan Cohen. Peptide compositions were evaluated by high performance liquid chromatography (HPLC) on an Agilent 1100 series HPLC equipped with a 4.6x150 mm (5μm) C18 column. All separations used mobile phases of 0.1% formic acid (v/v) in water (solvent A) and 0.1% formic acid (v/v) in acetonitrile (solvent B). A linear gradient of 0–80% solvent B in 30 minutes at room temperature with a flow rate of 1.0 mL min^{-1} was used. The eluent was monitored by UV absorbance at 220 nm unless otherwise noted. Mass spectrometry to check reaction mixtures was performed on an Agilent 1100 Series HPLC with MSD VL mass spectrometer using positive polarity electrospray ionization (+ESI).

LC-HRMS:

Analysis was done at Merck facilities at Rahway, NJ by Ryan Cohen. High resolution MS data were acquired using an Agilent 1290 UHPLC with a 6520 Q-ToF mass spectrometer under positive polarity electrospray ionization (+ESI) with capillary and fragmentor voltages set to 3.5 kV and 175 V, respectively. The instrument was calibrated prior to data acquisition using Agilent's reference standard solution, which provided accurate masses within 5 ppm. A reversed phase, linear gradient separation was performed at a flow rate of 1.0 mL min^{-1} on a C18 column (Waters Acquity UPHLC peptide BEH C18, 1.7 μm particle size, 50 mm x 2.1 mm I.D.) thermostatted at 45°C and gradient from 95% solvent A (0.1% formic acid in water) to 99% solvent B (0.1% formic acid in acetonitrile) in 4 minutes, followed by 1 minute re-equilibration.

MS/MS:

A portion of analysis was done was done at Merck facilities at Rahway, NJ by Ryan Cohen; the other portion was done at Princeton University by Chuhan Zong. Tandem MS analyses were performed either on an Agilent 1260 HPLC coupled with a 6530 Q-ToF or on an Agilent 1290 HPLC coupled with a 6520 Q-ToF mass spectrometer. Both mass spectrometers were operated under electrospray ionization in positive polarity (+ESI). Reverse phase C18 gradient separation conditions were used (mobile phase A = 0.1% formic acid in water and mobile phase B = 0.1% formic acid in acetonitrile). Spectra were acquired in profile mode with collision energies ranging from 25-90 V and nitrogen (N_2) as collision gas.

General Procedure for Fmoc Solid-Phase Peptide Synthesis:

Peptides were synthesized manually using standard protocols[42] with either Rink amide resin (0.25-mmol scale) for Ser-Phe-Arg-Tyr-Ala-Glu or 2-ClTrt resin (0.175-mmol scale) for *N*-methylated peptides. Fmoc-protected amino acids were sequentially coupled on the resin with an excess of HBTU and DIEA for 2 hrs at room temperature. Fmoc was deprotected with a 20% v/v piperidine in DMF solution for 20 min.

Macrocyclization of Linear Peptide, Ser-Phe-Arg-Tyr-Ala-Glu, on Solid Support:

To peptide, Ser-Phe-Arg-Tyr-Ala-Glu, containing O-allyl protected glutamic acid on solid support (50 mg, 0.69 mmol/g) was added a 3 mL solution of 20 mg $Pd(PPh_3)_4$ and 72 µL phenylsilane in DCM. This solution was incubated on a shaker at room temperature for 40 minutes. The resin was washed with DCM (3 X 2 min). The above reaction was repeated, and then the resin was washed with DCM (3 X 2 min), MeOH (3 X 2 min), and DMF (3 X 2 min). The palladium catalyst was removed from the resin by washing with DIEA in DMF (2% v/v) (2 X 2 min) and then with DMF (2 X 2 min). Next, the peptide was treated with 20% piperdine in DMF to remove the Fmoc

protection. Macrocyclization was achieved by exposing the resin to DIC, HOAt, and a catalytic amount of DMAP in DMF on a shaker for 15 hrs. The solution was drained, and the resin was washed with DMF. To confirm complete macrocyclization, the Kaiser test[43] was performed, which showed yellow bead coloration indicating absence of free amino groups. The peptide was then cleaved from the resin by 2 hrs incubation with a solution of TFA:TIPS:water (95/2.5/2.5, v/v/v). Resin was removed by filtration, and the resulting solution was evaporated. The peptide was purified by reversed-phase semi-preparative chromatography.

_N_-Methylation of Linear Peptides on Solid Support:

N-methylated residues were synthesized on solid support using a reported method involving *o*NBS protection.[43] A 155 mg *o*-NBS-Cl and 230 µL trimethylpyridine solution in 5 mL DMF was added to the resin and shaken at room temperature for 30 minutes, which was then filtered and washed with DMF. The protection step was then repeated, followed by 5X washing with DMF. The resin was soaked in THF and then dried under vacuum. A solution of 230 mg triphenylphosphine and 2.5 µL MeOH in 2.5 mL THF was added to the resin and shaken for 5 minutes. The resin was then treated with a solution of 175 µL DIAD in 2.4 mL THF (added portion-wise at 500 µL each time and shaken for 10 minutes). The resin was then filtered and washed with DMF. Mitsunobu conditions were repeated once more. Finally, *o*NBS deprotection proceeded by soaking the resin in DMF, followed by 10 minute exposure at room temperature while shaking in a 5 mL DMF solution containing 122 µL 2-mercaptoethanol and 131 µL DBU, and then filtering and washing with DMF. Deprotection was repeated once.

Macrocyclization of _N_-methylated Peptides in Solution:

Peptides were cleaved from the resin with a 20% v/v HFIP solution in DCM. The DCM was evaporated, and linear peptides were purified by chromatography. Cyclization occurred with 20

mg HOAt, 20 μL DIC in dry DMF per 2 mg of peptide. Side chain protecting groups were then deprotected by adding a solution of TFA:TIPS:water (95/2.5/2.5, v/v/v) and incubating the solution for 2 hrs. *N*-methylated macrocyclic peptides were then purified by reversed-phase semi-preparative chromatography.

Albusnodin Expression and Purification:

This portion of work was done by Chuhan Zong at Princeton University. For a typical heterologous production of albusnodin, *Streptomyces* strains were cultured in GYM medium with total volume of 1 L. All strains were cultured in 5 x 1000 mL Erlenmeyer flasks (each flask contained 200 mL of GYM medium) that were treated with 1 mL dimethyldichlorosilane in toluene to minimize the culture from sticking to the glass wall and a stainless steel spring was placed at the bottom of each flask to improve aeration. Starting cultures were inoculated with 2 x 10^9 spores in 150 mL 2YT medium and incubated at 30 °C with orbital shaking (350 rpm) for 8 h. The germlings were harvested by centrifugation at 2,000 x *g*, resuspended in 25 mL of GYM medium, dispersed by a quick sonication pulse, and used to inoculate production cultures at a starting OD_{450} of 0.005. Production cultures were incubated at 30°C with orbital shaking (250 rpm) for an additional 7 days. The cell lysate and supernatant of the samples were separated by centrifugation at 8,000 x *g* at 4 ˚C for 20 minutes. The cells were lysed in methanol. The insoluble fraction was removed by centrifugation at 8,000 x *g* at 4 ˚C for 10 minutes, and the methanol-soluble fraction was dried using rotary evaporation and then resuspended in 500 μL of 50% acetonitrile and 50% water solution. The supernatant was extracted using a Strata C8 column (Phenomenex 6 mL size). The crude extract was eluted from the column using methanol, and dried using rotary evaporation, and then reconstituted in 500 μL of a solution of acetonitrile / water (1:1, v/v).

Benenodin-1 ΔC5 Expression and Purification:

This portion of work was done by Chuhan Zong at Princeton University. Benenodin-1 ΔC5 was obtained as previously described.[32] Briefly, the lasso peptide was expressed using *E. coli* BL21 cells with the plasmid, pCZ35, containing the lasso peptide gene cluster. Cultures were grown at 37°C in 100 µg/mL LB-ampicillin broth. Cells were harvested by centrifugation and then lysed in MeOH. The lysate containing the peptide was purified by solid phase extraction followed by reversed-phase semi-preparative chromatography.

General Procedure for Activation of Ser, Thr, Cys, Glu, and Hse Residues within Macrocyclic Peptides and Lasso Peptides:

Per cleavage site in 1 mg of peptide, excess DSC (25 mg), DIEA (20 µl of a 20% vol/vol solution in DMF), and a crystal of DMAP were added to macrocyclic or lasso peptides in 450 µl DMF in a 5 mL round bottom flask. The reaction mixture was agitated using an orbital shaker for 16 hours to modify desired cleavage residues. The mixture was then removed and DMF was evaporated.

General Procedure for Ring-Opening of Macrocyclic and Lasso Peptides:

0.45 mL of phosphate buffer (pH 7.6) was added to the activated macrocyclic or lasso peptide. This mixture was then agitated on an orbital shaker at 37°C for 48 hours and then checked by LC-MS to ensure complete cleavage.

Procedure for Protease Incubation.

Incubation of macrocyclic and lasso peptides with chymotrypsin and/or trypsin were conducted using a 50 mM NH_4HCO_3 digestion buffer. The peptide substrates were treated with 50 µl of 0.1 µg/µL protease (sequencing grade) in a total volume of 800 µL for each protease alone or 1600 µL for both trypsin and chymotrypsin together. Incubation was carried out for 48 hours at 37°C and samples were then analyzed by LC-HRMS.

3.8 REFERENCES

1. Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A.J.; Raj, M. *Angew. Chem.* **2018**, DOI: 10.1002/anie.201801299.
2. Koehn, F. E.; Carter, G. T.; *Nat. Rev. Drug Discovery*, **2005**, *4*, 206.
3. Winn, M.; Fyans, J.K.; Zhuo, Y.; Michlefield, J. *Nat. Prod. Rep*. **2016**, *33*, 317.
4. Zorzi, A.; Deyle, K.; Heinis, C. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24.
5. Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Chen, V.; Crews, P.; Jacobson, M. P.; Lokey, R. S. *Future Med. Chem*. **2015**, *7*(16): 2121.
6. Cheung, W. L.; Pan, S. J.; Link, A. J. *J. Am. Chem. Soc.*, **2010**, *132* (8), 2514.
7. Colgrave, M. L.; Craik, D. J. *Biochemistry*, **2004**, *43* (20), 5965.
8. Getz, J. A.; Rice, J. J.; Daugherty, P. S. *ACS Chem. Biol.*, **2011**, *6* (8), 837.
9. Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. *ACS Chem. Bio.* **2007**, *2*, 625.
10. Howell, S. M.; Fiacco, S. V.; Takahashi, T. T.; Jalali-Yazdi, F.; Millward, S. W.; Hu, B.; Wang, P.; Roberts, R. W. *Sci. Rep.* **2014**, *4*, 6008.
11. Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1*, 80.
12. Simpson, L. S.; Kodadek, T. *Tetrahedron Lett*. **2012**, *53*, 2341.
13. King, J. L.; T. H. Jukes. *Science*, **1969**, *164*, 788.
14. Kaiser, R., Metzka, L. *Anal. Biochem.* **1999**, *266*, 1.
15. *Butler, M.S. J. Nat. Prod* **2004**, *67* (12), 2141.
16. Ngoka, L. C. M.; Gross, M. L. *J. Am. Soc. Mass Spec.* **1999**, *10*, 732.
17. Siegel, M. M.; Huang, J.; Lin, B.; Tsao, R.; Edmonds, C. G. *Bio. Mass Spec*. **1994**, *23*, 186.
18. Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. *J. Comb. Chem.* **2003**, *5*, 33.
19. Eckart, K.; Schwarz, H.; Tomer, K. B.; Gross, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 6765.

20. Schilling, B.; Wang, W.; McMurray, J. S.; Medzihradszky, K. F. *Rapid Commun. Mass Spec.* **1999**, *13*, 2174.

21. Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. *Acc. Chem. Res*. **2015**, *48*, 1909.

22. Caboche, S.; Pupin, M.; Leclére, V.; Fontaine, A.; Jacques, P.; Kucherov, G.; *Nucleic Acids Res.* 2008, *36*, D326.

23. De Marco, R.; Tolomelli, A.; Campitiello, M.; Rubini, P.; Gentilucci, L. *Org. Bio. Chem.* **2012**, *10*, 2307.

24. Elashal, H. E.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, ***8***, 117.

25. (a) Elashal, H. E.; Raj, M. *Chem. Commun.* **2016**, *52*, 6304. (b) Zong, C.; Cheung-Lee W. L.; Elashal, H. E.; Raj, M.; Link, A.J. *Chem. Commun.* **2018**, *54*, 1339.

26. Shinde, N. V.; Dhake, A. S.; Haval, K. P. *Orient. J. Chem*. **2016**, *32* (1).

27. Tanabe, K.; Taniguchi, A.; Matsumoto, T.; Oisaki, K.; Sohma, Y.; Kanai, M. *Chem. Sci*, **2014**, *5*, 2747.

28. Thorner, J.; Emr, S. D.; Abelson, J. N. *Methods Enzymol*. **2000**, *326*, 200.

29. Milovic´, N. M.; Kostic´, N. M., *J. Am. Chem. Soc.* **2003**, *125*, 781.

30. Chatterjee, J.; Laufer, B.; Kessler, H. *Nat. Protoc.* **2012**, *7*, 432.

31. Maksimov, M. O.; Link, A. J. *Nat. Prod. Rep*. **2012**, *29*, 996.

32. Zong, C.; Wu, M. J.; Qin, J. Z.; Link, A. J. *J. Am. Chem. Soc*. **2017**, *139*, 10403.

33. *Chekan, J. R.; Koos, J. D.; Zong, C.; Makimov, M. O.; Link, A. J.; Nair, S. K. J. Am. Chem. Soc.* **2016**, *138* (50), 16452.

34. Maksimov, M. O.; Link, A. J. *J. Am. Chem. Soc*. **2013**,*135*, 12038.

35. Maksimov, M. O.; Koos, J. D.; Zong, C.; Lisko, B.; Link, A. J. *J. Biol. Chem***. 2015**, *290*, 30806.

36. Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. *J. Am. Chem. Soc.* **2006**, *128* (23), 7486.

37. Zong, C.; Cheung-Lee, W.; Elashal, H. E.; Raj, M.; Link, A. J. *Chem. Commun.* **2018**, *54*, 1339.

38. Ogino, H.; Ohata, K. *Inorg. Chem.* **1984**, *23* (21), 3312.

39. Shen, Y. X.; Gibson, H. W.; *Macromolecules*, **1992**, *25* (7), 2058.

40. Griffiths, K. E.; Stoddart, J. F. *Pure Appl. Chem.* **2008**, *80* (3), 485.

41. *Xue, M.; Yang, Y.; Chi, X.; Yan, X.; Huang, F. Chem. Rev.* **2015**, *115* (15), 7398.

42. Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.

43. Chatterjee, J.; Laufer, B.; Kessler, H. *Nat. Protoc.* **2012**, *7*, 432.

CHAPTER 4: OXAZOLIDINONE MEDIATED SEQUENCE DETERMINATION OF ONE BEAD ONE COMPOUND CYCLIC PEPTIDE LIBRARIES

4.1 ABSTRACT

Cyclic peptides are attractive therapeutics for drug discovery due to high proteolytic stability, permeability, and high binding affinity. The therapeutic ability of cyclic peptides prompted the synthesis of cyclic peptide libraries by split and pool methods for High Throughput Screening (HTS) against various biological targets. A major obstacle associated with this approach is the sequencing difficulty of hit cyclic peptides obtained by HTS using traditional sequencing methods such as NMR, MS/MS, and enzymatic digestion. Therefore, we developed a one bead one compound (OBOC) dual ring opening/cleavage approach for sequencing of hit cyclic peptides by selective modification of a serine residue to an oxazolidinone moiety within the cyclic peptide and as a linker to allow for the liberation of the protected linear peptide into solution. Formation of the oxazolidinone moiety increases the susceptibility of the amide bond at the N-terminus of serine towards hydrolysis and leads to the opening of a cyclic peptide to its linear counterpart and simultaneous release of the resulting linear peptide from the solid support. The resulting linear peptide in solution was then sequenced in one minute by a liquid chromatography, ion mobility spectrometry, tandem mass spectrometry (LC-IMS-MS/MS) method, which provided faster sequencing of the ring opened peptides due to its inherently high peak capacity. To extend this approach, libraries of cyclic peptides (4-10 amino acids) containing different combinations of amino acids were successfully synthesized and sequenced with 98% overall accuracy, demonstrating the robustness and versatility of this methodology.

4.2 CHAPTER OBJECTIVES

In this chapter, *a simple and versatile one bead one compound approach for determining the sequence of libraries of cyclic peptides on solid support is described.* A methodology based on the selective-activation of the backbone amine chain of cyclic peptide libraries was developed, which will facilitate the sequencing of HIT peptides post biological screening. This technique can be revolutionary in decreasing the amount of time spent in determining HIT sequences, thereby paving the way for rapid development of lead compounds into therapeutic drugs.

The strategy entails the selective activation of the amide backbone chain at serine, threonine, and/or cysteine residues for the ring opening and sequencing of HIT cyclic peptides. The significance of this approach is that encoding tags or toxic reagents are not required for determining the sequence, which makes it distinct from the currently used methods. Moreover, this method is compatible with free amino acid side chains and does not lead to any side reactions. Another major advantage of this approach is that disulfide bonds remain stable under the reaction conditions, thus it has a potential to be applicable for the sequencing of cyclic peptides rich in disulfide bonds. In addition, the IMS-MS technique was used in an unprecedented manner for the rapid sequencing of cyclic peptides from the crude mixture; this portion of the study was done by Ryan Cohen. The use of IMS-MS reduces the time for sequencing of one cyclic peptide by 20-fold as compared to traditional MS/MS methods.

4.3 GRAPHICAL ABSTRACT

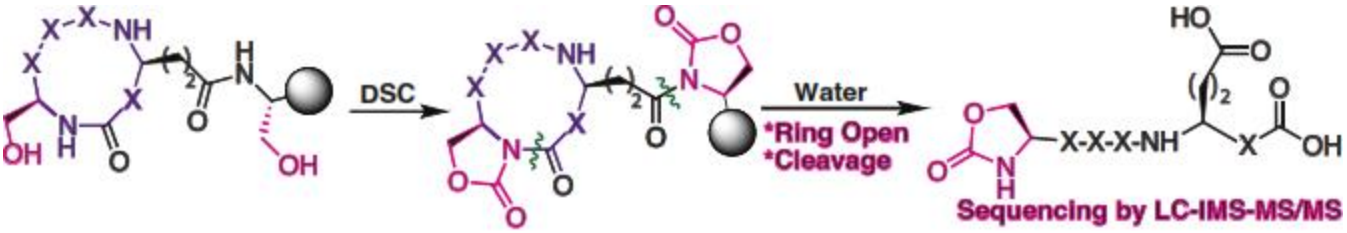


Figure 4.1 Graphical abstract for *Oxazolidinone Mediated Sequence Determination of One Bead One Compound Cyclic Peptide Libraries*. Figure adopted with permission from: Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Raj, M. *Org. Lett.* **2018**, *20* (8), 2374.[1]

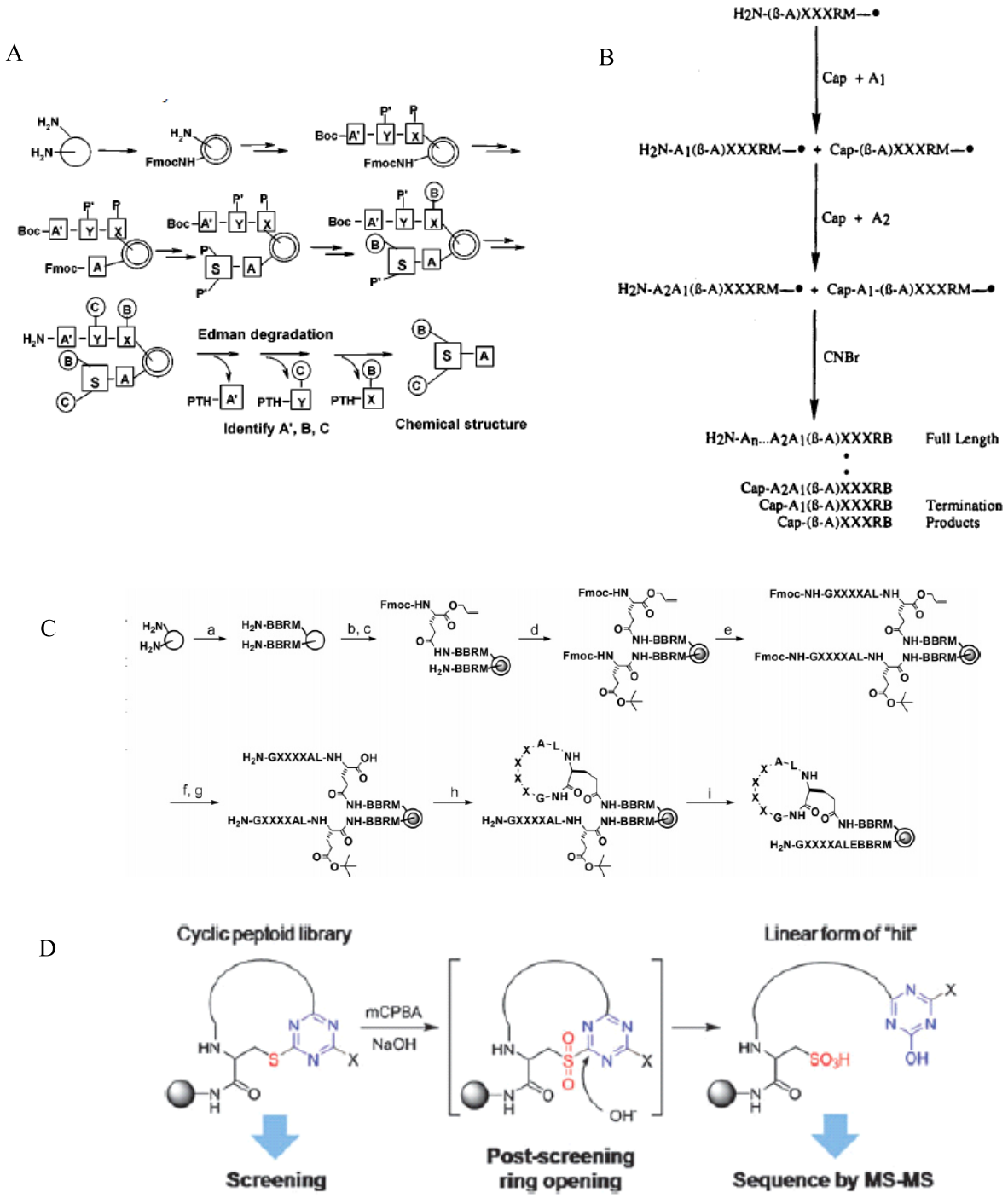
4.4 INTRODUCTION

Cyclic peptides are of considerable interest due to their powerful potential as therapeutic agents.[2-5] Compared to their linear counterparts, cyclic peptides exhibit high receptor-binding affinity, specificity, and stability due to their high conformational rigidity.[6] Owing to their unique character, cyclic peptides are suitable tools for drug discovery. The therapeutic ability of cyclic peptides has prompted the synthesis of cyclic peptide libraries by split and pool methods[7] followed by high-throughput screening against biological targets to determine "hit" peptides.[8] However, their utility is hindered by the challenges that accompany the sequencing of the cyclic peptides. While enzymatic cleavage is one of the most utilized methods for peptide sequencing, it is incompatible with cyclic peptides; cyclic peptides are resistant to enzymatic cleavage by 20-300 folds as compared to their linear conterparts.[9] Mechanistically, the decrease in cleavage results from both a lower rate of cleavage (k_{cat}) and a weaker interaction with proteases (K_m).[10] Since cyclic peptides lack a free N-terminal amine, they cannot be sequenced via Edman's degradation.[11] Meanwhile, tandem mass spectroscopy (MS/MS), another widely employed sequencing method, affords complex fragmentation patterns, which are difficult to interpret.[12-16] Many attempts have been made in developing methods to overcome the challenges associated with sequencing hit macrocyclic peptides. The first utilized method in sequencing of hit cyclic peptides consisted of one-bead one-compound (OBOC) libraries that employed genetic encoding tags to record

synthesis history (Scheme 4.1).[17-19] Nonetheless, OBOC is mainly compatible with natural amino acid containing peptides therefore its use in sequencing of cyclic peptides is of limited practicality. Meanwhile, the ladder synthesis involves capping small amounts of the peptide after each coupling before the final macrocyclization of the peptide to track the cyclic peptide sequence (Scheme 4.1). However, this approach is limited by interference from the small linear peptides (encoding tags) during the screening process.[20]

To combat the abovementioned limitations, Pei and co-workers developed a one-bead two-compound (OBTC) method consisting of segregated bilayer beads allowing each bead to constitute two elements.[21,22] A cyclic peptide is exposed on the surface for screening, while the inside contains its linear counterpart for sequencing (Scheme 4.1). The disadvantage of this approach is that the cyclization step cannot be monitored due to the presence of the linear peptide tag.[21] Another major limitation associated with this method is that it cannot be used for sequencing cyclic peptides obtained from nature. More recently, strategies based on opening the macrocyclic ring post screening have been developed as an alternative to bead-based assays.[23] These methods rely on incorporation of specific moieties into the cyclic peptide that can be chemically cleaved, resulting in the linear form which allows sequencing of hit cyclic peptides via MS/MS.

In this manner, Lim and co-workers introduced an alkylthioaryl bridge into cyclic peptides, which is cleavable upon oxidation of the thioether by mCPBA to generate the linear peptide.[24] However, the oxidation reaction is not specific to the thioether, consequently, other functional groups such as guanidines and primary amines are also oxidized during the ring-opening reaction leading to complicated MS/MS fragmentation. Moreover, this method utilizes toxic cyanogen bromide (CNBr) for the cleavage of cyclic peptides.

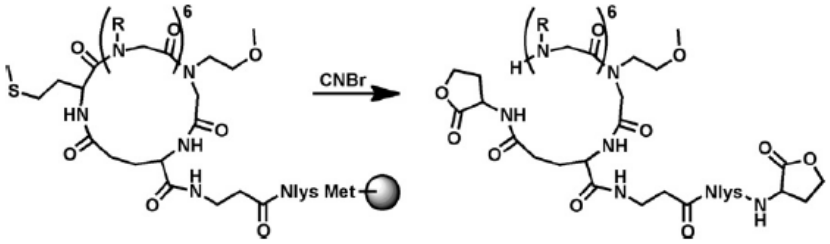


Scheme 4.1 Literature methods for sequencing of cyclic peptides. A) One-bead-one-compound method (Adopted from reference [22] with permission) B) Ladder synthesis method (Adopted from reference [20] with permission) C) One-bead-two-compound method (Adopted from reference [21] with permission) D) Alkylthioaryl bridge method (Adopted from reference [24] with permission)

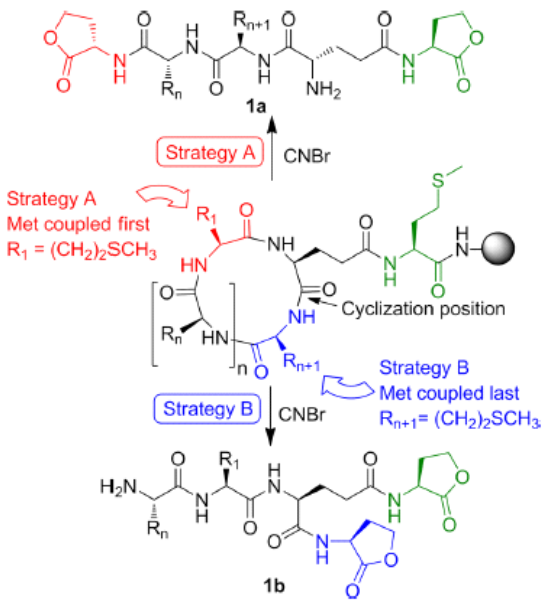
Recently, the Kodadek[25] and Biron[26] groups reported the use of a methionine residue for linearization and cleavage of cyclic peptides/peptoids from the solid support which is appealing in the sense that a natural amino acid is incorporated into different sequences for ring opening (Scheme 4.2). Nonetheless, complicated inverse peptide synthesis, side reactions, long cleavage times, and the use of toxic CNBr for cleavage detract from the attractiveness of this method.

Most reported methods require aggressive chemical reagents or post-screening reactions that could lead to side-chain modifications. Conversely, Biron and coworkers have reported the utilization of a photo-cleavable moiety, β-amino acid 3-amino-3-(2-nitrophenyl)propionic acid (ANP), for the ring opening and release of the peptide from the resin (Scheme 4.2).[27] Although this approach is reagent free and utilizes UV radiation for the opening of the cyclic ring, it remains limited by the formation of a mixture of various linear peptide side products, leading to complicated HPLC and MS/MS data interpretation. Another major limitation is that this method leads to the oxidation of the Met and Trp side chains. In addition, the synthesis and screening of these cyclic peptides must be done in the dark to prevent opening of the macrocycle or cleavage from the resin.[27] Due to the sensitivity of the ANP moiety towards light, the cyclic peptides containing the ANP unit cannot be stored for long term use. A major advantage of the ring-opening strategy is that the same chemical entity is displayed inside and on the surface of the bead thus eliminating the risk of interference by coding tags during the screening, which is in contrast to the OBTC and ladder synthesis methods. Moreover, the cyclization step can be easily monitored to ensure a complete conversion to the cyclic peptide. Based on these strategies, our research goal was related to the development of an efficient and single-step ring-opening approach that would be compatible with free amino acid side chains and would not require the use of harsh chemical reagents.

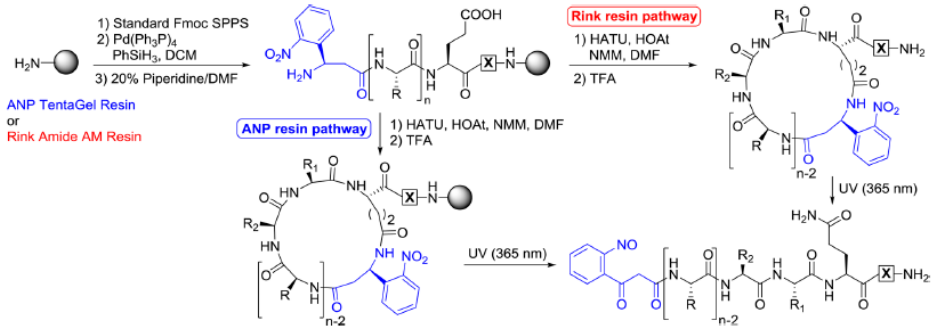
A



B



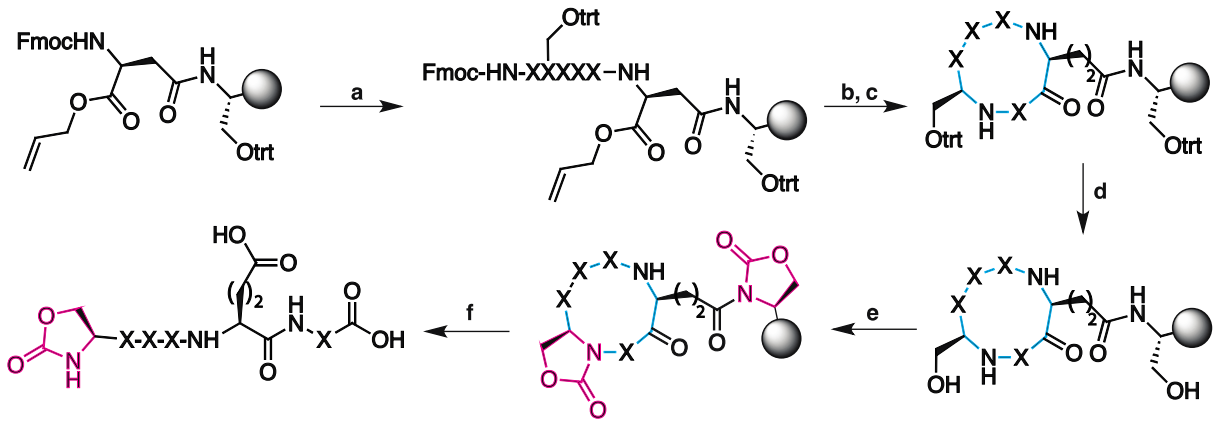
C



Scheme 4.2 Literature methods for cleavage of cyclic peptides. A) Kodadek's CNBr method (Adopted from reference [25] with permission) B) Biron's CNBr method (Adopted from reference [26] with permission) C) ANP method (Adopted from reference [27] with permission)

This chapter reports a simple and convenient approach for fast sequence determination of

cyclic peptides from OBOC libraries that overcome the challenges associated with existing methods. The methodology entails the selective activation of the amide backbone chain at the serine residue for the cleavage and sequencing of hit cyclic peptides. The significance of this approach is that encoding tags or toxic reagents are not required for determining the sequence, which makes it distinct from the currently used methods. Moreover, this method is compatible with free amino acid side chains and does not lead to any side reactions.



Scheme 4.3 Rationale for the sequencing of cyclic peptides by ring-opening/cleavage strategy. Reagents: (a) standard Fmoc/HBTU chemistry; (b) $Pd(PPh3)_4$, phenylsilane; 20% piperdine (c) HOAt, DIC; (d) TFA (25%); (e) DSC, DIEA, DMAP; and (f) H_2O: ACN (1:1), DIEA. Scheme adopted from reference (1) with permission.

4.5 RESULTS AND DISCUSSION

This approach utilizes two serine residues, one as part of the linker and one within the macrocycle (Scheme 4.3). The strategy involves the conversion of Ser to an oxazolidinone moiety, thereby making the amide bond susceptible to cleavage at the Ser N-terminus under hydrolytic conditions. Earlier attempts at the formation of this moiety on Fmoc/Boc protected peptides led to the formation of dehydroalanine.[28] Our group recently reported the application of this moiety in the synthesis of thioesters.[29] This methodology simultaneously converts cyclic peptides into their linear counterparts and releases them from the solid support. The linear peptides obtained from

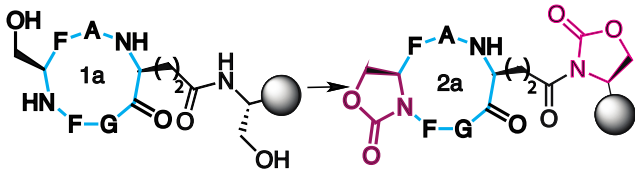
solid support could then be easily sequenced by mass spectrometry. Scheme 4.3 illustrates the ring-opening/cleavage reaction for sequencing of cyclic peptides. Notably, only those resin-bound peptides that complete both the activation and displacement steps would be released. Thus, the released linear peptides should exhibit high purity, eliminating the need for extensive purification.

To implement this strategy, a serine (trt) trityl residue with a selectively removable side-chain protecting group (pg) as a linker was anchored onto the solid support. To this linker, a linear peptide with another serine (trt) residue, with a selectively removable side-chain pg, was synthesized on solid support by standard Fmoc chemistry.[29] After removal of the allyl[21] and Fmoc protecting groups, the peptides were cyclized on resin by using HOAt and DIC.[30] The Kaiser test was used to monitor the macrocyclization reaction qualitatively.[31] After macrocyclization, the (trt) groups from both the serine residues were selectively removed with TFA/DCM (1:3).[32]

Initially, we assessed the efficiency of the activation of serine residues within a macrocyclic peptide and as a linker. The serine residues were activated with N,*N*'-disuccinimidyl carbonate (DSC) to generate two five-membered oxazolidinone rings on a model cyclic peptide *cyc*(GFSFAE)-S **1a** (Table 4.1). To optimize the formation of two oxazolidinone moieties on the model cyclic peptide *cyc*(GFSFAE)-S on solid support, various reaction conditions such as time, temperature, and equivalents of DSC were explored (entries 1-6, Table 4.1). A mixture of mono oxazolidinone moiety-containing cyclic peptide **2a'** and desired bis oxazolidinone moieties-containing cyclic peptide *cyc*(GF-*Oxd*-FAE)-*Oxd* **2a** was obtained by using 10-15 equiv. of DSC at room temperature in 15 hrs (entries 1-2, Table 4.1).

Table 4.1 Optimization of the activation of cyclic peptide *cyc*(Gly-Phe-Ser-Phe-Ala-Glu)-Ser-

Rink **1a** on solid support.[a]



Entry	DSC (equiv.)	Temp. (°C)	Time (h)	Conversion (%)[b]	
				2a	**2a'**
1	10	RT	15	50	50
2	15	RT	15	75	25
3	20	RT	15	>99	-
4	20	RT	6	90	10
5	20	65	6	ND	ND
6	20	65	15	ND	ND

[a]Cyclic peptide **1a** (50 mg) was reacted with DSC (10-20 equiv.), DIEA (10-20 equiv.) and a crystal of DMAP in DMF for (3-15 hrs). [b]Conversion to desired two oxazolidinone-containing cyclic peptide cyc(GF-Oxd-FAE)-Oxd **2a** and one oxazolidinone-containing cyclic peptide **2a'** was calculated from the absorbance at 220 nm using HPLC after release from solid support ND = not detected. Table adopted from reference (1) with permission.

When 20 equiv. of DSC was used at room temperature, the conversion to **2a** was 90% in 6 h, whereas when the reaction was run for 15 hrs, the conversion increased to 99% (entries 3-4, Table 4.1). Similar reactions at high temperature (65 °C) gave a mixture of various unidentified side-products (entries 5-6, Table 4.1). To confirm the absence of the unmodified cyclic peptide, a small amount of the resin was cleaved and analyzed by HPLC and MS (Figure 4.2).

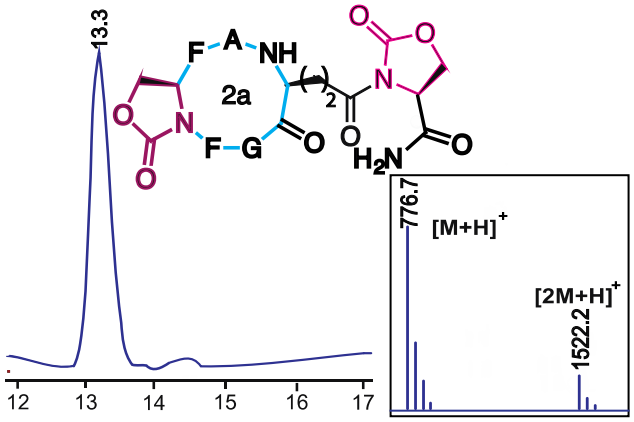
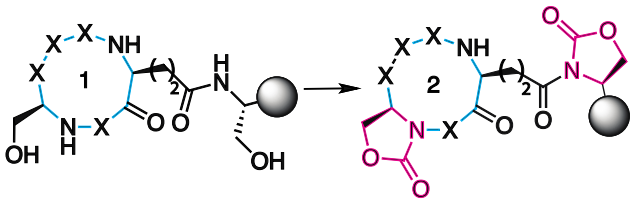


Figure 4.2 HPLC and MS of *cyc*(GF-*Oxd*-FAE*)-Oxd* **2a.** Figure adopted from reference (1) with permission.

Subsequently, serine activation was explored on solid support with various cyclic peptides of different sequences and different ring sizes (4-9 membered) (**1b-1f**, entries 1-5, Table 2). It is noteworthy that formation of the activated oxazolidinone moiety is independent of the nature of the amino acid preceding serine, and high conversion to the oxazolidinone moiety was obtained with cyclic peptides of various sizes. Moreover, the results in Table 2 showed that activation of serine does not depend on its position in the cyclic peptide. This is in contrast to the methionine-based ring cleavage method where a C-terminal Met lead to formation of two homoserine lactone residues after ring opening, resulting in difficult sequencing of the linearized peptide, while an N-terminal Met led to moderate cyclization efficiency.

Table 4.2 Substrate scope of the activation of cyclic peptides **1b-1i** on solid support to **2b-2i**.[a]



Entry	Cyclic Peptides **1**	Conversion (%)[b]
1	*cyc*(SFFAE)-S **1b**	85
2	*cyc*(SGVFAE)-S **1c**	89
3	*cyc*(GSFAE)-S **1d**	92
4	*cyc*(GFKSYGLE)-S **1e**	90
5	*cyc*(AFSIGFE)-S **1f**	88
6	*cyc*(AFSFE)-T **1g**	95
7	*cyc*(GFCE)-C **1h**	92
8	*cyc*(AMPFISFPE)-C **1i**	85

[a]Cyclic peptide **1b-1i** (50 mg) was reacted with DSC (20 equiv.), DIEA (20 equiv.) and a crystal of DMAP in DMF for (15 hrs) at RT. [b]Conversion to desired two oxazolidinone-containing cyclic peptide **2b-2i** was calculated from the absorbance at 220 nm using HPLC after release from solid support. Table adopted from reference (1) with permission.

Next, the peptide backbone activation on cyclic peptides was investigated with threonine

cyc(AFSFE)-T **1g**, whose side chain functionality exhibits strong similarity to serine. As expected, threonine underwent smooth cyclization under the optimized reaction conditions and generated a methylated oxazolidinone moiety *cyc*(AF-*Oxd*-FE)-^{Me}*Oxd* **2g** (entry 6, Table 4.2). In a similar manner, cysteine-containing cyclic peptides *cyc*(GFCE)-C **1h** and *cyc*(AMPFISFPE)-C **1i** generated thiazolidinone-activated cyclic peptides *cyc*(GF-*Thz*-E)-*Thz* **2h** and *cyc*(AMPFI-*Oxd*-FPE)-*Thz* **2i** under the reaction conditions (entries 7 and 8, Table 4.2).

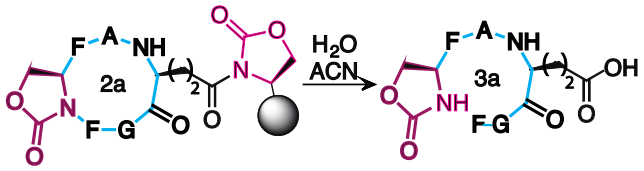
The stability of the oxazolidinone moiety on the resin-bound macrocyclic peptide was also evaluated. The resin-bound oxazolidinone moiety containing cyclic peptide was found to be stable in a desiccator for longer than a month, which is in contrast to the previously reported photo-cleavable ANP method.[27]

Nucleophilic displacement of the oxazolidinone moieties under basic hydrolytic conditions simultaneously opened the cyclic peptide to its linear variant and released the resulting linear peptide acid from the solid support. Initially, the reaction was performed at room temperature in water:ACN (1:1, v/v) without base (entry 1, Table 4.3). This resulted in release of 10% linear peptide acid **3a** from the resin after 15 hrs at room temperature.

To optimize the cleavage of the activated cyclic model peptide *cyc*(GF-*Oxd*-FAE)-*Oxd* **2a** to its linear counterpart and simultaneous release of the resulting linear peptide *Oxd*-FAEGF **3a** into solution, various reaction conditions, such as base and temperature, were investigated (entries 2-6, Table 4.3 and Figure 4.3). Hydrolytic cleavage under neutral conditions indicates that macrocyclic ring opening occurs more rapidly than peptide release from the resin (entries 1-2, Table 4.3). This was confirmed by acidic cleavage of the peptide from the resin after treatment with hydrolytic conditions for brief amount of time. Another advantage of this approach is that

only the serine activated peptide chain is released from the resin, thus delivering a highly pure linear peptide into the solution. The resin was then filtered, and the solvent was evaporated to obtain linear peptide *Oxd*-FAEGF **3a**, which was then analyzed by HPLC and MS/MS (Figure 4.3).

Table 4.3 Optimization of the reaction conditions for dual ring-opening/cleavage of cyclic peptide **2a** to its linear counterpart **3a** on solid support.[a]



Entry	Base	Temp. (ºC)	Time (h)	Conv.(%)[b]
1	-	RT	15	10
2	DIEA	RT	4	70
3	DIEA	RT	8	97
4	DIEA	RT	15	99
5	DIEA	65	4	>99
6	DIEA	65	2	99

[a]Activated cyclic peptide **2a** (50 mg) was treated with water:ACN (1:1) 1 mL and DIEA (40 µL). [b]Conversion and release of resulting linear peptide Oxd-FAEGF **3a** into the solution was calculated from the absorbance at 220 nm using HPLC. Table adopted from reference (1) with permission.

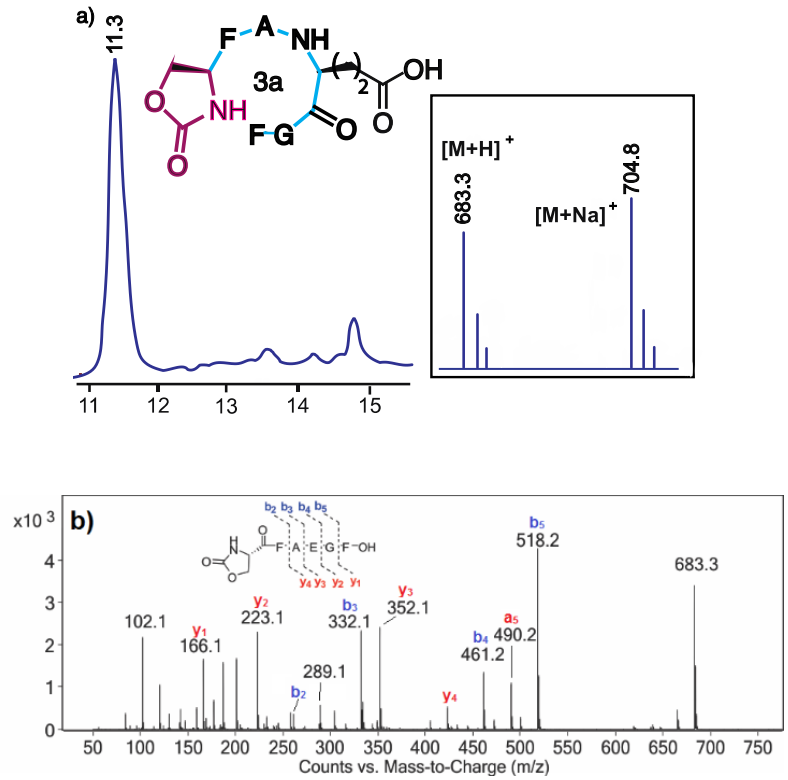
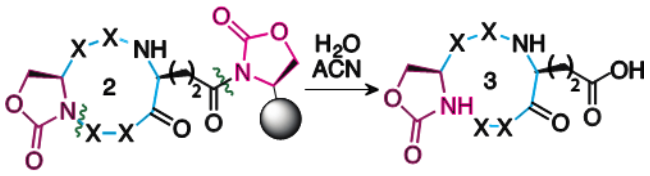


Figure 4.3 a) HPLC and MS of linear peptide *Oxd*-FAEGF **3a,** b) MS/MS spectra of peptide *Oxd*-FAEGF **3a.** Figure adopted from reference (1) with permission.

Next, various cyclic peptides with different amino acids and sizes were evaluated for the ring-opening and release strategy on solid support (Table 4.4).

Table 4.4 Substrate scope of the ring-open/cleavage strategy on activated cyclic peptides **2b-2i**.[a]



Entry	Activated Cyclic Peptides **2**	Conversion (%)[b]
1	*cyc*(*OxdFFAE*)-*Oxd* **2b**	85
2	*cyc*(*Oxd-GVFAE*)-*Oxd* **2c**	89

3	*cyc*(G-*Oxd*-FAE)-*Oxd* **2d**	92
4	*cyc*(GFK-*Oxd*-YGLE)-*Oxd* **2e**	90
5	*cyc*(AF-*Oxd*-IGFE)-*Oxd* **2f**	88
6	*cyc*(AF-*Oxd*-FE)-^{Me}*Oxd* **2g**	95
7	*cyc*(GF-*Thz*-E)-*Thz* **2h**	92
8	*cyc*(AMPFI-*Oxd*-FPE)-*Thz* **2i**	85

[a]Activated cyclic peptide **2b-2i** (50 mg) was treated with water:ACN (1:1) 1 mL and DIEA (40 µL). [b]Conversion and release of resulting linear peptide **3b-3i** into the solution was calculated from the absorbance at 220 nm using HPLC. Table adopted from reference (1) with permission.

It was noted that opening of the oxazolidinone activated cyclic peptides **2** on solid support to their linear counterparts **3** is independent of the nature of the amino acid residues preceding serine. For example, cyclic peptides with bulky amino acid residues next to serine such as Tyr(tBu), Phe, Lys(Boc), and a cyclic peptide with β-branched residues such as Ile generated linear peptides cleanly in high yields (entries 4-8, Table 4.4). Noteworthy, backbone activation is also possible at Thr and Cys due to similar side chain functionality exhibited by these resides. The possibility of modification at Ser, Cys, and/or Thr residues provides an additional advantage to the utility of this method affording multiple cleavage sites to aid in the sequencing process. Moreover, possible activation at any of the abovementioned residues allows variation in the generation of OBOC libraries generated; the peptide library is not limited to include only Ser residues. In the presence of Thr or Cys, sequencing post screening becomes feasible as demonstrated in peptides **2g, 2h, and 2i** which contain methylated oxazolidinone and thiazolidinone (entries 6-8, Table 4.4) By demonstrating an expanded substrate scope, oxazolidinone formation efficiently mimics enzyme function by providing multiple cleavage sites under mild conditions.

After the ring-opening/cleavage of a cyclic peptide, one oxazolidinone residue remained attached to the N-terminus of the linear peptide chain. Since only the activated serine residue can be released from the solid support after hydrolysis, the result is release of a highly pure linear peptide into solution. In contrast to other methods, we did not observe any oxidation of the unprotected side chain residues after dual ring opening/cleavage of cyclic peptides from the solid support.

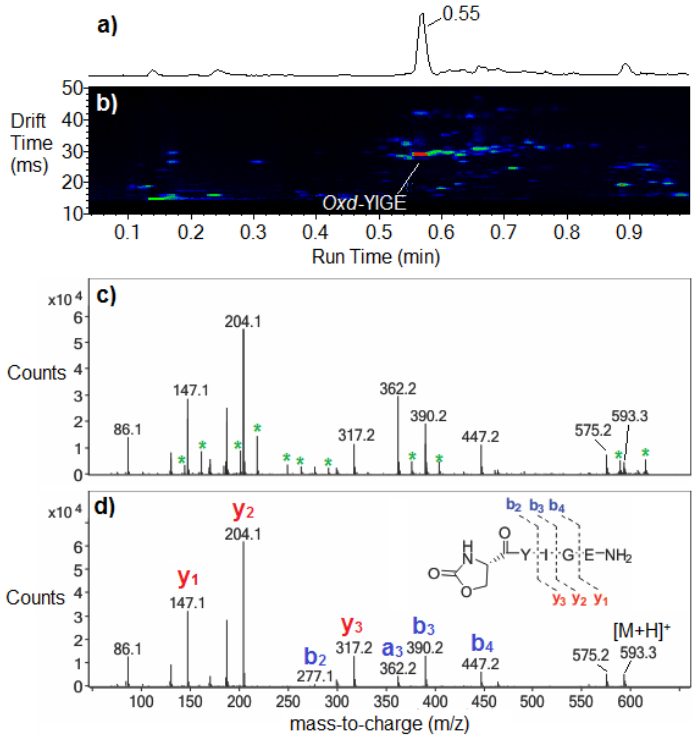


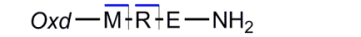
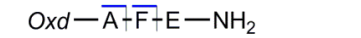
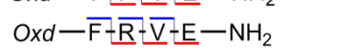
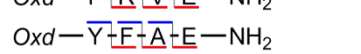
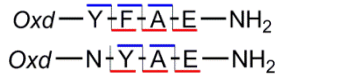
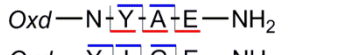
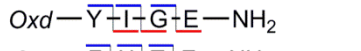
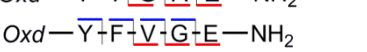
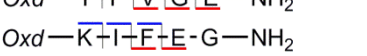
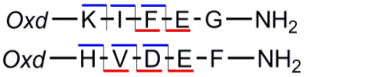
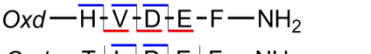
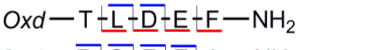
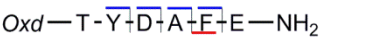
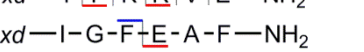
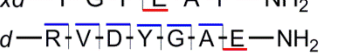
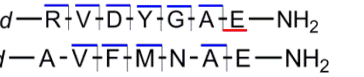
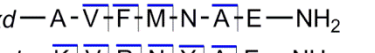
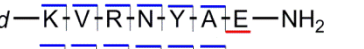
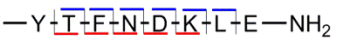
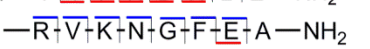
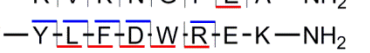
Figure 4.4 a) LC-MS separation of *Oxd*-YIGE, b) 2D separation – ion mobility drift time versus total ion chromatogram, c) MS/MS spectrum without quadrupole selection of parent ion (interfering signals marked with green asterisks), and d) MS/MS spectrum including ion mobility filtering to remove chemical interference. Figure adopted from reference (1) with permission.

Sequencing was then accomplished by tandem mass spectrometry on a quadrupole time-of-flight (Q-ToF) instrument incorporating a two-dimensional liquid chromatography, ion mobility spectrometry separation (LC-IMS-MS/MS) (Figure 4.3).[33] Peak capacity can be significantly increased by pairing IMS with ultra-high-pressure liquid chromatography; thereby,

allowing for a ballistic reversed gradient to be employed. The overall separation time was just one minute, which would enable the analysis of large combinatorial cyclic peptide libraries. Both HRMS identification of the precursor ion and MS/MS were performed in a single run by passing all ions through the quadrupole to the collision cell for collision induced dissociation (CID). For traditional LC-MS/MS instrument setups, this would result in significant interference from impurities and system contaminants; however, here IMS filtering in this method allows for the production of a clean MS/MS spectra (Figure 4.3).[34] Spectra were then interpreted either manually or in most cases via de novo sequencing software (*viz*., PEAKS Studio).

To demonstrate the compatibility of our strategy with OBOC libraries, a small cyclic peptide library of varying amino acid composition and ring sizes ranging from 4 to 10 amino acids was prepared on Tenta Gel resin with Glu(OAll) as the first amino acid on the solid support. The library was prepared by split-and-pool synthesis using standard Fmoc/tBu solid-phase peptide chemistry. Following Glu(OAll), the next positions within the peptide library were filled by a random combination of 20 *L*-amino acids. Serine was added randomly within varying positions within the peptide sequence, and the peptide was cyclized after selective deprotection of Fmoc and OAllyl groups as described above. This was followed by Ser activation using DSC. After removal of the side-chain protecting groups, 34 beads were randomly selected and individually hydrolyzed under basic conditions at 65 ^{o}C on a heated shaker for 4 hrs. The resulting crude linear peptides were then sequenced by the LC-IMS-MS/MS method (Figure 4.4, Table 4.5).[35-37] These results demonstrated the wide scope of the ring-opening OBOC strategy in sequencing of cyclic peptides of varying size and composition.

Table 4.5 High throughput screening of various cyclic peptides with varying ring sizes and composition on beads.

Entry	Cyclic Peptide Sequence	MS/MS Sequence Coverage[a]	Calculated[b] $[M+H]^+$	Observed[b] $[M+H]^+$	Conversion (%)[c]
1	SMRE		547.2293	547.2303	>99
2	SAFE		478.1932	478.1918	>99
3	STFVE		607.2722	607.2716	>99
4	SFRVE		662.3257	662.3265	>99
5	SYFAE		641.2566	641.2561	>99
6	SNYAE		608.2311	608.2335	>99
7	SYIGE		593.2566	593.2584	99
8	SFNTE		622.2467	622.2461	>99
9	SVRFAE		733.3628	733.3651	>99
10	SFRYAE		797.3577	797.3595	>99
11	SNYAAE		679.2682	679.2676	>99
12	SYIGAE		664.2937	664.2947	99
13	SYFVGE		726.3093	726.3115	99
14	SKIFEG		705.3566	705.3576	>99
15	SHVDEF		758.3104	758.3105	>99
16	STLDEF		736.3148	736.3156	>99
17	SRQFEA		762.3529	762.3531	99
18	SKIFAGE		776.3937	776.3944	99
19	SAHDVGE		739.3006	739.2987	99
20	STYDAFE		857.3312	857.3307	>99
21	STFKRVE		891.4683	891.4704	>99
22	SIGFEAF		795.3672	795.3683	>99
23	SRVDYGAE		921.4061	921.4053	>99
24	SAVFMNAE		893.3822	893.3811	>99
25	SKVRNYAE		496.2514[d]	496.2513[d]	99
26	SRYQVANE		496.2332[d]	496.2330[d]	99
27	SKARFGVE		459.7432[d]	459.7430[d]	99
28	SAQVFRFTE		1109.5374	1109.5393	99
29	SYTFNDKLE		1141.5160	1141.5174	>99
30	SRVKNGFEA		1032.5221	1032.5232	99
31	SYLFDWREK		1268.6058	1268.6067	>99
32	SGYVKDFRAE		1196.5695	1196.5672	99
33	SAYKFGPSAE		541.2511[d]	541.2519[d]	99
34	SATKFESRVE		589.7937[d]	589.7917[d]	>99

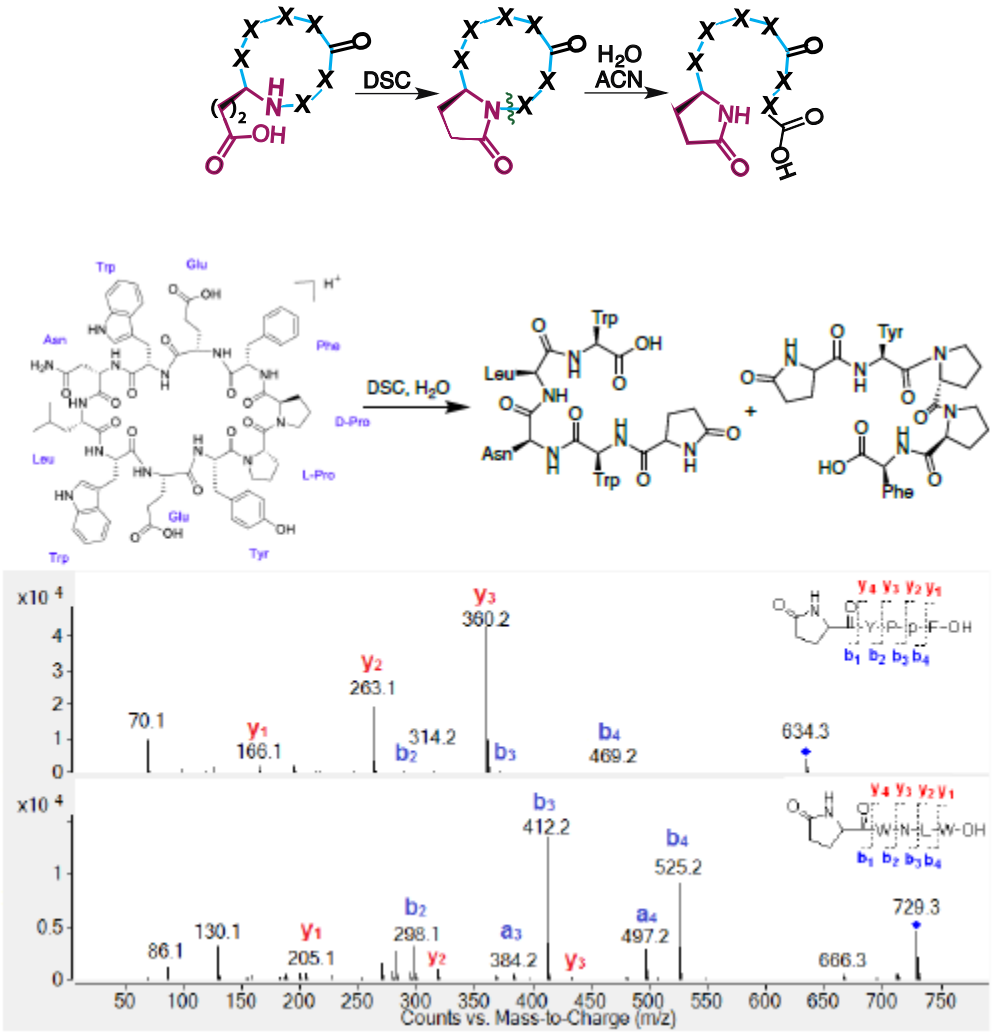
[a]MS/MS sequence coverage indicates observed y and b ions as right (red) and left (blue) cleavages, respectively

[b]Calculated and observed $[M+H]^+$ are for ring-opened peptides

[c]Reaction conditions: Conversion was determined by cleavage of the pre-treated resin under TFA cleavage conditions

[d]Doubly charged ions, $[M+2H]^2$ Table adopted from reference (1) with permission.

Another unique feature of this approach is that peptides with glutamic acid residues can also be selectively modified to a pyro-glutamyl amide moiety with the backbone peptide chain leading to the ring opening of the cyclic peptide at the N-terminus of Glu under hydrolytic conditions (Scheme 4.4). We synthesized and sequenced a potent p53 cyclic peptide analogue which is implicated in the p53/MDM2 interaction[38] and contains two glutamic acid residues.



Scheme 4.4 Rationale for the sequencing of cyclic peptides containing glutamic acid by the formation of pyro-glutamyl amide moiety. Reagents: DSC, DIEA, DMAP; followed by H_2O: ACN (1:1), DIEA. Scheme adopted from reference (1) with permission.

4.6 CONCLUSIONS

We developed a general one bead one compound (OBOC) method for sequencing cyclic peptides obtained from combinatorial libraries. Our method is compatible with both on-bead and solution-phase library screening. As compared to encoding tag methods, one bead two compound (OBTC) approaches or other one bead one compound (OBTC) methods, the major advantage of this method is that it is direct, without the need to tag each bead, which makes the synthesis of libraries of cyclic peptides easier. By excluding the encoding tag component, chances of interferences during the screening process are greatly reduced, if not abolished. Our method has high substrate scope and ring opening of cyclic peptides can be achieved at serine, threonine, cysteine, and glutamic acid residues, which is a significant advantage over the current OBOC approaches that rely either on a methionine residue or an unnatural entity. This method is compatible with free amino acid side chains due to the fact that they remain unmodified under reaction conditions.[39] Furthermore, our method does not require any special residues and is stable to light and air thus eliminating the requirement of special handling conditions. The activated peptides can be stored in a desiccator for a month without any degradation. Therefore, this methodology can be readily applied in any chemical or biochemical laboratory.

In addition, we utilized the higher separation power of ion mobility spectrometry paired with ultra-high-pressure liquid chromatography to develop a high throughput tandem MS sequencing method. Precursor ion selection was found to be unnecessary as interfering signals could be removed via ion mobility filtering, and this allowed for compound libraries to be analyzed in a single run. Using this method, over 1000 cyclic peptides can be sequenced in one day. Moreover, interpretation of the data has the capability of being automated as the spectra were found to be compatible with de novo sequencing software. The main attributes of our method are

its simplicity, facile synthesis without major side product formation, high substrate scope, and high-throughput capability. This robust method will better enable the search for cyclic peptides as novel therapeutics.

4.7 EXPERIMENTAL SECTION

General: All commercial materials (Sigma-Aldrich, Fluka, and Novabiochem) were used without further purification. All solvents were reagent or HPLC (Fisher) grade. Diethyl ether, CH_2Cl_2, and DMF were obtained from a dry solvent system (passed through column of alumina) and used without further drying. Conversions were obtained by comparison of HPLC peak areas of products and starting materials. HPLC was used to monitor reaction progress.

Materials: Fmoc-amino acids were obtained from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide resin was obtained from ChemPep Inc (Wellington, Florida). N,N,N′,N′-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate **(**HBTU) was obtained from CreoSalus (Louisville, Kentucky). N,*N′*-Disuccinimidyl carbonate (DSC) was obtained from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP) was obtained from Merck KGaA (Darmstadt, Germany). N,*N*-Dimethylformamide (DMF) was obtained from Macron Fine Chemicals (Center Valley, Pennsylvania). Acetonitrile, N,*N*- Diisopropylethylamine (DIEA), and N,*N*'-diisopropylcarbodiimide (DIC) were purchased from Novabiochem (EMD Millipore Corporation, Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR (100 Matsonford Road Radnor, Pennsylvania). Diethyl ether was obtained from Sigma-Aldrich (St. Louis, Missouri). Water was purified using a Millipore Milli-Q water purification system.

Purification: Semi-preparative chromatography was performed using a Beckman Coulter equipped with a System Gold 168 detector, a 125P solvent module, and a 10 mm C-18 reversed-phase column. All separations involved a mobile phase of 0.1% formic acid (v/v) in water (solvent A) and 0.1% formic acid (v/v) in acetonitrile (solvent B). The semi-preparative HPLC method employed a linear gradient of 0–80% solvent B over 30 minutes at ambient temperature with a flow rate of 3.0 mL min^{-1}. The separation was monitored by UV absorbance at both 220 and 254 nm unless otherwise noted.

HPLC:

Peptide compositions were evaluated by high performance liquid chromatography (HPLC) on an Agilent 1200 series HPLC equipped with a 4.6x150mm (5µm) C-18 reversed-phase column. All separations used mobile phases of 0.1% formic acid (v/v) in water (solvent A) and 0.1% formic acid (v/v) in acetonitrile (solvent B). A linear gradient of 0–80% solvent B in 30 minutes at room temperature with a flow rate of 1.0 mL min^{-1} was used. The eluent was monitored by UV absorbance at 220 nm unless otherwise noted.

LC-MS:

Mass spectrometry to check reaction mixtures was performed using an Agilent 1100 Series HPLC with MSD VL mass spectrometer using positive polarity electrospray ionization (+ESI).

HRMS:

Analysis was conducted by Ryan Cohen at Merck facilities in Rahway, NJ. HRMS data were recorded on an Agilent 1290 UHPLC with 6560 ion mobility Q-ToF mass spectrometer using positive polarity electrospray ionization (+ESI).

LC-IMS-MS/MS for sequencing peptides:

Analysis was conducted by Ryan Cohen at Merck facilities in Rahway, NJ. An Agilent 1290

UHPLC and 6560 ion mobility Q-ToF mass spectrometer using positive polarity electrospray ionization (+ESI) were used to sequence peptides. LC-IMS-MS conditions are provided in **Table 4.6**. Collision induced dissociation (CID) was used to generate MS/MS spectra using N_2 as collision gas and energies ramped from 20 to 60 V.

Table 4.6. LC-IMS-MS conditions

Parameter	Condition
Column	Waters Acquity CSH C18, 1.7 μm particle size, 50 mm x 2.1 mm I.D.
Column temperature	50°C
Mobile phase A	0.1% formic acid in water
Mobile phase B	Acetonitrile
Gradient	time (min) / % A / % B 0.00 / 95 / 5 0.65 / 1 / 99 0.75 / 1 / 99 0.76 / 95 / 5 1.00 / 95 / 5
Injection volume	1.0 μL
Flow rate	1.0 mL/min
Ionization	(+)ESI using Agilent's jet stream source
Voltages	capillary = 3.5 kV, fragmentor = 175 V, nozzle = 1100 V
Drying gas	N_2, 300°C, 10 L/min
Nebulizer	45 psig
Sheath gas	N_2, 275°C, 12 L/min
Mass range	m/z 150 to 3200
IM trap	fill time = 40,000 μs, release time = 150 μs
Acquisition rate	frames = 1/s, IM transients = 19/frame, max drift time = 50 ms

Fmoc Solid-Phase Peptide Synthesis:[40] Peptides were synthesized manually on a 0.25 mmol scale using Rink amide, Wang, Chem Matrix and TentaGel resins. Fmoc was deprotected using 20% piperidine–DMF for 20 min to obtain a deprotected peptide-resin. Fmoc-protected amino acids (0.75 mmol) were sequentially coupled on the resin using HBTU (0.75 mmol) and DIEA (0.75 mmol) for 2 hrs at room temperature. Peptides were synthesized using standard protocols.[1] Peptides were cleaved from the resin using a cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water for 2 hrs. The resin was removed by filtration and the resulting solution was

concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by semi-preparative chromatography.

General Procedure for Cyclization of Linear Peptides on Solid Support: To a peptide containing Glu(Oallyl)OH amino acid on the solid support (50 mg, 0.69 mm/g) was added a solution of tetrakis(triphenylphosphine)palladium(0), $Pd(PPh_3)_4$ (20 mg), phenylsilane (72 µL) in DCM (3 mL), and the peptide was left on the shaker for 40 minutes. The resin was washed with 10 mL DCM (3 X 2 min). The above reaction was repeated, and then the resin was washed with 10 mL DCM (3 X 2 min), 10 mL MeOH (3 X 2 min), and 10 mL DMF (3 X 2 min). The palladium catalyst was removed from the resin by washing it with 10 mL DIEA (2 X 2 min), followed by washing with 10 mL DMF (2 X 2 min). Next, the peptide was treated with 20% piperdine/DMF to remove the Fmoc protecting group. Macrocyclization was achieved by exposing the resin to 50µL DIC, 50 mg HOAt, and a catalytic amount of DMAP in DMF per 50mg of resin on a shaker for 15 hrs. The solution was drained, and the resin was washed with DMF. To confirm the complete macrocyclization, the Kaiser test was performed. Yellow bead coloration indicated absence of free amines, while blue coloration indicated incomplete macrocyclization. To further confirm formation of cyclic peptides, the resin was cleaved using a cocktail of 95:2.5:2.5 trifluoroacetic acid:triisopropyl silane:water (v/v/v) for 2 hrs. The resin was then removed by filtration, and the resulting solution was evaporated. Cyclic peptides were purified by semi-preparative chromatography and analyzed by MS.

General Procedure for Serine Activation to Cyclic Urethane Moiety within Cyclic Peptides on Solid Support: The trityl protecting groups on serine residues were selectively removed by exposing the peptide on solid support to a 30% TFA in DCM 10 mL solution (3 X 5 min). Both serine groups (one as a linker and the other within the macrocycle) were activated on solid support

(25-100 mg, 0.25-0.69 mm/g) using a solution of DSC (25 equiv.), DIEA (25 equiv.), and a catalytic amount of DMAP in DMF on shaker for 17 hrs. The solution was drained, and the resin was washed with DMF, followed by peptide cleavage using a cocktail of 95:2.5:2.5 trifluoroacetic acid:triisopropyl silane:water (v/v/v) for 2 hrs. The resin was removed by filtration, and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by semi-preparative chromatography for MS analysis.

General Procedure for Dual Ring-Opening/Cleavage of Cyclic Peptides on Solid Support for MS/MS Sequencing: A solution of 1 mL H_2O:ACN (1:1, v/v) and 40 µL of DIEA was added to the resin containing the activated cyclic urethane-containing macrocyclic peptide. This reaction was left on a heated shaker at 65 °C for 4 hrs. The resin was then filtered, followed by solvent removal under vacuum. The resulting ring-opened peptide acid was then sequenced by LC-IMS-MS/MS.

General procedure for macrocyclization in solution phase. Peptide was cleaved from the resin with a 20% v/v HFIP solution in DCM. The DCM was evaporated, and the linear peptide was purified by chromatography. Cyclization occurred with 20 mg HOAt, 20 µL DIC in dry DMF per 2 mg of peptide. Side chain protecting groups were then deprotected by adding a solution of TFA:TIPS:water (95/2.5/2.5, v/v/v) and incubating the solution for 2 hrs. The macrocyclic peptide was then purified by reversed-phase analytical chromatography.

General Procedure for the activation of glutamic acid in cyclic peptide in solution phase. To a 5-mL round-bottom flask containing peptide 2-20mg (1 equiv.) in 0.5-2 mL dimethylformamide (DMF) was added a solution of DSC (15 equiv.), DIEA (15 equiv.) and crystal of DMAP in DMF

(0.2-0.5 mL). The mixture was stirred at room temperature for 10 hrs. The reaction was concentrated under vacuum and resulting peptide was dissolved in 1:1 mixture of water and acetonitrile and purified by HPLC.

General Procedure for ring-opening of cyclic peptide in solution phase. To an activated cyclic peptide, 1 mL phosphate buffer (pH 7.4) was added. The reaction was stirred at 25 °C and monitored by analytical HPLC at regular intervals. The reaction mixture was lyophilized, purified by HPLC and isolated.

4.8 REFERENCES

1. Elashal, H.; Cohen, R. D.; Elashal, H. E.; Raj, M. *Org. Lett.* **2018**, *20* (8), 2374.
2. Joo, S. H. *Biomolecules & Therapeutics* **2012**, *20*, 19.
3. McGregor, D. P. *Curr. Opin. Pharma.* **2008**, *8*, 616.
4. Roberts, R. W.; Szostak, J. W. *Proc. Nat. Acad. Sci.* **1997**, *94*, 12297.
5. Millward, S. W.; Takahashi, T. T.; Roberts, R. W. *J. Am. Chem. Soc.* **2005**, *127*, 14142.
6. Yudin, A. K. *Chem. Sci.* **2015**, *6*, 30.
7. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82.
8. Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411.
9. Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. *ACS Chem. Bio.* **2007**, *2*, 625.
10. Howell, S. M.; Fiacco, S. V.; Takahashi, T. T.; Jalali-Yazdi, F.; Millward, S. W.; Hu, B.; Wang, P.; Roberts, R. W. *Sci. Rep.* **2014**, *4*, 6008.
11. Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1*, 80.

12. Eckart, K.; Schwarz, H.; Tomer, K. B.; Gross, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 6765.

13. Schilling, B.; Wang, W.; McMurray, J. S.; Medzihradszky, K. F. *Rapid Commun. Mass Spec.* **1999**, *13*, 2174.

14. Ngoka, L. C. M.; Gross, M. L. *J. Am. Soc. Mass Spec.* **1999**, *10*, 732.

15. Siegel, M. M.; Huang, J.; Lin, B.; Tsao, R.; Edmonds, C. G. *Bio. Mass Spec.* **1994**, *23*, 186.

16. Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. *J. Comb. Chem.* **2003**, *5*, 33.

17. Scott, C. P.; Abel-Santos, E.; Wall, M.; Wahnon, D. C.; Benkovic, S. J. *Proc. Nat. Acad. Sci.* **1999**, *96*, 13638.

18. Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. *ACS Chem. Bio.* **2007**, *2*, 625.

19. Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H. *Nat. Chem. Biol.* **2009**, *5*, 888.

20. Youngquist, R. S.; Fuentes, G. R.; Lacey, M. P.; Keough, T. *J. Am. Chem. Soc.* **1995**, *117*, 3900.

21. Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. *J. Am. Chem. Soc.* **2006**, *128*, 13000.

22. Liu, R.; Marik, J.; Lam, K. S. *J. Am. Chem. Soc.* **2002**, *124*, 7678.

23. Maria, C. M.-C.; Silvana, L. G.; Soledad, L. S.; Juan, M. G.-M.; Rosa, E.-B.; Fernando, A.; Osvaldo, C.; Silvia, A. C. *Curr. Pharma. Biotech.* **2016**, *17*, 449.

24. Lee, J. H.; Meyer, A. M.; Lim, H.-S. *Chem. Commun.* **2010**, *46*, 8615.

25. Simpson, L. S., Kodadek, T. *Tetrahedron Lett*. **2012**, *53*, 2341–2344.

26. Liang, X.; Girard, A.; Biron, E. *ACS Comb. Sci.* **2013**, *15*, 535-540.

27. Vézina-Dawod, S.; Bédard, F.; Porte, K.; Biron, E. *Org. Lett.* **2016**, *18*, 1174.

28. De Marco, R.; Tolomelli, A.; Campitiello, M.; Rubini, P.; Gentilucci, L. *Org. Bio. Chem.* **2012**, *10*, 2307.

29. Elashal, H. E.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2016**.

30. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84.

31. Furka, Á.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pep. Pro. Res.* **1991**, *37*, 487.

32. Wellings, D. A. A., E. *G. B. Academic Press, San Diego,* **1997**, p. 54.

33. Elashal, H. E.; Cohen, R. D.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699.

34. Valentine, S. J.; Kulchania, M.; Barnes, C. A. S.; Clemmer, D. E. *Int. J. Mass Spectrom.* **2001**, *212*, 97.

35. McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int. J. Mass Spectrom.* **2005**, *240*, 301.

36. May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A. *Anal. Chem.* **2014**, *86*, 2107.

37. Baker, E. S.; Clowers, B. H.; Li, F.; Tang, K.; Tolmachev, A. V.; Prior, D. C.; Belov, M. E.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1176.

38. Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2004**, *43*, 2109.

39. Elashal, H. E.; Raj, M. *Chem. Commun.* **2016**, *52*, 6304.

40. Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.

CHAPTER 5: FMOC SOLID PHASE SYNTHESIS OF PROTECTED C-TERMINAL MODIFIED PEPTIDES BY FORMATION OF A BACKBONE CYCLIC URETHANE MOIETY

5.1 ABSTRACT

C-terminally modified peptides are of high significance due to the therapeutic properties that accompany various C-terminal functional groups and the ability to manipulate them for further applications. In this chapter, a universal solid phase strategy for the synthesis of various C-terminal modified peptides is reported. The method is independent of the type of resins, linkers, and unnatural moieties typically needed for C-terminal modifications. The technique proceeds by the modification of a C-terminal serine to a cyclic urethane moiety which results in the activation of the backbone amide chain for the nucleophilic displacement by various nucleophiles to generate C-terminally modified acids, esters, N-alkyl amides, alcohols, and hydrazides. This cyclic urethane technique (CUT) also provides a general strategy for synthesis of C-terminal protected peptides which can be used for convergent synthesis of large peptides. The C-terminal protecting groups can be cleaved by hydrolysis to release the free peptide.

5.2 CHAPTER OBJECTIVES

This chapter describes *a universal method for the solid phase synthesis of various C-terminal modified peptides which is independent of the nature of resins, linkers, and unnatural moieties typically needed for C-terminal modifications*. This strategy is highly versatile and generates peptides with various C-terminal end groups such as acids, esters, alcohols, hydrazides, and N-alkyl amides. *Another unique feature of this methodology is that it provides a semi-permanent protecting group at the C-terminus which allows for various modifications in solution; hydrolysis can be performed to remove the protecting group and release the free peptide post-*

modification if desired.

C-terminal modified peptides are of high significance because of the therapeutic properties that accompany various C-terminal functional groups. As a result, there is a necessity for an effective solid-phase synthesis approach for the preparation of peptides with various C-terminal functional groups. Currently, these modifications are generally attainable by using special resins that are appropriately designed to allow the anchoring of the peptide's C-terminus and generate a specific end group at the C-terminus upon cleavage. Other methods include: use of unnatural linkers or handles, side-chain or backbone amide anchoring, and inverse solid-phase peptide synthesis. Despite their obvious attractiveness, all these methodologies have specific disadvantages that need to be addressed. This strategy provides a universal method that relies on a natural amino acid (e.g., serine), which is modified to provide a cyclic urethane moiety that undergoes nucleophilic displacement with various nucleophiles to generate peptides with the desired C-terminal end group.

In this chapter, we describe the Cyclic Urethane Technique (CUT) that relies on the modification of C-terminal serine to the cyclic urethane moiety on solid support and its potential for detachability by various nucleophiles under mild conditions to generate C-terminal modified protected peptides.

5.3 GRAPHICAL ABSTRACT

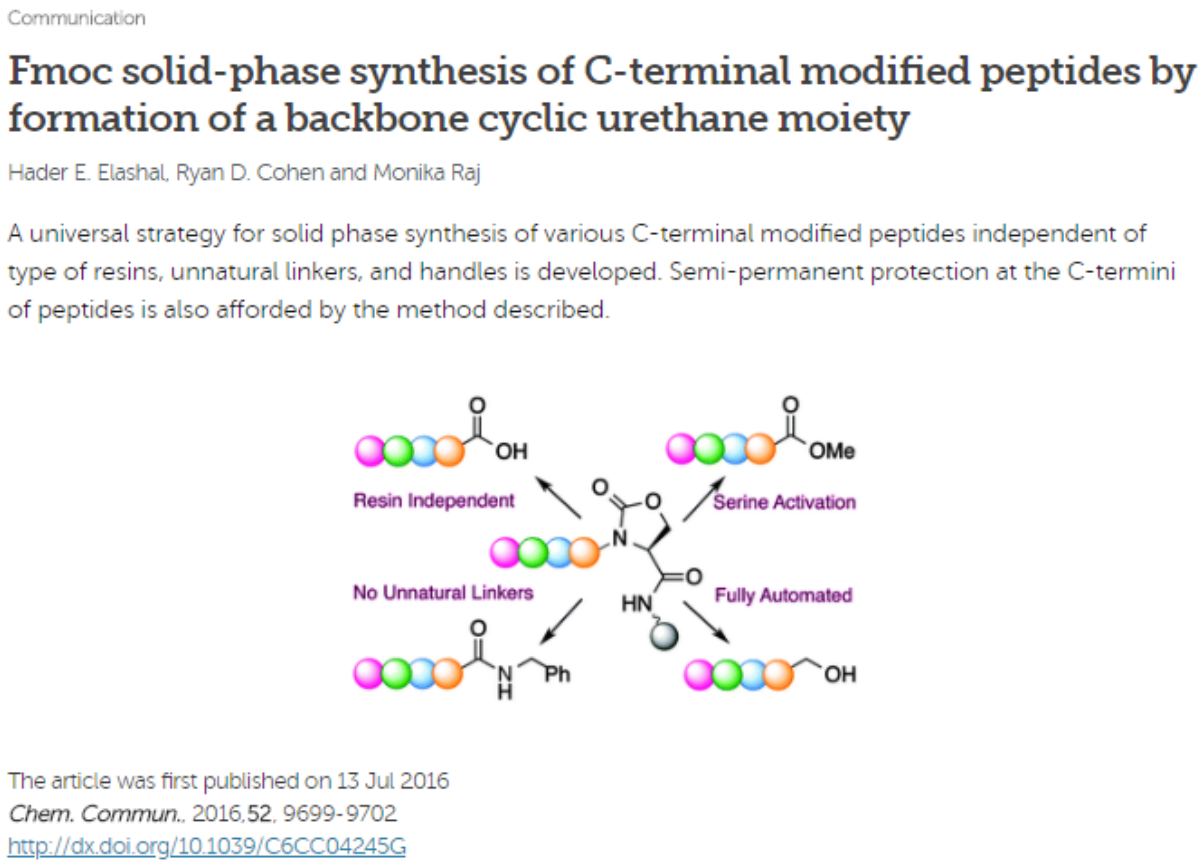


Figure 5.1 Graphical abstract for *Fmoc Solid-Phase Synthesis of Protected C-terminal Modified Peptides by Formation of Backbone Cyclic Urethane Moiety*. Figure adopted with permission from: Elashal, H. E.; Cohen, R. D.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699.[1]

5.4 INTRODUCTION

C-terminal modified peptides such as acids and amides have been extensively used in chemical synthesis of proteins, semi-synthesis, catalysis, and in various therapeutic applications.[2] Other C-terminal peptide modifications of potential interest in therapeutics include esters, N-alkyl amides, and alcohols. These C-terminally modified peptides exhibit a wide range of biological activities; for instance, N-alkyl amides enhance metabolic stability and biological activity.[3,4] Key processes of cell growth and differentiation embody ester groups at their C- terminus.[5] Moreover, peptide alcohols, such as peptaibols are well known antibiotics and exhibit a wide range of

biological activities.[6]

In general, C-terminal modified peptides are synthesized in solution after cleavage of peptides from solid support.[7] For subsequent biological investigations, however, a series of peptide derivatives are often required, which calls for the development of flexible solid phase synthesis techniques. Modifications at the N-terminus of peptides can be easily achieved on a solid support because synthesis of peptides takes place from the C to the N terminus. On the other hand, modification at the C-terminus is a challenge due to its direct attachment to the solid support. Therefore these modifications are generally attainable by using special resins that are appropriately designed to allow the anchoring of the peptide's C-terminus and generate a specific end group at the C-terminus upon cleavage.[8] Due to the therapeutic potential of C-terminally modified peptides, a growing number of studies reported alternative methods such as the use of (i) special linkers between the peptide chain and a solid support which rely on unique oxidative conditions to activate the linker for attack by various nucleophiles.[9] However, this approach is limited due to the undesirable oxidation of sensitive residues such as Cys, Met, and Trp. Other methods include (ii) side-chain or backbone amide anchoring approach, whereby the C-terminal carboxyl group is not involved as an attachment point to the resin.[10] And (iii) inverse solid-phase peptide synthesis (*viz*., from N to C-terminus).[11] One major drawback of this approach is potential epimerization at all coupling stages due to repeated resin-bound carboxyl activation. Another method is (iv) internal resin-capture technique which also involves the use of special linkers or handles.[12-14] All of the reported methods require special resins, unnatural linkers, or handles for the synthesis of C-terminally modified peptides.

This chapter describes the discovery of a novel general strategy that circumvents the use of special resins, linkers, or handles and affords peptides with various types of C-terminal

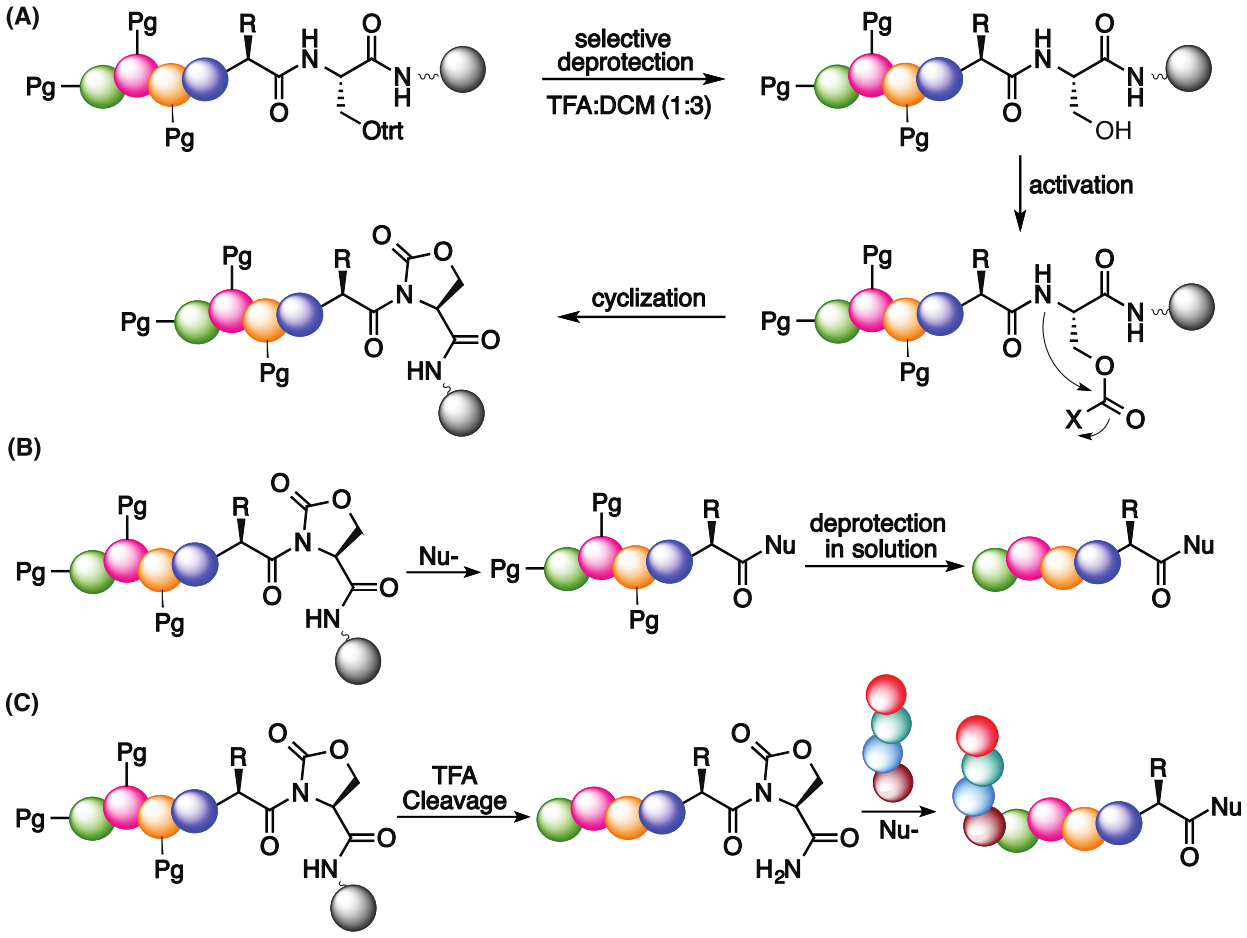
modifications such as acids, esters, N-alkyl amides, and alcohols. Another advantage of this approach is that it provides a semi-permanent protection at the C-terminus of peptides after cleavage of the peptide from the insoluble polymer resin. The semi-permanent protecting group at the C-terminus of a peptide is highly desirable for further manipulation in solution, such as for the synthesis of large peptides.[15]

5.5 RESULTS AND DISCUSSION

This method is based on the activation of the peptide backbone amide bond. We reasoned that the activation of a peptide backbone amide bond by increasing its nucleofugality could render the C-N bond susceptible to nucleophilic attack and provide peptides with various C-terminal modifications depending upon the nucleophile. The backbone amide chain is activated by the formation of a backbone cyclic urethane moiety, which, after nucleophilic displacement, provides the C-terminally modified peptide (Scheme 5.1A and 5.1B). Another application of the cyclic urethane technique (CUT) is to generate a peptide with semi-permanent protection at the C-terminus. Peptides with a C-terminal protecting group can potentially be used for further modifications in solution, followed by the removal of a protecting group under basic hydrolysis conditions (Scheme 5.1C).

Synthesis of C-terminally modified peptides by this strategy entails anchoring of a C-terminal serine residue with a selectively removable side-chain protecting group to a solid support (Scheme 5.1A). Upon peptide chain assembly, the serine side-chain is selectively deprotected. In the first key step, strong activation of the serine's hydroxymethyl group results in the formation of the cyclic urethane moiety on-resin (Scheme 5.1A). Earlier attempts at the formation of the cyclic urethane moiety on Fmoc/Boc peptides were unsuccessful and led to the formation of

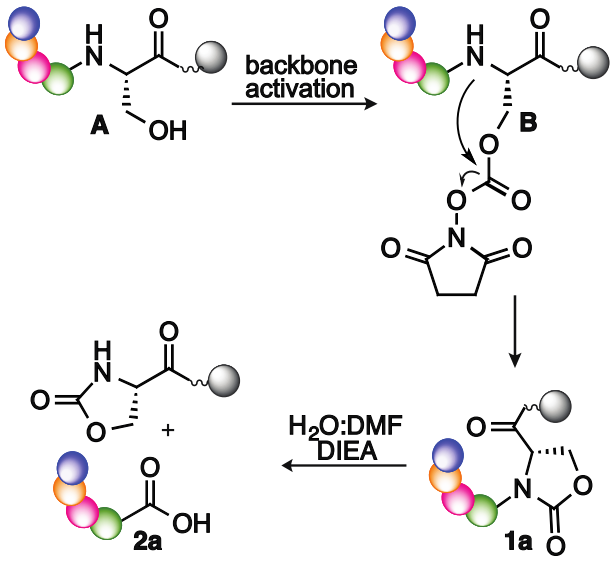
dehydroalanine.[16] Our group recently reported the solution phase synthesis of such cyclic urethane moiety between the side chain of serine and the amidic nitrogen of the backbone peptide chain.[17] In the second key step, nucleophilic displacement of the cyclic urethane moiety by treatment with various nucleophiles releases the protected C-terminal modified peptides from the solid support, which are subsequently deprotected in solution (Scheme 5.1B).



Scheme 5.1 Rationale for the development of cyclic urethane technique (CUT). (a) Methodology for the synthesis of cyclic urethane moiety on solid support. (b) Synthesis of C-terminally modified peptides by nucleophilic displacement of cyclic urethane moiety. (c) Synthesis of protected C-terminal peptides using (CUT) and synthesis of large peptides via convergent approach followed by nucleophilic displacement. Scheme adopted from reference (1) with permission.

This cyclic urethane moiety also acts as a semi-permanent protecting group (pg) for the C-terminus upon cleavage of the peptide from the solid support (Scheme 5.1C). Such protection is of significance due to the potential use of the C-terminally protected peptides for the formation of large peptides via convergent synthesis approach. Finally, the protecting group (pg) at the C-terminus can be easily removed by mild, hydrolysis conditions (Scheme 5.1C). It is noteworthy that special linkers or handles are not required for CUT, which is in contrast to current methods for synthesis of C-terminally protected peptides.[15]

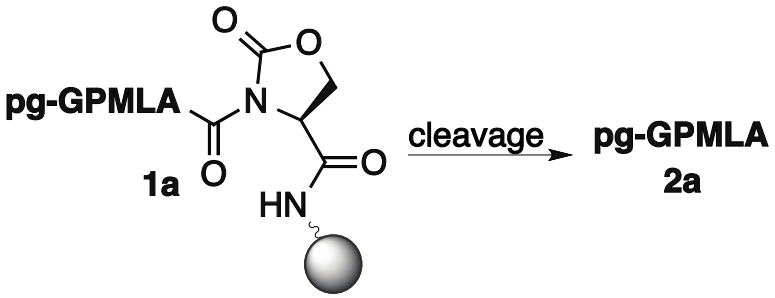
To implement this strategy for the synthesis of a C-terminally modified protected peptide, a model hexamer peptide, Ac-GPMLAS-Rink was synthesized on Rink Amide resin using the SPPS approach with serine(trt) as the C-terminal residue. The trityl (trt) group from serine was selectively removed with TFA/DCM (1:3) to generate peptide **A** (Scheme 5.2). This was followed by activation of serine's hydroxymethyl group with N,*N*'-disuccinimidyl carbonate (DSC) to generate activated intermediate **B**.[17]



Scheme 5.2 Synthesis of peptide acid by using CUT**.** Scheme adopted from reference (1) with permission.

Next, intramolecular nucleophilic attack of the amidic nitrogen of the peptide backbone on the activated intermediate **B** generated a five-membered cyclic urethane ring Ac-GPMLAOxd **1a** (Scheme 5.2). After formation of the cyclic urethane moiety **1a**, nucleophilic release of the protected peptide acid **2a** from the solid support was achieved by carrying out hydrolysis under basic conditions (Scheme 5.2). The solvent was then removed, and the protected peptide acid **2a** was analyzed by HPLC and LCMS (Scheme 5.2). Initially, hydrolysis was carried out at ambient temperature but longer reaction time (> 24hrs) was required to fully release peptide acid **2a** from the resin (entry 1, Table 5.1). To achieve maximum conversion, reaction conditions such as time, solvent, and temperature were optimized (Table 5.1). For the complete liberation of peptide acid **2a** from the solid support, DMF was used as a co-solvent with water due to beneficial resin swelling in DMF (entries 3-5, Table 5.1).

Table 5.1 Optimization of Reaction Conditions for Synthesis of Protected Peptide Acids.[a]



Entry	pg	Solvent	Temp (°C)	Time (h)	Conv (%)[b]
1	Ac	H_2O	RT	15	30
2	Ac	H_2O	50	15	45
3	Ac	H_2O: DMF (4:1)	50	15	77
4	Ac	H_2O: DMF (1:1)	50	15	99
5	Ac	H_2O: DMF (1:1)	65	4	59

6	**Ac**	**H_2O: ACN (1:1)**	**65**	**2**	**99**
7	NH_2	H_2O: ACN (1:1)	65	2	99
8[c]	Fmoc	H_2O: ACN (1:1)	65	2	99

[a]Reaction conditions: Peptide (25 mg, 0.7 mm/g) on solid support was reacted with DSC (100 mg), DIEA (70 uL) and a crystal of DMAP in DMF at room temperature followed by hydrolysis. [b]Conversion to **2a** was calculated from the absorbance at 220 nm using HPLC. The entry in bold represents the optimized reaction conditions for hydrolysis. [c]Mixture of completely protected peptide and Fmoc deprotected peptide was observed. DIEA = Diisopylethylamine, DMAP = 4-(N,*N*-dimethylamino)- pyridine. Table adopted from reference (1) with permission.

To decrease the release time of peptide acid **2a**, hydrolysis was performed on a heated shaker at temperatures ranging from RT to 65 ºC (entries 1-5, Table 5.1). Since it is difficult to remove DMF from the cleaved peptide, a more volatile solvent such as acetonitrile was preferred (entries 6-8, Table 5.1).

The extent of epimerization from formation of the cyclic urethane moiety and nucleophilic displacement was investigated. Diastereomer, Ac-GPML-(D)-A-OH (*D***2a**), was synthesized using the optimized conditions (entry 6, Table 5.1), and peptide acids **2a** and *D***2a'** were subsequently analyzed by HPLC (Figure 5.2). No detectable levels of epimerization were observed.

To evaluate the compatibility of all the amino acids and various protecting groups towards serine's activation as a cyclic urethane moiety and its nucleophilic displacement, various peptides with different sequences and protecting groups were synthesized and screened under the reaction conditions (Table 5.2).

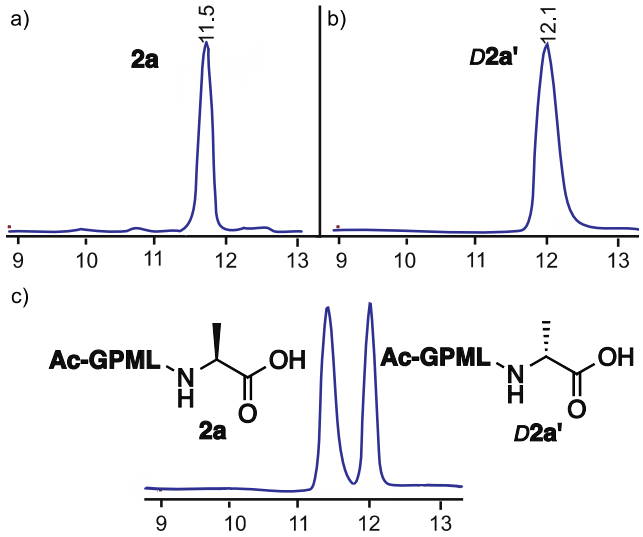
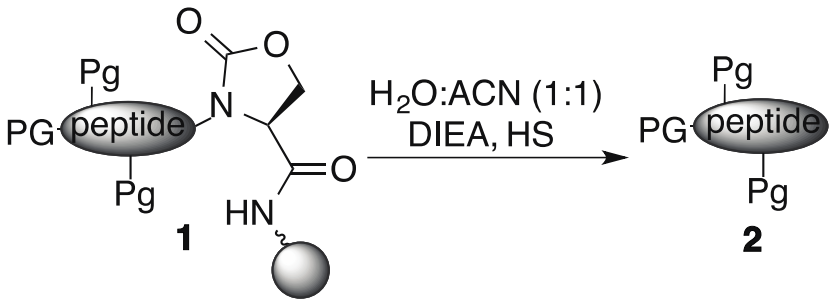


Figure 5.2 (a) HPLC trace of purified peptide acid **2a** Ac-GPML-(L)-A-OH (b) HPLC trace of purified peptide acid diastereoisomer *D***2a'** Ac-GPML-(D)-A-OH (c) HPLC trace of mixture containing both diastereoisomers **2a** and *D***2a'**. Figure adopted from reference (1) with permission.

The results indicate that all the peptides underwent smooth activation to generate activated peptides **1b**-**1k** followed by hydrolysis to corresponding protected peptide acids **2b**-**2k** with varying HPLC conversions ranging from 70% to greater than 99% (entries 1-8, Table 5.2). To determine the effect of the bulky neighboring residue on serine's activation and nucleophilic peptide release, the hexapeptides, FESQIS-Rink (**1e**), and TCDVS-Rink (**1k**) were synthesized. The results indicate that the activated peptides **1e** and **1k** with bulky amino acids (Ile and Val) adjacent to serine underwent complete conversion to peptide acids **2e** and **2k** (entries 4 and 10, Table 5.2). The reaction with peptide, FESQIS-Rink **1e**, also showed that the technique was applicable to peptides with multiple serine residues by selective activation and displacement of serine at the C-terminus (entry 4, Table 5.2).

