

Medical University of South Carolina

**MEDICA**

---

MUSC Theses and Dissertations

---

2020

## Venomics as a Drug Discovery Platform: Identifying Conopeptides with Pharmacological Activity

Meghan Kathleen Grandal  
*Medical University of South Carolina*

Follow this and additional works at: <https://medica-musc.researchcommons.org/theses>

---

### Recommended Citation

Grandal, Meghan Kathleen, "Venomics as a Drug Discovery Platform: Identifying Conopeptides with Pharmacological Activity" (2020). *MUSC Theses and Dissertations*. 536.  
<https://medica-musc.researchcommons.org/theses/536>

This Dissertation is brought to you for free and open access by MEDICA. It has been accepted for inclusion in MUSC Theses and Dissertations by an authorized administrator of MEDICA. For more information, please contact [medica@musc.edu](mailto:medica@musc.edu).

Venomics as a Drug Discovery Platform:  
Identifying Conopeptides with Pharmacological Activity

Meghan Kathleen Grandal

A dissertation submitted to the faculty of the Medical University of South Carolina in  
partial fulfillment of the requirements for the degree of Doctor of Philosophy in the  
College of Graduate Studies.

Department of Drug Discovery and Biomedical Sciences

2020

Approved by:

---

Frank Marí, Chairman

---

Mark Hamann, Co-chair

---

Patrick Woster, Committee Member

---

Christopher Davies, Committee Member

---

Yuri Peterson, Committee Member

## ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge Clay Davis and Ben Neely at NIST-Charleston for training me in every aspect of proteomic research used in this study. They selflessly provided expertise, advice, and assistance whenever it was requested from them. They were instrumental in my graduate training and dissertation research. I would also like to thank the venomics team at Hollings Marine Lab, Mickelene Hoggard and Carolina Moller, for their contribution and friendship during my PhD training.

Dr. Roger Papke and Clare Stokes at the University of Florida graciously allowed me to spend time in their lab where I learned so much from their expertise in and neuropharmacology nicotinic receptor research. They were both immeasurably welcoming and helpful and my time spent with them at UF was invaluable to the completion of this project.

Thank you to Yuri Peterson, research professor at MUSC and committee member, who dedicated his time and expertise to help with the molecular modeling in this project.

Lastly, I would like to thank my advisor, Frank Mari, who took a chance on a marine biologist with an interest in cell-based assays and fluorescent microscopy and turned her into a biochemist with a deep appreciation for mass spectrometry and bioinformatics. Thank you for always considering my ideas and for allowing me to be independent (and stubborn at times!). I have enjoyed learning from and working with you and am forever grateful for your mentorship.

## TABLE OF CONTENTS

Acknowledgements.....	ii
List of Tables.....	v
List of Figures.....	viii
Key to Symbols and Abbreviations.....	xii
Abstract.....	xiii
<b>Chapter 1: Introduction and Significance.....</b>	<b>1</b>
1.1 Venom in drug discovery.....	2
1.2 Conopeptides and their therapeutic applications.....	4
1.3 Nicotinic acetylcholine receptors: structure and function.....	6
1.4 Alpha-conotoxins: ligands of the nAChR.....	8
1.5 Impact.....	11
<b>Chapter 2: Materials and Methods.....</b>	<b>12</b>
2.1 Cone snail specimens.....	13
2.1.1 <i>C. purpurascens</i> specimens and venom collection	
2.1.2 Other <i>Conus</i> species	
2.2 Tissue dissection and RNA sequencing.....	13
2.3 Bioinformatics.....	14
2.3.1 Transcriptome assembly	
2.3.2 Conopeptide interrogation	
2.3.3 Insulin-like peptide interrogation	
2.4 Mass spectrometry analysis of venom samples.....	15
2.5 Database configuration and search parameters.....	16
2.6 Hierarchical cluster analysis and Principal Component Analysis.....	18
2.7 Alpha-conotoxin testing on nAChR subtypes.....	19
2.7.1 Alpha-conotoxin identification, selection, and synthesis	
2.7.2 Oocyte harvesting and injection	
2.7.3 Two-electrode voltage-clamp	
2.7.4 Data analysis and IC <sub>50</sub> calculation	
2.7.5 Homology modeling of the human $\alpha 7$ nAChR	
<b>Chapter 3: Proteogenomic Assessment of Intraspecific Venom Variability in <i>Conus purpurascens</i>.....</b>	<b>23</b>
3.1 Introduction.....	24
3.2 Results.....	26

3.2.1	Conopeptide identification	
3.2.2	Intraspecific venom comparison	
3.2.3	Novel S-superfamily conotoxin- PVIIIA	
3.3	Discussion.....	36
<b>Chapter 4:</b>	Insulin-like peptide from the injected venom of <i>Conus purpurascens</i> .....	43
4.1	Introduction.....	44
4.2	Results.....	46
4.2.1	Phylogenetic analysis of ILP expression	
4.2.2	ILP expression in <i>C. purpurascens</i> tissues	
4.2.3	Con-Ins P1, new ILP from injected venom	
4.3	Discussion.....	52
<b>Chapter 5:</b>	Functional screening of novel alpha-conotoxins on nicotinic acetylcholine receptor subtypes .....	58
5.1	Introduction.....	59
5.2	Results.....	62
5.2.1	Bioinformatic approach to identifying novel $\alpha$ -conotoxins	
5.2.2	Functional screening of novel $\alpha$ -conotoxins on nAChR subtypes	
5.2.3	Modeling $\alpha$ -PID and $\alpha$ -PIA binding to the $\alpha$ 7 nAChR	
5.3	Discussion.....	70
<b>Chapter 6:</b>	Conclusion and Future Directions.....	75
<b>Appendix A:</b>	Toxiforms and sites of modification for each conopeptide identified from the injected venom of <i>C. purpurascens</i> .....	79
<b>Appendix B:</b>	Chromatograms (TIC) of injected venom samples from <i>C. purpurascens</i> specimens.....	97
<b>Appendix C:</b>	Features of the New Conopeptides in <i>C. purpurascens</i> .....	112
<b>Appendix D:</b>	Alignment of gastropod insulin superfamily proteins.....	130
<b>Curriculum Vitae</b>	.....	133
<b>References</b>	.....	134

## LIST OF TABLES

**Table 1-** Molecular targets and bioactivity of known conopeptides.

**Table 2-** Conopeptides that reached human clinical trial as pain therapies.

**Table 3-** Search database file descriptions.

**Table 4-** nAChR subtypes with associated subunit RNA and ACh controls used for voltage-clamp experiments.

**Table 5-** Conopeptides identified from the injected venom of *Conus purpurascens*.

**Table 6-** Protein sequences of identified conopeptides.

**Table 7-** Comparison of transcriptome expression and MS identification from two specimens of *C. purpurascens*.

**Table 8-** RNA expression of conoinsulins in *C. purpurascens* tissues

**Table 9-** Inhibitory activity of Framework I  $\alpha$ -conotoxins from *C. purpurascens*

**Table 10-** Inhibitory activity of  $\alpha$ -PIA,  $\alpha$ -PID,  $\alpha$ -NuxIA, and  $\alpha$ -CedIA on nAChR subtypes

**Table 11-** Dose-response of  $\alpha$ -PIA and  $\alpha$ -PIDc on  $\alpha 4\alpha 6\beta 2\beta 3$  and  $\alpha 7$  nAChRs.

