The development of β-sheet promoting amino acid residues and fundamental studies of peptide structure

by

Stacy J. Maynard

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: March 10, 2015

The dissertation is approved by the following members of the final oral committee:
Samuel H. Gellman, Professor, Chemistry
Silvia Cavagnero, Professor, Chemistry
Tehshik Yoon, Professor, Chemistry
Randall Goldsmith, Assistant Professor, Chemistry
Sandro Mecozzi, Professor, Pharmaceutical Sciences

ACKNOWLEDGEMENTS

I would like to thank my advisor, Sam Gellman, for his patience and support during my years in graduate school. I would also like to thank the members of the Gellman group for their help and support.

Collaboration

Dr. Jay Steinkruger
Dr. Aaron Almeida
Dr. David Mortenson
Dr. Michael Giuliano
Dr. Yasuharu Yoshimi

Mass Spectrometry

Martha Vestling

NMR Spectroscopy

Charlie Fry
Monika Ivancic

Editorial Assistance

Brian Fisher
Ross Cheloha
Marlies Hager
Dale Kreitler

Undergraduate Students
Amanda Frankiewicz
Kaitlyn Mayer
Shirley Nwangwa
Robert Gruener

Funding
NIH Grant (GM-061238)

NMR spectrometers were purchased with partial support from NIH (Grant No. 1 S10 RR13866-01)

MALDI-MS was purchased with support from NIH (Grant NCRR 1S10RR024601-01)

Table of Contents

Chapter 1. Introduction

1.1 Peptides, unnatural amino acids, and foldamers

Protein-protein interactions play important roles in a large number of biological processes. The ability to regulate specific protein-protein interactions is useful for treating a wide variety of diseases including HIV[1,2] and cancer.[3] Traditional small-molecule therapeutics often have difficulty disrupting protein-protein interactions that have a large surface area.[3] Peptide drugs, of intermediate size, can often target these interactions (Figure 1.1).[3] For this reason, peptide drugs are a growing class of therapeutics. In the first decade of the 2000s, 151 peptide drug candidates entered clinical trials.[3] Peptides also provide an advantage over larger molecules (such as proteins) as they can, in some cases, cross the cell membrane and target intracellular binding partners.[3]

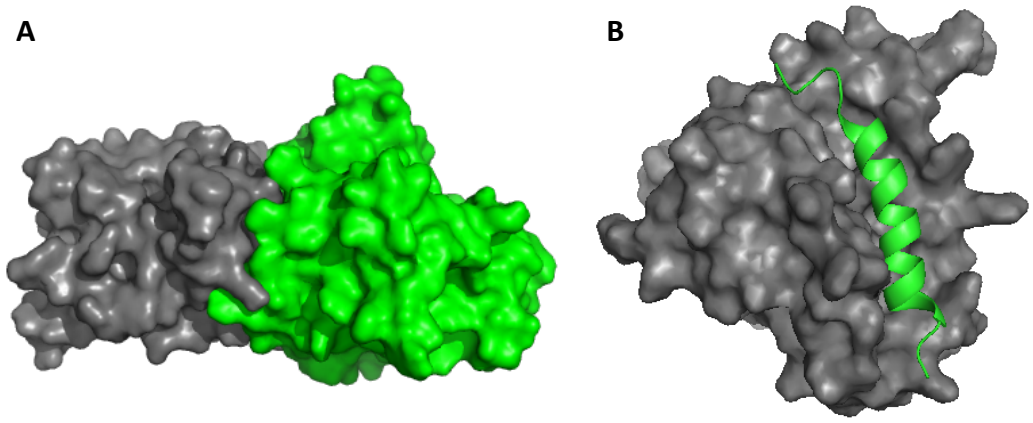
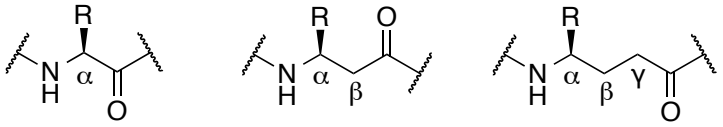


Figure 1.1. Example of a protein-protein interaction that can be disrupted by a peptide. (A) NMR structure of Bcl-XL-p53 complex (protein data bank (pdb) code 2MEJ).[4] (B) NMR structure of Bcl-XL and peptide inhibitor (pdb code 1G5J).[5]

Peptide drugs often adopt or mimic secondary structures found in naturally occurring peptides and proteins.[3] There are peptide or peptide-like therapeutic candidates that mimic common secondary structures such as α-helices and β-sheets.[3,6,7]

Peptides are well-studied and useful scaffolds for designing intermediate sized therapeutics, but we are not limited to naturally occurring amino acid residues as building blocks.[8,9] Intermediate-sized molecules built from unnatural scaffolds that fold into discrete secondary or tertiary structure are termed foldamers.[10] Often, foldamer scaffolds contain amino acids with expanded backbones (Figure 1.2).[8] Natural peptide scaffolds contain α-amino acids. β- and γ-amino acids are expanded by one and two backbone carbons, respectively (Figure 1.2).[8]



α-amino acids (proteinogenic) β-amino acids γ-amino acids

Figure 1.2 Classification of unnatural amino acids with expanded backbones.

Peptides that are made from mixtures of α- and β-amino acids can fold into helices that mimic the side-chain display and function of natural helices.[8] For example, gp41 is a protein from HIV that contains helices that mediate HIV-cell fusion through the formation of a six-helix bundle containing three helices from the CHR domain and three helices from the NHR domain.[2] Peptide 1.1, which mimics the structure of the CHR helices found in gp41, blocks the formation of HIV's six-helix bundle by binding the three NHR helices to form an alternative six-helix bundle that cannot mediate cell fusion (and therefore prevents cells from being infected by HIV).[1,2] When every three to four residues in peptide 1.1 are replaced with β-amino acids (peptide 1.2), the resulting foldamer has a similar structure as peptide 1.1 and HIV's CHR helices (Figure 1.3 and Figure 1.4).[6] Peptide 1.1 and 1.2 have a similar ability to prevent HIV-cell fusion. The

IC_{50} is concentration at which a peptide blocks half of cell-fusion events. Peptide 3.1 has an IC_{50} of cell fusion of 9 nM and peptide 1.2 has an IC_{50} of 5 nM.[6] Note that there are two types of β-residues used in peptide 1.2, β^3-residues and cyclic residues. β^3-residues can display the same side-chain groups as natural amino acids, but the extra methylene in the backbone increases the flexibility of the residue and generally destabilizes helical structure relative to the α-residue analogue.[6] The lost stability is regained by replacing residues in the helix with cyclic β-residues that are more rigid and enforce the helical structure.[6]

Ac-TTWEAWDRAIAEYAARIEALIRAAQEQQEKNEAALREL-NH_2 (Peptide 1.1)
Ac-TTWEXWDZAIAEYAXRIAXLIZAAQEQQEKNEXALZEL-NH_2 (Peptide 1.2)

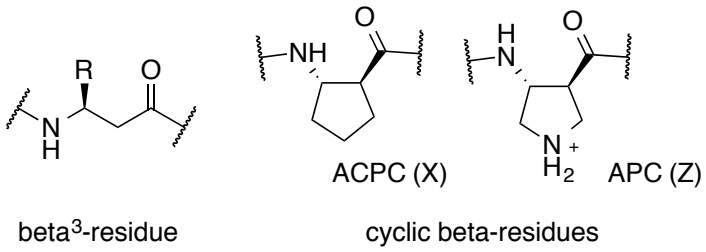


Figure 1.3. Primary sequences of two peptides derived from gp41.[6] β^3-amino acids are in blue. (β^3-amino acids residues are β-residues with the side-chain at the 3 position.) Cyclic β-amino acids are in orange. Amino acids are given by their one letter code. For β^3-amino acid residues, the one-letter codes mean the side chain group, R, is the same as the side-chain for the α-amino acid residue.

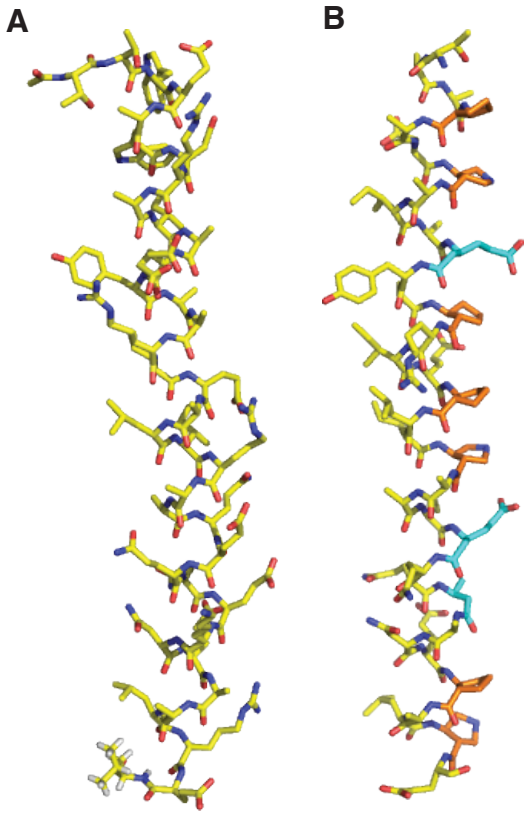


Figure 1.4. Crystal structures of (A) peptide 1.1, an α-peptide derived from gp41 (pdb code 3F4Y)[6] and (B) peptide 1.2, an α/β analogue of peptide 1.1. (pdb code 3F50)[6]. β^3-amino acids are in blue. Cyclic β-amino acids are in orange.

Proteinases are enzymes that break down peptides and proteins by hydrolyzing amide bonds. One major advantage of incorporating unnatural amino acids into peptides, or using entirely unnatural scaffolds, is that the residues usually increase protease resistance and therefore have longer half-lives *in vivo* relative to natural amino acid residues.[10] α-Peptide 1.1 has a half life of 0.7 minutes when incubated with Proteinase K, while peptide 1.2 has a half life of 200 minutes.[6] Note that the β-amino acids provide protease resistance for surrounding bonds between natural residues.[6] The increased protease resistance helps to address one of the major limitations of peptide therapeutics.[10]

In addition to increasing protease resistance, unnatural residues can promote structure that natural sequences do not readily adopt. For example, in helices formed by

proteinogenic amino acids, the directionality of the hydrogen bonds formed by the backbone is almost always uniform.[11] In other words, hydrogen-bond accepting backbone carbonyl groups point toward the C-terminus and hydrogen-bond donating backbone N-H groups point toward the N-terminus. However, peptides containing β and/or γ residues can promote mixed helices in which the direction of the hydrogen bonds alternates.[12–16] Natural amino acids are L-amino acids and their enantiomers are D-amino acids. As discussed in detail in section 1.4, DPro promotes two-residue turns compatible with β-sheet structures, while natural amino acids do not reliably do so.[17,18] Unnatural residues are useful as they can have properties natural residues lack, and the design of new amino acids to promote specific conformations is a rich area of new research. The design and synthesis of threonine derivatives for use in β-sheet secondary structure will be discussed in Chapters 2 and 3.

1.2 Methods of thermodynamic analysis of peptide structure

Methods that help us understand the structure of peptides and proteins are useful in the development of therapeutics. Several longstanding biophysical techniques have been used to characterize peptide and protein structure.[19] Herein, methods of thermodynamic characterization, both old and new, are briefly discussed. Many of these methods are utilized in the following chapters.

1.2.1 Chemical and thermal denaturation

Chemical and thermal denaturation have been used to explore the stability of folded peptide and protein structures.[20] For the purposes of this discussion, we will assume that the proteins and peptides of interest can either adopt a folded conformation

or an unstructured conformation (two-state approximation). In other words, folding is assumed to be cooperative. The folding equilibrium constant (K_F) between the folded and unstructured conformations is defined in Figure 1.5. The thermodynamic stability of folding (ΔG_F) can be calculated from the equation

$$\Delta G_F = -RT\ln K_F \quad \text{(Equation 1.1)}$$

where T is the temperature and R is the gas constant.

In a denaturation experiment, a spectroscopic signal is measured over increasing denaturant or temperature (Figure 1.5).[20] The spectroscopic signal must be sensitive to protein structure and must be different for the folded and unstructured states of the protein. Circular dichroism (CD) or fluorescence is often used. A typical denaturation curve has a region where most of the protein is in the folded state, a transition region where both the folded and denatured states are populated, and a region where the protein is in the denatured stated (Figure 1.5). The denatured state approximates the unstructured (random coil) conformation. The ΔG_F can be calculated for the transition region of the curve where the some of the peptide is in the folded state and some is in the denatured state.[20] The resulting denaturation curve for chemical denaturation can also be fitted to an equation

$$Y_{Obs} = \frac{Y_D + m_D[x] + (Y_N + m_N[x])e^{-(\Delta G^O{}_{H_2O} + m[x])/RT}}{1 + e^{-(\Delta G^O{}_{H_2O} + m[x])/RT}} \quad \text{(Equation 1.2)}$$

where Y_{Obs} is the observed spectroscopic signal, Y_N and Y_D are the y-intercepts of the folded and denatured states, m_N and m_D are the slope of the folded and denatured states with the variation of denaturant concentration, [x] is the concentration of the denaturant, and ΔG^o_{H2O} is the ΔG_F when no denaturant is present. An analogous equation can be

written for a thermal denaturation experiment. Equation 1.2 allows extrapolation of the ΔG_F to the physically relevant region of the curve (37ºC and no denaturant present) from the transition region of the curve.[21]

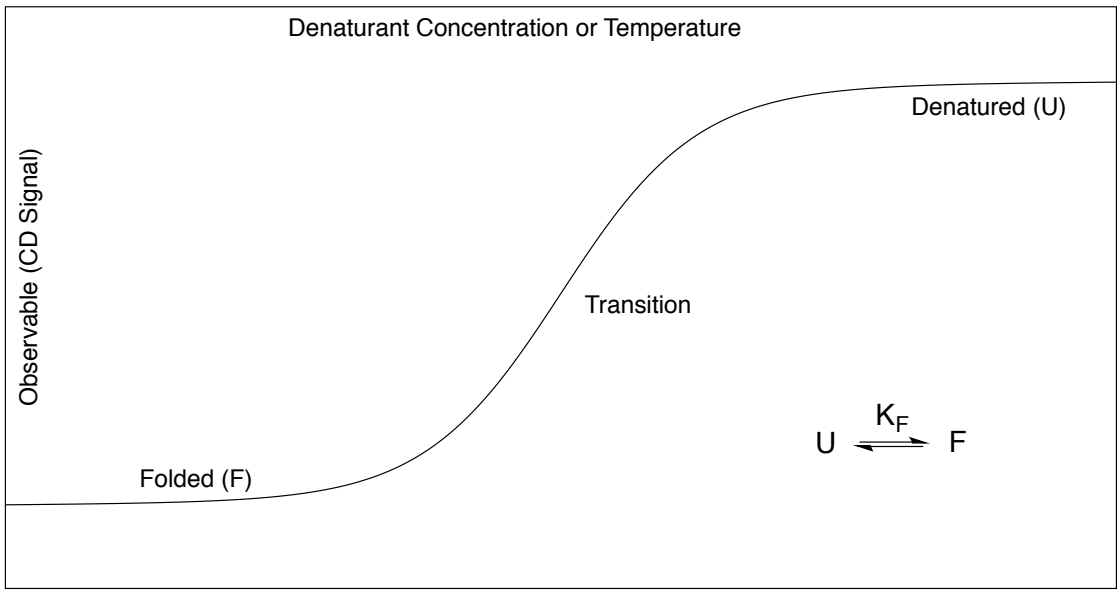


Figure 1.5. Theoretical denaturation curve for a two-state folding model

While chemical and thermal denaturation data can be acquired with minimal effort, most of the data acquired are taken under conditions that are not physiologically relevant (high temperature or high denaturant concentration).[22] Therefore, one must consider the ΔG_F measured from denaturation experiments with caution. Nevertheless, these experiments can be highly useful, provided the limitations are considered.[20]

1.2.2 NMR Analysis

NMR techniques allow the thermodynamic analysis of the folding of peptides and proteins. These studies can prove valuable data regarding stability and folding, particularly because NMR can provide residue and domain specific information for the

protein rather than making a global measurement of stability like denaturation experiments. However, it generally require time-consuming and/or complex data acquisition and data analysis.[23]

NMR and H/D Exchange

Amide protons in the backbone of a peptide or protein can potentially be exchanged for deuterium in a solvent that contains D_2O, MeOD, or other deuterated solvents with exchangeable deuteriums.[24] The rate at which various backbone amide hydrogens are exchanged for deuterium can provide information of the stability of a protein or peptide.[24] The slower amide hydrogens in the interior of the folded structure are exchanged for a deuterium, the higher the hydrogens' protection factors, and the more folded the protein. This method is particularly useful at obtaining information about the folding intermediates of proteins.[24] While folding is usually cooperative (which allows us to make the two-state approximation), evidence suggests that proteins fold through specific pathways where secondary structure generally forms faster than the tertiary structure. H/D exchange with partially denatured proteins provide residue by residue information on the folding intermediates of the protein. The residues that form structure early in the folding process will experience a rate of deuteration closer to that of the folded protein than residues that form structure late in the folding process. Techniques also exist that allow temporary denaturation of the protein. For examples, lasers can cause a temporary temperature increase in a sample that will denature the protein. Again, the residues that form structure early in the folding process with experience less deuteration than residues that form structure late in the folding process.[25]

One limitation of the using H/D exchange as a method for exploring stability is that it can be labor-intensive. Conditions under which the exchange takes place at a reasonable rate must be identified, and the peptide ^{1}H resonances must be assigned.[24] In many cases, peptides and proteins must be labeled with ^{15}N or ^{13}C in order to allow full assignment.[24] Labeling is usually cost prohibitive when peptides are synthesized chemically.[24]

Chemical shift analysis

Chemical shift analysis may provide qualitative thermodynamic data. In special cases, it can also provide quantitative thermodynamic analysis.[26,27] For more details regarding these calculations and for specific examples, see Chapters 2 and 3.

1.2.3 Calorimetry

Differential scanning calorimetry (DSC) measures the energy required to increase the temperature of a sample as it is heated.[28] When the peptide or protein of interest undergoes a thermal transition, the energy necessary to raise the temperature is increased. The curve that is generated can be used to obtain thermodynamic parameters such as ΔC_P, ΔH, and ΔG_F.[28] The thermodynamic parameters are measured at the transition temperature, and ΔG_F at physiologically relevant temperature is extrapolated.[28,29]

If the thermodynamic parameters of the interaction between two species are desired, isothermal titration calorimetry (ITC) may be utilized. The energy required to maintain a constant temperature as a substrate is added is measured and thermodynamic parameters are derived from the resulting curve.[30]

1.2.4 Disulfide exchange

Disulfide exchange is a technique that can be used to determine relative thermodynamic stabilities of the structured conformation of peptides for which the folded state contains disulfide bonds.[31–33] Disulfide bonds are labile at near-neutral pH in the presence of thiol groups.[31] If we add a thiol-containing peptide to a disulfide bond-containing peptide at pH 7.0, there can be disulfide exchange between the possible species (Figure 1.6). The equilibrium between the different possible species reflects the relative stability of the folded conformation of each disulfide-containing species. For example, suppose we wish to know which two-helix bundle is more stable, A or C (Figure 1.6).[33] If we allow disulfide exchange to occur and reach equilibrium, the equilibrium constant will favor the side of the reaction in which the more stable two-helix bundle is formed (Figure 1.6).[31] Quantitative analysis using disulfide exchange is complicated by the fact that even in a simple system, a large number of different disulfide complexes are possible.[31] However, qualitative interpretation of the data is useful.

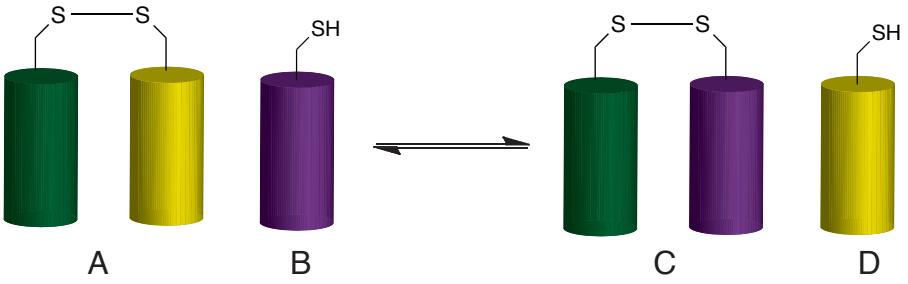


Figure 1.6. Disulfide exchange

1.2.5 Backbone thioester exchange (BTE)

Backbone thioester exchange (BTE) is similar to disulfide exchange in that it allows the measurement of thermodynamic stability from the relative concentration of

species at equilibrium, but it solves many of the difficulties associated with disulfide exchange.[34] This technique will be discussed in detail in Chapter 5.

1.3 α-Helices

1.3.1 Structure

Helices are one of the most common elements of secondary structure in peptides and proteins. Most helices are α-helices, which are right-handed helices with a heptad (seven-residue) repeat and 3.6 residues per turn of the backbone (Figure 1.7).[11] One heptad is shown in Figure 1.7b-d. The positions of the heptad repeat are designated a-g (Figure 1.7c).[11] The backbone amides groups form hydrogen bonds so that the donating N-H groups point toward the N-terminus of the helix (Figure 1.7d).[11] The alignment of the hydrogen bonds creates a net dipole over the length of the helix.[11]

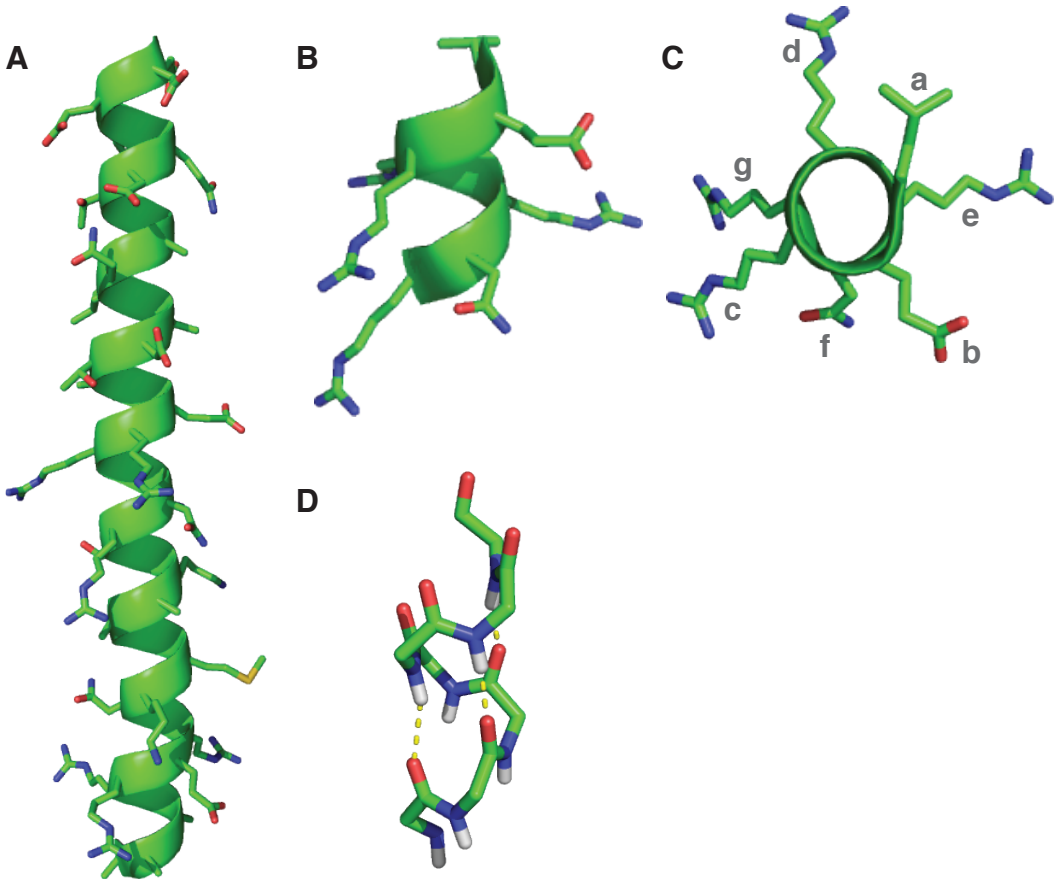


Figure 1.7. α-Helix structure from pdb (1FOS).[35] (A) Cartoon of an α-helix with its side-chain groups. (B) One heptad of an α-helix. The backbone of the helix is shown by the ribbon cartoon and the side chains are shown. (C) One heptad of an α-helix looking down the backbone of the helix. The positions in the heptad are labeled a-g. Each position in the heptad is located on a different side of the helix. The a-g positions in the next heptad of the helix will be approximately above the corresponding position in this heptad of the helix. (D) Backbone atoms of the α-helix. Hydrogen bonds are shown in yellow.

Certain amino acids stabilize helical structure more than others.[36] The ability of a residue to promote helical structure is known as helical propensity.[36] The classical propensity scale was developed by Chou and Fasman and is based on the occurrence of amino acids in helices in proteins of known structure (Table 1.1). [36]

Table 1.1. Helical propensity of naturally occurring amino acids.[36] Amino acids are named by their three letter codes.

Residue	Helical Propensity	Residue	Helical Propensity
Glu	1.51	Asp	1.01
Met	1.45	His	1.00
Ala	1.42	Arg	0.98
Leu	1.21	Thr	0.83
Lys	1.16	Ser	0.77
Phe	1.13	Cys	0.70
Gln	1.11	Tyr	0.69
Trp	1.08	Asn	0.67
Ile	1.08	Pro	0.57
Val	1.06	Gly	0.57

1.3.2 Coiled Coils

Two or more helices can associate to form coiled coils.[11] These structures have a left-handed superhelical twist that allows the a and d positions on each helix to remain in the core of the coiled coil and pack together in a 'knobs into holes' fashion.[11] Helices can associate in a parallel fashion, with the N termini aligned at one end, or antiparallel fashion, with the N-termini at opposite ends.[11] Figure 1.8 shows examples of dimeric, trimeric, and tetetrameric coiled coils from the literature. [35,37,38]

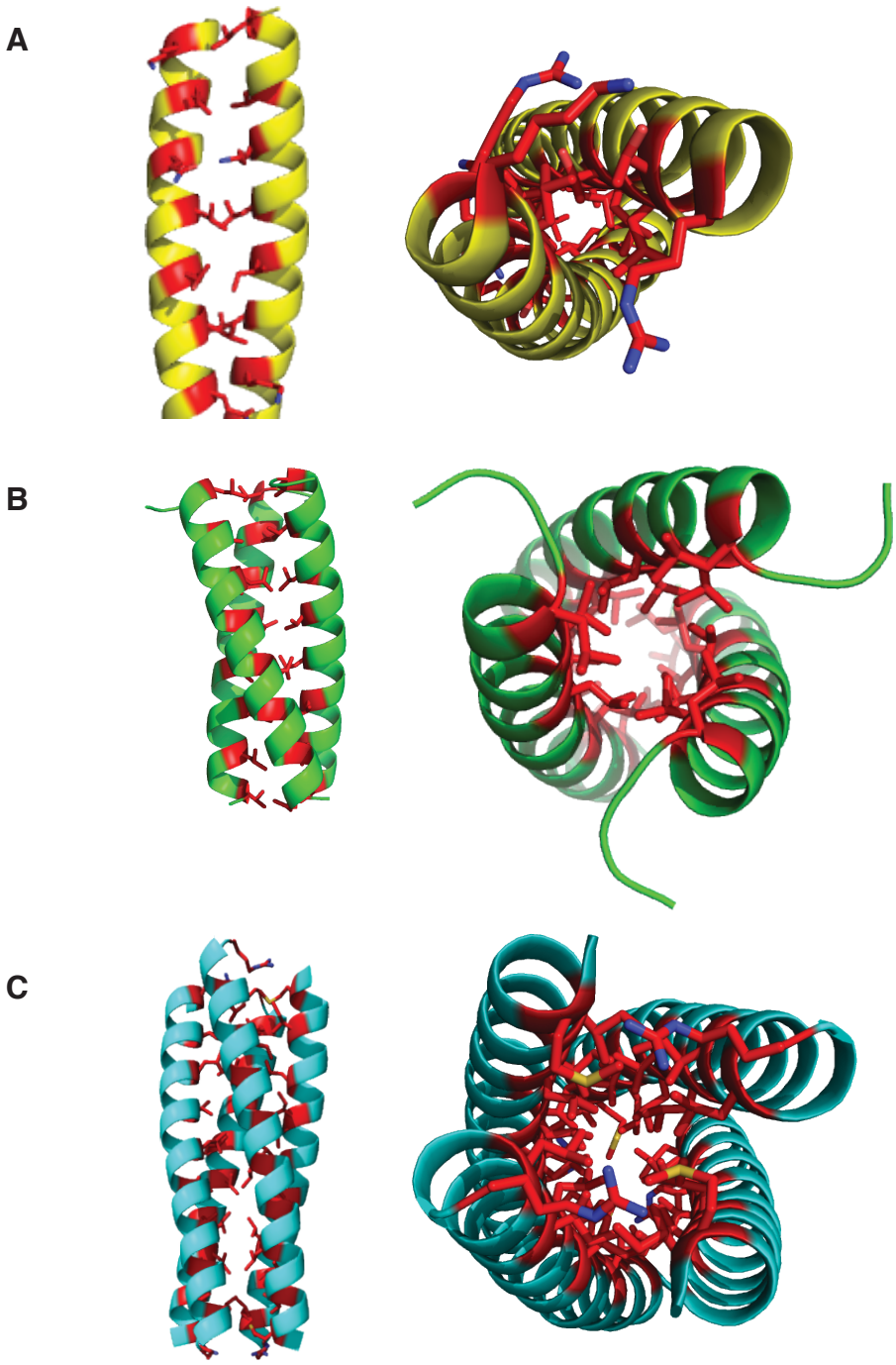


Figure 1.8. Coiled-coil (A) parallel dimer (pdb code 1FOS)[35], (B) parallel trimer (pdb code 1COI)[37], and (C) antiparallel tetramer (pdb code 1W5J)[38]. Core a and d positions are in red.

Coiled-coil interactions have been studied in great detail.[39–50] The stability of a coiled coil increases with length, provided that the residues in the added segments have a propensity to adopt the coiled-coil structure.[11] The alignment of helices (parallel or antiparallel) in a coiled coil can be controlled by electrostatic interactions at the e and g positions which flank the core a and d positions (Figure 1.7c).[32] Detailed pairing preferences of amino acids at the a, d, e, and g positions have been explored and are useful for the design of helical peptides.[39–50]

The specificities of natural coiled-coil sequences are essential to function.[11] For example, the bZIP family of proteins dimerize through coiled coil interfaces and regulate gene expression.[51] Computational studies from the Woolfson and Keating labs have focused on understanding and designing coiled-coil structures with specificity.[52–56] Work from the Woolfson lab has explored designing helix bundles with a specified number of helices.[57] The group has created a six-helix bundle that may be useful as a scaffold for catalysis as it has a hydrophobic interior with properties that can be tuned.[57]

1.3.3 Unnatural amino acids in helices

As mentioned in Section 1.1, peptides with unnatural amino acids residues can fold into helices that mimic α-helices. Figure 1.9 shows the structures of various non-α amino acids that can adopt helical structure.

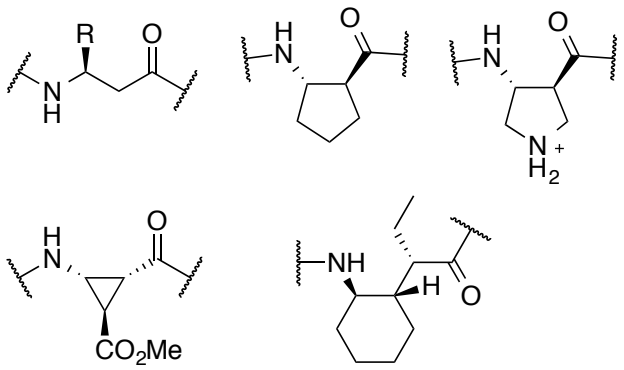
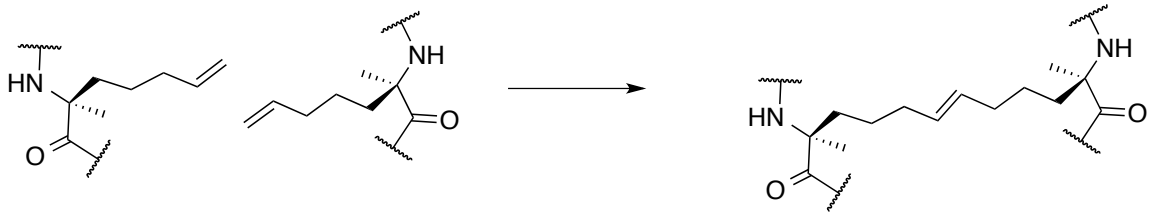


Figure 1.9. Examples of unnatural amino acids that adopt helical structure[6,8,58–60]

Natural and unnatural α-amino acids have been used to introduce 'staples' into peptides that create covalent bonds between side-chains of two different residues.[61–66] One of the most common methods for producing stapled peptides is to use alkene metathesis reactions (Scheme 1.1).[61,65,66] These staples are intended to stabilize helical structure by constraining it.[61–66]

Scheme 1.1. Metathesis reaction used to form staple peptides.[61,65,66]



1.4 β-sheets

1.4.1. Overview of β-sheet secondary structure

β-sheet secondary structure is common in peptides and proteins. β-sheets are formed when two or more strands of peptide associate and form a hydrogen bonding network through their backbone amide groups. A two-stranded β-sheet connected by a turn is a β-hairpin. β-sheet structures are parallel if backbones have the same orientation (the N-termini are at the same end of the hairpin) or antiparallel if the backbones have

opposite orientations (Figure 1.10 and 1.11). In both types of hairpins, the side-chains of the amino acids are alternately oriented above and below the plane of the backbone association (Figure 1.10c and 1.11c).[67]

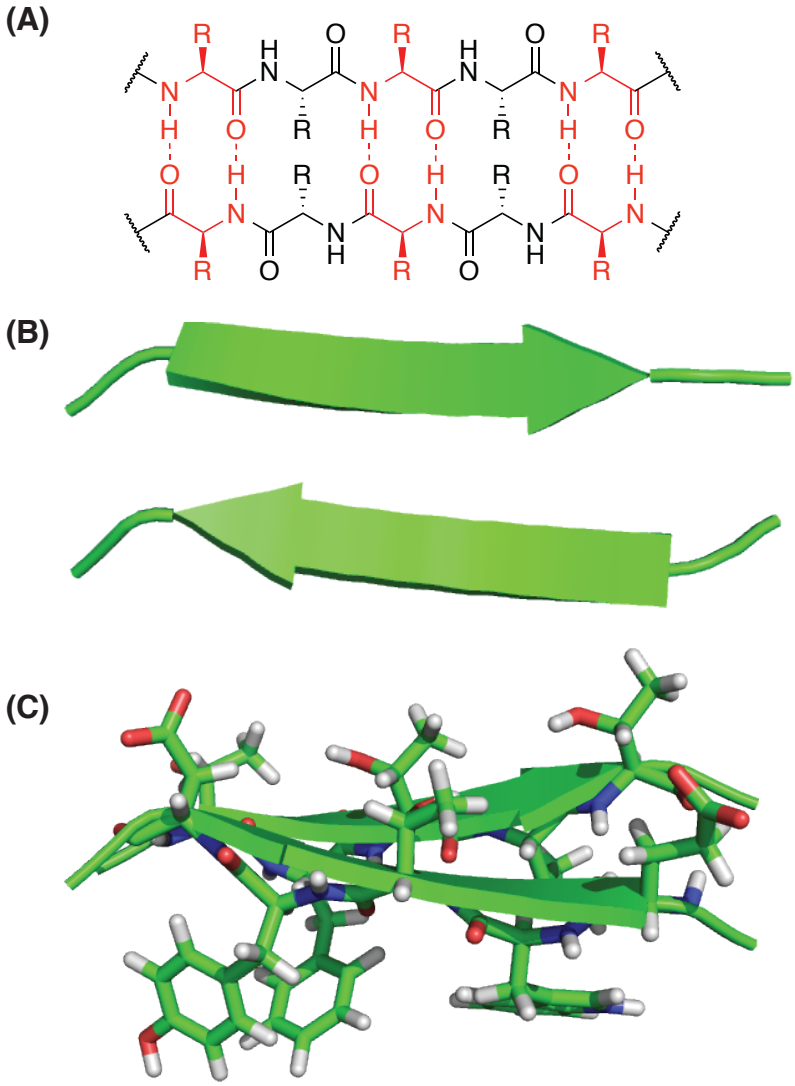


Figure 1.10. Antiparallel β-sheet structure. (A) Line representation. Hydrogen bonding pairs are shown in red. (B), (C) antiparallel sheet from GB1 (pdb code 2GB1).[68] Only backbone atoms are shown in (B) as indicated by the flat arrow, while (C) includes side-chain groups.

Antiparallel and parallel hairpins have different hydrogen-bonding patterns. In antiparallel sheets, residues alternate between those for which the backbone carbonyl and N-H hydrogen bond with the cross-strand partner residue (hydrogen bonding position)

and those for which backbone carbonyl and N-H are oriented in the opposite direction (non-hydrogen bonding position) (Figure 1.10). In parallel β-sheets, the backbone amides hydrogen bond with the amide groups from the residue on the opposing strand that is diagonal to its position (Figure 1.11).[69]

Strands of both antiparallel and parallel hairpins must be connected by intervening residues if they are part of the same peptide.[67] In the case of antiparallel hairpins, a minimum of two residues is needed to form the turn between the strands.[67] In the case of parallel hairpins, if only natural residues are present, longer intervening segments are necessary as the N-termini of the two strands must align.[67]

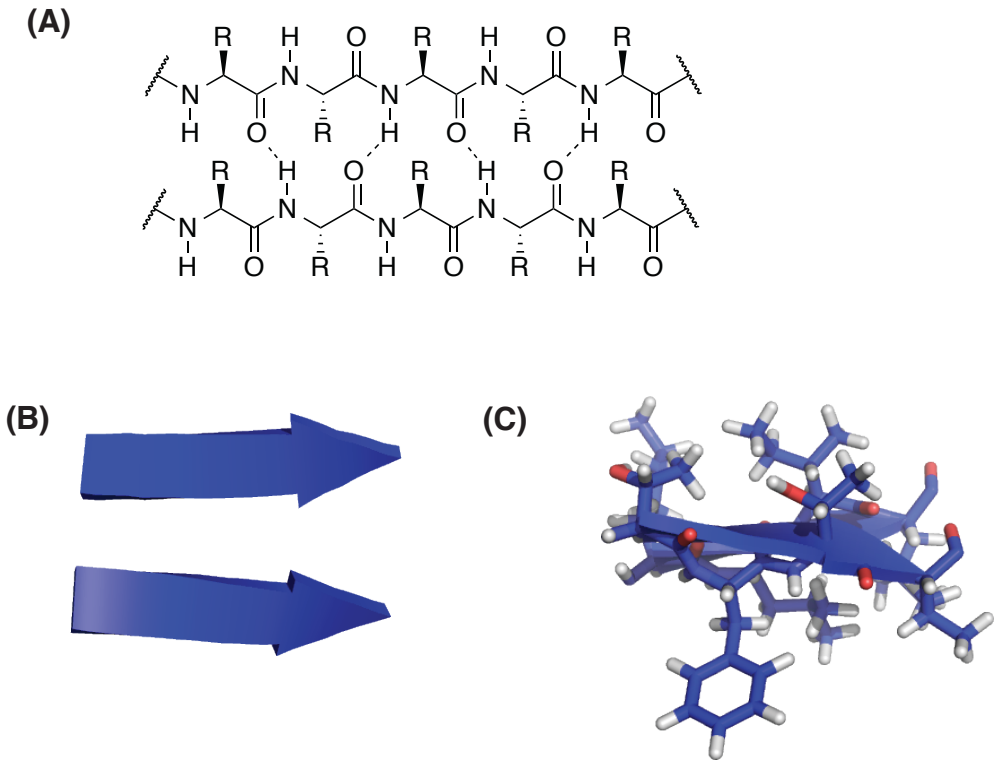


Figure 1.11. Parallel β-sheet structure. (A) Line representation. (B), (C) Parallel sheet from GB1 (pdb code 2GB1).[68] Only backbone atoms are shown in (B) as indicated by the flat arrow, while (C) includes side-chain groups.

In addition to β-strands associating in the lateral direction via the peptide backbone, β-sheets can pack on top of one another, an interaction mediated through side-chains.[69] These interactions form sandwich structures and are of particular interest due to their prevalence in the amyloid fibrils associated with diseases such as Alzheimer's and prions.[69] Crystal structures of the sandwich structures from sequences of amyloid-forming proteins have been obtained by the Eisenberg lab (Figure 1.12).[70,71]

In addition to their importance in amyloid fibrils, β-sheets are known to mediate a wide variety of protein-protein interactions. A database of β-sheet mediated interactions between different protein chains or different proteins was compiled by Dou *et al*.[72]

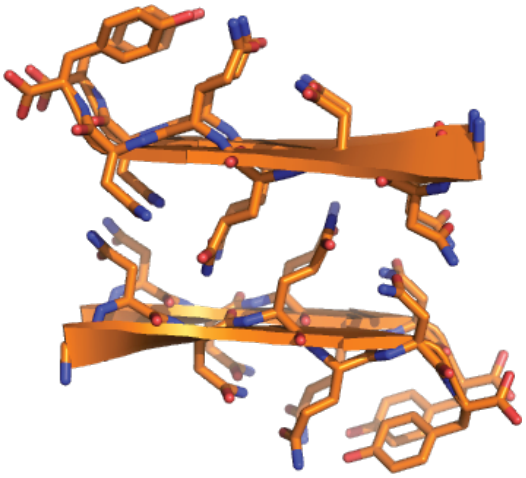


Figure 1.12 Crystal structure of a face to face β-sheet interaction (pdb code 1YJP)[70]

1.4.2 β-sheet propensity

The ability of an amino acid to promote β-sheet secondary structure is known as β-sheet propensity.[36] An understanding of propensity can aide in the design of β-sheet structures. Numerous propensity scales have been developed. The earliest, from Chou and Fasman, was based on the normalized occurrence of amino acids in β-sheets within

protein crystal structures.[36,73] Subsequent experimental propensity scales were based on model systems. Kim and Minor measured β-sheet propensity in the context of the GB1 protein.[74] Kim and Berg measured β-sheet propensity using a zinc finger binding protein. The relative β-sheet stabilities associated with different substitutions were calculated based on their ability to bind zinc, which is an indirect readout of stability.[75] The order and magnitude of the β-sheet propensity of various amino acids differs from scale to scale. However, most scales show the same general trends. Therefore, while propensity is not entirely context independent, propensity scales still provide useful information. Aromatic and β-branched residues tend to have high β-sheet propensities.[76]

Numerous theories have been proposed to explain β-sheet propensity in terms of the structure of the amino acid side chains.[77–83] One theory, from Bai and Englander, suggests that the ability of side-chain groups to block water from accessing the hydrogen bonding groups in the backbone is important in β-sheet formation. Therefore the better an amino acid is at preventing the solvation of the β-sheet backbone, the higher the β-sheet propensit.[79] A competing theory, from Swindells *et al*., suggests that the intrinsic Φ/Ψ angles of an amino acid, which are defined in Figure 1.13, are related to the β-sheet propensity of an amino acid.[76] This theory is discussed in more detail in Chapter 3.

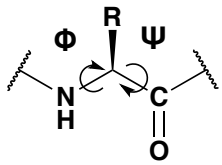


Figure 1.13. Φ/Ψ angles in a peptide backbone

1.4.3 Development of water-soluble β-sheet hairpins

Small autonomously folding models of secondary structure allow detailed analysis and understanding of the factors that influence the folding of peptides and proteins.[67] A β-hairpin is the minimal model for β-sheet secondary structure and consists of a two β-strands connected by a turn.[67] Small water-soluble β-hairpin model systems of antiparallel hairpins were first developed in the mid-nineties.[84–86] Prior to this time, many assumed it was not possible to form β-sheet secondary structure outside of the context of a specific tertiary structure.[84] Later, parallel hairpin model systems were developed using unnatural linkers for the turn segment.[87]

Antiparallel hairpin model systems

The first hairpin system was a nine-residue peptide developed by Nieto lab and had a two-residue turn. Including the adjacent residues, the turn sequence was Asn-Pro-Asp-Gly.[84] Subsequent studies demonstrated that controlling the nature of the turn was problematic with naturally occurring residues.[88] It was often difficult to predict whether a sequence would form a two-residue or three-residue turn, and in some cases a mixture of both hairpins was observed.[67] The inability to control the turn meant it was often difficult to predict the residue pairings, which is essential for a versatile β-sheet model system.[67]

The development of sequences containing DPro supplied a reliable method for designing a two-residue turn that promotes β-hairpin formation.[89] The design of the DPro-Gly turn was based on key features of turns in protein crystal structures. Thornton *et al*. analyzed hairpins found in protein structures.[90] Turns are classified by the number of residues in the turn and the Φ/Ψ angles of these residues around the backbone atoms of

the turn (Table 1.2).[90] The most common type of turns are Type I and Type II.[90] The inverse of Type I and II turns, Type I' and II', are much less common.[90] Based on computer modeling, Thornton *et al*. proposed that Type I' and II' turns are disfavored by steric clashes involving side chain groups of the natural amino acids.[90] Thornton *et al*. also analyzed the turns associated with antiparallel β-hairpins.[90] Even though Type I' and II' two-residue turns are uncommon in general, they were the most common turn structure in the context of a β-hairpin.[90] Thornton *et al*. explain this discrepancy with the observation β-hairpins are not flat, but have a twist that Type I' and II' turns accommodate but type I and II turns do not.[90]

Table 1.2. Common turn types[91–93]

Turn Type	Residue 1 Φ/Ψ (Degrees)	Residue 2 Φ/Ψ (Degrees)
Type I	-60, -30	-90, 0
Type I'	60, 30	90, 0
Type II	-60, 120	80, 0
Type II'	60, -120	-80, 0

Haque *et al*. explored the possibility that D-amino acids could be used to promote type I' and II' turns.[17] Given that natural L-amino acids prefer to adopt type I and II turns, D-amino acids should prefer to adopt turns that are the mirror image of type I and II turns.[17] These turns are by definition type I' and type II' turns.[17]

The turn sequences were initially explored in tetrameric depsipeptides and peptides.[17,89] It was found that DPro-Gly and DPro-DAla promoted the turn structure necessary for the formation of a hairpin with L-amino acids in the strands.[17,18,89] Turns with LPro-Gly and LPro-LAla did not promote hairpin structure with L-strands.[17,18,89]

Other linkers and scaffolds have been used to model antiparallel hairpins. The Nowick lab has developed an ornithine linker and an unnatural scaffold that enforces strand formation in water using amino acid Hao (Figure 1.14).[94] Hao replaces three residues in the strand of the β-sheet and mimics the antiparallel hydrogen bonding pattern displayed by the backbone atoms of natural amino acids.[94] Initial versions of this scaffold were not soluble in aqueous solvents and/or did not adopt sheet structure in aqueous solvents.[95,96] A crucial advance in this system was the ability to make it water-soluble as measurements in organic solvents may not be relevant to the aqueous environment.[94] Nowick and coworkers have since used this model system to study the packing of β-sheet interactions in face-to-face interactions (aka interactions mediated by side-chain groups).[71,97,98] It should be noted that the Hao residue does not display the side-chains of the residues it replaces. In addition, Hao does not display hydrogen bonding donors and acceptors that are properly positioned to hydrogen bond with the backbone of an external strand. Therefore, only one strand of the model system can bind another β-strand effectively.[94] Nowick exploited this feature of his model system to control and study both edge to edge and face to face (sandwich) β-sheet interactions (Figure 1.15).[71,94,96,97,99,100]

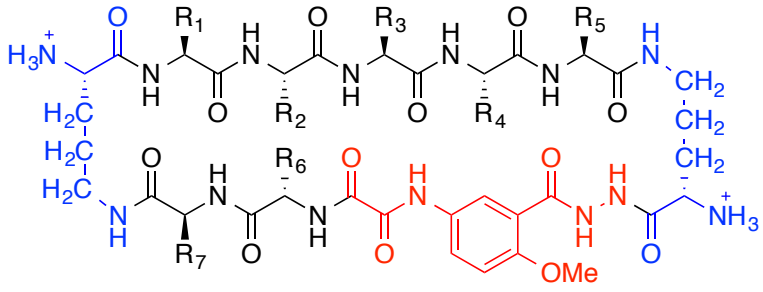


Figure 1.14. Nowick β-sheet scaffold.[94] The Hao residue is in red and the ornithine turns are in blue.

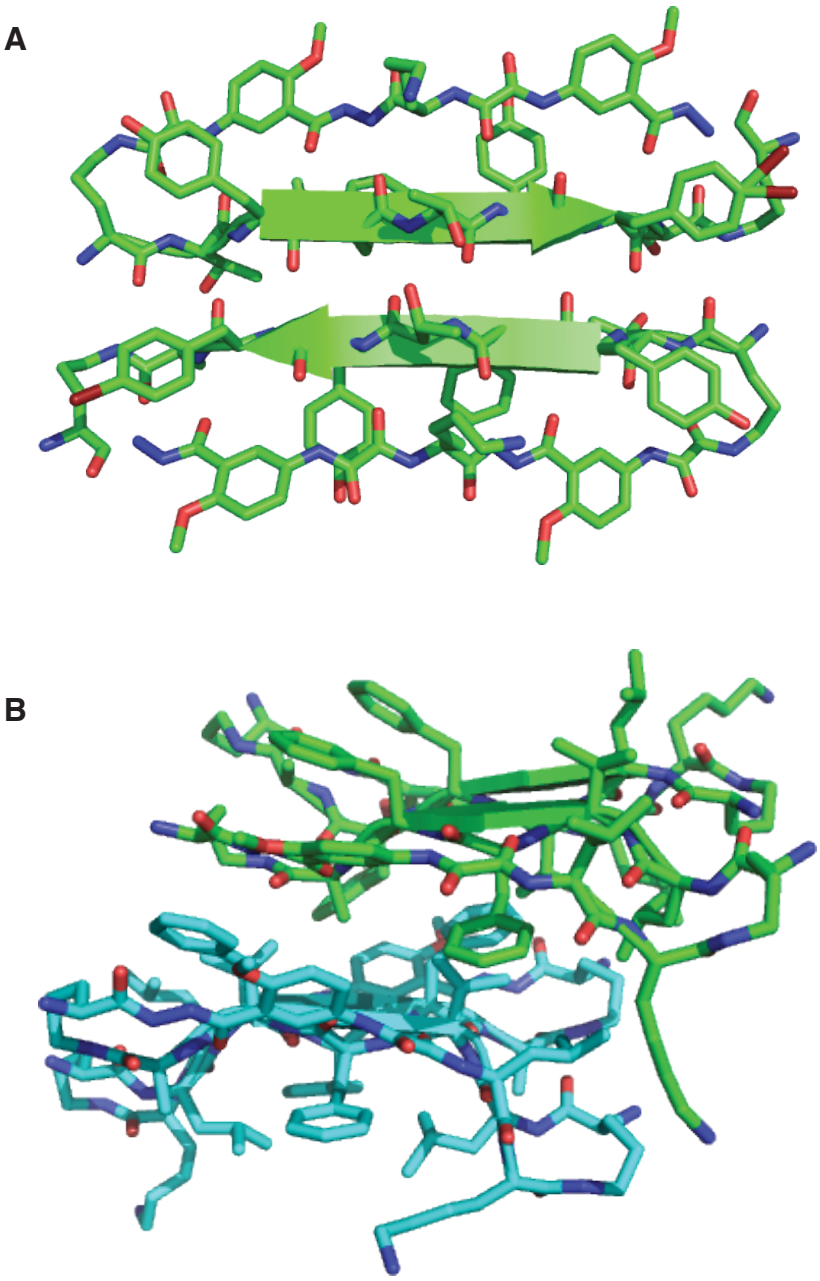


Figure 1.15. Crystal structures of Nowick model system that explore (A) edge to edge (antiparallel, pdb code 3NI3)[98] and (B) face to face (pdb code 3Q9H)[97] packing interactions of β-strands

Parallel hairpin model systems

Unlike antiparallel hairpins, the strands of parallel hairpins cannot be connect by short natural turn sequences as the N-termini align at one end of the parallel hairpin and

the C-termini align at the other end.[67] Initial attempts to develop parallel sheet-promoting turn sequences did not result in structures that were folded in aqueous solution.[95,101–106] Fisk *et al*.[107] noted that parallel sheet structure had a similar twist as antiparallel structure and therefore, since the Φ/Ψ angles of DPro favor a turn that accommodates the antiparallel twist for L-residue strands, the same amino acid might also promote a turn compatible with parallel β-sheets. A diamine, 1,2-diamino-1,1-dimethylethane (DADME), was also incorporated to link the two C termini of the parallel strands.[87,107] The resulting DPro-DADME turn promoted hairpin structure (Figure 1.16).[87,107]

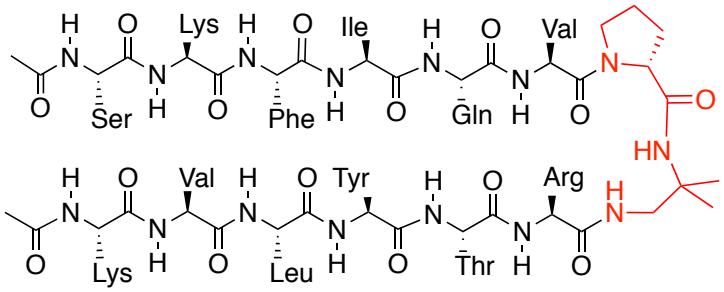


Figure 1.16. Parallel β-hairpin model system.[87] The two-residue turn sequence, DPro-DADME, is highlighted in red.

Unlike the antiparallel hairpins, parallel hairpins have two distinct ends: one with both N-termini and one with both C-termini. Therefore, two different linkers are required depending upon which end of the hairpin is used to connect the strands. A diacid linker, cis-cyclohexanedicarboxylic acid (CHDA), was developed to connect the N-termini (Figure 1.17).[108] The CHDA linker is more time-consuming and expensive to synthesize than the DPro-DADME linker. Therefore, use of the C-terminal DPro-Dadme linker is preferred when there is a choice. When a cyclized peptide is desired, the DPro-Dadme C-

terminal linker is sufficient to promote a stable β-sheet structure, so a simplified achiral diacid linker can be used in place of CHDA (Figure 1.18).[109]

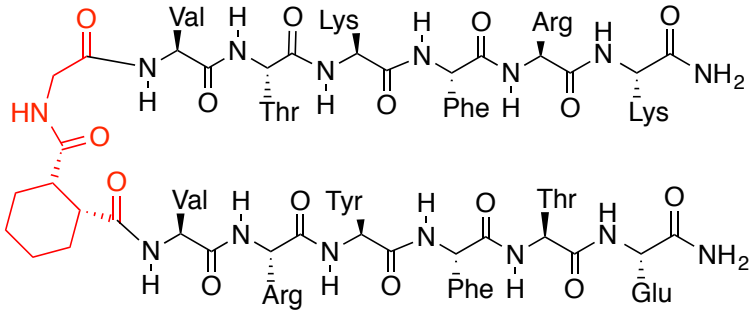


Figure 1.17 Parallel β-hairpin model system with a diacid linker (CHDA-Gly) that connects the N-termini of the strands, from Freire *et al*.[108] The linker is highlighted in red.

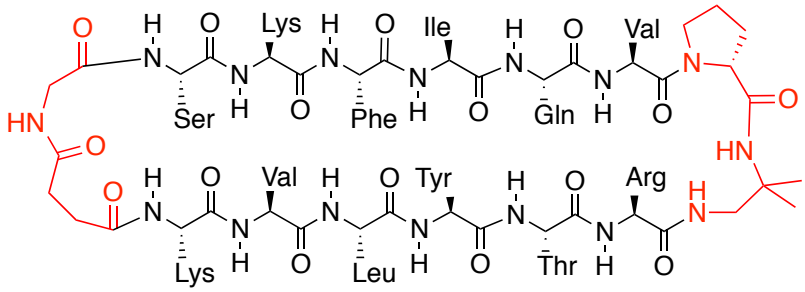


Figure 1.18 Parallel β-hairpin with DPro-Dadme C-termini linker and achiral N-termini linker.[109] Linkers are shown in red.

1.4.4. Brief summary of key thermodynamic observations

Model systems of parallel and antiparallel hairpins have lead to a better understanding of β-sheet structure. It has been shown that increasing the number of strands stabilizes the sheet.[110,111] This observation can be explained by the hypothesis that the formation of the initial hairpin has a higher entropic cost than the addition of a strand because the structure requires the organization of two strands of a β-sheet rather than one.[110,111] β-sheet structure can be extended by lengthening the strands. Evidence suggests that antiparallel and parallel hairpins may differ in the effect of length on secondary structure stability.[112] Cross-strand pairing preferences have also been explored.[113–117]

1.4.5. Unnatural amino acids in β-sheets

There are several examples in the literature of scaffolds that form structures reminiscent of β sheets.[118–127] Both β and γ residues can be used to form β-sheet like scaffolds. [118,122–124] Also, β hairpins constructed from both α and β residues have been reported.[121,125,127] It is unclear how well these structures mimic the presentation of side-chains by natural β-sheets as the expanded backbones affect side chain spacing. As mentioned above, Nowick has developed the unnatural amino acid Hao that can mimic the hydrogen-bonding pattern of a β-sheet and enforce β-sheet structure in the context of a hairpin.[94] While the Hao residue is extremely useful, it does not mimic the display of side-chains of natural amino acids.

1.4.6 β-sheets as therapeutics

There are several examples of peptides that adopt β-sheet structure and have useful therapeutic properties. Amyloid fibrils adopt β-sheet structures.[70] Several peptide hairpins that can bind to the proteins that form the fibrils have been shown to prevent fibril formation.[128,129] Nowick has used his unnatural β-hairpin scaffold for this purpose. [99,129] As the Hao residue on one face of his hairpin cannot associate with another strand, higher order oligomerization is prevented.[99,129]

Other potentially useful β-sheet peptide therapeutics have been explored.[130–133] For, example, a four-stranded antiparallel β-sheet peptide adopts a structure similar to that of gp120 from HIV.[134–136] This structure may be useful in the development of an HIV

vaccine.[136] PDZ domains are small modules involved in the regulation of signaling pathways. These domains bind hairpin substrates and have been inhibited by β haripins.[7]

1.5 References

(1) Matthews, T.; Salgo, M.; Greenberg, M.; Chung, J.; DeMasi, R.; Bolognesi, D. Enfuvirtide: The First Therapy to Inhibit the Entry of HIV-1 into Host CD4 Lymphocytes. *Nat. Rev. Drug Discov.* **2004**, *3*, 215–225.

(2) Moore, J. P.; Doms, R. W. The Entry of Entry Inhibitors: A Fusion of Science and Medicine. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 10598–10602.

(3) Tsomaia, N. Peptide Therapeutics: Targeting the Undruggable Space. *Eur. J. Med. Chem.* **2015**, In Press.

(4) Follis, A. V.; Llambi, F.; Ou, L.; Baran, K.; Green, D. R.; Kriwacki, R. W. The DNA-Binding Domain Mediates Both Nuclear and Cytosolic Functions of p53. *Nat. Struct. Mol. Biol.* **2014**, *21*, 535–543.

(5) Petros, a M.; Nettesheim, D. G.; Wang, Y.; Olejniczak, E. T.; Meadows, R. P.; Mack, J.; Swift, K.; Matayoshi, E. D.; Zhang, H.; Thompson, C. B.; et al. Rationale for Bcl-xL/Bad Peptide Complex Formation from Structure, Mutagenesis, and Biophysical Studies. *Protein Sci.* **2000**, *9*, 2528–2534.

(6) Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H. Structural and Biological Mimicry of Protein Surface Recognition by Alpha/beta-Peptide Foldamers. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14751–14756.

(7) Grillo-Bosch, D.; Choquet, D.; Sainlos, M. Inhibition of PDZ Domain-Mediated Interactions. *Drug Discov. Today Technol.* **2013**, *10*, 531–540.

(8) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A Field Guide to Foldamers. *Chem. Rev.* **2001**.

(9) Martinek, T. a; Fülöp, F. Peptidic Foldamers: Ramping up Diversity. *Chem. Soc. Rev.* **2012**, *41*, 687–702.

(10) Gellman, S. H. Foldamers : A Manifesto. **1998**, *31*, 173–180.

(11) Lupas, a. Coiled Coils: New Structures and New Functions. *Trends Biochem. Sci.* **1996**, *21*, 375–382.

(12) Seebach, D.; Gademann, K.; Schreiber, J. V; Matthews, J. L.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H. "Mixed" B-Peptides. A Unique Helical Secondary Structure in Solution. Preliminary Communication. *Helv. Chim. Acta* **1997**, *80*, 2033–2038.

(13) Rueping, M.; Schreiber, J. V.; Seebach, D. Mixed Beta-2/beta-3-Hexapeptides and Beta-2/beta-3-Nonapeptides Folding to (P)-Helices with Alternating Twelve- and Ten-Membered Hydrogen-Bonded Rings. *Helv. Chim. Acta* **2002**, *85*, 2577–2593.

(14) Sharma, G. V. M.; Jadhav, V. B.; Ramakrishna, K. V. S.; Jayaprakash, P.; Narsimulu, K.; Subash, V.; Kunwar, a C. 12/10- and 11/13-Mixed Helices in Alpha/gamma and Beta/gamma-Hybrid Peptides Containing C-Linked Carbo-Gamma-Amino Acids with Alternating Alpha- and Beta-Amino Acids. *J. Am. Chem. Soc.* **2006**, 14657–14668.

(15) Giuliano, M. W.; Maynard, S. J.; Almeida, A. M.; Reidenbach, A. G.; Guo, L.; Ulrich, E. C.; Guzei, I. a.; Gellman, S. H. Evaluation of a Cyclopentane-Based Gamma-Amino Acid for the Ability to Promote Alpha/gamma-Peptide Secondary Structure. *J. Org. Chem.* **2013**, *78*, 12351–12361.

(16) Giuliano, M. W.; Maynard, S. J.; Almeida, A. M.; Guo, L.; Guzei, I. A.; Spencer, L. C.; Gellman, S. H. A Gamma - Amino Acid That Favors 12/10-Helical Secondary Structure in Alpha / Gamma - Peptides. *J. Am. Chem. Soc.* **2014**, *136*, 15046–15053.

(17) Haque, T. S.; Little, J. C.; Gellman, S. H. “Mirror Image” Reverse Turns Promote Beta-Hairpin Formation. *J. Am. Chem. Soc.* **1994**, *116*, 4105–4106.

(18) Haque, T. S.; Gellman, S. H. Insights on Beta-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I′ and Type II′ Beta-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

(19) Privalov, P. L. Thermodynamic Problems of Protein Structure. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 47–69.

(20) Street, T. O.; Courtemanche, N.; Barrick, D. Protein Folding and Stability Using Denaturants. *Methods Cell Biol.* **2008**, *84*, 295–325.

(21) Pace, C. N.; Shaw, K. L. Linear Extrapolation Method of Analyzing Solvent Denaturation Curves. **2000**, *7*, 1–7.

(22) Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. Urea Promotes Polyproline II Helix Formation: Implications for Protein Denatured States. *Biochemistry* **2005**, *44*, 6269–6275.

(23) Wurthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(24) Englander, S. W.; Mayne, L. Using Hydrogen-Exchange Labeling and Two Dimensional Nmr. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 243–265.

(25) Englander, S. W. Protein Folding Intermediates and Pathways Studied by Hydrogen Exchange. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213–238.

(26) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quanfification of Beta-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

(27) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry* **1992**, *31*, 1647–1651.

(28) Castronuovo, G. Proteins in Aqueous Solutions. Calorimetric Studies and Thermodynamic Characterization. *Thermochim. Acta* **1991**, *193*, 363–390.

(29) Bruylants, G.; Wouters, J.; Michaux, C. Differential Scanning Calorimetry in Life Science: Thermodynamics, Stability, Molecular Recognition and Application in Drug Design. *Curr. Med. Chem.* **2005**, *12*, 2011–2020.

(30) Jelesarov, I.; Bosshard, H. R. Isothermal Titration Calorimetry and Differential Scanning Calorimetry as Complementary Tools to Investigate the Energetics of Biomolecular Recognition. *J. Mol. Recognit.* **1999**, *12*, 3–18.

(31) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wieter, J.-L.; Sanders, J. K. M.; Otto, S. Dynamic Combinatorial Chemistry. *Chem. Rev.* **2006**, *106*, 3652–3711.

(32) Oakley, M. G.; Kim, P. S. A Buried Polar Interaction Can Direct the Relative Orientation of Helices in a Coiled Coil. *Biochemistry* **1998**, *37*, 12603–12610.

(33) Lai, J. R.; Fisk, J. D.; Weisblum, B.; Gellman, S. H. Hydrophobic Core Repacking in a Coiled-Coil Dimer via Phage Display : Insights into Plasticity and Specificity at a Protein - Protein Interface. **2004**, 10514–10515.

(34) Woll, M. G.; Gellman, S. H. Backbone Thioester Exchange: A New Approach to Evaluating Higher Order Structural Stability in Polypeptides. *J. Am. Chem. Soc.* **2004**, *126*, 11172–11174.

(35) Glover, J. N.; Harrison, S. C. Crystal Structure of the Heterodimeric bZIP Transcription Factor c-Fos-c-Jun Bound to DNA. *Nature*, 1995, *373*, 257–261.

(36) Chou, P. Y.; Fasman, G. D. Secondary Structural Prediction of Proteins from Their Amino Acid Sequence. *Trends Biochem. Sci.* **1977**, *2*, 128–131.

(37) Yadav, M. K.; Leman, L. J.; Price, D. J.; Brooks, C. L.; Stout, C. D.; Ghadiri, M. R. Coiled Coils at the Edge of Configurational Heterogeneity. Structural Analyses of Parallel and Antiparallel Homotetrameric Coiled Coils Reveal Configurational Sensitivity to a Single Solvent-Exposed Amino Acid Substitution. *Biochemistry* **2006**, *45*, 4463–4473.

(38) Ogihara, N. L.; Weiss, M. S.; Degrado, W. F.; Eisenberg, D. The Crystal Structure of the Designed Trimeric Coiled Coil Coil-VaLd: Implications for Engineering Crystals and Supramolecular Assemblies. *Protein Sci.* **1997**, *6*, 80–88.

(39) Kohn, W. D.; Kay, C. M.; Hodges, R. S. Orientation, Positional, Additivity, and Oligomerization-State Effects of Interhelical Ion Pairs in Alpha-Helical Coiled-Coils. *J. Mol. Biol.* **1998**, *283*, 993–1012.

(40) Monera, O. D.; Kay, C. M.; Hodges, R. S. Electrostatic Interactions Control the Parallel and Antiparallel Orientation of Alpha-Helical Chains in Two-Stranded Alpha-Helical Coiled-Coils. *Biochemistry* **1994**, *33*, 3862–3871.

(41) Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S. The Role of Position a in Determining the Stability and Oligomerization State of Alpha-Helical Coiled Coils: 20 Amino Acid Stability Coefficients in the Hydrophobic Core of Proteins. *Protein Sci.* **1999**, *8*, 2312–2329.

(42) Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C. T.; Hodges, R. S. Effects of Side-Chain Characteristics on Stability and Oligomerization State of a de Novo-Designed Model Coiled-Coil: 20 Amino Acid Substitutions in Position "D". *J. Mol. Biol.* **2000**, *300*, 377–402.

(43) Wagschal, K.; Tripet, B.; Hodges, R. S. De Novo Design of a Model Peptide Sequence to Examine the Effects of Single Amino Acid Substitutions in the Hydrophobic Core on Both Stability and Oligomerization State of Coiled-Coils. *J. Mol. Biol.* **1999**, *285*, 785–803.

(44) Acharya, A.; Ruvinov, S. B.; Gal, J.; Moll, J. R.; Vinson, C. A Heterodimerizing Leucine Zipper Coiled Coil System for Examining the Specificity of a Position Interactions: Amino Acids I, V, L, N, A, and K. *Biochemistry* **2002**, *41*, 14122–14131.

(45) Acharya, A.; Rishi, V.; Vinson, C. Stability of 100 Homo and Heterotypic Coiled-Coil a-a' Pairs for Ten Amino Acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **2006**, *45*, 11324–11332.

(46) Krylov, D.; Mikhailenko, I.; Vinson, C. A Thermodynamic Scale for Leucine Zipper Stability and Dimerization Specificity: E and G Interhelical Interactions. *EMBO J.* **1994**, *13*, 2849–2861.

(47) Krylov, D.; Barchi, J.; Vinson, C. Inter-Helical Interactions in the Leucine Zipper Coiled Coil Dimer: pH and Salt Dependence of Coupling Energy between Charged Amino Acids. *J. Mol. Biol.* **1998**, *279*, 959–972.

(48) Phelan, P.; Gorfe, A. a.; Jelesarov, I.; Marti, D. N.; Warwicker, J.; Bosshard, H. R. Salt Bridges Destabilize a Leucine Zipper Designed for Maximized Ion Pairing between Helices. *Biochemistry* **2002**, *41*, 2998–3008.

(49) Marti, D. N.; Bosshard, H. R. Inverse Electrostatic Effect: Electrostatic Repulsion in the Unfolded State Stabilizes a Leucine Zipper. *Biochemistry* **2004**, *43*, 12436–12447.

(50) Engel, D. E.; DeGrado, W. F. Amino Acid Propensities Are Position-Dependent throughout the Length of Alpha-Helices. *J. Mol. Biol.* **2004**, *337*, 1195–1205.

(51) Chinenov, Y.; Kerppola, T. K. Close Encounters of Many Kinds: Fos-Jun Interactions That Mediate Transcription Regulatory Specificity. *Oncogene* **2001**, *20*, 2438–2452.

(52) Grigoryan, G.; Keating, A. E. Structure-Based Prediction of bZIP Partnering Specificity. *J. Mol. Biol.* **2006**, *355*, 1125–1142.

(53) Grigoryan, G.; Reinke, A. W.; Keating, A. E. Design of Protein-Interaction Specificity Gives Selective bZIP-Binding Peptides. *Nature* **2009**, *458*, 859–864.

(54) Reinke, A. W.; Grant, R. a.; Keating, A. E. A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering. *J. Am. Chem. Soc.* **2010**, *132*, 6025–6031.

(55) Chen, T. S.; Reinke, A. W.; Keating, A. E. Design of Peptide Inhibitors That Bind the bZIP Domain of Epstein-Barr Virus Protein BZLF1. *J. Mol. Biol.* **2011**, *408*, 304–320.

(56) Woolfson, D. N.; Bartlett, G. J.; Bruning, M.; Thomson, A. R. New Currency for Old Rope: From Coiled-Coil Assemblies to Alpha-Helical Barrels. *Curr. Opin. Struct. Biol.* **2012**, *22*, 432–441.

(57) Burton, A. J.; Thomas, F.; Agnew, C.; Hudson, K. L.; Halford, S. E.; Brady, R. L.; Woolfson, D. N. Accessibility, Reactivity, and Selectivity of Side Chains within a Channel of de Novo Peptide Assembly. *J. Am. Chem. Soc.* **2013**, *135*, 12524–12527.

(58) Koglin, N.; Zorn, C.; Beumer, R.; Cabrele, C.; Bubert, C.; Sewald, N.; Reiser, O.; Beck-Sickinger, A. G. Analogues of Neuropeptide Y Containing Beta-Aminocyclpropane Carboxylic Acids Are the Shortest Linear Peptides That Are Selective for the Y1 Receptor. *Angew. Chemie Int. Ed.* **2003**, *42*, 202–205.

(59) De Pol, S.; Zorn, C.; Klein, C. D.; Zerbe, O.; Reiser, O. Surprisingly Stable Helical Conformations in Alpha/beta-Peptides by Incorporation of Cis-Beta-Aminocyclopropane Carboxylic Acids. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 511–514.

(60) Appella, D. H.; Christianson, L. a.; Klein, D. a.; Richards, M. R.; Powell, D. R.; Gellman, S. H. Synthesis and Structural Characterization of Helix-Forming Beta-Peptides: Trans-2-Aminocyclopentanecarboxylic Acid Oligomers. *J. Am. Chem. Soc.* **1999**, *121*, 7574–7581.

(61) Blackwell, H. E.; Grubbs, R. H. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angew. Chemie - Int. Ed.* **1998**, *37*, 3281–3284.

(62) Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. General Approach to the Synthesis of Short .alpha.-Helical Peptides. *J. Am. Chem. Soc.* **1991**, *113*, 9391–9392.

(63) Bracken, C.; Gulyas, J.; Taylor, J. W.; Baum, J. Synthesis and Nuclear Magnetic Resonance Structure Determination of an .alpha.-Helical, Bicyclic, Lactam-Bridged Hexapeptide. *J. Am. Chem. Soc.* **1994**, *116*, 6431–6432.

(64) Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. A General Method for Constraining Short Peptides to an Alpha-Helical Conformation. *J. Am. Chem. Soc.* **1997**, *119*, 455–460.

(65) Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. Activation of Apoptosis in Vivo by a Hydrocarbon-Stapled BH3 Helix. *Science* **2004**, *305*, 1466–1470.

(66) Schafmeister, C. E.; Po, J.; Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* **2000**, *122*, 5891–5892.

(67) Gellman, S. H. Minimal Model Systems for Beta Sheet Secondary Structure in Proteins. *Curr. Opin. Chem. Biol.* **1998**, *2*, 717–725.

(68) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. A Novel, Highly Stable Fold of the Immunoglobulin Binding Domain of the Streptococcal Protein G. *Science (80-.).* **1991**, *253*, 657–661.

(69) Cheng, P. N.; Pham, J. D.; Nowick, J. S. The Supramolecular Chemistry of Beta-Sheets. *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.

(70) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the Cross-Beta Spine of Amyloid-like Fibrils. *Nature* **2005**, *435*, 773–778.

(71) Liu, C.; Zhao, M.; Jiang, L.; Cheng, P.-N.; Park, J.; Sawaya, M. R.; Pensalfini, A.; Gou, D.; Berk, A. J.; Glabe, C. G.; et al. Out-of-Register Beta-Sheets Suggest a Pathway to Toxic Amyloid Aggregates. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20913–20918.

(72) Dou, Y.; Baisnée, P. F.; Pollastri, G.; Pécout, Y.; Nowick, J.; Baldi, P. ICBS: A Database of Interactions between Protein Chains Mediated by Beta-Sheet Formation. *Bioinformatics* **2004**, *20*, 2767–2777.

(73) Chou, P. Y.; Fasman, G. D. Conformational Parameters for Amino Acids in Helical, Beta-Sheet, and Random Coil Regions Calculated from Proteins. *Biochemistry* **1974**, *13*, 211–222.

(74) Minor, D. L.; Kim, P. S. Measurement of the Beta-Sheet-Forming Propensities of Amino Acids. *Nature* **1994**, *367*, 660–663.

(75) Kim, C. A.; Berg, J. M. Thermodynamic Beta-Sheet Propensities Measured Using a Zinc-Finger Host Peptide. *Nature* **1993**, *362*, 267–270.

(76) Swindells, M. B.; MacArthur, M. W.; Thornton, J. M. Intrinsic Phi, Psi Propensities of Amino Acids, Derived from the Coil Regions of Known Structures. *Nat. Struct. Biol.* **1995**, *2*, 596–603.

(77) Stapley, B. J.; Doig, a J. Free Energies of Amino Acid Side-Chain Rotamers in Alpha-Helices, Beta-Sheets and Alpha-Helix N-Caps. *J. Mol. Biol.* **1997**, *272*, 456–464.

(78) Niwa, T. S.; Ogino, A. Multiple Regression Analysis of the Beta-Sheet Propensity of Amino Acids. *J. Mol. Struct. Theochem* **1997**, *419*, 155–160.

(79) Bai, Y.; Englander, S. W. Hydrogen Bond Strength and Beta-Sheet Propensities: The Role of a Side Chain Blocking Effect. *Proteins* **1994**, *18*, 262–266.

(80) Street, A. G.; Mayo, S. L. Intrinsic Beta-Sheet Propensities Result from van Der Waals Interactions between Side Chains and the Local Backbone. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9074–9076.

(81) Avbelj, F.; Baldwin, R. L. Role of Backbone Solvation in Determining Thermodynamic Beta Propensities of the Amino Acids. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1309–1313.

(82) Rossmeisl, J.; Kristensen, I.; Gregersen, M.; Jacobsen, K. W.; Nørskov, J. K. B-Sheet Preferences from First Principles. *J. Am. Chem. Soc.* **2003**, *125*, 16383–16386.

(83) Koh, E.; Kim, T.; Cho, H. S. Mean Curvature as a Major Determinant of Beta-Sheet Propensity. *Bioinformatics* **2006**, *22*, 297–302.

(84) Blanco, F. J.; Jimbnez, M. A.; Herranz, J.; Rim, M.; Santoro, J.; Nieto, J. L. NMR Evidence of a Short Linear Peptide That Folds Into a Beta-Hairpin in Aqueous Solution. *J. Am. Chem. Soc.* **1993**, 5887–5888.

(85) Blanco, F. J.; Rivas, G.; Serrano, L. A Short Linear Peptide That Folds into a Native Stable Beta-Hairpin in Aqueous Solution. *Nat. Struct. Biol.* **1994**, *1*, 584–590.

(86) Searle, M. S.; Williams, D. H.; Packman, L. C. A Short Linear Peptide Derived from the N-Terminal Sequence of Ubiquitin Folds into a Water-Stable Non-Native Beta-Hairpin. *Nat. Struct. Biol.* **1995**, *2*, 999–1006.

(87) Fisk, J. D.; Gellman, S. H. A Parallel Beta-Sheet Model System That Folds in Water. *J. Am. Chem. Soc.* **2001**, *123*, 343–344.

(88) De Alba, E.; Jiménez, M. A.; Rico, M. Turn Residue Sequence Determines Beta-Hairpin Conformation in Designed Peptides. *J. Am. Chem. Soc.* **1997**, *119*, 175–183.

(89) Haque, T. S.; Little, J. C.; Gellman, S. H. Stereochemical Requirements for B-Hairpin Formation: Model Studies with Four-Residue Peptides and Depsipeptides. *J. Am. Chem. Soc.* **1996**, *118*, 6975–6985.

(90) Sibanda, B. L.; Thornton, J. M. Beta-Hairpin Families in Globular Proteins. *Nature* **1985**, *316*, 170–174.

(91) Venkatachalam, C. M. Stereochemical Criteria for Polypeptides and Proteins. V. Conformation of a System of Three Linked Peptide Units. *Biopolymers* **1968**, *6*, 1425–1436.

(92) Chou, Y.; Fasman, G. D. Beta-Turns in Proteins. *J. Mol. Biol.* **1977**, *115*, 135–175.

(93) Sibanda, B. L.; Blundell, T. L.; Thornton, J. M. Conformation of Beta-Hairpins in Protein Structures. A Systematic Classification with Applications to Modelling by Homology, Electron Density Fitting and Protein Engineering. *J. Mol. Biol.* **1989**, *206*, 759–777.

(94) Nowick, J. S.; Lam, K. S.; Khasanova, T. V; Kemnitzer, W. E.; Maitra, S.; Mee, H. T.; Liu, R. An Unnatural Amino Acid That Induces Beta-Sheet Folding and Interaction in Peptides. *J. Am. Chem. Soc.* **2002**, *124*, 4972–4973.

(95) Nowick, J. S. Chemical Models of Protein ??-Sheets. *Acc. Chem. Res.* **1999**, *32*, 287–296.

(96) Nowick, J. S. Exploring Beta-Sheet Structure and Interactions with Chemical Model Systems. *Acc. Chem. Res.* **2008**, *41*, 1319–1330.

(97) Liu, C.; Sawaya, M. R.; Cheng, P. N.; Zheng, J.; Nowick, J. S.; Eisenberg, D. Characteristics of Amyloid-Related Oligomers Revealed by Crystal Structures of Macrocyclic Beta-Sheet Mimics. *J. Am. Chem. Soc.* **2011**, *133*, 6736–6744.

(98) Khakshoor, O.; Lin, A. J.; Korman, T. P.; Sawaya, M. R.; Tsai, S.-C.; Eisenberg, D.; Nowick, J. S. X-Ray Crystallographic Structure of an Artificial Beta-Sheet Dimer. *J. Am. Chem. Soc.* **2010**, *132*, 11622–11628.

(99) Cheng, P.-N.; Liu, C.; Zhao, M.; Eisenberg, D.; Nowick, J. S. Amyloid Beta-Sheet Mimics That Antagonize Protein Aggregation and Reduce Amyloid Toxicity. *Nat. Chem.* **2012**, *4*, 927–933.

(100) Khakshoor, O.; Demeler, B.; Nowick, J. S. Macrocyclic Beta-Sheet Peptides That Mimic Protein Quaternary Structure through Intermolecular Beta-Sheet Interactions. *J. Am. Chem. Soc.* **2007**, *129*, 5558–5569.

(101) Chitnumsub, P.; Fiori, W. R.; Lashuel, H. a.; Diaz, H.; Kelly, J. W. The Nucleation of Monomeric Parallel Beta-Sheet-like Structures and Their Self-Assembly in Aqueous Solution. *Bioorganic Med. Chem.* **1999**, *7*, 39–59.

(102) Kemp, D. S.; Bowen, R. Synthesis of Peptide-Functionalized Diacylaminopindolidiones As Templates for Beta-Sheet Formation. *Tetrahedron Lett.* **1988**, 5077–5080.

(103) Winningham, M. J.; Sogah, D. Y. A Modular Approach to Polymer Architecture Control via Catenation of Prefabricated Biomolecular Segments: Polymers Containing Parallel Beta-Sheets Templated by a Phenoxathiin-Based Reverse Turn Mimic. *Macromolecules* **1997**, *30*, 862–876.

(104) Skar, M. L.; Svendsen, J. S. Conformationally Restricted Diamines as Spacers for Parallel Beta-Sheet Formation. *Tetrahedron* **1997**, *53*, 17425–17440.

(105) Jones, I. G.; Jones, W.; North, M. Conformational Analysis of Peptides and Pseudopeptides Incorporating an Endo-(2S, 3R)-Norborn-5-Ene Residue as a Turn Inducer. **1998**, *3263*, 1505–1513.

(106) Ranganathan, D.; Haridas, V.; Kurur, S.; Thomas, A.; Madhusudanan, K. P.; Nagaraj, R.; Kunwar, a. C.; Sarma, a. V. S.; Karle, I. L. Demonstration of Endo-Cis-(2S,3R)-Bicyclo[2.2.1]hept-5-En-2,3 Dicarbonyl Unit as a Reverse-Turn Scaffold and Nucleator of Two-Stranded Parallel Beta-Sheets: Design, Synthesis, Crystal Structure, and Self-Assembling Properties of Norborneno Peptide Analogu. *J. Am. Chem. Soc.* **1998**, *120*, 8448–8460.

(107) Fisk, J. D.; Powell, D. R.; Gellman, S. H. Control of Hairpin Formation via Proline Configuration in Parallel Beta- Sheet Model Systems. *J. Am. Chem. Soc.* **2000**, *122*, 5443–5447.

(108) Freire, F.; Fisk, J. D.; Peoples, A. J.; Ivancic, M.; Guzei, I. a; Gellman, S. H. Diacid Linkers That Promote Parallel Beta-Sheet Secondary Structure in Water. *J. Am. Chem. Soc.* **2008**, *130*, 7839–7841.

(109) Freire, F.; Gellman, S. H. Macrocyclic Design Strategies for Small, Stable Parallel Beta-Sheet Scaffolds. *J. Am. Chem. Soc.* **2009**, *131*, 7970–7972.

(110) Sharman, G. J.; Searle, M. S. Cooperative Interaction between the Three Strands of a Designed Antiparallel Beta-Sheet. *J. Am. Chem. Soc.* **1998**, *120*, 5291–5300.

(111) Probe, T.; Schenck, H. L.; Gellman, S. H. Use of a Designed Triple-Stranded Antiparallel Beta-Sheet to Probe Beta-Sheet Cooperativity in Aqueous Solution. *J. Am. Chem. Soc.* **1998**, *120*, 4869–4870.

(112) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. Length-Dependent Stability and Strand Length Limits in Antiparallel Beta-Sheet Secondary Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12015–12020.

(113) Kuo, L. H.; Li, J. H.; Kuo, H. T.; Hung, C. Y.; Tsai, H. Y.; Chiu, W. C.; Wu, C. H.; Wang, W. R.; Yang, P. A.; Yao, Y. C.; et al. Effect of Charged Amino Acid Side Chain Length at Non-Hydrogen Bonded Strand Positions on ??-Hairpin Stability. *Biochemistry* **2013**, *52*, 7785–7797.

(114) Kuo, H. T.; Fang, C. J.; Tsai, H. Y.; Yang, M. F.; Chang, H. C.; Liu, S. L.; Kuo, L. H.; Wang, W. R.; Yang, P. A.; Huang, S. J.; et al. Effect of Charged Amino Acid Side Chain Length on Lateral Cross-Strand Interactions between Carboxylate-Containing Residues and Lysine Analogues in a Beta-Hairpin. *Biochemistry* **2013**, *52*, 9212–9222.

(115) Ramírez-Alvarado, M.; Blanco, F. J.; Serrano, L. De Novo Design and Structural Analysis of a Model Beta-Hairpin Peptide System. *Nat. Struct. Mol. Biol.* **1996**, *3*, 604–612.

(116) De Alba, E.; Rico, M.; Jiménez, M. A. Cross-Strand Side-Chain Interactions versus Turn Conformation in Beta-Hairpins. *Protein Sci.* **1997**, *6*, 2548–2560.

(117) Tatko, C. D.; Waters, M. L. Selective Aromatic Interactions in Beta-Hairpin Peptides. *J. Am. Chem. Soc.* **2002**, *124*, 9372–9373.

(118) Woll, M. G.; Lai, J. R.; Guzei, I. A.; Taylor, S. J. C.; Smith, M. E. B.; Gellman, S. H. Parallel Sheet Secondary Structure in Gamma-Peptides. *J. Am. Chem. Soc.* **2001**, *123*, 11077–11078.

(119) Lingard, H.; Han, J. T.; Thompson, A. L.; Leung, I. K. H.; Scott, R. T. W.; Thompson, S.; Hamilton, A. D. Diphenylacetylene-Linked Peptide Strands Induce Bidirectional B-Sheet Formation. *Angew. Chemie - Int. Ed.* **2014**, *53*, 3650–3653.

(120) Dutt, A.; Dutta, A.; Kar, S.; Koley, P.; Drew, M. G. B.; Pramanik, A. Stabilization of Two Smallest Possible Diastereomeric Beta-Hairpins in a Water Soluble Tetrapeptide Containing Non-Coded Beta-Amino Isobutyric Acid (Aib) and M-Amino Benzoic Acid. *J. Mol. Struct.* **2009**, *928*, 138–143.

(121) Lengyel, G. A.; Horne, W. S. Design Strategies for the Sequence-Based Mimicry of Side-Chain Display in Protein Beta-Sheets by Alpha/beta-Peptides. *J. Am. Chem. Soc.* **2012**, *134*, 15906–15913.

(122) Krauthauser, S.; Christianson, L. A.; Popwel, D. R.; Gellman, S. H. Antiparallel Sheet Formation in Beta-Peptide Foldamers: Effects of Beta-Amino Acid Substitution on Conformational Preference. *J. Am. Chem. Soc.* **1997**, *119*, 11719–11720.

(123) Martinek, T. a.; Tóth, G. K.; Vass, E.; Hollósi, M.; Fülöp, F. Cis-2-Aminocyclopentanecarboxylic Acid Oligomers Adopt a Sheetlike Structure: Switch from Helix to Nonpolar Strand. *Angew. Chemie - Int. Ed.* **2002**, *41*, 1718–1721.

(124) Ganesh Kumar, M.; Mali, S. M.; Raja, K. M. P.; Gopi, H. N. Design of Stable B-Hairpin Mimetics through Backbone Disulfide Bonds. *Org. Lett.* **2015**, *17*, 230–233.

(125) Karle, I. L.; Gopi, H. N.; Balaram, P. Peptide Hybrids Containing Alpha - and Beta-Amino Acids: Structure of a Decapeptide Beta-Hairpin with Two Facing Beta-Phenylalanine Residues. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3716–3719.

(126) Karle, I. Supramolecular Chemistry And Self-Assembly Special Feature: Infinite Pleated Beta -Sheet Formed by the Beta -Hairpin Boc-Beta -Phe-Beta -Phe-D-Pro-Gly-Beta -Phe-Beta -Phe-OMe. *Proc. Natl. Acad. Sci.* **2002**, *99*, 5160–5164.

(127) Seebach, D.; Abele, S.; Gademann, K.; Jaun, B. Pleated Sheets and Turns of ??-Peptides with Proteinogenic Side Chains. *Angew. Chemie - Int. Ed.* **1999**, *38*, 1595–1597.

(128) Sievers, S. a; Karanicolas, J.; Chang, H. W.; Zhao, A.; Jiang, L.; Zirafi, O.; Stevens, J. T.; Münch, J.; Baker, D.; Eisenberg, D. Structure-Based Design of Non-Natural Amino-Acid Inhibitors of Amyloid Fibril Formation. *Nature* **2011**, *475*, 96–100.

(129) Zheng, J.; Liu, C.; Sawaya, M. R.; Vadla, B.; Khan, S.; Woods, R. J.; Eisenberg, D.; Goux, W. J.; Nowick, J. S. Macrocyclic Beta-Sheet Peptides That Inhibit the Aggregation of a Tau-Protein-Derived Hexapeptide. *J. Am. Chem. Soc.* **2011**, *133*, 3144–3157.

(130) Khachatoorian, R.; Arumugaswami, V.; Ruchala, P.; Raychaudhuri, S.; Maloney, E. M.; Miao, E.; Dasgupta, A.; French, S. W. A Cell-Permeable Hairpin Peptide Inhibits Hepatitis C Viral Nonstructural Protein 5A-Mediated Translation and Virus Production. *Hepatology* **2012**, *55*, 1662–1672.

(131) Osmakov, D. I.; Kozlov, S. a.; Andreev, Y. a.; Koshelev, S. G.; Sanamyan, N. P.; Sanamyan, K. E.; Dyachenko, I. a.; Bondarenko, D. a.; Murashev, A. N.; Mineev, K. S.; et al. Sea Anemone Peptide with Uncommon Beta-Hairpin Structure Inhibits Acid-Sensing Ion Channel 3 (ASIC3) and Reveals Analgesic Activity. *J. Biol. Chem.* **2013**, *288*, 23116–23127.

(132) Kim, S.; Kim, D.; Jung, H. H.; Lee, I. H.; Kim, J. Il; Suh, J. Y.; Jon, S. Bio-Inspired Design and Potential Biomedical Applications of a Novel Class of High-Affinity Peptides. *Angew. Chemie - Int. Ed.* **2012**, *51*, 1890–1894.

(133) Fischer, S.; Geyer, A. Filaggrin Peptides with Beta-Hairpin Structure Bind Rheumatoid Arthritis Antibodies. *Angew. Chemie - Int. Ed.* **2014**, *53*, 3849–3853.

(134) Chakraborty, K.; Shivakumar, P.; Raghothama, S.; Varadarajan, R. NMR Structural Analysis of a Peptide Mimic of the Bridging Sheet of HIV-1 gp120 in Methanol and Water. *Biochem. J.* **2005**, *390*, 573–581.

(135) Sharma, D.; Balamurali, M. M.; Chakraborty, K.; Kumaran, S.; Jeganathan, S.; Rashid, U.; Ingallinella, P.; Varadarajan, R. Protein Minimization of the gp120 Binding Region of Human CD4+. *Biochemistry* **2005**, *44*, 16192–16202.

(136) Schiavone, M.; Fiume, G.; Caivano, A.; de Laurentiis, A.; Falcone, C.; Masci, F. F.; Iaccino, E.; Mimmi, S.; Palmieri, C.; Pisano, A.; et al. Design and Characterization of a Peptide Mimotope of the HIV-1 gp120 Bridging Sheet. *Int. J. Mol. Sci.* **2012**, *13*, 5674–5699.

Chapter 2. Synthesis and Evaluation of Cationic Threonine, an Amino Acid Designed to Promote β-Sheet Secondary Structure

Portions of this chapter appear in:

Maynard, S. J.; Almeida, A. M.; Yoshimi, Y.; Gellman, S. H. "New Charge-Bearing Amino Acid Residues that Promote β-Sheet Secondary Structure." *J. Am. Chem. Soc.* **2014**, 16683-16688.

2.1 Introduction

Peptides that adopt secondary or tertiary structures found in proteins are useful both as tools for exploring the fundamental properties of structure and as tools for binding proteins and blocking protein-protein interactions.[1–6] α-Helices and β-sheets are two of the most common types of secondary structure. α-Helices have been studied in detail and there are many helical peptide therapeutics.[7] On the other hand, β-sheets are not as well understood, and few β-sheet peptide inhibitors have been designed.

One reason fewer β-sheet peptides exist is that small water-soluble β-sheet peptides are fundamentally more difficult to design and not as well understood than α-helical peptides. The residues that promote the β-sheet secondary structure are hydrophobic or moderately polar (Figure 2.1).[8–10] Since there are no charged residues that promote sheet structure, there is often trade-off between the stability and solubility of β-sheet peptides. A hydrophilic residue that could promote β-sheet structure would in theory increase our ability to easily design β-sheet promoting peptides. For further detail on β-sheet structure, see Chapter 1.

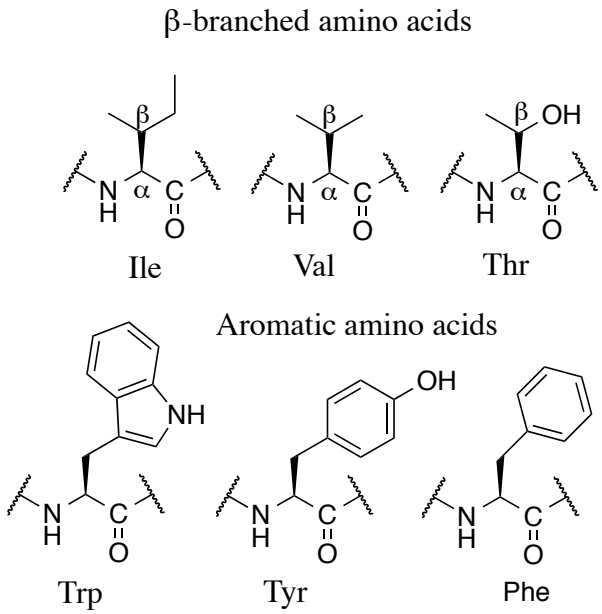
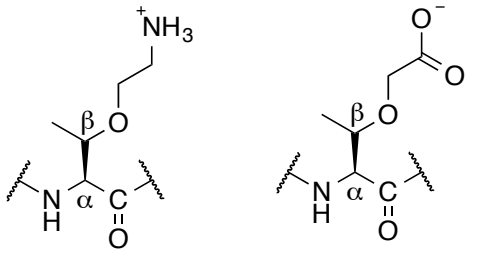


Figure 2.1 Natural amino acid residues that promote β-sheet secondary structure

2.2 Design and synthesis of cationic and anionic threonine

2.2.1 Design

Natural amino acids that promote β-sheet secondary structure are either aromatic or β-branched. β-branched amino acids have two substituents at the β position (Figure 2.1), and all natural β-branched residues promote β-sheet formation.[8,10] It has been proposed that β-branching in any amino acid, natural or unnatural, is a structural motif that should promote β-sheet secondary structure.[11–13] Therefore, we designed charged and β-branched amino acids based on threonine (Figure 2.2). In theory, the charge should increase the solubility of the peptide, and the β-branching should enable the amino acid residue to promote sheet structure.



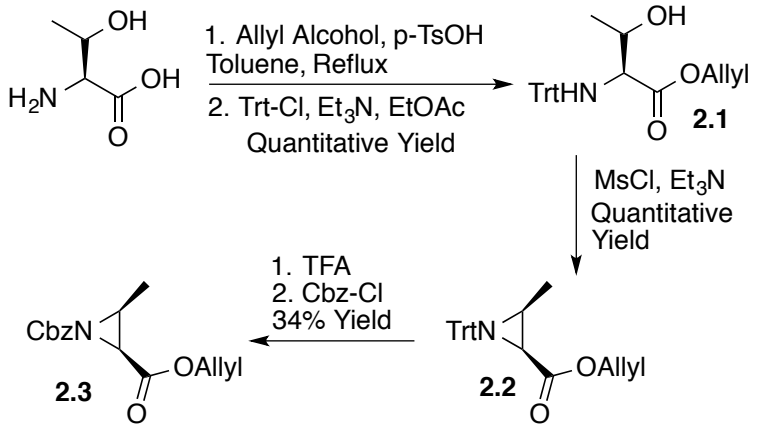
Cationic Threonine Anionic Threonine

Figure 2.2. Structure of cationic and anionic threonine, residues designed to promote β-sheet formation and aqueous solubility.

2.2.2 Synthesis

A protected form of cationic threonine was synthesized through an aziridine ring opening via methods derived from Vederas *et al*.[14] The aziridine precursor **2.3** was synthesized from threonine (Scheme 2.1). Aziridine **2.2** was synthesized as described by Vederas *et al*.[14] Briefly, threonine was protected with an allyl ester and a trityl protecting group to form threonine derivative **2.1**. Aziridine **2.2** was formed via a ring closing reaction. Aziridine **2.3** was synthesized from aziridine **2.2** by replacing the trityl protecting group with a carbamate protecting group (Scheme 2.1).

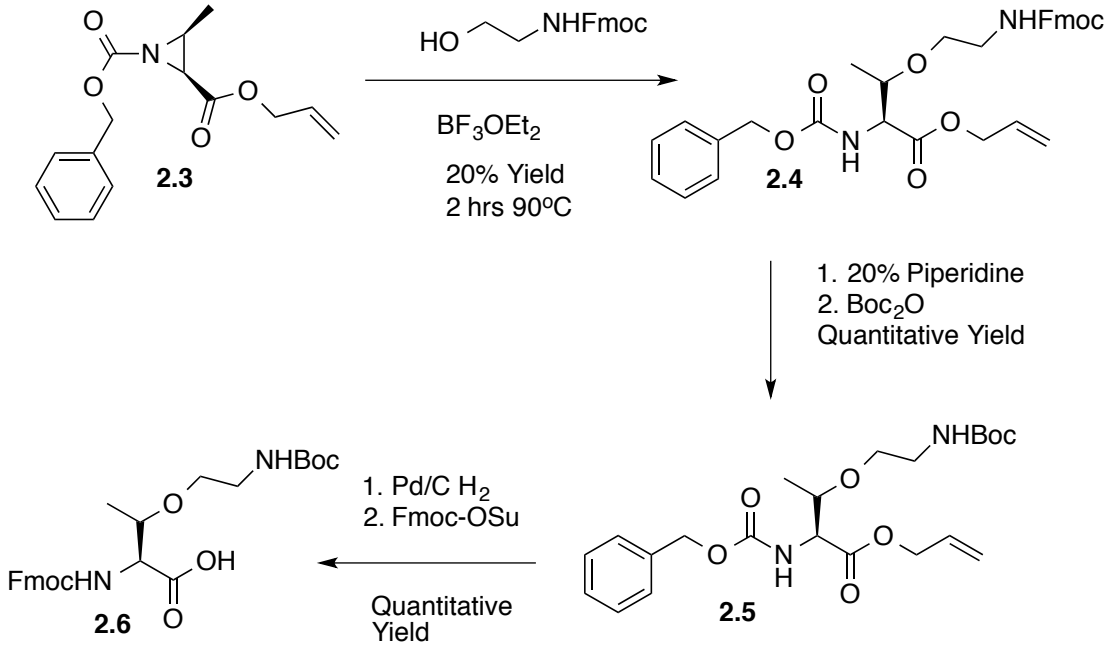
Scheme 2.1. Aziridine precursor synthesis from threonine.



Aziridine **2.3** was ring opened under Lewis acidic conditions to form a protected version of cationic threonine **2.4** (Scheme 2.2). After protecting group manipulation,

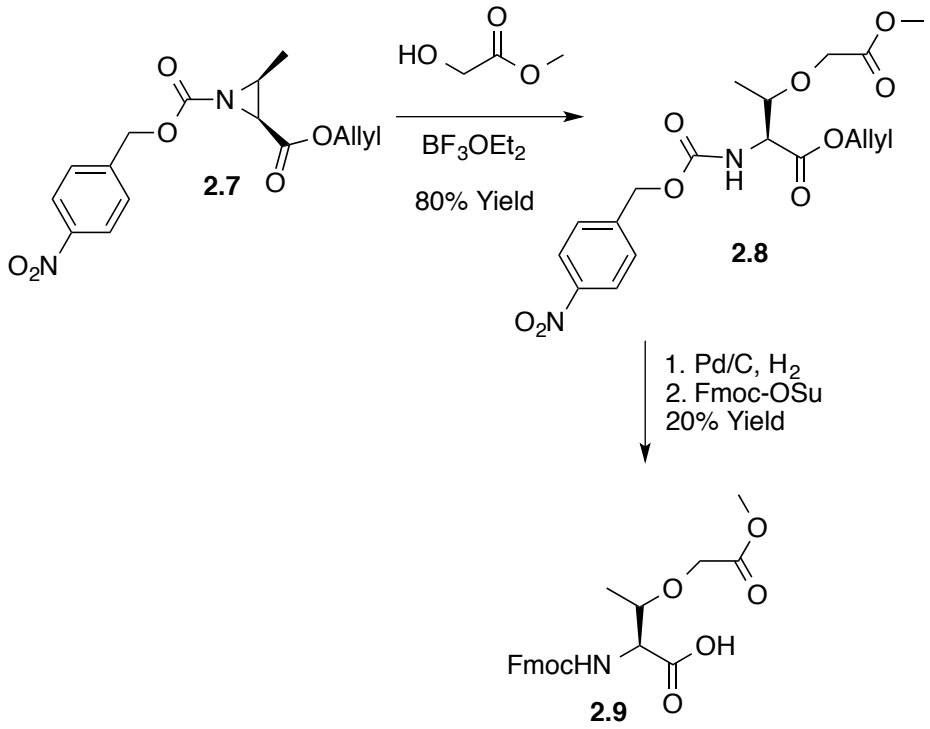
cationic threonine with the necessary protecting groups for Fmoc solid phase peptide synthesis (SPPS) was formed (compound **2.6**).

Scheme 2.2. Cationic threonine synthesis



A protected version of anionic threonine was synthesized through aziridine precursor **2.7** (Scheme 2.3). The aziridine was synthesized as previously described.[14] The aziridine ring was opened under Lewis acidic conditions by an alcohol nucleophile, and protecting group manipulation yielded Fmoc-protected amino acid **2.9**. The methyl ester is not a standard side chain protecting group used in Fmoc SPPS. However, previous work has shown methyl groups can be used in peptide synthesis and removed afterward by careful hydrolysis under basic conditions with LiOH.[15] However, in this case, SPPS did not yield any peptide of the expected mass. Possibly, the methyl protecting group was not bulky enough to prevent side-reactions of the ester group during SPPS synthesis.[15]

Scheme 2.3. Anionic threonine synthesis



2.3 Evaluation of cationic threonine (TO^+)

In order to evaluate the ability of cationic threonine (TO^+) to promote β-sheet secondary structure, cationic threonine was incorporated into a small water-soluble β-sheet hairpin structure. We chose a peptide, hairpin 2.10a, developed by Syud *et al*. that had previously been thermodynamically characterized as a parent system (Figure 2.3).[16] Hairpin 2.10a is a two-stranded β sheet connected by a two-residue DPro-Gly turn. The DPro-Gly turn has been shown to promote β-sheet secondary structure as the DPro can adopt the phi and psi angles necessary for a two-residue turn that is compatible with the β-hairpin structure.[17,18] Cationic threonine was substituted for lysine to generate peptide

2.11a, which allows us to compare positively charged cationic threonine to a naturally occurring positively charged residue.

Peptide 2.11a was characterized by a set of standard 2D NMR experiments. A COSY, TOCSY, and ROESY allowed the assignment of the peptide's ^{1}H resonances. The COSY experiment has cross-peaks for proton resonances that are on adjacent atoms (and coupled to one-another). The TOCSY experiment has through-bond cross-peaks for ^{1}H in the same spin system. For example, a ^{1}H in a Val residue will, in theory, have TOCSY cross-peaks with all other ^{1}H resonances in the same Val residue. Together with the ROESY, which has cross-peaks for ^{1}H resonances that are close in space, these experiments allow the systematic assignment of all non-exchangeable protons in the peptide.[19]

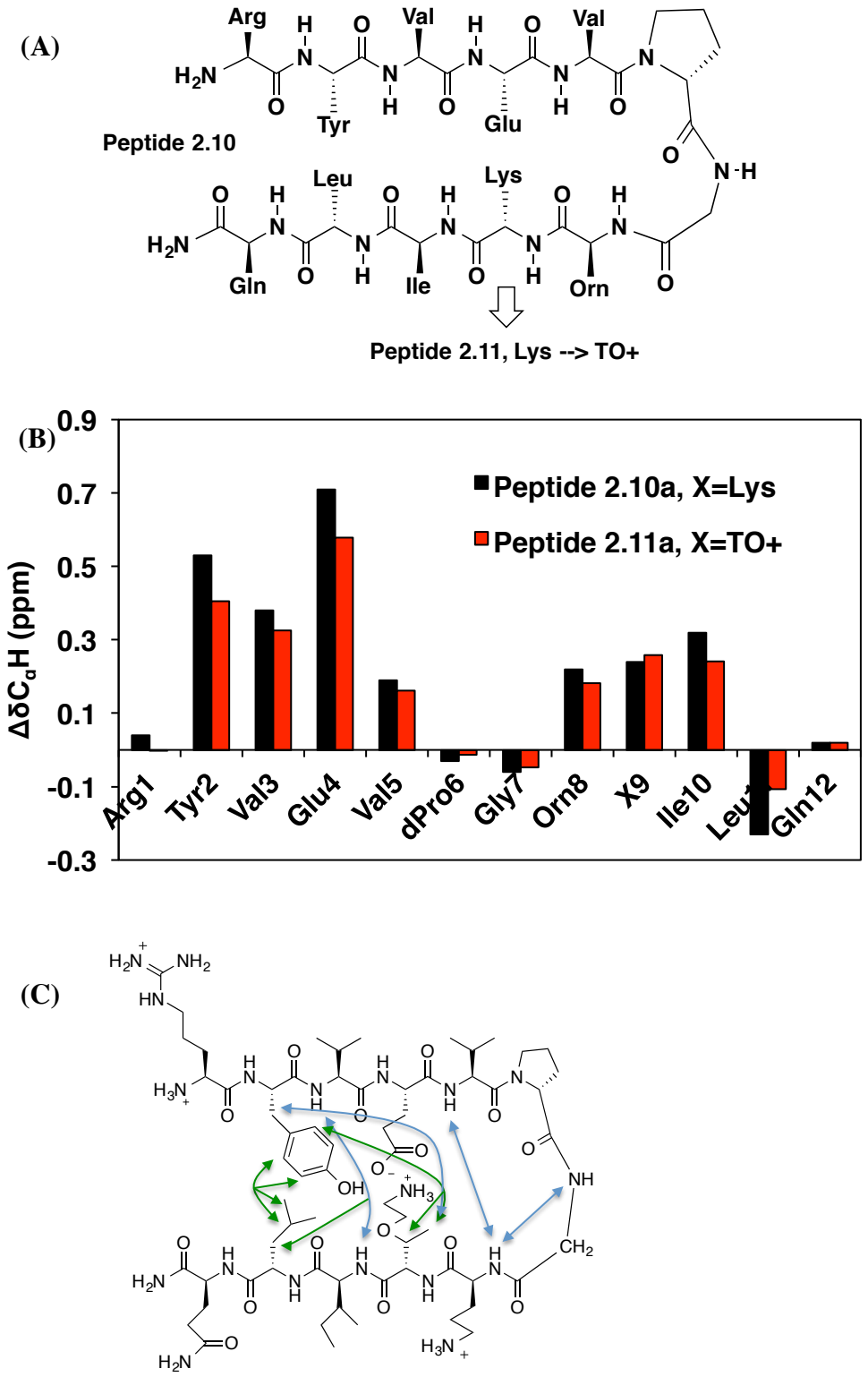


Figure 2.3 (A) Sequence of hairpin used to evaluate the ability of cationic threonine to promote sheet structure.[16] (B) Chemical shift deviations of peptide 2.10a and 2.11a. $\Delta\delta C_\alpha H$ at or adjacent to the substitution cannot be directly compared. (C) Cross-strand ROEs observed for peptide 2.11a. Medium ROEs (3.0-3.5 Å) are blue and weak ROEs (3.5-4.5 Å) are green.