Table 5.2 Substrate scope of CUT in synthesis of peptide acids.[a]



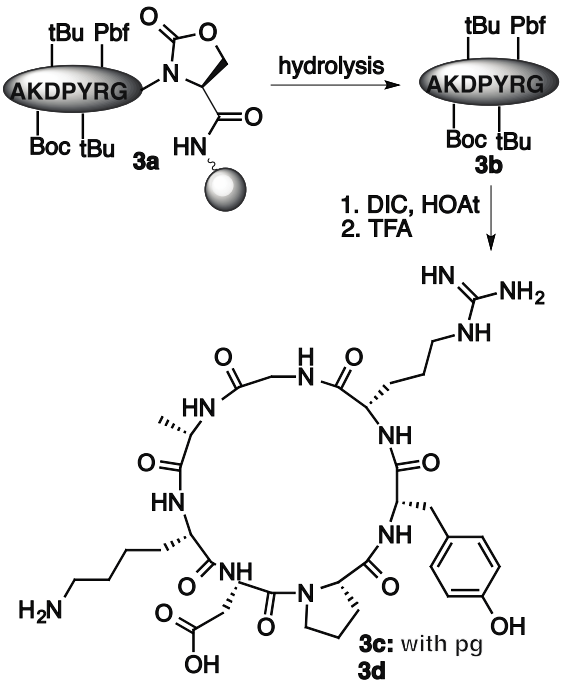
Entry	Substrate	Peptide	PG	Conv (%)[b]
1	**1b**	LFK(Boc)N(Trt)A	Ac	95
2	**1c**	LFK(Boc)N(Trt)A	NH_2	95
3[c]	**1d**	LFK(Boc)N(Trt)A	Fmoc	80
4	**1e**	FE(tBu)S(tBu)Q(Trt)I	NH_2	90
5	**1f**	R(Pbf)D(tBu)PMLG	Ac	95
6	**1g**	R(Pbf)D(tBu)PMLG	NH_2	90
7	**1h**	Y(tBu)LFK(Boc)N(Trt)A	Ac	95
8	**1i**	Y(tBu)LFK(Boc)N(Trt)A	NH_2	90
9	**1j**	VWR(Pbf)A	Ac	90
10	**1k**	T(tBu)C(tBu)D(tBu)V	Ac	90

[a]Reaction conditions: Peptide **1** (25 mg, 0.7 mmol/g) on solid support was reacted with DSC (100 mg), DIEA (70 uL) and a crystal of DMAP in DMF at room temperature followed by hydrolysis with H_2O:ACN in DIEA (20 uL) at 65 °C for 2-4 hrs. [b]Conversion to peptide acid was calculated from the absorbance at 220 nm using HPLC. [c]Mixture of completely protected peptide and Fmoc deprotected peptide was observed. Table adopted from reference (1) with permission.

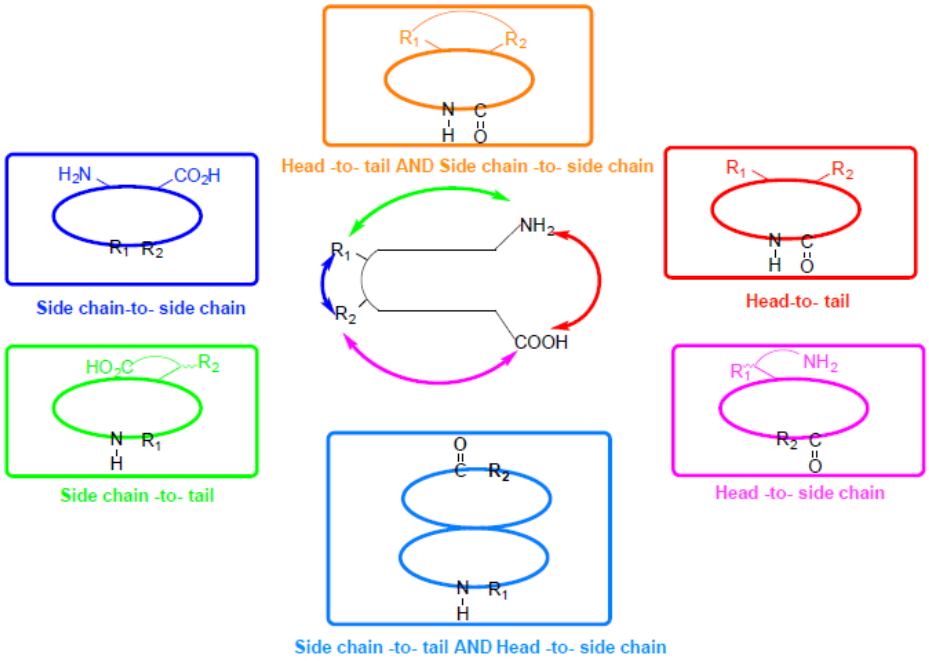
Next, to demonstrate that the cyclic urethane technique (CUT) is resin independent, model pentapeptide, WVRAS was synthesized on Rink Amide, Wang, Chem Matrix, and Tentagel resins. The results indicate that both the activation of serine and nucleophilic displacement gave high

yields of the protected peptide acid independent of the type of resin, which is in contrast to other methods.[8]

Protected peptide acids have been extensively useful for the synthesis of macrocyclic peptides of therapeutic importance.[18] To demonstrate the application of CUT in the synthesis of macrocyclic peptides, AKDPYRGS-Rink AM was synthesized on solid support, followed by serine selective deprotection, activation **3a,** and cleavage with water to obtain the protected linear peptide acid AK(Boc)D(tBu)PY(tBu)R(Pbf)G-OH **3b** (Scheme 5.3). Next, the protected peptide acid **3b** was subjected to coupling reagents to obtain the desired head to tail macrocyclized protected peptide cyc(AK(Boc)D(tBu)PY(tBu)R(Pbf)G) **3c**. This was followed by removal of the side-chain protecting groups of the macrocyclic peptide using the TFA cocktail to generate the deprotected macrocyclic peptide cyc(AKDPYRG) **3d** which was analyzed by NMR, HPLC, and MS (Scheme 5.3). In comparison, peptide acid AKDPYRG-OH generated by using Wang resin[19] gave an undesired double cyclized product. Double cyclization can arise from side-chain to head or side-chain to side-chain reactions (Scheme 5.4).

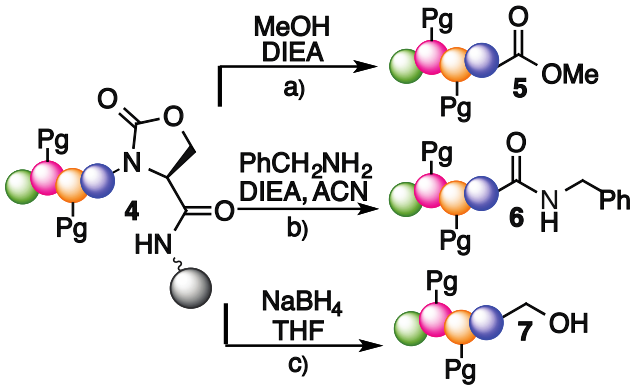


Scheme 5.3 Application of CUT in synthesis of macrocyclic peptide **3d**. Pg = protecting group. Scheme adopted from reference (1) with permission.



Scheme 5.4 Various possibilities for macrocyclization using an unprotected peptide. Scheme adopted from reference (1) with permission.

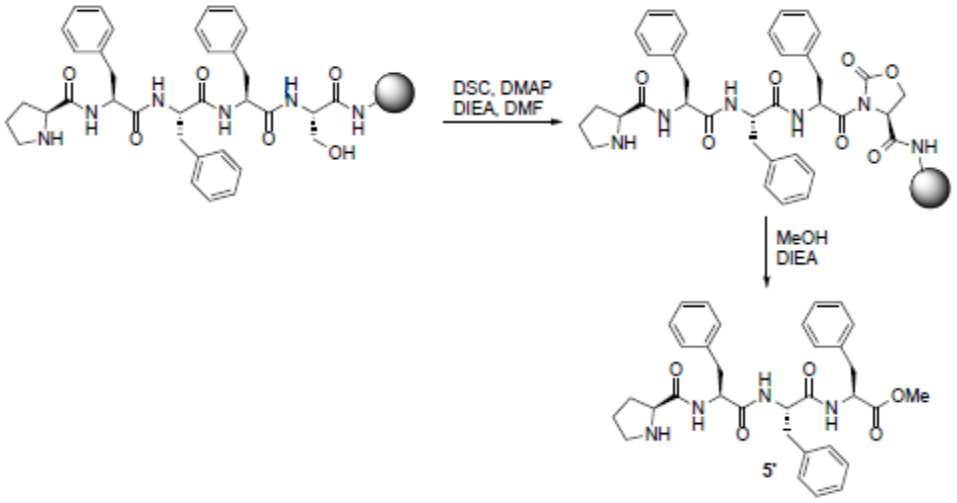
Next, we applied CUT for the synthesis of protected peptides with various C-terminal modifications such as esters, amides and alcohols (Scheme 5.5). To obtain the desired C-terminally modified peptide, the corresponding nucleophile was added to a solid support with the cyclic urethane moiety at the C-terminus (Scheme 5.5). The cyclic urethane moiety underwent nucleophilic displacement and generated the C-terminally modified protected peptide in solution. The peptide methyl ester, AVGPPGVA-OMe **5** was obtained by treatment of activated peptide **4** with methanol in the presence of a base (**a**, Scheme 5.5). These peptide methyl esters play a significant role in key processes of cell growth and differentiation.[5] This technique was also utilized for the synthesis of peptide methyl ester PFFFOMe **5'** (Scheme 5.6). Synthesis of these peptide methyl esters has previously been reported in solution-phase, which is a tedious process and requires purification at each step with the constant risk of epimerization.[20] It has been shown that the presence of a methyl ester group is critical for the catalytic activity of the peptide **5'**.[20]



Scheme 5.5 CUT for synthesis of peptide amides, esters and alcohols. peptide = **X**-AVGPPGVA-Oxd, **4. X** = NH_2 or Ac. Scheme adopted from reference (1) with permission.

Next, benzylamine was used as a nucleophile on the cyclic urethane moiety to obtain peptide an *N*-alkyl amide peptide **6**, an important class of biologically active peptides (**b**, Scheme 5.4). The presence of an alkyl amine group at the C-terminus enhances its biological activity,

metabolic stability, and hydrophobicity.[3,4] This technique was also applied for synthesis of peptides with a C-terminal alcohol functionality. Such peptides occur in nature among these, peptaibols are well-known antibiotic peptides that exhibit a wide range of biological activity.[6] To obtain an alcohol group at the C-terminus **7**, the cyclic urethane containing activated peptide **4** was treated with a solution of sodium borohydride in THF (**c**, Scheme 5.4).



Scheme 5.6 Application of CUT in synthesis of a catalyst **5'** for aldol reaction. Scheme adopted from reference (1) with permission.

Moreover, this methodology was implemented in the synthesis of hydrazides. The critical importance of a peptide hydrazide is to synthesize complex proteins and bioconjugates; therefore, it is an added advantage that the cyclic urethane moiety is capable of providing long peptide hydrazides with diverse functional groups that can be utilized directly in NCL. A multi-serine containing bioactive peptide hydrazide S(tBu)GIS(tBu)GPLS(tBu)-$CONHNH_2$ **3k**, a fragment of antimicrobial bovine β-defensin 13, was successfully synthesized from the corresponding cyclic urethane activated peptide S(tBu)GIS(tBu)GPLS(tBu)-Oxd **2k** by treatment of the peptide with hydrazine monohydrate under basic conditions using a heated shaker at 65 °C for 2 h in acetonitrile

as a solvent for optimal release (Figure 5.3).

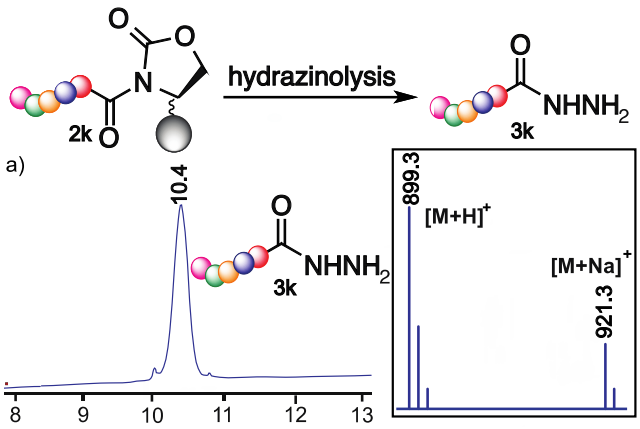
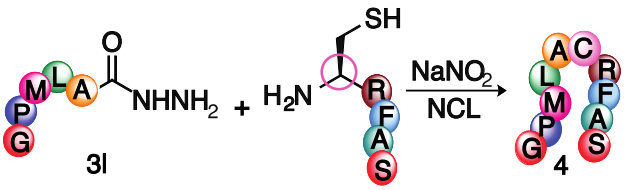


Figure 5.3 Activated serine for the synthesis of the fragment of antimicrobial bovine *β*-defensin 13 bioactive peptide hydrazide **3k.** (a) HPLC trace of pure bioactive peptide hydrazide S(tBu)GIS(tBu)GPLS(tBu)-$CONHNH_2$ **3k** and inset shows MS of protected peptide hydrazide **3k**.

Given the importance of peptide hydrazides in protein chemical synthesis, we envisioned the application of this methodology in an efficient one-pot strategy, comprising solid-phase synthesis, hydrazinolysis, and the direct ligation of crude peptide hydrazides to cysteine-containing peptide (Figure 5.4). Peptide hydrazide Ac-GPMLA-$CONHNH_2$ **3l** was cleanly released from Ac-GPMLA-Oxd **2l** by treatment with hydrazine monohydrate and DIEA in acetonitrile at 65 °C for 2 h and ligated with a cysteine containing peptide CRFAS-NH_2.



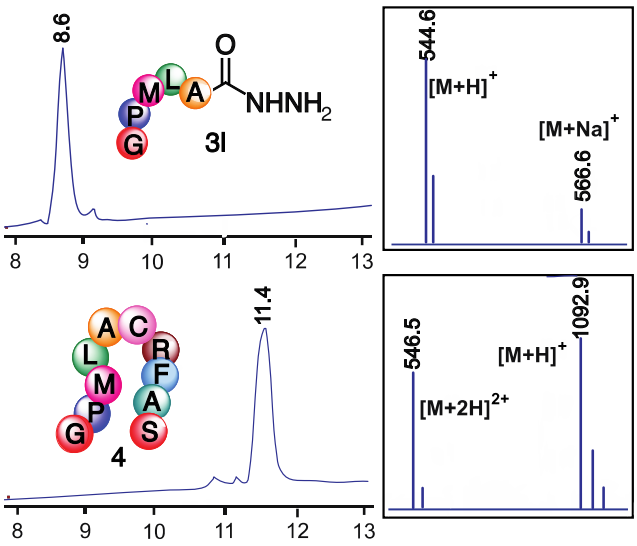
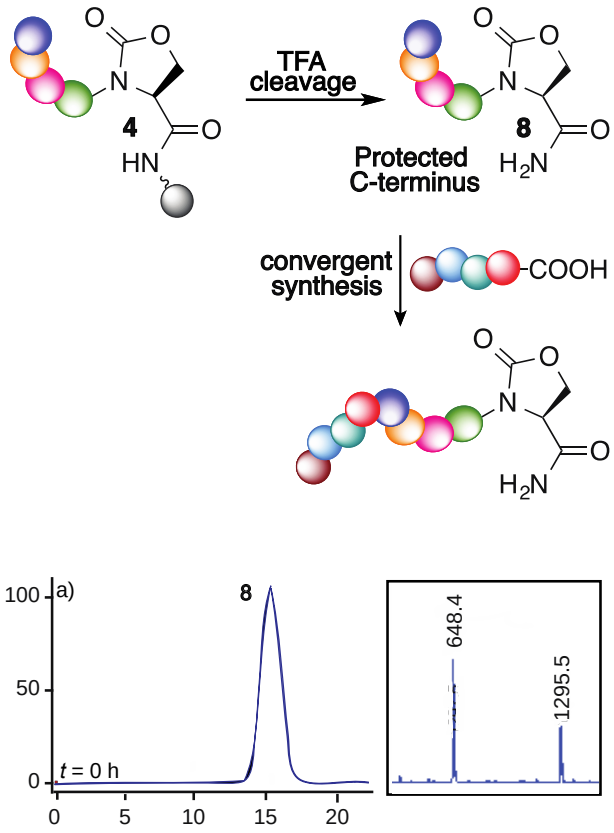


Figure 5.4 Native Chemical Ligation of peptide hydrazide Ac-GPMLA-CONHNH$_2$ **3l** with N-terminal cysteine peptide CRFAS-NH$_2$. HPLC/MS traces of peptide hydrazide **3l** and ligated product **4**.

The main advantage of the described approach is that the elongation of the peptide chain can be carried out by means of standard techniques after attachment of the first amino acid as serine. Another significant advantage of this approach is that only those resin-bound peptides that complete both activation and displacement steps are released. Thus, the released C-terminally modified peptides exhibit high purity, eliminating the need for extensive purification.

Next, we demonstrated the application of CUT for synthesis of peptides with a semi-permanent protecting group (pg) at the C-terminus (Scheme 5.7). This is of significance because C-terminally protected peptides can be used for various manipulations in solution. This technique is in contrast to the current methods that require reprotection of the C-terminal carboxyl group in solution, with a consequent risk of racemization and low yields.[21] Another method for synthesis of peptides with a semi-permanent protecting group at the C-terminus utilizes click chemistry based handles but is limited by the use of special linkers and the requirement of harsh deprotection conditions.[21] Our method does not require any special linkers or handles and it employs mild deprotection conditions for removal of pg at the C-terminus.

For the synthesis of peptides with a semi-permanent C-terminus pg, activated peptide Fmoc-AR(Pbf)FPPFR(Pbf)AOxd **4** on a solid support was cleaved from the resin by using the TFA cleavage cocktail to generate peptide Fmoc-ARFPPFRAOxd **8** with protection at the C-terminus (Scheme 5.7). The C-terminally protected peptide Fmoc-ARFPPFRAOxd **8** can be used for the synthesis of large peptides by a convergent approach. Subsequently, the protecting group at the C-terminus of the resulting large peptide can be easily removed under basic hydrolysis conditions as shown above.



Scheme 5.7 (a) CUT technique for synthesis of protected C-terminus peptide **8**. (b) HPLC and MS chromatogram of C-terminal protected peptide Fmoc-ARFPPFRA-Oxd, **8.** peptide = Fmoc-ARFPPFRA**.** Scheme adopted from reference (1) with permission.

5.6 CONCLUSIONS

A backbone amide activation approach has been developed for the synthesis of C-terminally modified peptides of therapeutic importance such as acids, amides, esters, N-alkyl

amides, alcohols, and hydrazides. This approach is resin and linker independent thus circumventing limitations associated with earlier methods. The main advantage of the described approach is that the elongation of the peptide chain can be carried out by means of standard methodologies after attachment of the first amino acid residue as serine. After construction of the peptide, the backbone chain is activated as a cyclic urethane moiety and cleavage is carried out under mild nucleophilic displacement conditions that allow easy formation of the desired C-terminally modified peptide. The target peptides are released from the resin without racemization. This cyclic urethane technique (CUT) has also been applied to the synthesis of peptides with the protected C-terminus, which can undergo various modifications in solution after release from solid support (e.g., large peptides synthesis by a convergent strategy). Moreover, the C-terminal protecting group can be easily removed under mild basic hydrolytic conditions. A universal solid phase strategy has been effectively developed and applied for the synthesis of various C-terminal modified peptides, which are of therapeutic importance due to biological properties that accompany the attachment of various C-terminal functional groups.

5.7 EXPERIMENTAL SECTION

General: All commercial materials (Aldrich, Fluka, Nova) were used without further purification. All solvents were reagent grade or HPLC grade (Fisher). Anhydrous THF, diethyl ether, CH_2Cl_2, and DMF were obtained from a dry solvent system (passed through column of alumina) and used without further drying. Conversion were obtained by comparison of HPLC peak areas of products and starting material.

Materials: Fmoc-amino acids were obtained from Nova Biochem (EMD Millipore Corporation) (Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide resin was obtained from ChemPep Inc (Wellington, Florida). N,N,N′ ,*N*′ -Tetramethyl-O (1H benzotriazol-1-

yl)uronium hexafluorophosphate (HBTU) was obtained from CreoSalus (Louisville, Kentucky). N,*N*-Disuccinimidyl carbonate (DSC) was obtained from Nova Biochem, under (EMD Millipore Corporation) (Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP): Merck KGaA (Darmstadt, Germany). N,*N*-Dimethylformamide (DMF): Macron Fine Chemicals (Center Valley, Pennsylvania). Dichloroethane (DCE), acetonitrile, N,N- Diisopropylethylamine (DIEA), N,*N*' diisopropylcarbodiimide (DIC), were purchased from (EMD Millipore Corporation)(Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR 100 Matsonford Road Radnor, PA. Diethyl Ether: Sigma Aldrich (St. Louis, Missouri). Water was purified using a Millipore MilliQ water purification system.

NMR:

This was done by Ryan Cohen at Merck facilities at Rayway, NJ. NMR spectra were recorded on a 600 MHz spectrometer and carbon NMR spectra on a 151 MHz, spectrometer at ambient temperature. All NMR chemical shifts (™) are referenced in ppm relative to residual solvent or internal tetramethylsilane. ^{1}H NMR chemical shifts referenced to residual DMSO-d_5 at 2.50 ppm, and ^{13}C NMR chemical shifts referenced to DMSO-d_6 at 39.52ppm. Carbon NMR spectra are proton decoupled. NMR spectral data are reported as chemical shift (multiplicity, coupling constants (*J*), integration). Multiplicity is reported as follows: singlet (s), broad singlet (bs), doublet (d), doublet of doubles (dd), doublet of triplet (td), triplet (t) and multiplet (m). Coupling constant (*J*) in Hertz (Hz).

HPLC

Semi-Preparative HPLC:

Preparative HPLC chromatography (HPLC) was performed on Beckman Coulter equipped with System Gold 168 detector and 125P solvent module HPLC with a 10 mm C-18 reversed-phase

column. All separations involved a mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1% FA (v/v) in acetonitrile (solvent B). Semi-preparative HPLC method using a linear gradient of 0–80% acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 3.0 mL min-1. The eluent was monitored by absorbance at 220 nm and 254 nm unless otherwise noted.

Analytical HPLC:
Analytical HPLC chromatography (HPLC) was performed on an Agilent 1200 series HPLC equipped with a 4.6x 150 mm (5µm) C-18 reversed-phase column. All separations involved mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1% FA (v/v) in acetonitrile (solvent B). Peptide compositions were evaluated by analytical reverse phase HPLC using a gradient of 0.1% FA in acetonitrile versus 0.1% FA in water. Analytical HPLC method using a linear gradient of 0–80% 0.1% FA (v/v) acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 1.0 mL min-1. The eluent was monitored by absorbance at 254 nm unless otherwise noted.

LCMS:
Mass spectrometry was performed using ultra high-performance liquid chromatography-mass spectrometry using the Agilent 1100 Series LCMSD VL MS Spectrometer.

HRMS:
This was done by Ryan Cohen at Merck facilities at Rayway, NJ. HRMS data were recorded on an Agilent 6520 Q-ToF mass spectrometer using positive polarity electrospray ionization (+ESI).

Fmoc Solid-Phase Peptide Synthesis.[22] Peptides were synthesized manually on a 0.25 mmol scale using Rink amide, Wang, Chem Matrix and TentaGel resins. Fmoc–group was deprotected using 20% piperidine–DMF for 20 min to obtain a deprotected peptide-resin. Fmoc-protected amino acids (1.25 mmol) were sequentially coupled on the resin using a HBTU (1.25 mmol) and DIEA (1.25 mmol) for 2 hrs at room temperature. Peptides were synthesized using standard

protocols.[22] The peptide was cleaved from the resin using a cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water for 2 hrs. The resin was removed by filtration and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by HPLC.

General procedure for the activation of serine to cyclic urethane moiety on solid support: To a peptide on the solid support (25-100 mg (0.25-0.69 mmol/g) was added a solution of DSC (15 equiv.), DIEA (15 equiv.) and catalytic amount of DMAP in dimethylformamide (DMF). The resin was left on shaker for 17 hrs. The solution was drained and resin was washed with DMF followed by cleavage using a cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water for 2 hrs. The resin was removed by filtration and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by HPLC for analysis by MS and NMR.

General procedure for the synthesis of peptide acids from solid support: To an activated peptide as cyclic urethane moiety on the solid support, 1 mL H_2O:ACN (1:1) and 20 µL of DIEA was added and resin was left on the heated shaker at 65 °C for 2 hrs. The resin was filtered which is followed by the removal of solvent under high vacuum. The resulting peptide acid was analyzed by MS and purified by HPLC to obtain white solid. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

General procedure for the synthesis of peptide esters from solid support: To an activated peptide as cyclic urethane moiety on the solid support, 1 mL MeOH and 100 µL of DIEA was

added and resin was left on the heated shaker at 65 °C for 4 hrs. The resin was filtered which is followed by the removal of solvent under high vacuum. The resulting peptide ester was analyzed by MS and purified by HPLC to obtain white solid. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

General procedure for the synthesis of peptide N-aryl amide from solid support: To an activated peptide as cyclic urethane moiety on the solid support, 1 mL ACN and 100 µL of benzyl amine was added and resin was left on the heated shaker at 65 °C for 4 hrs. The resin was filtered which is followed by the removal of solvent under high vacuum. The resulting peptide N-aryl amide was analyzed by MS and purified by HPLC to obtain white solid. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

General procedure for the synthesis of peptide alcohol from solid support: To an activated peptide as a cyclic urethane moiety on the solid support, 100 mg sodium borohydride in 1 mL THF was added and resin was left on the shaker overnight at room temperature. The resin was filtered which is followed by the removal of solvent under high vacuum. The resulting peptide alcohol was analyzed by MS and purified by HPLC to obtain white solid. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

General procedure for the synthesis of peptide hydrazides from solid support: To an activated peptide as a cyclic urethane moiety on the solid support, 20 µL hydrazine monohydrate in 1mL acetonitrile was added to 25 mg of resin and left on the shaker for 2 hrs at 65ºC. The resin was filtered which is followed by the removal of solvent under high vacuum. The resulting peptide hydrazide was analyzed by MS and purified by HPLC to obtain white solid. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

Deprotection of protecting groups from C-terminal modified peptides: C-terminally modified peptides were deprotected by using a cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water for 2 hrs in solution and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by HPLC and analyzed by MS.

General procedure for the synthesis of protected C-terminal peptides: To an activated peptide as cyclic urethane moiety on the solid support, cleavage cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water was added and left on shaker for 2 hrs. Resulting solution was concentrated and oily residue was triturated with diethyl ether to obtain a white suspension. The resulting protected C-terminal peptide was purified by HPLC and analyzed by MS. HPLC: 0.1% FA (v/v) in water (solvent A): 0.1% FA (v/v) acetonitrile (solvent B); gradient 0-80 %, 0.1% FA (v/v) acetonitrile in 25 min, flow rate = 1.0 mL/min, detection wavelength 220 nm.

Procedure for Native Chemical Ligation: Peptide hydrazide Ac-GPMLA-$CONHNH_2$ was cleanly released from Ac-GPMLA-Oxd by treatment with hydrazine monohydrate and DIEA in acetonitrile at 65 °C for 2 hrs. After the filtration of the resin and removal of acetonitrile by a

stream of nitrogen, an aqueous phosphate (0.2 M) buffer containing 6.0 M guanidinum chloride followed by addition of cysteine containing peptide CRFAS-NH_2 was added (final peptide concentration of 1.5 and 2.0 mM, respectively). At low pH (3.0) and temperature (−10 °C), an aqueous $NaNO_2$ solution was added to the ligation mixture. After 20 min, 4-mercaptophenylacetic acid (MPAA) was added, the pH value was adjusted to 7.0, and the reaction was left at room temperature for 16 hrs. The formation of the ligation product Ac-GPML<u>AC</u>RFAS-NH_2 was analyzed by injecting the sample in the HPLC and MS after regular intervals of time.

5.8 REFERENCES

1. Elashal, H. E.; Cohen, R. D.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699.
2. Songster, M. F.; Barany, G. *Methods in Enzymology*, Academic Press, **1997**, vol. Volume 289, pp. 126.
3. Woo, Y. H.; Mitchell, A. R.; Camarero, J.A. *Int. J. Pept. Res. Therap.* **2007**, *13*, 181.
4. Alsina, J., Albericio, F. *Pept. Science* **2003**, *71*, 454.
5. Peters, C.;Waldmann, H. *J. Org. Chem.* **2003**, *68*, 6053.
6. Chugh, J. K.; Wallace, B. A. *Biochem. Soc. Trans.* **2001**, *29*, 565.
7. T. W. G. a. P. G. M. Wuts, *John Wiley & Sons, New York*, **1999**.
8. Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F. *ACS Comb. Sci.* **2013**, *15*, 217.
9. Semenov, A. N.; Gordeev, K. *Int. J. Pept. Protein Res.* **1995**, *45*, 303.
10. Alsina, J.; Yokum, T. S.; Albericio, F.; Barany, G. *J. Org. Chem.* **1999**, *64*, 8761.
11. Thieriet,N.; Guibe, F; Albericio,F. *Org. Lett.* **2000**, *2*, 1815.
12. Hasinoff, L.; Takats, J.; Zhang, X. W.; Bond, A. J.; Rogers,R. D. *J. Am. Chem. Soc.* **1994**, *116*, 8833.
13. Davies, M.; Bradley, M. *Angew. Chem., Int. Ed.*, **1997**, *36*, 1097.

14. Davies, M.; Bradley, M. *Tetrahedron*, **1999**, *55*, 4733.

15. Gongora-Benitez, M.; Cristau, M.; Giraud, M.; Tulla-Puche, J.; Albericio, F. *Chem. Commun.* **2012**, *48*, 2313.

16. De Marco, R.; Tolomelli, A.; Campitiello, M.; Rubini, P.; Gentilucci, L. *Org. Biomol. Chem.* **2012**, *10*, 2307.

17. Elashal, H. E.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699.

18. Driggers, E. M.; Hale, S.P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discov.* **2008**, ***7***, 608.

19. Wang, W. W. *J. Am. Chem. Soc.* **1973**, *95*, 1328.

20. Tang, Z.; Yang, Z. H.; Cun, L. F.; Gong, L.Z.; Mi, A. Q.; Jiang, Y. Z. *Org. Lett.* **2004**, *6*, 2285.

21. Kolb, H.C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004.

22. Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.

CHAPTER 6: SERINE PROMOTED SYNTHESIS OF PEPTIDE THIOESTER-PRECURSOR ON SOLID SUPPORT FOR NATIVE CHEMICAL LIGATION

6.1 ABSTRACT

Fmoc solid phase peptide synthesis of thioesters for the chemical synthesis of proteins via native chemical ligation is a challenge. We have developed a versatile approach for the direct synthesis of peptide thioesters from a solid support utilizing Fmoc chemistry. Peptide thioester synthesis is performed by the formation of a cyclic urethane moiety via a selective reaction of the backbone amide chain with the side group of serine. The activated cyclic urethane moiety undergoes displacement by a thiol to generate the thioester directly from the solid support. Importantly, the method activates the serine residue for the synthesis of peptide thioesters; it is fully automated and free of the types of resins, linkers, handles, and unnatural amino acids typically needed for the synthesis of peptide thioesters using Fmoc chemistry. The resulting thioester is free of epimerization and is successfully applied for the synthesis of longer peptides using NCL.

6.2 CHAPTER OBJECTIVES

This chapter describes *a simple and versatile approach for the Fmoc solid phase synthesis of peptide thioesters for chemical synthesis of proteins by native chemical ligation (NCL). A novel methodology for the activation of serine to generate peptide thioesters from solid support was developed and resulting peptide thioesters can be directly used in NCL to yield the desired protein without further modifications.*

Fmoc SPPS of peptide thioesters are of high significance because they are precursors for convergent synthesis of proteins via NCL. As a result, there is a great necessity for an effective Fmoc solid-phase approach for the preparation of peptide thioesters. Currently, peptide thioesters

are synthesized by the use of sulfonamide safety catch linkers, O/N to S acyl shift methods, an activated N-acylurea linker, a backbone pyro-glutamyl imide linker, and peptide hydrazides. Despite these significant advances, current approaches for the synthesis of peptide thioesters are limited by a 'slow rate' of ligation, the requirement of acidic conditions, formation of various side products, and the need for unnatural linkers at the C-terminus. Our strategy provides a novel way for the modification of serine on a solid support which acts as a thioester surrogate. The modified serine as a thioester surrogate undergoes nucleophilic displacement with a thiol to generate a peptide thioester directly from the solid support. The technique relies on reactivity of the side chain of serine for the acylation of backbone amide chain to generate the cyclic urethane moiety on solid support. The cyclic urethane moiety has a potential to act as a thioester surrogate to generate peptide thioesters directly from a solid support and post synthetic steps prior to NCL are not required.

6.3 GRAPHICAL ABSTRACT

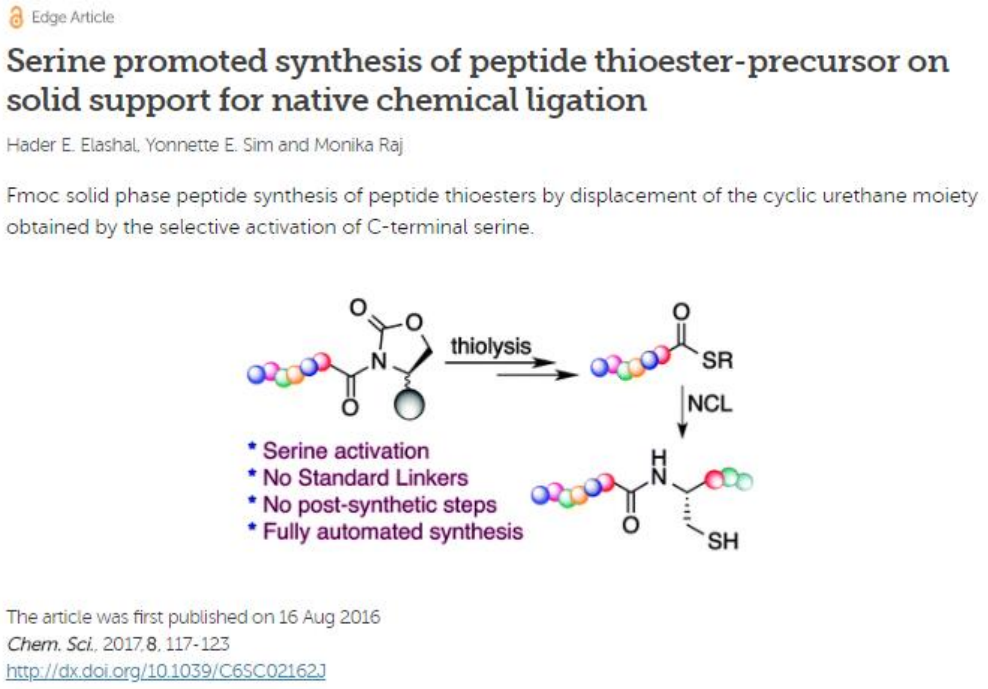


Figure 6.1 Graphical abstract for *Serine Promoted Synthesis of Peptide Thioester-Precursor on Solid Support for Native Chemical Ligation*. Figure adopted with permission from: Elashal, H.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, *8*, 117.[1]

6.4 INTRODUCTION

Total chemical synthesis of proteins provides easy access to various modified proteins with high stability and improved biological activity which can be key in understanding the importance of various post translational modifications.[2,3] Native Chemical Ligation (NCL) has revolutionized the field of chemical synthesis of proteins and relies on the reaction of a peptide thioester with a cysteinyl peptide.[4] Thus, considerable effort has been applied for synthesis of peptide thioesters using Fmoc or Boc solid- phase peptide synthesis (SPPS) approach.

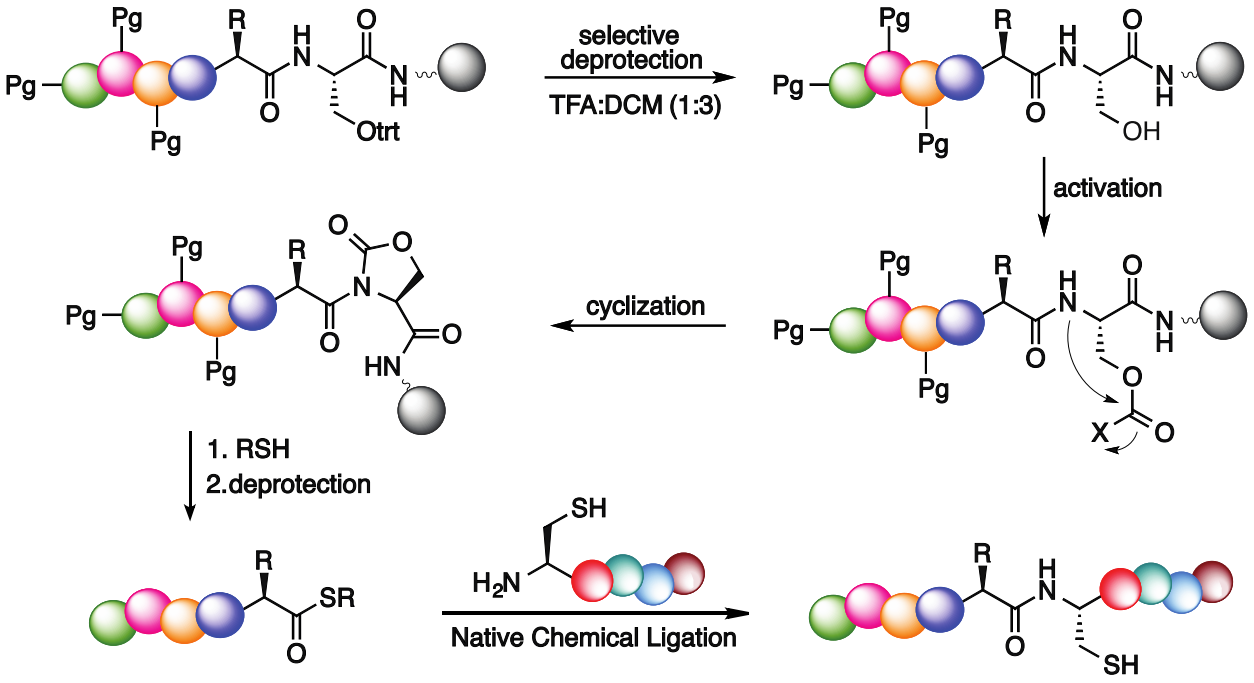
The common method for generation of peptide thioesters by Boc-based SPPS utilizes special thioester linkers but Boc chemistry is limited by the use of harsh cleavage conditions, usually hydrofluoric acid (HF), thus is not suitable for synthesis of post-translationally modified peptides such as glyco- or phospho-peptides.[5] Fmoc SPPS is not applicable on thioester resins because of the instability of the peptide thioesters towards the piperidine treatment used for the deprotection of Fmoc group. Substantial efforts have been utilized for the synthesis of peptide thioesters by Fmoc SPPS approach such as use of less basic Fmoc deprotection protocol,[6] use of sulfonamide safety catch linkers,[7] activation of protected peptides in solution,[8] O/N to S acyl shift methods,[9-15] and peptide hydrazides.[16-18] Activation of protected peptides in solution for synthesis of thioesters is limited by the susceptibility of the C-terminal residue to racemization. Despite the significant advances of O/N to S acyl shift methods, they are limited by slow ligation kinetics,[14] and peptide hydrazides require acidic conditions,[14] which is unsuitable for synthesis of acid sensitive post translationally modified peptides.

Another significant approach for the formation of a thioester is the activation of a relatively inert backbone amide bond followed by subsequent thiolysis. Examples include, an activated *N*-acylurea linker,[19,20] a backbone pyro-glutamyl imide linker.[21] An activated *N*-acylurea approach has been extensively used for synthesis of variety of proteins[19,20] but is limited by the use of an unnatural amino acid, formation of various side products,[22,23] and requirement of additional steps for the synthesis of *N*-methylated unnatural amino acid. Whereas, a backbone pyro-glutamyl imide linker is limited by the instability of the resin.[21] Thus, an efficient and versatile route to peptide thioesters via Fmoc-SPPS is highly desirable.

This chapter describes an approach for the Fmoc synthesis of peptide thioesters for use in NCL that is based on the activation of the amide backbone chain at a serine residue. The significance of this approach is that it utilizes a natural serine residue that can be automatically assembled onto the solid support and does not require special unnatural amino acids and loading procedures, which makes it distinct from the currently used methods.[19,20] We believe that this accessible and robust Fmoc-based thioesterication technique provides a significant advance to chemical protein synthesis due to its uncomplicated nature, therefore eliminating special precautions and additional steps typically needed for the synthesis of peptide thioesters.[19,20]

This approach relies on the site-selective activation of a backbone amide bond at the C-terminal serine residue for the generation of a cyclic urethane moiety on the solid support. Synthesis of peptide thioesters by this strategy entails anchoring of a C-terminal serine residue with a selectively removable side-chain protecting group to a solid support (Scheme 6.1). Next, the side-chain of serine is selectively deprotected and activated by an electrophile that results in the formation of a cyclic urethane moiety on resin (Scheme 6.1). Next, nucleophilic displacement of the cyclic urethane moiety by treatment with a thiol releases the peptide thioester from the solid

support, which can then be deprotected in the solution (Scheme 6.1). It is noteworthy that this approach utilizes a natural serine during chain assembly followed by acylation of the backbone amide chain and is compatible with various amino acids and protecting groups commonly utilized in Fmoc solid phase peptide synthesis. Moreover, the thioester is directly obtained from the solid support thus reduces the need for extra steps before proceeding with NCL.



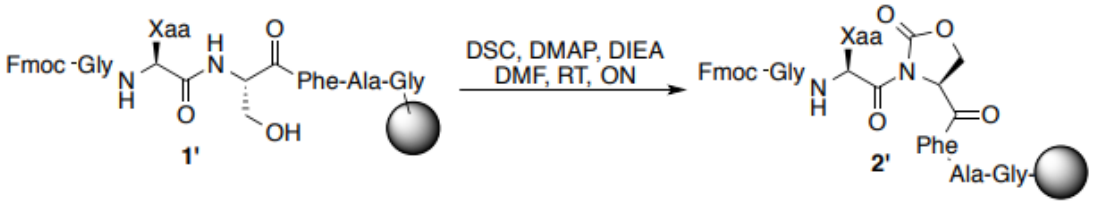
Scheme 6.1 Rationale for the development of cyclic urethane technique (CUT) for the synthesis of peptide thioesters and its application in native chemical ligation for chemical synthesis of proteins. Scheme adopted from reference (1) with permission.

6.5 RESULTS AND DISCUSSION

To test the proposed methodology for the synthesis of thioesters, the first step was the synthesis of the cyclic urethane moiety from a serine residue on the solid support. The serine activation reaction was explored on a solid support with the peptide Fmoc-Gly-Xaa-Ser-Phe-Ala-Gly, where various amino acids were substituted in the Xaa position.(Table 6.1) It is noteworthy that the formation of the activated cyclic urethane moiety is independent of the nature of the amino

acid preceding serine and a high conversion to cyclic urethane moiety was obtained even with bulky amino acid residues next to serine such as Tyr, Trp, and Val (Table 6.1, substrates 1e'-1g'). Moreover, activation of serine does not depend upon the relative position of serine and the resin.

Table 6.1 Substrate scope of Ser Cyclization on Fmoc-Gly-**Xaa-Ser**-Phe-Ala-Gly



Substrate	Xaa	Conversion (%)[b]
1a'	Ala	>99
1b'	Gly	>99
1c'	Met	>99
1d'	His	>99
1e'	Tyr	>99
1f'	Trp	>99
1g'	Val	>99

[a]Reaction conditions: Peptide **1'** (25 mg, 0.7 mm/g) on solid support was reacted with DSC (10 equiv.), DIEA (10 equiv.), and a crystal of DMAP in DMF (3 mL) at room temperature for 17 h. [b]Conversion to cyclic urethane moiety **2'** was calculated from the absorbance at 220 nm using HPLC.

Table 6.1 adopted from reference (1) with permission.

The stability of the cyclic urethane moiety on the resin bound peptide was evaluated. The resin bound cyclic urethane moiety is very stable in a desiccator for longer than a month, which is in contrast to previously reported pyroglutamyl imide precursor method.[21] Moreover, formation of the cyclo-urethane moiety on the peptide Fmoc-GFA(*L*)Oxd (**2f**) and the corresponding diastereoisomer Fmoc-GFA(*D*)Oxd (**2f'**) showed no sign of epimerization when analyzed by HPLC (Figure 6.2).

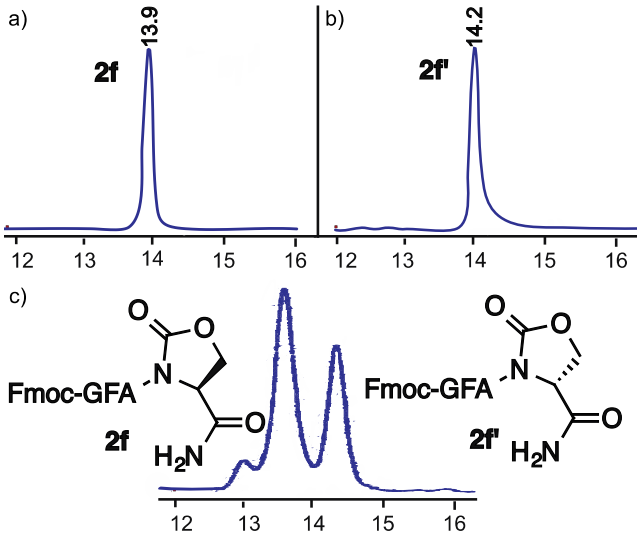
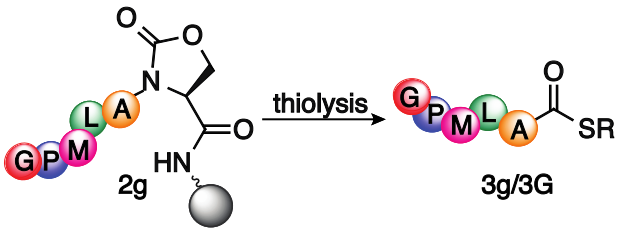


Figure 6.2 HPLC traces of (a) purified peptide with cyclic urethane moiety **2f** Fmoc-GFA(*L*)-Oxd (b) purified peptide diastereoisomer **2f'** Fmoc-GFA(*D*)-Oxd and (c) mixture containing both diastereoisomers of cyclic urethane activated peptides **2f** and **2f'**, demonstrating lack of detectable epimerization. Figure adopted from reference (1) with permission.

After activating the peptide backbone as a cyclic urethane moiety on model peptide Ac-GPMLA-*Oxd*-Rink AM (**2g**) on solid support, the thiolysis reaction was performed (Table 6.2). The resin was treated with varying amounts of thiol/base mixtures and solvents at different temperatures, in order to release the peptide into solution as a peptide thioester Ac-GPMLA-COSR **3g**. The resin was then filtered and solvent was evaporated to obtain peptide thioester **3g** which was analyzed by HPLC, MS and NMR.

Table 6.2 Synthesis of peptide thioester **3g/3G** from activated peptide Ac-GPMLA-Oxd-Rink AM **2g** on solid support.[a]



Entry	Base (equiv.)	Solvent	Temp. (°C)	Time (h)	Conv. (%)[b] **3g**	**3G**
1	-	DMF	RT	20	10	15
2	DBU (5 equiv.)	DMF	RT	20	40 [c]	60
3	DIEA (20 equiv.)	DMF	RT	20	60 [c]	65
4	**PhSNa (0.5 equiv.)**	**DMF**	**RT**	**20**	**85** [c]	**99**
5	PhSNa (0.5 equiv.)	ACN	RT	20	10 [c]	20
6	PhSNa (0.5 equiv.)	DMF	RT	5	30	35
7	PhSNa (0.5 equiv.)	DMF	HS, 60 °C	20	10 [c]	70 [c]
8	PhSNa (0.5 equiv.)	DMF	HS, 60 °C	5	30 [c]	99
9	PhSNa (0.5 equiv.)	DMF	HS, 60 °C	3	50 [c]	99

[a]Reaction conditions: Cyclic urethane activated peptide Ac-GPMLA-*Oxd*, **2g** (25 mg, 0.7 mmol/g) on solid support was reacted with thiol (100 µL), and base (0.5-20 equiv.), in DMF (1 mL). [b]Conversion to peptide thioester Ac-GPMLA-COSR **3g** or Ac-GPMLA-COSR **3G** was calculated from the absorbance at 220 nm using HPLC. [c]Activated peptide was completely released from the resin but hydrolysis product was observed along with the thioester **3g**. RSH = SH-$(CH_2)_2$-OH or R'SH = SH-$(CH_2)_2$-$COOC_2H_5$, HS = heated shaker, RT = room temperature, **Entries in bold**: optimized conditions. Thioesterification at HS leads to 5-10% epimerization. Table adopted from reference (1) with permission.

Initially, the reaction was performed at room temperature using DMF as a solvent in the presence of a thiol without use of any base (entry 1, Table 6.2). This resulted in the release of 10%

peptide thioester **3g** from the resin after treatment for 20 h. To increase the conversion to peptide thioester **3g**, various bases such as DIEA, DBU and sodium thiolate were investigated (entries 2-4, Table 2.3). Addition of a catalytic amount of sodium thiolate as a base significantly increased the yield of peptide thioester **3g** released from the resin (80%; entry 4, Table 6.2). Since it is difficult to remove DMF from the reaction mixture, a less volatile solvent such as ACN was explored, but resulted in very low conversion to a peptide thioester **3g** (20 %; entry 5, Table 6.2). The final yields in ACN were low due to the poor solubility of sodium thiolate in ACN. Temperatures higher than 60 °C resulted in the significant hydrolysis of peptide thioester **3g** into the corresponding acid (entries 7-9, Table 6.2). The formation of the hydrolyzed product was circumvented by the use of ethyl-3-mercaptopropionate, which gave stable peptide thioesters **3G** that are less susceptible to hydrolysis. All reactions were clean at room temperature irrespective of the nature of thiol used and did not show significant amounts of hydrolysis.

Potential epimerization during the thioesterification process was rigorously examined by synthesizing a peptide thioester Ac-GPMLA(*L*)-COS$(CH_2)_2$OH **3g** under optimized reaction conditions and epimer Ac-GPMLA(*D*)-COS$(CH_2)_2$OH **3g'**. No detectable levels of epimerization were observed as analyzed by HPLC (Figure 6.3). Next, we investigated epimerization studies on peptide thioester Ac-GVAL**F**(*L*)-$(CH_2)_2$-$COOC_2H_5$ **3j** with bulky Phe residue at the C-terminus by analyzing it with epimer Ac-GVAL**F**(*D*)-$(CH_2)_2$-$COOC_2H_5$ **3j'**. Less than 1 % of epimerization was observed as analyzed by HPLC. Indeed, Phe is known as an amino acid relatively prone to epimerization, and thus these results indicate that our new methodology did not afford substantial epimerization.

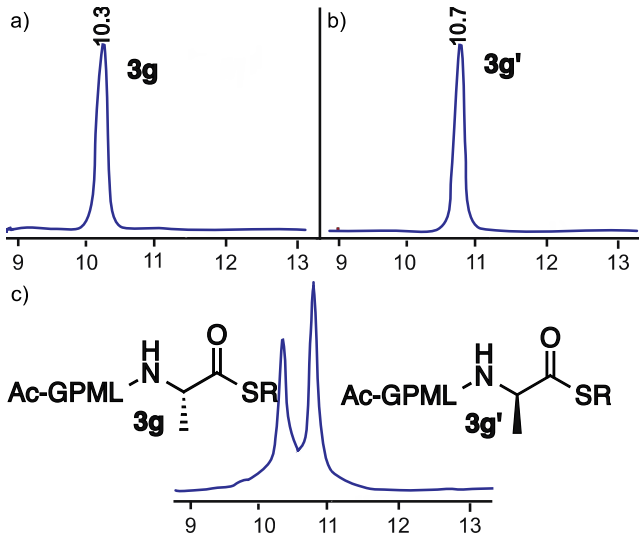
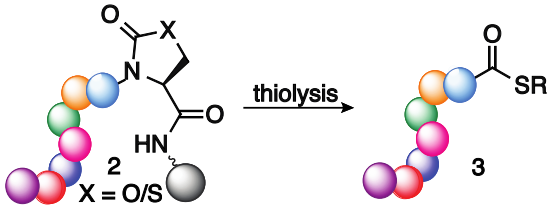


Figure 6.3 HPLC traces of (a) purified peptide thioester Ac-GPMLA(*L*)-COS$(CH_2)_2$OH **3g** (b) purified peptide thioester diastereoisomer Ac-GPMLA(*D*)-COS$(CH_2)_2$OH **3g'**and (c) mixture containing both diastereoisomers of peptide thioesters **3g** and **3g'**, demonstrating lack of detectable epimerization. RSH = HS$(CH_2)_2$OH. Figure adopted from reference (1) with permission.

Various peptides with different amino acids and protecting groups were evaluated for the synthesis of peptide thioesters (Table 6.3). Protecting groups from peptide thioesters were removed by using a TFA cleavage cocktail in solution. It was noted that the formation of peptide thioesters is independent of nature of amino acid residues preceding serine. For example, peptides with a bulky amino acid residue next to serine such as Val, Phe, Glu(tBu), and Ser(tBu) generated the protected peptide thioesters in a clean manner and with high yields (entries 2-5, Table 6.3). From these studies, we conclude that the *cyclic urethane (Oxd) peptide* will be a versatile precursor for synthesis of peptidyl thioesters.

Table 6.3 Fmoc SPPS synthesis of peptide thioesters *via* activated serine.[a]



Entry	Substrate	Peptide	Conv (%)[b]
1	**2h**	Ac-AVGPPGV**A**-*Oxd*	95
2	**2i**	Ac-R(Pbf)AFK(Boc)Y(tBu)GL**E(tBu)**-*Oxd*	95
3	**2j**	Ac-GVAL**F**-*Oxd*	95
4	**2k**	Ac-Y(tBu)FD(tBu)IR(Pbf)A**V**-*Oxd*	90
5	**2l**	Ac-S(tBu)GIS(tBu)GPL**S(tBu)**-Oxd	95
6	**2m**	Ac-GPMLA-^{Me}*Oxd*	90
7	**2n**	Ac-GPMLA-*Thz*	95

[a]Reaction conditions: Activated peptide, **2** (25 mg, 0.7 mm/g) on solid support was reacted with ethyl-3-mercaptopropionate (100 µL), and catalytic amount of sodium thiolate (0.5 equiv.), in DMF (1 mL). [b]Conversion to peptide thioester **3** was calculated from the absorbance at 220 nm using HPLC. SR = S-$(CH_2)_2$-$COOC_2H_5$. Table adopted from reference (1) with permission.

Next, the peptide backbone activation was investigated with threonine, because of its side chain functional group which exhibits similarity with serine. As expected, threonine underwent smooth cyclization under the optimized reaction conditions and generated ^{Me}Oxd activated intermediate (entry 6, Table 6.3). In a similar manner, cysteine generated cyclic thiozolidinone (Thz) intermediate under the reaction conditions (entry 7, Table 6.3). The activated intermediates; ^{Me}Oxd and Thz on thiolysis generated peptide thioester **3G**.

Importantly, our method was successfully applied for the synthesis of a multi-serine containing bioactive peptide thioester Ac-SGISGPLS-S$(CH_2)_2CO_2C_2H_5$ **3L**, a fragment of antimicrobial bovine beta-defensin 13,[24] from cyclic urethane activated peptide S(tBu)GIS(tBu)GPLS(tBu)-*Oxd* **2l** (Figure 6.4). This result and table 6.3 demonstrated the versatility of cyclic urethane technique in the synthesis of peptide thioesters.

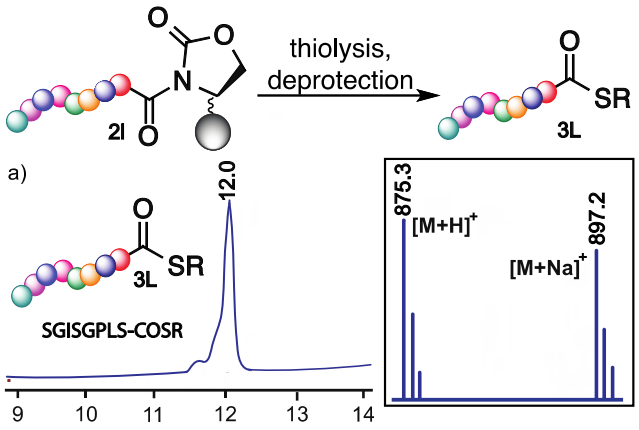


Figure 6.4 Cyclic urethane technique for the synthesis of the fragment of antimicrobial bovine beta-defensin 13 bioactive peptide thioester **3L** RSH = SH-$(CH_2)_2$-$COOC_2H_5$**.** (a) HPLC trace of pure bioactive peptide thioester **3L** and inset shows MS of peptide thioester **3L**. Figure adopted from reference (1) with permission.

Next, we utilized peptide thioesters Ac-GPMLA-COS-$(CH_2)_2$-$COOC_2H_5$ **3G** and Ac-AVGPPGVA-COS-$(CH_2)_2$-$COOC_2H_5$ **3h** generated by the cyclic urethane technique in Native Chemical Ligation (NCL) with an *N*-terminal cysteine containing peptide Ac-SFARC (Figures 6.5 and 6.6).

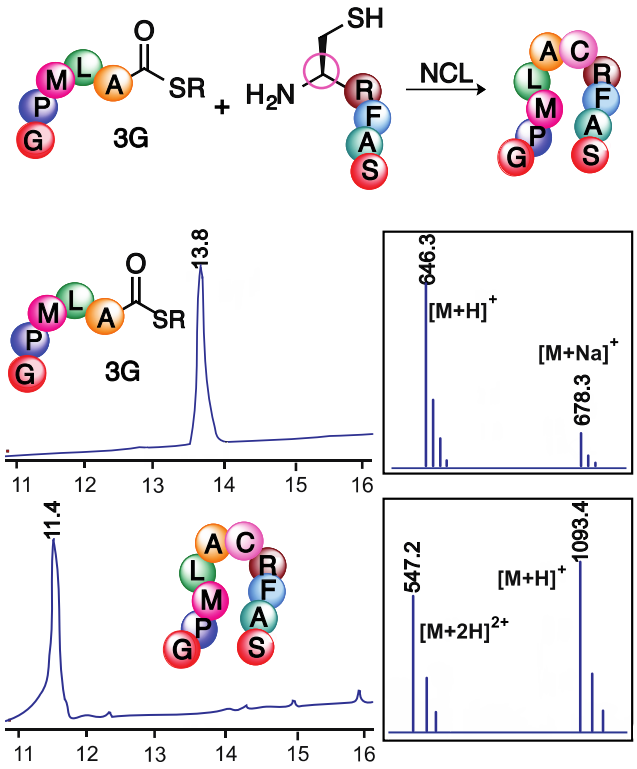


Figure 6.5 Native chemical Ligation of peptide thioester Ac-GPMLA-COSR **3G** with N-terminal cysteine peptide Ac-SFARC. RSH = HS-$(CH_2)_2$-$COOC_2H_5$. Figure adopted from reference (1) with permission.

Relatively dilute conditions were used for the peptide reactants (1.2 mM) as well as low concentrations of the 4-mercaptophenylacetic acid catalyst (25 mM). The ligation was completed after 24 hrs at 37 °C, demonstrating the utility of cyclic urethane technique in the synthesis of peptide thioesters for the formation of large peptides using NCL.

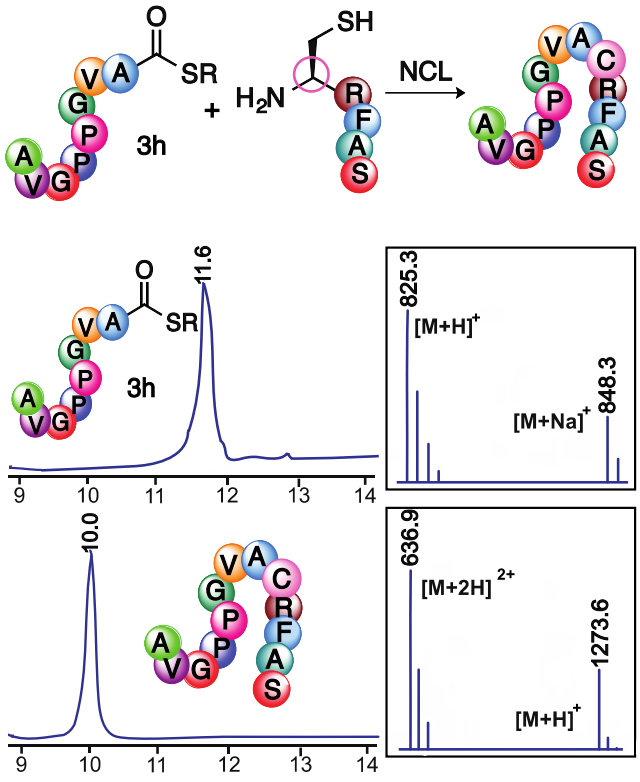


Figure 6.6 Native Chemical Ligation of peptide thioester Ac-AVGPPGVA-COSR **3h** with N-terminal cysteine peptide Ac-SFARC. RSH = SH-$(CH_2)_2$-$COOC_2H_5$. Figure adopted from reference (1) with permission.

6.6 CONCLUSIONS

We have developed a cyclic urethane activation technique for the rapid synthesis of peptide thioesters directly from solid support. The cyclic urethane technique carries out the acylation of the peptide backbone with Ser's side chain undergoing nucleophilic displacement by a thiol to release a corresponding thioester into the solution. The significant advantage of this technique is that unnatural handles or linkers are not required, and peptide synthesis is carried out by Fmoc-SPPS without a need for any special reaction conditions and can be fully automated, thus constituting a major advance in the field. It is noteworthy that only those peptides that undergo complete activation are displaced by the thiol into the solution thus delivering very pure thioesters in the solution. Moreover, activation of serine to the cyclic urethane moiety is compatible with both unhindered Ala and bulky Phe/Val amino acids next to serine and is resistant to epimerization. The results suggest that the cyclic urethane activation technique will be utilized in organic synthesis, including peptide and protein chemistry. Moreover, the cyclic urethane ring is stable and can be stored in the desiccator for long periods of time without any disintegration. Applicability of the strategy was demonstrated by the synthesis of long peptides using NCL. As a result, access to C-terminally activated cyclic urethane peptides will be applicable for the synthesis of complex post-translationally modified peptides and other biomolecules.

6.7 EXPERIMENTAL SECTION

General: All commercial materials (Aldrich, Fluka, Nova) were used without further purification. All solvents were reagent grade or HPLC grade (Fisher). Anhydrous THF, diethyl ether, CH2Cl2, and DMF were obtained from a dry solvent system (passed through column of alumina) and used without further drying. Conversion % were obtained by comparison of HPLC peak areas of products and starting material. HPLC was used to monitor the reaction progress. Materials. Fmoc-

amino acids were obtained from Nova Biochem (EMD Millipore Corporation)(Billerica, Massachusetts) and CreoSalus (Louisville, Kentucky). Rink amide resin was obtained from ChemPep Inc (Wellington, Florida). N,N,N′,*N*′-Tetramethyl-O-(1Hbenzotriazol-1-yl)uronium hexafluorophosphate (HBTU) was obtained from CreoSalus (Louisville, Kentucky). N,*N*′-Disuccinimidyl carbonate (DSC) was obtained from Nova Biochem, under (EMD Millipore Corporation) (Billerica, Massachusetts). 4-Dimethylaminopyridine (DMAP): Merck KGaA (Darmstadt, Germany). N,*N*-Dimethylformamide (DMF): Macron Fine Chemicals (Center Valley, Pennsylvania). Dichloromethane (DCM), acetonitrile, N,*N*- Diisopropylethylamine (DIEA), N,N'-diisopropylcarbodiimide (DIC), Triethylsilane (TES), were purchased from (EMD Millipore Corporation)(Billerica, Massachusetts). Piperidine was purchased from Alfa Aesar (Ward Hill, Massachusetts). Trifluoroacetic acid (TFA) was purchased from VWR 100 Matsonford Road Radnor, PA. Diethyl Ether, 4-mercaptophenylacetic acid (MPAA), sodium thiophenolate (NaSPh), and Tris(2-carboxyethyl)phosphine hydrochloride TCEP.HCl : Sigma Aldrich (St. Louis, Missouri). Water was purified using a Millipore MilliQ water purification system.

NMR: Analysis done by Ryan Cohen at Merck facilities located at Rahway, NJ. Proton NMR spectra were recorded on a 600 MHz spectrometer and carbon NMR spectra on a 151 MHz, spectrometer at ambient temperature. All NMR chemical shifts (δ) are referenced in ppm relative to residual solvent or internal tetramethylsilane. ^{1}H NMR chemical shifts referenced to residual DMSO-d_5 at 2.50 ppm, and ^{13}C NMR chemical shifts referenced to DMSO-d_6 at 39.52 ppm. Carbon NMR spectra are proton decoupled. NMR spectral data are reported as chemical shift (multiplicity, coupling constants (J), integration). Multiplicity is reported as follows: singlet (s), broad singlet (bs), doublet (d), doublet of doubles (dd), doublet of triplet (td), triplet (t) and multiplet (m). Coupling constant (J) in Hertz (Hz).

Semi-Preparative HPLC: Preparative HPLC chromatography (HPLC) was performed on Beckman Coulter equipped with System Gold 168 detector and 125P solvent module HPLC with C-18 reversed-phase column. All separations involved a mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1% FA (v/v) in acetonitrile (solvent B). Semi-preparative HPLC method using a linear gradient of 0– 80% acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 3.0 mL min-1 . The eluent was monitored by absorbance at 220 nm unless otherwise noted.

Analytical HPLC: Analytical HPLC chromatography (HPLC) was performed on an Agilent 1100 series HPLC equipped with a 4.6x150 mm (5µm) C-18 reversed-phase column. All separations involved mobile phase of 0.1% FA (v/v) in water (solvent A) and 0.1% FA (v/v) in acetonitrile (solvent B). Peptide compositions were evaluated by analytical reverse phase HPLC using a gradient of 0.1% FA in acetonitrile versus 0.1% FA in water. HPLC method using a linear gradient of 0–80% 0.1% FA (v/v) acetonitrile in 0.1% aqueous FA over 30 min at room temperature with a flow rate of 1.0 mL min-1 . The eluent was monitored by absorbance at 220 nm unless otherwise noted.

LCMS: Mass spectrometry was performed using ultra high-performance liquid chromatography-mass spectrometry using the Agilent 1100 Series LCMSD VL MS Spectrometer.

Fmoc Solid-Phase Peptide Synthesis:[25] Peptides were synthesized manually on a 0.25 mmol scale using Rink amide resin. The Fmoc–group was deprotected using 20% piperidine–DMF for 20 min to obtain a deprotected peptide-resin. Fmoc-protected amino acids (1.25 mmol) were sequentially coupled on the resin using a HBTU (1.25 mmol) and DIEA (1.25 mmol) for 2 hrs at room temperature. Peptides were synthesized using standard protocols.1 The peptide was cleaved from the resin using a cocktail of 95:2.5:2.5, trifluoroacetic acid:triisopropyl silane:water for 2 hrs.

The resin was removed by filtration and the resulting solution was concentrated. The oily residue was triturated with diethyl ether to obtain a white suspension. The resulting solid was purified by HPLC.

General procedure for the native chemical ligation for synthesis of large peptides: 900 μL of ligation buffer pH 7.1 (0.2 M sodium phosphate buffer and 6 M guanidine hydrochloride) was added to 1.0 mg of the peptide thioester 3G/3h (final concentration 1.3 mM) and 1.3 mg of peptide CRAFS (final concentration 2 mM) under argon. The mixture was treated with MPAA (3.8 mg, 20 mM) and TCEP.HCl (5.5 mg, 17 mM) and pH of solution was adjusted to 7.0 by using 2N NaOH solution. The ligation was carried out at 37°C and monitored by RP HPLC. After completion of reaction, TCEP.HCl (30 mM in water) was added into the reaction mixture and it was left for stirring at room temperature for 1 hr followed by purification of ligated peptide using RP-HPLC.

6.8 REFERENCES

1. Elashal, H.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, *8*, 117.
2. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.***2000**, *69*, 923.
3. Unverzagt, C.; Kajihara, Y. *Chem. Soc. Rev.* **2013**, *42*, 4408.
4. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent,S. B. *Science* **1994**, *266*, 776.
5. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci.* **1999**, *96*, 10068.
6. Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **1998**, *39*, 8669.
7. Shin, Y.; Winans,K. A.; Backes, B. J.;Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684.
8. Mezo, A. R.; Cheng, R. P.; Imperiali, B. *J. Am. Chem. Soc.* **2001**, *123*, 3885.

9. Kawakami, T.; Sumida, M.; Nakamura, K.; Vorherr, T.; Aimoto, S. *Tetrahedron Lett.* **2005**, *46*, 8805.
10. Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *Org. Lett.* **2009**, *11*, 823.
11. Kang, J.; Richardson, J. P.; Macmillan, D. *Chem. Commun.* **2009**, 407.
12. Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. *Org. Lett.* **2010**, *12*, 5238.
13. Zheng, J. S.; Chang, H. N.; Wang, F. L.; Liu, L. *J. Am. Chem. Soc.* **2011**, *133*, 11080.
14. Burlina, F.; Papageorgiou, G.; Morris, C.; White, P. D.; Offer, J. *Chem. Sci.* **2014**, *5*, 766.
15. Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas, A. F.; Aucagne, V. *Chem. Sci.* **2016**, *7*, 339.
16. Camarero, J. A.; Hackel, B. J.; De Yoreo, J. J.; Mitchell, A. R. *J. Org. Chem.* **2004**, *69*, 4145.
17. Zheng, J. S.; Yu, M.; Qi, Y. K.; Tang, S.; Shen, F.; Wang, Z. P.; Xiao, L.; Zhang, L.; Tian, C. L.; Liu, L. *J. Am. Chem. Soc.* **2014**, *136*, 3695.
18. Li, Y. M.; Li, Y. T.; Pan, M.; Kong, X. Q.; Huang, Y. C.; Hong, Z. Y.; Liu, L. *Angew. Chem. Int. Ed.* **2014**, *53*, 2198.
19. Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2008**, *47*, 6851.
20. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197.
21. Tofteng, A. P.; Sørensen, K. K.; Conde-Frieboes, K. W.; Hoeg-Jensen, T.; Jensen, K. J. *Angew. Chem. Int. Ed.* **2009**, *48*, 7411.
22. Mahto, S. K.; Howard, C. J.; Shimko, J. C.; Ottesen, J. J. *ChemBioChem* **2011**, *12*, 2488.
23. Mong, S. K.; Vinogradov, A. A.; Simon, M. D.; Pentelute, B. L. *ChemBioChem* **2014**, *15*, 721.

24. Selsted,M. E.; Tang, Y. Q.; Morris, W. L.; McGuire, P. A.; Novotny, M. J.; Smith, W.; Henschen, A. H.; Cullor, J. S. *J. Biol. Chem.* **1993**, *268*, 6641.

CHAPTER 7: CONCLUSIONS AND CONTRIBUTIONS TO KNOWLEDGE

7.1 CONCLUSIONS AND CONTRIBUTIONS TO KNOWLEDGE MADE IN THIS THESIS

7.1.1 SITE-SELECTIVE CHEMICAL CLEAVAGE OF PEPTIDE BONDS

The second chapter of this thesis focuses on the development of an artificial protease mimic by site-selective modification of Ser, Thr, Cys, and Glu residues. In this study, we optimized formation of the cyclic urethane-derived moieties and applied them to site-selective cleavage of peptide bonds. The substrate scope was explored to determine the effect of neighboring residues on the effectiveness of cleavage. Generally, neighboring residues had no effect; however, bulky residues such as Val required longer incubation times in buffer to yield the cleavage product. Notably, cleavage proceeded even in the presence of Pro which is of significance due to the resistance of amide bonds neighboring Pro to enzymatic cleavage. Furthermore, the methodology was successfully applied to peptides containing modifications such as D, β, and *N*-methylated residues which are known to hinder enzymatic cleavage in native sequences. Finally, the scope of the reaction was extended to the scission of bioactive peptides including an amyloid beta peptide and a fragment of antimicrobial Bovine β-defensin 13. The significance of this research lies in the ability to cleave peptide bonds even in the presence of post-translational modifications and mutations which is vital in probing age-related disease and the function of these modifications. This study was recognized by a publication in *Chemical Communications*, and it set the foundation for further exploration of oxazolidinone mediated peptide bond cleavage and its various applications.

7.1.2 SEQUENCING OF PROMINENT NATURAL PRODUCTS

The third chapter of this thesis focuses on expanding the application of cyclic urethane-mediated cleavage of peptide bonds to sequencing of cyclic peptides and lasso peptides which are natural products with unique therapeutic and structural properties. This study was accomplished in collaboration with the Link group at Princeton University, which facilitated expanding the scope of the study to lasso peptides. By modifying Ser to a cyclic urethane moiety, linearization of these unique peptides became a feasible task. The methodology was applied to a broad substrate scope including the ability to modify peptides containing Thr, Cys, Glu, and unnatural amino acids such as homoserine which account for 60% of naturally occurring cyclic peptides. Utilizing this methodology, the cleavage and sequencing of stable N-methylated cyclic peptides, Polycarponin C, and a Somatostatin analogue in addition to lasso peptides lariatin A, albusnodin, and benenodin-1 have also been validated. Additionally, we created a peptide rotaxane by the selective cleavage of benenodin-1. The impact of this work lies in application of a chemical cleavage methodology for the first time to effectively cleave lasso peptides which is a major milestone in simplifying sequence determination of lasso peptides, studying the lasso topology, monitoring the effect of cleavage on sterically locked residues, using lasso peptides as scaffolds for making rotaxane type interlocked machines, and providing a platform for discovering newly isolated biologically active peptides. This work was reported in a publication in *Angewandte Chemie*.

7.1.3 SEQUENCING OF ONE-BEAD-ONE-COMPOUND LIBRARIES

In chapter four, oxazolidinone mediated cleavage of peptide bond is explored on solid support to facilitate sequencing of HIT peptides post high throughput screening against biological targets. Since cyclic peptides are attractive therapeutics for drug discovery due to high proteolytic stability, permeability, and high binding affinity, simplifying the sequencing process is of utmost

importance in developing lead compounds. Toward this end, we developed a one bead one compound (OBOC) dual ring opening/cleavage approach for sequencing of hit cyclic peptides by selective modification of a serine residue to an oxazolidinone moiety within the cyclic peptide and as a linker to allow for the liberation of the protected linear peptide into solution. The major advantage of this method is that it is direct, without the need to tag each bead, which makes the synthesis of libraries of cyclic peptides easier. In addition, our method has high substrate scope and ring opening of cyclic peptides can be achieved at serine, threonine, cysteine, and glutamic acid residues, which is a significant advantage over the current OBOC approaches that rely either on a methionine residue or an unnatural entity. This work was published in *Organic Letters.*

7.1.4 SYNTHESIS OF C-TERMINALLY MODIFIED PEPTIDES

In this chapter, a universal solid phase strategy for the synthesis of various C-terminal modified peptides is reported; the method is independent of the type of resins, linkers, and unnatural moieties typically needed for C-terminal modifications. The technique proceeds by the modification of C-terminal serine to a cyclic urethane moiety which results in the activation of the backbone amide chain for the nucleophilic displacement by various nucleophiles to generate C-terminally modified acids, esters, *N*-alkyl amides, alcohols, and hydrazides. This cyclic urethane technique (CUT) also provides a general strategy for synthesis of C-terminal protected peptides which can be used for convergent synthesis of large peptides. A universal solid phase strategy has been effectively developed and applied for the synthesis of various C-terminal modified peptides, which are of therapeutic importance due to biological properties that accompany the attachment of various C-terminal functional groups. This cyclic urethane application has been reported in a publication in *Chemical Communications.*

7.1.5 SYNTHESIS OF THIOESTERS DERIVED FROM FMOC SPPS

In chapter six, we developed a versatile approach for direct synthesis of peptide thioesters from a solid support utilizing Fmoc chemistry. Fmoc SPPS of peptide thioesters are of high significance because they are precursors for convergent synthesis of proteins via NCL. As a result, there is a great necessity for an effective Fmoc solid-phase approach for the preparation of peptide thioesters. To this end, the modified serine acts as a thioester surrogate and undergoes nucleophilic displacement with a thiol to generate peptide thioesters directly from the solid support. Applicability of the strategy was demonstrated by the synthesis of long peptides using NCL. As a result, access to C-terminally activated cyclic urethane peptides will be applicable for the synthesis of complex post-translationally modified peptides and other biomolecules. The significance of this work was recognized by a publication in *Chemical Science.*

7.2 PUBLICATIONS, CONFERENCE PRESENTATIONS, AND AWARDS

Manuscripts Accepted for Publication

1. **Elashal, H. E.;** Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A. J.; Raj, M. *Angew. Chem.* **2018**, DOI: 10.1002/anie.201801299.

2. **Elashal, H. E**.; Cohen, R. D.; Elashal, H. E.; Raj, M. *Org. Lett.* **2018**, *20* (8), 2374.

3. Zong, C.; Cheung-Lee W. L.; **Elashal, H. E**.; Raj, M.; Link, A.J. *Chem. Commun.* **2018**, *54*, 1339.

4. **Elashal, H**.**E**.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, *8*, 117.

5. **Elashal, H. E**.; Cohen, R. D.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699

6. **Elashal, H**.**E**.; Raj, M. *Chem. Commun.* **2016,** *52*, 6304.

Poster Presentations

1. **Elashal, H**.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A. J.; Raj. M. Lasso Peptides: Sequence Determination, Topology Analysis, and Rotaxane Formation. NYAS Chemical Biology Year-End Symposium, New York, NY, USA, May 23rd, 2018.

2. **Elashal, H**.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A. J.; Raj, M. Sequence and Topology Analysis of Cyclic and Lasso Peptides. Chemistry and Biology of Peptides Gordon Research Conference, Ventura, CA, USA, February 11-16, 2018.

3. **Elashal, H**.; Raj, M. Oxazolidinone Mediated Cleavage of Peptide Bonds and its Applications. 25th American Peptide Symposium, Whistler, Canada, June 17-22, 2017.

4. **Elashal, H**.; Cohen, R. D.; Raj. M. Oxazolidinone Mediated Sequencing of Cyclic Peptides. NYAS Chemical Biology Year-End Symposium, New York, NY, USA, May 24th, 2017.

5. **Elashal, H**; **Cohen, R. D.;** Elashal, H. E.; Raj, M. Oxazolidinone Mediated Cleavage of Peptide Bonds. 21st Petersheim Academic Exposition, Seton Hall University, April 18, 2017.

6. **Elashal, H**.; Raj, M. Serine Selective Cleavage of Peptide Bonds and its Applications. NYAS Chemical Biology Year-End Symposium, New York, NY, USA, May 25th, 2016.

7. **Elashal, H**.; Raj, M. Site-Selective Chemical Cleavage of Peptide Bonds. 20th Petersheim Academic Exposition, Seton Hall University, April 19, 2016.

Oral Presentations

1. **Elashal, H**. Oxazolidinone Mediated Sequencing of Cyclic Peptides. NYAS Chemical Biology Year-End Symposium, New York, NY, USA, May 24th, 2017.

Awards

1. Robert De Simone Graduate Fellowship, Department of Chemistry and Biochemistry, Seton Hall University, September 2016-May 2018.

2. Petersheim Award, Department of Chemistry and Biochemistry, Seton Hall University, May 2018

3. Honorable Mention Poster Award Sequence and Topology Analysis of Cyclic and Lasso Peptides. Chemistry and Biology of Peptides Gordon Research Conference, Ventura, CA, USA, February 11-16, 2018.

4. Top Biochemistry Student Award, Department of Chemistry and Biochemistry, Seton Hall University, May 2017.

Bibliography

1. Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Chen, V.; Crews, P.; Jacobson, M. P.; Lokey, R. S. *Future Med. Chem*. **2015**, *7*(16): 2121.
2. Alsina, J., Albericio, F. *Pept. Science* **2003**, *71*, 454.
3. Alsina, J.; Yokum, T. S.; Albericio, F.; Barany, G. *J. Org. Chem*. **1999**, *64*, 8761.
4. Antonelli, G.; Turriziani, O. *Int J Antimicrob Agents* **2012**, *40* (2), 95.

5. Baker, E. S.; Clowers, B. H.; Li, F.; Tang, K.; Tolmachev, A. V.; Prior, D. C.; Belov, M. E.; Smith, R. D. *J. Am. Soc. Mass Spectrom*. **2007**, *18*, 1176.
6. Bergmann, M.; Zervas, L. *Berichte der deutschen chemischen Gesellschaft* **1932**, *65*, 1192.

7. Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2008**, *47*, 6851.
8. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc*. **2015**, *137*, 7197.
9. Bramson, H. N.; Thomas, N. E.; Kaiser, E. T. *J. Biol. Chem.* **1985,** *260*, 15452.

10. Burlina, F.; Papageorgiou, G.; Morris, C.; White, P. D.; Offer, J. Chem. Sci. 2014, 5, 766.
11. *Butler, M.S. J. Nat. Prod* **2004**, *67* (12), 2141.

12. Caboche, S.; Pupin, M.; Leclére, V.; Fontaine, A.; Jacques, P.; Kucherov, G.; *Nucleic Acids Res.* 2008, *36*, D326.
13. Camarero, J. A.; Hackel, B. J.; De Yoreo, J. J.; Mitchell, A. R. *J. Org. Chem*. **2004**, *69*, 4145.
14. Campbell, M. K.; Farrell, S. O. *Biochemistry*, 7th ed.; Brooks/Cole, Cengage Learning, 2012.
15. Carpino, L. A.; Han, G. Y. *J. Am. Chem. Soc.* **1970**, *92*, 5748.
16. Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1979**, *44*, 3739.
17. Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem Sci*. **2011**, *2*, 1666.
18. Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.
19. Chatterjee, J.; Laufer, B.; Kessler, H. *Nat. Protoc.* **2012**, *7*, 432.

20. Checco, J. W.; Lee, E. F.; Evangelista, M.; Sleebs, N. J.; Rogers, K.; Pettikiriarachchi, A.; Kershaw, N. J.; Eddinger, G. A.; Belair, D. G.; Wilson, J. L.; Eller, C. H.; Raines, R. T.; Murphy, W. L.; Smith, B. J.; Gellman, S. H.; Fairlie, W. D. *J. Am. Chem. Soc.* **2015**, *137* (35), 11365.
21. Chekan, J. R.; Koos, J. D.; Zong, C.; Makimov, M. O.; Link, A. J.; Nair, S. K. *J. Am. Chem. Soc.* **2016**, *138* (50), 16452.
22. Cheung, W. L.; Pan, S. J.; Link, A. J. *J. Am. Chem. Soc.*, **2010**, *132* (8), 2514.

23. Chugh, J. K.; Wallace, B. A. *Biochem. Soc. Trans*. **2001**, *29*, 565.
24. Colgrave, M. L.; Craik, D. J. *Biochemistry*, **2004**, *43* (20), 5965.

25. Davies, M.; Bradley, M. *Angew. Chem., Int. Ed.* **1997**, *36*, 1097.
26. Davies, M.; Bradley, M. *Tetrahedron* **1999**, *55*, 4733.
27. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.***2000**, *69*, 923.
28. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent,S. B. *Science* **1994**, *266*, 776.
29. De Marco, R.; Tolomelli, A.; Campitello, M.; Rubini, P.; Gentilucci, L. *Org. Biomol. Chem.* **2012**, *10*, 2307.
30. Degani, Y.; Patchornik, A. *Biochemistry.* **1974**, *13*, 1.

31. Diaz, G.; De Freitas, M. A. A.; Ricci-Silva, M. E.; Diaz, M. A. N. *Molecules* **2014**, *19* (6), 7429.
32. Driggers, E. M.; Hale, S.P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discov*. **2008**, *7*, 608.
33. Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc.* **1953**, *75*, 4879.
34. Dumy, P.; Keller, M.; Ryan, D. E.; Rohwedder, B.; Wöhr, T.; Mutter, M. *J. Am. Chem. Soc.* **1997**, *119*, 918.
35. Eckart, K.; Schwarz, H.; Tomer, K. B.; Gross, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 6765.

36. Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1*, 80.

37. Edman, P.; Högfeldt, Erik; Sillén, Lars Gunnar; Kinell, Per-Olof. *Acta Chem. Scand*. **1950**, *4*, 283.
38. Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A.J.; Raj, M. *Angew. Chem.* **2018**, DOI: 10.1002/anie.201801299.
39. Elashal, H. E.; Cohen, R. D.; Raj, M. *Chem. Commun*. **2016**, *52*, 9699.

40. Elashal, H.E.; Cohen, R. D.; Elashal, H. E.; Raj, M. *Org. Lett*. **2018**, *20* (8), 2374.

41. Elashal, H. E.; Raj, M. *Chem. Commun.* **2016**, *52*, 6304.
42. Elashal, H. E.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, ***8***, 117.

43. Esquivel, R.; Molina-Espíritu, M. In *Advances in Quantum Mechanics*; INTECH, 2013; pp 641–669.
44. Evans, D. A.; Helmchen, G. *Asymmetric Synth. Essentials* **2007**, 3.
45. Evans, D. A.; Wu, L. D.; Wiener, J. J. M.; Johnson, J. S.; Ripin, D. H. B.; Tedrow, J. S. *J. Org. Chem.* **1999**, *64* (17), 6411.
46. Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2004**, *43*, 2109.
47. Fields, G. B.; Simon, M. I.; Abelson, J. N. *Methods in Enzymology*; Academic Press, 1997; Vol. 448.
48. Fischer, E.; Fourneau, E. *Berichte der deutschen chemischen Gesellschaft* **1901**, *34*, 2868.
49. Fosgerau, K.; Hoffmann, T. *Drug Discov. Today* **2015**, *20*, 122.
50. Fujita, S. I.; Kanamaru, H.; Senboku, H.; Arai, M. *Int. J. Mol. Sci.* **2006**, *7*, 438.
51. Furka, Á.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pep. Pro. Res*. **1991**, *37*, 487.
52. Getz, J. A.; Rice, J. J.; Daugherty, P. S. *ACS Chem. Biol.*, **2011**, *6* (8), 837.

53. Gokhale, A. S.; Satyanarayanajois, S. *Immunotherapy* **2014**, *6*, 755.

54. Golias, C.; Charalabopoulos, A.; Stagikas, D.; Charalabopoulos, K.; Batistatou, A. *Hippokratia.* **2007**, *11*, 124.
55. Gongora-Benitez, M.; Cristau, M.; Giraud, M.; Tulla-Puche, J.; Albericio, F. *Chem. Commun*. **2012**, *48*, 2313.
56. Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F. *ACS Comb. Sci*. **2013**, *15*, 217.
57. Gorman, J. J.; Wallis, T. P.; Pitt, J. J. *Mass Spectrum Rev.* **2002**, *21*, 183.

58. Griffiths, K. E.; Stoddart, J. F. *Pure Appl. Chem.* **2008**, *80* (3), 485.

59. Gross, E. *Methods Enzymol*. **1967**, *11*, 238.
60. Guillier, F.; Orain, D.; Bradley, M. *Chem. Rev.* **2000**, *100*, 3859.
61. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci*. **1999**, *96*, 10068.
62. Hasinoff, L.; Takats, J.; Zhang, X. W.; Bond, A. J.; Rogers,R. D. *J. Am. Chem. Soc.* **1994,** *116*, 8833.
63. Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. *Acc. Chem. Res*. **2015**, *48*, 1909.
64. Holmes, T. J.; Lawton, R. G. *J. Am. Chem. Soc.* **1977**, *99*, 1984.

65. Houben-Weyl. *Methods of Organic Chemistry*, 4th ed.; Thieme, 1995.
66. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84.
67. Howell, S. M.; Fiacco, S. V.; Takahashi, T. T.; Jalali-Yazdi, F.; Millward, S. W.; Hu, B.; Wang, P.; Roberts, R. W. *Sci. Rep.* **2014**, *4*, 6008.
68. Hoyer, D.; Cho, H.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 3249.

69. Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. *J. Am. Chem. Soc.* **2006**, *128* (23), 7486.
70. Joo, S. H. *Biomolecules & Therapeutics* **2012**, *20*, 19.
71. Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. *J. Am. Chem. Soc*. **2006**, *128*, 13000.
72. Kaiser, R., Metzka, L. *Anal. Biochem.* **1999**, *266*, 1.

73. Kaminskaia, N. V.; Johnson, T. W.; Kostic, N. M. *J. Am. Chem. Soc.***1999**, *121*, 8663.

74. Kang, J.; Richardson, J. P.; Macmillan, D. *Chem. Commun.* **2009**, *0*, 407.
75. Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H. *Nat. Chem. Biol*. **2009**, *5*, 888.
76. Kawakami, T.; Sumida, M.; Nakamura, K.; Vorherr, T.; Aimoto, S. *Tetrahedron Lett*. **2005**, *46*, 8805.
77. King, J. L.; T. H. Jukes. *Science,* **1969**, *164*, 788.

78. Kita, Y.; Nishii, Y.; Higuchi, T.; Mashima, K. *Angew. Chem. Int. Ed.* **2012**, *51*, 5723.

79. Koehn, F. E.; Carter, G. T.; *Nat. Rev. Drug Discovery*, **2005**, *4*, 206.

80. Kolb, H.C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed*. **2001**, *40*, 2004.

81. Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411.
82. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82.
83. Lee, J. H.; Meyer, A. M.; Lim, H.-S. *Chem. Commun*. **2010**, *46*, 8615.
84. Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett*. **1998**, *39*, 8669.
85. Li, Y. M.; Li, Y. T.; Pan, M.; Kong, X. Q.; Huang, Y. C.; Hong, Z. Y.; Liu, L. *Angew. Chem. Int. Ed*. **2014**, *53*, 2198.
86. Liang, X.; Girard, A.; Biron, E. *ACS Comb. Sci*. **2013**, *15*, 535-540.
87. Liu, R.; Marik, J.; Lam, K. S. *J. Am. Chem. Soc.* **2002**, *124*, 7678.
88. Lundblad, R. L. Techniques in Protein Modification, CRC Press, Boca Raton, Florida, 1995.
89. Maeda, H.; Takata, T.; Fujii, N.; Sakaue, H.; Nirasawa, S.; Takahashi, S.; Sasaki, H.; Fujoo, N. *Anal. Chem.* **2015**, *87*, 561.
90. Mahoney, W. C.; Smith, P. K.; Hermodson, M. A. *Biochemistry*. **1981**, *20*, 443.
91. Mahto, S. K.; Howard, C. J.; Shimko, J. C.; Ottesen, J. J. *ChemBioChem.* **2011**, *12*, 2488.
92. Maksimov, M. O.; Koos, J. D.; Zong, C.; Lisko, B.; Link, A. J. *J. Biol. Chem.* **2015**, *290*, 30806.
93. Maksimov, M. O.; Link, A. J. *J. Am. Chem. Soc*. **2013**,*135*, 12038.
94. Maksimov, M. O.; Link, A. J. *Nat. Prod. Rep*. **2012**, *29*, 996.
95. Maria, C. M.-C.; Silvana, L. G.; Soledad, L. S.; Juan, M. G.-M.; Rosa, E.-B.; Fernando, A.; Osvaldo, C.; Silvia, A. C. Curr. *Pharma. Biotech*. **2016**, *17*, 449.
96. May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A. *Anal. Chem*. **2014**, *86*, 2107.
97. McGregor, D. P. *Curr. Opin. Pharma*. **2008**, *8*, 616.
98. McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int. J. Mass Spectrom*. **2005**,
99. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149.
100. Merrifield, R. B. *Recent Prog. Horm. Res.* **1967,** *23*, 451.
101. Mezo, A. R.; Cheng, R. P.; Imperiali, B. *J. Am. Chem. Soc*. **2001**, *123*, 3885.
102. Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. *ACS Chem. Bio.* **2007**, *2*, 625.
103. Millward, S. W.; Takahashi, T. T.; Roberts, R. W. *J. Am. Chem. Soc*. **2005**, *127*, 14142.
104. Milovic´, N. M.; Kostic´, N. M., *J. Am. Chem. Soc.* **2003**, *125*, 781.
105. Mong, S. K.; Vinogradov, A. A.; Simon, M. D.; Pentelute, B. L. *ChemBioChem* **2014**, *15*, 721.
106. Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827.
107. Mótyán, J.; Tóth, F.; Tőzsér, J. *Biomolecules* **2013**, *3* (4), 923.
108. Ngoka, L. C. M.; Gross, M. L. *J. Am. Soc. Mass Spec.* **1999**, *10*, 732.
109. Ogino, H.; Ohata, K. *Inorg. Chem.* **1984**, *23* (21), 3312.

110. Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. *Org. Lett*. **2010**, *12*, 5238.
111. Pandit, N.; Singla, R. K.; Shrivastava, B. *Int. J. Med. Chem.* **2012**, *2012*, 1.
112. Parac, T. N.; Ullmann, G. M.; Kostic, N.M. *J. Am. Chem. Soc.* **1999**, *121*, 3127.

113. Peters, C.;Waldmann, H. *J. Org. Chem*. **2003**, *68*, 6053.
114. Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. *J. Comb. Chem.* **2003**, *5*, 33.

115. Redman, J. E.; Wilcoxen, K. M.; Ghadiri, M. R. *J. Comb. Chem*. **2003**, *5*, 33.
116. Rich DH, Singh J. The carbodiimide method. In: *Gross* E, Meienhofer J, editors. The Peptides. Vol. 1. Academic Press; New York: 1979. pp. 241–314
117. Rink, H. *Tetrahedron Lett.* **1987**, *28* (33), 3787.
118. Roberts, R. W.; Szostak, J. W. *Proc. Nat. Acad. Sci*. **1997**, *94*, 12297.
119. Russell, P. J. *iGenetics: A Molecular Approach Third Edition*, 3rd ed.; Pearson Education, 2010.
120. Sabatino, D.; Proulx, C.; Klocek, S.; Bourguet, C. B.; Boeglin, D.; Ong, H.; Lubell, W. D. *Org. Lett.* **2009**, *11* (16), 3650.
121. Sakakibara, S.; Shimonishi, Y.; Kishida, Y.; Okada, M.; Sugihara, H *Bull Chem. Soc. Jpn.* **1967**, *40*, 2164.
122. Sanger, F. *Biochem. J.* **1945**, *39*, 507.

123. Sanger, F. *Les Prix Nobel* **1958**, 134.

124. Schacherl, M.; Pichlo, C.; Neundorf, I.; Baumann, U. *Structure.* **2015**, *23* (9), 1632.

125. Schepartz, A.; Cuenoud, B. *J. Am. Chem. Soc.* **1990**, *112*, 3247.

126. Schilling, B.; Wang, W.; McMurray, J. S.; Medzihradszky, K. F. *Rapid Commun. Mass Spec.* **1999**, *13*, 2174.
127. Scott, C. P.; Abel-Santos, E.; Wall, M.; Wahnon, D. C.; Benkovic, S. *J. Proc. Nat. Acad. Sci*. **1999**, *96*, 13638.
128. Seki, Y.; Tanabe, K.; Sasaki, D.; Sohma, Y.; Oisaki, K.; Kanai, M. *Angew. Chem.* **2014**, *126*, 6619. (*Angew Chem, Int. Ed.* **2014**, *53*, 6501).
129. Selsted,M. E.; Tang, Y. Q.; Morris, W. L.; McGuire, P. A.; Novotny, M. J.; Smith, W.; Henschen, A. H.; Cullor, J. S. *J. Biol. Chem*. **1993**, *268*, 6641.
130. Semenov, A. N.; Gordeev, K. *Int. J. Pept. Protein Res*. **1995**, *45*, 303.
131. Shen, Y. X.; Gibson, H. W.; *Macromolecules*, **1992**, *25* (7), 2058.

132. Shin, Y.; Winans,K. A.; Backes, B. J.;Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc*. **1999**, *121*, 11684.
133. Shinde, N. V.; Dhake, A. S.; Haval, K. P. *Orient. J. Chem*. **2016**, *32* (1).

134. Siegel, M. M.; Huang, J.; Lin, B.; Tsao, R.; Edmonds, C. G. *Bio. Mass Spec.* **1994**, *23*, 186.
135. Simpson, L. S.; Kodadek, T. *Tetrahedron Lett.* **2012**, *53*, 2341.

136. Smith, J. B. *ENCYCLOPEDIA OF LIFE SCIENCES*; Macmillan Publishers Ltd, Nature Publishing Group, 2001.
137. Songster, M. F.; Barany, G. Methods in Enzymology, Academic Press, 1997, vol. Volume 289, pp. 126.
138. Stawikowski, M.; Fields, G. B. *Curr. Protoc. Protein Sci.* **2002**, *26*, 1.
139. T. W. G. a. P. G. M. Wuts, John Wiley & Sons, New York, 1999.
140. Tanabe, K.; Taniguchi, A.; Matsumoto, T.; Oisaki, K.; Sohma, Y.; Kannai, M. *Chem. Sci.* **2014**, *5*, 2747.
141. Tang, Z.; Yang, Z. H.; Cun, L. F.; Gong, L.Z.; Mi, A. Q.; Jiang, Y. Z. *Org. Lett.* **2004**, *6*, 2285.
142. Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas, A. F.; Aucagne, V. *Chem. Sci.* **2016**, *7*, 339.
143. Thieriet,N.; Guibe, F; Albericio,F. *Org. Lett.* **2000**, *2*, 1815.
144. Thorner, J.; Emr, S. D.; Abelson, J. N. *Methods Enzymol.* **2000**, *326*, 200.

145. Tofteng, A. P.; Sørensen, K. K.; Conde-Frieboes, K. W.; Hoeg-Jensen, T.; Jensen, K. J. *Angew. Chemie - Int. Ed.* **2009**, *48* (40), 7411.
146. Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *Org. Lett.* **2009**, *11*, 823.
147. Uhlig, T.; Kyprianou, T.; Martinelli, F. G.; Oppici, C. A.; Heiligers, D.; Hills, D.; Calvo, X. R.; Verhaert, P. *EuPA Open Proteomics* **2014**, *4*, 58.
148. Unverzagt, C.; Kajihara, Y. *Chem. Soc. Rev.* **2013**, *42*, 4408.
149. Vagner, J.; Qu, H.; Hruby, V. J. *Curr. Opin. Chem.* **2008**, *12*, 292.
150. Valentine, S. J.; Kulchania, M.; Barnes, C. A. S.; Clemmer, D. E. *Int. J. Mass Spectrom.* **2001**, *212*, 97.
151. Vézina-Dawod, S.; Bédard, F.; Porte, K.; Biron, E. *Org. Lett.* **2016**, *18*, 1174.
152. Walker, J. M. The Protein Protocols Handbook, Humana Press, Totowa, NJ, 2002.

153. Wang, S. *J. Am. Chem. Soc.* **1973**, *95* (4), 1328.
154. Wellings, D. A. A., E. G. B. Academic Press, San Diego, 1997, p. 54.
155. White, C. J.; Yudin, A. K. *Nat. Chem.* **2011**, *3*, 509.
156. Winn, M.; Fyans, J.K.; Zhuo, Y.; Michlefield, J. *Nat. Prod. Rep.* **2016**, *33*, 317.

157. Wöhr, T.; Mutter, M. *Tetrahedron Lett.* **1995**, *36*, 3847
158. Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218.
159. Woo, Y. H.; Mitchell, A. R.; Camarero, J.A. *Int. J. Pept. Res. Therap.* **2007**, *13*, 181.
160. Xue, M.; Yang, Y.; Chi, X.; Yan, X.; Huang, F. *Chem. Rev.* **2015**, *115* (15), 7398.