**Table 12-**  $\alpha$ -PIA toxiforms

**Table 13-**  $\alpha$ -PIB toxiforms

**Table 14-**  $\alpha$ -PIC toxiforms

**Table 15-**  $\alpha$ -PID toxiforms

**Table 16-** PIE toxiforms

**Table 17-** PIF toxiforms

**Table 18-** PIG toxiforms

**Table 19-**  $\alpha$ -PIVA toxiforms

**Table 20-**  $\kappa$ -PIVE toxiforms

**Table 21-**  $\kappa$ -PIVF toxiforms

**Table 22-** PIVH toxiforms

**Table 23-** Linear-P toxiforms

**Table 24-** Ile-Contryphan-P toxiforms

**Table 25-** Contryphan-P3 toxiforms

**Table 26-** Contryphan-P4 toxiforms

**Table 27-**  $\psi$ -PIIIE toxiforms

**Table 28-** PIIIG toxiforms

**Table 29-** PIIIH toxiforms

**Table 30-** PIIII toxiforms

**Table 31-**  $\delta$ -PVIA toxiforms

**Table 32-** PVIB toxiforms

**Table 33-** PVIC sites of modification

**Table 34-** PVID toxiforms

**Table 35-** PVIE toxiforms

**Table 36-** PVIF toxiforms

**Table 37-** PVIG toxiforms

**Table 38-**  $\kappa$ -PVIIA toxiforms

**Table 39-** Contryphan-P toxiforms

**Table 40-** PIIA toxiforms

**Table 41-** PVIIIA toxiforms

**Table 42-** PVA toxiforms

**Table 43-** PVB toxiforms

**Table 44-** p21b sites of modification on the two identified peptides

## LIST OF FIGURES

**Figure 1-** Knottin peptide structure and disulfide pairing.

**Figure 2-** Overlay of AChBP subunit with muscle-type nAChR  $\alpha$ -subunit

**Figure 3-** Alpha-conotoxin selectivity for nAChR subtypes

**Figure 4-** Neuronal subtypes of nAChRs and their clinical applications

**Figure 5-** Venom milking from *C. purpurascens*

**Figure 6-** Workflow of the database search strategy

**Figure 7-** Conopeptide IDs for 27 *C. purpurascens* injected venom samples

**Figure 8-** Conopeptide frequency in injected venom of *C. purpurascens*

**Figure 9-** Conopeptide profiles from 27 *C. purpurascens* specimens

**Figure 10-** *C. purpurascens* expresses two distinct venom profiles

**Figure 11-** Alignment of PVIIIA with characterized S-superfamily conotoxins

**Figure 12-** PVIIIA toxiform expression from 27 *C. purpurascens* specimens

**Figure 13-** Phylogenetic tree of the insulin superfamily.

**Figure 14-** Conoinsulins that exhibit a vertebrate-like insulin cysteine framework

**Figure 15-** Conoinsulin sequences expressed in *C. purpurascens* tissue

**Figure 16-** Con-Ins P1 from the injected venom of two *C. purpurascens* specimens

**Figure 17-** Con-Ins P1 modifications identified by MS/MS

**Figure 18-** Con-Ins P1 displays cysteine framework homology to vertebrate insulins

**Figure 19-** Sequences of  $\alpha$ -conotoxins extracted from *Conus* venom duct RNA-seq transcripts

**Figure 20-** Sequences of the  $\alpha$ -conotoxins tested for inhibitory activity.

**Figure 21-** Inhibitory activity of  $\alpha$ -PIA,  $\alpha$ -PID,  $\alpha$ -NuxIA, and  $\alpha$ -CedIA on nAChR subtypes

**Figure 22-** Dose response curve for  $\alpha$ -NuxIA on  $\alpha$ 1 $\beta$ 2 $\delta$  $\epsilon$  nAChRs.

**Figure 23-** Dose-response curves for  $\alpha$ -PIA and  $\alpha$ -PID on  $\alpha$ 4 $\alpha$ 6 $\beta$ 2 $\beta$ 3 and  $\alpha$ 7 nAChRs

**Figure 24-** Alignment of  $\alpha$ -PID and  $\alpha$ -PIA with conotoxins that inhibit  $\alpha$ 7 receptors.

**Figure 25-** Model of  $\alpha$ -PID bound to human  $\alpha$ 7 receptor.

**Figure 26-** Total ion chromatogram from venom sample of specimen 1

**Figure 27-** Total ion chromatogram from venom sample of specimen 2

**Figure 28-** Total ion chromatogram from venom sample of specimen 3

**Figure 29-** Total ion chromatogram from venom sample of specimen 4

**Figure 30-** Total ion chromatogram from venom sample of specimen 5

**Figure 31-** Total ion chromatogram from venom sample of specimen 6

**Figure 32-** Total ion chromatogram from venom sample of specimen 7

**Figure 33-** Total ion chromatogram from venom sample of specimen 8

**Figure 34-** Total ion chromatogram from venom sample of specimen 9

**Figure 35-** Total ion chromatogram from venom sample of specimen 10

**Figure 36-** Total ion chromatogram from venom sample of specimen 11

**Figure 37-** Total ion chromatogram from venom sample of specimen 12

**Figure 38-** Total ion chromatogram from venom sample of specimen 13

**Figure 39-** Total ion chromatogram from venom sample of specimen 14

**Figure 40-** Total ion chromatogram from venom sample of specimen 15

**Figure 41-** Total ion chromatogram from venom sample of specimen 16

**Figure 42-** Total ion chromatogram from venom sample of specimen 17

**Figure 43-** Total ion chromatogram from venom sample of specimen 18

**Figure 44-** Total ion chromatogram from venom sample of specimen 19

**Figure 45-** Total ion chromatogram from venom sample of specimen 20

**Figure 46-** Total ion chromatogram from venom sample of specimen 21

**Figure 47-** Total ion chromatogram from venom sample of specimen 22

**Figure 48-** Total ion chromatogram from venom sample of specimen 23

**Figure 49-** Total ion chromatogram from venom sample of specimen 24

**Figure 50-** Total ion chromatogram from venom sample of specimen 25

**Figure 51-** Total ion chromatogram from venom sample of specimen 26

**Figure 52-** Total ion chromatogram from venom sample of specimen 27

**Figure 53-** Supplementary information for Linear-P.

**Figure 54-** Supplementary information for Ile-Contryphan-P.

**Figure 54-** Supplementary information for Contryphan-P3

**Figure 55-** Supplementary information for PID

**Figure 56-** Supplementary information for PIE

**Figure 57-** Supplementary information for PIF

**Figure 58-** Supplementary information for FIG

**Figure 59-** Supplementary information for PIIA

**Figure 60-** Supplementary information for PIIIG

**Figure 61-** Supplementary information for PIIIH

**Figure 62-** Supplementary information for PIIII

**Figure 63-** Supplementary information for PIVH

**Figure 64-** Supplementary information for PVB

**Figure 64-** Supplementary information for PVIB

**Figure 65-** Supplementary information for PVIC

**Figure 66-** Supplementary information for PVID

**Figure 67-** Supplementary information for PVIE

**Figure 68-** Supplementary information for PVIF

**Figure 69-** Supplementary information for PVIG

**Figure 70-** Supplementary information for PVIIIA

**Figure 71-** Supplementary information for p21b

**Figure 72-** Alignment of gastropod insulin superfamily proteins.

## KEY TO SYMBOLS AND ABBREVIATIONS

PTMs	Post-translational modifications
nAChRs	Nicotinic acetylcholine receptors
AChBP	Acetylcholine binding protein
$\alpha$ -CNTX	Alpha- conotoxin
LC-MS/MS	Liquid chromatography mass spectrometry
ILPs	Insulin-like peptides
MIPs	Molluscan insulin peptides
TPM	Transcripts per million

MEGHAN KATHLEEN GRANDAL. Venomics as a Drug Discovery Platform: Identifying Conopeptides with Pharmacological Activity. (Under the direction of FRANK MARÍ).

## ABSTRACT

Cone snail venom is a mixture of disulfide-constrained peptides (conotoxins), hormone-like peptides, and proteins that have been ‘weaponized’ for predation and defense. Venom peptides, or conopeptides, have efficiently evolved to bind receptors and ion channels that modulate the neuromuscular, cardiovascular, and central nervous systems in prey species. With over 850 species of cone snails, each with unique venom concoctions, cone snail venom is a valuable source of novel pharmacological probes and potential drug leads. However, the complexity of the venom poses a challenge for drug discovery. Contributing to the complexity is 1) a wide range in molecular weight 2) peptide hyper-variability by post-translational modifications and 3) many potential molecular targets to pursue. In this research, a ‘venomics’ approach was employed for the global identification of venom components. This ‘venomics’ methodology combines RNAseq data from the venom duct and proteomic data from raw injected venom to identify novel conopeptides.

This project was a data-driven effort to define the venom components of the cone snail, *Conus purpurascens*, and to stimulate further hypothesis-driven studies. First, 21 new base conopeptides were identified from the injected venom of *Conus purpurascens*, a fish-hunting cone snail native to the Pacific coast of Central America. The molecular targets were projected based on homology to previously characterized conopeptides. The newly identified conopeptides included  $\alpha$ -conotoxin,  $\alpha$ -PID. Alpha-conotoxins are inhibitory ligands of nicotinic acetylcholine receptors (nAChRs), and the most ubiquitous venom components across the *Conus* genus. Ligands of nAChRs are clinically important for addiction, cognitive disorders, neurodegenerative diseases, and pain. Functional characterization of  $\alpha$ -PID and three other  $\alpha$ -conotoxins was performed to test their activity on different nAChR subtypes using heterologous receptor expression and molecular modeling techniques. A unique insulin-like peptide (Con-Ins P1) was also identified and was the first instance of an insulin-like peptide identified directly from injected venom. This research demonstrates how discovery-based ‘venomics’ workflows can be used to yield novel peptides with pharmacological applications and stimulate further hypothesis-driven experiments.

CHAPTER 1:

INTRODUCTION AND SIGNIFICANCE

## 1.1 Venom in drug discovery

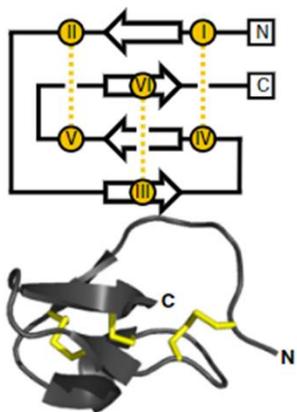
The world's most poisonous animals, which include snakes, scorpions, jellyfish, octopi, and cone snails, all use venom for predation and defense. Venom is a complex mixture of peptides and proteins that have evolved across the animal kingdom. While envenomation strategies may differ across clades, the venom protein scaffolds are highly conserved for the unified goal of paralyzing prey. Venom peptides and proteins target a wide range of pharmacological targets to accomplish this goal, and because of this, humans continue to use and study venom for medicinal purposes [2]. There are a few drugs on the market derived from venom that fall under the following therapeutic categories: ACE inhibitors (captopril), anti-platelet drugs (eptifibatide and tirofiban), thrombin inhibitors (lepirudin and bivalirudin), type-2 diabetes drugs (exenatide), and pain killers (ziconotide) [3].

