Structural and Functional Mimicry of Disulfide-Rich Peptides

by

Chino Cabasa Cabalteja

B.S. Microbiology, University of Hawaii at Manoa, 2008

M.S. Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, 2014

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Chino Cabasa Cabalteja

It was defended on

September 10, 2020

and approved by

Rieko Ishima, Associate Professor, University of Pittsburgh Department of Structural Biology

Kabirul Islam, Assistant Professor, University of Pittsburgh Department of Chemistry

Renã Robinson, Associate Professor, Vanderbilt University Department of Chemistry

Thesis Advisor: W. Seth Horne, Associate Professor, University of Pittsburgh Department of Chemistry

Structural and Functional Mimicry of Disulfide-Rich Peptides

Chino Cabasa Cabalteja, PhD

University of Pittsburgh, 2020

Disulfide-rich peptides have promising applications as bioactive agents, but their utility is attenuated by synthetic challenges and various pharmacological drawbacks including low proteolytic stability. Some of these limitations may be addressed by foldamers – sequence specific oligomers with a propensity to fold into discrete conformations. One established approach to the design of foldamer mimics of peptides with complex folds involves incorporation of unnatural building blocks at strategic sites within the natural sequence. Applied judiciously, these modifications can result in peptide variants with a heterogeneous backbone which can adopt complex native-like folds. These proteomimetics can also display augmented pharmacological properties including improved proteolytic resistance. This study demonstrates the application of this heterogeneous backbone substitution strategy to disulfide-rich peptides in the context of structural mimicry of a computationally designed miniprotein and explores the consequences of peptide backbone modification on the function of a bioactive peptide toxin. Lastly, this work delves into the experimental characterization of a series of model tripeptides for validation of a newly developed force field that can simulate artificial protein-like backbones.

Table of Contents

Preface xviii

1.0 Introduction 1

1.1 Small Molecule Drugs, Protein Biologics and Peptides Therapeutics 3

1.1.1 Overview of Peptide Chemical Modifications 7

1.1.1.1 Side Chain and Terminal Modifications 8

1.1.1.2 Peptide and Protein Backbone Modification 11

1.1.1.3 Cyclization 14

1.2 Disulfide-Rich Peptides 17

1.2.1 Bioactive Disulfide Rich Peptides 19

1.2.2 Disulfide-Rich Peptides as Scaffolds for Drug Development 22

2.0 Heterogeneous-Backbone Foldamer Mimics of a Computationally Designed, Disulfide-Rich Miniprotein 27

2.1 Introduction 27

2.2 Results and Discussion 31

2.2.1 Design and Synthesis of Turn Variants 31

2.2.2 Structural Analysis 33

2.2.3 Design and Synthesis of Variants with Modified Helix, Turns, and Hairpin 37

2.2.4 Structural Analysis of Variants with Modified Helix, Turns, and Hairpin 41

2.2.5 Proteolytic Susceptibility 43

2.3 Conclusion 47

2.4 Experimental Section 48

2.4.1 Peptide Synthesis and Cleavage 48

2.4.2 Oxidative Folding 49

2.4.3 Circular Dichroism Spectroscopy 49

2.4.4 Oxidative Folding Kinetics Analysis 50

2.4.5 Iodine Oxidation of $1a_{ox}$ and $1b_{ox}$ 50

2.4.6 NMR Data Acquisition, Analysis, and Structure Determination 51

2.4.7 Proteolytic Degradation by Trypsin 52

3.0 Heterogeneous Backbone Modification and Functional Mimicry of Lasiocepsin 54

3.1 Introduction 54

3.2 Results and Discussion 58

3.2.1 Las NH_2 and Design of First-Generation Las Analogs 58

3.2.2 Synthesis and Regioselective Disulfide Formation 59

3.2.3 Microbiological Assay of First-Generation Lasiocepsin Variants 60

3.2.4 NMR of Las Var 1 61

3.2.5 Design of Second-Generation Las Variants 64

3.2.6 Design of Second-Generation Las Variants 65

3.2.7 Microbiological Assays of Second-Generation Las Variants 66

3.2.8 Hemolytic Assay of Las Variants 67

3.2.9 Cytotoxicity Assay of Las Variants 68

3.2.10 NMR of Las Var 9 and 11 70

3.2.11 Proteolysis of Second-Generation Las Variants 73

3.2.12 Third Generation Lasiocepsin Variant 74

3.2.12.1 Las Var 12 Proteolytic Degradation 75

3.2.12.2 Microbiological Assay of Las Var 12 76

3.2.12.3 Hemolysis Assay of Las Var 12..... 77

3.2.12.4 Cytotoxicity of Las Var 12 78

3.2.12.5 NMR of Las Var 12..... 79

3.3 Conclusions and Future Directions..... 80

3.4 Methods 83

3.4.1 Solid Phase Synthesis and Cleavage 83

3.4.2 Regioselective Disulfide-Bond Formation.....84

3.4.3 NMR of Las Variants.....84

3.4.4 Proteolytic Degradation.....85

3.4.5 ESKAPE Pathogen Assay.....86

3.4.6 Hemolytic Assay of Mouse RBC 86

3.4.7 Cytotoxicity Assay Against RAW264.7 Cells 87

4.0 Experimental Validation of a Force Field for Simulation of Artificial Protein-Like Backbones..... 88

4.1 Introduction 88

4.1.1 Molecular Dynamics Simulations and Peptide Folding.....89

4.1.2 Design of Tripeptides for Force Field Validation.....91

4.2 Results and Discussion 92

4.3 Conclusions 95

4.4 Methods 97

4.4.1 Peptide Synthesis and Purification.....97

4.4.2 NMR and Data Acquisition for Peptides 1-4 98

Appendix A Supplementary Information for Section 2.0 99

Appendix B Supplementary Information for Section 3.0 103

Appendix C Supplementary Section for Section 4.0 111

Bibliography 112

List of Tables

Table 1. MALDI-TOF MS data for peptides 1-6 32

Table 2. Minimum inhibitory concentration of first-generation Las variants againts clinically-relevant ESKAPE pathogens..... 61

Table 3. Minimum inhibitory concentration of first and second-generation against clinically-relevant ESKAPE pathogens 66

Table 4. Direct comparison of Las var 12 and Las NH_2 against clinically-relevant ESKAPE pathogens 76

Table 5. ESI-MS data for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.. 92

Table 6. $^{3}J_{\text{HN-H}\alpha}$ coupling constants for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing. 95

Table 7. Comparison of experimental and calculated J-coupling constants for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein

mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing. 96

Table 8. MALDI-TOF MS data for peptides 1a and 1b 99

Table 9. Statistics for NMR structure calculations for peptides 3-6 99

Table 10. Observed m/z and putative sequences of peptide 1, 5, and 6 digests 101

Table 11. MALDI-TOF MS data of fully oxidized Las variants investigated in this work 104

Table 12. Chemical shifts involved in backbone walk of Las var 9 109

Table 13. Chemical shifts involved in backbone walk of Las var 11 110

List of Figures

Figure 1. Small molecule drugs with varied function. From Drews, J., Drug discovery: a historical perspective. Science 2000, 287 (5460), 1960-4. Repinted with permission from AAAS..... 4

Figure 2. The number of FDA approved biologics from the year 2000 to 2019. Adapted from de la Torre, B. G.; Albericio, F., The Pharmaceutical Industry in 2019. An Analysis of FDA Drug Approvals from the Perspective of Molecules. Molecules 2020, 25, (3). 6

Figure 3. Examples of common PEG conjugation performed on proteins and peptides. Adapted from Lawrence, P. B.; Price, J. L., How PEGylation influences protein conformational stability. Current Opinion in Chemical Biology 2016, 34, 88-94..... 9

Figure 4. Design considerations in the mimicry of peptide secondary structure through peptide backbone modification. Adapted from George, K. L.; Horne, W. S., Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. Accounts of Chemical Research 2018, 51 (5), 1220-1228..... 13

Figure 5. Schematic of the synthesis of cyclic kalata B1. Adapted from Cheneval, O.; Schroeder, C. I.; Durek, T.; Walsh, P.; Huang, Y.-H.; Liras, S.; Price, D. A.; Craik, D. J., Fmoc-Based Synthesis of Disulfide-Rich Cyclic Peptides. The Journal of Organic Chemistry 2014, 79 (12), 5538-5544..... 15

Figure 6. Two cysteines oxidize for form a disufide-bonded cystine unit..... 17

Figure 7. The three types of cysteine knots motifs all exhibit the same cysteine connectivity. Reproduced with permission from Iyer, S.; Acharya, K. R., Tying the knot: The cystine

signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. The FEBS Journal 2011, 278 (22), 4304-4322. 19

Figure 8. Various bioactive disulfide-rich peptides and respective organism of origin. Clockwise from green sector: sunflower trypsin inhibitor-1, θ-defensin-1, α-conotoxin Vc1.1, kalata B1 and chlorotoxin. Reproduced with permission from Wang, C. K.; Craik, D. J., Designing macrocyclic disulfide-rich peptides for biotechnological applications. Nature Chemical Biology 2018, 14 (5), 417-427. 20

Figure 9. Three structural isomers with different disulfide connectivity can form from the statistical oxidation of a linear unoxidozied precursor peptide with 4 cysteine residues. 25

Figure 10. 1 Sequence, secondary structure map, and NMR structure (PDB 5KVN) for peptide 1 and sequences of backbone modified variants 2-6. The identity of "R" groups, when present, match the side chain specified by the corresponding α-residue denoted by the single letter code..... 30

Figure 11. HPLC chromatograms showing the time course of the oxidative folding reactions to form peptides 1, 5, and 6 from the corresponding fully-reduced linear precursors. Conditions: 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3; int = single disulfide intermediate, prod = fully oxidized product. 33

Figure 12. CD spectra of 1-6 at 60 μM concentration in 10 mM phosphate, pH 7. 34

Figure 13. Summary of NMR data (chemical shift index values, vicinal HαHN coupling constants, and NOE correlations indicative of fold) for peptides 1-4. A putative secondary structure map is provided for each sequence based on the analysis of the aggregate data. 35

Figure 14. Comparison of the overall folding pattern and helix-to-hairpin turn from NMR structures of peptides 1 (PDB 5KVN), 3 (PDB 6E5H), and 4 (PDB 6E5I). Spheres in the cartoon representation indicate positions of a Cys residue or artificial backbone unit; residues are colored by type, matching the scheme in Fig. 10. 36

Figure 15. Time course of the oxidative folding to form 1, 5, and 6 from the corresponding fully reduced linear precursors. Reactions were held with 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3. 38

Figure 16. Scheme showing the preparation of 1 through sequential chemoselective disulfide bond formation from protected precursors 1a and 1b (top) as well as random aerobic oxidation of fully reduced precursor 1red (bottom). Key HPLC retention time changes in product vs. starting material for each reaction that informed the assignment of the structure of single-disulfide intermediate 1int are shown. Full HPLC chromatograms can be found in appendix A..... 39

Figure 17. (A) Comparison of the overall folding pattern (center) and zoomed views of identical sites in the helix (left) and hairpin (right) from the NMR structures of peptides 1 (PDB 5KVN), 5 (PDB 6E5J), and 6 (PDB 6E5K). Spheres in the cartoon representation indicate positions of a Cys residue or artificial backbone unit. (B) Packing of Tyr23 in the structure of 1 and 6. Residues in both panels are colored by type, matching the scheme in Fig. 10. 42

Figure 18. Time courses for proteolytic degradation of peptides 1, 5, and 6 in the presence of trypsin; each reaction was held with 100 μM peptide with 2.8 mg/mL enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5..... 44

Figure 19. Observed trypsin cleavage sites for peptides 1, 5 and 6. 45

Figure 20. Proposed models for AMP activity. Adapted from Mahlapuu, M.; Håkansson, J.; Ringstad, L.; Björn, C., Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. Frontiers in Cellular and Infection Microbiology 2016, 6, 194. 55

Figure 21. Structure of native lasiocepsin (PDB 2MBD). Disulfide bonds are depicted as spheres. 57

Figure 22. Sequences of first generation Lasiocepsin variants. 59

Figure 23. Synthetic scheme of Las variants. 60

Figure 24. Chemical shift deviation plot from random coil values of native Lasiocepsin as derived from published BMRB chemical shifts. 63

Figure 25. Hα chemical shift deviation plot of Las var 1. 63

Figure 26. Sequence of generation 1 and 2 Las variants. 64

Figure 27. Hemolysis of Lasiocepsin variants at various concentrations. 68

Figure 28. Cytotoxicity of first and second generation Lasiocepsin variants at 4- and 24-hour incubations. 69

Figure 29. Hα chemical deviation plot of Las var 9. 71

Figure 30. Hα chemical deviation plot of Las var 11. 72

Figure 31. Time courses for proteolytic degradation of Las NH_2, Las var 9, and Las var 11 in the presence of trypsin; each reaction contained 250 μM peptide with 120 nM enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5. 73

Figure 32. Sequence of Las var 12 as compared to Las var 11 and native Lasiocepsin. Las var 12 represents the third generation of Lasiocepsin variants and is modified from Las var 11. 75

Figure 33. Time courses for proteolytic degradation of Las var 11 and Las var 12 in the presence of trypsin; each reaction contained 250 μM peptide with 120 nM enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5. 76

Figure 34. Direct comparison of Las var 12 hemolysis against Las NH_2 and select Lasiocepsin analogs..... 78

Figure 35. Cytotoxicity of Las var 12 as compared with select Lasiocepsin variants. 79

Figure 36. Sequence and structure of synthesized tripeptides for force field validation. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing. 92

Figure 37. 1D NMR of amide NH region for various concentrations of peptide 1. A consistent lineshape and chemical shift was observed, indicating that the samples were not aggregating. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing. 93

Figure 38. Amide region of the ^{1}HNMR of peptides 1-4 showing peak assignments for H_N resonances. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for

protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing. 94

Figure 39. HPLC chromatograms showing the time course of the oxidation reactions to form peptides $1a_{ox}$ and $1b_{ox}$ from the corresponding reduced linear precursors 1a and 1b. Conditions: 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3. 100

Figure 40. HPLC chromatograms showing the time course of the proteolysis of peptides 1, 5, and 6by trypsin. Each reaction was 100 µM peptide and 2.8 mg/mL enzyme in 0.1 M ammonium bicarbonate, pH 8.5. 101

Figure 41. Summary of NMR data (chemical shift index values, vicinal HαHN coupling constants, and NOEs indicative of fold) for peptides 5and 6. A putative secondary structure map is provided for each sequence based on the analysis of the aggregate data. 102

Figure 42. Analytical HPLC traces of fully oxidized and purified Lasiocepsin variants investigated in this study. Las var 1 and 3 were run on a 20-30% Solvent B gradient. Las NH_2, var 2, and var 9 – 12 were run on a 20-25% Solvent B gradient. 103

Figure 43. MALDI-TOF MS spectra of Las NH_2..... 104

Figure 44. MALDI-TOF MS spectra of Las var 1. 105

Figure 45. MALDI-TOF MS spectra of Las var 2. 105

Figure 46. MALDI-TOF MS spectra of Las var 3. 106

Figure 47. MALDI-TOF MS spectra of Las var 9. 106

Figure 48. MALDI-TOF MS spectra of Las var 10. 107

Figure 49. MALDI-TOF MS spectra of Las var 11. 107

Figure 50. MALDI-TOF MS spectra of Las var 12. 108

Figure 51. Analytical HPLC traces for purified peptides 1-4. Gradient was 0-10% solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) over 30 minutes for peptides 2 and 4; the gradient was 0-5% solvent B over 30 minutes for peptide 1 and 3. 111

Preface

I would like to express my gratitude to everyone who has supported me throughout my academic journey. First and foremost, I would like to thank my graduate advisor Dr. W. Seth Horne for his guidance, trust, and encouragement throughout my time as a Ph.D. student. Working in his lab has exposed me to the field of foldamers and proteomimetics, which has sparked my passion for working with synthetic peptides. I would also like to thank my committee members: Dr. Rieko Ishima, Dr. Kabirul Islam, and Dr. Renã Robinson for their guidance and advice in many aspects of my studies.

Getting to this point in my life would not have been possible without the unconditional love and support of my family. I would like to thank my parents Felina and Chito Garduque for their sacrifices in order to provide me a chance at an education. They have backed me in all my pursuits especially in my decision to move to the mainland for school. Mom and Dad, thank you for always encouraging me. To my cousins Cristina and Katherine McLaughlin, mahalo for the care packages and for always checking up on me.

I would also like to thank my lab mates Dr. Haley Grimm, Jackie Santhouse, Shilpa Rao, Shelby Schettler, Tom Harmon, and Jacob Wolfe for the interesting scientific discussions and for making the lab an enjoyable place to work. To Austin Asari, Leila Terrab, Xing Yee Gan, and Serene Tai thanks for all the laughs and good times. You all have made Pittsburgh my home away from home and I am truly blessed to have your friendship.

Lastly, I would like to dedicate this dissertation to my grandmother Narcisa Cabalteja who took care of me and fostered my love for science at a young age. One of my earliest memories in life was of her showing me various plants in her garden and explaining how they can be used as a

form of medicine. Because of this, I became curious about the world around me, which ultimately directed me to a path in science.

1.0 Introduction

The canonical twenty proteinogenic amino acids give rise to peptides and proteins that participate in complex biological cascades that comprise fundamental physiological processes of life. The key to the diversity of proteins lies in how unique amino acid sequences are able to drive interactions to form intricate secondary and tertiary folded structures. These complex folded structures are the foundation of the interactions of proteins and peptides with other molecules in complex biological cascades. The deregulation of these pathways can lead to disease states, making the study of these aberrant interactions an important objective in biomedical research in order to determine the etiology of disease and develop therapeutic agents.[1] Small molecules have been a mainstay in the pharmaceutical field and have served as scaffolds for numerous drugs. However, these molecules can have low selectivity to their drug targets, leading to various side-effects. Peptides have found a resurgence of interest as scaffolds for drug development due to their involvement in these biological cascades and desirable properties including selectivity and potency.[2] However, peptide drugs also suffer various drawbacks that need to be surmounted in order to develop efficacious therapeutic agents.[3] To this end, peptide chemists have sought to apply chemical modifications to peptide structure with the objective of leveraging the pharmacological advantages of peptides and improving upon these drawbacks.[4]

Development of novel peptide-based therapeutics involves overcoming the pharmacological limitations of a given native peptide while maintaining the unique three-dimensional conformations responsible for its selectivity. These limitations may be addressed by foldamers, sequence-specific oligomers with a propensity to fold into discrete conformations.[5] An important aim in foldamer research is to utilize natural and unnatural building blocks to mimic a

diverse array of natural structures. A particularly powerful method involves modifying the peptide backbone through judicious substitution of unnatural amino acids at specific locations in the peptide sequence.[6] This heterogeneous backbone substitution approach has been successfully applied to mimic secondary and tertiary structures of various native peptides, thus creating artificial protein-like agents or "proteomimetics".[7] Importantly, this approach can offer varying degrees of protection against proteases, making it useful in designing peptide analogues with improved pharmacokinetic properties.[8] One area of interest is to broaden the scope of foldamer design principles to mimic diverse prototype protein structures in order to establish a more robust design methodology and access new functions. In this regard, disulfide-rich peptides offer a myriad of unique folds and structures[9] and represent an attractive class to expand the repertoire of potential foldamer scaffolds.

Peptides containing multiple disulfide linkages have diverse roles in biology, ranging from hormones, bactericidal defensins, and venoms.[10-11] Some classes of these peptides display efficient and specific binding to target receptors due to conformationally constrained architecture, while others present broad antimicrobial properties. As such, many of these peptides have been explored as potential pharmaceutical candidates.[12-13] However, synthesis of this peptide class can pose challenges due to complications in formation of the correct disulfide-bond connectivity, which can result in structural isomers that differ in receptor target affinity and bioactivity.[14-15] The importance of accurate disulfide connectivity has been extensively documented in peptide toxins, where many structural isomer products display no *in vivo* function.[14, 16] Substitution of various α-amino acids with post-translationally modified variants display mixed results[17], emphasizing the importance of sequence design in maintaining function. Interestingly, there is a dearth of information regarding

the systematic study of backbone-modifying amino acid mutations of disulfide-rich peptides for the purpose of structural mimicry and functional assessment.

This dissertation describes the advancement and application of heterogeneous backbone substitutions to disulfide-rich peptides, thus expanding the scope of foldamer design to an important new structural class. The central hypothesis underlying this body of work is that previously successful design strategies in mimicking other systems can be applied to recapitulate the structures of disulfide-rich peptides and that these proteomimetic analogs will exhibit bioactivity comparable to their native counterparts.

Although the diversity of folding patterns able to be mimicked by foldamers has continued to expand, through the work of myself and others, there is a lack of molecular insights into how backbone modification impacts dynamics and folding pathways. In an effort to address this gap in knowledge, my lab has undertaken a collaborative project to develop and apply a new force field able to treat artificial backbones in state-of-the art simulations. Separate but complementary to my main focus on disulfide-rich scaffolds, I have prepared and characterized model peptides to provide experimental validation of this newly developed force field. These efforts are also described in this dissertation.

1.1 Small Molecule Drugs, Protein Biologics and Peptides Therapeutics

The field of modern drug development has historically focused on small molecule scaffolds that are derived from or inspired by natural products. Historically, the development of refining methods in coal and tar processing led to an interest in the development of benzene derivatives and aromatic compounds as scaffolds for drug development which have been the foundation of

early antipyretics including aspirin and acetaminophen (paracetamol).[18] The discovery of other small molecules, notably antimicrobials, such as penicillin, macrolides, and quinolones provided an array of additional drug development starting points, and have served as the basis of several generations of broad-range antibiotics that are still in clinical use.[19-20] Moreover, these small molecules have also functioned as scaffolds for the development of other drugs with varied mechanisms of action. For example, sulfanilamide has not only been developed as an antibiotic, but has also served as the archetype of other drugs for hypo-glycemia, diuretics and hypertension[21] (**Error! Reference source not found.**), underscoring the diversity and importance of small molecules in the drug discovery process.

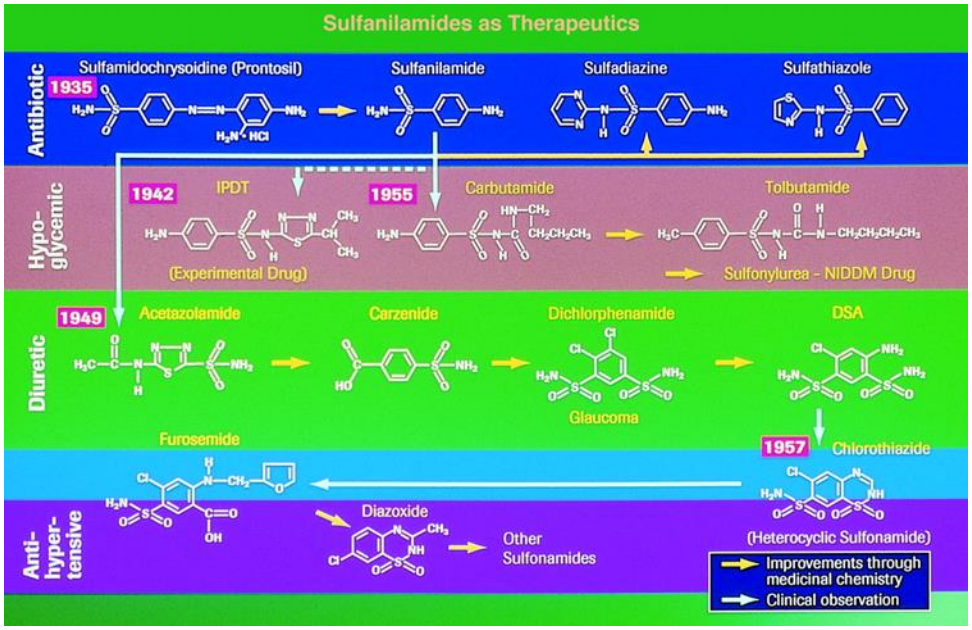


Figure 1. Small molecule drugs with varied function. From Drews, J., Drug discovery: a historical perspective. Science 2000, 287 (5460), 1960-4. Repinted with permission from AAAS.

Small molecule chemical entities are commonly characterized as having a molecular weight of <500 Daltons. Many small molecule drugs are stable in the acidic milieu of the stomach,

as well as the basic environment of the small and large intestines. These factors allow for small molecule drugs to have high oral bioavailability.[22-23] As such, small molecules have been a mainstay in the area of drug development. Despite their ubiquity in the drug discovery process, small molecules have several drawbacks. Small size and limited functional groups often lead to interaction with multiple targets, decreasing selectivity and resulting in the possibility of adverse effects.[24] This issue is particularly pressing when targeting protein-protein interactions, as the size limitation of small molecules may not recapitulate the extent of interaction that endogenous ligands can have with predefined binding pockets, or sites, necessary to evoke a biological response.[25-26] Consequently, there has been a push to obtain pharmaceutical candidates from other molecular classes.

The advancement of molecular biology and the advent of recombinant expression has spurred an interest in protein biologics: a class of therapeutics derived from biological sources including monoclonal antibodies, cytokines and growth factors.[27] These macromolecules are considerably greater in mass than small molecule drugs, with molecular weights >5 kDa, and contain many functional groups including hydrogen donors and acceptors, violating several of the contingencies of Lipinski's "rule of 5" for oral bioavailability.[13, 28] As such, members of this drug class are often administered parenterally.[27] Of the various biological molecules available, the advantages of proteins as pharmaceuticals include their specific and selective functions, which cannot be readily recapitulated by small molecule drugs, resulting in therapeutics that are more efficacious and have fewer off-target effects.[29] The specificity of proteins can be attributed to their ability to interact with a larger surface area of targets or receptors as compared with small molecule drugs.[27] This characteristic allows biologics to affect targets and pharmacophores that are inaccessible to small molecules.[29] The potential for this pharmaceutical class has been supported

by a greater number of biologics being approved by the Food and Drug Administration (FDA) within 2014-2019 as compared to the preceding decade from 2000-2009 (Figure 2). [30]

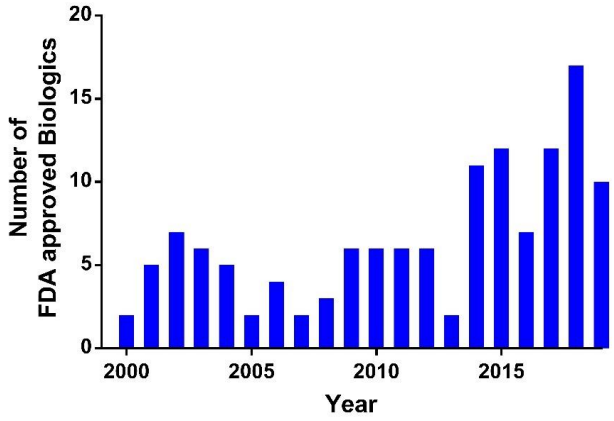


Figure 2. The number of FDA approved biologics from the year 2000 to 2019. Adapted from de la Torre, B. G.; Albericio, F., The Pharmaceutical Industry in 2019. An Analysis of FDA Drug Approvals from the Perspective of Molecules. Molecules 2020, 25, (3).

Despite gaining popularity, large biologics suffer several drawbacks, chief among which is susceptibility to enzymatic degradation. Protein biologics are susceptible to digestion by gastric and pancreatic enzymes including pepsin, trypsin, and chymotrypsin, precluding efficient oral administration.[27] Even if a protein biologic can withstand the harsh environment of the digestive system, poor transport across the intestinal membrane and mucus in the intestinal epithelium provide physical hindrances that result in diminished systemic bioavailability.[31] This low bioavailability is compounded by a short plasma half-life due to metabolic processing in the body, breaking down large biologics into non-bioactive peptide fragments and amino acids that are either excreted or reprocessed as building blocks for synthesizing endogenous proteins.[32-33]

The disadvantages of small molecule drugs and biologics have motivated researchers to look for alternative molecular frameworks that mitigate these drawbacks. This has led to a renewed

interest in the use of peptides as medicines. Like their larger protein counterparts, peptides are involved in many biological pathways and physiological functions which is why peptide hormones such as insulin and adrenocorticotrophic hormone (ACTH) have historically been used as therapeutics.[34] Many of these initial agents were isolated from natural resources such as animal tissues and organs, which could affect the amount and purity of peptide isolated. The advancement of chemical synthesis of peptides on solid support pioneered by Bruce Merrifield have allowed for the accessibility of peptides that may be difficult to isolate or deemed intractable, and has resulted in the production of synthetic peptide drugs such as oxytocin and enfuvirtide (Fuzeon).[34-35]

Many biologically active peptides range from 10 to 50 amino acids long and are approximately 500-5000 Da in molecular weight, occupying a size between small molecules and large biologics.[13] This size range allows for the desirable characteristics of both small molecules and large biologics. Like large biologics, peptides have high selectivity and potency towards their binding target due to the presence of multiple functional groups in a given sequence and an extended interface.[23] Similar to small molecules, synthetic peptides are less immunogenic and can be tuned to be more stable *in-vivo* just like small molecule drugs.[13] Tunability is an important factor to the development of this drug class, as naturally occurring peptides suffer many of the same drawbacks as their larger protein counterparts, including susceptibility to amide bond cleavage by proteases, low absorption, and decreased bioavailability.[2]

1.1.1 Overview of Peptide Chemical Modifications

The biological function of a peptide is contingent on adopting a proper fold. Folded structure (secondary or tertiary) dictates the spatial orientation of amino acid side chains and mediates the interaction of the macromolecule with its molecular target. Issues related to working

with peptides and developing drug leads involves overcoming the barriers related to synthesis, recapitulation of structure, and modulation of function. Numerous chemical approaches exist to overcome these challenges. The most common methods include side chain and terminal modifications, backbone modification, and structural constraint via macrocyclization; all with the goal of facilitating efficient synthesis and chemical property modulation.

1.1.1.1 Side Chain and Terminal Modifications

One of the most common methods for modifying peptides involves altering side chain structure. Each of the twenty canonical amino acids can be categorized based on the functionality of its side chains – aliphatic, polar non-ionizable, polar ionizable, and aromatic. Each of these side-chain types has been used as chemical handles for complex derivatizations, a detailed discussion of which is beyond the scope of this dissertation but is elegantly delineated by Baran and coworkers in a recently published review.[36] One of the most common side chain modifications to alter the property of a peptide *in vivo* is the attachment of polyethylene glycol (PEG) moieties. PEGylation of peptides has been shown to augment conformational stability, increase serum half-life as well as improve peptide solubility.[37-38] PEG units are typically introduced through functionalization of one of both of the polymer ends with reactive groups that are compatible with the side chain functional groups of the desired amino acid to be conjugated.[38] Early PEG derivatization approaches focused on alkylation of amine groups, typically the ε-amine of lysine residues, as well as the thiol groups of cysteine but has since been adapted to other canonical amino acids as well as residues not found in nature. For example, unnatural amino acids bearing azides or propargyl moieties have been derivatized via click chemistry, and residues containing ketone handles have been PEGylated through oxime conjugation (Figure 3). [38-39]

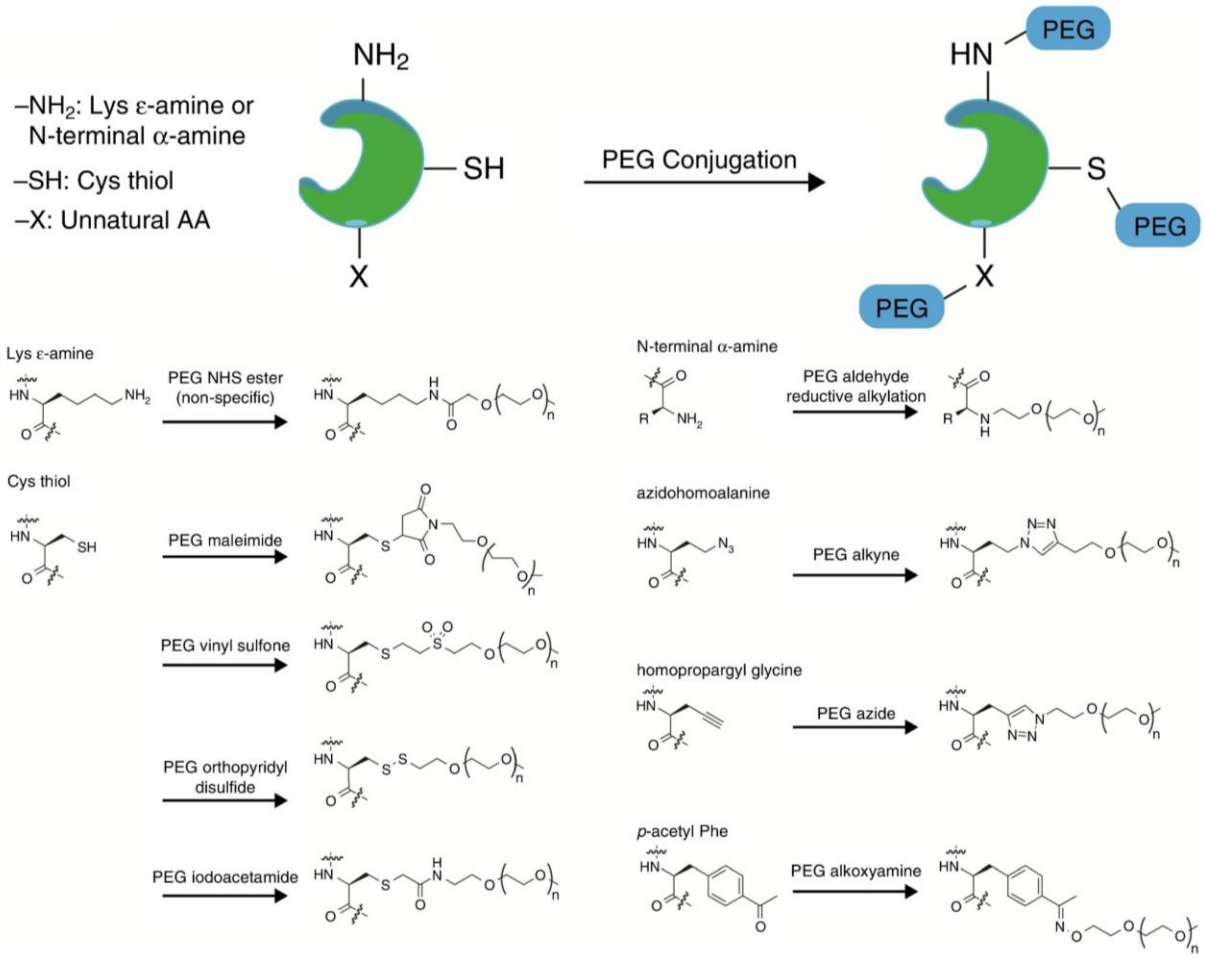


Figure 3. Examples of common PEG conjugation performed on proteins and peptides. Adapted from Lawrence, P. B.; Price, J. L., How PEGylation influences protein conformational stability. Current Opinion in Chemical Biology 2016, 34, 88-94.

The ability of PEG groups to be conjugated to amine groups allows them to be a common modification of the amino terminus of peptides and proteins, which has been used to improve enzymatic stability and circulation half-life in the blood.[37, 40] Other N-terminal modifications include acetylation to control peptide charge, modulate cell localization, and impart protection against proteases, the addition of pyroglutamic acid to improve target binding and inhibition, as well as the conjugation of fluorophores such as fluorescein for quantitative detection of peptides

in biophysical and biochemical assays.[41-44] Common modifications of the C-terminus include the addition of a carboxamide or thioester. C-terminal amidation is commonly found in bioactive peptides such as the hormone oxytocin and antimicrobial peptides such as sarcotoxin IA, and has been used to increase protection against exoproteases.[45] C-terminal thioesters are also common in peptide synthesis and participate in native chemical ligation, allowing for the coupling of two peptide fragments.[46]

More than just methods for improving chemical properties, side chain modifications of amino acids can have pronounced effects on peptide bioactivity. For example, derivatization of HIV-1 Tat cysteine side chains with methyl or tert-butyl moieties, and substitution of the side chain guanidinium of arginine with a urea moiety eliminated the peptide's capability to stimulate HIV replication.[47] Moreover, it was found that side chain length alteration, through incorporation of homologated arginine in synthetic peptide analogs of parathyroid hormone PTH(1-14)NH_2 generated variants that were thirty times more effective at stimulating downstream cAMP synthesis in than the parent peptide in cAMP activity assays.[48]

Other important side chain modifications in the scope of drug development include the introduction of post-translational modifications such as lipidation, glycosylation, hydroxylation and phosphorylation. Lysine residues can be acylated to improve lipophilicity and membrane permeability, allowing for more efficient absorption in the gut lumen.[49-50] Lipidation has also been used to enhance the bioactivity of glucagon for its receptor through stabilization of its secondary structure.[51] In addition, serine and asparagine residues can be glycosylated to enhance bioavailability by augmenting metabolic stability, lowering clearance, and even increasing biodistribution through site specific targeting.[52] The breadth of chemical modifications on peptide

side chains exhibits its high degree of modularity and further indicates why these macromolecules are desirable as pharmaceutical leads.