161. Yan, B. *Comb. Chem. High Throughput Screen.* **1998**, *1*, 215.
162. Youngquist, R. S.; Fuentes, G. R.; Lacey, M. P.; Keough, T. *J. Am. Chem. Soc.* **1995**, *117*, 3900.
163. Yudin, A. K. *Chem. Sci.* **2015**, *6*, 30.
164. Zheng, J. S.; Chang, H. N.; Wang, F. L.; Liu, L. *J. Am. Chem. Soc.* **2011**, *133*, 11080.
165. Zheng, J. S.; Yu, M.; Qi, Y. K.; Tang, S.; Shen, F.; Wang, Z. P.; Xiao, L.; Zhang, L.; Tian, C. L.; Liu, L. *J. Am. Chem. Soc.* **2014**, *136*, 3695.

166. Zhu, L.; Qin, L.; Parac, T. N.; Kostoc, N. M. *J. Am. Chem. Soc.* **1994**, *116*, 5218.

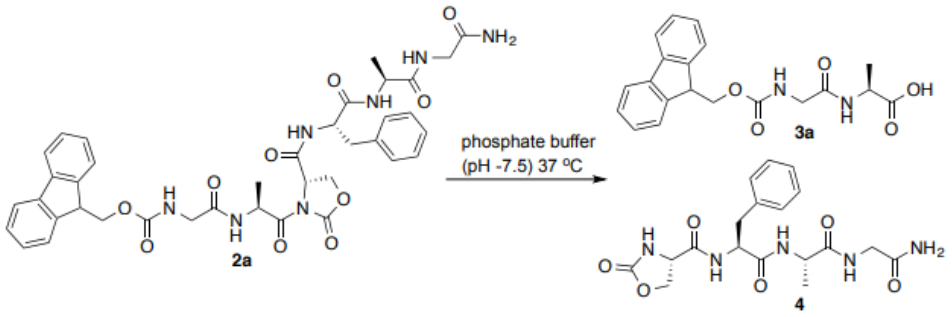
167. Zong, C.; Cheung-Lee, W.; Elashal, H. E.; Raj, M.; Link, A. J. *Chem. Commun.* **2018**, *54*, 1339.
168. Zong, C.; Wu, M. J.; Qin, J. Z.; Link, A. J. *J. Am. Chem. Soc*. **2017**, *139*, 10403.

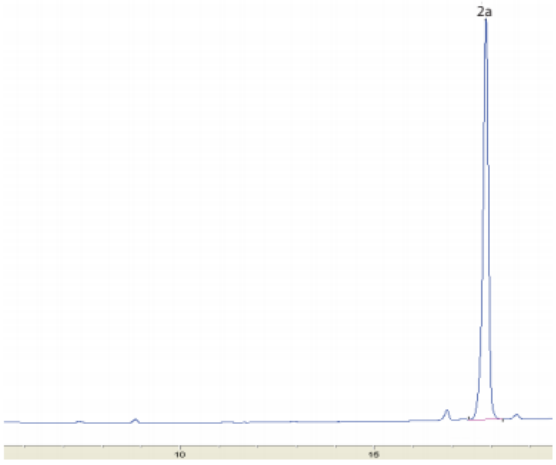
169. Zorzi, A.; Deyle, K.; Heinis, C. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24.

170. Zuckermann, R. N.; Kerr, J. M.; Moosf, W. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **1992**, *114* (26), 10646.

CHAPTER 2 SUPPLEMENTAL MATERIAL

All of the following figures have been reproduced with permission from: Elashal, H. E.; Raj, M. *Chem. Commun.* **2016,** *52*, 6304.





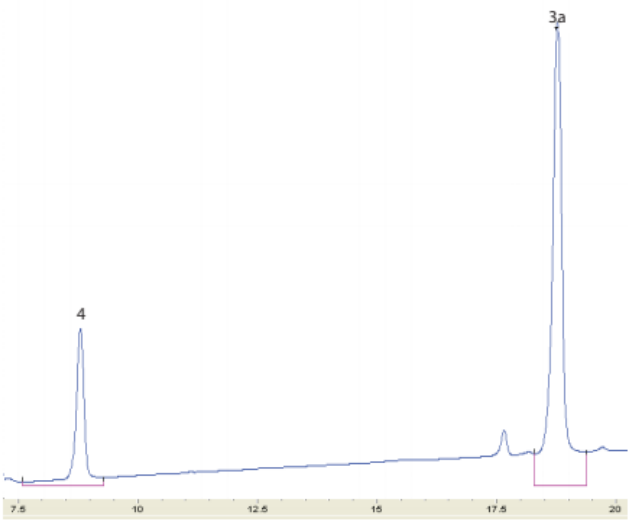


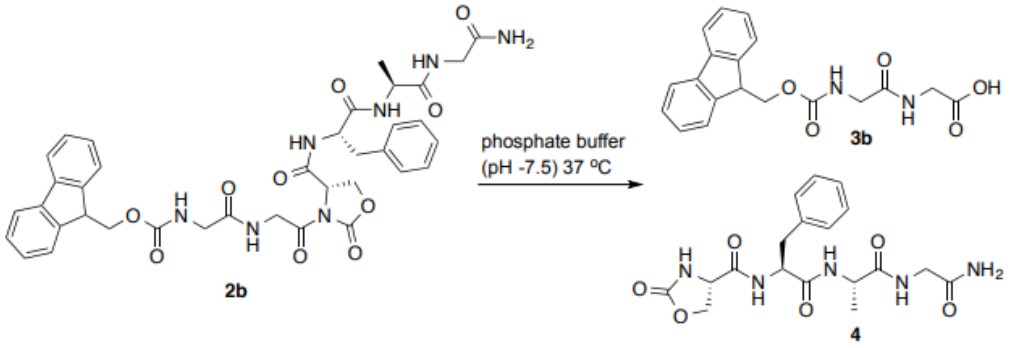
Figure A1 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Ala-Ser-Phe-Ala-Gly-NH2 (1a). LCMS: m/z 730.20 (calcd [M+H]+ = 730.01), Purity: >95% (HPLC analysis at 254 nm). Retention time: 17.85 min.

Fmoc-Gly-Ala-Oxd-Phe-Ala-Gly-NH2 (2a). LCMS: m/z 756.40 (calcd [M+H]+ = 756.10), 778.5 (calcd [M+Na]+ = 778.0). Purity: >95% (HPLC analysis at 254 nm). Retention time: 17.9 min.

Fmoc-Gly-Ala-OH (3a). LCMS: m/z 369.20 (calcd [M+H]+ = 369.13). Purity: >95% (HPLC analysis at 254 nm). Retention time: 18.6 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.20 (calcd [M+H]+ = 406.0), 428.2 (calcd [M+Na]+ = 428.3). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.9 min.


phosphate buffer
(pH -7.5) 37 ºC
2b
3b
4

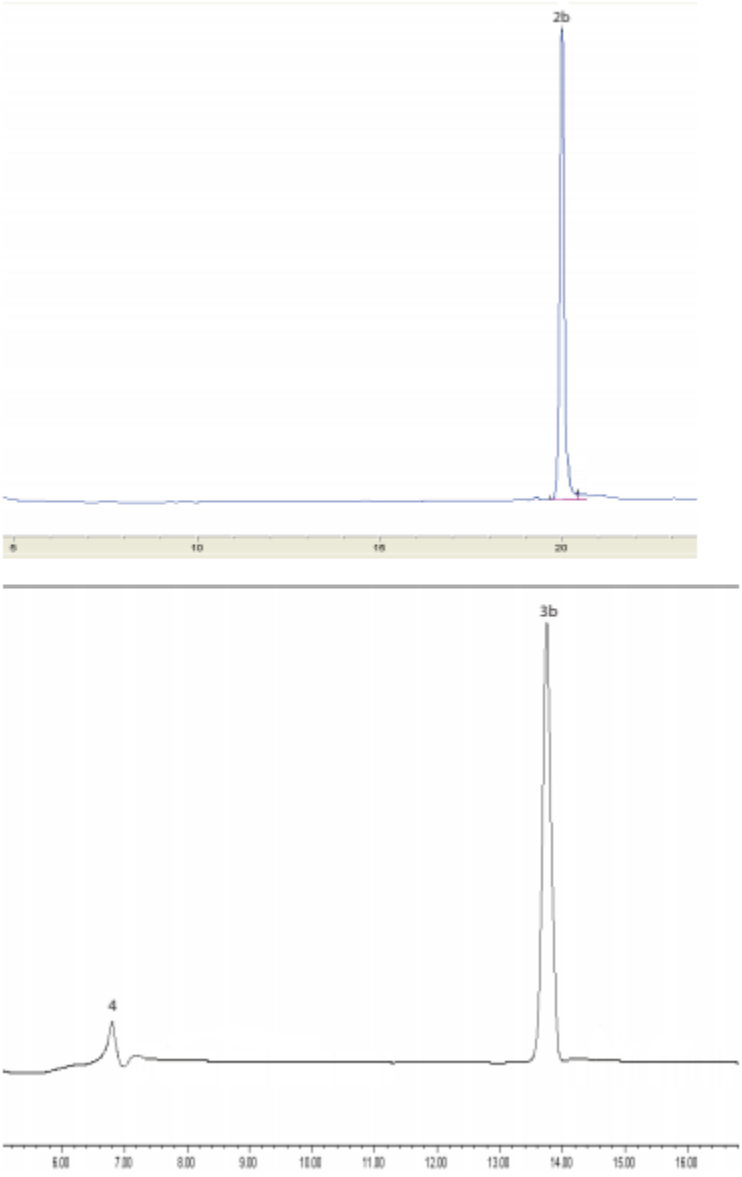
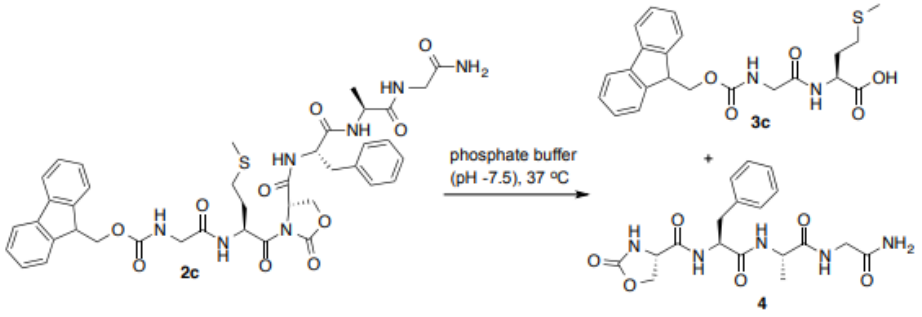

2b
3b
4

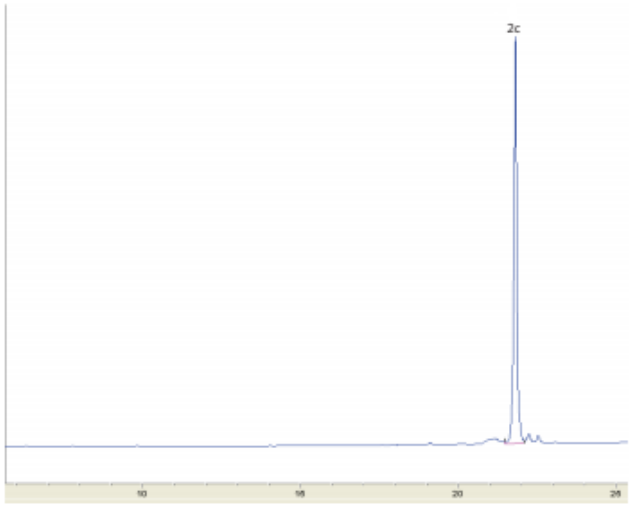
Figure A2 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Gly-Oxd-Phe-Ala-Gly-NH2 (2b). LCMS: m/z 742.40 (calcd [M+H]+ = 742.50), 371.9 (calcd [(M+2/)2] + = 371.75). Purity: >95% (HPLC analysis at 254 nm). Retention time: 20.1 min.

Fmoc-Gly-Gly-OH (3b). LCMS : m/z 355.20 (calcd [M+H]+ = 355.1). Purity: >95% (HPLC analysis at 254 nm). Retention time: 13.8 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 6.9 min.


phosphate buffer
(pH -7.5), 37 ºC
2c
3c
4


2c

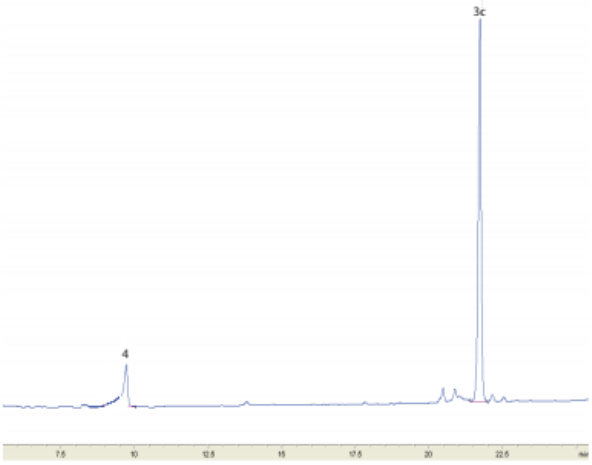

3c
4

Figure A3 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Met-Oxd-Phe-Ala-Gly-NH2 (2c). LCMS: m/z 816.0 (calcd [M+H]+ = 816.20), 408.2 (calcd [(M+2/)2] + = 408.8). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.8 min.

Fmoc-Gly-Met-OH (3c). LCMS : m/z 429.60 (calcd [M+H]+ = 429.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.6 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.1 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.8 min.

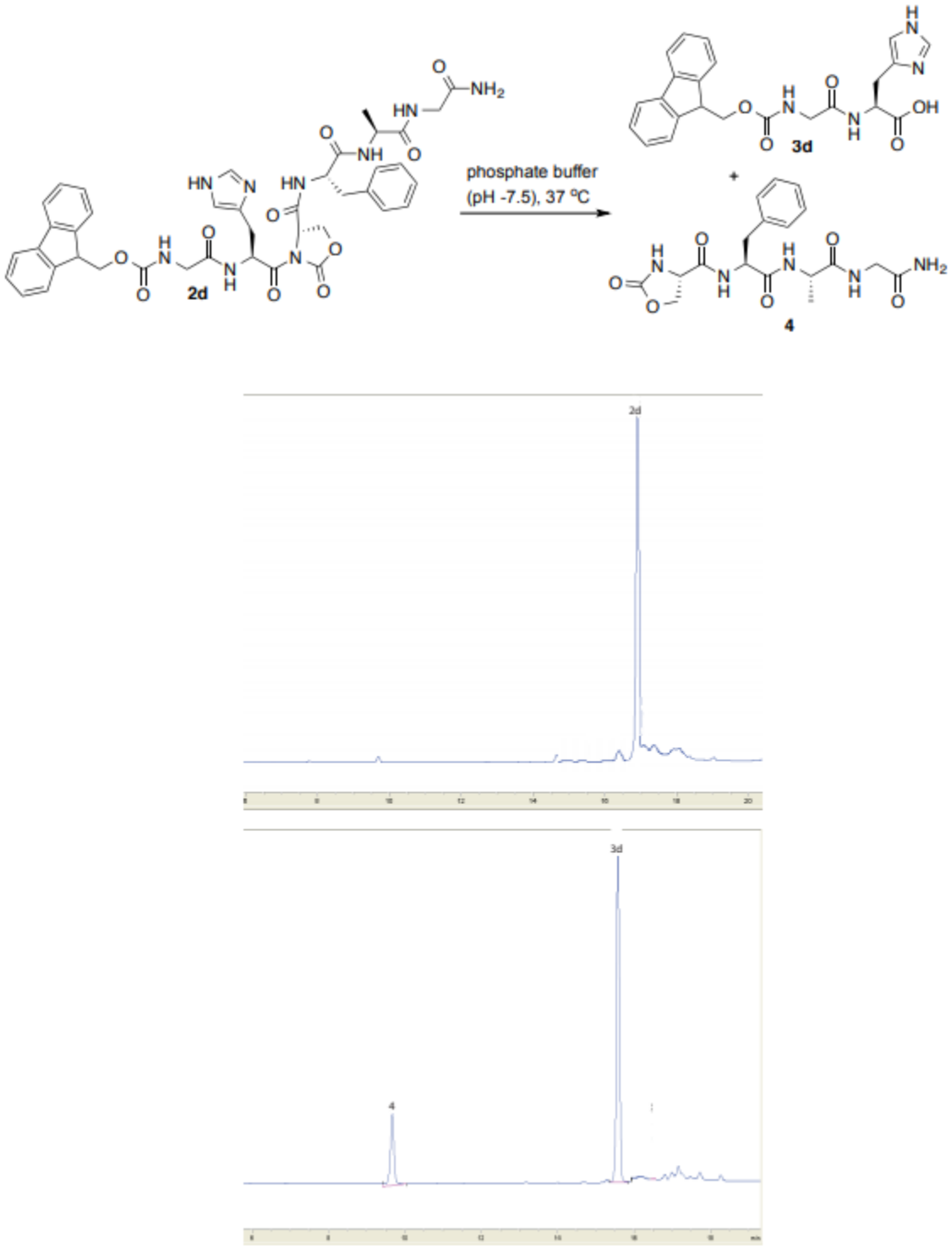
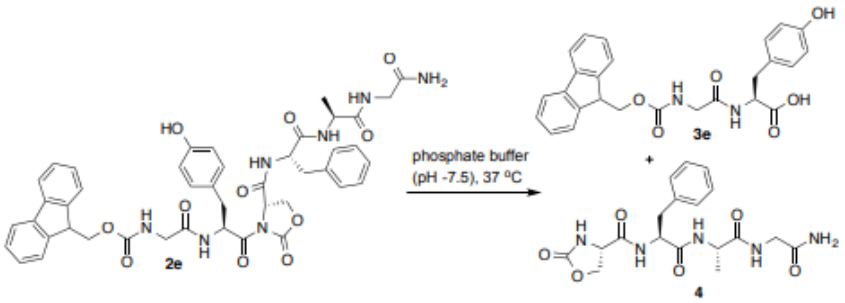


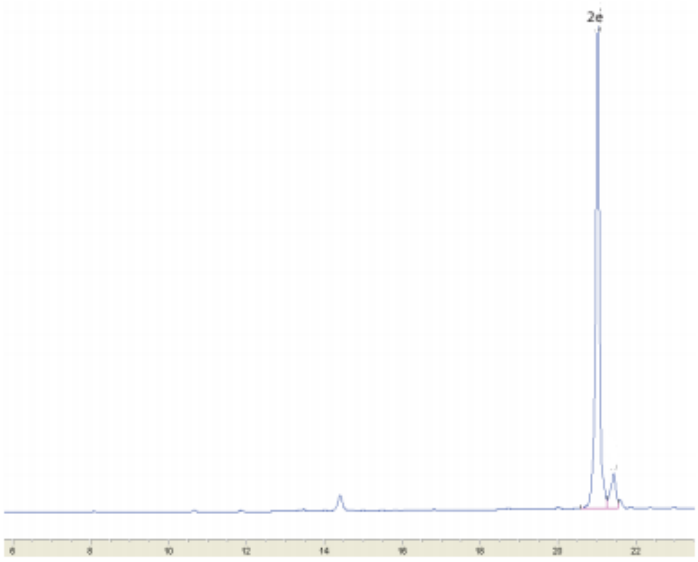
Figure A4 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-His-Oxd-Phe-Ala-Gly-NH2 (2d). LCMS: m/z 822.10 (calcd [M+H]+ = 822.60), 411.2 (calcd [(M+2/)2] + = 411.79). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.9 min.

Fmoc-Gly-His-OH (3d). LCMS : m/z 435.0 (calcd [M+H]+ = 435.19). Purity: >95% (HPLC analysis at 254 nm). Retention time: 15.8 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.3 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.8 min.


phosphate buffer
(pH ~7.5), 37 °C
2e
3e
4


2e

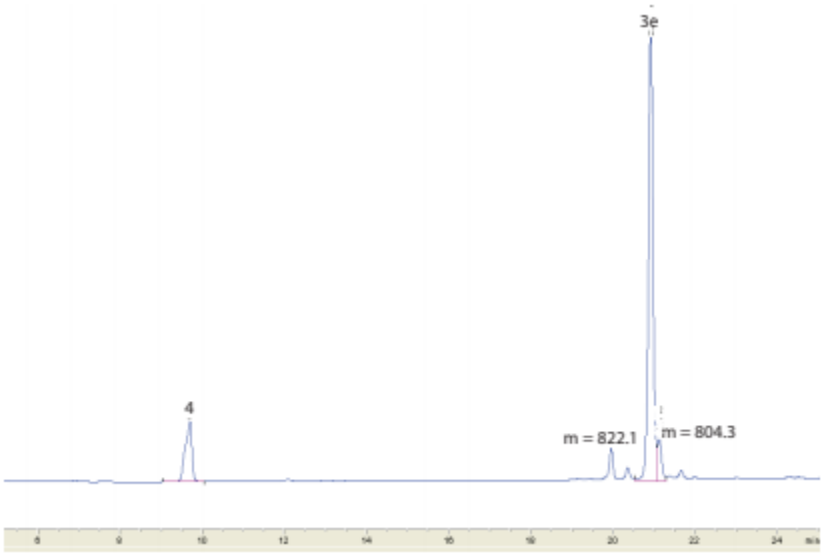
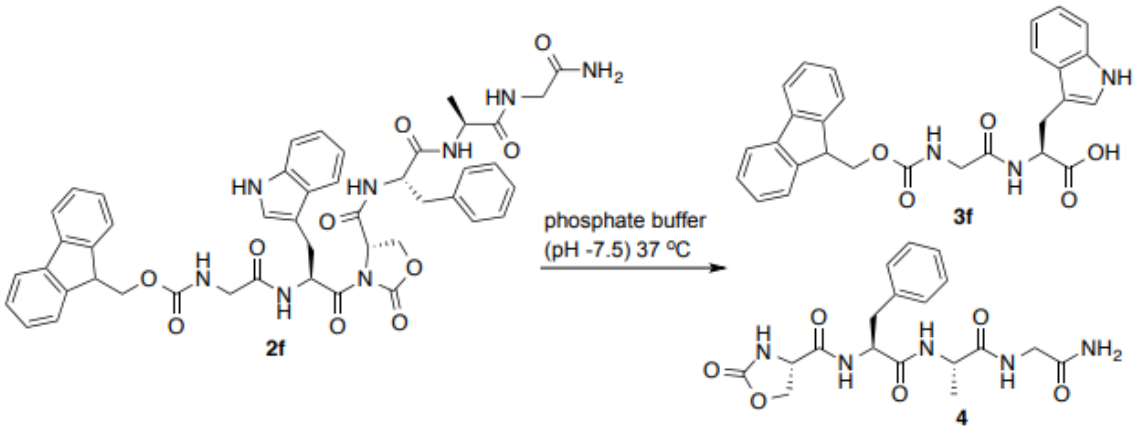

3e
4
m = 822.1
m = 804.3

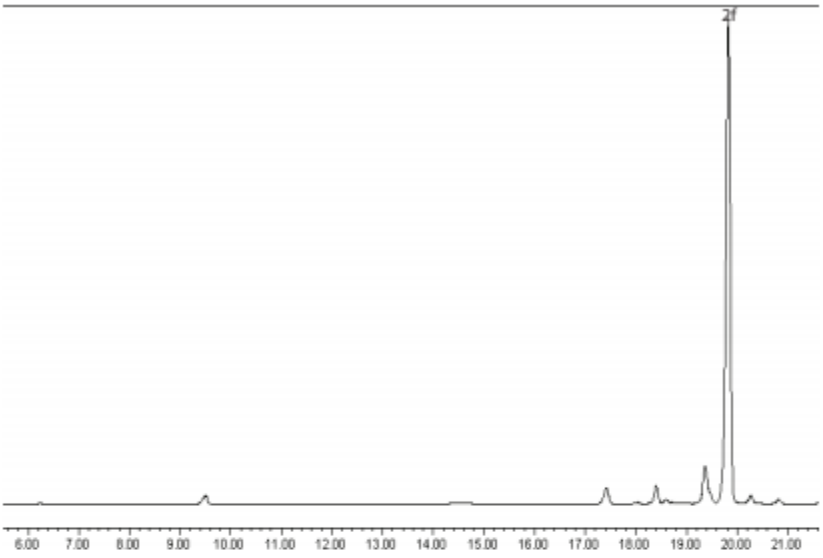
Figure A5 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Tyr-Oxd-Phe-Ala-Gly-NH2 (2e). LCMS: m/z 848.20 (calcd [M+H]+ = 848.62), 424.1 (calcd [(M+2/)2] + = 424.81). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.1 min.

Fmoc-Gly-Tyr-OH (3e). LCMS : m/z 461.1 (calcd [M+H]+ = 461.22). Purity: >95% (HPLC analysis at 254 nm). Retention time: 20.9 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.8 min.


phosphate buffer
(pH -7.5) 37 °C
2f
3f
4


2f
6.00 7.00 8.00 9.00 10.00 11.00 12.00 13.00 14.00 15.00 16.00 17.00 18.00 19.00 20.00 21.00

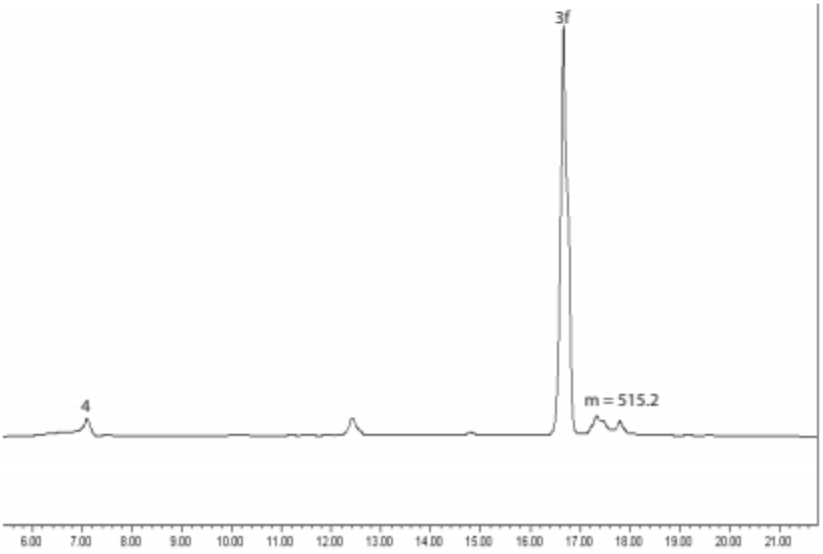
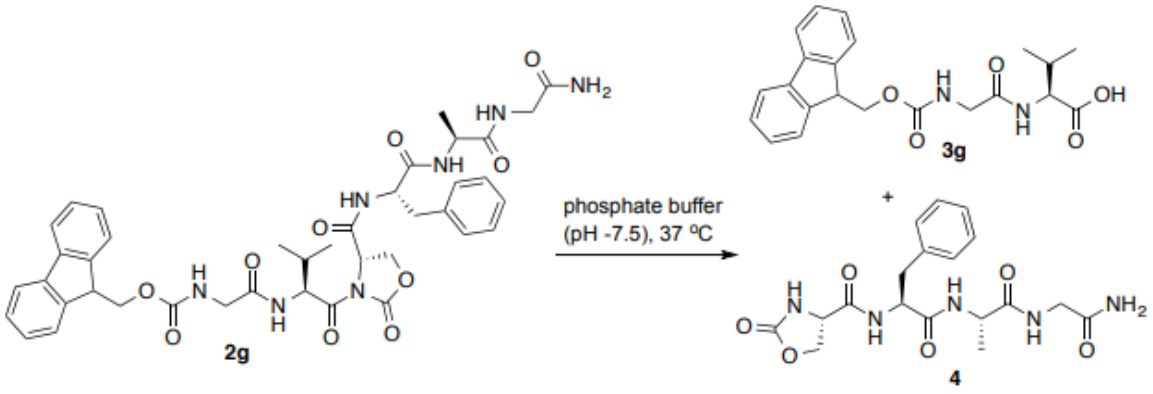

3f
4
m = 515.2
6.00 7.00 8.00 9.00 10.00 11.00 12.00 13.00 14.00 15.00 16.00 17.00 18.00 19.00 20.00 21.00

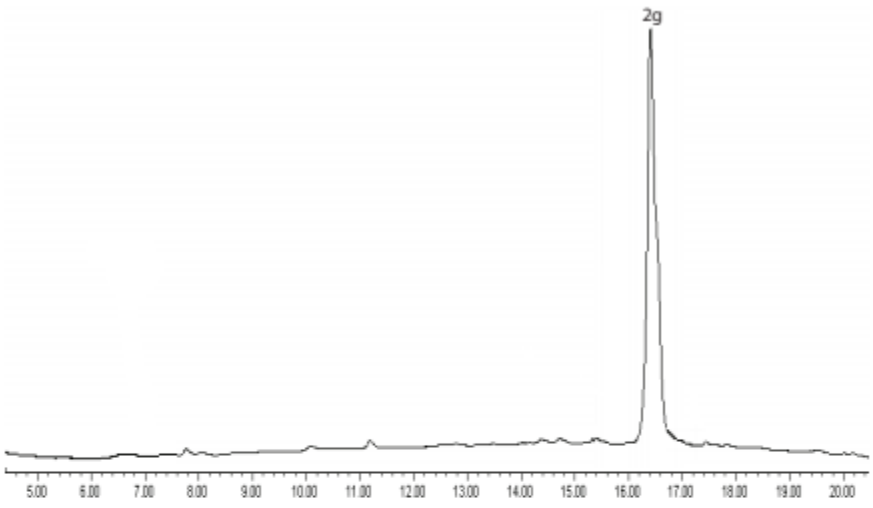
Figure A6 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Trp-Oxd-Phe-Ala-Gly-NH2 (2f). LCMS: m/z 871.20 (calcd [M+H]+ = 871.6), 436.0 (calcd [(M+2/)2] + = 436.33). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.8 min.

Fmoc-Gly-Trp-OH (3f). LCMS : m/z 484.1 (calcd [M+H]+ = 484.26). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.7 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 6.8 min.


phosphate buffer
(pH -7.5), 37 °C
2g
3g
4


2g
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
17.00
18.00
19.00
20.00

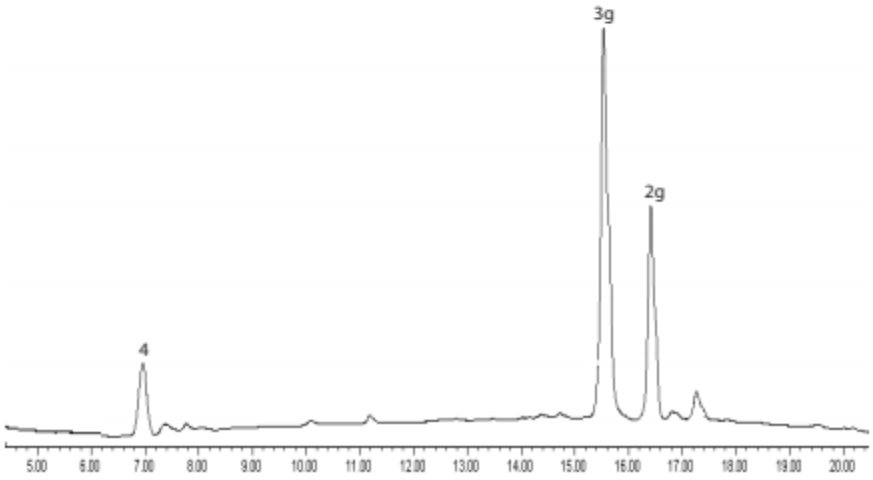
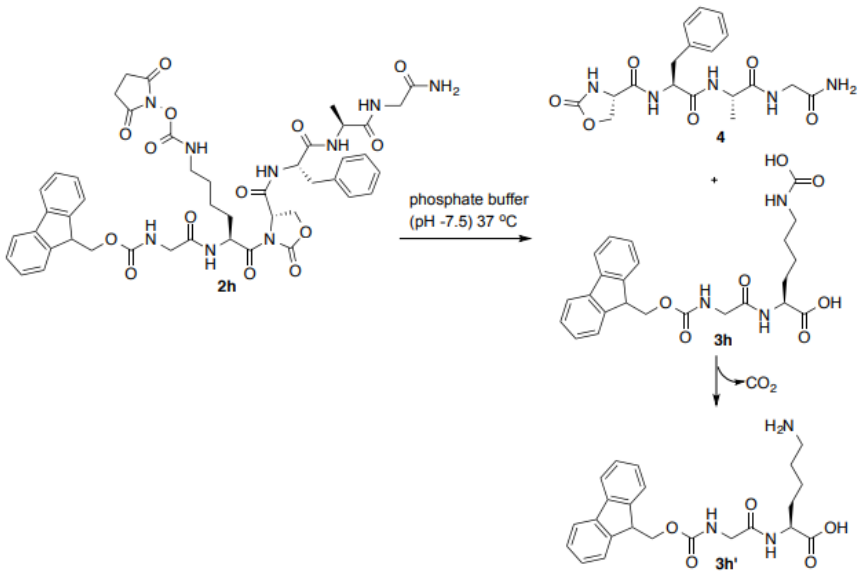

3g
2g
4
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
17.00
18.00
19.00
20.00

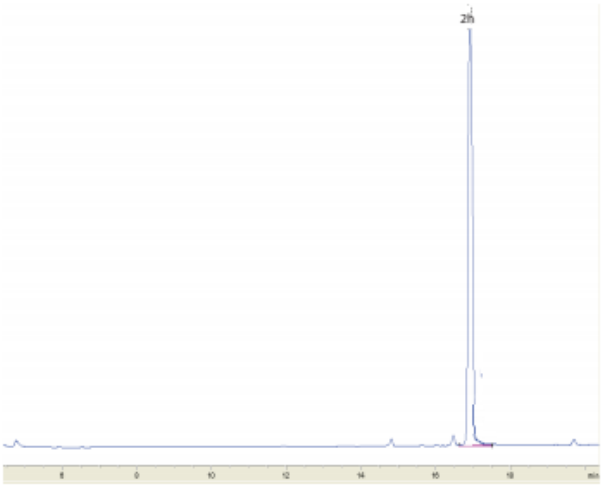
Figure A7 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

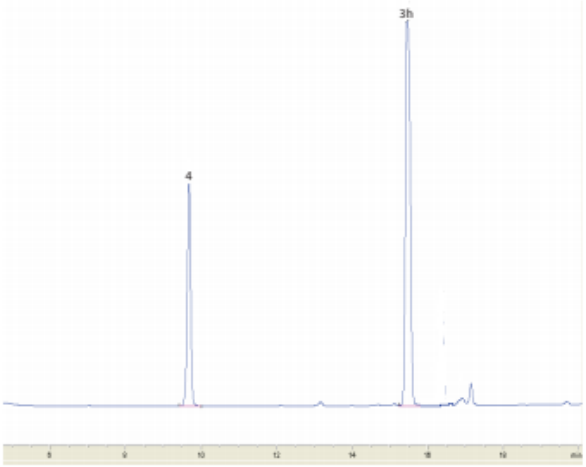
Fmoc-Gly-Val-Oxd-Phe-Ala-Gly-NH2 (2g). LCMS: m/z 784.60 (calcd [M+H]+ = 784.58), 392.5 (calcd [(M+2/)2] + = 392.79). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.6 min.

Fmoc-Gly-Val-OH (3g). LCMS : m/z 397.1 (calcd [M+H]+ = 397.18). Purity: >95% (HPLC analysis at 254 nm). Retention time: 15.6 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 6.8 min.


phosphate buffer
(pH -7.5) 37 ºC
2h
4
3h
3h'
CO2


2h


3h
4

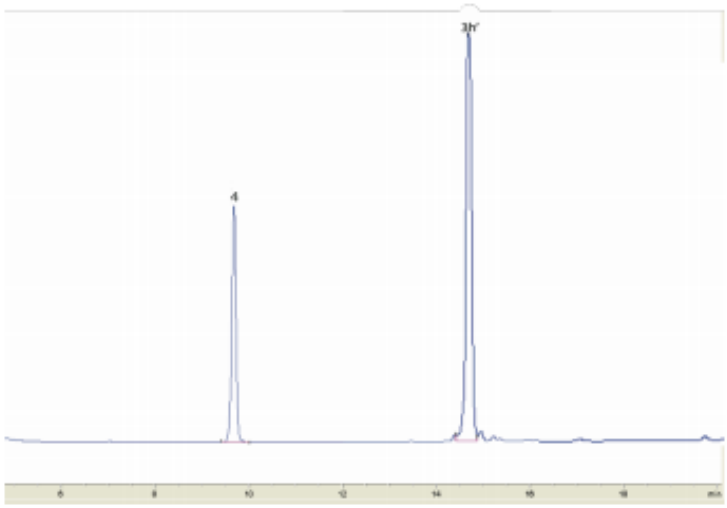


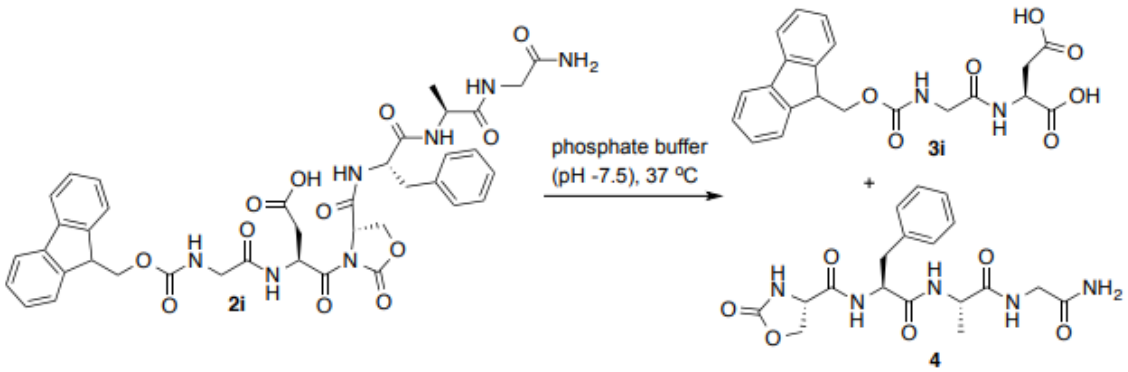
Figure A8 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

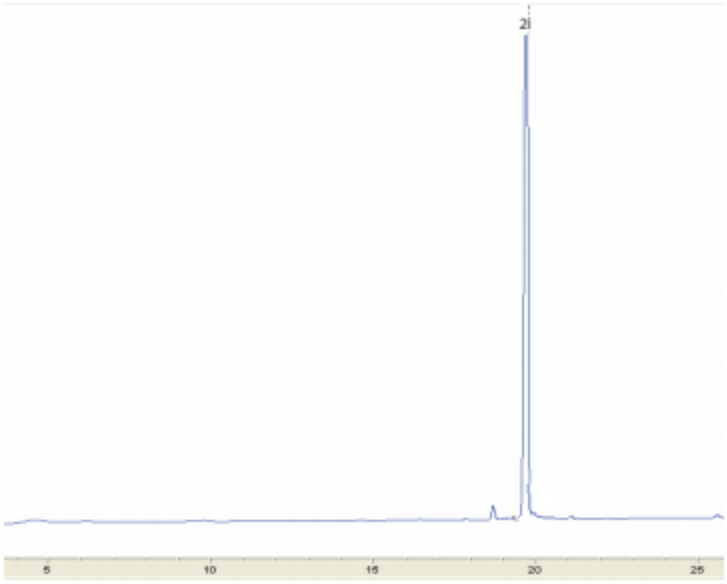
Fmoc-Gly-(N-hydroxysuccinimide-Lys)-Oxd-Phe-Ala-Gly-NH2 (2h). LCMS: m/z 928.10 (calcd [M+H]+ = 928.62), 464.2 (calcd [(M+2/)2] + = 464.81). Purity: >95% (HPLC analysis at 254 nm). Retention time: 17.1 min.

Fmoc-Gly-N-carboxy-Lys-OH (3h). LCMS : m/z 470.1 (calcd [M+H]+ = 470.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 15.6 min.

Fmoc-Gly-Lys-OH (3h'). LCMS : m/z 426.1 (calcd [M+H]+ = 426.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 14.7 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.8 min.


phosphate buffer
(pH ~7.5), 37 °C
2i
3i
4


2i

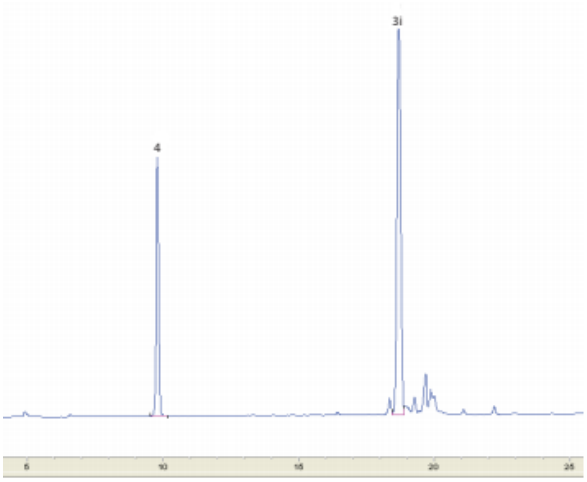
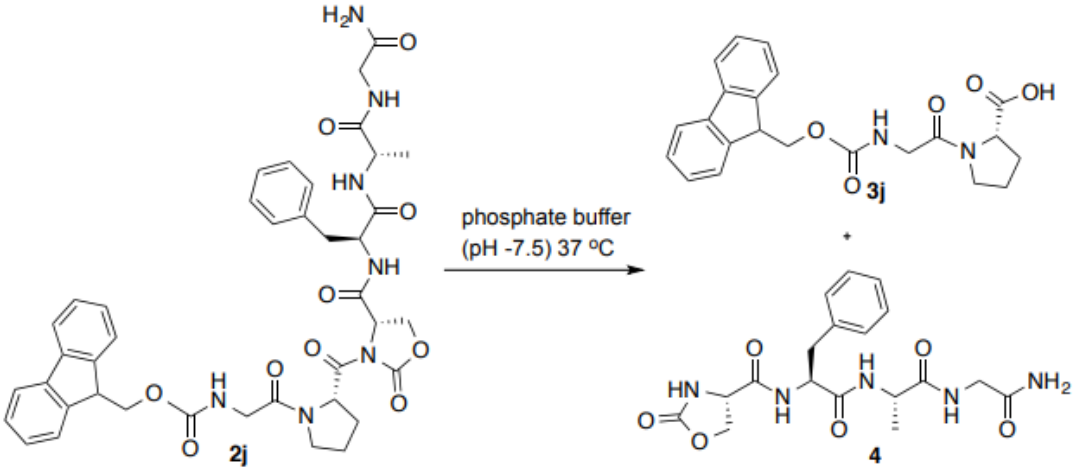

3i
4

Figure A9 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Asp(cyc)-Oxd-Phe-Ala-Gly-NH2 (2i). LCMS: m/z 782.10 (calcd [M+H]+ = 782.53), 391.1 (calcd [(M+2/)2] + = 391.76). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.8 min.

Fmoc-Gly-Asp-OH (3i). LCMS : m/z 413.1 (calcd [M+H]+ = 413.14). Purity: >95% (HPLC analysis at 254 nm). Retention time: 18.8 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.8 min.



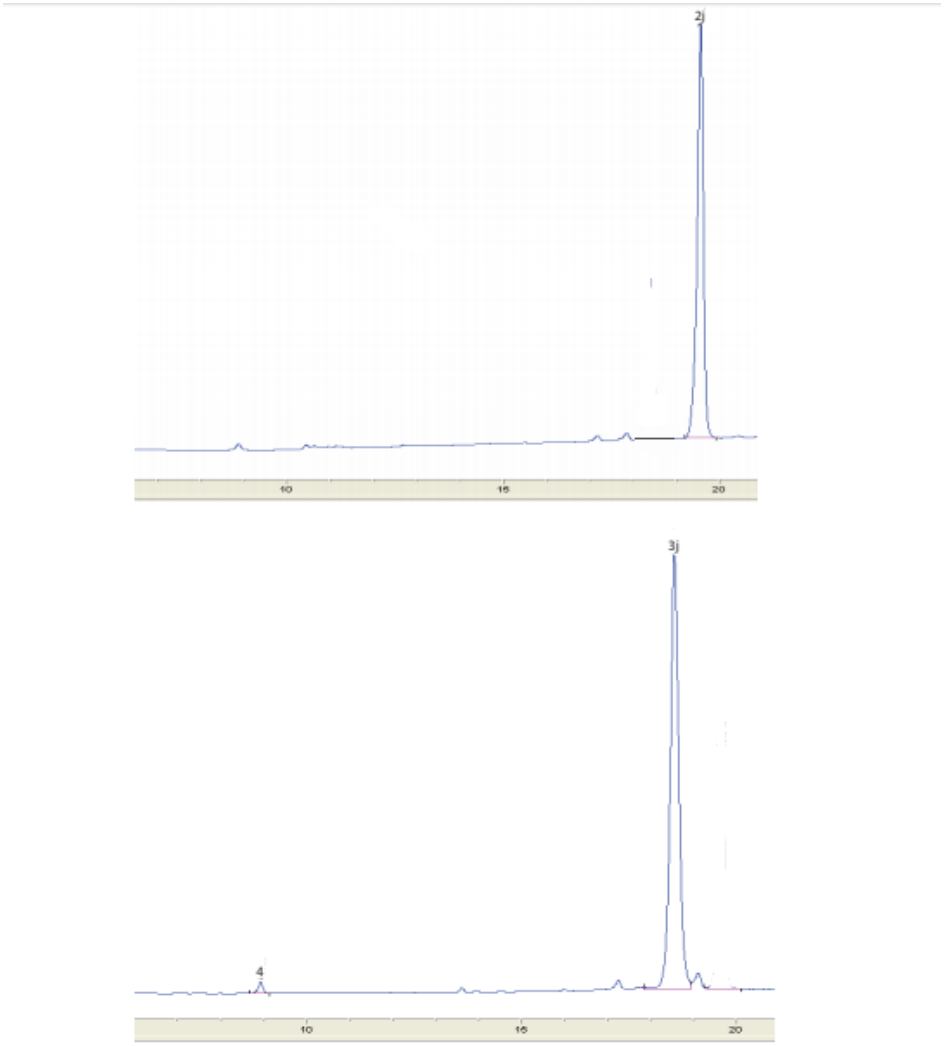
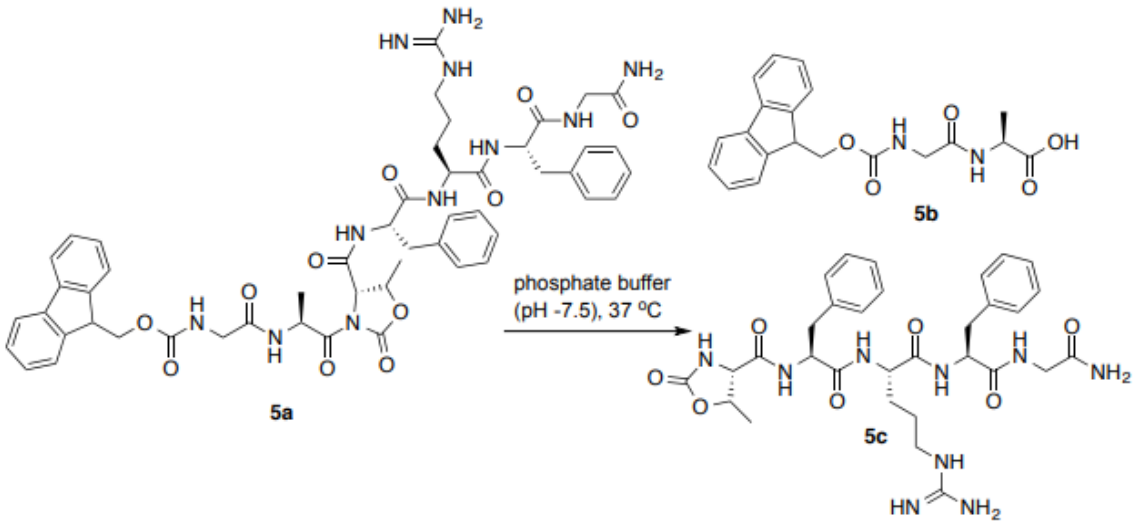


Figure A10 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Pro-Oxd-Phe-Ala-Gly-NH2 (2j). LCMS: m/z 782.20 (calcd [M+H]+ = 782.56), 391.5 (calcd [(M+2/)2] + = 391.78). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.55 min.

Fmoc-Gly-Pro-OH (3j). LCMS : m/z 395.1 (calcd [M+H]+ = 395.16). Purity: >95% (HPLC analysis at 254 nm). Retention time: 18.54 min.

Oxd-Phe-Ala-Gly-NH2 (4). LCMS: m/z 406.0 (calcd [M+H]+ = 406.1), 428.0 (calcd [M+Na]+ = 428.2). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.9 min.


phosphate buffer
(pH -7.5), 37 °C
5a
5b
5c

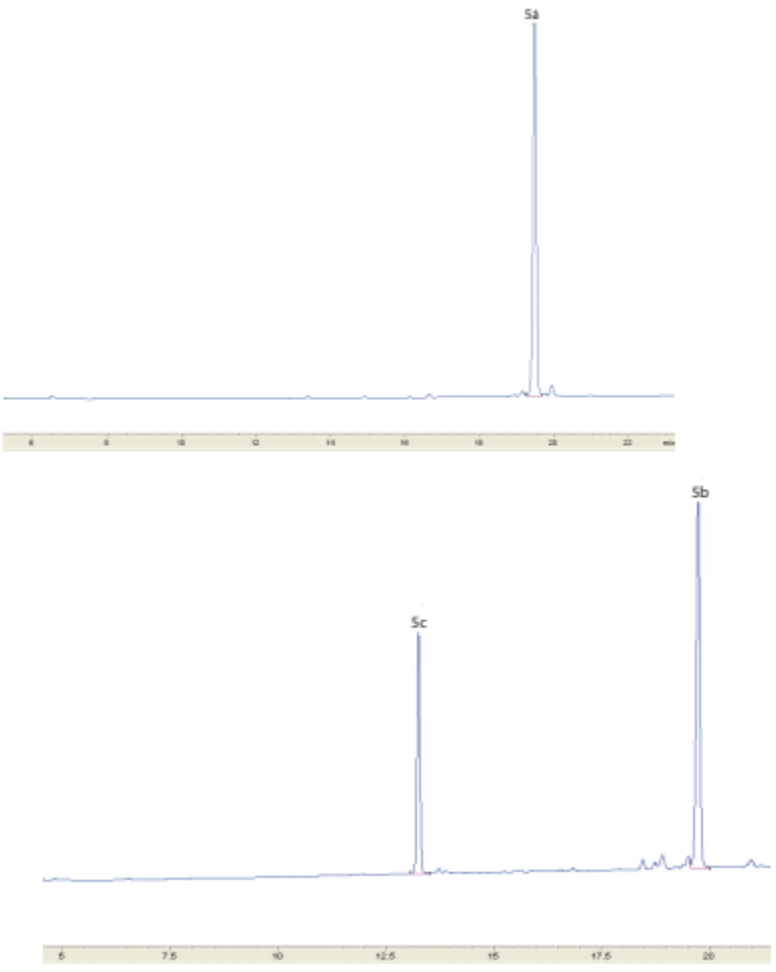
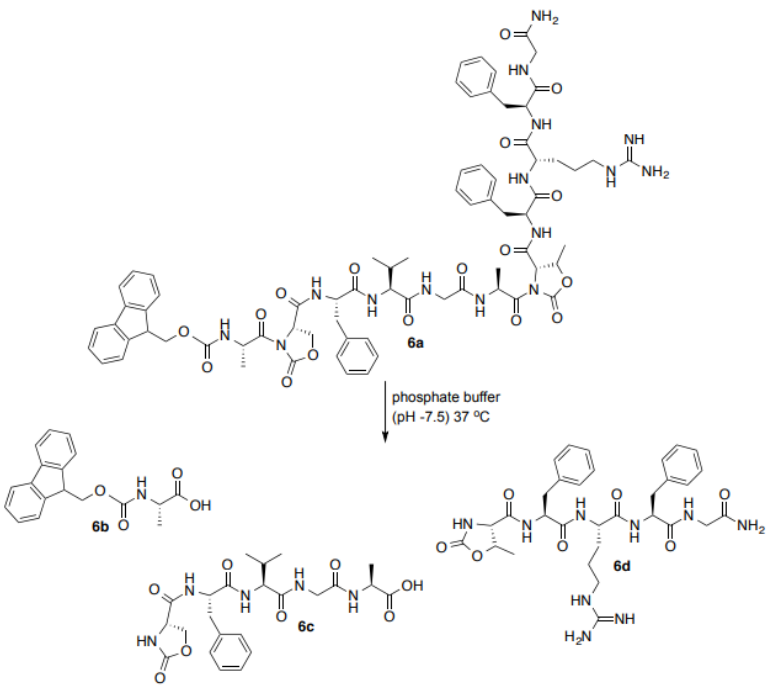

5a
5b
5c

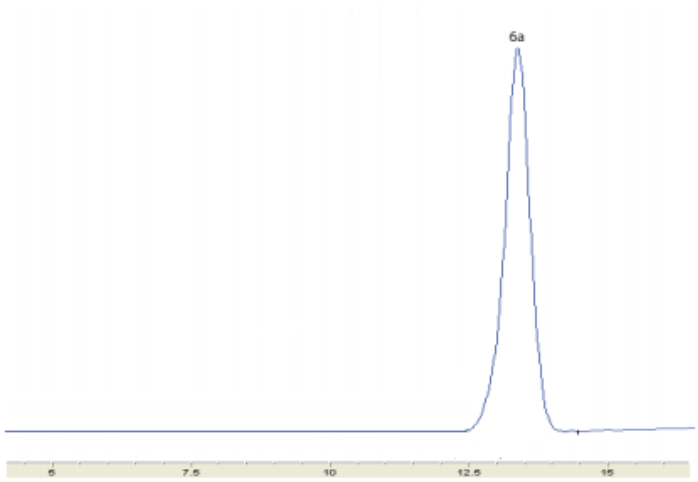
Figure A11 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Gly-Ala-Oxd-Phe-Arg-Phe-Gly-NH2 (5a). LCMS: m/z 1003.0 (calcd [M+H]+ = 1003.2), 502.0 (calcd [(M+2/)2] + = 501.91). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.6 min.

Fmoc-Gly-Ala-OH (5b). LCMS : m/z 369.2 (calcd [M+H]+ = 369.13). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.89 min.

Oxd-Phe-Arg-Phe-Gly-NH2 (5c). LCMS: m/z 652.3 (calcd [M+H]+ = 652.71), 674.2 (calcd [M+Na]+ = 674.71). Purity: >95% (HPLC analysis at 254 nm). Retention time: 13.3 min.


6a
phosphate buffer
(pH -7.5) 37 °C
6b
6c
6d


6a
5
7.5
10
12.5
15

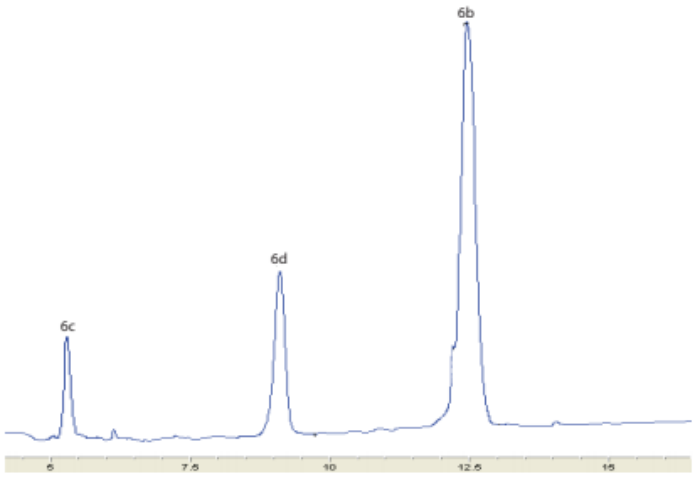

6b
6d
6c
5
7.5
10
12.5
15

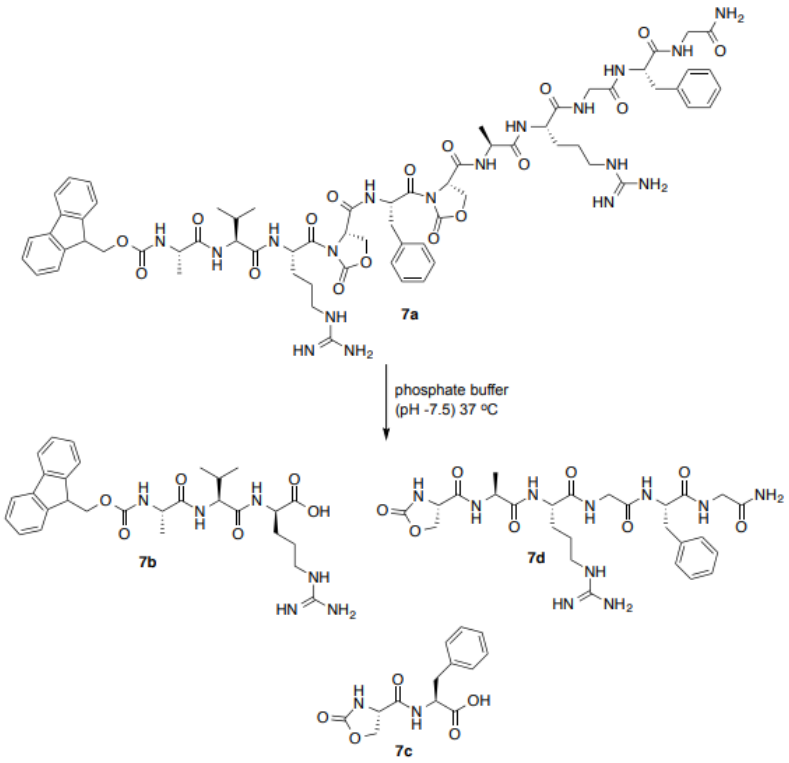
Figure A12 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Ala-Oxd-Phe-Val-Gly-Ala-(Oxd)-Phe-Arg-Phe-Gly-NH2 (6a). LCMS: m/z 1433.2 (calcd [M+H]+ = 1433.3), 717.1 (calcd [(M+2/)2] + = 717.15). Purity: >95% (HPLC analysis at 254 nm). Retention time: 13.04 min.

Fmoc-Ala-OH (6b). LCMS : m/z 312.0 (calcd [M+H]+ = 312.07). Purity: >95% (HPLC analysis at 254 nm). Retention time: 12.9 min.

Oxd-Phe-Val-Gly-Ala-OH (6c). LCMS: m/z 506.3 (calcd [M+H]+ = 506.51), 528.2 (calcd [M+Na]+ = 528.51). Purity: >95% (HPLC analysis at 254 nm). Retention time: 5.06 min.

Oxd-Phe-Arg-Phe-Gly-NH2 (6d). LCMS: m/z 653.1 (calcd [M+H]+ = 653.71), 675.2 (calcd [M+Na]+ = 675.71). Purity: >95% (HPLC analysis at 254 nm). Retention time: 9.12 min.


7a
phosphate buffer
(pH -7.5) 37 ºC
7b
7d
7c

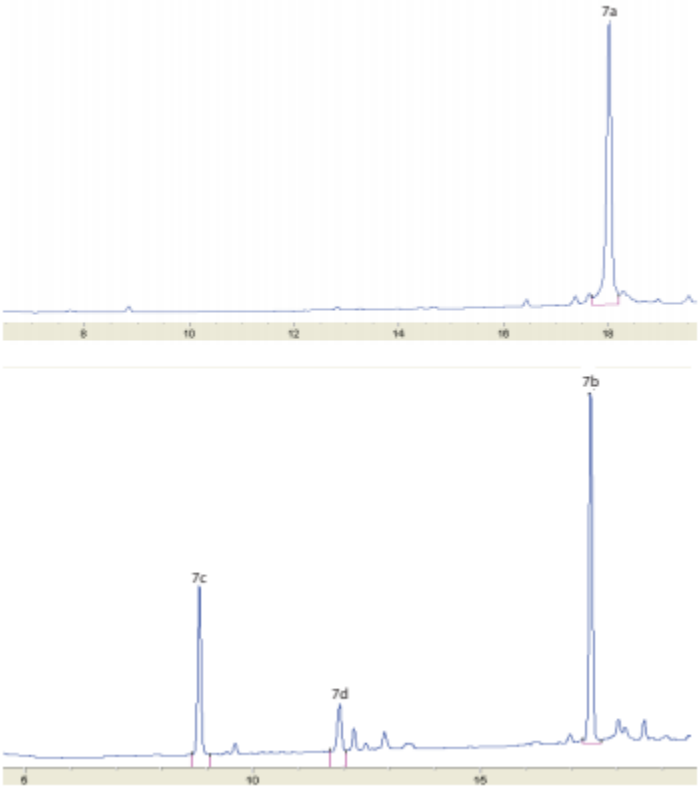

7a
8
10
12
14
16
18
7b
7c
7d
5
10
15

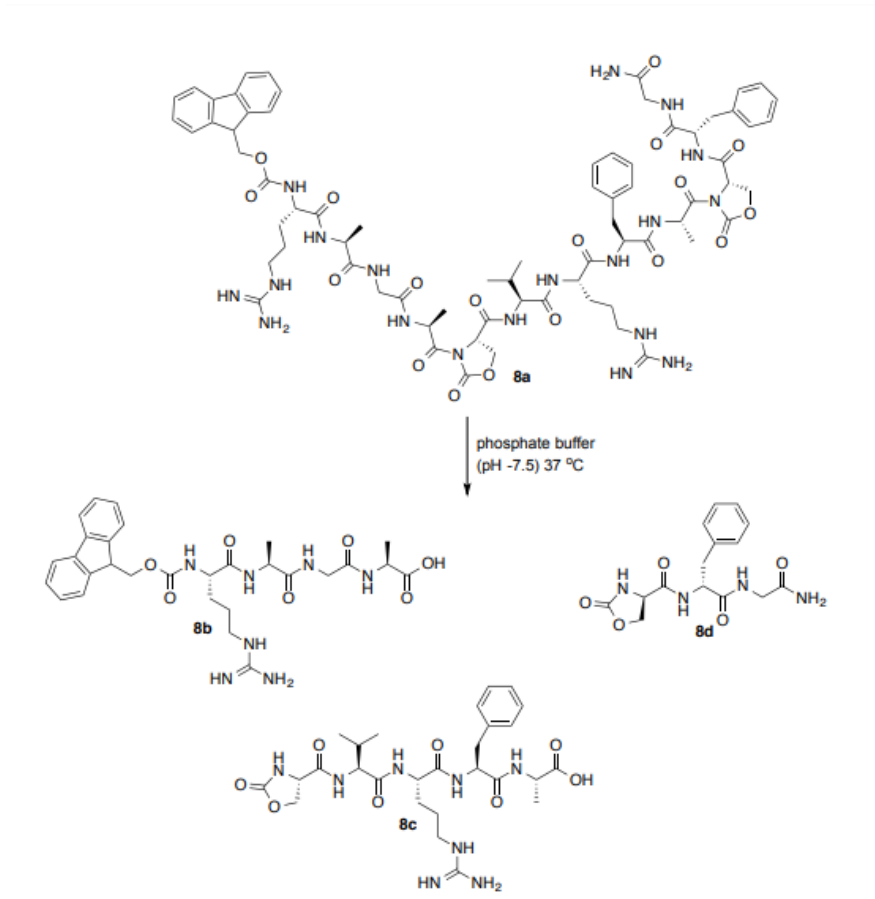
Figure A13 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Ala-Val-Arg-Oxd-Phe-Oxd-Ala-Arg-Gly-Phe-Gly-NH2 (7a). LCMS: m/z 1428.1 (calcd [M+H]+ = 1428.29), 714.3 (calcd [(M+2/)2] + = 714.64). Purity: >95% (HPLC analysis at 254 nm). Retention time: 18.2 min.

Fmoc-Ala-Val-Arg-OH (7b). LCMS : m/z 567.2 (calcd [M+H]+ = 567.39). Purity: >95% (HPLC analysis at 254 nm). Retention time: 17.3 min.

Oxd-Phe-OH (7c). LCMS: m/z 279.1 (calcd [M+H]+ = 279.25), 301.1 (calcd [M+Na]+ = 301.25). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.8 min.

Oxd-Ala-Arg-Gly-Phe-Gly-NH2 (7d). LCMS: m/z 619.5 (calcd [M+H]+ = 619.63), 641.4 (calcd [M+Na]+ = 641.63). Purity: >95% (HPLC analysis at 254 nm). Retention time: 11.87 min.


8a
phosphate buffer
(pH -7.5) 37 °C
8b
8d
8c

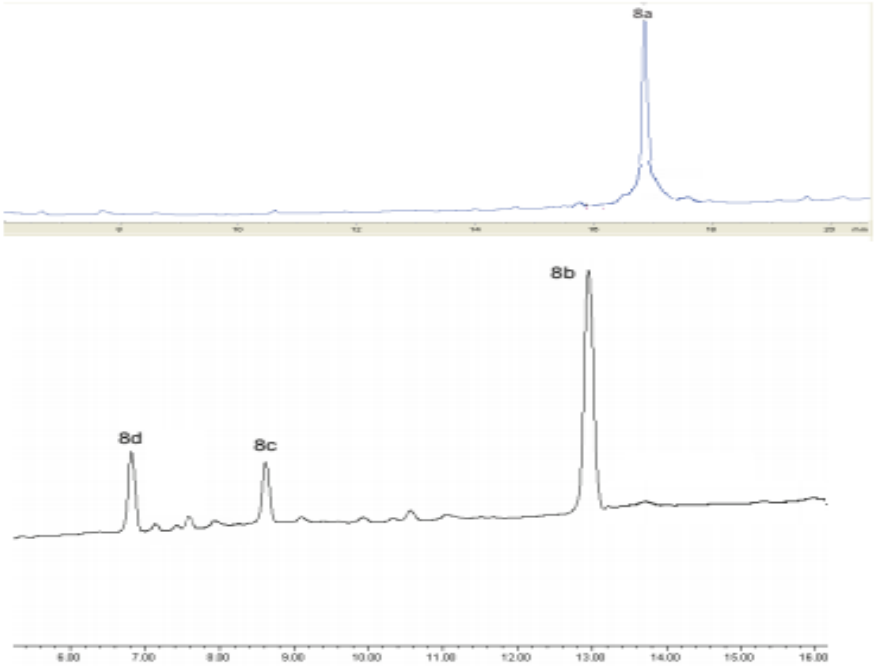

8a
8b
8d
8c
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00

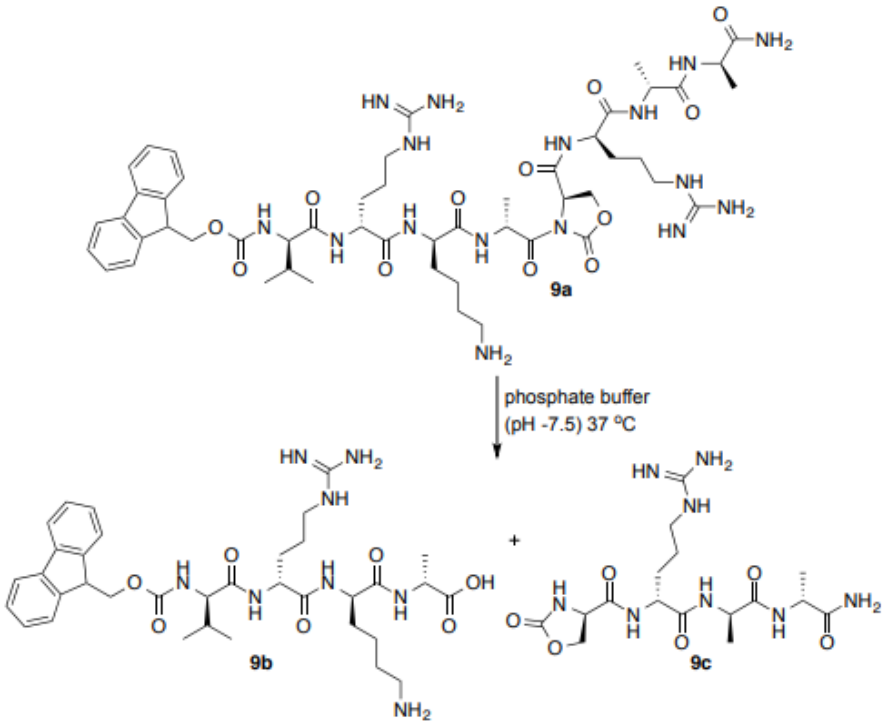
Figure A14 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Arg-Ala-Gly-Ala-Ser-Val-Arg-Phe-Ala-Ser-Phe-Gly-OH (8a). LCMS: m/z 1499.1 (calcd [M+H]+ = 1499.36), 750.0 (calcd [(M+2/)2] + = 750.18). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.8 min.

Fmoc-Arg-Ala-Gly-Ala-OH (8b). LCMS : m/z 596.2 (calcd [M+H]+ = 596.39). Purity: >95% (HPLC analysis at 254 nm). Retention time: 12.9 min.

Oxd-Val-Arg-Phe-Ala-OH (8c). LCMS: m/z 605.3 (calcd [M+H]+ = 605.65), 627.4 (calcd [M+Na]+ = 627.65). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.71 min.

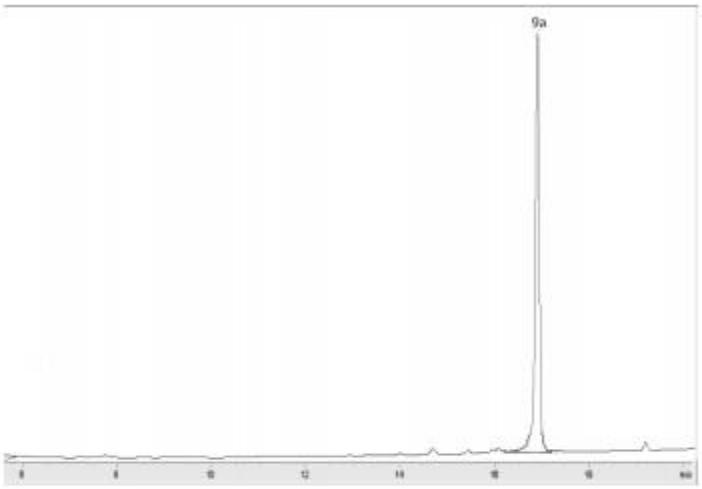
Oxd-Phe-Gly-OH (8d). LCMS: m/z 335.09 (calcd [M+H]+ = 335.32), 357.1 (calcd [M+Na]+ = 357.32). Purity: >95% (HPLC analysis at 254 nm). Retention time: 6.8 min.



9a

9b

9c



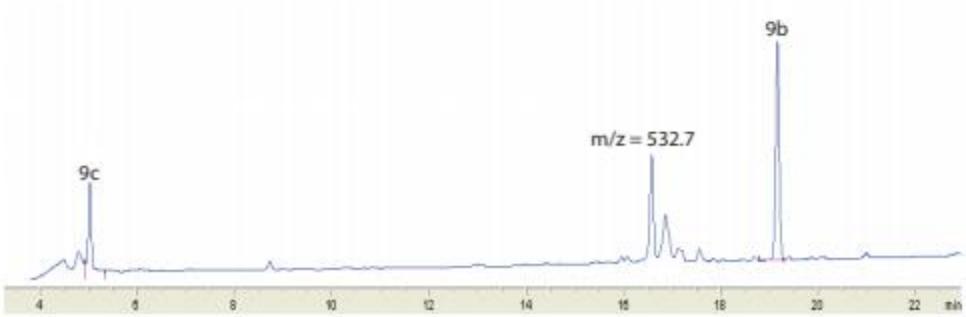
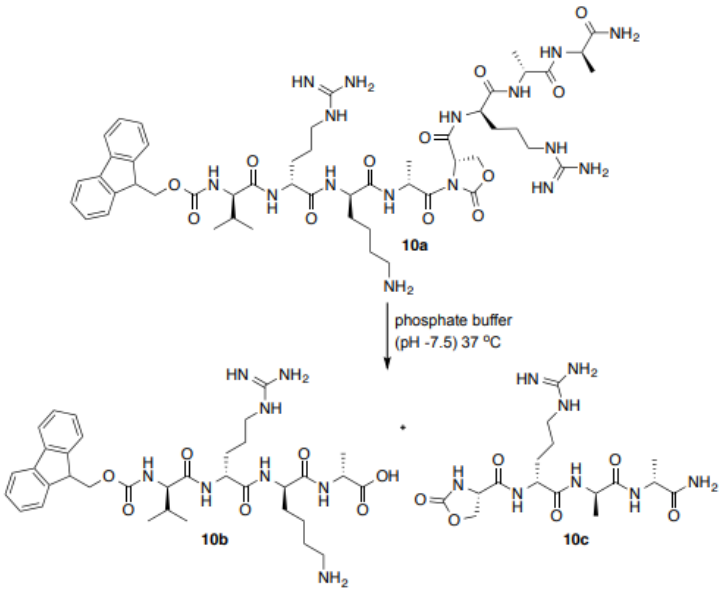


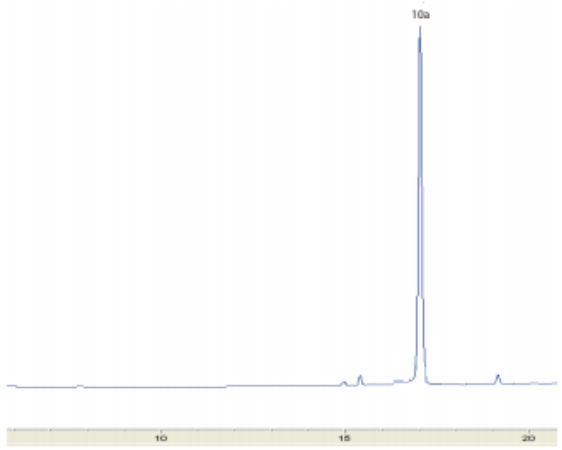
Figure A15 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-D-Val-D-Arg-D-Lys-D-Ala-D-Oxd-D-Arg-D-Ala- D-Ala-NH2 (9a). LCMS: m/z 1106.0 (calcd [M+H]+ = 1106.02), 553.1 (calcd [(M+2/)2] + = 553.5). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.9 min.

Fmoc-D-Val-D-Arg-D-Lys-D-Ala-OH (9b). LCMS : m/z 695.3 (calcd [M+H]+ = 695.55). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.2 min.

D-Oxd-D-Arg-D-Ala-D-Ala-NH2 (9c). LCMS: m/z 429.3 (calcd [M+H]+ = 429.43), 451.3 (calcd [M+Na]+ = 451.43), Purity: >95% (HPLC analysis at 254 nm). Retention time: 5.2 min.


10a
phosphate buffer
(pH -7.5) 37 °C
10b
10c


10a

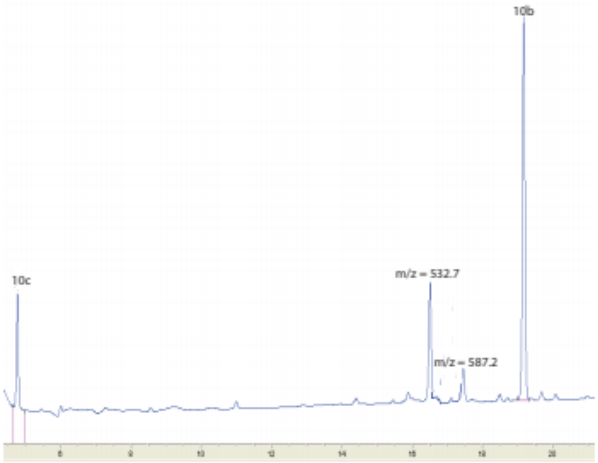
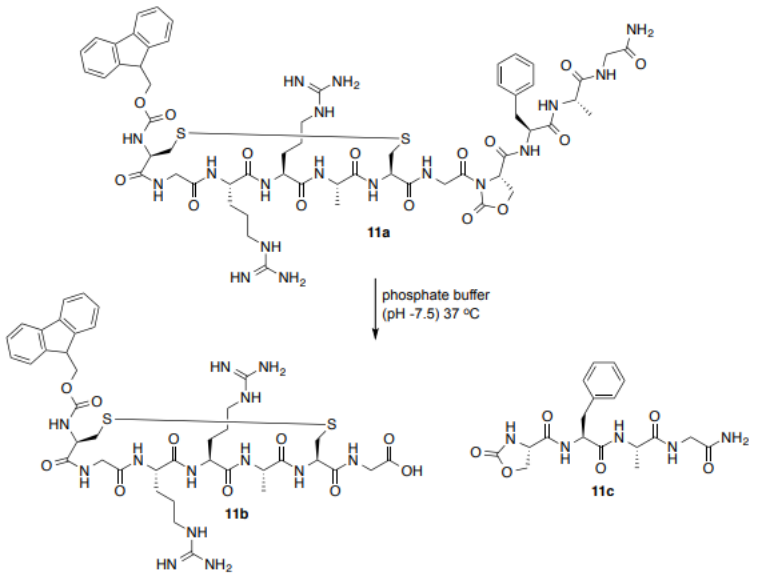

10b
10c
m/z = 532.7
m/z = 587.2

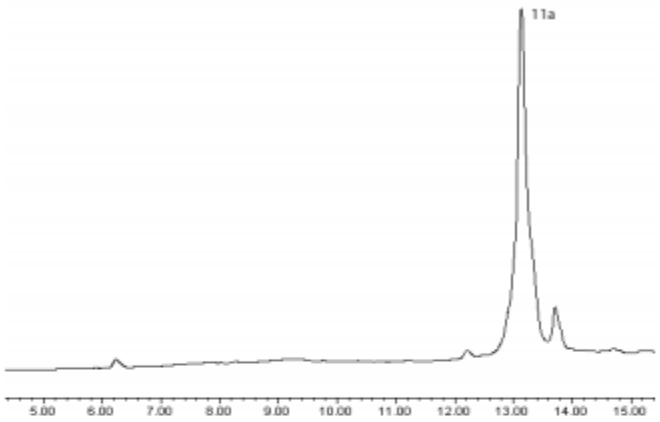
Figure A16 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-D-Val-D-Arg- D-Lys-D-Ala-L-Oxd-D-Arg-D-Ala- D-Ala-NH2 (10a). LCMS: m/z 1106.1 (calcd [M+H]+ = 1106.02), 553.0 (calcd [(M+2/)2] + = 553.5). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.89 min.

Fmoc-D-Val- D-Arg- D-Lys-D-Ala-OH (10b). LCMS : m/z 695.1 (calcd [M+H]+ = 695.55). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.18 min.

L-Oxd-D-Arg-D-Ala-D-Ala-NH2 (10c). LCMS: m/z 429.3 (calcd [M+H]+ = 429.43), 451.7 (calcd [M+Na]+ = 451.43.Purity: >95% (HPLC analysis at 254 nm). Retention time: 4.7 min.


11a
phosphate buffer
(pH -7.5) 37 °C
11b
11c


11a
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00

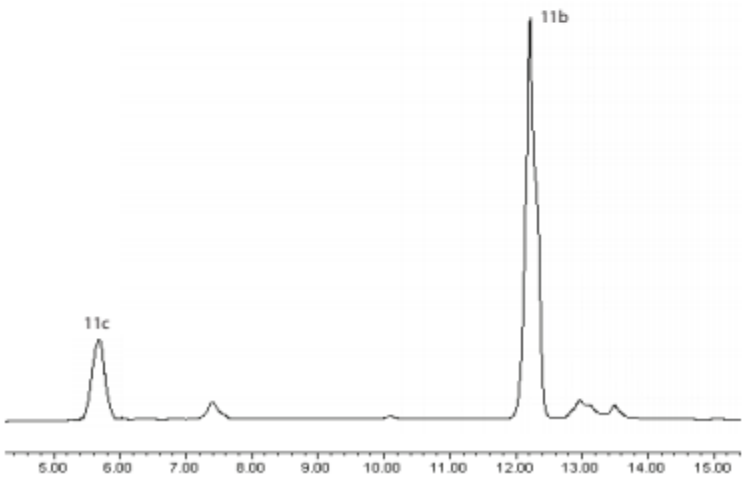

11b
11c
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00

Figure A17 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Cys-Gly-Arg-Arg-Ala-Cys-Gly-Oxd-Phe-Ala-Gly-NH2, disulfide bond (11a). LCMS: m/z 1330.1 (calcd [M+H]+ = 1330.24), 665.5 (calcd [(M+2/)2] + = 665.62). Purity: >95% (HPLC analysis at 254 nm). Retention time: 13.1 min.

Fmoc-Cys-Gly-Arg-Arg-Ala-Cys-Gly-OH, disulfide bond (11b). LCMS : m/z 942.5 (calcd [M+H]+ = 942.84), 471.7 (calcd [(M+2/)2] + = 471.92). Purity: >95% (HPLC analysis at 254 nm). Retention time: 12.2 min.

Oxd-Phe-Ala-Gly-NH2 (11c). LCMS: m/z 406.2 (calcd [M+H]+ = 406.39), Purity: >95% (HPLC analysis at 254 nm). Retention time: 5.8 min.

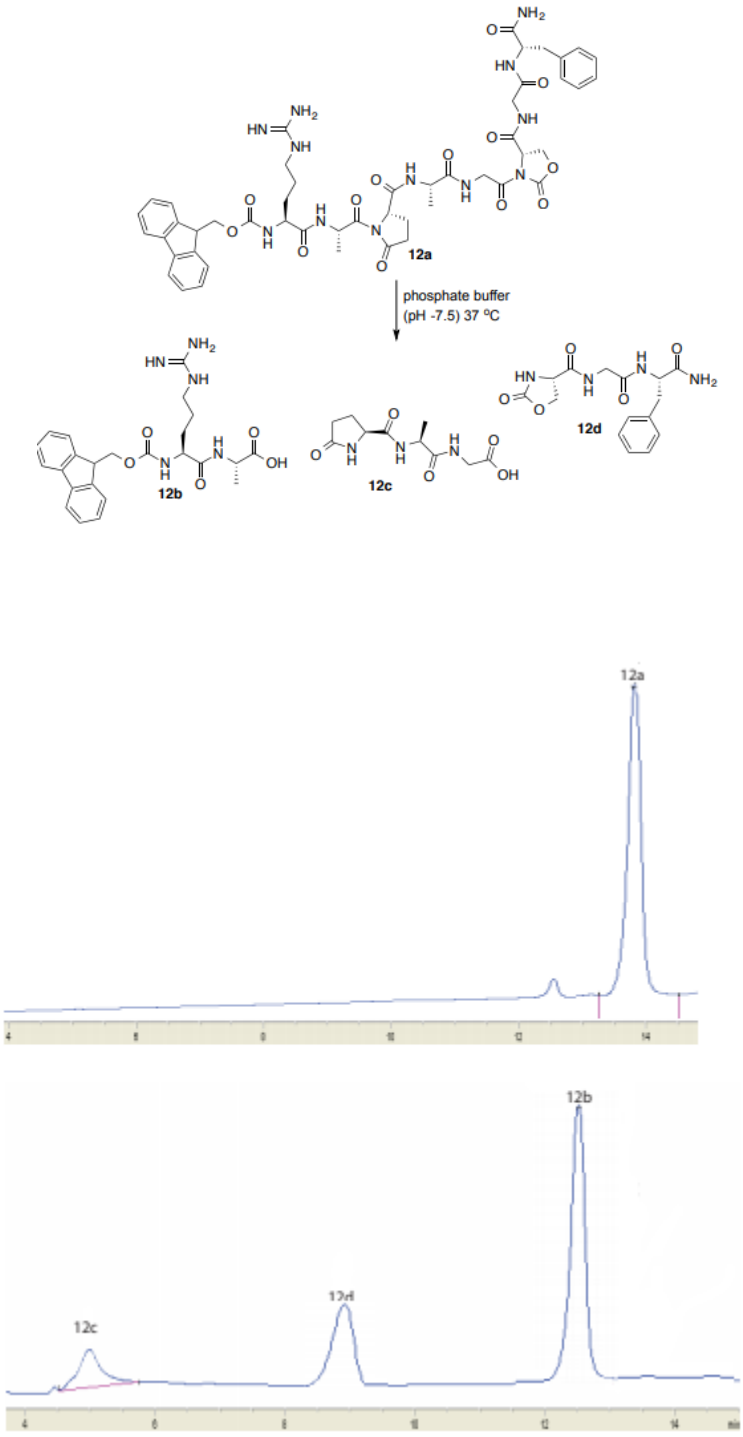


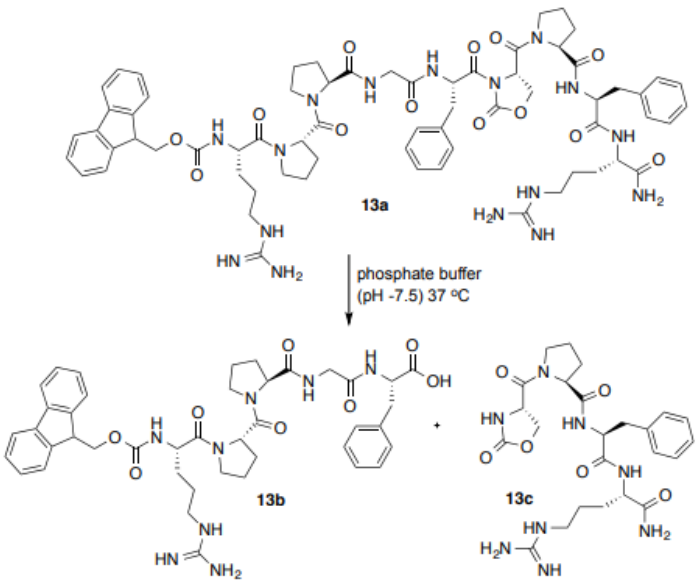
Figure A18 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

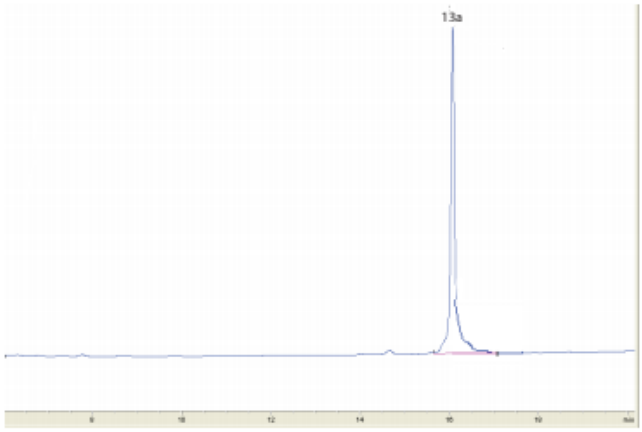
Fmoc-Arg-Ala-Pyr-Ala-Gly-Oxd-Gly-Phe-NH2 (12a). LCMS: m/z 1023.4 (calcd [M+H]+ = 1023.83), 512.1 (calcd [(M+2/)2] + = 512.41). Purity: >95% (HPLC analysis at 254 nm). Retention time: 13.9 min.

Fmoc-Arg-Ala-OH (12b). LCMS : m/z 468.09 (calcd [M+H]+ = 468.26). Purity: >95% (HPLC analysis at 254 nm). Retention time: 12.5 min.

Pyr-Ala-Gly-OH (12c). LCMS: m/z 258.14 (calcd [M+H]+ = 258.24), 280.0 (calcd [M+Na]+ = 280.24). Purity: >95% (HPLC analysis at 254 nm). Retention time: 4.98 min.

Oxd-Gly-Phe-NH2 (12d). LCMS: m/z 335.2 (calcd [M+H]+ = 335.32), 357.4 (calcd [M+Na]+ = 357.32). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.98 min.


13a
phosphate buffer
(pH -7.5) 37 °C
13b
13c


13a

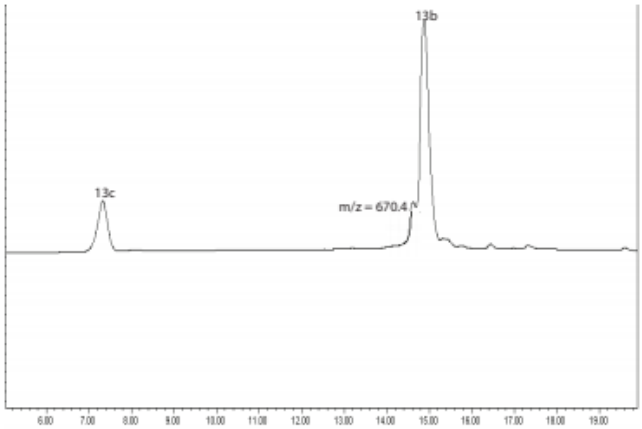
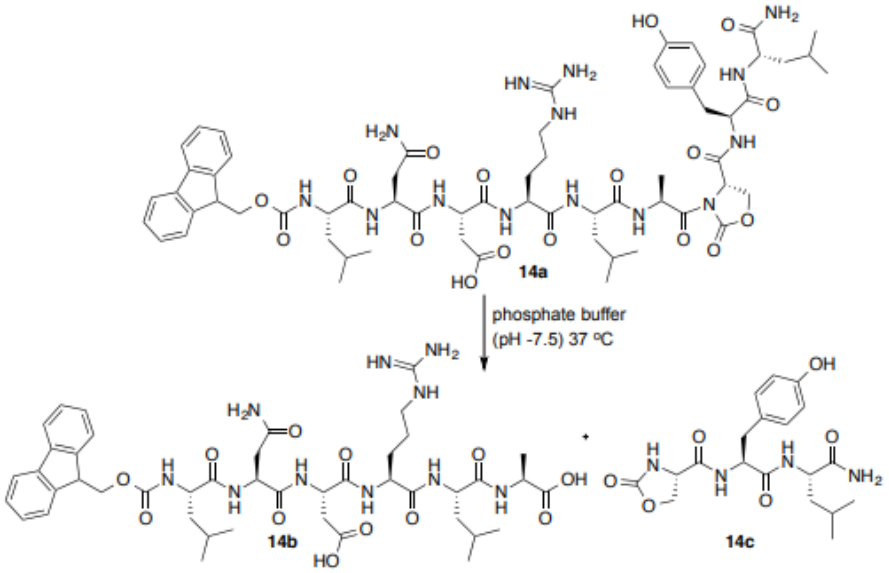

13b
13c
m/z = 670.4
6.00
7.00
8.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
17.00
18.00
19.00

Figure A19 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Arg-Pro-Pro-Gly-Phe-Oxd-Pro-Phe-Arg-NH2 (13a). LCMS: m/z 1308.3 (calcd [M+H]+ = 1308.22), 654.3 (calcd [(M+2/)2] + = 654.61). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.2 min.

Fmoc-Arg-Pro-Pro-Gly-Phe-OH (13b). LCMS : m/z 795.3 (calcd [M+H]+ = 795.64). Purity: >95% (HPLC analysis at 254 nm). Retention time: 14.9 min.

Oxd-Pro-Phe-Arg-NH2 (13c). LCMS: m/z 531.3 (calcd [M+H]+ = 531.57), Purity: >95% (HPLC analysis at 254 nm). Retention time: 7.28 min.


14a
phosphate buffer
(pH -7.5) 37 °C
14b
14c

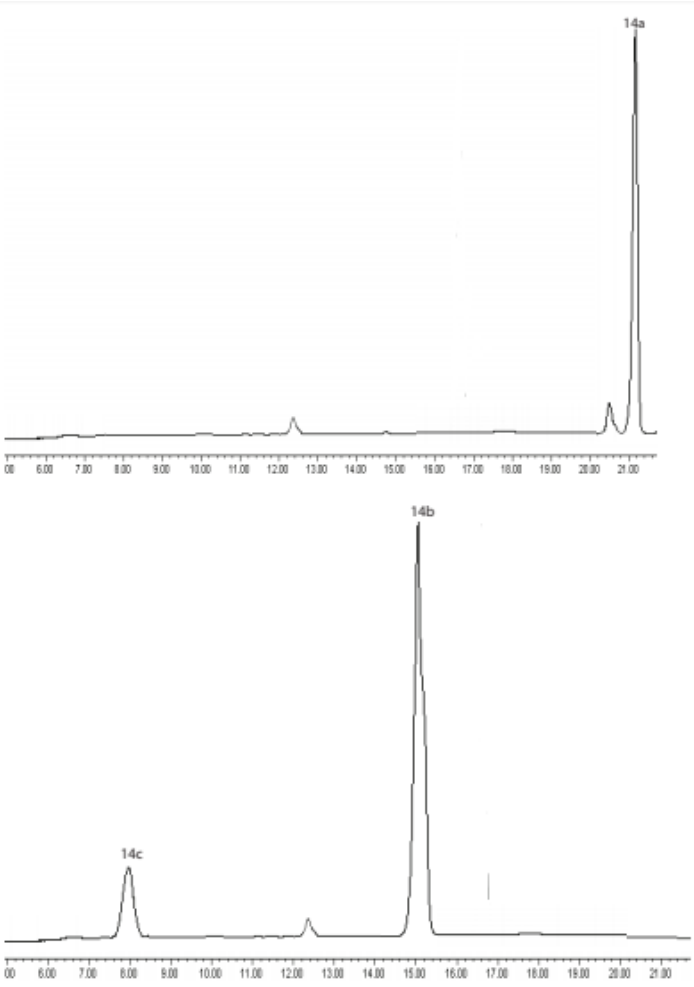
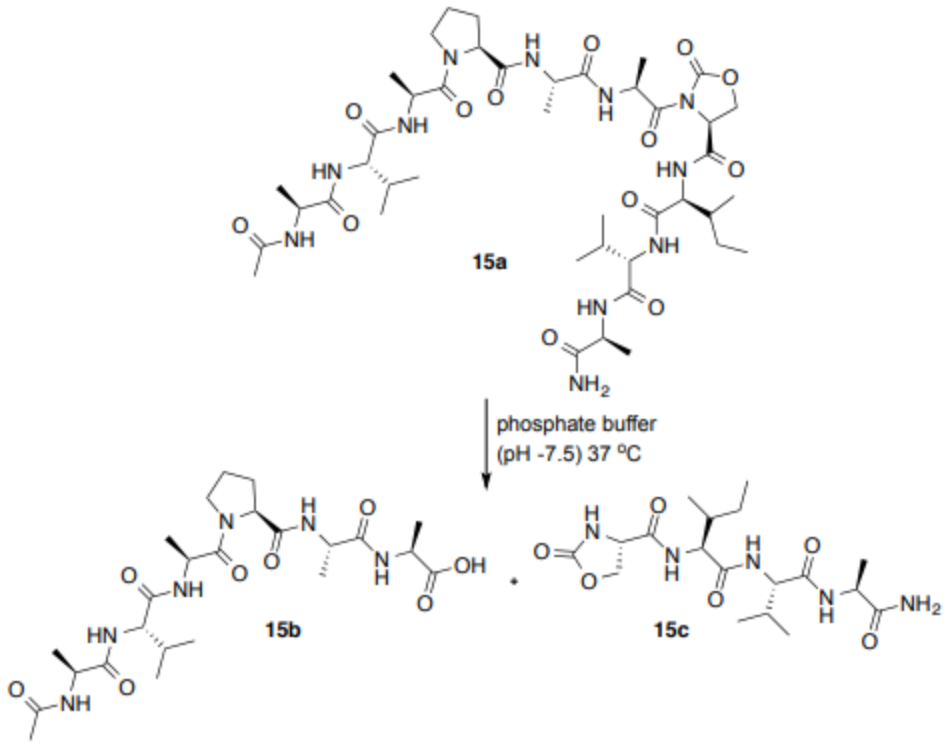

14a
14b
14c

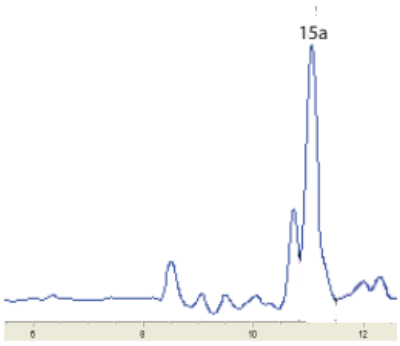
Figure A20 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Leu-Asn-Asp-Arg-Leu-Ala-Oxd-Tyr-Leu-NH2 (14a). LCMS: m/z 1312.0 (calcd [M+H]+ = 1312.2), 656.1 (calcd [(M+2/)2] + = 656.6). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.18 min.

Fmoc-Leu-Asn-Asp-Arg-Leu-Ala-OH (14b). LCMS : m/z 923.4 (calcd [M+H]+ = 923.77). Purity: >95% (HPLC analysis at 254 nm). Retention time: 15.1 min.

Oxd-Tyr-Leu-NH2 (14c). LCMS: m/z 407.2 (calcd [M+H]+ = 407.42), 429.3 (calcd [M+Na]+ = 429.42). Purity: >95% (HPLC analysis at 254 nm). Retention time: 7.9 min.


15a
phosphate buffer
(pH -7.5) 37 °C
15b
15c


15a

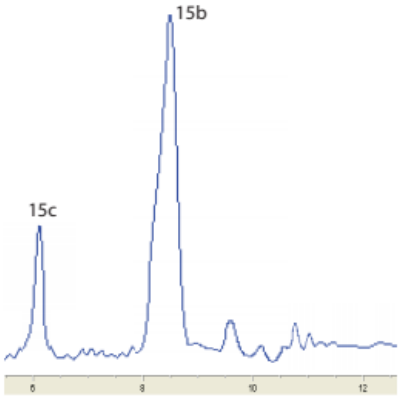
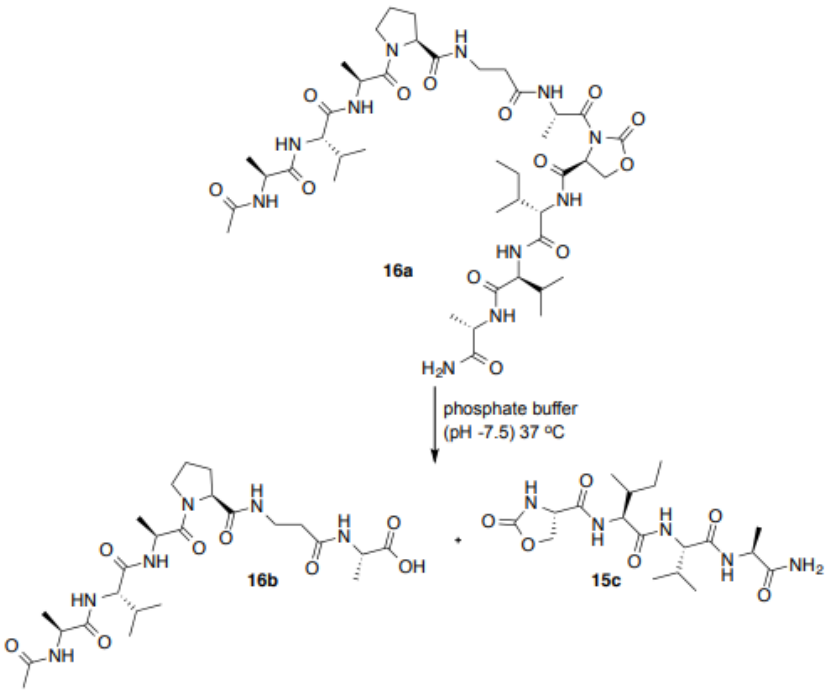

15b
15c

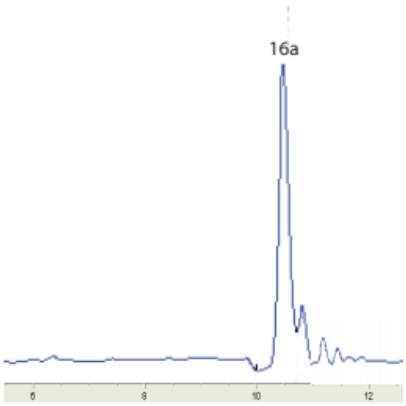
Figure A21 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

OAc-Ala-Val-Ala-Pro-Ala-Ala-Oxd-Ile-Val-Ala-NH2 (15a). LCMS: m/z 937.1 (calcd [M+H]+ = 937.02), Purity: >95% (HPLC analysis at 254 nm). Retention time: 10.9 min.

