CONOPEPTIDOMICS OF CONUS REGIUS

ALDO FRANCO

Conopeptidomics of *Conus regius*

By Aldo Franco

A Dissertation Submitted to the Faculty of

The Charles E. Schmidt College of Science

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Florida Atlantic University

Boca Raton, Fl

May 2006

Conopeptidomics of *Conus regius*

By Aldo Franco

This dissertation was prepared under the direction of the candidate's thesis advisor, Dr. Frank Mari, Department of Chemistry and Biochemistry and has been approved by the members of his supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

Dr. Frank Mari
Dissertation Advisor

Dr. Gregg B. Fields

Dr. Gary W. Perry

Dr. Victor Asirvatham

Chairman, Department of Chemistry and Biochemistry

Dean, Charles E. Schmidt College of Science

Dean, Graduate Studies and Programs

April 28, 2006
Date

ACKNOWLEDGEMENTS

Para mi familia con todo mi cariño y dedicacion

I would like to thank Dr. Frank Mari for his guidance, time and for giving me the opportunity to work with him.

I would also like to thank Dr. Gregg Fields, Dr. Gary Perry and Dr. Victor Asirvatham for their time, help, advice, and efforts as members of my dissertation committee.

Furthermore, I would like to thank Simon Bulley for all the help with bioassays, Herminsul Cano and Carolina Moller for their unconditional support, Sanaz for helping me with the pictures and to Jose and the rest of the research group for always being there when I needed you.

Most importantly I would like to thank Marcela, my wife and best friend; Papa Cesar, my father and mentor; and Mama Blanca, my mother and unconditional source of love for giving me the strength and determination that allowed me to finish this work. Thank you for always being there for me and helping me believe that I can accomplish my goals.

ABSTRACT

Author: Aldo Franco

Title: Conopeptidomics of *Conus regius*

Institution: Florida Atlantic University

Dissertation Advisor: Dr. Frank Mari

Degree: Doctor of Philosophy

Year: 2006

The main objective of this dissertation is the isolation and characterization of novel neuroactive peptides from *Conus regius*. The conopeptides targeted in this work have a MW of 3500 Da or less, in the hopes that they can become viable drug candidates. A total of 30 sequences were isolated and characterized from the venom of *Conus regius*, giving us a partial library of the conopeptides found in this species. Techniques such as size exclusion chromatography, reversed phase chromatography, mass spectrometry, nano-nuclear magnetic resonance, chemical modifications of peptides, peptide sequencing through Edman degradation and in some instances bioassays were used together in an effort to perform "conopeptidomics" of *Conus regius*. The first chapter deals with *Conus regius* M-superfamily conopeptides. The second chapter is about the A-superfamily conopeptides found in *Conus regius*. The third chapter deals with *Conus regius* P-superfamily conopeptides. Finally the fourth

chapter encompasses the T-superfamily conopeptides and all other small and linear peptides found in *Conus regius* that do not have a classification. This work is the first example reported, for any cone snail species, where most of the components of the venom have been sequenced directly for a single cone snail species. This work shows that a more realistic library of conopeptides can be obtained by direct analysis of the venom as opposed to cDNA libraries, which while useful; it does not reflect the post-translational modifications commonly found in conopeptides.

TABLE OF CONTENTS

LIST OF FIGURES

Page

Chapter II

Chapter III

Chapter IV

LIST OF TABLES

INTRODUCTION

While it is not difficult to overlook the deadly potential of the venom of some animals or insects, it is difficult to imagine the possible pharmaceutical potential of the venom of these organisms. This is the case of the beautiful cone snails, highly regarded as a shell collector's item, now also a valuable source of promising pharmaceutical compounds from the sea. Cone snails belong to the genus *Conus* and are venomous marine gastropods that utilize modified peptide toxins in an effective biochemical campaign to capture prey, deter competitors, defend against predators, and possibly in other biological functions not yet understood [1]. There are over 1000 cone snail species distributed mainly over tropical and subtropical regions of the Atlantic, Pacific and Indian oceans. These modified peptides (conopeptides), typically 6-30 amino acids in length, make up the venom of cone snails and are fast acting neuroactive compounds that immobilize the victim by affecting its neurophysiology. Once the cone snail injects its venom into the prey it is rapidly subdued due to the high-affinity binding of the venom's protein and peptide toxins to voltage and ligand gated ion channels essential for the proper function of the prey's nervous and muscular systems [2]. Since the first conopeptides were isolated a few decades ago, extensive systematic investigation has been conducted on these promising ion channel ligands. Originally conopeptides were used as tools for neuroscience research;

however, the ability of some conotoxins to selectively bind to some types of receptors and ion channels makes them attractive pharmacological candidates [3].

Cone snails can be broadly characterized by their choice of prey. Molluscivorous cone snails prey on other mollusks, vermivorous cone snails feed on polychaete worms, and piscivorous cone snails prey upon small fish. This classification based on prey preferences is not always followed rigidly by the snails. It has been observed in our research laboratory that in some instances piscivorous snails like *Conus purpurescens* can target other preys such as earth worms. The cone snails' lack of mobility and competitive marine environment demands them to have a fast acting venom to paralyze their agile prey quickly immobilizing their target [4]. Hence, the major components of the cone snail venom have been found to target key cell surface-signaling components of nervous systems such as ion channels and receptors.

ANATOMY OF THE VENOM APPARATUS

The venom apparatus in cone snails broadly consist of three parts: a venom bulb, a venom duct, and a radular sac containing harpoon-like radular teeth. Cone snails capture their prey by envenomation via a highly efficient venom apparatus and delivery system (Figure 1). The venom apparatus in all cone snails when dissected in greater detail consists of a venom bulb (vb), which is in charge of driving the venom throughout the venom apparatus; the venom duct (vd) is where the venom is synthesized by the epithelial cells lining the duct and where the venom is stored; a radula sac (rs) is where the harpoon-like teeth (h) are stored and the proboscis (pr), which is used to deliver the harpoon and venom to the prey. At the end of the secretory pathway, which is mainly unknown, conotoxins are packed into secretory granules that are secreted into the duct lumen [5]. Each harpoon is used only once and they act as disposable hypodermic needles to eject the venom [6]. The way the venom apparatus works, once the cone snail has targeted a prey, involves the following steps: a radular tooth is transported from the radular sac to the tip of the proboscis, then by extending the proboscis the preys' protective outer surface is penetrated by the radular tooth and then the cone snail venom is injected, paralyzing the prey and subduing it captive to the cone snail.

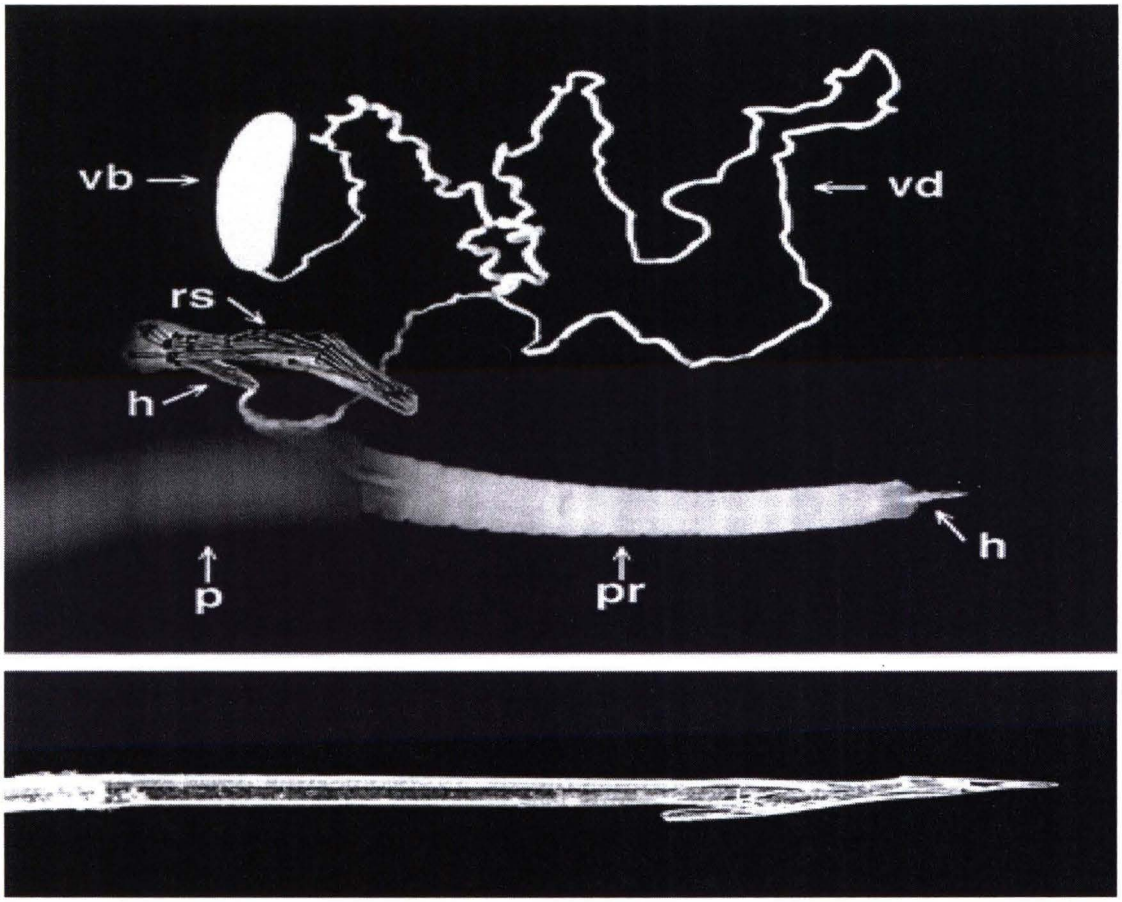


Figure 1. Structure of a typical Conus venom apparatus from *Conus purpurascens* [7]

Conus regius is a common Western Atlantic cone snail species that feeds on polychaete worms, more specifically the Caribbean fire worm *Hermodice carunculata* (Figure 2). These worms range in color from green, orange, red and could possibly explain the many variations of color of *Conus regius*. Two variants are present based on the color of the shell, *Conus regius regius* which has the typical red-burgundy marks, and *Conus regius citrinus* which lacks most of these red-burgundy marks and exhibits an orangey tint (Figure 3).

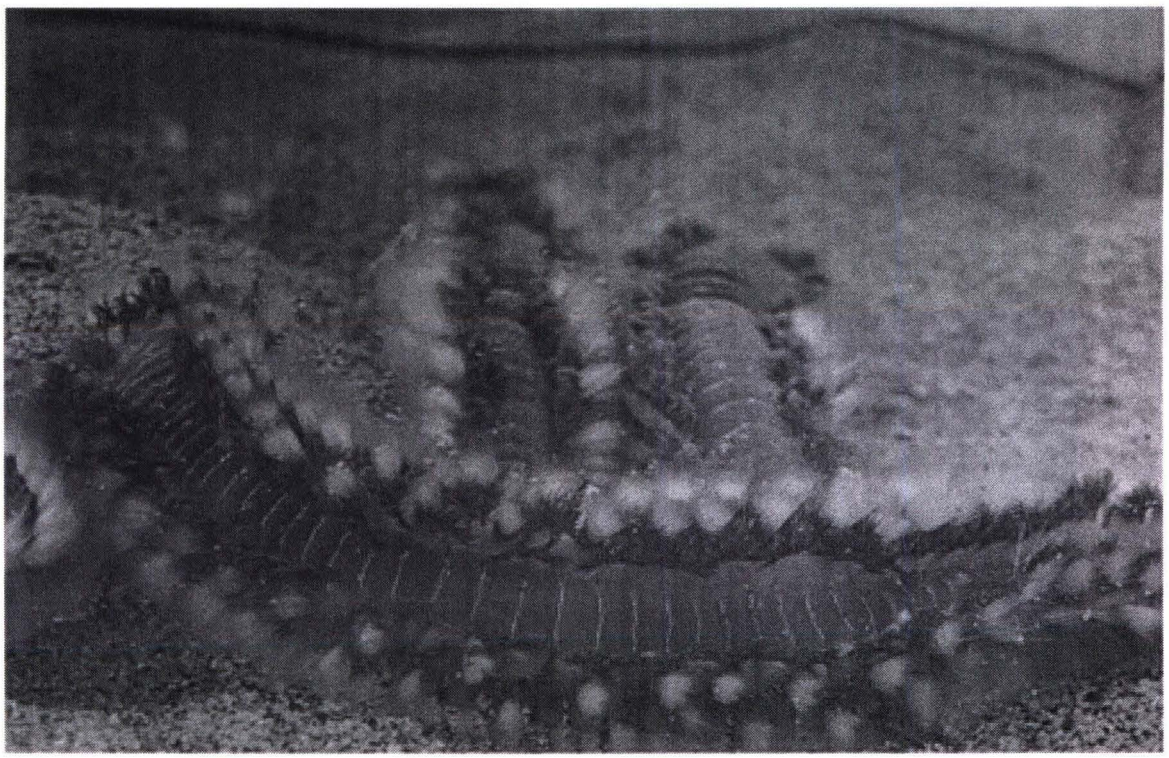


Figure 2. Different shades of the Caribbean fire worm *Hermodice carunculata.*

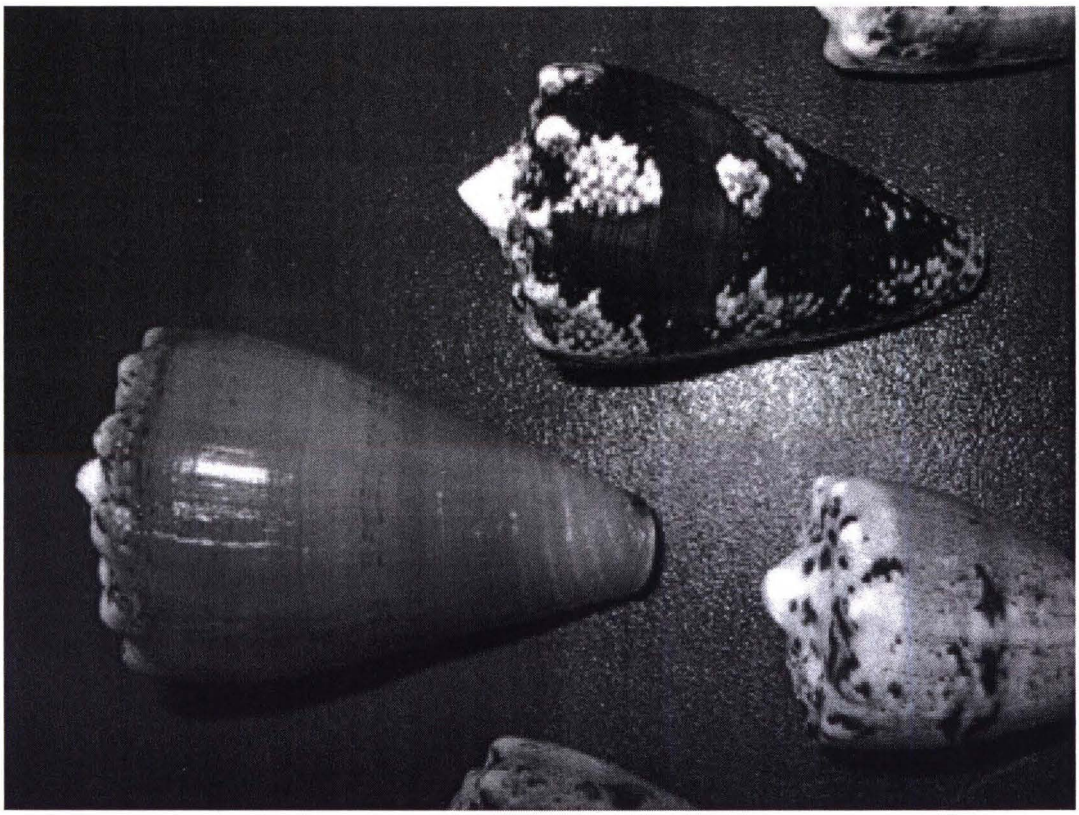


Figure 3. *Conus regius regius* and *Conus regius citrinus.*

Unpublished studies of the venom components of these two variants show that the components in their venom are the same thus exhibiting only a distinction based on the color of their shell. All *Conus* predatory gastropods hunt prey from one of five different phyla, and all of these preys have cholinergic synapses [8]. *Conus regius* is from the class gastropoda, order caenogastropoda, superfamily conacea, and family conidae. It is a nocturnal shallow water species found buried in gravel and under rocks. They thrive in open water regions and are commonly found from 5 to 40 feet in depth; however, it is not uncommon to find them beyond 40 feet. The cone resides in the entire Western Atlantic region from Florida, throughout the Caribbean and as far

south as the northern coast of Brazil. Figure 4 shows a photograph of a *Conus regius* shell.

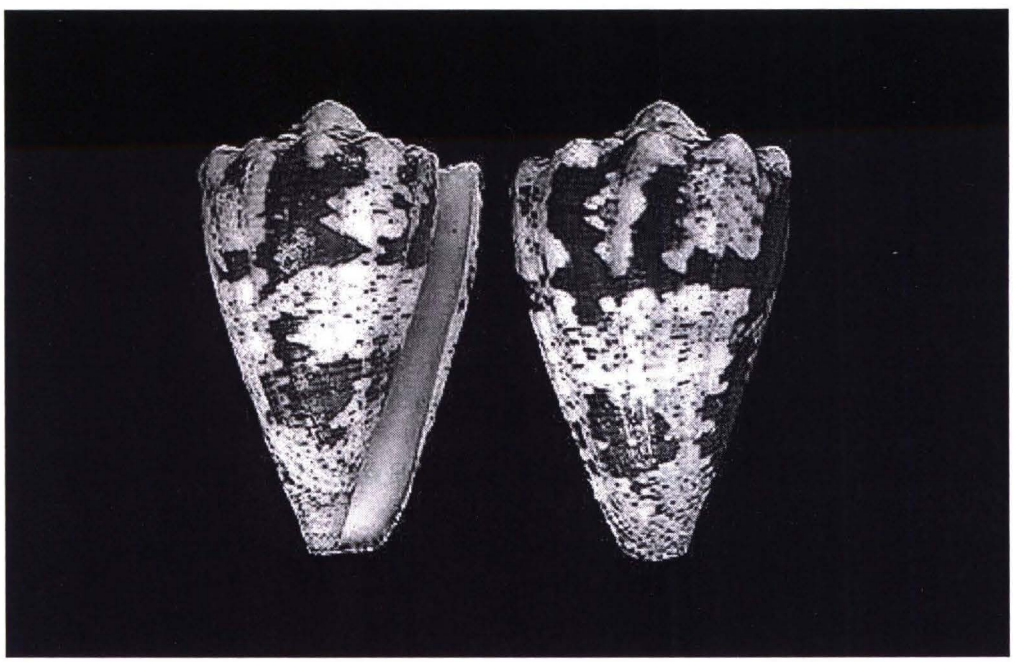


Figure 4. Photograph of *Conus regius*.

CONOPEPTIDES

The venom of animals belonging to the genus *Conus* have in the last two decades contributed to a new an exciting field, one in which physiologically active peptides are now considered as potential pharmaceuticals. This pharmacological potential was realized since the nineteen sixties when the potency of crude extracts of cone snail venom compared to that of scorpions and spiders [9]. Cone snail venom is composed of conopeptides which are the major paralytic components. *Conus* is one of the largest and most successful living genus of marine animals, comprised of about 1000 different species [3]. Each species of cone snails exhibits its own distinct repertoire of conopeptides, different from those of all other species. Similar conopeptides can be found among cone snail species that vary in two or more amino acids but none identical. The difference is thought to arise from divergent biotic interactions within the tropical marine habitats in which each cone snails thrive, that give rise to a corresponding divergence of cone snail venom. Each *Conus* species has a unique set of biotic interactions characteristic of that species that helps to rationalize why each species has a different complement of up to 300 conopeptides [10]. Although the specific details of the interactions between any *Conus* species with other animals is unknown, the fact that there are about 100 conopeptides per species (thus, over 100,000 unique active conotoxins) remains clear [4]. Although capable of producing an arsenal of peptides in excess of 300 or more, each species is believed to express only a subset of all of its peptides at one time, optimizing its' venom for that one particular prey. This diversity in the venom and their ability to produce more of one conopeptide, when compared to the crude venom of another member of the same

species, has given rise to much interest in their highly selective biological activity that show a range of physiological responses from shaking to depression and paralysis.

It has been shown that the active components of these *Conus* venoms contain various small structured peptides with many post translational modifications. These conopeptides are capable of each specifically targeting different ion channels and receptors. Conopeptides are broadly defined as all peptide components found in the venom. Conopeptides are initially biosynthesized as larger precursor peptides, usually 60-150 amino acids in length, consisting of the N-terminal signal sequence, an intervening propeptide region and a mature toxin section, which is later proteolytically cleaved [10]. These precursors may vary in sequence at the prepropeptide region portion; however, conopeptides that target the same receptor have in common the position of their disulfide bonds within the mature toxin portion around the C-terminus of the precursor protein. Members of a pharmacologically-related family exhibit homology in their precursor sequences; however, sequence divergence is observed for families of different pharmacology within the same genomic class [11].

The latest classification of conopeptides was reported by Terlau [4]. According to his scheme there are two major classes of conopeptides: non-disulfide rich and disulfide-rich. Disulfide-rich conopeptides contain two or more disulfide bonds and are termed conotoxins. Conantokins, Contulakins, Conorfamides and Conophans are linear peptides with no disulfide bridges. This work will show that in addition to these

families, there are other non-disulfide rich peptides present in the venom of cone snails that lack a classification. Conopressins have one disulfide bridge and have sequence homology to the vasopressin and oxytocin family of peptides hormones. The conotoxins are sub-divided into superfamilies in accordance genomic configuration. As mentioned above, the precursors of superfamily members share a highly conserved signal sequence in their precursors.

Many super families have already been identified and named, such as the O, M, A, S, T, P, and I superfamilies and many others remain to be described. Subclassifications within superfamilies are based on their pharmacology, and in some cases, the relative arrangement of cysteine residues and the number of residues between these cysteines (Figure 5).

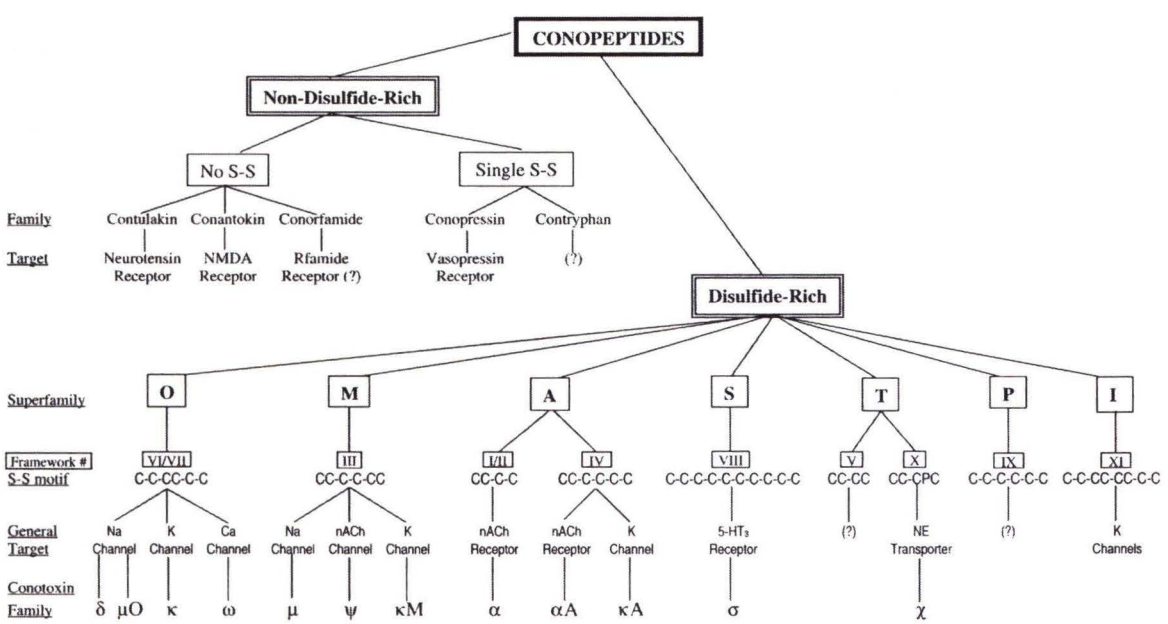


Figure 5. Classification of Conopeptides by Molecular Target and Loop Pattern [4].

Post-translational modifications are a common feature among conopeptides. Insights into the mechanism of post-translational modifications of conopeptides have been provided from studies of the γ-carboxylation of glutamate and other well-characterized post-translational modifications. Post-translational modifications enhance the molecular diversity of conotoxins. Among these we find O-glycosylation, bromination of tryptophan, γ-carboxylation of glutamate residues, hydroxylation of prolines, valines and lysines, or L- to D-epimerization that appears in many of the peptides [12]. Some of these modifications are quite common, while others are unusual, and some were first discovered in *Conus* peptides like D-γ-Hydroxyvaline [13].

NEUROPHYSIOLOGY AND NEUROPHARMACOLOGY OF VENOM

Venom peptide neurotoxins are found in many animals such as wasps, bees, scorpions, spiders, snakes and cone snails among others. These neurotoxins have become important investigative tools in neurophysiology. In cone snails, it is the diverse nature of these peptides which helps them exhibit a huge assortment of pharmacological activity, even within a given superfamily. The diversity is ascribed to differences in primary structure of the peptides, which give rise to variability in charge, structure and characteristics of each unique peptide. Conotoxins target a wide array of molecular targets. These include G-protein coupled receptors, neurotransmitter transporters, ligand–gated ion channels, and voltage–gated ion channel. The ligand–gated ion channels and voltage–gated ion channel are the best

understood as described in the review by Terlau [4]. The concept of Na^+ and K^+ channels as components of action potential generation has given way to an intricate molecular complexity characterized by the fact that the K^+ channel actually consists of over 80 genes that can arrange in a variety of combinations to yield an overwhelming diversity of isoforms of K^+ channels. The wide diversity in ion channel structure and function makes conopeptides important modulators of many biological activities.

In regards to the mode of action of conopeptides, one can only be amazed at their ability to impart the necessary physiological effect on their prey or predator. This is evident when we observe that the venom of cone snails is composed of hundreds of components, each acting synergistically with the other components in order to accomplish the same goal, which is immobilizing the targeted animal [4]. Terlau defines the mode of action of the venom components as a two step process. The first peptides to inflict their mode of action in the targeted animal act to immediately immobilize the victim by inhibiting voltage gated Na^+ channel inactivation and blocking K^+ channels. This type of activity leads to mass depolarization of axons in the vicinity of the site of venom injection resulting in a physiological state similar to that of electrocution. The second physiological effect is achieved by peptides that work to completely inhibit neuromuscular transmission. These peptides must be transported to sites, such as neuromuscular junctions, that are away from the site of venom injection, and are therefore required to reach those targets before imparting the desired physiological effect, making them slower acting. The combination of these

two steps have as a result the inhibition of presynaptic Ca^{2+} channels, postsynaptic nicotinic receptors, and Na^{+} channels involved in the muscle action potential [4]. Table 1 shows a listing of molecular targets and therapeutic potentials of different classes of peptides.

Class	**Target**	**Therapeutic Potential**
Contulakins - Linear	Neurotensin receptors	Neuropathic pain
Conatokins - Linear	NMDA receptors	Epilepsy, Parkinson's
Conopressins - CC	Vasopressin receptors	Regulates blood pressure
χ-Conotoxins - CC-CPC	Neuropathic pain	Neuropathic pain
ω-Conotoxins - C-C-CC-C-C	Ca channels	Analgesic, Stroke
κ-Conotoxins - C-C-CC-C-C	K channels	Arrhythmia, Hypertension
μ-Conotoxins – CC-C-C-CC	Skeletal muscle Na channels	Neuromuscular block
ψ-Conotoxins - CC-C-C-CC	Skeletal muscle nACh channels	Analgesic, Parkinson's, Hypertension
α-Conotoxins - CC-C-C	Skeletal muscle nACh receptor	Analgesic, Parkinson's, Hypertension

Table 1. Targets and therapeutic potential of different classes of conopeptides [14]

From the described actions of the different peptides in the venom in Table 1, one can clearly see that the physiological effect desired comes from a variety of peptides with different targets and activities, all acting in a synergistic fashion. This presumed strategic plan of attack by cone snails could very well explain the great number of components in the venom that has generated an enormous conopeptide library with an unlimited pharmacological potential. The first conotoxin to be used therapeutically is Ziconotide, the ω-conotoxin MVIIA, now known as Prialt™ from *Conus magus*. This conotoxin has obtained FDA-approval and is currently one of the strongest pain-killers on the market. It has been reported to be 100-1000 fold more potent than morphine as an analgesic but is not addictive [15]. Other therapeutic conopeptides being tested to date are listed in Table 2 [14].

Name of Toxin	**Class of Toxin**	**Species**	**Therapeutic Interest**	**Mode of Action**	**Company**	**Stage of Development**
MrIA/B	χ	*C. marmoreus*	Neuropathic pain	Targets noradrenaline transporter	Xenome	Preclinical
AM336	ω	*C. cactus*	Morphine resistant pain	Inhibit Ca channel	Amrad	Phase II
ACV1	α	*C. victoriae*	Neuropathic pain/ Nerve injury recovery	Nicotinic receptor agonist	Metabolic	Preclinical
Contulakin-G	Contulakin	*C. geographus*	Chronic pain	Targets neurotensin receptor	Cognetix	Phase II
Conatokin-G	Conatokin	*C. geographus*	Antiepileptic agent	NMDA receptor antagonist	Cognetix	Preclinical

Table 2. Potential therapeutic conopeptides [3, 14]

SEPARATION OF THE VENOM COMPONENTS

Cone snail venom can be obtained either by extraction from the venom ducts or by milking the venom out of the live snails. Extraction of the venom from the snail is accomplished by dissecting-out and homogenizing the venom ducts in an aqueous solution. The most widely used solution for extraction is 0.1% trifluoroacetic acid (TFA). However, solutions of 1 – 5% acetic acid and solutions of up to 60% acetonitrile (ACN) in 0.1% trifluoroacetic acid have also been reported. This aqueous mixture is then centrifuged at speeds up to 12000 rpm while kept at 4°C. The aqueous layer is then lyophilized and the crude venom is ready for chromatography. Milked venom is typically subjected to chromatography directly. Some publications report adding protease inhibitors like phenylmethylsulfonyl fluoride to the either the extracted or milked venom [16].

Whether dissected or milked venom is used, the numerous components of it have been typically separated by using different types of reversed phase HPLC columns. C18 columns, and to a lesser degree C8 columns, have been reported to be used for conopeptide separations, while C4 columns have been reported to be used for conoprotein work. Typical reversed phase separations are 45 – 100 minutes long with an increase of 1% B / min, using for solution A 0.1% TFA and for solution B 60 – 90% ACN in 0.085% TFA. It was these reversed phase chromatography separations that showed the complexity of cone snail venom. Great advances have been made in the production of materials that are able to separate complex mixtures of proteins or peptides. However, the great majority of early publications dealing with the

separation of compounds from cone snail venom showed reversed-phase chromatography as their first method of use and in some cases their only choice.

Researchers working with cone snail venom have for the longest time either overlooked or had limited success with size exclusion chromatography. It is finally now that publications are beginning to show consistently the use of size exclusion chromatography (SEC) as a valuable tool to first separate the venom [13]. Typical size exclusion materials used in the reported literature are Sephadex G-25 and G-50 [16]. The lengths of the columns vary from 85 to 100 cm and the inner diameters are in between 1.5 to 2 cm. The flow rate for these types of columns is usually set to 0.5ml / min. The most common reported buffers for SEC of snail venom are solutions of ammonium acetate (0.1 – 0.2 M) and 1% acetic acid. Aqueous solutions of ACN have also been reported to be used with SEC, showing less degree of success. Unfortunately, despite great efforts, the quality of the majority of these SEC separations is poor. These poor separations were mainly the result of two factors: (1) Size exclusion material not being able to separate peptides efficiently and (2) Poor choice of buffer solutions. Nowadays there are options that can be used for the separation of peptides using size exclusion materials. These new materials in combination with reversed phase methods can efficiently separate peptides from complex mixtures such as cone snail venom. Six years ago our group developed a step-wise separation method that combined both SEC and RP chromatography. This method effectively separates the different components found in cone snail venom with a degree of resolution not found in any publication to date. From this separation we obtain typically nanomoles of compound that allow us to perform a wide array of

experiments from 1D and 2D NMR to sequencing, MS and MS/MS experiments and bioassays of native peptides. This allows us to address important issues such as post-translational modifications (PTM), something that it cannot be addressed with sequences derived from cDNA libraries in combination with synthetic cone snail venom components. SEC in combination with RP chromatography is the most effective way of separating components from cone snail venom. The work performed on *Conus regius* follows this separation scheme.

NMR AS A TOOL IN CONOPEPTIDOMICS

Conopeptidomics is a term we use to describe the process of fully sequencing all peptides present in the venom of cone snails, whether the venom is dissected or milked. Traditionally, conopeptides sequences are obtained from an isolated native conopeptide template or cDNA libraries. Subsequently, these conopeptides are synthesized, allowing the use of these compounds for many types of assays and experiments. Then usually the last step is to obtain the three-dimensional structure of the conopeptide in question by using the synthetic toxin. However, in the case of cDNA sequences, it raises the question of whether all post-translational modifications are been addressed properly. In case were sequences are obtained from peptide sequencing; co-elution of native and synthetic peptides is generally a good indication of both peptides being the same; however, there is no in identity, as differences in the sequence and PTMs can easily be overlooked by co-eluting different concentrations of peptides or too much peptide for that matter.

Working with the native peptide does address this matter. Unlike traditional conotoxin research, we first screen every possible conopeptide fraction by NMR spectroscopy. NMR spectroscopy is a non-destructive technique that gives not only invaluable structural and sequential information, but also the concentration of the conopeptide. Typical concentrations of conopeptide structural studies performed by NMR spectroscopy range from 1-16 mM, 5 mM being the most common, and typical volumes range from 600 µl down to 200 µl if a Shigemi tube is used. Studies using volumes lower than 40 µl can be performed by using a nanoprobe from Varian.

However, sample preparation for the nanoprobe is difficult to set up and can not be used for routine screening of samples. While working on the conopeptidomics of *Conus regius* we have adopted and perfected a methodology that not only minimizes the chances of overlooking post-translational modifications, but also helps us routinely screen native peptides by NMR spectroscopy in minimal concentrations. Volumes of approximately 40 µl of native peptide solution can be routinely screened by NMR spectroscopy using a 1.7 mm NMR tube in combination with a 3 mm NMR probe. We try to obtain as much NMR data possible for a fraction before submitting it to more destructive techniques. ^{1}H 1D, 2D TOCSY, 2D NOESY, 2D-HSQC and 2D HMBC are experiments that can be performed given the right conditions. Concentration is the key factor for obtaining NMR spectra. The majority of the time peptide quantities are too low (< 1 nanomole) allowing us to obtain a ^{1}H 1D and perhaps a 2D TOCSY. The 2D TOCSY experiment allows us to double check discrepancies in the sequence obtained from Edman degradation that may arise from PTMs. With this technique we have recorded 2D TOCSY spectra from a 1 nanomole native conopeptide sample in 37 µl of solution. For 1D experiments we have been able to obtain spectra after 1000 scans of 100 picomoles of native peptide. We use TSP as an internal standard to obtain concentrations. However, problems with this compound have been reported; therefore, we use a standardized peptide solution to compare signals and verify concentrations based on the methyl groups from both spectra. NMR spectroscopy becomes an invaluable tool for conopeptidomics providing us with information that could have not otherwise been obtained from such small quantities of native conopeptides. NMR spectroscopy has enabled us to perform

conopeptidomics on *Conus regius* by providing us a tool to evaluate the concentrations, purity and sequence information of the peptides. This approach has allowed us to characterized 30 novel native conopeptides within existing families and new families of conopeptides.

3D STRUCTURE OF CONOPEPTIDES

NMR spectroscopy has been the primary tool to determine the three-dimensional structure of conopeptides as it accounts for 42 three-dimensional structures. Although mainly used in protein work, x-ray crystallography has been used to determine the three-dimensional structures of 3 conopeptides [17-19]. There are a total of 47 conotoxin 3D-structures reported to date and all but one of the three-dimensional structures were obtained by using the synthetic homologue of the conopeptide. 17 of the structures available today belong to the A-superfamily [17-31], 15 belong to the O-superfamily [32-43], 7 belong to the M-superfamily [44-49], 2 conantokins [50], 4 contryphans [51-54] and 2 that do not fit the current classification of conopeptides (TVIIA and GS) [55, 56].

A-Superfamily

From the 17 structures that belong to the A superfamily 15 structures are α-conotoxins and 2 structures are αA-conotoxins. Recently, a minireview series composed of 4 papers [57] thoroughly discussed structure and function of the neuronal active α-conotoxins. The structure of α-conotoxins targeting the muscle specific acetylcholine receptor (nAChR) like EI [58] were not discussed. It has been established that α-conotoxins have a conserved framework composed of a CC-Loop1-C-Loop2-C arrangement that make up the disulfide bridges and also conserved amino acids that may account for their affinity to the nAChR. These features give rise to a secondary structure composed of an α-helix and a loop with various degrees of flexibility common to all α-conotoxins. α-conotoxins share a similar three-

dimensional structure suggesting that it is those amino acids that are not conserved determined the selectivity to target different nAChRs. The 2 αA-conotoxin structures are EIVA and PIVA [23, 25]. These conotoxins have some degree of sequence homology. Their most striking feature is the good backbone superposition between the residues 17-20 of EIVA and residues 17-19 of PIVA, this area being called the handle. Beyond the "handle area" the secondary structure varies and no other common features can be observed. With only 2 structures of this type and the lack of superposition of disulfide bonds in the paper of EIVA with PIVA is difficult to determine what features distinctively characterize αA-conotoxins. As of now "the handle" is the only conserved secondary structure common to all αA-conotoxins.

O-Superfamily

There are 15 published structures from this superfamily. 10 structures are ω-conotoxins [32, 34-37, 39, 59], 2 are δ-conotoxins [41, 60], another 2 are μO-conotoxins [40] and 1 κ-conotoxin [43]. Despite the low amino acid sequence homology among ω-conotoxins the overall fold and molecular shape is very similar among them. The distribution of charges on the surface of ω-conotoxins varies among them in part by their low sequence homology, accounting for the differences in binding affinities to Calcium channels. The spatial arrangement of the cysteine residues and their respective disulfide-bond pattern play a crucial role in keeping the overall three-dimensional structure similar among the ω-conotoxins. A triple-stranded antiparallel β-sheet is the most striking common feature found in the members of this

family, with minor differences in the length of the β-sheet and the type of some of the turns [39].

δ-conotoxins are represented by EVIA and TxVIA. They share the cystine knot motif commonly found in members of this family. The major structural difference among them is the length of the second loop, being that of EVIA larger by three residues than that of TxVIA. A small β-sheet composed of three short antiparallel strands is common among them, giving rise to a similar three-dimensional fold. Both structures of μO-conotoxins come from *Conus marmoreus* MrVIA and MrVIB [40]. Even though the first sequence of μO conotoxins was published about 10 years ago, structural information is limited for this conotoxins. The three-dimensional fold of μO-conotoxins is almost the same as that of ω and δ conotoxins despite little sequence homology. MrVIA and MrVIB sequence differ by only 3 residues. It is this small difference that accounts for MrVIA acting as a calcium channel agonist and MrVIB as an antagonist at low micromolar concentrations. These conotoxin three-dimensional structures consist of a β–sheet, a cysteine knot and four backbone loops. Loop 2 is highly disordered and as large as the δ-conotoxin EVIA. This loop is known to have key residues that facilitate binding to voltage-gated calcium channels. There is only one three-dimensional structure available for κ-conotoxins, that of PVIIA which blocks the potassium channel. The structure is composed of two parallel loops that are stabilized by a triple stranded antiparallel β-sheet. There are three disulfide bonds that give rise to a three-dimensional fold similar to other members of

the O superfamily, GVIA, MVIIA and MVIIC [32, 33]. This one is a great example as to how similar three-dimensional folds have different targets.

M-Superfamily

There are 7 three-dimensional structures from members of the M-superfamily. 3 μ-conotoxins [44, 45], 2 ψ-conotoxins [46, 61], 1 κM-conotoxin [48] and 1 mini-M conotoxin [49]. μ-Conotoxins inhibit the voltage-sensitive sodium channels. PIIIA, GIIIA and GIIIB are the only three-dimensional examples of this class. These three conotoxins have high sequence homology. While GIIIA and GIIIB are structurally very similar, PIIIA in spite of the high sequence homology posses some structural differences. In the last paper published of a member of this class, it is clearly explained that this difference arises from a Hyp at position 8 in PIIIA that adopts the *trans* conformation. Members of this class have as a structural motif a distorted helix, a small beta hairpin and a series of turns. There are 2 three-dimensional structures of ψ-conotoxins, PIIIE and PIIIF, both from *Conus purpurascens*. The overall three-dimensional fold is very similar to one another as expected from the high sequence homology and the identical disulfide bond arrangement. Their differences arise from the flexibility at both termini. The overall three-dimensional fold of ψ-conotoxins consists of a series of beta turns without helices or β-sheet regions. The disulfide bond arrangement is identical to that of μ-conotoxins giving rise to similar folds without having sequence homology. κM-conotoxins have one three-dimensional structure published, RIIIK from *Conus radiatus* [48]. Structural comparisons to same members of this class can not be done until a new structure of this class is published.

However, the authors of the structure of RIIIK point out that despite the different amino acid sequence and different pharmacological activity the structure of RIIIK is very similar to that of μ-conotoxin GIIIA in the C-terminal area [44]. Structural features for this class are a γ-turn, a distorted helix, and a 3_{10} helix, giving κM-conotoxins a three-dimensional structure in the form of a disk. The last class of conotoxin with a three-dimensional structure from this superfamily is the mini-M mr3a from *Conus marmoreus* [49]. The key feature for this class is that despite having the same cysteine arrangement as the μ and ψ-conotoxins the connectivity is different. The three-dimensional fold is one of a tight globular structure that is composed of three turns.

Conantokins

There are two three-dimensional structures from this family, conantokin-T and G from *Conus geographus* and *Conus tulipa* [50, 62]. One of the key features of this family is the unusually high numbers of γ-carboxyglutamic acid (Gla) residues. The main three-dimensional structural feature found in members of this family is the presence of alpha and 3_{10} helix-like areas. Members of this family have shown so far the need to have present divalent cations to form stable helices.

Contryphans

Contryphans have four three-dimensional structures available, Contryphan-R from *Conus radiatus*, Contryphan-Vn from *Conus ventricosus*, Contryphan-P from *Conus purpurascens* and Contryphan-Sm from *Conus stercusmuscarum*. Key features of this

family are unusual post-translational modifications such as Leu and Trp isomerization and Trp bromination among others. These are cyclic peptides due to the presence of a single disulfide bond. There is high sequence homology among the three-dimensional structures available for this family. In most cases one amino acid is the only difference among them. Contryphan-Vn differs from the rest by three residues yet the characteristics of two of these residues are similar. All contryphan structures reported so far exist in solution as a mixture of two conformers due to *cis-trans* isomerization between a Cys-Hyp/Pro peptide bond.

This review of the current three-dimensional structures of conotoxins is the basis for the next step of this project, the structure elucidation of the *C. regius* conotoxins. The goal of this study was to isolate and characterize as many conopeptides from this previously untapped cone snail species (conopeptidomics). This work represents a partial peptide library of conotoxins with a MW less than 3500 Da, to be later submitted for biological assays to determine function and uses. There is huge potential in the venom of *Conus regius* and these peptides from the sea may one day be the key to elucidating several neuronal processes and treat an array of neurological disorders.

REFERENCES

1. Terlau, H., et al., *Strategy for rapid immobilization of prey by a fish-hunting marine snail.* Nature, 1996. **381**(6578): p. 148-151.
2. Newcomb, R. and G. Miljanich, *Neurotoxins of cone snail venoms.* Handbook of Neurotoxicology, 2002. **1**: p. 617-651.
3. Mari, F. and G.B. Fields, *Conopeptides: Unique Pharmacological Agents that Challenge Current Peptide Methodologies.* Chimica Oggi, 2003. **21**: p. 43-48.
4. Terlau, H. and B.M. Olivera, *Conus venoms: A rich source of novel ion channel-targeted peptides.* Physiological Reviews, 2004. **84**(1): p. 41-68.
5. Cruz, L.J., et al., *Conus Venoms - a Rich Source of Neuroactive Peptides.* Journal of Toxicology-Toxin Reviews, 1985. **4**(2): p. 107-132.
6. Olivera, B.M., *EE Just lecture, 1996 - Conus venom peptides, receptor and ion channel targets, and drug design: 50 million years of neuropharmacology.* Molecular Biology of the Cell, 1997. **8**(11): p. 2101-2109.
7. Olivera, B.M., J.S. Imperial, and D. Yoshikami, . *ω-Conotoxins (Conus spp.).* Guidebook to Protein Toxins and Their Use in Cell Biology, 1997: p. 170-172.
8. Olivera, B.M., et al., *Diversity of Conus Neuropeptides.* Science, 1990. **249**(4966): p. 257-263.
9. Endean, R.J., Izatt, *Pharmacological Study of the Venom of the Gastropod Conus magus.* Toxicon, 1965. **3**: p. 81-93.

