

Structural Characterisation and Engineering of the Single Disulfide-Directed β -Hairpin (SDH) Fold

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

By

Balasubramanyam Chittoor

Bachelor of Science (Botany, Zoology and Chemistry) Sri Venkateswara University, India
Master of Science (Biotechnology) Sri Venkateswara University, India



Monash University

Medicinal Chemistry
Monash Institute of Pharmaceutical Sciences
Melbourne, Victoria, Australia

2018

Copyright notice

© The author (2018). Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Monash University

Declaration for thesis based or partially based on conjointly published or unpublished work

General Declaration

In accordance with Monash University Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes one original paper published in peer reviewed journals and three traditional chapters. The core theme of the thesis is: Structural characterization and engineering the single disulfide-directed β -hairpin (SDH) fold. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Department of Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences under the supervision of Professor Raymond S. Norton.

[The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.]

In the case of Chapters 3 my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status (published, in press, accepted or returned for revision)	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
3	The single disulfide-directed β -hairpin fold. dynamics, stability and engineering	Accepted	60%. Concept and collecting data and writing first draft	1. Krishnarjuna Bankala, Collecting data, input into manuscript 10%	No
				2. Rodrigo A. V. Morales, intellectual input, manuscript preparation, input into manuscript 5%	No
				3. Christopher A. MacRaild, intellectual input, data analysis, manuscript preparation, input into manuscript 5%	No
				4. Maiada Sadek, Collecting data, input into manuscript 2%	Yes
				5. Eleanor W. W. Leung, Collecting data, input into manuscript 2%	No
				6. Samuel D. Robinson, Collecting data, input into manuscript 2%	No
				7. Michael W. Pennington, Collecting data, input into manuscript 2 %	No
				8. Raymond S. Norton, intellectual input, input into manuscript 12%	No


I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: (C. Balasubramanyam)

Date: 05/07/2018

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible

author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: ()

Date: 05/07/2018

Publications

Balasubramanyam Chittoor, Bankala Krishnarjuna, Rodrigo A. V. Morales, Christopher A. MacRaid, Maiada Sadek, Eleanor W. W. Leung, Samuel D. Robinson, Michael W. Pennington and Raymond S. Norton (2017) The single disulfide-directed β -hairpin fold. Dynamics, stability and engineering. *Biochemistry* **56**: 2455-2466.

Samuel D. Robinson, Sandeep Chhabra, Alessia Belgi, Balasubramanyam Chittoor, Helena Safavi-Hemami, Andrea J. Robinson, Anthony T. Papenfuss, Anthony W. Purcell, Raymond S. Norton. (2016) A naturally occurring peptide with an elementary single disulfide-directed β -hairpin fold. *Structure* **24**: 293–299.

Drane, S. B., Robinson, S. D., MacRaid, C. A., Chhabra, S., Chittoor, B., Morales, R. A.V., Leung, E. W. W., Belgi, A., Espino, S. S., Olivera, B. M., Robinson, A. J., Chalmers, D. K., and Norton, R. S. (2017) Structure and activity of contryphan-Vc2: Importance of the D-amino acid residue, *Toxicon* **129**, 113-122.

Acknowledgments

I would like to express my deepest gratitude to the following individuals for their valuable insight and contributions to this thesis, without whom none of this would have been possible.

First and foremost, I would like to express my sincere gratitude to Professor Raymond S. Norton for his willingness to take me into his lab and to guide me under his supervision. I would like to thank him for the leadership in directing the projects and providing enough resources and the care for completion of the project. He not only provided me with a lot of intellectual support and constructive criticism throughout my Ph.D., but was also a good role model as a highly enthusiastic, professional and successful scientist. His guidance and undivided support have allowed me to build myself into a more independent researcher and I am very grateful to him for giving me numerous opportunities during my PhD years to expand my skill sets and to improve myself.

I would like to especially thank Dr Krishnarjuna Bankala, for his help in recording NMR experiments in my project and for his support in my life. Without your help I could not have got this far. I would like to thank Dr Christopher A. MacRaild, Dr Rodrigo A. V. Morales, Dr Eleanor Leung and Dr Samuel Robinson for their recommendations and suggestions, which have been invaluable for my research progress.

I also wish to extend my gratitude to Professor Martin J. Scanlon, Professor Philip Thompson, and Dr David Chalmers, whose support has been invaluable. I would like to thank my colleagues Stephen Drane, Jeff Seow, Michela Mitchell, Maiada Sadek, Sreedam Das, Mansura Akther and Cael Debono for their support and help.

I would like to thank all the members of the Scanlon group and Philip Thompson's group for allowing me use the equipment and lab resources. I would like to thank Karen Drakatos, Jason Dang, Nicky Penny, Paul Dover, Adrian Whear and Jayn Lindholm for their timely responses and help.

Last but not least I would like to express a sense of gratitude and love to my friends and my beloved family members for their support, strength, help and for everything.

Table of Contents

Chapter 1. Introduction

1.1 Introduction

- 1.1.1 CS α / β Scorpion toxin fold
- 1.1.2 Inhibitor cystine knot scaffold
- 1.1.3 Cyclotide peptide scaffold

1.2 Animal venoms: A rich source of peptide scaffolds

- 1.2.1 Cone snails
- 1.2.2 Conopeptides: classification and their molecular organization

1.3 Diversity of conotoxins in the *Conus victoriae*

1.4 Contryphan family

1.5 Contryphan-Vc1

1.6 Scope of the thesis

Chapter 2. Materials and Methods

2.1 Recombinant expression of contryphan-Vc1 peptides

- 2.1.1 Preparation of the expression plasmid for thioredoxin-fused contryphan-Vc1 by cloning
- 2.1.2 Expression checking for thioredoxin-fused contryphan-Vc1

2.2 Purification of thioredoxin-fused contryphan-Vc1 by Nickel affinity chromatography

2.3 Reversed-phase HPLC (RP-HPLC) and LC-MS

2.4 Synthesis and purification of contryphan-Vc1 and its analogues

- 2.4.1 Solid phase peptide synthesis
- 2.4.2 Peptide folding by air oxidation and purification by reversed-phase HPLC

2.5 Nuclear magnetic resonance (NMR) spectroscopy

- 2.5.1 One-dimensional ^1H NMR experiments
- 2.5.2 Two-dimensional ^1H NMR experiments
- 2.5.3 ^1H - ^1H COSY
- 2.5.4 ^1H - ^1H TOCSY
- 2.5.5 ^1H - ^1H NOESY

2.6 Sequence-specific resonance assignments

2.7 Structure calculation

2.8 NMR relaxation measurements

2.9 Proteolysis assays

2.10 Surface plasmon resonance

2.11 Molecular dynamics simulations

Chapter 3. The Single Disulfide-Directed β -Hairpin Fold. Dynamics, Stability and Engineering

Chapter 4. Constraining the Single Disulfide-Directed β -hairpin (SDH) Fold with an Additional Disulfide bond

4.1 Introduction

4.2 Materials and methods

- 4.2.1 MD simulations
- 4.2.2 Peptide synthesis
- 4.2.3 NMR spectroscopy
- 4.2.4 Structure Calculation
- 4.2.5 Proteolysis Assays

4.3 Results

- 4.3.1 Analysis of the MD simulation trajectories
- 4.3.2 Synthesis and purification of contryphan-Vc1₁₋₂₂[Q1C, Y9C]
- 4.3.3 Temperature-dependent conformational averaging and pH stability
- 4.3.4 Sequence-Specific Resonance Assignments for contryphan-Vc1₁₋₂₂[Q1C, Y9C]
- 4.3.5 Solution Structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C]
- 4.3.6 Proteolytic stability of contryphan-Vc1

4.4 Discussion

Chapter 5. Membrane Binding Properties of Contryphan-Vc1

5.1 Introduction

5.2 Materials and methods

- 5.2.1 Peptide synthesis
- 5.2.2 Sample preparation and NMR spectroscopy
- 5.2.3 Surface Plasmon resonance

5.3 Results

- 5.3.1 Interaction of full-length contryphan-Vc1 with DPC micelles
- 5.3.2 Interaction with DPC micelles is mediated by the structured part of contryphan-Vc1
- 5.3.3 Sequence-specific resonance assignments for contryphan-Vc1₁₋₂₂[Z1Q] in the Presence of the DPC micelles
- 5.3.4 SDH fold of contryphan-Vc1 is maintained in the presence of DPC micelles
- 5.3.5 Contryphan-Vc1₁₋₂₂[Z1Q] binds weakly to POPC bilayers

5.4 Discussion

Chapter 6. Understanding the Roles of Key Residues in Maintenance of the SDH Fold

6.1 Introduction

6.2 Materials and methods

- 6.2.1 Peptide synthesis
- 6.2.2 Sample preparation and NMR spectroscopy

6.3 Results

- 6.3.1 Identification of contryphan-Vc1 homologous sequences
- 6.3.2 Disulfide bond is critical for the folding and maintenance of SDH fold
- 6.3.3 Full-length oyster-Vc1 is unfolded
- 6.3.4 Analogues contryphan-cg1₁₋₂₂(M9Y, P18I) and contryphan-cg1₁₋₂₂(R4Q, M9Y, R15I, P18I] restore the SDH fold

6.4 Discussion

Chapter 7. Conclusions and Future Directions

Chapter 8. References

Appendix I. A naturally occurring peptide with an elementary single disulfide- directed β -hairpin fold.

Appendix II. Structure and activity of contryphan-Vc2: Importance of the D-amino acid residue

Abbreviations

ACN	Acetonitrile
sACM	S-acetamido methyl
CTC	Chloro-trityl
COSY	Correlation spectroscopy
CCK	Cyclic cystine knot
DMSO	Dimethyl sulfoxide
DPC	Dodecylphosphocholine
DMF	Dimethylformamide
DIPEA	<i>N,N</i> -diisopropylethylamine
DODT	3, 6-dioxa-1, 8-octanedithiol
DMB	1,3-dimethoxy benzene
DCM	Dichloromethane
EDC	1-ethyl-3(3-diethylaminopropyl)-carbodiimide hydrochloride
FPLC	Fast protein liquid chromatography
Fmoc	<i>N</i> -(9-fluorenyl) methoxycarbonyl
HCTU	O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HSQC	Heteronuclear single-quantum coherence
ICK	Inhibitory cystine knot
iNOS	inducible nitric oxide synthase
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITT	Inverse temperature transition
kDa	Kilodalton
LB	Luria broth
LC-MS	Liquid chromatography-mass spectrometry
MD	Molecular dynamics
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
ORF	Open reading frame
PDB	Protein data bank

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
PPI	Protein-protein interaction
<i>R_g</i>	Radius of gyration
RMSD	Root-mean-square deviation
RU	Response unit
RP-HPLC	Reversed phase high performance liquid chromatography
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SDH	Single disulfide-directed β -hairpin
SPR	Surface plasmon resonance
SPSB	SPRY domain-containing SOCS (suppressor of cytokine signalling) box
TOCSY	Total correlation spectroscopy
TCEP	Tris(2-carboxyethyl)phosphine
TIPS	Triisopropylsilane
TFA	Tri-fluoro acetic acid
Trx	Thioredoxin
VMD	Visual molecular dynamics

Abstract

Peptide grafting, i.e. reproduction of the structural features and functional properties of active epitopes on an appropriate scaffold or a template with high conformational and proteolytic stability, provides one potential solution for overcoming poor stability and bioavailability of the peptide epitopes. Contryphan-Vc1, a 31-residue peptide from the venom of the cone snail *Conus victoriae*, has been shown to have a unique fold, designated the single disulfide-directed β -hairpin (SDH). The most conspicuous structural feature of contryphan-Vc1 is its two-stranded anti-parallel β -sheet stabilised by a single disulfide bond. The β -hairpin core of contryphan-Vc1 displays an ordered structure and remarkable thermal stability, and is very similar structurally to the inhibitory cystine knot (ICK) fold, which is stabilised by a minimum of three disulfide bonds. These properties suggest that contryphan-Vc1 may have potential as a useful scaffold for grafting. This thesis focuses on establishing and developing the SDH fold as a novel peptide scaffold.

The first part of my thesis examines different aspects of the stability of the contryphan-Vc1, including thermal, chemical, redox and proteolytic stability. Contryphan-Vc1 exhibited remarkable thermal stability, with the SDH fold of the peptide being maintained even at 95°C. Contryphan-Vc1 also exhibited remarkable chemical stability in the presence of urea, showing <30% unfolding even in 7 M urea. Contryphan-Vc1 is also quite stable over a broad pH range of pH 2 to 8. It is highly redox stable and required more than 2 days at room temperature for complete reduction in the presence of an excess of a strong reducing agent such as tris(2-carboxyethyl)phosphine (TCEP). The β -hairpin core structure was found to be resistant to cleavage by trypsin and chymotrypsin although it was susceptible to cleavage by pepsin. In addition, using backbone ^{15}N relaxation data we have shown that residues Trp2–Ile18 adopt a well-defined conformation, while residues beyond Thr19 are more disordered. The flexible C-terminus of contryphan-Vc1 could be truncated without loss of ordered structure.

The loops in the core structure of contryphan-Vc1 were replaced with a five-residue DINNN motif from inducible nitric oxide synthase (iNOS) that mediates the interaction of iNOS with the SPRY domain-containing SOCS (suppressor of cytokine signalling) box (SPSB) proteins, which facilitate the proteolytic degradation of iNOS. Even though both Vc1-DINNN analogues, sCon-Vc1₁₋₂₂[Z1Q, DINNN₄₋₈] and sCon-Vc1₁₋₂₂[Z1Q, DINNN₁₂₋₁₆], bound to SPSB2 with affinities of 25 ± 8 and 5.7 ± 3 nM respectively, ^1H NMR spectra of both peptides showed poor peak dispersion, indicating that substituting the loop 1 or loop 2 residues with the five-residue DINNN motif disrupted the native SDH fold. The analogue sCon-Vc1₁₋₂₂[NNN₁₂₋

¹⁴], which had a shorter three-residue NNN insert in place of the VLG residues in loop2, bound to human SPSB2 with an affinity of 1.3 μ M. Moreover, ¹H NMR spectra of this analogue showed good peak dispersion and native-like chemical shifts, indicating that the β -hairpin fold was maintained.

An attempt was made to further enhance the limited proteolytic stability and to overcome the limitation of incorporating only short peptide motifs by constraining the SDH fold of contryphan-Vc1 with an additional disulfide bond. An additional disulfide bond is introduced by changing Gln1 and Tyr9 of contryphan-Vc1 to Cys utilising orthogonal Cys protection. The introduction of an additional disulfide bond into contryphan-Vc1 did not disrupt the SDH fold. Contryphan-Vc1₁₋₂₂[Q1C, Y9C] was shown to be resistant to trypsin digestion and to be cleaved at Leu20 by chymotrypsin, similar to that of Con-Vc1₁₋₂₂[Z1Q]. However, contryphan-Vc1₁₋₂₂[Q1C, Y9C] was completely resistant to pepsin digestion, whereas native contryphan-Vc1 was fully degraded by pepsin.

The membrane binding properties of contryphan-Vc1 were also explored using dodecylphosphocholine micelles as a model membrane. Both full-length and truncated contryphan-Vc1 interact with DPC micelles, as seen in the significant changes in the ¹H NMR spectra of both peptides in presence of DPC. In contrast surface plasmon resonance (SPR) analysis showed that, Con-Vc1₁₋₂₂[Z1Q] bound very weakly to a flat bilayer formed by the zwitterionic POPC lipid and surprisingly full-length contryphan-Vc1 did not show any binding to POPC lipid.

Homologous sequences to contryphan-Vc1 were identified and their sequences were aligned with contryphan-Vc1 to identify conserved residues. As expected, the Cys residues are highly conserved in all sequences, consistent with our finding that the disulfide bond was critical maintenance of the SDH fold of contryphan-Vc1. We have undertaken structural characterisation of contryphan-Vc1 homologous peptide contryphan-cg1 which is 46% identical to that of contryphan-Vc1 and was expected to have similar structure. Surprisingly, full-length contryphan-cg1 peptide is found out to be unstructured. We have also shown that the SDH fold can be restored by introducing conserved amino acid residues as observed in the native contryphan-Vc1.

In summary, this thesis provides an understanding of the SDH fold of contryphan-Vc1 and explores its potential as a peptide scaffold. Further structural and functional characterisation of the newly identified members of this family should allow us to understand the functional significance of the SDH fold of contryphan-Vc1 and its role in the venom of *Conus victoriae* as well as other species in which homologues are found.

Chapter 1. Introduction

1.1 Introduction

Protein–protein interactions (PPIs) are essential for almost all intracellular and extracellular biological processes, including cell proliferation, cell division, signal transduction, transcription, translation, and programmed cell death. As PPIs maintain homeostasis and also play a critical role in various pathological conditions, PPI-modifying drugs have massive therapeutic potential.¹ Traditionally, small molecules are preferred drug candidates to modulate PPIs owing to their size, bioavailability and membrane permeability.² Targeting PPIs with small molecules has produced a mix of both exciting success stories and frustrating challenges that has been observed in several studies where the binding surfaces of the proteins are large and with less defined binding pockets.^{3–5} By contrast, peptides and peptidomimetics, even though having challenges to be developed as drugs, are good candidates to target PPIs because of their specificity and ability to mimic a protein surface to effectively compete for binding. Interestingly, it has been estimated that 15–40% of all PPIs are mediated by short linear peptide sequences.¹ Peptides can be rationally designed based on the natural sequences that mediate PPIs in the proteins, and therefore can mask a critical part of the binding surface; furthermore, peptidomimetics can be chemically modified to stabilize the bioactive conformation, mimicking the three-dimensional structure of the binding surface of the protein.¹

Most of the linear peptides occurring in nature are not directly suitable for use as therapeutics because of their intrinsic weaknesses, including poor chemical and physical stability, and a short circulating plasma half-life.⁶ These aspects must be addressed in developing peptides to use as therapeutics.⁷ Peptide grafting, i.e. reproduction of the structural features and therefore the functional properties of the active epitopes on an appropriate scaffold or a template with high conformational and proteolytic stability provides one potential solution for overcoming poor stability and bioavailability of the peptide epitopes.⁸

The use of naturally-occurring proteins as scaffolds is an attractive choice, as nature has evolved a spectra of well-defined and intrinsically stable templates. By using natural constraints such as covalent interaction in the form of disulfide bridges, and non-covalent interactions such as electrostatic interaction and salt bridges, nature has found effective ways of stabilizing short polypeptide chains as peptide scaffolds. Examples of natural protein scaffolds include three-finger toxin from snakes, charybdotoxin from scorpions, and various inhibitory cystine knot (ICK) motif-containing peptides from different sources such as cone snails, plants and spiders.

Many of these natural scaffolds present stable secondary structure elements such as α -helix and/or β -sheets, held together in a compact structure that is reinforced by disulfide bridges,

while leaving most of the functional side chains free to engage their biological target.^{9, 10} These natural scaffolds typically also exhibit a high level of tolerance to sequence mutation and functional adaptability. Most of the scaffolds are comparatively smaller in size, and are compact with well-defined secondary structures. Peptide scaffolds that are most utilized for grafting novel biological activities are described below.

1.1.1 CS α / β Scorpion toxin fold

Charybdotoxin, a scorpion toxin with CS α / β fold,¹¹ is the first peptide scaffold utilized to design a grafted molecule, in this case to inhibit the CD4-gp120 interaction associated with HIV infection.¹² This fold consists of a triple stranded anti-parallel β -sheet on one face and a short α -helix on the opposite face. The core structure is stabilized by three disulfide bonds. The core cysteine framework of CS α / β scorpion toxins is the C–C–C–C–C–C pattern, but some scorpion toxins like maurotoxin and chlorotoxin have additional disulfide bonds¹³(**Figure 1**).

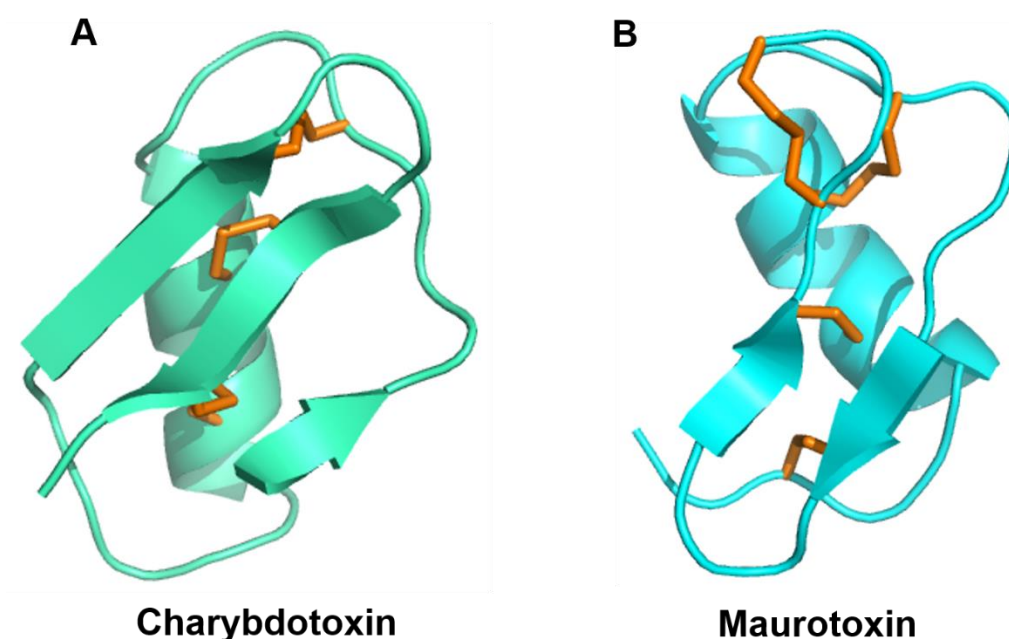


Figure 1. A and B Cartoon representation of charybdotoxin (PDB code 2CRD) and maurotoxin (PDB code 1TXM) as examples of peptides with CS α / β fold.

1.1.2 Inhibitor cystine knot scaffold

The smallest of all β -sheet-containing scaffolds known to date is the inhibitory cystine knot (ICK) motif containing peptides.^{14, 15} Peptides containing ICK motifs are found in myriad organisms, such as plants, insects, and mammals, and carry out diverse functions including protease inhibition, ion channel blockade, and antimicrobial activity. The ICK motif is a prominent structural motif among the different folds identified in disulfide rich protein domains.

ω -conotoxin MVIIA, which is used in chronic pain therapy under trade name PRIALT (Ziconotide), is an extensively-studied ICK peptide.¹⁶ The *in vivo* stability exhibited by ω -conotoxin MVIIA highlights the exceptional stability conferred upon this peptide by its cystine knot structure. The consensus sequence among the initially identified ICK peptides was CX₃₋₇ CX₄₋₆ CX₀₋₅ CX₁₋₄ CX₄₋₁₀ C and the presence of the ICK motif was predicted utilizing this half-cystine gaps in the sequence of the peptides.¹⁵ The number of ICK peptides identified has increased significantly and new members have expanded the lengths of the intercysteine loops. Schematic representation of the ICK fold and structures of ICK peptides ω -Conotoxin GVIA and agatoxin are shown in the figure 2.

These ICK molecules typically contain 25-35 residues and are characterized by a triple-stranded, anti-parallel β -sheet stabilized by a cystine knot. The cystine knot comprises a ring formed by two disulfide bridges and the interconnecting backbone, through which the third disulfide passes. The β -strands are reported as the minimal conserved secondary structural elements among the ICK peptides and the β -sheet is regarded as integral to this ICK motif. ICK peptides despite sharing a core structure exhibit significant sequence and functional diversity. Peptides with an ICK motif structure exhibit a wide variety of biological targets ranging from different ion channels to proteases. The high sequence diversity in the ICK peptides arises from the loop regions between the half-cystine residues, which are hypervariable and generate high functional diversity with the same underlying scaffold.^{15, 17, 18}

Owing to their small size, high thermal and proteolytic stability and tolerance to new sequences ICK peptides gradually became preferred molecular scaffolds as exemplified by utilization of several ICK peptides like EETI-II and agouti-related protein (AgRP) for protein engineering purposes. *Ecballium elaterium* trypsin inhibitor (EETI-II), an ICK peptide from the squash family of protease inhibitors, was the first ICK peptide to be modified to incorporate RGD-based integrin binding motifs in one of its surface-exposed loop.^{19, 20} AgRP, a neuropeptide that is involved in regulating metabolism and appetite, is another ICK peptide that is regularly used as a peptide scaffold.^{21, 22} EETI and AgRP analogues engineered for integrin binding were labelled with a variety of contrast agents and used in non-invasive tumour imaging. More recently agatoxin, an ICK peptide from spider venom, was also successfully engineered for tumour imaging purpose.^{23, 24}

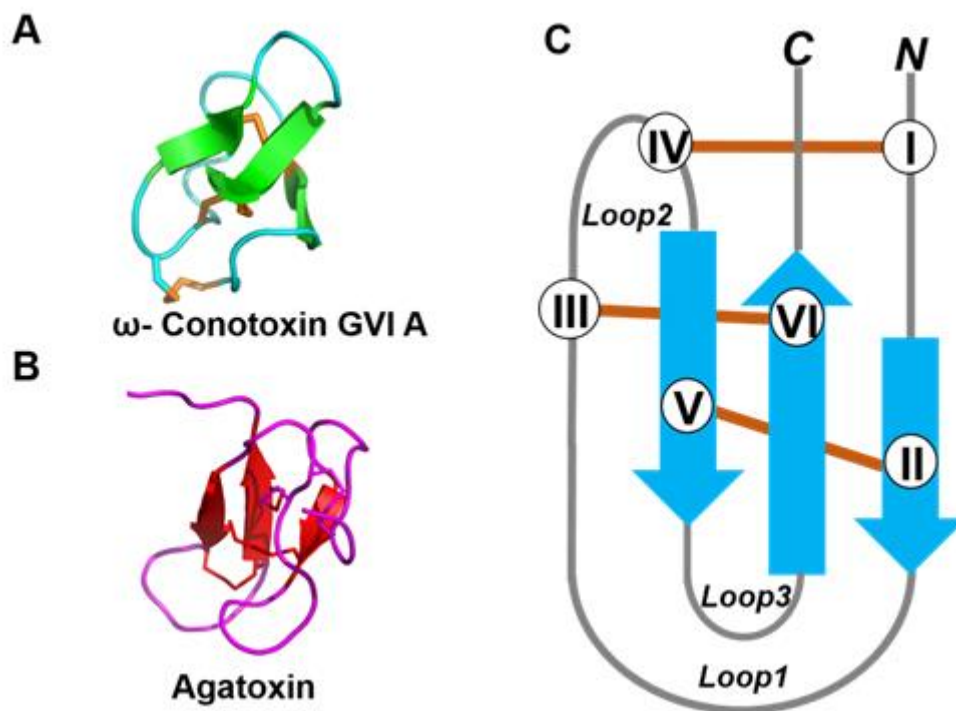


Figure 2. A and B Cartoon representation of ω -conotoxin GVI A (PDB code 1TTL) and agatoxin (PDB code 2N6N) as examples of peptides with ICK motif. C. Schematic representation of ICK motif.

1.1.3 Cyclotide peptide scaffold

Cyclotides are a large family of plant-derived peptides that contain approximately 28-37 amino acids and are characterized by a head-to-tail cyclic backbone and ICK knot arrangement of three disulfide bonds.¹⁵ This structural motif is referred to as the cyclic cystine knot (CCK) and provides exceptional stability to cyclotides. Kalata B1, (**Figure 3**) a cyclotide from the plant *Oldenlandia affinis* has been utilized as a scaffold to generate novel anti-cancer and anti-viral peptide-based lead molecules.^{25, 26} The trypsin inhibitor cyclotide MCoTI-II from *Momordica cochinchinensis* had been utilized to produce stable and potent angiogenic agents with promising potential in the design of drug leads for therapeutic angiogenesis.²⁷

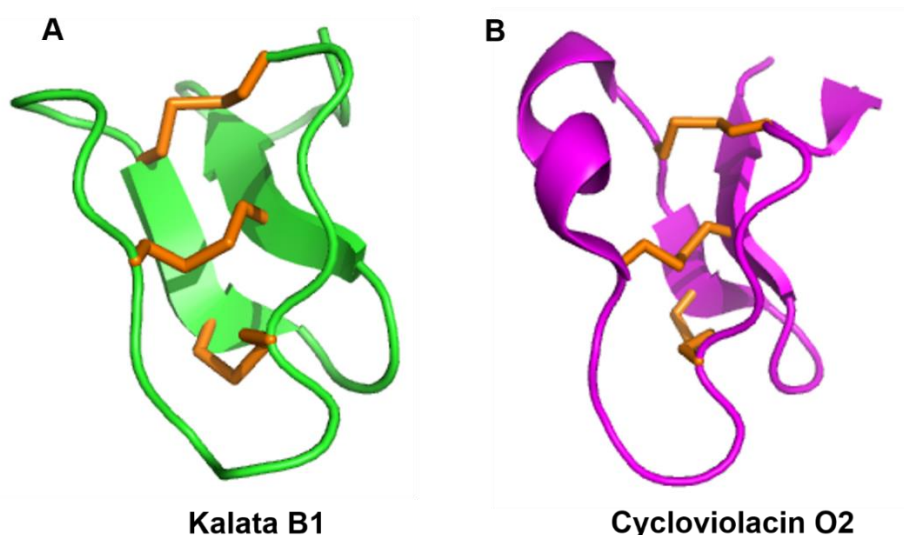


Figure 3. A and B Cartoon representation of kalata B1 (PDB code 1NB1) and cycloviolacin O2 (PDB code 2KCG) as examples of peptides with CCK motif.

1.2 Animal venoms: A rich source of peptide scaffolds

Animal venoms are a rich source of many peptide scaffolds. Venom is a complex cocktail of bioactive compounds which comprise a mixture of protein and peptides (commonly referred to as toxins), salts and organic components, such as amino acids and neurotransmitters.²⁸⁻³⁰ The proteinaceous components are usually the most abundant. The peptide toxins have attracted the attention of the researchers due to their extreme diversity of the biological targets and their mode of action. Apart from exhibiting remarkable structural and functional diversity some of the peptide toxins have high conformational and proteolytic stability making them attractive candidates for utilizing as peptide scaffolds.

1.2.1 Cone snails

Cone snails are a group of venomous predatory marine molluscs that belong to the genus *Conus* and family Conidae. Based on their prey, cone snails are divided into three groups: the vermivorous (worm-hunting) species, molluscivorous (snail hunting) species and the piscivorous (fish-hunting) cone snails. Short disulfide-rich peptides called conopeptides or conotoxins constitute the majority of the venom of the predatory cone snails. Conopeptides or conotoxins are produced in the venom duct attached to the venom bulb. The venom bulb contracts to push venom into a disposable, hypodermic needle like structure, which injects venom into the prey. These harpoons are evolutionary modified teeth that are stored in a radular sac.^{31, 32}

1.2.2 Conopeptides: classification and their molecular organization

Conopeptides target a wide variety of receptors, including both ligand-gated and voltage-gated ion channels, various G-protein coupled receptors and neuro-transmitter transporters, and have therefore been invaluable tools in understanding the structural and functional aspects of their molecular targets and are emerging as promising therapeutic molecules.³³ The ω -conotoxin Ziconotide (SNX-111; Prialt®), a synthetic equivalent of a conopeptide isolated from *Conus magus* is the first approved venom-derived drug that acts on nervous system and is being used in chronic pain therapy.¹⁶

Conopeptides are classified according to three schemes: Gene superfamily classification, based on the consensus conserved signal peptide sequence, cysteine framework classification, based on the occurrence of cysteine residues in a conopeptides primary sequence independent of their connectivity, and pharmacological families based on receptor specificities and the physiological activity (**Figure 4**). Pharmacological families are denoted by using Greek letters.^{31, 34, 35} Conopeptides are expressed as precursor proteins that are processed into mature peptides in the endoplasmic reticulum and Golgi apparatus. Each precursor of a conopeptide consists of a highly conserved and hydrophobic secretory signal sequence, a propeptide region and the mature peptide region. The signal sequence facilitates secretion of the peptide and the propeptide region contains binding sites for the enzymes that modify the peptide post-translationally and is abundant in charged residues. The mature peptide region is the functional peptide, and the sequence of the mature peptide is least conserved and is subjected to hyper mutations (**Figure 5**). Mature conopeptides exhibit a vast array of post translational modifications.^{34, 36}

tapped to date. The recent advent of proteomic/transcriptomic approaches combined with high-throughput screens has accelerated the identification of new conopeptides.³⁷⁻³⁹ Recently, one such study was undertaken for large-scale discovery of conopeptide sequences from the venom gland of the Australian cone snail *Conus victoriae* using cDNA library normalization, high-throughput sequencing and de novo transcriptome assembly yielded over 100 unique conopeptide sequences.⁴⁰

1.3 Diversity of conotoxins in the *Conus victoriae*

More than 100 unique conotoxin sequences from 20 gene superfamilies were discovered in a transcriptomic study of the venom gland of *Conus victoriae* using cDNA library normalization, high-throughput 454 sequencing, *de novo* transcriptome assembly and annotation with BLASTX and profile hidden Markov models; this was the highest diversity of conotoxins so far reported in a single study. Many of the sequences identified were new members of known conotoxin superfamilies, some help to redefine these superfamilies and others represent altogether new classes of conotoxins.⁴⁰

1.4 Contryphan family

The first contryphan peptide was identified in 1996 from the venom of the piscivorous cone snail *Conus radiatus* as a peptide containing D-tryptophan. Subsequently, many peptides belonging to the contryphan family were identified from various species of cone snails (Table 1).⁴¹⁻⁴⁵ Contryphans are considered the smallest peptide components in cone snail venom, with around 7-12 amino acid residues. Contryphan peptides are remarkable for their extensive post-translational modifications, including proline hydroxylation, C-terminus amidation as well as isomerization of tryptophan, and leucine, bromination of tryptophan, occurrence of gamma carboxy glutamate.⁴⁶ Contryphan peptides are cyclized via a single disulfide bridge.

All the members of the contryphan family of peptides characterized to date have the consensus sequence motif $X_1COwX_5PWC-NH_2$, where X_1 is either glycine or absent, O is 4-trans-hydroxyproline and X_5 is glutamate, aspartate or glutamine. Contryphan peptides exist as a mixture of two conformers in solution due to *cis-trans* isomerization around the *N*-terminal Cys-Pro peptide bond. Intra-cerebro-ventricular injection of contryphan peptides in mice causes stiff-tail syndrome at low doses and induces seizures as well as death at high doses.⁴¹⁻⁴⁵ However the exact molecular target(s) of these peptides are still unknown.

Table 1. List of different contryphan peptides with their sequence

Peptide	Sequence	References 40-44
Contryphan-R	GCOWEPWC-NH ₂	Jimenez et al. 1996
Contryphan-Sm	GCOWQPWC-NH ₂	Jacobsen et al. 1998
Contryphan-P	GCOWDPWC-NH ₂	Jacobsen et al. 1998
Contryphan-Tx	GCOWQPYC-NH ₂	Jimenez et al. 2001
Contryphan-Vn	GDCPWKPWC-NH ₂	Massilia et al. 2001
Contryphan-fib	GCOWMPWC-NH ₂	Rajesh et al. 2015

1.5 Contryphan-Vc1

Contryphan-Vc1 is identified in the transcriptome analysis performed on the venom gland of *Conus victoriae*. Contryphan-Vc1 is placed under the O2 /contryphan super family, based on significant similarities in its signal peptide sequence and the cysteine framework. However, beyond the signal peptide and single disulfide bond, contryphan-Vc1 shared no obvious sequence similarities with other previously characterised contryphans or indeed with any other conopeptides (Figure 6A).

Contryphan-Vc1 is the largest contryphan peptide, with 31 residues, and the first with an intercysteine loop length other than five residues. Contryphan-Vc1 shows a remarkably well ordered structure. The solution structure of contryphan-Vc1 shows that this peptide adopts a unique fold described as single disulfide-directed β -hairpin (SDH) which is entirely different from previous known contryphan fold.⁴⁷ The most conspicuous feature in the structure of contryphan-Vc1 is the presence of a double-stranded anti-parallel β -sheet that is stabilized by a single disulfide bond. Residues Tyr7 and Tyr9 form a mini hydrophobic core and the *N*-terminal pyro glutamic acid residue is in close association with the side chain of Tyr9. The four residues (QPGY) following Cys3 form a type II β -turn, leading into the first β -strand. The two strands are connected by a type IV β -turn encompassing four residues (PVLG). The second β -strand has Cys16 which links to Cys3 via the disulfide. This strand terminates following a final hydrogen-bond between the amide of Thr19 and the carbonyl of Gly6. The *C*-terminus of the peptide is less structured and highly flexible (Figure 6B).⁴⁷

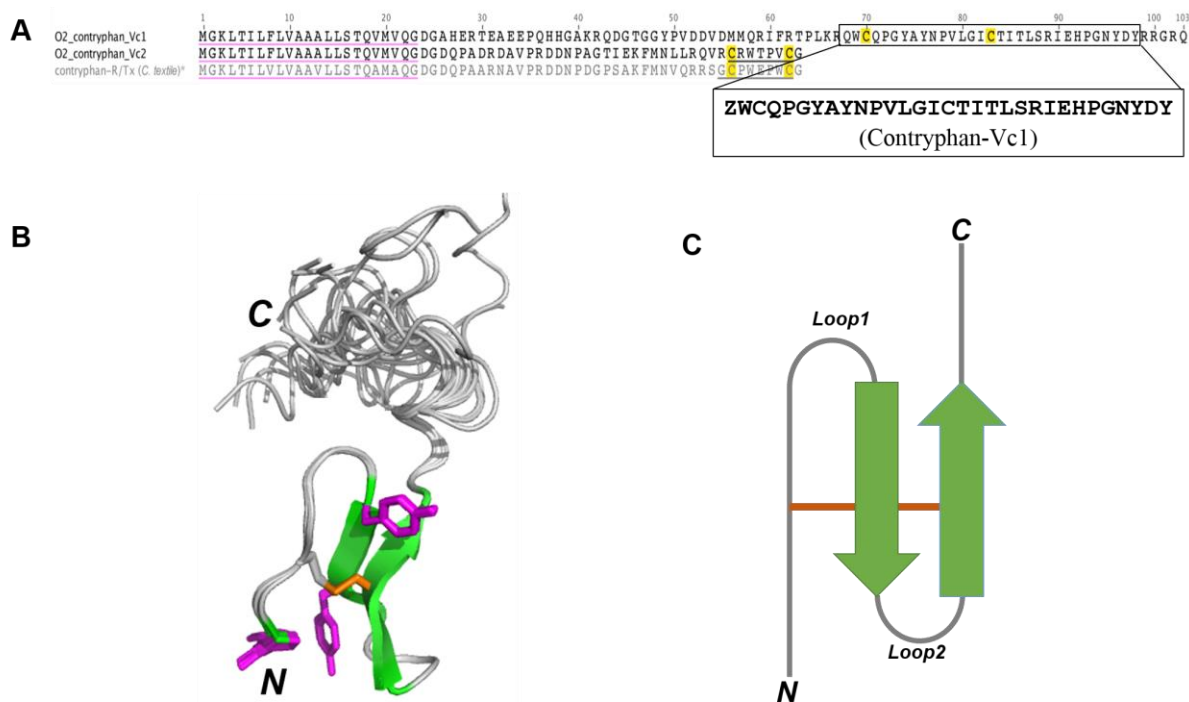


Figure 6. A. Amino acid sequence of mature peptide of contryphan-Vc1 B. Ensemble of the final 20 structures of synthetic contryphan-Vc1 superimposed over the backbone heavy atoms of residues 1-20. The *N*-terminal 20 amino acids of the peptide form a well-defined β -hairpin motif while the *C*-terminus is less well-defined but shows some helical propensity. C. Schematic representation of SDH fold.

1.6 Scope of the thesis

Contryphan-Vc1 exhibits a highly ordered structure despite the presence of only a single disulfide bond, and is the founding member of a hitherto unknown peptide fold, the single disulfide-directed β -hairpin (SDH) fold. Subsequently, many sequences with similarities to the primary structure of contryphan-Vc1 have been identified in both venomous and non-venomous molluscs, suggesting that contryphan-Vc1 may belong to a previously undescribed class of structurally-related mollusc peptides. Contryphan-Vc1 also exhibits remarkable thermal stability,⁴⁷ as seen in circular dichroism (up to 95°C) and NMR experiments (up to 70°C), where the SDH fold is maintained even at high temperatures.⁴⁷ The β -hairpin core of contryphan-Vc1 is very similar structurally to the ICK fold despite its reduced cysteine framework complexity (**Figure 7**). Its highly ordered structure, remarkable thermal stability, and similarities to the ICK fold suggested that contryphan-Vc1 may have potential as a useful scaffold for grafting. Therefore, a major aim of this study is to investigate the stability of the SDH fold and its potential as peptide scaffold and also understand the novel SDH fold by studying some of the contryphan-Vc1 homologous peptides.

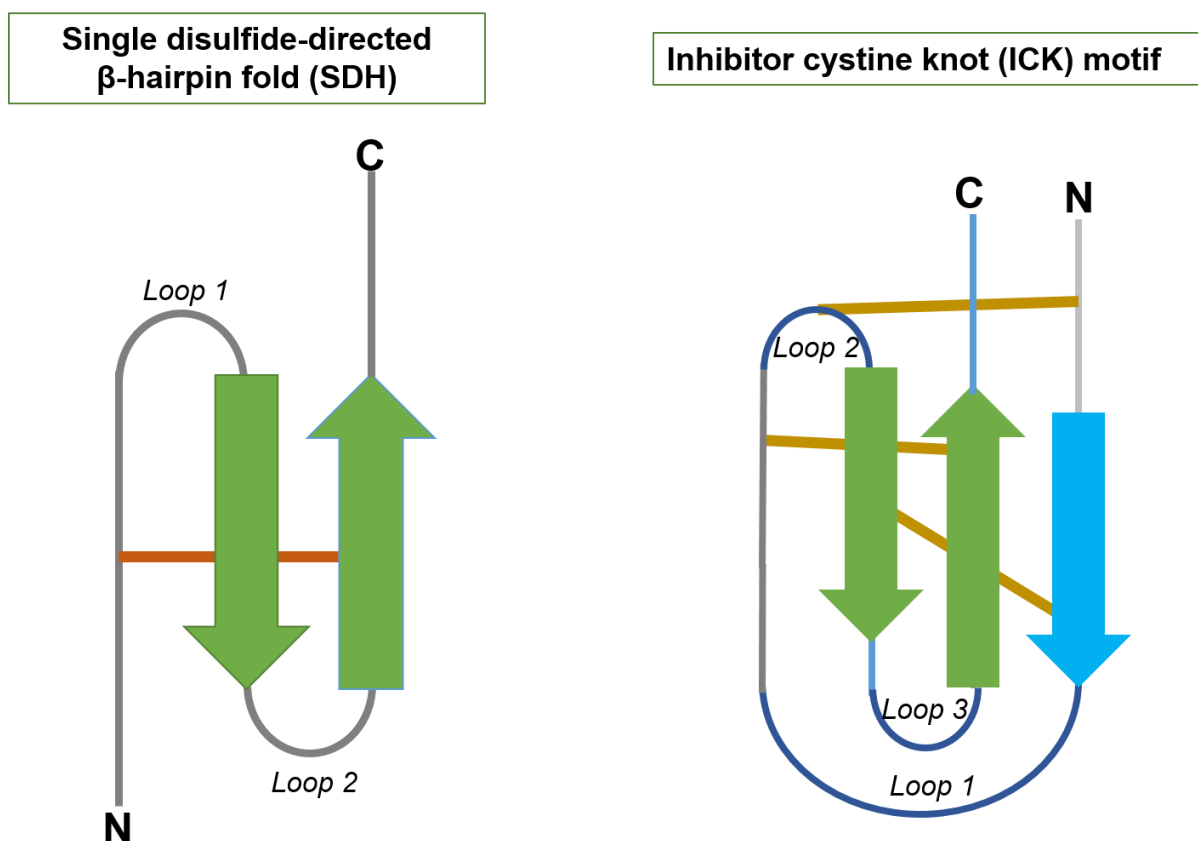


Figure 7. Comparison of schematic representations of SDH and ICK folds.

In **Chapter 2**, all general materials and methods used in the study are elaborated, which includes recombinant and synthetic production of the contryphan-Vc1 and its analogues, NMR data acquisition and structure calculation and refinement using CYANA and Xplor. Assays for establishing proteolytic stability, binding studies of Vc1-DINNN analogs utilizing SPR and molecular dynamics (MD) simulations studies are also described.

Chapter 3 describes studies of the stability of the SDH fold to chemical denaturation and proteolytic enzymes, as well as its resilience to truncation and the insertion of non-native peptide motifs in order to evaluate its potential as a scaffold. The backbone dynamics of contryphan-Vc1 were analysed in order to identify the ordered and flexible regions of the peptide, and showed that the flexible C-terminal region is highly flexible. We have investigated whether contryphan-Vc1 can incorporate a bioactive epitope by utilizing the DINNN motif of inducible nitric oxide synthase (iNOS) which is essential for recognition of iNOS by the SPRY domain-containing SOCS (suppressor of cytokine signalling) box (SPSB) proteins in circulating macrophages.⁴⁸⁻⁵⁰ The peptide analogues sCon-Vc1₁₋₂₂[Z1Q, DINNN₄₋₈], sCon-Vc1₁₋₂₂[Z1Q, DINNN₁₂₋₁₆] and sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄], were tested for their affinity for

human SPSB2 utilizing surface plasmon resonance (SPR) and analysed conformationally by NMR.

Chapter 4 outlines the design, synthesis and structural characterization of the two-disulfide bonded analogue of contryphan-Vc1, contryphan-Vc1₁₋₂₂[Q1C,Y9C). An additional disulfide bridge was added by replacing Gln1 and Tyr9 with cysteine residues in order to further enhance stability and flexibility of loops to receive longer epitopes.

Chapter 5 describes the lipid binding properties of contryphan-Vc1 as observed by NMR experiments in the presence of deoxyphosphocholine (DPC) micelles, and by SPR experiments utilizing an L1 chip in the presence of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid.

In **Chapter 6**, the design and structural characterization of analogues of contryphan-cg1-Vc1, a homologue of contryphan-Vc1 identified from pacific oyster *Crassostrea gigas*. NMR spectra recorded on the full-length contryphan-cg1 peptide showed that the peptide is not structured. A double mutant (contryphan-cg1₁₋₂₂[M9Y, P18I]) and a quadruple mutant contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] were designed and synthesized based on conserved residues in the primary structure of contryphan-Vc1 homologues and studied for the effect of those residues on the SDH fold. It is shown that the SDH fold can be restored by introducing conserved amino acid residues as observed in the native contryphan-Vc1.

Finally, all major findings from this study are summarised and potential future directions are discussed in Chapter 7.

Chapter 2

Materials and Methods

2.1 Recombinant expression of contryphan-Vc1 peptides

Chemical synthesis and recombinant production have been the most widely used methods for production of peptide toxins. Recombinant expression of peptide toxins from *E. coli* has become preferred common mode of production of peptide toxins owing to the advantages such as high yield and cost effectiveness.⁵¹⁻⁵³ However, the poor yield of correctly-folded disulfide-rich peptides in the reduced environment of the cytosol, the occurrence of peptide in inclusion bodies, and lack of post translational modifications in *E. coli* make it challenging to produce peptide toxins. Periplasmic expression of disulfide rich peptides in *E. coli* where the over expressing peptide is secreted into oxidizing environment of periplasm which contains foldases, disulfide oxidoreductase (**DsbA**) and disulfide isomerase (**DsbC**), that catalyse the formation and isomerisation of disulfide bonds is also an attractive option for production of disulfide rich peptides in *E. coli*. Nonetheless, the presence of a single disulfide bond and lack of complex post translational modifications in the sequence of the contryphan-Vc1 prompted us to undertake recombinant production of contryphan-Vc1 and its variant peptides.

2.1.1 Preparation of the expression plasmid for thioredoxin fused contryphan-Vc1 by cloning

DNA encoding the open reading frame (ORF) of the mature peptide sequence of contryphan-Vc1 was codon-optimized for *E. coli* expression, synthesized and cloned in pUC based vector (Genscript). The ORF of full-length contryphan-Vc1 was amplified using the primers F1 and R1 and the DNA fragment was subjected to restriction digestion using BamH1 and Xho1 enzymes. The DNA fragment was subsequently ligated to the BamH1 and Xho1 digested pET32a⁺ vector to generate a modified pET32a+ plasmid which expresses thioredoxin fused contryphan-Vc1 protein with an enterokinase cleavage site between thioredoxin and contryphan-Vc1 and a hexa histidine tag at C-terminus. Similarly a modified pET32a+ plasmid which expresses thioredoxin fused truncated contryphan-Vc1 is made utilizing primers F1 and R2 (**Figure 1**).

A Forward primer (F1): 5' CAGTCGACGATGACGATGACAAACAG 3'
Reverse primer (R1): 5' GGTGGATCCCTCGAGTTAGTAATCATAATTGCCCGG 3'
Reverse primer (R2): 5' CGGGATCCCTCGAGTCAACGGCTCAGCGTGATGGTAC 3'

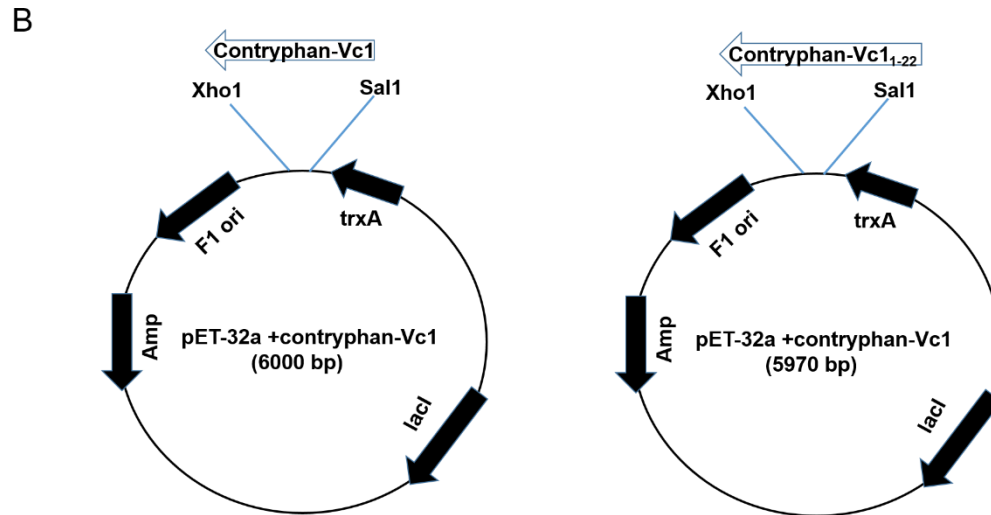


Figure 1. A. Primers used to amplify the DNA sequence of full-length contryphan-Vc1 and truncated contryphan-Vc1. Restriction enzyme cleavage sequence is coloured green and the stop codon is coloured red. **B.** Schematic representation of the modified pET-32a plasmids expressing thioredoxin fusion proteins of full-length and truncated contryphan-Vc1.

2.1.2 Expression checking for thioredoxin-fused contryphan-Vc1

The pET32a+ plasmids coding for thioredoxin fused to contryphan-Vc1 were transformed into *E. coli* BL21 cells and the *E. coli* colonies were selected on Luria-Bertani (LB) agar plates containing 100 µg/mL of ampicillin. In order to check the expression of the thioredoxin-fused contryphan-Vc1, a single colony was picked and inoculated into 5 mL of LB medium containing 100 µg/mL of ampicillin and allowed to grow at 30 °C till the OD₆₀₀ of the culture reached 0.6. The cells were then induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and allowed to grow for 4 h. Cells equivalent to 0.2 OD before and after IPTG induction were pelleted and resuspended in 50 µL of 1 x SDS gel loading dye and boiled at 90 °C for 10 min in order to lyse the cells. The cell lysate was loaded in equal amounts on a 12% SDS PAGE gel along with protein molecular weight marker (BioRad). Overexpression of thioredoxin fused contryphan-Vc1 proteins were confirmed initially by matching the molecular mass of the fusion proteins to the standards.

2.2 Purification of thioredoxin fused contryphan-Vc1 by Nickel affinity chromatography

For purification of the thioredoxin-fused contryphan-Vc1, the cultures were scaled up to 1 L, where the transformed *E. coli* BL21 cells were grown at 30 °C and induced for 4 h with 0.5 mM IPTG once the culture reached 0.6 OD. The cells were pelleted and were either frozen or

utilized immediately for the purification of thioredoxin-fused contryphan-Vc1 protein. The cells were resuspended in Novagen bug buster and incubated for 20 min in the presence of EDTA-free protease inhibitor cocktail from Roche, then clarified by centrifuging at 14,000 rpm for 30 min at 4 °C. The supernatant was loaded onto pre-equilibrated Ni-NTA Hitrap column (GE Healthcare, 5ml) at a flow rate of 0.25 ml/min. Unbound proteins were removed by extensive washing with buffer A (20 mM Tris-Cl, pH 8.0, 10 mM imidazole, 100 mM NaCl) followed by a wash with buffer A containing 0.5 M NaCl. The fusion protein was eluted with buffer A containing 250 mM imidazole and exchanged with enterokinase cleavage buffer (50 mM Tris-Cl, pH 8.0, 1mM CaCl₂ and 50 mM NaCl). The thioredoxin tag was removed by incubating the fusion protein in enterokinase at 23 °C for 16 h and passing the mixture back onto Ni-NTA Hitrap column (GE Healthcare, 5 mL). Unbound peptide was collected in buffer A.

2.3 Reversed-phase HPLC (RP-HPLC) and LC-MS

Reversed-phase HPLC (RP-HPLC) effectively separates molecules based on polarity and their interactions with a hydrophobic matrix. Peptides bind to the column under aqueous conditions and are eluted with increasing concentrations of organic solvent. All peptides described here were purified by reversed-phase high performance liquid chromatography (RP-HPLC) on a Phenomenex® Luna C18 column (100 Å, 5 µm, 100 x 10 mm) using a gradient of 5–95% B (A: 99.9% H₂O, 0.1% tri fluoro acetic acid (TFA); B: 80% acetonitrile (ACN), 19.9% H₂O, 0.1% TFA) over 30-60 min and the samples were lyophilized. The purity and molecular mass of the recombinant contryphan-Vc1 were confirmed with liquid chromatography-mass spectrometry (LC-MS) on a Shimadzu LCMS2020 instrument, incorporating a Phenomenex® Luna C8 column (100 Å, 3 µm, 100 x 2 mm) and using a linear gradient of 100% H₂O (0.05% TFA) for 4 min, followed by 0-60% ACN (0.05% TFA) in water over 10 min at a flow rate of 0.2 mL/min.

2.4 Synthesis and purification of contryphan-Vc1 and its analogues

2.4.1 Solid phase peptide synthesis (SPPS)

SPPS is an alternative to obtain natural peptides that are difficult to express in bacteria. SPPS can incorporate amino acids that do not occur naturally and modify the peptide/protein backbone. In this method, amino acids attach to polymer beads suspended in a solution to build peptides, which remain attached to beads until cleaved by a reagent such as TFA. This immobilizes the peptide during the synthesis so it can be captured during filtration, while liquid-phase reagents and by-products are simply flushed away. The benefits of solid-phase

synthesis are that it greatly speeds production of peptides since it is a relatively simple process; it is easier to scale, and it is more suitable for making peptide analogues quickly by skipping the cloning optimization of expression.

Contryphan-Vc1 and its truncated analogs were synthesized utilizing an automated Peptide Synthesizer *via* Fmoc chemistry. All the peptides were synthesized at 0.1 mmol scale. Full-length contryphan-Vc1 was synthesized using *N*-(9-fluorenyl) methoxycarbonyl (Fmoc)-Tyr (tBu)-Wang resin on a Prelude automated peptide synthesizer (Protein Technologies, Tucson, Arizona). All couplings were mediated with diisopropylcarbodiimide with 6-Cl-HOBT for 3 h. The peptide was cleaved from the resin support and simultaneously deprotected using a cocktail of trifluoroacetic acid, anisole, triisopropylsilane, thioanisole, water and 1,2-ethanedithiol (9:1:2:1:1:1 v/v) for 2 h at room temperature. The crude peptide was precipitated with ice-cold diethyl ether and washed three times with the same, then subsequently dissolved in 50% aqueous acetic acid.

Con-Vc1₁₋₂₂[Z1Q], con-Vc1₁₋₂₂[Q1C, Y9C], full-length contryphan-cg1, Vc1-DINNN, and the Vc1-NNN analogues (**Table 1**) were synthesized on a PTI Instruments PS3 peptide synthesizer, using Rink amide AM resin. Rink Amide AM resin (0.47 mmol/g loading) was weighed into a reaction vessel and swelled in dimethylformamide (DMF) (10 mL/g resin) for 30 min. 0.3 mmol (3 eq, 124 mg) of Fmoc-protected amino acids and O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) were weighed into vials (with one amino acid/HCTU powder mixtures per vial) and sealed using a PTFE/Silicone septa and cap. All vials were then queued (from C-terminus residue to N-terminus residue) on the carousel on PS3. To start the synthesis, each vial was assigned with Program 7, where the Fmoc protecting group on the resin was first removed by 20% (v/v) piperidine in DMF prior coupling to the amino acid residue for 50 min per coupling cycle under the activation of 0.3 mmol (3 eq) of HCTU and 7% (v/v) *N,N*-diisopropylethylamine (DIPEA) in DMF. Cleavage from the resin was performed for 3 h with a mixture of 3, 6-dioxo-1, 8-octanedithiol, triisopropylsilane, 1,3-dimethoxy benzene and trifluoroacetic acid (DODT:TIPS:DMB:TFA at ratios 2.5: 2.5: 5: 92.5 by volume) The cleavage mixture was purged with nitrogen and the crude peptide was precipitated with ice-cold diethyl ether and washed three times with the same and subsequently dissolved in 50% acetonitrile and freeze dried.

The double mutant of contryphan-cg1₁₋₂₂, contryphan-cg1₁₋₂₂[M9Y, P18I] and quadruple mutant contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] were synthesized on PTI Instruments PS3 peptide synthesizer, using chloro-trityl (CTC) resin. CTC resin (0.7 mmol/g loading) was

weighed into a reaction vessel, swelled in dichloromethane (DCM) for 30 min and coupled manually with the last amino acid in the sequence by dissolving two equivalents of the amino acid in a mixture of DCM and DMF (2.5 ml+0.5 ml) and 87.5 μ l of DIPEA for over 1 h. The uncoupled resin was then capped with methanol (0.8 ml/gram resin) for 30 min and the resin washed with an excess of DCM and then an excess of DMF. Synthesis and cleavage of the peptide from the resin were done as mentioned earlier.

Table 1. Peptide used in this study

Peptide	Sequence
sCon-Vc1 ₁	^a ZWCQPGYAYNPVLGICTITLSRIEHPGNYDY
rCon-Vc1[Z1Q]	QWCQPGYAYNPVLGICTITLSRIEHPGNYDY
rCon-Vc1 ₁₋₂₂ [Z1Q]	QWCQPGYAYNPVLGICTITLSR
sCon-Vc1 ₁₋₂₂ [NNN ₁₂₋₁₄]	ZWCQPGYAYNP <u>NNN</u> ICTITLSR
sCon-Vc1 ₁₋₂₂ [Z1Q,DINNN ₄₋₈]	QWCD <u>INNN</u> AYNPVLGICTITLSR
sCon-Vc1 ₁₋₂₂ [Z1Q,DINNN ₁₂₋₁₆]	QWCQPGYAYNP <u>DINNN</u> ICTITLSR
con-Vc1 ₁₋₂₂ [Q1C, Y9C]	CWCQPGYACNPVLGICTITLSR
Contryphan-cg1-22[R4Q, M9Y, R15I, P18I]	GWCQPGFTYNPVLNICVITMRS

^a Pyroglutamate modification is indicated as Z

2.4.2 Peptide folding by air oxidation and purification by reversed-phase HPLC

The linear peptide was allowed to fold by air oxidation in 0.1M ammonium bicarbonate by adjusting the concentration to 0.3 mg/mL and pH 8.0 with continuous stirring. For some peptides 5-10% DMSO was added to the folding reaction in order to enhance the oxidation process. The folded peptide was purified by reversed-phase HPLC on a Phenomenex® Luna C18 column (100 Å, 5 μ m, 100 x 10 mm) using a gradient of 5–95% B (A: 99.9% H₂O, 0.1% TFA; B: 80% acetonitrile (ACN), 19.9% H₂O, 0.1% TFA) over 30-60 min and the samples were lyophilized. The purity and molecular mass of the each peptide were confirmed by liquid chromatography-MS (LC-MS) on a Shimadzu LCMS2020 instrument, incorporating a Phenomenex® Luna C8 column (100 Å, 3 μ m, 100 x 2 mm) using a linear gradient of 100% H₂O (0.05% TFA) for 4 min, followed by 0-60% ACN (0.05% TFA) in water over 10 min at a flow rate of 0.2 mL/min.

2.5 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy is a powerful tool that enables the determination of structures of proteins in solution under near-physiological conditions. Even though proteins with molecular masses larger than about 50 kDa present a major challenge for structure determination, NMR spectroscopy is ideal for structure determination and dynamics for peptides and small proteins (5-25 kDa).⁵⁴ Freeze-dried peptide that was > 95% pure was dissolved in either 93% H₂O/7% ²H₂O or 100% ²H₂O to concentrations ranging from 0.2 to 1.5 mM, and pH was adjusted accordingly. A sufficiently low concentration of ²H₂O was added to the solvent to facilitate lock of the field (typically 10% ²H₂O), while minimizing the loss of signal by deuterium exchange for rapidly exchanging protons. Spectra were acquired using 600 MHz Bruker NMR spectrometer equipped with a cryogenically cooled triple-resonance probe.

2.5.1 One-dimensional ¹H NMR experiments

One-dimensional (1D) NMR spectra of a peptide are often recorded to check the spectral quality and signal dispersion, which gives information about the folded state, propensity of aggregation of the peptide and whether the peptide exists in single conformation.⁵⁴ In this study, one-dimensional NMR was used often to judge the spectral quality and signal dispersion to get an estimation of the folded state of the peptide, for temperature titration, and pH titration experiments. 1D NMR spectra recorded at temperatures ranging from 283 – 313 K, in 5 K increments is also regularly used to obtain information on the thermal stability conformation of the peptide and also to obtain amide temperature coefficients for the residues which will be helpful in identifying the residues that are in hydrogen bonding. 1D NMR spectra recorded at different temperatures also help to identify the optimal temperature where the signal dispersion and spectral quality are best for subsequent two dimensional (2D) NMR experiments.

pH titration experiments were performed to check the stability of the conformation over different pH ranges (pH 2.0 – 9.0), to identify a suitable pH for subsequent 2D NMR experiments, and to identify titratable amino acids based on pH-dependent changes in proton chemical shift.

2.5.2 Two-dimensional ¹H NMR experiments

Because of the large number of protons per residue and the resulting resonance overlap, even smaller peptides cannot be assigned completely from only 1D ¹H NMR spectra. The homonuclear 2D NMR experiments correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) are frequently used in conjunction with 1D ¹H NMR for obtaining the ¹H NMR assignments. The cross-peaks manifest through-bond relations between protons that are separated by not more than three covalent bonds, i.e. which are part of the same amino acid

residue. The connectivities of the protons are described below for the COSY, TOCSY and NOESY experiments.

The most important information for structure determination of peptides is obtained from either ^1H - ^1H ROESY (Rotating frame Overhauser Effect Spectroscopy) for peptides of molecular mass between 0.6- 1.2 kDa or ^1H - ^1H NOESY (Nuclear Overhauser Effect Spectroscopy) for peptides > 1.2 kDa. ^1H - ^1H NOESY was the experiment of choice in my study as the molecular mass of the peptides investigated here was > 2 kDa. A cross-peak between two hydrogen atoms is observed only if those two proton are separated by a shorter distance than approximately 5.0 Å. Since the NOE depends on the through- space distance, the locations of the two interacting protons in the primary structure may be far apart.^{54, 55}

2.5.3 ^1H - ^1H COSY

The COSY displays [^1H , ^1H]-correlations arising from scalar (through-bond) couplings (usually up to four bonds). The positions of cross peaks in the COSY are characteristic for particular amino acids. COSY cross peaks display a multiplet fine structure, which is caused by the scalar couplings. Active couplings (those that lead to the cross peak) display anti-phase components (components with positive and negative signal intensity) whereas all other couplings (passive couplings) are in-phase. ^1H - ^1H COSY was frequently used in my studies for obtaining proton resonance assignments.⁵⁶

2.5.4 ^1H - ^1H TOCSY

A TOCSY experiment contains all cross peaks from protons of the same spin system. Protons from different amino acids always belong to different spin systems, because there is no scalar coupling across the amide bond. Analysis of spin systems identifies which type of amino acids the spin system belongs. Possible criteria for identifying a particular spin system are: occurrence or absence of methyl groups, length of the spin system, positions of chemical shifts in the spin system. Although the exact amino acid can rarely be derived, many amino acids can be excluded from such an analysis. In the study described in this thesis, ^1H - ^1H TOCSY was regularly used to in conjunction with 1D NMR, 2D ^1H - ^1H COSY and ^1H - ^1H NOESY to obtain resonance assignments and sequential assignments. ^1H - ^1H TOCSY spectra with a spin lock time of 70 or 80 ms were acquired using the DIPSI-2 pulse sequence with excitation sculpting for water suppression.^{56, 57}

2.5.5 ^1H - ^1H NOESY

Cross peaks in the NOESY are due to dipolar couplings resulting from interactions of spins via space and hence only depend on the distance but not on the number of intervening bonds. Dipolar couplings are averaged to zero in solution but give rise to very important relaxation

phenomena, one of which is the NOE (nuclear Overhauser effect, NOE).⁵⁶ The strong dependence of the cross peak intensity on the proton separation explains why this parameter is the most useful for structure determination.⁵⁴ ^1H - ^1H NOESY was regularly used in this study to obtain sequential resonance assignments and distance restraints for structure calculations. ^1H - ^1H NOESY spectra were acquired at different mixing times ranging from 50 ms to 300 ms to determine the optimal mixing time for maximum cross peak intensity. For Con-Vc1₁₋₂₂[Z1Q] a mixing time of 200 ms, and for Con-Vc1₁₋₂₂[Q1C, Y9C] a mixing time of 300 ms was selected for cross-peak analyses.

2.6 Sequence-specific resonance assignments

The ^1H - ^1H COSY, ^1H - ^1H TOCSY and ^1H - ^1H NOESY spectra were overlaid in either Topspin or CCPNMR programs for the analysis and assignments of the cross peaks.⁵⁸ The sequence-specific assignment of proton resonances for all the peptides were performed according to procedures developed by Wüthrich and co-workers.⁵⁴ Residues containing methyl groups such as Ala, Thr, Leu and Ile served as good starting points for relatively simple identification of the spin systems and sequential assignment process. The spin systems for proline was identified by the observation of strong NOE cross peaks between the H^α of the preceding residue and Pro H^δ .

2.7 Structure calculation

Structure calculation requires both distance constraints and angle constraints as inputs to different structure calculation programs such as CYANA.⁵⁹ The intensities of cross peaks in the NOESY spectra were utilized to generate distance constraints. $^3J_{\text{HN-H}\alpha}$ coupling constants were measured from 1D ^1H spectra, which yielded ϕ angle constraints. These ϕ angles were restrained to $-120 \pm 30^\circ$ for $^3J_{\text{HN-H}\alpha} \geq 8.0$ Hz and $-65 \pm 25^\circ$ for $^3J_{\text{HN-H}\alpha} \leq 6.0$ Hz. Three distance constraints were added for the disulfide bridge as follows: 2.00, 3.00, and 3.00 Å for S(i)–S(j), S(i)–C β (j), and S(j)–C β (i), respectively. The initial structures were generated using the program CYANA (version 3.0) and then refined by energy minimization using Xplor-NIH.⁶⁰ Conventional simulated annealing protocols were used in Xplor-NIH to generate an ensemble of 100 structures, from which the 20 lowest-energy structures were chosen to represent the solution structure. The root-mean-square deviation (RMSD) values for these structures were assessed using MolMol (version 2K.1).⁶¹ Structural figures were prepared using PyMOL (version 1.5.0.4).