ROESY allows the analysis of through-space H-H contacts. (The ROESY experiment is similar to the better-known NOESY experiment.) Protons that are close in space have ROE cross-peaks. The integration of the ROEs is related to the distance between the protons and falls off by a power of six with increasing distance. Sequential ROEs are cross-peaks that arise between i and i+1 residues and are largely due to the proximity of the residues in the primary sequence. Non-sequential ROEs are cross peaks between residues that are two or more amino acids apart in the primary sequence. The non-sequential ROEs arise due to the conformation of a peptide. In the context of a β-hairpin structure, we expect to see ROEs from residues on adjacent strands. We define the strength of these ROEs as strong (2.5-3.0 Å), medium (3.0-3.5 Å), or weak (3.5-4.5 Å) based on the theoretical distance between the protons, which is calculated from the integration of the ROEs. Note that the integrations of the ROEs are effected both by the distance between the protons in the folded structure and by the stability of the folded structure.[19,20]

The ROEs for peptide 2.11a are summarized in Figure 2.3c. Medium and weak cross-strand ROEs are present, which strongly supports the hypothesis that the peptide adopts a β-hairpin conformation. All ROEs are consistent with β-sheet secondary structure.

The chemical shifts of the α-protons are indicative of secondary structure. The α-protons are the protons attached to the carbons in the backbone of the peptide (see Figure 2.4). Chemical shifts of the random coil (unstructured) peptides have been compared to the chemical shifts of structured peptides and the difference between the two is the chemical shift deviation. The deviations can be calculated for each residue in a peptide.

Downfield chemical shift deviations indicate β-sheet secondary structure, and upfield deviations indicate α-helix secondary structure.[21] A diastereomer of peptide 2.11a with an LPro substituted for the DPro residue, peptide 2.11b, was synthesized and characterized (Figure 2.4). When LPro is substituted for DPro, the peptide can no longer form the hairpin structure as LPro cannot adopt the necessary Phi/Psi angles for a two-residue turn.[18] The chemical shifts of LPro versions of peptides 2.10a and 2.11a can therefore be used to represent the random coil (unfolded) population and determine the chemical shift deviations ($\Delta\delta C_\alpha H$) of the peptides of interest (Figure 2.3b).[16,18] See Chapter 1 for further detail on LPro versus DPro turns. The $\Delta\delta C_\alpha H$ at the hydrogen-bonding positions in the core of the peptide are thought to be the most reliable reporters of β-sheet secondary structure.[22]

Since the folding and unfolding of the peptides usually happens faster than the NMR time scale, the chemical shifts of the α-protons will be the weighted average of the chemical shift in all conformations.[18,21] Therefore, the magnitude of the chemical shift deviation can be used to determine the relative stability of the folded conformation.[18] Note that chemical shift deviations cannot be directly compared at or adjacent to a substitution. For example, the chemical shift deviations at residue 9 in peptides 2.10a and 2.11a (Lys 9→TO^+) cannot be directly compared. The dynamic range is the difference in chemical shift between the folded hairpin population and the unstructured population. At Lys 9 in peptide 2.10a, the dynamic range is 0.52 ppm. At TO^+ 9 in peptide 2.11a, the dynamic range is 0.59 ppm. If 50% of the peptides adopt the folded conformation, the chemical shift deviation will be half its dynamic range, 0.26 ppm for peptide 2.10a and 0.30 ppm for peptide 2.11a. Therefore, if the folded conformations of the two peptides

were equally stable, the chemical shift deviation at residue 9 of peptide 2.11a would appear to be greater than the chemical shift deviation at residue 9 of peptide 2.10a. It should also be noted that we can account for changes in dynamic range in quantitative thermodynamic calculations (see below).

The graph of chemical shift deviation for peptide 2.10a and 2.11a by residue number is consistent with the expected structure (Figure 2.3b). The chemical shift deviations are above +0.1 ppm for residues that are thought to be part of the β-sheet structure. The residues at the end of the β-sheet structure may be 'frayed' as they have smaller chemical shift deviations. Surprisingly, The chemical shift deviations of peptide 2.10a are generally greater than the chemical shift deviations of TO^+-containing peptide 2.11a. The data suggest that the β-hairpin conformation of peptide 2.10a is more stable than the β-hairpin conformation of 2.11a and that TO^+ *does not* stabilize β-sheet structure relative to Lys.

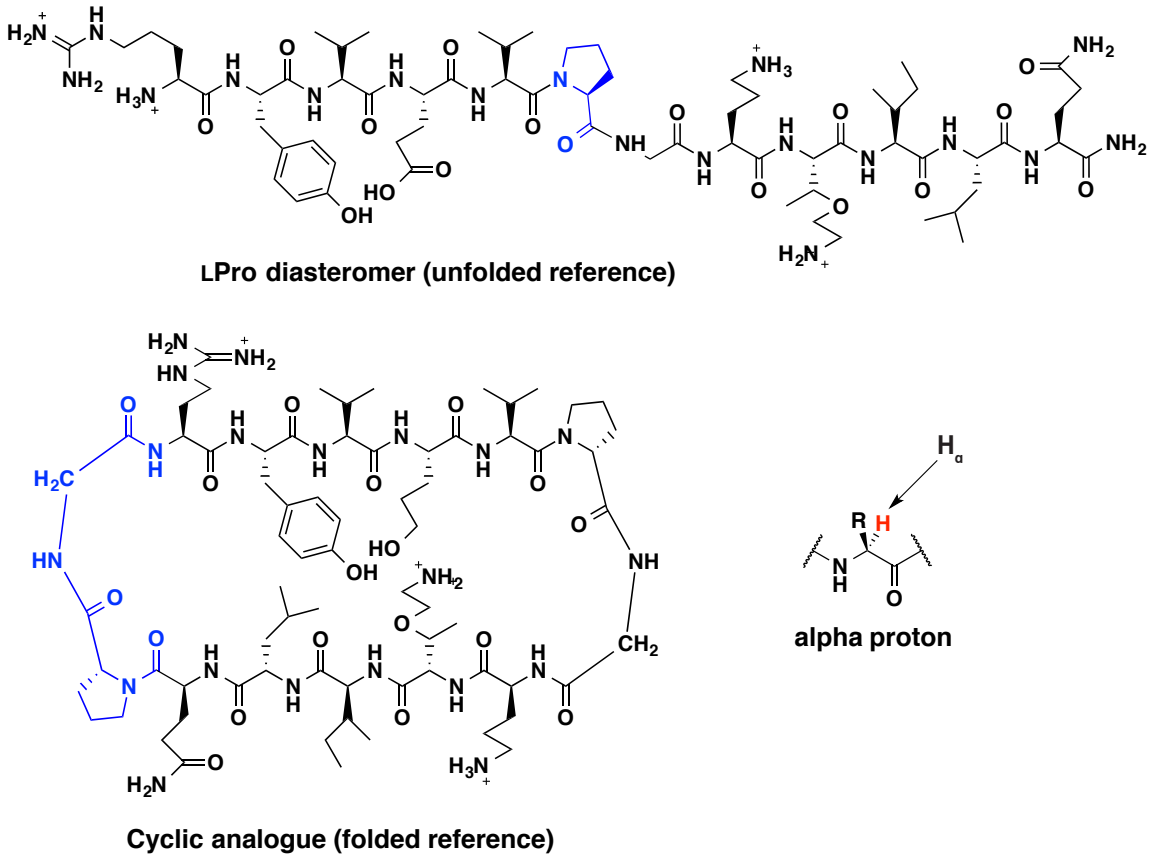


Figure 2.4. Control peptides for quantitative thermodynamic analysis and example of an α-proton in a residue.

Quantitative thermodynamic analysis of the β-hairpin folding of peptides was undertaken. Cyclic versions of peptides 2.10a and 2.11a were used as fully folded controls and were designed by connecting the C- and N-termini with a two residue DPro-Gly turn to give peptide 2.10c and 2.11c, respectively (Figure 2.4).[23] The DPro-Gly cyclic constraint enforces β-hairpin structure and forces the peptide to adopt the β-hairpin conformation. Therefore, α-proton chemical shifts of the peptide approximate the chemical shifts of the folded population of a linear peptide of a related sequence.[23]

2D NMR analysis of cyclic peptide 2.11c was performed and the chemical shifts of the α-protons were determined. NMR data from Syud *et al*. for peptide 2.10c was

used.[16] The values of the α-proton chemical shifts of the folded hairpin populations and unstructured populations were approximated from the chemical shifts from the cyclic and LPro reference peptides. These chemical shifts were then used to calculated the percent of the linear DPro hairpin peptides in the folded conformation with the equation

$$\beta\text{-sheet population}=\frac{\delta_{Obs}-\delta_U}{\delta_F-\delta_U} \quad \text{Equation 2.1}$$

where δ_{Obs} is the α-proton chemical shift in the peptide of interest, δ_U is the chemical shift in the LPro unfolded reference, and δ_F is the chemical shift in the cyclic folded reference. The folded population can be used to calculate the folding equilibrium constant (K_F) with the equation

$$K_F=\frac{beta\ sheet\ population}{1-beta\ sheet\ population} \quad \text{Equation 2.2}$$

ΔG_F is calculated with the equation

$$\Delta G_F = -RT\ln(K_F) \quad \text{Equation 2.3}$$

The ΔG_F value is calculated at four indicator positions (Val 3, Val 5, Lys 8, and Ile 10) and then averaged. These positions were chosen as they are all hydrogen-bonding residues in the core of the β-sheet interaction. Hydrogen bonding positions (as defined in Chapter 1) are used as they are thought to most accurately reflect the stability of the peptide. (The folded populations calculated from the hydrogen bonding positions are more self-consistent than the values calculated at non-hydrogen bonding positions.)[22] Note that when directly comparing chemical shift deviations, it was not possible to use positions at or adjacent to the substitution. However, in the thermodynamic calculation, folded and unfolded controls with the same sequence as the peptide of interest allow us to account for changes in the dynamic range. Therefore, these positions can be used in the

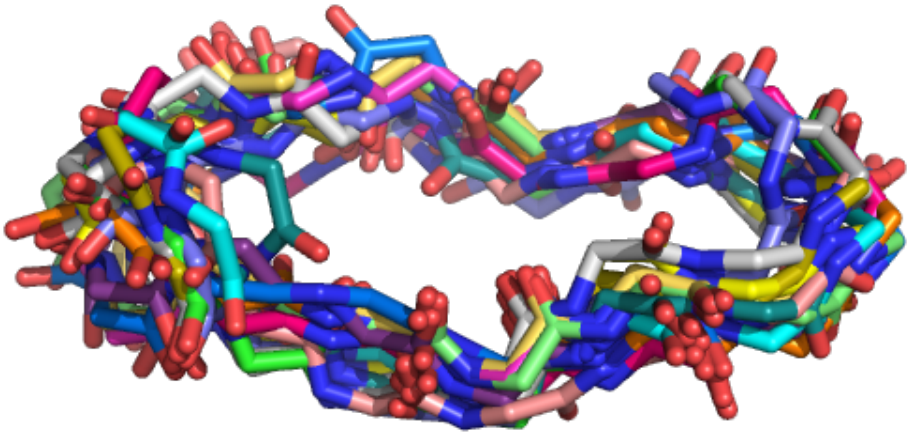
thermodynamic calculation. The $\Delta\Delta G_F$ between peptide 2.10a and 2.11a was calculated by measuring the $\Delta\Delta G_F$ at each position (Val 3, Val 5, Lys 8, and Ile 10) with the equation

$$\Delta\Delta G_{F(\text{Residue X})} = \Delta G_{F(2.11a,\ \text{Residue X})} - \Delta G_{F(2.10a,\ \text{Residue X})} \qquad \text{Equation 2.4}$$

and averaging the four values. The standard deviation of the values was calculated and the student's T test at the 95% confidence interval was used to determine the error. Peptide 2.10a is +0.1± 0.2 kcal/mol less stable than peptide 2.11a. In other words, the stability of the two peptides is not distinguishable.

In order to determine if the NOEs observed in the cyclic "fully-folded" peptide 2.11c were consistent with the expected β-sheet structure, a molecular dynamics calculation was performed using the observed sequential and non-sequential NOEs as constraints in CNS.[24] Only NOEs that could be definitively assigned were used in the calculation and use of ambiguous NOEs might have lead to a more homogenous set of lowest energy structures. The lowest energy structures demonstrate that the expected β-sheet structure is consistent with all the observed non-sequential NOEs as most of these structures adopt a β-hairpin conformation.

(A)



(B)

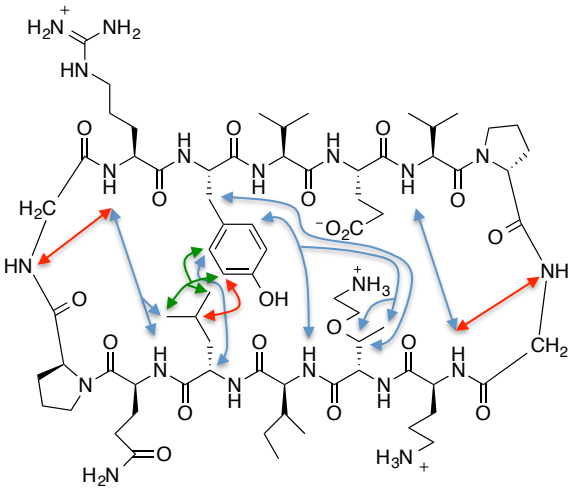


Figure 2.4. (A) Backbone atoms of the 16 lowest energy structures of the cyclic hairpin 2.11c (B) Observed cross-strand NOEs in cyclic hairpin 2.11c. Strong NOEs are red, medium NOEs are blue, and weak NOEs are green.

As the stability of peptide 2.10a and 2.11a (Lys 9→TO^+) were not distinguishable, Lys→TO^+ substitutions were tested in other systems to see if it was possible to differentiate between the two residues. In addition, there can be some context dependence in β-sheet peptides,[10] so it was important test if the unexpected effect of the Lys→TO^+ substitution was a general effect or a special case. In order to do so, the Lys→TO^+ substitution was made in a hairpin derived from Espinosa *et al* (Figure 2.5a).[25] Again, the hydrogen bonding positions most accurately reflect the secondary structure as the non-hydrogen bonding positions can be influenced by the presence of aromatic amino acids,

which accounts for the apparent negative chemical shift deviation of positions 9 and 11 (Figure 2.6).[25] The parent peptide 2.12 was compared to the Lys8→TO^+ peptide 2.13 (Figure 2.6a).

The NOEs observed in peptide 2.13 support the existence of β-sheet secondary structure (Figure 2.5). Numerous strong, medium, and weak cross-strand NOEs are observed. All NOEs are consistent with β-hairpin structure. Strong and medium cross-strand NOEs are observed between Trp 2 and Val 11. Strong and medium NOEs are observed between Trp 2 and Phe 9. Medium and weak NOEs are observed between Tyr 4 and Phe 9. A medium NOE is observed between TO^+ 8 and Val 6. Note that this hairpin contains three aromatic residues. As aromatic resonances fall into a well-dispersed region of the peptide's NMR spectrum, numerous NOEs can be unambiguously assigned.

(A) **(B)**

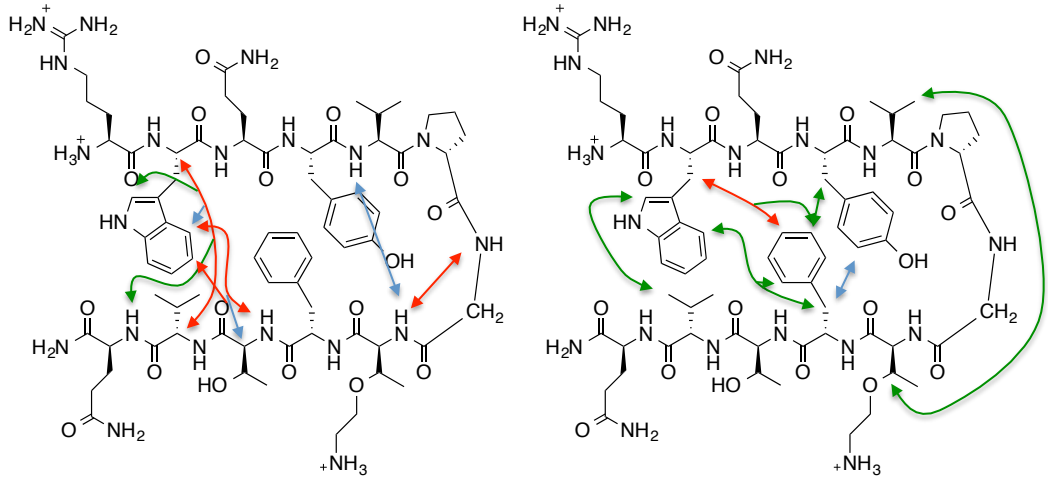
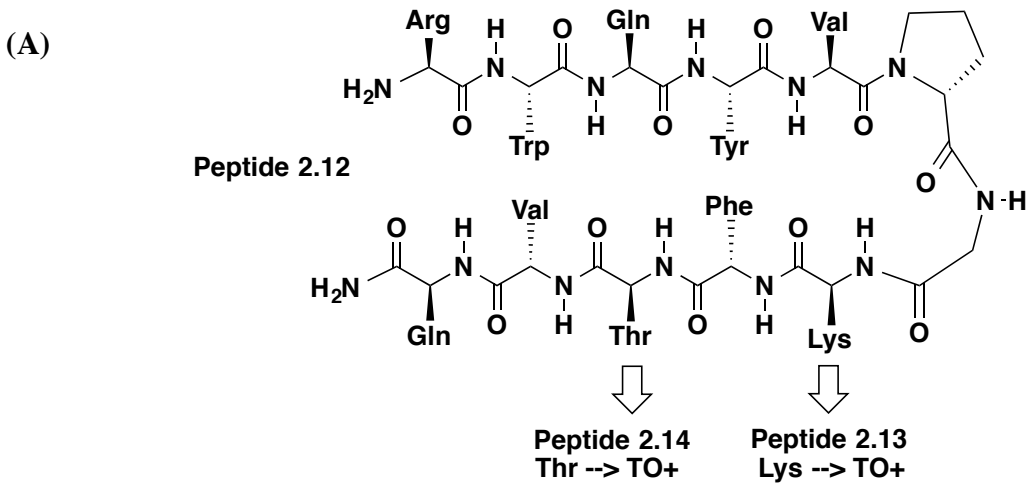


Figure 2.5. Cross-strand NOEs observed for peptide 2.13. Strong NOEs are red, medium NOEs are blue, and weak NOEs are green. (A) Backbone-backbone and backbone-side chain NOEs. (B) Side chain-side chain NOEs



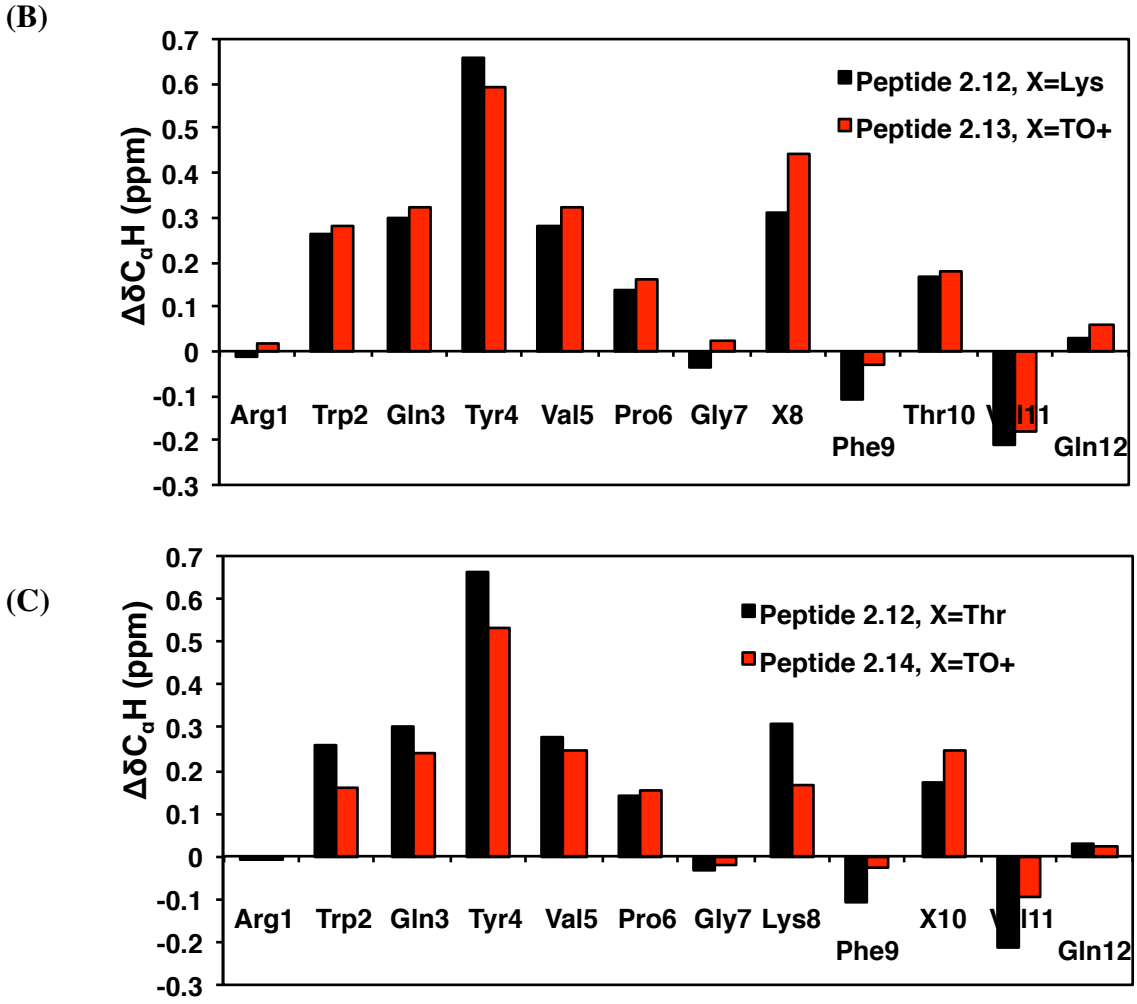


Figure 2.6 (A) Peptides derived from Espinosa *et al*.[25] and (B), (C) their chemical shift deviations. Chemical shift deviations at or adjacent to the substitution cannot be directly compared.

At hydrogen bonding positions, the chemical shifts deviations of the α-protons in peptide 2.13 (Lys 8→TO^+) appear to be slightly greater than the chemical shift deviations of the original Lys containing peptide. However, the chemical shift deviations of 2.12 and 2.13 are probably within error of each other (Figure 2.6b). Also, the Lys→TO^+ substitution was made next to the two-residue turn rather in the core of the peptide, which may not be an accurate test of β-sheet propensity in the core of a β-sheet.

A Thr→TO^+ substitution was made in hairpin 2.12. Threonine is a β-branched residue that promotes sheet structure and is polar in contrast to other β-sheet promoting residues, which are hydrophobic. Therefore, if charged TO^+ can promote sheet structure as well as Thr, it functions as designed. Peptide 2.14 was analyzed with 2D NMR and the observed ROEs were consistent with the expected β-sheet structure (Figure 2.7). However, the chemical shift deviations of Peptide 2.14 were generally smaller than the chemical shift deviations of peptide 2.12 (Figure 2.6c). Therefore, the Thr→TO^+ substitution appears to destabilize the peptide. It should be noted that the substitution places the positive charge in TO^+10 close in space to the positive charge in Lys8 when the peptide folds into the hairpin structure, which will destabilize the hairpin. In order to accurately compare TO^+ to Thr, one must choose a position where charge-charge attraction or repulsion to other side-chain groups is not likely to be a major factor. As small water-soluble hairpins must be highly charged in order to be water-soluble, it is difficult to find a system where one can substitute a neutral Thr for a positively charged TO^+ without packing it against a pre-existing positively charged residue.

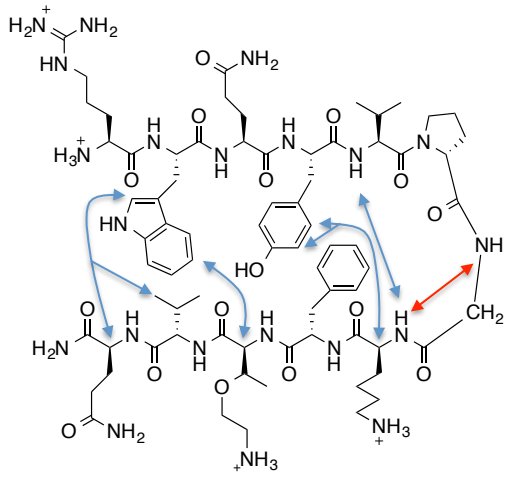


Figure 2.7. Backbone-backbone and backbone-side chain NOEs observed for peptide 2.14. Strong NOEs are red and medium NOEs are blue.

In order to obtain a clear answer about the ability TO^+ to stabilize β-sheet structure relative to the existing natural positively charged residues, a Lys→TO^+ substitution was made in another hairpin sequence (Figure 2.8).[18] The chemical shift deviations of peptide 2.15 and 2.16 were consistent with those of a β-sheet structure. The chemical shift deviations of parent hairpin 2.15 are within error of the chemical shift deviations of the TO^+-containing hairpin 2.16 (Figure 2.8). Note that the chemical shift deviations at or adjacent to the substitution cannot be included in the analysis for two reasons. First, both peptide 2.15 and 2.16 were referenced to the LPro diastereomer (unfolded control) with the sequence of peptide 2.15. Therefore the chemical shift deviations at or adjacent to the substitution may not be accurate for peptide 2.16. Secondly, the substitution may change the dynamic range of positions at or adjacent to the substitution. The chemical shift deviations were analyzed at all other positions in the core of the hairpin and are within error of each other. Therefore, it is not possible to distinguish between the stability of the hairpin conformation of peptide 2.15 and 2.16. TO^+ does not appear to stabilize β-sheet structure relative to lysine.

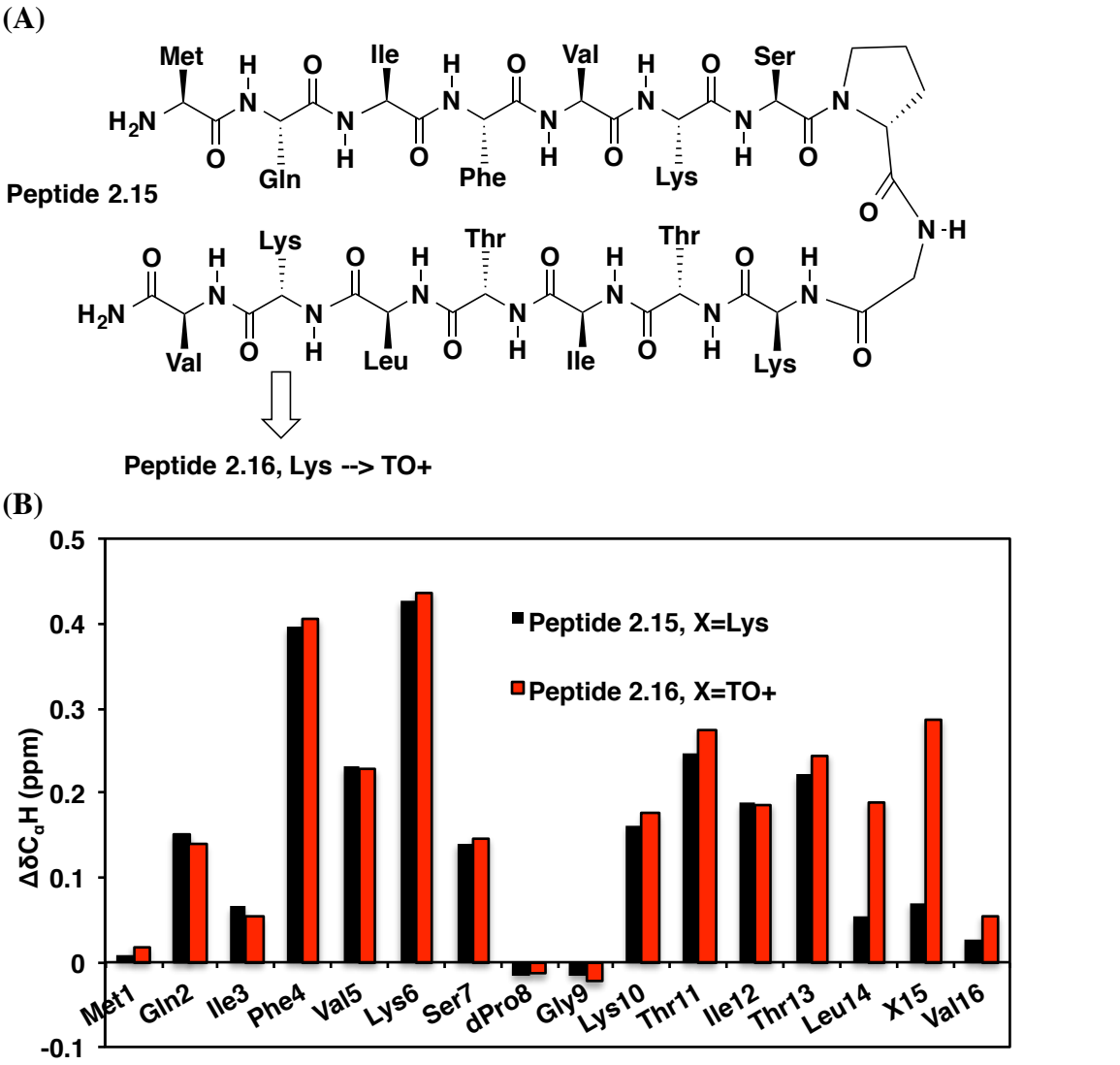


Figure 2.8. (A) Peptide structures (B) Chemical shift deviations for peptides derived from Haque *et al*.[18] Peptides are referenced to the LPro diastereomer of parent peptide 2.15. Substitutions at or adjacent to the substitution (positions 14-16) cannot be directly compared.

2.4 Conclusions

Cationic threonine does not appear to significantly stabilize β-sheet structure relative to Lys. This result is highly unexpected as β-branching is a feature of amino

acids thought to result in high β-sheet propensity.[8,11,26] The possible issues with the design of cationic threonine will be discussed in more detail in chapter 3.

2.5 Methods

2.5.1. Synthesis

General. Chemicals were purchased from Sigma- Aldrich, Chemimpex, or Fluka Analytical unless otherwise noted. NMR spectra were taken on a Bruker AC+300a (300 mHz), Bruker Avance-400 (400 mHz), Bruker Avance-500 with a DCH cryoprobe (500 mHz), Varian INOVA 600 (600 mHz), or Varian MercuryPlus 300 (300 mHz). Optical rotations were taken on a Rudolph Autopol III automatic polarimeter with a 1 mL cell and a 5 cm path length. NMR spectra can be found in the appendix.

2-Amino-*N*-(9-fluorenylmethoxycarbonyl)ethanol. 2-Aminoethanol (2.0 g, 32.7 mmol) was dissolved in 2:1 acetone: sat. aq. sodium bicarbonate (300 ml). *N*-(9-Fluorenylmethoxycarbonyloxy)succinimide (11.04 g, 32.7 mmol) was added. The reaction mixture was stirred at room temperature for 4 hours. The acetone was removed under reduced pressure. Water (200 ml) and ethyl acetate (200 ml) were added. The mixture was acidified with 1 M HCl. The layers were separated, and the organic layer was washed with 1 M HCl (3x200ml), saturated aqueous $NaHCO_3$ (1x200 ml), and brine (1x200 ml). The organic layer was dried with $MgSO_4$ and filtered. The solvent was removed under reduced pressure to yield a white solid in a quantitative yield. ESI-MS: Calc. $[C_{17}H_{18}NO_3]^+$=284.1, Obsd. $[C_{17}H_{18}NO_3]^+$=284.1; Melting Point: 144.5-144.9^oC; ^1H NMR ($CDCl_3$, 500 MHz): δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 7.5 Hz,

2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 2H), 5.15 (s, 1H), 4.43 (d, *J* = 6.7 Hz, 2H), 4.22 (t, *J* = 6.5 Hz, 1H), 3.77 – 3.67 (m, 2H), 3.40 – 3.30 (m, 2H), 2.06 (t, *J* = 4.8 Hz, 1H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 157.14, 143.86, 141.33, 127.70, 127.05, 125.00, 119.98, 66.76, 62.36, 47.24, 43.45.

(*2S, 3S*)-Allyl-3-methyl-1-tritylaziridine-2-carboxylate (2.2). Synthesized as previously described by Vederas *et al*. (Scheme 2.1).[14]

(2*S*, 3*S*)-Allyl-1-benzyl-3-methylcarbonylaziridine-2-carboxylate (2.3). Aziridine **2.2** was dissolved in methanol (32 ml) and dichloromethane (32 ml). Trifluoroacetic acid (32 ml) was added dropwise over 10 minutes. The reaction mixture was stirred for 10 minutes, and the solvent was removed under reduced pressure. The residue was dissolved in ether (150 ml) and H_2O (150 ml). The layers were separated, and the ether layer was extracted with H_2O (3 x 30 ml). The combined aqueous layers were cooled to 0°C and made alkaline with solid sodium bicarbonate. Benzyl chloroformate (3.1 g) and ethyl acetate (300 ml) were added, and the reaction mixture was stirred vigorously for 24 hours. The layers were separated. The organic layer was dried with $MgSO_4$, and the solvent was removed under reduced pressure. The crude product was purified by chromatography on silica gel (EtOAc:hexanes 1:4). The product was a clear liquid (3.2 g, 11.6 mmol, 31% yield). ESI-MS: Calc. $[C_{15}H_{18}NO_4]^+$=276.1231, Obsd. $[C_{15}H_{18}NO_4]^+$=276.1225; $[\alpha]_D^{25}$= +73 (*c* 1.15, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 500 mHz): δ 7.40-7.32 (m, 5H), 5.93 (ddt, *J* = 16.4, 10.5, 5.9 Hz, 1H), 5.35 (dd, *J*=17.2, 1.2 Hz, 1H), 5.27 (dd, *J*=10.4, 1.0 Hz) 5.13 (AB, *J*=13.8 Hz, 2H), 4.68 (d, *J* = 5.8 Hz, 2H), 3.20 (d, *J* = 6.7 Hz, 1H), 2.83 (dq, *J* = 6.5, 5.7 Hz, 1H), 1.36 (d, *J* = 5.6 Hz, 3H); ^{13}C NMR ($CDCl_3$,

125 mHz): δ 166.87, 161.61, 135.30, 131.47, 128.57, 128.49, 128.38, 119.15, 68.59, 66.13, 39.86, 38.99, 12.86.

(2*S*, 3*R*) Allyl-2-(benzylcarbonylamino)-3-(9-fluorenylmethoxycarbonyl aminoethyloxy)butane carboxylate (2.4). Aziridine **2.3** (5.53 g, 20.1 mmol) and 2-amino-*N*-(9-fluorenylmethoxycarbonyl)ethanol (38.5 g, 136 mmol) were suspended in 1000 ml of toluene. Boron trifluoride diethyl etherate (1.24 ml, 10.0 mmol) was added dropwise. The reaction mixture was refluxed for three hours. The solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (20% EtOAc in hexanes). The product was a clear oil (2.2 g, 4.0 mmol, 20% yield). ESI-MS: Calculated $[C_{32}H_{35}N_2O_7]^+$= 581.2259, Measured $[C_{32}H_{35}N_2O_7]^+$= 581.2254; ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.75 (d, *J* = 7.5 Hz, 2H), 7.60 (d, *J* = 7.1 Hz, 2H), 7.41-7.28 (m, 9H), 5.86 (ddt, *J* = 16.8, 10.6, 5.7 Hz, 1H), 5.50 (d, *J* = 9.3 Hz, 1H), 5.29 (d, *J*=17.4 Hz, 1H) 5.22 (d, *J* =10.3 Hz, 1H), 5.14-5.10 (m, 3H), 4.67-4.58 (m, 2H), 4.43-4.32 (m, 3H), 4.20 (t, *J* = 6.9 Hz, 1H), 4.11 – 3.99 (m, 1H), 3.63-3.60 (m, 1H), 3.33-3.28 (m, 3H), 1.22 (d, *J* = 6.1 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.64, 156.65, 143.94, 141.30, 136.17, 131.42, 128.57 , 128.25, 128.12, 127.69, 127.06, 125.12, 119.97, 119.15, 75.47, 68.26, 67.23, 66.76, 66.13, 58.60, 47.23, 49.97, 16.11.

(2*S*, 3*R*) Allyl-2-(benzylcarbonylamino)-3-(*tert*-butylmethoxycarbonyl aminoethyloxy)butane carboxylate (2.5). Compound **2.4** (1.6 g, 2.9 mmol) was dissolved in 20% piperidine in DMF (100 ml), and the solution was stirred for 15 minutes. Solvent was removed under reduced pressure by azeotroping with toluene. The residue was dissolved in dichloromethane (100 ml). Triethylamine (0.4 ml, 2.9 mmol)

and di-*tert*-butyl dicarbonate (4.1 g, 18.9 mmol) were added. The reaction mixture was stirred for 3 hours. The solvent was removed under reduced pressure, and the residue was dissolved in ethyl acetate (100 ml). This solution was washed with 1 M HCl (3x100 ml), saturated aqueous sodium bicarbonate (1x100 ml), and brine (1x100 ml). The organic layer was dried with magnesium sulfate and filtered. The solvent was removed under reduced pressure. The crude product was purified by silica column chromatography (20% EtOAc in hexanes) to yield a clear oil (1.3 g, 2.9 mmol, quantitative yield). ESI-MS: Calc. $[C_{22}H_{33}N_2O_7]^+$=437.2283, Obsd. $[C_{22}H_{33}N_2O_7]^+$= 437.2294; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38-7.32 (m, 5H), 5.91 (ddt, *J* = 17.0, 10.6, 5.7 Hz, 2H), 5.47 (d, *J* = 9.6 Hz, 1H), 5.35 (d, *J*=18.3 Hz, 1H), 5.27 (d, *J*=10.4 Hz, 1H), 5.14 (s, 2H), 4.77 (s, 2H), 4.67 (d, *J* = 5.8 Hz, 4H), 4.39 (dd, *J* = 11.9, 2.2 Hz, 1H), 4.05 (qd, *J* = 6.2, 2.2 Hz, 1H), 3.65-3.55 (m, 1H), 3.35 – 3.14 (m, 3H), 1.43 (s, 9H), 1.21 (d, *J* = 6.3 Hz, 6H). ^{13}C NMR (125 MHz, $CDCl_3$): δ170.56, 156.65, 155.85, 136.18, 131.50, 128.55, 128.21, 128.10, 119.22, 79.35, 75.42, 68.44, 67.18, 66.15, 58.65, 40.44, 28.39, 16.08.

(2*S*, 3*R*) 2-(9-fluorenylmethoxycarbonylamino)-3-(2-*tert*-butyloxycarbonylaminoethyloxy)butane carboxylic acid (2.6). Compound **2.4** (0.42 g, 1.0 mmol) was dissolved in methanol (100 ml). Ammonium formate (0.60 g) was added, and the solution was stirred under N_2 for 20 minutes. Palladium on carbon (0.36 g, 10% loading by weight on a wet support) was added, and the mixture was stirred under N_2 for 24 hours. The mixture was filtered through celite, and the solvent was removed from the filtrate under reduced pressure. The residue was dissolved in 2:1 acetone:saturated aqueous sodium bicarbonate (50 ml). *N*-(9-fluorenylmethoxycarbonyloxy) succinimide

was added, and the mixture was stirred for 5 hours. The acetone was removed under reduced pressure, and ethyl acetate was added (50 ml). The solution was acidified with 1 M HCl and extracted with ethyl acetate (3x50 ml). The combined organic layers were dried with magnesium sulfate and filtered. The solvent was removed under reduced pressure to give a white solid (0.5 g, 1.0 mmol, quantitative yield). EMM-MS: Calc. $[C_{26}H_{32}N_2O_7NH_4]^+$=502.2548, Obsd. $[C_{26}H_{32}N_2O_7NH_4]^+$= 502.2549; Melting Point: 108.5-111.0°C; ^{1}H NMR (500 MHz, CD_3OD): δ7.82 (d, *J* = 7.5 Hz, 2H), 7.76 – 7.68 (m, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.34 (t, *J* = 7.4 Hz, 2H), 4.41 (dd, *J* = 7.0, 2.2 Hz, 2H), 4.32 – 4.23 (m, 2H), 4.11 (qd, *J* = 6.2, 2.6 Hz, 1H), 3.61-3.57 (m, 1H), 3.41 (dt, *J* = 9.8, 5.3 Hz, 1H), 3.35-3.15 (m, 2H), 1.46 (s, 9H), 1.21 (d, *J* = 6.3 Hz, 3H); ^{13}C NMR (125 MHz, CD_3OD): δ 174.16, 159.22, 158.56, 145.38, 145.15, 142.60, 128.78, 128.19, 128.17, 126.31, 126.28, 120.92, 120.91, 80.13, 76.52, 69.45, 68.13, 60.10, 41.58, 28.78, 26.28, 16.75.

(2*S*, 3*S*)-Allyl-3-methyl-1-(4-nitrobenzyl)carbonylaziridine-2-carboxylate (2.7). Aziridine 2.7 was synthesized as previously described.[14]

(2*S*, 3*R*)-Allyl-3-(methoxycarbonylmethoxy)-2-(4-nitrobenzylcarbonylamino)-butanoate (2.8). Aziridine **2.7** (7.2 g, 26.1 mmol) was dissolved in toluene (260 ml). Methyl glycolate (7.0 g, 78 mmol) was added. Boron trifluoride diethyl etherate (1.5 ml, 13 mmol) was added dropwise. The reaction mixture was refluxed for 2 hours. The solvent was removed under reduced pressure. The crude oil was purified by column chromatography (1:1 ethyl acetate:hexanes), and pure product was obtained (7.6 g, 80% yield). ESI-MS: Calc. $[C_{18}H_{23}N_2O_9]^+$=411.1, Obsd. $[C_{18}H_{23}N_2O_9]^+$=411.1; ^{1}H NMR (300

MHz, $CDCl_3$): δ 8.22 (d, *J* = 8.7 Hz, 2H), 7.53 (d, *J* = 8.5 Hz, 2H), 6.08 – 5.83 (m, 2H), 5.42 – 5.15 (m, 4H), 4.70-4.68 (m, 2H), 4.32 (dd, *J* = 8.7, 3.1 Hz, 1H), 4.25 – 4.13 (m, 1H), 3.74 (s, 3H), 1.26 (d, *J* = 6.3 Hz, 3H).

(2*S*, 3*R*)-2-(9-fluorenylmethoxycarbonylamino)-3-(methoxycarbonylmethoxy)butyric acid (2.9). Ester **2.8** (2.0g, 4.9 mmol) was dissolved in methanol (100 ml), and Pd/C (0.2 g) was added. The mixture was stirred under H_2 for 24 hours. The mixture was filtered through celite to remove the Pd/C, and the solvent was removed under reduced pressure. The resulting oil was dissolved in methanol (100 ml). 9-Fluorenylmethyl *N*-succinimidyl carbonate (1.77g, 5.2 mmol) and diisopropylethylamine (1.8 ml, 10.3 mmol) were added. The mixture was stirred for 4 hours, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (80% ethyl acetate and 1% acetic acid in hexanes) to yield 0.4 g of product (0.96 mmol, 20% yield). ^{1}H NMR (300 MHz, DMSO): δ 7.90 (d, *J* = 7.6 Hz, 2H), 7.77 (d, *J* = 7.5 Hz, 2H), 7.50 – 7.18 (m, 4H), 4.29-4.22 (m, 3H), 4.16 (m, 2H), 4.10 – 3.91 (m, 2H), 3.65 (s, 3H), 1.15 (d, *J* = 5.1 Hz, 3H).

2.5.2. Peptide Synthesis and Purification

Peptides were synthesized as using NovaPEG rink amide resin or ChemMatrix rink amide resin as solid support.[27] Cationic threonine was synthesized as described in the previous section. All other amino acids were purchased from Chemimpex International or Novabiochem.

Standard Fmoc solid phase peptide synthesis (SPPS) was used to synthesize the peptides.[27,28] Synthesis was done on a 100 μmol scale with 4 equivalents of amino acid, 4

equivalents coupling reagent HCTU (2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophasphate) or PyBOP ((benzotriazol-1-yloxy)tripyrrolidinophosphoniumhexafluorophosphate), 4 equivalents Cl-HOBt (1-hydroxy-6-chloro-benzotriazole), and 8 equivalents diisopropylethylamine (DIEA) in 4 ml DMF for each coupling step. A CEM Mars microwave instrument was used to speed coupling reactions. The coupling program was a 2-minute ramp to 70ºC and a 4-minute hold at 70ºC. In some cases, only two equivalents of the protected unnatural amino acids TS^+, TS^-, and TO^+ were used for the coupling. Deprotections were done with 4 ml of 20% piperidine in DMF with a 2 minute ramp to 80ºC and a 2 minute hold at 80ºC.

Peptides were cleaved with a solution of ethane dithiol (200 μL), water (200 μL), triisopropylsilane (80 μL), and in trifluoroacetic acid (8 mL) for 2-6 hours. The deprotection solution was drained from the resin, and the peptide was precipitated with cold ether. The ether solution was centrifuged, and the solvent was decanted.

All peptides were purified on a Shimadzu reverse phase semi-preparative HPLC instraments with a C18 or C5 reverse phase columns (25x21.2 cm) and eluted with a gradient of acetonitrile and water with 0.1% trifluoroacetic acid. Purity was checked using a Shimadzu reverse-phase analytical HPLC with a 10-60% acetonitrile gradient over 50 minutes gradient. All peptides were greater than 95% pure. Analytical HPLC traces can be found in appendix A. MALDI-TOF MS data were taken on a Bruker REFLEX® II. MALDI-TOF MS data can be found in appendix A.

2.5.3. NMR Sample Preparation and Data Acquisition

Peptides were dissolved in 9:1 H_2O:D_2O sodium acetate buffer (100 mM, pH=3.8) with a 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) standard. Sample concentrations were determined with UV if a Tyr was present in the sequence, or by mass. For each peptide, COSY, TOCSY, ROESY, and 1D data were taken on a 600 mHz INOVA Varian NMR spectrometer with a 3 mm probe ($^{1}H/^{13}C/^{15}N$ with 3 axis PFG) with standard Varian pulse sequences gcosy, wgtocsy, wgroesy, and presat, respectively. The 1D and COSY experiments used solvent suppression with a 0.6-1.5 second solvent presaturation. The TOCSY and ROESY experiments used watergate solvent suppression. The COSY was taken in absolute value mode with gradient echo coherence selection. The TOCSY used a mixing time of 80 ms, and the ROESY had a mixing time of 250 ms. The TOCSY and ROESY were taken in the sensitive mode with hypercomplex phase cycling (States-Haberkorn method).[29,30] In the f2 direction, 2048 points were collected, and in the f1 direction, 220-512 point were collected. Varian VNMR 6.1 software was used to process the data, and sine bell window functions were used before Fourier transformation. All assignments were made with the Sparky program.[31]

2.5.4. Structure Calculations from NOE Restraints

Distance restraints were derived from the integration of NOEs in the 2D NMR spectra. Distances restraints where calculated using the approximation of $I=cr^{-6}$ where I is NOE intensity, c is a constant derived from an NOE of known distance (NOE between the protons on the aromatic ring of Tyr), and r is the distance between protons.[19] Distance restraints were classified as strong (≤3.0), medium (≤3.5), or weak (≤4.5). Simulated annealing was done in CNS version 1.1 with the default settings.[24] Restraints as entered

into CNS are shown below in Section 12. Only distance restraints were used; dihedral, chemical shift, and hydrogen bonding restraints were disabled. For each calculation, 1000 structures were generated, and 10-20 of the lowest energy structures were selected.

2.6 References

(1) Searle, M. S.; Ciani, B. Design of B-Sheet Systems for Understanding the Thermodynamics and Kinetics of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14*, 458–464.

(2) Freire, F.; Almeida, A. M.; Fisk, J. D.; Steinkruger, J. D.; Gellman, S. H. Impact of Strand Length on the Stability of Parallel-Beta-Sheet Secondary Structure. *Angew. Chemie Int. Ed.* **2011**, *50*, 8735–8738.

(3) Kier, B. L.; Shu, I.; Eidenschink, L. a; Andersen, N. H. Stabilizing Capping Motif for Beta-Hairpins and Sheets. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 10466–10471.

(4) Gao, J.; Bosco, D. A.; Powers, E. T.; Kelly, J. W. Localized Thermodynamic Coupling between Hydrogen Bonding and Microenvironment Polarity Substantially Stabilizes Proteins. *Nat. Struct. Mol. Biol.* **2009**, *16*, 684–690.

(5) Hughes, R. M.; Waters, M. L. Model Systems for Beta-Hairpins and Beta-Sheets. *Curr. Opin. Struct. Biol.* **2006**, *16*, 514–524.

(6) Cheng, P. N.; Pham, J. D.; Nowick, J. S. The Supramolecular Chemistry of Beta-Sheets. *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.

(7) McGregor, D. P. Discovering and Improving Novel Peptide Therapeutics. *Curr. Opin. Pharmacol.* **2008**, *8*, 616–619.

(8) Chou, P. Y.; Fasman, G. D. Secondary Structural Prediction of Proteins from Their Amino Acid Sequence. *Trends Biochem. Sci.* **1977**, *2*, 128–131.

(9) Smith, C. K.; Withka, J. M.; Regan, L. A Thermodynamic Scale for the Beta-Sheet Forming Tendencies of the Amino Acids. *Biochemistry* **1994**, *33*, 5510–5517.

(10) Minor, D. L.; Kim, P. S. Context Is a Major Determinant of Beta-Sheet Propensity. *Nature*, 1994, *371*, 264–267.

(11) Swindells, M. B.; MacArthur, M. W.; Thornton, J. M. Intrinsic Phi, Psi Propensities of Amino Acids, Derived from the Coil Regions of Known Structures. *Nat. Struct. Biol.* **1995**, *2*, 596–603.

(12) Muñoz, V.; Serrano, L. Intrinsic Secondary Structure Propensities of the Amino Acids, Using Statistical Phi-Psi Matrices: Comparison with Experimental Scales. *Proteins* **1994**, *20*, 301–311.

(13) Bai, Y.; Englander, S. W. Hydrogen Bond Strength and Beta-Sheet Propensities: The Role of a Side Chain Blocking Effect. *Proteins* **1994**, *18*, 262–266.

(14) Liu, H.; Pattabiraman, V. R.; Vederas, J. C. Stereoselective Syntheses of 4-Oxa Diaminopimelic Acid and Its Protected Derivatives via Aziridine Ring Opening. *Org. Lett.* **2007**, *9*, 4211–4214.

(15) Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Chemical Approaches to the Synthesis of Peptides and Proteins*; CRC Press: Boca Raton, FL, 1997.

(16) Syud, F. A.; Stanger, H. E.; Gellman, S. H. Interstrand Side Chain - Side Chain Interactions in a Designed Beta-Hairpin: Significance of Both Lateral and Diagonal Pairings. *J. Am. Chem. Soc.* **2001**, *123*, 8667–8677.

(17) Haque, T. S.; Little, J. C.; Gellman, S. H. "Mirror Image" Reverse Turns Promote Beta-Hairpin Formation. *J. Am. Chem. Soc.* **1994**, *116*, 4105–4106.

(18) Haque, T. S.; Gellman, S. H. Insights on Beta-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I′ and Type II′ Beta-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

(19) Wurthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(20) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy*; 2nd ed.; Elsevier Academic Press: Burlington, MA, 2007.

(21) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry* **1992**, *31*, 1647–1651.

(22) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. Length-Dependent Stability and Strand Length Limits in Antiparallel Beta-Sheet Secondary Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12015–12020.

(23) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quanfification of Beta-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

(24) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; et al. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. D. Biol. Crystallogr.* **1998**, *54*, 905–921.

(25) Espinosa, J. F.; Gellman, S. H. A Designed Beta-Hairpin Containing a Natural Hydrophobic Cluster. **2000**, *11*, 2330–2333.

(26) Chou, P. Y.; Fasman, G. D. Conformational Parameters for Amino Acids in Helical, Beta-Sheet, and Random Coil Regions Calculated from Proteins. *Biochemistry* **1974**, *13*, 211–222.

(27) Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. Methods and Protocols of Modern Solid Phase Peptide Synthesis. *Mol. Biotechnol.* **2006**, *33*, 239–254.

(28) El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **2011**, *111*, 6557–6602.

(29) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. Structure Determination of a Tetrasaccharide: Transient Nuclear Overhauser Effects in the Rotating Frame. *J. Am. Chem. Soc.* **1984**, *106*, 811–813.

(30) Bax, A.; Davis, D. G. MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. *J. Magn. Reson.* **1985**, *65*, 355–360.

(31) Goddard, T. D.; Kneller, D. G. Sparky 3, 1986.

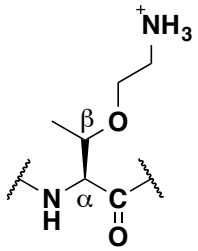
Chapter 3. Redesigning Cationic Threonine: Synthesis of Thioether Cationic and Anionic Threonine, Amino Acids that Promote β-Sheet Secondary Structure

Portions of this chapter appear in:

Maynard, S. J.; Almeida, A. M.; Yoshimi, Y.; Gellman, S. H. "New Charge-Bearing Amino Acid Residues that Promote β-Sheet Secondary Structure." *J. Am. Chem. Soc.* **2014**, 16683-16688.

3.1 Introduction

In the previous chapter, the design and synthesis of the novel amino acid cationic threonine was discussed. Cationic threonine was intended to stabilize β-sheet secondary structure relative to the natural, positively charged amino acids, but unexpectedly failed to do so. β-branching, where two functional groups are present at the β-position of the residue, is associated with high β-sheet propensity.[1] However, cationic threonine, a new β-branched amino acid, demonstrates that β-branching alone is not sufficient to ensure that an amino acid has high β-sheet propensity. Further analysis of the theories surrounding β-sheet propensity lead to insights into the failure of the original cationic threonine design and allowed the development of a charged amino acid that can stabilize β-sheet structure.



Cationic Threonine

3.1.1 β-sheet propensity scales

β-sheet propensity is a measure of a residue's ability to promote β-sheet secondary structure. The classic propensity scale, developed by Chou and Fasman, measures propensity based on the normalized occurrence of a given amino acid in β-sheet segments within proteins of known structure in the protein data bank (pdb).[2,3] Experimental propensity scales have also been developed (see chapter 1 for more detail).[4–7]

While the various scales may differ in the exact order and magnitude of various amino acids' β-sheet propensities, they all show the same general trends. Aromatic and β-branched amino acids all have high β-sheet propensities.

3.1.2. Understanding β-sheet propensities

The origins of β-sheet propensities have been explored by numerous authors, and a plethora of theories have been proposed.[1,8–15] Generally, the theories plot β-sheet propensity relative to a feature of the amino acid and draw a correlation. While an amino acid's β-sheet propensity is probably due to several complex factors and not just a single parameter, several theories do have some merit. (See Chapter 1 for a more detailed discussion of β-sheet propensity.)

One theory of note for predicting β-sheet propensity is the theory that propensity is related to an amino acid's intrinsic Φ/Ψ angles.[1] The Ramachadron plot is a graph of the possible angles around the flexible bonds in the backbone of a peptide (Figure 3.1). Due to steric interactions, only certain portions of the Φ/Ψ plot are actually accessible to amino acids.[16] Different secondary structures have Φ/Ψ angles in different regions of the Ramachadron plot.[16]

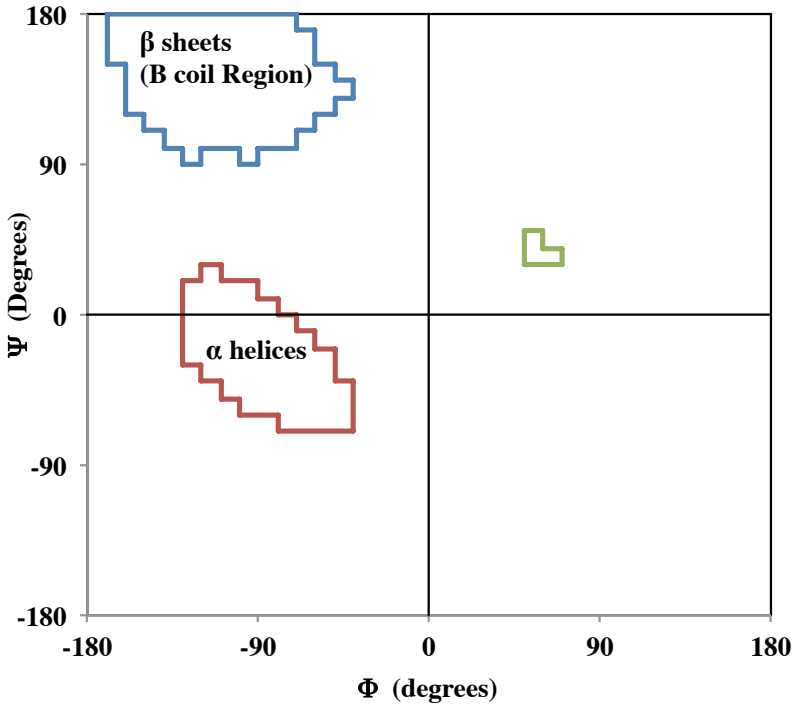


Figure 3.1 Ramachandron plot

According the to theory, if the intrinsic Φ/Ψ angles of a residue fall into the β-sheet region of the Φ/Ψ plot, then it will have a high β-sheet propensity and stabilize β-sheet secondary structure.[1] Features that stabilize the Φ/Ψ angles that are associated with helical secondary structure push the intrinsic Φ/Ψ angles away from those necessary for β-sheet formation and therefore lower the β-sheet propensity of the amino acid.[1]

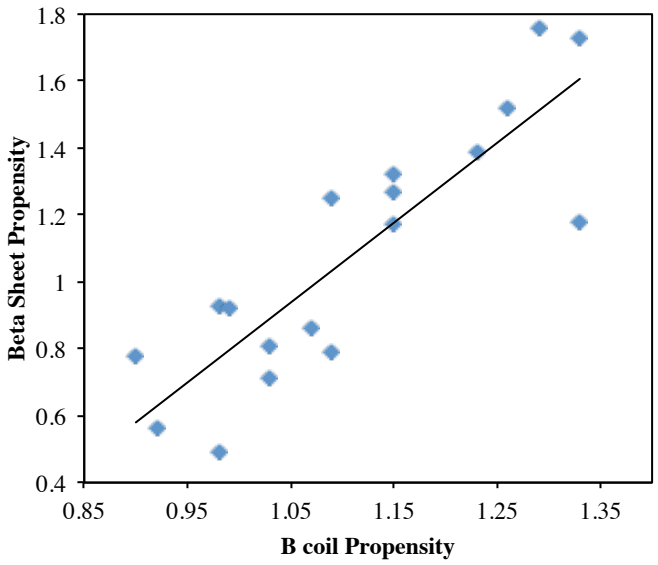
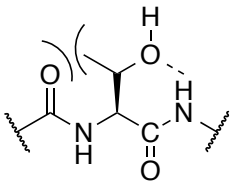


Figure 3.2. Correlation between β-sheet propensity of natural amino acids and the propensity for the Φ/Ψ angles of the amino acid to fall into the β-sheet region of the Φ/Ψ plot (B coil propensity)[1]

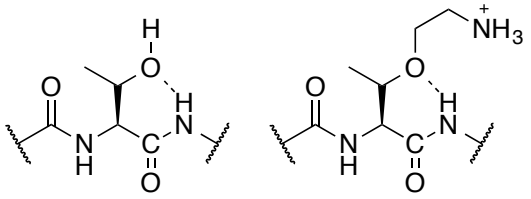
The theory provides an explanation for why β-branched amino acids have high β-sheet propensities.[1] When the β-branched amino acids adopt the Φ/Ψ angles in the α-helical region of the Φ/Ψ plot, there is a steric interaction between the β-methyl group and the backbone of the peptide (see inset).[1] This steric interaction pushes the Φ/Ψ angles of the amino acid away from the α-helical region of the Φ/Ψ plot and toward the β-sheet region of the Φ/Ψ plot.[1] The authors note that amino acids with side-chains that can hydrogen bond to the backbone of the peptide tend to have lower β-sheet propensities, theoretically because the hydrogen bond can only be formed when the amino acid adopts Φ/Ψ angles in the α-helical region of the Φ/Ψ plot.[1] Threonine is an interesting case because it is a β-branched residue with the potential to hydrogen bond to the backbone of the peptide. However, in the case of threonine, the steric interaction that disfavors the α-helical region appears to



be more energetically important than the hydrogen bond as its intrinsic Φ/Ψ angles still fall in the β-sheet region of the Φ/Ψ plot. It should be noted, however, that Ile and Val have slightly higher β-sheet propensities than Thr, possible due to the hydrogen bonding capabilities unique to Thr. [1]

3.1.3. Understanding the low β-sheet propensity of cationic threonine (TO^+)

The Φ/Ψ β-sheet propensity theory provides a possible explanation for the unexpectedly low β-sheet propensity of cationic threonine observed in Chapter 2. Like threonine, cationic threonine can hydrogen bond to the peptide backbone. However, the hydrogen bond accepter is an ether in TO^+, rather than an alcohol, and an ether is the better hydrogen bond acceptor.[17] It's also possible that in the context of a peptide, the ether is less solvent-exposed than the alcohol, making the potential hydrogen bond more energetically favorable. In theory, the stronger hydrogen bond pushes the Φ/Ψ angles away from the β-sheet region of the Φ/Ψ plot and leads to a lower β-sheet propensity.



In an attempt to prevent the potentially problematic hydrogen bond, cationic threonine was redesigned with a thioether linkage instead of an ether linkage as thioethers are not good hydrogen bond acceptors (Figure 3.3).[17] The new residue, thioether cationic threonine (TS^+) was synthesized and tested in the context of a hairpin. Likewise, a negatively charged version of anionic threonine, thioether anionic threonine (TS^-), was synthesized and tested.

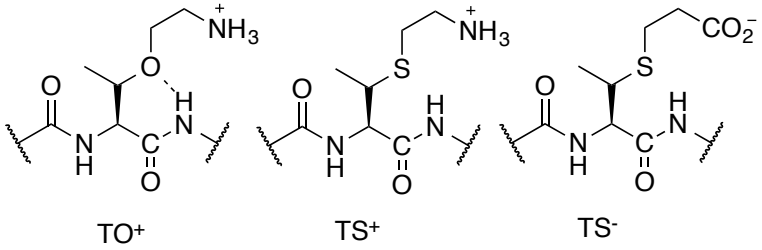
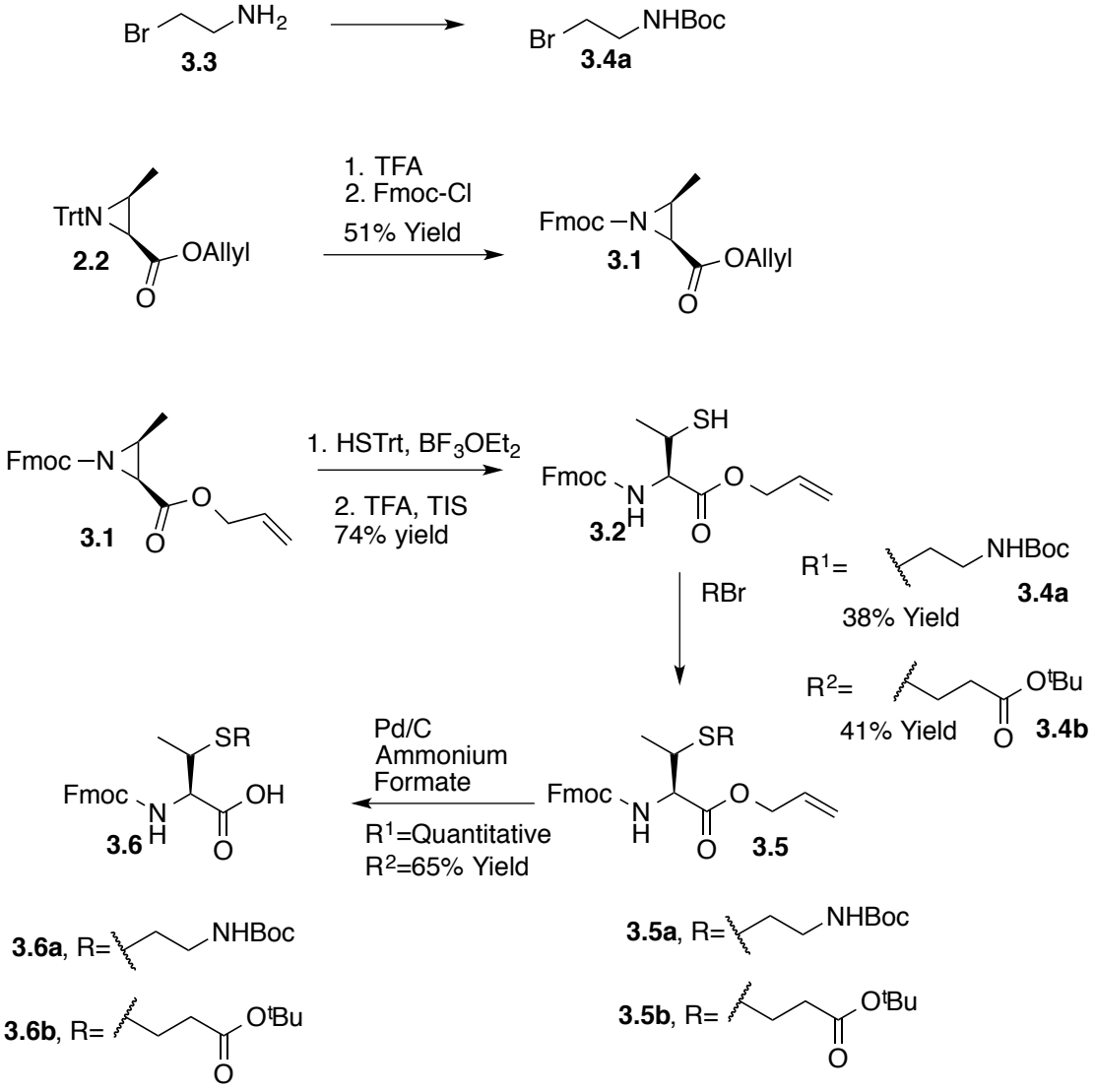


Figure 3.3 (A) Cationic threonine, (B) thioether cationic threonine, and (C) thioether anionic threonine (TS^-)

3.2 Synthesis of protected forms of thioether cationic threonine (TS^+) and thioether anionic threonine (TS^-)

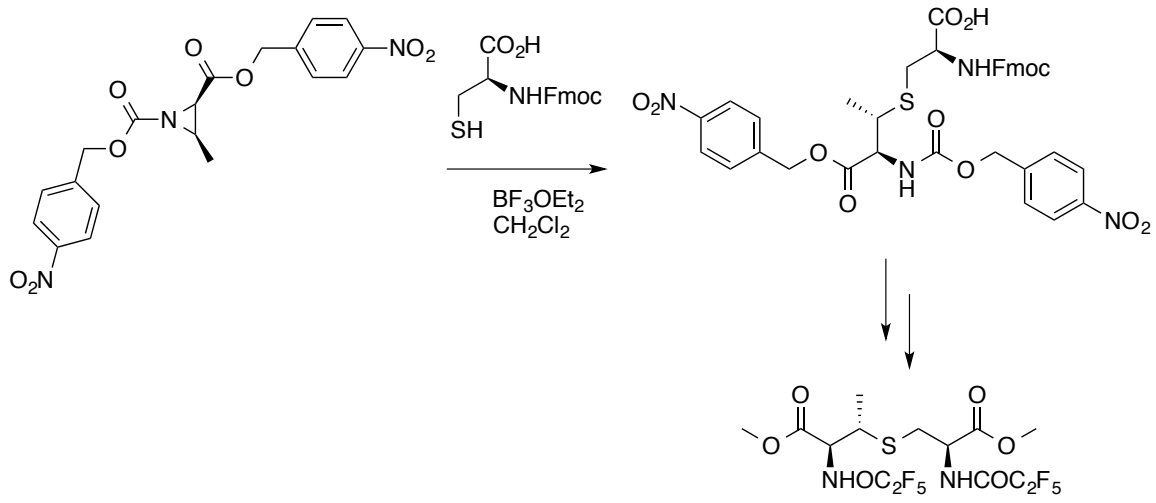
Thioether cationic threonine (TS^+) and thioether anionic threonine (TS^-) were synthesized by the ring opening of an aziridine using methods derived from VanNieuwenhze *et al*. (Scheme 3.1).[18,19] The aziridine route was chosen as it is known to proceed with high stereoselectivity. Aziridine **2.2** was synthesized from threonine, a cheap source of chiral centers, using previously described methods.[20] Aziridine **2.2** was protected with an Fmoc group to give aziridine **3.1** and ring opened with trityl thiol to give thiol **3.2** as previously described.[18,19] Alkyl bromide **3.3** was protected with a Boc group to give alkyl bromide **3.4a**. The functionalization of thiol **3.2** with alkyl bromide **3.4a** or **3.4b** yielded thioethers **3.5a** and **3.5b**, respectively. Protecting group manipulation gave thioethers **3.6a** and **3.6b**, which are protected for standard Fmoc solid phase peptide synthesis (SPPS).

Scheme 3.1 Synthesis of TS^+ and TS^-



Note that the stereochemistry of the aziridine ring opening was assigned based on literature precedent.[21–23] A specific example of a closely related reaction from the literature is given in Scheme 3.2.[21] The product of the ring-opening reaction was derivatized and its stereochemistry was assigned based on comparison of retention times on chiral GC/MS to standards of all possible diastereomers with known stereochemistry. (All diastereomers had different retention times and the product was coinjected with the molecule thought to have the same stereochemistry.)[21]

Scheme 3.2. Literature[21] example of an aziridine ring-opening reaction closely related to the aziridine ring opening used to form aziridine **3.2**.



3.3 Evaluation of the ability of TS^+ and TS^- to stabilize β-sheet structure

3.3.1 Evaluation of the ability of TS^+ to stabilize β-sheet secondary structure

In order to test the ability of TS^+ to stabilize β-sheet structure relative to a naturally occurring positively-charged residue, TS^+ was substituted for Lys in a previously characterized hairpin from Syud *et al*.[24] (Figure 3.4). The TS^+ containing hairpin 3.7a was synthesized by standard solid phase peptide synthesis (SPPS). Note that the TS^+ residue was incorporated into peptides with good yields using standard peptide coupling conditions. A set of 2D NMR spectra (COSY, TOCSY, and ROESY) was taken at 4ºC and the ^{1}H chemical shifts were assigned. The ROESY experiment shows signals between protons that are close in space, and these correlations provide information about the conformation (See Chapter 2 for further explanation).[25] As expected, the ROEs associated with peptide 3.7a are consistent with β-sheet secondary structure. The cross-

strand ROEs for peptide 3.7a are characteristic of a β-sheet conformation (Figure 3.4d). As cross-strand ROEs arise between residues distant from each other in the primary sequence, they are assumed to be the result of the folded conformation of the peptide. Peptide 3.7a has several medium and weak cross-strand ROEs. Due to overlap, many of the ROEs observed in the spectra could not be definitively assigned. Therefore, an additional set of spectra was taken at 10ºC (Figure 3.4c). Again, numerous medium and weak cross-strand ROEs were observed and all the ROEs were consistent with the expected β-sheet structure.

As described in detail in Chapter 2, the chemical shifts of the α-protons in residues ($\delta C_\alpha H$) are indicative of secondary structure.[24,26] Comparisons of α-proton chemical shifts in the peptides of interest to the chemical shifts in unstructured control peptides allow the calculation of chemical shift deviations ($\Delta\delta C_\alpha H$).[26] As discussed in detail in chapter 2, the unstructured control peptides are analogues of the peptides of interest where DPro is replaced with LPro. LPro cannot form the two-residue turn necessary to promote a hairpin and therefore serves as an unfolded control (Figure 3.5).[27–30] The LPro diastereomer of peptide 3.7a (peptide 3.7b) was analyzed by COSY, TOCSY, and ROESY experiments, and the ^{1}H signals were assigned. Non-sequential ROEs arise due to the proximity of ^{1}H resonances in the folded conformation. The lack of non-sequential ROEs in the 2D NMR analysis of LPro 3.7b supports the assumption that this peptide is unstructured and allows the determination of the random coil chemical shifts for the sequence. Positive chemical shift deviations for peptide 3.7a indicate β-sheet secondary structure, and chemical shift deviations can be

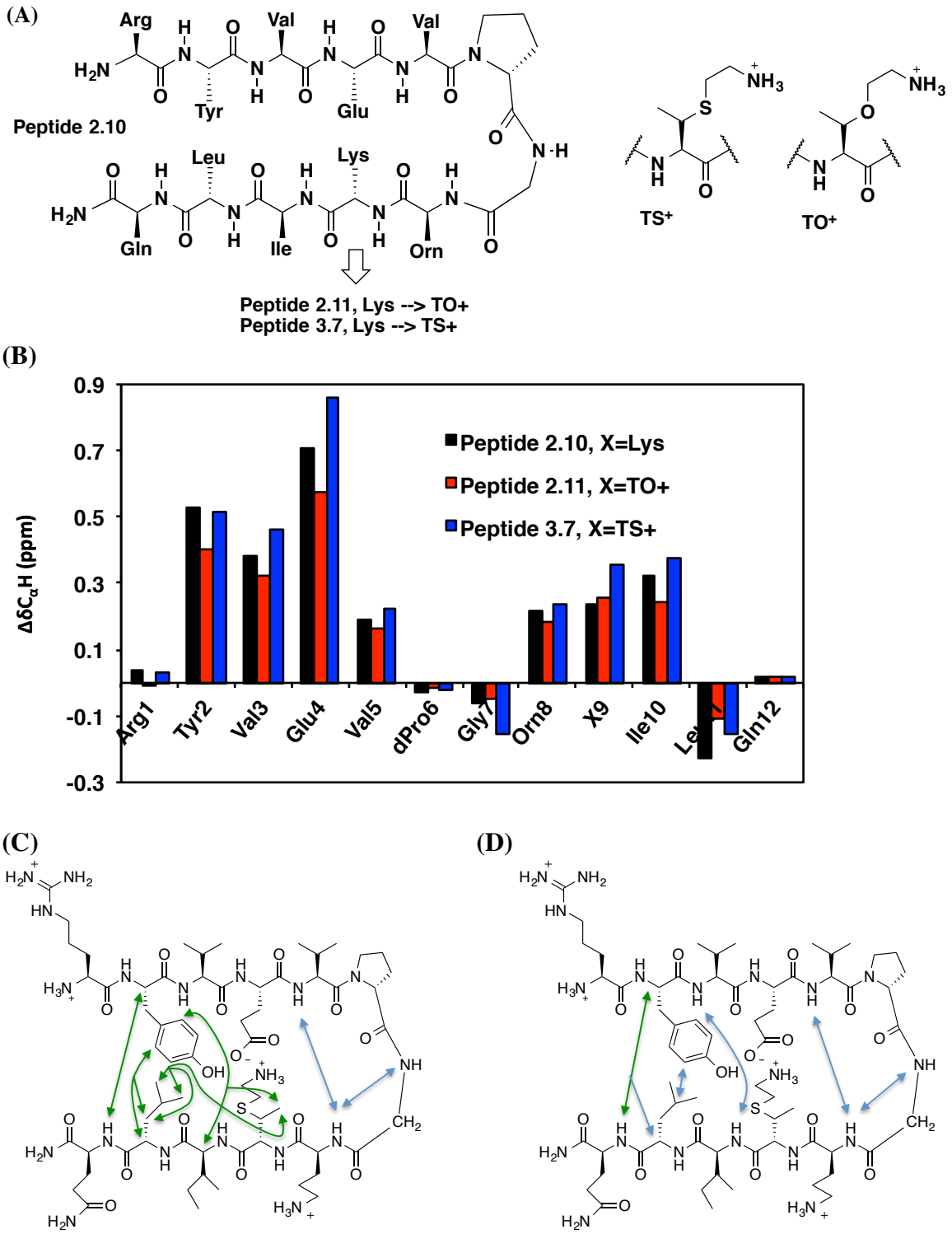
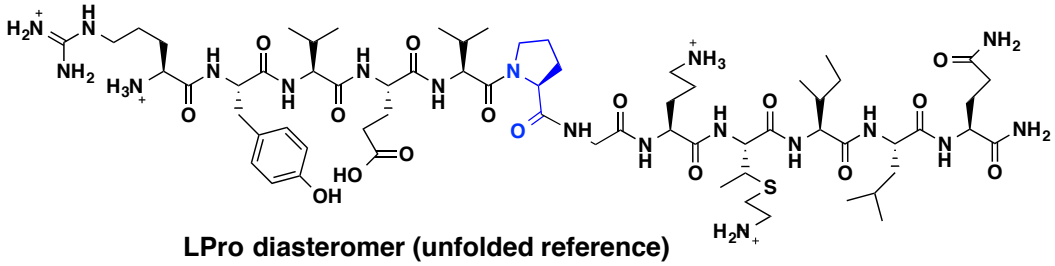


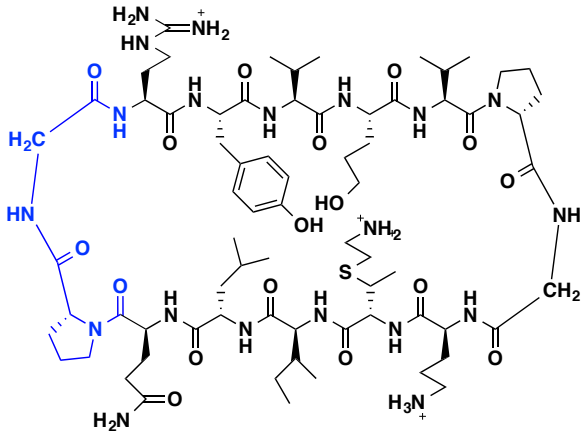
Figure 3.4. (A) Sequences of hairpins. (B) Chemical shift deviations calculated relative to DPro analogues of peptides of the same sequence. Parent peptide chemical shift were from Syud *et al*.[24] Chemical shifts at or adjacent to the substitution cannot be compared due to changes in dynamic range. (C) Summary of cross-strand ROEs observed for peptide 3.7a at 10°C and (D) at 4°C. Blue arrows are medium ROEs and green arrows are weak ROEs.

used to determine qualitative relative thermodynamic stability of hairpin peptides with closely related sequences.[26]

The chemical shift deviations ($\Delta\delta C_\alpha Hs$) of peptide 2.10a and 3.7a are consistent with the expected β-sheet structure (Figure 3.4b). At residues in the core of the β-sheet (Val 3-Val 5 and Orn8-Ile10), the $\Delta\delta C_\alpha Hs$ are greater for TS^+ containing peptide 3.7a, supporting the hypothesis that TS^+ stabilizes β-sheet structure relative to Lys. Note that chemical shift deviations of residues at or adjacent to the substitution cannot be directly compared to each other due to changes in dynamic range (see Chapter 2 for more detail). TS^+ was also compared to TO^+ (Figure 3.4b). Qualitatively, TS^+ stabilizes the peptide relative to TO^+.



LPro diasteromer (unfolded reference)



Cyclic analogue (folded reference)

Figure 3.5. Reference peptides for thermodynamic analysis.

In order to quantify the extent that TS^+ stabilizes the β-sheet relative to Lys, controls were synthesized and analyzed (Figure 3.5). A cyclized analogue of peptide 3.7a (cyclic peptide 3.7c) serves as a fully folded control (Figure 3.6). Cylic hairpin 3.7c was analyzed by 2D NMR (ROESY, TOCSY, COSY). All the observed ROEs are consistent with the expected β-sheet hairpin structure. Simulated annealing structural calculations were done for cyclic peptide 3.7c using distance restraints derived from the ROEs (see Appendix B for more detail). The resulting 18 lowest energy structures were consistent with the expected β-sheet conformation (Figure 3.6).

The $\Delta\delta C_\alpha H$ of the random coil and folded controls allow the calculation of the equilibrium constant (K_F) and ΔG_F (see Chapter 2 for further detail). Hydrogen bonding positions in the core of the peptide are chosen as indicator positions and used for ΔG_F calculations as they are thought to most accurately reflect the folded population of the peptide.[31] The $\Delta\Delta G_F$ value for the indicator positions (Val3, Val5, Orn8, Ile10) were calculated and then averaged (see Table 3.1 in Section 3.3 for the values). The $\Delta\Delta G_F$ between peptide 2.10a and 3.7a is -0.5±0.4 kcal/mol. For parent peptide 2.10a, a Lys → TS^+ substitution stabilizes the peptide. The $\Delta\Delta G_F$ for a TO^+ → TS^+ substitution can also be calculated by comparing peptides 2.11a (Lys→TO^+) and 3.7a (Lys→TS^+). The $\Delta\Delta G_F$ is -0.6±0.3 kcal/mol. Note that this value is determined by calculating $\Delta\Delta G_F$ from the chemical shifts of peptides 2.11 and 3.7 instead comparing the $\Delta\Delta G_F$ values for 3.7 and 2.11 relative to parent peptide 2.10. The direct comparison gives a more precise $\Delta\Delta G_F$ value as it does not include error associated with the measurement of the folded populations in parent peptide 2.10.

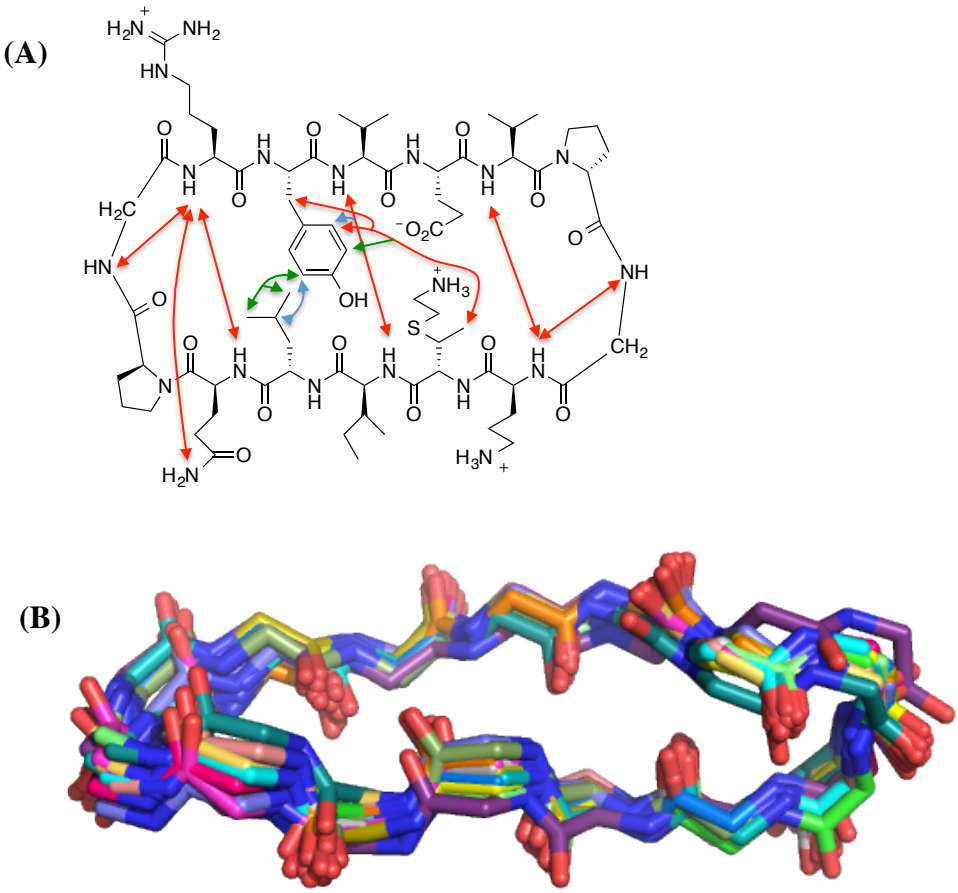


Figure 3.6. (A) Cross-strand ROEs observed for cyclic peptide 3.7c at 4ºC. Strong ROEs are red, medium ROEs are blue, and weak ROEs are green. (B) Backbone atoms for the 18 lowest energy structures calculated from ROE distance restraints.

As there can be some context dependence to the β-sheet propensity of an amino acid residue, it was important to test a Lys → TS^+ substitution in a completely different sequence. Peptide 3.8 (Lys → TS^+) was derived from parent peptide 2.15, a sequence developed by Haque *et al*.[29] (Figure 3.8a). Peptide 3.8 was analyzed by COSY, TOCSY, and ROESY NMR experiments. All ROEs observed are consistent with the expected hairpin structure (Figure 3.7). Note that fewer ROEs are observed for sequences derived from peptide 2.15 than peptide 2.10, possibly because the β-hairpin conformation of peptide 2.15 is not as stable.

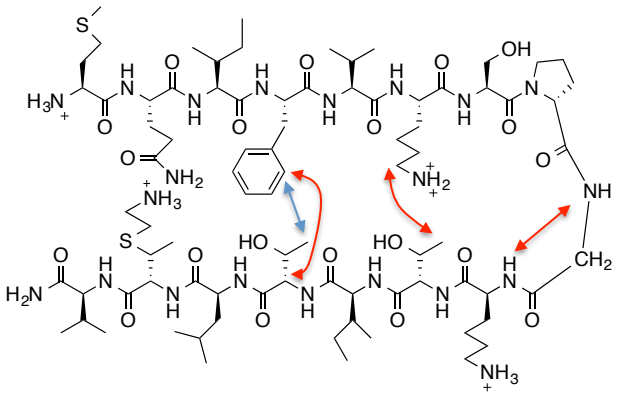


Figure 3.7. Summary of ROEs observed for peptide 3.8. Red arrows are strong ROEs. Blue arrows are medium ROEs.

The chemical shifts of parent peptide 2.15, peptide 2.16 (Lys →TO^+), and peptide 3.8 (Lys→TS^+) were compared. The chemical shift deviations ($\Delta\delta C_\alpha Hs$) were calculated using the LPro diastereomer of peptide 2.15 as a reference for the chemical shifts of random coil population (Figure 3.8). Chemical shifts at or adjacent to the substitution cannot be analyzed as the LPro of the parent peptide is not a suitable reference for the random coil chemical shifts of peptides 2.16 and 3.8. At all other positions in the core of the β-sheet interaction, the $\Delta\delta C_\alpha H$ for peptide 3.8 (Lys→TS^+) are greater than for parent peptide 2.15, consistent with the hypothesis that TS^+ stabilizes β-sheet structure relative to lysine. The $\Delta\delta C_\alpha H$ of peptide 2.16 (Lys15→TO^+) and peptide 3.8 (Lys15→TS^+) are consistent with the conclusion that TS^+ stabilizes sheet structure relative to TO^+.

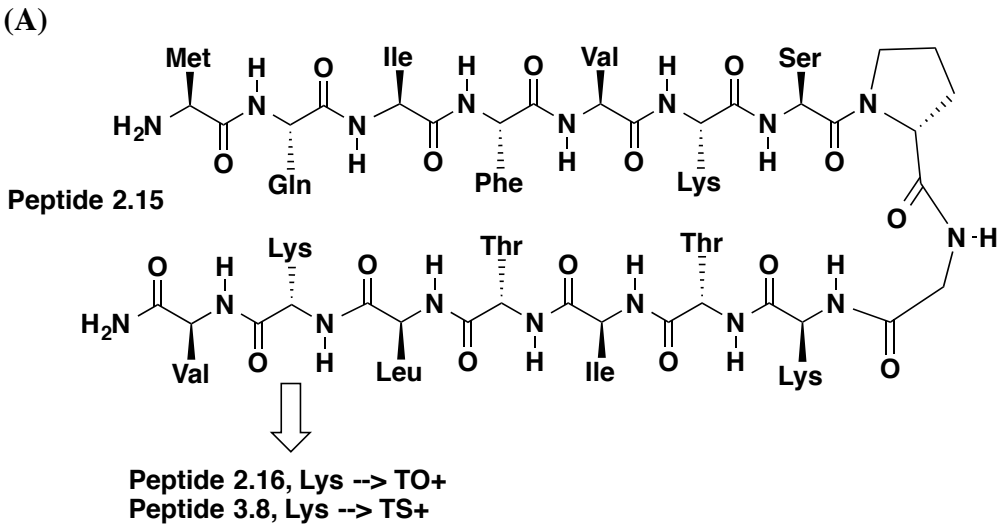


Figure 3.8. A) Sequences of hairpins. (B) Chemical shift deviations calculated relative to DPro analogues of peptide 2.15. Chemical shifts at or next to the substitution cannot be compared. Parent peptide chemical shifts are from Haque *et al*.[29]

In order to be useful, TS^+ need only be better at promoting β-sheet structure than the existing positively charged residues. However, it is also useful to know how well TS^+ promotes β-sheet structure relative to the existing amino acids with high β-sheet propensity. The substitution of a hydrophobic or polar residue for a positively charged residue has the potential to pack the new positively charge against an existing charged

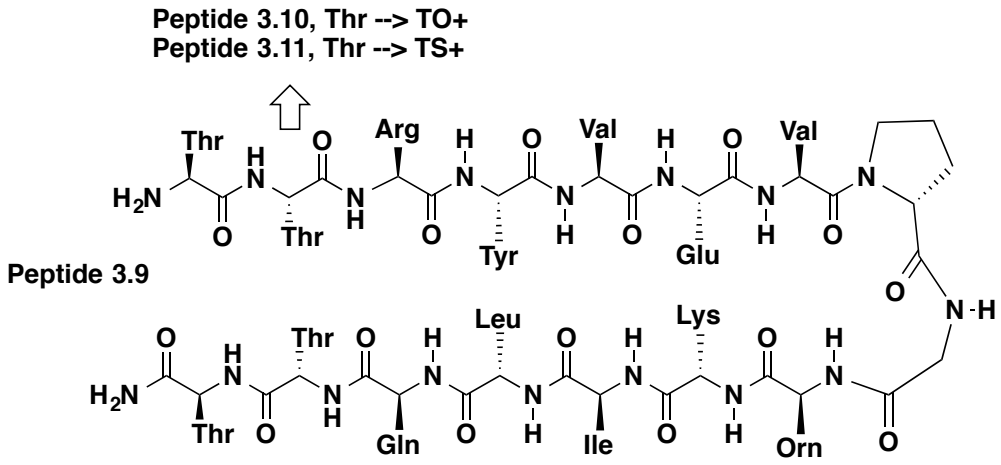
residue in the folded hairpin structure. Therefore, in order to have a meaningful comparison, one must choose a position where the new positive charge is not close in space to any existing charge in the folded hairpin structure. Given that all small water-soluble hairpins are highly charged, few options for such a substitution exist in the well-characterized hairpin peptides. However, the position of Thr2 in hairpin peptide 3.9 from Stanger *et al*.[31] should, in theory, allow a comparison of TS^+ to Thr without the added complication of extra charge-charge interactions (Figure 3.9). If TS^+ has a β-sheet propensity close to that of Thr, the $\Delta\delta C_\alpha H$ data for the two peptides should be similar. As the substitution is close to the terminus of the hairpin, it is possible that the position is part of the 'splayed ends' of the hairpin rather than strictly within the β-sheet. If this is the case, the Thr→TS^+ substitution may not impact the stability of the structure simply because it's not really within the hairpin. To control for this possibility, the $\Delta\delta C_\alpha H$s of peptide 3.10 (Thr→TO^+) were also measured. As TS^+→TO^+ destabilizes β-sheet structure, if position 2 is within the core of the hairpin, the substitution will impact the $\Delta\delta C_\alpha H$ data. Peptide 3.10 also helps control for the possibility that parent peptide 3.9 is too stable for the chemical shift deviations to be sensitive to small differences in stability. If the TS^+→TO^+ does not have a large impact on chemical shift, than nothing can be concluded from the chemical shift deviations of peptide 3.11 (Thr→TS^+).

Peptides 3.10 and 3.11 where synthesized and analyzed by 2D NMR (COSY, TOCSY, and ROESY). The proton resonances were assigned. All non-sequential ROEs were consistent with the expected β-sheet secondary structure (see appendix B).

The chemical shift deviations of peptides 3.9, 3.10, and 3.11 were analyzed. There is no significant difference among the $\Delta\delta C_\alpha H$ of peptides 3.9, 3.10, and 3.11. Therefore,

it is likely that the position of the substitution is not the core of the β-sheet structure and no conclusions can be drawn from this data set.

(A)



(B)

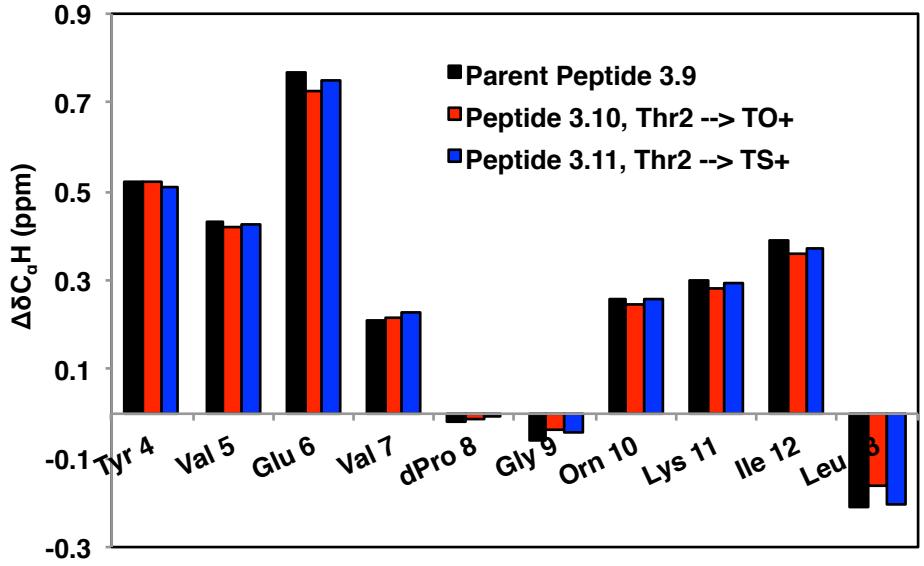
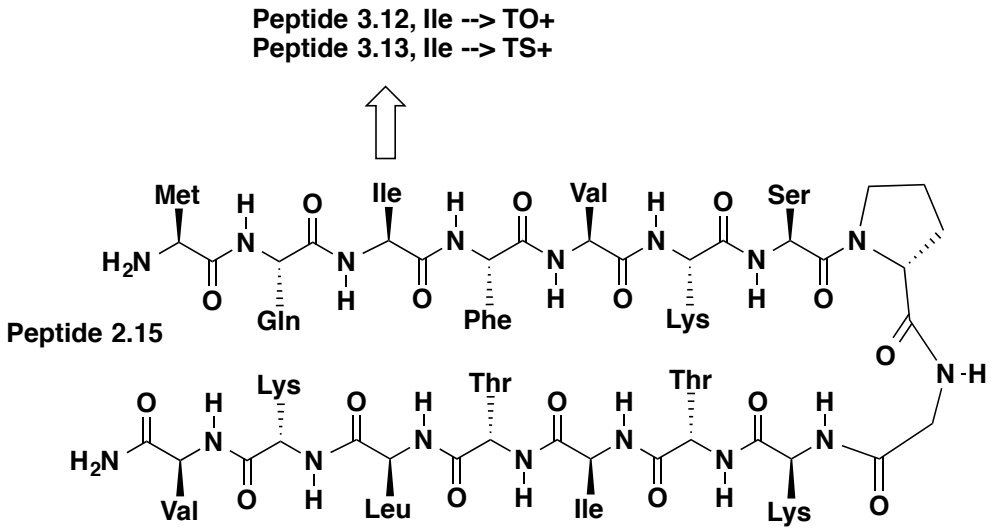


Figure 3.9. (A) Peptide sequences. (B) Chemical shift deviations. All $\Delta\delta C_\alpha H$ are measured relative to a unfolded control 2.10b.[31]

TO^+ and TS^+ were compared to Ile using parent peptide 2.15 (Figure 3.10).[29] Ile is a hydrophobic β-branched amino acid with a high β-sheet propensity.[1] Peptide 3.12 (Ile3→TO^+) and peptide 3.13 (Ile3→TS^+) were synthesized and analyzed by COSY, TOCSY, and ROESY. The ROEs observed in the resulting spectra were consistent with the expected structure (see appendix B).

The $\Delta\delta C_\alpha H$ data at each position of parent peptide 2.15 and peptide 3.12 (Ile3→TO^+) were compared (Figure 3.9). Given that TO^+ appears to have a β-sheet propensity close to that of Lys, TO^+ should destabilize sheet structure relative to an amino acid with a high β-sheet propensity like Ile. Peptide 3.12 (Ile→TO^+) had smaller

(A)



(B)

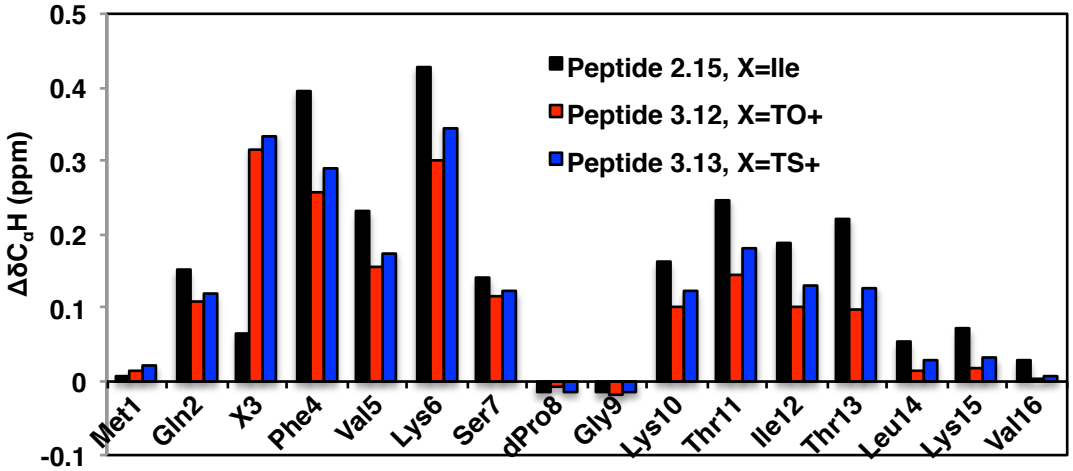


Figure 3.10. (A) Peptide sequences. (B) $\Delta\delta C_\alpha H$s of peptides derived from Haque *et al*.[29] Chemical shifts were referenced to an LPro analogue of peptide 2.15. $\Delta\delta C_\alpha H$s at or adjacent to the substitution cannot be directly compared.

$\Delta\delta C_\alpha H$s than parent peptide 2.15. Therefore, as expected, TO^+ destabilized the hairpin conformation relative to Ile.

The $\Delta\delta C_\alpha H$ data of peptides 2.15 and 3.13 (Ile3→TS^+) were compared (Figure 3.4). The $\Delta\delta C_\alpha H$s of peptide 3.13 (Ile3→TS^+) were smaller than the corresponding $\Delta\delta C_\alpha H$ data for peptide 2.15. Therefore, TS^+ appears to destabilize the β sheet relative to Ile. Ile is generally thought to have one of the highest β-sheet propensities,[2,4,5] so it's still possible TS^+ may stabilize sheet structure relative to Val or Thr. However, in order to be useful, TS^+ need only be better than existing charged residues.

3.3.2 Evaluation of the ability TS^- to stabilize β-sheet secondary structure

In addition to synthesizing a positively charged amino acid that stabilizes sheet structure, I also prepared a negatively charged amino acid, TS^- (Scheme 3.1). Negatively charged TS^- was compared to negatively charged glutamic acid in the context of peptide 2.10a (Figure 3.11).[24] Resulting hairpin 3.14a (Glu→TS^-) was synthesized and analyzed by 2D NMR experiments. The observed ROEs were all consistent with the expected β-sheet secondary structure. As expected for a β-hairpin conformation, several strong, medium, and weak cross-strand ROEs were observed (Figure 3.11c).

The chemical shifts of hairpin 3.14a were analyzed and referenced relative to the LPro analogue as described in Chapter 2. Non-sequential ROEs are cross-peaks between ^1H resonances in residues greater than i to i+1 positions apart in the primary sequence. These ROEs arise due to the folded conformation of the peptide and are not expected to be observed for peptide that are unstructured. No non-sequential ROEs were observed in

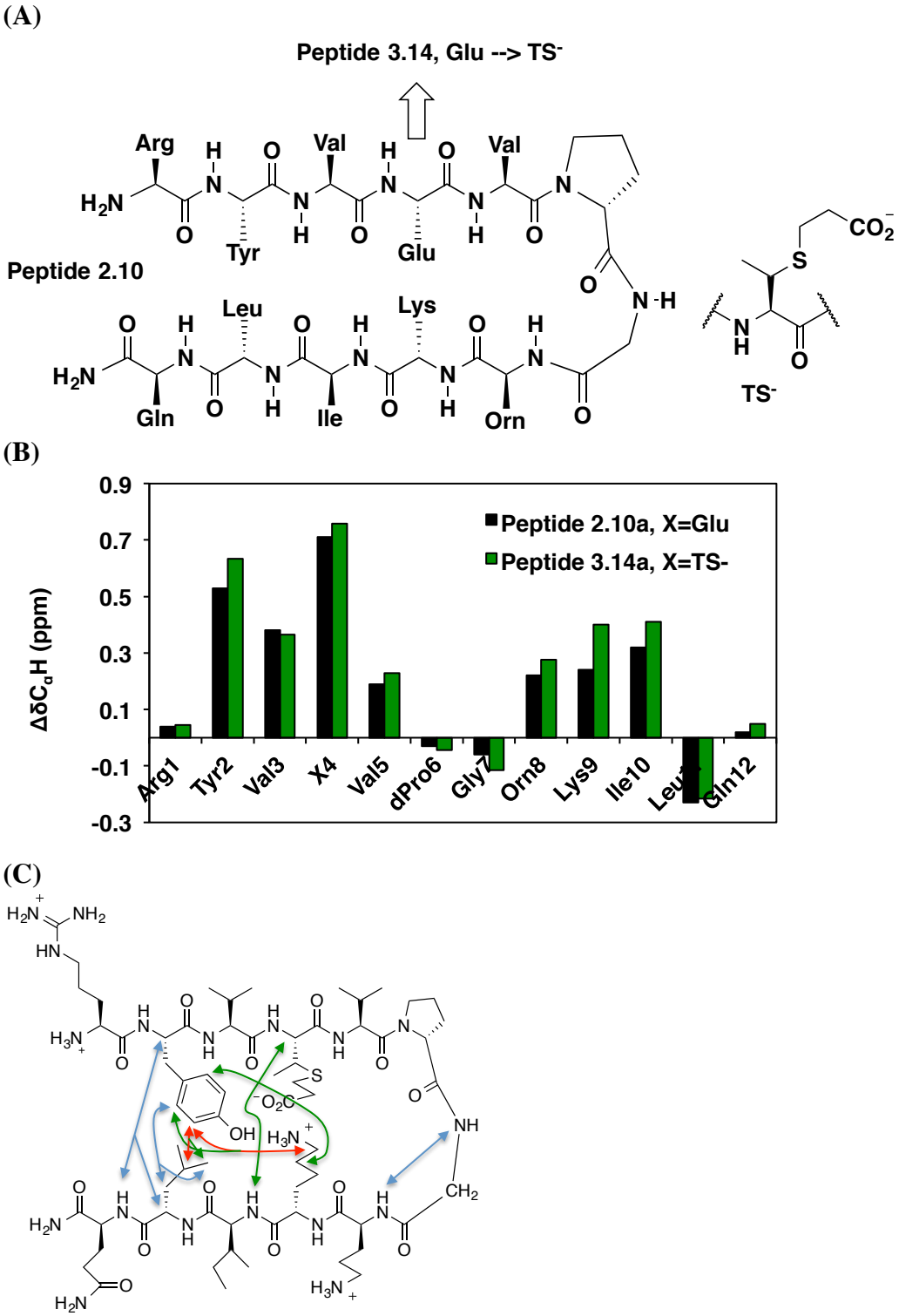
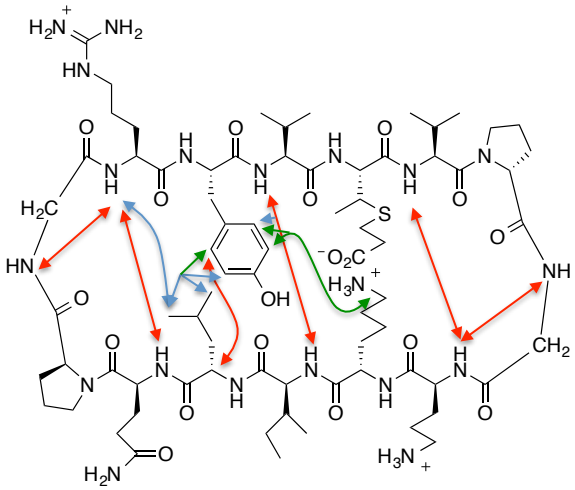


Figure 3.11. (A) Peptide sequences. (B) $\Delta\delta C_\alpha H$. Chemical shifts were referenced to an unfolded control (in which LPro is substituted for DPro) of a peptide of the same sequence. (See Chapter 2 for details.) Chemical shifts at or adjacent to the substitution cannot be directly compared. The parent peptide sequence and data are from Syud *et al*.[24] (C) Cross-strand ROEs observed for peptide 3.14a.

the 2D NMR spectra of the LPro peptide 3.14b, which is consistent with the assumption that the peptide is a suitable unfolded control. The chemical shift deviations ($\Delta\delta C_\alpha H$) for parent peptide 2.10a and peptide 3.14a (Glu→TS^-) were compared (Figure 3.11b). In general, the chemical shift deviations of peptide 3.14a are greater than those of peptide 2.10a, which is consistent with the hypothesis that TS^- stabilizes β-sheet structure relative to Glu.

In order to quantify the difference between Glu and TS^-, a control to represent the fully populated hairpin was synthesized for peptide 3.14a (cyclic peptide 3.14c). Cyclic peptide 3.14c was analyzed by COSY, TOCSY, and ROESY and ^1H resonances were assigned. The observed ROEs were all consistent with the expected β-sheet secondary structure (Figure 3.12). Simulated annealing calculations with CNS converged to the expected β-sheet secondary structure (Figure 3.11b).