10. Bulaj G., O.B., I. Goodsell, E.C. Jimenez, J. Kranski, J.S. Nielsen, J.E. Garrett and B.M. Olivera, *Efficient oxidative folding of conotoxins and the radiation of venomous cone snails.* Proc. Natl. Acad. Sci. USA, 2003. **Suppl. 2**(100): p. pp. 14562-14568.

11. Colledge, C.J., et al., *Precursor Structure of ω-Conotoxin Gvia Determined from a cDNA Clone.* Toxicon, 1992. **30**(9): p. 1111-1116.

12. Craig, A.G., P. Bandyopadhyay, and B.M. Olivera, *Post-translationally modified neuropeptides from Conus venoms.* European Journal of Biochemistry, 1999. **264**(2): p. 271-275.

13. Pisarewicz, K., et al., *Polypeptide Chains Containing D-γ-Hydroxyvaline.* Journal of the American Chemical Society, 2005. **127**(17): p. 6207-6215.

14. Livett, B.G., K.R. Gayler, and Z. Khalil, *Drugs from the sea: Conopeptides as potential therapeutics.* Current Medicinal Chemistry, 2004. **11**(13): p. 1715-1723.

15. Heading, C.E., *Conus peptides and neuroprotection.* Current Opinion in Investigational Drugs (PharmaPress Ltd.), 2002. **3**(6): p. 915-920.

16. Craig, A.G., *The characterization of conotoxins.* Journal of Toxicology-Toxin Reviews, 2000. **19**(1): p. 53-93.

17. HU, S.-H., *Hu, S.-H., The 1.1 A crystal structure of the neuronal acetylcholine receptor antagonist, α-conotoxin PnIA from Conus pennaceus.* Structure (London), 1996. **4**(4): p. 417-23.

18. HU, S.-H., *Crystal Structure at 1.1 A Resolution of α-conotoxin PnIB: Comparison with α-conotoxins PnIA and GI.* Biochemistry, 1997. **36**: p. 11323-30.

19. Hu, S.-H., *The 1.1 A Resolution Crystal Structure of [Tyr15]EpI, a novel α-conotoxin from Conus episcopatus, solved by Direct Methods.* Biochemistry, 1998. **37**: p. 11425-33.

20. Yun, J.-A., et al., *NMR solution structure of α-conotoxin FI, a specific antagonist of nAChR, isolated from Korean cone snail Chelyconus fulmen.* Peptide Science, 2003. **40th**: p. 279-282.

21. Chi, S.-W., et al., *Solution conformation of alpha-conotoxin GIC, a novel potent antagonist of α3β2 nicotinic acetylcholine receptors.* Biochemical journal: 2004 Jun 1. **380**(Pt 2): p. 347-52.

22. Nicke, A., et al., *Isolation, structure, and activity of GID, a novel α4/7-conotoxin with an extended N-terminal sequence.* Journal of Biological Chemistry, 2003. **278**(5): p. 3137-3144.

23. Chi, S.-W., et al., *Solution conformation of αA-conotoxin EIVA, a potent neuromuscular nicotinic acetylcholine receptor antagonist from Conus ermineus.* Journal of biological chemistry. 2003 Oct 24. **278**(43): p. 42208-13.

24. Pardi, A., et al., *Solution structures of alpha-conotoxin G1 determined by two-dimensional NMR spectroscopy.* Biochemistry. 1989 Jun 27. **28**(13): p. 5494-501.

25. Han, K.H., et al., *NMR structure determination of a novel conotoxin, [Pro 7,13] alpha A-conotoxin PIVA.* Biochemistry. 1997 Feb 18. **36**(7): p. 1669-77.

26. Shon, K.-J., et al., *Three-Dimensional Solution Structure of a-Conotoxin MII, an a3b2 Neuronal Nicotinic Acetylcholine Receptor-Targeted Ligand.* Biochemistry, 1997. **36**(50): p. 15693-15700.

27. Franco, A. and F. Mari, *Three-dimensional structure of α-conotoxin EI determined by 1H NMR spectroscopy.* Letters in Peptide Science, 1999. **6**(4): p. 199-207.

28. Gehrmann, J., et al., *Solution Structure of α-Conotoxin ImI by 1H Nuclear Magnetic Resonance.* Journal of Medicinal Chemistry, 1999. **42**(13): p. 2364-2372.

29. Benie, A.J., et al., *Solution structure of α-conotoxin SI.* FEBS Letters, 2000. **476**(3): p. 287-295.

30. Cho, J.H., et al., *Nuclear magnetic resonance solution conformation of α-conotoxin AuIB, an alpha(3)beta(4) subtype-selective neuronal nicotinic acetylcholine receptor antagonist.* Journal of biological chemistry. 2000 Mar 24. **275**(12): p. 8680-5.

31. Favreau, P., et al., *Biochemical Characterization and Nuclear Magnetic Resonance Structure of Novel α-Conotoxins Isolated from the Venom of Conus consors.* Biochemistry, 1999. **38**(19): p. 6317-6326.

32. Sevilla, P., et al., *Three-dimensional structure of omega-conotoxin GVIA determined by 1H NMR.* Biochemical and biophysical research communications. 1993 May 14. **192**(3): p. 1238-44.

33. Basus, V.J., et al., *Solution structure of ω-conotoxin MVIIA using 2D NMR spectroscopy.* FEBS letters. 1995 Aug 21. **370**(3): p. 163-9.

34. Kohno, T., et al., *Three-dimensional structure in solution of the calcium channel blocker ω-conotoxin MVIIA.* Biochemistry. 1995 Aug 15. **34**(32): p. 10256-65.

35. Nemoto, N., et al., *Solution structure of ω-conotoxin MVIIC determined by NMR.* Biochemical and biophysical research communications. 1995 Feb 15. **207**(2): p. 695-700.

36. Civera, C., et al., *Solution structure determination by two-dimensional 1H NMR of ω-conotoxin MVIID, a calcium channel blocker peptide.* Biochemical and biophysical research communications. 1999 Jan 8. **254**(1): p. 32-5.

37. Kobayashi, K., et al., *Three-Dimensional Solution Structure of ω-Conotoxin TxVII, an L-Type Calcium Channel Blocker.* Biochemistry, 2000. **39**(48): p. 14761-14767.

38. Lewis, R.J., et al., *Novel ω-conotoxins from Conus catus discriminate among neuronal calcium channel subtypes.* Journal of Biological Chemistry, 2000. **275**(45): p. 35335-35344.

39. Yan, Y., et al., *Three-dimensional solution structure of ω-conotoxin SO3 determined by 1H NMR.* Chinese Science Bulletin, 2003. **48**(11): p. 1097-1102.

40. Daly Norelle, L., et al., *Structures of μO-conotoxins from Conus marmoreus. Inhibitors of tetrodotoxin (TTX)-sensitive and TTX-resistant sodium channels in mammalian sensory neurons.* Journal of biological chemistry. 2004 Jun 11. **279**(24): p. 25774-82.

41. Volpon, L., et al., *NMR solution structures of delta-conotoxin EVIA from Conus ermineus that selectively acts on vertebrate neuronal Na+ channels.* Journal of biological chemistry. 2004 May 14. **279**(20): p. 21356-66.

42. Kohno, T., et al., *Three-dimensional solution structure of the sodium channel agonist/antagonist δ-conotoxin TxVIA.* Journal of biological chemistry. 2002 Sep 27. **277**(39): p. 36387-91.

43. Savarin, P., et al., *Three-dimensional structure of κ-conotoxin PVIIA, a novel potassium channel-blocking toxin from cone snails.* Biochemistry, 1998. **37**(16): p. 5407-5416.

44. Ott, K.H., et al., *Solution structure of μ-conotoxin GIIIA analyzed by 2D-NMR and distance geometry calculations.* FEBS letters. 1991 Jan 28. **278**(2): p. 160-6.

45. Hill, J.M., P.F. Alewood, and D.J. Craik, *Three-dimensional solution structure of μ-conotoxin GIIIB, a specific blocker of skeletal muscle sodium channels.* Biochemistry, 1996. **35**(27): p. 8824-8835.

46. Mitchell, S.S., et al., *Three-dimensional solution structure of conotoxin psi-PIIIE, an acetylcholine gated ion channel antagonist.* Biochemistry. 1998 Feb 3. **37**(5): p. 1215-20.

47. Van Wagoner Ryan, M., et al., *Characterization and three-dimensional structure determination of psi-conotoxin Piiif, a novel noncompetitive antagonist of nicotinic acetylcholine receptors.* Biochemistry. 2003 Jun 3. **42**(21): p. 6353-62.

48. Al-Sabi, A., et al., *kappa M-conotoxin RIIIK, structural and functional novelty in a K+ channel antagonist.* Biochemistry, 2004. **43**(27): p. 8625-8635.

49. McDougal, O.M. and C.D. Poulter, *Three-dimensional structure of the mini-M conotoxin mr3a.* Biochemistry, 2004. **43**(2): p. 425-429.

50. Skjaerbaek, N., et al., *Determination of the solution structures of conantokin-G and conantokin-T by CD and NMR spectroscopy.* Journal of biological chemistry. 1997 Jan 24. **272**(4): p. 2291-9.

51. Pallaghy, P.K., et al., *Solution structure of contryphan-R, a naturally occurring disulfide-bridged octapeptide containing D-tryptophan: Comparison with protein loops.* Biochemistry, 1999. **38**(35): p. 11553-11559.

52. Eliseo, T., *Solution Structure of the Cyclic Peptide Contryphan-Vn, a Ca2+ - Dependent K+ Channel Modulator.* Biopolymers, 2004. **74**: p. 189-98.

53. Pallaghy, P.K., et al., *Structures of the Contryphan Family of Cyclic Peptides. Role of Electrostatic Interactions in cis-trans Isomerism.* Biochemistry, 2000. **39**(42): p. 12845-12852.

54. Jacobsen, R., et al., *The contryphans, a D-tryptophan-containing family of Conus peptides: interconversion between conformers.* Journal of Peptide Research, 1998. **51**(3): p. 173-179.

55. Hill, J.M., P.F. Alewood, and D.J. Craik, *Conotoxin TVIIA, a novel peptide from the venom of Conus tulipa: 2. Three-dimensional solution structure.* European Journal of Biochemistry, 2000. **267**(15): p. 4649-4657.

56. Hill, J.M., P.F. Alewood, and D.J. Craik, *Solution structure of the sodium channel antagonist conotoxin GS: a new molecular caliper for probing sodium channel geometry.* Structure (London), 1997. **5**(4): p. 571-583.

57. Nicke, A., S. Wonnacott, and R.J. Lewis, *alpha-Conotoxins as tools for the elucidation of structure and function of neuronal nicotinic acetylcholine receptor subtypes.* European Journal of Biochemistry, 2004. **271**(12): p. 2305-2319.

58. Franco, A. and F. Mari, *Three-dimensional structure of alpha-conotoxin EI determined by H-1 NMR spectroscopy.* Letters in Peptide Science, 1999. **6**(4): p. 199-207.

59. Lewis, R.J., et al., *Novel w-conotoxins from Conus catus discriminate among neuronal calcium channel subtypes.* Journal of Biological Chemistry, 2000. **275**(45): p. 35335-35344.

60. Kohno, T., et al., *Three-dimensional Solution Structure of the Sodium Channel Agonist/Antagonist d-Conotoxin TxVIA.* Journal of Biological Chemistry, 2002. **277**(39): p. 36387-36391.

61. Van Wagoner Ryan, M. and M. Ireland Chris, *An improved solution structure for ψ-conotoxin PiiiE.* Biochemistry. 2003 Jun 3. **42**(21): p. 6347-52.

62. Warder, S.E., *The NMR solution structure of the NMDA receptor antagonist, conantokin-T, in the absence of divalent metal ions.* 411, 1997: p. 19-26.

Chapter I

Mini-M Conotoxins from the Venom of *Conus regius*: a Western Atlantic Worm-Hunting Cone Snail Species.

ABSTRACT

Venomous marine mollusks belonging to the genus *Conus* (cones snails) utilize a unique neurochemical strategy to capture their prey. Their venom is composed of a complex mixture of highly modified peptides (conopeptides) that interact with a wide range of neuronal targets. This work describes the isolation and characterization of twelve novel conopeptides belonging to the newly described mini-M conotoxins from the venom of *Conus regius*. These novel conopeptides exhibit a differential hydroxylation strategy that could affect the neuronal targeting of this set of mini-M-conotoxins. The observed preferential hydroxylation in these conopeptides could imply the need for cone snails to augment their venom arsenal for better neuronal targeting and prey capturing. These *C. regius* mini-M conotoxins can be further classified into subtypes m-1, m-2 and m-3. These subtypes show little sequence homology and their loop size variability is significant.

1. INTRODUCTION

Cone snails are predatory marine mollusks that utilize venom to capture their prey. These animals are among the most prolific and versatile peptide engineers known in nature. Their venom is an extremely complex concoction of 20-200 compounds (Figure. 1) mostly composed of modified peptides (conopeptides) which are part of a biochemical strategy used for predation. These unique marine organisms, which can prey upon fish (piscivorous), mollusks (molluscivorous) and worms (vermivorous), deliver their complex venom through a specialized radular tooth that serves as both harpoon and disposable hypodermic needle. Conopeptides elicit a wide range of strong neurophysiological responses [1-5]; in few instances, human fatalities have resulted from mishandling these animals [6]. The development of such extremely potent and biochemically diverse venom is likely to be an evolutionary adaptation designed to compensate for the lack of mobility of cone snails when compared to other marine predators. The precise composition of cone snail venom is species specific [1, 3, 5]. *Conus* venom is the product of 55 million years of evolutionary refinement that has yielded a complex library of over 100,000 neuroactive conopeptides, as this genus comprises over 1000 species distributed in the tropical and sub-tropical areas of Atlantic, Indian and Pacific Oceans [7].

Conopeptide precursors are ribosomally-expressed proteins that subsequently undergo post-translational modifications and proteolytic cleavage to form the mature conotoxin [8]. Conopeptides inherently contain high degrees of modified amino acids (usually combinations of them), such as cystines, hydroxyproline, γ-

carboxyglutamate, Br-Trp, D-Trp, D-Leu, D-Phe [9], pyro-Glu, glycosylated Ser/Thr, and sulfated Tyr [10-12]. These modifications confer conopeptides with unique stability and exquisite specificity towards neuronal targets [3, 11, 13], enabling cone snails to capture prey.

Among the predominant components of the venom of Cone snails are a variety of cysteine-stabilized (with two or more disulfide bridges) conotoxin scaffolds. Conotoxins target specifically ion channels and neuronal receptors as part of the neurochemical strategy for predation. Conotoxins are important tools for investigating ion channel and receptor function and have great potential pharmacological applications [2, 14]. Conotoxins are grouped into various superfamilies (O, M, A, S, T, P, I), each with highly conserved signal sequences in their precursor proteins and a characteristic cystine arrangement in the mature peptides. Within the superfamilies, conotoxins are further classified into families according to their pharmacological targets, which include voltage-gated ion channels (Na^+, K^+ and Ca^{2+}), ligand-gated ion channels (nAChR and 5-HT_3R), receptors (neurotensin type 1, α1 adrenergic, NMDA, RFamide and vassopressin) and neurotransmitter transporters (NE) [2, 15].

The M-superfamily of conotoxins is characterized by 3-loop/6-Cys (CC-loop1-C-loop2-C-loop3-CC) arrangement, where the disulfide pairing is variable as well as the size of the loops. Families within the M-superfamily (μ, κM, ψ, m1, m2 and m3) shows great variability in neuronal targeting and structural composition. Conotoxins belonging to the M-superfamily are found in all varieties of cone snail species (from

Piscivorous to Vermivorous); however, the structural diversity of these compounds is only starting to be explored.

As part of our efforts on the analysis and characterization of conopeptides isolated from cone snail species from the Americas, we decided to undertake the isolation and structural analysis of conopeptides from *Conus regius* (species code: reg), a widespread worm-hunting cone snail of the Western Atlantic Ocean. Here we present the results from the isolation and structural analysis of twelve novel conopeptides belonging to the newly described mini-M conotoxins from the venom of *Conus regius*. These *C. regius* mini-M conotoxins can be further classified into subtypes m-1, m-2 and m-3. These subtypes show little sequence homology and their loop size variability is significant. These peptides exhibit preferential proline hydroxylation strategy that is likely to affect their neuronal targeting.

2. MATERIALS AND METHODS

2.1 Specimen collection.

Specimens of *C. regius* (30-70 mm) were collected off the Florida Keys (Plantation Key), USA, using SCUBA at depths ranging from 2-20 m. All snails were kept in aquaria prior to transportation to the lab, where there were dissected and immediately frozen at –80°C.

2.2 Crude venom extraction.

Venom ducts dissected from specimens of *C. regius* were homogenized in 0.1% TFA at 4°C. Whole extracts were centrifuged at 10,000 x g for 20 min, at 4°C, and the resulting pellets were washed three times with 0.1% TFA and re-centrifuged under identical conditions. The supernatants containing the soluble peptides were pooled, lyophilized, and stored at -80°C until further use.

2.3 Peptide purification.

Crude venom was initially fractionated by RP-HLPC using a C18 semipreparative column (Vydac, 218TP510, 10 mm x 250 mm; 5µm particle diameter; 300 Å pore size); however, due to the complexity of the venom, a different fractionation scheme was required. 50 mg of crude venom were separated by SE-HPLC on a Pharmacia Superdex-30 column (2.5 cm x 100 cm) equilibrated and eluted with 0.1 M NH_4HCO_3 using a flow rate of 1.5 ml/min. Further separation of fractions obtained from the Superdex-30 column was performed on a Superdex Peptide (Amersham Biosciences) column (10 x 300 mm). Chromatographic fractions were monitored at λ

= 220, 250 and 280 nm. Additional purification of peptide-containing peaks was achieved by RP-HPLC on a C18 semipreparative column (Vydac, 218TP510, 10 x 250 mm; 5 µm particle diameter; 300 Å pore size) equipped with a C18 guard column (Upchurch Scientific, AC-43 4.6 mm) at a flow rate of 3.5 ml/min. Further peptide purification was carried out by re-chromatographing fractions on an analytical C18 column (Vydac, 238TP54, 4.6 x 250 mm; 5 µm particle diameter; 300 Å pore size), with a flow rate of 1 ml/min. For semipreparative and analytical RP-HPLC separation, the buffers were 0.1% TFA (buffer A) and 0.1% TFA in 60% acetonitrile (buffer B). Peptides were eluted with an incremental linear gradient of 1% B /min. Absorbances were monitored at λ = 220 and 280 nm. All HPLC fractions were manually collected, lyophilized and kept at -40°C prior to further use.

2.4 Reduction and alkylation of cysteyl residues.

Reduction and alkylation of cystine groups were carried out as previously described [16] with slight modifications. An aliquot of each peptide (~1 pmol) was dried, re-dissolved in 0.1 M Tris-HCl (pH 6.2), 5 mM EDTA, 0.1% sodium azide and reduced with 20 mM DTT. Following incubation at 60°C, for 30 min, peptides were alkylated in a final volume of 15 µl with 50 mM IAM and 2 µl of NH_4OH (pH 10.5), at room temperature, for 1 h, in the dark. The reduced and alkylated peptides were purified using a Zip Tip (C18, size P10, Millipore).

2.5 Peptide sequencing.

Alkylated peptides were adsorbed onto Biobrene-treated glass fiber filters and amino acid sequences were carried by Edman degradation using an Applied Biosystems Procise model 491A Sequencer. The concentration of the peptides was determined by using the calibrated intensities of the first five PTH-amino acids residues on samples that were not reduced and alkylated.

2.6 Molecular mass determination.

Positive ion MALDI-TOF mass spectrometry was carried out on an Applied Biosystems Voyager-DE STR spectrometer. Samples were dissolved in 0.1% TFA, 50% acetonitrile, and applied on α-cyano-4-hydroxycinnamic acid matrix. Spectra were obtained in the linear and reflector mode using Calmix 1 and Calmix 2 (Applied Biosystems) as external calibration standards.

2.7 Disulfide connectivity analysis.

The analysis of the disulfide connectivity was carried out on the native conopeptide. About 5 nmol of the lyophylized conopeptide was partially reduced using tris(2-carboxyethyl)phosphine (TCEP) in pH 3.0 citrate buffer to produce a mixture of partially reduced peptide isomers; the nascent sulfhydryls are immediately cyanylated by 1-cyano-4-dimethylamino-pyridinium tetrafluoroborate (CDAP) under the same buffered conditions [17]. The cyanylated peptide is then purified on an analytical C18 column (Vydac, 238TP54, 4.6 x 250 mm; 5 μm particle diameter; 300 Å pore size), with a flow rate of 1 ml/min. The purified cyanylated peptide is then cleaved at the

peptide bond adjacent to the N-terminal side of the cyanylated cysteines by exposing it to aqueous ammonia. This step forms peptide fragments that are still linked by the residual disulfide bonds; these bonds are then completely reduced with TCEP and the peptide fragments are analyzed by mass spectrometry.

2.8 NMR spectroscopy.

NMR spectra were acquired on a Varian Inova 500 MHz instrument equipped with PFG, 3xRF channels and waveform generators. Nanomolar quantities of the native conopeptides directly isolated from the venom (reg12e = 32 nmoles, reg12l = 19 nmoles, reg12i = 4 nmoles, the rest of the reg12 peptides were below 1 nmole) were dissolved in 40 µl of water with 10% D_2O (used for locking purposes) and 4 nanomoles of TSP and placed in 1.7 mm NMR tubes (Wilmad WG-1364-1.7). The pH was adjusted using 0.01 M solutions of HCl and NaOH and a Thermo micro-pH probe. Spectra were obtained using a Varian *g*HCN (generation 5) high performance 3 mm probe (pw90 = 3 µs, at the upper limit of the linear range of the RF amplifier) with a 1.7 mm capillary adaptor (Wilmad V-GFK-10/1.7). NMR experiments were recorded at pH 3.60 and at different temperatures (0, 10 and 25°C) in order to achieve the best chemical shift dispersion possible to aid the sequence specific assignments. For 1D NMR experiments, the water signal was suppressed by using either WET [18] or presaturation. In addition to the concentrations determined from sequencing, peptide concentrations were also evaluated by integrating the NMR signals of selected methyl groups and using the known concentration of TSP as an internal standard [19] or the signal of selected methyl groups from peptides with known

concentration as external standards. For 2D experiments, water suppression was carried out using WATERGATE(*wg*) [20] in combination with 3919 purge pulses with flipback [21], which were implemented in the TOCSY and NOESY pulse sequences. The *wg*TOCSY and *wg*NOESY experiments were used to obtain information on sequence-specific assignments [22]. All 2D-NMR spectra were recorded in the phase sensitive mode using the States-Haberkorn method [23] with a spectral width of 6000 Hz and 2K data points. For the *wg*TOCSY experiment, 160 scans for each of the 96 FIDs were acquired with relaxation delay of 1.7 s and a mixing time of 120 ms. The 2D *wg*NOESY spectra were recorded using 256 scans for each of the 128 FIDs acquired with a 1.7 s relaxation delay and a mixing time of 200 ms. All 2D NMR data were processed using VNMR 6.1C (Varian NMR Instruments) on Sun Blade 150 workstations. FIDs were apodized with a shifted sine bell window function and linearly predicted to 1K points in *t*1 and zero-filled to 2k x 2k data matrices. The data was baseline corrected in F2 by applying a polynomial function.

2.9 Molecular model of reg12e.

Molecular models were built by comparative modeling methods [24] based on the NMR structure of the mr3a conotoxin as template and using Modeller (version 8.0). Briefly, conotoxin sequences were aligned according to the standard routine in the program using the PDB entry as template. A set of 10 model structures was built accordingly and the selected structure was that of better target Modeller energy.

2.10 Nomenclature.

In this chapter we adopted a nomenclature of three letters to designate *Conus* species, because the one or two letter nomenclature currently in place will not be enough to describe the large number of different non-fish-hunting species, especially those with similar first letter names. We decided to use the three letters "reg" to name the peptides from *C. regius.* Arabic numbers were used to represent the disulfide framework; 12 has already been assigned to the mini-M conotoxins [25], in spite of the fact that some mini-M conotoxins were previously designated with framework 3 [26], which is the disulfide pairing of the maxi-M or m-4 conotoxins. The letter after the framework number indicates the order of elution on RP-HPLC.

3. RESULTS

3.1 Peptide purification.

The venom of *Conus regius* (Figure 1A) is an extremely complex mixture of peptides and proteins whose direct separation is shown in Figure 1B. More than 100 fractions can be obtained from this separation. However, most of these fractions show multiple components, for which we were required to adopt an improved separation scheme that includes a prefractionation step using size exclusion chromatography on a Superdex-30 column, followed by a refined peptide-optimized size exclusion step on a Superdex Peptide column (Figure 2). The resulting fractions are then separated by reversed phase on a peptide-optimized C18 Vydac Everest column. Most of the resulting fractions are single-component (Figure 3) and were subsequently analyzed by mass spectrometry (MALDI-TOF and ESI-Q-TOF), NMR spectroscopy and peptide sequencing by Edman degradation chemistry. Using this methodology, the most significant components of the venom of *C. regius* can be sequenced (Conopeptidome) and the components of the venom can be grouped in the different families of conopeptides. The reg12 peptides were separated by the Superdex-30 in spite of their similar size. This is not entirely unexpected, since this column is also known to partition analytes by hydrophobic interactions [27].

A

B

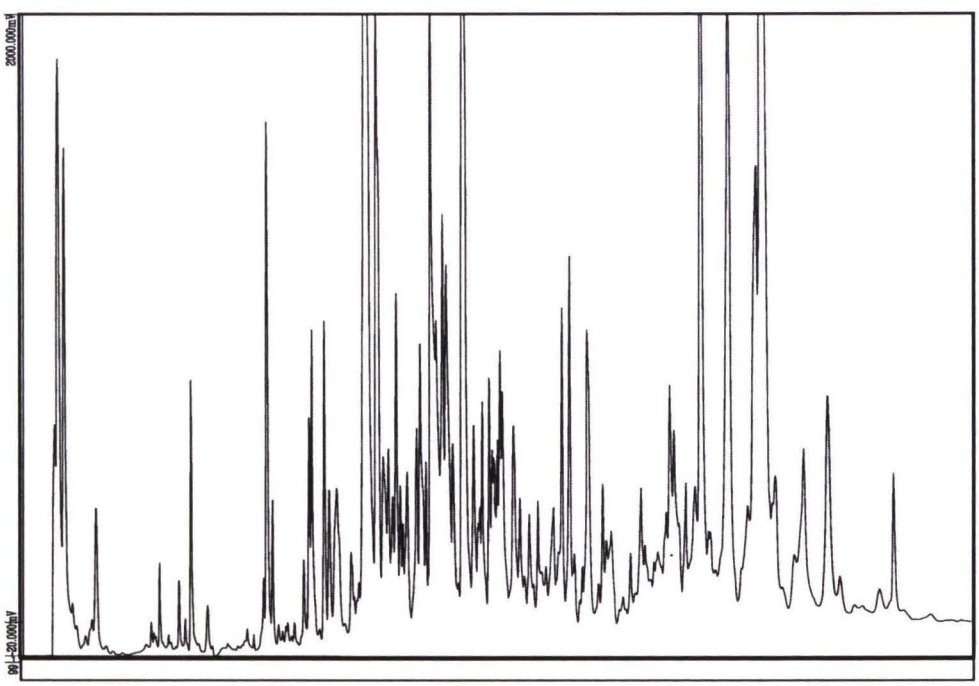


Figure 1. A. Live specimen of *C. regius* and **B.** the corresponding RP-HPLC chromatogram (C-18 Vydac, linear gradient 0-60% H_2O/ACN with 0.1% TFA in 100 min) of its crude venom. *C. regius* a widespread Western Atlantic worm-hunting species. This specimen was collected off the Florida Keys in 2-meter depth. The *C. regius* shows another shell variant known *C. regius citrinus* whose shell is bright orange. This variation has no effect in their venom composition.

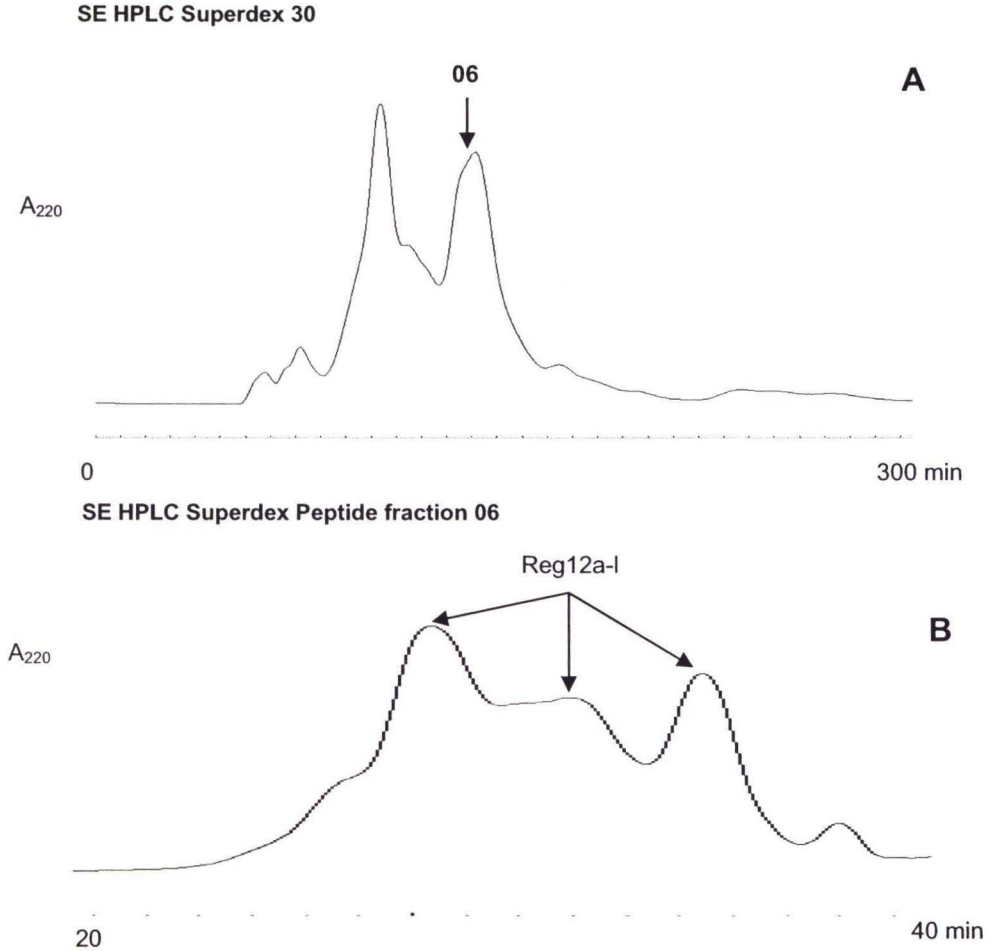


Figure 2. Size exclusion chromatogram of *C. regius* dissected venom. (A) 50 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 6 was chosen for further separation. (B) Separation of fraction 6 using a Superdex Peptide column using a 0.1 M ammonium bicarbonate solution.

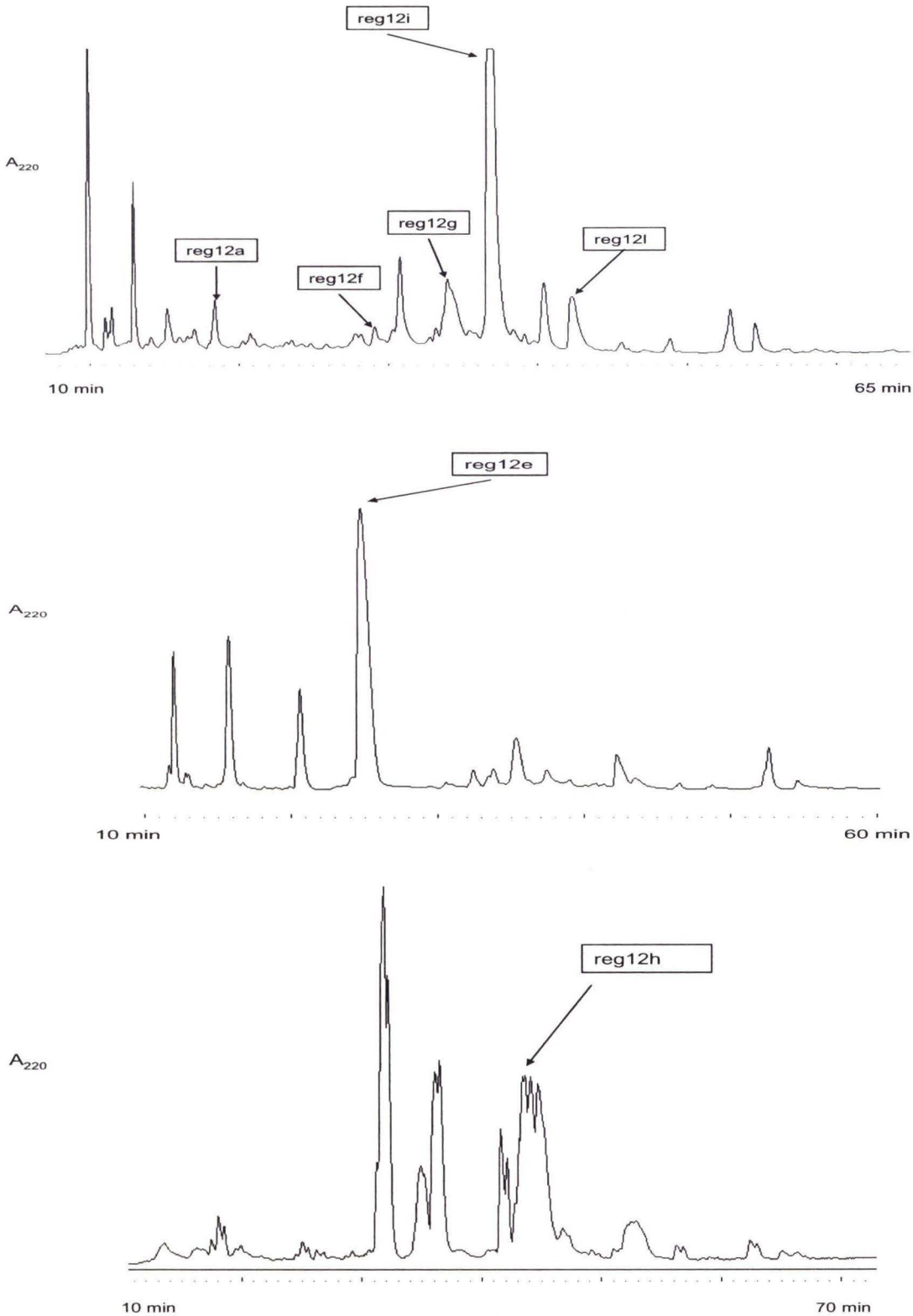


Figure 3. Semi-preparative RP-HPLC chromatograms showing the location of reg12 peptides coming from fraction 6 of size-exclusion chromatography.

3.2 Reduction/Alkylation and peptide sequence determination.

The purified peaks were subjected to reduction with DTT and alkylation with iodoacetamide. Mass spectrometry of the reduced/carboxymethylated peptides and of native peptides showed a mass difference consistent with the presence of six cysteine residues for each peptide. The reduced/alkylated conotoxins were sequenced to completion by Edman degradation. The sequences of the reg12 conotoxins are shown in Table 1. It is worth noticing the differences between peptides reg12a and reg12d since they belong to the same subtype m-3. reg12a and reg12d only differ by two amino acid residues. reg12a has at position 1 a glycine residue while reg12d has a leucine residue and reg12a has at position 13 a threonine residue while reg12d has an alanine residue. These conotoxins contained between 13 and 22 residues and the six cysteines were separated by three loops of variable sizes (loop1 = 3-5 amino acids, loop2 = 2-4 amino acids and loop3 = 1-3 amino acids). This arrangement of Cys residues within the M-conotoxin families has been classified according to the number of residues in loop3 as m-1, m-2 and m-3. Additionally, a BLAST search [28] of the databases (Swissprot/EMBL, PIR, PDB and nrdb95) did not show any significant sequence homology to reported proteins and peptides. Several mini-M (Table 1), with similar loop spacing, have been reported from mollusk-hunting and worm-hunting *Conus* species of Indo-Pacific origin [29-31].

Mini-M	Sequence	Loops	Subtype
reg12a	G**CC**OOQW**C**GOD--**C**TSO**CC**	4/3/3	M-3
reg12b	**CC**AIRL**C**NVYL-**C**GS-**CC**O	4/4/2	M-2
reg12c	-**CC**AFOQW**C**GAG-**C**IVO**CC**	5/3/3	M-3
reg12d	L**CC**OOQX**C**GOD--**C**ASO**CC**	4/3/3	M-3
reg12e	-**CC**TAL-**C**SRYH-**C**L-P**CC**	3/4/2	M-2
reg12f	K**CC**MRPI**C**T----**C**--O**CC**IGP	4/1/1	M-1
reg12g	G**CC**PFPA**C**THTII**C**R--**CC**	4/5/1	M-1
reg12h	-**CC**MAL—**C**SRYH-**C**L-P**CC**	3/4/2	M-2
reg12i	G**CC**SOWN**C**IQLRA**C**--O**CC**ON	4/5/1	M-1
reg12j	G**CC**SOWN**C**IQLRA**C**--G**CC**	4/5/1	M-1
reg12k	K**CC**MRPI**C**M----**C**--O**CC**IGP	4/1/1	M-1
reg12l	R**CC**PMPG**C**FAGPF**C**--P**CC**PV	4/5/1	M-1
Mr3a	G**CC**-GSFA**C**RFG-**C**VO--**CC**V	4/3/2	M-2
Bt12b	**CC**ELP--**C**HG--**C**VP--**CC**WP	3/2/2	M-2
RIIIK	LOS**CC**SLNNLRL**C**OVOA**C**KRNO**CC**T	7/4/4	

from cDNA of *C. radiatus*
Shaker K+ channel blocker

maxi-M (M-4)

μGIIIA	RD**CC**TOOKK**C**KDRQ**C**KOQR**CC**A	5/4/4
μPIIIA	RL**CC**GFOKS**C**RSRQ**C**KOHR**CC**	5/4/4
ψPIIIE	HO**CC**LYGK-**C**RRYOG**C**SSASAS**CC**QR	4/5/6

non comp. nAChR

Table 1. reg12 sequences from *Conus regius* and other mini M peptides

3.3 Mass spectrometry of purified peptides.

Mass spectrometry carried out using MALDI-TOF in the reflector mode (M/ΔM resolution ~ 10,000) yielded the following monoisotopic molecular ions: reg12a = 1814.0 Da, reg12b = 2060.45 Da, reg12c = 1795.0, reg12d = 1998.2 Da, reg12e = 1669.8 Da, reg12f = 1738.0 Da, reg12g = 1823.0 Da, reg12h = 1700.5 Da, reg12i = 2111.8 Da, reg12j=2145 Da, reg12k = 1768.0 Da and reg12l = 1981.6 Da. Mass analysis of the reduced/carboxymethylated peptides and the native peptides showed a mass difference consistent with the presence of six cysteine residues in each peptide. The masses obtained for the peptides were for the most part in agreement with the calculated theoretical monoisotopic values determined for the assigned sequences. reg12b, 12d, 12j and 12l did not fully sequence. Repeated attempts were carried out with no success so no theoretical calculated molecular weight was obtained. The difference in the calculated molecular weights and the experimental molecular weights from MALDI-TOF for the peptides in sequence agreement resulted in four peptides amidated at the C-terminus, reg12g, reg12h, reg12i, and reg12k. The rest of the reg12 conotoxins were found to be carboxylated at the C-terminal. This includes reg12e, the most abundant peptide (Figure 4). The calculated molecular weights (MWcal) were obtained using Protein Prospector [32].

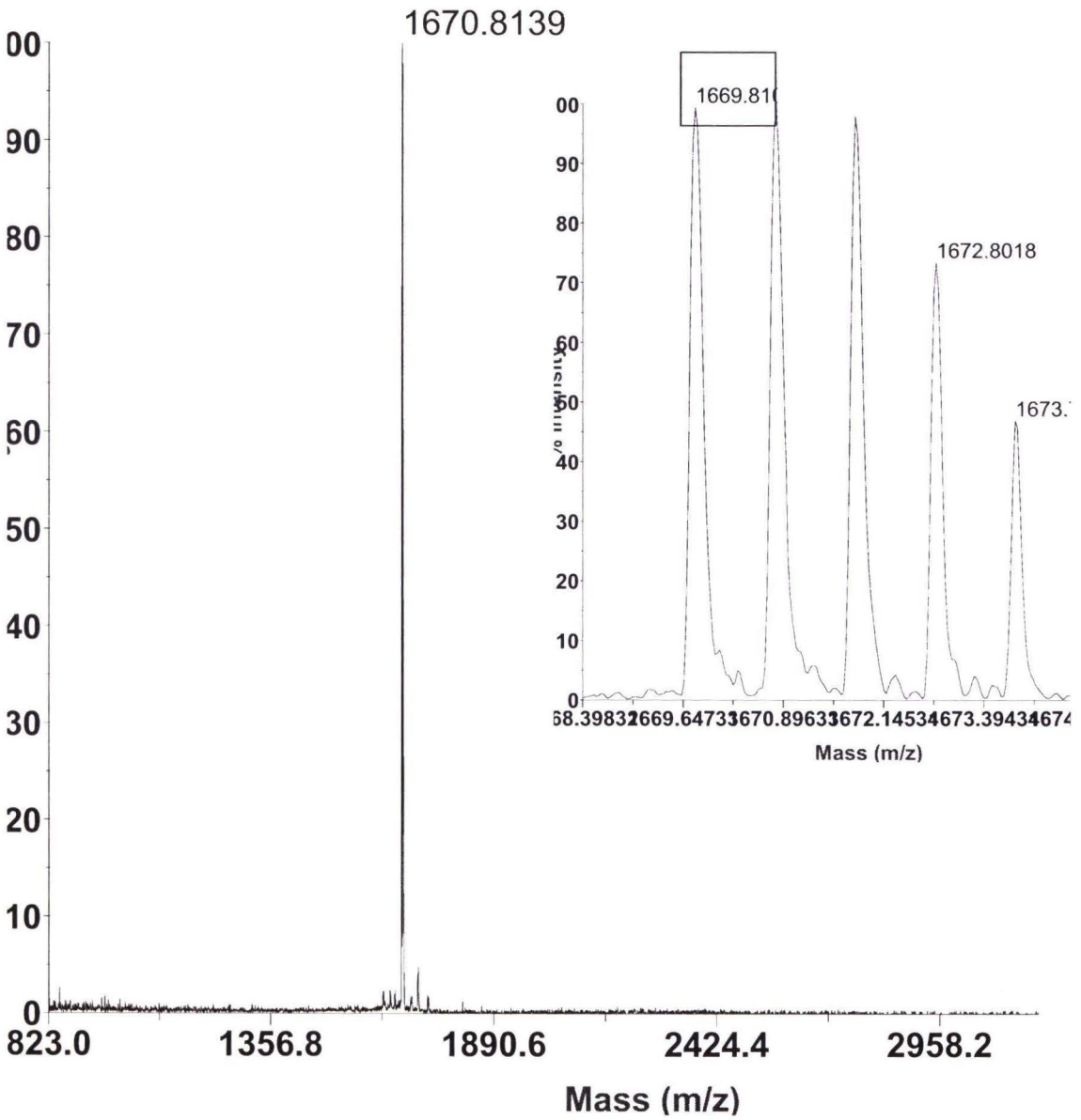


Figure 4. MALDI-TOF mass spectra for native reg12e. Insert shows the isotopic distribution with the correct molecular weight.

3.4 Disulfide connectivity analysis.

The presence of six cysteine residues in these sequences indicated the presence of three disulfide bridges. Due to the limited quantities of native conotoxins available after their purification, the disulfide connectivity was established by partial reduction and cyanylation using TCEP and CDAP, respectively [17], with subsequent RP-HPLC separation (Figure 5). Besides fraction b, which corresponds to the native reg12e the rest of the fractions indicate cyanylation at 2, 4 and 6 cysteines. Further base catalyzed cleavage at the N-terminal peptide bond of cyanylated cysteinyl residues was carried out for all fractions containing 2 and 4 cyano groups, resulting in the formation of 2-iminothiazolidine-4-carboxyl peptides. Fractions f (m/z 1721.7) and l (m/z 1774.8), addition of 2 and 4 cyano groups correspondingly, yielded the clearest results (Figure 6) in addition to being some of the most abundant peaks. The m/z 1596.9 in figure 6 corresponds to the fragment cyanylated and cleaved at the first and sixth cysteine, indicating a disulfide bond between them. The m/z 1077.6 in Figure 6 corresponds to the fragment cyanylated and cleaved at the second and fourth cysteine, indicating a disulfide bond between them. By process of elimination the third disulfide bond is between the third and fifth cysteine. This indicates a disulfide bond pattern between cysteines of 1-6, 2-4 and 3-5, as would be expected from an m2 mini-M as shown for mr3a [25].