OAc-Ala-Val-Ala-Pro-Ala-Ala-OH (15b). LCMS : m/z 541.4 (calcd [M+H]+ = 541.56). Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.2 min.

Oxd-Ile-Val-Ala-NH2 (15c). LCMS: m/z 414.3 (calcd [M+H]+ = 414.46), Purity: >95% (HPLC analysis at 254 nm). Retention time: 6.1 min.


16a
phosphate buffer
(pH -7.5) 37 °C
16b
15c


16a

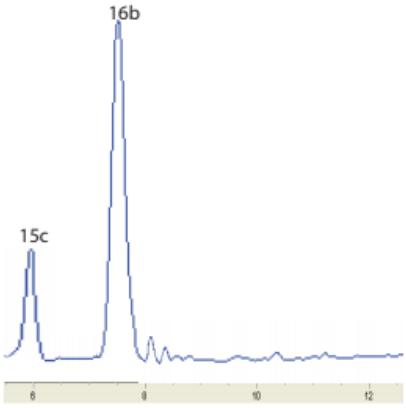
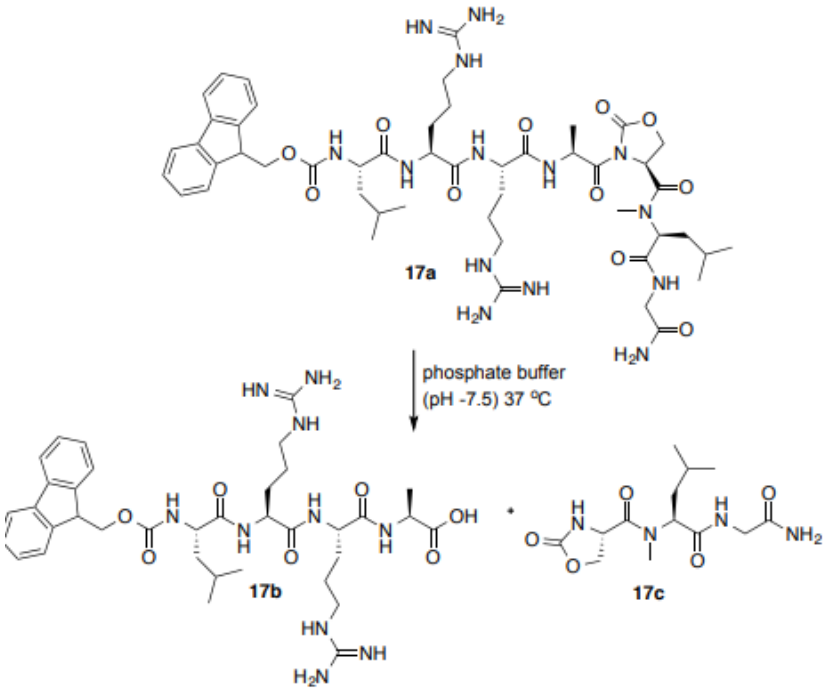

16b
15c

Figure A22 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

OAc-Ala-Val-Ala-Pro-βAla-Ala-Oxd-Ile-Val-Ala-NH2 (16a). LCMS: m/z 937.1 (calcd [M+H]+ = 937.02), 975.1 (calcd [M+K] + = 975.02). Purity: >95% (HPLC analysis at 254 nm). Retention time: 10.6 min.

OAc-Ala-Val-Ala-Pro-βAla-Ala-OH (16b). LCMS : m/z 541.4 (calcd [M+H]+ = 541.56). Purity: >95% (HPLC analysis at 254 nm). Retention time: 7.6 min.

Oxd-Ile-Val-Ala-NH2 (15c). LCMS: m/z 414.3 (calcd [M+H]+ = 414.46), Purity: >95% (HPLC analysis at 254 nm). Retention time: 5.9 min.


17a
phosphate buffer
(pH ~7.5) 37 °C
17b
17c

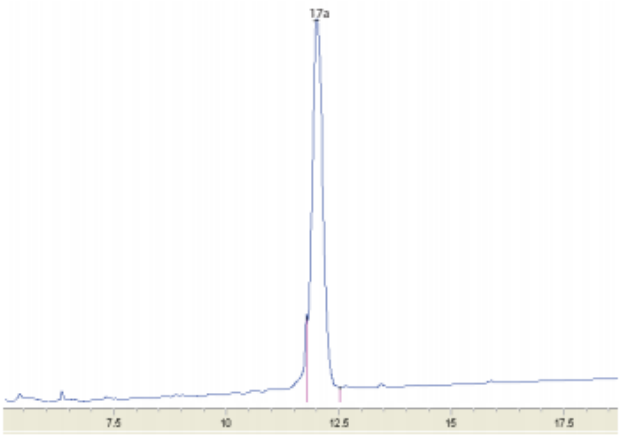

17a
17b
17c
7.5
10
12.5
15
17.5

Figure A23 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in buffer solution; the product peak is labeled.

Fmoc-Leu-Arg-Arg-Ala-Oxd-(N-methyl)Leu-Gly-NH2 (17a). LCMS: m/z 1033.5 (calcd [M+H]+ = 1033.91), 517.3 (calcd [(M+2/)2] + = 517.45). Purity: >95% (HPLC analysis at 254 nm). Retention time: 11.9 min.

Fmoc-Leu-Arg-Arg-Ala-OH (17b). LCMS : m/z 737.4 (calcd [M+H]+ = 737.61), 369.1 (calcd [(M+2/)2] + = 369.3)Purity: >95% (HPLC analysis at 254 nm). Retention time: 11.51 min.

Oxd-(N-methyl)Leu-Gly-NH2 (17c). LCMS: m/z 315.2 (calcd [M+H]+ = 315.3), 337.1 (calcd [M+Na]+ = 337.3) Purity: >95% (HPLC analysis at 254 nm). Retention time: 8.88 min.

CHAPTER 3 SUPPLEMENTAL MATERIAL

All of the following figures have been reproduced with permission from: Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Zong, C.; Link, A. J.; Raj, M. *Angew. Chem.* **2018**, DOI: 10.1002/anie.201801299.

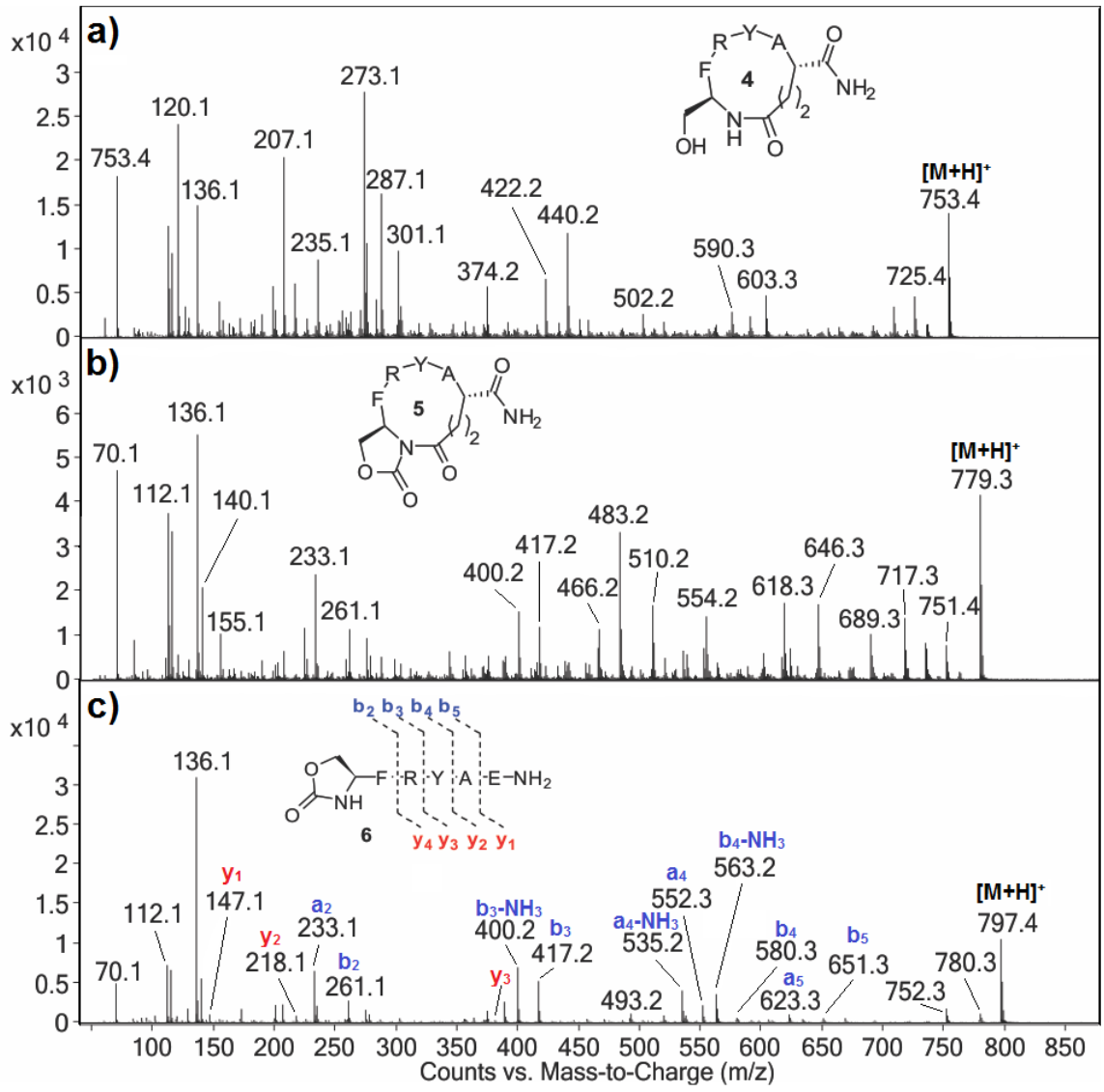


Figure A24 MS/MS data for a) *cyc*(SFRYAE) **4**, b) activated oxazolidinone (*Oxd*)-containing macrocyclic peptide, *cyc*(*Oxd*-FRYAE) **5**, and c) *Oxd*-FRYAE **6** showing that MS/MS of ring-opened peptide is much simpler to interpret.

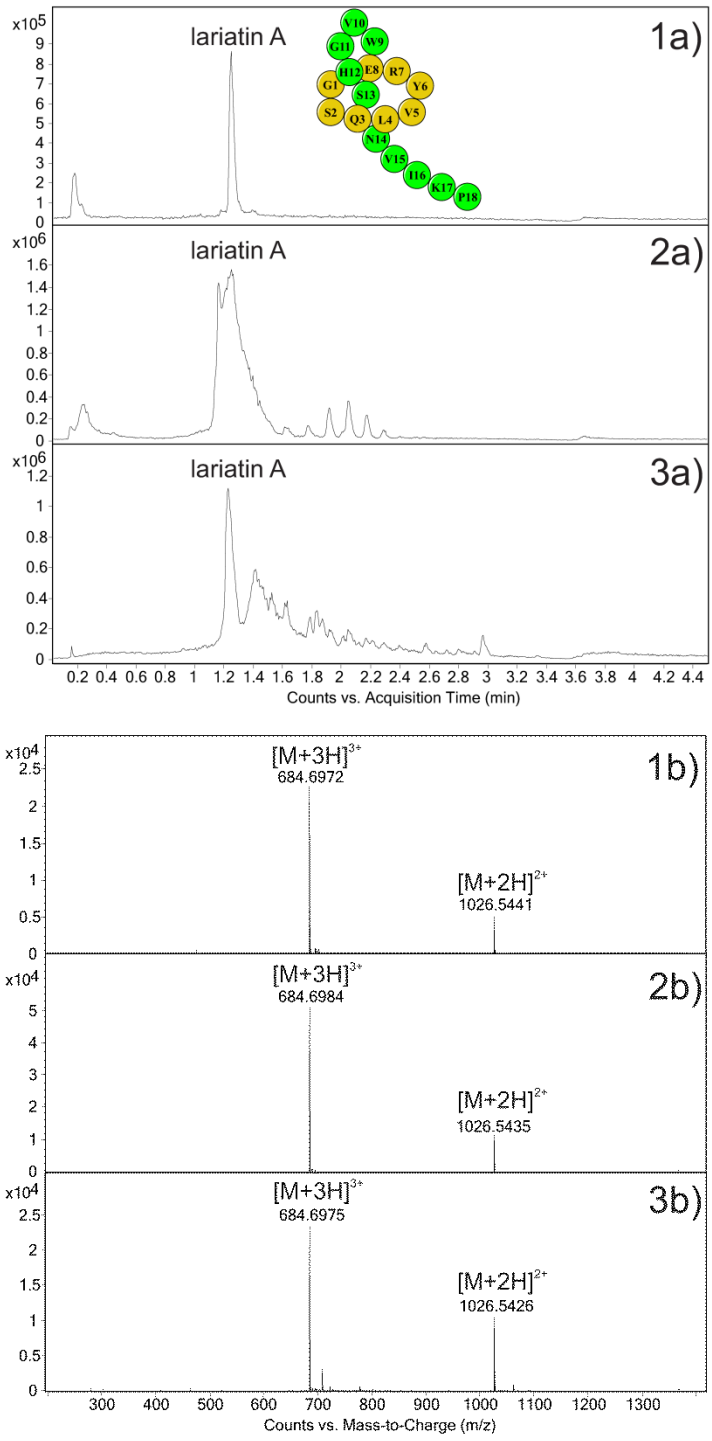


Figure A25 Total ion chromatograms and HRMS spectra for lariatin A (1a/b) incubated with trypsin (2a/b) and both trypsin and chymotrypsin (3a/b) for 48 h. Some protease degradation was observed due to self-cleavage. Peak intensities and ion abundances were similar for each sample indicating that lariatin A is stable to protealytic cleavage.

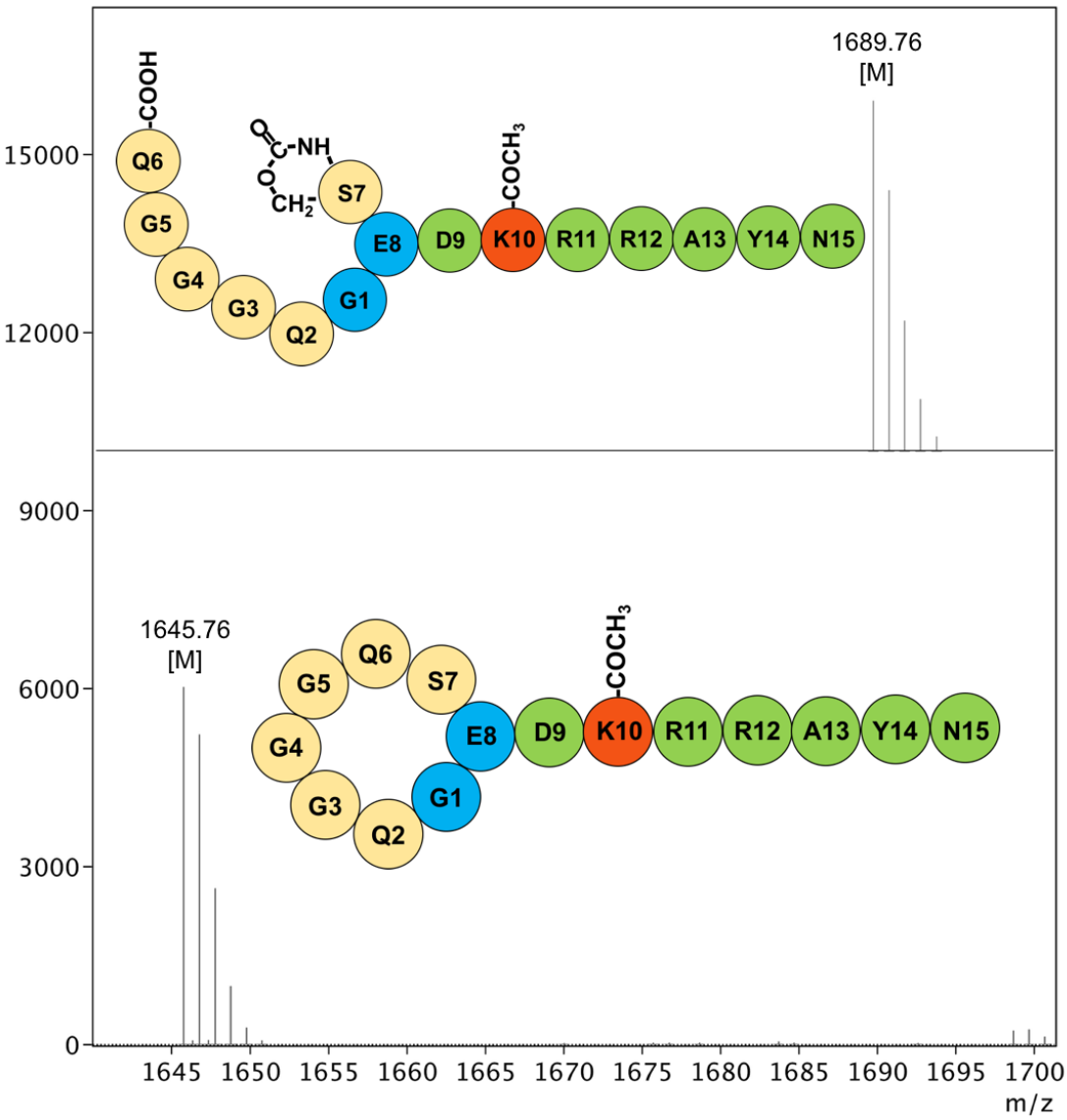


Figure A26 Deconvoluted mass spectra for albusnodin (below) and chemical cleavage product (above) at Ser-7.

CHAPTER 4 SUPPLEMENTAL MATERIAL

All of the following figures have been reproduced with permission from: Elashal, H. E.; Cohen, R. D.; Elashal, H. E.; Raj, M. *Org. Lett.* **2018**, *20* (8), 2374.

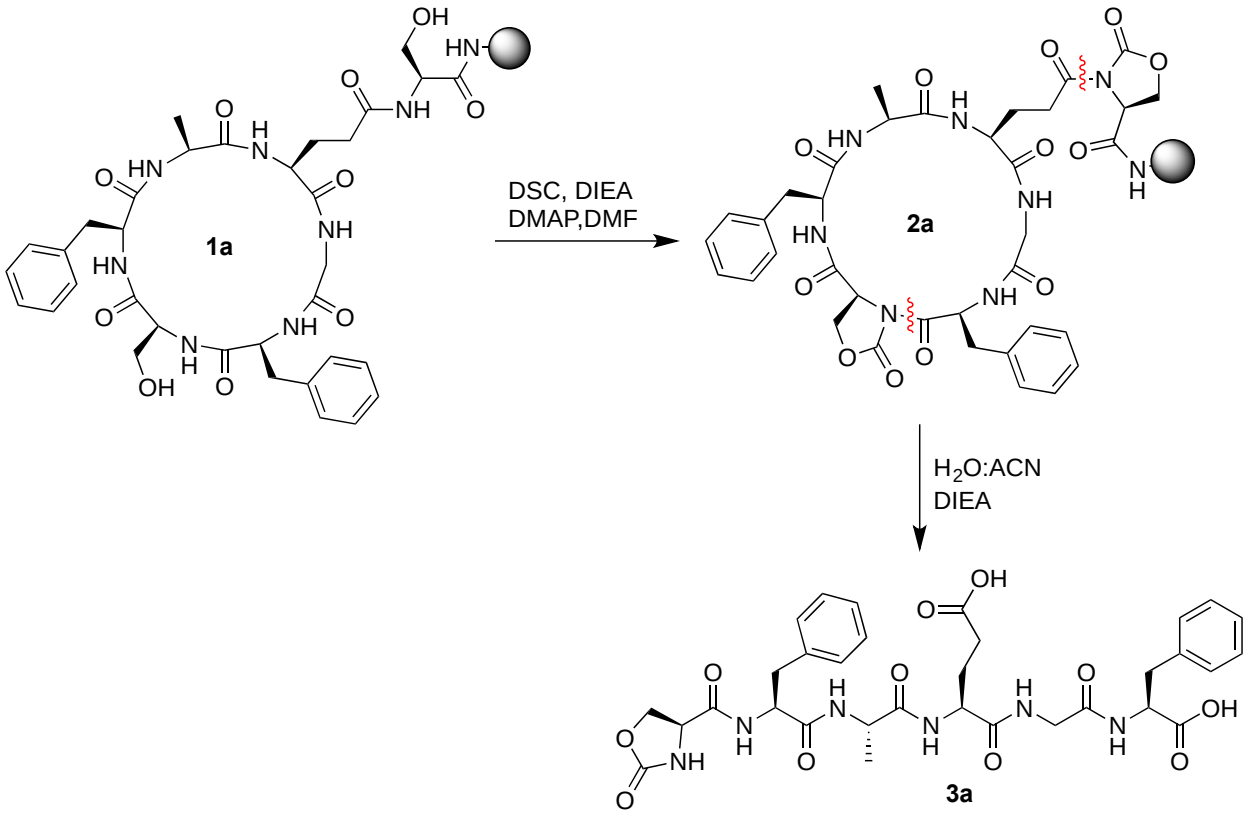


Figure A27 Macrocyclic ring-opening and resin-cleavage of peptide **1a**, *cyc*(GFSFAE)-S-*Rink*

***cyc*(Gly-Phe-Ser-Phe-Ala-Glu)-Ser-CONH$_2$ (1a).** LC-MS: *m/z* 724.8 (calcd $[M+H]^+$ = 725.3), *m/z* 746.8 (calcd $[M+Na]^+$ = 747.3), Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.19 min

***cyc*(Gly-Phe-*Oxd*-Phe-Ala-Glu)-*Oxd*-CONH$_2$ (2a).** LC-MS: *m/z* 776.7 (calcd $[M+H]^+$ = 777.2), *m/z* 798.7 (calcd $[M+Na]^+$ = 799.2), 1552.2 (calcd $[2M+H]^+$ = 1554.4), 1574.3 (calcd $[2M+Na]^+$ = 1576.4), Purity: >95% (HPLC analysis at 220 nm). Retention time: 13.5 min

***Oxd*-Phe-Ala-Glu-Gly-Phe (3a).** LC-MS: *m/z* 682.7 (calcd $[M+H]^+$ = 683.2), *m/z* 704.8 (calcd $[M+Na]^+$ = 705.2), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.3 min

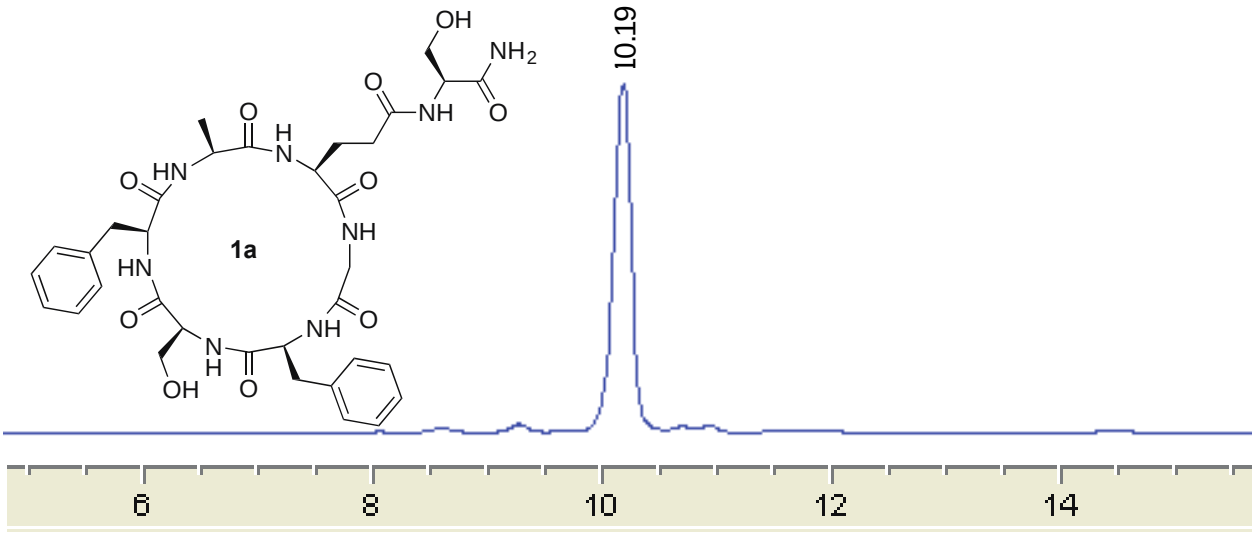


Figure A28 HPLC trace and mass spectrum of peptide **1a**, *cyc*(GFSFAE)-S-$CONH_2$ after resin cleavage

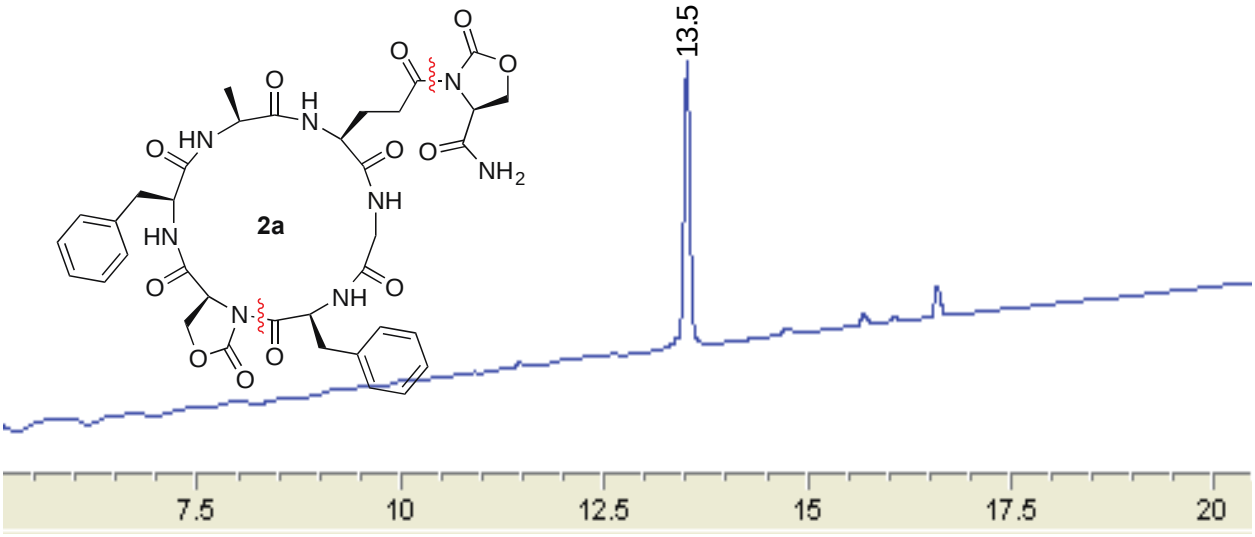


Figure A29 HPLC trace and mass spectrum of peptide **2a**, *Cyc*(GF-*Oxd*-FAE)-*Oxd*-$CONH_2$ after resin cleavage

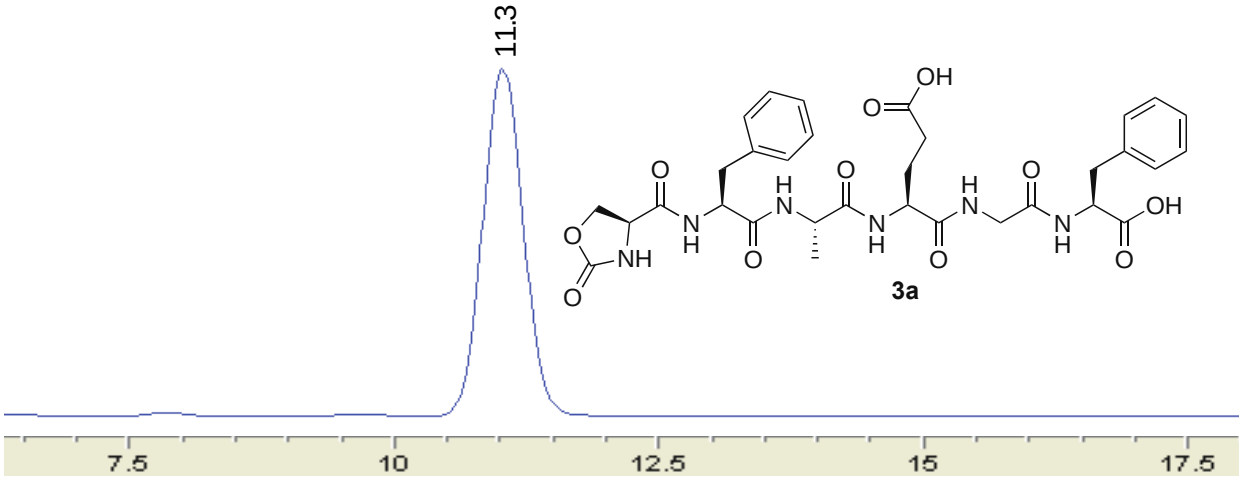


Figure A30 HPLC trace and mass spectrum of peptide **3a**, *Oxd*-FAEGF-OH

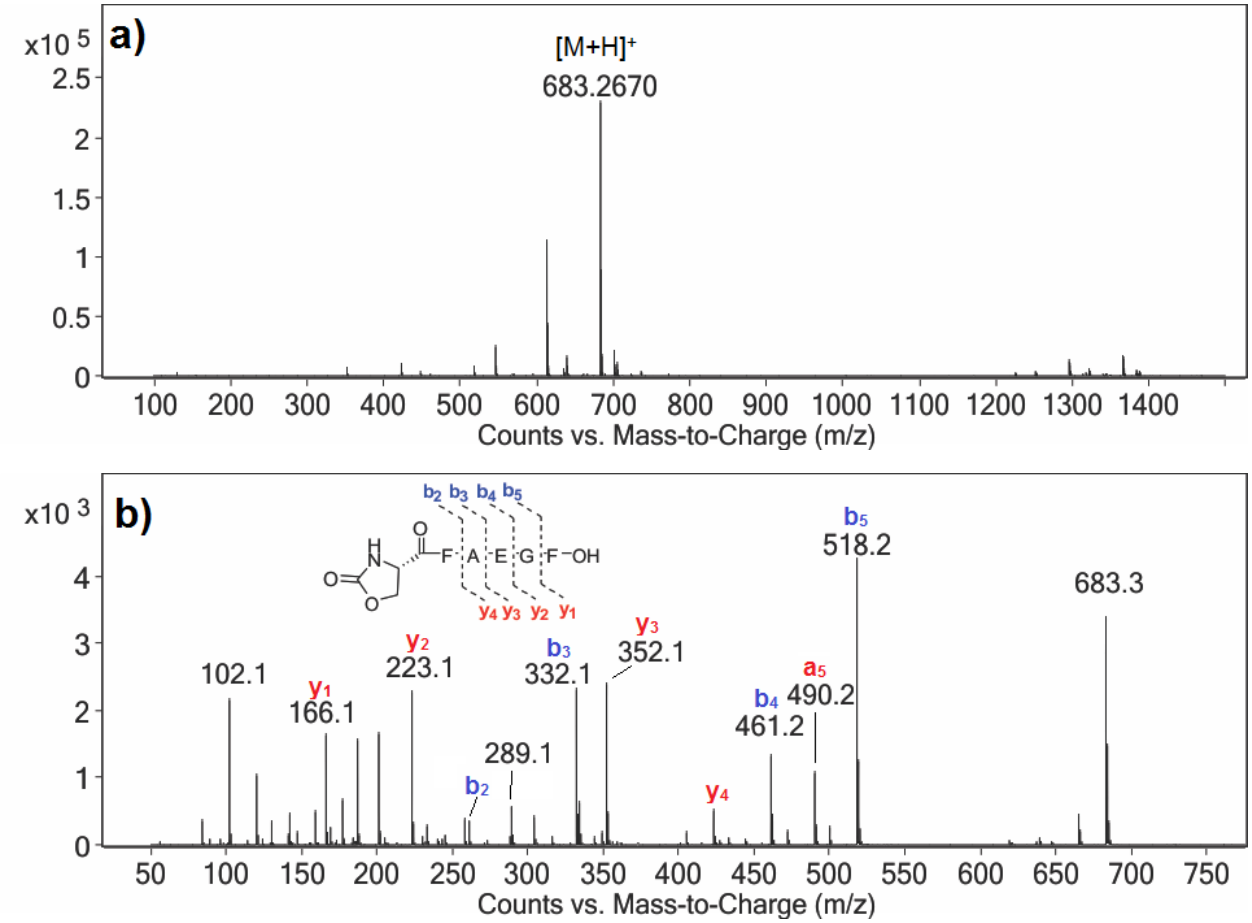


Figure A31 LC-IM-MS/MS data for linear peptide **3a**, *Oxd*-FAEGF-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time. Calcd $[M+H]^+$ = 683.2671, Found $[M+H]^+$ = 683.2670

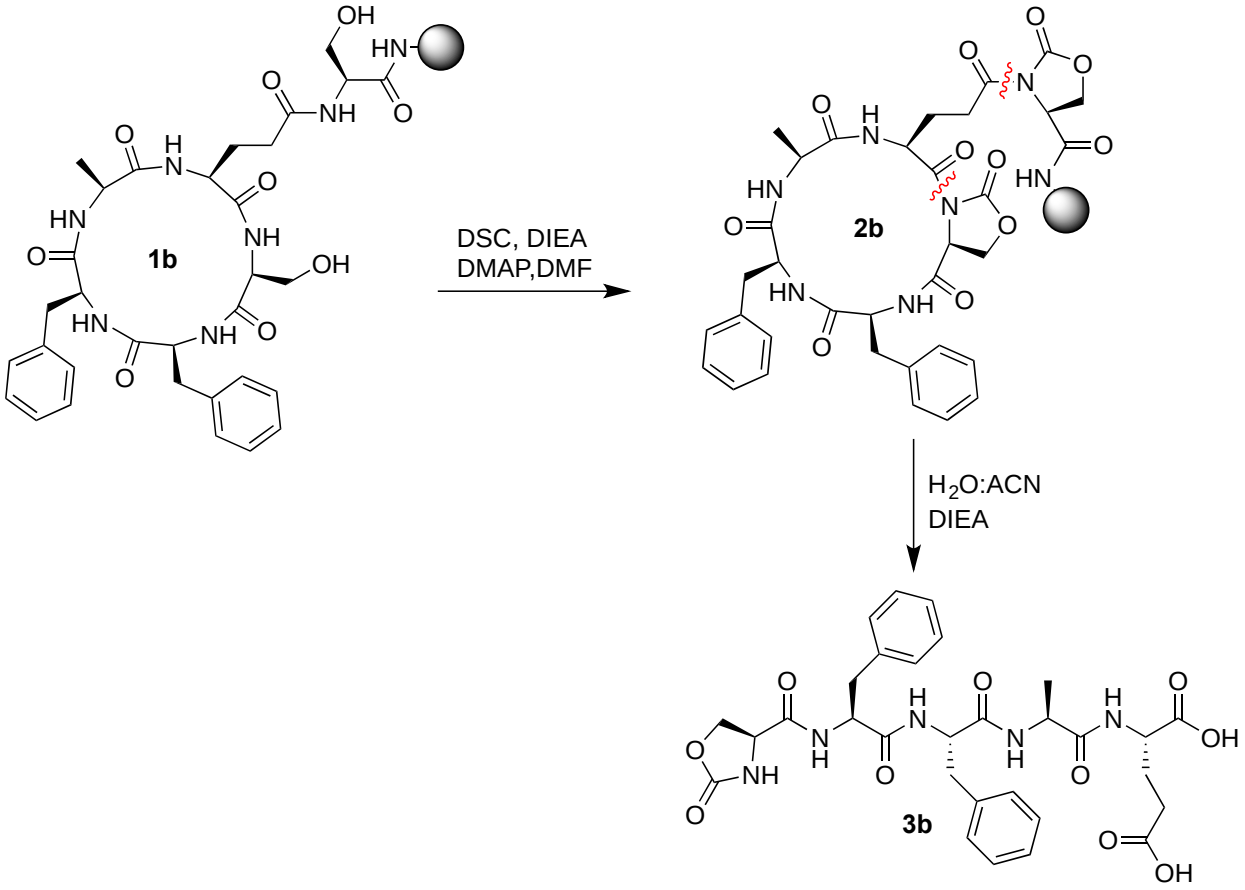


Figure A32 Macrocyclic ring-opening and resin-cleavage of peptide **1b**, *cyc*(SFFAE)-S-*Rink*

***cyc*(Ser-Phe-Phe-Ala-Glu)-Ser-CONH$_2$ (1b).** LCMS: *m/z* 668.0 (calcd $[M+H]^+$ = 668.3), *m/z* 690.3 (calcd $[M+Na]^+$ = 690.3), Purity: >95% (HPLC analysis at 220 nm). Retention time: 13.96

***Oxd*-Phe-Phe-Ala-Glu (3b).** LCMS: *m/z* 625.8 (calcd $[M+H]^+$ = 626.24), *m/z* 647.5 (calcd $[M+Na]^+$ = 648.24), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.71 min

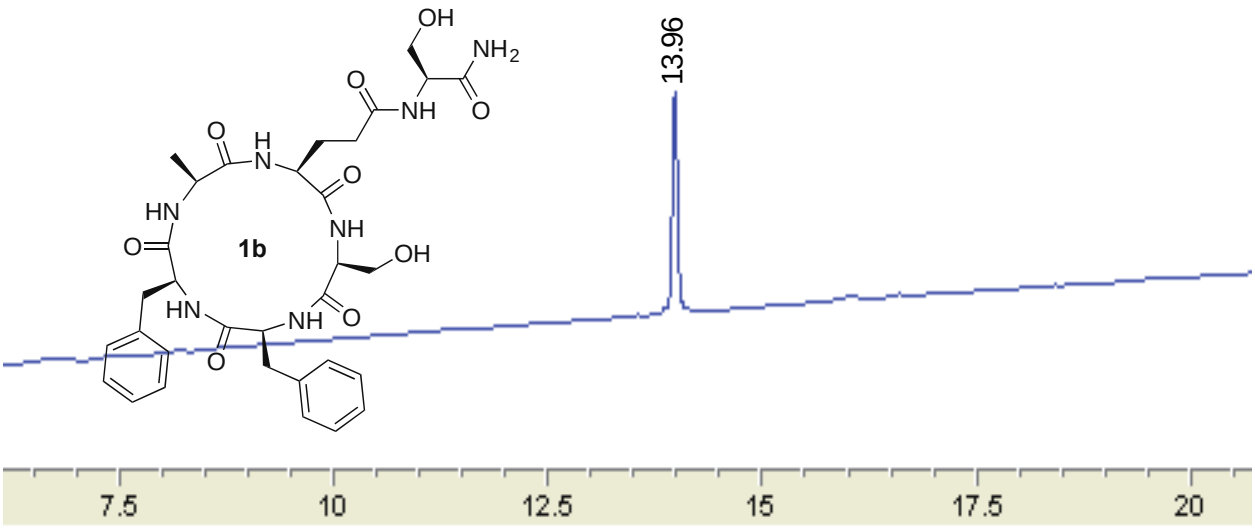


Figure A33 HPLC trace and mass spectrum of peptide **1b**, *cyc*(SFFAE)-S-$CONH_2$ after resin cleavage

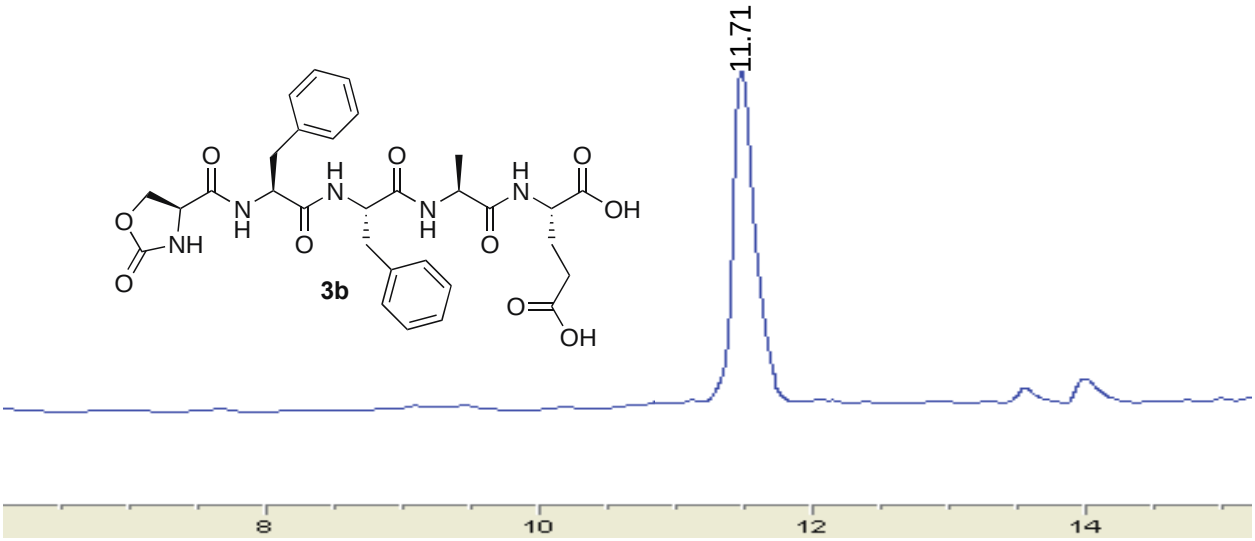


Figure A34 HPLC trace and mass spectrum of peptide **3b**, *Oxd*-SFFAE-OH

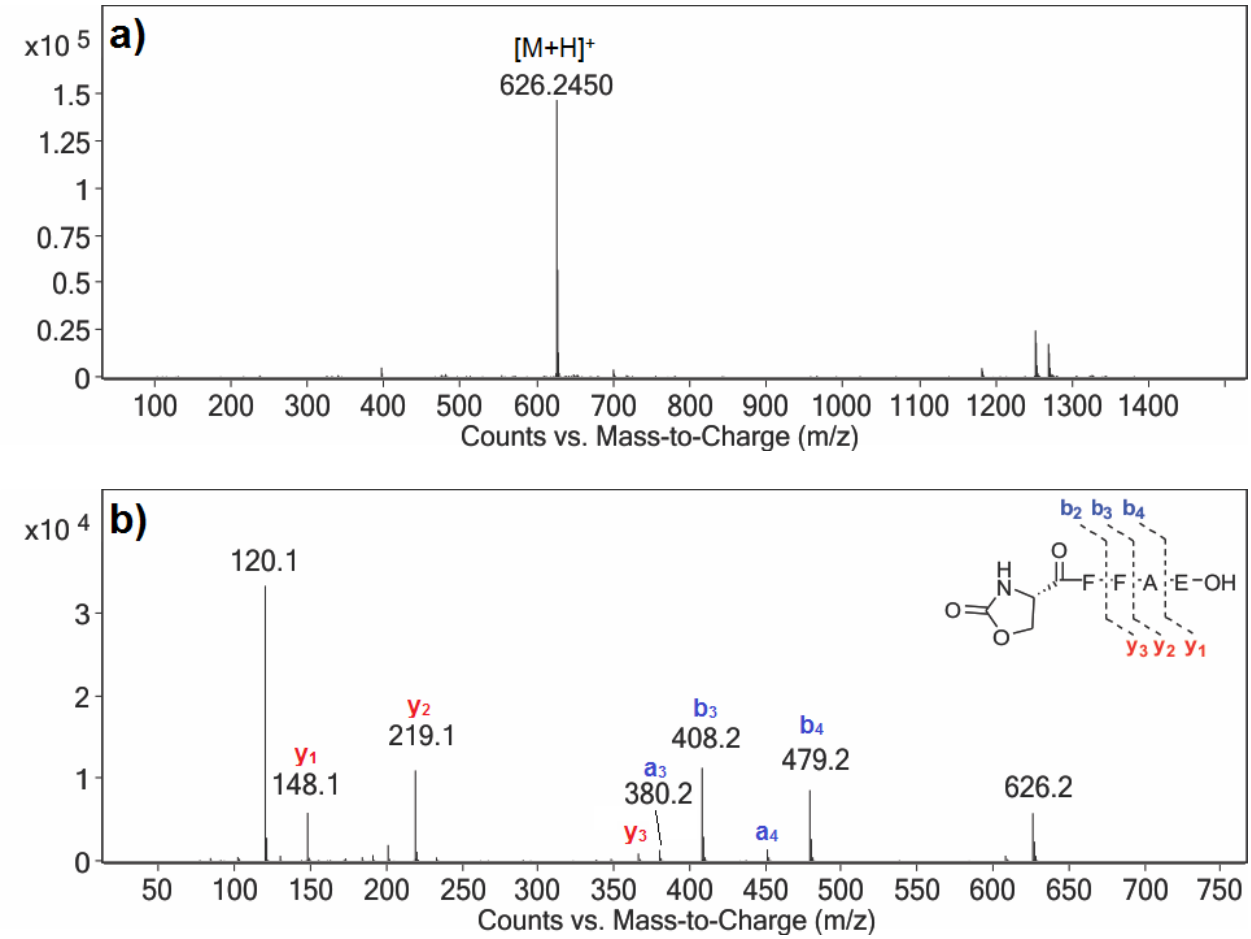


Figure A35 LC-IM-MS/MS data for linear peptide **3b**, *Oxd*-FFAE-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 626.2457, Found $[M+H]^+$ = 626.2450

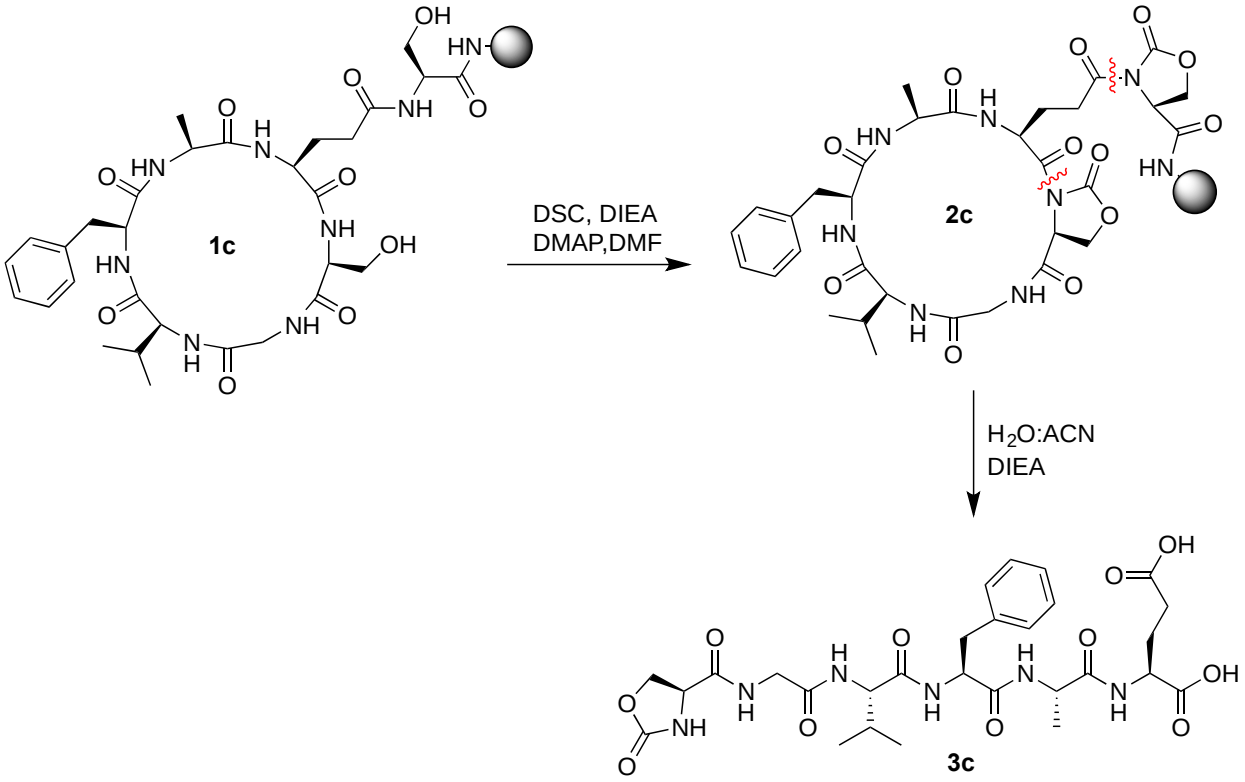


Figure A36 Macrocyclic ring-opening and resin-cleavage of peptide **1c**, *cyc*(SGVFAE)-S-*Rink* ***cyc*(Ser-Gly-Val-Phe-Ala-Glu)-Ser-CONH$_2$ (1c).** LCMS: *m/z* 676.70 (calcd $[M+H]^+$ = 677.3), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.5 min

***Oxd*-Gly-Val-Phe-Ala-Glu (3c).** LCMS: *m/z* 635.0 (calcd $[M+H]^+$ = 635.6), Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.9 min

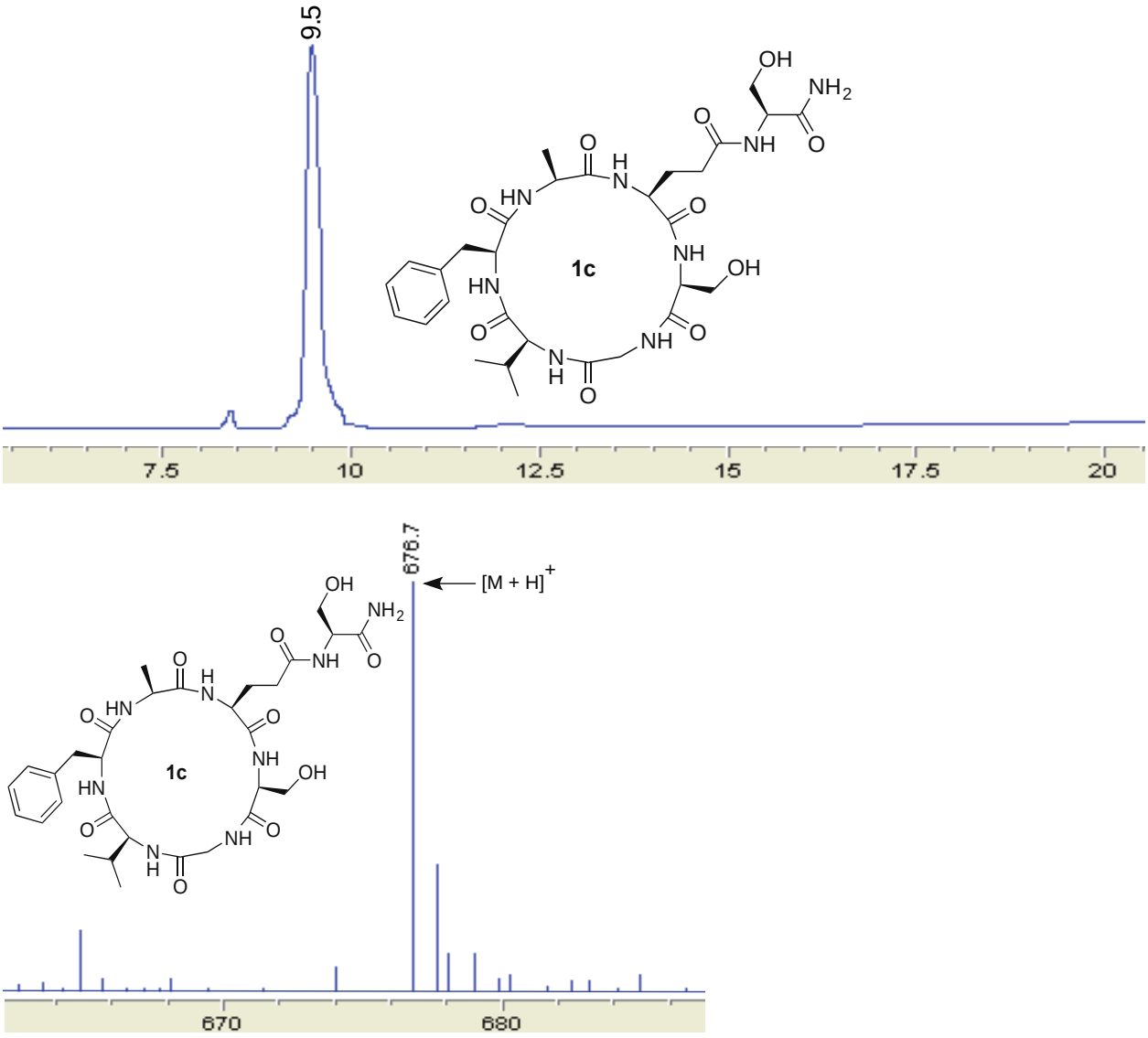


Figure A37 HPLC trace and mass spectrum of peptide **1c**, *cyc*(SGVFAE)-S-$CONH_2$ after resin cleavage

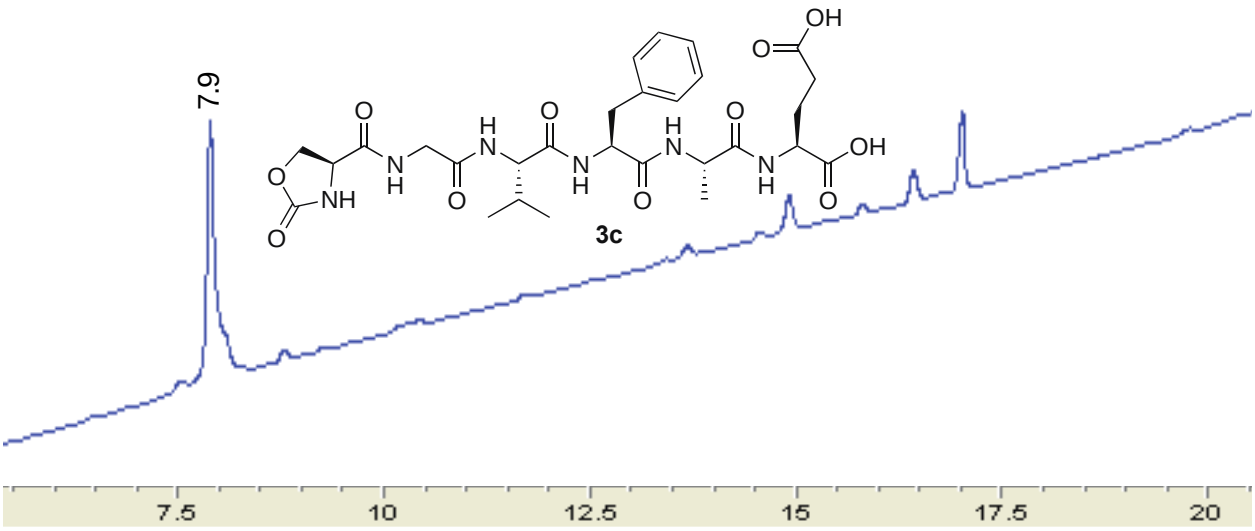


Figure A38 HPLC trace of peptide **3c**, *Oxd*-GVFAE-OH

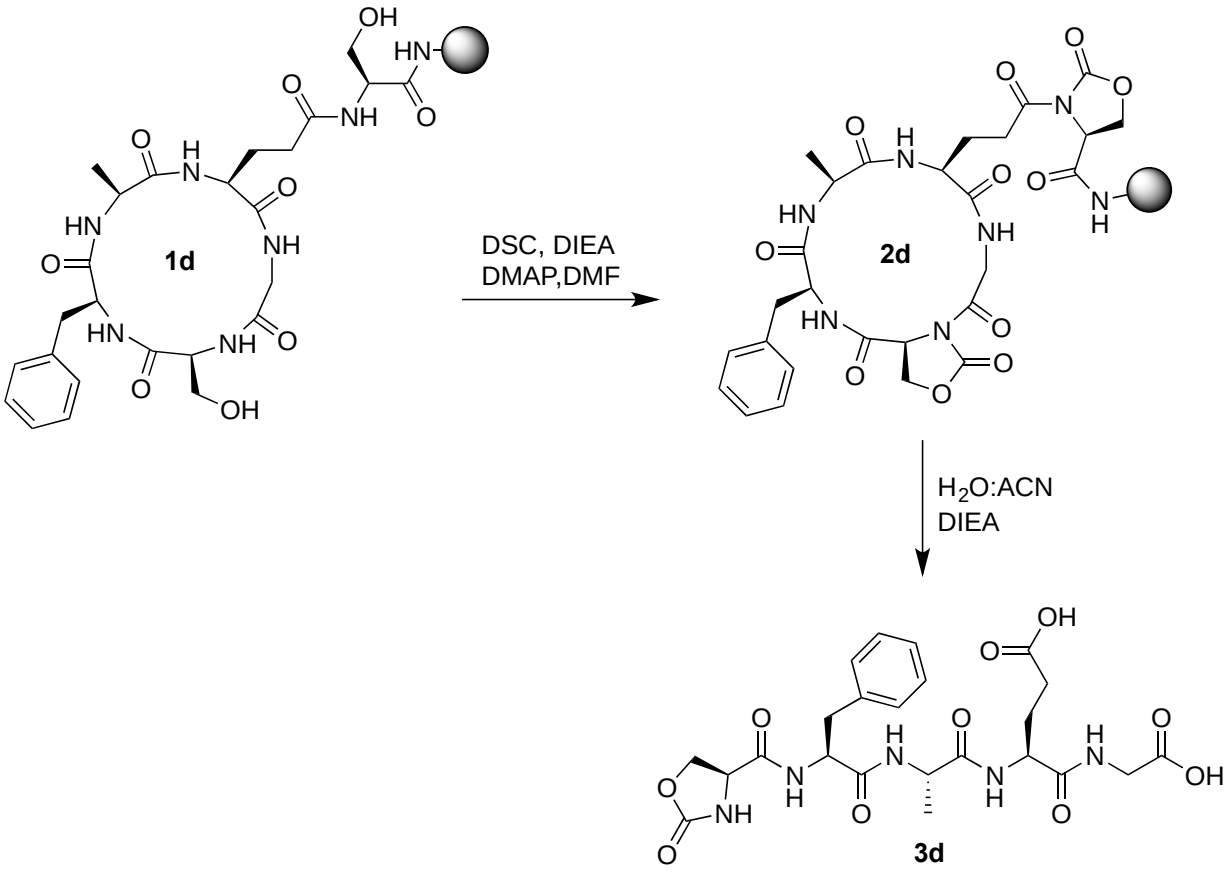
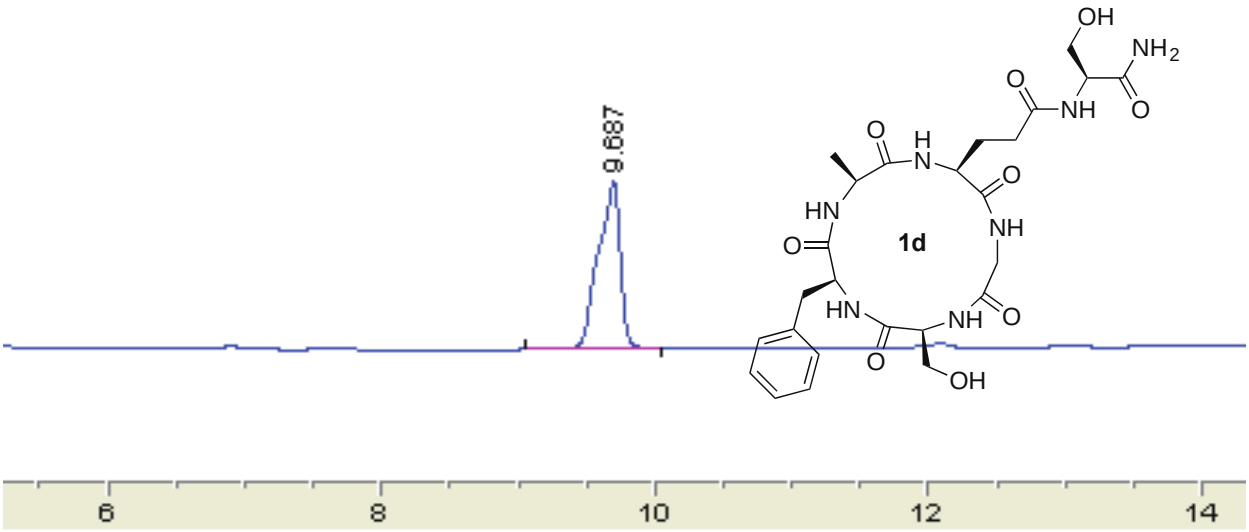


Figure A39 Macrocyclic ring-opening and resin-cleavage of peptide **1d**, *cyc*(GSFAE)-S-*Rink*

***cyc*(Gly-Ser-Phe-Ala-Glu)-Ser-CONH$_2$ (1d).** LCMS: *m/z* 578.0 (calcd $[M+H]^+$ = 578.2), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.68 min

***Oxd*-Phe-Ala-Glu-Gly (3d).** LCMS: *m/z* 536.0 (calcd $[M+H]^+$ = 536.1), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.55 min



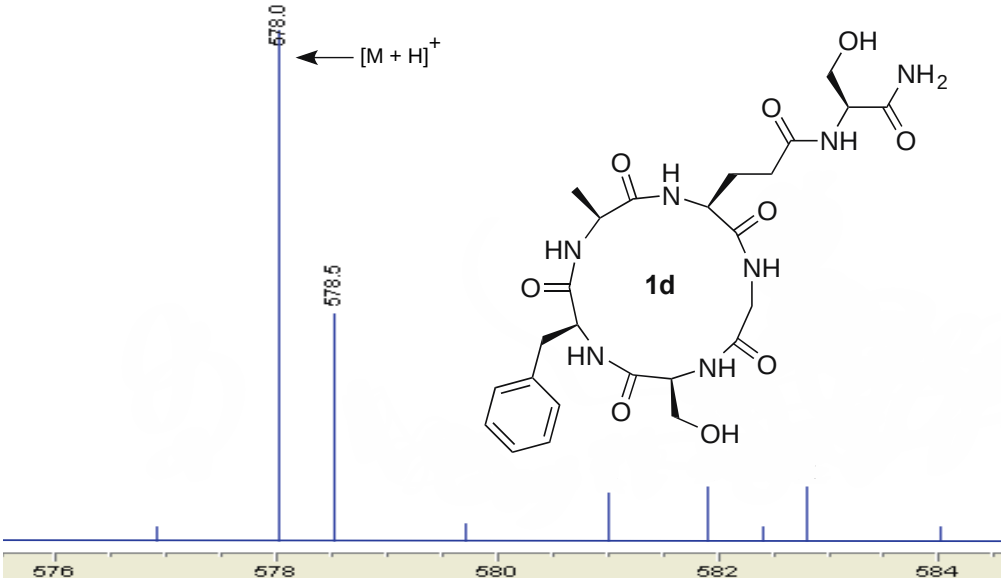


Figure A40 HPLC trace and mass spectrum of peptide **1d**, *cyc*(GSFAE)-S-CONH$_2$ after resin cleavage

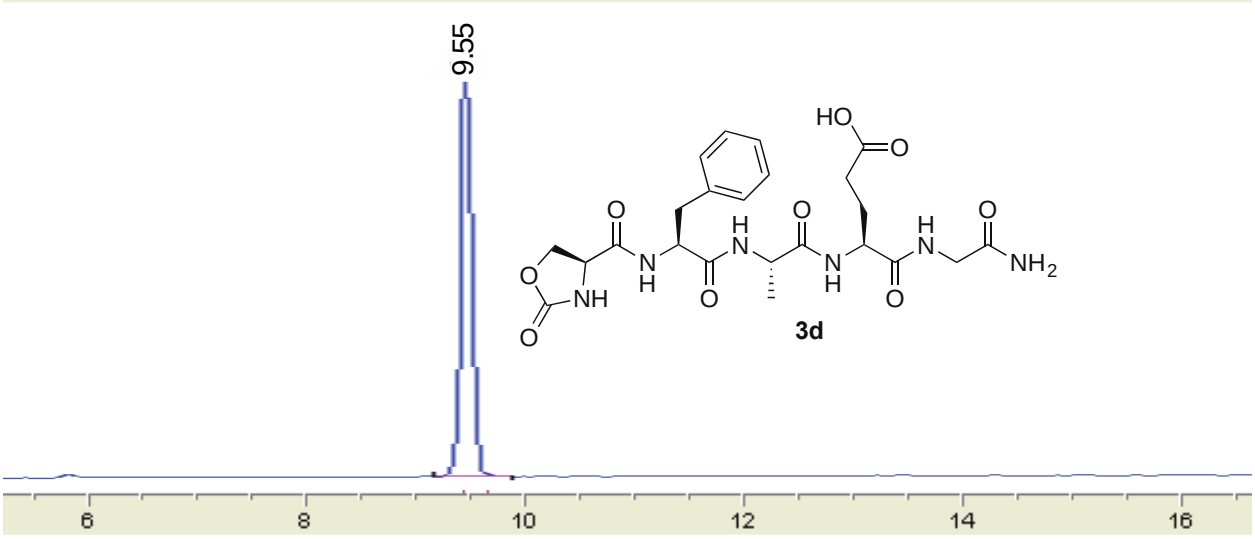


Figure A41 HPLC trace of peptide **3d**, *Oxd*-FAEG-NH_2

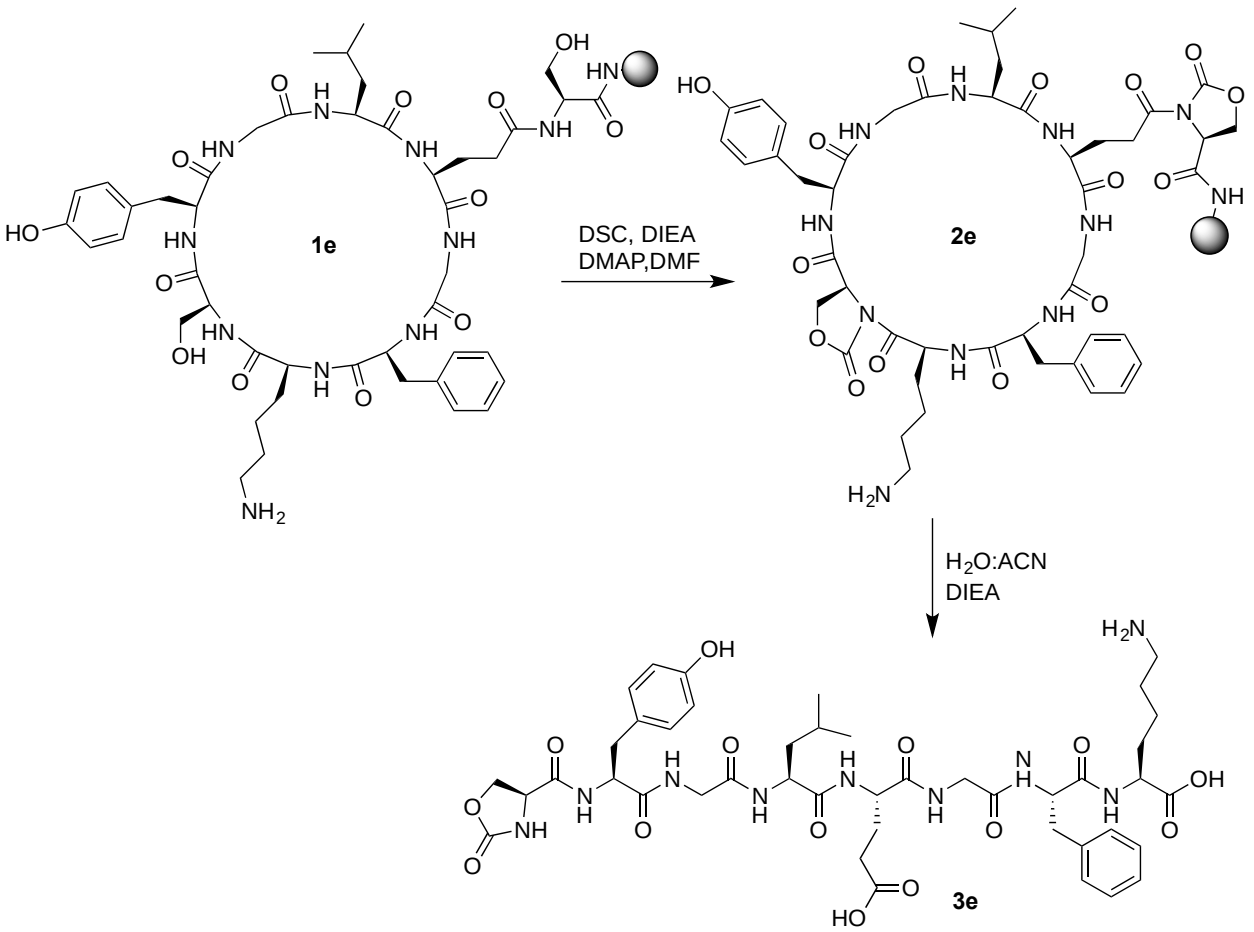


Figure A42 Macrocyclic ring-opening and resin-cleavage of peptide **1e**, *cyc*(GFKSYGLE)-S-*Rink*

***cyc*(Gly-Phe-Lys-Ser-Tyr-Gly-Leu-Glu)-Ser-CONH$_2$ (1e).** LCMS: *m/z* 967.7 (calcd $[M+H]^+$ = 968.4), *m/z* 989.7, (calcd $[M+Na]^+$ = 990.4), *m/z* 484.5 (calcd $[(M+2)/2]^+$ = 484.7), Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.12

***Oxd*-Tyr-Gly-Leu-Glu-Gly-Phe-Lys (3e).** LCMS: *m/z* 925.7 (calcd $[M+H]^+$ = 926.4), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.5 min

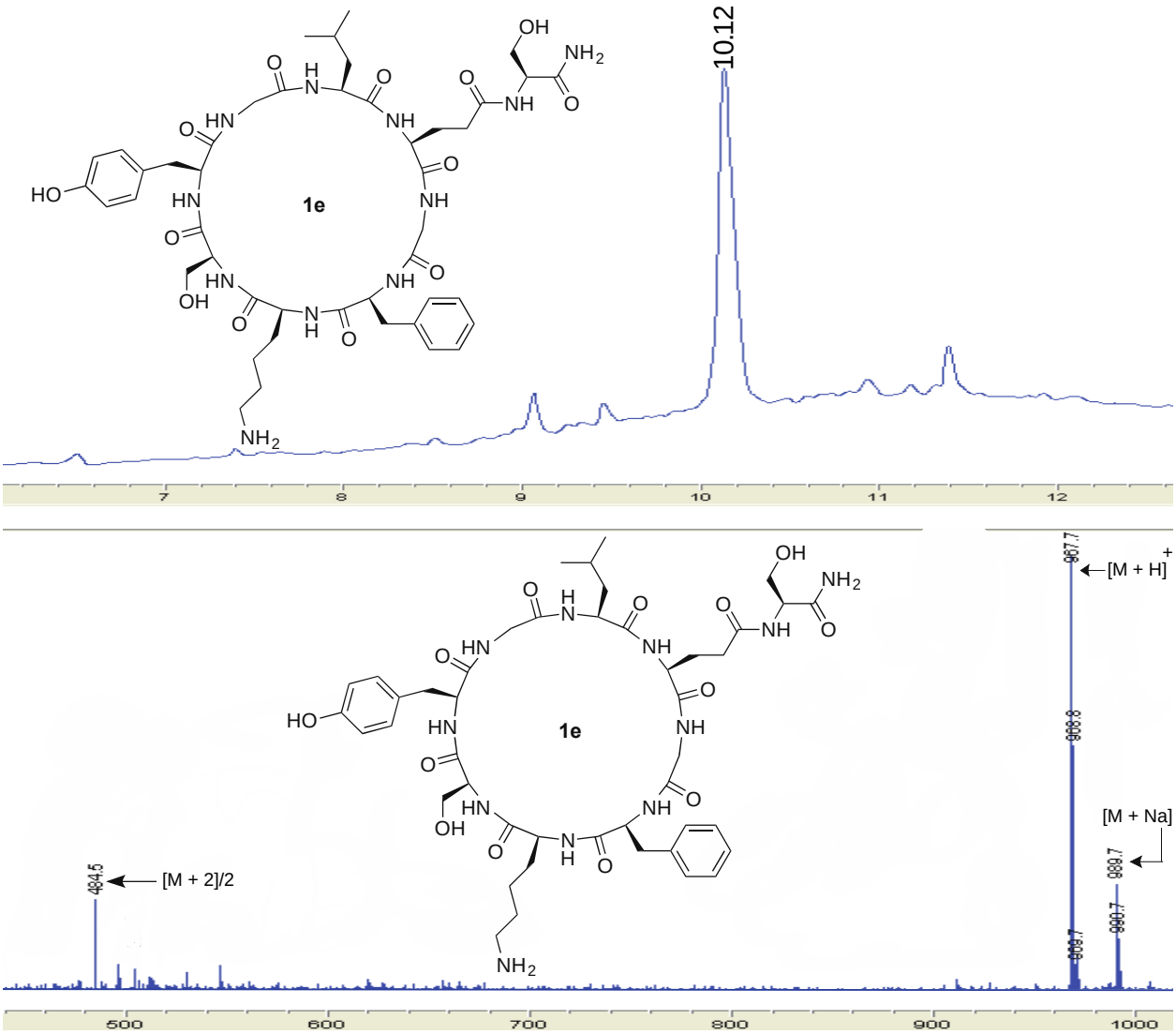


Figure A43 HPLC trace and mass spectrum of peptide **1e**, *cyc*(GFKSYGLE)-S-CONH$_2$ after resin cleavage

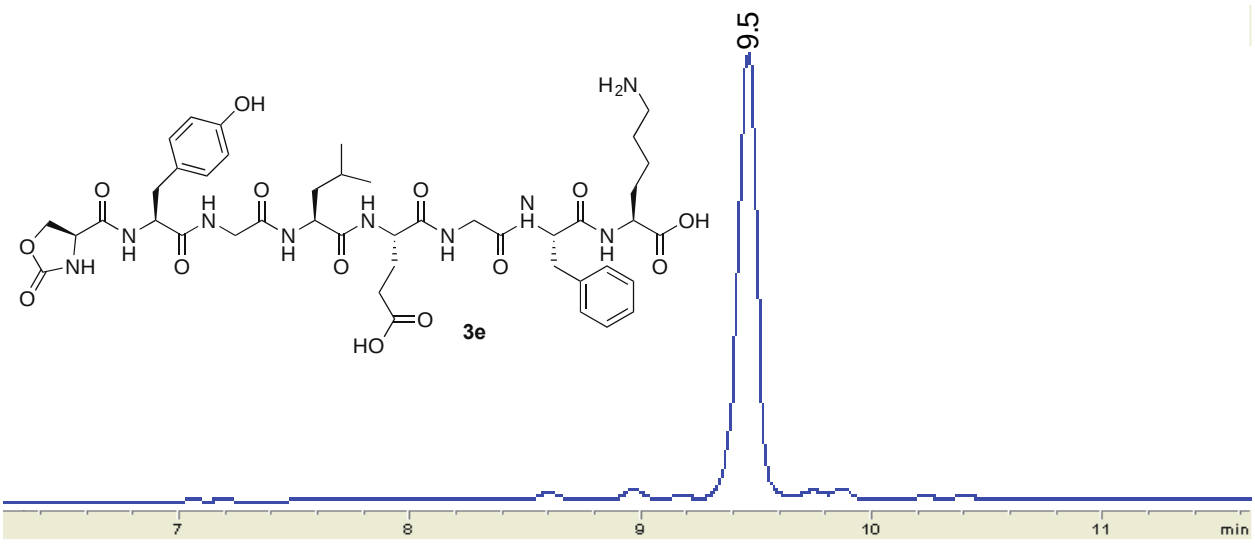


Figure A44 HPLC trace of peptide **3e**, *Oxd*-YGLEGFK-OH

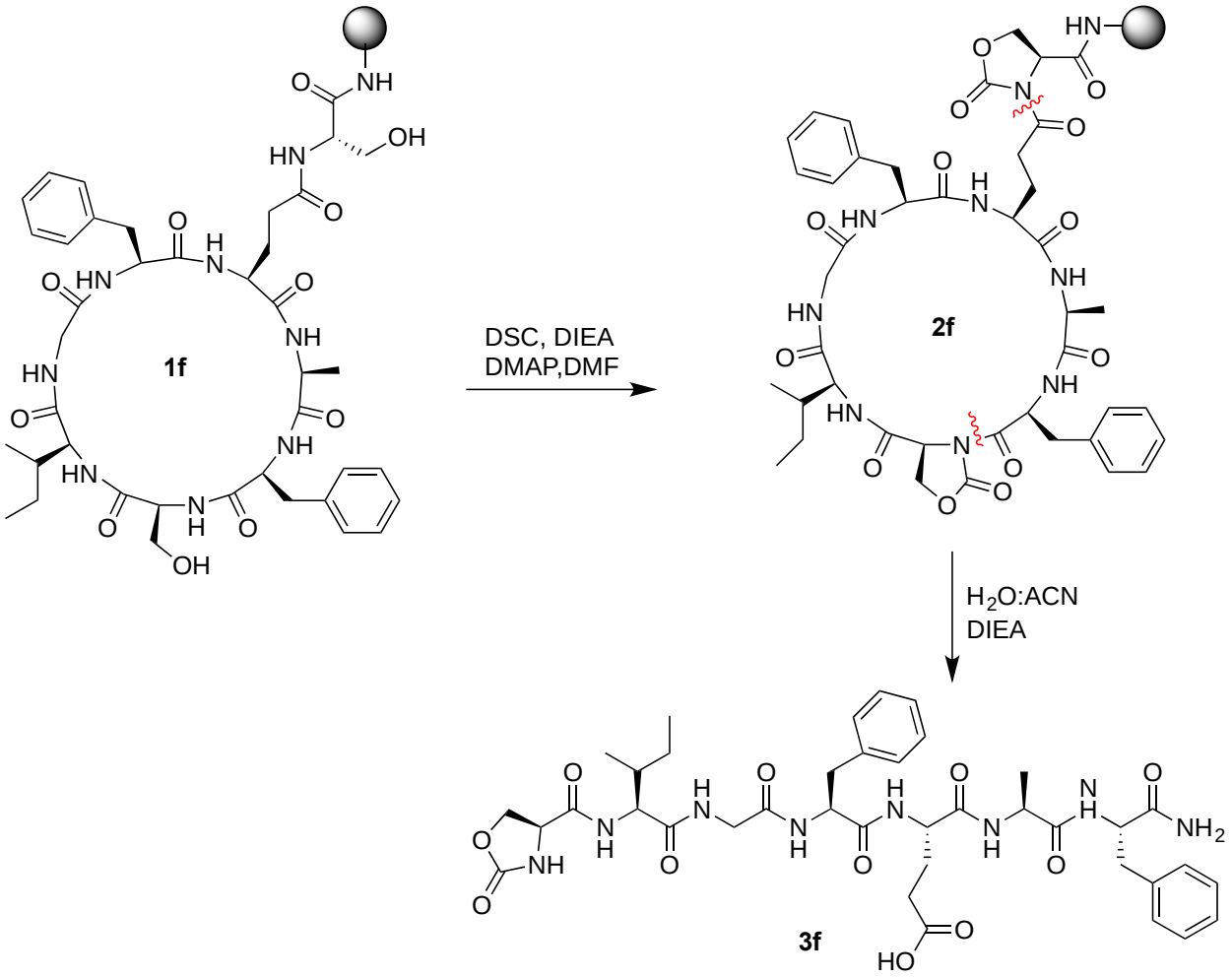


Figure A45 Macrocyclic ring-opening and resin-cleavage of peptide **1f**, *cyc*(AFSIGFE)-S-*Rink* ***cyc*(Ala-Phe-Ser-Ile-Gly-Phe-Glu)-Ser-CONH$_2$ (1f).** LCMS: *m/z* 837.8 (calcd $[M+H]^+$ = 838.4), *m/z* 859.7 (calcd $[M+Na]^+$ = 860.4), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.9

***Oxd*-Ile-Gly-Phe-Glu-Ala-Phe (3f).** LCMS: *m/z* 795.8 (calcd $[M+H]^+$ = 795.3), *m/z* 818.7, (calcd $[M+Na]^+$ = 817.3), Purity: >95% (HPLC analysis at 220 nm). Retention time: 12.8 min

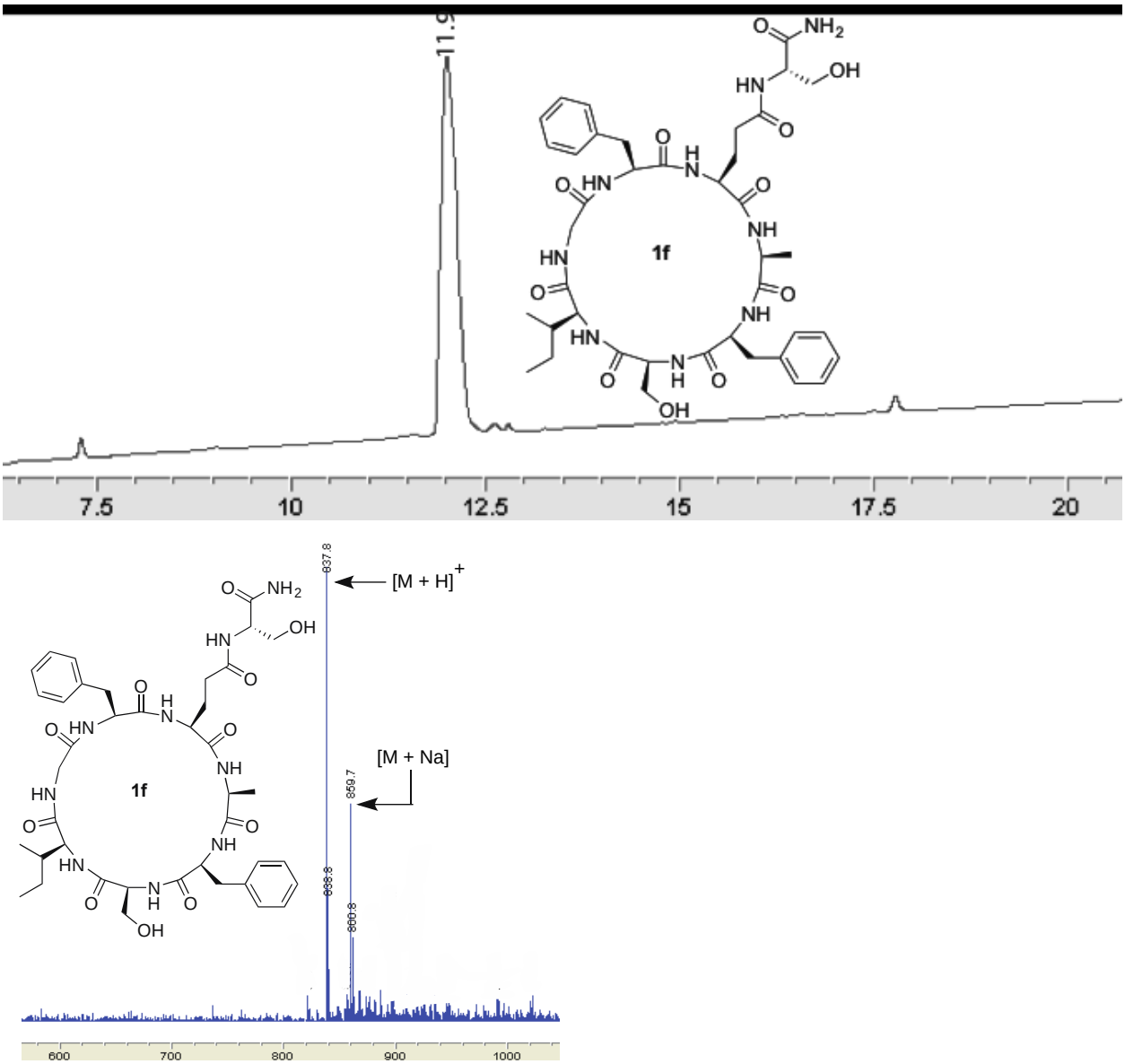


Figure A46 HPLC trace and mass spectrum of peptide **1f**, *cyc*(AFSIGFE)-S-$CONH_2$ after resin cleavage

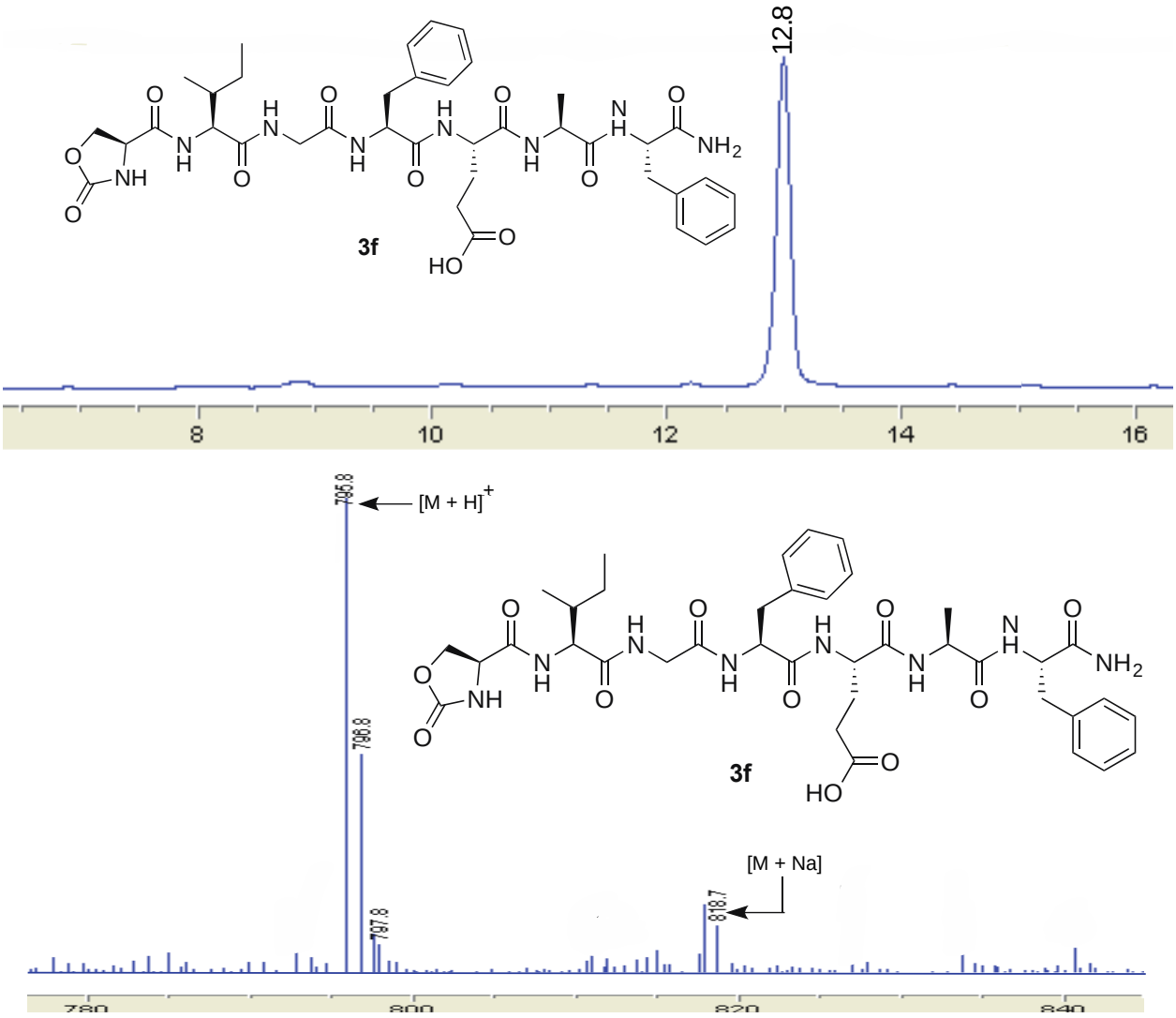


Figure A47 HPLC trace and mass spectrum of peptide **3f**, *Oxd*-IGFEAF-NH_2

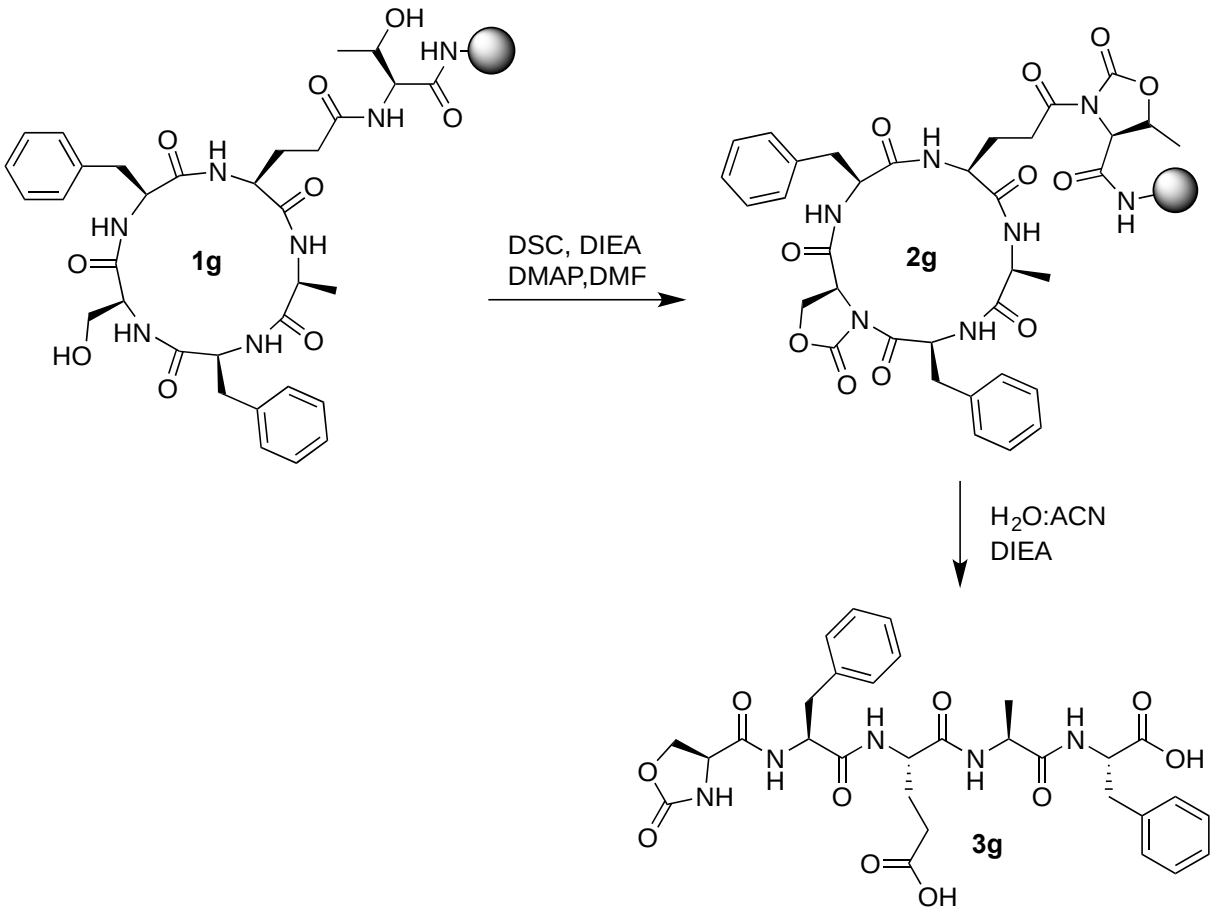


Figure A48 Macrocyclic ring-opening and resin-cleavage of peptide **1g**, *cyc*(AFSFE)-T-*Rink*

***cyc*(Ala-Phe-Ser-Phe-Glu)-Thr-CONH$_2$ (1g).** LCMS: *m/z* 681.90 (calcd $[M+H]^+$ = 682.3), *m/z* 1363.7 (calcd $[2M+H]^+$ = 1364.6), Purity: >95% (HPLC analysis at 220 nm). Retention time: 12.07

***Oxd*-Phe-Glu-Ala-Phe (3g).** LCMS: *m/z* 625.8 (calcd $[M+H]^+$ = 626.2), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.31 min

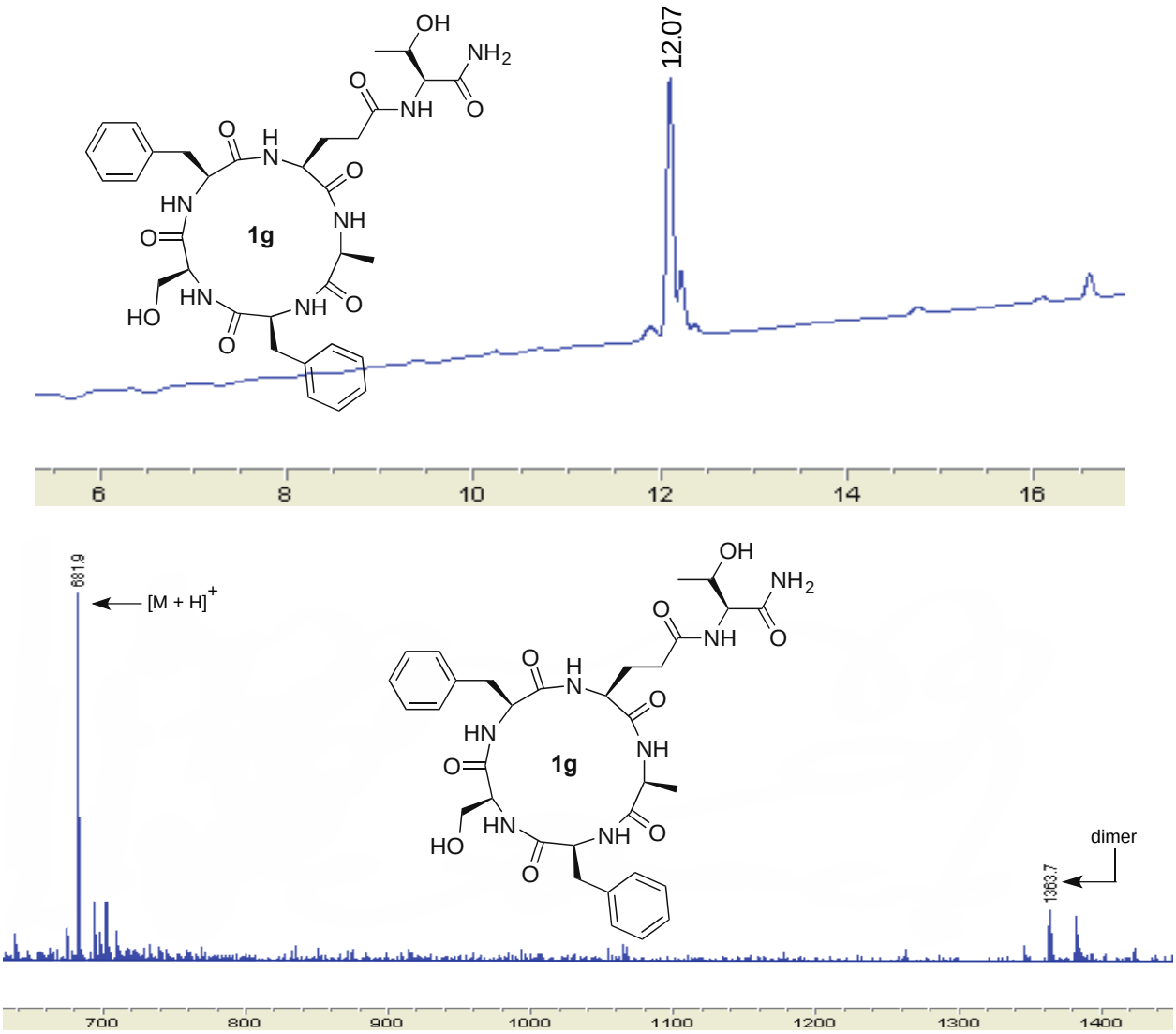


Figure A49 HPLC trace and mass spectrum of peptide **1g**, *cyc*(AFSFE)-T-CONH$_2$ after resin cleavage