2.8 NMR relaxation measurements

^{15}N relaxation experiments were performed on ^{15}N -labelled full-length contryphan-Vc1 at 20°C on a Bruker 600 MHz spectrometer. Spin-lattice relaxation times (T_1) were obtained from

a series of ^1H - ^{15}N correlation spectra with 0, 50, 100, 200, 400, 800, and 1600 ms relaxation delays. Spin-spin relaxation times (T_2) were obtained using 0, 20, 50, 80, 120, 150, 180, and 240 and 300 ms spin-lock periods. The T_1 and T_2 time constants were calculated using CcpNmr Analysis by plotting peak volumes versus relaxation delay times to an equation for single-order exponential decay. The reported errors are standard deviations derived from the fit of the data. Steady-state ^1H - ^{15}N NOE values were determined from the ratio of peak intensities for spectra collected with and without 3 s proton presaturation. The peak intensities of residues with and without proton presaturation were calculated utilizing CARA.

2.9 Proteolysis assays

A potentially important attribute of a peptide for use in pharmaceutical applications or as a scaffold is resistance to proteolytic degradation.⁶ Proteolytic stability for the peptides in this study was tested in the presence of the common digestive enzymes trypsin, chymotrypsin and pepsin. Proteolysis assays were performed at a 250:1 substrate (peptide)/enzyme ratio. at 37°C for up to 4 h. As a positive control to ensure that active enzyme was present, bovine serum albumin was used as substrate. All digestion assay products were analysed by LC-MS (0-60 % ACN gradient, 10 min). Trypsin (EC 3.4.21.4, Sigma) and α -chymotrypsin (EC 3.4.21.1, Sigma) stocks were prepared in 50 mM Tris, 100 mM NaCl (pH 7.4), and pepsin (EC 3.4.23.1, Sigma) stocks were prepared in 10 mM HCl (pH 2). The reactions for trypsin and α -chymotrypsin were performed in 50mM Tris, 100mM NaCl, 2mM CaCl_2 , pH 7.4. The reaction for pepsin was performed in 1 mM HCl, pH 2. The trypsin and α -chymotrypsin reactions were quenched with 0.1% TFA, and the pepsin reaction was quenched with 50% 1 M NaOH.

2.10 Surface plasmon resonance

Surface plasmon resonance (SPR) biosensors are optical sensors that use special electromagnetic waves (surface plasmon-polaritons) to probe interactions between an analyte in solution and a biomolecular recognition element immobilized on the SPR sensor surface.⁶² SPR detects changes in the refractive index of the medium adjacent to the thin metal film (in this case, the sensor chip with gold film), which is dependent on the mass of the surface. SPR has been used to study a diverse set of interaction partners of biological interest, such as protein–protein, protein–lipids, protein–nucleic acids, or protein and low molecular weight molecules such as drugs, substrates, and cofactors.

The binding affinities of Vc1 peptide analogues were analyzed by SPR, using a Biacore T200 instrument. All experiments were performed in degassed 25 mM HEPES buffer (pH 7.4) containing 150 mM NaCl, 3 mM EDTA, and 0.05% surfactant P-20 at 25°C. Human SPSB2 was immobilized on a Biacore CM5 biosensor chip by amine coupling as follows: the CM-

dextran matrix was activated with 0.2 M 1-ethyl-3(3-diethylaminopropyl)-carbodiimide hydrochloride (EDC) and 0.05 M *N*-hydroxysuccinimide. Human SPSB2 (at 100 µg/mL in 10 mM sodium acetate [pH 5.5]) was then flowed over the activated surface at 10 µL/min for 10 min. Finally, 1 M ethanolamine-HCl, pH 8.5, was injected into both target and reference flow cells to deactivate any remaining activated carboxyl groups on the surface. The binding of the Vc1 contryphan-Vc1 peptide analogues was investigated in a running buffer containing 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.05% Surfactant P-20, with and without 1m DTT, pH 7.4. The peptides were injected onto the surface with contact time of 100 s, flow rate of 100 µL/min, and dissociation time of 600 s. Sensorgrams were further corrected for nonspecific binding to the surface by subtracting the signals of the reference surface from those of the target protein surface. The corrected signal was then fitted to a steady state 1:1 interaction model with Biacore T200 Evaluation Software (version 2.0), in order to estimate the binding affinity (K_D), association rate, k_{on} , and dissociation rate, k_{off} .

Lipid binding by Con-Vc1₁₋₂₂[Z1Q] and full-length contryphan-Vc1 was analysed on a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) immobilized Biacore L1 chip using buffer containing degassed 20 mM HEPES buffer (pH 7.4) containing 50 mM NaCl, and 3 mM EDTA at 25°C. A POPC solution at 600 µM in SPR buffer was immobilized on the Biacore L1 chip and the binding of contryphan-Vc1 peptides was investigated in the same buffer.

2.11 Molecular dynamics simulations

Molecular Dynamics simulations (MD) is a computational method which has become one of the principal tools in the theoretical study of biological molecules. This method calculates the time-dependent behaviour of a molecular system, providing detailed information on the fluctuations and conformational changes of biological molecules. These methods are now routinely used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes.^{63, 64}

MD simulations were carried out using GROMACS version 5.0.4,⁶⁵ utilising the GROMOS 54a7 united-atom force field and a 2 fs time step.⁶⁶ Temperature coupling made use of the velocity rescale algorithm with a reference temperature of 298 K. Pressure coupling used the Parinello-Rahman algorithm, with a reference pressure of 1 bar and compressibility of 4.5×10^{-5} bar⁻¹. Initial structures were generated using the Maestro software package (v 10.3.014), by *in silico* mutation of the published structures of contryphan-Vc1 (PDB codes: 2N24 and 5KKM). These structures were solvated with SPC water and subjected to a steepest-descent minimisation of 2000 steps to remove bad van der Waals contacts between atoms. Temperature equilibration (without pressure coupling) was run for 10,000 steps. Isotropic pressure coupling

was then applied for 500,000 steps. Following equilibration, the simulation production runs were executed for 300 ns each. The resultant trajectories were visualised with the Visual Molecular Dynamics (VMD) software package (v 1.9.2)⁶⁷ and analysed with VMD and GROMACS built-in tools.

Chapter 3

**The Single Disulfide-Directed β -Hairpin
Fold. Dynamics, Stability and
Engineering**

Declaration by candidate

In the case of **Chapter 3**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Recombinant production and synthetic production of contryphan-Vc peptides and its analogues, recording and analysis of NMR data, structure calculation, performed stability assays, proteolysis assays, redox assays and manuscript preparation.	60

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Bankala Krishnarjuna	Performed experiments, data analysis, manuscript preparation	
Rodrigo A. V. Morales	Intellectual input	
Christopher A. MacRaild	Intellectual input, data analysis	
Maiada Sadek	Performed experiments	2
Eleanor W. W. Leung	Performed experiments, data analysis,	
Samuel D. Robinson	Manuscript preparation	
Michael W. Pennington	Manuscript preparation	
Raymond S. Norton	Intellectual input, manuscript preparation	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's

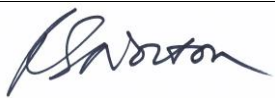
Signature

	Date 02/07/2018
---	---------------------------

Main

Supervisor's

Signature

	Date 02/07/2018
---	---------------------------

The Single Disulfide-Directed β -Hairpin Fold. Dynamics, Stability, and Engineering

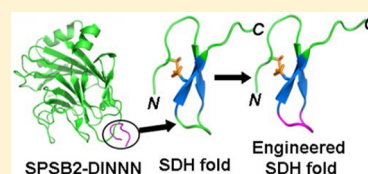
Balasubramanyam Chittoor,[†] Bankala Krishnarjuna,[†] Rodrigo A. V. Morales,[†] Christopher A. MacRaild,[†] Maiada Sadek,[†] Eleanor W. W. Leung,[†] Samuel D. Robinson,^{†,§} Michael W. Pennington,[‡] and Raymond S. Norton^{*,†,§}

[†]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

[‡]Peptides International, Louisville, Kentucky 40299, United States

S Supporting Information

ABSTRACT: Grafting bioactive peptide sequences onto small cysteine-rich scaffolds is a promising strategy for enhancing their stability and value as novel peptide-based therapeutics. However, correctly folded disulfide-rich peptides can be challenging to produce by either recombinant or synthetic means. The single disulfide-directed β -hairpin (SDH) fold, first observed in contryphan-Vc1, provides a potential alternative to complex disulfide-rich scaffolds. We have undertaken recombinant production of full-length contryphan-Vc1 (rCon-Vc1[Z1Q]) and a truncated analogue (rCon-Vc1_{1–22}[Z1Q]), analyzed the backbone dynamics of rCon-Vc1[Z1Q], and probed the conformational and proteolytic stability of these peptides to evaluate the potential of contryphan-Vc1 as a molecular scaffold. Backbone ¹⁵N relaxation measurements for rCon-Vc1[Z1Q] indicate that the N-terminal domain of the peptide is ordered up to Thr19, whereas the remainder of the C-terminal region is highly flexible. The solution structure of truncated rCon-Vc1_{1–22}[Z1Q] was similar to that of the full-length peptide, indicating that the flexible C-terminus does not have any effect on the structured domain of the peptide. Contryphan-Vc1 exhibited excellent proteolytic stability against trypsin and chymotrypsin but was susceptible to pepsin digestion. We have investigated whether contryphan-Vc1 can accept a bioactive epitope while maintaining the structure of the peptide by introducing peptide sequences based on the DINNN motif of inducible nitric oxide synthase. We show that sCon-Vc1_{1–22}[NNN_{12–14}] binds to the iNOS-binding protein SPSB2 with an affinity of 1.3 μ M while maintaining the SDH fold. This study serves as a starting point in utilizing the SDH fold as a peptide scaffold.



Grafting bioactive peptide sequences that target specific protein–protein interactions onto stable scaffolds is a promising approach for the development of peptide-based therapeutics.¹ Disulfide-rich peptides have proven to be valuable acceptor scaffolds, with applications as drugs, pharmacological probes, and imaging agents.^{2–7} A valuable scaffold in such applications is the inhibitor cystine knot (ICK) motif,^{8,9} which is found in peptides from plants, cone snails, scorpions, and spiders.^{10,11} The ICK motif peptides agatoxin and Agouti-related protein (AgRP) have been engineered for *in vivo* tumor imaging¹² and targeting tumor angiogenesis,¹³ respectively.

The production of correctly folded disulfide-rich peptides by chemical synthesis or recombinant expression can be challenging, especially for peptides containing three or four disulfide bridges.^{14–16} Recently, contryphan-Vc1, a 31-residue peptide identified in the venom of the marine cone snail *Conus victoriae*,^{17,18} has been shown to have a unique fold, designated as the single disulfide-directed β -hairpin (SDH).¹⁹ The core structure of contryphan-Vc1 is a double-stranded antiparallel β -sheet stabilized by just a single disulfide bridge. The β -hairpin core of contryphan-Vc1 consists of two loops, the first of which comprises residues Gln4, Pro5, Gly6, and Tyr7 (QPGY), which form a type II β -turn that leads to the first β -strand. The second loop comprises residues Pro11, Val12, Leu13, and Gly14, which

form a type IV β -turn. Residues from Thr19 to Tyr31 are flexible and unstructured. The β -hairpin core of contryphan-Vc1 displays an ordered structure and remarkable thermal stability and is very similar structurally to the ICK fold despite its simpler cysteine framework; these properties suggest that it may have potential as a useful scaffold for grafting.

In this study, we have investigated the stability of the SDH fold to chemical denaturation and proteolytic enzymes, as well as its resilience to truncation and the insertion of non-native peptide motifs, to evaluate its potential as a scaffold. We have undertaken recombinant production of full-length contryphan-Vc1 (rCon-Vc1[Z1Q]) by replacing the N-terminal pyroglutamic acid with a glutamine residue. We analyzed the backbone dynamics of contryphan-Vc1 to identify the ordered and flexible regions of the peptide and showed that the flexible C-terminal region could be deleted without affecting the ordered structure of the peptide. We also investigated whether contryphan-Vc1 can accept a bioactive epitope by utilizing the DINNN motif of inducible nitric oxide synthase (iNOS), which is essential for recognition of iNOS by the SPRY

Received: February 11, 2017

Revised: April 20, 2017

Published: April 24, 2017

Table 1. Peptide Sequences Investigated in This Study

Peptide	Sequence
sCon-Vc1 ₁	^a ZWCQPGYAYNPVLGICTITLSRIEHPGNYDY
rCon-Vc1[Z1Q]	QWCQPGYAYNPVLGICTITLSRIEHPGNYDY
rCon-Vc1 ₁₋₂₂ [Z1Q]	QWCQPGYAYNPVLGICTITLSR
sCon-Vc1 ₁₋₂₂ [NNN ₁₂₋₁₄]	ZWCQPGYAYNPNNICTITLSR
sCon-Vc1 ₁₋₂₂ [Z1Q,DINNN ₄₋₈]	QWCDINNNAYNPVLGICTITLSR
sCon-Vc1 ₁₋₂₂ [Z1Q,DINNN ₁₂₋₁₆]	QWCQPGYAYNPINNICTITLSR

^aPyroglutamate modification is indicated as Z.

domain-containing SOCS (suppressor of cytokine signaling) box (SPSB) proteins in circulating macrophages.^{20–22} The peptide analogues sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈], sCon-Vc1₁₋₂₂[Z1Q,DINNN₁₂₋₁₆], and sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄] were tested for their affinity for human SPSB2 utilizing surface plasmon resonance (SPR) and analyzed conformationally by NMR.

EXPERIMENTAL PROCEDURES

Construction of Plasmids Expressing Full-Length and Truncated Contryphan-Vc1. Codon-optimized DNA corresponding to the mature peptide sequence of contryphan-Vc1 was synthesized (GenScript) with SalI and XhoI restriction sites, respectively, at the 5' and 3' ends. An enterokinase cleavage site was included at the 5' end after the SalI restriction site. The gene fragment was ligated into the pET32a vector (Novagen) using the SalI and XhoI restriction sites to generate modified pET32a expressing a thioredoxin (Trx) fusion protein of full-length contryphan-Vc1. The DNA sequence of contryphan-Vc1 was then amplified with a forward primer for contryphan-Vc1 with a SalI restriction site and enterokinase cleavage site and a reverse primer with a stop codon after the Arg22 codon, and the gene fragment was ligated into the pET32a vector using the SalI and XhoI restriction sites to generate modified pET32a expressing the Trx fusion protein of truncated contryphan-Vc1 (rCon-Vc1₁₋₂₂[Z1Q]) (Table 1).

Peptide Expression and Purification. Trx fusion proteins of both full-length contryphan-Vc1 and the truncated analogue Con-Vc1₁₋₂₂ were expressed in *Escherichia coli* BL21(DE3) cells at 30 °C, induced for 4 h with 0.5 mM isopropyl β-D-1-thiogalactopyranoside. To generate ¹⁵N-labeled and ¹⁵N- and ¹³C-labeled contryphan-Vc1 peptides, *E. coli* BL21(DE3) cells were grown in M9 minimal medium supplemented with 1 g/L ¹⁵NH₄Cl and 3 g/L [¹³C]glucose. For purification of the Trx-fused peptides, cells were resuspended in bug buster (Novagen) at a concentration of 1 g of cell paste per mL and incubated for 20 min in the presence of EDTA-free protease inhibitor cocktail (Roche) and then clarified by centrifugation at 23660g for 30 min at 4 °C. The supernatant was loaded onto a pre-equilibrated Ni-NTA Hitrap column (GE Healthcare, 5 mL) at a flow rate of 0.25 mL/min. Unbound material was removed by extensive washing with buffer A [20 mM Tris-HCl (pH 8), 10 mM imidazole, and 100 mM NaCl] followed by a wash with buffer A containing 0.5 M NaCl. The fusion protein was eluted with buffer A containing 250 mM imidazole and buffer-exchanged with enterokinase cleavage buffer [50 mM Tris-HCl (pH 8), 1 mM CaCl₂, and 50 mM NaCl]. The Trx tag was removed by incubating the fusion protein with enterokinase (NEB) at a concentration of 0.005 μg/mg of protein at 23 °C

for 16 h and passing the mixture onto a Ni-NTA Hitrap column. Unbound supernatant was collected and purified by reversed-phase HPLC (RP-HPLC) on a Phenomenex Luna C18 column (100 Å, 5 μm, 100 mm × 10 mm) using a gradient from 5 to 95% B [A, 99.9% H₂O and 0.1% TFA; B, 80% acetonitrile (ACN), 19.9% H₂O, and 0.1% TFA] over 30–60 min, and then the samples were lyophilized. The peptides were further purified by peptide gel filtration using a Superdex peptide 10/300 GL column, and the purity and molecular mass of each peptide were confirmed by liquid chromatography and mass spectrometry (LC–MS) on a Shimadzu LCMS2020 instrument, incorporating a Phenomenex Luna C8 column (100 Å, 3 μm, 100 mm × 2 mm) using a linear gradient of water having 0.05% TFA for 4 min, followed by 0–60% ACN (0.05% TFA) in water over 10 min at a flow rate of 0.2 mL/min (Table 1 and Figures S1 and S2).

Peptide Synthesis. Full-length contryphan-Vc1 was synthesized using *N*-(9-fluorenyl) methoxycarbonyl (Fmoc)-Tyr (tBu)-Wang resin on a Prelude automated peptide synthesizer (Protein Technologies, Tucson, AZ). All couplings were mediated with diisopropylcarbodiimide with 6-chloro-1-hydroxybenzotriazole (6-Cl-HOBT) for 3 h. The peptide was cleaved from the resin support and simultaneously deprotected using a cocktail of trifluoroacetic acid, anisole, triisopropylsilane, thioanisole, water, and 1,2-ethanedithiol [9:1:2:1:1:1 (v/v)] for 2 h at room temperature. The crude peptide was precipitated and washed thrice with ice-cold diethyl ether, then dissolved in 50% aqueous acetic acid, and diluted in water to a concentration of 0.3 mg/mL. The pH of the peptide solution was adjusted to 7.8 with ammonium hydroxide (NH₄OH). Disulfide bond formation was accelerated by adding 1 mL of 3% hydrogen peroxide (H₂O₂), and the peptide solution was allowed to stir gently for 18 h. Other contryphan-Vc1 peptide analogues were synthesized on a PTI Instruments PS3 peptide synthesizer, using Rink amide AM resin. Peptides were deprotected with 20% piperidine in dimethylformamide (DMF), activated with 70 mL/L *N,N*-diisopropylethylamine (DIPEA) in DMF, and coupled with a 3-fold excess of 2-(6-chloro-1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethylammonium hexafluorophosphate (HCTU) for 50 min. Cleavage from the resin was performed over 2 h with a mixture of 3,6-dioxo-1,8-octanedithiol, triisopropylsilane, 1,3-dimethoxybenzene, and trifluoroacetic acid [2.5:2.5:5:92.5 (v/v) DODT/TIPS/DMB/TFA]. The cleavage mixture was purged with nitrogen, and the crude peptide was precipitated and washed thrice with ice-cold diethyl ether, then dissolved in 0.1 M ammonium bicarbonate (pH 8.0), and allowed to fold by adjusting the concentration to 0.3 mg/mL. The crude folded peptide was purified by RP-HPLC on a Phenomenex Luna C18 column (100 Å, 5 μm, 100

mm \times 10 mm) using a gradient of 5 to 95% B (A, 99.9% H₂O and 0.1% TFA; B, 80% ACN, 19.9% H₂O, and 0.1% TFA) over 30–60 min, and the samples were lyophilized. The purity and molecular mass of the each peptide were confirmed by LC–MS on a Shimadzu LCMS2020 instrument, incorporating a Phenomenex Luna C8 column (100 Å, 3 μ m, 100 mm \times 2 mm) using a linear gradient of water having 0.05% TFA for 4 min, followed by 0 to 60% ACN (0.05% TFA) in water over 10 min at a flow rate of 0.2 mL/min (Table 1 and Figures S1 and S3).

NMR Spectroscopy. All spectra were recorded on a Bruker 600 MHz spectrometer equipped with a cryogenically cooled triple-resonance probe. The lyophilized peptide was dissolved in either 93% H₂O with 7% ²H₂O or 100% ²H₂O (pH 4.0). One-dimensional ¹H spectra were recorded at different temperatures between 5 and 30 °C, at intervals of 5 °C, and between pH 3 and 9. Two-dimensional NMR spectra utilized for sequence-specific assignments and structure calculations were recorded at pH 4 and 20 °C. Two-dimensional homonuclear TOCSY spectra with a spin-lock time of 80 ms were recorded using the DIPSI-2 pulse sequence²³ with excitation sculpting for water suppression²⁴ at 15, 20, and 25 °C. Two-dimensional NOESY spectra were recorded at two different mixing times, 50 and 200 ms, to analyze the time dependence of NOE intensities. A DQF-COSY spectrum was recorded in 100% ²H₂O for measuring *J* couplings. ¹³C HSQC and ¹⁵N HSQC spectra were recorded for carbon and nitrogen chemical shifts, respectively. A sine-bell squared window function was used for processing spectra. All spectra were processed using Bruker TopSpin (version 3.2) and analyzed using CcpNmr Analysis (version 2.1.5).²⁵ The sequence-specific resonance assignments are summarized in Table S1 and deposited in BMRB²⁶ (entry 30124). For the determination of potential amide backbone hydrogen bonding in rCon-Vc1_{1–22}[Z1Q], a series of one-dimensional ¹H spectra was recorded at 10, 15, 20, 25, and 30 °C, and the amide resonance temperature coefficients were determined from the slope of linear least-squares fits to the data.

Structure Calculation. The intensities of cross-peaks in NOESY spectra with a mixing time of 200 ms were utilized to generate distance constraints. ³J_{HN–H α} coupling constants were measured from one-dimensional ¹H spectra, which yielded 11 ϕ angle constraints. These ϕ angles were restrained to $-120 \pm 30^\circ$ for ³J_{HN–H α} values of ≥ 8.0 Hz and $-65 \pm 25^\circ$ for ³J_{HN–H α} values of ≤ 6.0 Hz. Three distance constraints were added for the disulfide bridge as follows: 2.00, 3.00, and 3.00 Å for S(*i*)–S(*j*), S(*i*)–C β (*j*), and S(*j*)–C β (*i*), respectively. The initial structures of contryphan-Vc1_{1–22} were generated using CYANA (version 3.0)²⁷ and then refined utilizing Xplor-NIH.²⁸ Structure calculations were performed using 297 interproton distance constraints derived from the NOESY spectrum (88 intrasidue, 88 sequential, 66 medium-range, and 55 long-range NOE constraints), 11 dihedral angle constraints derived from ³J_{HN–H α} *J* coupling measurements from one-dimensional ¹H, two other dihedral angle constraints derived from two-dimensional DQF-COSY spectra, and three disulfide bond restraints. Conventional simulated annealing protocols were used in Xplor-NIH to generate an ensemble of 100 structures, from which the 20 lowest-energy structures were chosen to represent the solution structure of rCon-Vc1_{1–22}[Z1Q]. The root-mean-square deviation (RMSD) values for these structures were assessed using MolMol (version 2K.1).²⁹ Structural figures were prepared using PyMOL (version 1.5.0.4).

NMR Relaxation Measurements. ¹⁵N relaxation experiments were performed at 20 °C on a Bruker 600 MHz spectrometer. Spin–lattice relaxation times (*T*₁) were obtained from a series of ¹H–¹⁵N correlation spectra with relaxation delays of 0, 50, 100, 200, 400, 800, and 1600 ms. Spin–spin relaxation times (*T*₂) were obtained using spin-lock periods of 0, 20, 50, 80, 120, 150, 180, 240, and 300 ms. The *T*₁ and *T*₂ time constants were calculated using CcpNmr Analysis by plotting peak volumes versus relaxation delay times with an equation for single exponential decay. The reported errors are standard deviations derived from the fit of the data. Steady-state ¹H–¹⁵N nuclear Overhauser enhancement (NOE) values were determined from the ratio of peak intensities for spectra recorded with and without 3 s proton presaturation. The peak intensities of residues with and without proton presaturation were calculated utilizing CARA.³⁰

Reduction Assay. Tris(2-carboxyethyl)phosphine (TCEP) stock solutions (0.5 M) were prepared in Milli-Q water immediately prior to use. TCEP was added to 0.5 mM synthetic contryphan-Vc1 (sCon-Vc1) to a final concentration of 10 mM, and the sample was incubated at room temperature. Peptide samples were assayed at time intervals of 0, 1, 6, 18, 24, and 48 h using LC–MS on a Shimadzu LCMS2020 instrument incorporating a Phenomenex Luna C8 column (100 Å, 3 μ m, 100 mm \times 2 mm) with a linear gradient of 0.05% TFA in water for 4 min, followed by 0 to 60% ACN (0.05% TFA) in water over 10 min at a flow rate of 0.2 mL/min. One-dimensional ¹H NMR spectra were recorded on synthetic contryphan-Vc1 (sCon-Vc1) in 6 mM TCEP (pH 4 and 20 °C) after incubation for 1, 2, 16, and 24 h.

Proteolysis Assays. Proteolysis assays were performed at a 250:1 substrate (peptide):enzyme ratio with pepsin, trypsin, and α -chymotrypsin. For all assays, peptides were incubated with protease at 37 °C for up to 4 h. As a positive control to ensure that active enzyme was present, bovine serum albumin was used as a substrate. All digestion assay products were analyzed by LC–MS (0 to 60% ACN gradient, 10 min). Trypsin (EC 3.4.21.4, Sigma) and α -chymotrypsin (EC 3.4.21.1, Sigma) stocks were prepared in 50 mM Tris (pH 7.4) and 100 mM NaCl, and pepsin (EC 3.4.23.1, Sigma) stocks were prepared in 10 mM HCl (pH 2). The reactions for trypsin and α -chymotrypsin were performed in 50 mM Tris-HCl (pH 7.4), 100 mM NaCl, and 2 mM CaCl₂. The reaction for pepsin was performed in 1 mM HCl (pH 2). The trypsin and α -chymotrypsin reactions were quenched with 0.1% TFA, and the pepsin reaction was quenched with 50% 1 M NaOH.

Surface Plasmon Resonance. The binding affinities of the Con-Vc1 peptide analogues were analyzed by surface plasmon resonance (SPR), using a Biacore T200 instrument (GE Healthcare). All experiments were performed in degassed buffer containing 25 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.005% surfactant P-20 at 25 °C. Human SPSB2 was immobilized on a Biacore CM5 biosensor chip by amine coupling as follows. The CM-dextran matrix was activated with 0.2 M 1-ethyl-3-[3-(diethylamino)propyl]carbodiimide hydrochloride (EDC) and 0.05 M *N*-hydroxysuccinimide; 100 μ g/mL human SPSB2 in 10 mM sodium acetate (pH 5.5) was then passed over the activated surface at a rate of 10 μ L/min for 10 min. Finally, 1 M ethanolamine-HCl (pH 8.5) was injected into both target and reference flow cells to deactivate any remaining activated carboxyl groups on the surface. The binding of the contryphan-Vc1 peptide analogues to immobilized human SPSB2 was investigated in a running buffer containing 10

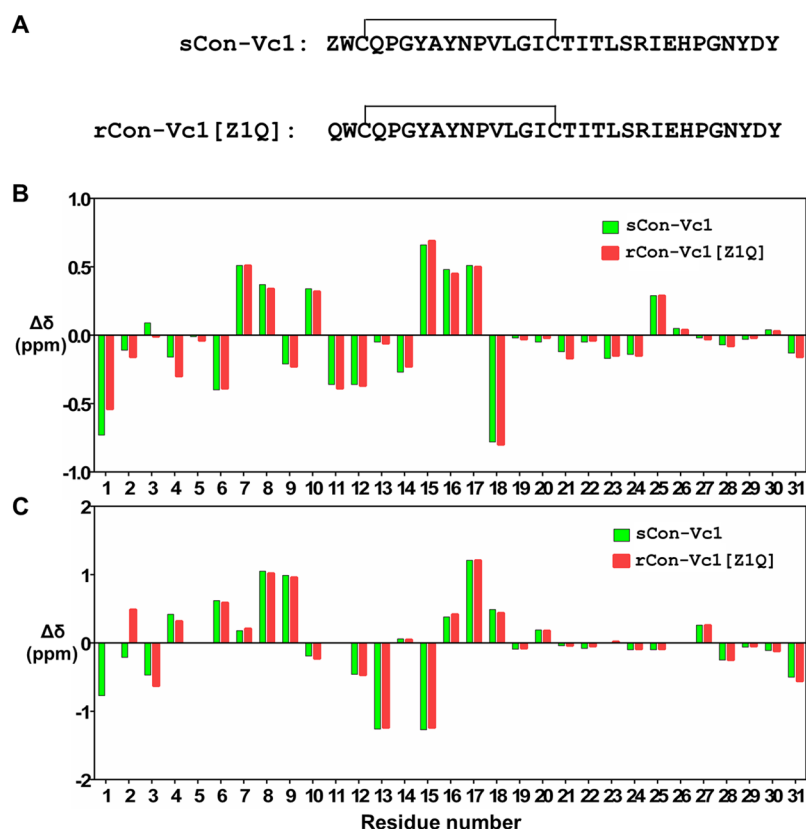


Figure 1. Secondary chemical shift difference plots for synthetic and recombinant contryphan-Vc1. (A) Sequences of sCon-Vc1 and rCon-Vc1[Z1Q] indicating disulfide connectivity. (B) H^α shift deviations from random coil chemical shifts. (C) H^β shift deviations from random coil chemical shifts.³¹

mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.005% surfactant P-20, with and without 1 mM DTT (pH 7.4). The peptides were injected onto the surface with a contact time of 100 s at a flow rate of 100 μ L/min and a dissociation time for 600 s. Sensorgrams were further corrected for nonspecific binding to the surface by subtracting the signals of the reference surface from that of the human SPSB2-bound surface. The corrected signal was then fitted to a 1:1 Langmuir binding model using Biacore T200 Evaluation Software (version 2.0) to estimate the association rate, k_{on} , and dissociation rate, k_{off} . In cases in which kinetics could not be measured, SPR data were evaluated using a 1:1 steady-state binding model to determine the binding affinity (K_D).

RESULTS

Recombinant Production of Contryphan-Vc1 Peptides. Full-length contryphan-Vc1 (rCon-Vc1[Z1Q]) and a truncated analogue, rCon-Vc1_{1–22}[Z1Q], were expressed as Trx fusion proteins in *E. coli* BL21(DE3). The Trx fusion proteins were expressed in M9 medium to produce 15 N-labeled and 15 N- and 13 C-labeled peptides. The Trx fusion proteins were highly soluble, with yields of 20 mg/L from expression in both Luria-Bertani broth and M9 medium. Clarified cell lysates containing His-tagged Trx fusion proteins were loaded onto a Ni-NTA column, and then bound protein was eluted from the Ni-NTA column and cleaved with enterokinase. The cleaved fusion protein was passed through a Ni-NTA column to remove His-tagged Trx, and the desired peptide was purified by RP-HPLC. The identities and purities of the peptides were confirmed by LC-MS. RP-HPLC and LC-MS showed that purified rCon-

Vc1[Z1Q] and rCon-Vc1_{1–22}[Z1Q] were essentially homogeneous and ~95% pure (Figure S2).³⁰

Comparison of Synthetic and Recombinant Contryphan-Vc1. Contryphan-Vc1 from the venom gland of *Conus victoriae* has an N-terminal pyroglutamate modification, which is in the proximity of the Tyr9 side chain in the structure of the peptide.¹⁹ To assess the role of the pyroglutamate modification in the structure of the peptide, we produced recombinant full-length contryphan-Vc1 having unmodified glutamine at the N-terminus and compared its chemical shifts with those of synthetic contryphan-Vc1 with pyroglutamate at its N-terminus. Plots of the deviations of H^α and backbone chemical shifts from random coil values were nearly identical for the two peptides, with minor differences observed only near the N-terminus (Figure 1). This confirms that synthetic and recombinant contryphan-Vc1 have similar structures and that the pyroglutamate modification has no significant effect on the structure of contryphan-Vc1.

Backbone Dynamics of Recombinant Contryphan-Vc1. Backbone 1H – ^{15}N NOEs and ^{15}N R_1 and R_2 relaxation rates for rCon-Vc1[Z1Q] are shown in Figure 2. Uniform heteronuclear NOE values of ~0.6 were observed for residues Trp2–Ile18, consistent with the dynamics of these residues being dominated by a single process on a time scale of approximately 2 ns, as would be expected for the overall tumbling of a molecule of this size. In contrast, most residues beyond Thr19 showed either negative or zero 1H – ^{15}N NOE values, indicating the presence of faster dynamic processes (Figure 2). Consistent with the heteronuclear NOE values, the R_1 and R_2 rates for residues Trp2–Ile18 are relatively uniform

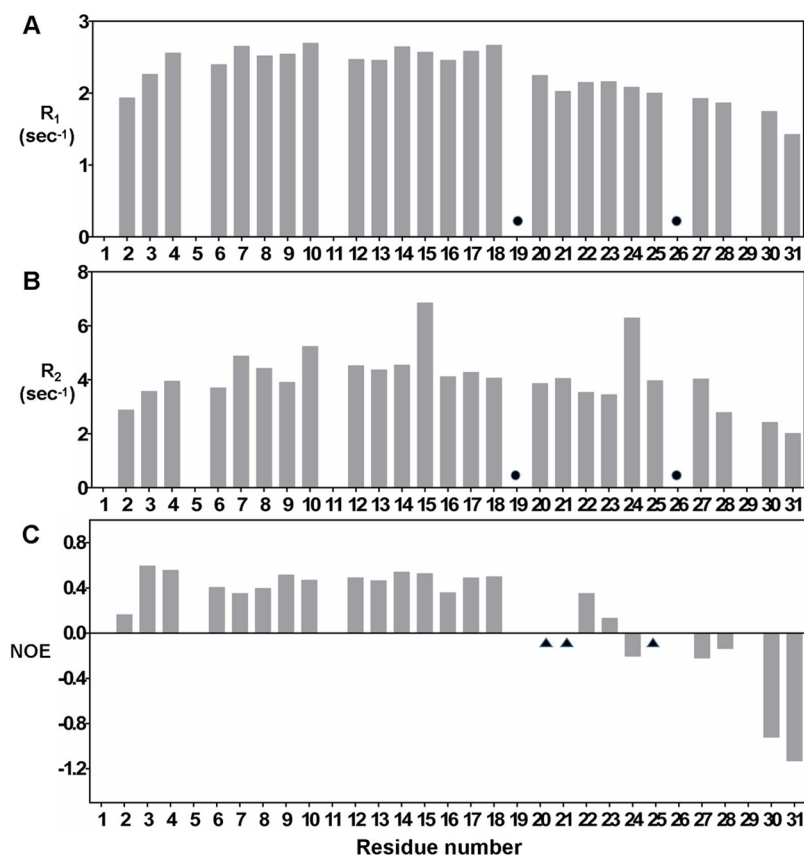


Figure 2. R_1 , R_2 , and ^1H – ^{15}N NOE parameters (A–C, respectively) for recombinant full-length contryphan-Vc1. Overlapping residues are denoted with circles, and the residues with zero intensity after NOE saturation are denoted with triangles.

and larger than those of the C-terminal residues Thr19–Tyr31. Together, these relaxation data indicate a well-defined conformation and limited internal mobility for residues Trp2–Ile18 and significant flexibility across the C-terminal tail.

Sequence-Specific Resonance Assignments for rCon-Vc1_{1–22}[Z1Q]. Spin systems were identified through combined analysis of DQF-COSY, TOCSY (80 ms spin-lock time), and NOESY (200 ms mixing time) spectra recorded at 293 K and pH 3.9. Sequential resonance assignments were made utilizing two-dimensional NOESY and three-dimensional HNCA and HNCACB spectra. Residues containing methyl groups such as Ala8, Thr17 and -19, Leu13 and -20, and Ile15 and -18 served as good starting points for the identification of spin systems and the sequential assignment process. The spin systems for Pro5 and Pro11 were identified by the observation of strong NOE cross-peaks between Gln4 α -Pro5 δ and Asn10 α -Pro11 δ ; the presence of strong $\alpha\delta$ NOEs confirmed the *trans* conformations of both X–Pro bonds. Complete backbone and side chain proton resonance assignments were obtained for all spin systems (Table S1). Backbone resonances of recombinant full-length contryphan-Vc1 were assigned by comparing the chemical shifts with those of synthetic contryphan-Vc1.¹⁹

Solution Structure of Con-Vc1_{1–22}. To investigate the influence of the unstructured and highly flexible C-terminal tail of contryphan-Vc1, the structure of truncated contryphan-Vc1 (rCon-Vc1_{1–22}[Z1Q]) was determined and compared with that of full-length contryphan-Vc1.¹⁹ Structural constraints are summarized in Table 2. rCon-Vc1_{1–22}[Z1Q] retained its ordered structure after deletion of the flexible C-terminal tail region and is almost identical to that of the full-length peptide,

Table 2. Structural Statistics for rCon-Vc1_{11–22}[Z1Q]

NMR Distance and Dihedral Constraints	
no. of distance constraints	
total NOEs	297
intraresidue	88
inter-residue	
sequential ($ i - j = 1$)	88
medium-range ($1 < i - j < 5$)	66
long-range ($ i - j > 5$)	55
hydrogen bond	4
total no. of dihedral angle restraints	
backbone (φ angle)	12
side chain (χ_1 angle)	3
Structural Statistics	
energy ^a	
E_{NOE} (kcal mol ⁻¹)	0.605 \pm 0.41
E_{vdw} (kcal mol ⁻¹)	3.02 \pm 0.20
RMSD among 20 conformers (residues 2–16), average pairwise RMSD ^b	
backbone (\AA) (N, C α , C)	0.36 \pm 0.12
all heavy atoms (\AA)	0.91 \pm 0.25
Ramachandran Analysis	
residues in most favored regions (%)	68.4
residues in additionally allowed regions (%)	31.6
residues in generously allowed regions (%)	0.0
residues in disallowed regions (%)	0.0

^aThe values for E_{NOE} were calculated from a square-well potential with force constants of 150 kcal mol⁻¹ \AA^2 . ^bThe pairwise RMSD was calculated across 20 refined structures.

as shown in Figure 3. Backbone superimposition of the full-length form and rCon-Vc1_{1–22}[Z1Q] yielded an RMSD of 0.74

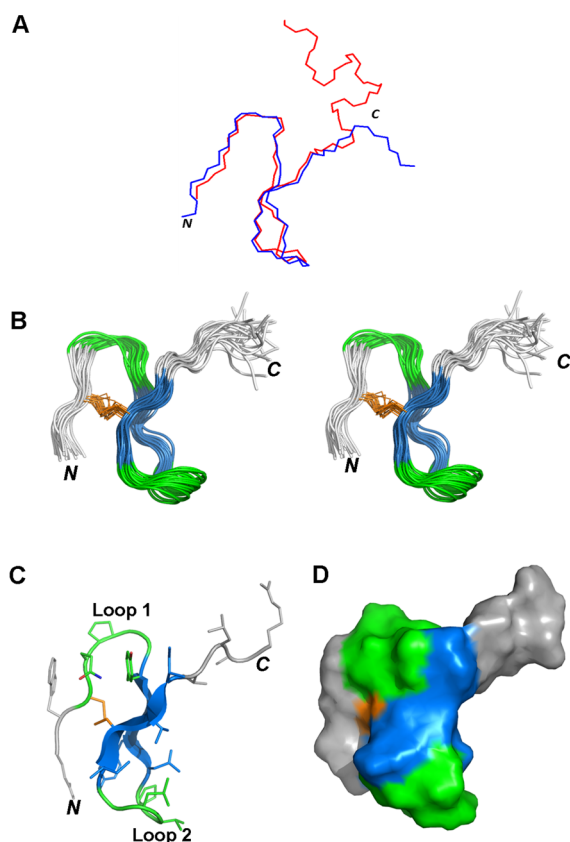


Figure 3. (A) Backbone superposition (using MolMol) of the closest-to-average structures of recombinant rCon-Vc1_{1–22}[Z1Q] (blue) with sCon-Vc1 (red).¹⁹ (B) Stereoview of the ensemble of the final 20 structures of rCon-Vc1_{1–22}[Z1Q] from Xplor-NIH superimposed over the backbone heavy atoms (N, C α , and C) of residues 3–16. (C) PyMOL representation of the closest-to-average structure of rCon-Vc1_{1–22}[Z1Q]. Loop 1 and loop 2 are colored green; β -strands are colored blue, and the unstructured C-terminus is colored gray. (D) Surface representation of structure of rCon-Vc1_{1–22}[Z1Q], with the single disulfide bridge colored bright orange.

Å over C α atoms of residues 3–16 (Figure 3A). The main feature of the structure of rCon-Vc1_{1–22}[Z1Q] is an antiparallel β -sheet with two β -strands connected by a β -turn (Figure 3B,C). The first two residues of rCon-Vc1_{1–22}[Z1Q], Glu1 and Trp2, were less well-defined and had R_1 and R_2 rates that were smaller than those of residues in the β -hairpin core of the peptide. Residues Gln4, Pro5, Gly6, and Tyr7 form a type II β -turn that leads to the first β -strand. Ala8, Tyr9, and Asn10 form the first β -strand, and Ile15, Cys16, and Thr17 make up the other strand. Pro11, Val12, Leu13, and Gly14 form a type IV β -turn. Residues after the second β -strand, Ile18–Arg22, are unstructured and highly flexible. The amides of Ala8, Asn10, Ile15, and Thr17 had temperature coefficients less negative than -4.75 ppb/K, indicating that these residues participate in hydrogen bonding. The hydrogen bonds inferred from the structure are between Ala8 H^N and Thr17 O, Asn10 H^N and Ile15 O, Ile15 H^N and Asn10 O, and Thr17 H^N and Ala8 O, which are all part of the antiparallel β -sheet.

Conformational Stability of Contryphan-Vc1. Contryphan-Vc1 has been shown to have remarkable thermal

stability.¹⁹ The conformational stability of the SDH fold was explored further by observing the effects of urea and pH on conformation. One-dimensional ¹H NMR spectra of sCon-Vc1 recorded in the presence of increasing concentrations of urea (≤ 7 M) showed subtle changes in the spectrum, indicating the interaction of urea with the peptide. Although subtle chemical shift changes were observed in the one-dimensional ¹H NMR, the chemical shift dispersion was maintained even at 7 M urea, indicating that the peptide was still folded (Figure S4). Limited unfolding of the peptide can be inferred from the appearance of a minor peak from the Trp2 indole proton around 9.6 ppm as the concentration of urea increased (Figure S4C). The degree of unfolding of the peptide was estimated from the relative intensity of this peak. The resulting urea unfolding curve (Figure S4D) shows that <30% of the peptide was unfolded at 7 M urea, confirming the exceptional conformational stability of contryphan-Vc1.

sCon-Vc1 was also examined over the pH range of 2–9. One-dimensional ¹H NMR spectra of sCon-Vc1 showed subtle chemical shift changes in the pH titration between pH 2 and 8, but the overall spectral dispersion was maintained. Most of the amide peaks were broadened at pH 9 because of solvent exchange (Figure S5). The titratable residues Glu24 and His25 are present in the highly flexible and unstructured C-terminal region, and their titrations have no effect on the conformation of the peptide.

Redox Stability. To test the redox stability of the SDH fold, 0.5 mM sCon-Vc1 was subjected to a high concentration of the reducing agent TCEP. Contryphan-Vc2, a seven-residue peptide that also has a single disulfide bridge, was used as control for the reduction assay.¹⁸ Unlike sCon-Vc1, contryphan-Vc2 was completely reduced within 1 h upon incubation with 10 mM TCEP at room temperature (Figure S6A). In contrast, only <5% of sCon-Vc1 was reduced under similar conditions (Figure 4A). The reduction kinetics of sCon-Vc1 were slow at room temperature, and even after being incubated for 48 h, the peptide was not completely reduced (Figure 4A). One-dimensional ¹H NMR spectra and ¹³C HSQC spectra of sCon-Vc1 (0.5 mM) showed no significant changes in the presence of a 12-fold excess of TCEP (6 mM) after incubation for 1 h at 20 °C. After incubation for 16 h with the same concentration of TCEP (Figure 4C and Figure S6C,D), a minor peak at 10.04 ppm appeared, corresponding to the Trp2 indole proton resonance of the reduced species of contryphan-Vc1. Even after incubation for 24 h with 6 mM TCEP at 20 °C, one-dimensional ¹H NMR spectra showed only a slight increase in the reduced species. One-dimensional ¹H NMR spectra recorded on contryphan-Vc1_{1–22}[Z1Q] incubated with 24 mM TCEP showed that the peaks disappeared completely after TCEP incubation for 7 days, showing that the disulfide bond contributes to the stability of the SDH fold (Figure S6E,F).

Proteolytic Stability of Contryphan-Vc1. *In vitro* proteolysis was performed on sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1_{1–22}[Z1Q] utilizing the enzymes trypsin, α -chymotrypsin, and pepsin to assess the proteolytic stability of contryphan-Vc1 and its analogues. In the presence of trypsin, sCon-Vc1 showed accumulation of a peak with a mass of 2464.8 Da, which corresponds to residues 1–22 (cleavage following Arg22) according to LC–MS. Tryptic digestion of rCon-Vc1[Z1Q] also showed accumulation of a peak having a mass of 2481.8 Da, also corresponding to residues 1–22 (the difference in the mass of the fragments between sCon-Vc1 and

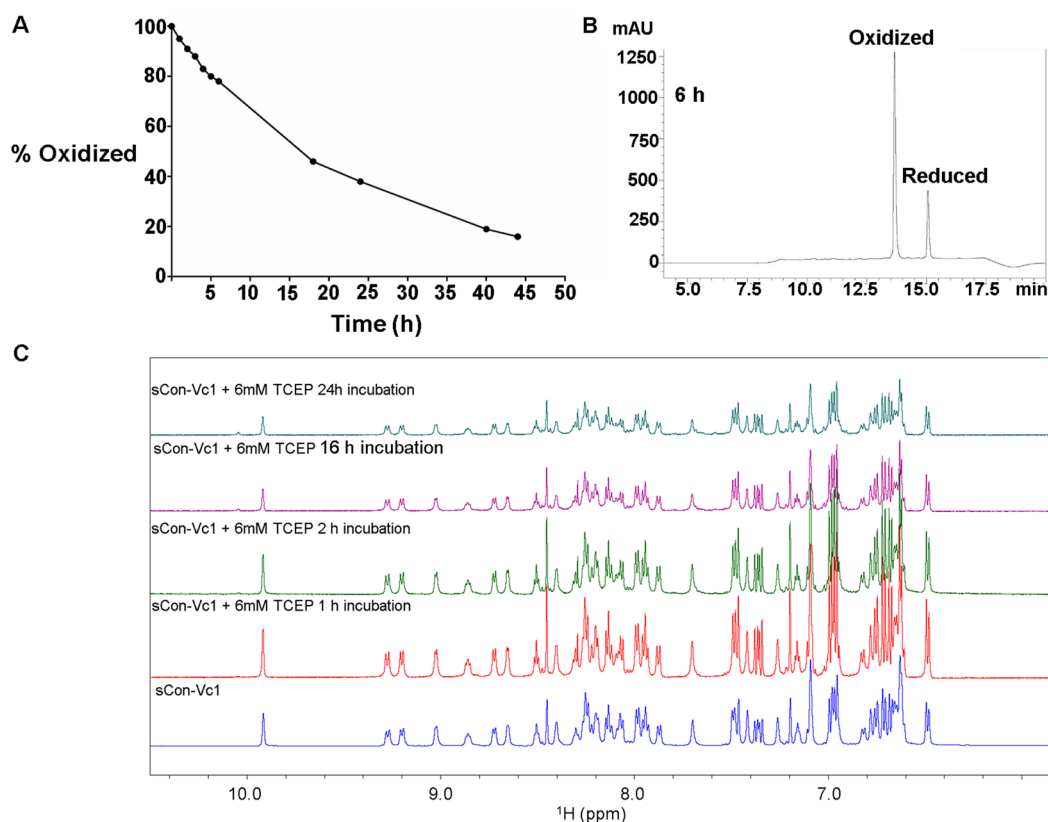


Figure 4. (A) Plot showing the percentage of oxidized species of sCon-Vc1 over the time course of 45 h in the presence of 10 mM TCEP at room temperature. (B) LC–MS chromatogram of sCon-Vc1 at the 6 h time point after incubation with 10 mM TCEP at room temperature, with observed mass differences for oxidized and reduced species of sCon-Vc1 marked. (C) Amide and aromatic regions of one-dimensional ¹H NMR spectra of sCon-Vc1 after incubation for 1 h, overnight, and for 24 h with 6 mM TCEP at pH 4.0 and 20 °C in water containing 7% ²H₂O.

rCon-Vc1[Z1Q] is due to the N-terminal pyroglutamate modification). The pattern of cleavage with trypsin is expected because this protease cleaves the peptide chain mainly C-terminal to lysine and arginine (Figure 5, Figure S7, and Table S2). In contrast, trypsin did not degrade rCon-Vc1_{1–22}[Z1Q] over the observed time range (0.5, 1, 2, and 4 h) (Figure S7C) as the peptide has no tryptic cleavage site. In the presence of α -chymotrypsin, sCon-Vc1 showed accumulation of two major peaks in the LC–MS chromatogram with masses of 2221.5 and 1350.4 Da, which correspond to fragments 1–20 and 21–31, respectively (Figure S7A). Similarly, α -chymotrypsin digestion of rCon-Vc1[Z1Q] also showed accumulation of two major peaks with masses of 2238.6 and 1350.4 Da, also corresponding to fragments 1–20 and 21–31, respectively (Figure S7B). rCon-Vc1_{1–22}[Z1Q] in the presence of α -chymotrypsin shows accumulation of a single major peak with a mass of 2238.6 Da, which corresponds to fragment 1–20 in rCon-Vc1_{1–22}[Z1Q] (Figure 5, Figure S7C, and Table S2). Thus, the major cleavage site for α -chymotrypsin is Leu20 (Figure 5). In contrast, incubation for 1 h with pepsin resulted in significant digestion of sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1_{1–22}[Z1Q] (Figure S7).

Evaluation of a Single Disulfide-Directed β -Hairpin (SDH) Fold as an Acceptor Scaffold. To test whether the SDH fold can accommodate foreign peptide sequences, we replaced loop 1 or 2 of rCon-Vc1_{1–22}[Z1Q] with the DINNN motif of iNOS based on the structural similarity between this epitope bound to its target SPSB proteins and the part of the scaffold it was replacing (Figure 6A,B).^{19–21} ¹H NMR spectra of both sCon-Vc1_{1–22}[Z1Q,DINNN_{4–8}] and sCon-

Vc1_{1–22}[Z1Q,DINNN_{12–16}] showed a significant decrease in the level of peak dispersion compared to that of native rCon-Vc1_{1–22}[Z1Q], indicating that these peptides did not adopt the SDH fold (Figures S8 and S9). Somewhat surprisingly, however, both peptides bound to human SPSB2 with K_D values of 25 ± 8 and 5.7 ± 3 nM, respectively, comparable to the affinity of disulfide-cyclized CVDINNNC (Figure 6 and Table 3).³² Because the majority of the direct interactions between the DINNN motif and SPSB2 are with the three Asn residues, we substituted loop 2 of sCon-Vc1_{1–22} with just the NNN sequence.^{21,33} ¹H NMR spectra of sCon-Vc1_{1–22}[NNN_{12–14}] showed good peak dispersion, and a comparison of backbone chemical shifts with those of sCon-Vc1_{1–22} showed that the peptide maintained the native fold (Figures S8 and S9). The sCon-Vc1_{1–22}[NNN_{12–14}] peptide bound to the human SPSB2 protein with an affinity of 1.3 μ M, which is almost 4-fold weaker than the affinity of the linear DINNN motif for the human SPSB2 protein. However, the binding affinity of sCon-Vc1_{1–22}[NNN_{12–14}] is almost 50-fold higher than that of the linear NNN epitope itself (Figure 6 and Table 3).

DISCUSSION

In this study, we have further characterized the stability of the SDH fold first described in contryphan-Vc1¹⁹ and shown that it can be truncated without loss of ordered structure. Because the molecule is stabilized by just a single disulfide bridge, it was possible to obtain good expression in *E. coli*. We further established that the N-terminal pyroglutamate modification has

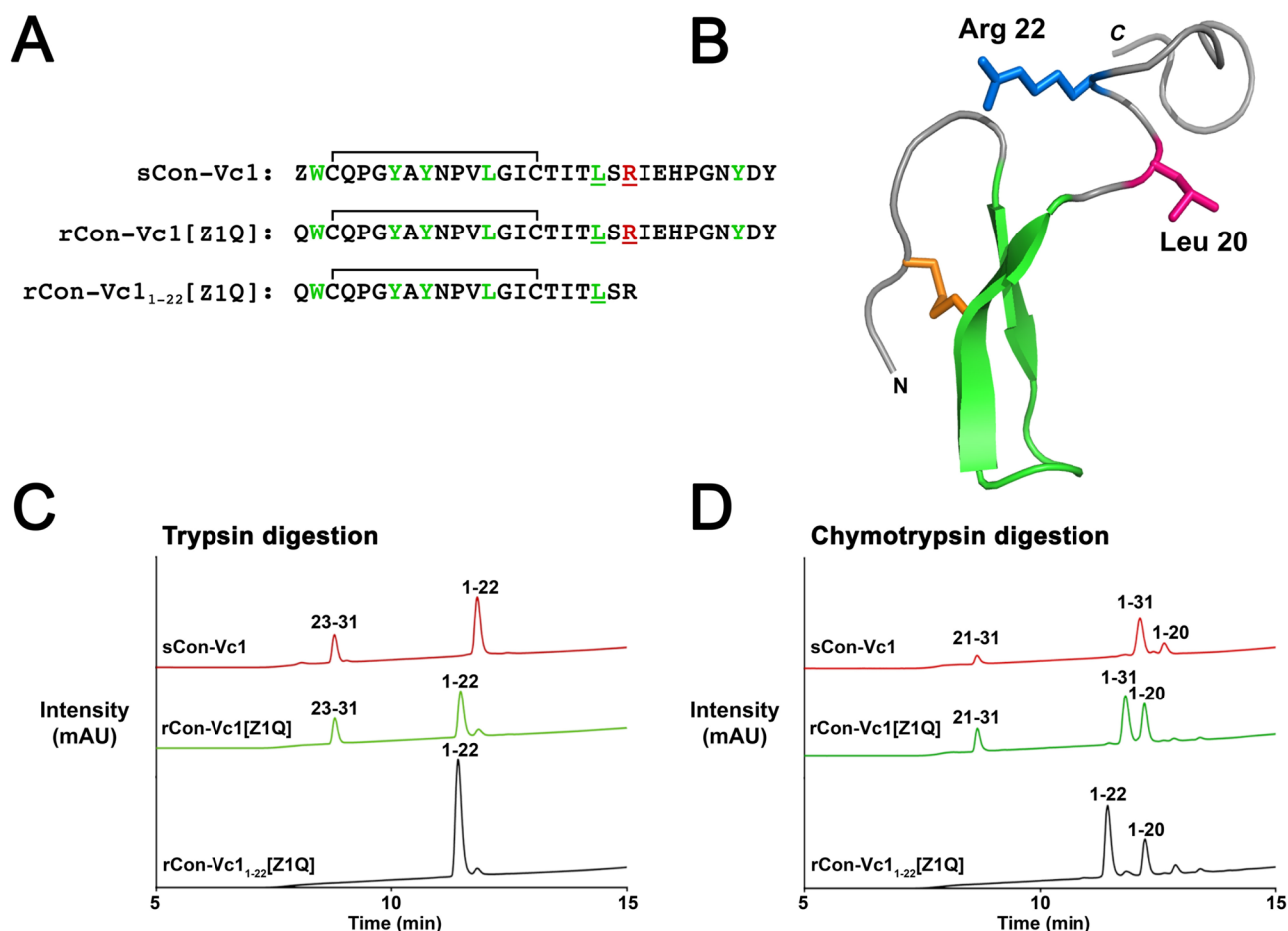


Figure 5. Reversed-phase HPLC analysis after treatment of sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1₁₋₂₂[Z1Q] with trypsin, α -chymotrypsin, and pepsin for 2 h. (A) Sequences of peptides sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1₁₋₂₂[Z1Q] used in the study with expected trypsin and α -chymotrypsin cleavage sites colored red and green, respectively, and the observed trypsin and α -chymotrypsin cleavage sites underlined with red and green lines, respectively. (B) Observed trypsin cleavage site Arg22 and α -chymotrypsin cleavage site Leu20 are mapped on the structure of full-length contryphan-Vc1. (C) Reversed-phase HPLC chromatogram for trypsin digestion of peptides sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1₁₋₂₂[Z1Q]. (D) Reversed-phase HPLC chromatogram for α -chymotrypsin digestion of peptides sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1₁₋₂₂[Z1Q].

a negligible effect on the core structure of the peptide, as monitored by NMR chemical shifts. Backbone ^{15}N relaxation data showed that residues Trp2–Ile18 adopt a well-defined, stable conformation. Residues beyond Thr19 showed either negative or zero ^1H – ^{15}N NOE values and relatively lower R_1 and R_2 values, indicating that the region from residue 19 to 31 is highly disordered, with the flexibility increasing toward the C-terminus.

Contryphan-Vc1 exhibited remarkable thermal stability.¹⁹ The secondary structure monitored by circular dichroism spectroscopy did not change even at 95 °C, and one-dimensional NMR spectra showed no change in peak dispersion up to 70 °C (the highest temperature examined).¹⁹ To further assess the conformational stability of contryphan-Vc1, we investigated the effects of urea, pH, and reducing agents on its structure, as monitored by NMR. The peptide exhibited remarkable chemical stability in the presence of urea, showing <30% unfolding even at 7 M urea (Figure S4). Contryphan-Vc1 is also quite stable over a broad pH range, with one-dimensional NMR spectra showing no significant changes to the core structure between pH 2 and 8 (Figure S5).

One of the potential limitations of peptides having disulfide bridges as therapeutic leads is their redox stability. Most cellular compartments are reducing environments,³⁴ and in the

circulation and extracellular space, thiol–disulfide exchange reactions³⁵ are performed by different members of the protein disulfide isomerase family.³⁶ Disulfide bridges are usually unstable in reducing environments and are susceptible to disulfide bond exchange reactions with biological thiols such as glutathione.^{37,38} This can partially or completely inactivate peptides that have disulfide bridges,^{39,40} rendering them more susceptible to proteolytic cleavage.³⁷ With less stable disulfide-cyclized peptides such as Ac-c[CVDINNNC]-NH₂,⁴¹ which was shown to be reduced within 0.5 h,⁴¹ and contryphan-Vc2¹⁸ [used in this study (Figure S6)], 5 mM TCEP at pH 4 was sufficient to reduce these peptides within 1 h at room temperature. In contrast, to completely reduce a highly stable peptide such as sCon-Vc1, incubation with 10 mM TCEP (pH 4) for more than 2 days at room temperature was required, emphasizing the remarkable redox stability of contryphan-Vc1.

Both the calculated structures and the backbone relaxation parameters concur in showing that the N-terminal region of contryphan-Vc1 up to Thr19 adopts an ordered structure while the C-terminal region from residues Thr19–Tyr31 is flexible. To identify the minimal region sufficient to maintain the structure and to analyze the influence of the highly flexible C-terminus, a truncated variant of the peptide encompassing residues 1–22 was generated. Even though residues Thr19–

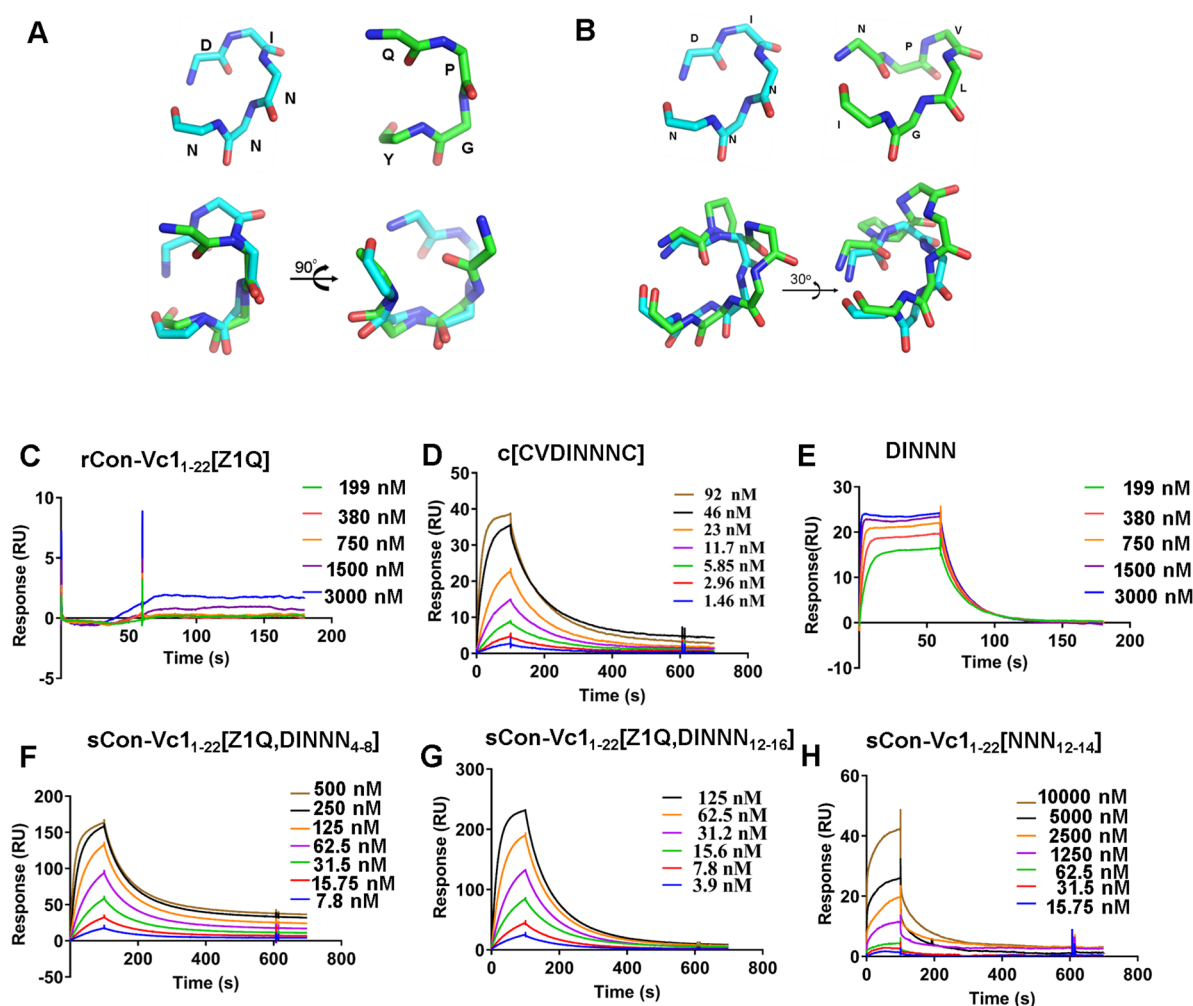


Figure 6. (A) PyMOL representation of the alignment of the DINNN motif (cyan) bound to SPSB2 (3EMW)²¹ and loop 1 (QPGY) (green) of rCon-Vc11–22[Z1Q]. (B) PyMOL representation of the alignment of the DINNN motif (cyan) bound to SPSB2 (3EMW)²¹ and loop 2 (PVLG) (green) of rCon-Vc11–22[Z1Q]. SPR sensorgrams of the interactions between immobilized human SPSB2 and the Vc1-DINNN analogues: (C) rCon-Vc11–22[Z1Q], (D) cystathionine-cyclized c[CVDINNNC], (E) linear DINNN, (F) sCon-Vc11–22[Z1Q,DINNN_{4–8}], (G) sCon-Vc11–22[Z1Q,DINNN_{12–16}], and (H) sCon-Vc11–22[NNN_{12–14}].

Table 3. Association Rates (k_a), Dissociation Rates (k_d), and Binding Affinities (K_D) for the Vc1-DINNN Peptide Analogues and Controls

peptide	k_a ($M^{-1} s^{-1}$)	k_d (s^{-1})	K_D (nM)
rCon-Vc11–22[Z1Q]	no binding	no binding	no binding
sCon-Vc11–22[NNN _{12–14}]	5.1×10^{-3}	0.007	1300
sCon-Vc11–22[Z1Q,DINNN _{4–8}]	7.9×10^{-5}	0.00064	25 ± 8
sCon-Vc11–22[Z1Q,DINNN _{12–16}]	5.7×10^{-5}	0.002695	5.7 ± 3
DINNN	1.1×10^{-6}	0.12	350
Ac-[CVDINNNC]	4.7×10^{-5}	0.0074	10
NNN ³³	<i>a</i>	<i>a</i>	74.6

^aAssociation and dissociation rates were too rapid to be determined.

Arg22 are flexible, with negative or zero 1H – ^{15}N NOE values, truncation was performed at Arg22 with a view to retaining polar residues Thr19 and Ser21 and the charged Arg22, to maintain aqueous solubility. The structure of rCon-Vc11–22[Z1Q] revealed that the peptide retained its structure even after deletion of the C-terminal tail region (Figure 3).

A key attribute of a peptide scaffold is its resistance to proteolytic degradation.⁴² Proteolysis of full-length contryphan-

Vc1 with trypsin yielded a major cleavage product, corresponding to cleavage at Arg22. In contrast, the truncated peptide rCon-Vc11–22[Z1Q], which lacks the tryptic cleavage site, is highly resistant to trypsin cleavage. Both synthetic and recombinant contryphan-Vc1 full-length peptides showed two cleavage products upon digestion with chymotrypsin, indicating that these peptides are cleaved at a major cleavage site, identified as Leu20 (Figure 5). Cleavage of rCon-Vc11–22[Z1Q] with chymotrypsin yielded a single major cleavage product corresponding to fragment 1–20. Observation of a single cleavage site for contryphan-Vc1, even though there are several residues in the core β -sheet structure that can serve as cleavage sites for chymotrypsin, indicates that this structure imparts resistance to proteolysis. In contrast, both full-length and truncated peptides were found to be susceptible to pepsin cleavage, with both peptides being digested completely by this protease.

Even some peptides with the highly stable ICK fold and multiple disulfide bonds are susceptible to proteolytic degradation. Derivatives of Agouti-related protein and two squash protease inhibitors were stable to pepsin and elastase but were extensively degraded by trypsin and chymotrypsin.⁴³

Similarly constrained peptides, including ω -conotoxin MVIIA and a hybrid of two squash trypsin inhibitors (EETI-II and MCoTI-II), were degraded by trypsin and chymotrypsin.^{44,45} Multiple strategies for increasing the proteolytic stability of the peptides are emerging, including backbone cyclization,^{46,47} where the N- and C-termini of the linear peptide are linked using peptide coupling agents.⁴⁸ Depending on the target of an engineered analogue of contryphan-Vc1, individual residues that are sites of proteolysis, for example, Leu20 in the case of chymotrypsin, could be replaced to enhance resistance.

To assess the capacity of the SDH fold to incorporate foreign peptide sequences and present them in a functional manner, the loops in the core structure of contryphan-Vc1 were replaced with a five-residue DINNN epitope from iNOS that interacts with SPSB proteins and mediates its proteolytic degradation.^{21,49} In the sCon-Vc1_{1–22}[Z1Q,DINNN_{4–8}] analogue, the four residues comprising loop 1 of rCon-Vc1_{1–22}[Z1Q] (QPGY) were replaced by the DINNN motif, while the three residues in loop 2 (VLG) were substituted with the DINNN motif in the sCon-Vc1_{1–22}[Z1Q,DINNN_{12–16}] analogue. Proline is preferred at some positions of β -turns and has been shown to increase their stability;^{50,51} hence, Pro11 was retained in sCon-Vc1_{1–22}[Z1Q,DINNN_{12–16}]. Even though both Vc1-DINNN analogues bound to SPSB2 with affinities of 25 ± 8 and 5.7 ± 3 nM, respectively (Figure 6), ¹H NMR spectra of both peptides showed poor peak dispersion, indicating that substituting loop 1 and loop 2 residues with the entire DINNN motif disrupted the native fold (Figures S8 and S9). While this indicates that there are limitations on the sequence variation that can be accommodated in either of these loops, the retention of the high binding affinity for SPSB2 implies that the incorporated DINNN sequences can adopt a conformation that is recognized by SPSB2 even as part of this larger, partially constrained peptide.²¹

In an effort to maintain the structure of the turn in the β -hairpin, we replaced the VLG residues of loop 2 with a shorter three-residue NNN epitope, noting that the majority of the direct contacts with SPSB2 are mediated by the three Asn residues.³³ The sCon-Vc1_{1–22}[NNN_{12–14}] analogue bound to human SPSB2 with an affinity of 1.3 μ M, which is almost 50-fold stronger than that of the linear NNN epitope itself (Figure 6 and Table 3).³³ Moreover, ¹H NMR spectra of this analogue showed good peak dispersion and native-like chemical shifts, indicating that the native fold was maintained. The binding affinity for SPSB2 implies that the NNN motif is constrained appropriately for binding to SPSB2, but the absence of the Asp residue of the original DINNN motif precludes tighter binding.