(A)



(B)

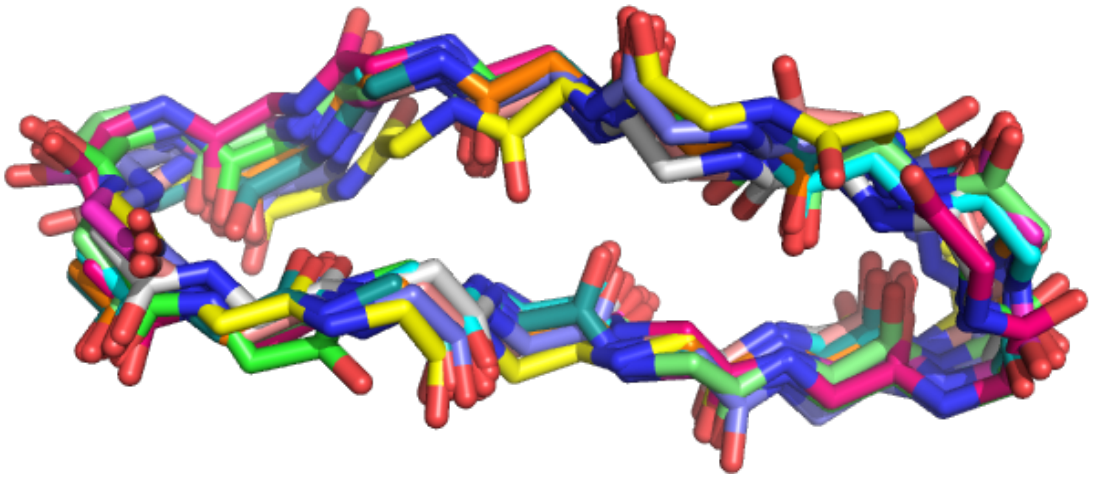


Figure 3.12. (A) ROEs observed for cyclic peptide 3.14c at 4°C. Strong ROEs are red, medium ROEs are blue, and weak ROEs are green. (B) Backbone atoms for the 11 lowest energy structures calculated from ROE distance restraints.

The thermodynamic stability of peptide 3.14c (Glu→TS$^-$) relative to parent peptide 2.10a was analyzed (see Chapter 2 for details). Cyclic peptide 3.14c and LPro peptide 3.14b served as chemical shift references for the folded hairpin populations and random coil states, respectively. The Glu→TS$^-$ substitution stabilized the peptide by -0.8±0.5 kcal/mol. Thermodynamic data for peptides derived from 2.10 are summarized in Table 3.1.

Table 3.1. Thermodynamic data for peptides in β-hairpins derived from Peptide 2.10a.

Peptide	Val 3	Val 5	Orn 8	Ile 10	$\Delta\Delta G_{fold}$ (kcal/mol)[a]
2.10a[b]	70%	76%	65%	59%	
2.11a (Lys9→TO^+)	62%	77%	63%	51%	+0.1±0.2
3.7a (Lys9→TS^+)[c]	85%	93%	76%	71%	-0.5±0.4
3.14a (Glu4→TS^-)	93%	96%	84%	77%	-0.8±0.5

[a]The $\Delta\Delta G_F$ was calculated for each of the hydrogen bonding positions in the core of the β sheet and then averaged. [b]Data from Syud et al.[24] [c]For peptide 2.11a relative to 3.7a, ΔΔG=-0.6±0.3 kcal/mol.

In order to account for context dependence, the effect of the Glu→TS^- substitution was tested in a second hairpin with an unrelated sequence from Haque *et al*. (Figure 3.13).[29] As the parent peptide 2.15 does not contain Glu, an Ile3→Glu substitution was made (Figure 3.13a). The resulting hairpin 3.15 was analyzed by 2D NMR experiments (ROESY, TOCSY, and COSY). The chemical shift deviations of hairpin 3.15 were determined by referencing it to the LPro analogue of peptide 2.15 (where LPro is substituted for DPro), whose chemical shifts represent the random coil chemical shifts. Most of the chemical shift deviations of peptide 3.15 are smaller than the chemical shift deviations for peptide 2.15 (Figure 3.13b). (Recall chemical shift deviations at or adjacent to the substitution cannot be compared.) As expected from the literature β-sheet propensity scales, the Ile→Glu substitution destabilizes the hairpin.

Peptide 3.16 (Ile3→TS^-) was synthesized and analyzed by 2D NMR. The resulting chemical shift deviations were measured. The chemical shift deviations for peptide 3.16 are greater than the chemical shift deviations for peptide 3.15. Therefore TS^- stabilizes the hairpin relative to Glu.

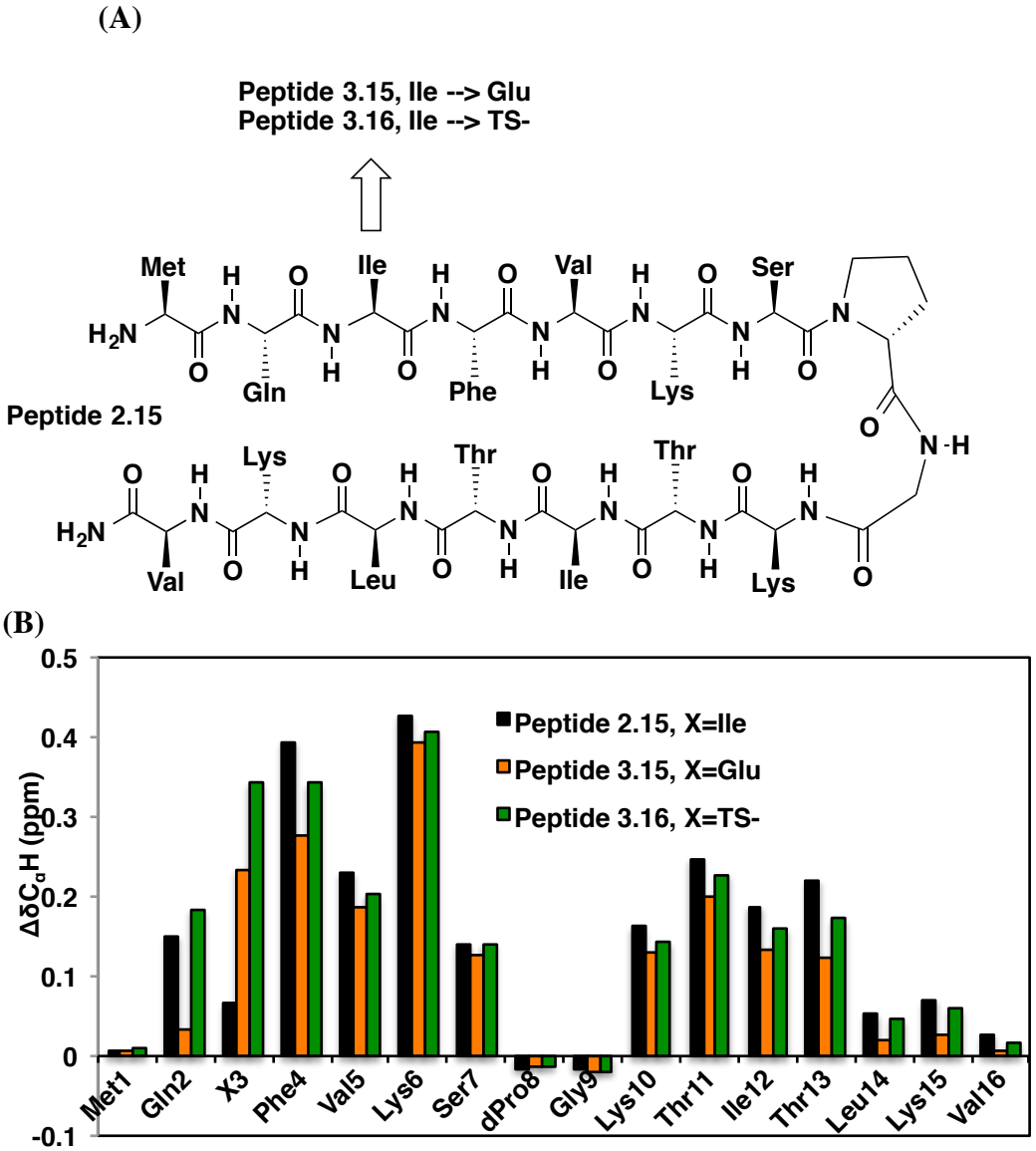


Figure 3.13. (A) Peptide sequences and their (B) $\Delta\delta C_\alpha H$. Chemical shifts were referenced to an unfolded LPro analogue of peptide 2.15. Chemical shift deviations at or adjacent to the substitution cannot be directly compared.

Note that peptide 3.16 (Ile→TS^-) can also be compared to parent hairpin 2.15. The chemical shift deviations of peptide 3.16 are smaller than those of peptide 2.15. Therefore, TS^- destabilizes the peptide relative to Ile. However, in order to be useful, TS^- only needs to have a higher β-sheet propensity than the natural negatively charged amino acids. Furthermore, Ile is thought to have the highest β-sheet propensity of the natural β-

branched amino acids,[2] so it is still possible that TS^- is as stabilizing as one of the other β-branched amino acids such as Thr or Val.

3.3.3 Notes on the effects temperature and pH

All NMR studies were done in acetate buffer at pH 3.8 and 4ºC. Although these conditions are common for the analysis of peptides, they differ from physiological pH (7.0) and temperature (37ºC). In order to probe the global effect of temperature and pH, I used circular dichroism spectroscopy (CD) to probe the secondary structure of my peptides.[32] β-sheet secondary structure is associated with a minimum near 215 nm.[32] Note that 2D NMR structural analysis provides more detailed and reliable information regarding β-sheet secondary structure than does CD, but CD can be a useful complementary technique to quickly gain information for a large range of experimental conditions. It should also be noted that CD analysis is considered more reliable for α-helical peptides than for β-sheet peptides as the CD signature of α-helical peptides is more intense.[32] Due to the lack of sensitivity of the CD signal to β-sheet structure, NMR is a more sensitive method for measuring differences in stability. Nevertheless, CD of β-sheet peptides can probe whether pH and temperature have a large effect on structure over a range of temperatures and conditions.

CD spectra were taken for parent peptide 2.10, peptide 3.7 (Lys9→TS^+), and peptide 3.14 (Glu4→TS^-) for a variety of conditions (Figure 3.14). Under the conditions used for NMR studies (pH 3.8, 4ºC), each peptide displays a minimum near 215 nm, as expected if β-sheet structure has formed. Changing from pH 3.8 to pH 7.0 did not have a large effect on the overall structure. Increasing the temperature from 4ºC to 37ºC did not

have a large impact on structure. In phosphate buffer at pH 7.0, the signal of the three peptides are indistinguishable (Figure 3.15).

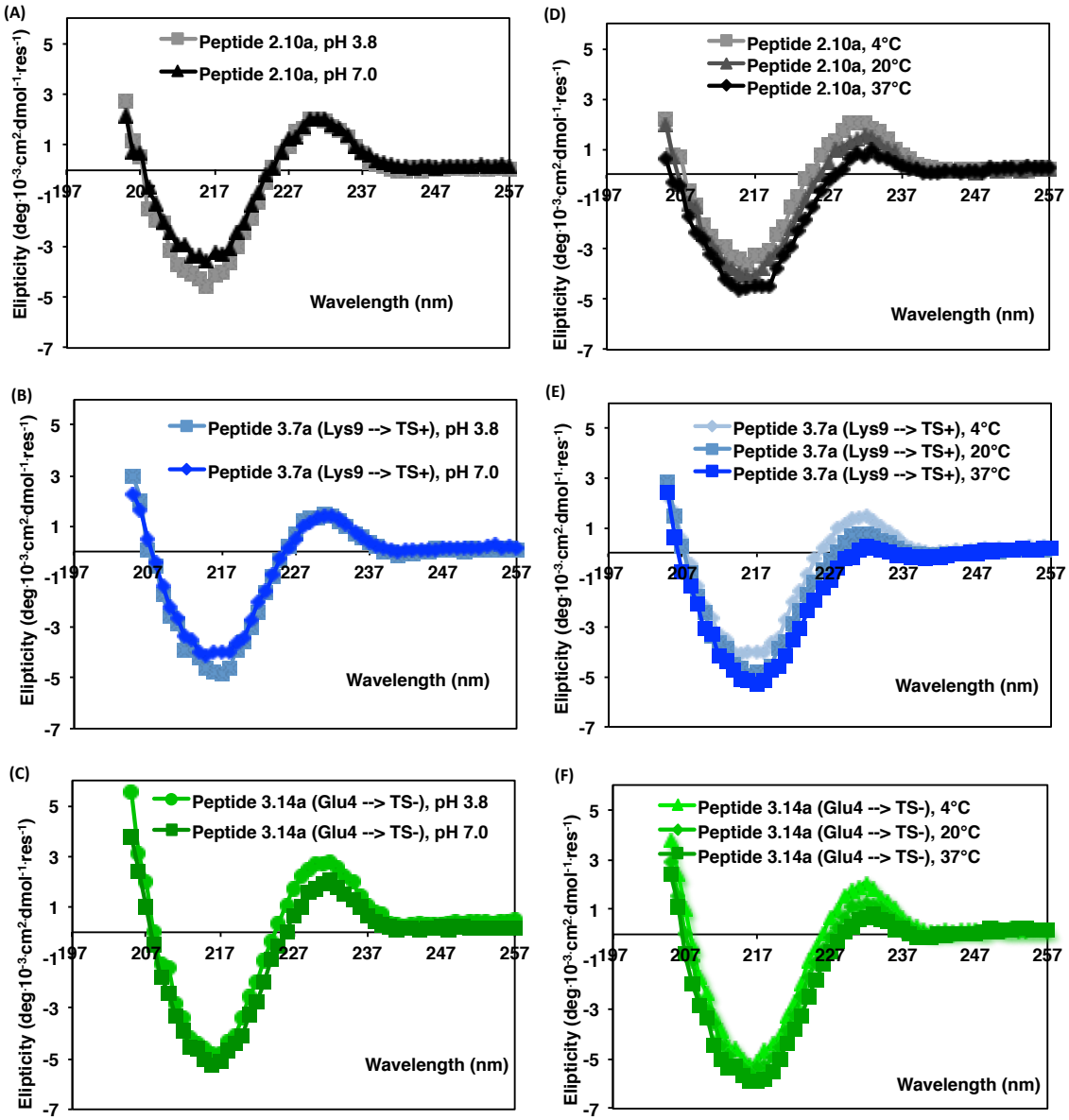


Figure 3.14. CD spectra of peptides 2.10a, 3.7a, and 3.14a at (A-C) 4ºC in phosphate buffer at pH 7.0 or acetate buffer at pH 3.8. These spectra compare the structure of the peptide at physiologically relevant pH 7.0 to the conditions used for the NMR analysis. (D-F) CD spectra of peptides 2.10a, 3.7a, and 3.14a in phosphate buffer at pH 7.0 at 4, 20 or 37ºC. The minima near 215 nm indicate the peptides adopt similar structures at the temperature used for NMR analysis (4ºC) and physiologically relevant temperature (37ºC).

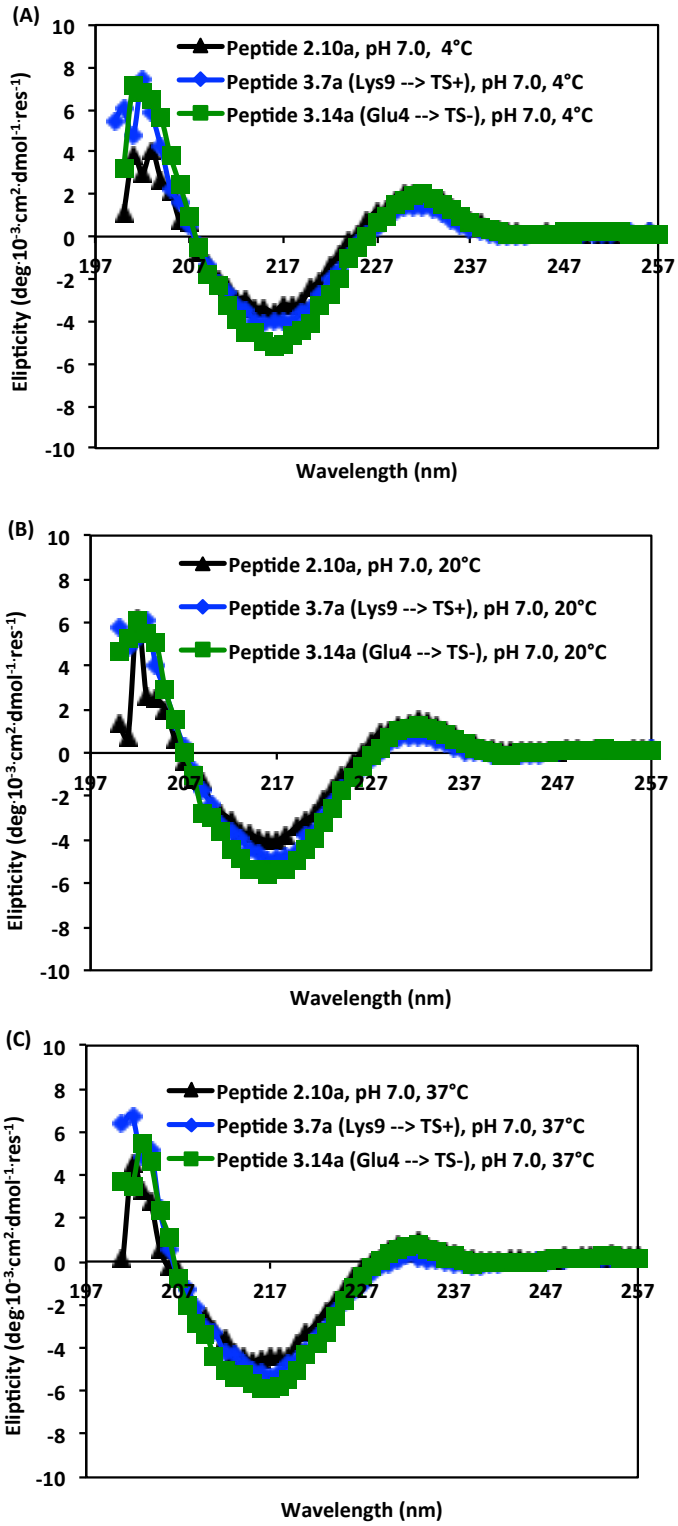


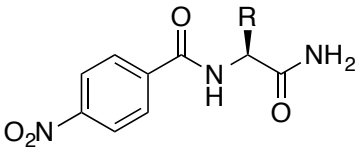
Figure 3.15. CD of peptide 2.10a, 3.7a, and 3.14a in phosphate buffer at (A) 4°C, (B) 20°C, and (C) 37°C.

3.3.4 Evaluating the hydrophilicity of TS^+ and TS^- relative to natural residues

Hydrophilic residues should increase the aqueous solubility of a peptide relative to hydrophobic residues. In order to determine the hydrophilicity of TS^+ and TS^- relative to natural amino acids, octanol/water distribution coefficients were measured for 4-nitrobenzoyl derivatives of the amino acid residues (see Table 3.2 for the amino acid structure).[33,34] The distribution coefficient is a measure of the relative concentration of a compound in the aqueous layer and the octanol layer of a mixture. The distribution coefficient (Π) is the log of the ratio of the concentration in the octanol layer to the concentration in the aqueous layer. The more hydrophobic the compound, the higher the concentration in the octanol layer, and the higher the distribution coefficient.[33,34] The 4-nitrobenzoyl derivatives were used to measure relative distribution coefficients as the UV absorbance allows the quantification of the relative concentration of amino acid derivatives (Table 3.2).[33,34]

Table 3.2 Selected octanol/water distribution coefficients (Π) of 4-nitrobenzoyl amino acid derivatives.

Amino Acid	Π
Lys	-2.44 ± 0.02
TS^+	-1.47 ± 0.04
Glu	-1.85 ± 0.4
TS^-	-1.75 ± 0.2
Gly	0
Thr	0.07 ± 0.01
Ile	1.59 ± 0.08



The distribution coefficients of both Thr and Ile are at least an order of magnitude larger than those of TS^+ and TS^-. Therefore, in the context of a peptide, one would expect TS^+ and TS^- to help solubilize the peptide relative to the natural β-branched residues, because TS^+ and TS^- are more hydrophilic than Thr or Ile.

3.4 Conclusions and future directions

TS^+ and TS^- are both able to promote β-sheet structure relative to natural, charged amino acids. We have successfully synthesized amino acids that can promote β-sheet structure without sacrificing the solubility of a peptide.

In the synthesis of the thioester threonine amino acids, the side-chain is functionalized in the second-to-last step of the synthesis. Therefore, this approach may be generalizable to create a variety of amino acids with high β-sheet propensity.

3.5 Methods

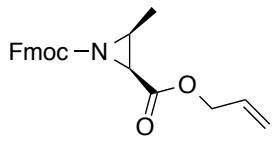
3.5.1 Amino Acid Synthesis

General. All chemicals were purchased from Sigma-Aldrich, Chemimpex, or Fluka Analytical. All NMR spectra were taken on a Bruker AC+300a (300 mHz), Bruker Avance-400 (400 mHz), Bruker Avance-500 with a DCH cryoprobe (500 mHz), Varian INOVA 600 (600 mHz), or Varian MercuryPlus 300 (300 mHz). All optical rotations were taken on a Rudolph Autopol III automatic polarimeter with a 1 mL cell and a 5 cm path length. NMR spectra can be found in Appendix B. The methods for synthesizing TO^+ can be found in chapter 2.

2-(*tert*-Butylmethoxycarbonylamino)ethyl bromide (3.4a)

Synthesized as previously described in Sawai *et. al.*[35]

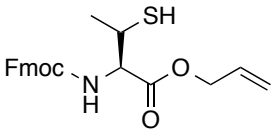
(2*S*, 3*S*)-Allyl-1-(9-fluorenylmethoxycarbonyl)-3-methylaziridine-2-carboxylate (3.1)



The synthesis was based on methods from Vederas *et al.*[20] Trityl protected azirdine **2.2** (65.5 g, 170.7 mmol) was dissolved in dichloromethane (100 ml) and methanol (100 ml). The solution was cooled to 0ºC. Trifluoroacetic acid was added dropwise over 10 minutes. The solution was stirred for 10 minutes. The solvent was removed under reduced pressure. The resulting brown oil was dissolved in ether (500 ml) and water (500 ml). The organic layer was extracted with water (3 x 100 ml). The

combined water layers were cooled to 0ºC. Sodium bicarbonate was added until the pH was basic. Ethyl acetate (1000 ml) and Fmoc-Cl (55.2 g) were added. The solution was stirred vigorously at room temperature for 20 hours. The organic and aqueous layers were separated, and the aqueous layer was extracted with ethyl acetate (3 x 200 ml). The organic layer was washed with brine (3 x 500 ml), dried with magnesium sulfate, and filtered. The solvent was removed under reduced pressure. The resulting brown oil was purified by chromatography on silica gel (10% ethyl acetate in hexanes). The product was a clear oil (22.1 g, 51% yield). ESI-MS: Calc. $[C_{22}H_{22}NO_4]^+$=364.1544, Obsd. $[C_{22}H_{22}NO_4]^+$=364.1555; ^{1}H NMR ($CDCl_3$, 500 mHz): δ 7.74 (d, *J*= 7.5 Hz, 2H), 7.57 (d, *J* = 7.5 Hz, 2H), 7.38 (t, *J* = 7.5 Hz, 2H), 7.30 (t, *J* = 7.5 Hz, 2H), 5.93 (ddt, *J* = 16.3, 10.6, 5.8 Hz, 1H), 5.36 (dq, *J*=17.1, 1.3 Hz, 1H), 5.27 (dq, *J*=10.4, 1.0 Hz, 1 H), 4.74 – 4.61 (m, 2H), 4.41 (d, *J* = 7.0 Hz, 2H), 4.21 (t, *J* = 7.0 Hz, 1H), 3.09 (d, *J* = 6.7 Hz, 1H), 2.68 (p, *J* = 5.7 Hz, 1H), 1.36 (d, *J* = 5.7 Hz, 3H); ^{13}C-NMR ($CDCl_3$, 125 mHz): δ 166.88, 161.62, 143.43, 141.36, 131.55, 127.92, 127.17, 127.18, 125.13, 125.17, 120.04, 120.06, 119.12, 68.47, 66.12, 46.90, 39.83, 39.07, 12.93.

(2*R*, 3*R*) Allyl-*N*-(9-fluorenylmethoxycarbonyl)-3-methylcysteine carboxylate (3.2):

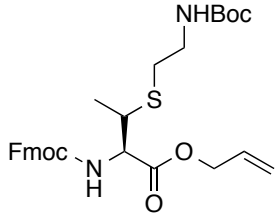


The synthesis was based on methods from Narayan and VanNieuwenhze.[18] Aziridine **3.1** (13.8 g, 38.0 mmol, 1 eq) was dissolved in dry dichloromethane (300 ml), and the solution was cooled to 0ºC. Boron trifluoride diethyl etherate (9.9 ml, 80.2 mmol, 2.1 eq)

was dissolved in dichloromethane (0ºC, 60 ml) under N_2. Triphenylmethanethiol (37 g, 133.9 mmol, 3.5 eq) was added to the aziridine solution. The boron trifluoride solution was added dropwise to the aziridine solution over 5 minutes. The mixture was stirred at 0ºC for 2.5 hours. The reaction was quenched with saturated aqueous sodium bicarbonate and stirred for an hour. The aqueous and organic layers were separated. The aqueous layer was extracted with dichloromethane. The combined organic layers were dried with magnesium sulfate and filtered. The solvent was removed under reduced pressure. The crude mixture was purified by chromatography on silica gel (5% ethyl acetate and 10% dichloromethane in hexanes). A crude yellow oil (11.9 g) was obtained. The oil was dissolved in triisopropylsilane (12 ml), dichloromethane (300 ml), and trifluoroacetic acid (30 ml). The solution was stirred for 20.5 hours. The solvent was removed under reduced pressure, and the product was purified by chromatography (5% dichloromethane, 5% ethyl acetate in hexanes). A white solid was obtained (11.3 g, 28.4 mmol, 74% yield). Two rotomers were observed in the ^{1}H NMR at room temperature. Peaks begin to coalesce at higher temperatures (45ºC). A COSY was taken, and the observed cross-peaks are consistent with the expected regioisomer. (Note that there is some concern that the thiol nucleophile could open the aziridine at the C2 carbon instead of the C3 carbon. The COSY experiment checks that the correct regioisomer has been isolated.) ESI-MS: Calc. $[C_{22}H_{24}NO_4S]^+$=398.1421, Obsd. $[C_{22}H_{24}NO_4S]^+$=398.1419; ^{1}H NMR ($CDCl_3$, 500 mHz) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.66-7.58 (m, 2H), 7.46-7.36 (m, 4H), 5.91 (ddt, *J* = 16.3, 11.0, 5.7 Hz, 1H), 5.50 (d, *J* = 9.6 Hz, 1H), 5.40 – 5.12 (m, 3H), 4.70 (d, *J* = 5.6 Hz, 2H), 4.62 – 4.37 (m, 2H), 4.25 (t, *J* = 7.1 Hz, 1H), 1.39 (2d rotomers, *J*=6.3 Hz, 3H).^{13}C-NMR ($CDCl_3$, 125 mHz): δ 171.17, 169.14, 169.12, 156.77, 156.60, 143.81, 143.58,

141.53 , 141.34, 141.32, 131.24, 127.78, 127.12, 125.12, 125.10, 124.74, 120.04, 120.03, 119.12, 90.78, 89.40, 67.41, 66.54, 60.42, 58.10, 57.94, 47.13, 17.41, 17.24, 14.22.

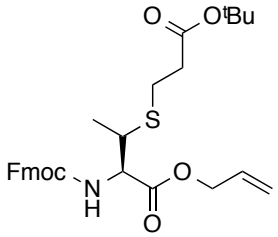
(2*R*, 3*R*) Allyl-2-(9-fluorenylmethoxycarbonylamino)-3-(2-*tert*-butyloxycarbonylaminoethylthio)butane carboxylate (3.5a):



Thiol **3.2** (0.55 g, 1.4 mmol, 1eq) was dissolved in ethyl acetate and saturated aqueous sodium bicarbonate (24 ml, 1:1). Tetrabutylammonium bisulfate (1.7 g) and excess solid sodium bicarbonate were added. Alkyl bromide **3.4a** (0.6g, 2.8 mmol, 2eq) was added. The mixture was stirred vigorously for 16 hours. The layers were separated, and the aqueous layer was extracted with ethyl acetate. The organic layers were dried with magnesium sulfate and filtered. The solvent was removed under reduced pressure. The crude product was purified by chromatography on silica gel (10-20% ethyl acetate in hexanes). The product was a clear oil (0.28 g, 38% yield). EMM-MS: Calc. $[C_{29}H_{37}N_2O_6S]^+$=541.2367, Obsd. $[C_{29}H_{37}N_2O_6S]^+$= 541.2371; ^{1}H NMR ($CDCl_3$, 500 mHz): δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.62 (d, *J* = 7.1 Hz, 2H), 7.41 (t, *J* = 7.3 Hz, 2H), 7.32 (t, *J* = 7.2 Hz, 2H), 5.94 (ddt, *J* = 16.5, 11.2, 5.9 Hz, 1H), 5.62 (d, *J* = 8.9 Hz, 1H), 5.38 (d, *J* = 17.2 Hz, 1H), 5.29 (d, *J* = 10.4 Hz, 1H), 4.88 (s, 1H), 4.68 (d, *J* = 5.6 Hz, 2H), 4.60 (dd, *J* = 9.2, 3.1 Hz, 1H), 4.46-4.33 (m, 2H), 4.25 (t, *J* = 7.1 Hz, 1H), 3.47-3.35 (m, 1H), 3.35-3.20 (m, 2H), 2.64 (ddt, *J* = 19.7, 13.1, 6.6 Hz, 2H), 1.45 (s, 9H), 1.37 (d, *J* = 7.0 Hz, 3H). ^{13}C-NMR ($CDCl_3$, 125 mHz): δ 170.53, 156.62, 155.92, 144.05, 143.87,

141.52 , 131.49 , 127.95 , 127.31 , 125.35 , 120.22, 120.21, 119.76 , 79.77, 67.52, 66.66, 58.80, 47.35, 42.91, 40.12, 31.95, 28.60, 19.94.

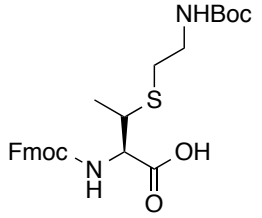
(2*R*, 3*R*) Allyl-2-(9-fluorenylmethoxycarbonylamino)-3-(2-*tert*-butyloxycarbonylpropylthio)butane carboxylate (3.5b):



Thiol **3.2** (0.55 g, 1.38 mmol) was dissolved in EtOAc:sat. aq. $NaHCO_3$ (1:1). Alkyl bromide **3.4b** (0.46 ml, 2.76 mmol) was added. Tetrabutylammonium bisulfate (1.7 g) was added. The mixture was stirred vigorously for 16 hours. The aqueous and organic layers were separated, and the aqueous layer was extracted with ethyl acetate. Combined organic layers were dried with magnesium sulfate and filtered. Solvent was removed under reduced pressure. Crude product was purified by chromatography on silica gel (10% ethyl acetate in hexanes). Note that the starting material has a similar R_F to the product. While the two compounds could not be resolved by TLC, a careful column allows separation. Careful analysis of the integration of the peaks in the NMR is necessary to determine the success of the column as all the peaks in the starting material overlay with peaks in the product. The product was a clear oil (0.3 g, 0.571 mmol, 41% yield). ESI-MS: Calc. $[C_{29}H_{36}NO_6S]^+$=526.2258, Obsd. $[C_{29}H_{36}NO_6S]^+$=526.2242; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.68 – 7.58 (m, 2H), 7.36 (dt, *J* = 41.8, 7.4 Hz, 4H), 5.94 (ddt, *J* = 16.5, 10.9, 5.9 Hz, 1H), 5.70 (d, *J* = 9.3 Hz, 1H), 5.38 (d, *J*=16.5 Hz, 1 H), 5.28 (d, *J* = 10.3 Hz, 1H), 4.74-4.61 (m, 1H), 4.59 (dd, *J* = 9.3, 3.1 Hz, 1H), 4.46 – 4.35 (m, 2H), 4.26 (t, *J* = 7.3 Hz, 1H), 3.47 (qd, *J* = 7.0, 3.2 Hz, 1H), 2.81-

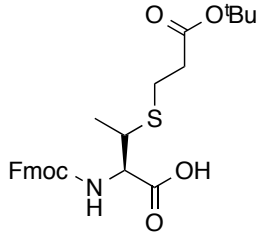
2.61 (m, 2H), 2.49 (t, *J* = 7.2 Hz, 2H), 1.46 (s, 9H), 1.37 (d, *J* = 7.1 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 141.30, 131.34, 127.71, 127.09, 125.18, 119.98, 119.36, 81.04, 67.30, 66.37, 58.64, 47.14, 42.90, 35.95, 28.09, 26.49, 19.70.

(2*R*, 3*R*) 2-(9-fluorenylmethoxycarbonylamino)-3-(2-*tert*-butyloxycarbonylaminoethylthio)butane carboxylic acid (3.6a):



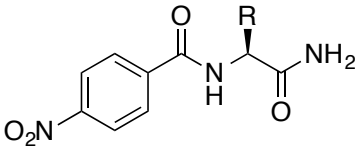
Thioether **3.5a** (0.28 g, 0.52 mmol) was dissolved in methanol (20 ml). Ammonium formate (2 g) was added, and the solution was stirred under N_2 for 10 minutes. Palladium on carbon (0.3 g, 10% loading by weight, wet support) was added, and the mixture was stirred overnight. The reaction was filtered through celite and the solvent was removed under reduced pressure. The product was a white solid (0.3 g, quantitative yield). ESI-MS: Calc. $[C_{26}H_{33}N_2O_6S]^+$=501.2054, Obsd. $[C_{26}H_{33}N_2O_6S]^+$=501.2063; Melting Point: 112.8-116.6ºC; ^{1}H NMR (CD_3OD, 500 mHz): δ 7.70 (d, *J* = 7.6 Hz, 2H), 7.64 – 7.55 (m, 2H), 7.29 (t, *J* = 7.5 Hz, 2H), 7.21 (t, *J* = 7.5 Hz, 2H), 4.32-4.24 (m, 2H), 4.16 (t, *J* = 7.0 Hz, 1H), 3.29-3.21 (m, 2H), 3.17-3.07 (m, 2H), 2.60-2.51 (m, 2H), 1.33 (s, 9H), 1.24 (d, *J* = 7.0 Hz, 3H). ^{13}C NMR (CD_3OD , 125 mHz): δ 172.37, 157.31, 143.91, 143.72, 141.17, 127.38, 126.78, 124.90, 119.50, 78.75, 66.76, 58.87, 42.25, 40.05, 30.84, 27.35, 18.73.

(2*R*, 3*R*) 2-(9-fluorenylmethoxycarbonylamino)-3-(2-*tert*-butyloxycarbonylethylthio)butane carboxylic acid (3.6b):



Thioether **3.5b** (0.61 g, 1.2 mmol) was dissolved in methanol (100 ml). Ammonium formate (2 g) was added, and the solution was stirred under N_2 for 10 minutes. Palladium on carbon (0.3 g , 10% loading by weight, wet support) was added, and the mixture was stirred overnight. The mixture was filtered through celite, and the solvent was removed under reduced pressure. The product was a white solid (0.36 g, 0.75 mmol, 65% yield). ESI-MS: Calc. $[C_{26}H_{31}NO_6SNH_4]^+$=503.2211, Obsd. $[C_{26}H_{31}NO_6SNH_4]^+$=503.2200; $[\alpha]_D^{25}$=+18 (*c* 0.5, MeOH); Melting Point: 131.5-133.1ºC; ^{1}H NMR (CD_3OD, 500 mHz): δ 7.80 (d, *J* = 7.5 Hz, 2H), 7.72 – 7.66 (m, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.5 Hz, 2H), 4.45 – 4.30 (m, 3H), 4.25 (t, *J* = 6.9 Hz, 1H), 3.42 – 3.33 (m, 1H), 2.85-2.72 (m, 2H), 2.51 (t, *J* = 7.0 Hz, 2H), 1.45 (s, 9H), 1.32 (d, *J* = 7.1 Hz, 3H). ^{13}C NMR (CD_3OD, 125 mHz): δ 173.74, 173.05, 158.72, 145.32, 145.14, 142.60, 128.79, 128.19, 126.33, 126.30, 120.93, 82.04, 68.14, 60.19, 43.79, 37.07, 28.34, 27.53, 20.04, 18.79.

N-4-Nitrobenzoyl carboxamides:



Amino acid derivatives were synthesized by standard Fmoc solid-phase peptide synthesis. For amino acid derivatives of Ile, Thr, Gly, Glu, Lys, and TS^-, protected amino

acids were couple to Chemmatrix rink amide resin on a 100 μmol scale with 4 eq PyBOP ((benzotriazol-1-yloxy)tripyrrolidinophosphoniumhexafluorophosphate), 4 eq Cl-HOBt (1-hydroxy-6-chloro-benzotriazole), and 8 equivalents diisopropylethylamine (DIEA) in 4 ml DMF for 40 minutes. The resin was rinsed with DMF (3x10 ml), and then the resin was stirred in 20% piperidine in DMF (4 ml) for 15 minutes. The resin was rinsed with DMF (3x10 ml). 4 eq 4-nitrobenzoyl chloride and 8 eq diisopropylethylamine in DMF were added to the resin and the mixture was stirred for 1 hour. The resin was rinsed with DMF (3x10 ml) and DCM (3x10 ml). The amino acid residue was cleaved from the resin by stirring the resin in a solution of 8 ml trifluoroacetic acid (TFA), 200 μl water, and 200 μl triisopropylsilane for 1 hour. The residue was precipitated from the resulting solution with water and acetonitrile and purified on a Shimadzu reverse-phase semiprep HPLC with a gradient acetonitrile and water. Fraction with product were combined, and the solvent was removed under reduced pressure. For the TS^+ derivative, the synthesis was done as described above with the following modifications. The synthesis was done on a 5 μmol scale and only 1 eq of monomer, 1 eq PyBOP, 1 eq Cl-HOBt, and 2 eq DIEA in 1 ml DMF were used. The coupling was done on a CEM microwave for 12 minutes at 70ºC. All amino acid derivatives were characterized by NMR and ESI-MS.

4-Nitrobenzoyl isoleucine carboxamide. ESI-MS: Calc. $[C_{13}H_{17}N_3O_4Na\ ]^+$=302.1112, Obsd. $[C_{13}H_{17}N_3O_4Na]^+$=302.1102; ^{1}H NMR (CD_3OD, 500 mHz): δ 8.24 (d, J = 8.9 Hz, 2H), 7.96 (d, J = 8.9 Hz, 2H), 4.37 (d, J = 8.1 Hz, 1H), 1.99 – 1.79 (m, 1H), 1.59-1.51 (multiplet, 1H), 1.26 – 1.11 (m, 1H), 0.94 (d, 6.8 Hz, 3H), 0.87 (t, J = 7.4 Hz, 3H); ^{13}C

NMR (CD_3OD, 125 mHz): 176.28, 168.50, 151.25, 141.42, 130.08, 124.74, 59.87, 37.98, 26.36, 16.12, 11.46.

4-Nitrobenzoyl threonine carboxamide. ESI-MS: Calc. $[C_{11}H_{13}N_3O_5Na\ ]^+$=290.0748, Obsd. $[C_{11}H_{13}N_3O_5Na]^+$=290.0755;^{1}H NMR (500 MHz, CD_3OD): δ 8.35 (d, *J* = 8.9 Hz, 1H), 8.11 (d, *J* = 8.9 Hz, 1H), 4.57 (d, *J* = 4.0 Hz, 1H), 4.29 (qd, *J* = 6.4, 4.1 Hz, 1H), 1.26 (d, *J* = 6.4 Hz, 3H); ^{13}C NMR (126 MHz, MeOD): δ 175.15, 168.47, 151.36, 141.12, 130.11, 124.79, 68.67, 60.75, 20.53.

4-Nitrobenzoyl glycine carboxamide. ESI-MS: Calc. $[C_9H_9N_3O_4Na]^+$=246.0486, Obsd. $[C_9H_9N_3O_4Na]^+$=246.0478; ^{1}H NMR (400 MHz, D_2O) δ 8.30 (d, *J* = 8.8 Hz, 2H), 8.04 (d, *J* = 9.2 Hz, 2H), 3.70 (s, 2H). (Insufficient material was obtained for ^{13}C-NMR analysis to be possible.)

4-Nitrobenzoyl glutamic acid carboxamide. ESI-MS: Calc. $[C_{12}H_{13}N_3O_6Na]^+$=318.0697, Obsd. $[C_{12}H_{13}N_3O_6Na]^+$=318.0696; ^{1}H NMR (500 MHz, Methanol-d4) δ 8.24 (d, *J* = 8.8 Hz, 2H), 8.00 (d, *J* = 8.9 Hz, 2H), 4.59 – 4.42 (m, 1H), 2.40 (t, *J* = 7.6 Hz, 2H), 2.26 – 2.07 (m, 1H), 2.05-1.97 (m, *J* = 14.3, 9.2, 7.0 Hz, 1H); ^{13}C NMR (126 MHz, Methanol-d4) δ 176.78, 176.36, 168.41, 151.31, 141.10, 130.13, 124.72, 54.92, 31.60, 28.39.

4-Nitrobenzoyl thioether anionic threonine carboxamide. ESI-MS: Calc. $[C_{14}H_{17}N_3O_6SNa\ ]^+$=378.0731, Obsd. $[C_{14}H_{17}N_3O_6SNa]^+$=378.0734; ^{1}H NMR (500 MHz,

Methanol-d4): δ 8.26 (d, *J* = 8.9 Hz, 2H), 7.99 (d, *J* = 8.9 Hz, 2H), 4.62 (d, *J* = 6.9 Hz, 1H), 3.30 (quint, *J* = 7.0 Hz, 1H), 2.87 – 2.71 (m, 2H), 2.62 – 2.41 (m, 2H), 1.32 (d, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, MeOD): δ 175.71, 174.57, 168.25, 141.21, 130.14, 124.79, 59.06, 43.75, 35.71, 26.72, 19.62.

4-Nitrobenzoyl lysine carboxamide: ESI-MS: Calc. $[C_{13}H_{19}N_4O_4]^+$=295.1401, Obsd. $[C_{13}H_{19}N_4O_4]^+$=295.1396; ^{1}H NMR (500 MHz, Methanol-d4) δ 8.25 (d, *J* = 8.9 Hz, 2H), 8.00 (d, *J* = 8.9 Hz, 2H), 4.51 (dd, *J* = 9.0, 5.4 Hz, 1H), 2.87-2.84 (m, 2H), 1.97 – 1.84 (m, 1H), 1.77 (dtd, *J* = 14.1, 9.4, 5.2 Hz, 1H), 1.65 (tt, *J* = 14.7, 6.5 Hz, 2H), 1.56 – 1.31 (m, 2H); ^{13}C NMR (126 MHz, MeOD) δ 176.64, 168.42, 151.34, 141.07, 130.12, 124.74, 54.99, 40.68, 32.64, 28.30, 24.19.

4-Nitrobenzoyl thioether cationic threonine carboxamide: ESI-MS: Calc. $[C_{13}H_{19}N_4O_4S\ ]^+$=327.1122, Obsd. $[C_{13}H_{19}N_4O_4S]^+$=327.1121; ^{1}H NMR (500 MHz, Deuterium Oxide) δ 8.37 (d, *J* = 9.2 Hz, 2H), 8.00 (d, *J* = 8.4 Hz, 2H), 3.24 (t, *J* = 6.9 Hz, 1H), 3.04 – 2.86 (m, 2H), 3.54 – 3.46 (m, 1H), 1.43 (d, *J* = 6.6 Hz, 3H).

3.5.2. Circular Dichroism (CD)

Peptides were dissolved in phosphate buffer (100 mM, pH=7.0), and an aliquot was used to determine the concentration by UV in 6 M guanidinium hydrochloride using the extinction coefficient of Tyr. CD samples were prepared by serial dilutions of the stock sample. All CD spectra were taken on an Aviv Model 420 Circular Dichroism Spectrometer with 1 mm pathlength cells.

3.5.3. Distribution Coefficient Studies

The 4-nitrobenzoyl carboxamide amino acid derivatives of Ile, Thr, Gly, Glu, Lys, TS^+, and TS^- were synthesized as described in the previous section. The hydrophobicities of TS^+ and TS^- were compared to those of natural amino acids by measuring the distribution cofficients of the 4-nitrobenzoyl amino acid derivatives as previously described.[33,34] Briefly, the amino acids were dissolved in equal volumes of octanol and phosphate buffer (pH 7.0, 100 mM) and rocked on a rocker for at least 24 hours. The phases were separated and the relative concentration of amino acid in each phase was determined by analytical HPLC monitored at 275 nm. At least three replicates were measured for each residue. The normalized distribution coefficient, Π, was calculated from the equation $\Pi=\log(D_{\text{amino acid}})-\log(D_{\text{glycine}})$, D is the ratio of the concentration of the residue in octanol layer and aqueous layer.

3.5.4. NMR Sample Preparation and Data Acquisition

NMR samples were prepared and data was acquired as described in the methods section of Chapter 2.

3.5.5. Dilution Studies

In order to check for aggregation, dilution studies were undertaken for all peptides in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8) at 4ºC. The most concentrated samples were at the concentration used for acquisition of the 2D NMR spectra. The most dilute sample was at lest 50 fold less concentrated than the NMR sample. When necessary, baselines were corrected by manually defining baseline regions.

Water suppression was performed with a presaturation pulse sequence as described in Chapter 2. The spectra of the concentrated and diluted peptide were the same. Therefore, the peptide aggregation state did not change over this concentration range. The concentration ranges are specified in the dilution spectra found in Appendix B. The most concentrated sample is generally around 5 mM. In addition, all parent peptides were previously characterized by sedimentation equilibrium analytical ultracentrifugation experiments and found to be monomeric.[24,31,29,36] Therefore, we can assume that each peptide is monomeric under the conditions of the NMR experiments.

3.5.6. Peptide Synthesis.

Peptides were synthesized and purified as described in Chapter 2. HPLC traces of pure peptides can be found in Appendix B.

3.6 References

(1) Swindells, M. B.; MacArthur, M. W.; Thornton, J. M. Intrinsic Phi, Psi Propensities of Amino Acids, Derived from the Coil Regions of Known Structures. *Nat. Struct. Biol.* **1995**, *2*, 596–603.

(2) Chou, P. Y.; Fasman, G. D. Conformational Parameters for Amino Acids in Helical, Beta-Sheet, and Random Coil Regions Calculated from Proteins. *Biochemistry* **1974**, *13*, 211–222.

(3) Chou, P. Y.; Fasman, G. D. Secondary Structural Prediction of Proteins from Their Amino Acid Sequence. *Trends Biochem. Sci.* **1977**, *2*, 128–131.

(4) Minor, D. L.; Kim, P. S. Measurement of the Beta-Sheet-Forming Propensities of Amino Acids. *Nature* **1994**, *367*, 660–663.

(5) Kim, C. A.; Berg, J. M. Thermodynamic Beta-Sheet Propensities Measured Using a Zinc-Finger Host Peptide. *Nature* **1993**, *362*, 267–270.

(6) Smith, C. K.; Withka, J. M.; Regan, L. A Thermodynamic Scale for the Beta-Sheet Forming Tendencies of the Amino Acids. *Biochemistry* **1994**, *33*, 5510–5517.

(7) Minor, D. L.; Kim, P. S. Context Is a Major Determinant of Beta-Sheet Propensity. *Nature*, 1994, *371*, 264–267.

(8) Muñoz, V.; Serrano, L. Intrinsic Secondary Structure Propensities of the Amino Acids, Using Statistical Phi-Psi Matrices: Comparison with Experimental Scales. *Proteins* **1994**, *20*, 301–311.

(9) Bai, Y.; Englander, S. W. Hydrogen Bond Strength and Beta-Sheet Propensities: The Role of a Side Chain Blocking Effect. *Proteins* **1994**, *18*, 262–266.

(10) Stapley, B. J.; Doig, a J. Free Energies of Amino Acid Side-Chain Rotamers in Alpha-Helices, Beta-Sheets and Alpha-Helix N-Caps. *J. Mol. Biol.* **1997**, *272*, 456–464.

(11) Niwa, T. S.; Ogino, A. Multiple Regression Analysis of the Beta-Sheet Propensity of Amino Acids. *J. Mol. Struct. Theochem* **1997**, *419*, 155–160.

(12) Street, A. G.; Mayo, S. L. Intrinsic Beta-Sheet Propensities Result from van Der Waals Interactions between Side Chains and the Local Backbone. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9074–9076.

(13) Avbelj, F.; Baldwin, R. L. Role of Backbone Solvation in Determining Thermodynamic Beta Propensities of the Amino Acids. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1309–1313.

(14) Rossmeisl, J.; Kristensen, I.; Gregersen, M.; Jacobsen, K. W.; Nørskov, J. K. B-Sheet Preferences from First Principles. *J. Am. Chem. Soc.* **2003**, *125*, 16383–16386.

(15) Koh, E.; Kim, T.; Cho, H. S. Mean Curvature as a Major Determinant of Beta-Sheet Propensity. *Bioinformatics* **2006**, *22*, 297–302.

(16) Efimov, A. V. Standard Structures in Proteins. *Prog. Biophys. Mol. Biol.* **1993**, *60*, 201–239.

(17) Kebarle, P. Ion Thermochemistry and Solvation From Gas Phase Ion Equilibria. *Annu. Rev. Phys. Chem.* **1977**, *28*, 445–476.

(18) Narayan, R. S.; VanNieuwenhze, M. S. Versatile and Stereoselective Syntheses of Orthogonally Protected Beta-Methylcysteine and Beta-Methyllanthionine. *Org. Lett.* **2005**, *7*, 2655–2658.

(19) Carrillo, A. K.; Vannieuwenhze, M. S. Synthesis of the AviMeCys-Containing D-Ring of Mersacidin. *Org. Lett.* **2012**, *14*, 1034–1037.

(20) Liu, H.; Pattabiraman, V. R.; Vederas, J. C. Stereoselective Syntheses of 4-Oxa Diaminopimelic Acid and Its Protected Derivatives via Aziridine Ring Opening. *Org. Lett.* **2007**, *9*, 4211–4214.

(21) Liu, W.; Chan, A. S. H.; Liu, H.; Cochrane, S. a.; Vederas, J. C. Solid Supported Chemical Syntheses of Both Components of the Lantibiotic Lacticin 3147. *J. Am. Chem. Soc.* **2011**, *133*, 14216–14219.

(22) Tarrade, A.; Dauban, P.; Dodd, R. H. Enantiospecific Total Synthesis of (-) - Polyoxamic Acid Using 2, 3-Aziridino-Gamma-Lactone Methodology. *J. Org. Chem.* **2003**, 9521–9524.

(23) Tanner, D. Chiral Aziridines—Their Synthesis and Use in Stereoselective Transformations. *Angew. Chemie Int. Ed. English* **1994**, *33*, 599–619.

(24) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quanfification of Beta-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

(25) Wurthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(26) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry* **1992**, *31*, 1647–1651.

(27) Haque, T. S.; Little, J. C.; Gellman, S. H. "Mirror Image" Reverse Turns Promote Beta-Hairpin Formation. *J. Am. Chem. Soc.* **1994**, *116*, 4105–4106.

(28) Haque, T. S.; Little, J. C.; Gellman, S. H. Stereochemical Requirements for B-Hairpin Formation: Model Studies with Four-Residue Peptides and Depsipeptides. *J. Am. Chem. Soc.* **1996**, *118*, 6975–6985.

(29) Haque, T. S.; Gellman, S. H. Insights on Beta-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I ′ and Type II ′ Beta-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

(30) Sibanda, B. L.; Thornton, J. M. Beta-Hairpin Families in Globular Proteins. *Nature* **1985**, *316*, 170–174.

(31) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. Length-Dependent Stability and Strand Length Limits in Antiparallel Beta-Sheet Secondary Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12015–12020.

(32) Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751*, 119–139.

(33) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. Stabilizing and Destabilizing Effects of Phenylalanine --> F5-Phenylalanine Mutations on the Folding of a Small Protein. **2006**, *53706*, 15932–15933.

(34) Lee, K. H.; Lee, H. Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. Fluorous Effect in Proteins: De Novo Design and Characterization of a Four-Alpha-Helix Bundle Protein Containing Hexafluoroleucine. *Biochemistry* **2004**, *43*, 16277–16284.

(35) Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. Modified DNA Aptamer That Binds the (R)-Isomer of a Thalidomide Derivative with High Enantioselectivity. *J. Am. Chem. Soc.* **2007**, *129*, 1456–1464.

(36) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. Analysis of the Factors That Stabilize a Designed Two-Stranded Antiparallel Beta-Sheet. *Protein Sci.* **2002**, 1492–1505.

Chapter 4: Toward a Peptide Inhibitor of the Interaction between Fibrinogen and *Streptococcal pyogenes* M1 Protein

4.1. Introduction

Streptococcus pyogenes is a bacterium responsible for some cases of toxic shock syndrome and rheumatic fever.[1] The M1 protein on the surface of *Streptocuccus* is associated with the virulence of the bacterium.[1] The M1 protein is known to form a complex with fibrinogen that causes neutrophils to release heparin-binding protein (HBP). HBP is responsible for a problematic inflammatory response. The complex between fibrinogen and M1 also prevents the destruction of the bacteria by phagocytosis. Therefore, an inhibitor that could prevent the interaction between M1 and fibrinogen would potentially be a useful therapeutic.[2]

4.1.1. Structure of M1

The M1 protein is 483 residues in length and self-associates to form a long helical dimer (Figure 4.1).[2,3] Regions of the M1 protein are defined in Figure 4.1. Fibrinogen binds the protein in the B repeat region.[2] A crystal structure of a construct containing the A and B regions of M1 ($M1^{AB}$) shows that the dimer is not a canonical coiled coil, but has irregularities.[3] An extra residue between the A and B regions disrupts the helical heptad repeat and causes the ends of the dimer to splay. These ends interact in an antiparallel fashion with another $M1^{AB}$ dimer (Figure 4.1).[3] The splaying is thought to be an artifact of crystallization. However, the antiparallel interactions in the splayed region reveal the positions of the helix that can form the core of a coiled coil. These positions might in the core of the parallel coiled coil in solution (Figure 4.1).[3]

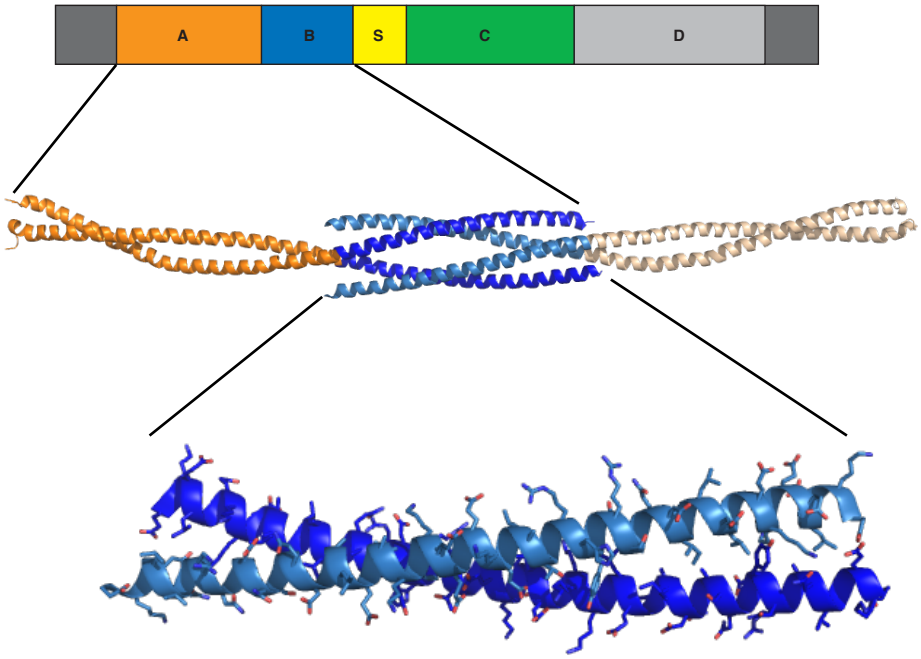


Figure 4.1. Scheme of the M1 protein and crystal structure of $M1^{AB}$ from McNamara *et al*. (pdb code 2OTO).[3]

A fragment of M1 that includes the B and S regions (M1*) was cocrystallized with a fibrinogen fragment D (FgD).[2] The crystal structure shows that each M1 binds four fibrinogen molecules (Figure 4.2).[2] Guided by the crystal structure, Ghosh *et al*. hypothesized that M1 and fibrinogen form a network they propose to be essential for neutrophil activation and ultimately causes toxic shock syndrome (Figure 4.2c).[2]

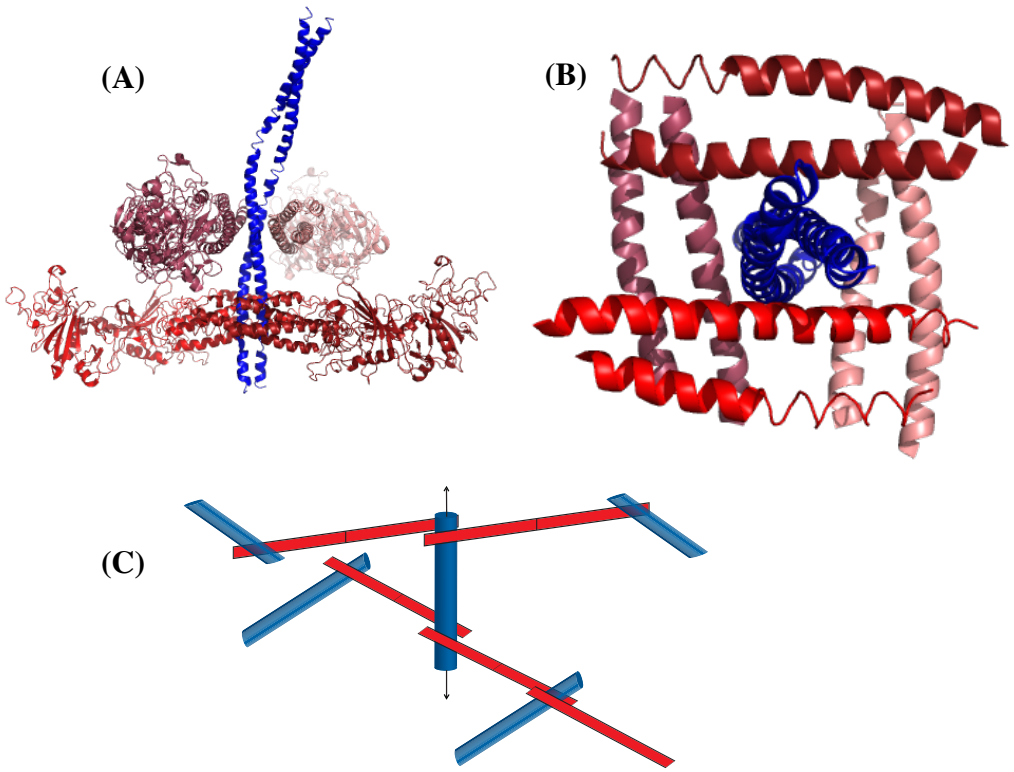


Figure 4.2. (A) Crystal structure from Ghosh *et al*.[2] of M1* bound to four copies of FgD. M1* is in blue and each copy of FgD is in a different shade of red (pdb code 2XNX). (B) Top view of the interface involving the four FgD proteins and M1* dimer. (C) Cartoon of the proposed network formed between fibrinogen and M1.[2]

In a canonical coiled-coil dimer, two positions of the heptad repeat, the a and the d positions, are in the core of the coiled-coil interaction. A helical wheel representation of a coiled coil shows the heptad repeat of the helices down the vertical axis (Figure 4.3). Hydrophobic residues at the a and d positions promote the formation of dimers or higher order oligomers.[4]

The M1 homodimer is unusual as it appears to have two possible modes of dimerization (Figure 4.3).[2] Register 1 was observed in the crystal structure of $M1^{AB}$ without the presence of FgD.[3] Register 2 differs from register 1 by the rotation of the helices about the vertical axis so that the a and a' positions in the core of the coiled coil

interaction in register 1 are adjacent to the core positions in register 2 (Figure 4.3). Register 2 was observed in the M1* and FgD crystal structure.[2] Evidence suggests there is an exchange between the two registers in solution, and register 2 binds fibrinogen.[2] Figure 4.3 is a helical wheel diagram that shows the relationship between the two registers. It should also be noted that in the absence of fibrinogen, there is exchange between labeled and unlabeled dimers of M1 at 37°C, suggesting that the dimer is labile.[2] In theory, disrupting the dimerization of M1 should prevent the formation of the fibrinogen M1 complex, and reduce the toxicity of a bacterial infection.[2]

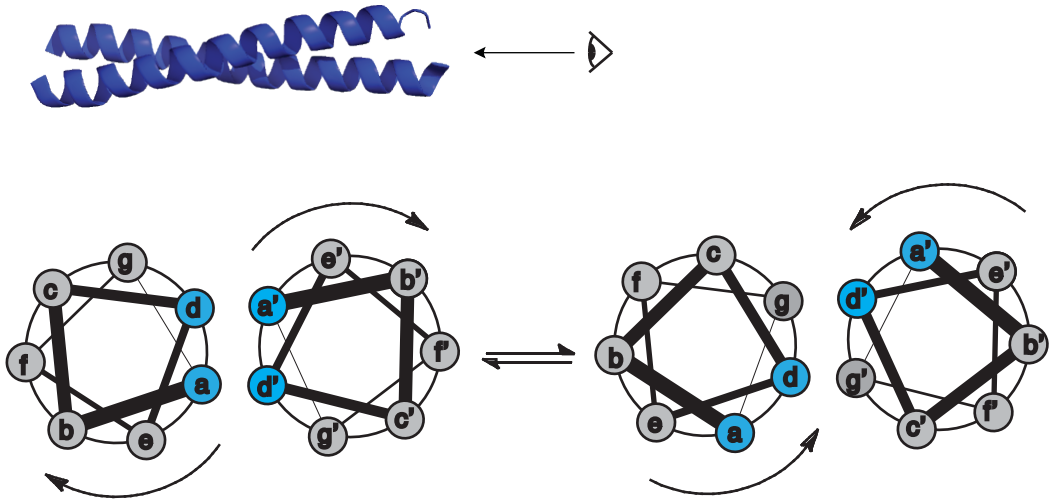


Figure 4.3. Helical wheel diagram of M1 showing the relationship between the two possible registers of dimerization.

4.2. Design and evaluation of M1 inhibitor candidates

Using rational design, I designed peptide sequences that were intended to bind the B repeats region of M1, block the dimerization of the protein, and prevent fibrinogen binding. Using detailed information about the specificities of different amino acids in coiled coils,[4–7] we analyzed the M1 sequence in the B repeats region for register 1 and register 2 (Figure 4.4). The helix net diagram in Figure 4.4 was used to visualize the

residues in the core of the coiled-coil homodimer. (For a more detailed explanation of helix net diagrams, see chapter 1.) In a coiled coil, residues in the core of the interactions pack in a 'knobs in holes fashion',[4] so that each residue from the first helix is surrounded by residues from the second helix. By convention, the a and d positions are in the core of the helix and the e and g positions flank the core positions. Residues from the first helix are at positions a-g and residues from the second helix are at positions a'-g'.

Note that both registers have unusual amino acids in the core of the coiled coil interface. Typically, hydrophobic residues form the interface, but in both registers, charged residues such as Glu, Asp, or Lys are found at the core positions (a and d). Also, in register 1, Tyr is found at a core position, which is unusual. The non-ideal nature of the interface is consistent with the labile nature of the M1 homodimer. Inhibitor candidates were designed to bind either register 1 or register 2.

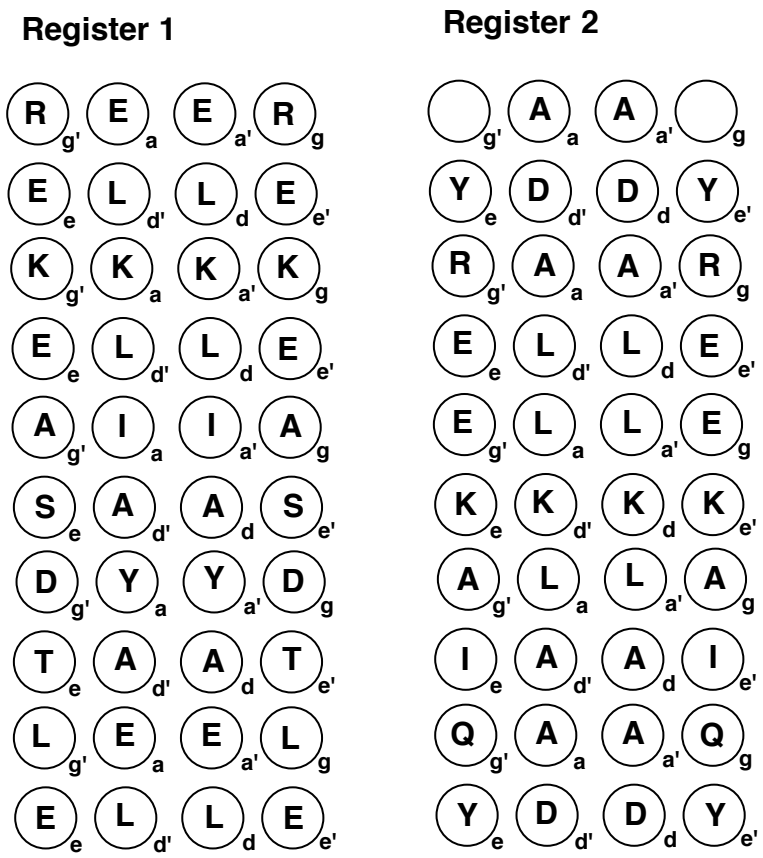


Figure 4.4. Helical net diagrams of register 1 residues 133-166 and register 2 residues 151-183. These were the segments inhibitor peptides were designed to bind.

4.2.1. Design and evaluation of register 1 M1 inhibitor candidates.

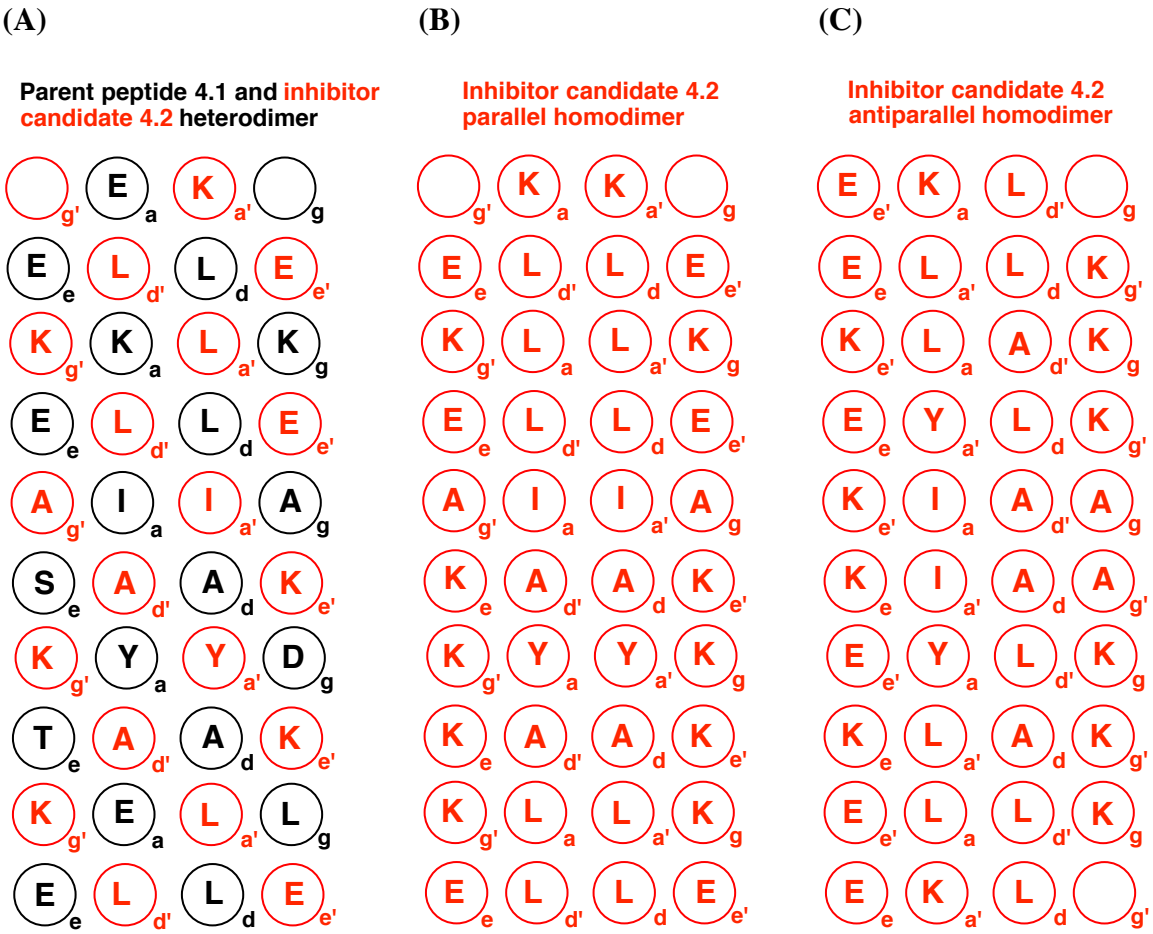
Ideally, an inhibitor should bind M1 more tightly than M1 binds itself. The black residues in Figure 4.5a show the surface presented by one helix of the M1 dimer in register 1 from residues 133-166. Inhibitor candidates were designed to bind this surface. Initially, the focus was on optimizing the interface to obtain the most ideal interface in the theoretical M1+inhibitor candidate heterodimer.

Inhibitor candidates were synthesized and tested against a fragment of M1 that contained the intended complementary sequence. For register 1 inhibitor candidates, the complementary sequence comprised residues 132-168 (parent peptide 3.1). Inhibitor candidates 4.2 and 4.3 were the initial designs (Figure 4.5). The helical net diagrams of

inhibitor candidate 4.2 is shown in Figure 4.5. Many of the residues in the core of inhibitor 4.2 are hydrophobic residues known to promote coiled-coil interactions. There is a Tyr in the core of the parent peptide. Little is known about the pairing preferences of Tyr. Peptide 4.2 was designed to pair the Tyr next to another Tyr and surround the bulky position with smaller Ala residues above and below to prevent steric clashes. Peptide 4.3 pairs the Tyr in the parent peptide with an Ala to prevent destabilizing steric clashes and places dimer-promoting Leu residues above and below the Tyr.

Self associatation of the inhibitor candidate was considered. If an inhibitor candidate strongly self-associates to form a homodimer that is more stable than the intended heterodimer between the inhibitor candidate and parent peptide, the candidate will not be an effective inhibitor. Elements of negative design were incorporated into the first round of inhibitor candidates to discourage self-association. The helical net digram (Figure 4.5) shows that the e and g positions of the inhibitor candidate homodimers have blocks of Lys, charges-charge interactions of which should destablize the homodimer.

Circular dichroism (CD) was used to assess the ability of the inhibitor candidates to bind the parent sequence. Helical peptides have a charateristic CD signature with minima at 208 and 222 nm. CD spectra of an inhibitor candidate and the parent peptide were compared to the CD signature of a 1:1 mixture of the two peptides. If the peptides do not interact, the signal of the measured mixture will be the average of the signal of the parent peptide and inhibitor candidate. If the inhibitor candidate binds the parent peptide as a helical dimer, the experimentally measured CD signature should be more helical than the calculated CD signal of the non-interacting mixture.



(D) RL**E**KE**L**EEK**K**EA**L**ELA**I**DQ**A**SRD**Y**HR**A**TAL**E**KE**L**EEK 4.1 (Parent)
KKE**L**EEK**L**EA**L**ELA**I**DQ**A**KRK**Y**HR**A**KAK**L**KE**L**EEK 4.2
KKE**L**EEK**L**EA**L**ELA**I**DQ**L**KRK**A**HR**L**KAK**L**KE**L**EEK 4.3

Figure 4.5. Helical net diagrams of (A) parent peptide 4.1 and inhibitor candidate 4.2 heterodimer, (B) inhibitor candidate 4.2 parallel homodimer, and (C) inhibitor candidate 4.2 antiparallel homodimer. (D) Sequences of parent peptide and inhibitor candidates. Core residues are bolded.

The CD spectrum of inhibitor 4.2 is shown in Figure 4.6. The parent peptide, shown in blue, does not have minima at 208 and 222 nm, which suggests it is not helical under the sample conditions. The inhibitor candidate 4.2, shown in red, has minima at 208 and 222 nm, which suggests it is helical under the sample conditions. The spectrum of the 1:1 mixture of the peptides was compared to the average of the individual spectra. The experimentally measured mixture (green) had a slightly stronger helical signature

than the calculated theoretical average (black). Therefore, there may be some weak interaction between the inhibitor candidate and parent peptide. However, the helical signature of the inhibitor candidate alone is greater than the helical signature of the mixture. If the inhibitor candidate tightly binds the parent peptide in a 1:1 fashion, the helical signature of the mixture should be at least as intense as the inhibitor candidate's helical signature alone.

In order to further understand the interaction between inhibitor candidate and parent peptide, variable temperature data was recorded (Figure 4.6c). The samples were monitored by CD at 208 and 222 nm. The parent peptide appears to lose all helical character by about 20ºC, and the inhibitor appears to lose its helicity at about 48ºC. The data for a 1:1 mixture were compared to the calculated theoretical CD signature of a non-interacting mixture. The two melting curves are approximately the same, supporting the idea that inhibitor candidate 4.2 does not bind the parent peptide.

(A) RL**E**KE**L**EEK**K**EA**L**ELA**I**DQ**A**SRD**Y**HR**A**TAL**E**KE**L**EEK (Parent peptide 4.1)
KKE**L**EEK**L**EA**L**ELA**I**DQ**A**KRK**Y**HR**A**KAK**L**KE**L**EEK (Inhibitor candidate 4.2)

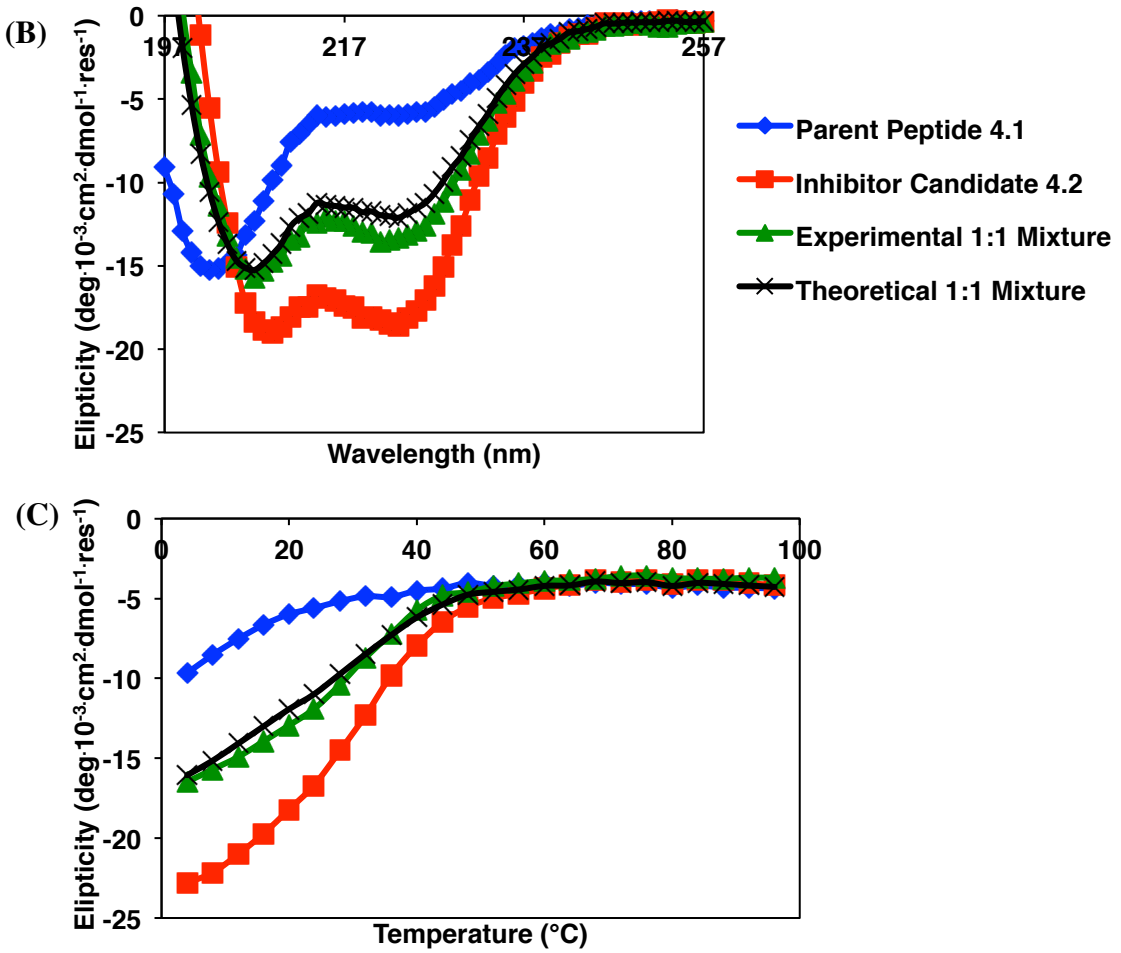


Figure 4.6. (A) Sequence of parent peptide 4.1 and inhibitor candidate 4.2. (B) CD spectra of parent peptide 4.1 and inhibitor candidate 4.2. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration for each measurement was 30 μM (phosphate buffer, pH 7.0, 20°C). (C) Variable temperature data for inhibitor candidate 4.2, parent peptide 4.1, and a 1:1 mixture (30 μM, phosphate buffer, pH 7.0). CD signal was monitored at 222 nm.

Inhibitor candidate 4.3 was also tested by CD (Figure 4.7). The experimental mixture of inhibitor 4.3 and parent peptide is more intense than the theoretical spectrum. Therefore, peptide 4.3 appears to weakly bind peptide 4.1. The variable temperature data support this conclusion. Below 48°C, the experimental CD signal at 222 nm is greater

than the theoretical CD signal. Therefore, below 48ºC, some of the inhibitor candidate binds the parent peptide. However, the inhibitor candidate alone appears to melt above 60ºC. If the inhibitor candidate is self-associating, this implies that it binds itself more strongly than it binds the parent peptide. Self-association of the inhibitor candidate is concerning as it may interfer with its ability to bind the target M1 sequence.

(A)

RL**E**KE**L**EEK**K**EA**L**ELA**I**DQ**A**SRD**Y**HR**A**TAL**E**KE**L**EEK (Parent Peptide 4.1)

KKE**L**EEK**L**EA**L**ELA**I**DQ**L**KRK**A**HR**L**KAK**L**KE**L**EEK (Inhibitor Candidate 4.3)

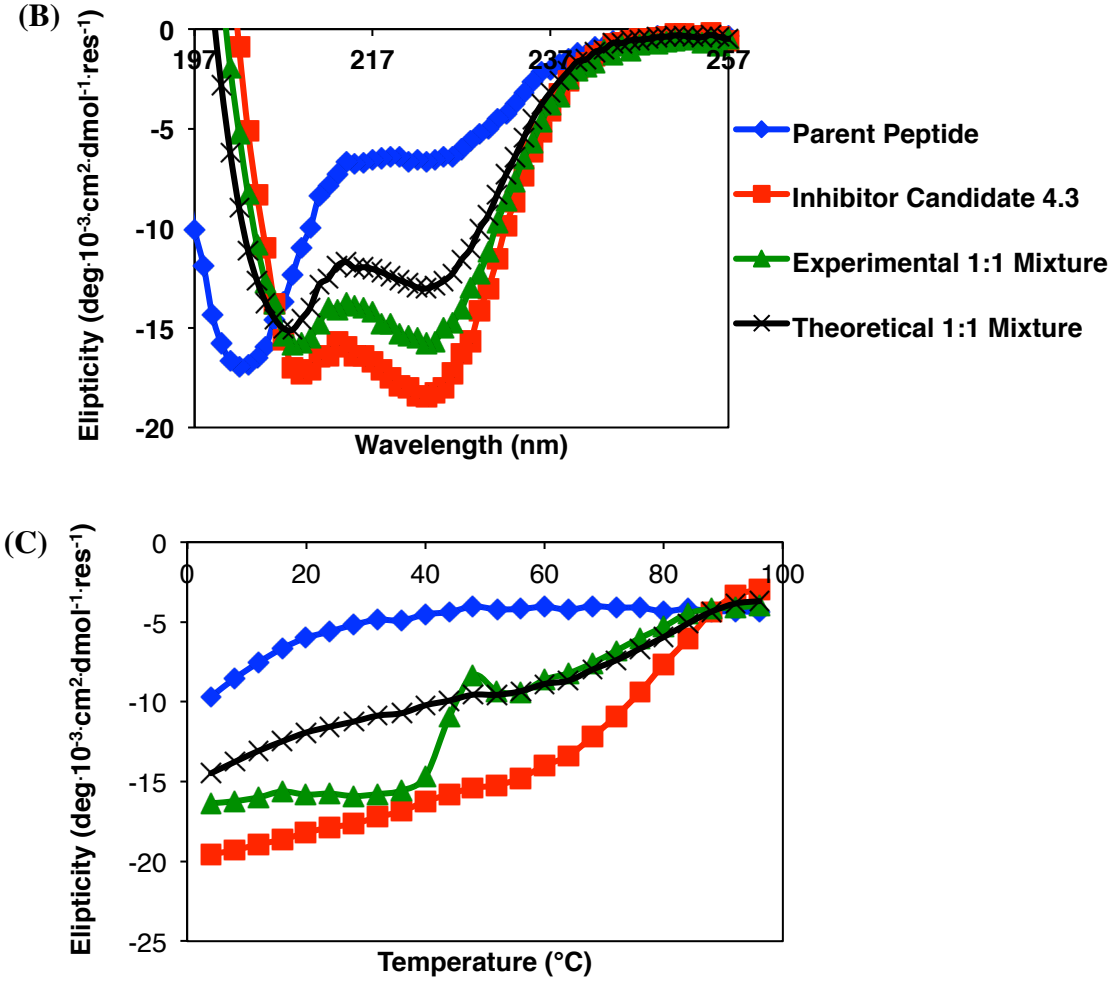


Figure 4.7. (A) Sequence of parent peptide 4.1 and inhibitor candidate 4.3. (B) CD spectra of parent peptide 4.1 and inhibitor candidate 4.3. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration of each spectra was 30 μM (phosphate buffer, pH 7.0, 20ºC). (C) Variable temperature data for inhibitor candidate 4.3, parent peptide 4.1, and a 1:1 mixture (30 μM, phosphate buffer, pH 7.0). CD signal was monitored at 222 nm.

In an attempt to prevent self-association of the inhibitor candidates, I synthesized a series of peptides with Lys at all e and g positions. In the inhibitor candidates' self-associated form, these Lys will be brought close to one another to form energetically

unfavorable interactions (Figure 4.8). In the heterodimer, the positively charged Lys in peptide 4.4 should pack against negatively charged or neutral side chains in parent peptide 4.1, which are energetically favorable interactions (Figure 4.8). The second-generation inhibitors were a heptad shorter than the first-generation inhibitors to avoid unfavorable interactions between positively charged parent peptide e and g positions and positively charged inhibitor candidate e and g positions. In addition, shorter peptides are more practical as they are generally easier to synthesize. A YGG tag was added to the N-termini of all peptides without a Tyr to allow the quantification of peptide concentration by UV-visible spectroscopy.

(A)

YGG**L**EA**L**KLK**I**DQ**L**KRK**A**HR**L**KAK**L**KE**L**KEK (4.4)
YGG**L**EA**L**KLK**I**DQ**A**KRK**A**HR**A**KAK**L**KE**L**KEK (4.5)
YGG**L**EA**L**KLK**I**DQ**L**KRK**L**HR**L**KAK**L**KE**L**KEK (4.6)
LEA**L**KLK**I**DQ**L**KRK**Y**HR**L**KAK**L**KE**L**KEK (4.7)

(B) (C)

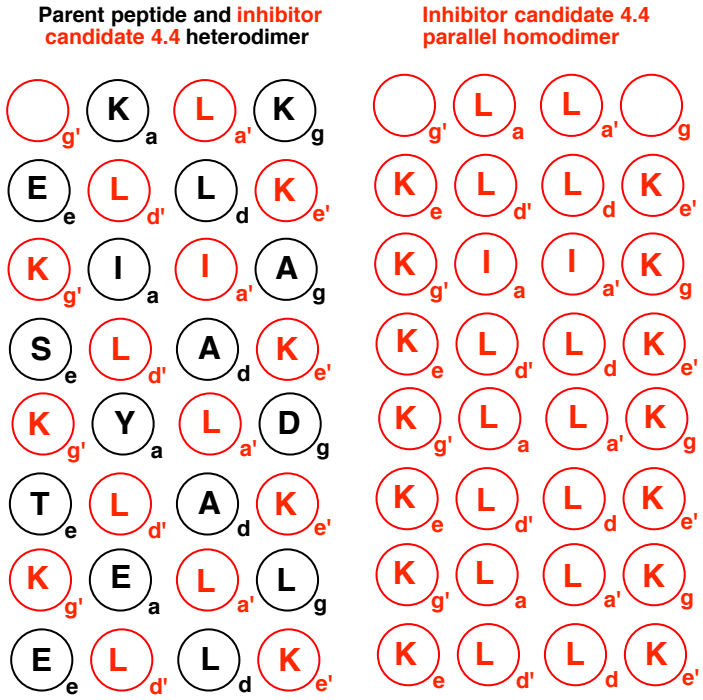
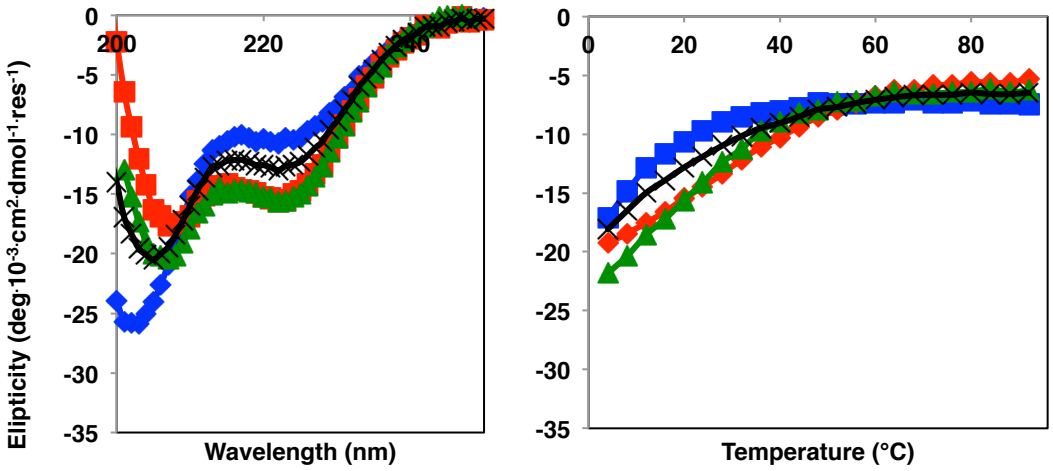


Figure 4.8. (A) Inhibitor candidate sequences. Helical net diagrams of (B) parent peptide 4.1 and inhibitor candidate 4.4 heterodimer and (C) inhibitor candidate 4.4 parallel homodimer.

Peptides 4.4 and 4.5, which had at least one Ala at a core position, did not have strong helical signatures and did not appear to strongly interact with the parent peptide (Figure 4.9).

(A) YGG**L**EA**L**KLK**I**DQ**L**KRK**A**HR**L**KAK**L**KE**L**KEK (4.4)



(B) YGG**L**EA**L**KLK**I**DQ**A**KRK**A**HR**A**KAK**L**KE**L**KEK (4.5)

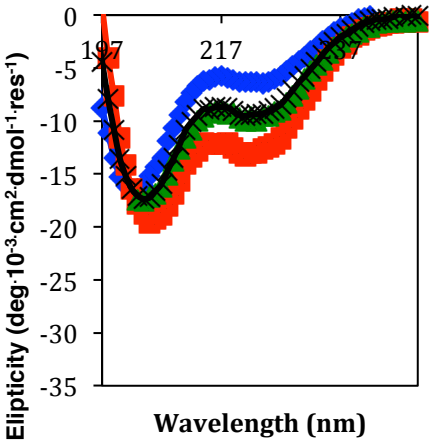
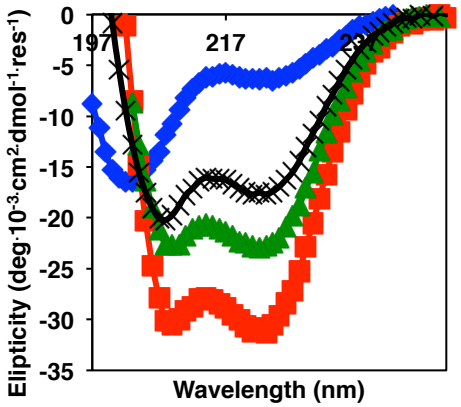


Figure 4.9. (A) CD spectra and variable temperature data of parent peptide 4.1 and inhibitor candidate 4.4. The variable temperature data was monitored at 222 nm. (B) CD spectra of parent peptide 4.1 and inhibitor candidate 4.5. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration for each measurement was 30 μM (phosphate buffer, pH 7.0, 20°C). Inhibitor candidates are shown in red, parent peptide 4.1 is shown in blue, theoretical 1:1 mixtures are shown in black. Experimental 1:1 mixtures are shown in green.

Peptides 4.6 and 4.7 had strong helical signatures and appeared to interact weakly with the parent peptide. A melting curve for peptide 4.7 suggests that by 24°C the

interaction between the inhibitor candidate and parent peptide is negligible. A successful inhibitor of the M1 interaction must bind M1 strongly at physiologically relevant temperature (37ºC). Furthermore, substituting three core positions in peptide 4.7 (Leu14, Tyr18, Leu21) for Ala, a residue that stabilizes helical secondary structure but does not promote dimerization, results in a peptide 4.5, which has no helical character. The lack of helical character in peptide 4.5 suggests that the helical character seen in peptide 4.7 is promoted by self-association. Since strong self-association may interfere with a peptide's ability to bind M1, designs with features that further discourage self-association may lead to improved inhibitor candidates.

(A) YGG**L**EA**L**KLK**I**DQ**L**KRK**L**HR**L**KAK**L**KE**L**KEK (4.6)



(B) **L**EA**L**KLK**I**DQ**L**KRK**Y**HR**L**KAK**L**KE**L**KEK (4.7)

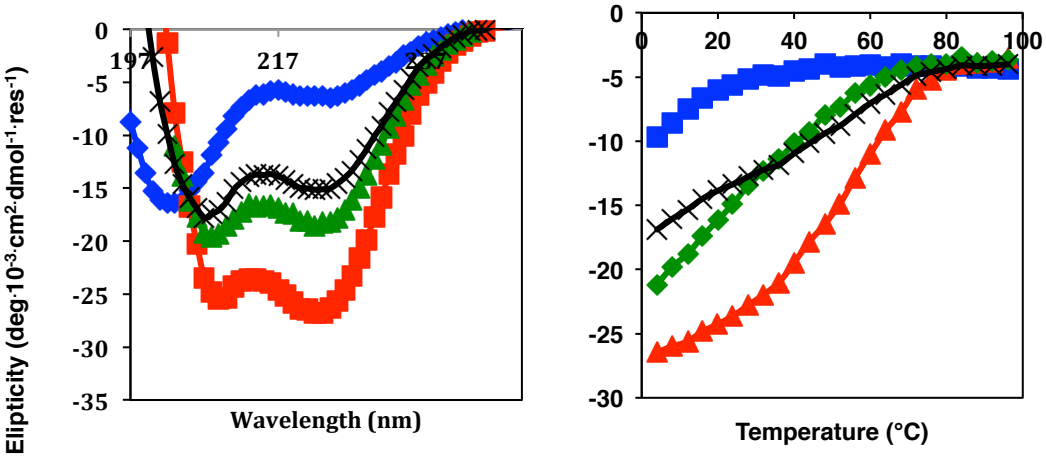


Figure 4.10. (A) CD spectra of parent peptide 4.1 and inhibitor candidate 4.6. (B) CD spectra and variable temperature data of parent peptide 4.1 and inhibitor candidate 4.7. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration was 30 μM for each measurement (phosphate buffer, pH 7.0, 20°C). The temperature melt was monitored at 222 nm. The inhibitor candidates are in red, the parent peptide is in blue, the experimental mixture is in green, and the theoretical mixture is in black.

In a further attempt to prevent self-association, core positions in the inhibitor candidates were substituted with positively charged Lys. In order to offset the destablization from the Lys, other positions on the back-side of the helix were substituted for residues with higher helical propensity. The resulting set of inhibitor candidates

(peptides 4.8-4.11) was tested by CD. None of the inhibitor candidates were helical, and none of the 1:1 mixtures of inhibitor canidate and parent peptide were helical (Figure 4.11). Given the failure to find an inhibitor candidate from the register 1 interface, the register 2 interface was explored.

(A)

YGG**L**EA**K**KAK**I**EQ**L**KRK**L**HR**L**KAK**L**KE**K**KEK (4.8)
YGG**L**EA**K**KAK**I**EQ**L**KRK**I**HR**L**KAK**I**KE**K**KEK (4.9)
YGG**L**EA**K**KAK**I**EQ**L**KRK**K**HR**L**KAK**I**KE**K**KEK (4.10)
YGG**L**EA**K**KAK**I**EQ**L**KRK**A**HR**L**KAK**I**KE**K**KEK (4.11)

(B)

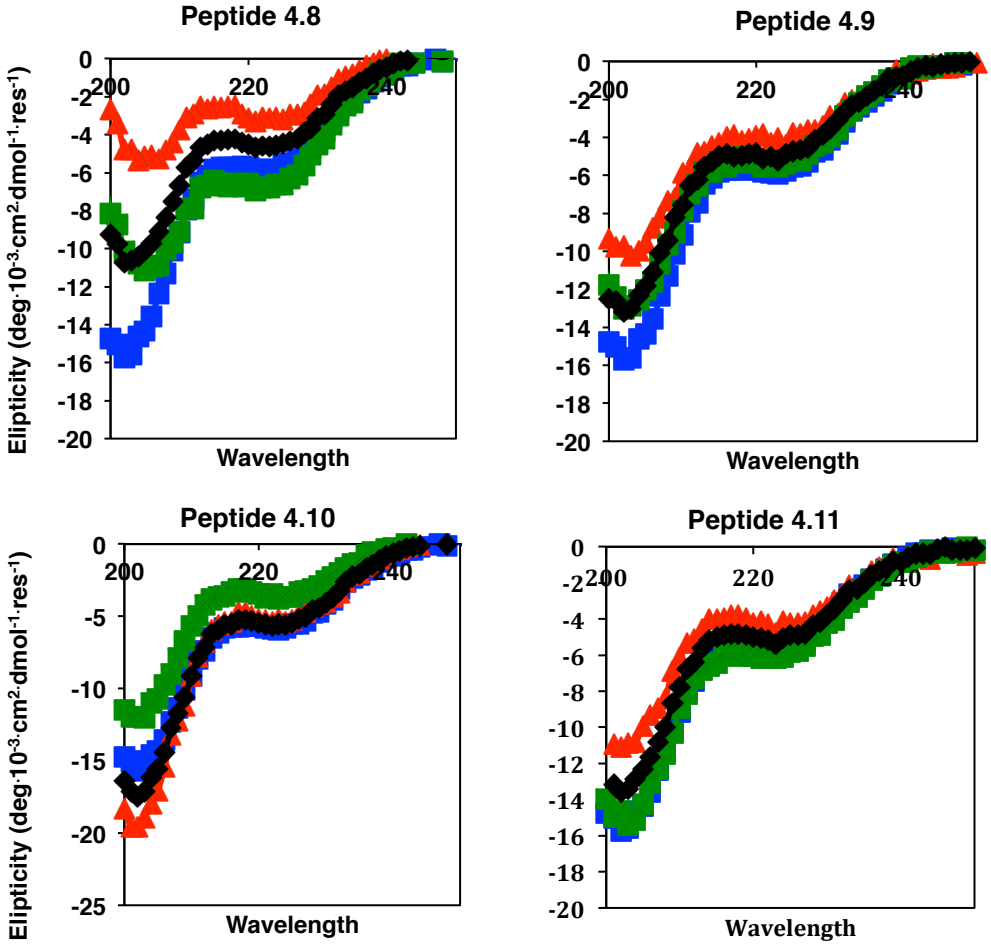


Figure 4.10. (A) Peptide sequences. (B) CD spectra of parent peptide 4.1 and inhibitor candidates. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration for each measurement was 30 μM (phosphate buffer, pH 7.0, 20°C). Inhibitor candidates are in red, parent peptide is in blue, experimental 1:1 mixtures are in green, and theoretical 1:1 mixtures are in black.

4.2.2. Design and evaluation of register 2 M1 inhibitor candidates.

Several sequences were designed to bind register 2 using similar design principles as explored for the register 1 inhibitor candidates (Figure 4.11). After several rounds of trial and error, two sequences where identified in which the mixture of the inhibitor candidate and parent peptide has a CD signature similar to that of the inhibitor candidate alone. The core a and d positions for the sucessful inhibitor canidates (4.17 and 4.18) are amino acids that promote the stability of the heterodimer and do not destabilize the homodimer. For both sequences, the mixture appears to melt at a lower temperature than the inhibitor candidate, but the results are still mildly promising (Figure 4.13).

IDQ**A**SR**D**YHR**A**TA**L**EKE**L**EE**K**KKA**L**EL**A**IDQ**A**SQ**D**YN Parent Peptide 4.12
YGGR**L**SR**L**KHR**I**TA**L**QKE**L**EE**E**EKQ**L**EL**L**RDR**I**SQ**L**R 4.13
YGGR**E**SR**L**KHR**L**TA**L**QKE**L**EE**E**EKQ**L**EL**L**RER**L**SQ**E**R 4.14
YGGR**R**SR**L**KHR**L**TA**L**QKE**L**EE**E**EKQ**L**EL**L**RER**L**SQ**R**R 4.15
YGGK**K**KK**L**EEE**E**EE**L**ELL**K**DK**I**SRL**K**HR**I**TAL**K**KK**L**E 4.16
YGGR**R**SR**L**KHR**L**TA**L**RKR**L**EE**L**EKR**L**EL**L**RER**L**SQ**R**R 4.17
YGGR**R**SR**L**KHR**I**TA**L**RKR**L**EE**L**EKR**L**EL**L**RER**I**SQ**R**R 4.18

Figure 4.11. Sequences of register 2 inhibitor candidates and parent peptide. The a and d positions are bolded.

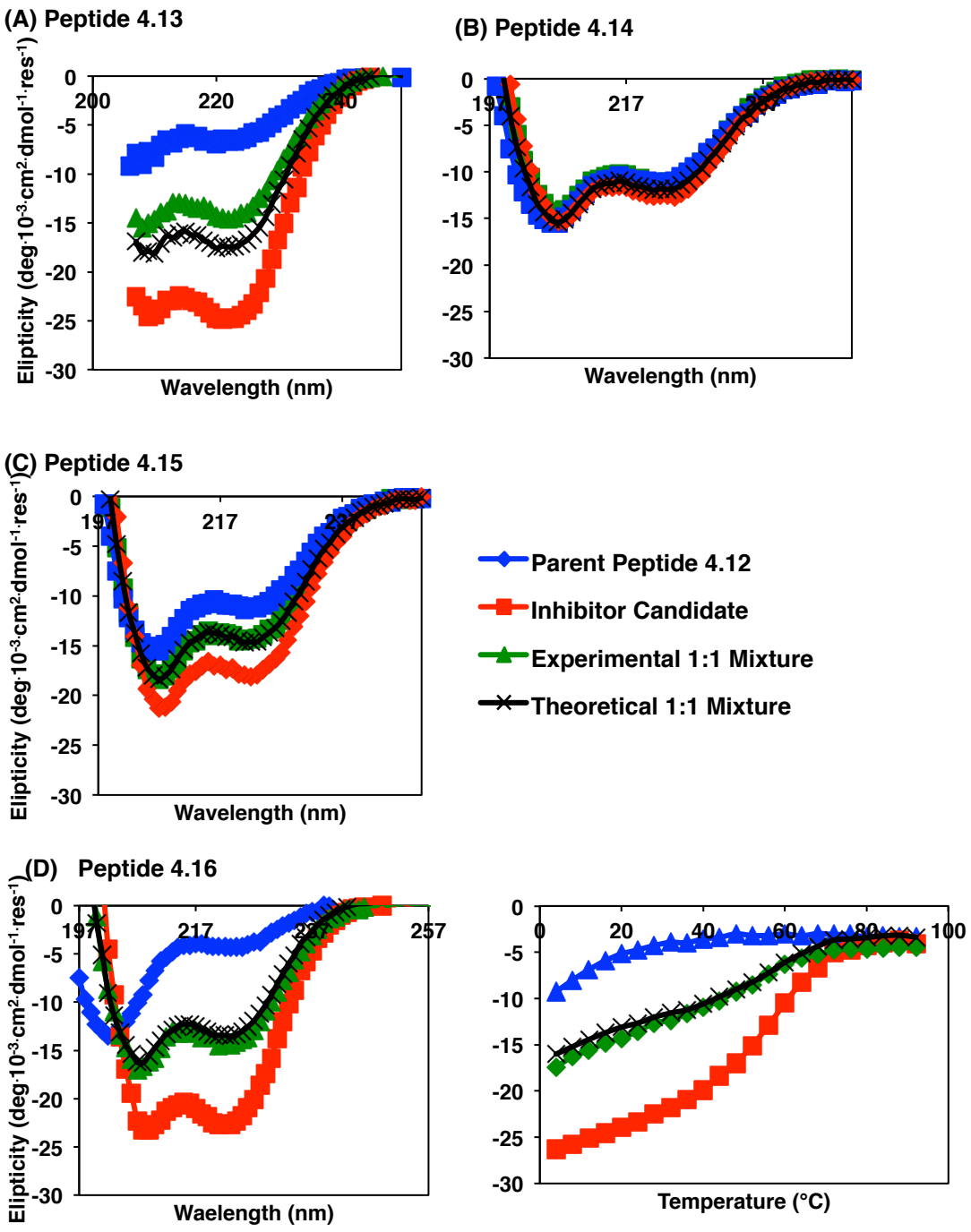
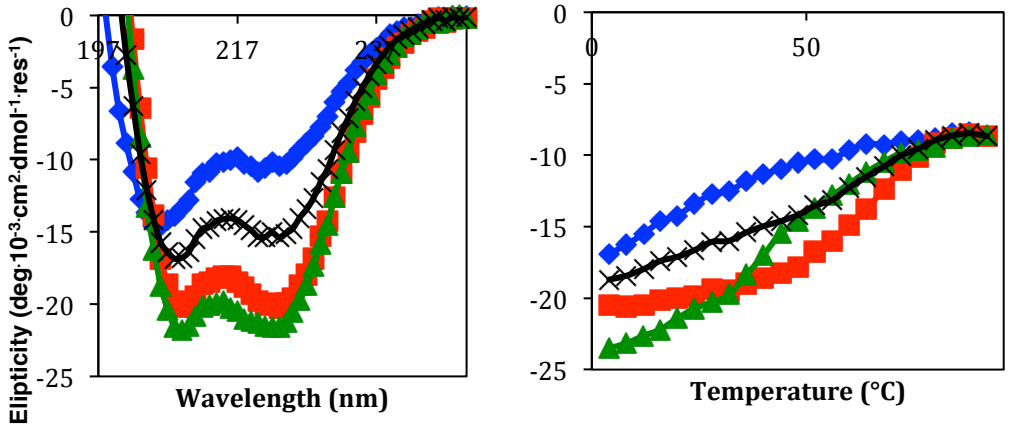


Figure 4.12. CD spectra of parent peptide 4.12 and inhibitor candidates. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. Unless otherwise noted, the total peptide concentration was 30 μM (phosphate buffer, pH 7.0, 20°C). The variable temperature data were monitored at 222 nm. (A) Inhibitor candidate 4.13 in HEPES (pH 7.0). Peptide 4.13 was not soluble in phosphate buffer. (B) Peptide 4.14. (C) Peptide 4.15 (D) Peptide 4.16.

(A)

YGGR**R**SR**L**KHR**L**TA**L**RKR**L**EE**L**EKR**L**EL**L**RER**L**SQ**R**R 4.17



(B)

YGGR**R**SR**L**KHR**I**TA**L**RKR**L**EE**L**EKR**L**EL**L**RER**I**SQ**R**R 4.18

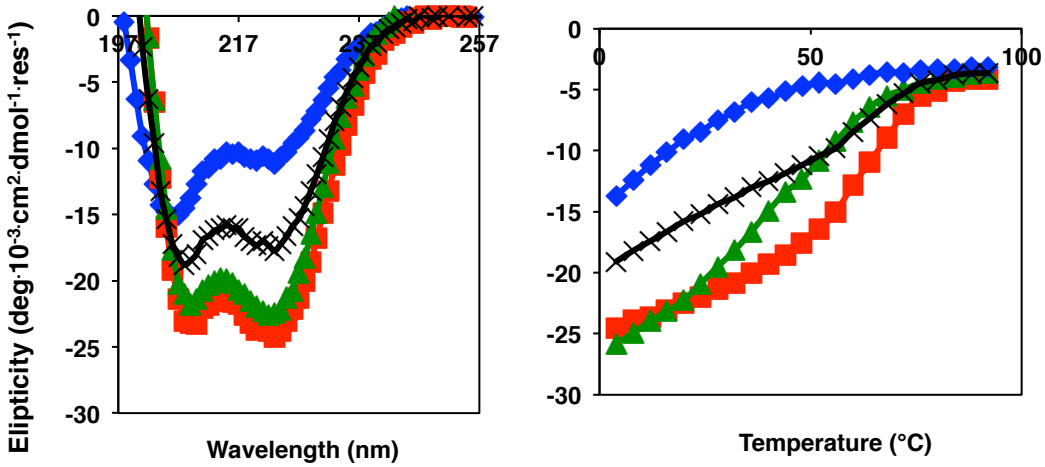


Figure 4.13. CD spectra and variable temperature data of parent peptide 4.12 and inhibitor candidates. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration was 30 μM for each measurement (phosphate buffer, pH 7.0, 20°C). (A) Inhibitor candidate 4.17 and (B) inhibitor candidate 4.18. Inhibitor candidates are in red, parent peptide is in blue, experimental 1:1 mixtures are in green, and theoretical 1:1 mixtures are in black.

4.2.3. M1 and FgD Bind assays of inhibitor candidates

Collaborators in the Ghosh lab at the University of California-San Diego tested our most promising inhibitor, peptide 4.18, for the ability to inhibit binding between M1 and FgD. In a co-precipitation assay, peptide 4.18 failed to block M1's association with FgD (Figure 4.14).