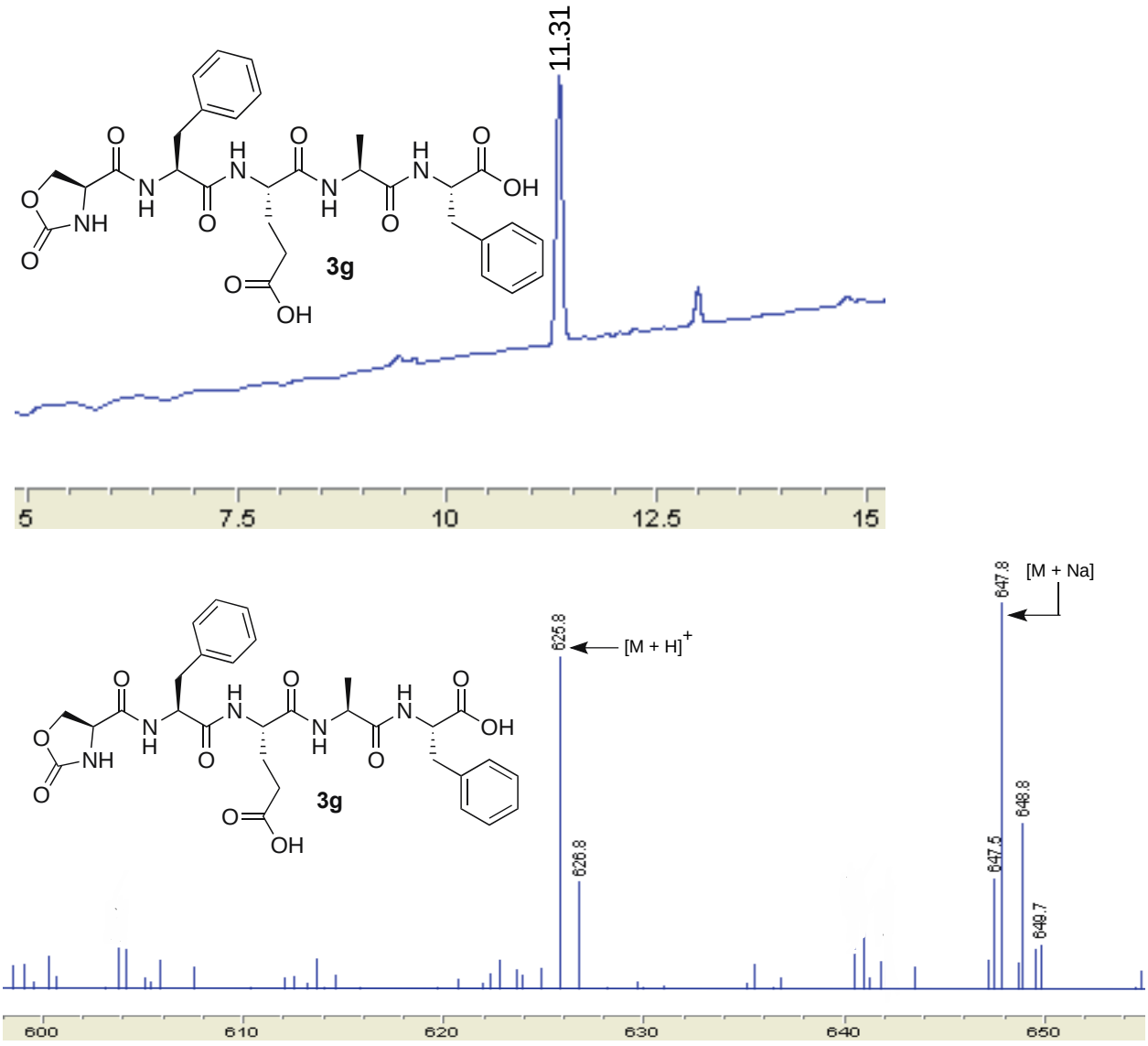


Figure A50 HPLC trace and mass spectrum of peptide **3g**, *Oxd*-FEAF-OH

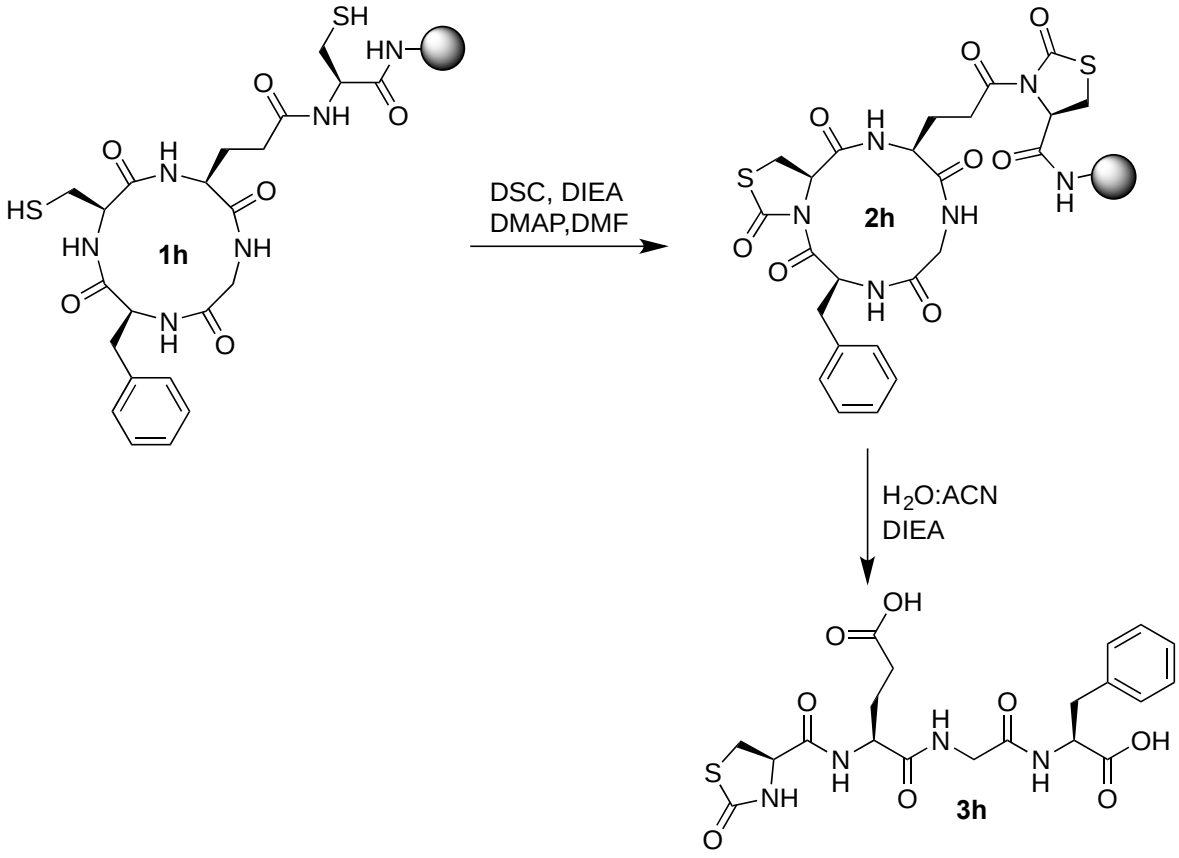


Figure A51 Macrocyclic ring-opening and resin-cleavage of peptide **1h**, *cyc*(GFCE)-C-*Rink* ***cyc*(Gly-Phe-Cys-Glu)-Cys-CONH$_2$ (1h).** LCMS: *m/z* 539.0 (calcd $[M+H]^+$ = 539.1), *m/z* 1076.0, (calcd $[dimer]^+$ = 1076.2) *m/z* 1098.5 (calcd $[dimer+Na]^+$ = 1099.2), Purity: >95% (HPLC analysis at 220 nm). Retention time: 12.8

***Thz*-Glu-Gly-Phe (3h).** LCMS: *m/z* 480.9 (calcd $[M+H]^+$ = 481.1), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.1 min

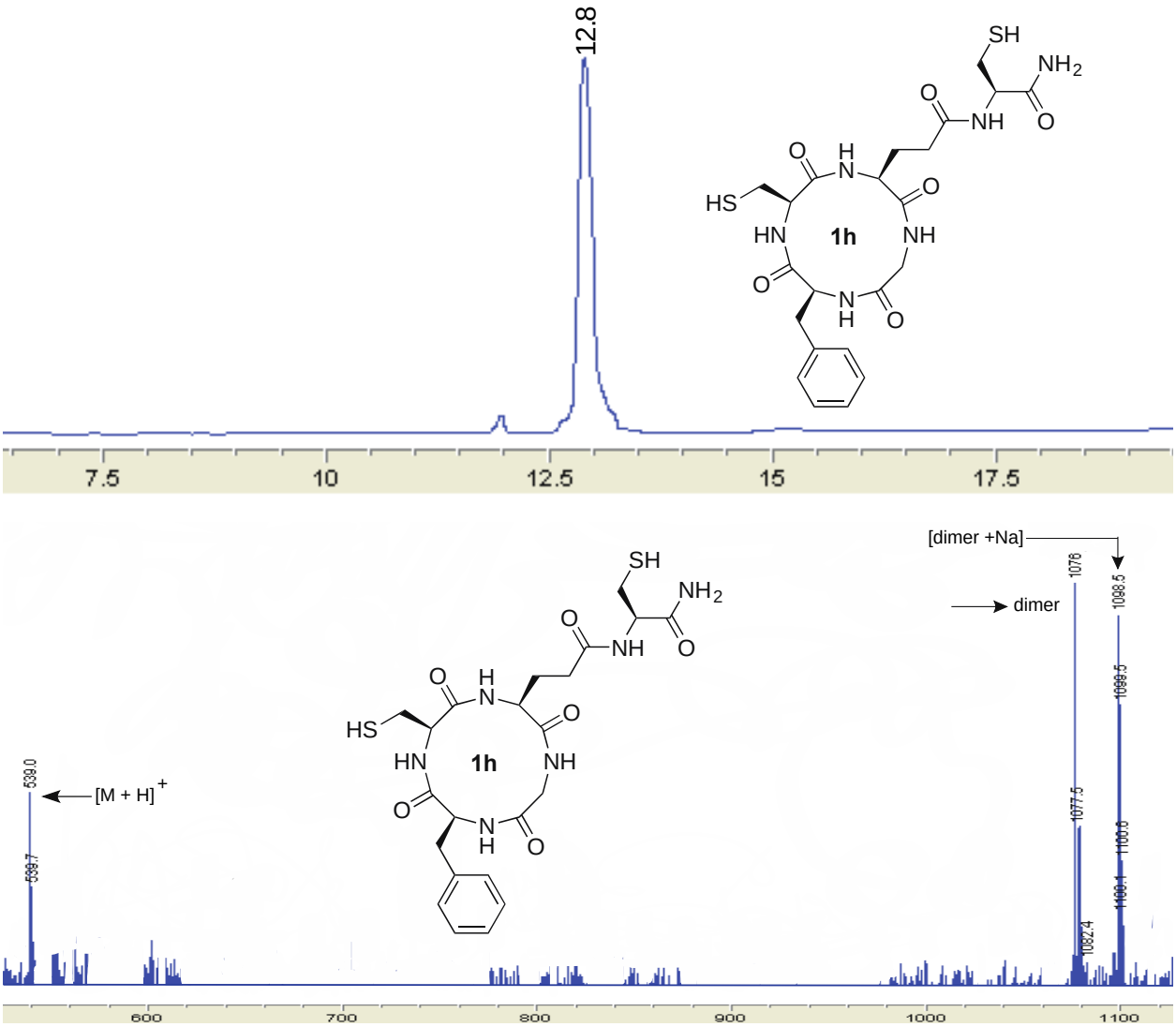


Figure A52 HPLC trace and mass spectrum of peptide **1h**, *cyc*(GFCE)-S-$CONH_2$ after resin cleavage

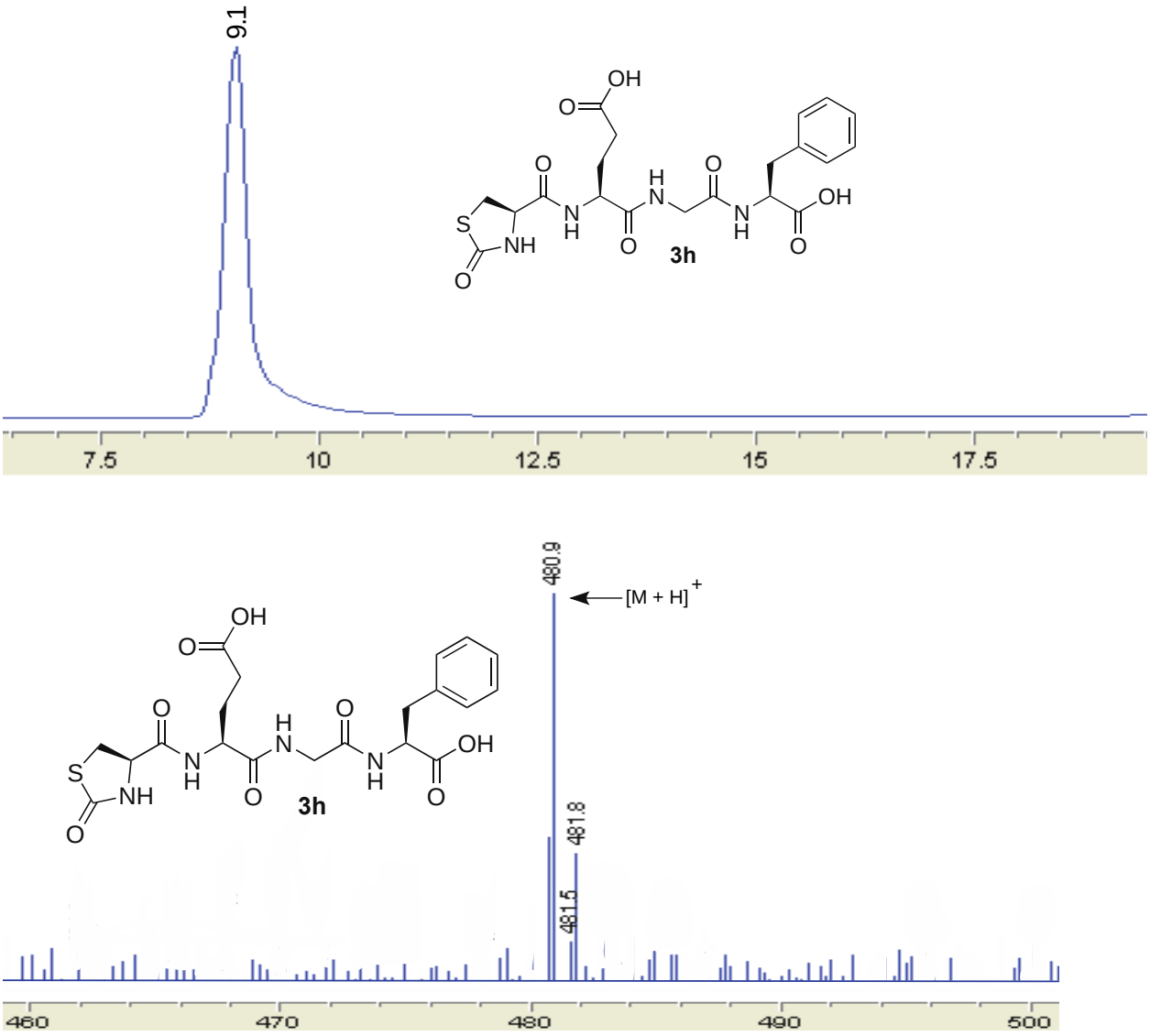


Figure A53 HPLC trace and mass spectrum of peptide **3h**, *Thz*-EGF-OH

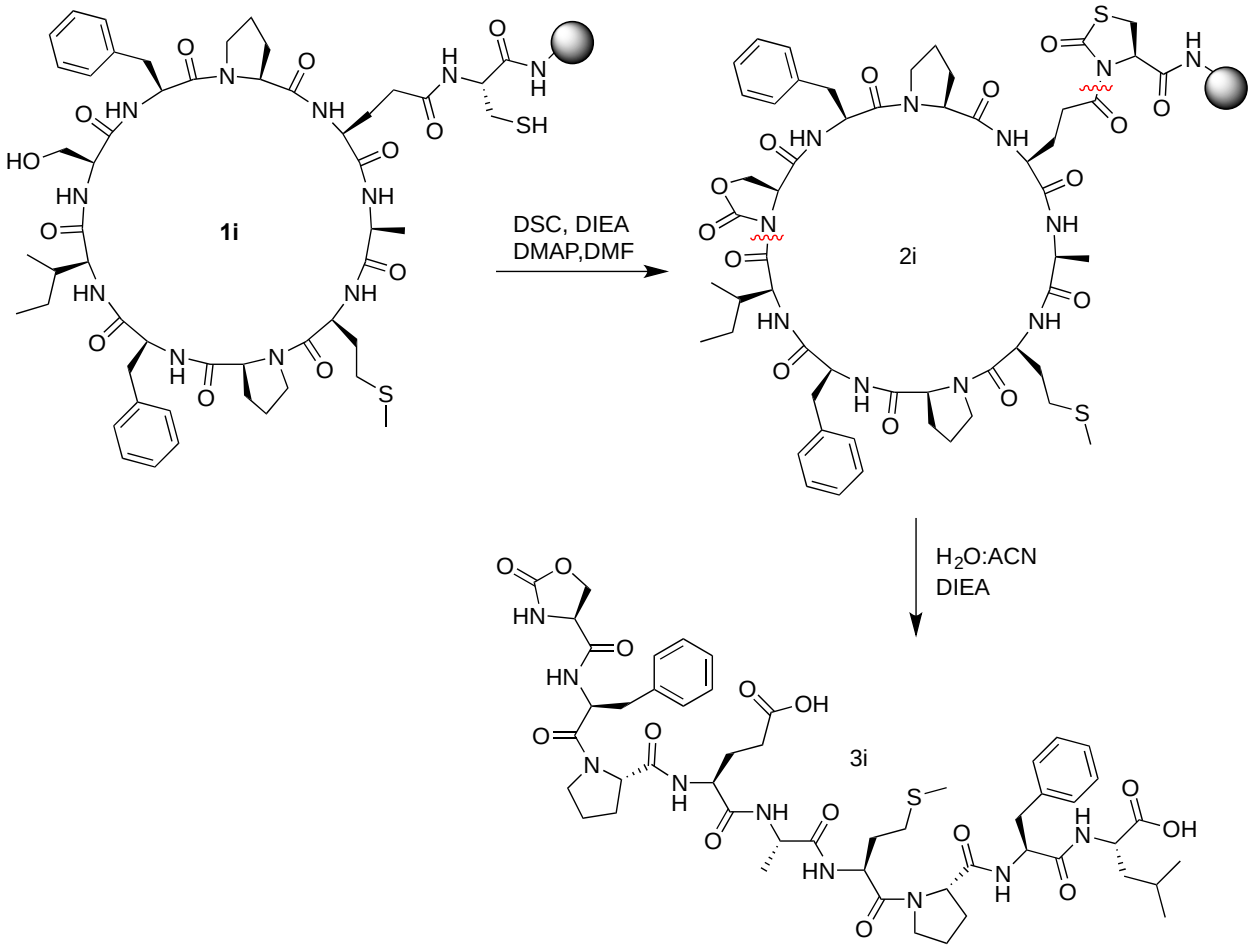


Figure A54 Macrocyclic ring-opening and resin-cleavage of peptide **1i**, *cyc*(AMPFISFPE)-C-*Rink*

***cyc*(Ala-Met-Pro-Phe-Ile-Ser-Phe-Pro-Glu)-Cys-CONH$_2$ (1i).** LCMS: *m/z* 1121.7 (calcd $[M+H]^+$ = 1122.5), *m/z* 1143.7, (calcd $[M+Na]^+$ = 1144.5) 561.5 (calcd $[M+Na]^+$ = 561.7), Purity: >95% (HPLC analysis at 220 nm). Retention time: 15.1 min

***Oxd*-Phe-Pro-Glu-Ala-Met-Pro-Phe-Ile (3i).** LCMS: *m/z* 1063.6 (calcd $[M+H]^+$ = 1064.4), *m/z* 1087.8, (calcd $[M+Na]^+$ = 1086.4) *m/z* 532.5 (calcd $[M+Na]^+$ = 532.7), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.5 min

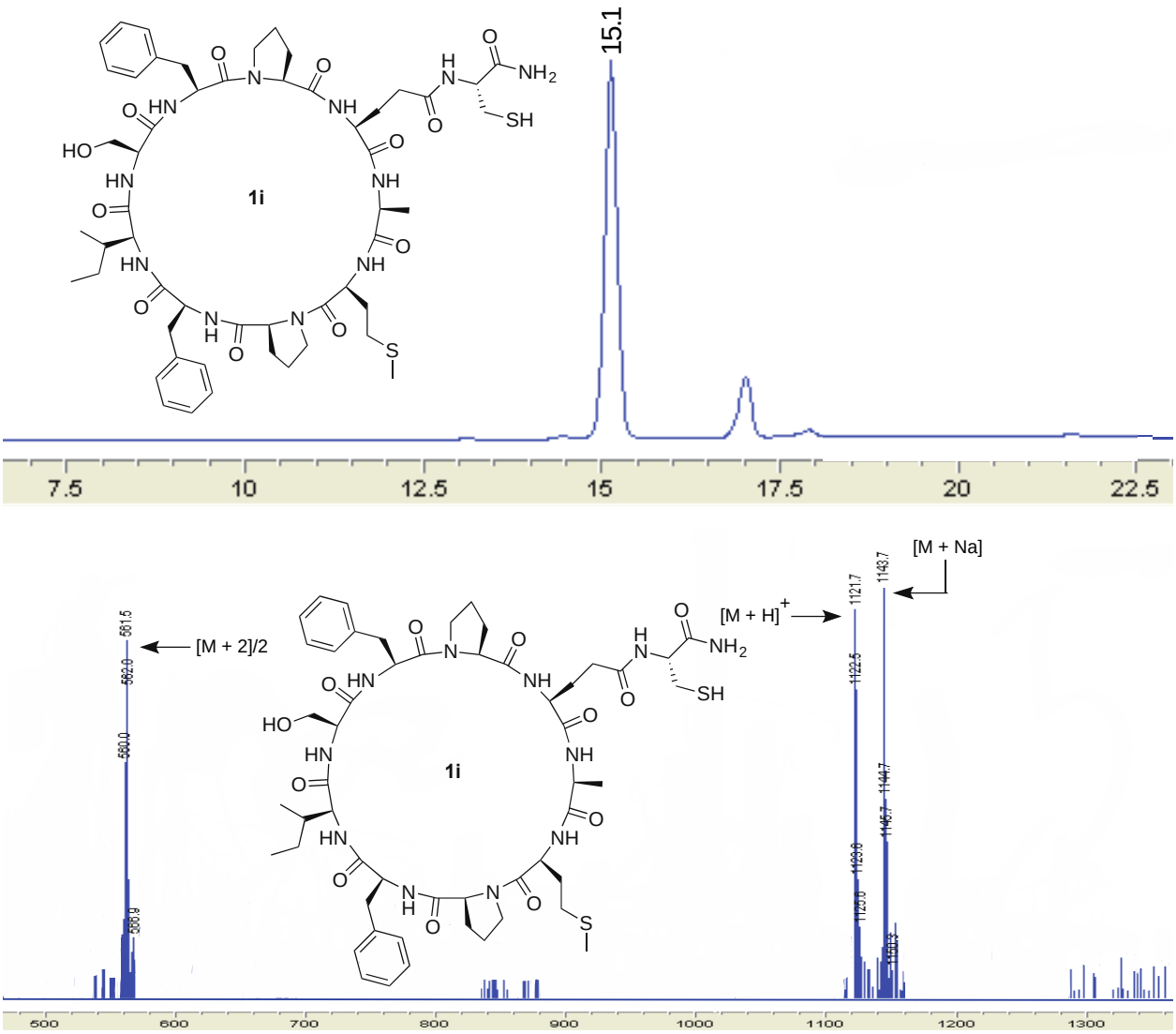


Figure A55 HPLC trace and mass spectrum of peptide **1i**, *cyc*(AMPFISFPE)-C-CONH$_2$ after resin cleavage

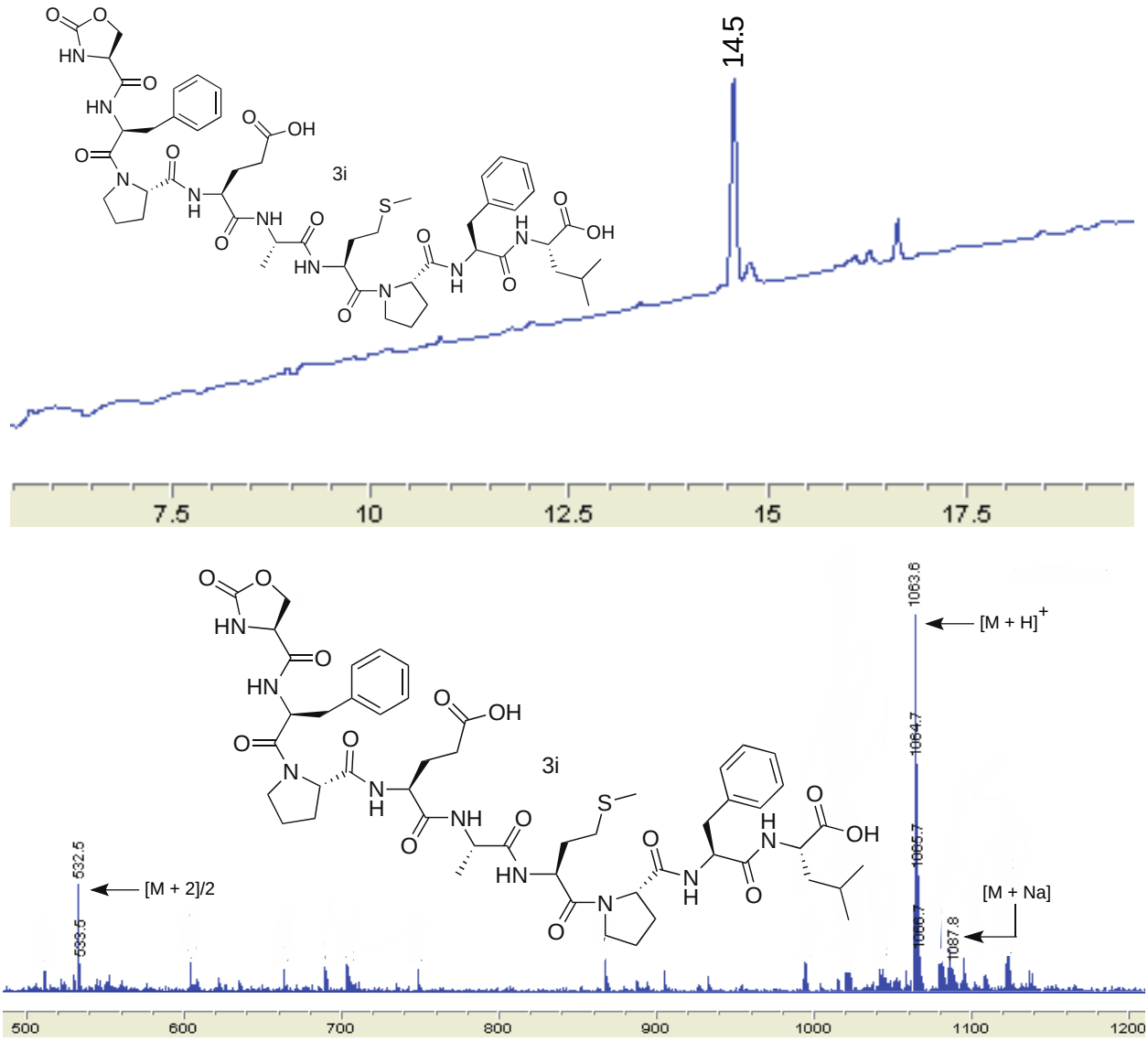


Figure A56 HPLC trace and mass spectrum of peptide **3i**, *Oxd*-FPEAMPFI-OH

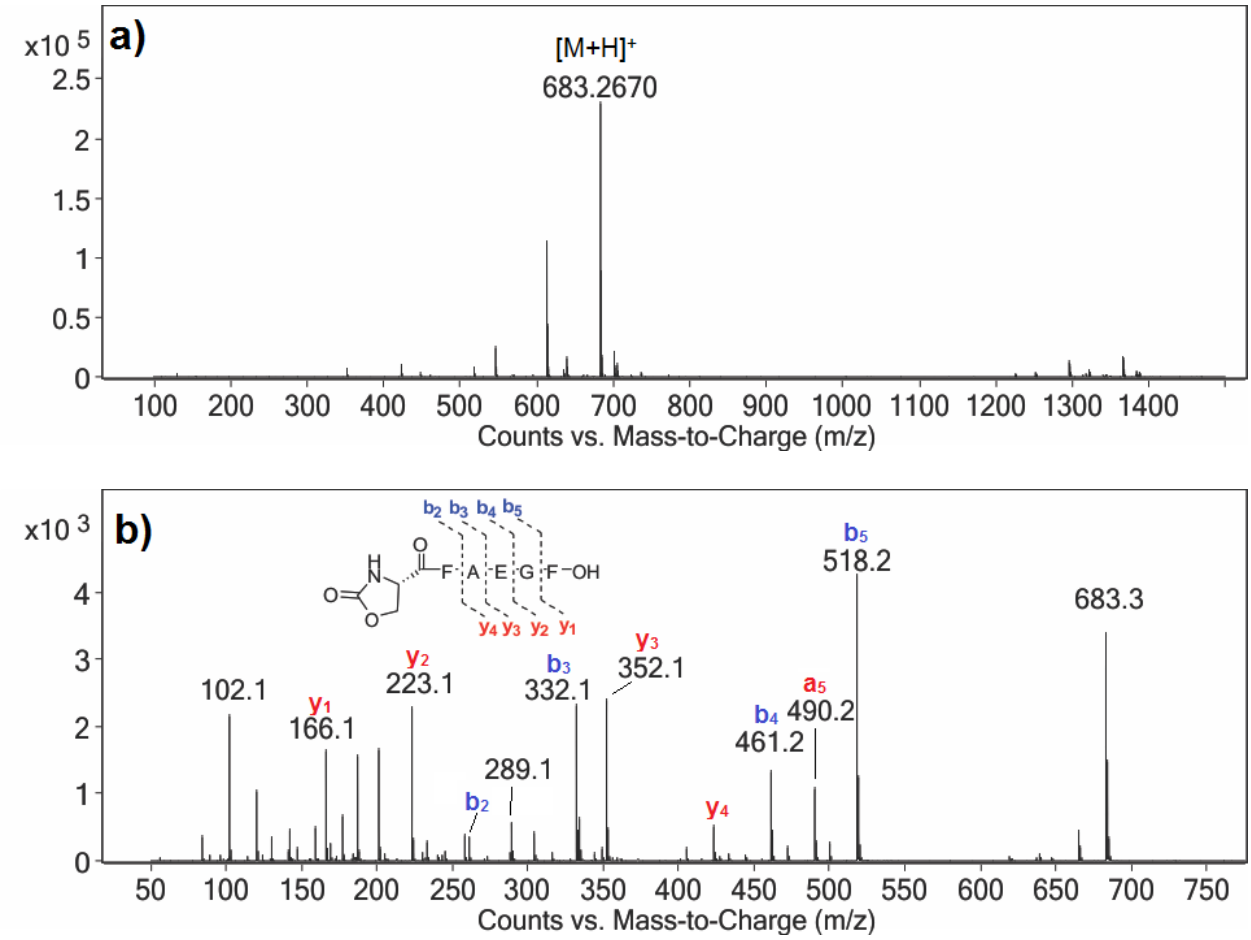


Figure A57 LC-IM-MS/MS data for linear peptide **3a**, *Oxd*-FAEGF-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 683.2671, Found $[M+H]^+$ = 683.2670

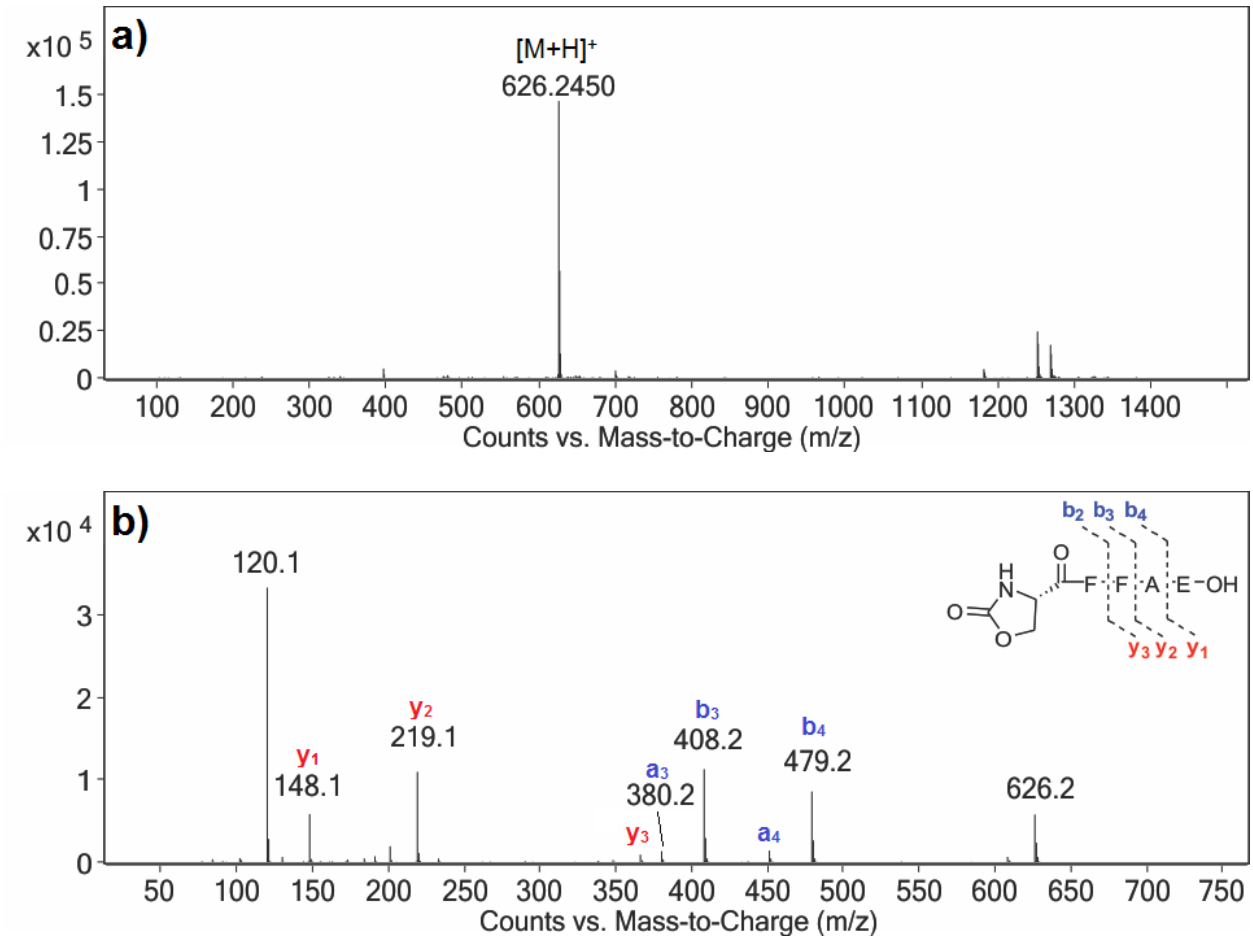


Figure A58 LC-IM-MS/MS data for linear peptide **3b**, *Oxd*-FFAE-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 626.2457, Found $[M+H]^+$ = 626.2450

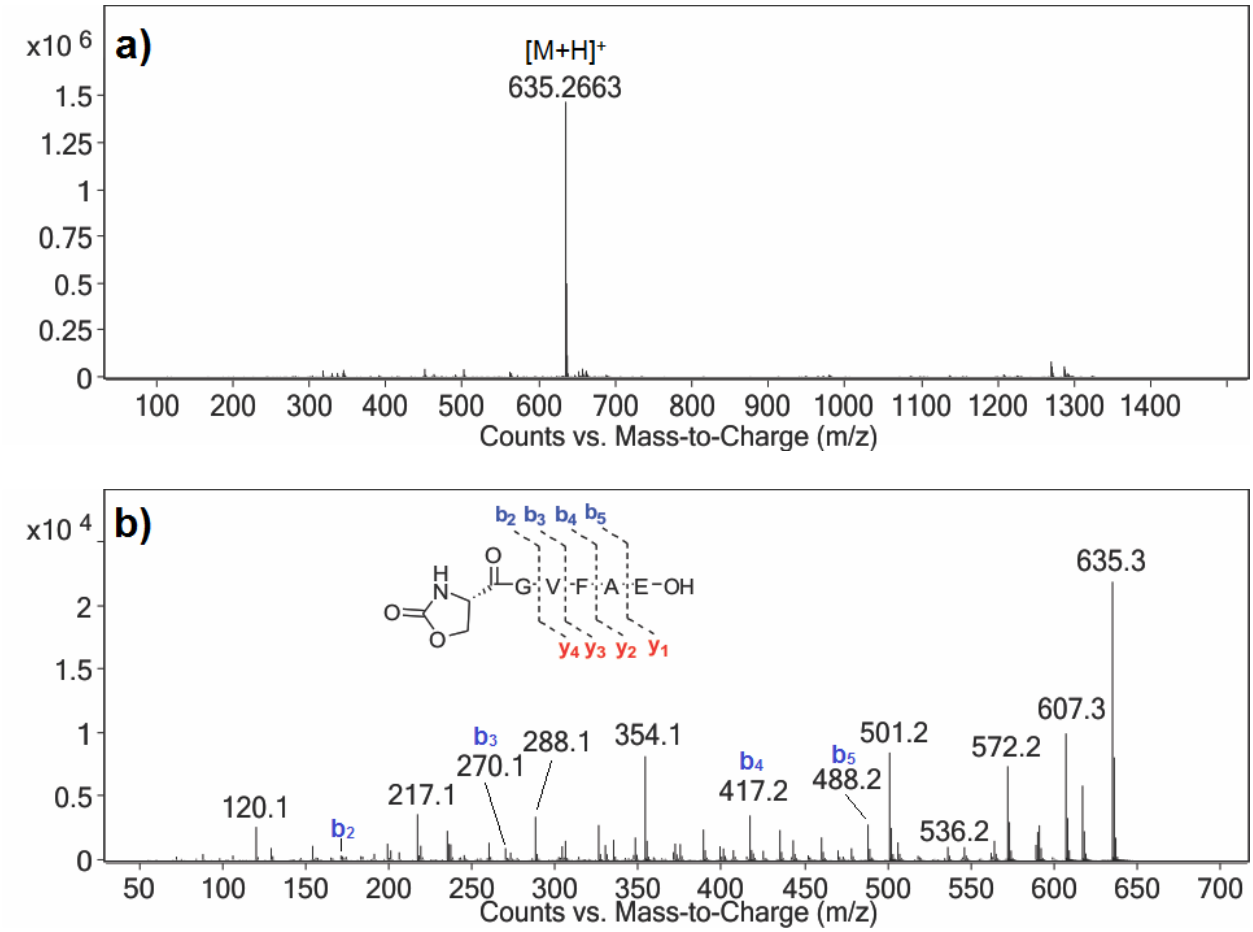


Figure A59 LC-IM-MS/MS data for linear peptide **3c**, *Oxd*-GVFAE-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 635.2671, Found $[M+H]^+$ = 635.2663

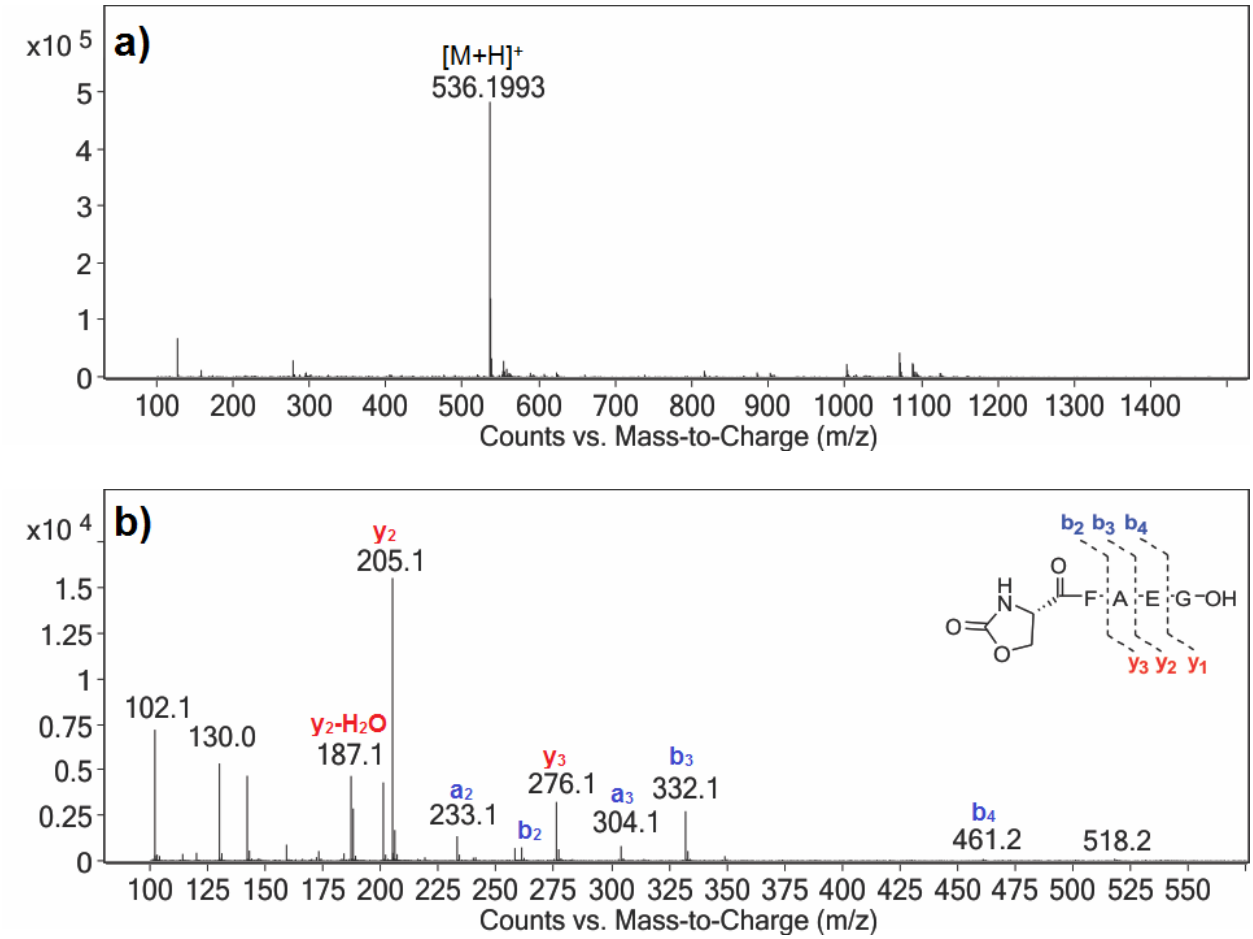


Figure A60 LC-IM-MS/MS data for linear peptide **3d**, *Oxd*-FAEG-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 536.1987, Found $[M+H]^+$ = 536.1993

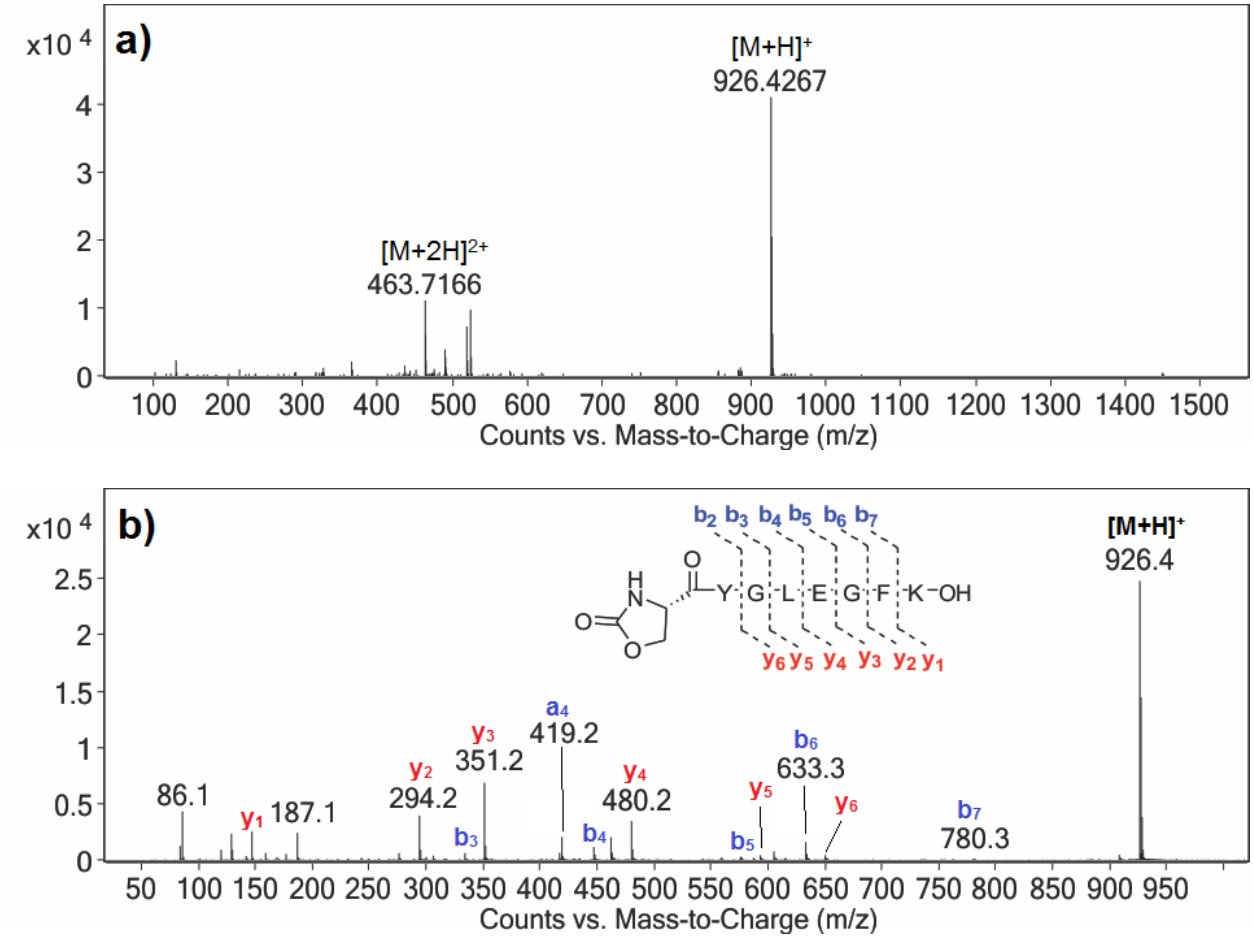


Figure A61 LC-IM-MS/MS data for linear peptide **3e**, *Oxd*-YGLEGFK-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 926.4254, Found $[M+H]^+$ = 926.4267

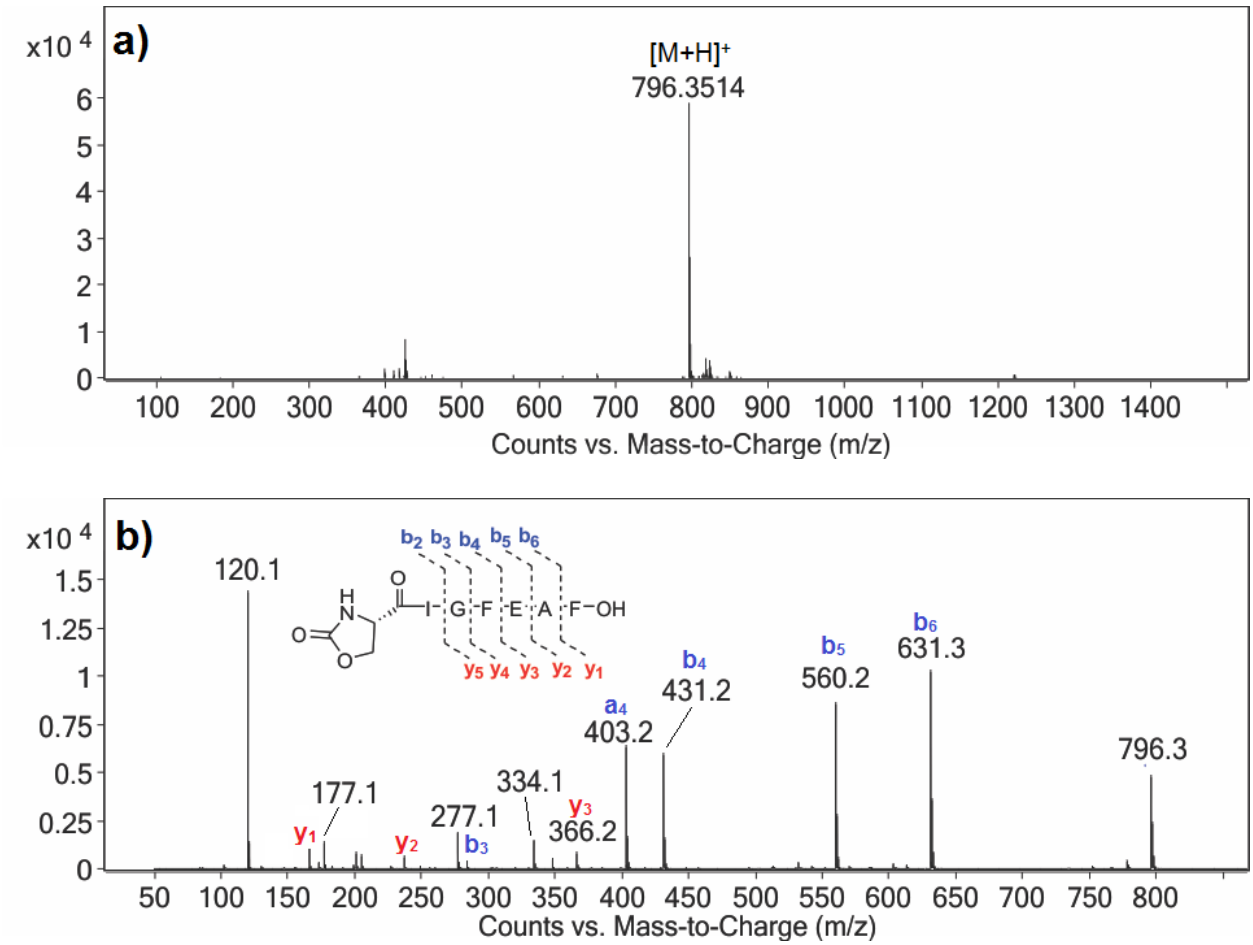


Figure A62 LC-IM-MS/MS data for linear peptide **3f**, *Oxd*-IGFEAF-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 796.3512, Found $[M+H]^+$ = 796.3514

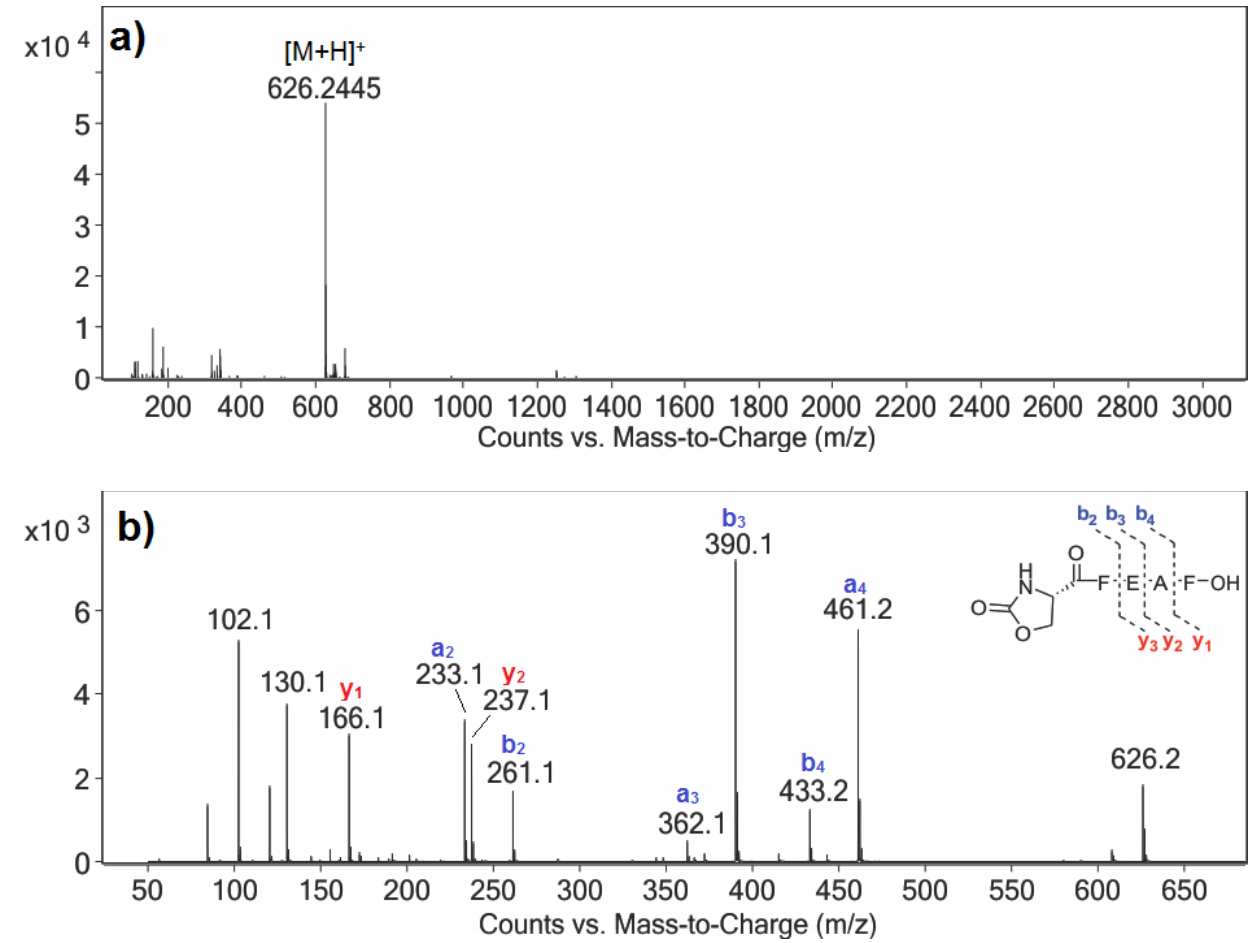


Figure A63 LC-IM-MS/MS data for linear peptide **3g**, *Oxd*-FEAF-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 626.2457, Found $[M+H]^+$ = 626.2445

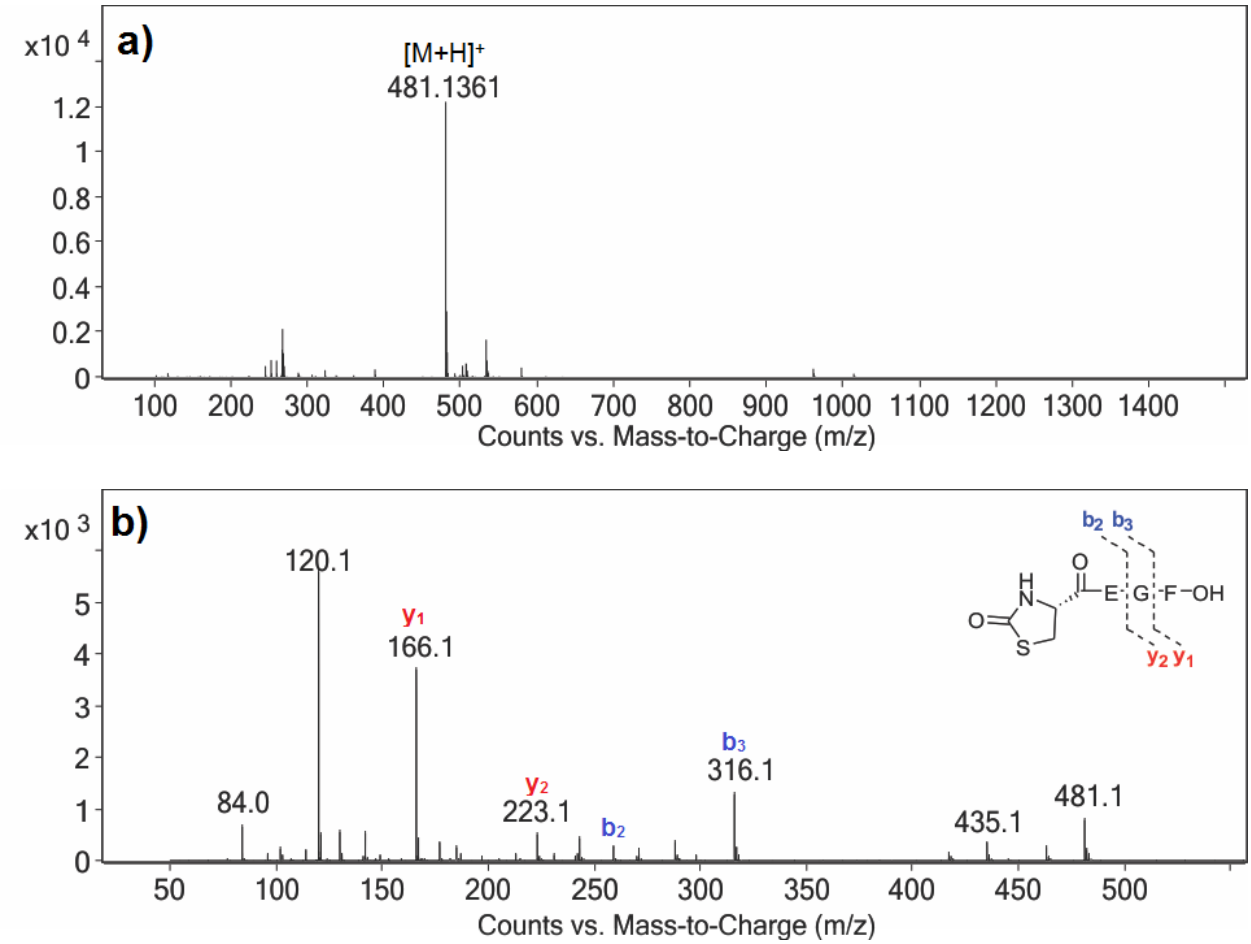


Figure A64 LC-IM-MS/MS data for linear peptide **3h**, *Thz*-EGF-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 481.1388, Found $[M+H]^+$ = 481.1361

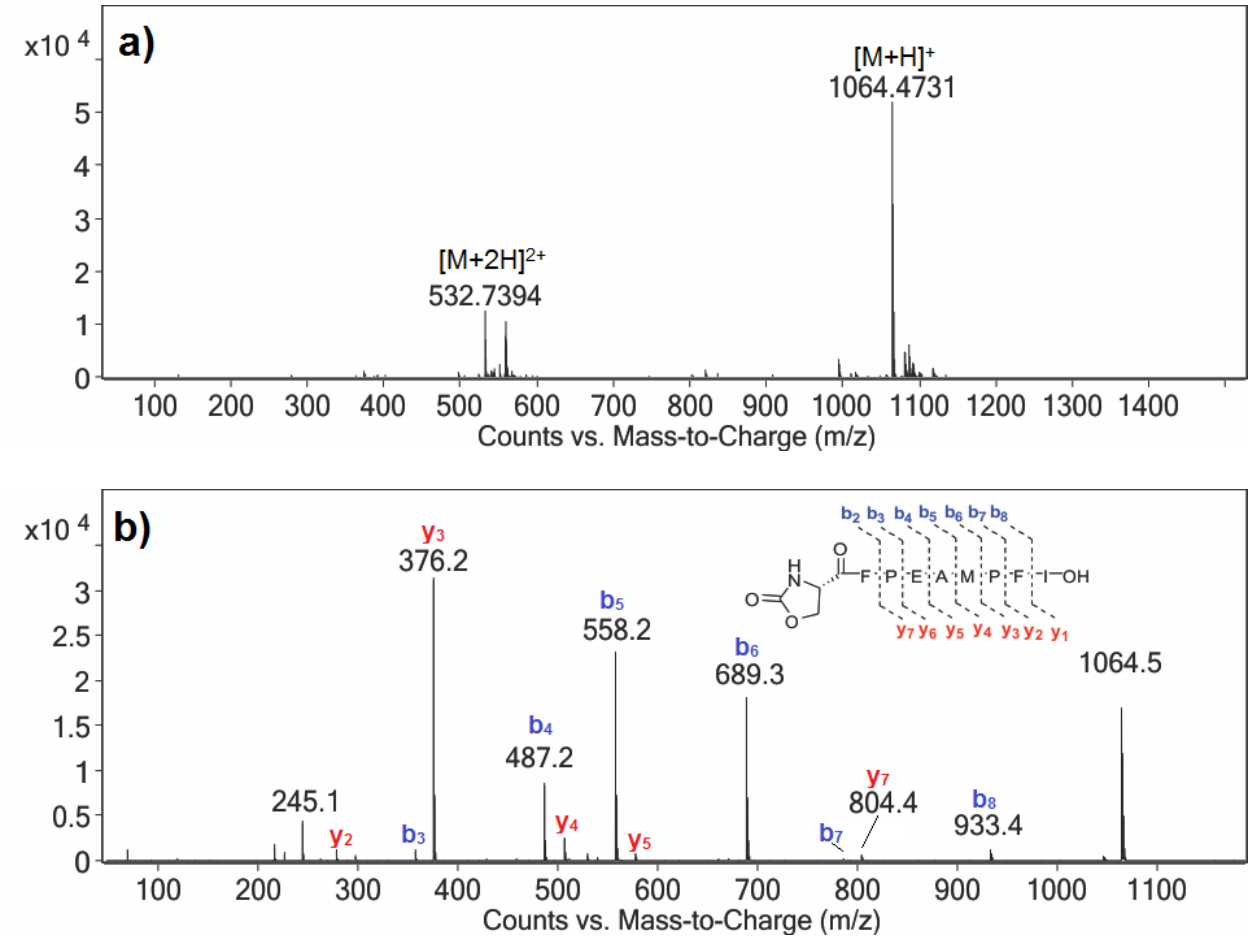


Figure A65 LC-IM-MS/MS data for linear peptide **3i**, *Oxd*-FPEAMPFI-OH. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time
Calcd $[M+H]^+$ = 1064.4757, Found $[M+H]^+$ = 1064.4731

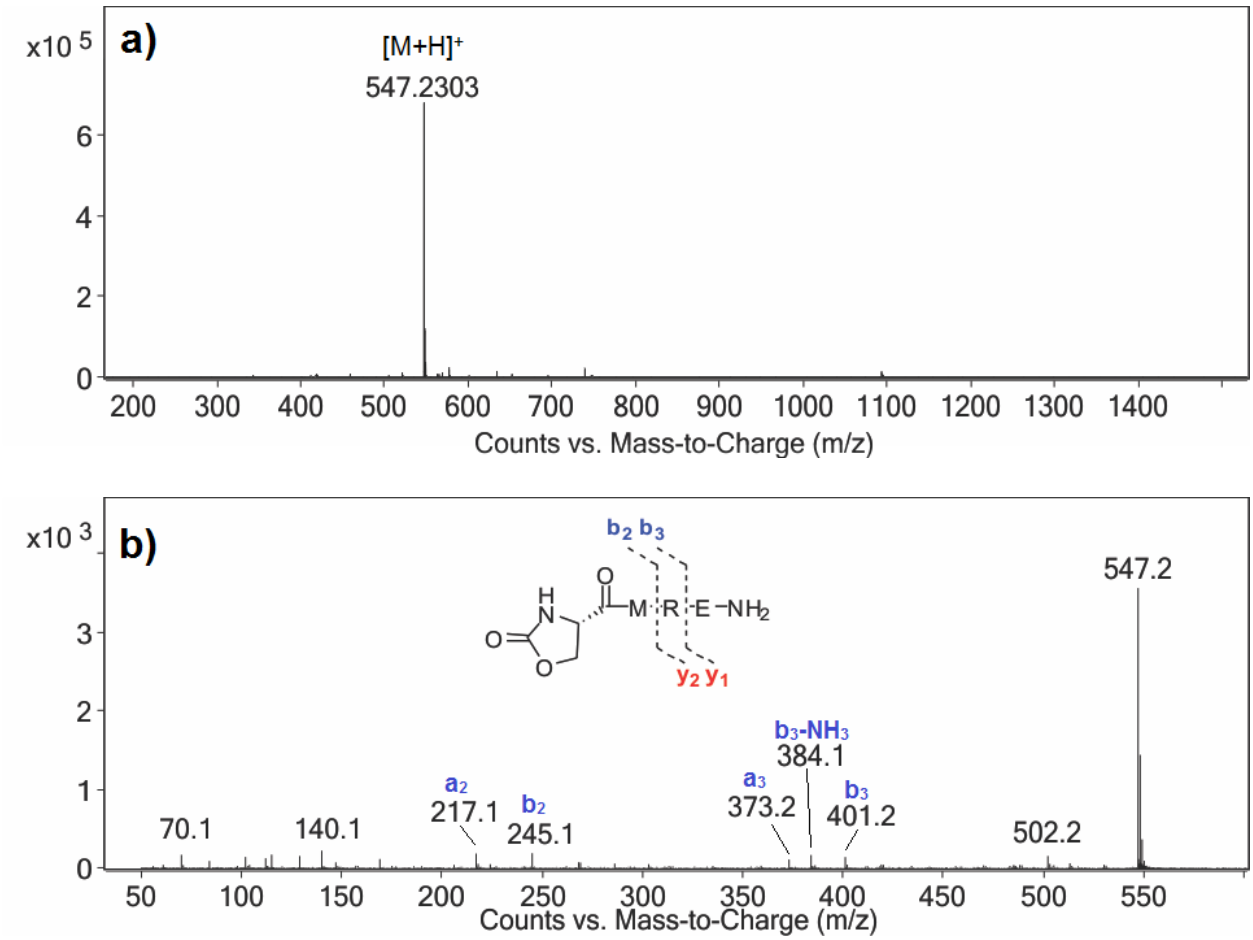


Figure A66 LC-IM-MS/MS data for linear peptide, *Oxd*-MRE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 547.2293, Found $[M+H]^+$ = 547.2303

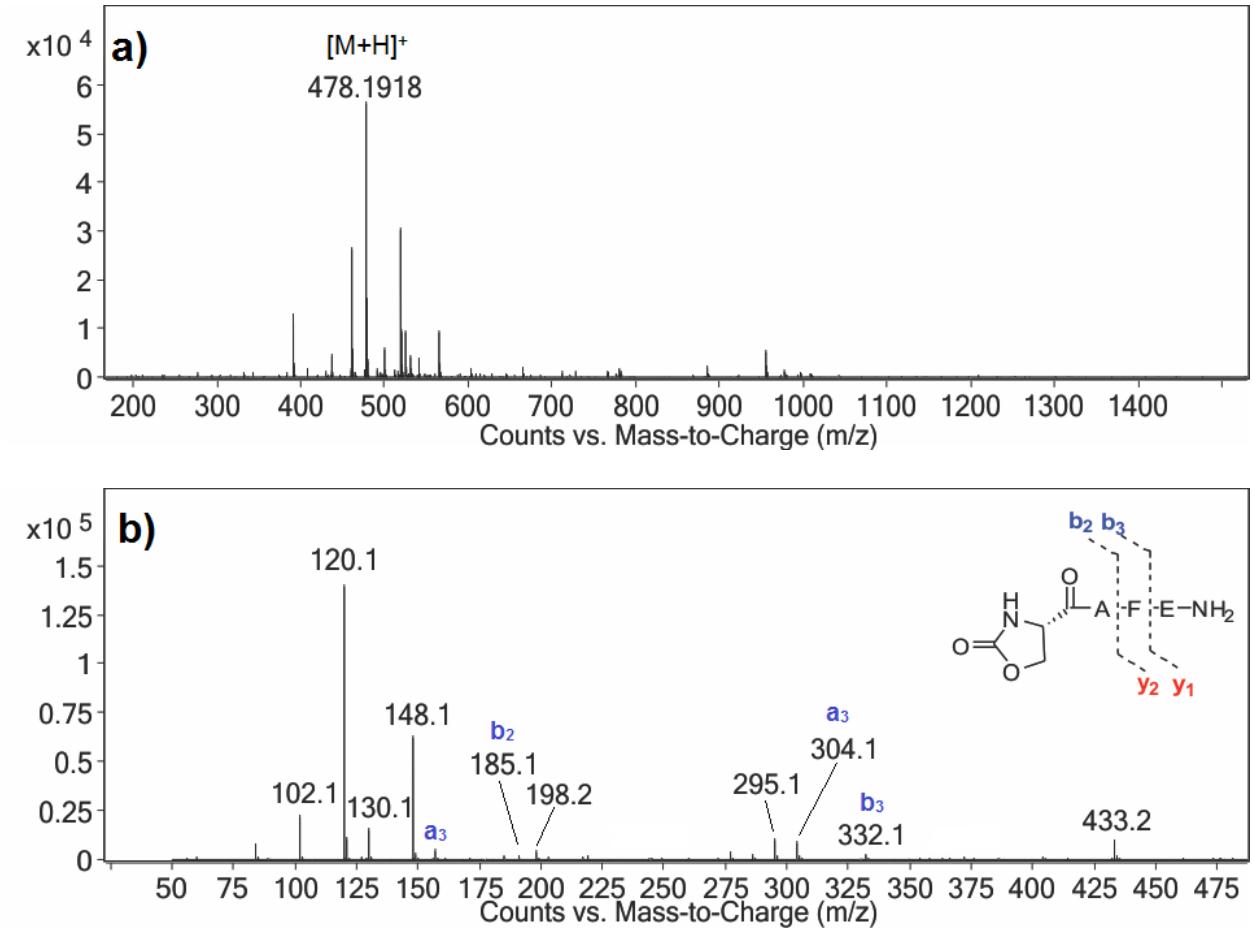


Figure A67 LC-IM-MS/MS data for linear peptide, *Oxd*-AFE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time Calcd $[M+H]^+$ = 478.1932, Found $[M+H]^+$ = 478.1918

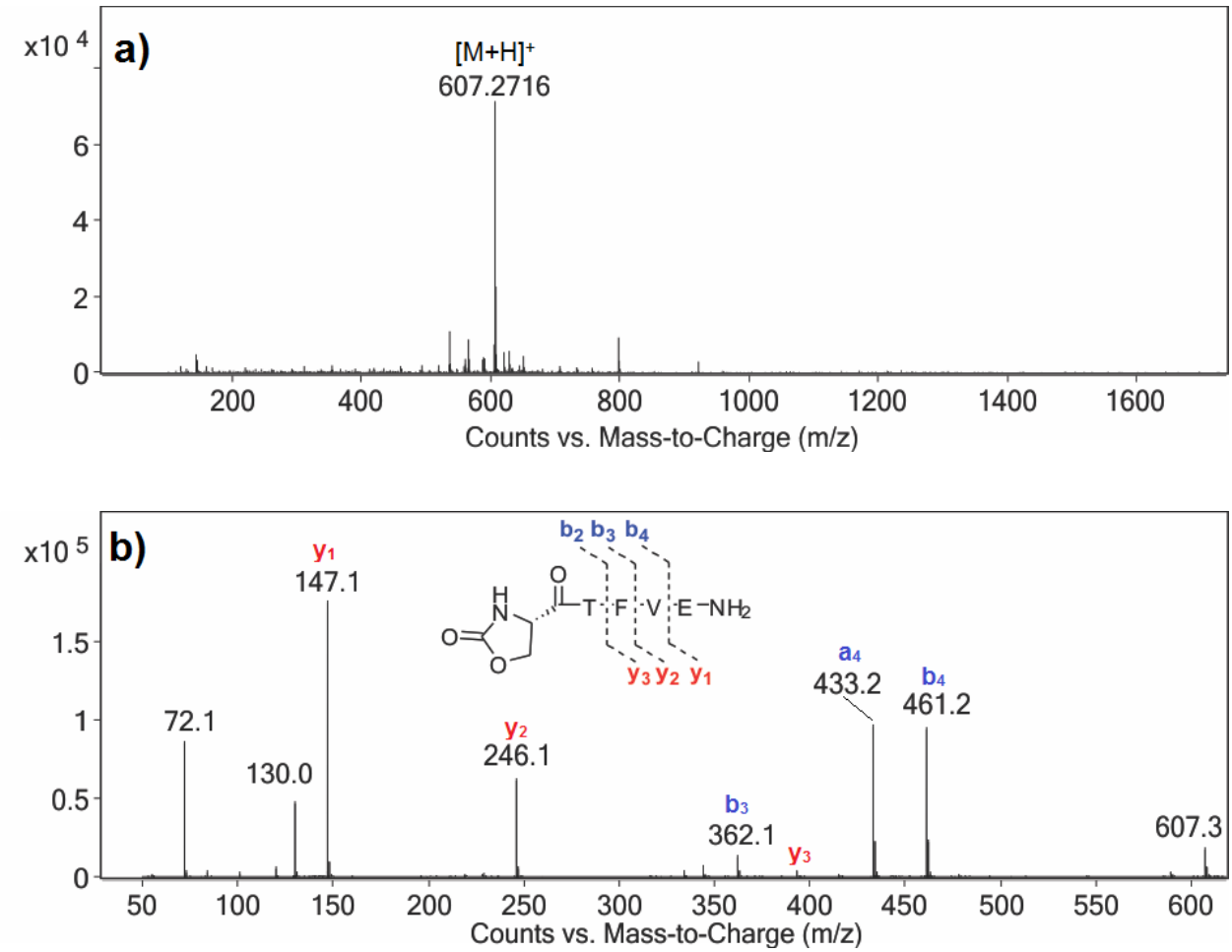


Figure A68 LC-IM-MS/MS data for linear peptide *Oxd*-TFVE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 607.2722, Found $[M+H]^+$ = 607.2716

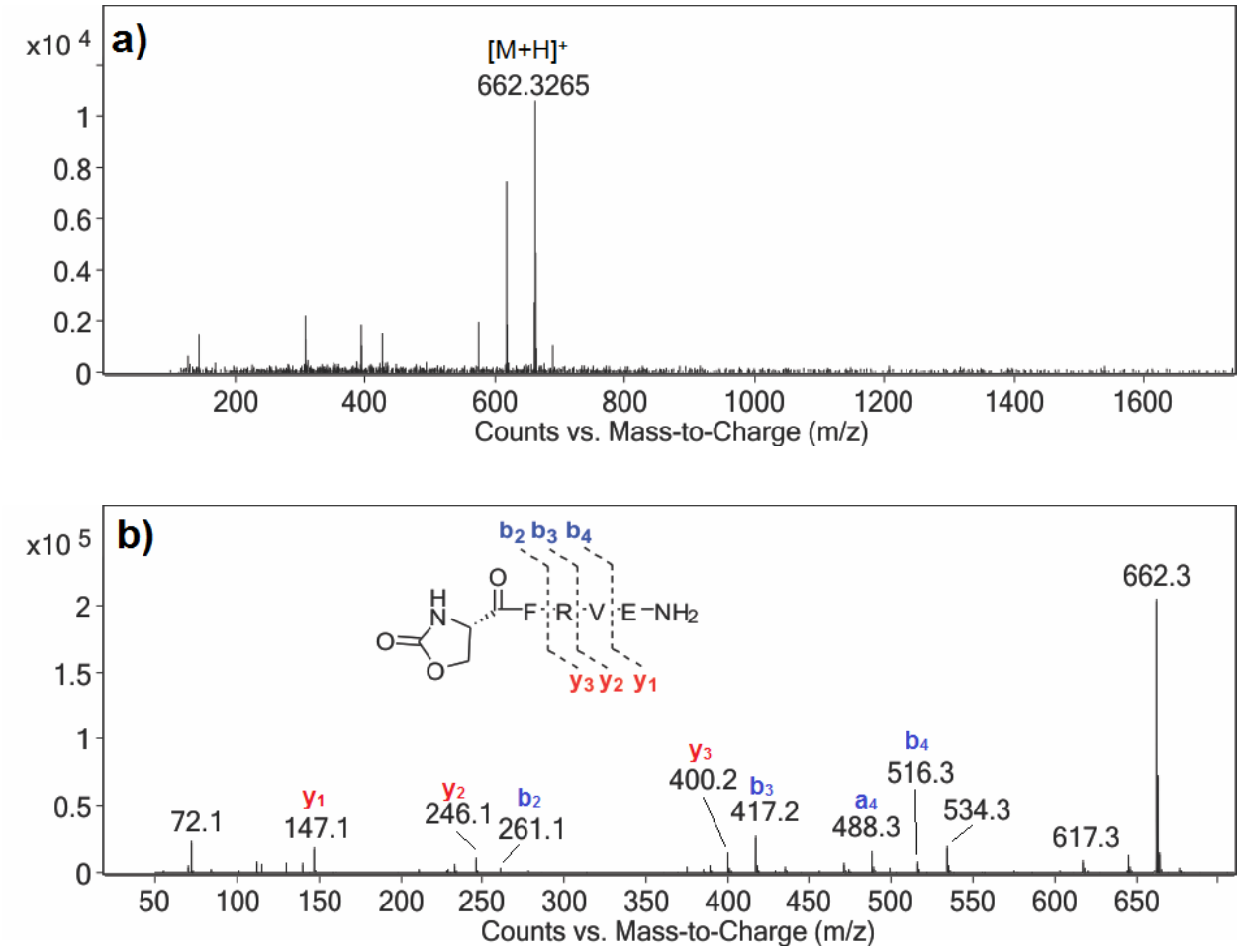


Figure A69 LC-IM-MS/MS data for linear peptide *Oxd*-FRVE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 662.3257, Found $[M+H]^+$ = 662.3265

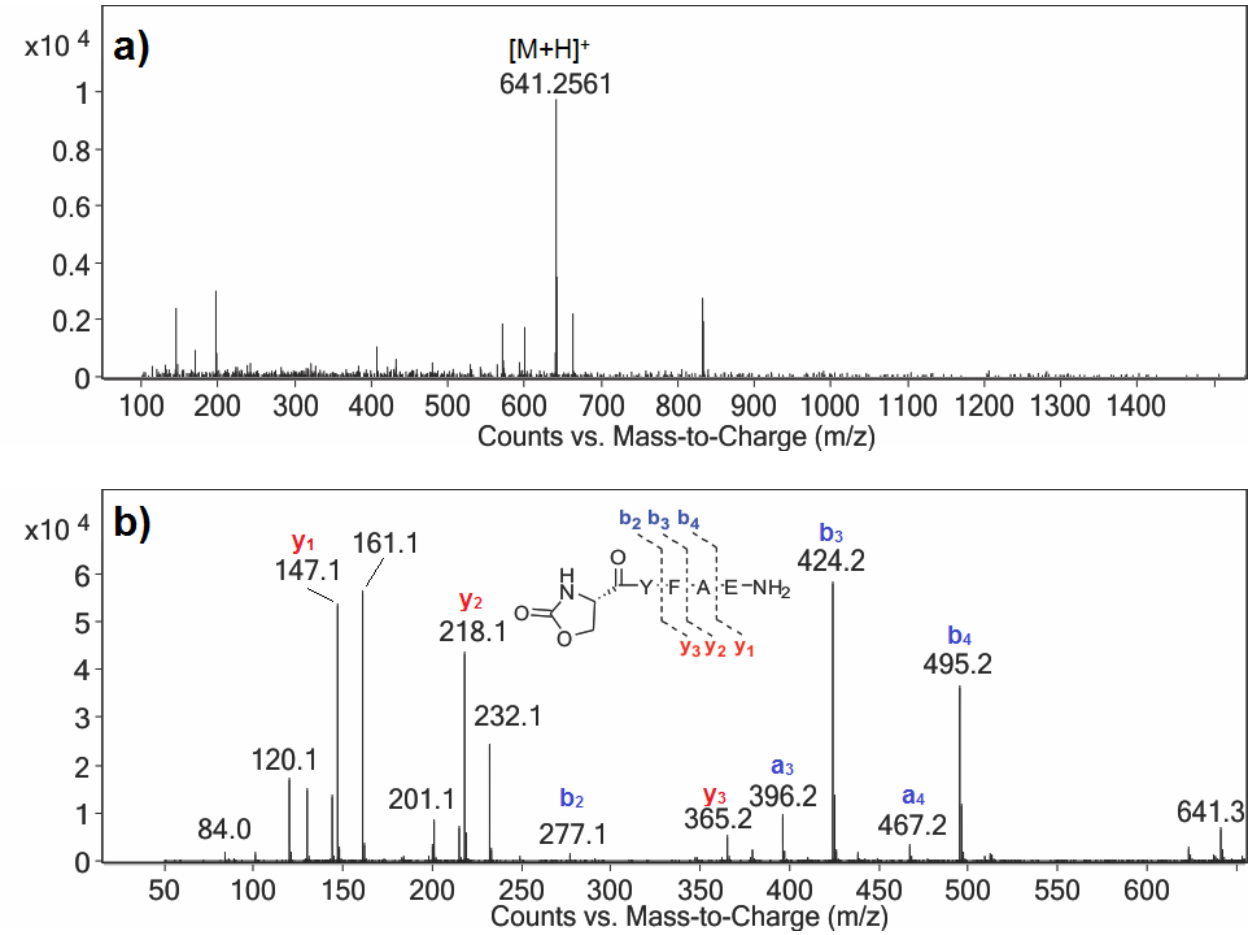


Figure A70 LC-IM-MS/MS data for linear peptide *Oxd*-YFAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 641.2566, Found $[M+H]^+$ = 641.2561

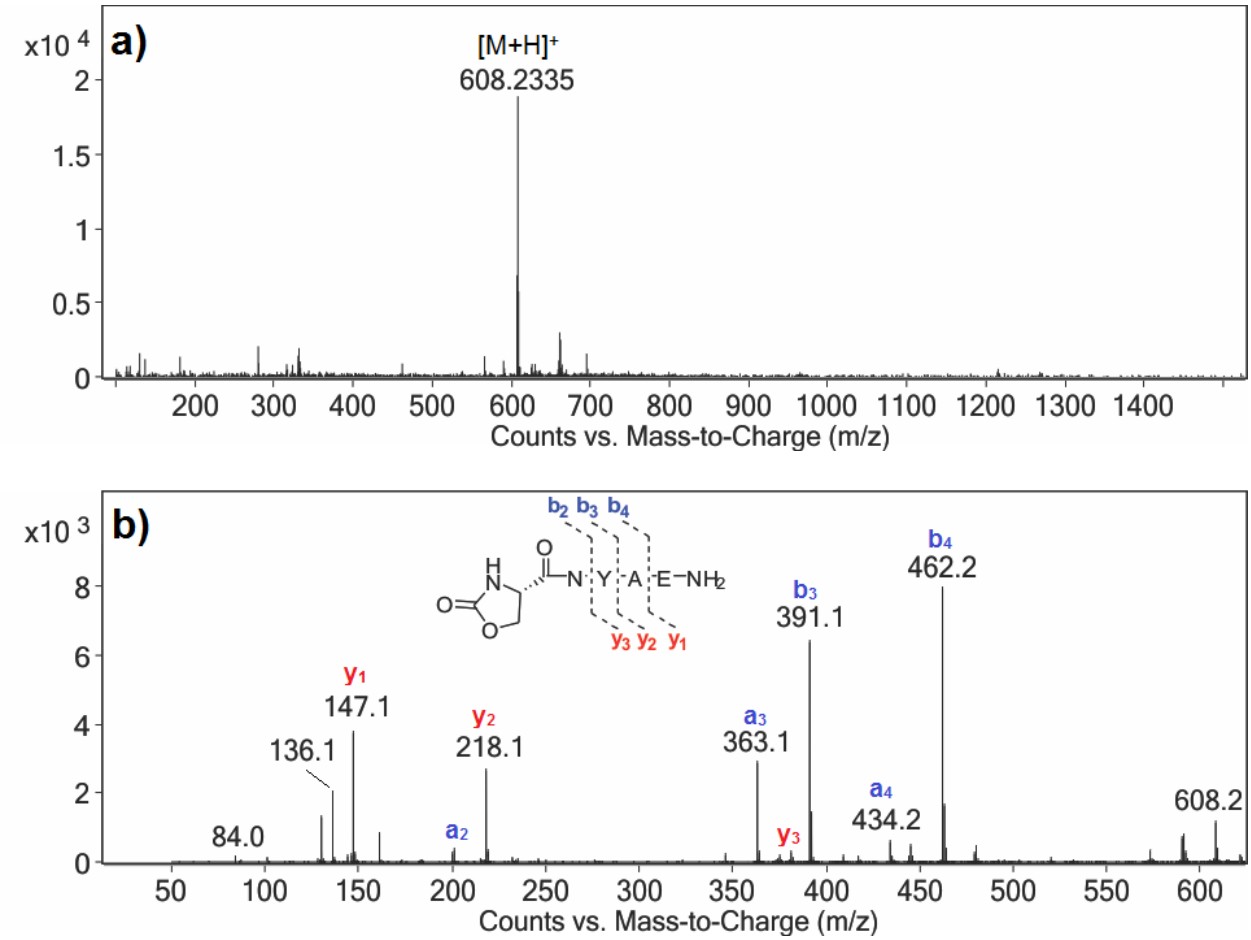


Figure A71 LC-IM-MS/MS data for linear peptide *Oxd*-NYAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 608.2311, Found $[M+H]^+$ = 608.2335

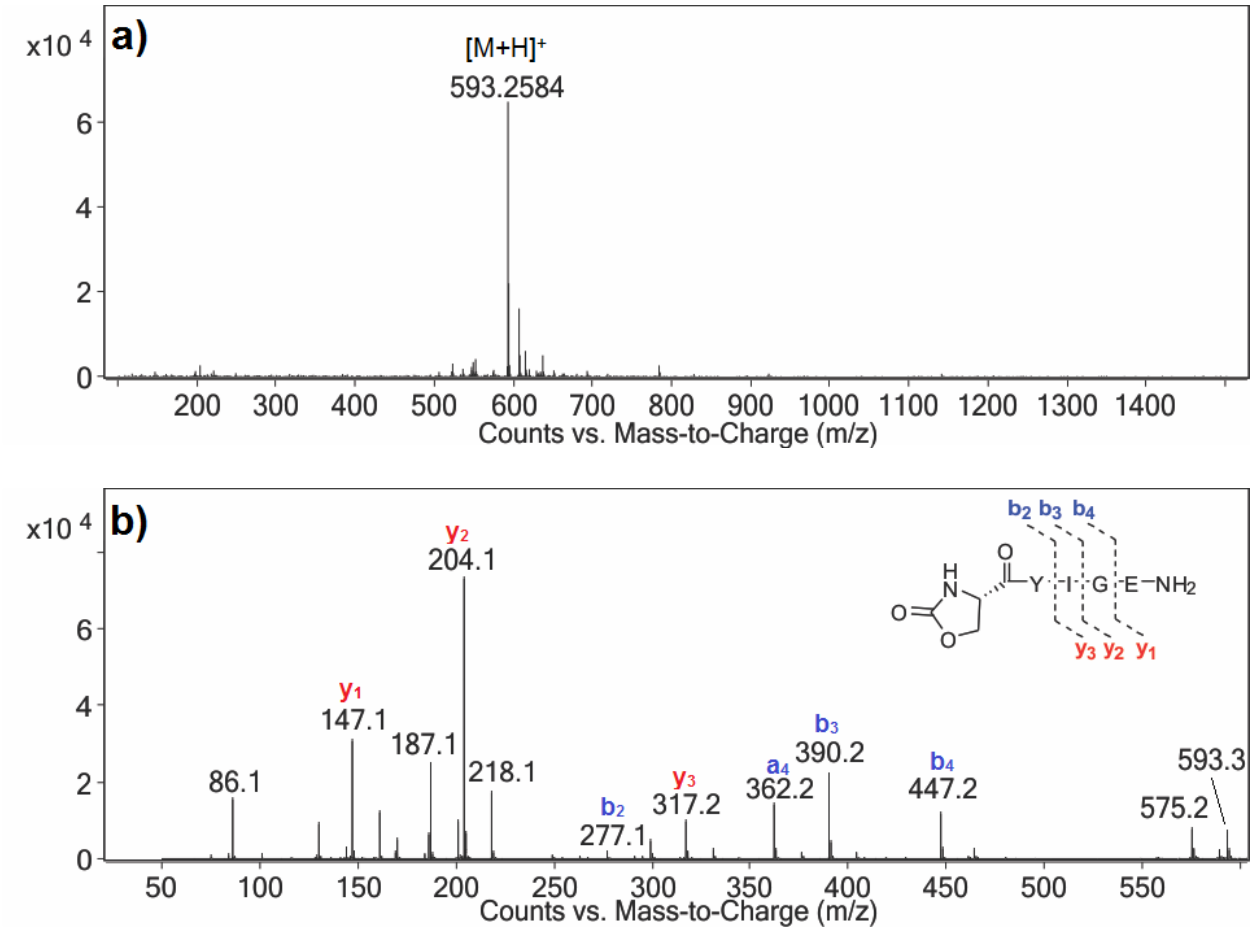


Figure A72 LC-IM-MS/MS data for linear peptide *Oxd*-YIGE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 593.2566, Found $[M+H]^+$ = 593.2584

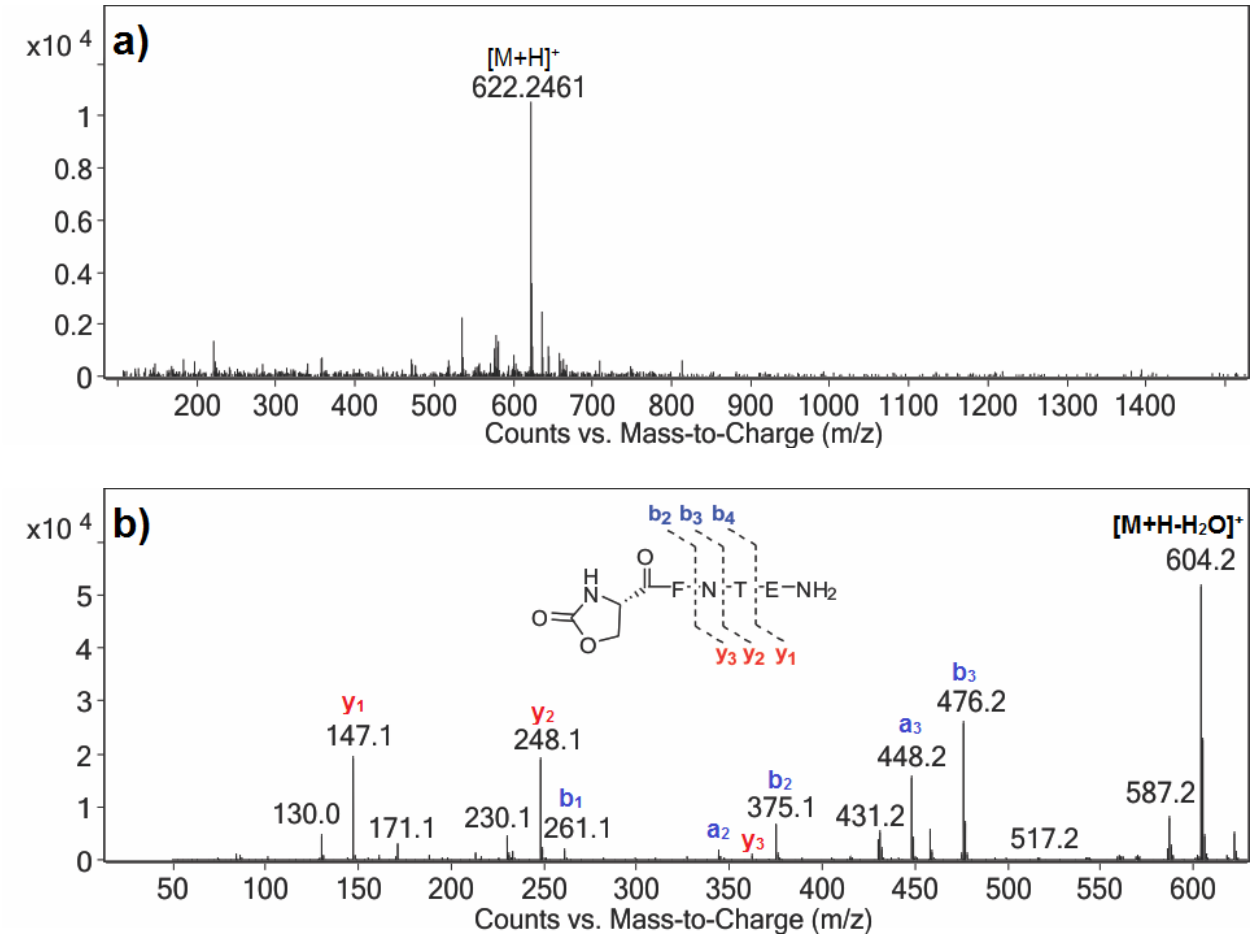


Figure A73 LC-IM-MS/MS data for linear peptide *Oxd*-FNTE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 622.2467, Found $[M+H]^+$ = 622.2461

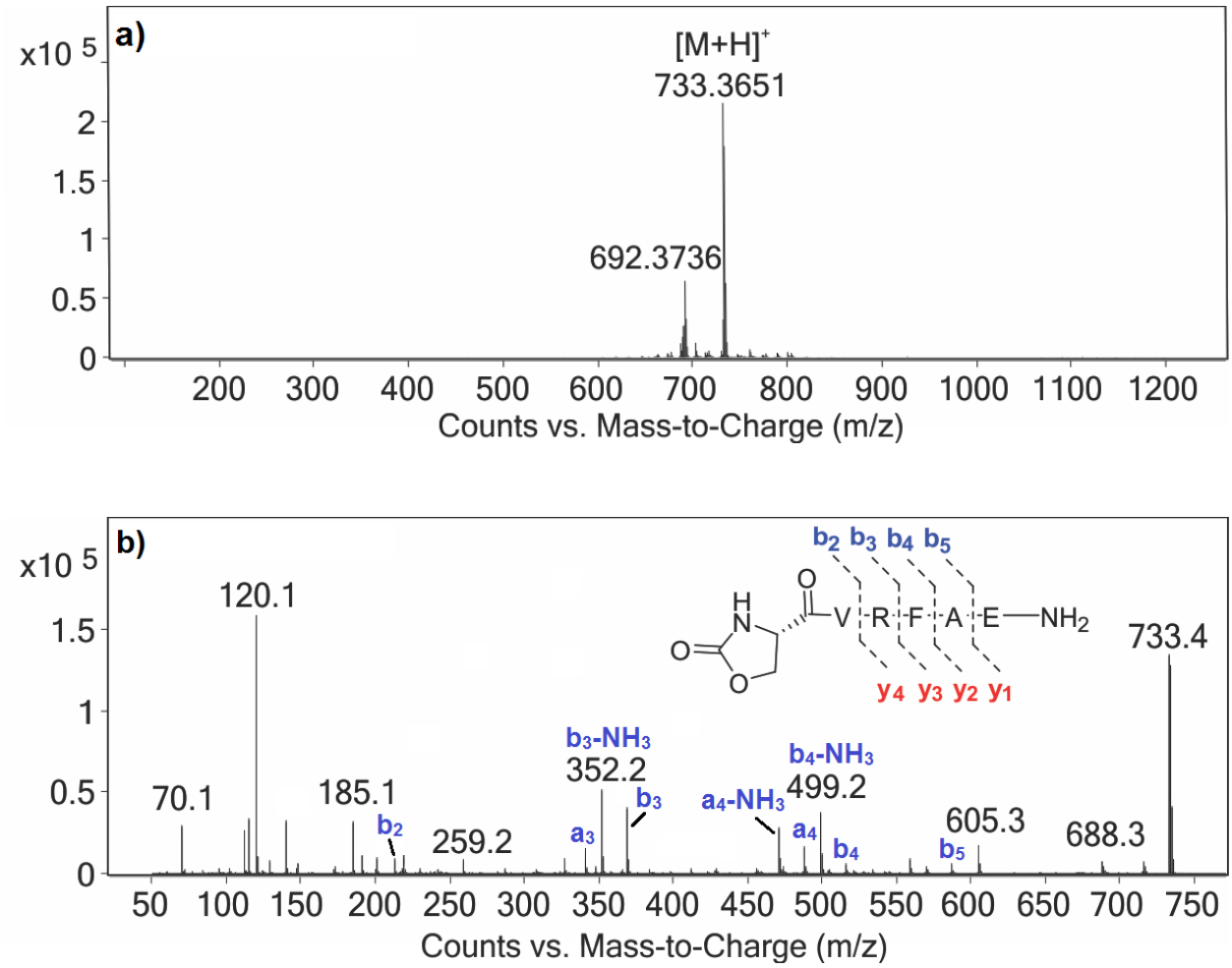


Figure A74 LC-IM-MS/MS data for linear peptide *Oxd*-VRFAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 733.3628, Found $[M+H]^+$ = 733.3651

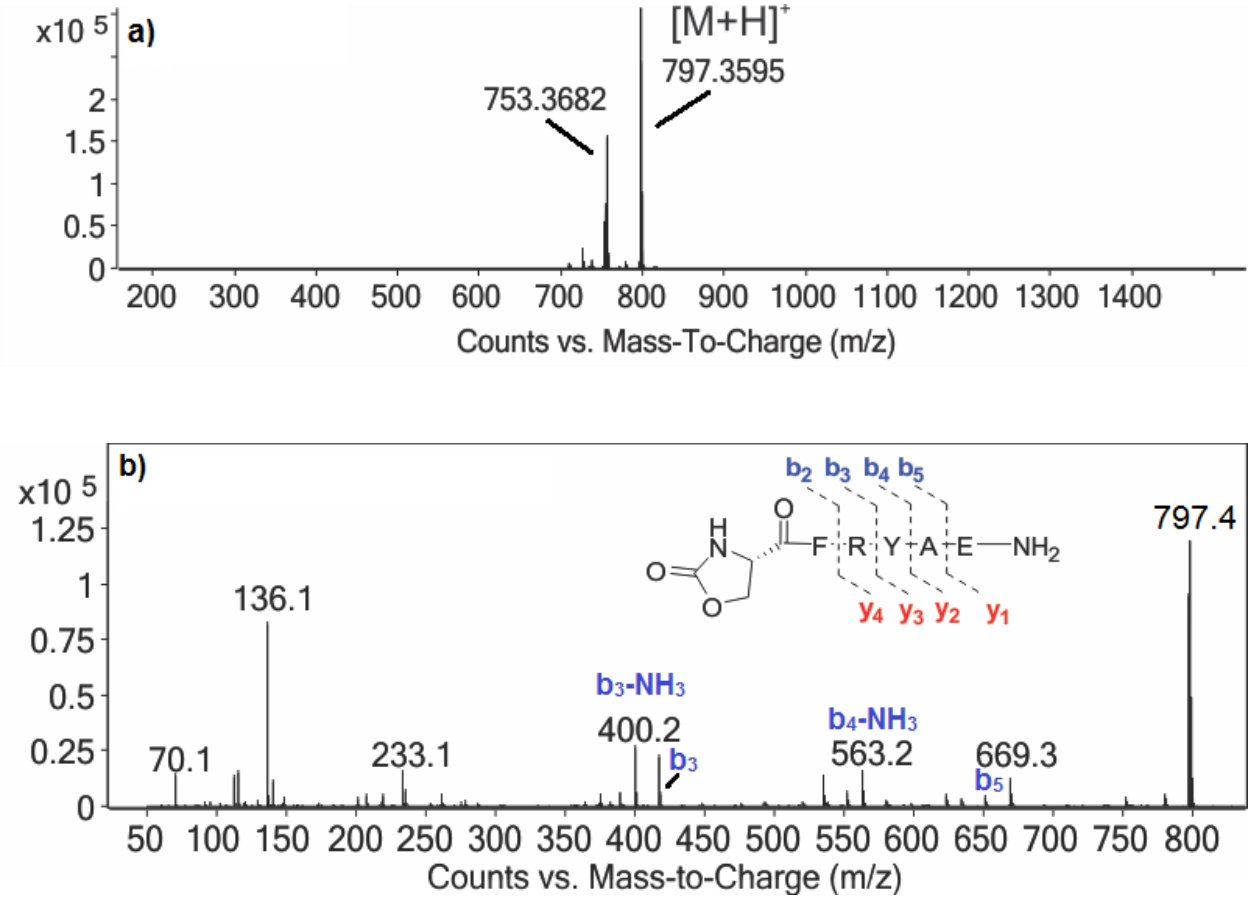


Figure A75 LC-IM-MS/MS data for linear peptide *Oxd*-FRYAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 797.3577, Found $[M+H]^+$ = 797.3595

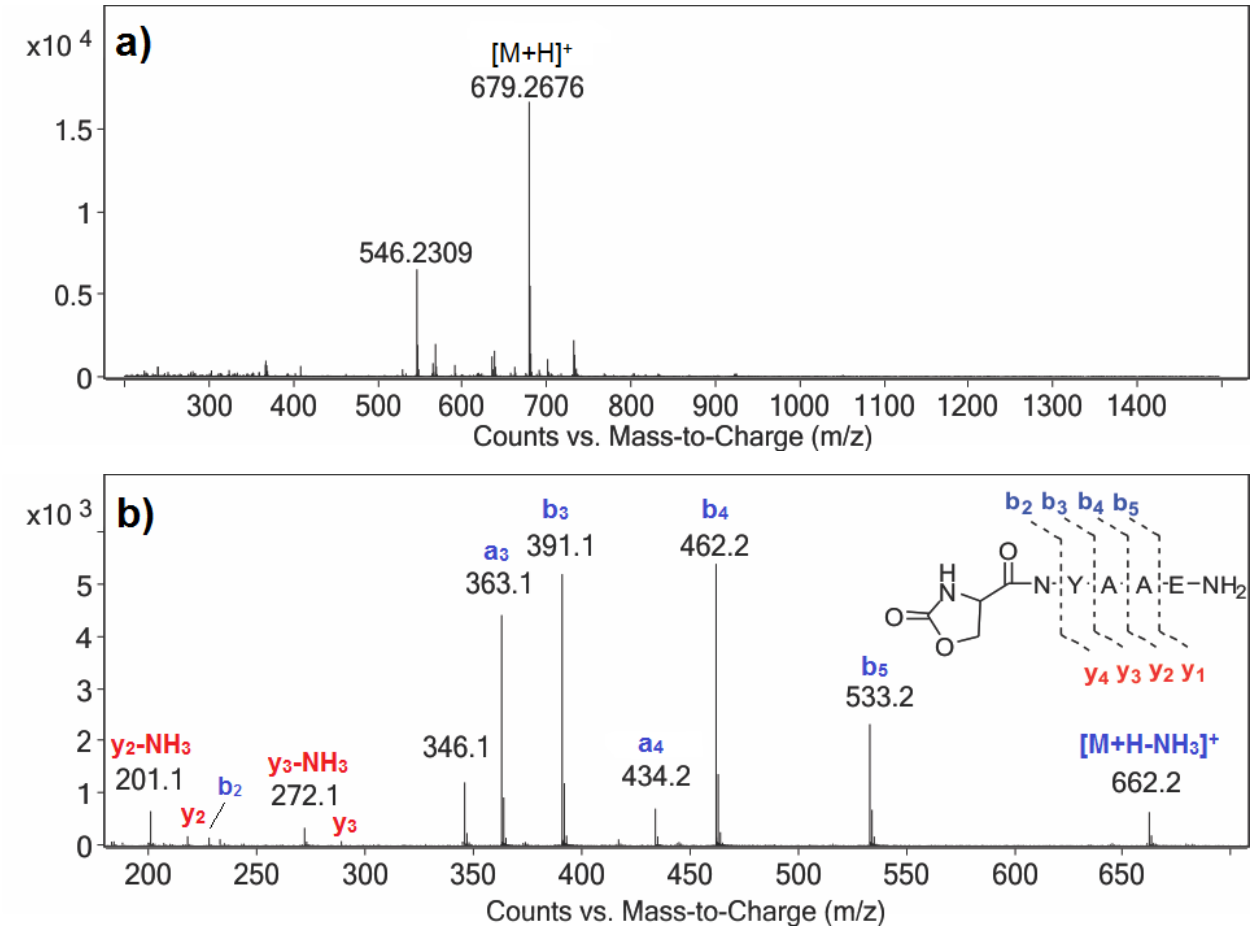


Figure A76 LC-IM-MS/MS data for linear peptide *Oxd*-NYAAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 679.2682, Found $[M+H]^+$ = 679.2676