In the study presented here, important structural, dynamic, and stability properties of the recently discovered contryphan-Vc1 have been evaluated. Contryphan-Vc1 offers a significant advantage in that it contains just a single disulfide bond, while exhibiting remarkable thermal, redox, and chemical stability and reasonable proteolytic stability. Contryphan-Vc1 can accept the NNN sequence without compromising the core structure of the peptide, although more substantial insertions did affect the structure. Further exploration of the capacity of rCon-Vc1_{1–22}[Z1Q] to accommodate other functional epitopes is clearly warranted.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00120.

Chemical shifts for truncated contryphan-Vc1 (rCon-Vc1_{1–22}[Z1Q]) at pH 3.9 and 293 K (Table S1), fragments and their masses observed after proteolytic digestion of sCon-Vc1, rCon-Vc1[Z1Q], and rCon-Vc1_{1–22}[Z1Q] with the enzymes trypsin and α -chymotrypsin (Table S2), RP-HPLC chromatogram and ESI-MS results of the sCon-Vc1 peptide (Figure S1), LC-MS profile of recombinant full-length contryphan-Vc1 and rCon-Vc1_{1–22}[Z1Q] (Figure S2), LC-MS profile of contryphan-Vc1 analogues (Figure S3), one-dimensional ¹H NMR spectra of synthetic full-length contryphan-Vc1 titrated with increasing concentrations of urea and the urea unfolding curve for the Trp2 indole proton peak of sCon-Vc1 (Figure S4), one-dimensional ¹H NMR spectra of synthetic full-length contryphan-Vc1 over the pH range of 2–9 (Figure S5), LC-MS chromatograms, ¹H NMR spectra, and ¹³C HSQC spectra of sCon-Vc1 incubated with TCEP (Figure S6), reversed-phase HPLC analysis after treatment of synthetic, recombinant full-length contryphan-Vc1 and rCon-Vc1_{1–22}[Z1Q] with trypsin, α -chymotrypsin, and pepsin (Figure S7), one-dimensional ¹H NMR spectra of rCon-Vc1_{1–22}[Z1Q], sCon-Vc1_{1–22}[NNN_{12–14}], sCon-Vc1_{1–22}[Z1Q,DINNN_{4–8}], and sCon-Vc1_{1–22}[Z1Q,DINNN_{12–16}] (Figure S8), and plots of the deviation of the chemical shifts from random coil values for Vc1_{1–22}[Z1Q,DINNN_{4–8}] and sCon-Vc1_{1–22}[NNN_{12–14}] (Figure S9) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville 3052, Australia. Telephone: (+61 3) 9903 9167. E-mail: ray.norton@monash.edu.

ORCID

Balasubramanyam Chittoor: 0000-0001-5460-4537

Raymond S. Norton: 0000-0001-8893-0584

Present Address

[§]S.D.R.: Department of Biology, University of Utah, Salt Lake City, UT 84112.

Funding

R.S.N. acknowledges fellowship support from the National Health and Medical Research Council of Australia. B.C. is supported by a Faculty of Pharmacy and Pharmaceutical Sciences Scholarship, and M.S. is supported by a Monash Graduate Scholarship.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Biswarajan Mohanty (Monash Institute of Pharmaceutical Sciences, Monash University) for helpful discussions regarding NMR relaxation studies.

■ ABBREVIATIONS

HSQC, heteronuclear single-quantum coherence; ICK, inhibitor cystine knot; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; RMSD, root-mean-square deviation; rCon-Vc1[Z1Q], recombinant contryphan-Vc1; rCon-Vc1_{1–22}[Z1Q], recombinant truncated contryphan-Vc1; RP-HPLC, reversed-phase high-performance liquid chromatography; RMSD, root-mean-square deviation; sCon-Vc1, synthetic contryphan-Vc1; TCEP, tris(2-carboxyethyl)phosphine.

■ REFERENCES

- (1) Wang, C. K., Gruber, C. W., Cemazar, M., Siatskas, C., Tagore, P., Payne, N., Sun, G., Wang, S., Bernard, C. C., and Craik, D. J. (2014) Molecular grafting onto a stable framework yields novel cyclic peptides for the treatment of multiple sclerosis. *ACS Chem. Biol.* 9, 156–163.
- (2) Kimura, R. H., Cheng, Z., Gambhir, S. S., and Cochran, J. R. (2009) Engineered knottin peptides: a new class of agents for imaging integrin expression in living subjects. *Cancer Res.* 69, 2435–2442.
- (3) Kimura, R. H., Miao, Z., Cheng, Z., Gambhir, S. S., and Cochran, J. R. (2010) A dual-labeled knottin peptide for PET and near-infrared fluorescence imaging of integrin expression in living subjects. *Bioconjugate Chem.* 21, 436–444.
- (4) Miao, Z., Ren, G., Liu, H., Kimura, R. H., Jiang, L., Cochran, J. R., Gambhir, S. S., and Cheng, Z. (2009) An engineered knottin peptide labeled with ¹⁸F for PET imaging of integrin expression. *Bioconjugate Chem.* 20, 2342–2347.
- (5) Terlau, H., and Olivera, B. M. (2004) Conus venoms: a rich source of novel ion channel-targeted peptides. *Physiol. Rev.* 84, 41–68.
- (6) Vita, C., Vizzavona, J., Drakopoulou, E., Zinn-Justin, S., Gilquin, B., and Ménez, A. (1998) Novel miniproteins engineered by the transfer of active sites to small natural scaffolds. *Biopolymers* 47, 93–100.
- (7) Wong, C. T., Rowlands, D. K., Wong, C. H., Lo, T. W., Nguyen, G. K., Li, H. Y., and Tam, J. P. (2012) Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chem., Int. Ed.* 51, 5620–5624.
- (8) Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) A common structural motif incorporating a cystine knot and a triple-stranded β -sheet in toxic and inhibitory polypeptides. *Protein Sci.* 3, 1833–1839.
- (9) Norton, R. S., and Pallaghy, P. K. (1998) The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* 36, 1573–1583.
- (10) Craik, D. J., Daly, N. L., and Waite, C. (2001) The cystine knot motif in toxins and implications for drug design. *Toxicon* 39, 43–60.
- (11) Norton, R. S., and McDonough, S. I. (2008) Peptides targeting voltage-gated calcium channels. *Curr. Pharm. Des.* 14, 2480–2491.
- (12) Moore, S. J., Leung, C. L., Norton, H. K., and Cochran, J. R. (2013) Engineering agatoxin, a cystine-knot peptide from spider venom, as a molecular probe for *in vivo* tumor imaging. *PLoS One* 8, e60498.
- (13) Jiang, H., Moore, S. J., Liu, S., Liu, H., Miao, Z., Cochran, F. V., Liu, Y., Tian, M., Cochran, J. R., Zhang, H., and Cheng, Z. (2013) A novel radiofluorinated agouti-related protein for tumor angiogenesis imaging. *Amino Acids* 44, 673–681.
- (14) Hannig, G., and Makrides, S. C. (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol.* 16, 54–60.
- (15) Klint, J. K., Senff, S., Saez, N. J., Seshadri, R., Lau, H. Y., Bende, N. S., Undheim, E. A., Rash, L. D., Mobli, M., and King, G. F. (2013) Production of recombinant disulfide-rich venom peptides for structural and functional analysis via expression in the periplasm of *E. coli*. *PLoS One* 8, e63865.
- (16) Salinas, G., Pellizza, L., Margenat, M., Fló, M., and Fernández, C. (2011) Tuned *Escherichia coli* as a host for the expression of disulfide-rich proteins. *Biotechnol. J.* 6, 686–699.

- (17) Robinson, S. D., and Norton, R. S. (2014) Conotoxin gene superfamilies. *Mar. Drugs* 12, 6058–6101.
- (18) Robinson, S. D., Safavi-Hemami, H., McIntosh, L. D., Purcell, A. W., Norton, R. S., and Papenfuss, A. T. (2014) Diversity of conotoxin gene superfamilies in the venomous snail, *Conus victoriae*. *PLoS One* 9, e87648.
- (19) Robinson, S. D., Chhabra, S., Belgi, A., Chittoor, B., Safavi-Hemami, H., Robinson, A. J., Papenfuss, A. T., Purcell, A. W., and Norton, R. S. (2016) A naturally occurring peptide with an elementary single disulfide-directed β -hairpin fold. *Structure* 24, 293–299.
- (20) Kuang, Z., Lewis, R. S., Curtis, J. M., Zhan, Y., Saunders, B. M., Babon, J. J., Kolesnik, T. B., Low, A., Masters, S. L., Willson, T. A., Kedzierski, L., Yao, S., Handman, E., Norton, R. S., and Nicholson, S. E. (2010) The SPRY domain-containing SOCS box protein SPSB2 targets iNOS for proteasomal degradation. *J. Cell Biol.* 190, 129–141.
- (21) Filippakopoulos, P., Low, A., Sharpe, T. D., Uppenberg, J., Yao, S., Kuang, Z., Savitsky, P., Lewis, R. S., Nicholson, S. E., Norton, R. S., and Bullock, A. N. (2010) Structural basis for Par-4 recognition by the SPRY domain- and SOCS box-containing proteins SPSB1, SPSB2, and SPSB4. *J. Mol. Biol.* 401, 389–402.
- (22) Lewis, R. S., Kolesnik, T. B., Kuang, Z., D’Cruz, A. A., Blewitt, M. E., Masters, S. L., Low, A., Willson, T., Norton, R. S., and Nicholson, S. E. (2011) TLR regulation of SPSB1 controls inducible nitric oxide synthase induction. *J. Immunol.* 187, 3798–3805.
- (23) Shaka, A., Lee, C., and Pines, A. (1988) Iterative schemes for bilinear operators; application to spin decoupling. *J. Magn. Reson.* 77, 274–293.
- (24) Hwang, T.-L., and Shaka, A. (1995) Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J. Magn. Reson., Ser. A* 112, 275–279.
- (25) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct., Funct., Genet.* 59, 687–696.
- (26) Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Kent Wenger, R., Yao, H., and Markley, J. L. (2008) BioMagResBank. *Nucleic Acids Res.* 36, D402–D408.
- (27) Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Protein NMR Techniques* 278, 353–378.
- (28) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Marius Clore, G. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* 160, 65–73.
- (29) Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51–55.
- (30) Keller, R. L. J. (2005) Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. Ph.D. Thesis, ETH, Zurich.
- (31) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* 5, 67–81.
- (32) Yap, B. K., Leung, E. W., Yagi, H., Galea, C. A., Chhabra, S., Chalmers, D. K., Nicholson, S. E., Thompson, P. E., and Norton, R. S. (2014) A potent cyclic peptide targeting SPSB2 protein as a potential anti-infective agent. *J. Med. Chem.* 57, 7006–7015.
- (33) Leung, E. W. W., Mulcair, M. D., Yap, B. K., Nicholson, S. E., Scanlon, M. J., and Norton, R. S. (2017) Molecular insights into the interaction between the SPRY domain-containing SOCS box protein SPSB2 and peptides based on the binding motif from iNOS. *Aust. J. Chem.* 70, 191–200.
- (34) Go, Y. M., and Jones, D. P. (2008) Redox compartmentalization in eukaryotic cells. *Biochim. Biophys. Acta, Gen. Subj.* 1780, 1273–1290.
- (35) Kozlov, G., Määttänen, P., Thomas, D. Y., and Gehring, K. (2010) A structural overview of the PDI family of proteins. *FEBS J.* 277, 3924–3936.

- (36) Butera, D., Cook, K. M., Chiu, J., Wong, J. W. H., and Hogg, P. J. (2014) Control of blood proteins by functional disulfide bonds. *Blood* 123, 2000–2007.
- (37) Armishaw, C. J., Daly, N. L., Nevin, S. T., Adams, D. J., Craik, D. J., and Alewood, P. F. (2006) α -Selenoconotoxins, a new class of potent $\alpha 7$ neuronal nicotinic receptor antagonists. *J. Biol. Chem.* 281, 14136–14143.
- (38) Aslund, F., and Beckwith, J. (1999) Bridge over troubled waters: sensing stress by disulfide bond formation. *Cell* 96, 751–753.
- (39) Pennington, M. W., Lanigan, M. D., Kalman, K., Mahnir, V. M., Rauer, H., McVaugh, C. T., Behm, D., Donaldson, D., Chandry, K. G., Kem, W. R., and Norton, R. S. (1999) Role of disulfide bonds in the structure and potassium channel blocking activity of ShK toxin. *Biochemistry* 38, 14549–14558.
- (40) Flinn, J. P., Pallaghy, P. K., Lew, M. J., Murphy, R., Angus, J. A., and Norton, R. S. (1999) Role of disulfide bridges in the folding, structure and biological activity of ω -conotoxin GVIA1. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* 1434, 177–190.
- (41) Yap, B. K., Harjani, J. R., Leung, E. W. W., Nicholson, S. E., Scanlon, M. J., Chalmers, D. K., Thompson, P. E., Baell, J. B., and Norton, R. S. (2016) Redox-stable cyclic peptide inhibitors of the SPSB2–iNOS interaction. *FEBS Lett.* 590, 696–704.
- (42) Craik, D. J., Fairlie, D. P., Liras, S., and Price, D. (2013) The future of peptide-based drugs. *Chem. Biol. Drug Des.* 81, 136–147.
- (43) Werle, M., Schmitz, T., Huang, H. L., Wentzel, A., Kolmar, H., and Bernkop-Schnurch, A. (2006) The potential of cystine-knot microproteins as novel pharmacophoric scaffolds in oral peptide drug delivery. *J. Drug Target* 14, 137–146.
- (44) Zheng, K., Lubman, D. M., Rossi, D. T., Nordblom, G. D., and Barksdale, C. M. (2000) Elucidation of peptide metabolism by on-line immunoaffinity liquid chromatography mass spectrometry. *Rapid Commun. Mass Spectrom.* 14, 261–269.
- (45) Werle, M., Kafedjiiski, K., Kolmar, H., and Bernkop-Schnurch, A. (2007) Evaluation and improvement of the properties of the novel cystine-knot microprotein McoEeTI for oral administration. *Int. J. Pharm.* 332, 72–79.
- (46) Borchardt, R. T. (1999) Optimizing oral absorption of peptides using prodrug strategies. *J. Controlled Release* 62, 231–238.
- (47) Li, P., and Roller, P. P. (2002) Cyclization strategies in peptide derived drug design. *Curr. Top. Med. Chem.* 2, 325–341.
- (48) Clark, R. J., Fischer, H., Dempster, L., Daly, N. L., Rosengren, K. J., Nevin, S. T., Meunier, F. A., Adams, D. J., and Craik, D. J. (2005) Engineering stable peptide toxins by means of backbone cyclization: stabilization of the α -conotoxin MII. *Proc. Natl. Acad. Sci. U. S. A.* 102, 13767–13772.
- (49) Kuang, Z., Lewis, R. S., Curtis, J. M., Zhan, Y., Saunders, B. M., Babon, J. J., Kolesnik, T. B., Low, A., Masters, S. L., Willson, T. A., Kedzierski, L., Yao, S., Handman, E., Norton, R. S., and Nicholson, S. E. (2010) The SPRY domain-containing SOCS box protein SPSB2 targets iNOS for proteasomal degradation. *J. Cell Biol.* 190, 129–141.
- (50) Fu, H., Grimsley, G. R., Razvi, A., Scholtz, J. M., and Pace, C. N. (2009) Increasing protein stability by improving β -turns. *Proteins: Struct., Funct., Genet.* 77, 491–498.
- (51) Watanabe, K., Masuda, T., Ohashi, H., Mihara, H., and Suzuki, Y. (1994) Multiple proline substitutions cumulatively thermostabilize *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase. *Eur. J. Biochem.* 226, 277–283.

SUPPLEMENTARY INFORMATION

The Single Disulfide-Directed β -Hairpin Fold. Dynamics, Stability and Engineering

Balasubramanyam Chittoor¹, Bankala Krishnarjuna¹, Rodrigo A. V. Morales¹, Christopher A. MacRaid¹, Maiada Sadek¹, Eleanor W. W. Leung¹, Samuel D. Robinson^{1#}, Michael W. Pennington² and Raymond S. Norton^{1*}

¹ Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia.

² Peptides International, Louisville, Kentucky 40299

Current address: Department of Biology, University of Utah, Salt Lake City, UT 84112, USA.

* Corresponding Author: E-mail: ray.norton@monash.edu

Table S1. Chemical shifts for truncated contryphan-Vc1 (rCon-Vc1₁₋₂₂[Z1Q]) at pH 3.9, 293K

Residue	H ^N	H ^α	H ^β	N	C ^α	C ^β	other
Gln1	NA	3.8	1.53,1.65	NA	52.6	27.1	H ^γ 2.07, H ^ε 6.91
Trp2	8.75	4.5	3.17	125.9	58.4	30.0	H ^δ 1 7.23; H ^ε 1 10.01; H ^ε 3 7.43; H ^ε 2 7.43; H ^ε 3 7.04,7.45; H ^η 2 7.20
Cys3	7.8	4.69	2.92,2.39	121.3	51.8	42.1	
Gln4	8.65	4.04	1.96,1.77	121.4	54.5	27.2	H ^γ 2.17,2.05 H ^ε 6.72,7.12
Pro5	-	4.37	2.37,1.91	-	64.6	31.7	H ^γ 2.08, 2.20; H ^δ 3.69, 3.83
Gly6	8.9	4.35,3.62	-	112.3	45.1		
Tyr7	8.31	5.05	3.35,2.39	120.1	56.4	41.6	H ^δ 6.70; H ^ε 6.69
Ala8	9.25	4.65	1.35	121.3	51.1	22.3	
Tyr9	9.08	4.31	2.80,3.02	124.6	59.7	39.1	H ^δ 6.85; H ^ε 6.57
Asn10	8.17	5.08	3.00,2.80	128.5	49.3	39.5	H ^δ 2 7.36, 7.76
Pro11	-	4.03	2.41	-	64.1	32.3	H ^γ 2.05, 2.11; H ^δ 3.90
Val12	7.57	3.75	2.11	119.5	65.4	32.1	H ^γ a 0.90; H ^γ b 1.00
Leu13	6.95	4.28	1.29	116.7	54.8	44.2	H ^γ 1.52 H ^δ a 0.82; H ^δ b 0.86
Gly14	8.4	3.96,3.73	-	108.5	46.3		

Ile15	6.77	4.85	1.91	109.8	57.5	42.2	H ^γ 1 1.172; H ^γ 2 0.89; H ^δ 1 0.82
Cys16	8.86	5.15	3.10,2.77	120.8	56.2	42.9	
Thr17	9.35	4.85	4.28	116.6		71.6	H ^γ 2 1.20
Ile18	8.42	3.43	1.3	124.5	62.2	38.9	H ^γ 1 0.79, 1.06; H ^γ 2 0.69; H ^δ 1 0.67
Thr19	8.19	4.29	3.96	121.3	61.7	69.6	H ^γ 2 1.07
Leu20	8.31	4.38	-	126.2	54.9	42.6	H ^γ 1.59; H ^δ a 0.83; H ^δ b 0.90
Ser21	8.33	4.42	3.83	117.8	58.3	63.8	
Arg22	8.06	4.21	1.72,1.86	127.6	57.4	31.5	H ^γ 1.58; H ^δ 3.18; H ^ε 7.18

30124 - BMRB accession number for the chemical shifts of rCon-Vc1₁₋₂₂[Z1Q]

Table S2. Fragments and their masses observed after proteolytic digestion of sCon-Vc1, rCon-Vc1[Z1Q] and rCon-Vc1₁₋₂₂[Z1Q] with the enzymes trypsin and α -chymotrypsin.

Peptide	Fragment sequence	¹ MH+1 (calculated)	¹ MH+1 (observed)
sCon-Vc1 1-31	ZWCQPGYAYNPVLGICTITLSRIEHPGNYDY	3554.7	3554.6
sCon-Vc1 1-22	ZWCQPGYAYNPVLGICTITLSR	2466.2	2464.8
sCon-Vc1 23-31	IEHPGNYDY	1107.5	1107.1
sCon-Vc1 1-20	ZWCQPGYAYNPVLGICTITL	2221.1	2224
sCon-Vc1 21-31	SRIEHPGNYDY	1350.4	1350.4
rCon-Vc1[Z1Q]1-31	QWCQPGYAYNPVLGICTITLSRIEHPGNYDY	3573	3573
rCon-Vc1[Z1Q]1-22	QWCQPGYAYNPVLGICTITLSR	2481.9	2482.2
rCon-Vc1[Z1Q] 23-31	IEHPGNYDY	1107.1	1107.1
rCon-Vc1[Z1Q]1-20	QWCQPGYAYNPVLGICTITL	2238.6	2240
rCon-Vc1[Z1Q] 21-31	SRIEHPGNYDY	1350.4	1350.4
rCon-Vc1 ₁₋₂₂ [Z1Q]	QWCQPGYAYNPVLGICTITLSR	2481.9	2482.2
rCon-Vc1 ₁₋₂₂ [Z1Q]	QWCQPGYAYNPVLGICTITL	2238.6	2240

¹ Mono Isotopic mass

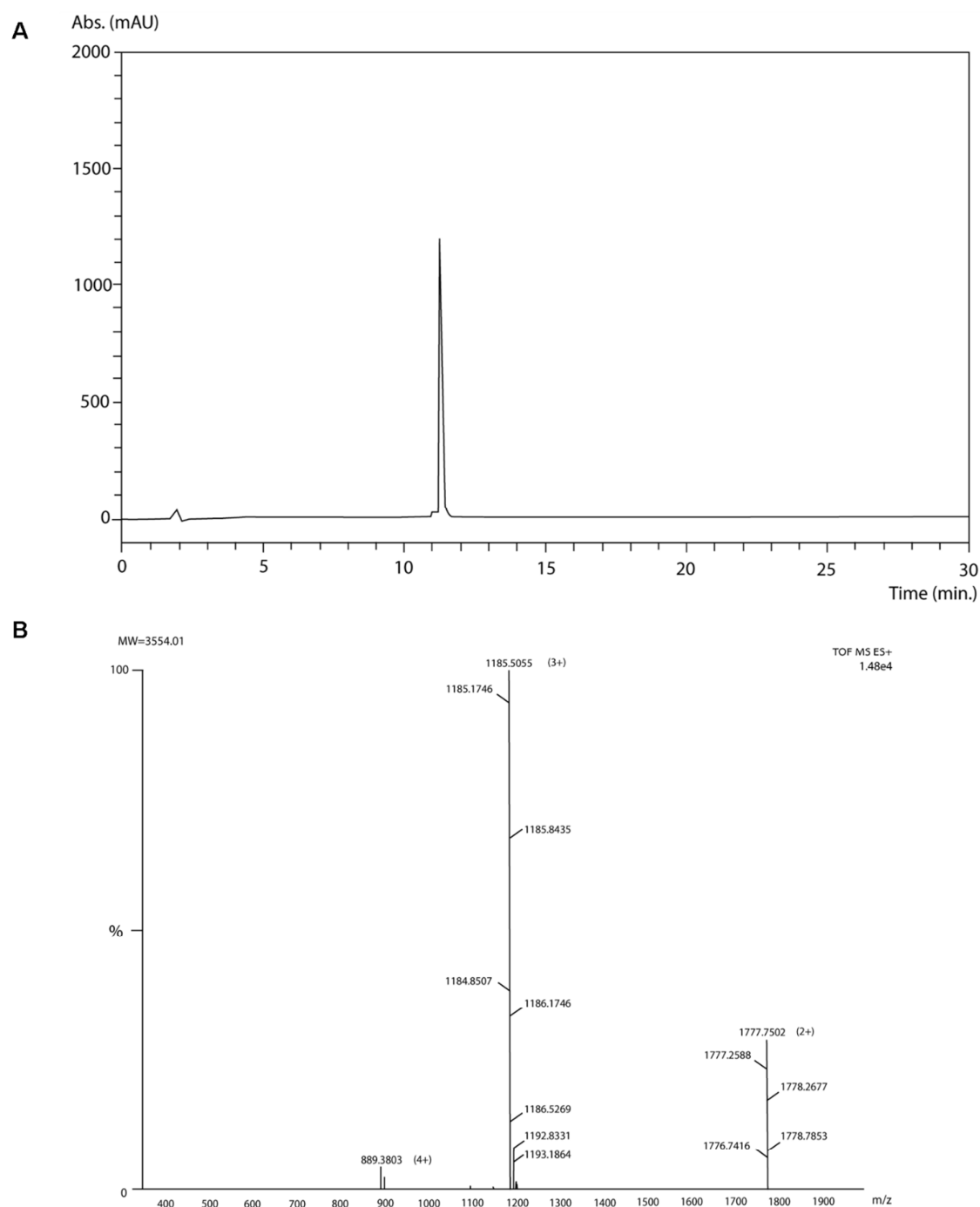


Figure S1. A. RP-HPLC chromatogram of sCon-Vc1 using an ODS-silica column gradient from 10-70% B in 30 min at 1 mL/min. Absorbance at 220 nm. **B.** ESI-MS results for purified sCon-Vc1 peptide. Molecular ion peaks of MH^{+2} , MH^{+3} and MH^{+4} , indicating a molecular mass of 3554.01.

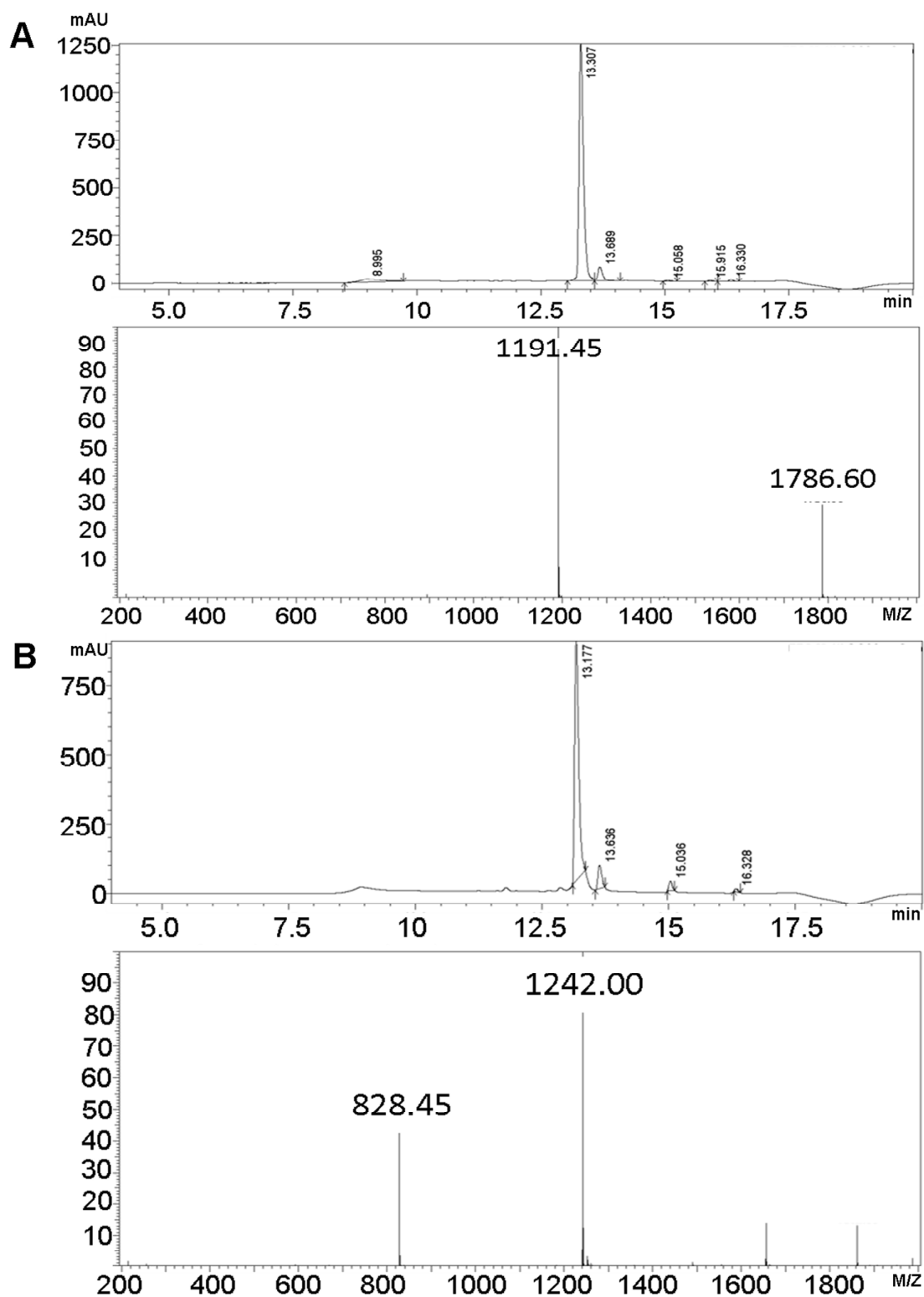
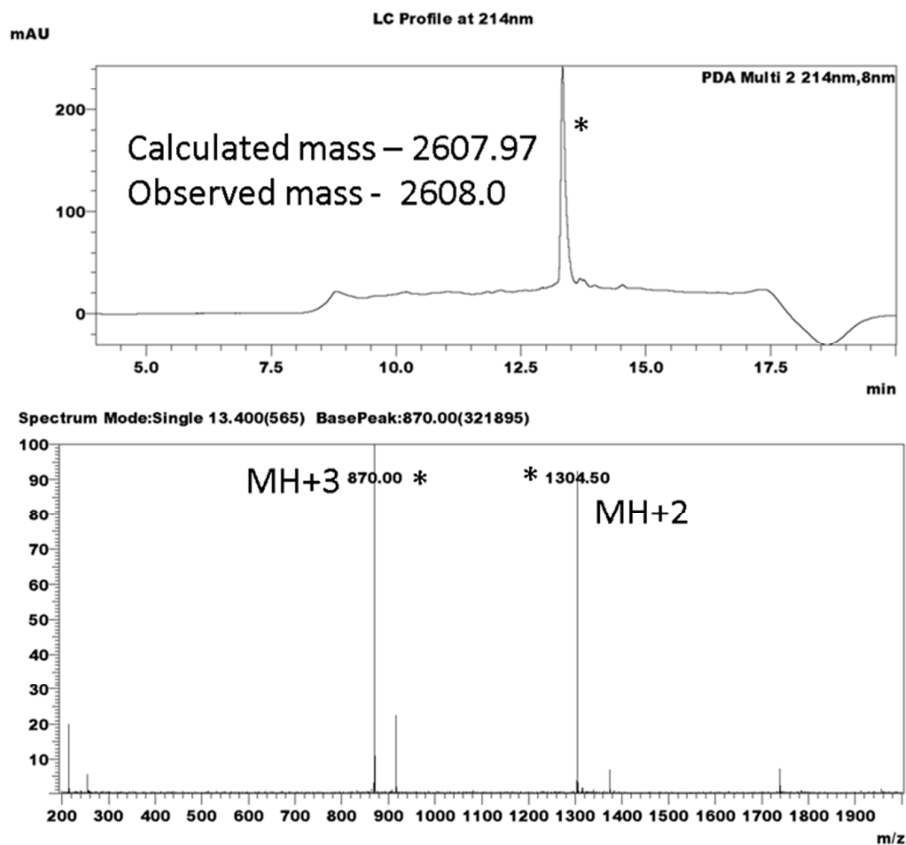
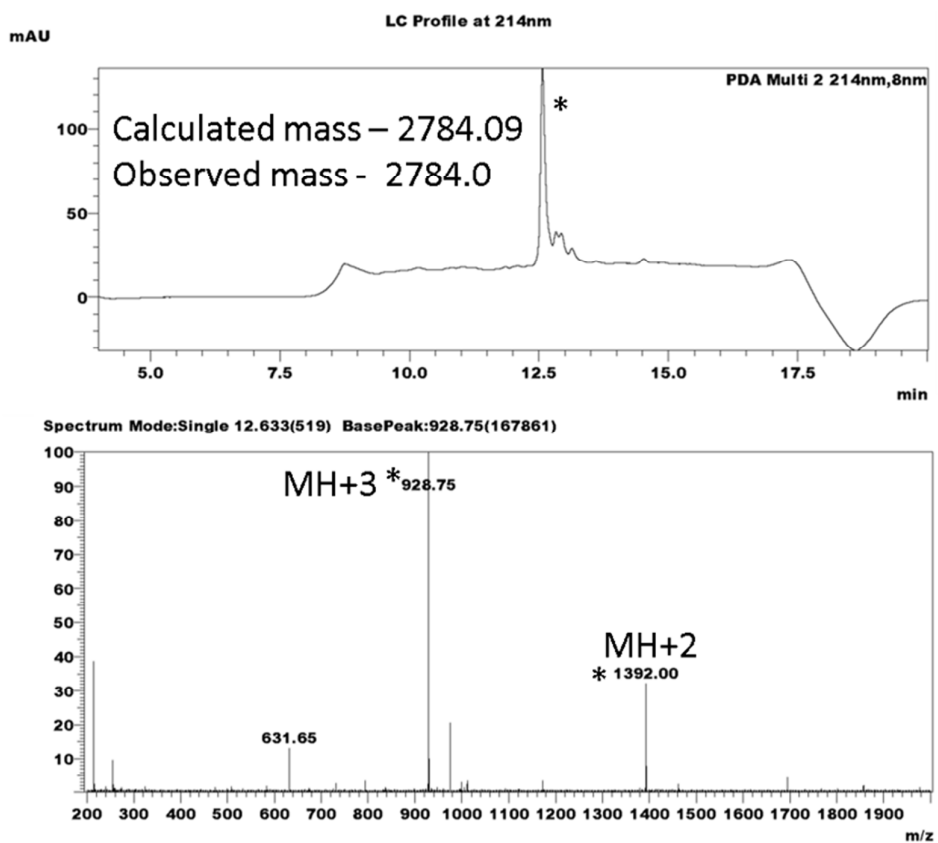


Figure S2. **A.** LC-MS profile of recombinant full-length contryphan-Vc1. **B.** LC-MS profile of recombinant rCon-Vc1₁₋₂₂[Z1Q].

A



B



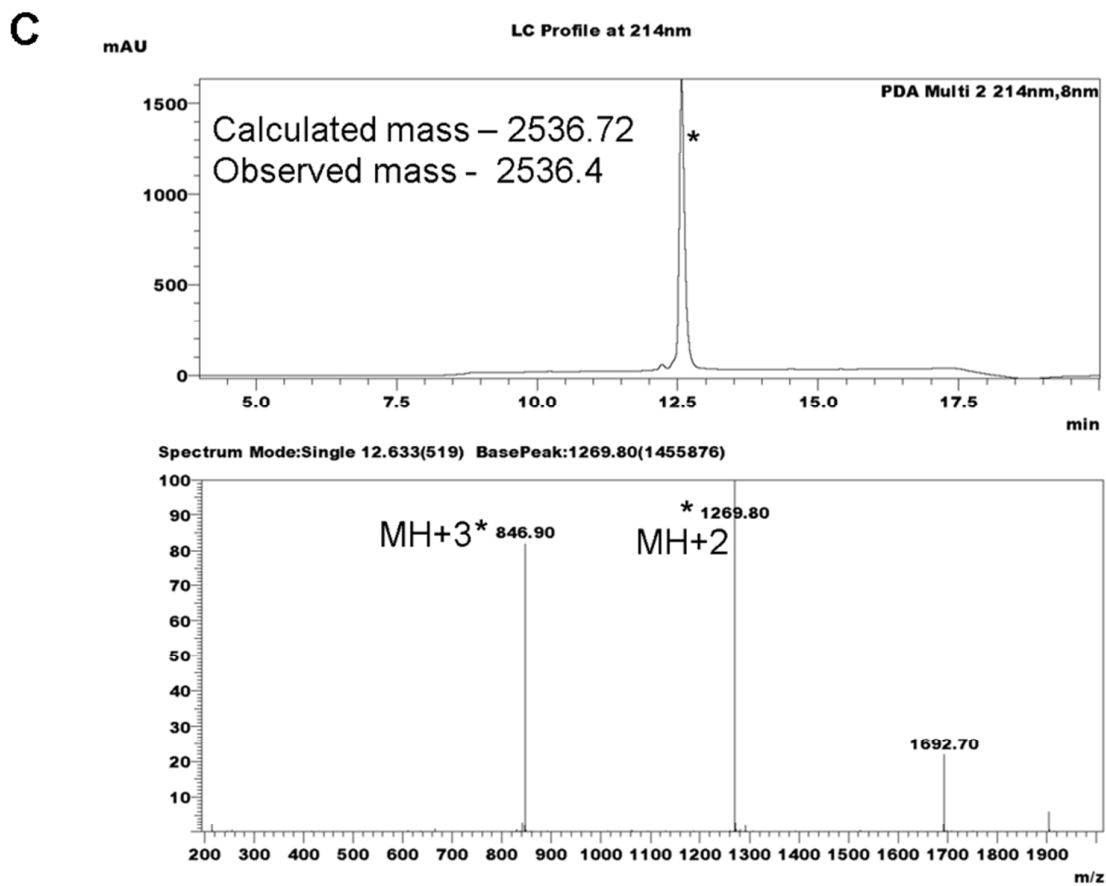
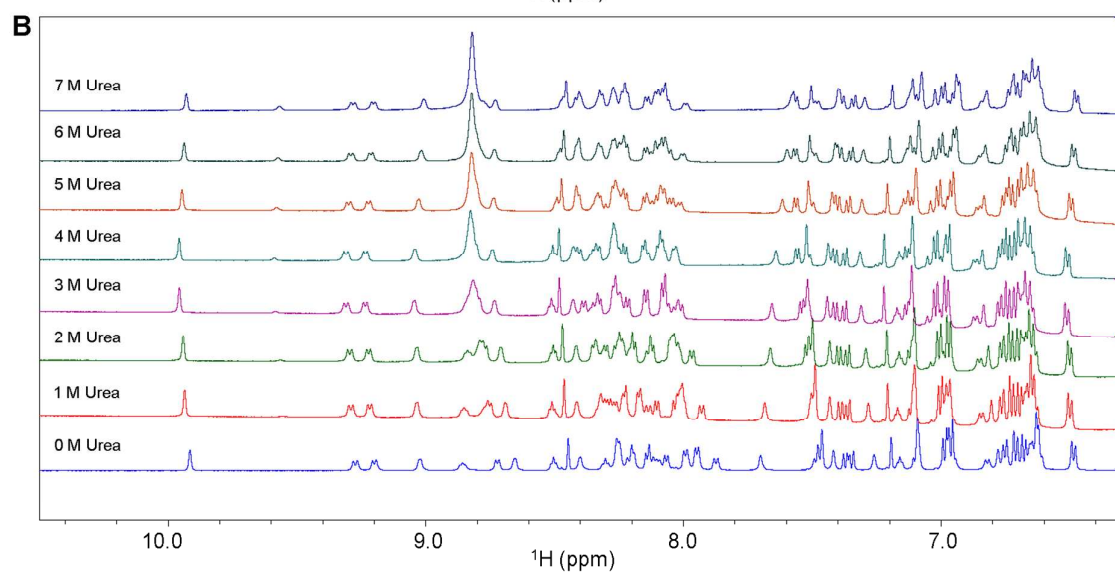
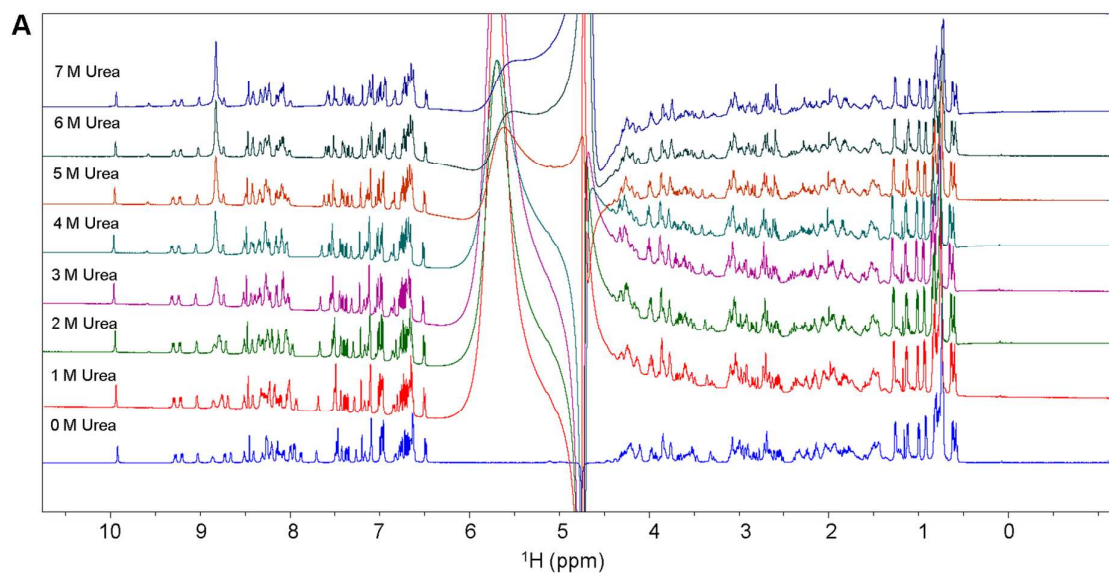


Figure S3. A. LC-MS profile of sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈]. B. LC-MS profile of sCon-Vc1₁₋₂₂[Z1Q,DINNN₁₂₋₁₆] and C. sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄].



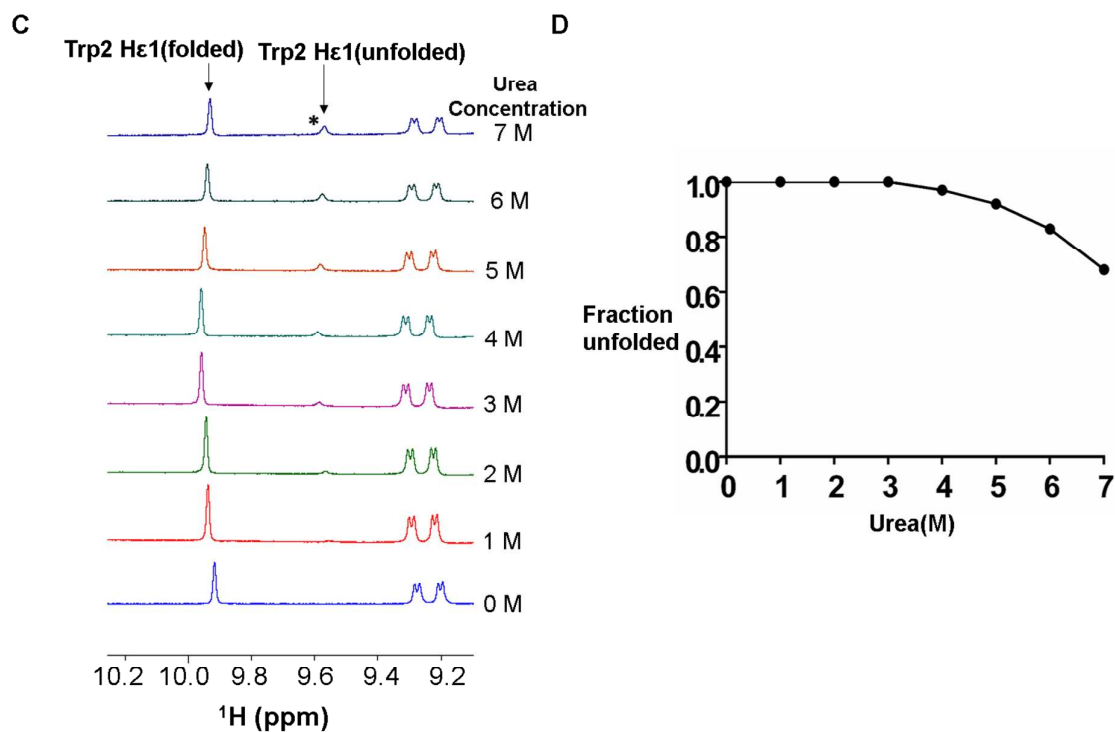
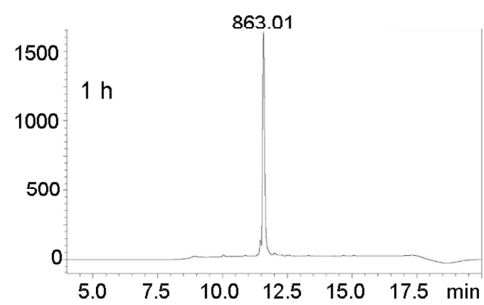
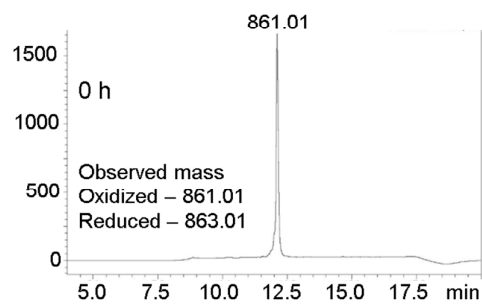
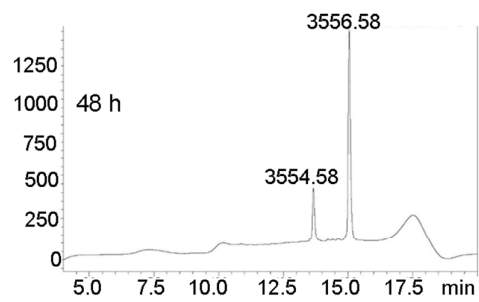
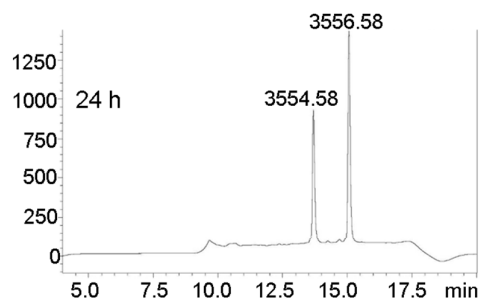
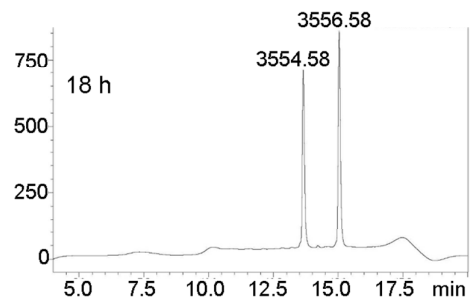
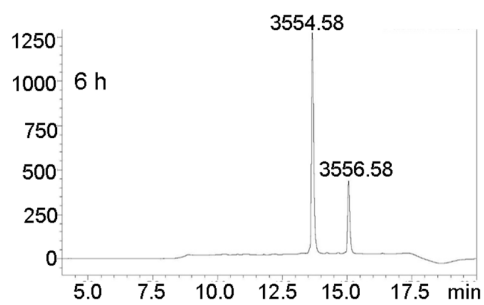
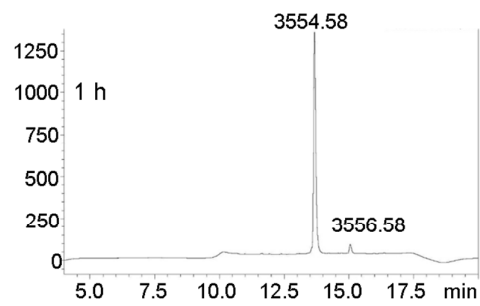
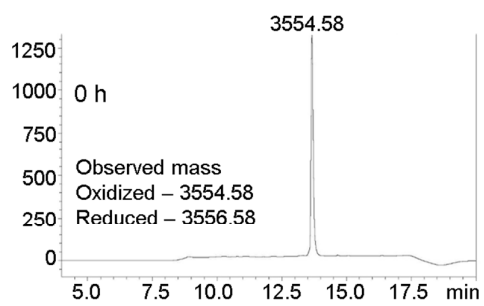
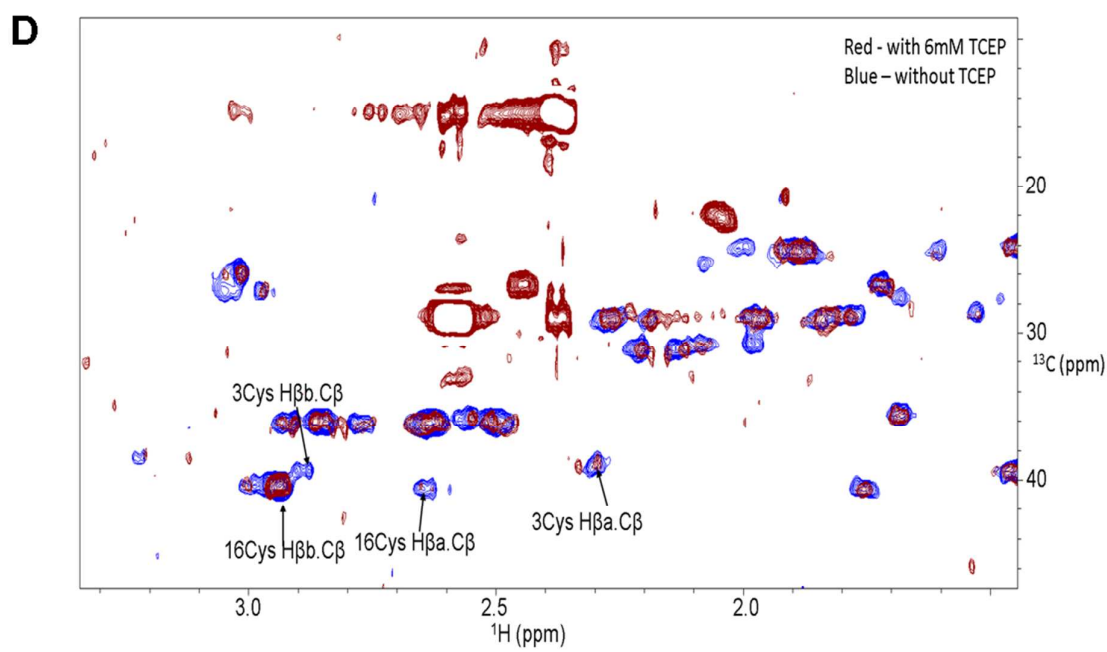
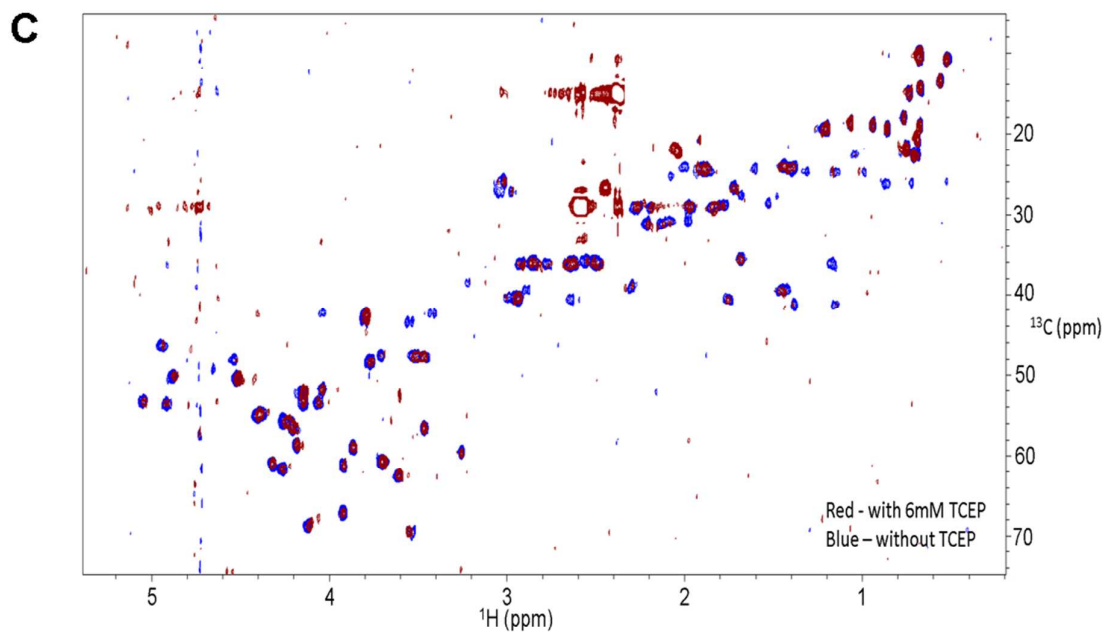


Figure S4. A. 1D ^1H NMR spectra of synthetic full-length contryphan-Vc1 with increasing concentrations of urea to 7 M at pH 4.0 and 20°C in water containing 7% $^2\text{H}_2\text{O}$. Broad signal at 5.5 ppm is due to urea. **B.** Amide and aromatic regions of spectra in A. **C.** Region of 1D ^1H NMR spectra showing Trp2 indole peak of both folded and unfolded conformations of sCon-Vc1. **D.** Urea unfolding curve for Trp2 indole proton peak of sCon-Vc1. Fraction of unfolded peptide was obtained from the relative intensities of the unfolded and folded resonances of Trp2 indole NH.



Figure S5. **A.** One-dimensional ^1H NMR spectra of synthetic full-length contryphan-Vc1 over the pH range 2-9 at 20°C in water containing 7% $^2\text{H}_2\text{O}$. **B.** Amide and aromatic region of spectra in A. **C.** Aliphatic region of spectra in A.

A**B**



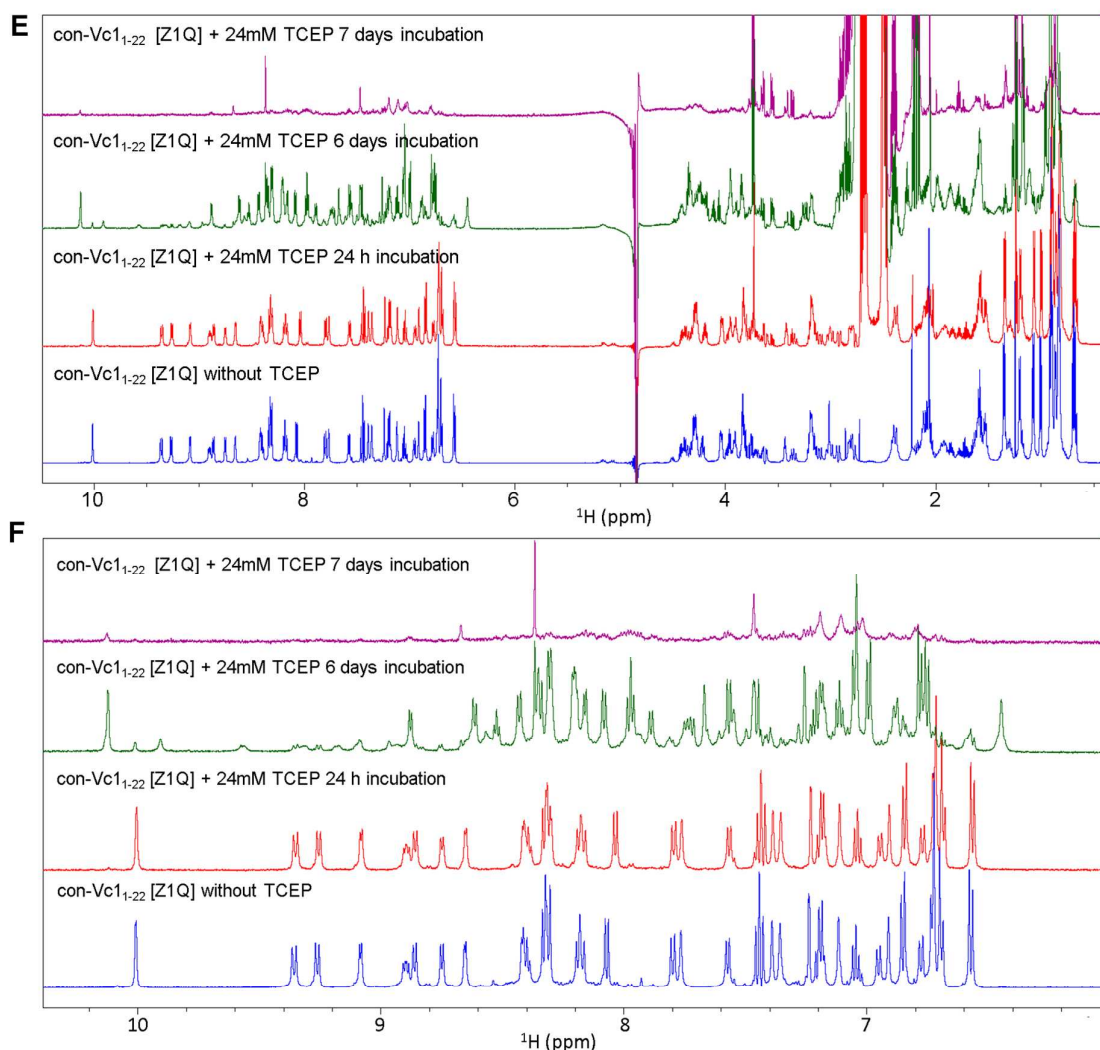
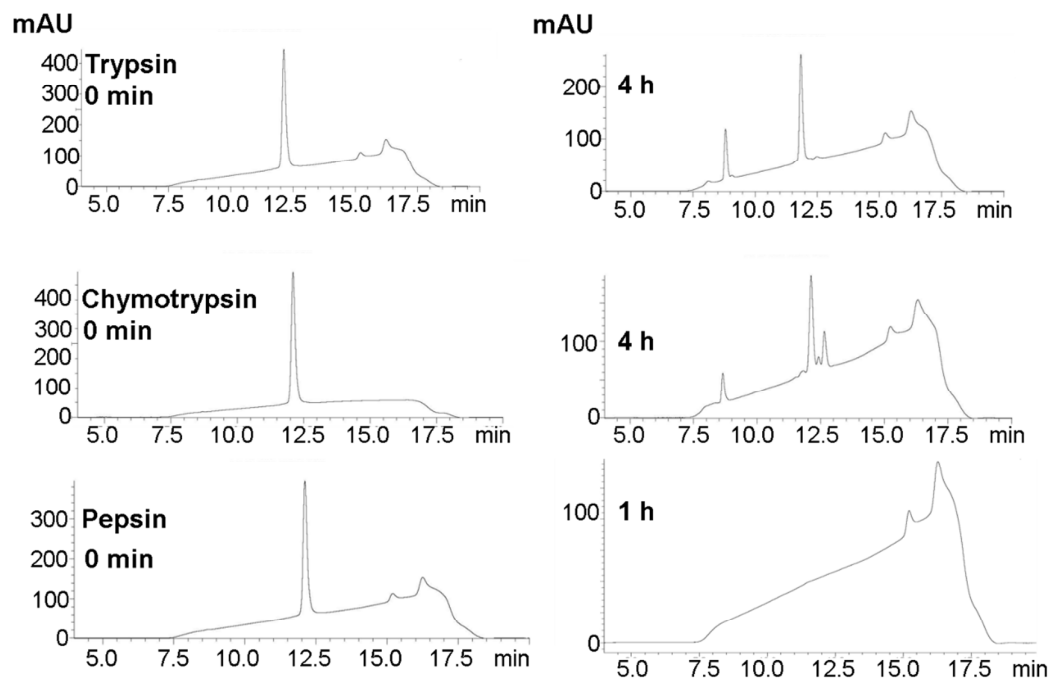
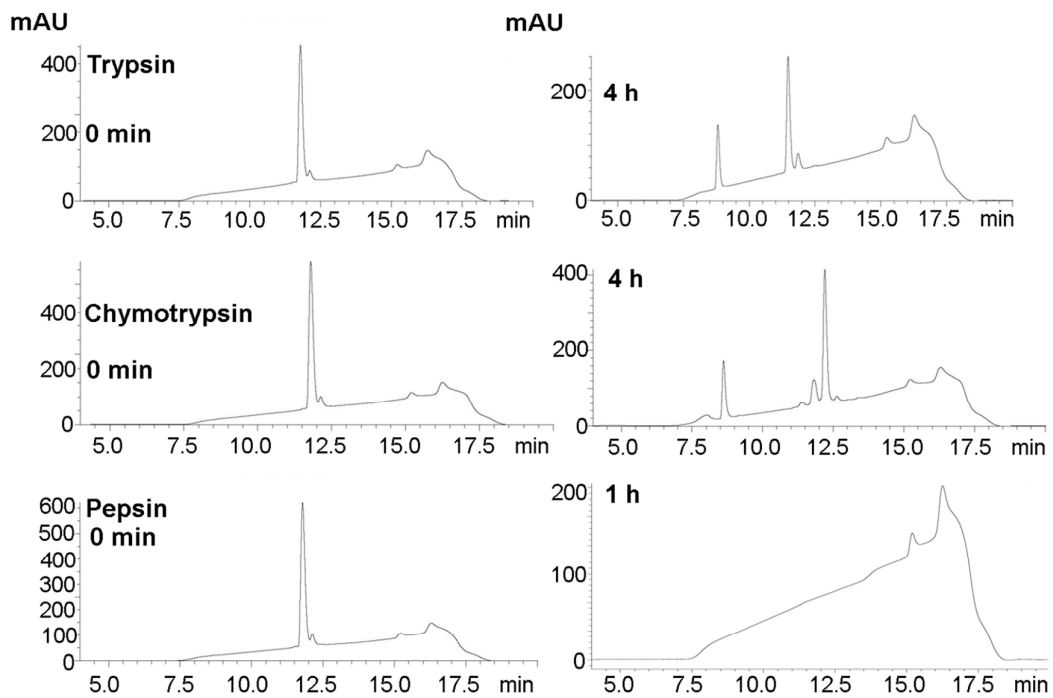


Figure S6. A. LC-MS chromatogram of contryphan-Vc2 before and after 1 h incubation with 10 mM TCEP at room temperature. **B.** LC-MS chromatogram of sCon-Vc1 at different time points after incubation with 10 mM TCEP at room temperature, with observed mass difference for oxidized and reduced species of sCon-Vc1marked. **C.** Overlay of ¹³C HSQC spectra of sCon-Vc1 in **presence** and **absence** of 6 mM TCEP. **D.** Region of ¹³C HSQC spectra of sCon-Vc1 showing ¹H peaks for Cys 3 and Cys16. **E.** One-dimensional ¹H NMR spectra of contryphan-Vc1₁₋₂₂[Z1Q] in the presence of 24 mM TCEP at 24 h, 6 days and 7 days at 20°C, pH 4.0 in water containing 7% ²H₂O. **F.** Amide and aromatic region of spectra in A.

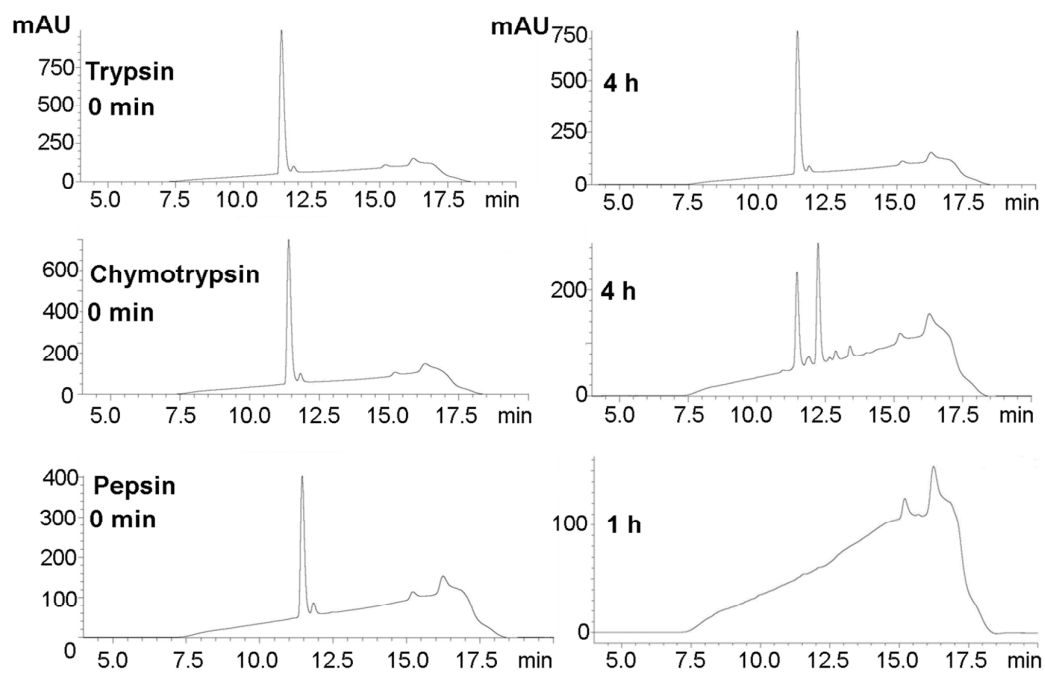
A sCon-Vc1



B rCon-Vc1[Z1Q]



C Con-Vc1₁₋₂₂[Z1Q]



D BSA

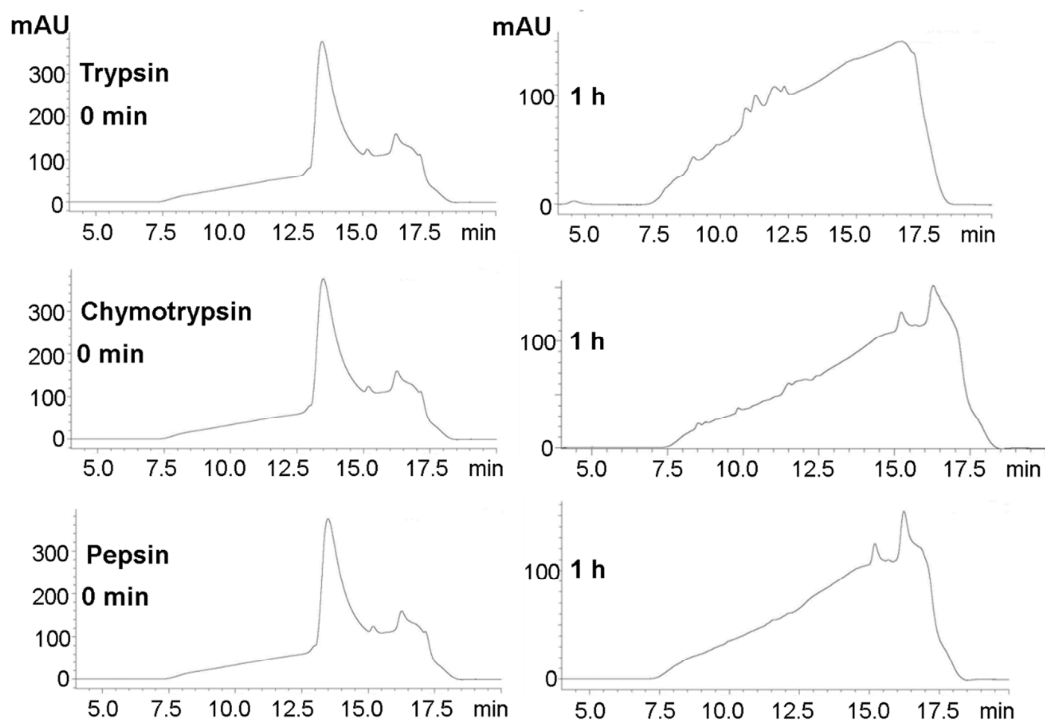


Figure S7. Reversed-phase HPLC analysis of synthetic, recombinant full-length contryphan-Vc1 and rCon-Vc1₁₋₂₂[Z1Q] treated with trypsin, α -chymotrypsin and pepsin. **A.** Chromatograms for synthetic full-length contryphan-Vc1 digested with trypsin, α -chymotrypsin and pepsin. **B.** Chromatograms for recombinant full-length contryphan-Vc1 digested with trypsin, α -chymotrypsin and pepsin. **C.** Chromatograms for rCon-Vc1₁₋₂₂ [Z1Q] digested with trypsin, α -chymotrypsin and pepsin. **D.** Chromatograms for bovine serum albumin control digested with trypsin, α -chymotrypsin and pepsin.