1.1.1.2 Peptide and Protein Backbone Modification

Another effective approach to augmenting the pharmacological properties of peptides involves the replacement of natural amino acids with unnatural moieties in order to alter the connectivity of the α-peptide backbone. These substitutions can offer local restrictions to a peptide structure, predisposing it to adopt a select conformation. Given that the function of a protein is a consequence of its three-dimensional structure and amino acid sequence, modification of the folded structure can have a direct effect on function and, thus, its pharmacological properties. The onus for engineering the peptide backbone is borne from four primary motivations: 1) facilitation of synthesis 2) surveying the effects of chirality 3) study and control of folding, and 4) generation of protein mimics.

Pioneering efforts in the alteration of the peptide backbone were carried out as means to facilitate extended protein chain synthesis. Kent and colleagues demonstrated the synthesis of HIV-I protease variants containing a thioester in place of an amide in the peptide backbone as a result of a non-native ligation of two peptide fragments; one having an N-terminal bromoacetyl group, and the other containing a C-terminal thioacid.[53] The resulting HIV-I proteases were found to be bioactive, demonstrating that proteins with unnatural backbones can recapitulate native function. Later endeavors from the same group demonstrated the effects of chirality on the peptide backbone, through the synthesis and characterization of the D-enantiomer of HIV-I protease, which catalyzed a similar reaction as the native enzyme, albeit to the D-enantiomer of the native substrate[54]. This undertaking paralleled the efforts of Berg and Zawadzke in the characterization of the D-enantiomer of rubredoxin[55]. In order to understand the properties of folding, researchers

have applied amide bond surrogates including N-methyl residues to abrogate hydrogen bonding, Cα-methylated residues and thioesters to control backbone flexibility, and dipeptides as well as small molecules to induce and stabilize turn motifs.[56-61]

The final motivation for engineering protein backbones focus on peptide and protein structural mimicry. Protein structure and pharmacological properties are part and parcel of each other, therefore, mimicry and modification of peptide structure can affect its pharmacology. One effective method in probing this relationship and designing peptide structural mimics involves the use of foldamers, unnatural oligomers with a propensity to fold into defined conformations or ordered states.[5] These ordered states are stabilized by noncovalent interactions between non-adjacent monomers of the linear sequence.[62]

A particularly robust method of synthesizing foldamers involves the judicious substitution of the peptide backbone with noncanonical amino acids, resulting in a different backbone connectivity. Some of the common unnatural amino acid classes used include, β^3, β^2, β-cyclic, N-methylated, Cα-methylated, ornithine and D-amino acids. Incorporation of these residues in known combinations and regular patterns can create peptides with a heterogeneous backbone, which have been shown to generate a myriad of structures that are reminiscent of natural peptide and protein secondary structures including α-helices, β-sheets, turns, β-hairpin turns, and loops (Figure 4).[63]

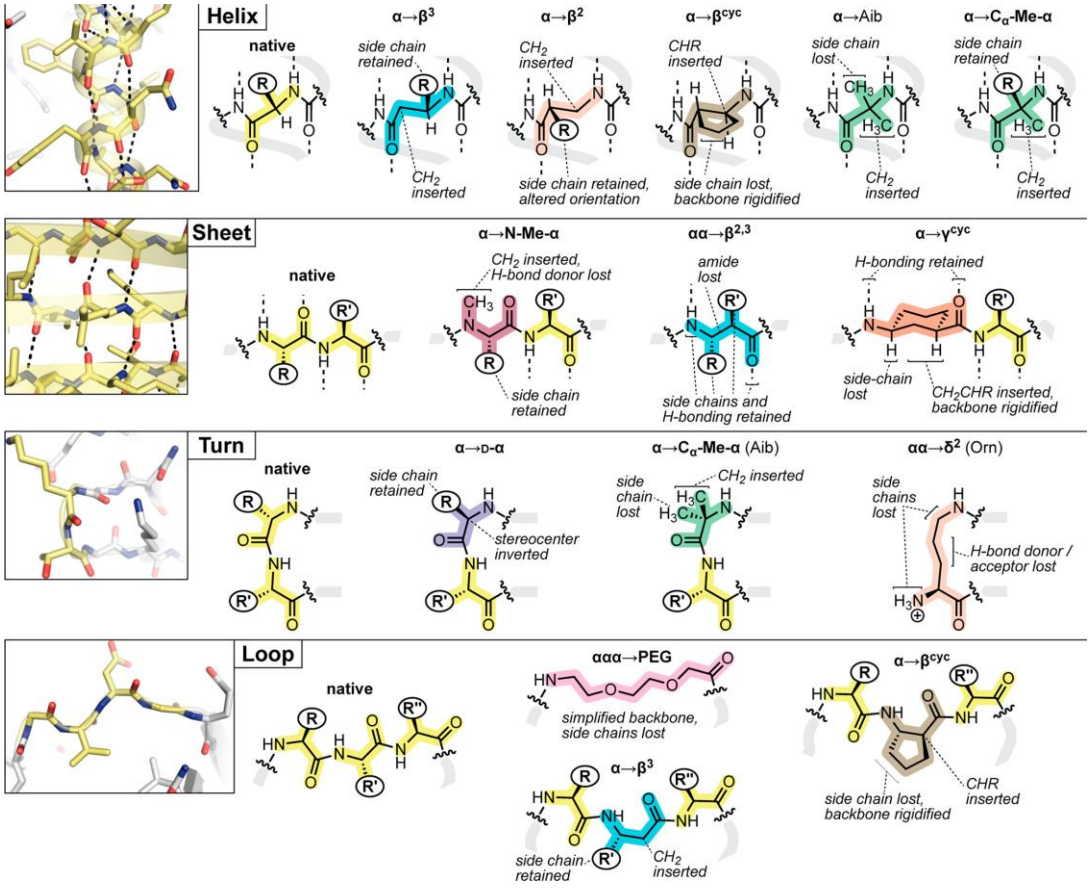


Figure 4. Design considerations in the mimicry of peptide secondary structure through peptide backbone modification. Adapted from George, K. L.; Horne, W. S., Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. Accounts of Chemical Research 2018, 51 (5), 1220-1228.

Heterogeneous backbone analogs can exhibit the same activity as their prototype peptides. Gellman and colleagues have demonstrated that foldamer counterparts of glucagon-like peptide 1 were able to retain native-like function and greater structural stability *in vivo* compared to the native peptide.[64-65] Due to the differences in backbone connectivity, foldamer peptides containing a high density of backbone substitutions display resistance to proteases.[6] Moreover, it has been demonstrated by our lab that relative positioning and combination of these unnatural amino acids can impart varying degrees of protease resistance[8]. This can, in principle, help in generating peptide drugs with greater oral bioavailability.

An important frontier in the field of peptide mimicry is to apply the principles obtained from precedent studies on foldamers to higher order architectures beyond peptide secondary structure to comprise tertiary folds, generating protein-like analogs or proteomimetics.[66] Using the natural peptide sequence as a template and rational substitution directed by secondary structure, heterogeneous backbone modification serves as a robust methodology for generating proteomimetics with a high degree of unnatural amino acids that work in concert to copy the global fold of a peptide or protein. This methodology has seen success in the synthesis of larger proteins with more complex folds, such as the B1 domain of Streptococcal Protein G (GB1).[67] Moreover, this approach also allowed for the synthesis of functional mimics of naturally occurring sequences such as the DNA-binding zinc finger SP1-3[68] and ubiquitin.[69]

1.1.1.3 Cyclization

Reducing the conformational freedom of a peptide requires a concerted effort of both local and global restriction strategies. Macrocyclization allows for the constraint of the general fold of a structure, lowering conformational entropy and stabilizing both the secondary and tertiary structure of a peptide, making the structure more defined, and in principle, more selective towards its molecular targets.[13] Additionally, cyclized peptides are often more proteolytically stable than linear counterparts.[70] Cyclization of a peptide involves the covalent cross-linking of distal parts of a sequence and can be accomplished in four ways: N-C terminus, side-chain to side-chain, head-to-side chain, and side chain-to-tail, depending on functional groups available within the peptide sequence.[13, 71]

N-C terminal cyclization, also referred to as backbone cyclization[72], was partly inspired by naturally occurring cyclic peptides often referred to as cyclotides found in bacteria and plants. One of the earliest examples of work with cyclic peptides involves Gramicidin S, an antimicrobial

peptide derived from *Bacillus brevis* and used to treat wounds and prevent infection during World War II. [71] Since then, many backbone cyclized peptides have been isolated or synthesized and used clinically including cyclosporine[73] and pasireotide[74]. Subsequent studies on cyclic peptides have shown that exceptional resistance to exoprotease degradation due to the lack of free N- and C-terminals.

The proteolytic stability of these peptides has inspired chemists to apply backbone cyclization to numerous bioactive peptides to decrease the chances of degradation in the digestive system and in serum. Early work by Tam and colleagues demonstrated the utility of native chemical ligation to the backbone cyclization of peptides Kalata, Circulin A and B, and cyclopsychotride where the linear precursor peptides were synthesized with an N-terminal cysteine and a C-terminal residue bearing a thioester.[75] Later work by Craik and colleagues demonstrated that cyclization can also be achieved through standard amide bond coupling of a free N- and C-terminus using HATU and DIEA. In this route, the linear peptide of Kalata B1 and its D-enantiomer were synthesized on acid-labile 2-chlorotrityl resin and cleaved under mild acidic conditions, to prevent side chain deprotection (Figure 5). Cyclized Kalata B1, was found to be resistant to proteases, but with comparable efficacy as its native counterpart.[76]

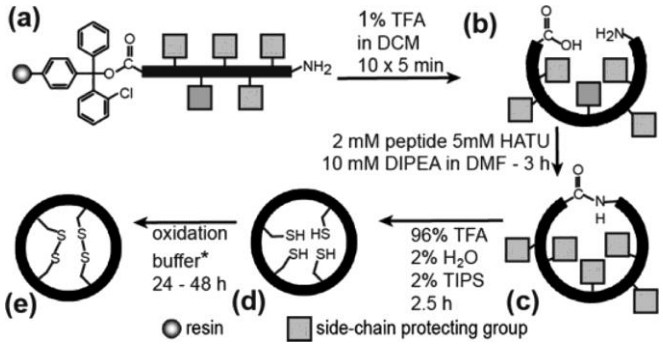


Figure 5. Schematic of the synthesis of cyclic kalata B1. Adapted from Cheneval, O.; Schroeder, C. I.; Durek, T.; Walsh, P.; Huang, Y.-H.; Liras, S.; Price, D. A.; Craik, D. J., Fmoc-Based Synthesis of Disulfide-Rich Cyclic Peptides. The Journal of Organic Chemistry 2014, 79 (12), 5538-5544.

Numerous cyclization methods exist, including the use of Sanger's reagent, glyceric esters as well as Staudinger ligation as detailed in a comprehensive review by Yudin.[71] An important consideration when applying N-to-C cyclization to a non-cyclotide is the distance between the N- and C- terminus in the three-dimensional fold of the peptide. If the termini are too distant from each other, cyclization can perturb the overall structure of the peptide. In the case of bioactive cyclic peptides, linkers comprised of extended sequences of glycine and alanine residues have been employed.[77]

Less common cyclization strategies involve the ligation of the N-terminus or the C-terminus to a side chain of an amino acid in the sequence. These ligation strategies, like N-to-C cyclization, usually involve the formation of a lactam ring. Head-to-side chain to ligation can be achieved by lactam ring formation between the N-terminal amine to a side chain of a glutamyl or aspartyl residue. However, this cyclization has also been performed using a lactam handle from the ε amine of lysine residue, as demonstrated by Aldrich et al. in the synthesis of dynorphin A analogs.[78-79] Side chain-to-tail ligation can be performed between a lysine ε-amino group and a free C-terminal thioester in aqueous imidazole solutions.[80]

Side chain-to-side chain cyclization or stapling is the most common method for generating peptide macrocycles. In nature, this can be achieved through disulfide bonds (discussed in greater detail in section **1.2**) or through an amide bond formation between the side chains of lysine or ornithine and distal aspartic acid or glutamic acid residues. The noncanonical amino acid selenocysteine has also be been used as a more stable, isosteric substitution for disulfide bonds. Other functional groups historically used to link amino acid side chains include thioethers, dicarba

linkages, ethers, triazoles and esters. These and other linking chemistries have been discussed in detail in precedent reviews, notably by Spring[81], Davies[82], and de Araujo.[83]

1.2 Disulfide-Rich Peptides

The thiol side chains of two cysteines can react with each other in the presence of an oxidizing agent in basic conditions to form a disulfide covalent linkage termed a cystine dimer.

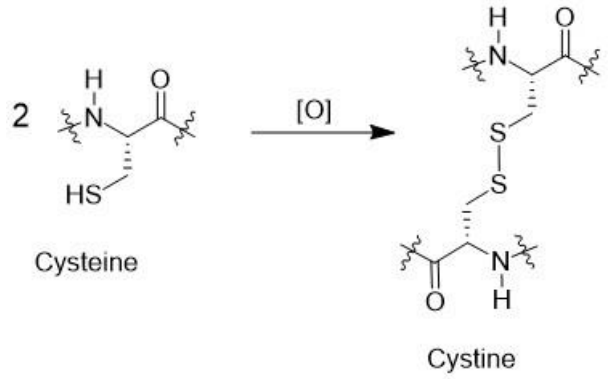


Figure 6. Two cysteines oxidize for form a disufide-bonded cystine unit.

The resulting disulfide bond formed from this interaction can be found intramolecularly within two cysteines of the same chain or intermolecularly between cysteine residues of different polypeptides. Both types play an important role in the stability of peptide tertiary structure. Structurally, the CH_2-S-S-CH_2 linkage between the cysteine residues are arranged in a non-linear, non-planar conformation.[84] The cystine dimer itself has limited rotation about the $\chi 1$ and $\chi 2$ angles of each respective cysteine residue, along with the $\chi 3$ angle of the S-S bond, adding further constraint to the peptide structure.[85-86] The prevalence of cysteine residues in nature is low compared to other proteinogenic amino acids.[87] However, a large percentage of cysteine residues

can be observed in excreted proteins with varied functions including venoms, hormones, defensins, and enzyme inhibitors that are abundant in nature and encompass a larger family of disulfide-rich peptides.[9]

Disulfide-rich peptides are comprised of less than a hundred amino acids but can contain up to 10% or more of cysteine residues within its sequence.[88] These residues are conserved and can be used to classify peptides belonging to the same group. Due to their short sequences, these peptides can be bereft of an extended hydrophobic core and canonical secondary structure motifs but still take on a folded structure due to the stabilization of one or more disulfide bonds.[9] Like cysteine placements in the sequence, these folds are typically conserved and are used to classify these peptides. Precedent studies have shown that the disulfide bond pattern and topology of bioactive peptides directly influence global fold and structural motifs[89].

One of the most common motifs found in bioactive disulfide-rich peptides the is the cystine knot motif characterized by a sequence containing six cysteine residues with a disulfide bond connectivity pattern of CysI-CysIV, CysII-CysV and CysIII-CysVI. These disulfide bonds are spatially arranged in a rotaxane-like structure such that the CysIII-CysVI bond is threaded through a macrocycle formed by the other two cystine moieties, and peptide backbone segments from CysI to CysII, and CysIV to CysV. There are three types of cysteine knots, based on the peptides from which these structures were originally observed, namely growth factors, peptide toxins and cyclic peptides, however despite the differences in structure, the cysteine connectivity of all three classes are the same.[90] This motif is widespread in nature; being present in a number of peptide toxins and has been shown to impart a degree of proteolytic protection as well as resistance to thermal and chemical denaturation.[91]

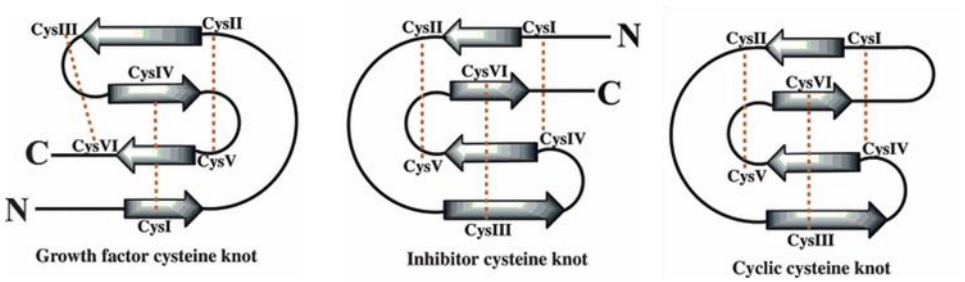


Figure 7. The three types of cysteine knots motifs all exhibit the same cysteine connectivity. Reproduced with permission from Iyer, S.; Acharya, K. R., Tying the knot: The cystine signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. The FEBS Journal 2011, 278 (22), 4304-4322.

Another common motif found in disulfide-rich peptides is the cysteine stabilized α-helix, which can be found in neurotoxins such as apamin and vasoconstrictors like sarafotoxin.[92] This motif is classified by disulfide bond pairings between two distal consensus sequences CXC and CXXC, where X is any amino acid. Other common motifs observed in nature include disulfide stabilized αβ motifs and the β-hairpin-like motif, that has a CysI-IV and CysII-III disulfide bond connectivity and a CXC-CXXXC spacing pattern.[89] A comprehensive bioinformatic investigation on 1595 disulfide-rich peptides in the PDB yielded a complex classification comprising of 41-fold groups and 98 families, based on common structural cores or motifs[9], many of which were peptides and miniproteins that have vastly disparate biological functions, showing the array of structural and functional diversity available to these domains.

1.2.1 Bioactive Disulfide Rich Peptides

Disulfide-rich peptides serve diverse biological functions and can be found in a variety of life forms in nature. In higher organisms such as primates and humans, disulfide-rich peptides

including as insulin and defensins function as hormones or antimicrobials. In plants, these molecules serve as insecticides and enzyme inhibitors as seen with kalata B1 and sunflower trypsin inhibitor. Other peptides including chlorotoxin in spiders and Vc1.1 from cone snails serve as potent neurotransmitters (Figure 8).[12] The array of functional diversity available to disulfide-rich peptides make them attractive starting points for drug discovery and have spurred several areas of research devoted to the study of these bioactive compounds.

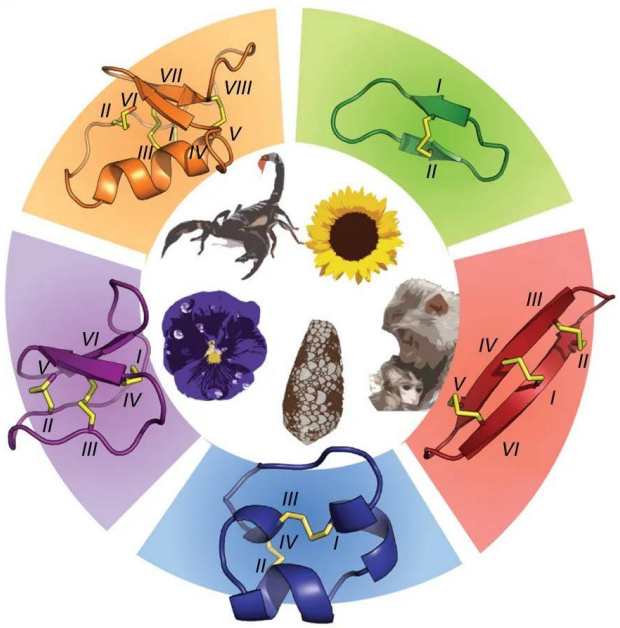


Figure 8. Various bioactive disulfide-rich peptides and respective organism of origin. Clockwise from green sector: sunflower trypsin inhibitor-1, θ-defensin-1, α-conotoxin Vc1.1, kalata B1 and chlorotoxin. Reproduced with permission from Wang, C. K.; Craik, D. J., Designing macrocyclic disulfide-rich peptides for biotechnological applications. Nature Chemical Biology 2018, 14 (5), 417-427.

A burgeoning field involves the investigation of disulfide-rich peptides from animal venoms including those isolated from arachnids, insects, reptiles and mollusks.[93] Interest in the study of these toxins stemmed from initial observations that the venom cocktails used to incapacitate the prey of these animals were comprised of a large majority of disulfide-rich peptides.

Isolation of these peptides showed that many were able to bind to neuroreceptors such as nicotinic acetylcholine receptors and voltage gated ion channels, blocking downstream effects and eventually leading to physiological responses including tremors, and muscle paralysis.[94]

Structure-activity relationship studies of these venom compounds have found that disulfides play a structural role in the bioactivity. This is unsurprising as many disulfide bonds are essential in maintaining the tertiary fold of a given sequence. Accordingly, reduction of disulfides can perturb structure and even abrogate function especially in peptides that exert a mechanism of action by binding to a specific receptor. Early work by Kent and colleagues demonstrated that one of three disulfide bridges (Cys17-80) in the cytokine interleukin is essential in bioactivity; and that analogs with all cysteine pairs substituted with alanine exhibited 2000-fold less activity than the native peptide.[95] Recent work on the venom component μ-conotoxin PIIIA* showed that elimination of key cysteine pairs of the peptide perturbed structure and reduced the capacity to inhibit $hNa_V1.4$ channels.[96] Since proper structure is required for function, it follows that non-native of disulfide pairings of these toxins can also have an impact on bioactivity. Pioneering investigations by Nishiuchi and Sakakiraba on structural isomers of α-conotoxin GI showed that analogs with non-native disulfide pairings have diminished ability to induce muscle paralysis in mice.[97] Later structural investigations showed that these analogs greatly differed in structure.[14]

The biological activity of disulfide-rich peptides, however, are not limited to venoms. Natural disulfide-rich peptides in the body, such as insulin and vasopressin play important roles in human metabolism. Insulin is involved in the regulation of blood sugar levels[98], while vasopressin is involved in control of plasma volume and osmolarity.[99] Deregulation in the expression of these peptides are associated with various metabolic diseases including diabetes mellitus and diabetes insipidus.[100] As such, these peptides have been used as hormone replacement therapeutics to

alleviate the symptoms of these diseases. Insulin has been invaluable in controlling diabetic ketoacidosis and acute hyperglycemia, while vasopressin has been utilized in the concentration of urine for patients with diabetes insipidus.[100-101]

Peptides found in other organisms have also been exploited as potential therapeutic agents. Defensins found in phagocytic and epithelial cells of the mammalian innate immune system have been studied for their bactericidal properties.[102] Sunflower trypsin inhibitor has been researched as a kallikrein inhibitor in blood coagulation[103], while Kalata B1 from *Oldenlandia affinis* has been researched for its capability to induce muscle contractions during childbirth and its potential as an insecticide.[104] These precedents highlight the intricate role, and distinct functions of disulfide-rich peptides in biology.

1.2.2 Disulfide-Rich Peptides as Scaffolds for Drug Development

Disulfide-rich peptides have been touted as privileged scaffolds for drug development as this peptide class exhibits several desirable properties including compact folded architectures, structural stability, and modularity. The presence of one or more disulfide bonds in the peptide architecture can impose local constraints, which limit the conformation flexibility of the folded structure of the peptide.[105] Formation of disulfide bonds at distal regions of the peptide, limit the conformational entropy of the unfolded state, thus predisposing the peptide to conform to the native folded state.[13, 106] Collectively, this allows the peptide to assume a structure closest to its native active conformation. With respect to bioactive peptides, such as venom components bind to receptors or ion channels, this can augment affinity and selectivity, translating to increased potency. Another desirable pharmacological property of disulfide-rich peptides is that their cystine-stapled architectures can impart a degree of protection against thermal and chemical

denaturation compared to unstapled counterparts, which may be leveraged in generating orally bioavailable peptide drugs. This has seen success in the disulfide-rich peptide drug Linzess (Linaclotide), which may be administered orally as a treatment for irritable bowel syndrome.[107] Lastly, the defined conformation of the folded states of disulfide-rich peptides allows for replacement of certain parts of a sequence with other bioactive sequences, a technique often referred to as molecular grafting.[108] This technique can allow for the bioactive portion of an otherwise unstable peptide to be grafted into a scaffold that is not only more stable but also leads to improved affinities and potencies of the guest.[13] This technique was utilized in generating vascular endothelial growth factor antagonists using Kalata B1 as a scaffold, as well as an ω-conotoxin MVIIA/MVIIC chimera to modulate affinity for N- and P/Q-Type calcium channels.[109]

Despite possessing desirable pharmacokinetic properties, disulfide-rich peptides also suffer several drawbacks that have curtailed their widespread application. These include, disulfide-bond scrambling, low membrane permeability and rapid renal clearance.[10, 110-111] Additionally, even though the constrained architectures of disulfide-rich peptides can improve structural stability, not all these peptides can be administered orally as many sequences are still susceptible to enzymatic degradation by proteases of the digestive system.[110] A comprehensive study by Basit and colleagues has shown that different disulfide-rich sequences have varying resistance against select gastrointestinal fluids. In this study, insulin was found to be unstable in both gastric and intestinal fluids while oxytocin was resistant to gastric fluid while being susceptible to intestinal fluid.[112] Furthermore, studies on ICK-containing peptides, showed varying degrees of susceptibility when incubated with select enzymes.[113]

Another drawback in developing disulfide-rich peptides as drugs involves disulfide bond scrambling. Precedent studies have shown that disulfide bond scrambling can occur in blood

plasma due to the presence of thiols in serum albumin which resulted in structural isomers with non-native disulfide bond pairing. This interaction had the potential to impact bioactivity.[111, 114] Mathematically, the number of structural isomers p that can form from intramolecular disulfide bonds is given by the formula $p = \frac{(2n)!}{(2^n)n!}$ where n is the number of disulfide bonds in a single peptide chain. As such, sequences containing 2 (Figure 9), 3, and 4 disulfide bonds can theoretically yield 3, 15, and 105 structural isomers respectively.[105] Each disulfide bond isomer has a different fold than the native peptide, and precedent investigations have shown that many of these isomers have no bioactivity, or in the case of some peptide toxins, target different receptors. A study on human β-defensin 3 has shown that disulfide bond topological isomers showed varying effects on the chemotactic activities of *Escherichia coli* but negligible effect on its antimicrobial activity.[115] Interestingly, substitutions of cysteines with isosteric homoalanine, had no effect on antimicrobial activity but completely eliminated chemotactic activity.[115] The presence of other disulfide bond topologies has also historically been a problem in the chemical synthesis of these peptides, especially in cases where the generation of disulfide bonds are achieved through statistical oxidation of unprotected cysteine residues.[15, 106] This issue is especially prevalent in synthesis of small peptides such as animal toxins that lack both a regular secondary structure, and an extended hydrophobic core that helps preorganize a peptide into a specific fold.

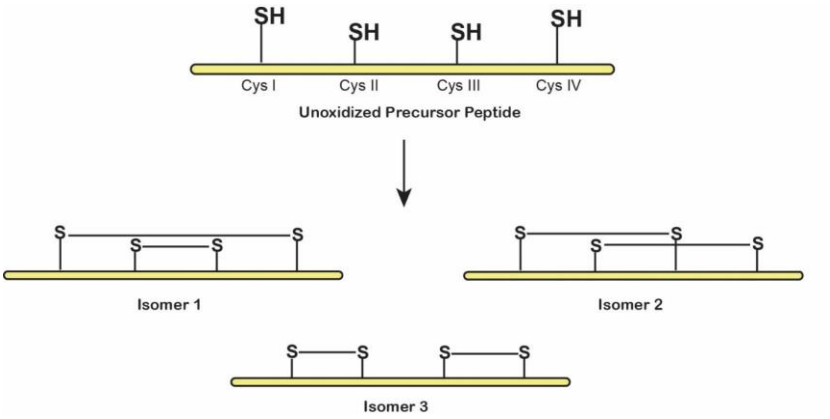


Figure 9. Three structural isomers with different disulfide connectivity can form from the statistical oxidation of a linear unoxidozied precursor peptide with 4 cysteine residues.

The drawbacks of disulfide-rich peptides as starting points for drug design has spurred an interest in the development of chemical strategies for adding increased stability to the peptide structure, by replacing disulfide linkages with more stable covalent bonds including diselenide bonds,[116] amide cross-linkages,[117] thioether bonds[118], and triazoles.[119] In many cases, peptide variants containing these replacements showed improved stability but with the caveat of a more complicated synthesis. The generation of other disulfide bond topologies have also provided an onus for research on the development of disulfide-rich peptide scaffolds that minimize the creation of these isomers. For example, through incorporation of unnatural residues penicillamine and a synthetic amino acid containing two thiol functional groups, Wu and colleagues were able to generate scaffolds containing six cysteines that oxidized with only two structural isomers.[15] Moreover, advances in computational techniques have allowed Baker and colleagues to design *de novo* disulfide-rich scaffolds that fold into a single isomer and has also enabled the construction of other stable peptide macrocycle scaffolds.[120-121]

A complementary method available to peptide chemists that can aid in circumventing drawbacks of disulfide-rich peptides is the incorporation of non-canonical amino acids with

modified backbones in the peptide sequence, creating a disulfide-rich proteomimetics with a heterogeneous backbone (See section 1.1.1.2). A body of work from our lab and others have shown that this strategy can be used to mimic complex yet delicate peptide folds and can be incorporated efficiently in solid phase peptide synthesis. Examples of precedent research on backbone substitutions in disulfide-rich contexts include the incorporation of amino-3-oxopentanoic acid and 6-aminohex-anoic acid into the backbone of μ-conotoxin SIIIA as a means of replacing constrained areas in the peptide. It was found that the generated "polytide" analogs exhibited comparable analgesic activity to the native peptide.[122] In another study, D-amino acids were incorporated in the synthesis of small molecule mimics of the exposed loops of ω-conotoxin CVID.[123] Although these precedents only involved a low density of modification, they demonstrate that backbone modifications can be used to generate analogs of natural peptides that have structural and biological activity. With increased utility of backbone modification, it may be possible to generate proteomimetic variants that have tunable pharmacokinetic properties, while still maintaining the structure and function of the native peptide.

2.0 Heterogeneous-Backbone Foldamer Mimics of a Computationally Designed, Disulfide-Rich Miniprotein

A portion of this chapter has been published in: Cabalteja, C. C.; Mihalko, D. S.; Horne, W. S., Heterogeneous-Backbone Foldamer Mimics of a Computationally Designed, Disulfide-Rich Miniprotein. ChemBioChem 2019, 20 (1), 103-110.

2.1 Introduction

Peptides and small proteins containing multiple disulfide linkages play diverse biological roles.[13] Many act through specific binding to a target receptor, a process facilitated by their complex topologies and constrained architectures. Extensive structure-activity relationship studies carried out on disulfide-rich sequences derived from animal venoms have led to approved drugs (e.g., ziconotide for chronic pain, hirudin for anticoagulation) as well as a number of compounds in clinical development.[124] Expanding beyond their natural functions, disulfide-rich scaffolds have also been engineered as *de novo* ligands for protein surfaces,[125-126] with sequences derived from rational design,[127-128] computational,[129] as well as directed evolution approaches.[130]

While peptides can exhibit improved target selectivity and potency compared to small molecules,[2] they often suffer pharmacological drawbacks including low metabolic stability, poor oral bioavailability, and rapid clearance.[23] Disulfide-rich sequences enjoy some advantages over similar size linear chains in this regard, as the conformational constraint imposed by multiple side-

chain linkages can mask efficient recognition by proteolytic enzymes; however, this protection is not complete and varies considerably with sequence. As an example, the spider toxin ProTx-I is degraded rapidly by trypsin, while the closely related ProTx-II from the same organism (<1.5 Å Cα rmsd)[131-132] is completely stable under the same conditions.[113] The need for improved properties in some disulfide-rich systems has motivated chemists to develop and apply various chemical alterations;[86, 123] examples include backbone cyclization or replacement of the disulfides with other moieties.[15, 133-137] A key challenge in such work is maintaining the precise 3-dimensional fold of a prototype after chemical modification.

One approach for generating analogues of peptides with improved biostability is the modification of the covalent structure of the peptide backbone. A broad body of precedent has shown that many well-defined folds are possible in sequence-specific oligomers containing artificial building blocks. Referred to as "foldamers,"[5] these agents have been shown to mimic an array of natural peptidic structures and functions.[138] A subset of work in this area has focused on "heterogeneous" backbones that combine natural and artificial building blocks in chains that display biologically derived primary side-chain sequences.[6] Our lab has a longstanding interest in the development of design principles for the construction of heterogeneous-backbone foldamers with complex protein-like tertiary folding patterns.[63] In prior work, we validated design rules toward such entities in the context of the B1 domain of Streptococcal protein G and a zinc finger domain from the human protein specificity factor 1.[68, 139]

Given the myriad of unique folds and functions available to disulfide-rich sequences, we saw the pursuit of mimicry of this class by heterogeneous backbones an attractive testbed for expanding the scope of existing design strategies. The effect of backbone alteration on disulfide-rich scaffolds has not been extensively explored, but some precedent exists that inspired the present

work. Bulaj showed the incorporation of ethylene glycol and methylene spacers in place of a Gly-Gly motif in a loop of the μ-conotoxin SIIIA led to hybrids exhibiting bioactivities comparable to the native toxin.[122] More recently, Pentelute reported heterochiral variants of *Ecballium elaterium* trypsin inhibitor II, a disulfide-rich miniprotein that has been retargeted toward cell-surface receptors for tumor imaging applications.[140] Replacement of a large exposed loop in the prototype with a D-α-peptide segment led to mimics of the natural protein with similar folds but improved stability to proteolytic degradation.

Here, we explore the effects of heterogeneous backbone substitution in the context of NC_HEE_D1 (**1**, Fig. 10), a computationally designed 27-amino acid de novo protein recently reported by Baker.[120] This sequence was created as part of a broader effort to develop a family of hyperstable miniproteins as scaffolds for biomedical applications.[129] The fold of 1 is comprised of an α-helix and a β-hairpin stapled by two disulfide bonds ($Cys^{4\text{-}18}$ and $Cys^{14\text{-}27}$). Of note, **1** already contains a backbone alteration at a single site in the form of a D-Pro introduced to stabilize the hairpin turn. This motivated us to assess the amenability of **1** to more extensive modification.

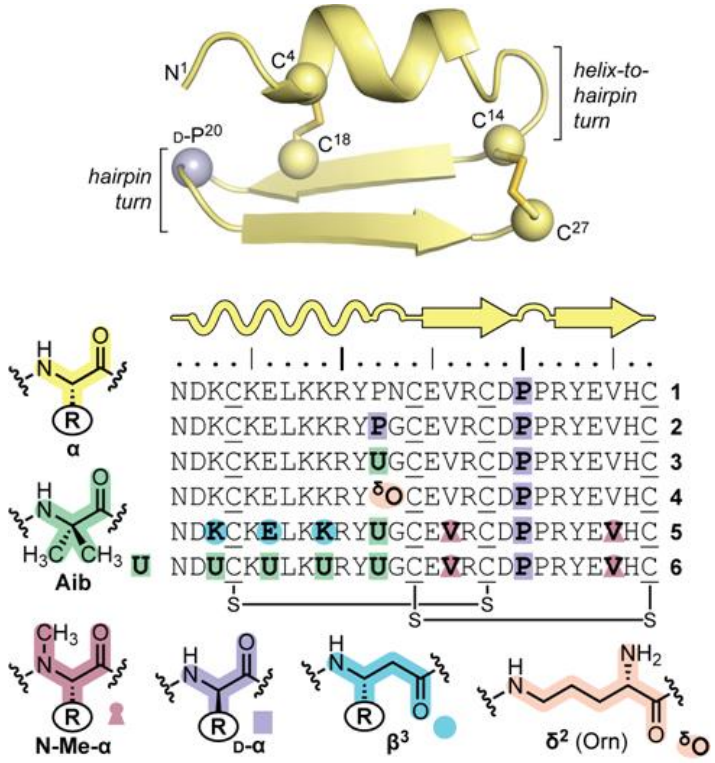


Figure 10. 1 Sequence, secondary structure map, and NMR structure (PDB 5KVN) for peptide 1 and sequences of backbone modified variants 2-6. The identity of "R" groups, when present, match the side chain specified by the corresponding α-residue denoted by the single letter code.