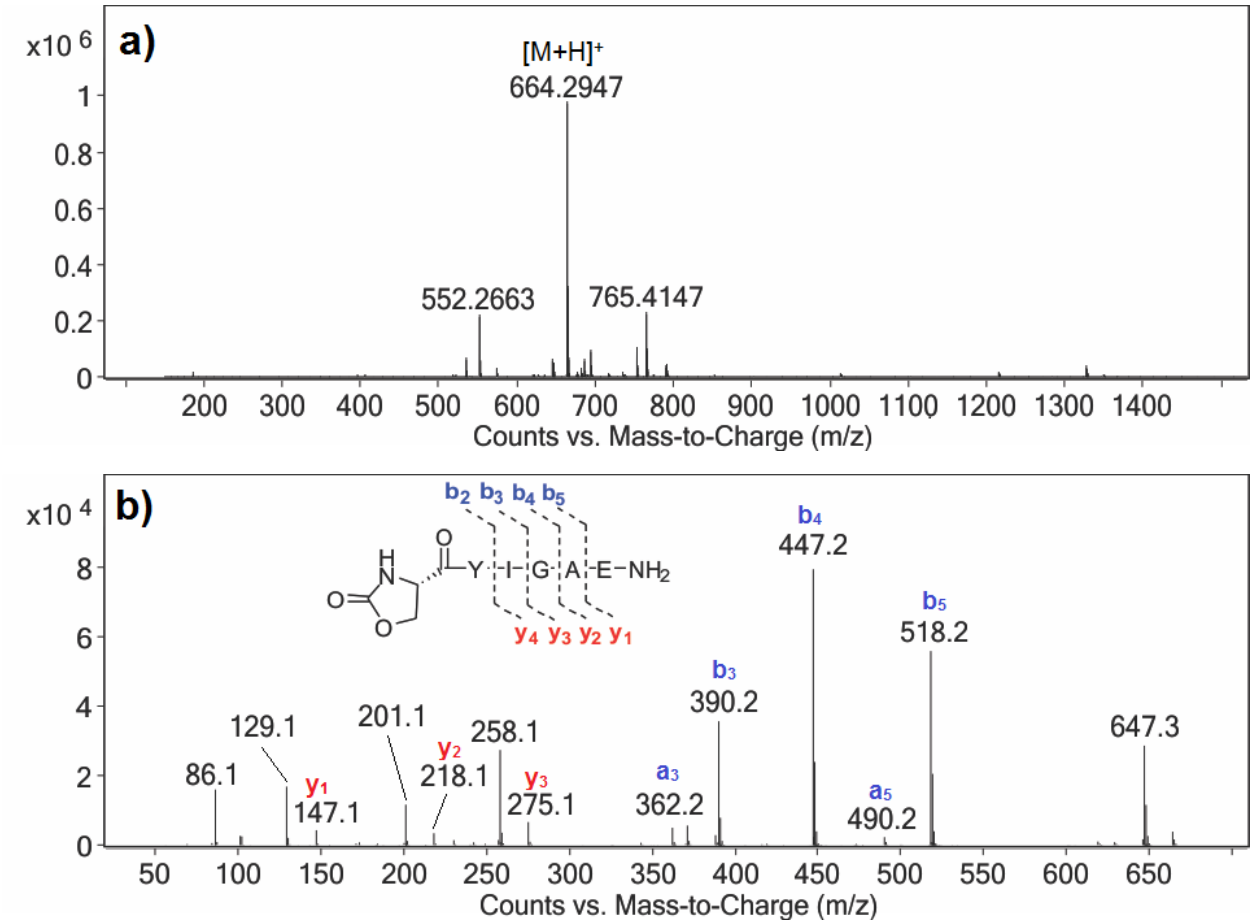


Figure A77 LC-IM-MS/MS data for linear peptide *Oxd*-YIGAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 664.2937, Found $[M+H]^+$ = 664.2947

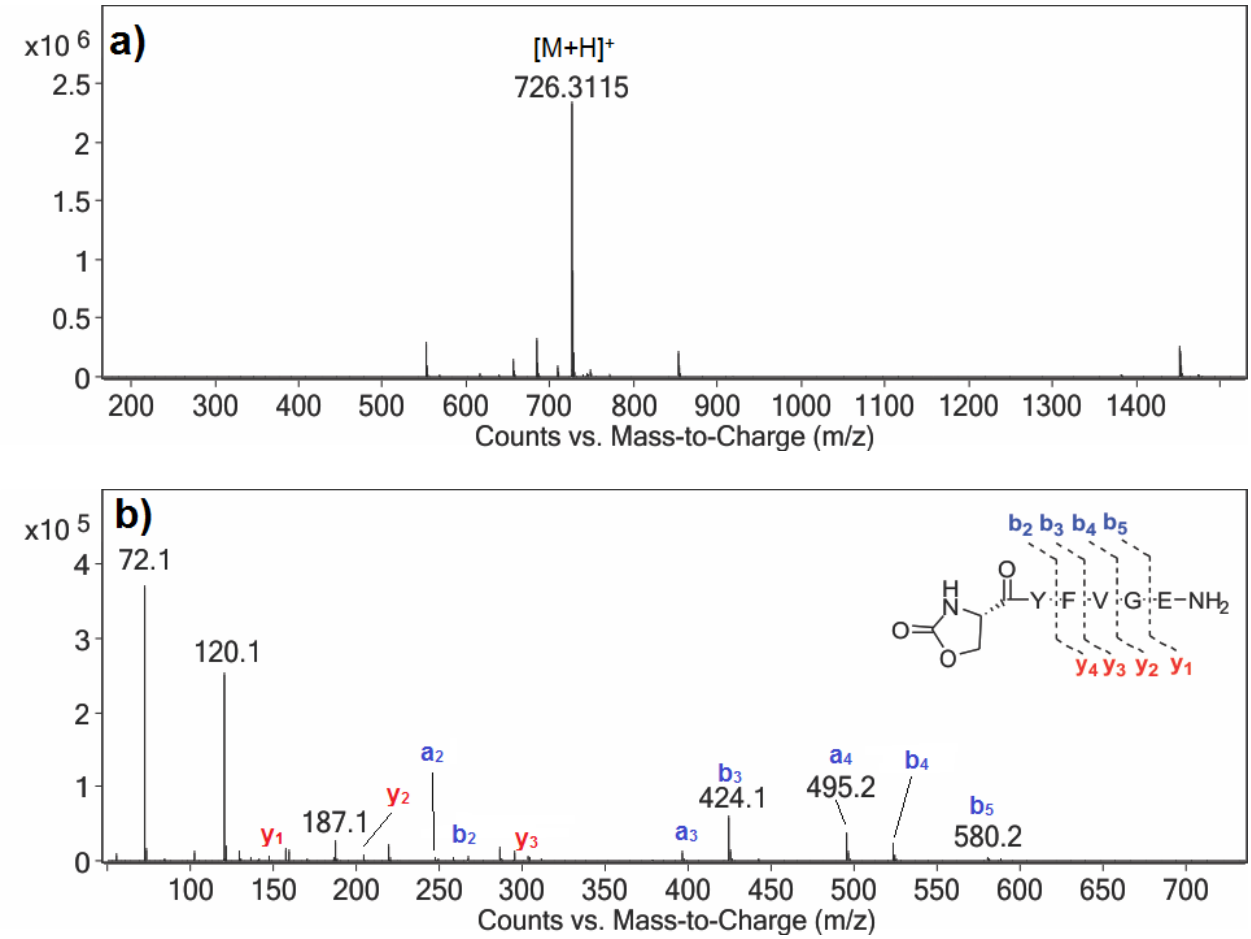


Figure A78 LC-IM-MS/MS data for linear peptide *Oxd*-YFVGE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 726.3093, Found $[M+H]^+$ = 726.3115

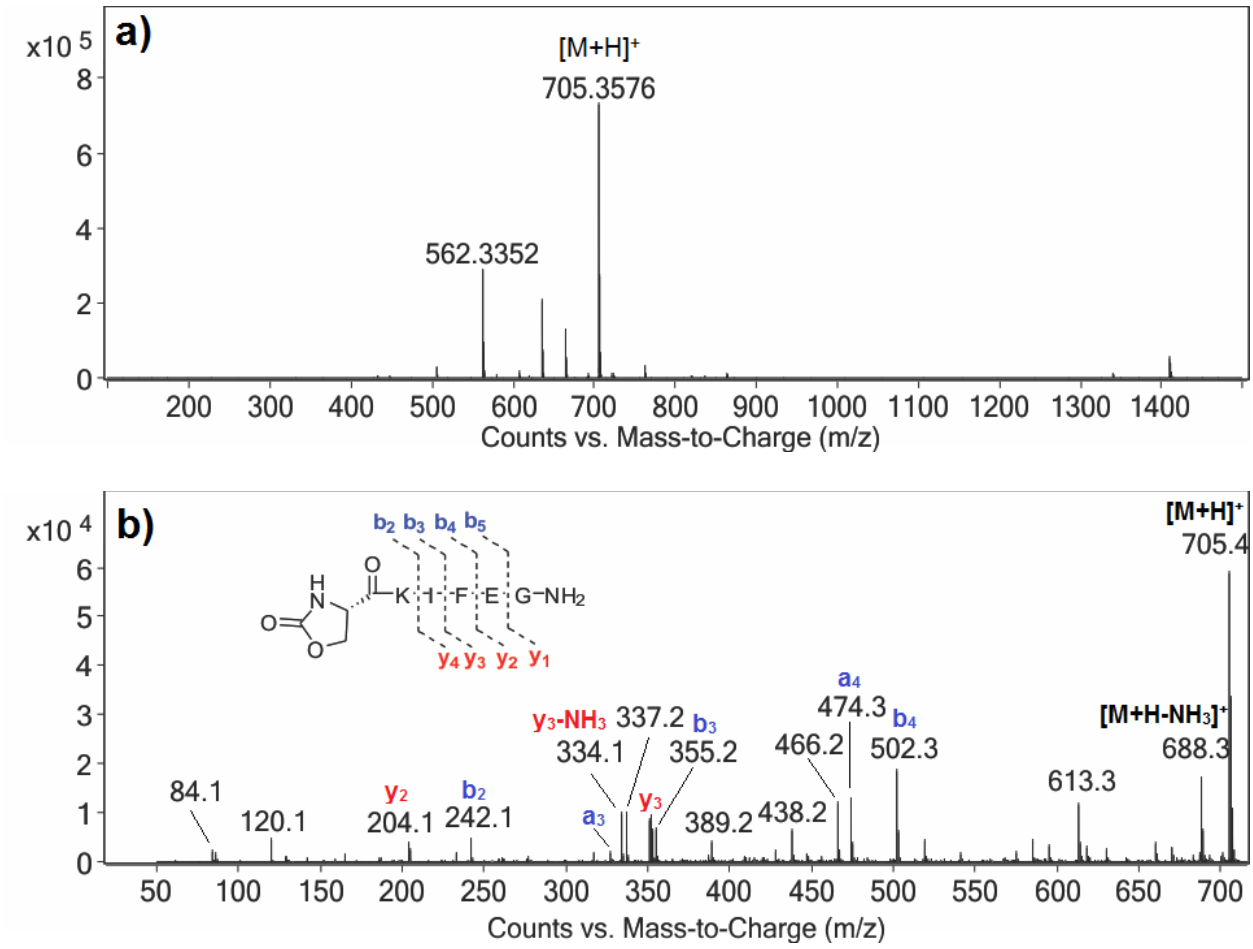


Figure A79 LC-IM-MS/MS data for linear peptide *Oxd*-KIFEG-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 705.3566, Found $[M+H]^+$ = 705.3576

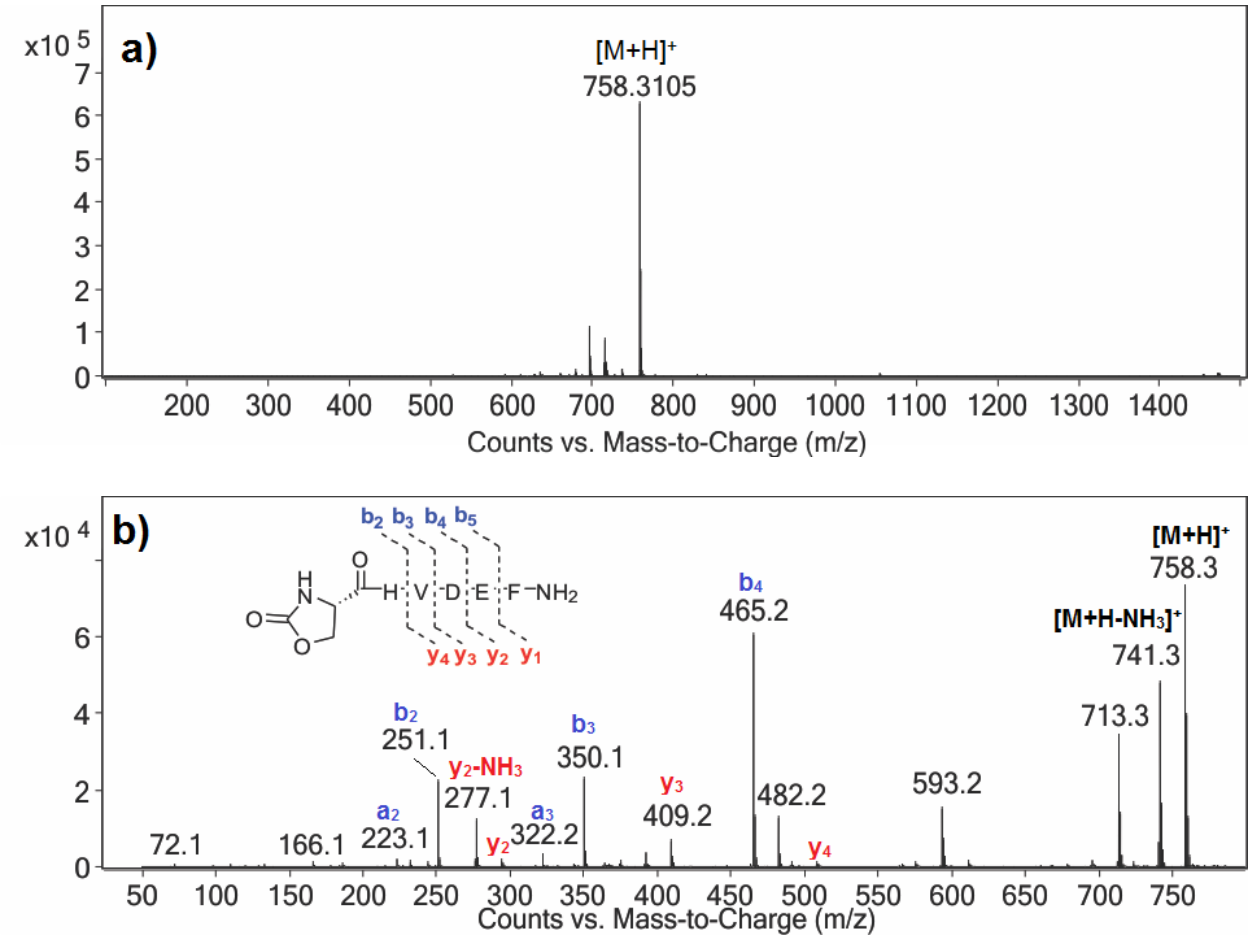


Figure A80 LC-IM-MS/MS data for linear peptide *Oxd*-HVDEF-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 758.3104, Found $[M+H]^+$ = 758.3105

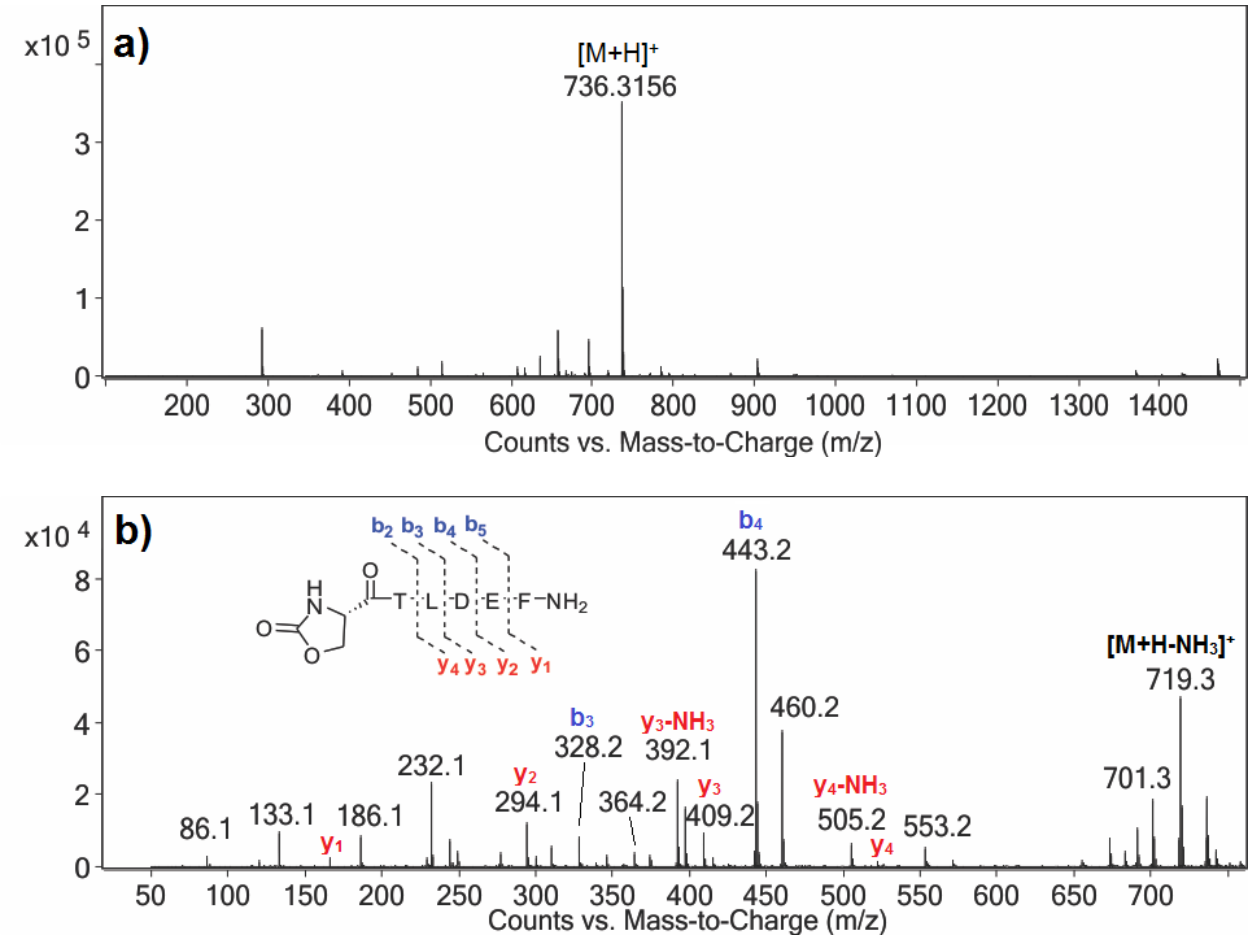


Figure A81 LC-IM-MS/MS data for linear peptide *Oxd*-TLDEF-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 736.3148, Found $[M+H]^+$ = 736.3156

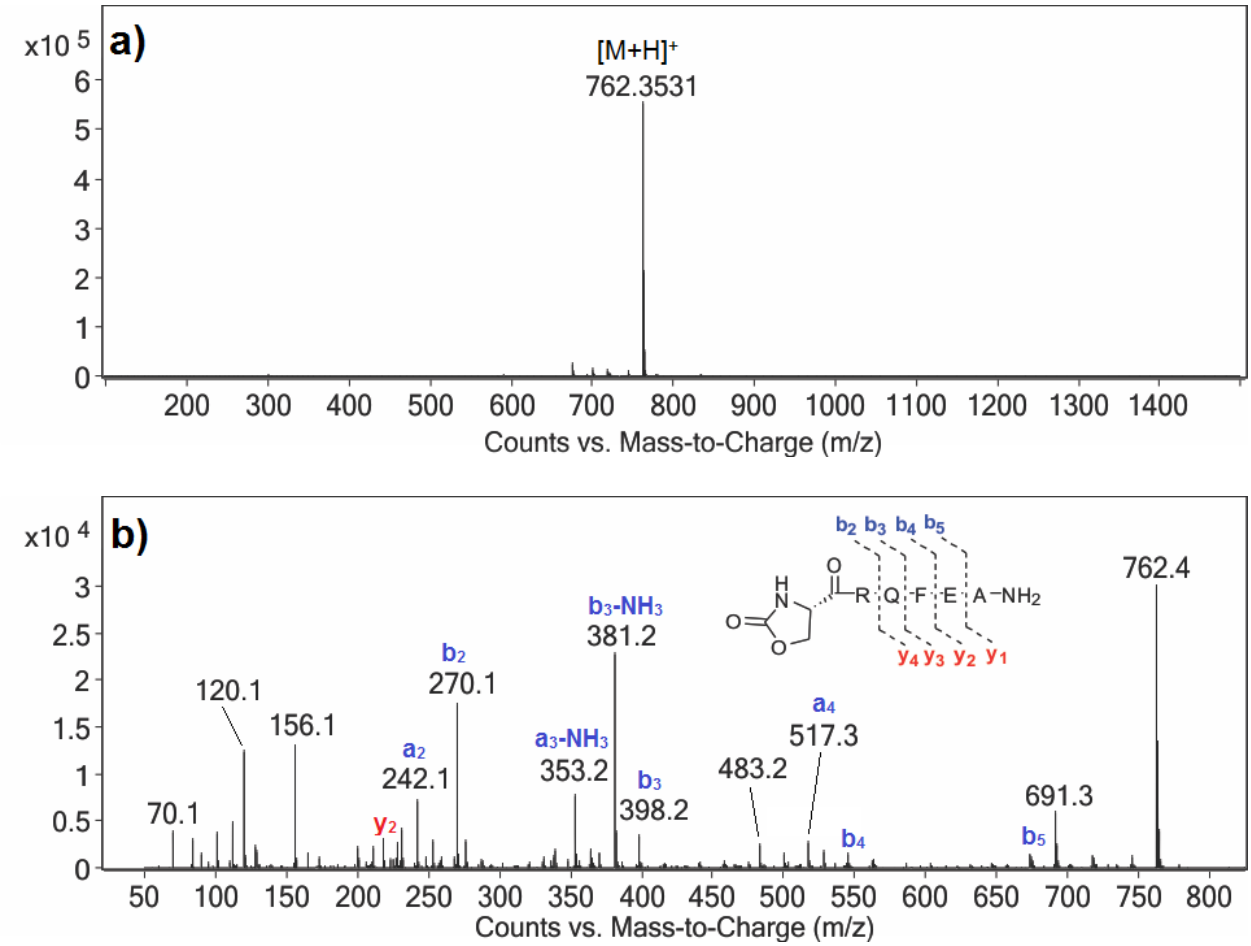


Figure A82 LC-IM-MS/MS data for linear peptide *Oxd*-RQFEA-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 762.3529, Found $[M+H]^+$ = 762.3531

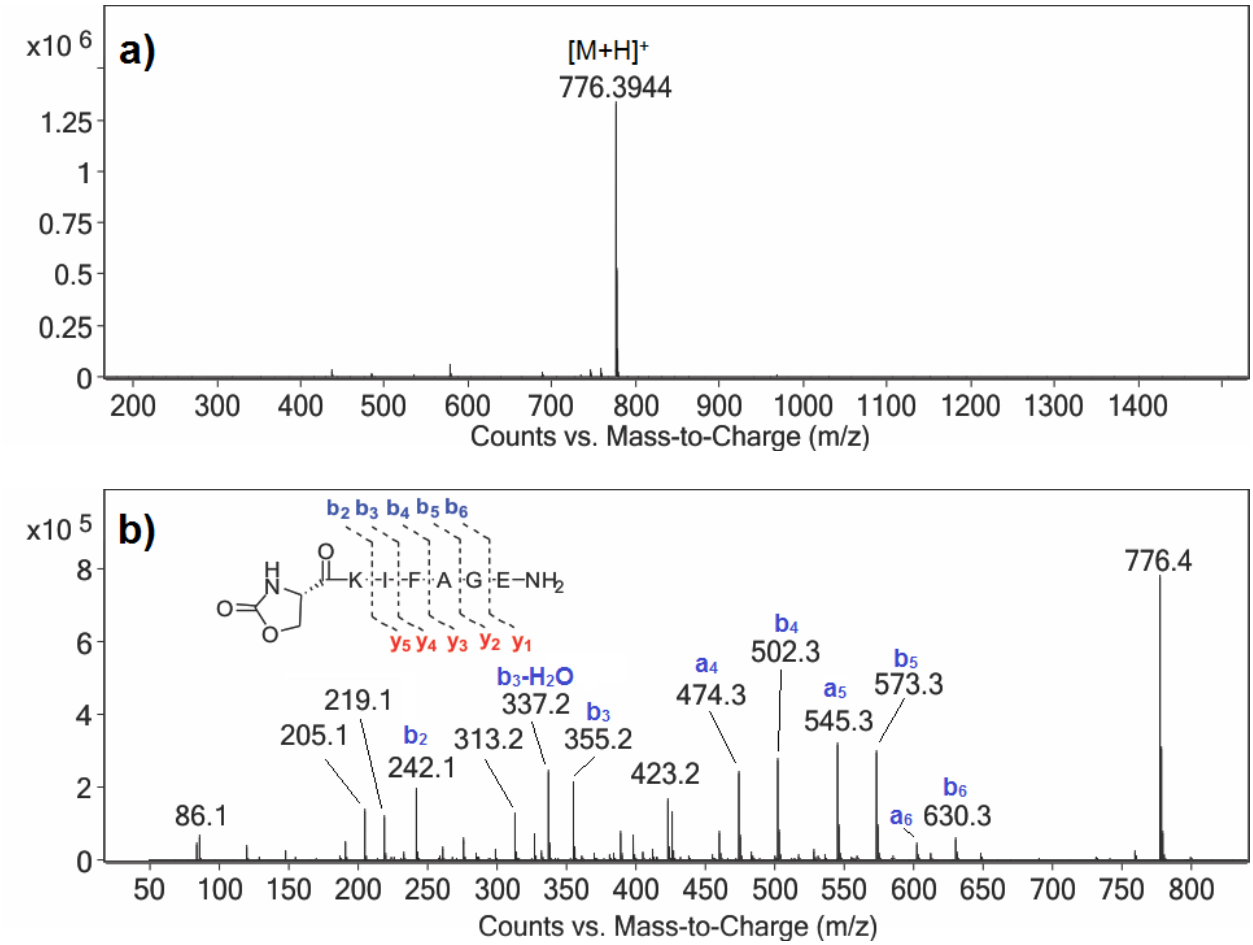


Figure A83 LC-IM-MS/MS data for linear peptide *Oxd*-KIFAGE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 776.3937, Found $[M+H]^+$ = 776.3944

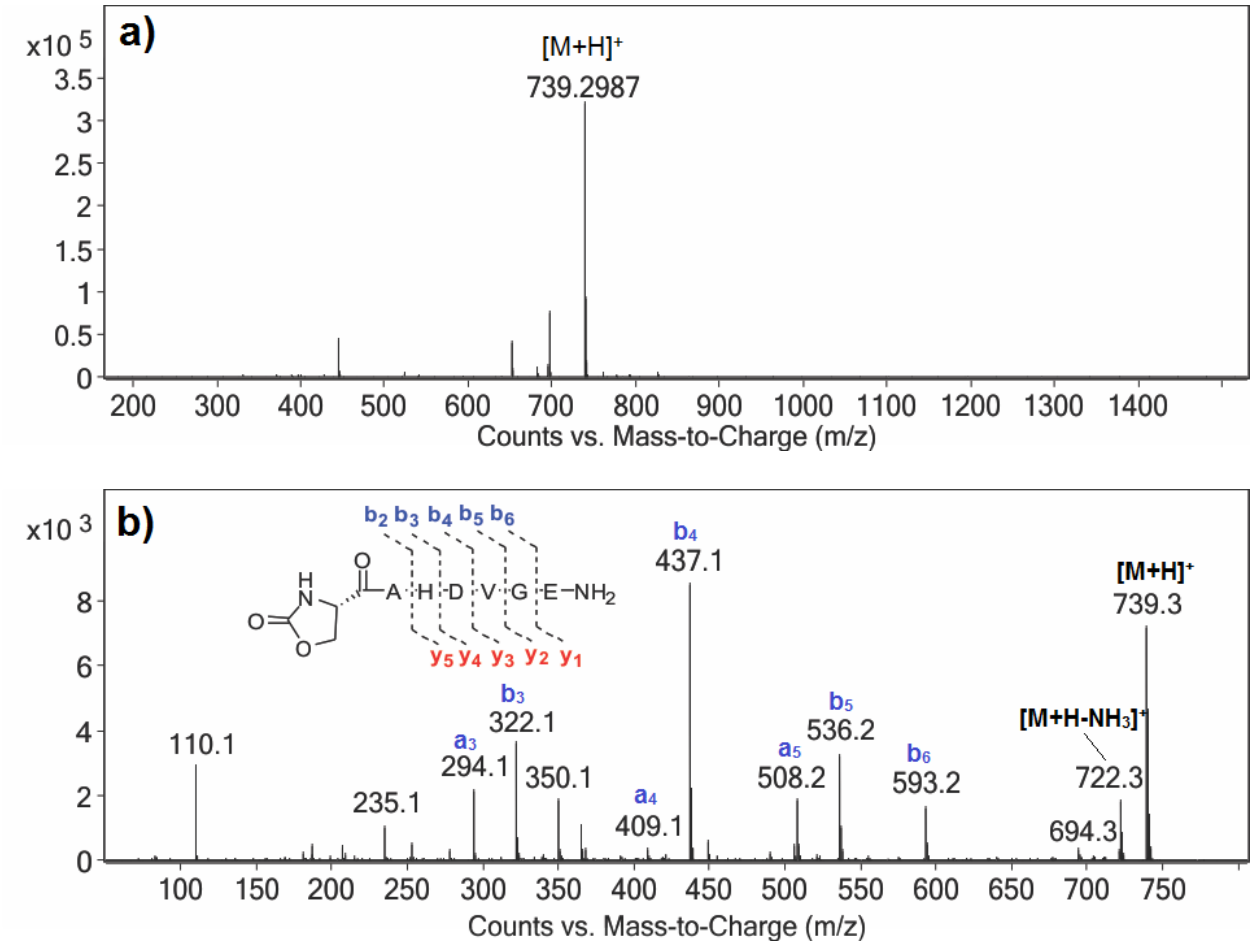


Figure A84 LC-IM-MS/MS data for linear peptide *Oxd*-AHDVGE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 739.3006, Found $[M+H]^+$ = 739.2987

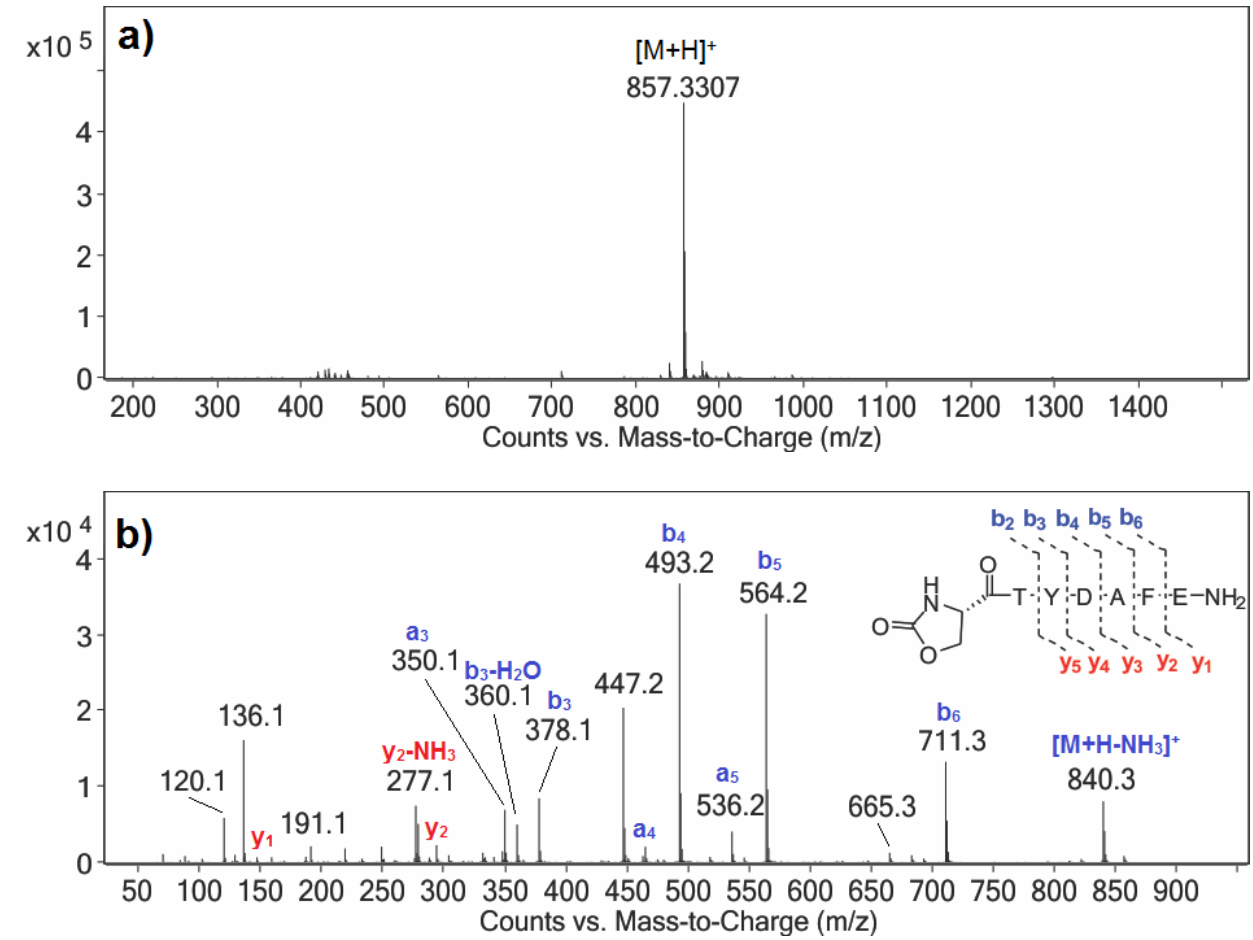


Figure A85 LC-IM-MS/MS data for linear peptide *Oxd*-TYDAFE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 857.3312, Found $[M+H]^+$ = 857.3307

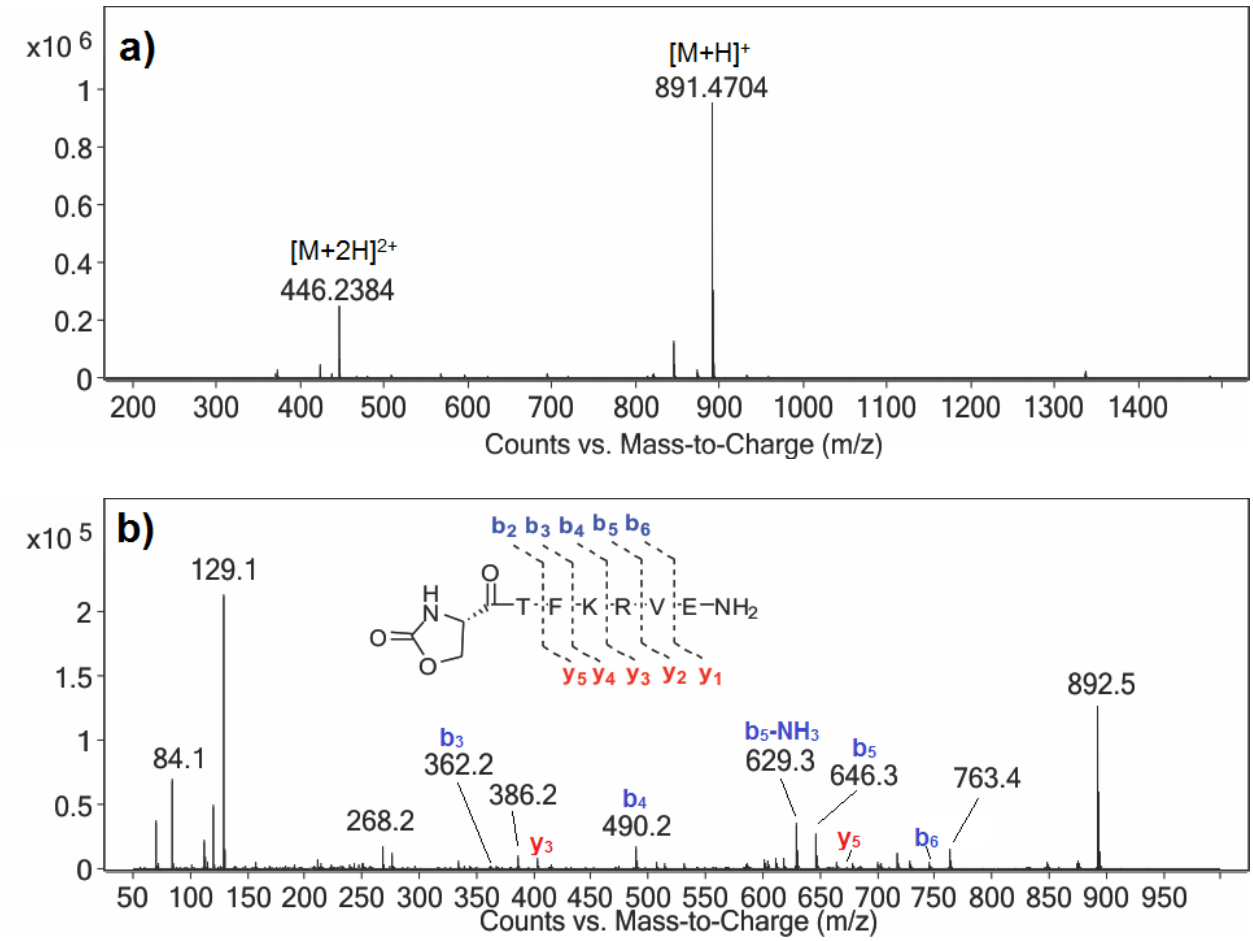


Figure A86 LC-IM-MS/MS data for linear peptide *Oxd*-TFKRVE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 891.4683, Found $[M+H]^+$ = 891.4704

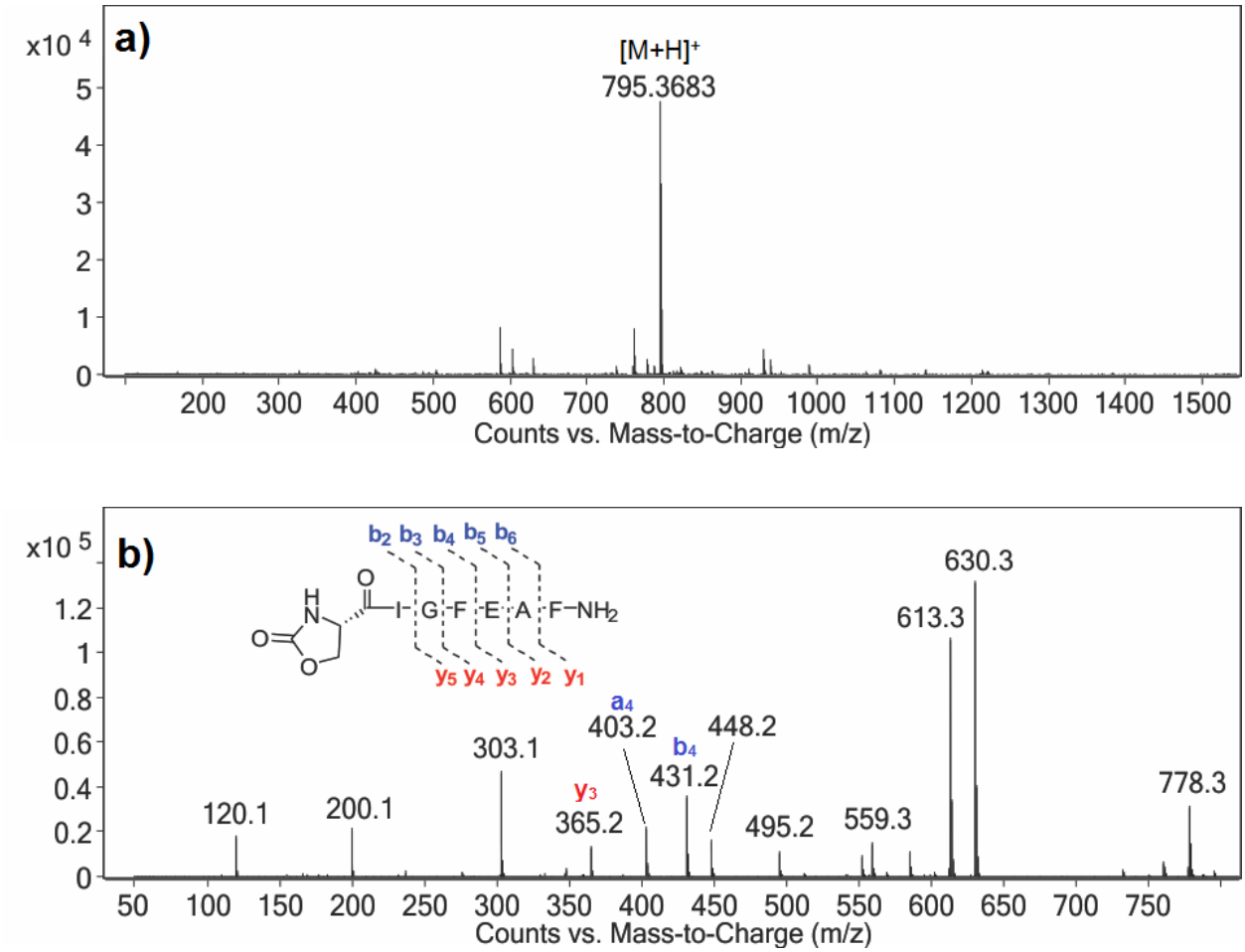


Figure A87 LC-IM-MS/MS data for linear peptide *Oxd*-IGFEAF-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 795.3672, Found $[M+H]^+$ = 795.3683

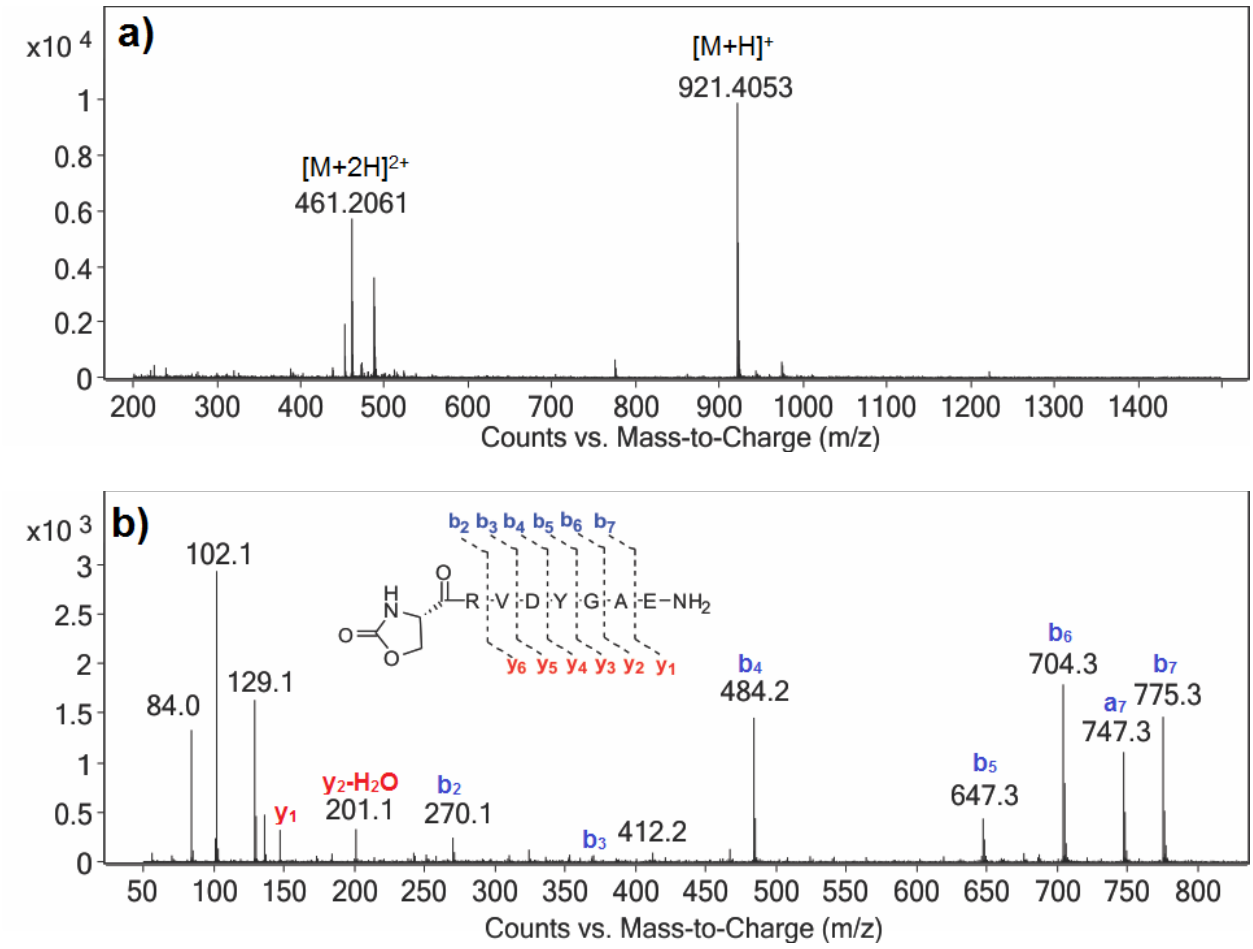


Figure A88 LC-IM-MS/MS data for linear peptide *Oxd*-RVDYGAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 921.4061, Found $[M+H]^+$ = 921.4053

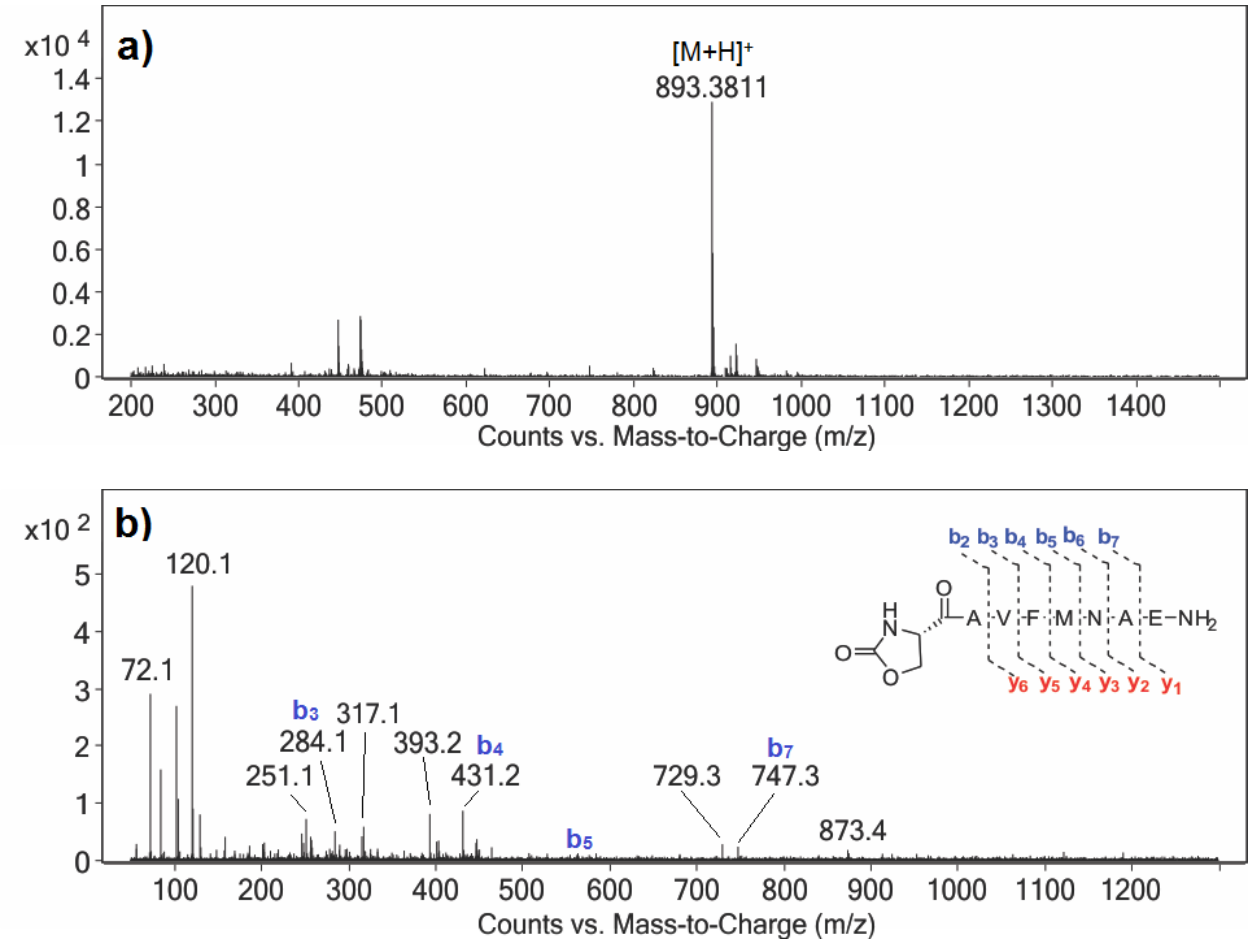


Figure A89 LC-IM-MS/MS data for linear peptide *Oxd*-AVFMNAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 893.3822, Found $[M+H]^+$ = 893.3811

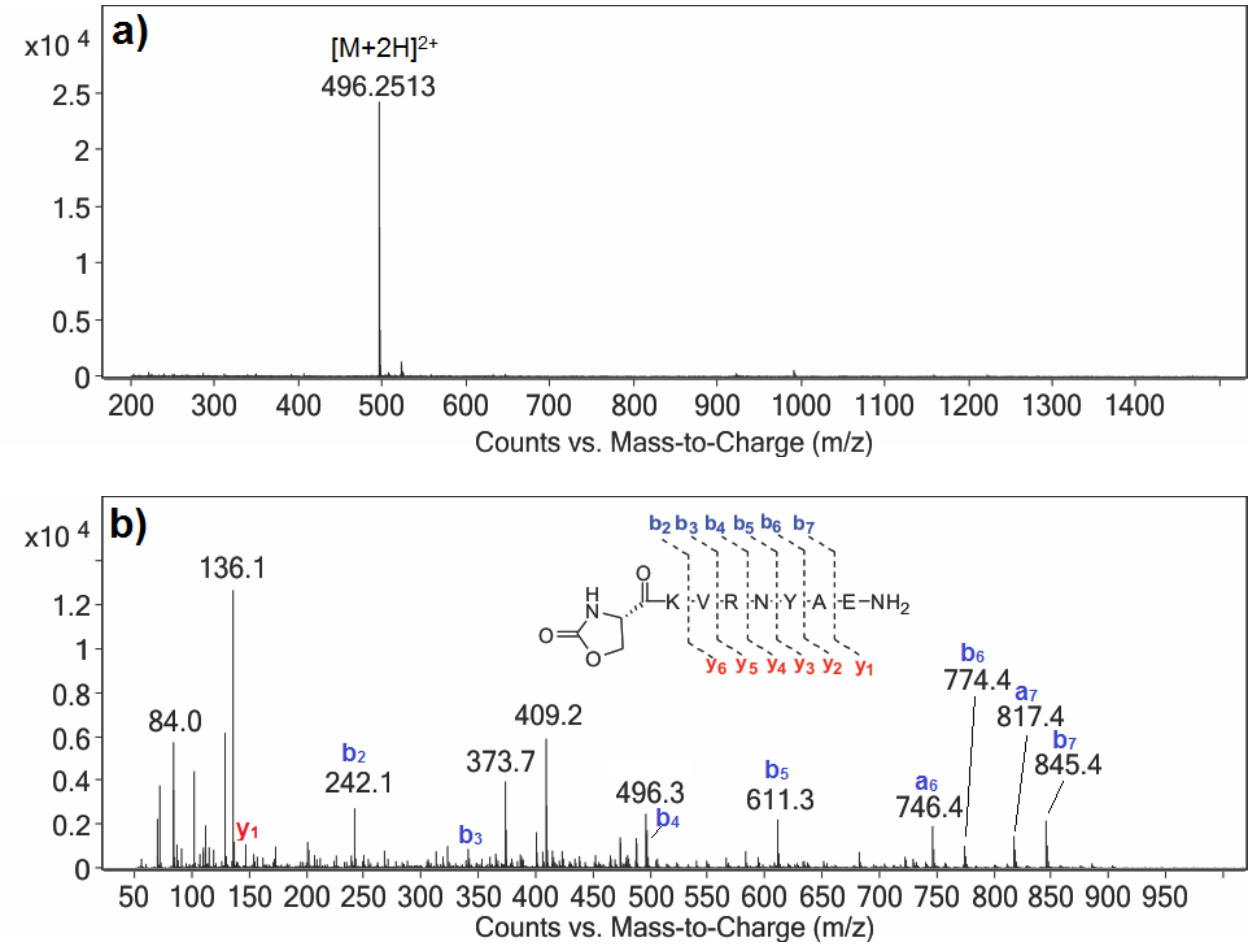


Figure A90 LC-IM-MS/MS data for linear peptide *Oxd*-KVRNYAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+2H]^{2+}$ = 496.2514, Found $[M+2H]^{2+}$ = 496.2513

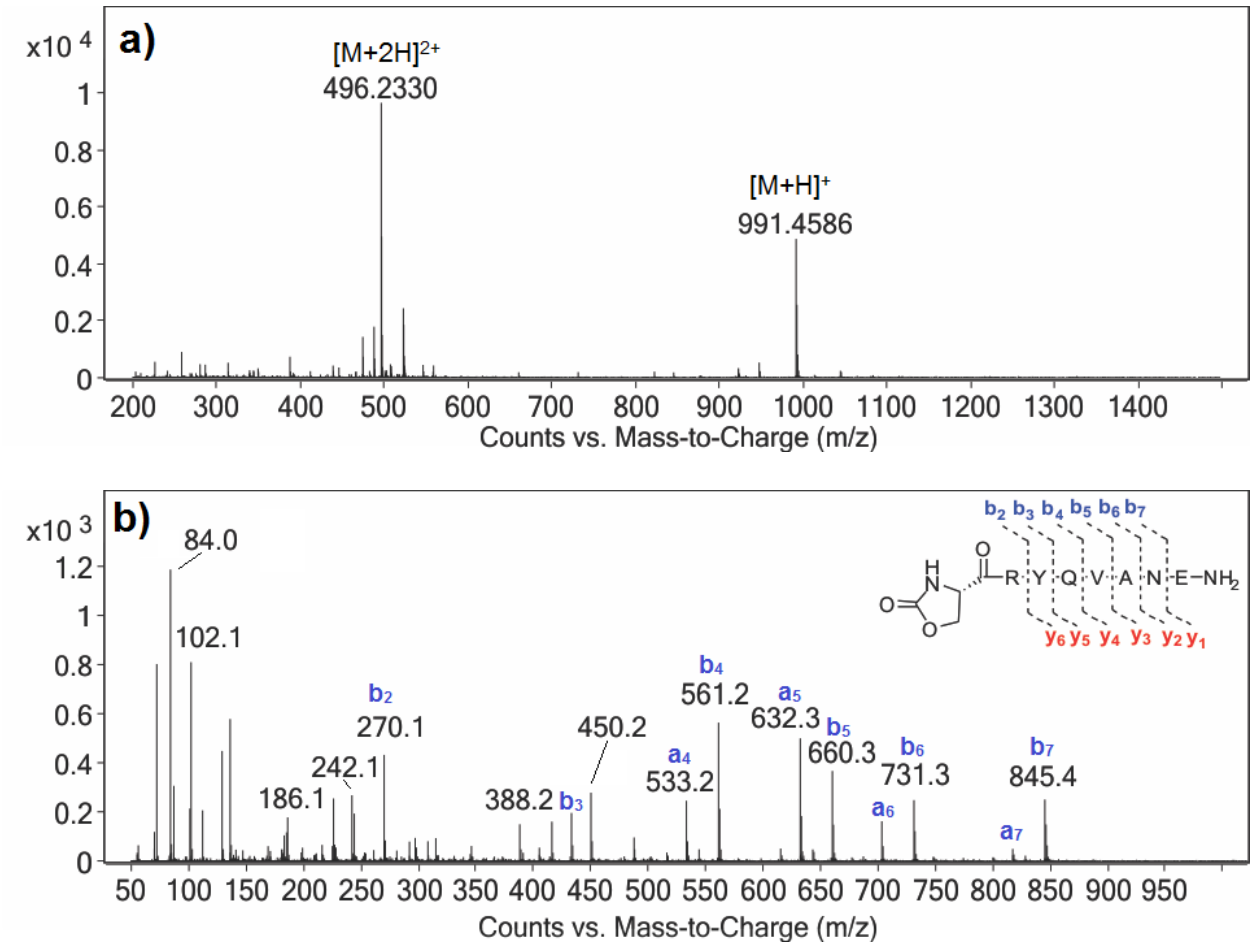


Figure A91 LC-IM-MS/MS data for linear peptide *Oxd*-RYQVANE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+2H]^{2+}$ = 496.2332, Found $[M+2H]^{2+}$ = 496.2330

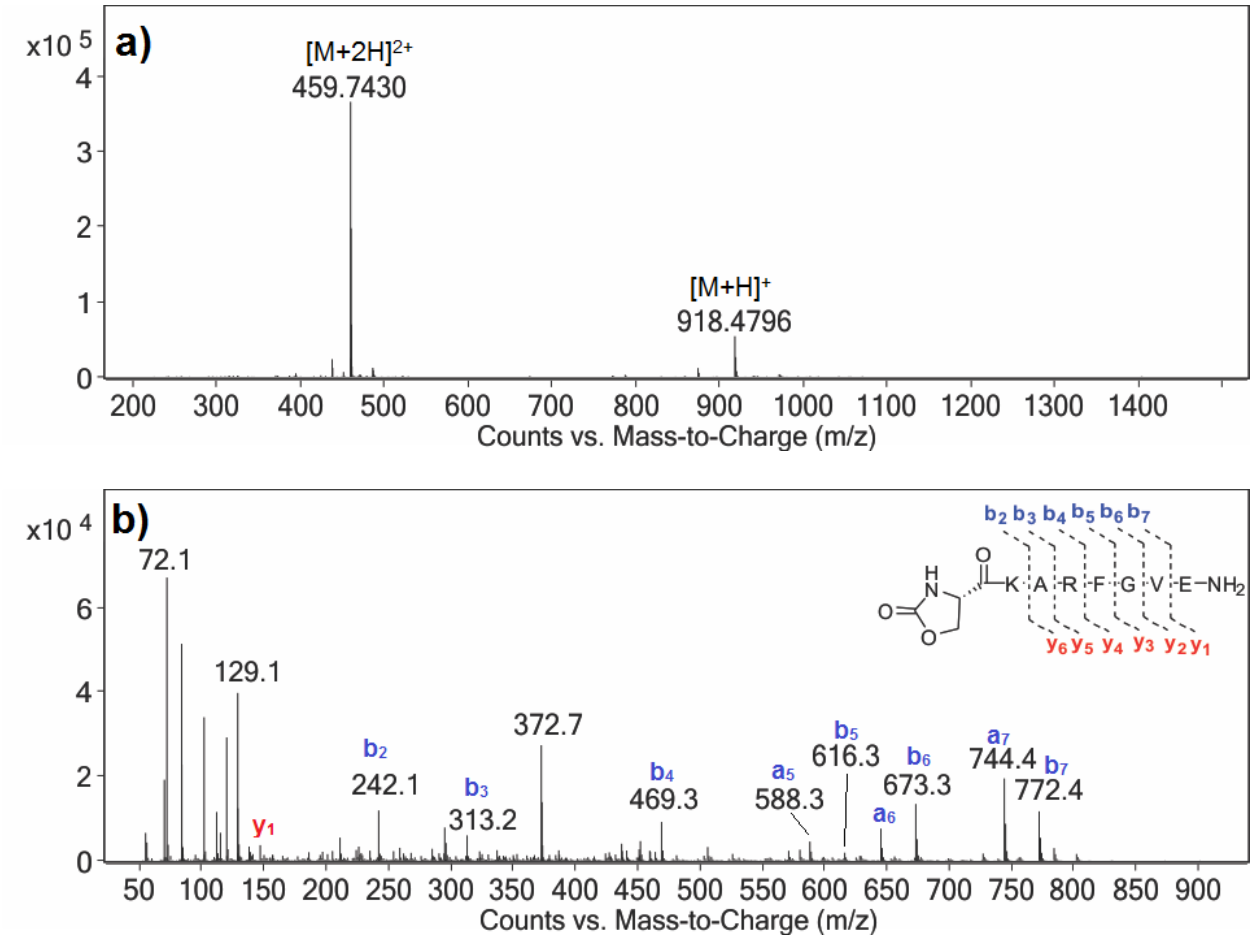


Figure A92 LC-IM-MS/MS data for linear peptide *Oxd*-KARFGVE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+2H]^{2+}$ = 459.7432, Found $[M+2H]^{2+}$ = 459.7430

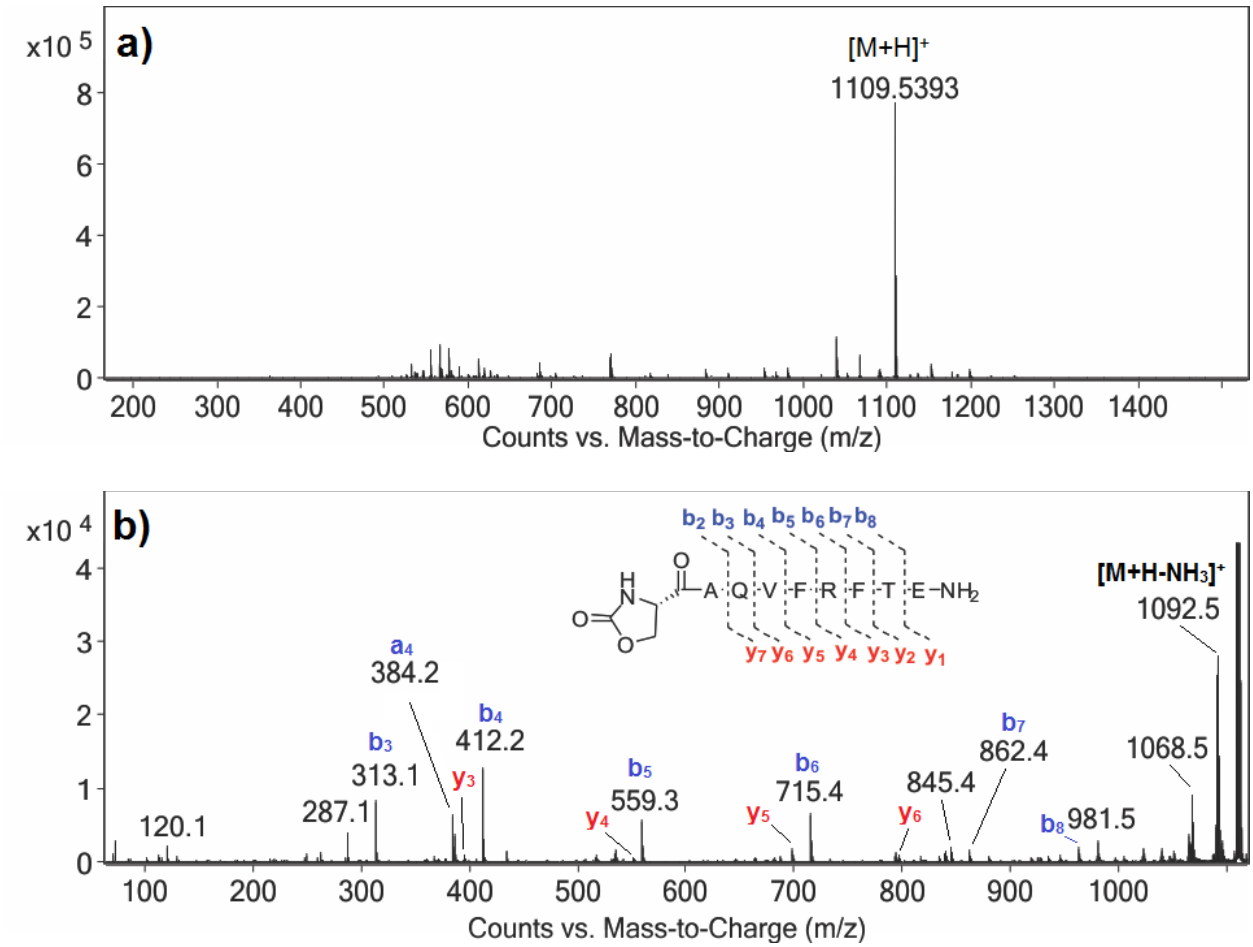


Figure A93 LC-IM-MS/MS data for linear peptide *Oxd*-AQVFRFTE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 1109.5374, Found $[M+H]^+$ = 1109.5393

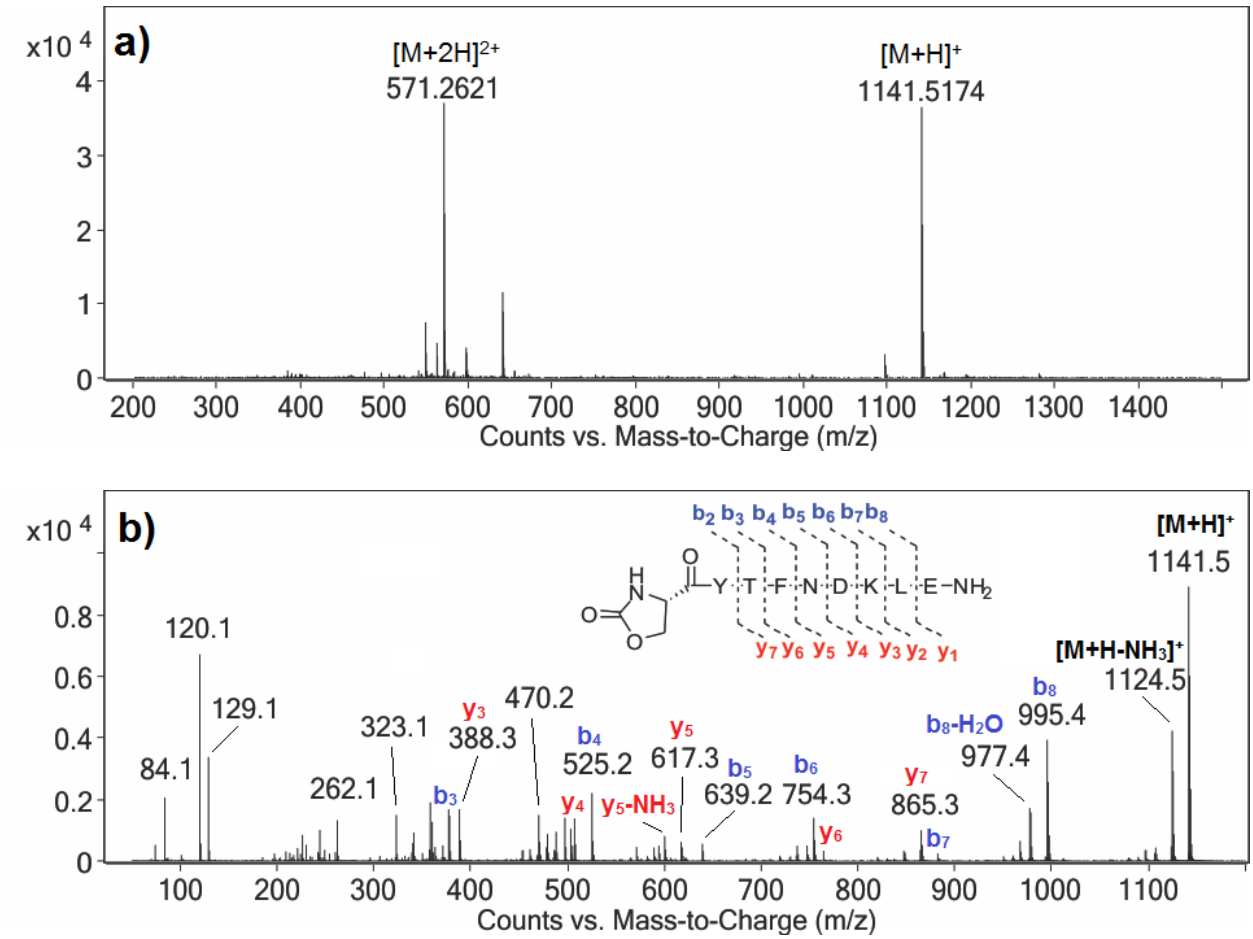


Figure A94 LC-IM-MS/MS data for linear peptide *Oxd*-YTFNDKLE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 1141.5160, Found $[M+H]^+$ = 1141.5174

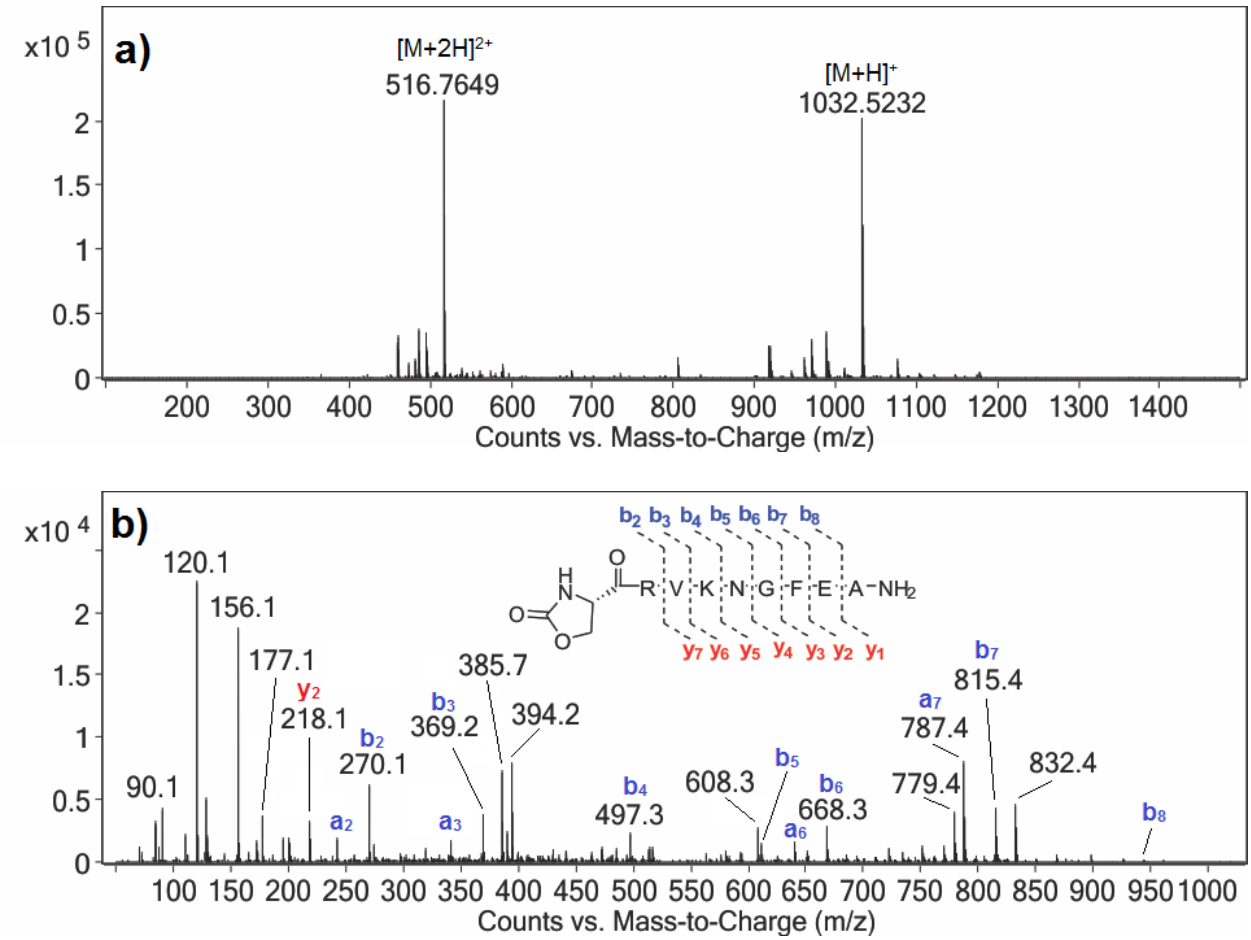


Figure A95 LC-IM-MS/MS data for linear peptide *Oxd*-RVKNGFEA-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 1032.5221, Found $[M+H]^+$ = 1032.5232

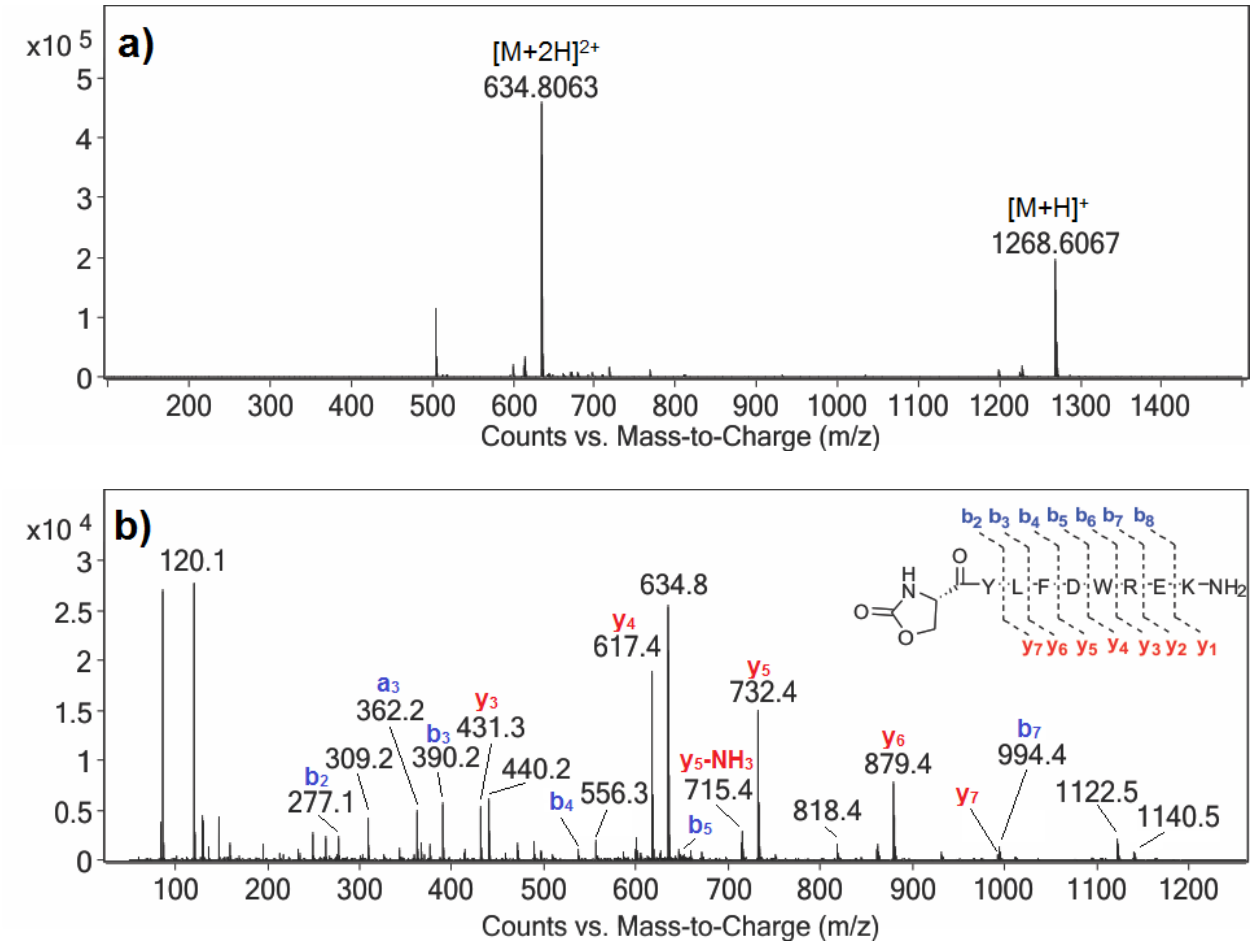


Figure A96 LC-IM-MS/MS data for linear peptide *Oxd*-YLFDWREK-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 1268.6058, Found $[M+H]^+$ = 1268.6067

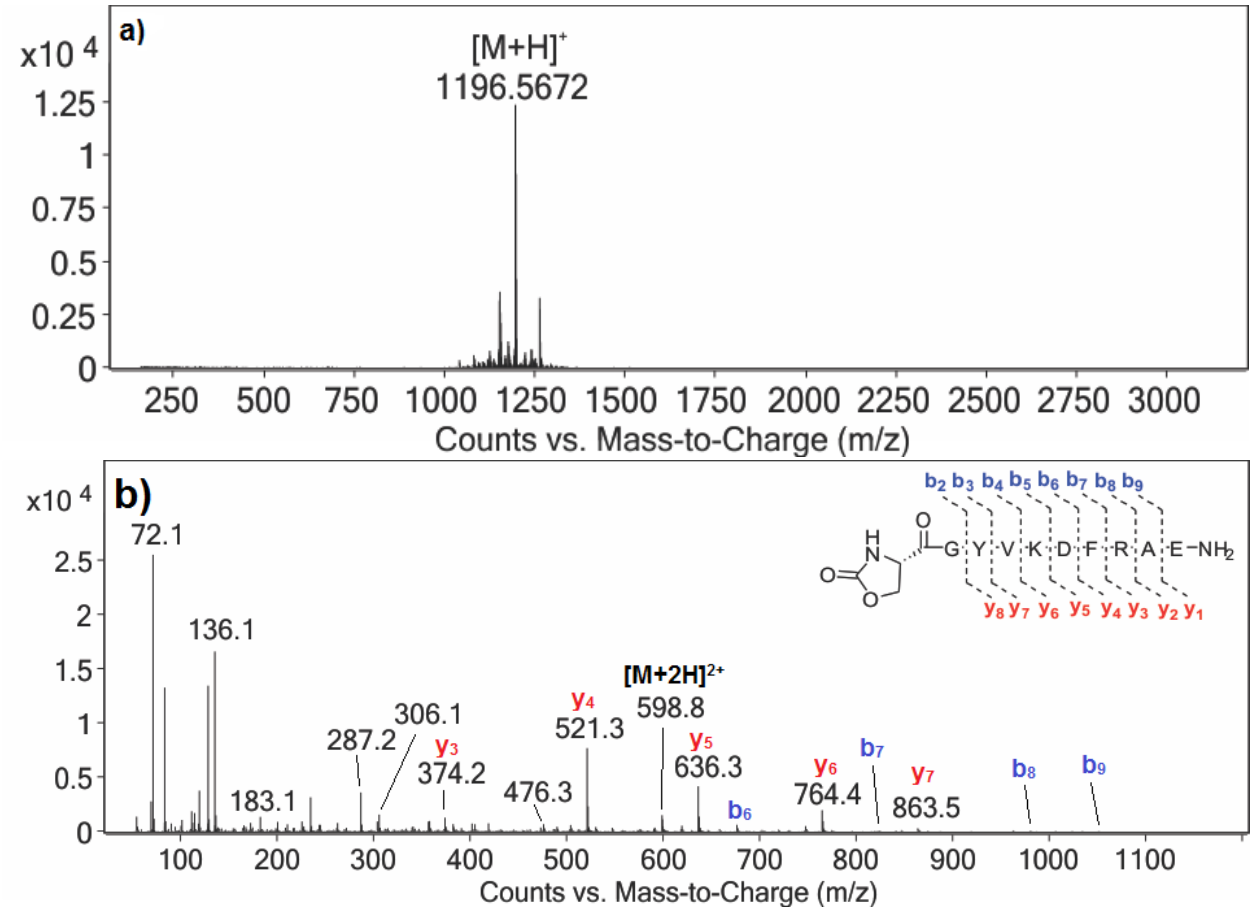


Figure A97 LC-IM-MS/MS data for linear peptide *Oxd*-GYVKDFRAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+H]^+$ = 1196.5695, Found $[M+H]^+$ = 1196.5672

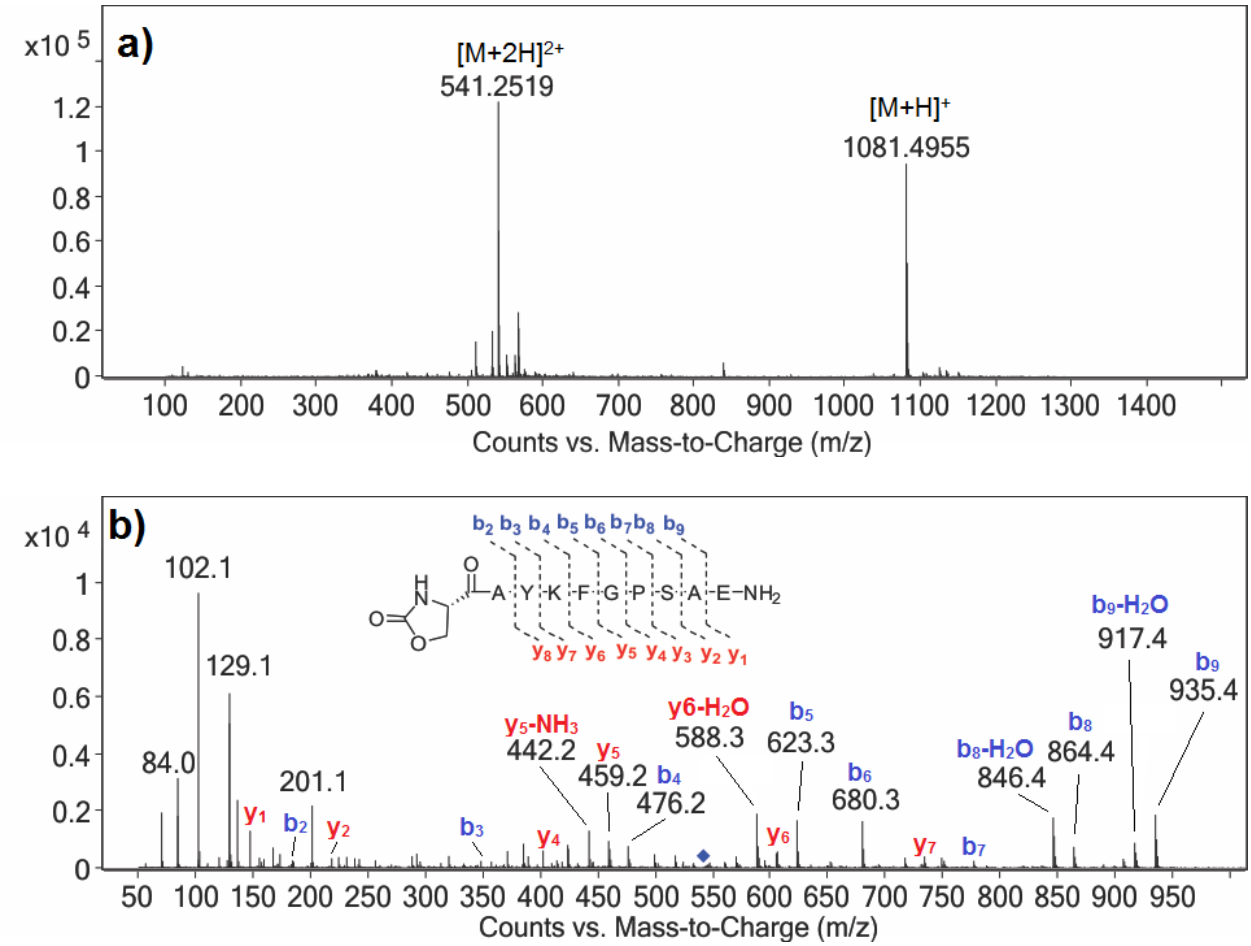


Figure A98 LC-IM-MS/MS data for linear peptide *Oxd*-AYKFGPSAE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+2H]^{2+}$ = 541.2511, Found $[M+2H]^{2+}$ = 541.2519

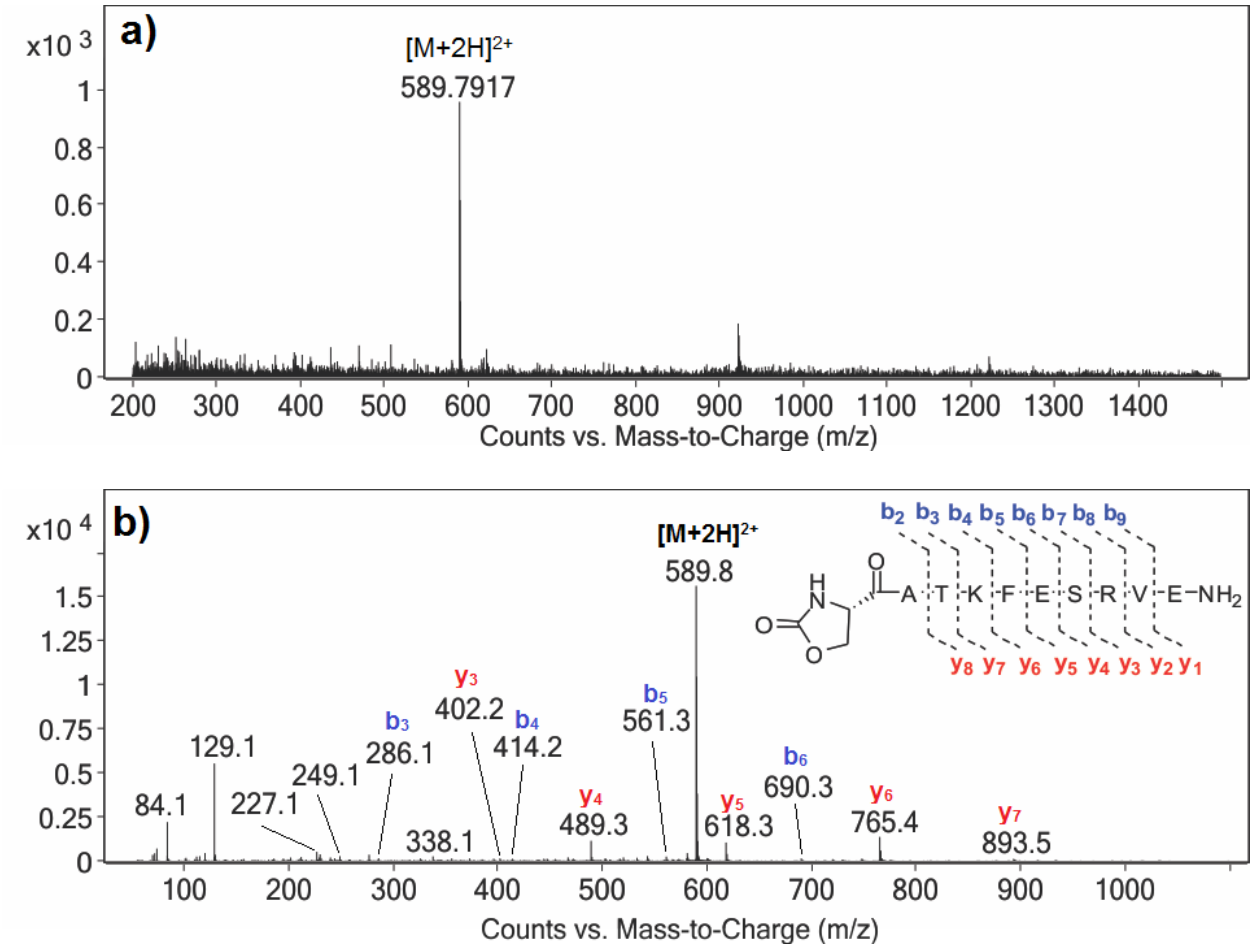


Figure A99 LC-IM-MS/MS data for linear peptide *Oxd*-ATKFESRVE-NH_2. a) extracted MS spectrum, and b) filtered MS/MS spectrum based on product peak drift time; Calcd $[M+2H]^{2+}$ = 589.7937, Found $[M+2H]^{2+}$ = 589.7917

CHAPTER 5 SUPPLEMENTAL MATERIAL

All of the following figures have been reproduced with permission from: Elashal, H. E.; Cohen, R. D.; Raj, M. *Chem. Commun.* **2016**, *52*, 9699.

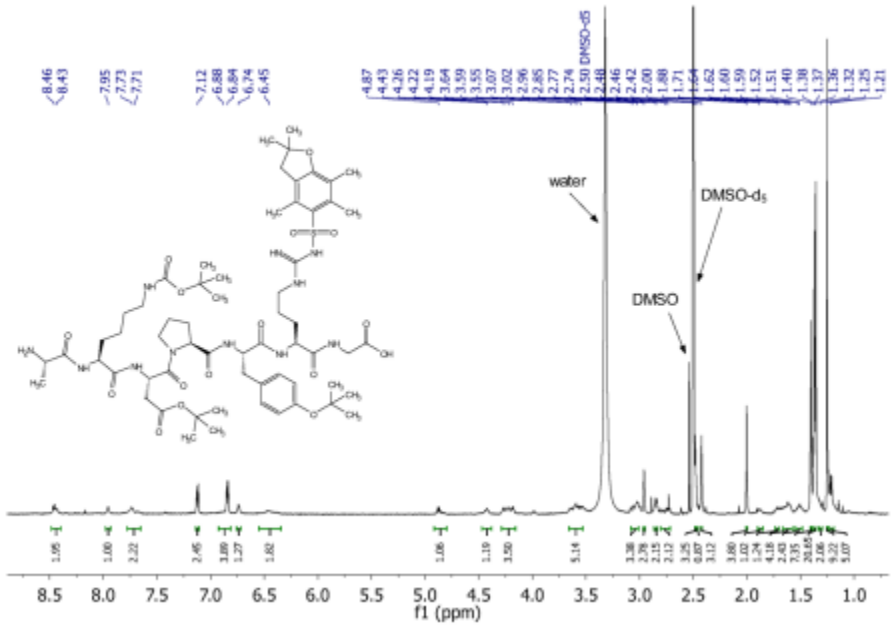


Figure A100 ^{1}H NMR spectrum for linear protected peptide 3b, AK(Boc)D(tBu)PY(tBu)R(Pbf)G-OH, recorded at 600 MHz

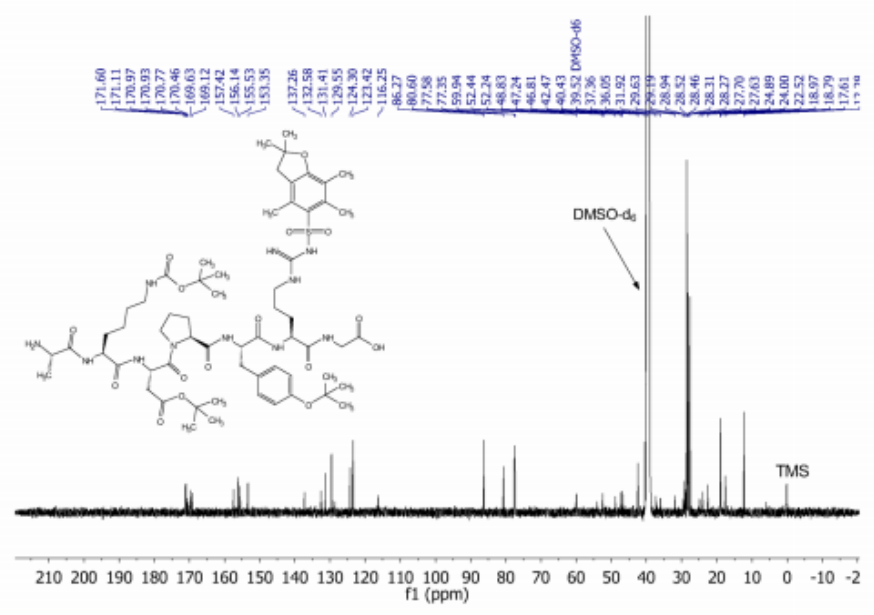


Figure A101 ^{13}C NMR spectrum for linear protected peptide 3b, AK(Boc)D(tBu)PY(tBu)R(Pbf)G-OH, recorded at 151 MHz

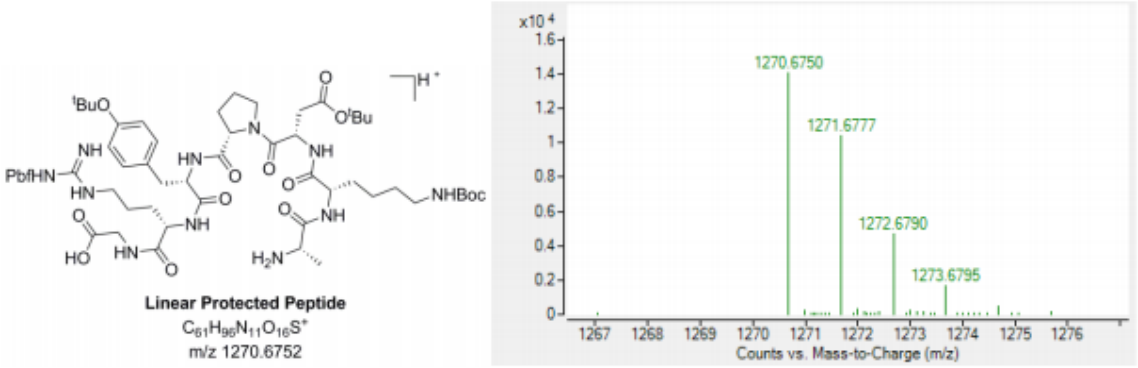


Figure A102 HRMS (+ESI) data for linear protected peptide 3b, AK(Boc)D(tBu)PY(tBu)R(Pbf)G-OH

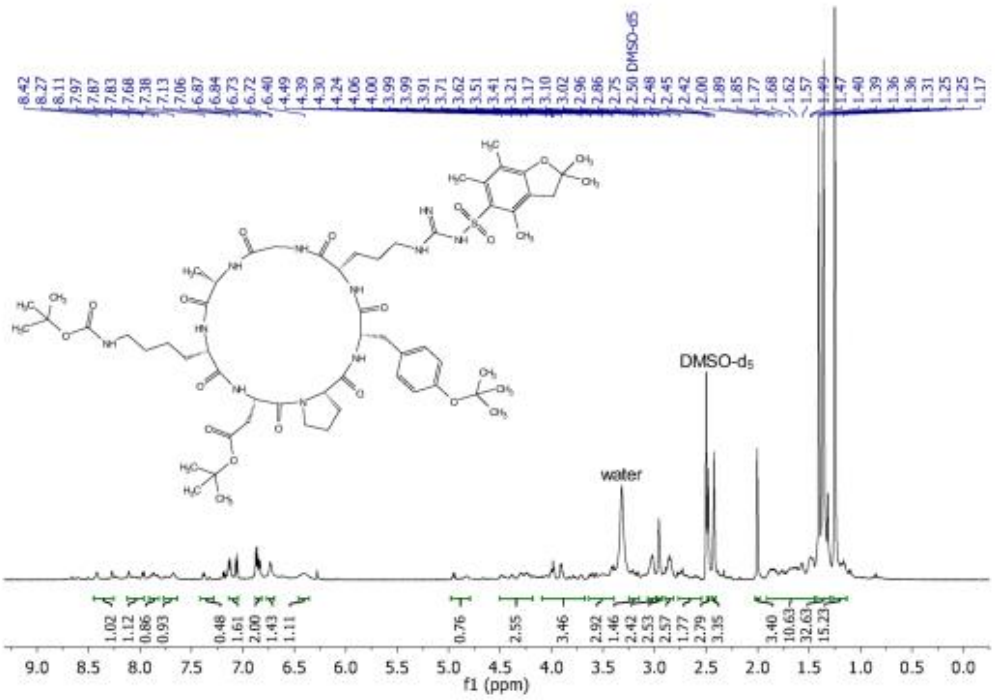


Figure A103 ^{1}H NMR spectrum for cyclic protected peptide 3c, cyc(AK(Boc)D(tBu)PY(tBu)R(Pbf)G), recorded at 600 MHz

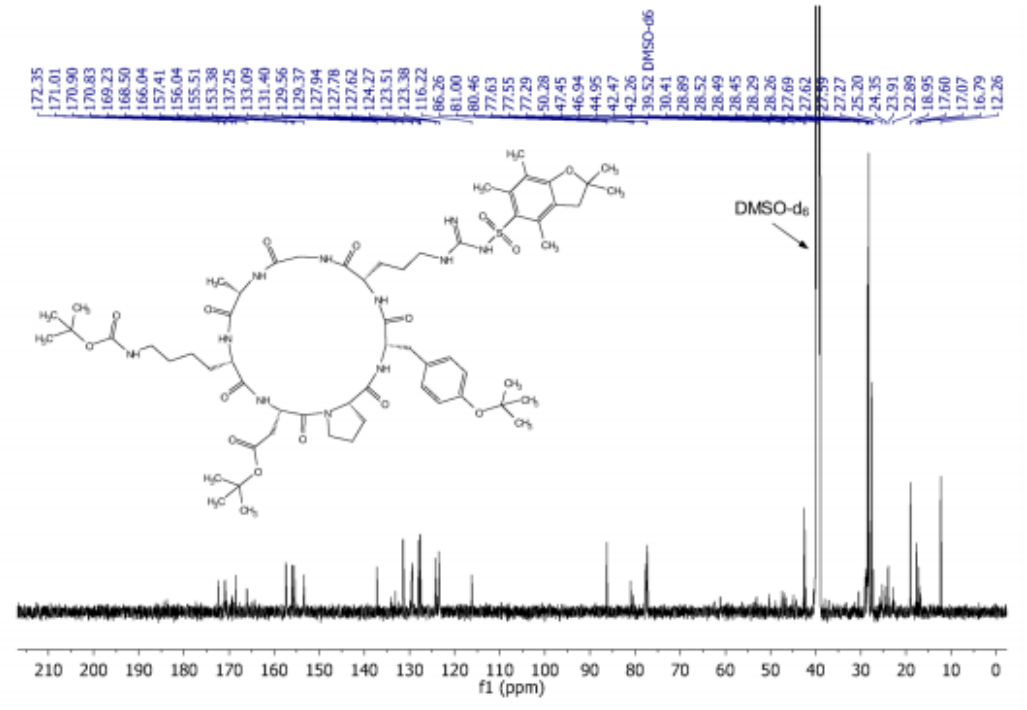


Figure A104 ^{13}C NMR spectrum for cyclized protected peptide 3c, cyc(AK(Boc)D(tBu)PY(tBu)R(Pbf)G), recorded at 151 MHz

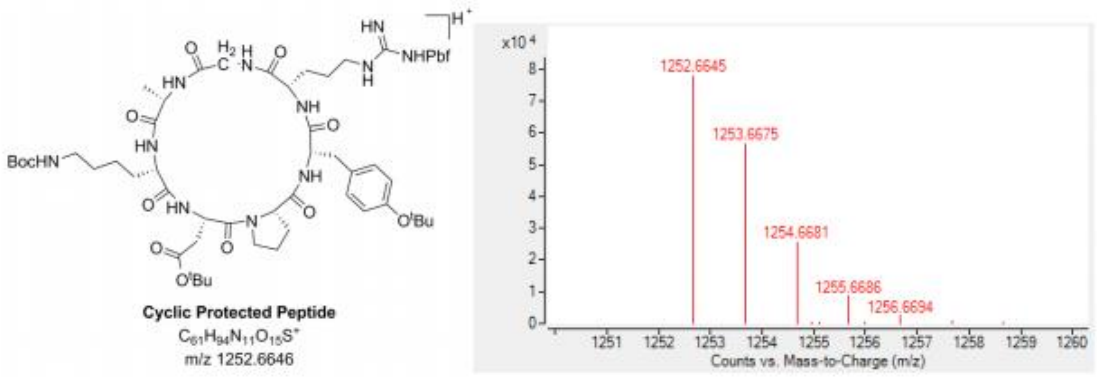


Figure A105 HRMS (+ESI) data for cyclized protected peptide 3c, cyc(AK(Boc)D(tBu)PY(tBu)R(Pbf)G)

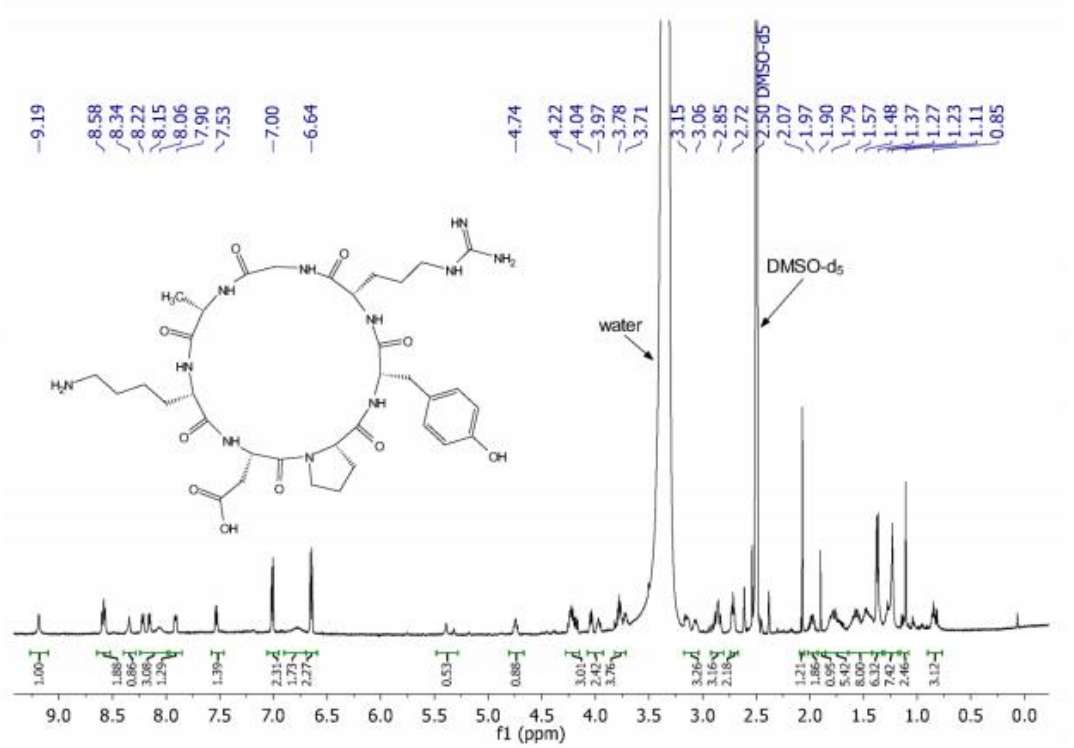


Figure A106 ^{1}H NMR spectrum for cyclic unprotected peptide 3d, cyc(AKDPYRG), recorded at 600 MHz

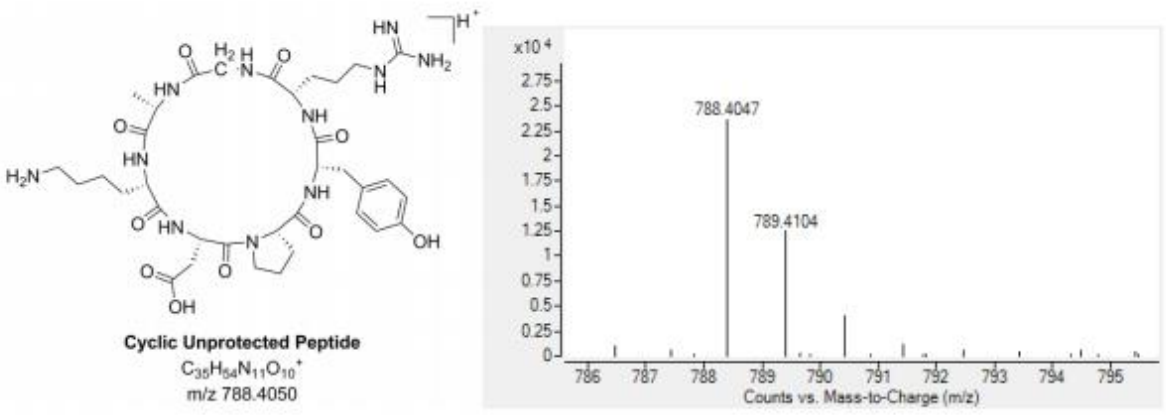


Figure A107 HRMS (+ESI) data for cyclized unprotected peptide 3d, cyc(AKDPYRG)

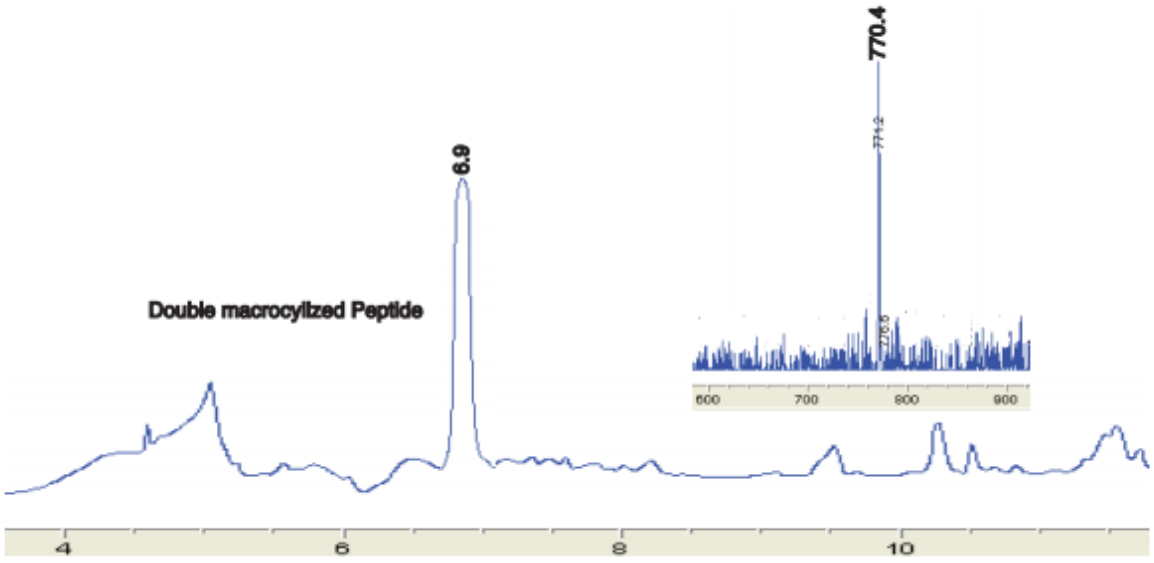


Figure A108 Synthesis of macrocyclic peptide from unprotected linear peptide acid AKDPYRG obtained from Wang resin. Various possibilities of the cyclization of linear unprotected peptide AKDPYRG. Double cyclization of unprotected peptide cyc2(AKDPYRG) was observed as shown by LCMS.