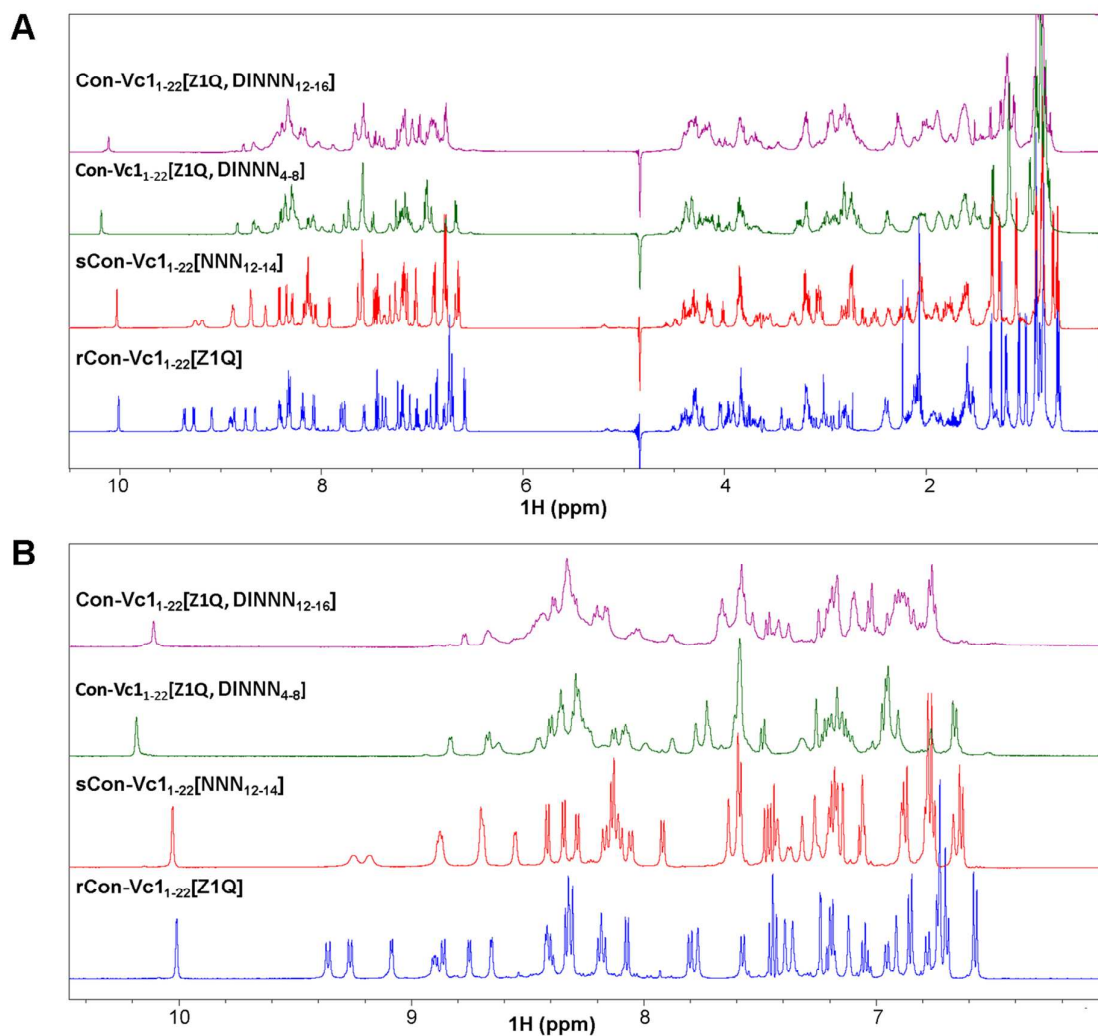


Figure S8. A. One-dimensional ^1H NMR spectra of rCon-Vc1₁₋₂₂[Z1Q], sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄], sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈] and sCon-Vc1₁₋₂₂[Z1Q,DINNN₁₂₋₁₆] at pH 4 at 20°C in water containing 7% $^2\text{H}_2\text{O}$. **B.** Amide and aromatic regions of spectra in A.

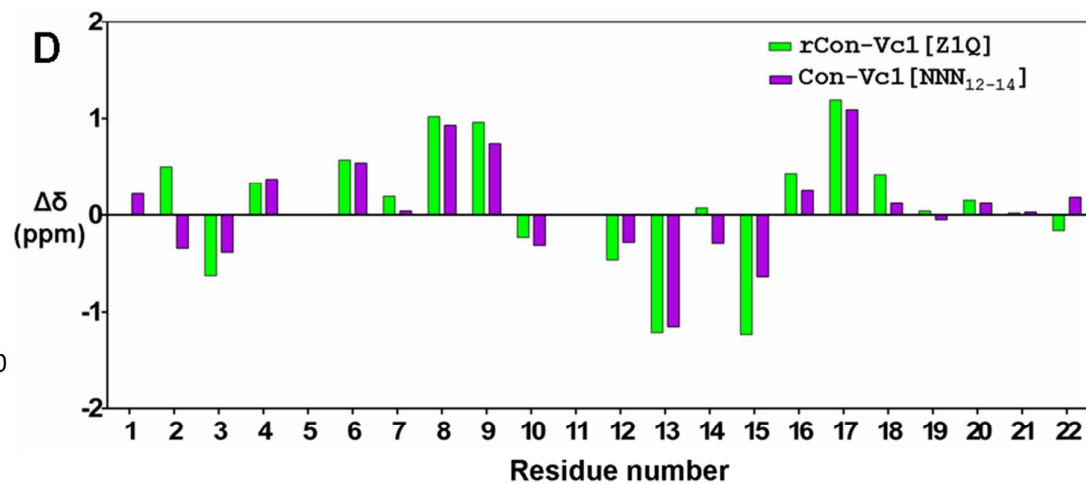
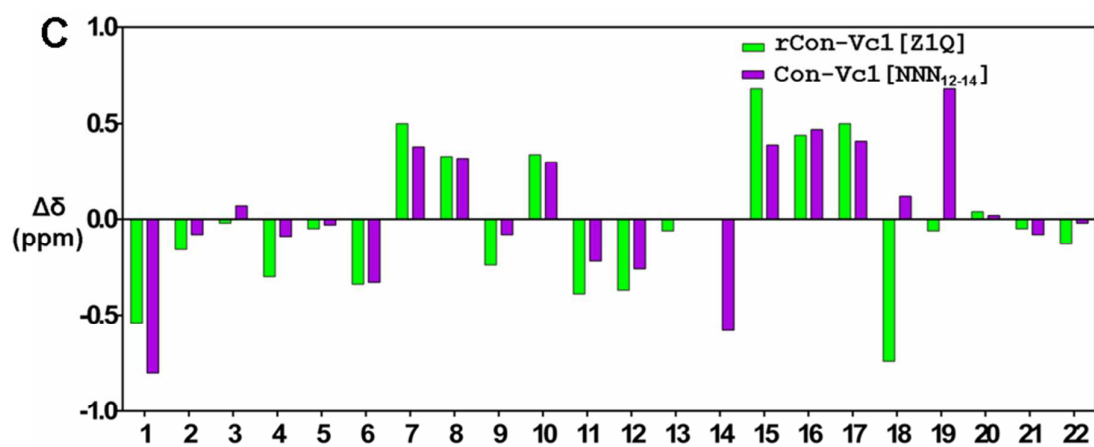
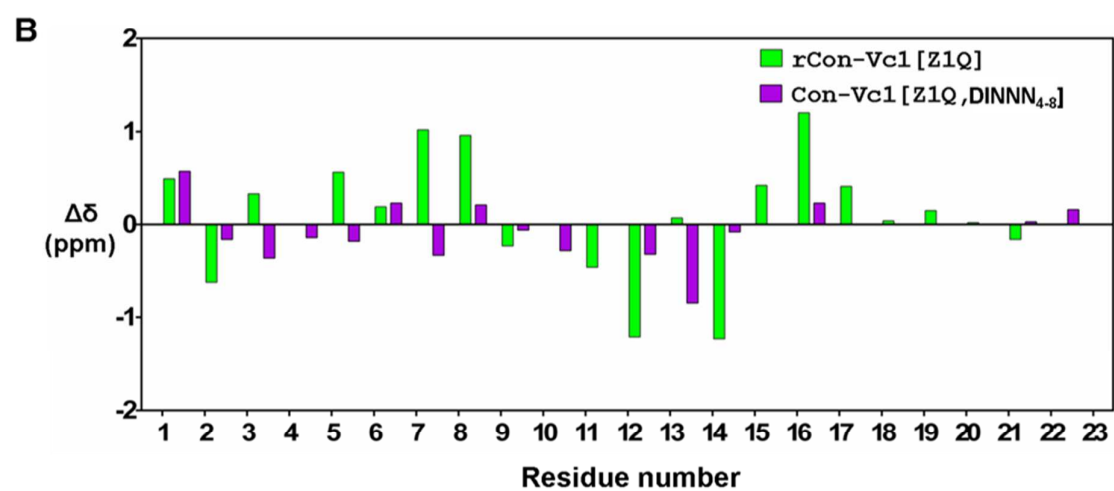
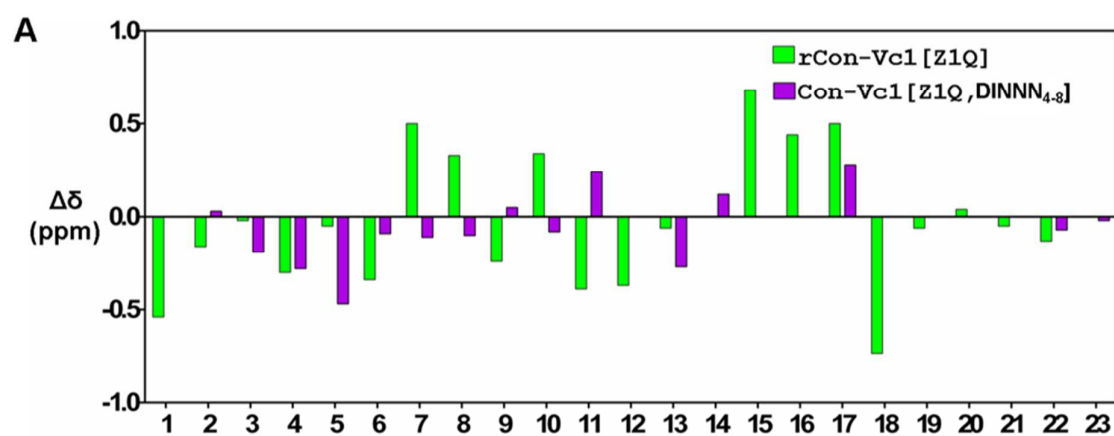


Figure S9. Deviations from random coil chemical shifts for rCon-Vc1₁₋₂₂[Z1Q], sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄] and sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈] . **A.** H^α chemical shift comparison between rCon-Vc1₁₋₂₂[Z1Q] and sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈] . **B.** H^N chemical shift comparison between rCon-Vc1₁₋₂₂[Z1Q] and sCon-Vc1₁₋₂₂[Z1Q,DINNN₄₋₈]. **C.** H^α chemical shift comparison between rCon-Vc1₁₋₂₂[Z1Q] and sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄]. **D.** H^N chemical shift comparison between rCon-Vc1₁₋₂₂[Z1Q] and sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄].¹

References

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

Chapter 4

Constraining the Single Disulfide-Directed β -Hairpin (SDH) Fold with an Additional Disulfide Bond

4.1 Introduction

The single disulfide-directed β -hairpin (SDH) fold observed first in contryphan-Vc1, has been shown to possess remarkable thermal, conformational and chemical stability and also shown to accept short bioactive epitope without compromising the core structure of the peptide.⁶⁸ Apart from maintaining the SDH fold of the peptide, the grafted molecule Vc1-NNN also enhanced the binding of the truncated NNN epitope of DINNN motif towards hSPSB2, suggesting the success in utilising the SDH fold as a peptide scaffold.^{48-50, 68} Even though the SDH fold of contryphan-Vc1 exhibits remarkable thermal, chemical and conformational stability, the peptide has shown only a limited proteolytic stability.⁶⁸ The β -hairpin core of the contryphan-Vc1 is shown to be resistant to the cleavage by the enzymes trypsin and chymotrypsin but is susceptible to the cleavage by pepsin.⁶⁸ Pepsin is an aspartic protease, using a catalytic aspartate in its active site which preferentially cleaves at Phe, Tyr, Trp and Leu in position P1 or P1'.⁶⁹ Selective residue replacement to make the peptide resistant to digestion by a particular protease is one of the methods to overcome the susceptibility of a peptide towards a protease. The sequence of the truncated contryphan-Vc1 consists of five sites that are susceptible to cleavage by pepsin, which are Trp2, Tyr7 and 9, Leu13 and 20 primary sequence rendering the method of selective residue replacement ineffective.

Apart from the susceptibility of contryphan-Vc1 towards pepsin, another aspect of the SDH fold that needs to be enhanced in order to use it as a robust scaffold is its ability to accommodate epitopes that are larger than three residues. Incorporation of the DINNN motif of inducible nitric oxide synthase (iNOS) which is essential for recognition of iNOS by the SPRY domain-containing SOCS (suppressor of cytokine signalling) box (SPSB) proteins in circulating macrophages⁴⁸⁻⁵⁰ disrupted the SDH fold of the contryphan-Vc1, suggesting a limitation of the SDH fold of contryphan-Vc1 in accommodating epitopes larger than three residues. Introducing disulfide bonds is one of the common methods utilized to increase the stability of proteins in biomedical and industrial applications, and various proteins have been successfully engineered for increasing thermal stability by introducing disulfide bridges.⁷⁰⁻⁷⁵ Interestingly, not all engineered disulfides have provided an increase in stability, as there are a number of reports of destabilizing disulfides.⁷⁶ Disulfide bonds are believed to decrease the conformational entropy and raise the free energy of the denatured state, thus increasing the stability of the folded protein conformation.⁷⁷ Given that native disulfide bonds provide considerable stability to proteins,⁷⁸ it follows that the addition of a new crosslink to a protein or peptide might increase the stability.⁷⁹

Introducing disulfide bonds, apart from increasing thermodynamic and kinetic stability of a protein, can also enhance proteolytic stability of a peptide or protein.^{80, 81} An engineered disulfide bridge in subtilisin E was shown to increase thermal stability considerably.⁸² Proteases such as thermolysin have been shown to gain stability by introduction of an additional disulfide bridge.⁸⁰ In order to enhance the proteolytic stability and to overcome the limitation of accepting only short epitopes, we decided to introduce an additional disulfide bond to further constrain the SDH fold of the contryphan-Vc1. The residues Gln1 and Tyr9 were chosen to be replaced by cysteines in order to introduce an additional disulfide bond between residues 1-9 based on their proximity in the structure of contryphan-Vc1 and based on inspection of many peptide sequences with the ICK fold, with which the SDH fold of contryphan-Vc1 shares many similarities.⁴⁷

Molecular dynamics (MD) simulations are routinely used to investigate the structure and dynamics of biological molecules and their complexes *in silico*. In this study we have utilized MD simulations in order to assess whether introduction of an additional disulfide bond would destabilize the SDH fold of contryphan-Vc1. We have synthesised contryphan-Vc1₁₋₂₂[Q1C, Y9C] on rink amide utilising orthogonal cysteine protection, where Cys at positions 1 and 9 were protected using S-acetamido methyl protective group and Cys 3 and 16 were protected using S-trityl protective group. The solution structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C] was determined using NMR spectroscopy and its proteolytic stability was assayed in the presence of trypsin, chymotrypsin and pepsin.

4.2 Materials and methods

4.2.1 MD simulations

MD simulations were carried out using GROMACS version 5.0.4, using the GROMOS 54a7 united-atom force field and a 2 fs time step. Temperature coupling made use of the velocity rescale algorithm with a reference temperature of 293 K. Pressure coupling used the Parinello-Rahman algorithm, with reference pressure of 1 bar and compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. Initial structures were generated using the Maestro software package (v 10.3.014), by *in silico* mutation of the published structure of contryphan-Vc1₁₋₂₂[Z1Q] (PDB code: 5KKM). These structures were solvated with SPC water and subjected to a steepest-descent minimisation of 2000 steps to remove bad van der Waals contacts between atoms. Temperature equilibration (without pressure coupling) was run for 10,000 steps. Isotropic pressure coupling was then applied for 500,000 steps. Following equilibration, the simulation production runs were executed for 300 ns each. The resultant trajectories were visualised with the Visual Molecular

Dynamics (VMD) software package (v 1.9.2) and analysed with VMD and GROMACS built-in tools. Root-mean-square deviation (RMSD) and radius of gyration were calculated based on data extracted using the *rms* command and the *gyrate* command of the GROMACS software and plotted using the gnuplot software package (v 5.0).

4.2.2 Peptide synthesis

The procedure for peptide synthesis used in this study is described in detail in the general materials and methods under section 2.4. Here the method specific for synthesis of contryphan-Vc11-22[Q1C, Y9C] is described briefly. Contryphan-Vc11-22[Q1C, Y9C] was synthesised on a PTI Instruments PS3 peptide synthesiser, using Rink amide AM resin. Orthogonal cysteine protection was utilised where cysteines 1 and 9 were protected using s-acetamido methyl protective group and the cysteines 3 and 16 were protected using s-trityl protective group. Cleavage from the resin was performed over 2 h with a mixture of 3,6-dioxa-1,8-octanedithiol, triisopropylsilane, 1,3-dimethoxy benzene and trifluoroacetic acid (DODT:TIPS:DMB:TFA at ratios 2.5: 2.5: 5: 92.5 by volume) The cleavage mixture was purged with nitrogen and the crude peptide was precipitated with ice-cold diethyl ether and washed three times with the same and subsequently folded in 0.1M ammonium bicarbonate by adjusting the concentration to 0.3 mg/mL and pH 8.0. The partially folded peptide where the disulfide bond between Cys 3 and 16 formed by air oxidation was purified by reversed-phase HPLC (RP-HPLC) on a Phenomenex® Luna C18 column (100 Å, 5 µm, 100 x 10 mm) using a gradient of 5–95% B (A: 99.9% H₂O, 0.1% TFA; B: 80% acetonitrile (ACN), 19.9% H₂O, 0.1% TFA) over 30-60 min and the sample was lyophilised. The disulfide bond between Cys 1 and 9 was formed by iodine oxidation where 0.5 mM of lyophilised partially folded peptide is incubated for 30 minutes in presence of 5 mM iodine dissolved in 50% acetonitrile. The reaction was stopped using 100 mM sodium ascorbate solution and the fully oxidised peptide was purified by reversed-phase HPLC and checked by LC-MS (**Figure 1**).

4.2.3 NMR spectroscopy

All spectra were acquired on a Bruker 600 MHz spectrometer equipped with a cryogenically-cooled triple-resonance probe. Lyophilised peptide was dissolved in either 93% H₂O/7% ²H₂O or 100% ²H₂O (pH 4.0). One-dimensional ¹H spectra were acquired at different temperatures between 5 and 40°C, at intervals of 5°C, and between pH 3 and 9. Two-dimensional NMR spectra utilised for sequence-specific assignments and structure calculations were acquired at pH 4 and 40°C. Two-dimensional homonuclear TOCSY spectra with a spin lock time of 80 ms were acquired using the DIPSI-2 pulse sequence with excitation sculpting for water

suppression at 15, 20 and 25°C. Two-dimensional NOESY spectra were acquired with mixing times of 50, 200 and 300 ms to analyse the time-dependence of NOE intensities. ^{13}C -HSQC and ^{15}N -HSQC spectra were acquired for carbon and nitrogen chemical shifts, respectively. A sine-bell squared window function was used for processing spectra. All spectra were processed using Bruker TopSpin (Version 3.2) and analysed using CcpNmr Analysis (Version 2.1.5). The sequence-specific resonance assignments are summarized in **Table 1**.

4.2.4 Structure Calculation

The intensities of cross peaks in the NOESY spectra with a mixing time of 300 ms were utilised to generate distance constraints. $^3J_{\text{HN-H}\alpha}$ coupling constants were measured from one-dimensional ^1H spectra, which yielded 7 ϕ angle constraints. These ϕ angles were restrained to $-120 \pm 30^\circ$ for $^3J_{\text{HN-H}\alpha} \geq 8.0$ Hz and $-65 \pm 25^\circ$ for $^3J_{\text{HN-H}\alpha} \leq 6.0$ Hz. Three distance constraints were added for the disulfide bridge as follows: 2.00, 3.00, and 3.00 Å for S(i)–S(j), S(i)–C β (j), and S(j)–C β (i), respectively. The structures of contryphan-Vc1₁₋₂₂[Q1C, Y9C] were generated using the program CYANA (version 3.0). Structure calculations were performed using 213 inter-proton distance constraints derived from the NOESY spectrum (65 intra-residue, 64 sequential, 34 medium-range and 50 long-range NOE constraints), seven dihedral angle constraints derived from $^3J_{\text{HN-H}\alpha}$ coupling constants measured from one-dimensional ^1H , and three disulfide bond restraints. The root-mean-square deviation (RMSD) values for these structures were assessed using MolMol (version 2K.1). Structural figures were prepared using PyMOL (version 1.5.0.4).

4.2.5 Proteolysis Assays

Proteolysis assays were performed at a 250:1 substrate (peptide)/enzyme ratio with pepsin, trypsin and α -chymotrypsin. For all assays, peptides were incubated with enzyme at 37°C for up to 4 h in the presence of protease. As a positive control to ensure that active enzyme was present, full-length contryphan-Vc1 was used as substrate. All digestion assay products were analysed by LC-MS (0-60 % ACN gradient, 10 min). Trypsin (EC 3.4.21.4, Sigma) and α -chymotrypsin (EC 3.4.21.1, Sigma) stocks were prepared in 50 mM Tris, 100 mM NaCl (pH 7.4), and pepsin (EC 3.4.23.1, Sigma) stocks were prepared in 10 mM HCl (pH 2). The reactions for trypsin and α -chymotrypsin are performed in 50mM Tris, 100mM NaCl, 2mM CaCl₂ pH 7.4. The reaction for pepsin is performed in 1 mM HCl pH 2. The trypsin and α -chymotrypsin reactions were quenched with 0.1% TFA, and the pepsin reaction was quenched with 50% 1 M NaOH.

4.3 Results

4.3.1 Analysis of the MD simulation trajectories

In order to check whether introduction of an additional disulfide bond between residues 1 and 9 of the contryphan-Vc1 is permissible we have performed MD simulations on contryphan-Vc1₁₋₂₂[Q1C, Y9C] and calculated global RMSD and radius of gyration (R_g) across the 300 ns simulation time of each trajectory. RMSD values indicated the extent to which the structure was changing, while radius of gyration allowed the compactness of the structure to be monitored. If the structure of the grafted construct were to unravel during the simulation, this would be indicated by an increase in R_g . The RMSD plot of the simulation trajectory over the course of 300 ns indicates that the RMSD of the molecule was fairly constant throughout the simulation time course with an initial surge before 50 ns (**Figure 1A**). The radius of gyration plot of the simulation trajectory indicates that the R_g values were generally constant throughout the simulation time course, implying that the structure of the molecule was stable and the peptide did not unfold (**Figure 1B**).

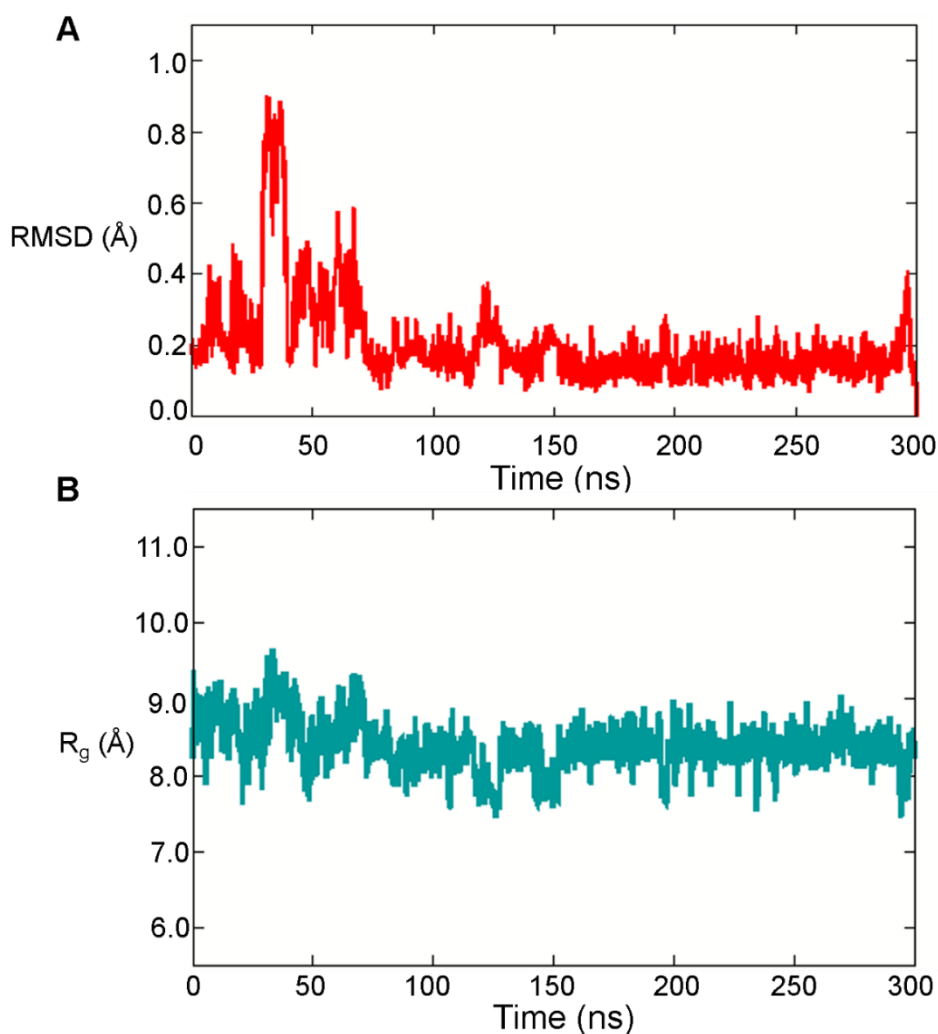


Figure 1 A. RMSD plot throughout the time course of simulation. **B.** Radius of gyration plot (R_g) of the simulation trajectory.

4.3.2 Synthesis and purification of contryphan-Vc1₁₋₂₂[Q1C, Y9C]

Contryphan-Vc1₁₋₂₂[Q1C, Y9C] was synthesized utilising solid phase peptide synthesis on rink amide AM resin using orthogonal cysteine protection. Cysteines 1 and 9 were protected using s-acetamido methyl protective group and the cysteines 3 and 16 were protected using s-trityl protective group. The partially folded peptide where the disulfide bond between Cys 3 and 16 formed by air oxidation was purified by reversed-phase HPLC (RP-HPLC). The disulfide bond between Cys 1 and 9 was formed by iodine oxidation. Purity and mass of the peptide was confirmed by LC-MS (**Figure 2**).

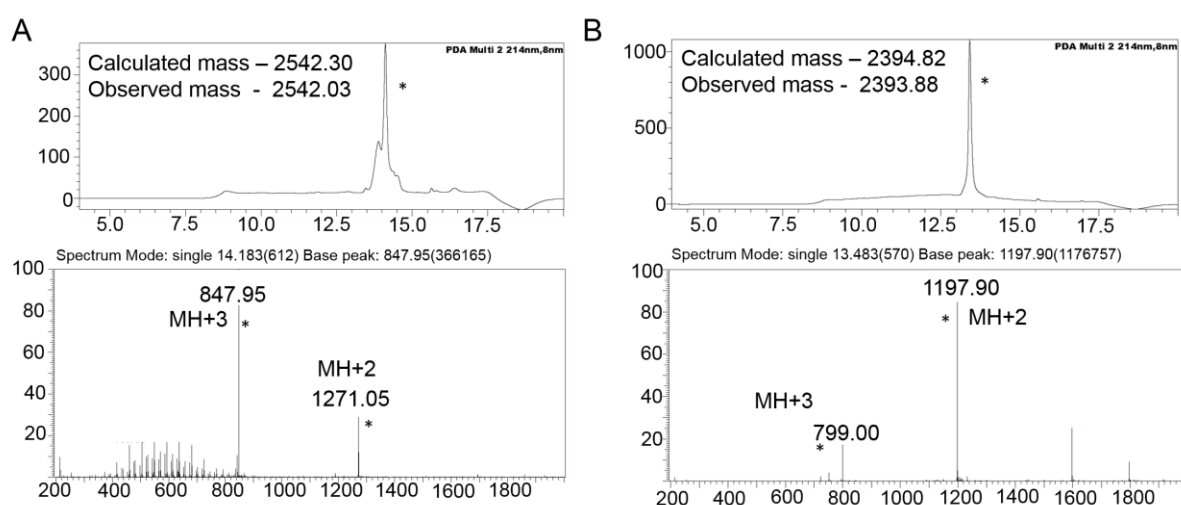


Figure 2. A. LC-MS profile of partially oxidized contryphan-Vc1₁₋₂₂[Q1C, Y9C]. **B.** LC-MS profile of fully oxidized and RP-HPLC purified contryphan-Vc1₁₋₂₂[Q1C, Y9C].

4.3.3 Temperature-dependent conformational averaging and pH stability

One-dimensional ¹H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] recorded at different temperatures with an interval of 5°C at pH 4.0 showed that the peptide exhibits temperature-dependent conformational averaging (**Figure 3A**). At temperatures of 5-25°C most of the peaks were broad, indicating the presence of different conformations in exchange. The folded conformation was stabilised at higher temperatures as the peaks became sharper and peak dispersion was similar to that of native contryphan-Vc1 (**Figure 3B**). One-dimensional ¹H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] showed subtle chemical shift changes pH 2 and 7, but the overall spectral dispersion was maintained. Most of the amide peaks were broadened by exchange with solvent at pH 8.0 (**Figure 3C**).

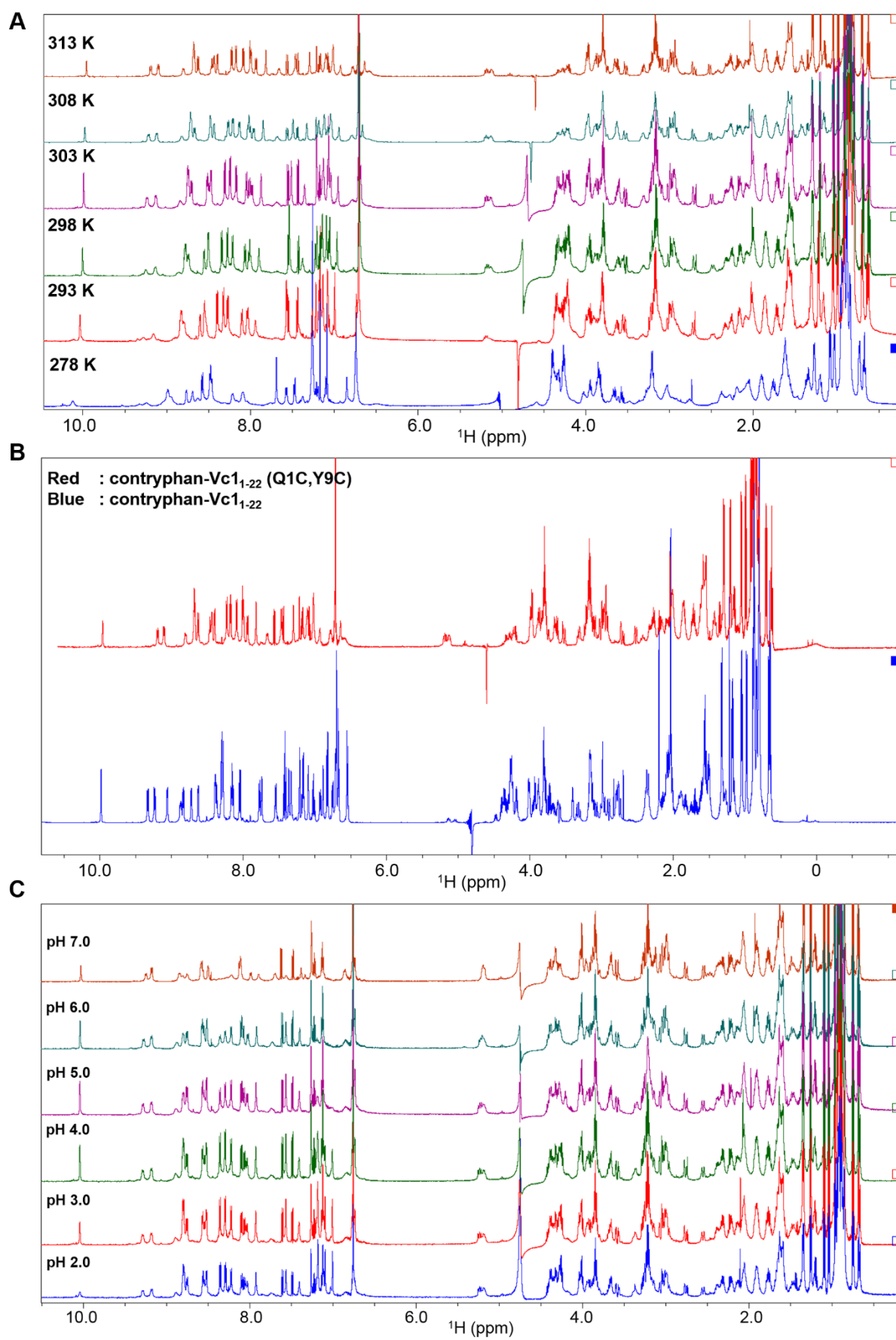


Figure 3. A. One-dimensional ^1H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] recorded at different temperatures ranging from 5°C – 40°C. **B.** Comparison of one-dimensional ^1H NMR spectra of **contryphan-Vc1₁₋₂₂[Q1C, Y9C]** with **con-Vc1₁₋₂₂[Z1Q]**. **C.** one-dimensional ^1H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] over the pH range 2-7 at 30°C in water containing 7% $^2\text{H}_2\text{O}$.

4.3.4 Sequence-Specific Resonance Assignments for contryphan-Vc1₁₋₂₂[Q1C, Y9C]

Spin systems were identified through combined analysis of DQF-COSY, TOCSY (80 ms spin-lock time) and NOESY (300 ms mixing time) spectra acquired at 313 K and pH 4.0. Sequential resonance assignments were made utilizing two-dimensional TOCSY and NOESY spectra. Residues containing methyl groups such as Ala8, Thr17 & 19, Leu13 & 20, and Ile15 & 18 served as good starting points for the relatively simple identification of spin systems and sequential assignment process (**Figure 4A**). The spin systems for Pro5 & 11 were identified by the observation of strong NOE cross peaks between Gln4 $^{\alpha}$ - Pro5 $^{\delta}$ and Asn10 $^{\alpha}$ - Pro11 $^{\delta}$; the presence of strong $\alpha\delta$ NOEs confirmed the *trans* conformations of both X-Pro bonds (**Figure 4B**). Complete backbone and side chain ^1H resonance assignments were obtained for all spin systems (**Figure 4C** and **Table 1**).

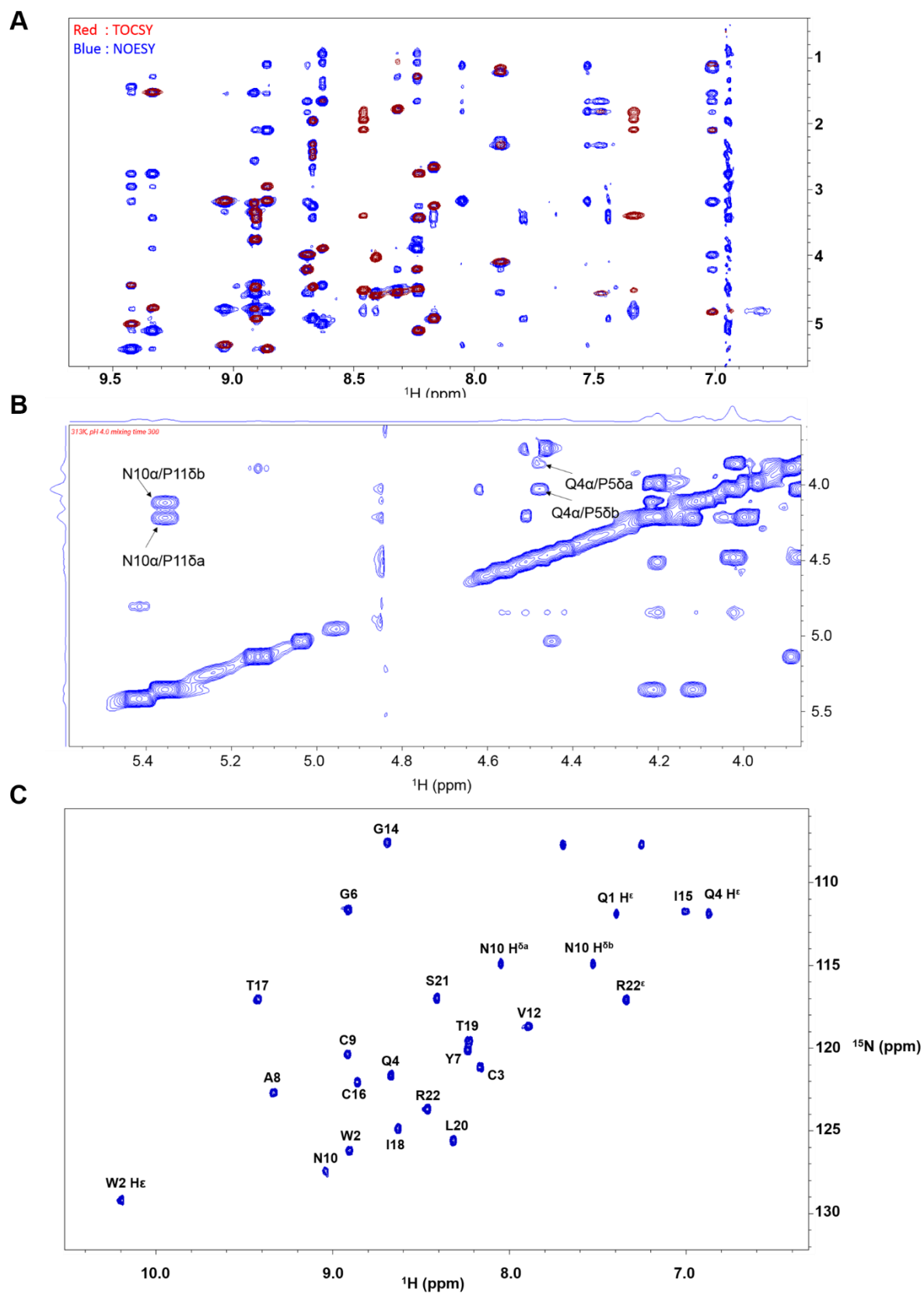


Figure 4 A. Overlay of two-dimensional ^1H NMR spectra TOCSY and NOESY of contryphan-Vc₁₁₋₂₂[Q1C, Y9C] recorded at 40°C. **B.** region of two-dimensional NOESY spectra showing cross peaks between Gln4 $^{\alpha}$ - Pro5 $^{\delta}$ and Asn10 $^{\alpha}$ - Pro11 $^{\delta}$ suggesting that both prolines are in

trans-conformation. **C.** ^{15}N -HSQC spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] at pH 4.0 and at 40°C in water containing 7% $^2\text{H}_2\text{O}$.

Table 1. Chemical shifts for contryphan-Vc1₁₋₂₂[Q1C, Y9C] at pH 4.0, 313K

Residue	H ^N	H ^α	H ^β	N	other
Cys1	ND	4.42	3.40,3.54	ND	
Trp2	8.90	4.96	3.44,3.37	126.19	H ^d 1 7.45 H ^e 1 10.19 H ^e 3 7.80 q H ² 2 7.63q H ³ 3 7.67 q N ^e 129.20
Cys3	8.17	4.95	2.66,3.25	121.18	
Gln4	8.67	4.48	1.96,2.31	121.64	H ^γ 2.42,2.50 H ^e 7.31,7.41 N ^e 111.90
Pro5	-	4.57	2.55,2.09	-	H ^γ 2.39, 2.25; H ^δ 3.86, 4.03
Gly6	8.91	4.49,3.76	-	111.66	
Tyr7	8.23	5.14	2.76,3.43	120.11	H ^δ 6.95; H ^e 6.94
Ala8	9.34	4.79	1.53	122.71	
Cys9	8.91	4.81	3.21,3.34	120.39	
Asn10	9.04	5.36	3.18	127.44	H ^δ 2 7.53, 8.05 N ^δ 2 114.92
Pro11	-	4.56	2.51	-	H ^γ 2.25 ; H ^δ 4.22,4.12
Val12	7.89	4.1	2.33	118.7	H ^γ a 1.14; H ^γ b 1.22
Leu13	7.47	4.57	1.81		H ^γ 1.66 H ^δ a 1.09
Gly14	8.70	3.99,4.21	-	107.60	
Ile15	7.01	4.85	2.10	111.76	H ^γ 1 1.54; H ^γ 1.17; H ^δ 1 1.10
Cys16	8.86	5.42	2.96,3.17	122.06	
Thr17	9.42	5.04	4.45	117.07	H ^γ 2 1.43
Ile18	8.63	3.89	1.65	124.86	H ^γ 1 0.94 ; H ^γ 2 1.34;H ^δ 1 0.86
Thr19	8.24	4.51	4.21	119.58	H ^γ 2 1.28
Leu20	8.32	4.56	1.78	125.58	H ^δ a 1.06; H ^δ b 1.04
Ser21	8.41	4.61	4.03	117.00	
Arg22	8.46	4.53	1.95,2.09	123.70	H ^γ 1.51; H ^δ 3.40; H ^e 7.34 ; N ^e 117.07

ND- not determined.

4.3.5 Solution Structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C]

In order to investigate the influence of the additional disulfide bond between the residues 1 and 9 on the structure of the contryphan-Vc1, the structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C] was determined and compared with that of native contryphan-Vc1. Contryphan-Vc1₁₋₂₂[Q1C, Y9C] retained the fold of the contryphan-Vc1 even after the introduction of an additional disulfide bond and is almost identical to that of the native contryphan-Vc1, as shown in **Figure 3**. Superimposition of the structure with the lowest target function of contryphan-Vc1₁₋₂₂[Q1C, Y9C] and con-Vc1₁₋₂₂[Z1Q] yielded an RMSD of 0.9 Å over C^α atoms of residues 3-16 (**Figure 4B**). The main feature of the structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C] is an anti-parallel β-sheet with two β-strands connected by a β-turn (**Figure 5A&C**). The four residues Gln4, Pro5, Gly6 and Tyr7 form a type-II β-turn that leads to the first β-strand. Ala8, Cys9 and Asn10 form

the first β -strand and Ile15, Cys16 and Thr17 make up the other strand. Pro11, Val12, Leu13 and Gly14 form a β -hairpin. Similar to that of con-Vc1₁₋₂₂[Z1Q] residues after the second β -strand, Ile18 - Arg22, are unstructured and highly flexible. Structural constraints are summarized in Table 2.

Table 2. Structure Statistics for contryphan-Vc1₁₋₂₂[Q1C, Y9C]

NMR Distance and Dihedral Constraints	
Distance constraints	
Total NOEs	213
Intra-residue	65
Inter-residue	
Sequential ($ i-j = 1$)	64
Short	129
Medium ($1 < i-j \leq 5$)	34
Long ($ i-j > 5$)	50
Hydrogen bond	4
Total dihedral angle restraints	
Backbone (ϕ angle)	7
Structure statistics	
RMSD between 20 conformers (residues 3-16)	
Average pairwise rmsd ^b (Å)	
Backbone (Å) (N, C ^{α} , C)	0.27 +/- 0.06
All heavy atoms (Å)	0.55 +/- 0.07
Ramachandran analysis	
Residues in most favoured regions (%)	68.1
Residues in additionally allowed regions (%)	31.9
Residues in generously allowed regions (%)	0.0
Residues in disallowed regions (%)	0.0

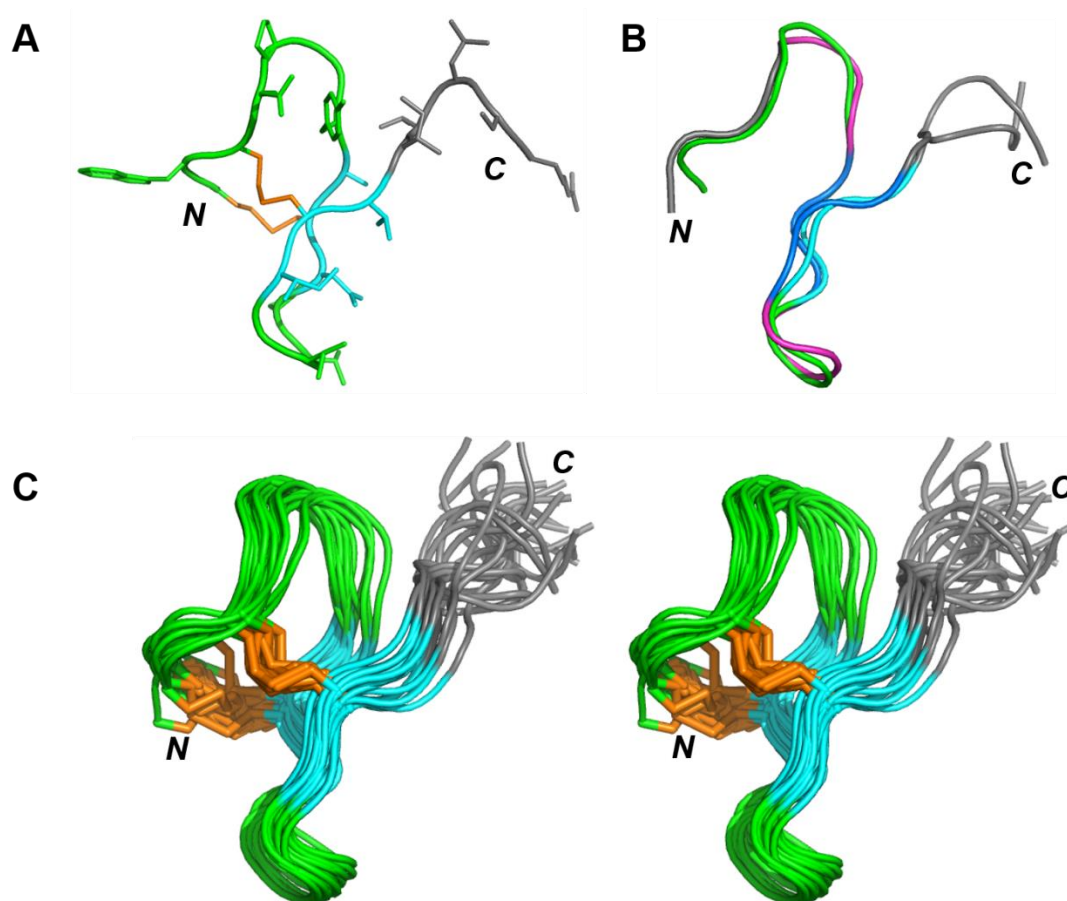


Figure 5. **A.** PYMOL representation of the closest-to-average structure of contryphan-Vc11-22[Q1C, Y9C] β-strands are shown in marine, loops are shown in green, residues 20-22 are shown in grey and the disulfide bonds are shown in orange. **B.** Backbone overlay of closest-to-average structures of recombinant con-Vc11-22[Z1Q] (loops in magenta and β-strands in blue) with contryphan-Vc11-22[Q1C, Y9C] (loops in green and β-strands in marine) using PYMOL. **C.** Ensemble of the final 20 structures of contryphan-Vc11-22[Q1C, Y9C] from CYANA superimposed over the backbone heavy atoms (N, C^α, C) of residues 3–16.

4.3.6 Proteolytic stability of contryphan-Vc1

In vitro proteolysis was performed on contryphan-Vc11-22[Q1C, Y9C] utilising the enzymes trypsin, α-chymotrypsin and pepsin in order to assess whether the proteolytic stability of contryphan-Vc1 was enhanced after the introduction of an additional disulfide bond. Contryphan-Vc11-22[Q1C, Y9C] did not show any peaks corresponding to the cleavage of trypsin as expected because the peptide has no tryptic cleavage site (**Figure 6A**). In the presence of α-chymotrypsin contryphan-Vc11-22[Q1C, Y9C] showed two major peaks in the LC-MS chromatogram, with masses of 2393.9 and 2153.5 Da, which correspond to uncleaved peptide and fragment 1-20, respectively (**Figure 6B**). Contryphan-Vc11-22[Q1C, Y9C] also did not show any cleavage peaks in presence of pepsin indicating that the peptide was now resistant to pepsin digestion (**Figure 6C**).

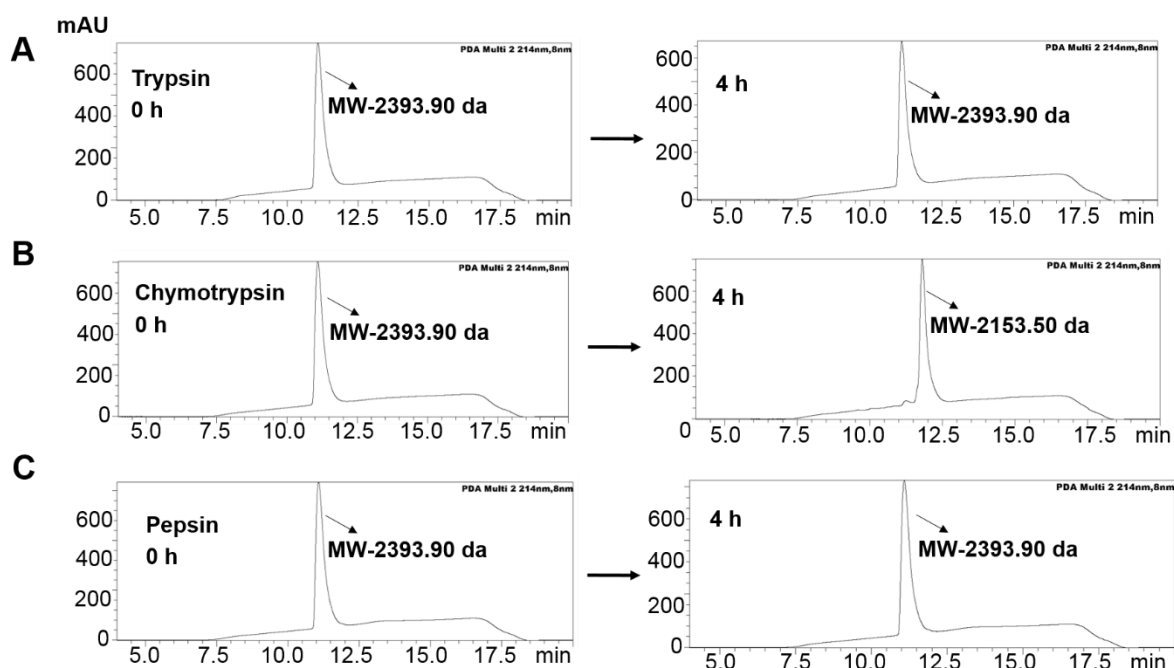


Figure 6. Reversed-phase HPLC analysis of contryphan-Vc1₁₋₂₂[Q1C, Y9C] treated with trypsin, α -chymotrypsin and pepsin. **A.** Chromatograms for contryphan-Vc1₁₋₂₂[Q1C, Y9C] digested with trypsin. **B.** Chromatograms for contryphan-Vc1₁₋₂₂[Q1C, Y9C] digested with α -chymotrypsin. **C.** Chromatograms for contryphan-Vc1₁₋₂₂[Q1C, Y9C] digested with pepsin.

4.4 Discussion

In the present study we have demonstrated that an additional disulfide bond can be introduced by replacing Gln1 and Tyr9 with Cys and using orthogonal Cys protection. We have shown that the introduction of an additional disulfide bond to contryphan-Vc1 did not disrupt its SDH fold. Temperature-dependent conformational averaging phenomenon was observed, where the folded conformation of the peptide was stabilised at higher temperatures. We have also shown that the introduction of an additional disulfide bond enhanced its proteolytic stability, with contryphan-Vc1₁₋₂₂[Q1C, Y9C] being completely resistant to pepsin digestion, in contrast to the native contryphan-Vc1.

Structural similarities of the SDH fold of the contryphan-Vc1 with the peptides having ICK motif enabled us to analyse the disulfide framework of different ICK peptides. Based on these analyses, Gln1 and Tyr9 were chosen to be replaced by Cys in order to introduce an additional disulfide bond between residues 1-9 based on the disulfide framework of the ICK peptides and also their proximity in the structure of contryphan-Vc1.⁶⁸ RMSD values from the backbone of structures throughout the simulation indicated the extent to which the structure was changing, while the radius of gyration allowed the compactness of the structure to be monitored. The RMSD plot of the simulation trajectory over the course of 300 ns indicated that the RMSD of the molecule was fairly constant throughout the simulation time course (**Figure**

2A). Moreover, the radius of gyration plot of the simulation trajectory indicated that the R_g values were generally constant throughout the simulation time course suggesting that the structure of the molecule was stable and the peptide did not unfold (**Figure 2B**).

One-dimensional ^1H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] recorded at 5-25°C showed that most of the peaks were broad, indicating presence of different conformations in exchange. As the temperature was increased the peaks sharpened and the dispersion became similar to that of native contryphan-Vc1, suggesting that the folded conformation was being stabilised at higher temperatures (**Figure 3**). This phenomenon of inverse temperature transition (ITT),⁸³ as seen in the case of elastin-like peptides where the peptides change from a disordered (extended) to an ordered (folded) conformation upon heating, is quite intriguing and further study is warranted to understand the reasons for this temperature-dependent conformational averaging. One-dimensional ^1H NMR spectra of contryphan-Vc1₁₋₂₂[Q1C, Y9C] showed subtle chemical shift changes in the pH titration between pH 2 and 7, but the overall spectral dispersion was maintained, indicating the stability of the peptide over this pH range.

The solution structure of contryphan-Vc1₁₋₂₂[Q1C, Y9C] was similar to that of native contryphan-Vc1, with superimposition of the structures of contryphan-Vc1₁₋₂₂[Q1C, Y9C] and con-Vc1₁₋₂₂[Z1Q] with the lowest target function showing an RMSD of 0.9 Å over C $^\alpha$ atoms of residues 3-16 (**Figure 5**). Thus ontryphan-Vc1₁₋₂₂[Q1C, Y9C] maintained the overall fold of the native contryphan-Vc1.

Another important driver for the introduction of an additional disulfide was to enhance the proteolytic stability of the native contryphan-Vc1. Contryphan-Vc1₁₋₂₂[Z1Q] was resistant to trypsin cleavage, has a single cleavage site at Leu20 for α -chymotrypsin and was completely degraded in presence of pepsin. Contryphan-Vc1₁₋₂₂[Q1C, Y9C] is shown to be resistant to trypsin digestion, but was cleaved at Leu20, in presence of chymotrypsin similar to that of con-Vc1₁₋₂₂[Z1Q]. In contrast contryphan-Vc1₁₋₂₂[Q1C, Y9C] was completely resistant to pepsin digestion whereas native contryphan-Vc1 was completely degraded by this enzyme (**Figure 6**).

In the current study, we have shown that the SDH fold of the contryphan-Vc1 can be further constrained by the introduction of an additional disulfide bond while maintaining the overall fold of native contryphan-Vc1. We have further shown that engineering contryphan-Vc1 with an additional disulfide bond enhanced the limited proteolytic stability of the native contryphan-Vc1. Further studies to check whether contryphan-Vc1₁₋₂₂[Q1C, Y9C] can accept epitopes larger than three residues, eg. The DINNN epitope, will establish whether the disulfide

engineering is successful or not in overcoming the limitation of accepting short epitopes (~ 3 residues) by con-Vc11-22[Z1Q].

Chapter 5

Membrane Binding Properties of Contryphan-Vc1

5.1 Introduction

Most peptide toxins exert their effects by targeting ion channels and receptors of both the central and/or peripheral nervous systems, interfering with action potential conduction and/or synaptic transmission.⁸⁴ An increasing body of evidence suggests that some peptide toxins act by forming a ternary complex with a membrane and receptor in concert, or even target the membrane directly, leading to lysis of the cells.⁸⁴⁻⁸⁸ Melittin, a peptide from bee venom, has been shown to bind to lipid membranes as a monomer and induce transient leakage, whereas at higher concentrations, it can form stable pores, and can disintegrate the membrane through detergent-like permeabilisation activity.⁸⁹ Other well-known classes of membrane active peptides are the snake venom cytotoxins and phospholipases, the activities of which have been reviewed extensively.⁸⁶

Alternatively, some venom peptides have been shown to bind voltage-sensing domains (VSDs) of voltage-gated ion channels (VGICs) and alter the kinetics and gating behaviour by changing the relative stability of the closed, open or inactivated states of the channel.^{14, 90, 91} Several gating modifier toxins (GMTs), including GsMTx-IV, HaTx-I, VsTx-I, ProTx-I, ProTx-II and SgTx-I, have been shown to bind to model membranes.⁹²⁻⁹⁶ Trimolecular lipid–peptide–channel complexes have subsequently been proposed to exist in interactions between GMTs, voltage-gated ion channels and the lipid membrane.^{87, 97} Interestingly, all these gating modifier toxins share a similar structural fold and have at least six cysteine residues arranged to form an ICK motif.¹⁵ In addition, these peptides share a conserved amphipathic surface profile characterized by a high proportion of hydrophobic amino acid residues such as Trp, Tyr and Phe, surrounded by a ring of cationic residues including Arg and Lys that typically promote peptide-membrane interactions,⁹² suggesting a common mode of access and binding to the VSDs.

Even though bilayer partitioning was considered to be a common mechanism for all the ion-channel gating modifiers, the ability of these different GMTs to interact with model lipid bilayers was shown to vary considerably.⁹⁸ GsMTx-IV and VsTx-I were shown to interact with both anionic and zwitterionic lipids, while SgTx-I and the GsMTx-I were shown to bind strongly to anionic lipids only, suggesting that the factors involved in membrane binding are still unclear.⁹⁹ Contryphan-Vc1, a 31-residue peptide identified in the venom of the marine cone snail *Conus victoriae*,^{40, 100} has been shown to have a unique fold, designated as the single disulfide-directed β -hairpin (SDH). The core structure of contryphan-Vc1 is a double-stranded anti-parallel β -sheet stabilized by just a single disulfide bridge. The SDH fold of contryphan-Vc1 is very similar structurally to the multiple disulfide containing ICK motif.^{47, 101} Apart from

the structural similarity, contryphan-Vc1 has also shown to possess remarkable conformational, chemical and thermal stability, which are the preferable qualities in many of the ICK peptides.^{23, 24, 68} The activity of contryphan-Vc1 remains elusive and the peptide did not show any reproducible activity in mice upon intracranial injection⁴⁷ The peptide did not produce any observable changes to normal or depolarization-induced intracellular Ca^{2+} levels in mouse dorsal root ganglion cells.⁴⁷

The surface of the β -hairpin core of contryphan-Vc1 is highly hydrophobic with Trp, Pro, Tyr, Ala, Leu, Val and Ile. Polar residues like Gln, Ser, Thr and Arg are on the edges and another hydrophilic residue asparagine is buried.^{47, 68} The similarity in structure to the ICK peptides and presence of a hydrophobic core of contryphan-Vc1 has posed the question of whether the peptide binds to lipid membranes. Dodecylphosphocholine (DPC) is a detergent often used to mimic zwitterionic membranes by forming micelles and DPC micelles have been successfully utilized as a model membrane in a number of studies.^{95, 102, 103} In this study we have utilized dodecylphosphocholine (DPC) micelles as a model membrane system and characterized the interactions of contryphan-Vc1 and DPC micelles utilising solution NMR. Zwitterionic lipid 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) is used as model bilayer and the interactions were analysed utilizing surface Plasmon resonance.

5.2 Materials and methods

5.2.1 Peptide synthesis

The procedure for peptide synthesis used in this study is described in the general Materials and Methods (Chapter 2 and section 2.4) quite elaborately.

5.2.2 Sample preparation and NMR spectroscopy

All spectra were acquired on a Bruker 600 MHz spectrometer equipped with a cryogenically-cooled triple-resonance probe. Lyophilised peptide was dissolved in 93% H_2O /7% $^2\text{H}_2\text{O}$ (pH 4.0). Working stocks of $^2\text{H}_{38}$ -DPC (Cambridge Isotope Laboratories) and LPPG (Avanti Polar Lipids) were freshly prepared in the Milli-Q water. For samples containing DPC micelles, 50 mM DPC solutions were prepared by dissolving 16 mg of dodecylphosphocholine or deuterated dodecylphosphocholine in milli-Q water and added to the lyophilized peptide. One-dimensional ^1H spectra of full-length contryphan-Vc1 was acquired at 20°C, at pH 4.0 in the presence of 50 mM DPC. One-dimensional ^1H spectra of truncated contryphan-Vc1 (Con-Vc1₁₋₂₂[Z1Q]) were recorded in the presence of different concentrations of DPC micelles ranging from 10-50 mM at temperatures 20°C and 40°C. Two-dimensional NMR spectra utilised for sequence-specific assignments were acquired at pH 4 and 40°C. Two-dimensional

homonuclear TOCSY spectra with a spin lock time of 80 ms were acquired using the DIPSI-2 pulse sequence with excitation sculpting for water suppression at 40°C in the presence of 20, 30, 40 and 50 mM DPC. Similarly, two-dimensional NOESY spectra were acquired at 200 ms in the presence of 20, 30, 40 and 50 mM DPC at 40°C. A sine-bell squared window function was used for processing spectra. All spectra were processed using Bruker TopSpin (Version 3.2).

5.2.3 Surface plasmon resonance

The lipid binding property of Con-Vc1₁₋₂₂[Z1Q] and full-length contryphan-Vc1 were analyzed by SPR, using a Biacore T200 instrument. All experiments were performed in degassed 25 mM HEPES buffer (pH 7.4) containing 150 mM NaCl, 3 mM EDTA, and 0.05% surfactant P-20 at 25°C. Around 700 µM of POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) was immobilised on a Biacore L1 chip using buffer containing degassed 20 mM HEPES buffer (pH 7.4) containing 50 mM NaCl, and 3 mM EDTA at 25°C. POPC lipid was prepared by dissolving in chloroform:methanol (1 : 1), dried under N₂ and vacuum, and resuspended to 26.3 mM lipid in buffer. The lipid mixture was sonicated until a clear solution was obtained. The peptides were injected onto the surface with a contact time of 100 s, flow rate of 100 µL/min, and dissociation time of 600 s. Sensorgrams were further corrected for nonspecific binding to the surface by subtracting the signals of the reference surface from those of the target protein surface.

5.3 Results

5.3.1 Interaction of full-length contryphan-Vc1 with DPC micelles

One-dimensional ¹H spectra of full-length contryphan-Vc1 in the presence of 50 mM DPC has shown significant changes when compared to the spectra recorded in the absence of DPC micelles (**Figure 1**). Intermediate exchange of peaks with DPC micelles was observed for the peaks between 8.8 and 10.0 ppm, which included the peak corresponding to the indole NH of Trp2 and the backbone amide peaks for Gly6, Ala8, Tyr9, Cys16, and Thr17, with peaks corresponding to those residues becoming very broad and not observed in the spectrum. Peak broadening and shifts were observed for peaks between 6.5 and 8.86 ppm. Peak exchange, shift of peaks and peak broadening for several residues confirmed that the peptide is interacting with the DPC micelles.

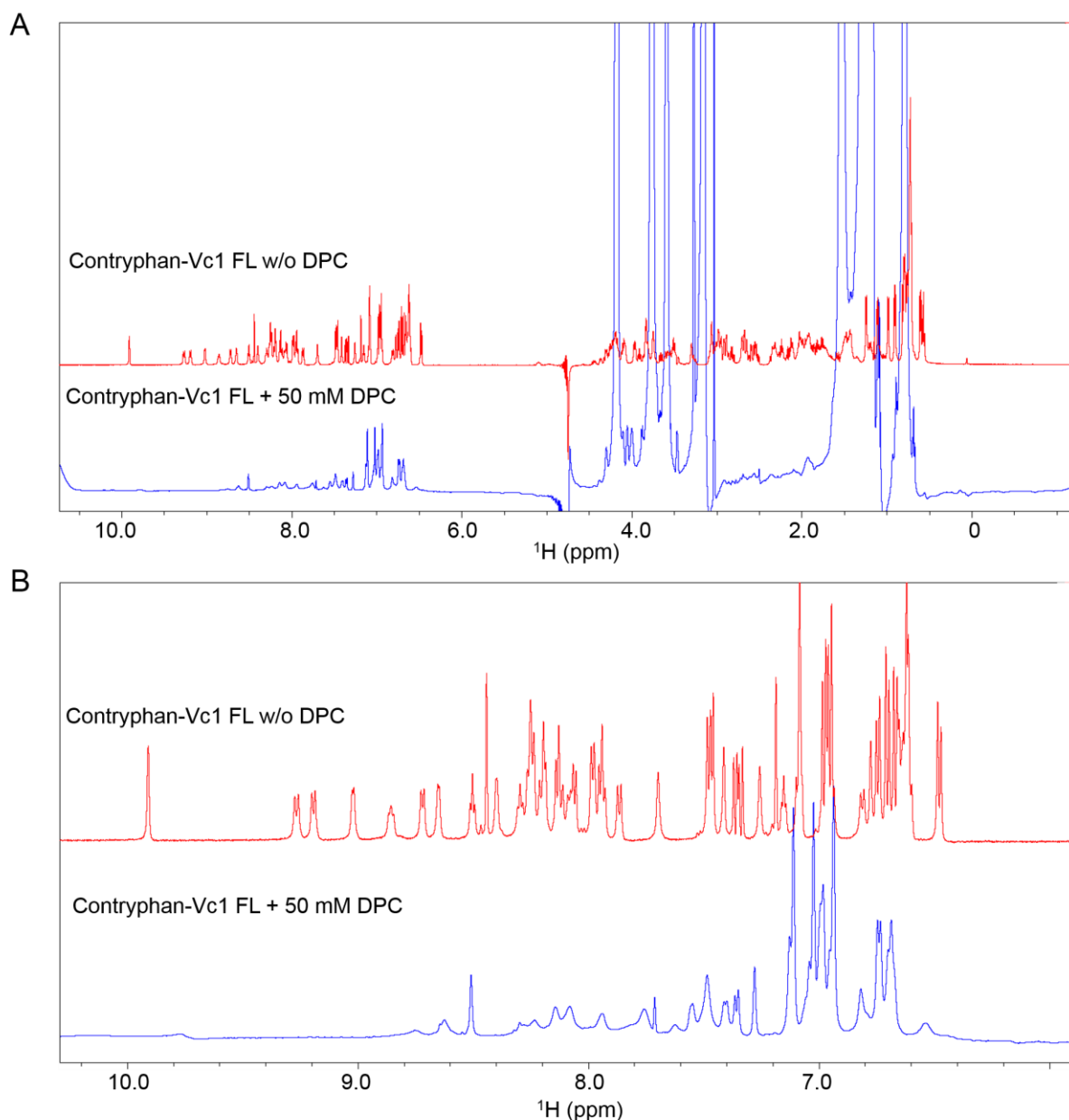


Figure 1. A. 1D ^1H NMR spectra of synthetic full-length contryphan-Vc1 overlayed with ^1H NMR spectra of synthetic full-length contryphan-Vc1 in presence of 50 mM DPC. Both the spectra were recorded at pH 4.0 and 20°C in water containing 7% $^2\text{H}_2\text{O}$. **B.** Amide and aromatic regions of spectra in A.

5.3.2 Interaction with DPC micelles is mediated by the structured part of contryphan-Vc1

In order to understand whether the core structure or the highly flexible *C*-terminus part of contryphan-Vc1 was interacting with the DPC micelles, we utilised con-Vc1₁₋₂₂[Z1Q] and performed titration with increasing concentrations of DPC up to 50 mM with intervals of 10 mM. At 50 mM DPC one-dimensional ^1H spectra of con-Vc1₁₋₂₂[Z1Q] also showed similar changes where the peaks between 8.8 and 10.0 ppm were in exchange. Peak broadening and

shifts were observed for many peaks between 8.8 and 6.5 ppm, similar to that of the full-length contryphan-Vc1, suggesting that the β -hairpin core of the contryphan-Vc1 was involved in the interaction with DPC micelles (**Figure 2**). Both two-dimensional TOCSY and NOESY spectra recorded in the presence of increasing concentrations of DPC micelles showed concentration-dependant changes in peaks, where the intensity of amide peaks of residues Ala8 and Tyr9 were considerably reduced at 20 mM DPC and are completely broadened at higher concentrations of DPC. (**Figure 3**).