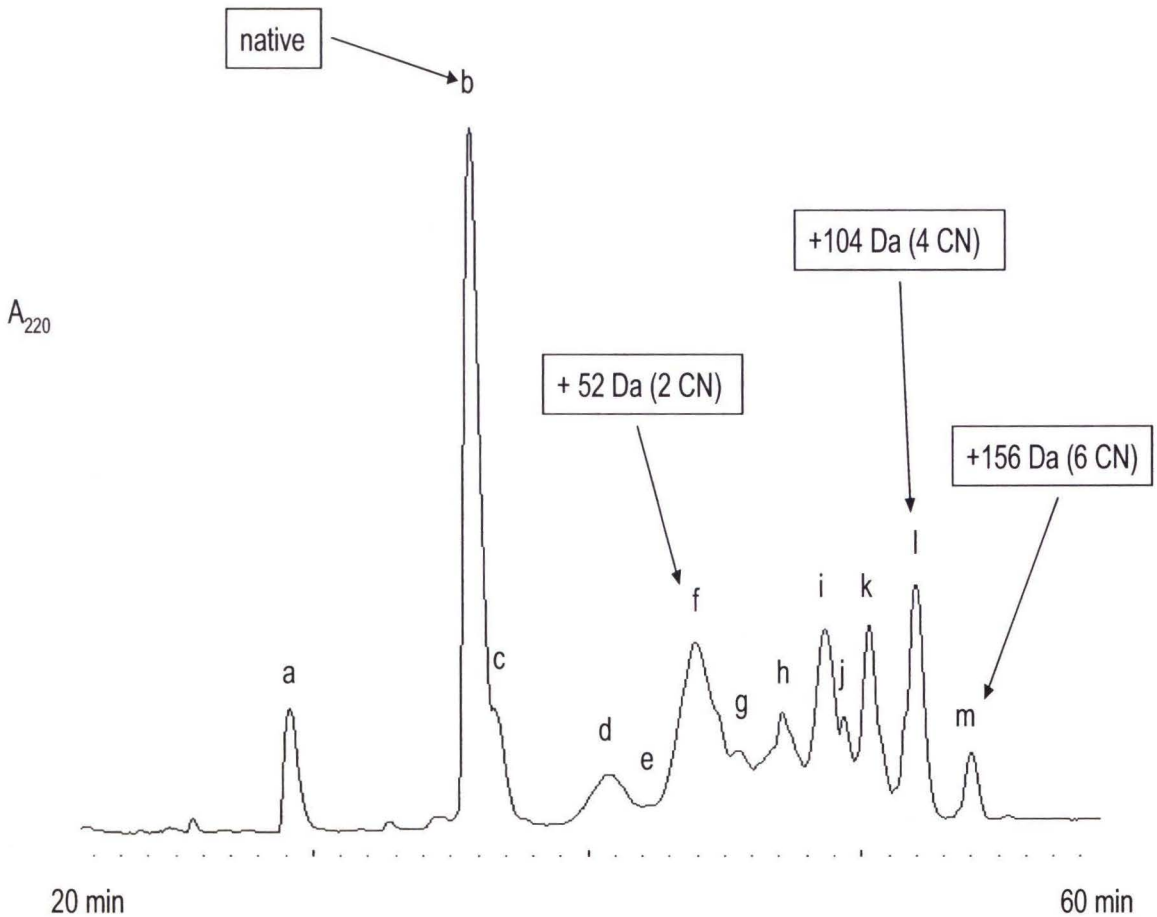


Figure 5. Analytical RP-HPLC separation of partially reduced and cyanylated reg12e isomers. The molecular weight of the peptide changes by 26Da for each cysteine cyanylated.

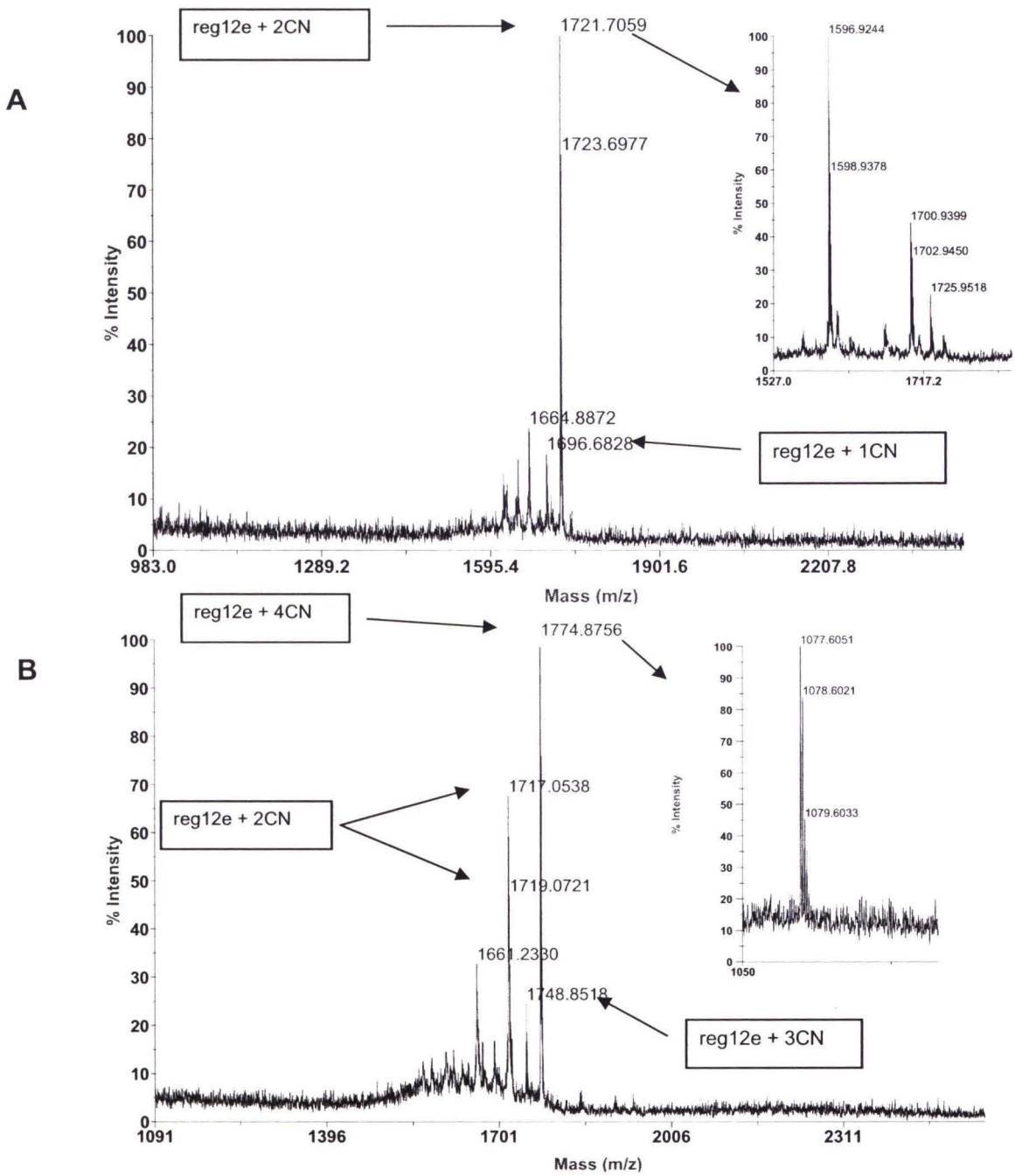


Figure 6. MALDI-TOF of (A) Fraction f (m/z 1721.7) and the resulting fragment (m/z 1596.9). (B) Fraction l (m/z 1774.8) and the resulting fragment (m/z 1077.6). Inserts represent the fragments obtained after cleavage with ammonium hydroxide. These fragments are the fingerprint for the disulfide bond connectivity.

3.5 NMR spectroscopy.

We were able to obtain NMR spectra (1D) of all conotoxins and (2D) of nanomolar quantities (nanoNMR) of two conotoxins directly isolated from the venom of the cone snails. 2D *wg*TOCSY and 2D *wg*NOESY spectra were obtained for reg12e and reg12l. Figure 7 shows the 1D of reg12e and reg12l, Figure 8 shows the *wg*TOCSY and *wg*NOESY spectra of reg12e and Figure 9 shows the 1D, *wg*TOCSY and *wg*NOESY spectra of reg12l at 25°C. In spite of their small size (15 and 17 residues), a significant number of NOE (>100) cross-correlations were found for these peptides. Spectra at lower temperatures (0°C and 10°C, data not shown) resulted in even higher number of NOE cross-correlations. The reg12e and reg12l conotoxins have very well defined structure in solution at room temperature. A preliminary table with the sequence specific assignments for reg12e is reported on Table 2. Structural analysis of conopeptides can be carried out directly from the venom of cone snails.

Residue	NH	αH	βH	γH	δH	Others
Cys1		4.27	3.18			
Cys2	9.07	4.84	3.25,3.02			
Thr3	8.69	4.41	4.31	1.24		
Ala4	8.13	4.35	1.43			
Leu5	8.44	4.34	1.78	1.59	0.92,0.80	
Cys6	8.27	4.61	3.41,3.27			
Ser7	8.59	4.39	4.03,3.90			
Arg8		4.00	1.65,1.46	1.18,1.05	2.97	NH 7.05
Tyr9	7.97	4.53	3.17,2.69			7.16,6.81
His10	8.06	4.65	3.36,3.25			8.62,7.28
Cys11	8.09	4.67	3.34,3.00			
Leu12	8.26	4.57	1.56	1.44	0.98,0.88	
Pro13		4.61	2.41,1.77	2.13,2.01	3.53	
Cys14	9.30	4.92	3.27,3.14			
Cys15	8.63	4.41	3.36,2.91			

Table 2. Sequence specific assignments for reg12e.

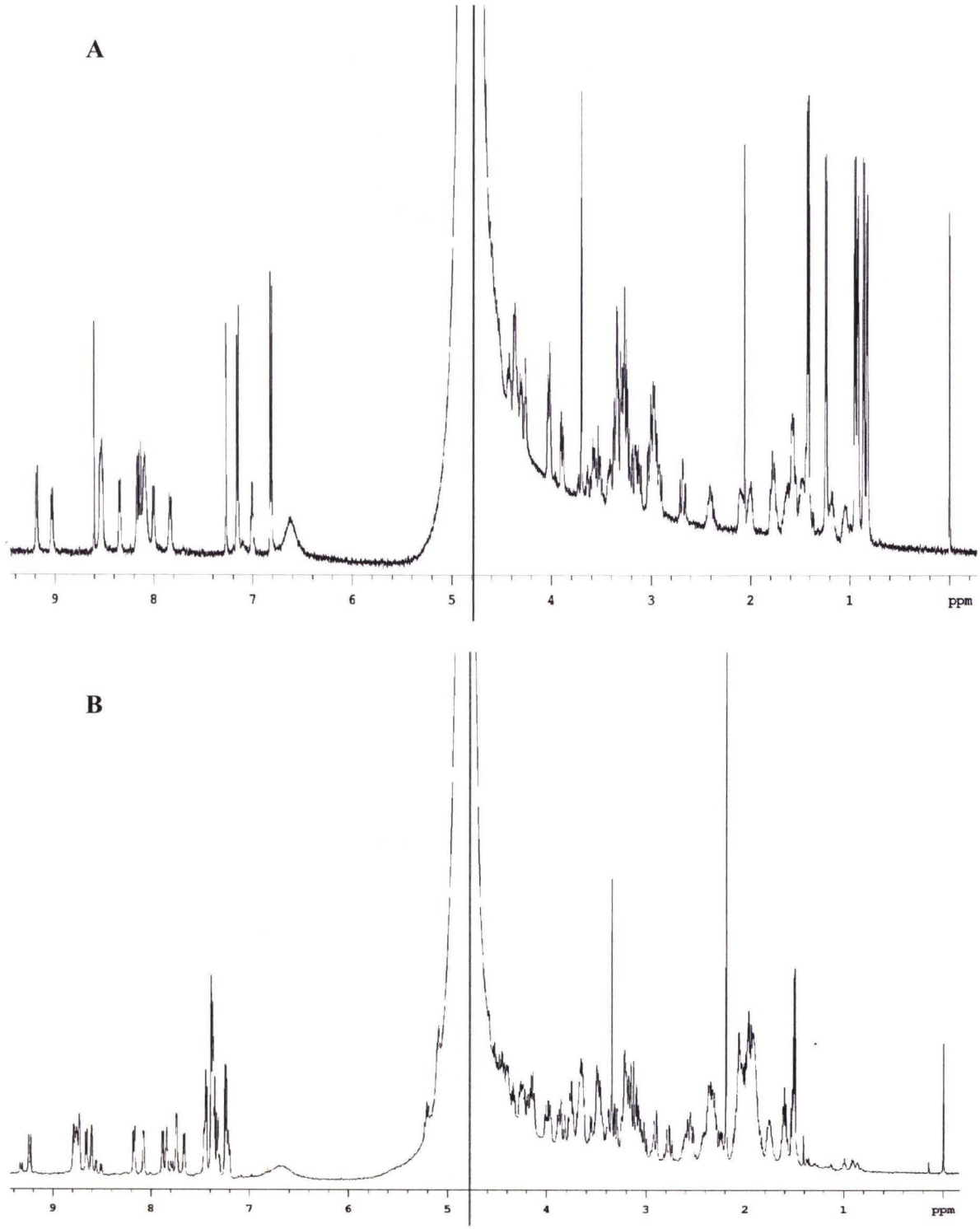


Figure 7. Proton 1D NMR spectra at 25°C using a 3 mm NMR tube in 5 mm gHCN probe of reg12e (A) and reg12l (B).

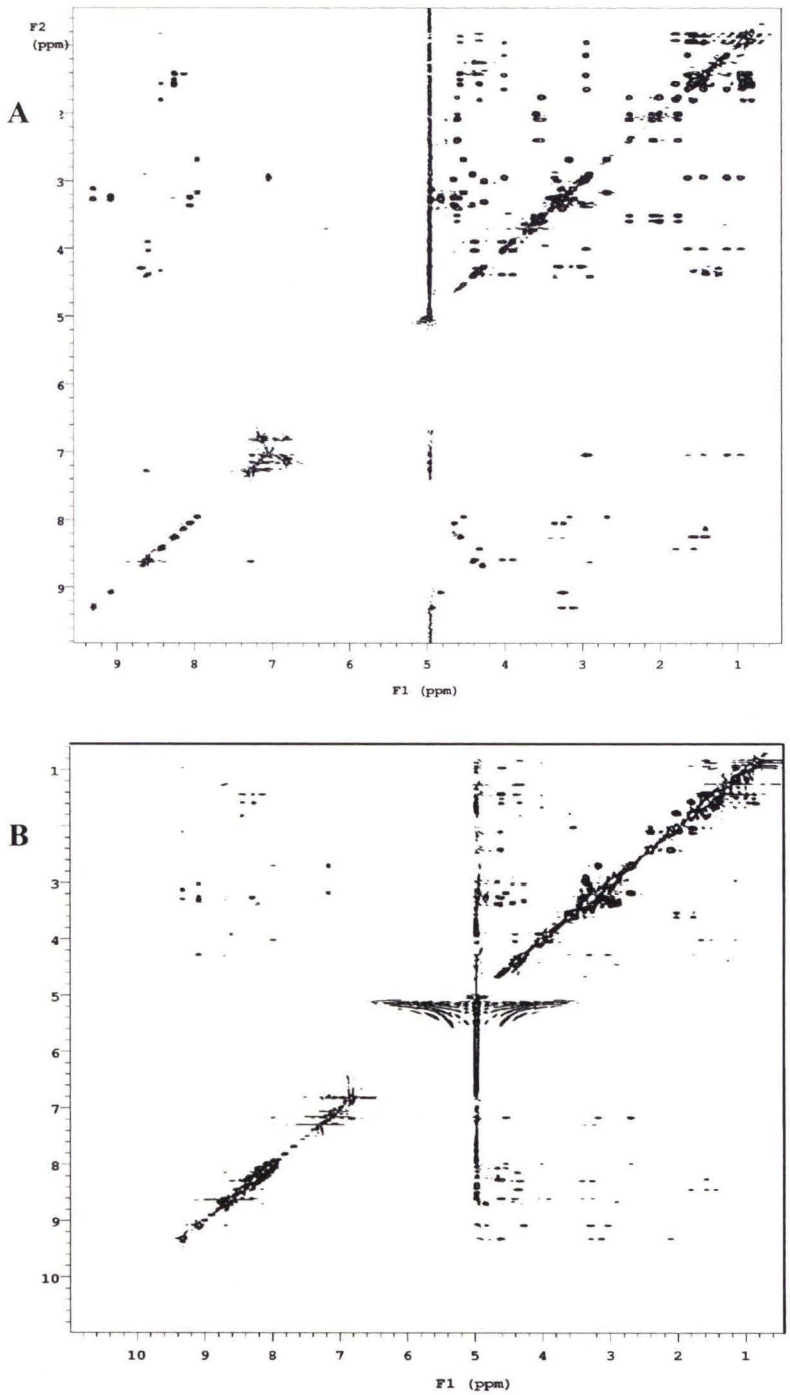


Figure 8. 2D NMR spectra of reg12e at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) wgTOCSY, (B) wgNOESY.

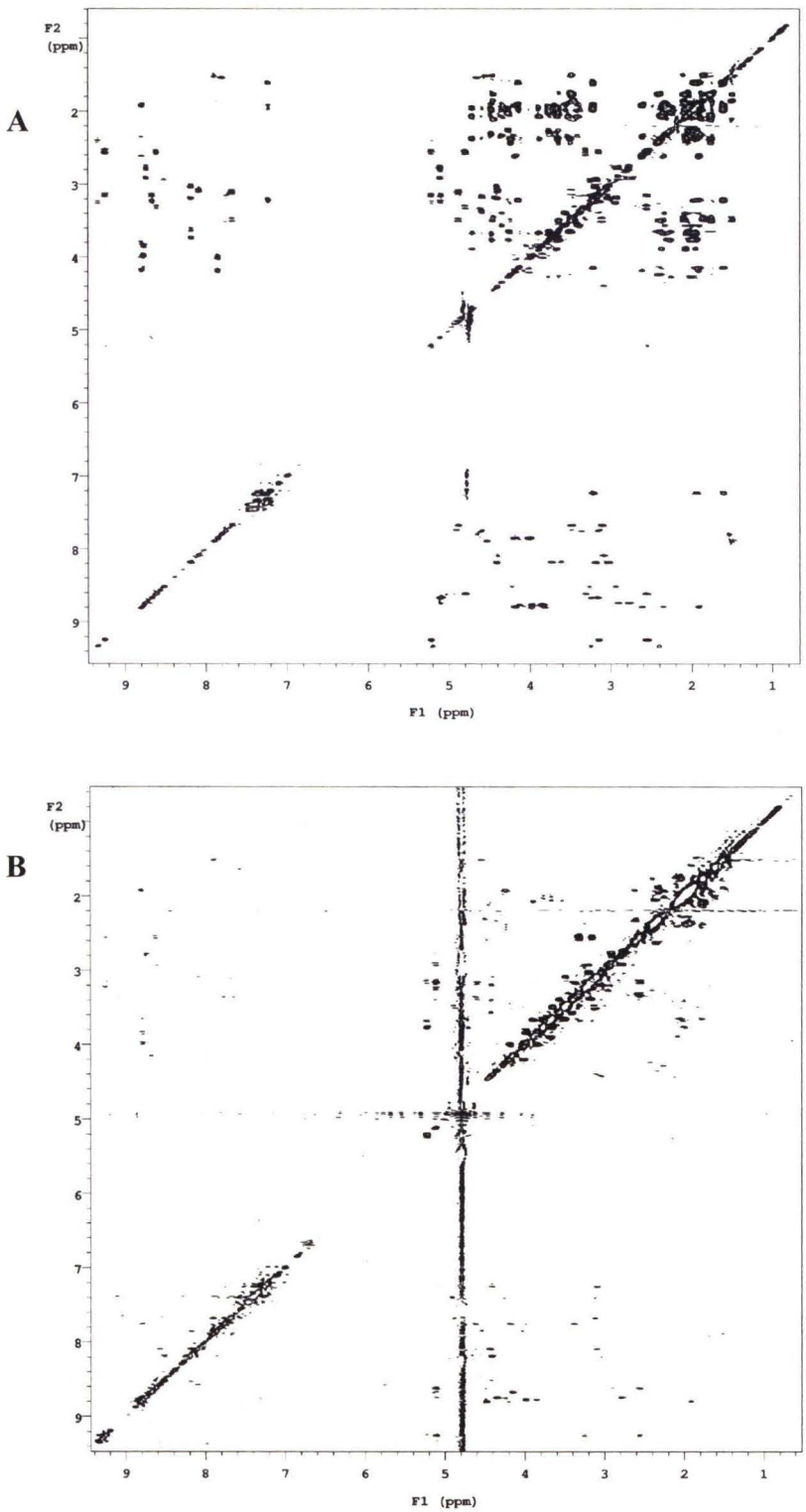
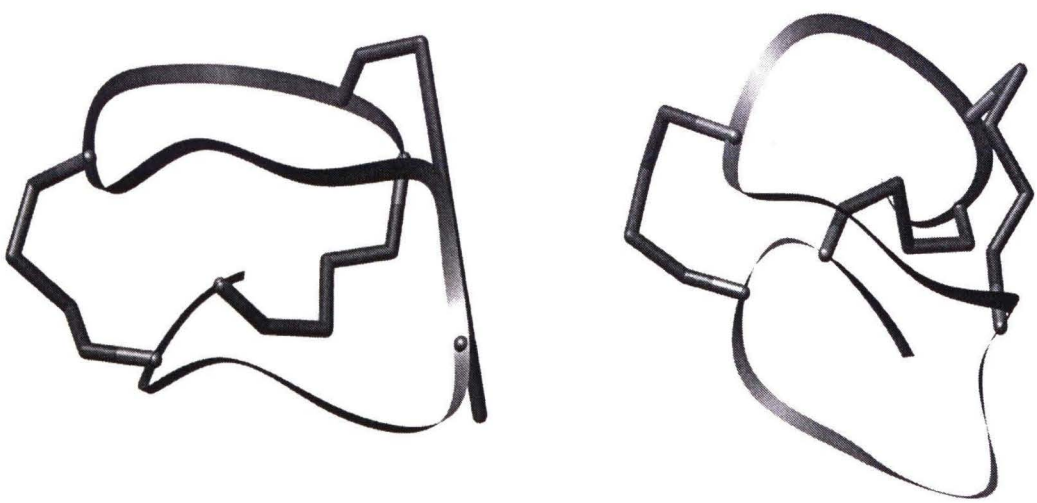


Figure 9. 2D NMR of reg12l at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) wgTOCSY, (B) wgNOESY.

3.6 Molecular model of reg12e.

A model of the six-Cys (1-6, 2-4, 3-5 Cys pairing) reg12e was built using the mr3a conotoxin three-dimensional structure as a template (Figure 10). As expected, the model of 12e has the same overall three-dimensional fold of mr3a; however, notable differences are observed in the precise details of the triple turn structure of these conotoxins.



reg12e:CCTALCSRYHCLPCC

mr3A:GCCGSFACRFGCVOCCV

Figure 10. Model of the lowest energy structure of reg12e based on mr3A. Both belong to the m-2 class.

DISCUSSION

The venom of *Conus regius* is an extremely complex mixture of peptides and proteins whose direct separation is shown in Figure 1. More than 100 fractions can be obtained from this separation. However, most of these fractions show multiple components, for which we were required to adopt an improved separation scheme that includes a prefractionation step using size exclusion chromatography on a Superdex-30 column, followed by a refined peptide-optimized size exclusion step on a Superdex Peptide column. The resulting fractions are then separated by reversed phase on a peptide-optimized C18 Vydac Everest column. Most of the resulting fractions are single-component and were subsequently analyzed by mass spectrometry (MALDI-TOF and ESI-Q-TOF), NMR spectroscopy and peptide sequencing by Edman degradation chemistry. Using this methodology, the most significant components of the venom of *C. regius* can be sequenced (Conopeptidome) and the components of the venom can be grouped in the different families of conopeptides. One of the predominant families of conopeptides found in the venom of *Conus regius* are the mini-M conotoxins.

We isolated twelve mini-M conotoxins from the venom of *C. regius*: reg12a-l. There is no sequence homology among these reg12 mini-M conotoxins except for reg12i and the partially characterized reg12j. When considering the size of their loops (a defining feature within the M-Superfamily), there are many variants within the reg12 conotoxins: 3/4/2, 4/1/1, 4/3/3 and the 4/5/1 subtypes. As with other members of the M-superfamily, reg12a, reg12d and reg12f are hydroxylated at all Pro residues of

their sequences, just as the known maxi-M conotoxins. On the other hand, the Pro residues of reg12e, reg12g and reg12h are unmodified; notably the latter did not exhibit hydroxylation in spite of having four Pro residues with its sequence. reg12f and reg12k are partially hydroxylated as one of their two prolines were found to be a Hyp residue. The sequence diversity found within these reg12 mini-M conotoxins suggests that their targeting of neuronal receptors and/or ion channels might be equally diverse. Perhaps in this case hydroxylation is part of a refinement strategy, where certain sequences within this family are required to have the additional polarity and hydrogen bonding capability, whereas others do not. Differential hydroxylation has been observed within the same conopeptide sequence [33]; this does not appear the case for the reg12 mini-M conotoxins.

Mini-M conotoxins share the same arrangement of Cys in their sequence as other members of the M-superfamily (CC-C-C-CC), such as the Maxi-M or m4 subclass. The loops in the Mini-M are much shorter and the Cys pairing can be different from the classical Cys knot observed in the Maxi-M subclass. Several targets have been identified in the Maxi-M subclass: voltage gated sodium channels (μ-Conotoxins) [34], voltage gated potassium channels (κM-Conotoxins) [35] and the nAChR (ψ-Conotoxins) [31]. Invariably, all of these conotoxins are fully hydroxylated at their Pro residues; to date, all reported maxi-M conotoxins are hydroxylated at the Pro residues of their sequences. In fact, Pro hydroxylation has been assumed to be a conserved feature of this superfamily [35]. While the Mini-M conotoxins are prevalent in many *Conus* species, details of their isolation and characterization are

only starting to be fully disclosed [26]; few reports describe the presence of these toxins in cone snails. Mini-M conotoxins have been isolated from *C. marmoreus*, *C. textile* [25] and *C. betulinus* [36]. These conotoxins are the major components in many mollusk-hunting and worm-hunting *Conus* species. This is the case for *Conus regius* where 12 of the 30 conotoxins isolated from the venom belong to the mini-M subclass. Their molecular target has not been identified; however, it appears to be different from the maxi-M families.

The most abundant mini-M in the venom of *Conus regius* is reg12e. Approximately 32 nmoles of this conotoxin were isolated from the crude venom. Disulfide bond connectivity determination was performed on reg12e. The amounts obtained from the rest of the reg12 conopeptides were not sufficient to carry this study. The mixture obtained from the partial reduction and cyanylation of reg12e contains residual intact peptide and partially reduced/cyanylated isoforms. Besides the main peak that corresponds to the native reg12e, there are other peaks not resolved that correspond to species having 52 Da, and 104 Da greater than the original mass of reg12e (1669.9 Da). This is reflected in the MALDI MS spectra where small fragments corresponding to the cleavage of other fractions are observed. All of these fractions were cleaved with ammonium hydroxide and fractions f and l gave positive matches. Theoretical masses were calculated for fragments based on the cleavage at cyanylated cysteinyl sites. The mass of each fragment is related to the position of the two cyanylated cysteinyl residues. This information was then used to deduce the disulfide bond linkage. Fragments 1596.9 Da and 1077.6 Da correspond to the theoretical

masses of fragments formed from the cleavage of Cys1-Cys6 and Cys2-Cys4 disulfide bonds.

Recently, the NMR structure of a mini-M conotoxin has been published [25]. Although ample time was spent performing NMR experiments on all reg12 conopeptides, the results would only be in accordance to the amounts of peptide originally isolated from the crude venom. In all instances a proton 1D NMR spectrum was obtained (appendix A). Even the lowest amounts (~200 pmoles) of a few reg12 conopeptides resulted in an NMR spectrum after 1000 scans. However, only three would yield TOCSY spectra (reg12e, reg12l, reg12i). Through out this research it has been established that in order to obtain a TOCSY spectrum a minimum of 2-3 nmoles must be present. This is in accordance to our NMR sample preparation technique, where the conopeptide is only dissolved in a total of 40µl of solution. Out of these three conopeptides, only reg12e and reg12l provided NOESY spectra. We believe that by using at least 10 nmoles of peptide a NOESY spectrum can be obtained; reg12e had 32 nmoles and reg12l had 19 nmoles. These amounts were established from the integration of the proton 1D NMR signal of TSP (see materials and methods) in conjunction with peptide sequencing (average of the first three amino acid peaks compared to the standard amino acids in sequencer). These NOESY spectra have enough information to carry out structural analysis on reg12e and reg12l isolated directly from the venom of the cone snail. A typical NMR structure of a synthetic peptide reported in the literature requires around 1-5 µmoles of peptide to obtain structural information. By using the 1.7 mm NMR capillary tube method,

optimized in our laboratory, structural analysis can be performed using significantly less quantities and more importantly by using the native peptide. This technique also allowed us to address post-translational modifications by NMR. The 2D *wg*NOESY spectra of both reg12e and reg12l conotoxins at room temperature showed well defined structures in solution by displaying a considerable number NOE cross-correlations. Although enough NMR data was acquired, a three-dimensional structure of neither reg12e nor reg12l was not attempted; however, this work will be continued at a later stage.

A model of reg12e was built using the NMR structure in solution of mr3A. 10 structures were obtained from MODELLER and the model that provided the lowest energy was used for comparison (Figure 10). Slight differences can be observed amongst the two structures, but the overall fold is very similar. This is expected since both conotoxins belong to m-2 subclass of the M-superfamily. The differences are found mainly at the beginning of the N-terminal and at the end of the C-terminal. mr3A is slightly larger than reg12e and this could account for those differences. The amino acids found between cysteines in both conotoxins are not conserved; therefore, there is no sequence homology except for the cysteine residues. However, the type of amino acids is very similar in each loop. In the first loop both conotoxins posses a hydroxylated amino acid (T,S) and an alanine. In the second loop both conotoxins posses an arginine and an aromatic amino acid (Y, F). In the last loop one contains a proline and the other a hydroxyproline, in essence the same amino acid. Even though the amino acid sequence is different, the nature of the amino acids and the same

disulfide bond connectivity could explain the similar three-dimensional fold results obtained by the modeling of the structure. It is important to mention that cDNA studies should be performed on these mini-M conopeptides from *C. regius* to further establish their relation to the M-superfamily. Likewise, NMR analysis should be performed for the structure elucidation of these peptides.

5. REFERENCES

1. Newcomb, R. and G. Miljanich, *Neurotoxins of cone snail venoms.* Handbook of Neurotoxicology, 2002. **1**: p. 617-651.
2. McIntosh, J.M. and R.M. Jones, *Cone venom - from accidental stings to deliberate injection.* Toxicon, 2001. **39**(10): p. 1447-1451.
3. Myers, R.A., et al., *Conus Peptides as Chemical Probes for Receptors and Ion Channels.* Chemical Reviews, 1993. **93**(5): p. 1923-1936.
4. Adams, D.J., et al., *Conotoxins and their potential pharmaceutical applications.* Drug Development Research, 1999. **46**(3/4): p. 219-234.
5. Olivera, B.M., *Conus venom peptides: Reflections from the biology of clades and species.* Annual Review of Ecology and Systematics, 2002. **33**: p. 25-47.
6. Fegan, D. and D. Andresen, *Conus geographus envenomation.* Lancet, 1997. **349**(9066): p. 1672-1672.
7. Filmer, R.M., *A Catalogue of Nomenclature and taxonomy in the Living Conidae 1758-1998.* 2001, Leiden: Backhuys Publisher.
8. Santos, A.D., et al., *The A-superfamily of conotoxins - Structural and functional divergence.* Journal of Biological Chemistry, 2004. **279**(17): p. 17596-17606.
9. Buczek, O., et al., *Posttranslational amino acid isomerization: A functionally important D-amino acid in an excitatory peptide.* Journal of Biological Chemistry, 2005.
10. Craig, A.G., et al., *Contulakin-G, an O-glycosylated invertebrate neurotensin.* Journal of Biological Chemistry, 1999. **274**(20): p. 13752-13759.

11. Craig, A.G., P. Bandyopadhyay, and B.M. Olivera, *Post-translationally modified neuropeptides from Conus venoms.* European Journal of Biochemistry, 1999. **264**(2): p. 271-275.

12. Craig, A.G., et al., *Enzymatic glycosylation of contulakin-G, a glycopeptide isolated from Conus venom, with a mammalian ppGalNAc- transferase.* Toxicon, 2001. **39**(6): p. 809-815.

13. McIntosh, J.M., A.D. Santos, and B.M. Olivera, *Conus peptides targeted to specific nicotinic acetylcholine receptor subtypes.* Annual Review of Biochemistry, 1999. **68**: p. 59-88.

14. Mari, F. and G. B. Fields, *Conopeptides: Unique Pharmacological Agents that Challenge Current Peptide Methodologies.* Chimica Oggi, 2003. **21**: p. 43-48.

15. Terlau, H. and B.M. Olivera, *Conus venoms: A rich source of novel ion channel-targeted peptides.* Physiological Reviews, 2004. **84**(1): p. 41-68.

16. Yen, T.Y., H. Yan, and B.A. Macher, *Characterizing closely spaced, complex disulfide bond patterns in peptides and proteins by liquid chromatography/electrospray ionization tandem mass spectrometry.* Journal of Mass Spectrometry, 2002. **37**(1): p. 15-30.

17. Jiang Wu, J.T.W., *A novel methodology for assignment of disulfide bond pairings in proteins.* Protein Science, 1997(6): p. 391-398.

18. Smallcombe, S.H., S.L. Patt, and P.A. Keifer, *WET solvent suppression and its applications to LC NMR and high-resolution NMR spectroscopy.* Journal of Magnetic Resonance Series A, 1995. **117**(2): p. 295-303.

19. Larive, C.K., D. Jayawickrama, and L. Orfi, *Quantitative analysis of peptides with NMR spectroscopy.* Applied Spectroscopy, 1997. **51**(10): p. 1531-1536.

20. Piotto, M., V. Saudek, and V. Sklenar, *Gradient-Tailored Excitation for Single-Quantum NMR - Spectroscopy of Aqueous-Solutions.* Journal of Biomolecular NMR, 1992. **2**(6): p. 661-665.

21. Price, W.S., *Water signal suppression in NMR spectroscopy*, in *Annual Reports on NMR Spectroscopy, Vol 38.* 1999. p. 289-354.

22. Wuthrich, K., *NMR of Proteins and Nucleic Acids.* 1986, New York: Wiley & Sons.

23. States, D.J., R.A. Haberkorn, and D.J. Ruben, *A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in 4 Quadrants.* Journal of Magnetic Resonance, 1982. **48**(2): p. 286-292.

24. Fiser, A.S. and A. Sali, *MODELLER: Generation and refinement of homology-based protein structure models.* 2003, Academic Press: San Diego. p. 463-493.

25. McDougal, O.M. and C.D. Poulter, *Three-dimensional structure of the mini-M conotoxin mr3a.* Biochemistry, 2004. **43**(2): p. 425-429.

26. Corpuz, G.P., et al., *Definition of the M-Conotoxin Superfamily: Characterization of Novel Peptides from Molluscivorous Conus Venoms.* Biochemistry, 2005. **44**(22): p. 8176-8186.

27. Aird, S.D., *Chromatographic behavior of Bothrops erythromelas phospholipase and other venom constituents on Superdex 75.* Preparative Biochemistry & Biotechnology, 2004. **34**(4): p. 345-364.

28. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.

29. Cruz, L.J., et al., *Conus Venoms - a Rich Source of Neuroactive Peptides.* Journal of Toxicology-Toxin Reviews, 1985. **4**(2): p. 107-132.

30. Olivera, B.M., J.M. McIntosh, and J. Kulak, *Conotoxins and neuronal calcium channels.* Naunyn-Schmiedebergs Archives of Pharmacology, 1998. **358**(1): p. W14.

31. Shon, K.J., et al., *A noncompetitive peptide inhibitor of the nicotinic acetylcholine receptor from Conus purpurascens venom.* Biochemistry, 1997. **36**(31): p. 9581-9587.

32. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching.* Analytical Chemistry, 1999. **71**(14): p. 2871-2882.

33. Hopkins, C., et al., *A New Family of Conus Peptides Targeted to the Nicotinic Acetylcholine-Receptor.* Journal of Biological Chemistry, 1995. **270**(38): p. 22361-22367.

34. Li, R.A. and G.F. Tomaselli, *Using the deadly μ-conotoxins as probes of voltage-gated sodium channels.* Toxicon, 2004. **44**(2): p. 117-122.

35. Al-Sabi, A., et al., *κM-conotoxin RIIIK, structural and functional novelty in a K+ channel antagonist.* Biochemistry, 2004. **43**(27): p. 8625-8635.

36. Zhao, T.Y., et al., *Purification, sequence and disulfide bonding pattern of a novel conotoxin BtIIIB.* Acta Chimica Sinica, 2005. **63**(2): p. 163-168.

Chapter II

α-Conotoxins from the Venom of *Conus regius*: a Western Atlantic Worm-Hunting Cone Snail Species.

ABSTRACT

We have isolated six novel α4/3 conotoxins and one novel α4/7 conotoxin from the venom of *Conus regius*, a widespread worm-hunting cone snail species of the Western Atlantic Ocean. Five of these α4/3 conotoxins, reg1a, reg1b, reg1d, reg1e and reg1f, have sequence homology with the previously reported α4/3 conotoxins from *Conus imperialis* ImI, a potent inhibitor of the α7 nAChR and a previously reported α4/3-conotoxin sequence translated from the cDNA library of *Conus regius* (α-RgIA). Three of these conotoxins are hydroxylated at the conserved Pro-6 and two keep the conserved Pro-6 while all have positively charged amino acids in the second intercystine loop. The other α4/3 conotoxin, reg1c, is unusual as it lacks the conserved Ser-Asp diad (replaced by Arg-Arg diad) and the positively charged amino acids in the second intercystine loop. reg1c has no sequence homology with any other known α-conotoxin. The α4/7 conotoxin, reg2a, has remarkable sequence homology the previously reported α4/7 conotoxins GIC and GID from *Conus geographus*, the

quintessential fish-hunting cone snail from the Indo-Pacific region. The reg1 and reg2 conotoxins were isolated from the crude venom by using a combination of the SE and RP HPLC chromatography. The resulting nanomolar quantities of conopeptides were directly characterized by 2D-NMR methods (*wg*TOCSY). Subsequently, sequencing by Edman degradation and MALDI Q-TOF MS/MS revealed the two sets of α4/3-conotoxins and an α4/7 conotoxin. The hydroxylation and replacement of the Ser-Asp diad (reg1c) of the reg1 conopeptides is unusual and represents the first examples of this modification within the α4/3-conotoxin sub-family. This can only be observed by studying the components directly from the venom of *Conus regius*. Recently, the sequence of an α4/3-conotoxin from *Conus regius* was determined from a cDNA clone. This sequence is similar to reg1d, in fact we believe α-RgIA is reg1d without post-translational modifications. This shows that it is essential to study the venom components directly from the venom and not from cDNA libraries, as one can overlook important post-translational modifications and place amino acids to the sequence that should not be there.

1. INTRODUCTION

Acetylcholine targets the nicotinic acetylcholine receptor to elicit an excitatory transmission. Anything that affects the nicotinic transmission could possibly be used for treatment of pain, memory disorders and other neuronal disorders. Neuronal nAChRs are pentameric proteins belonging to the major ligand-gated ion channel superfamily [1, 2]. Nicotinic acetylcholine receptors (nAChRs) are involved in rapid gating of ions elicited by acetylcholine. The nAChRs are homo- or heteromeric pentamers of structurally related subunits that encompass an extracellular N-terminal ligand-binding domain (LBD), four transmembrane-spanning regions that form the ion channel, and an extended intracellular region between spans M3 and M4 [3]. They exist in at least three conformational states with distinctive sensitivities to the nicotinic ligands that dictate channel gating and function: basal or resting (closed, but it could be rapidly activated), activated (open), and desensitized (closed). Indeed, ligand binding triggers conformational changes that are transmitted to the transmembrane-spanning region, leading to gating and changes in membrane potential [3]. The ligand binding sites are located at the interface of two subunits. To effectively understand how the nAChRs work different probes are needed to distinguish the different subunits present in the nAChR. Recently, the X-ray structure of α-conotoxin ImI bound to the AChBP (homologue of the nAChR extracellular domain) was reported [3]. This work reveals that ImI is buried in the equivalent of the α-subunit (Loop C) and that minor conformational changes are observed in the structure of the bound ImI when compared to the free ImI (Figure 1).

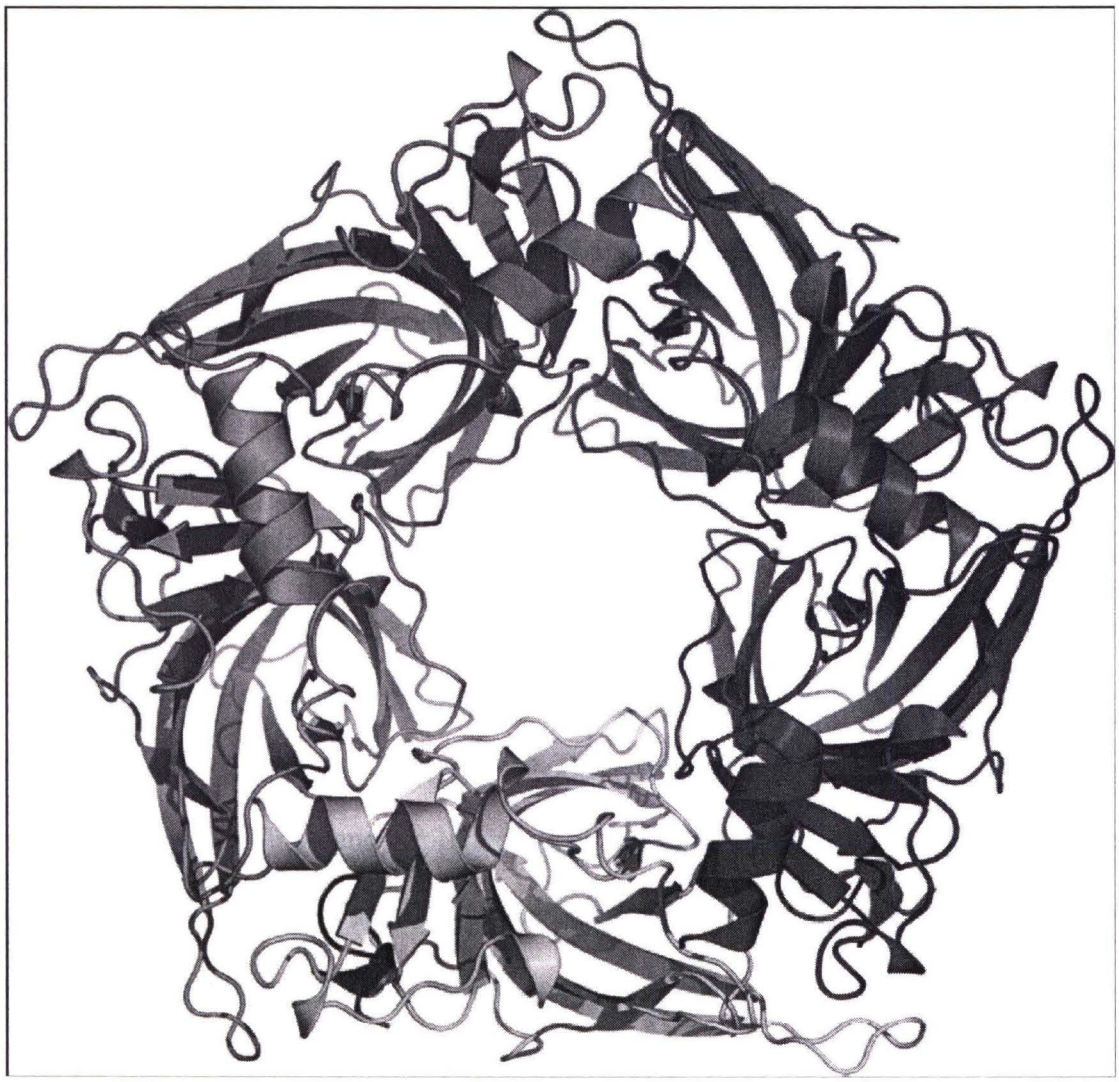


Figure 1. Crystal structure of α-conotoxin ImI bound to Ac-AChBP viewed along the fivefold axis. Conotoxins are in red.

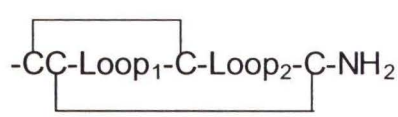
Several forms of disulphide-constrained conopeptides are known to target the different types of nicotinic acetylcholine receptors (nAChRs) [4]. Cone snails utilize nAChRs antagonists to immobilize and capture their prey. α-Conotoxins are the best-studied nAChR-specific conopeptide probes. Their distribution in *Conus* venom varies significantly, depending upon the prey preference of a particular species of cone snail, which consist of either fish (piscivorous), mollusks (molluscivorous) or worms (vermivorous). Conopeptide precursors are ribosomally-expressed proteins that subsequently undergo post-translational modifications and proteolytic cleavage to form the mature conotoxin [5]. Conopeptides inherently contain high degrees of modified amino acids (usually combinations of them), such as cystines, hydroxyproline, γ-carboxyglutamate, Br-Trp, D-Trp, D-Leu, D-Phe[6], pyro-Glu, glycosylated Ser/Thr, and sulfated Tyr [7-9]. These modifications confer conopeptides with unique stability and exquisite specificity towards neuronal targets [8, 10, 11], enabling cone snails to capture prey.

