OXIDATIVE FOLDING OF CONOTOXINS: CHEMICAL AND BIOLOGICAL DEVELOPMENTS TO EFFICIENCY OF PRODUCTION OF SHORT, CYSTEINE-RICH PEPTIDES

by

Andrew M. Steiner

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Medicinal Chemistry

The University of Utah

May 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of Andrew M. Steiner
has been approved by the following supervisory committee members:

Grzegorz Bulaj, Chair	15 Mar. 2013 Date Approved
Amy Barrios, Member	1 Apr. 2013 Date Approved
Thomas Cheatham, Member	1 Apr. 2013 Date Approved
Ryan Looper, Member	1 Apr. 2013 Date Approved
Eric Schmidt, Member	15 Mar. 2013 Date Approved

and by Darrell Davis, Chair of

the Department of Medicinal Chemistry

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

Oxidative folding is one of the key challenges hampering the development of peptide-based compounds as therapeutics. While disulfide-rich peptides are often thought to be more appealing drug lead compounds because of their stable, highly-crosslinked structure, their oxidative folding to the correct disulfide connectivity is often difficult, and is optimized in a peptide-specific way. This work advanced knowledge of chemical and biological means to improve oxidative folding of conotoxins. Herein I present a generalized folding protocol suitable for folding diverse disulfide-rich peptides. I also show that the incorporation of selenocysteines to replace a disulfide bridge with a diselenide effectively adds an intramolecular oxidative folding catalyst, where this bridge had previously been assumed to be static, with any folding improvements being a consequence of conformational effects. These are followed by a discussion of oxidative folding mechanisms *in vivo*, relating energy expenditure to directing disulfide isomerization to the desired connectivity, as well as a novel analysis tool to consider the codon conservation of the cysteine residues that comprise the disulfide scaffold. This work represents significant improvements to chemical strategies to efficiently produce disulfide-rich peptides and genetic analyses towards a better understanding of the role oxidative folding plays in the evolution of disulfide scaffolds of cysteine-rich peptides.

For my parents

CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge numerous people for their diverse contributions and assistance throughout the course of my doctorate work, as well as those who guided me to the University of Utah to pursue my doctorate:

Grzegorz Bulaj, for his mentorship, being a sounding-board to explore ideas, his assistance in manuscript preparation, teaching me to more effectively market my research, and for showing me what life as a career-scientist really means.

Baldomero Olivera, for his extensive comments and experienced perspectives regarding the course that my research took, as well as his unwavering dedication to both the students and research under his direction.

Joanna Gajewiak, for her technical assistance in peptide synthesis, as well as being a good resource to discuss ideas relating to oxidative folding, and a good friend.

Helena Safavi-Hemami and Brad Green, for uncountable discussions about the mechanisms of peptide folding, for being a sounding board for my crazier ideas, and for being good friends.

Kenneth Woycechowsky, both for being an excellent teacher during my coursework, and a thoughtful contributor to the diselenide-catalysis portion of this work.

Ryan Looper, Amy Barrios, Thomas Cheatham and Eric Schmidt, for taking the time to serve on my committee, and in that capacity, being supportive yet blunt; demanding but reasonable. Our discussions were invaluable: each of you brought a fresh perspective and challenging questions—allowing me to forget how nervous I was, and simply focus on the science.

David Goldenberg and David Blair, for their careful review and comments on one of the manuscripts included in this work.

Phoebe Rice, of the University of Chicago, for being my first real scientific mentor, teaching me what it meant to be a research scientist, inspiring the passion for research that brought me to graduate school, and fostering a lab environment that encouraged independence, discovery and friendship.

Jun Sun, for giving me my first lab job, and for always being supportive, even through the most difficult of times.

I would also like to thank the Bulaj, Olivera, McIntosh and Yoshikami lab members for sitting through countless unintelligible presentations, and for being such an effective and cohesive team.

My wife, Liz, for listening to innumerable crazy ideas, and teaching me how to explain and teach science.

My parents and brother, for encouraging and supporting me.

To all of you, and so many more—I wouldn't be here without your help. Thank you.

The research in this thesis was supported in part under NIH Program Project Grant GM 48677, as well as funding from Johnson & Johnson and the Willard Eccles Fellowship Fund.

CHAPTER 1

INTRODUCTION

1.1 Overview of Chapter 1

Oxidative folding is critical to the chemical and biological syntheses of disulfide-rich peptides, and is the focus of this thesis. The following sections of this chapter will provide a general overview of oxidative folding. Challenges inherent to conformation-driven thiol-disulfide exchange and traditional methods of oxidative folding are also discussed. Conotoxins are a useful model system to study oxidative folding, and the value of this model system—both to chemical methodologies and pharmaceutical development—are the topic of the ensuing sections. The concluding section of this chapter outlines the scope of this dissertation.

1.2 Oxidative Folding

Oxidative folding, or the concomitant formation of disulfide bridges and three-dimensional structure in peptides and proteins, is usually mediated by various enzymes *in vivo*, or an excess of disulfide transfer reagents *in vitro*. In both cases, the ultimate source of the oxidizing equivalents for the formation of disulfide bridges is molecular oxygen; however, the direct transfer of electrons between thiols and molecular oxygen is spin-forbidden. Consequently, intermediaries are used to facilitate the oxidation of thiols to disulfides. These intermediaries can vary significantly based on context, and include enzymatic machinery with flavin adenine dinucleotide ("FAD") cofactors [1], several transition metals (copper and iron are particularly effective) [2, 3, 4], and diselenides [5]. However, these intermediaries are most commonly used to form the disulfide transfer reagents, which themselves serve as a second layer of intermediary. Glutathione is one particularly common disulfide transfer reagent, and is used both *in vivo* and *in vitro* [6, 7, 8, 9, 10].

The proper formation of disulfide bridges is critical to the function of many proteins and peptides because disulfides stabilize active conformations; this was first shown by Anfinsen's group in 1957 by reduction of the disulfides of ribonuclease A [11], and further confirmed

with the reactivation of reduced ribonuclease A in subsequent work [12]. This work gave the first hints that in disulfide-containing peptides and proteins, the formation of disulfide bridges could provide insight into folding mechanisms and intermediates. The prototypical example of the ability to use disulfides to trap folding intermediates, whereby allowing the determination of the folding pathway, is bovine pancreatic trypsin inhibitor (BPTI), by work done in the laboratory of Thomas Creighton [13]; intermediates with native disulfides were shown to dominate the folding pathway of BPTI several years later [14].

Where proteins usually have enough sequence-encoded folding information to form the native disulfide bridges with high efficiency (first shown by Anfinsen, [12]), the same is not generally true of peptides due to their dramatically shorter sequence, particularly for disulfide-rich peptides, in which cysteine can comprise over 30% of the sequence (for instance μ-KIIIA, which is 37.5% cysteine [15]). While the shorter sequence of peptides may consequently seem burdensome, it also offers advantages because short peptides can be easily synthesized using solid phase synthesis methods [16, 17], allowing facile access to sequence modifications and incorporation of unnatural amino acids. Consequently, the efficient formation of native disulfide bridges in short, cysteine-rich peptides is the challenge addressed herein.

1.2.1 The Challenge of Oxidative Folding

As is intuitively obvious, the formation of the desired disulfide bridging pattern becomes more difficult with increasing numbers of disulfide bridges, as well as with decreasing sequence length (such that the cysteines comprise a larger fraction of the total peptide sequence). If all disulfide isomers are equally accessible and equally favored (which is rarely the case), then the total number of maximally oxidized disulfide isomers, p, is given by

$$p = n!! = \frac{n!}{\lfloor \frac{n}{2} \rfloor ! \cdot 2^{\lfloor \frac{n}{2} \rfloor}} \tag{1.1}$$

where n is the number of cysteine residues (the double factorial is used to indicate the product of all odd numbers between 1 and n). My derivation of this formula is presented as Appendix A. For a more empirical view of the relationship between the number of cysteines and the possible disulfide isomers, see Table 1.1. It is worth noting that while exponential growth is usually considered the "holy grail" of growth rates, factorial growth is considerably faster. In biology, exponential growth is the fastest rate at which organisms can replicate, and factorial growth only applies to increases in complexity. Exponential and factorial growth are compared in Figure 1.1, using bacterial growth (given by 2^n, since each cell division cycle doubles the number of cells) and the number of disulfide isomers possible.

Table 1.1. Empirical view of the number of possible disulfide isomers. The left column shows the number of cysteines in a peptide, and the right column shows the number of disulfide isomers that are possible with the specified number of cysteines. The table includes values up to 16 cysteines because genes for peptides containing up to 16 cysteines have been found in venomous marine gastropods.

Cysteines	**Isomers**
1	1
2	1
3	3
4	3
5	15
6	15
7	105
8	105
9	945
10	945
11	10,395
12	10,395
13	135,135
14	135,135
15	2,027,025
16	2,027,025

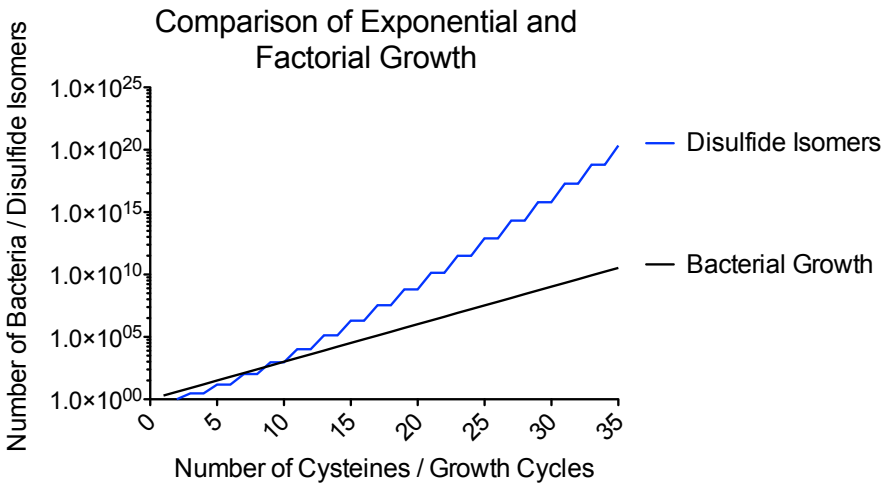


Figure 1.1. Comparison of factorial and exponential growth. The number of possible disulfide isomers is shown in blue, and increases according to Eq. 1.1 (factorial growth). Bacterial growth is used as a model for exponential growth, as each cell division cycle doubles the number of bacterial cells (modeled with 2^n). For the number of disulfide isomers, the horizontal axis gives the number of cysteines, and the vertical axis gives the number of disulfide isomers (excluding topological isomers); in the case of bacterial growth, the horizontal axis gives the number of cell division cycles, and the vertical axis gives the total number of bacterial cells. With these two examples, factorial growth first surpasses exponential growth with the addition of the ninth cysteine residue.

It is worth noting that not all of these disulfide isomers are necessarily going to be accessible to every peptide. Peptide bonds do offer some general structure, and only certain ϕ and ψ torsion angles are allowed (the ω torsion angle is virtually always planar, as a consequence of resonance structures of the peptide bond), due to spatial limitations imposed by the side chains of each residue, as well as structural limitations by the backbone bonds themselves. As an example, vicinal disulfide bonds (a disulfide between two adjacent cysteines) are usually highly disfavored. While this does impose some limitations to the number of disulfide isomers that can be accessed, it does not eliminate disulfide isomers in any predictable way. For instance, PIIIA has recently been shown to fold to three disulfide isomers which have activity, one of which includes two vicinal disulfides [18].

Additionally, a single disulfide isomer can have more than one three-dimensional form, such that bonds must be broken and reformed to convert between topological isomers. These have classically been referred to as 'knots' [19], despite not necessarily being knots in the formal mathematical sense (although they can be). Examples of 'knotting isomers' are shown in Figure 1.2, which shows both an 'effective' (top) and a 'real' (bottom) topological isomer of a single disulfide scaffold. The unknotted isomer is also shown (center). Conversion between the 'effective' topological isomers (the top two isomers in Figure 1.2) does not necessarily require breaking bonds, although the pathways available to short, disulfide-rich peptides (based on excluded volume, conformational limitations and loop lengths) would require breaking and reforming at least one disulfide. Conversion between the 'real' topological isomers requires breaking and reforming at least one disulfide.

Consequently, the oxidative folding of peptides remains an interesting challenge, and offers an excellent venue for the development of novel folding techniques and chemical approaches, many of which will be discussed further in the coming pages.

1.2.2 *In Vitro* Oxidative Folding Reactions

Oxidative folding of peptides and proteins has classically been done in solution, and requires very dilute peptide (20-50 μM) in order to avoid oligomerization [20]. There are two general approaches to *in vitro* oxidative folding: conformation-driven thiol-disulfide exchange, and stepwise disulfide bridge formation. These two methods are substantively distinct; while the former allows for single folding and purification steps, the folding yield can vary substantially, from below 5% to above 95% [21, 22], the latter theoretically only yields a single disulfide bridging pattern, but is limited by the number of orthogonal cysteine protecting groups and requires multiple purification steps. Because the challenges facing stepwise bridge formation are essentially limited to efficient purification of a single product,

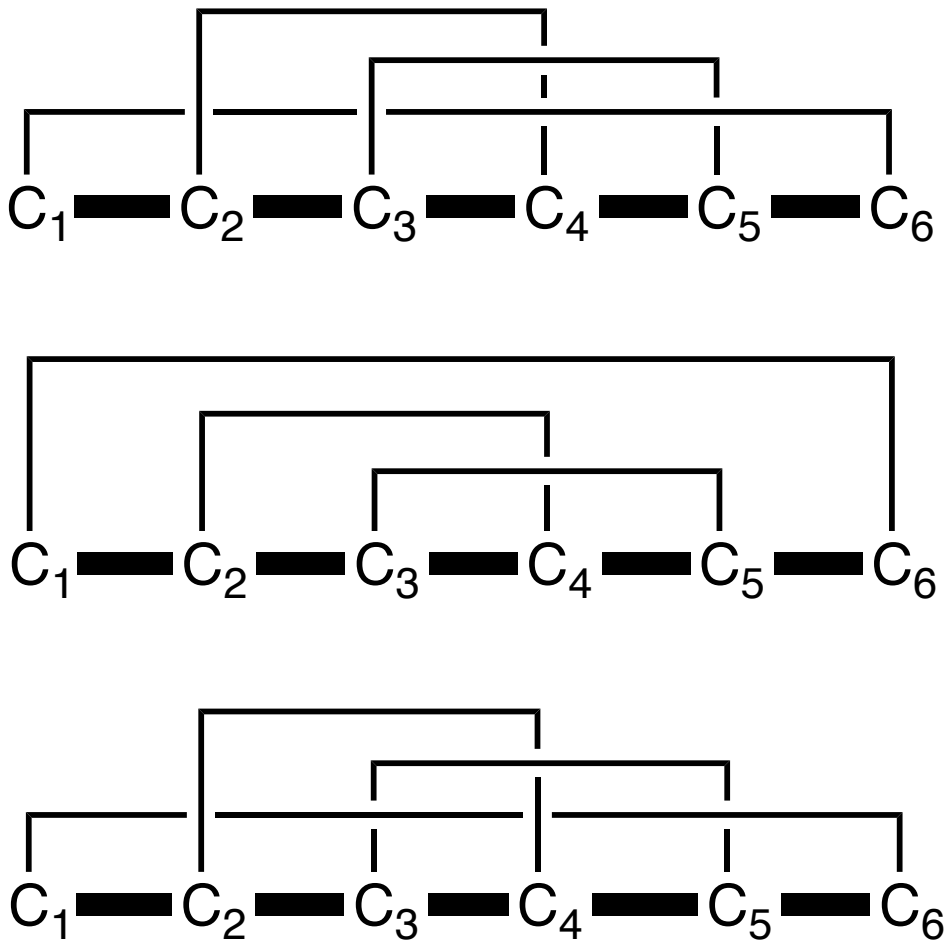


Figure 1.2. Topological isomers for a three-disulfide containing peptide. The top two isomers shown are 'effective' topological isomers, in that conversion between the two does not necessarily require breaking and reforming disulfides, although in a peptide containing less than 40 total residues, this type of conversion would only be accomplished by breaking and reforming one or more disulfide(s) due to conformational constraints. Such a conversion (top to middle) can be accomplished without breaking bonds by threading C_1 through the C_2-C_4 disulfide (back to front), and threading C_6 through the C_3-C_5 disulfide (front to back); the conversion is completed by 'lifting' the C_1-C_6 disulfide up from between the C_2-C_4 disulfide and the C_3-C_5 disulfide. The bottom isomer is a 'real' topological isomer of the upper two isomers, as any conversion between the upper two and the bottom isomer would require breaking and reforming disulfides.

I will focus primarily on conformation-driven thiol-disulfide exchange.

Numerous oxidants are employed to form disulfide bridges from free thiols, each with its own advantages and disadvantages. The most common oxidants for *in vitro* oxidative folding of peptides are disulfide transfer reagents, such as a mixture of oxidized and reduced glutathione, shown in Figure 1.3. Since its discovery in the endoplasmic reticulum [7], glutathione has become the prototypical disulfide transfer reagent because it is believed to better mimic oxidative folding *in vivo*. Other broadly-employed oxidative reagents include dimethyl sulfoxide (DMSO), iodine and transition metals, in particular iron and copper. The mechanism of DMSO-oxidation was reported in 1975, and established that the rate-limiting step was the formation of a sulfoxide-thiol adduct [23], although DMSO-based oxidation can also result in sulfoxide formation at methionine [24]. Iodine is a potent oxidant, which forms a sulfenyl iodide intermediate with thiols, which then further reacts with another equivalent of thiol to form a disulfide and two equivalents of hydrogen iodide [25]. However, disulfide formation by iodine can promote oxidation of other residues, including sulfoxide formation at methionine and oxidation of tryptophan to oxindole-tryptophan [26, 27]. The mechanism of copper-catalyzed disulfide formation is also known [3]; this mechanism is complex, and shows biphasic kinetics. Metal-catalyzed oxidation is often referred to as "air oxidation" because it usually employs trace metals already in the solution, rather than explicit addition of metals that are known to be redox-active.

1.3 *Conus* Venom Peptides as a Model System for Oxidative Folding

Conotoxins are a particularly appealing system to study oxidative folding of small, cysteine-rich peptides because they are therapeutically relevant and extremely diverse. Conotoxins exhibit a broad range of oxidative folding yields, hydrophobicities, number of disulfides, and disulfide scaffolds, making them an excellent sampling with which to lay groundwork for oxidative folding that can then be extended to virtually any cysteine-rich peptide.

Additionally, there are numerous sequences of conotoxins for which only genetic information is known, which has led to several studies on the specific effects posttranslational modifications can have on oxidative folding [28, 29]. The rate at which new sequences can be discovered by genetic means significantly exceeds the rate at which new peptides can be characterized from venom, and consequently the generation of cDNA libraries has become increasingly indispensable to conotoxin research; genomic DNA has also been employed, but

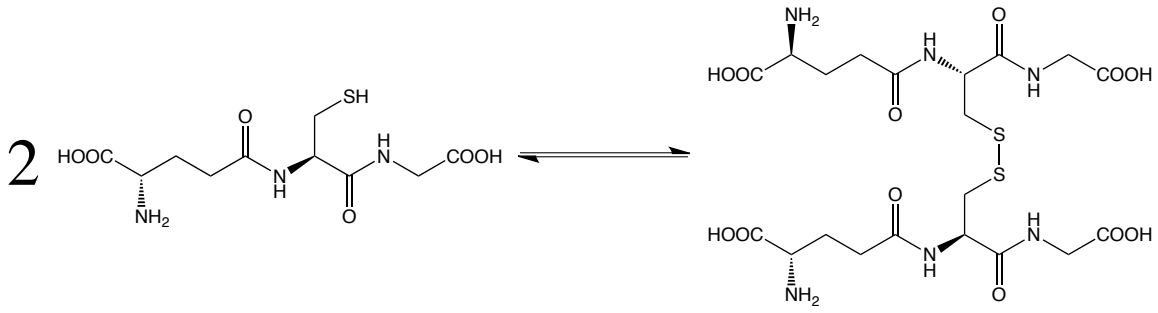


Figure 1.3. Glutathione, a typical disulfide transfer reagent used for the formation of disulfide bridges *in vitro*. To transfer the disulfide equivalent to a peptide, the peptide thiolate attacks the disulfide of oxidized glutathione (right), and the resulting mixed disulfide is subsequently attacked by another peptide thiolate; reduced glutathione (left) is added to increase the rate of disulfide isomerization. Note that the thiol hydrogen (as shown on reduced glutathione) is removed to generate the thiolate, which is what acts as the nucleophile to drive the reaction.

as the signal sequence and mature toxin are encoded in distinct exons [30, 31, 32], the correct classification of a genomic sequence into a family of conotoxins (usually accomplished by comparison of disulfide scaffolds as well as signal sequence homology) is made much more difficult with genomic DNA. Nonetheless, the peptide sequences of the mature toxins can be fairly easily identified, emphasizing the need for high-throughput synthesis and oxidative folding techniques.

Conus shows a great diversity of disulfide scaffolds (see Figure 1.4), which allows conotoxins to function as a model system for virtually any known peptide that is also short and disulfide rich (with respect to oxidative folding). Furthermore, because conotoxins are used pharmacologically by the snail to incapacitate prey, it is reasonable to expect that almost every conotoxin will show activity if presented with the correct receptor. While a typical conotoxin only has one active disulfide isomer, a peptide that has no biologically active isomers would consume energy without conferring any benefit to the snail, and so would be an evolutionary disadvantage to the snail, and selective pressures would drive its removal from the genome.

1.4 *Conus* as a Rich Source of Pharmaceutical Leads

As I have alluded to previously, conotoxins have broad therapeutic potential. Prialt was FDA approved in 2004 for the treatment of chronic pain [33], and remains the only conotoxin to have gained FDA approval to date. Nonetheless, several other conotoxins have entered clinical trials, including CVID, MrIA, Conantokin-G, Contulakin-G, and Vc1.1 [34]. While the majority have undergone clinical trials to treat intractable or neuropathic pain, Conantokin-G was being developed to treat epilepsy. Of these, only Prialt and CVID have three disulfide bridges, whereas MrIA and Vc1.1 have two disulfides and Conantokin-G and Contulakin-G have no disulfides. The bias towards peptides with fewer disulfides emphasizes the challenge that is presented by oxidative folding, and the relevance of that challenge to pharmaceutical development.

Peptide-based therapeutics have classically been less prevalent in the clinic than traditional small-molecule therapeutics because of their zwitterionic character (making them less likely to diffuse through a membrane to intracellular targets), as well as their high production cost. Conotoxins largely target extracellular receptors, whereby overcoming many of the traditional concerns about peptide-based therapeutics; however, effective delivery to the site of action remains a challenge–which was addressed by the use of

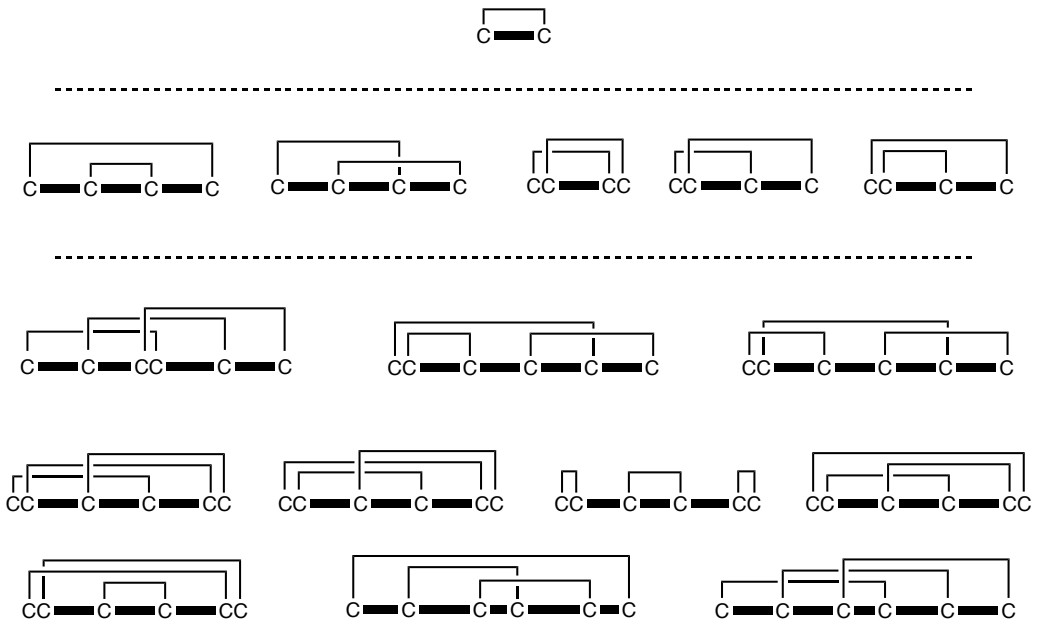


Figure 1.4. Disulfide Scaffolds in *Conus*. Only known disulfide scaffolds with three or fewer disulfide bridges are shown. Since there is only one possible scaffold for a peptide containing one disulfide bridge, only one is shown. There are several disulfide scaffolds in *Conus* which bear two disulfide bridges. The focus of my work has been conotoxins bearing three disulfide scaffolds, which are much more numerous. Only those with published examples are shown, although other connectivities and cysteine placements are possible. Intercysteine loops are shown as being of variable, nonzero length. There are also many known cysteine patterns with more than three disulfide bridges; however, in many cases the connectivity of the disulfide bonds remains unknown, and those peptides are beyond the scope of this work.

an intrathecal pump in the case of Prialt. The high production cost of peptide-based therapeutics derives partly from the cost of materials, which will be driven down as these become more prevalent; a large portion of the high production cost also comes from oxidative folding, which introduces additional steps with varying yields, further compounding the high cost of the base materials.

Despite the inherent challenges, disulfide-rich peptides are very promising drug leads; cross-linking has shown efficacy in increasing the metabolic half-life of peptides [35, 36], which likely functions by tightly packing the peptide structure, as well as 'locking-in' specific, active conformations, effectively making the backbone peptide bonds less accessible to proteases. When disulfide-rich peptides are eventually degraded, their component amino acids can simply be recycled by the patient's body, alleviating many concerns with undesired effects of drug metabolites. Additionally, the increased surface area of interaction with the target that accompanies a peptide's larger size (relative to a small-molecule therapeutic) often provides greater specificity, thereby reducing off-target effects. This last aspect has also been critical to the development of conotoxins as research tools [37, 38], largely due to their extraordinary subtype selectivity.

Oral bioavailability is one of the key challenges facing the development of peptide based drugs, although this can reportedly be overcome by backbone cyclization [39]. The advantages and disadvantages of peptide based therapeutics have been extensively reviewed, both generally [40, 41] and specific to conotoxins [33, 34].

1.5 Scope of the Thesis

While oxidative folding is often challenging, many of the inherent challenges can be overcome, as I show in the following chapters. Working with conotoxins allows the development of novel techniques to improve oxidative folding efficiency *in vitro*, as well as enhancing our understanding of the biological mechanisms involved in directing folding to the desired conformation. Furthermore, as several conotoxins are undergoing preclinical and clinical development [33, 34, 42], the tools discussed herein are directly relevant to the improvement of human health.

As much of the work with oxidative folding has been done *in vitro*, I begin with the improvements to *in vitro* oxidative folding. Chapter 2 presents a generalized oxidative folding protocol, which was a critical step in being able to assay the multitudinous sequences of unknown activity because the same methods could be used to fold multiple peptides in parallel, with reasonable yields. Chapter 3 outlines the discovery of diselenide catalysis of

oxidative folding, which allows further improvements to the throughput of *in vitro* folding, as well as allowing future work to bypass a purification step in initial screens for activity.

The discussion of evolutionary pressures involved in folding (Chapters 4 & 5) is presented in a fairly narrow scope; however, the work discussed therein has implications that reach to virtually every biological system. While it has been known for some time that the amino acid sequence of a protein encodes the necessary information to form its three-dimensional structure [43], it is all too often underappreciated that this information must also drive the evolution of the gene that encodes the protein because the correct formation of its three-dimensional structure is as indispensable as the protein itself.

1.6 References

[1] Tu, B. P. and Weissman, J. S. (2002) The FAD- and O(2)-dependent reaction cycle of Ero1-mediated oxidative protein folding in the endoplasmic reticulum. *Mol. Cell 10*, 983–994.

[2] Kachur, A. V., Koch, C. J., and Biaglow, J. E. (1998) Mechanism of copper-catalyzed oxidation of glutathione. *Free Rad. Res. 28*, 259–269.

[3] Kachur, A. V., Koch, C. J., and Biaglow, J. E. (1999) Mechanism of copper-catalyzed autoxidation of cysteine. *Free Rad. Res. 31*, 23–34.

[4] Wallace, T. J. (1966) Reactions of thiols with metals. II. Low-temperature oxidation by soluble metal salts. *J. Org. Chem. 31*, 3071–3074.

[5] Beld, J., Woycechowsky, K. J., and Hilvert, D. (2008) Catalysis of oxidative protein folding by small-molecule diselenides. *Biochemistry 47*, 6985–6987.

[6] Chakravarthi, S., Jessop, C. E., and Bulleid, N. J. (2006) The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep 7*, 271–275.

[7] Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992) Oxidized redox state of glutathione in the endoplasmic reticulum. *Science 257*, 1496–1502.

[8] Bulaj, G. (2005) Formation of disulfide bonds in proteins and peptides. *Biotechnol. Adv. 23*, 87–92.

[9] Saxena, V. P. and Wetlaufer, D. B. (1970) Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme. *Biochemistry 9*, 5015–5023.

[10] Mamathambika, B. S. and Bardwell, J. C. (2008) Disulfide-linked protein folding pathways. *Annu. Rev. Cell. Dev. Biol. 24*, 211–235.

[11] Sela, M., White, F. H., and Anfinsen, C. B. (1957) Reductive cleavage of disulfide bridges in ribonuclease. *Science 125*, 691–692.

[12] Anfinsen, C. B., Haber, E., Sela, M., and White, F. H. (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A. 47*, 1309–1314.

[13] Creighton, T. E. (1974) Intermediates in the refolding of reduced pancreatic trypsin inhibitor. *J. Mol. Biol. 87*, 579–602.

[14] Weissman, J. S. and Kim, P. S. (1991) Reexamination of the folding of BPTI: predominance of native intermediates. *Science 253*, 1386–1393.

[15] Bulaj, G., West, P. J., Garrett, J. E., Watkins, M., Marsh, M., Zhang, M. M., Norton, R. S., Smith, B. J., Yoshikami, D., and Olivera, B. M. (2005) Novel conotoxins from *Conus striatus* and *Conus kinoshitai* selectively block TTX-resistant sodium channels. *Biochemistry 44*, 7259–7265.

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

[17] Atherton, E., Fox, H., Harkiss, D., Logan, C. J., Sheppard, R. C., and Williams, B. J. (1978) A mild procedure for solid phase peptide synthesis: use of fluorenylmethoxycarbonylamino-acids. *J. Chem. Soc., Chem. Commun.* pp. 537–539.

[18] Tietze, A. A., Tietze, D., Ohlenschläger, O., Leipold, E., Ullrich, F., Kühl, T., Mischo, A., Buntkowsky, G., Görlach, M., Heinemann, S. H., and Imhof, D. (2012) Structurally diverse μ-conotoxin PIIIA isomers block sodium channel Na_V 1.4. *Angew. Chem. Int. Ed. Engl. 51*, 4058–4061.

[19] Liang, C. and Mislow, K. (1995) Topological features of protein structures: Knots and links. *J. Am. Chem. Soc. 117*, 4201–4213.

[20] Annis, I., Hargittai, B., and Barany, G. (1997) Disulfide bond formation in peptides. *Methods Enzymol. 289*, 198–221.

[21] Buczek, P., Buczek, O., and Bulaj, G. (2005) Total chemical synthesis and oxidative folding of δ-conotoxin PVIA containing an N-terminal propeptide. *Biopolymers 80*, 50–57.

[22] Cemazar, M., Zahariev, S., Lopez, J. J., Carugo, O., Jones, J. A., Hore, P. J., and Pongor, S. (2003) Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues. *Proc. Natl. Acad. Sci. U. S. A. 100*, 5754–5759.

[23] Snow, J., Finley, J., and Friedman, M. (1975) Oxidation of sulfhydryl groups to disulfides by sulfoxides. *Biochem. Biophys. Res. Comm. 64*, 441–447.

[24] Shechter, Y. (1986) Selective oxidation and reduction of methionine residues in peptides and proteins by oxygen exchange between sulfoxide and sulfide. *J. Biol. Chem. 261*, 66–70.

[25] Danehy, J. P., Doherty, B. T., and Egan, C. P. (1971) Oxidation of organic divalent sulfur by iodine. II. equilibrating thiol-iodine-disulfide-hydrogen iodide system in acetic acid and evidence for sulfenyl iodide intermediates. *J. Org. Chem. 36*, 2525–2530.

[26] Chikwana, E., Davis, B., Morakinyo, M. K., and Simoyi, R. H. (2009) Oxyhalogen-sulfur chemistry – Kinetics and mechanism of oxidation of methionine by aqueous iodine and acidified iodate. *Can. J. Chem. 87*, 689–697.

[27] Hartdegen, F. J. and Rupley, J. A. (1967) The oxidation by iodine of tryptophan 108 in lysozyme. *J. Am. Chem. Soc. 89*, 1743–1745.

[28] Bulaj, G., Buczek, O., Goodsell, I., Jimenez, E. C., Kranski, J., Nielsen, J. S., Garrett, J. E., and Olivera, B. M. (2003) Efficient oxidative folding of conotoxins and the radiation of venomous cone snails. *Proc. Natl. Acad. Sci. U. S. A. 100*, 14562–14568.

[29] Lopez-Vera, E., Walewska, A., Skalicky, J. J., Olivera, B. M., and Bulaj, G. (2008) Role of hydroxyprolines in the in vitro oxidative folding and biological activity of conotoxins. *Biochemistry 47*, 1741–1751.

[30] Yuan, D. D., Han, Y. H., Wang, C. G., and Chi, C. W. (2007) From the identification of gene organization of α conotoxins to the cloning of novel toxins. *Toxicon 49*, 1135–1149.

[31] Woodward, S. R., Cruz, L. J., Olivera, B. M., and Hillyard, D. R. (1990) Constant and hypervariable regions in conotoxin propeptides. *EMBO J. 9*, 1015–1020.

[32] Schoenfeld, R. A. (1999) *The Genomic Structure of Delta Conotoxins and Other O-Superfamily Conotoxins.* (Department of Biology, University of Utah, Salt Lake City, Utah), pp. 27–49.

[33] Lewis, R. J. (2009) Conotoxins: molecular and therapeutic targets. *Prog. Mol. Subcell. Biol. 46*, 45–65.

[34] 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.

[35] Muttenthaler, M., Andersson, A., de Araujo, A. D., Dekan, Z., Lewis, R. J., and Alewood, P. F. (2010) Modulating oxytocin activity and plasma stability by disulfide bond engineering. *J. Med. Chem. 53*, 8585–8596.

[36] Weber, S. J., Greene, D. L., Hruby, V. J., Yamamura, H. I., Porreca, F., and Davis, T. P. (1992) Whole body and brain distribution of [^{3}H]cyclic [D-Pen^2,D-Pen^5] enkephalin after intraperitoneal, intravenous, oral and subcutaneous administration. *J. Pharmacol. Exp. Ther. 263*, 1308–1316.

[37] Myers, R. A., Cruz, L. J., Rivier, J. E., and Olivera, B. M. (1993) *Conus* peptides as chemical probes for receptors and ion channels. *Chem. Rev. 93*, 1923–1936.

[38] Armishaw, C. J. and Alewood, P. F. (2005) Conotoxins as research tools and drug leads. *Curr. Protein Pept. Sci. 6*, 221–240.

[39] Clark, R. J., Jensen, J., Nevin, S. T., Callaghan, B. P., Adams, D. J., and Craik, D. J. (2010) The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chem. Int. Ed. Engl. 49*, 6545–6548.

[40] Griebel, G. and Holsboer, F. (2012) Neuropeptide receptor ligands as drugs for psychiatric diseases: the end of the beginning? *Nat. Rev. Drug Discov. 11*, 462–478.

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

[42] King, G. F. (2011) Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther. 11*, 1469–1484.

[43] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science 181*, 223–230.

CHAPTER 2

TOWARDS AN OPTIMIZED AND GENERALIZED METHOD FOR OXIDATIVE FOLDING

The majority of this chapter is a paper that was published in 2010;[1] I will provide a brief background for this work, as well as some additional concluding remarks, in light of discoveries that have happened since this paper was published.

2.1 Introduction

When I started working with conotoxins, there was no generalized protocol that could be used to efficiently fold a broad array of disulfide-rich peptides containing three or more disulfides. Rather, the oxidative folding had to be optimized for each peptide, from an original set of starting parameters. Occasionally, those starting parameters were sufficient to generate testable quantities of peptide, and this effect was only moderately predictable.

The starting parameters for oxidative folding of cysteine-rich peptides included (final concentrations) 0.1 M Tris, pH 7.5 with 1 mM EDTA, 1 mM oxidized glutathione and 1 mM reduced glutathione. The reduced, or "linear," peptide is added to this mixture such that the final peptide concentration is 20 μM. As discussed in Chapter 1, as well as in the Scope and Comments section of the following paper, the peptide concentration had to be kept dilute in order to promote intramolecularity of disulfide bond formation (effectively avoiding disulfide-linked aggregates).

The usual course of the optimization is to first do a folding timecourse using the starting parameters. These parameters can then be altered in response to the effects seen in the timecourse. For example, to promote disulfide isomerization, additional reduced glutathione can be added. If aggregation proves to be a significant issue (usually a

[1]Steiner, A., Bulaj, G. Optimization of Oxidative Folding Methods for Cysteine-rich Peptides: A Study of Conotoxins Containing Three Disulfide Bridges. *J. Pep. Sci.* 2010; **17**(1): 1-7. doi: 10.1002/psc.1283.

result of hydrophobicity of the peptide or appended groups, like a palmitoyl group on the ϵ-amine of lysine), then varying amounts of detergents or cosolvents can be added. The optimization continues until an acceptable oxidative folding yield is achieved (what constitutes "acceptable" can vary dramatically, from enough to generate testable material to fully-optimized for large-scale production). Other tangential directions can also be tested, for instance using dimethyl sulfoxide as the oxidant (in lieu of glutathione); these are also discussed in Section 2.2.

In the following paper, I created an oxidative folding method that was fairly generalizable, as well as being amenable to high-throughput oxidative folding in ways that were beyond the grasp of peptide-specific optimizations. This was accomplished by employing ClearOx, a reagent that was developed by George Barany's group [1, 2], and had been tested with conotoxins [3], but had never undergone extensive optimization or generalization. This was combined with diselenide-bridges within the peptide substrate, which were believed at the time to function by conformationally restricting the peptide, as well as reducing the number of possible outcomes (due to the favoring of the diselenide over mixed sulfur-selenium bonds, as demonstrated in [4]).

Additionally, one of the critical aspects of this paper was the introduction of "Apparent Folding Yield," as a term to describe the total amount of correctly folded peptide that results from a folding reaction, relative to the amount of peptide that was added to the reaction (effectively a mass/mass yield measurement on a scale too small to accurately measure mass). Because the high peptide concentrations used for ClearOx folding (3 mM) risked aggregation, as well as covalent linkage to the immobilized oxidation medium, the Apparent Folding Yield is distinct from the "Relative Folding Yield," which is the fraction of the recovered peptide that corresponds to the correctly folded form. Most work to date had employed Relative Folding Yield, referring to this as the yield of the folding reaction.

The General Optimized Procedure, Section 2.2.1.1, presents an abbreviated form of the full protocol outlined in this paper, which is detailed in Section 2.2.3.2. To expand on this abbreviation slightly, only linear peptide was used in the folding reactions. The acidification with formic acid resulted in a final pH of approximately 1.74. The number of molar equivalents of ClearOx was determined based on [2], and the resin was prepared by swelling in dichloromethane, followed by serial washes in dimethylformamide, methanol, 50% acetonitrile in water and finally 50% acetonitrile in 0.1 M MOPS, pH 7. The method of preparing the resin and subsequently dividing it into the reaction vessels likely does not merit the number of significant figures that is presented in the text, which was maintained

from [2].

The dimethyl sulfoxide concentrations used ranged from 0.7-1.4 M (approximately a 11,700 - 23,400-fold excess over the disulfides in the peptide); this overwhelming excess may be the cause of the slightly more random distribution of folding isomers that is seen with dimethyl sulfoxide-mediated folding. Despite the overwhelming excess, folding by dimethyl sulfoxide is very slow compared to folding with glutathione or ClearOx.

Additionally, some care should be taken with the terminology used in this paper. While glutathione is discussed as an oxidant, this is not entirely true; glutathione disulfide (or oxidized glutathione) is a disulfide transfer reagent, whereas reduced glutathione drives disulfide isomerization, but is not an oxidant in any capacity. Also, while Prialt is the only FDA-approved conotoxin, five other venom-derived peptides have also been approved by the FDA: Capoten, Integrilin, Aggrastat, Angiomax and Byetta [5].

For each of the folding reactions, as well as the yield determinations, the absorbance was monitored at 220 nm. This wavelength was chosen because it maintains a consistent baseline (allowing more meaningful integration of peaks), as well as detecting the peptide bonds with reasonable efficiency. This is also why the same peptide is used to determine the 'Apparent Folding Yield'–the number of peptide bonds must be the same, and preferably in a similar overall conformation.

The observed differences between the 'Apparent Folding Yield' (AFY) and the 'Relative Folding Yield' (RFY) could result from a number of factors. For the glutathione-containing reactions with SIIIA, the AFY is systematically higher than the RFY, suggesting that the total integral (in absorbance·minutes) of the folding chromatogram may be artificially high due to folding intermediates with glutathione-peptide disulfides, effectively increasing the number of peptide bonds detected. For the diselenide-containing peptides, the AFY is systematically lower than the RFY, suggesting that the diselenide may absorb 220 nm light differently than a disulfide in a comparable chemical environment (since the disulfide-only form was used as the reference to determine the AFY). Conformation-based effects on the absorbance were assumed to be negligible, although this may not be entirely true; any conformation-based deviations in absorbance would only affect the RFY. However, it is evident from the values of the AFY and RFY for ClearOx that the AFY does correct for disulfide adducts to the resin, as well as potential aggregation. Thus, provided that the correct peptide is used as the reference material, the AFY is likely to be the superior measure.

2.2 Optimization of Oxidative Folding Methods for Cysteine-rich Peptides: A Study of Conotoxins Containing Three Disulfide Bridges

Steiner, A., Bulaj, G. Optimization of oxidative folding methods for cysteine-rich peptides: A Study of conotoxins containing three disulfide bridges. *J. Pep. Sci.* 2010; **17**(1): 1-7. doi:10.1002/psc.1283.

2.2.1 Abstract

The oxidative folding of small, cysteine-rich peptides to selectively achieve the native disulfide bond connectivities is critical for discovery and structure-function studies of many bioactive peptides. Because the propensity to acquire the native conformation greatly depends on the peptide sequence, numerous empirical oxidation methods are employed. The context-dependent optimization of these methods has thus far precluded a generalized oxidative folding protocol, in particular for peptides containing more than two disulfides. Herein we compare the efficacy of optimized solution-phase and polymer-supported oxidation methods using three-disulfide bridged conotoxins, namely μ-SIIIA, μ-KIIIA and ω-GVIA. The use of diselenide bridges as proxies for disulfide bridges is also evaluated. We propose the ClearOx-assisted oxidation of selenopeptides as a fairly generalized oxidative folding protocol. The generalized reaction scheme for μ-SIIIA is shown in Figure 2.1

2.2.1.1 General Optimized Procedure(s)

2.2.1.1.1 Glutathione: Ten-fold concentrated solutions were used to prepare a folding reaction with 1 mM oxidized glutathione, 1 mM reduced glutathione, 0.1 M Tris, pH 7.5, 0.1 mM EDTA, 20 μM μ-Conotoxin SIIIA. The reaction was allowed to proceed for 16 hours, and was quenched by acidification with formic acid to a final concentration of 8%.

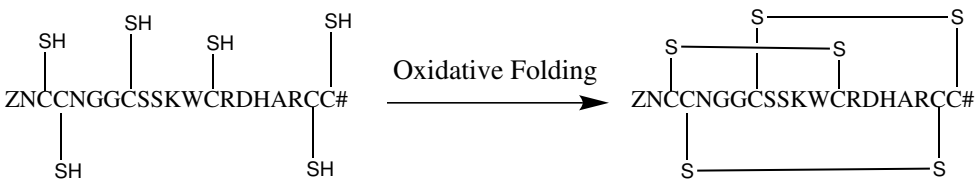


Figure 2.1. Generalized reaction scheme, shown with μ-SIIIA. Z denotes pyroglutamate, and # denotes amidation at the C-terminus.

2.2.1.1.2 Dimethyl Sulfoxide: To a dilute solution of dimethyl sulfoxide (11.63%) were added 10-fold concentrated solutions to a final concentration of 0.1 M Tris, pH 7.5, 0.1 mM EDTA and 20 μM μ-Conotoxin SIIIA, making the final DMSO concentration 10%. The reaction proceeded for 16 hours, and was quenched by acidification with formic acid to a final concentration of 8%.

2.2.1.1.3 ClearOx: To 12 molar equivalents of swelled and washed resin was added a solution of 3 mM μ-Conotoxin SIIIA in 50% acetonitrile, 50% 0.1 M MOPS, pH 7. The reaction proceeded for 1 hour, and was quenched by 100-fold dilution with 0.1% TFA in water.

2.2.2 Scope and Comments

Disulfide bonds conformationally restrict a peptide, which is crucial to their biological activity. However, achieving specificity in the topology of disulfide formation during chemical synthesis is difficult, in particular for peptides containing more than two disulfide bridges (for two disulfide bridged peptides, the orthogonal side chain protection of pairs of cysteines and regioselective folding is the method of choice [6]). The challenge in chemical synthesis of disulfide-rich peptides is reflected in their oxidative folding yields, ranging from prohibitively low to quantitative [7]. Coincidently, a majority of cysteine-rich bioactive peptides contain three or more disulfide bridges, including plant-derived protease inhibitors, defensins from both vertebrates and invertebrates, and neurotoxins from spiders, scorpions and mollusks. Despite the fact that these peptides comprise a pool of millions of bioactive compounds, many of which have already become research tools and even therapeutics, their discovery and structure-function studies have progressed at a relatively slow pace. Due to their small size, chemical synthesis has been the method of choice over recombinant expression, but efficient oxidative folding methods remain the bottleneck in exploring the full potential of cysteine-rich peptides [8, 9, 10].

Herein we compare oxidation methodologies in the context of conotoxins, a large class of cysteine-rich peptides that includes one of the only venom-derived, FDA-approved biopharmaceuticals to date [11, 12, 13, 14]. These peptides are very short (10-40 residues), usually with 2-4 disulfide bonds, and exhibit a wide range of folding yields (see Table 2.1). *Conus* peptides are an excellent model system for the study of disulfide bond formation because they are natural product peptides that are sufficiently complex and diverse to address the myriad variables of oxidative folding.

There are several methods used to oxidize cysteine-rich peptides to their folded products; these methods fall into three broad categories: solution phase, polymer-bound oxidants

Table 2.1. * indicates that these values are approximated from the graph provided in the publication. † indicates that this is thermally optimized, but without folding additives, and folding additives improve folding yield. ‡ indicates that Tween-40 was used as a folding additive. ♣ indicates that Methanol was used as a folding additive. ♠ indicates that a special preparation of a ternary matrix (miniGroEL/DsbA/PPI/Agarose) was used to fold the peptide. ♢ indicates that the peptide is present in humans, but was originally discovered in teleost fish.

Comparing Oxidative Folding Yield for Various Cysteine-Rich Peptides				
Peptide	Origin	Folding Method	Yield	Ref.
μ-SIIIA	*C. striatus*	Solution	18%* ± 5%*	[15]
μ-GIIIA	*C. geographus*	Solution	49.3% ± 3%*	[16]
ω-MVIIC	*C. magus*	Solution†	8%* ± 1%*	[16]
ω-MVIIC	*C. magus*	ClearOx†	13%* ± 2%*	[16]
δ-PVIA	*C. purpurascens*	Solution†	2.1% ± 0.5%*	[17]
δ-PVIA	*C. purpurascens*	Solution‡	6.4% ± 0.8%*	[17]
α-ImI	*C. imperialis*	Solution♣	73% ± 4%	[18]
α-GI	*C. geographus*	Solution	68%	[3]
α-GI	*C. geographus*	Hydrocarbon resin	34%	[3]
α-GI	*C. geographus*	ClearOx	80%-88%	[3]
μ-PIIIA	*C. purpurascens*	Solution	20% ± 1%*	[3]
μ-PIIIA	*C. purpurascens*	Hydrocarbon resin	15% ± 1%*	[3]
μ-PIIIA	*C. purpurascens*	ClearOx	32.5% ± 2.5%*	[3]
Cn5	Scorpion	Resin-bound Media♠	87%	[19]
α-AahI	Scorpion	Solution	0.3%-2%	[20]
β-CssII	Scorpion	Solution	34.1%	[21]
AAI	*A. hypocondriacus*	Solution	$> 95\%$	[22]
Urotensin II	Human♢	ClearOx	54%	[2]

and on-resin oxidation. To form the disulfide, deprotonation of the thiol is required to make the reactive nucleophilic thiolate [23, 24]; consequently, the pH must be buffered (near neutral) to control the thiol/thiolate equilibrium (see Table 2.S1 for pH dependence). The protonation/deprotonation of cysteine residues is sequence dependent [25], and this property could be used to modulate reactivity of individual cysteines. The other variables in the folding reaction, such as peptide concentration, temperature, and folding time should be optimized. Further optimization of peptide folding can be accomplished using folding additives. These factors can often be predicted based on the sequence of the peptide. Salt can suppress electrostatic interactions [26], and organic solvents that are miscible with water, detergents, or 1-ethyl-3-methylimidazolium acetate can significantly increase the relative folding yield of hydrophobic peptides [18, 27, 28].

Solution phase oxidation is the most commonly employed method to form intramolecular disulfide bonds in peptides [18, 26, 27, 29, 30]. Solution-phase oxidation requires very dilute peptide (20-50 μM), in order to ensure that the disulfide bonds form intramolecularly. At higher peptide concentrations, the effective concentration of the intramolecular thiols is surpassed by the concentration of intermolecular thiols, leading to oligomers and resulting in lower folding yields [6]. Because solution-phase oxidation has been thoroughly explored, there are numerous possible oxidants (glutathione, selenoglutathione, cysteine, cystamine, DMSO, (+/-)-trans-1,2-bis(2-mercaptoacetamido)cyclohexane (BMC) [31], and molecular oxygen are common examples), with context-dependent variations in the folding efficiency. While selenoglutathione can oxidize peptides at mildly acidic pH ($\sim$ 5), to our knowledge, it has not been used to fold peptides with three or more disulfides [32]. Glutathione-based redox buffers are common, as they are believed to better reflect oxidation pathways *in vivo* [33, 34]. The folding efficiency and extent of oxidation are also influenced by the ratio of oxidized to reduced glutathione (see Figure 2.2). Dilute dimethyl sulfoxide (final concentration 5-10%) is another common oxidant. The product ratios from glutathione and DMSO methods are often very similar, although dimethyl sulfoxide oxidation can give a slightly more random distribution of folding isomers (see Figure 2.3), and often requires longer folding times (see Table 2.S4). Molecular oxygen is also used to oxidize thiols to disulfides. Drawing oxygen from either air or oxygen atmosphere, this method relies on low-levels of oxygen dissolved in solution; consequently, these methods are relatively slow (usually requiring several days) [6].

Folding at high dilution presents problems with both scaled-up and higher throughput folding; these problems can often be overcome using polymer-bound oxidation. With this

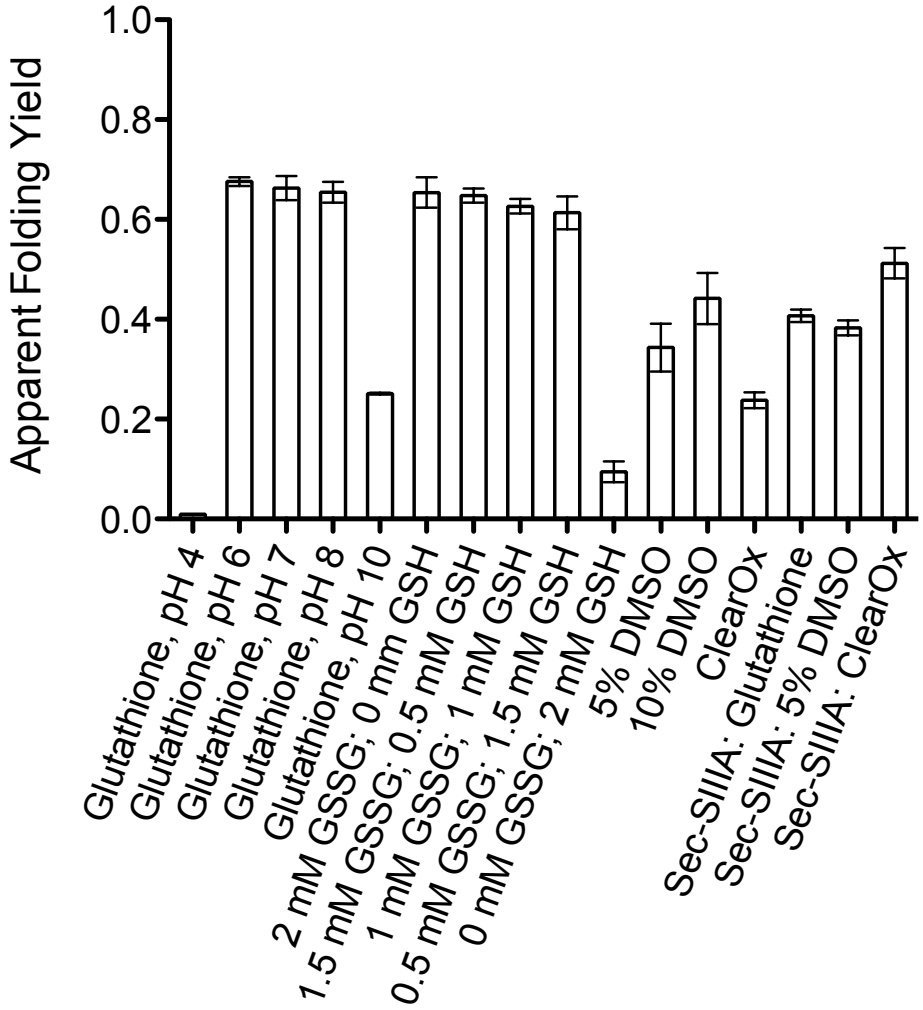


Figure 2.2. Apparent folding yield of SIIIA oxidation under various oxidation conditions. Error bars show standard deviation from 3 analytical HPLC separations. Unless otherwise stated, "Glutathione" indicates 1 mM oxidized (GSSG) and 1 mM reduced (GSH) glutathione. All reactions with glutathione or DMSO proceeded for 16 hours. ClearOx folding reactions were carried out with 12 equivalents of ClearOx resin, and proceeded for 1 hour. HPLC separations were monitored by measuring UV absorbance at 220 nm.