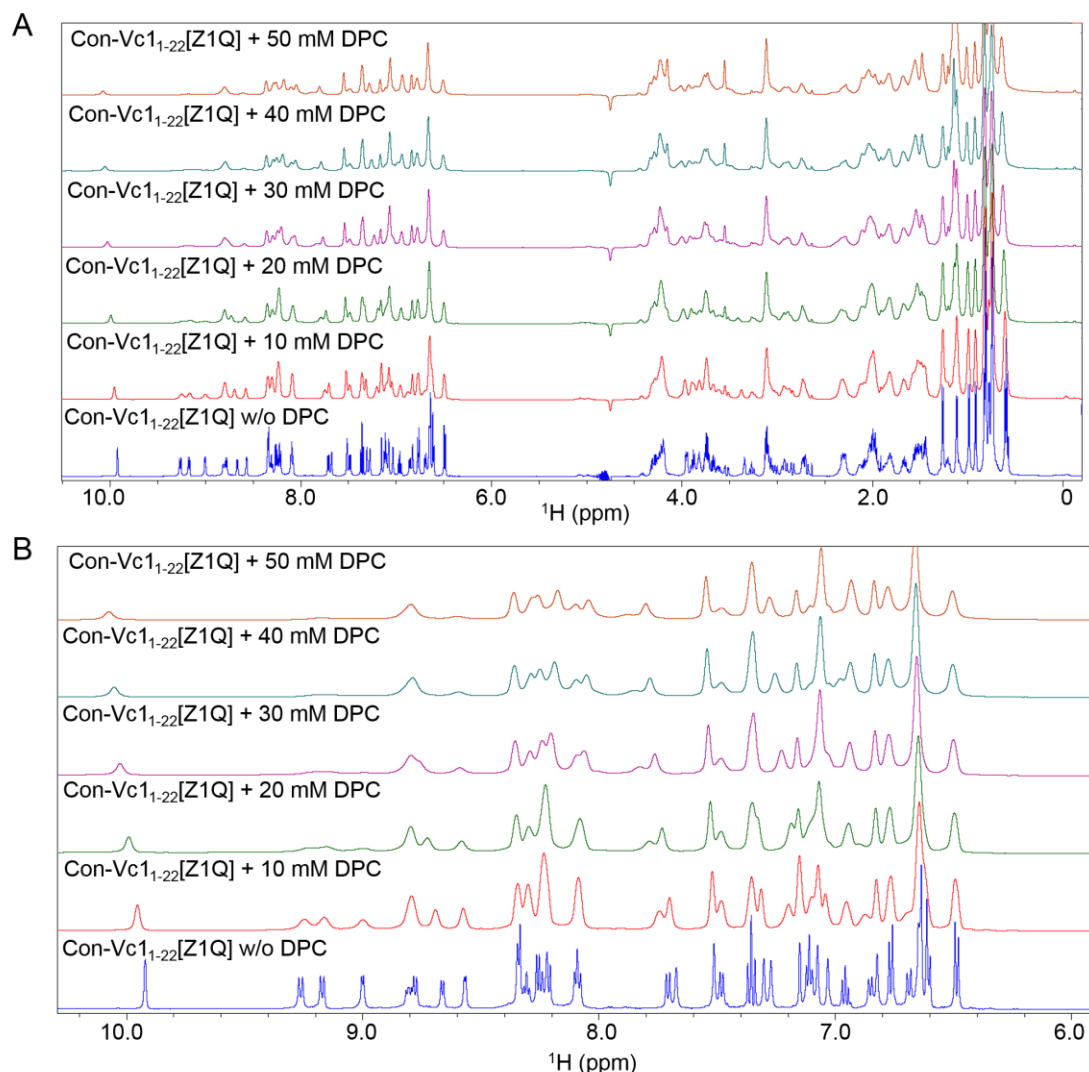


Figure 2. **A.** 1D ^1H NMR spectra of con-Vc1₁₋₂₂[Z1Q] recorded in the presence of increasing concentrations of DPC micelles ranging from 10 mM to 50 mM at pH 4.0 and 20°C in water containing 7% $^2\text{H}_2\text{O}$. **B.** 1D ^1H NMR spectra of con-Vc1₁₋₂₂[Z1Q] recorded in the presence of increasing concentrations of DPC micelles ranging from 10 mM to 50 mM at pH 4.0 and 40°C in water containing 7% $^2\text{H}_2\text{O}$.

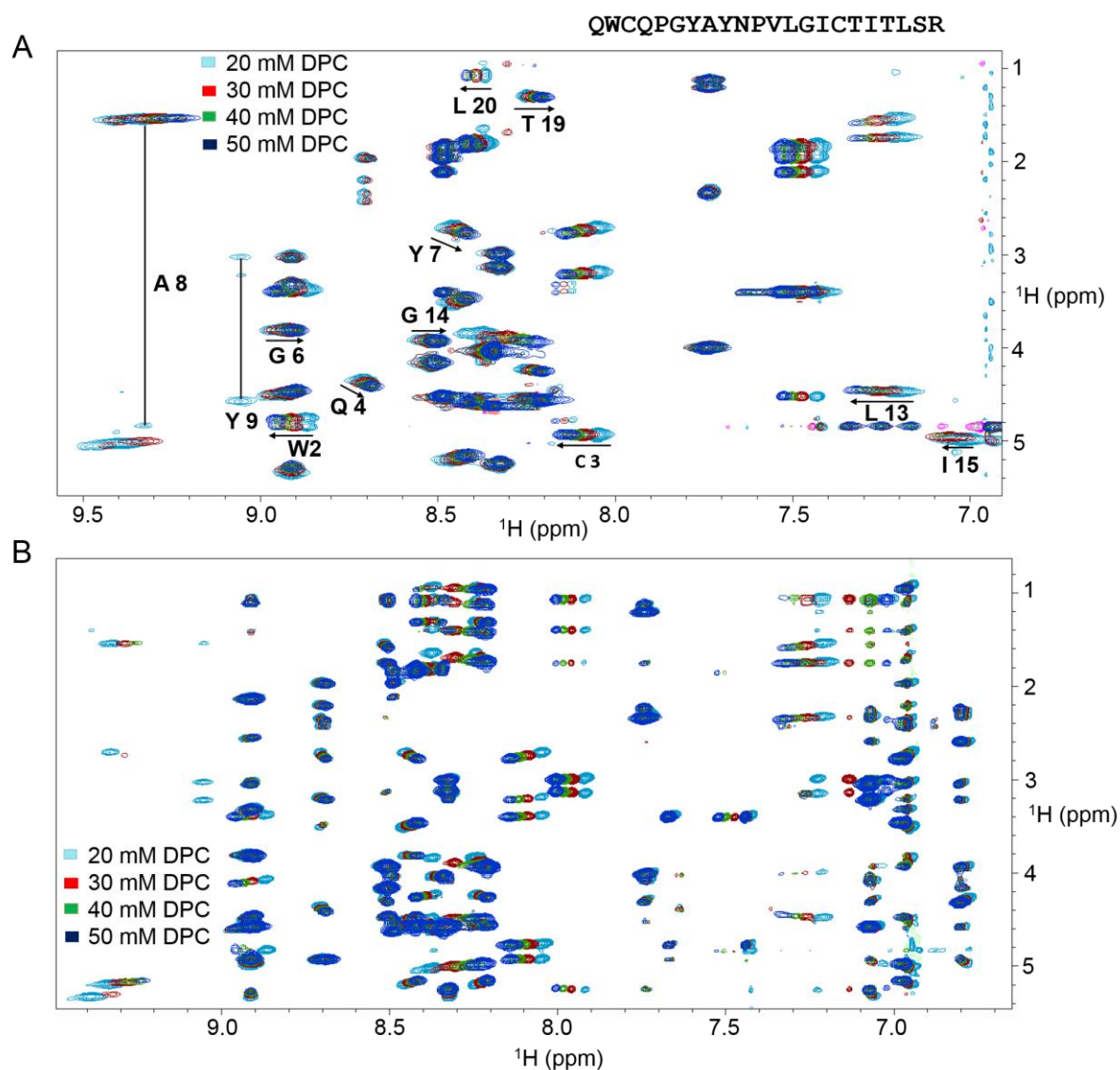


Figure 3. A. Two-dimensional ^1H NMR TOCSY spectra of con-Vc1-22[Z1Q] recorded in the presence of increasing concentrations of DPC ranging from 20 to 50 mM at pH 4.0 and 40°C in water containing 7% $^2\text{H}_2\text{O}$. Residues that showed considerable changes in the presence of DPC micelles are labelled. **B.** Two-dimensional ^1H NMR NOESY spectra of con-Vc1-22[Z1Q] recorded in the presence of increasing concentrations of DPC ranging from 20 to 50 mM at pH 4.0 and 40°C in water containing 7% $^2\text{H}_2\text{O}$.

5.3.3 Sequence-specific resonance assignments for contryphan-Vc1-22[Z1Q] in the presence of DPC micelles

Spin systems were identified through combined analysis of TOCSY (80 ms spin-lock time) and NOESY (200 ms mixing time) spectra acquired at 313 K and pH 4.0 in the presence of increasing concentrations of DPC micelles. Sequential resonance assignments were made utilizing two-dimensional COSY, TOCSY and NOESY spectra. Residues containing methyl groups such as Ala8, Thr17 & 19, Leu13 & 20, and Ile15 & 18 served as good starting points

for the relatively simple identification of spin systems and sequential assignment process. Complete backbone assignments were obtained for all spin systems (**Figure 3**).

Table 1. Chemical shifts for contryphan-Vc1₁₋₂₂ at pH 4.0, in presence of 50 mM DPC at 313K.

Residue	H ^N	H ^α	H ^β
Gln1	ND	4.106	3.40,3.54
Trp2	8.95	4.76	3.34
Cys3	8.14	4.93	2.77,3.21
Gln4	8.68	4.42	1.96,2.31
Pro5	-	4.57	2.55,2.09
Gly6	8.90	4.46,3.81	-
Tyr7	8.41	5.14	2.77,3.458
Ala8	9.22	4.79	1.53
Tyr9			
Asn10	8.32	5.23	3.12, 2.99
Pro11	-	4.31	2.51
Val12	7.74	4.004	2.34
Leu13	7.32	4.44	
Gly14	8.50	3.92,4.15	-
Ile15	7.07	4.92	2.10
Cys16	8.91	5.26	3.046,3.31
Thr17	9.25	4.977	4.45
Ile18	8.22	3.94	1.65
Thr19	8.20	4.55	4.26
Leu20	8.42	4.57	1.07
Ser21	8.34	4.61	4.03
Arg22	8.49	4.52	3.39

ND- Not determined

5.3.4 SDH fold of contryphan-Vc1 is maintained in the presence of DPC micelles

There are several instances where changes in conformation and oligomeric status of peptides and proteins were reported in the presence of DPC micelles.¹⁰³⁻¹⁰⁶ In order to understand the effect of DPC micelles on the SDH fold of the contryphan-Vc1 we have analysed both the amide and H^α secondary chemical shift plots of the con-Vc1₁₋₂₂[Z1Q] in presence of 50 mM DPC. Both amide and H^α secondary chemical shifts showed similar deviations from random coil values when compared to the secondary chemical shifts of con-Vc1₁₋₂₂[Z1Q] without DPC micelles, with minor differences at the *N*- and *C*-termini of the peptide, suggesting that the structure of contryphan-Vc1 is not altered significantly, however significant peak broadening for residues Ala8 and Tyr9 and subtle chemical shift differences for several residues warrants

an in depth analysis of the structure of contryphan-Vc1 in presence of DPC in order to understand the influence of DPC on the SDH fold of the contryphan-Vc1. (**Figure 4**).

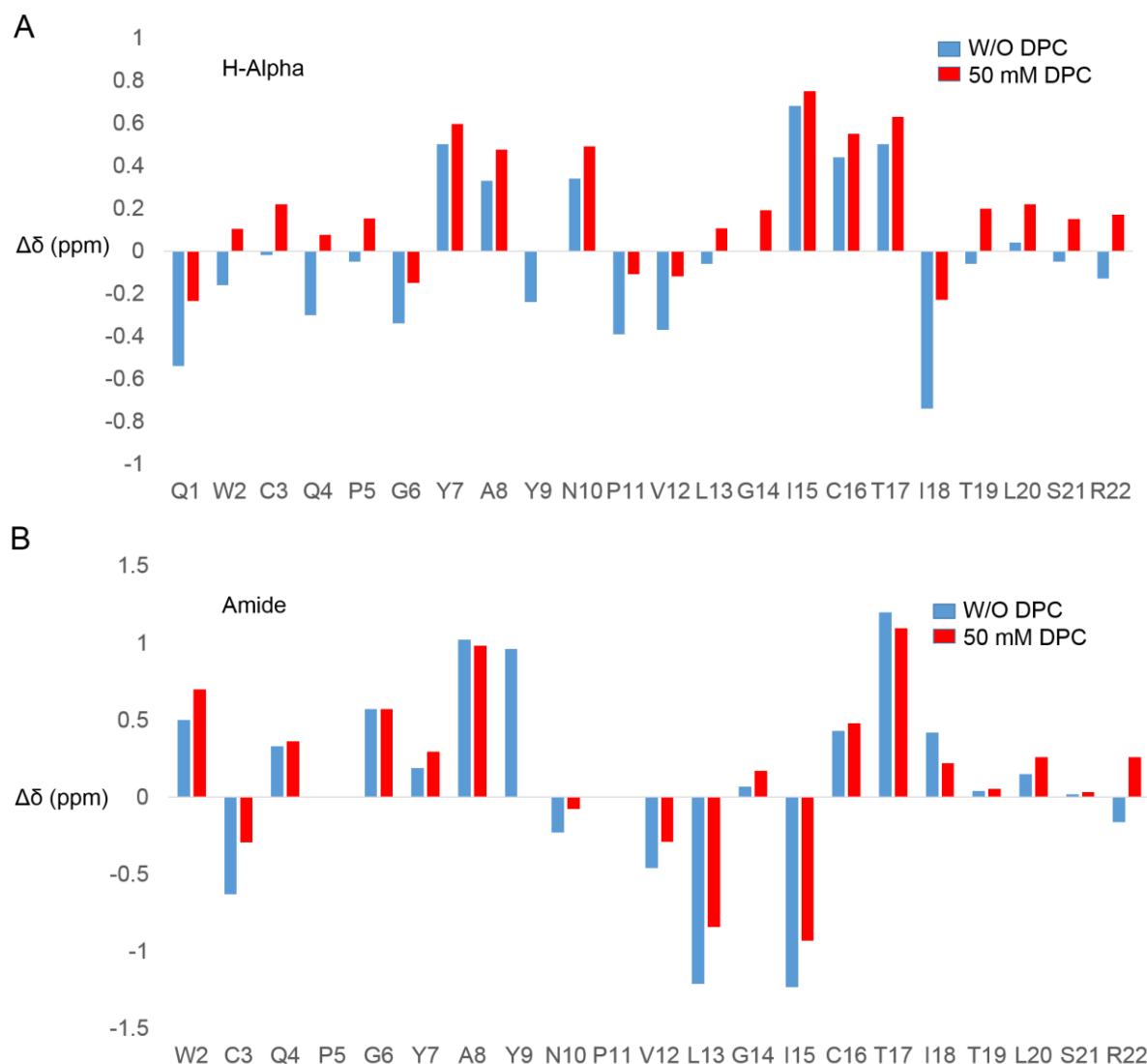


Figure 4. Secondary chemical shift difference plots for con-Vc1₁₋₂₂[Z1Q] in presence of 50 mM DPC and con-Vc1₁₋₂₂[Z1Q] without (W/O) DPC. **A.** H α secondary chemical shift plots of the con-Vc1₁₋₂₂[Z1Q] in presence of 50 mM DPC. **B.** HN secondary chemical shift plots of the con-Vc1₁₋₂₂[Z1Q] in presence of 50 mM DPC.

5.3.5 Contryphan-Vc1₁₋₂₂[Z1Q] binds weakly to POPC bilayers

DPC micelles are roughly spheres in solution above the critical micellar concentration of 1.0-1.5 mM.^{102, 107} Owing to the strong curvature of the micelles, hydrophobic peptides will more readily interact with DPC micelles than flat bilayers. POPC bilayers were used to study the interaction of contryphan-Vc1 towards bilayer-forming lipids. SPR sensorgrams for the interaction of the con-Vc1₁₋₂₂[Z1Q] to the POPC bilayers immobilized on the surface of a L1-biacore chip suggest that con-Vc1₁₋₂₂[Z1Q] interacts with the zwitterionic POPC with a weak affinity in the high μ M range, where a response is shown starting from 2.5 μ M of peptide

concentration (**Figure 5**). Surprisingly, the full-length contryphan-Vc1 did not show any response in when passed over the POPC bilayers, suggesting that the full-length contryphan-Vc1 does not interact with the bilayer forming POPC lipid.

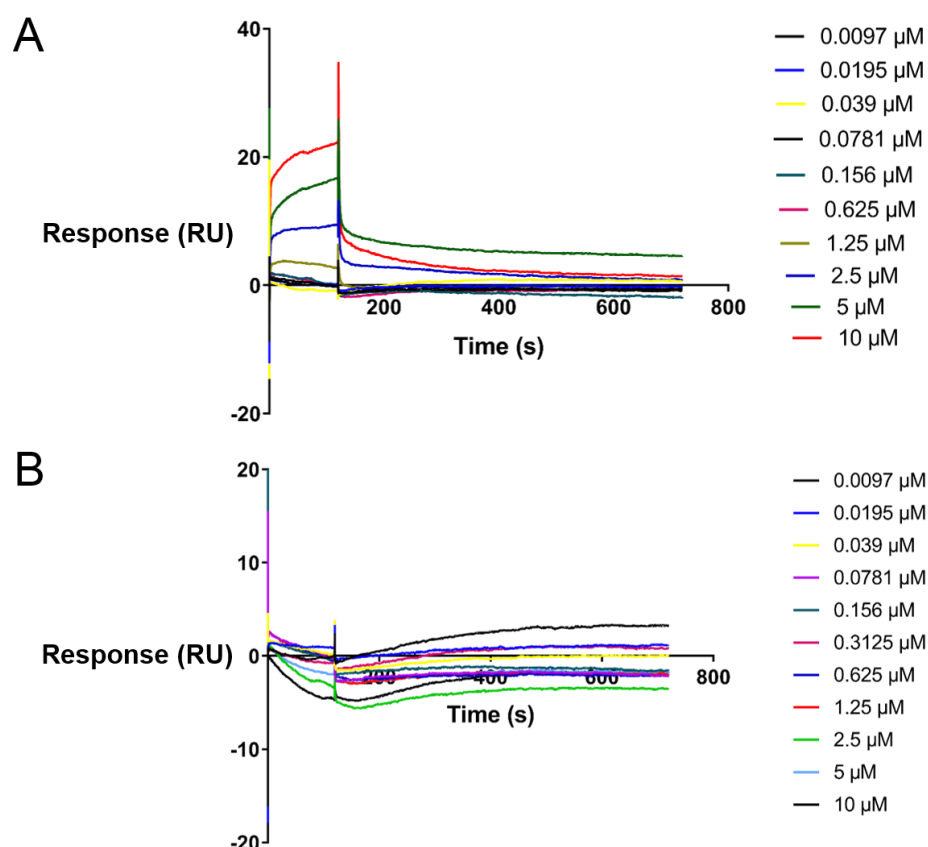


Figure 5. SPR sensorgrams of the interactions between immobilized POPC lipid on L1-biacore chip and the peptides con-Vc1₁₋₂₂[Z1Q] and full-length contryphan-Vc1. **A.** Con-Vc1₁₋₂₂[Z1Q]. **B.** Full-length contryphan-Vc1.

5.4 Discussion

In this study the lipid binding property of contryphan-Vc1 is studied utilising DPC micelles as a model membrane system. One-dimensional ^1H spectra of full-length contryphan-Vc1 acquired in the presence of 50 mM DPC showed significant changes when compared to spectra recorded in the absence of DPC micelles, suggesting that contryphan-Vc1 does interact with DPC micelles. Many studies have been shown that DPC micelles induce some unstructured peptides to adopt helical conformation.¹⁰⁸ Unstructured plantaricin peptides from *Lactobacillus plantarum* C11 have been shown to adopt a helical conformation in presence of DPC micelles.¹⁰⁸ DPC micelles also has been shown to cause a switch between a β -strand and an α -helix.^{109, 110}

In order to determine whether contryphan-Vc1 interacts with the DPC micelles through its unstructured C-terminus or the structured region we have used the truncated con-Vc1₁₋₂₂[Z1Q] peptide and performed a concentration-dependent titration with DPC micelles. Similar peak broadening and peak shifts to those seen in the spectra of full-length contryphan-Vc1 in the presence of DPC micelles suggested that the interaction of contryphan-Vc1 with DPC micelles is mediated by the structured region. In order to observe the influence of DPC micelles on the SDH fold of the peptide we have analyzed the secondary chemical shift plots for both H^{α} and H^N . Similar deviation from the random coil values for con-Vc1₁₋₂₂[Z1Q] when compared to the contryphan-Vc1 in absence of DPC micelles observed in the amide and 1H secondary chemical shift plots suggests that the structure of contryphan-Vc1 is not altered significantly, however significant peak broadening for residues Ala8 and Tyr9 and subtle chemical shift differences for several residues warrants an in depth analysis of the structure of contryphan-Vc1 in presence of DPC.⁶⁸

We also investigated the interaction of contryphan-Vc1 with the bilayer-forming lipid POPC. SPR sensorgrams for the interaction of the con-Vc1₁₋₂₂[Z1Q] to the POPC bilayers suggest that con-Vc1₁₋₂₂[Z1Q] interacts with the zwitterionic POPC, albeit with a weak affinity in the high μM range. Surprisingly, the full-length contryphan-Vc1 did not show any response in the presence of POPC lipid, suggesting the full-length contryphan-Vc1 does not interact with POPC bilayers. POPC is shown to form vesicles in the diameter of 20–28 nm range, which falls in the range expected for small unilamellar vesicles. The existence of separate inner and outer choline $N(CH_3)^+$ methyl signals at 3.2 ppm also provided positive evidence for the formation of vesicles.¹¹¹ It is expected that the partitioning of the peptides in POPC will be considerably less when compared to the highly curved DPC micelles as the surface of the vesicles of POPC is quite flat, but the lack of interaction of full-length contryphan-Vc1 with POPC vesicles is quite surprising. It will be interesting to test whether contryphan-Vc1 binds to other anionic vesicles or a mixture of zwitterionic and anionic lipids. Analyzing the structure of the contryphan-Vc1 in the presence of DPC using ^{15}N - ^{13}C double labelled peptide will help in understanding the influence of DPC on the SDH fold. Further characterization of lipid interactions of contryphan-Vc1 utilizing various spin-labelled lipids,¹¹² and obtaining diffusion coefficient rates in different lipids will shed more light on the membrane binding and membrane partitioning ability of contryphan-Vc1.

Chapter 6

Understanding the Roles of Key Residues in Maintenance of the SDH Fold

6.1 Introduction

The single disulfide-directed β -hairpin (SDH) fold, first observed in contryphan-Vc1, is a novel and unique peptide fold which represents an ideal privileged peptide scaffold.^{47, 68} The SDH core structure is a double-stranded anti-parallel β -sheet stabilised by just a single disulfide bridge as observed in contryphan-Vc1.^{47, 68} The SDH core is common to other multiple disulfide-containing peptide folds such as the inhibitory cystine knot fold (ICK), disulfide-directed β hairpin (DDH) fold and boundless β hairpin (BBH) fold.^{15, 47, 101, 113} The SDH fold is a small, highly stable, elementary independent folding unit, with potential as a new peptide scaffold where one of the loops of the SDH core of the contryphan-Vc1 is amenable to replacement by a short bioactive NNN epitope.⁶⁸

Two predicted proteins CGI_10002631 and LOC101860216 with relatively high sequence identities (46% and 41%), and sequence similarity for residues predicted to be structurally important for contryphan-Vc1, were identified from the genomes of the non-venomous molluscs *Crassostrea gigas* (Pacific oyster) and *Aplysia californica* (Californian sea hare), respectively. Both sequences are clearly prepropeptides, with a 20-residue secretory signal peptide, followed by a propeptide region, a dibasic cleavage site, and one or more mature peptides encoded near the C terminus. In CGI_10002631 from *C. gigas*, the predicted mature peptide is 32 residues in length and, as with contryphan-Vc1, is followed by a dibasic cleavage site and a three-residue propeptide region. The sequence LOC101860216 from *A. californica* encodes two distinct mature peptides (29 and 30 residues in length) separated by an additional dibasic cleavage site.⁴⁷

Homologous sequences are expected to have similar structures, and potentially similar functions as well. Structural and functional characterisation of homologous sequences for a given protein or a peptide can provide valuable insights into the structure and function of all the members of the similar family.¹¹⁴ In this study we have mined different cone snail transcriptomic data and UniProt/GenBank for sequences similar to contryphan-Vc1, we have made an analogue of the truncated contryphan-Vc1 (con-Vc1₁₋₂₂[Z1Q]) where both cysteines, Cys3 and Cys16 were replaced with Ser residues in order to study the effect of the disulfide bond on the structure of contryphan-Vc1, and synthesised the mature peptide CGI_10002631 of the *C. gigas* homologous to contryphan-Vc1 and characterised the peptide to further understand the SDH fold.

6.2 Materials and methods

6.2.1 Peptide synthesis

The procedure for peptide synthesis used in this study is described in Chapter 2 under section 2.4. Here methods specific for synthesis of peptides used in this study are described briefly. The peptides full-length contryphan-cg1, contryphan-Vc1₁₋₂₂ [C3S, C16S] were synthesised using Rink Amide AM resin (0.47 mmol/g loading) on PTI Instruments PS3 peptide synthesiser as described in section 2.4. A double mutant of contryphan-cg1₁₋₂₂, contryphan-cg1₁₋₂₂[M9Y, P18I] and quad mutant contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] were synthesised on PTI Instruments PS3 peptide synthesiser using chloro-trityl (CTC) resin. CTC resin (0.7 mmol/g loading) was weighed into a reaction vessel and swollen in dichloromethane (DCM) for 30 min and coupled manually with the last amino acid in the sequence by dissolving two equivalents of the amino acid in a mixture of DCM and DMF (2.5 ml+0.5 ml) and 87.5 μ l of DIPEA for over an hour. The uncoupled resin was then capped with methanol (0.8 ml/gram resin) for 30 min and the resin washed with an excess of DCM and then an excess of DMF. Cleavage of the peptide from the resin and folding were as described in section 2.4.

6.2.2 Sample preparation and NMR spectroscopy

All spectra were acquired on a Bruker 600 MHz spectrometer equipped with a cryogenically-cooled triple-resonance probe. Lyophilised peptide was dissolved in either 93% H₂O/7% ²H₂O or 100% ²H₂O (pH 4.0). One-dimensional ¹H spectra of contryphan-Vc1₁₋₂₂ [C3S,C16S], full-length contryphan-cg1 and the truncated analogues, double mutant of contryphan-cg1₁₋₂₂, contryphan-cg1₁₋₂₂[M9Y, P18I] and quad mutant contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] were acquired at 20°C, at pH 4.0. Two-dimensional NMR spectra for sequence-specific assignments were acquired at pH 4 and 20°C. Two-dimensional homonuclear TOCSY spectra with a spin lock time of 80 ms were acquired using the DIPSI-2 pulse sequence with excitation sculpting for water suppression. Similarly, two-dimensional NOESY spectra were acquired with 200 ms mixing time. A DQF-COSY spectrum was acquired in 100% ²H₂O for measuring *J*-couplings. ¹³C-HSQC and ¹⁵N-HSQC spectra were acquired for carbon and nitrogen chemical shifts, respectively. A sine-bell squared window function was used for processing spectra. All spectra were processed using Bruker TopSpin (Version 3.2).

6.3 Results

6.3.1 Identification of contryphan-Vc1 homologous sequences

A search of UniProt/GenBank using BLASTp for sequences similar to mature contryphan-Vc1 and careful examination of transcriptomic data obtained on different cone snails yielded an additional six sequences homologous to contryphan-Vc1 apart from the existing three

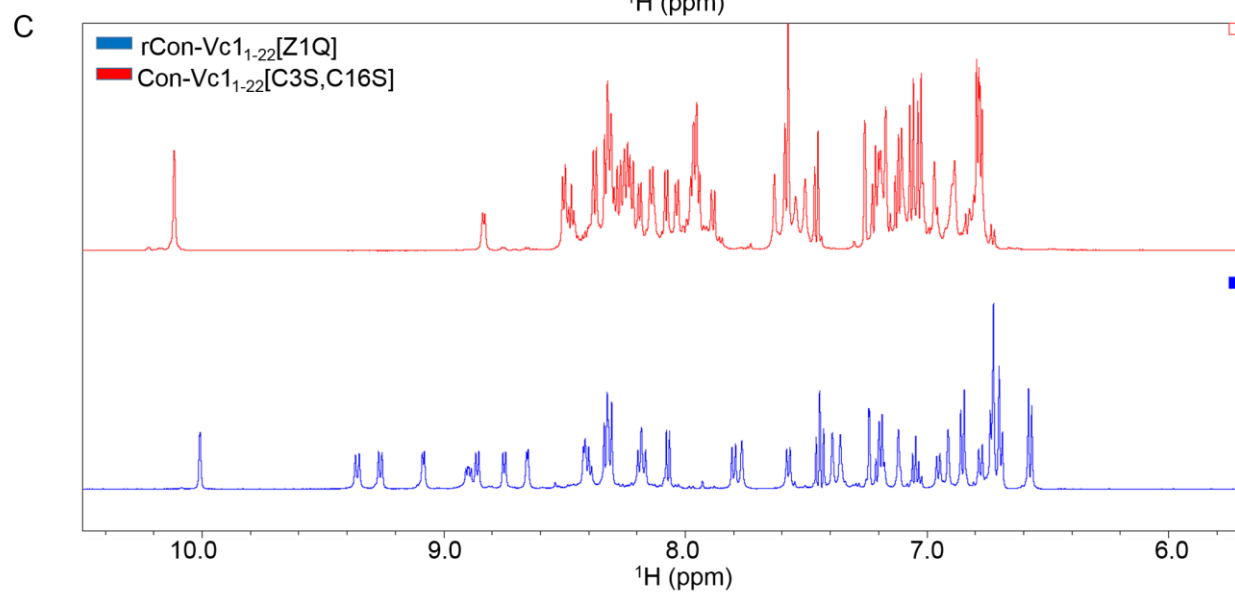
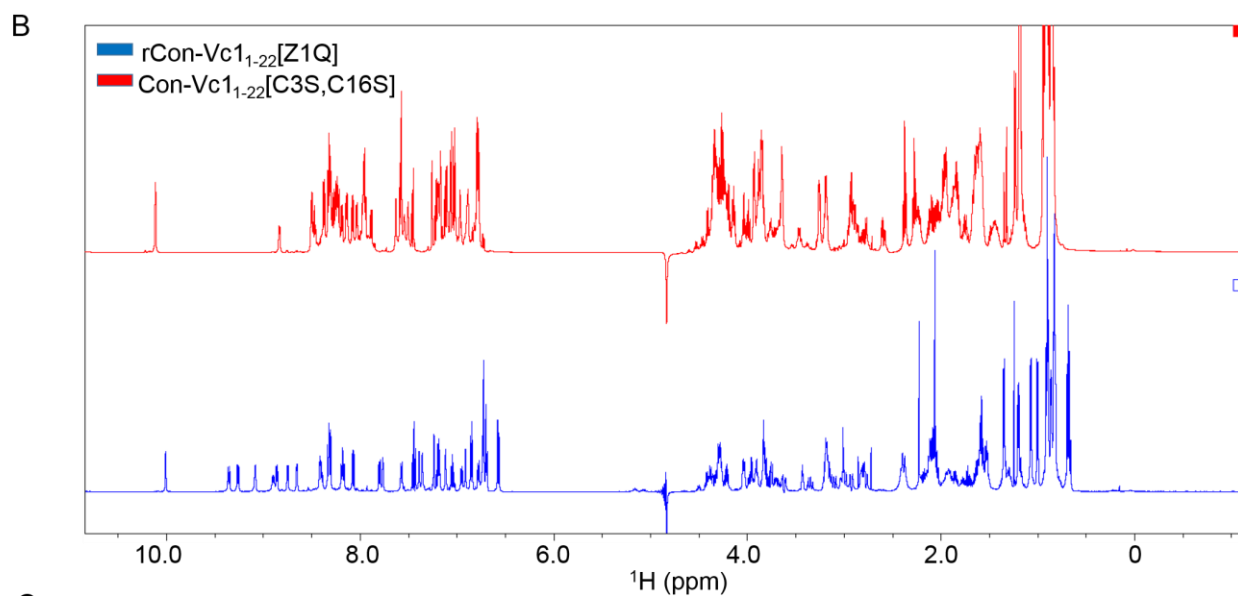
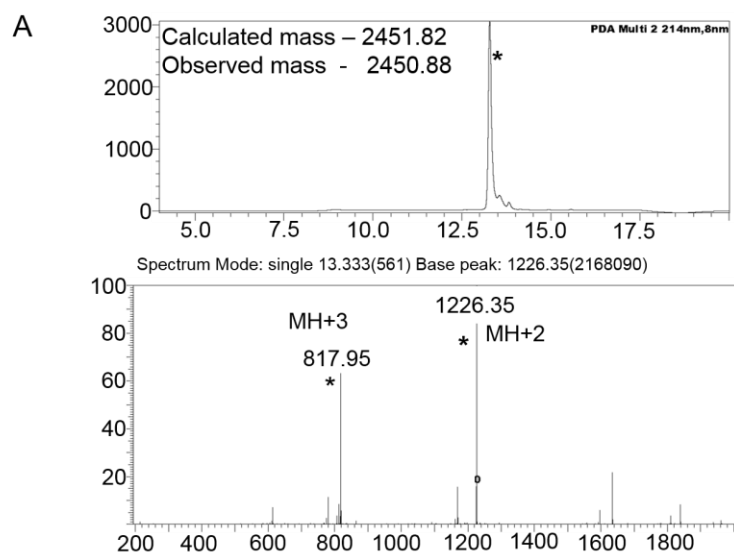
sequences.⁴⁷ Three of the nine sequences are from the venom glands of mollusc-hunting cone snails (*victoriae*, *ammiralis*, *crocatus*). Another two sequences are from worm-hunting cone snails (*lenavati* and *tessulatus*) and another sequence is from the nervous system of a mollusc-hunting cone snail (*textileNR*). The mature peptide sequences are aligned and the residues that are conserved are coloured based on the extent of the conservation across the available nine sequences (**Figure 1**).

<i>A. californica</i> (LOC101860216)	·G W C R P N M T F N S I L G R C T Y V Y S K L K G G R G K R -----
<i>C. gigas</i> (CGI_10002631)	·G W C R P G F T M N P V L N R C V P T M R S L I K R N Y N R F R R F G K -
<i>C. lenavati</i> (A0A0K8TUR9)	·Q W C R R G M A Y N P A L G T C T F S V A A M R G R G R K Y R--RH---
<i>C. textileNR2</i>	·Q W C R R G M A Y N P A L G T C T F S V A A M R G R G R --KYRRH---
<i>C. textileNR</i>	·Q W C R S G M S Y N P V L G T C T L S L A A L R G R G R S F R----GV-
<i>C. tessulatus</i>	·Q W C R S G M S F N P V L G T C T L S L A A L R G R G R S F R----GV-
<i>C. crocatus</i>	·Q W C Q P G W A Y N P V L G T C I K S L A I K Y P G L Y E N --SRGHQ
<i>C. victoriae</i> (W4VSF6)	·Q W C Q P G Y A Y N P V L G I C T I T L S R I E H P G N Y D Y--RRGRQ
<i>C. ammiralis</i>	·Q W C R P G W A Y N P A L G R C T I S K S R I E H P G N Y E Y--RRGRQ

Figure 1. Sequence alignment of contryphan-Vc1 homologous sequences identified from different cone snails and non-venomous molluscs like *Aplysia californica* and *Crassostrea gigas*. Residues that are completely conserved throughout the nine sequences are coloured red and the residues that are fairly conserved are coloured green, blue and purple. (CGI_10002631- contryphan-cg1 from *Crassostrea gigas*, LOC101860216- contryphan-ac1 from *Aplysia californica*, A0A0K8TUR9 – contryphan-li from *conus lenavati*, W4VSF6- contryphan-Vc1 from *conus victoriae*.)

6.3.2 Disulfide bond is critical for the folding and maintenance of SDH fold

Contryphan-Vc1₁₋₂₂ [C3S, C16S], an analogue of Con-Vc1₁₋₂₂[Z1Q] where both Cys 3 and 16 were replaced with Ser, was made to analyse the importance of the disulfide bond in the folding and maintenance of the SDH fold of the contryphan-Vc1. One-dimensional ¹H spectra of Con-Vc1₁₋₂₂ [C3S, C16S] recorded at 20°C at pH 4.0 showed that the peptide was not folded as the peak dispersion is lost when compared to the native contryphan-Vc1 (**Figure 2B & C**). Both amide and ¹³C secondary chemical shift plots confirmed that the contryphan-Vc1₁₋₂₂ [C3S, C16S] had not formed the SDH fold (**Figure 2D**).



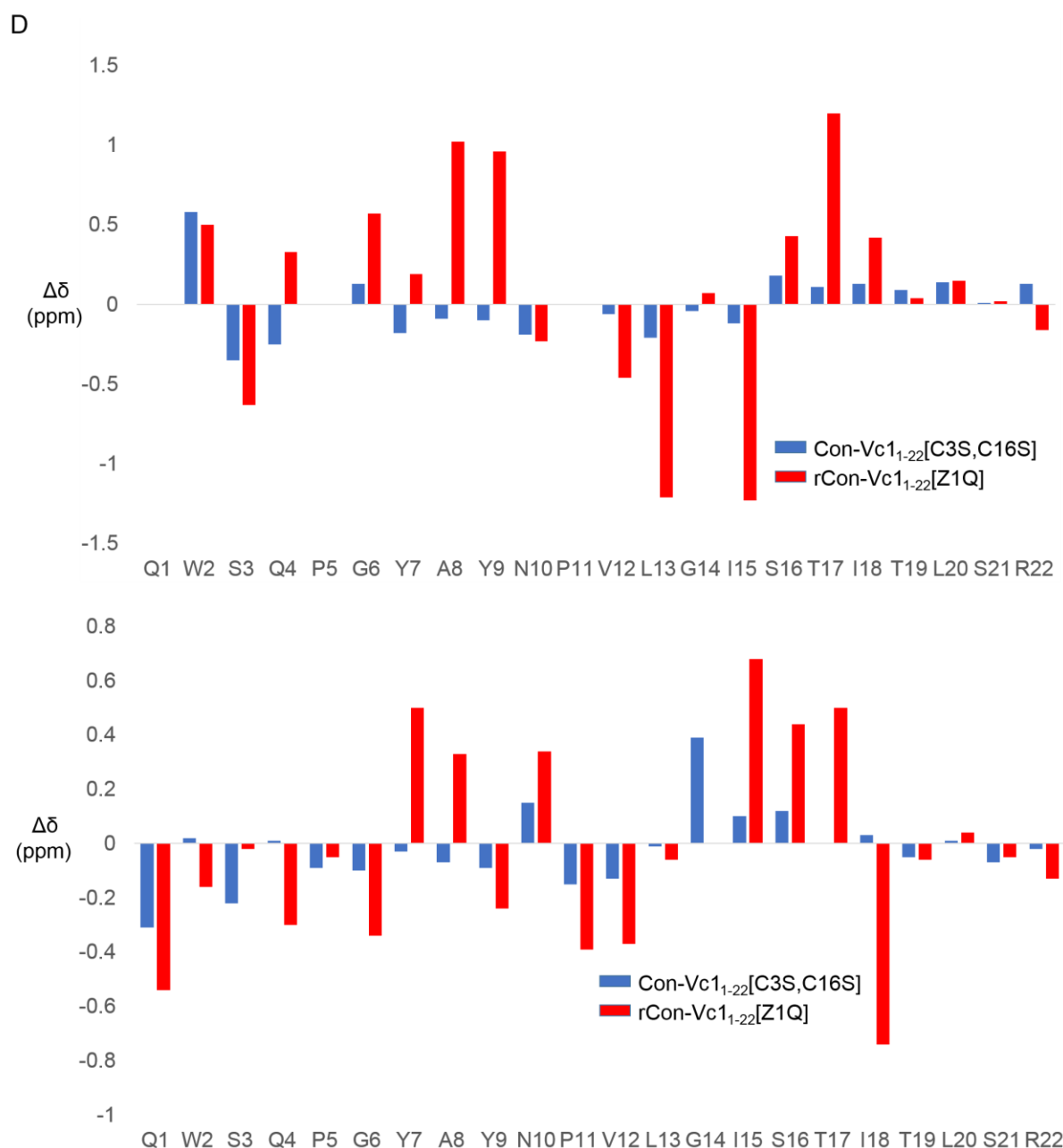


Figure 2. **A.** LC-MS profile of contryphan-Vc1₁₋₂₂ [C3S, C16S]. **B.** Comparison of one-dimensional ¹H NMR spectra of contryphan-Vc1₁₋₂₂ [C3S, C16S] at 20°C in water containing 7% ²H₂O with the ¹H NMR spectra of contryphan-Vc1₁₋₂₂. **C.** Amide and aromatic region of spectra in B. **D.** Comparison of H^α and H^N shift deviations from random coil chemical shifts¹¹⁵ between rCon-Vc1₁₋₂₂[Z1Q] and contryphan-Vc1₁₋₂₂ [C3S, C16S].

6.3.3 Full-length contryphan-cg1 is unfolded

Full-length contryphan-cg1 was made using solid-phase peptide synthesis and the mass and purity were checked by LC-MS (**Figure 3A**). One-dimensional ¹H spectra of full-length contryphan-cg1 recorded at 20°C at pH 4.0 confirmed that the peptide was unfolded, as evident from the lack of peak dispersion of the spectra when compared to the spectrum of native contryphan-Vc1 (**Figure 2B**).

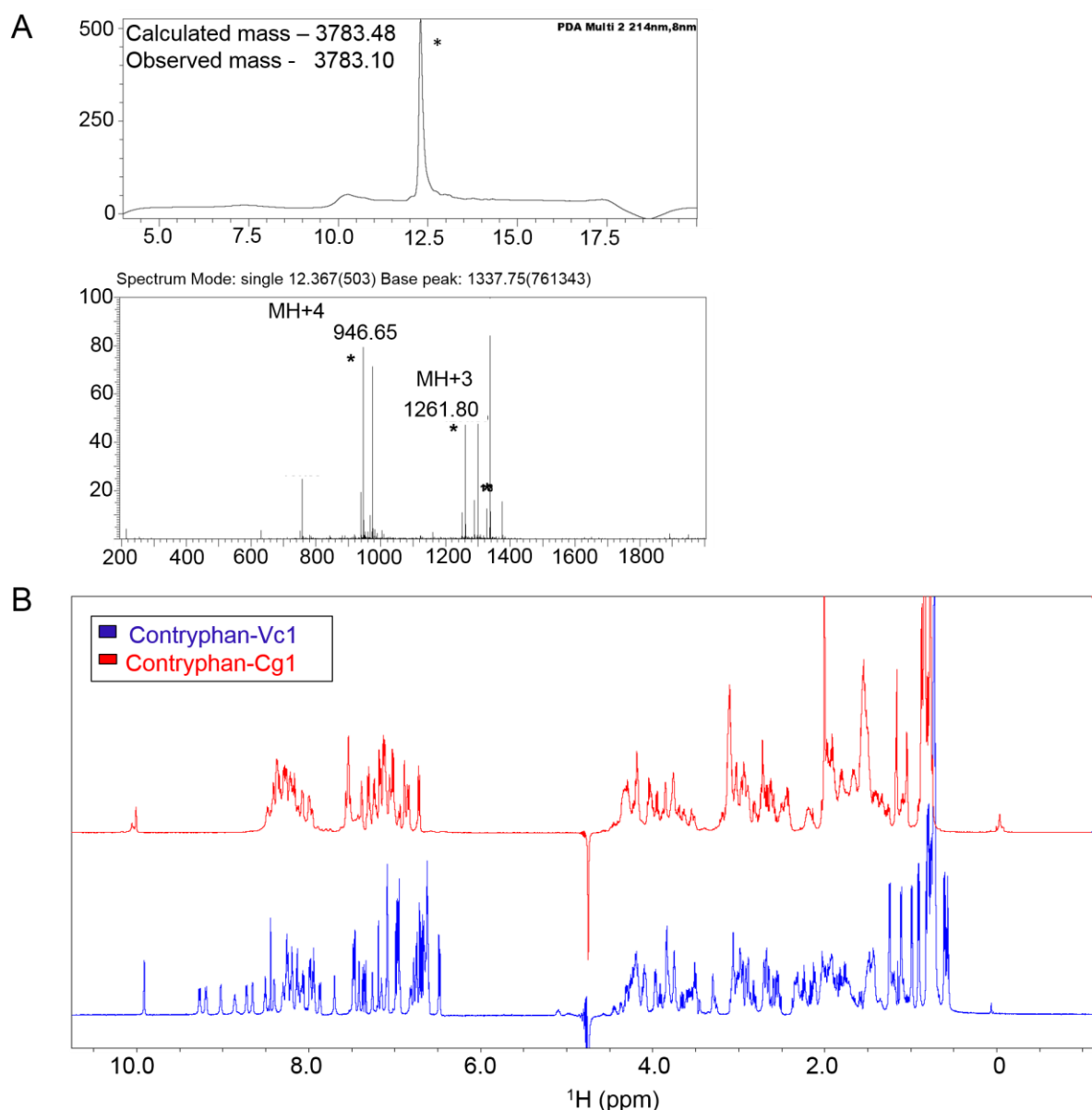
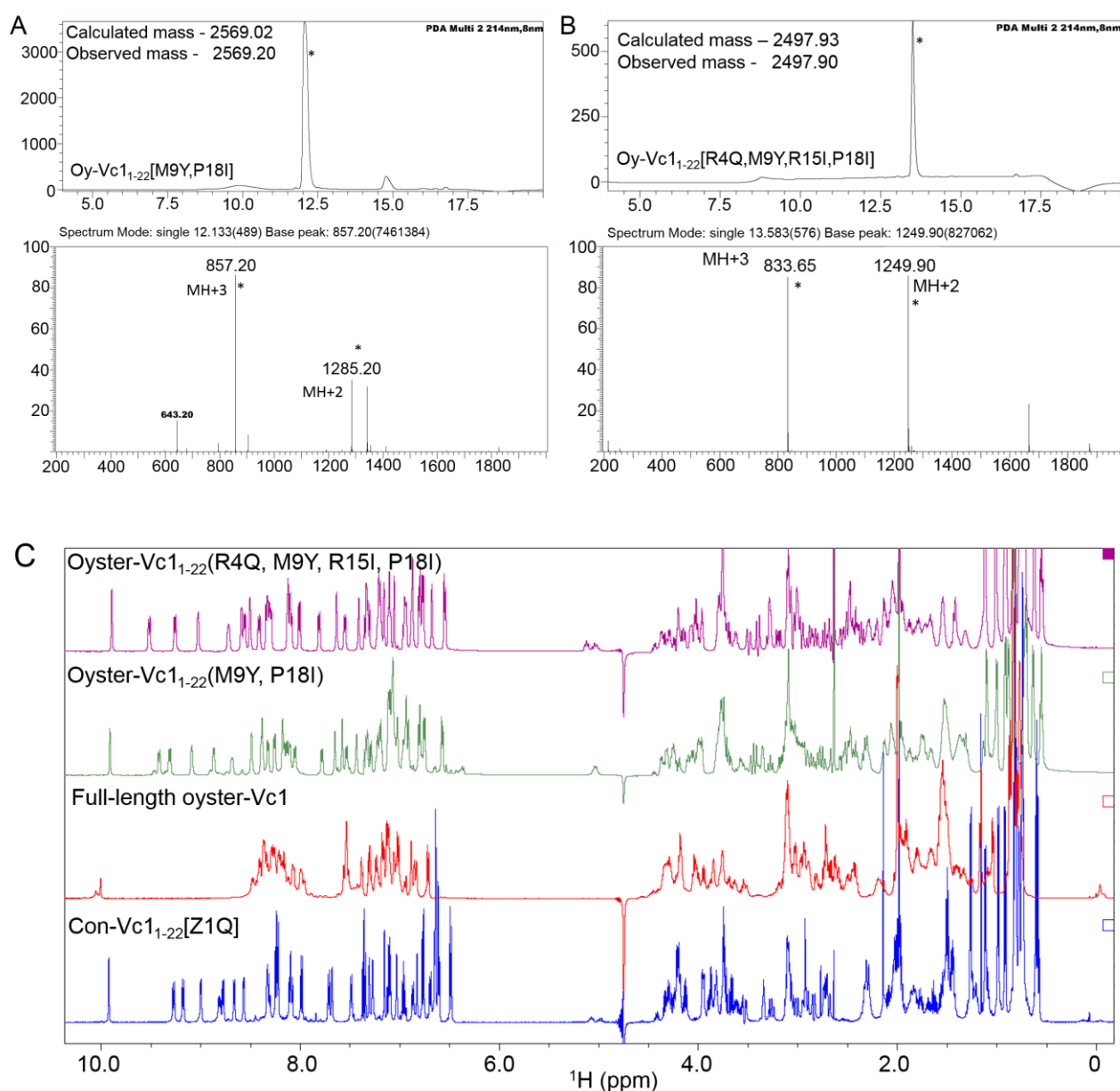


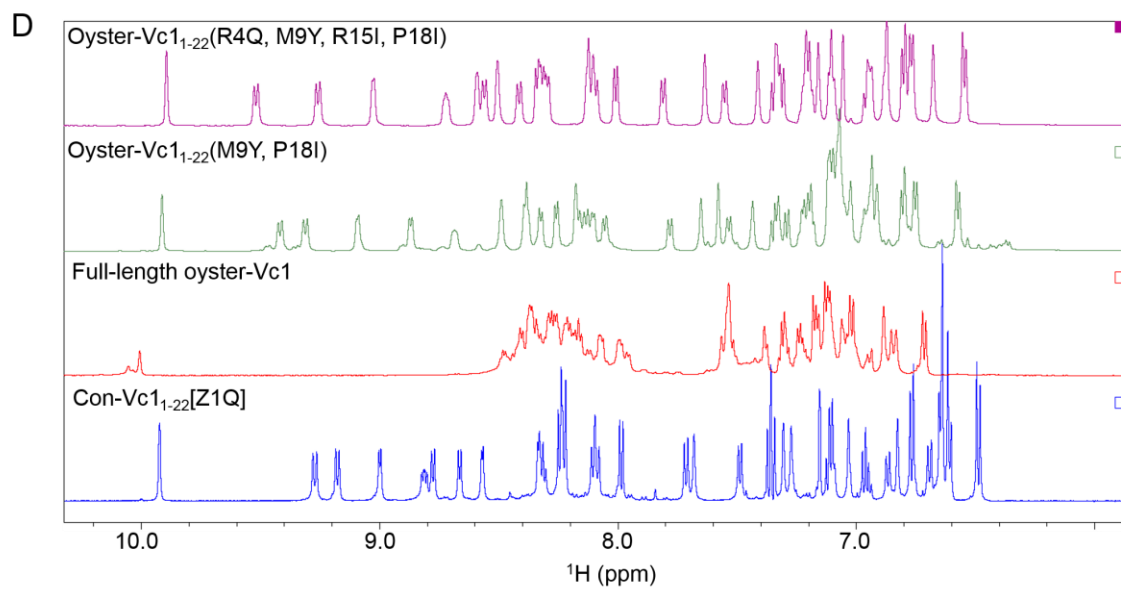
Figure 3. A. LC-MS profile of full-length contryphan-cg1 **B.** Comparison of one-dimensional ¹H NMR spectrum of full-length contryphan-cg1 at 20°C in water containing 7% ²H₂O with ¹H NMR spectrum of contryphan-Vc1.

6.3.4 Analogues contryphan-cg1₁₋₂₂[M9Y, P18I] and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] restore the SDH fold

In order to check the influence of the residues Met and Pro on the folding of the peptide, a double mutant of contryphan-cg1₁₋₂₂, contryphan-cg1₁₋₂₂[M9Y, P18I] is made where both Met and Pro were replaced with Tyr and Ile as observed in contryphan-Vc1. Another analogue a quad mutant, contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I], was made to check the influence of positively charged Arg4 and 15 residues on the SDH fold. Both Arg4 and 15 were replaced with Gln and Ile, respectively, as found in contryphan-Vc1. Both contryphan-cg1 analogues were made using solid-phase peptide synthesis and their mass and purity were checked by LC-

MS. One-dimensional ^1H spectra of contryphan-cg1₁₋₂₂[M9Y, P18I] and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] recorded at 20°C at pH 4.0 suggest that both of the analogues are folded based on the well-dispersed peaks observed in the spectra. Both amide and $^{\alpha}\text{H}$ secondary chemical shift plots showed similar deviations from random coil values when compared to the secondary chemical shifts of con-Vc1₁₋₂₂[Z1Q], confirming that the both the contryphan-cg1 analogue adopt the SDH fold.





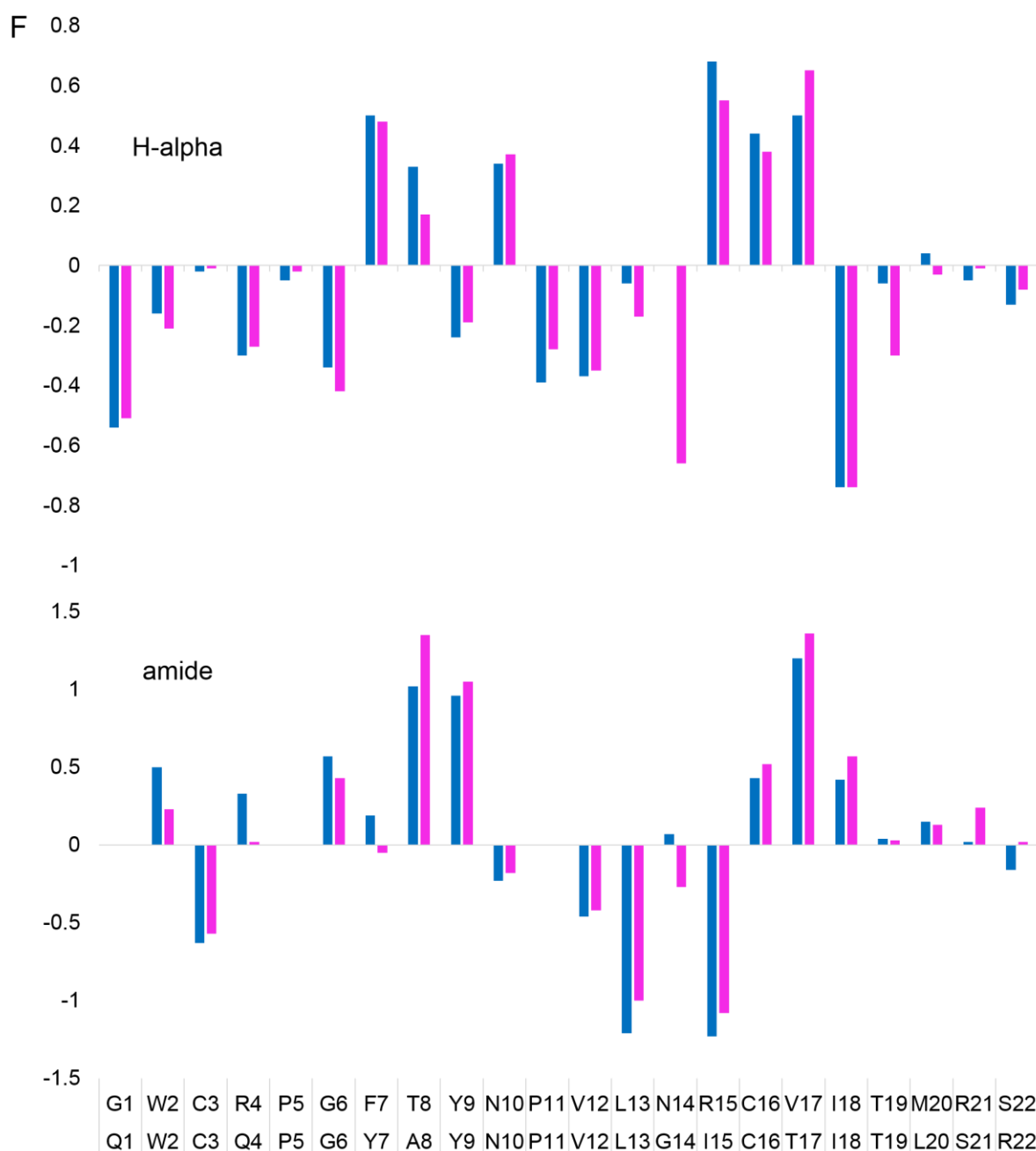


Figure 4. **A.** LC-MS profile of contryphan-cg1₁₋₂₂[M9Y, P18I]. **B.** LC-MS profile of and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I]. **C.** Comparison of one-dimensional ¹H NMR spectra of contryphan-cg1₁₋₂₂(M9Y, P18I) and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I] at 20°C in water containing 7% ²H₂O with the ¹H NMR spectra of full-length contryphan-cg1 and contryphan-Vc1₁₋₂₂. **D.** Amide and aromatic region of spectra in C. **E.** Comparison of H^α and H^N shift deviations from random coil chemical shifts¹¹⁵ between rCon-Vc1₁₋₂₂[Z1Q] and contryphan-cg1₁₋₂₂[M9Y, P18I]. **F.** Comparison of H^α and H^N shift deviations from random coil chemical shifts¹¹⁵ between rCon-Vc1₁₋₂₂[Z1Q] and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, P18I].

6.4 Discussion

In this study I performed a database search for sequences similar to that of contryphan-Vc1 in order to further understand the SDH fold of the contryphan-Vc1 and the possible functions the

peptides might carry out. Sequence alignment and database searching are essential tools in biology because a protein's structure and function can often be inferred from homologous proteins or peptides.¹¹⁶ A search of databases like UniProt and GenBank using BLASTp has previously helped in identifying two predicted proteins CGI_10002631 and LOC101860216 with relatively high sequence identities (46% and 41%), and sequence similarity for residues predicted to be structurally important for contryphan-Vc1, from the genomes of the non-venomous molluscs *Crassostrea gigas* (Pacific oyster) and *Aplysia californica* (Californian sea hare), respectively.⁴⁷ Further data mining utilising transcriptomic sequences of different cone snails yielded an additional six sequences that are homologous to the contryphan-Vc1. Three of the nine sequences are from the venom glands of mollusc-hunting cone snails (*victoriae*, *ammiralis*, *crocatus*). Another two sequences are from worm-hunting cone snails (*lenavati* and *tessulatus*) and another sequence is from the mollusc-hunting cone snail (textileNR).

Disulfide bonds are often responsible for stabilising the structure of a protein or a peptide, and therefore have an important role in protein folding and stability.¹¹⁷⁻¹¹⁹ In order to check the importance of the disulfide bond for maintenance of the SDH fold, we created an analogue of contryphan-Vc1₁₋₂₂ where both the cysteines were replaced with serine. Serine is often used to replace cysteine as they differ only in the swap of a sulfur atom with an oxygen, so the side chain remains the same size and retains its hydrophilic properties, while losing only its redox activity. Both ¹H spectra of Con-Vc1₁₋₂₂[C3S, C16S] and the amide and ⁹H secondary chemical shift plots confirmed that the contryphan-Vc1₁₋₂₂ [C3S, C16S] did not maintain the SDH fold of native contryphan-Vc1, suggesting the importance of the disulfide bond in stabilising the SDH fold.

The predicted protein CGI_10002631, from the genome of the non-venomous mollusc *Crassostrea gigas* (Pacific oyster), showed 46% similarity with contryphan-Vc1.⁴⁷ In order to check whether the mature peptide sequence from the Pacific oyster adopted the SDH fold, the full-length contryphan-cg1 peptide was made using solid-phase peptide synthesis. Surprisingly, one-dimensional ¹H spectra of full-length contryphan-cg1 peptide showed that the peptide was poorly folded. Close observation of the sequence of contryphan-cg1 showed the presence of a Met in the sequence rather than Tyr or Phe in other homologous sequences. Contryphan-cg1 also had a Pro where Leu or Ile the preferred residues. The sequence of contryphan-cg1 was also found to be highly cationic and is Arg-rich, with a predicted pI of 12.0.

In order to check the influence of the Met and Pro on the folding of the peptide, we replaced them with Tyr and Ile as observed in contryphan-Vc1. Another analogue was made where Arg4 and 15 were replaced with Gln and Ile, respectively, as found in contryphan-Vc1.

Both analogues contryphan-cg1₁₋₂₂[M9Y, P18I] and contryphan-cg1₁₋₂₂[R4Q, M9Y, R15I, and P18I] showed similar NMR spectral dispersion to that of native contryphan-Vc1 and similar deviations from random coil values as in con-Vc1₁₋₂₂[Z1Q], confirming that the SDH fold of the contryphan-Vc1 was restored in these analogues. Restoration of SDH fold after replacing selected residues like Met and Pro suggests that there are critical residues apart from Cys contribute to the stability of the SDH fold and also suggests that a consensus motif might be present for a peptide to exhibit the SDH fold. It also raises the possibility that the reported *Crassostrea gigas* sequence may be incorrect, but this needs to be explored further. The availability of new genomic and transcriptomic data for related species will help clarify this point. Determining the structures of these analogues will provide further information on the SDH fold and the roles of different residues. Further examination of the influence of other conserved residues on the β -hairpin core by selective residue replacement should also help define a consensus motif for the SDH fold.

Chapter 7

Conclusions and Future Directions

Conclusions

Grafting bioactive peptide sequences that target specific protein-protein interactions onto stable scaffolds has become a promising approach for the development of peptide-based therapeutics. Animal venoms are a rich source of potentially useful peptide scaffolds. Cone snails have arguably the most complex venoms, consisting of small and highly selective peptides for prey capture and defence. The venoms of cone snails are vast combinatorial-like libraries of evolutionarily selected bioactive peptides called conotoxins or conopeptides, which have huge therapeutic potential. The recent identification of contryphan-Vc1, a 31-residue peptide from the venom of the cone snail *Conus victoriae*, has been shown to have a unique fold, designated as the single disulfide-directed β -hairpin (SDH). The most conspicuous structural feature of contryphan-Vc1 is its two-stranded anti-parallel β -sheet stabilised by a single disulfide bond. The β -hairpin core of contryphan-Vc1 displays an ordered structure and remarkable thermal stability, and is very similar structurally to the ICK fold, which is stabilised by a minimum of three disulfide bonds. These properties suggest that contryphan-Vc1 may have potential as a useful scaffold for grafting.

In Chapter 3 of this thesis my focus was to characterize the SDH fold in order to evaluate its stability. Based on backbone ^{15}N relaxation data I showed that residues Trp2–Ile18 adopt a well-defined conformation, while residues beyond Thr19 are highly disordered, with the flexibility increasing towards the C-terminus. I have shown that the flexible C-terminus of contryphan-Vc1 can be truncated without loss of ordered structure. I have determined the thermal, chemical, redox and proteolytic stability of the contryphan-Vc1. Contryphan-Vc1 exhibited remarkable thermal stability, with the SDH fold of the peptide being maintained even at 95°C. Contryphan-Vc1 also exhibited remarkable chemical stability in the presence of urea, showing <30% unfolding even in 7 M urea. Contryphan-Vc1 is also quite stable over a broad pH range of pH 2 to 8. It is highly redox stable and required more than 2 days at room temperature for complete reduction in the presence of an excess of a strong reducing agent such as TCEP. The β -hairpin core structure was found to be resistant to cleavage by trypsin and chymotrypsin although it was susceptible to cleavage by pepsin.

In order to assess the capacity of the SDH fold to incorporate foreign peptide sequences and present them in a functional manner, the loops in the core structure of contryphan-Vc1 were replaced with a five-residue DINNN motif from iNOS, which interacts with SPSB proteins and mediates its proteolytic degradation. Even though both Vc1-DINNN analogues, sCon-Vc1₁₁₋₂₂[Z1Q, DINNN₄₋₈] and sCon-Vc1₁₁₋₂₂[Z1Q, DINNN₁₂₋₁₆], bound to SPSB2 with affinities of 25 ± 8 and 5.7 ± 3 nM respectively, ^1H NMR spectra of both peptides showed poor

peak dispersion, indicating that substituting the loop 1 and loop 2 residues with the five-residue DINNN motif disrupted the native SDH fold. I have shown that the analogue sCon-Vc1₁₋₂₂[NNN₁₂₋₁₄] with a shorter three-residue NNN insert in the place of the VLG residues in loop2 bound to human SPSB2 with an affinity of 1.3 μ M, which is almost fifty-fold stronger than that of the linear NNN epitope itself. Moreover, ¹H NMR spectra of this analogue showed good peak dispersion and native-like chemical shifts, indicating that the β -hairpin fold was maintained.

This study indicated that the SDH fold of contryphan-Vc1 has limited proteolytic stability and limitations on the length of sequence variation that can be accommodated in either of these loops. In order to further enhance the limited proteolytic stability and to overcome the limitation of incorporating only short peptide motifs we decided to constrain the SDH fold of the contryphan-Vc1 with an additional disulfide bond. Chapter 4 describes our effort in obtaining the two disulfide bonded analogue of contryphan-Vc1, contryphan-Vc1₁₋₂₂[Q1C,Y9C], and its characterisation using NMR. I demonstrated that an additional disulfide bond can be introduced by changing Gln1 and Tyr9 of contryphan-Vc1 to Cys utilising orthogonal Cys protection. The introduction of an additional disulfide bond into contryphan-Vc1 did not disrupt the SDH fold. Contryphan-Vc1₁₋₂₂[Q1C,Y9C] was shown to be resistant to trypsin digestion and to is cleaved at Leu20 by chymotrypsin, similar to that of Con-Vc1₁₋₂₂[Z1Q]. In contrast contryphan-Vc1₁₋₂₂[Q1C,Y9C] was completely resistant to pepsin digestion, whereas native contryphan-Vc1 was completely susceptible to pepsin.

The target of contryphan-Vc1 remains elusive as the peptide did not show any reproducible activity in mice upon intracranial injection. The peptide did not produce any observable changes to normal or depolarization-induced intracellular Ca²⁺ levels in mouse dorsal root ganglion cells. As the surface of the β -hairpin core of the contryphan-Vc1 is highly hydrophobic, with the presence of Trp, Tyr Pro, Leu, Ile, Val and Ala residues, we explored the membrane binding properties of contryphan-Vc1 using dodecylphosphocholine micelles as a model membrane. In Chapter 5 I showed that both full-length and truncated contryphan-Vc1 interact with DPC micelles, as seen in the significant changes in the ¹H NMR spectra of both peptides in presence of DPC. In contrast, Con-Vc1₁₋₂₂[Z1Q] bound very weakly to a flat bilayer formed by the zwitterionic POPC lipid and surprisingly full-length contryphan-Vc1 did not show any binding towards POPC lipid.

The concept of homology is central to analyses of protein and DNA sequences. Homologous sequences are expected to have similar structures, and frequently, similar functions as well. Structural and functional characterisation of the homologous sequences for

a given protein or a peptide will provide valuable insights on the structure and function of all the members of the similar family. In Chapter 6 we presented the alignment of available contryphan-Vc1 homologous sequences and noted the different residues based on the degree of the conservation. As expected, the Cys residues are highly conserved in all sequences, and showed that the disulfide bond is critical for folding and the maintenance of the SDH fold of the contryphan-Vc1 using the analogue contryphan-Vc1₁₋₂₂ [C3S, C16S]. I showed that full-length contryphan-cg1 peptide is unstructured and the SDH fold can be restored by introducing conserved amino acid residues.

Future directions

The two-disulfide bonded analogue of contryphan-Vc1, contryphan-Vc1₁₋₂₂[Q1C, Y9C] not only maintained the SDH fold of the peptide but also showed enhanced proteolytic stability. Further studies to check whether contryphan-Vc1₁₋₂₂[Q1C, Y9C] can accept motifs larger than three residues will establish whether the disulfide engineering was successful in overcoming the limitation of accepting short epitopes (~ 3 residues) and will provide insights into the directed evolution of disulfide-rich folds from the elementary SDH fold. Contryphan-Vc1₁₋₂₂[Q1C,Y9C] exhibits conformational averaging, and the unusual temperature dependence, where the peptides becomes more ordered upon heating is quite intriguing. Further studies are warranted to understand the reasons for this temperature-dependent conformational averaging.

Contryphan-Vc1 was shown to interact strongly with DPC micelles through its structured region. In contrast, only weak binding was observed for truncated contryphan-Vc1 in the presence of a flat bilayer composed of zwitterionic POPC lipid, and surprisingly full-length contryphan-Vc1 did not show any binding towards POPC lipid. Determining the structure of contryphan-Vc1 in DPC micelles will help in understanding whether the peptide is able to maintain the SDH fold and the influence of DPC micelles on the structure of the peptide. It will be interesting to assess whether contryphan-Vc1 binds to anionic vesicles or a mixture of zwitterionic and anionic lipids. Further characterisation of lipid interactions of contryphan-Vc1 utilising different spin-labelled lipids and obtaining diffusion coefficient rates in different lipid environments and observing whether labelled-contryphan-Vc1 interacts with cell membranes using microscopy will shed more light on the membrane binding and membrane partitioning ability of contryphan-Vc1.