In seeking to expand the scope of protein tertiary structure mimicry by heterogeneous-backbone foldamers to disulfide-rich scaffolds, we were faced with a key question unique to the new target class. Would extensive modification to the backbone have an impact on the oxidative folding pathways? Synthesis of peptides with multiple disulfide bonds can be complicated by the formation of products with the incorrect disulfide-bond connectivity that differ in receptor target or biological activity.[141-144] Though significant work has been put toward developing stable disulfide-rich peptide scaffolds that minimize such byproducts, these problems can be exacerbated by even subtle alterations to peptide sequence.[123]

2.2 Results and Discussion

2.2.1 Design and Synthesis of Turn Variants

Our initial goal was to establish the applicability of the previously reported synthesis of **1** by statistical aerobic oxidation to variants modified in both turn regions. The prototype contains two backbone reversals: a type-II' turn in the center of the β-hairpin and a type-I turn connecting the α-helix to the β-hairpin. As noted above, the hairpin turn already has a backbone alteration, so it was left unchanged. As we evaluated potential modifications to apply in the other turn, we were faced with a conundrum. There are many known backbone modifications that stabilize turns in hairpins; however, these moieties remain largely unexplored in turns from other structural contexts. Due to the role of the helix-to-hairpin turn in **1** juxtaposing the secondary structures that pack to form the core of the tertiary fold, an incompatible substitution would likely disrupt the structure considerably. Given the similarity of the type-I turn to that commonly found in hairpins, we hypothesized that hairpin-stabilizers might be well accommodated. Thus, we synthesized three variants of **1** (Figure 10) in which Pro12-Asn13 was replaced with D-Pro-Gly (peptide **2**), Aib-Gly (peptide **3**), or δ-linked Orn (peptide **4**). In terms of known folding behavior, D-Pro-Gly tends to nucleate type II' or type I' turns,[145] Aib-Gly is more diverse in its conformational preferences,[146-147] and δ^2-Orn promotes a backbone reversal similar to that formed by D-Pro-Gly.[148]

Reduced linear precursors corresponding to **1**-**4** were synthesized by microwave-assisted Fmoc solid-phase methods as C-terminal carboxamides, purified by preparative HPLC, and the identities of each confirmed by MALDI-TOF MS (Table 1). Following literature precedent for **1**,[120] the disulfide bonds were generated in solution through random aerobic oxidation (~75 μM

peptide, pH 8.3 ammonium bicarbonate); these reactions were monitored by analytical HPLC and MS.

Table 1. MALDI-TOF MS data for peptides 1-6

#	$[M+H]^+$ of reduced linear precursor *m/z* (monoisotopic)		$[M+H]^+$ of fully oxidized product *m/z* (monoisotopic)	
	Calculated	**Observed**	**Calculated**	**Observed**
1	3321.6	3321.8	3317.6	3317.6
2	3264.6	3264.7	3260.5	3260.7
3	3252.6	3252.8	3248.5	3248.6
4	3224.6	3324.7	3220.5	3220.8
5	3322.6	3322.9	3318.6	3318.7
6	3150.5	3150.7	3146.5	3146.6

In the case of prototype **1**, conversion of the linear precursor to a single major intermediate and final product was observed, accompanied by characteristic retention time shifts for each new species (Figure 11). A peak eluting later than the starting material was observed in the early phase of the reaction with mass corresponding to the presence of a single disulfide. At later time points, a peak eluting earlier than starting material was formed with a mass corresponding to the fully oxidized peptide. Aerobic oxidation of turn-modified variants **2-4** proceeded in a similar manner as **1**, yielding a single disulfide bond intermediate that transitioned to a single major fully oxidized product (Figure 11). The observation of similar oxidative folding behavior among prototype **1** and turn variants **2-4** suggests identical disulfide topologies and similar folded structures. Purification by preparative HPLC generated samples that were subjected to further characterization to provide direct evidence bearing on this hypothesis.

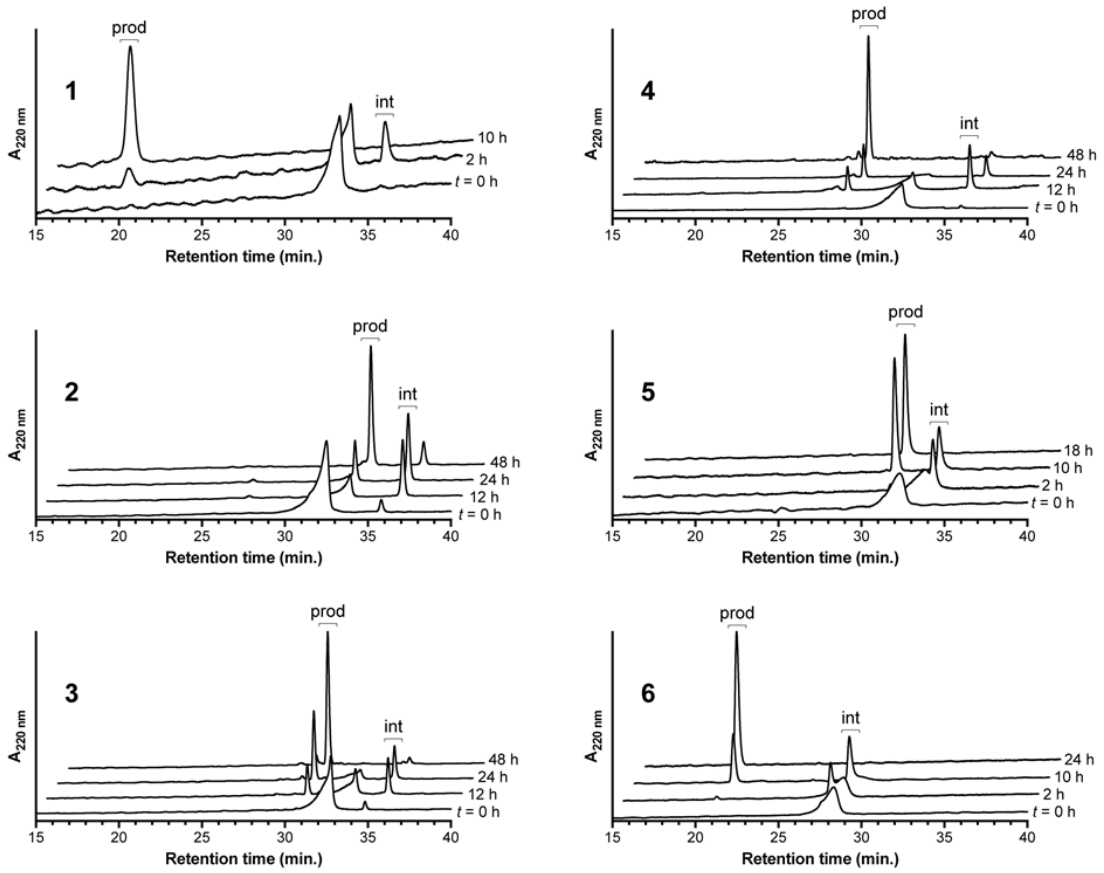


Figure 11. HPLC chromatograms showing the time course of the oxidative folding reactions to form peptides 1, 5, and 6 from the corresponding fully-reduced linear precursors. Conditions: 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3; int = single disulfide intermediate, prod = fully oxidized product.

2.2.2 Structural Analysis

To assess the impact of the turn modifications in 2-4 on the overall fold, we subjected each to circular dichroism (CD) spectroscopy (60 μM peptide in 10 mM phosphate, pH 7.0). The spectrum for prototype **1** showed a minimum at 218 nm and a shoulder at 210 nm (Figure 12), consistent with published results.[120] Spectra obtained for **3** and **4** were similar in shape and magnitude to **1**, suggesting a similar fold. The spectrum for peptide **2** differed more significantly from the prototype, with a reduction in intensity of the negative peaks. This may indicate a change

in overall secondary structure content; however, the complex relationship between CD spectral shape and backbone composition in heterogeneous-backbone foldamers make these results somewhat inconclusive.[139, 149-151]

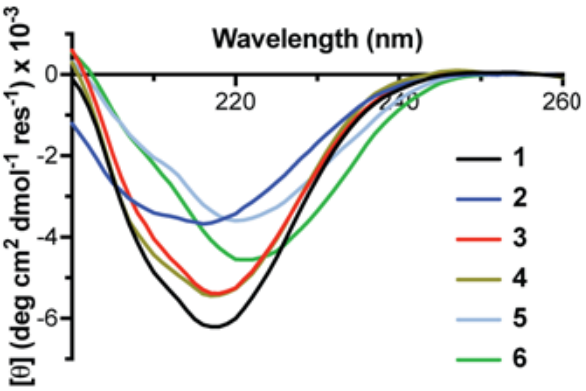


Figure 12. CD spectra of 1-6 **at 60 µM concentration in 10 mM phosphate, pH 7.**

To obtain more incisive data on the folded structures of **1-4**, we subjected each to multidimensional NMR spectroscopy. We acquired $^{1}H/^{1}H$ NOESY, COSY and TOCSY spectra (90% H2O / 10% D2O, pH 4.0, 298 K), enabling full assignment of proton chemical shifts. Qualitative analysis of the NMR results (Figure 13) provides insights into secondary structure content and disulfide connectivity. The α-helix in prototype **1** (Asp^{2}-Tyr^{11}) is indicated by strong sequential $i \rightarrow i+1$ H_N-H_N NOEs, medium-range i→i+3 Hα-HN NOEs, small $^{3}J_{H\alpha\text{-}HN}$ coupling constants, and contiguous negative chemical shift index (CSI) values vs. random coil. The hairpin region in 1 (Cys^{14} to Cys^{27}) is apparent from cross-strand H_N-H_N or H_α-H_α NOEs, large $^{3}J_{H\alpha\text{-}HN}$ coupling constants, and contiguous positive CSI values. Finally, disulfide connectivity can be gleaned from long-range NOEs between remote Cys residues. Analyzing the above factors for **2-4** indicates that each retains the native disulfide topology as well as a defined hairpin region (Figure 13). The NMR observables in the putative helix region for **3** and **4** are very similar to those for

prototype **1**; however, data for **2** show little evidence supporting a well-formed helix. Indeed, the grouping of three positive CSI values preceding the D-Pro may point to a partial strand character.[152]

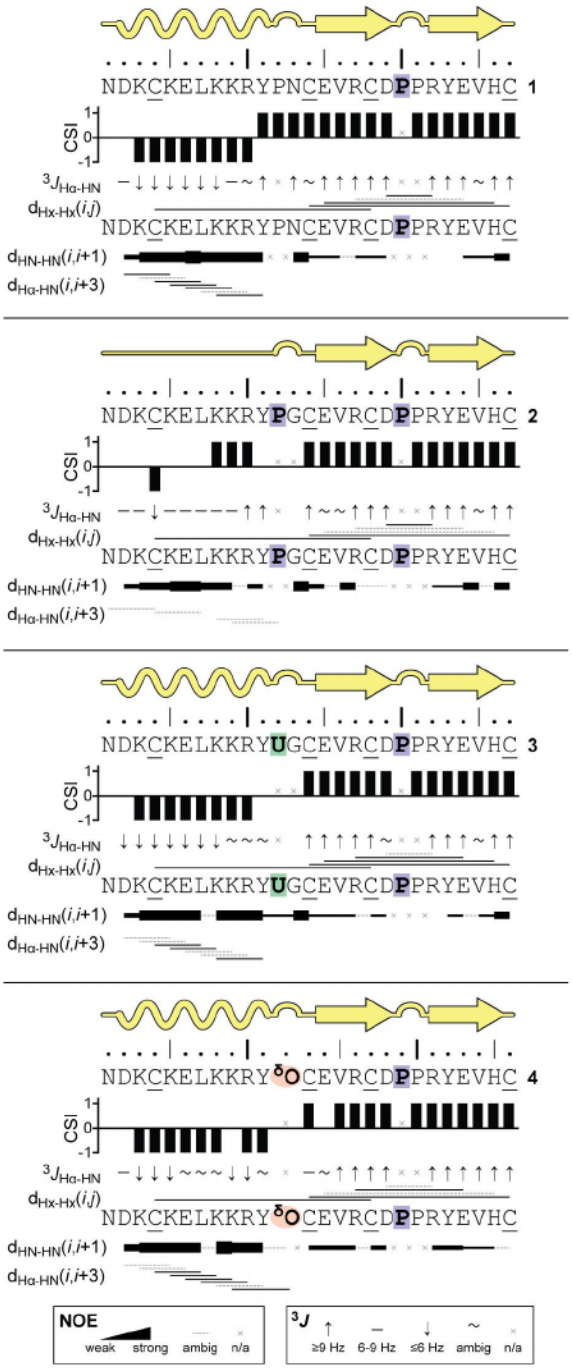


Figure 13. Summary of NMR data (chemical shift index values, vicinal HαHN coupling constants, and NOE correlations indicative of fold) for peptides 1-4. A putative secondary structure map is provided for each sequence based on the analysis of the aggregate data.

We generated high-resolution structures of well-folded variants **3** and **4** by simulated annealing with NMR-derived restraints (Figure 14, Appendix A Table 9). The ensemble of low energy structures obtained showed excellent internal agreement (backbone rmsd < 0.75 Å); we used the lowest energy model from each ensemble in the structural analysis below. As the qualitative NMR data obtained for 1 (a C-terminal carboxamide) were entirely consistent with the previously reported structure of the prototype sequence as a terminal carboxylic acid,[120] we used the latter published coordinates (PDB 5KVN) as the basis for comparison. Analysis of the NMR structures obtained for **1**, **3**, and **4** shows that the turn modifications have no significant impact on the overall tertiary fold (Figure 14); however, closer comparison of the region of the helix-to-hairpin turn bearing the backbone modification reveals some subtle differences.

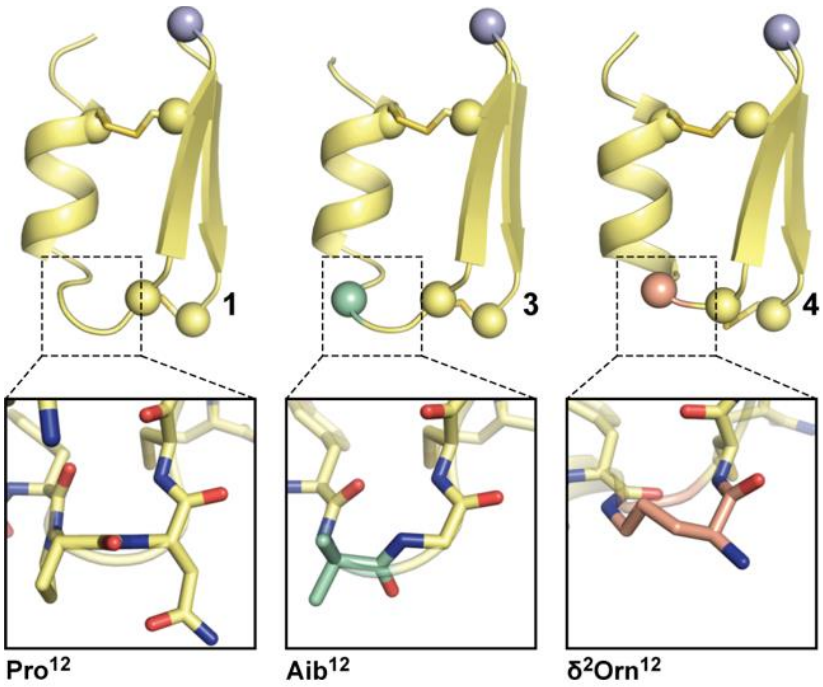


Figure 14. Comparison of the overall folding pattern and helix-to-hairpin turn from NMR structures of peptides 1 (PDB 5KVN), 3 (PDB 6E5H), and 4 (PDB 6E5I). Spheres in the cartoon representation indicate positions of a Cys residue or artificial backbone unit; residues are colored by type, matching the scheme in Fig. 10.

As mentioned above, the Pro^{12}-Asn^{13} segment in **1** adopts a canonical type-I turn. Examination of the same region in the NMR structure of **3** shows that the Aib^{12}-Gly^{13} segment adopts a type II turn. The high-resolution structure of peptide **4** shows that δ^2-Orn as a dipeptide replacement, while not able to be categorized as a canonical turn type, still positions the flanking residues in the helix-to-hairpin turn at similar positions as seen in **1**.

Taken together, the CD and NMR results obtained for **1**-**4** suggest that Aib-Gly and δ^2-Orn are well-accommodated modifications in the helix-to-hairpin turn and that D-Pro-Gly, although a well-established stabilizer of turns in β-hairpins, is unable to effectively mimic this region in disulfide-rich scaffold **1**. One possible origin of this finding is an incompatibility of the mirror image turn types (I', II') available to D-Pro with the helix-to-hairpin connection. This hypothesis is further supported by the observation that Aib-Gly, which has four turn types available to it, preferentially adopts a type-II turn in this context.

2.2.3 Design and Synthesis of Variants with Modified Helix, Turns, and Hairpin

The successful application of backbone modification in both turns of the prototype spurred us to seek variants with a greater density of artificial residue content. Given the short sequence, complex topology, and compact folded architecture of **1**, choices of potential modification sites were limited. Cys residues were left untouched to minimize the potential impact on disulfide bond formation. We incorporated Aib-Gly in the helix-to-hairpin turn based on the results above. For modification to the hairpin, we made side-chain retaining α→N-Me-α substitutions at cross-strand non-hydrogen bonding positions (Val^{16}, Val^{25}). Looking to the helix, prior work has shown that both β^3-residues and Cα-Me-α-residues like Aib are well accommodated as helix modifications in

proteins[67]—even stabilizing to the tertiary fold in some contexts.[68] As the differences in flexibility between Aib and β^3-residues impact folding thermodynamics,[149-150] we sought to compare these two residue types side-by-side in the present system. Guided by the above considerations, we designed two heterogeneous-backbone variants of prototype **1**. Peptides **5** and **6** each have ~25% artificial residue content and differ only in the composition of the helix segment. Peptide **5** incorporates side-chain-retaining $\alpha \rightarrow \beta^3$ substitutions in place of Lys^3, Glu^6 and Lys^9; $\alpha \rightarrow$Aib substitutions were made at the corresponding positions in peptide **6**.

Linear precursors to 5 and 6 were synthesized and purified as described above. We analyzed the time course of oxidative folding reactions and compared the results to prototype **1** (Figure 15, Figure 11) in an effort to ascertain if backbone modification led to a change in the rate and order in which the disulfide bonds form. Each reaction progressed similarly in terms of observable intermediates, with initial formation of a single intermediate followed by clean conversion to a single fully oxidized product. The rates of oxidative folding varied somewhat with backbone composition: oxidation of **1** was complete in 10 hours vs. 18 and 24 hours for heterogeneous-backbone counterparts **5** and **6**, respectively.

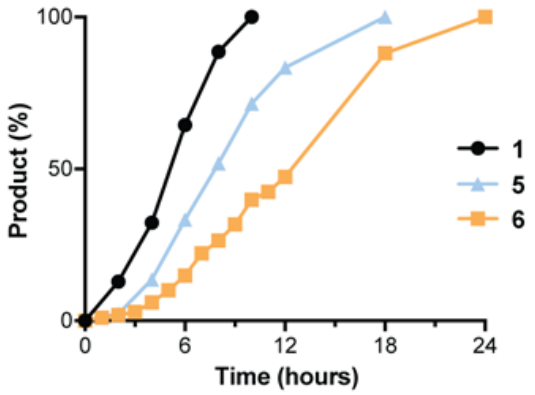


Figure 15. Time course of the oxidative folding to form 1, 5, and 6 from the corresponding fully reduced linear precursors. Reactions were held with 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3.

The above observations led us to seek a better understanding of the folding mechanism. Thus, we prepared parent peptide 1 by two parallel routes in which the disulfide bonds were formed sequentially. Two linear precursors were prepared with the same sequence as **1**: one with acetamidomethyl (Acm) protecting groups at Cys^{4} and Cys^{18} (**1a**) and the other with Acm groups at Cys^{14} and Cys^{27} (**1b**). Aerobic oxidation of these peptides produced two structural isomers with different disulfide bond connectivity, $\mathbf{1a_{ox}}$ and $\mathbf{1b_{ox}}$ (Figure 16).

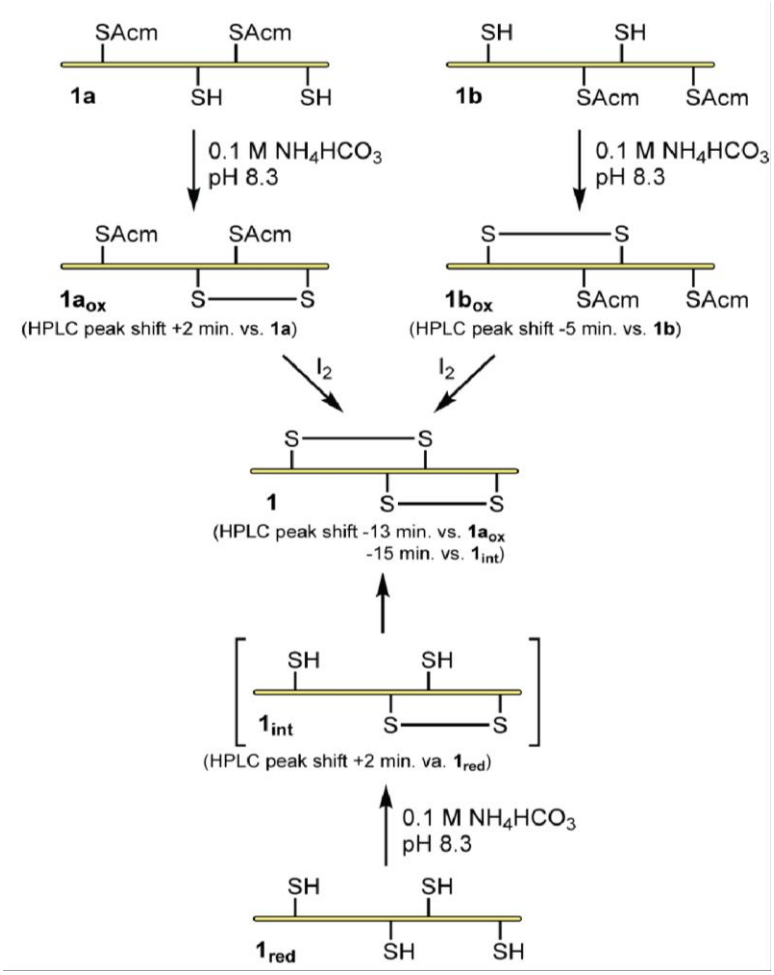


Figure 16. Scheme showing the preparation of 1 through sequential chemoselective disulfide bond formation from protected precursors 1a and 1b (top) as well as random aerobic oxidation of fully reduced precursor 1red (bottom). Key HPLC retention time changes in product vs. starting material for each reaction that informed the assignment of the structure of single-disulfide intermediate 1int are shown. Full HPLC chromatograms can be found in appendix A.

Comparison of the HPLC properties of the cyclic products to the corresponding linear precursors (Appendix A Figure 39) shows that formation of the Cys^{4-18} disulfide shifts retention time earlier, while formation of the Cys^{14-27} disulfide shifts retention time later. The above difference is due to the altered conformational preferences of $\mathbf{1a_{ox}}$ and $\mathbf{1b_{ox}}$ compared to their linear precursors due to the difference in disulfide-bond connectivity between the two species.

Recall, the single disulfide intermediate ($\mathbf{1_{int}}$) observed in the aerobic oxidation of the fully reduced precursor to **1** ($\mathbf{1_{red}}$) eluted later than starting material. Based on the above analysis, we assigned $\mathbf{1_{int}}$ as corresponding to initial closing of the hairpin to form the linkage at Cys^{14-27}. Of note, both $\mathbf{1a_{ox}}$ and $\mathbf{1b_{ox}}$ could be cleanly converted to **1** by treatment with iodine to cleave the Acm groups and form the second disulfide.

The oxidative folding to form **5** and **6** proceeded in a similar manner as **1**, with the observation of a single intermediate. In the case of **5**, the intermediate eluted slightly later than starting material as observed for the prototype, while the intermediate en route to **6** eluted slightly earlier than the linear precursor. Despite this difference, the fully oxidized final products of both foldamer variants eluted much earlier than starting material. We interpret this as indicating the extensive burial of residues in the hydrophobic core resulting from the juxtaposition of the helix and β-hairpin upon formation of the Cys^{4-18} disulfide bond second.

Collectively, the above results demonstrate that even extensively-modified heterogeneous-backbone foldamer analogues of a disulfide-rich peptide sequence can undergo oxidative folding with similar mechanism and efficiency as their native counterpart.

2.2.4 Structural Analysis of Variants with Modified Helix, Turns, and Hairpin

We assessed the folded structure of **5** and **6** by CD and NMR spectroscopy, as detailed above for the turn variants. CD spectra of **5** and **6** were similar to each other but qualitatively different than the spectrum of **1**, with a red shift of the minimum from 218 nm to ~222 nm and a decrease in signal magnitude (Figure 12). Qualitative interpretation of the NMR data is complicated somewhat in these sequences by the density of artificial building blocks; however, the results support the hypothesis that the secondary structure content and disulfide topology of the prototype sequence are both retained in the variants (Figure 41). High-resolution structures determined by simulated annealing with NMR-derived restraints bear out this hypothesis (Fig. 6A). Heterogeneous-backbone variants 5 and 6 show overall tertiary folds that are virtually identical to 1 (backbone rmsd values of 1.1 Å and 1.5 Å, respectively).

In both sequences, the added methyl moieties of N-Me-Val^{16} and N-Me-Val^{25} orient toward solvent as designed and do not disrupt the hairpin or the packing of the side chains at these sites into the hydrophobic core. A type-I turn conformation is preferred for the Aib-Gly motif in the helix-to-hairpin loop in **5** and **6**, matching prototype **1** but contrasting with the type-II turn preferred for this moiety in peptide **3**. The β^3-residues in **5** and the Aib residues in **6** are well accommodated in the helix, leading to a native-like fold and hydrogen-bonding pattern. Residues Asn^1 and Asp^2 are somewhat disordered in each structure; however, this region in peptide **6** does not appear to be as dynamic as in the prototype.

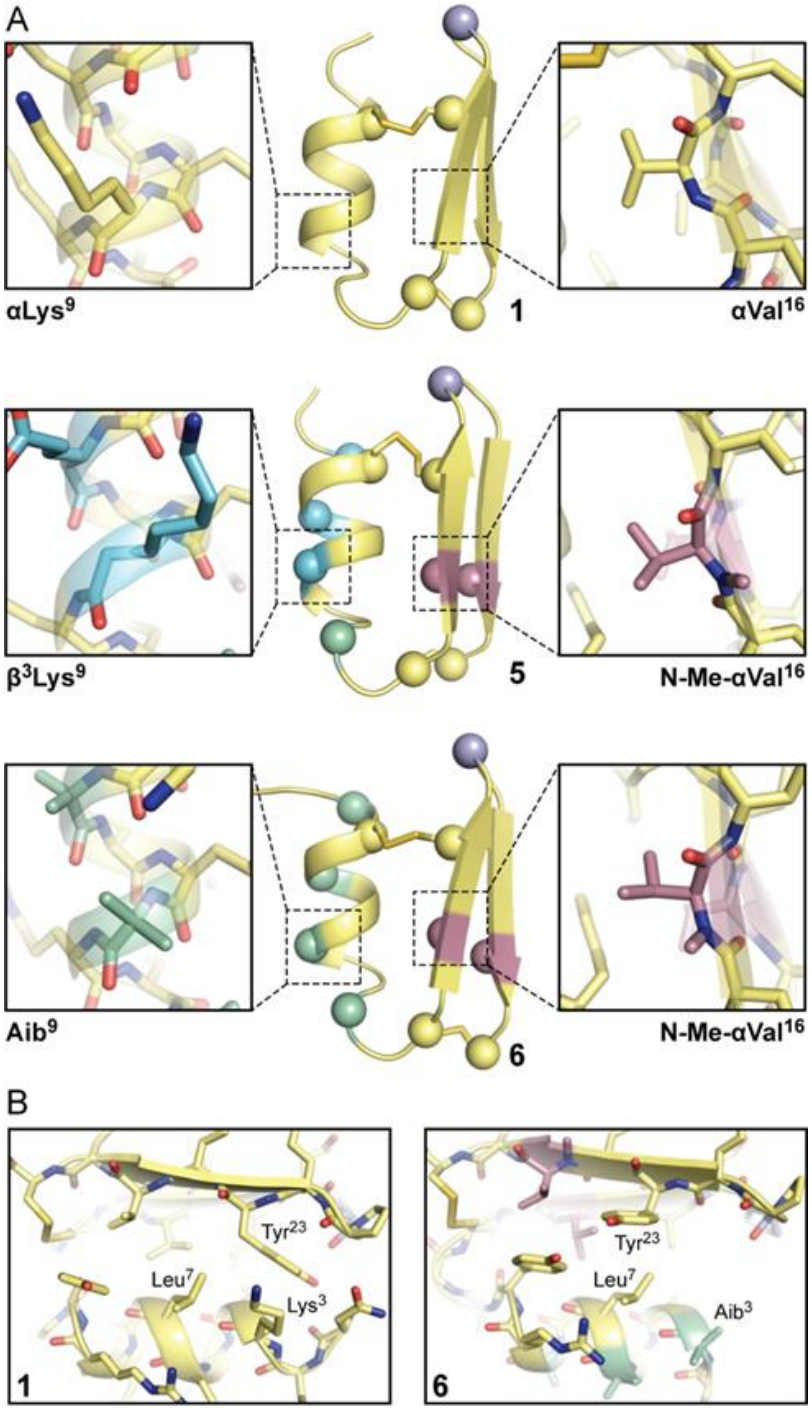


Figure 17. (A) Comparison of the overall folding pattern (center) and zoomed views of identical sites in the helix (left) and hairpin (right) from the NMR structures of peptides 1 (PDB 5KVN), 5 (PDB 6E5J), and 6 (PDB 6E5K). Spheres in the cartoon representation indicate positions of a Cys residue or artificial backbone unit. (B) Packing of Tyr23 in the structure of 1 and 6. Residues in both panels are colored by type, matching the scheme in Fig. 10.

In terms of subtle differences, inspection of the structure of **6** showed that the helix is juxtaposed at a slight angle in relation to the hairpin—an orientation not pronounced in the prototype peptide or other variants examined in this work. This unique packing may be partially caused by the loss of a side chain resulting from substitution of Lys^3 with Aib (Figure 17 B). In the structure of the prototype **1**, the side chain of Tyr^{23} orients toward the hairpin turn and forms hydrophobic contacts with the side chains of Lys^3. Replacement of Lys^3 with Aib eliminates this interaction, leading to a change in the orientation of Tyr^{23} in the structure such that it engages with Leu^7.

2.2.5 Proteolytic Susceptibility

One practical motivation for developing foldamer mimics of peptides and small proteins is the prospect of such modification reducing susceptibility to degradation by protease enzymes.[138] Disulfide-rich sequences as a class raise an interesting question in this regard, as one characteristic driving their use in biomedical contexts is high intrinsic proteolytic stability.[88] With the above considerations in mind, does backbone alteration provide a measurable functional benefit in an already topologically constrained scaffold? In order to probe this question, we subjected prototype **1** and variants **5** and **6** to digestion in the presence of trypsin (Figure 18, Figure 40).

Prototype **1** was degraded by the enzyme, albeit slowly, with ~20% of the starting material remaining after 24 h. In comparison of the proteolysis profiles for **5** and **6** under identical conditions, the heterogeneous-backbone variants were degraded to a lesser degree—~75% and ~40% remaining after 24 h for **5** and **6**, respectively. The observation of greater protection in the case of the helix modified with β^3 residues vs. Aib contrasted previous results from a side-by-side comparison of these residue types in an intrinsically disordered chain;[153] the behavior in the present

system may be due to subtle differences in hydrophobic core packing (Figure 18). Overall, the above results show that backbone modification provides an added degree of protection above and beyond that inherent to the disulfide-rich scaffold.

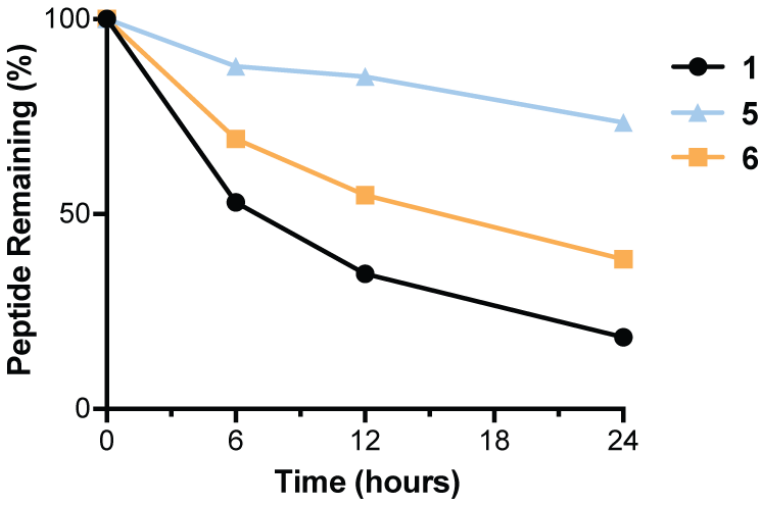


Figure 18. Time courses for proteolytic degradation of peptides 1, 5, and 6 in the presence of trypsin; each reaction was held with 100 µM peptide with 2.8 mg/mL enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5.

A portion of each remaining digested sample was subjected to both MALDI-TOF-MS and high-resolution ESI-MS (HRMS) to determine potential enzmatic cut sites for each peptide. Trypsin cleaves at the carboxyl terminus of charged amino acids arginine and lysine. Based on number of lysine and arginine residues in the sequence of peptide **1**, there are seven predicted cleavage sites located at residues Lys^{3}, Lys^{5}, Lys^{8}, Lys^{9}, Arg^{10}, Arg^{17}, and Arg^{22} respectively. Analysis of observed masses from MS reveal peptide fragments (Table 10) correlating with various cut sites on the peptide chain (Figure 19). For the peptide **1** sample, three putative fragments could be observed by MALDI, with *m/z* of 880.609, 1036.763, and 748.421 respectively, corresponding with peptide fragments Tyr^{11}-Pro-Asn-Cys-Glu-Val-Arg^{22} (calc. MW = 879.3 Da), Arg^{10}-Tyr-Pro-

Asn-Cys-Glu-Val-Arg^{17} (calc. MW = 1034.5 Da), and Cys^{4}-Lys-Glu-Leu-Lys-Lys^{9} (calc. MW = 747.4 Da) correlating with cut sites at after Lys^{3}, Lys^{9}, Arg^{10}, Arg^{17} (Figure 19). For peptide 5, two *m/z* were observed in HRMS: 413.20202 (*z* = 2) and 825.39080 (*z* = 1) indicating the presence of fragment Tyr^{11}-Aib-Gly-Cys-Glu-Nmev-Arg^{17} (calc. MW = 824.4 Da). Subjecting the same sample to MALDI-TOF MS showed the presence of peak with an *m/z* of 1304.050 implicating the presence of the peptide fragment Asn^{1}-Asp-Blys-Cys-Lys-β^{3}Glu-Leu-Lys-β^{3}Lys-Arg^{10} (calc. MW = 1302.7 Da). Collectively, this points to cut sites after Arg^{10} and Arg^{12} for peptide **5**. HRMS data for peptide **6** displayed several peaks, the first *m/z* of 377.88405 (*z* = 3), 566.3188 (*z* = 2), and 1131.62857 (*z* = 1) corresponding with the peptide fragment Asn^{1}-Asp-Aib-Cys-Lys-Aib-Leu-Lys-Aib-Arg^{10} (calc. MW = 1130.6 Da). Another set of peaks with *m/z* values of 413.20235 (*z* = 2) and 825.38954 (z = 1) point to the presence of the peptide fragment Tyr^{11}-Aib-Gly-Cys-Glu-NMEtVal-Arg^{17} (calc. MW = 1130.6 Da). Together, these values suggest that peptide **6** was cleaved after residues Arg^{10} and Arg^{17}.

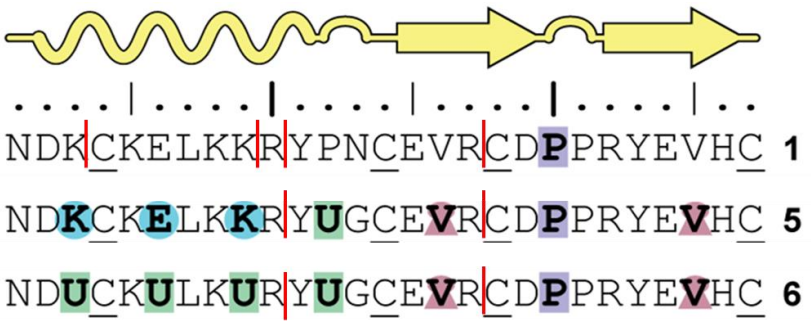


Figure 19. Observed trypsin cleavage sites for peptides 1, 5 and 6.

Out of the possible seven cut sites in peptide **1** only four were observed. It is precedented that trypsin cleavage may be affected by positioning of scissile residues in the hydrophobic core or secondary structure motifs of a protein or peptide.[154] The lack of observed cut sites may be due

to the presence of a defined secondary structures in peptide **1** making certain cut sites inaccessible to the enzyme. The lack of data implicating cleavage at all sites may also be affected by experimental factors including differences in ionization for the peptides under different MS sources, sample volume, and concentration subjected to analysis. From the observed information however, it can be gleaned that positions after residues Arg^{10} and Arg^{17} are particularly susceptible to tryptic degradation, as fragments corresponding with these cut sites are observed in all three peptides investigated. Interestingly, fragments implicating cut sites after residues 3 and 9 were not observed for variants **5** and **6**. In peptide **5** these positions come after a β^3-lysine residues, suggesting that placing a β^3 modified residue at a position before a putative cleavage site in this peptide scaffold imparts a degree of proteolytic protection against trypsin. This agrees with previous observations in our lab that β^3 residues at positions P1 prior to a cleavage site can confer proteolytic protection, with the caveat that in the precedent study the peptide sample had a random coil structure, and that the enzyme digest was performed with chymotrypsin instead of trypsin.[8] For peptide **6**, residues Lys^3 and Lys^9 are substituted with Aib, replacing long positively-charged side chains, as well as an α-hydrogen at this position. It may be hypothesized that the removal of the lysine side chain likely eliminates interactions with the conserved residue Asp 189 in the specificity pocket of trypsin and trypsin-like enzymes that aids in the recognition and positioning of peptide substrates for cleavage by the catalytic triad[155-156], thus decreasing the likelihood of scission at these positions in peptide **6**.