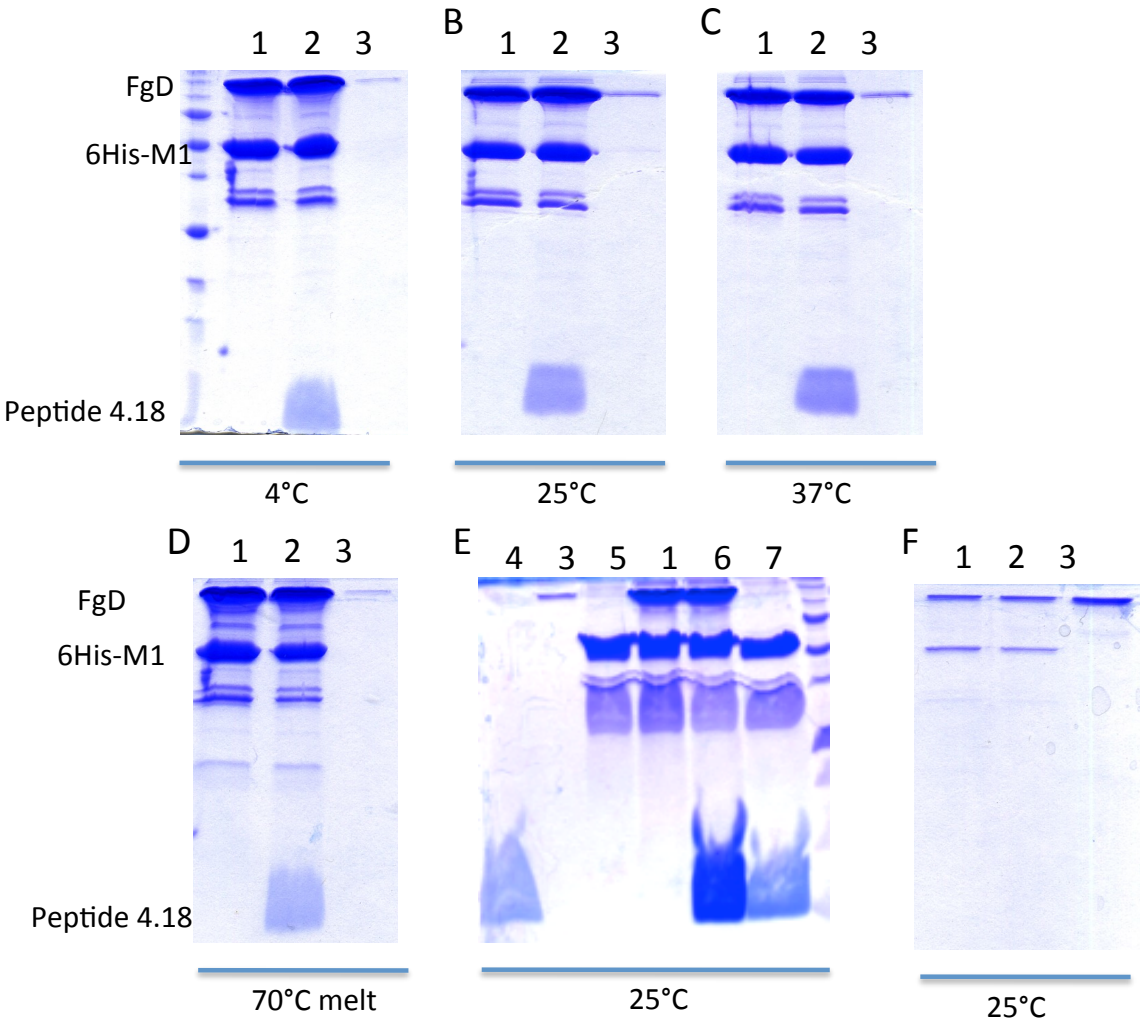


Figure 4.14. FgD and M1 co-precipitation assay. Column 1 is His tagged M1 (6His-M1), column 2 is 6His-M1, FgD, and peptide 4.18, and column 3 is FgD. Solutions were incubated at (A) 4ºC, (B), 25ºC, (C) 37ºC, and 70ºC with 100 μM inhibitor candidate and coprecipitated with a Ni^{2+}-nitrilotriacetic acid (NTA) agarose. Bound fractions were visualized on an SDS-PAGE gel with non-reducing coomassie stain. 6His-M1 binds FgD even in the presence of inhibitor. (E) Additional controls where run at 25ºC. Column 1-3 are as described above. Column 4 is peptide 4.18, column 5 is 6His-M1, column 6 is 6His-M1, FgD, and 800 μM peptide 4.18, and column 7 is 6His-M1 and FgD. (F) Unbound fractions for the 25ºC experiment. These experiments were performed by Cosmo Buffalo and Chelsea Stewart.

4.2.4. Attempted computational design of inhibitor candidates with PLUG

PLUG is a computational program from the Woolfson lab at the University of Bristol that designs peptide sequences to bind defined knobs in holes helical interfaces. The program attempts to find sequences that will form a coiled-coil association with the target sequence more strongly than they will self-associate. Using register 2 of the M1 sequence, collaborater Gail Bartlett used PLUG to find two sequences designed to bind M1. The PLUG program defines only the residues present at the coiled coil interface. Two difference sets of residues were used for the backside positions, leading to four possible designs. None of the designs showed any interactions with the parent peptide when measured by CD (Figure 4.15).

YGGR**T**SR**L**KHE**R**TA**L**DKK**I**EE**L**EKQ**L**EL**L**EDE**V**SQ**L**Q (4.19, A)
YGGK**A**SR**A**EHR**A**TA**L**EKK**I**EE**L**EKR**L**EL**A**EDK**V**SQ**A**Q (4.20, B)
YGGR**I**SR**L**KHE**R**AA**L**DKK**I**EE**L**EKQ**L**EA**L**EKE**V**SQ**L**Q (4.21, C)
YGGK**A**SR**A**EHR**A**AA**L**EKK**I**EE**L**EKR**L**EA**A**EKK**V**SQ**A**Q (4.22, D)

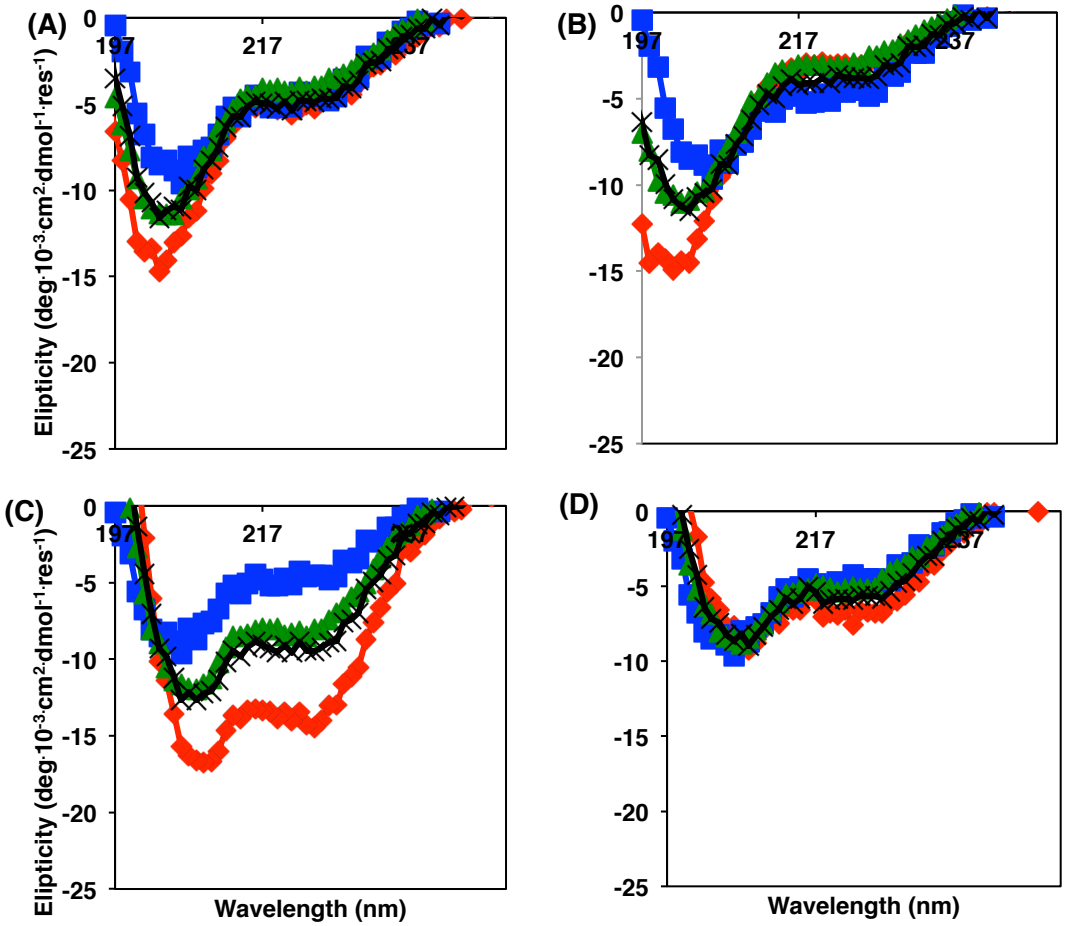


Figure 4.15. CD spectra of parent peptide 4.12 and inhibitor candidates. The CD spectrum of 1:1 mixture of each inhibitor candidate and parent peptide was measured experimentally. The theoretical spectra of the 1:1 mixture was calculated from the average of the inhibitor candidate and parent peptide spectra. The total peptide concentration of each spectrum was 30 μM (phosphate buffer, pH 7.0, 37°C).

4.3. Conclusions

Many of the residues at the core of M1's coiled coil interaction destabilize the interface. Inhibitors that can tightly bind M1 suffer from self-association because the

same residues that stabilize the heterodimer will often stabilize the undesired inhibitor homodimer. Therefore, M1 is a challenging target.

4.4. Methods

4.4.1. Peptide Synthesis

General. Unless otherwise noted, all materials were obtained from Chemimpex, Novabiochem, Fischer, Peptech or Sigma-Aldrich. Shimadzu reverse phase HPLCs were used with C5 or C18 reverse phase columns and a gradient of acetonitrile/water containing 0.1% TFA. The analytical HPLC had a flow rate of 1 ml/min and the preparative HPLC had a flow rate of 12-15 ml/min.

Peptide Synthesis. All peptides were synthesized by standard Fmoc solid phase peptide synthesis (SPPS) on NovaPEG rink amide resin, NovaPEG TGR resin, or Chemmatrix rink amide resin. All amino acids were obtained from Chemimpex, Peptech, or Novabiochem. All peptides were synthesized on a CEM Mars microwave with a temperature probe or on a symphony automated peptide synthesizer at room temperature. Residues were coupled onto the resin from the C-terminus to the N-terminus with iterative coupling and deprotection steps. In each coupling step, 4 eq of amino acid was activated with 4 eq coupling reagent, 4 eq HOBt or Cl-HOBt, and 8 eq DIEA in DMF or NMP (1ml for a 25 μmol scale) for at least 1 minute. The coupling reagent was PyBOP, HCTU or HBTU. With the microwave, the coupling solution was added to the resin and heated on the microwave (2 minute ramp to 70ºC, 4 minute hold at 70ºC). The resin was washed with DMF. In each deprotection step, 20% piperidine in DMF was added (>1

ml/25 umol). The solution was microwaved (2 min ramp to 80ºC, 2 min hold at 80ºC). Peptides were cleaved from the resin and deprotected with an acidic deprotection solution of H_2O (0.1 mL/50 μmol scale), triisopropylethylsilane (0.1 mL/50 μmol scale), and trifluoroacetic acid (3.8 ml/50 μmol scale) for 1-6 hours. The peptides were precipitated with cold ether and purified by reverse phase HPLC with gradients of acetonitrile and water with 0.1% trifloroacetic acid on a Shimadzu prep HPLC with a C5 or C18 column. All peptides were characterized by analytical HPLC and found to be greater than 90% pure. The masses of all peptides were checked by MALDI-MS on a TOF Bruker Reflex II with a 337 nm nitrogen laser or a TOF/TOF Bruker Ultraflex III with a SmartBeam laser. All purity checks and MALDI spectra can be found in the appendix.

4.4.2. CD spectra and temperature melts

All CD spectra were taken on an Aviv Model 420 Circular Dichroism Spectrometer with 1 mm pathlength cells with a bandwidth of 1.00 nm and an averaging time of 3 sec. Unless otherwise noted, all peptide solutions were 30 μM in 20 mM phosphate buffer, pH=7.0. Experimental mixtures were 1:1 inhibitor candidate: parent peptide with a total peptide concentration 30 μM. Unless otherwise noted, wavelength scans were taken at 20ºC. Measurements were taken at 300, 220 and 208 nm every 4ºC starting at 0ºC and ending at 96ºC. The sample was allowed to equilibrate for 600 seconds at each temperature before measurements were taken. Theoretical mixture spectra were the average of the spectra for the parent peptide and for the inhibitor candidate.

4.5. References

(1) Herwald, H.; Cramer, H.; Mörgelin, M.; Russell, W.; Sollenberg, U.; Norrby-Teglund, A.; Flodgaard, H.; Lindbom, L.; Björck, L. M Protein, a Classical Bacterial Virulence Determinant, Forms Complexes with Fibrinogen That Induce Vascular Leakage. *Cell* **2004**, *116*, 367–379.

(2) Macheboeuf, P.; Buffalo, C.; Fu, C.; Zinkernagel, A. S.; Cole, J. N.; Johnson, J. E.; Nizet, V.; Ghosh, P. Streptococcal M1 Protein Constructs a Pathological Host Fibrinogen Network. *Nature* **2011**, *472*, 64–68.

(3) McNamara, C.; Zinkernagel, A. S.; Macheboeuf, P.; Cunningham, M. W.; Nizet, V.; Ghosh, P. Coiled-Coil Irregularities and Instabilities in Group A Streptococcus M1 Are Required for Virulence. *Science* **2008**, *319*, 1405–1408.

(4) Steinkruger, J. D.; Woolfson, D. N.; Gellman, S. H. Side-Chain Pairing Preferences in the Parallel Coiled-Coil Dimer Motif: Insight on Ion Pairing between Core and Flanking Sites. *J. Am. Chem. Soc.* **2010**, *132*, 7586–7588.

(5) Krylov, D.; Mikhailenko, I.; Vinson, C. A Thermodynamic Scale for Leucine Zipper Stability and Dimerization Specificity: E and G Interhelical Interactions. *EMBO J.* **1994**, *13*, 2849–2861.

(6) Acharya, A.; Ruvinov, S. B.; Gal, J.; Moll, J. R.; Vinson, C. A Heterodimerizing Leucine Zipper Coiled Coil System for Examining the Specificity of a Position Interactions: Amino Acids I, V, L, N, A, and K. *Biochemistry* **2002**, *41*, 14122–14131.

(7) Acharya, a; Acharya, a; Rishi, V.; Rishi, V.; Vinson, C.; Vinson, C. Stability of 100 Homo and Heterotypic Coiled-Coil a-a' Pairs for Ten Amino Acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **2006**, *45*, 11324–11332.

Chapter 5. Toward Backbone Thioester Exchange (BTE) Systems to Study the Thermodynamic Stability of Peptides

5.1 Introduction

The folding and stability proteins and peptides has long been of interest. The folded state of large proteins may only be stabilized by a few kcal/mol relative to the denatured state.[1] In some cases, the posttranslational modification of a single amino acid can have large global effects on structure.[2] It is useful to be able to understand, in detail, the effects of different amino acids on structure and stability.

Several methods have been used to obtain the stabilities of the folded conformations of peptides over the years.[1] As discussed in detail in Chapter 1, classical methods involve the chemical or thermal denaturation of peptides.[1] The limitation of these methods are that they are not performed under physiologically relevant conditions as they require the addition of harsh denaturants or an increase in temperature.[1,3]

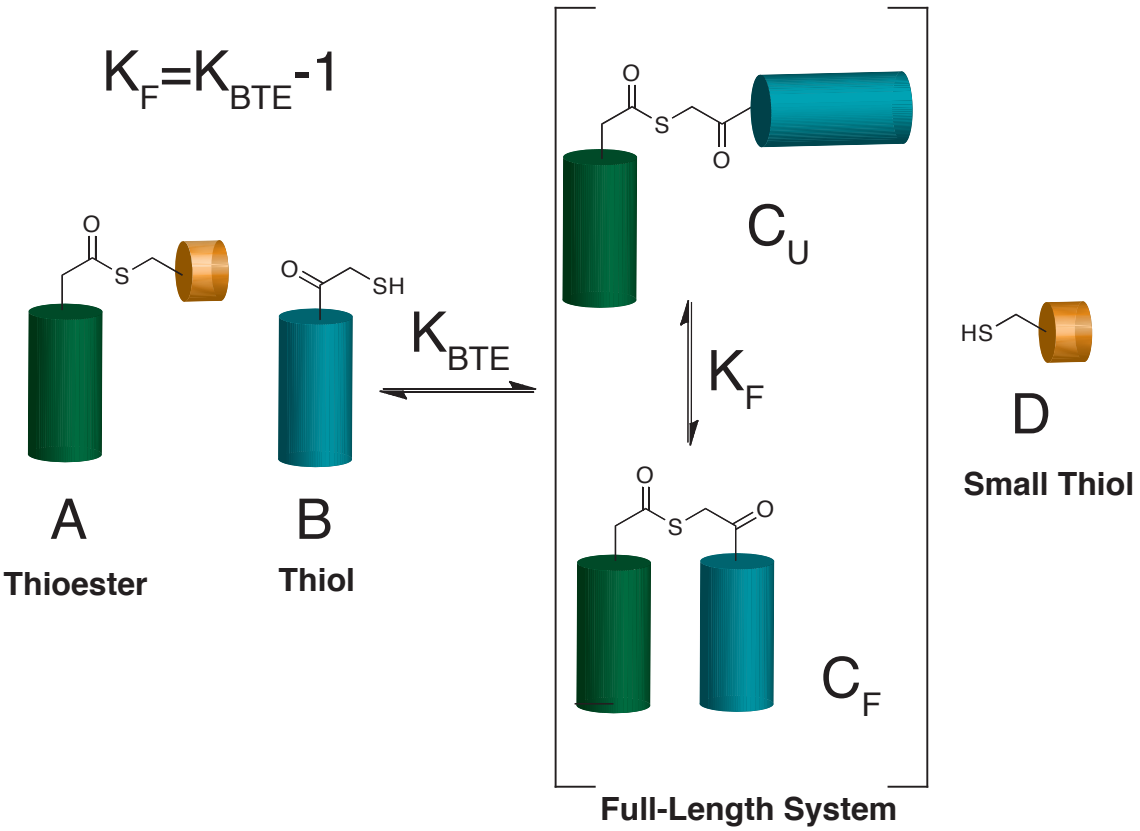
Backbone thioester exchange (BTE) is a method that allows the measurement of the thermodynamic stability of the folded conformations of peptides at physiologically relevant temperature and pH.[3] BTE systems have been used to study the interaction of two helices in detail.[4–7] In these systems, two helices are connected by a flexible linker of glycine residues.[7] A thioester bond replaces an amide bond in the middle of the flexible linker.[7] Thioester bonds are labile at neutral pH.[8] Figure 5.1 shows the transthioesterification that occurs in the presence of the small molecule thiol. The system equilibrates between the full-length thioester-containing peptide plus small molecule thiol and shorter thioester- and thiol-containing peptides.[7] The structure of the small molecule thiol is usually the tyrosine derived structure (Thiol **5.0**) shown in Figure 5.1.[7] In the context of the thioester (Species A, Figure 5.1), the small thiol should not be able form energetically significant interactions with the peptide. This assumption is tested by

allowing thioester A to equilibrate with different small molecule thiols and measuring the relative concentrations of the resulting species.

In the BTE system, the equilibrium constant between the shorter peptides and the full-length peptide plus small thiol is given by

$$K_{BTE} = \frac{[C][D]}{[A][B]} = \frac{([C_F]+[C_U])[D]}{[A][B]} \qquad \text{Equation 5.1}$$

where A, B, C, and D are the concentration of the species as defined in Figure 5.1. The concentration of the species is generally measured by the area under each of the four peaks in analytical HPLC,[3] but can also be measured using ^{19}F-NMR.[9]



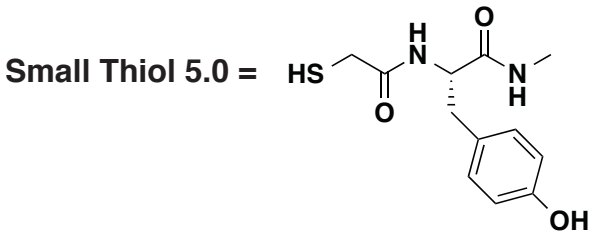


Figure 5.1 Backbone Thioester Exchange

The folding equilibrium constant (K_F) is the value of interest as it allows us to calculate a ΔG_F for the peptide of interest.[3] We can calculate K_F from K_{BTE} using the following derivation.[3]

$$K_F=\frac{[C_F]}{[C_U]} \qquad \text{Equation 5.2}$$

Combining equations 5.1 and 5.2 gives

$$K_{BTE}=\frac{(1+K_F)[C_U][D]}{[A][B]} \qquad \text{Equation 5.3}$$

Assuming the thioester bonds in species A and C are isoenergetic (an assumption which can be tested), then

$$\frac{[C_U][D]}{[A][B]} = 1 \qquad \text{Equation 5.4}$$

If you combine equations 5.3 and 5.4 you get

$$K_{BTE} = 1 + K_F \qquad \text{Equation 5.5}$$

which rearranges to give

$$K_F = K_{BTE} - 1 \qquad \text{Equation 5.6}$$

Using equation 5.6, we calculate K_F from integrations of HPLC peaks. K_F can be used to calculate ΔG_F from the equation

$$\Delta G_F = -RT\ln(K_F) \qquad \text{Equation 5.7}$$

5.2. Toward a BTE system to study the interactions α/β helices

5.2.1. Background

As discussed in Chapter 1, unnatural amino acids in peptides are useful both to enforce secondary structure and prevent proteolysis.[10] A detailed understanding of the structural propensities of unnatural amino acids can aid the rational design of foldamers.[5,10] Previously, BTE has been used to study the interaction of a helix of all natural α amino acids with a helix of alternating α and β amino acids (α/β helix).[5] I attempted to extend this work to study the interaction of an α/β helix with another α/β helix. The BTE system was intended to aid in the design of foldamers with more complex tertiary structure.

5.2.2. Design, synthesis, and evaluation of a BTE system for studying the association of α/β helices

An α/β helix-loop-helix BTE system was designed based on the previous α helix-loop-α/β helix BTE system (Figure 5.2).[5]

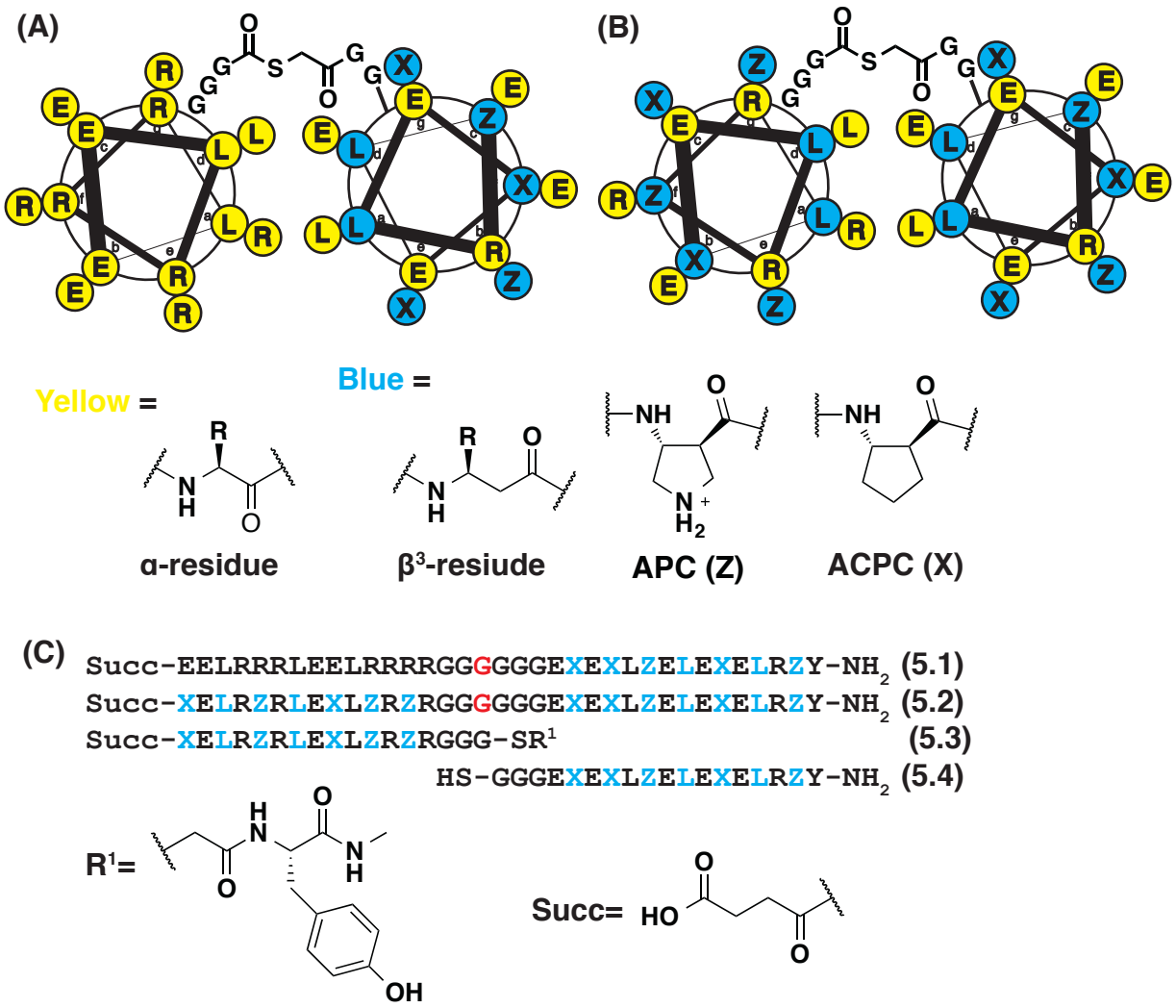


Figure 5.2. Helical wheel diagrams of a (A) chimeric α helix loop α/β helix BTE system[5] and (B) α/β helix loop helix BTE system. (C) Primary sequences.

The N-terminal helix (thioester 5.3) was synthesized via solid phase peptide synthesis (SPPS) as a C-terminal thioester using the sulfonamide method[11] or trithioorthoester method.[12] The trithioorthoester method was more labor intensive as it required the synthesis of a linker and trithioorthoester, but qualitatively seemed to produce much higher yields. The C-terminal helix

(thiol 5.4) was synthesized using standard SPPS. A BTE assay was performed by mixing thiol 5.4 (Species A, Figure 5.1) and thioester 5.3 (Species B, Figure 5.1) in buffer and allowing the system to equilibrate between the starting materials and the small thiol (Species C, Figure 5.1) plus full-length depsipeptide (Species D, Figure 5.1). A preliminary ΔG_F was calculated to be -1.0 kcal/mol. In a well-behaved BTE system, the concentration of the full-length depsipeptide should be approximately the concentration of the tyrosine-derived small thiol. However, the HPLC trace of the BTE experiment showed the concentration of the small thiol was more than double the concentration of the full-length depsipeptide, suggesting the thioester bond had undergone hydrolysis or another side reaction. Since the only Tyr or Trp residue in thioester 5.3 is part of the C-terminal thiol, hydrolysis of the thioester bond results in a peptide that cannot be detected at 275 nm in the HPLC trace. Normally, one would expect to see evidence of a hydrolysis peak at 220 nm, but no additional peak was detected at this wavelength.

In order to check for degradation, thioester 5.3 was dissolved in BTE buffer and monitored over time by analytical HPLC. After 2.5 hours, a small thiol peak was observed on the analytical HPLC trace, and the integration of the thioester had decreased. A fine white precipitate was visible in the buffered thioester solution, even though the thioester had initially dissolved completely. MALDI analysis of the precipitate showed it was 70 Da lighter than the thioester, consistent with an intramolecular acyl transfer reaction with a secondary amine of an APC residue (Figure 5.4a). It is also possible some thioester is degraded by hydrolysis of the thioester bond. The precipitate was dissolved in water, acetonitrile, and trifluoroacetic acid and run on the analytical HPLC. Three peaks were present in the 220 nm trace (Figure 5.4). Note that there are three APC residues in the thioester peptide and therefore three possible acyl transfer products (Figure 5.3).

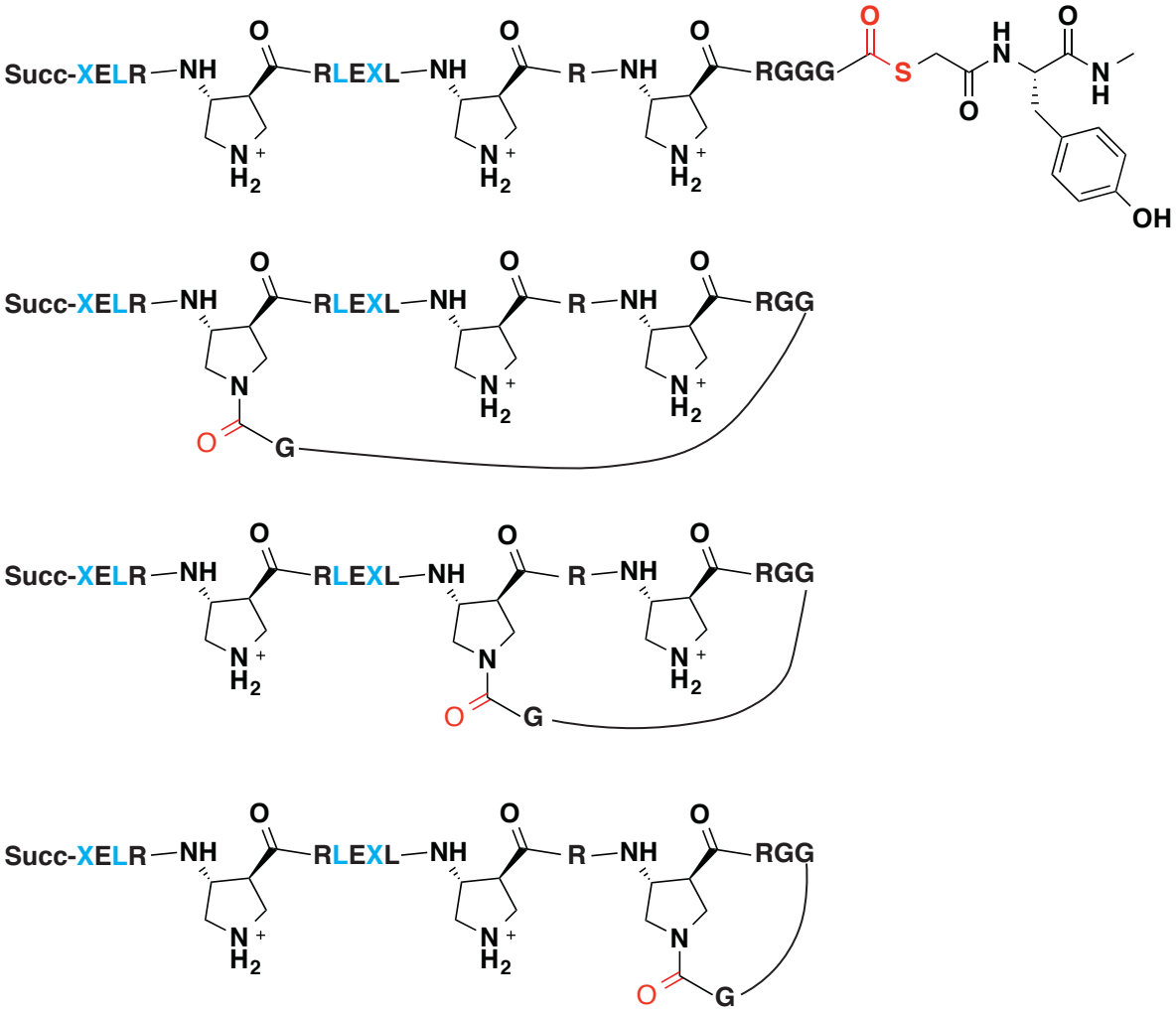


Figure 5.3. Peptide 5.3 and acyl transfer degradation products.

In order to get an approximate rate of degradation, we can look at the 275 nm traces of the thioester solution. Assuming the concentration of the small thiol is equal to the amount of thioester that has been degraded, by 2.4 hours, 39% of the thioester has been lost. In an attempt to slow the side reaction, BTE experiments were done at pH 6.5, but the system was not able to equilibrate in a reasonable time frame under these conditions.

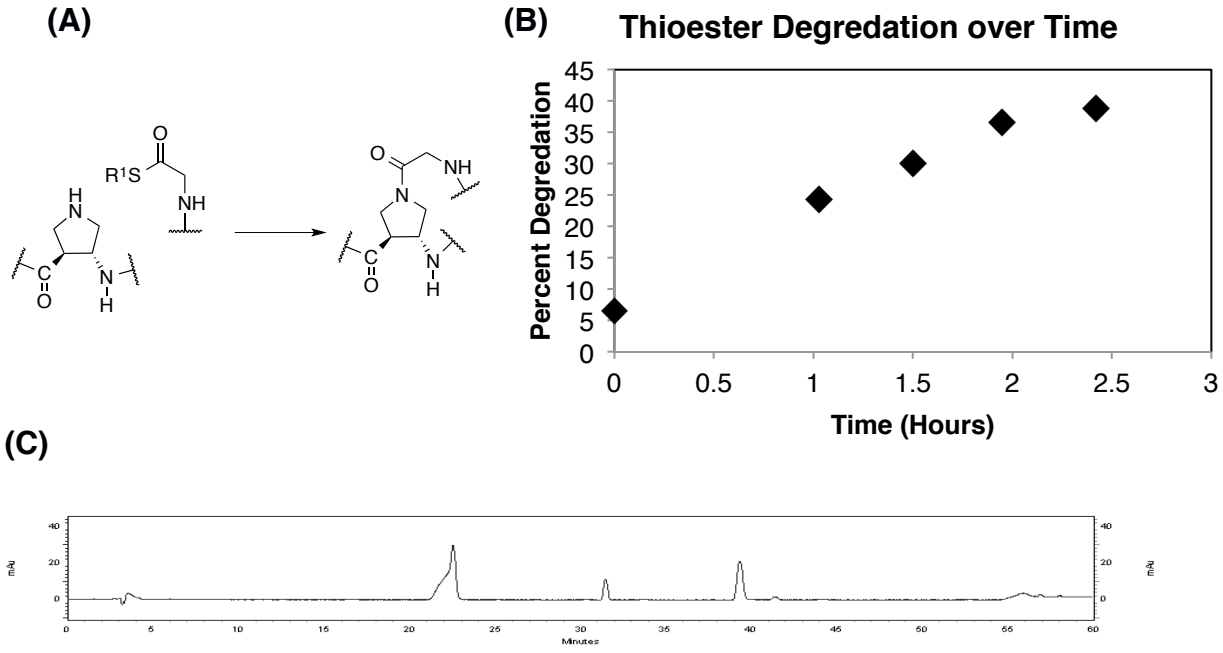


Figure 5.4. Thioester degradation study (A) Acyl transfer reaction (B) Thioester degradation over time. (C) HPLC trace of precipitate in thioester solution

Given that the thioester is degraded rapidly, reliable thermodynamic measurements cannot be made from thioester exchange. The APC residues in the thioester are probably the cause of the degradation, and the system cannot be used to determine the thermodynamic stability of an α/β helix-loop-helix tertiary structure.

5.2.3. Conclusions and Future Directions

The two APCs in the chimeric α and α/β BTE system from Price *et al*. did not cause significant thioester degradation within the time frame of the BTE measurements (Figure 5.2a).[5] One difference between the original system and the all α/β system is the presence of APCs in the N-terminal helix. These APCs can cause thioester degradation through intramolecular acyl transfer reactions in both the thioester (Species A, Figure 5.1) and full-length (Species C, Figure 5.1) depsipeptides of the BTE system. In contrast, the APCs in the system from Price *et al*. were

all in the C-terminal helix.[5] These APCs are not present in the shorter thioester (Species A, Figure 5.1) and therefore are only susceptible to intramolecular acyl transfer in the full-length depsipeptide. The full-length peptide with two helices may be more folded than the shorter thioester-containing peptide. When the depsipeptides are folded, the APCs are probably not correctly positioned to attack the thioester bond. Therefore, the acyl transfer reaction is slow enough for BTE measurements to be obtained. On the other hand, when APC is present in the thioester peptide, the degradation reactions can happen more quickly. It is also possible that the position of the APC in the thioester peptide is important for rate of the reaction. Regardless, peptide 5.2 cannot be used in a BTE assay.

The system must be redesigned. When the work was done, positively charged cyclic residues without an amine had not yet been developed. Now, derivatives of APC with a non-nucleophilic guanidino group instead of an amine could be used to develop a functional BTE system (Figure 5.5).[13] It is possible that a redesigned system involving anionic APC instead of APC might also be fruitful (Figure 5.5). Note that replacing the APCs with β^3-residues is also a possibility, but may not yield a well-folded structure as the cyclically restrained residues stabilize helical structure.[14]

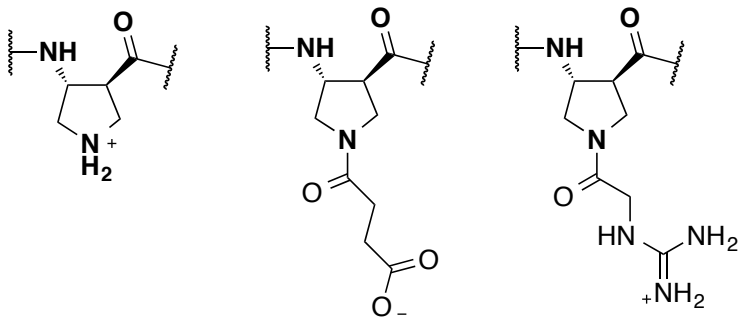


Figure 5.5. Structures of APC derivatives

5.3. Toward a BTE System to Study Intrinsically Disordered Peptides and Proteins

5.3.1. Background

Intrinsically disordered proteins (IDPs) lack a well-defined tertiary structure.[15] IDPs are prevalent in cells and involved in numerous cellular functions including cell signaling.[15] The traditional central dogma of biochemistry is that the primary sequence of a protein determines the structure that in turn determines the function of the protein. However, with disordered proteins or protein domains, the protein maintains function in the absence of a well-defined tertiary structure.[15] In some cases, the disorder is thought to be important in the recognition of multiple binding partners.[15]

The lack of well-defined tertiary structure makes IDPs difficult to characterize and to establish a detailed understanding of the relationship between sequence and function.[15] In addition, while IDPs do not adopt a single well-defined structure, they do have residual structure that may be important.[15] It is difficult to understand this residual structure with traditional biochemical techniques.[15] Therefore, we attempted to apply BTE in order to get detailed thermodynamic data from an ID system.

5.3.2. Truncated Villin Headpiece Subdomain (VHP) as an ID model system

VHP is a 35-residue peptide that folds into a discrete secondary structure of three helices with a hydrophobic core (Figure 5.6).[16] When the C-terminal helix is removed, the resulting peptide (tVHP) becomes intrinsically disordered, but NMR studies have shown the disordered state retains some native-like character.[17–19] BTE has previously been implemented with VHP to explore the effect Phe → F_5-Phe substitutions in the hydrophobic core of the peptide.[20]

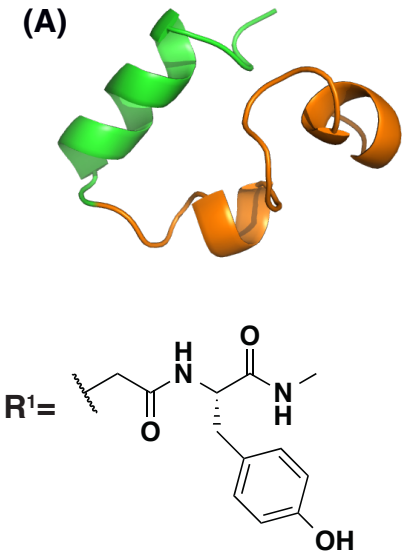


Figure 5.6. (A) Solution structure of VHP (PDB 1VII).[16] N-terminal helices are in orange. C-terminal helix is in green. (B) Sequence of VHP (5.5), truncated VHP (tVHP) (5.6), truncated VHP variant 5.7 for BTE, and tVHP variant 5.8 for BTE. The positions of the thioester bonds are highlighted in red.

In order to study the disordered tVHP, the full-length sequence (peptide 5.5) was truncated as done in previous studies to give intrinsically disordered peptide 5.6.[17–19] Arg was substituted for Lys at position 7 and a thioester bond was substituted for an amide bond between Phe10 and Gly11 as done in previous BTE studies of VHP (Depsipeptide 5.7, Figure 5.6).[20] The Arg to Lys substitution prevents degradation of the thioester through acyl transfer reactions. The N-terminal thioester and C-terminal thiol were synthesized (Depsipeptides 5.8 and 5.9). An initial BTE assay suggested a ΔG_F of around 0.8 kcal/mol. In order for the thermodynamic measurements from a BTE system to be meaningful, the peptide of interest must be soluble and monomeric in the BTE buffer. In order to quickly check for aggregation, Peptide 5.7 was synthesized with an amide bond in place of the thioester bond because the thioester bond can be hydrolyzed when left in solution for long periods of time. Turbidity studies were performed by dissolving the peptide in BTE buffer at 200, 400, or 800 μM and measuring the UV absorbance

spectra (200-350 nm) at 0 and 15 hours. The 400 and 800 μM samples showed lower UV absorbance at 15 hours than at 0 hours, indicating aggregation had occurred.[I] While it may have been possible to use the sequence by making BTE measurements at concentrations under 200 μM, concerns about aggregation in longer analytical ultracentrifugation experiments lead to the decision to modify the sequence to make it more soluble. Analytical ultracentrifugation (AU) is a method used to determine the aggregation state of peptides. In BTE assays, the data is analyzed with the assumption that the system is monomeric. AU is a reliable method to test this assumptions, and an AU study takes several days. In order to make the tVHP system more soluble and less prone to aggregation over time, the N-terminal Met residue was removed and Arg replaced Met4 in peptide 5.8. Studies with a version of peptide 5.8 were the thioester bond was replaced with an amide bond showed no loss of UV signal over 24 hours at 100, 200, and 400 μM.[I]

It should be noted that BTE systems are usually monitored by the absorbance of Tyr or Trp at 275 or 280 nm, respectively.[3] Since the extinction coefficients of Tyr and Trp are known, the integrations of the HPLC peaks can be correlated to the relative concentrations of the four species (Figure 5.1).[3] These concentrations can then in turn be used to calculate K_{BTE}.[3] However, the tVHP system lacks Trp or Tyr in the thiol peptide or full-length sequence. Therefore, absorbance must be measured at 220 nm, where amide bonds absorb. Since peptides of different lengths have different extinction coefficients at 220 nm, the integrations of the HPLC peaks alone cannot be used to calculate K_{BTE}.[21] Instead, a correction factor, E, must first be determined[21] where

$$K_{BTE}=E*K_{App}=E*\frac{CD}{AB} \quad \text{Equation 5.8}^{21}$$

[I] See methods section

and A, B, C and D are the integration at 220 nm of the HPLC peaks of species A, B, C, and D as defined in Figure 5.1.[I] E is calculated with the equation

$$E=\frac{\Delta Int_A \times \Delta Int_B}{\Delta Int_C \times \Delta Int_D} \qquad \text{Equation 5.9}^{21}$$

where ΔInt is the change in integration of the specified peak over time and is generally calculated by analyzing equal volumes of a BTE assay by HPLC directly after mixing the peptides and again after the system has equilibrated. The amount of the BTE solution injected into the HPLC should be approximately the same for the measurement at approximately 0 minutes and the measurement after the system has been equilibrated. Since slight differences in the volume of the BTE solution injected into the HPLC will increase the error in the measurement of ΔInt, we control for the volume of solution injected by adding the total integration of the Tyr-containing peaks at 275 nm. Since the total integration of the Tyr containing peaks should be the same if the same volume of solution was injected, we can weigh the measured integrations of the HPLC peaks accordingly. Determining the correction factor is problematic because it introduces systematic error into the BTE measurement. However, we are still able to account for that error if we measure the correction factor from multiple replicates. So, in theory, it is a feasible method.

The more concerning source of error in the tVHP system is the error associated with the high ΔG_F values. ΔG_F values are expected to be positive for systems that are not well folded as is probably the case in ID systems. Any system that is too stable or too unstable will be prone to error due to the logarithmic nature of the relationship between K_F and ΔG_F (Figure 5.7). In a BTE experiment, we measure K_F and convert it to ΔG_F using equation 5.7. The precision to which we can measure K_F is limited by the accuracy with which we can measure the integrations of the resulting HPLC peaks. Assuming the error in K_F remains approximately constant at values above

[I] See methods section

0 kcal/mol, the error in the ΔG_F will increase exponentially. Due to the logarithmic nature of the conversion of K_F to ΔG_F, you cannot take a standard deviation associated with K_F and convert it to a standard deviation around ΔG_F. However, it is possible to understand why error increases exponentially with ΔG_F using Figure 5.7, which shows the correlation between ΔG_F and K_F. As ΔG_F increases, the slope of the line grows increasingly shallow. In other words, as ΔG_F increases, smaller and smaller changes in K_F result in larger changes in ΔG_F.

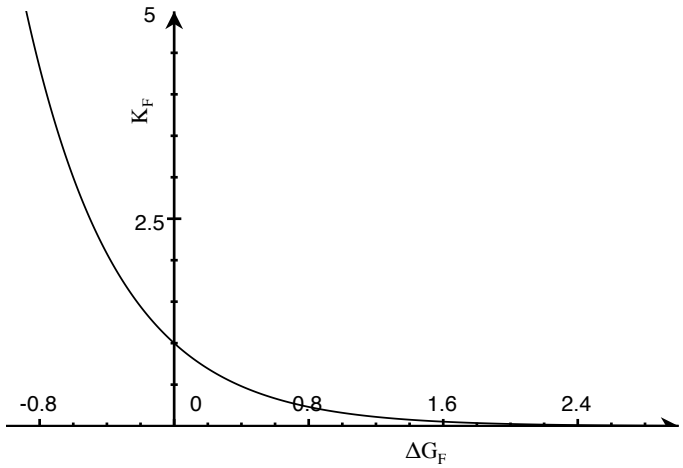


Figure 5.7. K_F as a function of ΔG_F at 275 K.

Between 0 kcal/mol and 0.5 kcal/mol, the change in the value of K_F is 0.6 (Table 5.1). However, between 1 and 1.5 kcal/mol, also a $\Delta\Delta G_F$ of 0.5 kcal/mol, the change in the value of K_F is only 0.1 (Table 5.1). Therefore, as the ΔG_F increases, one must measure K_F more and more accurately to get the same amount of error around the ΔG_F value. Assuming one can only measure K_F to an accuracy of about 0.1, one cannot distinguish a difference in the stability of two peptides with a $\Delta\Delta G_F$ of 0.5 kcal/mol when the ΔG_F of the peptides are above 1 kcal/mol. The problem of high error in ΔG_F grows exponentially as ΔG_F increases.

Table 5.1. ΔG_F and K_F values

ΔG_F (kcal/mol)	K_F	ΔK_F*
4	0.00067	-
3.5	0.0017	0.0010
3	0.0041	0.0025
2.5	0.010	0.006
2	0.026	0.015
1.5	0.064	0.039
1	0.16	0.10
0.5	0.40	0.24
0	1.0	0.6
-0.5	2.5	1.5
-1	6.2	3.7

*$\Delta K_F = K_F(\Delta G_F) - K_F(\Delta G_F + 0.5$ kcal/mol)

Given that ID peptides are likely to have high ΔG_F, often above 1 kcal/mol, ID BTE cannot be done with a traditional BTE assay. Instead, we explored the possibility of competition BTE.[21] Competition BTE was originally proposed by Matthew Woll.[21] Instead of comparing the stability of the desired full-length peptide to the stability of a thioester with a non-interacting thiol, one compares two full length peptides. The first peptide has a known stability (ΔG_1) and the second peptide has an unknown stability (ΔG_2) (Figure 5.8).

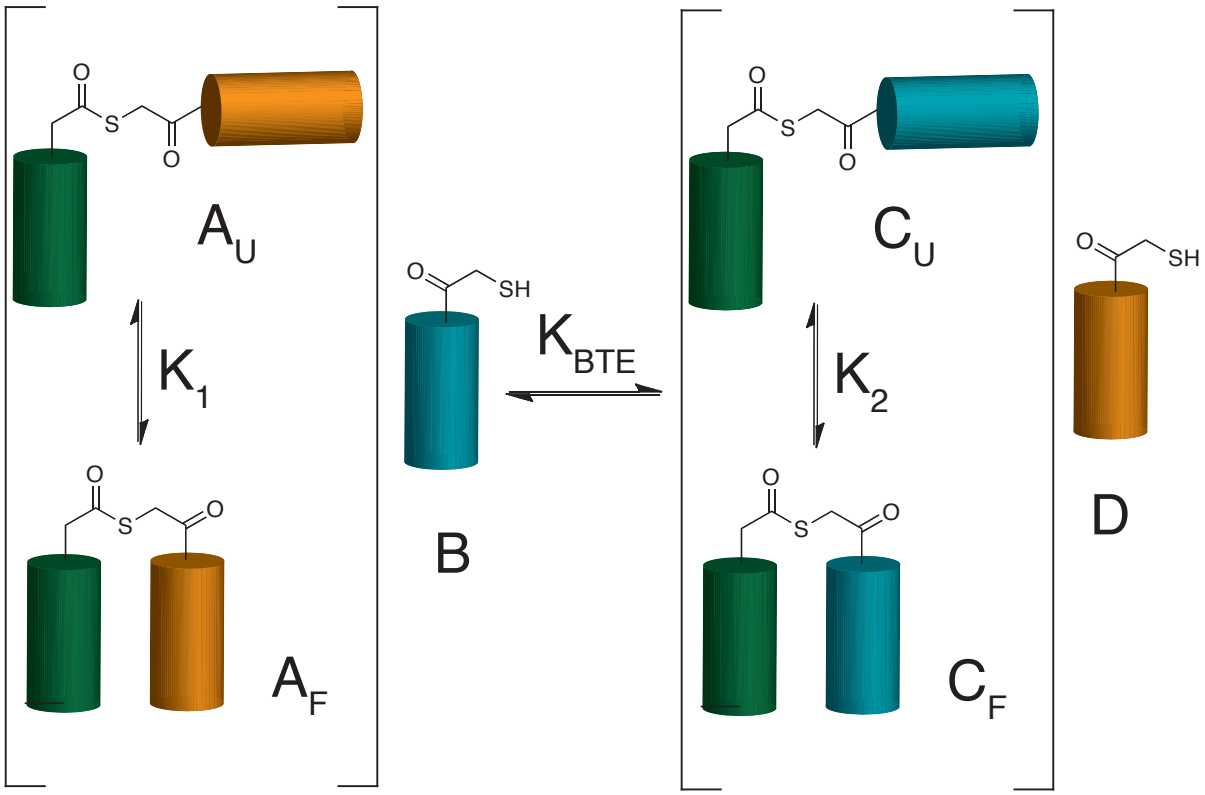


Figure 5.8 Competition BTE

The equation for K_{BTE} in a competition assay is[I]

$$K_{BTE}=\frac{K_2+1}{K_1+1} \qquad \text{Equation 5.10}^{21}$$

We can also write an equation for the difference between ΔG_1 and ΔG_2,

$$\Delta\Delta G_{2,1} = \Delta G_2 - \Delta G_1$$
$$= -RT\ln(K_2/K_1) \qquad \text{Equation 5.11}^{21}$$

Equation 5.11 allows us to define $\Delta\Delta G_F$ in terms of the ratio of K_2/K_1. If we can define K_2/K_1 in terms of the variable K_{BTE} alone, the absolute value of the unstable ΔG_2 will not impact the amount of error. Unfortunately, equation 5.10 does not rearrange to allow us to define K_2/K_1 independently of K_1. To illustrate this point, theoretical K_{BTE} were calculated (Table 5.2).

[I] See methods section for details

Table 5.2. Theoretical Competition K_{BTE} values

ΔG_1	ΔG_2	$\lvert\Delta\Delta G_{2,1}\rvert$	Comp. K_{BTE}	ΔK_{BTE}*
0	1	1	0.60	0.40
1	2	1	0.87	0.13
2	3	1	0.97	0.03
3	4	1	1	0.00

*ΔK_{BTE} is the difference between the competition K_{BTE} when $\Delta\Delta G_{2,1}$=0 and the specified $\Delta\Delta G_{1,2}$

If the competition K_{BTE} value was independent of ΔG_2 and ΔG_1, then the absolute competition K_{BTE} value would be the same for the absolute values $\Delta\Delta G_{2,1}$ that are the same. However, this is not the case. As the values of ΔG_1 and ΔG_2 increase, the dynamic range of K_{BTE} for $\Delta\Delta G_{2,1}$ between 0 kcal/mol and 1 kcal/mol becomes increasing small. So, as ΔG_2 becomes increasing large, one must measure K_{BTE} more and more precisely to get the same error in $\Delta\Delta G_{1,2}$. Similar theoretical calculations show similarly high uncertainties arise when holding the known ΔG_1 constant and increasing ΔG_2. In other words, competition does not solve the problem that inherently high $\Delta G_{Folding}$ values cannot be accurately measured by BTE.

5.3.3. Conclusions

It is unlikely that BTE can be used to measure meaningful thermodynamic data from intrinsically disordered systems. However, it should be noted that we are making the assumption that ID systems are inherently unstable. It is possible one could find a system in which the amorphous interactions are additively stabilizing to a degree that would allow the application of BTE. However, I would hypothesize that even if such a system exists, the size of it would be challenging in the context of BTE.

5.4. Toward a BTE System to Study the Interactions between D-Peptide Helices and L-Peptide Helices

5.4.1 Background

D-amino acids are the enantiomers of naturally occurring L-amino acids. Peptides made of D-amino acids (D peptides) have been known to bind to natural L proteins and L peptides. They are also protease-resistant, and therefore may have longer half lives in vivo. Therefore, there is great interest in developing D peptides as therapeutics.[22]

One challenge to developing D peptides as therapeutics is determining sequences that will bind to a desired target protein.[22,23] With a limited understanding of the interactions between L proteins and D peptides, it's difficult to rationally design a D-peptide inhibitor that binds to a specified protein target and blocks a protein-protein interaction, even when a known L-peptide inhibitor sequence exists.[22] Initially, Muller *et al*. proposed that one could find a D-peptide sequence from an L-peptide sequence by inverting the sequence and replacing the D amino acids with L amino acids.[24] However, this technique does not appear to work in most cases.[25] A deeper and more detailed understanding of the interactions of L and D peptides would potentially allow one to rationally design D-peptide inhibitors for L-protein targets.

An alternative approach to rational design is to determine the sequence of an inhibitor peptide with phage display.[26] In phage display, random peptide sequences are encoded and displayed on the surface of a phage. Phages with sequences that bind to the desired target protein are separated through a selection process and then amplified. The process of selection and amplification is repeated several times to identify sequences that tightly bind the target.[26]

Phages cannot encode D peptides. So, in mirror image phage display, D-protein targets are synthesized and used to determine the sequence of an L peptide that will bind. L peptides that

bind to a D protein target will have the same sequence as D peptide that bind to an L protein target. Therefore mirror image phage display indirectly allows the determination of the D peptide inhibitor sequence.[23] The major downside of this approach is that the protein target must be chemically synthesized. This greatly limits the size of the target protein and often requires intense synthetic efforts.[23]

Given the limitations of mirror image phage display and the potential utility of D-peptide inhibitors, understanding the interactions between D and L peptides and proteins would be useful. Given that helices are common in proteins, understanding the interaction between a D-peptide helix and an L-peptide helix would be useful.

The work of Kim *et al*. has demonstrated that D and L peptides helices can interact. Presumably, these peptides interact as a four-helix bundle, though no structure has been determined.[27] Some structural information exists for the interaction of D and L proteins or peptides. Currently there are a handful of crystal structures with D/L interfaces and a few show the interaction of a D-peptide helix with an L-peptide helix (Figure 5.9).[28–30]

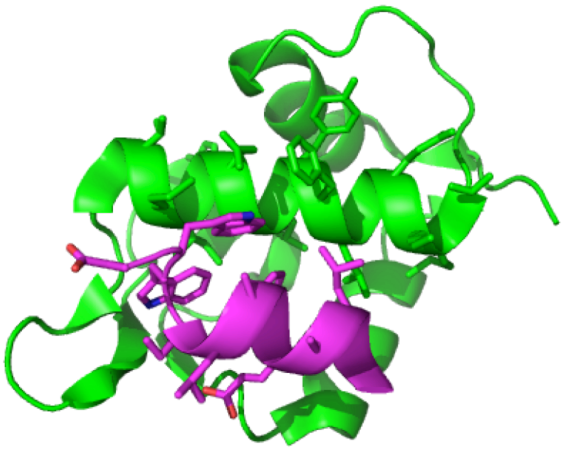


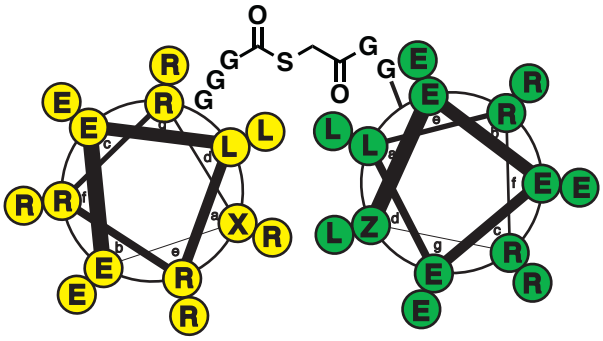
Figure 5.9 Crystal structure of D peptide interacting with an L protein MDM2 from Liu *et al*.[30]

As mentioned previously, one common way of exploring peptide-peptide interactions is chemical or thermal denaturation.[31] The limits of this technique have been discussed at length in chapter 1. However, when applied to D peptide and L peptide interactions, there is an additional limitation. Chemical and thermal denaturations are often monitored by CD.[31] The CD signatures of helical D and L peptides have almost opposite CD signatures.[27] In a solution with equal amounts of D-peptide and L-peptide helices, the signatures will cancel, or almost cancel each other.[27] Though denaturations of D and L peptides have been monitored with CD,[27] the results are questionable.

5.4.2 Toward a Backbone Thioester Exchange system to study interactions between D-peptide and L-peptide helices

Attempted Heptad Repeat Design

The initial helix-loop-helix system was designed based on a previous system in which a helix-loop-helix BTE system was used to quantify the energetic contributions of different residues in the core of a natural coiled coil.[7,6] The original system was modified by replacing the residues in the second helix with D-amino acids (Peptide 5.11, Figure 5.10).



Succ-EELRRRXEELRRRRGGGGGGLEEELRRZEEELRRY (5.11)
Succ-EELRRRXEELRRRRGGGGGGLRRLEEEZRRLEEEY (5.12)
X, Z= Variable Positions (A, I, L, N, or V); Green= D amino acids

Figure 5.10. Helical wheel and sequences of peptides for D/L BTE.

As discussed in detail in Chapter 1, helices have a heptad repeat and 3.6 residues per turn of the helix. The positions on the helix are designated a-g. Natural helices that associate as antiparallel dimers do so through their a and d positions. In the dimer, the positions on the first helix are designated a-g and the positions on the second helix are designated a'-g'. The a and d residues of the first helix fit into a 'holes' created by the display of the residues on the second helix. This packing is often called 'knobs into holes'. For example, every a residue has one d' residue next to it and a' residues above and below it.[32]

In an antiparallel helix-loop-helix where one helix is made of L-residues and the other of D-residues, Kim *et al*. suggests that there is a vertical offset of the D-peptide helix by one turn of the L-peptide helix.[27] In other words, instead of the a position packing across from the d' position, it is across from the a' position. The association between the two helices is still a 'knobs into holes' interaction. The predicted offset is observed in crystal structures from David Mortenson of the helical M2 protein associating with its enantiomer.[28] In order to account for this offset in the BTE system, a second set of D peptides were designed (Peptide 5.12, Figure 5.10).

Synthesis and Evaluation of Heptad Repeat Designs

Peptides were prepared by solid phase peptide synthesis (SPPS) as described be in the methods section. Synthesis of the thioesters by the Dawson[33] and sulfonamide[11] methods were not successful. Thioesters were synthesized using the backbone amide linker (oBAL) method.[12]

Before any meaningful thermodynamic measurements can be made, time studies must be done to determine when the BTE system reaches equilibrium. At each time point, the ratio between the concentration of the products and the concentration of the reactions (Q) is measured, and log(Q) is plotted. The systems reach equilibrium within an hour.

The next control is a concentration study. If the system is monomeric and the linker allows the favored association, then the K_{BTE} will not be concentration dependent. However, if peptides self-associate, K_{BTE} will probably be concentration dependent and the assumptions that allow the calculation of ΔG_F from K_{BTE} no longer apply. For example, if the helices fold into a four-helix bundle rather than a two-helix bundle, two copies of Peptide 5.11 will be present in the folded structure. As peptide concentration increases, a greater percentage of Peptide 5.11 will self-associate, which increases the concentration of C_F as defined in Figure 1.1. This will, in theory, increase the value of K_{BTE}. Results from all concentration studies show a dependence (Table 5.3), suggesting that the system is not monomeric.

Table 5.3 Concentration Dependence of D/L BTE Systems

Parent Sequence	Variable Position (X, Z)*	Total Peptide Concentration (μM)	Apparent $\Delta G_{Folding}$ (kcal/mol)
5.11	L, i	30	-0.7
		100	-0.7
		200	-0.6
		300	-0.4
5.11	L, n	30	0.1
		100	0.3
		200	0.5
		250	-0.1
5.11	L, v	30	-0.5
		100	-0.3
		200	-0.2
		250	-0.2
5.12	L, a	30	-0.5
		100	-0.2
		200	-0.2
5.12	L, n	30	0.3
		200	0.7
		300	1.8
5.12	L, v	30	1.0
		300	0.4

*Lowercase letters are D-amino acids

The results of the concentration study suggest that the system is not monomeric. Therefore, we cannot determine an accurate ΔG_F value for the helix-loop-helix structure. However, if there was no interaction between the D-residues and L-residues, K_{BTE} would be 1 for the BTE systems based on Peptide 5.11. (Note that interactions that could perturb K_{BTE} through interactions between all L-residues or all D-residues have been ruled out by controls in previous studies.)[6] The fact that K_{BTE} is greater than one demonstrates that there is an interaction between the D-residues and the L-residues.

Attempted 11 Residue Repeat Design and Evaluation

The original design of the BTE system assumes that a helix of D-residues and a helix L-residues can associate as a two-helix bundle with a heptad repeat. (The heptad repeat places hydrophobic residues at a and d positions. See Chapter 1 for further details). However, Kim *et al*. suggests D- and L-helices will prefer to associate as a four-helix bundle in aqueous solution. To my knowledge, the only existing crystal structures of D/L-two-helix bundles are David Mortenson's crystal structures of M2 with its enantiomer.[28] However, these structures were crystallized in lipids or micelles, which is a hydrophobic environment. If you map the heptad repeat interfaces, the a and d positions drift out of the core of the interaction between the two helices.[28] Due to this drift, if D/L-peptides adopted the same structure in aqueous solution with a heptad repeat and hydrophobic residues at a and d positions, hydrophobic residues would be exposed to the aqueous solvent, which would destabilize the structure. The four-helix bundle structure proposed by Kim *et al*. would keep the a and d positions in the core of the interaction. Therefore, forming a four-helix structure rather than a two-helix structure might be preferred.[27]

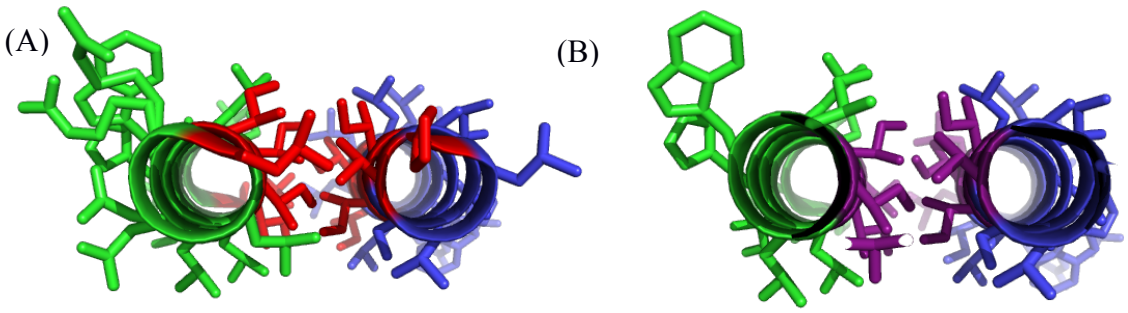


Figure 5.11 Crystal Structure of D/L helical interface from Mortensen.[28] (A) Residues in the core of the helix assigned a heptad repeat are in red. (B) Residues in the core of the helix assigned an 11-residue repeat are highlighted in purple.

The heptad (7-residue) repeat is the most common repeat found in natural two-helix bundle interfaces (see Chapter 1 for further details). However, 11-residue repeats have also been observed.[32] In these structures, the a, d and h positions are the residues in the core of the interaction between helices. If an 11-residue repeat is mapped onto the crystal structure of M2 and its enantiomer, the a, d, and h positions remain in the core of the interface between the helices (Figure 5.11b).[28] (Note that the a positions in the 11 residue repeat seem to be slightly more solvent exposed the rest of the core positions.) Therefore, an 11-residue repeat D/L helical BTE system was designed.

```
   hijkabcdefghijka       hijkabcdefghijka
Ac-IEERIRRIRAEIEERRGGGGGGIRREIEEIEARIRREEY   (Peptide 5.13)
```

Figure 5.10. Sequence for 11 residue repeat D/L Helical BTE system. The core positions are bolded. Green residues are D amino acids. The red residue indicates the position of the thioester bond.

A time study showed that the system came to equilibrium within an hour. A concentration study showed that the apparent ΔG_F was concentration dependent (Table 5.4). The fact that the K_{BTE} is greater than 1 suggests that the D and L residues may associate in some

manner. However, quantitative thermodynamic data cannot be derived from the system due to the fact that the full-length peptide does not appear to be monomeric.

Table 5.4 Concentration Dependence of D/L 11-residue repeat BTE System

Total Peptide Concentration (uM)	Apparent $\Delta G_{Folding}$ (kcal/mol)
100	-1.3
200	-1.9
300	-3.2

5.4.3 Conclusions and Future Directions

Due to the fact that the D/L BTE systems tested showed concentration dependence, no thermodynamic data could be obtained. Kim *et al*.[27] have suggested the D/L systems cannot be coiled-coil dimers. While Mortenson's crystal structure proves D/L two-helix bundles can form in lipids or micelles,[28] it does not demonstrate that structure can be formed in aqueous solution. Therefore, a successful BTE system my need to be a four-helix bundle rather that a two-helix bundle.

5.5 Conclusions from BTE experiments

BTE is a method well-suited to generating large amounts of data quickly, but there are many potential pitfalls. BTE systems must be carefully designed, and can be useful in the correct context.

5.6 Methods

General. Unless otherwise noted, all materials were obtained from Chemimpex, Novabiochem, Fischer, Peptech or Sigma-Aldrich. All HPLC data was acquired on Shimadzu

reverse phase instraments with C5 or C18 revers-phase columns with a gradient of acetonitrile/water with 0.1% TFA. The analytical HPLC had a flow rate of 1 ml/min, and the preparative HPLC had a flow rate of 12-15 ml/min.

5.5.1. Synthesis

Peptide Synthesis. Unless otherwise noted, all peptides were synthesized by standard Fmoc solid phase peptide synthesis (SPPS) on NovaPEG rink amide resin, NovaPEG TGR resin, or Chemmatrix rink amide resin. Unless otherwise noted, all amino acids were obtained from Chemimpex, Peptech, or Novabiochem. Fmoc protected versions APC and ACPC for SPPS were synthesized as previously described.[34] Unless otherwise noted, peptides were synthesized on a CEM Mars microwave with a temperature probe. Residues were coupled onto the resin from the C-terminal to the N-terminal with iterative coupling and deprotection steps. In each coupling step, 4 eq of amino acid was activated with 4 eq coupling reagent, 4 eq HOBt or Cl-HOBt, and 8 eq DIEA in DMF or NMP (1 ml for a 25 μmol scale) for at least 1 minute. The coupling reagent was PyBOP, HCTU or HBTU. The coupling solution was added to the resin and to react on the microwave (2 minute ramp to 70ºC, 4 minute hold at 70ºC). The resin was washed with DMF. In each deprotection step, 20% piperidine in DMF was added (>1 ml/25 umol). The solution was microwaved (2 min ramp to 80ºC, 2 min hold at 80ºC). (Tritylsulfanyl)acetic acid was coupled onto the C-terminus of the thiol peptides using the same procedure as the protected amino acids. (Tritylsulfanyl)acetic acid was synthesized as previously described.[3] Peptides were cleaved from the resin and deprotected with an acidic deprotection solution for 1-6 hours. Unless the peptide contained a thiol, the deprotection solution was H_2O (0.1mL/50 μmol scale), triisopropylethylsilane (0.1 mL/50 μmol scale), and trifluoroacetic acid (3.8 ml/50 μmol scale).

For the thiol peptides, the deprotection solution was H_2O (50 μl/25 μmol scale), triisopropylethylsilane (40 μl/25 μmol scale), 1,2-ethanedithiol (50 μl/25 μmol scale), and trifluoroacetic acid (2 ml/25 μmol scale). The peptides were precipitated with cold ether and purified by reverse phase HPLC with gradients of acetonitrile and water with 0.1% trifloroacetic acid on a Shimadzu prep HPLC with a C5 or C18 column. All peptides were characterized by analytical HPLC and found to be greater than 90% pure. The masses of all peptides were checked by MALDI-MS on a TOF Bruker Reflex II with a 337 nm nitrogen laser or a TOF/TOF Bruker Ultraflex III with a SmartBeam laser. All purity checks and MALDI spectra can be found in the appendix.

Thioester Peptide Synthesis. All thioester peptides were synthesized by the sulfonamide[11] or backbone amide-linker methods[12] as described below.

For the sulfonamide "Safety Catch" method, peptides were synthesized by standard SPPS as described above on H-Gly-Sulfamylbutyryl AM resin. Iodoacetonitrile was filtered over basic alumina. A solution of diisopropylethyamine (3 M) and iodoacetonitrile (0.44 M) in DMF (3 ml/25 μmol scale) was added to the resin, and the resin was microwaved for one hour at 50ºC. The resin was washed with DMF. Methylthioglycolate (27.3 μl/25 μmol scale) and DMF (3 ml/25 μmol scale) were added to the resin, and the resin was microwaved for 1 hour at 50ºC. The solution was drained into a round bottom flask, and the resin was rinsed with DCM. The solvent was removed under reduced pressure in the hood. The resulting residue was dissolved in a solution of trifloroacetic acid (3.8 ml), triisopropylethylamine (0.1 ml), and water (0.1 ml) and stirred for 2-4 hours. The peptide was precipitated with cold ether and purified by HPLC as described above. The resulting thioester was dissolved in BTE buffer (50 mM HEPES, 2 mM

TCEP, pH=7.0) and excess tyrosine small thiol 5.1 (synthesized as previously described[3]) was added. Thioester exchanged was allowed to occur and the resulting thioester was purified by HPLC.

For the trithioorthoester method, the backbone amide linker was synthesized and loaded onto resin as previously described.[12] The thioester then synthesized as previously described.[12]

5.5.2 Methods for α/β BTE

Backbone Thioester Exchange Assay. A freshly prepared solution of thioester and thiol peptides (species A and B in Figure 5.1) in buffer (50 mM HEPES and 2 mM TCEP, pH=7.0) was injected onto an analytical HPLC. The relative concentration of the four species (Figure 5.1) was measured by taking the integration of the four peaks 275 nm and adjusting for the number of tyrosine residues present. During the first BTE run for a given set of peptides, BTE measurements are taken at various time points in order to determine the approximate time the system needs to equilibrate. After the initial BTE run, all BTE measurements are taken after the system has reach equilibrium. The K_{BTE}, K_F, and ΔG_F values of the system are calculated with equations 5.1, 5.6, and 5.7, respectively.

Thioester Degradation Assay. Thioester 5.2 was dissolved in buffer (50 mM HEPES, 2 mM TCEP, pH=7.0) and monitored over time by analytical HPLC. The integration of the thioester peak decreased over time and the integration of the small thiol peak increased (Table 5.4). The percent of thioester degraded was estimated by dividing the integration of the thioester peak at 275 nm over the integration of both peaks at 275 nm and multiplying by 100 (Table 5.4).

Table 5.4. Thioester Degredation at pH=7.0

Time (hours)	Thioester Integration	Small Thiol Integration	Thioester Degraded (%)
0	248011	17402	6.6
1.03	198465	63722	24
1.5	180827	77480	30
1.95	170023	98105	37
2.42	159268	100858	39

The remaining BTE solution was centrifuged and a white precipitate was isolated. The precipitate was dissolved in acetonitrile and injected onto the analytical HPLC. Three peaks were observed. The thioester degradation experiment was also done at pH=6.5 and no degradation was observed.

5.5.3. Methods for ID BTE

Backbone Thioester Exchange Assays. Unless otherwise noted, BTE assays for the tVHP system were performed as described in section 5.5.2. Phosphate buffer (50 mM, pH 7.0, 2 mM TCEP) was used instead of HEPES buffer. Integrations were measured at 220 nm instead of 275 nm as the thiol and full length peptides do not contain a Tyr. K_{BTE} was calculated using equation 5.8. The correction factor E can be calculated using equation 5.9. The derivation for equation 5.8 was originally developed by Matt Woll[21] and is as follows. Using the Beer-Lambert Law

$$A=\varepsilon cl \qquad \text{Equation 5.12}$$

where A is absorbance, ε is the extinction coefficient, c is the concentration, and l is the pathlength. Also,

$$A=\text{Area} \times B \qquad \text{Equation 5.13}$$

where area is the integration of the HPLC peaks and B is a coefficient that converts absorbance to area. If we combine equation 5.12 and 5.13 we get

$$c=\frac{\text{Area} \times B}{\varepsilon \times l} \qquad \text{Equation 5.14}$$

If we measure the integration of each HPLC peak at two time points during an HPLC experiment:

$$\Delta c=\frac{\Delta\text{Area} \times B}{\varepsilon \times l} \qquad \text{Equation 5.15}$$

which rearranges to give

$$\varepsilon=\frac{\Delta\text{Area} \times B}{\Delta c \times l} \qquad \text{Equation 5.16}$$

Equation 5.14 and 5.16 combine to give

$$c=\frac{\text{Area} \times \Delta c}{\Delta\text{Area}} \qquad \text{Equation 5.17}$$

Recall

$$K_{BTE}=\frac{[C][D]}{[A][B]} \qquad \text{Equation 5.1}$$

Also, in a thioester exchange reaction,

$$\Delta[A]=\Delta[B]=-\Delta[C]=-\Delta[D] \qquad \text{Equation 5.18}$$

where species A, B, C, and D as defined in Figure 5.1. Equations 5.1, 5.17, and 5.18 give

$$K_{BTE}=\frac{\frac{\text{Area}_C}{\Delta\text{Area}_C}\times\frac{\text{Area}_D}{\Delta\text{Area}_D}}{\frac{\text{Area}_A}{\Delta\text{Area}_A}\times\frac{\text{Area}_B}{\Delta\text{Area}_B}} \qquad \text{Equation 5.19}$$

Define E as

$$E=\left(\frac{\Delta\text{Area}_A\times\Delta\text{Area}_B}{\Delta\text{Area}_C\times\Delta\text{Area}_D}\right) \qquad \text{Equation 5.20}$$

Combine Equation 5.19 and 5.20 to give

$$K_{BTE}=\left(\frac{\text{Area}_C\times\text{Area}_D}{\text{Area}_A\times\text{Area}_B}\right) \text{ x } E$$

$$= K_{meas} \times E \qquad \text{Equation 5.21}$$

Note that E is a constant that can be measured (Equation 5.20) by taking the difference in integration between HPLC peaks measured directly after the peptides were mixed and again after the system has reached equilibrium.

Turbidity Assays. Versions of Peptides 5.7 and 5.8 were the thioester bond was replaced with an amide bond were dissolved in phosphate buffer (50 mM, pH 7.0) at various concentrations (see Figure 5.11). A UV spectrum from 200-375 nm was recorded and the sample sat at room temperature for 15 to 24 hours. The samples were centrifuged and the solution was decanted. The UV spectrum of each solution was recorded again.

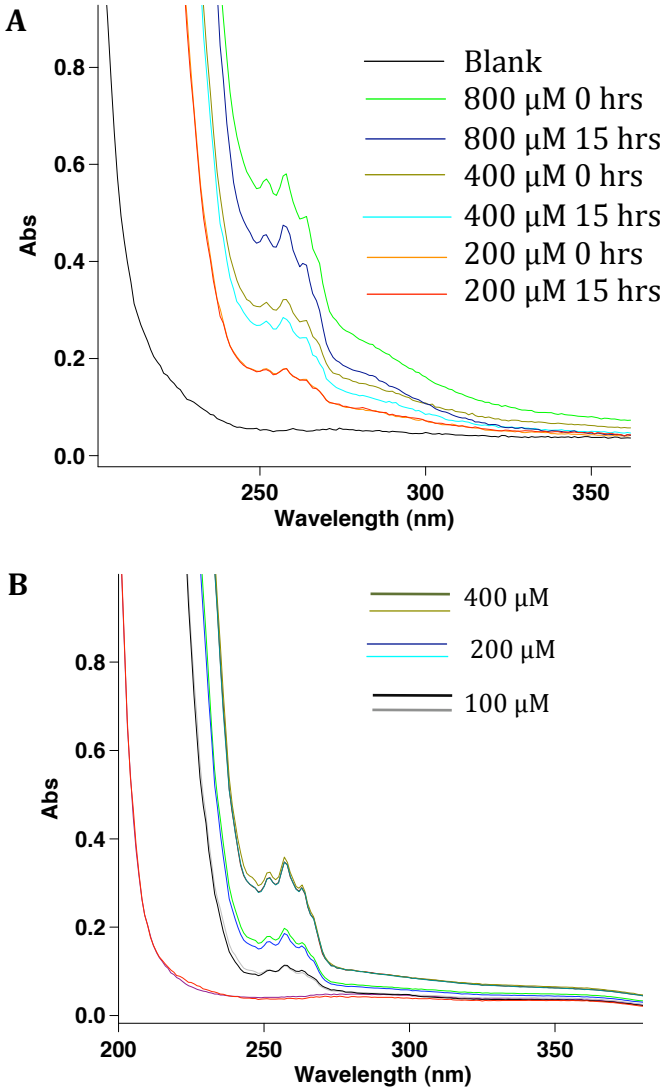


Figure 5.11. Turbidity experiment. (A) Peptide 5.7 at 0 and 15 hours (B) Peptide 5.8 at 0 and 24 hours.

Theoretical Competition BTE Calculations. Competition BTE was originally proposed by Matt Woll.[21] By definition (Figure 5.6),

$$K_{BTE}=\frac{[C_U+C_F][D]}{[A_U+A_F][B]} \qquad \text{Equation 5.22}$$

$$K_1=\frac{[A_F]}{[A_U]} \qquad \text{Equation 5.23}$$

$$K_2=\frac{[C_F]}{[C_U]} \qquad \text{Equation 5.24}$$

Equation 5.22-5.24 combine to give

$$K_{BTE}=\frac{(K_2+1)[C_U][D]}{(K_1+1)[A_U][B]} \qquad \text{Equation 5.25}$$

Assume the thioester bonds are isoenergetic so that

$$\frac{[C_U][D]}{[A_u][B]}=1 \qquad \text{Equation 5.26}$$

Combine equations 5.25 and 5.26 to give

$$K_{BTE}=\frac{K_2+1}{K_1+1} \qquad \text{Equation 5.10}$$

Given a known value of K_1, K_2 can be calculated by measuring a competition K_{BTE}.

5.5.4. Methods for D/L BTE

Backbone thioester exchange assays. BTE assays were done as described in section 5.5.2. with the following modifications. Assays were done phosphate buffer (50 mM, 2 mM TCEP, pH=7.0).

5.7 References

(1) Maxwell, K. L.; Wildes, D.; Zarrine-Afsar, A.; De Los Rios, M. A.; Brown, A. G.; Friel, C. T.; Hedberg, L.; Horng, J.; Bona, D.; Miller, E. J.; et al. Protein Folding : Defining a "standard" Set of Experimental Conditions and a Preliminary Kinetic Data Set of Two-State Proteins. *Protein Sci.* **2005**, *14*, 602–616.

(2) Johnson, L. N.; Barford, D. The Effects of Phosphorylation on the Structure and Function of Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 199–232.

(3) Woll, M. G.; Gellman, S. H. Backbone Thioester Exchange: A New Approach to Evaluating Higher Order Structural Stability in Polypeptides. *J. Am. Chem. Soc.* **2004**, *126*, 11172–11174.

(4) Steinkruger, J. D.; Woolfson, D. N.; Gellman, S. H. Side-Chain Pairing Preferences in the Parallel Coiled-Coil Dimer Motif: Insight on Ion Pairing between Core and Flanking Sites. *J. Am. Chem. Soc.* **2010**, *132*, 7586–7588.

(5) Price, J. L.; Hadley, E. B.; Steinkruger, J. D.; Gellman, S. H. Detection and Analysis of Chimeric Tertiary Structures by Backbone Thioester Exchange: Packing of an Alpha Helix against an Alpha/beta-Peptide Helix. *Angew. Chemie Int. Ed.* **2010**, *49*, 368–371.

(6) Hadley, E. B.; Testa, O. D.; Woolfson, D. N.; Gellman, S. H. Preferred Side-Chain Constellations at Antiparallel Coiled-Coil Interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 530–535.

(7) Hadley, E. B.; Gellman, S. H. An Antiparallel Alpha-Helical Coiled-Coil Model System for Rapid Assessment of Side-Chain Recognition at the Hydrophobic Interface. *J. Am. Chem. Soc.* **2006**, *128*, 16444–16445.

(8) Dawson, P. E.; Kent, S. B. H. Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.

(9) Pomerantz, W. C.; Hadley, E. B.; Fry, C. G.; Gellman, S. H. In Situ Monitoring of Backbone Thioester Exchange by 19-F NMR. *Chembiochem* **2009**, *10*, 2177–2181.

(10) Martinek, T. a; Fülöp, F. Peptidic Foldamers: Ramping up Diversity. *Chem. Soc. Rev.* **2012**, *41*, 687–702.

(11) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A.; Km, V. P. Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc / T-Bu Chemistry. *J. Am. Chem. Soc.* **1999**, 11369–11374.

(12) Brask, J.; Albericio, F.; Jensen, K. J. Fmoc Solid-Phase Synthesis of Peptide Thioesters by Masking as Trithioorthoesters. *Org. Lett.* **2003**, *5*, 2951–2953.

(13) Shin, Y. Thesis, University of Wisconsin-Madison, 2015.

(14) Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H. Structural and Biological Mimicry of Protein Surface Recognition by Alpha/beta-Peptide Foldamers. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14751–14756.

(15) Oldfield, C. J.; Dunker, A. K. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584.

(16) McKnight, J.; Matsudaira, P. T.; Kim, P. S. NMR Structure of the 35-Residue Villin Headpiece Subdomain. *Nat. Struct. Biol.* **1997**, *4*, 180–184.

(17) Tang, Y.; Rigotti, D. J.; Fairman, R.; Raleigh, D. P. Peptide Models Provide Evidence for Significant Structure in the Denatured State of a Rapidly Folding Protein: The Villin Headpiece Subdomain. *Biochemistry* **2004**, *43*, 3264–3272.

(18) Tang, Y.; Goger, M. J.; Raleigh, D. P. NMR Characterization of a Peptide Model Provides Evidence for Significant Structure in the Unfolded State of the Villin Headpiece Helical Subdomain. *Biochemistry* **2006**, *45*, 6940–6946.

(19) Meng, W.; Shan, B.; Tang, Y.; Raleigh, D. P. Native like Structure in the Unfolded State of the Villin Headpiece Helical Subdomain, an Ultrafast Folding Protein. *Protein Sci.* **2009**, *18*, 1692–1701.

(20) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. Stabilizing and Destabilizing Effects of Phenylalanine --> F5-Phenylalanine Mutations on the Folding of a Small Protein. **2006**, *53706*, 15932–15933.

(21) Woll, M. G. Thesis, University of Madison-Wisconsin, 2004.

(22) Zhao, L.; Lu, W. Mirror Image Proteins. *Curr. Opin. Chem. Biol.* **2014**, *22C*, 56–61.

(23) Sun, N.; Funke, S. A.; Willbold, D. Mirror Image Phage Display--Generating Stable Therapeutically and Diagnostically Active Peptides with Biotechnological Means. *J. Biotechnol.* **2012**, *161*, 121–125.

(24) Guichard, G.; Benkirane, N.; Zeder-lutz, G.; Regenmortel, M. H. V. Van; Briand, J.-P.; Muller, S. Antigenic Mimicry of Natural L-Peptides with Retro-Inverso- Peptidomimetics. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 9765–9769.

(25) Fischer, P. M. The Design, Synthesis and Application of Stereochemical and Directional Peptide Isomers: A Critical Review. *Curr. Protein Pept. Sci.* **2003**, *4*, 339–356.

(26) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E.; Panganiban, C. Random Peptide Libraries : A Source of Specific Protein Binding Molecules. **2015**, *249*, 404–406.

(27) Sia, S. K.; Kim, P. S. A Designed Protein with Packing between Left-Handed and Right-Handed Helices †. *Biochemistry* **2001**, *40*, 8981–8989.

(28) Mortenson, D. Thesis, University of Wisconsin-Madison, 2014.

(29) Liu, M.; Pazgier, M.; Li, C.; Yuan, W.; Li, C.; Lu, W. A Left-Handed Solution to Peptide Inhibition of the p53-MDM2 Interaction. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 3649–3652.

(30) Liu, M.; Li, C.; Pazgier, M.; Li, C.; Mao, Y.; Lv, Y.; Gu, B.; Wei, G.; Yuan, W.; Zhan, C.; et al. D-Peptide Inhibitors of the p53-MDM2 Interaction for Targeted Molecular Therapy of Malignant Neoplasms. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14321–14326.

(31) Street, T. O.; Courtemanche, N.; Barrick, D. Protein Folding and Stability Using Denaturants. *Methods Cell Biol.* **2008**, *84*, 295–325.

(32) Lupas, a. Coiled Coils: New Structures and New Functions. *Trends Biochem. Sci.* **1996**, *21*, 375–382.

(33) Blanco-Canosa, J. B.; Dawson, P. E. An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 6851–6855.

(34) Appella, D. H.; Christianson, L. a.; Klein, D. a.; Richards, M. R.; Powell, D. R.; Gellman, S. H. Synthesis and Structural Characterization of Helix-Forming Beta-Peptides: Trans-2-Aminocyclopentanecarboxylic Acid Oligomers. *J. Am. Chem. Soc.* **1999**, *121*, 7574–7581.

Appendix A
Supporting Information for Chapter 2

Table of Contents

Supplementary Section A1
Summary of Cross-Strand NOEs

General. For details on sample preparation and data acquisition, please see the methods section of Chapter 2.

Figure A1.1. Cross-strand ROEs observed in parent peptide 2.10a at 4ºC in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8) from Syud et al.[1] Strong and medium ROEs are black and weak ROEs are green.

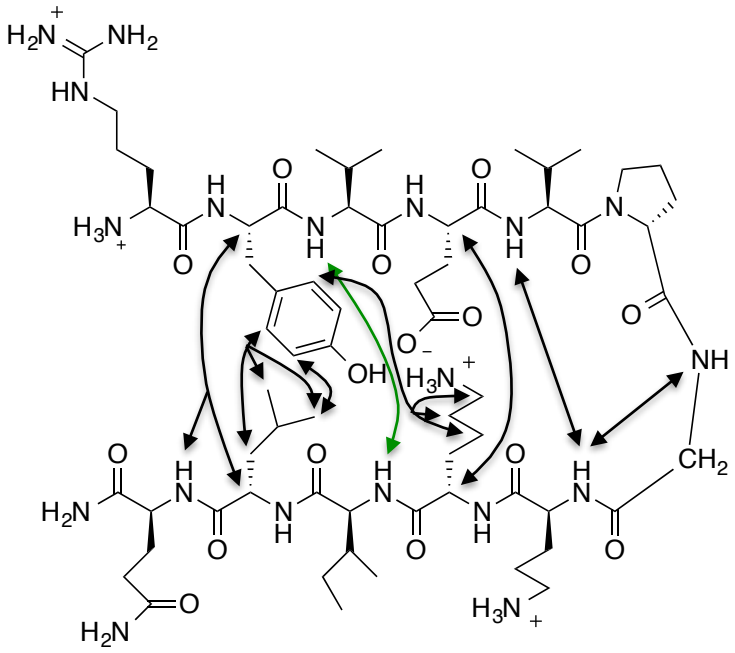


Figure A1.2. Cross-strand ROEs observed in parent hairpin 2.15a derived from Haque *et al.*[2] at 4°C in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8). Medium NOEs are blue. There is an additional strong or medium NOE(s) from Lys10 to Ser7 and/or Gly9, which is consistent with the backbone hydrogen bonding pattern of the expected hairpin structure. In general, fewer ROEs are observed for this parent peptide and its derivatives than the other parent sequences. It is probable that this peptide is not as well folded as the sequence derived from Syud et al.[1] However, the observed ROEs and $\Delta\delta C_\alpha H$ values support that the sequence does form a hairpin.

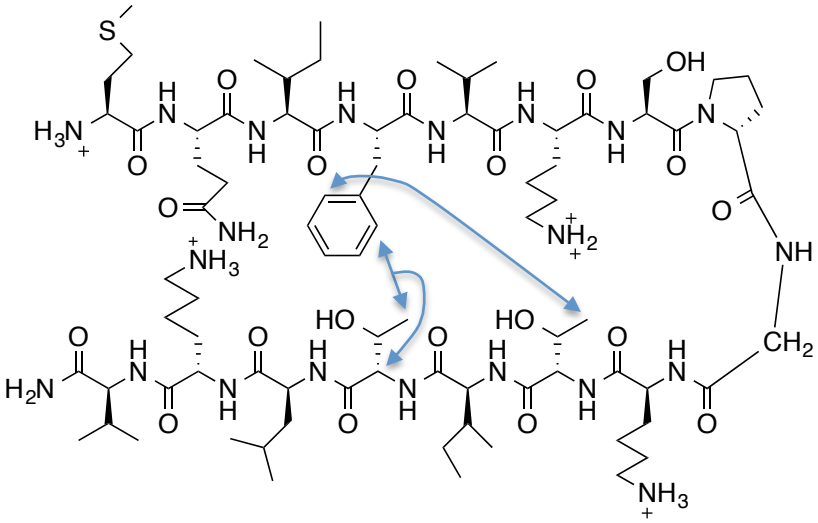
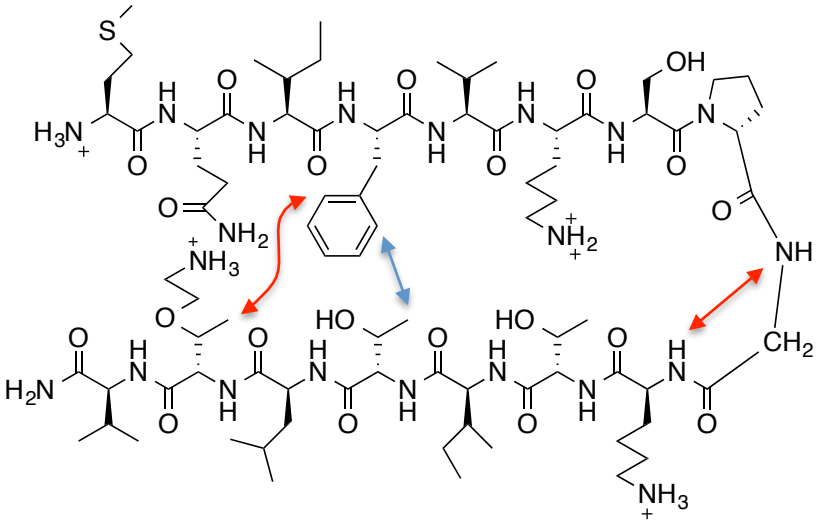


Figure A1.3. Cross-strand ROEs observed for peptide 2.16a (Lys15 → TO^+, 3.6 mM) at 4°C in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8). Strong ROEs are red and medium ROEs are blue.



Supplementary Section A2
Chemical Shifts of Peptides from 2D NMR

General. Details for the sample preparation and acquisition are in the methods section of Chapter 2. All unfolded controls are the peptide sequence of interest with LPro in place of DPro. See Chapter 2 for further details. All cyclic peptides (fully folded controls) are the sequence of interest cyclized with a DPro-Gly turn (See Chapter 2 for further details).

Table A2.1. ^{1}H chemical shifts of peptide 2.11a (Lys9 → TO^{+}, 0.6 mM) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Sidechain
Arg1		4.021	1.917	1.56	3.199		7.256
Tyr2	9.003	5.027	2.864				6.832
			2.916				7.035
Val3	8.775	4.285	1.994	0.843			
Glu4	8.596	4.787	1.914	2.28			
Val5	8.768	4.565	1.988	0.934			
DPro6		4.396	1.993	2.044	3.853		
			2.374	2.124			
Gly7	8.678	3.815					
		3.997					
Orn8	8.028	4.643	1.865	1.713	3.022		7.684
TO^{+}9	8.725	4.742	3.758	1.094	3.365	3.047	
					3.628		
Ile10	8.933	4.417	1.849	0.886	0.8138		
				1.160			
Leu11	8.667	4.276	1.607	1.491	0.7218		
Gln12	8.555	4.301	2.317	1.887			7.283
				2.049			7.565
Cterm	7.247						
	7.835						

Table A2.2. ^{1}H chemical shifts of cyclic peptide 2.11c (cyclic Lys9 → TO^{+}, 6.5 mM) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1	7.789	4.627	1.813	1.566	3.182		7.184
Tyr2	8.744	5.033	2.656				6.781
			2.811				6.957
Val3	9.321	4.483	2.025	0.847			
				0.885			
Glu4	8.689	4.985	1.915	2.281			
			1.997				
Val5	9.009	4.614	1.943	0.889			
				0.914			
DPro6		4.385	1.983	2.063	3.827		
			2.386	2.157	3.897		
Gly7	8.812	3.819					
		4.032					
Orn8	7.934	4.752	1.822	1.702	3.015		7.697
			1.876				
TO^{+}9	8.84	5.074	3.714	1.102	3.368	3.034	
					3.657		
Ile10	9.365	4.645	1.849	0.887	0.817		
				1.105			
				1.348			
Leu11	8.579	4.329	1.182	1.115	0.512		
			1.578		0.178		
Gln12	9.232	4.868	1.813	2.194			7.464
			2.003	2.241			
DPro13		4.305	1.941	2.009	3.715		
			2.325	2.141	3.786		
Gly14	8.555	3.8135					

Table A2.3. ^{1}H chemical shifts of LPro diasteromer (unfolded control) peptide 2.11b (unfolded Lys9 → TO^{+}, 15 mM) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.024	1.891	1.582	3.195		7.258
Tyr2	8.893	4.622	2.913				6.803
			3.029				7.103
Val3	8.149	3.959	3.959	0.843			
Glu4	8.443	4.209	1.984	2.367			
				2.421			
Val5	8.491	4.403	2.062	0.953			
LPro 6		4.409	1.933	1.991	3.696		
			2.319	2.053	3.877		
Gly 7	8.618	3.954					
Orn 8	8.398	4.461	1.796	1.718	3.027		7.705
			1.887	1.782			
TO^{+}9	8.576	4.483	3.97	1.187	3.631	3.179	7.749
					3.813		
Ile 10	8.45	4.176	1.821	1.167	0.886		
				1.485			
Leu 11	8.589	4.383	1.634	1.581	0.867		
					0.930		
Gln 12	8.553	4.282	1.976	2.371			6.984
			2.094				7.662
C Term.	7.232						
	7.745						

Table A2.4. ^{1}H chemical shifts of parent peptide 2.15 (2 mM) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	NH	Ha	Hb	Hg	Hd	He	Side Chain
Met1		4.15	2.153	2.558			
Gln2	8.840	4.512	1.806	2.110			7.011
			1.857	2.179			7.570
Ile3	8.516	4.189	1.755	0.822	0.789		
				1.081			
				1.381			
Phe4	8.562	5.064	2.883				7.187
			2.965				7.296
							7.236
Val5	8.757	4.243	1.998	0.877			
Lys6	8.621	4.661	1.751	1.402	1.646	2.965	7.566
DPro7		4.868	1.981	2.030	3.791		
			2.334	2.083			
Ser7	8.725	4.422	3.725				
Gly9	8.717	3.926					
Lys10	8.131	4.565	1.828	1.421	1.683	2.983	7.618
Thr11	8.587	4.56	4.007	1.116			
Ile12	8.933	4.436	1.892	0.881	0.824		
				1.408			
Thr13	8.503	4.534	4.022	1.080			
Leu14	8.634	4.414	1.528	1.524	0.827		
					0.885		
Lys15	8.524	4.409	1.731	1.344	1.658	2.976	7.617
			1.788				
Val16	8.346	4.095	1.998	0.926			
C Term	7.229						
	7.883						

Table SA2.5. ^{1}H chemical shifts of LPro diasteromer (unfolded control, 4 mM) of 2.15 at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.142	2.158	2.582			
Gln2	8.865	4.361	1.911	2.182			7.008
							7.631
Ile3	8.45	4.123	1.763	0.837	0.8404		
				1.757			
Phe4	8.643	4.669	2.979				7.217
			3.049				7.310
							7.264
Val5	8.276	4.012	1.934	0.8833			
Lys6	8.534	4.234	1.781	1.442	1.705	2.997	7.614
Ser7	8.597	4.727	3.853				
LPro8		4.438	1.976	2.035	3.745		
			2.295		3.814		
Gly9	8.511	3.898					
		3.986					
Lys10	8.367	4.403	1.828	1.434	1.742	2.994	7.613
Thr11	8.431	4.313	4.108	1.184			
Ile12	8.527	4.248	1.858	0.889	0.8525		
				1.189			
				1.488			
Thr13	8.486	4.313	4.103	1.185			
Leu14	8.502	4.361	1.627	1.559	0.858		
					0.919		
Lys15	8.512	4.338	1.801	1.406	1.672	2.991	7.611
Val16	8.359	4.067	2.027	0.9537			
Cterm	7.24						
	7.871						

Table A2.6. ^{1}H chemical shifts of peptide 2.16 (Lys15 → TO^{+}, 3.6 mM) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	NH	Ha	Hb	Hg	Hd	He	Side Chain
Met1		4.161	2.151	2.567			
Gln2	8.839	4.501	1.805	2.116			7.018
			1.850	2.189			7.571
Ile3	8.5	4.176	1.754	0.827	0.7956		
				1.076			
				1.372			
Phe4	8.564	5.073	2.908				7.190
			2.984				7.248
							7.301
Val5	8.765	4.24	2.004	0.8885			
Lys6	8.624	4.671	1.747	1.419	1.642	2.94	7.569
Ser7	8.73	4.874	3.733				
DPro8		4.424	1.993	2.035	3.801		
			2.342	2.081			
Gly9	8.726	3.897					
		3.941					
Lys10	8.135	4.578	1.801	1.429	1.688	2.993	7.621
			1.859				
Thr11	8.6	4.588	4.015	1.125			7.712
Ile12	8.954	4.435	1.902	0.890	0.8329		
				1.143			
Thr13	8.519	4.558	4.021	1.079			
Leu14	8.716	4.55	1.573	1.568	0.8538		
					0.902		
TO^{+}15	8.545	4.625	4.018	1.182	3.605	3.155	
					3.822		
Val16	8.311	4.122	2.008	0.942			
C Term.	7.260						
	7.918						

Table A2.7. ^{1}H chemical shifts of peptide 2.13 (Lys8 → TO^{+}, 5.6 mM) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.06	1.908	1.598	3.16		7.225
Trp2	8.866	4.9	2.999				10.190, 7.164
			3.061				7.320, 7.049
							7.217, 7.453
Gln3	8.761	4.47	1.877	2.115			6.973, 7.469
			1.974				
Tyr4	8.65	4.79	2.732				6.800 6.707
			2.870				
Val5	8.62	4.54	1.995	0.898			
				0.926			
DPro6		4.39	1.966	2.05	3.783		
			2.359				
Gly7	8.291	4.114					
		3.733					
Thr8	7.95	4.71	3.951	1.176	3.598	3.14	7.651
					3.780		
Phe9	8.893	4.71	2.813				6.898
			2.901				7.075
							7.029
TO^{+}10	8.605	4.51	4.028	1.147			
Val11	8.359	3.9	1.779	0.732			
				0.785			
Gln12	8.444	4.28	1.850	2.282			7.266
			2.045				7.763
Cterm	6.952						
	7.530						

Table A2.8. ^{1}H chemical shifts of peptide 2.14 (Thr10 → TO^{+}, 1.9 mM) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.039	1.895	1.587	3.166		7.221
Trp2	8.902	4.781					10.16 7.124
							7.382 7.061
							7.208 7.461
Gln3	8.641	4.388	1.884	2.084			6.935 7.463
			1.967				
Tyr4	8.421	4.733	2.722				6.883 6.712
			2.898				
Val5	8.68	4.469	1.995	0.9015			
DPro6		4.381	1.942	2.009	3.709		
			2.322				
Gly7	8.333	3.782					
		3.982					
Lys8	7.93	4.438	1.735	1.347	1.638	2.94	7.611
Phe9	8.561	4.714	2.943				6.988 7.182
			2.837				7.063
TO^{+}10	8.507	4.577	3.791	1.067	3.582	3.119	7.709
					3.722		
Val11	8.136	3.983	1.852	0.806			
				0.726			
Gln12	8.512	4.241	1.860	2.282			6.937 7.556
			2.01				
C Term	7.250						
	7.797						

Section A3
NMR Structure Calculations

Table A3.1. NMR statistics from CNS calculations for the cyclic hairpin 2.11c (cyclic Lys9 → TO^{+}). See Chapter 2 methods section for further detail.

NMR Distance and Dihedral Constraints	Protein
Distance Constraints	
Total NOEs	34
Intra-residue	0
Inter-residue	34
Sequential (|*i-j*|=1)	17
Medium-range (|*i-j*|<4)	4
Long-range (|*i-j*|>5)	13
Intermolecular	0
Hydrogen bonds	0
Total dihedral angle restraints	
ϕ	0
φ	0
Structure statistics	
Violations (mean and s.d.)	0
Distance constraints (A)	0
Dihedral angle constraints (^{O})	0
Max. dihedral angle violation (^{O})	0
Max. distance constraint violation (A)	0
Deviations from idealized geometry	
Bond lengths (A)	±0.0016
Bond angles (^{O})	±0.314
Impropers (^{O})	±0.49
Average pairwise r.m.s. deviation (A) (16 structures)	
Heavy	3.3 ± 0.7
Backbone	1.3 ± 0.4

Supplementary Section A4
Dilution Studies

General. Dilution Studies were performed for all peptides in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8) at 4ºC to check for aggregation. The most concentrated samples were the concentration used for acquisition of the 2D NMR spectra. Water suppression was performed with a presaturation pulse sequence as described in the methods section of Chapter 2. Where needed, baselines were corrected by manually defining baseline regions. In all cases, chemical shifts in the spectra of the concentrated and diluted peptide were the same, which indicates that the peptide aggregation state did not change over this concentration range. We assume that each peptide was therefore monomeric (i.e., not self-associated) throughout the range. All parent peptides were previously characterized by sedimentation equilibrium analytical ultracentrifugation experiments.[1–3] All unfolded analogues are diastereomers of the peptides of interest with LPro instead of DPro. See Chapter 2 for further explanation. All cyclic peptides, used to represent the fully populated β-hairpin, have the sequence of interest cyclized with a DPro-Gly turn.

Figure A4.1. Dilution Study of Peptide 2.11, Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TO^+-Ile-Leu-Gln-NH_2

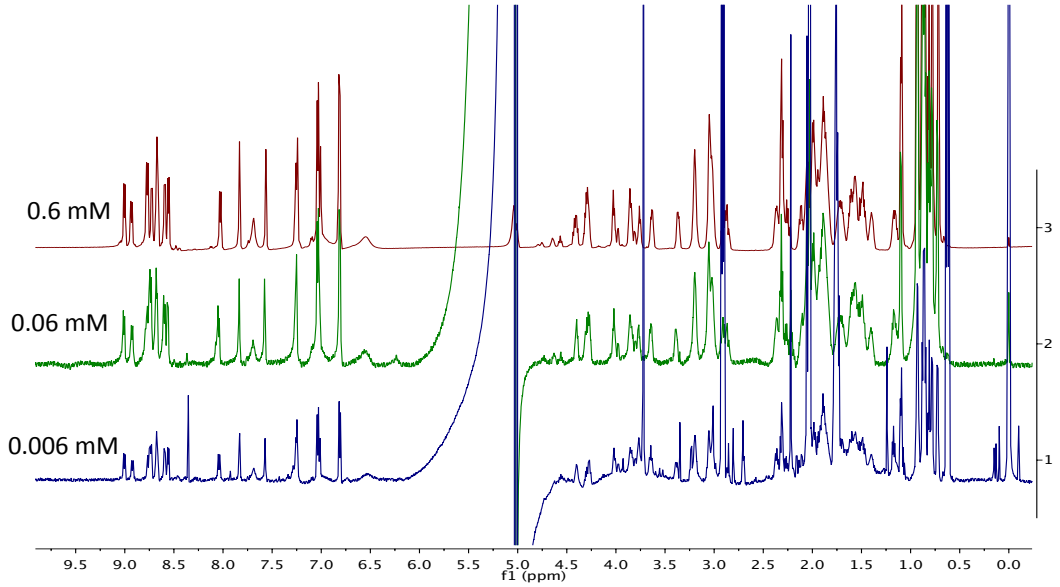
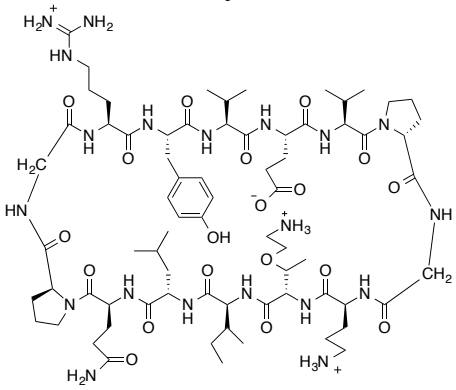


Figure A4.2. Dilution study of cyclic peptide 2.11c (fully folded control), Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TO^{+}-Ile-Leu-Gln-DPro-Gly



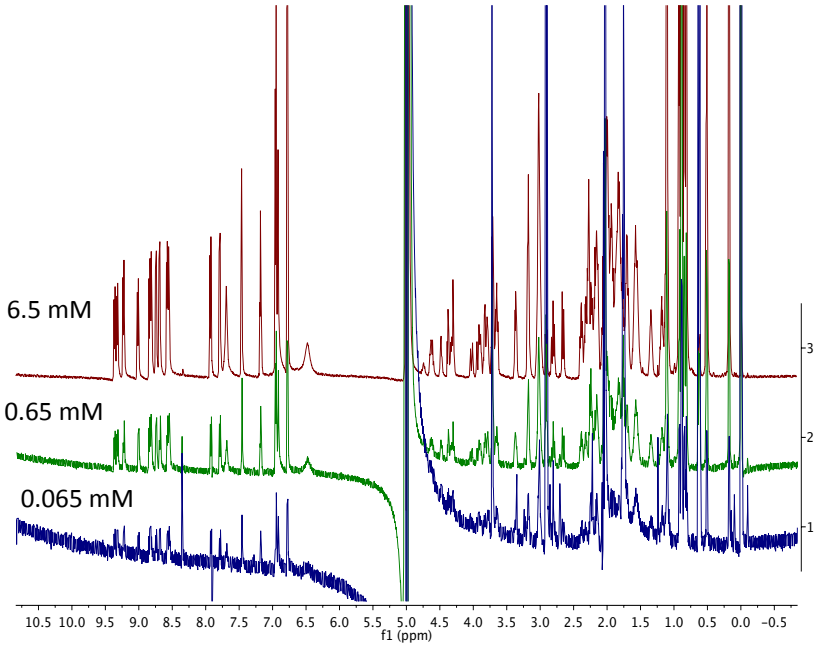


Figure A4.3. Dilution study of LPro diasteromer 2.11b (unfolded control), Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-TO^+-Ile-Leu-Gln-NH_2. A small impurity from the buffer is visible in the spectrum of the most dilute sample.