α-conotoxins of the 3/5 subtype, such as GI, MI and SI are found only in fish-hunting *Conus*. α3/5 conotoxins bind to the muscular type nAChR imparting paralytic characteristics to the cone snail venom by inhibiting neuromuscular transmission [12]. All other of α-conotoxins, the related 4/3 and 4/7 subtypes, bind to the neuronal type nAChR. Notably, the only exception is the α4/7-Conotoxin EI, isolated from *C. ermineus*, the only known fish-hunting species of the Atlantic Ocean. EI targets the muscular nAChR; however, its sequence differs significantly to all other α4/7-Conotoxins. To date, α4/3 conotoxins have only been found in worm hunting species

(*C. imperialis and C. regius*); whereas α4/7 conotoxins have been found in all types *Conus* (piscivorous, such as GID and GIC in *C. geographus*; molluscivorous such as Vc1.1 in *C. victorae* and vermivorous such as AnIC in *C. anemone*).

The α4/7-conotoxin MII was the first described conopeptide inhibitor of the neuronal nAChR and it was found to specifically bind to the α3β2 subtype. Since then, the binding specificity of other α4/7 conotoxins to several subtypes of the neuronal nAChR has been studied. Neuronal nAChRs have been implicated in wide variety of synaptic dysfunctions [13] and several neuropharmacological applications have been outlined for modulators of this ligand-gated ion channel[14].

Structurally, α-conotoxins are single polypeptide chain less than twenty amino acids long. The main chain is restricted by two disulfide bonds, which constraints the peptide into a "folded" scaffold in spite of its very small size. The covalent structure of α-conotoxins can be schematically represented by



The disulfide bonds confine the covalent structure to a bicyclic moiety with Loop1 as a common segment between the cyclic rings. The length of the first and second loops determines the particular subclass of α-conotoxin [12]. The best characterized α4/3-conotoxin, isolated from *Conus imperialis*, the α-conotoxin ImI (ACCSDPRCAWRC*), and several structures of this particular conopeptide have

reported. The best characterized α4/7-conotoxin is MII isolated from *Conus magus* (GCCSNPVCHLEHSNLC*) and the structure of this particular conopeptide has been reported when in aqueous solutions and in TFE mixtures [15]. The structure of MII is characterized by a helical loop at the central part of sequence that is stabilized by the ^{1}C-^{3}C disulfide bond and turns involving ^{1}C-^{2}C. Some regions of the peptide appear to be flexible [15]. The overall fold of MII is similar to that of other α4/7 conotoxins (PnIA/B, EI, EpI, GIC and GID). However, all of these α4/7 conotoxins exhibit differential targeting towards the nAChR. The differences in selectivity can be attributed to differences in the surface charge distribution among these α4/7 conotoxins [16]. One amino acid difference in the sequence of α-conotoxins is known to produce a selectivity switch of sites within same nAChR subtype [17] or a change of binding selectivity towards a subtype [18].

As part of our efforts to obtain the conopeptidome of *Conus regius,* a widespread worm-hunting cone snail of the Western Atlantic Ocean, we took on the isolation and characterization of α-conotoxins from *Conus regius.* In this chapter, we present the results from the isolation and characterization of seven novel α-conotoxins. A 16-residue α4/7-conotoxin from *Conus regius* (species code: reg) reg2a and six α4/3 conotoxins (reg1a-e') also from *Conus regius.* The sequence of reg2a is 81% homologous with GIC and 94% homologous with GID; the latter two isolated from *C. geographus.* The sequences of reg1a-e' with the exception of reg1c display the typical characteristics of most α4/3 conotoxins, with the important difference that some present hydroxylation at the conserved residue Pro-6. reg1c is the most striking

α4/3 conotoxin from *C. regius* since it does not follow any of the traditional features for this family. The sequences of reg2a α4/7-conotoxin and reg1a-conotoxins were modeled against the NMR structures of conotoxins α4/7 GIC and α4/3 ImI. The models obtained were compared with structures of other α4/7 and α4/3 conotoxins.

2. MATERIALS AND METHODS

2.1 Specimen collection.

Refer to materials & methods Chapter I.

2.2 Crude venom extraction.

Refer to materials & methods Chapter I.

2.3 Peptide purification.

Refer to materials & methods Chapter I.

2.4 Reduction and alkylation of cysteyl residues.

Refer to materials & methods Chapter I.

2.5 Peptide sequencing.

Refer to materials & methods Chapter I.

2.6 Molecular mass determination.

Refer to materials & methods Chapter I.

2.7 Expression of neuronal nAChRs and eletrophysiological methods for reg2a.

In collaboration with Dr. Charles Luetje from the University of Miami, neuronal α3β2 nAChRs were expressed in frogs (*X. laevis)* oocytes, see reference [19] for details. Current responses were measured for reg2a, reg1b and reg1d (100nM) under a two-

electrode voltage clamp, at a holding potential of -70 mV, using an OC-725C voltage clamp unit (Warner Instruments, Handem, CT). Oocytes were perfused at room temperature. Perfusion was continuous (except during toxin incubations 5 min) at a rate of 2 ml/min. ACh (70 uM) was applied diluted in perfusion solution. The extent of receptor blockade was determined by comparing the ACh induced peak current response following 5 minute incubation with toxin, to the average of three ACh induced peak current responses preceding the toxin incubation. Refer to [19] materials and method section for more detailed information.

2.8 NMR spectroscopy.

NMR spectra were acquired on a Varian Inova 500 MHz instrument equipped with PFG, 3xRF channels and waveform generators. Nanomolar quantities of the native conopeptides directly isolated from the venom (reg2a = 7 nmoles, reg1c = 6 nmoles, the rest of the α-conotoxin peptides were below 1 nmole) were dissolved in 40 µl of water with 10% D_2O (used for locking purposes) and 4 nanomoles of TSP and placed in 1.7 mm NMR tubes (Wilmad WG-1364-1.7). For further details on this procedure refer to Materials & Methods Chapter I.

2.9 Molecular Model of reg2a and reg1d. Molecular models were built by comparative modeling methods [20] based on the NMR structure of the GIC and ImI conotoxins as template and using Modeller (version 8.0). Briefly, conotoxin sequences were aligned according to the standard routine in the program using the

PDB entry as template. A set of 10 model structures was built accordingly for each and the selected structures were that of better target Modeller energy.

2.10 Nomenclature.

In this chapter we adopted a nomenclature of three letters to designate *Conus* species, because the one or two letter nomenclature currently in place will not be enough to describe the large number of different non-fish-hunting species, especially those with similar first letter names. We decided to use the three letters "reg" to name the peptides from *C. regius*. Arabic numbers were used to represent the disulfide framework; 1 and 2 have already been assigned to the alpha conotoxins. The letter after the framework number indicates the order of elution on RP-HPLC. For the α4/3 conotoxins the code reg1 has been assigned, and for the α4/7 conotoxin the code reg2 has been assigned.

3. RESULTS

3.1 Peptide purification.

The venom of *Conus regius* is an extremely complex mixture of peptides and proteins whose direct separation is shown in Figure 1B, Chapter I. More than 100 fractions can be obtained from this separation. However, most of these fractions show multiple components, for which we were required to adopt an improved separation scheme that includes a prefractionation step using size exclusion chromatography on a Superdex-30 column (Figure 2), followed by a refined peptide-optimized size exclusion step on a Superdex Peptide column (Figure 3). The resulting fractions are then separated by reversed phase on a peptide-optimized C18 Vydac Everest column. Most of the resulting fractions are single-component (Figure 4 and 5) and were subsequently analyzed by mass spectrometry (MALDI-TOF and ESI-Q-TOF), NMR spectroscopy and peptide sequencing by Edman degradation chemistry. Using this methodology, the complete analysis of the most significant components of the venom of *C. regius* can be sequenced (conopeptidome) and the components of the venom can be grouped in the different families of conopeptides. The reg1 and 2 peptides were separated by the Superdex-30 in spite of their similar size. This is not entirely unexpected, since this column is also known to partition analytes by hydrophobic interactions [21].

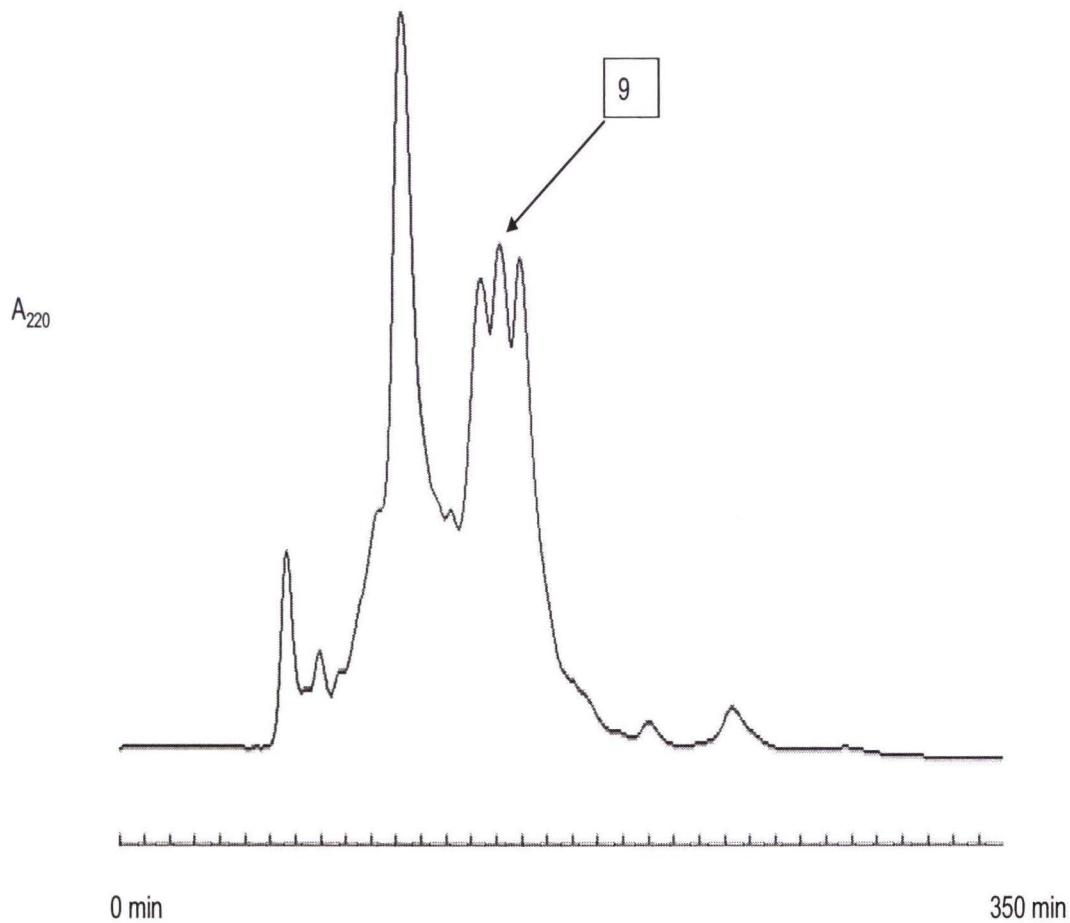


Figure 2. Size exclusion chromatogram of *C. regius* dissected venom. 60 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 9 was chosen for further separation.

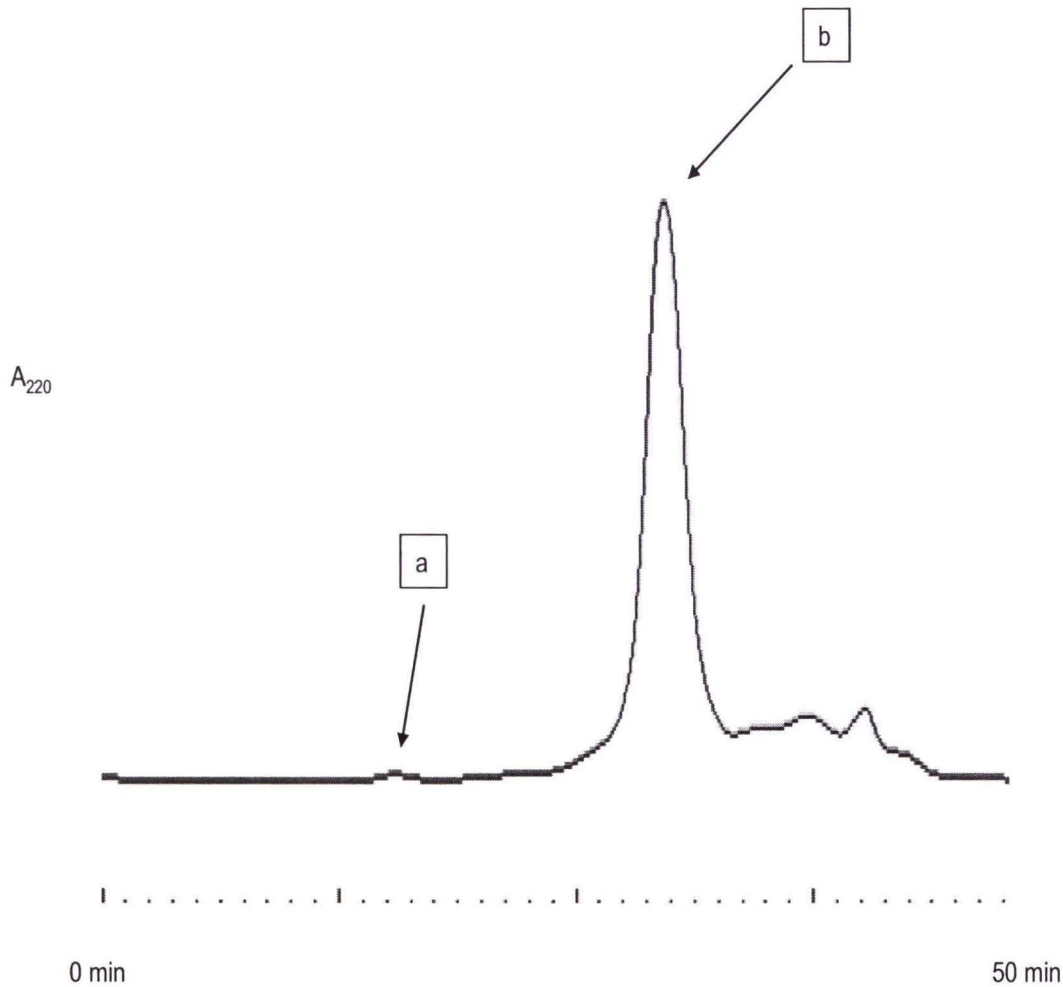


Figure 3. Size exclusion chromatogram of *C. regius* fraction 9 from regG dissected venom in a Superdex Peptide column using a 0.1 M ammonium bicarbonate solution. Fraction b was chosen for further separation.

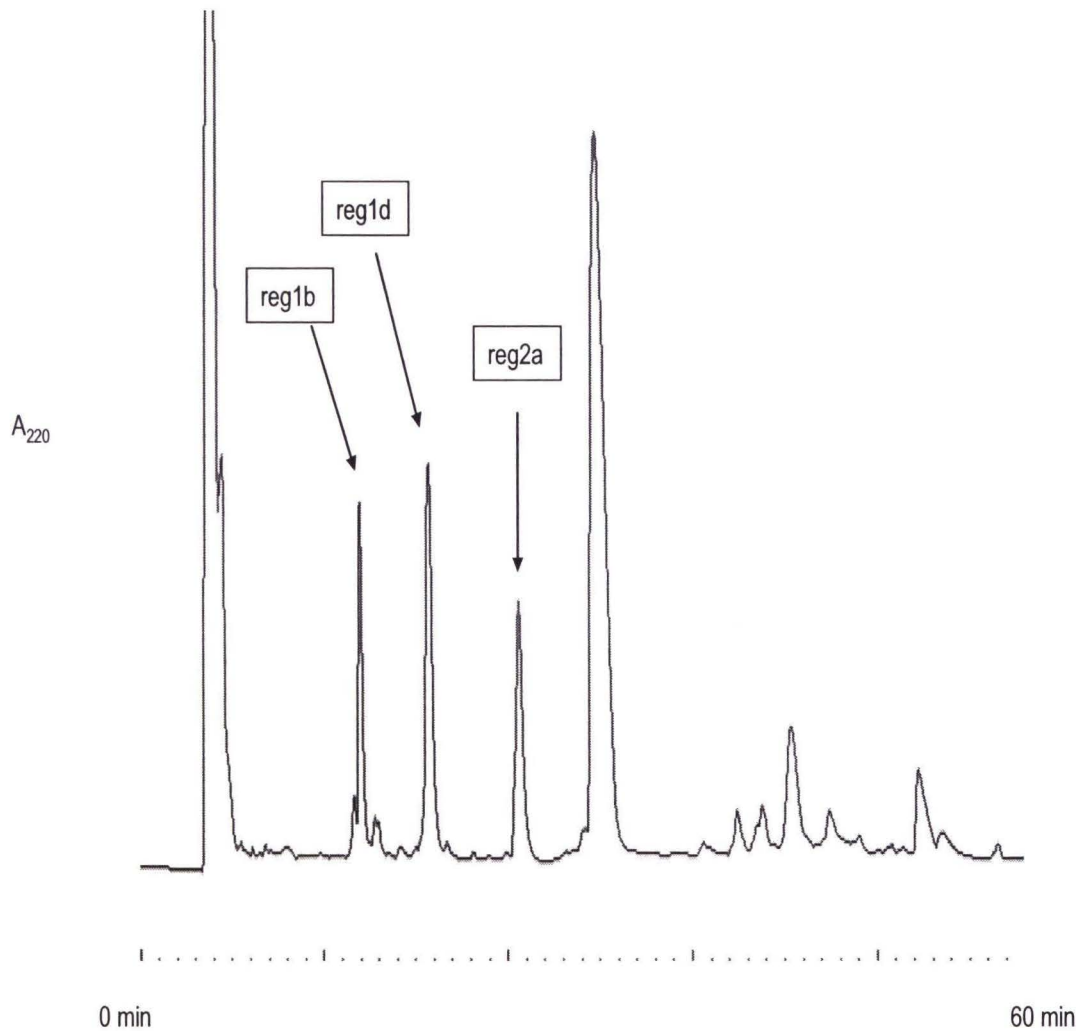


Figure 4. Semi-preparative RP-HPLC chromatogram showing the location of reg1b, reg1d and reg2a peptides from fraction 9 SEC. A linear gradient from 0 to 100% solution B in 100 min at 1ml/min was employed. HPLC solutions were: 0.1% (v/v) TFA in water (A) and 0.1% (v/v) TFA in 60% (v/v) aqueous MeCN (B).

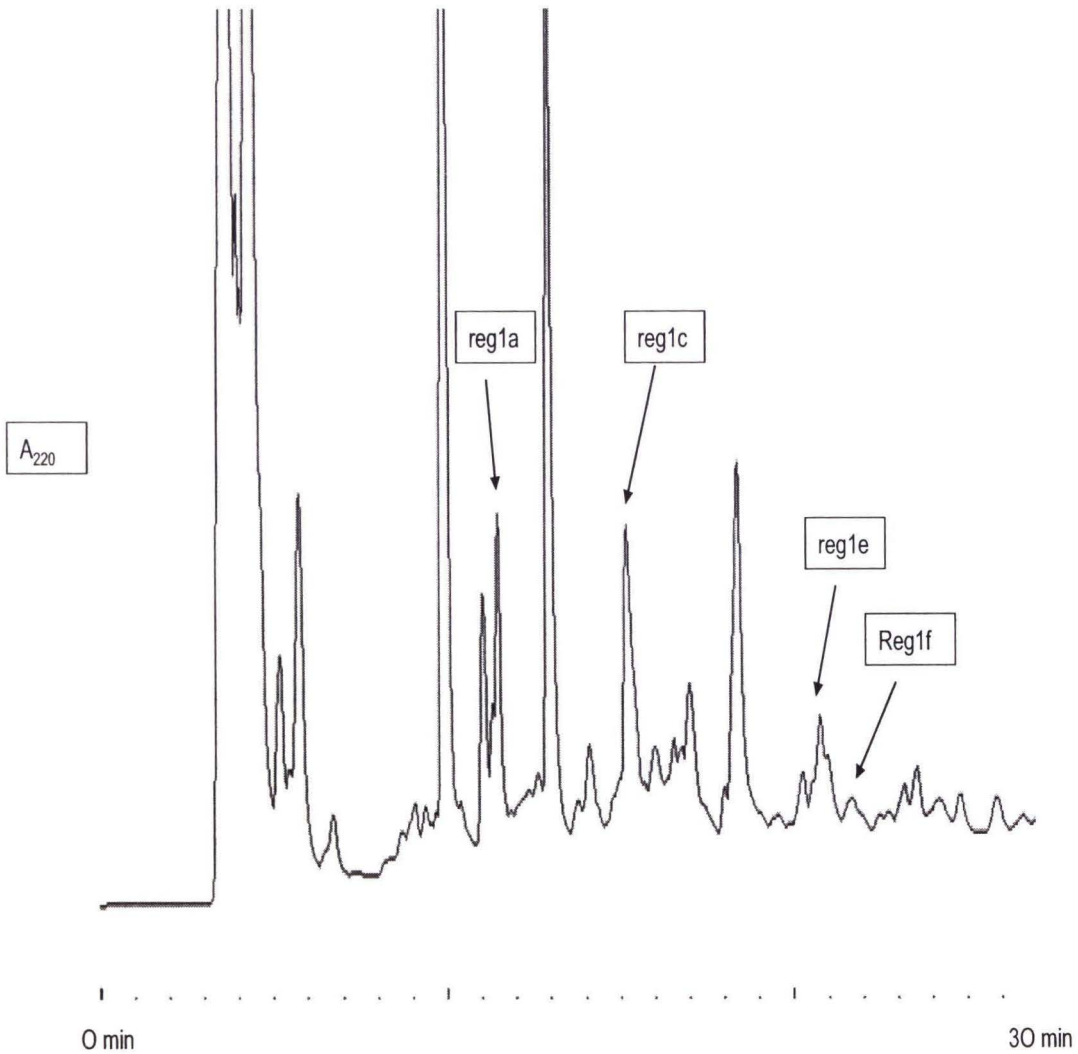


Figure 5. Semi-preparative RP-HPLC chromatogram showing the location of reg1a, reg1c, reg1e and reg1e' peptides. A linear gradient from 0 to 100% solution B in 100 min at 1ml/min was employed. HPLC solutions were: 0.1% (v/v) TFA in water (A) and 0.1% (v/v) TFA in 60% (v/v) aqueous MeCN (B).

3.2 Reduction/Alkylation and peptide sequence determination.

The purified peaks were subjected to reduction with DTT and alkylation with iodoacetamide. Mass spectrometry of the reduced/carboxymethylated peptides and of native peptides showed a mass difference consistent with the presence of four cysteine residues for each peptide. The reduced/alkylated conotoxins were sequenced to completion by Edman degradation. The sequences of the reg1 and 2 conotoxins are shown in Table 1. It is worth noticing that the differences between peptides reg1e and reg1e' is just one amino acid at position 11, where a glutamine is substituted by a glutamate. reg1b differs from reg1e in that the proline at position 6 is hydroxylated. reg1d and reg1a differ by only one amino acid at position 11. In reg1a the amino acid at position 11 has not been identified (ΔM = 175.5 Da), while in reg1d position 11 is filled with an Arg residue. Hydroxyprolines were present in sequences reg1a-d. This residue was evident by presence of two new peaks in the same Edman degradation cycle, corresponding to the different isomers of hydroxyproline. These conotoxins contained between 12 and 17 residues and the four cysteines were separated by two loops of variable sizes (loop1 = 4 amino acids, loop2 = 3 - 7 amino acids). Additionally, a BLAST search [22] of the databases (Swissprot/EMBL, PIR, PDB and nrdb95) did show some significant sequence homology to reported proteins and peptides, except for reg1c that had no matches. The sequences of α-conotoxins from *Conus regius* are reported in Table 1.

α-4/3	Sequence	Target
reg1a	--G**CC**SDOR**C**RYX**C***	
reg1b	--G**CC**SDOR**C**KHQ**C***	
reg1c	-DY**CC**RROO**C**TLI**C***	
reg1d	--G**CC**SDOR**C**RYR**C***	
reg1e	--G**CC**SDPR**C**KHQ**C***	
reg1f	--G**CC**SDPR**C**KHE**C***	
ImI	--A**CC**SDPR**C**AWR**C***	α7
ImII	--G**CC**SDRR**C**AWR**C***	α7
Bn1.3 (cDNA)	-DY**CC**HRGP**C**MVW**C***	
regIA (cDNA)	--G**CC**SDPR**C**RYR**C**R	α9, α10

α-4/7	Sequence	Target
reg2a	--G**CC**SHPA**C**NVNNPHI**C***	α3β2
GIC	--G**CC**SHPA**C**AGNNQHI**C***	α3β2
GID	- IRDγ**CC**SNPA**C**RVNNOHV**C**	α3β2, α7 >α4β2
n1.1 (cDNA)	--G**CC**SHPA**C**SVNNPDI**C***	
Bn1.2(cDNA)	--E**CC**THPA**C**HVSHPEL**C***	
Vc1.1	G**CC**SDPR**C**NYDHPEI**C***	α3α7β4/α3α5β4
vc1a	G**CC**SDOR**C**NYDHPγI**C***	?
EpI	--G**CC**SDPR**C**NMNNPDŶ**C***	α3β2, α3β4; α7
PnIA	--G**CC**SLPP**C**AANNPDŶ**C***	α3β2 > α7
AuIA	--G**CC**SYPP**C**FATNSDŶ**C***	α3β4
PIA	RDP**CC**SNPV**C**TVHNPQI**C***	α6β2β3
MII	--G**CC**SNPV**C**HLEHSNL**C***	α3β2
AnIC	GG**CC**SHPA**C**AANNQDŶ**C***	α3β2, α7
AnIC	GG**CC**SHPA**C**AANNQDŶ**C***	α3β2, α7
EI	RD**OCC**YHPT**C**NMSNPQI**C***	α1γ= α1δβ2 musc.

Ŷ indicates sulfation of Tyrosine, γ indicates Gla

Table 1. reg1 and 2 sequences from *Conus regius* and other α-conotoxins.

3.3 Mass spectrometry of purified peptides.

Mass spectrometry carried out using MALDI-TOF in the reflector mode (M/ΔM resolution ~ 10,000) yielded the following monoisotopic molecular ions: reg1a = 1449.6 Da, reg1b = 1348.8 Da, reg1c = 1571.9, reg1d = 1432.0 Da, reg1e = 1332.5 Da, reg1e' = 1333.6 Da, reg2a = 1664.8 Da (Figures 6-9). Mass analysis of the reduced/carboxymethylated peptides and the native peptides showed a mass difference consistent with the presence of four cysteine residues in each peptide. The masses obtained for the peptides were for the most part in agreement with the calculated theoretical monoisotopic values determined for the assigned sequences. reg1a did not produce a full sequence since amino acid residue at position 11 could not be identified from the sequencing cycle. Repeated attempts were carried out with no success. The difference in the calculated molecular weights and the experimental molecular weights from MALDI-TOF MS for the peptides in sequence agreement resulted in all seven peptides amidated at the C-terminus. The calculated molecular weights (MWcal) were obtained using Protein Prospector [23].

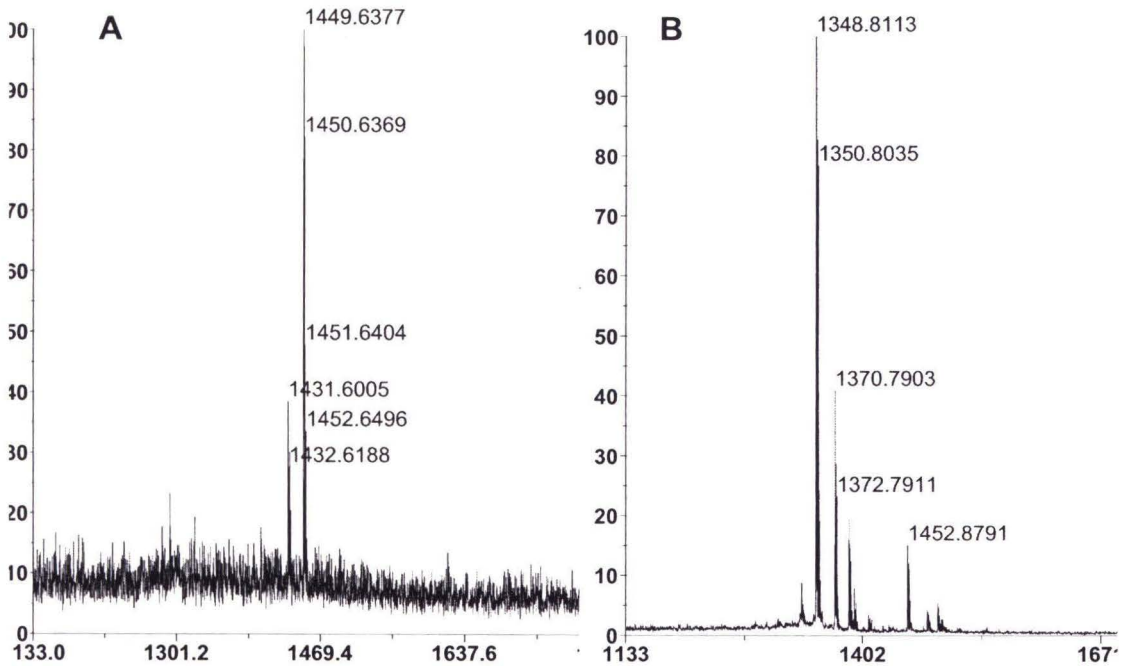


Figure 6. MALDI-TOF mass spectra for reg1a (A) and reg1b (B). Spectra is from initial RP-HPLC separation of these conotoxins as evident by the presence of other ions.

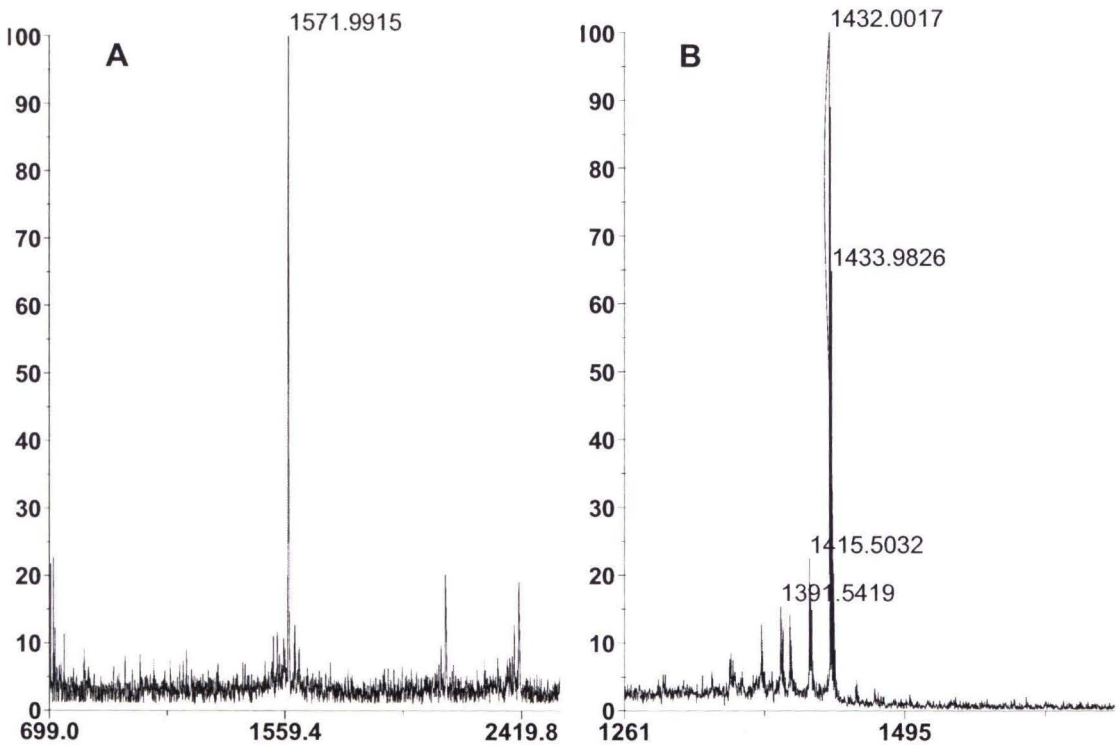


Figure 7. MALDI-TOF mass spectra for reg1c (A) and reg1d (B). Spectra are from initial RP-HPLC separations of these conotoxins as evident by the presence of other ions.

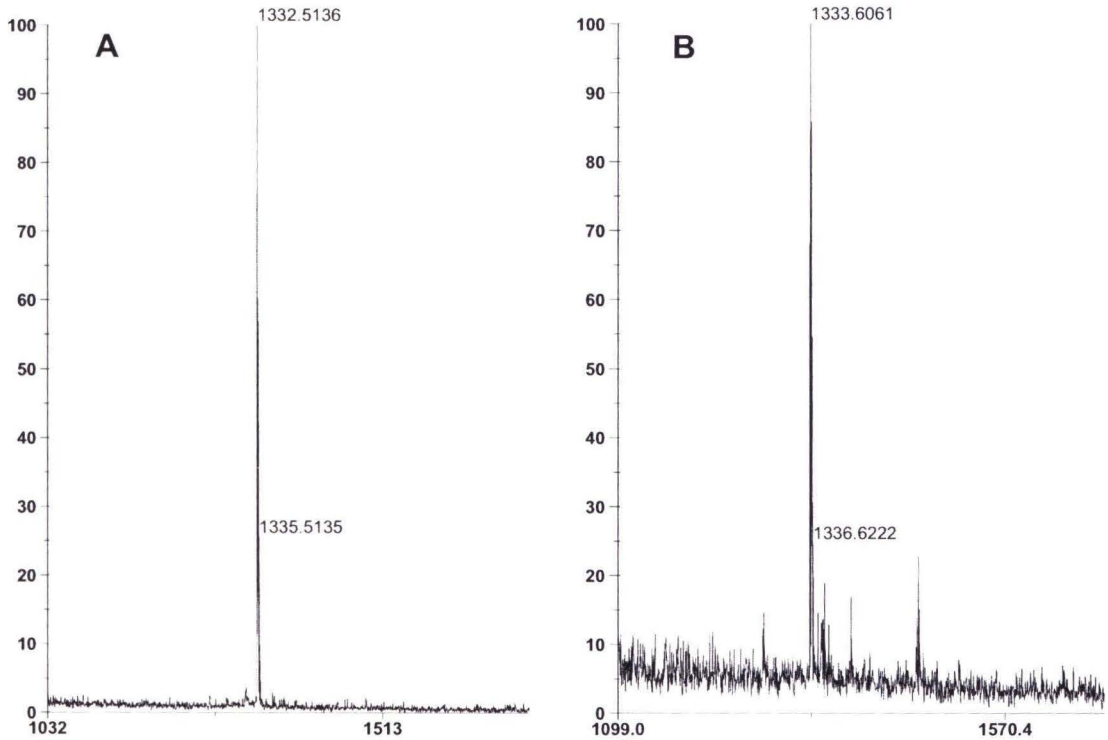


Figure 8. MALDI-TOF mass spectra for reg1e (A) and reg1f (B). Spectra are from initial RP-HPLC separations of these conotoxins as evident by the presence of other ions.

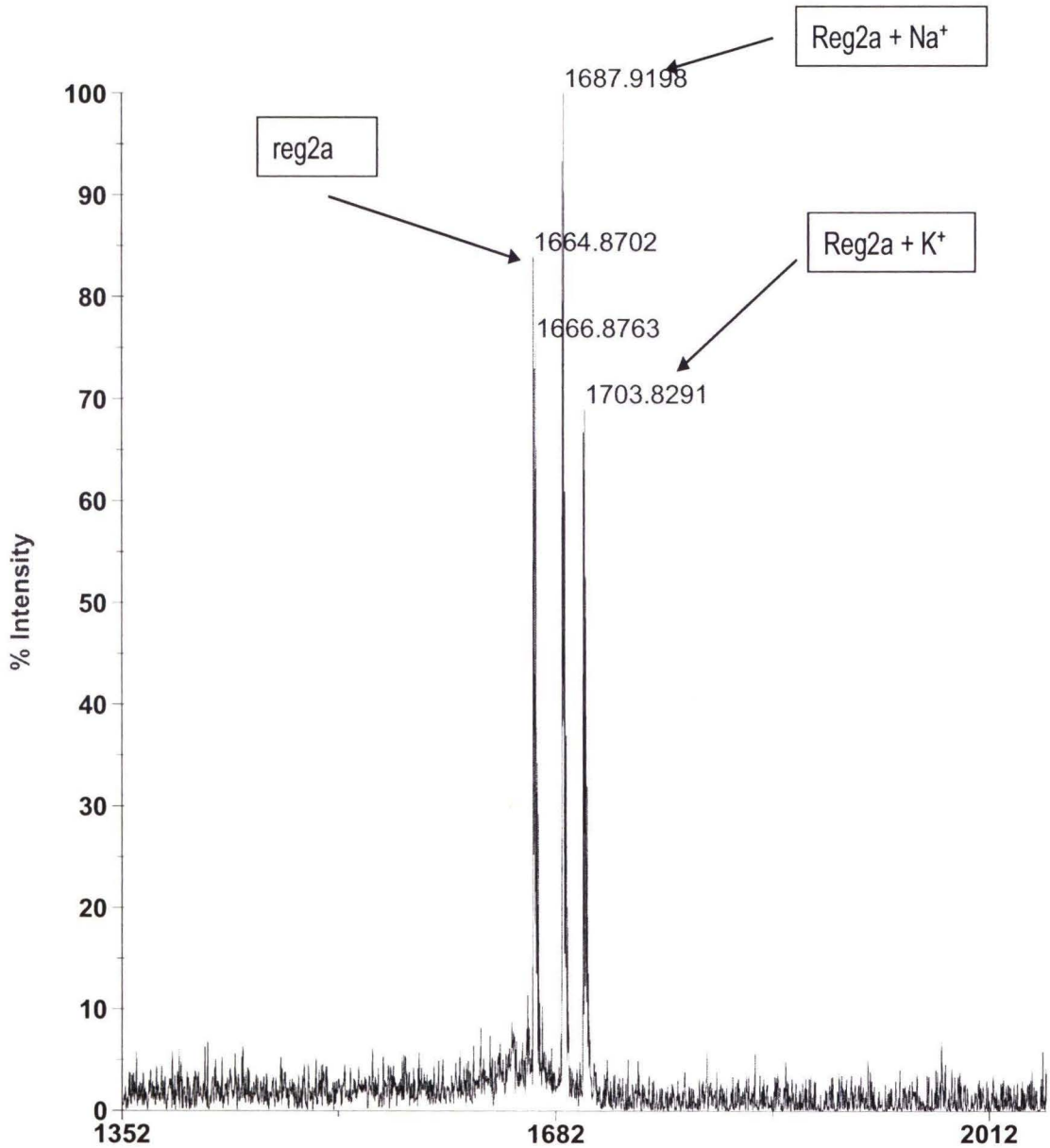


Figure 9. MALDI-TOF mass spectra for reg2a. Two more peaks are present in the spectra and correspond to the native peptide reg2a and addition of Na^+ and K^+ ions.

3.4 Electrophysiology results of reg2a, reg1b and reg1d.

100 nM of reg2a, reg1b and reg1d were tested against the α3β2 nAChRs expressed in frogs (*X. laevis)* oocytes. 100 nM of toxin was used in combination with 5 minutes of incubation. reg2a blocked the response of α3β2 nAChRs to 70 μM acetylcholine (EC50) by about 85% (Figure 10). This is expected of reg2a as it is an α4/7 with high sequence homology to GIC and GID. These two conotoxins have been previously reported to block the neuronal α3β2 nAChRs. reg1b and reg1d had no effect on α3β2 nAChRs (Figure 11). These conotoxins belong to the α4/3 class of α-conotoxins and have high sequence homology to the recently reported regIA which is known to block the α9 and α10 nAChRs; therefore, these results were expected. At the time of this assay the information for rgIA was not available, we were then inclined to believe that reg1b and reg1d would block the α7 nAChRs instead of the α3β2 nAChRs. Due to sample quantity limitations a detailed study was not carried out on the target receptor. Based on the information obtained an estimation of the IC50 for reg2a would be 10-30 nM, similar to that of PnIA and PnIB from *Conus pennaceus*. All the electrophysiological data is presented in Table 2.

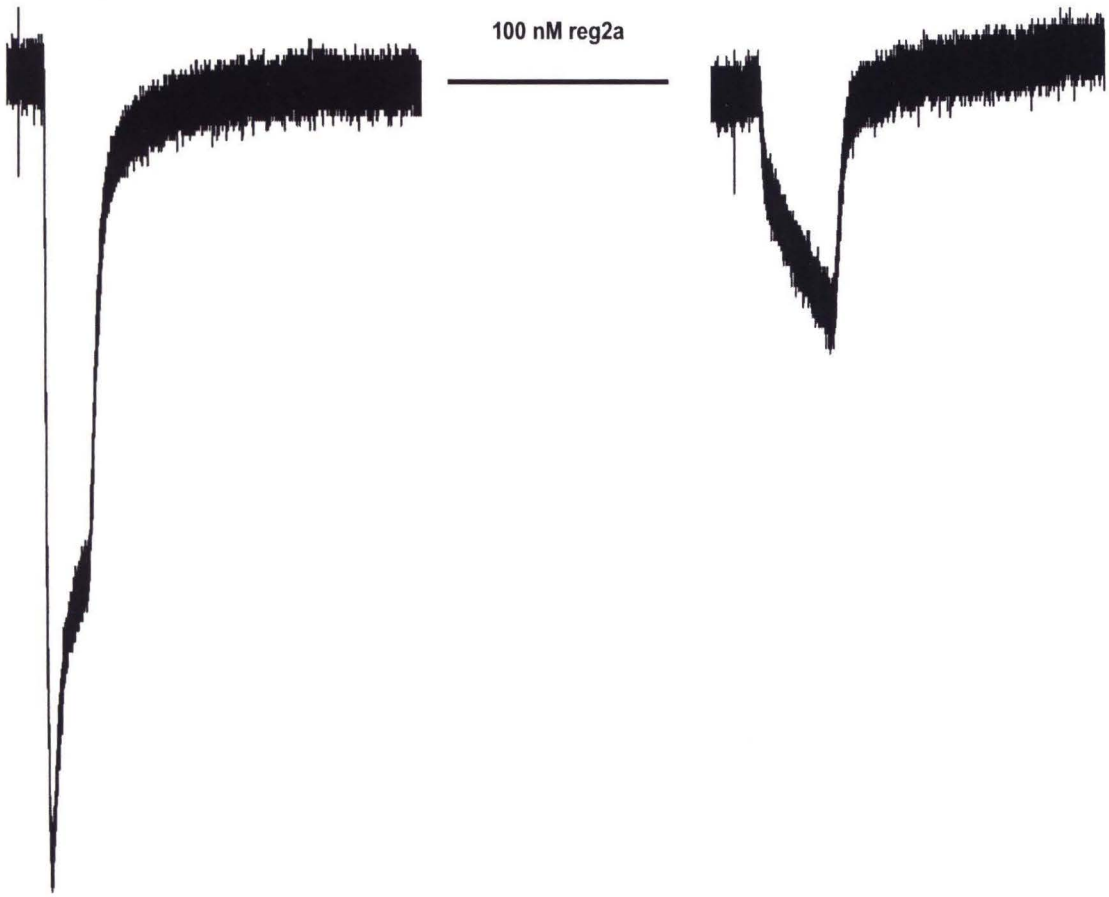


Figure 10. Current responses of *Xenopus* oocytes expressing α3β2 neuronal nicotinic receptors to 70 μM Ach before and after 5 min incubation with reg2a
(scale: y = 100 nA, x = 20s)

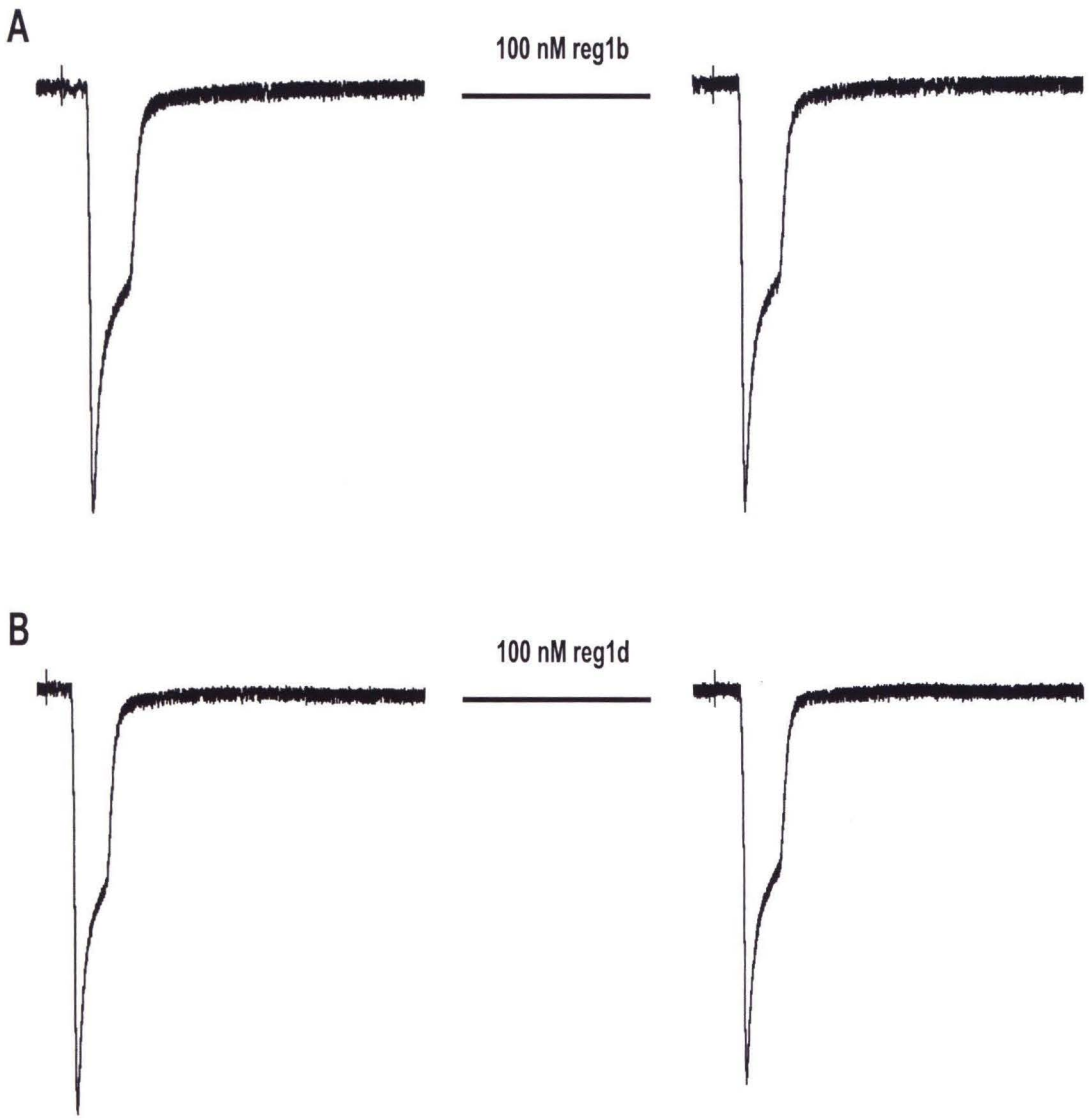


Figure 11. Current responses of *Xenopus* oocytes expressing α3β2 neuronal nicotinic receptors to 70 μM Ach before and after 5 min incubation with (A) reg1b (scale: y = 250 nA, x = 20 s), (B) reg1d (scale: y = 250 nA, x = 20s).