Chlorotoxin from the deathstalker scorpion, *Quinquestriatus hebraeus*, binds chloride channels with high affinity. Chlorotoxin has unique selectivity for tumor cells, and when attached to a fluorescent dye (Tumor Paint®, Blaze Biotech, Seattle) allows surgeons to visualize tumors using infrared glasses [4]. Tumor Paint® is currently undergoing clinical trials for use during surgical removal of brain, breast, and skin cancers. Dalazatide (Kineta, Seattle), a voltage-gated potassium channel blocker ( $K_v1.3$ ) from the Caribbean sun anemone (*Stichodactyla helianthus*), is currently being developed for the treatment of autoimmune diseases. Phase I trials showed positive results against psoriasis [5] and is now being tested for rheumatoid arthritis, lupus, and type 1 diabetes. There are also several venom peptides in preclinical development for pain, including tarantula-derived sodium channel blockers (Nav1.7) [6] and conotoxin Rg1A [7]. Therapeutic applications of conotoxins will be discussed in-depth in the following section.

The biggest obstacle when developing venom components as drugs is their bioavailability. This results from their relatively large size, compared to small molecule therapeutics, and their hydrophilic nature, rendering them incapable of crossing biological

membranes to reach their molecular targets. These barriers include the blood-brain barrier and intestinal walls [2, 8]. Venom toxins under drug development commonly require peptide engineering to improve bioavailability. One successful mechanism is the cyclization of conotoxins by linking the N- and C- termini [9-11]. Another option is to synthesize smaller toxin analogs that maintain the functional group of the peptide but improve bioavailability. Post-translational modifications (PTMs), such as glycosylation, have also been shown to improve peptide permeability through biological membranes [12].

One group of peptides that have had success as pharmaceuticals are the cysteine knot peptides, miniprotein scaffolds restrained by multiple disulfide bonds [13]. Included in this



**Figure 1- Knottin peptide structure and disulfide pairing.**  
Adapted from Kintzing et al. 2016

family are the knottins, which have six cysteine residues and a specific disulfide pairing between cysteines 1 and 4, 2 and 5, and 3 and 6 (Figure 1). Cysteine knot peptides are naturally produced by animals, plants, and fungi, but the knottins are found most prominently in cone snail and spider venom [14]. More recently, knottin peptides have

also been described from anemone venom and marine sponge tissue [15, 16].

Two FDA approved drugs are derived from naturally occurring cysteine knot peptides, linaclotide (Linzess<sup>®</sup>, AbbVie and Ironwood Pharmaceuticals, Inc.) from the endogenous hormone guanylin [17] and ziconotide (Prialt<sup>®</sup>, TerSera Therapeutics, LLC) from conotoxin  $\omega$ -MVIIA [18]. Factors contributing to cysteine knot peptide success as drugs include their thermal, chemical, and proteolytic stability [19]. Their resistance to proteases allows these peptides to remain intact in biological environments, including the gastrointestinal tract, thereby increasing their bioavailability and likelihood of availability through oral administration [20]. This is not the case for ziconotide, a conotoxin-based drug for chronic pain that is administered

intrathecally through an infusion pump due to its low oral bioavailability. As we continue to discover venom peptides with clinically relevant targets, the knottin miniprotein scaffolds will require substantial engineering efforts and improved drug delivery mechanisms.

## **1.2 Conopeptides and their therapeutic applications**

Conopeptides are a diverse group of rapidly evolving gene products found in cone snail venom that range in size, structure, and molecular target. They work synergistically to immobilize prey/predators. Positive selection through point mutations, alternative splicing, and post-translational modifications (PTMs) has created a rich source of bioactive peptides that target membrane receptors with high specificity [21-23]. The molecular targets are used to classify the conopeptides into pharmacological families (Table 1) [24, 25]. Their targets include voltage-gated and ligand-gated ion channels, G-protein-coupled receptors, and neurotransmitter transporters, all with important clinical implications.

Several conopeptides have successfully reached clinical trials (Table 2). Although conopeptides have a broad range of therapeutic targets, their high specificity for neuronal receptors make them particularly apt for pain treatment [26, 27]. Overuse and overprescribing of opiate-based pain killers have resulted in a current opiate crisis [28]. There is a severe need for alternative, non-opiate pain management treatments. Conopeptides are a promising source of novel pain medications without the risk of addiction and other negative side effects caused by opiate use.

**Table 3- Molecular targets and bioactivity of known Conopeptides.**

Pharmacological Family	Target	Activity	Framework	Cysteine pattern
$\alpha$ - conotoxin	nAChR	inhibit channel	I	CC-C-C
$\alpha$ A- conotoxin			II	CCC-C-C-C
$\alpha$ S- conotoxin			IV	CC-C-C-C-C
$\alpha$ D- conotoxin			VIII	C-C-C-C-C-C-C-C-C-C
$\Psi$ - conotoxin			XX	C-CC-C-CC-C-C-C-C
$\sigma$ - conotoxin	5HT <sub>3</sub> R	allosteric inhibitor	III	CC-C-C-CC
$\mu$ - conotoxin	Na <sup>+</sup> channel	inhibit channel	VIII	C-C-C-C-C-C-C-C-C-C
$\mu$ O- conotoxin		inhibit conductance	III	CC-C-C-CC
$\delta$ - conotoxin		delay inactivation	VI/VII	C-C-CC-C-C
$\omega$ - conotoxin	Ca <sup>2+</sup> channel	inhibit channel	VI/VII	C-C-CC-C-C
$\kappa$ A- conotoxin	K <sup>+</sup> channel	inhibit conductance	IV	CC-C-C-C-C
$\kappa$ J- conotoxin		inhibit channel	XIV	C-C-C-C
$\kappa$ M- conotoxin		inhibit channel	III	CC-C-C-CC
$\kappa$ O- conotoxin	shaker K <sup>+</sup>	inhibit channel	VI/VII	C-C-CC-C-C
$\chi$ - conotoxin	NE transporter	inhibit transporter	I/X	CC-C-C
Conantokin	NMDA receptor	---		
Contulakin	Neurotensin receptor	Agonist		
Conopressin	Vasopressin receptor	Agonist		C-C
conoCAPS		cardioactive peptide		C-C

**Table 4- Conopeptides that reached human clinical trial as pain therapies.**

<b>Conopeptide</b>	<b>Species</b>	<b>Target</b>	<b>Application</b>	<b>Status</b>
MVIIA	<i>C. magus</i>	Ca <sub>v</sub> 2.2	Pain	FDA approved, 2004
CBID	<i>C. catus</i>	Ca <sub>v</sub> 2.2	Pain	discontinued
MrlA	<i>C. marmoreus</i>	Norepinephrine transporter	Pain	discontinued
Contulakin-G	<i>C. geographus</i>	Neurotensin receptor	Pain	discontinued
Conantokin-G	<i>C. geographus</i>	NMDA receptor	Pain, epilepsy	discontinued
Vc1.1	<i>C. victoriae</i>	α9α10 nAChR	Pain	discontinued
RgIA4	<i>C. regius</i>	α9α10 nAChR	Pain	preclinical

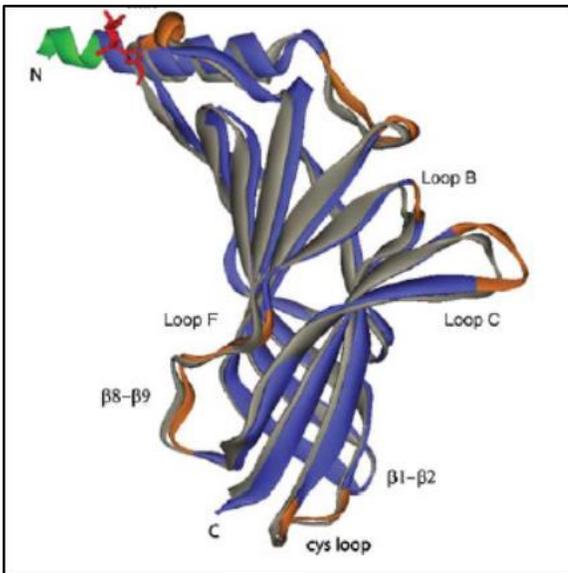
Thus far, Ziconotide is the only conotoxin-derived drug approved by the FDA (Prialt™) and is the only venom peptide approved for the management of intractable pain. Ziconotide is a non-opioid, non-NSAID analgesic that blocks Ca<sub>v</sub>2.2 subtype calcium channels [29]. It was developed from the conotoxin MVIIA from *Conus magus*. Ziconotide requires an intrathecal drug delivery system because of its inability to pass the blood-brain barrier and is therefore not an optimal alternative for chronic pain management.

Contulakin-G, a neurotensin receptor antagonist from *Conus geographus*, was tested previously in a clinical trial for severe chronic pain. Vc1.1 from *Conus victoriae*, a nicotinic receptor inhibitor, went to a Phase 2 clinical trial for neuropathic pain associated with sciatica and diabetic neuropathy [30, 31]. Another nicotinic receptor inhibitor, RgIA4, an analog of α-conotoxin RgIA from the venom of *Conus regius*, is currently in preclinical development for neuropathic pain [3, 32]. Vc1.1 and RgIA both inhibit nicotinic receptors [33], but conflicting evidence suggests that their analgesic effect is through the GABA<sub>B</sub> receptor [34, 35]. In this proposed study, we aim to identify venom peptides that target nicotinic receptors.