2.3 Conclusion

In summary, we have reported here that sequence-guided incorporation of artificial backbone units into a computationally designed disulfide-rich miniprotein can generate variants with ~25% artificial residue content and identical tertiary folds to the prototype. The helical region was effectively mimicked by replacing every third α-amino acid with a β^3 analogue or Aib, and N-methyl amino acid substitutions were effectively incorporated at cross-strand positions in the β-hairpin. Comparison of three artificial turn inducers in the helix-to-hairpin loop showed that Aib-Gly and δ^2-Orn were both effective mimics of the type-I turn in the prototype, while incorporation of D-Pro-Gly significantly disrupted the fold. All variants showed similar oxidative folding behavior as the starting sequence, converting cleanly to a single disulfide isomer under aerobic conditions. Investigations into the mechanism suggest the order of disulfide formation is the same in each case. Backbone modification provided a measurable improvement in stability to proteolytic degradation relative to the starting scaffold. Further MS investigations on these reactions show peptide fragments implicating cleavage after residues Arg^{10} and Arg^{17} for all variants and suggests that these positions may be particularly scissile to trypsin for this scaffold design.

Disulfide-rich peptides have garnered growing interest as starting points for the generation of pharmaceutical leads,[124] particularly given their potential role in the development of novel treatments for chronic pain.[157] The approach we demonstrate here adds to the arsenal of methods for tuning the properties of disulfide-rich sequences through chemical alteration, [109, 145-146] particularly given their potential role in the development of novel treatments for chronic pain.[157] Importantly, sequence-guided backbone modification of the type described is fully complementary to existing approaches such as disulfide replacement and cyclization. From the standpoint of protein tertiary structure mimicry by foldamers, the present work represents an important

expansion in the demonstrated scope of existing design principles through their application to a new target class (disulfide-rich sequences) as well as a new sequence origin (*de novo* designed rather than taken from nature).

2.4 Experimental Section

2.4.1 Peptide Synthesis and Cleavage

All reduced linear precursor peptides were synthesized via microwave-assisted Fmoc solid-phase methods (CEM MARS microwave reactor) on NovaPEG rink amide resin. The resin was swelled in DMF overnight prior to the initiation of synthesis. Amino acids were activated *in situ* by mixing N-α-Fmoc-protected amino acid (4 equiv), HCTU (4 equiv), and DIEA (6 equiv) in NMP (final concentration 0.1 M amino acid) unless otherwise specified. Microwave couplings were performed at 70 °C for 4 min. Fmoc-Cys(Trt)-OH, Fmoc-Cys(Acm)-OH, and Fmoc-His(Trt)-OH were coupled at room temperature for 30 min to prevent racemization. Fmoc-N-Me-Val-OH and Fmoc-Aib-OH were coupled using PyAOP (4 equiv) in place of HCTU. Microwave-assisted deprotection of the Fmoc group was achieved using 20% v/v 4-methylpiperidine in DMF at 80 °C for 2 min. Post-synthesis, the resin was sequentially washed with DMF, CH_2Cl_2 and methanol, then dried under vacuum. Cleavage of the peptide from resin was achieved by treatment with a cocktail composed of TFA/TIS/anisole/EDT/H_2O (85:4:4:4:3 by volume) for a period of 3 h. After filtration, peptide was precipitated with cold diethyl ether and pelleted by centrifugation at 6000 rpm for 2 min. Crude peptides were purified by preparative HPLC on a Phenomenex Jupiter Prep C_{18} column (particle size 10 μm, 300 Å pore size) using gradients between water and acetonitrile

containing 1% TFA (solvent A and solvent B, respectively). The identity of each peptide was confirmed by MALDI-TOF MS and purity was assessed by analytical HPLC with a Phenomenex Jupiter Analytical C_{18} column (particle size 10 μm, 300 Å pore size). Pure peptides were lyophilized prior to oxidative folding, as detailed below.

2.4.2 Oxidative Folding

Lyophilized peptides were dissolved in 0.1 M NH_4HCO_3 (pH 8.3) at a 0.25 mg/mL concentration and allowed to stir at room temperature open to air. The reaction was monitored by analytical HPLC and the identity of the intermediates and final products confirmed by MALDI-TOF MS. After completion, the reaction was quenched by addition of 50% acetic acid / H_2O to a pH ~4. The quenched samples were then diluted with H_2O, lyophilized, and purified by preparative HPLC as detailed above.

2.4.3 Circular Dichroism Spectroscopy

Stock solutions of each peptide were prepared in H_2O, and the concentration quantified UV-vis spectroscopy (ε_{276} = 3190 M^{-1} cm^{-1}). CD spectra were acquired on an Olis 17 UV/VIS/NIR spectrophotometer. Samples consisted of 60 μM peptide in 10 mM sodium phosphate buffer at pH 7.0 at 20 °C. Each spectrum was acquired with wavelength scans from 200-260 nm in 1 nm increments. Sample data were exported in ASCII format and visualized in GraphPad. A 2nd order polynomial smoothing function was applied to the nearest 5 neighboring points for each spectrum.

2.4.4 Oxidative Folding Kinetics Analysis

Lyophilized linear precursors of peptides **1, 5** and **6** were diluted to a concentration of 0.25 mg/mL NH_4HCO_3 buffer (pH 8.3) and stirred open to atmosphere at room temperature. At each time point, 15 μL of the reaction mixture was aliquoted and quenched with 5 μL of 50% acetic acid. Evolution of the oxidation reaction was monitored by analytical HPLC using a Jupiter Analytical C_{18} column (particle size 5 μm, 300 Å pore size). Each sample was run at unique gradients used in the initial purification of the peptides (peptide **1**: 10-30 %B, peptide **5**: 15-30 %B, and peptide **6**: 20-33 %B over 40 min). Identity of the fully oxidized final product was determined by HPLC retention time shift and MALDI-TOF MS. The final time points were determined by the complete disappearance of the linear precursor and the single disulfide intermediate.

2.4.5 Iodine Oxidation of $\mathbf{1a_{ox}}$ and $\mathbf{1b_{ox}}$

Purified and lyophilized $\mathbf{1a_{ox}}$ or $\mathbf{1b_{ox}}$ were dissolved in a solution of 4:1 acetic acid / H_2O at a concentration of 2 mg/mL. To this was added a solution of 0.1 M iodine in acetic acid (25 equiv). The reaction was vortexed for 10 min and then quenched by addition of 0.2 M ascorbic acid in water until the solution turned clear. A small sample of the quenched reaction was desalted by a C_{18} resin ZipTip, and the identity of the fully-oxidized peptide species confirmed by MALDI-TOF MS. The oxidized peptides were diluted with H_2O and lyophilized prior to being purified by preparative HPLC as detailed above.

2.4.6 NMR Data Acquisition, Analysis, and Structure Determination

NMR samples were prepared with 1-5.7 mM peptide in 1:9 D_2O/H_2O and 0.2 mM DSS at pH 4 (uncorrected). Spectra were obtained on a Bruker Avance 700 MHz spectrometer at 298 K. Two dimensional experiments were comprised of NOESY (200 ms mixing time, 32 scans), TOCSY (80 ms mixing time, 16 scans), and COSY (16 scans). Spectra were acquired with 2048 data points in the direct dimension and 512 data points in the indirect dimension. Each spectrum was processed in TOPSPIN, using DSS as an internal standard for calibration. NMR data were analyzed using the NMRFAM-SPARKY software package,[158] and the proton resonances in each sequence assigned by standard methods. H_α chemical shifts for canonical α-residues were compared against random-coil chemical shift values generated through the Poulsen IDP/IUP Random Coil Chemical Shift Server [https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/][159] to generate a chemical shift index (CSI) plots for each peptide.[152]

Determination of high-resolution structures was carried out by simulated annealing using ARIA (Ambiguous Restraints for Iterative Assignment, version 2.3)[160] in conjunction with CNS (Crystallography & NMR System, version 1.2).[161] The software was modified to handle nomenclature and geometric restraints for the unnatural residues, with parameter and topology definitions based on analogous atom types already present for the natural α-residues. Base parameters for each ARIA run were modified from program defaults to improve model quality and convergence, as previously described.[162] Each sequence showed unambiguous long-range NOEs consistent with the native disulfide topology, so corresponding covalent bonds were introduced in the calculation. A set of initial H-bond restraints was generated based on qualitative inspection of the NOESY spectrum: existence of long-range $H_N(i) \rightarrow H_N(j)$ cross peaks for the hairpin, strength

of $H_N(i) \rightarrow H_N(i+1)$ cross peaks and existence of $H_\alpha(i) \rightarrow H_N(i+3)$ cross peaks for the helix. Backbone φ dihedral restraints were prepared based on measured $^3J_{H\alpha\text{-}HN}$ coupling constants for well-resolved amide doublets in the 1D ^{1}HNMR spectrum (φ = -65°±25° for $J \leq 6.0$ Hz and φ = -120°±40° for $J \geq 8.0$ Hz). NOE distance restraints were generated automatically by ARIA in iterative fashion over the course of the calculation, starting from an unassigned set of integrated peaks from the NOESY spectrum and a list of proton chemical shift values. The ensemble of 10 lowest energy structures resulting from the above calculation was then used as the input structure for a second ARIA run. Parameters used were the same as above, with additional restraints added for any new H-bonds observed in the first ensemble. The final set of 10 lowest energy structures resulting from the second run was taken as the NMR ensemble for that peptide. The lowest energy entry from each ensemble was used for structure comparisons, figure generation, and analysis presented in the main text. Ensemble coordinates and additional experimental data are deposited in the PDB (accession codes **3**: 6E5H, **4**: 6E5I, **5**: 6E5J, **6**: 6E5K) and BMRB (accession codes **3**: 30496, **4**: 30497, **5**: 30498, **6**: 30499).

2.4.7 Proteolytic Degradation by Trypsin

Stock solution of trypsin was prepared by dissolving 2.8 mg bovine trypsin (MW 23,290 Da) in 1 mL of 0.1 M NH_4HCO_3 solution with 1 mM $CaCl_2$ (pH 8.5). Lyophilized samples of peptide **1**, **5**, and **6** were dissolved in deionized H_2O, and their concentrations determined by UV spectroscopy (ε_{276} = 3190 M^{-1} cm^{-1}). Each reaction contained 10 μM enzyme and 100 μM peptide in 0.1 M NH_4HCO_3 (pH 8.5), with a total volume of 120 μL. Reactions were mixed and allowed to incubate at room temperature. At 6, 12, and 24 h time points, 40 μL of the reaction solution was quenched with 5 μL of 10% TFA and diluted by addition of 55 μL of 0.1 M NH_4HCO_3. The

quenched samples were analyzed by analytical HPLC (90 μL injection) using a Jupiter Analytical C_{18} column (particle size 5 μm, 300 Å pore size) and a 10-30% B gradient. Identity of the peptides of interest were confirmed by MALDI-TOF MS, and degradation monitored by integration of the corresponding peak on the HPLC chromatogram. Peak areas were normalized to the area of the starting material and the amount of peptide remaining over time was plotted via GraphPad Prism. Identity of the peptide fragments resulting from enzymatic degradation was examined by High Resolution ESI-MS and MALDI-TOF MS.

3.0 Heterogeneous Backbone Modification and Functional Mimicry of Lasiocepsin

3.1 Introduction

The excessive use of antibiotics has contributed to the spread and prevalence of antibiotic resistance in microorganisms, which has become a serious concern to public health.[163] These medications are frequently administered to patients with active infections or as a means of prophylaxis, inducing a selective pressure for bacteria to evolve certain mechanisms of resistance.[163-165] As such, resilient strains predominate and are increasingly more difficult to eradicate with available treatments. Due to the paucity of new antibiotic agents there is increasing concern that the arsenal currently available to battle infection may not be sufficient in the future as the number of clinically significant antibiotic resistant bacterial strains rise.[166] As such, researchers have looked for other sources of potential microbicidal agents to complement the existing compendium of antibiotics available.

A group of compounds that have been investigated for their broad range bactericidal effects are antimicrobial peptides (AMPs), low molecular weight proteins that serve in host immunity of numerous flora and fauna.[167] These peptides are typically positively charged and amphipathic allowing them to be soluble in both aqueous milieu and in hydrophobic environments.[168] These properties allow AMPs to interact with the cell membrane of target microbes and exert cidal effects through membrane disruption. Specifically for bacteria, it is proposed that AMPs contact bacterial cell membranes and form amphipathic secondary structures such that the positively charged residues interact with the negatively charged phospholipid heads, while the hydrophobic regions embed into the lipid core of the membrane, ultimately resulting in membrane permeabilization and

cell lysis.[169] Though the exact mechanism of membrane permeabilization is contested, four popular models have been proposed. The barrel stave model posits that peptides insert perpendicular to the membrane, allowing for sequential assembly of a pore through the cell membrane. Similarly, the toroidal pore model also posits a similar mechanism but with the caveat that formation of the pore causes bending of the cell membrane such that the pore is lined with both phospholipid head groups and peptide. Lastly the carpet model supposes that accumulation of peptides on the membrane surface leads to pinching of the membrane to form small micelles, disrupting membrane continuity. Other posited AMPs mechanisms of action involve translocation across the cell membrane, followed by downstream impacts on biochemical pathways including DNA and protein synthesis (Figure 20).

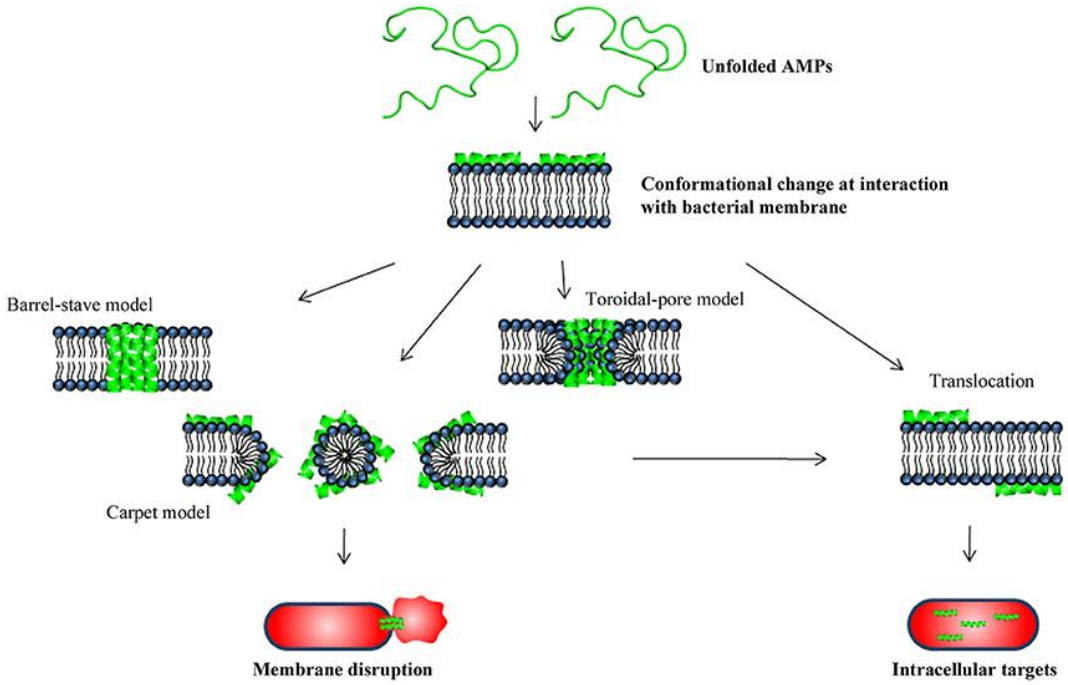


Figure 20. Proposed models for AMP activity. Adapted from Mahlapuu, M.; Håkansson, J.; Ringstad, L.; Björn, C., Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. Frontiers in Cellular and Infection Microbiology 2016, 6, 194.

Unlike the many small molecule antibiotics that target a particular cellular process, such as cell wall elongation, AMPs display a general membrane-disrupting mechanism, which can potentially be leveraged to attack even bacteria that are resistant to traditional antibiotics.[170] Clinically important drug-resistant bacteria include members of the *Enterococcus*, *Staphylococcus*, *Klebsiella*, *Acinetobacter*, *Pseudomonas* and *Enterobacteriaceae* groups, referred to as ESKAPE pathogens. These strains are common causes of hospital-acquired infections, are multi-drug resistant, and are associated with the high mortality among patients.[171] As such, a task of finding antimicrobials for these bacterial strains is an important area of research in the scope of public health.

Early work has shown that engineered *de novo* cationic AMPs can have had success in both being effective in targeting members of the ESKAPE pathogens groups and that resistance to these bioactive peptides were harder to mount even at sublethal concentrations.[172-173] Paralleling these efforts, researchers have investigated a variety of natural host-defense peptides such as disulfide-rich defensins from the innate immune system of animals and plants as potential sources of AMPs to combat these drug-resistant strains.[115, 174-175] Bioactive peptides from animal venoms such as scorpions can have broad range antimicrobial activity[176], and rational modification of these sequences can be used to alter properties making them effective against certain ESKAPE strains *in vitro*[177]. Additionally, it has been demonstrated that modification of a helical antimicrobial peptide through incorporation of D-amino acids had effect on target selectivity and proteolytic stability.[178-179] Sequence modification holds promise as several known AMPs in clinical use including vancomycin and polymyxin B contain nonproteinogenic amino acids including N-methylated and N-formylated residues[167]. These precedents highlight the capability of modulating

AMP properties, which inspired us to ask the question – Could a bioinspired heterogeneous backbone analog of a naturally-occuring AMP recapitulate native biological properties?

We decided to explore this question in the context of Lasiocepsin (Las), a 27 amino acid peptide derived from the venom of the bee *Lasioglossum laticeps*. Las contains two helices connected by a 6-residue loop and stabilized by two disulfide bonds between Cys^{8}-Cys^{25} and Cys^{17}-Cys^{25} (Figure 21).

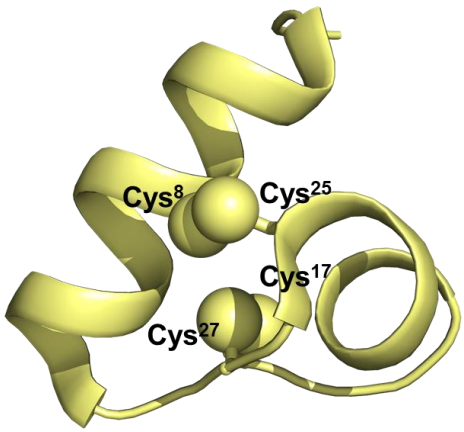


Figure 21. Structure of native lasiocepsin (PDB 2MBD). Disulfide bonds are depicted as spheres.

The larger helix extends from residues Arg^{4} to Lys^{13} and is positioned orthogonal to the smaller helix which spans residues Pro^{20} to Val^{24}. Las does not have notable sequence homology with other antimicrobial peptides. Of interest from the standpoint of protein mimicry work outlined in Chapter 2, Las is noteworthy of having a defined tertiary fold unlike many AMPs which rely on secondary structure for activity. This peptide has been shown to exhibit bactericidal activity against Gram positive bacteria such as *Bacillus subtilis*, *Staphylococcus aureus*, and *Micrococcus luteus* as well as Gram negative bacteria such as *Escherichia coli* and *Pseudomonas aeruginosa*. Interestingly, it does not display significant hemolytic activity at 200 μM. The broad-range

antimicrobial property of Las made it an ideal candidate to examine the functional consequences of heterogeneous backbone substitution.

3.2 Results and Discussion

3.2.1 Las NH_2 and Design of First-Generation Las Analogs

Lasiocepsin bearing a C-terminal amide group (Las NH_2) was synthesized to serve as a control alongside a series of first-generation analogs (Las var 1-3). Las NH_2 contains the same amino acid sequence as the native peptide but was found to display increased antimicrobial activity[143] making it a good point of comparison against its proteomimetic counterparts. A series of first-generation analogs were designed to vary in the density of artificial monomers from ~30% (Var 1) to 37% (Var 2 and 3). The diversity of artificial residue types was also progressively increased from Las Var 1 to 3, with Las Var 1 bearing mostly β^3 substitutions to retain the side chain each residue in the helices. Las var 2 and 3 contain increasingly more complex substitutions including Aib, D-, and the β-cyclic amino acid 2-aminocyclopentane carboxylic acid (ACPC). These modifications alter the side chains of replaced residues. In precedent experiments, it was found that helical features of Las were retained upon binding to sodium dodecyl sulfate micelles.[180] Furthermore, mutants that lacked or had incorrect disulfide bond connectivity exhibited significantly reduced activity or were completely inactive.[143] Thus, each successive peptide in these generation 1 variants were designed with the intention of rigidifying the helical domains (Figure 22). Lastly, D-Ala was incorporated at the N-terminus of all three sequences in an effort to impart proteolytic resistance against exopeptidases.

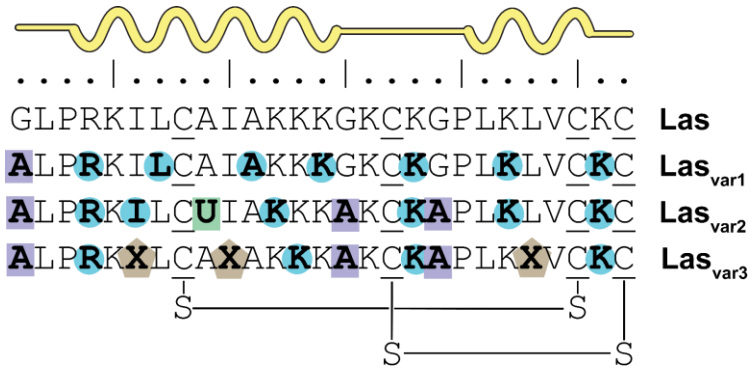


Figure 22. Sequences of first generation Lasiocepsin variants.

3.2.2 Synthesis and Regioselective Disulfide Formation

All peptides were synthesized using manual or automated microwave-assisted solid phase peptide synthesis on Rink Amide Resin affording a C-terminal amide to all peptides. A sequential method for generating disulfide bonds was adopted from precedent studies by Monincová as statistical oxidation of the linear sequence was determined to produce two inactive conformational isomers.[143] For each sequence, Cys^{17} and Cys^{27} were protected with Trt to afford the generation of the initial disulfide bond, and Cys^{8} and Cys^{25} were protected with Acm groups to produce the second disulfide by iodine oxidation. The linear peptide cleanly oxidized to a monodisulfide intermediate through aerobic oxidation in ammonium acetate (pH 7.8). This accompanied a distinct retention time shift via HPLC. Generation of the second disulfide was achieved through iodine-mediated oxidation (Figure 23) of the monodisulfide intermediate for 1.5 hours. For each peptide synthesized, MALDI of final iodine-oxidized products revealed the presence of the desired bisdisulfide species, with a minor product showing the addition of a single Acm group. This Acm side product is documented by Monincova and colleagues as an Acm transfer from Cys to one of

several Lys residues in the Las sequence.[143] Despite this, the desired products with ≥ 95% purity were able to be purified by analytical HPLC.

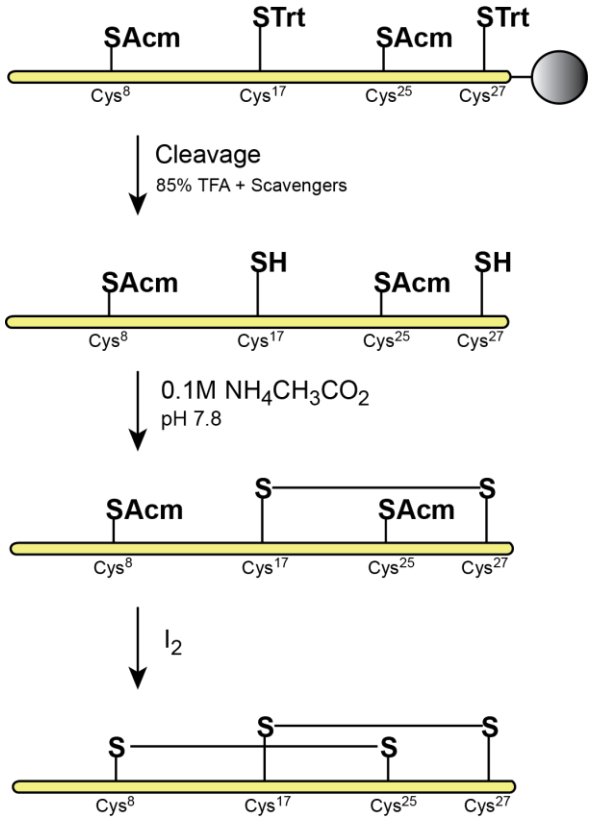


Figure 23. Synthetic scheme of Las variants.

3.2.3 Microbiological Assay of First-Generation Lasiocepsin Variants

In a collaborative effort with the lab of Dr. Peter Di at the University of Pittsburgh School of Public Health, the first generation of Lasiocepsin variants were subjected to microbiological assays against drug resistant ESKAPE pathogens comprising of *Enterobacter faecalis*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Escherichia coli* (a representative of the Enterobacteriaceae family), commonly

found in acquired nosocomial infections. The control peptide Las NH_2 exhibited bactericidal activity for the panel of ESKAPE pathogens tested, with single digit for μM minimum inhibitory concentration for all strains (Table 2). The native sequence was especially active towards *E. faecium*, *A. baumannii* and *E.coli* with a minimum inhibitory concentration (MIC) of 1 μM for each respective strain and moderately effective toward methicillin sensitive *S. aureu*s (4 μM MIC). Las Var 1, 2, and 3 did not exhibit any measurable bioactivity towards the tested ESKAPE strains evidenced by high MIC values.

Table 2. Minimum inhibitory concentration of first-generation Las variants againts clinically-relevant ESKAPE pathogens

	MIC (μM)					
	Gram Positive		Gram Negative			
	E. faecium	*S. aureus (MSSA)*	*K. pneumoniae*	*A. baumannii*	*P. aeruginosa*	*E. coli*
Las NH_2	1	4	8	1	8	1
var 1	>32	>32	>32	>32	>32	>32
var 2	>45	>45	NT	>23	>23	NT
var 3	>32	>32	>32	>32	>32	>32

NT = Not Tested

3.2.4 NMR of Las Var 1

The observation that simple substitutions of select α-amino acid residues with their β^3 counterparts in Las var 1 significantly lowered bioactivity compared to the native peptide was surprising. These substitutions retained the same side chain and charge as the prototype peptide. Moreover, it is precedented that an ααβαααβ pattern can recapitulate a helical fold in helix bundles as well as globular protein contexts.[149, 181] Thus, it was initially postulated that this variant would be effective mimic of Las NH_2. Given lack of bioactivity, we became curious to see if Las var 1 had a similar folded structure as the native peptide. Thus, we subjected Las var 1 to

multidimensional NMR spectroscopy (pH 7 at 300 K). We acquired 1H, TOCSY, NOESY and COSY experiments which allowed us to assess the chemical shifts and backbone connectivity of each residue, as well as their spatial interactions. From the published assigned chemical shifts (BMRB entry 19396), we generated a Hα chemical shift deviation (CSD) plot of the native Lasiocepsin peptide (Figure 24). Chemical shift deviation is contingent on observed values of Hα chemical shifts in model peptides that are determined to be in a random coil.[182] These values are not currently determined for backbone modified amino acids especially in unnatural sequences that adopt discrete conformations. This is further complicated by increased density of substitution of unnatural amino acids, as well as nearest neighbor effects which could affect the chemical shift of adjoining amino acids.[152] This underscores the need for a complementary method for determining peptide structure, which is the impetus for work described in Chapter 4 of this dissertation. We used CSD plots for a qualitative evaluation of peptide fold in our analysis of the synthesized Las variants.

CSD values for the native are consistent with its known fold, evidenced by successive negative values from residues 3-7 and 9-14 for the N-terminal helix, and residues 21 to 25 for the C-terminal helix. Residues 15-20 which comprise the unstructured loop of Las contains a mix of positive and negative chemical shifts deviations.

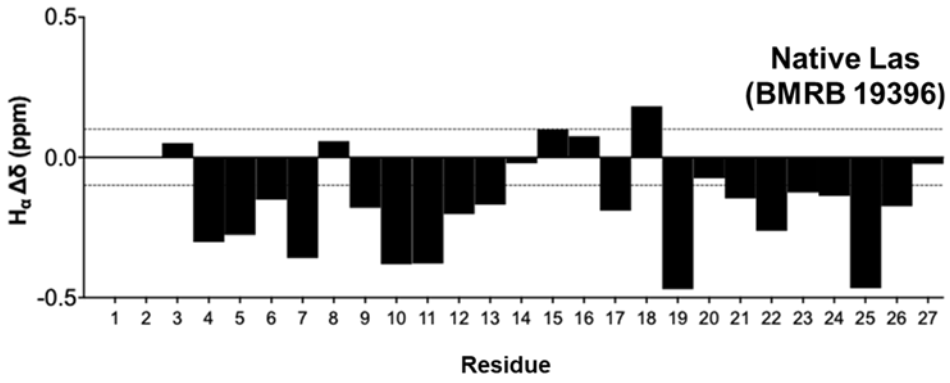


Figure 24. Chemical shift deviation plot from random coil values of native Lasiocepsin as derived from published BMRB chemical shifts.

Hα chemical shift deviation plots of Las var 1 differ qualitatively from the native. No region of contiguous negative values is seen along the entire chain, suggesting that this analog did not have the defined N-terminal and C-terminal helical structures of the native (Figure 25). The Hα chemical shift deviations at residues 1, 4, 7, 11, 14, 18, 22 and 26 were not calculated since the suitable random coil values fort these unnatural residues are currently undefined. Taken together the NMR and bioactivity results suggest that Las var 1 is not folded.

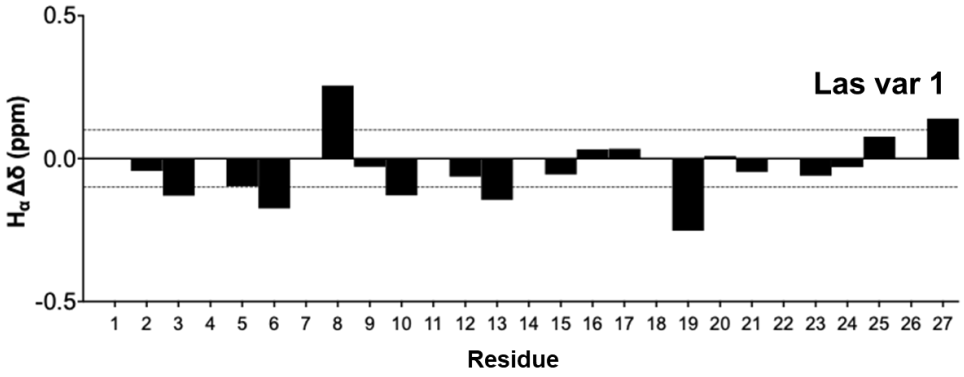


Figure 25. Hα chemical shift deviation plot of Las var 1.

3.2.5 Design of Second-Generation Las Variants

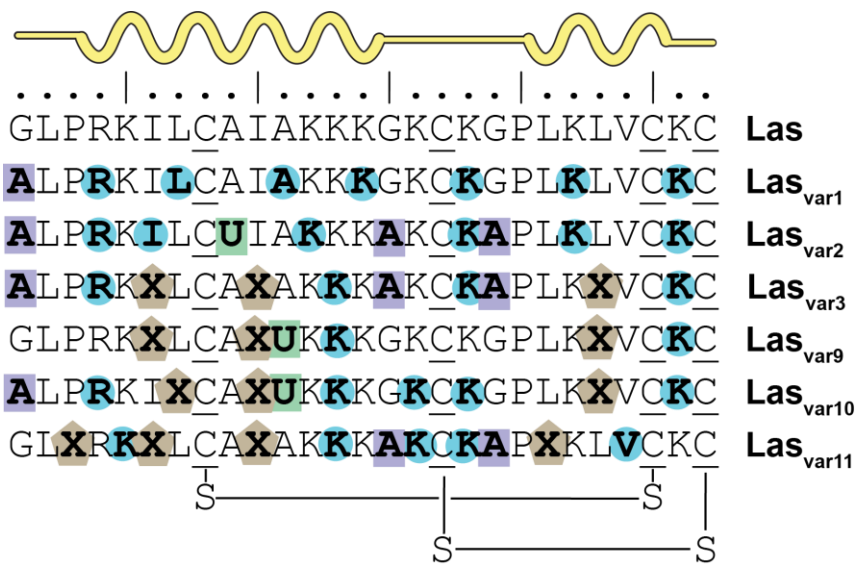


Figure 26. Sequence of generation 1 and 2 Las variants.

Motivated by the results of the first generation, we designed a second series of Las analogs to test specific hypotheses related to the role of the different regions of the native sequence in imparting bioactivity as well as screen different strategies for backbone alteration for one or more able to maintain the native fold and activity. Given that Las var 1 did not exhibit measurable helical content, we tried to incorporate substitutions that would rigidify the N-terminal and C-terminal helices. Given the precedent of helix rigidification using 2-aminocyclopentane carboxylic acid (ACPC)[150], we incorporated it as an additional building blocks in second generation designs. Substitution with the unnatural amino ACPC was hypothesized to rigidify the helical regions, especially the N-terminus, as this region was previously reported to be important in bioactivity in the prototype.[180]

3.2.6 Design of Second-Generation Las Variants

The first of the second-generation analogs, Las var 9 is reminiscent of the design of Las var 3, but with the first five N-terminal residues left unmodified to test if these residues are needed for bioactivity. ACPC was substituted in the same positions as in Las var 3 to probe if substitution at these positions will maintain the native helical structure. Moreover, the loop region was left unmodified to establish if the native sequence and length is necessary for proper fold. There were no modifications in the C-terminal helix relative to Las var 3 (Figure 22).

Las var 10 is similar in design to Las var 3 but the ACPC was shifted to position 7 such that every third amino acid from position 4 to 10 is a β amino acid. This is a pattern that we found to be effective in mimicking helices in precedent investigation described in Chapter 2.[183] Moreover, Lys^{15} and Lys^{18} were substituted with their β^3 counterparts to determine if elongation of the loop region would enable orthogonal packing of the modified helices. The C-terminal helix was also left unaltered, just as in Las var 3 and 9. D-Ala was again substituted at position 1 to impart proteolytic stability (Figure 26).

Las var 11 was designed with a greater percentage of substitution compared to other peptides in the series. Pro3 was substituted with ACPC, as this position resides in the region that was found to interact with SDS micelles[24] and we wanted to assess if extending the helix, without the kink induced by proline would still be conducive to bioactivity. ACPC was also substituted at position 6 and 10 to examine if the N-terminal helix could be supported by these two rigid residues. Additionally, Gly^{15} and Gly^{19} of the loop region were substituted with D-Ala to impart resistance against endopeptidases. In the native domain, these two Gly residues exist in the D-residue region of the Ramachandran plot. The C-terminus was modified to shift the two β residues that are located within the defined C-terminal helix region in the native peptide.

3.2.7 Microbiological Assays of Second-Generation Las Variants

Purified Las var 9, 10 and 11 were subjected to the same ESKAPE pathogen assays as with the first-generation peptides. Out of the three peptides tested, two were found to be bioactive. Briefly, Las var 9 exhibited bioactivity toward *E. faecium* and *A. baumannii* with MIC values of 4 μM for each respective strain. Variant 11 exhibited greater bioactivity towards the same bacterial strains with a 2 μM MIC. Interestingly, the same two strains were most susceptible to control Las NH_2, with MIC values of 1 μM toward *E. faecium* and *A. baumannii* (Table 3) respectively. Interestingly, *E. faecium* and *A. baumannii* differ in cell wall composition, with *E. faecium* categorized as Gram positive and *A. baumannii* as Gram negative Both Las var 9 and 11 also exhibited a degree of bioactivity towards *E.coli* (MIC 8 μM) albeit less than *E. faecium* and *A. baumannii*. Interestingly, these three strains were also observed to be the most susceptible to the native peptide, suggesting that the substitutions made in Las var 9 and Las var 11 may help in both the recapitulation of function as well as increased selectivity towards select bacterial targets. This is corroborated by Las var 10 was found to be non-bioactive, with MIC values greater than 32 μM for all strains tested.