We have identified some crucial residues which can stabilise or destabilise the SDH fold as seen in the analogues of contryphan-cg1. Further structural characterization of these analogues will help define the fine details and nuances of the SDH fold and the influence of

different residues on that fold. Further examination of the influence of other conserved residues on the β -hairpin core by selective residue replacement should define the minimum consensus motif necessary for the SDH fold. Contryphan-Vc1 is the first member identified in this new family of peptides with SDH fold whose function is still elusive. Further structural and functional characterisation of the newly identified members of this family allows us to understand the functional significance of contryphan-Vc1 and its role in the venom of *Conus victoriae*.

Chapter 8

References

1. Cunningham, A. D., Qvit, N., and Mochly-Rosen, D. (2017) Peptides and peptidomimetics as regulators of protein-protein interactions. *Curr Opin Struct Biol* 44, 59-66.
2. Arkin, M. R., and Wells, J. A. (2004) Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* 3, 301-317.
3. Fischer, G., Rossmann, M., and Hyvonen, M. (2015) Alternative modulation of protein-protein interactions by small molecules. *Curr Opin Biotechnol* 35, 78-85.
4. London, N., Raveh, B., and Schueler-Furman, O. (2013) Druggable protein-protein interactions--from hot spots to hot segments. *Curr Opin Chem Biol* 17, 952-959.
5. Arkin, M. R., Tang, Y., and Wells, J. A. (2014) Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem Biol* 21, 1102-1114.
6. Craik, D. J., Fairlie, D. P., Liras, S., and Price, D. (2013) The future of peptide-based drugs. *Chem Biol Drug Des* 81, 136-147.
7. Fosgerau, K., and Hoffmann, T. (2015) Peptide therapeutics: current status and future directions. *Drug Discov Today* 20, 122-128.
8. Vita, C., Drakopoulou, E., Vizzavona, J., Garnier, P., and Menez, A. (1999) Engineering novel mini-proteins by the transfer of active sites to small disulfide stabilized natural scaffolds. *J Pep Sci*, 121-122.
9. Martin, L., and Vita, C. (2000) Engineering novel bioactive mini-proteins from small size natural and *de novo* designed scaffolds. *Curr Protein Peptide Sci* 1, 403-430.
10. Hosse, R. J., Rothe, A., and Power, B. E. (2006) A new generation of protein display scaffolds for molecular recognition. *Protein Sci* 15, 14-27.
11. Masefski, W., Jr., Redfield, A. G., Hare, D. R., and Miller, C. (1990) Molecular structure of charybdotoxin, a pore-directed inhibitor of potassium ion channels. *Science* 249, 521-524.
12. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., Roumestand, C., Yang, Y. S., Ylisastigui, L., Benjouad, A., and Gluckman, J. C. (1999) Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. *Proc Natl Acad Sci U S A* 96, 13091-13096.
13. Lavergne, V., Alewood, P. F., Mobli, M., and King, G. F. (2015) CHAPTER 2 The structural universe of disulfide-rich venom peptides, In *venoms to drugs: venom as a source for the development of human therapeutics*, pp 37-79, The Royal Society of Chemistry.
14. Norton, R. S., and Pallaghy, P. K. (1998) The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* 36, 1573-1583.
15. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) A common structural motif incorporating a cystine knot and a triple-stranded beta-sheet in toxic and inhibitory polypeptides. *Protein Sci* 3, 1833-1839.
16. Atanassoff, P. G., Hartmannsgruber, M. W., Thrasher, J., Wermeling, D., Longton, W., Gaeta, R., Singh, T., Mayo, M., McGuire, D., and Luther, R. R. (2000) Ziconotide, a new N-type calcium channel blocker, administered intrathecally for acute postoperative pain. *Reg Anesth Pain Med* 25, 274-278.
17. Isaacs, N. W. (1995) Cystine Knots. *Curr Opin Struct Biol* 5, 391-395.
18. Kolmar, H. (2009) Biological diversity and therapeutic potential of natural and engineered cystine knot miniproteins. *Curr Opin Pharmacol* 9, 608-614.
19. Kimura, R. H., Levin, A. M., Cochran, F. V., and Cochran, J. R. (2009) Engineered cystine knot peptides that bind α v β 3, α v β 5, and α 5 β 1 integrins with low-nanomolar affinity. *Proteins* 77, 359-369.

20. Kimura, R. H., Cheng, Z., Gambhir, S. S., and Cochran, J. R. (2009) Engineered knottin peptides: A new class of agents for imaging integrin expression in living subjects. *Cancer Res* 69, 2435-2442.
21. Ollmann, M. M., Wilson, B. D., Yang, Y.-K., Kerns, J. A., Chen, Y., Gantz, I., and Barsh, G. S. (1997) Antagonism of central melanocortin receptors in vitro and in vivo by agouti-related protein. *Science* 278, 135-138.
22. Jackson, P. J., McNulty, J. C., Yang, Y.-K., Thompson, D. A., Chai, B., Gantz, I., Barsh, G. S., and Millhauser, G. L. (2002) Design, pharmacology, and NMR structure of a minimized cystine knot with agouti-related protein activity. *Biochemistry* 41, 7565-7572.
23. Miao, Z., Ren, G., Liu, H., Kimura, R. H., Jiang, L., Cochran, J. R., Gambhir, S. S., and Cheng, Z. (2009) An engineered knottin peptide labeled with ¹⁸F for PET imaging of integrin expression. *Bioconjug Chem* 20, 2342-2347.
24. Moore, S. J., Leung, C. L., Norton, H. K., and Cochran, J. R. (2013) Engineering agatoxin, a cystine-knot peptide from spider venom, as a molecular probe for *in vivo* tumor imaging. *PLoS One* 8, e60498.
25. Gunasekera, S., Daly, N. L., Clark, R. J., and Craik, D. J. (2009) Dissecting the oxidative folding of circular cystine knot miniproteins. *Antioxid Redox Signal* 11, 971-980.
26. Thongyoo, P., Roque-Rosell, N., Leatherbarrow, R. J., and Tate, E. W. (2008) Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org Biomol Chem* 6, 1462-1470.
27. Chan, L. Y., Gunasekera, S., Henriques, S. T., Worth, N. F., Le, S.-J., Clark, R. J., Campbell, J. H., Craik, D. J., and Daly, N. L. (2011) Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds. *Blood* 118, 6709-6717.
28. King, G. F. (2011) Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin Biol Ther* 11, 1469-1484.
29. Fry, B. G., Roelants, K., Champagne, D. E., Scheib, H., Tyndall, J. D., King, G. F., Nevalainen, T. J., Norman, J. A., Lewis, R. J., Norton, R. S., Renjifo, C., and de la Vega, R. C. (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet* 10, 483-511.
30. Olivera, B. M. (2002) Conus venom peptides: Reflections from the biology of clades and species. *Annu Rev Ecol Syst* 33, 25-47.
31. Norton, R. S., and Olivera, B. M. (2006) Conotoxins down under. *Toxicon* 48, 780-798.
32. Akondi, K. B., Muttenthaler, M., Dutertre, S., Kaas, Q., Craik, D. J., Lewis, R. J., and Alewood, P. F. (2014) Discovery, synthesis, and structure-activity relationships of conotoxins. *Chem Rev* 114, 5815-5847.
33. Han, T. S., Teichert, R. W., Olivera, B. M., and Bulaj, G. (2008) Conus venoms - A rich source of peptide-based therapeutics. *Curr Pharm Des* 14, 2462-2479.
34. Kaas, Q., Westermann, J. C., and Craik, D. J. (2010) Conopeptide characterization and classifications: an analysis using ConoServer. *Toxicon* 55, 1491-1509.
35. Robinson, S. D., and Norton, R. S. (2014) Conotoxin gene superfamilies. *Mar Drugs* 12, 6058-6101.
36. Terlau, H., and Olivera, B. M. (2004) Conus venoms: a rich source of novel ion channel-targeted peptides. *Physiol Rev* 84, 41-68.
37. Safavi-Hemami, H., Hu, H., Gorasia, D. G., Bandyopadhyay, P. K., Veith, P. D., Young, N. D., Reynolds, E. C., Yandell, M., Olivera, B. M., and Purcell, A. W. (2014) Combined proteomic and transcriptomic interrogation of the venom gland of *Conus geographus* uncovers novel components and functional compartmentalization. *Mol Cell Proteomics* 13, 938-953.

38. Laverigne, V., Dutertre, S., Jin, A.-h., Lewis, R. J., Taft, R. J., and Alewood, P. F. (2013) Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. *BMC Genomics* 14, 708.
39. Gao, B., Peng, C., Yang, J., Yi, Y., Zhang, J., and Shi, Q. (2017) Cone snails: A big store of conotoxins for novel drug discovery. *Toxins (Basel)* 9, 397.
40. Robinson, S. D., Safavi-Hemami, H., McIntosh, L. D., Purcell, A. W., Norton, R. S., and Papenfuss, A. T. (2014) Diversity of conotoxin gene superfamilies in the venomous snail, *Conus victoriae*. *PLoS One* 9, e87648.
41. Jimenez, E. C., Olivera, B. M., Gray, W. R., and Cruz, L. J. (1996) Contryphan is a D-tryptophan-containing *Conus* peptide. *J Biol Chem* 271, 28002-28005.
42. Jimenez, E. C., Watkins, M., Juszczak, L. J., Cruz, L. J., and Olivera, B. M. (2001) Contryphans from *Conus textile* venom ducts. *Toxicon* 39, 803-808.
43. Jacobsen, R. B., Jimenez, E. C., De la Cruz, R. G., Gray, W. R., Cruz, L. J., and Olivera, B. M. (1999) A novel D-leucine-containing *Conus* peptide: diverse conformational dynamics in the contryphan family. *J Pept Res* 54, 93-99.
44. Massilia, G. R., Schinina, M. E., Ascenzi, P., and Polticelli, F. (2001) Contryphan-Vn: a novel peptide from the venom of the mediterranean snail *Conus ventricosus*. *Biochem Biophys Res Commun* 288, 908-913.
45. Rajesh, R. P. (2015) Novel M-Superfamily and T-Superfamily conotoxins and contryphans from the vermivorous snail *Conus figulinus*. *J Pept Sci* 21, 29-39.
46. Hansson, K., Ma, X., Eliasson, L., Czerwicz, E., Furie, B., Furie, B. C., Rorsman, P., and Stenflo, J. (2004) The first gamma-carboxyglutamic acid-containing contryphan. A selective L-type calcium ion channel blocker isolated from the venom of *Conus marmoreus*. *J Biol Chem* 279, 32453-32463.
47. Robinson, S. D., Chhabra, S., Belgi, A., Chittoor, B., Safavi-Hemami, H., Robinson, A. J., Papenfuss, A. T., Purcell, A. W., and Norton, R. S. (2016) A naturally occurring peptide with an elementary single disulfide-directed beta-hairpin fold. *Structure* 24, 293-299.
48. Filippakopoulos, P., Low, A., Sharpe, T. D., Uppenberg, J., Yao, S., Kuang, Z., Savitsky, P., Lewis, R. S., Nicholson, S. E., Norton, R. S., and Bullock, A. N. (2010) Structural basis for Par-4 recognition by the SPRY domain- and SOCS box-containing proteins SPSB1, SPSB2, and SPSB4. *J Mol Biol* 401, 389-402.
49. Kuang, Z., Lewis, R. S., Curtis, J. M., Zhan, Y., Saunders, B. M., Babon, J. J., Kolesnik, T. B., Low, A., Masters, S. L., Willson, T. A., Kedzierski, L., Yao, S., Handman, E., Norton, R. S., and Nicholson, S. E. (2010) The SPRY domain-containing SOCS box protein SPSB2 targets iNOS for proteasomal degradation. *J Cell Biol* 190, 129-141.
50. Lewis, R. S., Kolesnik, T. B., Kuang, Z., D'Cruz, A. A., Blewitt, M. E., Masters, S. L., Low, A., Willson, T., Norton, R. S., and Nicholson, S. E. (2011) TLR regulation of SPSB1 controls inducible nitric oxide synthase induction. *J Immunol* 187, 3798-3805.
51. Lee, S. Y. (1996) High cell-density culture of *Escherichia coli*. *Trends Biotechnol* 14, 98-105.
52. Ihssen, J., Kowarik, M., Dilettoso, S., Tanner, C., Wacker, M., and Thony-Meyer, L. (2010) Production of glycoprotein vaccines in *Escherichia coli*. *Microb Cell Fact* 9, 61.
53. Hannig, G., and Makrides, S. C. (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol* 16, 54-60.
54. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley, New York.
55. Wüthrich, K. (1995) *NMR in structural biology : a collection of papers by Kurt Wüthrich*, World Scientific, Singapore ; River Edge, NJ.

56. Claridge, T. D. W. (1999) Chapter 5 Correlations through the chemical bond I: Homonuclear shift correlation, In *Tetrahedron Organic Chemistry Series*, pp 147-220, Elsevier.
57. Shaka, A., Lee, C., and Pines, A. (1988) Iterative schemes for bilinear operators; application to spin decoupling. *J Magn Reson* 77, 274-293.
58. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687-696.
59. Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Protein NMR Techniques*, 353-378.
60. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160, 65-73.
61. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14, 51-55, 29-32.
62. Homola, J. (2008) Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species. *Chemical Reviews* 108, 462-493.
63. Karplus, M., and McCammon, J. A. (2002) Molecular dynamics simulations of biomolecules. *Nat Struct Mol Biol* 9, 646.
64. Hansson, T., Oostenbrink, C., and van Gunsteren, W. (2002) Molecular dynamics simulations. *Curr Opin Struct Biol* 12, 190-196.
65. Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91, 43-56.
66. Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E., and van Gunsteren, W. F. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* 40, 843-856.
67. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J Mol Graphics* 14, 33-38.
68. Chittoor, B., Krishnarajuna, B., Morales, R. A. V., MacRaild, C. A., Sadek, M., Leung, E. W. W., Robinson, S. D., Pennington, M. W., and Norton, R. S. (2017) The single disulfide-directed beta-hairpin fold. Dynamics, stability, and engineering. *Biochemistry* 56, 2455-2466.
69. Keil, B. (2012) *Specificity of proteolysis*, Springer Science & Business Media.
70. Jo, B. H., Park, T. Y., Park, H. J., Yeon, Y. J., Yoo, Y. J., and Cha, H. J. (2016) Engineering de novo disulfide bond in bacterial α -type carbonic anhydrase for thermostable carbon sequestration. *Sci Rep* 6, 29322.
71. Yu, H., and Huang, H. (2014) Engineering proteins for thermostability through rigidifying flexible sites. *Biotechnol Adv* 32, 308-315.
72. Liu, L., Deng, Z., Yang, H., Li, J., Shin, H.-d., Chen, R. R., Du, G., and Chen, J. (2014) In silico rational design and systems engineering of disulfide bridges in the catalytic domain of an alkaline α -amylase from *Alkalimonas amylolytica* to improve thermostability. *Appl Environ Microbiol* 80, 798-807.
73. Matsumura, M., Becktel, W. J., Levitt, M., and Matthews, B. W. (1989) Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc Natl Acad Sci U S A* 86, 6562-6566.
74. Le, Q. A. T., Joo, J. C., Yoo, Y. J., and Kim, Y. H. (2012) Development of thermostable *Candida antarctica* lipase B through novel in silico design of disulfide bridge. *Biotechnol Bioeng* 109, 867-876.
75. Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G., and Eijssink, V. G. (1998) Engineering an enzyme to resist boiling. *Proc Natl Acad Sci U S A* 95, 2056-2060.

76. Craig, D. B., and Dombkowski, A. A. (2013) Disulfide by design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* 14, 346.
77. Flory, P. J. (1956) Theory of elastic mechanisms in fibrous proteins. *J Am Chem Soc* 78, 5222-5235.
78. Mamathambika, B. S., and Bardwell, J. C. (2008) Disulfide-linked protein folding pathways. *Annu Rev Cell Dev Biol* 24, 211-235.
79. Dombkowski, A. A., Sultana, K. Z., and Craig, D. B. (2014) Protein disulfide engineering. *FEBS Lett* 588, 206-212.
80. Mansfeld, J., Vriend, G., Dijkstra, B. W., Veltman, O. R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R., and Eijssink, V. G. H. (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J Biol Chem* 272, 11152-11156.
81. Roesler, K. R., and Rao, A. G. (2000) A single disulfide bond restores thermodynamic and proteolytic stability to an extensively mutated protein. *Protein Sci* 9, 1642-1650.
82. Takagi, H., Takahashi, T., Momose, H., Inouye, M., Maeda, Y., Matsuzawa, H., and Ohta, T. (1990) Enhancement of the thermostability of subtilisin E by introduction of a disulfide bond engineered on the basis of structural comparison with a thermophilic serine protease. *J Biol Chem* 265, 6874-6878.
83. Rousseau, R., Schreiner, E., Kohlmeyer, A., and Marx, D. (2004) Temperature-dependent conformational transitions and hydrogen-bond dynamics of the elastin-like octapeptide GVG(VPGVG): a molecular-dynamics study. *Biophys J* 86, 1393-1407.
84. Hung, A., Kuyucak, S., Schroeder, C. I., and Kaas, Q. (2017) Modelling the interactions between animal venom peptides and membrane proteins. *Neuropharmacology* 127, 20-31.
85. Tosteson, M. T., and Tosteson, D. C. (1981) The sting. Melittin forms channels in lipid bilayers. *Biophys J* 36, 109-116.
86. Gasanov, S. E., Dagda, R. K., and Rael, E. D. (2014) Snake venom cytotoxins, phospholipase A₂s, and Zn⁽²⁺⁾-dependent metalloproteinases: Mechanisms of action and pharmacological relevance. *J Clin Toxicol* 4, 1000181.
87. Suchyna, T. M., Tape, S. E., Koeppe II, R. E., Andersen, O. S., Sachs, F., and Gottlieb, P. A. (2004) Bilayer-dependent inhibition of mechanosensitive channels by neuroactive peptide enantiomers. *Nature* 430, 235.
88. Lee, S. Y., and MacKinnon, R. (2004) A membrane-access mechanism of ion channel inhibition by voltage sensor toxins from spider venom. *Nature* 430, 232-235.
89. Ladokhin, A. S., and White, S. H. (2001) 'Detergent-like' permeabilization of anionic lipid vesicles by melittin. *Biochim Biophys Acta* 1514, 253-260.
90. Swartz, K. J. (2007) Tarantula toxins interacting with voltage sensors in potassium channels. *Toxicon* 49, 213-230.
91. Klint, J. K., Smith, J. J., Vetter, I., Rupasinghe, D. B., Er, S. Y., Senff, S., Herzig, V., Mobli, M., Lewis, R. J., Bosmans, F., and King, G. F. (2015) Seven novel modulators of the analgesic target NaV 1.7 uncovered using a high-throughput venom-based discovery approach. *Br J Pharmacol* 172, 2445-2458.
92. Jung, H. J., Lee, J. Y., Kim, S. H., Eu, Y.-J., Shin, S. Y., Milescu, M., Swartz, K. J., and Kim, J. I. (2005) Solution structure and lipid membrane partitioning of VSTx1, an inhibitor of the KvAP potassium channel. *Biochemistry* 44, 6015-6023.
93. Phillips, L. R., Milescu, M., Li-Smerin, Y., Mindell, J. A., Kim, J. I., and Swartz, K. J. (2005) Voltage-sensor activation with a tarantula toxin as cargo. *Nature* 436, 857-860.
94. Milescu, M., Vobecky, J., Roh, S. H., Kim, S. H., Jung, H. J., Kim, J. I., and Swartz, K. J. (2007) Tarantula toxins interact with voltage sensors within lipid membranes. *J Gen Physiol* 130, 497-511.

95. Smith, J. J., Alphy, S., Seibert, A. L., and Blumenthal, K. M. (2005) Differential phospholipid binding by site 3 and site 4 toxins. Implications for structural variability between voltage-sensitive sodium channel domains. *J Biol Chem* 280, 11127-11133.
96. 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., and Schroeder, C. I. (2016) 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* 291, 17049-17065.
97. Lau, C. H. Y., King, G. F., and Mobli, M. (2016) Molecular basis of the interaction between gating modifier spider toxins and the voltage sensor of voltage-gated ion channels. *Sci Rep* 6, 34333.
98. Garcia, M. L. (2004) Ion channels: gate expectations. *Nature* 430, 153-155.
99. Posokhov, Y. O., Gottlieb, P. A., Morales, M. J., Sachs, F., and Ladokhin, A. S. (2007) Is lipid bilayer binding a common property of inhibitor cysteine knot ion-channel blockers? *Biophys J* 93, L20-L22.
100. Robinson, S. D., and Norton, R. S. (2014) Conotoxin gene superfamilies. *Mar Drugs* 12, 6058-6101.
101. Li, D., Xiao, Y., Hu, W., Xie, J., Bosmans, F., Tytgat, J., and Liang, S. (2003) Function and solution structure of hainantoxin-I, a novel insect sodium channel inhibitor from the Chinese bird spider *Selenocosmia hainana*. *FEBS Lett* 555, 616-622.
102. Kallick, D. A., Tessmer, M. R., Watts, C. R., and Li, C. Y. (1995) The use of dodecylphosphocholine micelles in solution NMR. *J Magn Reson B* 109, 60-65.
103. Shin, K., Sarker, M., Huang, S. K., and Rainey, J. K. (2017) Apelin conformational and binding equilibria upon micelle interaction primarily depend on membrane-mimetic headgroup. *Sci Rep* 7, 15433.
104. Roumestand, C., Louis, V., Aumelas, A., Grassy, G., Calas, B., and Chavanieu, A. (1998) Oligomerization of protegrin-1 in the presence of DPC micelles. A proton high-resolution NMR study. *FEBS Lett* 421, 263-267.
105. Theisgen, S., Thomas, L., Schroder, T., Lange, C., Kovermann, M., Balbach, J., and Huster, D. (2011) The presence of membranes or micelles induces structural changes of the myristoylated guanylate-cyclase activating protein-2. *Eur Biophys J* 40, 565-576.
106. Akashi, S., and Takio, K. (2001) Structure of melittin bound to phospholipid micelles studied using hydrogen-deuterium exchange and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 12, 1247-1253.
107. le Maire, M., Champeil, P., and Moller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* 1508, 86-111.
108. Hauge, H. H., Mantzilas, D., Eijssink, V. G. H., and Nissen-Meyer, J. (1999) Membrane-mimicking entities induce structuring of the two-peptide bacteriocins plantaricin E/F and plantaricin J/K. *J Bacteriol* 181, 740-747.
109. Héctor, Z. C., Beatriz, M., Erik, S., S., U. A., M., S. J., and Ángeles, J. M. (2015) Micelle-triggered β -hairpin to α -helix transition in a 14-residue peptide from a choline-binding repeat of the pneumococcal autolysin LytA. *Chemistry – A European Journal* 21, 8076-8089.
110. Sauv  , S., and Aubin, Y. (2016) Dodecylphosphocholine micelles induce amyloid formation of the PrP(110-136) peptide via an α -helical metastable conformation. *PLoS One* 11, e0168021.
111. Bystrov, V. F., Dubrovina, N. I., Barsukov, L. I., and Bergelson, L. D. (1971) NMR differentiation of the internal and external phospholipid membrane surfaces using paramagnetic Mn^{2+} and Eu^{3+} ions. *Chem Phys Lipids* 6, 343-350.

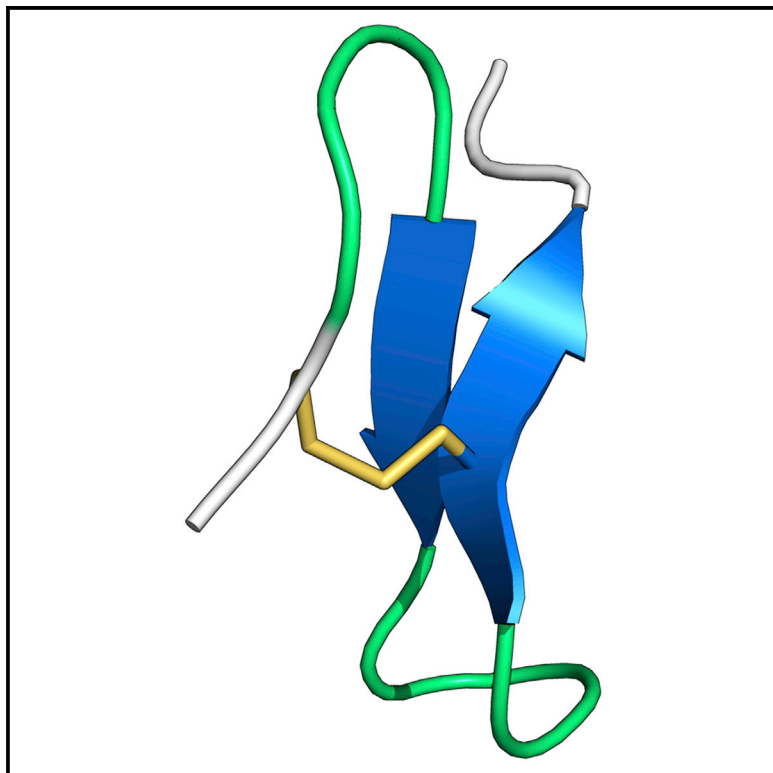
112. MacRaild, C. A., Pedersen, M. Ø., Anders, R. F., and Norton, R. S. (2012) Lipid interactions of the malaria antigen merozoite surface protein 2. *Biochim Biophys Acta* 1818, 2572-2578.
113. Smith, J. J., Hill, J. M., Little, M. J., Nicholson, G. M., King, G. F., and Alewood, P. F. (2011) Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif. *Proc Natl Acad Sci U S A* 108, 10478-10483.
114. Cui, X., Vinař, T., Brejová, B., Shasha, D., and Li, M. (2007) Homology search for genes. *Bioinformatics* 23, i97-i103.
115. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* 5, 67-81.
116. Biegert, A., and Soding, J. (2009) Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci U S A* 106, 3770-3775.
117. Pace, C. N., Grimsley, G. R., Thomson, J. A., and Barnett, B. J. (1988) Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J Biol Chem* 263, 11820-11825.
118. Pace, C. N. (1990) Conformational stability of globular proteins. *Trends Biochem Sci* 15, 14-17.
119. Betz, S. F. (1993) Disulfide bonds and the stability of globular proteins. *Protein Sci* 2, 1551-1558.

Appendix I

Structure

A Naturally Occurring Peptide with an Elementary Single Disulfide-Directed β -Hairpin Fold

Graphical Abstract



Authors

Samuel D. Robinson,
Sandeep Chhabra, Alessia Belgi, ...,
Anthony T. Papenfuss,
Anthony W. Purcell,
Raymond S. Norton

Correspondence

ray.norton@monash.edu

In Brief

Robinson et al. report the discovery of a new peptide fold they have called the single disulfide-directed β hairpin (SDH). The SDH is a naturally occurring, small, highly stable, independent folding unit and is the common elementary motif underlying several peptide folds, including the inhibitor cystine knot.

Highlights

- The single disulfide-directed β hairpin is an elementary peptide fold
- It forms the core of the inhibitor cystine knot motif
- It is highly resistant to thermal denaturation
- It appears to be common to a previously undescribed class of molluscan neuropeptide

Accession Numbers

W4VSF6
2N24



A Naturally Occurring Peptide with an Elementary Single Disulfide-Directed β -Hairpin Fold

Samuel D. Robinson,¹ Sandeep Chhabra,¹ Alessia Belgi,² Balasubramanyam Chittoor,¹ Helena Safavi-Hemami,³ Andrea J. Robinson,² Anthony T. Papenfuss,⁴ Anthony W. Purcell,⁵ and Raymond S. Norton^{1,*}

¹Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia

²School of Chemistry, Monash University, Clayton, VIC 3800, Australia

³Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

⁴Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia

⁵The Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, VIC 3800, Australia

*Correspondence: ray.norton@monash.edu

<http://dx.doi.org/10.1016/j.str.2015.11.015>

SUMMARY

Certain peptide folds, owing to a combination of intrinsic stability and resilience to amino acid substitutions, are particularly effective for the display of diverse functional groups. Such “privileged scaffolds” are valuable as starting points for the engineering of new bioactive molecules. We have identified a precursor peptide expressed in the venom gland of the marine snail *Conus victoriae*, which appears to belong to a hitherto undescribed class of molluscan neuropeptides. Mass spectrometry matching with the venom confirmed the complete mature peptide sequence as a 31-residue peptide with a single disulfide bond. Solution structure determination revealed a unique peptide fold that we have designated the single disulfide-directed β hairpin (SDH). The SDH fold is highly resistant to thermal denaturation and forms the core of several other multiple disulfide-containing peptide folds, including the inhibitor cystine knot. This elementary fold may offer a valuable starting point for the design and engineering of new bioactive peptides.

INTRODUCTION

A limited number of topologically distinct protein and peptide folds is found in nature (Orengo et al., 1993), some of which are utilized for diverse functions, for example the inhibitor cystine knot (ICK; also known as knottin) fold adopted by many disulfide-rich peptides. The ICK fold can be described as a β -hairpin motif stabilized by a cystine knot, in which a ring formed by two disulfides and their connecting backbone is threaded by a third disulfide (Pallaghy et al., 1994). The peptide chain (usually 25–50 amino acids) contains six cysteine residues (denoted Cys I–VI), and can be divided into five intercysteine loops (although Cys III and IV are often contiguous in the amino acid sequence). A resilience to chemical changes in these intercysteine loops, combined with exceptional intrinsic stability, has allowed the ICK fold to be used in a range of naturally occurring peptides with diverse biological functions (Norton and Pallaghy, 1998). ICK peptides

have been found in a variety of organisms, including plants, fungi, and in the venoms of animals, where they serve as protease inhibitors, antimicrobials, or ion channel modulators, respectively (Gelly et al., 2004).

The concept of a “privileged” chemical scaffold that is capable of interacting productively with multiple targets was proposed originally for certain classes of small organic molecules (Evans et al., 1988). It is clearly equally applicable to the ICK and several other peptide folds. As with their small organic counterparts, privileged peptide scaffolds represent valuable starting points for the design and engineering of new bioactive peptides (Moore et al., 2012).

Animal venoms are complex mixtures, often containing more than 100 unique bioactive peptides, and, owing to the requirements for high stability and diverse functionality, are a rich source of privileged peptide scaffolds (Reeks et al., 2015). The venoms of predatory marine cone snails are composed of hundreds of unique, small, mostly disulfide-rich peptides, known as conotoxins (Norton and Olivera, 2006; Olivera et al., 1990). Here, we report the discovery and structural characterization of the cone snail venom peptide contryphan-Vc1, which adopts a novel peptide motif that we have designated the single disulfide-directed β hairpin (SDH). This elementary structural motif forms the core of several other peptide folds and represents a new stable and synthetically accessible scaffold.

RESULTS

Discovery

Contryphan-Vc1 was identified as a prepropeptide in the venom gland transcriptome of *Conus victoriae* (Robinson et al., 2014). Conotoxins can be classified into gene superfamilies based on their signal peptide sequence (Jiménez et al., 1996), and by this criterion contryphan-Vc1 was defined as a member of the contryphan superfamily. However, beyond its signal peptide sequence and its single pair of cysteines, contryphan-Vc1 shared no obvious sequence similarity with other peptides of the contryphan superfamily (Figure S1), or indeed any other previously described conotoxin.

By searching a high-resolution tandem mass spectrometry (MS/MS) profile of reduced, alkylated, and trypsinized *C. victoriae* venom against the translated venom gland transcriptome, the mature peptide of contryphan-Vc1 and its associated

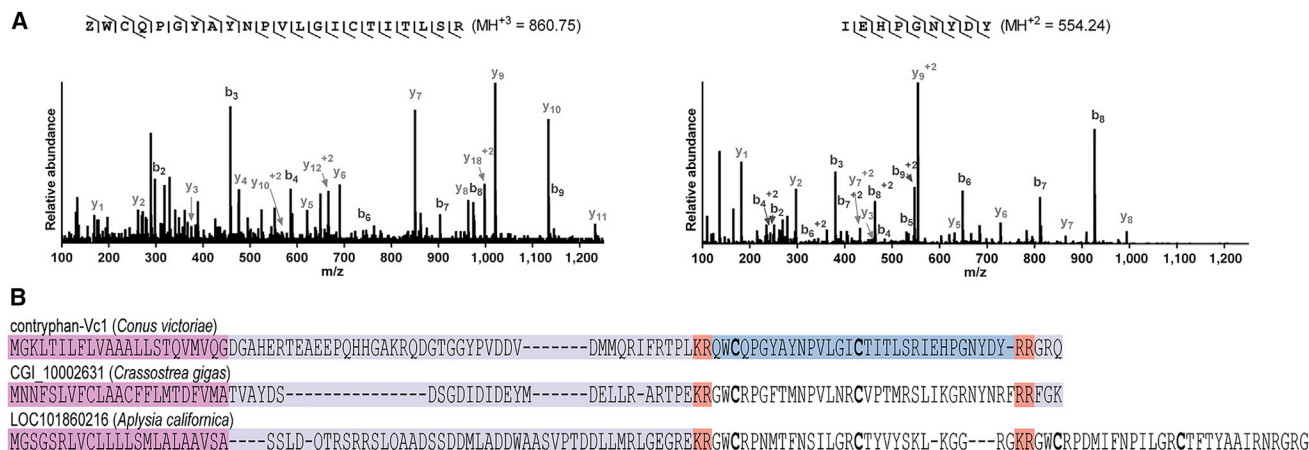


Figure 1. Discovery of Contryphan-Vc1

(A) MS/MS spectrum of contryphan-Vc1 tryptic peptides. Precursor ions selected for MS/MS had monoisotopic m/z of 860.75 ($z = 4$, predicted m/z 860.75) and 554.24 ($z = 2$, predicted m/z 554.24). For clarity, only b and y ions are labeled. Inset: b/y ladder diagrams summarize observed b and y ions.

(B) The precursor sequences of contryphan-Vc1, CGI_10002631, and LOC101860216. Precursor signal peptides are highlighted in purple, dibasic cleavage sites in red, and the confirmed encoded mature peptide of contryphan-Vc1 in blue.

post-translational modifications were confirmed. Two tryptic peptide fragments encoded by the contryphan-Vc1 precursor were detected by MS matching (Figure 1A), both of which were derived from the C-terminal region of the precursor and followed one another in the precursor sequence. The first 22-residue fragment was identified as an $[\text{M} + 3\text{H}]^{3+}$ ion of m/z 860.75 (theoretical $m/z = 860.75$), where the N-terminal Gln was post-translationally modified to pyroglutamic acid. This post-translational modification cannot occur mid-chain, indicating that this residue must represent the N terminus of the complete mature peptide. This residue also followed a dibasic cleavage site in the precursor, supporting this conclusion. This peptide fragment terminated in an Arg residue, which, as described below, was the result of tryptic cleavage of the complete mature peptide.

The second peptide fragment, nine residues in length, directly followed the first fragment in the precursor sequence. It was identified as an $[\text{M} + 2\text{H}]^{2+}$ ion of m/z 554.24 (theoretical $m/z = 554.24$). Importantly, the C terminus of this peptide was not the result of tryptic cleavage (it did not occur C-terminal to a Lys or Arg), indicating that this was the C terminus of the complete mature peptide. A short, positively-charged C-terminal propeptide is therefore removed during maturation, a process not unusual for conotoxins (Robinson and Norton, 2014). To further confirm that these two tryptic peptide fragments make up the complete contryphan-Vc1 mature peptide, the unprocessed venom was analyzed for the corresponding mass and collision-induced dissociation (CID) fragmentation ions consistent with those expected of the complete peptide. In the total ion chromatogram of the unprocessed crude venom, a peak of m/z 1,184.89 was identified, corresponding to the $[\text{M} + 3\text{H}]^{3+}$ of the full-length oxidized contryphan-Vc1 (theoretical $m/z = 1184.89$). Furthermore, when fragmented under CID, this peak yielded several strong ions diagnostic of the contryphan-Vc1 peptide sequence, including 728.29, 547.21, and 666.29 m/z , corresponding to the y6 (PGNYDY) and the internal fragment ions y6|b30 (PGNYD) and y27|b10 (PGYAYN),

respectively, reflecting expected preferential cleavage N-terminal to Pro5, 11, and 26. These data confirm the primary structure of native contryphan-Vc1 as a single-chain, single-disulfide-containing 31-residue peptide with an N-terminal pyroglutamic acid modification and the amino acid sequence ZWCQPGYAYNPVLGICTITLSRIEHPGNYDY.

A search of UniProt/GenBank using BLASTp for sequences similar to the mature contryphan-Vc1 peptide yielded two hits; predicted proteins CGI_10002631 and LOC101860216. Despite low alignment probability scores ($E = 0.41$ and 1.8), several factors led us to further explore these hits: relatively high sequence identities were observed (46% and 41%), sequence similarity was observed for residues predicted to be structurally important for contryphan-Vc1 (see Discussion), and both sequences were derived from other molluscs. Both sequences are uncharacterized and had been predicted using automated computational analysis, namely CGI_10002631 and LOC101860216 from the genomes of the non-venomous molluscs *Crassostrea gigas* (Pacific oyster) (Zhang et al., 2012) and *Aplysia californica* (California sea hare), respectively. Analysis of the full-length sequences and comparison with that of contryphan-Vc1 revealed some striking similarities (Figure 1B). Both sequences are clearly prepropeptides, with a 20-residue secretory signal peptide, followed by a propeptide region, a dibasic cleavage site, and one or more mature peptides encoded near the C terminus. In CGI_10002631 from *C. gigas* the predicted mature peptide is 32 residues in length and, as with contryphan-Vc1, is followed by a dibasic cleavage site and a three-residue propeptide region. The sequence LOC101860216 from *A. californica* encodes two distinct mature peptides (29 and 30 residues in length) separated by an additional dibasic cleavage site. All three predicted mature peptides have an inter-cysteine loop size of 11 residues and share several residues, some of which appear to be structurally important in contryphan-Vc1 (see Discussion). A search of the GenBank/EMBL/DBJ expressed sequence tag database using tblastn revealed that the sequence LOC101860216, predicted

Table 1. NMR and Refinement Statistics for Contryphan-Vc1

NMR Distance and Dihedral Constraints	
Distance constraints	
Total NOE	489
Intra-residue	132
Inter-residue	357
Sequential ($ i - j = 1$)	150
Medium-range ($ i - j < 4$)	112
Long-range ($ i - j > 5$)	95
Hydrogen bonds	5
Total dihedral angle restraints	34
ϕ	24
ψ	10
Structure Statistics	
Energies ^a	
E_{NOE} (kcal mol ⁻¹)	1.81 ± 0.33
E_{vdW} (kcal mol ⁻¹)	0.703 ± 0.27
Violations (mean and SD)	
Distance constraints (Å)	0.011 ± 0.001
Dihedral angle constraints (°)	0.312 ± 0.195
Max. dihedral angle violation (°)	3.242
Max. distance constraint violation (Å)	0.159
Deviations from idealized geometry	
Bond lengths (Å)	0.003 ± 0.000
Bond angles (°)	0.511 ± 0.011
Impropers (°)	0.317 ± 0.013
Average pairwise RMSD ^b (Å)	
Heavy	0.397 (2–20)
Backbone heavy	0.095 (2–20)

^aThe values for E_{NOE} were calculated from a square-well potential with force constants of 150 kcal mol⁻¹ Å².

^bPairwise RMSD was calculated among 20 refined structures.

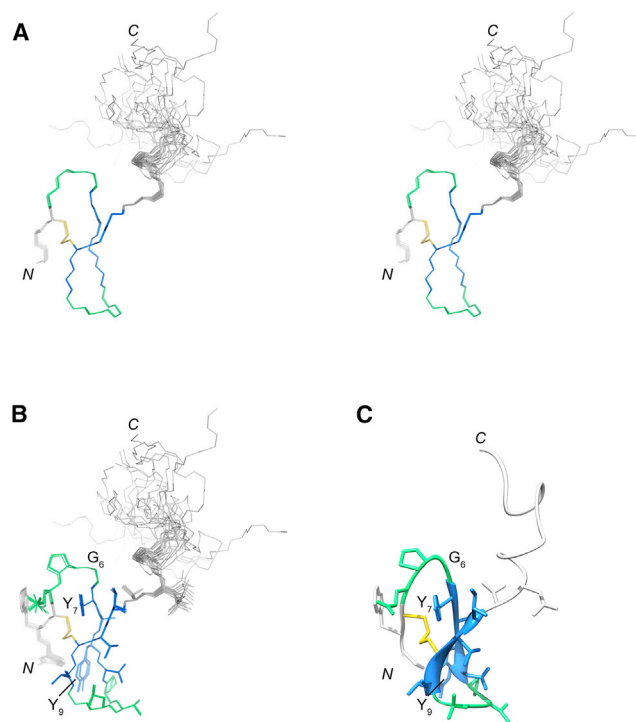
from the *Aplysia* genome, is present as a transcript, CNSN01-C-010033, in the neuronal transcriptome of *A. californica* (Moroz et al., 2006). Thus, contryphan-Vc1 may belong to a previously undescribed class of structurally related molluscan neuropeptides.

Peptide Synthesis

As contryphan-Vc1 was only available in limited amounts from its natural source, it was necessary to produce the peptide synthetically for further characterization. The linear peptide was synthesized by microwave-assisted solid-phase peptide synthesis (SPPS) using commercially available L-amino acids and subjected to oxidative folding in 0.1 M ammonium bicarbonate buffer to induce formation of the disulfide bond. Co-elution experiments (data not shown) and sedimentation velocity analysis (Table S1), suggested that the synthetic contryphan-Vc1 was chemically identical to that found in the venom.

NMR Spectroscopy

Initial examination of one- and two-dimensional ¹H nuclear magnetic resonance (NMR) spectra (Figure S2) indicated that contry-

**Figure 2. Structure of Contryphan-Vc1**

(A) Stereo view of the ensemble of the final 20 structures superimposed over the backbone heavy atoms of residues 2–20 (showing backbone heavy atoms). The N-terminal 20 residues form a well-defined β-hairpin motif while the C terminus is less well defined but shows some helical propensity. β turns are shown in green, β sheets in blue, disulfide in yellow, and the N and C termini are labeled.

(B) Ensemble of the final 20 structures superimposed over the backbone heavy atoms of residues 2–20 (showing backbone and side-chain heavy atoms). The side chains of Tyr7, Tyr9 (labeled), and the disulfide (yellow) make up the small hydrophobic core of the peptide.

(C) Closest-to-average structure of contryphan-Vc1.

phan-Vc1 adopted a single major conformation and was highly structured in solution. Chemical shifts were assigned by standard sequential assignment of total correlation spectroscopy (TOCSY), nuclear Overhauser effect spectroscopy (NOESY), and double quantum filtered correlation spectroscopy (DQF-COSY) spectra (293 K, pH 3.9). Complete ¹H, ¹³C, and ¹⁵N chemical-shift assignments have been deposited in the BMRB (accession code BMRB: 25585). Strong $d_{\alpha\beta}$ NOEs for Gln4-Pro5, Asn10-Pro11, and His25-Pro26 were diagnostic of *trans* conformations about each X-Pro bond.

Structures were generated in CYANA then refined in Xplor-NIH using a total of 489 NOE-derived distance constraints, 34 dihedral angle constraints (from ³J_{HN-Hα} J-coupling measurements and TALOS-N predictions), and five hydrogen bond constraints (from amide temperature coefficients and ²H exchange experiments). Structural constraints are summarized in Table 1. Determination of the solution structure of contryphan-Vc1 revealed that this peptide had a unique fold, not only among peptide toxins but among peptides in general, which can be described as an SDH (Figure 2). The most conspicuous feature of the structure is a double-stranded anti-parallel β sheet connected

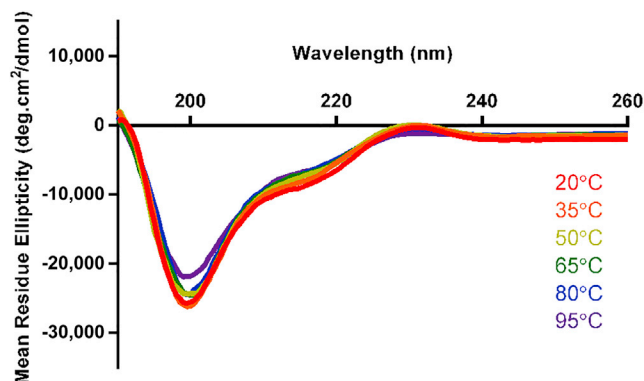


Figure 3. The SDH Fold of Contryphan-Vc1 Is Highly Resistant to Thermal Denaturation

Far-UV CD spectra of contryphan-Vc1 acquired from 20°C to 95°C (at 15°C intervals).

by a β turn, otherwise known as a β -hairpin motif. This region of the peptide (residues 1–20) is exceptionally well defined (backbone heavy atom pairwise root-mean-square deviation [RMSD] of 0.09 Å). While the single disulfide is likely to be critical in maintaining the high degree of structure, the side chains of residues Tyr7 and Tyr9 also appear to be important as they contribute the remainder of the small hydrophobic core of the peptide. The N-terminal pyroglutamic acid residue is well defined and in close proximity to the side chain of Tyr9. The four residues following Cys3 (Gln-Pro-Gly-Tyr) form a type-II β turn, leading into the first β strand. This β turn is stabilized by an $i - i + 3$ hydrogen bond between the carbonyl of Gln4 and the amide of Tyr7. The two strands are connected by a type-IV β turn around the four residues (Pro-Val-Leu-Gly). The second β strand incorporates Cys16 linking back to the N-terminal region via the disulfide. This strand terminates following a final hydrogen bond between the amide of Thr19 and the carbonyl of Gly6. Additional hydrogen bonds predicted as part of the anti-parallel β sheet were Ala8 H^N-Thr17 O, Thr 17 H^N-Ala8 O, and 10 H^N-15 O. All of the above amides exhibited slow ²H exchange, had temperature coefficients less negative than -4.75 ppb/K (Baxter and Williamson, 1997), and participated in hydrogen bonds (as defined in MOLMOL) in the initial structure calculations, and were therefore incorporated as structural restraints in the final structure calculations. The C-terminal tail of the peptide is less well defined but has regions with some helical propensity (Figure S3), with residues 20–22 and 28–30 forming short helical segments in 19 and 11 of the final structures, respectively. Temperature coefficients less negative than -4.75 ppb/K were observed for several residues of the C-terminal region (Arg22, His25, Asn28, Tyr29, and Asp30), indicating some degree of structure in this region.

Analysis of one-dimensional ¹H NMR spectra acquired at pH 7.0 (Figure S4A) indicated that this peptide fold was stable at physiological pH.

Thermal Stability

The thermal stability of contryphan-Vc1 was assessed using circular dichroism (CD) spectroscopy (Figure 3). Far-UV (260–190 nm) CD data acquired at 20°C (37 μ M) were consistent with a high content of β -strand and β -turn secondary structure

(Figure 3). Estimation of secondary structure composition using the CDSSTR algorithm (Whitmore and Wallace, 2004, 2008) (34% strand, 28% turn, 6% helix, and 31% unordered; normalized RMSD value of 0.016) was consistent with the solution structure determined on the basis of NMR data. Increasing the temperature to 95°C had little effect on the CD spectrum of contryphan-Vc1, demonstrating that the secondary structural features of this peptide are highly resistant to thermal denaturation. To probe the effect of temperature on the tertiary structure, we recorded one-dimensional ¹H NMR spectra of contryphan-Vc1 in water as a function of temperature. The excellent spectral dispersion evident at room temperature was fully maintained at 70°C (Figure S4B), the highest temperature accessible on the spectrometer, indicating that any thermally-induced change in the tertiary structure begins well above this temperature.

Activity

The intracranial injection (Olivera et al., 1990) of contryphan-Vc1 in mice did not produce any strong reproducible behavioral changes compared with control mice injected with saline (data not shown). Similarly, the peptide did not produce any observable changes to normal or depolarization-induced intracellular Ca²⁺ levels in mouse dorsal root ganglion cells (data not shown). It is possible that contryphan-Vc1 interacts with a receptor target in mice in such a way as to not produce observable responses in these assays, or, as *C. victoriae* is a molluscivorous species of *Conus*, it is possible that this peptide is selective for one or more molluscan receptors. Our observation that similar peptides are encoded in the genomes of *C. gigas* and *A. californica*, and transcribed in the CNS of the latter, suggests that these may represent a new class of molluscan neuropeptide for which a cognate receptor is likely to exist, and its identification is an exciting prospect for future studies.

There have been several reports of the presence of hormone/neuropeptide-like components in the venoms of cone snails, where they play a role in prey capture by disrupting neuronal and/or neuroendocrine function (Craig et al., 1999; Robinson et al., 2015; Safavi-Hemami et al., 2015). Contryphan-Vc1 may represent another such example.

DISCUSSION

Here we report the discovery of a unique peptide, contryphan-Vc1, from the venom of a cone snail. This peptide is highly structured and adopts a novel fold that we have designated the SDH. Even with just a single disulfide bridge, this structure is highly resistant to thermal denaturation.

To our knowledge, contryphan-Vc1 is the first example of an SDH, either naturally occurring or engineered. While a search for similar structures using Dali (Holm and Rosenström, 2010) did not yield any significant matches, it was clear that the fold of this single disulfide-containing peptide strongly resembled that of the multiple-disulfide-containing ICK structural motif. Of the many ICK peptides characterized structurally, hainantoxin-I (μ -theraphotoxin-Hhn2b), a sodium-channel blocker from the venom of the Chinese bird spider *Haplopelma hainanum*, is used here for sequence and structural comparisons. The SDH core motif of contryphan-Vc1 (residues 3–20) and hainantoxin-I (residues 16–33) align with a remarkably low RMSD of 0.75 Å

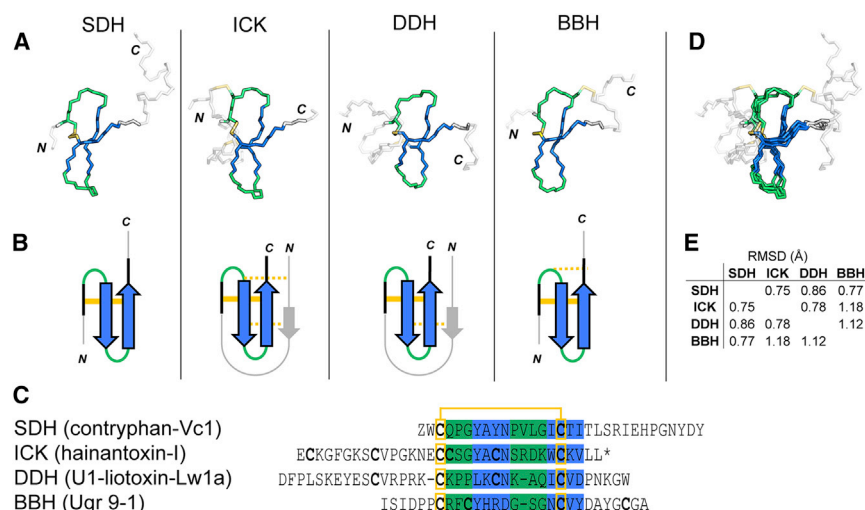


Figure 4. The SDH Fold Is a Basic Core Unit of Several Peptide Folds Including the Inhibitor Cysteine Knot

(A) Structures of contryphan-Vc1, hainantoxin-I (PDB: 1NIX), U₁-liotoxin-Lw1a (PDB: 2KYJ), and Ugr 9-1 (PDB: 2LZO) representing the SDH, ICK, DDH, and BBH folds, respectively. The core SDH unit of each fold is colored as follows: β turns, green; β strands blue; disulfides, yellow.

(B) Schematic illustrating the core SDH unit of each fold (using the same color scheme as in A).

(C) Structure-based sequence alignment of contryphan-Vc1, hainantoxin-I, U₁-liotoxin-Lw1a, and Ugr 9-1 representing the SDH, ICK, DDH, and BBH folds, respectively (using the same color scheme as in A).

(D) The SDH, ICK, DDH, and BBH folds, aligned over the 17 (or 16) residues of each fold comprising the basic SDH unit.

(E) The core SDH region of each of the four folds can be overlaid with backbone heavy atom root-mean-square deviation values of 0.75 Å (SDH

[residues 3–19] versus ICK [16–32]), 0.86 Å (SDH [3–10, 16–19] versus DDH [17–24, 29–32]), 0.77 Å (SDH [3–10, 16–19] versus BBH [7–14, 19–22]), 0.78 Å (ICK [16–23, 29–32] versus DDH [17–24, 29–32]), 1.18 Å (ICK [16–23, 29–32] versus BBH [7–14, 19–22]), and 1.12 Å (DDH [17–32] versus BBH [7–22]).

(backbone heavy atoms) (Figures 4D and 4E). A sequence alignment based on the structural overlay (Figure 4C) indicates that hainantoxin-I aligns well with contryphan-Vc1; 7 of the 20 aligned residues are shared and they share the same equivalent inter-cysteine loop-3, -4, and -5 lengths.

The shared SDH core of each fold consists of a single disulfide (Cys3–16 in contryphan-Vc1, Cys III–VI in the ICK fold), two anti-parallel β strands (residues 7–10 and 15–18 of contryphan-Vc1, residues 20–23 and 28–31 of hainantoxin-I) separated by a β turn, and, following the first half-cystine and preceding the first β strand, another β turn. In addition, the structural similarity between contryphan-Vc1 and hainantoxin-I extends to the backbone heavy atoms of several residues immediately following the second β strand. Both peptides share a Gly residue at the central position of the first β turn, and a Tyr in the following position. A turn-facilitating Gly (or Pro) residue at the former position is observed in the majority of ICK peptides. Similarly, the hydrophobic side chain (e.g. Tyr) at the latter position makes up part of the hydrophobic core of many ICK peptides.

Beyond the SDH core, the peptides differ in that the ICK fold of hainantoxin-I has an extended N terminus, whereas contryphan-Vc1 has an extended C terminus (Figure 4). Importantly, the N-terminal region of the ICK fold, which is missing in contryphan-Vc1, includes Cys I and II of the cystine knot. Therefore, contryphan-Vc1 does not have the Cys I–IV or Cys II–V disulfides of the ICK fold, and the equivalent positions in contryphan-Vc1 of Cys IV and V of the ICK fold are Gln4 and Tyr9, respectively.

The SDH core is common to other multiple disulfide-containing peptide folds (Figures 4A and 4B). It is present in the disulfide-directed β hairpin (DDH) fold found in U₁-liotoxin-Lw1a, a peptide from the venom of the scorpion *Liocheles waigiensis* (Smith et al., 2011), and as a subdomain of several eukaryotic body proteins such as the cellulose-binding domain of cellobiohydrolase I from the fungus *Trichoderma reesei* (Wang et al., 2000). The SDH core is also found in the recently reported boundless β hairpin (BBH) fold of the class9a sea anemone toxin Ugr9-1 (π -AnmTX Ugr 9a-1) (Osmakov et al., 2013).

A structural alignment of the shared SDH core of contryphan-Vc1 and the ICK, DDH, and BBH folds (Figure 4D) reveals remarkable similarity. The SDH core of contryphan-Vc1 overlays with the ICK (hainantoxin-I), DDH (U₁-liotoxin-Lw1a), and BBH (Ugr9-1) folds with backbone heavy atom RMSD values of 0.75, 0.86, and 0.77 Å, respectively (Figure 4E). Similarly, low RMSD values are observed between each of the ICK, DDH, and BBH folds. As well as sharing an overall fold, several key amino acids are shared; in each of the folds Pro or Gly is favored as the central residue of the first β turn, followed by an amino acid with a hydrophobic side chain, which forms part of the small hydrophobic core of the shared motif. It is clear that the SDH is the common core unit of these peptide folds.

Small, stable, synthetically accessible peptide scaffolds are useful as templates in drug design and for the engineering of novel bioactive peptides. The ICK fold, for example, has attracted attention for numerous such applications (Moore et al., 2012). Further efforts to identify the underlying core or minimal autonomous folding unit of the ICK fold have been made, with the hypothesis that such a peptide would represent a privileged peptide scaffold, valuable for the applications described above. Several lines of evidence had demonstrated that the N-terminal region of ICK peptides, including the I–IV disulfide, was not essential in maintaining the stability or three-dimensional structure of the ICK fold (Chiche et al., 1993; Le-Nguyen et al., 1993). Based on this, Min-23, an N-terminally truncated version of the ICK peptide EETI II, was engineered (Heitz et al., 1999). This peptide was shorter and, with only two disulfides, retained the high stability and tertiary structure of the ICK fold. It was termed the cystine-stabilized β sheet (CSB) motif and proposed as the elementary structural motif underlying the ICK fold, and indeed the potential of Min-23 as a scaffold for the development of new peptide ligands has since been confirmed (Souriau et al., 2005). It was, however, later recognized that the CSB motif could be viewed as just one example of a broader class of naturally occurring proteins/peptides defined by the two-disulfide DDH fold (Smith et al., 2011; Wang et al., 2000). As a result, the less

rigidly defined DDH fold was considered as the elementary structural motif underlying the ICK fold.

With the discovery of contryphan-Vc1 we show that a new structural motif, the SDH, forms the basic common core of not only the ICK and DDH folds but also the recently reported BBH fold. The SDH is a naturally occurring, small, highly stable, independent folding unit and is the common elementary motif underlying each of these peptide folds. As demonstrated elegantly in nature, the SDH motif can function as a scaffold for a vast array of functionalities. The exploration of this motif as a starting point for the design and engineering of new bioactive peptides will form the basis of future studies.

EXPERIMENTAL PROCEDURES

Detailed experimental procedures are provided in [Supplemental Experimental Procedures](#). In brief, venom was obtained by manual extrusion from freshly dissected *C. victoriae* venom glands (several specimens), and three samples (unprocessed venom, reduced and alkylated venom, and reduced, alkylated, and trypsinized venom) were prepared for MS. The mature peptide of contryphan-Vc1 and its associated post-translational modifications were confirmed by searching a high-resolution MS/MS profile of reduced, alkylated, and trypsinized *C. victoriae* venom against the translated venom gland transcriptome of *C. victoriae*, the preparation of which has been described previously (Robinson et al., 2014).

Contryphan-Vc1 was generated by SPPS and analyzed using NMR spectroscopy. Chemical-shift assignments for backbone and side-chain protons were made by conventional analysis of two-dimensional homonuclear TOCSY, NOESY, and DQF-COSY spectra. For structure calculations, assigned cross-peaks of NOESY spectra (200 ms mixing time) were used to generate distance constraints, $^3J_{\text{HN-H}\alpha}$ J -coupling measurements and TALOS-N predictions were used to generate dihedral angle constraints, and amide ^2H exchange rates and temperature coefficients were used in defining hydrogen bond constraints. Structures were optimized for a low target function in CYANA before the final constraint set was entered into Xplor-NIH, where conventional simulated annealing protocols were used to generate a new ensemble of 200 structures, from which the 20 lowest-energy structures were chosen to represent the solution structure of contryphan-Vc1.

Thermal stability was assessed by analysis of far-UV CD spectra and one-dimensional ^1H NMR spectra acquired over a range of temperatures.

ACCESSION NUMBERS

The accession number for contryphan-Vc1 are Uniprot: W4VSF6 and PDB: 2N24.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2015.11.015>.

AUTHOR CONTRIBUTIONS

Conceptualization: S.D.R. and R.S.N.; investigation: S.D.R., S.C., A.B., B.C., and H.S.-H.; writing—original draft: S.D.R.; writing—review and editing: S.D.R. and R.S.N.; resources: A.T.P., A.W.P., A.J.R., and R.S.N.

ACKNOWLEDGMENTS

We thank Johan Pas for specimen collection and Dr. Nicholas Williamson for technical assistance with MS. We thank Dr. Yee-Foong Mok for technical assistance with analytical ultracentrifugation and CD spectroscopy. The authors acknowledge financial support from a Discovery Grant (DP110101331) from the Australian Research Council (A.W.P.). A.W.P. and R.S.N. acknowl-

edge fellowship support from the Australian National Health and Medical Research Council (NHMRC). A.T.P. was supported by an NHMRC Career Development Fellowship (1003856) and a Program Grant (1054618), and benefited from support by the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC Independent Research Institute Infrastructure Support Scheme. H.S.-H. is supported by a Marie Curie Fellowship from the European Commission (CONBIOS 330486). S.D.R. received support from a Monash University Postgraduate Publications Award.

Received: June 22, 2015

Revised: November 11, 2015

Accepted: November 23, 2015

Published: January 7, 2016

REFERENCES

- Baxter, N., and Williamson, M. (1997). Temperature dependence of ^1H chemical shifts in proteins. *J. Biomol. NMR* 9, 359–369.
- Chiche, L., Heitz, A., Padilla, A., Le-Nguyen, D., and Castro, B. (1993). Solution conformation of a synthetic bis-headed inhibitor of trypsin and carboxypeptidase A: new structural alignment between the squash inhibitors and the potato carboxypeptidase inhibitor. *Protein Eng.* 6, 675–682.
- Craig, A.G., Norberg, T., Griffin, D., Hoeger, C., Akhtar, M., Schmidt, K., Low, W., Dykert, J., Richelson, E., Navarro, V., et al. (1999). Conatulakin-G, an O-glycosylated invertebrate neurotensin. *J. Biol. Chem.* 274, 13752–13759.
- Evans, B.E., Rittle, K.E., Bock, M.G., DiPardo, R.M., Freidinger, R.M., Whitter, W.L., Lundell, G.F., Veber, D.F., and Anderson, P.S. (1988). Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 31, 2235–2246.
- Gelly, J.C., Gracy, J., Kaas, Q., Le-Nguyen, D., Heitz, A., and Chiche, L. (2004). The KNOTTIN website and database: a new information system dedicated to the knottin scaffold. *Nucleic Acids Res.* 32, D156–D159.
- Heitz, A., Le-Nguyen, D., and Chiche, L. (1999). Min-21 and Min-23, the smallest peptides that fold like a cystine-stabilized β -sheet motif: Design, solution structure, and thermal stability. *Biochemistry* 38, 10615–10625.
- Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38, W545–W549.
- Jiménez, E.C., Olivera, B.M., Gray, W.R., and Cruz, L.J. (1996). Contryphan is a D-tryptophan-containing *Conus* peptide. *J. Biol. Chem.* 271, 28002–28005.
- Le-Nguyen, D., Heitz, A., Chiche, L., Hajji, M.E., and Castro, B. (1993). Characterization and 2D NMR study of the stable [9–21, 15–27] 2 disulfide intermediate in the folding of the 3 disulfide trypsin inhibitor EETI II. *Protein Sci.* 2, 165–174.
- Moore, S.J., Leung, C.L., and Cochran, J.R. (2012). Knottins: disulfide-bonded therapeutic and diagnostic peptides. *Drug Discov. Today Technol.* 9, e3–e11.
- Moroz, L.L., Edwards, J.R., Puthanveetil, S.V., Kohn, A.B., Ha, T., Heyland, A., Knudsen, B., Sahni, A., Yu, F., Liu, L., et al. (2006). Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. *Cell* 127, 1453–1467.
- Norton, R.S., and Pallaghy, P.K. (1998). The cystine knot structure of ion channel toxins and related polypeptides. *Toxicol.* 36, 1573–1583.
- Norton, R.S., and Olivera, B.M. (2006). Conotoxins down under. *Toxicol.* 48, 780–798.
- Olivera, B., Rivier, J., Clark, C., Ramilo, C.A., Corpuz, G.P., Abogadie, F.C., Edward Mena, E., Woodward, S.R., Hillyard, D.R., and Cruz, L.J. (1990). Diversity of *Conus* neuropeptides. *Science* 249, 257–263.
- Orengo, C.A., Flores, T.P., Jones, D.T., Taylor, W.R., and Thornton, J.M. (1993). Recurring structural motifs in proteins with different functions. *Curr. Biol.* 3, 131–139.
- Osmakov, D.I., Kozlov, S.A., Andreev, Y.A., Koshelev, S.G., Sanamyan, N.P., Sanamyan, K.E., Dyachenko, I.A., Bondarenko, D.A., Murashev, A.N., Mineev, K.S., et al. (2013). Sea anemone peptide with uncommon β -hairpin structure inhibits acid-sensing ion channel 3 (ASIC3) and reveals analgesic activity. *J. Biol. Chem.* 288, 23116–23127.

- Pallaghy, P.K., Nielsen, K.J., Craik, D.J., and Norton, R.S. (1994). A common structural motif incorporating a cystine knot and a triple-stranded β -sheet in toxic and inhibitory polypeptides. *Protein Sci.* 3, 1833–1839.
- Reeks, T.A., Fry, B.G., and Alewood, P.F. (2015). Privileged frameworks from snake venom. *Cell. Mol. Life Sci.* 72, 1939–1958.
- Robinson, S.D., and Norton, R.S. (2014). Conotoxin gene superfamilies. *Mar. Drugs* 12, 6058–6101.
- Robinson, S.D., Safavi-Hemami, H., McIntosh, L.D., Purcell, A.W., Norton, R.S., and Papenfuss, A.T. (2014). Diversity of conotoxin gene superfamilies in the venomous snail, *Conus victoriae*. *PLoS One* 9, e87648.
- Robinson, S.D., Safavi-Hemami, H., Raghuraman, S., Imperial, J.S., Papenfuss, A.T., Teichert, R.W., Purcell, A.W., Olivera, B.M., and Norton, R.S. (2015). Discovery by proteogenomics and characterization of an RF-amide neuropeptide from cone snail venom. *J. Proteomics* 114, 38–47.
- Safavi-Hemami, H., Gajewiak, J., Karanth, S., Robinson, S.D., Ueberheide, B., Douglass, A.D., Schlegel, A., Imperial, J.S., Watkins, M., Bandyopadhyay, P.K., et al. (2015). Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc. Natl. Acad. Sci. USA* 112, 1743–1748.
- Smith, J.J., Hill, J.M., Little, M.J., Nicholson, G.M., King, G.F., and Alewood, P.F. (2011). Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif. *Proc. Natl. Acad. Sci. USA* 108, 10478–10483.
- Souriau, C., Chiche, L., Irving, R., and Hudson, P. (2005). New binding specificities derived from Min-23, a small cystine-stabilized peptidic scaffold. *Biochemistry* 44, 7143–7155.
- Wang, X.-h., Connor, M., Smith, R., Maciejewski, M.W., Howden, M.E.H., Nicholson, G.M., Christie, M.J., and King, G.F. (2000). Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat. Struct. Mol. Biol.* 7, 505–513.
- Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 32, W668–W673.
- Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89, 392–400.
- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., et al. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49–54.

Supplemental Information

A Naturally Occurring Peptide with an Elementary

Single Disulfide-Directed β -Hairpin Fold

Samuel D. Robinson, Sandeep Chhabra, Alessia Belgi, Balasubramanyam Chittoor, Helena Safavi-Hemami, Andrea J. Robinson, Anthony T. Papenfuss, Anthony W. Purcell, and Raymond S. Norton

SUPPLEMENTAL DATA:

contryphan-Vc1 (*Conus victoriae*)
MGKLTILFLVAAALLSTQVMVQGDGAHERTEAEEPQHHGAKRQDGTGGYPVDDVDMQRIERTPLKRWQCQPGYAYNPVLGICTITLSRIEHPGNYDYRRGRQ
contryphan-R (*Conus radiatus*)
MGKLTILVLVAAVLLSAQVMVQGDGDQPADRNAVPRDDNPGGASGKFMNVLRRSGCPWEPCG
contryphan-Sm (*Conus stercusmuscarum*)
MGKLTILVLVAAVLLSTQVMVQGDADQPADRDVPRDDNPSGTDGKFMNVLRRFGCPWQPCG

Figure S1, related to Figure 1: Contryphan-Vc1 is member of the contryphan superfamily of conotoxins but shares no obvious sequence similarity with other contryphans. A comparison of the precursor sequence of contryphan-Vc1 with those of contryphan-R (Uniprot: P58786) and contryphan-Sm (Uniprot: P58787) from the venoms of *Conus radiatus* and *Conus stercusmuscarum*, respectively. Precursor signal peptides are highlighted in purple, dibasic cleavage sites in red and the mature peptides in blue.