### 1.3 Nicotinic acetylcholine receptors: structure and function

Nicotinic acetylcholine receptors (nAChRs) are ligand-gated ion channels responsible for neurotransmitter signaling at the synaptic gap and the neuromuscular junction. They belong to the Cys-loop superfamily of ligand-gated ion channels, which includes receptors gated by acetylcholine, serotonin, GABA, and glycine. All contain a signature cysteine loop formed by a disulfide bond joining adjacent subunits [36, 37]. Initial studies on nAChRs utilized the receptor-dense electrical organ tissue of *Torpedo marmorata* (electric stingray) [38]. This later allowed cloning and functional characterization of seventeen different genes for nAChR subunits ( $\alpha$ 1-10,  $\beta$ 1-4,  $\gamma$ ,  $\delta$ , and  $\epsilon$ ). Any five of these subunits join to form functional pentameric proteins that can be either heteromeric or homomeric.

Nicotinic receptors are classified into the muscle or neuronal subtypes depending on their subunit composition and localization [39]. Muscle subtypes are comprised of  $\alpha$ 1,  $\beta$ 1,  $\gamma$ ,  $\delta$ , and  $\epsilon$  subunits, and are expressed at the neuromuscular junction. The muscle-type nAChRs are critical for skeletal muscle contraction and voluntary movement. Inhibitors of muscle-type nAChRs include local anesthetics, such as lidocaine [40], and  $\alpha$ -neurotoxins that cause paralysis, such as the snake venom toxin,  $\alpha$ -bungarotoxin [41]. Neuronal subtypes are comprised of combinations of  $\alpha$  and  $\beta$  subunits ( $\alpha$ 2-6 and  $\beta$ 2-4) or are homomeric ( $\alpha$ 7-10). Neuronal subtypes are heterogeneously expressed throughout the central and peripheral nervous systems where they are involved in neuronal transmission and the dopaminergic pathway [42].



**Figure 2- Overlay of AChBP subunit with muscle-type nAChR  $\alpha$ -subunit.** AChBP subunit (blue) with muscle-type nAChR  $\alpha$ -subunit (gray) (Hansen, 2005). N-terminal (N), C-terminal (C), ligand-binding domain (Loops B,C,F).

The ligand-binding properties of nAChRs have been studied using the X-ray crystal structure of the soluble acetylcholine binding protein (AChBP) of mollusks *Lymnaea stagnalis* and *Aplysia californica* [43]. AChBPs are not functional ion channels, however; they form stable homopentamers that preserve features of the ligand-binding domain of nAChRs and thereby serve as useful binding models. The ligand-binding pocket is 'gated' by a loop structure (Loop C) that controls ligand activity [44]. Crystallized AChBP

complexes with select nicotinic agonists and antagonists revealed that agonists bind below the C loop causing it to close over the ligand-bound pocket and open the channel pore. Antagonists, such as  $\alpha$ -conotoxins, cause Loop C to be held in an extended conformation away from the ligand-binding pocket [45, 46] (Figure 2). Heteromeric receptors have two ligand-binding sites located between adjacent  $\alpha$  and  $\beta$  subunits. Homomeric receptors have five ligand-binding sites located between each  $\alpha$  subunit [47].

Ligands of neuronal nAChR subtypes are clinically important as treatments for addiction, cognitive disorders, neurodegenerative diseases, and pain [48]. However, there are few compounds available that target neuronal subtype nAChRs with high selectivity. To explore the physiological role of neuronal-type receptors, we need highly selective pharmacological probes. Alpha-conotoxins are the first group of peptide  $\alpha$ -neurotoxins that show selectivity for neuronal subtype nAChRs and are therefore extremely valuable research tools and pharmacological agents [49, 50].

## 1.4 Alpha-conotoxins: ligands of the nAChR

Alpha-conotoxins are inhibitors of the nAChR and are the most ubiquitous venom components across the *Conus* genus. All species analyzed express one or more  $\alpha$ -conotoxins in their venom that work synergistically to paralyze prey. Unique from other nAChR ligands,  $\alpha$ -conotoxins exhibit remarkable subtype selectivity, specifically among neuronal receptor subtypes

Alpha-conotoxins can display different cysteine frameworks, or patterns of cysteine residues within a sequence (Table 1). Framework I  $\alpha$ -conotoxins belong to the A Superfamily, as defined by a conserved gene signal sequence. This is the largest group of characterized  $\alpha$ -conotoxins with the greatest diversity in subtype selectivity. The post-translational modifications commonly found on these conotoxins (C-terminal amides and hydroxyprolines) are important for peptide stability and bioactivity [51]. In our study, we will be focusing on framework I  $\alpha$ -conotoxins because of their relatively small size (<22 amino acids) and because of their well-established disulfide connectivity (C1-C3, C2-C4) [52]. The fact that they have only two disulfide bonds with established pairing makes these peptides good candidates for chemical synthesis, which is necessary to perform functional assays.

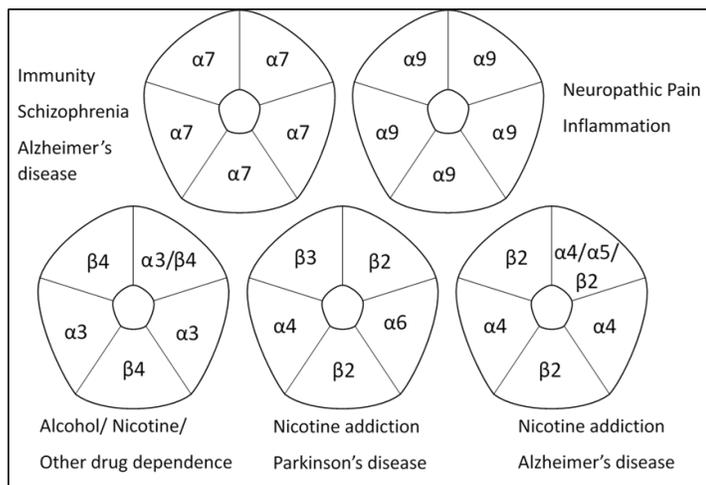
Inter-cysteine loops		
$\alpha 3/5$	$CC(X_{aa})_3C(X_{aa})_5C$	} Muscle subtype
$\alpha 4/4$	$CC(X_{aa})_4C(X_{aa})_4C$	
$\alpha 4/3$	$CC(X_{aa})_4C(X_{aa})_3C$	} Neuronal subtype
$\alpha 4/7$	$CC(X_{aa})_4C(X_{aa})_7C$	

**Figure 3- Alpha-conotoxin selectivity for nAChR subtypes.**

Within the framework I  $\alpha$ -conotoxins, there is significant diversity in amino acid composition and the size of the inter-cysteine loops. Inter-cysteine loop size affects affinity toward either muscular or neuronal nAChR subtypes (Figure 3) [53, 54]. In general,  $\alpha$ -conotoxins exhibiting a 3/5 inter-cysteine loop pattern are inhibitors of neuromuscular junction subtypes, and 4/3 and 4/7 toxins inhibit neuronal subtypes [55]. To date, there are very few 4/4  $\alpha$ -conotoxins

characterized. Of the ones known,  $\alpha$ -BuIA inhibits neuronal receptors containing  $\alpha 3$ ,  $\alpha 6$ , and  $\beta 2$  subunits, whereas  $\alpha$ -PIB and  $\alpha$ -PIC preferentially inhibit muscle subtypes.

Neuronal nAChR subtypes are expressed heterogeneously throughout the central nervous system [56] and are implicated in a range of neurological conditions (Figure 4).



**Figure 4-Neuronal subtypes of nAChRs and their clinical applications.**

The most widely expressed subtype in the mammalian brain,  $\alpha 4\beta 2$ , is heavily involved in the dopaminergic pathway and nicotine addiction. This receptor can exhibit different subtype stoichiometry;  $(\alpha 4)_2(\beta 2)_3$  is more sensitive to nicotine than  $(\alpha 4)_3(\beta 2)_2$  [57]. Because of its

role in nicotine addiction,  $\alpha 4\beta 2$  is the clinical target for smoking cessation therapies, including the partial agonist, varenicline (Chantix, Pfizer, Inc.). Very few  $\alpha$ -conotoxins inhibit  $\alpha 4\beta 2$ , with  $\alpha$ -GID being the most potent inhibitor. Amino acid substitution studies on  $\alpha$ -GID deemed A10, V13, and V18 as critical residues for  $\alpha 4\beta 2$  selectivity [58].

Of the neuronal subtypes, the  $\alpha 3\beta 2$  receptor is the most common  $\alpha$ -conotoxin target. LvIA is the first  $\alpha 3\beta 2$  subtype-selective conotoxin. The Asp11 residue is responsible for selectivity over  $\alpha 6$ -containing subtypes [59]. The selectivity of  $\alpha$ -PnIA for  $\alpha 3\beta 2$  over its alternative  $\alpha 7$  target is dependent on Ala10 [60].

The  $\alpha 3\beta 4$  subtype is the predominant nAChR in the sensory and autonomic ganglia neurons. It is expressed in the mesolimbic dopamine circuitry of the midbrain where it modulates addiction to nicotine and potentially other drugs of abuse [61, 62]. AulB is an  $\alpha 3\beta 4$  selective toxin with an uncommon 4/6 inter-cysteine loop size. The Phe9 residue regulates subtype selectivity [63]. Ligands with  $\alpha 3\beta 4$  selectivity will help better understand

the role of these receptors in the midbrain and may be useful therapeutics for addiction [64].