Table 3. Minimum inhibitory concentration of first and second-generation against clinically-relevant ESKAPE pathogens

	MIC (uM)					
	Gram Positive		Gram Negative			
	E. faecium	*S. aureus (MSSA)*	*K. pneumoniae*	*A. baumannii*	*P. aeruginosa*	*E. coli*
Las NH_2	1	4	8	1	8	1
var 1	>32	>32	>32	>32	>32	>32
var 2	>45	>45	NT	>23	>23	NT
var 3	>32	>32	>32	>32	>32	>32
var 9	4	>32	>32	4	>32	8
var 10	>32	>32	>32	>32	>32	32
var 11	2	>32	>32	2	16	8

NT = Not Tested

3.2.8 Hemolytic Assay of Las Variants

Antimicrobial peptides have promising characteristics to be developed for systemic applications, but their progress is usually hampered by their hemolytic and cytotoxic activities.[184-186] Native Lasiocepsin was found to have low hemolytic activity, making it attractive as a potential antibacterial agent.[143, 180] However, it is precedented that alteration of peptide sequences of AMPs can have pronounced effects on both hemolytic and cytotoxic activities.[186] In collaboration with the lab of Dr. Peter Di of the University of Pittsburgh Department of Public health, we endeavored to determine the effects of backbone modifications on Lasiocepsin, we subjected the select analogs to both hemolytic and murine macrophage RAW264.7 cell cytotoxicity assays. RAW264.7 are models for host leukocytes and serve as standards for assessing the capability of antimicrobials in harming host cells. We included Las var 1 and 3 as representatives of nonbactericidal peptides from the first generation as they bear the simplest and most complicated modifications in this series. It was found that all the backbone modified Lasiocepsin variants exhibited less hemolysis at all tested concentrations when compared with Las NH_2. At the highest concentration tested (128 μM) Las NH_2 exhibited ~73-79% hemolysis while all the analogs tested exhibited less than ~22% hemolysis. This result is interesting as the bioactive Las variants 9 and 11 are much less hemolytic at greater peptide concentrations, than the native sequence (Figure 27).

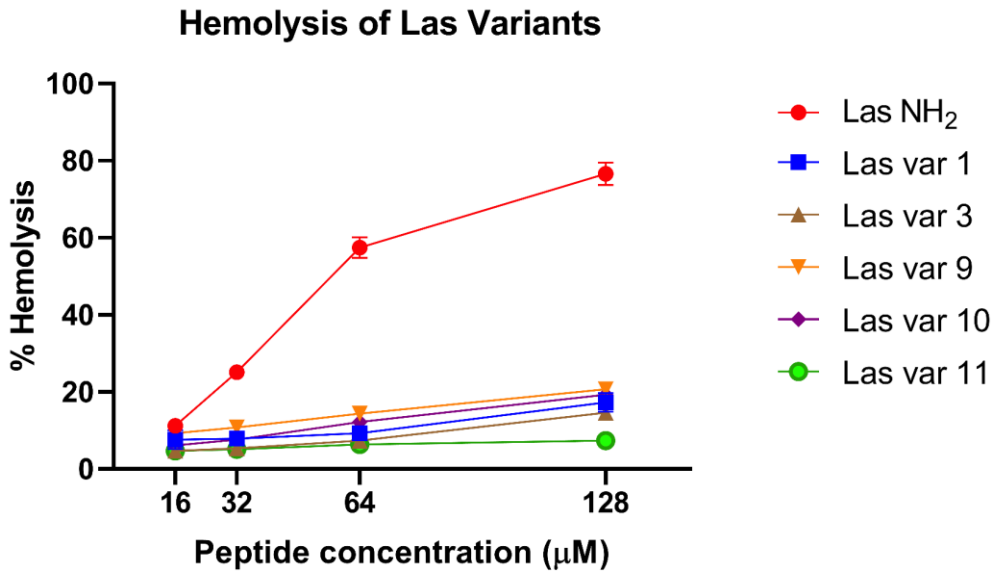


Figure 27. Hemolysis of Lasiocepsin variants at various concentrations.

3.2.9 Cytotoxicity Assay of Las Variants

Las NH_2 exhibited concentration dependent cytotoxicity, with significant RAW264.7 cell death occurring after 4 hours of incubation using 64μM of peptide. In contrast, the Lasiocepsin variants generally exhibited minimal cytotoxicity, at concentrations 8, 16 and 32 μM which is comparable to the untreated positive control. Las var 10 showed greater cytotoxicity compared to the other variants tested, but still less than the native peptide. Most samples incubated 64 μM of peptide, showed increased cell death compared to the positive control, although Las var 11 values are within error of the positive control.

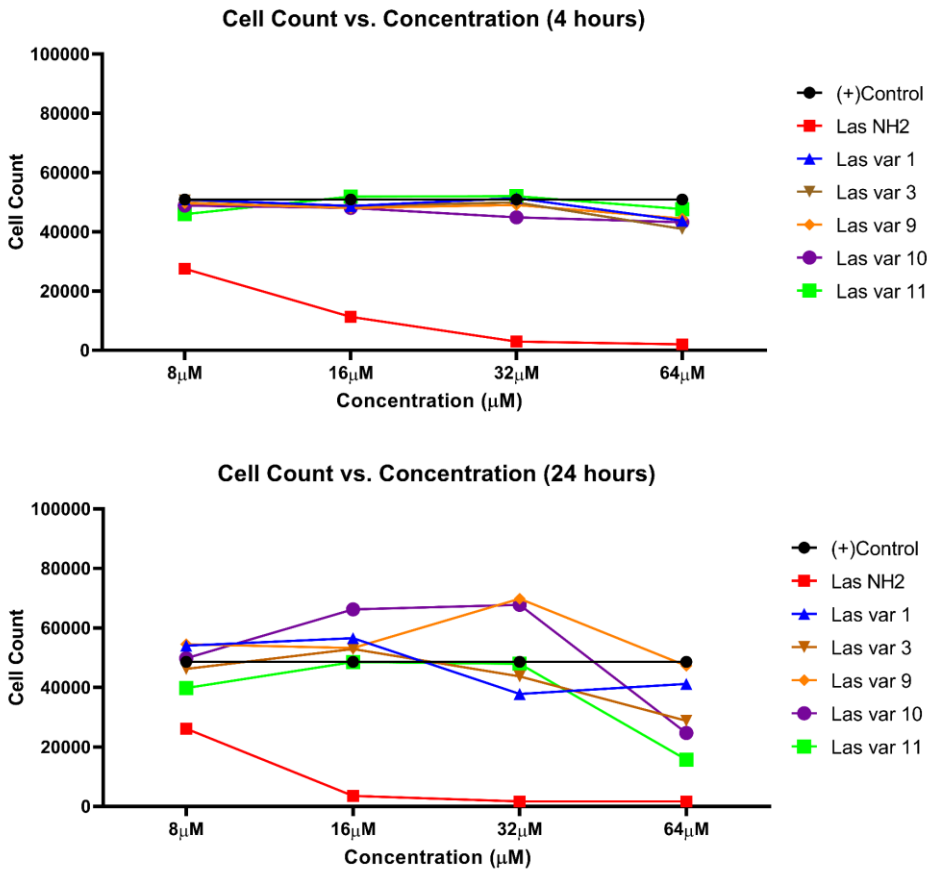


Figure 28. Cytotoxicity of first and second generation Lasiocepsin variants at 4- and 24-hour incubations.

After 24 hours, samples containing Las NH_2 displayed lower cell counts than was observed at the 4-hour time point, with significant toxicity upon incubation with 16 μM peptide. Interestingly, during this time point, samples incubated with 8 and 16 μM of the nonbactericidal variants (Las var 1, 3 and 10) generally displayed comparable or even greater cell numbers compared to the positive control, suggesting that these peptides are non-cytotoxic at lower concentrations. These bioactive variants however displayed lower cell numbers at higher (32 and 64 μM) peptide concentrations apart from Las var 10 which displayed cell counts greater than the control at 32 μM concentration. Las var 9 displayed cell counts greater or comparable to that of

control at 32 and 64 μM. At 32 μM, Las var 11 displayed similar cell counts as the control, but significantly less cell numbers at 64 μM. Taken together, these observations suggest that both bactericidal variants are non-cytotoxic except at higher concentrations. Moreover, these two variants are significantly less cytotoxic than Las NH_2 at all concentrations tested. Lastly, the observed differences in cell counts from 4 to 24 hours for 64μM samples indicate that this cytotoxicity may be time dependent.

3.2.10 NMR of Las Var 9 and 11

Lasiocepsin variants 9 and 11 were subjected to multidimensional COSY, TOCSY and NOESY NMR experiments, to gain a better understanding of their structure. The peptide variants yielded complex spectra containing several overlapping peaks. Backbone walks of each peptide sample revealed sets of peaks which corresponded to an isomer caused by *cis-trans* isomerization of proline. Integration of peaks suggests that these alternate forms were approximately 10% of the total population in the Las var 9 NMR sample, and approximately 50% of the Las var 11 sample. Due to the complexity of the NMR spectra, we were unable to obtain a high-resolution structure of the two variants. As such, we turned to qualitative NOE data to inform us of the relative fold of the peptides.

We first examined if the two variants contained a degree of helical structure. Chemical shift deviation values of Las var 9 and 10 do not show consecutive negative values for the change in chemical shift values compared to a random coil, which may suggest a lack of ordered helical structure (Figure 29). However, due to the density of modifications in the two peptides, complementary evidence is needed to obtain a clearer picture of the peptide secondary structure and fold.

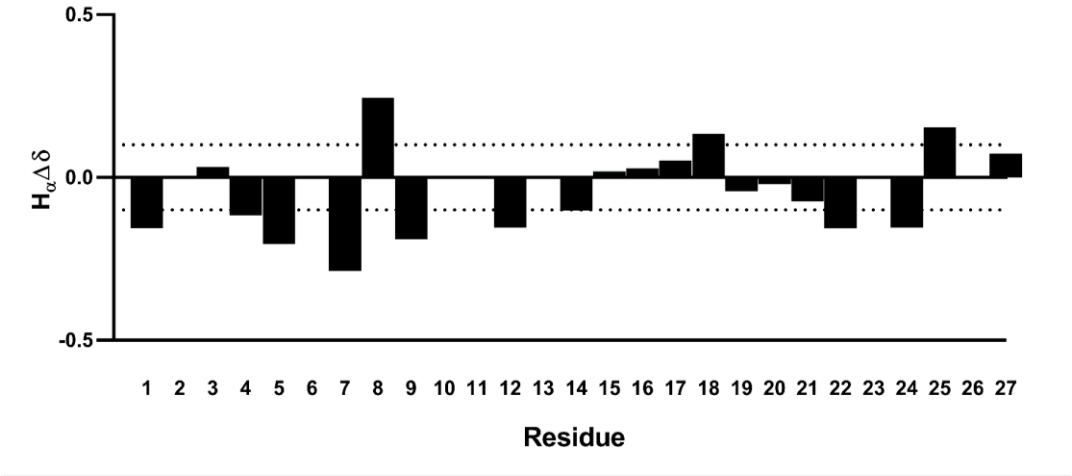


Figure 29. Hα chemical deviation plot of Las var 9.

Signal correlations pertaining to the backbone residues (backbone walk) of the Las var 9 NMR spectra revealed an unbroken connectivity of residues from Gly^{1}-Cys^{27} and a native disulfide bond connectivity of Cys^{8}-Cys^{25} and Cys^{17}-Cys^{27}. Examining helical contacts reveal primarily ambiguous $d_{H\alpha\text{-}HN}(i\text{-}i+3)$ NOE signals within Arg^{4}-Lys^{14} and Leu^{21}-Cys^{24} suggestive of two regions with putative helical characteristics, corresponding with the two helical regions observed in the native peptide. Inspection of H_α-H_N coupling constants suggest that Las var 9 only has partial helical characteristic and does not adopt the same fold as Lasiocepsin. This is corroborated by the absence of long-range NOE signals from side chains of residues observed to be proximal in the PDB structure of the native peptide.

Examination of the NMR data for Las var 11 yielded similar results. Chemical shift deviation of Las var 11 shows several residues with negative values, which may hint at partial helical characteristic, however negative Hα chemical values for consecutive residues cannot be observed due to the peptide's high density of unnatural amino acid residues (Figure 22).

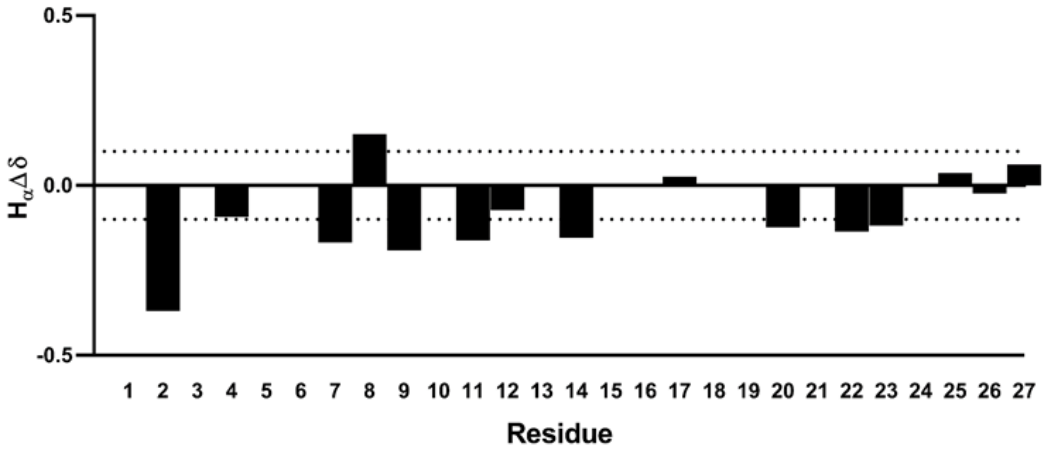


Figure 30. Hα chemical deviation plot of Las var 11.

Like Las var 9 $d_{H\alpha\text{-}HN}(i\text{-}i+3)$ NOEs reveal ambiguous signals within Cys^{8}-Cys^{25} and Cys^{17}-Cys^{27} which point to two possible helical domains. A closer examination of the 1D NMR reveals a spectrum with largely convoluted peaks, making it difficult to obtain Hα-HN coupling constant values for a particular residue. Long range NOE signals between Hβ of $Ala^{11}$to Cys^{25} and Cys^{8}-Ala^{11} were observed hint at these residues are proximal into each other in Las var 11 as is in the native peptide structure. However, the presence of a second set of peaks corresponding with the alternate *cis*-Pro conformation crowds the 2D spectra, precluding the complete assignment of sidechains and an unequivocal assessment of tertiary contacts germane to the overall fold of Las var 11. The presence of secondary conformations in Las var 11 and Las var 9 is postulated to be caused by a *cis-trans* isomerization of Pro^{20} which may greatly affect the formation of the helical region close to the C-terminus of these peptides. This is corroborated by the preponderance of alternate conformations of approximately half the residues for Las var 11 spanning Cys^{17} to Cys^{27}.

3.2.11 Proteolysis of Second-Generation Las Variants

Backbone substitution has been found to impart increased protection against various proteolytic enzymes, and can be augmented by the position and type of amino acids used in the sequence alteration.[8] As mentioned in Chapter 2, disulfides can also impart proteolytic resistance, making some disulfide-rich peptides desirable scaffolds for chemical modification. To qualitatively assess if the substitutions in the second generation of Las variants have an effect on proteolytic susceptibility, the native and bioactive variants they were subjected to proteolytic assays with trypsin. Lasiocepsin contains several lysine residues as well as an arginine residue at position 3 which may be labile trypsin cleavage. Las NH_2, var 9 and var 11 were subjected to tryptic digests at room temperature for an hour. With each 15-minute time point, an aliquot of the reaction solution was quenched via acidification with 2%TFA and monitored by analytical HPLC.

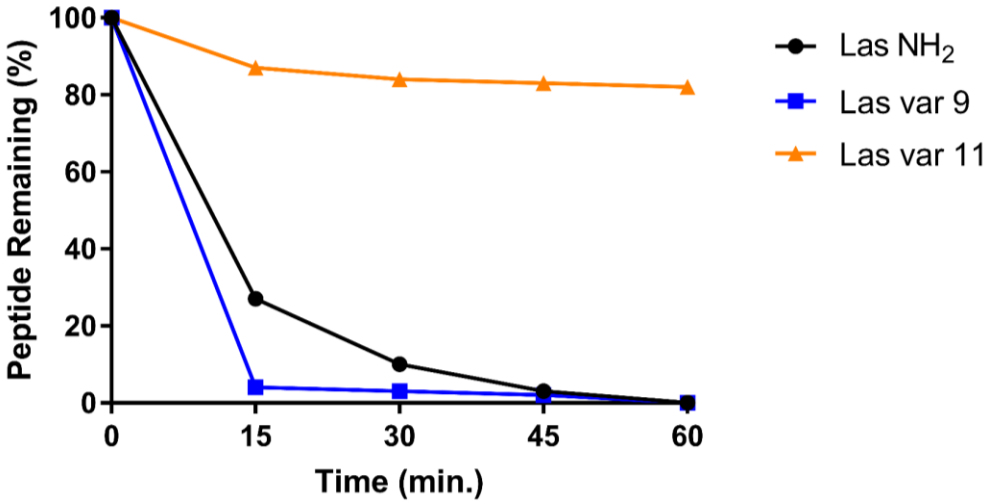


Figure 31. Time courses for proteolytic degradation of Las NH_2, Las var 9, and Las var 11 in the presence of trypsin; each reaction contained 250 µM peptide with 120 nM enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5.

Analytical traces of Las NH_2 showed that less than 50% of the peptide was remaining within 15 minutes of the assay and that the entire starting material was degraded by 60 minutes. This observation suggests that not every disulfide-rich peptide scaffold is inherently resistant to proteolysis. Interestingly, Las var 9 was qualitatively more susceptible to proteolysis compared to the native sequence, with less than 5% of the starting material remaining after 15 minutes of digestion. This observation underscores the importance of judicious placement of modifications when designing backbone-modified peptides. Recall that the first five residues as well as the loop region of Las var 9 were unmodified. These areas contain most of the basic residues within the sequence that are susceptible to tryptic cleavage. Given the unfolded structure of Las var 9, it is likely that these unmodified regions are spatially projected such that they are more accessible to cleavage by trypsin compared to the native. In contrast, Las var 11 which contained modifications in these two regions was observed to have proteolytic resistance against trypsin as evidenced by a high percentage of starting material remaining after an hour of incubation. The disparate properties of these samples underscore the challenge in designing peptide variants that are capable of structural and biological mimicry while augmenting other pharmacokinetic properties such as proteolytic resistance.

3.2.12 Third Generation Lasiocepsin Variant

Out of the synthesized second-generation Lasiocepsin variants, Las var 11 was more bioactive than Las var 9 and was also found to be more resistant to tryptic cleavage. As discussed in previous sections, the NMR of this variant contained a second set of chemical shifts which was postulated to be an alternate conformation caused by *cis-trans* isomerization about proline 20. We hypothesized that replacing this proline with another amino acid to prevent this structural

isomerism improve the bioactivity observed in Las var 11 but would and the alternate conformation observed via NMR. Moreover, to improve order in the N-terminal helix, we decided to mutate the β^3 lysine at position 5 back to its native residue. These modifications became the basis of our design for a third generation las variant, Las var 12 (Figure 32).

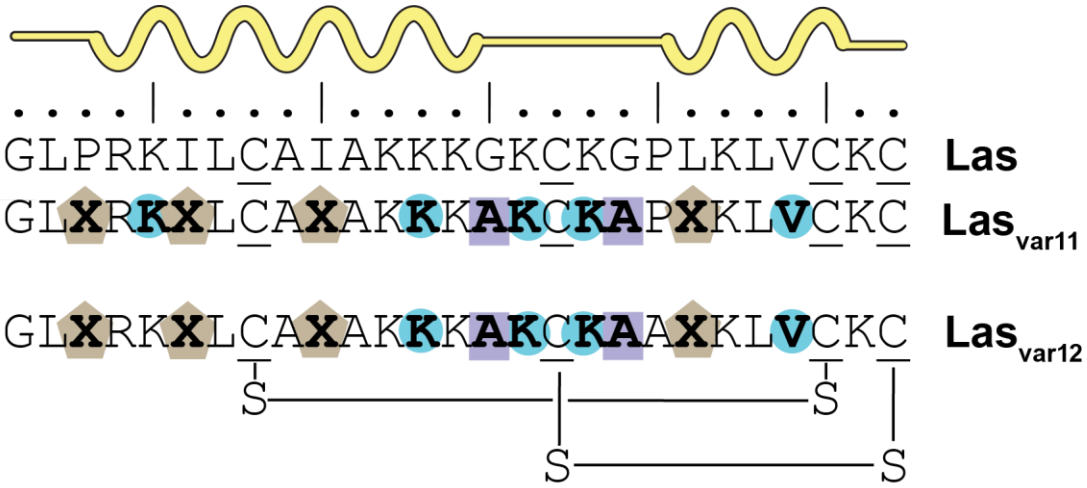


Figure 32. Sequence of Las var 12 as compared to Las var 11 and native Lasiocepsin. Las var 12 represents the third generation of Lasiocepsin variants and is modified from Las var 11.

3.2.12.1 Las Var 12 Proteolytic Degradation

Las var 12 was synthesized, purified, and subjected to proteolysis by trypsin. Given the stability of Las var 11 to tryptic degradation, Las var 12 was subjected to a trypsin digest and sampled at several time points from 15 minutes to 8 hours. In these runs, it was found Las var 12 had a half-life of approximately 158 minutes (Figure 33), which was as significant improvement from its native counterpart which completely degraded by 60 minutes under the same conditions (Figure 31). The proteolytic stability of Las var 12 was also compared with Las var 11 over an 8-hour time span. It was observed that both peptides were not completely degraded after this time span. It was observed that Las var 11 was slightly more resistant to proteolysis than Las var 12,

with an estimated half-life of approximately 292 minutes. It should be noted that this measured half-life may be overestimated as integration of the peak representing the Las var 11 starting material is complicated by overlapping peaks with digestion products.

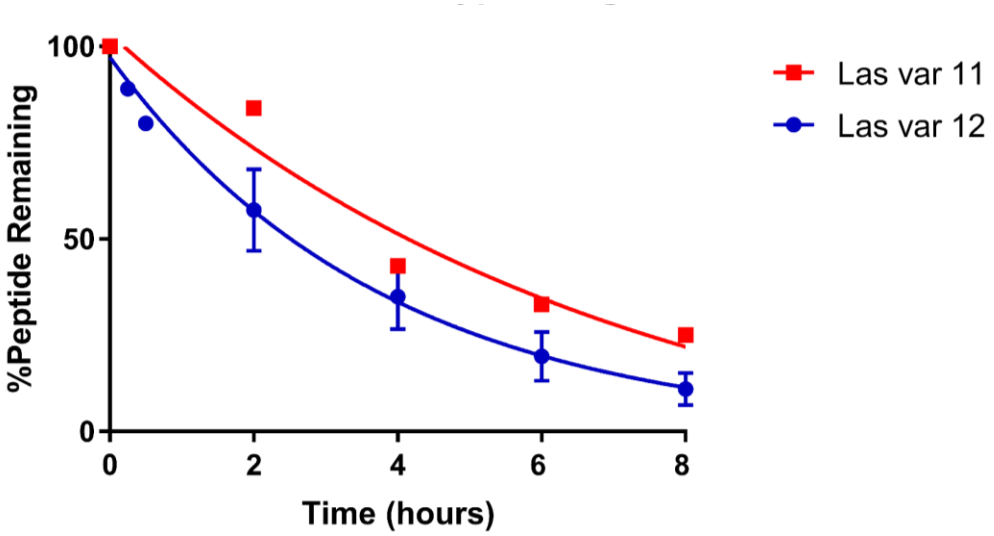


Figure 33. Time courses for proteolytic degradation of Las var 11 and Las var 12 in the presence of trypsin; each reaction contained 250 µM peptide with 120 nM enzyme in 0.1 M ammonium bicarbonate buffer, pH 8.5.

3.2.12.2 Microbiological Assay of Las Var 12

The antimicrobial properties of Las var 12 against various ESKAPE pathogens was directly evaluated alongside Las NH_2. It was found that Las var 12 exerted bactericidal activity against most of the ESKAPE pathogens in the panel apart from methicillin susceptible *S. aureus* (Table 4).

Table 4. Direct comparison of Las var 12 and Las NH_2 against clinically-relevant ESKAPE pathogens

	MIC (µM)					
	Gram Positive		Gram Negative			
	E. faecium	*S. aureus (MSSA)*	*K. pneumoniae*	*A. baumannii*	*P. aeruginosa*	*E. coli*
Las NH_2	1	4	4	1	8	1
Las var 12	1	>32	8	1	8	2

Moreover, the MIC values from this direct comparison suggests that Las var 12 exhibits similar potency as the native peptide towards the susceptible ESKAPE pathogens tested. Interestingly, Las var 12 shows increased activity towards *K. pneumoniae* and *P. aeruginosa* when compared to Las var 11. This suggests that the mutations made in the design of Las var 12, specifically the replacement of Pro^{20} with Alanine and reversion of β^3 Lysine to Lysine at position 5, effectively increases bioactivity towards a wider range of Gram negative ESKAPE pathogens. It is notable that Las var 12 inactive against methicillin susceptible *S. Aureus* (MSSA). It is precedented that Gram positive strains develop resistance to antimicrobial peptides through modification of the bacterial cell wall components, resulting the neutralization of negative charge needed for initial interaction of a the cationic AMPs.[187] It is probable that the MSSA strains tested in these assays have adapted a similar mode of resistance, making it difficult for Las variants to penetrate the bacterial cell wall, especially if these analogs are improperly folded.

3.2.12.3 Hemolysis Assay of Las Var 12

The hemolytic properties of Las var 12 was compared to the native sequence and Las variants from previous generations. It was found Las var 12 was considerably less hemolytic than Las NH_2 at all AMP concentrations tested (Figure 34). These observations indicate that the modifications made on Las var 12 did not affect hemolysis, as it showed similar activity to the weakly-hemolytic first and second-generation Las analogs (Figure 34).

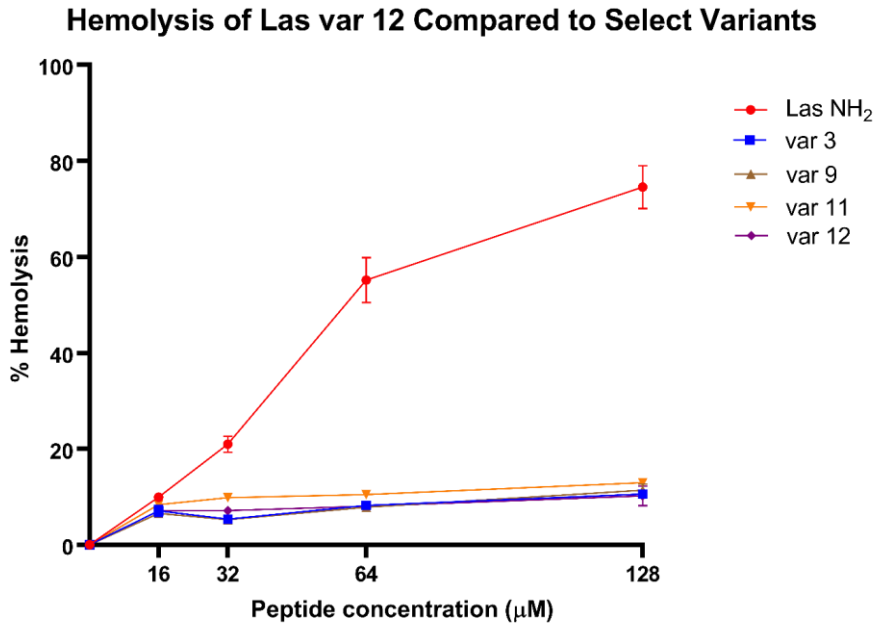


Figure 34. Direct comparison of Las var 12 hemolysis against Las NH_2 and select Lasiocepsin analogs.

3.2.12.4 Cytotoxicity of Las Var 12

The cytotoxicity of Las var 12 against RAW264.7 cells was compared to select first and second-generation Lasiocepsin variants and evaluated after 4- and 24-hour incubations. It was observed that was significantly less cytotoxic compared to Las NH_2 in all concentrations tested at the 4-hour and 24-hour time points (Figure 35). Curiously, after 24 hours of incubation, increased cytotoxicity could be observed in cell samples when incubated at 64 μM of Las var 12. This trend could also be observed with Las var 11 in previous cytotoxicity assays (Figure 28) suggesting that these two peptides of similar sequence may share a common process by which they elicit cytotoxicity. Conversely incubations with lower concentrations (8 and 16 μM) of the peptide did not exhibit cytotoxicity, corroborating previous observations that cytotoxicity may be elicited only at higher concentrations.

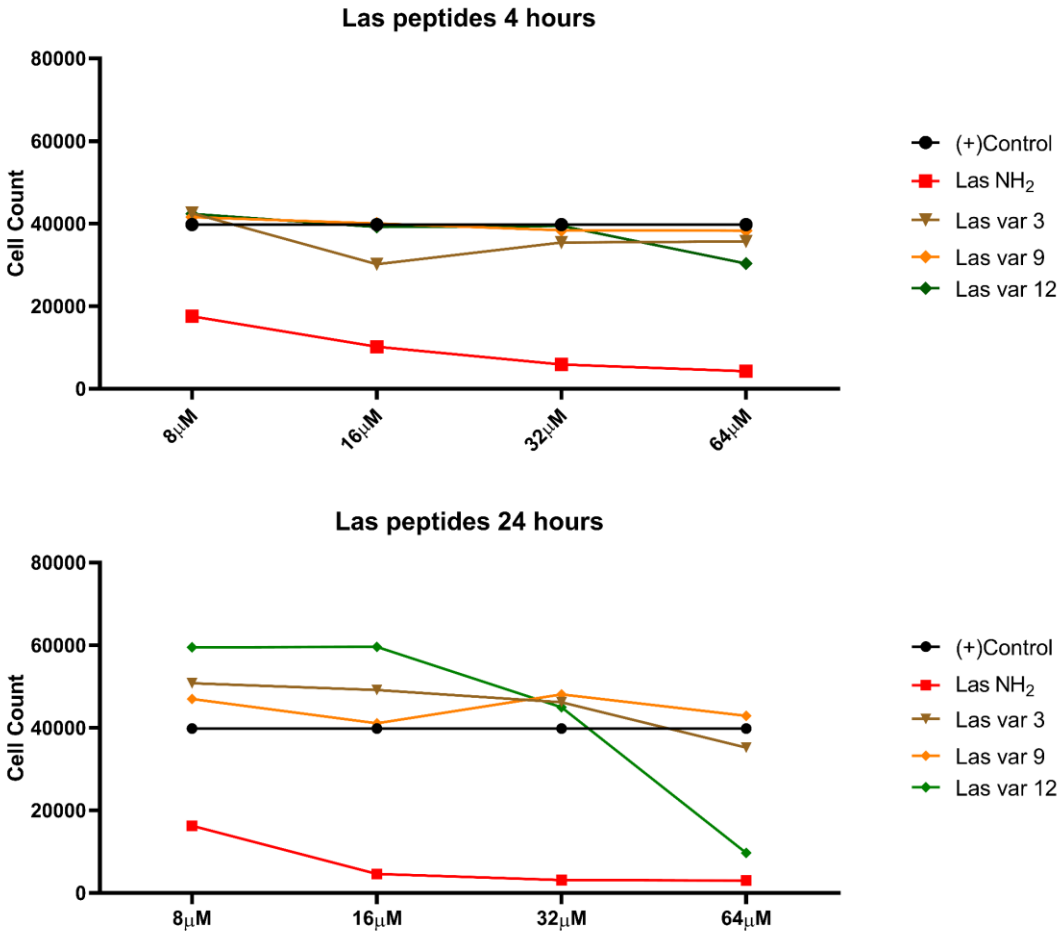


Figure 35. Cytotoxicity of Las var 12 as compared with select Lasiocepsin variants.

3.2.12.5 NMR of Las Var 12

Las var 12 was subjected to multidimensional NMR consisting of NOESY, COSY and TOCSY experiments. Qualitative comparison of the spectra of Las var 12 against that of Las var 11 shows that secondary chemical shifts associated with each residue are not present, indicating that Las var 12 does not assume an alternate conformation in solution. Efforts are ongoing to determine the structure of this variant.

3.3 Conclusions and Future Directions

Lasiocepsin is a 27 amino acid disulfide-rich peptide isolated from bee venom. This peptide exhibits notable activity against several clinically relevant bacteria comprising the ESKAPE pathogens which are responsible for increased antibiotic resistance and nosocomial infections. Through syntheses of three generations of Lasiocepsin analogs containing varying degrees of backbone substitutions, we demonstrated that the activity of this peptide could be mimicked, and that recapitulation of peptide bioactivity can be progressively tuned with strategic placement of unnatural amino acids.

Las var 12, the latest iteration of these heterogeneous backbone analogs exhibits antimicrobial properties comparable to that of the native peptide. Moreover, this variant displays proteolytic resistance above and beyond that of the native peptide. However, proteolytic resistance in not inherent with every backbone modified disulfide-rich peptide. Las var 9 in the second series of analogs, exhibited similar susceptibility to trypsin as the native sequence, completely degrading after an hour of incubation with trypsin. Progressive modification of the peptide backbones also yielded analogs that are less cytotoxic and hemolytic than the native peptide, pharmacological properties important in designing potential antimicrobials that have selectivity for target pathogens but are non-harmful to hosts.

Although natural peptides have structures that influence their function, the mimicry of function does not necessarily mean the recapitulation of structure. Investigation of bioactive peptides Las var 9 and 11 showed that these samples were qualitatively not folded and did not have the structure of the native peptide. This hints at the possibility that the variants are exhibiting a different mechanism of action than the parent sequence. Taken together, the observations in this study highlight the modularity of peptides and the importance of judicious placement of backbone

modifications to generate peptide analogs with both desired structural and functional characteristics.

Despite the success of Las var 12, 11 and 9 several open questions remain: What is a possible application of these sets of peptides? Can we tune the specificity of Las variants? and Do these analogs have similar mechanisms of action as native Las NH_2? These issues are important in guiding the design of future Las variants. Regarding the possible application of Lasiocepsin, one can envisage using these proteomimetics as a complementary therapy to broad-spectrum antibiotics during empirical treatment of patients suspected of being infected with bacteria from the ESKAPE panel. As mentioned, ESKAPE pathogens are involved in hospital-acquired infections and are increasingly becoming resistant to various drugs. Coupling Las variants with these antibiotics may decrease chance of mounting antibiotic resistance towards these treatments since AMPs have a different mechanism of action than small molecule antibiotics.

Although development of Las as an antibiotic is non-trivial, experiments can be conducted to determine the feasibility of delivering these peptides into *in vivo* systems. In this study, we have shown that Las variants can have longer half-lives than the native sequence in tryptic digests, which shows some promise with regards to oral administration of these peptides. Future investigations could involve analytical studies within the context of Las resistance to degradation in various biological fluids including gastric fluid, intestinal fluid, serum, and plasma. If found that Las and its variants are quickly degraded by these media, it may be reasonable to envision these peptides being incorporated as a topical antibiotic since these disulfide-rich peptides do not readily decompose when left at room temperatures. Lastly, even if Las and its variants are not utilized for therapeutic purposes, these proteomimetics still may be utilized as a guide for designing analogs of disulfide-rich peptides with antimicrobial properties.

Another open question involves tuning the specificity of Las variants towards certain combinations of bacteria in the ESKAPE pathogen panel. A trend observed in these ESKAPE assays is that bacteria that are most susceptible to the native peptide are also the first bacteria that become vulnerable to the Las variants. Two simple substitutions in Las var 12 increased its activity against *K. pneumoniae* and *P. aeruginosa* as compared to Las var 11. This indicates the possibility of tuning Las specificity towards certain bacterial strains through backbone substitutions. Future work may include synthesizing a series of peptides in which only a single motif of Las is modified at a time and determining its correlation with bioactivity. Once certain modifications can be associated with an effect on MIC, work can begin in optimizing specificity toward certain bacterial strains or in combining sequence substitutions to increase broad spectrum activity towards the entire ESKAPE panel. These structure-activity relationship studies would best be guided by high-resolution structures of these peptides, with the caveat that they may be challenging to obtain.

Precedent studies have tried to elucidate the structural origin of Lasiocepsin antimicrobial activity.[180] Despite being capable of inhibiting bacterial growth, it is an open question whether bioactive Las variants have a similar mechanism of action as the native sequence especially since the structures of Las var 9, 10 and 11 are currently unknown. It would thus be important to replicate the experiments outlined by Monincová and colleagues[180] including tryptophan fluorescence to determine the membrane composition preference as well as membrane permeabilization assays. A recent publication suggested that the Las NH_2 and other peptide toxins interact with the peptide binding site of *E.coli* ATPase and other possible cellular structures.[188] Determining if bioactive Las variants also interact with these targets would aid in elucidating their mechanism of action.