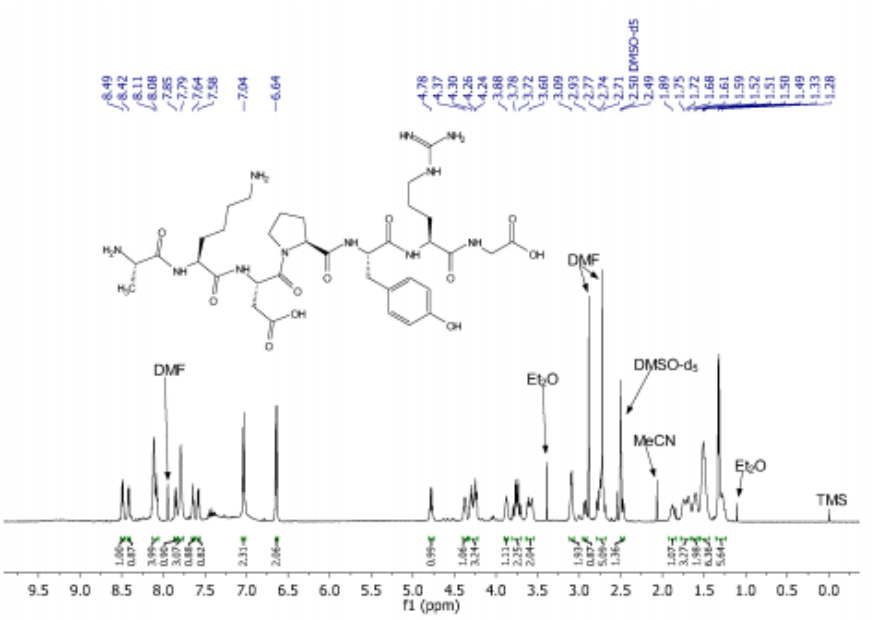


Figure A109 1 H NMR spectrum for linear unprotected peptide, AKDPYRG, recorded at 600 MHz

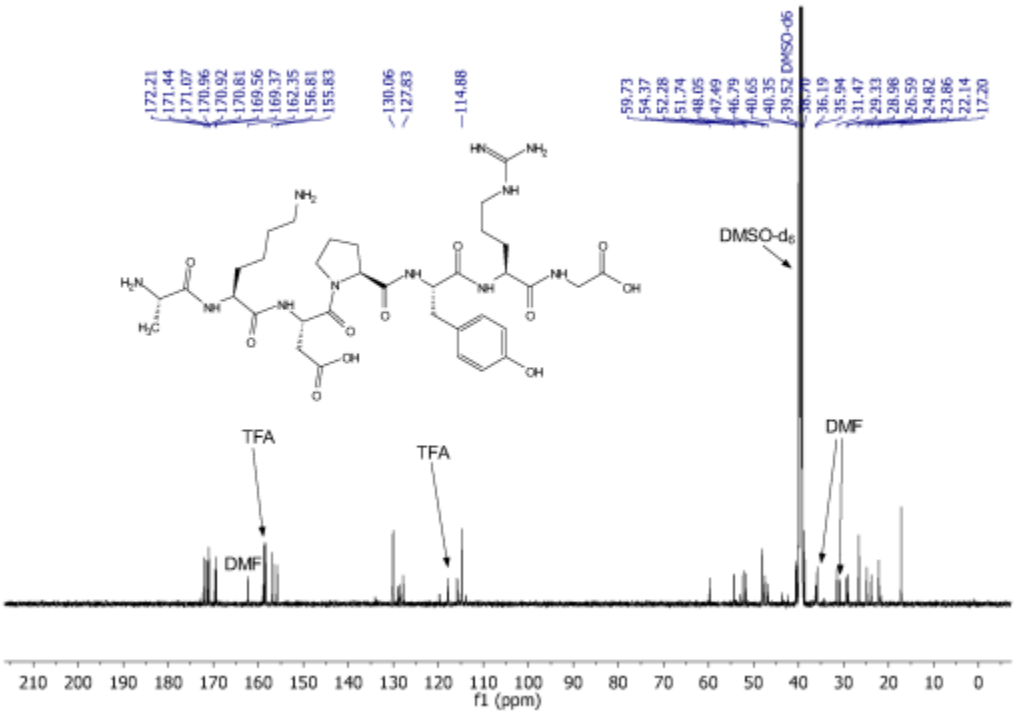


Figure A110 ^{13}C NMR spectrum for linear unprotected peptide, AKDPYRG, recorded at 151 MHz

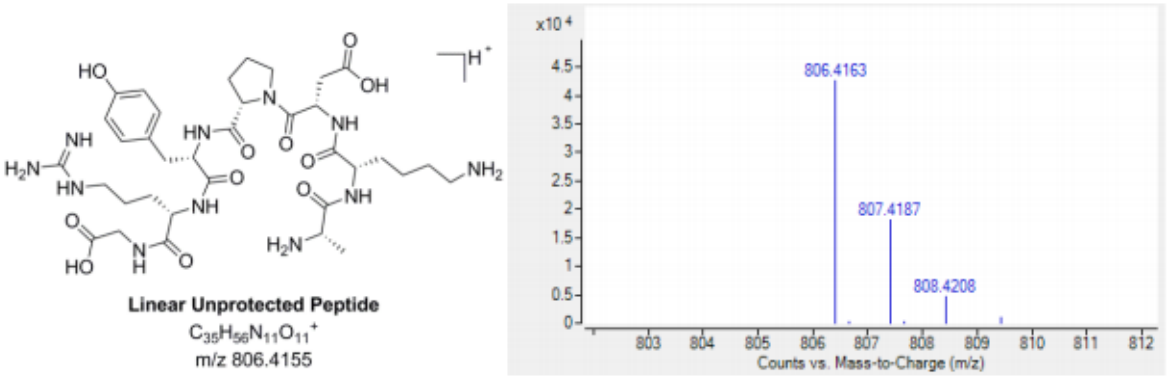


Figure A111 HRMS (+ESI) data for linear unprotected peptide, AKDPYRG

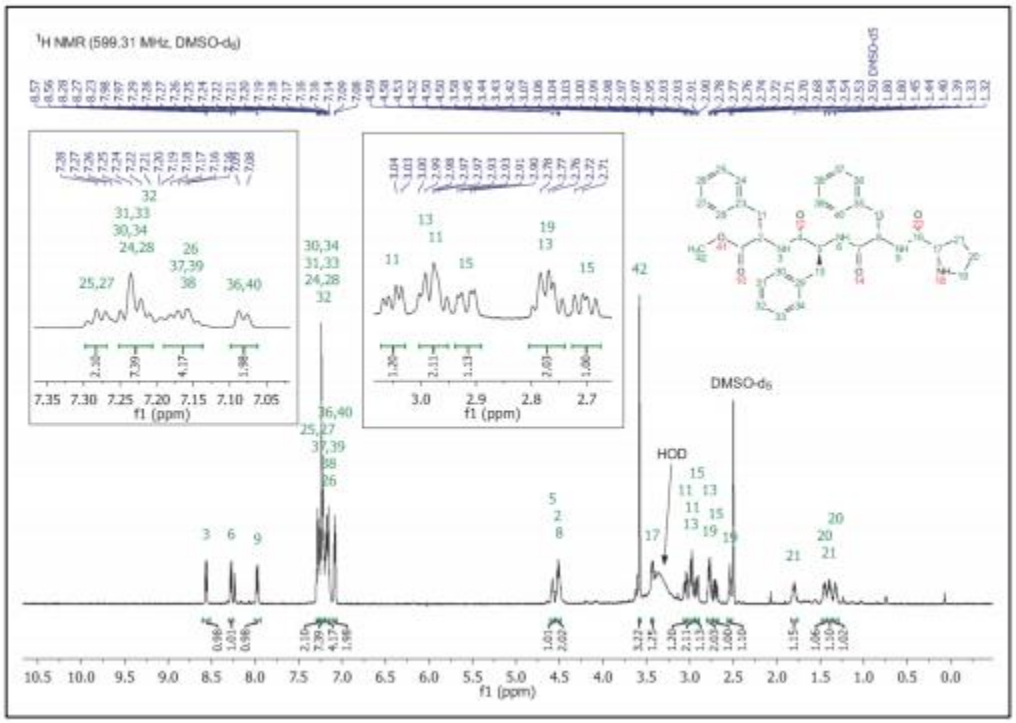


Figure A112 ^{1}H NMR spectrum for catalyst PFFFOMe 5', recorded at 600 MHz

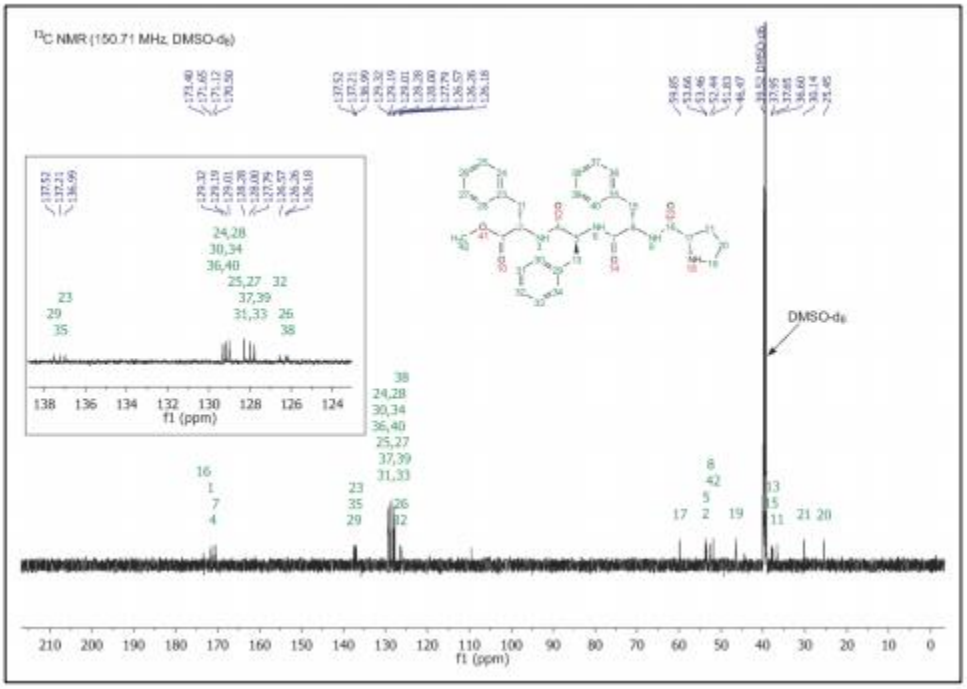


Figure A113 ^{13}C NMR spectrum for catalyst PFFFOMe 5', recorded at 151 MHz

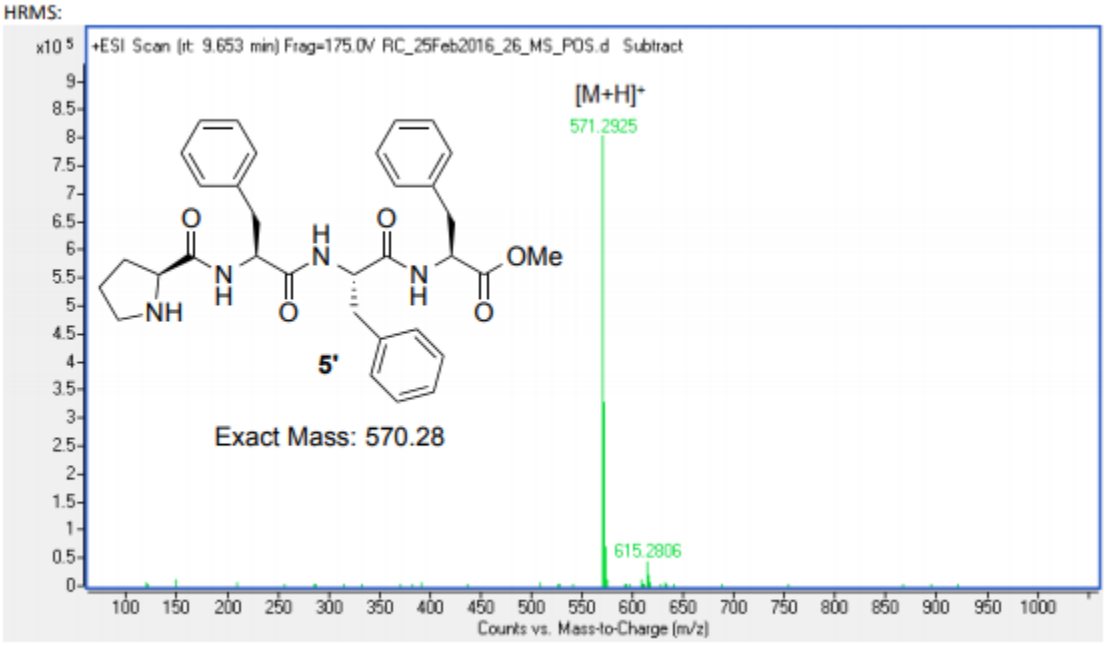


Figure A114 HRMS (+ESI) data for catalyst PFFFOMe 5'

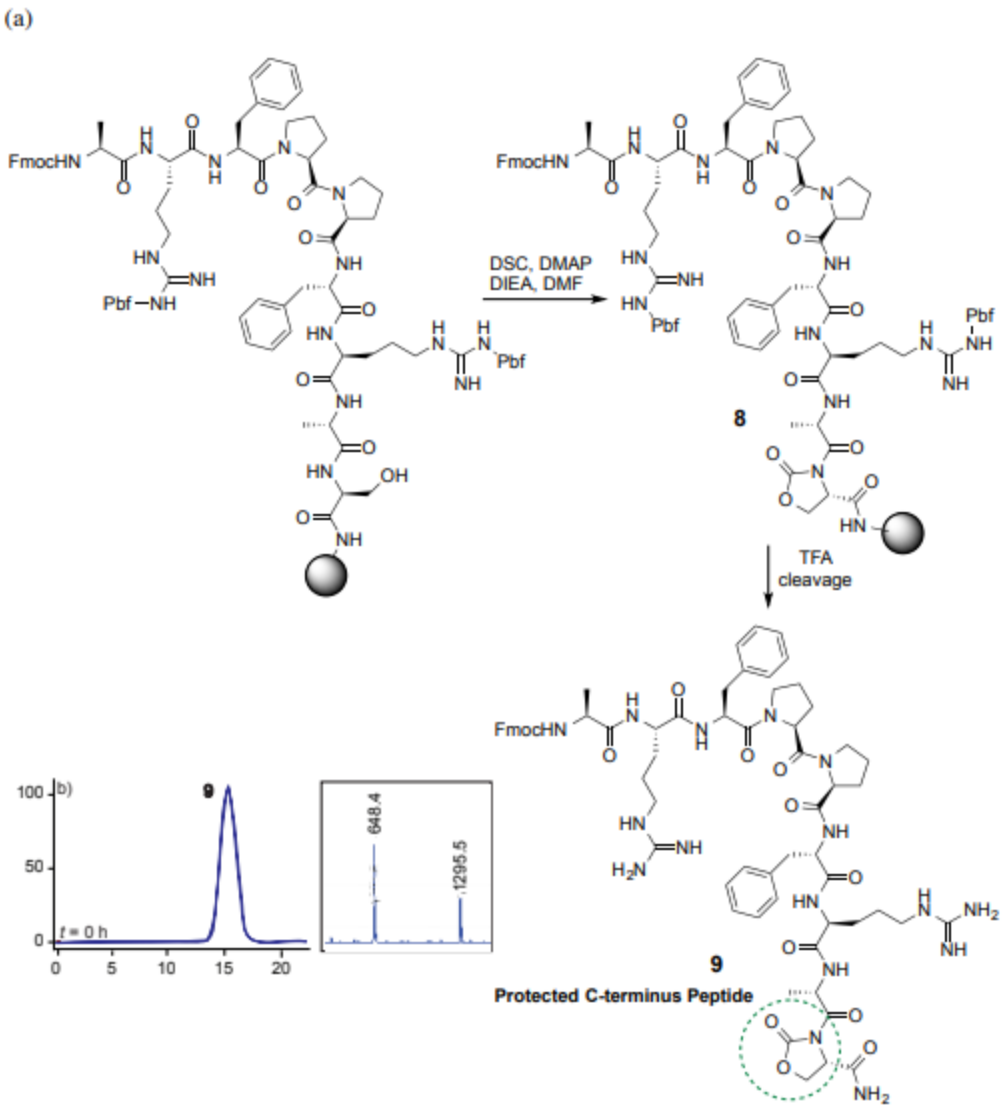


Figure A115 (a) Synthesis of peptide Fmoc-ARFPPFRAOxd, 9 with semi-permanent protecting group on the C-terminus using cyclic urethane technique (CUT). (b) HPLC and MS chromatogram of C-terminal protected peptide Fmoc ARFPPFRA-Oxd, 9.

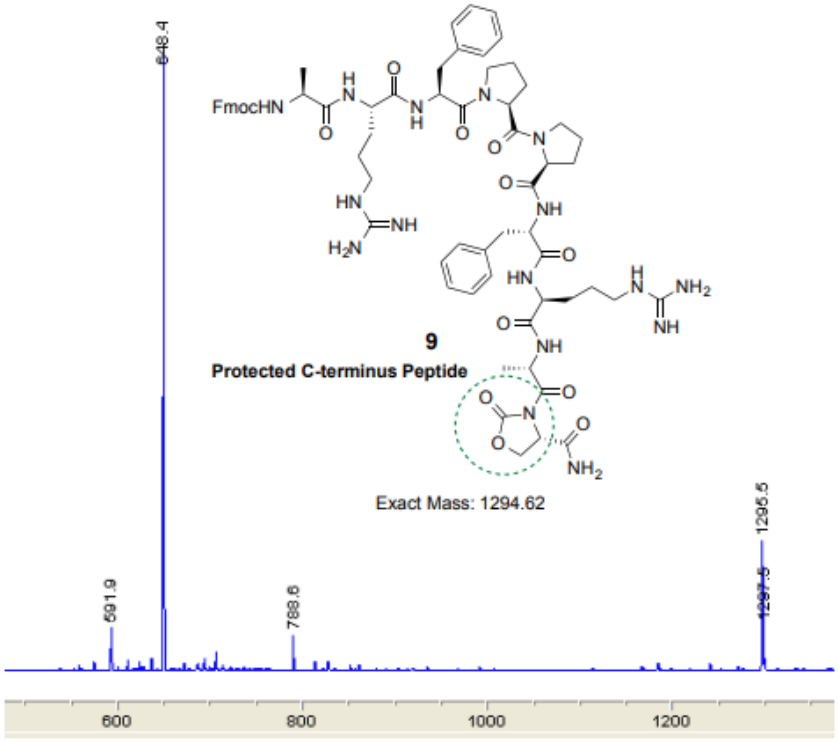


Figure A116 MS data for protected C-terminal peptide Fmoc-ARFPPFRA-Oxd, 9.

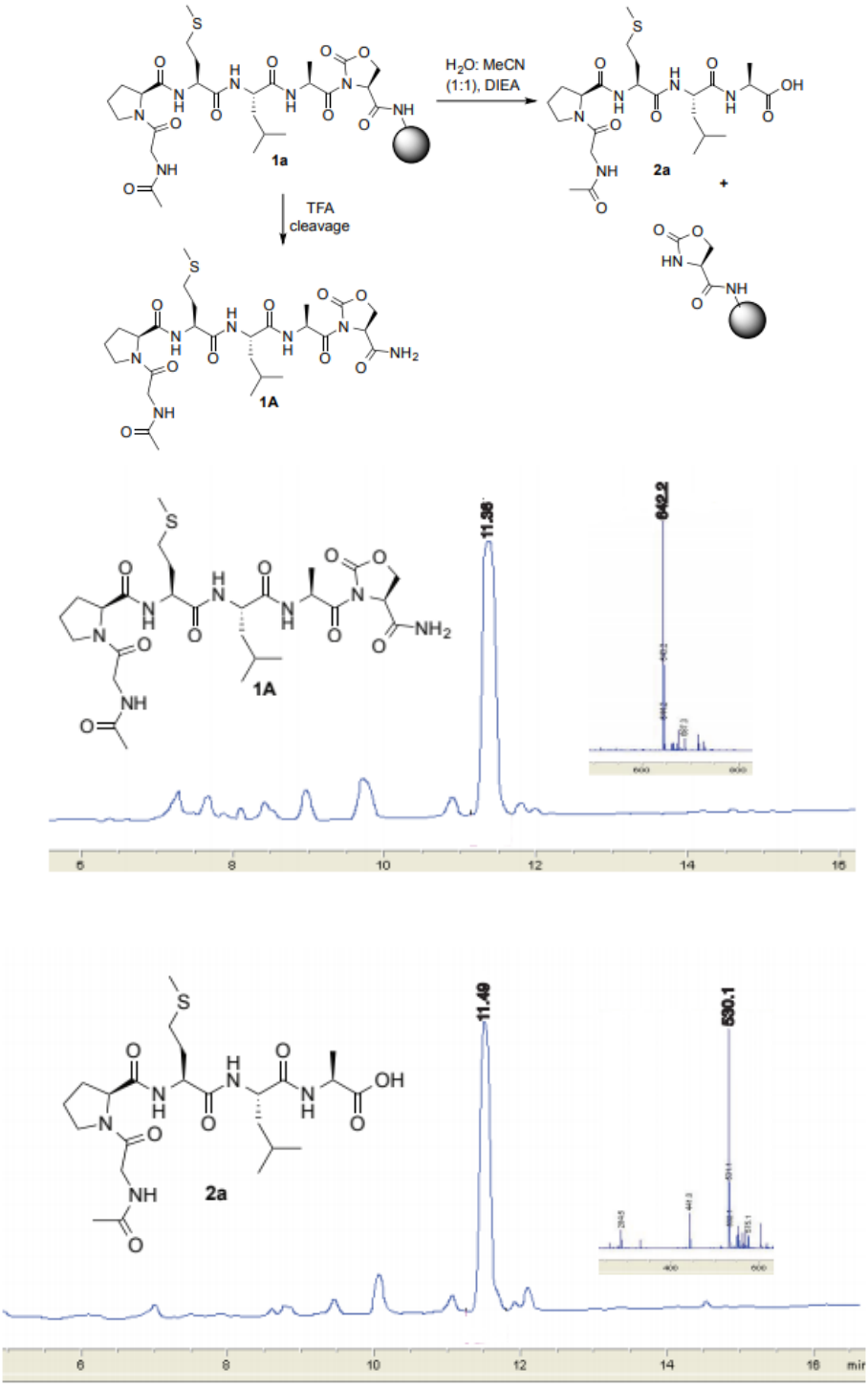


Figure A117 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Gly-Pro-Met-Leu-Ala-Oxd (1A). LCMS: m/z 642.20 (calcd [M+H]+ = 642.28), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.36
Ac-Gly-Pro-Met-Leu-Ala (2a). LCMS: m/z 530.1 (calcd [M+H]+ = 530.26), Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.49 min

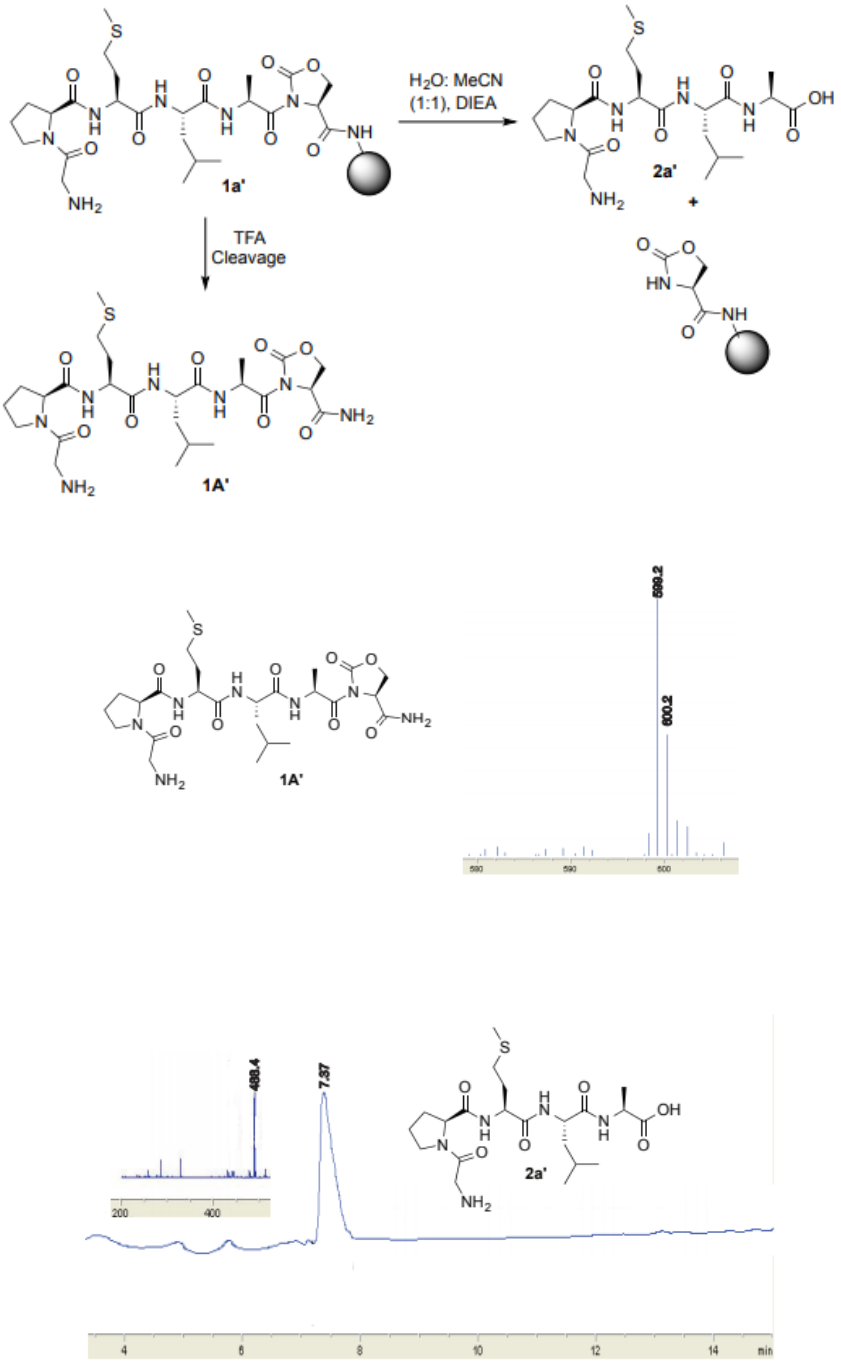


Figure A118 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Gly-Pro-Met-Leu-Ala-Oxd (1A'). LCMS: m/z 600.2 (calcd [M+H]+ = 600.27), Purity: >95% (HPLC analysis at 220 nm). Retention time: 8.11 min
Gly-Pro-Met-Leu-Ala (2a'). LCMS: m/z 488.40 (calcd [M+H]+ = 488.25), Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.37 min

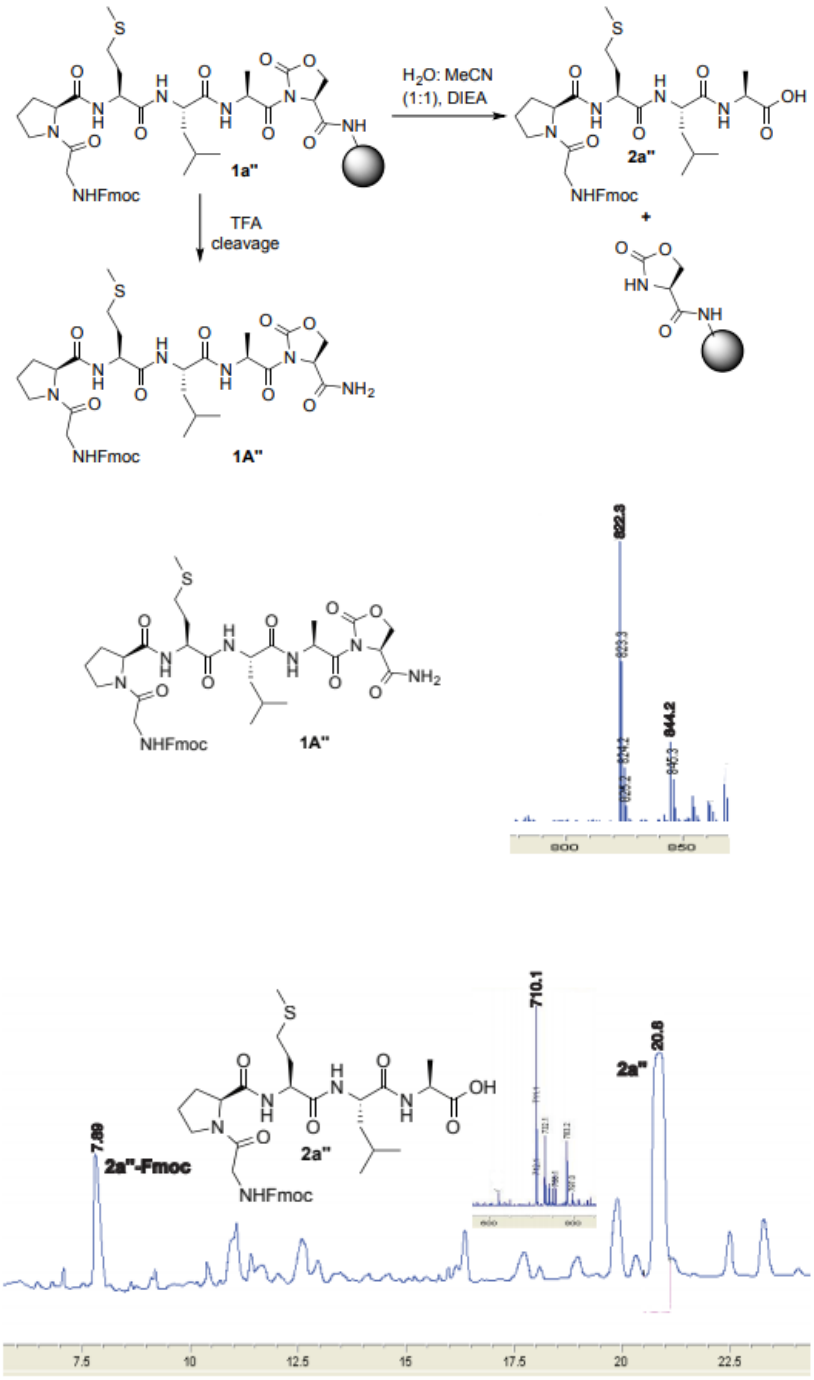


Figure A119 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.

Fmoc-Gly-Pro-Met-Leu-Ala-Oxd (1A''). LCMS: m/z 822.3 (calcd [M+H]+ = 822.34), 844.2 (calcd [M+Na]+ = 844.34); Purity: >95% (HPLC analysis at 220 nm). Retention time: 18.2 min. Fmoc-Gly-Pro-Met-Leu-Ala (2a''). LCMS: m/z 710.1 (calcd [M+H]+ = 710.31), (HPLC analysis at 220 nm). Retention time: 20.8 min. Gly-Pro-Met-Leu-Ala (2a''-Fmoc or 2a'). LCMS: m/z 488.4 (calcd [M+H]+ = 488.25), (HPLC analysis at 220 nm). Retention time: 7.89 min.

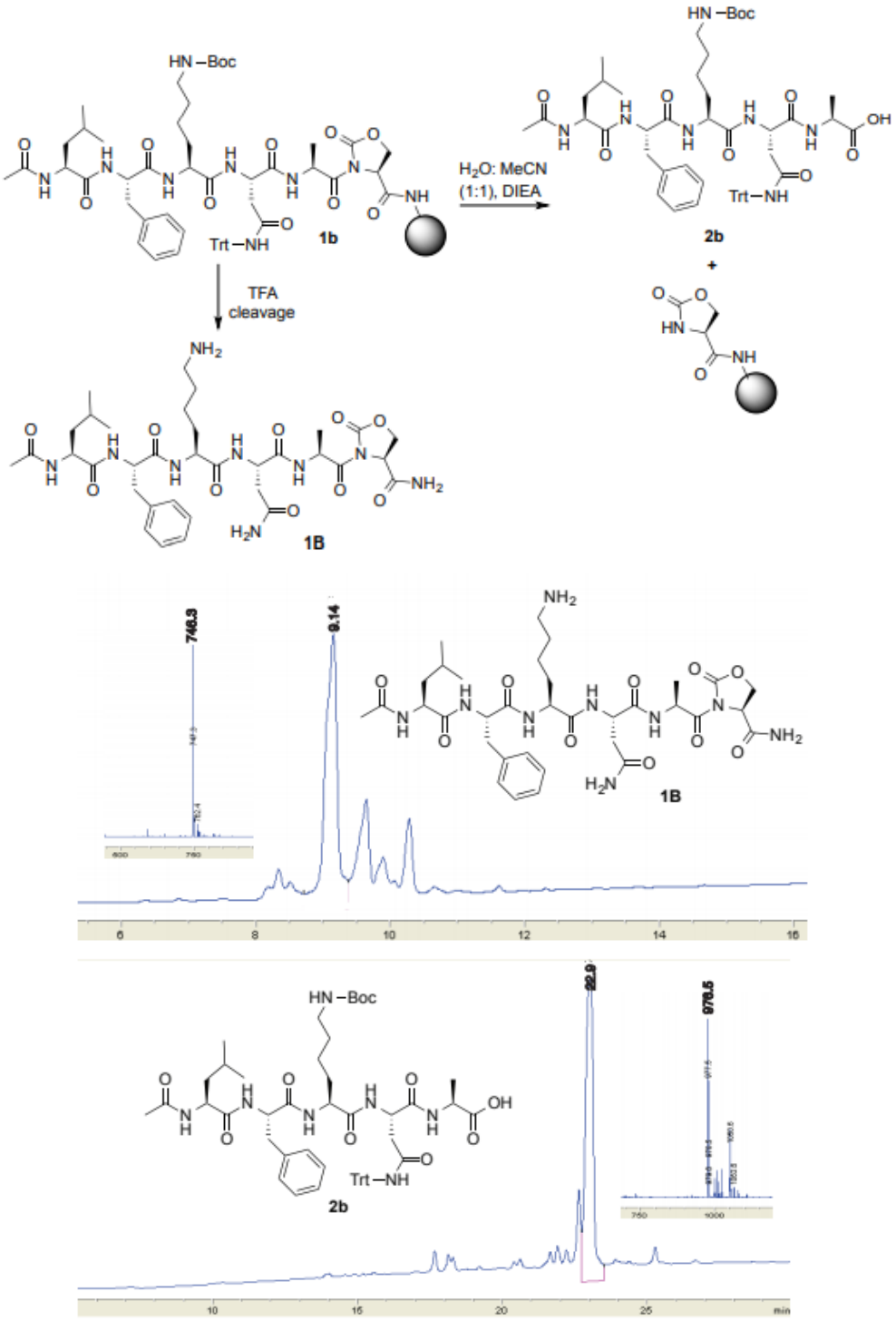


Figure A120 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Leu-Phe-Lys-Asn-Ala-Oxd (1B). LCMS: m/z 746.3 (calcd [M+H]+ = 746.38), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.14 min.
Ac-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2b). LCMS: m/z 976.5 (calcd [M+H]+ = 976.51), Purity: >95% (HPLC analysis at 220 nm). Retention time: 22.97 min.

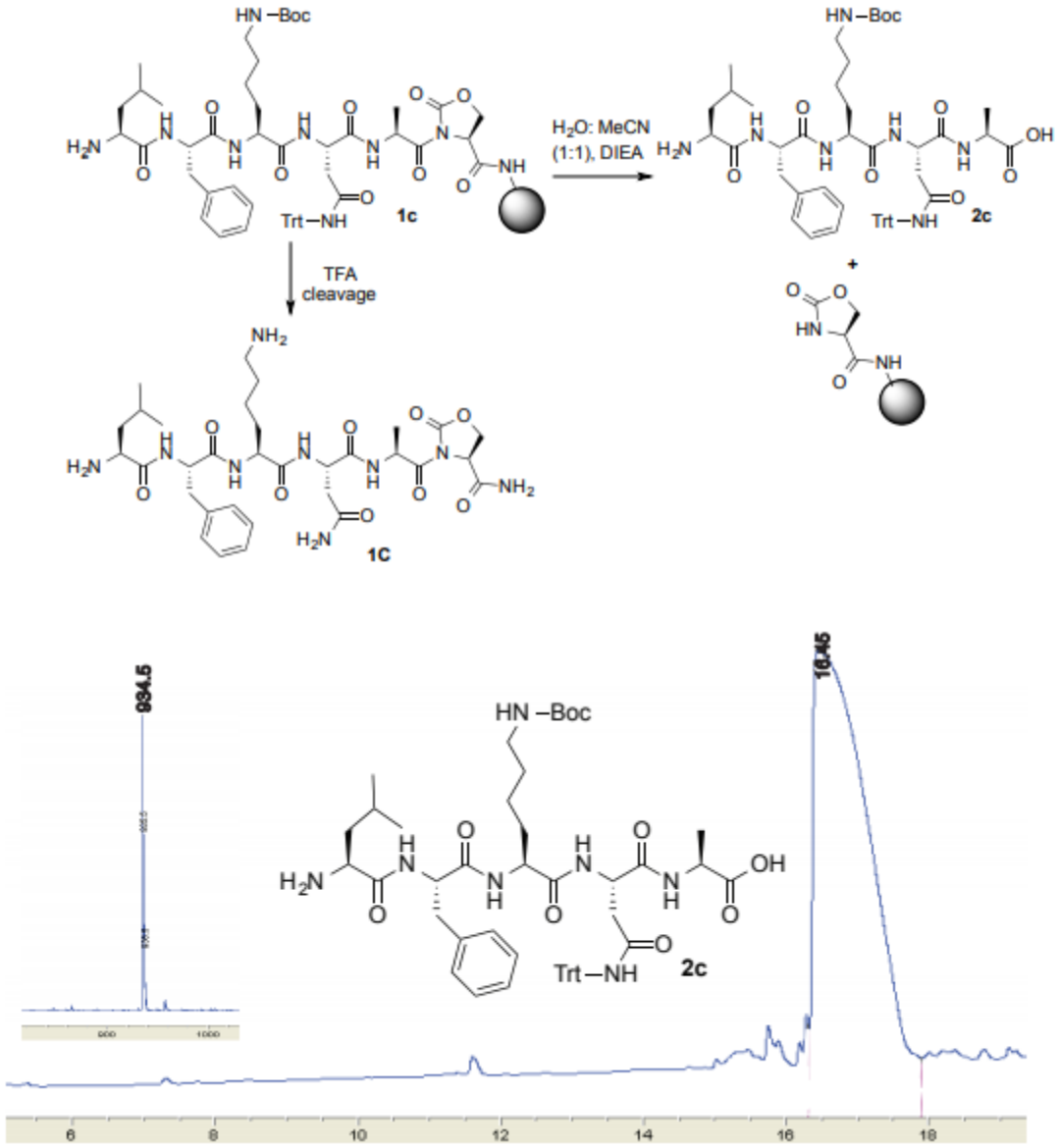


Figure A121 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.

Leu-Phe-Lys-Asn-Ala-Oxd (1C). LCMS: m/z 704.3 (calcd [M+H]+ = 704.37), Purity: >95% (HPLC analysis at 220 nm).

Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2c). LCMS: m/z 934.5 (calcd [M+H]+ = 934.57), Purity: >95% (HPLC analysis at 220 nm). Retention time: 16.45 min.

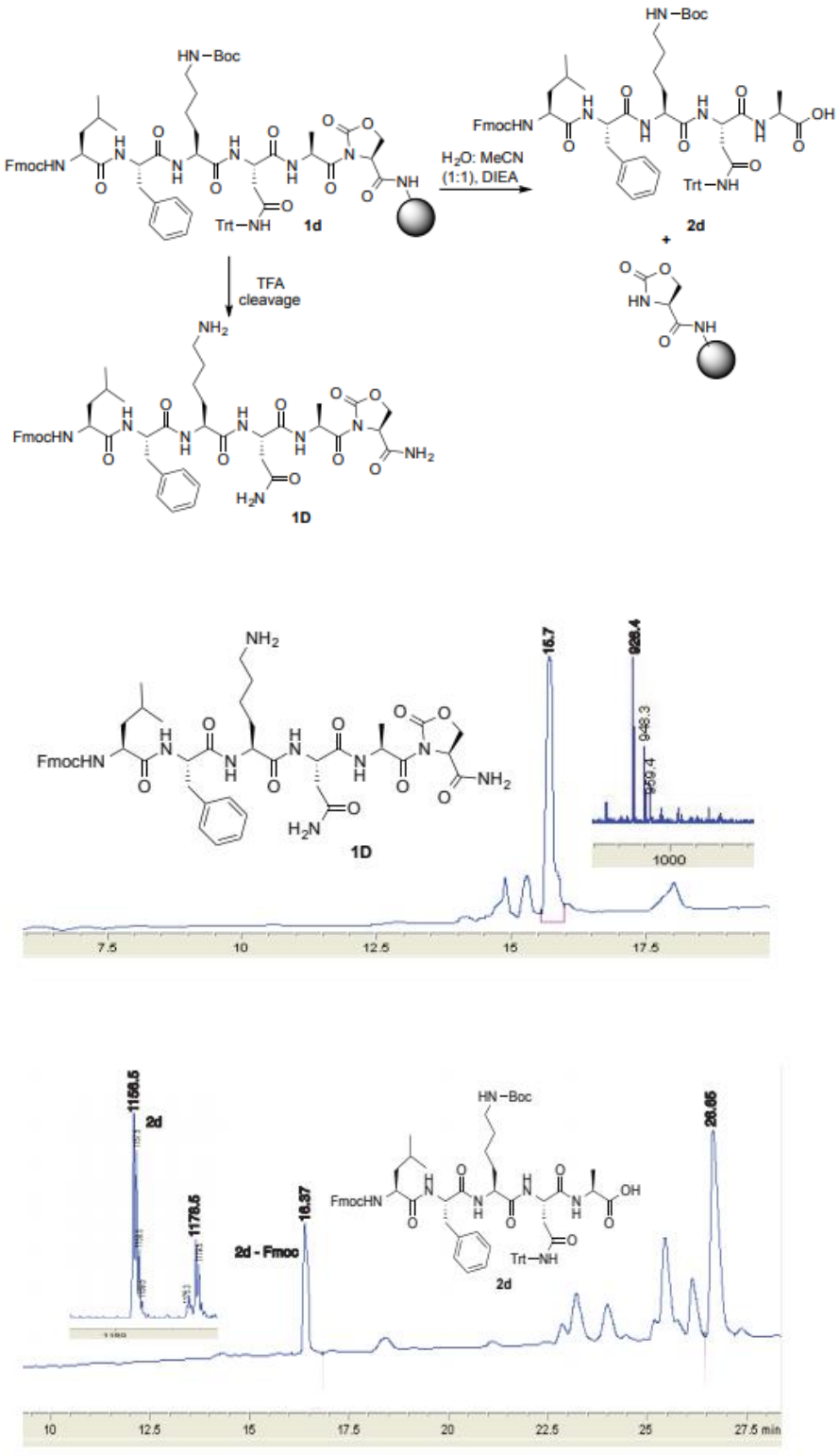


Figure A122 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Fmoc-Leu-Phe-Lys-Asn-Ala-Oxd (1D). LCMS: m/z 926.4 (calcd [M+H]+ = 926.43), 948.3 (calcd [M+Na] + = 948.43), Purity: >95% (HPLC analysis at 220 nm). Retention time: 15.7 min.

Fmoc-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2d). LCMS: m/z 1156.5 (calcd [M+H]+ = 1156.57), 1178.5 (calcd [M+Na] + = 1178.57), Purity: >95% (HPLC analysis at 220 nm). Retention time: 26.65 min.
Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2d-Fmoc or 2c). LCMS: m/z 934.5 (calcd [M+H]+ = 934.57), Purity: >95% (HPLC analysis at 220 nm). Retention time: 16.37 min.

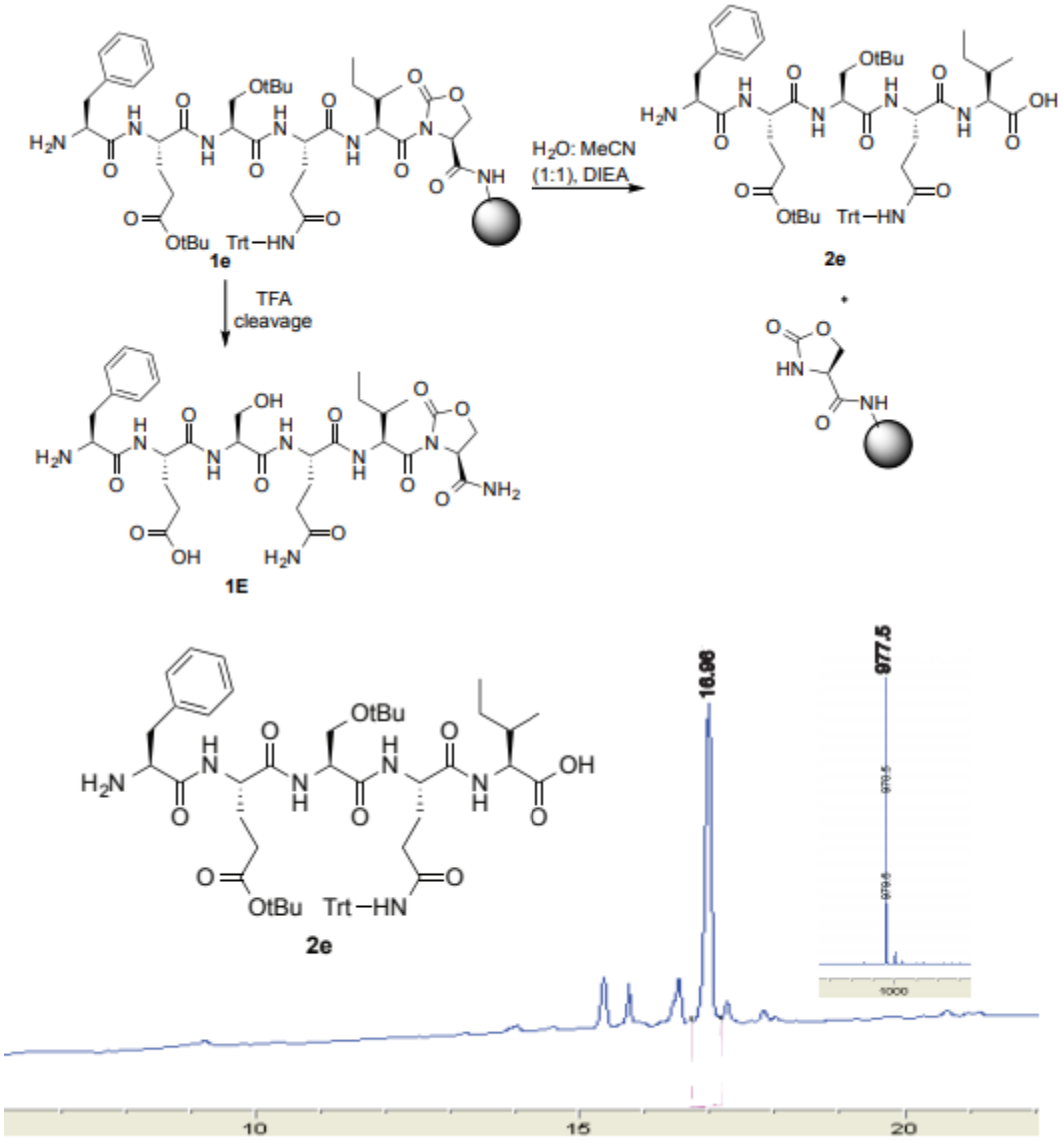


Figure A123 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Phe-Glu-Ser-Gln-Ile-Oxd (1E). LCMS: m/z 735.3 (calcd [M+H]+ = 735.32), Purity: >95% (HPLC analysis at 220 nm).
Phe-Glu(tBu)-Ser(tBu)-Gln(Trt)-Ile (2e). LCMS: m/z 977.5 (calcd [M+H]+ = 977.53), Purity: >95% (HPLC analysis at 220 nm). Retention time: 16.96 min.

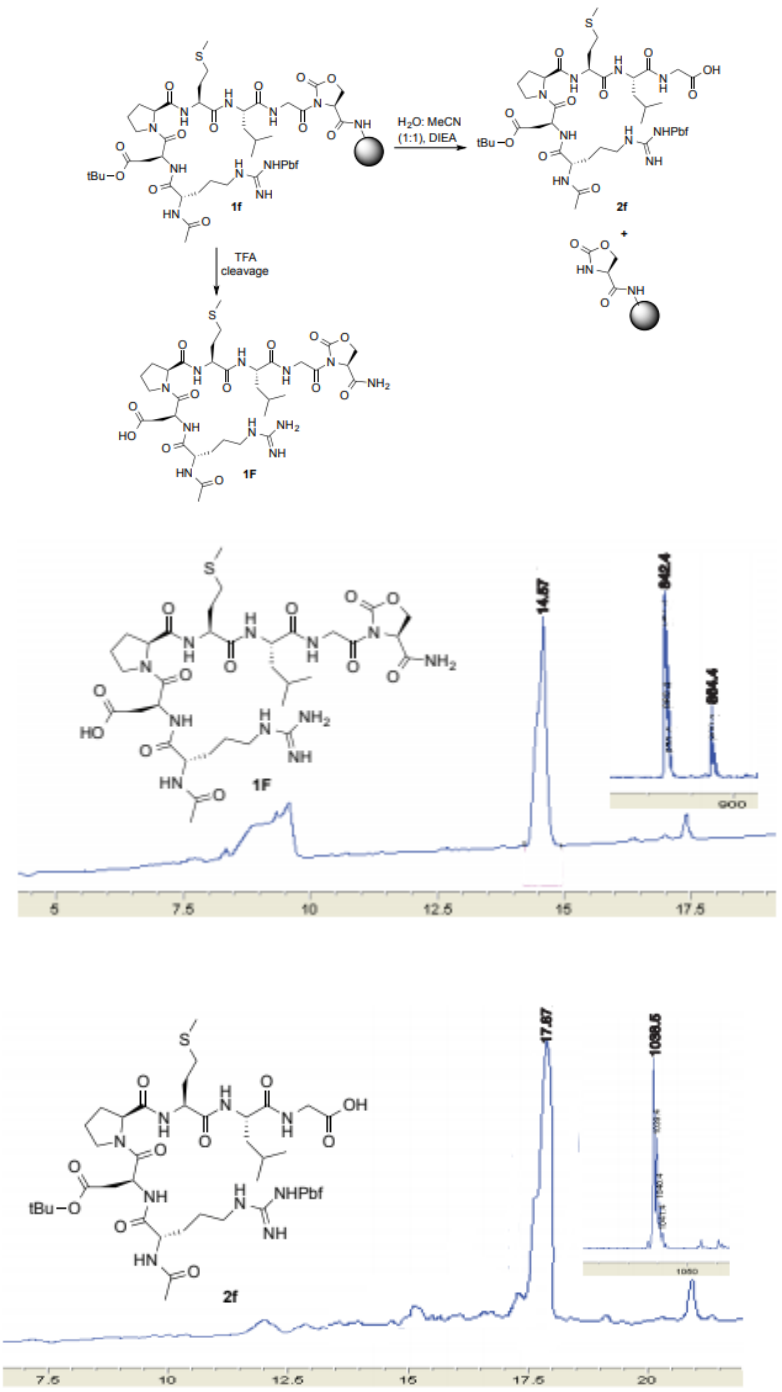


Figure A124 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Arg-Asp-Pro-Met-Leu-Gly-Oxd (1F). LCMS: m/z 842.4 (calcd [M+H]+ = 842.38), 864.4 (calcd [M+Na] + = 864.38), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.57 min.
Ac-Arg(Pbf)-Asp(tBu)-Pro-Met-Leu-Gly (2f). LCMS: m/z 1038.5 (calcd [M+H]+ = 1038.49), Purity: >95% (HPLC analysis at 220 nm). Retention time: 17.87 min.

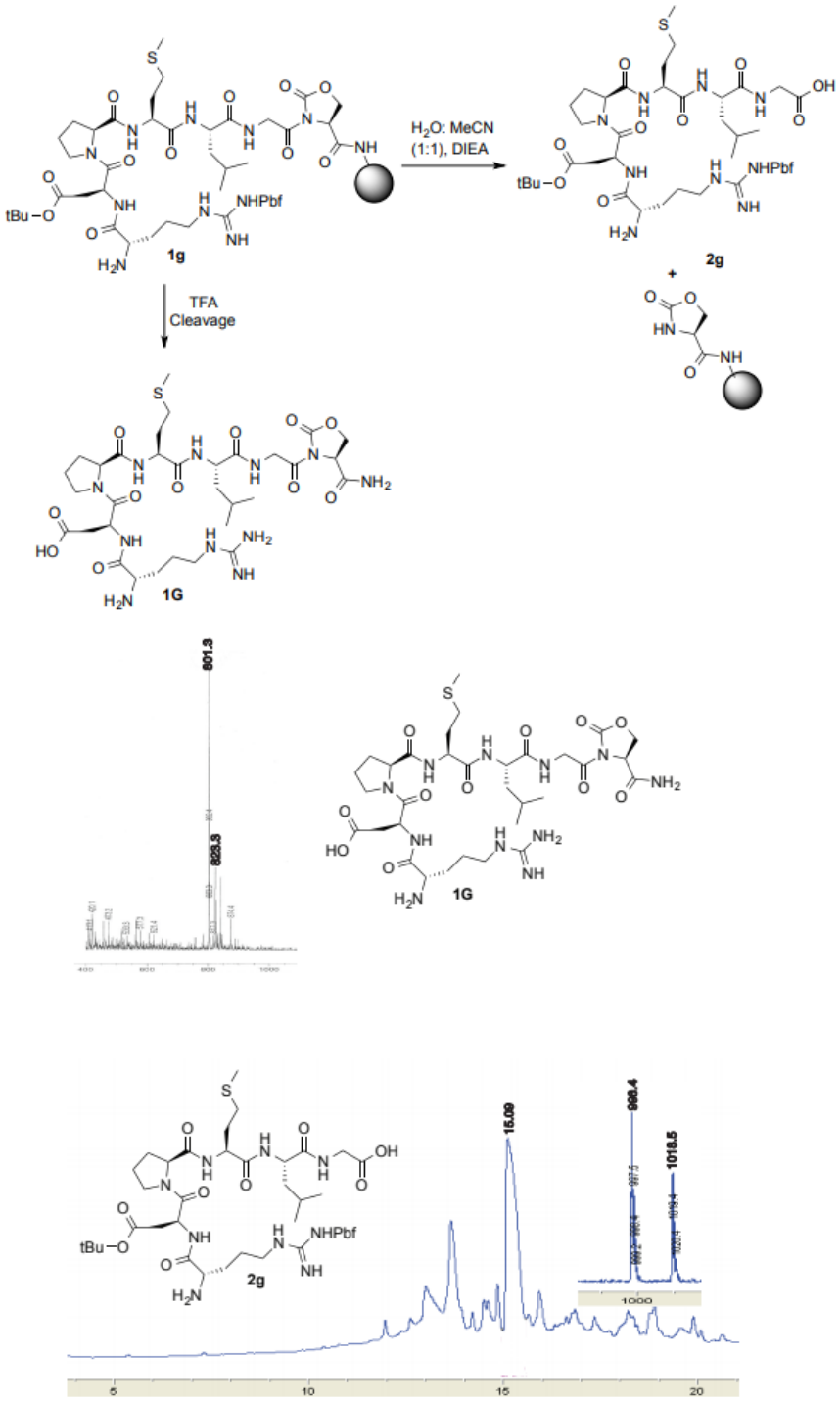


Figure A125 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Arg-Asp-Pro-Met-Leu-Gly-Oxd (1G). LCMS: m/z 801.3 (calcd [M+H]+ = 801.36), 823.3 (calcd [M+Na]+ = 823.36) Purity: >95% (HPLC analysis at 220 nm).
Arg(Pbf)-Asp(tBu)-Pro-Met-Leu-Gly (2g). LCMS: m/z 996.4 (calcd [M+H]+ = 996.48), 1018.5 (calcd [M+Na] + = 1018.48), Purity: >95% (HPLC analysis at 220 nm). Retention time: 15.09 min

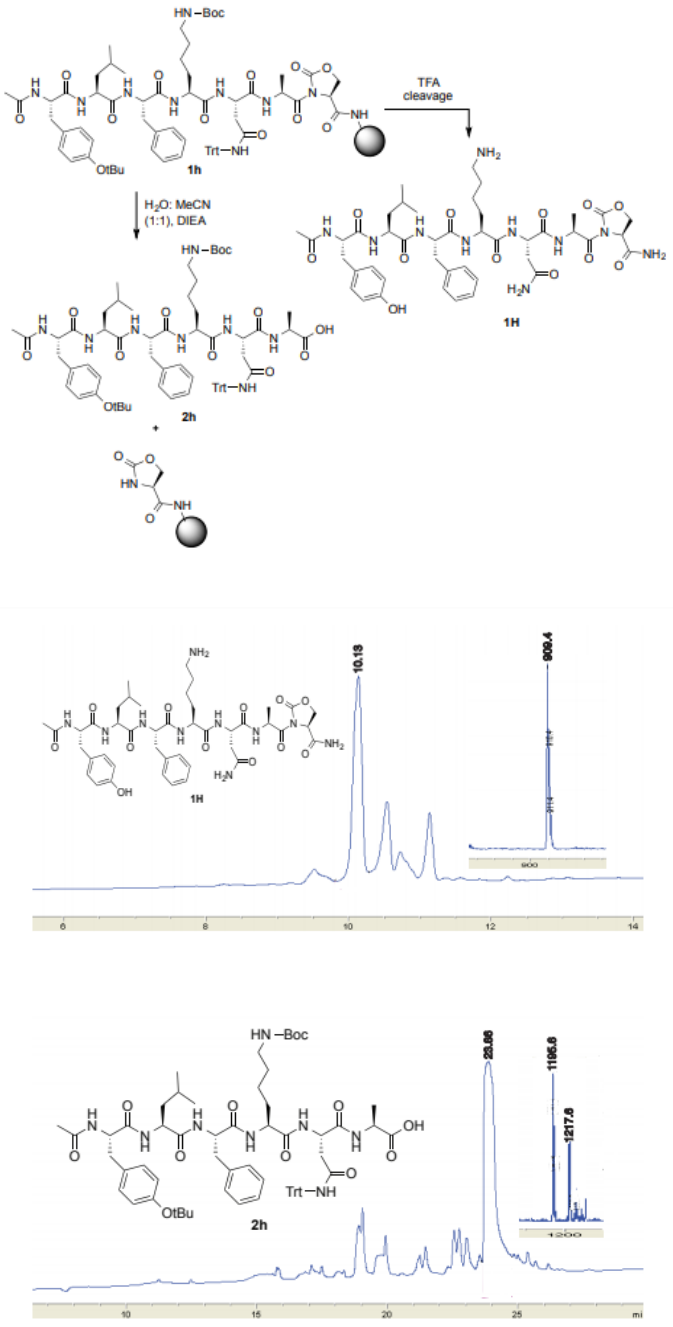


Figure A126 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Tyr-Leu-Phe-Lys-Asn-Ala-Oxd (1H). LCMS: m/z 909.4 (calcd [M+H]+ = 909.44) Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.13 min.
Ac-Tyr(tBu)-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2h). LCMS: m/z 1195.6 (calcd [M+H]+ = 1195.64), 1217.6 (calcd [M+Na]+ = 1217.64) Purity: >95% (HPLC analysis at 220 nm). Retention time: 23.86 min

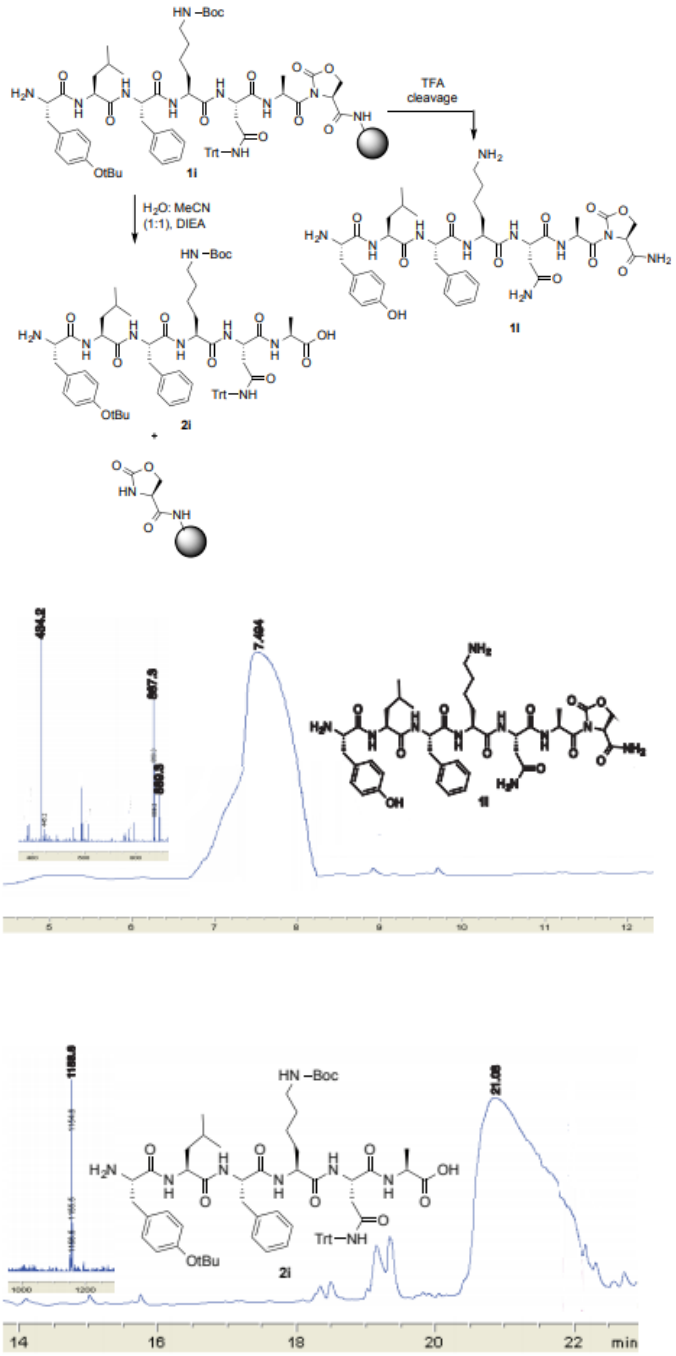


Figure A127 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Tyr-Leu-Phe-Lys-Asn-Ala-Oxd (1I). LCMS: m/z 867.3 (calcd [M+H]+ = 867.43), 434.2 (calcd [M+2H]2+ = 434.21), 889.3 (calcd [M+Na]+ = 889.43), Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.49 min.
Tyr(tBu)-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2i). LCMS: m/z 1153.6 (calcd [M+H]+ = 1153.63), Purity: >95% (HPLC analysis at 220 nm). Retention time: 21.08 min.

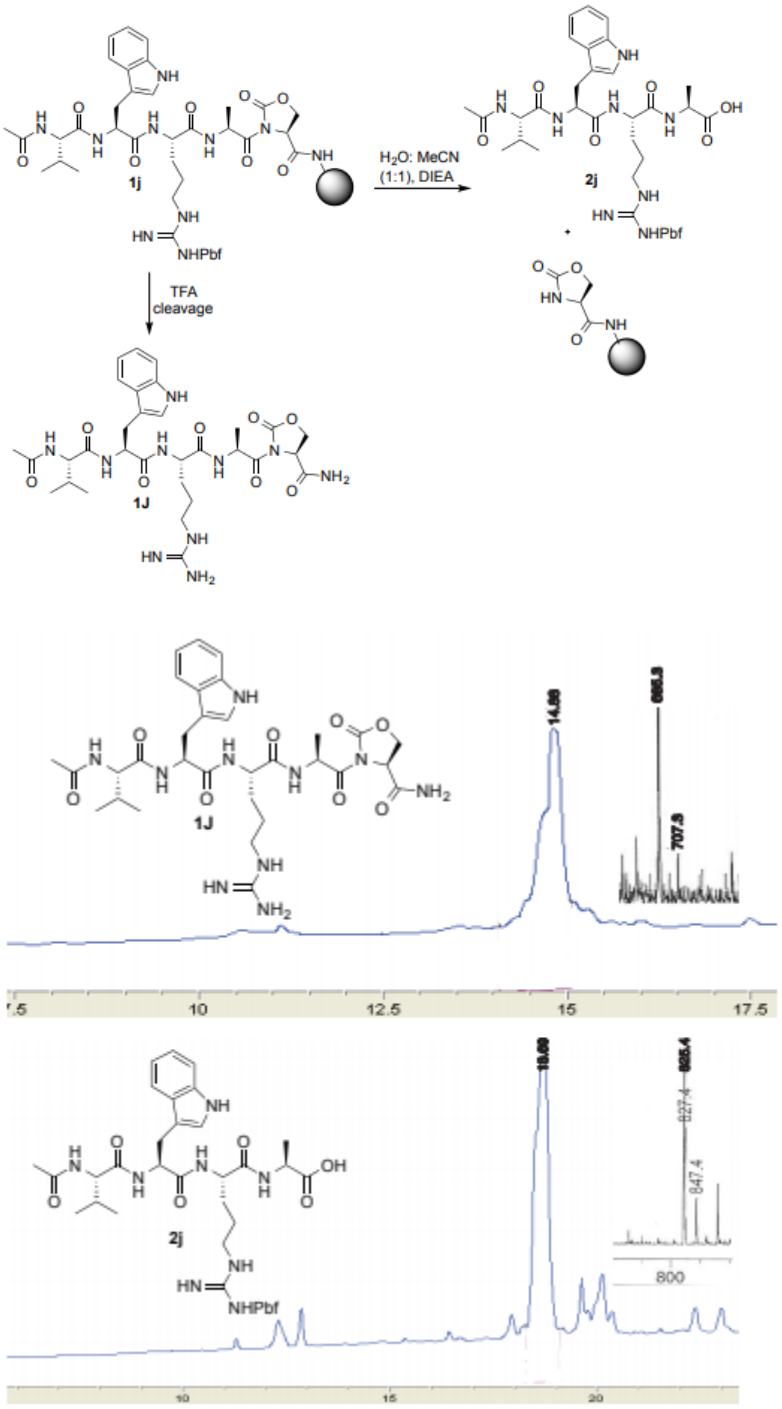


Figure A128 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Val-Trp-Arg-Ala-Oxd (1J). LCMS: m/z 685.3 (calcd [M+H]+ = 685.76), 707.4 (calcd [M+Na] + = 707.76), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.88 min.
Ac-Val-Trp-Arg(Pbf)-Ala (2j). LCMS: m/z 825.4 (calcd [M+H]+ = 825.63), 847.4 (calcd [M+Na] + = 847.63), Purity: >95% (HPLC analysis at 220 nm). Retention time: 18.69 min.

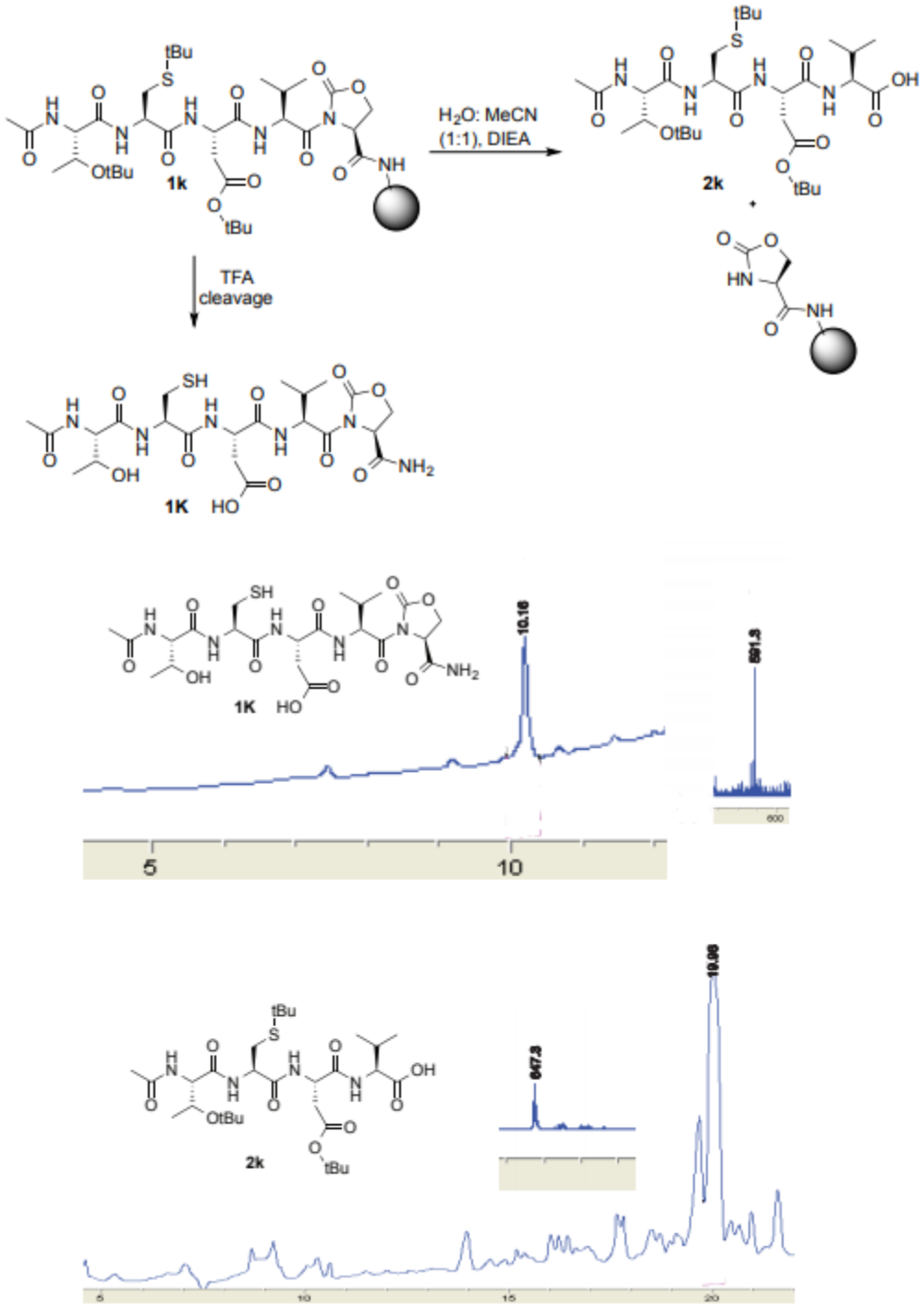


Figure A129 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Thr-Cys-Asp-Val-Oxd (1K). LCMS: m/z 591.3 (calcd [M+H]+ = 591.5), Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.16 min.
Ac-Thr(tBu)-Cys(tBu)-Asp(tBu)-Val (2k). LCMS: m/z 647.3 (calcd [M+H]+ = 647.36), Purity: >95% (HPLC analysis at 220 nm). Retention time: 19.98 min.

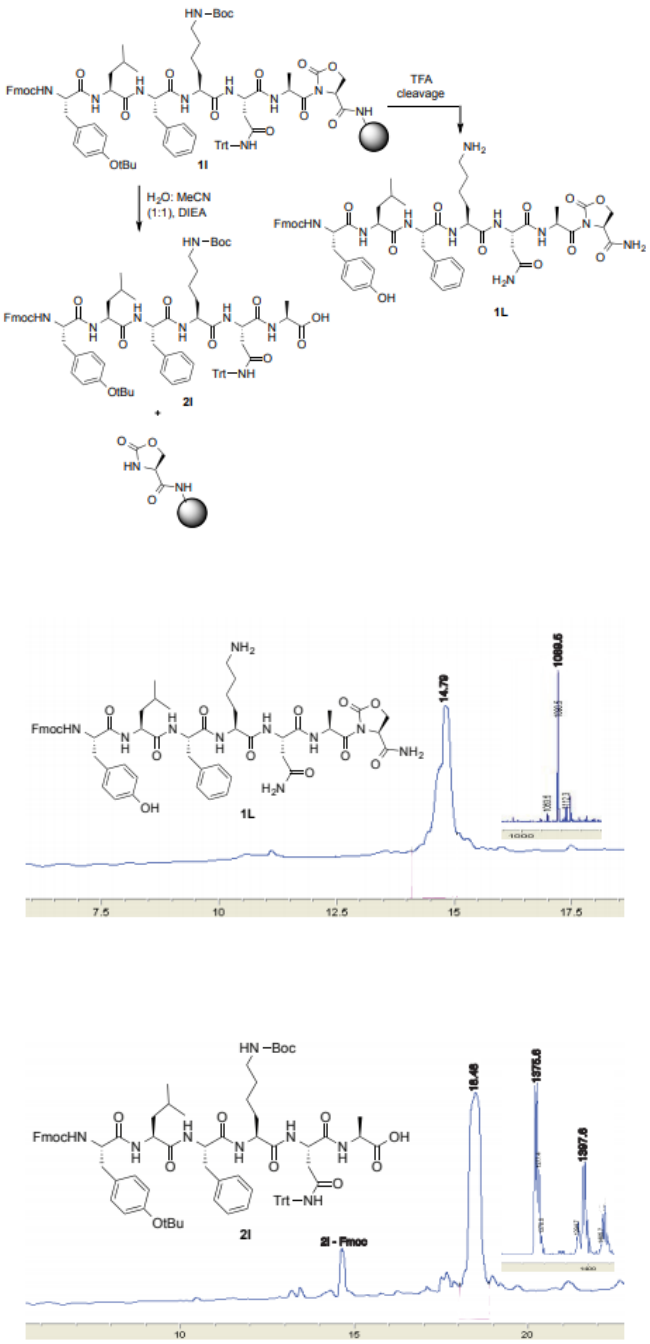


Figure A130 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.

Fmoc-Tyr-Leu-Phe-Lys-Asn-Ala-Oxd (1L). LCMS: m/z 1089.5 (calcd [M+H]+ = 1089.50), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.79 min.

Fmoc-Tyr(tBu)-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2l). LCMS: m/z 1375.6 (calcd [M+H]+ = 1375.69), 1397.6 (calcd [M+Na]+ = 1397.69) Purity: >95% (HPLC analysis at 220 nm). Retention time: 18.48 min

Tyr(tBu)-Leu-Phe-Lys(Boc)-Asn(Trt)-Ala (2l-Fmoc). LCMS: m/z 1153.6 (calcd [M+H]+ = 1153.69), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.9 min

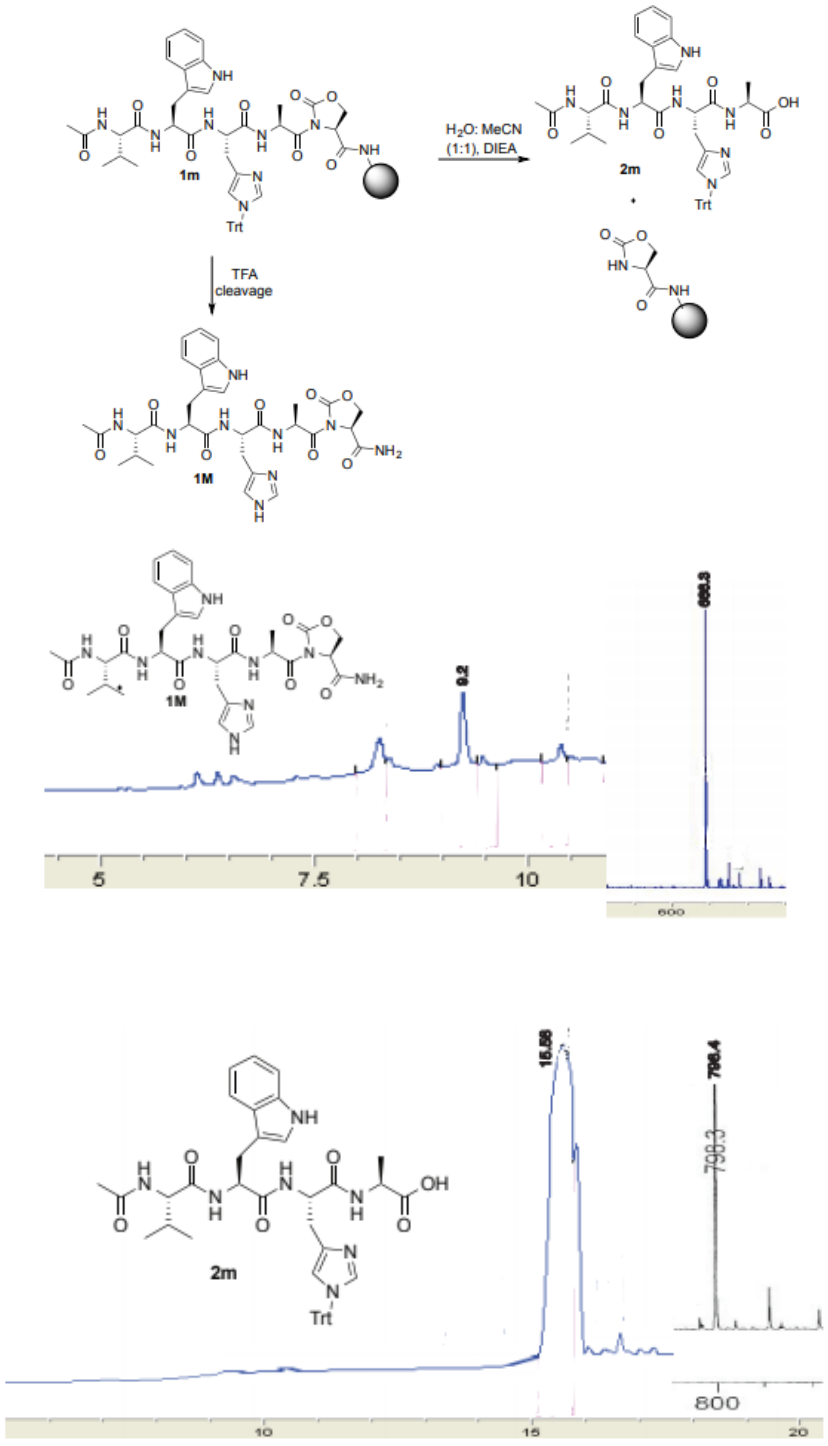


Figure A131 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Val-Trp-His-Ala-Oxd (1M). LCMS: m/z 666.3 (calcd [M+H]+ = 666.6), Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.2 min
Ac-Val-Trp-His(Trt)-Ala (2m). LCMS: m/z 796.4 (calcd [M+H]+ = 796.5), Purity: >95% (HPLC analysis at 220 nm). Retention time: 15.58 min

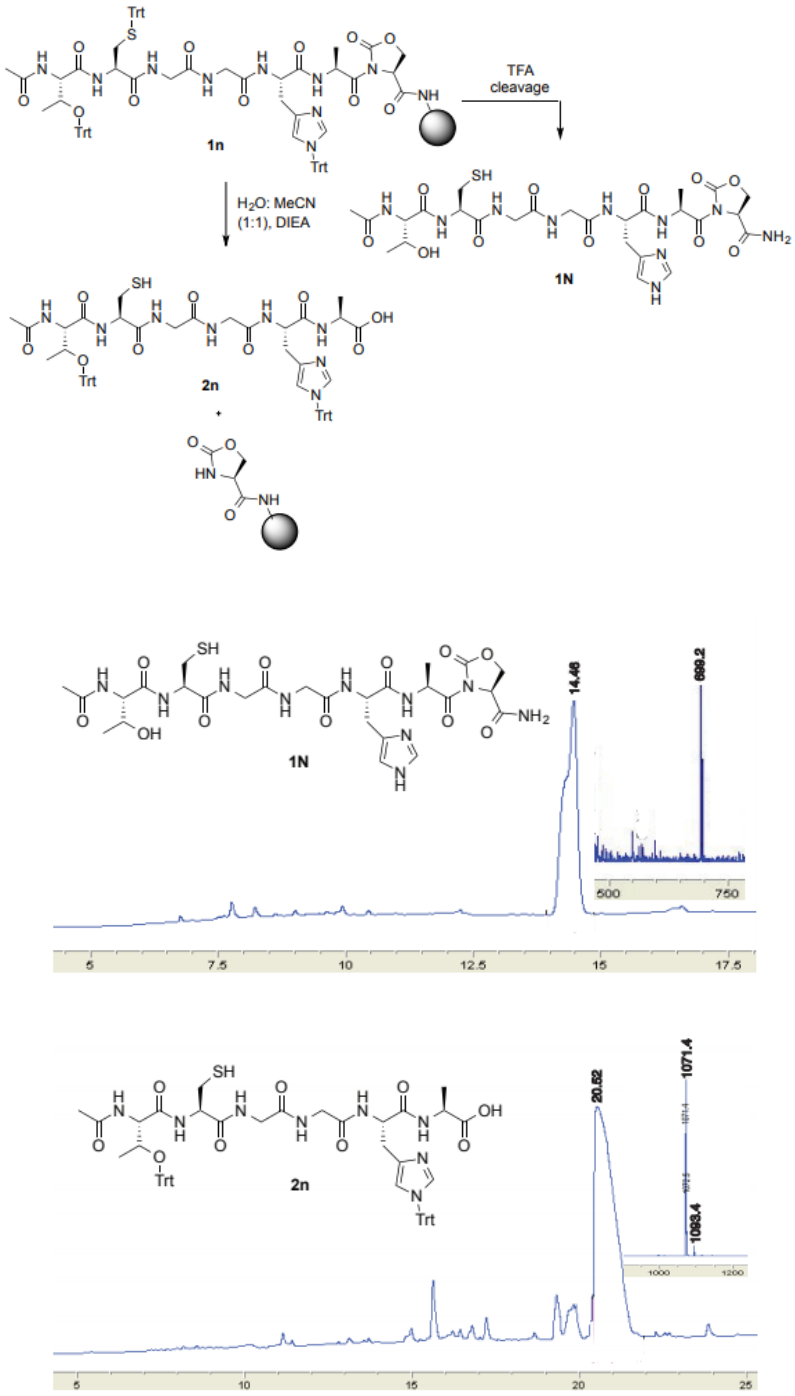


Figure A132 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Ac-Thr-Cys-Gly-Gly-His-Ala-Oxd (1N). LCMS: m/z 699.2 (calcd [M+H]+ = 699.66), Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.46 min
Ac-Thr(Trt)-Cys-Gly-Gly-His(Trt)-Ala (2n). LCMS: m/z 1071.4 (calcd [M+H]+ = 1071.5), 1093.4 (calcd [M+Na]+ = 1093.5) Purity: >95% (HPLC analysis at 220 nm). Retention time: 20.52 min

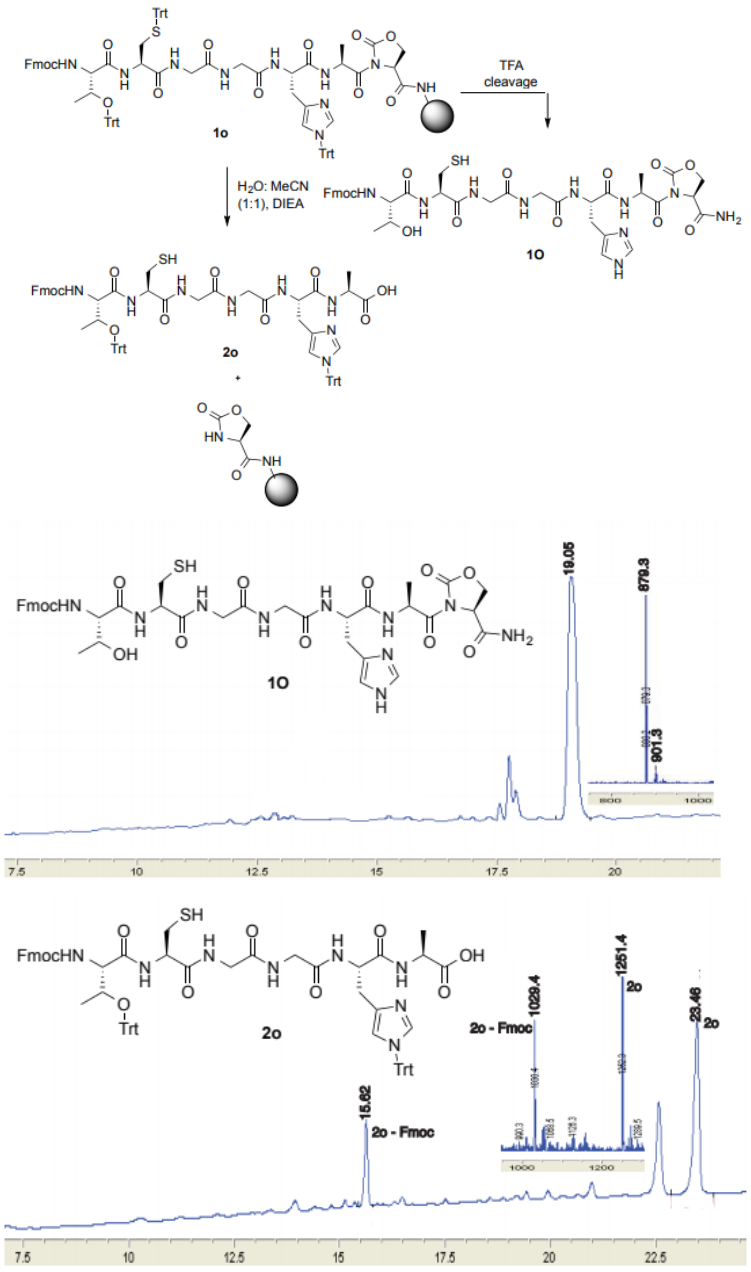


Figure A133 HPLC spectra of the cyclized starting material and the crude reaction mixture after hydrolysis; the product peak is labeled.
Fmoc-Thr-Cys-Gly-Gly-His-Ala-Oxd (1O). LCMS: m/z 879.3 (calcd [M+H]+ = 879.66), 901.3 (calcd [M+Na] + = 900.66), Purity: >95% (HPLC analysis at 220 nm). Retention time: 19.05 min
Fmoc-Thr(Trt)-Cys-Gly-Gly-His(Trt)-Ala (2o). LCMS: m/z 1251.4 (calcd [M+H]+ = 1251.57, Purity: >95% (HPLC analysis at 220 nm). Retention time: 23.46 min
Thr(Trt)-Cys-Gly-Gly-His(Trt)-Ala (2o-Fmoc). LCMS: m/z 1029.4 (calcd [M+H]+ = 1029.57), Purity: >95% (HPLC analysis at 220 nm). Retention time: 15.62 min

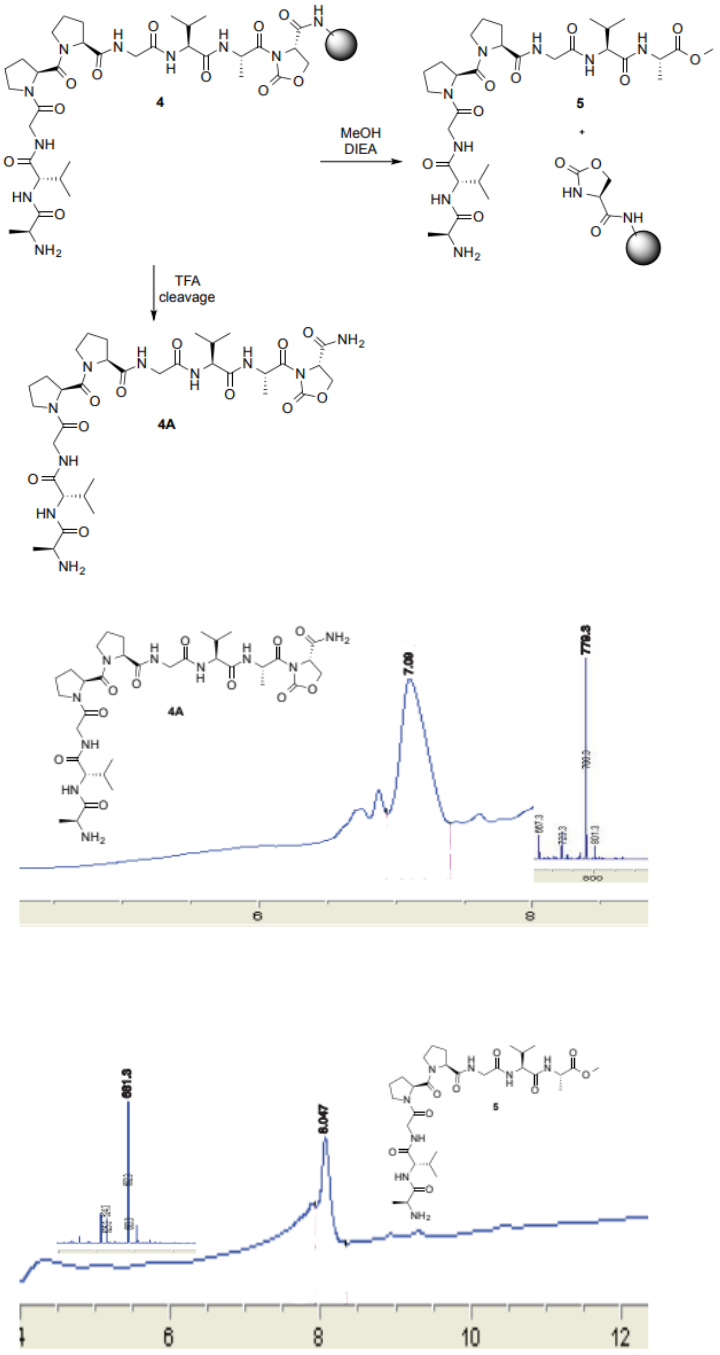


Figure A134 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in methanol solution; the product peak is labeled.
Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-Oxd (4A). LCMS: m/z 779.3 (calcd [M+H]+ = 779.37), Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.09 min
Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-OMe (5). LCMS: m/z 681.3 (calcd [M+H]+ = 681.46, Purity: >95% (HPLC analysis at 220 nm). Retention time: 8.047 min

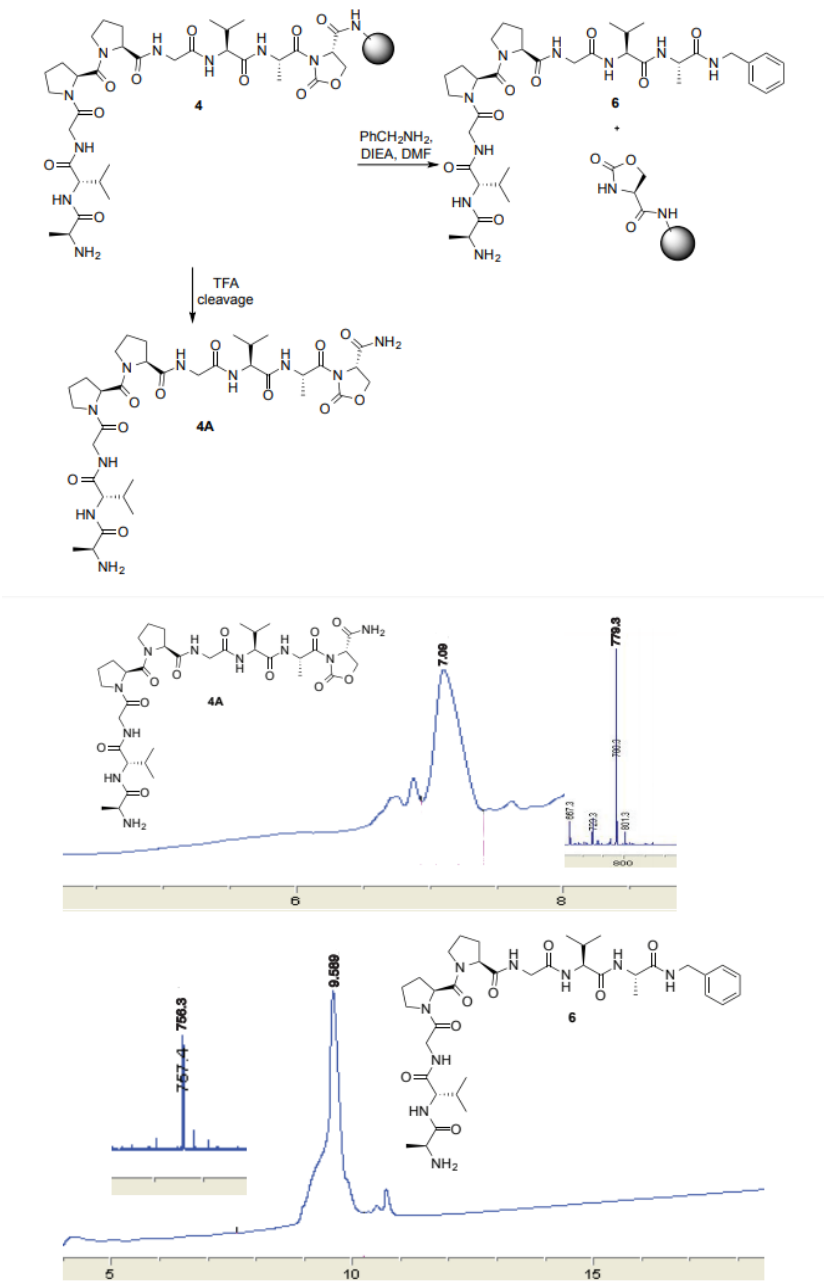


Figure A135 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in solution of benzyl amine in DMF; the product peak is labeled.
Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-Oxd (4A). LCMS: m/z 779.3 (calcd [M+H]+ = 779.37), Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.09 min
Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-NHBz (6). LCMS: m/z 756.3 (calcd [M+H]+ = 756.57, Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.58 min

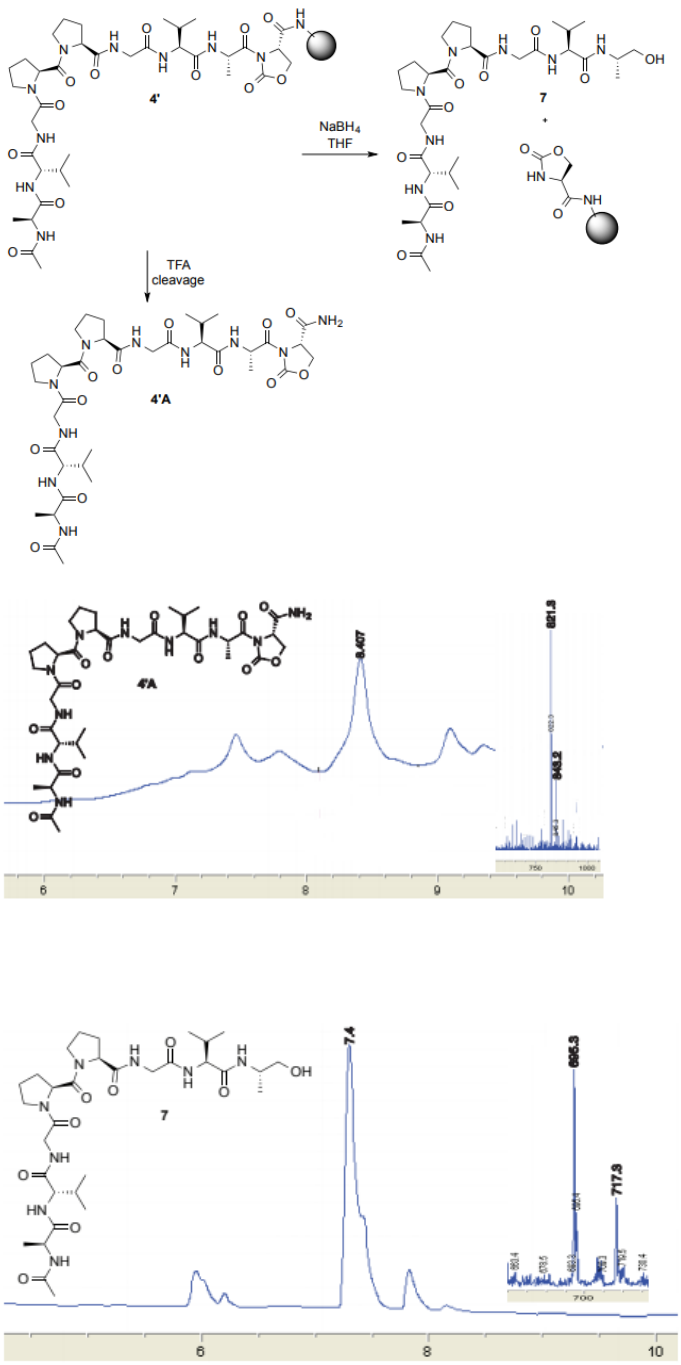


Figure A136 HPLC spectra of the cyclized starting material and the crude reaction mixture after incubation in sodium borohydride solution in THF; the product peak is labeled.
Ac-Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-Oxd (4'A). LCMS: m/z 821.3 (calcd [M+H]+ = 821.66), m/z 843.2 (calcd [M+Na] + = 843.66), Purity: >95% (HPLC analysis at 220 nm). Retention time: 8.407 min
Ac-Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-CH2OH (7). LCMS: m/z 695.3 (calcd [M+H]+ = 695.37); 717.3 (calcd [M+Na]+ = 717.37) Purity: >95% (HPLC analysis at 220 nm). Retention time: 7.4 min

CHAPTER 6 SUPPLEMENTAL MATERIAL

All of the following figures have been reproduced with permission from: Elashal, H. E.; Sim, Y. E.; Raj, M. *Chem. Sci.* **2017**, *8*, 117.