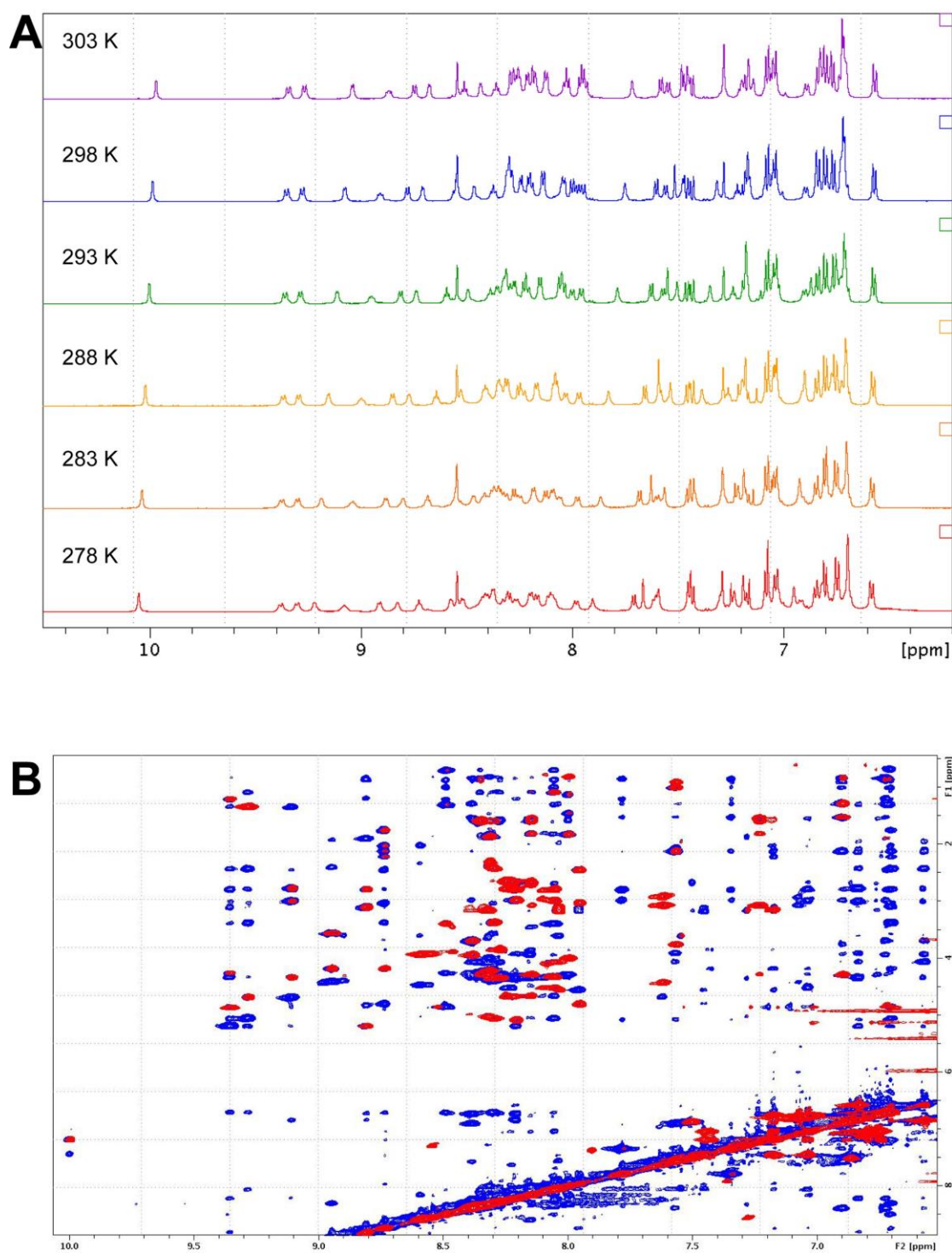


Figure S2, related to Figure 2: Contryphan-Vc1 adopts a single major conformation and is highly structured in solution. (A) Amide/aromatic region of one-dimensional ^1H NMR spectra as a function of temperature for contryphan-Vc1 in 93% H_2O /7% $^2\text{H}_2\text{O}$ at pH 3.9, acquired on a Bruker 600 MHz spectrometer. (B) Amide/aromatic region of NOESY (200 ms) (blue) and TOCSY (80 ms) (red) spectra of contryphan-Vc1 in 93% H_2O /7% $^2\text{H}_2\text{O}$ (293 K, pH 3.9).

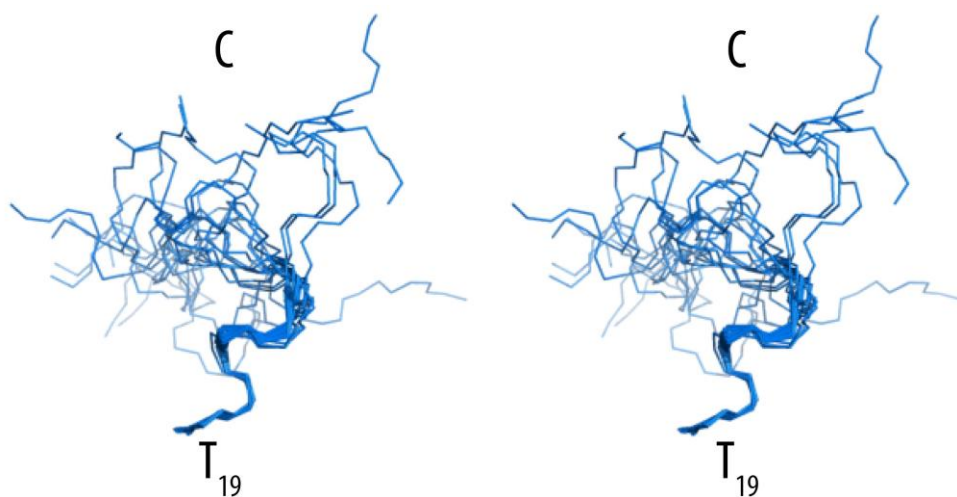


Figure S3, related to Figure 2: The C-terminal tail of contryphan-Vc1 (residues 19-31) shows some helical propensity. Stereo view of residues 19-31 of the final 20 structures of contryphan-Vc1, superimposed over the C_α atoms of residues 19-23.

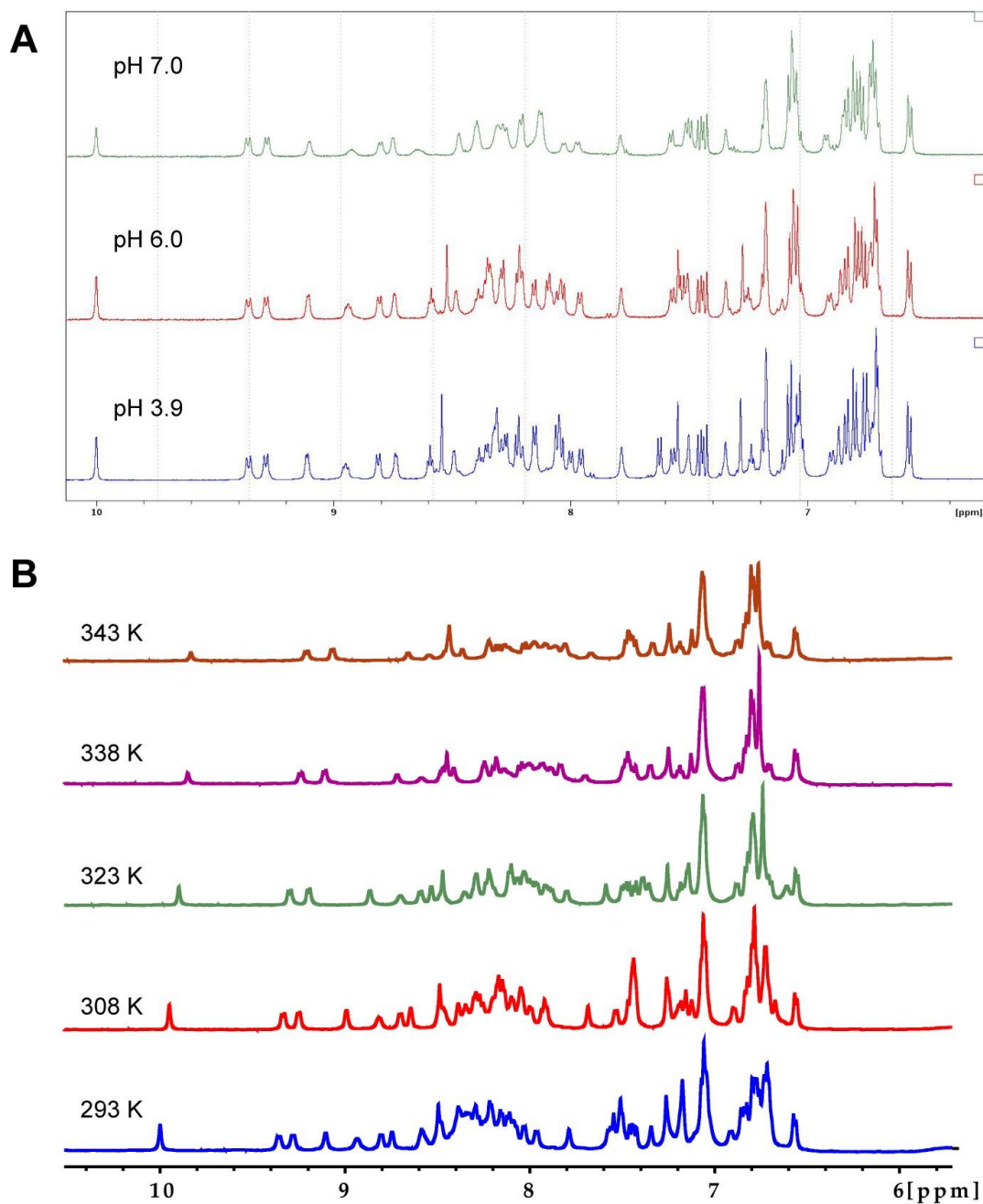


Figure S4, related to Figure 3: Contryphan-Vc1 is stable over a range of pH values and at high temperature. (A) Amide/aromatic region of one-dimensional ^1H NMR spectra of contryphan-Vc1 at pH 3.9, 6.0 and 7.0, acquired on a Bruker 600 MHz spectrometer. (B) Amide/aromatic region of one-dimensional ^1H NMR spectra of contryphan-Vc1, acquired on a Bruker 600 MHz spectrometer up to 343 K (70 °C) in water at pH 3.9. Most amide resonances shift upfield at higher temperature as expected and some are broader at pH 7 and at higher temperatures because of exchange with solvent water.

Table S1, related to Figure 3: Hydrodynamic properties of contryphan-Vc1 calculated from sedimentation velocity analysis.

Sample concentration	Weight average sedimentation coefficient of peaks (S)	Estimated M _r ^a	Frictional ratio
37 μM	0.51 (96 %)	3.3 kDa	1.40
110 μM	0.55 (92 %)	3.5 kDa	1.71
	1.72 (8 %)	18.7 kDa	

^a Theoretical M_r of monomeric contryphan-Vc1 is 3.55 kDa.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Venom gland transcriptome. Specimens of *C. victoriae* were collected from Broome, Western Australia. Preparation of the venom gland tissue and subsequent sequencing of the transcriptome has been described previously (Robinson et al., 2014). Briefly, whole venom glands of live specimens were dissected, snap-frozen in liquid nitrogen and stored at -80°C. Frozen venom glands were pulverized and homogenized prior to extraction of total RNA with Trizol (Invitrogen, Life Technologies). cDNA library preparation, normalization and 454 sequencing were performed by Eurofins, MWG Operon (Budendorf, GER). *De novo* transcriptome assembly was performed using MIRA3 (Chevreux et al., 2004). Conotoxin sequences were annotated using profile hidden Markov models built from known conotoxin superfamily sequences (Robinson et al., 2014).

Venom extraction & preparation. Venom samples obtained by manual extrusion from freshly dissected venom glands were snap-frozen in liquid nitrogen and stored at -80°C. Venom (from several specimens) was reconstituted in 0.1% trifluoroacetic acid (TFA), pooled and homogenized using a glass Dounce tissue grinder. Insoluble material was pelleted by centrifugation, and the supernatant was collected and lyophilized. Pellets were resuspended in 0.1% TFA / 20% acetonitrile (MeCN) then centrifuged, and the supernatant was collected and lyophilized. This process was repeated with 40% and 60% MeCN. Lyophilized venom was resuspended in 2% acetonitrile (MeCN), 0.1% TFA and pooled. Protein concentration was determined using a Bradford assay with ovalbumin as the standard. An aliquot of the venom sample was reduced in 20 mM tris(2-carboxyethyl)phosphine (TCEP) (pH 8) for 30 min at 60°C, then alkylated by incubating in 40 mM iodoacetamide for 30 min. The reduced and alkylated venom was further processed by tryptic digestion, essentially according to the manufacturer's instructions (proteomics-grade trypsin, Sigma-Aldrich), with an incubation time of 4 h at 37°C.

Mass spectrometry (MS) and MS matching. Aliquots (0.5 µg) of the unprocessed venom and the reduced, alkylated and trypsinized venom, were centrifuged at 13,000 x g for 10 min and loaded onto a microfluidic trap column packed with ChromXP C18-CL 3 µm particles (300 Å nominal pore size; equilibrated in 0.1% formic acid/5 % MeCN) at 5 µL/min using an Eksigent NanoUltra cHiPLC system. An analytical (15 cm x 75 µm ChromXP C18-CL 3) microfluidic column was then switched in line and venom separated using a linear gradient elution of 0-80 % MeCN/ 0.1 % formic acid over 90 min (300 nL/min). Separated venom peptides were analysed using an AB SCIEX 5600 TripleTOF mass spectrometer equipped with a Nanospray III ion source and accumulating 30 tandem mass spectrometry (MS/MS) spectra per second. Data were processed in ProteinPilot software (version Beta 4.1.46) using the Paragon algorithm. The search databases comprised a six-frame translation of the *C. victoriae* venom gland transcriptome database, as described previously (Robinson et al., 2014). Hits identified by ProteinPilot were then validated by comparison of experimentally-derived peaks against a theoretical peak list (Protein Prospector MS-Product, University of California, San Francisco).

Peptide synthesis and folding. Contryphan-Vc1 was synthesized on Wang resin (loading 1.1 mmol/g, 0.091 g, 0.1 mmol); the C-terminal amino acid was double-coupled manually to the resin (Fmoc-protected amino acid: dimethylamino pyridine (DMAP): diisopropylcarbodiimide, 3:0.1:3 equivalents respectively) and full chain synthesis was completed using microwave-assisted solid-phase peptide synthesis on a CEM Liberty-Discovery™ synthesiser. Resin was swollen for 60 min in *N,N*'-dimethylformamide (DMF): dichloromethane (DCM) (1:1). Coupling conditions were as follows: Fmoc-amino acids (0.2 M in DMF), activator (0.5 M *O*-(7-azabenzotriazol-1-yl)-*N,N,N,N*'-tetramethyluronium hexafluorophosphate (HATU) in DMF) and activator base (2 M *N,N*'-diisopropylethylamine (DIPEA) in *N*-methylpyrrolidone (NMP)). Fmoc-deprotection was accomplished using 20 % v/v piperidine in DMF. The default microwave conditions used in the synthesis of the linear peptide were as follows: initial deprotection (40 W, 75 °C, 0.5 min), deprotection (40 W, 75 °C, 3 min), and coupling (20 W, 75°C, 5 min). In the case of cysteine and histidine, the coupling was carried out at 50 °C with a multistep microwave method consisting of pre-activation (0 W, 50 °C, 2 min) and coupling (25 W, 50 °C, 4 min). Acid-mediated cleavage was performed at room temperature for 4 h, using a solution of 95:2:2:1 (v/v/v/v) TFA: triisopropylsilyl (TIPS):water:thioanisole. Air oxidation of the crude linear peptide was carried out at a concentration of 0.3 mg/mL in 0.1 M ammonium bicarbonate, according to the procedure described by Clark et al. (Clark et al., 2006). The reaction progress was monitored by reversed-phase RP-HPLC and ESI-MS. Folding was quenched by acidification to pH 2 with glacial acetic acid and the oxidised peptide was purified by preparative RP-HPLC to > 95% purity. Analytical and preparative RP-HPLC were performed on Agilent 1200 series instruments using a Vydac C18 analytical column (4.6 x 250 mm, 5 µm) at 1.5 mL/min and preparative column (22 x 250 mm, 10 µm) at 10 mL/min, respectively.

Co-elution. While we were able to confirm, by MS, the complete primary structure of the peptide, including the *N*-terminal pyroglutamic acid modification and disulfide bond formation, one post-translational modification potentially remained undetected: D-amino acid epimerization, which has been reported in other contryphans on the basis of co-elution experiments (Jiménez et al., 1996). This modification does not cause a shift in mass and cannot be observed by MS, so in order to determine whether the native venom peptide contained any D-amino acids, a simple co-elution experiment was performed with the synthetic peptide. A 100 µg sample of *C. victoriae* crude venom was loaded onto a Luna C-8 column (100 Å, 2 x 100 mm) coupled to a Shimadzu LCMS2020 instrument, and separated using a linear gradient of 0-60% MeCN/0.05% TFA over 30 min. An extracted ion chromatogram for the expected $[M+3H]^{+3}$ ion of contryphan-Vc1 indicated a peak eluting at 24.3 min. This was repeated with a ~500 ng sample of synthetic contryphan-Vc1 and an extracted ion chromatogram for the same ion indicated a peak eluting at the identical retention time (24.3 min) (data not shown).

Sedimentation velocity analysis. Several conotoxins are known to be present in venom as dimers (Loughnan et al., 2006; Quinton et al., 2009; Walker et al., 2009; Wu et al., 2010). Confirmation of monomeric contryphan-Vc1 in the unprocessed venom ruled out the

possibility of covalent dimerization. To explore the possibility of non-covalent homodimerisation we performed sedimentation velocity analysis of the peptide.

Samples were analyzed using an XL-I analytical ultracentrifuge (Beckman Coulter, Fullerton, CA) equipped with an AnTi-60 rotor. Contryphan-Vc1 samples at 37 μ M or 110 μ M (in water, pH 3.9) were added to double-sector Epon-filled centerpieces, with water in the reference compartment. Radial absorbance data was acquired at 20 °C using a rotor speed of 50,000 rpm and a wavelength of 280 nm, with radial increments of 0.003 cm in continuous scanning mode. The sedimenting boundaries were fitted to a model that describes the sedimentation of a distribution of sedimentation coefficients with no assumption of heterogeneity ($c(s)$) using the program SEDFIT (Schuck and Rossmanith, 2000). Data were fitted using a regularization parameter of $p = 0.95$, floating frictional ratios, and 150 sedimentation coefficient increments in the range of 0.1–10 S.

At both concentrations tested (37 and 110 μ M), the peptide was monomeric (**Table S1**), with a small percentage (8 %) of higher order aggregate present at the higher concentration (110 μ M).

NMR spectroscopy. Samples were prepared by dissolving the freeze-dried peptide in either 93% $H_2O/7\%$ 2H_2O or 100% 2H_2O (pH 3.9). The peptide did not dissolve completely under these conditions; insoluble material was pelleted by centrifugation and the supernatant used for further analysis (with a final concentration, as determined by A_{280} , of ~ 110 μ M). All spectra were acquired on a Bruker Avance 600 MHz instrument. One-dimensional 1H spectra were acquired at temperatures over the range 278–303 K, at intervals of 5 K, as well as 293 – 343 K at intervals of 15 K, and pH 3.9 and 4.8 (293 K). Two-dimensional homonuclear TOCSY spectra (spin lock time 80 ms) were acquired at 288 K, 293 K, and 298 K (pH 3.9) and at 293 K (pH 4.8). Two-dimensional homonuclear NOESY spectra (mixing times 50 and 200 ms) were acquired at 293 K, pH 3.9 in 93% $H_2O/7\%$ 2H_2O and 100% 2H_2O . DQF-COSY spectra were acquired at 293 K, pH 3.9 in 100% 2H_2O . ^{15}N SOFAST-HMQC and ^{13}C HSQC spectra were acquired at 293 K, pH 3.9 in 93% $H_2O/7\%$ 2H_2O and 100% 2H_2O , respectively. Chemical shift assignments for backbone and side chain protons were made by conventional analysis of two-dimensional TOCSY, NOESY and DQF-COSY spectra (293 K, pH 3.9). 1,4-Dioxane (chemical shift of 3.75) was used as a reference for 1H chemical shifts, while ^{15}N and ^{13}C chemical shifts were determined indirectly. All spectra were processed in TOPSPIN (version 3.2) and analyzed using CcpNmr-Analysis (version 2.1.5). Amide deuterium (2H) exchange rates were monitored by the acquisition of a series of one-dimensional 1H spectra and two-dimensional homonuclear TOCSY spectra following reconstitution of the freeze-dried peptide in 100% 2H_2O . Amide temperature coefficients were determined by analysis of one-dimensional 1H spectra and TOCSY spectra acquired over temperature ranges 278–303 K and 288–298 K, respectively.

Assigned cross-peaks of NOESY spectra (200 ms mixing time) acquired at 293 K, pH 3.9, were used to generate distance constraints for structure calculations. A combination of $^3J_{HN-H\alpha}$ J -coupling measurements from one-dimensional 1H spectra and two-dimensional homonuclear NOESY spectra (acquired with 32K data points) as well as TALOS-N predictions (Shen and Bax, 2013) were used to generate dihedral angle constraints. Hydrogen bond constraints, consistent with measured amide 2H exchange rates and temperature

coefficients, and where hydrogen bond acceptors could be determined from initial rounds of structure calculation (present and consistent in $\geq 80\%$ of structures), were included for subsequent calculations. Structures were optimized for a low target function in CYANA (Güntert, 2004) and initial structures were used to resolve the assignment of several ambiguous inter-residue NOEs. Once optimized in CYANA, the final constraint set was entered into XPLOR-NIH (Schwieters et al., 2003) and with conventional simulated annealing protocols a new ensemble of 200 structures was generated. The 20 lowest energy structures generated were chosen to represent the solution structure of contryphan-Vc1. PROCHECK-NMR (Laskowski et al., 1996) was used for the validation of final calculated structures and the final ensemble of 20 structures (residues 2-20) had 78.6% of residues in favored and 21.4% in allowed Ramachandran-plot regions. Structures were analyzed in MOLMOL (Koradi et al., 1996) and molecular representations were prepared using PyMOL [DeLano WL. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC (<http://www.pymol.org>)] and the UCSF Chimera package.

Circular dichroism spectroscopy. Far UV CD spectra were recorded for contryphan-Vc1 (37 μM , in water, pH 3.7) using an Aviv 410SF CD spectrometer (Lakewood, NJ). Spectra were acquired between 190 and 260 nm (far UV) from 20 to 95 °C, using a 1 mm path-length quartz cuvette, 0.5 nm step size, and 4 s averaging. CD spectroscopy data were analysed for secondary structure content using the CDSSTR program (Whitmore and Wallace, 2004, 2008) reference database 4 (Sreerama and Woody, 2000).

Data availability. The nucleotide sequence of the contryphan-Vc1 precursor has been deposited at DDBJ/EMBL/GenBank [Accession: GAIH01000059] and Uniprot [Accession: W4VSF6]. The solution structure and chemical shifts of contryphan-Vc1 have been deposited in the PDB [accession code: 2N24] and BMRB [accession code: 25585], respectively.

SUPPLEMENTAL REFERENCES

- Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E.G., Wetter, T., and Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res.* *14*, 1147-1159.
- Clark, R.J., Fischer, H., Nevin, S.T., Adams, D.J., and Craik, D.J. (2006). The synthesis, structural characterization, and receptor specificity of the α -conotoxin Vc1.1. *J. Biol. Chem.* *281*, 23254-23263.
- Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* *278*, 353-378.
- Jiménez, E.C., Olivera, B.M., Gray, W.R., and Cruz, L.J. (1996). Contryphan is a D-tryptophan-containing *Conus* peptide. *J. Biol. Chem.* *271*, 28002-28005.
- Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* *14*, 51-55.
- Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* *8*, 477-486.
- Loughnan, M., Nicke, A., Jones, A., Schroeder, C.I., Nevin, S.T., Adams, D.J., Alewood, P.F., and Lewis, R.J. (2006). Identification of a novel class of nicotinic receptor antagonists: Dimeric conotoxins VxXIIA, VxXIIB, and VxXIIC from *Conus vexillum*. *J. Biol. Chem.* *281*, 24745-24755.
- Quinton, L., Gilles, N., and De Pauw, E. (2009). TxXIIIA, an atypical homodimeric conotoxin found in the *Conus textile* venom. *J. Proteomics* *72*, 219-226.
- Schuck, P., and Rossmannith, P. (2000). Determination of the sedimentation coefficient distribution by least-squares boundary modeling. *Biopolymers* *54*, 328-341.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* *160*, 65-73.
- Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* *56*, 227-241.
- Sreerama, N., and Woody, R.W. (2000). Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* *287*, 252-260.
- Walker, C.S., Jensen, S., Ellison, M., Matta, J.A., Lee, W.Y., Imperial, J.S., Duclos, N., Brockie, P.J., Madsen, D.M., Isaac, J.T.R., *et al.* (2009). A novel *Conus* snail polypeptide causes excitotoxicity by blocking desensitization of AMPA receptors. *Curr. Biol.* *19*, 900-908.
- Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* *32*, W668-W673.
- Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* *89*, 392-400.
- Wu, X.-C., Zhou, M., Peng, C., Shao, X.-X., Guo, Z.-Y., and Chi, C.-W. (2010). Novel conopeptides in a form of disulfide-crosslinked dimer. *Peptides* *31*, 1001-1006.

Appendix II



Structure and activity of contryphan-Vc2: Importance of the D-amino acid residue



Stephen B. Drane^a, Samuel D. Robinson^a, Christopher A. MacRaild^a, Sandeep Chhabra^a, Balasubramanyam Chittoor^a, Rodrigo A.V. Morales^a, Eleanor W.W. Leung^a, Alessia Belgi^b, Samuel S. Espino^c, Baldomero M. Olivera^c, Andrea J. Robinson^b, David K. Chalmers^a, Raymond S. Norton^{a,*}

^a Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville 3052, Victoria, Australia

^b School of Chemistry, Monash University, Clayton 3800, Victoria, Australia

^c Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

ARTICLE INFO

Article history:

Received 13 October 2016

Received in revised form

11 January 2017

Accepted 16 February 2017

Available online 17 February 2017

Keywords:

Conotoxin

Contryphan-Vc2

Nuclear magnetic resonance

Solution structure

Molecular dynamics

D-amino acid

ABSTRACT

In natural proteins and peptides, amino acids exist almost invariably as L-isomers. There are, however, several examples of naturally-occurring peptides containing D-amino acids. In this study we investigated the role of a naturally-occurring D-amino acid in a small peptide identified in the transcriptome of a marine cone snail. This peptide belongs to a family of peptides known as contryphans, all of which contain a single D-amino acid residue. The solution structure of this peptide was solved by NMR, but further investigations with molecular dynamics simulations suggest that its solution behaviour may be more dynamic than suggested by the NMR ensemble. Functional tests in mice uncovered a novel bioactivity, a depressive phenotype that contrasts with the hyperactive phenotypes typically induced by contryphans. Trp3 is important for bioactivity, but this role is independent of the chirality at this position. The D-chirality of Trp3 in this peptide was found to be protective against enzymatic degradation. Analysis by NMR and molecular dynamics simulations indicated an interaction of Trp3 with lipid membranes, suggesting the possibility of a membrane-mediated mechanism of action for this peptide.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The contryphans are a family of small disulfide-cyclised peptides found in the venoms of marine cone snails of the genus *Conus*.

Abbreviations: BMRB, Biological Magnetic Resonance Data Bank; BSA, bovine serum albumin; DIPEA, *N,N*-diisopropylethylamine; DMF, *N,N*-dimethylformamide; DPC, dodecylphosphocholine; Hyp, hydroxyproline; LC-MS, liquid chromatography/mass spectroscopy; MD, molecular dynamics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PDB, Protein Data Bank; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; POPG, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol; ROESY, rotating-frame Overhauser effect spectroscopy; RP-HPLC, reversed-phase high performance liquid chromatography; TFA, trifluoroacetic acid; TOCSY, total correlation spectroscopy. The abbreviations for the common amino acids (L-isomers unless indicated otherwise) are in accordance with the recommendations of the IUPAC-IUB Joint Commission on Biochemical Nomenclature (*Eur. J. Biochem.* 1984, 138:9–37).

* Corresponding author. Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Victoria 3052, Australia.

E-mail address: Ray.Norton@monash.edu (R.S. Norton).

All contryphans investigated to date produce strong behavioural effects when administered by intracranial injection in mice (Jacobsen et al., 1999; Jimenez et al., 1997, 1996, 2001). The molecular basis for this activity is unknown, although target receptors for four contryphans have been proposed: voltage-gated Ca²⁺ channels (contryphan-Lo and contryphan-Am) (Sabareesh et al., 2006), L-type Ca²⁺ channels (glacontryphan-M) (Hansson et al., 2004) and Ca²⁺-dependent K⁺ channels (contryphan-Vn) (Massilia et al., 2003). Contryphans range in length from 7 to 11 residues and in most cases share the consensus sequence CO(DW)XPWC. A sequence alignment of the known contryphans that have been investigated at the protein level is presented in Table 1. Consensus features include two Pro residues at positions 2 and 5 of the intercytine loop, with the first proline usually modified to hydroxyproline. A C-terminal Pro-Trp-Cys tripeptide is also present in all family members characterised to date, except contryphan-Tx (in which Trp is replaced by Tyr) (Jimenez et al., 2001). C-terminal amidation is almost ubiquitous, being absent only in Leu-contryphan-P (Jacobsen et al., 1999). Several members also have

Table 1

Sequence alignment of members of the contryphan family. W = D-tryptophan, L = D-leucine, W = L-6-bromo-tryptophan, O = hydroxyproline, γ = gamma-carboxyglutamic acid, * = C-terminal amidation. Adapted from (Thakur and Balarum, 2007).

Designation	Sequence	Ref
Contryphan-Vc2	----CRWTPVC*	(Robinson et al., 2014)
Leu-contryphan-P	---GCVLLPWC*	(Jacobsen et al., 1998)
Leu-contryphan-Tx	---CVLYPWC*	(Jimenez et al., 2001)
Contryphan-In (In936)	---GCVLYPWC*	(Thakur and Balarum, 2007)
Glacontryphan-M	NYSYCPWHPWC*	(Hansson et al., 2004)
Contryphan-Vn	--GDCPWKPWC*	(Massilia et al., 2001)
Contryphan-Lo (Lo959)	---GCPWDPWC*	(Sabareesh et al., 2006)
Contryphan-Tx	---GCOWQPYC*	(Jimenez et al., 2001)
Contryphan-Sm	---GCOWQPYC*	(Jacobsen et al., 1998)
Contryphan-P	---GCOWDPWC*	(Jacobsen et al., 1998)
Contryphan-R	---GCOWEPWC*	(Jimenez et al., 1996)
Des(Gly1) contryphan-R	---GCOWEPWC*	(Jimenez et al., 1996)
Bromocontryphan-R	---GCOWEPWC*	(Jimenez et al., 1997)
Contryphan-Am (Am975)	---GCOWDPWC*	(Sabareesh et al., 2006)
Contryphan-fib	---GCOWMPWC*	(Rajesh, 2015)
Unnamed ^b	--VVGCOQPYC*	(Thakur and Balarum, 2007)

^a Detected by mass spectrometry; isomerism of residue 3 not investigated.

^b Detected in venom of *C. zeylanicus*, *C. betulinus* and *C. figulinus*.

N-terminal extensions to the consensus motif.

One noteworthy feature of contryphans is the presence of a D-amino acid at position 3 of the intercystine loop, which is either Trp or (rarely) Leu (Hansson et al., 2004; Jacobsen et al., 1998, 1999; Jimenez et al., 1997, 1996, 2001; Massilia et al., 2003; Sabareesh et al., 2006). The role of this D-amino acid in contryphans has not been explored in depth. The inclusion of a D-amino acid residue in a natural peptide is uncommon, but not unknown; in 2009, just over 30 examples of D-amino acid-containing peptides in animals were reported, of which nine were contryphans (Bai et al., 2009). In other peptides, D-amino acids have a range of effects. The first D-amino acid-containing peptide to be discovered in vertebrates was dermorphin, which was isolated from the skin of the frog *Phyllomedusa sauvagei* and contains D-Ala in position 2 (Montecucchi et al., 1981). The D-residue was found to be crucial for the potent opiate-like activity, which was lacking in a synthetic all-L analogue (Montecucchi et al., 1981). In the P-type Ca²⁺ channel antagonist ω-agatoxin IVB, found in spider venom, D-Ser46 conferred protection against the proteolytic effect of carboxypeptidase P (Heck et al., 1994). The excitotoxic conopeptide L-RXIA contains D-Phe, and the analogue with L-Phe, L-RXIA[L-Phe44], had a two-fold lower affinity and two-fold faster off rate than L-RXIA on Nav1.6 channels. In addition, the L-analogue was inactive at Nav1.2 channels (Fiedler et al., 2008).

In this study, we investigate a newly-discovered contryphan, contryphan-Vc2, identified in the transcriptome of *Conus victoriae* (Robinson et al., 2014). Mouse bioassays were used to define a novel bioactivity, which is distinct from that seen previously for other contryphans. We assessed the effect of epimerising Trp3 on the activity, solution structure, proteolytic stability and lipid binding of this peptide. Homonuclear 2D ¹H NMR was used to calculate a solution structure, and molecular dynamics (MD) simulations were used to reveal further detail of the behaviour of the peptide in solution. Both NMR and MD indicated an interaction with lipid membranes.

2. Materials and methods

2.1. Chemical synthesis

Contryphan-Vc2 was identified in the venom gland transcriptome of *C. victoriae* as reported previously (Robinson et al.,

2014). Contryphan-Vc2 peptides containing D-Trp, L-Trp or L-Ala in position 3 were prepared by conventional N-(9-fluorenyl)methyl-oxycarbonyl (Fmoc) chemistry on Rink amide resin at 0.1 mmol scale. Briefly, deprotection was performed in 20% piperidine (in DMF), followed by activation and elongation with 70 mL/L DIPEA (in DMF) and 3 equivalents of HCTU with Fmoc-protected amino acid for 50 min. Cleavage from the resin was performed over 2 h with a mixture of trifluoroacetic acid, triisopropylsilane, 1,3-dimethoxybenzene and 3,6-dioxo-1,8-octanedithiol (TFA:-TIPS:DMB:DODT, 92.5:2.5:2.5:2.5 by volume). The cleavage mixture was evaporated and precipitated with ice-cold diethyl ether. The crude product was lyophilised and stored at −20 °C until further purification.

Disulfide formation was achieved by stirring ~0.5 mg/mL crude peptide in 0.1 M ammonium bicarbonate (pH 8.0) for 17 h at room temperature. The cyclised peptides were purified on a Vydac 10 μm C18 (250 × 10 mm) column using a gradient of 40–70% buffer B over 30 min (buffer A: 0.1% TFA in MilliQ water; buffer B: 0.1% TFA in 80% acetonitrile). Sample purity was assessed by LC-MS to be greater than 95%. Further samples of synthetic [D-Trp3]- and [L-Trp3]-contryphan-Vc2 were purchased from Purar Chemicals (Victoria, Australia) and used for the proteolysis and DPC assays.

2.2. Proteolysis assays

Peptide resistance to proteolysis by trypsin, α-chymotrypsin and pepsin was measured by incubating a 250:1 peptide:enzyme mixture at 37 °C for 4 h. BSA was used as a positive control of protease activity. Stock solutions of trypsin and α-chymotrypsin were prepared in 1 mM HCl/2 mM CaCl₂ and reactions were run in 50 mM Tris, 100 mM NaCl (pH 7.4). Pepsin was prepared in 10 mM HCl and reactions run in 10 mM acetic acid/10 mM HCl (pH 2.0). Trypsin and α-chymotrypsin reactions were halted using 2.5% volume of acetic acid solution (25% v/v), and pepsin reactions were halted using 2.5% volume of 200 mM glycine-NaOH buffer (pH 11.4). The extent of digestion was analysed by LC-MS using a Jupiter 5 μm C4 300 Å column (50 × 2.0 mm) (buffer A: 0.1% formic acid in MilliQ water; buffer B: 0.1% formic acid in acetonitrile). Samples were eluted with a gradient of 0–60% B over 10 min.

2.3. Behavioural assay

Swiss Webster mice (15–21 days old; 6.6–10.1 g) were injected intracranially with different doses of synthetic peptides dissolved in 10 μL 0.9% NaCl, as described previously (McIntosh et al., 1994). Control mice were injected with 10 μL 0.9% NaCl solution. Following intracranial injection, mouse behaviour was observed for 2 h to determine differences between treated and control animals. All experiments involving the use of animals were approved by the Institutional Animal Care and Use Committee of the University of Utah.

2.4. Nuclear magnetic resonance spectroscopy

NMR spectra were acquired using a Bruker Avance III 600 MHz instrument. Lyophilised [D-Trp3]-contryphan-Vc2 was dissolved in 93% H₂O/7% ²H₂O at pH 4.0 to a concentration of 2 mM and 300 μL samples were placed in Shigemi tubes. TOCSY and ROESY experiments were recorded at 5 °C using mixing times of 80 and 350 ms, respectively. An additional ROESY experiment was recorded with a mixing time of 50 ms to assist in χ₁ angle determination. 1D ¹H experiments were recorded at temperatures ranging from 10 to 30 °C in 5 °C steps to calculate amide proton temperature coefficients, and at 37 °C to test stability under physiological conditions. Hydrogen-deuterium exchange rates were measured by

dissolving lyophilised [D-Trp3]-contryphan-Vc2 in 100% $^2\text{H}_2\text{O}$ and recording 1D ^1H spectra at 5 min intervals at 10 °C. Spectra were referenced using dioxane at 3.75 ppm. ^1H - ^{13}C HSQC and ^1H - ^{15}N HMQC spectra were collected at natural abundance for assignment of heavy atom chemical shifts, and were referenced indirectly using resonance frequency ratios (Wishart et al., 1995). [L-Trp3]-Contryphan-Vc2 was dissolved in 93% H_2O /7% $^2\text{H}_2\text{O}$ at pH 4.1 to a concentration of 3 mM, and TOCSY and ROESY spectra were collected on a 550 μL sample in a standard 5 mm NMR tube at 5 °C. Titration with DPC micelles was carried out by dissolving lyophilised [D-Trp3]- and [L-Trp3]-contryphan-Vc2 separately in 100 mM deuterated acetate buffer, pH 4.4 to a concentration of 0.5 mM. 1D ^1H NMR spectra were acquired at 5 °C both before and after addition of deuterated dodecylphosphocholine (DPC, Sigma-Aldrich), in concentrations of 1, 5, 10, 20 and 40 mM. Peaks were monitored for shifting and broadening. TopSpin (v3.2, Bruker, USA) was used for spectral processing and analysis.

2.5. Structure calculation

Distance constraints for structure determination were derived from intensities of NOE cross-peaks in the ROESY spectrum of [D-Trp3]-contryphan-Vc2 recorded at 5 °C and pH 4.0. Dihedral angle restraints were obtained by measuring the $^3J_{\text{HNHA}}$ scalar coupling constants at 5 °C, noting the magnitude of the splitting between doublets in amide resonances. Values below 6.0 Hz were taken to indicate α -helical conformation ($-90^\circ < \phi < -40^\circ$), while values above 8.0 Hz indicated β -strand conformation ($-160^\circ < \phi < -80^\circ$). Amide proton temperature coefficients were calculated by recording 1D ^1H spectra over the range 10–30 °C, and plotting the chemical shift values for each residue. Residues with coefficients more positive than -4.5 ppb/K were flagged as potential hydrogen bond donors (Baxter and Williamson, 1997). χ_1 angles were investigated by matching $^3J_{\alpha\beta}$ coupling constants and NOE intensities from a short mixing time (50 ms) NOESY spectrum to the patterns identified by Wagner et al. (1987), but no restraints were generated. The possibility of chemical shift-based restraints was investigated by submitting data to the TALOS-N web server (available at <http://spin.niddk.nih.gov/bax/nmrserver/talosn/>), but no useful constraints were returned. NMR solution structures were calculated using Cyana v3.0 and refined with XPLOR-NIH v2.40 based on distance restraints (from NOE intensities) and ϕ angle restraints (from $^3J_{\text{HNHA}}$ coupling constants).

2.6. Molecular dynamics (MD) simulations

Molecular dynamics simulations were carried out using GRO-MACS version 5.0.4 with the GROMOS 54a7 united-atom forcefield (modified to allow inclusion of D-Trp). Simulations used a 2 fs time step. Temperature coupling made use of the velocity rescale algorithm with a reference temperature of 298 K. Pressure coupling used the Berendsen or Parrinello-Rahman algorithms, with reference pressure of 1 bar and compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$.

Starting models of [D-Trp3]-, [L-Trp3]- and [W3A]-contryphan-Vc2 were built from the NMR-derived solution structures of [D-Trp3]-contryphan-Vc2 with the proline in the *trans* conformation. For membrane simulations, the model system was constructed by placing an appropriate number of lipid molecules in the simulation box and solvating with SPC water. A steepest-descent minimisation of 2000 steps was used to remove bad van der Waals contacts between atoms, followed by a temperature equilibration (without pressure coupling) for 10,000 steps. Semiisotropic pressure coupling was applied using the Berendsen barostat for 500,000 steps. The simulation was then run for 200 ns using the Parrinello-Rahman barostat to allow the lipid molecules to relax. Following

this equilibration procedure, the peptide was introduced to the system either in the aqueous phase or buried in the bilayer, with the disulfide bond in the $+z$ direction. The minimisation and equilibration steps were then repeated, with the peptide coordinates restrained. The simulation production runs for [D-Trp3]-, [L-Trp3]- and [W3A]-contryphan-Vc2 were executed for 2 μs . [D-Trp3]-Contryphan-Vc2 simulations were extended for an additional 1 μs (total 3 μs). A simulation of [D-Trp3]-contryphan-Vc2 in water used isotropic Parrinello-Rahman pressure coupling, and peptide coordinates were not restrained during equilibration; otherwise the same equilibration procedure was used and the production run was executed for 100 ns. Simulation trajectories were visualised with VMD (v1.9.2) and analysed using VMD and GROMACS built-in tools.

3. Results

3.1. Synthesis and purification of contryphan-Vc2

Solid-phase peptide synthesis using Rink amide resin was used to prepare samples of contryphan-Vc2 containing either D-Trp, L-Trp or L-Ala at position 3. Analysis of each of the linear peptides by LC-MS confirmed that the synthetic Trp-containing peptides had masses of 863.6 Da, closely matching the expected molecular mass for reduced peptide ($\text{MH}^+ m/z = 863.0$ Da). Oxidation in ammonium bicarbonate buffer produced a mass loss of 2 Da to 861.6 Da, consistent with formation of the single disulfide bond (theoretical $\text{MH}^+ m/z = 861.0$ Da). The oxidised Ala-containing peptide contained ions of 746.7 Da, matching the expected LC-MS profile (theoretical $\text{MH}^+ m/z = 746.9$ Da). Purification of peptides by RP-HPLC resulted in samples of >96% purity (Fig. S1, Supporting Information).

3.2. D-isomerism changes peptide elution times

Samples of [D-Trp3]- and [L-Trp3]-contryphan-Vc2 (both singly and as a mixture) were injected on a LC-MS system fitted with a Luna 3u C8(2) 100 Å column and eluted with a gradient of 0–60% acetonitrile over 90 min. [D-Trp3]-Contryphan-Vc2 eluted at 37.0 min, one minute before [L-Trp3]-contryphan-Vc2, which eluted at 38.1 min. The sample containing a mixture of both peptides showed two distinct peaks, with elution times consistent with the single injections. (Fig. S2, Supporting Information). This shows that changing the chirality at Trp3 alters the properties of the peptide sufficiently to affect the chromatographic behaviour.

3.3. Trp3 is important for bioactivity, regardless of chirality

Bioactivity of other contryphans has been demonstrated by intracranial injection in mice (Jacobsen et al., 1999; Jimenez et al., 1997, 1996, 2001). In this study, we employed the same assay to screen for broad biological activity of contryphan-Vc2 and its analogues. Control mice injected with 10 μL normal saline solution showed active exploratory behaviour over the 2 h period of observation, with short periods of rest and grooming. In contrast, [D-Trp3]-contryphan-Vc2 administered by intracranial injection in mice at a dose of 10 nmol produced a strong and reproducible phenotype characterised by limited movement, splaying of hind limbs, dragging of hind limbs and flattening of the lower body when moving (Table 2). At the lowest dose tested (1 nmol) the same behaviour was observed but with reduced severity and duration. At the highest dose tested (20 nmol), the same behaviour was again observed and was increased in both severity and duration. [L-Trp3]-contryphan-Vc2 also produced a strong phenotype which, within the resolution of the assay, was indistinguishable from that elicited

Table 2

Summary of results from intracranial mouse injections. Mice were between 15 and 21 days old and weighed between 6.6 and 10.1 g.

Dose (nmol)	Observed Behaviour (time = approximate post injection)
[D-Trp3]-contryphan-Vc2	
20 (n = 1)	0 min – 1 h: Both hind and front legs splayed and body flat. 1 h–2 h: Limited movement. Hind legs splayed, dragging hind legs and lower body flat when moving. Mouse had not recovered by the end of observation at 2 h.
10 (n = 5)	From 0 min: Limited movement. Hind legs splayed when still, hind legs dragging and lower body flat when moving, intermittent episodes of body tremor. Recovery was observed at 15 min and 45 min for three of the mice, while the remaining two mice had not recovered by the end of observation at 2 h.
1 (n = 2)	5–20 min: Moving around cage, but hind legs dragging when moving. 20 min–2 h: no difference from control.
[L-Trp3]-contryphan-Vc2	
20 (n = 1)	0–10 min: Limited movement. Both hind and front legs splayed and body flat. 10–20 min: Limited movement. Recovered by 20 min.
10 (n = 3)	0–15 min: Limited movement. Hind legs splayed and body flat. 15 min–1 h: Increased movement with intermittent episodes of body tremor. The peptide produced a stronger response in one mouse with hind legs splayed, dragging hind legs and lower body flat when moving for the entire observation period of 2 h (similar to what was observed for the 20 nmol dose of [D-Trp3]-contryphan-Vc2).
[W3A]-contryphan-Vc2	
10 (n = 3)	Active exploratory behaviour with short periods of rest and grooming.
Control	
(n = 5)	Active exploratory behaviour with short periods of rest and grooming.

by [D-Trp3]-contryphan-Vc2 (Table 2). However, the behaviour of mice injected with [W3A]-contryphan-Vc2 at a dose of 10 nmol could not be distinguished from that of control mice injected with normal saline solution. Together, these data demonstrate that contryphan-Vc2 is bioactive, and that the Trp sidechain in position 3 is critical for this bioactivity. Somewhat surprisingly, however, the chirality of Trp3 has no differential effect.

3.4. NMR spectroscopy

NMR spectra of [D-Trp3]-contryphan-Vc2 showed two sets of resonances, a major and minor species. The resonances of the major species showed good chemical shift dispersion, implying that the peptide adopts a well-defined conformation in solution (Fig. 1A).

Chemical shift assignments were made for ^1H and heavy atoms in the major conformer and ^1H in the minor conformer (Tables S1, S2 and S3, respectively, of the Supporting Information). Chemical shift data for the major conformer have been deposited in the BMRB (Ulrich et al., 2008), ID: 30152. The conformers differed in the geometry of the Thr4-Pro5 peptide bond, which was in the *trans* orientation in the major form, as shown by the strong NOE cross-peak between the Thr4 H α and Pro5 H δ resonances (Fig. S3, Supporting Information). Further corroborating this finding, the ^{13}C chemical shift difference between the C β and C γ resonances of Pro5, 4.61 ppm, conformed to the published value typical for the *trans* conformation (4.51 ppm) (Schubert et al., 2002). In the minor conformer, the Thr4-Pro5 bond was in the *cis* orientation, as shown by the NOE cross-peak between the H α resonances of Thr4 and Pro5 (Fig. S3, Supporting Information). The signals in the ^{13}C spectrum were too weak to determine the $\Delta\beta\gamma$ value, which for *cis* bonds is typically 9.64 ppm (Schubert et al., 2002). The *cis:trans* ratio was determined to be 1:5.

3.5. Structure determination

An ensemble of 20 structures was calculated for the major conformer of [D-Trp3]-contryphan-Vc2. The experimental constraints used for structure calculation and the structural statistics for the ensemble generated are summarised in Table 3.

Dihedral angle constraints were obtained for Arg2 and Val6, both of which possess $^3J_{\text{HNHA}}$ above the 8.0 Hz threshold for β -

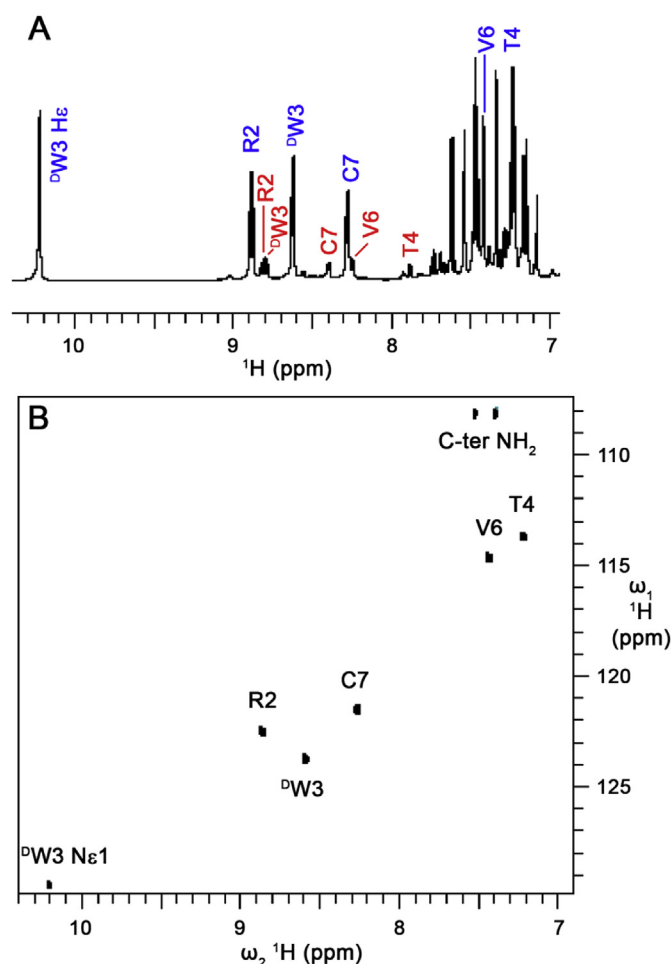


Fig. 1. 600 MHz NMR spectra of [D-Trp3]-contryphan-Vc2. (A) 1D ^1H spectrum, collected at 5 °C on a 2 mM sample at pH 4.0. Peaks are labelled from the major (blue) and minor (red) conformers. (B) 2D heteronuclear ^1H - ^{15}N -HMQC spectrum, collected at 10 °C on a 500 μM sample at pH 4.1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Structural statistics for 20 lowest-energy structures of contryphan-Vc2.

Distance restraints	
Intra-residue ($ i-j = 0$)	30
Sequential ($ i-j = 1$)	33
Medium-range ($2 \leq i-j \leq 4$)	10
Long-range ($ i-j > 4$)	0
Total	73
Dihedral restraints	
Backbone (ϕ angle)	2
Sidechain (χ_1 angle)	0
RMSD over 20 structures (all residues)	
Backbone (Å) (N, C α , C, O)	0.54 ± 0.29
All heavy atoms (Å)	0.80 ± 0.29
RMSD over 20 structures (residues 2–6)	
Backbone (Å) (N, C α , C, O)	0.11 ± 0.06
All heavy atoms (Å)	0.39 ± 0.20
Ramachandran analysis	
Residues in most favoured regions (%)	13.8
Residues in additionally allowed regions (%)	70.0
Residues in generously allowed regions (%)	16.2
Residues in disallowed regions (%)	0.0
Energies (XPLOR energy units)	
E_{NOE}	2.8 ± 0.1
$E_{\text{bond}} + E_{\text{angle}} + E_{\text{improper}}$	10.2 ± 0.1
RMSDs from idealised geometry	
Bonds (Å)	0.0026 ± 0.0004
Angles ($^\circ$)	0.515 ± 0.006
Impropers ($^\circ$)	0.25 ± 0.01

strand conformations ($-160^\circ < \phi < -80^\circ$). Analysis of amide temperature coefficients (Fig. S4, Supporting Information) suggested that the amide proton of Thr4 may be the donor of a hydrogen bond ($\Delta\delta/\Delta T = -2.0$ ppb/K, above the threshold of -4.5 ppb/K) (Baxter and Williamson, 1997), although this value could also be due to ring current effects from the adjacent aromatic D-Trp3 (Cierpicki and Otlewski, 2001). This amide (like all other amides) exchanged rapidly with deuterium when the peptide was dissolved in $^2\text{H}_2\text{O}$, making a hydrogen bond unlikely. The calculated structures showed no consistent acceptor for this hypothetical bond and so it was not used as a constraint. Fig. 2A presents a stereo view of an overlay of the ensemble of final structures calculated for the major conformer of [D-Trp3]-contryphan-Vc2, while Fig. 2B shows the closest-to-average structure. The structural ensemble has been deposited with the PDB (Berman et al., 2003), ID: 5L34.

The NMR-derived structure of [D-Trp3]-contryphan-Vc2 in water reveals that the peptide adopts a well-defined structure consisting of two turn-like regions (formed by residues 1–4 and residues 4–7). The C α atoms of Cys1 and Thr4 are separated by less than 7 Å, as are the C α atoms of Thr4 and Cys7. However, the dihedral angles of the internal residues do not match the definitions of the formal types of β -turn (Richardson, 1981; Wilmot and Thornton, 1988). Both turns share Thr4 as a common residue, and are joined by a disulfide bond at the termini. The disulfide bond is the least constrained region of the structure, and makes hydrophobic contact with Thr4. The RMSD values calculated across the structure ensemble show that all structures are similar, especially when the terminal residues are excluded. Ramachandran values are favourable, with most residues falling in the additionally allowed regions of the Ramachandran plot and none present in the disallowed regions (Table 3). This structure is maintained at

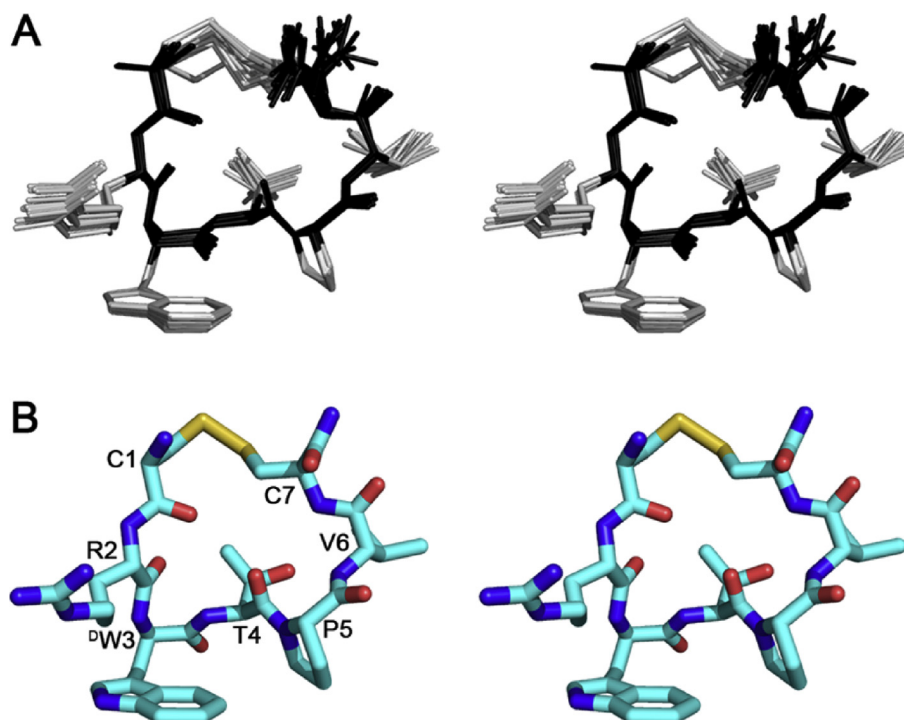


Fig. 2. Stereo images of the calculated solution structure of the major [D-Trp3]-contryphan-Vc2 conformer. (A) Ensemble of 20 structures of [D-Trp3]-contryphan-Vc2 calculated from distance and dihedral angle restraints. Structures have been superimposed over the backbone atoms of residues 2–6. Backbone is coloured black, residue side chains are grey. (B) Closest-to-average structure from the ensemble. The disulfide bond is yellow, nitrogen atoms are blue, oxygen atoms are red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

physiological temperature (Fig. S5, Supporting Information).

In order to probe the dynamics of [D-Trp3]-contryphan-Vc2 in solution, MD simulations were undertaken. These simulations revealed dynamic behaviour in the conformation associated with rotation of Trp3 around the χ_1 bond (Fig. 3; see also Fig. S6, Supporting Information). Fig. 3A plots the orientation of the Trp3 sidechain throughout the length of the simulation trajectory, revealing the adoption of two primary positions. Fig. 3B illustrates the major conformer (Trp3 $\chi_1 = +60^\circ$), in which the Trp3 and Arg2 sidechains are in close proximity to each other, projecting from a turn-like structure. In the minor conformer (Trp3 $\chi_1 = -60^\circ$, Fig. 3C), the sidechains of Trp3 and Pro5 are in close proximity. Neither of these conformers matches the sidechain positions of the calculated NMR structure, although both maintain the Thr4-Pro5 bond in the *trans* conformation and are similar in terms of the backbone conformation, with the turns intact and the disulfide bond still making hydrophobic contact with Thr4.

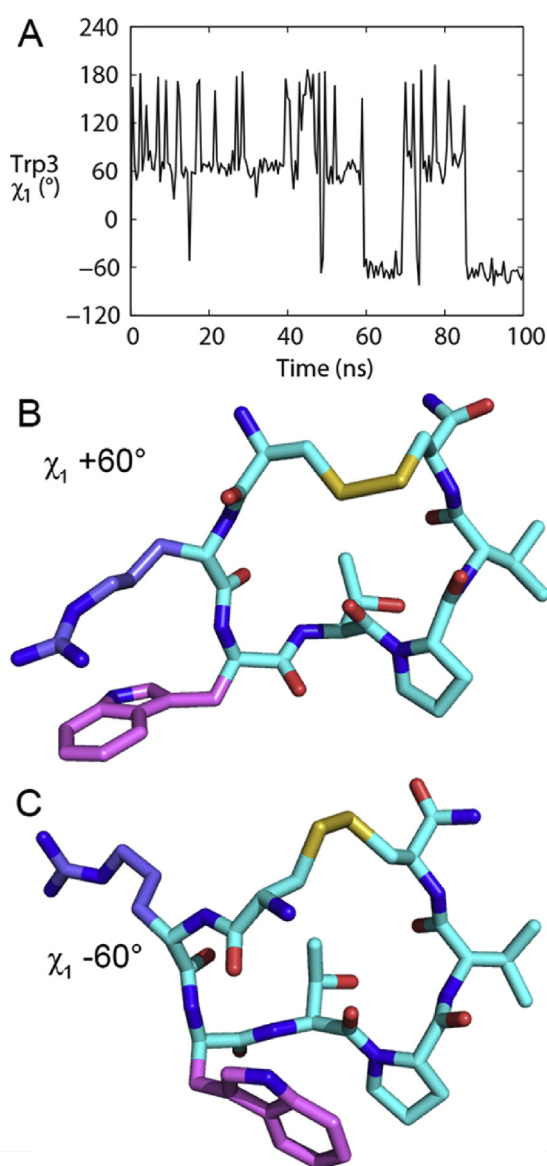


Fig. 3. Conformational data arising from MD simulations of [D-Trp3]-contryphan-Vc2 in water. (A) Plot of Trp3 χ_1 angle vs time, showing presence of two distinct states. (B) Representative structure of high occupancy state (Trp3 $\chi_1 = +60^\circ$), in which Arg2 and Trp3 sidechains are in close proximity. (C) Representative structure of low occupancy state (Trp3 $\chi_1 = -60^\circ$), in which Trp3 and Pro5 sidechains are in close proximity.

3.6. Chemical shift comparison between [D-Trp3]- and [L-Trp3]-containing peptides suggests structural similarity

The spectra of [L-Trp3]-contryphan-Vc2 were well-dispersed, and chemical shift assignments were made for all ^1H resonances of the major species (Table S4 of the Supporting Information). Comparison of the ^1H resonance chemical shifts of [D-Trp3]-contryphan-Vc2 and [L-Trp3]-contryphan-Vc2 revealed minor differences. The amide proton of Trp3 is shifted upfield by over 1 ppm in [L-Trp3]-contryphan-Vc2, with a smaller downfield shift for the amide proton of Thr4 of around 0.5 ppm. There is a comparable 0.5 ppm downfield shift for the H α resonance of Pro5, and minor shifts in both directions for other resonances, both H α and HN (Fig. S7, Supporting Information). These results suggest that the structure of [L-Trp3]-contryphan-Vc2, while slightly different, is largely intact and similar to that of [D-Trp3]-contryphan-Vc2.

3.7. D-isomerism protects against proteolytic degradation

One possible reason for the presence of D-amino acids in peptides is that the atypical chirality protects the molecule from proteolytic degradation. To test this notion in the context of contryphan-Vc2, the digestive enzymes trypsin, α -chymotrypsin and pepsin were used in proteolytic assays. Trypsin degraded [L-Trp3]-contryphan-Vc2 within 4 h, while having no effect on [D-Trp3]-contryphan-Vc2. Neither [D-Trp3]- nor [L-Trp3]-contryphan-Vc2 was susceptible to cleavage by pepsin or α -chymotrypsin over the 4 h duration of the assay. LC-MS traces are presented in Fig. S8 of the Supporting Information. These results suggest that the D-chirality of Trp3 is protective against proteolytic degradation.

3.8. Contryphan-Vc2 interacts with a model lipid system

The insensitivity of the bioactivity assay to the chirality of Trp3 suggested that the peptide may not interact with a protein receptor in a traditional manner. Titration of [D-Trp3]-contryphan-Vc2 with DPC micelles showed clear signs of an interaction between peptide and lipid, with amide peaks broadening and shifting as the concentration of lipid was increased. A similar effect was observed for titration of [L-Trp3]-contryphan-Vc2 into DPC micelles. Plots showing changes in chemical shift with DPC concentration are presented in Fig. 4 (Panels A–C), and spectra of [D-Trp3]- and [L-Trp3]-contryphan-Vc2 at 0 and 40 mM DPC are shown in Fig. S9 of the Supporting Information. These results demonstrate that both [D-Trp3]-contryphan-Vc2 and [L-Trp3]-contryphan-Vc2 interact with this model lipid system. When tested at a peptide concentration of 0.5 mM these effects stabilised upon addition of 20–40 mM DPC and no further alterations to the spectra were apparent in samples on addition of 100 mM DPC (data not shown). Hence saturation was reached around a ratio of 80:1 lipid:peptide. MD simulations were run to gain insight into specific molecular interactions between peptide and lipid bilayers, as described in the Supporting Information, and a representative structure of [D-Trp3]-contryphan-Vc2 in a lipid bilayer is shown in Fig. 4D.

4. Discussion

Contryphan-Vc2 differs from the consensus contryphan sequence (Sonti et al., 2013) at positions 2 and 6 (replacing Pro/Hyp2 with Arg and Trp6 with Val). The behaviour in RP-HPLC is also different from that typically seen for contryphan in containing a single sharp peak and lacking the later-eluting peak observed in elution profiles of other contryphanes (Jacobsen et al., 1999). The sequence changes in contryphan-Vc2 are evidently sufficient to abolish the appearance of a second species on the chromatographic

timescale. However, the NMR spectra of [D-Trp3]-contryphan-Vc2 clearly show two sets of peaks, indicating that this peptide does exist in two distinct conformations in solution, as has been observed previously for other members of the contryphan family (Eliseo et al., 2004; Pallaghy et al., 1999, 2000). The difference between the two conformations of contryphan-Vc2 was found to be *cis/trans* isomerism about the Thr4-Pro5 bond, with a *cis:trans* ratio of 1:5. All but one of the contryphans characterised to date display *cis/trans* isomerism (the exception being glacontryphan-M) (Grant et al., 2004). Of the two Pro residues in the consensus sequence, isomerism is usually observed in the conformation of the Cys-Pro bond at the start of the intercystine loop (where Pro can also be modified to Hyp), but in contryphan-Vc2, Pro2 is replaced by Arg and *cis/trans* isomerism occurs at the second Pro residue. This is unusual, as in all but one of the previously published solution structures of contryphans this second Xaa-Pro bond has been exclusively in the *trans* conformation (Eliseo et al., 2004; Grant et al., 2004; Pallaghy et al., 1999, 2000). The sole exception is contryphan-In (In936), which a recent study identified as also possessing isomerism at the second Xaa-Pro bond and which, like contryphan-Vc2, lacks the first Pro residue (containing Val instead) (Sonti et al., 2013) (Table 1). The *cis* bond in contryphan-In may be promoted by the Tyr residue preceding the second Pro, as it has been noted that preceding aromatic residues promote the formation of *cis* peptide bonds to Pro through an interaction between the partially positive Pro ring face and the negatively charged aromatic π -face (Zondlo, 2012). The residue preceding Pro5 in contryphan-Vc2 is Thr4, so this type of interaction is not present in this peptide.

Unlike the structures of contryphans solved previously, there is no *i/i+4* electrostatic interaction between the N-terminal ammonium group and an acidic sidechain, an interaction that has been hypothesised to stabilise *cis* Xaa-Pro bonds in contryphans (Pallaghy et al., 2000). Despite these differences, the orientation of the backbone in contryphan-Vc2 is broadly similar to other members of the contryphan family, as can be seen in an overlay of selected structures with the calculated NMR structure of [D-Trp3]-contryphan-Vc2 (Fig. 5).

Based on the NMR data, a model of the solution structure for [D-Trp3]-contryphan-Vc2 was calculated that satisfied the observed NOE and dihedral angle restraints. However, MD simulations of the peptide in aqueous solvent cast doubt on the adequacy of this single NMR model as a representation of the solution conformation of [D-Trp3]-contryphan-Vc2. The NMR-derived structure positions the sidechains of Arg2, Trp3 and Pro5 in close proximity to each other, satisfying the restraints generated by the observed NOE cross-peaks between Trp3 and the other two residues. However, the MD data suggest that the peptide actually adopts two distinct conformations, distinguished by rotation about the Trp3 χ_1 bond. In one of these, the sidechain of Trp3 is in close proximity to that of Arg2, while in the other it is proximal to Pro5 (as illustrated in Fig. 3). If the peptide does adopt these two distinct conformations in solution, then the assumption made in calculating the solution structure, that a single conformation was present, is compromised. The atomic positions implied by NOE results are assumed to be at full occupancy, with weaker cross-peaks interpreted as arising from an interaction between more distant atoms, rather than being a

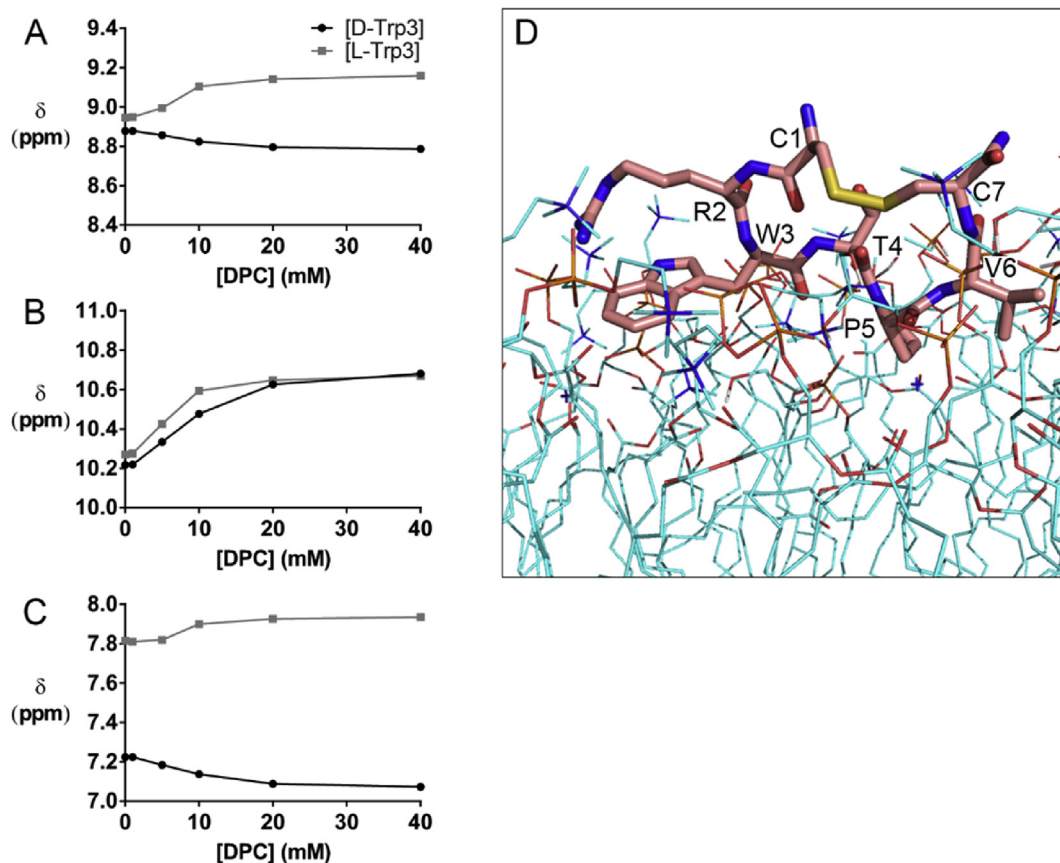


Fig. 4. Interactions of contryphan-Vc2 with lipid. (A) Chemical shift of Arg2 HN resonance in presence of increasing concentrations of DPC. (B) DPC-induced chemical shift changes for Trp3 Hε1 resonance. (C) DPC-induced chemical shift changes for Thr4 HN resonance. (D) Representative frame of [D-Trp3]-contryphan-Vc2 from MD simulation (see also Supporting Information, Figs. S10 and S11), showing position and orientation of the peptide in the polar head group region of the lipid bilayer.

consequence of the peptide sampling multiple conformations. In the latter case, the calculated NMR structure would be the result of averaging the interactions observed in all sampled conformations, and would not necessarily correspond to a pose that the peptide ever actually adopts. Such appears to be the case for contryphan-Vc2, as the MD simulations never show the sidechains of Arg2 and Pro5 in close proximity as in the structures calculated from the NMR data. This is an intriguing finding, as the NMR spectra of this peptide in aqueous solution are sharp and well-dispersed, properties that would typically be interpreted as indicating a single major conformation, or ensemble of very similar conformations. Spectra at substantially lower temperatures or in the solid state could potentially resolve the resonances from these two major conformations detected by the MD simulations.