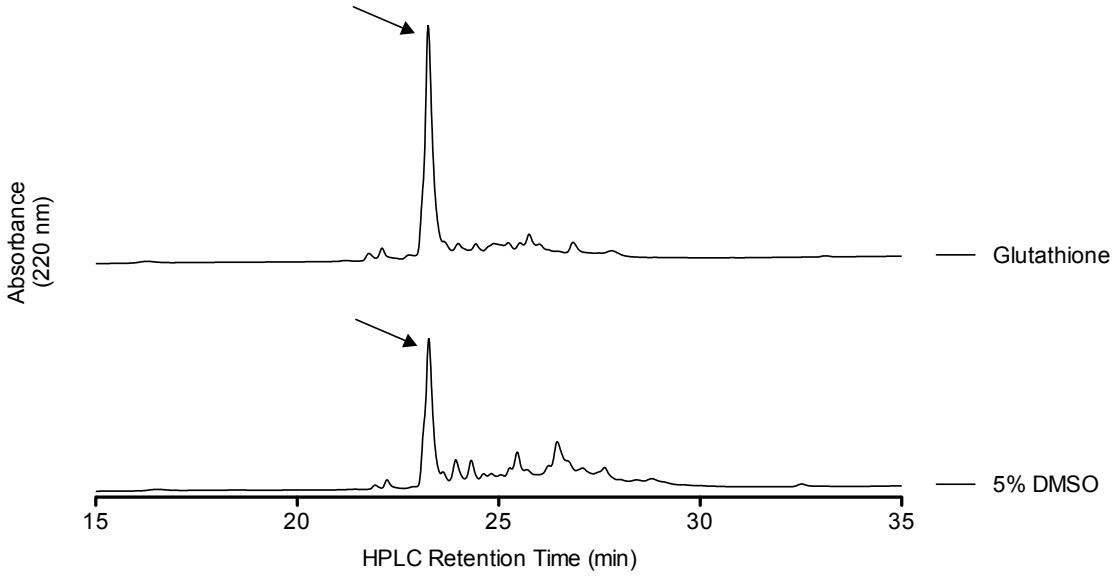


Figure 2.3. Representative HPLC traces of glutathione- and DMSO-assisted folding of SIIIA. Both folding reactions were carried out at pH 7.5 for 16 hours at room temperature. Also, both HPLC separations were done with the same gradient, from 4.5% to 31.5% acetonitrile, with 0.9% change per minute, and maintaining 0.1% TFA throughout.

type of oxidation, the oxidant is bound to a solid support at low loading levels, exploiting the pseudodilution effect to fold peptides at significantly higher concentrations; nonetheless, intramolecular disulfide bonding remains preferred. The low loading levels ensure that each molecule of the peptide is isolated, and intramolecular disulfide bonds are favored. With respect to polymer-bound folding methods, we evaluated the use of ClearOx resin. Solution-phase oxidants to fold peptides isolated on a hydrocarbon-based resin has been previously explored [3]; however, the (relative) folding yields did not justify further consideration of these methods (see Table 2.1). On more complicated or difficult folding substrates, agarose-bound folding media that includes folding chaperones and oxidant has also been employed [19].

With peptides containing one or two disulfides, oxidative folding before cleavage from the resin has been reported [10, 35, 36]. While on-resin folding methods are quite effective for folding one or two disulfide bonds, successful folding of three or more disulfide bonds with resin-bound peptide has not been reported to our knowledge.

μ-Conotoxin SIIIA (shown in Figure 2.1) has three disulfide bridges and has been the subject of several folding studies [15, 29, 30]. Noteworthy, μ-SIIIA is a potent blocker of neuronal subtypes of sodium channels, and exhibits analgesic activity in animal models of pain [37]. Also, SIIIA has a known three-dimensional structure [38]; it is used here as the model peptide to compare various oxidation methods. The wider applicability of these folding methods is demonstrated with additional *Conus* peptides, μ-KIIIA and ω-GVIA, and these results are presented in the Supplemental Information.

To describe efficiency of the folding methods, we employed two terms: 'relative folding yield' (accumulation of the native species, relative to all folding species, determined from HPLC separations of the quenched folding mixtures) and 'apparent folding yield' (recovery of the native species, relative to the amount of linear form used in the folding reaction). As illustrated in Figure 2.4, relative folding yield is taken to mean the fraction of the integrated HPLC peak corresponding to the natively folded species, relative to the total integrated HPLC area corresponding to all folding species in a given HPLC chromatogram. This is the canonical measurement of refolding efficiency on the analytical scale. Apparent folding yield is the fraction of correctly folded peptide to peptide in the reaction, accomplished by comparison of the HPLC peak corresponding to the natively folded species in the reaction with the integral of the same peak of a known quantity of very pure ($\geq$95% by analytical HPLC) reference folded peptide. In selenocysteine-containing peptides, the non-selenocysteine form was used for calculation of apparent folding yield. Interestingly,

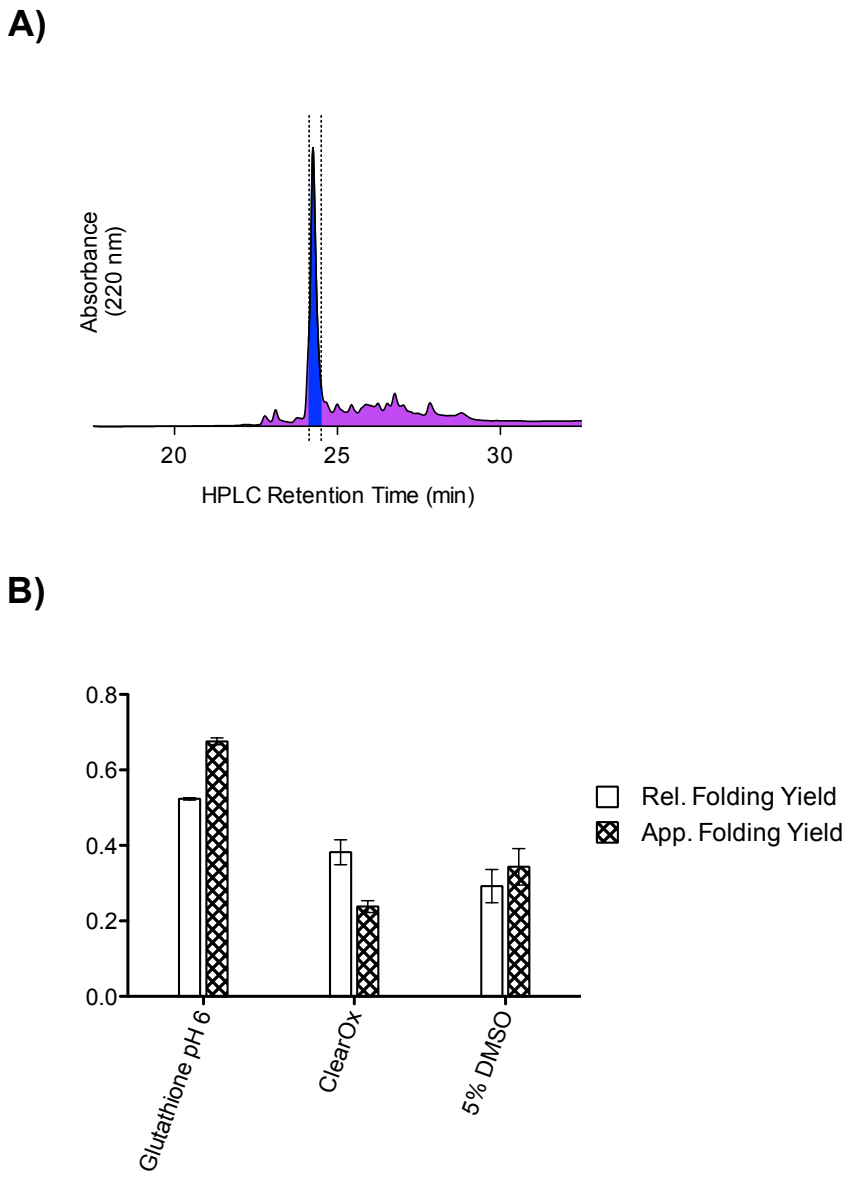


Figure 2.4. Methods for determining folding yields used in this study. **A)** Sample analytical HPLC separation of a quenched SIIIA folding reaction. Relative folding yield is calculated by taking the peak area of the natively folded peak (blue), and dividing by the total area (sum of blue and purple). To calculate apparent folding yield, the peak area of a folded and purified sample (with the same total amount of peptide as the folding reaction) is used as a reference, representing 100% yield. The apparent folding yield is then the area of the peak in the folding reaction (blue) divided by the average area of reference peaks (mean of 3 runs). **B)** A comparison between relative folding yield and apparent folding yield of selected conditions for oxidation of SIIIA.

these are quite distinct measurements of folding efficiency (Figure 2.4).

Both relative folding yield and apparent folding yield are useful measurements of folding efficiency; their difference lies in what is used as the theoretical value for 100% conversion. The canonical measurement (relative folding yield) takes this value as the integral of the entire folding area, hence assuming that all oxidation products are represented in the HPLC chromatogram (e.g., there is no aggregation). Conversely, apparent folding yield uses a separate HPLC chromatogram of the same peptide that has been folded, purified and quantified (for example, by spectrophotometry), hence comparing with an amount of the folded species that would represent 100% conversion. While neither is equivalent to the gold standard of a mass-to-mass calculation of yield, we use apparent folding yield throughout in order to meaningfully compare between oxidation methods with different propensities for aggregation. Also of note, the proposed intermediate for the polymer-supported oxidation method that we employ is resin-bound, and hence is not present in the HPLC separation. Canonical folding yield data is available in the Supplemental Information.

2.2.3 Experimental Procedures

2.2.3.1 Reduced Peptides

Peptides were synthesized on a single-channel automated peptide synthesizer using Rink Amide resin and standard Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry. All cysteine residues were trityl protected. Selenocysteine residues were 4-methoxybenzyl protected. Peptides without selenocysteine residues were cleaved from resin with Reagent K (trifluoroacetic acid (TFA) : phenol : water : thioanisole : 1,2-ethanedithiol; 33 : 2 : 2 : 2 : 1 by volume). Selenocysteine-containing peptides were cleaved from resin with enriched Reagent K† (TFA : phenol : water : thioanisole : 2,2'-dithiobis(5-nitropyridine); 178 : 14 : 5 : 1.6 : 1 by volume). DTNP is included in order to remove the *p*-methoxybenzyl protecting group from selenocysteine, which produces an adduct of selenocysteine with 2-thio-5-nitro-pyridine [39, 40]. Cleaved peptides were filtered and precipitated with methyl *tert*-butyl ether. Selenocysteine-containing peptides were then treated with 50 mM DTT (in 0.1 M Tris, pH 7.5 with 0.1 mM EDTA) for 2 hours, in order to remove the 2-thio-5-nitro-pyridine adduct. The diselenide-containing peptide is recovered following HPLC purification [41]. Peptides without selenocysteine were also purified by HPLC on a Waters 600 chromatograph with a dual-wavelength absorbance detector using Vydac C_{18} semipreparative and preparative columns. All SIIIA HPLCs (linear and folded) were run with a linear gradient from 4.5% to 31.5% acetonitrile, with 0.9% change per minute, maintaining 0.1% TFA throughout. Masses of the linear/diselenide peptides were validated by electrospray ionization mass

spectrometry.

2.2.3.2 Solution-Phase Oxidation

2.2.3.2.1 Glutathione based oxidation. Oxidative folding with glutathione-based redox buffers was accomplished with varying ratios of oxidized and reduced glutathione. Typical reaction conditions were 0.1 M Tris, 0.1 mM EDTA, 0.5-2 mM oxidized glutathione (GSSG), 0-2 mM reduced glutathione (GSH) and 20 μM μ-SIIIA (final concentration), pH 7.5. The pH dependence was investigated using MOPS in lieu of Tris/EDTA as the buffer (see Supplemental Information). The buffered glutathione redox system was first established, and the peptide was added as the final component. Reactions were set up in triplicate on an analytical scale (14 nmol of peptide), and were allowed to proceed for 16 hours. The temperature dependence of solution-phase folding was evaluated at 4° C, 22° C and 37° C using a mixture of 1 mM GSSG, 1 mM GSH. In order to manipulate product ratios in a solution-phase oxidation reaction, folding temperature and time are commonly optimized in parallel. The effect of pH was tested using three buffer systems. Folding at acidic pH was assayed with 0.1 M sodium acetate/acetic acid buffer (pH 4); mid-range pH's were considered with 0.1 M MOPS (pH 6-8); 0.1 M sodium carbonate/bicarbonate buffer (pH 10) was used to assess folding under basic conditions. Quenching by acidification was accomplished by addition of formic acid to a final concentration of 8% (final pH $\approx$ 1.78).

2.2.3.2.2 Dimethyl sulfoxide mediated oxidation. Folding reactions contained 0.1 M Tris, pH 7.5, 0.1 mM EDTA, 5-10% dimethyl sulfoxide and 20 μM μ-SIIIA (final concentrations). Because the reaction begins when the peptide contacts DMSO, the buffered, dilute DMSO was prepared, and the peptide was added as the final component. Reactions were set up in triplicate on an analytical scale (14 nmol of peptide). After 16 hours, formic acid was added to a final concentration of 8% to quench the folding reaction (final pH $\approx$ 1.83). A final concentration of 10% dimethyl sulfoxide was found to be most effective to fold μ-SIIIA.

A summary of various folding conditions investigated and the resulting apparent folding yields is shown in Figure 2.2, with representative HPLC chromatograms of the folding reactions are shown in Figure 2.3.

2.2.3.3 Polymer-Supported Oxidation

ClearOx is a commercially available resin that has been preloaded with Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid), DTNB), immobilizing the peptide by reaction of the peptide thiolate with the on-resin disulfide, producing a peptide-resin disulfide intermediate

[2]. Because ClearOx is preloaded, the low-loading conditions are preestablished to favor intramolecular disulfide bonding at millimolar peptide concentrations (indicated by the apparent folding yield, shown in Figure 2.2).

2.2.3.3.3 ClearOx based folding. Twelve molar-equivalents of resin per disulfide (1 equivalent = 2.768 μg resin/nmol peptide/disulfide to be oxidized) were swelled in dichloromethane for 30 minutes. For folding 30 nmoles of SIIIA, 2.93 mg of ClearOx was employed. Twenty to thirty milligrams of ClearOx was prepared in a single vessel, using 300-700 μl of each of the swelling and washing solvents. To begin washing, excess dichloromethane was pipetted off, and replaced with dimethylformamide. The same process was repeated with centrifugation to wash the resin with methanol, then 50% acetonitrile in water and finally 50% acetonitrile, 0.05 M MOPS, pH 7. To ensure that the resin will remain swelled, the solution is never entirely removed, enforcing the low-loading conditions that allow efficient folding while the solution is changed from dichloromethane to one that is more amenable to oxidative folding of peptides. The swelled and washed resin was then divided into separate vessels for each folding reaction. Fifty percent acetonitrile, 0.05 M MOPS, pH 7.0 was used to redissolve μ-SIIIA to 3 mM. To minimize folding and aggregation due to air-mediated oxidation, the peptides were resuspended as shortly before use as possible. Following resin preparation, all excess wash solution was pipetted off and replaced with solution containing 3 mM μ-SIIIA. Reactions were done in triplicate on an analytical (30 nmol) scale. Folding reactions proceeded for 1 hour in 50% acetonitrile in 0.1 M MOPS, pH 7.0. Reactions were quenched by acidification, accomplished by 100-fold dilution with 0.1% TFA in water (final pH $\approx$ 2.17). Air oxidation and disulfide shuffling are possible unless the reaction is acidified; for preparative work, this can also be done by addition of formic acid in order to maintain a small reaction volume. The peptide solution was then removed from the polymer-bound oxidant following centrifugation.

The amount of cosolvent present during ClearOx-based oxidative folding may affect both the relative folding yield and apparent folding yield; while this effect tends to vary based on the oxidation substrate, 50% acetonitrile is sufficiently generalizable to be a starting point for subsequent optimization. Optimization remains necessary, although the variables to be optimized are rather limited to the amounts of cosolvent and ClearOx resin in the folding reaction (see Supplemental Information on folding of ω-conotoxin GVIA). Noteworthy, further optimization of ClearOx-assisted folding may involve changing to more acidic pH, as described previously [2, 9].

Folding efficiency with ClearOx was found to be optimal at room temperature, although

the folding time varies substantially among peptides (see Supplemental Information, Tables 2.S3, 2.S4, 2.S5). Due to generally shorter folding times and higher peptide concentrations (which provides for preferable acidification conditions), ClearOx has significant potential for higher throughput peptide folding.

2.2.3.4 Selenocysteine

Selenocysteine residues have recently been shown to simplify oxidative folding and facilitate disulfide mapping in peptides [10, 15, 41, 42, 43, 44, 45]. The replacement of both cysteines of a single disulfide bridge with selenocysteines allows one bridge to be formed directly after deprotection and treatment with DTT [41]. The strong reduction potential of selenium prevails over encoded sequence information [43], and consequently the preformed bridge reduces the number of possible disulfide isomers. The preformed bridge serves as both a combinatorial constraint with respect to the number of possible folding isomers as well as an entropic constraint in the oxidative folding process, both of which benefit the folding efficiency. The combinatorial effect remains constant; however, the entropic effect depends on the location of the diselenide bridge [15, 41]. The use of diselenide bridges has been shown to complement solution-phase oxidative folding [15, 41, 42]. Since replacement of the first disulfide of μ-SIIIA with a diselenide gave the largest increase in relative folding yield [15], this substrate, namely [C3U;C13U] SIIIA (Sec-SIIIA) was used in the folding experiments.

2.2.3.4.4 Folding of selenocysteine-containing peptides. In solution, selenocysteine-containing peptides were folded under the same conditions as the same peptide without the diselenide. Briefly, folding reactions contained 0.1 M Tris, pH 7.5, 0.1 mM EDTA, 20 μM peptide, and either 1 mM each of GSSG and GSH or 5-10% DMSO. With ClearOx, the resin preparation was identical; however, the amount of resin used was decreased in order to maintain 12 equivalents of resin per disulfide bond to be closed in the folding reaction. Apparent folding yield calculations of selenocysteine-containing peptides were accomplished by normalizing to a known quantity of the folded form of the same peptide without the diselenide proxy. Figure 2.5 shows the effect of the diselenide, with analytical HPLC traces of parallel-foldings of SIIIA with Sec-SIIIA, both at 3 mM peptide concentration using ClearOx resin as the oxidant. The effect on folding efficiency is summarized in Figures 2.2 and 2.S1. When combined with a similar observation that the ClearOx-assisted oxidation of a selenopeptide analogue of GVIA (Table 2.S5 and Figure 2.S5) also produced relatively high folding yields, these results suggest that the above strategy could be considered as a more generalized folding protocol.

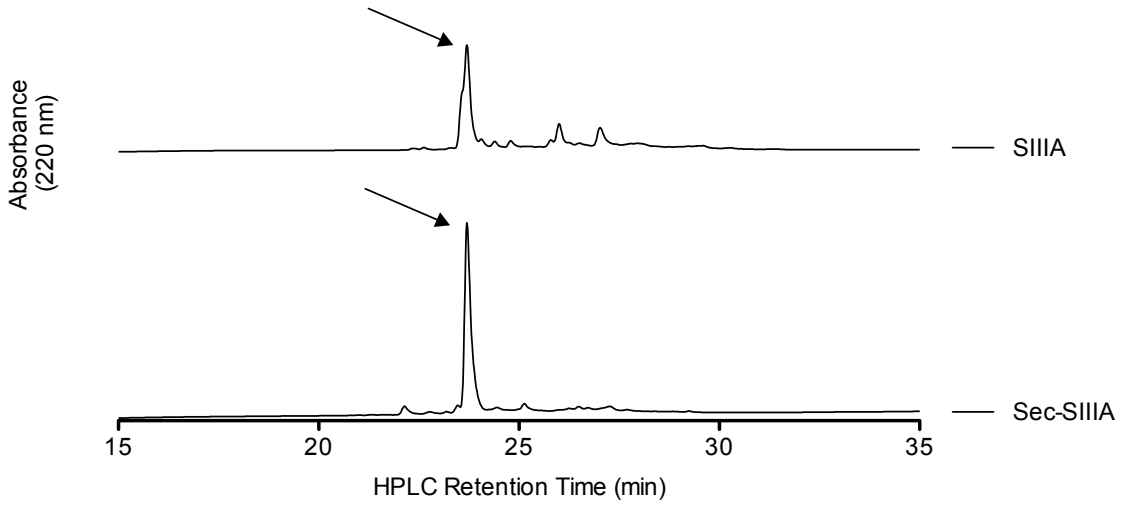


Figure 2.5. Analytical HPLC traces of SIIIA (top) and Sec-SIIIA (bottom) foldings with ClearOx. Arrows indicate the native species. Both reactions were carried out at 3 mM peptide using 12 eq. of ClearOx in 50% acetonitrile in 0.1 M MOPS, pH 7.0 for 1 hour, and were quenched by 100-fold dilution with 0.1% TFA.

2.2.4 Summary and Limitations

Optimized solution phase folding often produces the highest reaction yields; however, there are intrinsic costs to oxidative folding in the solution phase. The requirement for high dilution conditions presents problems for large-scale and high-throughput applications. However, with relatively minor costs, these problems may largely be overcome using resin-bound oxidants, such as ClearOx. For example, to oxidize 1 gram of μ-SIIIA using solution phase methods would require a reaction volume of 22.7 liters. However, to oxidize 1 gram of μ-SIIIA using ClearOx would require a reaction volume of only 151 ml. The optimization of folding time with ClearOx is significantly more reliable when separate reactions for each timepoint are carried out because swelled ClearOx resin is not readily pipetted.

Selenocysteine proxies for disulfides can significantly increase the relative folding yield, and simplify the HPLC folding pattern, allowing for easier purification. While the apparent folding yield for some peptides is not necessarily significantly increased with a diselenide proxy (Figures 2.2 and 2.S1), the benefit to apparent folding yield for other selenopeptides can be quite substantial (see Supplemental Information, Figure 2.S5, Table 2.S5). We suggest that ClearOx-assisted oxidation of selenopeptides (Figure 2.2 and 2.S5) should be considered for higher throughput folding of cysteine-rich peptides.

Given the unprecedented molecular diversity of cysteine-rich peptides with respect to both disulfide scaffolds and primary amino acid sequences, a word of caution about the generality of the recommended folding methods should be made. The selection of oxidation method and extent of optimization of the reaction conditions will depend on the peptides' structure, quantities required for subsequent studies, as well as on the purpose of studying the peptides. When smaller quantities (less than 1-2 mg) of a peptide are needed, it is recommended to start with optimization of the solution phase folding methods. For folding larger quantities of a peptide, the resin-bound oxidants should be considered. For folding difficult peptides, or for higher throughput oxidative folding of multiple peptides or combinatorial libraries, the use of selenopeptide analogues and resin-bound oxidants is suggested.

2.2.5 Acknowledgements

This work was supported by N.I.H. Program Project Grant GM 48677, and the Willard Eccles Fellowship Fund (A.M.S.). We thank Drs. Robert Shackmann and Scott Endicott from the DNA/Peptide Synthesis Core Facility at the University of Utah for their help in peptide synthesis. We also thank Peptides International, Inc. for kindly providing us with ClearOx^{TM} resin (ClearOx^{TM} is a trademark of Peptides International, Inc.).

2.3 Supplemental Information

2.3.1 Scope of Supplement

This supplemental information describes additional folding experiments, addressed only tangentially in the body of the paper. To expand optimization of the folding methods for μ-SIIIA, the effects of both temperature and pH in solution phase folding reactions were evaluated. For the μ-SIIIA folding reactions with ClearOx, effects of temperature, cosolvent and peptide concentrations were also studied.

To generalize the folding methods described here, we carried out comparable oxidation experiments using two additional conotoxins containing three disulfide bridges, namely μ-KIIIA and ω-GVIA. μ-Conotoxin KIIIA shares the same cysteine framework as SIIIA, whereas GVIA is an ω-conotoxin from *C. geographus* that differs in the cysteine framework from that of SIIIA and KIIIA. Lastly, we present results on ClearOx-assisted oxidation of a selenopeptide analogue of GVIA, confirming the effectiveness of this strategy to produce the folded species with relatively high efficiency.

2.3.2 SIIIA

Figure 2.S1 shows the relative (canonical) folding yields for SIIIA under identical conditions to those shown in Figure 2.2 in the main paper. Although the values for relative folding yields (Figure 2.S1) and apparent folding yields (Figure 2.2) differ from each other, the trends within a single oxidation system (glutathione, DMSO, ClearOx) are consistent enough to be used for optimization (of that oxidation system). Error bars show standard deviation from 3 analytical HPLC runs, based on calculation of the integral of the absorbance at 220 nm. Unless otherwise stated, "Glutathione" indicates 1 mM oxidized (GSSG) and 1 mM reduced (GSH) glutathione. All reactions with glutathione or DMSO proceeded for 16 hours. ClearOx folding reactions were done with 12 equivalents of ClearOx resin, and proceeded for 1 hour.

The pH dependence of SIIIA folding using solution-phase methods is shown in Table 2.S1. All reactions in Table 2.S1 were carried out with 20 μM peptide concentration, in the presence of 1 mM of each of oxidized and reduced glutathione, and were allowed to fold for 16 hours before the oxidation was quenched by addition of formic acid. To sample a sufficiently large range of pH's, it was necessary to use more than one buffer system. The reactions at pH 6-8 were buffered with 0.1 M MOPS. While the typical solution-phase oxidation conditions buffer with 0.1 M Tris, pH 7.5, MOPS was selected instead because MOPS buffers effectively in more acidic solution. A sodium acetate/acetic acid buffer system was used for folding reactions pH 4, and a sodium carbonate/bicarbonate buffer system was

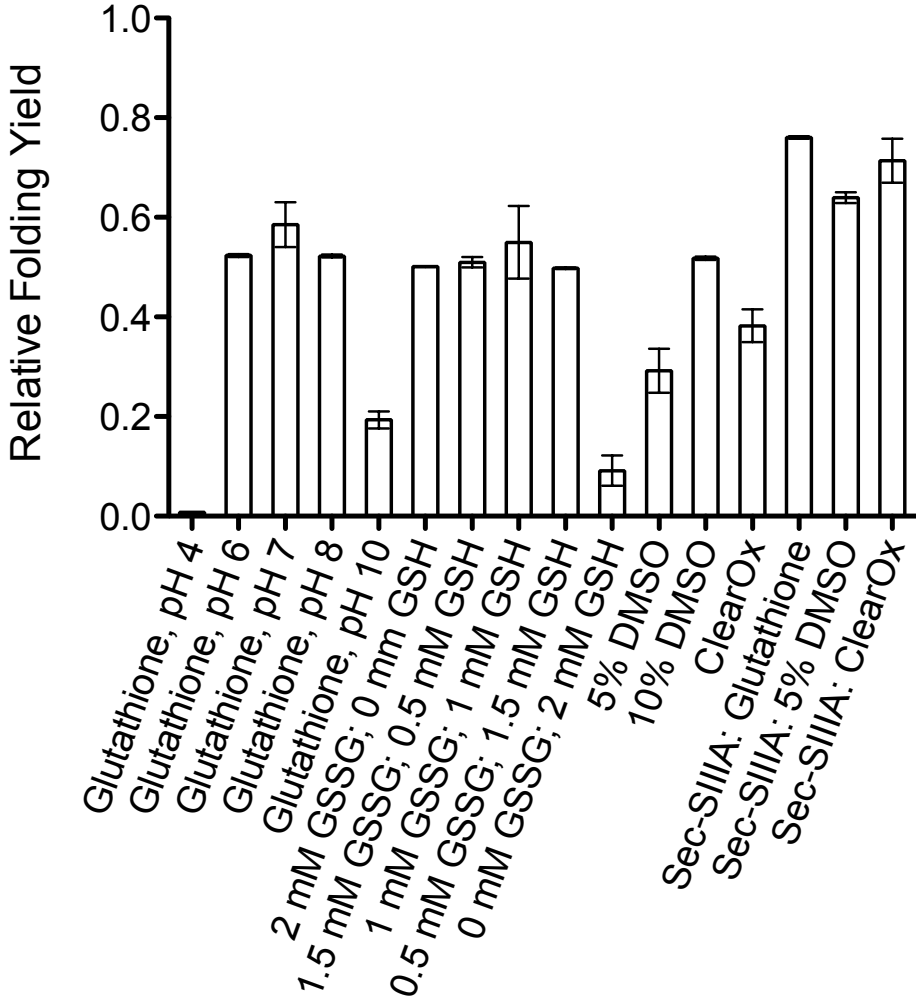


Figure 2.S1. Relative (canonical) folding yield (fraction of integral due to natively folded species)

Table 2.S1. $^{\square}$ The "relative folding yield" is expressed as the mean value of 3 experiments, taking the fraction of the total integral that is attributable to the native foldamer. $^{\wp}$ The standard deviation of the relative folding yield, based on 3 experiments. $^{\nabla}$ The "apparent folding yield" is expressed as the mean value of 3 experiments, taking the integral of the region of the folding chromatogram attributable to the native foldamer, and normalizing with a known quantity of very pure folded peptide (see Figure 2.4 in text). $^{\circledast}$ The standard deviation of the apparent folding yield, based on 3 experiments.

Table S1: Role of Buffer pH in SIIIA Folding Yield				
pH	Rel. Yield$^{\square}$	$\sigma_Y{}^{\wp}$	App. Yield$^{\nabla}$	$\sigma_A{}^{\circledast}$
4.0	0.78%	0.01%	0.94%	0.02%
6.0	52.34%	0.28%	67.59%	0.89%
7.0	58.52%	4.50%	66.27%	2.41%
8.0	52.25%	0.31%	65.43%	2.08%
10.0	19.34%	1.70%	25.11%	0.19%

employed for folding at pH 10. It is apparent that the folding efficiency is highest when the peptide is folded in the pH range 6-8.

Temperature dependence of folding in solution phase oxidation is shown in Table 2.S2. All reactions in Table 2.S2 were carried out at 20 μM peptide concentration. Reactions were buffered with 0.1 M Tris, 0.1 mM EDTA, pH 7.5, using 1 mM GSSG and 1 mM GSH, and were quenched by addition of formic acid to a final concentration of 8%. No major changes in the folding yields were observed when the reactions were carried out at 4° C and 22° C, whereas folding at higher temperature (37° C) produced less favorable results. Noteworthy, oxidative folding at low temperatures (0-4° C) might be recommended for peptides that exhibit hydrophobic character. For example, folding yields for very hydrophobic δ-conotoxins were significantly improved when peptides were oxidized in an ice bath [46].

The effect of temperature on the folding of SIIIA using ClearOx as the oxidant are shown in Table 2.S3. All reactions in Table 2.S3 were carried out at 3 mM peptide concentration, in triplicate. Mean and standard deviation are presented for both relative folding yield and apparent folding yield. All reactions were run in 50% acetonitrile in 0.1 M MOPS, pH 7.0. Only 1 timepoint (1 hour) was analyzed for ClearOx-based reactions. Reactions were quenched by 100-fold dilution with 0.1% TFA in water. Interestingly, the temperature dependence observed for ClearOx-based oxidation was much more significant than that of solution phase oxidation.

The solvent and peptide concentrations were also found to be important for ClearOx-assisted folding of SIIIA (Figure 2.S2). All reactions were buffered with 0.05 M MOPS, pH 7, were carried out with 12 eq. of ClearOx, and were allowed to proceed for 1 hour. The peptide concentration and amount of acetonitrile were varied, and are shown next to the corresponding HPLC trace. The possibility of air oxidation precluded analysis of longer timepoints with the aqueous and dilute conditions, and a single folding time was selected for comparison across all conditions. For dilute reactions (20 μM SIIIA), quenching was accomplished by addition of formic acid to a final concentration of 8%. For concentrated folding reactions, 100-fold dilution with 0.1% TFA was used to quench the folding reactions.

At high peptide concentrations (3 mM) and 50% acetonitrile, SIIIA folds quite effectively (top trace in Figure 2.S2). However, when the folding solution was made to parallel those used to fold with glutathione or DMSO, using dilute peptide (20 μM) in aqueous solution, ClearOx failed to fold SIIIA (bottom trace in Figure 2.S2). Based on this finding, we suggest that these two factors (cosolvent and peptide concentration) are very important factors to

Table 2.S2. Temperature dependence of SIIIA folding in solution

Table S2: Temperature Dependence of Folding of SIIIA in Solution					
Temp.	Time	Rel. Yield	σ_Y	App. Yield	σ_A
4 ℃	30 min	58.23%	0.76%	66.63%	2.65%
	60 min	56.18%	2.48%	68.98%	0.65%
	180 min	54.49%	0.40%	69.89%	1.57%
	16 hrs.	51.94%	0.81%	68.65%	1.29%
22 ℃	30 min	53.89%	3.43%	64.99%	0.82%
	60 min	52.14%	1.56%	65.62%	0.31%
	180 min	51.86%	4.27%	58.76%	0.35%
	16 hrs.	48.25%	2.56%	58.26%	1.33%
37 ℃	30 min	41.43%	0.23%	51.70%	1.04%
	60 min	40.97%	0.08%	50.17%	0.05%
	180 min	41.24%	0.57%	48.59%	1.94%
	16 hrs.	40.52%	0.42%	45.26%	0.77%

Table 2.S3. Temperature dependence of SIIIA folding with ClearOx

Table S3: Temperature Dependence of Folding of SIIIA with ClearOx				
Temp.	Rel. Yield	σ_Y	App. Yield	σ_A
4 ℃	19.67%	9.43%	6.22%	0.41%
22 ℃	38.21%	3.29%	23.80%	1.59%
37 ℃	26.89%	10.33%	22.20%	2.09%

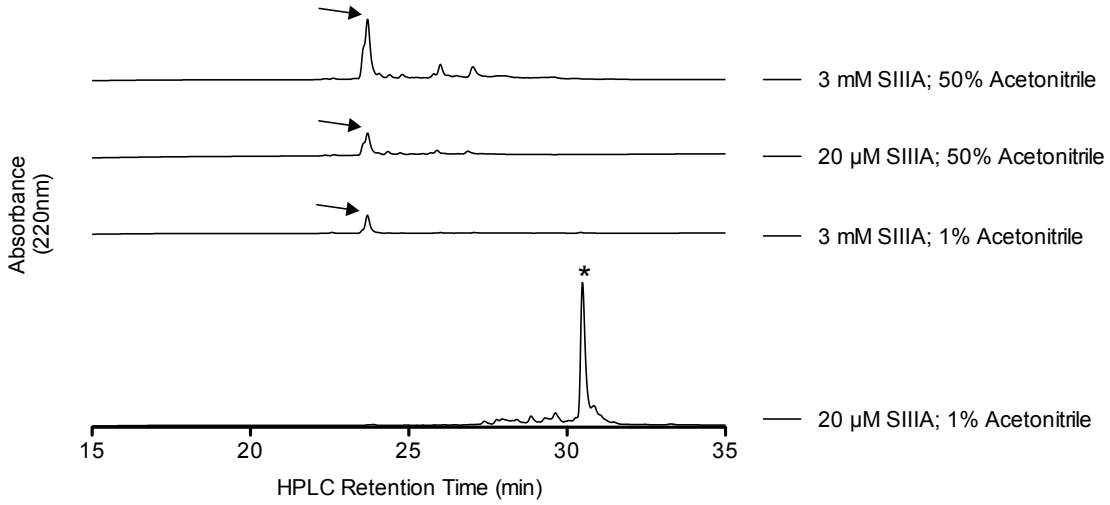


Figure 2.S2. Solvent System and Peptide Concentration in Folding SIIIA with ClearOx. Arrows indicate the correctly folded species, and * indicates the linear form.

consider for optimization of ClearOx-assisted folding.

2.3.3 KIIIA

μ-Conotoxin KIIIA has a similar structure to μ-SIIIA, and has been the subject of several folding studies [47, 48, 49]. The reaction scheme for the folding is shown in Figure 2.S3. Three folding conditions for KIIIA were examined: glutathione-, DMSO- and ClearOx-assisted oxidation. Because the conditions for solution-phase folding of KIIIA were established previously [48], these conditions were applied here for glutathione-mediated oxidation. Table 2.S4 shows the folding yields of KIIIA folding conditions. All folding reactions proceeded at room temperature. Mean and standard deviation are presented for both relative folding yield and apparent folding yield (reactions were done in triplicate). The glutathione folding reactions were done with 1 mM GSSG, 1 mM GSH. Glutathione- and DMSO-mediated folding were performed at 20 μM μ-KIIIA, and were buffered with 0.1 M Tris, 0.1 mM EDTA, pH 7.5; solution-phase folding was quenched by addition of formic acid to a final concentration of 8%. ClearOx folding reactions were done with 3 mM μ-KIIIA, 12 eq. ClearOx, and were buffered with 0.1 M MOPS, pH 7; quenching was accomplished by 100-fold dilution in 0.1% TFA in water.

Folding results for KIIIA suggest that the ClearOx-based method was more effective than the glutathione-assisted oxidation. Interestingly, this result is somewhat opposite for SIIIA, emphasizing the importance of screening various strategies and conditions for optimization of folding, even when working with related peptides.

2.3.4 GVIA

In order to compare solution-phase and polymer-supported folding methods using a peptide that differs in its cysteine scaffold from SIIIA and KIIIA, we also studied the folding of ω-conotoxin GVIA and its selenopeptide analogue (reaction scheme shown in Figure 2.S4). Folding of GVIA has been investigated previously [41, 50, 51, 52]. The structure of GVIA is shown in Figure 2.S4; this conotoxin is known to fold to an ICK (inhibitory-cystine-knot) motif. Noteworthy, GVIA is a potent blocker of voltage-gated calcium channels.

The folding reactions shown in Table 2.S5 were done as follows. Glutathione and DMSO mediated oxidation reactions were buffered with 0.1 M Tris, 0.1 mM EDTA, pH 7.5, and contained 20 μM peptide. The oxidant was either 5% DMSO or a redox buffer of 1 mM GSSG, 2 mM GSH, as previously established [50]. Solution phase reactions were quenched by addition of formic acid to a final concentration of 8%. ClearOx reactions were run in 50% acetonitrile in 0.1 M MOPS, pH 6.0, with a peptide concentration of 3 mM and 12 eq.

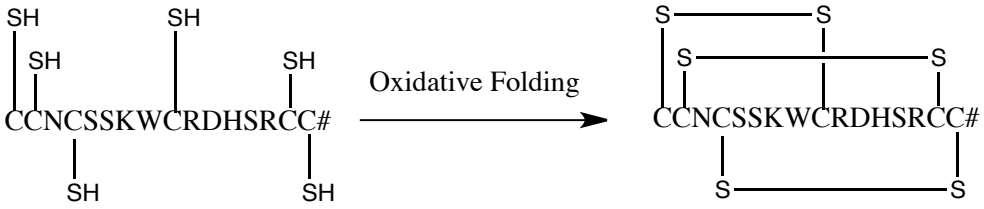


Figure 2.S3. Reaction scheme for KIIIA (# = amidated C-terminus)

Table 2.S4. Folding Yields of KIIIA; values shown are Mean ± Std. Dev.

Table S4: Oxidative Folding of KIIIA in Various Conditions					
Condition	Time	Rel. Yield	σ_Y	App. Yield	σ_A
Glutathione	30 min	42.04%	0.40%	36.19%	1.15%
	60 min	42.22%	0.27%	35.50%	1.58%
5% DMSO	30 min	2.30%	0.14%	2.04%	0.17%
	60 min	3.49%	0.16%	3.12%	0.19%
	16 hrs.	44.64%	1.76%	30.84%	0.41%
12 eq. ClearOx	30 min	60.37%	8.93%	33.04%	3.06%
	60 min	63.79%	1.95%	38.42%	0.61%

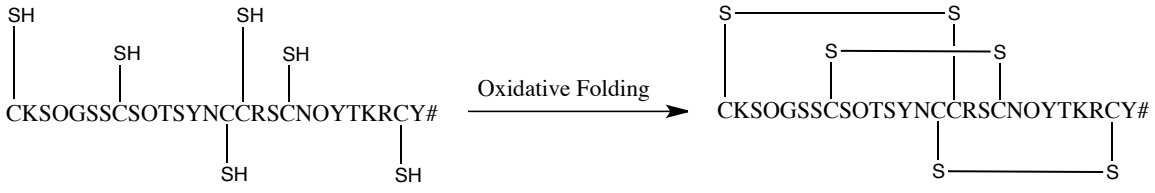


Figure 2.S4. Reaction scheme for GVIA (# = amidated C-terminus, O = Hydroxyproline)

Table 2.S5. Folding yields of GVIA; values shown are mean ± std. dev.

Table S5: Oxidative Folding of GVIA in Various Conditions					
Condition	Time	Rel. Yield	σ_Y	App. Recov. Yield	σ_A
Glutathione	60 min	38.88%	0.30%	40.02%	1.53%
	90 min	43.92%	5.89%	40.76%	2.00%
5% DMSO	16 hrs.	25.97%	3.58%	25.48%	1.84%
	24 hrs.	30.97%	3.53%	26.09%	5.24%
12 eq. ClearOx	180 min	28.24%	1.40%	21.25%	2.70%
24 eq. ClearOx	180 min	33.29%	0.89%	24.53%	0.98%
Effect of a Diselenide Proxy (Sec-GVIA)					
Glutathione	90 min	68.87%	0.21%	82.40%	2.65%
24 eq. ClearOx	180 min	60.87%	2.41%	53.54%	2.44%

of ClearOx, and were quenched by 100-fold dilution with 0.1% TFA in water. All folding reactions proceeded at room temperature, in triplicate. Mean and standard deviation are shown for both relative folding yield and apparent folding yield. The apparent folding yield for both GVIA and Sec-GVIA was calculated based on the integral of the folded peak of a known amount of very pure folded GVIA ($\geq$95% by analytical HPLC).

Figure 2.S5 shows 3 analytical HPLC traces of GVIA folding. Arrows indicate the correctly folded species, as determined previously [41]. The top trace shows the folding with ClearOx, based on standard conditions with 12 eq. ClearOx per disulfide bond, folding at 3 mM GVIA in 50% acetonitrile, 0.05 M MOPS, pH 6 for 3 hours. The center trace shows glutathione-folded GVIA, using the conditions that were established previously for folding GVIA [50], with 1 mM oxidized glutathione, 2 mM reduced glutathione, 0.1 M Tris, pH 7.5 with 0.1 mM EDTA, 20 μM GVIA, and proceeded for 90 minutes at room temperature. The bottom trace shows a selenocysteine-containing GVIA folding trace, where folding was accomplished at 3 mM Sec-GVIA with 24 eq. ClearOx resin, folding in 50% acetonitrile, 0.05 M MOPS, pH 6 for 3 hours. ClearOx folding reactions were quenched by 100-fold dilution with 0.1% TFA in water. The glutathione folding reaction was quenched by addition of formic acid.

Consistent with the previously published results, GVIA folded with relatively high yields using solution-phase oxidation (Table 2.S5 and Figure 2.S5). However, folding with ClearOx proved to be ineffective, as indicated in Table 2.S5 and Figure 2.S5. Altered levels of cosolvent did not ameliorate the folding efficiency. While increased levels of ClearOx do not substantially benefit the folding efficiency of GVIA, it did increase the reproducability of the folding reactions (see standard deviations, Table 2.S5). Lowering the pH of the ClearOx-assisted folding reactions to 6 (rather than 7, as used for the other peptides) was found to slightly improve the folding of GVIA.

To determine whether the simplification of the oxidative folding using a diselenide proxy could ameliorate the folding efficiency, selenocysteine-containing GVIA was tested in a folding reaction with ClearOx. All 3 selenopeptide analogues have been previously studied with solution-phase folding [41], and [C8U;C19U] GVIA (Sec-GVIA) showed the largest increase in relative folding yield. Therefore, [C8U;C19U] GVIA was selected for further folding experiments. Figure 2.S5 shows an analytical HPLC trace of ClearOx-mediated folding of Sec-GVIA, from which it is evident that the major product is the correctly folded species. The apparent improvement of the folding yields of the selenopeptide analogue of GVIA with ClearOx is similar to that observed for SIIIA (Figure 2.5), suggesting that this

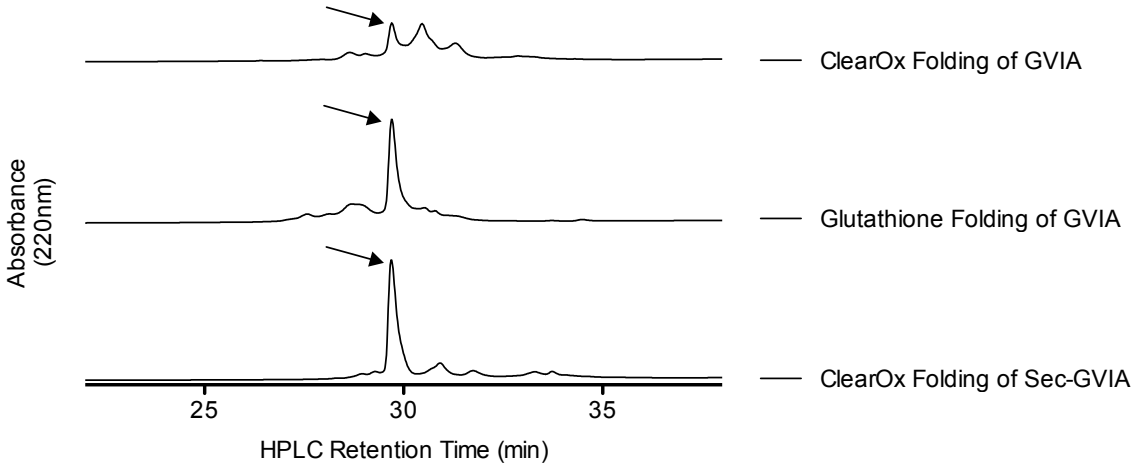


Figure 2.S5. Effect of a diselenide proxy on efficiency of ClearOx-mediated GVIA folding

oxidation strategy may be of more general use for cysteine-rich peptides varying in their disulfide scaffolds. However, more studies are needed to validate this generalized oxidative folding method.

2.4 Conclusion

The use of the pseudodilution effect to increase the concentration of the folding reaction dramatically aids high-throughput folding efforts, as the sheer fluid volume required to perform high-throughput oxidations at typical peptide concentrations (20 μM) is prohibitive. Additionally, the use of 50% acetonitrile allows the use of a fairly high level of organic solvent, which allows much better solvation of hydrophobic peptides, but remains miscible with water. These serve to make the method presented much more generalizable than anything presented previously; however, these methods should still be improved further before being employed for high-throughput oxidation because efficient high-throughput work requires sufficiently high Relative Folding Yields to also streamline chromatographic purification, not merely the oxidation step.

One aspect of this work deserves further comment: the role that diselenide bridges were playing in these reactions was likely underestimated. As detailed in the following chapter, diselenide bridges are capable of functioning as both inter- and intramolecular catalysts for the formation of disulfides, which was not known at the time of publication. Consequently, the diselenide bridges likely altered the course of the oxidation, as well as allowing redox chemistry to take place in the highly-concentrated solution phase. As a result, I cannot necessarily ensure that all disulfide formation employed the pseudodilution effect, and this may have led to aggregation, whereby modulating the Apparent Folding Yield. Conversely, allowing some disulfide formation to occur in the solution phase may also have decreased the extent of adduct formation with the immobilized Ellman's Reagent. Nonetheless, the method presented allows efficient formation of disulfide bridges using a generalized protocol that is much more broadly applicable than anything employed at the time.

In order to consider the relative efficiency of each of the presented methods, Figure 2.S6 shows HPLC chromatograms of folding with glutathione, ClearOx and air-oxidation, as well as folding of diselenide-containing SIIIA with ClearOx and air-oxidation. The absorbance measurements were not scaled, indicating the relative efficacy of each oxidation method. It is evident from Figure 2.S6 that the effect of incorporating a diselenide bridge provides a universal benefit, regardless of which method is used to form the disulfide bridges.

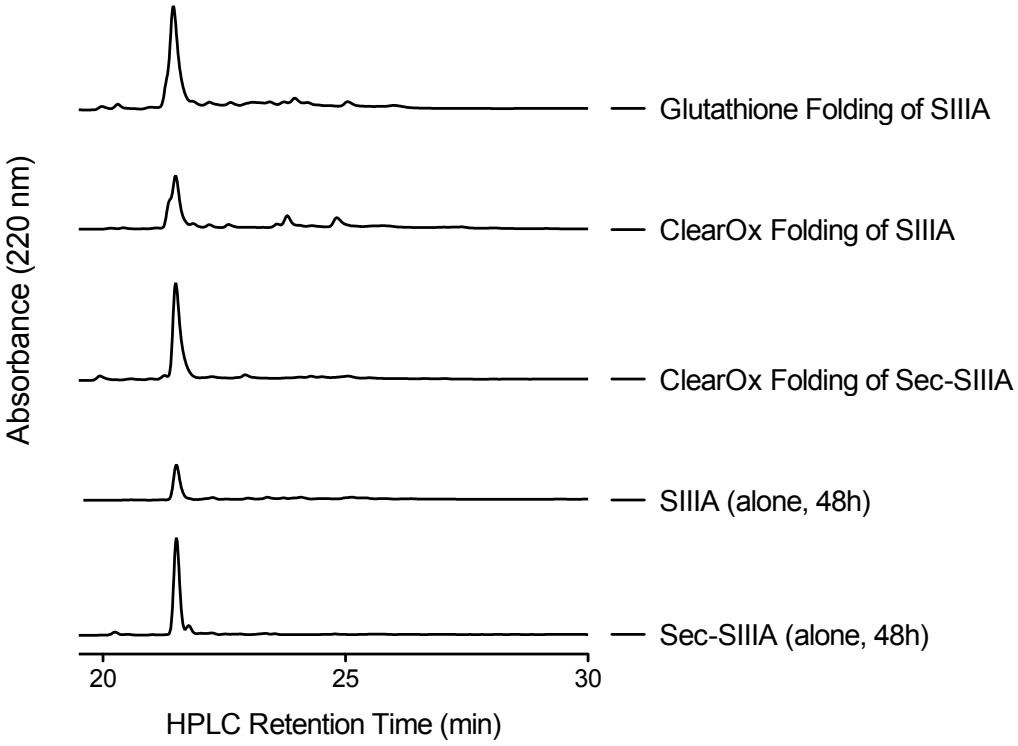


Figure 2.S6. HPLC Chromatograms of SIIIA folding conditions. Ready comparison can be made between SIIIA folding with glutathione, ClearOx and air-oxidation; diselenide-containing SIIIA ([C3U;C13U] SIIIA) is also shown for folding with ClearOx and air-oxidation. This compiles the data from Figures 2.3, 2.5, as well as the diselenide catalysis data from Chapter 3.

2.5 References

[1] Annis I., Chen L., Barany G. Novel solid-phase reagents for facile formation of intramolecular disulfide bridges in peptides under mild conditions. *J. Am. Chem. Soc.* 1998; **120**(29): 7226–7238. doi:10.1021/ja981111p.

[2] Darlak K., Wiegandt Long D., Czerwinski A., Darlak M., Valenzuela F., Spatola A.F., Barany G. Facile preparation of disulfide-bridged peptides using the polymer-supported oxidant CLEAR-OX^{TM}. *J. Pept. Res.* 2004; **63**(3): 303–312. doi:10.1111/j.1399-3011.2004.00153.x.

[3] Green B.R., Bulaj G. Oxidative folding of conotoxins in immobilized systems. *Protein Pept. Lett.* 2006; **13**(1): 67–70. doi:10.2174/092986606774502162.

[4] Besse D., Siedler F., Diercks T., Kessler H., Moroder L. The redox potential of selenocystine in unconstrained cyclic peptides. *Angew. Chem. Int. Ed. Engl.* 1997; **36**(8): 883–885. doi:10.1002/anie.199708831.

[5] King G.F. Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther.* 2011; **11**(11): 1469–1484. doi:10.1517/14712598.2011.621940.

[6] Annis I., Hargittai B., Barany G. Disulfide bond formation in peptides. *Methods Enzymol.* 1997; **289**: 198–221. doi:10.1016/S0076-6879(97)89049-0.

[7] Bulaj G., Walewska A. Oxidative folding of single-stranded disulfide-rich peptides. In *Oxidative Folding of Peptides and Proteins* (Eds. Buchner J., Moroder L.). RSC Biomolecular Sciences, 2009, pp. 279–286. doi:10.1039/9781847559265.

[8] Bulaj G., Olivera B.M. Folding of conotoxins: formation of the native disulfide bridges during chemical synthesis and biosynthesis of *Conus* peptides. *Antioxid. Redox Signal.* 2008; **10**(1): 141–155. doi:10.1089/ars.2007.1856.

[9] Zhang J., Diamond S., Arvedson T., Sasu B.J., Miranda L.P. Oxidative folding of hepcidin at acidic pH. *Biopolymers* 2010; **94**(2): 257–264. doi:10.1002/bip.21383.

[10] Muttenthaler M., Nevin S.T., Grishin A.A., Ngo S.T., Choy P.T., Daly N.L., Hu S.H., Armishaw C.J., Wang C.I., Lewis R.J., Martin J.L., Noakes P.G., Craik D.J., Adams D.J., Alewood P.F. Solving the α-conotoxin folding problem: efficient selenium-directed on-resin generation of more potent and stable nicotinic acetylcholine receptor antagonists. *J. Am. Chem. Soc.* 2010; **132**(10): 3514–3522. doi:10.1021/ja910602h.

[11] Armishaw C.J., Alewood P.F. Conotoxins as research tools and drug leads. *Curr. Protein Pept. Sci.* 2005; **6**(3): 221–240. doi:10.2174/1389203054065437.

[12] Han T.S., Teichert R.W., Olivera B.M., Bulaj G. *Conus* venoms - a rich source of peptide-based therapeutics. *Curr. Pharm. Des.* 2008; **14**(24): 2462–2479. doi: 10.2174/138161208785777469.

[13] Livett B.G., Gayler K.R., Khalil Z. Drugs from the sea: conopeptides as potential therapeutics. *Curr. Med. Chem.* 2004; **11**(13): 1715–1723. doi:10.2174/0929867043364928.

[14] Halai R., Craik D.J. Conotoxins: natural product drug leads. *Nat. Prod. Rep.* 2009; **26**(4): 526–536. doi:10.1039/b819311h.

[15] Walewska A., Zhang M.M., Skalicky J.J., Yoshikami D., Olivera B.M., Bulaj G. Integrated oxidative folding of cysteine/selenocysteine containing peptides: improving chemical synthesis of conotoxins. *Angew. Chem. Int. Ed. Engl.* 2009; **48**(12): 2221–2224. doi:10.1002/anie.200806085.

[16] Lopez-Vera E., Walewska A., Skalicky J.J., Olivera B.M., Bulaj G. Role of hydroxyprolines in the *in vitro* oxidative folding and biological activity of conotoxins. *Biochemistry* 2008; **47**(6): 1741–1751. doi:10.1021/bi701934m.

[17] Buczek P., Buczek O., Bulaj G. Total chemical synthesis and oxidative folding of δ-conotoxin PVIA containing an N-terminal propeptide. *Biopolymers* 2005; **80**(1): 50–57. doi:10.1002/bip.20211.

[18] Nielsen J.S., Buczek P., Bulaj G. Cosolvent-assisted oxidative folding of a bicyclic α-conotoxin ImI. *J. Pept. Sci.* 2004; **10**(5): 249–256. doi:10.1002/psc.531.

[19] Altamirano M.M., García C., Possani L.D., Fersht A.R. Oxidative refolding chromatography: folding of the scorpion toxin Cn5. *Nat. Biotechnol.* 1999; **17**(2): 187–191. doi:10.1038/6192.

[20] M'Barek S., Fajloun Z., Cestèle S., Devaux C., Mansuelle P., Mosbah A., Jouirou B., Mantegazza M., Van Rietschoten J., El Ayeb M., Rochat H., Sabatier J.M., Sampieri F. First chemical synthesis of a scorpion α-toxin affecting sodium channels: the Aah I toxin of *Androctonus australis hector*. *J. Pept. Sci.* 2004; **10**(11): 666–677. doi:10.1002/psc.582.

[21] Estrada G., Garcia B.I., Schiavon E., Ortiz E., Cestele S., Wanke E., Possani L.D., Corzo G. Four disulfide-bridged scorpion beta neurotoxin CssII: heterologous expression and proper folding *in vitro*. *Biochim. Biophys. Acta* 2007; **1770**(8): 1161–1168. doi:10.1016/j.bbagen.2007.04.006.

[22] Čemažar M., Zahariev S., Lopez J.J., Carugo O., Jones J.A., Hore P.J., Pongor S. Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues. *Proc. Natl. Acad. Sci. U.S.A.* 2003; **100**(10): 5754–5759. doi:10.1073/pnas.2225470100.

[23] Bach R.D., Dmitrenko O., Thorpe C. Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J. Org. Chem.* 2008; **73**(1): 12–21. doi:10.1021/jo702051f.

[24] Snow J., Finley J., Friedman M. Oxidation of sulfhydryl groups to disulfides by sulfoxides. *Biochem. Biophys. Res. Comm.* 1975; **64**(1): 441–447. doi:10.1016/0006-291X(75)90272-7.