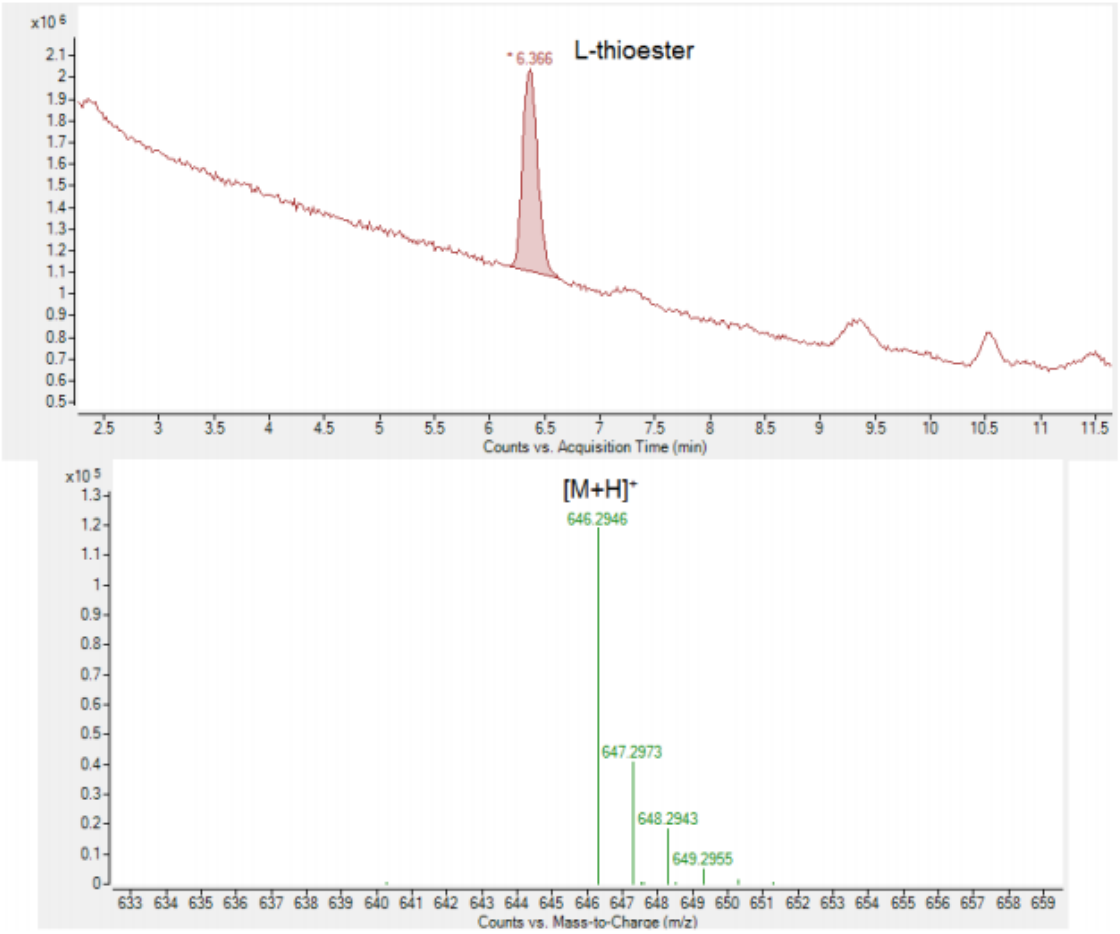


Figure A137 HRMS Spectra Ac-Gly-Pro-Met-Leu-Ala(L)-COS(CH2)2CO2C2H5 - 3G

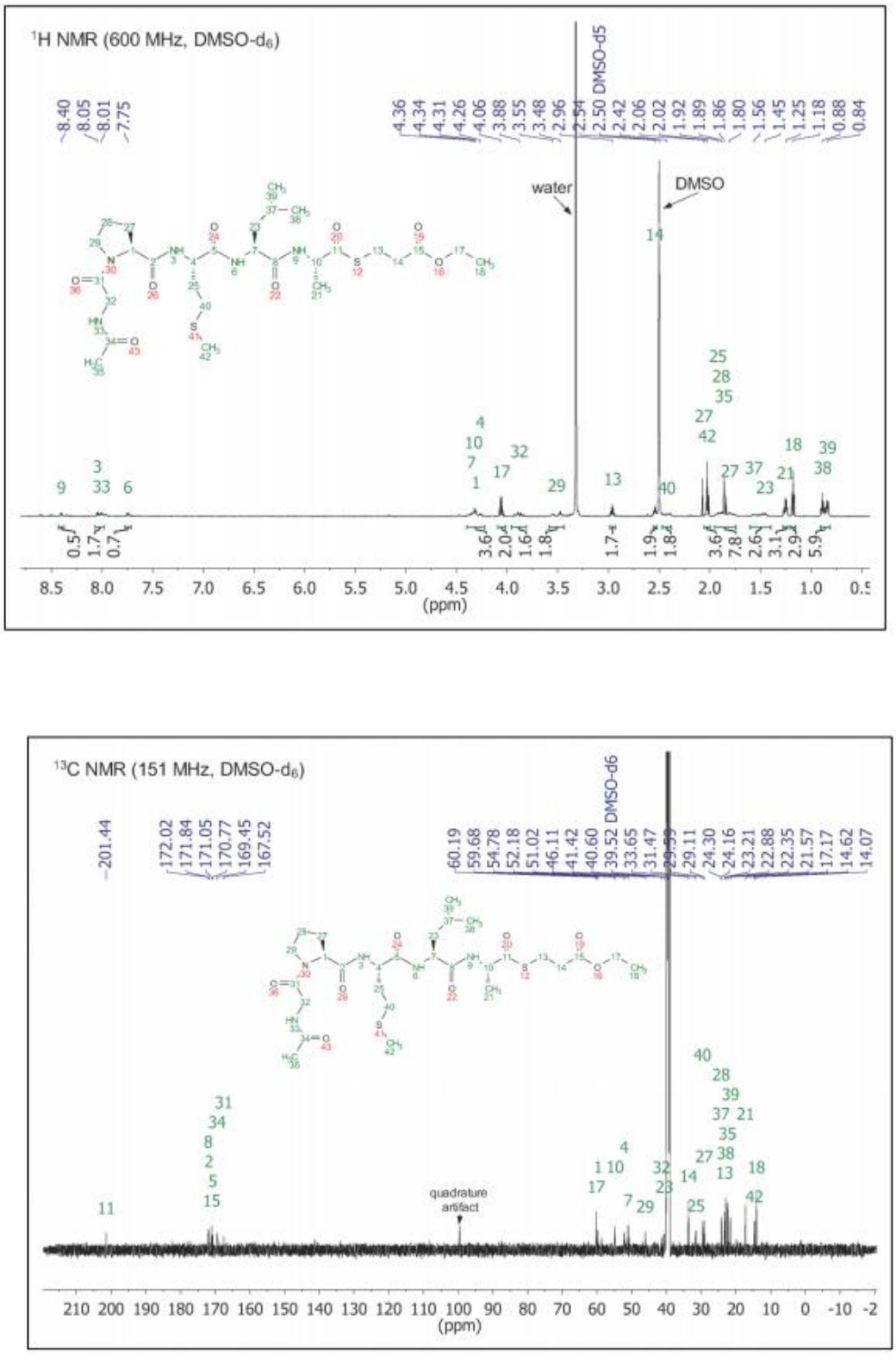


Figure A138 NMR spectra Ac-Gly-Pro-Met-Leu-Ala(L)-COS(CH2)2CO2C2H5 - 3G

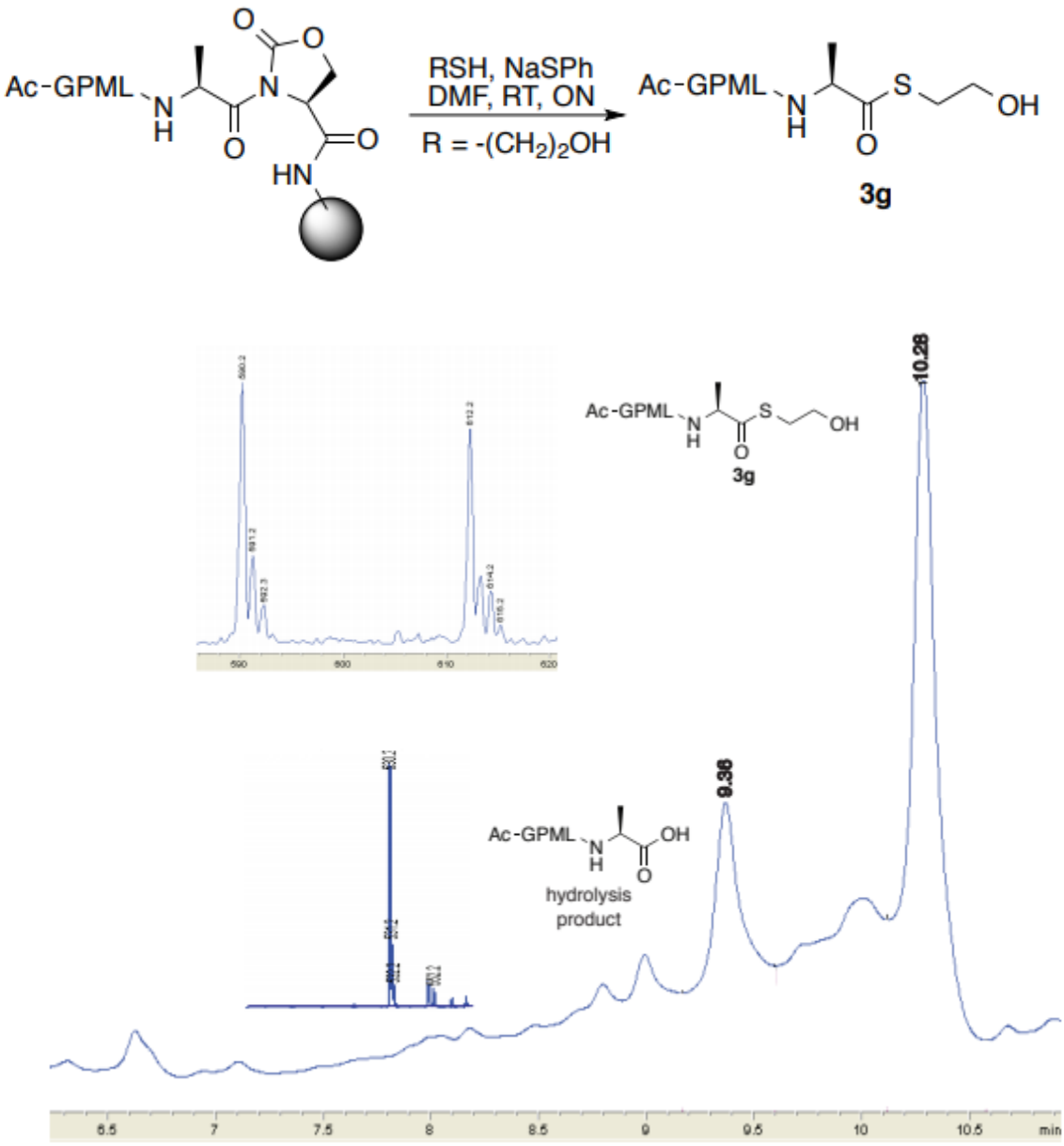


Figure A139 HPLC and MS traces for thiolysis product.
Ac-Gly-Pro-Met-Leu-Ala-COS(CH2)2OH (3g). LCMS: m/z 590.2 (calcd [M+H]+ = 590.6), m/z 612.2 (calcd [M+ Na] + = 612.6). Purity: >95% (HPLC analysis at 220 nm). Retention time:10.28 min
Ac-Gly-Pro-Met-Leu-Ala-OH. LCMS: m/z 530.2 (calcd [M+H]+ = 530.5), m/z 552.2 (calcd [M+Na]+ = 552.5). Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.36 min

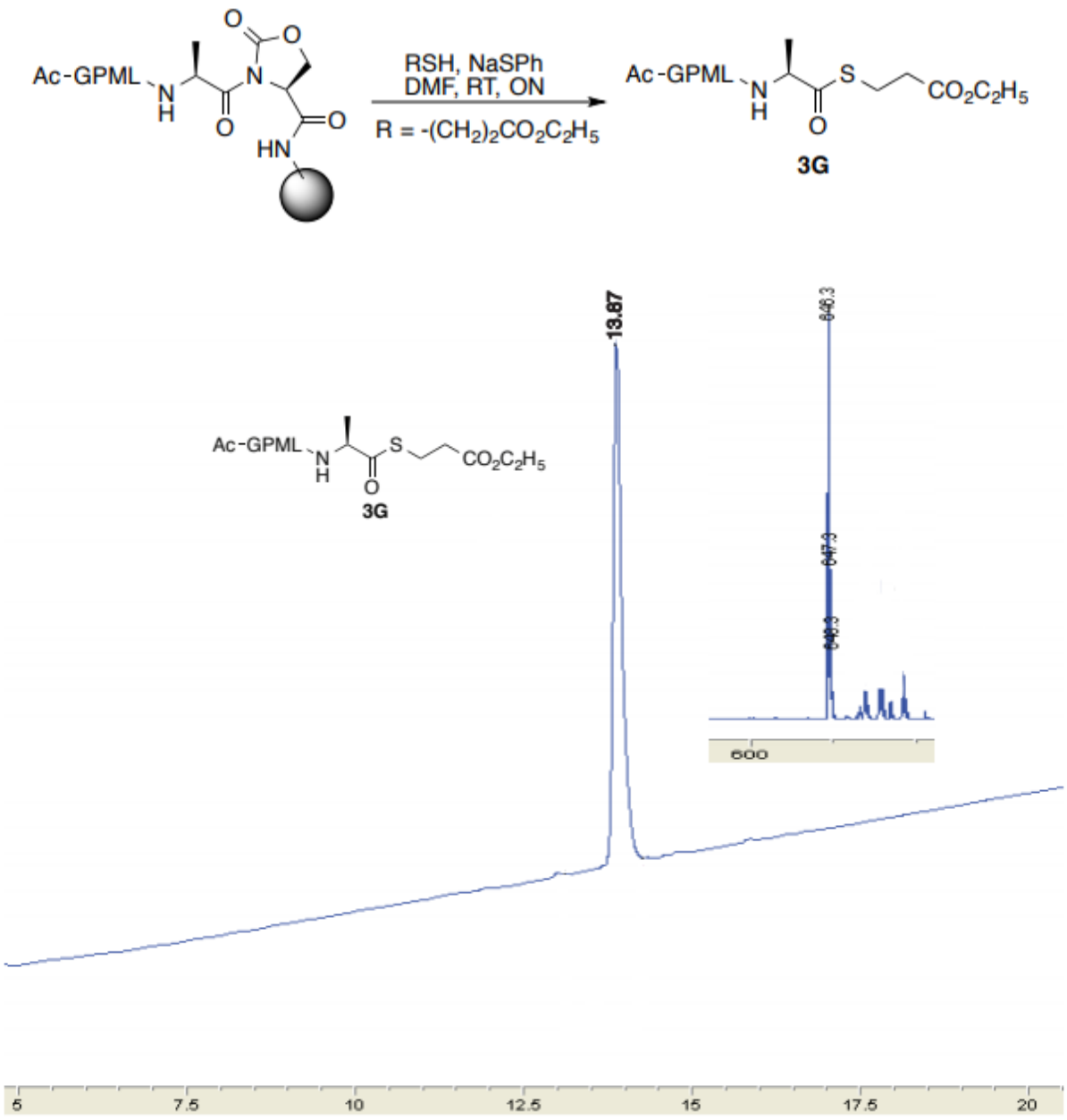


Figure A140 HPLC and MS traces for thiolysis product.
Ac-Gly-Pro-Met-Leu-Ala-COS(CH2)2CO2C2H5 (3G). LCMS: m/z 646.3 (calcd [M+H]+ = 646.7). Purity: >95% (HPLC analysis at 220 nm). Retention time: 13.87 min

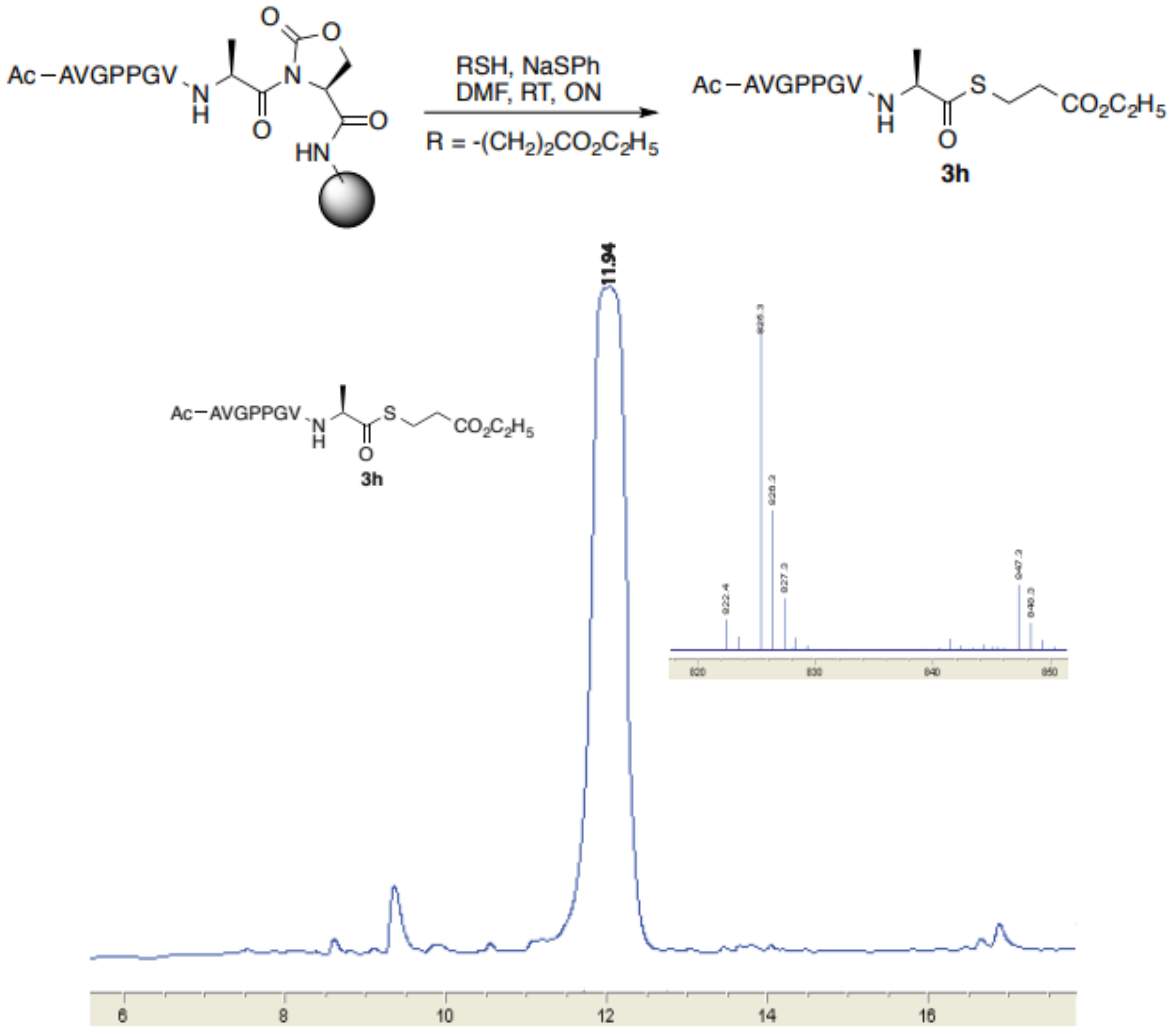


Figure A141 HPLC and MS traces for thiolysis product.
Ac-Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-COS(CH2)2CO2C2H5 (3h). LCMS: m/z 825.3 (calcd [M+H]+ = 825.7), m/z 847.3 (calcd [M+Na] + = 847.7). Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.94 min

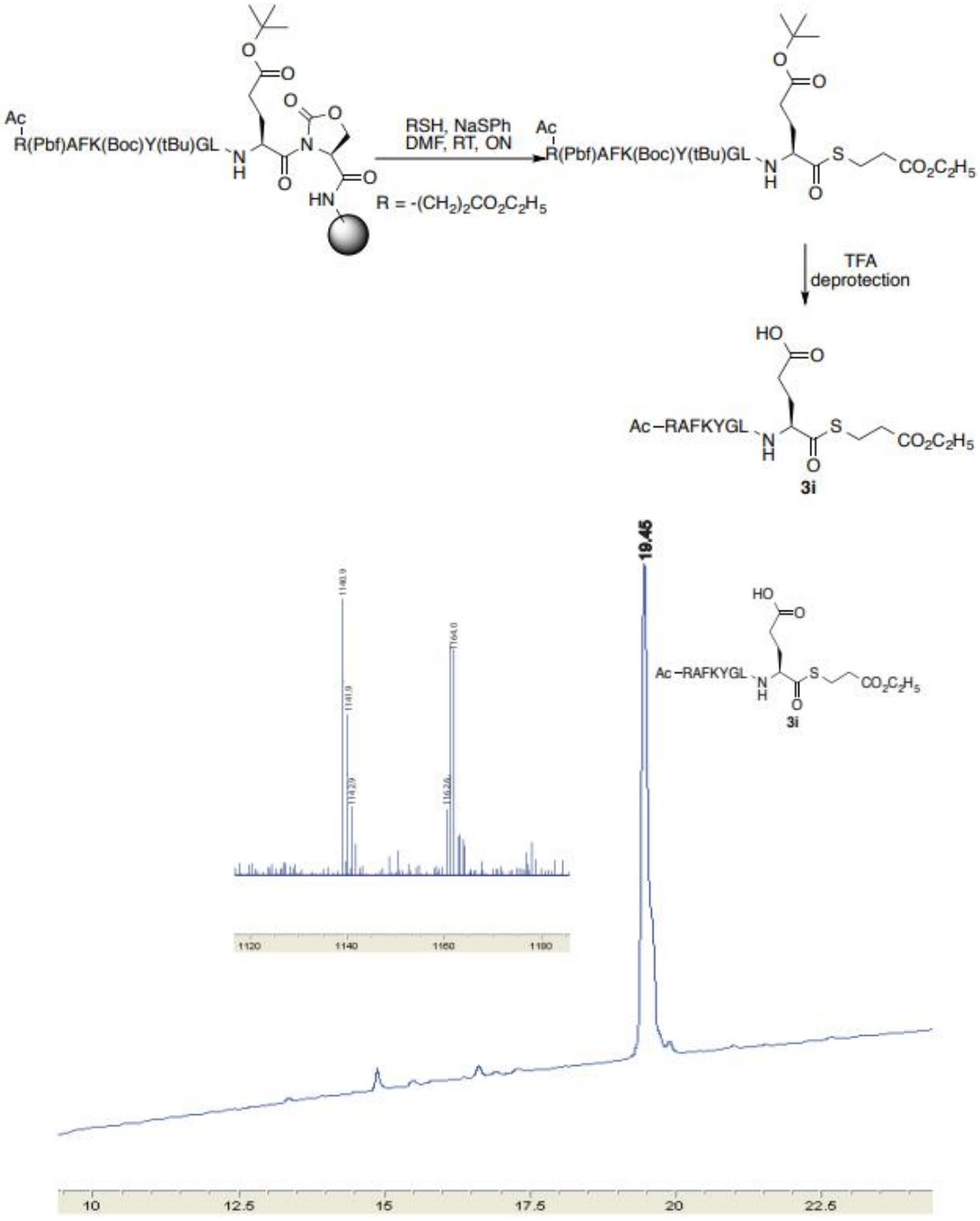


Figure A142 HPLC and MS traces for thiolysis product.
Ac-Arg-Ala-Phe-Lys-Tyr-Gly-Leu-Glu-COS(CH2)2CO2C2H5 (3i). LCMS: m/z 1140.9 (calcd [M+H]+ = 1140.5), m/z 1163.9 (calcd [M+Na] + = 1163.5). Purity: >95% (HPLC analysis at 220 nm). Retention time: 19.45 min

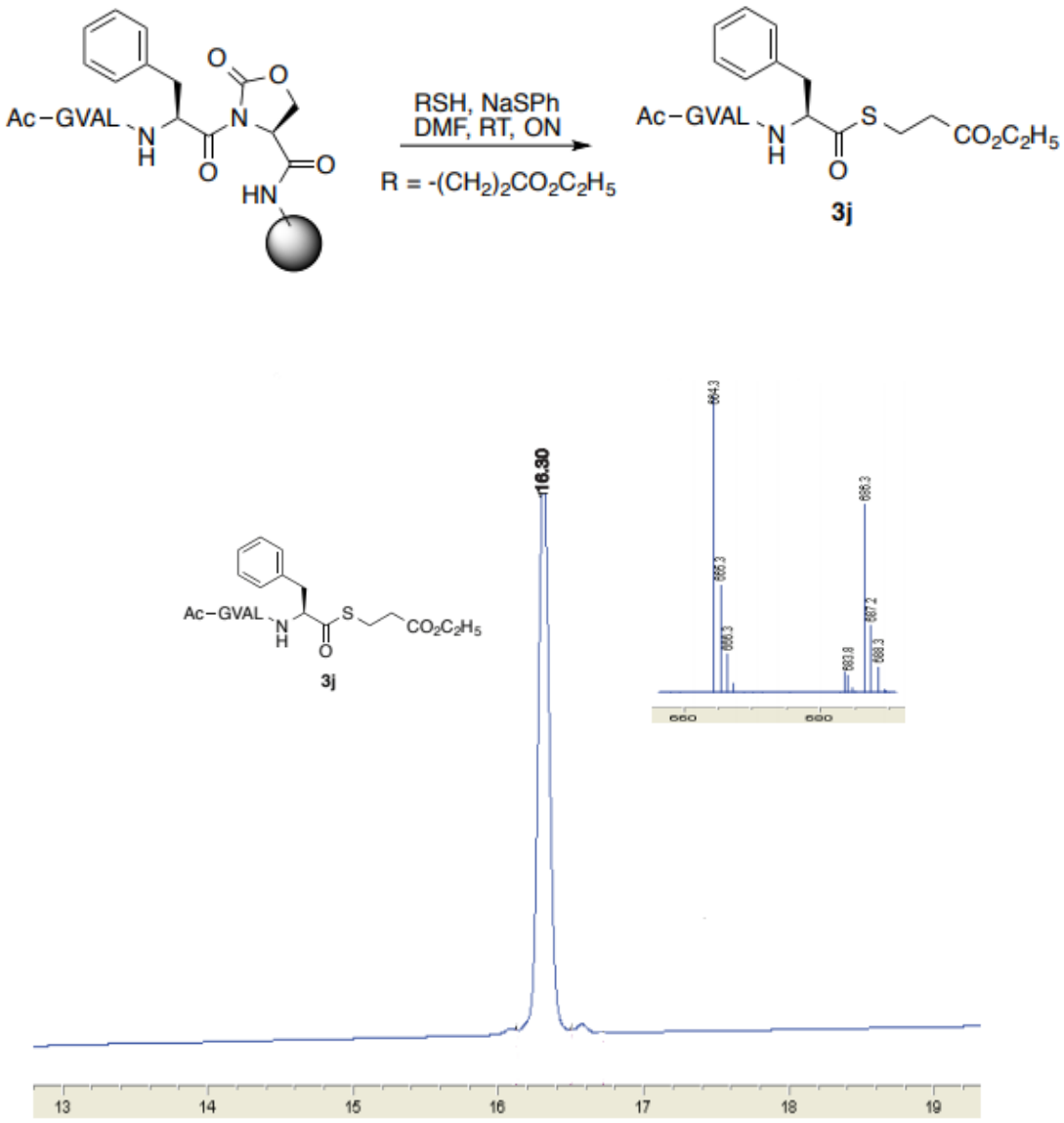
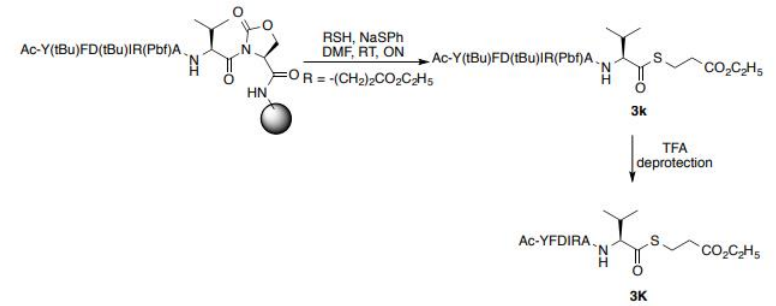
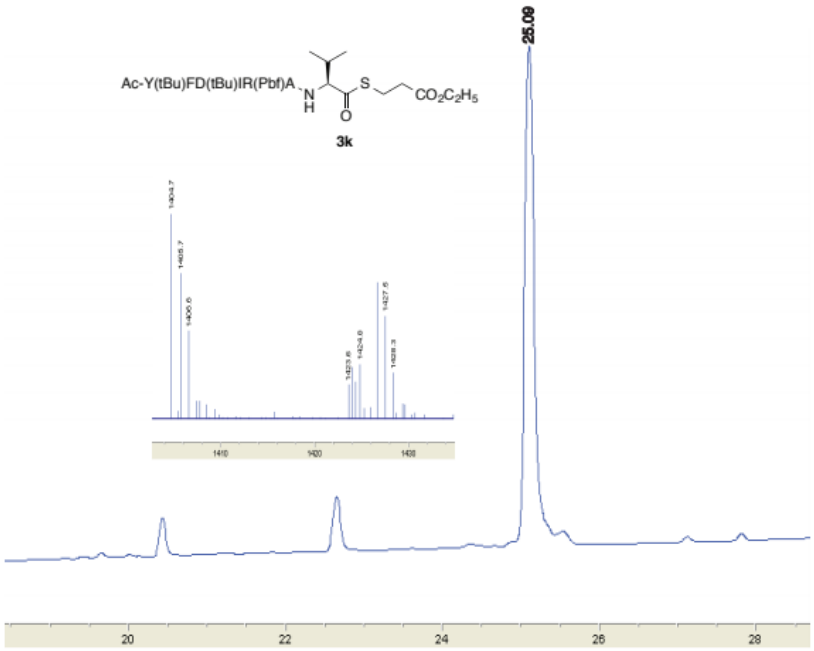


Figure A143 HPLC and MS traces for thiolysis product.
Ac-Gly-Val-Ala-Leu-Phe-COS(CH2)2CO2C2H5 (3j). LCMS: m/z 664.3 (calcd [M+H]+ = 664.6), m/z 686.3 (calcd [M+Na] + = 686.6). Purity: >95% (HPLC analysis at 220 nm). Retention time: 16.30 min.


Ac-Y(tBu)FD(tBu)IR(Pbf)A
RSH, NaSPh
DMF, RT, ON
R = -(CH2)2CO2C2H5
Ac-Y(tBu)FD(tBu)IR(Pbf)A
CO2C2H5
3k
TFA
deprotection
Ac-YFDIRA
CO2C2H5
3K


Ac-Y(tBu)FD(tBu)IR(Pbf)A
CO2C2H5
3k
25.09
20
22
24
26
28

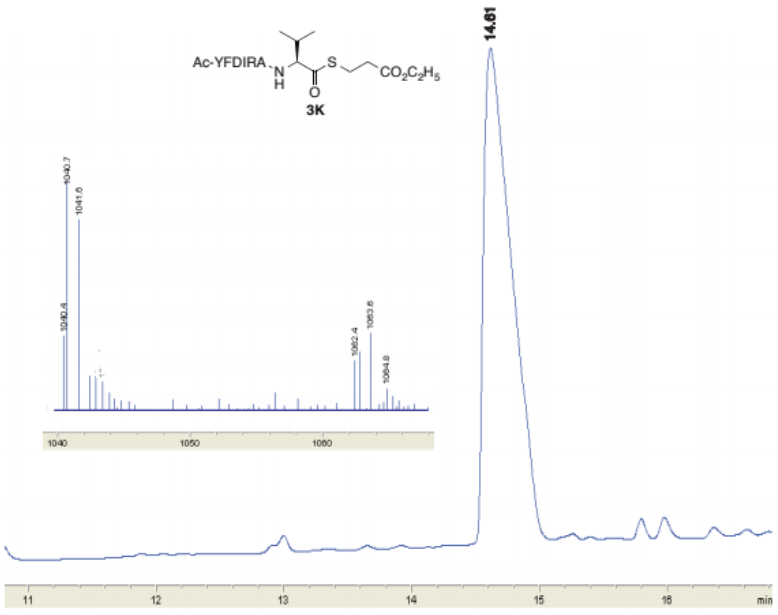

Ac-YFDIRA
CO2C2H5
3K
14.61
11
12
13
14
15
16
min

Figure A144 HPLC and MS traces for thiolysis product.
Ac-Tyr(tBu)-Phe-Asp(tBu)-Ile-Arg(Pbf)-Ala-Val-COS(CH2)2CO2C2H5 (3k). LCMS: m/z 1404.7 (calcd [M+H]+ = 1404.3), m/z 1426.5 (calcd [M+Na] + = 1426.3). Purity: >95% (HPLC analysis at 220 nm). Retention time: 25.09 min.
Ac-Tyr-Phe-Asp-Ile-Arg-Ala-Val-COS(CH2)2CO2C2H5 (3K). LCMS: m/z 1040.7 (calcd [M+H]+ = 1040.8), m/z 1063.6 (calcd [M+Na] + = 1063.8). Purity: >95% (HPLC analysis at 220 nm). Retention time: 14.61 min.

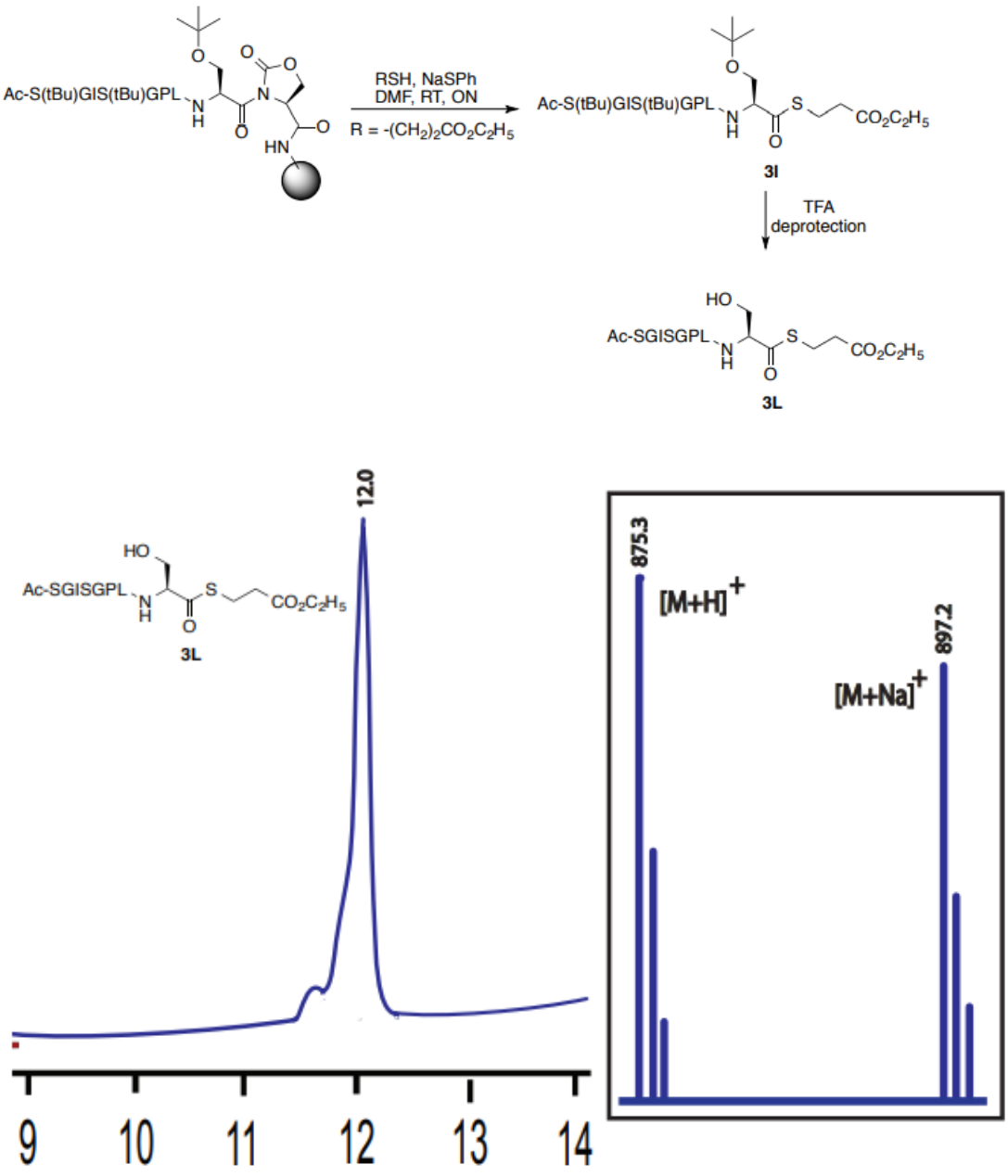


Figure A145 HPLC and MS traces for thiolysis product.
Ac-Ser-Gly-Ile-Ser-Gly-Pro-Leu-Ser-COS(CH2)2CO2C2H5 (3L). LCMS: m/z 875.3 (calcd [M+H]+ = 875.7), m/z 897.2 (calcd [M+Na] + = 897.7). Purity: >95% (HPLC analysis at 220 nm). Retention time: 12.0 min.

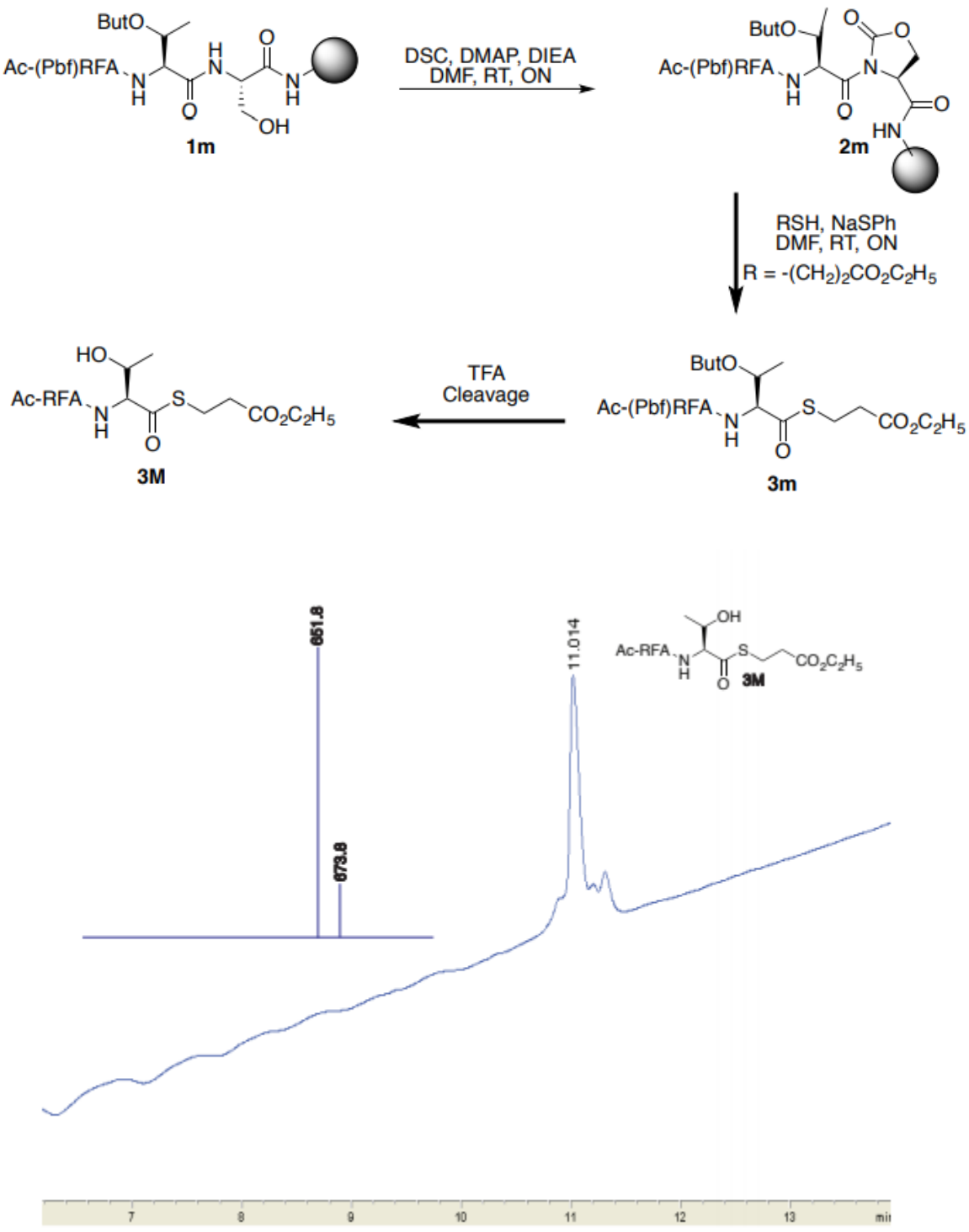


Figure A146 HPLC and MS traces for thiolysis product.
Ac-Arg-Phe-Ala-Thr-COS(CH2)2CO2C2H5 (3M). LCMS: m/z 651.8 (calcd [M+H]+ = 651.5), m/z 673.8 (calcd [M+Na]+ = 673.5). Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.01 min. Isolated yield = 50 %

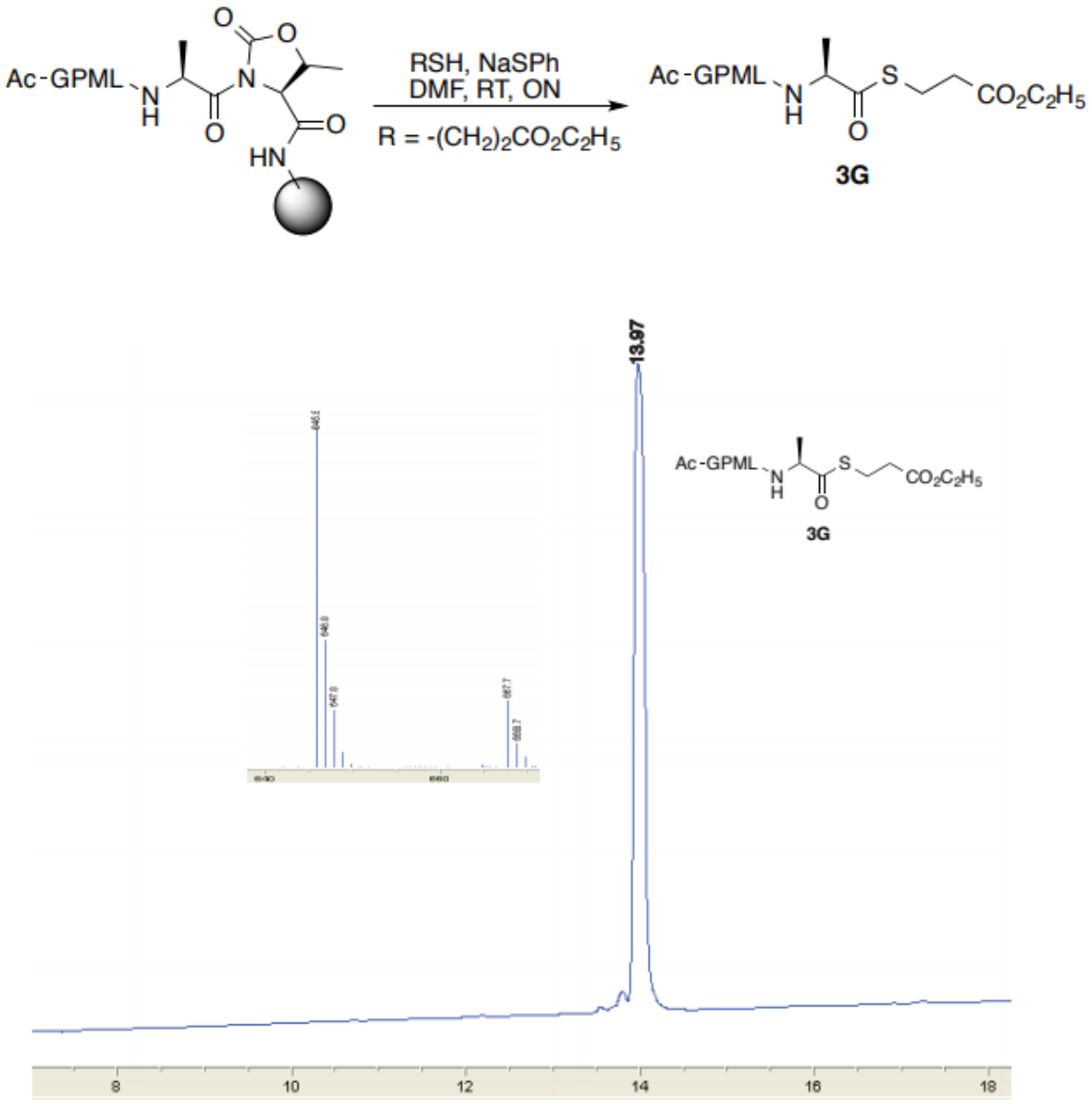


Figure A147 HPLC and MS traces for thiolysis product.
Ac-Gly-Pro-Met-Leu-Ala-COS(CH2)2CO2C2H5 (3G). LCMS: m/z 645.8 (calcd [M+H]+ = 645.5), m/z 667.7 (calcd [M+Na] + = 667.5). Purity: >95% (HPLC analysis at 220 nm). Retention time: 13.97 min

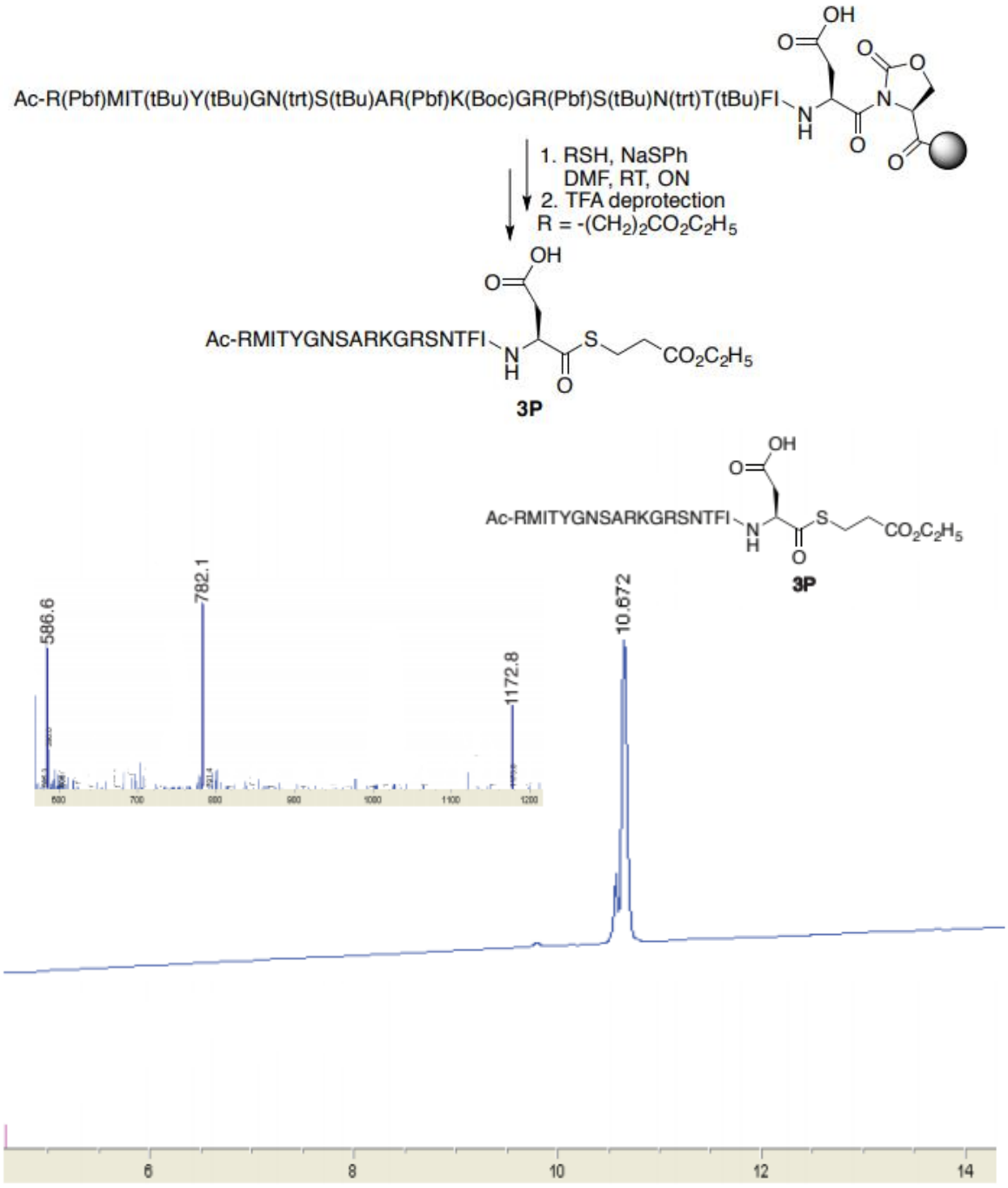


Figure A148 HPLC and MS traces for thiolysis product.
Ac-RMITYGNSARKGRSNTFID-COS(CH2)2CO2C2H5 (3P). LCMS: m/z 1172.8 (calcd [M+2H] 2+ = 1172.5), m/z 782.1 (calcd [M+3H] 3+ = 782.0), m/z 586.6 (calcd [M+4H] 4+= 586.7). Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.67 min. Isolated yield = 31%

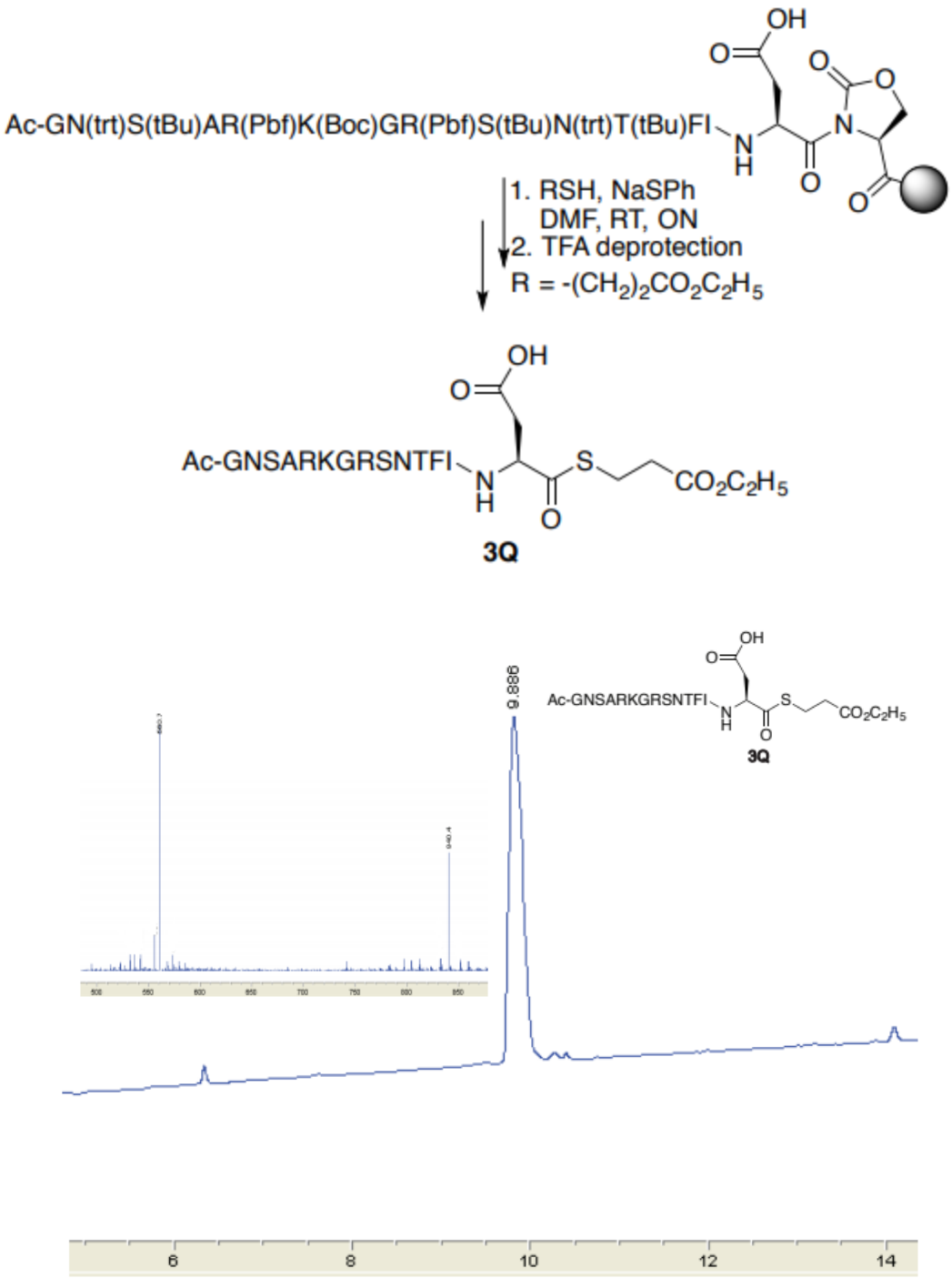
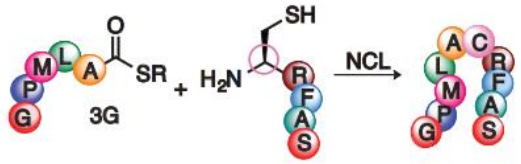
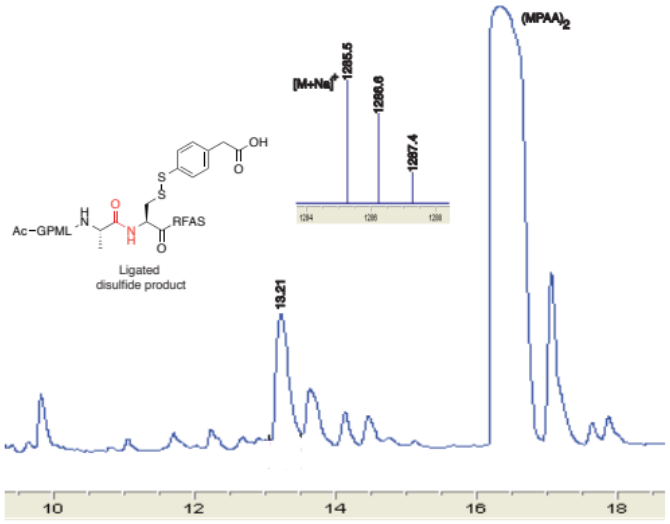
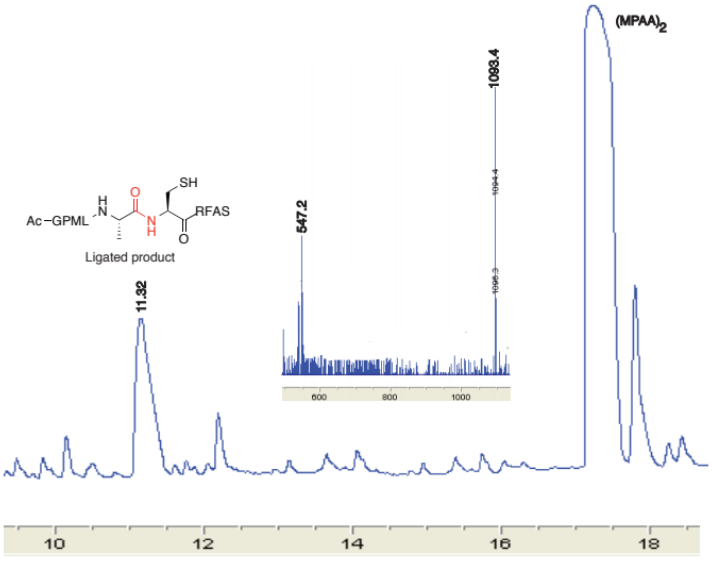


Figure A149 HPLC and MS traces for thiolysis product.
Ac-GNSARKGRSNTFID-COS(CH2)2CO2C2H5 (3Q). LCMS: m/z 840.4 (calcd [M+2H] 2+ = 840.9), m/z 560.7 (calcd [M+3H] 3+ = 560.9). Purity: >95% (HPLC analysis at 220 nm). Retention time: 9.88 min. Isolated yield = 39 %



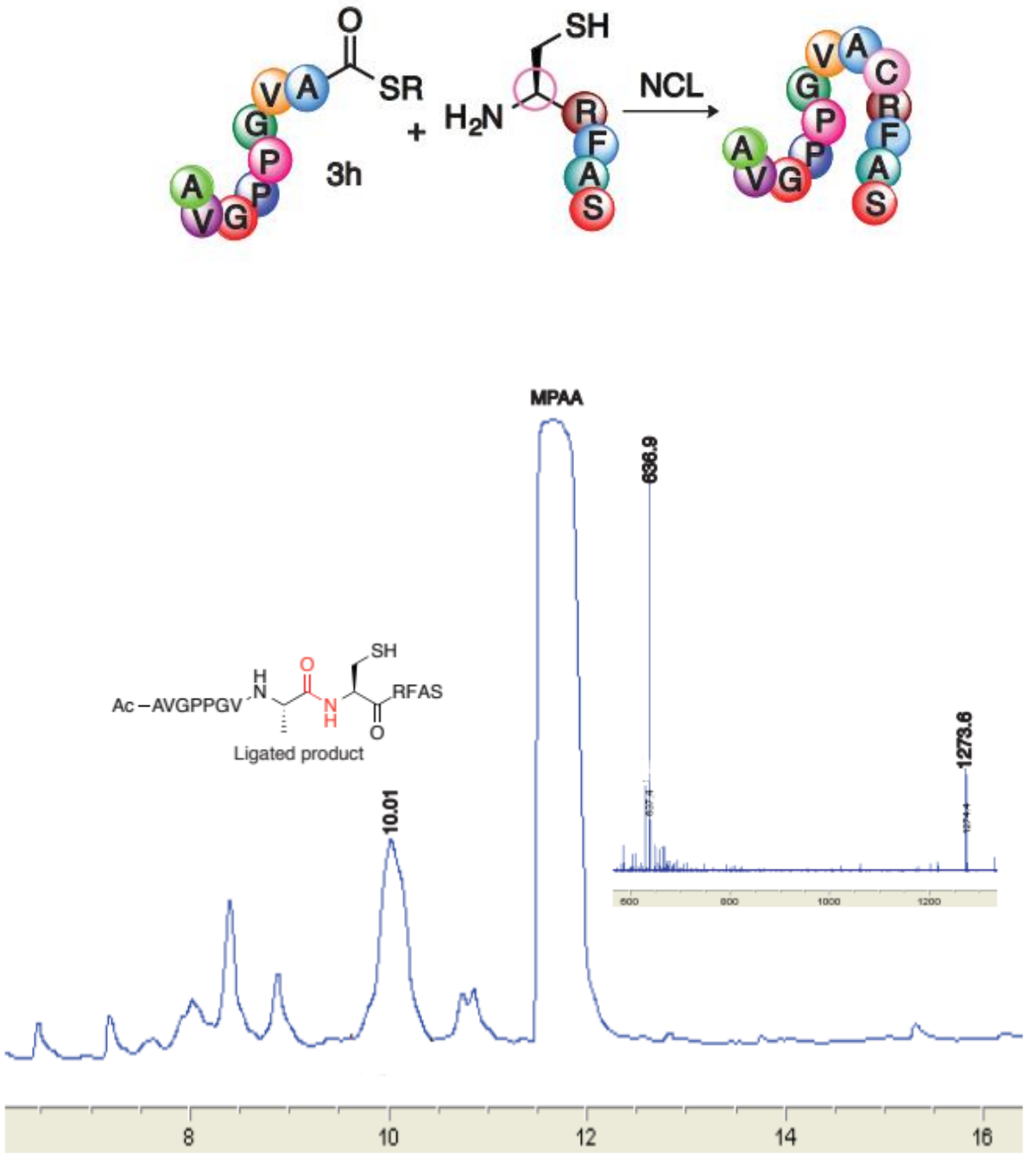


Native Chemical Ligation reaction before quenching it with TCEP.HCl



Native Chemical Ligation reaction after quenching it with TCEP.HCl

Figure A150 HPLC and MS traces for ligated product.
Ac-Gly-Pro-Met-Leu-Ala-Cys(MPAA)-Arg-Phe-Ala-Ser. LCMS: m/z 1285.5 (calcd [M+Na] + = 1285.8). Purity: >95% (HPLC analysis at 220 nm). Retention time: 13.2 min.
Ac-Gly-Pro-Met-Leu-Ala-Cys-Arg-Phe-Ala-Ser. LCMS: m/z 1093.4 (calcd [M+H]+ = 1093.3), m/z 547.2 (calcd [M+2] 2+ = 547.6). Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.32 min.



Native Chemical Ligation reaction after quenching it with TCEP.HCl

Figure A151 HPLC and MS traces for ligated product.
Ac-Ala-Val-Gly-Pro-Pro-Gly-Val-Ala-Cys-Arg-Phe-Ala-Ser. LCMS: m/z 1273.6 (calcd [M+H]+ = 1273.4), m/z 636.9 (calcd [M+2] 2+ = 637.2). Purity: >95% (HPLC analysis at 220 nm). Retention time: 10.01 min.

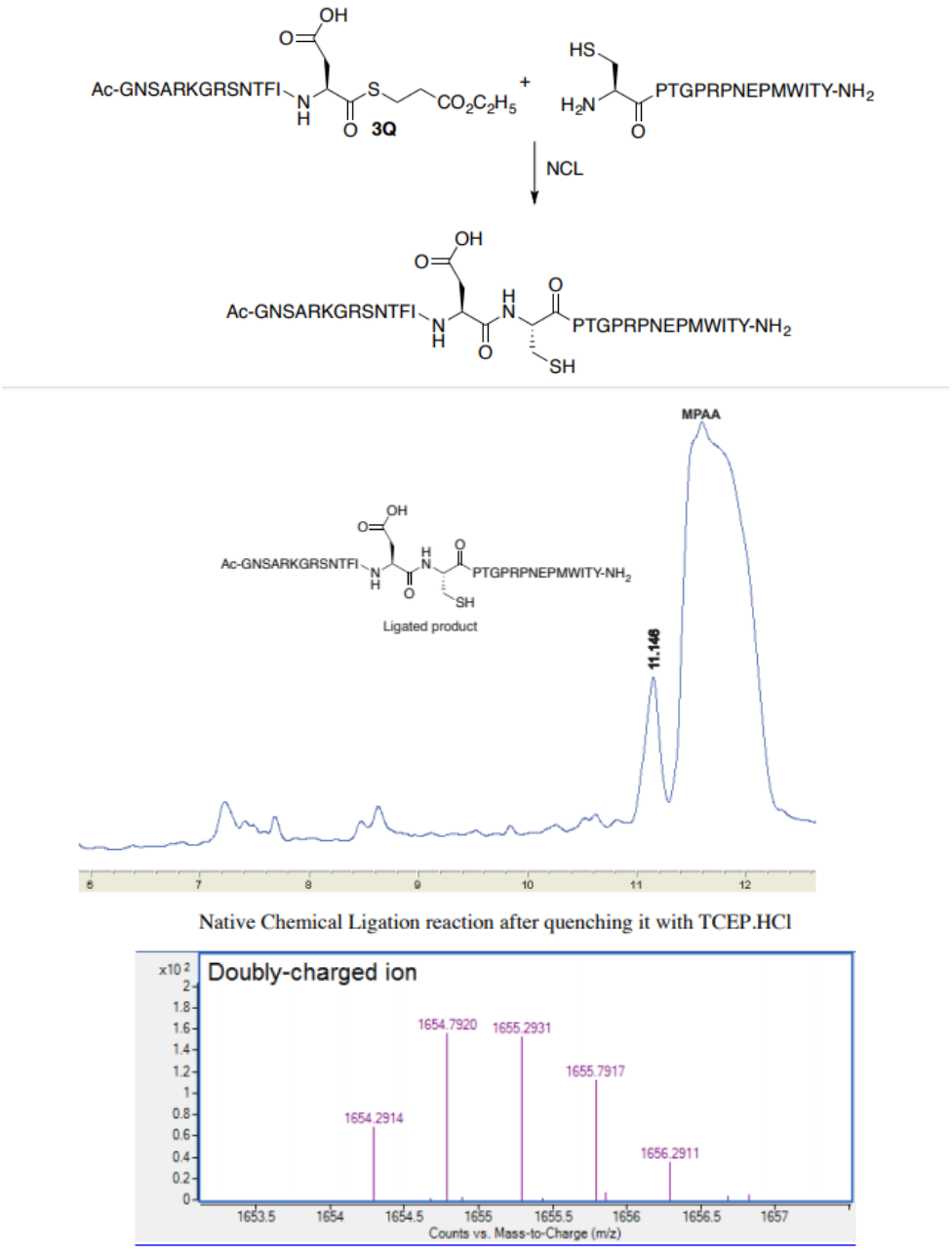


Figure A152 HPLC and MS traces for ligated product.
Ac-GNSARKGRSNTFIDCPTGPRPNEPMWITY-NH2. LCMS: m/z 3308.6 (calcd [M+H]+ = 3308.4), m/z 1654.8 (calcd [M+2]2+ = 1654.7). Purity: >95% (HPLC analysis at 220 nm). Retention time: 11.14 min

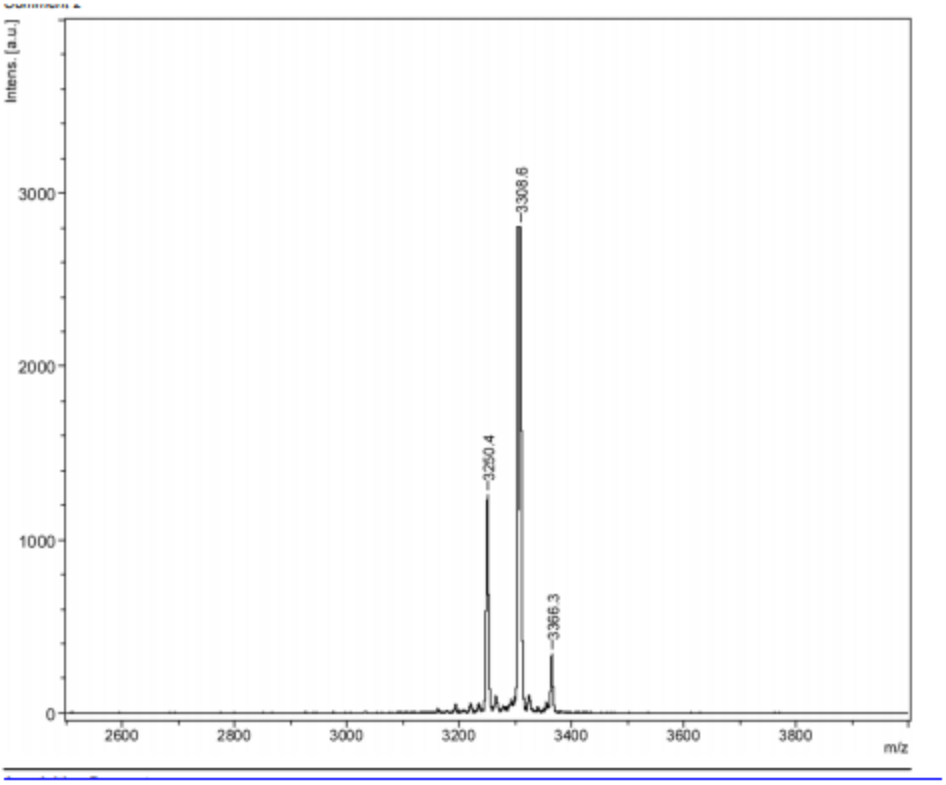
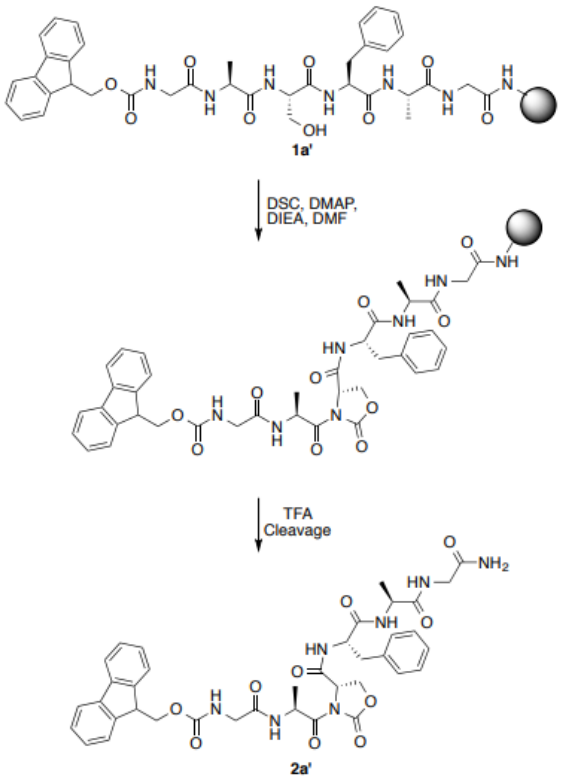


Figure A153 MALDI of ligated Product 29 amino acid long peptide Ac-GNSARKGRSNTFIDCPTGPRPNEPMWITY-NH_2



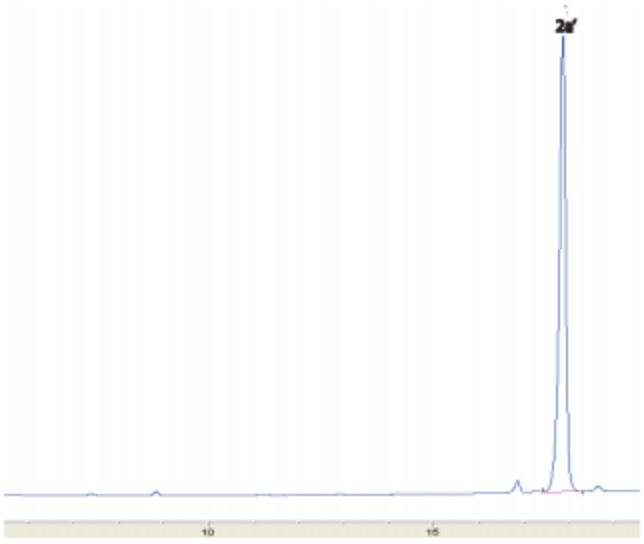


Figure A154 HPLC spectra of the cyclic urethane containing peptide 2a' Fmoc-Gly-Ala-Oxd-Phe-Ala-Gly-NH2 (2a'). LCMS: m/z 756.40 (calcd [M+H]+ = 756.10), 778.5 (calcd [M+Na]+ = 778.0). Purity: >95% (HPLC analysis at 254 nm). Retention time: 17.9 min.

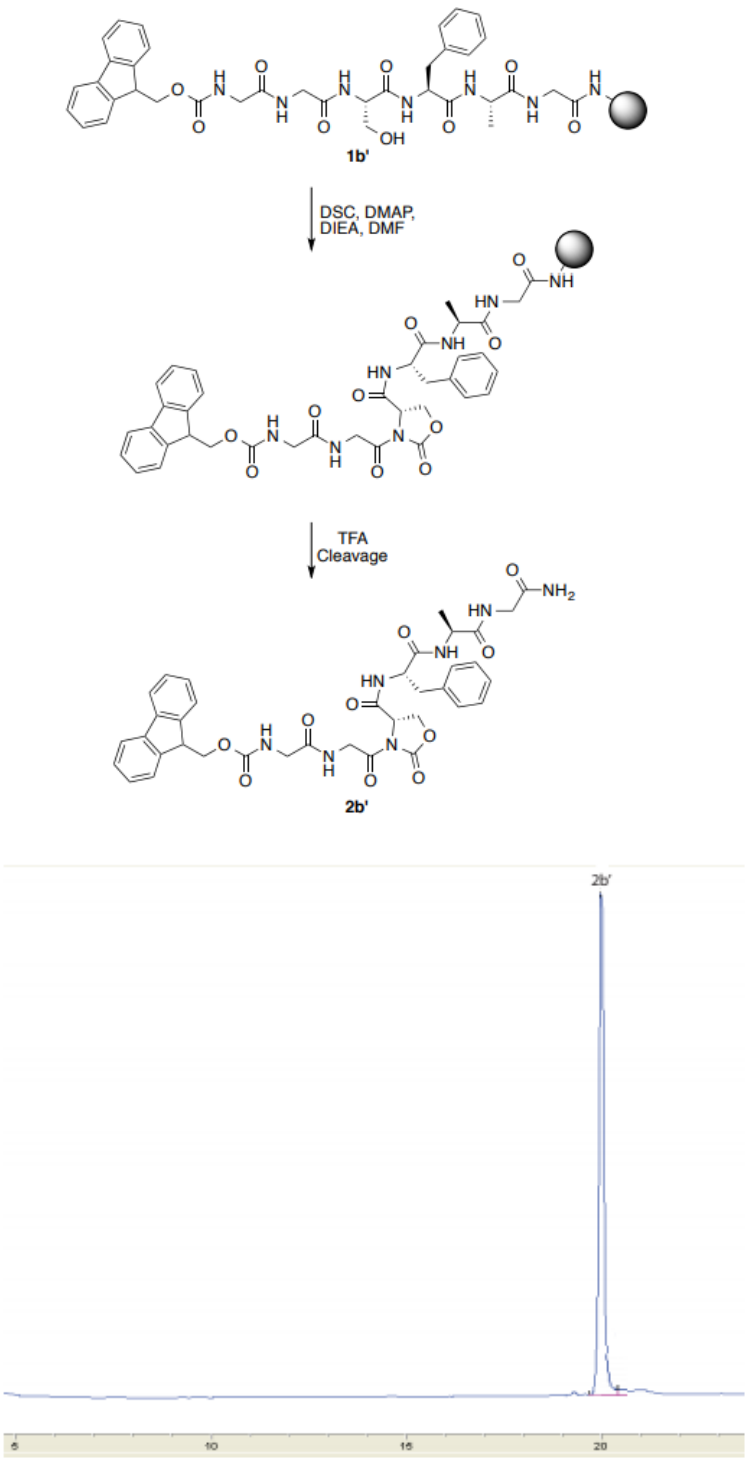
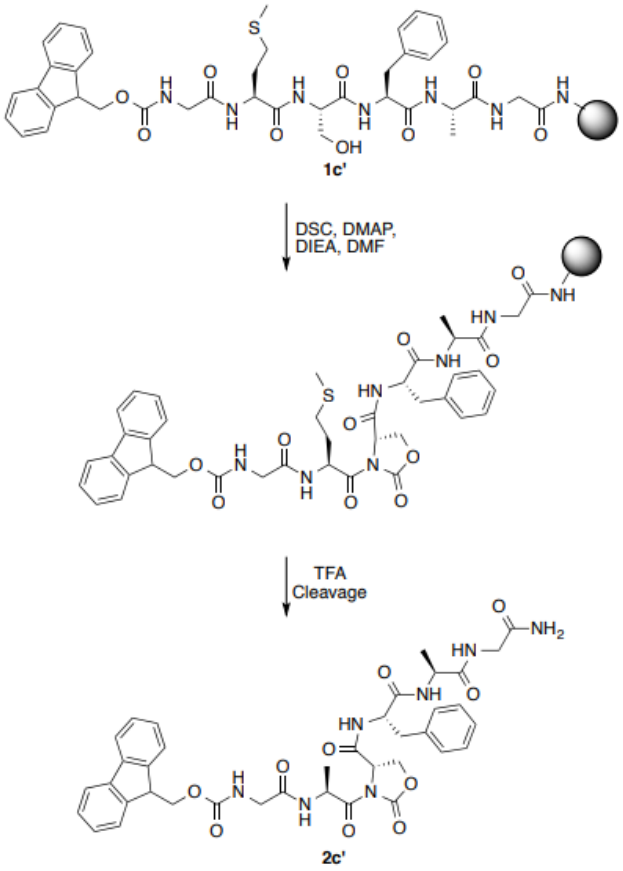


Figure A155 HPLC spectra of the cyclic urethane containing peptide 2b'
Fmoc-Gly-Gly-Oxd-Phe-Ala-Gly-NH2 (2b'). LCMS: m/z 742.40 (calcd [M+H]+ = 742.50), 371.9 (calcd [(M+2] 2+ = 371.75). Purity: >95% (HPLC analysis at 254 nm). Retention time: 20.1 min.



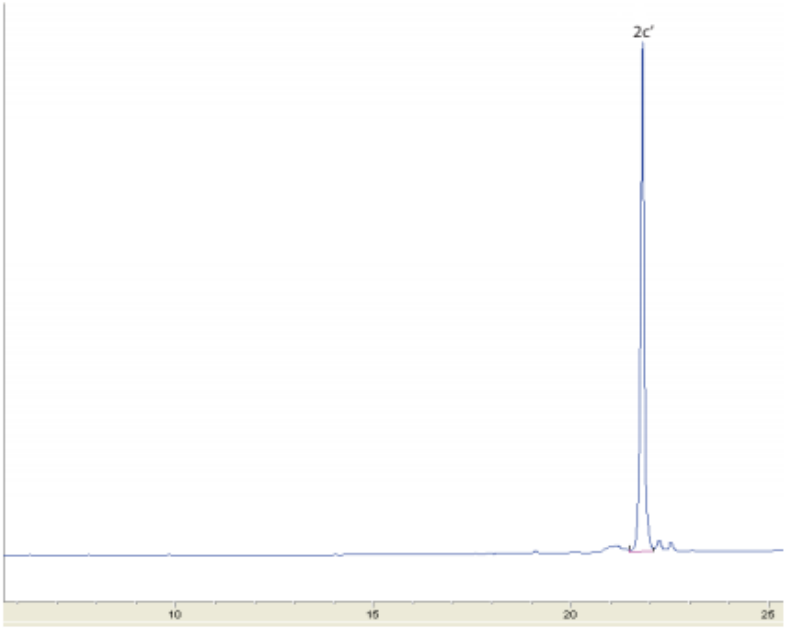


Figure A156 HPLC spectra of the cyclic urethane containing peptide 2c' Fmoc-Gly-Met-Oxd-Phe-Ala-Gly-NH2 (2c'). LCMS: m/z 816.0 (calcd [M+H]+ = 816.20), 408.2 (calcd (M+2) 2+ = 408.8). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.8 min.

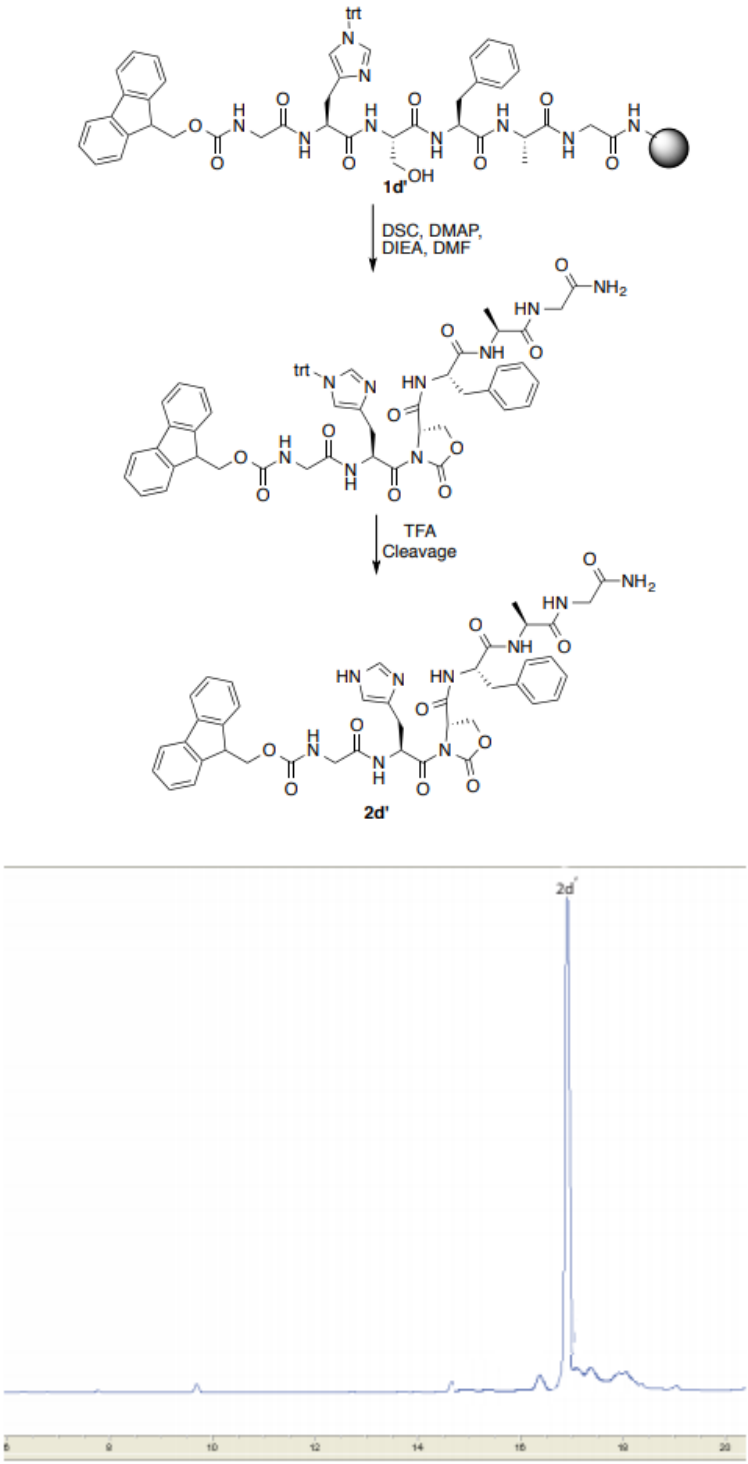
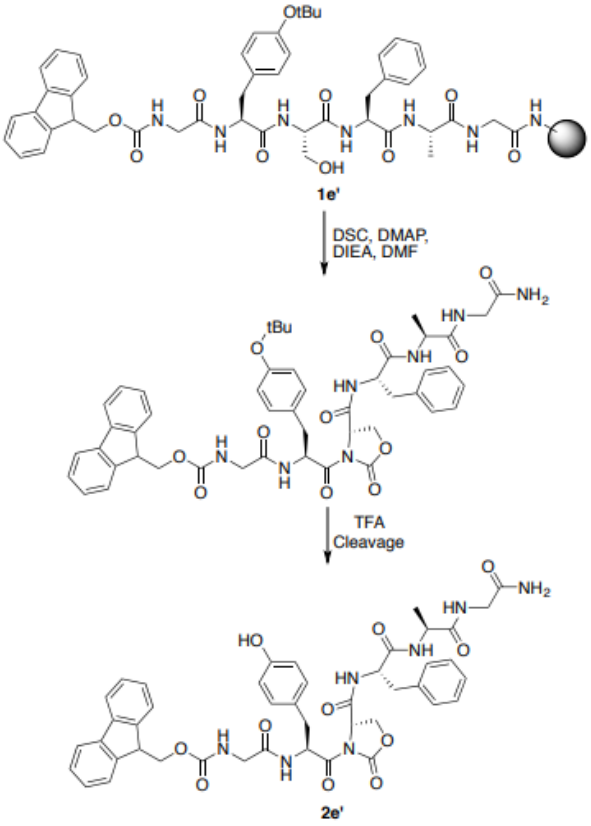


Figure A157 HPLC spectra of the cyclic urethane containing peptide 2d'
Fmoc-Gly-His-Oxd-Phe-Ala-Gly-NH2 (2d'). LCMS: m/z 822.10 (calcd [M+H]+ = 822.60), 411.2 (calcd (M+2)2+ = 411.79). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.9 min.



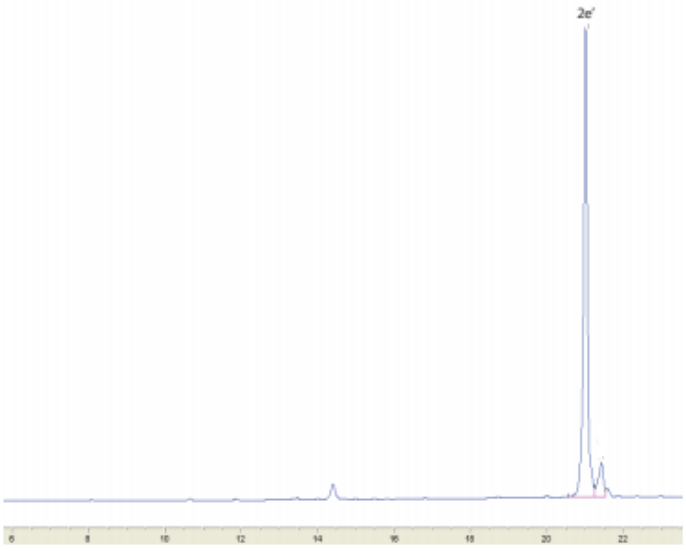


Figure A158 HPLC spectra of the cyclic urethane containing peptide 2e'
Fmoc-Gly-Tyr-Oxd-Phe-Ala-Gly-NH2 (2e'). LCMS: m/z 848.20 (calcd [M+H]+ = 848.62), 424.1 (calcd [(M+2) 2+ = 424.81). Purity: >95% (HPLC analysis at 254 nm). Retention time: 21.1 min.

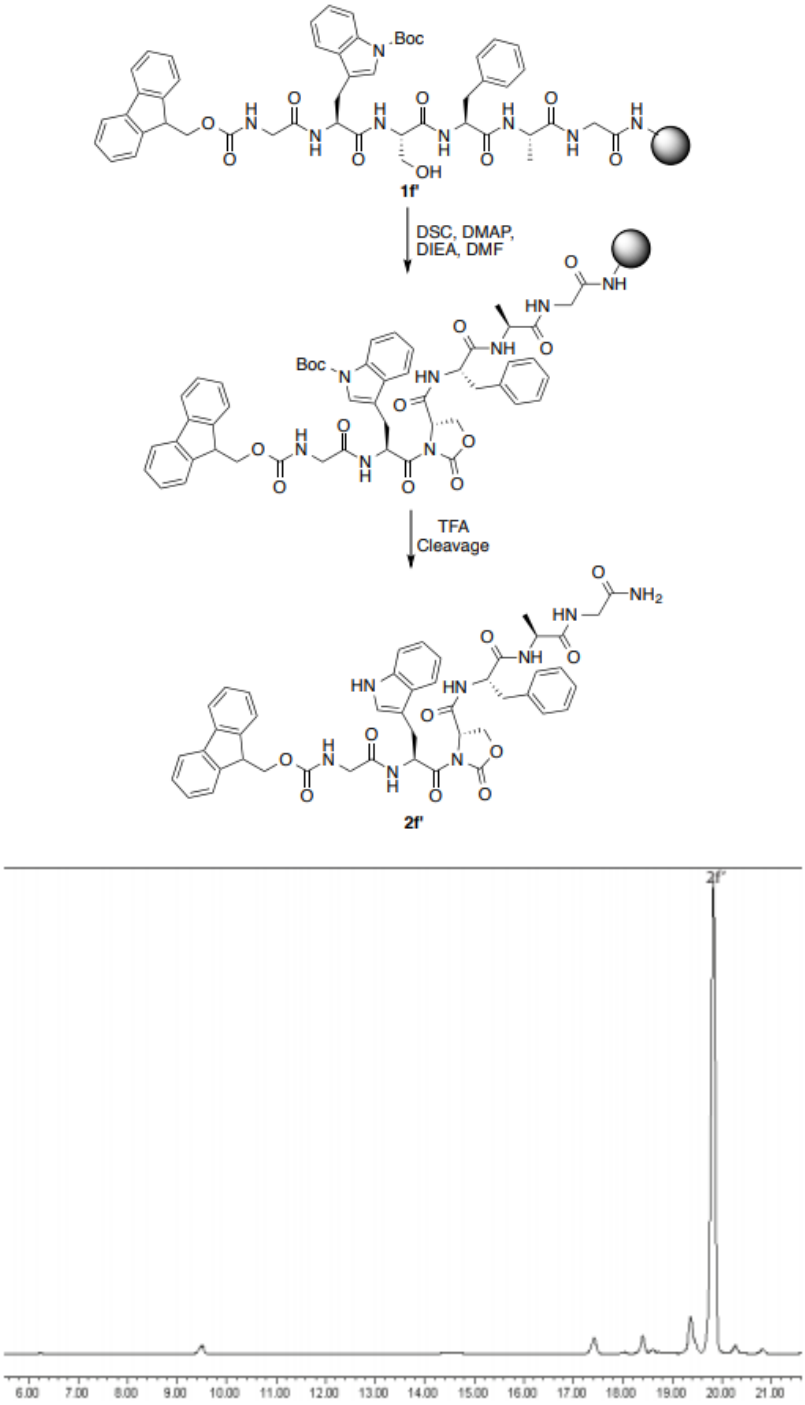


Figure A159 HPLC spectra of the cyclic urethane containing peptide 2f'
Fmoc-Gly-Trp-Oxd-Phe-Ala-Gly-NH2 (2f'). LCMS: m/z 871.20 (calcd [M+H]+ = 871.6), 436.0 (calcd (M+2) 2+ = 436.33). Purity: >95% (HPLC analysis at 254 nm). Retention time: 19.8 min.

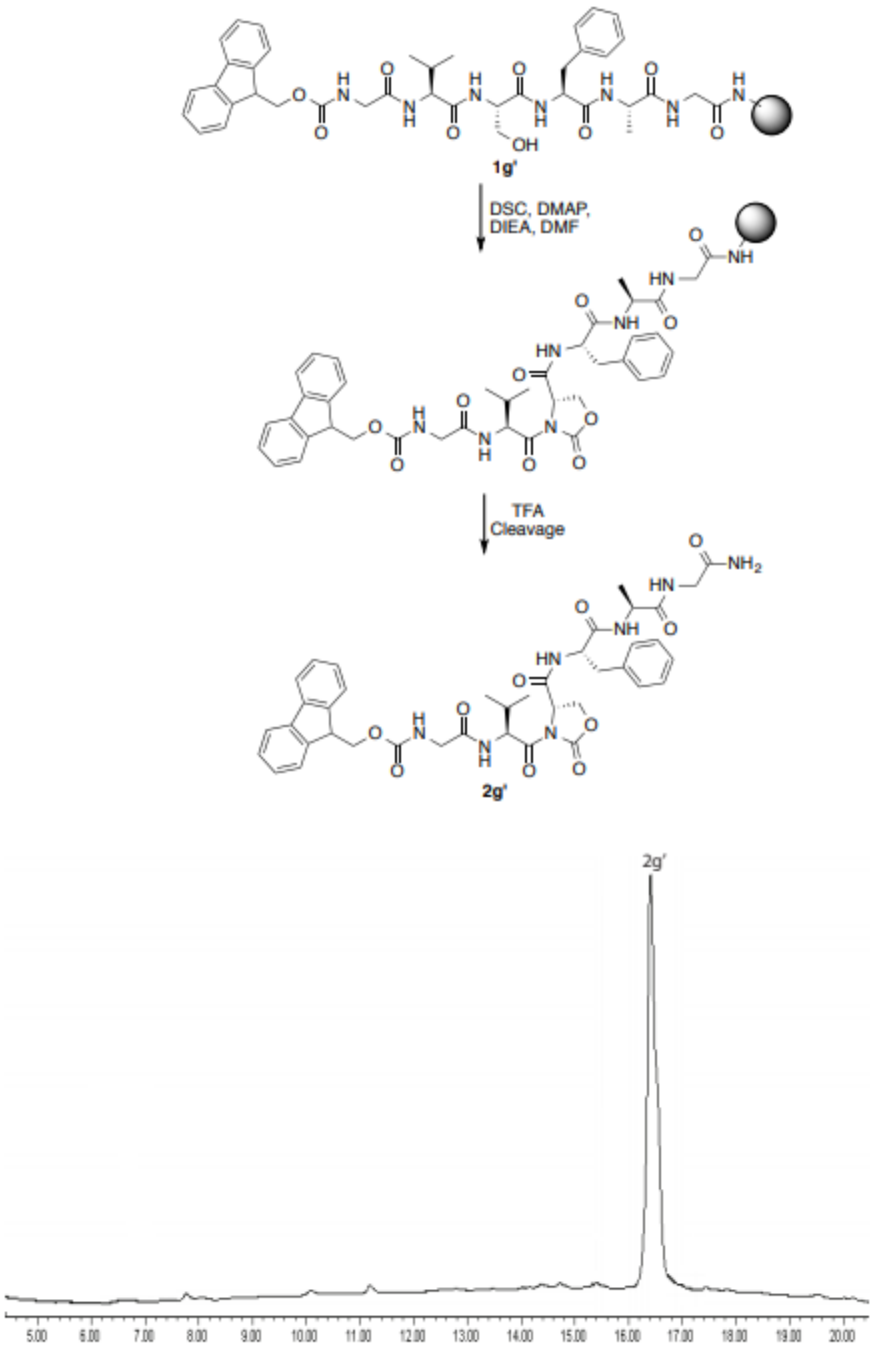


Figure A160 HPLC spectra of the cyclic urethane containing peptide 2g'
Fmoc-Gly-Val-Oxd-Phe-Ala-Gly-NH2 (2g'). LCMS: m/z 784.60 (calcd [M+H]+ = 784.58), 392.5 (calcd (M+2) 2+ = 392.79). Purity: >95% (HPLC analysis at 254 nm). Retention time: 16.6 min.