3.4 Methods

3.4.1 Solid Phase Synthesis and Cleavage

Linear Las NH_2, Las var 1 and Las var 2 were synthesized by microwave assisted solid-phase peptide synthesis on NovaPEG Rink Amide Resin (Novabiochem). Resin was washed with DMF and left overnight to facilitate swelling prior to synthesis. Each amino acid was activated *in situ* through mixing of N-α-Fmoc-protected amino acids (4 equivalents) and HCTU (4 equivalents) and DIEA (6 equivalents) in NMP with a final concentration of 0.1M amino acid unless otherwise specified. Coupling reactions for Fmoc-Cys(Trt)-OH and Fmoc-Cys(Acm)-OH residues were carried out at room temperature for 30 minutes to prevent racemization. Coupling reactions of other amino acids were performed on a CEM MARS microwave reactor at 70°C for 4 minutes. Microwave-assisted deprotection of each amino acid was performed using 20% v/v 4-methylpipperidine in DMF for 2 minutes at 80°C. Post-synthesis, the resin was sequentially washed with DMF, DCM and methanol prior to being dried under vacuum. Las var 3, 9, 10 and 11 as well as a second batch of Las NH_2 achieved via automated synthesis using a Biotage Alstra Automated Synthesizer using the same conditions as above. In this series, amino acids were double coupled to ACPC residues to decrease chances of peptide truncation. Cleavage of linear peptides from resin was performed using a cleavage cocktail comprised of TFA/TIS/thioanisole/EDT/H2O (90:2:3:2.5:2.5 by volume) and allowed to agitate at room temperature for 3.5 hours before filtration. Cleaved peptides were precipitated in cold diethyl ether and centrifuged at 6000 RPM for 2 minutes to pellet the precipitate. Crude linear peptides were purified by preparatory HPLC using a Phenomenex Jupiter C18 Preparatory Column (particle size 10 μm, 300 Å pore size) utilizing gradients of water and acetonitrile with 1% TFA. The identity of each peptide was

confirmed by analytical HPLC using a Phenomenex Jupiter Analytical C18 column (particle size 5 μm, 300 Å pore size) and by MALDI-TOF MS. Purified peptides were lyophilized prior to post-synthetic manipulations.

3.4.2 Regioselective Disulfide-Bond Formation

Lyophilized peptides were dissolved in a solution of 0.1M ammonium acetate (pH 7.8) in a concentration of 0.2 mg/mL. The solution was stirred at room temperature for 48 hours. Formation of 1 D.S. bond intermediate was confirmed by MALDI-TOF MS. The oxidation solution subsequently was acidified with 50% acetic acid in water, to a pH of 4 and lyophilized._

Lyophilized peptides bearing a single disulfide bond were dissolved in an aqueous solution of 50% methanol containing 20mM HCl at a concentration of 0.2 mg/mL. 0.1M iodine in acetic acid (5 equivalents/Acm) was added dropwise to the solution. The reaction was then allowed to stir at room temperature for 1.5 hours. Formation of the second disulfide bond was confirmed by MALDI-TOF MS prior to quenching by dropwise addition of 0.2M ascorbic acid. Excess methanol was then rotary evaporated from the solution prior to lyophilization. All bisdisulfide oxidized peptides were purified by preparatory HPLC and the purity checked by analytical HPLC prior to biological assays.

3.4.3 NMR of Las Variants

Purified oxidized peptide samples were dissolved in 500 μL of NMR solution comprised of 1:9 D_20:H_2O and 0.2mM DSS at pH 7 (uncorrected). The spectra of each sample were collected on a Bruker Avance 700 MHz NMR spectrometer at 300K. Two dimensional experiments

consisted of NOESY (200 ms mixing time), TOCSY (55 ms mixing time), and magnitude COSY with water suppression. Collected 1D and 2D spectra were processed in TOPSPIN with DSS as an internal standard and chemical shift assignments performed on SPARKY NMR Assignment Integration SOFTWARE (https://nmrfam.wisc.edu/nmrfam-sparky-distribution). Chemical shift deviation and index plots were generated through comparison of assigned Hα chemical shifts, compared to random chemical shift values obtained from the Poulsen IDP/IUP Random Coil Chemical Shift Server.

3.4.4 Proteolytic Degradation

Lyophilized peptides were weighed and dissolved in deionized water to a concentration of 1.716 mM. A stock solution of bovine trypsin (MW 23,290 Da) was prepared by dissolving 2.8 mg of enzyme in digest solution comprised of 0.1M NH_4HCO_3 (pH 8.5) containing 1mM $CaCl_2$. 1.2 μM working solutions of trypsin were generated by serial dilution. Samples representing the starting material consisted of 250 μM peptide in digest solution with 50 μL of 2% TFA. Each digestion reaction consisted of 250 μM peptide and 120 nM enzyme in digest solution. Reactions were incubated at room temperature. At desired time points, 50 μL of reaction mixture was aliquoted and quenched with 50 μL of 2% TFA solution. Each sample was by analytical HPLC (95 uL injection) using a Phenomenex Jupiter Analytical C18 column (particle size 5 μm, 300 Å pore size) with a linear 20-35% Solvent B gradient.

3.4.5 ESKAPE Pathogen Assay

These experiments were performed by our collaborator Qiao Lin from the lab of Dr. Peter Di at the University of Pittsburgh School of Public Health.

The panel of ESKAPE pathogens used in these experiments were obtained from isolated cultures of pediatric patients with chronic pulmonary infections from Seattle Children's Hospital or from adult patients at the University of Pittsburgh Medical Center. The minimum inhibitory concentration was assessed through standard procedures outlined by the Clinical Laboratory Standards Institute (CLSI) and from precedent studies by Deslouches and colleagues[172]. Assay modifications included using a cation-adjusted Muller Hinton broth 2 as described in procedures by Weigand and coworkers[189]. Assays were performed on a 96-well plate with an initial bacterial concentration for each treatment group was 10^6 CFU/mL. Samples were incubated in a microplate reader for 18 hours at 37°C. Post incubation, the optical density was measured at 570 nm and was used to evaluate the minimum inhibitory concentration.

3.4.6 Hemolytic Assay of Mouse RBC

These experiments were performed by our collaborator Dr. Qiao Lin from the lab of Dr. Peter Di at the University of Pittsburgh School of Public Health.

Assays were performed using mouse whole blood, stored in lithium heparin-coated tubes. The whole mouse blood was centrifuged for 30 minutes at 1500 x g to separate erythrocytes from other blood cells. The serum and buffy coat were subsequently removed and the remaining erythrocytes diluted in PBS to a 2.5% concentration of erythrocytes. 50 μL of this 2.5% RBC sample was aliquoted into a 96-well plate. A 100% RBC lysate standard was generated by dilution

of Triton X-100 to 1%. 50 μL of Las AMPs at various concentrations up to 128 μM were added to each 1% erythrocyte sample and incubated for 1 hour at 37°C followed by centrifugation of the whole plate for 5 minutes. The supernatant was carefully transferred to another 96-well plate without disturbing the pelleted erythrocytes. Results were assessed by observing the optical density at 570 nM.

3.4.7 Cytotoxicity Assay Against RAW264.7 Cells

These experiments were performed by our collaborator Dr. Qiao Lin from the lab of Dr. Peter Di at the University of Pittsburgh School of Public Health.

RAW264.7 cells were aliquoted into a 96-well plate at a volume of 100μL/well and a cell count of $5x10^4$ cells/well. Cells were incubated for 24 hours to allow adhesion to the wells. Las peptides were diluted in cell culture medium comprising of DMEM D5796, 10%FBS and 1% penicillin-streptomycin). After incubation 50 μL of the supernatant was removed prior to adding 50 μL of Las AMP solution (total volume of 100 μL per well). Treated samples were incubated for 4 or 24 hours prior to addition of 10% volume TetraZ Cell Counting Kit reagent. Plates were then incubated at 37°C for 1.5 hours to allow color development. Cytotoxicity was determined by measuring the optical density at 450 nm.

4.0 Experimental Validation of a Force Field for Simulation of Artificial Protein-Like Backbones

A portion of this chapter has been reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101[190] with permission from AIP Publishing. https://doi.org/10.1063/5.0019054

4.1 Introduction

The field of structural biology is fast expanding to encompass a panoply of protein and peptide entities containing residues beyond the canonical 20 amino acids. Chemical syntheses of these macromolecules can involve a large density of various backbone modifications. These may be derived from parent peptides of natural origin such as Lasiocepsin (Chapter 3), or *de novo* designed like NC_HEE_D1 (Chapter 2) and are required to experimentally study how such changes affect folding behavior. In contrast to the plethora of existing synthetic modifications for peptide backbones, it has been observed that cells *in vivo* readily alter the peptide backbone through post-translational modifications that yield natural proteins with complicated heterogeneous backbone alterations that are difficult to replicate via chemical synthesis. This class of macromolecules with unique structures resulting from extensive alteration of peptide backbones are referred to as Ribosomally-synthesized and Post-Translationally modified peptide natural

products (RiPPs).[191] The study of the structure of both synthetic and endogenous proteins are required to advance the current understanding of peptide folding behavior and stability that result from synergistic effects of backbone modifications.

4.1.1 Molecular Dynamics Simulations and Peptide Folding

Computational modeling provides a method of understanding peptide folding behavior that is complementary to experimental biophysical techniques such as NMR, X-ray crystallography and Cryo-electron microscopy. Accurate modeling of protein folding necessitates accounting for the disposition of each atom in the system and how it evolves over time. Atomic behavior is best described through quantum mechanics but applying these sets of rules to each component of a large system is computationally expensive and time consuming. Molecular dynamics simulations circumvent this limitation through approximation of quantum mechanical behavior using classical mechanics.

In its simplest form, this technique describes molecular forces acting on each component through force fields, a set of parameters derived from Newton's equations of force and motion, which can delineate the position and velocity of each atom with respect to time.[192] The total potential energy of the system can be viewed as the sum of all the potential energies of the atoms involved. Since atoms of a system can interact through bonds and through space, classical mechanics equations involved in MD simulations are generally organized into bonded and nonbonded terms. Bonded terms account for stretching of atomic bonds and bond angles which are described by equations for simple harmonic motion. Additionally, rotation through dihedrals are expressed as a sinusoidal function correlating the energy difference between eclipsed and staggered conformations. Non-bonded behavior simulates Van der Waals and electrostatic forces

and are described through terms that combining the Lennard-Jones potential and Coulomb's law that account for the attractive and repulsive forces acting on charged and neutral bodies with respect to distance.[192-193] Several force field packages have been developed for MD simulations of peptides and proteins including Amber, CHARMM, GROMOS and OPLS-AA. Although each of these force field suites simulate similar physical phenomena, they can differ in the treatment of their bonded and nonbonded terms, including increased emphasis on parameterization of improper dihedral angles for accurate modeling of chirality, as well as differential application of the Lennard-Jones potential for dissimilar atom types.[194]

Early work on MD simulations demonstrated the capabilities and adaptability of force-fields in recapitulating the structure of short dipeptides as well as large proteins ubiquitin and other macromolecules.[195-196] Further refinement and application of these force fields in MD simulations of proteins, have allowed researchers to monitor the evolution of polypeptide structure in short time frames, providing insights into folding kinetics, thermodynamics and topological propensities.[197] A notable study investigated octamer peptide fragment sequences taken from parts of larger proteins and showed that certain peptide sequences adopt defined motifs whether found as part of a larger protein, or as explicit polypeptides.[198] Moreover, MD simulations have provided insights on common phenomena such as salt bridge formation, and amphiphilicity which are involved in both peptide folding and biomolecular interactions.[199-201] Despite the insights gained from MD simulations of proteins, very few studies have focused on peptides with heterogeneous backbones. A notable early study on a backbone-modified peptide investigated the stability of a heptapeptide comprised for regular β^3-amino acids flanking an α-methylated β^3Alanine which was found to assume a 3_{10}-helical conformation even after thermal denaturation.[202] A recent study has shown that MD simulations are unable to accurately model *cis/trans* isomers of cyclic N-

methylated hexapeptides from unless the correct isomer is first delineated.[203] Moreover, MD simulations of cyclic α/β peptides have shown that popular eight popular force fields from the AMBER, CHARMM, GROMOS and OPLCS suites could not reproduce experimental NMR observables.[204] Collectively, these investigations point to a need in parameterization of force fields to account for unique conformations resulting from the backbone modification of peptides.

4.1.2 Design of Tripeptides for Force Field Validation

Simulating the folding behavior of peptides and proteins with heterogeneous backbones via molecular dynamics is currently difficult due to the lack of rigorous force fields that are parameterized specifically for unnatural amino acids. In a collaborative effort with the lab of Dr. Lillian Chong in the Department of Chemistry, we completed development and validation of a modified version of the AMBER ff15ipq force field capable of treating such entities. Initial efforts towards this goal required application and of the Amber ff15ipq force fields to a series of short peptide sequences containing specific backbone modification types. Precedent studies have used short polyalanine peptides to conformation-dependent experimental observables and compare them to φ and ψ angle values obtained from MD simulations, which have been used in force field optimization.[205] In pursuit of this goal, a series of tripeptides containing two Ala residues flanking a central residue comprised of either an Ala, D-Ala, β^3Ala, or ACPC residue was synthesized (Figure 36). Each peptide was synthesized with a C-terminal amide, and acetyl capped at the N-terminus. This peptide series was designed to be examined by 1D ^{1}HNMR to determine backbone H_N-Hα (or H_N-H_β) J-coupling values. These values are diagnostic of the distributions of relevant torsional angles about the corresponding N-C bond and can be readily extracted from MD simulation ensembles.

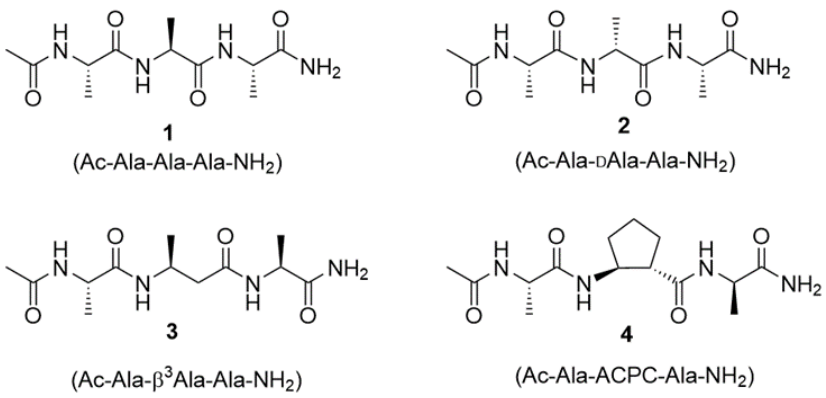


Figure 36. Sequence and structure of synthesized tripeptides for force field validation. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

4.2 Results and Discussion

Peptides **1-4** were synthesized and purified, and the identity of each peptide confirmed by ESI-MS (Table 5).

Table 5. ESI-MS data for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

Peptide	$[M+H]^+$ Calc.	Observed *m/z*
1	273.16	273.16
2	273.16	273.16
3	287.17	287.17
4	313.19	313.19

Purified peptides were subjected to NMR analysis. Precedent studies have suggested that polyalanine peptides can aggregate in solution[206], making accurate J-coupling values difficult to obtain. Thus, initial control experiments were performed to establish that peptide **1** did not aggregate in solution, validating the use of the peptide series for obtaining accurate J-coupling values. Solutions of **1** at various concentrations (1, 2, 5, and 10 mM) were subjected to 1D NMR and inspected for peak shifts and broadening. It was found that chemical shifts and peak shapes were consistent and well-resolved. The concentration-independent spectra suggest the peptide is behaving as an isolated species in solution and, thus, data from it can be correlated to a simulation ensemble involving a single molecule.

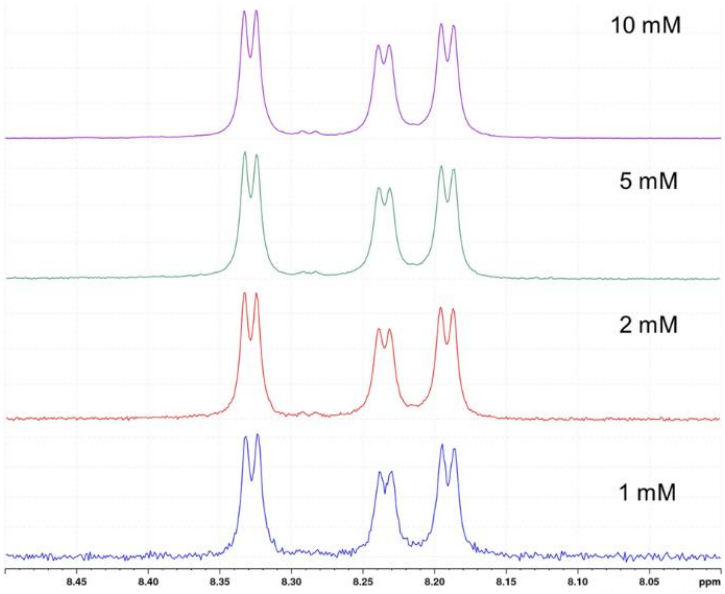


Figure 37. 1D NMR of amide NH region for various concentrations of peptide 1. A consistent lineshape and chemical shift was observed, indicating that the samples were not aggregating. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field

for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

Peptides **1**-**4** were then subjected to a series of 2D homonuclear experiments to enable complete sequence-specific resonance assignment, including the amide H_N signals. The peaks correlating to the identity of the HN signals of each respective residue are delineated in the 1D spectra are shown in (Figure 38). ^3JH_N-Hα coupling constants ($^3J_{HN\text{-}H\beta}$ for β^3Ala and ACPC) from the 1D spectra were measured and tabulated (Table 6). The observed values for α-Ala peptide at positions 1 and 3 closely match precedent experimental coupling constants observed for other polyalanine peptides.[205] This suggests that these tripeptides can serve as good experimental systems to derive accurate ψ and φ angles for comparison with the parameterized AMBER ff15ipq.

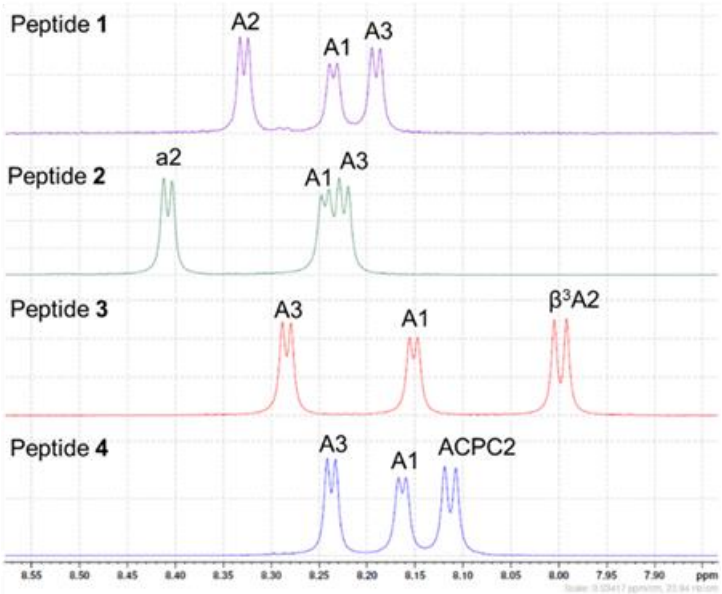


Figure 38. Amide region of the ^{1}HNMR of peptides 1-4 showing peak assignments for H_N resonances. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The

AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

Table 6. $^3J_{HN\text{-}H\alpha}$ coupling constants for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

Peptide **1**		Peptide **2**		Peptide **3**		Peptide **4**	
Peak	J[Hz]	Peak	J[Hz]	Peak	J[Hz]	Peak	J[Hz]
A1	5.9 ± 0.4	A1	5.9 ± 0.4	A1	6.3 ± 0.4	A1	6.1 ± 0.4
A2	6.2 ± 0.4	a2	6.4 ± 0.4	B3A2	9.2 ± 0.4	ACPC2	8.4 ± 0.4
A3	6.4 ± 0.4	A3	6.8 ± 0.4	A3	6.5 ± 0.4	A3	6.4 ± 0.4

4.3 Conclusions

The observed *J*-coupling constants for each residue of peptides **1**-**4** served as a point of comparison to assess the accuracy of the peptide backbone conformations derived from the AMBER ff15ipq-m force field parameterized by our collaborators. These NMR observables were compared to the computed φ torsion angles of simulated short peptides of similar sequence. $^3J_{HN\text{-}H\alpha}$ and $^3J_{HN\text{-}H\beta}$ values were derived for α residues and β residues respectively using various density functional theory (DFT) based Karplus coefficients.

It was observed that each simulation yielded *J*-coupling constants that were within error of the experimental observables (Table 7) with the DFT-3 coefficients being the most accurate as it

was specifically derived by our collaborators to treat the backbone-modified amino acids β^3-Ala and ACPC residues.

Table 7. Comparison of experimental and calculated J-coupling constants for peptides 1-4. Reproduced from Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. The Journal of Chemical Physics 2020, 153 (6), 064101. with permission from AIP Publishing.

		Simulation[a]			
Peptide	Residue	DFT-1	DFT-2	DFT-3	Experiment[b]
1	Ala^1	6.1 ± 0.1	6.7 ± 0.1	6.4 ± 0.1	5.9 ± 0.4
	Ala^2	6.2 ± 0.2	6.7 ± 0.1	6.5 ± 0.2	6.2 ± 0.4
	Ala^3	6.5 ± 0.1	7.0 ± 0.1	6.8 ± 0.1	6.4 ± 0.4
2	Ala^1	6.3 ± 0.3	6.8 ± 0.3	6.8 ± 0.4	5.9 ± 0.4
	D-Ala^2	6.1 ± 0.1	6.7 ± 0.1	6.4 ± 0.1	6.4 ± 0.4
	Ala^3	6.8 ± 0.2	7.4 ± 0.3	7.1 ± 0.2	6.8 ± 0.4
3	Ala^1	6.4 ± 0.2	6.4 ± 0.2	6.7 ± 0.3	6.3 ± 0.4
	β^3Ala^2	(7.5 ± 0.1)	(8.0 ± 0.1)	8.6 ± 0.1	9.2 ± 0.4
	Ala^3	6.5 ± 0.1	7.0 ± 0.1	6.8 ± 0.1	6.5 ± 0.4
4	Ala^1	6.4 ± 0.2	7.0 ± 0.2	6.7 ± 0.2	6.1 ± 0.4
	$ACPC^2$	(6.9 ± 0.1)	(7.5 ± 0.1)	8.5 ± 0.1	8.4 ± 0.4
	Ala^3	6.6 ± 0.1	7.2 ± 0.1	6.9 ± 0.1	6.4 ± 0.4

[a]Reported uncertainties are twice the standard error of the mean from ten simulations.
[b]Reported uncertainties are estimated as half the spectral resolution of the experiment.

Interestingly, the simulated *J*-coupling constant of β^3-Ala is outside of the error of measured values for the same residue in peptides **3**. This may be a result of a conformational ensemble in the simulations wherein resultant structures are in a more extended conformation as compared to the physical experiment. This is in accord with the NMR data obtained in the study of a computationally-designed disulfide-rich peptide scaffold discussed in chapter 2, where α-amino acids within less-extended structures such as the helical region contained smaller $^3J_{HN\text{-}H\alpha}$ compared to their counterparts observed in the extended β-hairpin.

4.4 Methods

4.4.1 Peptide Synthesis and Purification

All peptides were synthesized via microwave-assisted Fmoc solid-phase peptide synthesis (CEM MARS microwave reactor) on Sieber Amide Resin (0.05 mM scale). All amino acids were activated *in situ* by combining N-α-Fmoc-protected amino acid (4 equiv), HCTU (4 equiv), and DIEA (6 equiv) in N-Methyl-2-pyrrolidone (final concentration 0.1 M amino acid). Coupling of each amino acid was performed at 70°C for 4 minutes and deprotection of the Fmoc group was performed at 80°C for 2 minutes using 20% v/v 4-methylpiperidine in DMF. Each peptide was capped at the N-terminus with an acetyl group. The capping solution consisted of a mixture of DMF / DIEA / acetic anhydride (8:2:1 by volume). Capping solution (2 mL) was added to resin and allowed to stir at room temperature for 20 minutes. Post synthesis, the resin was successively washed with DMF, DCM, and methanol and allowed to dry in a vacuum desiccator for 30 minutes. Cleavage of the peptide from resin was performed by treatment of the resin with 1% TFA in DCM for 20-30 minutes with gentle stirring. The cleavage flowthrough was collected in a scintillation vial and excess DCM evaporated under nitrogen. For peptide **4**, the resin was washed an additional five times with cleavage solution post incubation to maximize cleavage yield. Cleaved peptides were purified by preparative HPLC on a Phenomenex Luna Prep C18 Column (particle size 10 μM, 100 Å pore size). The identity each peptide was confirmed by ESI-MS and purity determined by analytical HPLC. Purified peptides were lyophilized prior to NMR analysis.

4.4.2 NMR and Data Acquisition for Peptides 1-4

Lyophilized peptides were weighed and dissolved in 1:9 D2O:H2O solution (pH 2), with 0.2 mM DSS as a standard. NMR spectra of each sample were acquired on a Bruker Avance 700 MHz spectrometer at 300K, with 1D spectra comprising of 128 scans. Two dimensional experiments consisted of NOESY (200 ms mixing time, 16 scans) and COSY (8 scans). All spectra were processed using TOPSPIN, employing DSS as an internal standard. Concentrations of each sample were determined through integration a known peptide peak in the 1D spectrum and referencing it to the integration of the DSS standard (final concentrations were 10 mM, 9 mM, 17 mM, and 20 mM for peptides **1**-**4**, respectively). J-coupling values were measured on TOPSPIN through deconvolution and peak fitting using a mixture of Lorentzian and Gaussian curves; and using ½ the spectral resolution as an estimate of experimental uncertainty (0.4 Hz). Variable-concentration 1D spectra acquired for peptide **1** (1, 2, 5, and 10 mM) showed spectra were concentration-independent in the range of the experiment.

Appendix A Supplementary Information for Section 2.0

Table 8. MALDI-TOF MS data for peptides 1a and 1b

#	$[M+H]^+$ of reduced linear precursor *m/z* (monoisotopic)		$[M+H]^+$ of single disulfide species *m/z* (monoisotopic)		$[M+H]^+$ of fully oxidized product *m/z* (monoisotopic)	
	Calculated	**Observed**	**Calculated**	**Observed**	**Calculated**	**Observed**
1a	3463.7	3463.9	3461.6	3461.8	3317.6	3317.8
1b	3463.7	3463.8	3461.6	3461.8	3317.6	3317.8

Table 9. Statistics for NMR structure calculations for peptides 3-6

	3	**4**	**5**	**6**
Experimental restraints				
Unambiguous NOEs	356	434	389	435
Ambiguous NOEs	35	77	51	46
Total NOEs	391	511	440	481
H-bonds	26	24	24	26
Dihedrals	17	17	14	11
Violations				
NOE >0.5 Å	3.5 ± 1.3	5.0 ± 2.3	6.5 ± 2.5	5.5 ± 2.2
NOE rmsd (Å)	0.08 ± 0.01	0.08 ± 0.03	0.12 ± 0.03	0.09 ± 0.01
H-bond >0.5 Å	0	0	1.6 ± 2.1	0
Dihedral >5°	0	0	0	0
Ensemble rmsd				
Backbone heavy atoms	0.74 ± 0.23	0.63 ± 0.16	0.67 ± 0.14	0.56 ± 0.11
All heavy atoms	1.33 ± 0.18	1.22 ± 0.13	1.23 ± 0.11	1.05 ± 0.19
Geometry analysis				
rmsd bonds (Å)	0.0032 ± 0.0001	0.0033 ± 0.0001	0.0039 ± 0.0002	0.0038 ± 0.0001
rmsd angles (°)	0.41 ± 0.01	0.41 ± 0.02	0.55 ± 0.02	0.60 ± 0.01
rmsd impropers (°)	0.91 ± 0.08	1.02 ± 0.19	1.27 ± 0.07	1.20 ± 0.10
Ramachandran analysis[a]				
Favored (%)	93.8	99.1	86.8	93.2
Allowed (%)	6.2	0.9	12.1	6.8
Disallowed (%)	0	0	1.1	0

[a] Ramachandran analysis was done using the MolProbity server;[207] unnatural residues excluded.

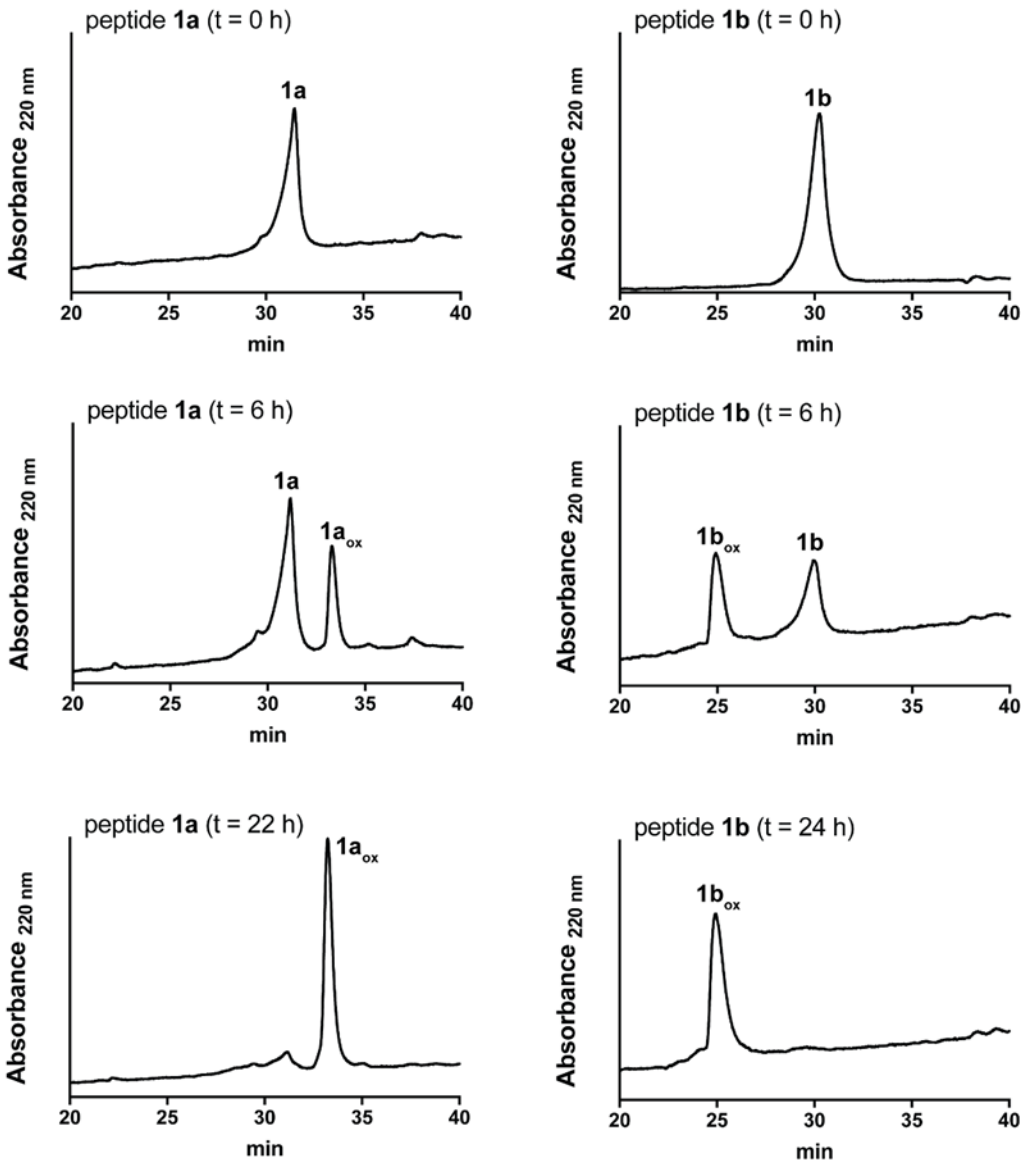


Figure 39. HPLC chromatograms showing the time course of the oxidation reactions to form peptides $1a_{ox}$ and $1b_{ox}$ from the corresponding reduced linear precursors 1a and 1b. Conditions: 0.25 mg/mL peptide in 0.1 M ammonium bicarbonate, pH 8.3.

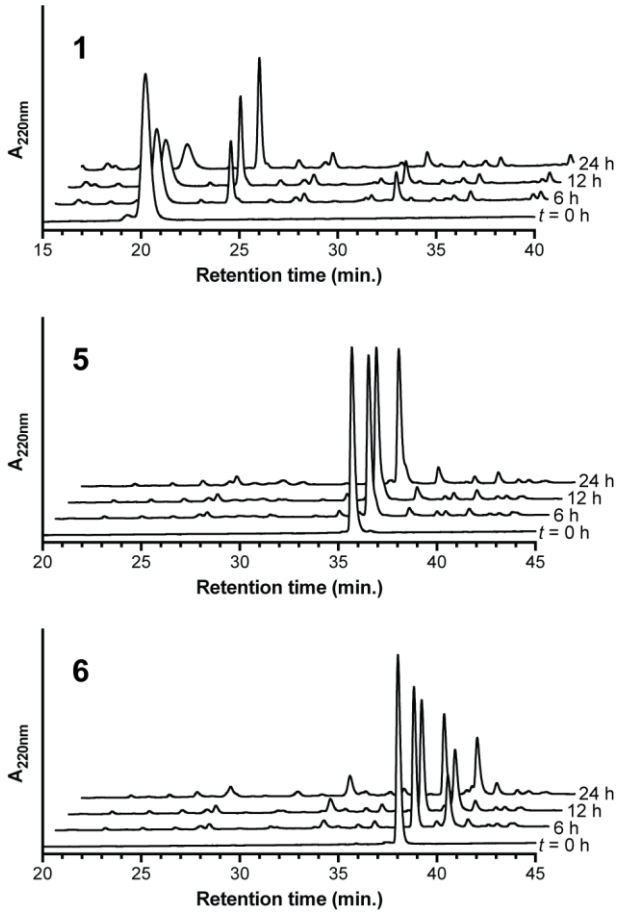


Figure 40. HPLC chromatograms showing the time course of the proteolysis of peptides 1, 5, and 6by trypsin. Each reaction was 100 µM peptide and 2.8 mg/mL enzyme in 0.1 M ammonium bicarbonate, pH 8.5.

Table 10. Observed m/z and putative sequences of peptide 1, 5, and 6 digests

Peptide	m/z	Sequence	Data Source
1	880.609	H-Tyr-Pro-Asn-Cys-Glu-Val-Arg-OH	MALDI
	1036.763	H-Arg-Tyr-Pro-Asn-Cys-Glu-Val-Arg-OH	MALDI
	748.421	H-Cys-Lys-Glu-Leu-Lys-Lys-OH	MALDI
5	413.20202 (z = 2)	H-Tyr-Aib-Gly-Cys-Glu-Nmev-Arg-OH	HRMS
	825.39080 (z = 1)	H-Tyr-Aib-Gly-Cys-Glu-Nmev-Arg-OH	HRMS
	1304.050	H-Asn-Asp-Blys-Cys-Lys-Bglu-Leu-Lys-Blys-Arg-OH	MALDI
6	377.88405 (z = 3)	H-Asn-Asp-Aib-Cys-Lys-Aib-Leu-Lys-Aib-Arg-OH	HRMS
	566.3188 (z = 2)	H-Asn-Asp-Aib-Cys-Lys-Aib-Leu-Lys-Aib-Arg-OH	HRMS
	1131.62857 (z = 1)	H-Asn-Asp-Aib-Cys-Lys-Aib-Leu-Lys-Aib-Arg-OH	HRMS
	413.20235 (z = 2)	H-Tyr-Aib-Gly-Cys-Glu-Nmev-Arg-OH	HRMS
	825.38954 (z = 1)	H-Tyr-Aib-Gly-Cys-Glu-Nmev-Arg-OH	HRMS

*Nmev = N-methyl valine, Blys = β^3- Lysine

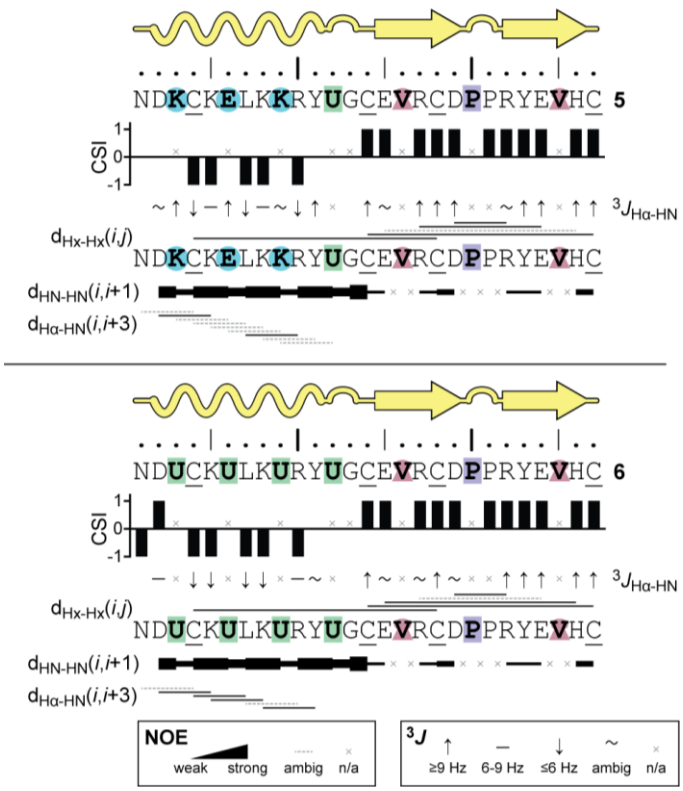


Figure 41. Summary of NMR data (chemical shift index values, vicinal HαHN coupling constants, and NOEs indicative of fold) for peptides 5and 6. A putative secondary structure map is provided for each sequence based on the analysis of the aggregate data.