[25] Bulaj G., Kortemme T., Goldenberg D.P. Ionization-reactivity relationships for cysteine thiols in polypeptides. *Biochemistry* 1998; **37**(25): 8965–8972. doi:10.1021/bi973101r.

[26] Kubo S., Chino N., Kimura T., Sakakibara S. Oxidative folding of ω-conotoxin MVIIC: effects of temperature and salt. *Biopolymers* 1996; **38**(6): 733–744. doi:10.1002/(SICI)1097-0282(199606)38:6<733::AID-BIP5>3.0.CO;2-S.

[27] DeLa Cruz R., Whitby F.G., Buczek O., Bulaj G. Detergent-assisted oxidative folding of δ-conotoxins. *J. Pept. Res.* 2003; **61**(4): 202–212. doi:10.1034/j.1399-3011.2003.t01-1-00048.x.

[28] Miloslavina A.A., Leipold E., Kijas M., Stark A., Heinemann S.H., Imhof D. A room temperature ionic liquid as convenient solvent for the oxidative folding of conopeptides. *J. Pept. Sci.* 2009; **15**(2): 72–77. doi:10.1002/psc.1106.

[29] Wang C.Z., Zhang H., Jiang H., Lu W., Zhao Z.Q., Chi C.W. A novel conotoxin from *Conus striatus*, μ-SIIIA, selectively blocking rat tetrodotoxin-resistant sodium channels. *Toxicon* 2006; **47**(1): 122–132. doi:10.1016/j.toxicon.2005.10.008.

[30] Bulaj G., West P.J., Garrett J.E., Watkins M., Marsh M., Zhang M.M., Norton R.S., Smith B.J., Yoshikami D., Olivera B.M. Novel conotoxins from *Conus striatus* and *Conus kinoshitai* selectively block TTX-resistant sodium channels. *Biochemistry* 2005; **44**(19): 7259–7265. doi:10.1021/bi0473408.

[31] Woycechowsky K.J., Wittrup K.D., Raines R.T. A small-molecule catalyst of protein folding *in vitro* and *in vivo*. *Chem. Biol.* 1999; **6**(12): 871–879. doi:10.1016/S1074-5521(00)80006-X.

[32] Beld J., Woycechowsky K.J., Hilvert D. Catalysis of oxidative protein folding by small-molecule diselenides. *Biochemistry* 2008; **47**(27): 6985–6987. doi:10.1021/bi8008906.

[33] Saxena V.P., Wetlaufer D.B. Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme. *Biochemistry* 1970; **9**(25): 5015–5023. doi:10.1021/bi00827a028.

[34] Hwang C., Sinskey A.J., Lodish H.F. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* 1992; **257**(5076): 1496–1502. ISSN 0036-8075. doi:10.1126/science.1523409.

[35] Galande A.K., Weissleder R., Tung C.H. An effective method of on-resin disulfide bond formation in peptides. *J. Comb. Chem.* 2005; **7**(2): 174–177. doi:10.1021/cc049839r.

[36] Galanis A.S., Albericio F., Grøtli M. Enhanced microwave-assisted method for on-bead disulfide bond formation: synthesis of α-conotoxin MII. *Biopolymers* 2009; **92**(1): 23–34. doi:10.1002/bip.21116.

[37] Green B.R., Catlin P., Zhang M.M., Fiedler B., Bayudan W., Morrison A., Norton R.S., Smith B.J., Yoshikami D., Olivera B.M., Bulaj G. Conotoxins containing nonnatural backbone spacers: cladistic-based design, chemical synthesis, and improved analgesic activity. *Chem. Biol.* 2007; **14**(4): 399–407. doi:10.1016/j.chembiol.2007.02.009.

[38] Yao S., Zhang M.M., Yoshikami D., Azam L., Olivera B.M., Bulaj G., Norton R.S. Structure, dynamics, and selectivity of the sodium channel blocker μ-conotoxin SIIIA. *Biochemistry* 2008; **47**(41): 10940–10949. doi:10.1021/bi801010u.

[39] Harris K.M., Flemer S., Hondal R.J. Studies on deprotection of cysteine and selenocysteine side-chain protecting groups. *J. Pept. Sci.* 2007; **13**(2): 81–93. doi: 10.1002/psc.795.

[40] Schroll A.L., Hondal R.J. Further development of new deprotection chemistry for cysteine and selenocysteine side chain protecting groups. *Adv. Exp. Med. Biol.* 2009; **611**: 135–136.

[41] Gowd K., Yarotskyy V., Elmslie K., Skalicky J., Olivera B., Bulaj G. Site-specific effects of diselenide bridges on the oxidative folding of a cystine knot

peptide, ω-selenoconotoxin GVIA. *Biochemistry* 2010; **49**(12): 2741–2752. doi: 10.1021/bi902137c.

[42] Armishaw C.J., Daly N.L., Nevin S.T., Adams D.J., Craik D.J., Alewood P.F. α-selenoconotoxins, a new class of potent α_7 neuronal nicotinic receptor antagonists. *J. Biol. Chem.* 2006; **281**(20): 14136–14143. doi:10.1074/jbc.M512419200.

[43] Fiori S., Pegoraro S., Rudolph-Böhner S., Cramer J., Moroder L. Synthesis and conformational analysis of apamin analogues with natural and non-natural cystine/selenocystine connectivities. *Biopolymers* 2000; **53**(7): 550–564. doi: 10.1002/(SICI)1097-0282(200006)53:7<550::AID-BIP3>3.0.CO;2-O.

[44] Han T.S., Zhang M.M., Gowd K.H., Walewska A., Yoshikami D., Olivera B.M., Bulaj G. Disulfide-depleted selenoconopeptides: simplified oxidative folding of cysteine-rich peptides. *ACS Med. Chem. Lett.* 2010; **1**(4): 140–144. doi:10.1021/ml900017q.

[45] Mobli M., de Araújo A.D., Lambert L.K., Pierens G.K., Windley M.J., Nicholson G.M., Alewood P.F., King G.F. Direct visualization of disulfide bonds through diselenide proxies using ^{77}Se NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* 2009; **48**(49): 9312–9314. doi:10.1002/anie.200905206.

[46] Bulaj G., DeLaCruz R., Azimi-Zonooz A., West P., Watkins M., Yoshikami D., Olivera B.M. δ-Conotoxin structure/function through a cladistic analysis. *Biochemistry* 2001; **40**(44): 13201–13208. doi:10.1021/bi010683a.

[47] Bulaj G. Formation of disulfide bonds in proteins and peptides. *Biotechnol. Adv.* 2005; **23**(1): 87–92. doi:10.1016/j.biotechadv.2004.09.002.

[48] Zhang M.M., Green B.R., Catlin P., Fiedler B., Azam L., Chadwick A., Terlau H., McArthur J.R., French R.J., Gulyas J., Rivier J.E., Smith B.J., Norton R.S., Olivera B.M., Yoshikami D., Bulaj G. Structure/function characterization of μ-conotoxin KIIIA, an analgesic, nearly irreversible blocker of mammalian neuronal sodium channels. *J. Biol. Chem.* 2007; **282**(42): 30699–30706. doi:10.1074/jbc.M704616200.

[49] Han T.S., Zhang M.M., Walewska A., Gruszczynski P., Robertson C.R., Cheatham T.E., Yoshikami D., Olivera B.M., Bulaj G. Structurally minimized μ-conotoxin analogues as sodium channel blockers: implications for designing conopeptide-based therapeutics. *ChemMedChem* 2009; **4**(3): 406–414. doi:10.1002/cmdc.200800292.

[50] Price-Carter M., Gray W.R., Goldenberg D.P. Folding of ω-conotoxins. 1. Efficient disulfide-coupled folding of mature sequences *in vitro*. *Biochemistry* 1996; **35**(48): 15537–15546. doi:10.1021/bi961574c.

[51] Flinn J.P., Murphy R. Effect of buffer system on the refolding of synthetic ω-conotoxin GVIA. *Lett. Pep. Sci.* 1996; **3**(3): 113–116. doi:10.1007/BF00132972.

[52] Flinn J.P., Pallaghy P.K., Lew M.J., Murphy R., Angus J.A., Norton R.S. Roles of key functional groups in ω-conotoxin GVIA: synthesis, structure and functional assay of selected peptide analogues. *Eur. J. Biochem.* 1999; **262**(2): 447–455. doi: 10.1046/j.1432-1327.1999.00383.x.

CHAPTER 3

AN UNEXPECTED FUNCTION FOR DISELENIDE BRIDGES

The majority of this chapter is a paper that was published in Angewandte Chemie International Edition in English in 2012.[1] As with the preceding chapter, I will provide a brief introduction and concluding remarks, with a particular emphasis on the aspects of this work that are not discussed in the included paper.

3.1 Introduction

At its core, this work combines two bodies of literature that had been developed independently of one another. The first body of literature used intramolecular diselenide bonds as an alternative to a disulfide bond [1], which took advantage of the difference in redox potentials between disulfide bonds, diselenide bonds and mixed sulfur-selenium bonds [2]. Because the diselenide is so strongly favored over mixed sulfur-selenium bonds, it was believed to be a static bridge that structurally constrained the peptide during the remainder of the folding process.

The second body of literature used diselenide bridges as catalysts for the formation of disulfides [3, 4, 5], and noted that the free selenol is never observed because the diselenide is reformed very quickly by reacting with molecular oxygen in solution.

This work linked these two bodies of literature, and observed that intramolecular diselenide bridges are actually potent intramolecular catalysts for the formation of the remaining disulfide bridges in a peptide. We propose a mechanism, and present evidence showing that diselenide bridges are thermodynamically quite stable, but are kinetically very reactive.

---

[1](a) A.M. Steiner, K.J. Woycechowsky, B.M. Olivera, G. Bulaj, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 5580. doi: 10.1002/anie.201200062; (b) A.M. Steiner, K.J. Woycechowsky, B.M. Olivera, G. Bulaj, *Angew. Chem.* **2012**, *124*, 5678-5682. doi: 10.1002/ange.201200062.

In doing this research, I had a few observations that were not included in the paper. First, this reaction is quite temperature sensitive, and required extensive thermal control to ensure reproducibility of rate constants. Second, the level of oxygen dissolved in solution is a critical factor; while we were unable to remove enough oxygen to completely stop the reaction, the dependence of the reaction rate on dissolved oxygen was readily visible as a function of the amount of other solute. With more total solute (for instance, Urea or NaCl), there is less oxygen dissolved in solution, which causes the reaction to slow. Conversely, when the folding mixture is pipetted directly onto dried peptide, the reaction rate increases dramatically.

Additionally, there is a short discussion regarding whether this is intramolecular or intermolecular catalysis. There are several lines of evidence suggesting that it is intramolecular catalysis, but one intriguing result is that when a folding substrate with a diselenide is combined with a folding substrate without a diselenide, the reaction rate appears to be the average of the two, suggesting that there is crossover between the two substrates. I believe that this occurs because the only oxidative catalyst present is the diselenide, which can act both intramolecularly and intermolecularly; the intramolecular reaction should predominate, but the intermolecular reaction occurs when peptide that lacks the intramolecular catalyst is mixed with peptide bearing the intramolecular catalyst.

3.2 Reagentless Oxidative Folding of Disulfide-Rich Peptides Catalyzed by an Intramolecular Diselenide

(a) A.M. Steiner, K.J. Woycechowsky, B.M. Olivera, G. Bulaj, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 5580. doi: 10.1002/anie.201200062; (b) A.M. Steiner, K.J. Woycechowsky, B.M. Olivera, G. Bulaj, *Angew. Chem.* **2012**, *124*, 5678-5682. doi: 10.1002/ange.201200062.

3.2.1 Abstract

In cysteine-rich peptides, diselenides can be used as a proxy for disulfide bridges as the energetic preference for Se–Se bonds over mixed Se–S bonds simplifies folding (see Figure 3.1). An intramolecular diselenide bond efficiently catalyzes the oxidative folding of selenopeptide analogues of conotoxins, and serves as a reagentless method to accelerate formation of various native disulfide bridging patterns.

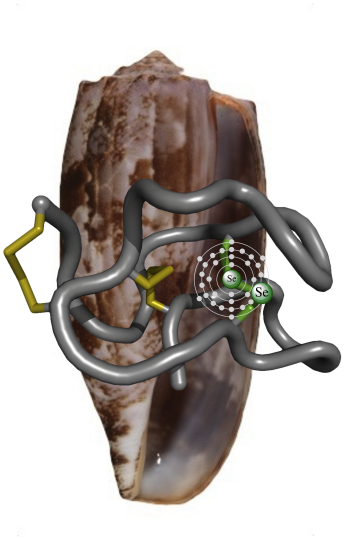


Figure 3.1. Graphical abstract

3.2.2 Main Text

Disulfide bonds are critical to stabilizing the native structure of many peptides and proteins. These crosslinks are formed during oxidative folding, and machinery to assist formation of the correct cysteine pairings is used by all walks of life, from prokaryotes to humans [6, 7]. However, in the chemical synthesis of disulfide-containing biomolecules alternative methods are employed to arrive at the native disulfide connectivity [8], and usually involves redox buffers and requires peptide-specific optimization [9].

Selenocysteine is often referred to as nature's 21st proteinogenic amino acid [10], and it can be readily incorporated into chemically synthesized peptides. Since diselenide bonds have greater thermodynamic stability than disulfide bonds, selenocysteine pairings can prevail over sequence-encoded folding information in short, cysteine-rich peptides [11]. By serving as a proxy for a disulfide bridge, selenocysteine incorporation into peptides may reduce the complexity of oxidative folding [9, 11, 12, 13, 14]. Consequently, diselenide bonds in peptides have been assumed to act as a fixed structural element during oxidative folding as a result of their high stability.

Diselenide bridges were incorporated into only a handful of cysteine-rich peptides including two-disulfide-containing apamin [11, 15] and α conotoxins [12, 16], as well as three-disulfide-bridged conotoxins [1, 9, 13, 17]. Diselenides have also been useful to facilitate disulfide mapping [1, 18]. While several groups have noted the enhanced kinetics

of disulfide formation in selenopeptides [13, 16], to date, the rate enhancement has been assumed to be due to conformational effects.

Diselenide additives have been found to catalyze oxidative protein folding for a wide range of substrates [4, 5]. With these additives, catalysis stems from rapid regeneration of the diselenide by reaction of the selenols formed in situ with atmospheric oxygen, even under (mildly) reducing conditions [9, 13]. In contrast, thiols will not readily oxidize without a catalyst, such as a metal ion. Here we tested a hypothesis that if a diselenide can act as an intermolecular catalyst of cysteine oxidation, then disulfide-rich selenopeptides should be able to catalyze their own aerobic oxidative folding.

Conotoxins, which are derived from the venom of marine gastropods, represent an important target for improvements in oxidative folding efficiency. The venom of each species of *Conus* has a large and distinct array of disulfide-rich peptides, many of which are used as neuropharmacological tools [19]. While conotoxins have broad applicability, their oxidative folding is often quite challenging [9, 13, 16, 20, 21], thus making conotoxins an attractive system for studying the mechanisms of oxidative folding.

To examine whether diselenides enable the reagentless oxidative folding of cysteine-rich peptides, we selected previously characterized selenocysteine-containing analogues of ω-conotoxin GVIA and μ-conotoxin SIIIA (Figure 3.2) [9, 13].

Comparisons of the folding behavior for GVIA and [C8U;C19U] GVIA in Tris-buffered water–without any oxidative reagents–are shown in Figures 3.3 a-c. Both k_{ox} (the rate of formation of the first disulfide) and k_{native} (the rate of appearance of the native connectivity) show significant enhancement in the presence of a diselenide, thus indicating its role in directing both oxidation and isomerization. To test the generality of this effect, similar experiments were also carried out using μ-conotoxin SIIIA and [C4U;C19U] SIIIA (Figures 3.3 d-f). The significant increase in both k_{ox} (by approximately 12- to 20-fold) and k_{native} (by approximately 34- to 175-fold) indicated that diselenides are efficient oxidative catalysts independent of the disulfide scaffold. Details of the kinetic analysis are presented in the Supporting Information.

The reagent-free formation of a disulfide bond by an intramolecular diselenide likely involves three chemical processes. First, the diselenide transfers an oxidizing equivalent to a pair of thiols to form a disulfide. Second, the selenol(ate) acts as a nucleophile to promote disulfide isomerization. Third, the selenol(ate)s are reoxidized by molecular oxygen in solution.

Both copper(II) and selenocystine are known to catalyze the oxidation of thiols to disulfides [3, 22]. For both disulfide scaffolds considered, the addition of these oxidative

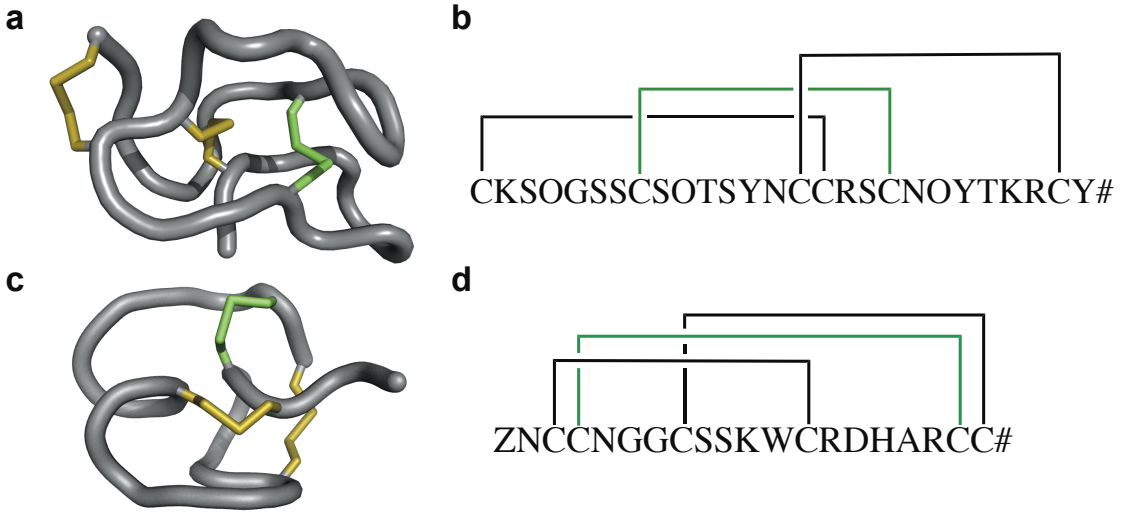


Figure 3.2. Folded products of peptides used in this study. a) Three-dimensional structure of GVIA (PDBID: 2CCO). b) Schematic representation of GVIA, where the bridge indicated in green was replaced with a diselenide in [C8U;C19U] GVIA. c) Three-dimensional structure of SIIIA (BMRB #20023). d) Schematic representation of SIIIA; the bridge indicated in green was replaced with selenocysteines in [C4U;C19U] SIIIA. Z = pyroglutamate, O = hydroxyproline, and # = C-terminal amidation.

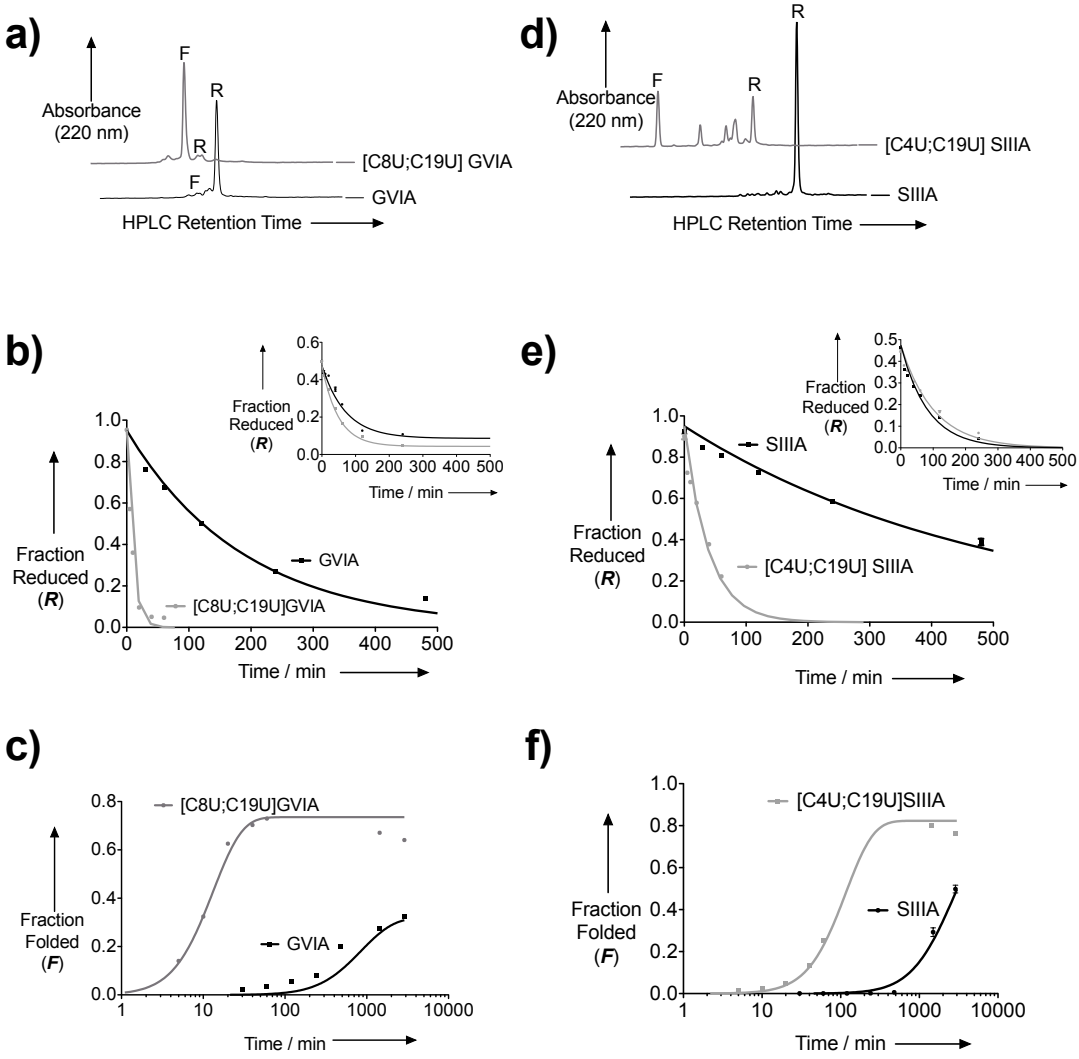


Figure 3.3. Folding kinetics of GVIA, SIIIA, and their diselenide-containing analogues. HPLC chromatograms after 1 hour of folding for comparison of GVIA (a) and SIIIA (b) to their respective diselenide-containing analogues. R = reduced form, F = natively folded form. Full folding timecourses are available in Figure 3.S4 of the Supporting Information. c,d) Kinetic traces of the disappearance of the starting material (bearing either six thiols or four thiols and a diselenide), with insets showing the depletion of each reduced peptide (maintaining colors) when the all-thiol and diselenide-containing peptides are folded in a 1:1 mixture. e,f) Kinetic traces of the formation of the natively bridged peptide. Equivalent traces using a linear timescale are shown in Figures 3.S2 and 3.S3 of the Supporting Information.

catalysts afforded a substantial increase in the rate of oxidation (k_{ox}). Similarly, the rate of native formation (k_{native}) was also enhanced, albeit to a much lower extent. A discussion of the error values for the rate of native formation (k_{native}) is provided in the Supporting Information. However, for diselenide-containing peptides, neither oxidative catalyst provided a significant change in either the rate of oxidation or the rate of native formation (Figure 3.4). Interestingly, the folding of wild-type GVIA in the presence of a selenocystine additive remains less efficient than that of reagentless folding with [C8U;C19U] GVIA.

The kinetic advantage of an intramolecular diselenide over added intermolecular diselenide in GVIA suggests a unimolecular mechanism. However, further analysis suggests that the reaction shows both inter- and intramolecular character. On one hand, the initial oxidation of [C8U;C19U] GVIA appears concentration-independent across a 10-fold range in concentrations (see the Supporting Information). However, a 1:1 mixture of diselenide-containing peptide with its all-thiol counterpart suggests that intermolecular catalysis also contributes (Figures 3.3 b and e, insets). This mixture of inter- and intramolecular catalysis is particularly pronounced for SIIIA, where the initial oxidation step (k_{ox}) was slightly faster with added selenocystine (intermolecular catalysis), but isomerizations and subsequent oxidations leading to the native disulfide connectivity (k_{native}) were slightly faster with an intramolecular diselenide (see the Supporting Information). Further discussion is provided in the Supporting Information.

To consider the role of diselenide placement within a peptide scaffold, we compared [C4U;C19U] SIIIA with [C3U ;C13U] SIIIA and [C8U ;C20U] SIIIA, which were chosen for this analysis because the reduced and folded forms are more readily separable by HPLC than the comparable analogues of GVIA. This data is shown in Figure 3.4 and in the Supporting Information, and demonstrates that the presence of a diselenide–regardless of its position in the peptide–greatly accelerates folding. While always beneficial, the position of the diselenide within the peptide does lead to some variation in the folding rate enhancement. This variation could result from either conformation-induced orientation effects or differential stability and activation energies conferred to intermediate conformations by the location of the diselenide bridge.

Thus, in addition to directing oxidative folding towards the native disulfide connectivity, intramolecular diselenide bridges also serve as catalysts for the formation of diuslfide bridges. A diselenide bridge in a peptide is therefore able to catalyze disulfide formation using dissolved oxygen as the stoichiometric oxidant, thus effectively enabling reagentless oxida-

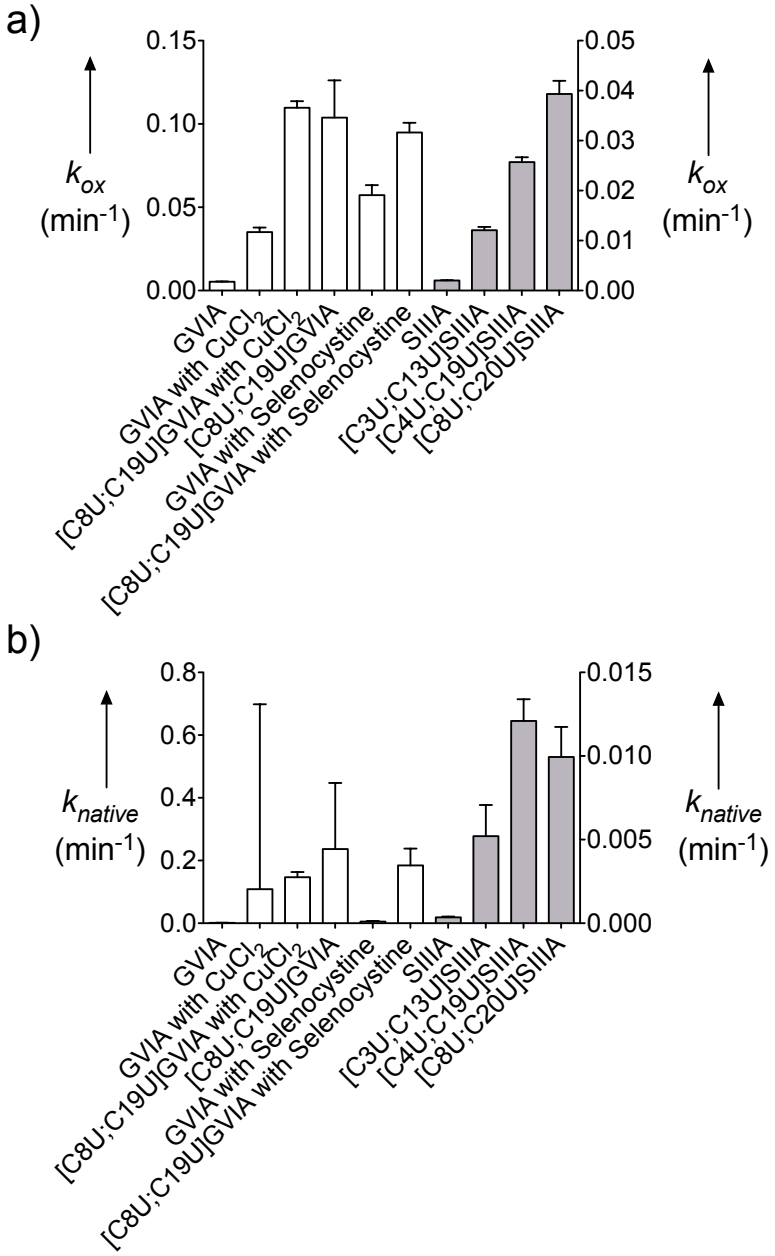


Figure 3.4. Rate constant comparison. a) Rate constants for initial oxidation. b) Rate constants for the formation of the native connectivity. On each graph are compared GVIA, GVIA with 100 nM cupric chloride, [C8U;C19U] GVIA with 100 nM cupric chloride, [C8U;C19U] GVIA, GVIA with equimolar selenocystine, [C8U;C19U] GVIA with equimolar selenocystine, as well as variation resulting from the location of the diselenide within SIIIA. GVIA and GVIA analogues (white bars) are graphed on the left axis; SIIIA and SIIIA analogues (grey bars) are graphed on the right axis. In all cases, n = 3 and values shown are mean ± SEM.

tion. Using only molecular oxygen as the oxidant further simplifies the oxidative folding of cysteine-rich peptides. We envision the use of reagentless oxidation of selenopeptides, whether in solution or on-resin [9] will be effective, even for peptides that are difficult to synthesize [21].

In the folding of selenopeptides, Se–S bonds are likely to be necessary intermediates for catalysis of intramolecular electron-transfer reactions and isomerization steps. Examples of stable Se–S bonds are rare [23]. Non-native Se–S bonds have also been shown to exist in synthetic α-conotoxin and oxytocin analogues bearing a native pair of selenocysteines [12, 16, 24]. The high reactivity of selenocysteine is likely the reason why nature has selected disulfide bridges over diselenide bridges to form covalent crosslinks in polypeptides. To this end, our findings warrant a cautious approach towards exploring selenopeptides as drug leads. However, this does not preclude the added value of selenocysteine in the chemical synthesis of cysteine-rich peptides, as well as in manipulating oxidative folding pathways [25].

Our work shows that incorporation of diselenides into conotoxins may significantly improve their folding rates because the oxidative folding of the peptide is catalyzed by the peptide itself. In combination with previous reports, in which diselenides showed superior effects on simplifying and improving folding yields and disulfide mapping, this approach provides a significant advantage to chemical syntheses of cysteine-rich peptides. Since molecular biology techniques have already revealed thousands of bioactive peptides from plants, snails, spiders, scorpions, and vertebrate animals, selenopeptide technologies will accelerate discoveries of novel research tools and biotherapeutics based on cysteine-rich peptides.

3.2.3 Experimental Section

Detailed methods are provided in the Supporting Information. Briefly, peptides were synthesized with standard Fmoc chemistry; all cysteine residues were trityl protected, and all selenocysteine residues were 4-methoxybenzyl protected. After cleavage from the resin, selenocysteine-containing peptides were treated with DTT. Peptides were then purified by HPLC prior to oxidative folding.

Unless otherwise indicated, all oxidations were air-mediated, using molecular oxygen as the only oxidant. Dried, reduced peptides were redissolved in 5% acetonitrile, 0.01% TFA. The peptide solution was then added to the folding mixture, which contained (final) 0.1 M Tris, pH 7.5, 0.1 mM EDTA, as well as any other folding additives for the specific condition. Folding reactions proceeded at 20℃. Timepoints were analyzed by analytical HPLC.

HPLC peak integration was used to quantify the extent of peptide folding for each timepoint. DynaFit was used to derive kinetic parameters by fitting the data to a model of consecutive irreversible first-order reactions [26]. Details of the kinetic analysis are presented in the Supporting Information.

3.2.4 Acknowledgements

We thank Professors David Blair, David Goldenberg, and Donald Hilvert for helpful discussions and comments on the manuscript. We also acknowledge the assistance of the Peptide Synthesis and Mass Spectrometry Core Facilities at the University of Utah. This work was supported by the NIH Program Project Grant GM 48677.

3.3 Supporting Information

3.3.1 Abbreviations

TFA, trifluoroacetic acid; DTNP, 2,2'-dithiobis(5-nitropyridine); MTBE, methyl *tert*-butyl ether; HPLC, high performance liquid chromatography.

3.3.2 Results and Discussion

3.3.2.1 Inter- vs. Intramolecularity

Several routes were considered to determine whether the catalysis of disulfide formation by an intramolecular diselenide was an inter- or intramolecular effect. Interestingly, oxidation reactions appear to show both intermolecular and intramolecular character.

k_{ox} measures the rate of formation of the first disulfide bond, as this was derived from the depletion of the reduced peptide (bearing either six thiols *or* four thiols and a diselenide), and therefore represents the first oxidation step. k_{ox} is used exclusively to determine the inter- or intramolecularity, as k_{native} also represents isomerization events as well as oxidation (see Data Analysis, 3.3.3.3).

The invariance of k_{ox} over a 10-fold range in concentration suggests an intramolecular mechanism (Figure 3.S1). However, if the reaction is exclusively intramolecular catalysis, then a 1:1 mixture of unmodified peptide with its diselenide-containing analog would give two distinct rates of depletion of the reduced peptide, both being unchanged from their individual rates when folded separately. However, this was not the case, as is shown in Figure 3.3b and e (and insets), for GVIA and SIIIA, respectively.

Thus, the catalysis of disulfide bond formation by an intramolecular diselenide demonstrates both inter- and intramolecular character. However, the co-folding reactions (combining an unmodified peptide with its diselenide-containing analog) may contain an artificial

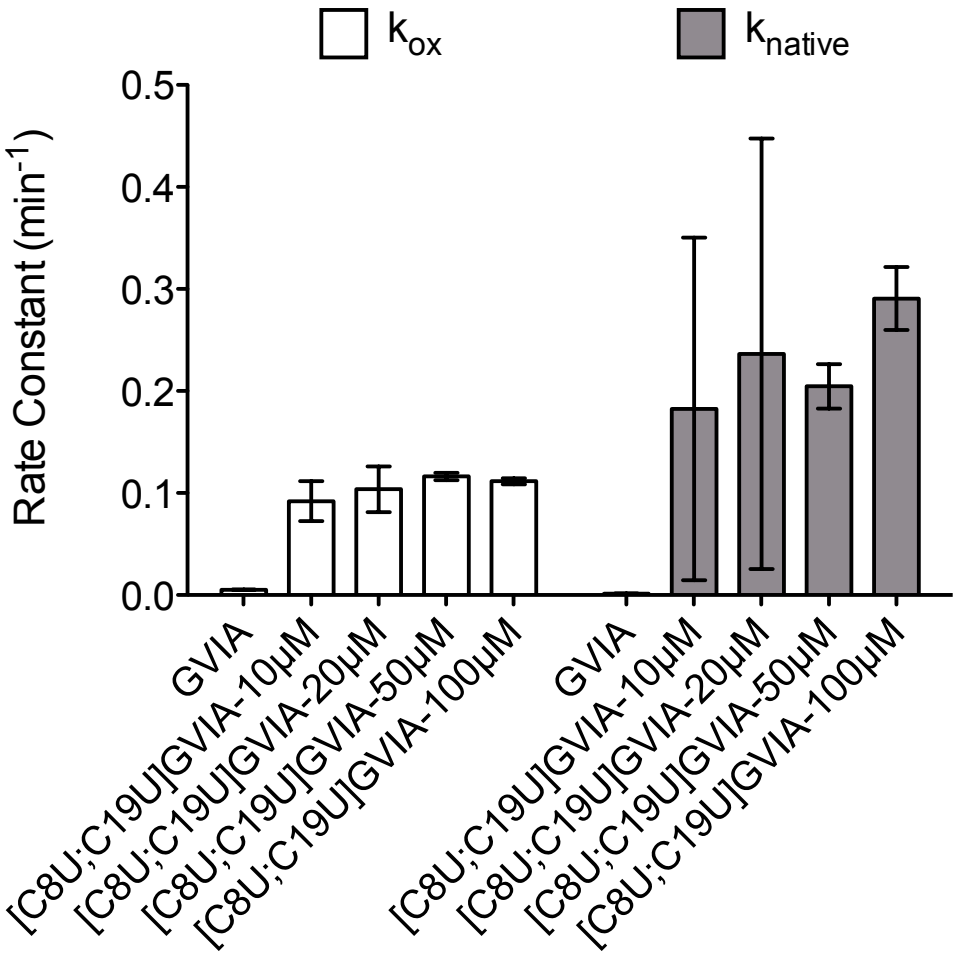


Figure 3.S1. The effect of peptide concentration on folding rate and disappearance of the "linear" form for [C8U;C19U] GVIA. Equivalent rates for GVIA (at 20 μM) are also shown for comparison. In all cases, $n = 3$ and values shown are mean $\pm$ SEM.

bias towards intermolecular catalysis, as the uncatalyzed folding of the unmodified peptides is extremely slow. Also, intermolecular diselenide catalysis differs in rate from catalysis with an intramolecular diselenide, despite the total diselenide and peptide concentrations being constant across conditions (Figure 3.4a and Table 3.S1).

3.3.2.2 Numerical Values for Rate Constants

In Table 3.S1, we present the numerical values for the parameters involved in the analysis of the oxidative folding of GVIA and SIIIA and their selenopeptide analogs.

The results in Table 3.S1 fit the data to the following equation, charting the extent of folding with respect to time, in minutes (given in Equation 3.1).

$$[F](t) = \gamma \cdot (1 - \frac{k_{native}}{k_{native} - k_{ox}} e^{-k_{ox} \cdot t} + \frac{k_{ox}}{k_{native} - k_{ox}} e^{-k_{native} \cdot t})$$

γ is the final yield, accounting for the fact that not all of the terminally oxidized peptide achieves the native disulfide connectivity. k_{ox} provides the rate constant for the initial oxidation step, and k_{native} provides the rate constant for all subsequent oxidation and isomerization steps leading towards native disulfide formation.

Unless otherwise indicated, the peptide concentration was 20 μM. γ is 'unitless,' and represents the fraction of total peptide that achieves the native disulfide connectivity. The units for k_{ox} and k_{native} are both min^{-1}. Provided error ranges are standard errors.

3.3.2.3 Kinetic Traces of Folding with Linear Timescales

Kinetic traces with a linear timescale are shown in Figures 3.S2 and 3.S3; they emphasize the rapidity of oxidative folding by autocatalysis with an intramolecular diselenide relative to the uncatalyzed condition.

3.3.2.4 Folding Chromatograms

Figure 3.S4 shows the HPLC chromatograms of reagentless folding reactions with respect to time for GVIA, [C8U;C19U] GVIA, SIIIA, [C3U;C13U] SIIIA, [C4U;C19U] SIIIA and [C8U;C20U] SIIIA. In all chromatograms, * indicates the native peak, and x indicates the reduced peptide.

3.3.2.5 Discussion of Error Values of k_{native}

In some cases, the standard error values for the rate of native formation, k_{native} are artificially large; in particular, the copper-catalyzed folding of GVIA and the lower

Table 3.S1. Numerical values for thermodynamic and kinetic parameters of the oxidation of SIIIA, GVIA and their selenopeptide analogs. In all cases, $n = 3$ and values shown are mean ± SEM. *Note: The kinetic model used for these analyses is not valid for the copper(II)- and selenocystine-catalyzed oxidation reactions, but the model was used for these reaction systems for effective comparison with catalysis by an intramolecular diselenide.*

Table S1: Kinetic Parameters			
Folding Substrate	γ	k_{ox} (min^{-1})	k_{native} (min^{-1})
SIIIA	0.850	0.00201 ± 0.0000595	0.000355 ± 0.0000416
[C3U;C13U] SIIIA	0.801	0.0121 ± 0.000654	0.00521 ± 0.001863
[C4U;C19U] SIIIA	0.823	0.0257 ± 0.000960	0.0121 ± 0.00130
[C8U;C20U] SIIIA	0.625	0.0393 ± 0.00263	0.009942 ± 0.001793
GVIA	0.319	.00526 ± .000210	0.00132 ± 0.000279
10 μM [C8U;C19U] GVIA	0.673	0.0921 ± 0.0196	0.183 ± 0.1679
20 μM [C8U;C19U] GVIA	0.736	0.1038 ± 0.02238	0.236 ± 0.211
50 μM [C8U;C19U] GVIA	0.771	0.116 ± 0.00373	0.205 ± 0.0218
100 μM [C8U;C19U] GVIA	0.805	0.112 ± 0.00298	0.291 ± 0.0310
GVIA + 100 nM $CuCl_2$	0.169	0.0351 ± 0.00268	0.109 ± 0.590
[C8U;C19U] GVIA + 100 nM $CuCl_2$	0.723	0.110 ± 0.00395	0.147 ± 0.0165
SIIIA + 20 μM Selenocystine	0.799	0.0831 ± 0.00539	0.00516 ± 0.000479
GVIA + 20 μM Selenocystine	0.615	0.0573 ± 0.00598	0.00542 ± 0.00181
[C8U;C19U] GVIA + 20 μM Selenocystine	0.630	0.0948 ± 0.00590	0.185 ± 0.0536

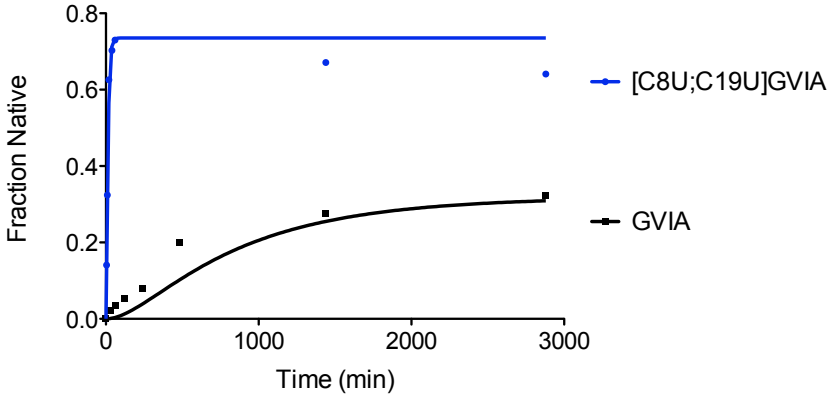


Figure 3.S2. Kinetic traces of folding of GVIA and [C8U;C19U] GVIA, shown with a linear timescale.

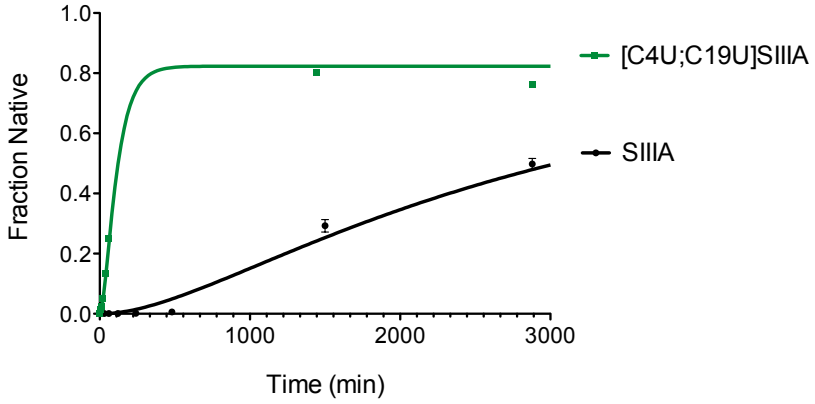


Figure 3.S3. Kinetic traces of folding of SIIIA and [C4U;C19U] SIIIA, shown with a linear timescale.

concentrations (10 μM and 20 μM) of [C8U;C19U] GVIA. In both cases, these are artifacts of the kinetic model.

In the case of [C8U;C19U] GVIA, the large standard error of k_{native} is an artifact of the model with respect to the calculation of yield. As shown in Figure 3.3(c), there is a decrease in the overall yield of natively folded peptide from 1 hour to 24 and 48 hours, causing uncertainty in the output of the final yield. As discussed in the Data Analysis section, a third rate constant, k_{non}, was used to account for alternative folding isomers, and

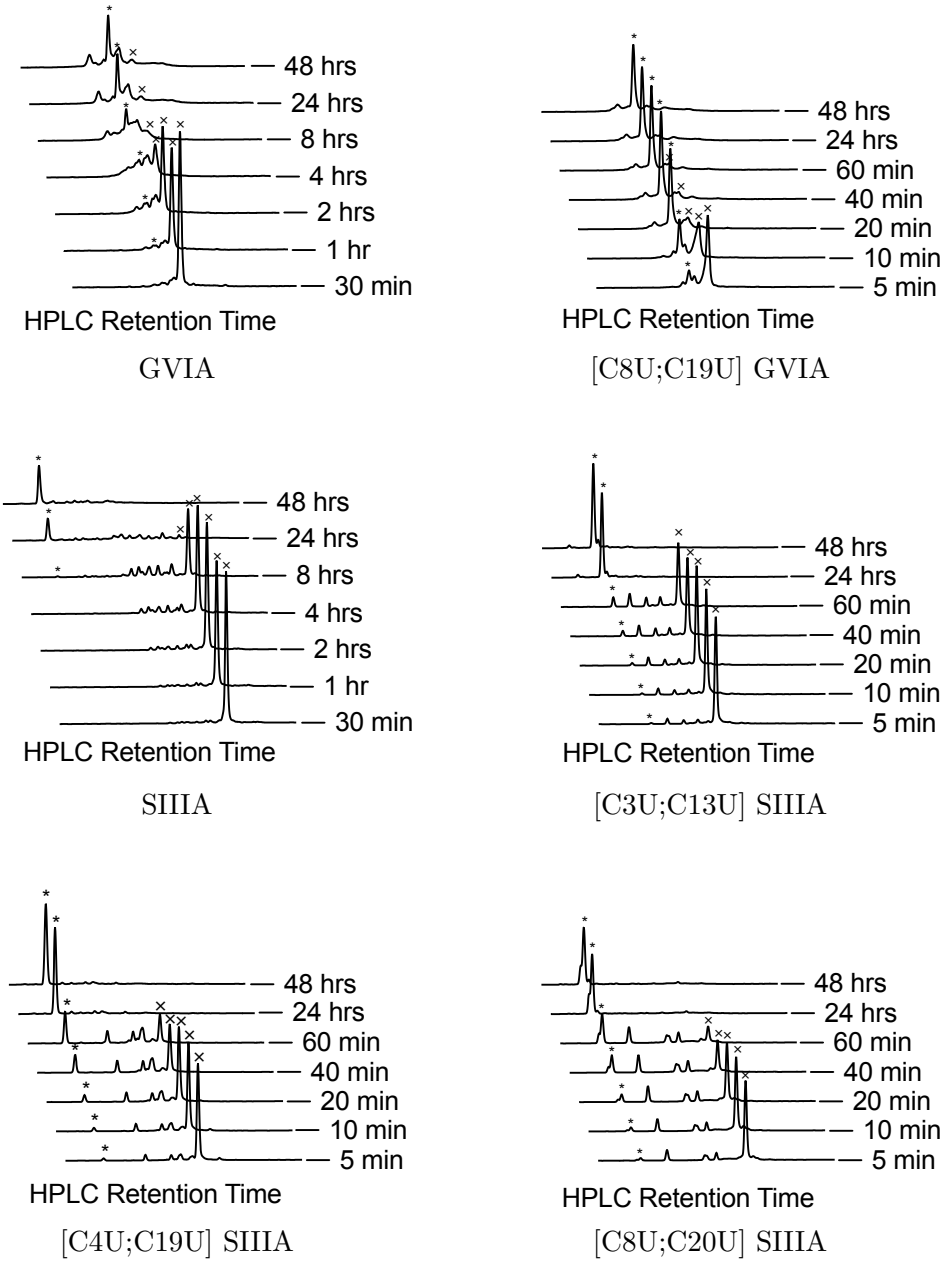


Figure 3.S4. Representative HPLC of folding chromatograms for each timepoint. * indicates the natively folded form, and x indicates the linear/diselenide-only form. Peaks were integrated and analyzed as described in the Data Analysis section.

the yield is a function of k_{native} and k_{non}. Therefore, the variation in yield from 1 hour to 48 hours appears as error in the determination of k_{native} and k_{non}.

The decrease in the fraction of natively folded peptide at extended timepoints could result from slow isomerization driven by a (very) small fraction of peptide thiols remaining in solution or minor reversibility. Interestingly, both suggest that the intramolecular folding mechanism is a higher-yield reaction pathway than intermolecular folding.

While the calculations and fitting to the kinetic model is not ideal due to the convolution of yield with k_{native}, it was selected in order to better assess errors due to covariability of k_{ox} with k_{native}. Alternatively, if all of the kinetic parameters are determined sequentially—k_{ox} from the depletion of the reduced peptide, then k_{native} and yield using the provided value of k_{ox} with direct fitting to Equation (1)—the kinetic parameters are effectively deconvoluted from the yield, resulting in dramatically decreased standard error values. For instance, with [C8U;C19U] GVIA folded at 20 μM, this deconvolution of the kinetic parameters and the yield decreases the standard error of k_{native} from 0.236 ± 0.211 min^{-1} (89% error) to 0.285 ± 0.0364 min^{-1} (13% error). The full set of sequentially-determined rate data is shown in Table 3.S2.

In the case of the copper-catalyzed folding of GVIA, the kinetic model fails to capture the complexity of the established copper-catalyzed reaction mechanism, wherein the formation of disulfide bridges proceeds through two distinct phases, each with separate reaction rates (20). However, the model was used in this instance in order to allow effective comparison with catalysis by an intramolecular diselenide. Also, the addition of copper to the folding reactions complicates the interpretation of HPLC chromatograms. Beginning at 2 hours, there is a peak that could be either a 15-second shift in HPLC elution time of the reduced peptide or a copper-specific folding intermediate, as it was not present in any other folding conditions. These two outcomes are compared in Table 3.S3. In our analyses, we used the interpretation that gave larger standard error values in order to avoid understatement of errors.

It is also worth noting that for all of these conditions, the variability represented in the determination of k_{native} is largely due to the fitting to the kinetic model (and not due to variability in the raw data). To demonstrate the consistency of the raw data, we can also fit each of the three replicates separately to the kinetic model (allowing for covariability of k_{ox} and k_{native}), and then determine values for γ (yield), k_{ox} and k_{native} for each replicate. The traces for each replicate are shown in Figure 3.S5. Also, by determining the mean ± SD from the kinetic parameters for each replicate, we show that relatively little of

Table 3.S2. Numerical values for the yield and kinetic parameters when the parameters are determined sequentially, effectively deconvoluting k_{native} from the yield, but also eliminating the covariability of k_{ox} with k_{native}. In all cases, $n = 3$ and values shown are mean ± SEM. γ represents the final yield. In this case, errors are shown for γ because it was determined separately from the rate constants.

Table S2: Kinetic Parameters Determined Sequentially

Folding Substrate	γ	k_{ox} (min^{-1})	k_{native} (min^{-1})
SIIIA	0.763 ± 0.116	0.00179 ± 0.0000552	0.000480 ± 0.000123
[C3U;C13U] SIIIA	0.761 ± 0.00707	0.00931 ± 0.000667	0.00828 ± 0.00116
[C4U;C19U] SIIIA	0.782 ± 0.00530	0.0225 ± 0.00102	0.0165 ± 0.000610
[C8U;C20U] SIIIA	0.594 ± 0.00634	0.0294 ± 0.00193	0.0199 ± 0.000954
GVIA	0.303 ± 0.00836	0.00581 ± 0.000162	0.00380 ± 0.000424
10 μM [C8U;C19U] GVIA	0.636 ± 0.00587	0.106 ± 0.00357	0.217 ± 0.0160
20 μM [C8U;C19U] GVIA	0.697 ± 0.00971	0.112 ± 0.00318	0.285 ± 0.0364
50 μM [C8U;C19U] GVIA	0.731 ± 0.00854	0.124 ± 0.00189	0.241 ± 0.0226
100 μM [C8U;C19U] GVIA	0.764 ± 0.00738	0.116 ± 0.00211	0.338 ± 0.0327
GVIA + 100 nM $CuCl_2$	0.157 ± 0.00279	0.0518 ± 0.00215	0.150 ± 0.0255
[C8U;C19U] GVIA + 100 nM $CuCl_2$	0.686 ± 0.00879	0.122 ± 0.00355	0.176 ± 0.0154
SIIIA + 20 μM Selenocystine	0.759 ± 0.0173	0.0924 ± 0.00864	0.00632 ± 0.000498
GVIA + 20 μM Selenocystine	0.584 ± 0.0152	0.0642 ± 0.0136	0.00842 ± 0.00104
[C8U;C19U] GVIA + 20 μM Selenocystine	0.596 ± 0.0150	0.0958 ± 0.00763	0.292 ± 0.0723

Table 3.S3. Comparison of two alternative interpretations of the copper-catalyzed folding data. Note that if there is a shift in the elution time of the reduced peptide, the error value for k_{native} suggests a 541% error, whereas if the elution time of the reduced peptide is assumed to not shift (indicating the existence of a copper-specific folding intermediate), the error value for k_{native} suggests a 147% error.

Table S3: Fitting of Copper Catalysis Data			
Identity of Peak	γ	k_{ox} (min^{-1})	k_{native} (min^{-1})
Elution time shift	0.169	0.0351 ± 0.00268	0.109 ± 0.590
Copper-specific intermediate	0.168	0.0359 ± 0.000903	0.0915 ± 0.135

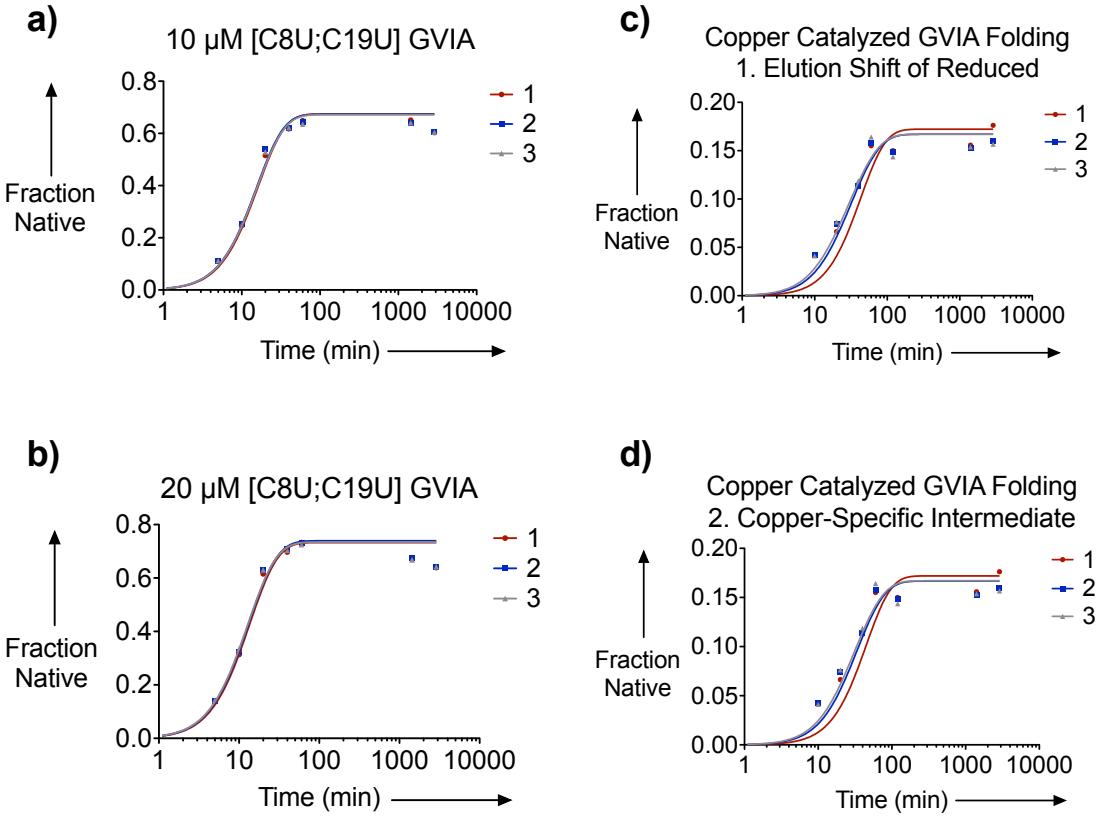


Figure 3.S5. Kinetic traces for the individual replicates of (a) 10 μM folding of [C8U;C19U] GVIA, (b) 20 μM folding of [C8U;C19U] GVIA, (c) copper-catalyzed folding, assuming the shift of elution time of the reduced peak, and (d) copper-catalyzed folding, assuming no shift of elution time for the reduced peak (and that there is a copper-specific folding intermediate). Each trace was determined allowing for the covariability of k_{ox} with k_{native}.

the error shown in Tables 3.S1 and 3.S2 is due to variability in the raw data. This is shown in Table 3.S4.