The  $\alpha 6$ -containing subtypes are also expressed abundantly in the midbrain dopaminergic neurons and are mediators of the nicotine reward pathway [65, 66]. Ligands with selectivity for  $\alpha 6$ -subtype nAChRs are important molecular probes to study the pathophysiology of addiction and other dopamine-related disorders, such as Parkinson's disease. However,  $\alpha 6$  selectivity is rare because of its high homology with the  $\alpha 3$  subunit. There is one  $\alpha 6$ -biased ligand,  $\alpha$ -PIA from *C. purpurascens*, known thus far.  $\alpha$ -PIA preferentially inhibits  $\alpha 6$ -containing receptors with a 75-fold greater affinity than  $\alpha 3$  receptors [67, 68].  $\alpha$ -MII from *C. magus* is an  $\alpha 3$ -selective ligand, but alanine substitution studies distinguished  $\alpha$ -MII analogs with significantly increased affinity  $\alpha 6$  over  $\alpha 3$  [69], shedding light on structural determinants of  $\alpha 6$  selectivity.

Homomeric  $\alpha 7$  nAChRs are a unique subtype expressed throughout the brain and in non-neuronal tissues, like immune cells [70]. Their distribution in the brain includes the hippocampus and cerebral cortex involved in learning and memory [71]. The  $\alpha 7$  receptor is unique in that agonist binding elicits a relatively low ion current and becomes easily desensitized, compared to heteromeric subtypes. The  $\alpha 7$  receptors also bind highly-selective molecules called positive allosteric modulators (PAMs) that significantly enhance agonist-induced ion currents, but have no effect alone (reviewed in [39]). Conotoxin  $\alpha$ -MrlC is the first described selective  $\alpha 7$  agonist that can activate, rather than inhibit,  $\alpha 7$  receptors in the presence of PAMs (65). It is thereby a useful tool in studying the kinetics and function of this unique receptor subtype.

The  $\alpha 9$  homomeric receptor is another unique, and perhaps the most elusive, nAChR subtype. It is expressed in cochlear hair cells and periphery epithelial and immune cells [42]. The  $\alpha 9$  subunit can co-express with  $\alpha 10$  to form functional ion channels, and elicit ACh-evoked currents 100-fold stronger than homomeric  $\alpha 9$  receptors [72]. Both  $\alpha 9$  and

$\alpha 9\alpha 10$  are thought to be involved in pain pathways [73]. The  $\alpha$ -conotoxins Vc1.1 (4/7) and RgIA (4/3) both elicit analgesia through inhibition of  $\alpha 9\alpha 10$  nAChRs [32, 73]. As an effective anti-nociceptive target,  $\alpha 9$  nAChRs offer an alternate non-opioid pathway for pain treatment that needs to be explored.

## 1.5 Impact

The identification of venom components has increased rapidly with the advancement in next-generation sequencing technologies and bioinformatics tools. Venomics is the hybridization of RNAseq and proteomic analysis to study venom and identify its components. Venomics enables high-throughput discovery of venom peptide and protein sequences [74-76].

Venomics methodology supersedes previous chemical-based or bioassay-guided fractionation for natural product discovery. These traditional discovery methods require multiple biological assays or elaborate chemical elucidation schemes to find a single compound of interest. The workflow established in this proposed research project can be modified to discover novel active peptides from any venomous animal. This project will also provide information on novel  $\alpha$ -conotoxin activity on multiple nAChR subtypes with important clinical applications. Specifically,  $\alpha$ -conotoxins have proven to be good candidates for **non-opiate** pain treatment and have made it to clinical trials. To combat our current opiate crisis and regulate the over-prescription of opiate medications, we must have safe and effective alternatives for pain management.

CHAPTER 2:

MATERIALS AND METHODS

## 2.1 Cone snail specimens

### 2.1.1 *C. purpurascens* specimens and venom collection

Specimens of *C. purpurascens* (n = 27) were collected from the Pacific coast of Costa Rica. *C. purpurascens* was chosen for this intraspecific venom analysis because it is a fish-hunting species that uses a hook-and-pull strategy to capture prey allowing venom collection through a “milking” procedure [31]. Briefly, venom is



**Figure 5- Venom milking from *C. purpurascens*.**

collected into Eppendorf tubes that have a piece of latex glove stretched over the opening and are baited with a piece of goldfish fin on the latex. When the snail senses the fin, it spears the latex and injects venom into the tube (Figure 5). After the venom is released, the snail is fed with a live fish. Snails were kept in an aquarium facility where they were fed and milked regularly. The injected venom samples were stored at -80 °C until used for further analysis.

### 2.1.2 Other *Conus* species

Venom ducts were also dissected for downstream RNASeq from the following *Conus* species: *C. striatus*, *C. vanhyningi*, *C. brunneus*, *C. spurius*, *C. princeps*, *C. regius*, *C. gladiator*, *C. lindae*, *C. arangoi*, *C. moreleti*, *C. nussatella*, *C. richardbinghami*, *C. leopardus*, *C. caysalenensis*, *C. granulatus*, *C. dalli*, *C. anabathrum*.

## 2.2 Tissue dissection and RNA sequencing

Venom ducts were dissected from two *C. purpurascens* snails, specimens 1 and 14. The venom ducts were immediately placed in RNAlater (Invitrogen), and stored at -80° C. mRNA was extracted from the venom duct using an RNeasy Lipid Tissue mini kit (Qiagen), and mRNA quality was confirmed with a 2100 Bioanalyzer (Agilent). Illumina libraries were

prepared with a NEBNext Ultra Directional RNA Library prep kit (New England BioLabs). Sequencing was performed on a NextSeq 500 platform (Illumina, Inc.) and each venom duct yielded approximately 28 million paired-end reads (75 bp).

## **2.3 Bioinformatics**

### *2.3.1 Transcriptome assembly*

Raw reads for each venom duct were assembled with Trinity *de novo* transcript assembler (v. 2.2.0) using default parameters; group pairs distance: 500 bp, path reinforcement distance: 75 bp [33]. The resulting transcriptomes (A and B) were translated with EMBOSS applications, transeq (6-frame) and getorf (between start and stop codons) [34, 35] (Table 3). Transcriptomes were interrogated for conotoxin expression.

### *2.3.2 Conopeptide interrogation*

The Trinity *de novo* assemblies (transcriptomes A and B) were translated with transeq and getorf, and resulting transcripts were blasted (blastp,  $e = 10^{-5}$ ) against the UniProt Animal Toxin Annotation database (ToxProt) and all UniProt *Conus* entries (TaxID: 6490) to extract toxin-like sequences. The resulting sequences were developed into databases for proteomic searches (described in detail in Section 2.5). Transcript quantification (TPM) for conopeptides from the *C. purpurascens* venom duct transcriptomes was performed with Salmon [30] using the Trinity assembly as the reference transcriptome (kmer length = 31).

### *2.3.3 Insulin-like peptide interrogation*

The translated FASTA files were interrogated for conoinsulin sequences using BLASTp search function ( $e = 10$ ) with a conoinsulin query database that included all conoinsulin sequences from UniProt. Transcript quantification (TPM) for conoinsulins from *C. purpurascens* tissues was performed with Salmon [30] using the Trinity assembly as the reference transcriptome (kmer length = 31).

A phylogenetic analysis of ILPs was performed using insulin superfamily proteins obtained from the InterPro database (IPR036438). Taxonomy IDs were extracted and submitted to NCBI Taxonomy Browser- Common Tree to generate a phylip tree. The phylip tree file was imported into Geneious 2020.1.1 (Biomatters, Ltd., Auckland, NZ) to create a phylogenetic tree. ILPs from all *Conus sp.* were extracted from the InterPro database (IPR036438) and aligned using Clustal omega within Geneious software.

#### **2.4 Mass spectrometry analysis of venom samples**

An aliquot of each venom sample (5  $\mu$ L) was diluted in ammonium bicarbonate buffer (50 mM). Cysteine bonds were reduced with dithiothreitol (7 mM) for 1 h at 60 °C and alkylated with iodoacetamide (18 mM) for 1 h at 21 °C in the dark. Following reduction and alkylation, the samples were desalted using C18 spin columns (ThermoFisher Pierce) and lyophilized before LC-MS/MS analysis.

Samples were reconstituted in water/0.1% formic acid and analyzed by LC-MS/MS on an Orbitrap Fusion Lumos trihybrid mass spectrometer (ThermoFisher Scientific) coupled with an UltiMate 3000 RSLCnano System (ThermoFisher Scientific). A 160 min gradient with solutions A (5% acetonitrile/ 0.1% formic acid) and B (80% acetonitrile/ 0.1% formic acid) on an Acclaim PepMap 2 $\mu$ m C18 column (75  $\mu$ m x 25 cm) (ThermoFisher Scientific) was used. The flow rate was set at 0.3  $\mu$ L/min with the following gradient steps: 0 min at 5% B, 10 min at 5% B, 115 min at 27.5% B, 130 min at 40% B, 140 min at 95% B, 150 min at 5% B, 160 min at 5% B.

MS1 scans (200-2000 m/z) were collected with an Orbitrap mass analyzer at a resolution of 120,000 using quadrupole isolation; RF lens 30%, AGC target  $4.0e^5$ , and a 50 ms injection time. Precursor ions were fragmented using HCD (32%). MS2 scans were collected with an Orbitrap resolution of 30,000 using quadrupole isolation and AGC target  $2e^4$ . A charge state filter was used (+2-6) and the intensity threshold was set to  $2e^4$ .