ACh applied at 70 μM (EC_{50} for α3β2) for ~15 sec

Receptor	Application	Trace#	Current (nA)	Avg	% remaining
α3β2	70 μM ACh	5	86		
	70 μM ACh	6	89		
	70 μM ACh	7	89		
	70 μM ACh	8	99	90.75	
	100 nM reg2a + 100 μg/ml BSA 5 min incubation				
	70 μM ACh	9	14		15%
α3β2	70 μM ACh	17	226		
	70 μM ACh	18	196		
	70 μM ACh	19	244	222	
	100 nM reg1b + 100 μg/ml BSA 5 min incubation				
	70 μM ACh	20	257		116%
α3β2	70 μM ACh	21	250		
	70 μM ACh	22	280	265	
	100 nM reg1d + 100 μg/ml BSA 5 min incubation				
	70 μM ACh	23	255		96%

Table 2. Electrophysiological data from the current responses of *Xenopus* oocytes expressing α3β2 neuronal nicotinic receptors to reg2a, reg1b and reg1d.

3.5 NMR spectroscopy.

We were able to obtain ^{1}H NMR spectra of conotoxins reg2a, reg1b, reg1c, reg1d and 2D spectra of nanomolar quantities (nanoNMR) of reg2a and reg1c conotoxins directly isolated from the venom of the cone snails.2D wgTOCSY spectra were obtained for reg2a and reg1c. Figure 12 shows the 1D NMR of reg2a at 25 °C (See Appendix A for the rest of the spectra). The 1D of reg2a shows the TSP peak at zero ppm, this is the equivalent to 4 nmoles of TSP in protons which indicates that there was enough concentration to try a 2D experiment. There seems to be some impurities in the sample (peaks around 7 ppm and 8.6 ppm). Figure 13 represents the *wg*TOCSY spectra of reg2a and reg1c at 25°C. In spite of the best efforts NOESY spectra were not obtained due to concentration limitations.

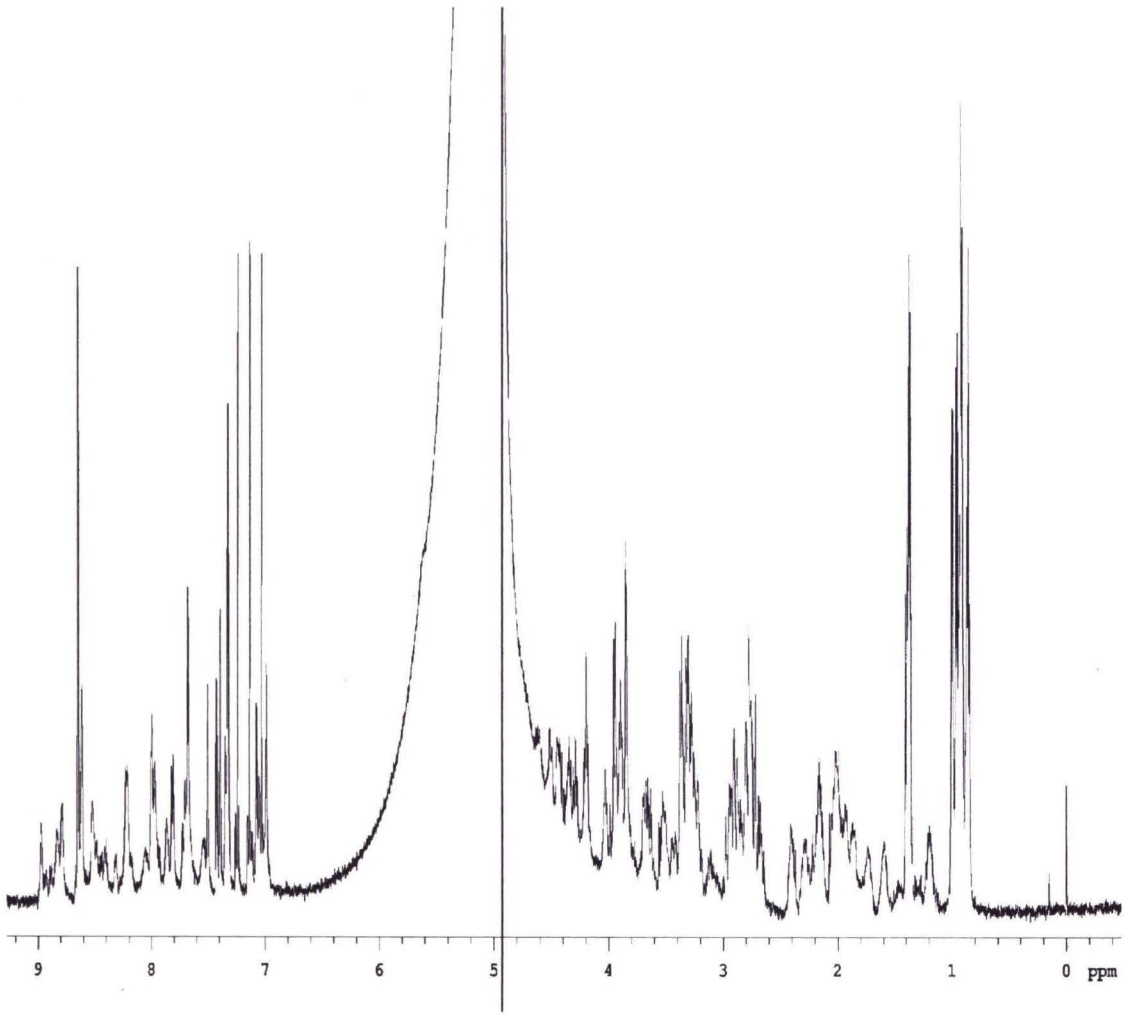


Figure 12. Proton 1D NMR spectrum at 25°C using a 3 mm NMR tube in 5 mm gHCN probe of reg2a.

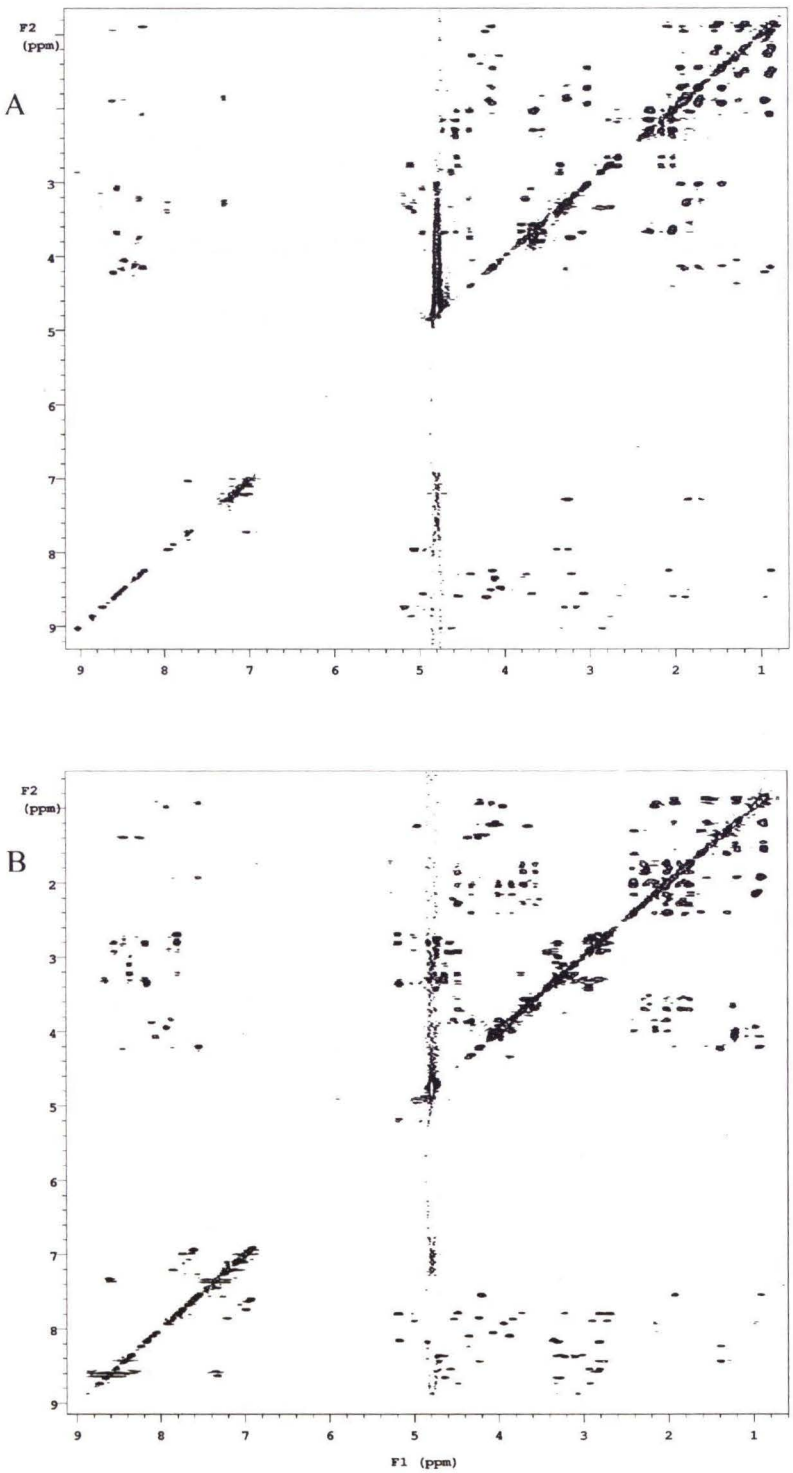
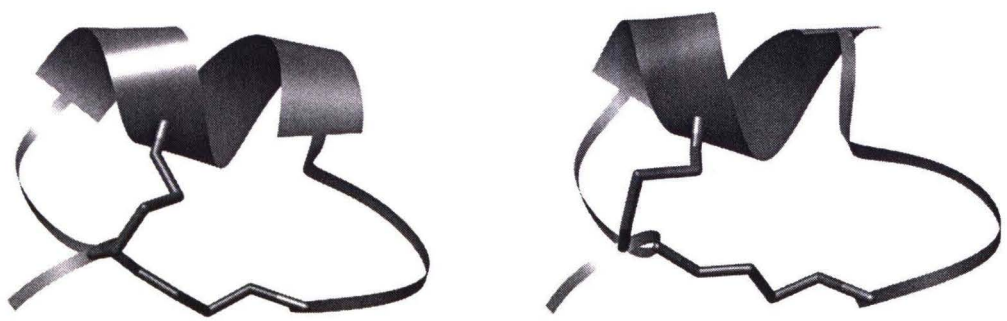


Figure 13. 2D NMR spectra *wg*TOCSY at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) reg2a (B) reg1c.

3.6 Molecular Model of reg2a and reg1d.

A model of the four-Cys (1-3, 2-4 Cys pairing) reg2a and reg1d was built using the MII and ImI conotoxin three-dimensional structure as a template (Figure 14 and 15). Slight differences are found among these two structures; however, the overall three-dimensional fold remains the same for these two conotoxins.



reg2a: GCCSHPACNVNNPHIC

GIC: GCCSHPACAGNNQHIC

Figure 14. Model of the lowest energy structure of reg2a based on GIC. Both belong to the α4/7 conotoxin subclass.

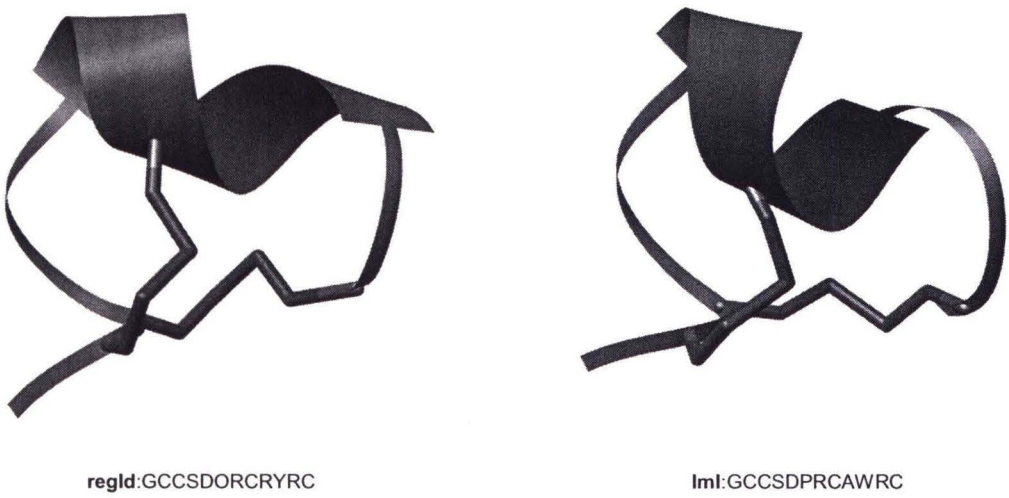


Figure 15. Model of the lowest energy structure of reg1d based on ImI. Both belong to the α4/3 conotoxin subclass.

4. DISCUSSION

The venom of *Conus regius*, a widespread cone snail species that inhabits the tropical Atlantic region and preys upon marine worms, is a complex mixture of conopeptides, which is the typical case of most *Conus* species. The SE-HPLC of the crude venom of *Conus regius* yielded several fractions (Fig. 2). The first seven fractions of the venom were shown to be proteins and peptides of molecular weights that exceeded 3500 Da. All other fractions (8-12) were fractions that contained conopeptides of lesser molecular weight. We chose from these low molecular weight fractions 8 and 9 to be further re-fractionated with the objective to give smaller clusters of conopeptides with even closer molecular weights and shape. RP-HPLC of fraction nine (Figure 4), revealed the presence of several well-resolved components from which we isolated several α-conotoxins including a set of α4/3-conotoxins reg1b, reg1d and an α4/7-conotoxin reg2a. One of the largest components of this fraction is the α4/7 conotoxin reg2a. There are other components present in this fraction; however, the molecular weight and apparently the molecular shape of them all are similar to each other. The combination of SE and RP HPLC provides a facile and efficient procedure for the separation of α-conotoxins that is far superior to the customary direct separation of the venom by RP HPLC, proving that adding a chromatographic step like Size-Exclusion improves the separation of the components dramatically. RP-HPLC of fraction eight (Figure 5) revealed the presence of additional components. Hidden among these components we isolated four more α-conotoxins, reg1a, reg1c, reg1e and reg1e', all belonging to the α4/3 class. 1D and 2D NMR data, when possible, was acquired on all these fractions prior to any further characterization steps. All these

conopeptides were isolated in nanomolar or picomolar quantities as determined by 1D NMR spectrometry. The use of a High Performance 3 mm probe combined with a 1.7 mm capillary insert was necessary for the acquisition of their NMR spectra (See experimental Section). The nano-NMR spectra revealed similar compositions for the reg conopeptides: 9-12 HNs resonances with variable content of readily identifiable residues such as His, Arg, Ala (Fig. 10 and Appendix A). 2D *wg*TOCSY experiments were only possible for reg2a and reg1d. This information was highly valuable in clarifying the presence of some residues that were not clearly identified from the sequencing data obtained. Amino acids have unique spin patterns shown in TOCSY experiments that can aid in identifying certain amino acids. The MALDI-TOF MS spectra of these fractions (Figure 6-9) revealed that these were single components fractions and that these fractions revealed little fragmentation for the molecular components of these fractions, even at high collision energies, which is the typical behavior of disulfide bond-containing peptides. This holds true with components having 2 or more disulfide bonds. It has been observed that components having one disulfide bond can fragment like a linear peptide given the right conditions, such as when the laser energies are high in MALDI MS experiments. Upon reduction of ~150 picomoles of these conopeptides with DTT and alkylation with iodoacetamide, the presence of four Cys residues was revealed for all peptides. Sequencing of the reduced and alkylated conopeptides identified the components as an α4/7-conotoxin and a set of α4/3-conotoxins (Table 1).

In table 1 we compare the sequences of these novel reg1 and reg2 conotoxins with other α4/3-conotoxins (ImI, ImII and Bn1.3) as well as with other α4/7-conotoxins. The reg1e and reg1f α4/3-conotoxins have high sequence homology among them, they differ only in the second to last amino acid where glutamate (reg1f) replaces Glutamine (reg1e). One could consider Glutamine deamidation of reg1e as a product of the isolation and characterization methods. However, it is possible that reg1f is part of the venom arsenal of *C. regius*. reg1e and reg1f have sequence homology with ImI in the first loop between Cys2 and Cys3, but not in the second loop between Cys3 and Cys4. The similarity in this loop lies not in the identity of the residue but in the nature of the residue. ImI and reg1e and f all have a positively charged amino acid in the second loop (K, R). They also share the presence of a bulky amino acid (W, H). Despite these differences, one would expect reg1e and reg1f to target the same subtype nAChR as ImI. The next set of peptides with an identity of their own is reg1a, reg1b, reg1c and reg1d. These represent the outcome of a strategy to optimize the venom components for the capturing of prey, a true conopeptide factory. Even though reg1b is highly homologous to reg1e, we will consider them separate since reg1b is hydroxylated at Pro-6 and appears to be part of a different subclass of α4/3 conotoxins. reg1a, reg1b and reg1d are part of the new arsenal of weapons of *Conus regius*, they all share the conserved Ser-Asp diad in the first loop and they all have a positively charged amino acid and a bulky amino acid in loop two. The difference is that they have hydroxylated Pro-6, a residue that is highly conserved in other members of the α4/3 conotoxins. This has not been seen in the α4/3 conotoxins; however, it has been observed in vc1a, an α4/7 conotoxin from *Conus victoriae*

where the Pro in the first loop (SDPR) has been hydroxylated. Another example of hydroxylation of Pro residues is EI, also an α4/7 conotoxin. In this case the hydroxylation occurred at position Pro-3, outside the conserved sequence of loop one (SDPR). This hydroxylation could be a strategy to go after less hydrophobic receptors where hydrogen bonding could play a more important role in binding and blocking the receptor. Electrophysiology studies were carried out using reg1b and reg1d on *Xenopus* oocytes expressing α3β2 nicotinic receptors. These two conotoxins showed little or no affinity towards this receptor. The difference in the sequence among this reg1d and reg1b is in the second loop where reg1b has (KHQ) and reg1d has (RYR). This difference could account for the better affinity showed by reg1d towards the α3β2 nicotinic receptor. A structure activity relation study should be carried out in the second loop of these conotoxins to determine the amino acid(s) responsible for this slight affinity. Modeling of the structure of reg1d was carried out using as a base model the NMR structure of the α4/3 conotoxin ImI. Slight differences were found; however, the overall three-dimensional fold remains the same. This is mainly a result of sequence homology and disulfide bond constrains, which forces the molecule into a predetermined structure.

The last α4/3 conotoxin left to discuss is reg1c. This molecule does not follow the traditional characteristics of members of the α4/3 conotoxin family. It does not have the SDPR diad in the first loop nor does it have a positively charged amino acid or a bulky amino acid in the second loop. Studies against other nicotinic receptors will be carried out in the future. One could immediately suggest that based upon the sequence

of reg1c that it targets a completely different receptor; thus belonging to another family of conotoxins. reg1c still has the 4/3 characteristics and it conserves the Hyp residue and a positively charged amino acid in the first loop as other members of the reg1 conotoxins. reg1c also has amino acids with long side chains in the second loop (L,I) similar in length to (R,K) found in reg1b and reg1d. Sufficient amounts of reg1c are available for a preliminary bioassay screening; however, more sample would be needed to establish an IC_{50} for this conotoxin.

Studies on RgIA determined from the cDNA of *Conus regius* have shown that it inhibits the α9α10 nicotinic receptors. The IC_{50} reported for RgIA is 6.6 nM. Similar results can be expected for reg1a, reg1b, reg1d, reg1e and reg1e' as they have high sequence homology with RgIA. RgIA is reported to have 1000 fold more affinity towards the α9α10 nicotinic receptors than the α7 nicotinic receptor subtype. reg1 conotoxins could follow this example; however, the presence of post-translational modifications like hydroxylation of Pro residues and amidation of C-terminus could in theory provide specificity towards other of the nicotinic receptor subtypes.

reg2a was the only α4/7 conotoxin isolated from *Conus regius*. It is most striking of this conotoxin that it has high sequence homology with conotoxins GIC and GID isolated from *Conus geographus*. *Conus regius* is a worm-hunter from the Western Atlantic Ocean, while *Conus geographus* is a fish-hunter from the Indo-Pacific region; yet they have both managed to produce conotoxins with high sequence homology. In the first loop between Cys-2 and Cys-3 reg2a has a Ser residue which is

highly conserved among other α4/7 conotoxins that act on the neuronal nicotinic receptors. The second amino acid in this loop is a His residue. 70% of the conotoxins in this class have a His or an Asn residue in this position. The third amino acid in this loop is a Pro residue, highly conserved among members of this family. The fourth residue in the first loop is the amino acid Ala, which is found in approximately 25% of the sequences reported for this class, amino acids like R, P and V are also present in this position. The major differences among α4/7 conotoxins arise in the second loop, although there is a common amino acid (N or H) in position 4 of the second loop the rest of the amino acid residues vary significantly. It is here where GIC and GID appear to be closely related to reg2a, something unexpected since both cone snails come from different places and target different prey. One would expect a sequence like PnIA coming from another worm-hunter like *Conus pennaceus* to have similarities with reg2a, not a sequence from a fish-hunter. The three-dimensional structure of reg2a was modeled from the NMR structure of conotoxin GIC. It is remarkable to observe that toxins from opposite sides of the world can have similar sequences and three-dimensional fold. Although slight differences can be found from one structure to the other, the overall three-dimensional fold is quite similar. The main difference among the structures of these two conotoxins can be attributed to the presence of a Pro residue in reg2a instead of a Gln residue found in GIC. The presence of this Pro residue forces the sequence into a more pronounced turn in the second loop, GIC does have a turn but not as pronounced as reg2a. The structure of reg2a displays a more organized helix than GIC, this could be due to the presence of an Asn residue in the beginning of the second loop that can help stabilize the helix

through hydrogen bonding. GIC has an Ala residue instead of the Asn residue. In principle, the difference in specificity among these α4/7 conotoxins that have the same structural framework can be attributed to shape, charge distribution and the hydrophobicity displayed by these conopeptides as a result of their sequence variability. reg2a has been found to block the α3β2 nicotinic receptor expressed in *Xenopus* oocytes (Figure 10). The response of α3β2 nicotinic receptors to ACh was blocked by about 85%, the block was washed off after about one minute. This indicates that the strength of the affinity of reg2a to this nicotinic receptor subtype is comparable to PnIA. Conotoxins like MII can bind the α3β2 nicotinic receptors for tenths of minutes, thus having the binding can be considered irreversible. While reg2a targets the α3β2 nicotinic receptors one can not rule out that reg2a antagonizes other nicotinic receptor subtypes. reg1 and reg2 conotoxins are great examples of *Conus* venom evolution within a species. These conotoxins represent the second most abundant set of peptides present in the venom of *Conus regius* (miniM conotoxins are the most abundant, see chapter one).

5. REFERENCES

1. McIntosh, J.M., B.M. Olivera, and L.J. Cruz, *Conus peptides as probes for ion channels.* Methods in Enzymology, 1999. **294**(Ion Channels, Part C): p. 605-624.
2. McIntosh, J.M., et al., *A Nicotinic Acetylcholine-Receptor Ligand of Unique Specificity, α-Conotoxin-ImI.* Journal of Biological Chemistry, 1994. **269**(24): p. 16733-16739.
3. Hansen, S.E., *Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations.* The EMBO Journal, 2005. **24**: p. 3635-46.
4. Olivera, B.M., et al., *Conotoxins.* Journal of Biological Chemistry, 1991. **266**(33): p. 22067-22070.
5. Santos, A.D., et al., *The A-superfamily of conotoxins - Structural and functional divergence.* Journal of Biological Chemistry, 2004. **279**(17): p. 17596-17606.
6. Buczek, O., et al., *Posttranslational amino acid isomerization: A functionally important D-amino acid in an excitatory peptide.* Journal of Biological Chemistry, 2005.
7. Craig, A.G., et al., *Contulakin-G, an O-glycosylated invertebrate neurotensin.* Journal of Biological Chemistry, 1999. **274**(20): p. 13752-13759.
8. Craig, A.G., P. Bandyopadhyay, and B.M. Olivera, *Post-translationally modified neuropeptides from Conus venoms.* European Journal of Biochemistry, 1999. **264**(2): p. 271-275.

9. Craig, A.G., et al., *Enzymatic glycosylation of contulakin-G, a glycopeptide isolated from Conus venom, with a mammalian ppGalNAc- transferase.* Toxicon, 2001. **39**(6): p. 809-815.

10. Myers, R.A., et al., *Conus Peptides as Chemical Probes for Receptors and Ion Channels.* Chemical Reviews, 1993. **93**(5): p. 1923-1936.

11. McIntosh, J.M., A.D. Santos, and B.M. Olivera, *Conus peptides targeted to specific nicotinic acetylcholine receptor subtypes.* Annual Review of Biochemistry, 1999. **68**: p. 59-88.

12. Olivera, B.M. and J.M. McIntosh, α-*Conotoxins (Conus spp.).* Guidebook to Protein Toxins and Their Use in Cell Biology, 1997: p. 202-203.

13. McIntosh, J.M., α-*Conotoxins and nicotinic acetylcholine receptors.* Biochemical Aspects of Marine Pharmacology, 1996: p. 28-33.

14. Adams, D.J., et al., *Conotoxins and their potential pharmaceutical applications.* Drug Development Research, 1999. **46**(3/4): p. 219-234.

15. Hill, J.M., et al., *Three-Dimensional Solution Structure of α-Conotoxin MII by NMR Spectroscopy: Effects of Solution Environment on Helicity.* Biochemistry, 1998. **37**(45): p. 15621-15630.

16. Franco, A. and F. Mari, *Three-dimensional structure of α-conotoxin EI determined by 1H NMR spectroscopy.* Letters in Peptide Science, 1999. **6**(4): p. 199-207.

17. Lewis, R.J. and M.L. Garcia, *Therapeutic potential of venom peptides.* Nature Reviews Drug Discovery, 2003. **2**(10): p. 790-802.

18. Terlau, H. and B.M. Olivera, *Conus venoms: A rich source of novel ion channel-targeted peptides.* Physiological Reviews, 2004. **84**(1): p. 41-68.

19. Everhart, D., Edward, C.; Luetje, C., *Determinants of Potency on α-conotoxin MII, a Peptide Antagonist of Neuronal Nicotinic Receptors.* Biochemistry, 2004. **43**: p. 2732-37.

20. Fiser, A.S. and A. Sali, *MODELLER: Generation and refinement of homology-based protein structure models*. 2003, Academic Press: San Diego. p. 463-493.

21. Aird, S.D., *Chromatographic behavior of Bothrops erythromelas phospholipase and other venom constituents on Superdex 75.* Preparative Biochemistry & Biotechnology, 2004. **34**(4): p. 345-364.

22. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.

23. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching.* Analytical Chemistry, 1999. **71**(14): p. 2871-2882.

Chapter III

P-Superfamily Conotoxins from the Venom of *Conus regius*: a Western Atlantic Worm-Hunting Cone Snail Species

ABSTRACT

We have isolated five novel P-Superfamily conotoxins (reg9a-e) from the venom of *Conus regius*, a wide spread worm-hunting cone snail from the Western Atlantic Ocean. None of the reg9 conopeptides have sequence homology with previously reported P-Superfamily conotoxins. reg9d and reg9b have high sequence homology with each other; however, they are quite different from the rest of the reg9 conopeptides. The rest (reg9a, reg9c and reg9e) have little or no sequence homology among themselves. The position of the cysteine residues throughout the sequence and the number of residues present between the second and third cysteines (five), third and fourth cysteines (three) and fourth and fifth cysteines (one) are conserved. The only feature that reg9 conopeptides have in common with other members of the P-superfamily such as tx9a from *Conus textile*, ama9a from *Conus ammiralis* and gm9a from *Conus glorimaris* is the number of residues before, after and in between cysteines. Sequence analysis of the reg9 conopeptides produced 3 complete

sequences for reg9c, reg9d and reg9e; reg9a and reg9b did not sequence completely even after repeated attempts. The receptor target for the P-superfamily members is not known. reg9d was submitted for neuroprotection assays and there is preliminary evidence that it offers cellular neuroprotection. NMR information was not obtained for the majority of reg9 conopeptides; however, reg9d was the only reg9 conopeptide isolated in sufficient amounts to obtain 1D and 2D NMR spectra. Despite the lack of sequence homology the three-dimensional structure of reg9d was modeled based on the NMR structure of gm9a from *C. glorimaris* in an attempt to observe the restraining effect of the disulfide bonds.

1. INTRODUCTION

The first conopeptides to be neuropharmacologically characterized were the α-conotoxins, which act as competitive antagonists at nicotinic acetylcholine receptors at the neuromuscular junction [1]. Additional classes of conopeptides were subsequently found to act at voltage-dependent Ca^{2+} channels, Na^{+} channels, K^{+} channels, 5-HT3receptors, NMDA receptors, vasopressin receptors and neurotensin receptors [2]. Conotoxins are grouped into various superfamilies (O, M, A, S, T, P, I), each with highly conserved signal sequences in their precursor proteins and a characteristic cystine arrangement in the mature peptides [3]. Within the superfamilies, conotoxins are further classified into families according to their specific pharmacological targets [4]. This highly specific action of neurotoxins has enabled neuroscientists to study key components of both vertebrate and invertebrate nervous systems [5]. Their small size, relative ease of synthesis, structural stability and target specificity make them important pharmacological probes. Many neurotoxins and their derivatives have been used as therapeutic agents [6]. As the number of ion channels and receptors grows, so does the need for additional target-specific neurotoxins. Among the members of the P-superfamily (spastic peptides) reported to date, we find tx9a and from *Conus textile* [7], ama9a from *Conus ammiralis,* BeTXIIb from *Conus betulinus* and gm9a from *Conus glorimaris* [7]. These conotoxins have in common a distinctive cysteine framework, -$CX_3CX_5CX_3CXCX_4C$-, where X can be any amino acid. The molecular target of these conotoxins has not been determined; however, the physiological responses of tx9a and gm9a have been observed upon intracranial injection to mice. These

peptides produced uncontrollable spasms in mice; hence the term spasmodic peptides. As part of our efforts to discover the molecular target of this superfamily we carried out neuroprotection studies on one member of the reg9 conopeptides. The mechanism by which neuronal damage is brought about is unclear; however, it seems to involve a cascade of events triggered by energy deficiency in brain tissue. Large increases in the release of neurotransmitters are seen. Glutamate, acting through NMDA and AMPA receptors, allows calcium to enter cells which leads to calcium release from intracellular stores. Cellular depolarization also allows calcium to enter neurons via voltage-gated calcium channels. A variety of enzymatic and degradation events follow, leading ultimately to rapid or delayed cell death. Clearly, interfering with any stage of these events might protect neurons [8]. We followed methods established for the neuronal protection offered by taurine [9] and replaced taurine for one of the reg9 conopeptides. Taurine regulates intracellular calcium levels in neurons. Excess calcium influx can be achieved through glutamate-induced toxicity and neuroprotective compounds like taurine can help reduce the neuronal damage from excess calcium levels and the cascade of events it may produce.

In this chapter we present the isolation of five novel conotoxins from the venom of *Conus regius*, a worm-hunting cone snail species from the Western Atlantic. These conopeptides reg9a-e display high sequence divergence among them for the most part, only reg9b and reg9d have a high degree of sequence homology. We will discuss the sequence divergence of these peptides, their similarities in their cysteine framework to other members of this superfamily. We will observe the three-

dimensional structure of reg9d modeled after the NMR structure of gm9a to establish if sequence divergence can affect the overall three-dimensional fold, and we will evaluate the indications that a member of this family has neuroprotection properties.

2. MATERIALS & METHODS

2.1 Specimen collection.

Refer to materials & methods Chapter I.

2.2 Crude venom extraction.

Refer to materials & methods Chapter I.

2.3 Peptide purification.

Refer to materials & methods Chapter I.

2.4 Reduction and alkylation of cysteyl residues.

Refer to materials & methods Chapter I.

2.5 Peptide sequencing and amino acid analysis.

Refer to materials & methods Chapter I.

Amino acid analysis was carried out at Yale University, W. M. Keck Foundation Biotechnology Resource Laboratory.

2.6 Molecular mass determination.

Refer to materials & methods Chapter I.

2.7 Neuroprotection assays for reg9d.

In collaboration with Dr. Jang Ye Wu from the Biomedical Sciences Department at Florida Atlantic University and Dr. Ian Mellor from the Department of Biology of the University of Nottingham, potential neuroprotective effects were tested in primary cultured rat brain neuronal cell cultures. Cell death by necrosis was compared using the lactate dehydrogenase (LDH) assay [9]. Mitochondrial activity was monitored using the 3-(4,5-diethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay. Calcium influx analyzed using a Calcium-45 assay, cells were exposed to 50μM glutamate for 10 minutes. Two electrode voltage-clamping of *Xenopus* oocytes expressing NR1A and NR2A NMDA receptor subtypes was performed to determine receptor inhibition.

2.8 NMR Spectroscopy.

NMR spectra were acquired on a Varian Inova 500 MHz instrument equipped with PFG, 3xRF channels and waveform generators. Nanomolar quantities of the native conopeptide directly isolated from the venom (reg9d = 72 nmoles, the rest of the reg9 conopeptides were well below 1 nmole and no NMR spectra obtained even after 4000 scans) were dissolved in 40 μl of water with 10% D_2O (used for locking purposes) and 4 nanomoles of TSP and placed in 1.7 mm NMR tubes (Wilmad WG-1364-1.7). For further details on this procedure refer to Materials & Methods Chapter I.

2.9 Molecular model of reg9d.

Molecular models were built by comparative modeling methods [10] based on the NMR structure of the gm9a conotoxin as template and using Modeller (version 8.0). Briefly, conotoxin sequences were aligned according to the standard routine in the program using the PDB entry as template. A set of 10 model structures was built accordingly for each and the selected structures were that of better target Modeller energy.

2.10 Nomenclature.

In this chapter we adopted a nomenclature of three letters to designate *Conus* species, because the one or two letter nomenclature currently in place will not be enough to describe the large number of different non-fish-hunting species, especially those with similar first letter names. We decided to use the three letters "reg" to name the peptides from *C. regius.* Arabic numbers were used to represent the disulfide framework; 9 has already been assigned to the members of the P-superfamily. The letter after the framework number indicates the order of elution on RP-HPLC.

3. RESULTS

3.1 Peptide purification.

The venom of *Conus regius* is an extremely complex mixture of peptides and protein. More than 100 fractions can be obtained from this separation (Figure 1B, Chapter I). However, most of these fractions show multiple components, for which we were required to adopt an improved separation scheme that includes a prefractionation step using size exclusion chromatography on a Superdex-30 column (Figure 1), followed by a refined peptide-optimized size exclusion step on a Superdex Peptide column. The resulting fractions are then separated by reversed phase on a peptide-optimized C18 Vydac Everest column. Most of the resulting fractions are single-component (Figure 2) and were subsequently analyzed by mass spectrometry (MALDI-TOF and ESI-Q-TOF), NMR spectroscopy and peptide sequencing by Edman degradation chemistry. Using this methodology, the complete analysis of the most significant components of the venom of *C. regius* can be sequenced (conopeptidome) and the components of the venom can be grouped in the different families of conopeptides. The reg9 peptides were separated by the Superdex-30 in spite of their similar size. This is not entirely unexpected, since this column is also known to partition analytes by hydrophobic interactions [11].

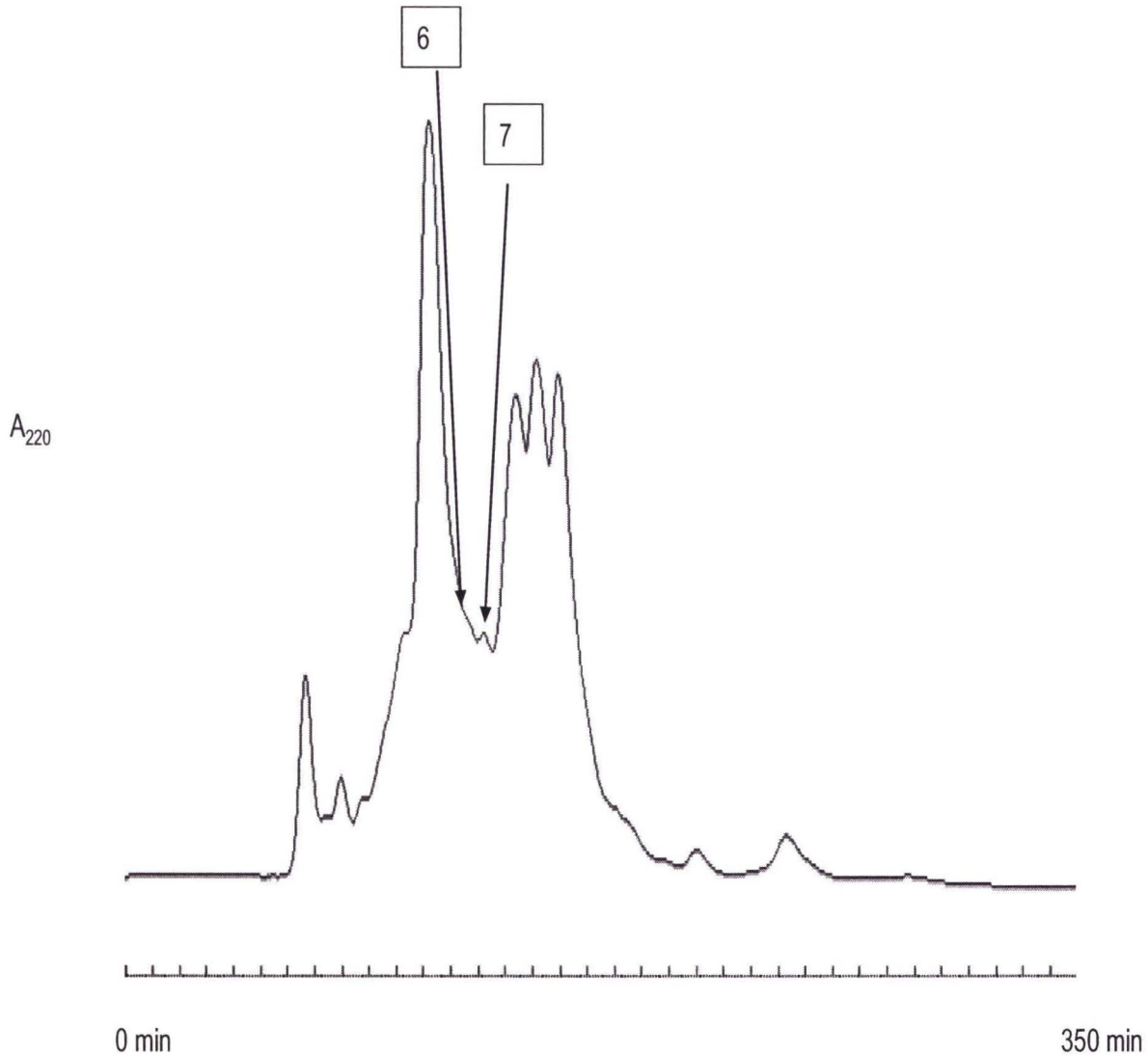


Figure 1. Size exclusion chromatogram of *C. regius* dissected venom. 60 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 6 and 7 were chosen for further separation.

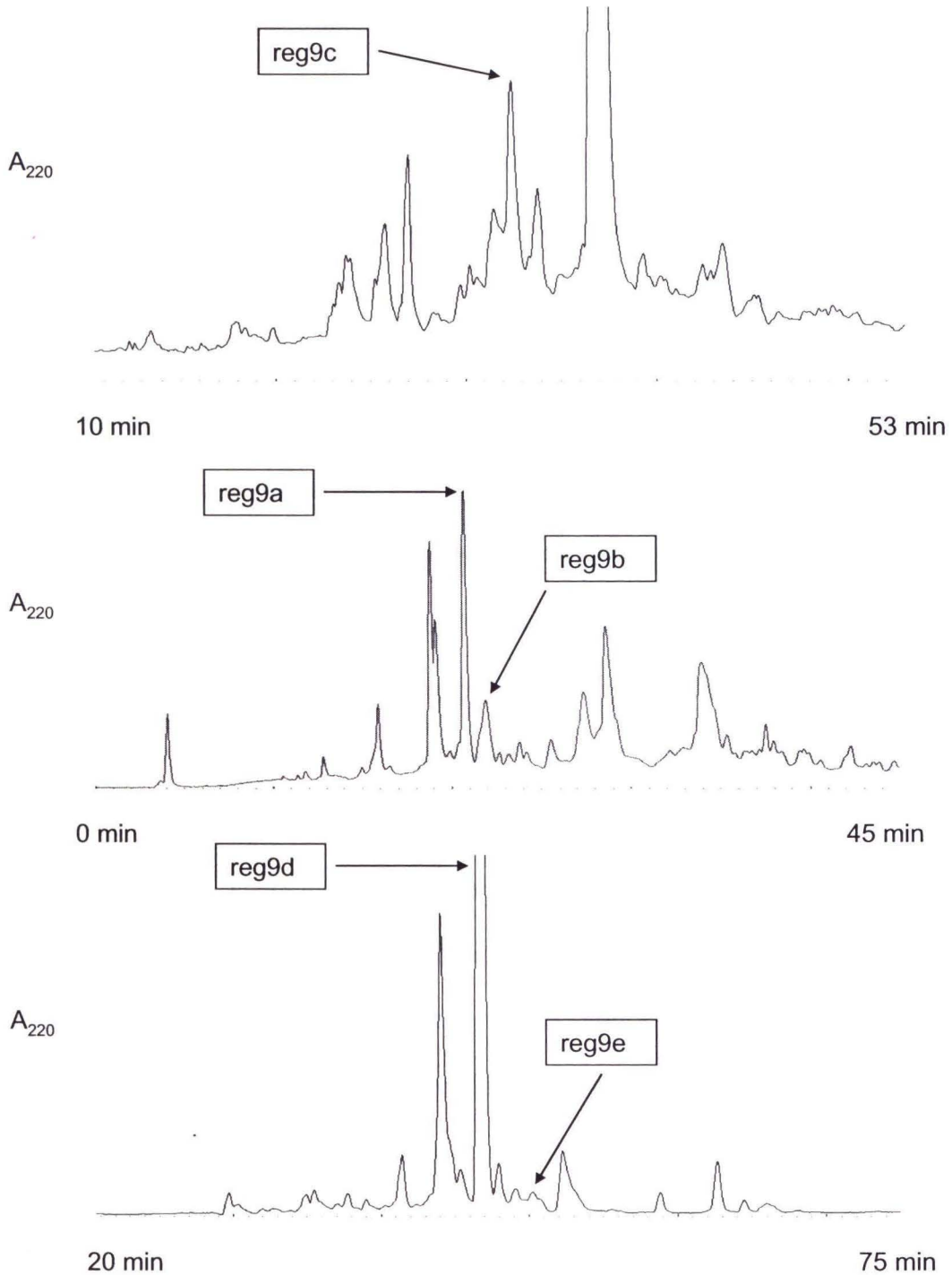


Figure 2. Semi-preparative RP-HPLC chromatogram (SE fractions 6 and 7) showing the location of reg9a, reg9b, reg9c, reg9d and reg9e peptides. A linear gradient from 0 to 100% solution B in 100 min at 1ml/min was employed. HPLC solutions were: 0.1% (v/v) TFA in water (A) and 0.1% (v/v) TFA in 60% (v/v) aqueous MeCN (B).