As indicated by the values in Tables 3.S1, 3.S2, 3.S3 and 3.S4, we elected to use the data analysis method that generally produced the largest error ranges, in order to avoid understatement of error.

3.3.3 Methods

3.3.3.1 Reduced Peptides

Peptides were synthesized on a single-channel automated peptide synthesizer using Rink amide resins and standard Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry. All cysteine residues were trityl protected, and all selenocysteine residues were 4-methoxybenzyl protected. Peptides were cleaved from the resin with either Reagent K (for non-selenocysteine-containing peptides) or Reagent K‡ (for selenocysteine containing peptides). Reagent K contained a 33:2:2:2:1 mixture of TFA : Phenol : water : thioanisole : 1,2-ethanedithiol; Reagent K‡ contained a 178:14:5:1.6:1 mixture of TFA : Phenol : water : thioanisole : DTNP. All peptides were cleaved for 3 hours. Cleaved peptides were filtered and precipitated with MTBE. Selenocysteine-containing peptides were then subjected to 2 hours of reduction (incubation in 50mM DTT, 0.1M Tris, pH 7.5 with 0.1mM EDTA), which was quenched by addition of formic acid to 8% (final). Peptides were then purified by HPLC on either a Waters 600 chromatograph or a Waters 2535 Quaternary Gradient Module, using a semi-preparative Vydac C18 column (218TP510). All peptides (reduced and oxidized) were purified using a gradient that ran from 5% to 35% B90 (89.9% acetonitrile, 10% water, 0.1% TFA), with the remainder being 0.1% TFA in water. The identity of the reduced peptides was then validated by electrospray ionization mass spectrometry.

3.3.3.2 Oxidative Folding

All oxidations were air-mediated, using molecular oxygen as the only oxidant, except as indicated.

Dried, reduced peptides were redissolved in 5% acetonitrile, 0.01% TFA (94.99% water) to 200 or 400 μM. In most cases, the peptide was dissolved to 200 μM; however, for folding reactions with a final peptide concentration of 100 μM, a more concentrated solution was used in order to limit the effect of the TFA on the final buffered pH.

Folding mixtures were then prepared, containing everything except the reduced peptides. Concentrations were chosen so that upon addition of the reduced peptides, the folding reactions would contain 0.1 M Tris, pH 7.5, 0.1 mM EDTA, as well as any other additives

Table 3.S4. Averaged kinetic parameters from fitting each of the three replicates, treated separately. Data is shown for conditions that showed a large standard error: 10 and 20 μM foldings of [C8U;C19U] GVIA, and the two interpretations of the copper-catalysis data. For all conditions, $n = 3$ and values shown are mean ± SD.

Table S4: Averaged Kinetic Parameters from Individual Replicates			
Folding Substrate	γ	k_{ox} (min^{-1})	k_{native} (min^{-1})
10 μM [C8U;C19U] GVIA	0.673 ± 0.00136	0.0921 ± 0.00155	0.183 ± 0.0100
20 μM [C8U;C19U] GVIA	0.736 ± 0.00323	0.104 ± 0.00125	0.237 ± 0.0149
GVIA+100 nM $CuCl_2$-Elution Shift	0.169 ± 0.00298	0.0351 ± 0.000386	0.131 ± 0.0623
GVIA+100 nM $CuCl_2$-Copper Specific Intermediate	0.169 ± 0.00293	0.0359 ± 0.000357	0.104 ± 0.0428

(the amount of EDTA was adjusted down to 1 nM for folding reactions containing 100 nM $CuCl_2$ as an oxidative catalyst). This mixture was allowed to thermally equilibrate for at least 5 minutes in a block set to 20^o C.

Peptide was then added to each folding reaction, and timepoints were taken as necessary. For selenocysteine-containing peptides, timepoints were at 5, 10, 20, 40 and 60 minutes, and 24 and 48 hours. For peptides that did not contain selenocysteine (or an oxidative catalyst), timepoints were taken at 0.5, 1, 2, 4, 8, 24 and 48 hours. For folding reactions with copper(II) chloride as an oxidative catalyst, timepoints were taken at 10, 20, 40 and 60 minutes, and 2, 24 and 48 hours. For reactions with selenocystine as an oxidative catalyst, timepoints were taken at 5, 10, 20, 40, 60 minutes, 24 and 48 hours. Each timepoint contained 2 nmol of peptide, and was added to a tube containing 15% of the volume of the timepoint of 1 M HCl. This was then transferred to a pre-chilled HPLC vial, and 1.67 nmol was injected onto an analytical Vydac C18 column (218TP54) with a Waters e2695 Separations Module. For more concentrated folding reactions, 0.1% TFA (in water) was added to a final volume of 120 μl for increased precision in folding analysis. Samples are kept at 4^o C until they are run on the analytical HPLC. All timepoints were analyzed by analytical HPLC within 24 hours of the timepoint being taken.

All folding reactions were performed in 0.1 M Tris, pH 7.5. Upon quenching with HCl, the pH is lowered to approximately 1.45. Any further dilution that occurred with 0.1% TFA in water (only in more concentrated folding reactions) maintained the pH below 2.0.

Unless otherwise indicated, all folding reactions were performed with 20 μM peptide.

3.3.3.3 Data Analysis

The first step in data analysis was to use Empower 3 to perform HPLC peak integration. This provided the percent of the total peptide that was present as either reduced or native peptide at each timepoint. The elution times of native peptide are known from our previous work (4), and the elution times of reduced peptides were shown using reduced peptides that had been kept under acidic conditions, prohibiting disulfide formation.

Data were fit to a slightly modified kinetic model of the chemical reaction shown in Figure 3.S6, where R denotes the reduced peptide (bearing either six thiols *or* four thiols and a diselenide), I denotes any folding intermediates and F denotes folded peptide with the native disulfide connectivity.

This assumes that every oxidation is irreversible. Disulfide isomerization following the formation of the first disulfide bridge is not explicitly included in this model because any

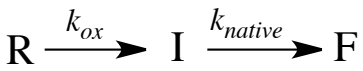


Figure 3.S6. Reaction scheme used for kinetic analyses

species bearing one disulfide bridge is an intermediate. The classical model for product formation in this reaction sequence is provided below, and is given in [26].

$$[F](t) = (1 - \frac{k_{native}}{k_{native} - k_{ox}} e^{-k_{ox} \cdot t} + \frac{k_{ox}}{k_{native} - k_{ox}} e^{-k_{native} \cdot t})$$

However, this had to be modified, since one of the assumptions is that all terminally oxidized forms are equivalent, which is not the case for our material (there are numerous folding isomers). Consequently, we also incorporated a multiplicative factor of γ, representing the fraction of the material that actually proceeds to the native disulfide connectivity (final yield). Thus, our analyses were done with Equation 3.1, below.

$$[F](t) = \gamma \cdot (1 - \frac{k_{native}}{k_{native} - k_{ox}} e^{-k_{ox} \cdot t} + \frac{k_{ox}}{k_{native} - k_{ox}} e^{-k_{native} \cdot t}) \tag{3.1}$$

DynaFit was used to solve Equation (1) for γ, k_{ox} and k_{native} simultaneously, using both the depletion of the reduced peptide ($\frac{d[R]}{dt}$) and the native formation ($\frac{d[F]}{dt}$). The maximum value was set to 0.95, as this was the approximate purity of the reduced peptides. The yield of the reaction (γ) was solved for by adding another component, wherein I goes to M (non-native or mis-folded folding products) at a rate of k_{non}. Because both reactions from intermediates to native product or intermediates to non-native folding products are assumed irreversible,

$$\gamma = \frac{k_{native}}{k_{native} + k_{non}}$$

Error values provided are formal standard errors, as given by DynaFit. Because the fitting software only used information on the depletion of the linear form (R) and the appearance of the native fold (F), the value of k_{non} is determined relative to the value of k_{native} in order to maintain a reasonable value for γ; consequently, errors were not calculated on the derivation of γ.

Most computations were performed with DynaFit; sequential determination of kinetic parameters (Table 3.S2) was performed with Prism 5. Graphical preparations were done with PyMOL, ChemBioDraw and Prism 5.

3.4 Conclusion

In addition to the paper that is included here, another paper with similar conclusions was also presented side-by-side [25], based on our previous arrangement with the authors of the other paper. This side-by-side publication used diselenide catalysis as well, but the core focus of their discovery was the use of non-native diselenide bridges to drive oxidative folding through productive folding pathways, whereby avoiding kinetic traps. To accomplish this goal, Metanis and Hilvert used Bovine Pancreatic Trypsin Inhibitor (BPTI), which has been extensively studied with respect to molecular dynamics [27], mechanism of action [28], folding pathways [29, 30], enzymatically-assisted folding [31], as well as folding with an intermolecular diselenide catalyst *in vitro* [3].

As a follow-up to these two papers, a commentary was published to discuss the implications to oxidative folding, the potential importance of selenium and selenocysteine to future biotechnological developments, and the enhancements to production–both chemical and recombinant–that may result from this work [32].

In addition to the rate increase, the inclusion of an intramolecular folding catalyst also has implications to the synthesis and testing of conotoxins. As was noted in both Chapters 1 and 2, oxidative folding is classically considered the rate-limiting-step to the synthesis and folding of conotoxins, and consequently to the exploration of the vast number of conotoxin sequences that are known, but for which peptide has never been synthesized, folded or tested for activity. In such cases, it may be advantageous to assay folding mixtures. Consequently, the observation that oxidative folding will occur spontaneously and with reasonable yields offers great promise to streamlining the discovery process for novel venom-based therapeutics.

3.5 References

[1] (a) A. Walewska, M. M. Zhang, J. J. Skalicky, D. Yoshikami, B. M. Olivera, G. Bulaj, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 2221–2224, doi:10.1002/anie.200806085; (b) A. Walewska, M.M. Zhang, J.J. Skalicky, D. Yoshikami, B.M. Olivera, G. Bulaj, *Angew. Chem.* **2009**, *121*, 2255-2258, doi: 10.1002/ange.200806085.

[2] (a) Dörthe Besse, F. Siedler, T. Diercks, H. Kessler, L. Moroder, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 883–885, doi:10.1002/anie.199708831; (b) D. Besse, F. Siedler, T. Diercks, H. Kessler, L. Moroder, *Angew. Chem.* **1997**, *109*, 915-917, doi: 10.1002/ange.19971090828.

[3] J. Beld, K. J. Woycechowsky, D. Hilvert, *Biochemistry* **2008**, *47*, 6985–6987, doi: 10.1021/bi8008906.

[4] J. Beld, K. J. Woycechowsky, D. Hilvert, *Biochemistry* **2007**, *46*, 5382–5390, doi: 10.1021/bi700124p.

[5] J. Beld, K. J. Woycechowsky, D. Hilvert, *ACS Chem. Biol.* **2010**, *5*, 177–182, doi: 10.1021/cb9002688.

[6] H. Nakamoto, J. C. Bardwell, *Biochim. Biophys. Acta* **2004**, *1694*, 111–119, doi: 10.1016/j.bbamcr.2004.02.012.

[7] L. Ellgaard, L. W. Ruddock, *EMBO Rep.* **2005**, *6*, 28–32, doi: 10.1038/sj.embor.7400311.

[8] I. Annis, B. Hargittai, G. Barany, *Methods Enzymol.* **1997**, *289*, 198–221.

[9] A. M. Steiner, G. Bulaj, *J. Pept. Sci.* **2011**, *17*, 1–7, doi:10.1002/psc.1283.

[10] C. Allmang, L. Wurth, A. Krol, *Biochim. Biophys. Acta* **2009**, *1790*, 1415–1423, doi: 10.1016/j.bbagen.2009.03.003.

[11] S. Fiori, S. Pegoraro, S. Rudolph-Böhner, J. Cramer, L. Moroder, *Biopolymers* **2000**, *53*, 550–564, doi:10.1002/(SICI)1097-0282(200006)53:7<550::AID-BIP3>3.0.CO;2-O.

[12] C. J. Armishaw, N. L. Daly, S. T. Nevin, D. J. Adams, D. J. Craik, P. F. Alewood, *J. Biol. Chem.* **2006**, *281*, 14136–14143, doi:10.1074/jbc.M512419200.

[13] K. H. Gowd, V. Yarotskyy, K. S. Elmslie, J. J. Skalicky, B. M. Olivera, G. Bulaj, *Biochemistry* **2010**, *49*, 2741–2752, doi:10.1021/bi902137c.

[14] A. Walewska, A. Jaśkiewicz, G. Bulaj, K. Rolka, *Chem. Biol. Drug Des.* **2011**, *77*, 93–97, doi:10.1111/j.1747-0285.2010.01046.x.

[15] S. Pegoraro, S. Fiori, J. Cramer, S. Rudolph-Böhner, L. Moroder, *Protein Sci.* **1999**, *8*, 1605–1613, doi:10.1110/ps.8.8.1605.

[16] M. Muttenthaler, S. T. Nevin, A. A. Grishin, S. T. Ngo, P. T. Choy, N. L. Daly, S. H. Hu, C. J. Armishaw, C. I. Wang, R. J. Lewis, J. L. Martin, P. G. Noakes, D. J. Craik, D. J. Adams, P. F. Alewood, *J. Am. Chem. Soc.* **2010**, *132*, 3514–3522, doi:10.1021/ja910602h.

[17] T. S. Han, M.-M. Zhang, K. H. Gowd, A. Walewska, D. Yoshikami, B. M. Olivera, G. Bulaj, *ACS Med. Chem. Lett.* **2010**, *1*, 140–144, doi:10.1021/ml900017q.

[18] (a) M. Mobli, A. D. de Araújo, L. K. Lambert, G. K. Pierens, M. J. Windley, G. M. Nicholson, P. F. Alewood, G. F. King, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 9312–9314, doi:10.1002/anie.200905206; (b) M. Mobli, A.D. de Araújo, L.K. Lambert, G.K. Pierens, M.J. Windley, G.M. Nicholson, P.F. Alewood, G.F. King, *Angew. Chem.* **2009**, *121*, 9476-9478, doi: 10.1002/ange.200905206.

[19] R. S. Norton, B. M. Olivera, *Toxicon* **2006**, *48*, 780–798, doi:10.1016/j.toxicon.2006.-07.022.

[20] T. S. Han, M. M. Zhang, A. Walewska, P. Gruszczynski, C. R. Robertson, T. E. Cheatham, D. Yoshikami, B. M. Olivera, G. Bulaj, *ChemMedChem* **2009**, *4*, 406–414, doi:10.1002/cmdc.200800292.

[21] (a) A. D. de Araujo, B. Callaghan, S. T. Nevin, N. L. Daly, D. J. Craik, M. Moretta, G. Hopping, M. J. Christie, D. J. Adams, P. F. Alewood, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 6527–6529, doi:10.1002/anie.201101642; (b) A.D. de Araujo, B. Callaghan, S.T. Nevin, N.L. Daly, D.J. Craik, M. Moretta, G. Hopping, M.J. Christie, D.J. Adams, P.F. Alewood, *Angew. Chem.* **2011**, *123*, 6657-6659, doi: 10.1002/ange.201101642.

[22] A. V. Kachur, C. J. Koch, J. E. Biaglow, *Free Rad. Res.* **1999**, *31*, 23–34, doi: 10.1080/10715769900300571.

[23] G. T. Mullenbach, A. Tabrizi, B. D. Irvine, G. I. Bell, R. A. Hallewell, *Nucleic Acids Res.* **1987**, *15*, 5484–5484.

[24] K. H. Gowd, K. D. Blais, K. S. Elmslie, A. M. Steiner, B. M. Olivera, G. Bulaj, *Biopolymers* **2012**, *98*, 212–223, doi:10.1002/bip.22047.

[25] (a) N. Metanis, D. Hilvert, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 5585–5588, doi: 10.1002/anie.201109129; (b) N. Metanis, D. Hilvert, *Angew. Chem.* **2012**, *124*, 5683-5686, doi: 10.1002/ange.201109129.

[26] C. G. Swain, *J. Am. Chem. Soc.* **1944**, *66*, 1696–1700, doi:10.1021/ja01238a027.

[27] J. A. McCammon, B. R. Gelin, M. Karplus, *Nature* **1977**, *267*, 585–590.

[28] U. Quast, J. Engel, H. Heumann, G. Krause, E. Steffen, *Biochemistry* **1974**, *13*, 2512–2520.

[29] T. E. Creighton, *J. Mol. Biol.* **1974**, *87*, 563–577.

[30] J. S. Weissman, P. S. Kim, *Science* **1991**, *253*, 1386–1393, doi:10.1126/science.1716783.

[31] J. S. Weissman, P. S. Kim, *Nature* **1993**, *365*, 185–188, doi:10.1038/365185a0.

[32] D. J. Craik, *Nat. Chem.* **2012**, *4*, 600–602, doi:10.1038/nchem.1417.

CHAPTER 4

IN VIVO FOLDING: ENZYME CATALYSIS AND EVOLUTION OF DISULFIDE BRIDGES

The venom of cone snails is produced by specialized secretory cells in the venom duct. Because *Conus* venom duct cells have never been cultured, there is very little direct evidence for the mechanisms of production, or even the order of events in the biosynthesis of conotoxins. For example, it is unknown whether posttranslational modifications occur before, after or during folding. Nonetheless, it can be inferred from cDNA sequences that conotoxins use the typical pathway for secreted proteins: mRNA synthesis, excision of introns, binding of mRNA to a ribosome, ribosomal peptide synthesis on the rough endoplasmic reticulum, posttranslational processing (including folding) and secretion into the venom duct.

Given the final structure of conotoxins, this method of production is exactly what would be expected. However, it remains unclear which enzymes facilitate the posttranslational processing, whether there are cofactors that also play a role, what the order of steps are (or even if this order is significant or controlled). It is also unknown whether energy is expended to drive oxidative folding to a specific disulfide connectivity. There are other questions regarding conotoxin evolution as well: what are the evolutionary pressures that drive conservation of disulfide scaffolds? Is there significance to the codon choice at each site (which is conserved in each conotoxin family [1, 2, 3, 4]), or does this choice only signify the evolutionary relationship between these sequences? Some of these questions were addressed in collaborations that will be discussed further below. Others are discussed in the following chapter, although many of these questions remain open.

4.1 Enzyme-Assisted Folding

There are many critical aspects that distinguish *in vivo* from *in vitro* oxidative folding; most critical among these is likely the involvement of enzymes such as protein disulfide

isomerase (PDI) [5, 6] and the heat shock protein family of chaperones [7, 8, 9]. Post-translational modifications also likely play a critical role, especially in cases where the modification alters the folding efficiency [10, 11].

Given the broad array of *in vitro* folding efficiencies (see Table 2.1), it is likely that certain conotoxins employ chaperones to assist their folding *in vivo*; however, it is unlikely that every conotoxin employs such a chaperone system. Additionally, the particular chaperones involved in conotoxin folding are likely to differ from one peptide to the next. For instance, peptidylprolyl cis-trans isomerase (PPI) facilitates the folding of μ-conotoxin GIIIA [12], but is unlikely to act on peptides without proline residues, or peptides whose folding pathway does not require prolyl isomerization (meaning that PPI acts only on a subset of peptides containing proline residues).

4.1.1 Protein Disulfide Isomerase

Because the 'correct' disulfides are critical to the action of any cysteine-rich peptide [13, 14, 15], it is likely that cone snails employ PDI to fold conotoxins. While the precise mechanisms behind the substrate specificity of PDI have remained elusive, one common belief is that PDI identifies misfolded proteins by exposed hydrophobic surface area [16]. However, numerous conotoxins have significant exposed hydrophobic surface area in their correctly folded form, including μO-MrVIB (see Figure 4.1) and many δ-conotoxins, suggesting that PDI in *Conus* may use a different motif for substrate recognition.

Because conotoxins are short, cysteine-rich peptides, they may not contain sufficient sequence-encoded folding information to drive oxidative folding to the correct disulfide isomer, and likely employ PDI to assist in forming the correct disulfide bonds. However, with many substrates, the effect of PDI on folding efficiency is fairly modest [6, 17], with bovine pancreatic trypsin inhibitor (BPTI) being one of the first examples to show a significant benefit to PDI-catalyzed folding [6], suggesting that one of the primary functions of PDI is to assist in disulfide isomerization in kinetically trapped intermediates. Some conotoxins may also be small enough to not entirely bury their disulfides, allowing reduced glutathione to function as the primary isomerization agent. As PDI is the β-subunit of prolyl 4-hydroxylase [18], and many conotoxins contain hydroxylated proline residues [10], PDI may also be serving a noncanonical role in conotoxin biosynthesis.

4.1.2 Endoplasmic-Reticulum-Resident Chaperones

In addition to disulfide isomerization and posttranslational modifications, there are many other chaperones that are likely to be involved in the biosynthesis of conotoxins. The

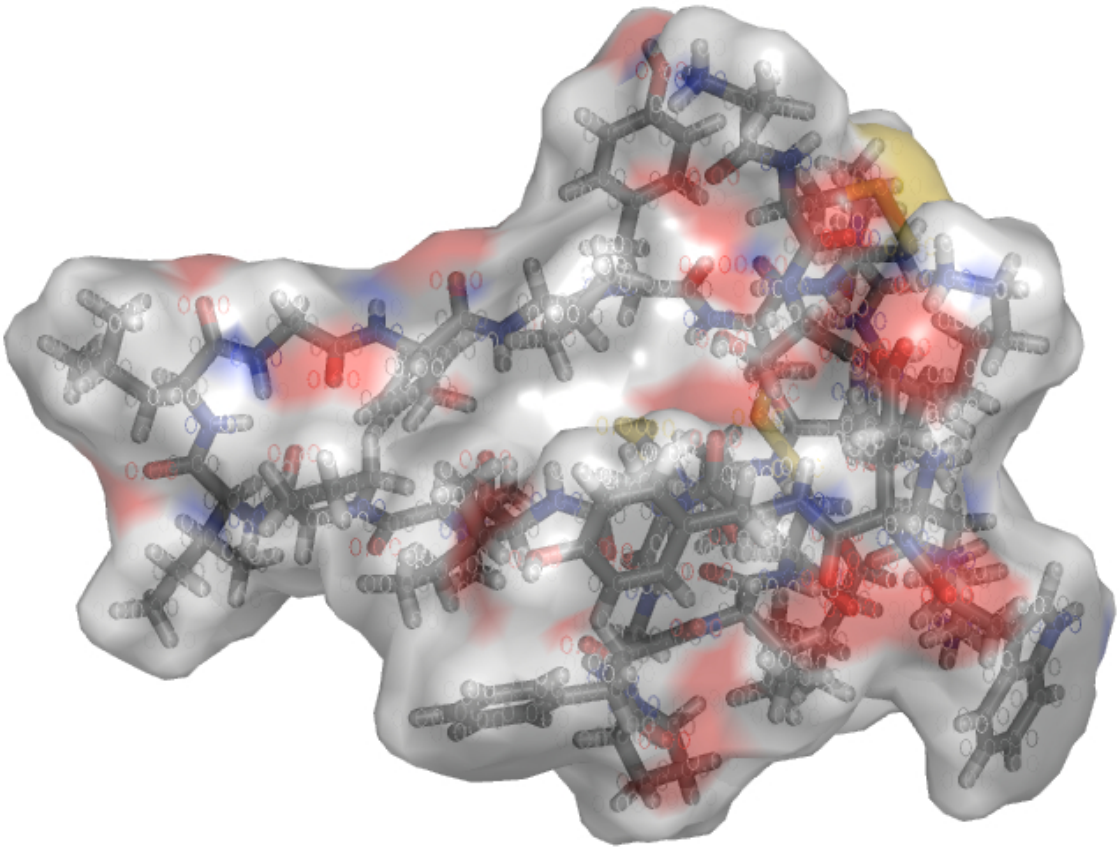


Figure 4.1. Three-dimensional structure of μO-MrVIB (PDBID: 1RMK), showing a surface representation, with transparency to see the peptide model underneath. Grey/white indicate hydrophobic regions, and red/blue indicate charged regions; disulfides are shown in yellow.

Hsp40/70 system is a canonical chaperone network that assists efficient protein folding, and is therefore likely involved in the folding of conotoxins as well.

BiP is the only 70-kDa heat shock protein (Hsp70) known to reside in the endoplasmic reticulum [19]; this class of chaperone has an ATPase domain, which controls the affinity for client proteins [7]. The ATPase activity of BiP is postulated to explain the observed ATP-dependence of oxidative folding of certain proteins [20, 21]. However, BiP does not directly bind client proteins; instead, the unfolded client proteins bind to a 40-kDa heat shock protein (a "J protein," from the homology to DnaJ in *E. coli*), which then carry the client protein to BiP. While there is only one ER-resident Hsp70, the J proteins are more numerous and believed to control the substrate specificity of BiP by determining which clients are shuttled to BiP [7].

4.1.3 Chaperone Activity in Oxidative Folding of Conotoxins

As mentioned above, the role of PPI in conotoxin folding has been investigated [12], and demonstrated that PPI does improve the folding of μ-GIIIA, although since many conotoxins do not contain proline residues, the role of PPI is unlikely to be universal. PDI has been shown to assist in the folding of tx3a and sTx3.1, showing notable enhancement to both oxidation and isomerization relative to glutathione [22]. It is worth noting, though, that the reports showing a role for PDI in assisting conotoxin folding have all used a *Conus* PDI, which allows for the possibility that conotoxins employ a specialized PDI that is more efficient with such small, cysteine-rich substrates.

Appendix B is a publication that discusses the role of chaperones in conotoxin folding [23]. Using a *Conus* BiP, the addition of BiP and PDI (also from *Conus*) together showed a greater effect than either individually. Because which J protein interacts with newly-synthesized conotoxins to carry them to BiP is unknown, microsomes derived from rat liver were added in order to provide an array of J proteins to fulfill this role. Together, these indicate that there may well be a multienzyme complex that assists the folding of conotoxins *in vivo*. Another interesting result of this work was the discovery of misfolded ImI in the venom of *C. imperialis*, which established that the involvement of chaperones does not necessarily direct the folding of conotoxins to a single, active disulfide connectivity.

4.2 Conservation of Disulfide Scaffolds

Disulfide scaffolds in *Conus* are extraordinarily conserved, with notable evidence for conservation at the codon level [1, 2, 3, 4]. Given the hypervariability of the adjacent

residues in the mature toxin, the reason for this extensive conservation is unclear. However, there are several examples of changes to disulfide scaffolds in *Conus*. In particular, the conkunitzins show a wide variety of disulfide scaffolds (see Figure 4.2), all of which are variants of the classical kunitz domain [24]. Other examples of changes to disulfide scaffolds can be found in the α- vs. αA-conotoxins (differing by one disulfide) [25], and in the M-superfamily (changes to the length of intercysteine loops with associated changes to disulfide connectivity) [4]. Given the wide diversity of disulfide scaffolds in *Conus*, these few isolated examples are the exception rather than the rule, and the conservation of such a great number of disulfide scaffolds (see Figure 1.4) is quite striking.

Appendix C is a publication that discusses potential reasons for the conservation of disulfide scaffolds in conotoxins [26]. Disulfide-depleted analogs of ω-GVIA were used to show the significance of the first disulfide, showing that it significantly impacted the efficiency of oxidative folding, but that this disulfide does not modulate the activity of the peptide. One of the most striking results of this paper is that using disulfide-depleted, diselenide-containing ω-GVIA does not fold as efficiently as unmodified ω-GVIA, despite only having one possible folding product (accounting for the strong preference for diselenide bridges over mixed sulfur-selenium bonds, as discussed in Chapter 3), as compared to the fifteen possible folding products for the unmodified peptide. These results are taken to indicate that the first disulfide in ω-GVIA is conserved to preserve efficient folding, although it is not known whether this bridge functions by conformationally constricting the peptide as the remaining bridges form, or if it acts by driving disulfide isomerization through more productive folding pathways.

4.3 Energetics of Venom Production

Venom is energetically expensive to produce; in one snake, the cost of venom regeneration was estimated at 26% of the cost of digesting small prey [27], although this may be an artificially low estimate [28]. To limit expenditure of venom, some venomous animals use 'dry bites' (no envenomation), which have been proposed to be due to kinematic constraints [28]. In the case of the cone snail, the venom is critical to effective prey capture, as *Conus* does not have an alternate (physical) means of prey immobilization. Consequently, in order to conserve energy, each envenomation event by *Conus* uses a very small amount of venom; *C. purpurascens* uses an average of 56 μl of venom for each envenomation [29].

The production of venom is very expensive at every step–and this is especially true for disulfide-rich, peptidic venom because *in vivo* oxidative folding yields are not quantitative

Figure 4.2. Disulfide scaffolds of conkunitzins. The traditional kunitz scaffold, bearing three disulfide bridges, is also used in *Conus* venom, as is a reduced version in which one disulfide has been deleted (top pannel). Conkunitzins can also have two–or even three–kunitz domains in a single toxin; the disulfide scaffold of two tandem domains has a fourth disulfide pair in the first kunitz domain (center pannel). Examples of conkunitzins with three kunitz domains are scarce, but the available example has three disulfides in each kunitz domain (bottom pannel).

[23]. Table 4.1 shows the energetic cost of venom production, from DNA to secretion. It is worth noting, though, that the per residue/vesicle values presented in Table 4.1 should only be used as approximations; these values do not account for introns; furthermore, the cost for secretion is expressed on a per vesicle basis, whereas the cost for synthesis is expressed on a per molecule/residue basis. Nonetheless, it is evident that venom is extraordinarily expensive to produce.

One critical question that is raised by the huge expense of venom production is why there is so much functional redundancy in venom [28]. There are several possible reasons for functional redundancy: (1) each peptide has its own metabolic half-life once injected into prey, and functional redundancy would increase the likelihood that the desired function is retained long enough for the snail to capture the prey; (2) by having multiple peptides (potentially with drastically different structures), the snail has an increased ability to adapt to changes in the target receptor; (3) in a venomic system with hypervariable mature toxin regions (as in *Conus*–see Chapter 5), the oxidative folding efficiency (and hence energetic cost of biosynthesis) for each peptide could vary wildly in each generation, and functional redundancy in such a system would allow buffering of the volatility in the energetic requirements to achieve the desired/necessary activity. It is worth noting that these potential benefits to functional redundancy in the venom are not mutually exclusive. This is discussed further in Chapter 5, although the focus of that discussion is buffering the energetic cost of venom production.

Table 4.1. Energy use in protein and peptide synthesis and secretion, quantifying energy expenditure from DNA to secretion into the venom duct. To compare the energetic cost of each step, the cost is also shown on a per residue or per vesicle basis. The cost of oxidative folding is unknown, and presumably varies immensely from one substrate to another. † This does not account for introns, which make this step less energetically efficient. ‡ This is debated, as some evidence exists that more than 1 molecule of GTP is hydrolyzed by EF-Tu. ◇ GroEL/ES folding often requires several cycles, with each cycle only accommodating one folding substrate, but the ATP demands are fairly independent of the size of the folding substrate.

Energy Use in Protein Synthesis and Secretion

Process	**Energetic Cost**	**Energy Per Residue/Vesicle**	**Ref(s)**
DNA → RNA	1 NTP / base	3 NTP†	[30]
Nuclear Export	1 ATP / Dbp5 cycle	1 ATP/mRNA	[31]
RNA → Protein	1 ATP, 2 GTP	3 NTP‡	[32, 33]
GroEL/ES assisted Folding	7 ATP/cycle	N/A◇	[34]
Oxidative Folding	Unk.	Unk.	
COPII Coating	1 GTP/ARF (1:1 to COP)	48 GTP/vessicle	[35]
SNARE Regeneration	1 ATP/complex/cycle	$\geq$ 2 per vesicle	[36]
Clathrin Uncoating	1 ATP / Clathrin	28/36/60 per vesicle	[37, 38]

4.4 References

[1] Conticello, S. G., Pilpel, Y., Glusman, G., and Fainzilber, M. (2000) Position-specific codon conservation in hypervariable gene families. *Trends Genet. 16*, 57–59.

[2] Conticello, S. G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z., and Fainzilber, M. (2001) Mechanisms for evolving hypervariability: the case of conopeptides. *Mol. Biol. Evol. 18*, 120–131.

[3] Santos, A. D., McIntosh, J. M., Hillyard, D. R., Cruz, L. J., and Olivera, B. M. (2004) The A-superfamily of conotoxins: structural and functional divergence. *J. Biol. Chem. 279*, 17596–17606.

[4] Wang, Q., Jiang, H., Han, Y. H., Yuan, D. D., and Chi, C. W. (2008) Two different groups of signal sequence in M-superfamily conotoxins. *Toxicon 51*, 813–822.

[5] Ellgaard, L. and Ruddock, L. W. (2005) The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep. 6*, 28–32.

[6] Weissman, J. S. and Kim, P. S. (1993) Efficient catalysis of disulphide bond rearrangements by protein disulphide isomerase. *Nature 365*, 185–188.

[7] Kampinga, H. H. and Craig, E. A. (2010) The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat. Rev. Mol. Cell Biol. 11*, 579–592.

[8] Gething, M. J. and Sambrook, J. (1992) Protein folding in the cell. *Nature 355*, 33–45.

[9] Weitzmann, A., Baldes, C., Dudek, J., and Zimmermann, R. (2007) The heat shock protein 70 molecular chaperone network in the pancreatic endoplasmic reticulum - a quantitative approach. *FEBS J. 274*, 5175–5187.

[10] Lopez-Vera, E., Walewska, A., Skalicky, J. J., Olivera, B. M., and Bulaj, G. (2008) Role of hydroxyprolines in the in vitro oxidative folding and biological activity of conotoxins. *Biochemistry 47*, 1741–1751.

[11] Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) Folding of ω-conotoxins. 2. Influence of precursor sequences and protein disulfide isomerase. *Biochemistry 35*, 15547–15557.

[12] Safavi-Hemami, H., Bulaj, G., Olivera, B. M., Williamson, N. A., and Purcell, A. W. (2010) Identification of *Conus* peptidylprolyl cis-trans isomerases (PPIases) and assessment of their role in the oxidative folding of conotoxins. *J. Biol. Chem. 285*, 12735–12746.

[13] Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) Folding of ω-conotoxins. 1. Efficient disulfide-coupled folding of mature sequences *in vitro*. *Biochemistry 35*, 15537–15546.

[14] Price-Carter, M., Hull, M. S., and Goldenberg, D. P. (1998) Roles of individual disulfide bonds in the stability and folding of an ω-conotoxin. *Biochemistry 37*, 9851–9861.

[15] Chang, S. G., Choi, K. D., Jang, S. H., and Shin, H. C. (2003) Role of disulfide bonds in the structure and activity of human insulin. *Mol. Cells 16*, 323–330.

[16] Hatahet, F. and Ruddock, L. W. (2007) Substrate recognition by the protein disulfide isomerases. *FEBS J. 274*, 5223–5234.

[17] Schönbrunner, E. R. and Schmid, F. X. (1992) Peptidyl-prolyl cis-trans isomerase improves the efficiency of protein disulfide isomerase as a catalyst of protein folding. *Proc. Natl. Acad. Sci. U. S. A. 89*, 4510–4513.

[18] Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhtala, M. L., Koivu, J., and Kivirikko, K. I. (1987) Molecular cloning of the β-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *EMBO J. 6*, 643–649.

[19] Gething, M. J. (1999) Role and regulation of the ER chaperone BiP. *Semin. Cell Dev. Biol. 10*, 465–472.

[20] Braakman, I., Helenius, J., and Helenius, A. (1992) Role of ATP and disulphide bonds during protein folding in the endoplasmic reticulum. *Nature 356*, 260–262.

[21] Mirazimi, A. and Svensson, L. (2000) ATP is required for correct folding and disulfide bond formation of rotavirus VP7. *J. Virol. 74*, 8048–8052.

[22] Wang, Z. Q., Han, Y. H., Shao, X. X., Chi, C. W., and Guo, Z. Y. (2007) Molecular cloning, expression and characterization of protein disulfide isomerase from *Conus marmoreus*. *FEBS J. 274*, 4778–4787.

[23] Safavi-Hemami, H., Gorasia, D. G., Steiner, A. M., Williamson, N. A., Karas, J. A., Gajewiak, J., Olivera, B. M., Bulaj, G., and Purcell, A. W. (2012) Modulation of conotoxin structure and function is achieved through a multienzyme complex in the venom glands of cone snails. *J. Biol. Chem. 287*, 34288–34303.

[24] Olivera, B., Bulaj, G., Garrett, J., Terlau, H., and Imperial, J. (2009) *Peptide Toxins from the Venoms of Cone Snails and other Toxoglossan Gastropods*, ed. de Lima, M. E. (EDITORA UFMG, Belo Horizonte/MG, Brazil), pp. 30–38.

[25] Jacobsen, R., Yoshikami, D., Ellison, M., Martinez, J., Gray, W. R., Cartier, G. E., Shon, K. J., Groebe, D. R., Abramson, S. N., Olivera, B. M., and McIntosh, J. M. (1997) Differential targeting of nicotinic acetylcholine receptors by novel αA-conotoxins. *J. Biol. Chem. 272*, 22531–22537.

[26] Gowd, K. H., Blais, K. D., Elmslie, K. S., Steiner, A. M., Olivera, B. M., and Bulaj, G. (2012) Dissecting a role of evolutionary-conserved but noncritical disulfide bridges in cysteine-rich peptides using ω-conotoxin GVIA and its selenocysteine analogs. *Biopolymers 98*, 212–223.

[27] Pintor, A. F., Krockenberger, A. K., and Seymour, J. E. (2010) Costs of venom production in the common death adder (*Acanthophis antarcticus*). *Toxicon 56*, 1035–1042.

[28] Morgenstern, D. and King, G. F. (2013) The venom optimization hypothesis revisited. *Toxicon 63*, 120–128.

[29] Chun, J. B., Baker, M. R., Kim, d. o. H., Leroy, M., Toribo, P., and Bingham, J. P. (2012) Cone snail milked venom dynamics–a quantitative study of *Conus purpurascens*. *Toxicon 60*, 83–94.

[30] Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., and Block, S. M. (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature 438*, 460–465.

[31] Linder, P. and Jankowsky, E. (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol. 12*, 505–516.

[32] Chinali, G. and Parmeggiani, A. (1980) The coupling with polypeptide synthesis of the GTPase activity dependent on elongation factor G. *J. Biol. Chem. 255*, 7455–7459.

[33] Weijland, A. and Parmeggiani, A. (1993) Toward a model for the interaction between elongation factor Tu and the ribosome. *Science 259*, 1311–1314.

[34] Xu, Z., Horwich, A. L., and Sigler, P. B. (1997) The crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature 388*, 741–750.

[35] Fath, S., Mancias, J. D., Bi, X., and Goldberg, J. (2007) Structure and organization of coat proteins in the COPII cage. *Cell 129*, 1325–1336.

[36] Jahn, R. and Scheller, R. H. (2006) SNAREs–engines for membrane fusion. *Nat. Rev. Mol. Cell Biol. 7*, 631–643.

[37] Rothman, J. E. (1996) The protein machinery of vesicle budding and fusion. *Protein Sci. 5*, 185–194.

[38] Fotin, A., Cheng, Y., Sliz, P., Grigorieff, N., Harrison, S. C., Kirchhausen, T., and Walz, T. (2004) Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature 432*, 573–579.

CHAPTER 5

CONSERVATION OF CYSTEINE CODONS IN CONOTOXINS

This chapter is a paper that has been submitted and its accompanying supplemental information.

5.1 On the Importance of Oxidative Folding in the Evolution of Conotoxins: Cysteine Codon Preservation Through Gene Duplication and Adaptation

Steiner, A., Bulaj, G., Puillandre, N. On the Importance of Oxidative Folding in the Evolution of Conotoxins: Cysteine Codon Preservation Through Gene Duplication and Adaptation. *J. Mol. Biol.*, *Submitted.*

5.1.1 Abstract

Conotoxin genes are among the most rapidly evolving genes currently known; however, despite the well-established hypervariability of the intercysteine loops, the cysteines demonstrate significant conservation, with a site-specific codon bias for each cysteine in a family of conotoxins. Herein we present a novel rationale behind the codon-level conservation of the cysteines that comprise the disulfide scaffold. We analyze cysteine codon conservation using an internal reference and phylogenetic tools; our results suggest that the established codon conservation can be explained as the result of selective pressures linked to the production efficiency and folding of conotoxins, which drive the conservation of cysteine at the amino-acid level. The preservation of cysteine has resulted in maintenance of the ancestral codon in most of the daughter lineages, despite the hypervariability of adjacent residues. We propose that the selective pressures acting on the venom components of cone snails involve an interplay of biosynthetic efficiency, activity at the target receptor, and the

importance of that activity to effective prey immobilization. Functional redundancy in the venom can thus serve as a buffer for the energy expenditure of venom production. The graphical abstract is shown as Figure 5.1.

5.1.2 Main Text

Conotoxins are peptides from the venom of *Conus*, a genus of predatory marine gastropod. When a cone snail envenomates its prey, a very small volume (6-480 μl; avg. 56.21 μl for *C. purpurascens*, [1]) of a highly-diverse venom is injected to paralyze the prey. This highly-diverse venom is largely composed of disulfide-rich peptides, which primarily target ion channels on the cell surface [2]. The extraordinary diversity of *Conus* venom is genetically-encoded within the hypervariable mature toxin region [3] and arises from gene duplication and positive selection [4], which are believed to be at constant work in *Conus*, allowing the snail to compete effectively in the venom-target molecular arms race [5]. This results in exceptionally potent and highly specific toxins, such that only a small volume of venom is necessary to efficiently paralyze the prey. The high potency of conotoxins minimizes the venom requirements (thus, energetic expenditure) of each envenomation event, while also allowing some variation in targeted prey.

Conotoxins have been extensively studied with respect to function [6], oxidative folding [7], gene structure [8, 9, 10], and their evolutionary patterns [3, 4, 11]. Conotoxin genes consist of a signal sequence, propeptide, and a toxin region containing highly conserved cysteine residues separated by hypervariable intercysteine loops. In this work we focused on the δ- and ω- families of conotoxins, which bear the intercysteine knot (ICK) motif. The ICK motif is defined by the cysteine scaffold, and is broadly employed by many organisms outside *Conus*, including insects, plants, fungi and spiders [12]; previous work has also referred to conotoxins bearing the ICK motif as 'four-loop conotoxins' [11, 13, 14].

Highly conserved cysteine residues are interspersed with the hypervariable intercysteine loops, with a site-specific codon preference for each cysteine in the disulfide scaffold (Figure 5.2). While closely related conotoxin families have similar codon usage patterns for cysteine, differences between two families are also conserved within each family (see Cysteine 6 in Figure 5.2). The driving force behind the extensive site-specific conservation of cysteine codons–which are flanked by hypervariable intercysteine loops–has not yet been fully elucidated. Proposed mechanisms for the codon-level conservation of cysteine involve a 'block substitution' mechanism [15] and a macromolecule that binds specifically to TGC or TGT [16, 13]; the latter has received support from frame-shifted pseudogenes which retained characteristic cysteine codons [13, 17]. However, our analyses indicated that

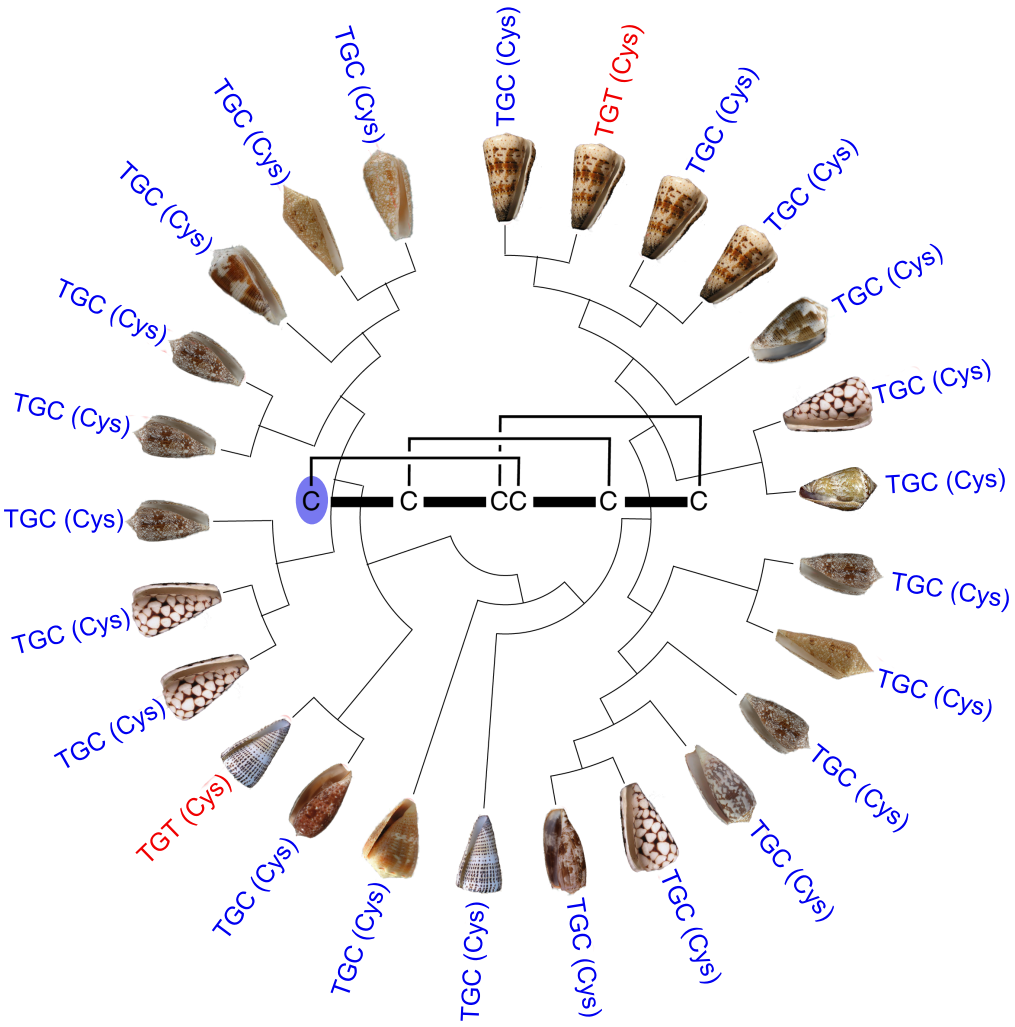


Figure 5.1. Graphical abstract, showing cysteine codon conservation in a group of δ-conotoxins.

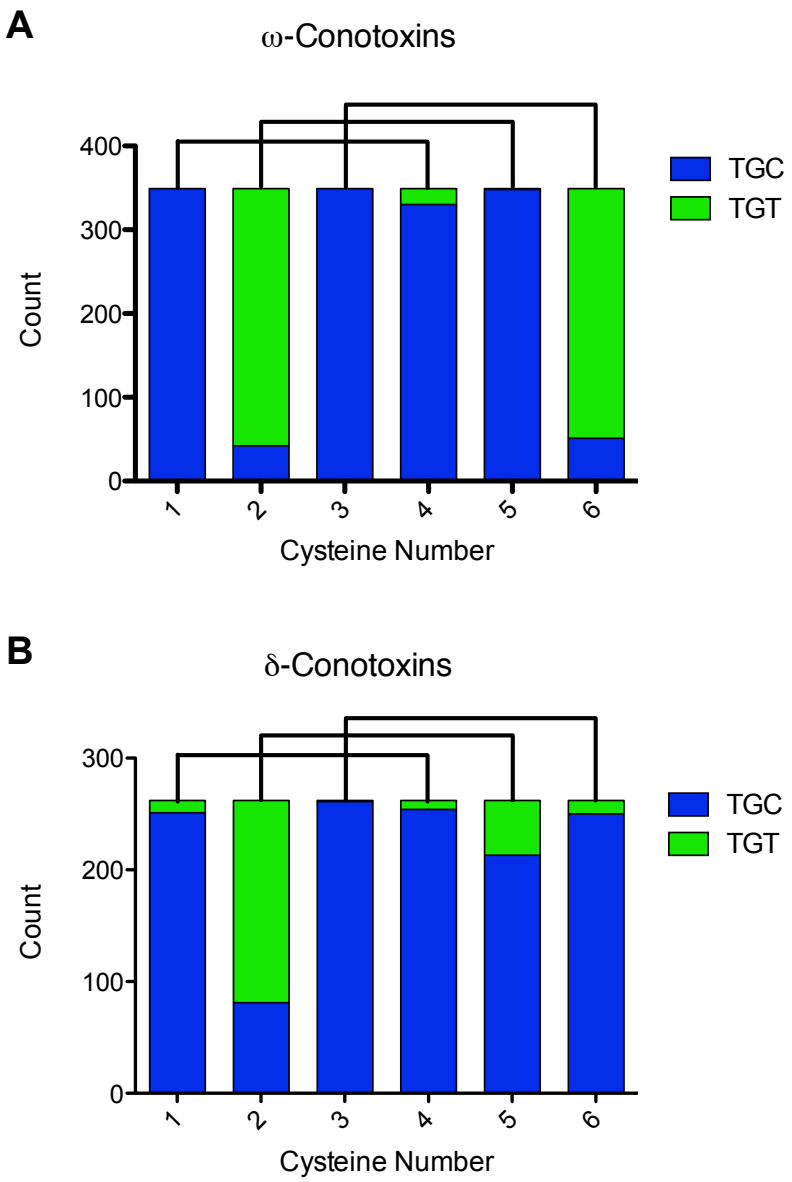


Figure 5.2. Codon usage statistics for the cysteine residues in two classes of four-loop conotoxin genes. **(a)** shows the codon bias for the ω-conotoxins, and **(b)** shows the codon bias for the δ-conotoxins. The lines above show the characteristic disulfide connectivity of the cysteine residues. While both ω- and δ-conotoxins maintain the same general cysteine scaffold, they are readily distinguished by the precursor sequences (both signal and propeptide), length of the intercysteine loops, as well as the tendency of the δ-conotoxins to be considerably more hydrophobic.

out-of-frame cysteine codons are not conserved (see Supplemental Information), suggesting that the frame-shifted psuedogenes have retained characteristic codons as a marker of their evolutionary history, not necessarily as the result of a direct mechanism. The 'block substitution' mechanism for generating hypervariability is consistent with feedback from folding efficiency, as it promotes hypervariability of the intercysteine loops, but avoids drastic perturbations to the overall length or cysteine scaffold of the toxin; furthermore, since both the template and target likely share a common ancestor, the cysteine codons (and disulfide scaffold) would remain preserved through numerous block substitution events. However, block substitution requires a template, and therefore is not sufficient to explain the extensive hypervariability observed in the intercysteine loops of conotoxins.

It is widely believed that the hypervariability of the intercysteine loops results from positive selection after gene duplication [18, 19, 20], based on the activity of the conotoxin at the target receptor [21, 18, 4, 22]; it follows that a mutation which confers a selective advantage with respect to activity will be preferentially retained over a mutation which is either neutral or detrimental to the activity. While activity does play a major role, the production/folding efficiency is also a critical factor; both act to direct the evolution of the amino acid sequences of the intercysteine loops, as well as the preservation of disulfide scaffolds and posttranslational modifications. Occasionally, the latter two are significant only to folding efficiency [23, 24, 25], although many changes affect both production efficiency and activity. Consequently, we propose that both activity and production efficiency be considered in models of conotoxin evolution.

In addition to holding the peptide in an active conformation [26, 27, 28], disulfides can also be critical to folding efficiency. In ω-GVIA, removing the Cys1-Cys16 disulfide bridge negligibly affects its activity on N-type calcium channels, but greatly reduces the peptide's propensity for the native connectivity [23]. Consequently, a disulfide can be conserved in order to preserve activity, or to preserve efficient folding, or both. It remains unclear whether a disulfide that is valuable only to folding efficiency acts by conformationally constraining the peptide as the other bridges form, or whether that cysteine pair drives thiol/disulfide exchange through a more productive folding pathway.

Post-translational modifications can also function to increase folding efficiency, without any measurable change in activity of the peptide. For instance, hydroxylation of proline in ω-MVIIC provides a twofold increase in folding rate, as well as a threefold increase in yield, but does not change the activity (as measured by *i.c.v.* injection into mice) [24]. Also, the γ-carboxylation of glutamate can introduce calcium-dependent improvements to

folding efficiency, without measurable effects on activity (also measured by *i.c.v.* injection into mice) [25]. While chaperones can offer some benefit to folding efficiency *in vivo*, their enhancement is largely limited to kinetics, as enzymes lower energy barriers, but do not change equilibria, although it is possible that kinetic enhancement to productive folding pathways could be coupled to secretion from the ER to effect a nonequilibrium folding outcome. The recent discovery of misfolded ImI in *Conus imperialis* venom [29] established that *in vivo* folding faces many of the same challenges to production efficiency as *in vitro* folding. Thus, chaperones do not necessarily drive *in vivo* oxidative folding to a single, "native" disulfide connectivity, further highlighting the importance of efficient oxidative folding.

In order to assess the extent to which chaperones could modulate the outcome of oxidative folding of conotoxins, we considered the ATP-dependence of chaperone-assisted folding. There is an established role for energy expenditure in efficient oxidative folding of proteins [30, 31, 32], which likely functions through BiP, the ER-resident 70 kDa heat shock protein [33]; although the extent to which this applies to small substrates—such as conotoxins—is unclear. Using microsome-assisted folding reactions to supply chaperones, we assayed the role that ATP expenditure plays in directing conotoxin folding. The results (Table 5.1) show a minor ATP dependence for the folding kinetics of certain conotoxins, suggesting that the presence of chaperones and expenditure of energy do not qualitatively change conotoxin folding. Consequently, the innate folding efficiency is likely to be the dominant factor in the efficient biosynthesis of conotoxins.

In light of these results, we propose that folding efficiency is one of the critical factors directing the evolution of conotoxins, due to its importance in an existing disulfide scaffold. This would imply that the only selective pressure–combined with evolutionary diversification from one (or a few) ancestral gene(s)–is at the amino acid level, and that the observed codon conservation [16, 13, 15, 17] is an artifact of the combination of cysteine conservation through selective pressure and phylogenetic conservatism.

Loss of a cysteine residue due to a mutation or frameshift could be significantly more detrimental than mutations that result in changes to the residues in the intercysteine loops. For instance, if one cysteine is removed, then the peptide will either create a disulfide dimer (homo- or hetero-dimer) [34], or it will retain a free thiol, leading to inter- and intramolecular disulfide shuffling in the venom duct, after secretion away from ER-based quality control machinery. If a pair of cysteines is lost, the resulting disulfide scaffold may loose activity, or may have a dramatically decreased folding efficiency [23]. Conversely, if a mutation

Table 5.1. ATP-dependence of microsome-assisted folding. As a model of *in vivo* oxidative folding in the ER, microsomes derived from rat liver were used to fold linear peptide (prepared as in [29]), using 40 μg total microsome protein per 1.5 nmol peptide timepoint. Only the timepoint with the largest difference is shown. ATP-depletion was accomplished with potato apyrase, which has been used previously to deplete ATP from *in vitro* reactions [32] (and references therein). Values given are Mean ± Standard Deviation. † No native peptide could be detected in the indicated condition. * Indicates that this peptide showed a statistically significant difference between microsomes that were ATP-depleted and untreated microsomes (paired, one-tailed t-test of triplicate data; $P < 0.05$).

Microsome-Assisted Folding				
Peptide	Folding Time	Amount Native (%)		
		ATP-Depleted Micros	Untreated Micros	No Micros
δ-PVIA*	24 hours	4.63 ± 0.09%	5.07 ± 0.33%	0%†
μ-SIIIA	30 min	2.090 ± 0.7727%	2.887 ± 0.2113%	1.02 ± 0.3143%
α-ImI*	8 min	1.617 ± 0.1804%	1.930 ± 0.0458%	0.4200 ± 0.01732%

changes a non-cysteine residue, the probability of retaining the important chemical attribute(s) is much greater. For example, if a given site must have a negatively charged amino acid for folding/activity, then there are two possibilities, encoded by four codons (GAT, GAC, GAA and GAG). Considering only single-base substitutions, then the probability to retain an amino acid that is negatively charged is one in three, or 33.3%. Since cysteine is the only residue capable of forming crosslinks, and can be encoded by two codons (TGT and TGC), the probability of retaining this functionality through a single-base substitution event is one in nine, or 11.1%. Hence, if a point mutation occurred, it would be much more likely to be detrimental at a cysteine site than elsewhere in the peptide because only silent mutations are viable, effectively selecting against non-silent mutation at cysteine sites.