Appendix B Supplementary Information for Section 3.0

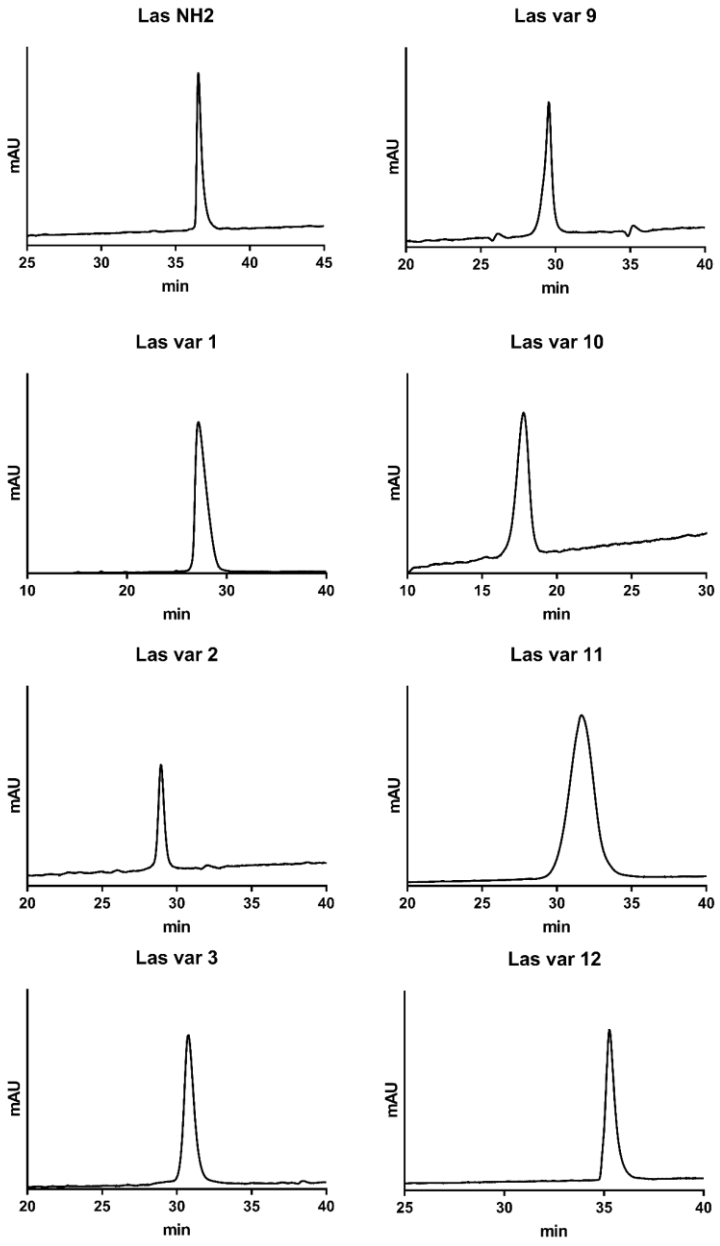
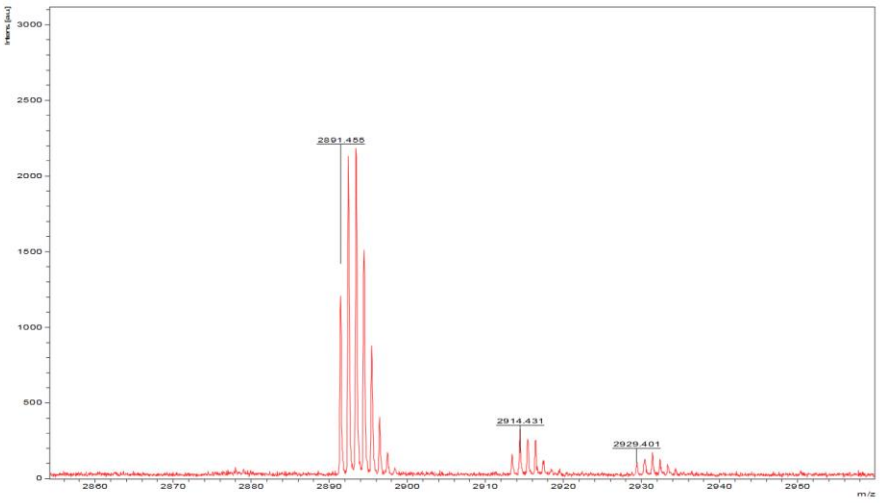


Figure 42. Analytical HPLC traces of fully oxidized and purified Lasiocepsin variants investigated in this study. Las var 1 and 3 were run on a 20-30% Solvent B gradient. Las NH_2, var 2, and var 9 – 12 were run on a 20-25% Solvent B gradient.

Table 11. MALDI-TOF MS data of fully oxidized Las variants investigated in this work

	[M+H]+ calc.	*m/z* obs.
Las NH_2	2891.716	2891.455
Las var 1	3003.842	3004.066
Las var 2	3031.873	3032.008
Las var 3	2983.780	2983.947
Las var 9	2927.717	2927.846
Las var 10	2983.780	2983.882
Las var 11	2997.795	2997.814
Las var 12	2957.764	2957.856



m/z	S/N	Quality Fac.	Res.	Intens.	Area
802.813	12	1701	7258	504	93.2
804.820	9	1216	6419	375	75.9
816.817	8	926	5821	341	79.4
829.387	22	4485	10162	874	130
829.835	8	1438	7056	344	63.1
830.833	9	958	6405	371	74.7
856.853	11	252	7359	445	82.2
906.952	6	1323	6009	281	71.7
932.963	7	1724	7288	319	72.3
2891.455	121	53207	13613	1411	1164
2914.431	13	5567	13545	159	126
2929.401	7	1077	10940	93.4	88.6

Figure 43. MALDI-TOF MS spectra of Las NH_2.

m/z	S/N	Quality Fac.	Res.	Intens.	Area
3004.066	178	8689	36559	891	280
3026.048	10	935	41048	50.8	14.8

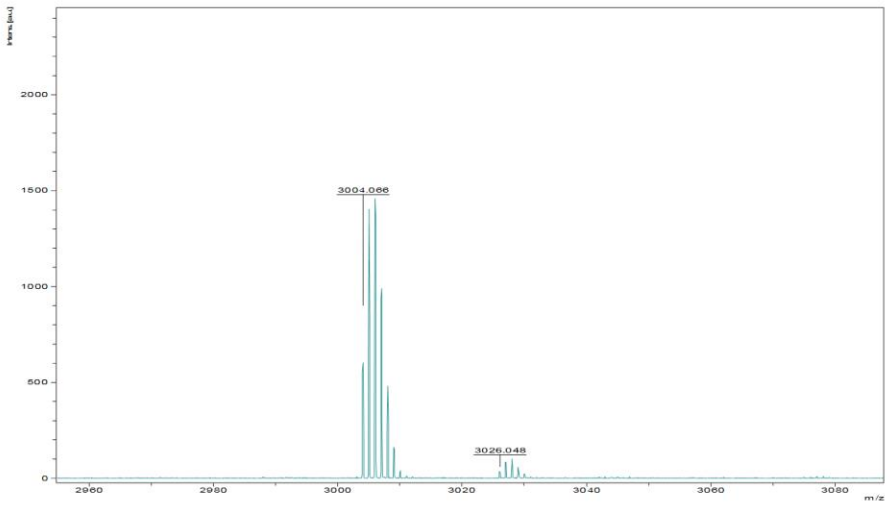


Figure 44. MALDI-TOF MS spectra of Las var 1.

m/z	S/N	Quality Fac.	Res.	Intens.	Area
3032.008	285	11653	41702	1330	369
3054.000	14	493	42609	63.2	17.4

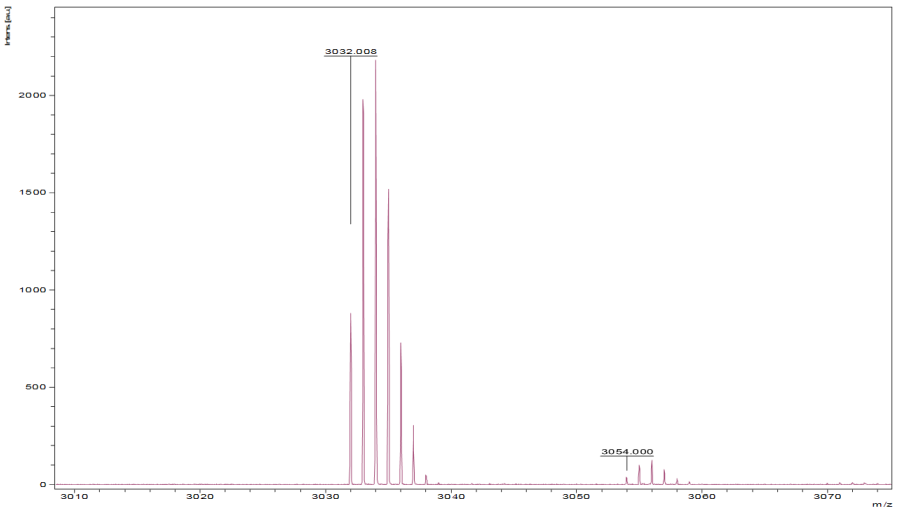


Figure 45. MALDI-TOF MS spectra of Las var 2.

m/z	S/N	Quality Fac.	Res.	Intens.	Area
2983.947	142	11587	42013	781	208
2986.954	15	295	32256	83.0	34.9
3005.928	16	3711	33156	89.8	32.3

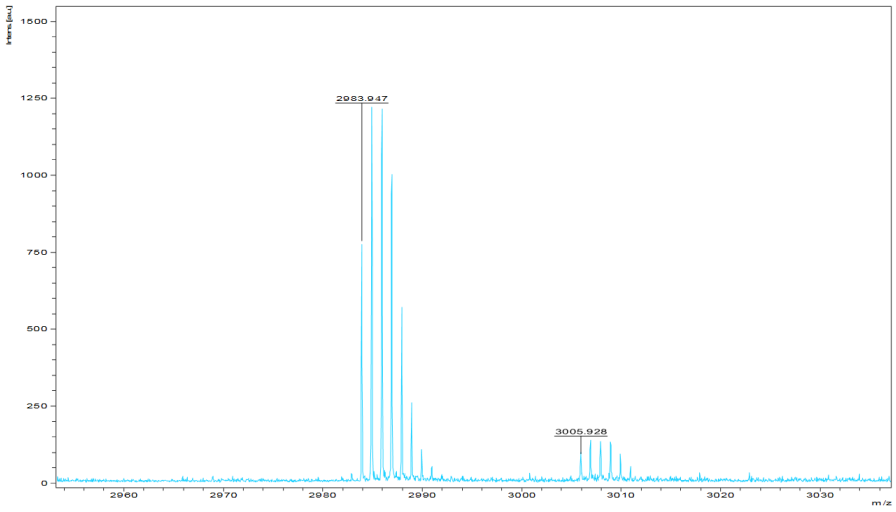


Figure 46. MALDI-TOF MS spectra of Las var 3.

m/z	S/N	Quality Fac.	Res.	Intens.	Area
1038.571	7	862	8678	345	68.9
1066.599	7	836	13940	367	49.8
1296.812	29	12258	16662	1244	228
1464.523	21	889	15208	920	213
1465.019	22	9083	15076	962	231
1475.497	11	331	21491	498	77.7
1475.997	11	1749	17359	505	99.8
1483.483	9	349	15364	431	96.4
1483.986	9	513	15445	438	96.4
1673.065	43	57523	17221	1740	503
1695.043	7	1487	15487	350	98.5
1711.012	7	1425	16870	347	91.4
2439.454	7	1323	13885	235	117
2471.438	7	1796	13514	229	122
2913.371	7	632	3954	200	342
2927.846	650	28262	11795	12499	12093
2930.761	67	472	6394	1358	2201
2942.839	19	2037	12510	429	351
2949.825	235	37127	14055	4514	3681
2965.800	115	31091	14741	2241	1716
2998.883	25	5965	15632	531	354
3020.847	6	658	18197	176	75.5

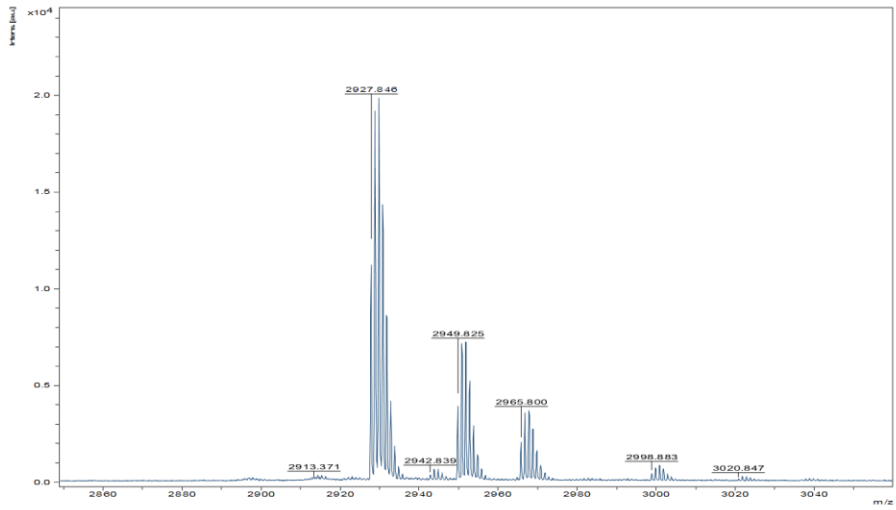
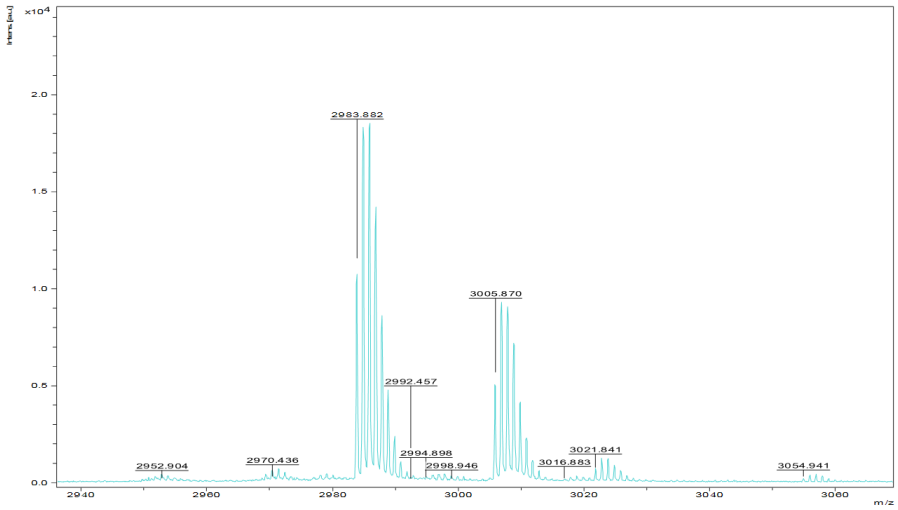
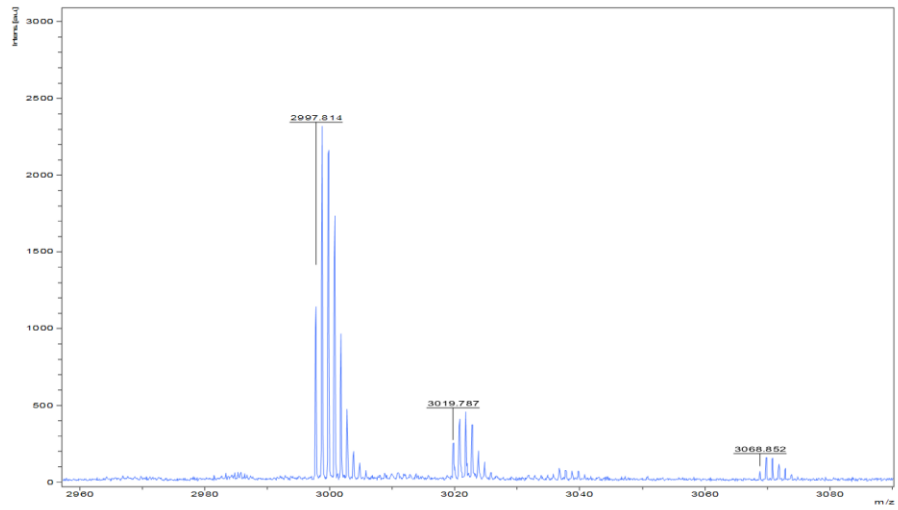


Figure 47. MALDI-TOF MS spectra of Las var 9.



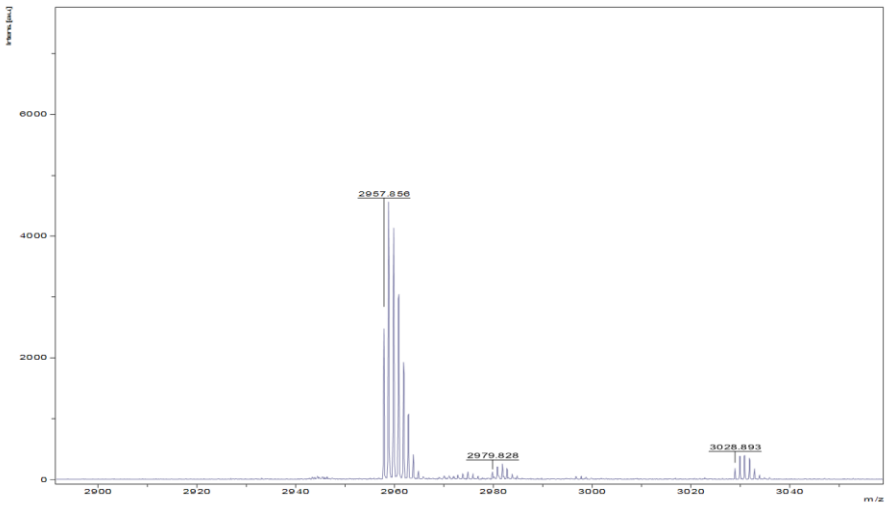
m/z	S/N	Quality Fac.	Res.	Intens.	Area
1492.474	9	500	35707	176	19.4
1503.467	7	905	24642	138	21.8
1503.973	9	3191	22328	165	28.0
2891.799	10	3966	20143	142	66.1
2927.805	20	3963	19051	270	136
2949.964	8	329	4115	142	265
2952.904	10	568	4822	157	261
2968.487	11	237	4968	168	279
2970.436	18	761	5208	248	442
2977.022	16	44.1	4464	208	495
2982.416	7	4.81	10363	122	85.9
2983.882	992	54699	12906	11483	10338
2988.875	46	896	34798	568	181
2992.457	8	216	4193	132	239
2994.898	15	937	5259	197	403
2998.946	8	221	5300	130	212
3005.870	488	47127	13764	5616	4896
3011.864	16	1902	20679	227	118
3016.883	10	471	5306	151	282
3021.841	58	27811	16494	698	486
3054.941	19	6192	20331	252	126
3120.869	10	7056	14435	132	96.9

Figure 48. MALDI-TOF MS spectra of Las var 10.



m/z	S/N	Quality Fac.	Res.	Intens.	Area
1097.831	7	388	25050	129	11.7
2997.814	176	13433	23868	1405	663
3000.807	19	495	14300	158	133
3019.787	33	4367	19578	265	167
3068.852	12	831	24497	94.0	42.8

Figure 49. MALDI-TOF MS spectra of Las var 11.



m/z	S/N	Quality Fac.	Res.	Intens.	Area
1480.012	19	1163	31815	133	17.5
1481.513	7	129	34683	51.4	6.03
2957.856	596	5753	19648	2812	1580
2961.843	57	334	25429	269	117
2972.848	14	1881	26078	67.9	30.3
2979.828	29	7130	25505	139	65.9
2980.104	6	369	17435	30.4	22.2
3028.893	54	1475	27957	249	102

Figure 50. MALDI-TOF MS spectra of Las var 12.

Table 12. Chemical shifts involved in backbone walk of Las var 9

	HN	Hα	Hβ
G1	7.561	3.852 (Hα1), 4.773 (Hα2)	-
L2	8.589	4.648	-
P3	-	4.48	-
R4	8.543	4.18	-
K5	8.311	4.119	-
ACPC6	7.826	2.697	4.212
L7	8.233	4.109	-
C8	8.257	4.735	-
A9	7.914	4.173	-
ACPC10	7.897	2.609	4.172
AIB11	8.244	-	-
K12	7.694	4.107	-
BK13	7.67	-	4.216
BK13'	7.734	-	4.21
K14	8.309	4.208	-
G15'	8.337	3.809	-
G15	8.378	-	-
K16	7.987	4.359	-
K16'	8.081	4.328	-
C17	8.522	4.573	-
C17'	8.458	4.642	-
K18	8.226	4.489	-
G19	8.489	3.943 (Hα1), 4.229 (Hα2)	-
P20	-	4.425	-
L21	8.019	4.227	-
K22	7.802	4.169	-
ACPC23	7.922	5.368	4.056
V24	7.945	3.965	2.117
C25	8.273	4.681	-
BK26'	8.004	2.545	4.102
BK26	7.937	-	4.075
C27	8.408	4.579	-
C27'	8.402	-	-

Table 13. Chemical shifts involved in backbone walk of Las var 11

	HN	Hα	Hβ
G1	-	3.844 (Qα)	-
L2	8.512	4.275	-
ACPC3	8.316	2.666	4.295
R4	8.181	4.204	-
BK5	7.889	-	4.203
ACPC6	8.144	2.643	4.203
L7	8.286	4.228	-
C8'	8.527	4.634	-
C8	8.497	4.641	-
A9	8.129	4.172	-
ACPC10	7.892	2.647	-
A11	8.227	4.159	-
K12	8.132	4.189	-
BK13	7.712	-	4.216
K14	8.375	4.155	-
a15	8.623	4.262	1.387
BK16	7.635	2.541	4.175
C17'	8.454	4.478	-
C17	8.413	4.547	-
BK18'	8.261	-	-
BK18	8.226	-	4.173
a19	8.176	4.624	-
a19'	8.264	4.183	-
P20	-	4.322	-
P20'	-	4.818	-
ACPC21'	8.376	2.717	4.308
ACPC21	8.047	2.63	4.195
K22'	8.337	4.241	-
K22	8.226	4.19	-
L23'	8.031	4.34	-
L23	8.165	4.277	-
BV24	7.805	-	4.012
BV24'	7.881	-	3.942
C25	8.32	4.654	-
C25'	8.722	4.594	-
K26'	8.37	4.362	-
K26	8.281	4.323	-
C27	8.446	4.568	-
C27'	8.347	4.589	-

Appendix C Supplementary Section for Section 4.0

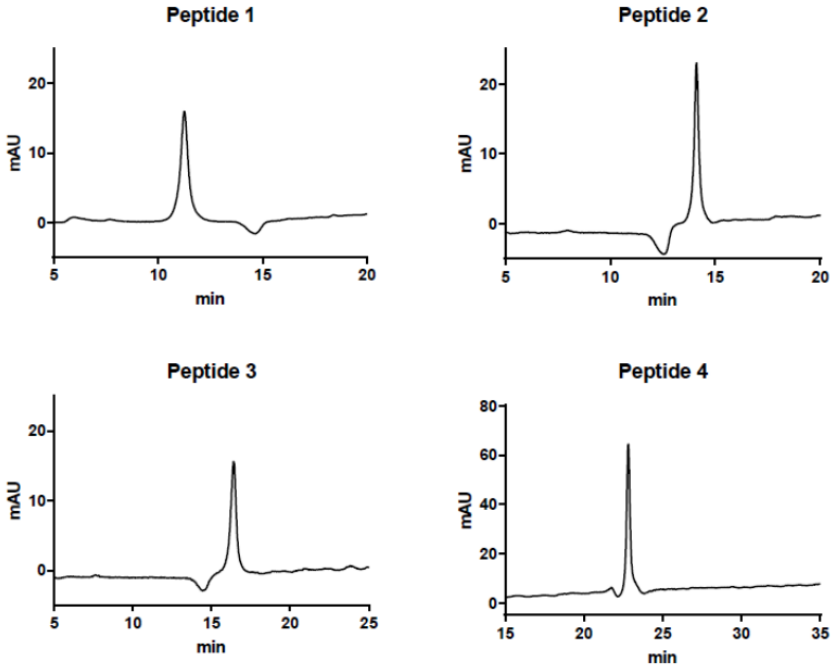


Figure 51. Analytical HPLC traces for purified peptides 1-4. Gradient was 0-10% solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) over 30 minutes for peptides 2 and 4; the gradient was 0-5% solvent B over 30 minutes for peptide 1 and 3.

Bibliography

1.Khan, S.; Vihinen, M., Spectrum of disease-causing mutations in protein secondary structures. *BMC Structural Biology* **2007,** *7*, 56.

2.Fosgerau, K.; Hoffmann, T., Peptide therapeutics: current status and future directions. *Drug Discovery Today* **2015,** *20* (1), 122-128.

3.McGregor, D. P., Discovering and improving novel peptide therapeutics. *Current Opinion in Pharmacology* **2008,** *8* (5), 616-619.

4.Tsomaia, N., Peptide therapeutics: Targeting the undruggable space. *European Journal of Medicinal Chemistry* **2015,** *94*, 459-470.

5.Gellman, S. H., Foldamers: A Manifesto. *Accounts of Chemical Research* **1998,** *31* (4), 173-180.

6.Horne, W. S.; Gellman, S. H., Foldamers with Heterogeneous Backbones. *Accounts of Chemical Research* **2008,** *41* (10), 1399-1408.

7.Horne, W. S.; Grossmann, T. N., Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns. *Nature Chemistry* **2020,** *12* (4), 331-337.

8.Werner, H. M.; Cabalteja, C. C.; Horne, W. S., Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution. *ChemBioChem* **2016,** *17* (8), 712-8.

9.Cheek, S.; Krishna, S. S.; Grishin, N. V., Structural Classification of Small, Disulfide-rich Protein Domains. *Journal of Molecular Biology* **2006,** *359* (1), 215-237.

10.Akondi, K. B.; Muttenthaler, M.; Dutertre, S.; Kaas, Q.; Craik, D. J.; Lewis, R. J.; Alewood, P. F., Discovery, synthesis, and structure-activity relationships of conotoxins. *Chemical Reviews* **2014,** *114* (11), 5815-47.

11.Mookherjee, N.; Anderson, M. A.; Haagsman, H. P.; Davidson, D. J., Antimicrobial host defence peptides: functions and clinical potential. *Nature Reviews Drug Discovery* **2020,** *19* (5), 311-332.

12.Wang, C. K.; Craik, D. J., Designing macrocyclic disulfide-rich peptides for biotechnological applications. *Nature Chemical Biology* **2018,** *14* (5), 417-427.

13.Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F., Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chemical Reviews* **2014,** *114* (2), 901-26.

14.Gehrmann, J.; Alewood, P. F.; Craik, D. J., Structure determination of the three disulfide bond isomers of α-conotoxin GI: a model for the role of disulfide bonds in structural stability11Edited by P. E. Wright. *Journal of Molecular Biology* **1998,** *278* (2), 401-415.

15.Zheng, Y.; Li, Z.; Ren, J.; Liu, W.; Wu, Y.; Zhao, Y.; Wu, C., Artificial disulfide-rich peptide scaffolds with precisely defined disulfide patterns and a minimized number of isomers. *Chemical Science* **2017,** *8* (4), 2547-2552.

16.Jin, A.-H.; Brandstaetter, H.; Nevin, S. T.; Tan, C. C.; Clark, R. J.; Adams, D. J.; Alewood, P. F.; Craik, D. J.; Daly, N. L., Structure of α-conotoxin BuIA: influences of disulfide connectivity on structural dynamics. *BMC Structural Biology* **2007,** *7* (1), 28.

17.Buczek, O.; Bulaj, G.; Olivera, B. M., Conotoxins and the posttranslational modification of secreted gene products. *Cell and Molecular Life Sciences* **2005,** *62* (24), 3067-79.

18.Jones, A. W., Early drug discovery and the rise of pharmaceutical chemistry. *Drug Testing and Analysis* **2011,** *3* (6), 337-344.

19.Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-36.

20.Pina, A. S.; Hussain, A.; Roque, A. C., An historical overview of drug discovery. *Methods Molecular Biology* **2009,** *572*, 3-12.

21.Drews, J., Drug discovery: a historical perspective. *Science* **2000,** *287* (5460), 1960-4.

22.Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2001,** *46* (1-3), 3-26.

23.Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The future of peptide-based drugs. *Chemical Biology and Drug Design* **2013,** *81* (1), 136-47.

24.Peters, J.-U.; Schnider, P.; Mattei, P.; Kansy, M., Pharmacological Promiscuity: Dependence on Compound Properties and Target Specificity in a Set of Recent Roche Compounds. *ChemMedChem* **2009,** *4* (4), 680-686.

25.Fauman, E. B.; Rai, B. K.; Huang, E. S., Structure-based druggability assessment—identifying suitable targets for small molecule therapeutics. *Current Opinion in Chemical Biology* **2011,** *15* (4), 463-468.

26.Lawson, A. D. G.; MacCoss, M.; Heer, J. P., Importance of Rigidity in Designing Small Molecule Drugs To Tackle Protein–Protein Interactions (PPIs) through Stabilization of Desired Conformers. *Journal of Medicinal Chemistry* **2018,** *61* (10), 4283-4289.

27.Baumann, A., Early development of therapeutic biologics--pharmacokinetics. *Curr. Drug Metab.* **2006,** *7* (1), 15-21.

28.Lipinski, C. A., Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* **2000,** *44* (1), 235-49.

29.Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nature Reviews Drug Discovery* **2008,** *7* (1), 21-39.

30.de la Torre, B. G.; Albericio, F., The Pharmaceutical Industry in 2019. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2020,** *25* (3).

31.Anselmo, A. C.; Gokarn, Y.; Mitragotri, S., Non-invasive delivery strategies for biologics. *Nature Reviews Drug Discovery* **2019,** *18* (1), 19-40.

32.Reed, G. A., Stability of Drugs, Drug Candidates, and Metabolites in Blood and Plasma. *Current Protocols in Pharmacology* **2016,** *75* (1), 7.6.1-7.6.12.

33.Xu, X.; Vugmeyster, Y., Challenges and opportunities in absorption, distribution, metabolism, and excretion studies of therapeutic biologics. *AAPS J* **2012,** *14* (4), 781-91.

34.Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* **2018,** *26* (10), 2700-2707.

35.Bray, B. L., Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nature Reviews Drug Discovery* **2003,** *2* (7), 587-93.

36.deGruyter, J. N.; Malins, L. R.; Baran, P. S., Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry* **2017,** *56* (30), 3863-3873.

37.Dozier, J. K.; Distefano, M. D., Site-Specific PEGylation of Therapeutic Proteins. *Int. J. Mol. Sci.* **2015,** *16* (10), 25831-64.

38.Roberts, M. J.; Bentley, M. D.; Harris, J. M., Chemistry for peptide and protein PEGylation. *Advanced Drug Delivery Reviews* **2012,** *64*, 116-127.

39.Lawrence, P. B.; Price, J. L., How PEGylation influences protein conformational stability. *Current Opinion in Chemical Biology* **2016,** *34*, 88-94.

40.Piedmonte, D. M.; Treuheit, M. J., Formulation of Neulasta (pegfilgrastim). *Advanced Drug Delivery Reviews* **2008,** *60* (1), 50-8.

41.Mikami, T.; Takao, T.; Yanagi, K.; Nakazawa, H., N (alpha) Selective Acetylation of Peptides. *Mass Spectrometry (Tokyo)* **2012,** *1* (2), A0010.

42.Weber, P. J. A.; Bader, J. E.; Folkers, G.; Beck-Sickinger, A. G., A fast and inexpensive method for N-terminal fluorescein-labeling of peptides. *Bioorganic & Medicinal Chemistry Letters* **1998,** *8* (6), 597-600.

43.Shcherbatko, A.; Rossi, A.; Foletti, D.; Zhu, G.; Bogin, O.; Galindo Casas, M.; Rickert, M.; Hasa-Moreno, A.; Bartsevich, V.; Crameri, A.; Steiner, A. R.; Henningsen, R.; Gill, A.; Pons, J.; Shelton, D. L.; Rajpal, A.; Strop, P., Engineering Highly Potent and Selective Microproteins against Nav1.7 Sodium Channel for Treatment of Pain. *J. Biol. Chem.* **2016,** *291* (27), 13974-86.

44.Marino, G.; Eckhard, U.; Overall, C. M., Protein Termini and Their Modifications Revealed by Positional Proteomics. *ACS Chemical Biology* **2015,** *10* (8), 1754-64.

45.Kim, K.-H.; Seong, B. L., Peptide amidation: Production of peptide hormonesin vivo andin vitro. *Biotechnology and Bioprocess Engineering* **2001,** *6* (4), 244-251.

46.Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B., Synthesis of proteins by native chemical ligation. *Science* **1994,** *266* (5186), 776-9.

47.Devadas, K.; Boykins, R. A.; Hardegen, N. J.; Philp, D.; Kleinman, H. K.; Osa, E.-O.; Wang, J.; Clouse, K. A.; Wahl, L. M.; Hewlett, I. K.; Rappaport, J.; Yamada, K. M.; Dhawan, S., Selective side-chain modification of cysteine and arginine residues blocks pathogenic activity of HIV-1-Tat functional peptides. *Peptides* **2006,** *27* (4), 611-621.

48.Shimizu, M.; Carter, P. H.; Khatri, A.; Potts, J. T., Jr.; Gardella, T. J., Enhanced activity in parathyroid hormone-(1-14) and -(1-11): novel peptides for probing ligand-receptor interactions. *Endocrinology* **2001,** *142* (7), 3068-74.

49.Witt, K. A.; Gillespie, T. J.; Huber, J. D.; Egleton, R. D.; Davis, T. P., Peptide drug modifications to enhance bioavailability and blood-brain barrier permeability. *Peptides* **2001,** *22* (12), 2329-43.

50.Buckley, S. T.; Hubalek, F.; Rahbek, U. L., Chemically modified peptides and proteins - critical considerations for oral delivery. *Tissue Barriers* **2016,** *4* (2), e1156805.

51.Ward, B. P.; Ottaway, N. L.; Perez-Tilve, D.; Ma, D.; Gelfanov, V. M.; Tschop, M. H.; Dimarchi, R. D., Peptide lipidation stabilizes structure to enhance biological function. *Molecular Metabolism* **2013,** *2* (4), 468-79.

52.Moradi, S. V.; Hussein, W. M.; Varamini, P.; Simerska, P.; Toth, I., Glycosylation, an effective synthetic strategy to improve the bioavailability of therapeutic peptides. *Chemical Science* **2016,** *7* (4), 2492-2500.

53.Schnolzer, M.; Kent, S., Constructing proteins by dovetailing unprotected synthetic peptides: Backbone-engineered HIV protease. *Science* **1992,** *256* (5054), 221-225.

54.Milton, R. C.; Milton, S. C.; Kent, S. B., Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity [corrected]. *Science* **1992,** *256* (5062), 1445-8.

55.Zawadzke, L. E.; Berg, J. M., A racemic protein. *Journal of the American Chemical Society* **1992,** *114* (10), 4002-4003.

56.Beligere, G. S.; Dawson, P. E., Design, Synthesis, and Characterization of 4-Ester CI2, a Model for Backbone Hydrogen Bonding in Protein α-Helices. *Journal of the American Chemical Society* **2000,** *122* (49), 12079-12082.

57.Chapman, E.; Thorson, J. S.; Schultz, P. G., Mutational Analysis of Backbone Hydrogen Bonds in Staphylococcal Nuclease. *Journal of the American Chemical Society* **1997,** *119* (30), 7151-7152.

58.Baca, M.; Alewood, P. F.; Kent, S. B., Structural engineering of the HIV-1 protease molecule with a beta-turn mimic of fixed geometry. *Protein Science* **1993,** *2* (7), 1085-91.