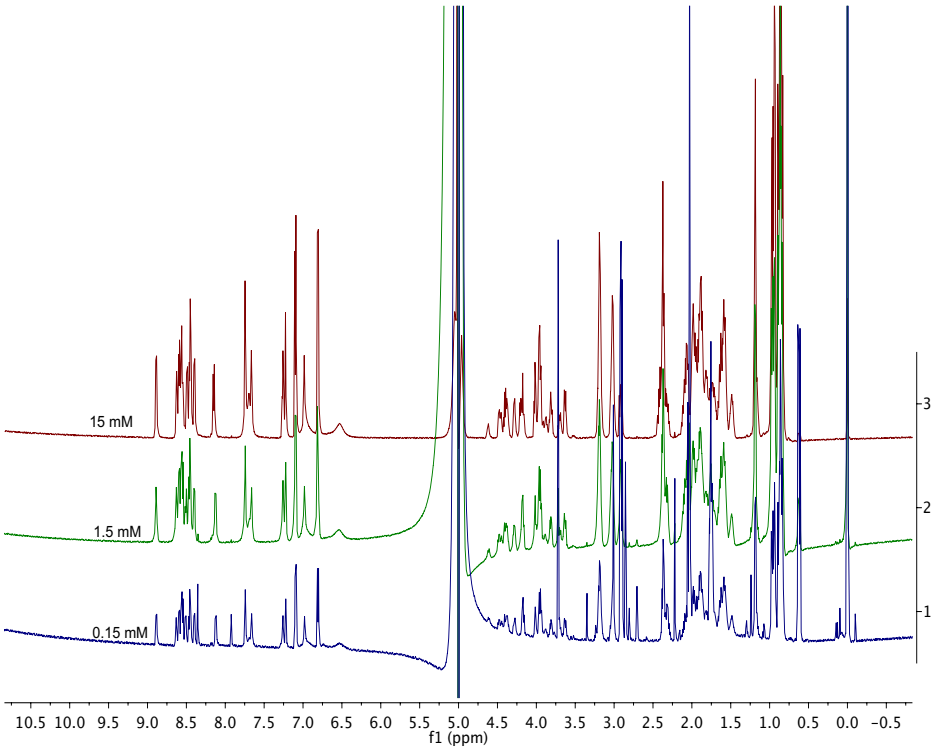


Figure A4.4. Dilution study of parent peptide 2.15, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

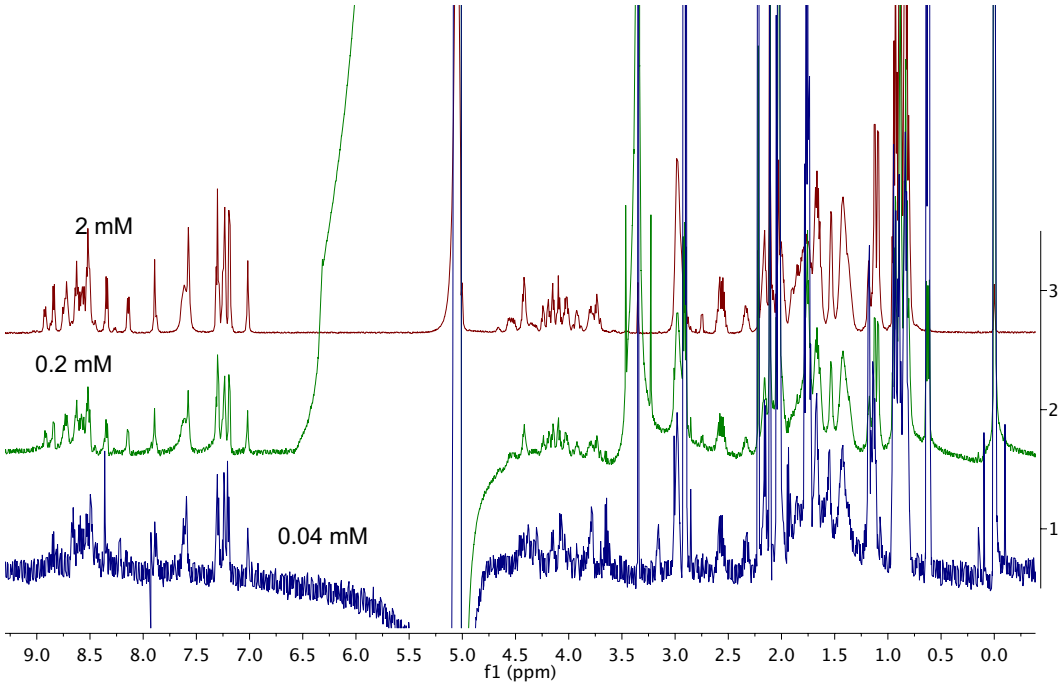


Figure A4.5. Dilution study of LPro diasteromer of 2.15 (unfolded control), Met-Gln-Ile-Phe-Val-Lys-Ser-LPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

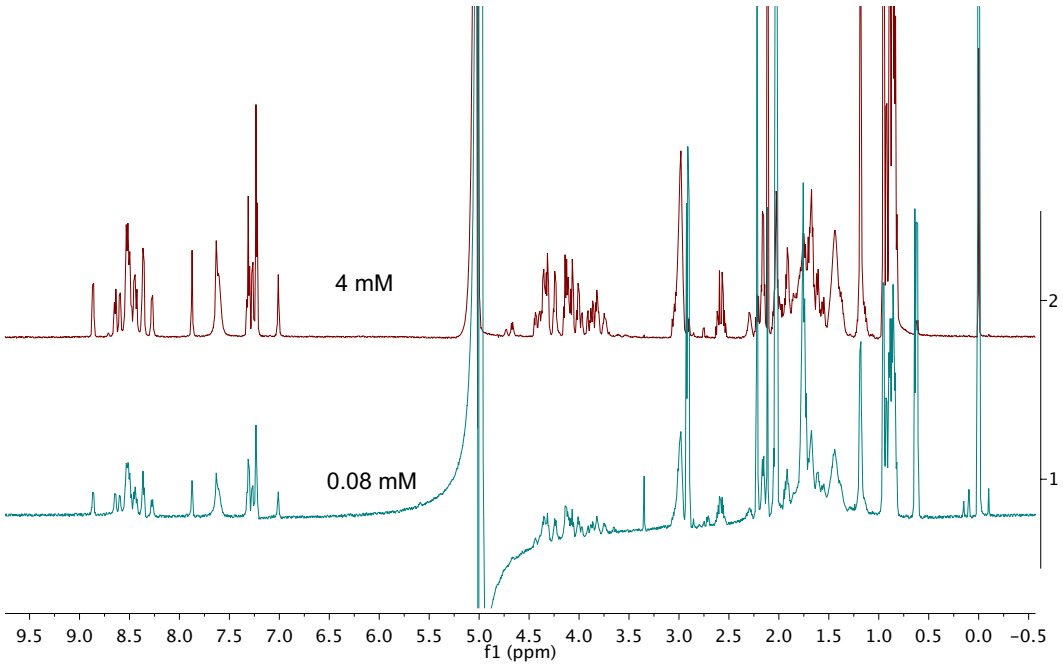


Figure A4.6. Dilution study of peptide 2.16, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-TO^+-Val-NH_2

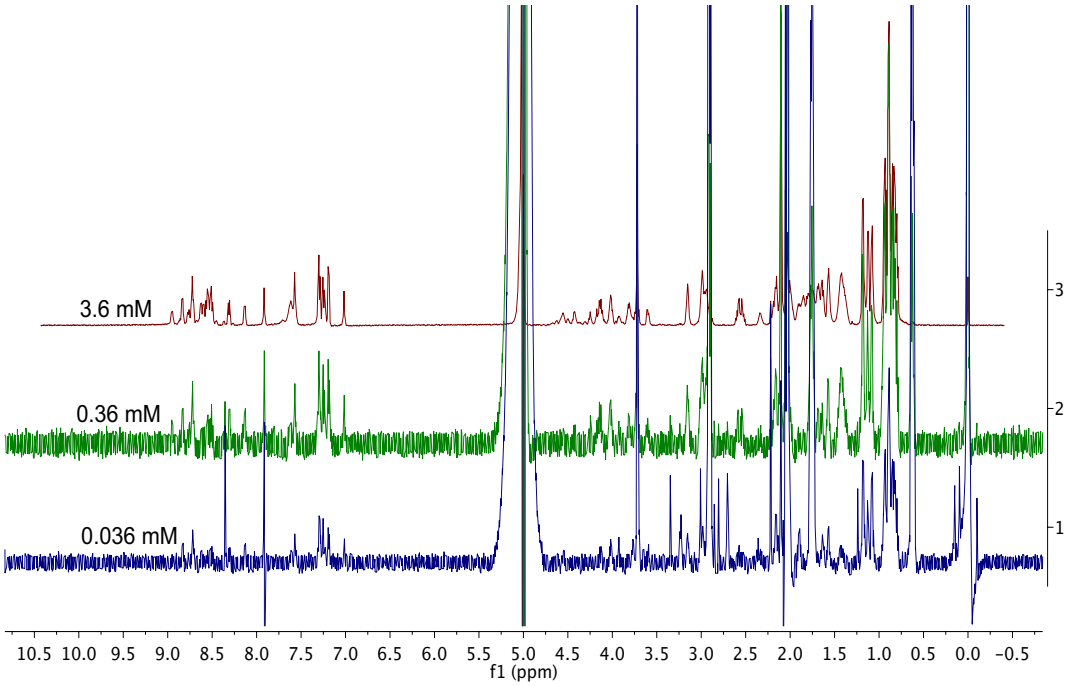


Figure A4.7. Dilution study of peptide 2.14, Arg-Trp-Gln-Tyr-Val-DPro-Gly-Lys-Phe-TO^+-Val-Gln-NH_2

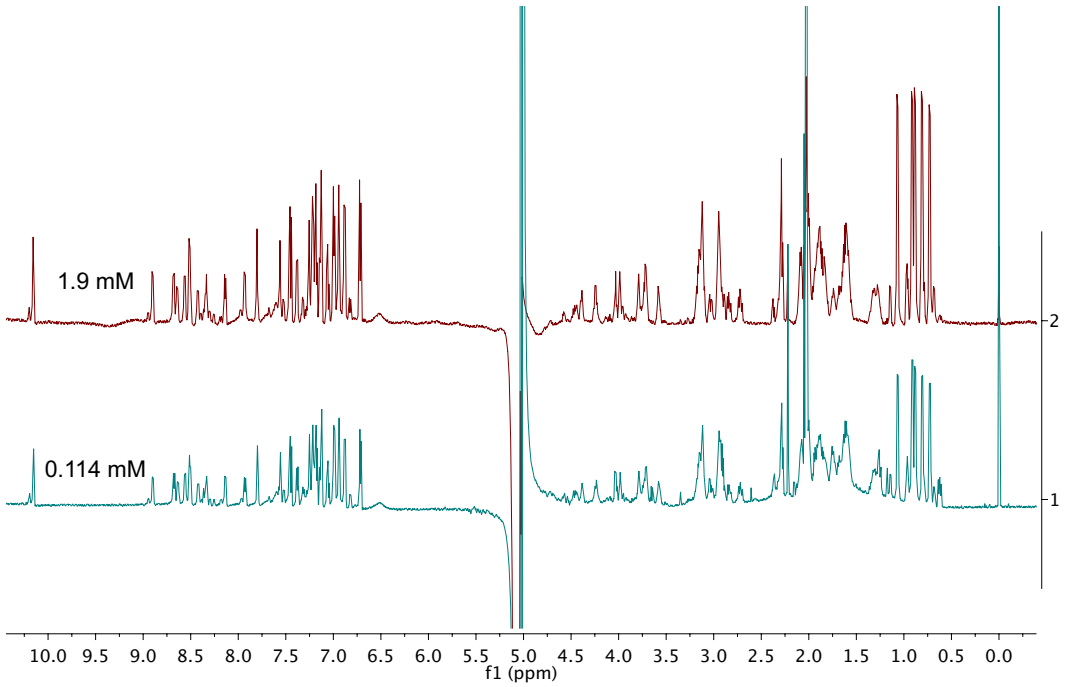
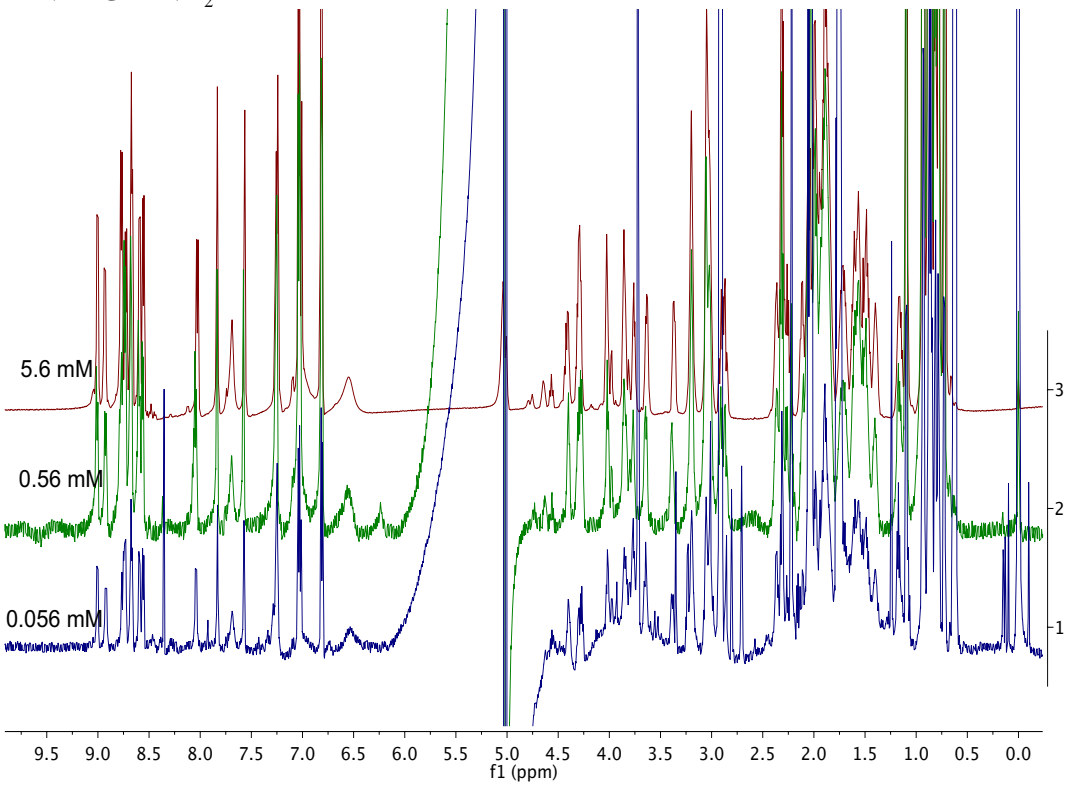


Figure A4.8. Dilution study of peptide 2.13, Arg-Trp-Gln-Tyr-Val-DPro-Gly-TO^{+}-Phe-Thr-Val-Gln-NH_2



Supplementary Section A5
Peptide Purity Check

General. For detailed information on data acquisition see the methods section of Chapter 2.

Figure A5.1. Purity check of peptide 2.11a, Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TO$^+$-Ile-Leu-Gln-NH_2

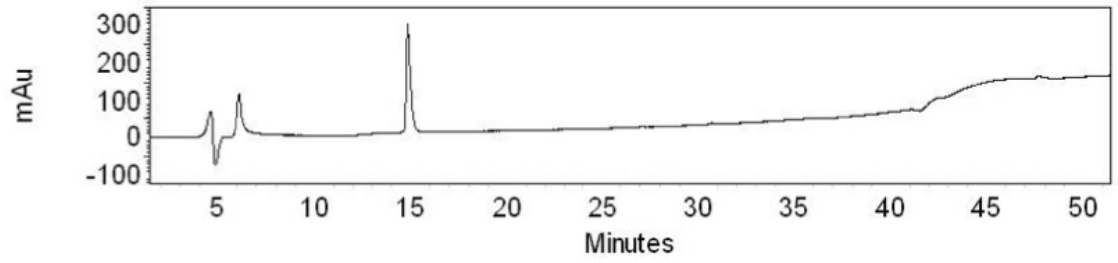


Figure A5.2. Purity check of peptide 2.11c, a cyclic peptide. Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TO$^+$-Ile-Leu-Gln-DPro-Gly

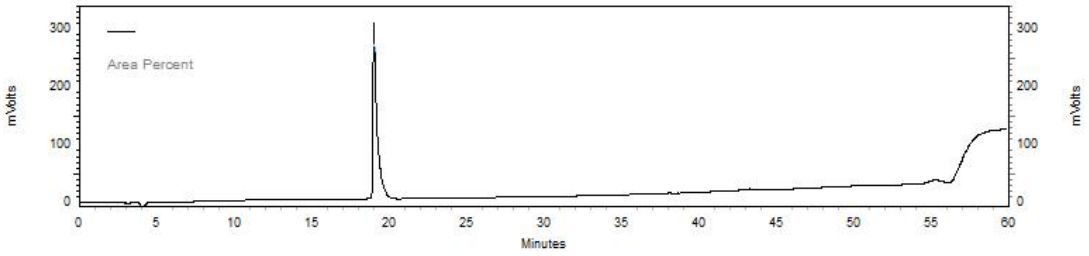


Figure A5.3. Purity check of LPro diasteromer 2.11b, Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-TO$^+$-Ile-Leu-Gln-NH_2

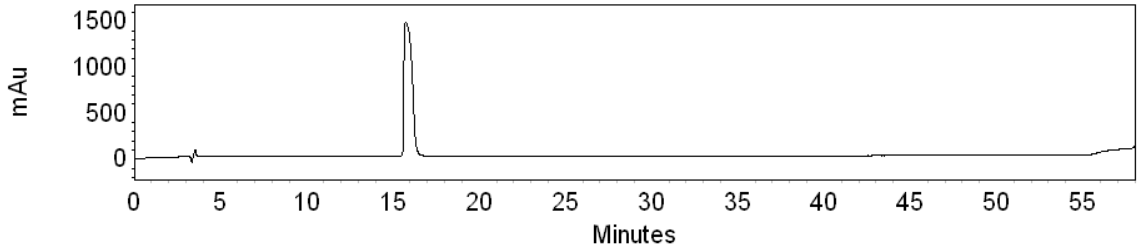


Figure A5.4. Purity check of parent peptide 2.15, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

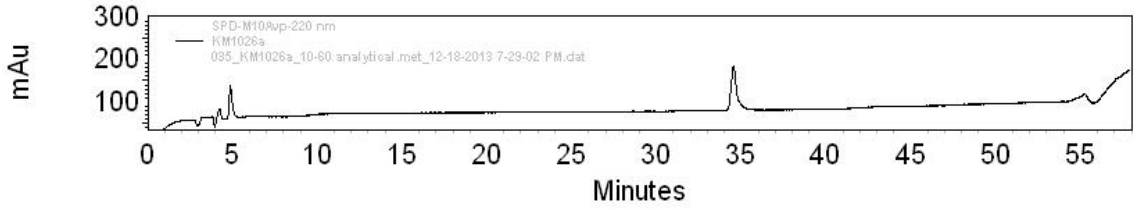


Figure A5.5. Purity check of LPro diasteromer of 2.15 (unfolded control), Met-Gln-Ile-Phe-Val-Lys-Ser-LPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

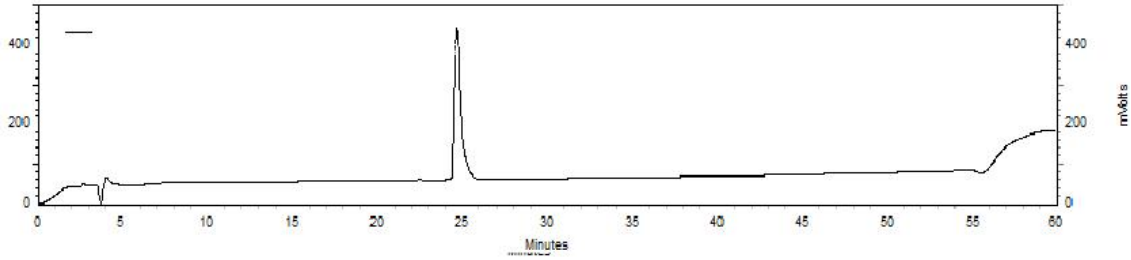


Figure A5.6. Purity check of peptide 2.16, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-TO^+-Val-NH_2

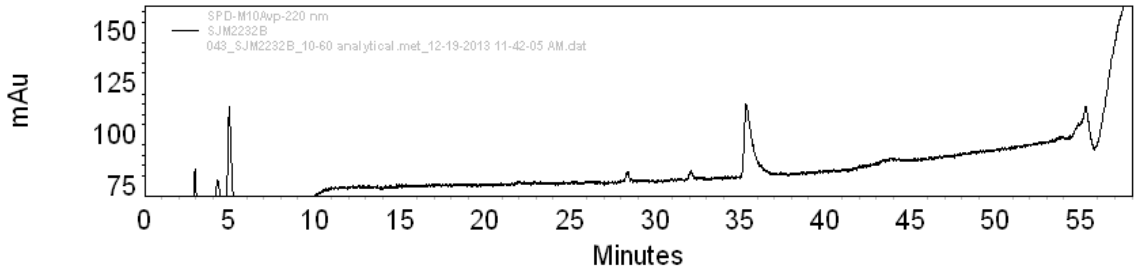


Figure A5.7. Purity check of peptide 2.14, Arg-Trp-Gln-Tyr-Val-DPro-Gly-Lys-Phe-TO^+-Val-Gln-NH_2

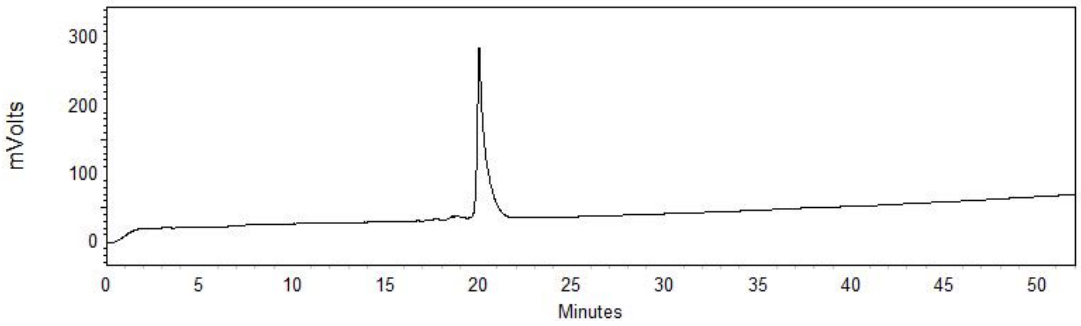
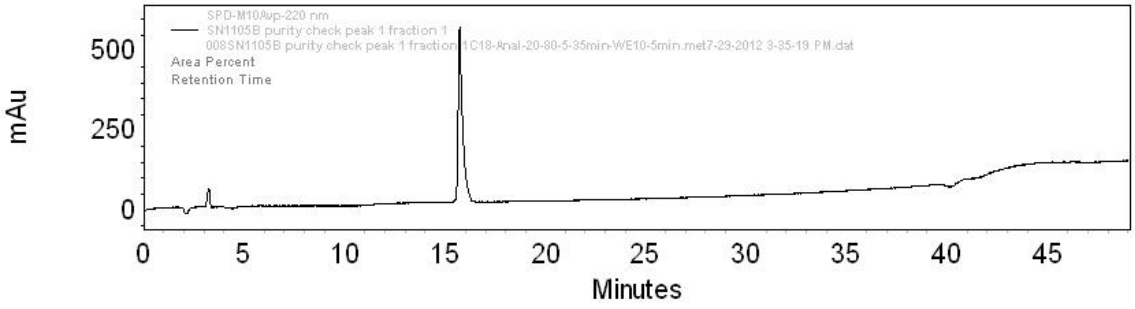
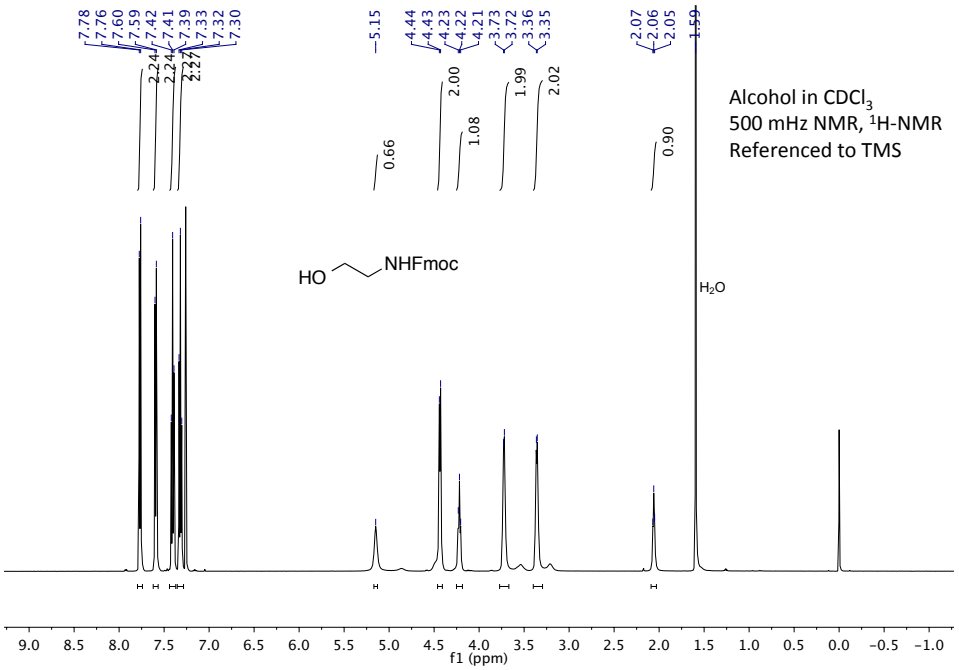
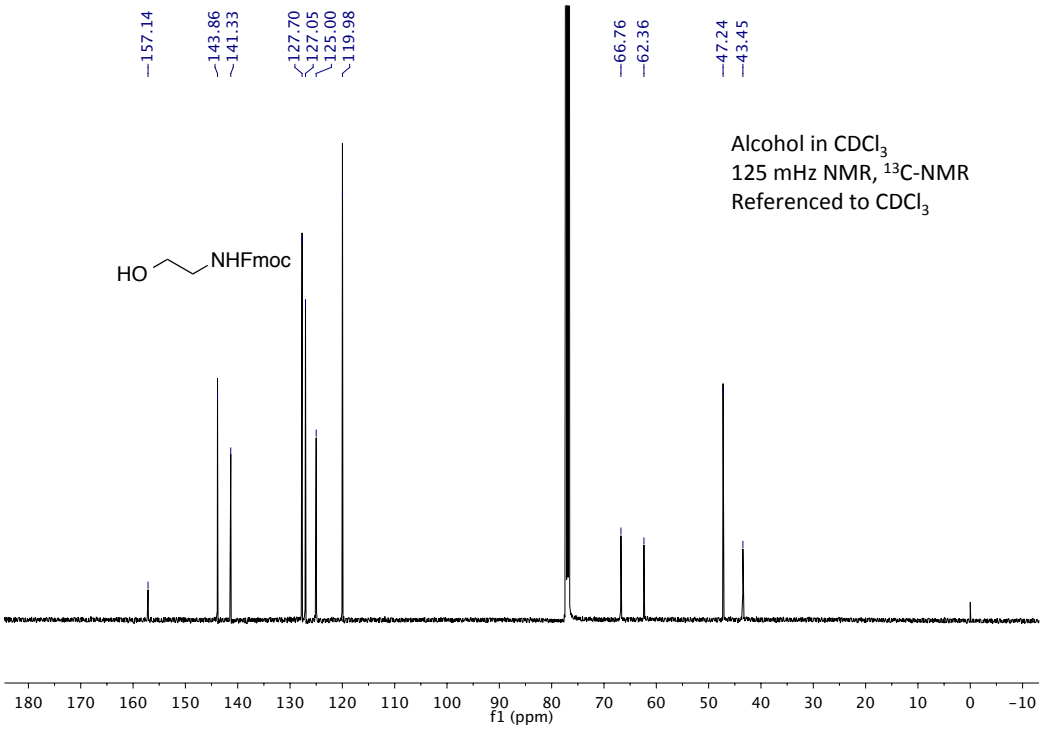


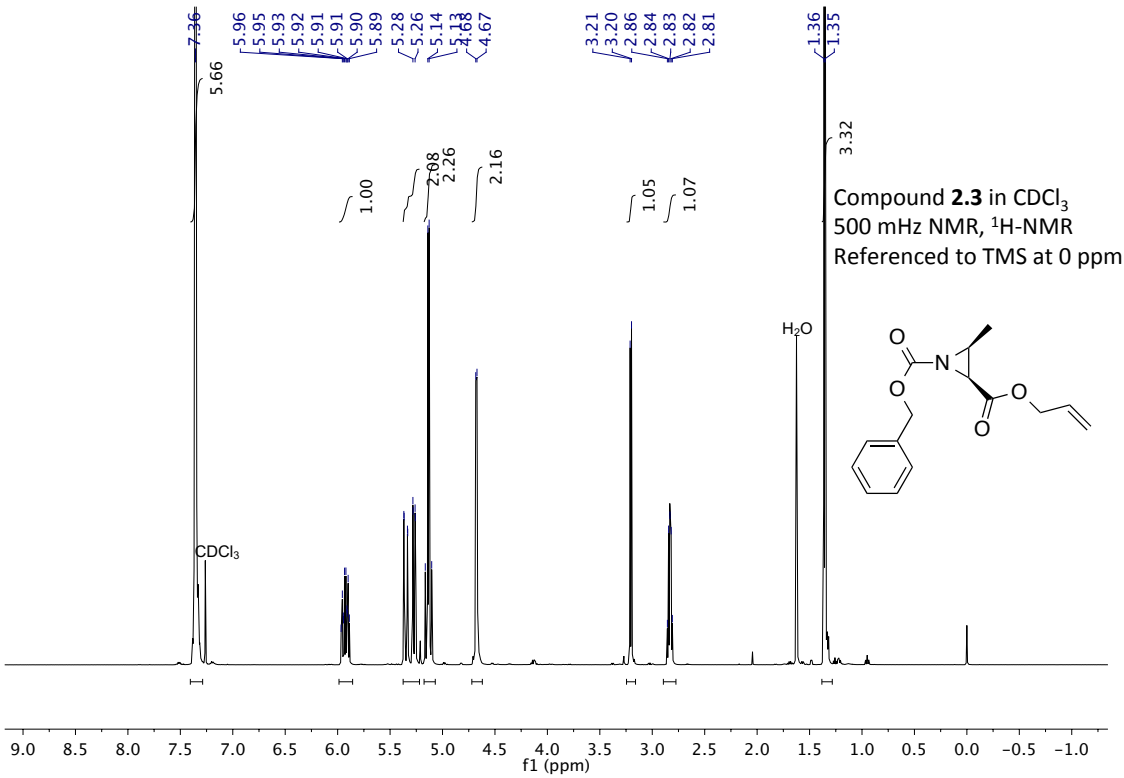
Figure A5.8. Purity check of peptide 2.13, Arg-Trp-Gln-Tyr-Val-DPro-Gly-TO^{+}-Phe-Thr-Val-Gln-NH_2

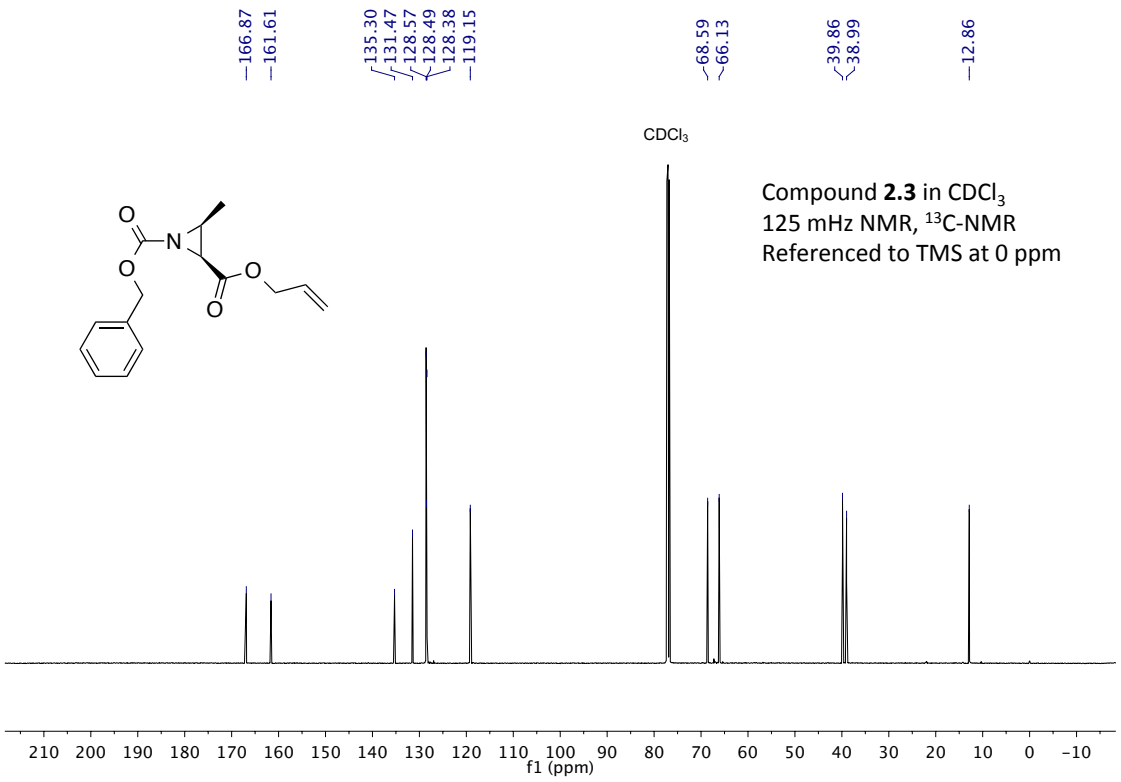


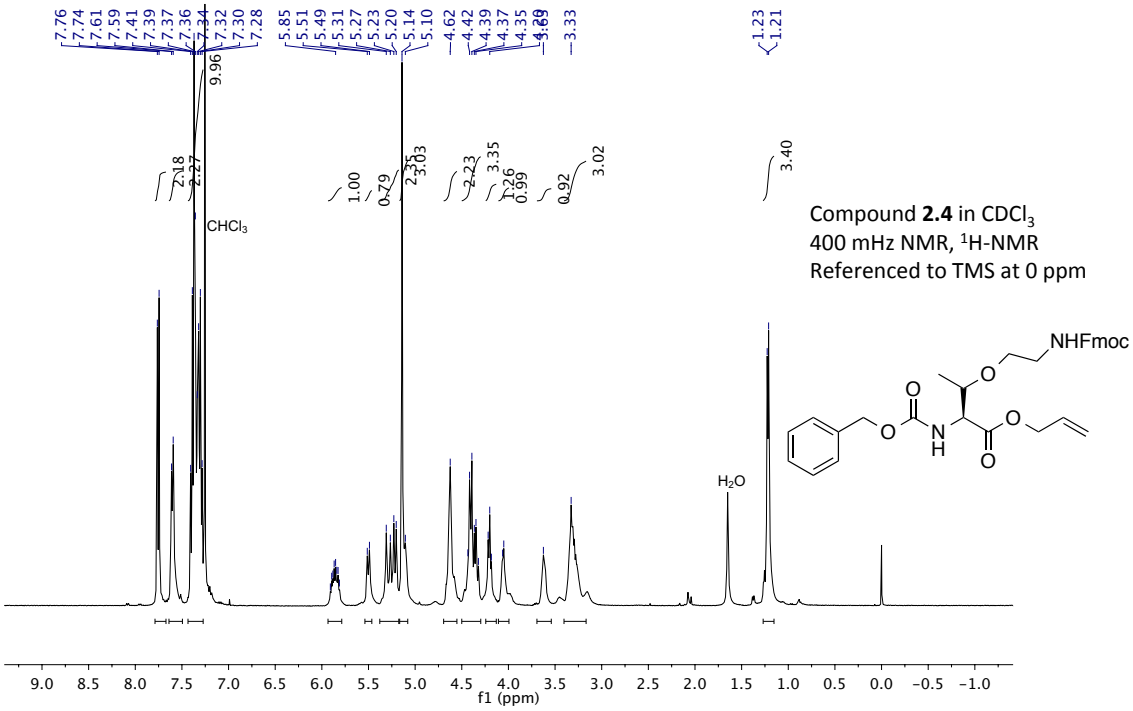
Supplementary Section A6
NMR Spectra



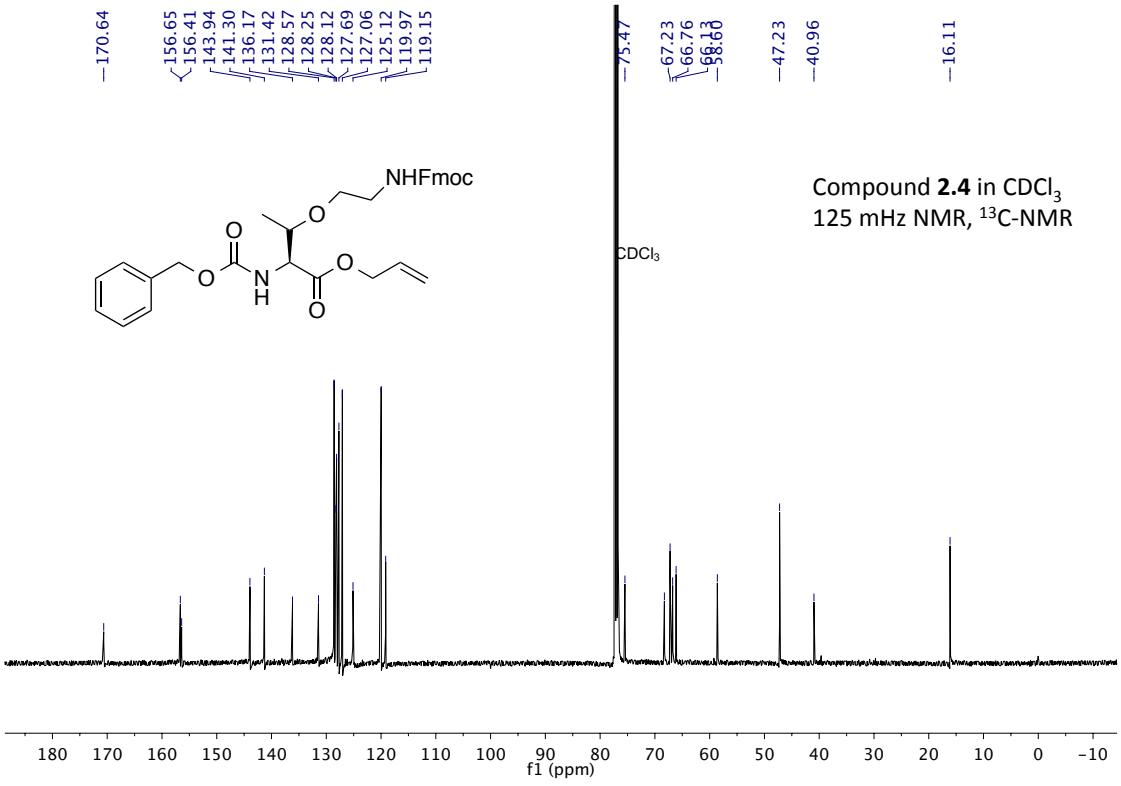



7.36
5.96
5.95
5.93
5.92
5.91
5.91
5.90
5.89
5.66
5.28
5.26
5.14
5.13
4.68
4.67
3.21
3.20
2.86
2.84
2.83
2.82
2.81
1.36
1.35
3.32
2.08
2.26
2.16
1.00
1.05
1.07
H2O
CDCl3
Compound 2.3 in CDCl3
500 mHz NMR, 1H-NMR
Referenced to TMS at 0 ppm
f1 (ppm)

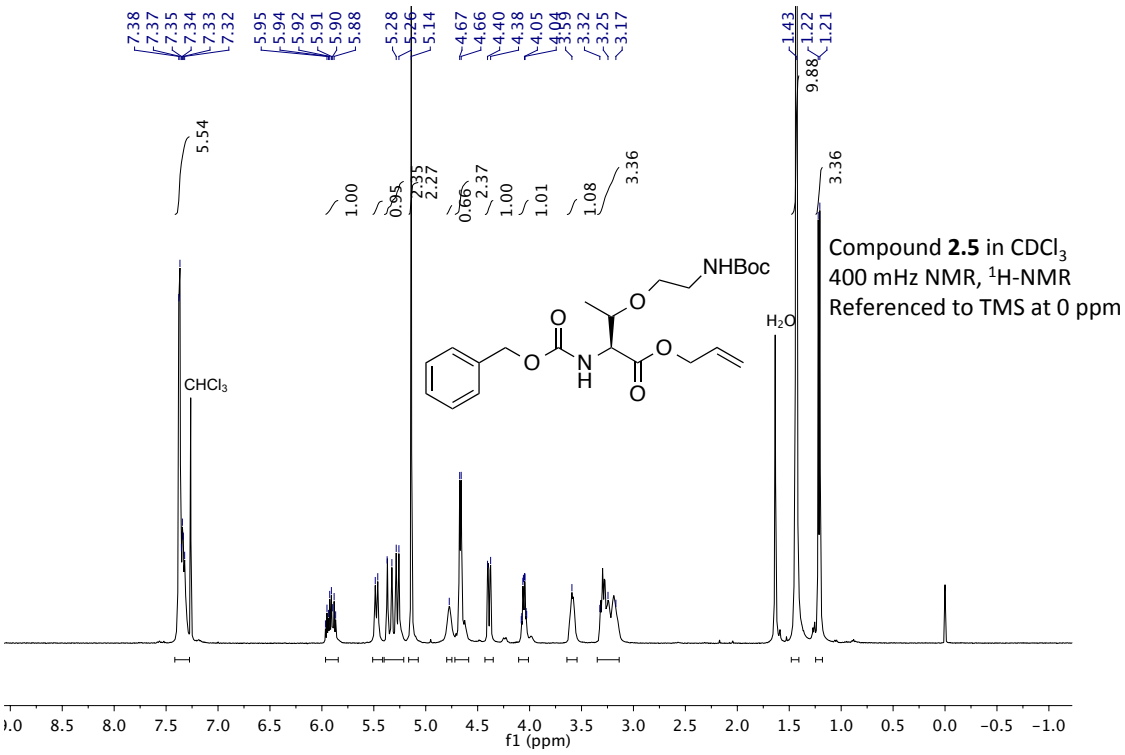

166.87
161.61
135.30
131.47
128.57
128.49
128.38
119.15
68.59
66.13
39.86
38.99
12.86
CDCl3
Compound 2.3 in CDCl3
125 mHz NMR, 13C-NMR
Referenced to TMS at 0 ppm
f1 (ppm)



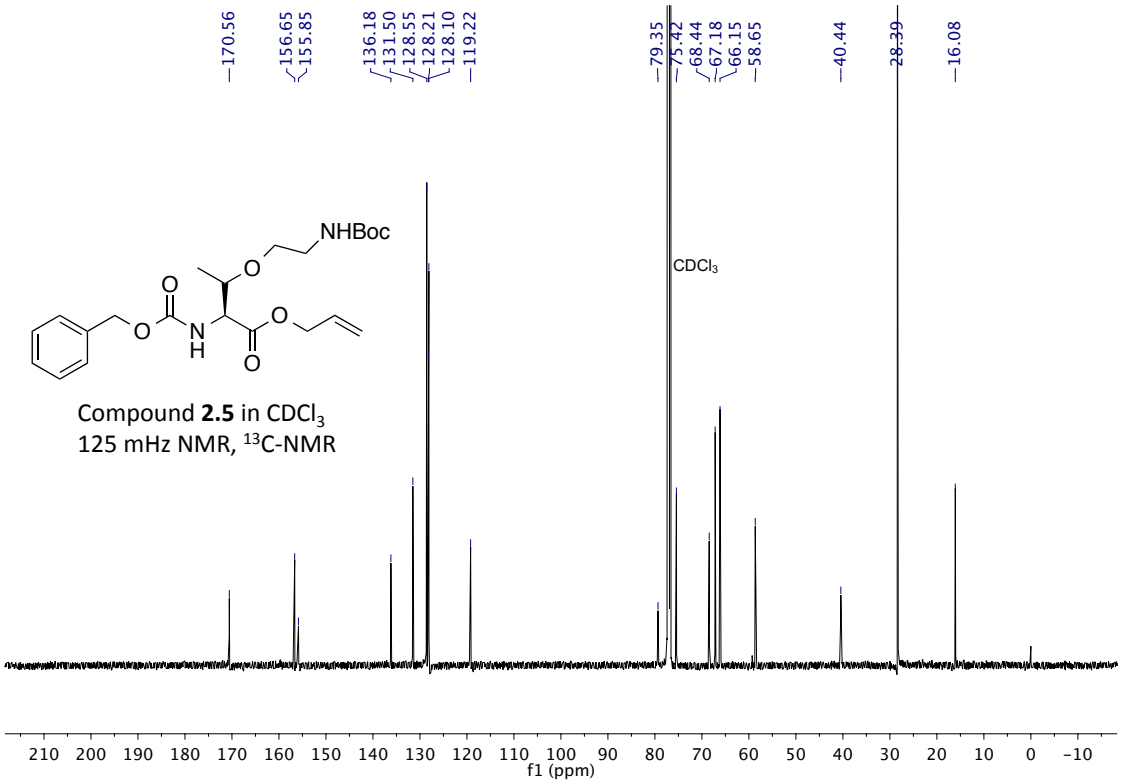
Compound **2.4** in $CDCl_3$
400 mHz NMR, ^{1}H-NMR
Referenced to TMS at 0 ppm



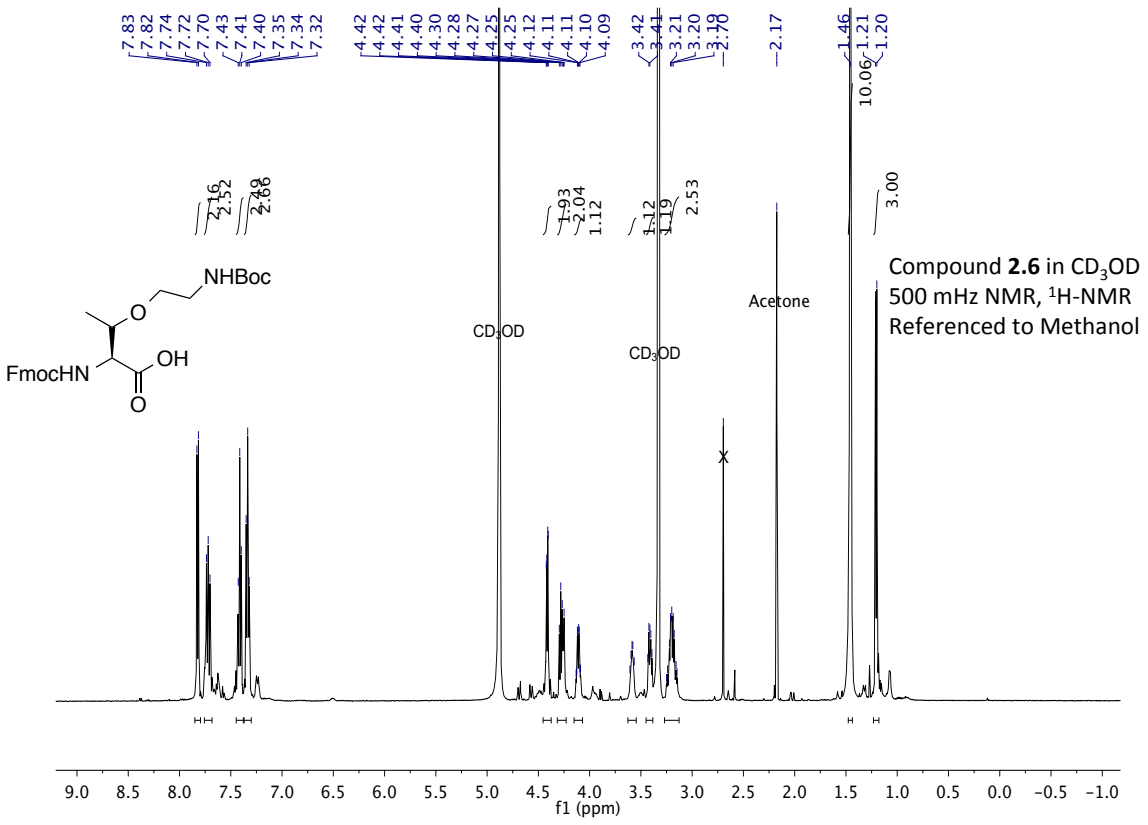
Compound **2.4** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR

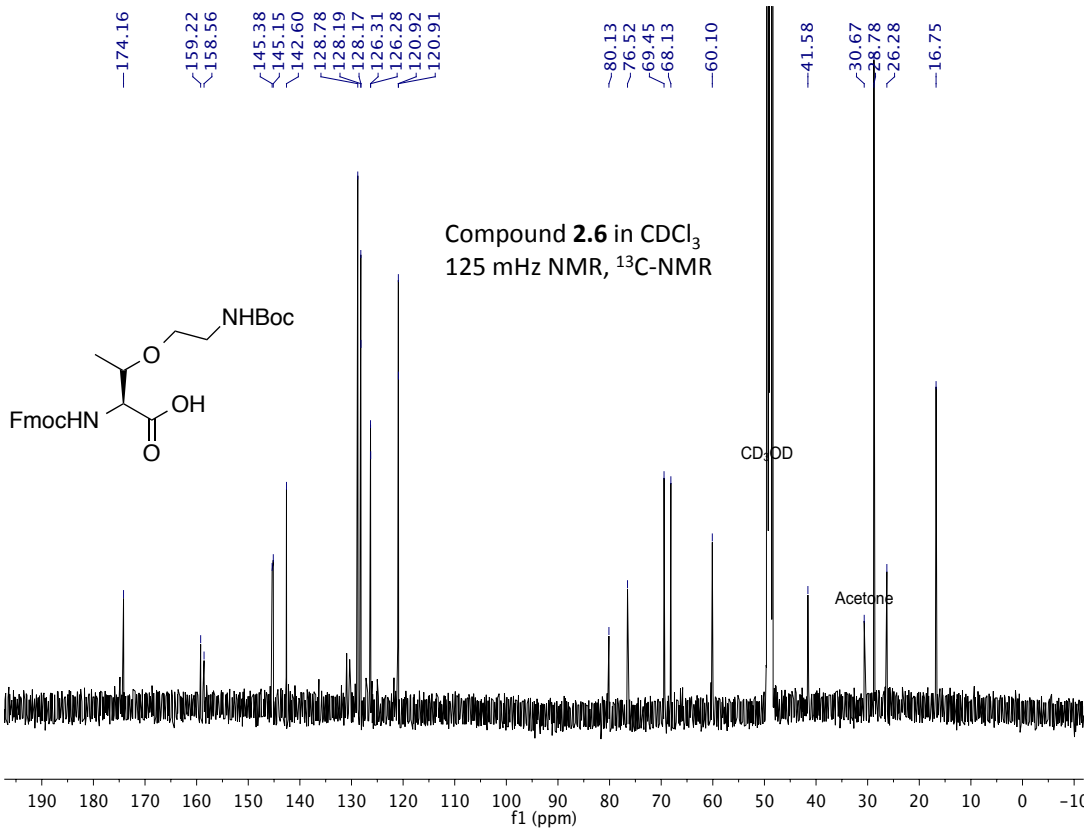


Compound **2.5** in $CDCl_3$
400 mHz NMR, ^{1}H-NMR
Referenced to TMS at 0 ppm



Compound **2.5** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR


NHBoc
O
FmocHN
OH
O
CD3OD
CD3OD
Acetone
Compound 2.6 in CD3OD
500 mHz NMR, 1H-NMR
Referenced to Methanol
f1 (ppm)


NHBoc
O
FmocHN
OH
O
Compound 2.6 in CDCl3
125 mHz NMR, 13C-NMR
CD3OD
Acetone
f1 (ppm)

Supplementary Section A7
MALDI-MS

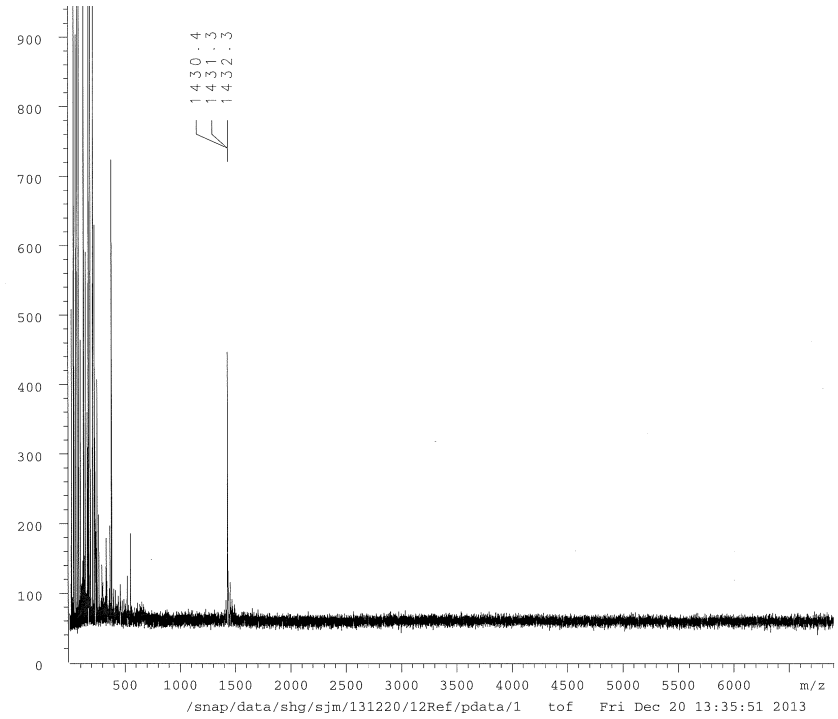


Figure A7.1. MALDI of peptide 2.11a in CHCA matrix. Expected M+1=1430.8 m/z.

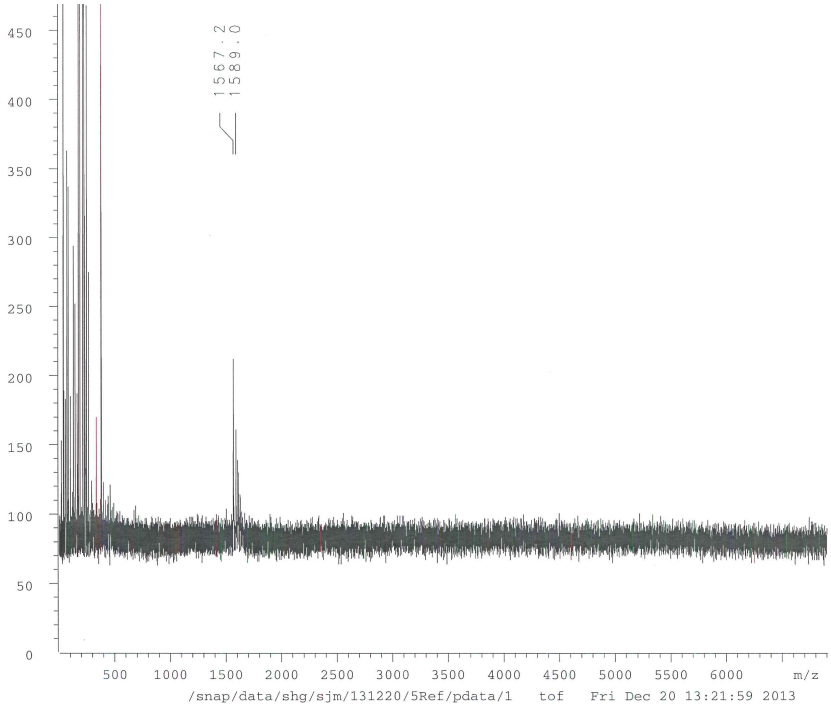


Figure A7.2. MALDI of cyclic peptide 2.11c in CHCA matrix Expected M+1=1567.9 m/z, M+Na=1589.9 m/z.

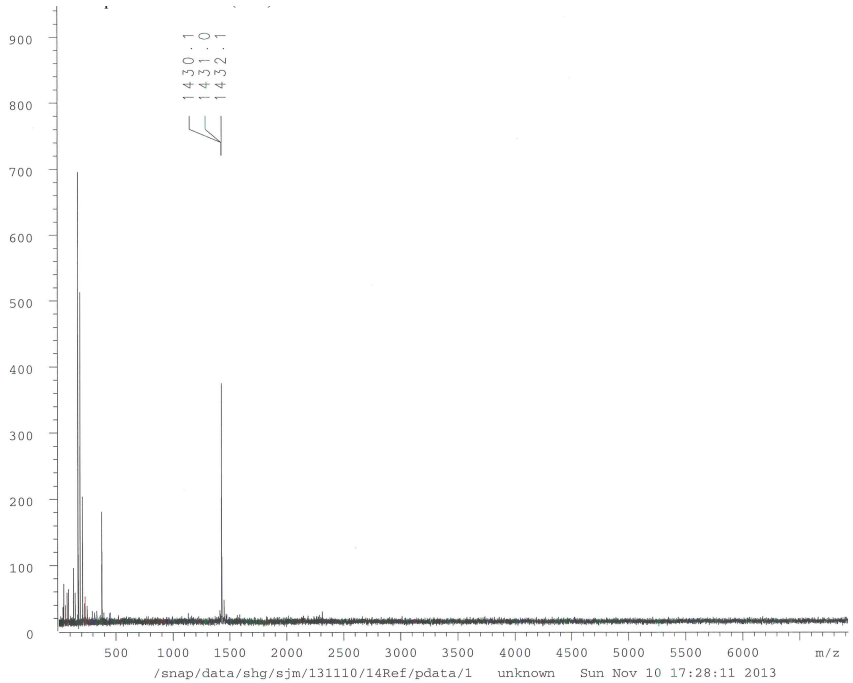


Figure A7.3. MALDI of LPro peptide 2.11b in CHCA matrix. Expected M+1=1430.8 m/z.

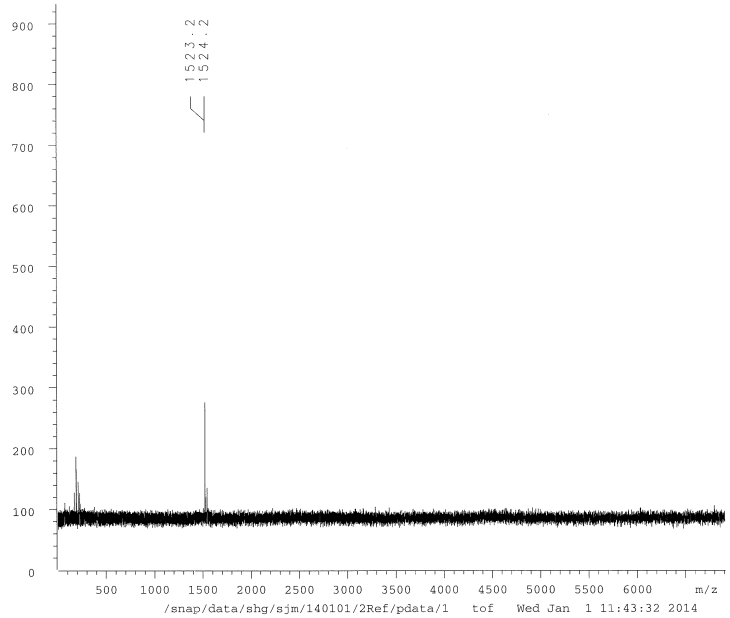


Figure A7.4. MALDI of peptide 2.13 in CHCA matrix. Expected: 1523.8 m/z (M+H).

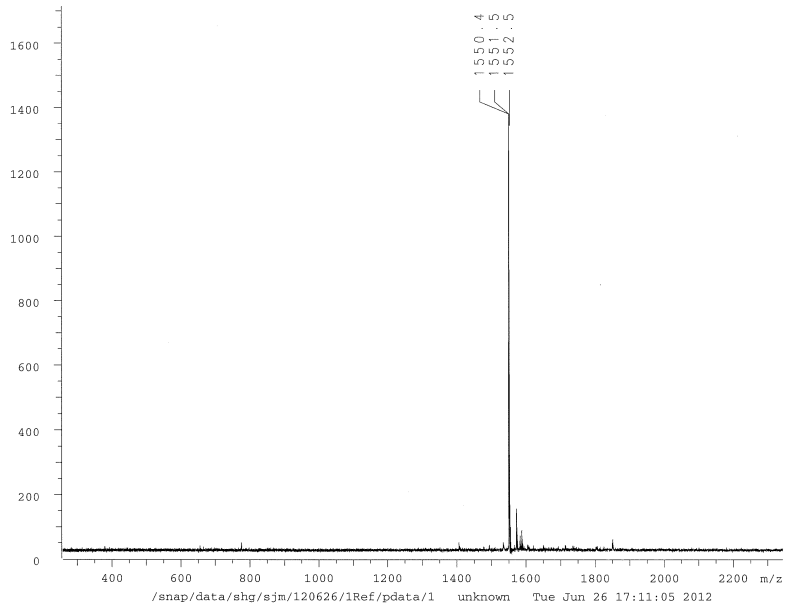


Figure A7.5. MALDI of peptide 2.14 in CHCA matrix. Expected: 1550.9 m/z (M+H).

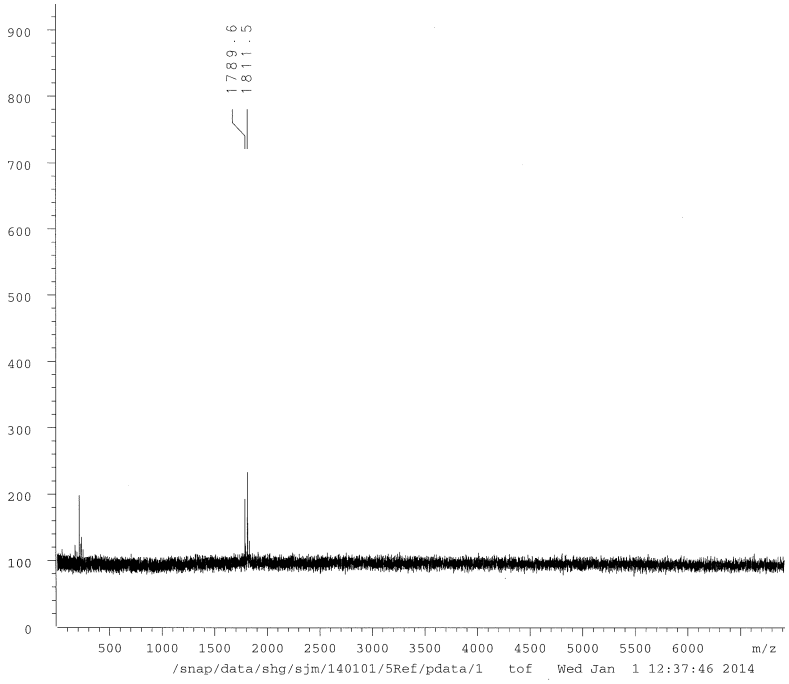


Figure A7.6. MALDI of peptide 2.15 in CHCA matrix. Expected M+H=1789.1 m/z. Expected M+Na=1811.5 m/z.

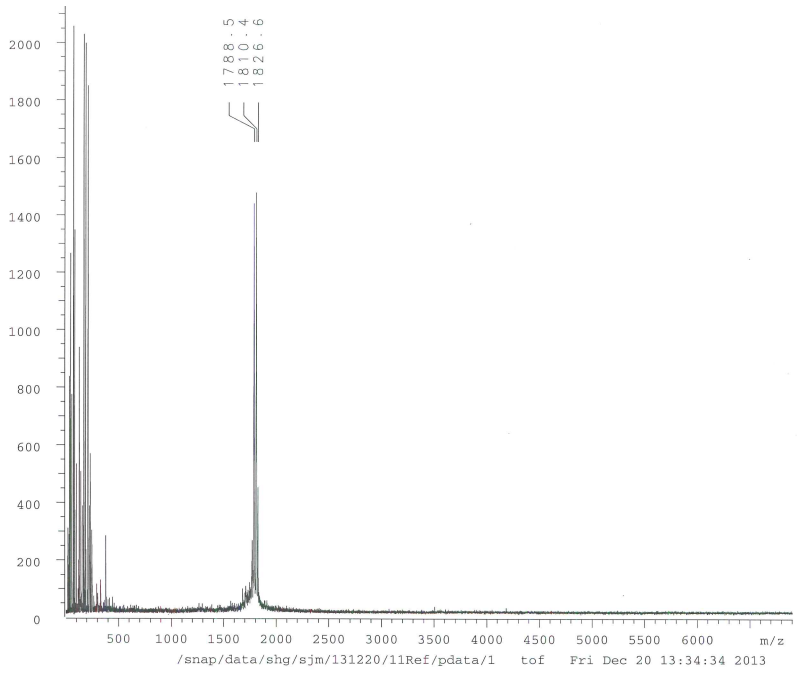


Figure A7.7. MALDI of LPro analogue of peptide 2.15 in CHCA matrix. Expected: 1789.1 m/z (M+H), 1811.5 m/z (M+Na).

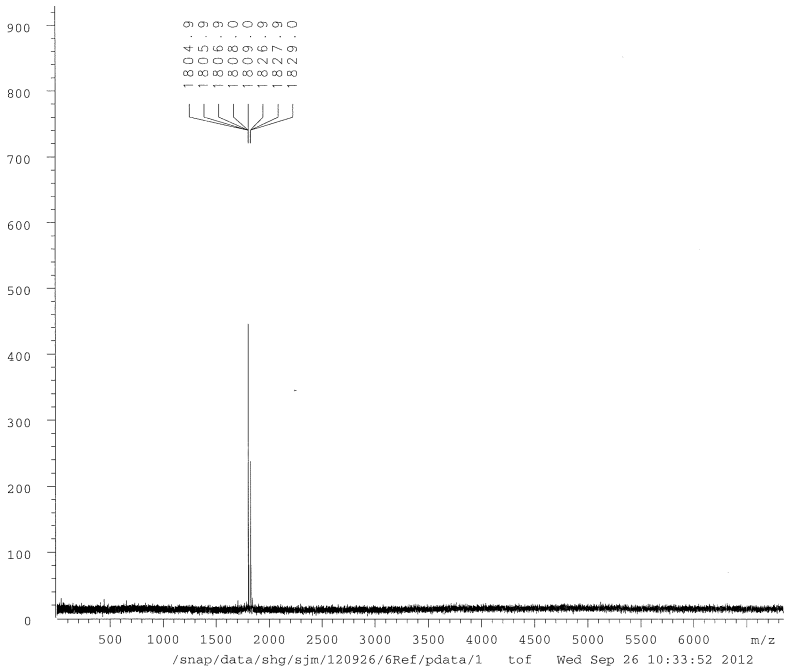


Figure A7.8. MALDI of peptide 2.16 in CHCA matrix. Expected: 1805.1 (M+1), 1827.1 m/z (M+Na).

Supplementary Section A8
NMR Structure Calculation Restraint Files

General. Detailed information on NMR data acquisition and structural calculations can be found in the methods section of Chapter 2. Here are the restraint files as they were entered into CNS. The protons are named and the distance restrains are assigned using the standard CNS format.[4]

Restraints as entered in CNS for cyclic peptide 2.11c (Lys9 → TO^+):

!!! Sequential and Short Range NOEs

assign (residue 1 and name HA) (residue 2 and name HN) 2.2 1.0 0.9
assign (residue 1 and name HB#) (residue 2 and name HN) 2.6 1.0 0.9
assign (residue 1 and name HN) (residue 14 and name HN) 2.6 1.0 0.9
assign (residue 1 and name HN) (residue 14 and name HA#) 2.8 1.0 0.9
assign (residue 2 and name HA) (residue 3 and name HN) 2.4 1.0 0.9
assign (residue 3 and name HA) (residue 4 and name HN) 2.2 1.0 0.9
assign (residue 4 and name HA) (residue 5 and name HN) 3.0 1.0 0.9
assign (residue 6 and name HA) (residue 7 and name HN) 2.6 1.0 0.9
assign (residue 7 and name HN) (residue 8 and name HN) 2.6 1.0 0.9
assign (residue 8 and name HA) (residue 9 and name HN) 2.2 1.0 0.9
assign (residue 8 and name HB#) (residue 9 and name HN) 1.5 1.0 0.9
assign (residue 9 and name HA) (residue 10 and name HN) 2.1 1.0 0.9
assign (residue 9 and name HB) (residue 10 and name HN) 2.9 1.0 0.9
assign (residue 10 and name HA) (residue 11 and name HN) 2.2 1.0 0.9
assign (residue 10 and name HB) (residue 11 and name HN) 1.8 1.0 0.9
assign (residue 11 and name HA) (residue 12 and name HN) 2.2 1.0 0.9
assign (residue 11 and name HD##) (residue 12 and name HN) 3.5 1.0 0.9

!!! Medium and Long Range NOEs

assign (residue 1 and name HN) (residue 11 and name HD##) 3.0 1.0 0.9
assign (residue 1 and name HN) (residue 12 and name HN) 2.7 1.0 0.9
assign (residue 2 and name HB1) (residue 9 and name HB) 2.9 1.0 0.9
assign (residue 2 and name HD#) (residue 10 and name HN) 2.8 1.0 0.9
assign (residue 2 and name HD#) (residue 11 and name HN) 3.1 1.0 0.9
assign (residue 2 and name HD#) (residue 11 and name HA) 2.9 1.0 0.9
assign (residue 2 and name HD#) (residue 11 and name HD##) 3.1 1.0 0.9
assign (residue 2 and name HD#) (residue 11 and name HD##) 3.1 1.0 0.9
assign (residue 2 and name HE#) (residue 11 and name HN) 3.1 1.0 0.9
assign (residue 2 and name HE#) (residue 11 and name HD##) 2.9 1.0 0.9

assign (residue 2 and name HE#) (residue 11 and name HD##) 3.0 1.0 0.9
assign (residue 5 and name HN) (residue 8 and name HN) 2.7 1.0 0.9
assign (residue 9 and name HB) (residue 11 and name HD##) 2.6 1.0 0.9
assign (residue 9 and name HB) (residue 11 and name HD##) 2.7 1.0 0.9

!! Second Round
assign (residue 2 and name HD#) (residue 9 and name HB) 2.8 1.0 0.9
assign (residue 2 and name HD#) (residue 9 and name HG2#) 2.5 1.0 0.9
assign (residue 2 and name HE#) (residue 11 and name HB) 2.5 1.0 0.9

References

(1) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quanfification of Beta-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

(2) Haque, T. S.; Gellman, S. H. Insights on Beta-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I′ and Type II′ Beta-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

(3) Espinosa, J. F.; Gellman, S. H. A Designed Beta-Hairpin Containing a Natural Hydrophobic Cluster. **2000**, *11*, 2330–2333.

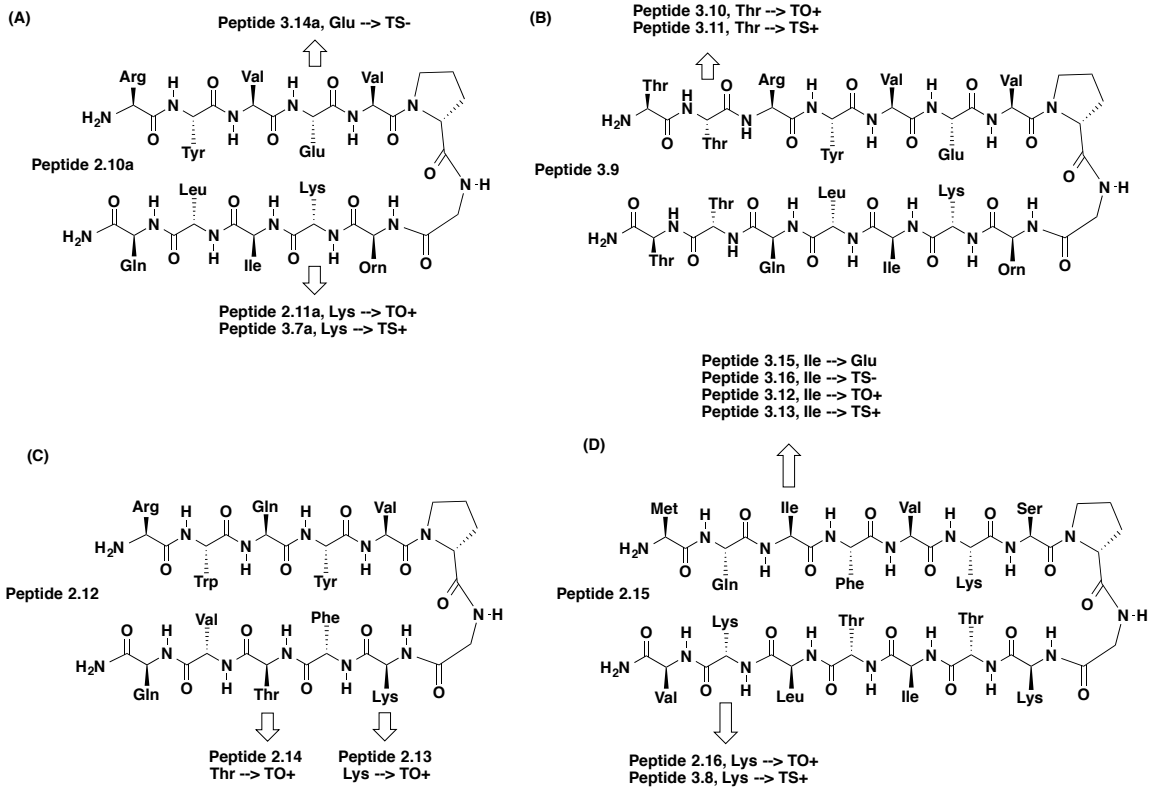
(4) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; et al. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. D. Biol. Crystallogr.* **1998**, *54*, 905–921.

Appendix B
Supporting Information for Chapter 3

Table of Contents

Supplementary Section B1
Peptide Sequence Summary

Figure B1. Summary of peptides synthesized and characterized. (A) Peptides derived from the parent sequence described by Syud et al.[1] (B) Peptides derived from the parent sequence described by Stanger et al.[2] The sequence is an extended version of the sequence described by Syud et al.[1] with two extra Thr residues at each terminus. (C) Peptides derived from the parent sequence described by Espinosa et al.[3] (D) Peptides derived from the parent sequence from Haque et al.[4]



Supplementary Section B2
Additional Summaries of Cross-Strand ROEs

General. For detailed information on sample preparation and data acquisition, see Chapter 3 methods section.

Figure B2.1. Cross-strand ROEs observed in peptide 3.10 (Thr2 → TO^+, 14 mM) at 4°C in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Backbone-backbone ROEs in the turn are also included. Strong ROEs are red, medium ROEs are blue, and weak ROEs are green.

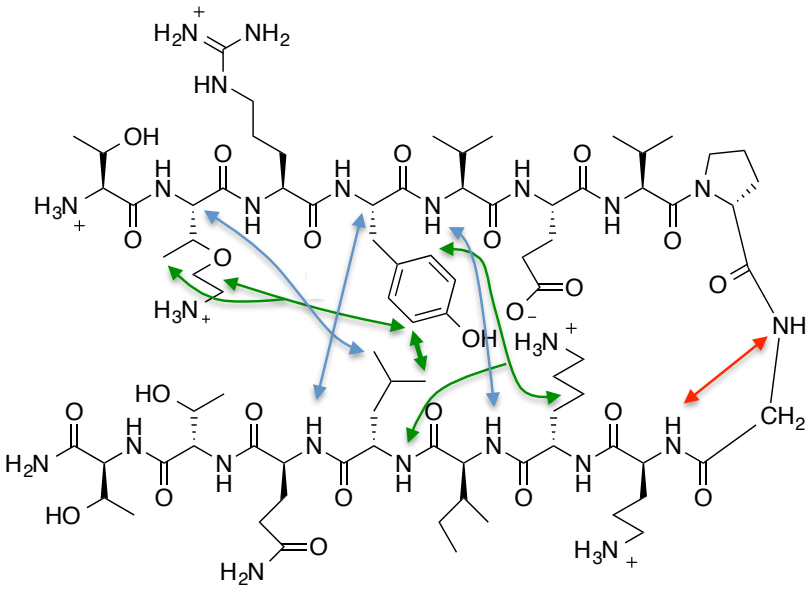


Figure B2.2. Cross-strand ROEs observed in peptide 3.11 (Thr2 → TS^+, 4 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Backbone-backbone ROEs in the turn are also included. Medium ROEs are blue and weak ROEs are green. Chemical shifts of Val 5 and 7-NH overlap, so cross-strand backbone Val5-NH to Ile12-NH and Val7-NH to Orn10-NH cannot be definitively assigned, but are detected.

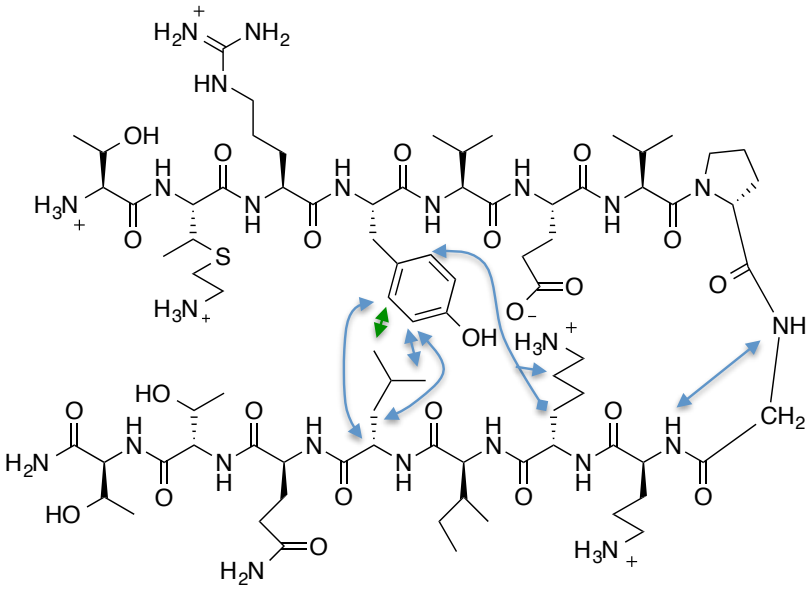


Figure B2.3. Cross-strand NOEs observed in peptide 3.8 (Lys15 → TS^+, 0.9 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Backbone-backbone ROEs in the turn are also included. Strong ROEs are red and medium ROEs are blue.

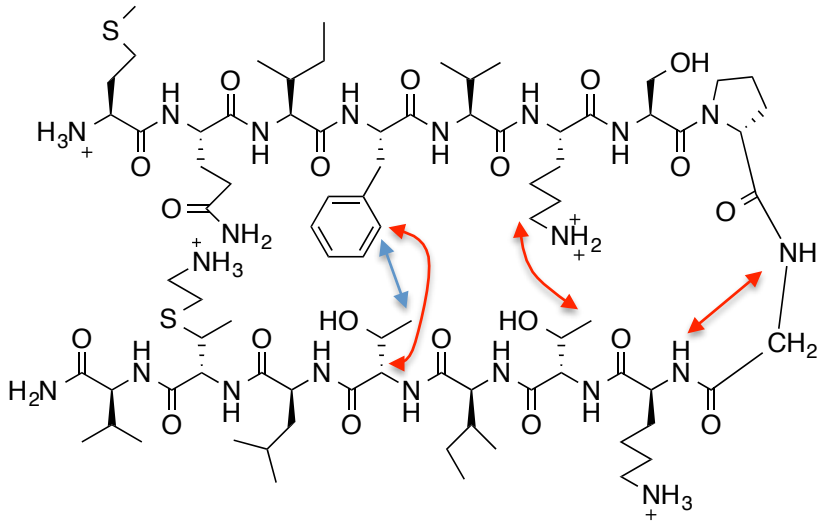


Figure B2.4. Cross-strand NOEs observed in peptide 3.12 (Ile3 → TO^+, 4.4 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Backbone-backbone ROEs in the turn are also included. Strong ROEs are red and weak ROEs are green.

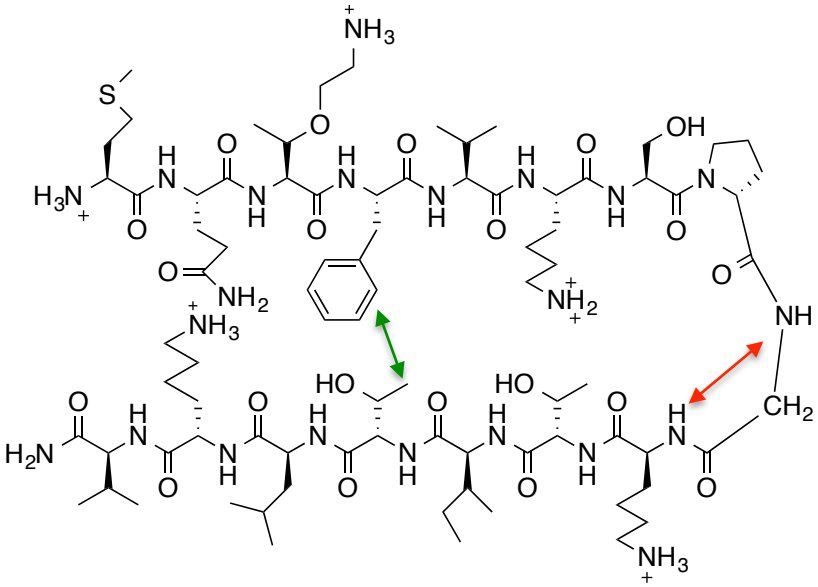


Figure B2.5. Cross-strand NOEs observed in peptide 3.13 (Ile3 → TS^+, 2 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Backbone-backbone ROEs in the turn are also included. Strong ROEs are red and medium ROEs are blue.

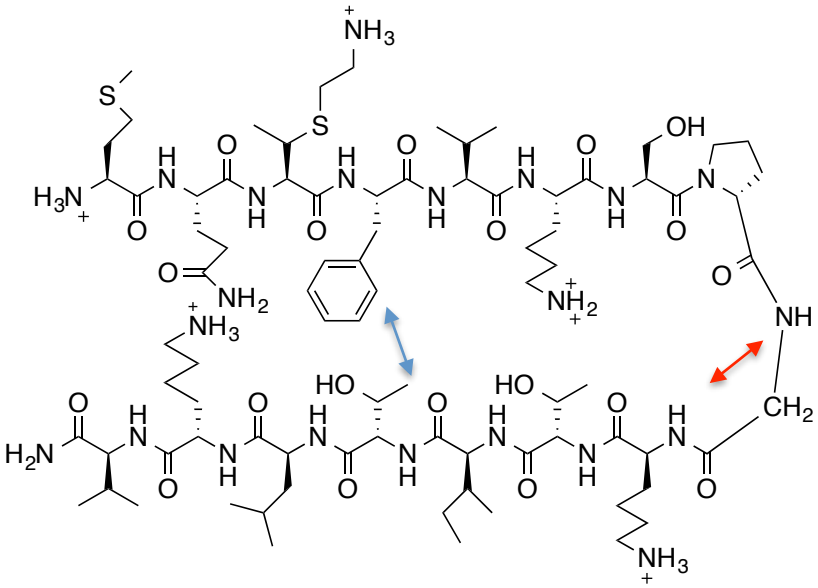


Figure B2.6. Cross-strand ROEs observed in peptide 3.16 (Ile3 → TS^-, 4 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Strong ROEs are red and medium ROEs are blue.

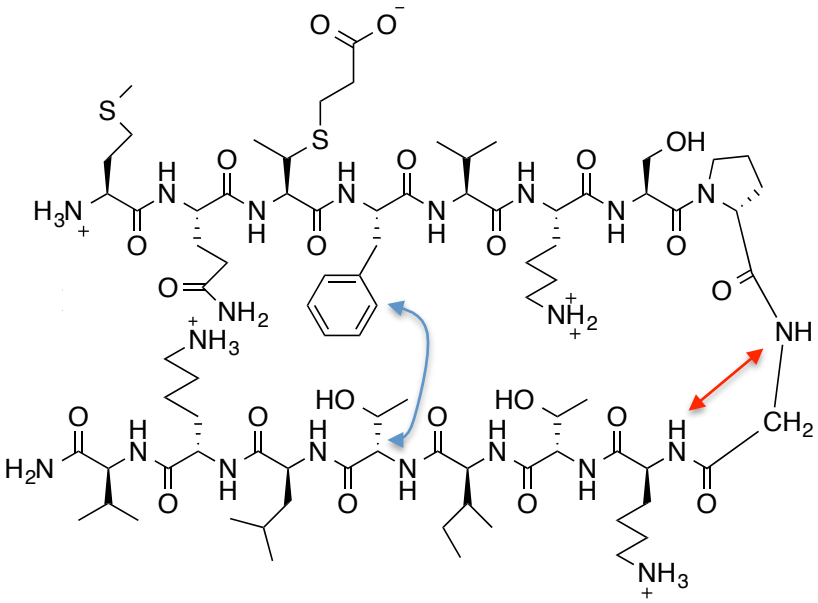
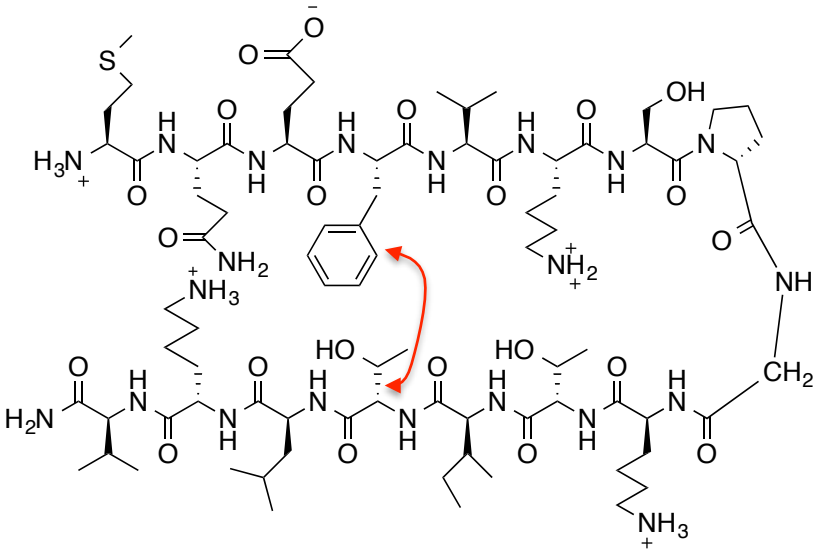


Figure B2.7. Cross-strand ROEs observed in peptide 3.15 (Ile3 → TS^-, 4 mM) at 4ºC in NMR buffer (9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate). Strong ROEs are red. Additional strong or medium ROE(s) are observed from Lys10 to Ser7 and/or Gly9 which is consistent with the expected hydrogen bonding pattern.



Supplementary Section B3
Chemical Shifts of Peptides from 2D NMR

General. Details for the sample preparation and acquisition of NMR spectra can be found in the methods section of Chapter 3. All unfolded controls are the peptide sequence of interest with LPro in place of DPro. See Chapter 3 for further explanation. All cyclic peptides (fully folded controls) are the sequence of interest cyclized with a two-residue DPro-Gly turn.

Table B3.1. ^{1}H chemical shifts of peptide 3.7a (Lys9 → TS^{+}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Sidechain
Arg1		4.049	1.899	1.556	3.204		7.246
			1.94				
Tyr2	9.031	5.134	2.845				6.867
							6.928
Val3	9.019	4.405	2.006	0.85			
Glu4	8.661	5.052	2.291	1.913			
				2.018			
Val5	8.882	4.615	1.973	0.9302			
DPro6		4.394	2.389	2.059	3.88		
			1.996	2.145			
Gly7	8.743	3.9195					
Orn8	8.032	4.65	1.849	1.680	3.02		7.7
				1.768			
TS^{+}9	8.97	4.85	3.11	1.192	2.793	3.091	
Ile10	9.155	4.557	1.892	0.876	0.806		
				1.118			
Leu11	8.649	4.23	1.614	1.425	0.599		
					0.722		
Gln12	8.57	4.303	2.027	2.283			7.007
							7.553
Cterm	7.25						
	7.878						

Table B3.2. ^{1}H chemical shifts of peptide 3.7a (Lys9 → TS^{+}) at 10ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.046	1.903 1.957	1.557	3.195		7.224
Tyr2	8.993	5.12	2.837				6.815 7.004
Val3	8.986	4.393	1.999	0.8568			
Glu4	8.619	5.034	1.914 2.005	2.27			
Val5	8.846	4.612	1.969	0.933			
DPro6		4.38	1.99 2.384	2.053 2.151	3.864		
Gly7	8.71	3.82, 4.019					
Orn8	8.021	4.642	1.841	1.667 1.728	3.009		7.667
TS^{+}9	8.921	4.84	3.106	1.19	2.788	3.095	
Ile10	9.12	4.536	1.889	0.972 1.125 1.359	0.8166		
Leu11	8.611	4.251	1.442 1.613	1.394	0.609 0.727		
Gln12	8.533	4.321	1.836 2.022	2.269	2.285		6.955 7.516
Cterm	7.205 7.825						

Table B3.3. ^{1}H chemical shifts of cyclic peptide 3.7c (cyclic Lys9 → TS^{+}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1	7.791	4.622	1.751	1.565	3.178		7.181
			1.813				
Tyr2	8.719	5.005	2.648				6.775
			2.770				6.957
Val3	9.351	4.486	2	0.846			
				0.882			
Glu4	8.711	5.067	1.925	2.269			
			2.050				
Val5	9.055	4.633	1.946	0.9194			
DPro6		4.376	1.991	2.053	3.878		
			2.398	2.154			
Gly7	8.809	3.850					
		4.053					
Orn8	7.971	4.724	1.835	1.694	3.009		7.694
TS^{+}9	9.036	5.061	3.117	1.238	2.827	3.113	
Ile10	9.302	4.717	1.878	0.865	0.8085		
				1.053			
				1.248			
Leu11	8.516	4.357	1.173	1.101	0.150		
			1.550	1.152	0.489		
Gln12	9.238	4.875	1.802	2.205			6.932
			1.998				7.499
Pro13		4.303	1.932	2.013	3.718		
			2.324	2.139	3.787		
Gly14	8.541	3.645					
		3.938					

Table B3.4. ^{1}H chemical shifts of LPro diastereomer (unfolded control) peptide 3.7b (unfolded Lys9 → TS^{+}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.018	2.324	1.682	3.198		7.263
Tyr2	8.893	4.619	2.908				6.807
			3.029				7.095
Val3	8.133	3.951	1.864	0.824			
				0.866			
Glu4	8.454	4.192	1.957	2.371			
Val5	8.505	4.4	2.066	0.958			
LPro6		4.419	1.932	1.982	3.694		
			2.219	2.059	3.879		
Gly7	8.636	3.966					
		4.177					
Orn8	8.414	4.415	1.784	1.707	1.876		
TS^{+}9	8.646	4.492	3.215	1.299	2.897	3.222	7.819
Ile10	8.633	4.178	1.834	0.888	0.8437		
				1.212			
				1.482			
Leu11	8.606	4.382	1.645	1.571	0.867		
					0.927		
Gln12	8.551	4.285	1.975	2.377			6.985
			2.092				7.665
C Term	7.226						
	7.740						

Table B3.5. ^{1}H chemical shifts of peptide 3.14a (Glu4 → TS^{-}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.056	1.891	1.548	3.201		7.27
			1.972				
Tyr2	9.015	5.343	2.802				6.819
							6.962
Val3	9.006	4.496	2.049	0.843			
				0.892			
TS^{-}4	8.934	5.19	3.105	1.093			2.587
							2.833
Val5	8.938	4.659	1.972	0.942			
DPro6		4.359	1.972	2.060	3.876		
			2.399	2.172	3.927		
Gly7	8.483	3.673					
		4.005					
Orn8	7.988	4.622	1.821	1.650	3.006		7.695
				1.723			
Lys9	8.663	4.721	1.548	1.118	1.226	2.357	7.376
					1.276	2.462	
Ile10	9.4	4.538	1.926	0.876	0.795		
				1.178			
Leu11	8.79	4.176	1.457	1.432	0.600		
			1.655		0.737		
Gln12	8.543	4.33	1.816	2.273			7.084
			2.048				7.448
C Term.	7.256						
	7.898						

Table B3.6. ^{1}H chemical shifts of cyclic peptide 3.14c (cyclic Glu4 → TS^{-}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1	7.779	4.662	1.798	1.524	3.183		
			1.832	1.611			
Tyr2	8.688	5.164	2.649				6.782
			2.809				6.944
Val3	9.325	4.525	2.048	0.843			
				0.894			
TS^{-}4	8.904	5.167	3.103	1.099			2.523
							2.593
							2.807
Val5	8.994	4.665	1.96	0.900			
				0.9435			
DPro6		4.347	1.967	2.057	3.869		
			2.397	2.172	3.929		
Gly7	8.487	3.67					
		3.99					
Orn8	7.946	4.672	1.822	1.639	3.01		7.688
				1.698			
Lys9	8.668	4.898	1.614	1.183	1.36	2.399	7.364
			1.677	1.297		2.498	
Ile10	9.46	4.655	1.88	0.8876	0.8065		
				1.123			
				1.366			
Leu11	8.59	4.241	1.209	1.131	0.214		
			1.556		0.489		
Gln12	9.184	4.86	1.811	2.213			6.915
			2.003				7.476
DPro13		4.314	1.936	2.005	3.686		
			2.305	2.138	3.769		
Gly14	8.576	3.74					
		3.915					

Table B3.7. ^{1}H chemical shifts of LPro diastereomer (unfolded control) peptide 3.14b (Glu4 → TS^{-}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.004	1.888	1.577	3.195		7.264
Tyr2	8.827	4.688	2.987				6.820
							7.122
Val3	8.296	4.12	1.945	0.8813			
TS^{-}4	8.362	4.41	3.169	1.275			2.551
							2.793
Val5	8.614	4.414	2.066	0.9653			
LPro6		4.392	1.926	1.973	3.698		
			2.318	2.063	3.899		
Gly7	8.616	3.957					
Orn8	8.359	4.355	1.858	1.737	3.012		7.693
Lys9	8.556	4.311	1.763	1.376	1.679	2.989	7.617
				1.447			
Ile10	8.477	4.122	1.833	0.888	0.8497		
				1.194			
				1.500			
Leu11	8.558	4.381	1.646	1.586	0866		
					0.938		
Gln12	8.519	4.283	1.977	2.376			6.989
			2.107				7.660
C Term.	7.234						
	7.737						

Table B3.8. ^{1}H chemical shifts of peptide 3.10 (Thr2 → TO^{+}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Thr 1		3.991	4.148	1.123			
TO^{+}2	8.959	4.708	3.809	1.123	3.159		
					3.751		
					3.466		
Arg 3	8.671	4.475	1.726	1.484	3.139	7.178	
			1.803				
Tyr 4	8.571	5.124	2.795				6.961, 6.777 (Aromatic)
Val 5	8.905	4.380	1.991	0.852			
Glu 6	8.655	4.907	1.977	2.267			
			1.881				
Val 7	8.942	4.595	1.968	0.943			
DPro 8		4.378	1.990	2.066,	3.881		
			2.385	2.129			
Gly 9	8.712	3.790					
		4.021					
Orn 10	7.977	4.589	1.831	1.676	3.010	7.697	
Lys 11	8.604	4.573	1.570	1.146	1.368	2.635	7.464
			1.672				(NH_3)
Ile 12	9.105	4.468	1.920	1.231	0.802		
				1.218			
				0.892			
Leu 13	8.669	4.229	1.520	1.432	0.701		
			1.590		0.738		
Gln 14	8.817	4.509	4.509	2.295			7.021, 7.491
			4.513				(NH_2)
Thr 15	8.587	4.576	4.214	1.19			
Thr 16	8.29	4.36	4.241	1.184			
C-Term	7.284						
	7.781						

Table B3.9. ^{1}H chemical shifts of peptide 3.11 (Thr2 → TS^{+}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Thr1		3.926	4.11	1.273			
TS^{+}2	9.046	4.581	3.138	1.163	2.876	3.203	
Arg3	8.726	4.498	1.700	1.504	3.142		7.179
			1.786				
Tyr4	8.599	5.109	2.781				6.769
							6.944
Val5	8.918	4.365	2.005	0.852			
				0.872			
Glu6	8.643	4.928	1.878	2.241			
			1.971				
Val7	8.921	4.609	1.983	0.9342			
DPro8		4.382	1.990	2.061	3.875		
			2.384	2.126			
Gly9	8.711	4.015					
		3.779					
Orn10	7.97	4.598	1.827	1.703	3.016		7.687
Lys11	8.593	4.582	1.562	1.145	1.358	2.626	7.449
			1.645				
Ile12	9.095	4.48	1.923	0.893	0.8036		
				1.220			
				1.402			
Leu13	8.678	4.189	1.518	1.426	0.7101		
			1.565				
Gln14	8.768	4.491	1.846	2.282			7.023
			2.045				7.473
Thr15	8.562	4.576	4.256	1.193			
Thr16	8.215	4.364	4.281	1.191			
C Term	7.287						
	7.770						

Table B3.10. ^{1}H chemical shifts of peptide 3.8 Lys15 → TS^{+} at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.16	2.155	2.53			
Gln2	8.881	4.791	1.748	2.148			7.014
			1.850	2.211			7.403
Ile3	8.652	4.241	1.76	0.826	0.7843		
				1.029			
				1.336			
Phe4	8.573	5.223	2.917				7.197
							7.297
							7.269
Val5	8.936	4.331	2.015	0.8829			
Lys6	8.688	4.788	1.761	1.418	1.63	2.936	7.554
Ser7	8.766	4.911	3.709				
Pro8		4.421	1.985	2.042	3.776		
			2.356	2.100	3.844		
Gly9	8.807	3.926					
Lys10	8.085	4.62	1.827	1.428	1.69	3.001	7.62
			1.865				
Thr11	8.636	4.692	3.987	1.123			
Ile12	9.138	4.485	1.912	0.903	0.8395		
				1.150			
				1.410			
Thr13	8.561	4.674	4.007	1.033			
Leu14	8.894	4.563	1.514	1.506	0.828		
					0.864		
TS^{+}15	8.743	4.734	3.212	1.272	2.897	3.194	
Val16	8.556	4.199	2.016	0.877			
				0.923			
C Term	7.290						
	7.967						

Table B3.11. ^{1}H chemical shifts of peptide 3.12 (Ile3 → TO^{+}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.161	2.164	2.598			
Gln2	8.937	4.479	1.951	2.269			7.027
							7.633
TO^{+}3	8.62	4.441	3.927	1.14			3.128
							3.445
							3.766
Phe4	8.45	5.306	2.953				7.202
			3.044				7.314
							7.264
Val5	8.582	4.17	1.99	0.8923			
Lys6	8.579	4.539	1.791	1.444	1.671	2.981	7.608
Ser7	8.695	4.835	3.771				
DPro8		4.435	1.983	2.051	3.794		
			2.325				
Gly9	8.639	3.925					
Lys10	8.199	4.511	1.819	1.438	1.675	2.978	
Thr11	8.562	4.46	4.039	1.125			
Ile12	8.757	4.349	1.876	0.895	0.848		
				1.155			
				1.4476			
Thr13	8.52	4.418	4.076	1.152			
Leu14	6.207	4.376	4.376	1.562	0.852		
					0.905		
Lys15	8.538	4.356	1.791	1.437	1.673	2.978	
Val16	8.34	4.065	2.025	0.948			
Cterm	7.234	7.873					

Table B3.12. ^{1}H chemical shifts of peptide 3.13 (Ile3 → TS^{+}) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.164	2.155	2.586			
Gln2	8.916	4.479	1.87	2.216			7.031
							7.604
TS^{+}3	8.663	4.458	3.165	1.261	2.837	3.17	
Phe4	8.66	4.959	2.923				7.200
			3.015				7.298
							7.254
Val5	8.669	4.187	1.995	0.8931			
Lys6	8.601	4.58	1.752	1.417	1.655	2.972	7.606
Ser7	8.705	4.849	3.76				
DPro8		4.423	1.988	2.062	3.795		
			2.330				
Gly9	8.661	3.927					
Lys10	8.185	4.525	1.82	1.432	1.679	2.977	7.606
Thr11	8.571	4.495	4.034	1.125			
Ile12	8.808	4.378	1.888	0.887	0.8414		
				1.159			
				1.424			
Thr13	8.527	4.441	4.06	1.131			
Leu14	8.549	4.39	1.554	1.562	0.843		
					0.894		
Lys15	8.546	4.369	1.793	1.414	1.673	2.982	7.603
Val16	8.332	4.076	2.019	0.9429			
C Term	7.233						
	7.878						

Table B3.13. ^{1}H chemical shifts of peptide 3.16 (Ile3 → TS^{-}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.153	2.159	2.569			
Gln2	8.847	4.545	1.871	2.2			7.021
							7.596
TS^{-}3	8.568	4.466	3.153	1.24			2.536
							2.734
Phe4	8.596	5.011	2.912				7.196
			3.028				7.307
							7.244
Val5	8.688	4.215	2.004	0.8951			
Lys6	8.617	4.642	1.756	1.416	1.651	2.976	7.604
Ser7	8.706	4.847	3.748				
DPro8		4.425	1.991	2.032	3.794		
			2.335	2.080			
Gly9	8.679	3.923					
Lys10	8.157	4.546	1.837	1.427	1.686	2.994	7.615
Thr11	8.577	4.541	4.023	1.117			
Ile12	8.874	4.408	1.895	0.894	0.8353		
				1.146			
				1.426			
Thr13	8.516	4.486	4.061	1.12			
Leu14	8.579	4.408	1.554	1.559	0.838		
					0.881		
Lys15	8.522	4.398	1.753	1.411	1.668	2.979	7.599
Val16	8.333	4.084	2.015	0.9386			
C Term	7.235						
	7.886						

Table B3.14. ^{1}H chemical shifts of peptide 3.15 (Ile3 → Glu) at 4°C in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Met1		4.148	2.162	2.584			
Gln2	8.868	4.395	1.915	2.224			7.009
							7.669
Glu3	8.653	4.358	1.863	2.240			
			1.956	2.329			
Phe4	8.509	4.946	2.920				7.193
			3.047				7.314
							7.248
Val5	8.622	4.2	1.998	0.8947			
Lys6	8.606	4.627	1.756	1.414	1.659	2.944	7.604
Ser7	8.696	4.854	3.75				
DPro8		4.425	1.992	2.032	3.798		
			2.330	2.071			
Gly9	8.665	3.923					
Lys10	8.167	4.533	1.826	1.462	1.686	2.99	7.616
Thr11	8.564	4.514	4.025	1.119			
Ile12	8.837	4.383	1.891	0.885	0.8375		
				1.152			
				1.428			
Thr13	8.523	4.436	4.058	1.137			
Leu14	8.536	4.38	1.551	1.546	0.8623		
Lys15	8.524	4.365	1.742	1.366	1.667	2.97	7.596
Val16	8.341	4.075	2.015	0.9385			
C Term	7.233						
	7.877						

Table B3.15. ^{1}H chemical shifts of peptide 3.10 (Thr10 → TO^{+}) at 4ºC in 9:1 H_2O:D_2O, pH 3.8, 100 mM sodium acetate buffer (pH not corrected for isotope effects)

	H-N	H α	H β	H γ	H δ	H ε	Side Chain
Arg1		4.039	1.895	1.587	3.166		7.221
Trp2	8.902	4.781					10.16 7.124
							7.382 7.061
							7.208 7.461
Gln3	8.641	4.388	1.884	2.084			6.935 7.463
			1.967				
Tyr4	8.421	4.733	2.722				6.883 6.712
			2.898				
Val5	8.68	4.469	1.995	0.9015			
DPro6		4.381	1.942	2.009	3.709		
			2.322				
Gly7	8.333	3.782					
		3.982					
Lys8	7.93	4.438	1.735	1.347	1.638	2.94	7.611
Phe9	8.561	4.714	2.943				6.988 7.182
			2.837				7.063
Thr10	8.507	4.577	3.791	1.067	3.582	3.119	7.709
					3.722		
Val11	8.136	3.983	1.852	0.806			
				0.726			
Gln12	8.512	4.241	1.860	2.282			6.937 7.556
			2.01				
C Term	7.250						
	7.797						

Supplementary Section B4
NMR Structure Calculations Statistics

Table B4.1. NMR Statistics from CNS calculations for the cyclic hairpin 3.7c (cyclic Lys9 → TS^{+})

NMR Distance and Dihedral Constraints	Protein
Distance Constraints	
Total NOEs	33
Intra-residue	0
Inter-residue	33
Sequential (|*i-j*|=1)	16
Medium-range (|*i-j*|<4)	6
Long-range (|*i-j*|>5)	11
Intermolecular	0
Hydrogen bonds	0
Total dihedral angle restraints	
ϕ	0
φ	0
Structure statistics	
Violations (mean and s.d.)	0
Distance constraints violations (A)	0
Dihedral angle constraints (^{O})	0
Max. dihedral angle violation (^{O})	0
Max. distance constraint violation (A)	0
Deviations from idealized geometry	
Bond lengths (A)	±0.0018
Bond angles (^{O})	±0.302
Impropers (^{O})	±0.474
Average pairwise r.m.s. deviation (A) (18 structures)	
Heavy	1.6 ± 0.3
Backbone	0.5 ± 0.1

Table B4.2. NMR statistics from CNS calculations for the cyclic hairpin 3.14c (cyclic Lys9 → TS^-)

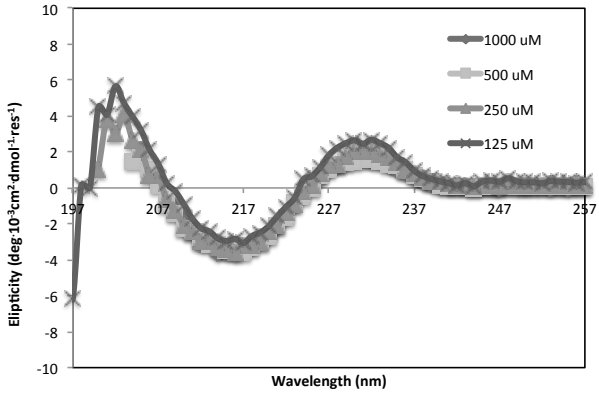
NMR Distance and Dihedral Constraints	Protein
Distance Constraints	
Total NOEs	50
Intra-residue	0
Inter-residue	50
Sequential (|*i-j*|=1)	29
Medium-range (|*i-j*|<4)	7
Long-range (|*i-j*|>5)	14
Intermolecular	0
Hydrogen bonds	0
Total dihedral angle restraints	
ϕ	0
φ	0
Structure statistics	
Violations (mean and s.d.)	0
Distance constraints (A)	0
Dihedral angle constraints (^{O})	0
Max. dihedral angle violation (^{O})	0
Max. distance constraint violation (A)	0
Deviations from idealized geometry	
Bond lengths (A)	± 0.0024
Bond angles (^{O})	± 0.362
Impropers (^{O})	±0.794
Average pairwise r.m.s. deviation (A) (11 structures)	
Heavy	1.9 ± 0.4
Backbone	0.7 ± 0.2

Supplementary Section B5
Circular Dichroism (CD)

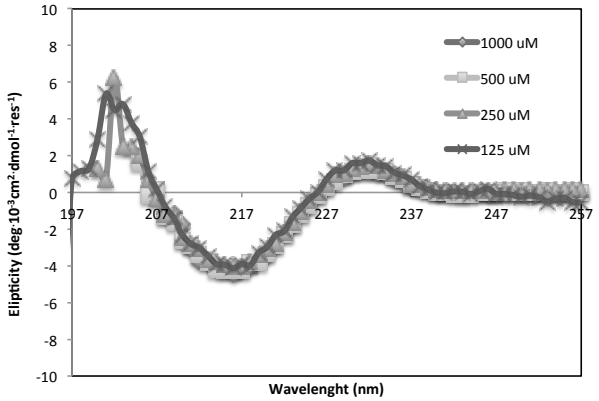
General. All peptides were dissolved in phosphate buffer (100 mM, pH=7.0), and an aliquot was used to determine the concentration by UV in 6 M guanidinium hydrochloride using the Tyr absorbance at 275 nm. CD samples were prepared by serial dilutions of the stock sample. All CD spectra were taken on an Aviv Model 420 Circular Dichroism Spectrometer with 1 mm pathlength cells. The parent peptide is shown in black or gray, the TS^+ peptide is shown in blue, and TS^- peptide is shown in green.