3.2 Reduction/Alkylation, peptide sequence determination and amino acid analysis. The purified peaks were subjected to reduction with DTT and alkylation with iodoacetamide. Mass spectrometry of the reduced/carboxymethylated peptides and of native peptides showed a mass difference consistent with the presence of six cysteine residues for reg9a, reg9c and reg9e peptides. reg9b and reg9d did not reduce or alkylate after repeated attempts. TCEP was used as alternative reducing agent without any success. The reduced/alkylated (reg9a, 9c, 9e) and unreduced (reg9b, 9d) conotoxins were sequenced by Edman degradation. reg9b and reg9d were then submitted for amino acid analysis to verify the presence of cysteine residues through performic acid oxidation (Table 1). The sequences of the reg9 conotoxins are shown in Table 2. It is worth noticing that the differences between the sequences of peptides reg9b and reg9d are at amino acid position 7, where a methionine is substituted by a γ-carboxyglutamate; at position 22, where a tryptophan is substituted for a histidine and at position 24, where an arginine is substituted by a leucine residue. Little or no sequence homology is shared by the rest of reg9 peptides (reg9a, reg9c, reg9e); however, it seems as residues with hydroxyl group are common among these members of the P-superfamily. Amino acid residues such as Thr, Ser, Tyr and Hyp that have side chains with OH groups in them are distributed throughout the many loops of these conopeptides. reg9a, reg9b and reg9c did not fully sequence, only in the case of reg9c the last amino acid residue was found to be alanine and evidence of Hyp eluting with Pro were present in the sequencer cycles. In the case of reg9a and reg9b, their complete sequences were not obtained: reg9a is missing an amino acid in position 12 and reg9b is missing at least 4 amino acids. These missing residues in the

sequencing procedure could be due to PTMs. Hydroxyprolines were present in sequences reg9a, reg9b and reg9d, this residue was evident by presence of two peaks in the same cycle corresponding to the two isomers of hydroxyproline. These conotoxins contained 27 or more residues and the six cysteines were separated by five loops of variable sizes (loop1 = 3-4 amino acids, loop2 = 5 amino acids, loop3 = 3 amino acids, loop4 = 1 amino acid, loop5 = 4-5 amino acids). Additionally, a BLAST search [12] of the databases (Swissprot/EMBL, PIR, PDB and nrdb95) did not show any significant sequence homology to reported proteins and peptides. The sequences of these conotoxins from *Conus regius* are reported in Table 2.

Amino Acid	aaa665	nmoles aa	µgrams	mole percent	# residues
cysac	16.1367	16.137	2.439	26.5%	6.4
cmcys					
asx	8.5535	8.554	0.984	14.1%	3.8
thr	0.3917	0.392	0.040	0.6%	0.2
ser	0.4632	0.463	0.040	0.8%	0.2
glx	1.9358	1.936	0.248	3.2%	1.0
pro	0.7159	0.716	0.070	1.2%	2.3
gly	8.7947	8.795	0.502	14.5%	2.9
ala	2.8882	2.888	0.205	4.8%	1.3
val	5.2291	5.229	0.518	8.6%	2.3
met					
ileu	0.3174	0.317	0.036	0.5%	0.1
leu	5.1968	5.197	0.588	8.5%	2.3
tyr	destroyed				
phe	0.0622	0.062	0.009	0.1%	0.0
his	2.5365	2.537	0.348	4.2%	1.1
lys	0.1508	0.151	0.019	0.2%	0.1
trp					
arg	7.4090	7.409	1.157	12.2%	3.3
% injected	100%				27

Table 1. Performic acid oxidation and amino acid analysis of ~ 750 pmoles of reg9d.

Peptide	Sequence	# residues between Cys
reg9a	--G**C**TGOK**C**TKDNX**C**ASO**C**K-**C**GYYSL**C**H	4,5,3,1,5
reg9b	RVL**C**OGM-**C**DODVG**C**OAG**C**Y-**C**HHLL-**C**XXXX	3,5,3,1,4
reg9c	--H**C**GSK-**C**FSDDH**C**PAS**C**P-**C**AAHFR**C**VRSA	3,5,3,1,5
reg9d	RVL**C**OGγ-**C**DODVG**C**OAG**C**Y-**C**HWLR-**C**RW	3,5,3,1,4
reg9e	--V**C**SGV-**C**YPAIT**C**NAN**C**K-**C**GKYFN**C**IPSS	3,5,3,1,5
Tx9a	G**C**NNS-**C**QγHSD**C**γSH**C**I-**C**TFRG-**C**GAVN*	3,5,3,1,4
Ama9a	S**C**NNS-**C**QQHSQ**C**ASH**C**V-**C**LLNK-**C**RTVN	3,5,3,1,4
BeTXIIb	G**C**GGV-**C**AYGES**C**PSS**C**NT**C**YSAQ-**C**TAQ	3,5,3,2,4
gm9a	S**C**NNS-**C**QSHSD**C**ASH**C**I-**C**TFRG-**C**GAVN	3,5,3,1,4
tx7	G**C**SSV-**C**NSHTD**C**VTH**C**I-**C**TFRG-**C**GAVN	3,5,3,1,4

Table 2. reg9 sequences from *Conus regius* and other reported members of the P-superfamily.

3.3 Mass spectrometry of purified peptides.

Mass spectrometry carried out using MALDI-TOF in the reflector mode (M/ΔM resolution ~ 10,000) yielded the following monoisotopic molecular ions: reg9a = 2797.8 Da, reg9b = 3354.6 Da, reg9c = 3000.7 Da, reg9d = 3180.31 Da, reg9e = 2937.8 Da (Figures 3 and 4). Mass analysis of the reduced/carboxymethylated peptides and the native peptides showed a mass difference consistent with the presence of six cysteine residues for reg9a, reg9c and re9e conopeptides. The masses obtained for the peptides were for the most part not in agreement with the calculated theoretical monoisotopic values determined for the assigned sequences. reg9a, reg9b and reg9c did not fully sequence. Repeated attempts were carried out with no success. reg9a has a mass difference of 126.8 Da from the theoretical MW and the experimental MW. reg9b has a 644.6 Da mass difference from the theoretical MW and the experimental MW. reg9c had a mass difference of 71 Da, which corresponds to an alanine at the last residue. The MALDI-TOF spectra of reg9c also shows smaller peaks with a difference of +16 Da and +32 Da which indicates the possibility of hydroxylation at the first and second Pro residue. reg9d had evidence from sequencing that a γ-carboxyglutamate residue was present at position 7. This shows in the experimental MW obtained from MALDI-TOF where the m/z peak is that of 3138.8 Da , 42 Da less than the MW obtained from ESI (3180.31 Da). Although it is not a full 44 Da loss that one would expect from the loss of a CO_2 molecule coming from the decarboxylation of γ-carboxyglutamate, it is highly likely that it comes from this residue, which will then be in agreement with the calculated MW. Also present in the MALDI-TOF spectra of reg9d are small peaks with a difference of -16 Da, which

would indicate the presence of the same peptide with a Pro residue instead of one of the Hyp residues present in the sequence. reg9e had a 1.4 Da difference between the experimental and calculated MW; however, the same situation occurs were a smaller peak with a +16 Da is present, indicating the possibility of hydroxylation of one of the Pro residues. From the MALDI-TOF MS data obtained and without the full sequence giving us the theoretical MW, the difference in the calculated molecular weights and the experimental molecular weights from MALDI-TOF could not be calculated for the reg9a, reg9b and reg9c conopeptides. This results in the inability to determine if the peptides were amidated at the C-terminus. The calculated MW for reg9d and reg9e were determined, and it was established that the peptides were not amidated at the C-terminus. The calculated molecular weights (MWcal) were obtained using Protein Prospector [13].

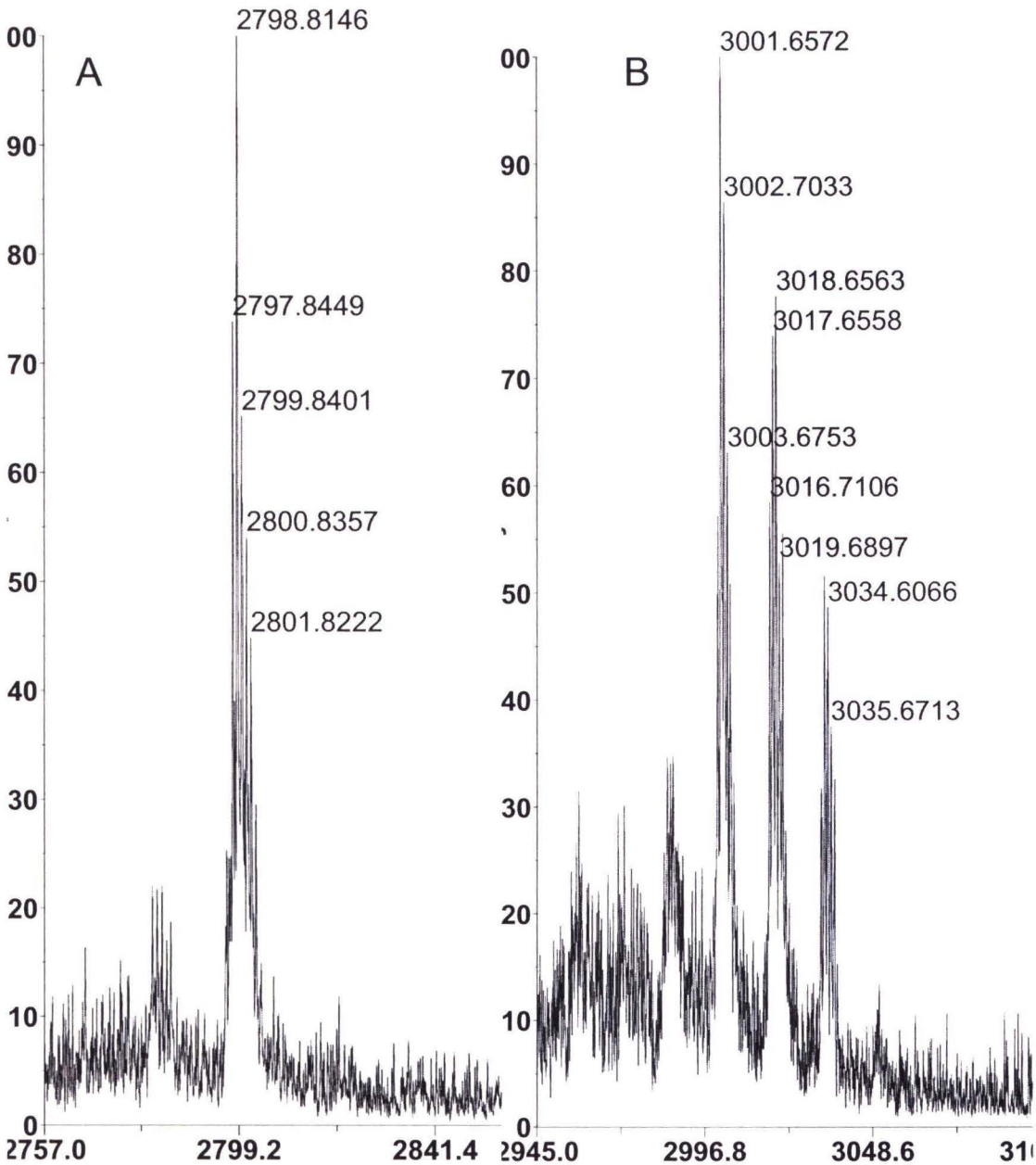


Figure 3. MALDI-TOF mass spectra for reg9a (A) and reg9c (B).

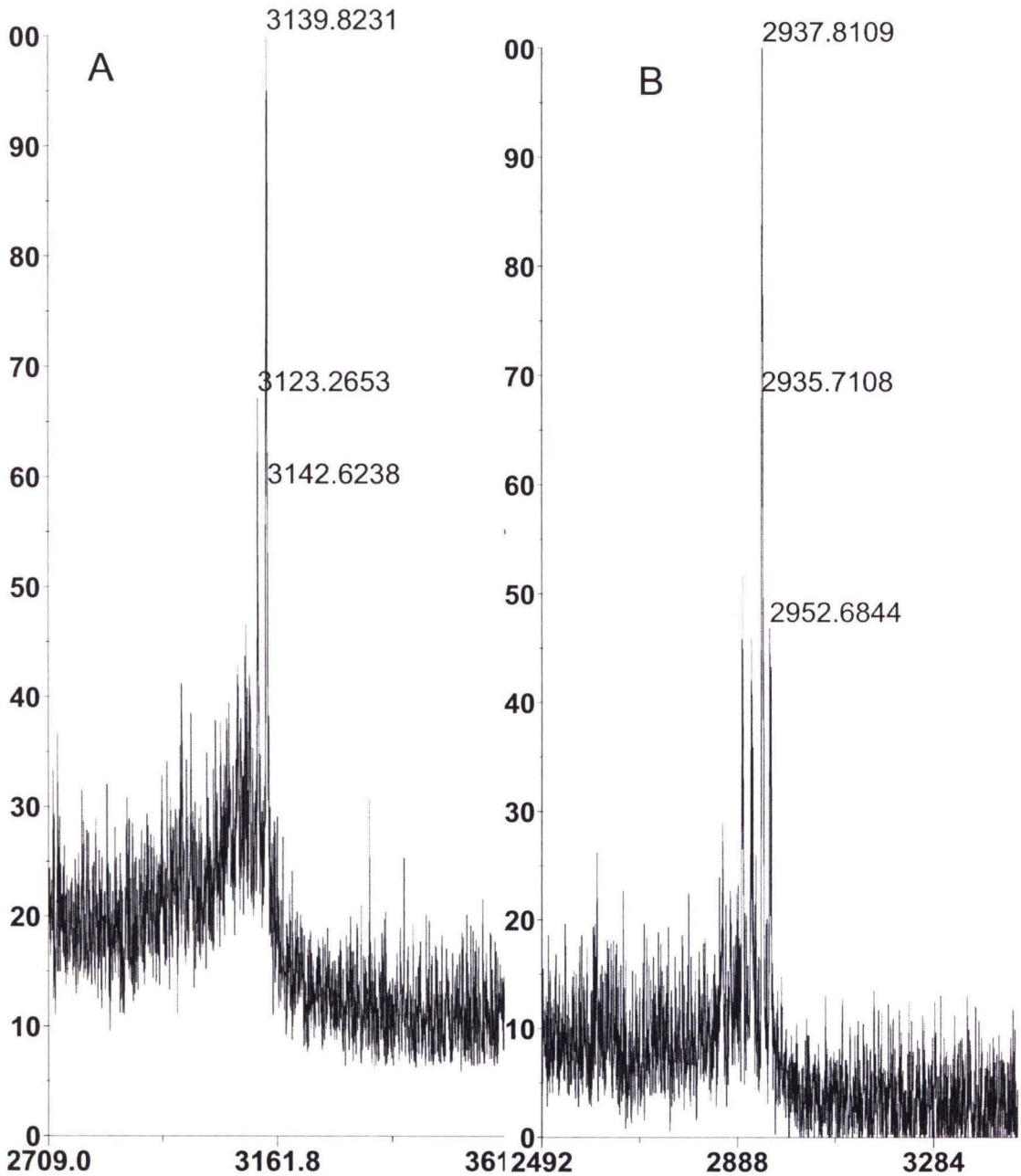


Figure 4. MALDI-TOF mass spectra for reg9d (A) and reg9e (B).

3.4 Neuroprotection assays for reg9d.

After pre-treatment with 0.25 μM Glu + 1 μM of reg9d peptide a significant increase in mitochondrial activity (Figure 5) and a decrease in mortality of cells exposed to glutamate was observed (Figure 6). Significant reduction in glutamate-induced calcium influx (83%) (Figure 7) was also seen suggesting that inhibition of a glutamate-gated ion channel occurs. Due to quantity limitations we were unable to test the peptide at a concentration greater than 100 nM. Other conopeptides reported to have neuroprotective potential have IC_{50} values in the μM range so grater inhibition may be observed at these higher concentrations.

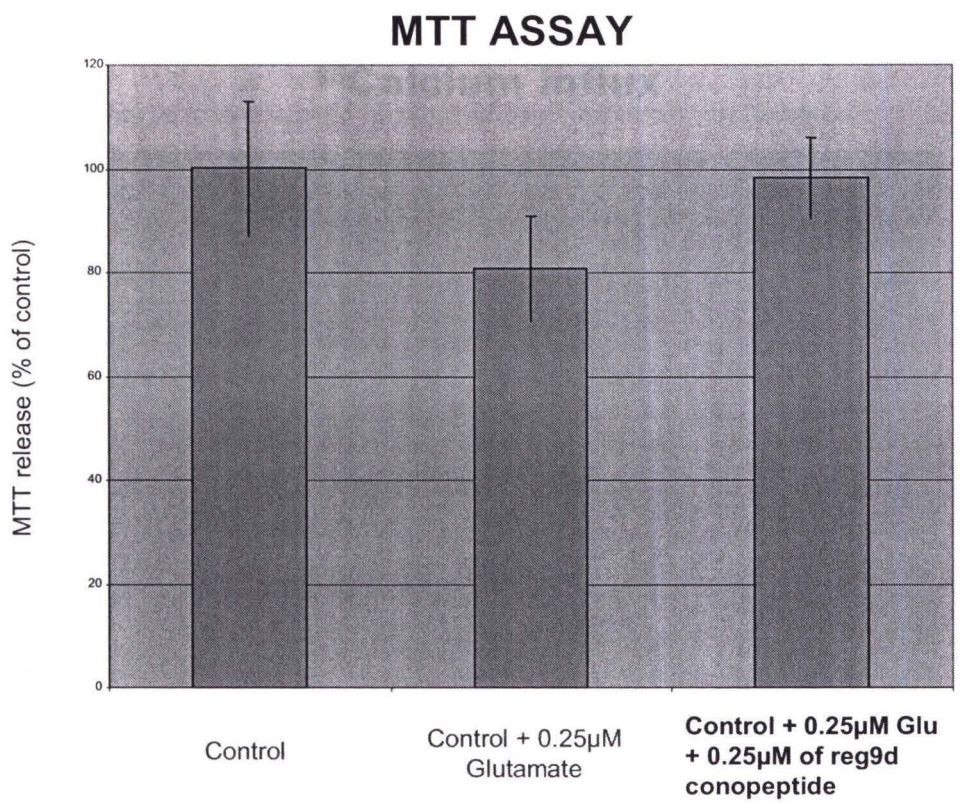


Figure 5. MTT assay for reg9d.

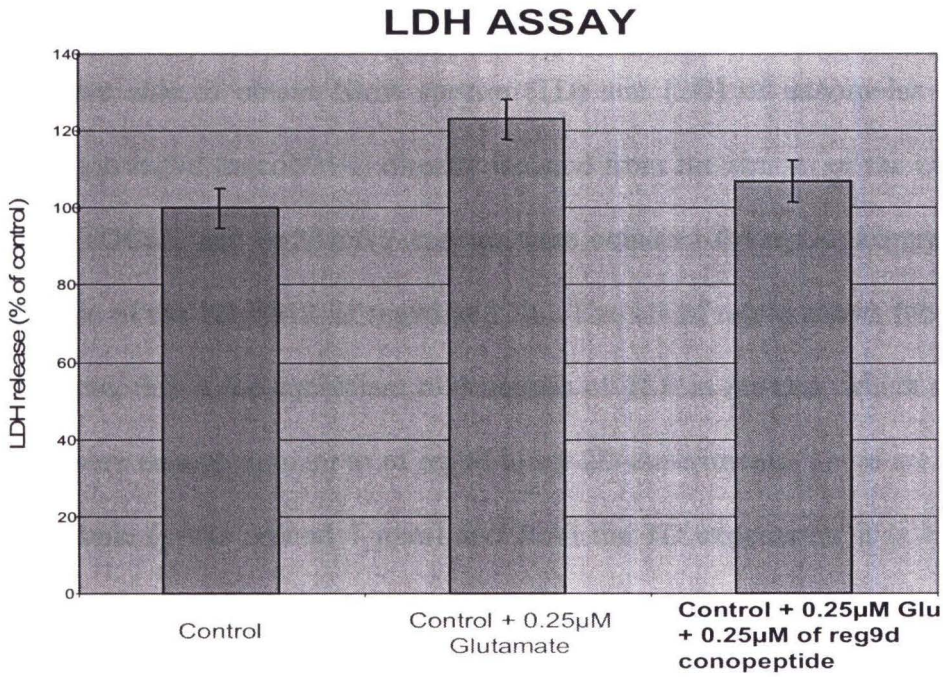


Figure 6. LDH assay for reg9d.

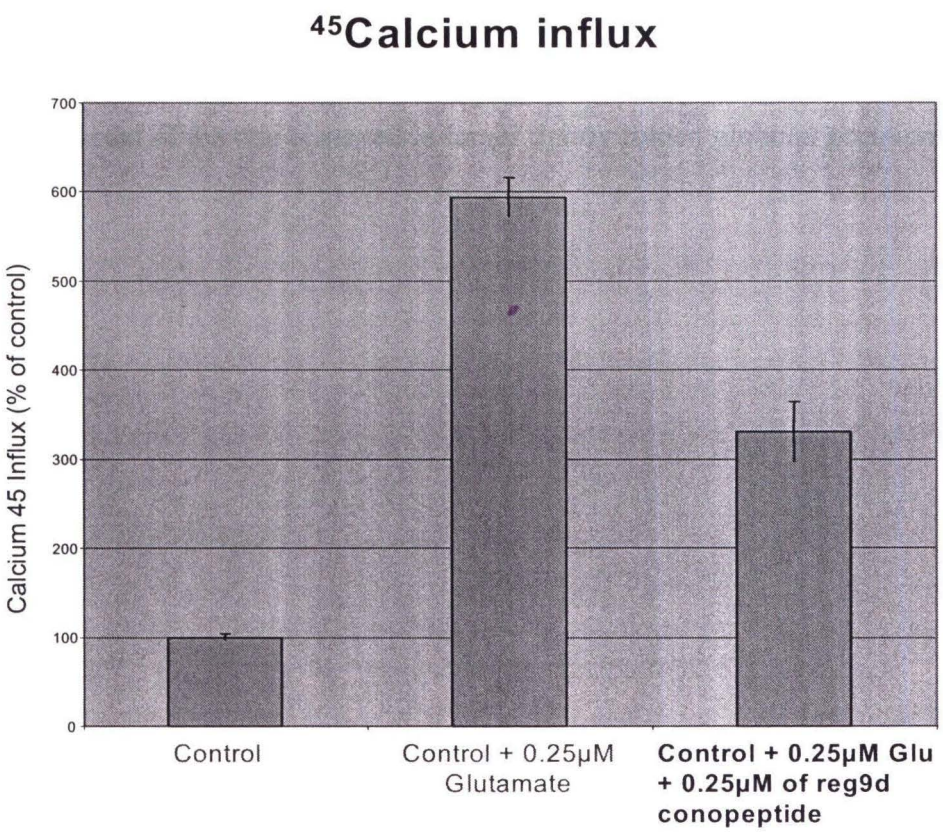


Figure 7. Calcium influx assay for reg9d.

3.5 NMR spectroscopy.

We were able to obtain NMR spectra (1D) and (2D) of nanomolar quantities of conotoxin reg9d (nanoNMR) directly isolated from the venom of the cone snails. A 2D *wg*TOCSY and *wg*NOESY spectra were obtained for reg9d. Figure 8 shows an example of the 1D NMR of reg9d at 25°C. The 1D of reg9d shows the TSP peak at zero ppm, this is the equivalent to 4 nmoles of TSP in protons, which indicates that there were enough amount to of reg9d to try 2D experiments. There are impurities in the sample (peaks around 7 ppm) and from the 1D experiment it is evident that a tryptophan residue is part of the sequence (peak around 10.2 ppm). Figure 9 represents the 2D experiments of reg9d. Both, the *wg*TOCSY spectra (Figure 9A) and *wg*NOESY spectra of reg9d display spectra suitable for structural analysis. The 2D data of reg9d shows a very well-defined structure in solution, in a way that is reminiscent of the one observed in larger tightly folded globular proteins.

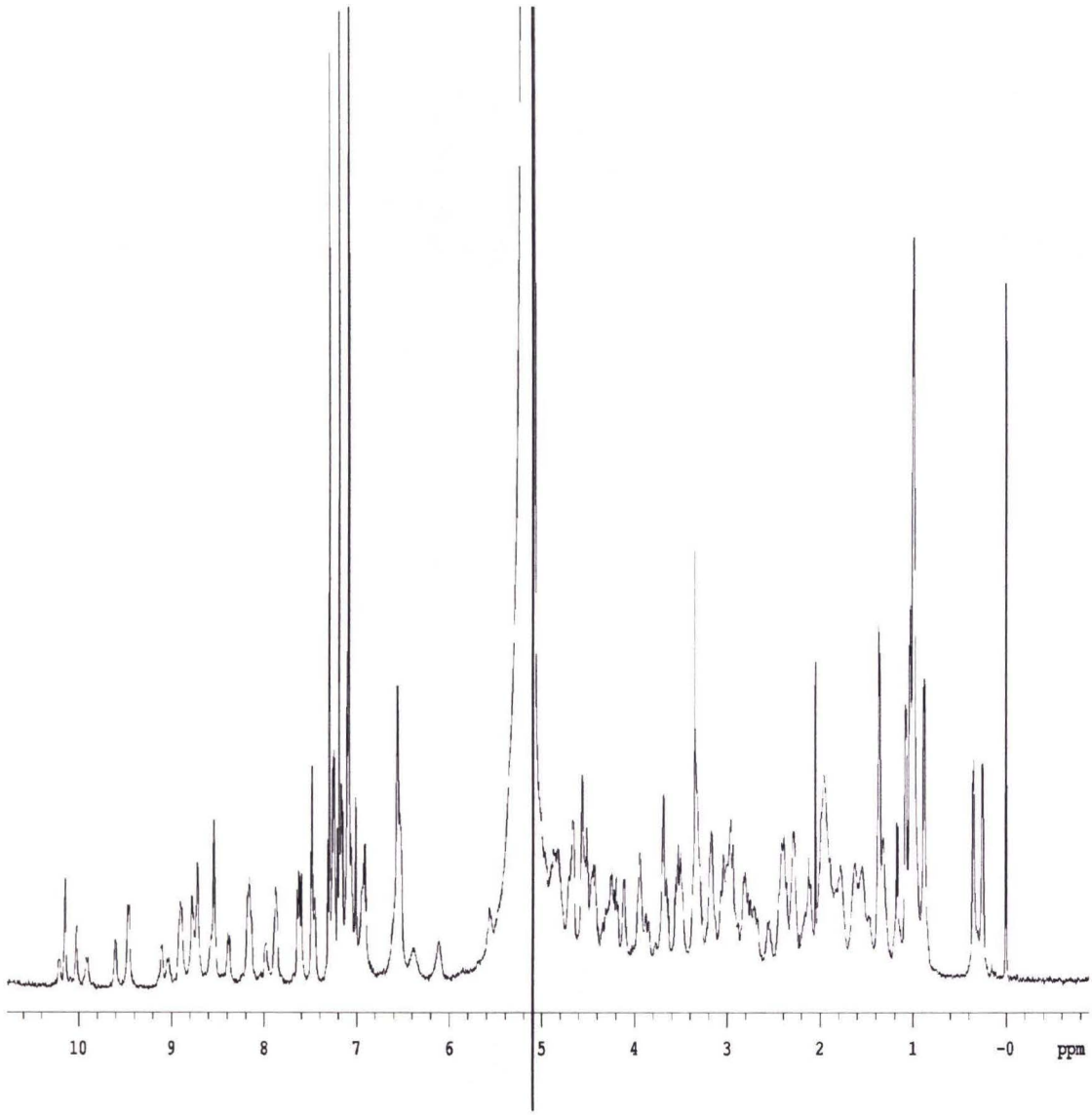


Figure 8. Proton 1D NMR spectrum at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe of reg9d.

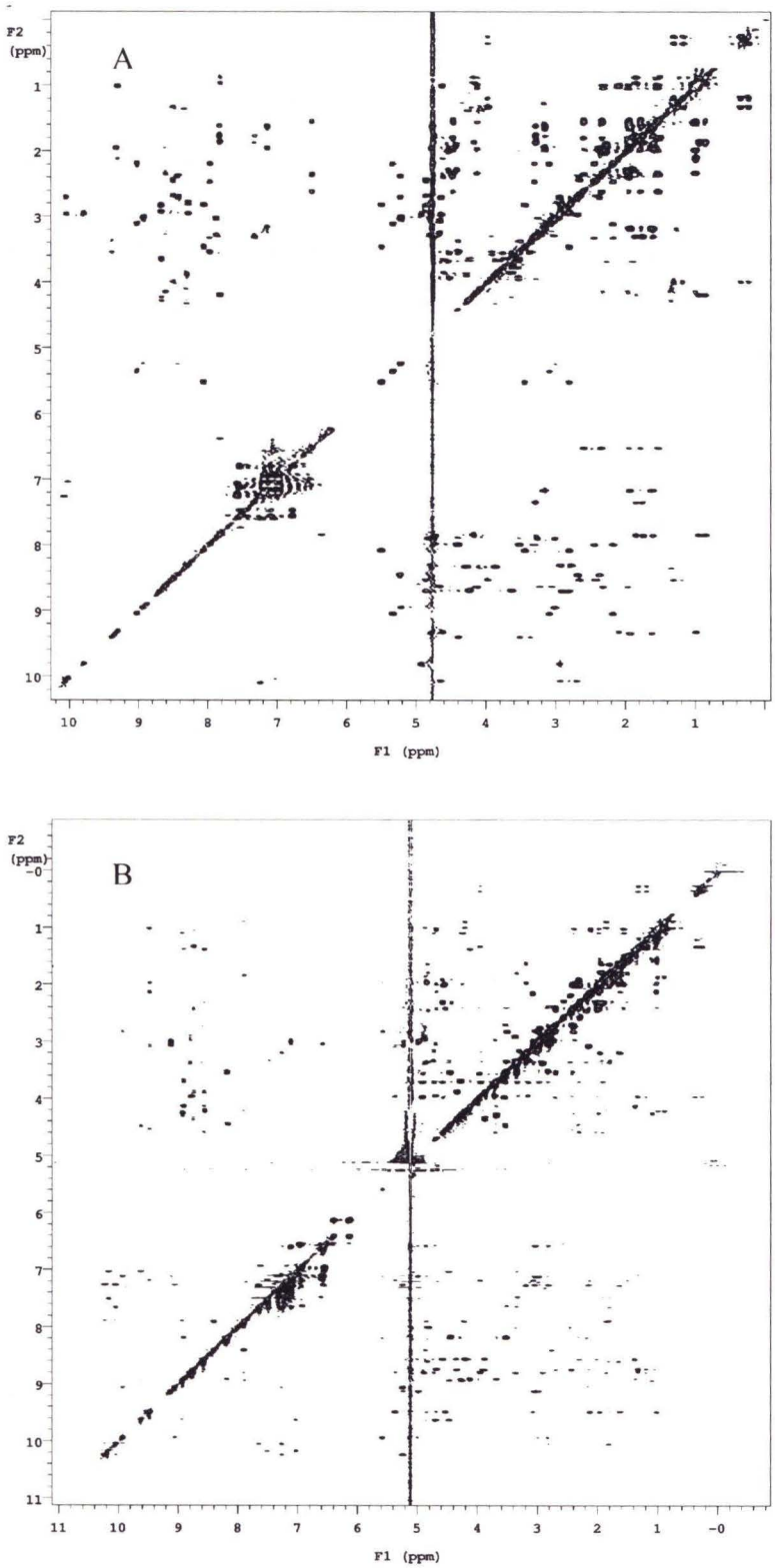


Figure 9. 2D NMR spectra of reg9d at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) *wg*TOCSY, (B) *wg*NOESY.

3.6 Molecular model of reg9d.

A model of the six-Cys (1-4 2-5 3-6 Cys pairing) reg9d was built using the gm9a conotoxin NMR structure as a template (Figure 10). Structural differences are found among these two structures; however, the overall three-dimensional fold remains remarkably similar for these two conotoxins.

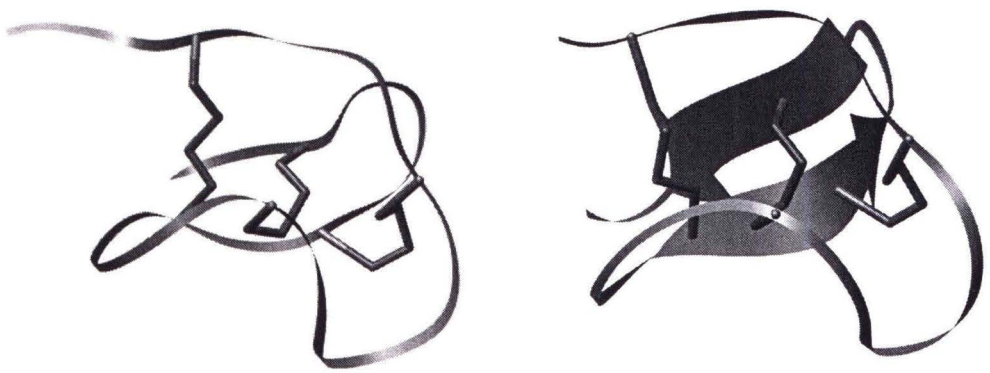


Figure 10. Model of the lowest energy structure of reg9d based on gm9a. Both belong to the P-superfamily of conotoxins.

4. DISCUSSION

The venom of *Conus regius*, a widespread cone snail species that inhabits the tropical Atlantic region and preys upon marine worms, is a complex mixture of conopeptides, which is the typical case of most *Conus* species. The SE-HPLC of the crude venom of *Conus regius* yielded several fractions (Fig. 1). The first seven fractions of the venom were shown to be proteins and peptides of molecular weights that exceeded 2700 Da. All other fractions (8-12) were fractions that contained conopeptides of lesser molecular weight. We chose from these first higher molecular weight fractions 6 and 7 to be further re-fractionated with the objective to give smaller clusters of conopeptides with even closer molecular weights and shape. RP-HPLC of fraction 6 and 7 (Figure 2) revealed the presence of several components from which we isolated several P-superfamily conotoxins. Only reg9a and reg9d were well-resolved components, the rest of the reg9 conotoxins were not well-resolved. The most abundant P-superfamily conotoxin from *Conus regius* is reg9d. The rest of the reg9 conopeptides were isolated in sufficient quantities to reduce/alkylate them and sequence them only. One of the largest components of this fraction is the conotoxin reg9b a peptide with high sequence homology to reg9d. There are other components present in these fractions in lower quantities that could represent other members of the P-superfamily, their molecular weight are similar to each other as evident in the MALDI-TOF MS data, which gives molecular weights ranging from 2800-3300 Da. This combination of SE and RP HPLC provides a facile and efficient procedure for the separation of P-superfamily conotoxins that is far superior to the customary direct separation of the venom by RP HPLC proving that adding a chromatographic step

like Size-Exclusion improves the separation of the components dramatically. Judging from the retention times of these reg9 conopeptides they exhibit different degrees of hydrophobic residues on the outer surface of their molecular shape.

We attempted to acquire 1D and 2D NMR data on all reg9 conopeptides prior to any further characterization steps. All but one of these conopeptides were isolated in low picomolar quantities, as attempts to obtain 1D NMR data even after 4000 scans were unsuccessful. reg9d was the only P-superfamily conotoxin from *Conus regius* isolated in sufficient quantities to obtain TOCSY and NOESY spectra. Enough data was acquired to perform structural analysis on reg9d; this work should be continued at a later time. The use of a high performance 3 mm probe combined with a 1.7 mm capillary insert was necessary for the acquisition of the NMR spectra (See experimental Section) of reg9d. The nano-NMR spectra revealed resonances of readily identifiable residues such as Trp, Arg, His and Ala (Fig. 8). 2D *wg*TOCSY and *wg*NOESY experiments were only possible for reg9d. This information was highly valuable in helping us resolve some residues that were not very clear from the sequencing data obtained. γ-carboxyglutamate (Gla) was one of those amino acids, having a unique spin pattern shown in TOCSY experiments, it was determined that only one Gla was present. Originally, reg9d was thought to contain 7 Gla residues. This came from the fact that there were 7 sequencing cycles that did not show any residues. It has been established that Gla residues can not withstand Edman degradation chemistry [14] and that a small glutamate peak shows up instead. This was the case for one of these sequencing cycles only. This information, in

combination with the conopeptide reg9d not reducing or alkylating, made us believe on the presence of more than one Gla in this peptide. This led the way for the neuroprotection assay studies, since Gla containing peptides (such as conG) have been observed to confer neuroprotection properties. Further analysis of the TOCSY spectra suggested the presence of Cys residues. Amino acid analysis confirmed this fact.

The MALDI-TOF MS spectra of the reg9 conopeptides (Figures 3 and 4) revealed that the majority of these peptides were not single components. With the exception of reg9a, the rest of the reg9 conopeptides showed the presence of the same peptide either hydroxylated (possibly at a Pro residue) or the presence of same peptide with a Pro residue instead of a Hyp residue as minor components. In addition to this some of the peptides did not sequence completely. reg9a has a difference of 126.8 Da for the sequence to be in agreement with the experimental molecular weight. This difference does not match any amino acid, combination of amino acids or post-translational modifications of any amino acid known; therefore, the identity of the residue in position 12 was not elucidated. reg9b sequencing cycles stopped producing amino acid signals after residue 25. The peptide was sequenced again and the same results were obtained. This indicates the possibility that residue 26 is interfering with Edman degradation; however, the difference of 644.6 Da is too great for one amino acid, there must be at least 4 amino acids missing from the sequence and the cycles after residue 26 did not show any discernable peaks to identify any amino acid. reg9b showed in the MALDI-TOF MS spectra peaks with a -16 Da and -32 Da difference,

suggesting the presence of the same peptide perhaps without hydroxylation of the Pro residues. reg9c sequenced to completion; however, the calculated molecular weight did not correlate to the experimental molecular weight. The difference of 71 Da matched the residue alanine; however, evidence of this residue in the sequencing cycle is not clear and unfortunately no NMR data was available that could show the resonances for Ala. reg9c MALDI-TOF MS spectra showed the opposite from reg9b: the presence of minor peaks with +16 Da, +32 Da suggested the possibility of having the same peptide with the Pro residues hydroxylated. reg9d sequence was completed with the combination of amino acid analysis, sequencing data and NMR data. This conopeptide also shows in the MALDI-TOF MS spectra the presence of minor peaks with -16 Da and -32 Da that could be from the same peptide but having Pro residues instead of Hyp. reg9e did sequence to completion. The disulfide bond connectivity was not established due to the difficulty on reducing and alkylating the peptide as observed with reg9d. Samples that did alkylate were not in sufficient quantities to undergo this study.

The reg9 sequences do not have any homology with previously reported P-superfamily conotoxins. They share the cysteine arrangement and the number of residues present in between cysteines. The number of residues in between cysteines seems to be the uniting feature for members of this superfamily. $CX_3CX_5CX_3CX_1CX_{4-5}$ is the preferred interloop composition, as they is shared by tx9a, gm9a, ama9a, BeTXIIb and all the reg9 conopeptides.

Neuroprotection assays were performed using reg9d conopeptide and it is evident that neuroprotective properties are not only conceded to peptides with high Gla content. reg9d only has one Gla and it is highly constrained by disulfide bonds; yet there are indications by the results of the LDH, MTT and calcium influx assays that it can prevent some damage to neuronal cells.

The modeled structure of reg9d against gm9a shows that despite significant sequence diversity these conotoxins have remarkably similar structure. This has been observed in ω-conotoxins like MVIIA and GVIA that share substantial three-dimensional structure similarities despite their sequence divergence. reg9d model has similarities with gm9a NMR structure and even though the overall fold seems similar there are differences in the three-dimensional fold. This could also be explained by the fact that the program used (MODELLER) does not recognize modified amino acids such as Gla and Hyp. These amino acids were intentionally left out of the modeling scheme.

The biological target of members of the P-superfamily has not been established yet. This research has been carried out in hopes that it can contribute to the understanding of the biochemistry and pharmacology of this superfamily.

5. REFERENCES

1. Nicke, A., S. Wonnacott, and R.J. Lewis, *α-Conotoxins as tools for the elucidation of structure and function of neuronal nicotinic acetylcholine receptor subtypes.* European Journal of Biochemistry, 2004. **271**(12): p. 2305-2319.
2. Lewis, R.J. and M.L. Garcia, *Therapeutic potential of venom peptides.* Nature Reviews Drug Discovery, 2003. **2**(10): p. 790-802.
3. McIntosh, J.M., B.M. Olivera, and L.J. Cruz, *Conus peptides as probes for ion channels.* Methods in Enzymology, 1999. **294**(Ion Channels, Part C): p. 605-624.
4. Harvey, A.L., *Toxins 'R' Us: more pharmacological tools from nature's superstore.* Trends in Pharmacological Sciences, 2002. **23**(5): p. 201-203.
5. Miljanich, G.P. and J. Ramachandran, *Antagonists of Neuronal Calcium Channels - Structure, Function, and Therapeutic Implications.* Annual Review of Pharmacology and Toxicology, 1995. **35**: p. 707-734.
6. Newcomb, R. and G. Miljanich, *Neurotoxins of cone snail venoms.* Handbook of Neurotoxicology, 2002. **1**: p. 617-651.
7. Miles, L.E., *Structure of a novel P-superfamily spasmodic conotoxin reveals an inhibitory cysteine knot motif.* J. Biol. Chem., 2002. **277**(45): p. 43033-40.
8. Heading, C.E., *Conus peptides and neuroprotection.* Current Opinion in Investigational Drugs (PharmaPress Ltd.), 2002. **3**(6): p. 915-920.

9. O'Byrne, M., *Taurine-Induced Attenuation of MPP+ Neurotoxicity in vitro: A Possible Role for the GABA-A Subclass of GABA receptors.* J. Neurochem., 2000. **74**: p. 2087-93.

10. Fiser, A.S. and A. Sali, *MODELLER: Generation and refinement of homology-based protein structure models*. 2003, Academic Press: San Diego. p. 463-493.

11. Aird, S.D., *Chromatographic behavior of Bothrops erythromelas phospholipase and other venom constituents on Superdex 75.* Preparative Biochemistry & Biotechnology, 2004. **34**(4): p. 345-364.

12. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.

13. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching.* Analytical Chemistry, 1999. **71**(14): p. 2871-2882.

14. McIntosh, J.M.e., *γ-carboxyglutamate in a neuroactive toxin.* J. Biol. Chem., 1984. **259**(23): p. 14343-46.

Chapter IV

T-Superfamily Conotoxins and Linear Conopeptides from the Venom of *Conus regius*: a Western Atlantic Worm-Hunting Cone Snail Species

ABSTRACT

We report the isolation of two novel T-1- conotoxin-like peptides (reg5a and reg5b), a annular conopeptide with a single disulfide bond (regJ1204) and three linear conopeptides (regG0708, regH0510b, regH0704) from the venom of *Conus regius,* a wide spread worm-hunting cone snail from the Western Atlantic Ocean. reg5a and reg5b (amidated) share a conserved arrangement of cysteine residues (--CC--CC--); however, they share little or no sequence homology among them and in between other reported members of the T-superfamily. The one feature that is shared by reg5b and tx5a from *C. textile* is the presence of two γ-carboxyglutamate residues, before and after the first and second cysteines. The rest of the sequences are non-disulfide-rich conopeptides; such as regJ1204, which contains a single disulfide bond similar to reported contryphans and conopressins; however, it shares no sequence homology to them. The linear conopeptides isolated have no sequence homology to any conopeptide reported and no hits were observed when trying to match these

sequences to any sequence reported. One of these linear conopeptides (regH0704) contains a bromotryptophan residue (W^+) as well of a γ-carboxyglutamate residue. This group of T-superfamily and non-disulfide-rich conopeptides surprisingly shares a large number of hydroxylated residues among them. This was observed in other conopeptides families within *Conus regius*, and it provides an insight into the strategy for venom optimization of this snail. NMR information was obtained for reg5a, reg5b and regH0704 only, the amounts obtained for the rest of the conopeptides were not sufficient for NMR experiments.