Dynamic exclusion was set to exclude precursor ions for 60 s after collecting 10 MS2 scans within 30 s.

## 2.5 Database configuration and search parameters

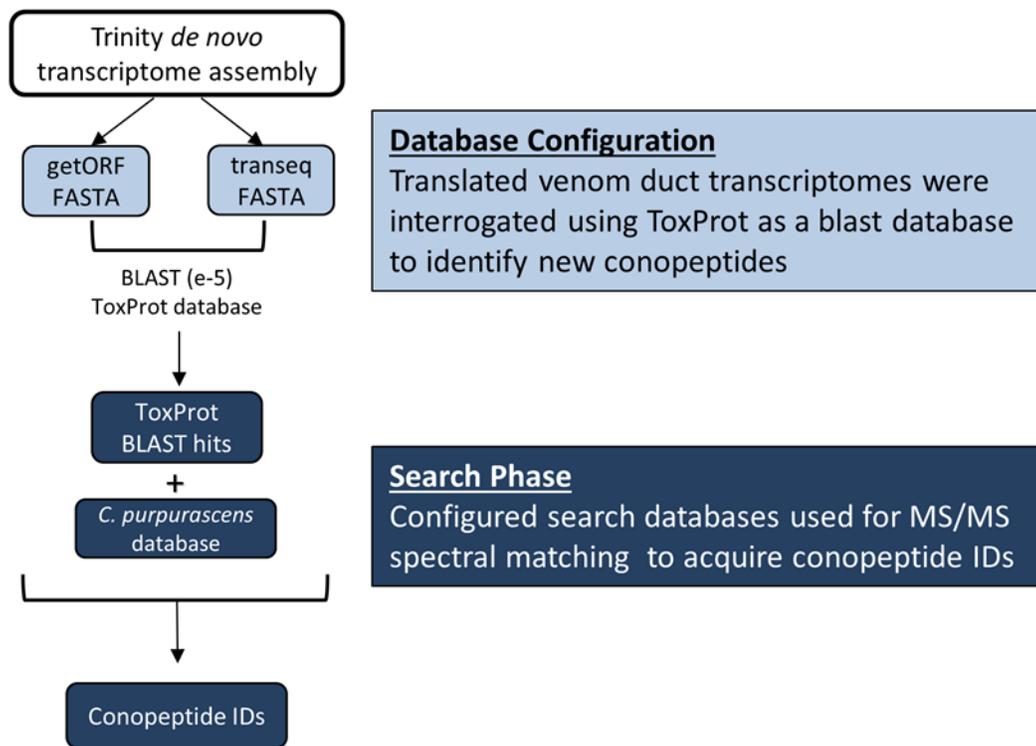
Several databases were configured and assessed for completeness before choosing the best search database for the 27 venom samples. The database was optimized for time-intensive non-enzymatic searches with many PTMs using the following criteria; inclusivity of conopeptide-encoding transcripts and the total number of entries. We compared the following four database configurations, all from the *de novo* transcriptome assemblies of venom duct RNAseq data. 1) The *de novo* assembly was blasted (blastX,  $e = 10^{-5}$ ) against the UniProt Animal Toxin Annotation database (ToxProt) and all UniProt *Conus* entries (TaxID: 6490), then translated ORFs were extracted with getorf, and complete transcripts with signal sequences were extracted with SignalP v4.0 [77]. 2) The *de novo* assembly was blasted as previously described using blastX, then the hits were translated with transeq, and only transcripts containing > 4 cysteines were extracted. 3) The *de novo* assembly was translated, the ORFs were extracted with getorf, and complete transcripts with signal sequences were extracted with SignalP (this configuration did not include a blast step). 4) Trinity assembly was translated with transeq and getorf, and resulting transcripts were blasted (blastp,  $e = 10^{-5}$ ) against the ToxProt database to extract toxin-like sequences. We chose the ToxProt-guided configuration (4) as the optimal search database, to which we added a customized *C. purpurascens* database that included conopeptide sequences not present in the transcriptomes (Table 3, Figure 6). The additional *C. purpurascens* database included previously identified peptides from UniProt (taxid: 41690) and unpublished conopeptide sequences identified in-house using the PEAKS *de novo* search algorithm (Bioinformatics Solutions Inc., Ontario, Canada) [78]. PEAKS can deduce peptide sequences from MS/MS spectra without a database.

PEAKS scored the predicted sequences with an average local confidence (ALC) score. In the in-house *C. purpurascens* database, only predicted conopeptide sequences with ALC scores greater than 98% were included.

Database searches were performed with the Sequest algorithm within Proteome Discoverer v2.0 (Thermo Fisher Scientific). Search parameters included a mass error of 10 ppm for precursor peptides and 0.02 Da for fragment ions. Fixed modification, Carbamidomethyl (C), was introduced and several previously reported conotoxin PTMs were introduced as variable modifications: oxidation (M/P), carboxylation (E), bromination (W), deamidation (N/Q), pyroglutamate (N-terminus), and amidation (C-terminus). The false discovery rate (FDR) threshold was set to 1% using a decoy database. Only high and medium confidence protein matches were considered for downstream analysis.

**Table 3- Search database file descriptions.**

<b>Sequence File</b>	<b>Description</b>	<b># Sequences</b>
1. Transcriptome A	Assembly of RNA-seq data from specimen 1	83,051
2. Transcriptome B	Assembly of RNA-seq data from specimen 14	84,410
3. Translated– getorf	Files 1 and 2 translated with getorf	231,175
4. Translated– transeq	Files 1 and 2 translated with transeq	1,004,766
5. ToxProt BLAST hits	Files 3 and 4 blasted against ToxProt database	1,652
6. <i>C. purpurascens</i> Database	<i>C. purpurascens</i> conopeptides including UniProt entries (taxid: 41690) and <i>de novo</i> sequences	40



**Figure 6- Workflow of the database search strategy.**

## 2.6 Hierarchal cluster analysis and Principal Component Analysis

Total intensities for each conopeptide were normalized to the highest intensity within each sample. Total intensities were normalized in this analysis to account for differences in protein concentration between venom samples. Hierarchal clustering and Principal Component Analysis were performed using ClustVis online software (v. 2018-12-20) [38]. Normalized intensities were log-transformed ( $\ln(x + 1)$ ) prior to hierarchal cluster analysis. Hierarchal clustering was employed on both x- and y- axes using Pearson correlation distance with average linkage.

## 2.7 Alpha-conotoxin testing on nAChR subtypes

### 2.7.1 Alpha-conotoxin identification, selection, and synthesis

To identify putative  $\alpha$ -conotoxin sequences, a conserved gene signal sequence was obtained from Conoserver.org and used as a search query against the 17 transcriptomes (BLASTp, e=10). Mature peptide sequences were predicted using SignalP software and the predicted sequences were aligned using Geneious Prime (Biomatters, Auckland NZ). From the alignment, we were able to assess inter-cysteine loop sizes and homology to other functionally characterized  $\alpha$ -conotoxins. Three peptides ( $\alpha$ -PID,  $\alpha$ -NuxIA,  $\alpha$ -CedIA) were chosen for functional assays on nAChRs and were synthesized by solid-state synthesis through a third-party company to yield milligram quantities (GenScript, Piscataway, NJ). The following post-translational modifications were incorporated: disulfide bonds between C1-C3 and C2-C4 and amidation of the C-terminal. Once received, peptide purity was confirmed by RP-HPLC separation on a C18 column.  $\alpha$ -PIA is commercially available and was purchased for functional screening (Alamone Labs, Jerusalem, Israel).

### 2.7.2 Oocyte harvesting and injection

Nicotinic receptor subtypes were heterologously expressed in *Xenopus laevis* oocytes. Frogs were maintained in the Animal Care Service facility of the University of Florida, and all protocols were approved by the University of Florida Institutional Animal Care and Use Committee. First, oocytes were surgically removed from female frogs. The oocytes were then injected with cRNA for the nAChR subunits required for the expression of one of seven different receptor subtypes (Table 4). After injection, the oocytes were stored in Barth's solution (88 mM NaCl, 1 mM KCl, 2.38 mM NaHCO<sub>3</sub>, 0.82 mM MgSO<sub>4</sub>, 15 mM

HEPES, and 12 mg/l tetracycline, pH 7.6) at 16 °C until electrophysiological experiments were performed.

**Table 4- nAChR subtypes with associated subunit RNA and ACh controls used for voltage-clamp experiments.**

nAChR subtype	Subunit RNA	ACh Control ( $\mu\text{M}$ )
$\alpha 3\beta 4$	h $\alpha 3$ h $\beta 4$	100
$\alpha 4\alpha 5\beta 2$	h $\beta 2$ -6- $\alpha 4$ h $\alpha 5$	10
$\alpha 4\beta 2$	h $\beta 2$ -6- $\alpha 4$ h $\alpha 4$	100
$\alpha 4\beta 2$	h $\beta 2$ -6- $\alpha 4$ h $\beta 2$	10
$\alpha 4\alpha 6\beta 2\beta 3$	h $\beta 3\alpha 4\beta 2\alpha 6\beta 2^*$	30
$\alpha 1\beta 1\epsilon\delta$	m $\alpha 1$ m $\beta 1$ m $\epsilon$ m $\delta$	30
$\alpha 7$	h $\alpha 7$ hRIC3	60

### 2.7.3 Two-electrode voltage-clamp recordings

Two-electrode voltage-clamp experiments were performed on an OpusXpress 6000A using pClamp software (Molecular Devices, San Jose, CA). Oocyte membrane potential was clamped at -60 mV and bath-perfused with Ringer's solution (115 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl<sub>2</sub>, 10 mM HEPES, and 1  $\mu\text{M}$  atropine, pH 7.2). Membrane current ( $\mu\text{A}$ ) was recorded throughout the following experimental protocol: pre-control (ACh), treatment application ( $\alpha$ -CNTX), co-application (ACh +  $\alpha$ -CNTX), and post-control (ACh). ACh controls were administered before and after  $\alpha$ -conotoxin treatments to establish the oocyte baseline response to ACh and the oocyte viability and receptor desensitization after treatment. The ACh control concentration for each receptor subtype is reported in Table 4.