Figure B5.1. CD spectra of parent peptide 2.10a (Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2) in phosphate buffer (pH=7.0, 100 mM) at variable concentrations at **(A)** 4ºC, **(B)** 20ºC, and **(C)** 37ºC. All spectra have a minimum near 215 nm, which is characteristic of β-sheet secondary structure. The CD spectra at different concentrations are approximately the same, as expected if there is no self-association.

(A)



(B)



(C)

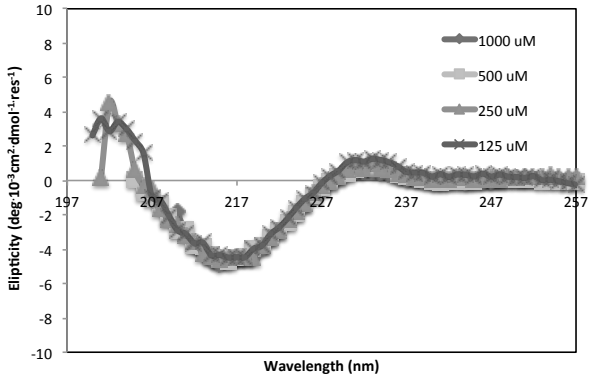
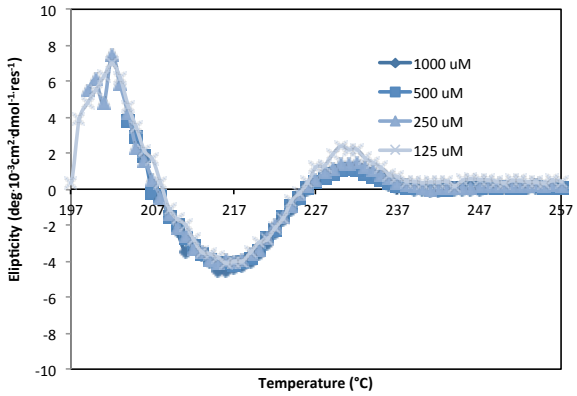
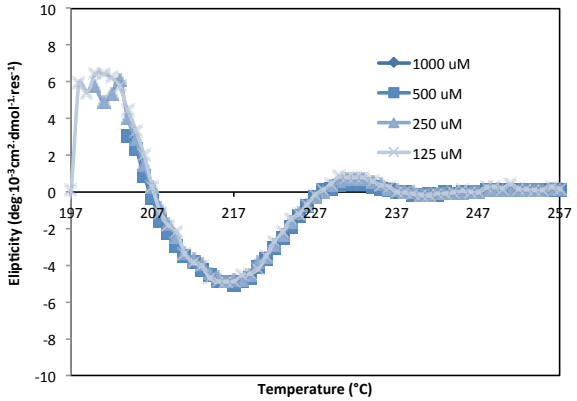


Figure B5.2. CD spectra of peptide 3.7a (Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^+-Ile-Leu-Gln-NH_2, Lys9→TS^+) in phosphate buffer (pH=7.0, 100 mM) at variable concentrations at **(A)** 4ºC, **(B)** 20ºC, and **(C)** 37ºC. All spectra have a minimum near 215 nm, which is characteristic of β-sheet secondary structure. The CD spectra at different concentrations are approximately the same, as expected if there is no self-association.

(A)



(B)



(C)

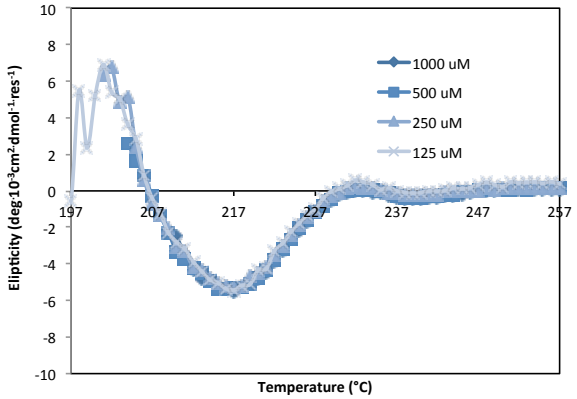
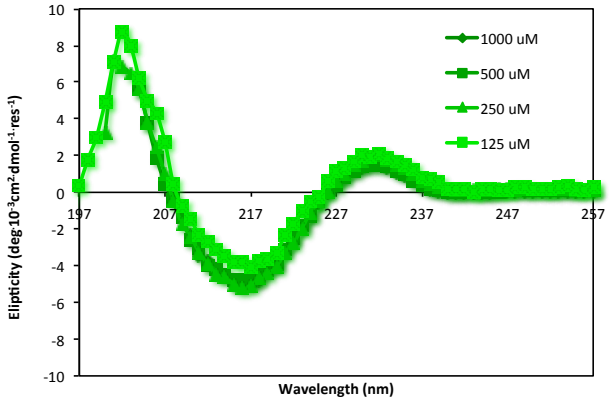
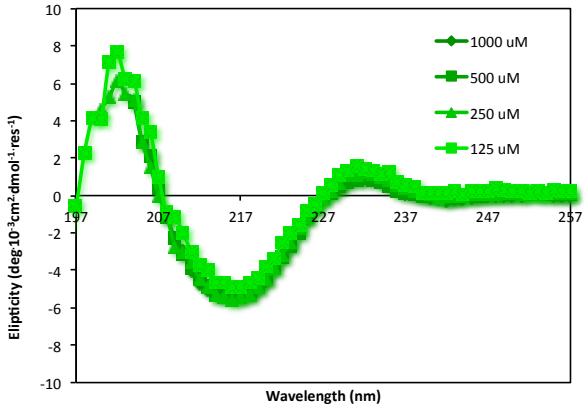


Figure B5.3. CD spectra of peptide 3.14a (Arg-Tyr-Val-TS$^-$-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2, Lys9→TS$^-$) in phosphate buffer (pH=7.0, 100 mM) at variable concentrations at **(A)** 4ºC, **(B)** 20ºC, and **(C)** 37ºC. All spectra have a minimum near 215 nm, which is characteristic of β-sheet secondary structure. The CD spectra at different concentrations are approximately the same, as expected if there is no self-association. (Note that the CD signature at 125 μM is less accurate due to the poor S/N.)

(A)



(B)



(C)

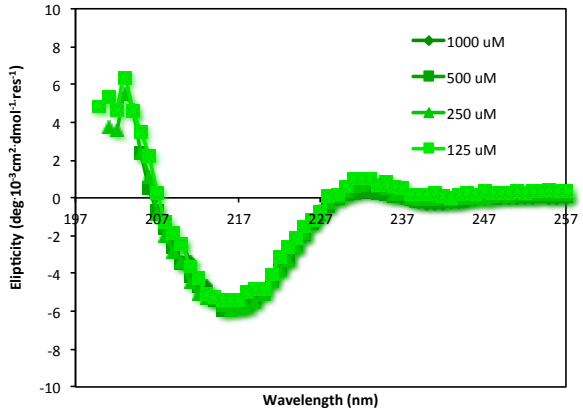
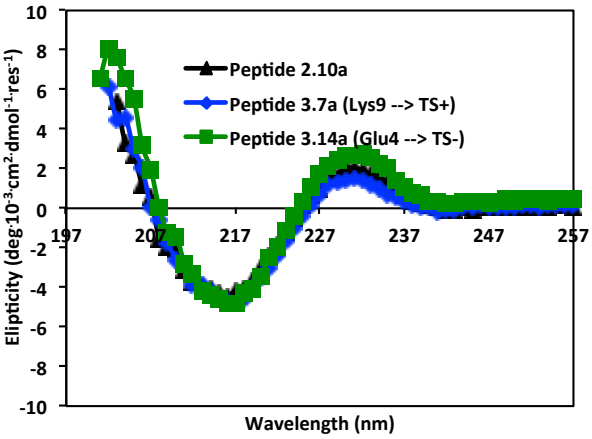


Figure B5.4. CD spectra of peptides 2.10a, 3.7a, and 3.14a (parent sequence Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2, 250 μM) in acetate buffer (pH=3.8, 100mM, 4ºC). All spectra have a minimum near 215 nm, which is characteristic of β-sheet secondary structure. The relative intensity of the minima at 215 nm is similar for all three peptides, though slightly less intense for the parent peptide. The data do not show a clear difference among the conformational stability of the peptides. The data does not contradict the relative stabilities determined by NMR because chemical shift deviations are a more accurate and precise method of measuring relative conformational stability. Unlike NMR chemical shift deviations, the magnitude of the minima in the CD spectra is sensitive to error in the measurement of the concentration of the samples. (Additionally, note that β-sheet CD signatures are less intense than α-helix signatures and therefore less useful in determining the relative stability of peptides as they have smaller dynamic ranges.)



Supplementary Section B6
Dilution Studies

General. Dilution Studies were performed for all peptides in NMR buffer (100 mM sodium acetate in 9:1 H_2O:D_2O at pH=3.8) at 4ºC to check for aggregation. The most concentrated samples were the concentration used for acquisition of the 2D NMR spectra. Water suppression was performed with a presaturation pulse sequence as described in Chapter 2. Where needed, baselines were corrected by manually defining baseline regions. In all cases, chemical shifts of the concentrated and diluted peptide were the same, which indicates that the peptide aggregation state did not change over this concentration range. We assume that each peptide was therefore monomeric (i.e., not self-associated) throughout the range. All parent peptides were previously characterized by sedimentation equilibrium analytical ultracentrifugation experiments.[2–5] All unfolded analogues are diastereomers of the peptides of interest with LPro instead of DPro. See supplementary section B1 for further explanation. All cyclic peptides, used to represent the fully populated β-hairpin, have the sequence of interest cyclized with a DPro-Gly turn (see main text Figure 3 for an example). For further sequence information refer to supplementary section B1.

Figure B6.1. Dilution study of peptide 3.7a, Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2

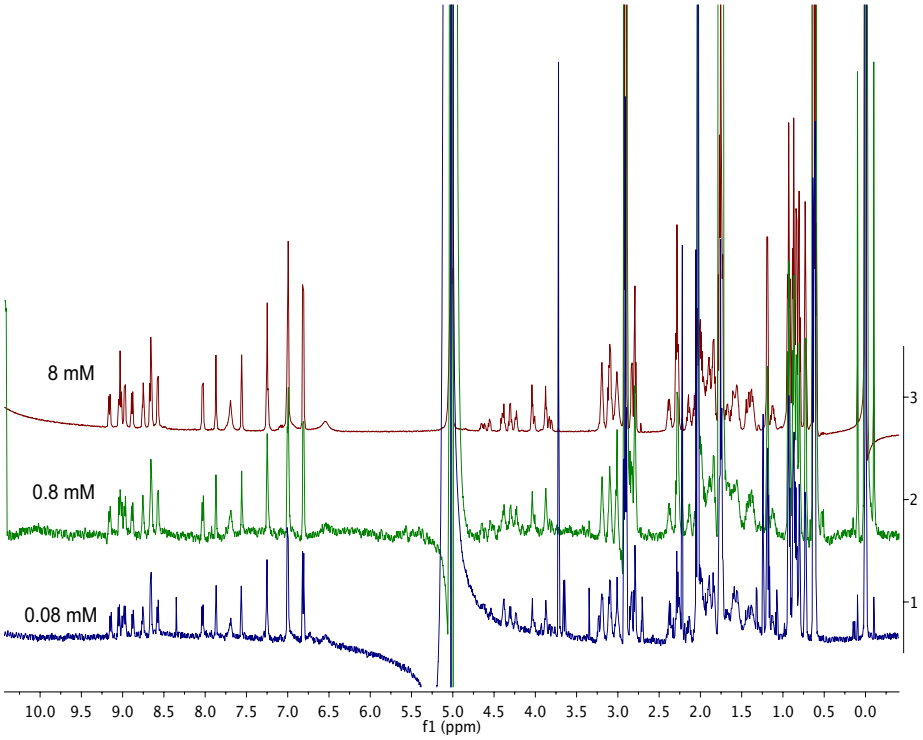
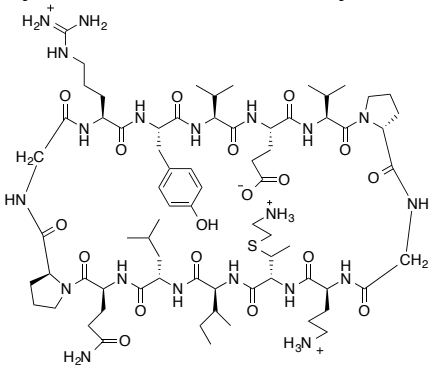


Figure B6.2. Dilution study of cyclic peptide 3.7c (fully folded control), cyclized Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-DPro-Gly



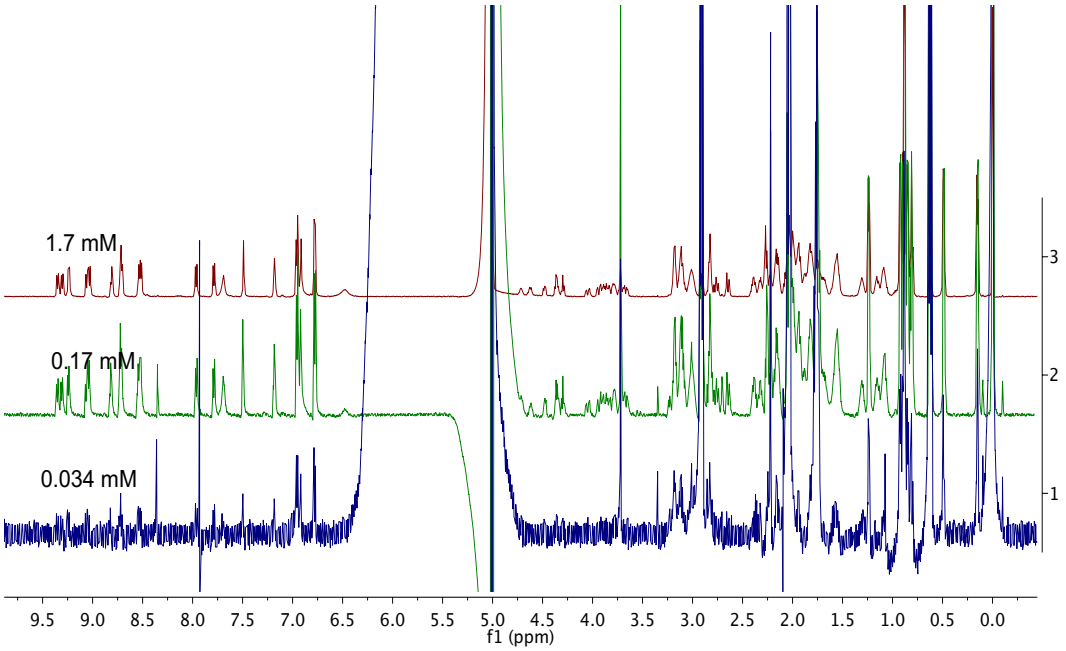


Figure B6.3. Dilution study of LPro diasteromer 3.7b (unfolded control), Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2

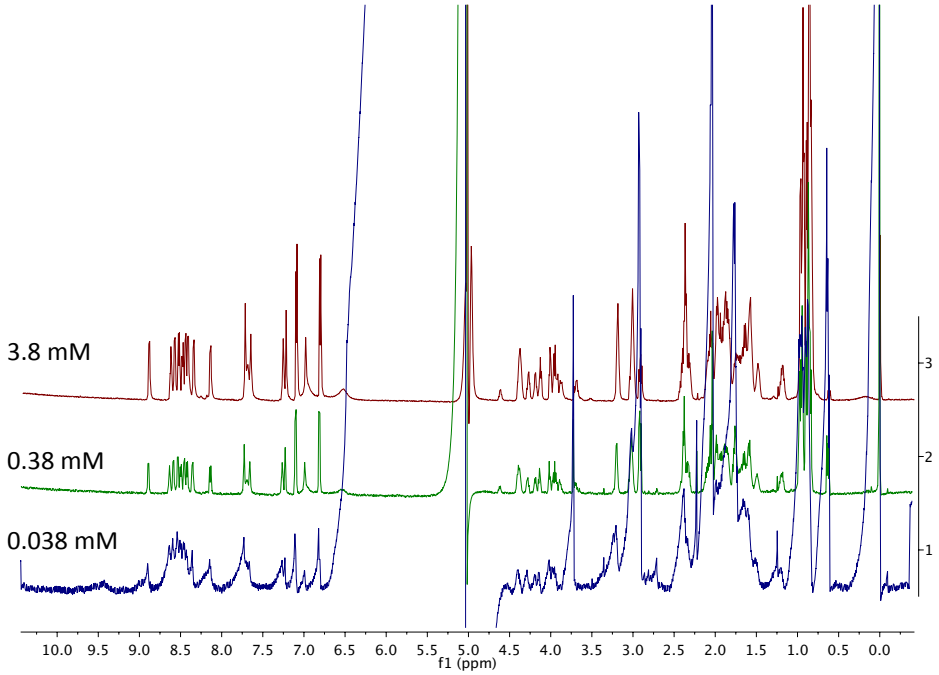


Figure B6.4. Dilution study of peptide 3.14a, Arg-Tyr-Val-TS^{-}-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2

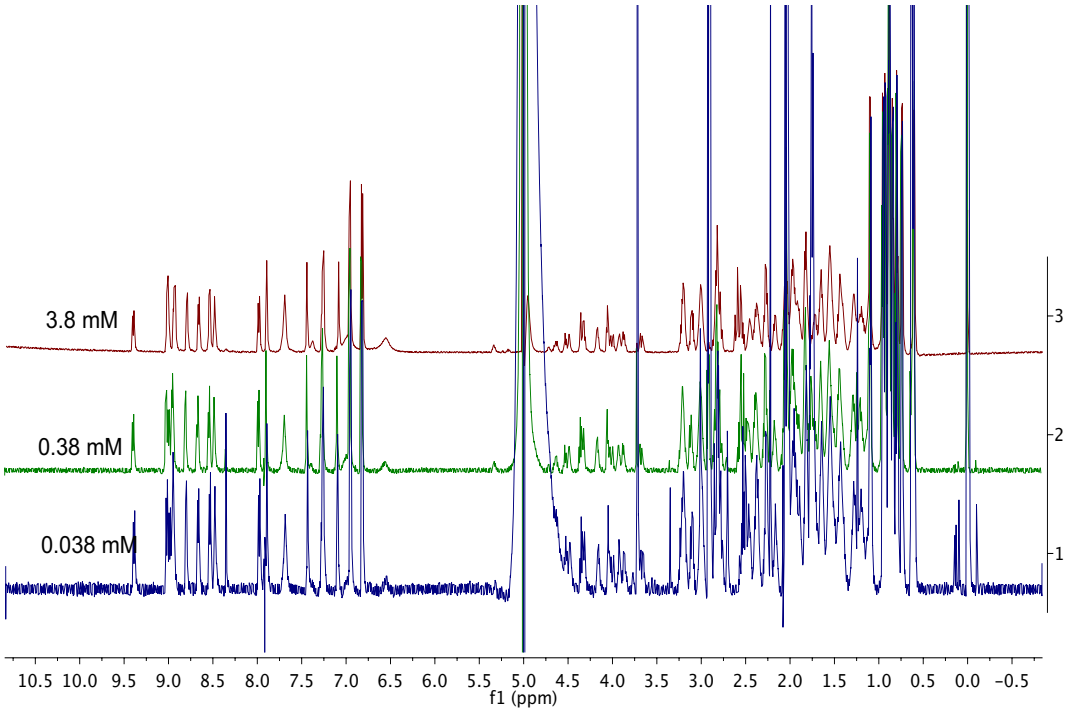
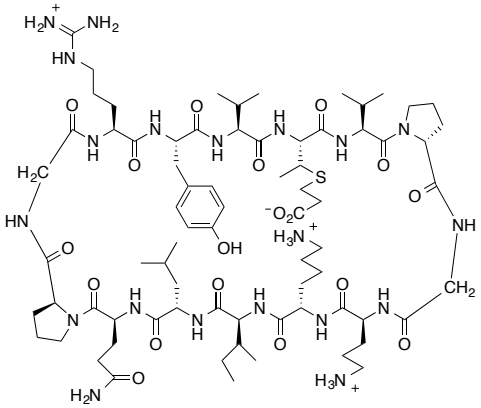


Figure B6.5. Dilution study of cyclic 3.14c (fully folded control), cyclized Arg-Tyr-Val-TS^{-}-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-DPro-Gly



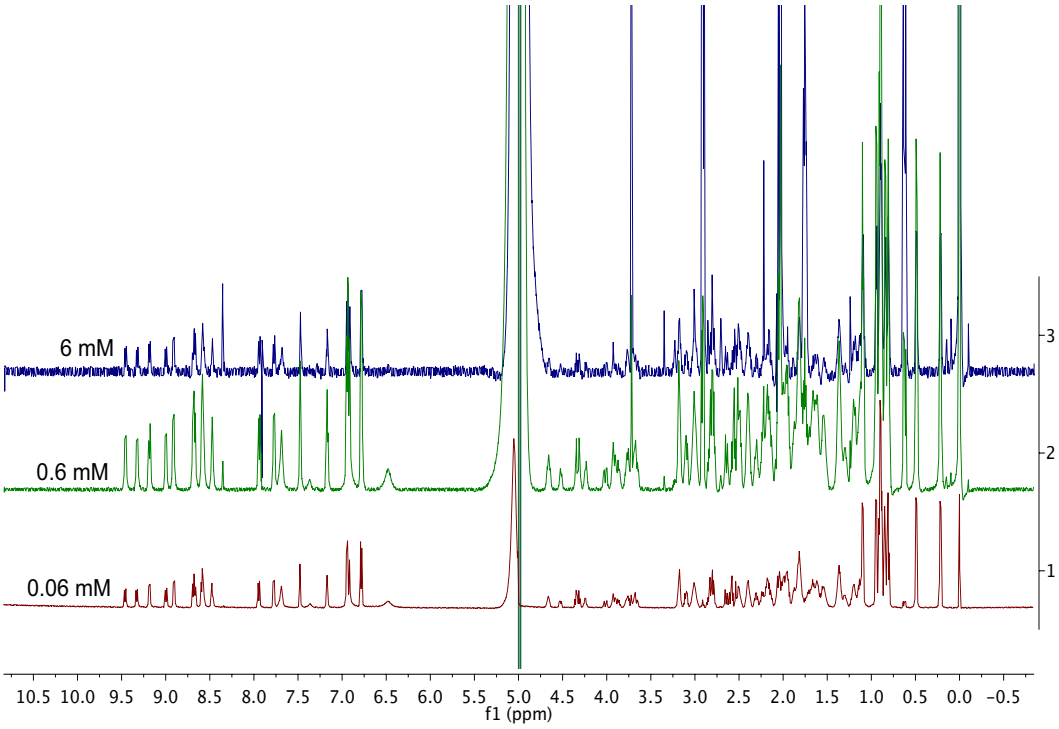


Figure B6.6. Dilution study of LPro diasteromer 3.14b (unfolded control), Arg-Tyr-Val-TS^{-}-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2

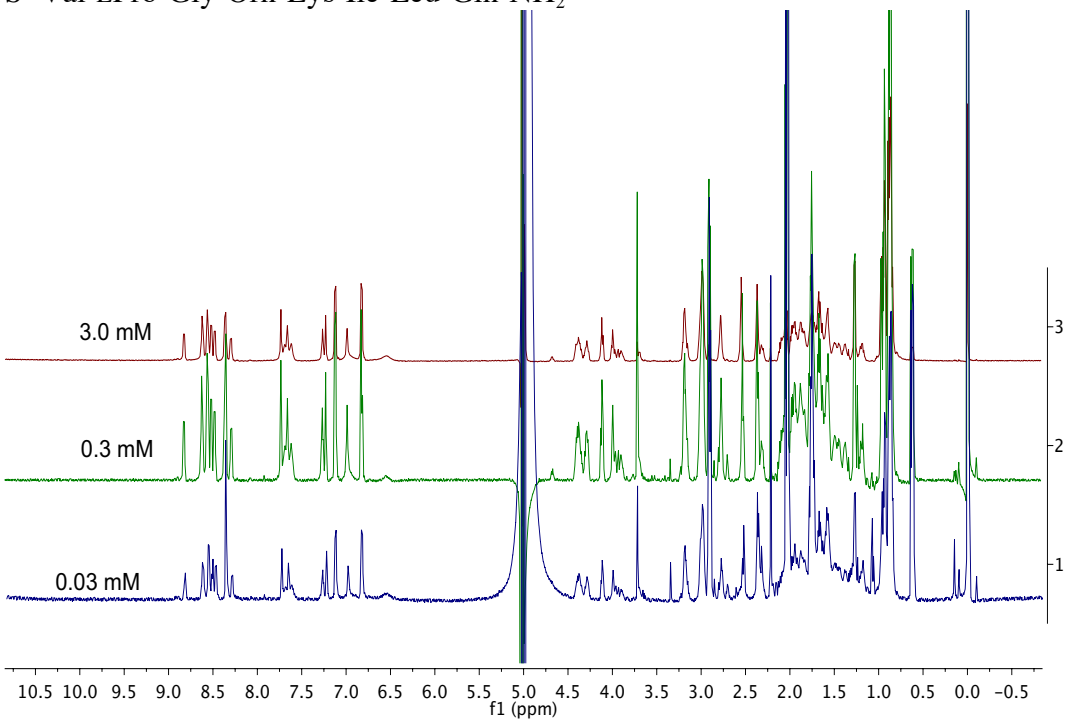


Figure B6.7. Dilution study of peptide 3.10, Thr-TO^{+}-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2

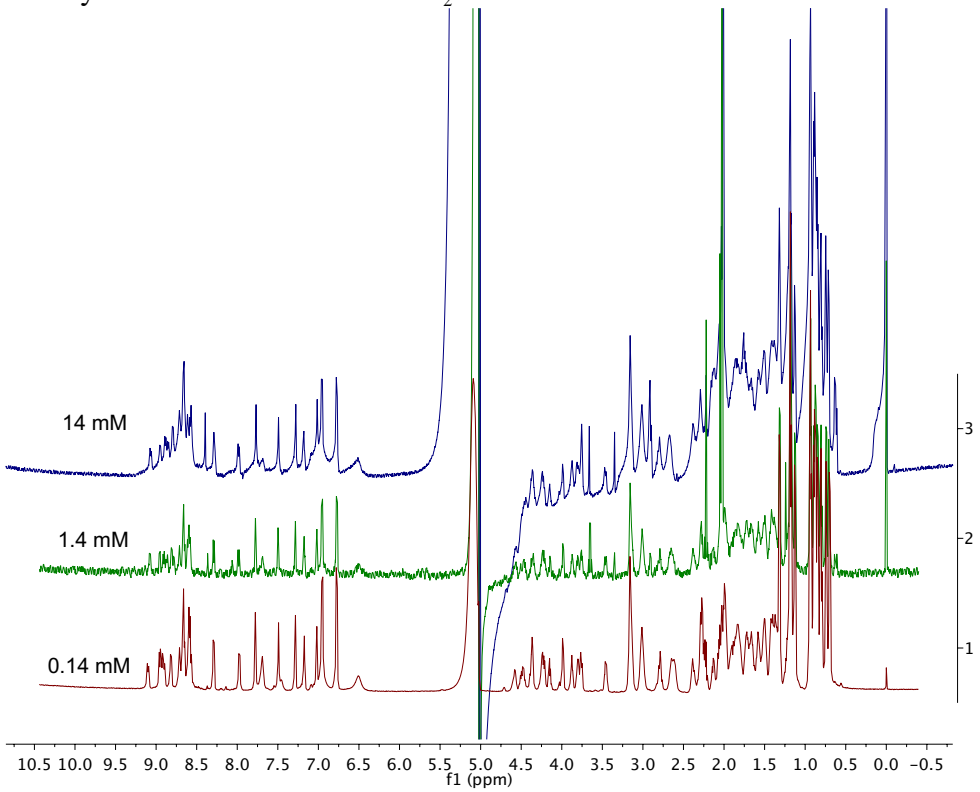


Figure B6.8. Dilution study of peptide 3.11, Thr-TS^{+}-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2

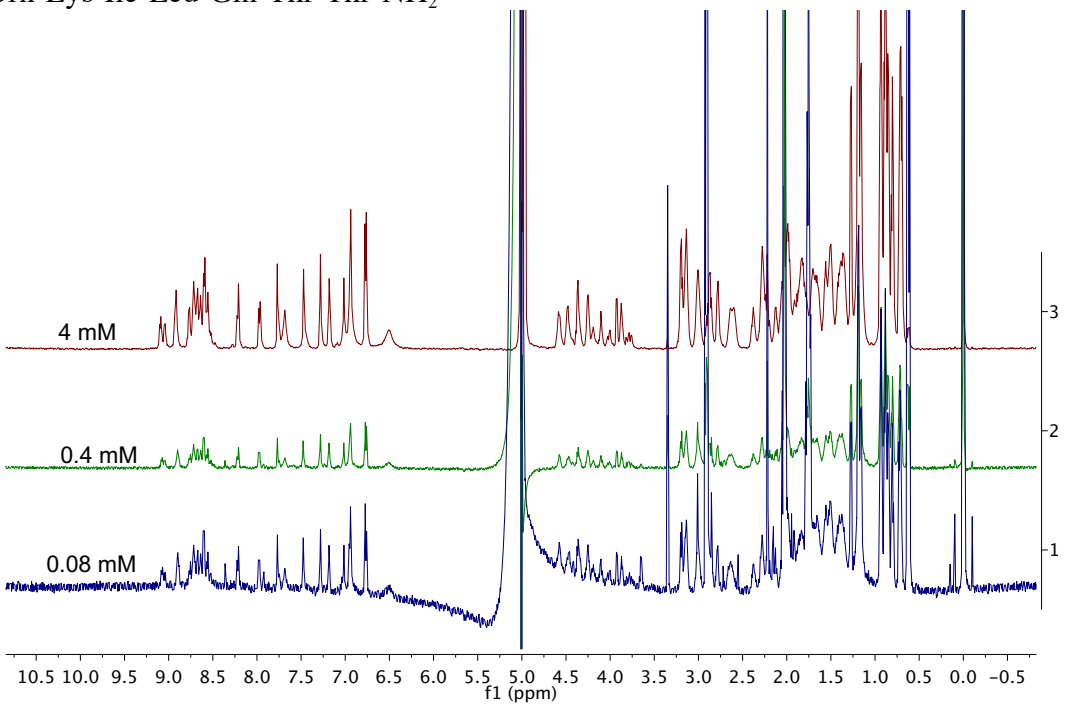


Figure B6.9. Dilution study of peptide 3.8, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-TS^{+}-Val-NH_2

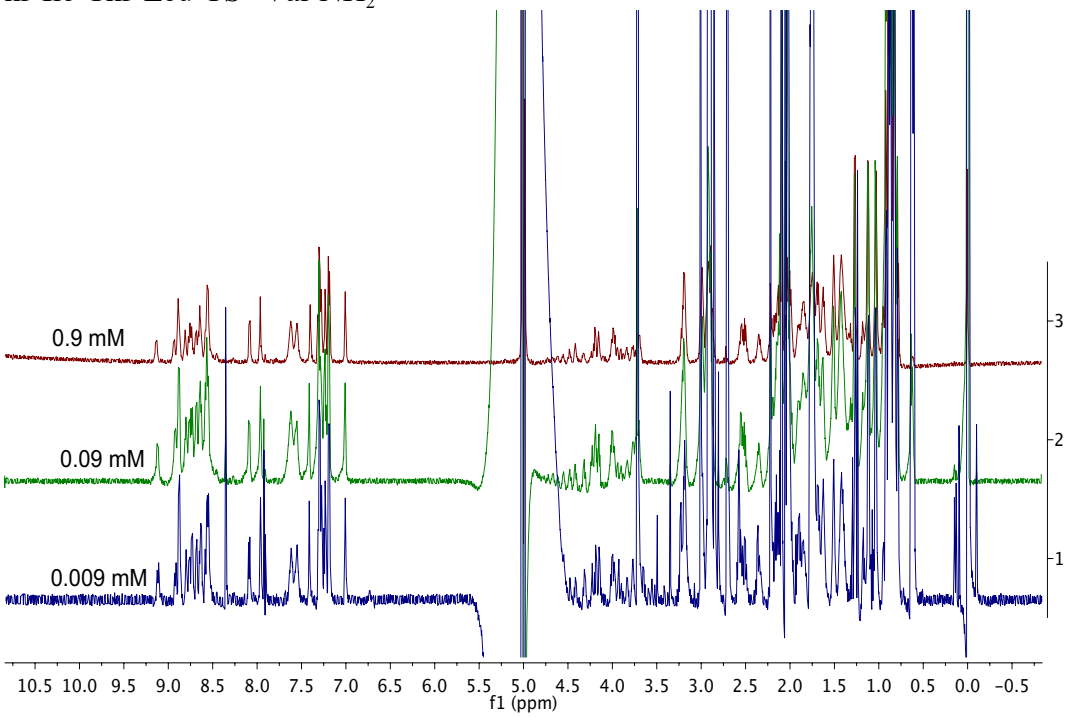


Figure B6.10. Dilution study of peptide 3.12, Met-Gln-TO^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

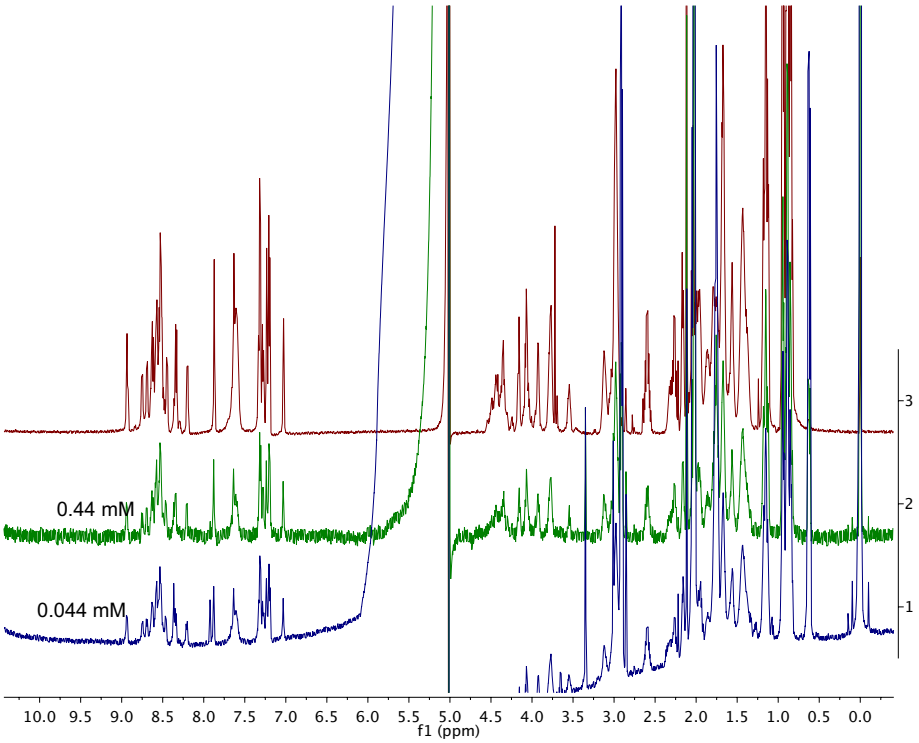


Figure B6.11. Dilution study of peptide 3.13, Met-Gln-TS^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

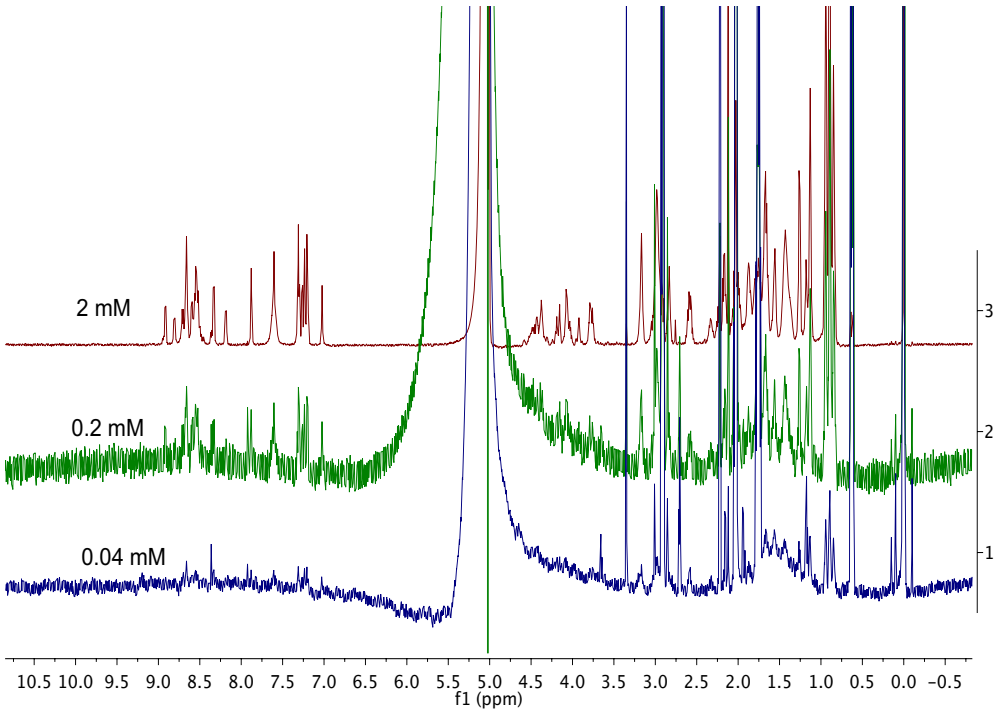


Figure B6.12. Dilution study of peptide 3.16, Met-Gln-TS^{-}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

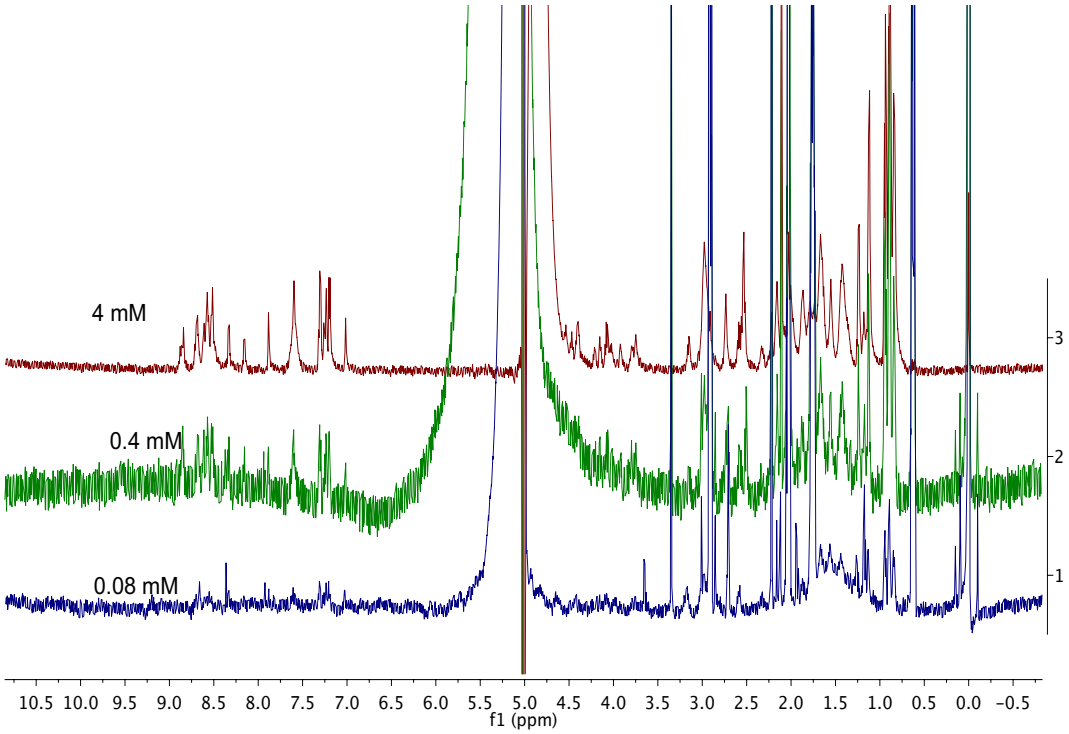
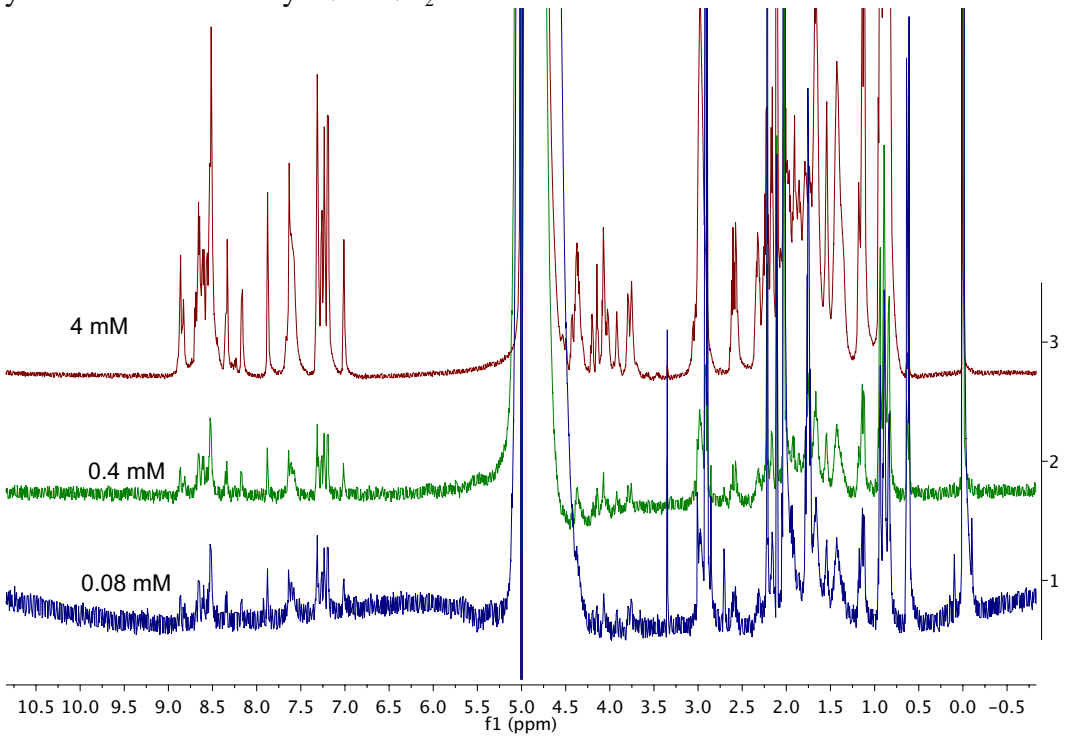


Figure B6.13. Dilution study of peptide 3.15, Met-Gln-Glu-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2



Supplementary Section B7
Peptide Purity Checks

General. For details on data acquisition, see Chapter 3 methods section.

Figure B7.1. Purity check of peptide 3.7a, Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2

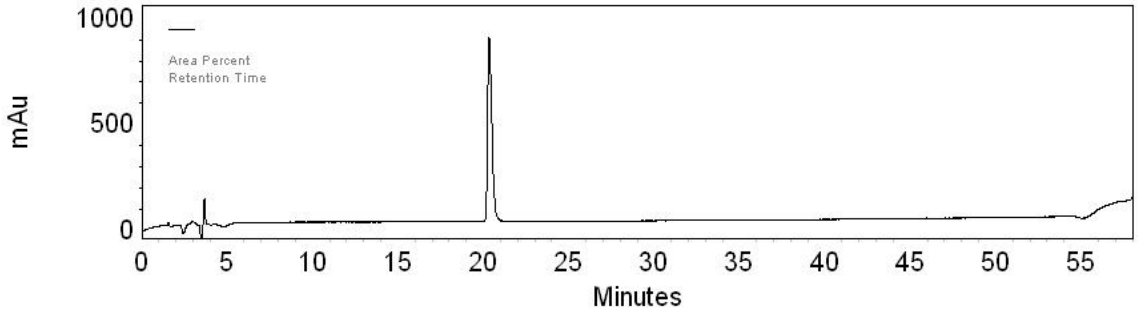


Figure B7.2. Purity check of cyclic peptide 3.7c, cyclized Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-DPro-Gly

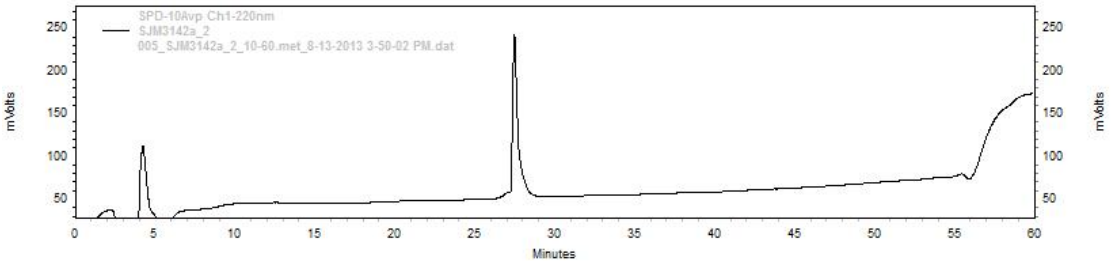


Figure B7.3. Purity check of LPro diasteromer 3.7b (unfolded control), Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2

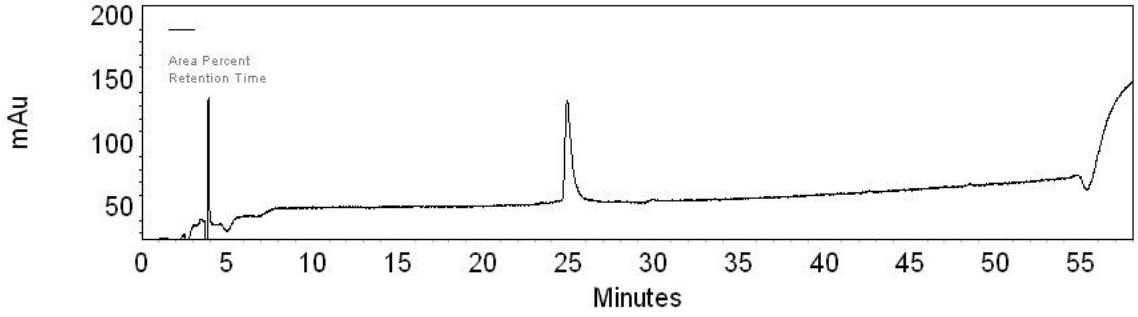


Figure B7.4. Purity check of peptide 3.14, Arg-Tyr-Val-TS$^-$-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2

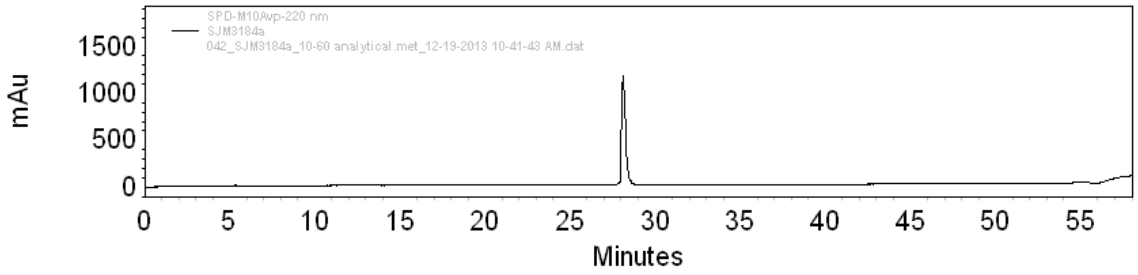


Figure B7.5. Purity check of cyclic peptide 3.14c (fully folded control), cyclized Arg-Tyr-Val-TS$^-$-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-DPro-Gly

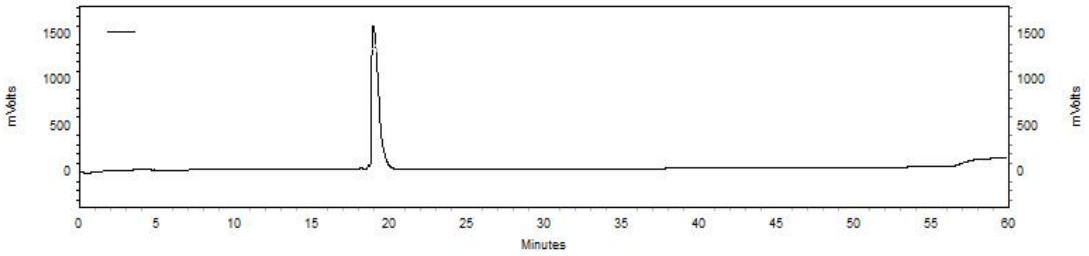


Figure B7.6. Purity check of LPro diastereomer 3.14b (unfolded control), Arg-Tyr-Val-TS$^-$-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2

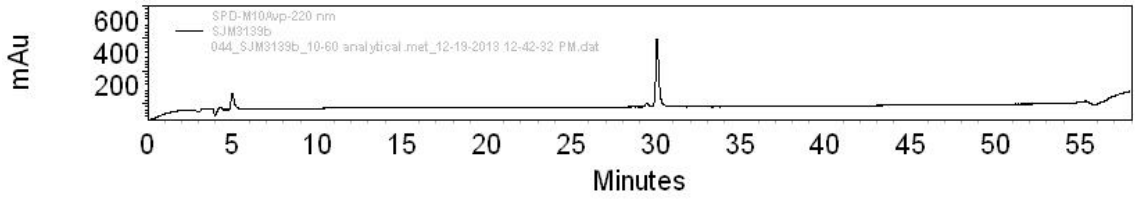


Figure B7.8. Purity check of peptide 3.10, Thr-TO$^+$-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2

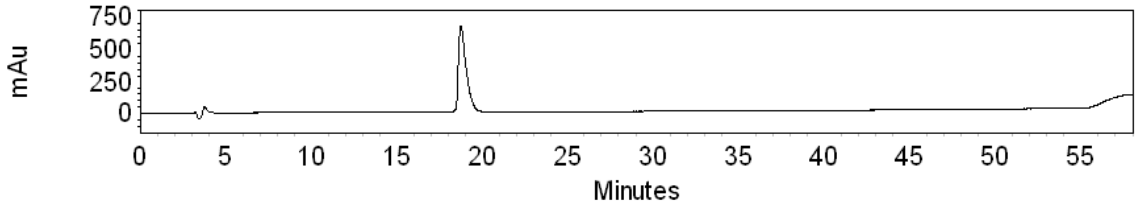


Figure B7.9. Purity check of peptide 3.11, Thr-TS^{+}-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2

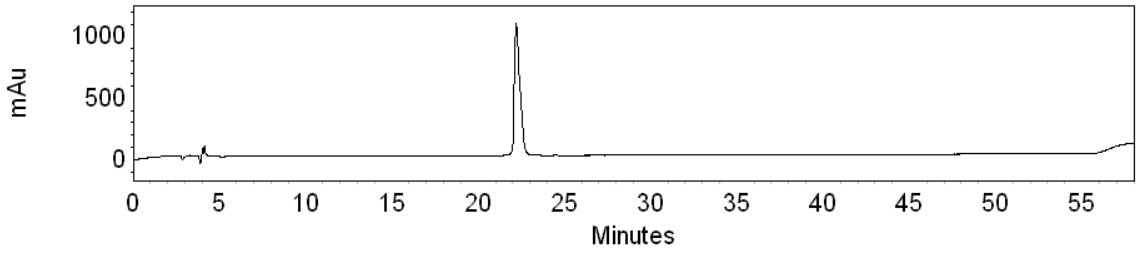


Figure B7.10. Purity check of peptide 3.8, Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-TS^{+}-Val-NH_2

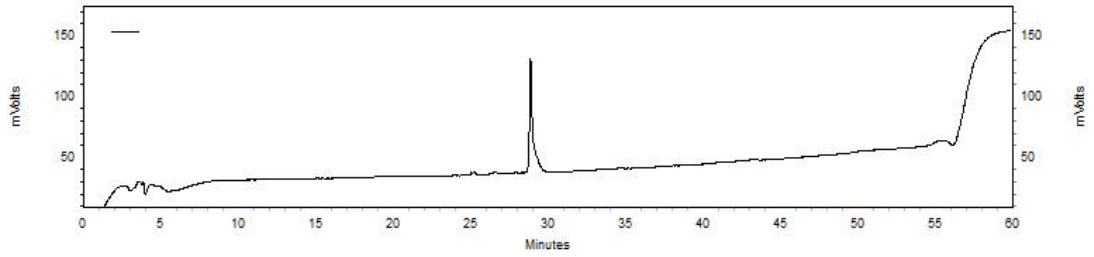


Figure B7.11. Purity check of peptide 3.12, Met-Gln-TO^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

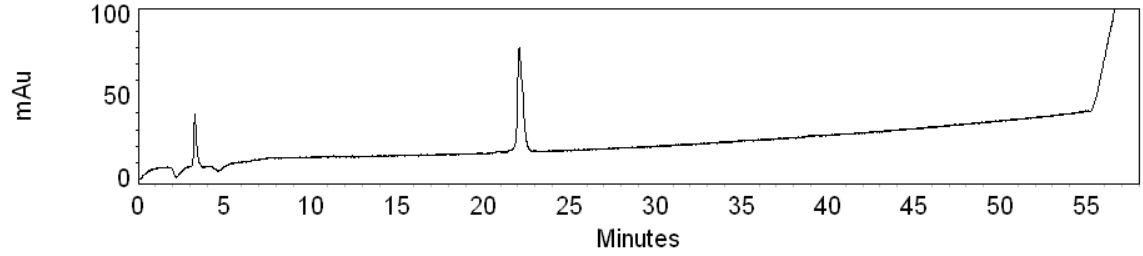


Figure B7.12. Purity check of peptide 3.13, Met-Gln-TS^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

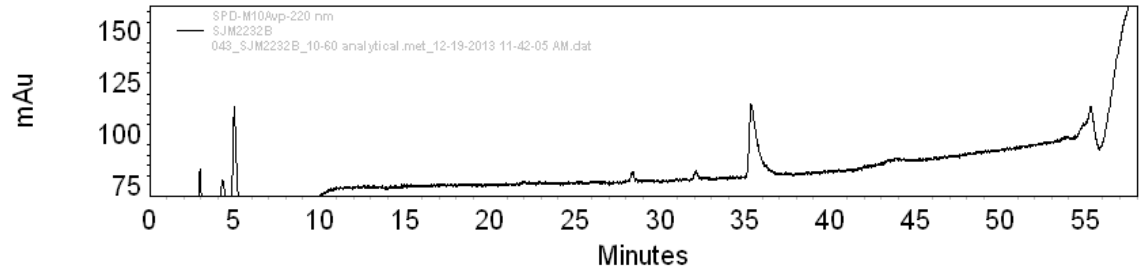


Figure B7.13. Purity check of peptide 3.16, Met-Gln-TS^{-}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2

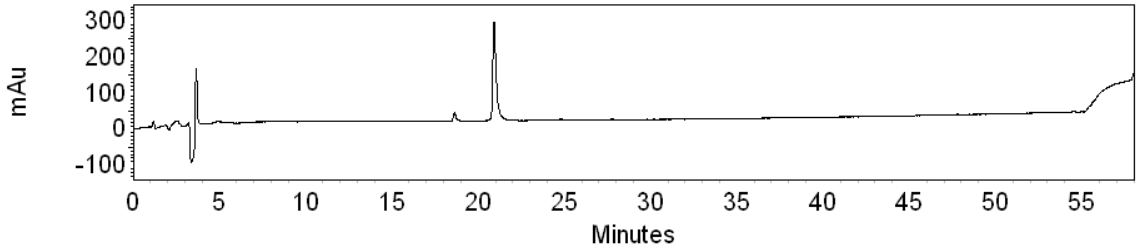
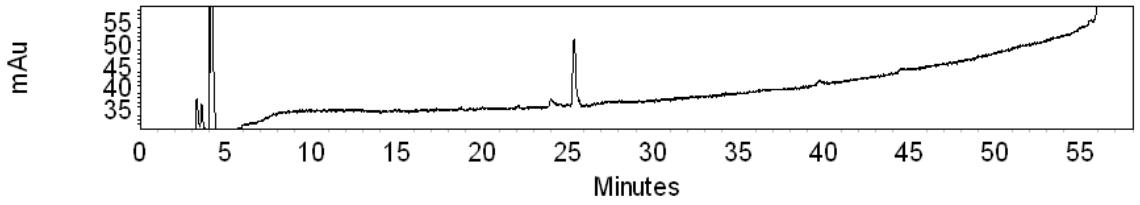
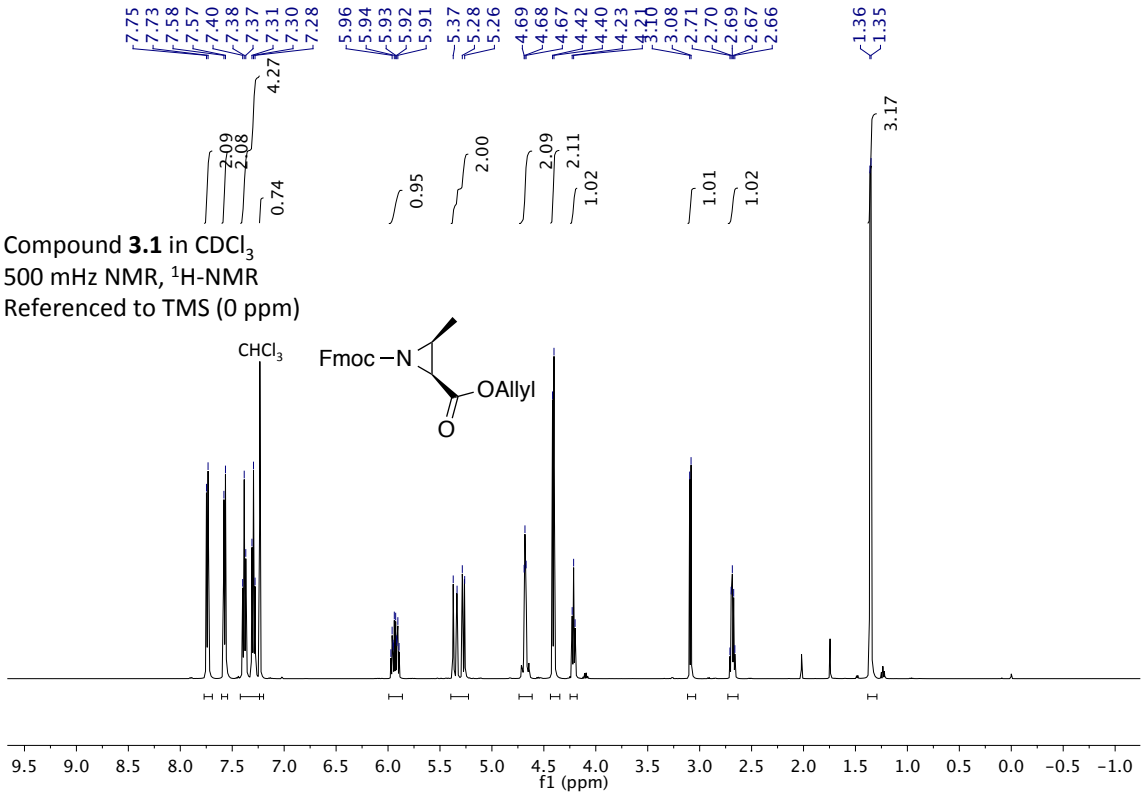


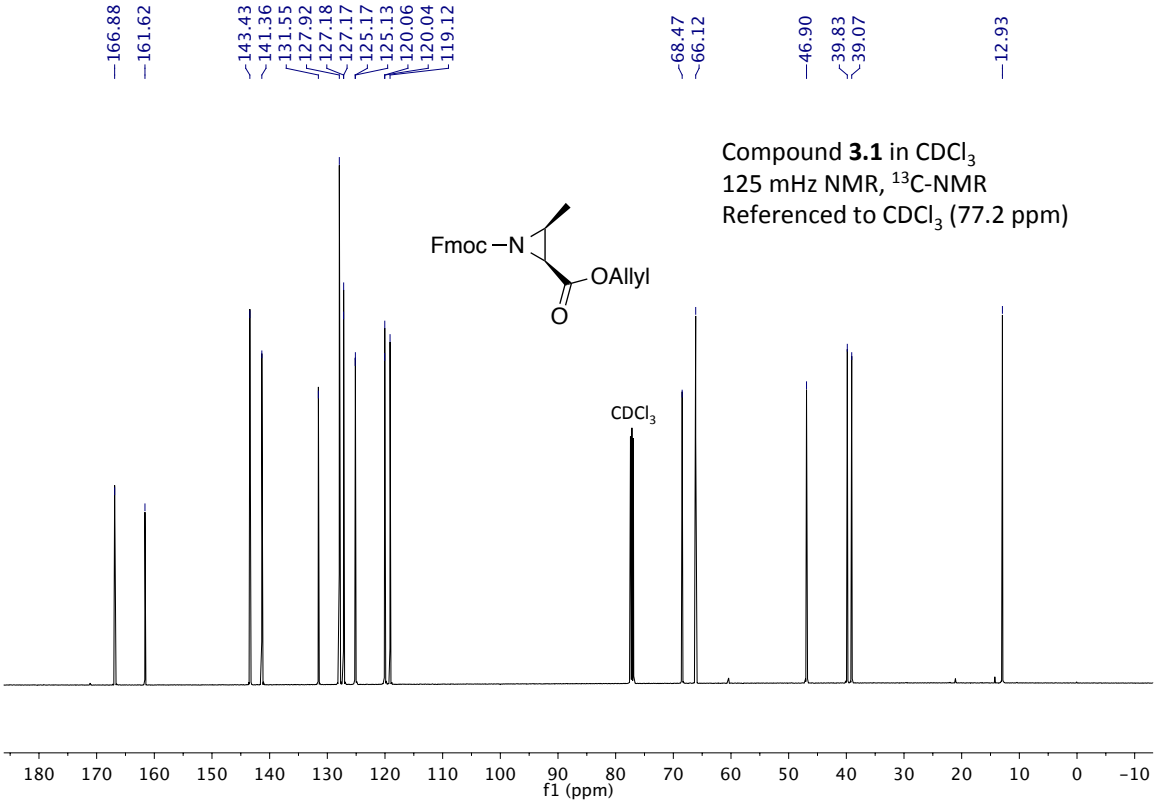
Figure B7.14. Purity check of peptide 3.15, Met-Gln-Glu-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2



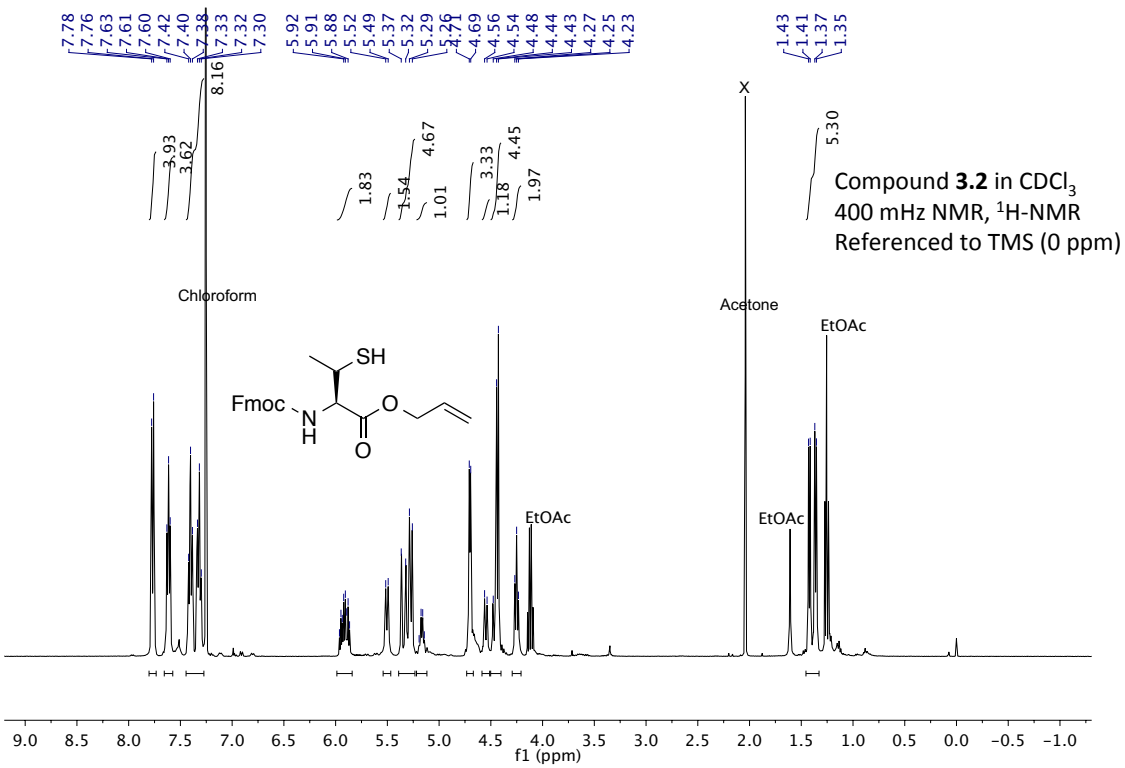
Supplementary Section B8
NMR Spectra



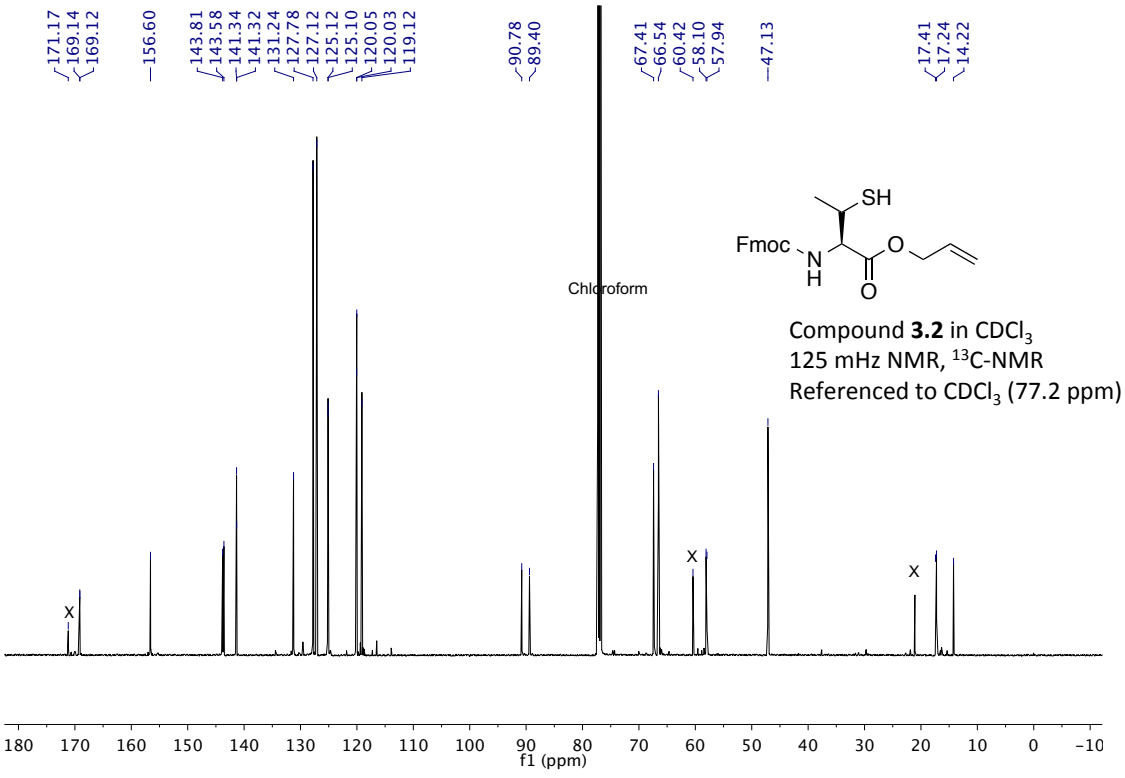
Compound **3.1** in $CDCl_3$
500 mHz NMR, ^{1}H-NMR
Referenced to TMS (0 ppm)



Compound **3.1** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR
Referenced to $CDCl_3$ (77.2 ppm)

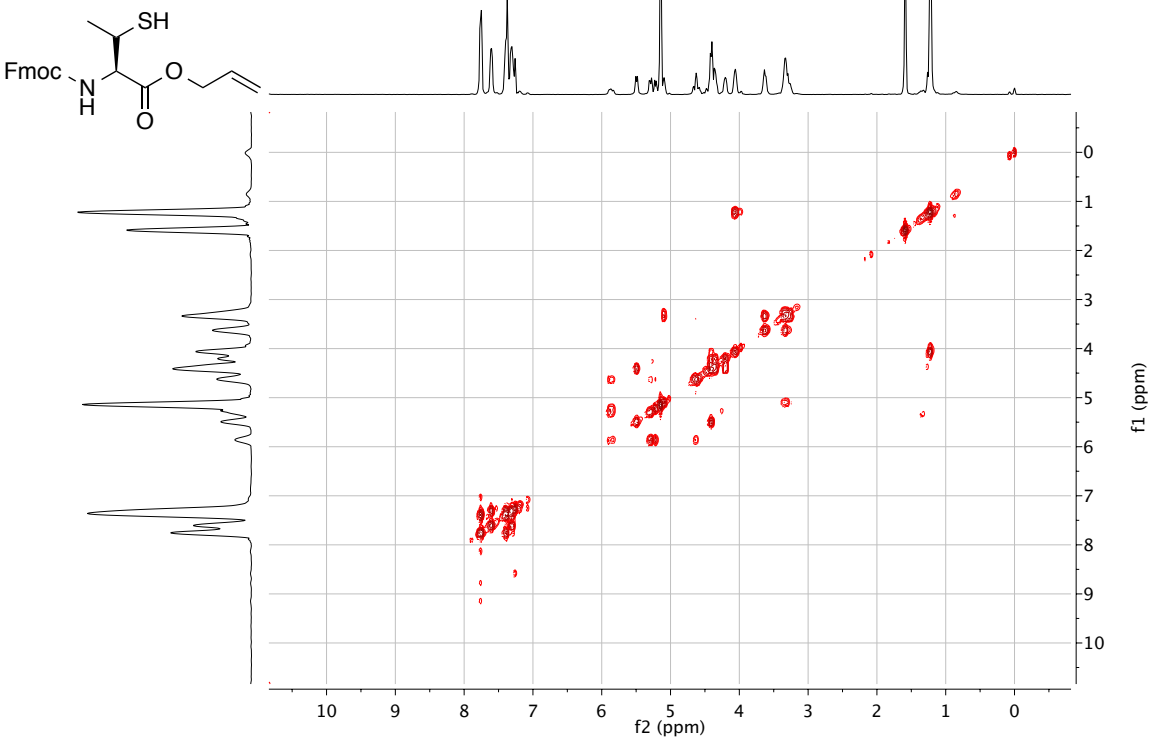


Compound **3.2** in $CDCl_3$
400 mHz NMR, ^{1}H-NMR
Referenced to TMS (0 ppm)

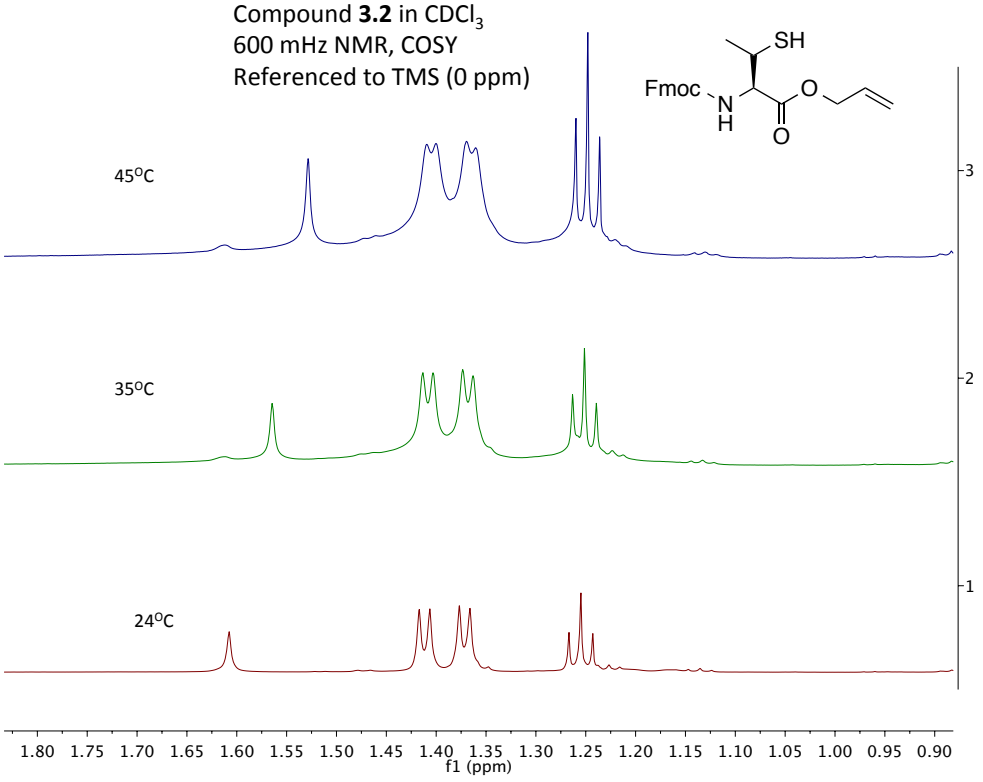


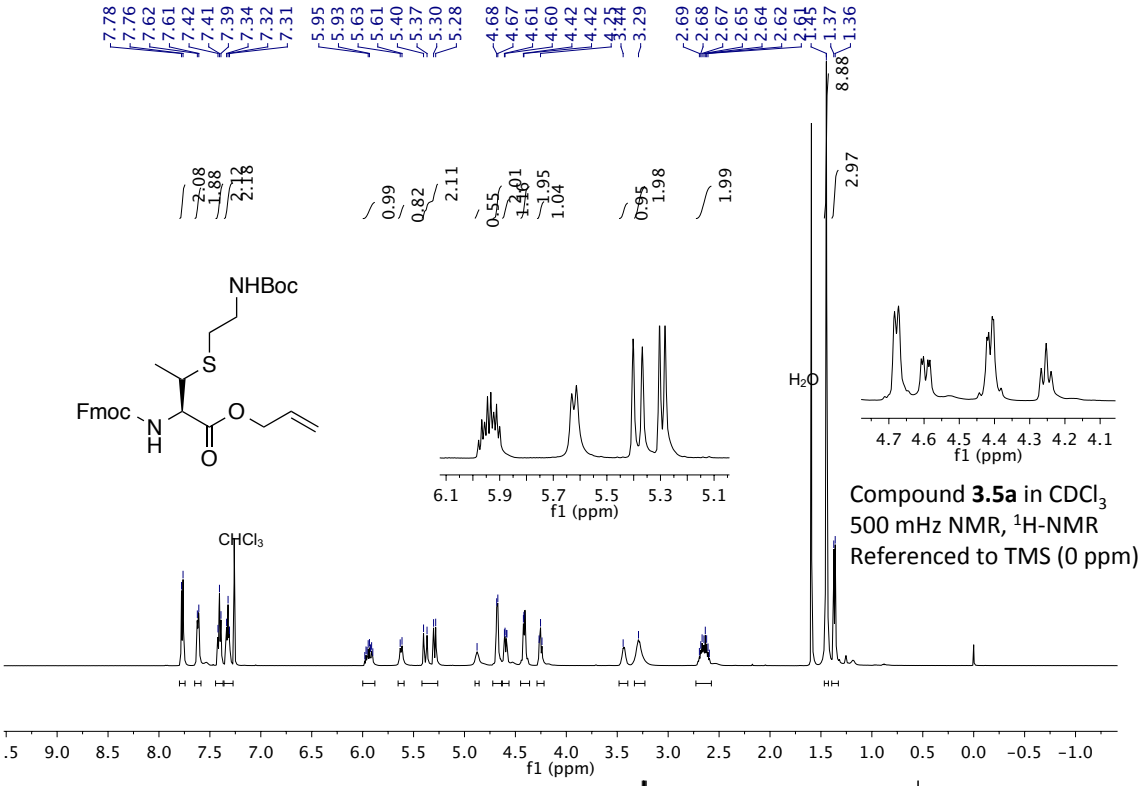
Compound **3.2** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR
Referenced to $CDCl_3$ (77.2 ppm)

Compound **3.2** in $CDCl_3$
600 mHz NMR, COSY
Referenced to TMS (0 ppm)

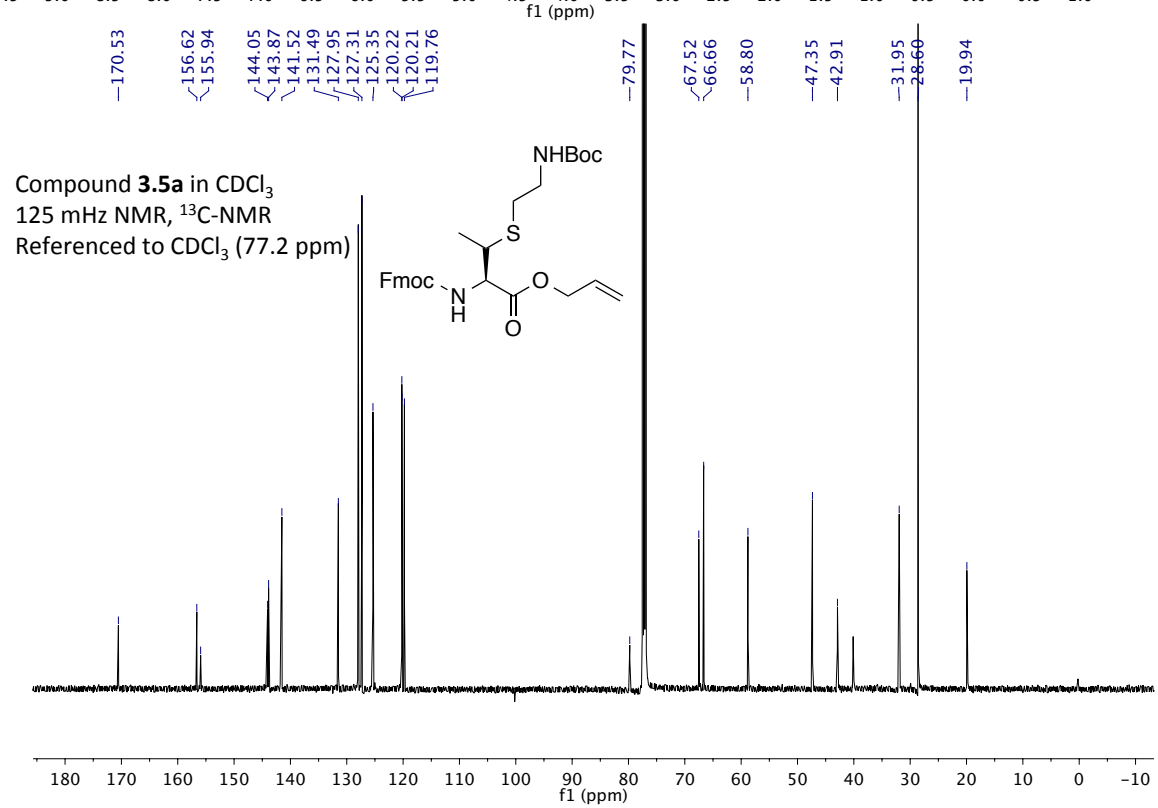


Compound **3.2** in $CDCl_3$
600 mHz NMR, COSY
Referenced to TMS (0 ppm)

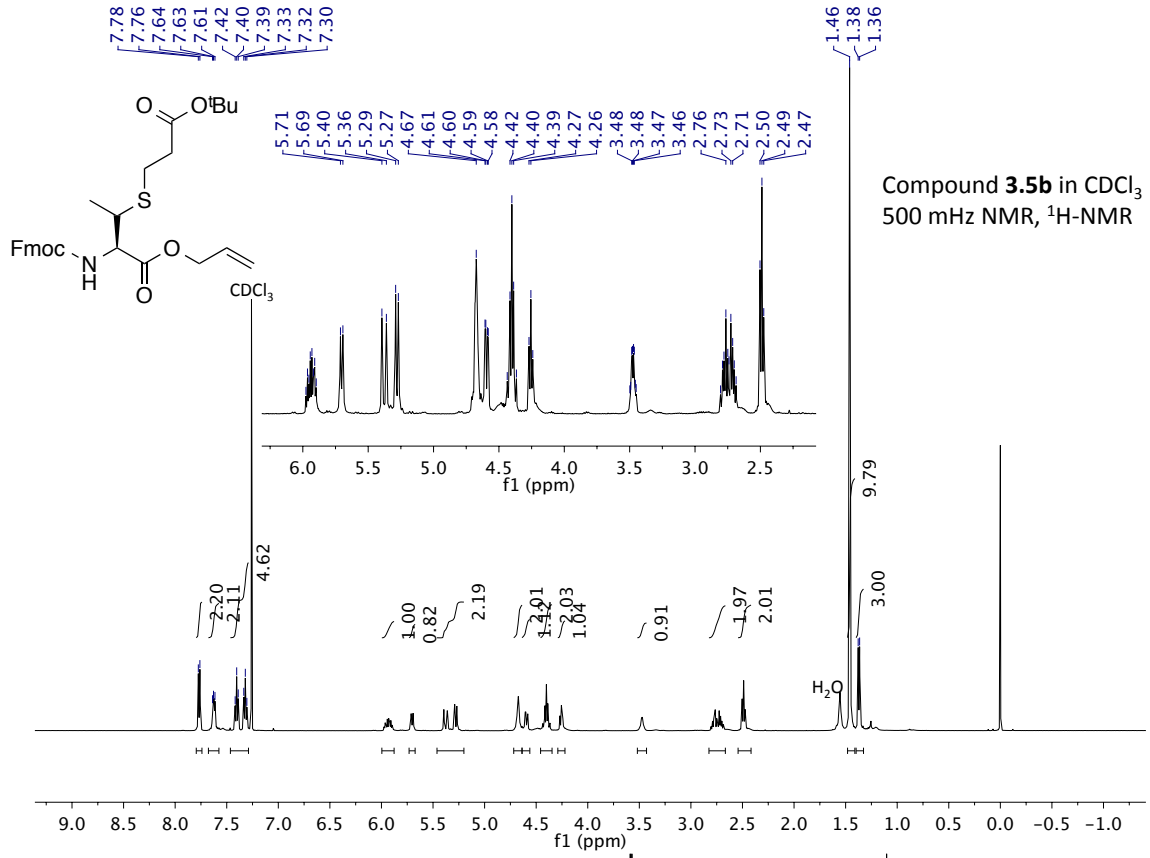


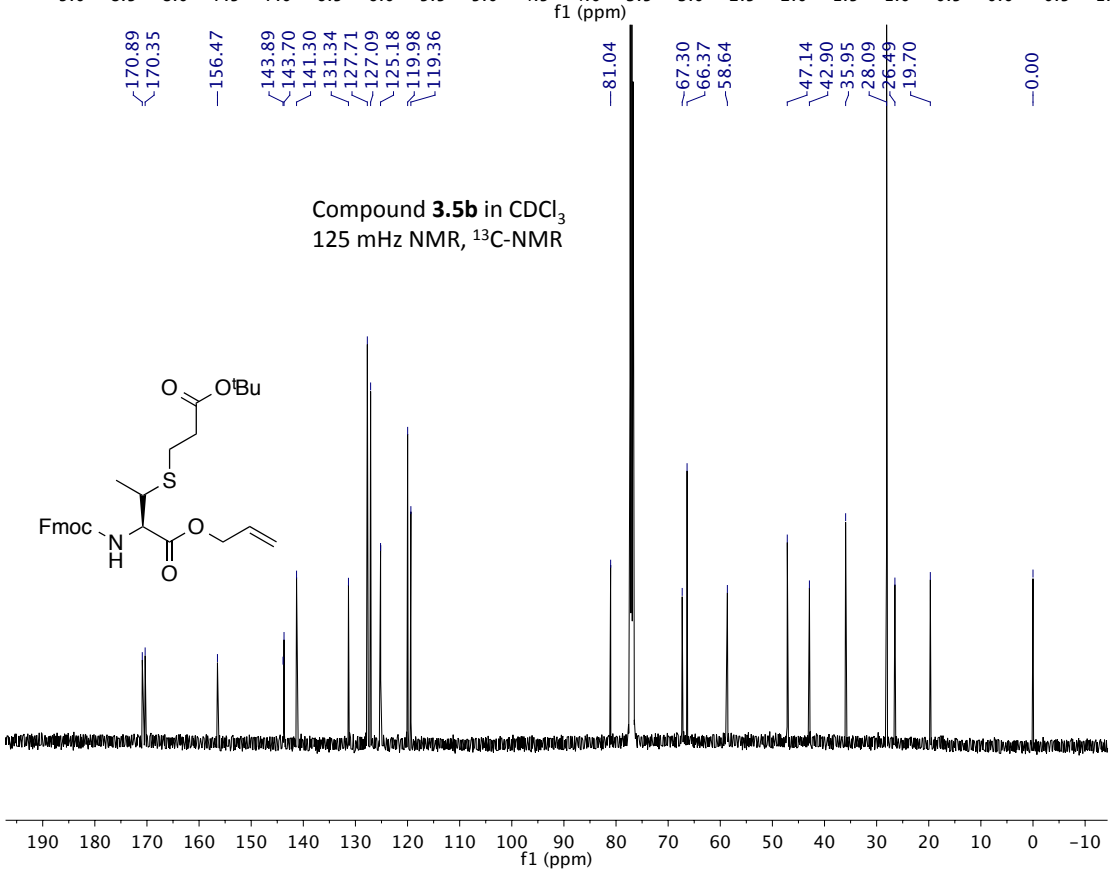


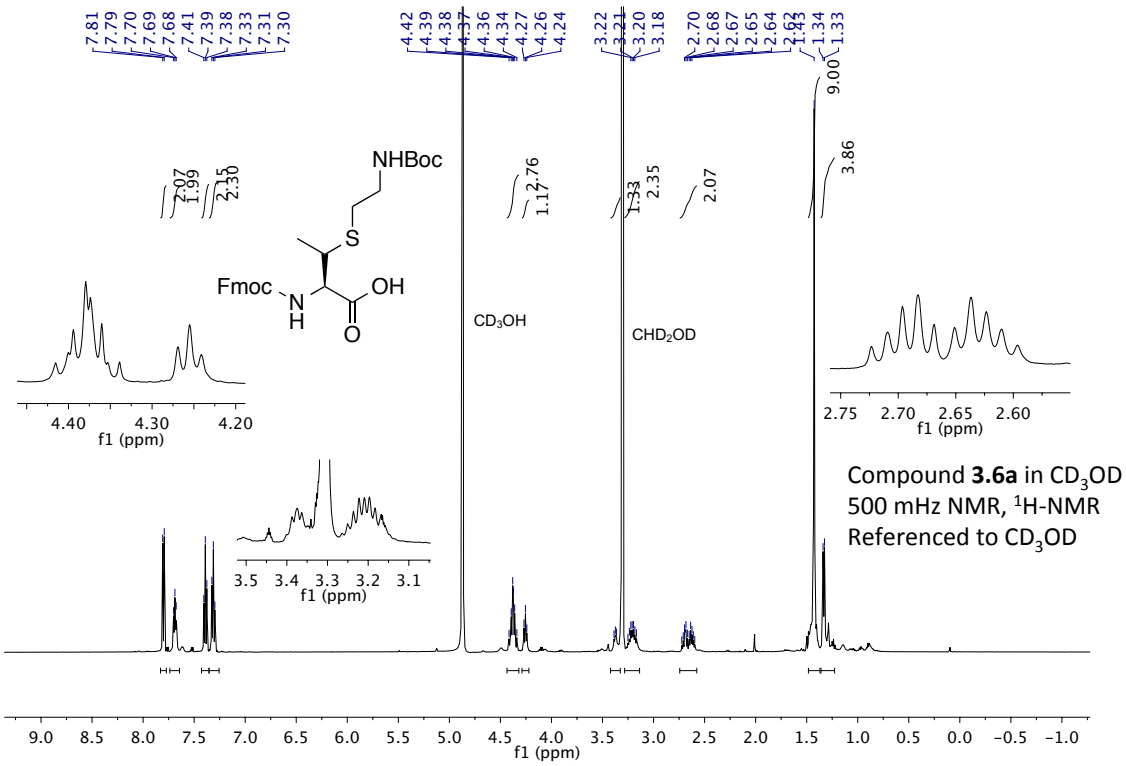
Compound **3.5a** in $CDCl_3$
500 mHz NMR, ^{1}H-NMR
Referenced to TMS (0 ppm)

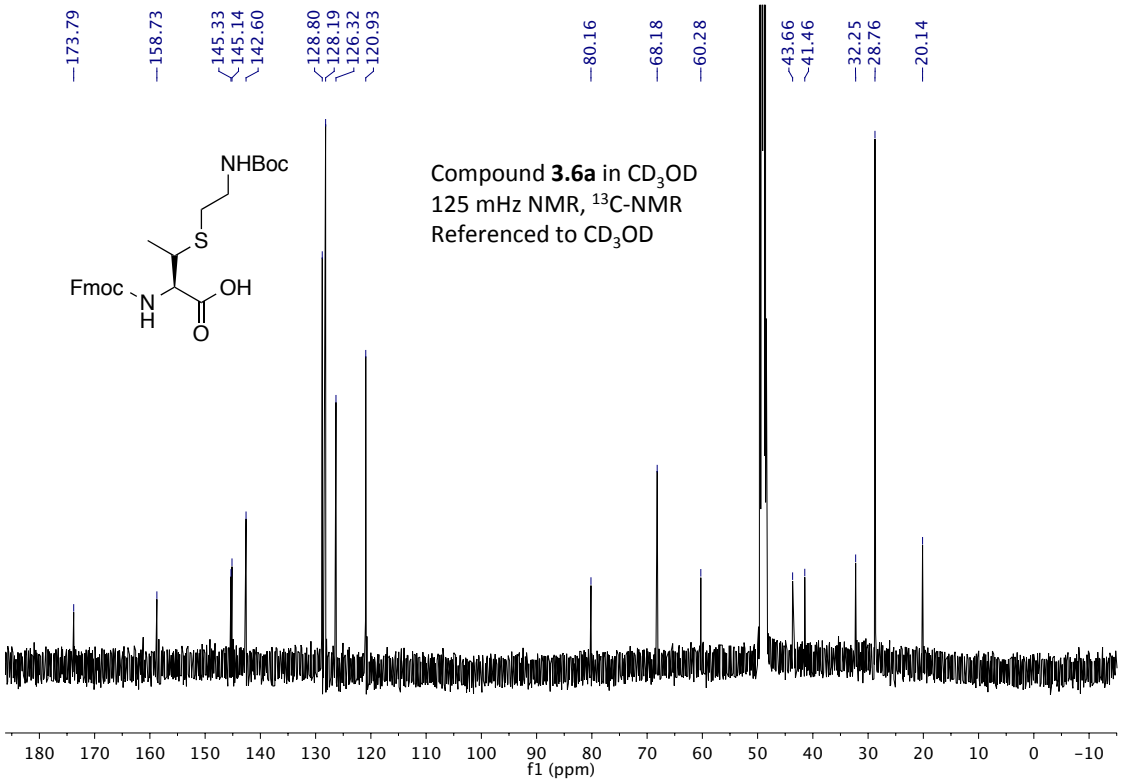


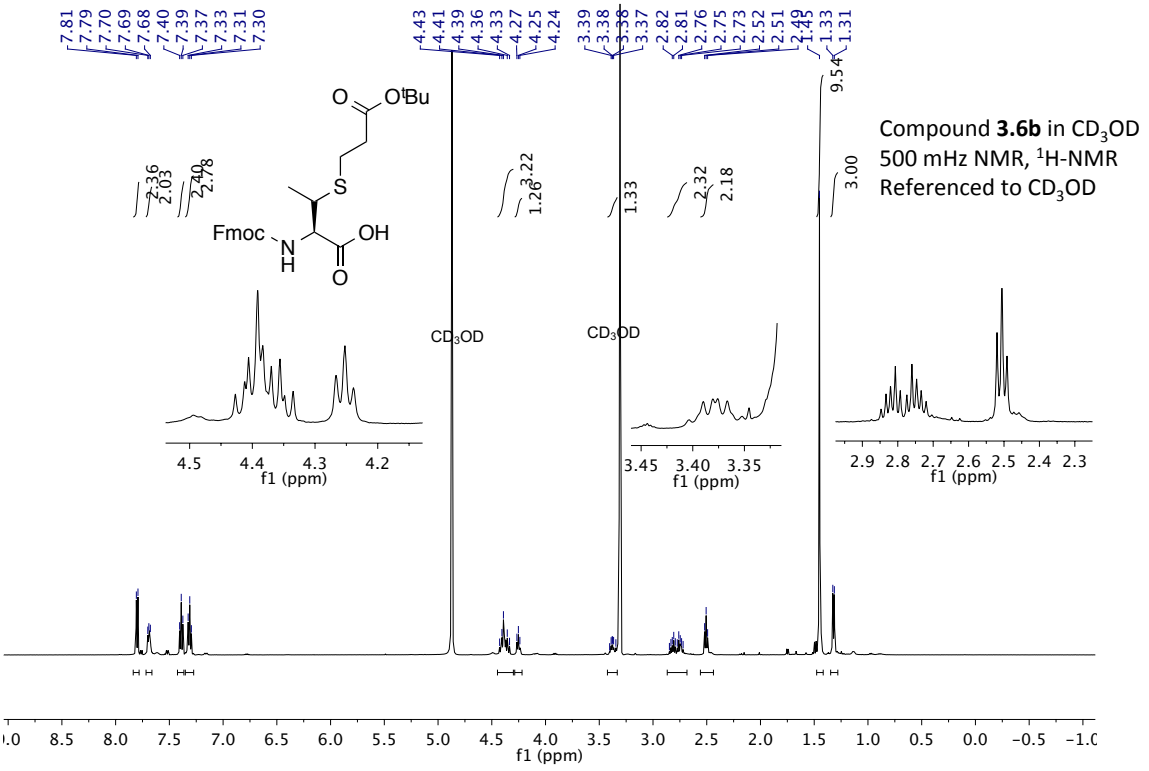
Compound **3.5a** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR
Referenced to $CDCl_3$ (77.2 ppm)

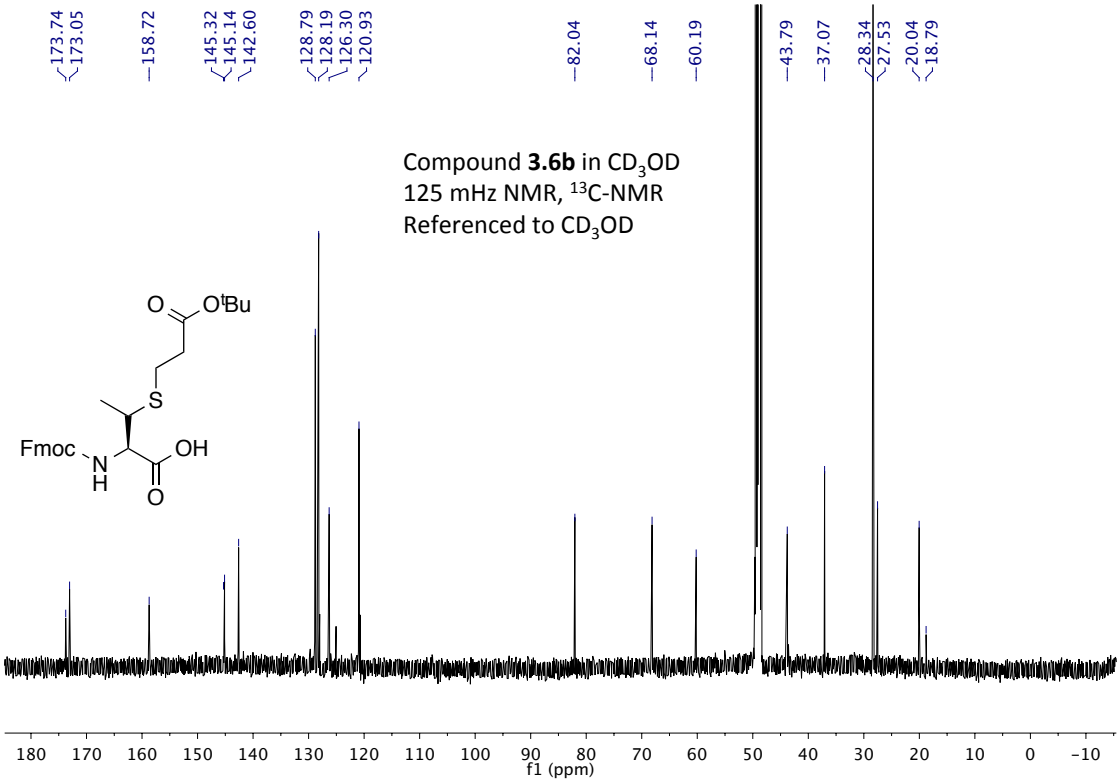

Compound **3.5b** in $CDCl_3$
500 mHz NMR, ^{1}H-NMR
f1 (ppm)

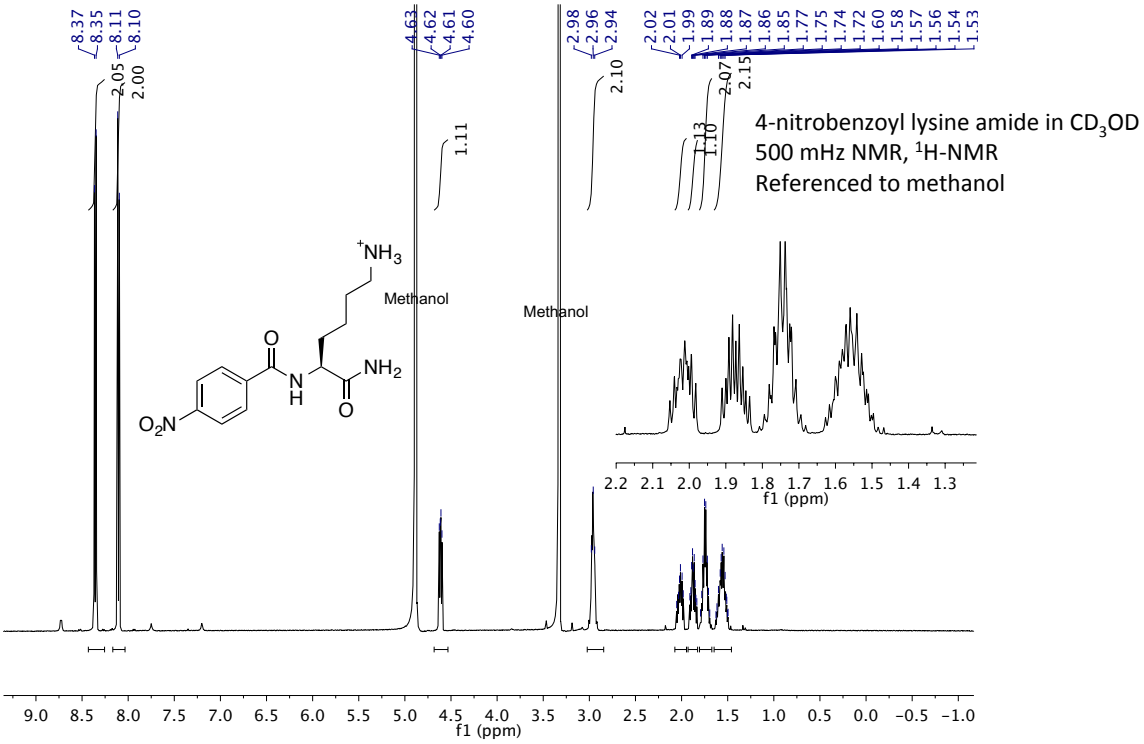

Compound **3.5b** in $CDCl_3$
125 mHz NMR, ^{13}C-NMR
f1 (ppm)

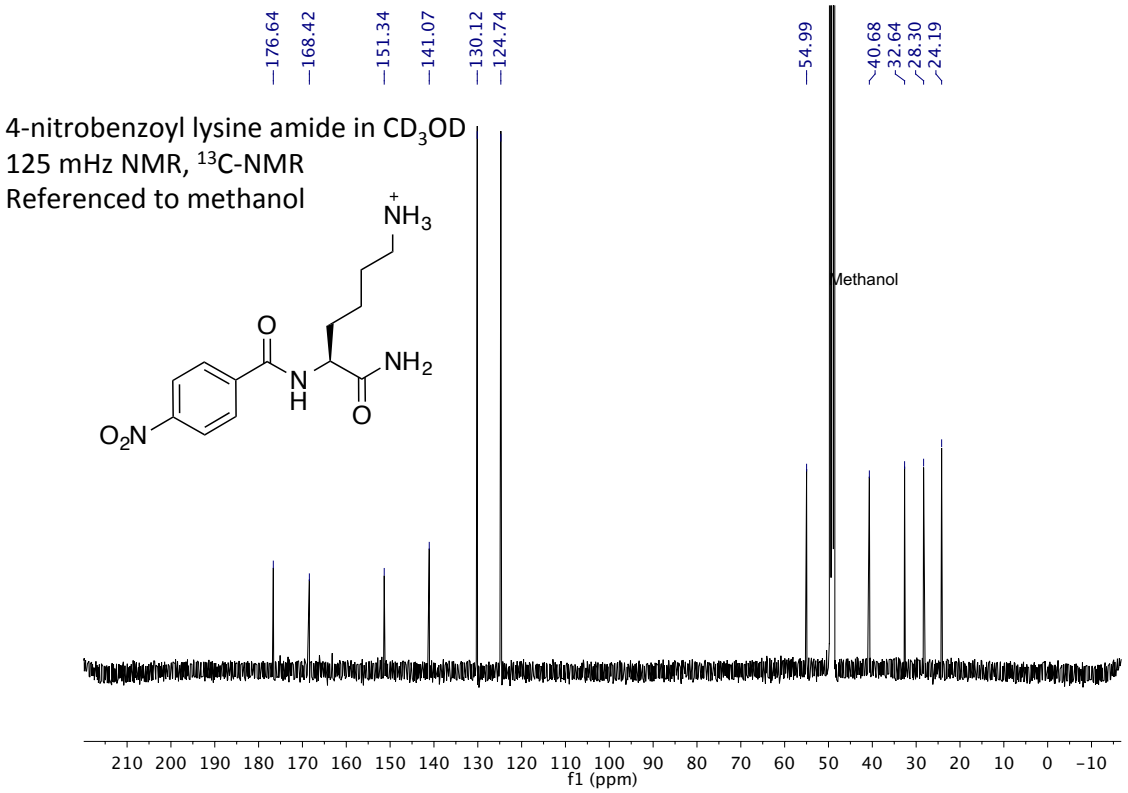

NHBoc
S
Fmoc
N
H
OH
O
CD3OH
CHD2OD
Compound 3.6a in CD3OD
500 mHz NMR, 1H-NMR
Referenced to CD3OD
f1 (ppm)


NHBoc
S
Fmoc
N
H
OH
O
Compound 3.6a in CD3OD
125 mHz NMR, 13C-NMR
Referenced to CD3OD
f1 (ppm)

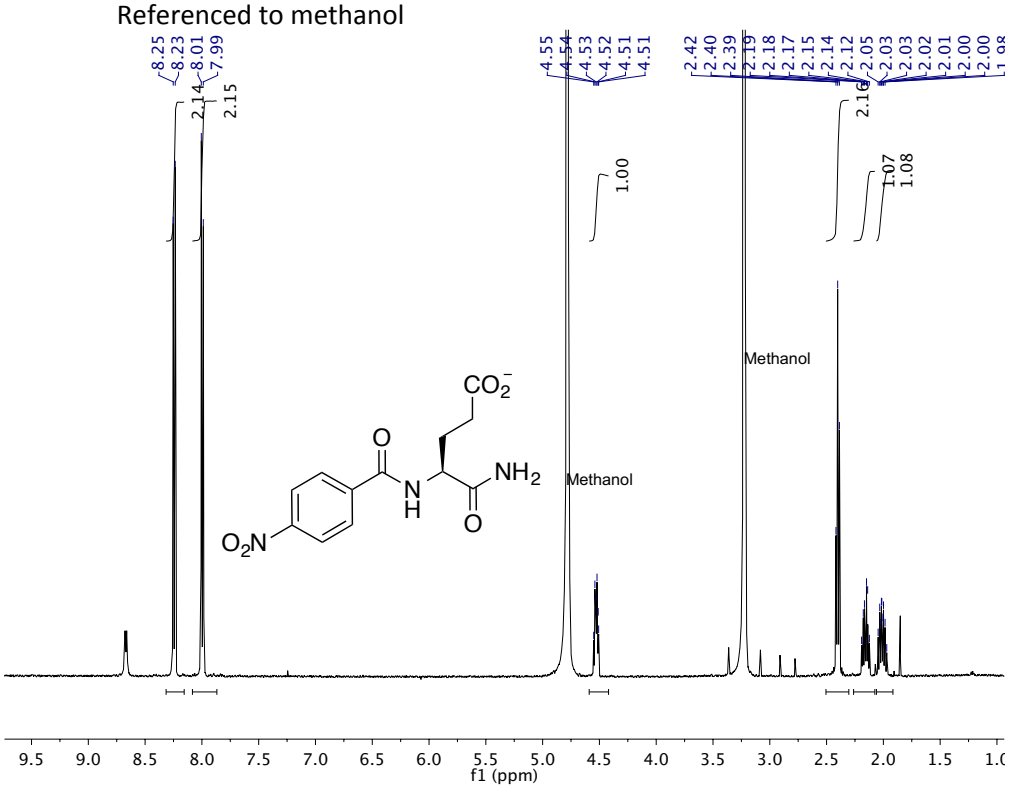

Compound **3.6b** in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to CD_3OD
CD_3OD
CD_3OD
f1 (ppm)