1. INTRODUCTION

The T-superfamily conotoxins are found in all three major feeding types of cone snails. In the T-superfamily we find two types of peptides that contain two disulfide bonds but different Cys frameworks [4]. The T-1 conotoxins were the first ones described in this superfamily (framework V in the original nomenclature, scaffold T-1 in the new nomenclature [5]). These conotoxins contains two pairs of adjacent Cys residues separated by five amino acids (four in toxin tx5a/e-TxIX): $CC(X)_5CC$, where ''X'' denotes one non-Cys residue, with disulfide connectivity of the type I–III, II–IV [9, 10]. The T-2 conotoxins (framework X in the original nomenclature, scaffold T-2 in the new nomenclature [5]) have the structure: $CC(X)_4C(X)_2C$, with disulfide pairing of the type I–IV, II–III [11, 12]. T-1 conotoxins have been found in piscivorous (*C. purpurascens*), vermivorous (*C. imperialis*), and molluscivorous (*C. textile, C. gloriamaris, C. aulicus, C. marmoreus*) and now in vermivorous cone snails *(C. regius)*. These conotoxins have remarkably divergent sequences, and some of them have a variety of post-translational modifications [9, 12]. Most of the T-1 conotoxins isolated so far produce biological effects on fish [10], and one of them (tx5a/e-TxIX, from *C. textile*) also has an effect on mice [9, 10] and mollusks [9]. This latter toxin may target pre-synaptic Ca^{2+} channels or G protein-coupled receptors [9]. To date, T-2 conotoxins have been found only in the molluscivorous species *C. marmoreus* [11, 12]. Taking into account that the four T-2 conotoxins known so far come from a single species, they seem to be a less heterogeneous group; although, all of its members have additional post-translationally modifications (beyond the cysteine bridges).

The widespread distribution of T-superfamily conotoxins suggests that they may have important physiological functions. The post-translational modifications seem to play a crucial role in this superfamily. tx5a from *C. textile* has a high numbers of post-translationally modified amino acids and when injected into mice (intra-cranially) elicits hyperactivity and spasticity; however, when analogs of the members of this superfamily (p5a, au5a) lacking these post-translational modifications were injected into mice they had no effect [10].

An assortment of non-disulfide-rich conopeptides are typically found in *Conus* venom. Some of these compounds have been found to target the neurotensin receptor (contulakins), Vasopressin receptors (conopressins), NMDA receptors (conantokin) and RFamide receptors (RFamide receptor), showing vast target diversity [4]. Two linear conopeptides, Conantokin-G and Contulakin-G are currently undergoing clinical trials. Therefore, these conopeptides, while not as studied as their conotoxin counterpart, can provide important neuropharmacological insight in the mode of action of neuronal receptors and ion channels.

In this chapter, we describe the purification and amino acid sequence of two T-1 conotoxin peptides (reg5a and reg5b) and four non-disulfide-rich conopeptides (one single disulfide-bonded and three linear conopeptides) from the venom of the vermivorous species *Conus regius*. Recently, the venom of this species has been shown to contain peptides belonging to the A-superfamily [8] and I-superfamily of conopeptides; however, no information on other conotoxins or conopeptides is

available for this species. These results contribute to broadening of the knowledge of the conopeptides and conotoxins found in the venom of *Conus regius*, and provides an insight into the capacity of this cone snail to modify its venom components to selectively target its prey.

available for this species. These results contribute to broadening of the knowledge of the conopeptides and conotoxins found in the venom of *Conus regius*, and provides an insight into the capacity of this cone snail to modify its venom components to selectively target its prey.

available for this species. These results contribute to broadening of the knowledge of the conopeptides and conotoxins found in the venom of *Conus regius*, and provides an insight into the capacity of this cone snail to modify its venom components to selectively target its prey.

2. MATERIALS & METHODS

2.1 Specimen collection.

Refer to materials and methods used in Chapter I

2.2 Crude venom extraction.

Refer to materials and methods used in Chapter I

2.3 Peptide purification.

Refer to materials and methods used in Chapter I

2.4 Reduction and alkylation of cysteyl residues.

Refer to materials and methods used in Chapter I

2.5 Peptide sequencing.

Refer to materials and methods used in Chapter I

2.6 Molecular mass determination.

Refer to materials and methods used in Chapter I

2.7 NMR spectroscopy.

NMR spectra were acquired on a Varian Inova 500 MHz instrument equipped with PFG, 3xRF channels and waveform generators. Nanomolar quantities of the native conopeptide directly isolated from the venom (reg5a = 6.8 nmoles, regH0704 = 3.2

2. MATERIALS & METHODS

2.1 Specimen collection.

Refer to materials and methods used in Chapter I

2.2 Crude venom extraction.

Refer to materials and methods used in Chapter I

2.3 Peptide purification.

Refer to materials and methods used in Chapter I

2.4 Reduction and alkylation of cysteyl residues.

Refer to materials and methods used in Chapter I

2.5 Peptide sequencing.

Refer to materials and methods used in Chapter I

2.6 Molecular mass determination.

Refer to materials and methods used in Chapter I

2.7 NMR spectroscopy.

NMR spectra were acquired on a Varian Inova 500 MHz instrument equipped with PFG, 3xRF channels and waveform generators. Nanomolar quantities of the native conopeptide directly isolated from the venom (reg5a = 6.8 nmoles, regH0704 = 3.2

nmoles, reg5b, regH0510b, regJ1204 and regH0704 conopeptides were below 1 nmole and no NMR spectra was obtained even after 4000 scans). For further details please refer to materials and methods used in Chapter I.

2.8 Nomenclature.

In this publication we adopted a nomenclature of three letters to designate *Conus* species, because the one or two letter nomenclature currently in place will not be enough to describe the large number of different non-fish-hunting species, especially those with similar first letter names. We decided to use the three letters "reg" to name the peptides from *C. regius.* Arabic numbers were used to represent the disulfide framework; 5 has already been assigned to the members of the T-superfamily. The letter after the framework number indicates the order of elution on RP-HPLC. For those conopeptides that do not fall into any classification we have maintained the nomenclature system used in our laboratory. The first three letters correspond to the species "reg", the next letter corresponds to the different batches, i.e. "H". The numbers that follow belong to the fraction number from SE chromatography, i.e. "07" and the next pair of numbers belongs to the fraction number from RP chromatography, i.e. "04". Any letters in between these numbers correspond to fractions coming from further separations coming from SE or RP chromatography. The example conopeptide would then read "regH0704".

3. RESULTS

3.1 Peptide purification.

The venom of *Conus regius* is an extremely complex mixture of peptides and proteins whose direct separation is shown (Figure 1B Chapter I). More than 100 fractions can be obtained from this separation. However, most of these fractions show multiple components, for which we were required to adopt an improved separation scheme that includes a prefractionation step using size exclusion chromatography on a Superdex-30 column, followed by a refined peptide-optimized size exclusion step on a Superdex Peptide column. The resulting fractions are then separated by reversed phase on a peptide-optimized C18 Vydac Everest column. Most of the resulting fractions are single-component and were subsequently analyzed by mass spectrometry (MALDI-TOF and ESI-Q-TOF), NMR spectroscopy and peptide sequencing by Edman degradation chemistry. Size exclusion chromatography of reg5 conopeptides is shown on Figure 1; subsequent RP-HPLC proceeded as shown in previous chapters. Size exclusion and RP-HPLC for regG0708 are revealed in Figure 2. The last two linear conopeptides were obtained from batch H, whose size exclusion and RP-HPLC chromatograms are shown in Figure 3 and Figure 4. Using this methodology, the analysis of the most significant components of the venom of *C. regius* can be sequenced (Conopeptidome) and the components of the venom can be grouped in the different families of conopeptides. The reg5 peptides were separated by the Superdex-30 in spite of their similar size. This is not entirely unexpected, since this column is also known to partition analytes by hydrophobic interactions [13]. The non-disulfide-rich conopeptides were also separated by size exclusion; however, a

wider range of molecular weights is observed for this group and separation was expected.

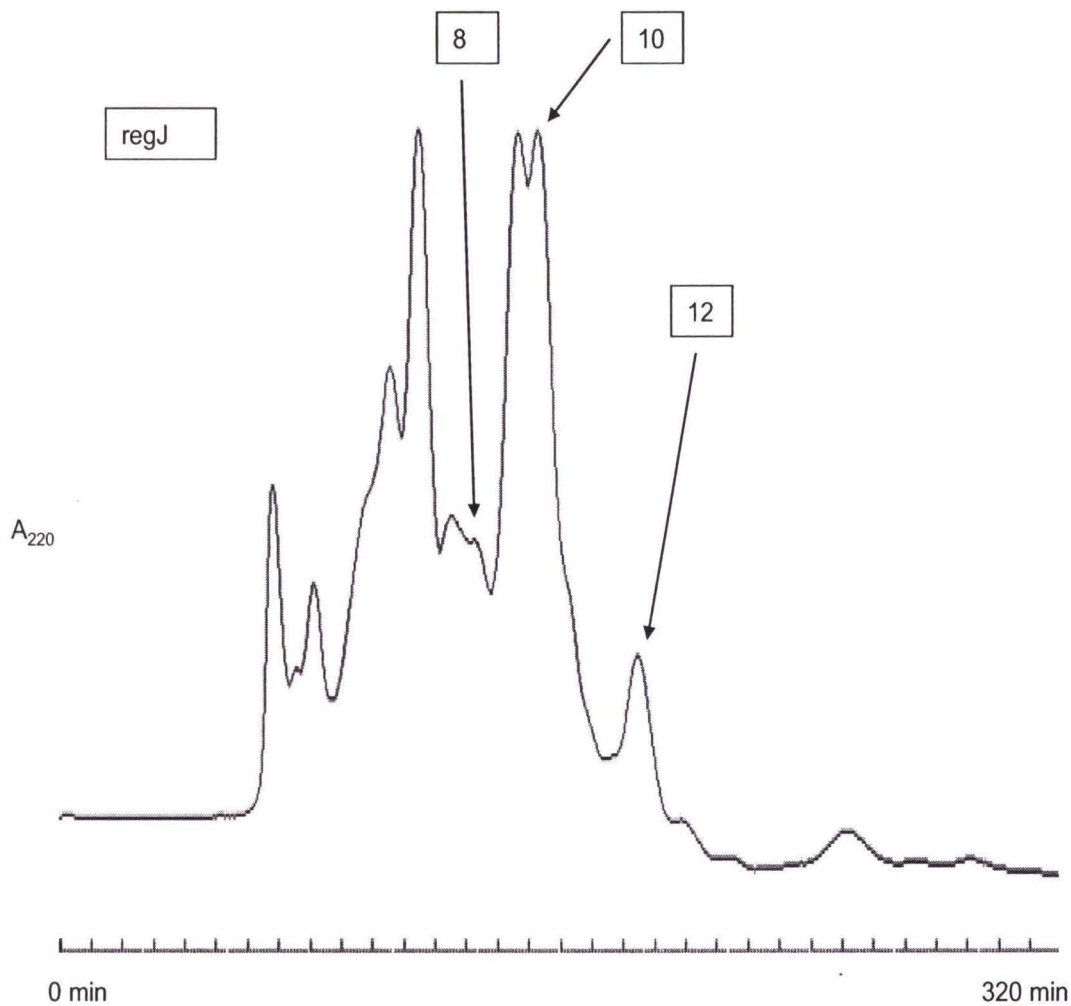


Figure 1. Size exclusion chromatogram of *C. regius* batch J dissected venom. 60 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 8, 10 and 12 were chosen for further separation.

wider range of molecular weights is observed for this group and separation was expected.

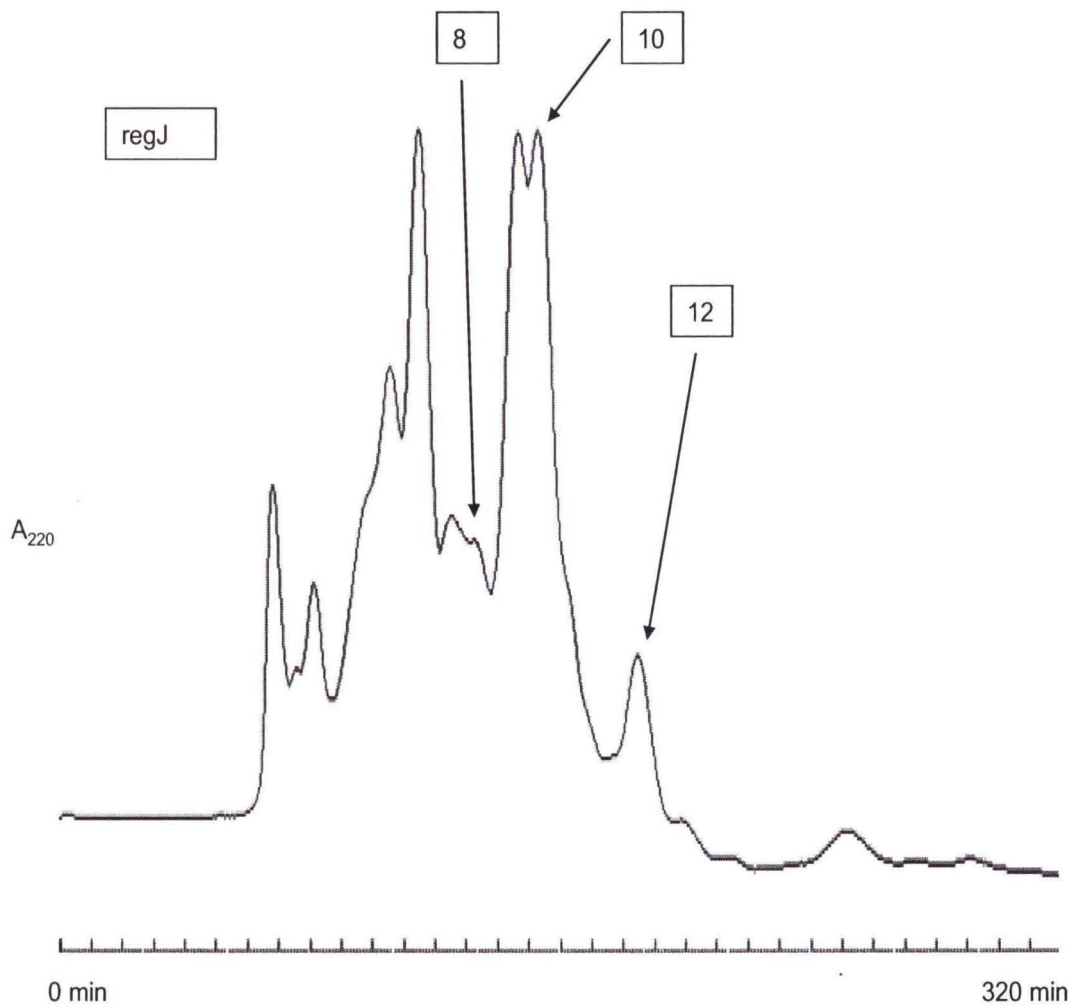


Figure 1. Size exclusion chromatogram of *C. regius* batch J dissected venom. 60 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 8, 10 and 12 were chosen for further separation.

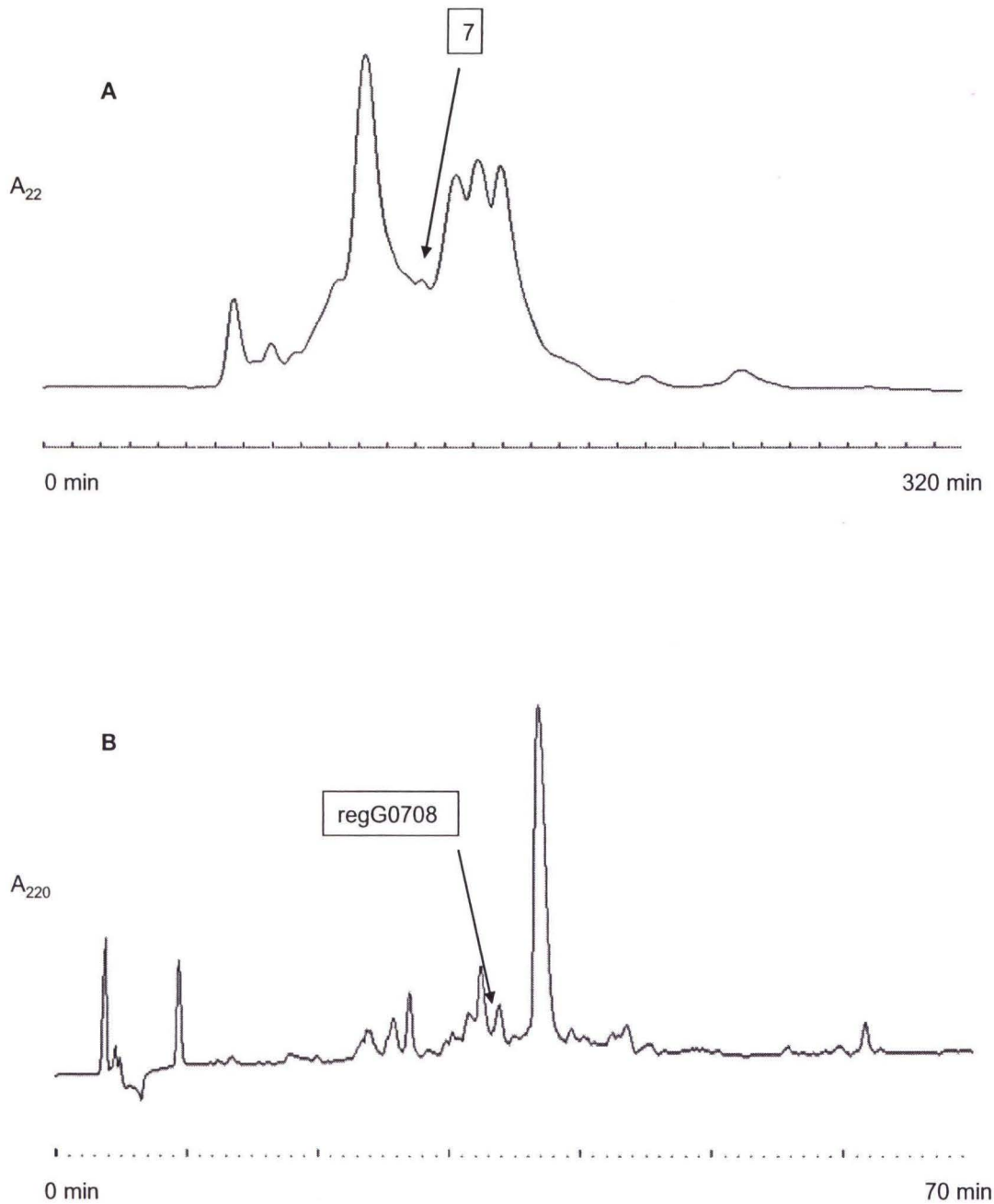


Figure 2. regG0708 conopeptide isolation. Crude venom from *C. regius* batch G was separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution (A), fraction 7 was further refined under RP-HPLC chromatography (B). A linear gradient from 0 to 100% solution B in 100 min at 1ml/min was employed. HPLC solutions were: 0.1% (v/v) TFA in water (A) and 0.1% (v/v) TFA in 60% (v/v) aqueous MeCN (B).

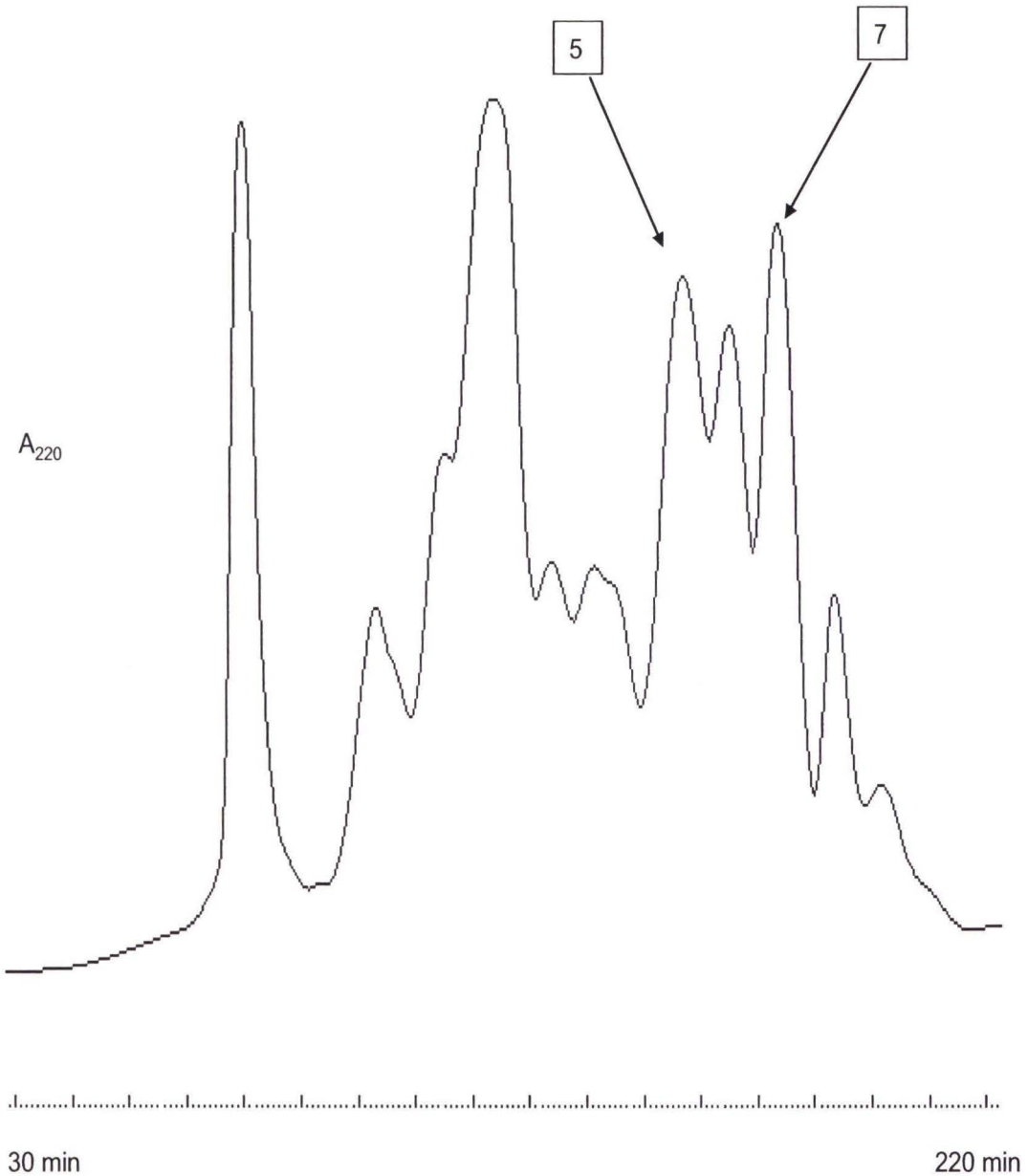


Figure 3. Size exclusion chromatogram of *C. regius* batch H dissected venom. 60 mg of crude venom separated on a Superdex 30 column using a 0.1 M ammonium bicarbonate solution. Fraction 5 and 7 were chosen for further separation.

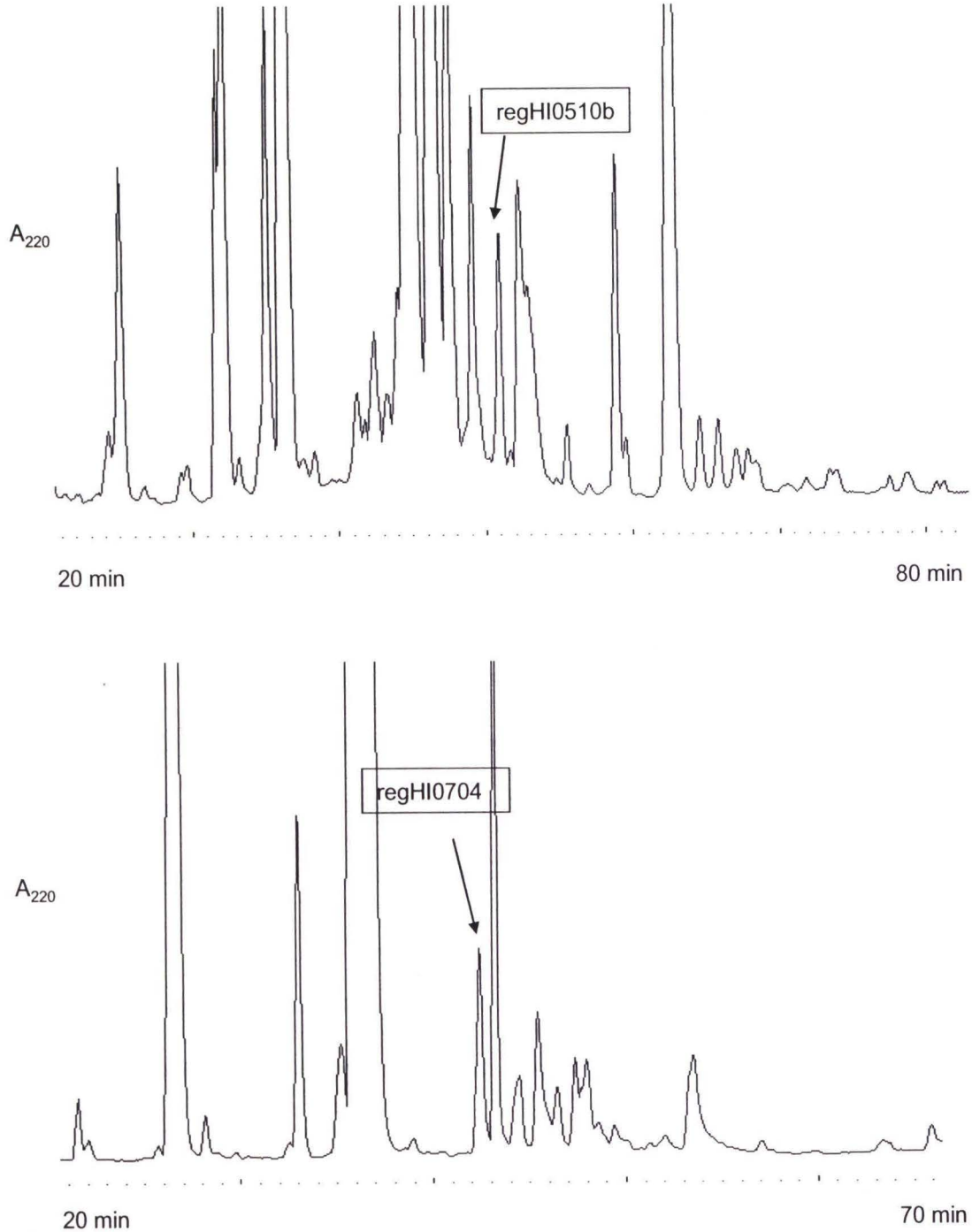


Figure 4. Semi-preparative RP-HPLC chromatogram showing the location of regH0510b and regH0704 peptides from SE fractions 5 and 7. A linear gradient from 0 to 100% solution B in 100 min at 1 ml/min was employed. HPLC solutions were: 0.1% (v/v) TFA in water (A) and 0.1% (v/v) TFA in 60% (v/v) aqueous MeCN (B).

3.2 Reduction/Alkylation and peptide sequence determinations.

The purified peaks were subjected to reduction with DTT and alkylation with iodoacetamide. Mass spectrometry of the reduced/carboxymethylated peptides and of native peptides showed a mass difference consistent with the presence of four cysteine residues for reg5a and reg5b. regJ1204 showed a mass difference consistent with the presence of two cysteine residues. regG0708, regH0510b and regH0704 did not reduce or alkylate after repeated attempts. The reduced/alkylated (reg5a, reg5b and regJ1204) and unreduced (regG0708, regH0510b and regH0704) conopeptides were sequenced by Edman degradation. It is worth noticing that there is no sequence homology between conotoxins reg5a and reg5b; however, they share the same cysteine arrangement. reg5a has three post-translationally modified amino acids (Hyp in positions 2, 4 and 11) and it is not amidated at the C-terminus. reg5b has three post-translationally modified amino acids (Hyp in position 3 and 2 Gla residues in position 2 and 8) and it is amidated at the C-terminus.

regJ1204 did not sequence to completion as residue seven could not be identified. regG0708 sequenced to completion and two post-translationally modified amino acids (γ-carboxyglutamate) were found at positions 17 and 18. This peptide is not amidated at the C-terminus. regH0510b sequenced completely and what is remarkable about this conopeptide is the high percentage (>50%) of residues that have a hydroxyl group attached to the side chain. regH0704 was not fully sequenced, as residues in positions 6, 12, 13 and 18 were not identified. We did, however, identify a γ-carboxyglutamate residue in position 3 and a bromotryptophan in position

9. The bromotryptophan residue was identified by a peak with a retention time higher to that of any of the standards in the sequencing cycle. Once again hydroxylation of residues is present among the peptide components in the venom of *C. regius*. Amino acid residues such as Thr, Ser, Tyr and Hyp that have side chains with OH groups in them are distributed throughout the sequence of these conopeptides.

Hydroxyprolines were present in sequences reg5a, reg5b and regH0510b, these residues were evident by presence of two new peaks in the same cycle corresponding to the different isomers of hydroxyproline. Additionally, a BLAST search [14] of the databases (Swissprot/EMBL, PIR, PDB and nrdb95) did not show any significant sequence homology to reported proteins and peptides. The sequences of the reg5 conotoxins and non-disulfide-rich conopeptides from *Conus regius* are reported in Table 1.

Name		Sequence	Species	Prey
T-SUPERFAMILY				
reg5a		DOVOY**CC**IRNOL**CC**	*C. regius*	v
reg5b		IγO**CC**PRγDH**CC***	*C. regius*	v
p5a		G**CC**PKQMR**CC**TL*	*C. purpurascens*	p
au5a		F**CC**PFIRY**CC**W	*C.aulicus*	
au5b		F**CC**PVIRY**CC**W	*C.aulicus*	
sr5a		IINW**CC**LIFYQ**CC**	*C.spurius*	v
tx5a		γ**CC**γDGW^{+}**CC**T$^{\$}$AAO	*C. textile*	m
Tx5.1	(cDNA)	**CC**QTFYW**CC**VQGK	*C. textile*	m
Gm5.1	(cDNA)	L**CC**VTEDW**CC**EWW	*C. gloriamaris*	m
Gm5.2	(cDNA)	V**CC**RPVQD**CC**SGK	*C. gloriamaris*	m
SINGLE S-S				
regJ1204		**C**YYNS-**C**XK	*C. regius*	v
Contryphans				
Contryphan-R		G**C**O̲W̲EPW**C***	*C. radiatus*	p
Bromocontryphan		G**C**O̲W̲EP~~W~~**C***	*C.radiatus*	p
Contryphan-Tx		G**C**O̲W̲QPY**C***	*C.textile*	m
Leu-Contryphan-Tx		-**C**VL̲YPW**C**	*C.textile*	m
Contryphan-Vn		GD**C**P̲W̲KPW**C***	*C. ventricosus*	v
Conopressins				
Lys-conopressin-G		**C**FIRN-**C**PKG*	*C. geographus*	p
Arg-conopressin-S		**C**IIRN-**C**PRG*	*C. striatus*	p
LINEAR CONOPEPTIDES				
regG0708		SDRGKTGVLSAGTARTγγVAFTA*	*C. regius*	v
regH0510b		TFSGQAYTOGSFOEGYFGYGYTYSLS	*C. regius*	v
regH0704		WLγAGXFGW^{+}LMAXXSPAEXYSNNKDQIIGG	*C. regius*	v

Table 1. T-superfamily and linear conopeptides sequences from *C. regius* and other *Conus* species.

3.3 Mass spectrometry of purified peptides.

Mass spectrometry carried out using MALDI-TOF in the reflector mode (M/ΔM resolution ~ 10,000) yielded the following monoisotopic molecular ions: reg5a = 1639.8 Da (Figure 5), reg5b = 1503.9 Da (Figure 6), regJ1204 = 1029.3 Da (Figure 7A), regG0708 = 2408.6 Da (Figure 7B), regH0510b = 2879.2 Da (Figure 8A) and regH0704 = 3302.8 Da (Figure 8B). Mass analysis of the reduced/carboxymethylated peptides and the native peptides showed a mass difference consistent with the presence of four cysteine residues for reg5a, reg5b. The masses obtained for the peptides were for the most part in agreement with the calculated theoretical monoisotopic values determined for the assigned sequences. regJ1204 and regH0704 did not fully sequence. Repeated attempts were carried out with no success. regJ1204 has a mass difference of 151.1 Da from the theoretical MW and the experimental MW. Its MALDI-TOF MS spectra also shows the presence of the molecular ion $+Na^+$ and $+K^+$. regH0704 has a 381.6 Da mass difference from the theoretical MW and the experimental MW. The MALDI-TOF spectra of regH0704 shows smaller peaks of the molecular ion +16 Da suggesting the possibility of having hydroxylation of a residue like Proline. regG0708 MALDI-TOF spectra shows a peak of the molecular ion -18 Da indicating a loss of H_2O. It also shows a smaller peak of the molecular ion -42 Da, indicating the presence of one of its γ-carboxyglutamate residues. Although it is not a full 44 Da loss that one would expect from the loss of a CO_2 molecule coming from the decarboxylation of γ-carboxyglutamate, it is highly likely that it comes from this residue, which will then be in agreement with the calculated MW. reg5a MALDI-TOF MS spectra shows the molecular ion +22 Da and +39 Da indicating the

presence of Na^+ and K^+ ions. The MALDI-TOF MS spectra of reg5b shows as the most prominent ion the molecular ion -88 Da, which is consistent with the loss of $2CO_2$ molecules coming from the two γ-carboxyglutamate residues. regH0510b MALDI-TOF spectra also shows smaller peaks with a difference of +14 Da which indicates the possibility of methylation of a residue like serine. From the MALDI-TOF MS data obtained and without the full sequence giving us the theoretical MW, the difference in the calculated molecular weights and the experimental molecular weights from MALDI-TOF could not be calculated for the regJ1204 and regH0704 conopeptides. This results in the inability to determine if the peptides were amidated at the C-terminus. The calculated MW for reg5a and regH0510b were determined, and it was established that the peptides were not amidated at the C-terminus. On the other hand, regG0708 and reg5b were determined to be amidated at the C-terminus. The calculated molecular weights (MWcal) were obtained using Protein Prospector [15].

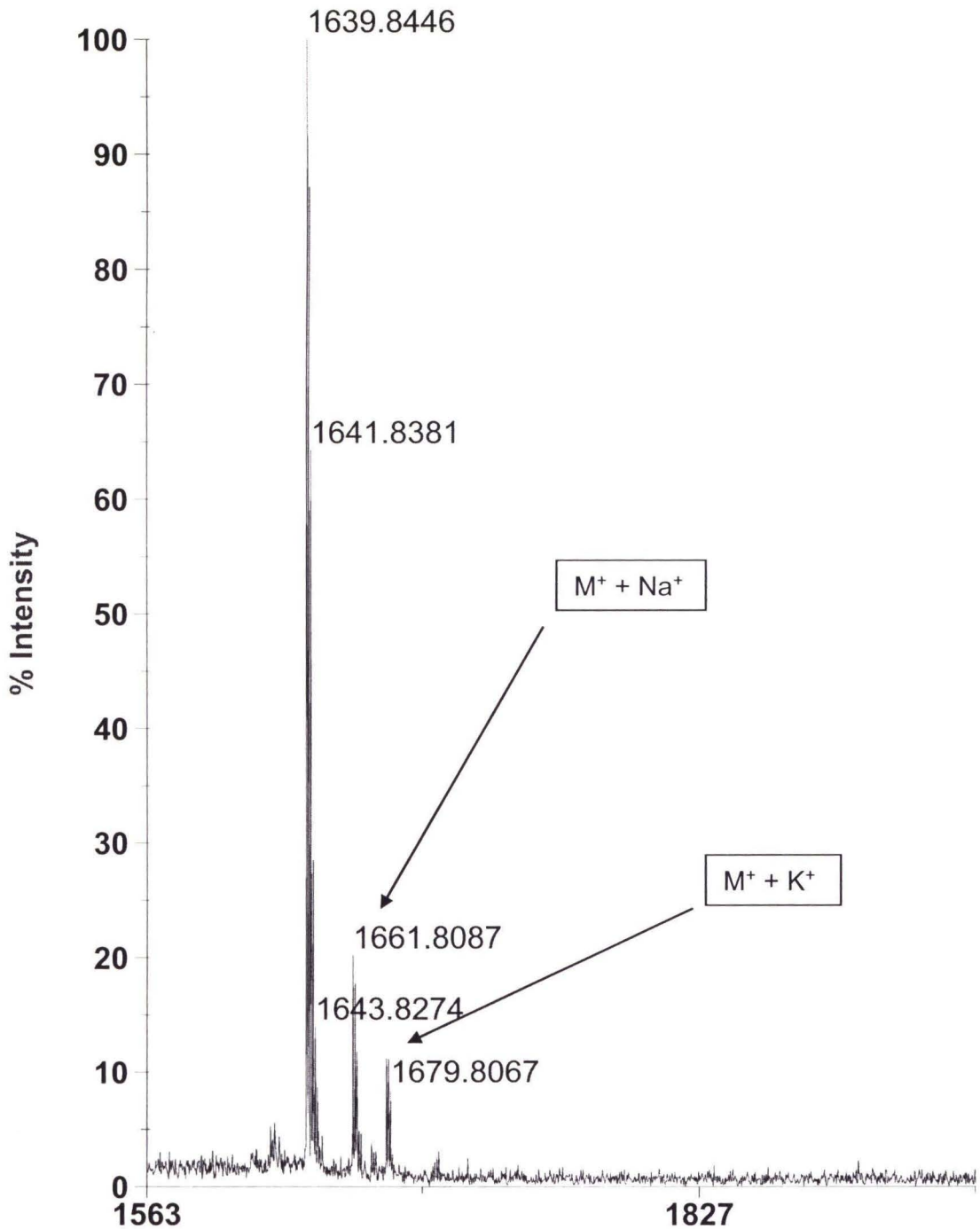


Figure 5. MALDI-TOF MS of reg5a.

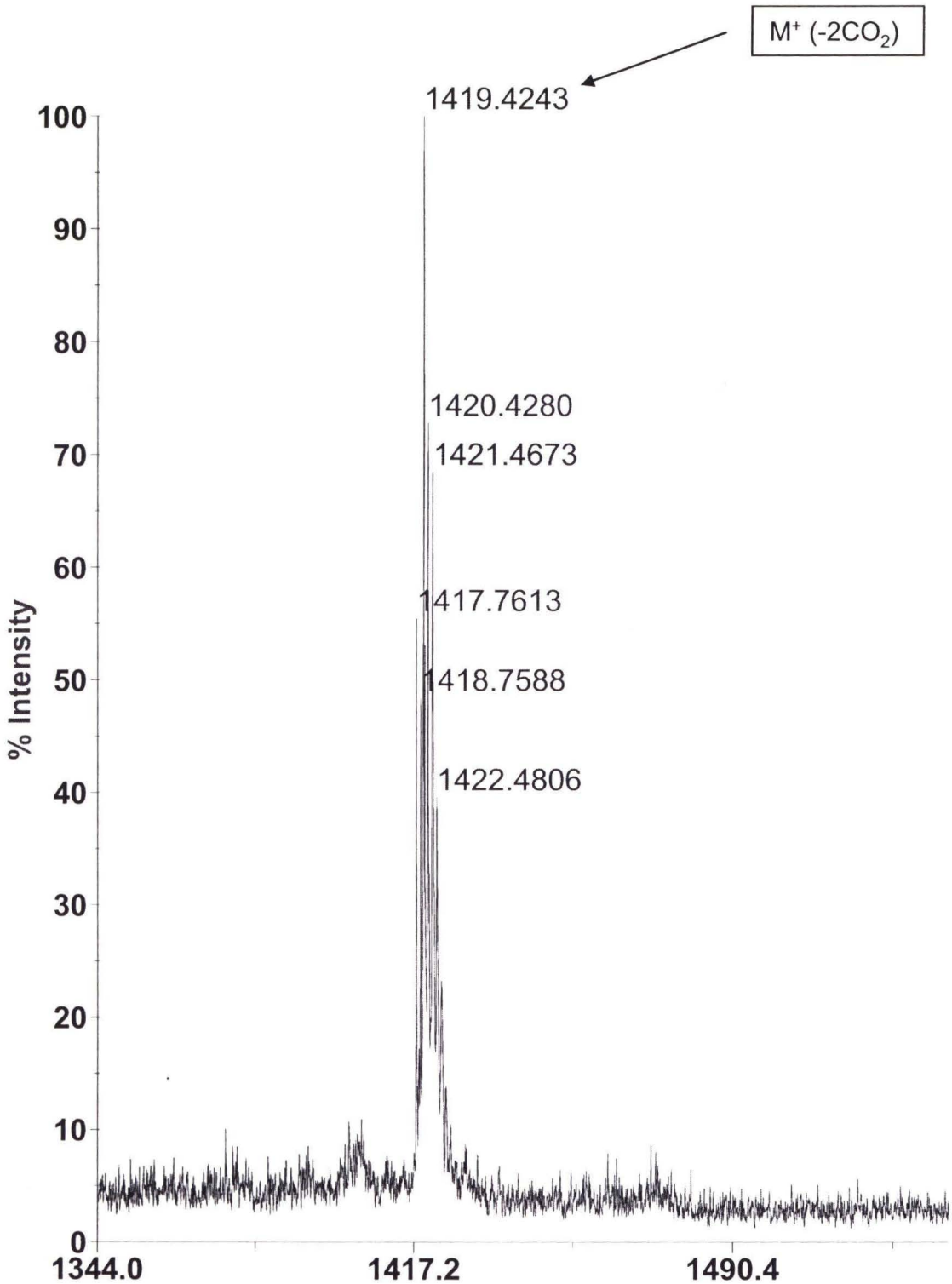


Figure 6. MALDI-TOF MS of reg5b.