An effective method for initially assessing the bioactivity of peptide toxins has been administration by intracranial injection in mice. Members of the contryphan family that have been assayed in this way have been reported to induce a 'stiff-tail' syndrome at low doses, while higher doses led to hyperactive behaviours (such as circular motion and barrel rolling), and eventually seizures and death (Jacobsen et al., 1998, 1999; Jimenez et al., 1997, 1996, 2001). By contrast, contryphan-Vc2 appears to induce a depressed phenotype with minimal movement; the mice adopt a splayed posture and drag their hind limbs when moving. This may indicate that the sequence differences in contryphan-Vc2 are sufficient to direct it to a different biological target, as has been demonstrated for several other toxin folds (Froy et al., 1999; Norton and Pallaghy, 1998).

It is intriguing that both [D-Trp3]-contryphan-Vc2 and [L-Trp3]-contryphan-Vc2 elicit indistinguishable biological effects. The lack

of activity observed with [W3A]-contryphan-Vc2 suggests that the side chain of Trp3 is directly involved in bioactivity, and presumably receptor binding. However, it would be expected that a change in chirality of this crucial sidechain would disrupt the interaction between peptide and receptor, leading to a loss of or reduction in activity (Dawson et al., 1999; Kreil et al., 1989).

While the similar activities of both D- and L-isomers may point to a mechanism of action mediated by a receptor other than a protein, there are examples of protein receptors that possess a more permissive binding site, for example CXCR4, a chemokine receptor, that was found to interact with all-L, all-D and mixed L/D analogues of a peptide derived from the N-terminal region of viral macrophage inflammatory protein II (vMIP-II) (Zhou et al., 2002).

An alternative to direct receptor binding could be that contryphan-Vc2 interacts with membranes, and then influences a biological target in a membrane-mediated manner. This phenomenon has been reported before; for example, in the interaction between the spider toxin GsMTx4 and gramicidin A channels (Suchyna et al., 2004). The gramicidin A channel is formed when two monomers in opposing leaflets of a lipid bilayer associate together, an event dependent on thinning of the bilayer in a localised area. GsMTx4 alters the lipid packing adjacent to the channel and induces this thinning, and has been observed to enhance the appearance of dimeric channels by a factor of 10–25-fold and extend the lifetime of an open channel by a factor of two. An all-D analogue of GsMTx4 was tested in the same study and exhibited identical activity to the native toxin, mirroring the situation with contryphan-Vc2. It is therefore conceivable that the effects of contryphan-Vc2 are due to a channel which is similarly sensitive to the state of the bilayer.

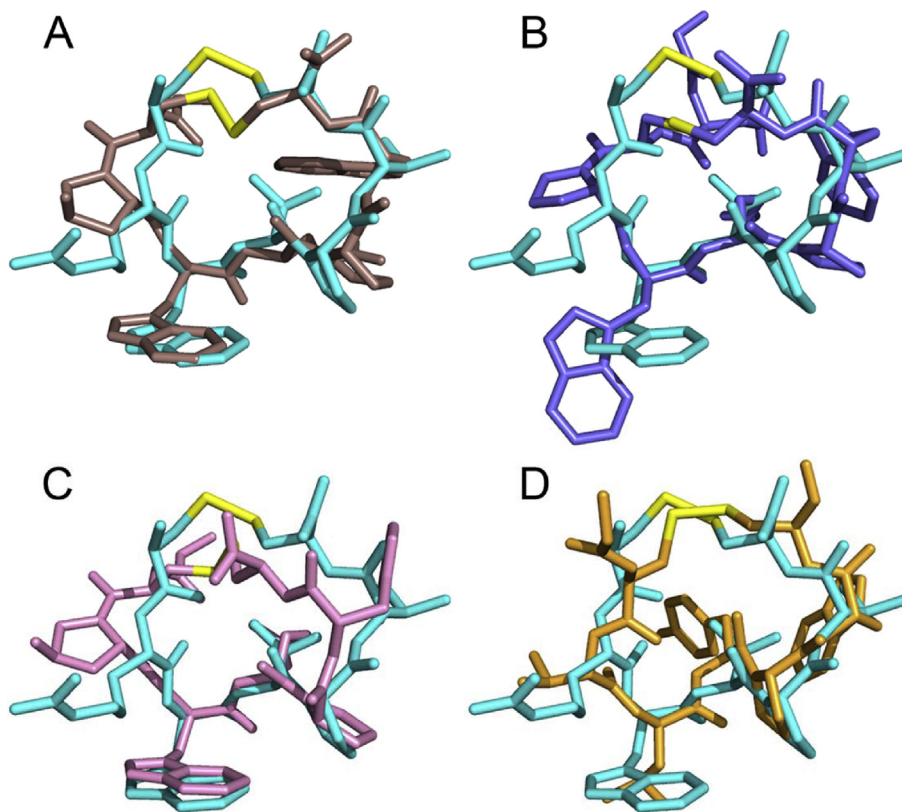


Fig. 5. Superimposed structures of [D-Trp3]-contryphan-Vc2 (cyan) with other contryphan family members. (A) Contryphan-R, PDB 1QFB (Pallaghy et al., 1999). (B) Contryphan-Vn, PDB 1NXN (Eliseo et al., 2004). (C) Contryphan-Sm, PDB 1DFY (Pallaghy et al., 2000). (D) Contryphan-In, PDB 2M6D (Sonti et al., 2013). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Titration of [D-Trp3]- or [L-Trp3]-contryphan-Vc2 with DPC micelles led to peak shifting and broadening in NMR spectra, so to gain insight into possible interactions with lipid bilayers, a series of MD simulations was performed with [D-Trp3]-, [L-Trp3]- and [W3A]-contryphan-Vc2 in the presence of mixed POPC/POPG bilayers. Two simulations were run for each peptide, one with the peptide starting in the aqueous phase and the other with the peptide already embedded in the membrane. In all simulations, the peptide migrated from its starting point to the headgroup region of the bilayer (Fig. S10, Supporting Information), although the final equilibrium position of each peptide was influenced by the starting point and did not converge for any pair of simulations. This suggests that there is a significant barrier to transfer of these peptides from aqueous solution to complete burial within the lipid phases. There were no apparent differences between the simulations of [D-Trp3]- and [L-Trp3]-contryphan-Vc2; both migrate to a region near the head group of the lipid and remain in that position for the duration of the simulation.

5. Conclusions

Contryphan-Vc2 is distinct from the rest of the contryphan peptide family in having a novel bioactivity in mice. The chirality of Trp3 did not affect the activity of the peptide in mice, with both D- and L-isomers eliciting the atypical depressive phenotype, although D-Trp did confer protection against proteolytic degradation. The solution structure determined on the basis of NMR data was well defined, but MD simulations cast doubt on whether the calculated structure is an adequate representation of the ensemble of conformations accessible to this peptide in aqueous solution. Although the backbone is well defined, MD suggested much greater flexibility around the Trp3 C α -C β bond.

NMR studies show that contryphan-Vc2 interacts with the membrane, a finding supported by MD simulations. Whether membrane interactions are crucial to the bioactivity or if the presumed receptor has a permissive binding site capable of accepting either D- or L-Trp3 remains an open question. A role for lipid in contryphan action may also account for the difficulty to date in identifying a molecular target for this class of peptides. If the molecular entities responsible for the activity of contryphan-Vc2 could be identified, it would be interesting to determine the precise structure-activity relationship of the interaction, especially with respect to the isomerism of Trp3 and the apparent inability of other contryphans to elicit a similar effect in mice.

Funding

R.S.N. acknowledges fellowship support from the NHMRC.

Notes

The authors declare no competing financial interest.

Acknowledgements/Author contributions

AB and AJR synthesised the initial sample, further samples were synthesised by SBD. SDR and SSE performed the behavioural assays. SBD collected NMR data and assigned the NMR spectra. SBD, SC and CAM calculated the solution structure. BC performed proteolysis assays. DKC advised on MD simulations. SBD, SDR and RSN wrote the manuscript, with input from all authors. RSN conceived of the study. RSN acknowledges fellowship support from the Australian National Health and Medical Research Council. We thank the Multimodal Australian Sciences Imaging and Visualisation Environment (MASSIVE) and Victorian Life Sciences Computation Initiative

(VLSI) for provision of CPU time and technical support.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.toxicon.2017.02.012>.

Transparency document

Transparency document related to this article can be found online at <http://dx.doi.org/10.1016/j.toxicon.2017.02.012>.

References

- Bai, L., Sheeley, S., Sweedler, J.V., 2009. Analysis of endogenous D-amino acid-containing peptides in Metazoa. *Bioanal. Rev.* 1, 7–24.
- Baxter, N.J., Williamson, M.P., 1997. Temperature dependence of ^1H chemical shifts in proteins. *J. Biomol. NMR* 9, 359–369.
- Berman, H., Henrick, K., Nakamura, H., 2003. Announcing the worldwide protein data bank. *Nat. Struct. Mol. Biol.* 10, 980–980.
- Cierpicki, T., Otlewski, J., 2001. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* 21, 249–261.
- Dawson, D.W., Volpert, O.V., Pearce, S.F.A., Schneider, A.J., Silverstein, R.L., Henkin, J., Bouck, N.P., 1999. Three distinct D-amino acid substitutions confer potent antiangiogenic activity on an inactive peptide derived from a thrombospondin-1 type 1 repeat. *Mol. Pharmacol.* 55, 332–338.
- Eliseo, T., Cicero, D.O., Romeo, C., Schinina, M.E., Massilia, G.R., Politicelli, F., Ascenzi, P., Paci, M., 2004. Solution structure of the cyclic peptide contryphan-Vn, a Ca^{2+} -dependent K^+ channel modulator. *Biopolymers* 74, 189–198.
- Fiedler, B., Zhang, M.-M., Buczek, O., Azam, L., Bulaj, G., Norton, R.S., Olivera, B.M., Yoshikami, D., 2008. Specificity, affinity and efficacy of iota-conotoxin RXIA, an agonist of voltage-gated sodium channels Nav1.2, 1.6 and 1.7. *Biochem. Pharmacol.* 75, 2334–2344.
- Froy, O., Sagiv, T., Poreh, M., Urbach, D., Zilberberg, N., Gurevitz, M., 1999. Dynamic diversification from a putative common ancestor of scorpion toxins affecting sodium, potassium, and chloride channels. *J. Mol. Evol.* 48, 187–196.
- Grant, M.A., Hansson, K., Furie, B.C., Furie, B., Stenflo, J., Rigby, A.C., 2004. The metal-free and calcium-bound structures of a γ -carboxyglutamic acid-containing contryphan from *Conus marmoreus*, glacontryphan-M. *J. Biol. Chem.* 279, 32464–32473.
- Hansson, K., Ma, X., Eliasson, L., Czerwicz, E., Furie, B., Furie, B.C., Rorsman, P., Stenflo, J., 2004. The first γ -carboxyglutamic acid-containing contryphan: a selective L-type calcium ion channel blocker isolated from the venom of *Conus marmoreus*. *J. Biol. Chem.* 279, 32453–32463.
- Heck, S.D., Kelbaugh, P.R., Kelly, M.E., Thadeio, P.F., Saccomano, N.A., Stroth, J.G., Volkmann, R.A., 1994. Disulfide bond assignment of ω -agatoxins IVB and IVC: discovery of a D-Serine residue in ω -agatoxin IVB. *J. Am. Chem. Soc.* 116, 10426–10436.
- Jacobsen, R., Jimenez, E.C., Grilley, M., Watkins, M., Hillyard, D., Cruz, L.J., Olivera, B.M., 1998. The contryphans, a D-tryptophan-containing family of *Conus* peptides: interconversion between conformers. *J. Pept. Res.* 51, 173–179.
- Jacobsen, R.B., Jimenez, E.C., De la Cruz, R.G.C., Gray, W.R., Cruz, L.J., Olivera, B.M., 1999. A novel D-leucine-containing *Conus* peptide: diverse conformational dynamics in the contryphan family. *J. Pept. Res.* 54, 93–99.
- Jimenez, E.C., Craig, A.G., Watkins, M., Hillyard, D.R., Gray, W.R., Gulyas, J., Rivier, J.E., Cruz, L.J., Olivera, B.M., 1997. Bromocontryphan: Post-translational bromination of tryptophan. *Biochemistry* 36, 989–994.
- Jimenez, E.C., Olivera, B.M., Gray, W.R., Cruz, L.J., 1996. Contryphan is a D-tryptophan-containing *Conus* peptide. *J. Biol. Chem.* 271, 28002–28005.
- Jimenez, E.C., Watkins, M., Juszczak, L.J., Cruz, L.J., Olivera, B.M., 2001. Contryphans from *Conus textile* venom ducts. *Toxicon* 39, 803–808.
- Kreil, G., Barra, D., Simmaco, M., Erspamer, V., Falconieri Erspamer, G., Negri, L., Severini, C., Corsi, R., Melchiorri, P., 1989. Deltorphan, a novel amphibian skin peptide with high selectivity and affinity for δ opioid receptors. *Eur. J. Pharmacol.* 162, 123–128.
- Massilia, G.R., Eliseo, T., Grolleau, F., Lapiet, B., Barbier, J., Bournaud, R., Molgo, J., Cicero, D.O., Paci, M., Schinina, M.E., Ascenzi, P., Politicelli, F., 2003. Contryphan-Vn: a modulator of Ca^{2+} -dependent K^+ channels. *Biochem. Biophys. Res. Commun.* 303, 238–246.
- Massilia, G.R., Schinina, M.E., Ascenzi, P., Politicelli, F., 2001. Contryphan-Vn: a novel peptide from the venom of the Mediterranean snail *Conus ventricosus*. *Biochem. Biophys. Res. Commun.* 288, 908–913.
- McIntosh, J.M., Yoshikami, D., Mahe, E., Nielsen, D.B., Rivier, J.E., Gray, W.R., Olivera, B.M., 1994. A nicotinic acetylcholine receptor ligand of unique specificity, α -conotoxin Iml. *J. Biol. Chem.* 269, 16733–16739.
- Montecuccchi, P.C., De Castiglione, R., Piani, S., Gozzini, L., Erspamer, V., 1981. Amino acid composition and sequence of dermorphin, a novel opiate-like peptide from the skin of *Phyllomedusa sauvagii*. *Int. J. Pept. Protein Res.* 17, 275–283.
- Norton, R.S., Pallaghy, P.K., 1998. The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* 36, 1573–1583.

- Pallaghy, P.K., He, W.L., Jimenez, E.C., Olivera, B.M., Norton, R.S., 2000. Structures of the contryphan family of cyclic peptides. Role of electrostatic interactions in *cis-trans* isomerism. *Biochemistry* 39, 12845–12852.
- Pallaghy, P.K., Melnikova, A.P., Jimenez, E.C., Olivera, B.M., Norton, R.S., 1999. Solution structure of contryphan-R, a naturally occurring disulfide-bridged octapeptide containing D-tryptophan: comparison with protein loops. *Biochemistry* 38, 11553–11559.
- Rajesh, R.P., 2015. Novel M-Superfamily and T-Superfamily conotoxins and contryphans from the vermivorous snail *Conus figulinus*. *J. Pept. Sci.* 21, 29–39.
- Richardson, J.S., 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34, 167–339.
- Robinson, S.D., Safavi-Hemami, H., McIntosh, L.D., Purcell, A.W., Norton, R.S., Papenfuss, A.T., 2014. Diversity of conotoxin gene superfamilies in the venomous snail, *Conus victoriae*. *PLoS One* 9, e87648.
- Sabareesh, V., Gowd, K.H., Ramasamy, P., Sudarshani, S., Krishnan, K.S., Sikdar, S.K., Balaram, P., 2006. Characterization of contryphans from *Conus lorioisii* and *Conus amadis* that target calcium channels. *Peptides* 27, 2647–2654.
- Schubert, M., Labudde, D., Oschkinat, H., Schmieder, P., 2002. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *J. Biomol. NMR* 24, 149–154.
- Sonti, R., Gowd, K.H., Rao, K.N.S., Ragothama, S., Rodriguez, A., Perez, J.J., Balaram, P., 2013. Conformational diversity in contryphans from *Conus* venom: *cis-trans* isomerisation and aromatic/proline interactions in the 23-membered ring of a 7-residue peptide disulfide loop. *Chem. Eur. J.* 19, 15175–15189.
- Suchyna, T.M., Tape, S.E., Koeppe, R.E., Andersen, O.S., Sachs, F., Gottlieb, P.A., 2004. Bilayer-dependent inhibition of mechanosensitive channels by neuroactive peptide enantiomers. *Nature* 430, 235–240.
- Thakur, S.S., Balaram, P., 2007. Rapid mass spectral identification of contryphans. Detection of characteristic peptide ions by fragmentation of intact disulfide-bonded peptides in crude venom. *Rapid Commun. Mass Spectrom.* 21, 3420–3426.
- Ulrich, E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannidis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C.F., Tolmie, D.E., Kent Wenger, R., Yao, H., Markley, J.L., 2008. BioMagResBank. *Nucleic Acids Res.* 36, D402–D408.
- Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Gö, N., Wüthrich, K., 1987. Protein structures in solution by nuclear magnetic resonance and distance geometry: the polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* 196, 611–639.
- Wilmot, C.M., Thornton, J.M., 1988. Analysis and prediction of the different types of β -turn in proteins. *J. Mol. Biol.* 203, 221–232.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., Sykes, B.D., 1995. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* 6, 135–140.
- Zhou, N., Luo, Z., Luo, J., Fan, X., Cayabyab, M., Hiraoka, M., Liu, D., Han, X., Pesavento, J., Dong, C.-Z., Wang, Y., An, J., Kaji, H., Sodroski, J.G., Huang, Z., 2002. Exploring the stereochemistry of CXCR4-peptide recognition and inhibiting HIV-1 entry with D-peptides derived from chemokines. *J. Biol. Chem.* 277, 17476–17485.
- Zondlo, N.J., 2012. Aromatic–proline interactions: electronically tunable CH/π interactions. *Acc. Chem. Res.* 46, 1039–1049.

Structure and activity of contryphan-Vc2: Importance of the D-amino acid residue

Stephen B. Drane^a, Samuel D. Robinson^a, Christopher A. MacRaild^a, Sandeep Chhabra^a, Balasubramanyam Chittoor^a, Rodrigo A. V. Morales^a, Eleanor W. W. Leung^a, Alessia Belgi^b, Samuel S. Espino^c, Baldomero M. Olivera^c, Andrea J. Robinson^b, David K. Chalmers^a, Raymond S. Norton^{a*}

^aMedicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville 3052, Victoria, Australia

^bSchool of Chemistry, Monash University, Clayton 3800, Victoria, Australia

^cDepartment of Biology, University of Utah, Salt Lake City, Utah 84112, USA

* Corresponding author: Raymond S. Norton, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, VIC 3052, Australia, Tel: (+61 3) 9903 9167; E-mail: Ray.Norton@monash.edu

Table S1. ^1H resonance assignments for [D-Trp3]-contryphan-Vc2 major conformer (*trans* Thr4-Pro5 peptide bond).

Residue	Proton chemical shift (ppm)		
	H^{N}	H^{α}	H (other)
Cys1	-	4.33	$\text{H}\beta 1$ 3.22; $\text{H}\beta 2$ 3.27
Arg2	8.88	4.52	$\text{H}\beta 1$ 1.34; $\text{H}\beta 2$ 1.42; $\text{H}\gamma 1$ 0.99; $\text{H}\delta 1$ 2.79; $\text{H}\delta 2$ 2.88; $\text{H}\epsilon$ 6.90
Trp3	8.61	4.82	$\text{H}\beta 1$ 3.21; $\text{H}\beta 2$ 3.41; $\text{H}\delta 1$ 7.33; $\text{H}\epsilon 1$ 10.21; $\text{H}\epsilon 3$ 7.61; $\text{H}\zeta 2$ 7.47; $\text{H}\zeta 3$ 7.14; $\text{H}\eta 2$ 7.23
Thr4	7.22	4.67	$\text{H}\beta$ 4.40; $\text{H}\gamma 2^*$ 1.13
Pro5		4.23	$\text{H}\beta 1$ 1.98; $\text{H}\beta 2$ 2.38; $\text{H}\gamma 1$ 2.06; $\text{H}\delta 1$ 3.72; $\text{H}\delta 2$ 3.85
Val6	7.45	4.18	$\text{H}\beta$ 2.07; $\text{H}\gamma 1^*$ 0.88; $\text{H}\gamma 2^*$ 0.89
Cys7	8.27	4.57	$\text{H}\beta 1$ 2.96; $\text{H}\beta 2$ 3.16
Nh2	7.41,7.53		

Table S2. $^{13}\text{C}/^{15}\text{N}$ resonance assignments for [D-Trp3]-contryphan-Vc2 major conformer (*trans* Thr4-Pro5 peptide bond).

Residue	Heavy atom chemical shift (ppm)			
	N^{H}	N (other)	C^{α}	C (other)
Cys1	-		54.8	$\text{C}\beta$ 40.9
Arg2	122.4		55.5	$\text{C}\beta$ 31.5; $\text{C}\gamma$ 26.0; $\text{C}\delta$ 43.3
Trp3	123.7	$\text{N}\epsilon$ 129.4	57.1	$\text{C}\beta$ 29.2; $\text{C}\delta 1$ 127.4; $\text{C}\epsilon 3$ 120.9; $\text{C}\zeta 2$ 114.5; $\text{C}\zeta 3$ 121.9; $\text{C}\eta 2$ 124.7
Thr4	113.6		58.7	$\text{C}\beta$ 69.5; $\text{C}\gamma 2$ 21.3
Pro5			64.7	$\text{C}\beta$ 32.0; $\text{C}\gamma$ 27.4; $\text{C}\delta$ 51.1
Val6	114.6		61.8	$\text{C}\beta$ 32.7; $\text{C}\gamma 1$ 20.2; $\text{C}\gamma 2$ 21.3
Cys7	121.5		55.6	$\text{C}\beta$ 40.0
Nh2	108.1			

Table S3. ¹H resonance assignments for [D-Trp3]-contryphan-Vc2 minor conformer (*cis* Thr4-Pro5 peptide bond).

Residue	Proton chemical shift (ppm)		
	H ^N	H ^α	H (other)
Cys1	-	4.30	Hβ1 3.37
Arg2	8.81	4.27	Hβ1 1.44; Hγ1 0.73; Hγ2 1.00; Hδ1 2.80; Hδ2 2.85; Hε 6.82
Trp3	8.78	4.90	Hβ1 3.18; Hβ2 3.40; Hδ1 7.28; Hε1 10.25; Hε3 7.73; Hζ2 7.47; Hζ3 7.16
Thr4	7.88	4.35	Hβ 3.91; Hγ2* 1.07
Pro5		4.91	Hβ1 2.25; Hβ2 2.32; Hγ1 1.82; Hγ2 1.99; Hδ1 3.56; Hδ2 3.66
Val6	8.24	4.13	Hβ 2.19; Hγ1* 1.00
Cys7	8.39	4.57	Hβ1 3.09; Hβ2 3.22
Nh2	7.38,7.68		

Table S4. ¹H resonance assignments for [L-Trp3]-contryphan-Vc2.

Residue	Proton chemical shift (ppm)		
	H ^N	H ^α	H (other)
Cys1	-	-	-
Arg2	8.84	4.15	Hβ1 1.31; Hγ1 0.96; Hδ1 2.79; Hε 6.85
Trp3	7.28	4.78	Hβ1 3.30; Hβ2 3.43; Hδ1 7.28; Hε1 10.27; Hε3 7.68; Hζ2 7.47; Hζ3 7.17; Hη2 7.22
Thr4	7.75	4.42	Hβ 4.43; Hγ2* 1.21
Pro5	-	4.87	Hβ1 2.30; Hγ1 1.82; Hγ2 2.07; Hδ1 3.60; Hδ2 3.69
Val6	7.57	4.30	Hβ 2.12; Hγ1* 0.94
Cys7	8.37	4.73	Hβ1 2.96; Hβ2 3.29
Nh2	-		

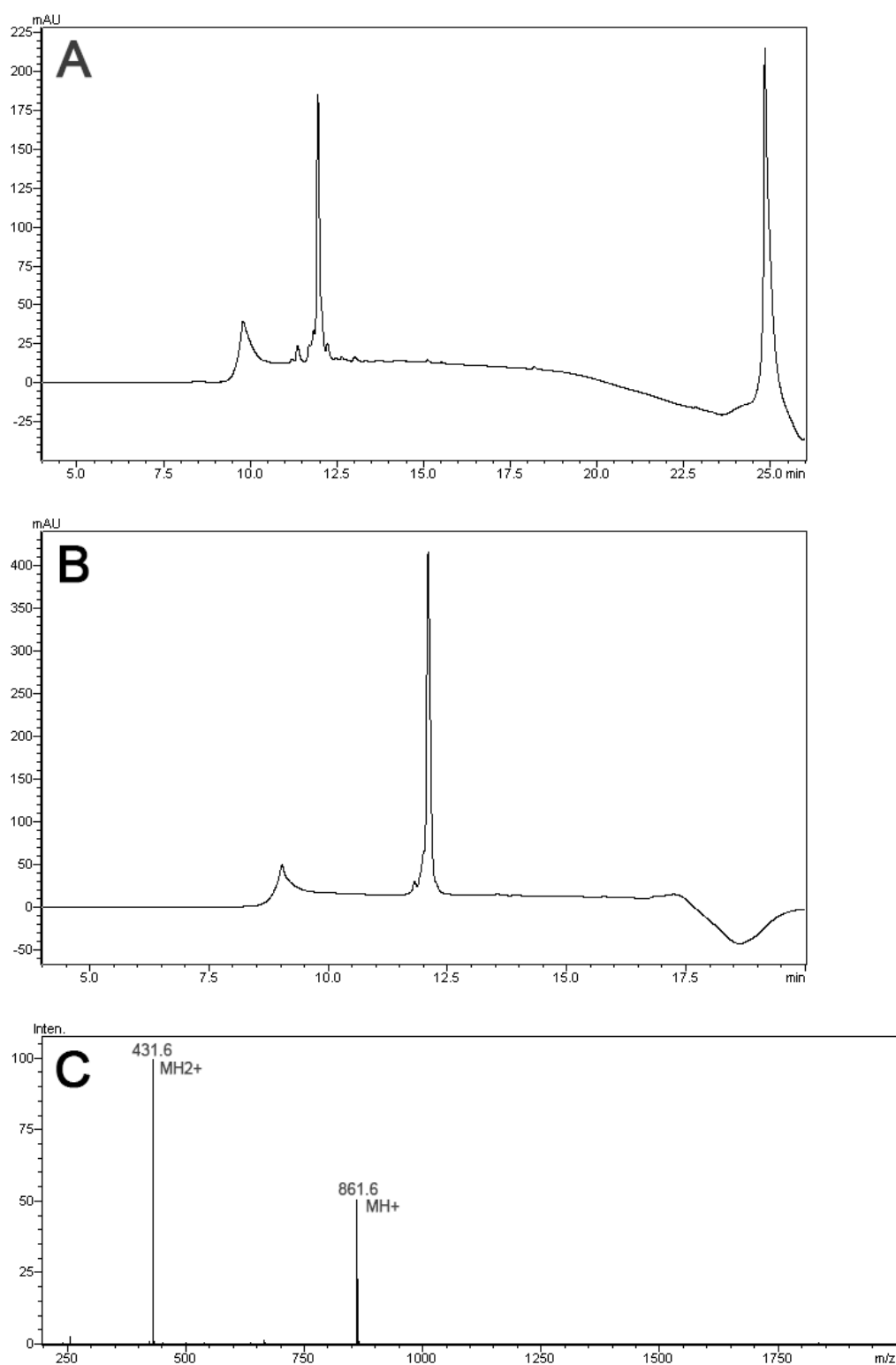


Figure S1. Purification of [L-Trp3]-contryphan-Vc2. **(A)** Oxidative folding mixture, run on LC-MS system with 3 μ m C8(2) 100 Å column, flowrate 0.2 mg/mL, gradient 0-100% B over 15 min; Buffer A: 0.05% TFA in H₂O, Buffer B: 0.05% TFA in acetonitrile. **(B)** Final purification, run on LC-MS system with 3 μ m C8(2) 100 Å column, gradient 0-60% B over 10 min; Buffer A: 0.05% TFA in H₂O, Buffer B: 0.05% TFA in acetonitrile. **(C)** Ions present in MS of pure peak.

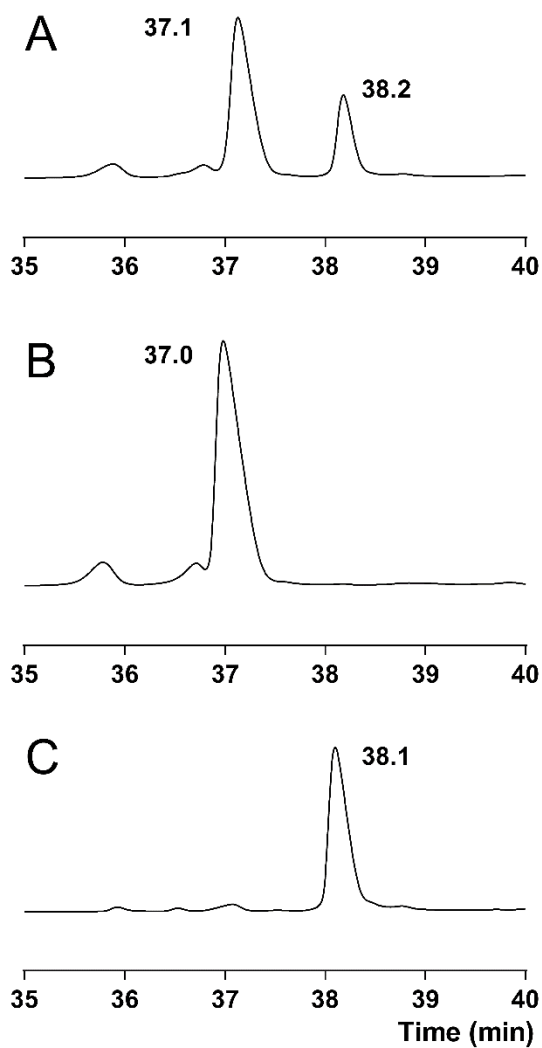


Figure S2. Co-elution of [D-Trp3]-contryphan-Vc2 and [L-Trp3]-contryphan-Vc2, run on LC-MS with a 3 μ m C8(2) 100Å column. Gradient 0-60% B over 90 min; Buffer A: 0.05% TFA in H₂O, Buffer B: 0.05% TFA in acetonitrile. (A) 50:50 mixture of [D-Trp3]-contryphan-Vc2 and [L-Trp3]-contryphan-Vc2. (B) [D-Trp3]-contryphan-Vc2. (C) [L-Trp3]-contryphan-Vc2.

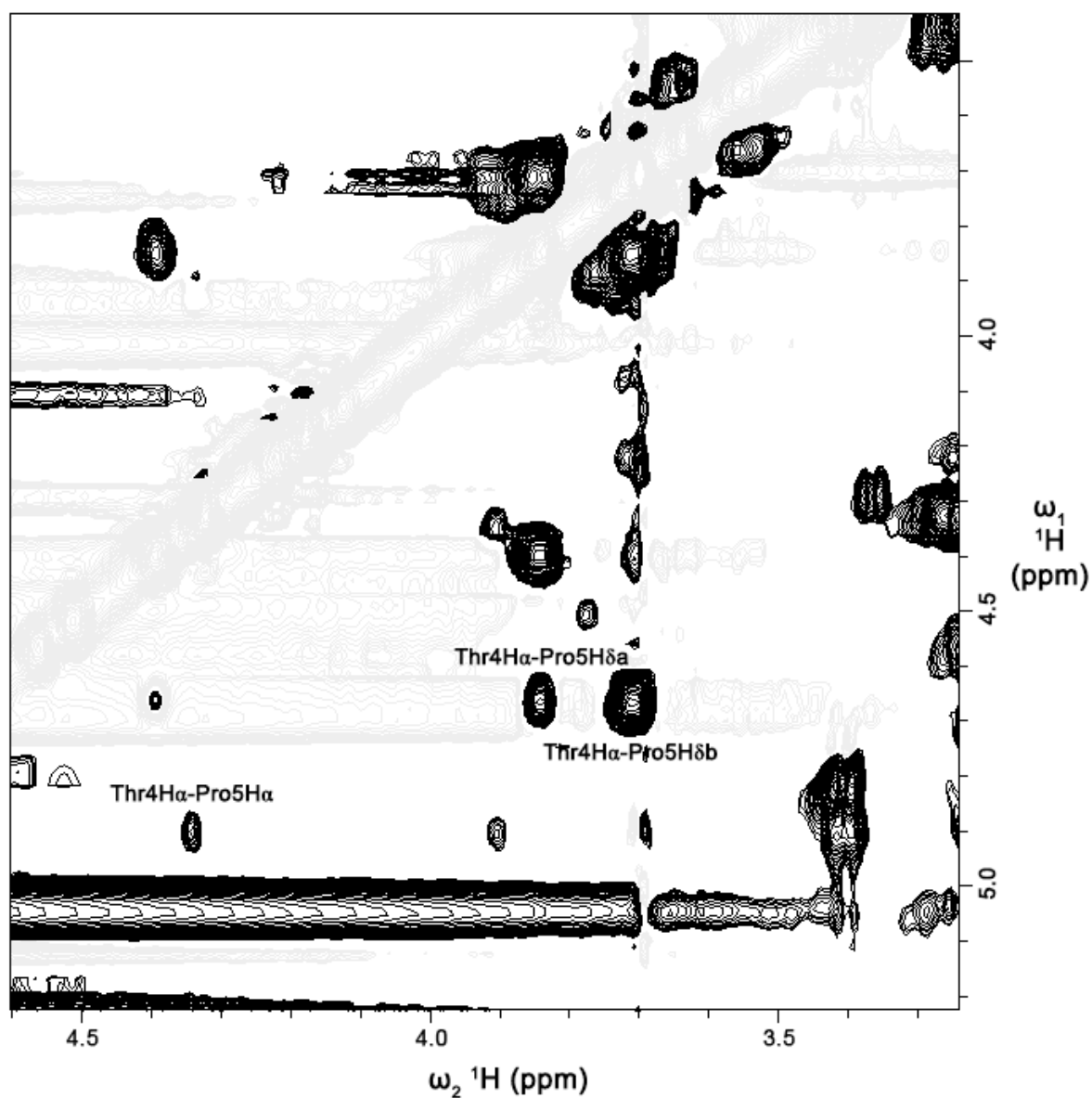


Figure S3. Section of 2D homonuclear ^1H ROESY spectrum, showing diagnostic NOE cross peaks indicating *cis/trans* isomerism of Thr4-Pro5 peptide bond. The strong Thr4H α -Pro5H δ cross peak indicates this bond is *trans* in the major form. The weaker Thr4H α -Pro5H α cross peak confirms a *cis* bond is present in the minor conformer.

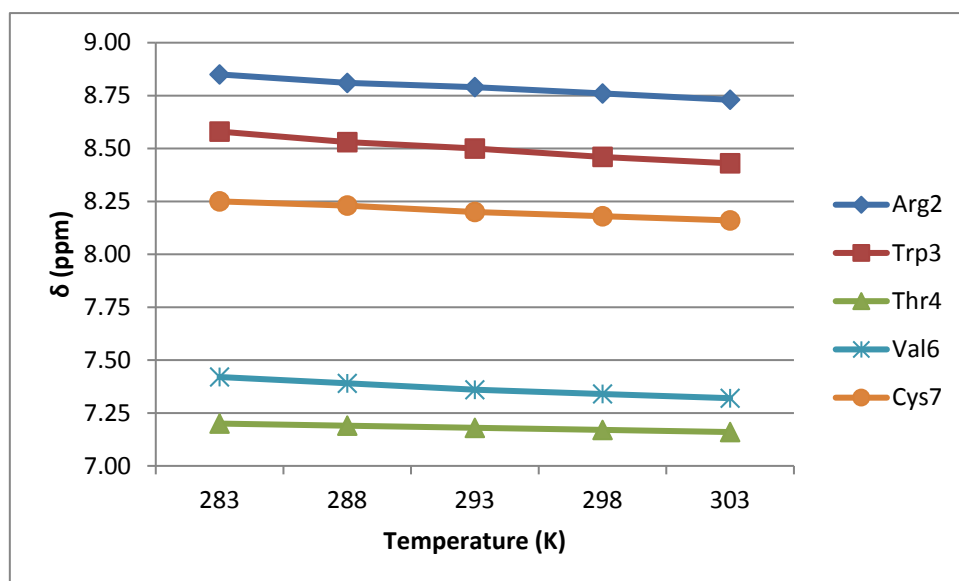


Figure S4. Plot of backbone amide chemical shifts against temperature. The gradients of the lines correspond to amide temperature coefficients.

Contryphan-Vc2 is stable at physiological temperatures

Spectra recorded on [D-Trp3]-contryphan-Vc2 at 37 °C (**Figure S5**) showed amide chemical shifts consistent with those expected from the linear temperature coefficients previously determined over the range 10-30 °C (**Figure S4**). A disruption in structure would disrupt the linearity of these relationships. Additionally, the aliphatic peaks of the spectra exhibited only minor shifts when compared at 10 °C and 37 °C (**Figure S5**), as would be expected for a stable structure. These results confirm that [D-Trp3]-contryphan-Vc2 maintains its structure at physiological temperatures.

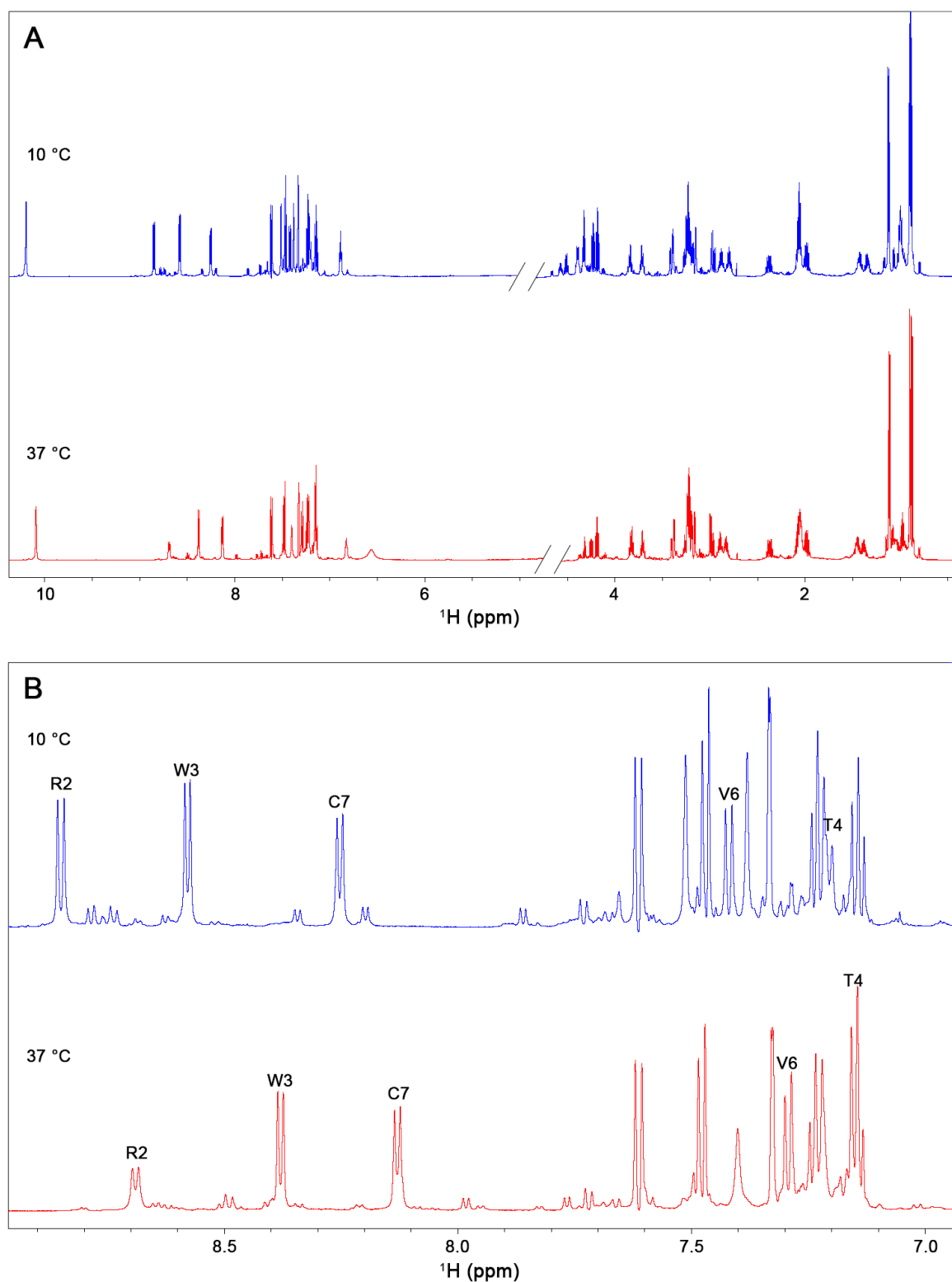


Figure S5. Comparison of NMR spectra of [D-Trp3]-contryphan-Vc2 recorded at 10 °C and 37 °C. (A) Full spectra, showing similarity in aliphatic region. The region of the spectrum in the vicinity of the water resonance has been truncated. (B) Amide region, showing shifts in amide peaks that are consistent with the linear temperature coefficient values measured over the range 10 to 30 °C (see main text and **Figure S4**), indicating that there is no conformational change over this temperature range.

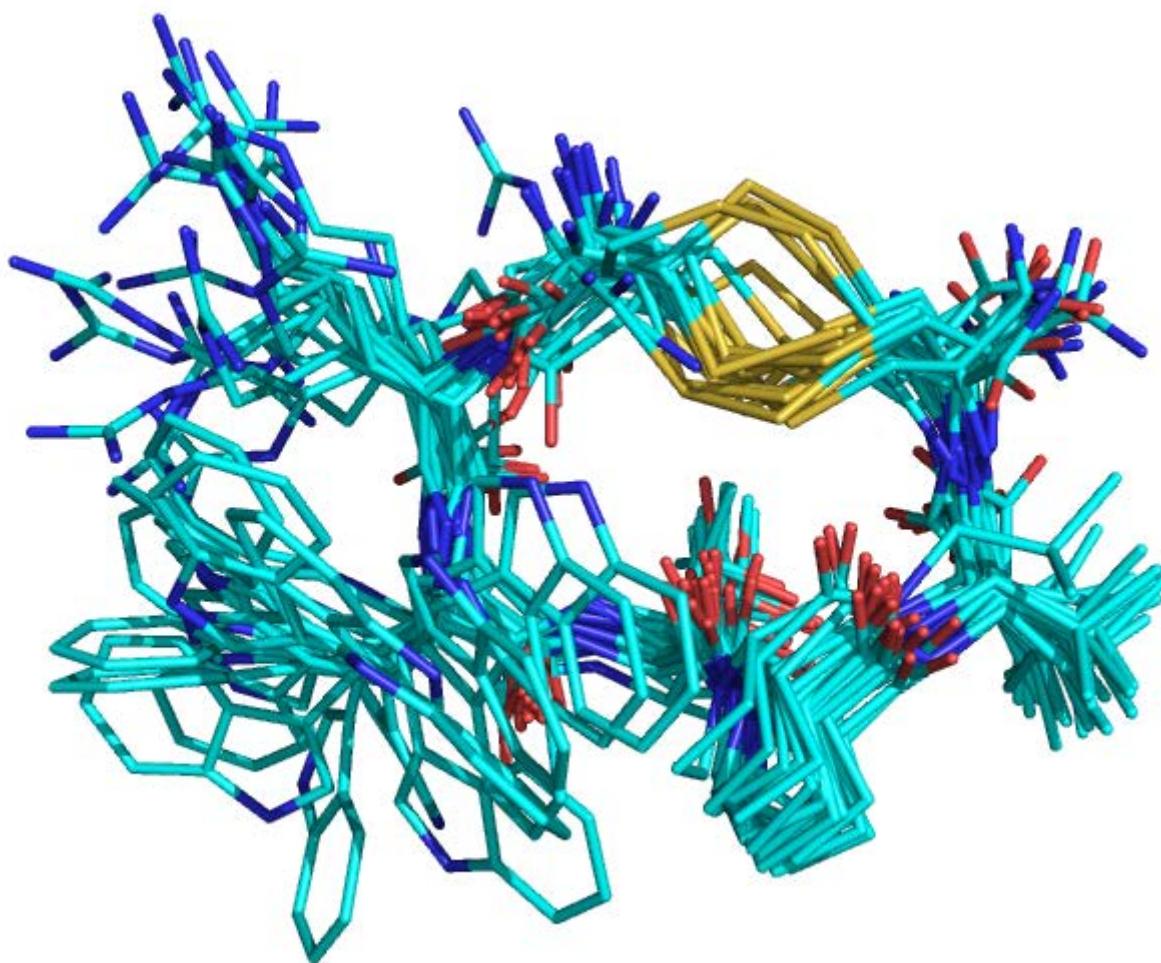


Figure S6. Aligned frames of MD simulation of [D-Trp3]-contryphan-Vc2 in water. Frames taken along length of trajectory at 5 ns intervals. Backbone is largely static, with some deviation at the disulfide bond and notable flexibility in sidechains of Arg2 and Trp3.

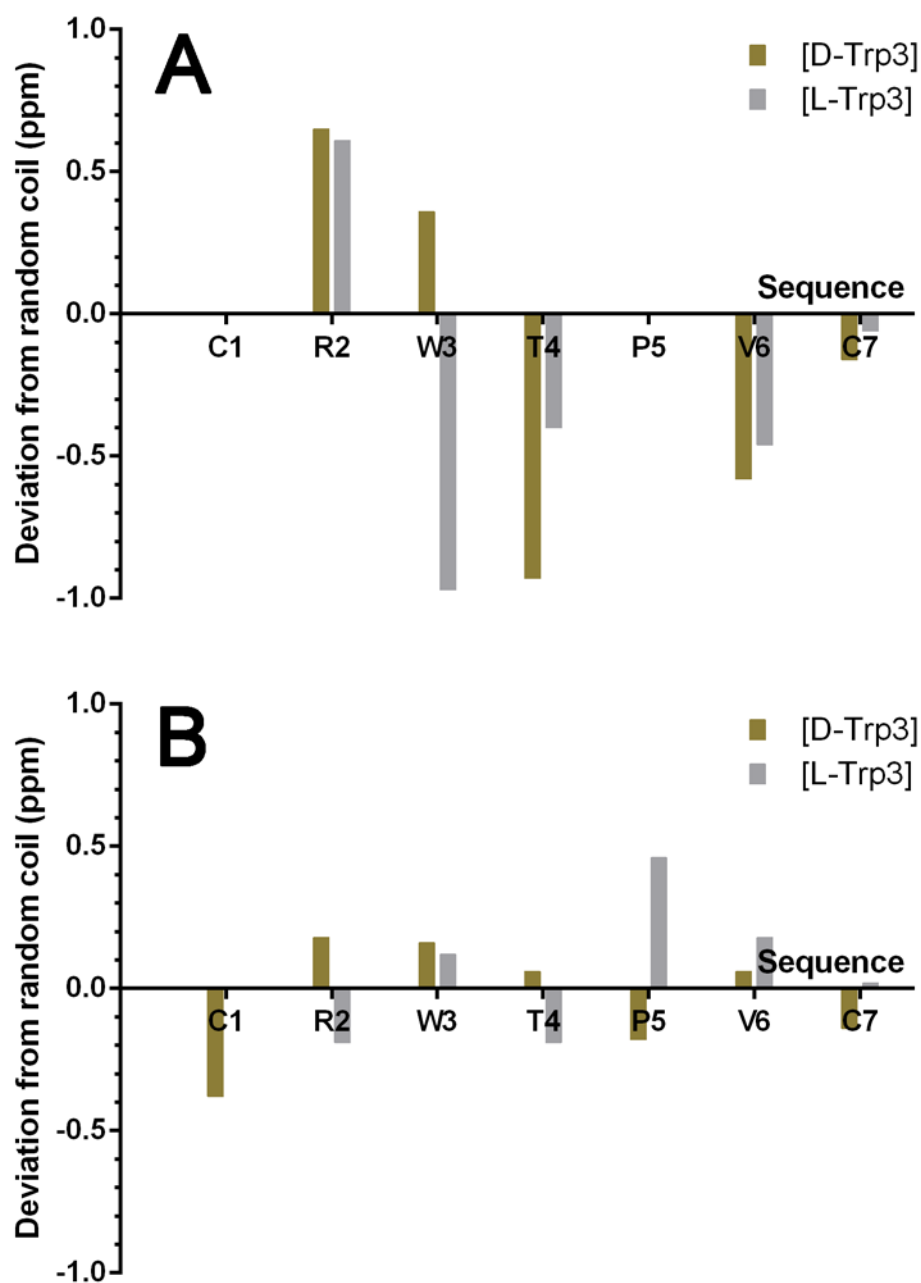


Figure S7. Deviation from random coil chemical shift plots for [D-Trp3]- and [L-Trp3]-contryphan-Vc2. Positive values indicate downfield shifts. (A) Amide proton resonances. (B) H α proton resonances.

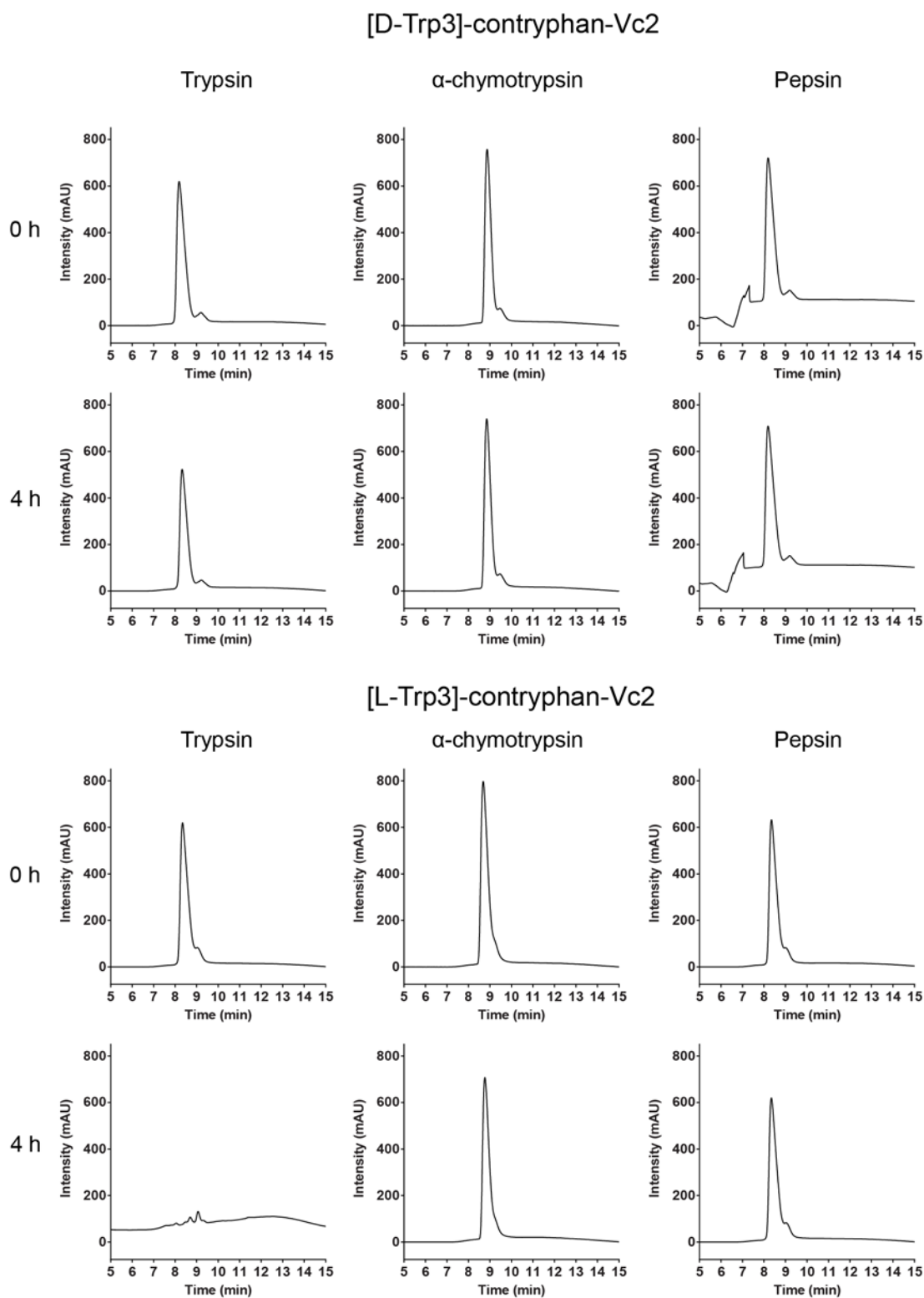


Figure S8. LC-MS traces of contryphan-Vc2 samples showing level of susceptibility to digestive enzymes after 4 h incubation at 37 °C.

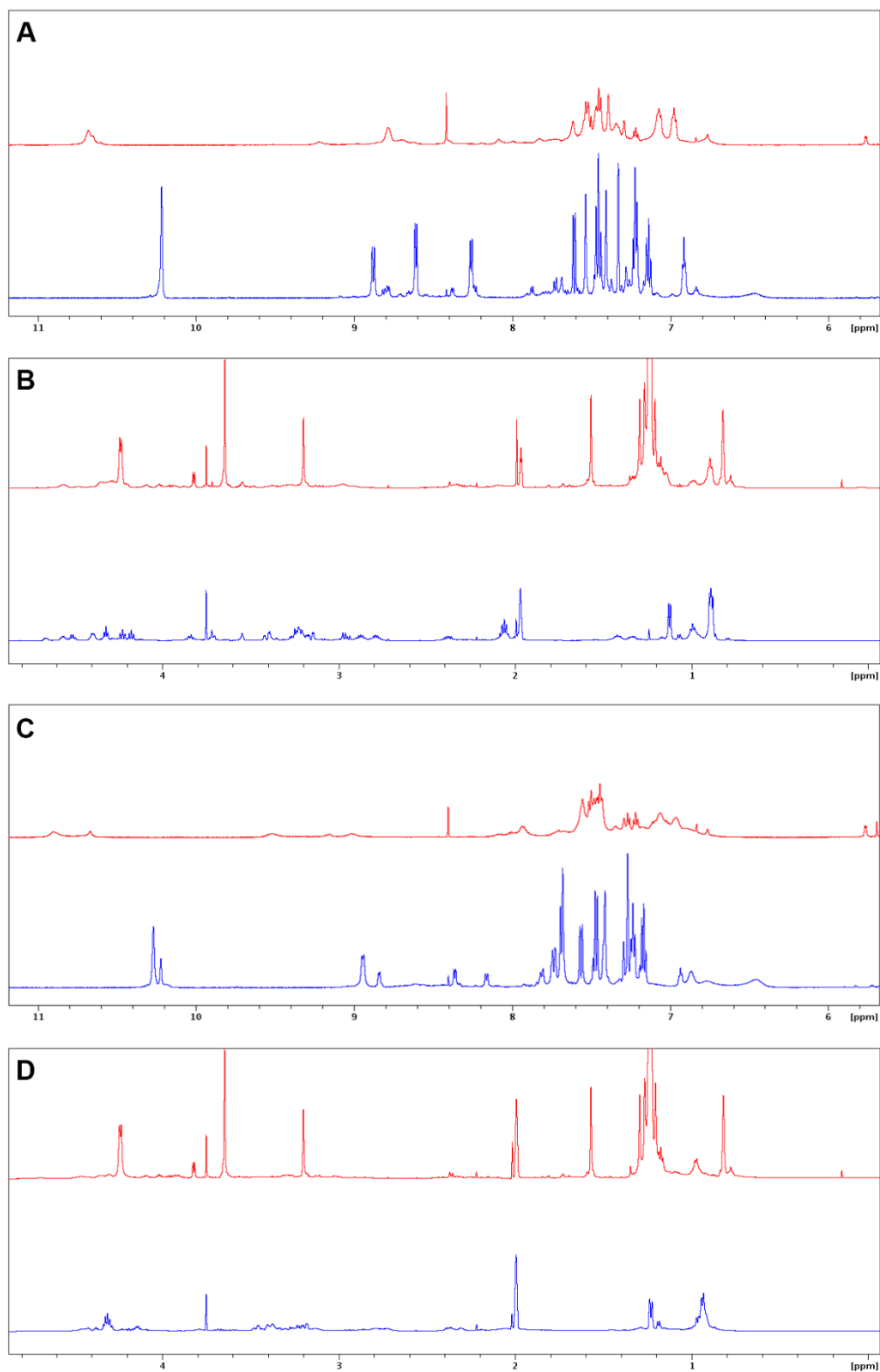


Figure S9. ^1H NMR spectra of [D-Trp3]- and [L-Trp3]-contryphan-Vc2 in absence (blue) and presence (red) of 40 mM DPC. (A) [D-Trp3]-contryphan-Vc2 amide region (B) [D-Trp3]-contryphan-Vc2 aliphatic region (C) [L-Trp3]-contryphan-Vc2 amide region (D) [L-Trp3]-contryphan-Vc2 aliphatic region.

Molecular dynamics simulations

MD simulations were run to gain insight into specific molecular interactions between peptides and lipid bilayers. Simulations of [L-Trp3]- and [W3A]-contryphan-Vc2 in the presence of mixed POPC/POPG bilayers were run for 2 μ s, while simulations of [D-Trp3]-contryphan-Vc2 were run for 3 μ s. Pairs of simulations were set up, varying in the initial position of the peptide. ‘Aqueous’ simulations started with the peptide in the water phase, while ‘buried’ simulations began with the peptide within the bilayer. **Figure S10** plots the depth of insertion of the peptide into the bilayer, comparing this value between the pairs of simulations. None of the pairs of simulations had converged after 2 μ s, and [D-Trp3]-contryphan-Vc2 had not converged within the extended 3 μ s timeframe either. The first 1 μ s of each MD trajectory was designated as equilibration time and discounted from further analysis; comparisons and inferences were drawn from the remaining portion of the trajectory.

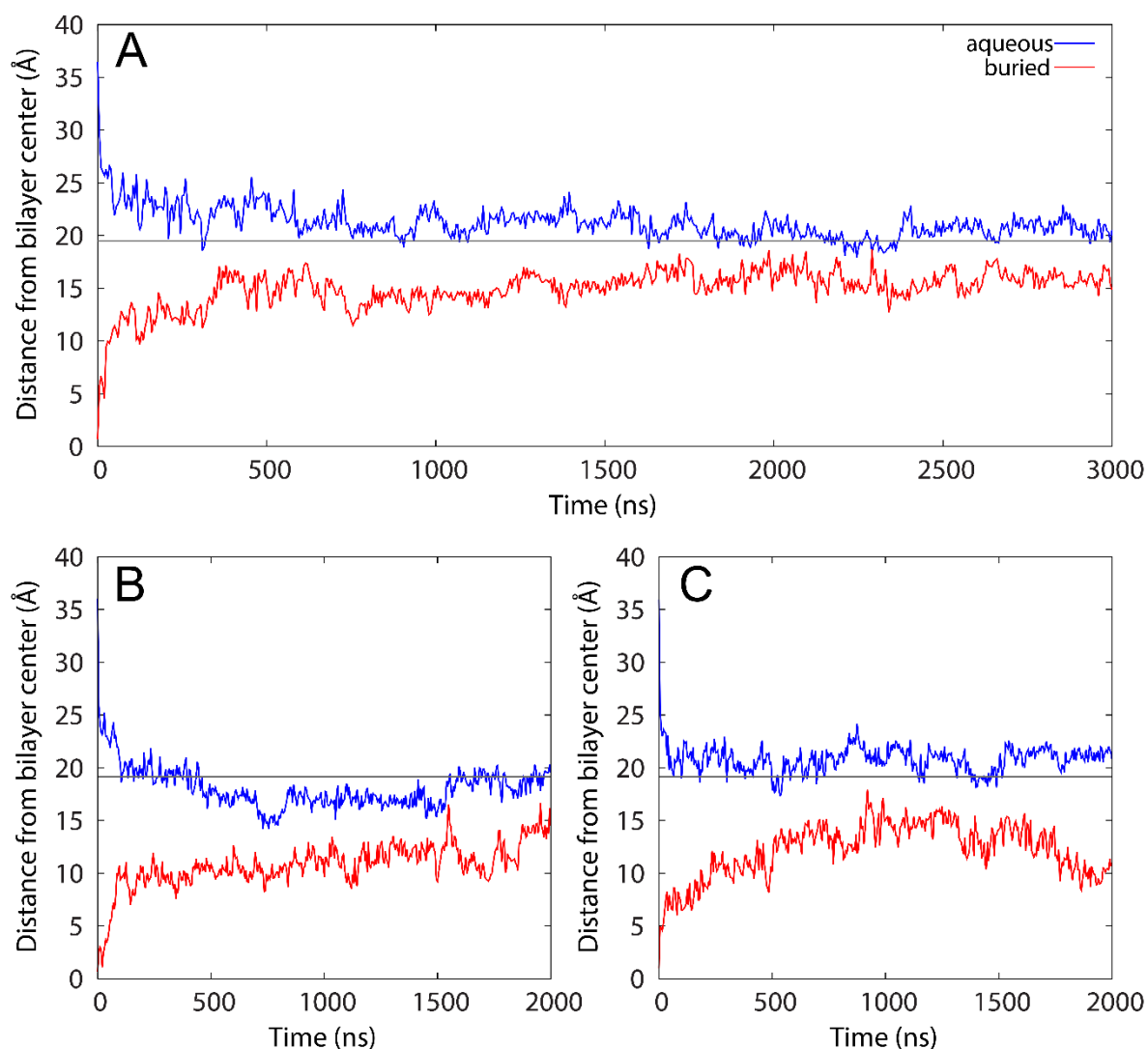


Figure S10. Insertion depth of contryphan-Vc2 centre of mass in lipid bilayer, calculated from trajectories for aqueous (*blue*) and buried (*red*) molecular dynamics simulations. *Grey* line indicates position of head group phosphate. (A) [D-Trp3]-contryphan-Vc2. (B) [L-Trp3]-contryphan-Vc2. (C) [W3A]-contryphan-Vc2.

The solvent-accessible surface area of each residue was calculated as a measure of the extent of interaction with the lipid. Plots for [D-Trp3]-contryphan-Vc2 are presented in **Figure S11**, while average values for all simulations are shown in **Table S5**.

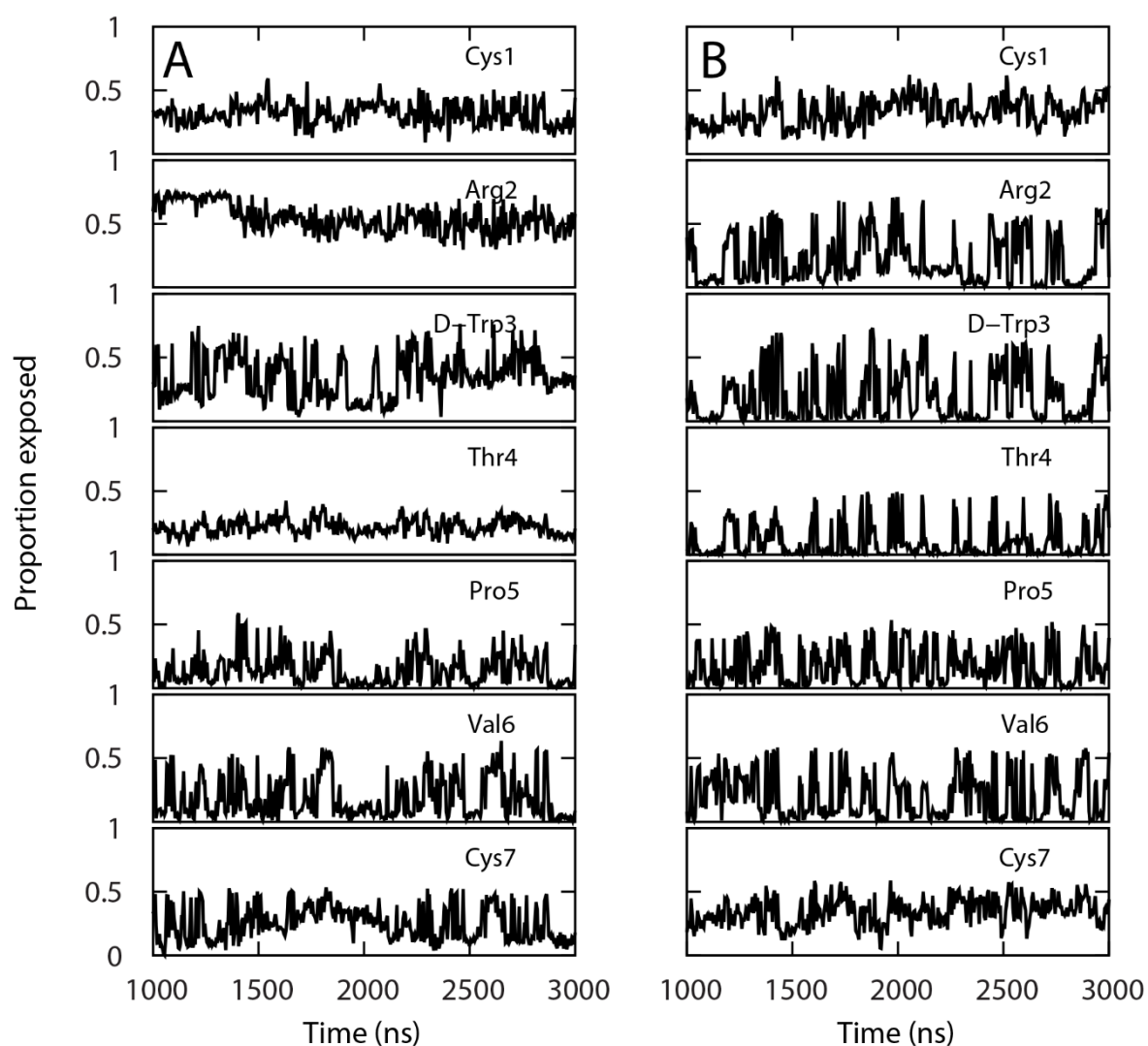


Figure S11. Solvent-accessible surface area of [D-Trp3]-contryphan-Vc2, by residue, calculated from molecular dynamics simulation trajectories. (A) Data from simulation with peptide starting position in aqueous phase. (B) Data from simulation with peptide starting position buried in bilayer.

Table S5. Average values of solvent-accessible surface area (percentage of total area) for each residue, calculated across length of MD simulation.

Residue	Simulation					
	[D-Trp3]		[L-Trp3]		[W3A]	
	aqueous	buried	aqueous	buried	aqueous	buried
Cys1	31	33	14	18	19	21
Arg2	55	22	24	30	13	38
Xaa3	35	20	28	18	8	15
Thr4	22	10	8	13	6	8
Pro5	14	17	42	25	36	17
Val6	19	20	38	22	36	15
Cys7	25	34	19	18	37	22

Thr4 is the most buried residue in all but one simulation (that being ‘aqueous’ [D-Trp3]-contryphan-Vc2). There is also no clear distinction between the relative solvent exposure of [D-Trp3]-, [L-Trp3]- and [W3A]-contryphan-Vc2, although Ala3 in [W3A]-contryphan-Vc2 is less accessible than D/L-Trp3 in the other peptides.