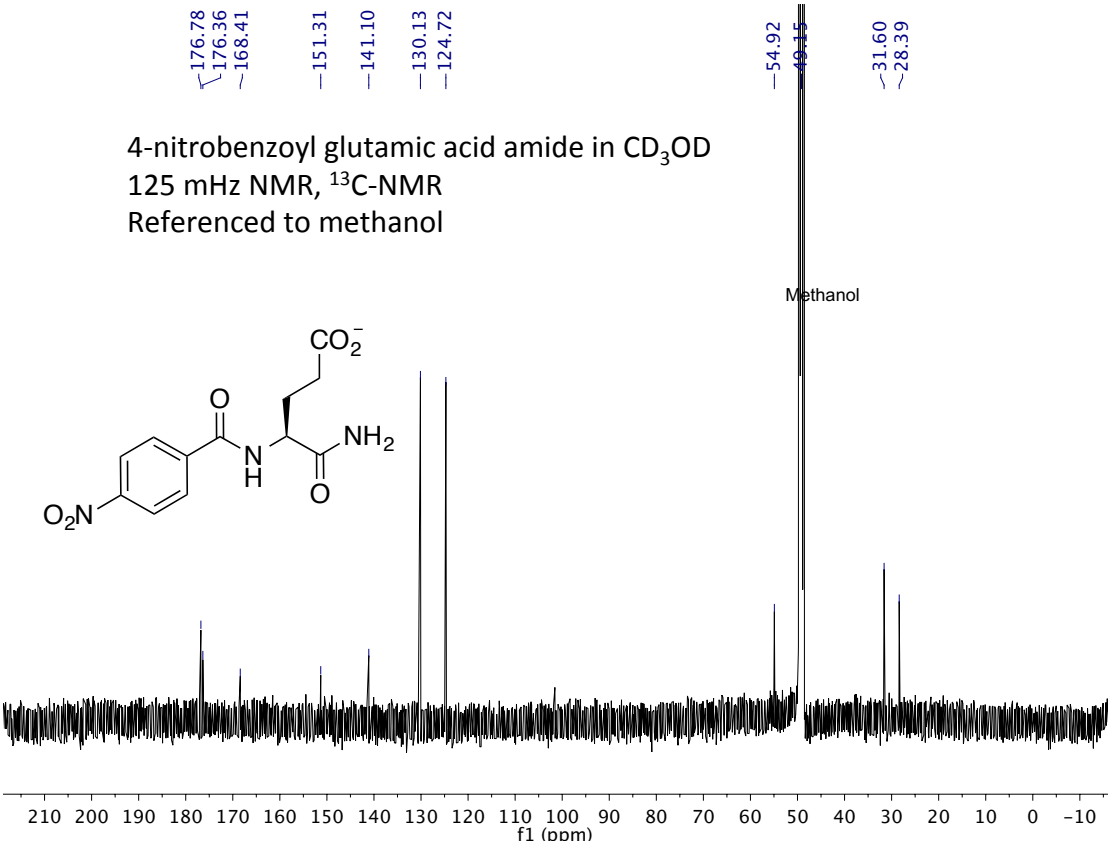

Compound **3.6b** in CD_3OD
125 mHz NMR, ^{13}C-NMR
Referenced to CD_3OD
f1 (ppm)


4-nitrobenzoyl lysine amide in CD3OD
500 mHz NMR, 1H-NMR
Referenced to methanol
Methanol
Methanol
f1 (ppm)
f1 (ppm)

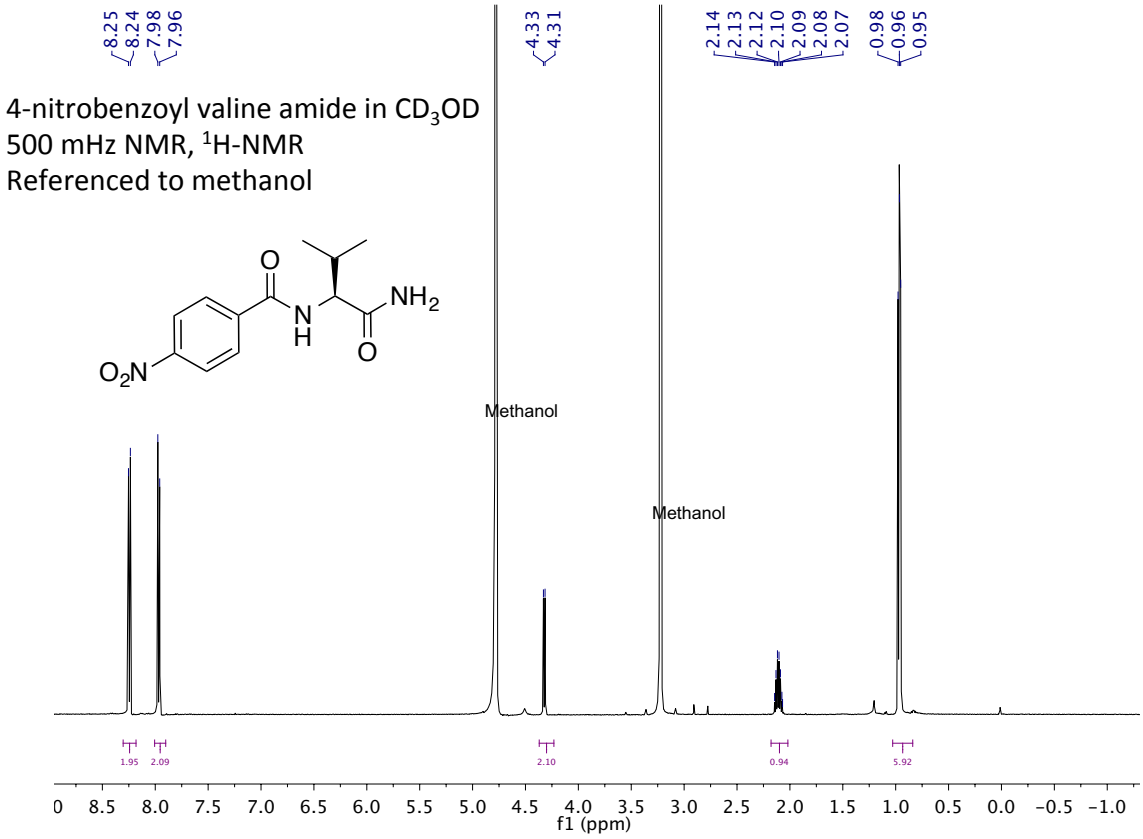

4-nitrobenzoyl lysine amide in CD3OD
125 mHz NMR, 13C-NMR
Referenced to methanol
Methanol
f1 (ppm)

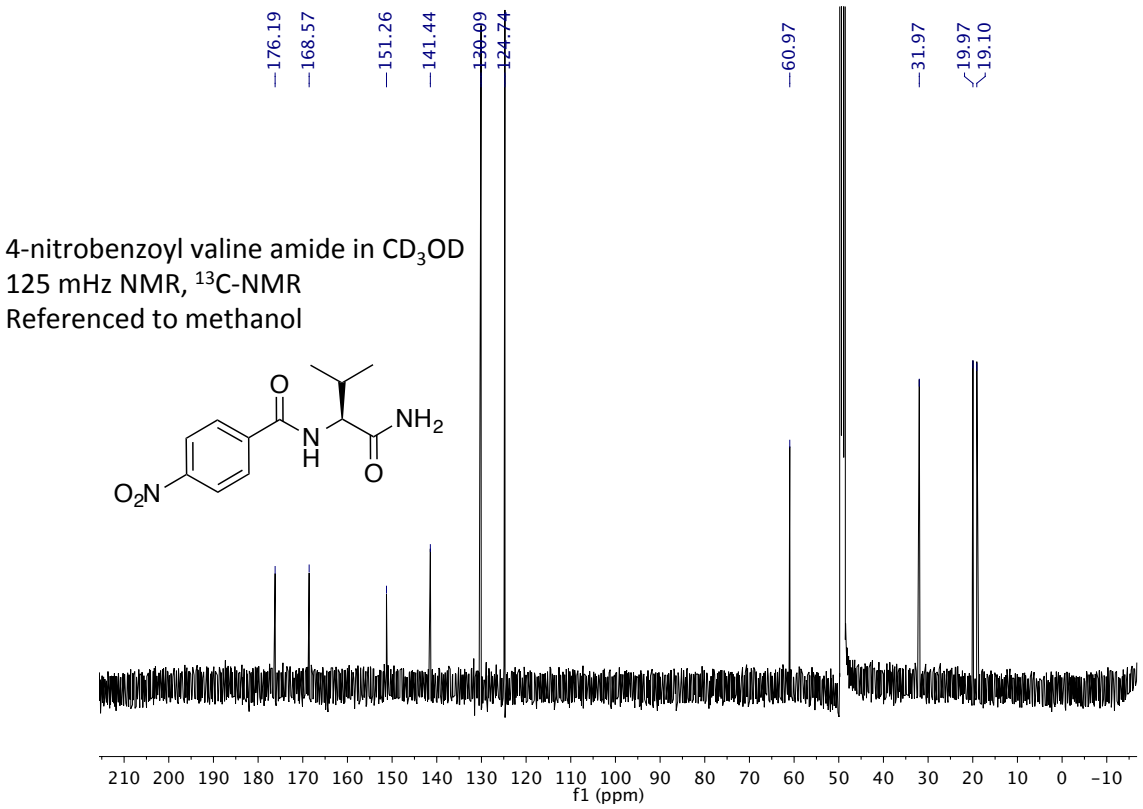
4-nitrobenzoyl glutamic acid amide in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to methanol



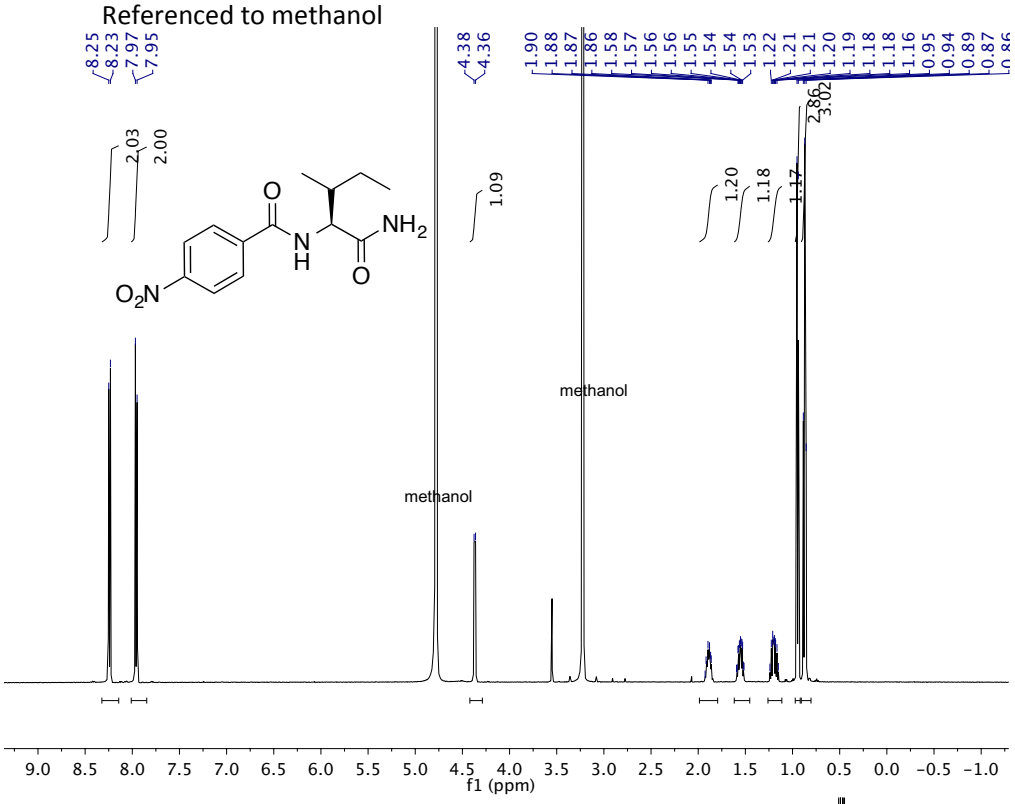


4-nitrobenzoyl glutamic acid amide in CD_3OD
125 mHz NMR, ^{13}C-NMR
Referenced to methanol

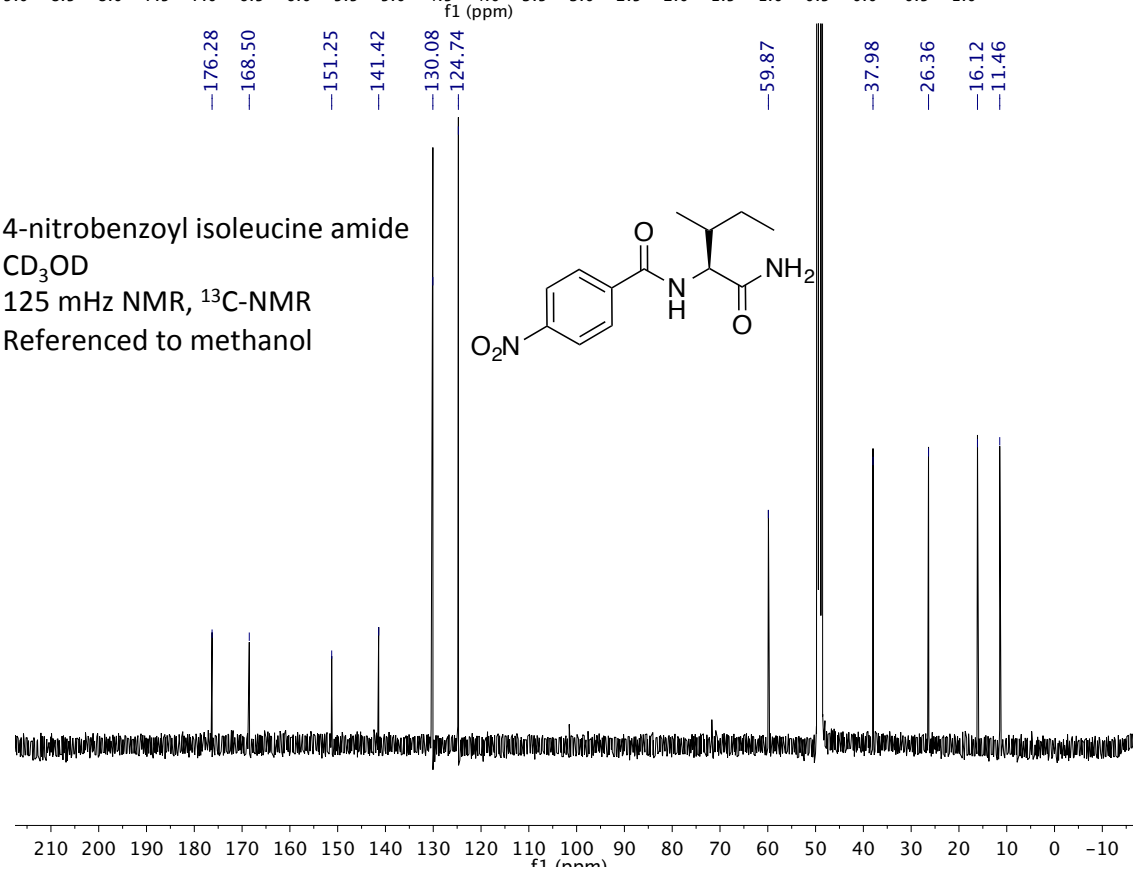

8.25
8.24
7.98
7.96
4.33
4.31
2.14
2.13
2.12
2.10
2.09
2.08
2.07
0.98
0.96
0.95
4-nitrobenzoyl valine amide in CD3OD
500 mHz NMR, 1H-NMR
Referenced to methanol
O
N
H
NH2
O
O2N
Methanol
Methanol
1.95
2.09
2.10
0.94
5.92
0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
f1 (ppm)


176.19
168.57
151.26
141.44
130.09
124.74
60.97
31.97
19.97
19.10
4-nitrobenzoyl valine amide in CD3OD
125 mHz NMR, 13C-NMR
Referenced to methanol
O
N
H
NH2
O
O2N
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

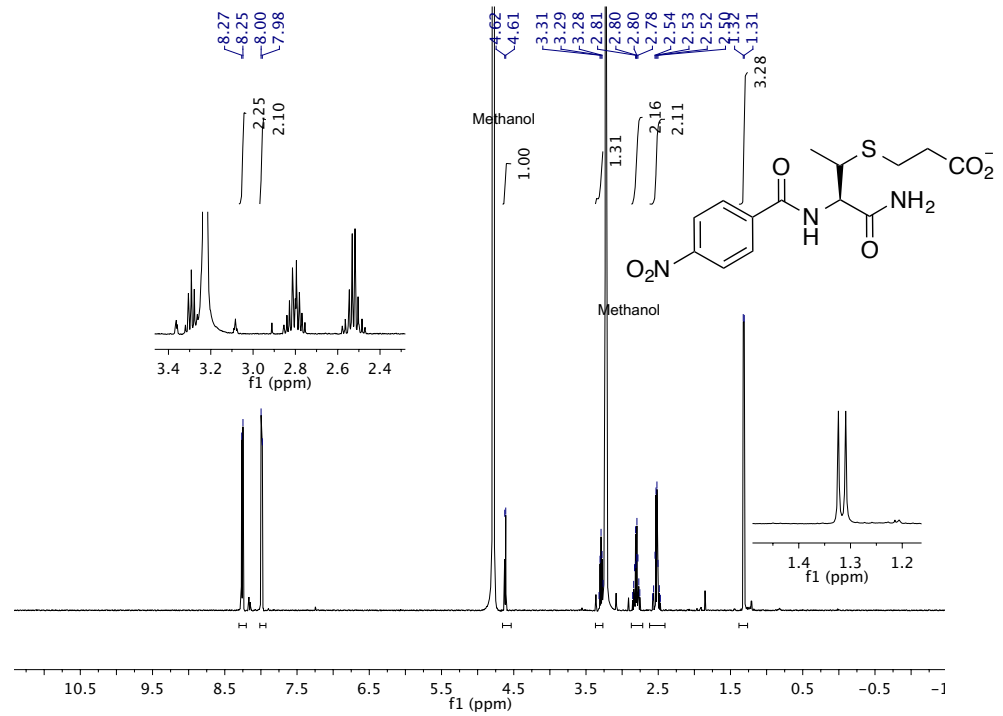
4-nitrobenzoyl isoleucine amide in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to methanol

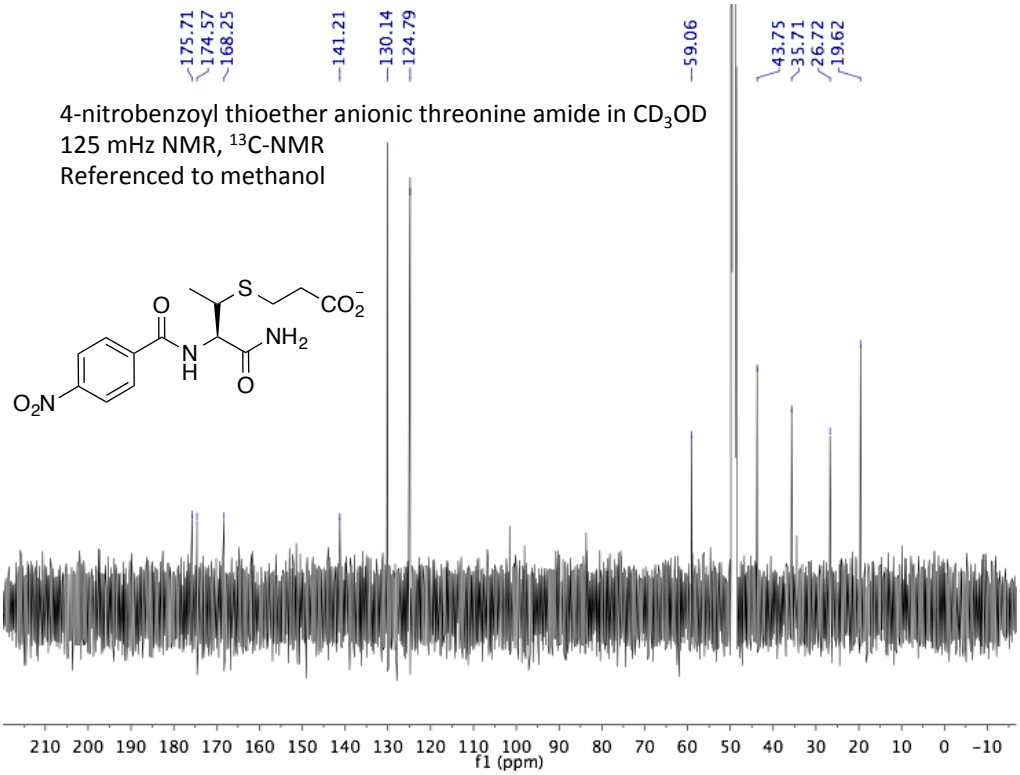


4-nitrobenzoyl isoleucine amide
CD_3OD
125 mHz NMR, ^{13}C-NMR
Referenced to methanol



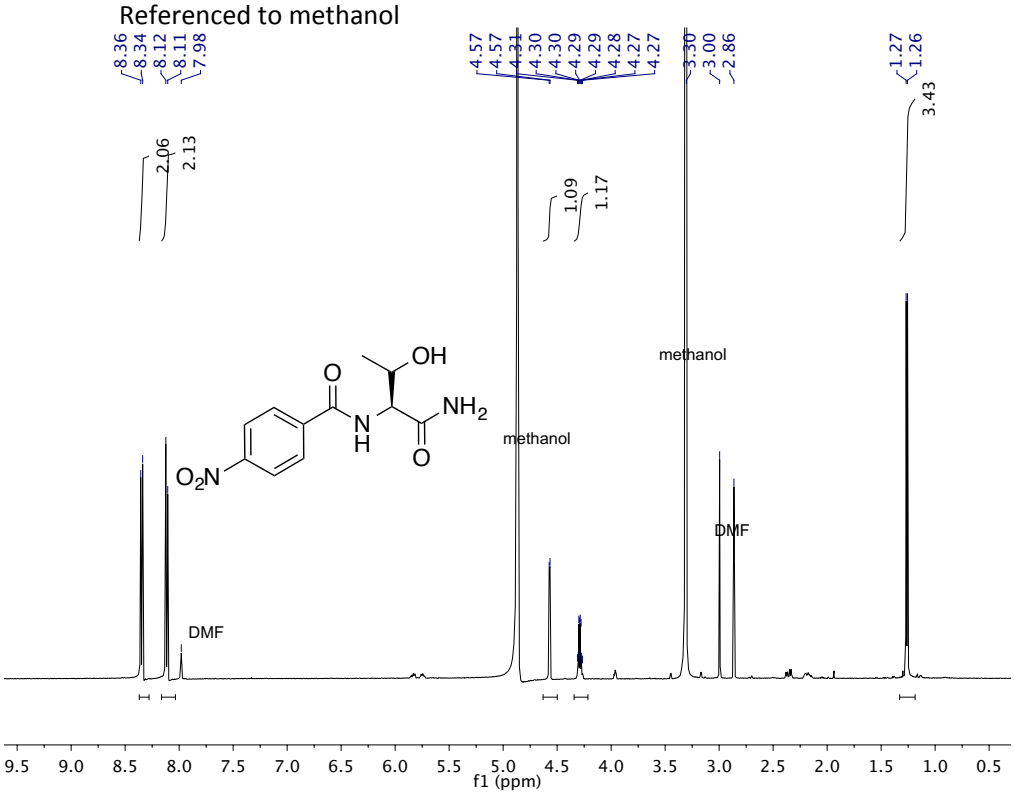
4-nitrobenzoyl thioether anionic threonine amide in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to methanol



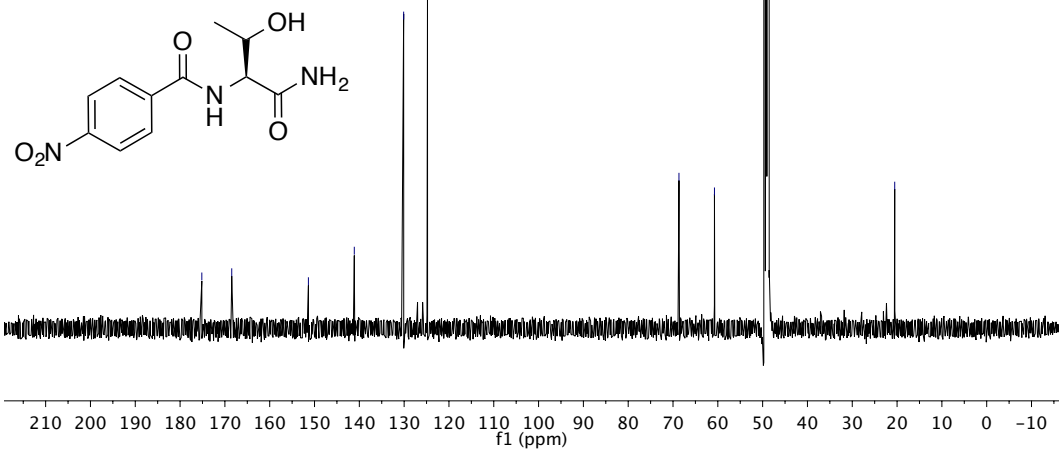


4-nitrobenzoyl thioether anionic threonine amide in CD_3OD
125 mHz NMR, ^{13}C-NMR
Referenced to methanol

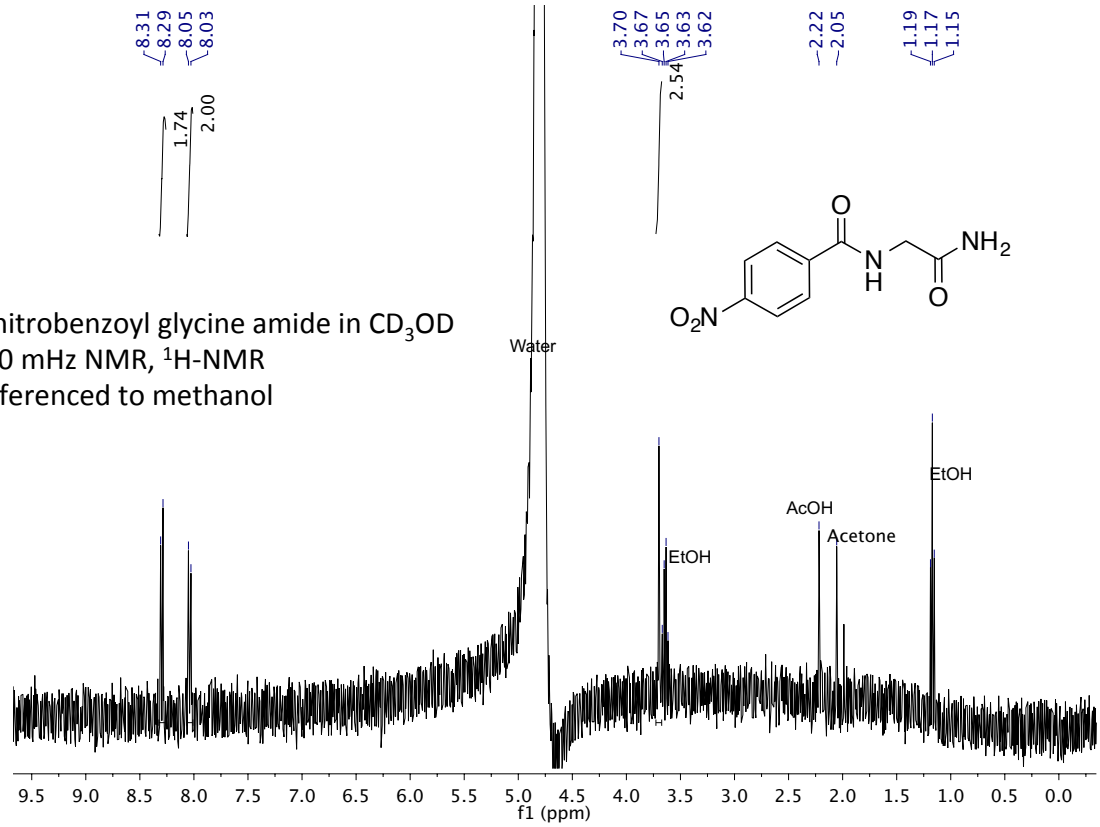
4-nitrobenzoyl threonine amide in CD_3OD
500 mHz NMR, ^1H-NMR
Referenced to methanol



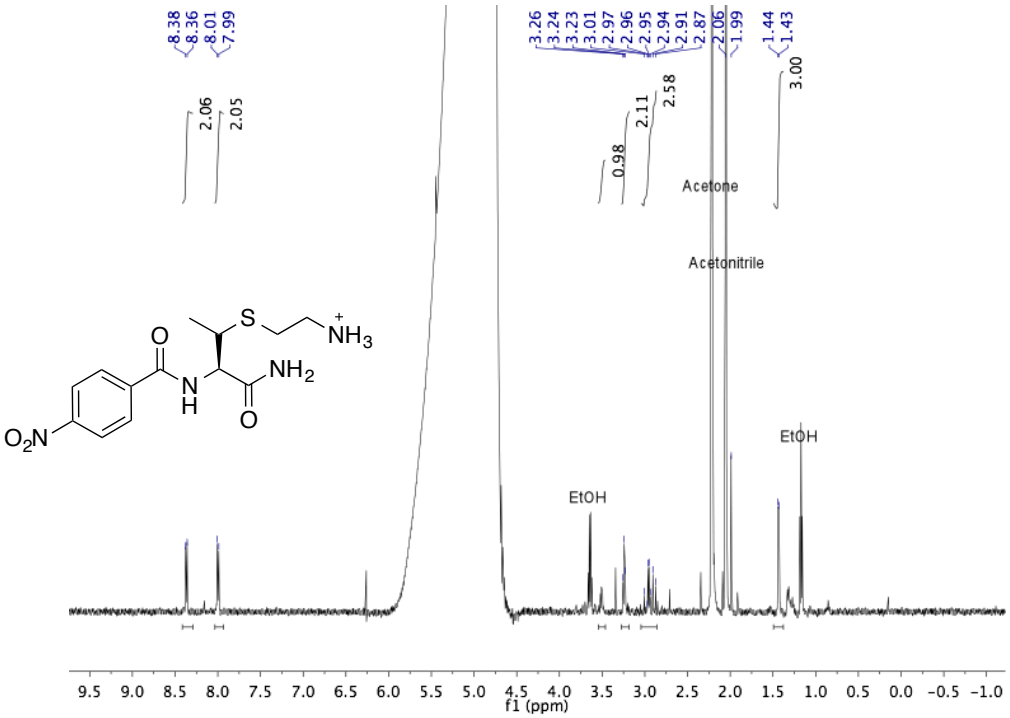
4-nitrobenzoyl threonine amide in CD_3OD
125 mHz NMR, ^{13}C-NMR
Referenced to methanol



4-nitrobenzoyl glycine amide in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to methanol



4-nitrobenzoyl thioether cationic threonine amide in CD_3OD
500 mHz NMR, ^{1}H-NMR
Referenced to methanol



Supplementary Section B9
MALDI-MS

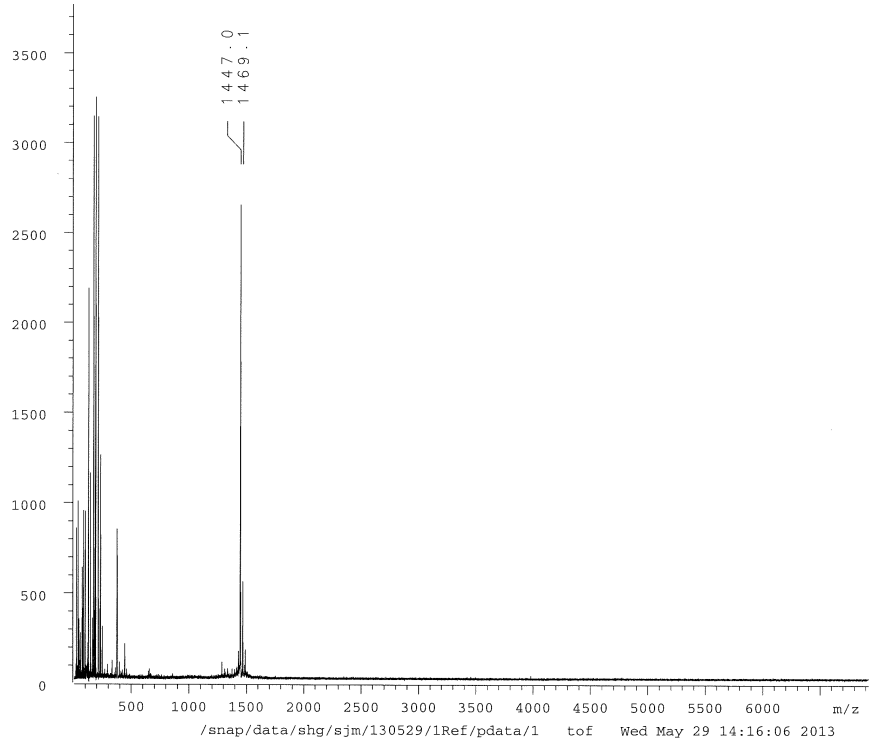


Figure B9.1. MALDI-MS of peptide 3.7a in CHCA matrix. Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2 Expected: 1446.8 m/z (M+H), 1468.9 (M+Na).

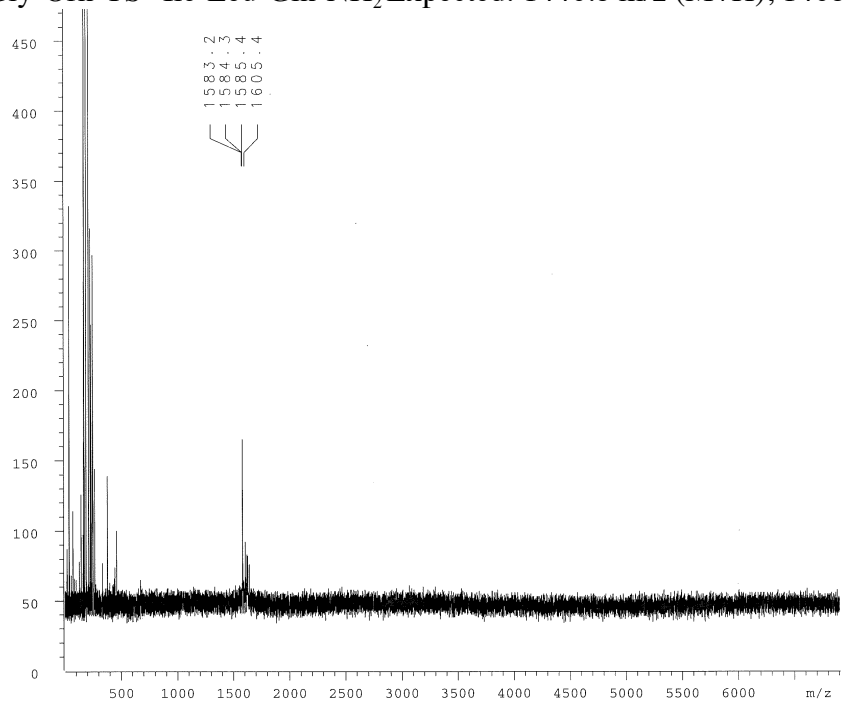


Figure B9.2. MALDI-MS of cyclized peptide 3.7c in CHCA matrix. Cyclized Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-DPro-Gly Expected: 1583.8 m/z (M+H), 1605.8 m/z (M+Na).

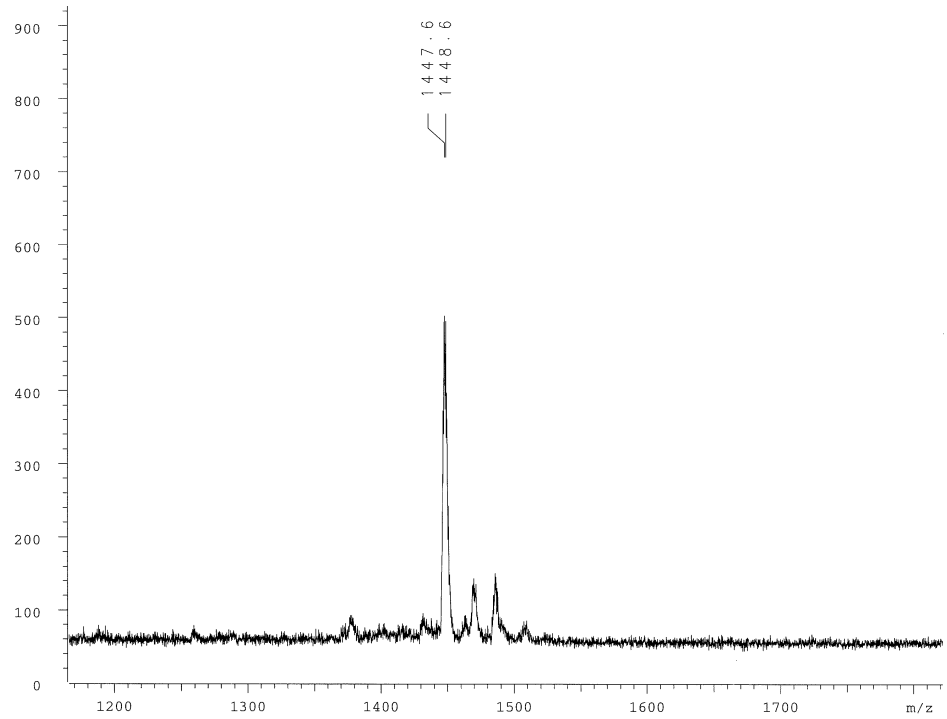


Figure B9.3. MALDI-MS of LPro diastereomer peptide 3.7b in CHCA matrix. Arg-Tyr-Val-Glu-Val-DPro-Gly-Orn-TS^{+}-Ile-Leu-Gln-NH_2 Expected: 1446.8 m/z (M+H), 1468.8 m/z (M+Na).

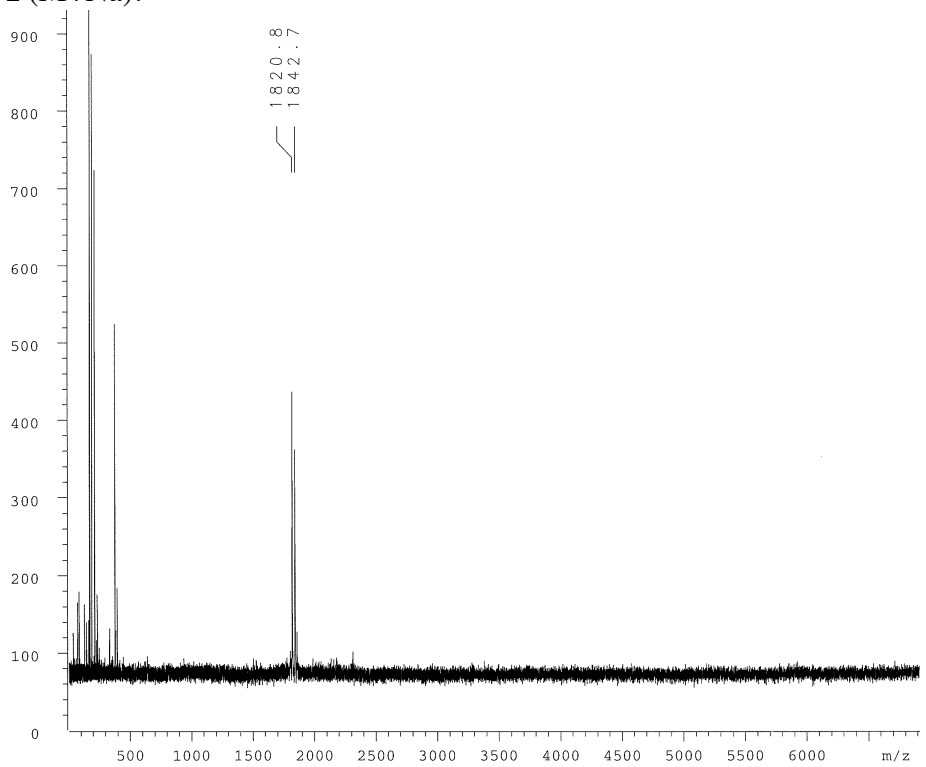


Figure B9.4. MALDI-MS of peptide 3.8 in CHCA matrix. Met-Gln-Ile-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-TS^{+}-Val-NH_2Expected: 1821.0 m/z (M+H), 1843.0 m/z (M+Na).

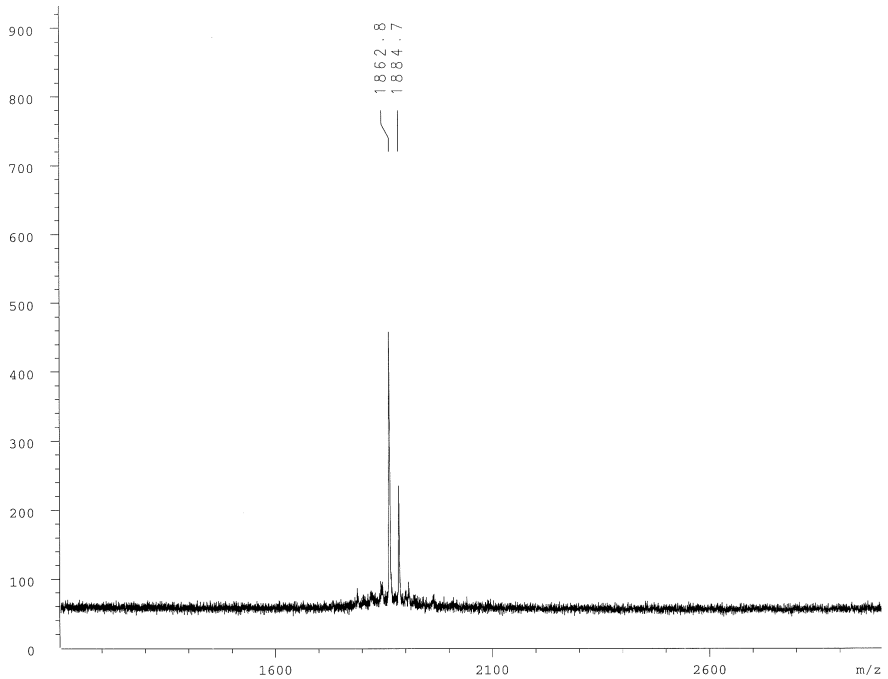


Figure B9.5. MALDI-MS of peptide 3.10 in CHCA matrix. Thr-TO^{+}-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2 Expected: 1861.1 m/z (M+H), 1884.1 m/z (M+Na).

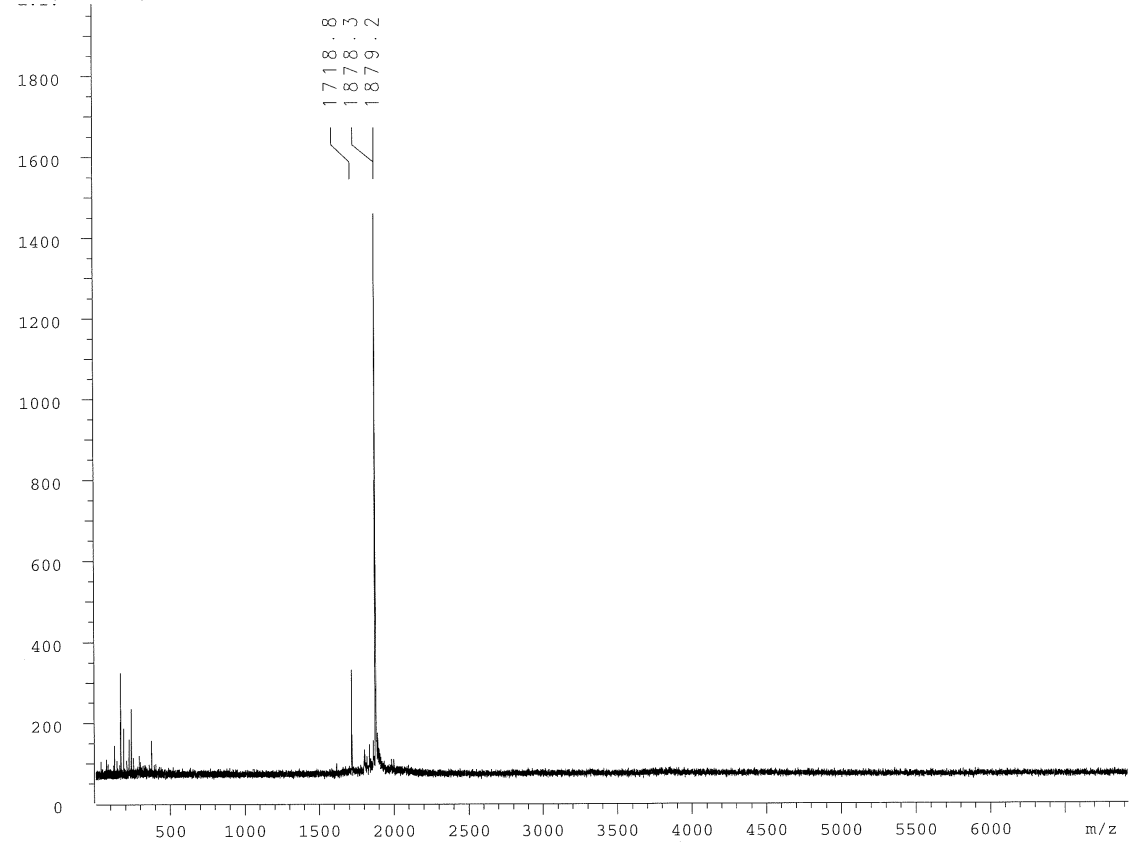


Figure B9.6. MALDI-MS of peptide 3.11 in CHCA matrix. Thr-TS^{+}-Arg-Tyr-Val-Glu-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-Thr-Thr-NH_2 Expected: 1878.1 m/z (M+H).

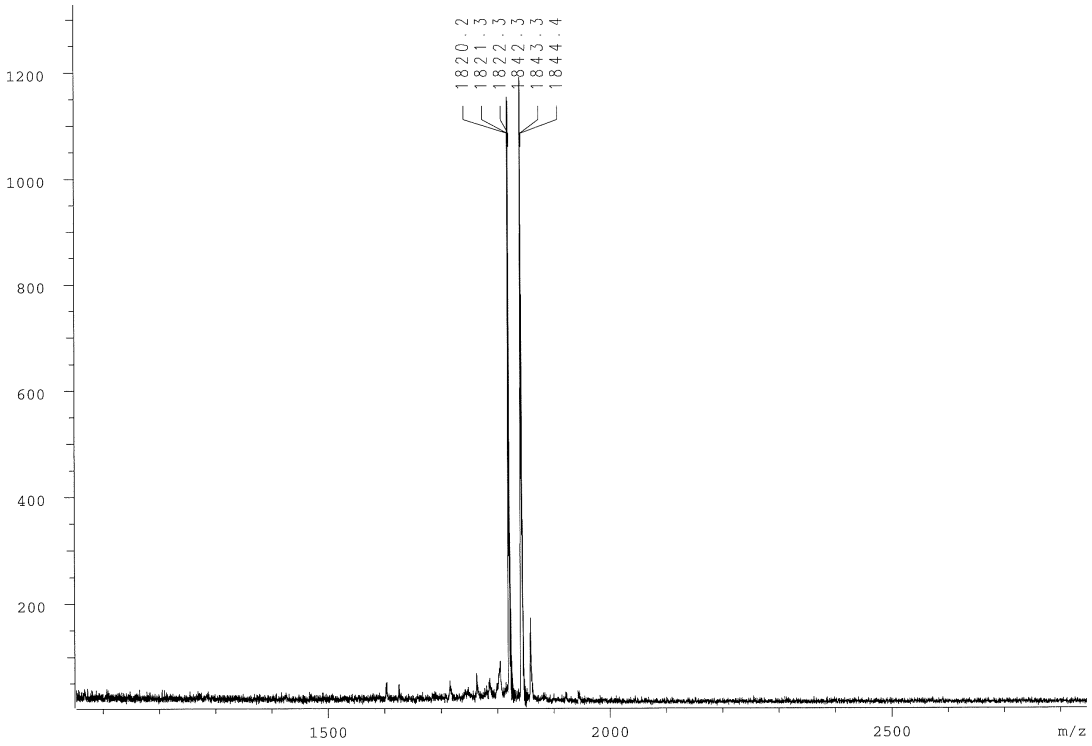


Figure B9.7. MALDI-MS of peptide 3.12 in CHCA matrix. Met-Gln-TO^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2 Expected: 1821.0 m/z (M+H), 1843.0 m/z (M+Na).

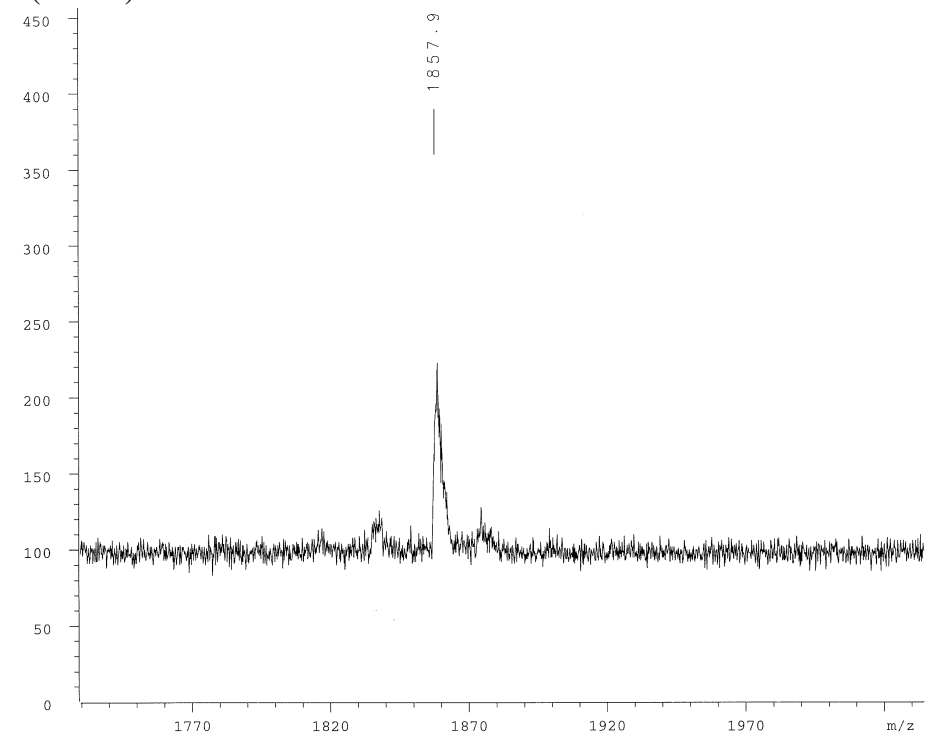


Figure B9.8. MALDI-MS of peptide 3.13 in CHCA matrix. Met-Gln-TS^{+}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2 Expected: 1836.1 m/z (M+H), 1858.1 m/z (M+Na).

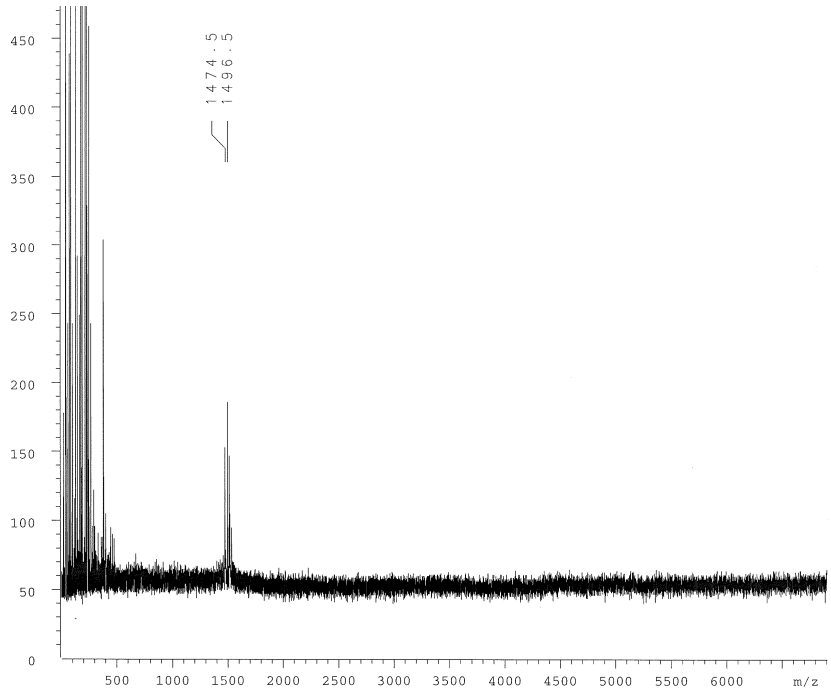


Figure B9.9. MALDI-MS of peptide 3.14a in CHCA matrix. Arg-Tyr-Val-TS^{-}-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2 Expected: 1474.8 m/z (M+H), 1496.8 m/z (M+Na).

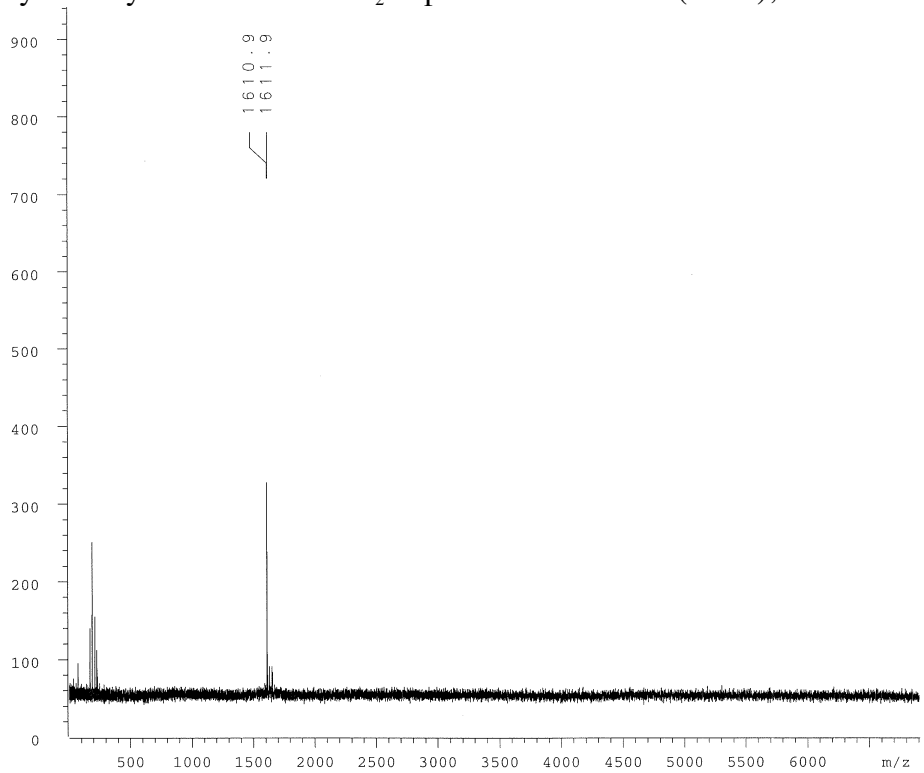


Figure B9.10. MALDI-MS of the cyclic peptide 3.14c in CHCA matrix. Cyclized Arg-Tyr-Val-TS^{-}-Val-DPro-Gly-Orn-Lys-Ile-Leu-Gln-DPro-Gly Expected: 1610.9 m/z (M+H).

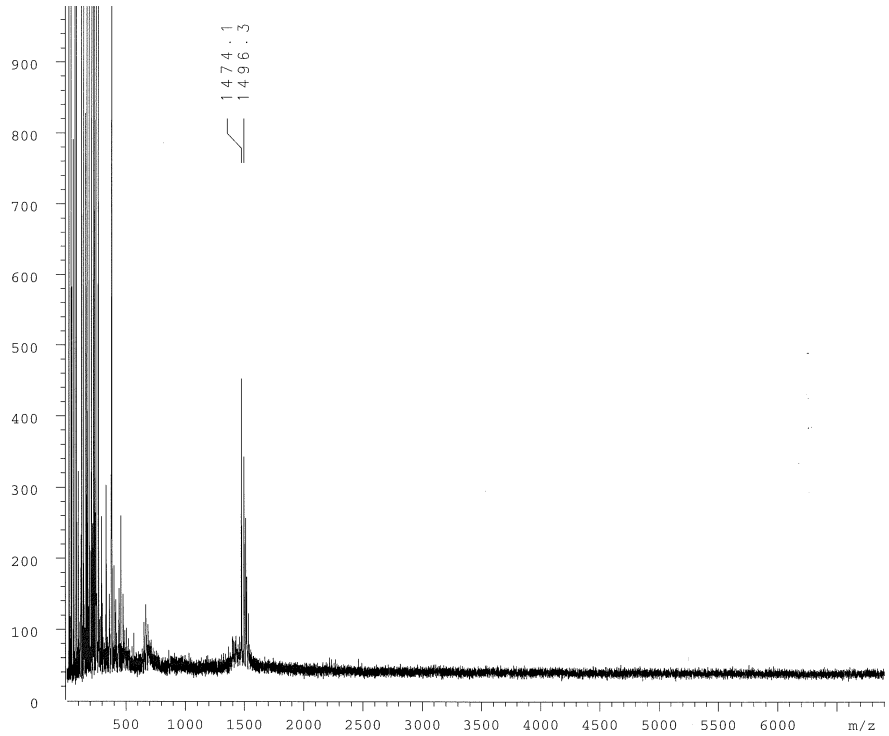


Figure B9.11. MALDI-MS of the LPro diasteromer peptide 3.14b in CHCA matrix. Arg-Tyr-Val-TS^{-}-Val-LPro-Gly-Orn-Lys-Ile-Leu-Gln-NH_2 Expected: 1474.8 m/z (M+H), 1496.8 m/z (M+Na).

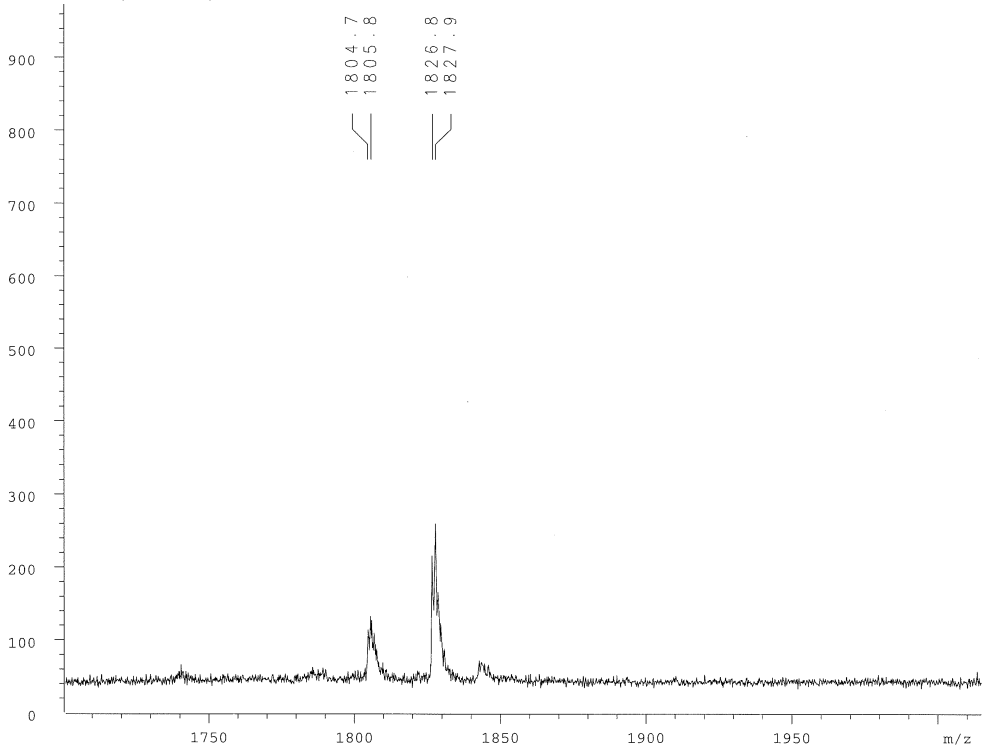


Figure B9.12. MALDI-MS of peptide 3.15 in CHCA matrix. Met-Gln-Glu-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2 Expected: 1805.0 m/z (M+H), 1827.0 m/z (M+Na).

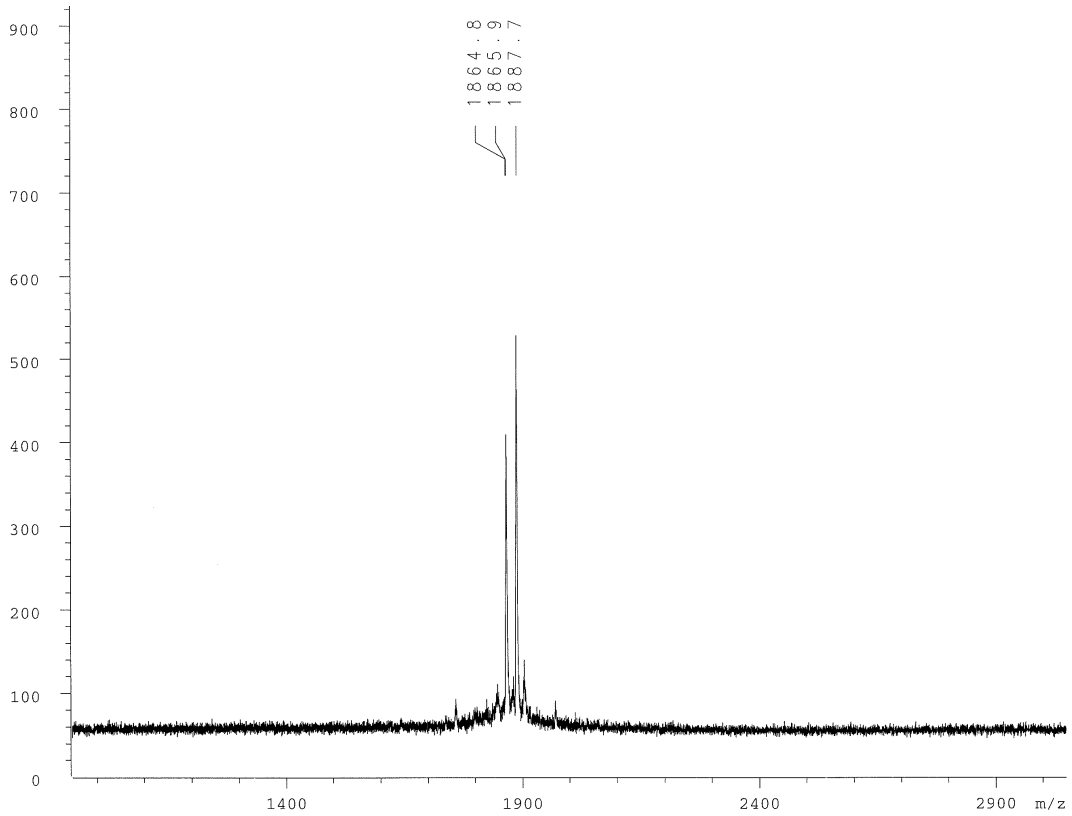


Figure B9.13. MALDI-MS of peptide 3.16 in CHCA matrix. Met-Gln-TS^{-}-Phe-Val-Lys-Ser-DPro-Gly-Lys-Thr-Ile-Thr-Leu-Lys-Val-NH_2 Expected: 1865.0 m/z (M+H), 1887.0 m/z (M+Na).

Supplementary Section B10
NMR Structure Calculation Restraint Files

General. Detailed information on NMR data acquisition and structural calculations can be found in the methods section of Chapter 3. Here are the restraint files as they were entered into CNS. The protons are named and the distance restrains are assigned using the standard CNS format.[6]

Restraints as entered in CNS for cyclic peptide 3.7c (Lys9 → TS^+):

!!Sequential and Short Range NOEs

assign (residue 1 and name HA) (residue 2 and name HN) 2.2 0.4 0.9
assign (residue 1 and name HN) (residue 14 and name HN) 2.5 0.4 0.9
assign (residue 2 and name HA) (residue 3 and name HN) 1.2 0.4 0.9
assign (residue 3 and name HA) (residue 4 and name HN) 1.6 0.4 0.9
assign (residue 5 and name HA) (residue 6 and name HG#) 2.6 0.4 0.9
assign (residue 5 and name HA) (residue 6 and name HD#) 2.0 0.4 0.9
assign (residue 5 and name HG##) (residue 6 and name HG#) 2.4 0.4 0.9
assign (residue 6 and name HA) (residue 7 and name HN) 2.6 0.4 0.9
assign (residue 7 and name HN) (residue 8 and name HN) 2.5 0.4 0.9
assign (residue 8 and name HA) (residue 9 and name HN) 2.1 0.4 0.9
assign (residue 8 and name HB#) (residue 9 and name HN) 2.4 0.4 0.9
assign (residue 9 and name HA) (residue 10 and name HN) 2.5 0.4 0.9
assign (residue 10 and name HA) (residue 11 and name HN) 2.2 0.4 0.9
assign (residue 11 and name HA) (residue 12 and name HN) 2.2 0.4 0.9
assign (residue 12 and name HA) (residue 13 and name HD2) 2.3 1.0 0.4
assign (residue 12 and name HA) (residue 13 and name HD1) 2.3 1.0 0.4

!!! Medium and Long Range NOEs

assign (residue 1 and name HN) (residue 12 and name HN) 2.6 0.4 0.9
assign (residue 1 and name HN) (residue 12 and name HE2#) 2.5 0.4 0.9
assign (residue 2 and name HB#) (residue 4 and name HG#) 2.6 0.4 0.9
assign (residue 2 and name HB#) (residue 9 and name HG2#) 2.4 0.4 0.9
assign (residue 2 and name HD#) (residue 9 and name HG2#) 2.4 0.4 0.9
assign (residue 2 and name HE#) (residue 9 and name HG2#) 2.9 0.4 0.9
assign (residue 2 and name HD#) (residue 10 and name HN) 2.7 0.4 0.9
assign (residue 2 and name HA) (residue 11 and name HA) 2.1 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HG) 2.4 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HD##) 4.1 0.4 0.9

assign (residue 2 and name HE#) (residue 11 and name HD##) 3.1 0.4 0.9
assign (residue 3 and name HN) (residue 10 and name HN) 2.1 0.4 0.9
assign (residue 4 and name HG#) (residue 9 and name HD1#) 2.2 0.4 0.9
assign (residue 4 and name HB#) (residue 9 and name HD1#) 2.5 0.4 0.9
assign (residue 5 and name HN) (residue 8 and name HN) 2.6 0.4 0.9
assign (residue 5 and name HG##) (residue 8 and name HG#) 2.6 0.4 0.9
assign (residue 11 and name HD##) (residue 14 and name HA#) 2.7 0.4 0.9

Restraints as entered in CNS for cyclic peptide 3.14c (Glu4 → TS^-):

!!Sequential and Short Range NOEs
assign (residue 1 and name HN) (residue 14 and name HN) 2.4 0.4 0.9
assign (residue 1 and name NH) (residue 14 and name HA1) 3.1 0.4 0.9
assign (residue 2 and name HN) (residue 2 and name HD#) 2.5 0.4 0.9
assign (residue 2 and name HB#) (residue 2 and name HD#) 2.4 0.4 0.9
assign (residue 2 and name HB#) (residue 2 and name HD#) 2.3 0.4 0.9
assign (residue 2 and name HB#) (residue 2 and name HE#) 3.2 0.4 0.9
assign (residue 2 and name HA) (residue 3 and name HN) 2.1 0.4 0.9
assign (residue 2 and name HA) (residue 3 and name HG##) 2.8 0.4 0.9
assign (residue 2 and name HD#) (residue 3 and name HN) 2.7 0.4 0.9
assign (residue 3 and name HA) (residue 4 and name HN) 2.5 0.4 0.9
assign (residue 4 and name HA) (residue 5 and name HN) 2.4 0.4 0.9
assign (residue 4 and name H32#) (residue 5 and name HN) 2.7 0.4 0.9
assign (residue 5 and name HA) (residue 6 and name HG#) 2.8 0.4 0.9
assign (residue 5 and name HA) (residue 6 and name HD#) 2.4 0.4 0.9
assign (residue 5 and name HA) (residue 6 and name HD#) 2.3 0.4 0.9
assign (residue 5 and name HG##) (residue 6 and name HD#) 1.8 0.4 0.9
assign (residue 5 and name HG##) (residue 6 and name HD#) 2.4 0.4 0.9
assign (residue 6 and name HA) (residue 7 and name HN) 2.6 0.4 0.9
assign (residue 7 and name HN) (residue 8 and name HN) 2.5 0.4 0.9
assign (residue 9 and name HA) (residue 10 and name HN) 2.4 0.4 0.9
assign (residue 10 and name HA) (residue 11 and name HN) 2.1 0.4 0.9
assign (residue 10 and name HB) (residue 11 and name HN) 2.6 0.4 0.9
assign (residue 11 and name HA) (residue 12 and name HN) 2.6 0.4 0.9
assign (residue 12 and name HB#) (residue 12 and name HE2#) 3.1 0.4 0.9
assign (residue 12 and name HB#) (residue 12 and name HE2#) 2.9 0.4 0.9
assign (residue 12 and name HG#) (residue 12 and name HE2#) 2.4 0.4 0.9
assign (residue 12 and name HA) (residue 13 and name HD#) 2.4 0.4 0.9
assign (residue 12 and name HA) (residue 13 and name HD#) 2.3 0.4 0.9
assign (residue 13 and name HA) (residue 14 and name HN) 2.2 0.4 0.9

!!Medium and Long Range NOEs
assign (residue 1 and name HD#) (residue 3 and name HG##) 2.5 0.4 0.9
assign (residue 1 and name HN) (residue 11 and name HD##) 3.0 0.4 0.9

assign (residue 1 and name HN) (residue 12 and name HN) 2.6 0.4 0.9
assign (residue 2 and name HB#) (residue 4 and name H32#) 2.3 0.4 0.9
assign (residue 2 and name HB#) (residue 9 and name HD#) 2.4 0.4 0.9
assign (residue 2 and name HD#) (residue 9 and name HE#) 3.1 0.4 0.9
assign (residue 2 and name HD#) (residue 9 and name HE#) 2.4 0.4 0.9
assign (residue 2 and name HE#) (residue 9 and name HE#) 3.0 0.4 0.9
assign (residue 2 and name HD#) (residue 10 and name HN) 2.8 0.4 0.9
assign (residue 2 and name HD#) (residue 11 and name HN) 2.7 0.4 0.9
assign (residue 2 and name HD#) (residue 11 and name HA) 2.3 0.4 0.9
assign (residue 2 and name HD#) (residue 11 and name HD##) 3.0 0.4 0.9
assign (residue 2 and name HD#) (residue 11 and name HD##) 3.0 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HN) 2.8 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HD##) 2.8 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HD##) 2.9 0.4 0.9
assign (residue 2 and name HE#) (residue 11 and name HD##) 3.5 3.0 0
assign (residue 5 and name HN) (residue 8 and name HN) 2.5 0.4 0.9
assign (residue 8 and name HD#) (residue 10 and name HD##) 2.6 0.4 0.9
assign (residue 11 and name HD##) (residue 14 and name HA#) 2.7 0.4 0.9
assign (residue 11 and name HD##) (residue 14 and name HA#) 2.8 0.4 0.9

Supplementary Section B11
Bibliography

(1) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quanfification of Beta-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577–11578.

(2) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. Length-Dependent Stability and Strand Length Limits in Antiparallel Beta-Sheet Secondary Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12015–12020.

(3) Espinosa, J. F.; Gellman, S. H. A Designed Beta-Hairpin Containing a Natural Hydrophobic Cluster. **2000**, *11*, 2330–2333.

(4) Haque, T. S.; Gellman, S. H. Insights on Beta-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I′ and Type II′ Beta-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

(5) Syud, F. A.; Stanger, H. E.; Gellman, S. H. Interstrand Side Chain - Side Chain Interactions in a Designed Beta-Hairpin: Significance of Both Lateral and Diagonal Pairings. *J. Am. Chem. Soc.* **2001**, *123*, 8667–8677.

(6) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; et al. Crystallography & NMR System: A New Software Suite for Macromolecular

Structure Determination. *Acta Crystallogr. D. Biol. Crystallogr.* **1998**, *54*, 905–921.

Appendix C: Supporting information for Chapter 4

Supplementary Section C1
Peptide purity checks

General. For information regarding sample acquisition, see the methods section of Chapter 4.

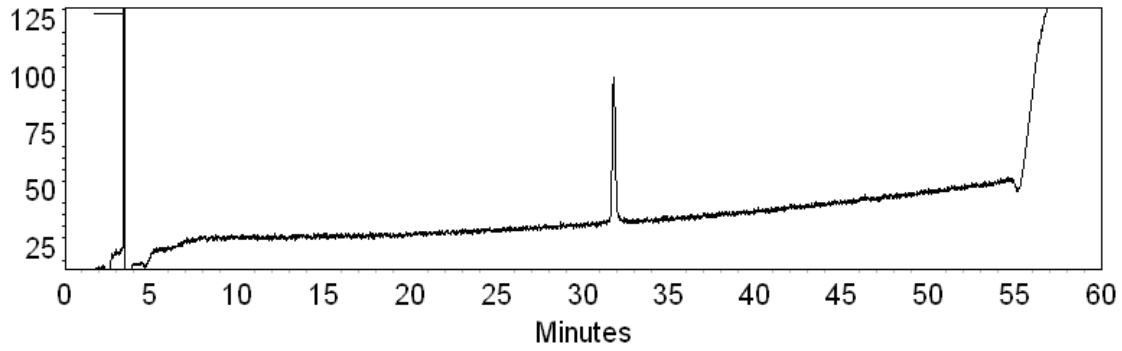


Figure C1.1. Purity check of peptide 4.1 monitored at 220 nm, Ac-RLEKELEEKKEALELAIDQASRDYHRATALEKELEEK-NH_2

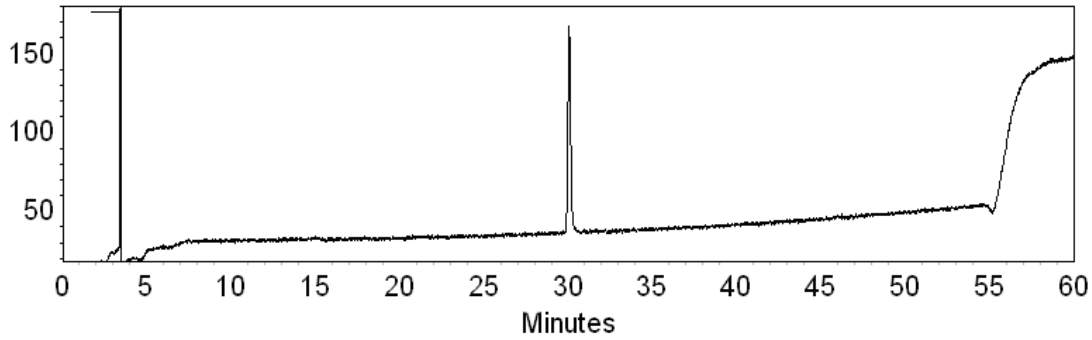


Figure C1.2. Purity check of peptide 4.2 monitored at 220 nm, Ac-KKELEEKLEALELAIDQAKRKYHRAKAKLKELEEK-NH_2

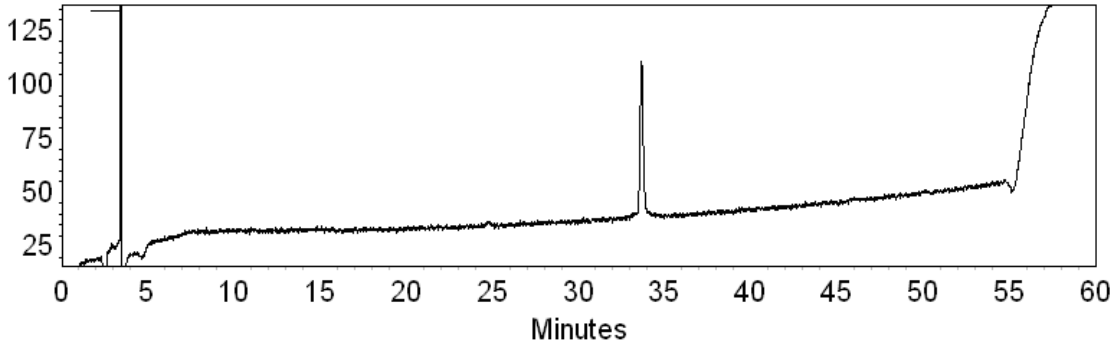


Figure C1.3. Purity check of peptide 4.3 monitored at 220 nm, Ac-KKELEEKLEALELAIDQLKRKAHRLKAKLKELEEK-NH_2

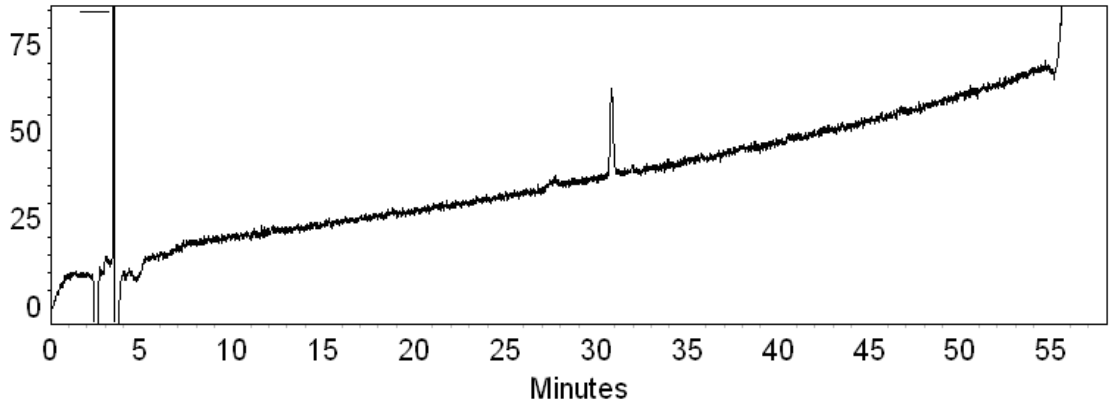


Figure C1.4. Purity check of peptide 4.4 monitored at 220 nm, Ac-YGGLEALKLKIDQLKRKAHRLKAKLKELKEK-NH_2

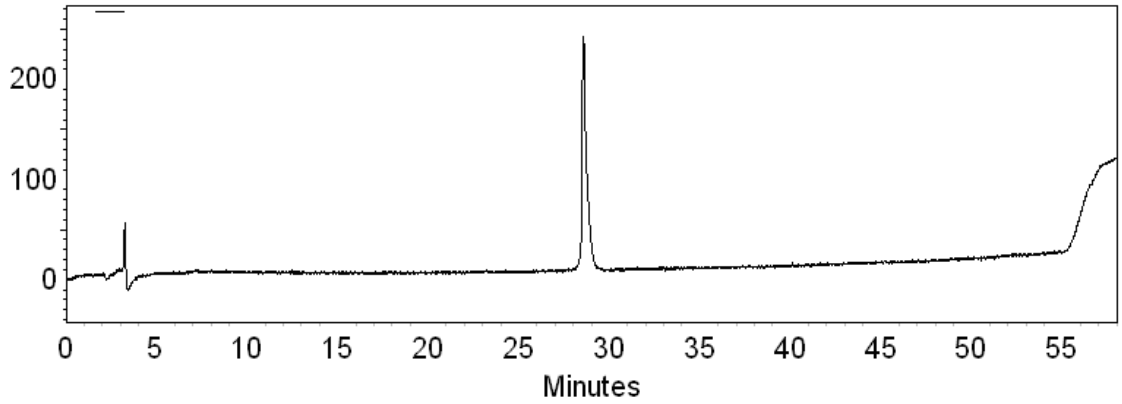


Figure C1.5. Purity check of peptide 4.5 monitored at 220 nm, Ac-YGGLEALKLKIDQAKRKAHRAKAKLKELKEK-NH_2

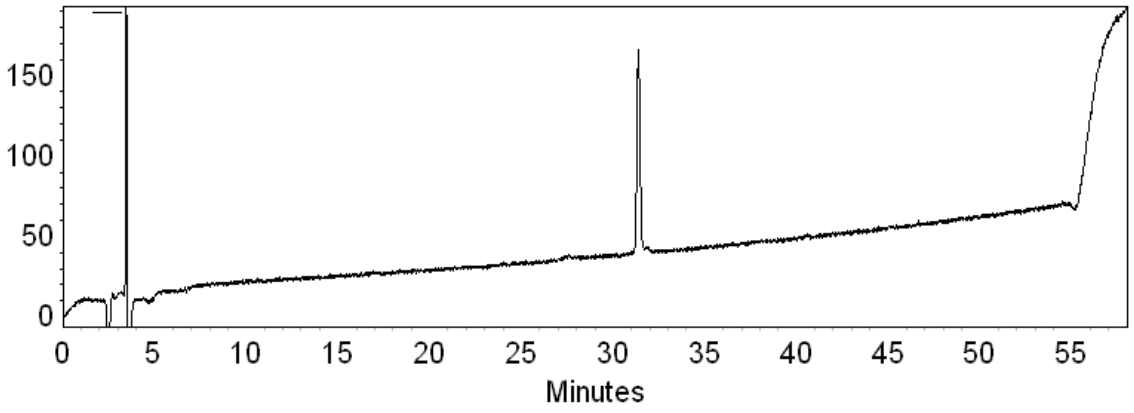


Figure C1.7. Purity check of peptide 4.7 monitored at 220 nm, Ac-LEALKLKIDQLKRKYHRLKAKLKELKEK-NH_2

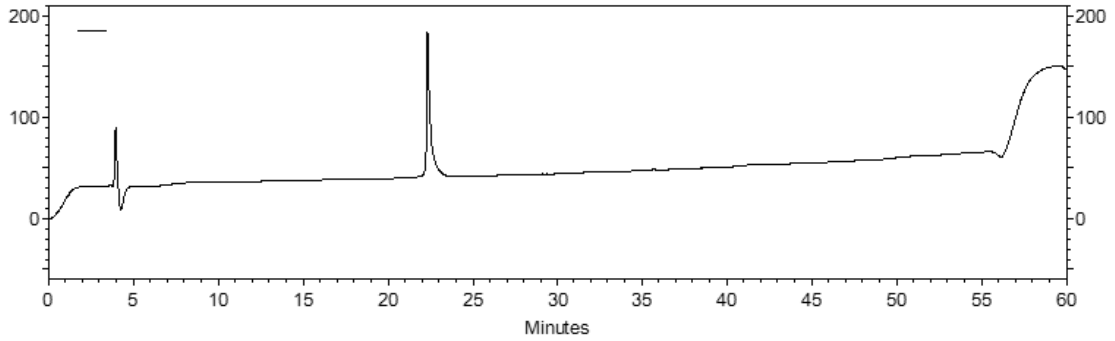


Figure C1.8. Purity check of peptide 4.8 monitored at 220 nm, Ac-YGGLEAKKAKIEQLKRKLHRLKAKLKEKKEK-NH_2

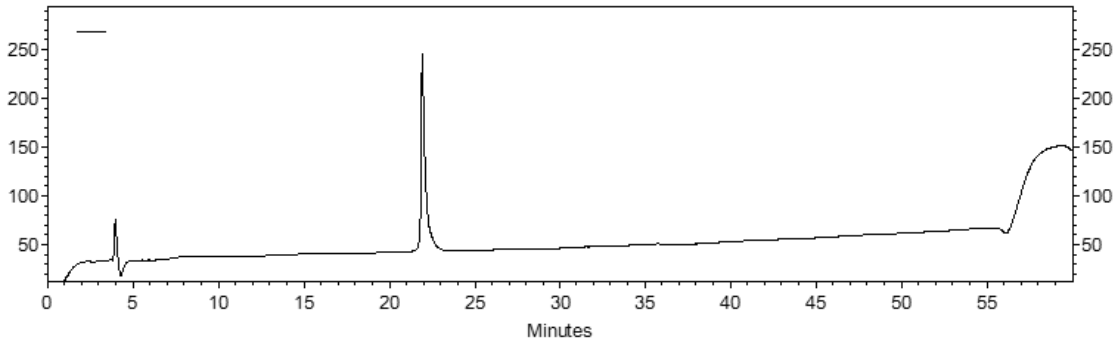


Figure C1.9. Purity check of peptide 4.9 monitored at 220 nm, Ac-YGGLEAKKAKIEQLKRKIHRLKAKIKEKKEK-NH_2

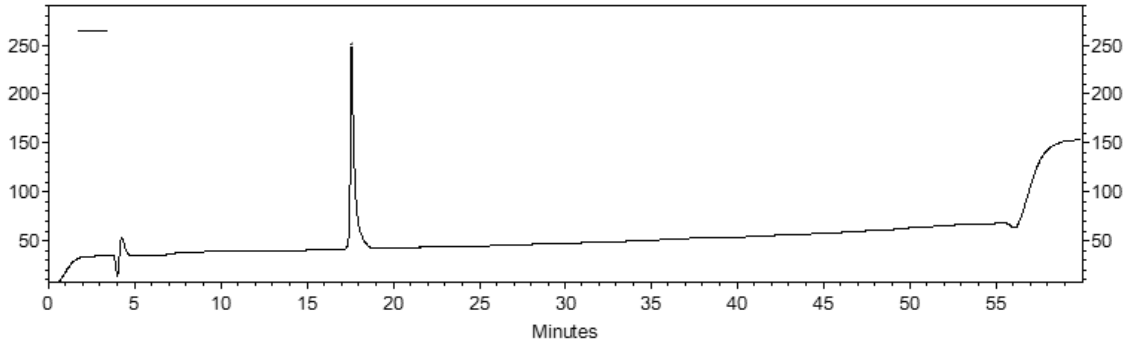


Figure C1.10. Purity check of peptide 4.10 monitored at 220 nm, Ac-YGGLEAKKAKIEQLKRKKHRLKAKIKEKKEK-NH_2

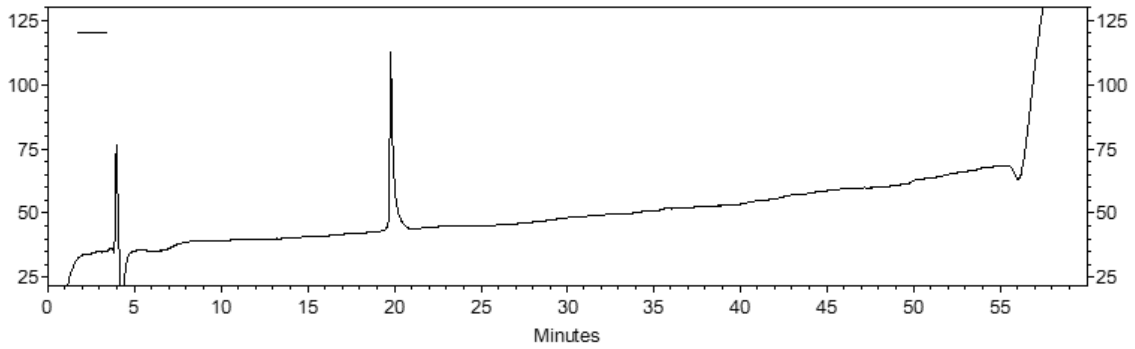


Figure C1.11. Purity check of peptide 4.11 monitored at 220 nm, Ac-YGGLEAKKAKIEQLKRKAHRLKAKIKEKKEK-NH_2

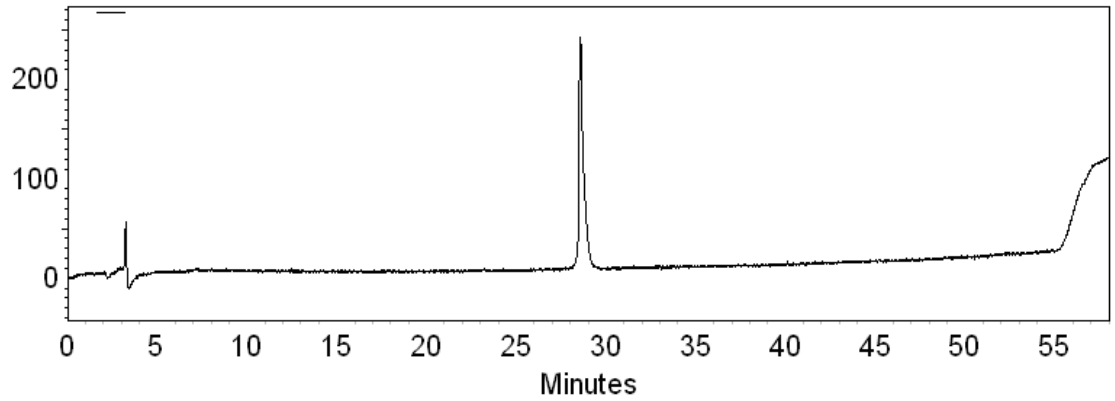


Figure C1.12. Purity check of peptide 4.12 monitored at 220 nm, Ac-IDQASRDYHRATALEKELEEKKKALELAIDQASQDYN-NH_2

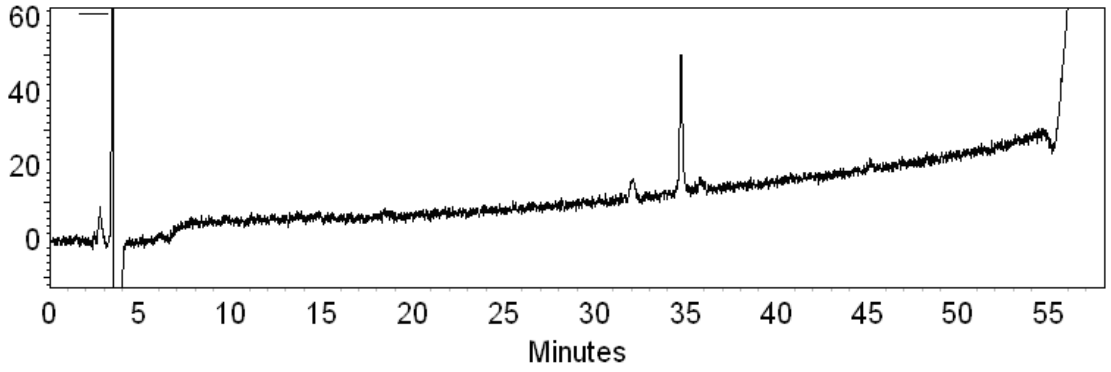


Figure C1.13. Purity check of peptide 4.13 monitored at 220 nm, Ac-YGGRLSRLKHRITALQKELEEEEKQLELLRDRISQLR-NH_2

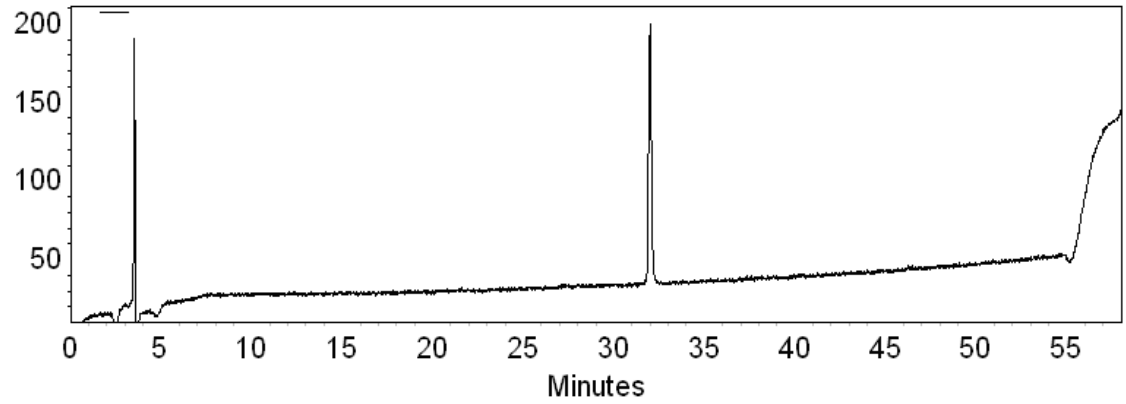


Figure C1.14. Purity check of peptide 4.14 monitored at 220 nm, Ac-YGGRESRLKHRLTALQKELEEEEKQLELLRERLSQER-NH_2

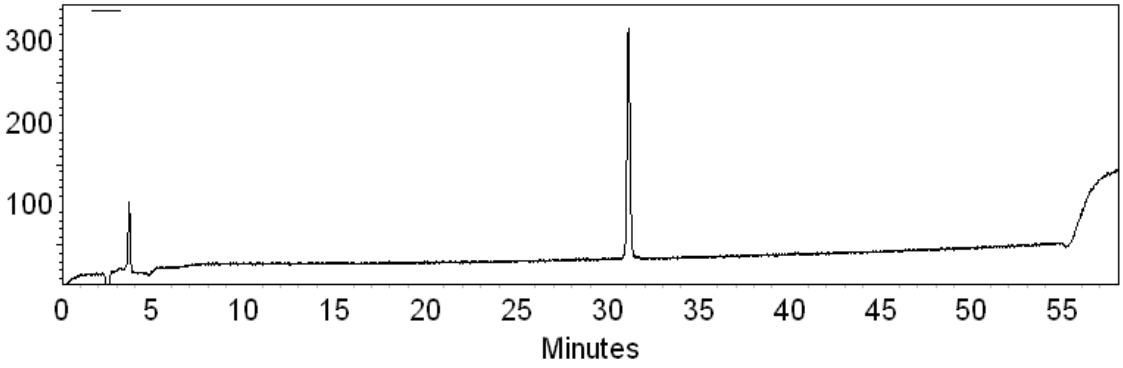


Figure C1.15. Purity check of peptide 4.15 monitored at 220 nm, Ac-YGGRRSRLKHRLTALQKELEEEEKQLELLRERLSQRR-NH_2

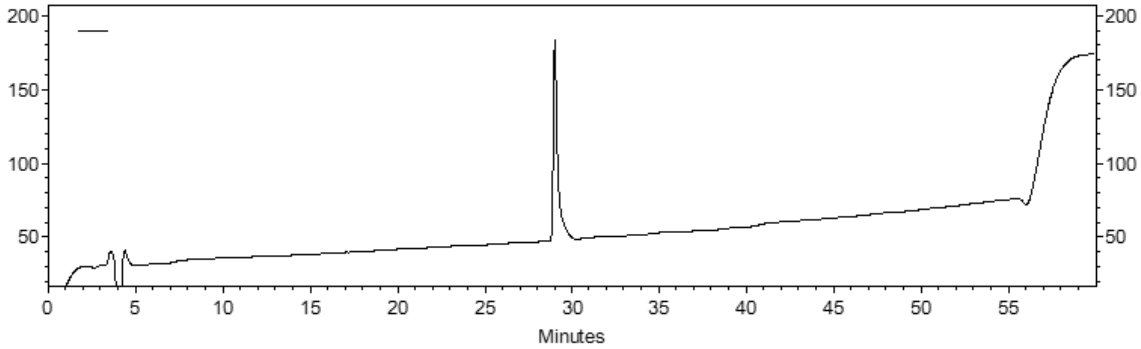


Figure C1.16. Purity check of peptide 4.16 monitored at 220 nm, Ac-YGGKKKKLEEEEEELELLKDKISRLKHRITALKKKLE-NH_2

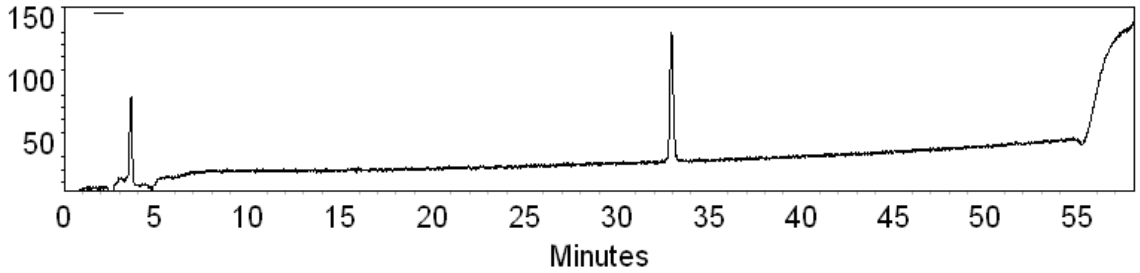


Figure C1.17. Purity check of peptide 4.17 monitored at 220 nm, Ac-YGGRRSRLKHRLTALRKRLEELEKRLELLRERLSQRR-NH_2

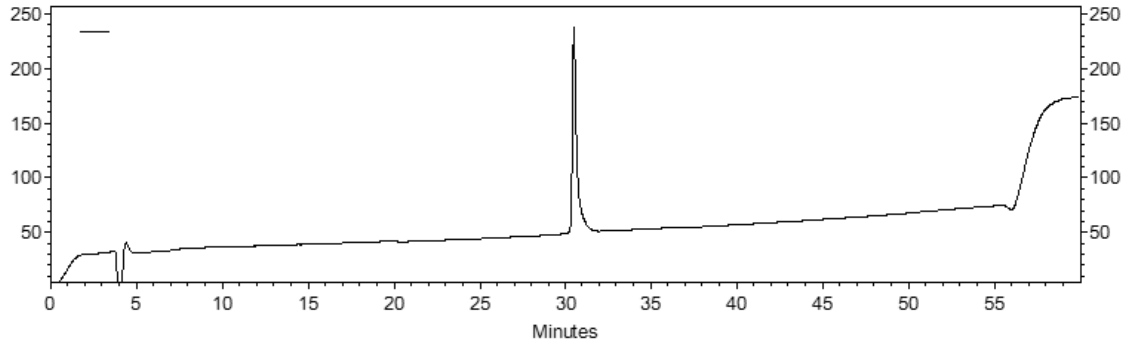


Figure C1.18. Purity check of peptide 4.18 monitored at 220 nm, Ac-YGGRRSRLKHRITALRKRLEELEKRLELLRERISQRR-NH_2

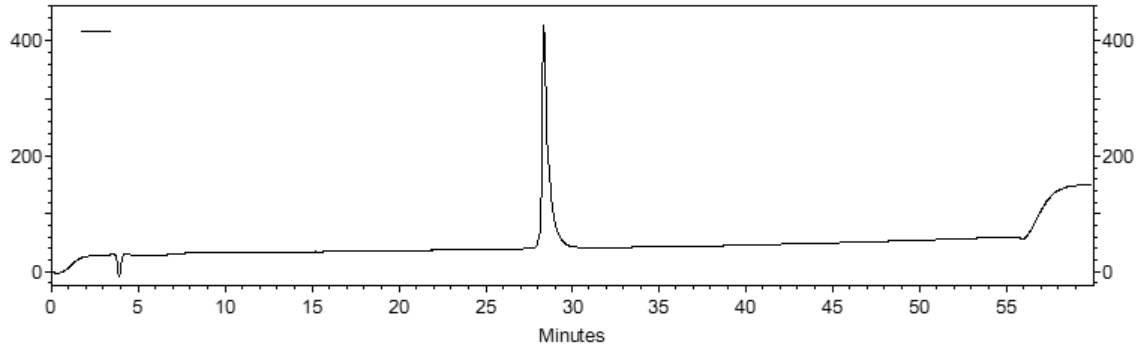


Figure C1.19. Purity check of peptide 4.19 monitored at 220 nm, Ac-YGGRTSRLKHERTALDKKIEELEKQLELLEDEVSQLQ-NH_2

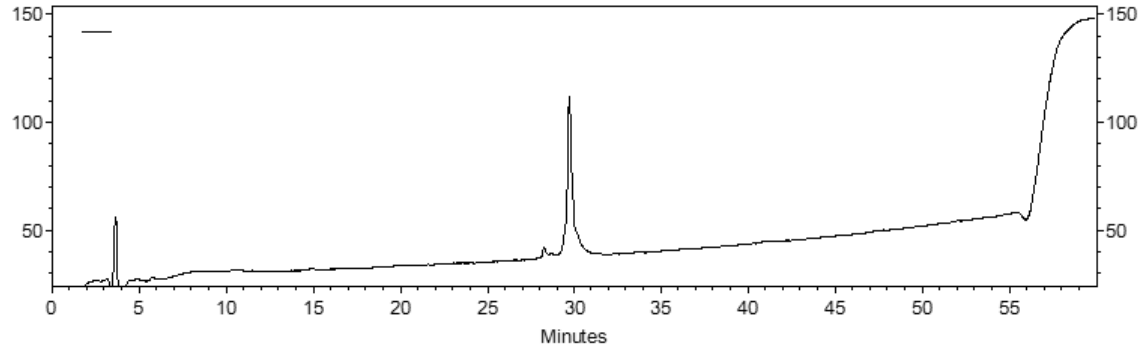


Figure C1.20. Purity check of peptide 4.20 monitored at 220 nm, Ac-YGGKASRAEHRATALEKKIEELEKRLELAEDKVSQAQ-NH_2

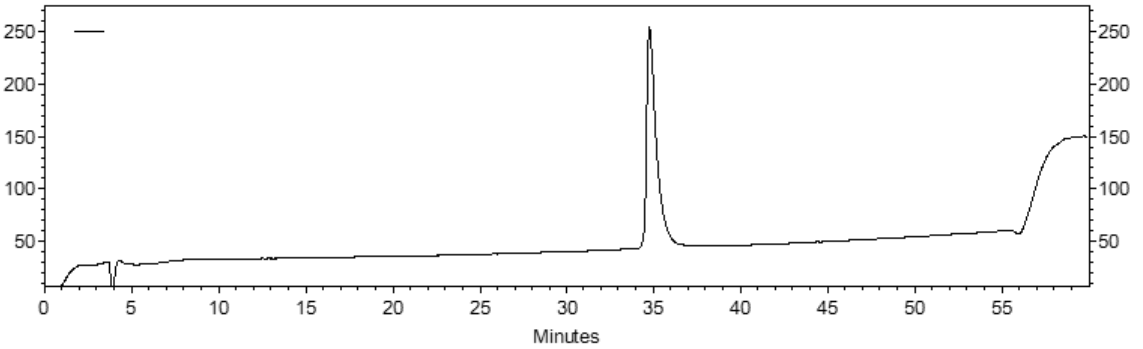


Figure C1.21. Purity check of peptide 4.21 monitored at 220 nm, Ac-YGGRISRLKHERAALDKKIEELEKQLEALEKEVSQLQ-NH_2

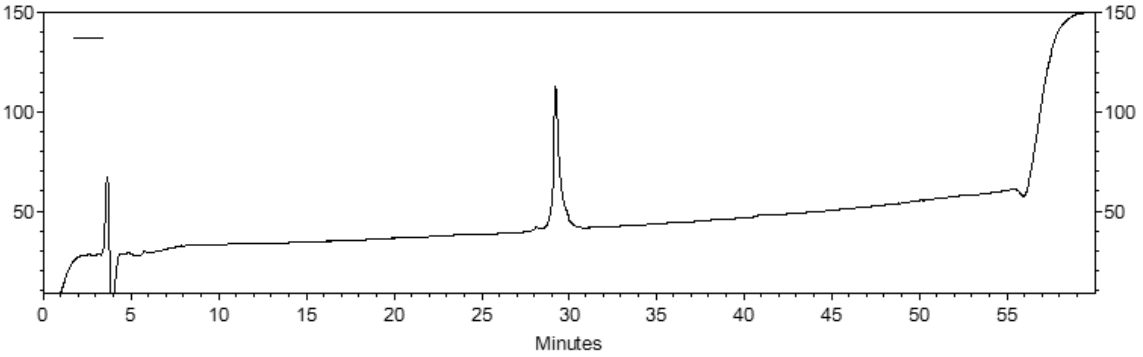


Figure C1.22. Purity check of peptide 4.22 monitored at 220 nm, Ac-YGGKASRAEHRAAALEKKIEELEKRLEAAEKKVSQAQ-NH_2

Supplementary Section C2
Mass spectrometry (MALDI) data for peptides

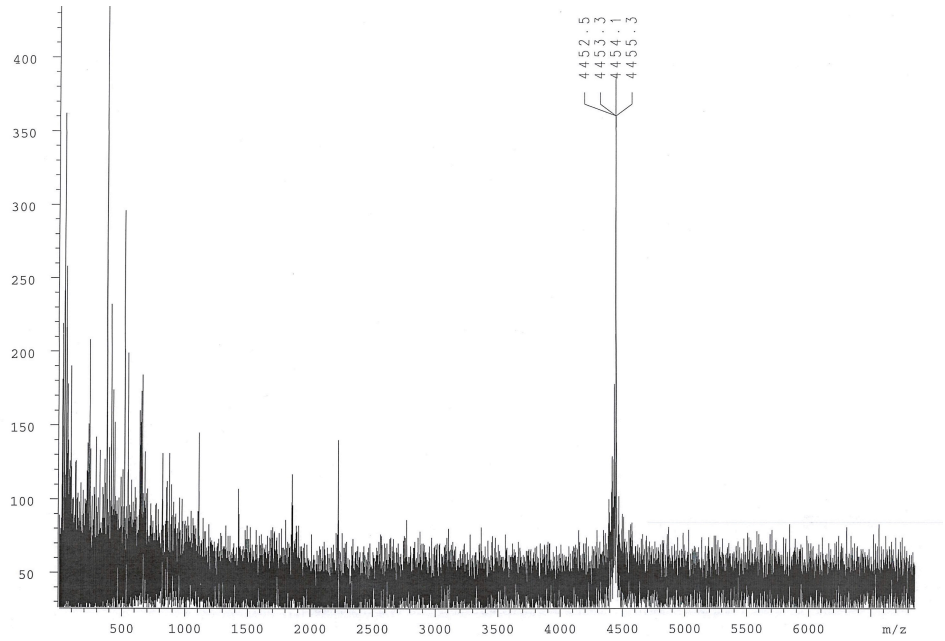


Figure C2.1. MALDI-MS of peptide 4.1 in CHCA matrix. Sequence: Ac-RLEKELEEKKEALELAIDQASRDYHRATALEKELEEK-NH_2 Expected: 4452.3 m/z (M+H)

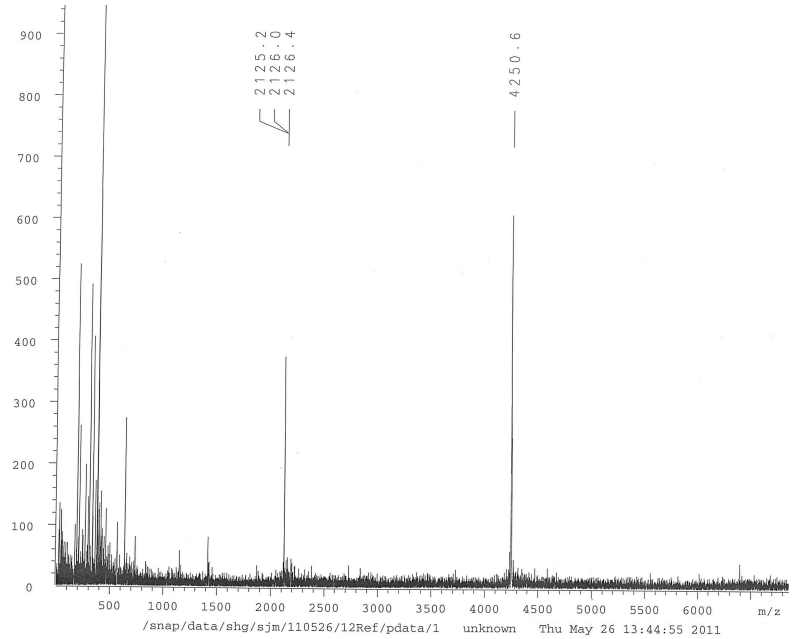


Figure C2.2. MALDI-MS of peptide 4.2 in CHCA matrix. Sequence: Ac-KKELEEKLEALELAIDQAKRKYHRAKAKLKELEEK-NH_2 Expected: 4247.3 m/z (M+H), 2124.2 m/z (M+2H)