To reexamine the codon conservation of cysteine in conotoxins, we sought a means to assess codon bias at each site that did not account for the encoded amino acid, in order to meaningfully compare codon conservation at cysteine to neighboring hypervariable residues. By using an internal reference for codon conservation, we accounted for the homology between sequences, allowing ready distinction between codon-level and amino-acid-level conservation. To accomplish this, we adapted the "effective number of codons" [35] to consider a group of related sequences site-by-site. In order to avoid assumptions regarding amino acids that were missing or underrepresented, we took a weighted average of the codon bias among amino acids that can be represented by the same number of codons (a 'redundancy class'), allowing analysis of codon bias independent of the amino acid encoded by each codon (details are in the Supplemental Information). The codon bias for each cysteine is compared to the codon bias in the intercysteine loops for other residues that can be encoded by two codons in Figure 5.3. Thus, when we accounted for homology among the members of a conotoxin family, the codon bias for cysteine became rather subtle, indicative of amino-acid-level conservation, rather than site-specific bias for a particular cysteine codon.

Phylogenetic analyses can also be used to account for homology among members of a conotoxin family. Under the assumption that the O1 superfamily of conotoxin genes evolved from a single ancestral sequence, we constructed a phylogenetic tree to chart the divergence of these genes. Figure 5.4 shows the phylogenetic tree for a representative subclade of the δ-conotoxins; the trees for all sequences used in this study are shown in the Supplemental Information. Mapping cysteine codon usage onto the tree results in six copies of the tree, each marking codon usage for a particular cysteine (Figure 5.4, C1-C6). The trees for cysteines 2 and 5, in particular (but see also Supplemental Information), clearly show that

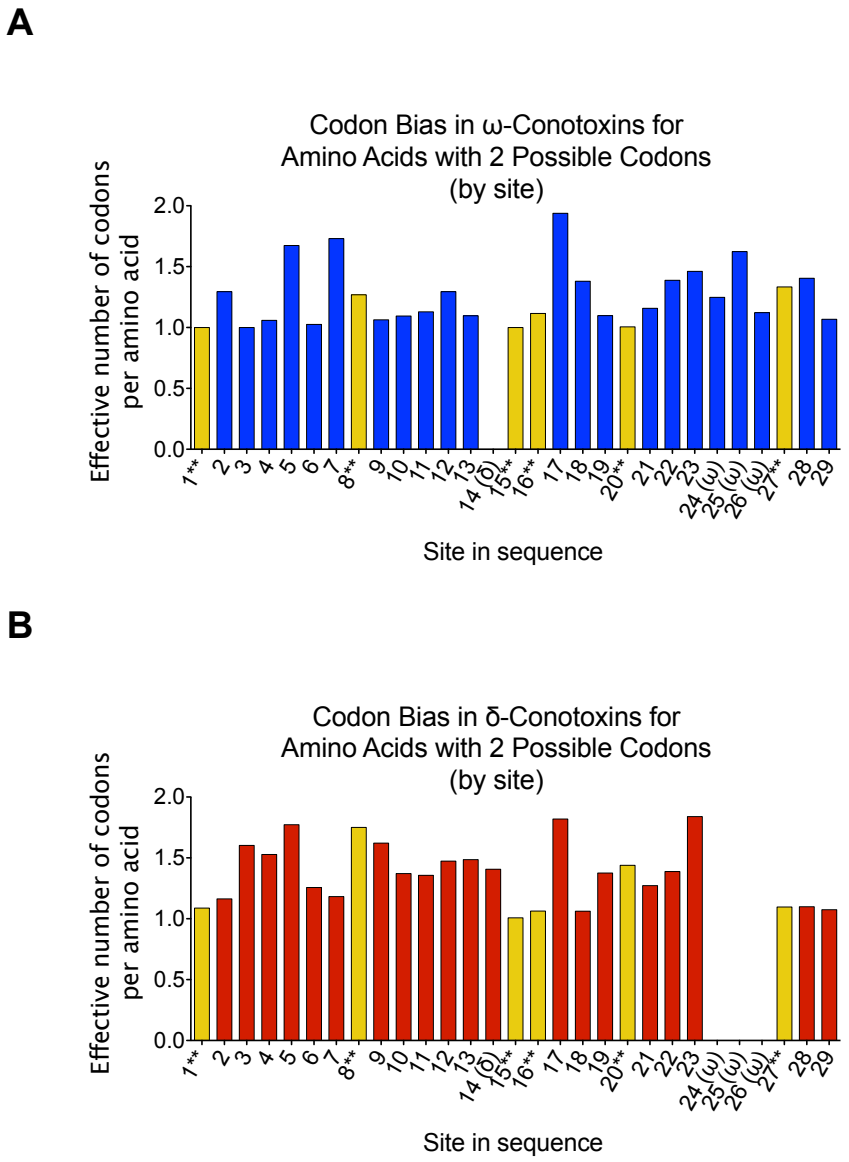


Figure 5.3. Codon bias for ω- **(a)** and δ- **(b)** Conotoxins. Values presented are the weighted average of N_e values (*sensu* [35], Eq. 2) for codons encoding amino acids that can be represented by two (and only two) codons. Because the number of residues in the intercysteine loops differ between the ω- and δ-Conotoxins, sites that are specific to one class of conotoxins are followed by the class to which they are specific (e.g., site 14(δ) only shows data for δ-conotoxins). Yellow bars and ** indicate cysteine sites. Analysis was only conducted on sites with at least 100 sequences in the alignment, at least 30 of which were residues that could be encoded by two codons. A value of 1 indicates that only a single codon is used, and a value of 2 indicates that there is no codon bias among the two possible codons for each residue. The same data for the other redundancy classes are presented in the Supplemental Information.

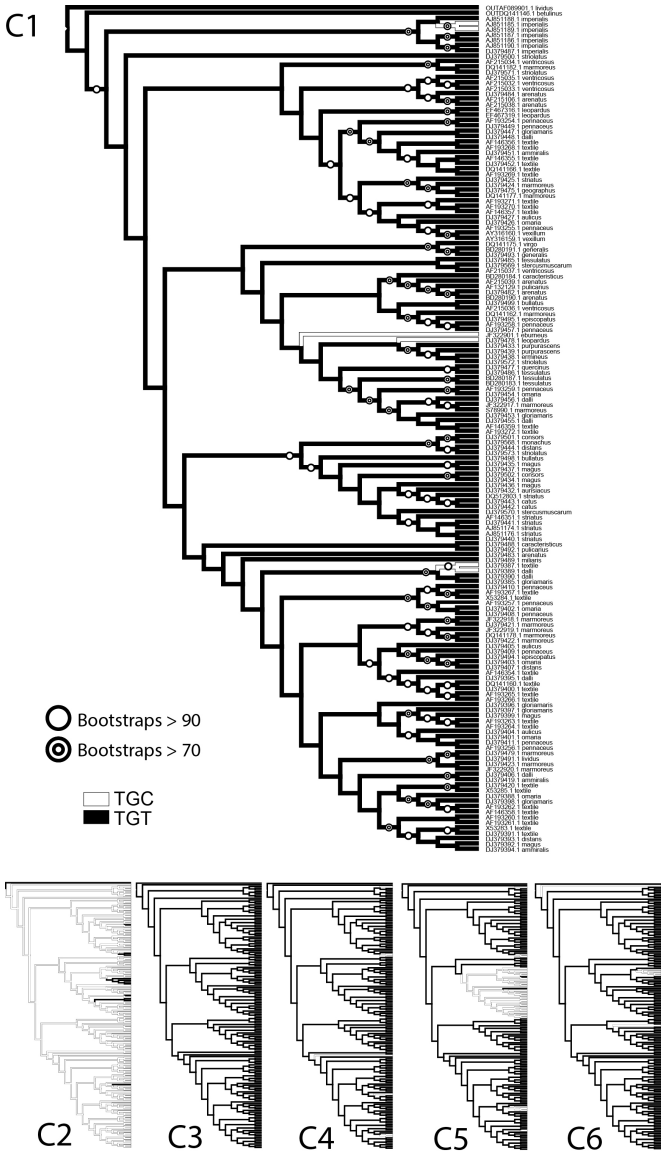


Figure 5.4. Phylogenetic trees of a represenatitve subclade of δ-conotoxin genes used in this study (all O1 conotoxins with signal, propeptide and mature sequence with a C—C—CC—C—C cysteine scaffold published in GenBank), showing the codon usage patterns for each cysteine. Tree reconstruction was performed on the Cipres Science Gateway (http://www.phylo.org/portal2), using the RAxML-HPC2 on TG tool; robustness of the nodes was assessed with a bootstrap analysis (1000 replicates). The evolution of the cysteine codons was assessed with Meqsuite V2.74 [36], using the option "tracing character history" and the parsimony ancestral reconstruction method. Each tree shows the codon usage for a single cysteine, and is labeled accordingly. Shaded lines indicate TGT, and hollow lines indicate TGC. Bootstrap values are also shown on the tree for Cysteine 1. Each branch in the tree for Cysteine 1 is labeled with the GenBank accession number of the sequence and the *Conus* species. Phylogenetic trees for the entire O1 superfamily are shown in the Supplemental Information.

mutations from one codon to the other arose independently several times; furthermore, in several cases (e.g., Figure 5.4 C5 tree, TGC clade, in white), once a mutation occurs, it is transmitted to the descendant sequences, suggesting that there is no counter-selection to revert to the canonical cysteine codon at that site. Taken together, these results are compatible with a scenario where cysteine is conserved at the amino acid level, and where the observed, but weak, codon level conservation is the result of a phylogenetic conservatism linked to the strength of the evolutionary pressure to preserve cysteine. In the M-Superfamily of conotoxins, it has been proposed that the divergence into distinct branches with distinct disulfide connectivities (and length of intercysteine loops) was concurrent with a switch in codon-preference at each site [37], further supporting that the cysteine codons are retained by evolutionary constraints on folding, and that the evolutionary pressure acts to retain cysteine residues, but does not act directly on the codon level.

We propose that the evolutionary pressures acting on the venom components of cone snails involve an interplay of production/folding efficiency, activity at the target receptor, and the relative importance of that activity to effective prey immobilization. Mutations that enhance or detriment folding efficiency or activity can be assessed with respect to energetic cost per activity unit, allowing a combined analysis of the potential cost and value of a mutation with respect to evolutionary pressures that cone snails experience. For example, if a peptide with an indispensable activity is mutated such that it is twofold more active, but folds one fourth as efficiently, then the snail must produce two times as much of that peptide (costing twice the energy) to effect the same result upon envenomation. However, in reality, the functional redundancy of *Conus* venom dampens this variability, effectively buffering the energetic cost of venom production against drastic perturbations that result from changes to the folding efficiency of individual conotoxins. Because mutations at cysteine residues are more likely to result in decreased folding efficiency and/or activity, there is significant pressure to retain cysteine residues, which in turn results in substantial codon bias over the course of evolutionary diversification of conotoxin gene families associated to phylogenetic conservatism. Using a measure that corrects for homology among conotoxin genes, we have shown that the established codon bias speaks more to the evolutionary history of the conotoxin genes than explicit codon-level conservation of cysteine. Thus, the site-specific codon-level conservation of cysteine is driven by amino-acid level conservation of cysteine's unique properties, and the importance of efficient oxidative folding in conotoxin biosynthesis.

5.1.3 Acknowledgements

We thank Prof. Toto Olivera for his generous support and discussions. This work was funded by NIH Program Project Grant GM 48677. G.B. acknowledges support from the Johnson & Johnson Pharmaceutical Research and Development funding.

5.2 Supplemental Information

5.2.1 Out of Frame Cysteine Codons

Figure 5.S1 shows the extent of conservation of out of frame cysteine codons, as compared to the cysteine codons that are in frame. This is shown separately for the δ- and ω-conotoxins. Interestingly, there does appear to be a minor conservation of out of frame cysteine codons at sites 9-11 in the ω-conotoxins; however, this constitutes a small fraction of the total number of sequences.

5.2.2 Determination of the Effective Number of Codons per Amino Acid by Site

5.2.2.1 Rationale

In order to consider codon-level conservation within a family of homologous sequences (especially from a single genus, such as *Conus*), it is important to be able to account for the homology that exists among this group of sequences. However, accounting for this homology is somewhat nontrivial, as the ancestral sequence is usually unknown, and in some cases cannot be readily determined.

While previous studies have not accounted for homology [13], and consequently resulted in extreme site-specific codon bias ($P < 1 \cdot 10^{-4}$), we decided to reassess the level of codon bias in a manner that accounts for homology between the sequences considered. However, because we do not necessarily know the origins or extent of that homology, we do not want to directly include any statements of what that homology may be in our analyses.

Herein we present a tool to assess codon bias site-by-site across a family of sequences, normalizing out the identity of any encoded amino acid, and without incorporating any assumptions about amino acids that are missing or underrepresented. However, in removing assumptions about missing amino acids, we cannot readily compile the redundancy classes (amino acids that can be encoded by the same number of codons, either 1, 2, 3, 4 or 6) into a single numeric output, and therefore must consider each redundancy class separately.

We have based this novel tool on the 'Effective Number of Codons,' originally introduced by Frank Wright [35]. In his work, the objective was to look at codon bias in a gene, in order to assess translational efficiency; other codon bias analyses have taken the same approach

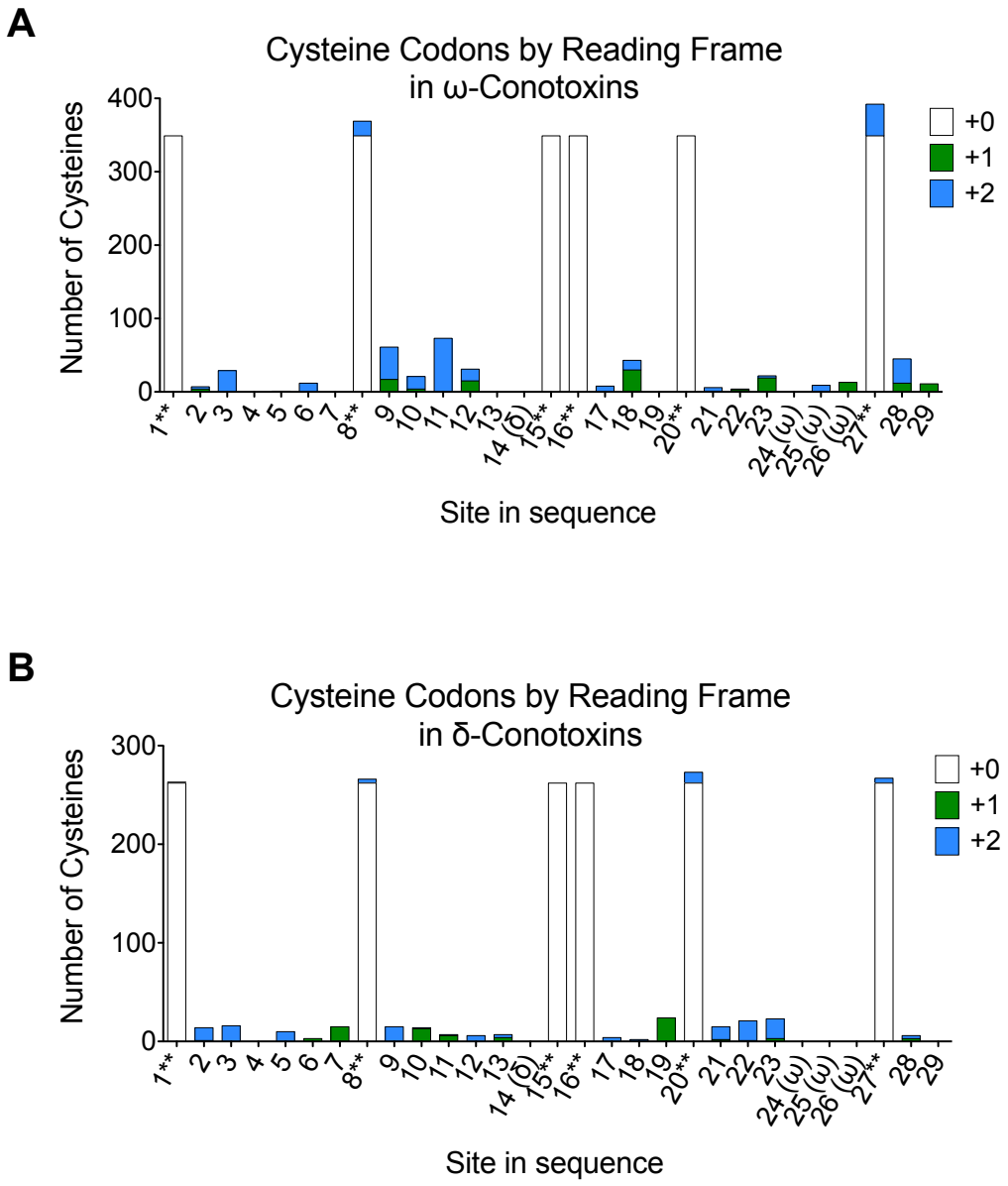


Figure 5.S1. Comparison of in-frame and out-of-frame cysteine codons for ω- (A) and δ- (B) Conotoxins. White bars show in-frame cysteine codons, green bars show cysteine codons in the +1 position, and blue bars show cysteine codons in the +2 position. 'Site in sequence' uses the same numeric designations as in Figure 5.3, and are similarly designated, with sites that are specific to one class of sequence being marked with that class (e.g., 14(δ) contains data for only the δ-conotoxins). ** indicates cysteine sites, which are also readily identifiable by the large white bars that indicate in-frame cysteine codons.

[38]. One concern is whether these methods can be used to analyze a relatively small sample set; of five prominent methods, the 'Effective Number of Codons' was among the least biased by short gene sequences [38], analogous to a relatively small set of sequences when considered site-by-site.

While our method seems intuitively similar to measuring a D_n/D_s ratio because both measure the correlative capacity of amino acid to a codon at each site, the two differ in a few critical ways. D_n/D_s is used to determine the evolutionary pressures on a gene, and what type of selection (purifying or adaptive) existed to result in the changes in that gene [39]. In contrast, our method does not have this power; rather, it is a means to assess codon bias by site. Our method should serve as a bridge between measures of evolutionary pressures and conventional measures of codon bias; by allowing a site-specific measure of codon variability, it can provide evolutionary hints as to which portions of a peptide sequence are under different selective pressures.

5.2.2.2 Overview of Adaptations to Effective Number of Codons

In the original treatment, the 'Effective Number of Codons' was determined by first finding the 'homozygosity' of each amino acid, which was then averaged within each redundancy class, and then the sum of the reciprocals is taken, weighted by the number of amino acids in each redundancy class. The result is a number between 20 and 61, which indicates the number of codons that are effectively used within that gene, accounting for the extent of bias for each of the 20 amino acids. If an amino acid is not present or underrepresented in a gene, then it is assumed that the codon bias for that amino acid would be the same as for the other amino acids within that redundancy class.

In our treatment, we sought to eliminate any assumptions about amino acids that were not present or underrepresented, as the sequence alignment for conotoxins contains numerous cysteine-only sites. Consequently, when considering each site, we instead used the effective number of codons *for each amino acid*, and then averaged these, weighting by the prevalence of that amino acid at that site (within that redundancy class). This results in a set of numbers for each redundancy class, at each site, allowing comparisons of codon bias between two sites using values that are intrinsically meaningful (for instance, for Site 4 in ω-conotoxins, amino acids that can be encoded by two codons are actually encoded by 1.059 codons, indicating marked codon bias at this site).

However, it is worth noting that comparisons between different redundancy classes cannot be readily made, as a consequence of our removing any assumptions about amino

acids that are not present or underrepresented. Consequently, we cannot readily compare the codon bias at Site 4 in ω-conotoxins for amino acids that can be encoded by two codons (effectively, 1.059 codons) with the codon bias at the same site for amino acids that can be encoded by six codons (effectively, 2.145 codons), despite both showing notable codon bias and their representing the same site in the same family of sequences.

5.2.2.3 Theoretical Buildup

For clarity, we use different indices to designate different 'levels' of information: i is for codon-level information; j is for amino acid level information; k is for redundancy class information. Although we consider only one at a time, in some cases more than one index is shown in order to clearly show the origins of each calculation.

To begin, each amino acid is placed into a "redundancy class," which designates the number of distinct, synonymous codons that can encode that amino acid. In the standard genetic code, two amino acids (M, W) can be encoded by one codon; nine amino acids (C, N, D, Q, E, H, K, F, Y) can be encoded by two codons; isoleucine can be encoded by three codons; five amino acids (A, G, P, T, V) can be encoded by four codons, and three amino acids (R, L, S) can be encoded by six codons. This method can be readily applied to any non-standard genetic code, as long as the amino acids encoded by each codon are known.

The number of occurrences of each codon are then counted, giving n_i. Thus, $n = \sum_i n_i$ is the total number of occurrences of the selected amino acid, and the probability of codon i is simply the number of occurrences of that codon divided by the total number of occurrences of that amino acid, $p_i = \frac{n_i}{n}$. So, by definition, $\sum_i p_i = 1$.

For each amino acid, we then determine the 'homozygosity' of amino acid j, $\hat{F}_j$, as in [35]:

$$\hat{F}_j = \frac{n \cdot \sum_{i=1}^{k} p_i^2 - 1}{n - 1} \tag{5.S1}$$

It is worth noting at this point that because $p_i \in [0, 1]$, $p_i^2 \leq p_i$, and thus $\sum_i p_i^2 \leq 1$, indicating that $0 \leq \hat{F}_j \leq 1$; however, if $\hat{F}_j = 0$, then it is disregarded, as this can only result from very few instances of amino acid j at that site ($n \leq 2$), indicating that amino acid is underrepresented and should be excluded from the analysis at this site.

The homozygosity of each amino acid ($\hat{F}_j$) is readily translated into the number of equally frequent codons that would show the same degree of homozygosity for that amino acid ($\hat{N}_e(j)$) by taking the reciprocal (as in [35]),

$$\hat{N}_e(j) = \frac{1}{\hat{F}_j}. \tag{5.S2}$$

Here is where we diverge from Wright's treatment (except that we were grouping sets of codons by site across a family of genes, rather than in a gene). Since there is significant bias towards certain residues at certain sites in our alignment (for instance, cysteine residues at cysteine sites, or the site used in the sample calculation), we decided to use a weighted average of $\hat{N}_e(j)$ values, averaged within each redundancy class (however, the redundancy classes cannot be compared, in order to avoid assumptions about amino acids that are not present or underrepresented).

In order to determine a weighted average, we define a new coefficient, p_j, as the fraction of redundancy class k that encodes amino acid j (this is analogous to p_i, except that it uses amino acid level information, indicated by the j index), in terms of counts of the codons at that site. Using this, we can determine the number of equally frequent codons that would show the same level of homozygosity for that redundancy class, or the effective number of codons per amino acid at that site (within a redundancy class),

$$\Lambda_k = \sum_{j=1}^{l} (p_j \cdot \hat{N}_e(j)) \tag{5.S3}$$

where l is the number of nonexcluded members of the redundancy class (exclusion as described for Eq. 5.S1).

Λ_k can be interpreted as the effective number of (or, number of equally frequent) codons that are actually used at that site, for that redundancy class. So for redundancy class k, a value of 1 would indicate that only one codon was used for each amino acid, whereas a value of k would indicate that every available codon was used indiscriminately for each amino acid.

5.2.2.4 Sample Calculation

For the sample calculation, we use the data from site 10 (counting from the first cysteine, as in Figure 5.3 in the main text) of the published ω-conotoxin genes in a multiple sequence alignment.

First, the codons for each amino acid are counted, and placed into bins based on the redundancy classes. We will deal with each redundancy class separately; we will not discuss the non-redundant class (M and W) because these can only be encoded by one codon, and therefore $\Lambda_1 = 1$ in all cases.

We will first address redundancy class 2, whose values are shown in Table 5.S1. Not all amino acids are shown, but if there is an occurrence of that amino acid, all possible codons for that amino acid are shown.

The next step is to determine the homozygosity for each amino acid. Using the data for glutamine because there are instances of both codons, $p_{CAA} = \frac{9}{10}$ and $p_{CAG} = \frac{1}{10}$, so,

$$\hat{F}_Q = \frac{10 \cdot \sum_{i=1}^{2} p_i^2 - 1}{10 - 1} = \frac{10((\frac{9}{10})^2 + (\frac{1}{10})^2) - 1}{10 - 1} = \frac{10(\frac{82}{100}) - 1}{10 - 1} = \frac{7.2}{9} = 0.8$$

For any amino acid that has occurrances of only one of the codons, the homozygosity is one, since one of the p_i values is one, and the other(s) are 0. So, doing the calculation for tyrosine will tell us the homozygosity values for all of the twofold redundant amino acids at site 10 in ω-conotoxins:

$$\hat{F}_Y = \frac{24 \cdot \sum_{i=1}^{2} p_i^2 - 1}{24 - 1} = \frac{24((\frac{22}{24})^2 + (\frac{2}{24})^2) - 1}{24 - 1} = \frac{24(\frac{122}{144}) - 1}{24 - 1} = 0.84058.$$

Now, adding the homozygosity, the number of equally frequent codons (the reciprocal of $\hat{F}_j$), and the fraction of each redundancy class that is the given amino acid (p_j) to Table 5.S1 results in Table 5.S2.

And, taking the product of the last two columns, and summing the results, we get the number of equally frequent codons per amino acid at site 10 of the ω-conotoxins, for the amino acids that can be represented by two (and only two) codons,

$$\Lambda_2 = \sum_{j=1}^{l} (p_j \cdot \hat{N}_e(j)) = 1(\frac{7}{75} + \frac{2}{75} + \frac{11}{75} + \frac{13}{75} + \frac{8}{75}) + 1.25(\frac{10}{75}) + 1.1897(\frac{24}{75}) = 1.094.$$

To do the same calculation for the other redundancy classes, we show only the resultant final table; the calculations at each step are exactly the same as those above, but the sums are longer because of the increased number of codons for each amino acid.

For the redundancy class that includes only isoleucine, only one codon was used. These calculations are not shown because they reduce to the trivial case.

Note that there is only one instance of threonine, so this data point is excluded from the analysis.

The data for amino acids that can be encoded by four possible codons is shown in Table 5.S3.

The same can be done for the amino acids encoded by six possible codons, and is shown in Table 5.S4.

Table 5.S1. Codon usage counts for site 10 in multiple sequence alignment of published ω-conotoxin genes.

Amino Acid	Codon	Count
N	AAT	7
	AAC	0
D	GAT	2
	GAC	0
Q	CAA	9
	CAG	1
H	CAT	11
	CAC	0
K	AAA	0
	AAG	13
F	TTT	8
	TTC	0
Y	TAT	22
	TAC	2

Table 5.S2. $\hat{N}_e(j)$ at site 10 of a multiple sequence alignment of published ω-conotoxins for amino acids that can be represented by two possible codons and their relative weights.

Amino Acid	Codon	Count	$\hat{F}_j$	$\hat{N}_e(j)$	p_j
N	AAT	7	1	1	$\frac{7}{75}$
	AAC	0			
D	GAT	2	1	1	$\frac{2}{75}$
	GAC	0			
Q	CAA	9	0.8	1.25	$\frac{10}{75}$
	CAG	1			
H	CAT	11	1	1	$\frac{11}{75}$
	CAC	0			
K	AAA	0	1	1	$\frac{13}{75}$
	AAG	13			
F	TTT	8	1	1	$\frac{8}{75}$
	TTC	0			
Y	TAT	22	0.84058	1.1897	$\frac{24}{75}$
	TAC	2			

Table 5.S3. Determination of the effective number of codons per amino acid for amino acids with four possible codons (Λ_4) at site 10 in a multiple sequence alignment of ω-conotoxins.

Redundnacy Class 4						
Amino Acid	**Codon**	**Count**	$\hat{F}_j$	$\hat{N}_e(j)$	p_j	Λ_4
A	GCT	25	0.92308	1.0833	$\frac{26}{120}$	1.1883
	GCC	0				
	GCA	1				
	GCG	0				
G	GGT	55	1	1	$\frac{55}{120}$	
	GGC	0				
	GGA	0				
	GGG	0				
P	CCT	8	0.49474	2.0213	$\frac{20}{120}$	
	CCC	0				
	CCA	0				
	CCG	12				
T	ACT	1	N/A	N/A	Not Incl.	
	ACC	0				
	ACA	0				
	ACG	0				
V	GTT	19	1	1	$\frac{19}{120}$	
	GTC	0				
	GTA	0				
	GTG	0				

Table 5.S4. Determination of the effective number of codons per amino acid for amino acids with six possible codons (Λ_6) at site 10 in a multiple sequence alignment of ω-conotoxins.

Redundnacy Class 6						
Amino Acid	**Codon**	**Count**	$\hat{F}_j$	$\hat{N}_e(j)$	p_j	Λ_6
R	CGT	28	0.51800	1.9305	$\frac{72}{130}$	1.9601
	CGC	0				
	CGA	0				
	CGG	0				
	AGA	0				
	AGG	0				
L	TTA	3	0.33846	2.9545	$\frac{26}{130}$	
	TTG	5				
	CTT	4				
	CTC	14				
	CTA	0				
	CTG	0				
S	TCT	2	0.82056	1.2187	$\frac{32}{130}$	
	TCC	0				
	TCA	1				
	TCG	0				
	AGT	29				
	AGC	0				

5.2.2.5 Data for Redundancy Classes 4 and 6

In order to compare the relative conservation of cysteine to the other amino acids in the intercysteine loops, only the twofold redundant codons were used for analysis in the main paper. However, every redundancy class was analyzed for each site, in order to assess the extent of codon conservation within each redundancy class.

For almost every site, isoleucine was not sufficiently prevalent to merit analysis of the redundancy class with three possible codons. While these calculations were performed, they are excluded from this presentation because we believe them to be artificially skewed by small sample size effects.

In Figures 5.S2 and 5.S3, the axes are scaled to the maximum possible value, which would represent indiscriminate usage of every possible codon for that amino acid. In most cases, the actual values are markedly lower than the maximum possible value, indicating that there is a global codon bias across all redundancy classes at almost every site. The bars are also color-coded by the number of sequences included in the analysis at that site, to indicate the risk of small sample size effects at the specified site for the specified redundancy class. Red or blue indicates that at least 30 sequences were analyzed; grey indicates that between 15 and 30 sequences were analyzed; white indicates that fewer than 15 sequences were analyzed.

5.2.3 Phylogenetic Trees for O1 Superfamily

The following phylogenetic trees show the toxin diversification for the O1 superfamily, using all published sequences (see Section 5.2.4). As mentioned in the main text, it is evident from these trees that there is no counter-selection to revert to the canonical cysteine residue at that site. Tree reconstruction was performed on the Cipres Science Gateway (http://www.phylo.org/portal2), using RAxML-HPC2 on TG tool; robustness of the nodes was assessed with a bootstrap analysis (1000 replicates). The evolution of cysteine codons was assessed with Mesquite V2.74 [36], using the option "tracing character history" and the parsimony ancestral reconstruction method. Each tree shows the codon usage for a single cysteine.

The tree for cysteine 1 is shown in Figure 5.S4; cysteine 2 is shown in Figure 5.S5; cysteine 3 is shown in Figure 5.S6; cysteine 4 is shown in Figure 5.S7; cysteine 5 is shown in Figure 5.S8; cysteine 6 is shown in Figure 5.S9. Each of these trees differs from the others only in which lines are hollow/shaded, indicating which cysteine codon is used at that site for the indicated conotoxin, allowing facile comparison of the evolutionary relationships of the codon usage patterns for each cysteine.

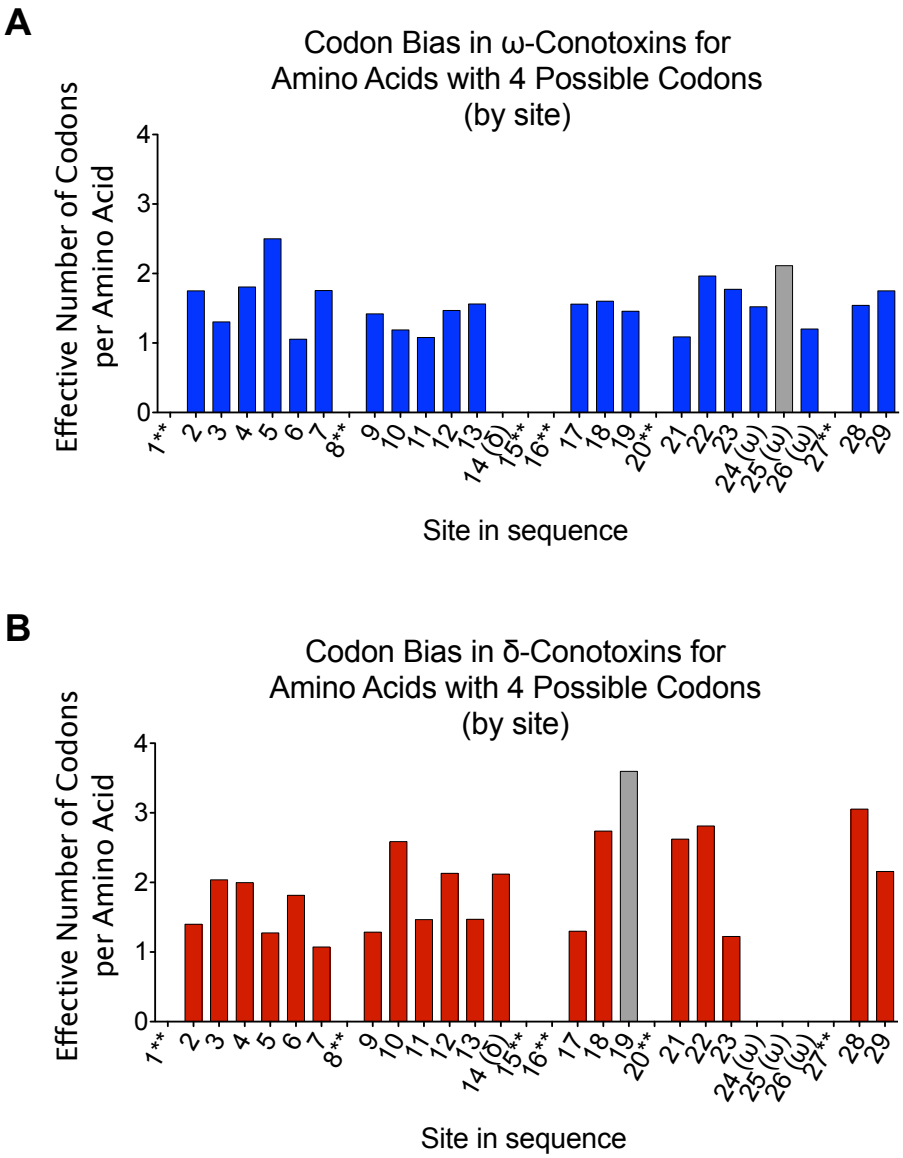


Figure 5.S2. Codon bias in the ω- (A) and δ- (B) Conotoxins, showing the bias for amino acids that can be represented by four possible codons. Because the length of the intercysteine loops differ between these two classes of conotoxins, residues that are specific to each class are marked with the class to which they are specific, as in Figure 5.3. ** indicates cysteine sites, which are blank because there are two possible codons for cysteine. A value of 1 indicates that only one of the four possible codons is used for each residue, and a value of 4 would indicate that each of the four possible codons for each amino acid are used indiscriminately. Axes are scaled to the maximum possible value. Bars are color-coded to the size of the sample set being analyzed (the number of sequences with fourfold redundant codons at that site): blue/red indicates at least 30 sequences and grey indicates between 15 and 30 sequences.

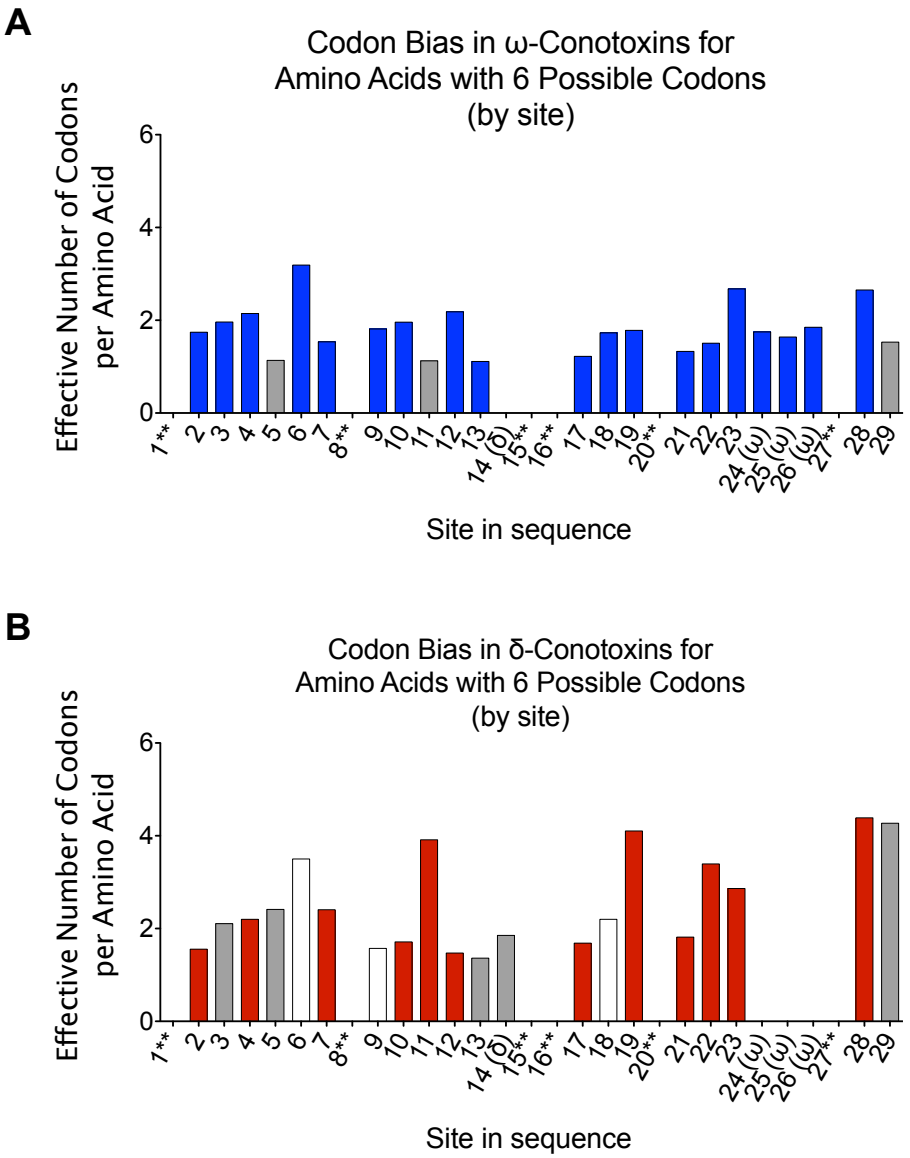


Figure 5.S3. Codon bias in the ω- (A) and δ- (B) Conotoxins, showing the bias for amino acids that can be represented by six possible codons. Because the length of the intercysteine loops differ between these two classes of conotoxins, residues that are specific to each class are marked with the class to which they are specific, as in Figure 5.3. ** indicates cysteine sites, which are blank because there are two possible codons for cysteine. A value of 1 indicates that only one of the six possible codons is used for each residue, and a value of 6 would indicate that each of the six possible codons for each amino acid are used indiscriminately. Axes are scaled to the maximum possible value. Bars are color-coded to the size of the sample set being analyzed (the number of sequences with sixfold redundant codons at that site): blue/red indicates at least 30 sequences, grey indicates between 15 and 30 sequences, and white indicates fewer than 15 sequences.

Figure 5.S4. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 1. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

Figure 5.S5. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 2. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

Figure 5.S6. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 3. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

Figure 5.S7. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 4. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

Figure 5.S8. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 5. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

Figure 5.S9. Phylogenetic tree for the O1 superfamily, showing codon usage for Cysteine 6. Hollow bars indicate TGC; shaded bars indicate TGT. Bootstrap values are shown at each branching point. GenBank Accession numbers and species of origin are indicated for each sequence.

5.2.4 Sequences Used for Analysis

5.2.4.1 ω-Conotoxin Sequences

We used published sequences of ω-conotoxins for these analyses. The GenBank accession numbers for these sequences are shown in Table 5.S5 (349 total sequences).

5.2.4.2 δ-Conotoxin Sequences

We used published sequences of δ-conotoxins for these analyses. The GenBank accession numbers for these sequences are shown in Table 5.S6 (262 total sequences).

Table 5.S5. GenBank accession numbers of all sequences of ω-conotoxins used in this study. This table continues on the following page.

GenBank Accession Numbers of ω-Conotoxin Sequences Analyzed					
DQ141146.1	DQ141149.4	AF090041.1	AF090042.1	AF090043.1	AF090044.1
AF090045.1	AF090046.1	AF090047.1	AF090048.1	AF090050.1	AF090051.1
AF090052.1	AF090054.1	AF090073.1	DQ644547.1	EF108269.1	EF108272.1
AF090055.1	AF090056.1	AF090057.1	AF090058.1	AF090059.1	AF090060.1
AF090061.1	AF090063.1	AF090064.1	AF090065.1	AF090066.1	AF090067.1
AF090068.1	AF090069.1	AF090071.1	AF090072.1	DQ644546.1	EU423370.1
EU423371.1	EU423372.1	EU423373.1	EU423374.1	EU423375.1	EU423376.1
EU423377.1	EU423378.1	EU423379.1	EU423380.1	EU423381.1	EU423382.1
EU423383.1	AF215040.1	AF215041.1	AF215042.1	DJ379450.1	AF089971.1
AF089972.1	AF089973.1	AF089974.1	AF089975.1	AF089976.1	AF089977.1
AF089978.1	AF089979.1	AF089980.1	AF089981.1	AF089982.1	AF089983.1
AF089984.1	AF089985.1	AF089986.1	AF089987.1	AF089988.1	AF089989.1
AF089990.1	AF089991.1	AF089992.1	AF089993.1	AF089995.1	AF089996.1
AF089997.1	AF089998.1	AF089999.1	AF090000.1	AF090001.1	AF090002.1
AF090003.1	AF090004.1	AF090005.1	AF090006.1	AF090007.1	AF090008.1
AF090009.1	AF090010.1	AF090011.1	AF090012.1	AF090013.1	AF090014.1
AF090015.1	AF090016.1	AF090017.1	AF090018.1	AF090019.1	AF090020.1
AF090021.1	AF090022.1	AF090023.1	AF090024.1	AF090025.1	AF090026.1
AF090027.1	AF090028.1	AF090029.1	AF090030.1	AF090031.1	AF090032.1
AF090033.1	AF090035.1	AF090036.1	AF090037.1	AF090038.1	AF090039.1
AF090053.1	AF090049.1	AF174268.1	AF174269.1	AF174271.1	AF174272.1
AF174273.1	AF174274.1	AF174275.1	AF174276.1	AF174277.1	AF174278.1
AF174279.1	AF174280.1	AF174281.1	AF174282.1	AF174283.1	AF174284.1
AF174285.1	AF174286.1	AF174287.1	FJ613506.1	FJ613512.1	FJ613513.1
FJ613517.1	FJ716817.1	FJ716824.1	FJ804530.1	FJ804531.1	FJ804532.1
FJ804533.1	FJ804534.1	FJ804535.1	FJ804536.1	FJ834437.1	AF090074.1
AF090075.1	AF090076.1	AF090077.1	AF090078.1	AF090079.1	AF090080.1
DQ644543.1	DQ644544.1	DQ644545.1	DQ644548.1	DQ644549.1	EF108267.1
EF108268.1	EF108270.1	EF108271.1	EF108273.1	EF108274.1	EF108275.1
EF108276.1	EF108277.1	EF108278.1	EF108279.1	EF108280.1	EF108281.1
EF108282.1	EF108283.1	EF108284.1	EF108285.1	EF108286.1	EF108287.1
AF132130.1	AF215060.1	AF215057.1	DD012749.1	DD012687.1	DD012688.1
DD012685.1	DD012686.1	DD012689.1	DD012778.1	DD012779.1	DD012780.1
DQ141148.1	DQ345369.1	AJ851170.1	AF146349.1	AJ851172.1	DD012762.1
DD012714.1	AF174214.1	AF174215.1	AF174216.1	AF174217.1	AF174218.1
AF174219.1	AF174220.1	AF174221.1	AF174222.1	AF174223.1	AF174224.1
AF174225.1	AF174226.1	AF174227.1	AF174228.1	AF174229.1	AF174230.1
AF174231.1	AF174232.1	AF174233.1	AF174235.1	AF174236.1	AF174237.1
AF174238.1	AF174239.1	AF174240.1	AF174241.1	AF174242.1	AF174243.1
AF174244.1	AF174245.1	AF174246.1	AF174247.1	AF174248.1	AJ851175.1
AF146350.1	DD012769.1	DD012696.1	DD012700.1	DD012775.1	DD012711.1

Table 5.S5, *continued*. *Continued from previous page.* GenBank accession numbers of all sequences of ω-conotoxins used in this study.

GenBank Accession Numbers of ω-Conotoxin Sequences Analyzed					
DD012719.1	DJ379574.1	DD012709.1	AF480312.1	DD012713.1	AF174267.1
DD012697.1	AF146348.1	AF146346.1	DD012710.1	DD012751.1	M84612.1
FJ959111.1	DD012715.1	DD012741.1	DD012745.1	DD012746.1	DD012764.1
DD012767.1	DD012726.1	DD012699.1	DD012701.1	DD012776.1	DD012708.1
DD012755.1	DD012759.1	DD012773.1	DD012716.1	DD012718.1	DD012724.1
DD012725.1	DD012781.1	BD241823.1	DD012772.1	DD012771.1	DD012774.1
DD012750.1	DD012722.1	DD012752.1	AF480313.1	AF480314.1	AF480315.1
AY236862.1	AY236863.1	AF174249.1	AF174250.1	AF174251.1	AF174252.1
DD012760.1	DD012720.1	DD012698.1	AJ851171.1	DD012712.1	DD012742.1
DD012777.1	DD012736.1	DD012717.1	DD012765.1	DD012761.1	EF467318.1
DD012733.1	DD012734.1	DD012735.1	DQ345370.1	DD012753.1	DD012754.1
AF174253.1	AF174254.1	AF174255.1	AF174256.1	AF174257.1	AF174258.1
AF174259.1	AF174260.1	AF174262.1	AF174263.1	AJ851173.1	DD012768.1
AF146347.1	AF174264.1	AF146360.1	AF146361.1	AF174265.1	AF174266.1
AF215043.1	AF215044.1	AF215045.1	AF215056.1	AF215058.1	DD012690.1
DD012691.1	DD012702.1	DD012684.1	AF215059.1	DD012721.1	DD012756.1
DD012757.1	DD012758.1	AF548000.1	DQ141151.1	DQ141169.1	DQ141173.1
DD012744.1					

Table 5.S6. GenBank accession numbers of all sequences of δ-conotoxins used in this study. This table continues on the following page.

GenBank Accession Numbers of δ-Conotoxin Sequences Analyzed					
AF089901.1	AF089902.1	AF089903.1	AF089904.1	AF089905.1	AF089906.1
AF089907.1	AF089908.1	AF089910.1	AF089911.1	AF089912.1	AF089913.1
AF089914.1	AF089915.1	AF089916.1	AF089917.1	AF089918.1	AF089919.1
AF089920.1	AF089921.1	AF089922.1	AF089923.1	AF089924.1	AF089925.1
AF089926.1	AF089927.1	AF089928.1	AF089929.1	AF089930.1	AF089931.1
AF089932.1	AF089933.1	AF089934.1	AF089935.1	AF089936.1	AF089937.1
AF089938.1	AF089939.1	AF089940.1	AF089941.1	AF089942.1	AF089943.1

Table 5.S6, *continued*. *Continued from previous page.* GenBank accession numbers of all sequences of δ-conotoxins used in this study.

GenBank Accession Numbers of δ-Conotoxin Sequences Analyzed					
AF089944.1	AF089945.1	AF089946.1	AF089947.1	AF089948.1	AF089949.1
AF089950.1	AF089951.1	AF089952.1	AF089953.1	AF089954.1	AF089955.1
AF089956.1	AF089957.1	AF089958.1	AF089959.1	AF089960.1	AF089961.1
AF089962.1	AF089963.1	AF089964.1	AF089965.1	AF089966.1	AF089967.1
AF089968.1	AF089969.1	AF089970.1	AF215054.1	EF467317.1	AJ851177.1
AJ851178.1	AJ851179.1	AJ851180.1	AJ851181.1	AJ851182.1	AJ851183.1
DQ141147.1	DQ345368.1	DQ141179.1	AF215051.1	AF215052.1	AF215046.1
AF215047.1	AF215048.1	AF215049.1	AF215050.1	AF215053.1	DQ141150.1
DQ141171.1	DQ141168.1	DQ141167.1	AJ851184.1	AF215055.1	DQ141174.1
DQ345371.1	DQ345372.1	HQ897690.1	DQ141170.1	AY151285.1	DQ141172.1
EF108288.1	AF215061.1	AF193254.1	AF146357.1	AF193255.1	AF193268.1
AF193269.1	DQ141166.1	DQ141177.1	AY316159.1	AY316160.1	AF193270.1
AF193271.1	DJ379427.1	DJ379425.1	DJ379448.1	DJ379475.1	DJ379447.1
DJ379424.1	DJ379426.1	DJ379452.1	AF146355.1	AF146356.1	DJ379449.1
DJ379451.1	AF146351.1	AF146354.1	AF193259.1	AF193272.1	AF146359.1
AF193260.1	AF193261.1	AF193263.1	AF193264.1	AF193265.1	AF193266.1
AF193267.1	AF215033.1	AF215034.1	AF215035.1	AF193256.1	AJ851190.1
AF193257.1	BD280191.1	AF215039.1	AJ851174.1	AJ851176.1	AJ851185.1
AJ851186.1	AJ851187.1	AJ851188.1	AJ851189.1	DJ379400.1	DJ379437.1
DJ379444.1	DJ379453.1	DJ379454.1	DJ379455.1	DJ379456.1	DQ141175.1
DJ379483.1	DJ379484.1	DJ379488.1	DJ379493.1	DJ379489.1	DJ379492.1
DJ379571.1	DQ141160.1	DQ141178.1	DQ512803.1	JF322901.1	JF322917.1
JF322918.1	JF322919.1	JF322920.1	S78990.1	X53283.1	X53284.1
X53285.1	DJ379432.1	DJ379419.1	DJ379394.1	DJ379404.1	DJ379405.1
DJ379498.1	DJ379442.1	DJ379443.1	BD280184.1	DJ379502.1	DJ379501.1
DJ379395.1	DJ379406.1	DJ379393.1	DJ379407.1	DJ379438.1	DJ379494.1
AF215106.1	AF215038.1	EF467316.1	EF467319.1	DJ379396.1	DJ379397.1
DJ379385.1	DJ379487.1	DJ379420.1	DJ379421.1	DJ379422.1	DJ379423.1
DJ379478.1	DJ379491.1	DJ379434.1	DJ379392.1	DJ379399.1	DJ379435.1
DJ379436.1	DJ379568.1	DJ379479.1	DJ379401.1	DJ379402.1	DJ379403.1
DJ379388.1	DJ379439.1	DJ379408.1	DJ379409.1	DJ379410.1	DJ379411.1
DJ379433.1	DJ379570.1	DJ379441.1	DJ379440.1	DJ379500.1	DJ379572.1
DJ379573.1	DJ379485.1	DJ379391.1	DJ379486.1	BD280183.1	BD280187.1
AF215032.1	AF132129.1	DJ379482.1	BD280190.1	DQ141182.1	DJ379398.1
AF146358.1	AF193262.1	DJ379495.1	DJ379457.1	AF193258.1	DJ379480.1
DJ379481.1	DQ141162.1	AF215036.1	AF215037.1	DJ379499.1	DJ379569.1
DJ379477.1	DJ379389.1	DJ379390.1	DJ379387.1		

5.3 References

[1] Chun, J. B., Baker, M. R., Kim, d. o. H., Leroy, M., Toribo, P., and Bingham, J. P. (2012) Cone snail milked venom dynamics–a quantitative study of *Conus purpurascens*. *Toxicon 60*, 83–94.

[2] Myers, R. A., Cruz, L. J., Rivier, J. E., and Olivera, B. M. (1993) *Conus* peptides as chemical probes for receptors and ion channels. *Chem. Rev. 93*, 1923–1936.

[3] Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R., Shetty, R., Watkins, M., Bandyopadhyay, P., and Hillyard, D. R. (1999) Speciation of cone snails and interspecific hyperdivergence of their venom peptides. potential evolutionary significance of introns. *Ann. N. Y. Acad. Sci. 870*, 223–237.

[4] Duda, T. F. and Palumbi, S. R. (2000) Evolutionary diversification of multigene families: allelic selection of toxins in predatory cone snails. *Mol. Biol. Evol. 17*, 1286–1293.

[5] Duda, T. F. (2008) Differentiation of venoms of predatory marine gastropods: divergence of orthologous toxin genes of closely related *Conus* species with different dietary specializations. *J. Mol. Evol. 67*, 315–321.

[6] Terlau, H. and Olivera, B. M. (2004) *Conus* venoms: a rich source of novel ion channel-targeted peptides. *Physiol. Rev. 84*, 41–68.

[7] Bulaj, G. and Olivera, B. M. (2008) Folding of conotoxins: formation of the native disulfide bridges during chemical synthesis and biosynthesis of *Conus* peptides. *Antioxid. Redox Signal. 10*, 141–155.

[8] Yuan, D. D., Han, Y. H., Wang, C. G., and Chi, C. W. (2007) From the identification of gene organization of α conotoxins to the cloning of novel toxins. *Toxicon 49*, 1135–1149.

[9] Woodward, S. R., Cruz, L. J., Olivera, B. M., and Hillyard, D. R. (1990) Constant and hypervariable regions in conotoxin propeptides. *EMBO J. 9*, 1015–1020.

[10] Schoenfeld, R. A. (1999) *The Genomic Structure of Delta Conotoxins and Other O-Superfamily Conotoxins*. (Department of Biology, University of Utah, Salt Lake City, Utah), pp. 27–49.

[11] Duda, T. F. and Palumbi, S. R. (1999) Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc. Natl. Acad. Sci. U. S. A. 96*, 6820–6823.

[12] Craik, D. J., Daly, N. L., and Waine, C. (2001) The cystine knot motif in toxins and implications for drug design. *Toxicon 39*, 43–60.

[13] Conticello, S. G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z., and Fainzilber, M. (2001) Mechanisms for evolving hypervariability: the case of conopeptides. *Mol. Biol. Evol. 18*, 120–131.

[14] Olivera, B. M., Rivier, J., Scott, J. K., Hillyard, D. R., and Cruz, L. J. (1991) Conotoxins. *J. Biol. Chem. 266*, 22067–22070.

[15] Santos, A. D., McIntosh, J. M., Hillyard, D. R., Cruz, L. J., and Olivera, B. M. (2004) The A-superfamily of conotoxins: structural and functional divergence. *J. Biol. Chem. 279*, 17596–17606.

[16] Conticello, S. G., Pilpel, Y., Glusman, G., and Fainzilber, M. (2000) Position-specific codon conservation in hypervariable gene families. *Trends Genet. 16*, 57–59.

[17] Zhangsun, D., Luo, S., Wu, Y., Zhu, X., Hu, Y., and Xie, L. (2006) Novel o-superfamily conotoxins identified by cdna cloning from three vermivorous *Conus* species. *Chem. Biol. Drug Des. 68*, 256–265.

[18] Puillandre, N., Watkins, M., and Olivera, B. M. (2010) Evolution of *Conus* peptide genes: duplication and positive selection in the A-superfamily. *J. Mol. Evol. 70*, 190–202.