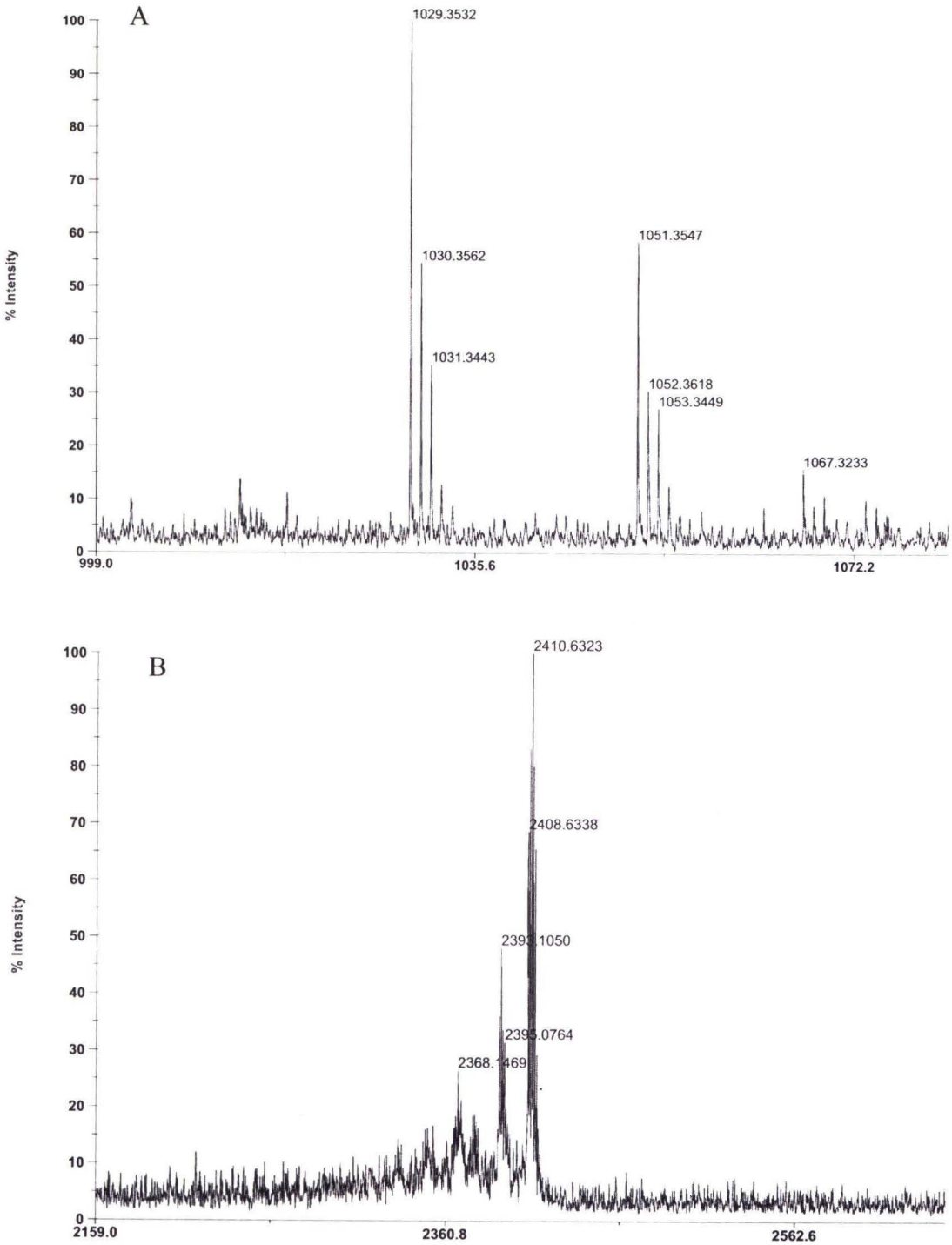


Figure 7. MALDI-TOF MS of regJ1204 (A) and regG0708 (B)

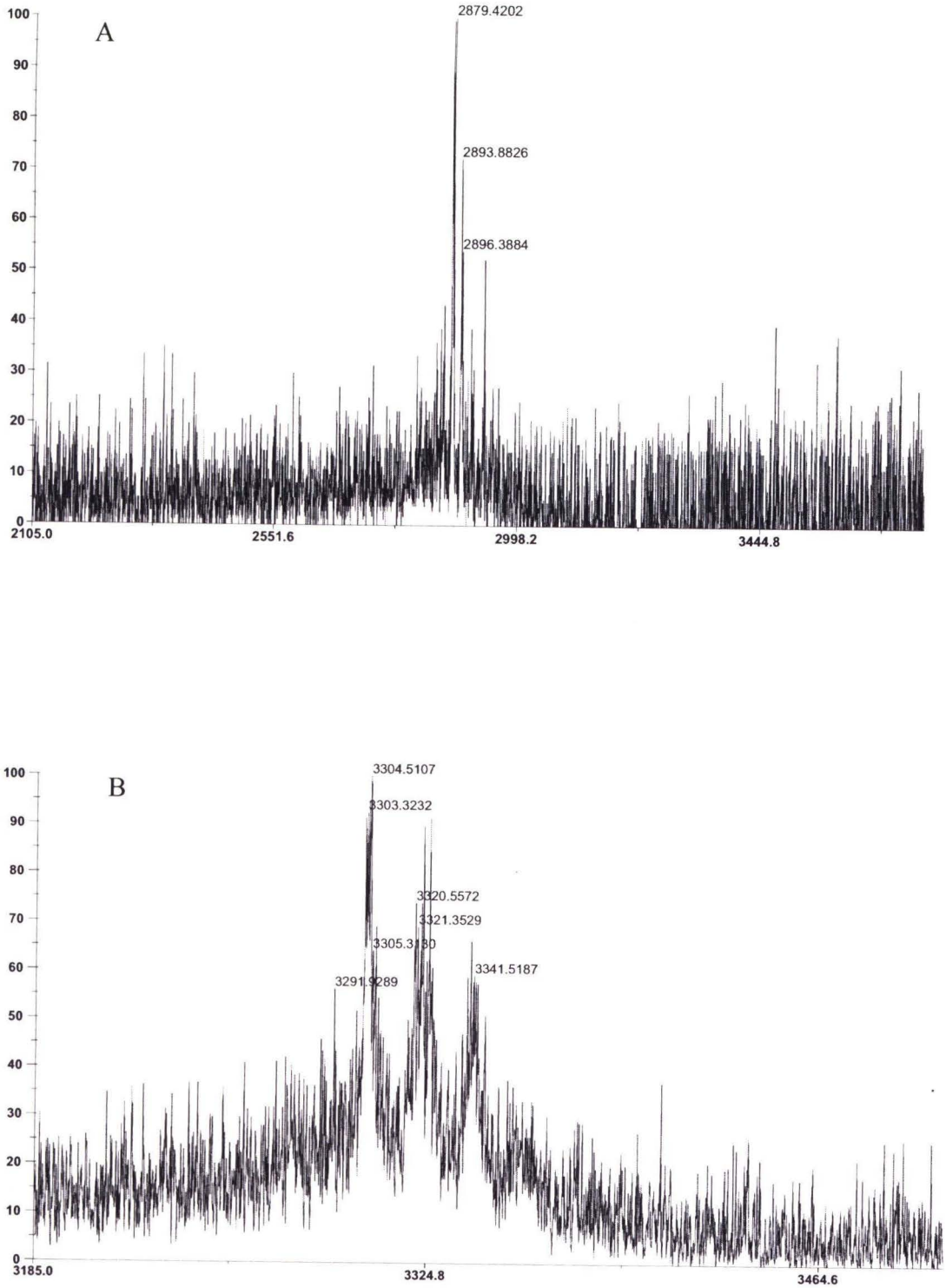


Figure 8. MALDI-TOF MS of regH0510b (A) and regH0704 (B)

3.4 NMR spectroscopy.

We were able to obtain 1D ^{1}H NMR spectra for reg5a, reg5b and regH0704, and 2D *wg*TOCSY spectra for reg5a and regH0704. These compounds were directly isolated from the venom of the cone snails in nanomolar quantities. A 2D *wg*TOCSY spectrum was obtained for reg5b and regH0704. Figure 9 shows the ^{1}H NMR spectra of reg5b at 25°C. It is evident that there are contaminants and that there are low quantities of the sample. The TSP peak at zero ppm is the equivalent to 4 nmoles of TSP in protons, which indicates that there is not enough concentration to try 2D experiments. Figure 10 represents the ^{1}H NMR and 2D *wg*TOCSY spectra experiments of reg5a. Both the ^{1}H NMR (Figure 10A) and *wg*TOCSY spectra (Figure 10B) of reg5a reflect spectra not suitable for structural analysis. The 2D *wg*TOCSY data obtained for reg5a is not good enough to show some of the characteristic cross peaks that identify different amino acids. 2D *wg*NOESY spectra could not be acquired due to low concentration of sample.

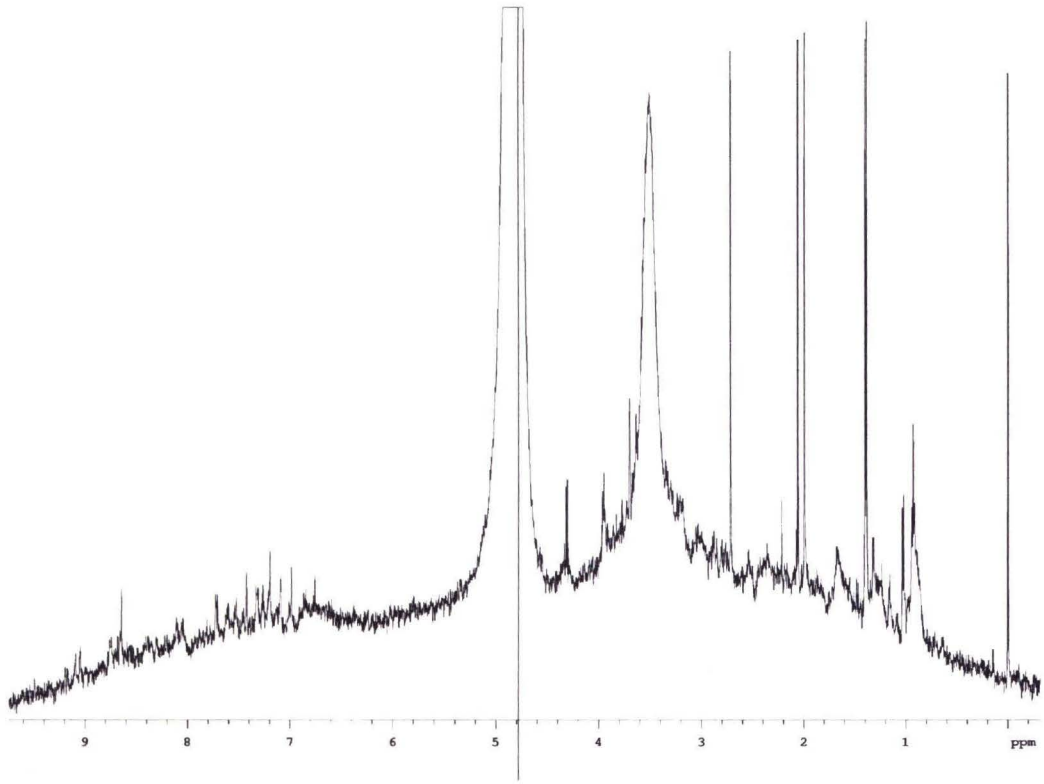


Figure 9. ^{1}H NMR spectrum at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe of reg5b.

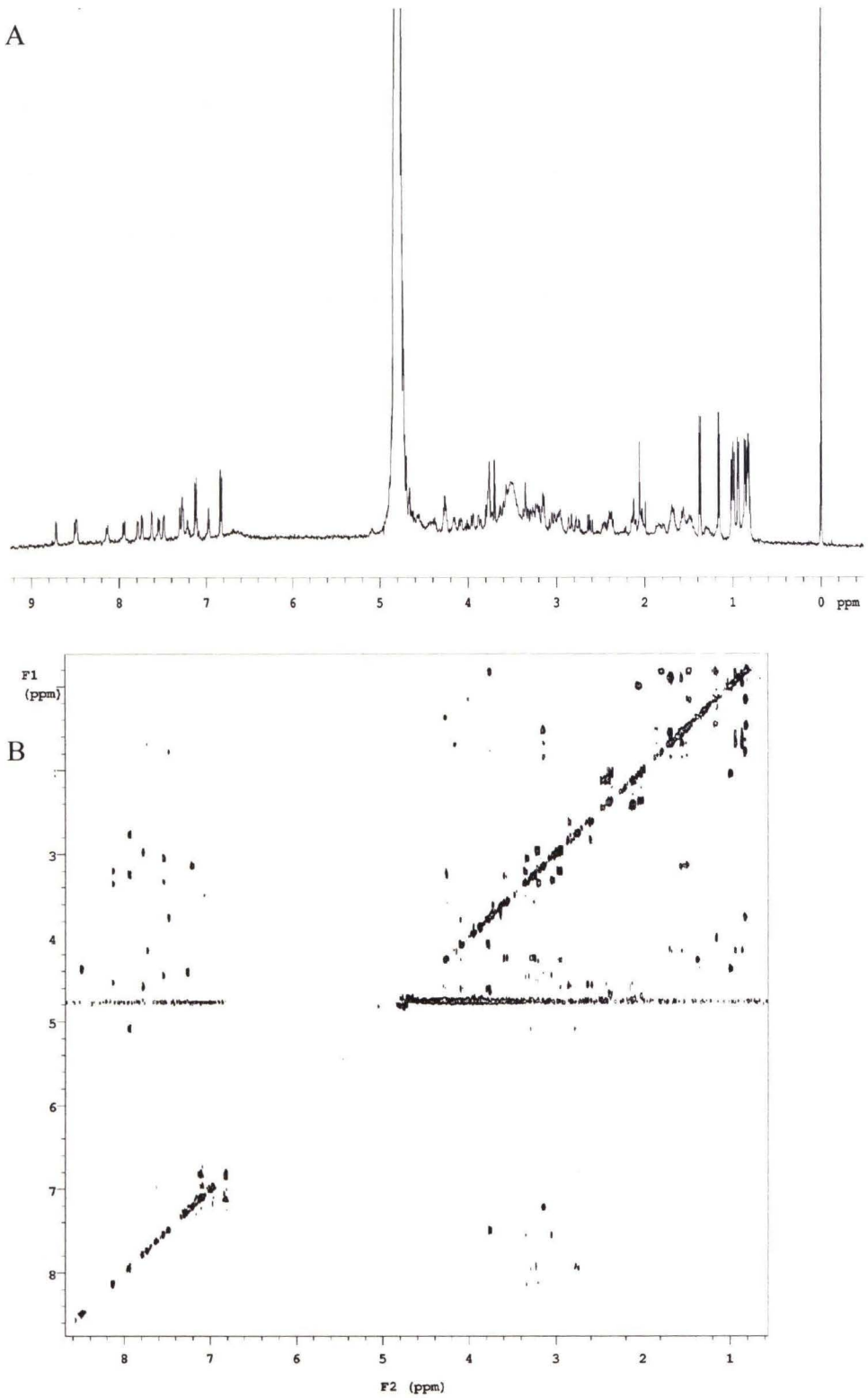


Figure 10. NMR spectra of reg5a at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) 1D ^{1}H NMR (B) *wg*TOCSY.

^{1}H NMR (Figure 11A) and *wg*TOCSY (Figure 11B) spectra were acquired for regH0704. The ^{1}H NMR spectra of regH0704 shows that there was enough sample to obtain a *wg*TOCSY spectrum; however, this amount was not sufficient to acquire a *wg*NOESY spectra. The 1D also shows the characteristic peaks around 10 ppm for the amino acid tryptophan, suggesting that more than one tryptophan residue may be present in the sequence. The *wg*TOCSY spectrum has very weak signal and little information on the amino acids present can be gathered.

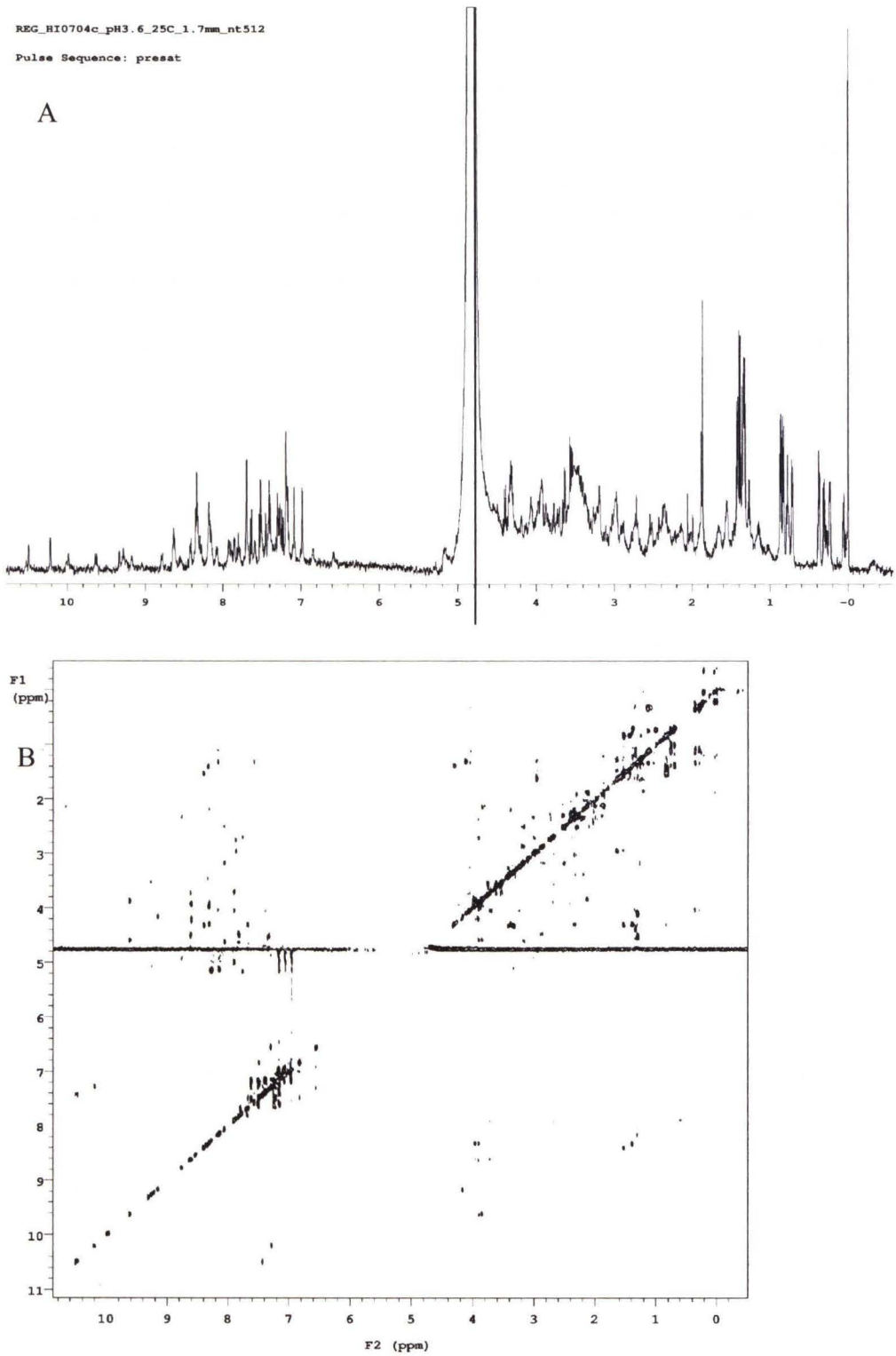


Figure 11. NMR spectra of regH0704 at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe. (A) ^{1}H NMR (B) *wg*TOCSY.

4. DISCUSSION

The reported sequences of reg5 conopeptides have little or no sequence homology to other members of the T-1 superfamily; however, they do share their cysteine arrangement. The two T-superfamily conotoxins found in *Conus regius* are quite different. reg5a has three Hyp residues while reg5b only has one and it is not in the same position. reg5a has no Gla residues, while reg5b has two Gla residues. It is evident that post-translational modifications are important features of *Conus regius'* biochemical arsenal of and that their success as predators depends on them. T-superfamily members are found throughout all three types of feeding habits, fish-hunters (*C. purpurascens*), worm-hunters (*C. regius*) and mollusk-hunters (*C. textile*). The reported sequences of the T-superfamily come from cone snail species that target different prey; this could explain the sequence divergence and little homology displayed by the conotoxins of this superfamily. Even among neighboring species as is the case for *C. regius* and *C. spurius* that share the Florida coast, their T-superfamily peptides share no sequence homology. Both target worms, but not the same kind. T-1 conotoxins are extremely diversified in both their sequences and post-translational modifications, so one would expect a wide array of biological functions.

The regJ1204 conopeptide from *Conus regius* has only one disulfide bond. This disulfide arrangement is also shared by contryphans and conopressins. There is no sequence homology among members of these families and regJ1204. Contryphans have characteristic post-translational modifications that are not present in regJ1204. There are only two reported members of the conopressin family; they share no

sequence homology with regJ1204. Conopressins and contryphans have amidated C-terminus; amidation regJ1204 at the C-terminus was not been established since one residue has not been identified.

regG0708 is a linear peptide that has 2 vicinal Gla residues in their sequence. Conantokins are an example of linear peptides that contain Gla residues side by side in the sequence. However, the similarities end there as they share no sequence homology. regG0708 has a good number of residues with hydroxyl groups (S,T) through out the sequence, a characteristic feature of the venom of *Conus regius*. The regH0510b conopeptide might be part of the preferential targeting strategy used by *Conus regius,* where hydroxyl groups are prevalent in its venom components, presumably to refine neuronal targeting. More than 50% of the residues of regH0510b conopeptide contain hydroxyl group (T, S, Y, O), two of which are post-translationally modified (O). A BLAST search revealed no sequence homology to any peptide reported in literature.

regH0704 is a larger linear conopeptide with at least two post-translational modifications, a Gla residue in position 3 and a bromotryptophan in position 9. The presence of bromotryptophan in combination with γ-carboxyglutamate residues has been observed in peptides that induce sleep [16]. regH0704 could be one of those peptides as there are four residues that have not been identified in this sequence. It is evident that the presence of all these post-translational modifications has a specific purpose in the efficiency of the venom, whether to capture prey or defend themselves

from predators. Hydroxylation of residues and the presence of γ-carboxyglutamate residues is a constant feature among the reg5 T-superfamily conotoxins as well as in the linear peptides of *Conus regius*. These features are key characteristics that could lead the way for a better understanding of the way the venom and all of its components work together to effectively reach a goal.

5. REFERENCES

1. Olivera, B.M., et al., *Conotoxins - Targeted Peptide Ligands from Snail Venoms.* ACS Symposium Series, 1990. **418**: p. 256-278.
2. Adams, D.J., et al., *Conotoxins and their potential pharmaceutical applications.* Drug Development Research, 1999. **46**(3/4): p. 219-234.
3. Olivera, B.M. and L.J. Cruz, *Conotoxins, in retrospect.* Toxicon, 2001. **39**(1): p. 7-14.
4. Terlau, H. and B.M. Olivera, *Conus venoms: A rich source of novel ion channel-targeted peptides.* Physiological Reviews, 2004. **84**(1): p. 41-68.
5. Olivera, B.M., *Conus venom peptides: Reflections from the biology of clades and species.* Annual Review of Ecology and Systematics, 2002. **33**: p. 25-47.
6. Livett, B.G., K.R. Gayler, and Z. Khalil, *Drugs from the sea: Conopeptides as potential therapeutics.* Current Medicinal Chemistry, 2004. **11**(13): p. 1715-1723.
7. Jones, R.M. and G. Bulaj, *Conotoxins - New vistas for peptide therapeutics.* Current Pharmaceutical Design, 2000. **6**(12): p. 1249-1285.
8. Ellison, M., *α-regIA: A novel conotoxin that specifically blocks the α9α10 nAChR.* Biochemistry, 2006.
9. Rigby, A.C., et al., *A conotoxin from Conus textile with unusual posttranslational modifications reduces presynaptic Ca^{2+} influx.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(10): p. 5758-5763.

10. Walker, C.S., et al., *The T-superfamily of conotoxins.* Journal of Biological Chemistry, 1999. **274**(43): p. 30664-30671.

11. Balaji, R.A., et al., *λ-Conotoxins, a new family of conotoxins with unique disulfide pattern and protein folding: Isolation and characterization from the venom of Conus marmoreus.* J. Biol. Chem., 2000. **275**: p. 39516-39522.

12. Sharpe, I.A., et al., *Two new classes of conopeptides inhibit the α1-adrenoceptor and noradrenaline transporter.* Nature Neuroscience, 2001. **4**(9): p. 902-907.

13. Aird, S.D., *Chromatographic behavior of Bothrops erythromelas phospholipase and other venom constituents on Superdex 75.* Preparative Biochemistry & Biotechnology, 2004. **34**(4): p. 345-364.

14. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.

15. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching.* Analytical Chemistry, 1999. **71**(14): p. 2871-2882.

16. Jimenez, E.C.e., *Multiple 6-bromotryptophan residues in a sleep-inducing peptide.* Biochemistry, 2004. **43**(38): p. 12343-8.

General Conclusions

As part of our efforts to study the venom composition of cone snail species from the Atlantic Ocean we took on the challenge of establishing the conopeptidome of *Conus regius*, the "crown cone". During this study many batches of *Conus regius* were obtained, even variants of the shell of *Conus regius* like *Conus regius citrinus*, and they were all studied under the same conditions. Originally, direct separation of the venom components was carried out using reversed-phase chromatography. Although we obtained impressive chromatograms due to number of peaks present, the separation left much to be desired as the majority of the single peaks obtained from reversed-phase chromatography were not single components as shown in MALDI-TOF spectra.

We needed an improved separation for the crude venom of *Conus regius*. As a result from this need a crucial size-exclusion chromatography step was added into the separation scheme for the crude venom of cone snail species. A Superdex Peptide (Pharmacia) column was used to evaluate the separation of the crude venom. Immediately, the results showed enhanced performance, as the column was able to fractionate the components of the crude venom in as many as 15 fractions. This fractionation step is primarily based on the molecular weight of the components; however, molecular shape plays a big factor in the separation of peptides by size-

exclusion methods as there are linear peptides that have co-eluted with higher molecular weight components. A third factor that affected the separation of the components through size-exclusion chromatography is hydrophobicity, as hydrophobic interactions have been known to occur between the column material and the peptide delaying the elution of the peptide. Now we had a column that would separate the venom components but can only separate 10 mg at a time as the maximum load. This was not sufficient, as typical amounts of crude venom obtained from one *Conus regius* snail was around 30-70 mg. We decided to utilize a Superdex 30 size-exclusion column that could handle 50-60 mg of crude venom separation at one time. The separation was almost as good as the Superdex Peptide column but could handle higher amounts, so we had found that extra step needed to clean out the high molecular weight components from the lower molecular weight components. Any further refinement of a fraction from the Superdex 30 column would be carried out in the Superdex Peptide column. The next step after size-exclusion chromatography would be reversed-phase chromatography. We carried out reversed-phase chromatography at the semi-preparative level for all fractions coming from size-exclusion chromatography. Any refinement of any fraction from semi-preparative reversed-phase chromatography would be carried out in an analytical reversed-phase column.

The next step was to perform NMR experiments on all fractions separated by reversed-phase chromatography. The problem we encountered here was that the amount of water present in the 3 mm NMR tubes was diluting the peptides

excessively; thus overwhelming the signal corresponding to the peptides. We decided to improve the concentration of our samples and lower the amount of water present in the sample by using a 1.7 mm capillary tube in combination with a high performance 3 mm probe. The results were excellent and samples containing as little as 300 pmoles of peptide were giving ^{1}H NMR spectra, samples over 2 nmoles of peptide were producing TOCSY spectra and samples over 10 nmoles of peptide were producing NOESY spectra with sufficient quality for structural analysis work.

The optimization of the venom separation chromatography and the improved detection of the NMR setup in combination with a MALDI-TOF and peptide sequencer paved the way for the next step, conopeptidomics. Conopeptidomics is a term that was coined to make an analogy with the word Proteomics. "Cono" from *Conus*, "peptid" from peptide and "omics" from the analogy of genomics. Conopeptidomics became the focus of this study, the chance to identify as many sequences from the venom of *Conus regius* and possibly elucidate the structure and function of each sequence. Through the systematic application of the outlined steps we were able to isolate and characterize a total of 30 sequences directly from the venom of *Conus regius* (Table 1). This is the first reported case where that many conopeptides have been isolated and sequenced from the crude venom of a single cone snail species. Members of the M-superfamily, A-superfamily, P-superfamily, T-superfamily, one disulfide-bond and linear peptides have been identified in this study directly from the venom, and in addition to these families others have reported members of the I-superfamily as well. *Conus regius* venom is truly a great example of

peptide diversity and a rich source for post-translational modifications, one could say a "model cone".

The peptides found throughout the venom exhibit many post-translational modifications like hydroxylation of residues, γ-carboxyglutamate residues, bromination of tryptophan, amidation of the C-terminus and disulfide bonds. The most striking feature of the venom of *Conus regius* is the unusual amount of hydroxylated residues distributed throughout the many sequences and families found in the venom. This gives us valuable insight as to how the different components work together to efficiently achieve their goal.

The conopeptidome of *Conus regius* increases the existing library of conopeptides and conotoxins reported in literature by 22%. This contribution is significant since the reported sequences in this work come from a single cone snail species (Table 1).

```
Non-Disulfide-Rich
regG0708          SDRGKTGVLSAGTARTγγVAFTA*
regH0510b       TFSGQAYTOGSFOEGYFGYGYTYSLS
regH0704     WLγAGXFGW+LMAXXSPAEXYSNNKDQIIGG
regJ1204                  CYYNSCXK

A-superfamily
reg1a                      GCCSDORCRYXC*
reg1b                      GCCSDORCKHQC*
reg1c                     DYCCRROOCTLIC*
reg1d                      GCCSDORCRYRC*
reg1e                      GCCSDPRCKHQC*
reg1e'                     GCCSDPRCKHEC*
reg2a                     GCCSHPACNVNNPHIC*

T-superfamily
reg5a              DOVOYCCIRNOLCC
reg5b                IγOCCPRγDHCC*

P-superfamily
Reg9a         GCTGOKCTKDNXCASOCK-CGYYSLCH
Reg9b       RVLCOGM-CDODVGCOAGCY-CHHLL-CXXXX
Reg9c         HCGSK-CFSDDHCPASCP-CAAHFRCVRSA
Reg9d       RVLCOGγ-CDODVGCOAGCY-CHWLR-CRW
Reg9e         VCSGV-CYPAITCNANCK-CGKYFNCIPSS

M-superfamily
reg12a             GCCOOQWCGOD--CTSOCC
reg12b              CCAIRLCNVYL-CGS-CCO
reg12c              CCAFOQWCGAG-CIVOCC
reg12d             LCCOOQXCGOD--CASOCC
reg12e              CCTAL-CSRYH-CL-PCC
reg12f             KCCMRPICT----C--OCCIGP
reg12g             GCCPFPACTHTIICR--CC
reg12h              CCMAL—CSRYH-CL-PCC
reg12i             GCCSOWNCIQLRAC--OCCON
reg12j          GCCSOWNCIQLRAC-G-CC
reg12k             KCCMRPICM----C--OCCIGP
reg12l             RCCPMPGCFAGPFC--PCCPV
```

Table 1. *Conus regius* sequences isolated directly from the crude venom.

The future work for this project would be the structure elucidation and target determination of these conotoxins and conopeptides. Special consideration should be given to elucidating the disulfide-bond connectivity of these conotoxins and identifying the missing residues. The principles used for the conopeptidomics of *Conus regius* can be applied effectively to other cone snail species, making it possible to work diligently with small amounts of venom with the realization that cDNA libraries that often overlook important post-translational modifications.

Appendix A

^{1}H NMR spectra of *Conus regius* sequences

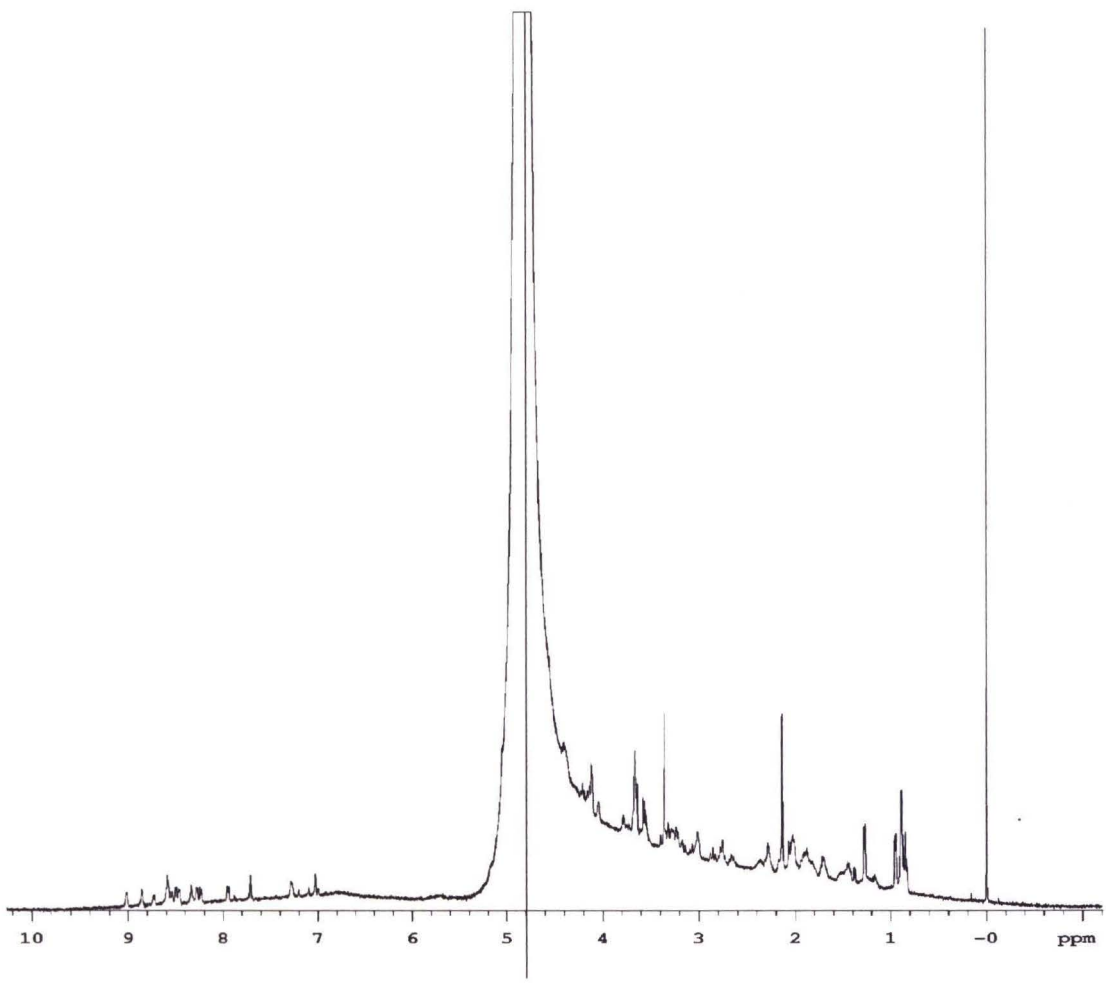


Figure 1. ^{1}H NMR spectrum at 25°C using a 3 mm NMR tube in 3 mm gHCN probe of reg12f

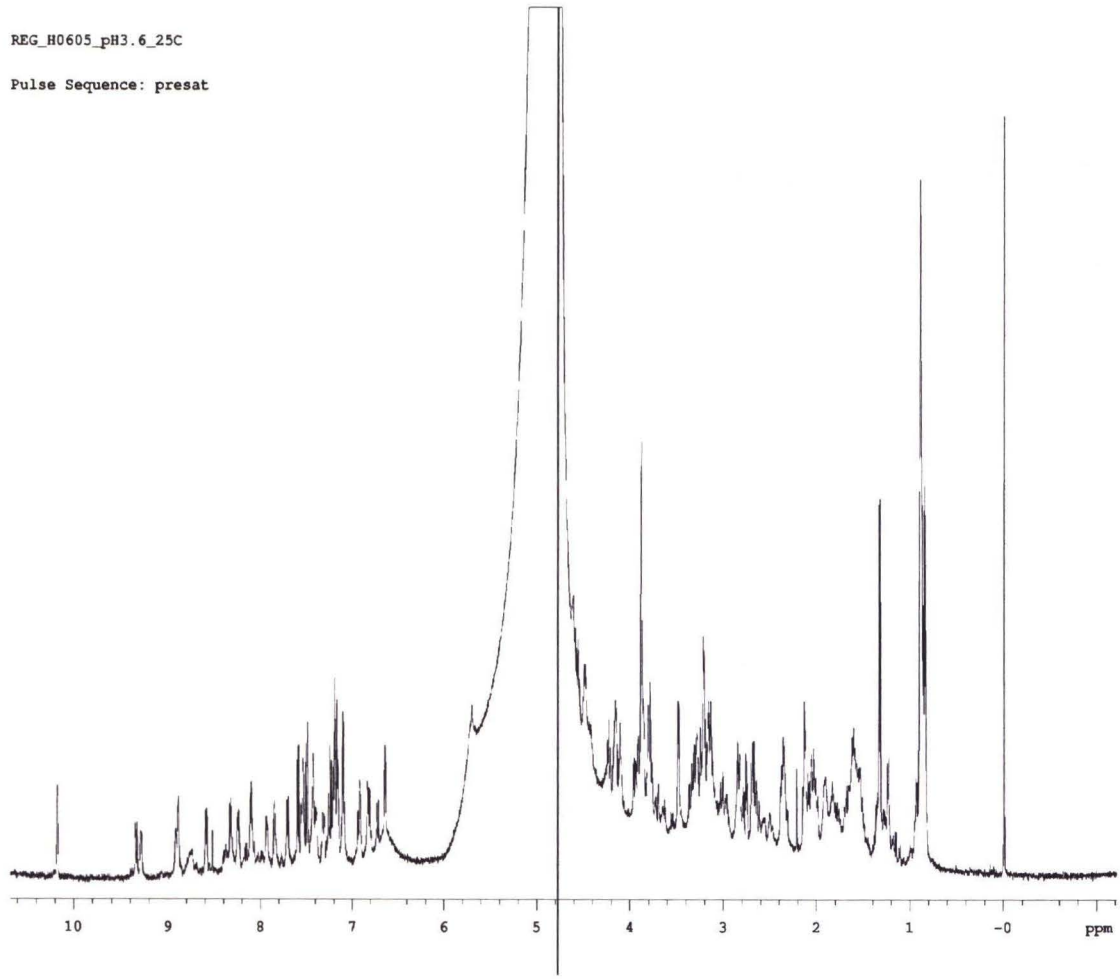


Figure 2. ^{1}H NMR spectrum at 25°C using a 3 mm NMR tube in 3 mm gHCN probe of reg12i

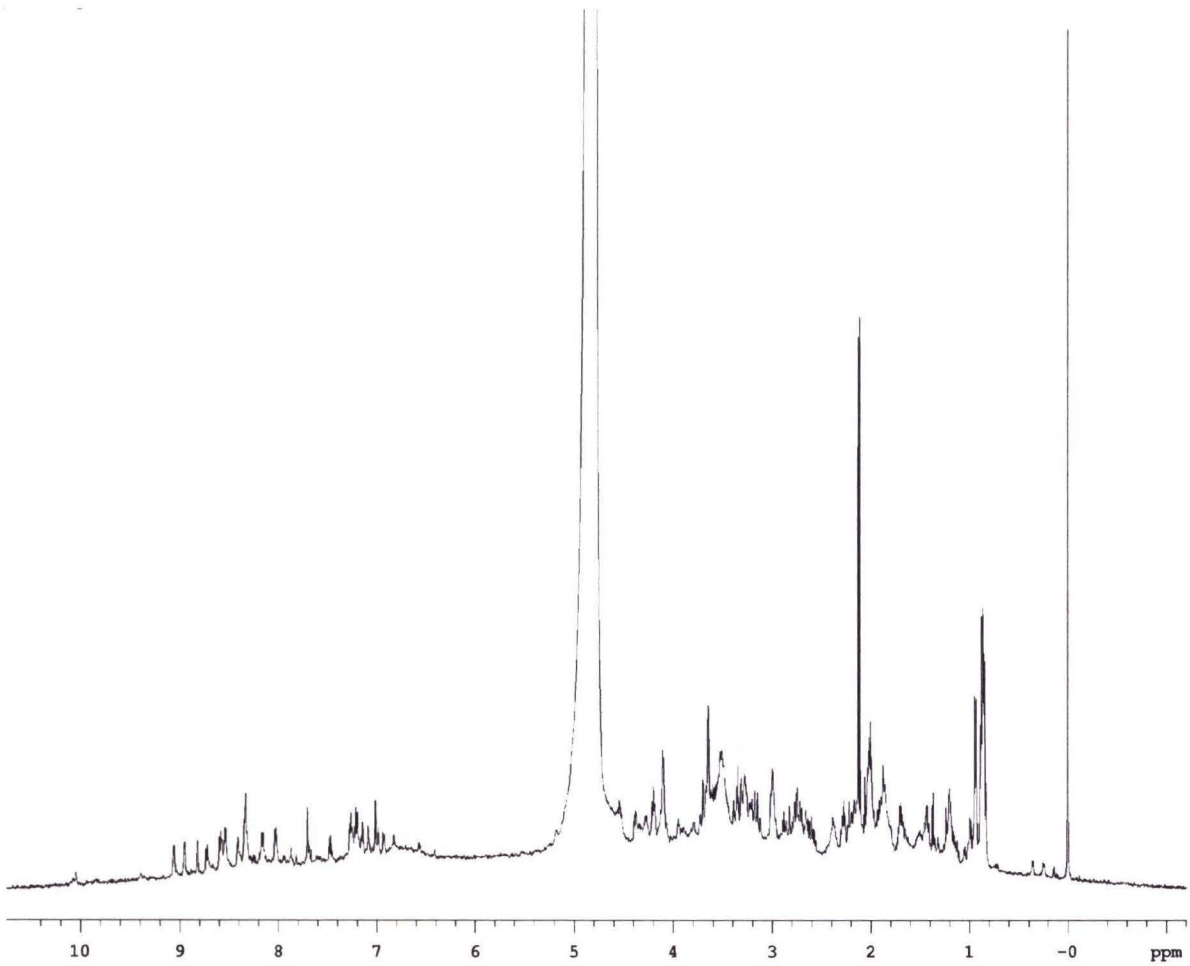


Figure 3. ^{1}H NMR spectrum at 25°C using a 3 mm NMR tube in 3 mm gHCN probe of reg12k

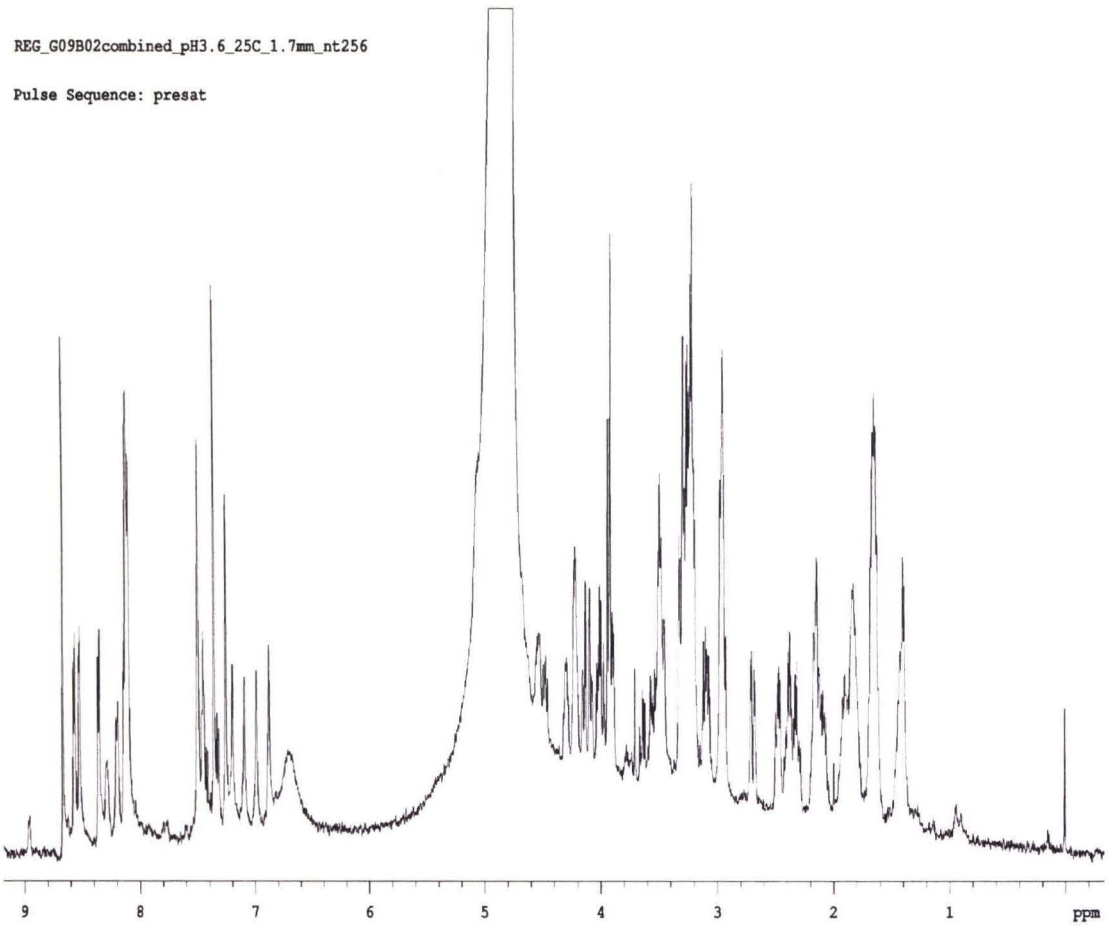


Figure 4. ^{1}H NMR spectrum at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe of reg1b

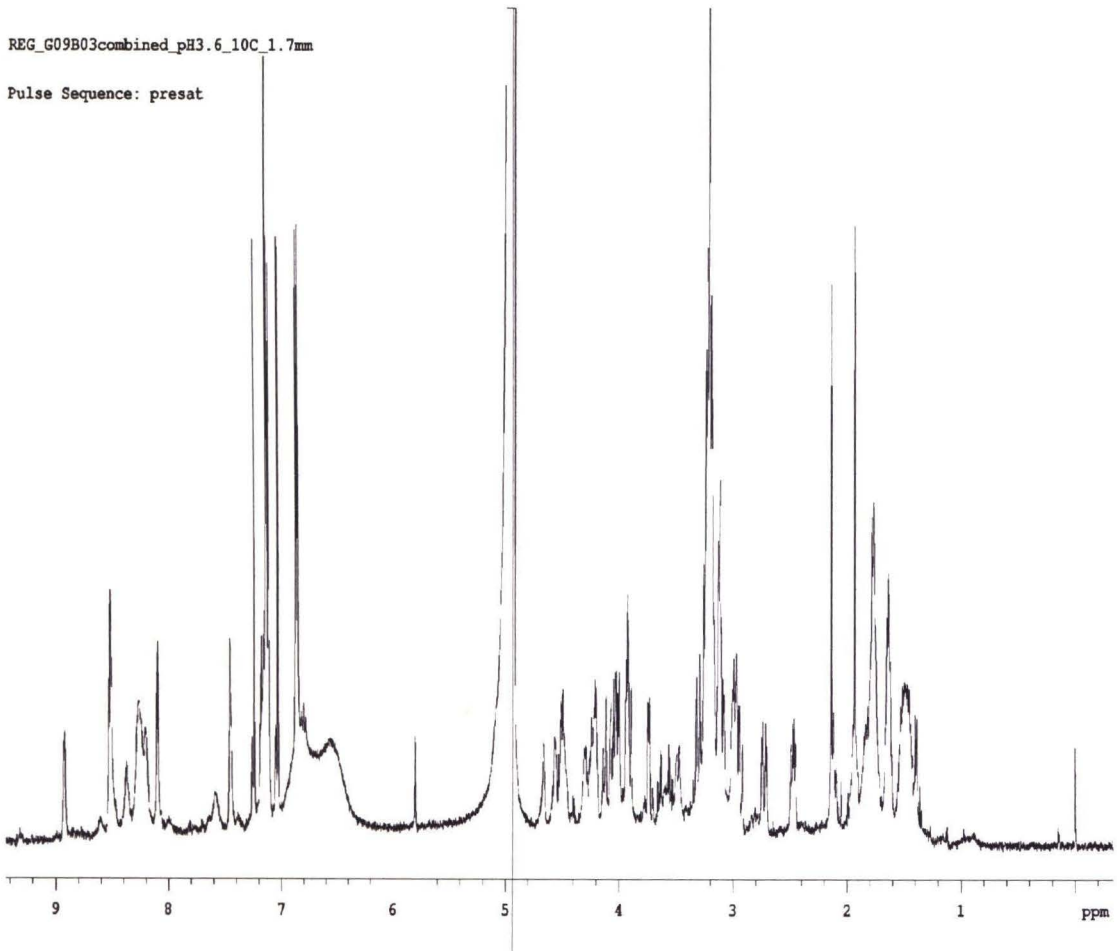


Figure 5. ^{1}H NMR spectrum at 25°C using a 1.7 mm NMR tube in 3 mm gHCN probe of reg1d