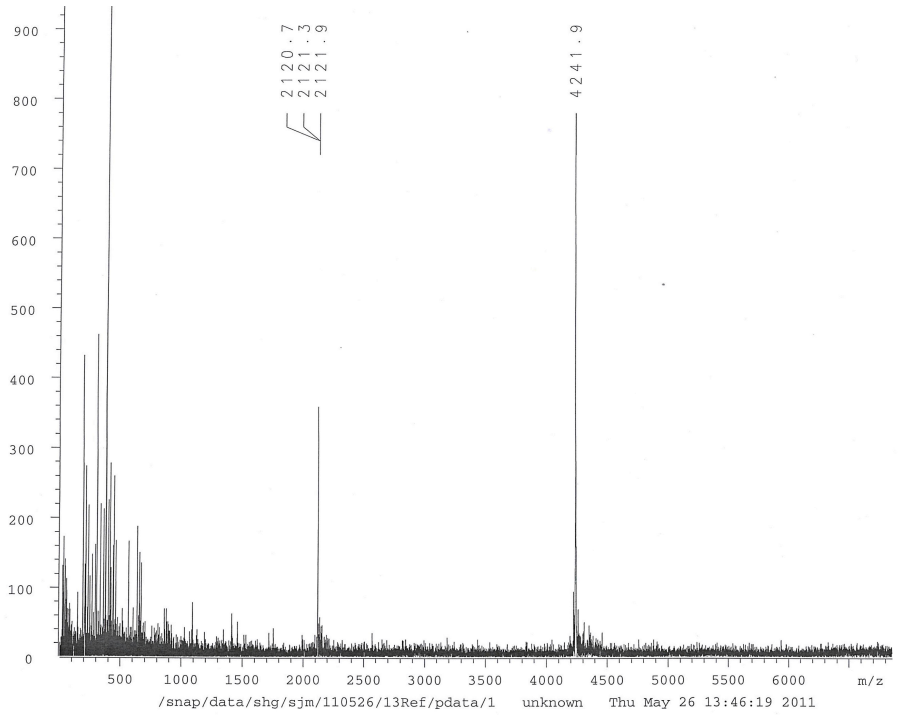


Figure C2.3. MALDI-MS of peptide 4.3 in CHCA matrix. Sequence: Ac-KKELEEKLEALELAIDQLKRKAHRLKAKLKELEEK-NH_2 Expected: 4239.3=4 m/z (M+H), 2120.7 m/z (M+2H)

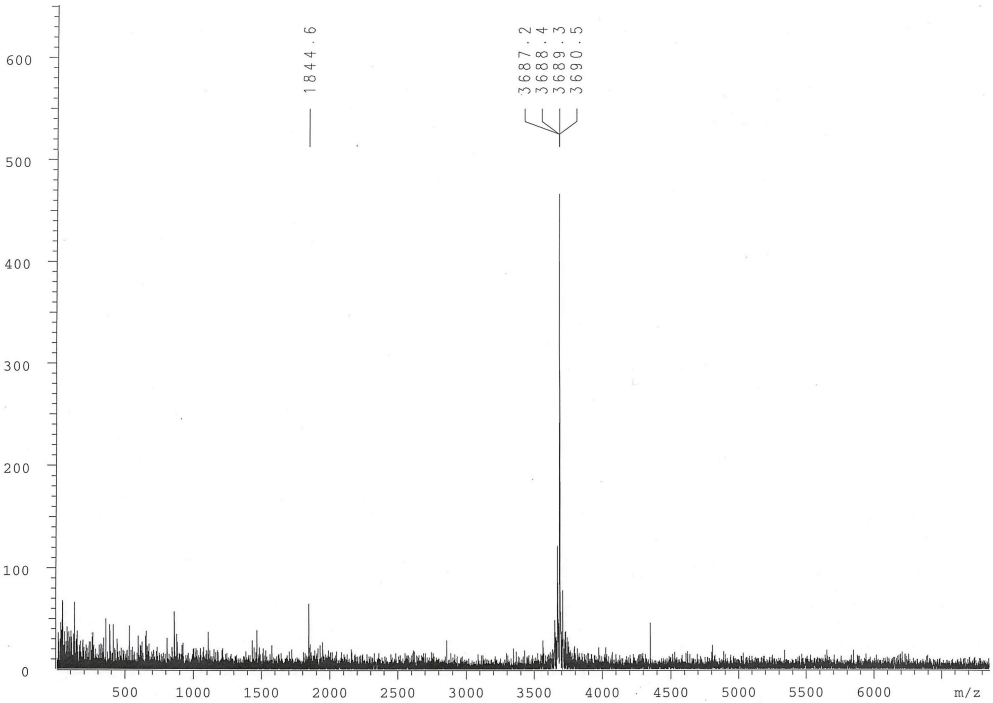


Figure C2.4. MALDI-MS of peptide 4.4 in CHCA matrix. Sequence: Ac-YGGLEALKLKIDQLKRKAHRLKAKLKELKEK-NH_2 Expected: 3687.2=4 m/z (M+H), 1844.1 m/z (M+2H)

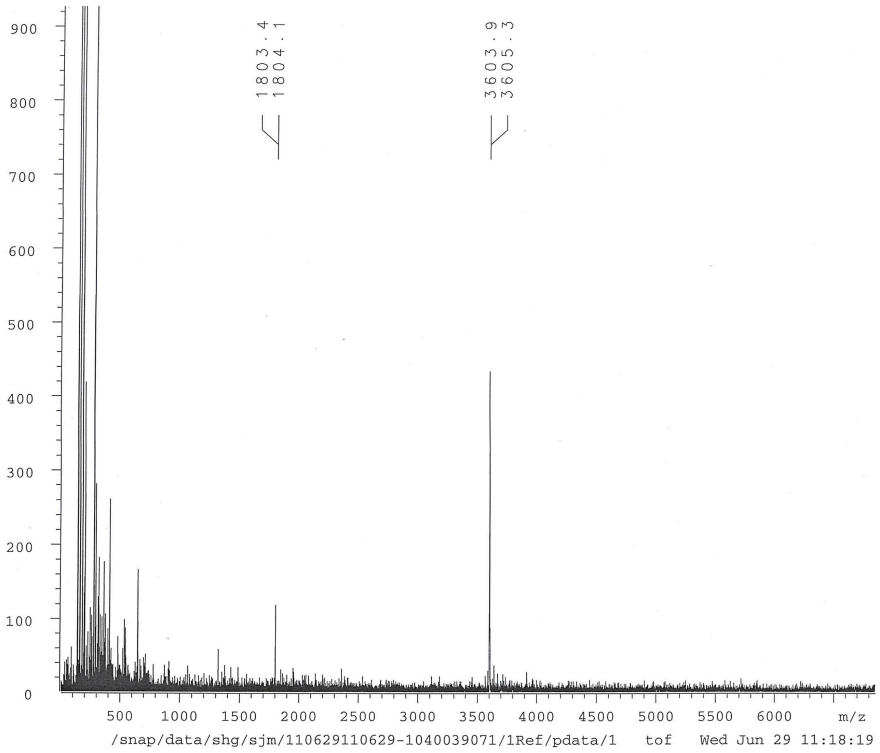


Figure C2.5. MALDI-MS of peptide 4.5 in DHB matrix. Sequence: Ac-YGGLEALKLKIDQAKRKAHRAKAKLKELKEK-NH_2 Expected: 3603.1 m/z (M+H), (M+2H)

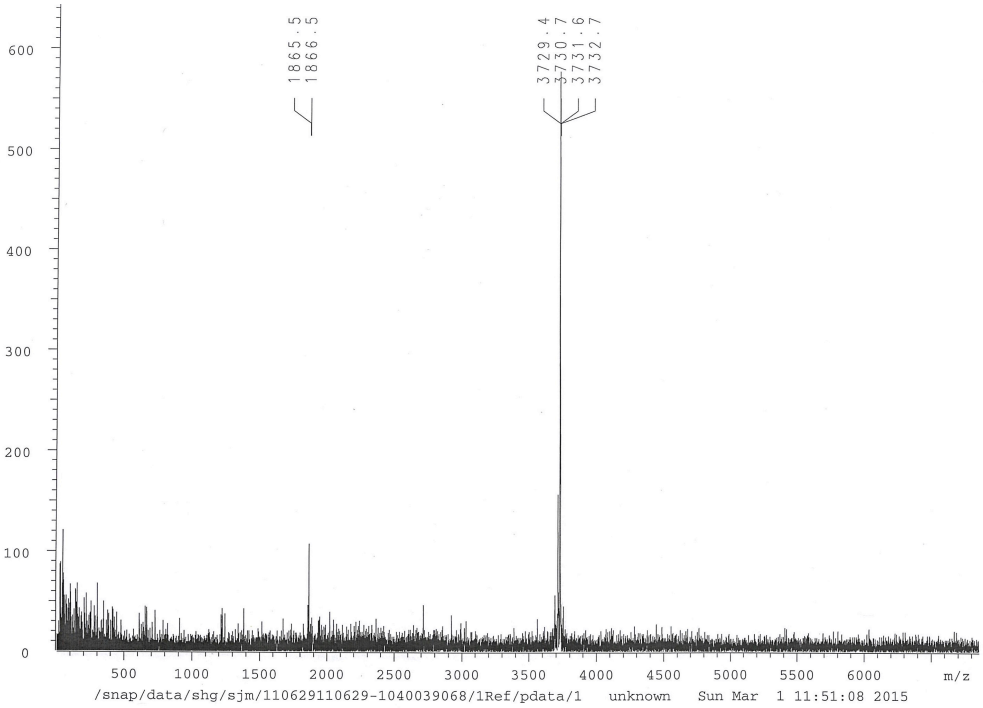


Figure C2.6. MALDI-MS of peptide 4.6 in CHCA matrix. Sequence: Ac-YGGLEALKLKIDQLKRKLHRLKAKLKELKEK-NH_2 Expected: 3729.2 m/z (M+H), 1865.1 (M+2H)

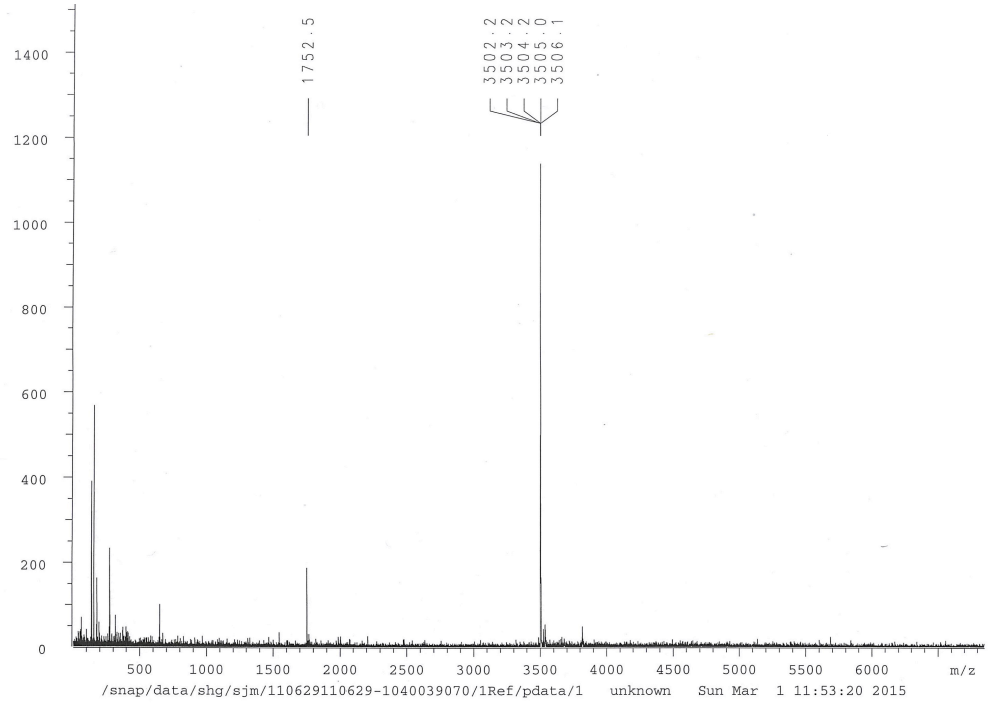


Figure C2.7. MALDI-MS of peptide 4.7 in CHCA matrix. Sequence: Ac-LEALKLKIDQLKRKYHRLKAKLKELKEK-NH_2 Expected: 3502.1 m/z (M+H)

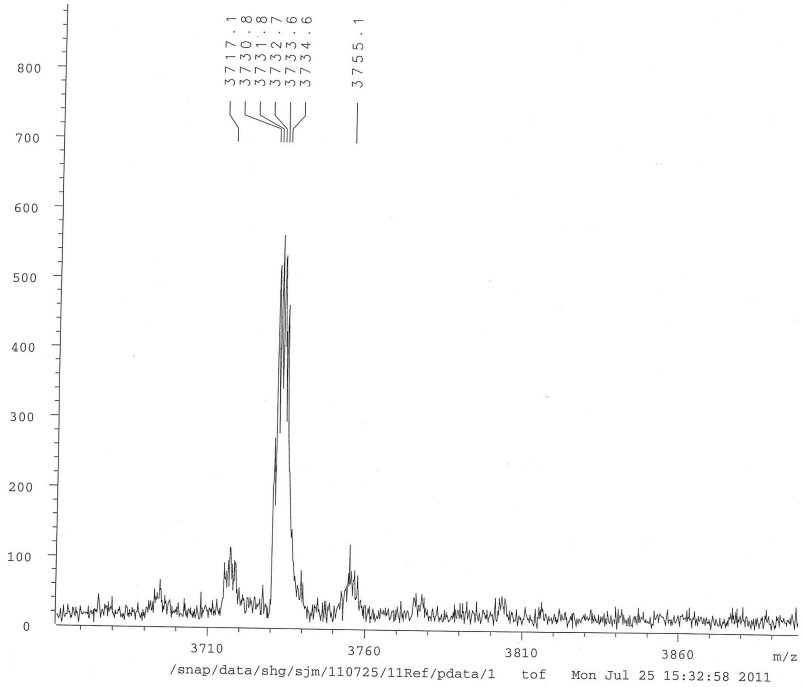


Figure C2.8. MALDI-MS of peptide 4.8 in CHCA matrix. Sequence: Ac-YGGLEAKKAKIEQLKRKLHRLKAKLKEKKEK-NH_2 Expected: 3731.2 m/z (M+H)

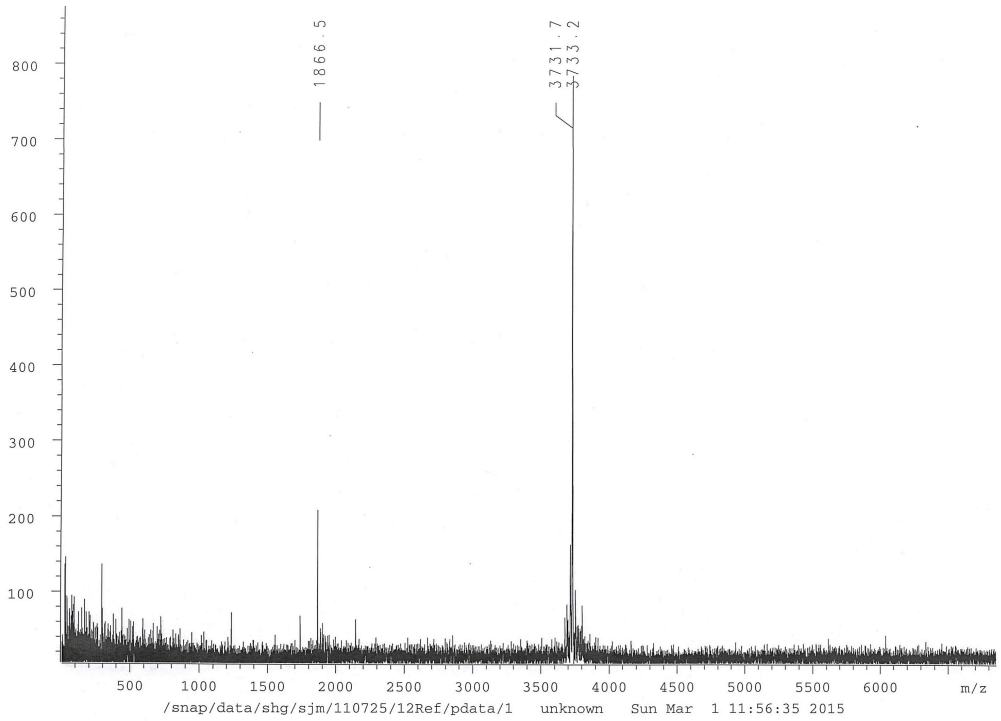


Figure C2.9. MALDI-MS of peptide 4.9 in CHCA matrix. Sequence: Ac-YGGLEAKKAKIEQLKRKIHRLKAKIKEKKEK-NH_2 Expected: 3731.2 m/z (M+H)

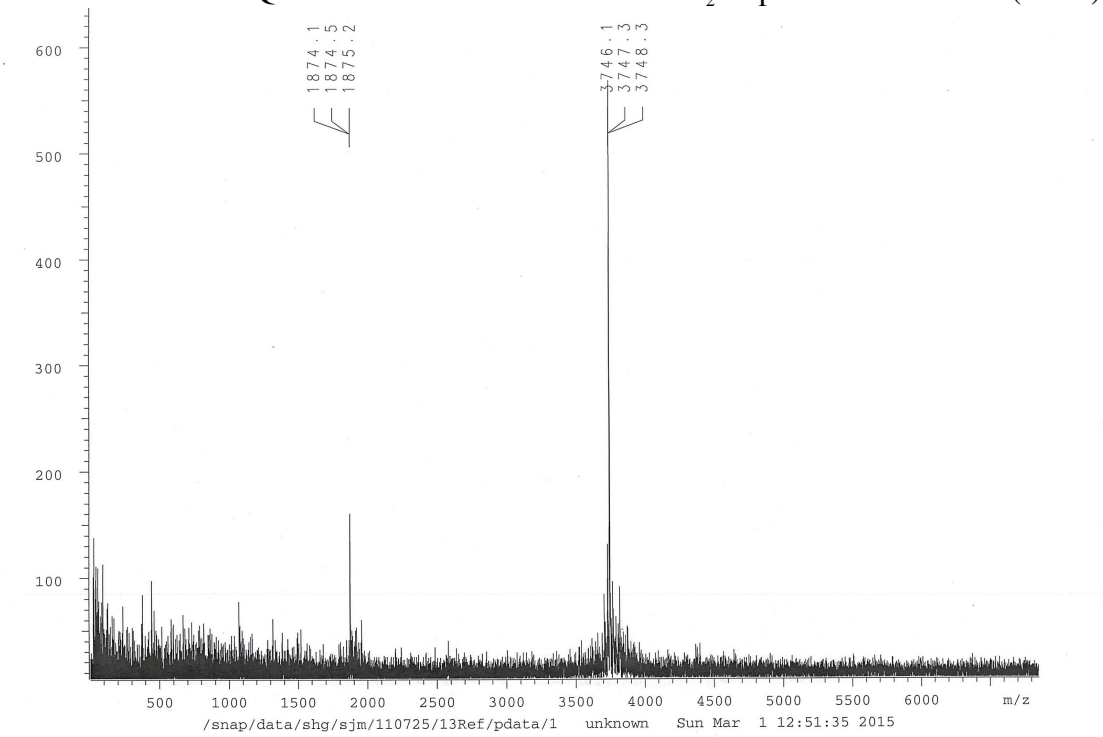


Figure C2.10. MALDI-MS of peptide 4.10 in CHCA matrix. Sequence: Ac-YGGLEAKKAKIEQLKRKKHRLKAKIKEKKEK-NH_2 Expected: 3746.1m/z (M+H)

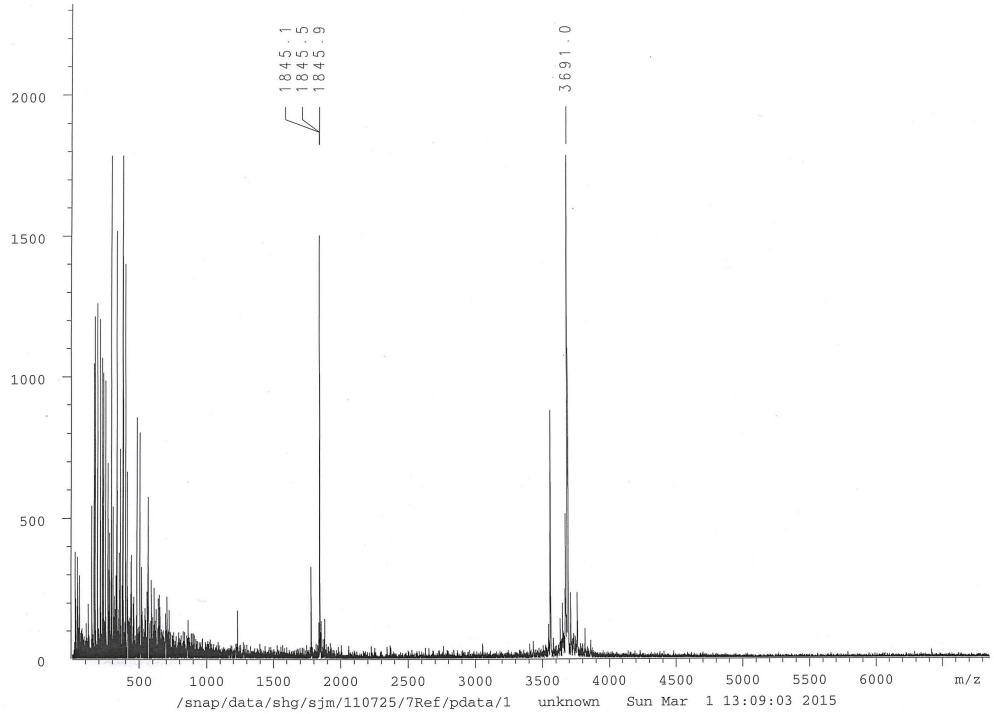


Figure C2.11. MALDI-MS of peptide 4.11 in CHCA matrix. Sequence: Ac-YGGLEAKKAKIEQLKRKAHRLKAKIKEKKEK-NH_2 Expected: 3689.1 (M+H), 1845.1 (M+2H).

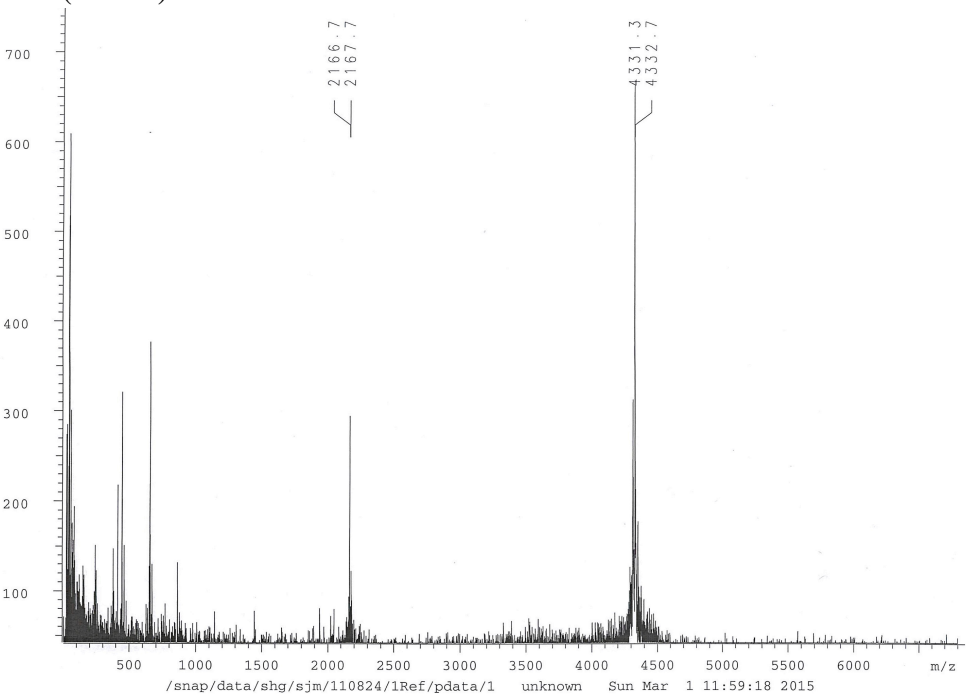


Figure C2.12. MALDI-MS of peptide 4.12 in CHCA matrix. Sequence: Ac-IDQASRDYHRATALEKELEEKKKALELAIDQASQDYN-NH_2 Expected: 4331.2 (M+H), 2166.1 (M+2H).

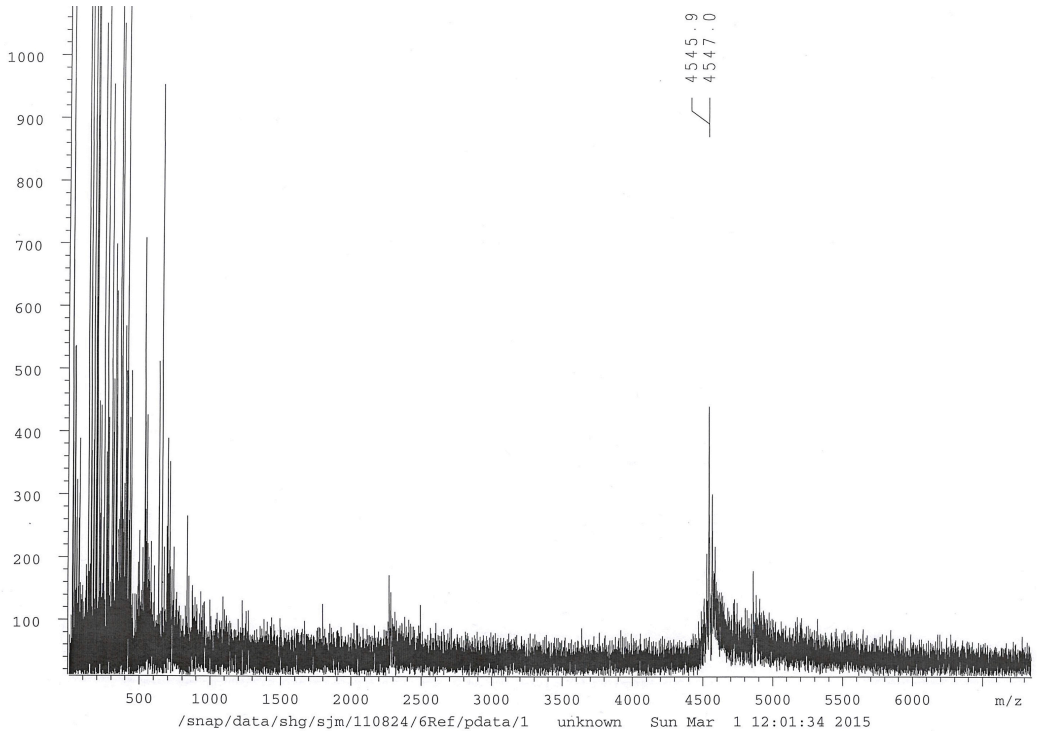


Figure C2.13. MALDI-MS of peptide 4.13 in CHCA matrix. Sequence: Ac-YGGRLSRLKHRITALQKELEEEEKQLELLRDRISQLR-NH_2 Expected: 4545.5 (M+H)

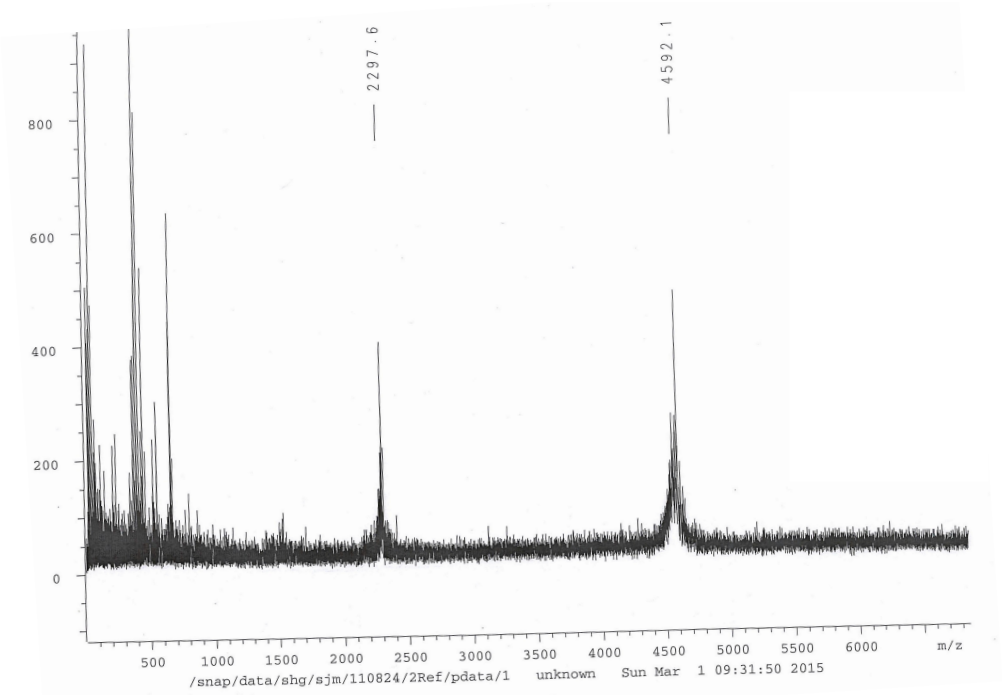


Figure C2.14. MALDI-MS of peptide 4.14 in CHCA matrix. Sequence: Ac-YGGRESRLKHRLTALQKELEEEEKQLELLRERLSQER-NH_2 Expected: 4591.4 (M+H), 2296.2 (M+2H)

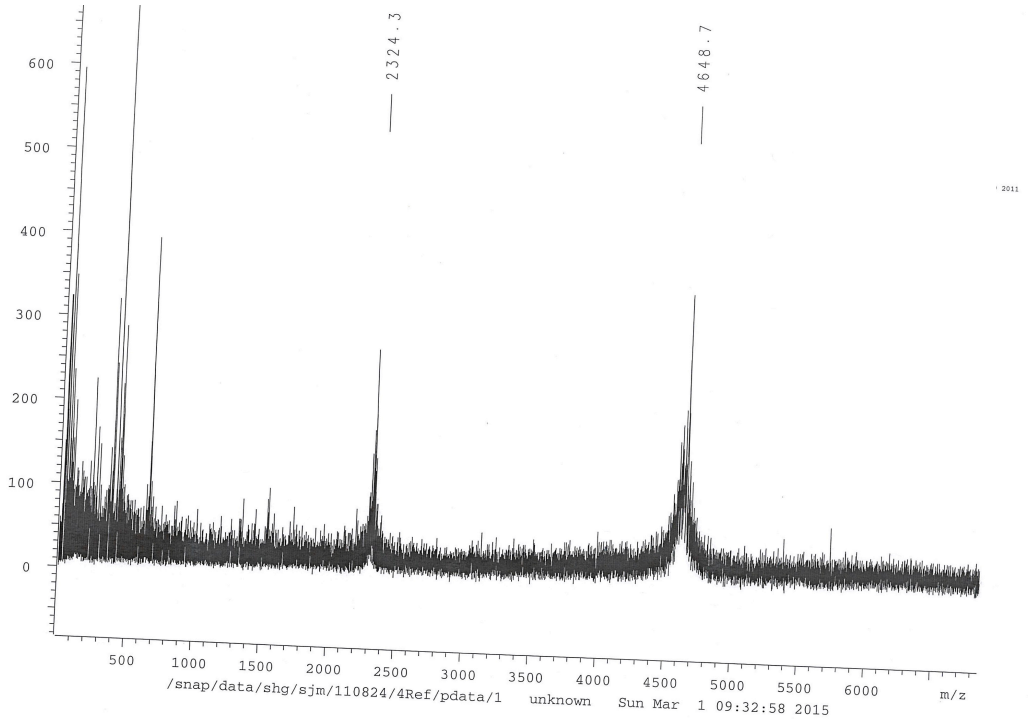


Figure C2.15. MALDI-MS of peptide 4.15 in CHCA matrix. Sequence: Ac-YGGRRSRLKHRLTALQKELEEEEKQLELLRERLSQRR-NH_2 Expected: 4645.5 (M+H), 2323.2 (M+2H)

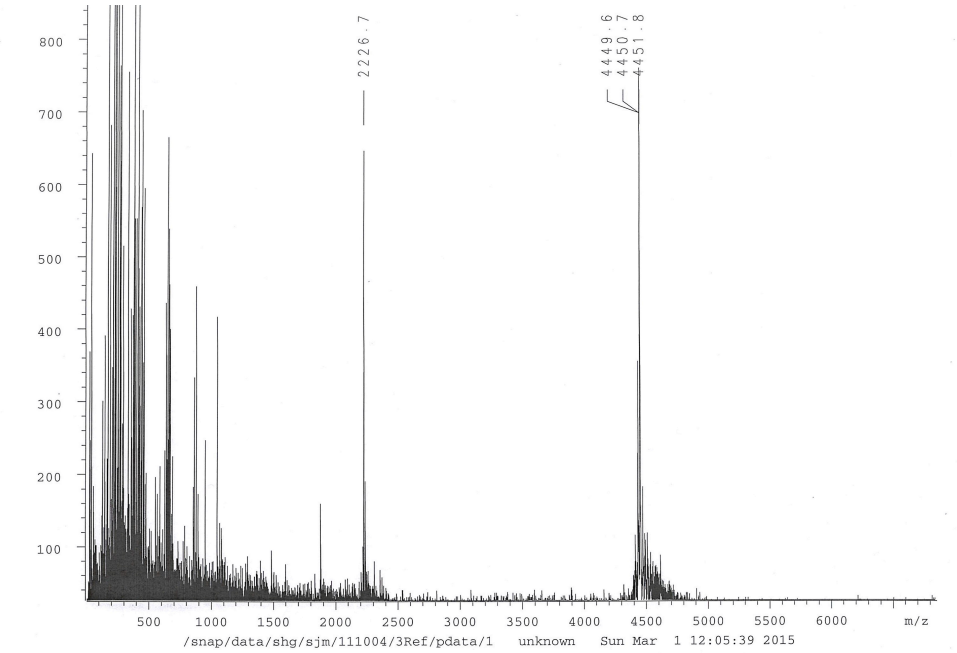


Figure C2.16. MALDI-MS of peptide 4.16 in CHCA matrix. Sequence: Ac-YGGKKKKLEEEEEELELLKDKISRLKHRITALKKKLE-NH_2 Expected: 4449.5 (M+H)

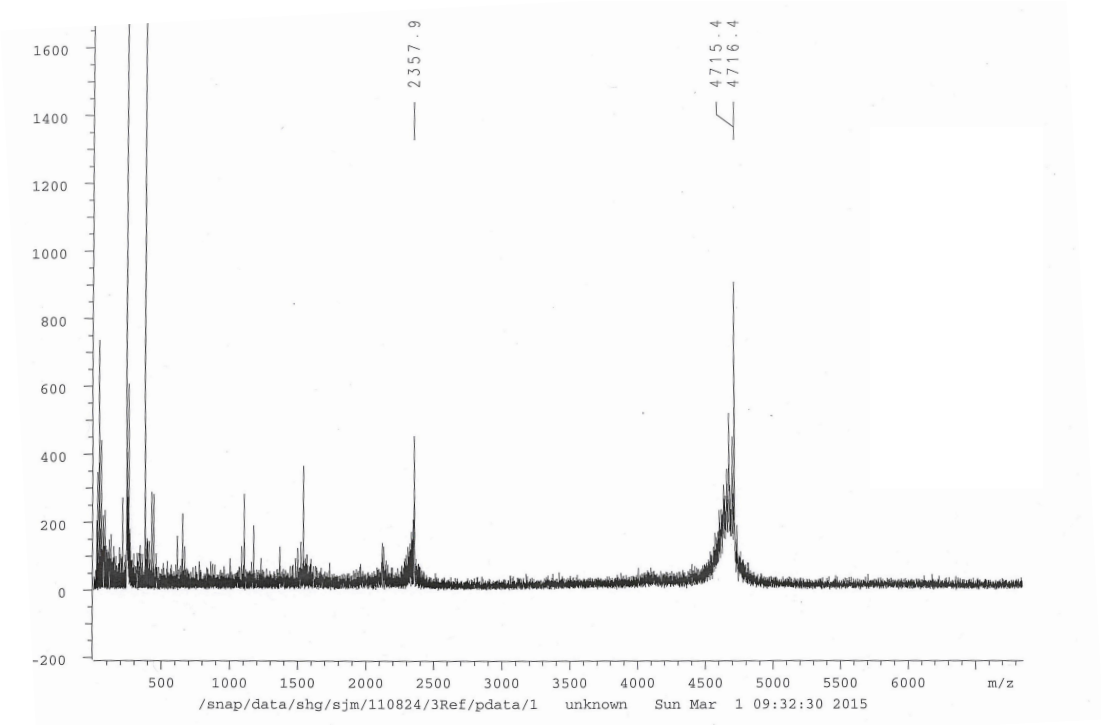


Figure C2.17. MALDI-MS of peptide 4.17 in CHCA matrix. Sequence: Ac-YGGRRSRLKHRLTALRKRLEELEKRLELLRERLSQRR-NH_2 Expected: 4712.7 m/z (M+H), 2356.9 m/z (M+2H).

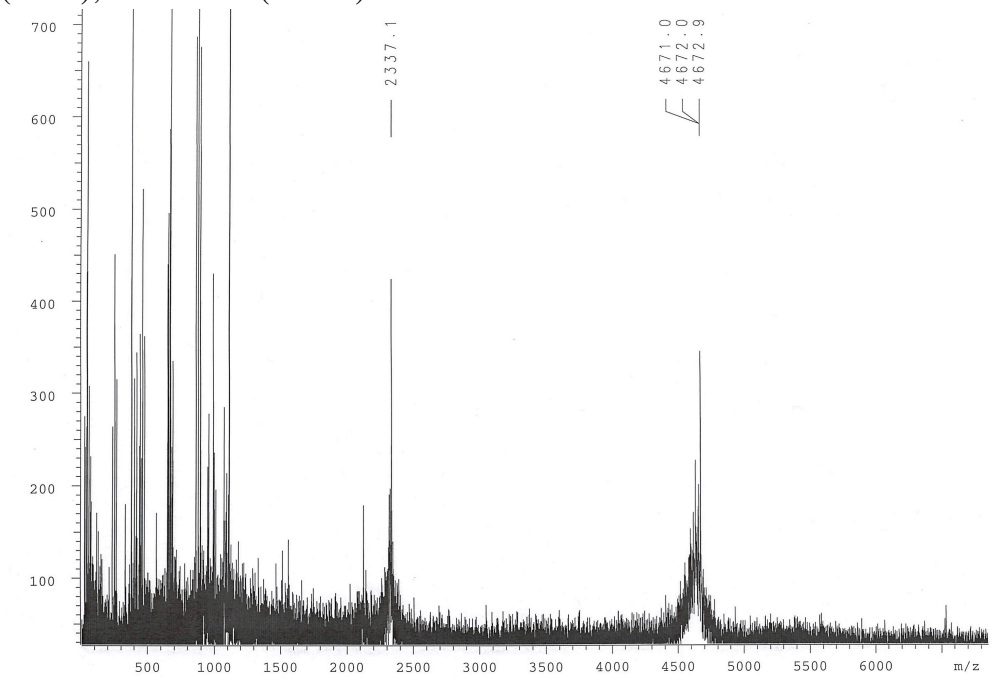


Figure C2.18. MALDI-MS of peptide 4.18 in CHCA matrix. Sequence: Ac-YGGRRSRLKHRITALRKRLEELEKRLELLRERISQRR-NH_2 Expected: 4670.7 (M+H)

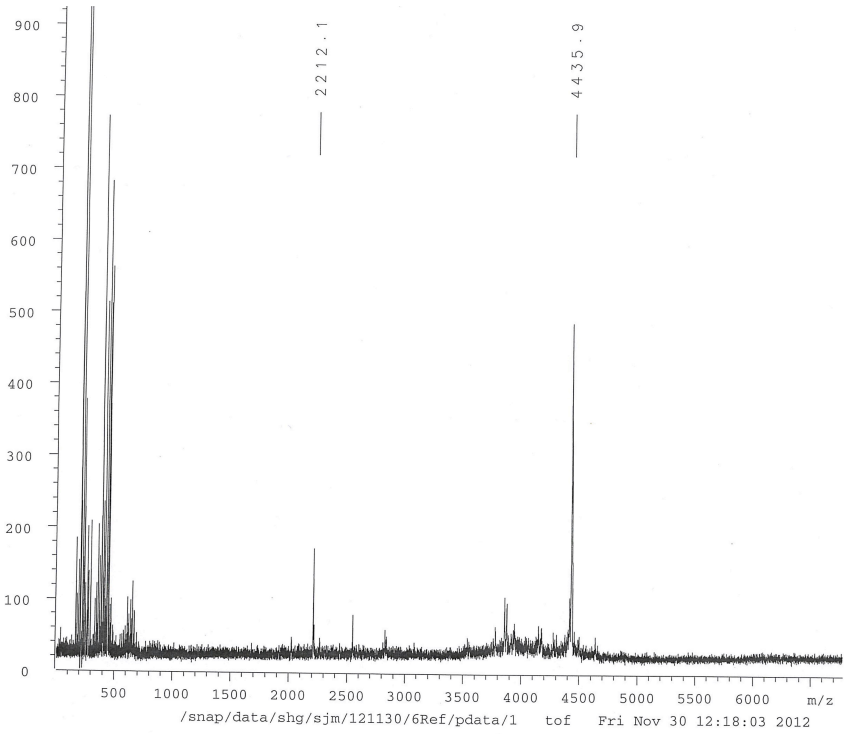


Figure C2.19. MALDI-MS of peptide 4.19 in CHCA matrix. Sequence: Ac-YGGRTSRLKHERTALDKKIEELEKQLELLEDEVSQLQ-NH_2 Expected:4435.3 (M+H)

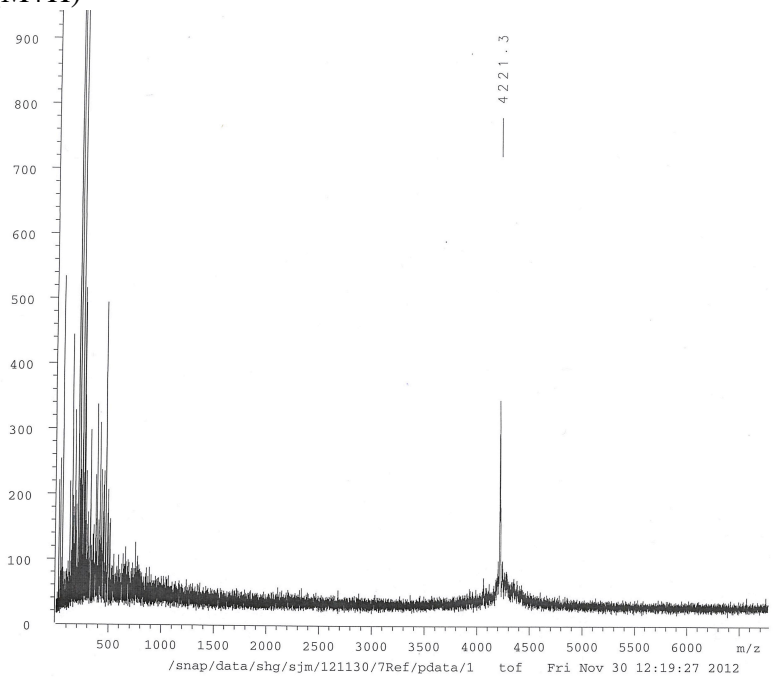


Figure C2.20. MALDI-MS of peptide 4.20 in CHCA matrix. Sequence: Ac-YGGKASRAEHRATALEKKIEELEKRLELAEDKVSQAQ-NH_2 Expected: 4223.2 (M+H)

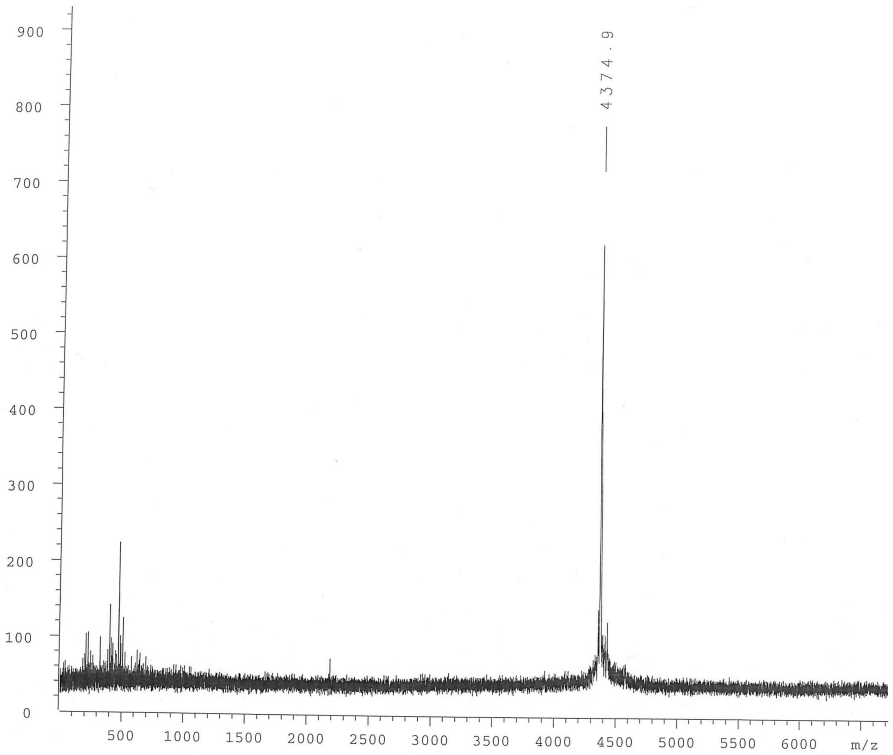


Figure C2.21. MALDI-MS of peptide 4.21 in CHCA matrix. Sequence: Ac-YGGRISRLKHERAALDKKIEELEKQLEALEKEVSQLQ-NH_2 Expected: 4376.3 (M+H)

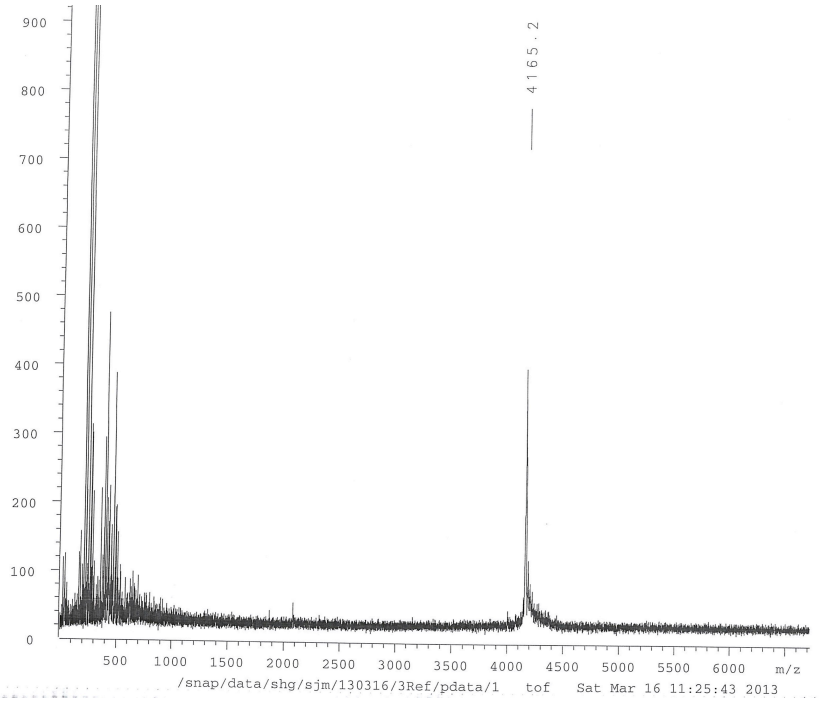


Figure C2.22. MALDI-MS of peptide 4.22 in CHCA matrix. Sequence: Ac-YGGKASRAEHRAAALEKKIEELEKRLEAAEKKVSQAQ-NH_2 Expected: 4164.2 (M+H)

Appendix D: Supporting information for Chapter 5

Supplementary Section D1
HPLC traces

General. For details regarding data acquisition, see the methods section of Chapter 5.

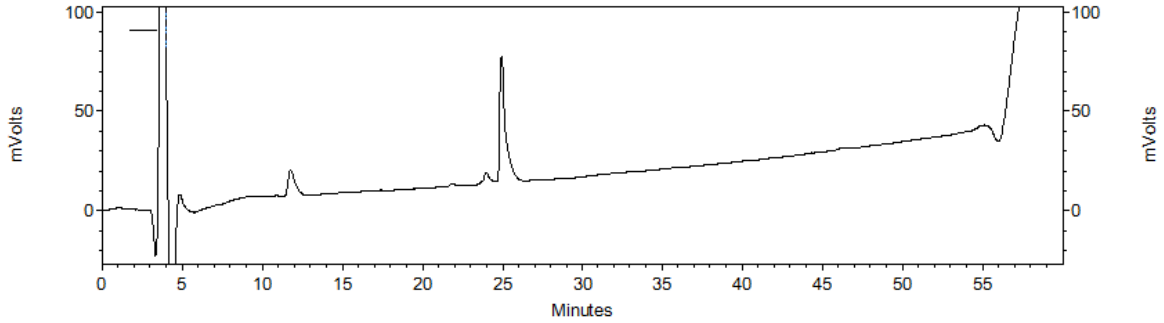


Figure D1.1. Purity check of thioester-containing peptide 5.3. HPLC trace monitored at 220 nm.

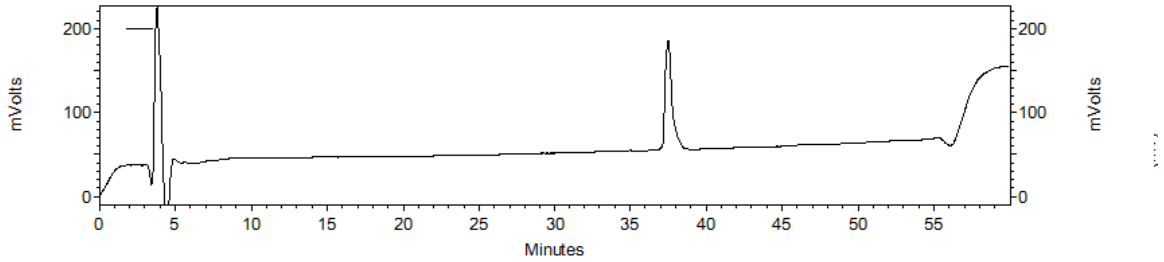


Figure D1.2. Purity check of thiol-containing peptide 5.4. HPLC trace monitored at 220 nm.

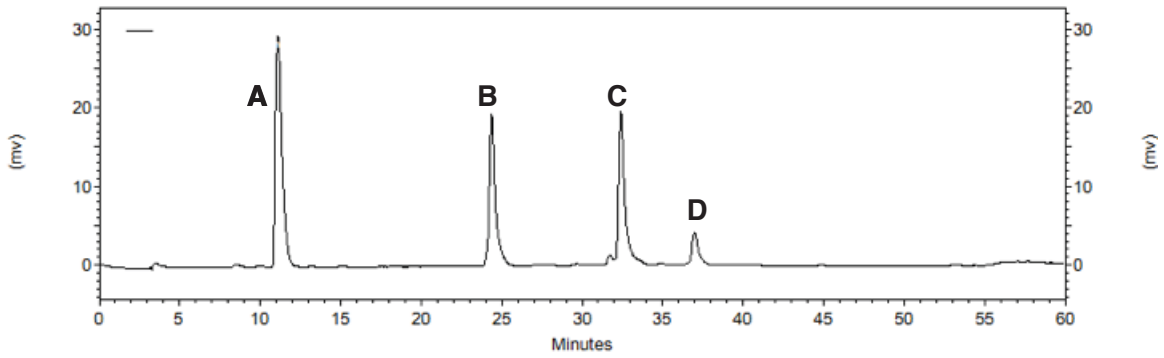


Figure D1.3. BTE assay of sequence 5.2 monitored at 275 nm. Peak A is small thiol 5.0, peak B is full-length depsipeptide 5.2, peak C is thioester depsipeptide 5.3, and peak D is thiol depsipeptide 5.4.

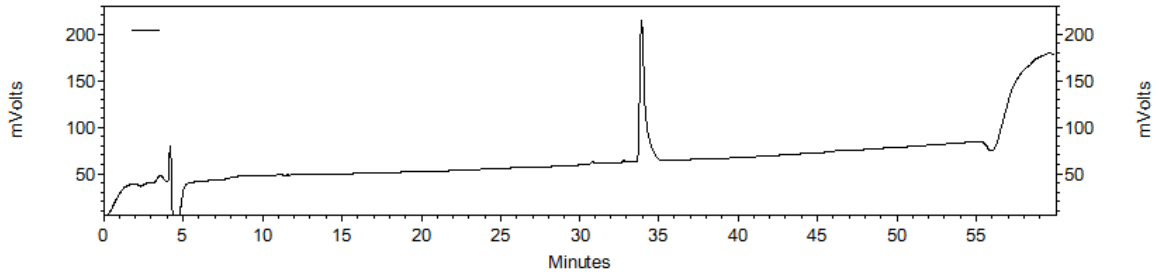


Figure D1.4. Purity check of peptide 5.7 with an amide bond in place of the thioester bond, monitored at 220 nm.

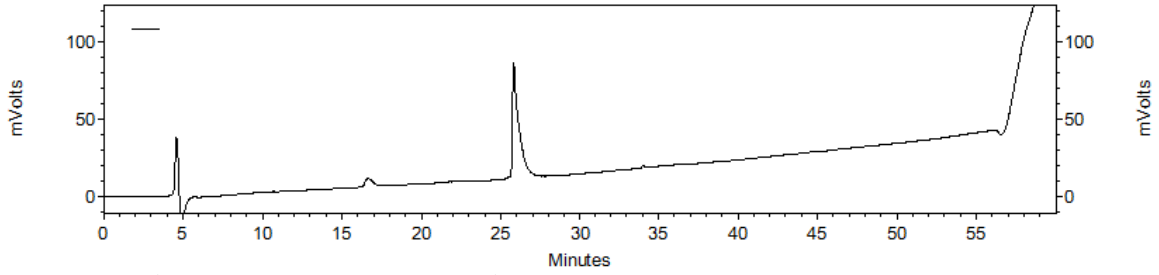


Figure D1.5. Purity check of peptide 5.9 monitored at 220 nm.

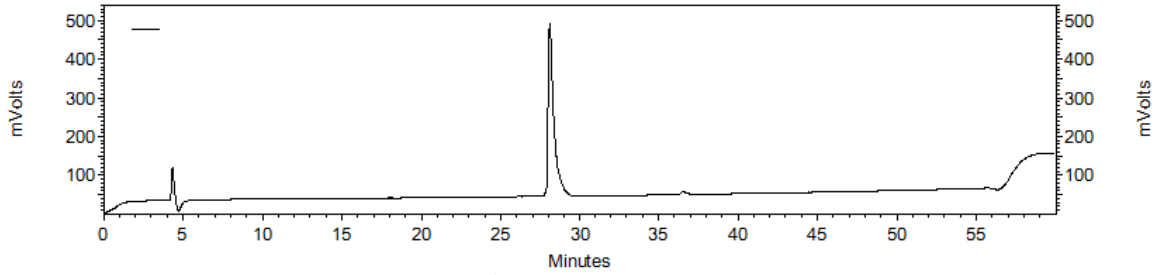


Figure D1.6. Purity check of peptide 5.10, with an amide bond in place of a thioester bond, monitored at 220 nm.

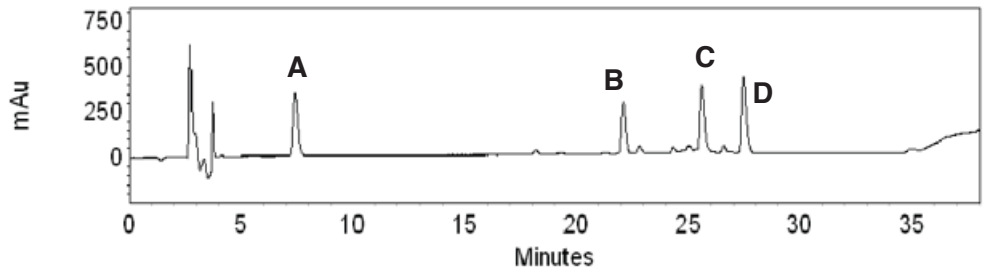


Figure D1.7. BTE assay of 5.7 monitored at 220 nm. Peak A is small thiol 5.0, peak B is thiol depsipeptide 5.9, peak C is thioester depsipeptide 5.8, and peak D is full-length depsipeptide 5.7.

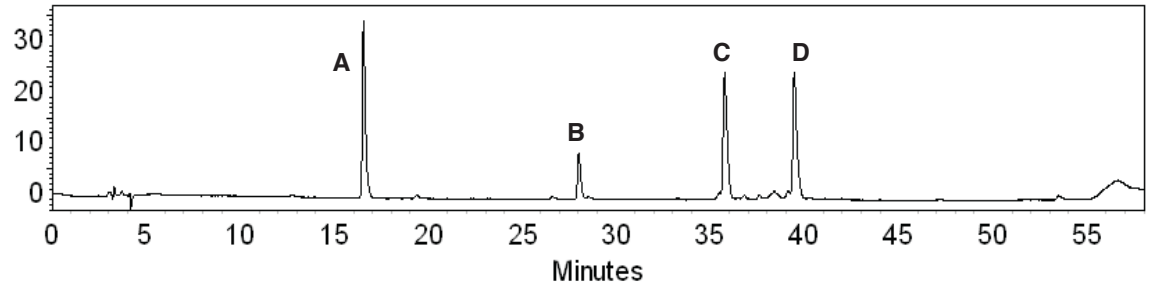


Figure D1.8. BTE assay of sequence 5.11 (X=L, Z=i) monitored at 275 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide.

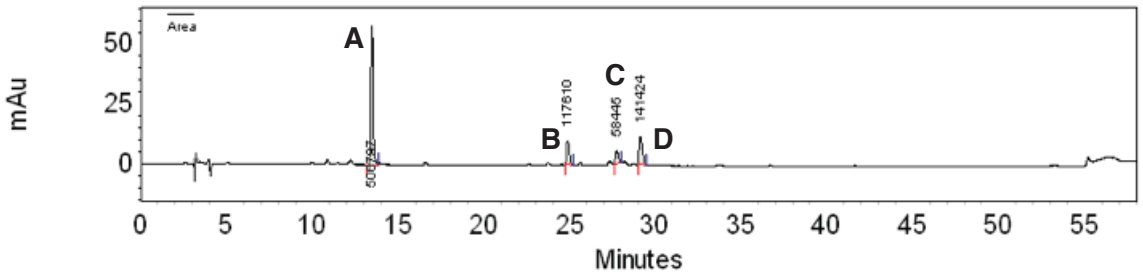


Figure D1.9. BTE assay of sequence 5.11 (X=L, Z=n) monitored at 275 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is C-terminal thiol depsipeptide, and peak D is the full-length thiol depsipeptide.

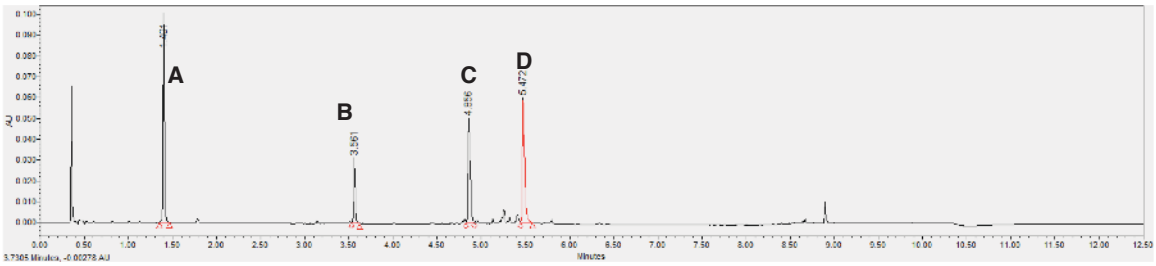


Figure D1.10. BTE assay of sequence 5.11 (X=L, Z=v) monitored at 275 nm. Data were collected on an ultra performance liquid chromatography (UPLC) instrament rather than an HPLC. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide.

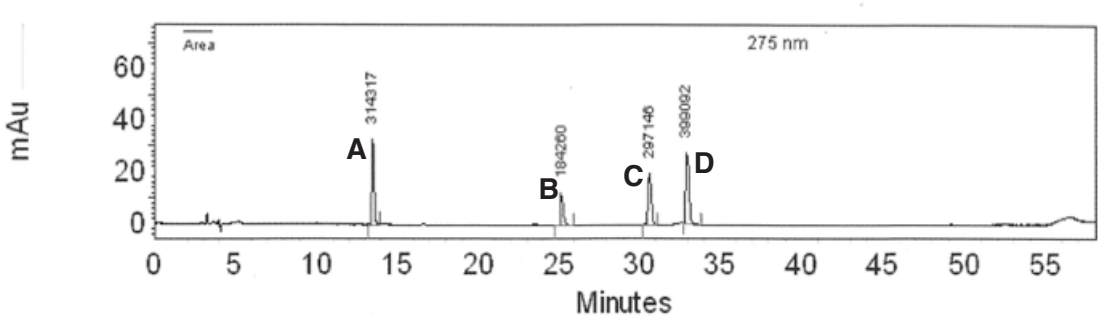


Figure D1.11. BTE assay of sequence 5.12 (X=L, Z=a) monitored at 275 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide.

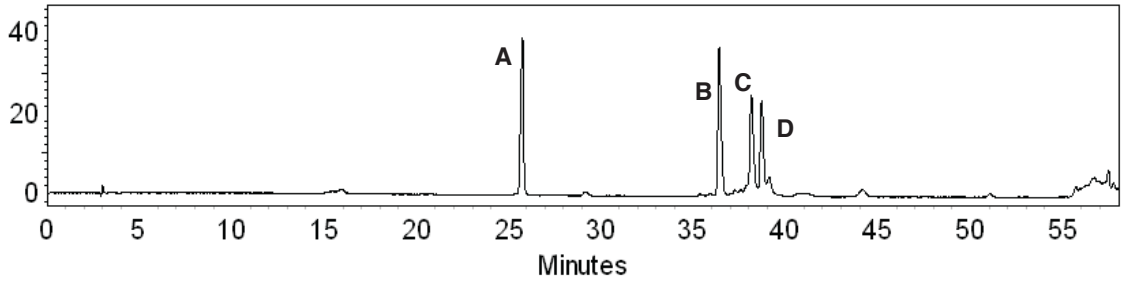


Figure D1.12. BTE assay of sequence 5.12 (X=L, Z=n) monitored at 275 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide. (Note the retention times of the peaks are later than normal due to issues with the HPLC. Integrations of the peaks are not affected and therefore the data is still usable.)

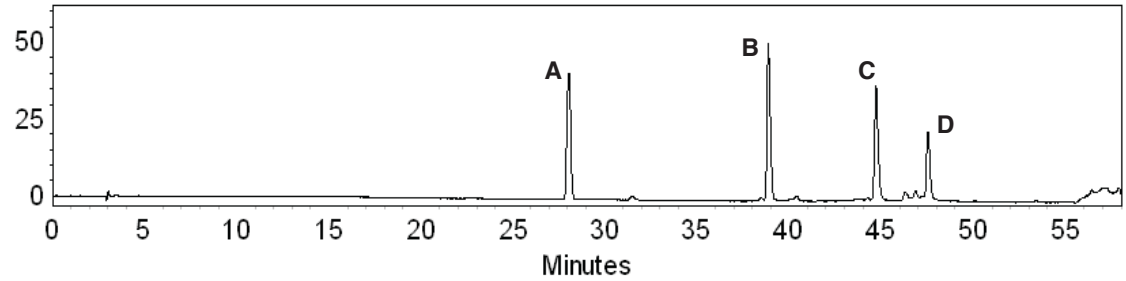


Figure D1.13. BTE assay of sequence 5.12 (X=L, Z=v) monitored at 275 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide. (Note the retention times of the peaks are later than normal due to issues with the HPLC. Integrations of the peaks are not affected and therefore the data is still usable.)

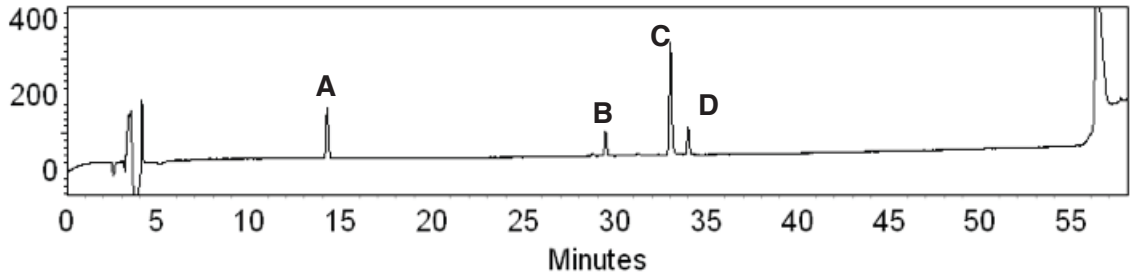


Figure D1.14. BTE assay of sequence 5.13 monitored at 220 nm. Peak A is small thiol 5.0, peak B is the N-terminal thioester depsipeptide, peak C is the full-length depsipeptide, and peak D is the C-terminal thiol depsipeptide. Note that two small peaks are present in the 275 nm trace between peak B and C. These peaks are present when no sample is injected into the HPLC and thought to be background noise. They are not present in the 220 nm trace where amide bonds absorb and therefore not peptidic.

Supplementary Section D2
MALDI data for peptides

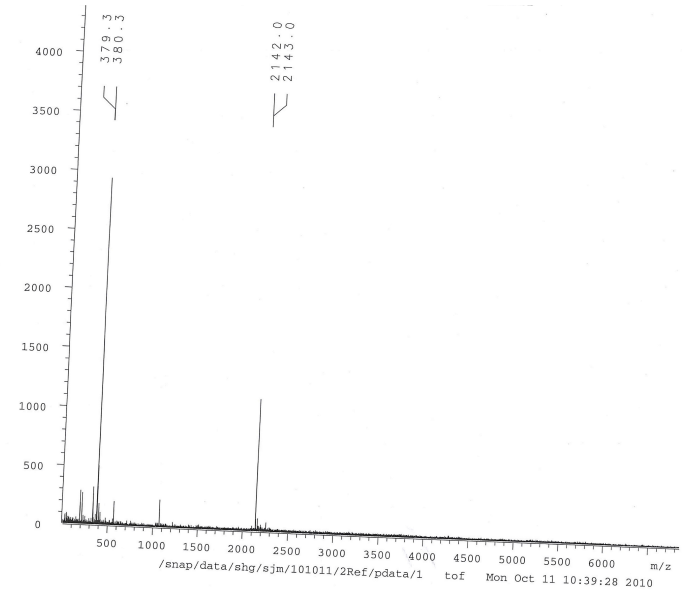


Figure D2.1. MALDI-MS of thioester peptide 5.3. Sequence is as shown in Figure 5.2, except R^1=Et. Expected: 2132.2 m/z (M+1). (Small thiol 5.0 was added to this peptide in BTE buffer to yield the thioester peptide used for BTE assays.)

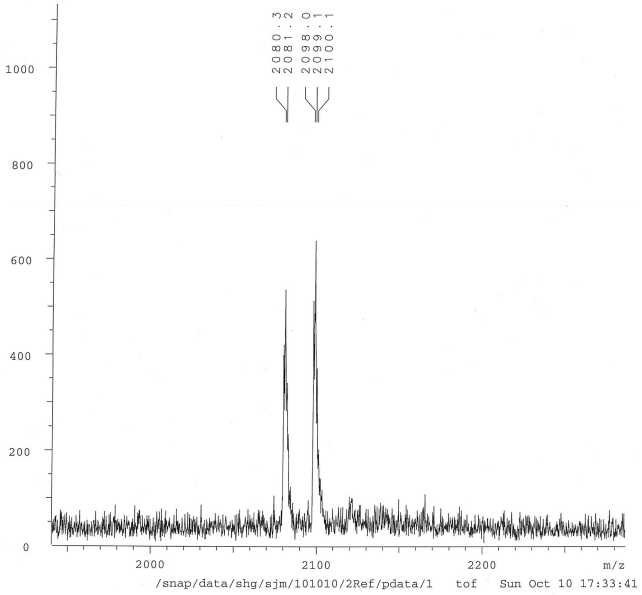


Figure D2.2. MALDI-MS of precipitate formed during a thioester degredation study with peptide 5.3 (See main text Figure 5.4c). Expected: hydrolyzed thioester 2098.2 m/z (M+1). Intramolecular acyl transfer 2080.2 m/z (M+1).

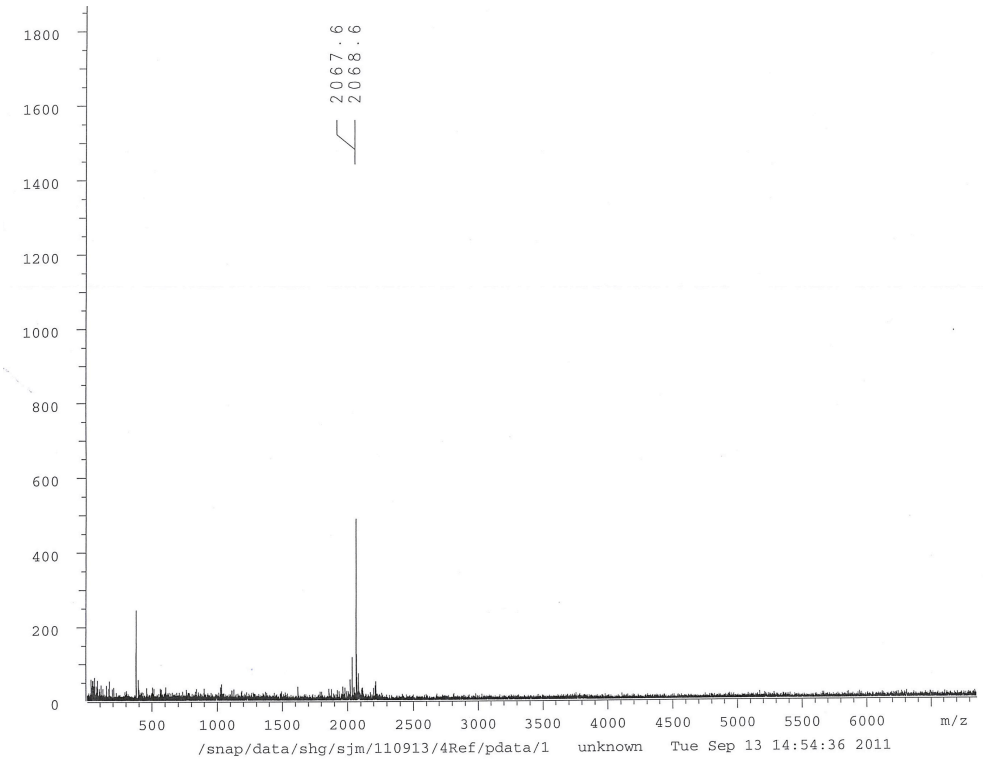


Figure D2.3. MALDI-MS of Peptide 5.4. Expected: 2068.0 m/z (M+1).

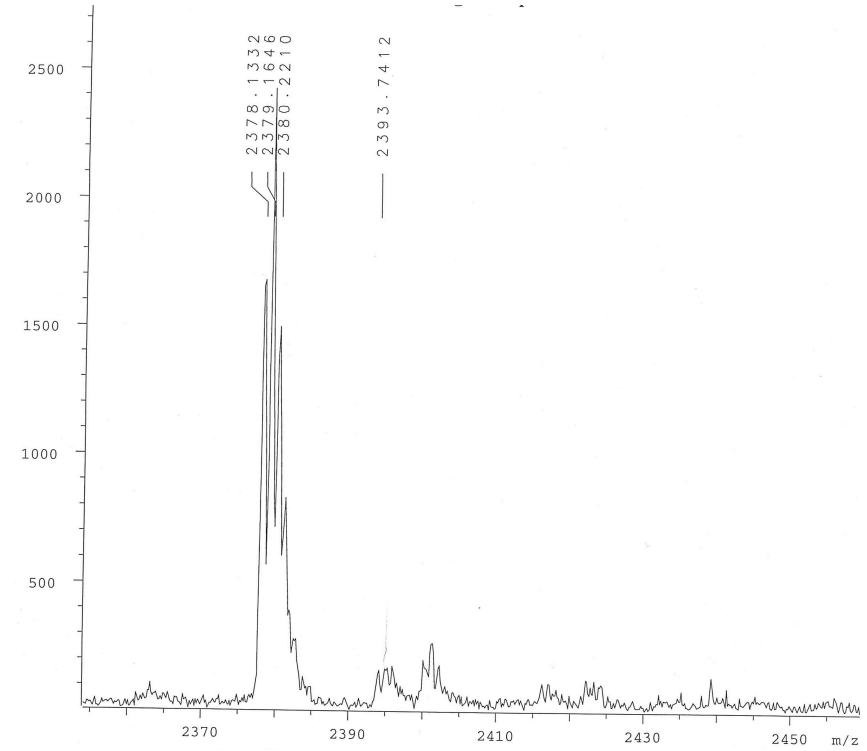


Figure D2.4. MALDI-MS of Peptide 5.7. Expected: 2378.1 m/z (M+1).

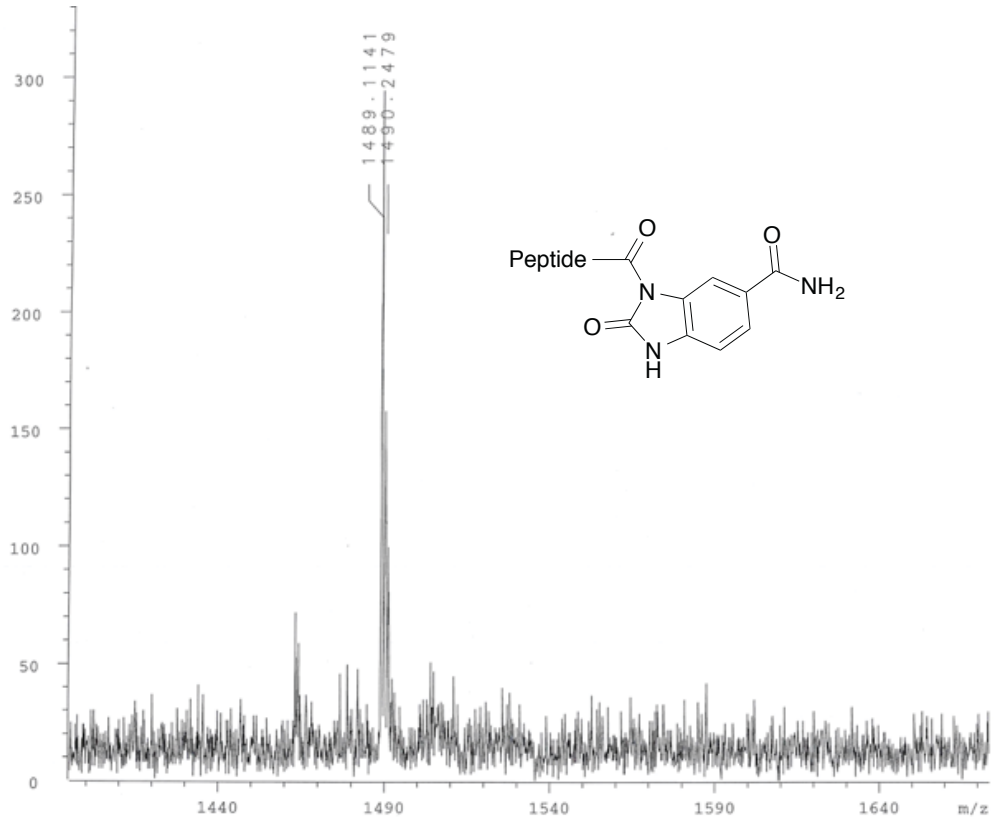


Figure D2.5. MALDI-MS of the Dawson adduct of Peptide 5.8. Expected: 1488.7 m/z (M+1). The Dawson adduct is the precursor to the thioester-containing peptide 5.8.

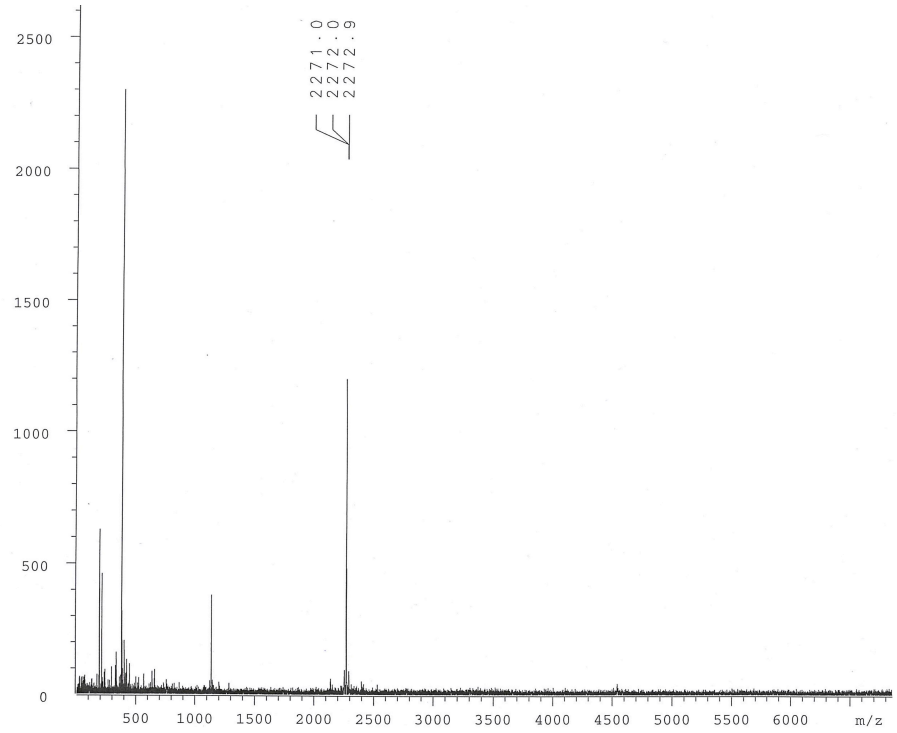


Figure D2.6. MALDI-MS of Peptide 5.10. Expected: 2271.2 m/z (M+1).

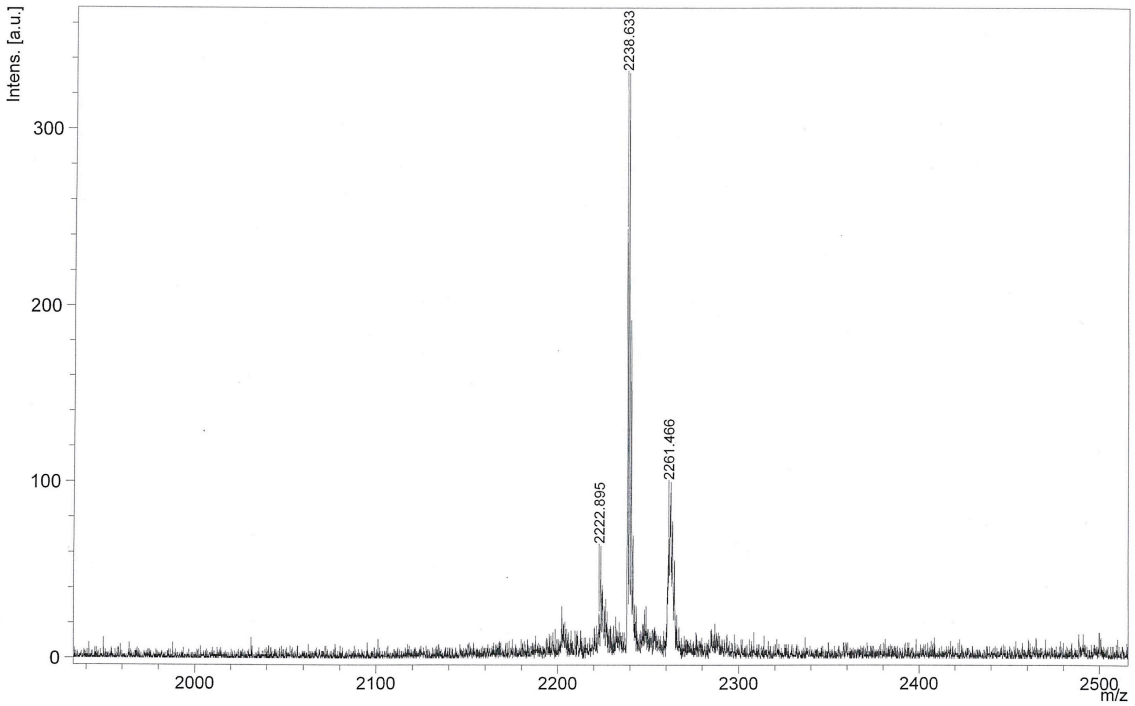


Figure D2.7. MALDI-MS of Succ-EELRRRLEELRRRRGGG-SR[1] where R[1] is the Tyr derived small thiol and Succ is succinic acid. Note that this thioester sample is hydrolyzed by the laser during the ionization. Therefore, we see the mass of the C-terminal carboxylic acid peptide rather than the C-terminal thioester peptide. (Due to the type of laser in the Bruker Ultraflex III MALDI, thioester bonds are often not stable during the ionization process. They typically degrade to give carboxylic acids.) The breadth of the MALDI peaks compared to other MALDI spectra acquired on the same instrument is evidence that the thioester is hydrolyzed during ionization rather than the original peptide sample has been degraded. Further evidence that the correct product has been obtained is the fact that the peptide has an absorbance in the HPLC trace monitored at 275 nm as the hydrolyzed peptide would not absorb at this wavelength. Expected (hydrolyzed peptide): 2238.2 m/z (M+1).

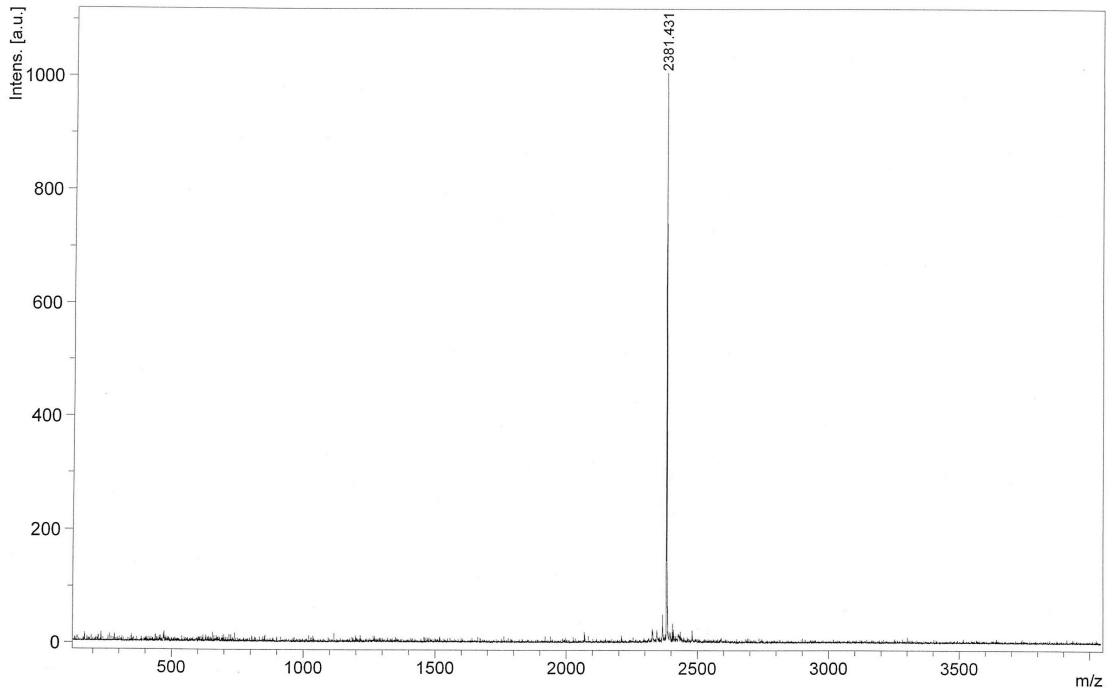


Figure D2.8. MALDI-MS of Ac-IEERIRRIRAEIEERRGGG-SR[1] where R[1] is an ethyl group. Expected: 2381.3 m/z (M+1)

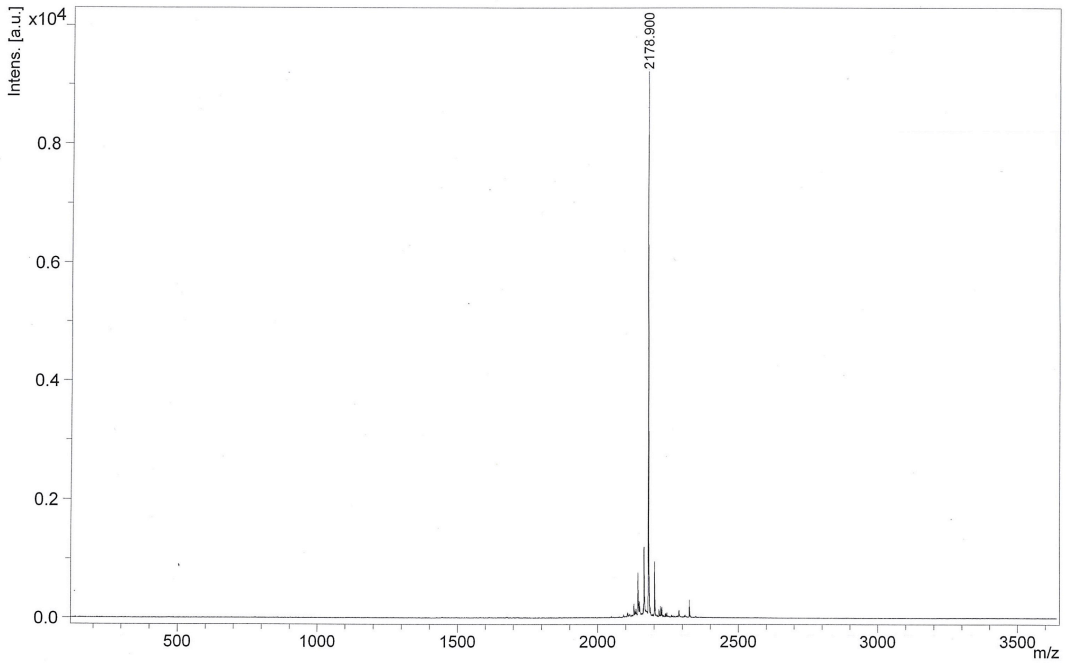


Figure D2.9. MALDI-MS of HS-GGGleeelrraeeelrry-NH_2. The lowercase letters are D-amino acids residues. Expected: 2178.0 m/z (M+1).

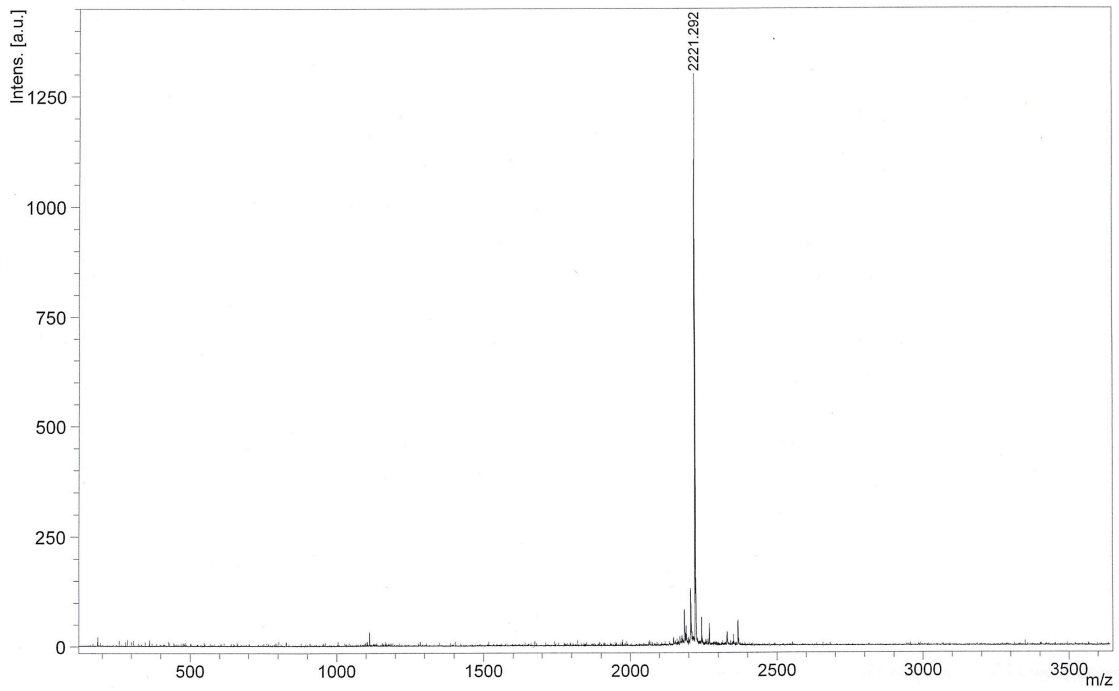


Figure D2.10. MALDI-MS of HS-GGGleeelrrieeelrry-NH_2. The lowercase letters are D-amino acids residues. Expected: 2220.1 m/z (M+1).

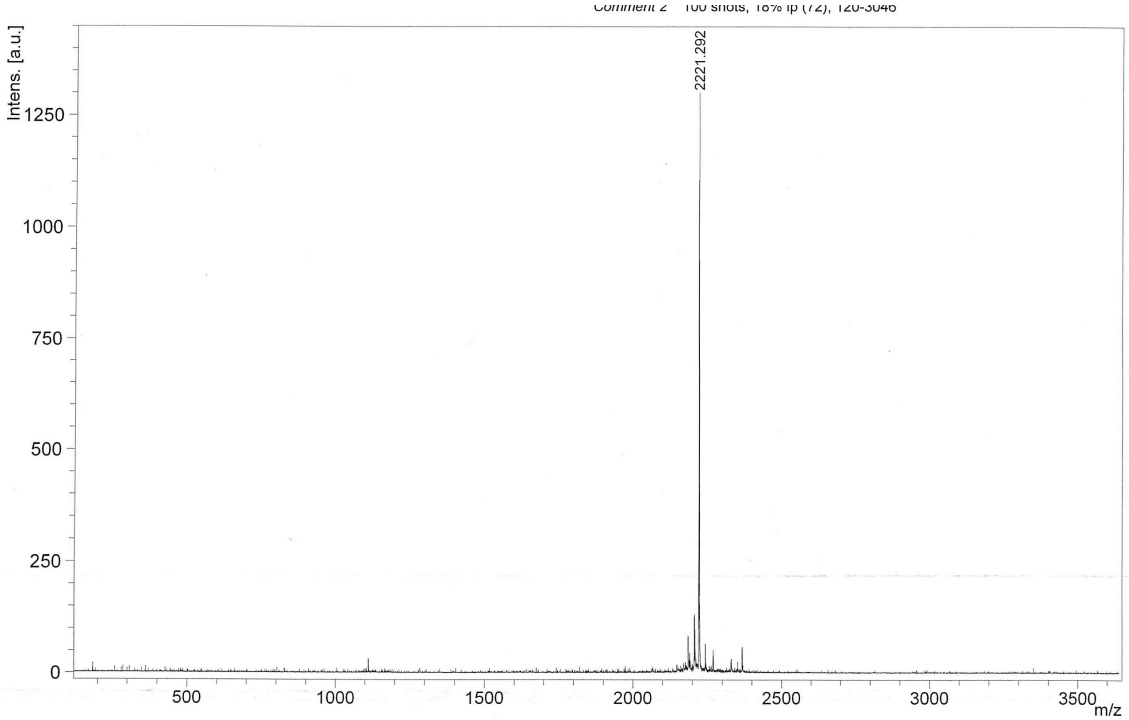


Figure D2.11. MALDI-MS of HS-GGGleeelrrleeelrry-NH_2. The lowercase letters are D-amino acids residues. Expected: 2220.1 m/z (M+1)

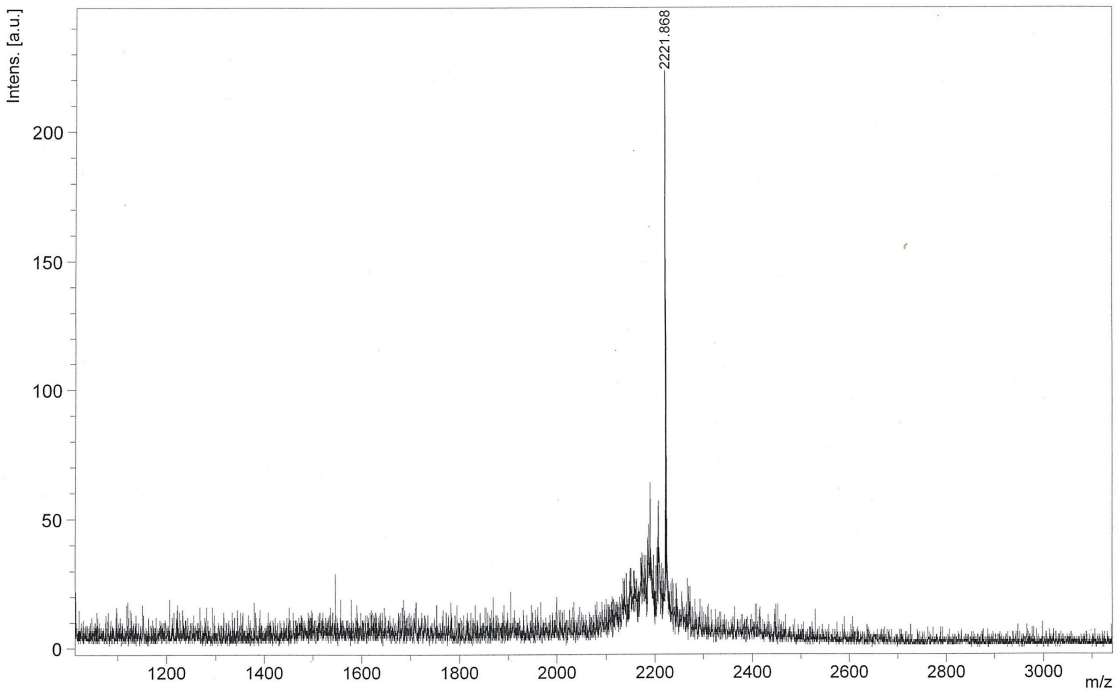


Figure D2.12. MALDI-MS of HS-GGGleeenrrneeelrry-NH_2. The lowercase letters are D-amino acids residues. Expected: 2221.0 m/z (M+1)

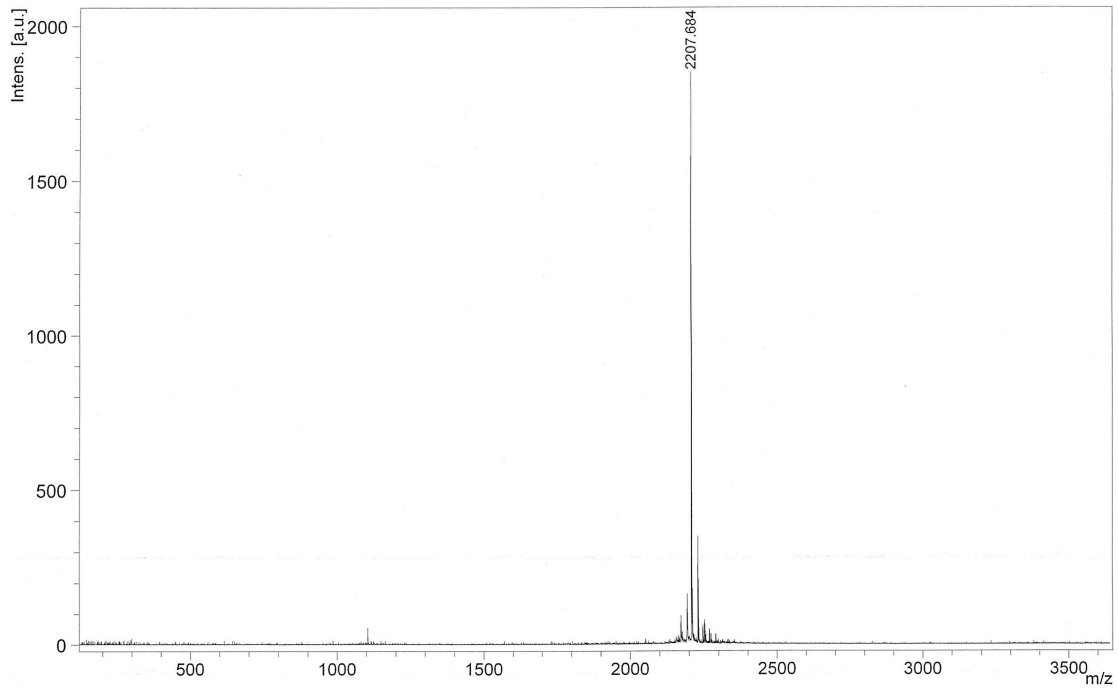


Figure D2.13. MALDI-MS of HS-GGGleeelrrveeelrry-NH_2. The lowercase letters are D-amino acids residues. Expected: 2206.1 m/z (M+1)

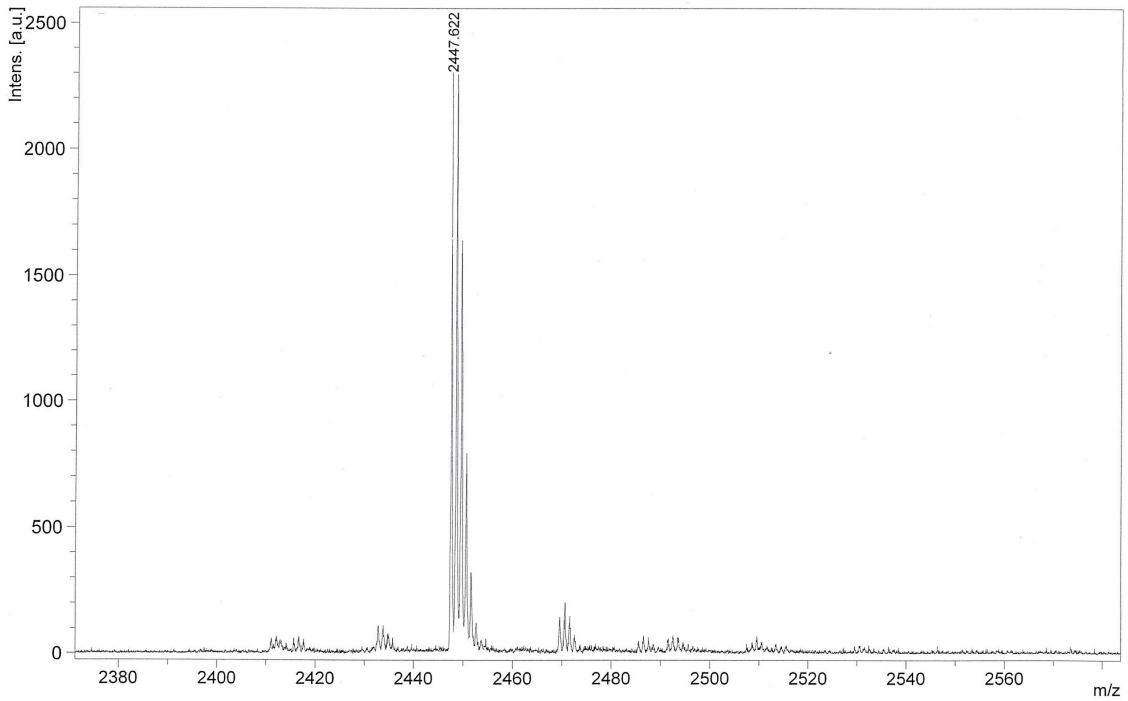


Figure D2.14. MALDI-MS of HS-GGGirreieeeiearirreey-NH_2. The lowercase letters are D-amino acids residues. Expected: 2447.2 m/z (M+1).

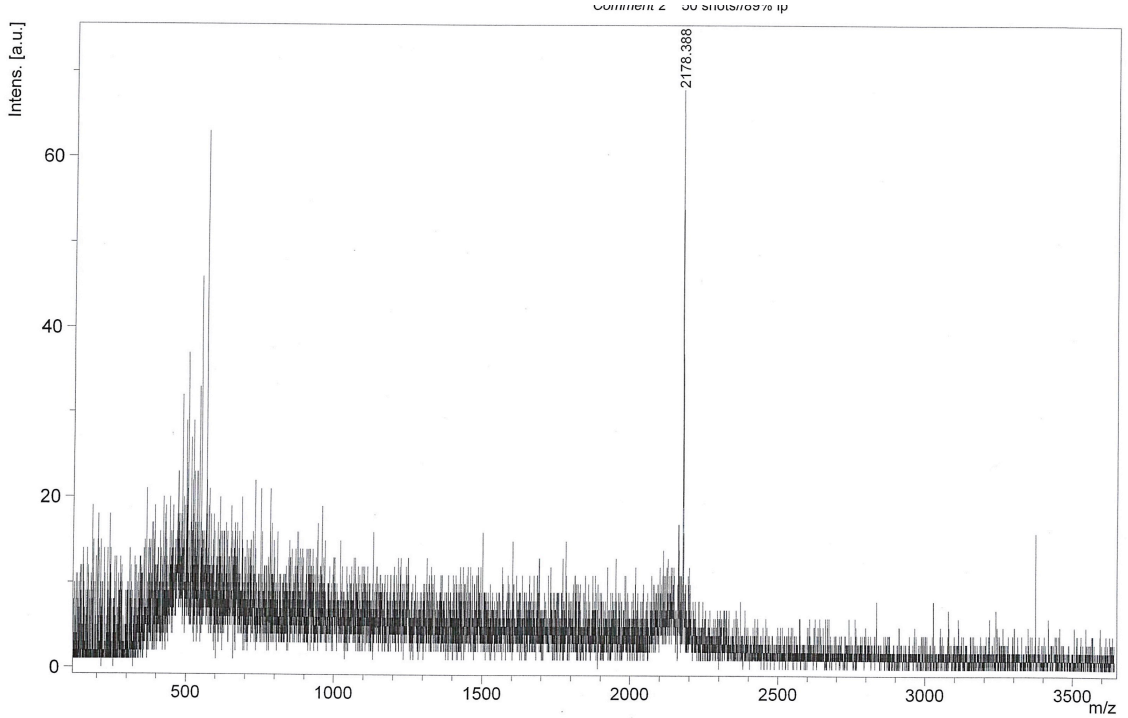


Figure D2.15. MALDI-MS of HS-GGGlrrleeearrleeey. The lowercase letters are D-amino acids residues. Expected: 2178.0 m/z (M+1).

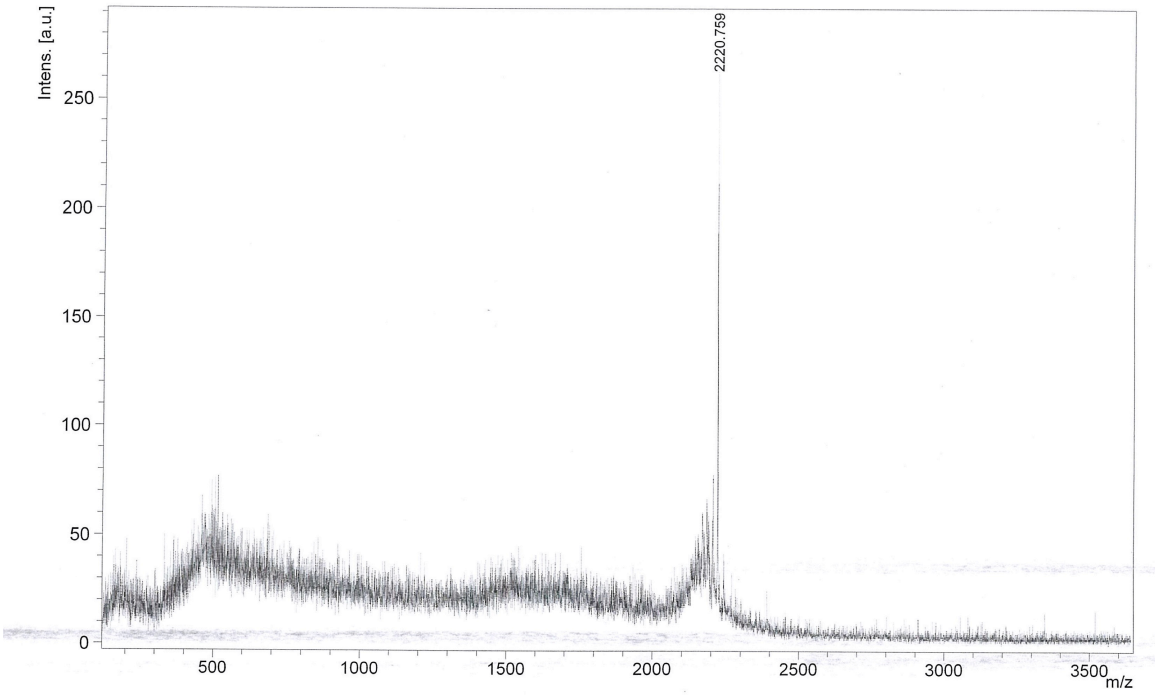


Figure D2.16. MALDI-MS of HS-GGGlrrleeeirrleeey. The lowercase letters are D-amino acids residues. Expected: 2220.1 m/z (M+1).

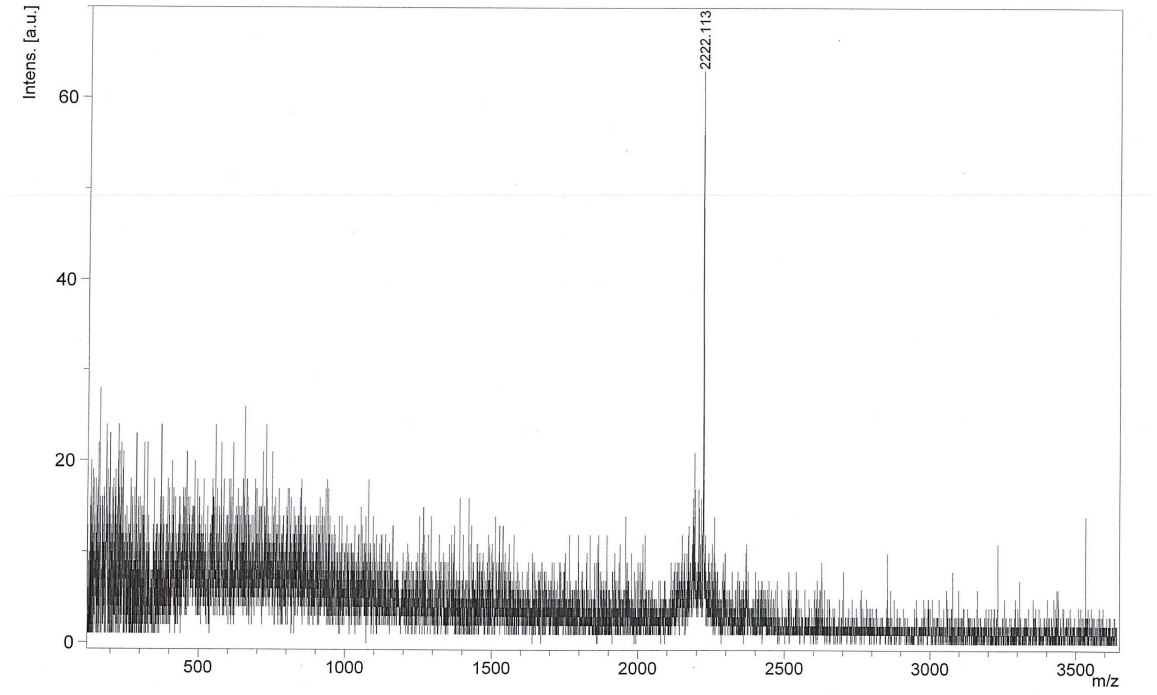


Figure D2.17. MALDI-MS of HS-GGGlrrleeedrrleeey. The lowercase letters are D-amino acids residues. Expected: 2221.0 (M+1).