[19] Chang, D. and Duda, T. F. (2012) Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol. 29*, 2019–2029.

[20] Ohno, S. (1970) *Evolution by Gene Duplication.* (Springer-Verlag, New York), pp. 59–88.

[21] Olivera, B. M. (2002) *Conus* venom peptides: Reflections from the biology of clades and species. *Annu. Rev. Ecol. Syst. 33*, 25–47.

[22] Bergthorsson, U., Andersson, D. I., and Roth, J. R. (2007) Ohno's dilemma: evolution of new genes under continuous selection. *Proc. Natl. Acad. Sci. U. S. A. 104*, 17004–17009.

[23] Gowd, K. H., Blais, K. D., Elmslie, K. S., Steiner, A. M., Olivera, B. M., and Bulaj, G. (2012) Dissecting a role of evolutionary-conserved but noncritical disulfide bridges in cysteine-rich peptides using ω-conotoxin GVIA and its selenocysteine analogs. *Biopolymers 98*, 212–223.

[24] Lopez-Vera, E., Walewska, A., Skalicky, J. J., Olivera, B. M., and Bulaj, G. (2008) Role of hydroxyprolines in the *in vitro* oxidative folding and biological activity of conotoxins. *Biochemistry 47*, 1741–1751.

[25] Bulaj, G., Buczek, O., Goodsell, I., Jimenez, E. C., Kranski, J., Nielsen, J. S., Garrett, J. E., and Olivera, B. M. (2003) Efficient oxidative folding of conotoxins and the radiation of venomous cone snails. *Proc. Natl. Acad. Sci. U. S. A. 100*, 14562–14568.

[26] Price-Carter, M., Hull, M. S., and Goldenberg, D. P. (1998) Roles of individual disulfide bonds in the stability and folding of an ω-conotoxin. *Biochemistry 37*, 9851–9861.

[27] Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) Folding of ω-conotoxins. 1. Efficient disulfide-coupled folding of mature sequences *in vitro*. *Biochemistry 35*, 15537–15546.

[28] Chang, S. G., Choi, K. D., Jang, S. H., and Shin, H. C. (2003) Role of disulfide bonds in the structure and activity of human insulin. *Mol. Cells 16*, 323–330.

[29] Safavi-Hemami, H., Gorasia, D. G., Steiner, A. M., Williamson, N. A., Karas, J. A., Gajewiak, J., Olivera, B. M., Bulaj, G., and Purcell, A. W. (2012) Modulation of conotoxin structure and function is achieved through a multienzyme complex in the venom glands of cone snails. *J. Biol. Chem. 287*, 34288–34303.

[30] Braakman, I., Helenius, J., and Helenius, A. (1992) Role of ATP and disulphide bonds during protein folding in the endoplasmic reticulum. *Nature 356*, 260–262.

[31] Mirazimi, A. and Svensson, L. (2000) ATP is required for correct folding and disulfide bond formation of rotavirus VP7. *J. Virol. 74*, 8048–8052.

[32] Lévy, F., Gabathuler, R., Larsson, R., and Kvist, S. (1991) ATP is required for in vitro assembly of MHC class I antigens but not for transfer of peptides across the ER membrane. *Cell 67*, 265–274.

[33] Gething, M. J. (1999) Role and regulation of the ER chaperone BiP. *Semin. Cell Dev. Biol. 10*, 465–472.

[34] 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.

[35] Wright, F. (1990) The 'effective number of codons' used in a gene. *Gene 87*, 23–29.

[36] Maddison, W. and Maddison, D. (2010) Mesquite: a modular system for evolutionary analysis. Version 2.74. http://mesquiteproject.org.

[37] Wang, Q., Jiang, H., Han, Y. H., Yuan, D. D., and Chi, C. W. (2008) Two different groups of signal sequence in M-superfamily conotoxins. *Toxicon 51*, 813–822.

[38] Comeron, J. M. and Aguadé, M. (1998) An evaluation of measures of synonymous codon usage bias. *J. Mol. Evol. 47*, 268–274.

[39] Yang, Z. and Bielawski, J. P. (2000) Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol. 15*, 496–503.

APPENDIX A

DERIVATION OF THE NUMBER OF POSSIBLE DISULFIDE ISOMERS

To derive the number of possible disulfide isomers (assuming that the maximum number of possible disulfide bridges are present in the final product), we will take an arbitrary peptide containing n cysteine residues. The intended goal will be to show that the number of possible disulfide isomers, p, can be expressed as the product of all odd numbers from 1 to n, arriving at the final formula

$$p(n) = n!! = \frac{n!}{\lfloor \frac{n}{2} \rfloor ! \cdot 2^{\lfloor \frac{n}{2} \rfloor}}$$

which is also given in the text as Eq. 1.1 (except here, p is stated explicitly as a function of n, where this was implicit in the form given in Eq. 1.1).

The derivation will focus on chemical interpretations, rather than using the most straightforward mathematical explanation. The derivation will be divided into two parts, comprising the two sub-cases of having an even or odd total number of cysteine residues.

The general approach will be a slight twist on proof by induction, solving first the trivial cases, then showing for odd values of n, $p(n) = n \cdot p(n-1)$, and for even values of n, $p(n) = p(n-1)$, with the resultant expression that for odd n, $p(n) = n \cdot p(n-2) = n \cdot (n-2) \cdot ...1$. This is equivalent to showing that $p(n) = n!!$.

A.1 Trivial Cases

For $n = 1$, there are no possible disulfides to form, since each disulfide requires two cysteine residues. Consequently, there is 1 possible "disulfide isomer."

For $n = 2$, there is only one possible disulfide; consequently, there is 1 possible disulfide isomer.

A.2 Odd Values of n

If n is odd, then any disulfide isomer containing a maximum number of disulfides must contain one cysteine residue that is not part of a disulfide bond. Since there are n total cysteines, there are thus n possible choices for the cysteine that is not disulfide-bonded.

For the remaining $n-1$ cysteines, there are $p(n-1)$ possible disulfide isomers (by definition).

Thus, for odd values of n, the number of possible disulfide isomers, $p(n)$ is given by

$$p(n) = n \cdot p(n-1)$$

A.3 Even Values of n

Intuitively, adding one more cysteine to a peptide with an odd number of cysteines will not change the number of possible disulfides because the free cysteine now has a disulfide bonding partner.

To show this more rigorously, I will seek to show that for n even, $p(n) = p(n-1)$. I will use proof by contradiction to show this, disproving first that $p(n) < p(n-1)$, then disproving $p(n) > p(n-1)$. Since $p(n)$ is neither greater nor less than $p(n-1)$, we can deduce that $p(n) = p(n-1)$.

A.3.1 Case 1

Assume $p(n) < p(n-1)$. This would mean that there is a disulfide isomer that can be formed with $n-1$ cysteines that cannot be formed by n cysteines. However, since all our cysteines are being treated as equivalent, any disulfide isomer that can be formed by $n-1$ cysteines could also be formed by n cysteines. Contradiction.

A.3.2 Case 2

Assume $p(n) > p(n-1)$. This would mean that there is at least one disulfide isomer that can be formed by n cysteines which cannot be formed by $n-1$ cysteines. Because n is even, and we are assuming the maximum number of disulfide bonds present, there are no free thiols.

Consider one such disulfide isomer that can be formed by n cysteines, but cannot be formed by $n-1$ cysteines. Removing one cysteine from this system would break a disulfide bond, resulting in one free cysteine. By our assumption, this disulfide isomer is not in the set of disulfide isomers for $n-1$ cysteines. However, the set of disulfide isomers for $n-1$ cysteines is complete, by definition (with $p(n-1)$ members). Contradiction.

A.4 Tying Together

Having shown that for odd values of n, $p(n) = n \cdot p(n-1)$, and that for even values of n, $p(n) = p(n-1)$, we can thus deduce:

1. For odd values of n, $p(n) = n \cdot p(n-2)$.

2. For even values of n, $p(n) = p(n-1)$.

Combining these together, we get that for any n, $p(n)$ is given by the product of all odd numbers between 1 and n. Thus, $p(n) = n!!$.

To show equivalence to the other statement ($p(n) = \frac{n!}{\lfloor \frac{n}{2} \rfloor ! \cdot 2^{\lfloor \frac{n}{2} \rfloor}}$), the cases of n being even and odd will be considered separately again.

For even values of n, there is some m such that $n = 2m$, allowing me to rewrite the expression as

$$p(2m) = \frac{(2m)!}{\lfloor \frac{2m}{2} \rfloor ! \cdot 2^{\lfloor \frac{2m}{2} \rfloor}} = \frac{(2m)!}{\lfloor m \rfloor ! \cdot 2^{\lfloor m \rfloor}} = \frac{(2m)!}{m! \cdot 2^m}$$

And since $m! \cdot 2^m$ is equivalent to the product of the even numbers between 2 and $2m$, we can simply factor these out from $(2m)!$, giving $(2m)!!$, or $n!!$.

For odd values of n, the floor functions are critical. Since n is odd, there is some r such that $n = 2r + 1$. Thus,

$$p(2r+1) = \frac{(2r+1)!}{\lfloor \frac{(2r+1)}{2} \rfloor ! \cdot 2^{\lfloor \frac{(2r+1)}{2} \rfloor}} = \frac{(2r+1)!}{\lfloor \frac{2r}{2} + \frac{1}{2} \rfloor ! \cdot 2^{\lfloor \frac{2r}{2} + \frac{1}{2} \rfloor}} = (2r+1) \cdot \frac{(2r)!}{r! \cdot 2^r}$$

Thus, applying the same principles as for the case of n being even, it follows that $\frac{(2r)!}{r! \cdot 2^r}$ is equivalent to the product of all odd numbers between 1 and $2r$, so $(2r+1) \cdot \frac{(2r)!}{r! \cdot 2^r}$ is equivalent to the product of all odd numbers between 1 and $2r + 1 = n$.

APPENDIX B

ENZYME-ASSISTED OXIDATIVE FOLDING

The policy for the Journal of Biological Chemistry allows reuse of journal articles in a thesis or dissertation with no explicit permission needed.

Citation: Safavi-Hemami, H., Gorasia, D.G., Steiner, A.M., Williamson, N.A., Karas, J.A., Gajewiak, J., Olivera, B.M., Bulaj, G., Purcell, A.W. (2012) Modulation of Conotoxin Structure and Function is Achieved through a Multienzyme Complex in the Venom Glands of Cone Snails. *J. Biol. Chem.* **287**, 34288-34303. Doi: 10.1074/jbc.M112.366781.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 41, pp. 34288–34303, October 5, 2012

Modulation of Conotoxin Structure and Function Is Achieved through a Multienzyme Complex in the Venom Glands of Cone Snails*

Received for publication, March 27, 2012, and in revised form, August 12, 2012 Published, JBC Papers in Press, August 13, 2012, DOI 10.1074/jbc.M112.366781

Helena Safavi-Hemami[‡§], Dhana G. Gorasia[‡§], Andrew M. Steiner[¶], Nicholas A. Williamson[§], John A. Karas[§], Joanna Gajewiak[‖], Baldomero M. Olivera[‖], Grzegorz Bulaj[¶], and Anthony W. Purcell[1]**

*From the [‡]Department of Biochemistry and Molecular Biology and the [§]Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, 3010 Victoria, Australia, the Departments of [¶]Medicinal Chemistry and [‖]Biology, University of Utah, Salt Lake City, Utah 84112, and the **Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia*

Background: Conotoxins can be utilized to investigate enzyme-assisted folding of disulfide-rich peptides.
Results: Various ER-resident cone snail enzymes act in concert to accelerate the oxidative folding of conotoxins and modulate their conformation by reconfiguring disulfide connectivities.
Conclusion: The folding of conotoxins is a tightly regulated multienzyme-assisted process.
Significance: Modulation of the conformation of conotoxins increases their molecular and functional diversity.

The oxidative folding of large polypeptides has been investigated in detail; however, comparatively little is known about the enzyme-assisted folding of small, disulfide-containing peptide substrates. To investigate the concerted effect of multiple enzymes on the folding of small disulfide-rich peptides, we sequenced and expressed protein-disulfide isomerase (PDI), peptidyl-prolyl *cis-trans* isomerase, and immunoglobulin-binding protein (BiP) from *Conus* venom glands. *Conus* PDI was shown to catalyze the oxidation and reduction of disulfide bonds in two conotoxins, α-GI and α-ImI. Oxidative folding rates were further increased in the presence of *Conus* PPI with the maximum effect observed in the presence of both enzymes. In contrast, *Conus* BiP was only observed to assist folding in the presence of microsomes, suggesting that additional co-factors were involved. The identification of a complex between BiP, PDI, and nascent conotoxins further suggests that the folding and assembly of conotoxins is a highly regulated multienzyme-assisted process. Unexpectedly, all three enzymes contributed to the folding of the ribbon isomer of α-ImI. Here, we identify this alternative disulfide-linked species in the venom of *Conus imperialis*, providing the first evidence for the existence of a "non-native" peptide isomer in the venom of cone snails. Thus, ER-resident enzymes act in concert to accelerate the oxidative folding of conotoxins and modulate their conformation and function by reconfiguring disulfide connectivities. This study has evaluated the role of a number of ER-resident enzymes in the folding of conotoxins, providing novel insights into the enzyme-guided assembly of these small, disulfide-rich peptides.

During synthesis on the ribosome, the nascent polypeptide chain enters the crowded environment of the ER,[2] where it engages in a variety of intra- and intermolecular interactions. The newly synthesized protein associates with a number of ER-resident chaperones and foldases that accelerate rate-limiting steps along the folding pathway. The three enzymes that have repeatedly been implicated in the folding of protein substrates in the ER are protein-disulfide isomerase (PDI), peptidyl-prolyl *cis-trans* isomerase (PPI), and immunoglobulin-binding protein (BiP).

Polypeptides containing one or more disulfides in their native state are likely to become substrates for PDI, the enzyme that catalyzes the oxidation, isomerization, and reduction of disulfide bonds. PDI contains four thioredoxin-like domains, two of which have the catalytic C*XX*C motif (1). The thiol-disulfide redox state of the catalytic sites determines the direction of the reaction. Conversion of the disulfide site to a dithiol leads to oxidation of a substrate protein, whereas the alteration from a dithiol to a disulfide results in reduction of the substrate (2). Isomerization occurs either directly or through cycles of reduction and reoxidation (3). To date, 19 members of the PDI family have been identified in humans that significantly differ in their primary amino acid sequence, domain organization, catalytic activity, or mass (4, 5).

In addition to disulfide bond formation, the *cis-trans* isomerization of peptidyl-prolyl bonds can impede folding of proline-containing proteins. Peptide bonds to proline are synthesized in the *trans* conformation. Although the majority of these bonds remain in *trans*, some peptidyl-prolyl bonds need to undergo *trans-cis* isomerization. The enzyme that catalyzes this other-

* This work was supported, in whole or in part, by National Institutes of Health, NIGMS, Program Project GM48677. This work was also supported by Australian Research Council Grant DP110101331.

[1] A National Health and Medical Research Council of Australia Senior Research Fellow. To whom correspondence should be addressed. Tel.: 613-9902-9265; Fax: 613-9902-9500; E-mail: anthony.purcell@monash.edu.

[2] The abbreviations used are: ER, endoplasmic reticulum; PDI, protein-disulfide isomerase; PPI, peptidyl-prolyl *cis-trans* isomerase; BiP, immunoglobulin-binding protein; IP, immunoprecipitation; ACN, acetonitrile; RP-HPLC, reversed-phase HPLC; ESI, electrospray ionization; Fmoc, *N*-(9-fluorenyl)methoxycarbonyl; BN, blue native; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)-ethyl]glycine.

TABLE 1
Diversity of disulfide-rich peptides and enzymes implicated in their biosynthesis

Peptide	Sequence/Disulfide Scaffold	Species, Common name	Activity	Folding enzymes
Conotoxin TxIIIA	CCSWDVCDHPSCTCCG	*Conus textile* cloth of gold cone	induces excitatory behaviour in mice (50)	PDI (18)
Conotoxin GIIIA	RDCCTOOKKCKDROCKOORCCA*	*Conus geographus* geography cone	Na^+ channel inhibitor (51)	PDI (36), PPI (17)
Attracotoxin	LLACLFGNGRCSSNRDCCELTPVCKRGSCVSSGPGLVGGILGGIL	*Hadronyche versuta* funnel web spider	Ca^{2+} channel inhibitor (52)	
Maurotoxin	VSCTGSKDCYAPCRKQTGCPNAKCINKSCKCYGC*	*Scorpio maurus palmatus* chactoid scorpion	K^+ channel blocker (53)	PDI, PPI (8)
Kalata B1	CGETCVGGTCNTPGCTCSWPVCTRNGLPV	*Oldenlandia affinis*	membrane disruption, plant defence (54)	PDI (55)
Defensin Hbd-2	PVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKP	*Homo sapiens* human	antimicrobial peptide (56)	
Insulin	GIVEQCCTSICSLYQLENYCN (a-chain) FVNQHLCGSHLVEALYLVCGERGFFYTPKT (b-chain)	*Homo sapiens* human	pancreatic hormone (57)	PDI (48), BiP (49)
Guanylin	PGTCEICAYAACTGC	*Homo sapiens* human	intestinal hormone (58)	
Bovine pancreatic trypsin inhibitor (BPTI)	RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGA	*Bos taurus* cow	trypsin inhibitor (59)	PDI (60), BiP and PDI (61)#
Obtustatin	CTTGPCCRQCKLKPAGTTCWKTSLTSHYCTGKSCDCPLYPG	*Viperidae lebetina obtuse* blunt-nosed viper	integrin inhibitor (62)	

* C-terminal amidation, O = hydroxyproline, prolines and hydroxyprolines are highlighted in gray, cysteine residues are shown in boldface type.
\# Findings were based on kinetic modeling.

wise slow reaction is PPI, a ubiquitous protein that is present in almost all cellular compartments (6). In mammalian cells, a number of ER-resident PPIs have been identified, including PPI B. Interestingly, in addition to accelerating folding rates via isomerization of peptide bonds to proline, PPI was also shown to improve the efficiency of PDI-mediated folding of ribonuclease T1 by providing partially folded protein chains with the correct proline isomers (7). Similarly, oxidative folding rates of maurotoxin, a disulfide-rich peptide from scorpion venom, were highest in the presence of PDI and FKBP-12, a PPI isoform located in the cytosol (8).

Another enzyme known to cooperate with PDI in the folding of disulfide-containing proteins is BiP, a member of the heat shock protein 70 (Hsp70) family. BiP binds to unfolded and partially misfolded proteins upon their entry into the lumen of the ER and limits protein misfolding and aggregation. It is also a key enzyme in the retrograde transport of misfolded proteins from the ER to the proteasome (9). In recent years, a number of studies have addressed the potential cooperative folding of disulfide-containing proteins by PDI and BiP. Investigations into the folding of antibody chains revealed that both enzymes cooperatively act in the refolding of Fab fragments *in vitro* (10). It has been suggested that BiP binds the unfolded polypeptide chains and keeps them in a conformation in which the cysteine residues are accessible for PDI (10). Immunoprecipitation experiments further revealed that PDIA6, a member of the PDI family, forms a non-covalent complex with BiP and shows specificity toward BiP client proteins in human fibroblast cells (11). Subsequently, the *in vitro* interaction between another PDI family member, PDIA3, and BiP was shown to improve folding rates of ribonuclease B and α-lactalbumin (12).

Together, these findings indicate that members of the PDI family can recruit or are themselves recruited by other ER-resident enzymes and chaperones, such as BiP and/or PPI B, to cooperatively act in the oxidative folding of at least some client substrates. Recently, an ER-resident complex comprising PDI, PPI B, and BiP was identified in human hepatoma and mouse lymphoma cells (13), reflecting tightly regulated compartmentalization of protein folding in the ER.

Although our understanding of enzyme-guided folding of proteins is constantly improving, comparatively little is known about the folding and assembly of small, disulfide-rich peptides. Disulfide-rich peptides are widely distributed in the plant and animal kingdoms, where they serve diverse functions. Examples include antimicrobial peptides, such as the defensins; various protease inhibitors; hormones, including insulin; and a wide range of neurotoxins found in animal venoms (Table 1). Predatory marine snails of the genus *Conus* synthesize a great diversity of disulfide-rich peptides that often carry additional post-translational modifications, most commonly proline hydroxylations, γ-carboxylations, and C-terminal amidations (14). The diversity of peptides generated by cone snails is astonishing. Each of the 500–700 species of cone snail is estimated to synthesize hundreds of different peptides with distinct disulfide connectivities (15). With their vast structural diversity, conotoxins can therefore be regarded as model disulfide-rich peptides for understanding general mechanisms of peptide folding, including the formation of correct disulfide bonds.

The concerted effect of PDI, PPI, and BiP on the folding of these and other disulfide-rich peptides has not been systematically explored. Although PDI and PPI have been identified in the venom gland of *Conus* (16) and are known to play a role in the *in vitro* folding of conotoxins (17, 18), their combined effect has not been investigated. We recently sequenced *Conus* BiP from the venom gland of *Conus novaehollandiae* and demonstrated it to be highly expressed in the venom glandular cells (16). Whether BiP plays a functional role in the folding of conotoxins was not addressed in our earlier study. To investigate

multienzyme-assisted folding of conotoxins, we cloned and expressed PDI, PPI B, and BiP from *C. novaehollandiae* and tested their potential in assisting in the oxidative folding of α-conotoxins GI and ImI. Here, we show that all three enzymes play a role in the oxidative folding of these two peptides and, in the case of PDI, also in the unfolding of disulfide bonds under reducing conditions. Importantly, a direct interaction between PDI and conotoxin α-GI was confirmed by affinity purification. We further reveal the presence of a complex between PDI and BiP, suggesting a concerted action of these enzymes, similar to what has been observed for the folding of larger proteins. Our findings indicate that the proper assembly of conotoxins relies upon the action of a number of ER-resident enzymes, including PDI, PPI B, and BiP. Importantly, these enzymes appear to modulate the disulfide isomers of conotoxins, a process that not only generates additional conformational diversity but also has the potential to enhance the functional diversity of these molecules.

EXPERIMENTAL PROCEDURES

Depending on the experimental requirements, a number of different cone snail species were utilized in this study. *Conus* PPI B and BiP were previously sequenced from the venom gland of *C. novaehollandiae*. Consequently, *C. novaehollandiae* was selected for molecular sequencing of PDI. Conotoxin α-ImI was originally isolated from the venom of *Conus imperialis*. Investigations into the presence of different isoforms of α-ImI were carried out on the venom of this species. Investigations of protein complexes were performed on the venom gland of *Conus victoriae*. This species has a particularly large venom gland, allowing the isolation of sufficient material. Conotoxin α-GI was originally isolated from the venom of *Conus geographus*. Co-immunoprecipitation (co-IP) experiments using the biotin-labeled propeptide of α-GI were performed on venom glands from this species.

Specimen Collection and Tissue Preparation—Specimens of *C. novaehollandiae* and *C. victoriae* were collected from Broome, Western Australia. Live animals were shipped to Melbourne, Australia, and processed within 24 h of collection. Venom ducts were extracted, snap-frozen and stored at −80 °C. Specimens of *C. imperialis* and *C. geographus* were collected from Marinduque Island, the Philippines. Venom ducts were prepared as described above. Prior to freezing, venom was extruded from isolated venom ducts of *C. imperialis*, air-dried, and stored at −80 °C.

Venom Extraction and Analysis—Twenty mg of dried *C. imperialis* venom was reconstituted in 1 ml of 40% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA) and homogenized using a Dounce tissue grinder. Insoluble material was pelleted by centrifugation at 12,000 × *g* for 10 min. The supernatant was collected, dried by vacuum centrifugation, and stored at −80 °C until processing. Four mg of crude venom was separated by reversed-phase high performance liquid chromatography (RP-HPLC) on a semipreparative C18 column (5-μm particle size, 10 × 250 mm, Vydac-Grace) using a linear gradient from 5 to 100% buffer B (90% ACN, 0.1% TFA) over 80 min. Peaks were manually collected and further analyzed by electrospray ionization tandem mass spectrometry (ESI-MS/MS) and matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). For ESI-MS/MS, samples were analyzed using a high resolution linear ion trap mass spectrometer (LTQ XL Orbitrap, Thermo Scientific). For MALDI-MS, peptides were spotted onto matrix-coated MALDI plates and analyzed using a MALDI-time of flight (TOF) mass spectrometer in positive reflector mode (Voyager, AB SCIEX).

Preparation of Rat Liver Microsomal Proteins—Microsomal proteins were prepared as described previously (19). Briefly, 20 mg of frozen rat liver tissue was homogenized in 20 ml of ice-cold homogenization buffer (10 mM Tris, 150 mM NaCl, pH 7.5, 1× protease inhibitor mixture (Roche Applied Science)) using a Dounce tissue grinder. Cellular debris was removed by centrifugation at 12,000 × *g* for 15 min. Membranous material containing microsomes was pelleted by ultracentrifugation at 100,000 × *g* for 1 h. Following ultracentrifugation, pellets were washed twice with homogenization buffer. Closed vesicles were converted to open membranes to release microsomal proteins by incubating with ice-cold 0.1 M sodium carbonate buffer (pH 11) for 20 min. Membranous material was removed by ultracentrifugation as described above. The pH of the supernatant was adjusted to 7.5 with 2 N HCl solution. Microsomal protein preparations were examined by SDS-PAGE, and protein concentrations were determined using Bradford reagent (Sigma-Aldrich). Protein-bound and free sulfhydryl groups were determined spectrophotometrically using Ellman's reagent (5,5′-dithiobis(2-nitrobenzoic acid)) as described previously (20).

Molecular Sequencing—We previously obtained full-length sequences of PPI B and BiP from the venom gland of *C. novaehollandiae* (accession numbers GU046312 and HM627497, respectively). In order to sequence *Conus* PDI, frozen venom glands from *C. novaehollandiae* were ground under liquid nitrogen, and total RNA was extracted using TRIzol® reagent (Invitrogen) and DNase I treated with Turbo DNase (Ambion Inc). RNA extraction and DNase treatment were performed according to the manufacturers' instructions. cDNA was reverse transcribed from 1 μg of DNase-treated RNA using the Superscript VILO cDNA synthesis kit (Invitrogen) with a 1:1 mixture of oligo(dT)s and random hexamers. Primary reverse transcription PCR (RT-PCR) was performed in volumes of 50 μl containing 4 μl of cDNA (200 ng), 0.5 μl of Velocity *Taq* DNA polymerase (Bioline), 1× PCR buffer (Bioline), 200 μM each deoxynucleoside triphosphate (dNTPs; Invitrogen), and 0.2 μM PDI sense (CGA CCA TAT GGA TGA TAT CAA ACA GGA GGA A) and antisense oligonucleotide primers (CGC TCG AGC AGT TCA TCT CTT GGC AGA TC). PCR amplicons were analyzed by gel electrophoresis, cloned, and sequenced as described previously (17). The novel *Conus* PDI sequence was deposited in GenBank™ (National Center for Biotechnology Information, United States National Library of Medicine, Bethesda, MD). Nucleotide sequences were translated into the predicted amino acid residues, and the putative signal peptide was predicted using SignalP software (21).

Cloning, Expression, and Purification of Recombinant Enzymes—*Conus* PPI B was expressed and purified as described previously (17). *Conus* PDI and BiP lacking the N-terminal signal sequences were cloned into the pET22b+ and pET30c+

expression vector, respectively (Novagen). Briefly, transcripts were PCR-amplified from venom gland cDNA prepared as described above and ligated into pET22b+ using the NdeI (5′) and XhoI (3′) restriction sites and into pET30c+ with the endonucleases NdeI (5′) and HindIII (3′) (New England Biolabs). Sequences were verified by nucleotide sequencing. The constructs, containing a C-terminal His_6 tag, were transformed into *Escherichia coli* (Rosetta strain, Novagen) by heat treatment. For expression of recombinant proteins, LB broth containing antibiotics (100 μg/ml ampicillin for pET22b+ and 50 μg/ml kanamycin for pET30c+) was inoculated with overnight cultures and incubated at 37 °C with shaking until the A_{600} spectrophotometric reading was 0.6. Expression was induced by adding 0.1 mM isopropyl-β-D-thiogalactopyranoside followed by incubation for 3 h at 25 °C with shaking. Bacteria were harvested and resuspended in native lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole, pH 8.0) containing 1× protease inhibitor mixture (EDTA-free; Roche Applied Science). Bacterial cells were lysed by probe tip sonication. Cellular debris and insoluble protein were pelleted by centrifugation at 15,000 × *g* for 20 min, and the supernatants were used for subsequent protein purifications.

Recombinant proteins were purified on a 1-ml immobilized metal affinity column under native conditions (Bio-Scale Mini Profinity IMAC cartridge, Bio-Rad). Purification was performed using the Bio-Rad Profinia protein purification system. Protein lysates were loaded onto the column at 0.5 ml/min, and nonspecifically bound proteins were removed with native buffer (300 mM KCl, 50 mM KH_2PO_4, pH 8.0) containing 5 mM imidazole at 2 ml/min for 6 min, followed by a second wash with native buffer containing 15 mM imidazole for 5 min. His-tagged proteins were eluted with 250 mM imidazole in native buffer at 1 ml/min for 4 min. Further purification and buffer exchange into 10 mM Tris-HCl, 150 mM NaCl, pH 8, was accomplished by size exclusion chromatography (Superdex 200, HiLoad 16/60, GE Healthcare) at 1 ml/min. The purified recombinant proteins were analyzed by SDS-PAGE and sequence-verified by in-gel digestion and analysis of proteotypic tryptic peptides as described previously (17). Protein concentrations were determined spectrophotometrically using the proteins' molar absorption coefficients (22).

Functional Enzyme Analyses—The insulin reduction assay was performed using a modification of the insulin turbidity assay described previously (23). Briefly, recombinant PDI (0.5 and 2.5 μM final concentration) was incubated with 2 mM dithiothreitol (DTT) for 15 min prior to adding the assay mixture (100 mM potassium phosphate, 2 mM EDTA, 2 mM DTT, and 0.85 mg/ml insulin, pH 6.5). The absorbance at 650 nm was measured for 1 h at 30 °C.

The peptidyl-prolyl *cis-trans* isomerase activity of recombinant PPI B was determined using the coupled chymotrypsin assay (24) with modifications as described previously (17). Briefly, 0.3 or 3 μM PPI B was incubated in 100 mM Tris-HCl, pH 8, at 25 °C for 2 min. Samples were transferred to 10 °C for 5 min before the addition of 30 μl of 600 mM chymotrypsin (Sigma-Aldrich). After 5 min at 10 °C, reactions were initiated by adding the substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide at a final concentration of 78 μM (Bachem AG; stock made in trifluoroethanol in the presence of 0.45 M LiCl (Sigma-Aldrich)). The *cis-trans* isomerization was measured by following the absorbance at 390 nM for 3 min.

The ATPase activity of recombinant BiP was verified using a modification of the NADH-coupled photometric assay described previously (25). The assay relies on the hydrolysis of ATP to ADP by the ATPase activity of BiP. Upon ATP hydrolysis, phosphoenolpyruvate is converted to pyruvate by pyruvate kinase, followed by further conversion into lactate by lactate dehydrogenase. The latter reaction requires oxidation of NADH, resulting in a decrease in absorbance at 340 nm. Briefly, reaction buffer containing 30 mM Tris (pH 7.6), 30 mM KCl, 5 mM $MgCl_2$, 0.15 mM NADH, 4 mM phosphoenolpyruvate, 25 μg/ml lactate dehydrogenase, and 25 μg/ml pyruvate kinase was incubated at 25 °C for 10 min. For control reactions, ADP was added to a final concentration of 10 mM, and NADH depletion was monitored at 340 nm. To determine the ATPase activity of BiP, ATP was added to a final concentration of 10 mM. After base-line stabilization, BiP was added at a final concentration of 2.5 and 5 μM, and the absorbance was monitored for 35 min at 25 °C.

Peptide Synthesis—Based on their well characterized folding properties, the two α-conotoxins GI and ImI were selected for oxidative folding studies. Conotoxin α-ImI was obtained from the peptide synthesis core facility at the University of Utah. Conotoxin α-GI was synthesized with an Apex 396 automated peptide synthesizer (AAPPTec) using a standard solid-phase Fmoc protocol. The peptide was constructed on preloaded Fmoc-Rink Amide MBHA resin (substitution, 0.4 mmol/g; Peptides International Inc.) on a 50-μmol scale. All amino acids were purchased form AAPPTec, and side-chain protection for the following amino acids was as follows: Glu, *O-tert*-butyl; Arg, 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl; Tyr and Ser, *tert*-butyl; Asn, Cys, and His, trityl. Coupling activation was achieved with 1 eq of 0.4 M benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate and 2 eq of 2 M *N,N*-diisopropylethyl amine in *N*-methyl-2-pyrrolidone as the solvent. A 10-fold excess of each amino acid was used, and each coupling reaction was conducted for 60 min. The Fmoc deprotection reaction was carried out for 20 min with 20% piperidine in dimethylfluoride. Peptides were cleaved from the resin by treatment with reagent K (TFA/thioanisole/ethanedithiol/water/phenol (82.5:5:2.5:5:5 by volume)) for 3 h and 3.5 h for α-GI and α-ImI, respectively. The peptides were subsequently filtered, precipitated, and washed with cold methyl *tert*-butyl ether. Linear peptides were purified by RP-HPLC on a semipreparative C18 column (5-μm particle size, 10 × 250 mm, Vydac-Grace) using a linear gradient from 5 to 40% buffer B (90% ACN, 0.1% TFA) over 20 min. Buffer A was 0.1% TFA, water. Absorbance was monitored at 220 nm. Concentrations were determined spectrophotometrically using the peptides' molar absorption coefficient at 280 nm. For co-elution experiments, correctly folded α-GI and α-ImI were obtained from the peptide synthesis facility at the Salk Institute.

The biotin-labeled propeptide of α-conotoxin GI (pro-GI; biotin-FPSERASDGRDDTAKDEGSDMEKLVEKKECCNPACGRHYSC-NH_2) was synthesized as described previously (26) with the exception that the peptide was C-terminally tagged with biotin followed by 6-aminohexanoic acid (GL Biochem)

and not oxidized prior to purification. The biotin-labeled control peptide (biotin-YARFQSQTTLKQKT) was purchased from the peptide synthesis facility at Monash University.

Oxidative Folding Studies—Oxidative folding reactions were carried out at room temperature in the presence of 0.1 M Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 10 mM ATP, 0.1 mM GSSG, 0.1 mM GSH, 3 μM PPI B and/or 1 μM PDI and/or 2 μM BiP. Because the exact concentrations of PDI, PPI B, and BiP in the ER are not known, the enzyme/substrate ratios were chosen to be in agreement with previous studies (7, 8). Folding was initiated by adding the linear peptide at a final concentration of 20 μM. Aliquots were taken at various time points, and reactions were quenched by acidification with formic acid to a final concentration of 8%. Folding reactions were analyzed by RP-HPLC on a C18 column (5-μm particle size, 4.6 × 250 mm, Vydac-Grace) using the conditions described above. For samples spiked with rat liver microsomal proteins, separation was performed on a diphenyl column (5-μm particle size, 4.6 × 250 mm, Vydac-Grace) using the same conditions. Native peptides were distinguished from linear forms based on characteristic elution profiles (27–29), co-elution experiments, and mass spectrometric (MS) analyses of manually collected reversed-phase fractions (MALDI-TOF mass spectrometer, positive reflector mode, Voyager, AB SCIEX). Statistical analyses were performed using two-tailed Student's *t* tests with unequal variance.

Conotoxin Reductase Assay—For reductase assays, 1 μM PDI was preincubated in reductase buffer (0.1 M Tris-HCl, 1 mM $MgCl_2$, 10 mM ATP, 5 mM GSH, pH 7.5) for 1 h at room temperature. Peptide reduction was initiated by adding the folded peptide at a final concentration of 20 μM. Quenching and analysis were performed as described above.

Blue Native PAGE—Blue Native (BN)-PAGE techniques were adapted from those described previously (30). Briefly, frozen venom glands from *C. victoriae* were homogenized in ice-cold buffer (20 mM Tris-HCl, 50 mM NaCl, 10% glycerol, 1 mM PMSF, pH 7.5, containing 1× protease inhibitor mixture) using a Dounce tissue grinder. Cellular debris was removed by centrifugation at 6,000 × *g* for 8 min at 4 °C. The supernatant was collected and ultracentrifuged at 100,000 × *g* for 1 h at 4 °C to separate soluble proteins from membranous material. Supernatants and pellets were collected, and proteins were solubilized by adding between 1 and 3% of digitonin or CHAPS followed by incubation for 20 min on ice. An aliquot of the soluble and membranous protein sample was mixed with 5× BN-PAGE loading buffer (5% Coomassie Brilliant Blue G-250, 500 mM (*E*)-aminocaproic acid in 100 mM BisTris, pH 7) and loaded onto NuPAGE Native 4–16% gels (Invitrogen) along with an in-house molecular weight marker (kindly provided by Srgjan Civciristov). Samples were separated in the presence of blue cathode buffer (50 mM Tricine, 15 mM BisTris, 0.02% Coomassie Brilliant Blue G-250, pH 7) and anode buffer (50 mM BisTris, pH 7) for 50 min at 150 V on ice. The blue cathode buffer was then replaced with cathode buffer without Coomassie dye, and electrophoresis was completed at 200 V. Gels were either stained with Coomassie Blue followed by in-gel tryptic digest or second dimension SDS-PAGE or further analyzed by immunoblotting.

For second dimension SDS-PAGE, gel bands were excised and soaked in 1% SDS and 1% mercaptoethanol for 2 h followed by two washes in water. Gel bands were placed into wells on 12% NuPAGE BisTris gels (Invitrogen). Gels were run for 20 min at 100 V followed by 40 min at 200 V. Following SDS-PAGE, gels were either stained with Coomassie Blue or analyzed by in-gel tryptic digest or immunoblotting.

Immunoblotting—Proteins were transferred from gels onto polyvinylidene fluoride (PVDF) membranes and blocked with 5% skim milk in phosphate-buffered saline, 0.05% Tween 20 (PBS-T) followed by immunoblotting. Primary antibodies were rabbit polyclonal anti-BiP antibody raised against a synthetic peptide sequence derived from human BiP (anti-GRP78 BiP antibody, ab21685, Abcam) and rabbit polyclonal anti-PDI antibody raised against recombinant rat liver PDI (courtesy of Prof. M. Hubbard); The secondary antibody was sheep anti-rabbit, HRP-conjugated (Abcam). Incubations were carried out at room temperature for 1 h. Antibody stocks were diluted 1:2,000 in 2.5% skim milk, PBS-T. Detection was performed with ECL reagent (PerkinElmer Life Sciences). For immunoblotting following BN-PAGE, gels were briefly soaked in transfer buffer containing 0.1% SDS prior to transfer onto PVDF membranes. After transfer, membranes were washed in methanol for 3 min to remove residual Coomassie dye. For reprobing, membranes were washed 5 min in PBS-T, followed by incubation in 62 mM Tris-HCl, 100 mM mercaptoethanol, 2% SDS, pH 6.8, for 30 min at 70 °C. Membranes were washed three times for 10 min each in PBS-T, and immunoblotting was performed as described above.

Co-immunoprecipitation and Affinity Pull-down Studies—Co-IP assays were performed using either a polyclonal anti-PDI or anti-BiP antibody or a biotin-labeled propeptide of the α-conotoxin GI. Briefly, venom glands were homogenized in 1% digitonin in Tris-buffered saline (TBS; 100 mM Tris, 150 mM NaCl, pH 7.5) containing 1× protease inhibitor mixture by Dounce homogenization followed by incubation on ice for 30 min. Cellular debris was removed by centrifugation at 12,000 × *g* for 20 min, and supernatants were precleared with either 70 μl of Protein A-Sepharose pre-equilibrated in lysis buffer, Mab-Select, GE Healthcare) or 70 μl of streptavidin-agarose (Sigma-Aldrich) for 1 h at 4 °C with mixing. Precleared supernatants were incubated with 40 μl of anti-PDI- or anti-BiP-conjugated Protein A-Sepharose slurry or pro-GI-conjugated biotin/streptavidin-agarose for 1 h at 4 °C with mixing. Non-immunized rabbit serum (Sigma-Aldrich) and a random biotin-labeled control peptide (YARF) were used to establish nonspecific binding. Following incubation, beads were washed six times with 0.1% digitonin in TBS containing 1× protease inhibitor mixture. Bound proteins were eluted by incubating with Laemmli sample buffer with and without 100 mM DTT for 10 min at 100 °C. Proteins were analyzed by SDS-PAGE followed by in-gel tryptic digest or immunoblotting as described above.

In-gel Digestion and Protein Identification—In-gel digestion was performed as described previously (31) with minor modifications. Briefly, gel spots were excised, washed in 50% ACN, triethylammonium bicarbonate, and reduced with 20 mM DTT, followed by alkylation in 100 mM iodoacetamide. In-gel digestion was performed using sequencing grade trypsin (Sigma-Aldrich) at a final concentration of 10 μg/ml in 25 mM triethylammonium bicarbonate. Peptides extracted after overnight digestions were separated on a C18 reversed-phase column

(ProteCol nanocolumn, pore size 300 Å, particle size 3 μm, dimensions 75 μm × 100 mm, SGE Analytical Sciences) and analyzed using a Hybrid Quadrupole-TOF LC-MS/MS mass spectrometer (QSTAR Elite, AB SCIEX). Acquired data were analyzed using Analyst QS software (version 2.0, AB SCIEX). MS/MS data were searched against an in-house database that contained all molluscan proteins submitted to UniProt ($n = 47{,}252$) using Protein Pilot software (version 3.0, AB SCIEX) with the following selections: iodoacetamide, trypsin gel-based identification, biological modifications, thorough ID. The false discovery rate cut-off was set to 5%. Peptides with confidence scores of ≥99 were selected for protein identifications.

RESULTS

cDNA Sequencing of PDI—Conus PDI was successfully sequenced from the venom gland of *C. novaehollandiae*. Mature *C. novaehollandiae* PDI is encoded by 1449 nucleotides translating to a putative protein of 54,682 Da (GenBank™ accession number JQ745296). The predicted protein is most homologous to PDI sequenced from *Conus virgo* (92% amino acid sequence identity, 97% homology, GenBank™ accession number ADZ76590) and shares 61% sequence identity (77% homology) with the human homologue (GenBank™ accession number NP_000909). *C. novaehollandiae* PDI exhibits the classical PDI domain organization with two catalytic a and a′ domains and the non-catalytic b and b′ domains. The active site motif is -CGHC- in both catalytic domains.

Characterization of Recombinant Proteins—All three enzymes were expressed without their N-terminal signal sequence and purified to ~90% purity by metal affinity and size exclusion chromatography (Fig. 1). The identity of recombinant proteins was confirmed by in-gel tryptic digestion of SDS-polyacrylamide gel bands (data not shown). Functional integrity was validated by enzyme-specific activity assays (Fig. 1). To assess enzymatic activity of *Conus* PDI, the insulin turbidity assay was performed. In the absence of PDI, insulin was slowly reduced by DTT, as observed by an increase in absorbance at 650 nm (Fig. 1*A*). The addition of PDI led to a substantial increase of insulin reduction, confirming that the active site cysteines efficiently unfolded the disulfide bonds of insulin (Fig. 1*A*).

Enzymatic activity of PPI B was tested using the coupled chymotrypsin assay. In the absence of enzyme, the peptidyl-prolyl bond of the oligopeptide *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide slowly underwent *cis* to *trans* isomerization, resulting in proteolytic cleavage by chymotrypsin. Accumulation of liberated *p*-nitroanilide led to an increase in absorbance at 390 nm (Fig. 1*B*). This reaction was accelerated in the presence of *Conus* PPI B, confirming the *cis-trans* isomerization activity of the recombinant enzyme (Fig. 1*B*).

To determine the ATPase activity of BiP, the NADH depletion assay was employed (25). In the absence of BiP, the addition of ADP led to a rapid depletion of NADH, verifying the functionality of the assay (Fig. 1*C*, *black arrows*). In the presence of BiP and ATP, NADH was depleted via BiP-catalyzed ATP hydrolysis over the time course of the experiment, whereas depletion was not observed in the no enzyme control (Fig. 1*C*).

Enzyme-assisted Oxidative Folding Studies—To investigate the role of *Conus* PDI, PPI B, and BiP in the folding of disulfide-rich peptides, the two α-conotoxins GI and ImI were tested as folding substrates (Fig. 2). Peptides were selected based on their well characterized *in vitro* oxidative folding properties (27–29). All α-conotoxins contain 4 cysteine residues that can potentially adopt three disulfide conformations (Fig. 2). The Cys1-3 and Cys2-4 disulfide pairing has been confirmed for a number of α-conotoxins and is generally referred to as the native, globular fold. Although the three-dimensional conformation and bioactivity of the other two isomers, the ribbon and the bead form (Cys1-4 plus Cys2-3 and Cys1-2 plus Cys3-4, respectively), have been studied in detail, these isoforms have not been found in the venom (28, 32).

Conotoxins α-GI and α-ImI are 13 and 12 amino acids in length, respectively. Both peptides are C-terminally amidated and contain one proline. In α-GI, the proline is in position 5, and the Asn^4–Pro^5 bond most likely adopts a trans configuration (32). This peptide belongs to the 3/5 subfamily of α-conotoxins with 3 residues between the first and second cysteine and 5 amino acids between the third and fourth. While sharing the characteristic disulfide pattern with GI, conotoxin ImI belongs to the 4/3 subgroup of the α-conotoxins with 4 and 3 residues between the first and second set of cysteines, respectively (Fig. 2). ImI contains one proline in position 6, and the Asp^5–Pro^6 bond has a *trans* orientation (33).

All reversed-phase elution profiles of oxidative folding reactions concurred with those previously reported for these two peptides (27–29), and the molecular masses of the collected peaks agreed with the calculated masses as determined by MALDI-TOF MS analysis (data not shown). Identities of linear and native peptides were further confirmed by co-elution experiments using commercially purchased peptides as a reference (data not shown). Oxidative folding reactions were quenched by acidification. Previous studies have reported insufficient quenching by acidification of enzyme-assisted folding reactions in some cases (34). To ensure efficient termination of the reaction, replicate samples were analyzed in random order within 12 h following folding assays. No differences were observed between replicate samples run between 1 and 12 h following acidification, indicating efficient termination of the folding assay.

Effect of PDI on the Folding of α-GI and α-ImI—All oxidative folding reactions were carried out in 0.1 M Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 10 mM ATP, 0.1 mM GSH, and 0.1 mM GSSG. Under these conditions, steady state accumulation of native α-GI occurred after ~64 min (Fig. 3*A*) with a final yield of 63.2% ± 2.0 (mean ± S.D.). The addition of PDI did not affect final folding yields but significantly increased folding rates (Fig. 3*A*). This effect was most apparent between 8 and 32 min of folding (Fig. 3*A*). The effect of PDI on the folding of α-GI was abolished when PDI was inactivated by denaturation for 20 min at 95 °C prior to the folding assay (data not shown).

Conotoxin α-ImI exhibited faster folding kinetics with steady state conditions reached after ~16 min (Fig. 3, *B* and *C*). Unlike α-GI, this peptide folds into two distinct isomers *in vitro*, the globular and the ribbon form, with almost equal yields at steady state conditions (Fig. 3, *B* and *C*). Interestingly, PDI had little effect on the folding rates of globular α-ImI (Fig. 3*B*) but significantly increased the folding rates of the ribbon isomer (Fig. 3*C*).

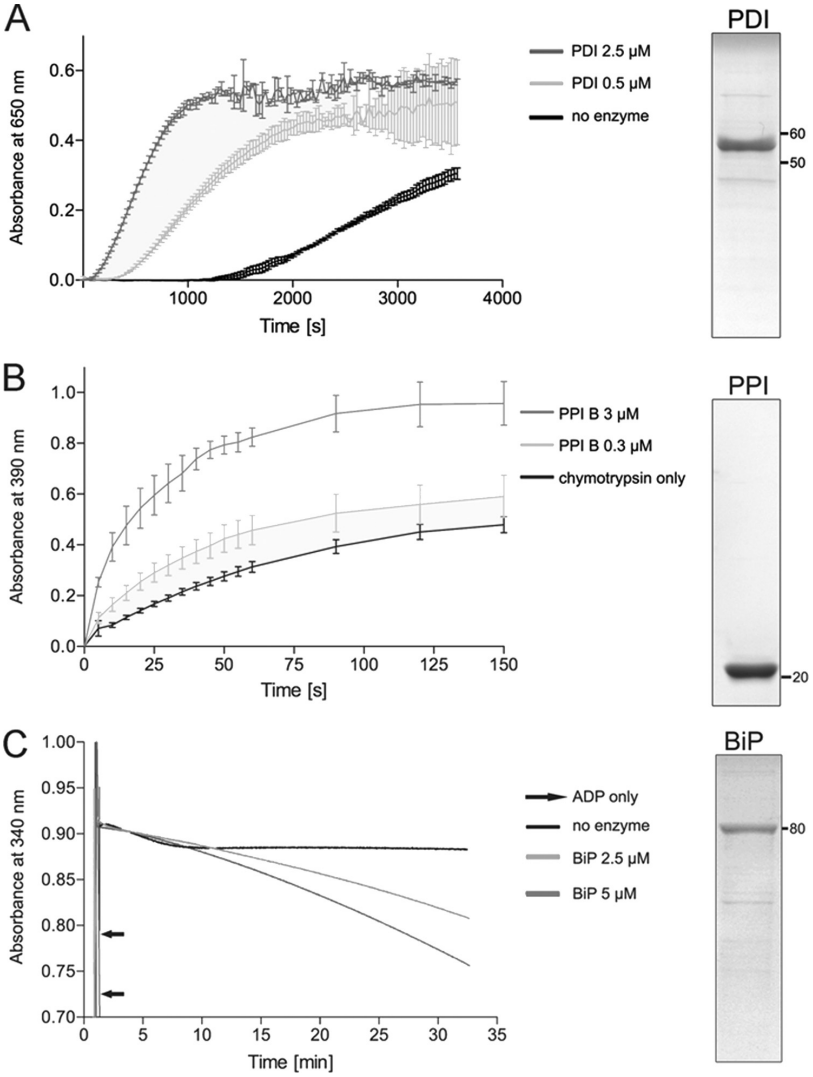


FIGURE 1. **Purity and integrity of recombinant *Conus* enzymes as determined by SDS-PAGE and enzyme-specific activity assays.** The purity of recombinant enzymes was determined by SDS-PAGE (*right panels*). A total of 3 μg of recombinant protein was separated on 12% Tris-glycine gels under reducing conditions. Proteins were visualized by Coomassie staining. *A*, PDI activity was measured using the insulin reductase assay. The addition of PDI (0.5 and 2.5 μM) caused rapid reduction of insulin, as observed by an increase in absorbance at 650 nm. This effect was concentration-dependent. *B*, the *cis-trans* isomerase activity of PPI B was measured using the coupled chymotrypsin assay. The addition of PPI B (0.3 or 3 μM) resulted in an increase in the rate of cis-trans isomerization of the oligopeptide *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide, a substrate of chymotrypsin (mean $\pm$ S.D.). Enzymatic cleavage leads to liberation of *p*-nitroanilide and an increase in absorbance at 390 nm. This reaction was concentration-dependent. *C*, the ATPase activity of BiP was determined using the NADH-coupled photometric assay. Hydrolysis of ATP was followed by monitoring depletion of NADH at 340 nm. The addition of ADP led to a sudden decrease in absorbance, confirming the functionality of the assay. In the presence of BiP (2.5 and 5 μM), absorbance slowly decreased over the time course of the assay, whereas no effect was observed for the no enzyme control. This effect was concentration-dependent.

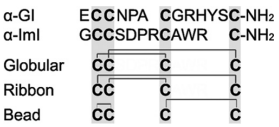


FIGURE 2. **Amino acid sequences and possible disulfide connectivities of α-conotoxins used in this study.** Cysteine residues are highlighted in *gray*. Both peptides are C-terminally amidated. Potential disulfide linkages of the three isomers, the globular, ribbon, and bead form, are depicted.

In addition to its oxidation and isomerization activity, PDI is known to catalyze the reduction of disulfide bonds (4). To assess its reductase activity on α-GI and α-ImI, reductase assays were carried out in the presence of 5 mM GSH with and without PDI. The addition of PDI significantly accelerated unfolding rates of α-GI with a 6-fold decrease in the half-time of disappearance of the folded form when PDI was present (Fig. 3*D*). A similar effect, although less distinct, was observed for the rib-

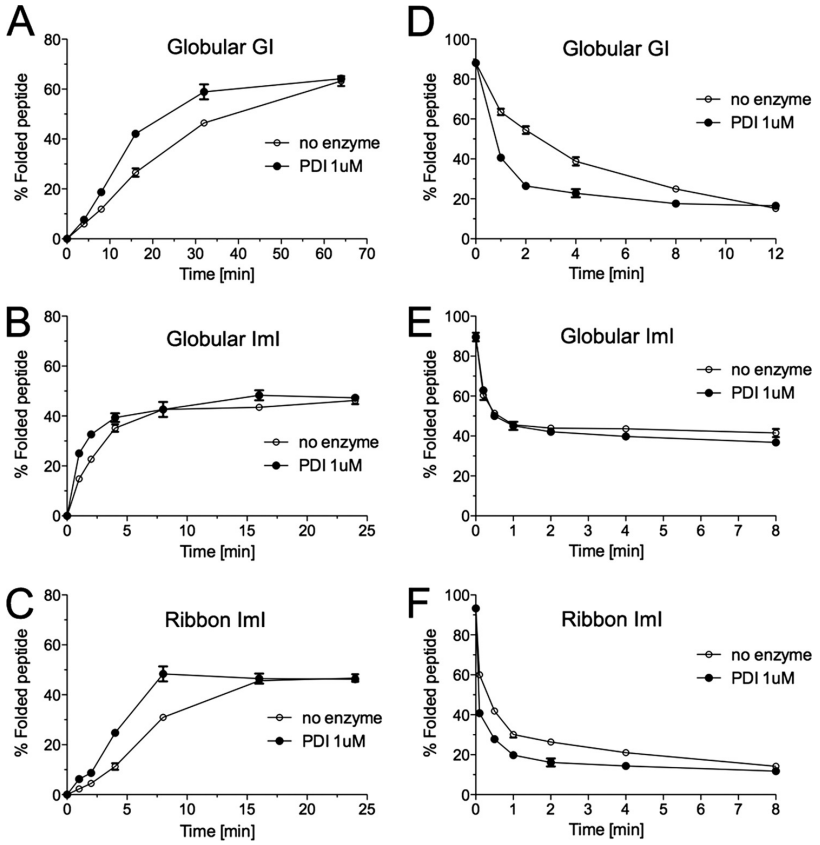


FIGURE 3. **Analysis of PDI-assisted folding and reduction of α-GI and α-ImI.** *A–C*, oxidative folding was carried out in the presence of 0.1 mM GSH, 0.1 mM GSSG, and 20 μM linear peptide with and without *Conus* 1 μM PDI. *D–F*, peptide reduction studies were performed in the presence of 5 mM GSH, 20 μM folded peptide with and without *Conus* PDI. Folding and reduction reactions were terminated by acid quenching and analyzed by reversed-phase chromatography (C18 column, Vydac-Grace). Relative abundances of fully folded globular α-GI (*A* and *D*), globular α-ImI (*B* and *E*), and ribbon α-ImI (*C* and *F*) were determined from three independent experiments (mean ± S.D. (*error bars*)).

bon isomer of α-ImI (Fig. 3*F*). The addition of PDI led to a 2-fold decrease in the half-time of the disappearance of this peptide. Unfolding kinetics of the globular form remained unchanged (Fig. 3*E*).

Effects of PDI and PPI B—To elucidate the effect of PDI in combination with PPI B, folding assays were carried out as described above in the presence of *Conus* PDI, PPI B, or both PDI and PPI B. The fastest folding rates for α-GI were observed in the presence of both PDI and PPI B, although both enzymes individually afforded a rate increase over the uncatalyzed reaction (Fig. 4). At 16 min, accumulation of the native peptide increased from 28.0 ± 0.33 to 32.0 ± 0.57% ($p < 0.01$) and 42.6 ± 1.35% ($p < 0.01$) in the presence of PPI B or PDI, respectively (Fig. 4 and Table 2). This further increased to 48.2 ± 0.25% in the presence of both enzymes.