#### 2.7.4 Data analysis and IC50 calculation

Each experiment will have a sample size of 4-8 oocytes. Individual oocyte recordings were eliminated from analysis if they were 1) unresponsive to the ACh pre-control 2) lost their 'clamp' or holding potential (60mV) during the experiment or 3) had significant desensitization to the post-control. Analysis of recordings was performed with Clampfit software (Molecular Devices, San Jose, CA). Treatment response was measured as peak current amplitude ( $\mu$ A). For each oocyte, the values for treatment response were normalized to the ACh controls to account for differences in receptor expression levels and ACh responsiveness between oocytes. T-tests were performed ( $n \geq 3$ ) for each treatment on each nAChR subtype to test the hypothesis that the response of ACh controls will be different from  $\alpha$ -conotoxin + ACh treatments ( $p \leq 0.05$ ). IC50 values were calculated using an inhibitory dose-response curve nonlinear regression model with the following equation:  $Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{((\text{LogIC50} - X) * \text{HillSlope}))}$ . When comparing  $\alpha$ -PIA and  $\alpha$ -PID dose-response curves the hill slope was constrained to -1 with the assumption that their binding mechanisms to the nAChR are similar.

#### 2.7.5 Homology modeling of the human $\alpha 7$ nAChR

The structure of the  $\alpha$ -PIA and  $\alpha$ -PID bound to the human  $\alpha 7$  nAChR (h $\alpha 7$ ) was modeled based on the co-crystal structure of *Aplysia californica* AChBP bound to conotoxin LvIA (PDB: 5XGL). First, the sequence of the extracellular domain of h $\alpha 7$  was obtained from Uniprot (ID: P36544). An alignment file was created by aligning the h $\alpha 7$  and  $\alpha$ -PIA/ $\alpha$ -PID sequence to the AChBP and LvIA sequence using the Clustal Omega alignment algorithm within Geneious Prime (Biomatters, Auckland, NZ). Homology modeling was performed using MODELLER (version 9.24) (University of California San Francisco)[79]. The resulting PDB files were visualized using UCSF ChimeraX software [80] and analyzed for molecular energy and residue contact distances using Molecular Operating Environment (MOE) (Chemical Computing Group, Montreal, Canada).

CHAPTER 3:

PROTEOGENOMIC ASSESSMENT OF INTRASPECIFIC VENOM VARIABILITY IN  
*CONUS PURPURASCENS*

### 3.1 Introduction

Venomous animals comprise over 200,000 species across several taxa and display a variety of mechanisms for venom production, delivery, and use [81]. For most animals, venom is proteinaceous; however, different taxonomic groups independently evolved to produce highly adapted venom as a solution to environmental pressures, a clear example of convergent evolution. Most venoms are complex mixtures of peptides, proteins, and small molecules that might act in concert to immobilize prey or deter predators. The specific molecular content of these composites varies from phyla, class, order, family, and genus. There can be also significant venom variability within the same species [82-90]. In some cases, venom varies within the individual specimens [83, 91-95], as some animals can switch their venom from predatory to defensive concoctions. Intraspecific venom plasticity expands the molecular adaptations of venomous animals and in doing so augments the remarkable repository of compounds with numerous applications that include the development of pharmaceuticals, such as Captopril from the Brazilian pit viper venom, Exenatide from the Gila monster, and Ziconotide/Prialt™ from cone snail venom [3].

The venom found in marine predatory snails belonging to the genus *Conus* (cone snails) has been intensely studied in terms of content and pharmacological properties. Most notable are the conotoxins, a diverse group of disulfide-constrained (two or more disulfide bonds) peptides that target ion channels, ligand-gated receptors, and transporters with high affinity and selectivity [96, 97]. *Conus* venom can also contain linear (no disulfides) and one-disulfide bond peptides [98], which along with conotoxins define the conopeptides, the full small peptidic complement of the venom of cone snails. Conopeptide diversity occurs at both the sequence and post-translational modification (PTM) level, resulting in thousands of conopeptides that range in size, chemistry, structure, and activity. Conopeptides/conotoxins are classified according to gene

superfamilies based on conserved signal sequences, and each superfamily can encode hundreds of mature conopeptide sequences [99-101]. Mature conotoxins have displayed a plethora of cysteine frameworks and disulfide-bonding patterns, which in turn affects activity. Conopeptide complexity also results from a high rate of PTMs [22, 102, 103]. The same base peptide can have many differentially modified forms [104, 105], or 'toxiforms'. Conopeptide hyper-modification can be viewed as an evolutionary expansion of venom genes used to engineer highly effective and potent toxins.

The molecular diversity of cone snail venom is extraordinary as its expression is species-specific with little overlap of components among the more than 800 extant species [106]. This complexity is compounded by intraspecific and intraspecimen venom variations due to predatory or defensive venom profiles [90, 91, 93]. This complexity provides a rich source of bioactive peptides [107, 108], but it also presents a challenge for venom characterization. Intraspecies studies have relied heavily upon comparisons of venom chromatography and mass-matching to known venom components, rather than global MS/MS spectral matching, to identify venom components. A limitation of this approach is that a single base conopeptide sequence can have many toxiforms with different masses. This makes it difficult and rather uninformative to assess intraspecific venom variation through molecular mass lists alone, and in doing so, it can lead to overestimates of the extent of venom variability. Next-generation sequencing technology for RNAseq and advances in high-resolution LC-MS/MS have mitigated the challenges associated with the analysis of complex venoms and have allowed assessment of the venom peptidome/proteome through 'venomic' approaches [109, 110].

A comprehensive analysis of the venom composition is crucial to assess venom plasticity and to determine synergistic mechanisms of envenomation used to immobilize prey or deter predators. Here, we present a large-scale intraspecific venom analysis of *Conus purpurascens*, the only fish-hunting species of the tropical Eastern Pacific region.

Earlier groundwork revealed that *C. purpurascens* had two distinct venom ‘cabals’, or groups of conopeptides acting synergistically to paralyze their prey [89, 111, 112]. The cabals act as either 1) a neuromuscular block (motor cabal), targeting nicotinic acetylcholine receptors (nAChRs,  $\alpha$ - and  $\psi$ -conotoxins) and skeletal muscle sodium channels ( $\mu$ - conotoxins) or 2) an excitotoxic neuronal block (lightning-strike cabal), targeting neuronal sodium ( $\delta$ -conotoxins) and potassium channels ( $\kappa$ -conotoxins). Previous works, however, were based on mass lists obtained from the venom of a limited number of specimens [83, 89].

We present a comprehensive venom analysis by utilizing high-resolution LC-MS/MS-based peptide identification to analyze and compare injected venom from 27 individual specimens of *C. purpurascens*. In doing so, we sought to maximize the identifications of conopeptides and their toxiforms. We also assessed the biochemical diversity of the venom arsenal by comparing conopeptide expression patterns to gain a more refined view of synergistic relationships among the venom components.

## **3.2 Results**

### *3.2.1 Conopeptide identification*

Milked venom samples from 27 specimens of *C. purpurascens* were analyzed by LC-MS/MS (Methods Section 2.4) and resulting MS/MS RAW files were searched against a FASTA database that consisted of conopeptide sequences either acquired from RNA-seq data or deposited in UniProt ([www.uniprot.org](http://www.uniprot.org)) (Methods section 2.5). The venom analysis yielded 543 unique conopeptide identifications, which included 33 base (or nascent) conopeptides and their associated toxiforms (modified versions). Of these 33 base conopeptides, 21 sequences were identified here for the first time (Table 5). Detailed descriptions of each new conopeptide are provided in Appendix C. Twenty-six of the conopeptides were identified through the transcriptome search database (RNA-seq-assembled transcripts). However, seven conopeptides were identified in the venom but

were not found in either of the transcriptomes. Four of these peptides were identified from *C. purpurascens* UniProt entries ( $\alpha$ -PIA,  $\alpha$ -PIB,  $\kappa$ -PIVF, PVIF), and three conopeptides were sequenced *de novo* and manually added to our in-house search database (Contryphan-P4, PIF, and PIG). For the conopeptides that were identified from RNA-seq data, full or partial transcripts were used to assign superfamilies through their corresponding signal sequence (Table 6).

The number of base conopeptide IDs per sample of injected venom ranged from 5-17 (mean=  $10.6 \pm 2.6$ ) (Figure 7). The most prevalent conopeptide in this population of snails was Ile-contryphan-P, identified in 25 of the 27 venom samples. This was followed closely by conopeptides  $\kappa$ -PVIIA, PVIIIA,  $\psi$ -PIIIE,  $\alpha$ -PIVA, and PVIB, all identified in more than 75% ( $n > 20$ ) of the venom samples (Figure 8).