59.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. *Journal of the American Chemical Society* **2002,** *124* (18), 4972-3.

60.Chalmers, D. K.; Marshall, G. R., Pro-D-NMe-Amino Acid and D-Pro-NMe-Amino Acid: Simple, Efficient Reverse-Turn Constraints. *Journal of the American Chemical Society* **1995,** *117* (22), 5927-5937.

61.Marshall, G. R.; Ballante, F., Limiting Assumptions in the Design of Peptidomimetics. *Drug Development Research* **2017,** *78* (6), 245-267.

62.Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., A Field Guide to Foldamers. *Chemical Reviews* **2001,** *101* (12), 3893-4012.

63.George, K. L.; Horne, W. S., Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Accounts of Chemical Research* **2018,** *51* (5), 1220-1228.

64.Johnson, L. M.; Barrick, S.; Hager, M. V.; McFedries, A.; Homan, E. A.; Rabaglia, M. E.; Keller, M. P.; Attie, A. D.; Saghatelian, A.; Bisello, A.; Gellman, S. H., A potent alpha/beta-peptide analogue of GLP-1 with prolonged action in vivo. *Journal of the American Chemical Society* **2014,** *136* (37), 12848-51.

65.Cary, B. P.; Hager, M. V.; Gellman, S. H., Impact of Substitution Registry on the Receptor-Activation Profiles of Backbone-Modified Glucagon-like Peptide-1 Analogues. *ChemBioChem* **2019,** *20* (22), 2834-2840.

66.Horne, W. S.; Grossmann, T. N., Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns. *Nature Chemistry* **2020**.

67.Reinert, Z. E.; Horne, W. S., Protein backbone engineering as a strategy to advance foldamers toward the frontier of protein-like tertiary structure. *Organic & Biomolecular Chemistry* **2014,** *12* (44), 8796-8802.

68.George, K. L.; Horne, W. S., Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *Journal of the American Chemical Society* **2017,** *139* (23), 7931-7938.

69.Werner, H. M.; Estabrooks, S. K.; Preston, G. M.; Brodsky, J. L.; Horne, W. S., Exploring the Functional Consequences of Protein Backbone Alteration in Ubiquitin through Native Chemical Ligation. *ChemBioChem* **2019,** *20* (18), 2346-2350.

70.Zhang, Y.; Zhang, Q.; Wong, C. T. T.; Li, X., Chemoselective Peptide Cyclization and Bicyclization Directly on Unprotected Peptides. *Journal of the American Chemical Society* **2019,** *141* (31), 12274-12279.

71.White, C. J.; Yudin, A. K., Contemporary strategies for peptide macrocyclization. *Nature Chemistry* **2011,** *3* (7), 509-524.

72.Gilon, C.; Halle, D.; Chorev, M.; Selincer, Z.; Byk, G., Backbone cyclization: A new method for conferring conformational constraint on peptides. *Biopolymers* **1991,** *31* (6), 745-750.

73.Wu, X.; Stockdill, J. L.; Wang, P.; Danishefsky, S. J., Total Synthesis of Cyclosporine: Access to N-Methylated Peptides via Isonitrile Coupling Reactions. *Journal of the American Chemical Society* **2010,** *132* (12), 4098-4100.

74.Ma, C.; Chen, M.; Chu, W.; Tao, J.; Kong, D.; Zhang, M.; Feng, W., A Practical and Total Synthesis of Pasireotide: Synthesis of Cyclic Hexapeptide via a Three-Component Condensation. *Molecules* **2019,** *24* (11).

75.Tam, J. P.; Lu, Y.-A.; Yang, J.-L.; Chiu, K.-W., An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proceedings of the National Academy of Sciences* **1999,** *96* (16), 8913-8918.

76.Cheneval, O.; Schroeder, C. I.; Durek, T.; Walsh, P.; Huang, Y.-H.; Liras, S.; Price, D. A.; Craik, D. J., Fmoc-Based Synthesis of Disulfide-Rich Cyclic Peptides. *The Journal of Organic Chemistry* **2014,** *79* (12), 5538-5544.

77.Clark, R. J.; Craik, D. J., Chapter three - Engineering Cyclic Peptide Toxins. In *Methods Enzymol.*, Wittrup, K. D.; Verdine, G. L., Eds. Academic Press: 2012; Vol. 503, pp 57-74.

78.Vig, B. S.; Murray, T. F.; Aldrich, J. V., A Novel N-Terminal Cyclic Dynorphin A Analogue cycloN,5[Trp3,Trp4,Glu5] Dynorphin A-(1−11)NH2 That Lacks the Basic N-Terminus. *Journal of Medicinal Chemistry* **2003,** *46* (8), 1279-1282.

79.Vig, B. S.; Murray, T. F.; Aldrich, J. V., Synthesis of novel basic head-to-side-chain cyclic dynorphin A analogs: Strategies and side reactions. *Peptide Science* **2003,** *71* (6), 620-637.

80.Li, Y.; Yongye, A.; Giulianotti, M.; Martinez-Mayorga, K.; Yu, Y.; Houghten, R. A., Synthesis of Cyclic Peptides through Direct Aminolysis of Peptide Thioesters Catalyzed by Imidazole in Aqueous Organic Solutions. *Journal of Combinatorial Chemistry* **2009,** *11* (6), 1066-1072.

81.Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R., Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **2015,** *44* (1), 91-102.

82.Davies, J. S., The cyclization of peptides and depsipeptides. *Journal of Peptide Science* **2003,** *9* (8), 471-501.

83.Fairlie, D. P.; Dantas de Araujo, A., Stapling peptides using cysteine crosslinking. *Peptide Science* **2016,** *106* (6), 843-852.

84.Cantor, C. R.; Schimmel, P. R., *The conformation of biological macromolecules*. W. H. Freeman: San Francisco, 1980; p xxvii, 341, 24 p.

85.Grant, G. A., *Synthetic peptides : a user's guide*. 2nd ed.; Oxford University Press: Oxford ; New York, 2002; p x, 390 p.

86.Gori, A.; Gagni, P.; Rinaldi, S., Disulfide Bond Mimetics: Strategies and Challenges. *Chemistry* **2017,** *23* (60), 14987-14995.

87.Marino, S. M.; Gladyshev, V. N., Analysis and functional prediction of reactive cysteine residues. *J. Biol. Chem.* **2012,** *287* (7), 4419-25.

88.Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F., Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014,** *114* (2), 901-926.

89.Moroder, L.; Musiol, H.-J.; Götz, M.; Renner, C., Synthesis of single- and multiple-stranded cystine-rich peptides. *Peptide Science* **2005,** *80* (2-3), 85-97.

90.Iyer, S.; Acharya, K. R., Tying the knot: The cystine signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. *The FEBS Journal* **2011,** *278* (22), 4304-4322.

91.Daly, N. L.; Craik, D. J., Bioactive cystine knot proteins. *Current Opinion in Chemical Biology* **2011,** *15* (3), 362-368.

92.Kobayashi, Y.; Takashima, H.; Tamaoki, H.; Kyogoku, Y.; Lambert, P.; Kuroda, H.; Chino, N.; Watanabe, T. X.; Kimura, T.; Sakakibara, S.; Moroder, L., The cystine-stabilized α-helix: A common structural motif of ion-channel blocking neurotoxic peptides. *Biopolymers* **1991,** *31* (10), 1213-1220.

93.Wilson, D.; Daly, N. L., Venomics: A Mini-Review. *High Throughput* **2018,** *7* (3).

94.Ramírez, D.; Gonzalez, W.; Fissore, R. A.; Carvacho, I., Conotoxins as Tools to Understand the Physiological Function of Voltage-Gated Calcium (Ca(V)) Channels. *Mar Drugs* **2017,** *15* (10), 313.

95.Clark-Lewis, I.; Hood, L. E.; Kent, S. B., Role of disulfide bridges in determining the biological activity of interleukin 3. *Proceedings of the National Academy of Sciences* **1988,** *85* (21), 7897.

96.Xu, X.; Xu, Q.; Chen, F.; Shi, J.; Liu, Y.; Chu, Y.; Wan, S.; Jiang, T.; Yu, R., Role of the disulfide bond on the structure and activity of μ-conotoxin PIIIA in the inhibition of NaV1.4. *RSC Advances* **2019,** *9* (2), 668-674.

97.Nishiuchi, Y.; Sakakibara, S., Primary and secondary structure of conotoxin GI, a neurotoxic tridecapeptide from a marine snail. *FEBS Letters* **1982,** *148* (2), 260-262.

98.Wilcox, G., Insulin and insulin resistance. *The Clinical Biochemist Reviews* **2005,** *26* (2), 19-39.

99.Sharman, A.; Low, J., Vasopressin and its role in critical care. *Continuing Education in Anaesthesia Critical Care & Pain* **2008,** *8* (4), 134-137.

100.Qureshi, S.; Galiveeti, S.; Bichet, D. G.; Roth, J., Diabetes Insipidus: Celebrating a Century of Vasopressin Therapy. *Endocrinology* **2014,** *155* (12), 4605-4621.

101.Kidson, W.; Casey, J.; Kraegen, E.; Lazarus, L., Treatment of Severe Diabetes Mellitus by Insulin Infusion. *British Medical Journal* **1974,** *2* (5921), 691.

102.Ganz, T., Defensins: antimicrobial peptides of innate immunity. *Nature Reviews Immunology* **2003,** *3* (9), 710-720.

103.Shariff, L.; Zhu, Y.; Cowper, B.; Di, W.-L.; Macmillan, D., Sunflower trypsin inhibitor (SFTI-1) analogues of synthetic and biological origin via N→S acyl transfer: potential inhibitors of human Kallikrein-5 (KLK5). *Tetrahedron* **2014,** *70* (42), 7675-7680.

104.Huang, Y. H.; Colgrave, M. L.; Daly, N. L.; Keleshian, A.; Martinac, B.; Craik, D. J., The biological activity of the prototypic cyclotide kalata b1 is modulated by the formation of multimeric pores. *J. Biol. Chem.* **2009,** *284* (31), 20699-707.

105.Benham, C. J.; Saleet Jafri, M., Disulfide bonding patterns and protein topologies. *Protein Science* **1993,** *2* (1), 41-54.

106.Fass, D., Disulfide Bonding in Protein Biophysics. *Annual Review of Biophysics* **2012,** *41* (1), 63-79.

107.Thomas, R. H.; Allmond, K., Linaclotide (Linzess) for Irritable Bowel syndrome With Constipation and For Chronic Idiopathic Constipation. *P & T : a peer-reviewed journal for formulary management* **2013,** *38* (3), 154-60.

108.Wang, C. K.; Gruber, C. W.; Cemazar, M.; Siatskas, C.; Tagore, P.; Payne, N.; Sun, G.; Wang, S.; Bernard, C. C.; Craik, D. J., Molecular Grafting onto a Stable Framework Yields Novel Cyclic Peptides for the Treatment of Multiple Sclerosis. *ACS Chemical Biology* **2014,** *9* (1), 156-163.

109.Sato, K.; Raymond, C.; Martin-Moutot, N.; Sasaki, T.; Ohtake, A.; Minami, K.; Van Renterghem, C.; Takahashi, M.; Seagar, M. J., Binding of Six Chimeric Analogs of ω-Conotoxin MVIIA and MVIIC to N- and P/Q-Type Calcium Channels. *Biochemical and Biophysical Research Communications* **2000,** *269* (1), 254-256.

110.King, G. F., Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opinion on Biological Therapy* **2011,** *11* (11), 1469-1484.

111.Muttenthaler, M.; Nevin, S. T.; Grishin, A. A.; Ngo, S. T.; Choy, P. T.; Daly, N. L.; Hu, S.-H.; Armishaw, C. J.; Wang, C.-I. A.; Lewis, R. J.; Martin, J. L.; Noakes, P. G.; Craik, D. J.; Adams, D. J.; Alewood, P. F., Solving the α-Conotoxin Folding Problem: Efficient Selenium-Directed On-Resin Generation of More Potent and Stable Nicotinic Acetylcholine Receptor Antagonists. *Journal of the American Chemical Society* **2010,** *132* (10), 3514-3522.

112.Wang, J.; Yadav, V.; Smart, A. L.; Tajiri, S.; Basit, A. W., Toward Oral Delivery of Biopharmaceuticals: An Assessment of the Gastrointestinal Stability of 17 Peptide Drugs. *Molecular Pharmaceutics* **2015,** *12* (3), 966-973.

113.Kikuchi, K.; Sugiura, M.; Kimura, T., High Proteolytic Resistance of Spider-Derived Inhibitor Cystine Knots. *International Journal of Peptides* **2015,** *2015*, 537508.

114.Armishaw, C. J.; Daly, N. L.; Nevin, S. T.; Adams, D. J.; Craik, D. J.; Alewood, P. F., Alpha-selenoconotoxins, a new class of potent alpha7 neuronal nicotinic receptor antagonists. *J. Biol. Chem.* **2006,** *281* (20), 14136-43.

115.Wu, Z.; Hoover, D. M.; Yang, D.; Boulègue, C.; Santamaria, F.; Oppenheim, J. J.; Lubkowski, J.; Lu, W., Engineering disulfide bridges to dissect antimicrobial and chemotactic activities of human beta-defensin 3. *Proc Natl Acad Sci U S A* **2003,** *100* (15), 8880-8885.

116.Besse, D.; Moroder, L., Synthesis of selenocysteine peptides and their oxidation to diselenide-bridged compounds. *Journal of Peptide Science* **1997,** *3* (6), 442-53.

117.Aoki, K.; Maeda, M.; Nakae, T.; Okada, Y.; Ohya, K.; Chiba, K., A disulfide bond replacement strategy enables the efficient design of artificial therapeutic peptides. *Tetrahedron* **2014,** *70* (42), 7774-7779.

118.Huang, D.-L.; Bai, J.-S.; Wu, M.; Wang, X.; Riedl, B.; Pook, E.; Alt, C.; Erny, M.; Li, Y.-M.; Bierer, D.; Shi, J.; Fang, G.-M., Non-reducible disulfide bond replacement implies that disulfide

exchange is not required for hepcidin–ferroportin interaction. *Chemical Communications* **2019,** *55* (19), 2821-2824.

119.Knuhtsen, A.; Whitmore, C.; McWhinnie, F. S.; McDougall, L.; Whiting, R.; Smith, B. O.; Timperley, C. M.; Green, A. C.; Kinnear, K. I.; Jamieson, A. G., α-Conotoxin GI triazole-peptidomimetics: potent and stable blockers of a human acetylcholine receptor. *Chemical Science* **2019,** *10* (6), 1671-1676.

120.Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V. S. R. K.; Kaas, Q.; Eletsky, A.; Huang, P.-S.; Johnsen, W. A.; Greisen, P., Jr.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsias, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D., Accurate de novo design of hyperstable constrained peptides. *Nature* **2016,** *538*, 329.

121.Hosseinzadeh, P.; Bhardwaj, G.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Rettie, S. A.; Kim, D. E.; Silva, D.-A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D., Comprehensive computational design of ordered peptide macrocycles. *Science (New York, N.Y.)* **2017,** *358* (6369), 1461-1466.

122.Green, B. R.; Catlin, P.; Zhang, M.-M.; Fiedler, B.; Bayudan, W.; Morrison, A.; Norton, Raymond S.; Smith, B. J.; Yoshikami, D.; Olivera, B. M.; Bulaj, G., Conotoxins Containing Nonnatural Backbone Spacers: Cladistic-Based Design, Chemical Synthesis, and Improved Analgesic Activity. *Chemistry & Biology* **2007,** *14* (4), 399-407.

123.Craik, D. J.; Adams, D. J., Chemical Modification of Conotoxins to Improve Stability and Activity. *ACS Chemical Biology* **2007,** *2* (7), 457-468.

124.Pennington, M. W.; Czerwinski, A.; Norton, R. S., Peptide therapeutics from venom: Current status and potential. *Bioorganic & Medicinal Chemistry* **2018,** *26* (10), 2738-2758.

125.Northfield, S. E.; Wang, C. K.; Schroeder, C. I.; Durek, T.; Kan, M.-W.; Swedberg, J. E.; Craik, D. J., Disulfide-rich macrocyclic peptides as templates in drug design. *European Journal of Medicinal Chemistry* **2014,** *77*, 248-257.

126.Vazquez-Lombardi, R.; Phan, T. G.; Zimmermann, C.; Lowe, D.; Jermutus, L.; Christ, D., Challenges and opportunities for non-antibody scaffold drugs. *Drug Discovery Today* **2015,** *20* (10), 1271-1283.

127.Krause, S.; Schmoldt, H.-U.; Wentzel, A.; Ballmaier, M.; Friedrich, K.; Kolmar, H., Grafting of thrombopoietin-mimetic peptides into cystine knot miniproteins yields high-affinity thrombopoietin antagonists and agonists. *The FEBS Journal* **2007,** *274* (1), 86-95.

128.Harris, M. M.; Coon, Z.; Alqaeisoom, N.; Swords, B.; Holub, J. M., Targeting anti-apoptotic Bcl2 proteins with scyllatoxin-based BH3 domain mimetics. *Organic & Biomolecular Chemistry* **2016,** *14* (2), 440-446.

129.Chevalier, A.; Silva, D.-A.; Rocklin, G. J.; Hicks, D. R.; Vergara, R.; Murapa, P.; Bernard, S. M.; Zhang, L.; Lam, K.-H.; Yao, G.; Bahl, C. D.; Miyashita, S.-I.; Goreshnik, I.; Fuller, J. T.; Koday, M. T.; Jenkins, C. M.; Colvin, T.; Carter, L.; Bohn, A.; Bryan, C. M.; Fernández-Velasco, D. A.; Stewart, L.; Dong, M.; Huang, X.; Jin, R.; Wilson, I. A.; Fuller, D. H.; Baker, D., Massively parallel de novo protein design for targeted therapeutics. *Nature* **2017,** *550* (7674), 74-79.

130.Kimura, R. H.; Levin, A. M.; Cochran, F. V.; Cochran, J. R., Engineered cystine knot peptides that bind αvβ3, αvβ5, and α5β1 integrins with low-nanomolar affinity. *Proteins: Structure, Function, and Bioinformatics* **2009,** *77* (2), 359-369.

131.Henriques, S. T.; Deplazes, E.; Lawrence, N.; Cheneval, O.; Chaousis, S.; Inserra, M.; Thongyoo, P.; King, G. F.; Mark, A. E.; Vetter, I.; Craik, D. J.; Schroeder, C. I., Interaction of Tarantula Venom Peptide ProTx-II with Lipid Membranes Is a Prerequisite for Its Inhibition of Human Voltage-gated Sodium Channel NaV1.7. *J. Biol. Chem.* **2016,** *291* (33), 17049-65.

132.Gui, J.; Liu, B.; Cao, G.; Lipchik, Andrew M.; Perez, M.; Dekan, Z.; Mobli, M.; Daly, Norelle L.; Alewood, Paul F.; Parker, Laurie L.; King, Glenn F.; Zhou, Y.; Jordt, S.-E.; Nitabach, Michael N., A Tarantula-Venom Peptide Antagonizes the TRPA1 Nociceptor Ion Channel by Binding to the S1–S4 Gating Domain. *Current Biology* **2014,** *24* (5), 473-483.

133.Clark, R. J.; Fischer, H.; Dempster, L.; Daly, N. L.; Rosengren, K. J.; Nevin, S. T.; Meunier, F. A.; Adams, D. J.; Craik, D. J., Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the α-conotoxin MII. *Proc Natl Acad Sci U S A* **2005,** *102* (39), 13767-13772.

134.Clark, R. J.; Jensen, J.; Nevin, S. T.; Callaghan, B. P.; Adams, D. J.; Craik, D. J., The Engineering of an Orally Active Conotoxin for the Treatment of Neuropathic Pain. *Angewandte Chemie International Edition* **2010,** *49* (37), 6545-6548.

135.Gori, A.; Wang, C.-I. A.; Harvey, P. J.; Rosengren, K. J.; Bhola, R. F.; Gelmi, M. L.; Longhi, R.; Christie, M. J.; Lewis, R. J.; Alewood, P. F.; Brust, A., Stabilization of the Cysteine-Rich Conotoxin MrIA by Using a 1,2,3-Triazole as a Disulfide Bond Mimetic. *Angewandte Chemie International Edition* **2015,** *54* (4), 1361-1364.

136.Wright, Z. V. F.; McCarthy, S.; Dickman, R.; Reyes, F. E.; Sanchez-Martinez, S.; Cryar, A.; Kilford, I.; Hall, A.; Takle, A. K.; Topf, M.; Gonen, T.; Thalassinos, K.; Tabor, A. B., The Role of Disulfide Bond Replacements in Analogues of the Tarantula Toxin ProTx-II and Their Effects on Inhibition of the Voltage-Gated Sodium Ion Channel Nav1.7. *Journal of the American Chemical Society* **2017,** *139* (37), 13063-13075.

137.Arai, K.; Takei, T.; Okumura, M.; Watanabe, S.; Amagai, Y.; Asahina, Y.; Moroder, L.; Hojo, H.; Inaba, K.; Iwaoka, M., Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue. *Angewandte Chemie International Edition* **2017,** *56* (20), 5522-5526.

138.Gopalakrishnan, R.; Frolov, A. I.; Knerr, L.; Drury, W. J.; Valeur, E., Therapeutic Potential of Foldamers: From Chemical Biology Tools To Drug Candidates? *Journal of Medicinal Chemistry* **2016,** *59* (21), 9599-9621.

139.Reinert, Z. E.; Lengyel, G. A.; Horne, W. S., Protein-like Tertiary Folding Behavior from Heterogeneous Backbones. *Journal of the American Chemical Society* **2013,** *135* (34), 12528-12531.

140.Mong, S. K.; Cochran, F. V.; Yu, H.; Graziano, Z.; Lin, Y.-S.; Cochran, J. R.; Pentelute, B. L., Heterochiral Knottin Protein: Folding and Solution Structure. *Biochemistry* **2017,** *56* (43), 5720-5725.

141.Grishin, A. A.; Wang, C. I.; Muttenthaler, M.; Alewood, P. F.; Lewis, R. J.; Adams, D. J., Alpha-conotoxin AuIB isomers exhibit distinct inhibitory mechanisms and differential sensitivity to stoichiometry of alpha3beta4 nicotinic acetylcholine receptors. *J. Biol. Chem.* **2010,** *285* (29), 22254-63.

142.Safavi-Hemami, H.; Gorasia, D. G.; Steiner, A. M.; Williamson, N. A.; Karas, J. A.; Gajewiak, J.; Olivera, B. M.; Bulaj, G.; Purcell, A. W., Modulation of conotoxin structure and function is achieved through a multienzyme complex in the venom glands of cone snails. *J. Biol. Chem.* **2012,** *287* (41), 34288-303.

143.Monincová, L.; Slaninová, J.; Fučík, V.; Hovorka, O.; Voburka, Z.; Bednárová, L.; Maloň, P.; Štokrová, J.; Čeřovský, V., Lasiocepsin, a novel cyclic antimicrobial peptide from the venom of eusocial bee Lasioglossum laticeps (Hymenoptera: Halictidae). *Amino Acids* **2012,** *43* (2), 751-761.

144.Zhangsun, D.; Zhu, X.; Kaas, Q.; Wu, Y.; Craik, D. J.; McIntosh, J. M.; Luo, S., αO-Conotoxin GeXIVA disulfide bond isomers exhibit differential sensitivity for various nicotinic acetylcholine receptors but retain potency and selectivity for the human α9α10 subtype. *Neuropharmacology* **2017,** *127*, 243-252.

145.Stanger, H. E.; Gellman, S. H., Rules for Antiparallel β-Sheet Design: d-Pro-Gly Is Superior to l-Asn-Gly for β-Hairpin Nucleation1. *Journal of the American Chemical Society* **1998,** *120* (17), 4236-4237.

146.Gessmann, R.; Brueckner, H.; Kokkinidis, M., Structural flexibility of Aib-containing peptides: The N-terminal tripeptide of trichotoxin. *Biochemical and Biophysical Research Communications* **1991,** *174* (2), 878-884.

147.Rai, R.; Raghothama, S.; Sridharan, R.; Balaram, P., Tuning the β-turn segment in designed peptide β-hairpins: Construction of a stable type I′ β-turn nucleus and hairpin–helix transition promoting segments. *Peptide Science* **2007,** *88* (3), 350-361.

148.Nowick, J. S.; Brower, J. O., A New Turn Structure for the Formation of β-Hairpins in Peptides. *Journal of the American Chemical Society* **2003,** *125* (4), 876-877.

149.Reinert, Z. E.; Horne, W. S., Folding Thermodynamics of Protein-Like Oligomers with Heterogeneous Backbones. *Chemical Science* **2014,** *5* (8), 3325-3330.

150.Tavenor, N. A.; Reinert, Z. E.; Lengyel, G. A.; Griffith, B. D.; Horne, W. S., Comparison of design strategies for alpha-helix backbone modification in a protein tertiary fold. *Chemical Communications (Camb)* **2016,** *52* (19), 3789-92.

151.Lengyel, G. A.; Reinert, Z. E.; Griffith, B. D.; Horne, W. S., Comparison of backbone modification in protein beta-sheets by alpha-->gamma residue replacement and alpha-residue methylation. *Organic & Biomolecular Chemistry* **2014,** *12* (29), 5375-81.

152.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* (6), 1647-51.

153.Werner, H. M.; Cabalteja, C. C.; Horne, W. S., Peptide backbone composition and protease susceptibility: Impact of modification type, position, and tandem substitution. *ChemBioChem* **2016,** *17* (8), 712-718.

154.Hubbard, S. J., The structural aspects of limited proteolysis of native proteins. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998,** *1382* (2), 191-206.

155.Perona, J. J.; Craik, C. S., Structural basis of substrate specificity in the serine proteases. *Protein Science* **1995,** *4* (3), 337-60.

156.Siepen, J. A.; Keevil, E.-J.; Knight, D.; Hubbard, S. J., Prediction of Missed Cleavage Sites in Tryptic Peptides Aids Protein Identification in Proteomics. *Journal of Proteome Research* **2007,** *6* (1), 399-408.

157.Yekkirala, A. S.; Roberson, D. P.; Bean, B. P.; Woolf, C. J., Breaking barriers to novel analgesic drug development. *Nat. Rev. Drug. Discov.* **2017,** *16*, 545.

158.Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015,** *31* (8), 1325-1327.

159.Kjaergaard, M.; Poulsen, F. M., Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of Biomolecular NMR* **2011,** *50* (2), 157-165.

160.Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, A.; Malliavin, T. E.; Nilges, M., ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **2007,** *23* (3), 381-382.

161.Brunger, A. T., Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2007,** *2*, 2728.

162.Mareuil, F.; Malliavin, T. E.; Nilges, M.; Bardiaux, B., Improved reliability, accuracy and quality in automated NMR structure calculation with ARIA. *Journal of Biomolecular NMR* **2015,** *62* (4), 425-438.

163.Lomazzi, M.; Moore, M.; Johnson, A.; Balasegaram, M.; Borisch, B., Antimicrobial resistance – moving forward? *BMC Public Health* **2019,** *19* (1), 858.

164.Das, S.; Adler, A. L.; Miles-Jay, A.; Kronman, M. P.; Qin, X.; Weissman, S. J.; Burnham, C.-A. D.; Elward, A.; Newland, J. G.; Selvarangan, R.; Sullivan, K. V.; Zaoutis, T.; Zerr, D. M., Antibiotic Prophylaxis Is Associated with Subsequent Resistant Infections in Children with an Initial Extended-Spectrum-Cephalosporin-Resistant Enterobacteriaceae Infection. *Antimicrob Agents Chemother* **2017,** *61* (5), e02656-16.

165.MacLean, R. C.; San Millan, A., The evolution of antibiotic resistance. *Science* **2019,** *365* (6458), 1082.

166.Ventola, C. L., The antibiotic resistance crisis: part 1: causes and threats. *P & T : a peer-reviewed journal for formulary management* **2015,** *40* (4), 277-283.

167.Zhang, L.-j.; Gallo, R. L., Antimicrobial peptides. *Current Biology* **2016,** *26* (1), R14-R19.

168.Izadpanah, A.; Gallo, R. L., Antimicrobial peptides. *Journal of the American Academy of Dermatology* **2005,** *52* (3), 381-390.

169.Mahlapuu, M.; Håkansson, J.; Ringstad, L.; Björn, C., Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. *Frontiers in Cellular and Infection Microbiology* **2016,** *6*, 194.

170.Lei, J.; Sun, L.; Huang, S.; Zhu, C.; Li, P.; He, J.; Mackey, V.; Coy, D. H.; He, Q., The antimicrobial peptides and their potential clinical applications. *Am J Transl Res* **2019,** *11* (7), 3919-3931.

171.Ma, Y.-X.; Wang, C.-Y.; Li, Y.-Y.; Li, J.; Wan, Q.-Q.; Chen, J.-H.; Tay, F. R.; Niu, L.-N., Considerations and Caveats in Combating ESKAPE Pathogens against Nosocomial Infections. *Advanced Science* **2020,** *7* (1), 1901872.

172.Deslouches, B.; Steckbeck, J. D.; Craigo, J. K.; Doi, Y.; Burns, J. L.; Montelaro, R. C., Engineered cationic antimicrobial peptides to overcome multidrug resistance by ESKAPE pathogens. *Antimicrob Agents Chemother* **2015,** *59* (2), 1329-1333.

173.Mohamed, M. F.; Abdelkhalek, A.; Seleem, M. N., Evaluation of short synthetic antimicrobial peptides for treatment of drug-resistant and intracellular Staphylococcus aureus. *Scientific Reports* **2016,** *6* (1), 29707.

174.Tai, K. P.; Kamdar, K.; Yamaki, J.; Le, V. V.; Tran, D.; Tran, P.; Selsted, M. E.; Ouellette, A. J.; Wong-Beringer, A., Microbicidal effects of α- and θ-defensins against antibiotic-resistant Staphylococcus aureus and Pseudomonas aeruginosa. *Innate Immun* **2015,** *21* (1), 17-29.

175.Tam, J. P.; Wang, S.; Wong, K. H.; Tan, W. L., Antimicrobial Peptides from Plants. *Pharmaceuticals (Basel)* **2015,** *8* (4), 711-757.

176.Cesa-Luna, C.; Muñoz-Rojas, J.; Saab-Rincon, G.; Baez, A.; Morales-García, Y. E.; Juárez-González, V. R.; Quintero-Hernández, V., Structural characterization of scorpion peptides and their bactericidal activity against clinical isolates of multidrug-resistant bacteria. *PLOS ONE* **2019,** *14* (11), e0222438.

177.Pedron, C. N.; Araújo, I.; da Silva Junior, P. I.; Dias da Silva, F.; Torres, M. D. T.; Oliveira Junior, V. X., Repurposing the scorpion venom peptide VmCT1 into an active peptide against Gram-negative ESKAPE pathogens. *Bioorganic Chemistry* **2019,** *90*, 103038.

178.Wang, P.; Nan, Y. H.; Yang, S.-T.; Kang, S. W.; Kim, Y.; Park, I.-S.; Hahm, K.-S.; Shin, S. Y., Cell selectivity and anti-inflammatory activity of a Leu/Lys-rich α-helical model antimicrobial peptide and its diastereomeric peptides. *Peptides* **2010,** *31* (7), 1251-1261.

179.Pfalzgraff, A.; Brandenburg, K.; Weindl, G., Antimicrobial Peptides and Their Therapeutic Potential for Bacterial Skin Infections and Wounds. *Frontiers in Pharmacology* **2018,** *9* (281).

180.Monincová, L.; Buděšínský, M.; Čujová, S.; Čeřovský, V.; Veverka, V., Structural Basis for Antimicrobial Activity of Lasiocepsin. *ChemBioChem* **2014,** *15* (2), 301-308.

181.Horne, W. S.; Price, J. L.; Gellman, S. H., Interplay among side chain sequence, backbone composition, and residue rigidification in polypeptide folding and assembly. *Proceedings of the National Academy of Sciences* **2008,** *105* (27), 9151.

182.Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D., 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR* **1995,** *5* (1), 67-81.

183.Cabalteja, C. C.; Mihalko, D. S.; Horne, W. S., Heterogeneous-Backbone Foldamer Mimics of a Computationally Designed, Disulfide-Rich Miniprotein. *ChemBioChem* **2019,** *20* (1), 103-110.

184.Ruiz, J.; Calderon, J.; Rondón-Villarreal, P.; Torres, R. In *Analysis of Structure and Hemolytic Activity Relationships of Antimicrobial Peptides (AMPs)*, Cham, Springer International Publishing: Cham, 2014; pp 253-258.

185.Oddo, A.; Hansen, P. R., Hemolytic Activity of Antimicrobial Peptides. In *Antimicrobial Peptides: Methods and Protocols*, Hansen, P. R., Ed. Springer New York: New York, NY, 2017; pp 427-435.

186.Bradshaw, J. P., Cationic Antimicrobial Peptides. *BioDrugs* **2003,** *17* (4), 233-240.

187.Maria-Neto, S.; de Almeida, K. C.; Macedo, M. L. R.; Franco, O. L., Understanding bacterial resistance to antimicrobial peptides: From the surface to deep inside. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2015,** *1848* (11, Part B), 3078-3088.

188.Amini, A.; Raheem, S.; Steiner, A.; Deeba, F.; Ahmad, Z., Insect venom peptides as potent inhibitors of Escherichia coli ATP synthase. *Int. J. Biol. Macromol.* **2020,** *150*, 23-30.

189.Wiegand, I.; Hilpert, K.; Hancock, R. E., Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature Protocols* **2008,** *3* (2), 163-75.

190.Bogetti, A. T.; Piston, H. E.; Leung, J. M. G.; Cabalteja, C. C.; Yang, D. T.; DeGrave, A. J.; Debiec, K. T.; Cerutti, D. S.; Case, D. A.; Horne, W. S.; Chong, L. T., A twist in the road less traveled: The AMBER ff15ipq-m force field for protein mimetics. *The Journal of Chemical Physics* **2020,** *153* (6), 064101.

191.Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013,** *30* (1), 108-60.

192.Lamberti, V. E.; Fosdick, L. D.; Jessup, E. R.; Schauble, C. J. C., A Hands-On Introduction to Molecular Dynamics. *Journal of Chemical Education* **2002,** *79* (5), 601.

193.Durrant, J. D.; McCammon, J. A., Molecular dynamics simulations and drug discovery. *BMC Biol.* **2011,** *9*, 71.

194.Guvench, O.; MacKerell, A. D., Comparison of Protein Force Fields for Molecular Dynamics Simulations. In *Molecular Modeling of Proteins*, Kukol, A., Ed. Humana Press: Totowa, NJ, 2008; pp 63-88.

195.Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P., A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society* **1984,** *106* (3), 765-784.

196.Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995,** *117* (19), 5179-5197.

197.Daura, X., Molecular Dynamics Simulation of Peptide Folding. *Theoretical Chemistry Accounts* **2006,** *116* (1), 297-306.

198.Ho, B. K.; Dill, K. A., Folding very short peptides using molecular dynamics. *PLoS Comp. Biol.* **2006,** *2* (4), e27.

199.Debiec, K. T.; Gronenborn, A. M.; Chong, L. T., Evaluating the Strength of Salt Bridges: A Comparison of Current Biomolecular Force Fields. *The Journal of Physical Chemistry B* **2014,** *118* (24), 6561-6569.

200.Ulmschneider, J. P.; Ulmschneider, M. B., Molecular Dynamics Simulations Are Redefining Our View of Peptides Interacting with Biological Membranes. *Accounts of Chemical Research* **2018,** *51* (5), 1106-1116.

201.Manandhar, A.; Kang, M.; Chakraborty, K.; Tang, P. K.; Loverde, S. M., Molecular simulations of peptide amphiphiles. *Organic & Biomolecular Chemistry* **2017,** *15* (38), 7993-8005.

202.Daura, X.; van Gunsteren, W. F.; Rigo, D.; Jaun, B.; Seebach, D., Studying the Stability of a Helical β-Heptapeptide by Molecular Dynamics Simulations. *Chemistry – A European Journal* **1997,** *3* (9), 1410-1417.

203.Slough, D. P.; Yu, H.; McHugh, S. M.; Lin, Y.-S., Toward accurately modeling N-methylated cyclic peptides. *Physical Chemistry Chemical Physics* **2017,** *19* (7), 5377-5388.

204.Paissoni, C.; Nardelli, F.; Zanella, S.; Curnis, F.; Belvisi, L.; Musco, G.; Ghitti, M., A critical assessment of force field accuracy against NMR data for cyclic peptides containing β-amino acids. *Physical Chemistry Chemical Physics* **2018,** *20* (23), 15807-15816.

205.Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H., Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. *Journal of the American Chemical Society* **2007,** *129* (5), 1179-1189.

206.Bernacki, J. P.; Murphy, R. M., Length-Dependent Aggregation of Uninterrupted Polyalanine Peptides. *Biochemistry* **2011,** *50* (43), 9200-9211.

207.Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, I. I. I. W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007,** *35* (suppl_2), W375-W383.