PDI and PPI B significantly increased the folding rates of the ribbon form of α-ImI (Fig. 5, *B* and *C*, and Table 2) but had little effect on the appearance of the globular form (Fig. 5*A*). Again, the fastest folding rates were achieved when both enzymes were present. At 4 min, accumulation of ribbon α-ImI increased from 11.5 ± 0.28% to 15.6 ± 0.29% ($p < 0.001$) and 22.8 ± 1.18% ($p < 0.001$) in the presence of PPI B and PDI, respectively (Fig. 5, *B* and *C*, and Table 2). This further increased to 31.1 ± 0.18% ($p < 0.001$) when both enzymes were present. The addition of PDI and PPI B significantly affected the folding rates of α-GI and ribbon α-ImI but did not alter their final folding yields (Figs. 4*A* and 5*B* and Table 2). Enzyme-mediated changes in folding rates were abolished when PDI and PPI B were heat-inactivated for 20 min at 95 °C prior to folding assays (data not shown).

Effects of PDI and BiP—To investigate the effect of PDI in combination with BiP, oxidative folding assays were carried out in the presence of *Conus* PDI, BiP, or both PDI and BiP. *Conus* BiP alone did not exhibit any apparent effect on the oxidative folding rates of α-GI and α-ImI under various enzyme concentrations and redox conditions tested. Although PDI maintained the ability to accelerate folding rates of α-GI and the ribbon isomer of α-ImI, no additional effect was observed when BiP was added (data not shown). To investigate whether recombinant BiP required additional co-factors for proper action, folding reactions were supplemented with rat microsomal extracts. Changes in the accumulation of folded peptides were deter-

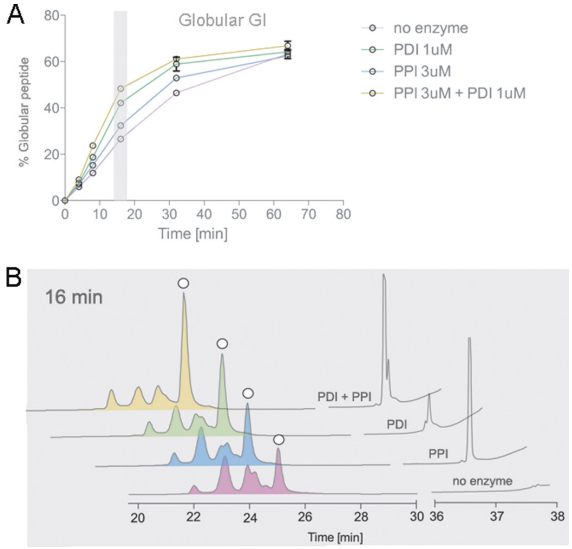


FIGURE 4. **Analysis of PDI- and PPI-assisted folding of α-GI.** Oxidative folding studies were performed in the presence of 0.1 mM GSH, 0.1 mM GSSG, and 20 μM linear peptide with and without *Conus* PDI and/or PPI. Folding reactions were acid-quenched at 0, 4, 8, 16, 32, and 64 min and analyzed by reversed-phase chromatography. *A*, relative abundance of the fully folded, globular peptide was determined by reversed-phase chromatography as shown in *B* and plotted against time points of folding. Plotted values are averages from three independent experiments (mean ± S.D. (*error bars*)). *B*, reversed-phase chromatograms of folding reactions quenched after 16 min. *White circles* denote the folded, globular peptide as determined by its characteristic elution profile (27, 29) and co-elution experiments.

mined at early time points of folding. Reactions were quenched at 4 and 8 min for α-ImI and α-GI, respectively. The addition of microsomal proteins alone led to an increase in the accumulation of folded α-GI (13.1 ± 0.19% *versus* 5.8 ± 0.16% without microsomes) and more rapid disappearance of the linear form after 8 min of folding (Fig. 6*A*), indicating the presence of foldases and/or chaperones in these preparations. This effect was only slightly reduced when microsomal proteins were heat-treated prior to folding (10.4 ± 0.07%), suggesting incomplete denaturing or renaturing of foldases and chaperones or the presence of non-protein co-factors. In the presence of microsomal proteins, the addition of PDI or BiP led to a significant increase in the accumulation of folded α-GI (28.7 ± 0.62 and 18.9 ± 0.3%, respectively ($p < 0.001$); Fig. 6*A*). Folding was most efficient when both enzymes were present in the same reaction (32.2 ± 0.27% ($p < 0.05$); Fig. 6*A*). For α-ImI, the addition of microsomal extracts alone had little effect on the accumulation of fully folded peptide but affected the disappearance of the linear form (Fig. 6*B*). However, the addition of PDI or BiP led to a significant increase in the accumulation of the folded ribbon isomer (20.1 ± 1.28 and 24.6 ± 0.75%, respectively, *versus* 10.9 ± 0.48% without enzyme ($p < 0.01$); Fig. 6*B*). This effect was further enhanced when both enzymes were present in the same reaction (30.1 ± 0.98% ($p < 0.01$); Fig. 6*B*). Interestingly, accumulation of the globular form was affected to a lesser extent. Although PDI and BiP individually increased the accumulation of the globular form (28.4 ± 0.89 and 28.8 ± 1.37%, respectively, *versus* 17.8 ± 0.60% without enzyme ($p < 0.01$); Fig. 6*B*), no additive effect was observed in the presence of both enzymes (28.5 ± 0.71%; Fig. 6*B*).

Gel electrophoresis of microsomal extracts revealed a complex mixture of proteins characteristic of these preparations (Fig. 6*C*). To determine the concentration of protein-bound and free sulfhydryl groups that could potentially affect PDI-mediated thiol-disulfide exchange reactions, microsomal extracts were analyzed using Ellman's reagent. The concentration of protein-bound thiols was 67 ± 1.4 μM, and that of free sulfhydryl groups was below the detection limit of the assay (<15 μM).

In summary, all three recombinant enzymes affected the oxidative folding kinetics of α-GI and the ribbon isomer of α-ImI, a disulfide species believed to be absent from the venom. Folding rates of globular, so-called native α-ImI remained unchanged. These observations led us to subsequently investigate the presence of the ribbon isomer of α-ImI in the venom of *C. imperialis*.

Detection of α-ImI Isomers in the Venom of C. imperialis—To determine the presence of different isomers of α-ImI in the venom of *C. imperialis*, reversed-phase elution profiles of synthetic ribbon and globular α-ImI were compared with those of the crude venom. Using a linear gradient from 5 to 100% buffer B (90% ACN, 0.1% TFA) over 80 min, the synthetic ribbon and globular isomers eluted at 28.1 min (23.8% ACN) and 29.5 min (25.3% ACN), respectively (Fig. 7, *A* and *B*, *yellow line*). Analysis of the venom of *C. imperialis* revealed the presence of two corresponding peaks (Fig. 7, *A* and *B*, *green line*). Both peaks strongly increased in size upon the addition of synthetic peptides (Fig. 7, *A* and *B*, *blue line*), suggesting the presence of both isomers in the crude venom. The two corresponding venom fractions were collected and subjected to MALDI-MS analysis. Peptides with identical masses to folded α-ImI were identified in both peaks (m/z = 1351.5; Fig. 7, *C* and *D*), whereas no mass matches were observed in interjacent fractions (data not shown). To further verify the presence of the two isomers in the venom, venom fractions were subjected to ESI-MS/MS analysis. Comparison of the venom peptides with their co-eluting synthetic analogues revealed the presence of two characteristic fragment ions only present in the co-eluting synthetic and venom-derived globular isomer (m/z = 668.3 and 1099.3; data not shown). The absence of these peaks in both the synthetic and venom-derived co-eluting ribbon fractions strongly suggests a different disulfide connectivity of these two isobaric peptides. Collectively, reversed-phase chromatography and mass spectrometric analysis led to the identification of the ribbon isomer in the venom of *C. imperialis* and confirmed the presence of its globular form. To our knowledge, this represents the first evidence for the existence of a "non-native" peptide isoform in the venom of *Conus*.

Investigation of Protein Associations in the Venom Glandular Cells—Our findings on enzyme-assisted folding and unfolding of conotoxins strongly indicate that their proper assembly is a multienzyme-assisted process. To determine whether PDI, PPI B, and BiP form a complex in the ER of the secretory cells of the venom gland, BN-PAGE and co-IP experiments were performed on the venom gland of *C. victoriae*. Soluble complexes were separated from membrane-associated ones by ultracen-

TABLE 2

Folding yields of globular α-GI and globular and ribbon α-ImI obtained at early and late time points of folding in the presence and absence of PDI and/or PPI B

Values represent mean ± S.D. calculated from three independent experiments.

Peptide	Time point	Percentage of folded peptide			
		No enzyme	PDI	PPI	PDI + PPI
	min	%	%	%	%
Globular α-GI	16	28.0 ± 0.33	42.6 ± 1.35	32.0 ± 0.57	48.2 ± 0.25
	64 (end point)	63.2 ± 1.98	64.2 ± 0.18	62.3 ± 0.26	66.8 ± 1.94
Globular α-ImI	4	35.4 ± 0.72	35.3 ± 0.93	38.8 ± 1.43	35.0 ± 0.47
	16 (end point)	43.5 ± 0.17	43.9 ± 0.19	40.3 ± 0.6	43.2 ± 1.94
Ribbon α-ImI	4	11.5 ± 0.28	22.8 ± 1.18	15.6 ± 0.29	31.1 ± 0.18
	16 (end point)	45.6 ± 2.0	45.12 ± 0.19	44.2 ± 0.6	45.3 ± 1.94

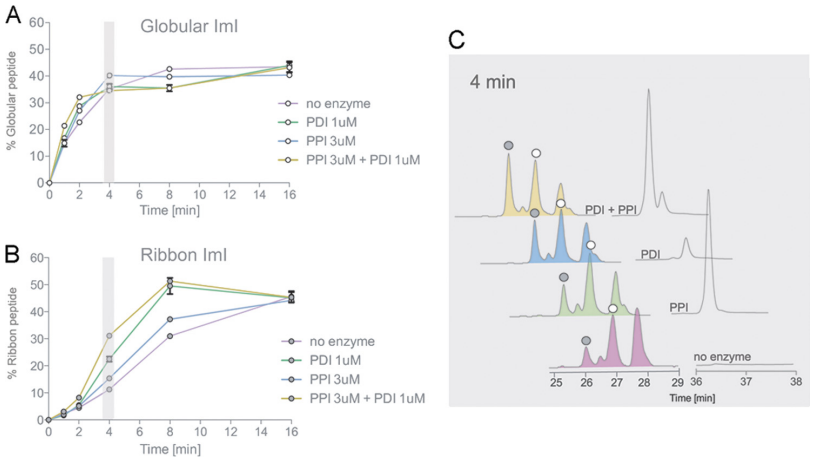


FIGURE 5. **Analysis of PDI- and PPI-assisted folding of α-ImI.** Oxidative folding studies were performed in the presence of 0.1 mM GSH, 0.1 mM GSSG, and 20 μM linear peptide with and without *Conus* PDI and/or PPI B. Folding reactions were acid-quenched at 0, 2, 4, 8, 16, and 32 min and analyzed by reversed-phase chromatography. Relative abundances of the fully folded globular (*A*) and fully folded ribbon peptide (*B*) were determined by reversed-phase chromatography as shown in *C* and plotted against time points of folding. Plotted values are averages from three independent experiments (mean ± S.D. (*error bars*)). *C*, reversed-phase chromatograms of folding reactions quenched after 4 min. *White circles* denote the folded globular peptide, and *gray circles* show the ribbon isomer, as determined by their characteristic elution profiles (28) and co-elution experiments.

trifugation. Initial experiments showed that the best solubilization of protein complexes was achieved with 1% digitonin for soluble and 2% CHAPS for membrane preparations. Under these conditions, a complex with immunopositive staining for BiP and PDI was identified in both protein preparations (Fig. 8*A*, *black arrows*). To further verify the identity of BiP and PDI, bands were excised from BN-polyacrylamide gels (Fig. 8*A*, *box Bi* and *Bii*) and subjected to either in-gel tryptic digest combined with mass spectrometric identification or second dimension denaturing SDS-PAGE followed by immunoblotting. Both techniques confirmed the presence of PDI and BiP in the soluble and membrane protein preparations (Fig. 8, *B* and *E*). The soluble complex appeared smaller (200–300 kDa) and may have separated from the larger, membrane-bound complex (300–400 kDa) during protein preparation (Fig. 8*A*). Interestingly, besides the 55-kDa PDI protein, an additional PDI isoform of ~120 kDa was observed in the soluble fraction (Fig. 8*Bi*). This 120-kDa protein is unlikely to represent a dimer of PDI because BN-polyacrylamide gel slices were incubated in denaturing buffer under strongly reducing conditions prior to second dimension electrophoresis (30 min in 2% SDS and 100 mM mercaptoethanol). A faint band of ~120 kDa was also visible in the membrane fraction (Fig. 8*Bii*), with the main PDI isoform migrating at ~90 kDa. In a previous study (16), immunoblotting of two-dimensional gels revealed the presence of larger molecular weight isoforms of PDI in the venom glands of *C. victoriae* and *C. novaehollandiae*. Whether the 90 kDa band represents a different PDI isoform remains to be determined because proteins excised from BN-PAGE can exhibit deviating migration patterns due to the presence of encasing BN gel matrices.[3] Second dimension SDS-PAGE illustrated the presence of additional proteins of unknown identity, particularly for the membrane complex (Fig. 8, *Bi* and *Bii*). Bands corresponding to the size of PPI B (~20 kDa) and Hsp40 (~40 kDa) were observed. Unfortunately, due to low protein concentrations and/or a lack of sequence homology to published protein sequences, mass spectrometry could not resolve the identity of these species. Furthermore, immunoblotting using an anti-PPI B antibody was unsuccessful because commercial polyclonal antibodies raised against human PPI B did not cross-react with the *Conus* enzyme. Future studies using antibodies raised against *Conus* PPI B will determine whether PPI B forms part of the BiP-PDI complex.

[3] H. Safavi-Hemami, unpublished observations.

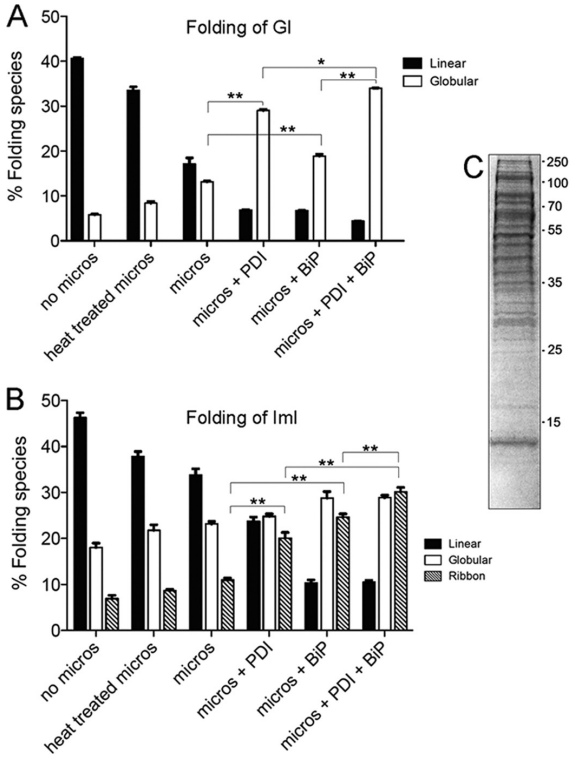


FIGURE 6. *A* and *B*, analysis of PDI and BiP-assisted folding of α-GI (*A*) and α-ImI (*B*) in the presence of rat liver microsomal proteins. Folding was carried out in the presence of 40 μg of microsomal proteins, 0.1 mM GSH, 0.1 mM GSSG, 1 mM $MgCl_2$, 10 mM ATP, and 20 μM linear peptide with and without *Conus* PDI and/or BiP. Folding reactions were acid-quenched at 8 and 4 min for α-GI and α-ImI, respectively, and analyzed by reversed-phase chromatography. Plotted values are averages from three independent experiments (mean ± S.D. (*error bars*)). Statistical analysis was performed using two-tailed Student's *t* tests with unequal variance (*, $p < 0.05$; **, $p < 0.001$). *C*, microsomal proteins were analyzed by SDS-PAGE. A total of 20 μg was separated on a 12% Tris-glycine gel under reducing conditions.

Because proteins that co-migrate on BN-polyacrylamide gels do not necessarily interact, co-IP experiments were performed to confirm the direct association between PDI and BiP. Initial IP approaches using the polyclonal anti-PDI antibody did not result in efficient immunoprecipitation of PDI, rendering this antibody unsuitable for co-IP experiments. However, the interaction between PDI and BiP was successfully confirmed using the anti-BiP-specific antibody (Fig. 8*C*). Co-immunoprecipitated proteins were separated by SDS-PAGE alongside whole lysates under reducing and non-reducing conditions (Fig. 8*C*). A band that was subsequently identified as *Conus* BiP was observed in co-IP samples, illustrating that the BiP antibody was suitable for IP experiments (Fig. 8*Ci*, *bands 1* and *3*). A very faint band corresponding to the molecular weight of PDI could also be observed in BiP co-IP lanes but not in non-immunized rabbit serum controls (Fig. 8*Ci*, *bands 2* and *4*). In-gel tryptic digestion unambiguously identified this protein as *Conus* PDI (Fig. 8*E*). A number of additional gel bands were observed in BiP co-IP lanes. However, these proteins could not be identified by in-gel tryptic digest. The identity of the PDI protein band was further confirmed by immunopositive staining of co-IP samples using the anti-PDI antibody (Fig. 8*Cii*, *black arrow*). Immunostaining of this band was also observed in whole lysates but not in the serum control (Fig. 8*Cii*). Both anti-PDI and anti-BiP were raised in rabbits, resulting in high background staining. To rule out the possibility that PDI staining resulted from the secondary anti-rabbit antibody, membranes were stripped and reblotted using secondary antibody only. PDI-associated staining could not be observed in BiP co-IP lanes. Together, in-gel tryptic digest and immunoblotting of BiP co-IP samples confirmed the association between PDI and BiP in the venom glandular cells of *C. victoriae*.

In Vitro Interaction between Conus PDI and Conotoxin α-GI—Several studies have proposed PDI-assisted folding of conotoxins (18, 27, 35–37). However, a direct interaction between PDI and conotoxins has never been demonstrated. To determine whether PDI directly interacts with a conotoxin *in vitro*, the biotin-labeled propeptide of α-GI was incubated with whole venom gland lysates of *C. geographus* and subsequently affinity-purified using streptavidin beads. Affinity-purified proteins were analyzed by SDS-PAGE under reducing conditions (Fig. 8*Di*). A number of proteins specifically associated with pro-GI but not with the random YARF control peptide (Fig. 8*Di*). Bands were excised and subjected to in-gel tryptic digest and mass spectrometric analysis. Due to a lack of *Conus* protein sequences represented in public databases, only a limited number of proteins could be identified. One of these proteins was *Conus* PDI (Fig. 8, *Di* (*white arrow*) and *E*). PDI was associated with pro-GI but not YARF (Fig. 8*Di*). Immunoblotting using anti-PDI antibody further confirmed that PDI was successfully affinity-purified with biotin-pro-GI but not biotin-YARF (Fig. 8*Cii*, *white arrow*).

DISCUSSION

The *Conus* venom gland can be regarded as one of the most sophisticated organs of disulfide-rich peptide biosynthesis. At any given time, hundreds of biologically diverse and structurally complex peptides are expressed, translated, folded, and secreted in the epithelial cells of this highly specialized organ. It is now well understood that the folding and assembly of larger polypeptides is guided by the action of a number of different foldases and chaperones. Enzymes known to cooperate in the folding of protein substrates include various members of the PDI family; heat shock proteins, such as Hsp90 and BiP; PPI-like enzymes; and lectins, such as calreticulin and calnexin (38). Although our understanding of these events is steadily improving, comparably little is known about the contribution of these enzymes to the assembly of small peptides, especially those comprising multiple disulfide bonds. Given the diversity of peptides generated by cone snails, we believe that conotoxins can serve as model peptides for understanding general mechanisms of peptide folding. In this study, we investigated the contribution of a number of enzymes to the folding of the two α-conotoxins GI and ImI. Our findings suggest that, similar to the assembly of larger polypeptide substrates, the proper folding of conotoxins is a multienzyme-assisted process.

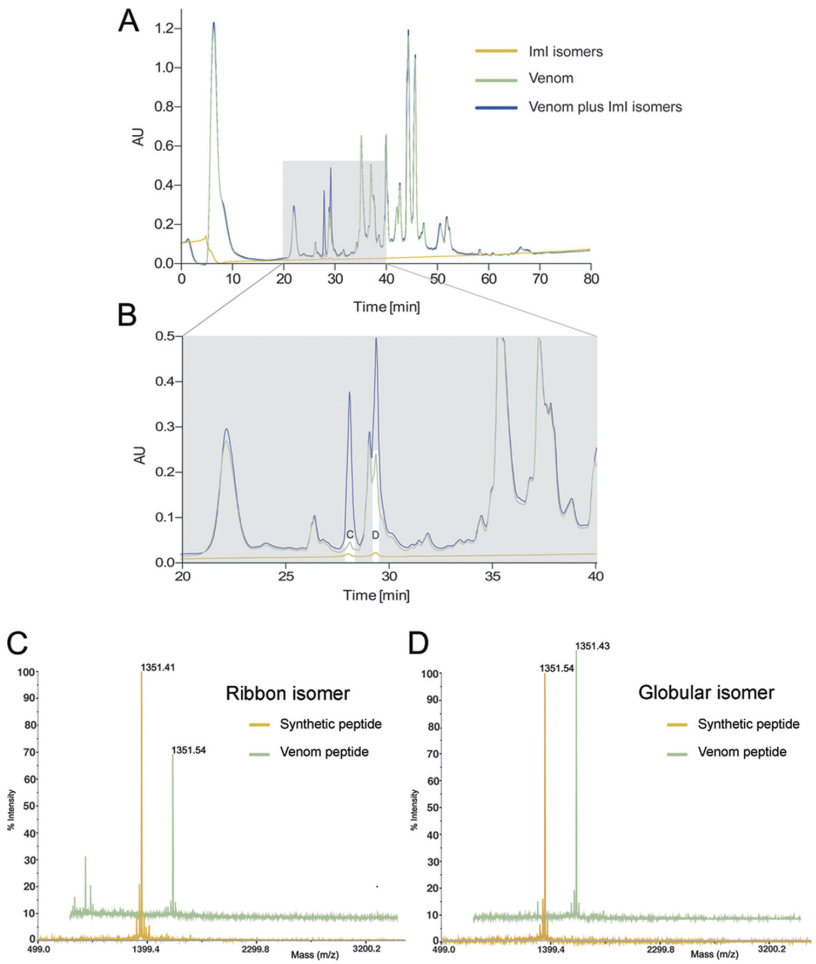
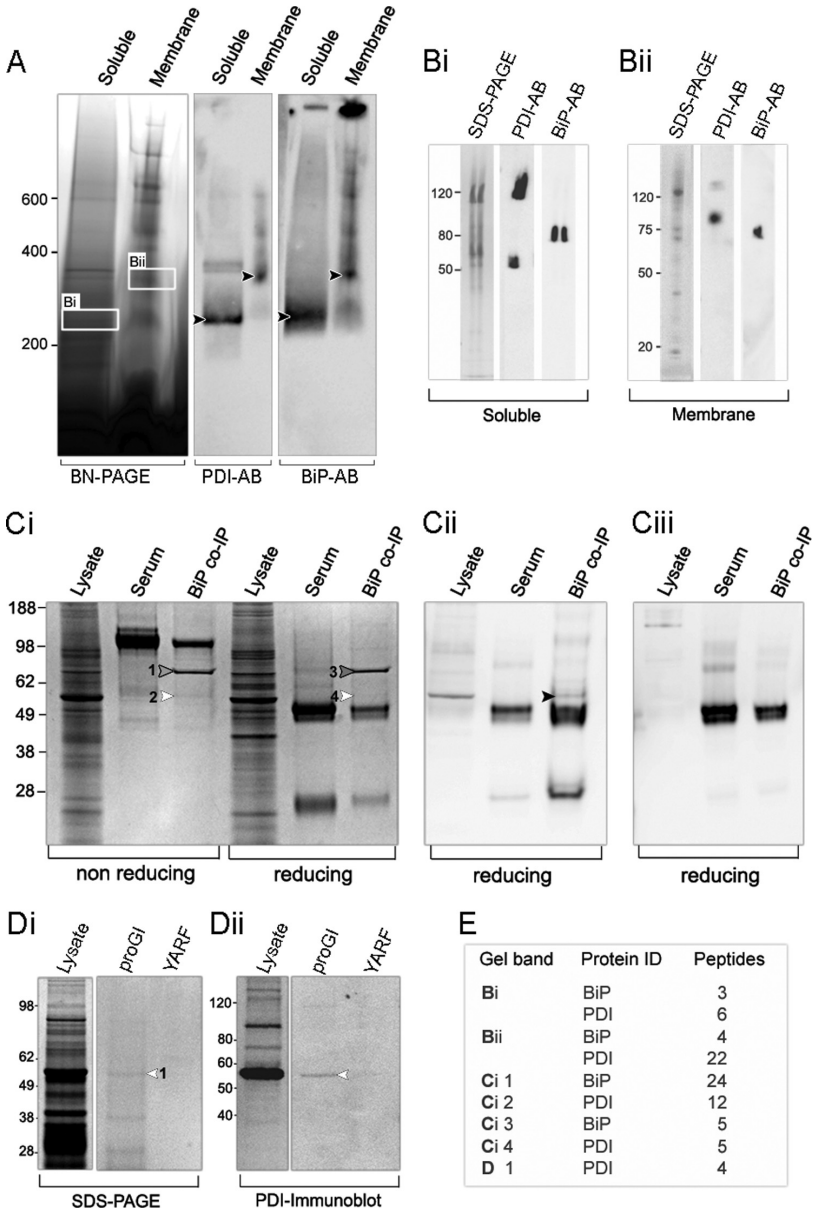


FIGURE 7. **Analysis of α-Iml isomers in the venom of *C. imperialis*.** *A*, overlay of reversed-phase chromatograms of crude *C. imperialis* venom (*green*; 4 mg) with the ribbon and globular isomer of Iml (*orange*; 2 nmol each) and crude venom spiked with 6 nmol of each Iml isomer (*blue*). Samples were separated on a semipreparative C18 column (5-μm particle size, 10 × 250 mm, Vydac-Grace) using a linear gradient from 5 to 100% buffer B (90% ACN, 0.1% TFA) over 80 min. *B*, *enlargement* of reversed-phase chromatogram depicted in *A*. Venom fractions co-eluting with ribbon (*C*) and globular Iml (*D*) were collected and subjected to MALDI-MS. Synthetic isomers were analyzed for comparison. MALDI-MS analysis shows the presence of peptides with the same mass/charge ratio as Iml in venom fractions that co-eluted with the ribbon (*C*) and globular (*D*) peptide (*m/z* 1351.5).

Three enzymes were selected based on their well established roles in the folding of protein substrates and their high expression levels in the venom gland of *Conus*. We showed that all three enzymes, PDI, PPI B, and BiP, efficiently assist in the oxidative folding of α-GI and α-ImI. PDI represents one of the most abundant soluble proteins in the venom gland of several cone snail species (16, 37) and has previously been shown to improve *in vitro* folding rates of conotoxin substrates (18). We confirm this finding and further demonstrate the ability of PDI to efficiently unfold disulfide bonds in conotoxins under reducing conditions. The role of this enzyme in oxidative protein folding is well established. In contrast, *in vivo* PDI-mediated protein unfolding has only been demonstrated for cholera toxin. PDI catalyzes the reduction of the A subunit of this protein, which enables its retrograde transport into the cytosol (39). Whether the PDI-assisted unfolding of conotoxins observed *in vitro* here also plays a role *in vivo* remains to be determined.

PDI has been repeatedly implicated in the oxidative folding of conotoxins (18, 36), but a direct interaction has never been demonstrated. Here, we provide evidence for the direct interaction between PDI and the α-conotoxin GI by affinity purification, ultimately establishing the role of this enzyme in the folding of conotoxins. Furthermore, by investigating different folding isomers of α-ImI, we reveal a substrate selectivity of PDI for the ribbon isomer, which was previously believed to be absent from the venom (28, 32). Similar observations were made for the folding of the ribbon isomer in the presence of *Conus* PPI B. Both enzymes significantly accelerated the folding of ribbon α-ImI, a peptide isomer with otherwise slow cysteine oxidation rates. These observations suggest that both disulfide bond formation and isomerization of the peptide bond to pro-

line are rate-limiting steps in the folding of this isomer. The three-dimensional structure of the ribbon isomer is unknown. However, PPI-assisted folding of ribbon α-ImI suggests that the Asp^5–Pro^6 bond has a *cis* orientation or transiently adopts this conformation during folding. For the globular form, the Asp^5–Pro^6 bond adopts a *trans* orientation (33) that most likely does not undergo transient *trans-cis* isomerization during folding. Previous studies have predominantly focused on the pharmacological properties of globular, so-called native conotoxin isomers. The few published studies on receptor selectivity of different α-conotoxin isomers all demonstrated a shift in target affinity upon disulfide rearrangement (40–42). For example, the ribbon isomer of α-Vc1.1 is a weaker antagonist of the bovine nicotinic receptor than the globular peptide (42). In contrast, the ribbon form of AuIB has significantly higher potency at rat nicotinic receptors than the globular peptide (40). These observations together with our findings of ribbon α-ImI in the venom of *C. imperialis* suggest that cone snails synthesize a variety of different disulfide isomers in order to extend their repertoire of biologically active neuropeptides. We



Gel band	Protein ID	Peptides
Bi	BiP	3
	PDI	6
Bii	BiP	4
	PDI	22
Ci 1	BiP	24
Ci 2	PDI	12
Ci 3	BiP	5
Ci 4	PDI	5
D 1	PDI	4

show that whereas the globular isomer of α-ImI rapidly folds under oxidizing conditions, the *in vitro* folding of the ribbon form is guided by ER-resident foldases.

Folding studies on α-GI and α-ImI carried out in the presence of both enzymes showed an additive rather than a synergistic effect, suggesting that these two enzymes act independently in the folding of the two peptides. Consequently, PPI B is unlikely to present conotoxins containing the correct proline isomer to PDI as previously suggested for ribonuclease T1 (7). Investigations on the concerted folding by PDI and BiP resulted in similar observations. An additive rather than a synergistic effect was observed when PDI and BiP were simultaneously present in folding reactions supplemented with microsomal extracts. BiP does not directly contribute to disulfide-exchange reactions (38). Instead, it may have promoted faster folding by preventing aggregation and/or increasing the solubility of the toxins and/or of PDI (9, 38). The BiP-mediated increase in folding rates was abolished when microsomal proteins were omitted, strongly suggesting that this enzyme requires one or more ER-resident co-factors. Free thiols present in microsomal preparations are unlikely to have contributed to faster folding rates because their concentration was very low. Macromolecular crowding is another factor that is known to contribute to alterations in rates and equilibria of protein interactions. However, the total protein concentrations used in this study were too low to have caused crowding (43). It is now well understood that members of the Hsp70 family interact with a variety of Hsp40 co-chaperones. The major function of Hsp40 is to stimulate the otherwise weak ATPase activity of BiP by increasing the rate of hydrolysis of bound ATP (44). Microsomal extracts are likely to have contained ER-resident Hsp40 co-chaperones that could have contributed to BiP-assisted folding of conotoxins. Because endogenous ATP was not depleted from microsomal preparations, no conclusions can be drawn on the ATP dependence of this reaction.

Although PDI- and BiP-assisted folding did not appear to be synergistic for α-GI and α-ImI, this could be the case for conotoxins that contain very hydrophobic regions and more than two disulfide bonds. We acknowledge that more studies utilizing a wider range of conotoxin substrates are needed to elucidate this possibility in the future.

However, our findings on the additive effect of ER-resident enzymes on the folding of α-conotoxins led us to investigate whether these events may occur in close spatial and/or temporal proximity. This would probably be corroborated by the existence of multienzyme complexes. Immunoprecipitation studies alongside with BN-PAGE experiments clearly identified a complex between PDI and BiP in the venom glandular cells of *Conus*. Because folding intermediates are highly susceptible to degradation, this multienzyme complex may decrease the time needed to generate conotoxins in their final three-dimensional structure. BN-PAGE followed by second dimension SDS-PAGE revealed that the PDI-BiP complex contained additional proteins, including those migrating at the same molecular weight as PPI B and Hsp40. Attempts to identify these proteins by mass spectrometry and immunoblotting failed due to low concentrations and/or lack of homology with published protein sequences. Future studies utilizing *Conus*-specific antibodies will determine whether proteins such as PPI B and Hsp40 form part of the BiP-PDI complex. A recent investigation into the existence of multiprotein complexes in mammalian cells identified a complex comprising, among other proteins, PDI, PPI B, BiP, and the Hsp40 protein ERdj3 (13). The authors suggested that a large fraction of ER-resident chaperones forms a network that can bind to unfolded protein substrates rather than existing as free proteins that assemble onto nascent polypeptide chains (13). Based on the contribution of PDI, PPI B, and BiP to conotoxin folding and the existence of a PDI-BiP complex in venom glandular cells, it may be hypothesized that this proposed ER network has been utilized by cone snails for toxin assembly for more than 55 million years and has remained conserved throughout evolution. Notably, high expression levels of PDI, BiP, and PPI were also found in the venom glands of snakes and wasps, animals that secrete disulfide-rich peptides as part of their venoms (45–47). The existence of multienzyme folding complexes in the venom glands of these and other venomous animals has not been investigated. However, studies on the oxidative folding of the scorpion toxin maurotoxin revealed most efficient folding in the presence of PDI and the cytosolic peptidyl-prolyl isomerase FKBP-12 (8). PDI has also been shown to accelerate the folding rates of insulin, an archetypal peptide hormone that contains multiple disulfide bonds (48). Interestingly, recent studies also implicated BiP in the folding and secretion of insulin (49). Based on these observations and our findings on the multienzyme-assisted folding of conotoxins, we hypothesize that small, disulfide-rich peptides abide by similar rules and experience folding environments comparable with those of larger polypeptide chains.

FIGURE 8. **Investigations of protein-protein interactions in the venom gland of *Conus*.** *A*, blue native polyacrylamide gel electrophoresis (BN-PAGE) of soluble and membrane-associated protein complexes isolated from the venom gland of *C. victoriae*. Gels were immunoblotted using anti-PDI- and anti-BiP-specific antibodies (*AB*). A band with immunopositive staining for PDI and BiP was identified in both preparations (*black arrows*). Corresponding bands were excised from BN-polyacrylamide gels and further analyzed by immunoblotting and second dimension SDS-PAGE under denaturing conditions (*Bi* and *Bii*) followed by in-gel tryptic digestion (see *E* for results). Both techniques confirmed the presence of PDI and BiP as part of the protein complex. *Ci*, co-IP of proteins in the venom gland of *C. victoriae* using an anti-BiP antibody. Immunoprecipitated proteins were separated by SDS-PAGE and analyzed by in-gel tryptic digestion (*bands 1*, *2*, *3*, and *4*; see *E* for results) and immunoblotting using an anti-PDI-specific antibody. In-gel tryptic digestion identified BiP and PDI in BiP co-IP samples but not in rabbit serum controls. Immunoblotting further revealed the presence of PDI in BiP Co-IP samples (*black arrow*), confirming the association between BiP and PDI. *Ciii*, to rule out the possibility that PDI-positive staining was derived from the secondary antibody, membranes were stripped and reblotted using secondary antibody alone. No staining of the PDI band was observed for reblotted membranes. *Di*, affinity purification of PDI using the biotin-labeled propeptide of α-GI. Biotin-labeled pro-GI was incubated with venom gland protein lysates of *C. geographus*. Proteins interacting with pro-GI were pulled down using streptavidin beads and analyzed by SDS-PAGE, in-gel tryptic digestion (see *E* for results), and immunoblotting using an anti-PDI-specific antibody. In-gel tryptic digestion demonstrated that PDI was successfully affinity-purified using biotin-labeled pro-GI but not using a random biotin-labeled control peptide (YARF). *Dii*, immunoblotting using an anti-PDI antibody further confirmed the identity of the PDI band (*white arrow*). *E*, results of in-gel tryptic digestions. Tryptic peptides were analyzed by ESI-MS/MS on the QStar Elite Hybrid Q-TOF mass spectrometer (AB SCIEX). MS/MS data were searched against an in-house molluscan protein database using Protein Pilot software (version 3.0, AB SCIEX). Peptides with a confidence of ≥99 and a false discovery rate of ≤5 were accepted for protein identifications.

Multienzyme-assisted Folding of Conotoxins

In conclusion, this study is the first to investigate the concerted contribution of a number of highly expressed, ER-resident *Conus* enzymes to the folding of conotoxins. All enzymes were shown to accelerate oxidative folding rates of two α-conotoxins *in vitro*, with the most efficient rates in the presence of PDI and PPI B or PDI, BiP, and microsomes. We were able to demonstrate multienzyme-assisted folding of the thermodynamically less favored ribbon isomer of α-ImI and reveal its presence in the venom of *C. imperialis*. We further provide the first evidence for the direct interaction between PDI and a nascent conotoxin, unequivocally establishing the role of this enzyme in the folding of cone snail peptides. Our subsequent investigations into the presence of multienzyme complexes led to the discovery of a novel *Conus* ER protein complex, strongly suggesting that conotoxins interact with a variety of enzymes to acquire their correct three-dimensional conformation. In effect, these enzymes form a molecular machine to process and fold these disulfide-rich toxins. Future studies addressing the full composition of this ER complex are likely to extend our understanding of the mechanisms underlying disulfide-rich peptide synthesis.

Acknowledgments—We thank Dr. Robyn Bradbury and John Ahern for specimen collection and maintenance, Dr. Yee-Foong Mok and Zachary Rosenes for assistance with the ATPase assay, Prof. Mike Hubbard for kindly providing the PDI antibody, Dr. Julita Imperial for advice on venom extraction, Dr. William Low for help with mass spectrometry, Srgjan Civciristov and Dr. Philippa Saunders for assistance with BN-PAGE, and Dr. Robert Schlaberg for reviewing the manuscript.

REFERENCES

1. Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A., and Rutter, W. J. (1985) Sequence of protein-disulfide isomerase and implications of its relationship to thioredoxin. *Nature* **317,** 267–270
2. Hawkins, H. C., and Freedman, R. B. (1991) The reactivities and ionization properties of the active-site dithiol groups of mammalian protein-disulfide isomerase. *Biochem. J.* **275,** 335–339
3. Schwaller, M., Wilkinson, B., and Gilbert, H. F. (2003) Reduction-reoxidation cycles contribute to catalysis of disulfide isomerization by protein-disulfide isomerase. *J. Biol. Chem.* **278,** 7154–7159
4. Ellgaard, L., and Ruddock, L. W. (2005) The human protein-disulfide isomerase family. Substrate interactions and functional properties. *EMBO Rep.* **6,** 28–32
5. Wilkinson, B., and Gilbert, H. F. (2004) Protein disulfide isomerase. *Biochim. Biophys. Acta* **1699,** 35–44
6. Galat, A. (2003) *Curr. Top. Med. Chem.* **3,** 1313–1347
7. Schönbrunner, E. R., and Schmid, F. X. (1992) Peptidyl-prolyl *cis-trans* isomerase improves the efficiency of protein-disulfide isomerase as a catalyst of protein folding. *Proc. Natl. Acad. Sci.* **89,** 4510–4513
8. di Luccio, E., Azulay, D. O., Regaya, I., Fajloun, Z., Sandoz, G., Mansuelle, P., Kharrat, R., Fathallah, M., Carrega, L., Estève, E., Rochat, H., De Waard, M., and Sabatier, J. M. (2001) Parameters affecting *in vitro* oxidation/folding of maurotoxin, a four-disulfide-bridged scorpion toxin. *Biochemical Journal* **358,** 681–692
9. Kleizen, B., and Braakman, I. (2004) Protein folding and quality control in the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **16,** 343–349
10. Mayer, M., Kies, U., Kammermeier, R., and Buchner, J. (2000) BiP and PDI cooperate in the oxidative folding of antibodies *in vitro*. *J. Biol. Chem.* **275,** 29421–29425
11. Jessop, C. E., Watkins, R. H., Simmons, J. J., Tasab, M., and Bulleid, N. J. (2009) Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins. *J. Cell. Sci.* **122,** 4287–4295
12. Okudo, H., Kato, H., Arakaki, Y., and Urade, R. (2005) Cooperation of ER-60 and BiP in the oxidative refolding of denatured proteins *in vitro*. *J. Biochem.* **138,** 773–780
13. Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol. Biol. Cell* **13,** 4456–4469
14. Buczek, O., Bulaj, G., and Olivera, B. M. (2005) Conotoxins and the post-translational modification of secreted gene products. *Cell. Mol. Life Sci.* **62,** 3067–3079
15. Davis, J. M., Jones, A., and Lewis, R. J. (2009) Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides* **30,** 11222–11227
16. Safavi-Hemami, H., Siero, W. A., Gorasia, D. G., Young, N. D., Macmillan, D., Williamson, N. A., and Purcell, A. W. (2011) Specialization of the venom gland proteome in predatory cone snails reveals functional diversification of the conotoxin biosynthetic pathway. *J. Proteome Res.* **10,** 3904–3919
17. Safavi-Hemami, H., Bulaj, G., Olivera, B. M., Williamson, N. A., and Purcell, A. W. (2010) Identification of *Conus* peptidylprolyl *cis-trans* isomerases (PPIases) and assessment of their role in the oxidative folding of conotoxins. *J. Biol. Chem.* **285,** 12735–12746
18. Wang, Z. Q., Han, Y. H., Shao, X. X., Chi, C. W., and Guo, Z. Y. (2007) Molecular cloning, expression, and characterization of protein-disulfide isomerase from *Conus marmoreus*. *FEBS J.* **274,** 4778–4787
19. Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment. Application to endoplasmic reticulum. *J. Cell Biol.* **93,** 97–102
20. Sedlak, J., and Lindsay, R. H. (1968) Estimation of total, protein-bound, and nonprotein sulfhydryl groups in tissue with Ellman's reagent. *Anal. Biochem.* **25,** 192–205
21. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. *Nat. Protoc.* **2,** 953–971
22. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4,** 2411–2423
23. Lambert, N., and Freedman, R. B. (1983) Kinetics and specificity of homogeneous protein-disulfide isomerase in protein disulfide isomerization and in thiol-protein-disulfide oxidoreduction. *Biochem. J.* **213,** 235–243
24. Fischer, G., Bang, H., and Mech, C. (1984) [Determination of enzymatic catalysis for the *cis-trans* isomerization of peptide binding in proline-containing peptides]. *Biomed. Biochim. Acta* **43,** 1101–1111
25. Kreuzer, K. N., and Jongeneel, C. V. (1983) *Escherichia coli* phage T4 topoisomerase. *Methods Enzymol.* **100,** 144–160
26. Safavi-Hemami, H., Siero, W. A., Kuang, Z., Williamson, N. A., Karas, J. A., Page, L. R., MacMillan, D., Callaghan, B., Kompella, S. N., Adams, D. J., Norton, R. S., and Purcell, A. W. (2011) Embryonic toxin expression in the cone snail *Conus victoriae*. Primed to kill or divergent function? *J. Biol. Chem.* **286,** 22546–22557
27. Buczek, O., Olivera, B. M., and Bulaj, G. (2004) Propeptide does not act as an intramolecular chaperone but facilitates protein-disulfide isomerase-assisted folding of a conotoxin precursor. *Biochemistry* **43,** 1093–1101
28. Kang, T. S., Radić, Z., Talley, T. T., Jois, S. D., Taylor, P., and Kini, R. M. (2007) Protein folding determinants. Structural features determining alternative disulfide pairing in α- and χ/λ-conotoxins. *Biochemistry* **46,** 3338–3355
29. Lopez-Vera, E., Walewska, A., Skalicky, J. J., Olivera, B. M., and Bulaj, G. (2008) Role of hydroxyprolines in the *in vitro* oxidative folding and biological activity of conotoxins. *Biochemistry* **47,** 1741–1751
30. Wittig, I., Braun, H. P., and Schägger, H. (2006) Blue native PAGE. *Nat. Protoc.* **1,** 418–428
31. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1,** 2856–2860
32. Gehrmann, J., Alewood, P. F., and Craik, D. J. (1998) Structure determi-

nation of the three disulfide bond isomers of α-conotoxin GI. A model for the role of disulfide bonds in structural stability. *J. Mol. Biol.* **278,** 401–415
33. Maslennikov, I. V., Shenkarev, Z. O., Zhmak, M. N., Ivanov, V. T., Methfessel, C., Tsetlin, V. I., and Arseniev, A. S. (1999) NMR spatial structure of α-conotoxin ImI reveals a common scaffold in snail and snake toxins recognizing neuronal nicotinic acetylcholine receptors. *FEBS Lett.* **444,** 275–280
34. Hatahet, F., and Ruddock, L. W. (2009) Protein-disulfide isomerase. A critical evaluation of its function in disulfide bond formation. *Antioxid. Redox Signal.* **11,** 2807–2850
35. Bulaj, G., Buczek, O., Goodsell, I., Jimenez, E. C., Kranski, J., Nielsen, J. S., Garrett, J. E., and Olivera, B. M. (2003) Efficient oxidative folding of conotoxins and the radiation of venomous cone snails. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 14562–14568
36. Fuller, E., Green, B. R., Catlin, P., Buczek, O., Nielsen, J. S., Olivera, B. M., and Bulaj, G. (2005) Oxidative folding of conotoxins sharing an identical disulfide bridging framework. *FEBS J.* **272,** 1727–1738
37. Gowd, K. H., Krishnan, K. S., and Balaram, P. (2007) Identification of *Conus amadis* disulfide isomerase. Minimum sequence length of peptide fragments necessary for protein annotation. *Mol. Biosyst.* **3,** 554–566
38. Brodsky, J. L., and Skach, W. R. (2011) Protein folding and quality control in the endoplasmic reticulum. Recent lessons from yeast and mammalian cell systems. *Curr. Opin. Cell Biol.* **23,** 464–475
39. Tsai, B., Rodighiero, C., Lencer, W. I., and Rapoport, T. A. (2001) Protein-disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* **104,** 937–948
40. Dutton, J. L., Bansal, P. S., Hogg, R. C., Adams, D. J., Alewood, P. F., and Craik, D. J. (2002) A new level of conotoxin diversity, a non-native disulfide bond connectivity in α-conotoxin AuIB reduces structural definition but increases biological activity. *J. Biol. Chem.* **277,** 48849–48857
41. Nicke, A., Samochocki, M., Loughnan, M. L., Bansal, P. S., Maelicke, A., and Lewis, R. J. (2003) α-Conotoxins EpI and AuIB switch subtype selectivity and activity in native versus recombinant nicotinic acetylcholine receptors. *FEBS Lett.* **554,** 219–223
42. Townsend, A., Livett, B. G., Bingham, J. P., Truong, H. T., Karas, J. A., O'Donnell, P., Williamson, N. A., Purcell, A. W., and Scanlon, D. (2009) Mass spectral identification of Vc1.1 and differential distribution of conopeptides in the venom duct of *Conus victoriae*. Effect of post-translational modifications and disulfide isomerisation on bioactivity. *Int. J. Pept. Res. Ther.* **15,** 195–203
43. Ellis, R. J. (2001) Macromolecular crowding. Obvious but underappreciated. *Trends Biochem. Sci.* **26,** 597–604
44. Szabo, A., Langer, T., Schröder, H., Flanagan, J., Bukau, B., and Hartl, F. U. (1994) The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system DnaK, DnaJ, and GrpE. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 10345–10349
45. Rioux, V., Gerbod, M. C., Bouet, F., Ménez, A., and Galat, A. (1998) Divergent and common groups of proteins in glands of venomous snakes. *Electrophoresis* **19,** 788–796
46. Birrell, G. W., Earl, S. T., Wallis, T. P., Masci, P. P., de Jersey, J., Gorman, J. J., and Lavin, M. F. (2007) The diversity of bioactive proteins in Australian snake venoms. *Mol. Cell. Proteomics* **6,** 973–986
47. Fernandes-Pedrosa Mde, F., Junqueira-de-Azevedo, I. D., Gonçalves-de-Andrade, R. M., Kobashi, L. S., Almeida, D. D., Ho, P. L., and Tambourgi, D. V. (2008) Transcriptome analysis of *Loxosceles laeta* (*Araneae*, *Sicariidae*) spider venomous gland using expressed sequence tags. *BMC Genomics* **9,** 279
48. Winter, J., Klappa, P., Freedman, R. B., Lilie, H., and Rudolph, R. (2002) Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin. *J. Biol. Chem.* **277,** 310–317
49. Zhang, L., Lai, E., Teodoro, T., and Volchuk, A. (2009) GRP78, but not protein-disulfide isomerase, partially reverses hyperglycemia-induced inhibition of insulin synthesis and secretion in pancreatic β-cells. *J. Biol. Chem.* **284,** 5289–5298
50. Corpuz, G. P., Jacobsen, R. B., Jimenez, E. C., Watkins, M., Walker, C., Colledge, C., Garrett, J. E., McDougal, O., Li, W., Gray, W. R., Hillyard, D. R., Rivier, J., McIntosh, J. M., Cruz, L. J., and Olivera, B. M. (2005) Definition of the M-conotoxin superfamily. Characterization of novel peptides from molluscivorous *Conus* venoms. *Biochemistry* **44,** 8176–8186
51. Cruz, L. J., Kupryszewski, G., LeCheminant, G. W., Gray, W. R., Olivera, B. M., and Rivier, J. (1989) μ-Conotoxin GIIIA, a peptide ligand for muscle sodium channels. Chemical synthesis, radiolabeling, and receptor characterization. *Biochemistry* **28,** 3437–3442
52. Wang, X., Smith, R., Fletcher, J. I., Wilson, H., Wood, C. J., Howden, M. E., and King, G. F. (1999) Structure-function studies of ω-atracotoxin, a potent antagonist of insect voltage-gated calcium channels. *Eur. J. Biochem.* **264,** 488–494
53. Kharrat, R., Mansuelle, P., Sampieri, F., Crest, M., Oughideni, R., Van Rietschoten, J., Martin-Eauclaire, M. F., Rochat, H., and El Ayeb, M. (1997) Maurotoxin, a four-disulfide bridge toxin from *Scorpio maurus* venom. Purification, structure, and action on potassium channels. *FEBS Lett.* **406,** 284–290
54. Huang, Y. H., Colgrave, M. L., Daly, N. L., Keleshian, A., Martinac, B., and Craik, D. J. (2009) The biological activity of the prototypic cyclotide kalata b1 is modulated by the formation of multimeric pores. *J. Biol. Chem.* **284,** 20699–20707
55. Gruber, C. W., Cemazar, M., Clark, R. J., Horibe, T., Renda, R. F., Anderson, M. A., and Craik, D. J. (2007) A novel plant protein-disulfide isomerase involved in the oxidative folding of cystine knot defense proteins. *J. Biol. Chem.* **282,** 20435–20446
56. Hamanaka, Y., Nakashima, M., Wada, A., Ito, M., Kurazono, H., Hojo, H., Nakahara, Y., Kohno, S., Hirayama, T., and Sekine, I. (2001) Expression of human β-defensin 2 (hBD-2) in *Helicobacter pylori*-induced gastritis. Antibacterial effect of hBD-2 against *Helicobacter pylori*. *Gut* **49,** 481–487
57. Murray, I. (1971) Paulesco and the isolation of insulin. *J. Hist. Med. Allied Sci.* **26,** 150–157
58. Forte, L. R., and Currie, M. G. (1995) Guanylin. A peptide regulator of epithelial transport. *FASEB J.* **9,** 643–650
59. Ascenzi, P., Bocedi, A., Bolognesi, M., Spallarossa, A., Coletta, M., De Cristofaro, R., and Menegatti, E. (2003) The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor). A milestone protein. *Curr. Protein Pept. Sci.* **4,** 231–251
60. Ostermeier, M., De Sutter, K., and Georgiou, G. (1996) Eukaryotic protein-disulfide isomerase complements *Escherichia coli* dsbA mutants and increases the yield of a heterologous secreted protein with disulfide bonds. *J. Biol. Chem.* **271,** 10616–10622
61. Gonzalez, R., Andrews, B. A., and Asenjo, J. A. (2002) Kinetic model of BiP- and PDI-mediated protein folding and assembly. *J. Theor. Biol.* **214,** 529–537
62. Marcinkiewicz, C., Weinreb, P. H., Calvete, J. J., Kisiel, D. G., Mousa, S. A., Tuszynski, G. P., and Lobb, R. R. (2003) Obtustatin. A potent selective inhibitor of α1β1 integrin *in vitro* and angiogenesis *in vivo*. *Cancer Res.* **63,** 2020–2023

APPENDIX C

THE EVOLUTIONARY BASIS FOR CONSERVATION OF DISULFIDE SCAFFOLDS

Citation: Gowd, K.H.; Blais, K.D.; Elmslie, K.S.; Steiner, A.M.; Olivera, B.M.; Bulaj, G. (2012) Dissecting a role of evolutionary-conserved but noncritical disulfide bridges in cysteine-rich peptides using ω-conotoxin GVIA and its selenocysteine analogs. Biopolymers, 98: 212-223. doi: 10.1002/bip.22047.

Dissecting a Role of Evolutionary-Conserved but Noncritical Disulfide Bridges in Cysteine-Rich Peptides Using ω-Conotoxin GVIA and Its Selenocysteine Analogs

Konkallu Hanumae Gowd,[1,2] Kirk D. Blais,[3] Keith S. Elmslie,[3] Andrew M. Steiner,[1] Baldomero M. Olivera,[4] Grzegorz Bulaj[4]

[1]*Department of Biology, University of Utah, Salt Lake City, UT 84112*

[2]*Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India*

[3]*Department of Pharmacology, Kirksville College of Osteopathic Medicine, AT Still University, Kirksville, MO 63501*

[4]*Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84108*

Received 14 November 2011; revised 23 January 2012; accepted 4 February 2012
Published online 9 March 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22047

ABSTRACT:

Conotoxins comprise a large group of peptidic neurotoxins that use diverse disulfide-rich scaffolds. Each scaffold is determined by an evolutionarily conserved pattern of cysteine residues. Although many structure–activity relationship studies confirm the functional and structural importance of disulfide crosslinks, there is growing evidence that not all disulfide bridges are critical in maintaining activities of conotoxins. To answer the fundamental biological question of what the role of noncritical disulfide bridges is, we investigated function and folding of disulfide-depleted analogs of ω-conotoxin GVIA (GVIA) that belongs to an inhibitory cystine knot motif family and blocks N-type calcium channels. Removal of a noncritical Cys1–Cys16 disulfide bridge in GVIA or its selenopeptide analog had, as predicted, rather minimal effects on the inhibitory activity on calcium channels, as well as on in vivo activity following intracranial administration. However, the disulfide-depleted GVIA exhibited significantly lower folding yields for forming the remaining two native disulfide bridges. The disulfide-depleted selenoconotoxin GVIA analog also folded with significantly lower yields, suggesting that the functionally noncritical disulfide pair plays an important cooperative role in forming the native disulfide scaffold. Taken together, our results suggest that distinct disulfide bridges may be evolutionarily preserved by the oxidative folding or/and stabilization of the bioactive conformation of a disulfide-rich scaffold. © 2012 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 98: 212–223, 2012.

Keywords: disulfide bridges; conotoxins, structure–function; oxidative folding; calcium channels

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

Correspondence to: Grzegorz Bulaj, Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84108; e-mail: bulaj@pharm.utah.edu
Conflicts of interest: B.M.O. is a cofounder of Cognetix Inc; G.B. is a cofounder of NeuroAdjuvants Inc.

INTRODUCTION

Nature uses disulfides in the construction of stable structural motifs in peptides that are secreted by the cell. Peptides containing multiple disulfide bonds are better protected from stress and unfolding conditions found in extracellular environments.[1–6] One

such class of disulfide-rich natural peptides having well-defined structures are conotoxins, neuroactive modulators derived from marine cone snails.[7–10] Conotoxins are short peptides of varying length from 6 to 50 residues and containing 1–5 disulfide bonds that provide characteristic scaffolds to the peptides and stabilize the biologically active conformation.[10] Although multiple disulfide bonds generally contribute to the stability of native conotoxin fold, the removal of one of the native disulfide bond was found to not affect their structure or function in conkunitzin-S1,[11] μ-conotoxin KIIIA,[12] ω-conotoxin MVIIA,[13] α-conotoxin ImI,[14] and α-conotoxin GI.[15] These findings indicate that disulfides are differentially associated with structure and function of disulfide rich conopeptides (a disulfide can be critical for structure or function or both), which could be dissected by studying the corresponding disulfide-depleted peptides. The best example of one such study is ω-conotoxin MVIIA, an N-type Ca^{2+} channel blocker isolated from the marine snail *Conus magus* having the inhibitory cystine knot (ICK) motif.[13,16] The fundamental question of why nature conserves functionally noncritical disulfide bridges still remains unanswered.

ICK motif is a common structural motif found in peptides of diverse origin including toxins from plants, antimicrobial peptides from insects, and in venoms of scorpions, spiders and cone snails.[17–21] They share a unique knotted topology consisting of three disulfide bridges, with one disulfide crossing a loop formed by the two other disulfides and interconnecting peptide backbone. The ICK motif containing conotoxins with the cysteine pattern C—C—CC—C—C are widely characterized from *Conus* venom; ω-conotoxin (inhibitor of Ca^{2+} channel),[22] κ-conotoxin (modulator of K^+ channel),[23] δ-conotoxin (delay inactivation of Na^+ channel),[24] and μO-conotoxin (inhibitor of Na^+ channel activation).[25,26] ω-Conotoxins are the best characterized ICK motif containing conopeptides (Figure 1a), including ω-conotoxin MVIIA, which is an Food and Drug Administration-approved drug for the treatment of intractable pain (ziconotide, Prialt).[27–30] The role of disulfide bonds in structure–activity relationships of ω-conotoxins were investigated by Sabo et al.,[31] Norton and coworkers[32] and the Goldenberg group.[13] As summarized in Table I, the first disulfide bond has a minimal contribution to the activity, whereas removal of either the second or third disulfide bond drastically affected the activity of peptide.

ω-Conotoxin GVIA (Figure 1b) is a standard pharmacological tool for inhibiting synaptic transmission.[29,30] Previous studies by the Sabo group and the Norton group showed that the disulfide bond between Cys1 and Cys16 in GVIA had minimal contribution to the activity.[31,32] Based on these findings, we selected GVIA to ask a more general fundamental question: what is the role of a functionally noncritical disulfide bridge? Keeping in mind that the functional role was assessed based on in vitro assays, one possibility was that the disulfide bridge might be important for in vivo activity (when injected into an animal). We also hypothesized a potential role of a functionally noncritical disulfide in oxidative folding, which affects efficient biosynthesis of a disulfide-rich scaffold, rather than in stabilizing the "end-point" bioactive conformation. To address the above questions, we reassessed the role of the functionally noncritical disulfide in folding and activity of ω-conotoxin GVIA using disulfide-depleted analogs, namely dd-GVIA and dd-Sec-GVIA (Figure 2). Replacing disulfide bridges with diselenide bridges was previously shown to be an effective strategy to simplify the oxidative folding of cysteine-rich peptides, without affecting their native conformation or biological activity.[33–43] Our findings indicate that the first disulfide bridge (Cys1–Cys16) in ω-conotoxin GVIA is important for directing the correct formation of the other two native disulfides. The results are discussed in the evolutionary context of conserving the functionally noncritical disulfides in cysteine-rich natural peptides.

MATERIALS AND METHODS

Peptide Synthesis

Peptides were synthesized using solid phase peptide synthesis methodology, with Fmoc chemistry and activated OPfp esters of protected amino acids.

Natively Folded dd-GVIA. Side chains of one pair of natively connected cysteines were protected with trityl (Trt) group, and the other pair of natively connected cysteines was protected with the acetamidomethyl group. Peptide cleaved from the resin and side chains were deprotected using reagent K (trifluoroacetic acid [TFA]/thianisol/phenol/water/ethanedithiol [82.5:5:5:5:2.5]), precipitated with MTBE, and washed several times with MTBE. The peptide was purified using C_{18} RP-HPLC over linear gradient of 10–40% buffer B (90% acetonitrile containing 0.1% TFA) for 40 min. The purified peptide contains two acetamidomethyl-protected cysteines and two free thiols. Disulfide bond formation between the free thiols was achieved by incubating 20 μM peptide in a folding buffer containing 100 m*M* Tris-HCl (pH 7.5), 1 m*M* oxidized glutathione (GSSG), and 2 m*M* reduced glutathione (GSH). A 20-μM single disulfide containing peptide was incubated with 10 m*M* iodine dissolved in 50% of acetonitrile containing 1% TFA for simultaneous deprotection of acetamidomethyl groups and formation of disulfide bond between corresponding cysteines. The reaction proceeded at room temperature for 5 min by gentle shaking in a glass vial and subsequently quenched by drop wise addition of 1 *M* ascorbic acid. The natively folded peptide was purified by RP-HPLC and characterized using mass spectrometry {$[M + H]^+_{oxi}$ = 2974.2 Da (calculated), 2974.3 Da (observed)}.

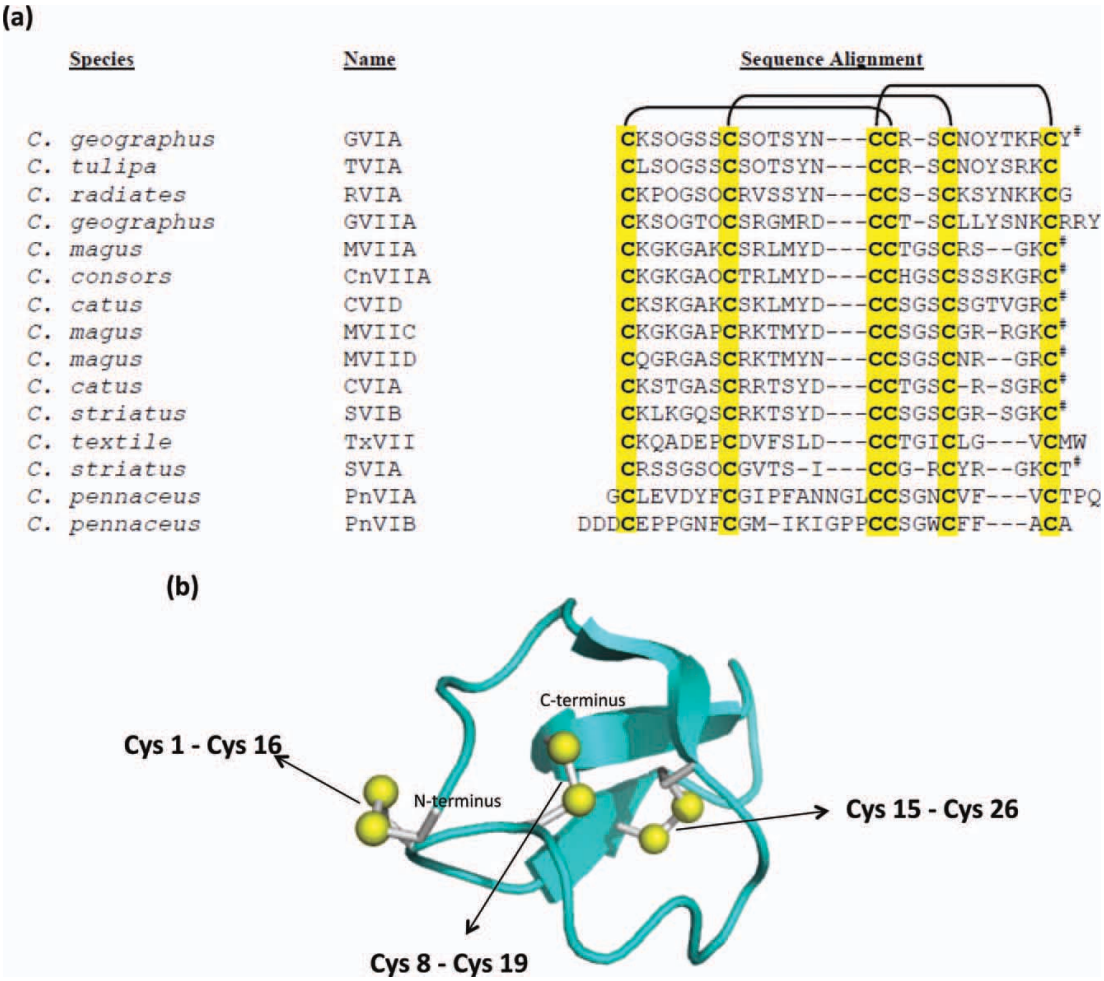


FIGURE 1 ω-Conotoxins derived from the marine cone snails. a: Multiple sequence alignment of diverse ω-conotoxins. Absolutely conserved residues were shaded and they include Cys residues. Observed modes of disulfide connectivity in ω-conotoxins were also highlighted. Sequences were retrieved from conoserver, alignments were achieved using ClustalW software and observed posttranslational modifications were manually incorporated (Note: Number sign [#] indicates C-terminal amidation). b: The orientation in three-dimensional space of the conserved Cys residues of ω-conotoxins, depicted using ω-conotoxin GVIA as an example. The structure of GVIA (PDBID: 2CCO) was modeled using PyMOL software.

Natively Folded dd-Sec-GVIA. The side chains of selenocysteines were protected with a *p*-methoxybenzyl (Mob) group and cysteines were protected with trityl (Trt) groups. Peptides were removed from resin, and side chains were deprotected using enriched reagent K (TFA/thianisol/phenol/water [90:2.5:7.5:5] and 1.3 equiv of DTNP [2,2′-dithiobis(5-nitropyridine)]) and precipitated using MTBE. Crude peptide was treated with 50 m*M* dithiothreitol (DTT) in 100 m*M* Tris-HCl (pH 7.5) and subsequently purified using preparative RP-HPLC with a C_{18} column over a linear gradient of 10–40% buffer B (90% acetonitrile containing 0.1% TFA) for 40 min. Observed mass of the peptide was 2 Da less than the predicted mass, confirming the presence of a preformed diselenide in the linear peptide $\{[M + H]^+_{red.}= 3074.2$ Da (calculated), 3072.4 Da (observed)$\}$. This was further supported by an alkylation reaction using iodoacetamide; the resulting species was shown by mass spectrometry to contain a diselenide bridge.[38,39] A 20-μ*M* of diselenide containing peptide was treated with 10 m*M* iodine dissolved in 50% of acetonitrile containing 1% TFA for formation of disulfide bond between free thiols. The reaction proceeded at room temperature for 5 min by gentle shaking in a glass vial and subse-

Table I Summary of the Role of Disulfide Bonds on the Activity of ω-Conotoxins

Peptide	Disulfide Connectivity	Biological Activity	References
GVIA	1-16; 8-19; 15-26	63[a]	31
[C1A,C16A] GVIA	ND	933[a]	31
[C8A,C19A] GVIA	ND	NA[a]	31
[C15A,C26A] GVIA	ND	NA[a]	31
GVIA	1-16; 8-19; 15-26	10 n*M*[b]	32
[C1S,C16S] GVIA	8-15; 19-26	> 1 μM[b]	32
[C8S,C19S] GVIA	ND	> 1 μM[b]	32
[C15S,C26S] GVIA	1-16; 8-19	> 10 μM[b]	32
ω-MVIIA-Gly	1-16; 8-20; 15-25	2.5 n*M*[c]	13
$[8\text{-}20, 15\text{-}25]_{1,16\ Ala}$	8-20; 15-25	170 n*M*[c]	13
$[1\text{-}16, 15\text{-}25]_{8,20\ Ala}$	1-16; 15-25	517 n*M*[c]	13
$[1\text{-}16, 8\text{-}20]_{15,25\ Ala}$	1-16; 8-20	13000 n*M*[c]	13
RCM ω-MVIIA-Gly	Linear peptide	37000 n*M*[c]	13

ND, not determined; NA, not active (LD_{50} > 6000 μg/kg).

[a] Biological activity of the peptide was determined by their LD_{50} (μg/kg) in the gold fish assay.

[b] Activity of the peptides was accessed using the twitch responses of the rat vas deferens to sympathetic nerve stimulation.

[c] Functions of the peptides was established by competitive binding assay using ^{125}I-labeled ω-GVIA on chick brain synaptosomes.

quently quenched by drop wise addition of 1 *M* ascorbic acid. The natively folded peptide was purified by RP-HPLC and characterized using mass spectrometry {$[M + H]^{+}_{oxi.}$= 3070.2 Da (calculated), 3070.3 Da (observed)}.

Oxidative Folding

Free thiol containing linear peptide dd-GVIA was obtained by incubating natively folded dd-GVIA in 10 m*M* DTT containing 10 m*M* Tris-Hcl (pH 7.5) at 37°C, purified by RP-HPLC and characterized using mass spectrometry. Note that the linear peptide dd-Sec-GVIA contains a preformed diselenide bridge. The folding reaction (or) stability assay was initiated by resuspending 5 nmol of linear peptide or natively folded peptide into 200 μL of folding buffer containing 100 m*M* Tris-HCl (pH 7.5), 1 m*M* EDTA, 1 m*M* GSSG, and 2 m*M* GSH. The reaction was quenched after 2 h by acidification with formic acid (10% final concentration). Samples were analyzed using C_{18} analytical HPLC over a linear gradient of 10–40% buffer B (90% acetonitrile containing 0.1% TFA) in 40 min. Accumulation of natively folded peptide at the steady state was calculated by integrating the HPLC chromatogram.

Proteolytic Digestion and Mass Spectrometry

The disulfide/selenosulfide connectivity in non-natively folded dd-GVIA and dd-Sec-GVIA were studied by coupling tryptic digestion with mass spectrometry. Five nanomoles of non-natively folded dd-GVIA (or) dd-Sec-GVIA were dissolved in 100 m*M* Tris-Hcl (pH 7.5), 5 μL of trypsin (25ng/μL) was added and reaction mixture was incubated at 37°C for 4–8 h. The resulting peptide fragments were separated using analytical RP-HPLC over a linear gradient of 10–80% buffer B (90% acetonitrile containing 0.1% TFA) in 60 min and the fractions were subsequently analyzed using mass spectrometry. The MALDI-mass spectra were acquired at the Mass Spectrometry and Proteomic Core facility of the University of Utah.

Measurement of N-Type Calcium Currents (CaV2.2)

HEK Cell Transfection. HEK293 cells were maintained in DMEM/Glutamax medium containing 10% fetal bovine serum and 1% antibiotic-antimycotic (Invitrogen, Carlsbad, CA), at 37°C in 5% CO_2 incubator. The cells were transfected using Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. The complementary DNA plasmids were 1.93 μg $\alpha_{1B\text{-}CFP}$, 1.55 μg $\alpha_2\delta$, and 0.97 μg β_{2a} subunits. N-channel expressing cells were visualized by CFP, which was attached to the N-terminus of α_{1B} ($\alpha_{1B\text{-}CFP}$; from Dr. Blaise Peterson, Penn State College of Medicine). Transfected cells were split into 35 mm dishes that also served as the recording chamber.

N-Current Measurement and Analysis. HEK cells were voltage-clamped as described previously[44,45] in an external solution containing (in m*M*): 145 *N*-methyl-D-glucamine (NMG) · Cl, 10 NMG · HEPES, 5 $BaCl_2$. The internal (pipette) solution contained (in m*M*): 104 NMG · Cl, 14 creatine · PO_4, 6 $MgCl_2$, 10 NMG · HEPES, 10 NMG_2 · EGTA, 5 Tris · ATP, and 0.3 Tris · GTP. The osmolarity of the external solution was 300 mOsm, while that of the internal solution was 280 mOsm. Both the internal and external solutions were adjusted to pH 7.4 using NMG base. Leak current was subtracted online using a -P/4 protocol for ionic currents. All recordings were carried out at room temperature and the holding potential was −120 mV. Ionic currents were digitized at 50 kHz after analog filtering at 5–10 kHz. Solutions were applied using a gravity-fed perfusion system with an exchange time of 1–2 s. All toxin concentrations were 1 μM. Data were analyzed using IgorPro (WaveMetrics, Lake Oswego, OR) as described previously.[38]

Behavioral Assay

The natively folded GVIA, Sec-GVIA, dd-GVIA, and dd-Sec-GVIA and the misfolded dd-GVIA and dd-Sec-GVIA were injected intracranially to 21- to 23-day-old Swiss Webster mice using a syringe with a 29-gauge needle. 1 nmol of each peptide was dissolved separately in 15 μL of saline and subsequently injected in mice; an equal volume of normal saline was used as a control. After injection, mice were placed in a cage for observation. All of the native peptides exhibited the shaking syndrome, which is characterized by a persistent body tremor, and this behavior was maintained for a long time in GVIA and Sec-GVIA, as compared to dd-GVIA and dd-Sec-GVIA. Mice injected with misfolded dd-GVIA and dd-Sec-GVIA did not exhibit the shaking syndrome; rather, the injected mice were quite passive, occasionally shaking their bodies.

RESULTS

Design of dd-GVIA and dd-Sec-GVIA

To investigate the role of the functionally noncritical disulfide in the folding and activity of ω-conotoxin GVIA, dd-GVIA, and dd-Sec-GVIA were designed and synthesized. Table I

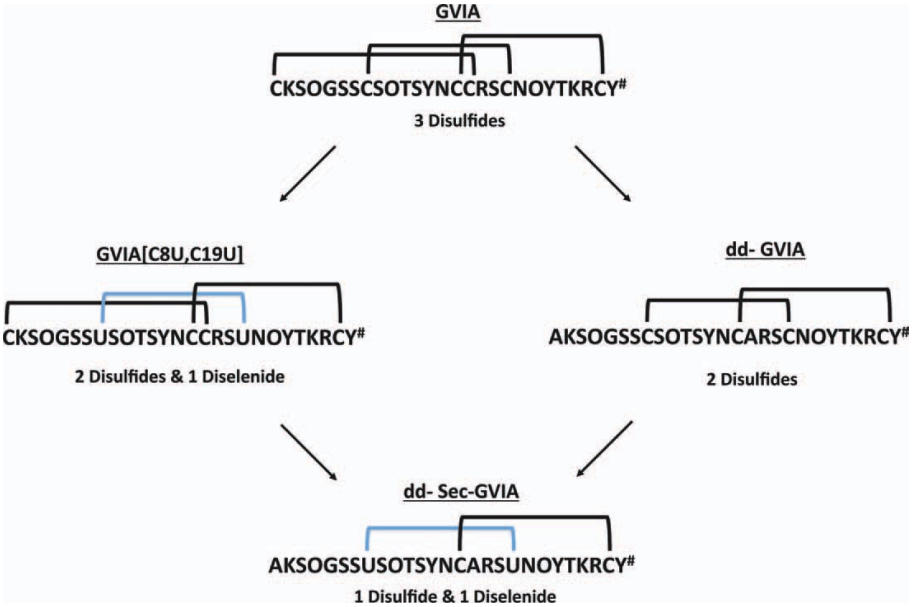


FIGURE 2 Design of disulfide-depleted seleno-ω-conotoxin GVIA was accomplished using redox favored selenocysteine and removing the functionally noncritical disulfide. Systematic disulfide-to-diselenide scanning in ω-conotoxin GVIA suggested that the second disulfide bridge is ideal for replacement with selenocysteines.[38] Structure/function studies on the role of disulfide bridges in ω-conotoxin GVIA suggested that the removal of first disulfide bridge had relatively small effects on the functional properties of ω-conotoxin GVIA.[31,32] Combination of disulfide deletion and incorporation of a diselenide in ω-conotoxin GVIA resulted in dd-Sec-GVIA, which requires the formation of only one disulfide, thus removing the need for disulfide mapping of the folded peptides (Note: Number sign [#] indicates C-terminal amidation).

summarizes the effect of removal of a disulfide bond on the activity of ω-conotoxins. It is apparent from Table I that removal of the first disulfide bridge in ω-conotoxins resulted in decreased activity by 15- to 70-fold. The loss of activity observed by removal of any other native disulfides is 200–5000 fold. Hence, the removal of first disulfide bridge in ω-conotoxin GVIA results in peptide that retains the ability to modulate N-type calcium currents. It is worth mentioning that the previous studies on the role of removing the first disulfide in ω-conotoxin GVIA are not conclusive because the disulfide connectivity of the peptides used in these studies were unknown[31] or misfolded.[32,46]

The removal of the first disulfide bridge in ω-conotoxin GVIA resulted in two-disulfide containing dd-GVIA. The removal of first disulfide bridge in ω-conotoxin GVIA reduces the possible number of disulfide isomers from 15 to 3. Figure 2 shows a schematic representation of the design of disulfide-depleted-seleno-ω-conotoxin GVIA (dd-Sec-GVIA). The design of dd-Sec-GVIA involves incorporation of selenocysteine into dd-GVIA. Systematic replacement of native disulfide bridges with a diselenide bridge in ω-conotoxin GVIA has resulted in increased yield of natively folded peptide in all Sec-GVIA analogs. The most prominent effect was observed when the second native disulfide bridge was replaced with a diselenide bridge (net yield enhancement of natively folded peptide is 19%).[38] Thus, removal of the functionally noncritical first disulfide bridge and incorporation of a diselenide bridge in the second native disulfide position in ω-conotoxin GVIA resulted in biologically active dd-Sec-GVIA. The incorporation of a diselenide bridge in dd-GVIA has further simplified the folding problem and eliminated the need for disulfide mapping in the folded peptide (dd-Sec-GVIA needs to form only one disulfide bond to achieve the natively folded state).

Synthesis of Natively Folded dd-GVIA and dd-Sec-GVIA

The natively folded dd-GVIA was accessed using an orthogonal protection scheme; one pair of natively connected

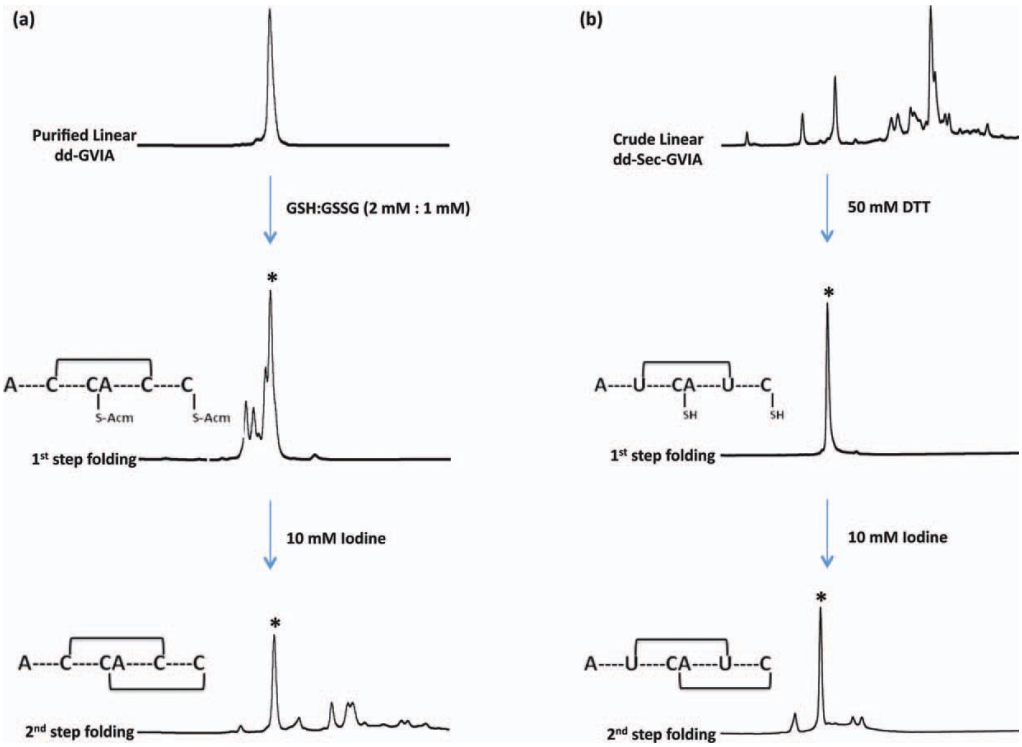


FIGURE 3 Comparison of regioselective folding of dd-GVIA and dd-Sec-GVIA. a: Two-step oxidative folding of dd-GVIA. The first disulfide bond formation was mediated by redox buffer (GSH/GSSG) and second disulfide bond formation was achieved using iodine. b: One-step oxidative folding of dd-Sec-GVIA. Crude peptide was treated with DTT to remove the DTNP adduct, which is followed by spontaneous diselenide bond formation. The diselenide-directed folding requires only one folding step (disulfide bond formation using iodine) to achieve the natively folded peptide. (Note: Asterisk [*] indicates the desired product).

cysteines were protected with the trityl (Trt) group (Cys8 and Cys19) and the other pair of natively connected cysteines were protected with the acetamidomethyl (Acm) group (Cys15 and Cys26). Figure 3a shows the regioselective folding of dd-GVIA. Upon cleavage from the resin, TFA-labile trityl groups yielded free thiols, and TFA-resistant acetamidomethyl groups retain protected thiols, as described in Materials and Methods section. The disulfide bond formation between free thiols was achieved using a GSSG/GSH buffer, resulting in formation of one native disulfide bridge. The deprotection of the Acm group and simultaneous disulfide bond formation between corresponding thiols was achieved using molecular iodine, yielding natively folded dd-GVIA. The identity of natively folded dd-GVIA was confirmed by mass spectrometric studies following the two successive folding steps, each of which resulted in a single major peak. Synthesis of natively folded dd-Sec-GVIA was achieved as described in Materials and Methods section, Sec residues were incorporated into the sequence using *p*-methoxybenzyl-protected selenocysteine and side chain deprotection was achieved using DTNP. Figure 3b shows the regioselective folding of dd-Sec-GVIA. The crude peptide obtained following cleavage from the resin contains two DTNP-selenocysteine adducts, which were subsequently removed by thiolysis using 50 m*M* DTT. The spontaneous oxidation of the selenols to a diselenide yields a linear dd-Sec-GVIA containing a preformed diselenide bridge, which was actually detected.[37–41] Observed mass of linear dd-Sec-GVIA analogue was 2 Da less than the predicted mass and had two alkylation sites (iodoacetamide was used as alkylating agent), confirming the presence of a preformed diselenide bridge. The formation of a disulfide bond between free thiols was achieved using molecular iodine, which leads to rapid and irreversible disulfide bond formation yielding natively folded dd-Sec-GVIA.

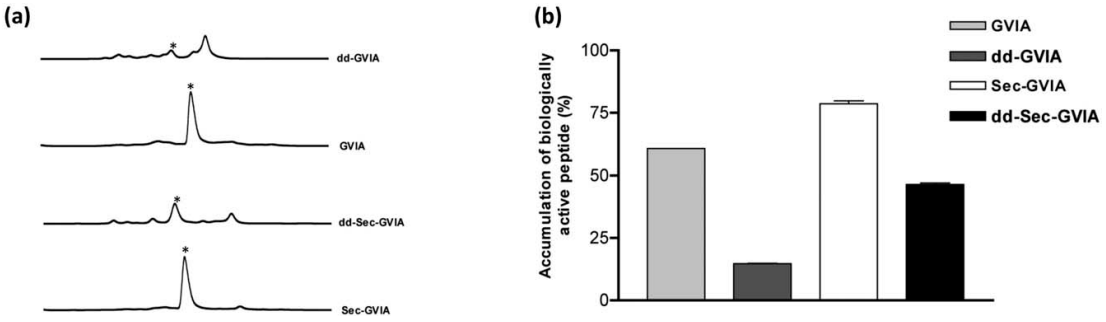


FIGURE 4 The role of the biologically noncritical disulfide in the oxidative folding of ω-conotoxin GVIA and Sec-GVIA. a: RP-HPLC elution profiles of steady state folding mixtures of GVIA, Sec-GVIA, and the corresponding disulfide-depleted analogues. Incubation of linear (or) natively folded peptides in the folding mixture containing GSH/GSSG (2:1) at steady-state yielded an identical elution profile (Note: Asterisk [*] indicates the biologically active and natively folded peptides). b: Steady state accumulation of biologically active and natively folded peptide during oxidative folding (at room temperature). Error bars represent the standard error of the mean, derived from four independent experiments.

Similar to dd-GVIA, the elution profiles of the folding pathway of dd-Sec-GVIA yielded a single major peak corresponding to folded peptides. Mass spectrometric studies confirmed the identity of natively folded dd-Sec-GVIA. It is evident from Figure 3 that the regioselective folding of dd-GVIA consists of two folding steps and that of dd-Sec-GVIA consists of only one folding step, demonstrating that the incorporation of selenocystine simplifies the oxidative folding.

Oxidative Folding Studies of dd-GVIA and dd-Sec-GVIA

Incubation of the native forms of disulfide-rich peptides in redox buffers, such as a mixture of GSSG/GSH, can be used as a measure of the stability of the corresponding disulfide isomer at the given redox potential.[36] The glutathione mediated folding of disulfide rich peptides involves thiol-disulfide exchange reactions and a steady-state folding mixture reflects the thermodynamic stability of the resulting disulfide isomers of cysteine-rich peptides. To investigate the stability of natively folded GVIA analogs, natively folded dd-GVIA and dd-Sec-GVIA were incubated in redox-buffer containing GSSG (1 m*M* GSSG) and GSH (2 m*M* GSH); these experimental conditions were shown previously to promote efficient folding of GVIA.[47,48]

Figure 4a shows elution profiles of steady-state folding mixtures of dd-GVIA, dd-Sec-GVIA and the corresponding peptide GVIA and Sec-GVIA. Similar elution profiles were obtained by incubating corresponding linear peptides in the folding buffer under identical experimental conditions. The elution profile of dd-GVIA at the steady state, in the presence of the glutathione mixture retains about 15% of natively folded peptide and that of GVIA retains about 60% of natively folded peptide. The major peak accumulated at the steady state in dd-GVIA corresponds to the non-natively folded species (observed mass is 2974 Da). The tryptic digest of non-natively folded dd-GVIA yielded fragments of mass 1745 Da and 1247/1265Da, which could be ascribed to peptide fragment AKSOGSSCSOTSYNCAR and SCNOYTKRCY$^{\#}$ (Note: O denotes hydroxyproline, $^{\#}$ denotes an amidated C-terminus and cysteines are disulfide bonded), respectively. The disulfide connectivity in non-natively folded dd-GVIA is Cys8–Cys15 and Cys19–Cys26. These results were further supported by the folding studies of [C1S, C16S]GVIA by Flinn et al.,[32] where the major folded fraction corresponds to the Cys8–Cys15 and Cys19–Cys26 connectivity.

The elution profile of dd-Sec-GVIA at the steady state, in the presence of the glutathione mixture, retains about 45% of natively folded peptide and that of Sec-GVIA retains about 78% of natively folded peptide. The major satellite peak accumulated at the steady state in dd-Sec-GVIA has the mass 3070 Da, indicating the formation of non-natively folded and seleno-sulfide containing dd-Sec-GVIA. The satellite peak was attributed to the isomer Sec8–Cys15 and Sec19–Cys26, since its retention time and shift in retention time (as compared to that of natively folded peptide) was similar to non-natively folded dd-GVIA having the Cys8–Cys15 and Cys19–Cys26 connectivity. To further confirm the presence of seleno-sulfide bonds in the satellite peak, the peptide was subjected to tryptic digestion and the resulting fragments were analyzed using mass spectrometry. Incubation of the satellite peak of dd-Sec-GVIA with trypsin yielded fragments

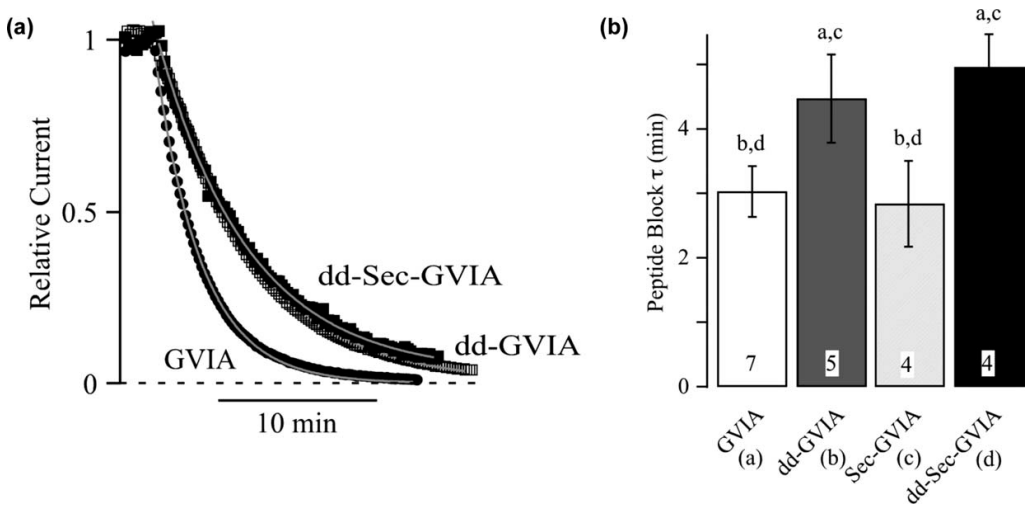


FIGURE 5 The blocking effect of GVIA, Sec-GVIA, and their corresponding disulfide-depleted analogues on the N-type calcium channel. a: The blocking time courses from three N-type calcium channel-expressing HEK cells exposed to 1 μM of either GVIA (closed circles), dd-GVIA (open squares), or dd-Sec-GVIA (closed squares). The smooth gray lines are single exponential fits that begin with the toxin onset. b: Peptide blocking time constant (block τ) measured from N-type calcium channels expressed in HEK293 cells. The bar graph shows mean $\pm$ standard deviation for data from 4 to 7 cells (indicated within each bar). The lower case letters indicate significant differences ($P < 0.05$, analysis of variance with Tukey HSD post hoc analysis).

of mass 1793 Da & 1295/1313 Da, which corresponds to the fragment AKSOGSSUSOTSYNCAR and SUNOYTKRCY[#] (Note: cysteine and selenocysteine are bonded), further confirming the presence of Sec8–Cys15 and Sec19–Cys26 connectivity in non-natively folded dd-Sec-GVIA. Note also that the natively folded and non-natively folded dd-Sec-GVIA were sensitive to dissolved oxygen species, and tended to redistribute among all three the possible disulfide/selenosulfide isomers of dd-Sec-GVIA over time.

Figure 4b shows accumulation of the natively folded and biologically active conformer of GVIA, dd-GVIA, Sec-GVIA, and dd-Sec-GVIA at steady state in the presence of a GSH/GSSG mixture. It is evident from Figure 4b that incorporation of selenocystine at native disulfide connectivity in GVIA (Sec-GVIA) and dd-GVIA (dd-Sec-GVIA) has a profound influence on improving the yield of natively folded peptide. The net increase in yield of natively folded peptide by incorporation of selenocystine in GVIA is 20%, and that of dd-GVIA is 29%. The greater redox stability and more rapid formation of diselenide bonds, as compared to disulfide bonds, significantly improves folding yields, and preferential formation of diselenides over selenosulfides further simplifies disulfide mapping in the folded peptide. Impairment in folding pattern of GVIA and Sec-GVIA upon removing the functionally noncritical disulfide confirms the role that the Cys1–Cys16 disulfide bond plays in the folding of ω-conotoxin GVIA.

Biological Activities of dd-GVIA and dd-Sec-GVIA

We previously demonstrated that GVIA and Sec-GVIA produced an equivalent block of N-type calcium currents following a 16 minute application.[38] The disulfide deficient peptides, dd-GVIA and dd-Sec-GVIA, also block N-type calcium currents expressed in HEK293 cells, but with a slower time course (Figure 5a). The blocking time course for each peptide was well described by a single exponential function to yield the time constant, τ. The τ values for dd-GVIA and dd-Sec-GVIA were significantly different from that of GVIA and Sec-GVIA (Figure 5b). As we had previously suggested,[38] the substitution of Cys8 and Cys19 by selenocysteine had no significant effect on the kinetics of N-current block since τ of GVIA and Sec-GVIA were not statistically different. The selenocysteine change also did not impact the block by the disulfide deficient peptide, since the block τ was statistically equivalent between dd-GVIA and dd-Sec-GVIA. We conclude that the removal of the first disulfide pair does not prevent block of N-type channels, but does slow that block by roughly 50%.

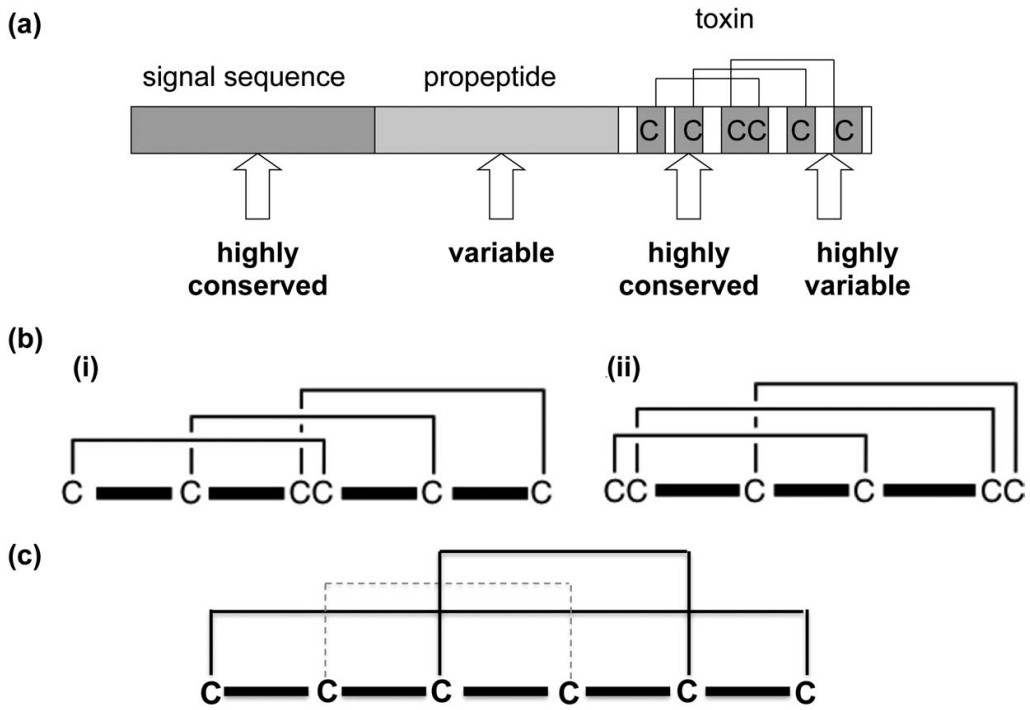


FIGURE 6 Biochemical synthesis of cysteine rich conotoxins and disulfide scaffolds found among conotoxins. a: The precursor prepropeptide with highlighted conserved and variable regions. b: Two disulfide scaffolds found in conotoxins (i) shows the disulfide scaffold for ω, κ, δ, and μO pharmacological families and (ii) shows the disulfide scaffold for the μ and Ψ pharmacological families. The scaffold shown in (i) can act as a calcium channel blocker, potassium channel blocker, delayer of sodium channel inactivation, or a sodium channel blocker. Conversely, the scaffold shown in (ii) can act as a sodium channel blocker or a noncompetitive inhibitor of nicotinic acetylcholine receptors. c: Conkunitizin scaffold with a naturally deleted disulfide bridge marked by a dashed line. Conserved Cys–Cys motifs in Kunitz domain were replaced by Gly-Gln motif in conkunitizin.

To investigate the role of removal of the functionally noncritical disulfide in the in vivo activity of ω-conotoxin GVIA, dd-GVIA, Sec-GVIA, dd-Sec-GVIA and the non-natively folded dd-GVIA and dd-Sec-GVIA were injected intracranially into mice. GVIA and Sec-GVIA were previously shown to induce shaking syndrome in mice upon intracranial injection, which is characterized by a persistent body tremor, and this behavior was maintained for a long time in a dose-dependent manner.[22,38] Both dd-GVIA and dd-Sec-GVIA also exhibit shaking syndrome in mice, confirming that the removal of the functionally noncritical disulfide did not affect the biological activity of ω-conotoxin GVIA. However, the shaking syndrome induced by dd-GVIA and dd-Sec-GVIA in mice was not as long-lasting as either GVIA or Sec-GVIA. The non-natively folded dd-GVIA and dd-Sec-GVIA did not exhibit shaking syndrome; rather, the injected mice were passive, occasionally shaking their bodies. The distinct behavioral symptoms elicited by natively folded and non-natively folded disulfide-depleted peptide analogues demonstrate that the topology of disulfide connectivity is critical to the biological activity. Similar behavioral symptoms were elicited by natively folded GVIA and dd-GVIA and their corresponding selenocysteine analogues, confirming that the removal of the disulfide between Cys1 and Cys16 did not substantially affect the biologically active conformation of ω-conotoxin GVIA.

DISCUSSION

Although cone snails have a very efficient mechanism to incapacitate their prey, they are not as mobile as other species,

and therefore rely on high-efficiency hunting mechanisms. Consequently, the implications of a failed envenomation event are dramatically larger for snails than they are for other venomous animals that are more able to pursue their prey as it attempts to escape. Therefore, it is likely that the evolutionary pressures on cone snails are stronger than those on other venomous animals. For example, if an injection of venom takes 2 min to paralyze the prey, the cone snail would remain hungry, where a snake or scorpion would likely not. This is probably the explanation for the atypically high evolutionary rate of venom peptides in *Conus*.[49] However, as discussed below, conotoxins offer a unique opportunity to study evolutionary and molecular mechanisms by which genetic information is translated into folding and activity of disulfide-rich polypeptides.

As illustrated in Figure 6a, the high rate of evolution of *Conus* venom toxins is in striking contrast with discrete conservation of various portions of a toxin precursor gene.[49] From a mechanistic point of view, even more puzzling is the fact that the cysteine scaffold remains highly conserved at the level of Cys codon usage (TGC or TGT), while adjacent codons belong to hypervariable sequences of intercysteine loops.[50–52] Despite this extreme conservation of the cysteine scaffold, this does not imply a conservation of pharmacological properties.[10] For example, Figure 6b shows two common disulfide scaffolds found in conotoxins blocking sodium channels,[53,54] but one scaffold (i) also blocks calcium channels and potassium channels (ω- and κ-conotoxins, respectively), as well as delaying the inactivation of sodium channels (δ-conotoxins), whereas the other scaffold (ii) is also capable of functioning as a noncompetitive inhibitor of nicotinic acetylcholine receptors.[55] Noteworthy, despite an evolutionary conservation of Cys positions and codons, both the scaffolds provide examples of conotoxins containing functionally noncritical disulfide bridges.

However, despite the highly conserved disulfide scaffolds in *Conus*, there are several instances of evolutionary divergence of disulfide scaffolds in these marine snails. Although it is often difficult to reliably determine which is the "ancestral" fold, the conservation of cysteine codons and propeptide region strongly suggests evolutionary linkage between the α-conotoxins (bearing two disulfide bridges) and the αA-conotoxins (bearing three disulfide bridges).[50] Additionally, the examples of Kunitz domains in conotoxins indicate that the evolution of disulfide bridges does not bear an inherent directionality (gain/loss of bridges). The Kunitz domain is a very ancient fold, being present in both plants[56] and animals,[57] with bovine pancreatic trypsin inhibitor (BPTI) being the prototypical example. From this ancient fold/cysteine scaffold, cone snails have developed a class of conotoxins, referred to as conkunitzins. The conkunitzin family has sequences that show a deletion of a disulfide bridge (illustrated in Figure 6c) that is highly conserved in non-Conus Kunitz domains,[58] as well as in other conkunitzins that contain two Kunitz domains in tandem, with addition of a fourth disulfide bridge on the N-terminal Kunitz scaffold.[59]

One outstanding question that remains is that the mechanism by which the cysteine scaffold is so highly conserved remains unknown. The extent of conservation decreases from the signal peptide to the mature toxin region (Figure 6a). However, the placement of cysteine, and even the cysteine codons within the mature toxin region are much more highly conserved than the remainder of the toxin region. Although it has been proposed that this is due to a polymerase that preferentially preserves cysteine codons,[60] a DNA polymerase capable of recognizing frame shifts has yet to be discovered. Consequently, it is possible that the cysteine scaffold is preserved by another mechanism, likely relying on evolutionary pressures. A given disulfide bond may be either critical to oxidative folding or to biological activity (or both), or may even be expendable. Thus, the extent to which the sequence encodes folding information to direct the folding to the native disulfide connectivity is critical to the ability to add and/or remove disulfide bridges to an existing scaffold. Herein, we explore the ability to delete a native disulfide bridge from ω-conotoxin GVIA, and consider the evolutionary pressures present to preserve a given disulfide scaffold.

The potential to dissect the roles of individual disulfide bridges in a short, disulfide-rich peptide allows for a much more thorough analysis of the evolutionary pressures present to maintain each bridge. In Figure 4, it is evident that the first disulfide bridge is critical to the efficient folding of GVIA, while Figure 5 shows that the same disulfide has a minimal effect on the biological activity of GVIA. Other work on deleting the second and third disulfide bridges demonstrated significant loss of the activity.[32] Using ω-conotoxin GVIA as a model for ICK peptides, it is evident that each of the disulfides serves a critical purpose in generating the final, biologically active compound. Consequently, the evolutionary pressure to efficiently produce active peptide has caused the retention of the disulfide bridges in almost all ICK peptides, with the only potential exception being U1-liotoxin-Lw1a, from the scorpion, *Liocheles waigiensis*,[61] which may also be a precursor that never acquired a third disulfide bridge.

The authors thank Drs. Robert Schackmann and Scott Endicott from the DNA/Peptide Synthesis Core Facility at the University of Utah for the synthesis of peptides as well as Pranav Mathur for critical reading of the manuscript and his helpful suggestions.

REFERENCES

1. Ryan, C. A. Hand Book of Biologically Active Peptides; Academic Press, Elsevier Inc.: CA, 2006.
2. Lehrer, R. I.; Lichtenstein, A. K.; Ganz, T. Annu Rev Immunol 1993, 11, 105–128.
3. Rodríguez de la Vega, R. C.; Schwartz, E. F.; Possani, L. D. Toxicon 2010, 56, 1155–1161.
4. Escoubas, P.; Sollod, B. L.; King, G. F. Toxicon 2006, 47, 650–663.
5. Gray, W. R.; Olivera, B. M.; Cruz, L. J. Annu Rev Biochem 1988, 57, 665–700.
6. Mouhat, S.; Jouirou, B.; Mosbah, A.; De Waard, M.; Sabatier, J. M. Biochem J 2004, 378, 717–726.
7. Olivera, B. M.; Rivier, J.; Clark, C.; Ramilo, C. A.; Corpuz, G. P.; Abogadie, F. C.; Mena, E. E.; Woodward, S. R.; Hillyard, D. R.; Cruz, L. J. Science 1990, 249, 257-263.
8. Olivera, B. M. Mol Biol Cell 1997, 8, 2101–2109.
9. Olivera, B. M.; Teichert, R. W. Mol Interv 2007, 7, 251–260.
10. Terlau, H.; Olivera, B. M. Physiol Rev 2004, 84, 41–68.
11. Bayrhuber, M.; Vijayan, V.; Ferber, M.; Graf, R.; Korukottu, J.; Imperial, J.; Garrett, J. E.; Olivera, B. M.; Terlau, H.; Zweckstetter, M.; Becker, S. J Biol Chem 2005, 280, 23766–23770.
12. Han, T. S.; Zhang, M. M.; Walewska, A.; Gruszczynski, P.; Robertson, C. R.; Cheatham, T. E., 3rd.; Yoshikami, D.; Olivera, B. M.; Bulaj, G. ChemMedChem 2009, 4, 406–414.
13. Price-Carter, M.; Hull, M. S.; Goldenberg, D. P. Biochemistry 1998, 37, 9851–9861.
14. Lamthanh, H.; Jegou-Matheron, C.; Servent, D.; Ménez, A.; Lancelin J. M. FEBS Lett 1999, 454, 293–298.
15. Kaerner, A.; Rabenstein, D. L. Biochemistry 1999, 38, 5459–5470.
16. Price-Carter, M.; Bulaj, G.; Goldenberg, D. P. Biochemistry 2002, 41, 3507–3519.
17. Rees, D. C.; Lipscomb, W. N. J Mol Biol 1982, 160, 475–498.
18. Pallaghy, P. K.; Nielsen, K. J.; Craik, D. J.; Norton, R. S. Protein Sci 1994, 3, 1833–1839.
19. Gelly, J. C.; Gracy, J.; Kaas, Q.; Le-Nguyen, D.; Heitz, A.; Chiche, L. Nucleic Acids Res 2004, 32, 156–159.
20. Gracy, J.; Le-Nguyen, D.; Gelly, J. C.; Kaas, Q.; Heitz, A.; Chiche, L. Nucleic Acids Res 2008, 36, 314–319.
21. Norton, R. S.; Pallaghy, P. K. Toxicon 1998, 36, 1573–1583.
22. Olivera, B. M.; McIntosh, J. M.; Cruz, L. J.; Luque, F. A.; Gray, W. R. Biochemistry 1984, 23, 5087–5090.
23. Shon, K. J.; Stocker, M.; Terlau, H.; Stühmer, W.; Jacobsen, R.; Walker, C.; Grilley, M.; Watkins, M.; Hillyard, D. R.; Gray, W. R.; Olivera, B. M. J Biol Chem 1998, 273, 33–38.
24. Shon, K. J.; Hasson, A.; Spira, M. E.; Cruz, L. J.; Gray, W. R.; Olivera, B. M. Biochemistry 1994, 33, 11420–11425.
25. McIntosh, J. M.; Hasson, A.; Spira, M. E.; Gray, W. R.; Li, W.; Marsh, M.; Hillyard, D. R.; Olivera, B. M. J Biol Chem 1995, 270, 16796–16802.
26. Fainzilber, M.; van der Schors, R.; Lodder, J. C.; Li, K. W.; Geraerts, W. P.; Kits, K. S. Biochemistry 1995, 34, 5364–5371
27. Olivera, B. M. In Drugs from the Sea; Fusetani N., Ed. Karger: Basel, Switzerland, 2000.
28. Miljanich, G. P. Curr Med Chem 2004, 11, 3029–3040.
29. Horne, A. L.; Kemp, J. A. Br J Pharmacol 1991, 103, 1733–1739.
30. Takahashi, T.; Momiyama, A. Nature 1993, 366, 156–158.
31. Sabo, T.; Gilon, C.; Shafferman, A.; Elhanaty, E. In 12th American Peptide Symposium; Smith J.A., Ed.; 1992, 1559–1560.
32. Flinn, J. P.; Pallaghy, P. K.; Lew, M. J.; Murphy, R.; Angus, J. A.; Norton, R. S. Biochim Biophys Acta 1999, 1434, 177–190.
33. Fiori, S.; Pegoraro, S.; Rudolph-Böhner, S.; Cramer, J.; Moroder, L. Biopolymers 2000, 53, 550–564.
34. Pegoraro, S.; Fiori, S.; Cramer, J.; Rudolph-Böhner, S.; Moroder, L. Protein Sci 1999, 8, 1605–1613.
35. Pegoraro, S.; Fiori, S.; Rudolph-Böhner, S.; Watanabe, T. X.; Moroder, L. J Mol Biol 1998, 284, 779–792.
36. Armishaw, C. J.; Daly, N. L.; Nevin, S. T.; Adams, D. J.; Craik, D. J.; Alewood, P. F. J Biol Chem 2006, 281, 14136–14143.
37. Walewska, A.; Zhang, M. M.; Skalicky, J. J.; Yoshikami, D.; Olivera, B. M.; Bulaj, G. Angew Chem Int Ed Engl 2009, 48, 2221–2224.
38. Gowd, K. H.; Yarotskyy, V.; Elmslie, K. S.; Skalicky, J. J.; Olivera, B. M.; Bulaj G. Biochemistry 2010, 49, 2741–2752.
39. Han, T. S.; Zhang, M. M.; Gowd, K. H.; Walewska, A.; Yoshikami, D.; Olivera, B. M.; Bulaj, G. ACS Med Chem Letter 2010, 4, 140–144.
40. Steiner, A. M.; Bulaj, G. J Pept Sci 2011, 17, 1–7.
41. Walewska, A.; Jaskiewicz, A.; Bulaj, G.; Rolka K. Chem Biol Drug Des 2011, 77, 93–97.
42. Muttenthaler, M.; Nevin, S. T.; Grishin, A. A.; Ngo, S. T.; Choy, P. T.; Daly, N. L.; Hu, S. H.; Armishaw, C. J.; Wang, C. I.; Lewis, R. J.; Martin, J. L.; Noakes, P. G.; Craik, D. J.; Adams, D. J.; Alewood, P. F. J Am Chem Soc 2010, 132, 3514–3522.
43. de Araujo, A. D.; Callaghan, B.; Nevin, S. T.; Daly, N. L.; Craik, D. J.; Moretta, M.; Hopping, G.; Christie, M. J.; Adams, D. J.; Alewood, P. F. Angew Chem Int Ed Engl 2011, 50, 6527–6529.
44. Yarotskyy, V.; Elmslie, K. S. J Neurophysiol 2009, 101, 332–340.
45. Yarotskyy, V.; Elmslie, K. S. Br J Pharmacol 2007, 152, 386–395.
46. Flinn, J. P.; Pallaghy, P. K.; Lew, M. J.; Murphy, R.; Angus, J. A.; Norton, R. S. Eur J Biochem 1999, 262, 447–455.
47. Price-Carter, M.; Gray, W. R.; Goldenberg, D. P. Biochemistry 1996, 35, 15537–15546.
48. Price-Carter, M.; Gray, W. R.; Goldenberg, D. P. Biochemistry 1996, 35, 15547–15557.
49. Duda, T. F.; Palumbi, S. R. Proc Natl Acad Sci USA 1999, 96, 6820–6823.
50. Santos, A. D.; McIntosh, J. M.; Hillyard, D. R.; Cruz, L. J.; Olivera, B. M. J Biol Chem 2004, 279, 17596–17606.
51. Wang, Q.; Jiang, H.; Han, Y. H.; Yuan, D. D.; Chi, C. W. Toxicon 2008, 51, 813–822.
52. Zhangsun, D.; Luo, S.; Wu, Y.; Zhu, X.; Hu, Y.; Xie, L. Chem Biol Drug Des 2006, 68, 256–265.
53. Eckberg, J.; Jayamanne, A.; Vaughan, C. W.; Aslan, S.; Thomas, L.; Mould, J.; Drinkwater, R.; Baker, M. D.; Abrahamsen, B.; Wood, J. N.; Adams, D. J.; Christie, M. J.; Lewis, R. J. Proc Natl Acad Sci USA 2006, 103, 17030–17035.
54. Zhang, M. M.; Green, B. R.; Catlin, P.; Fiedler, B.; Azam, L.; Chadwick, A.; Terlau, H.; McArthur, J. R.; French, R. J.; Gulyas, J.; Rivier, J. E.; Smith, B. J.; Norton, R. S.; Olivera, B. M.; Yoshikami, D.; Bulaj, G. J Biol Chem 2007, 282, 30699–30706.
55. Olivera, B. M.; Cruz, L. J. Toxicon 2001, 39, 7–14.

56. Troncoso, M. F.; Biron, V. A.; Longhi, S. A.; Retegui, L. A.; Wolfenstein-Todel, C. Int Immunopharmacol 2007, 7, 625–636.
57. Ascenzi, P.; Bocedi, A.; Bolognesi, M.; Spallarossa, A.; Coletta, M.; De Cristofaro, R.; Menegatti, E. Curr Protein Pept Sci 2003, 4, 231–251.
58. Dy, C. Y.; Buczek, P.; Imperial, J. S.; Bulaj, G.; Horvath, M. P. Acta Crystallogr D Biol Crystallogr 2006, 62, 980–990.
59. Olivera, B.; Bulaj, G.; Garrett, J.; Terlau, H.; Imperial, J. Editora UFMG 2009, 25–48.
60. Conticello, S. G.; Gilad, Y.; Avidan, N.; Ben-Asher, E.; Levy, Z.; Fainzilber, M. Mol Biol Evol 2001, 18, 120–131.
61. Smith, J. J.; Hill, J. M.; Little, M. J.; Nicholson, G. M.; King, G. F.; Alewood, P. F. Proc Natl Acad Sci USA 2011, 108, 10478–10483.