Differentially modified toxiforms were identified for 27 of the 33 base conopeptides (Table 5). Toxiforms were only considered if the peptide maintained a complete cysteine framework. The PTMs identified through MS/MS analysis included amidated C-terminal, hydroxyproline, oxidized methionine, deamidated asparagine/glutamine, carboxyglutamate, brominated tryptophan, N-terminal pyroglutamate, and truncations from both terminals (Appendix A). The most abundant PTMs were C-terminal amidation and hydroxyproline, which occurred on 75% of the base conopeptides identified. The same modification(s) occurred on different residues of the same peptide, generating unique toxiforms with the same molecular weight. This is the case of hydroxyproline, which occurred on up to three residues simultaneously on four peptides:  $\alpha$ -PIVA,  $\psi$ -PIIIE, PVIE, and PVIG. Differential hydroxylation patterns are seen for these conopeptides (Appendix A). The greatest PTM variability was observed on A-superfamily conotoxins  $\alpha$ -PIVA (98

toxiforms) and  $\kappa$ -PIVE (69 toxiforms), and new O1-superfamily conopeptide PVIB (67 toxiforms) (Table 5).

**Table 5 Conopeptides identified from the injected venom of *Conus purpurascens*.** Conopeptides reported for the first time here are in bold. Sequence alignment was performed manually for each cysteine framework. \*Conopeptides identified with the PEAKS *de novo* software. \*\*These base conopeptides were previously described from cDNA libraries as P2B-D [1]; since these designations do not conform with current nomenclature they were renamed accordingly. The number of toxiforms only includes peptides identified with the full cysteine framework. ND = not determined because there were no spectral matches associated with the complete base peptide.

Superfamily	Conopeptide	Sequence	Toxiforms
A	$\alpha$ -PIA	RDPCCSNPVC <del>TVHNPQIC</del>	18
A	$\alpha$ -PIB	QSPGCCWNPAC-VKNR-C	6
A	$\alpha$ -PIC	TSGCCKHPAC-GKNR-C	1
A	<b>PID</b>	DPCCSNPACNVNNPQICG	11
A	<b>PIE</b>	NAAAKAFDLTAPTAGEGCCFNPA <del>CAVNNPNIC</del>	2
A	<b>PIF*</b>	QEPGCCRNPA <del>C-VKHR-C</del>	13
A	<b>PIG*</b>	PCCSNPVC <del>TVHGGPQLC</del>	2
A	$\alpha$ A-PIVA	GCCGSYPNAACHPC <del>CKDRPSYCGQ</del>	98
A	$\kappa$ -PIVE	DCCGVKLEM-CHP <del>CLDNSCKNYGK</del>	69
A	$\kappa$ -PIVF	DCCGVKLEM-CHP <del>CLDNSCKKSGK</del>	32
A	<b>PIVH</b>	DCCGVVMEE-CHK <del>CLCNQTC</del> KKK	45
B2	<b>Linear-P</b>	QPSAENE <del>E</del> GKFRFFDKQ	6
M	<b>Ile-Contryphan-P</b>	GCVIWPWC	7
M	<b>Contryphan-P3</b>	CAIWTKC	3
ND	<b>Contryphan-P4*</b>	CVYWRKC	1
M	$\psi$ -PIIIIE	HPPCCLY <del>GK</del> CRRYPG <del>CSSASCCQR</del>	27
M	<b>PIIIG</b>	QWGCCPVNACRS <del>CHCC</del>	2
M	<b>PIIIH</b>	KCCPLTACKL <del>GSGCKCCE</del>	7
M	<b>PIIII</b>	CCQA-YCSRYH <del>CLPCC</del>	1
O1	$\delta$ -PVIA	EACYPAGTF <del>CGIKPGLCCSEFCLPGVCFG</del>	3
O1	<b>PVIB</b>	QCTPYGGS <del>CGVD-STCCGR</del> CNVPRNK <del>CE</del>	67
O1	<b>PVIC</b>	EACYPAGTF <del>CGIKPGLCCSALCLPAVCID</del>	ND
O1	<b>PVID**</b>	PCKKSGRK <del>CFPHQK</del> DCCGRACIITICP	3
O1	<b>PVIE</b>	VGEFRGCAHINQA <del>CNPP-QCCRGYT</del> CQSSYIPS <del>CQL</del>	16
ND	<b>PVIF**</b>	ATSNRP <del>CKKTGRK</del> CFPHQK <del>DCCGRACIITICP</del>	3
O1	<b>PVIG**</b>	GATSNRP <del>CKIPGRK</del> CFPHQK <del>DCCGRACIITICP</del>	16
O1	$\kappa$ -PVIIA	CRIPNQK <del>CFQHLD</del> DCCSRK <del>CNRFN</del> CKV	18
O2	Contryphan-P	GCPWDPWC	1
O3	<b>PIIA</b>	CCCI <del>RS</del> DGPK <del>CSRK</del> CLSSFFC	2
S	<b>PVIIIA</b>	GCSGSP <del>CFKNKTCRDEC</del> ICGGLSN <del>CW</del> CGYGGSRG <del>CKCT</del> CRE	33
T	PVA	GCCPKQMR <del>CCTL</del>	2
T	<b>PVB</b>	DCCPEKM <del>WCPL</del>	11
Con-ikot-ikot	<b>p21b</b>	FELLPSQDRS <del>CCIRK</del> TLE <del>C</del> LENYPGQESQRAHY <del>CQ</del> QDATTN CPDTYDFG <del>CCPGYATCMS</del> INAQNNVRPAHDT <del>CINRL</del> CFDPG F	ND

**Table 6- Protein sequences of identified conopeptides.** Sequences were translated from RNA-seq transcripts from venom duct transcriptomes A (unshaded) and B (shaded). Mature peptides, as determined by MS/MS, are in bold.

Superfamily	Conopeptide	Transcript
A	PIC	MGMRMMFIVLLVVLATTVGSFTLDRVLGLASEGRNAEAIQNALDQRDPKRR <b>TSGCCKHPACGKNRC</b>
	PID	MGMRMMFIVLLVVLATTVGSFTLDRASDRDAAANDKASDLIALTARR <b>PCCSNPACNVNNPQICG</b>
	PIVA	MGMRMMFIVLLVVLATTVGSFTSDRASDDRNTNDKASRLLSHVVR <b>GCCGSYPNAACHPCCKDRPSYCGQGR</b>
		MGMRMMFIVLLVVLATTVGSFTSDRASDDRNTNDKASRLLSHVVR <b>GCCGSYPNAACHPCCKDRPSYCGQGR</b>
	PIVE	MGMRMMFIVLLVALATTVGSFTSDRASDGRNAAVNDKASHLIDNVIR <b>DCCGVKLEMCHPCLCDNSCKNYGKGKKEYGK</b>
B2	PIVH	VVLATTVGSFTSDRASDGRNAAVNDKASPLIAKAVIR <b>DCCGVVMEECHKCLCNQTCKKKGKELWEMMTATDKRNT</b>
	B2 linear	MLRLIIAAVLASACLAFPERRDGVPAEQANLQGFDPAAQAMPAMAGMQMPGMAGGQFLPFNPNFGMAYKRDMDEILEKRKQHSQFNADNESPFEAGDNLGDFM
		NFMKGNNGNVPFANMDSATDLGN <b>FQPSAENEEGKFRFFDKQQ</b>
M	Ile-Contryphan-P	MLKMGVLLFTFLVLFPLATLQLDADQPVVERYAEKQDLNPDERTKTLHALRPPSVDKRATSL <b>GCVIWPWC</b>
	Contryphan P3	MLKMGVLLFIFLVLPLATLQLDADQPVVERYAENKQDLKPDERRRILPALGPPSVDKRATSL <b>ACAIWTKC</b>
	PIIIE	MMSKLGALLTICLLLPITALLMDGDQPADRPAERMEDDISSEVHRLLERR <b>HPPCCLYGKCRRYPGCSSASCCQRG</b>
		MMSKLGALLTICLLLPITALLMDGDQPADRPAERMEDDISSEVHRLLERR <b>HPPCCLYGKCRRYPGCSSASCCQRG</b>
	PIIIG	LITLQLDADQPVVERYAEDKQDLNPNERMGFILPALR <b>QWGCCPVNACRSCHCCGRSTVALCWADSTATAVVDHVYRAHVSLRMTN</b>
O1	PIIHH	MLKMGVLLFTFLVLFPLATLQLDADQPVVERYAENKQDLKPDERRRILPALGQR <b>KCCPLTACKLGSCKCCE</b>
	PIIHH	MMFKLVLLTICLLLPITAIPLDGDQPVQPAERMEDGKSTPNHPWFDPVKR <b>CCQAYCSRYHCLPCC</b>
	PVIA	MKLTVCVMIVAVLFLTAWTFVTADDSKNGLENHFVKARDEMNREASKLDDK <b>EACYAPGTFCGIKPGLCCSEFLPGVCFGG</b>
		MKLTVCVMIVAVLFLTAWTFVTADDSKNGLENHFVKARDEMNREASKLDDK <b>EACYAPGTFCGIKPGLCCSEFLPGVCFGG</b>
	PVIB	MKLTVCVIVAVLFLTAQLITADDSRRTQKHRALRSTTKLSMLTR <b>QCTPYGGSCGV DSTCCGRCNVPRNKCE</b>
	PVIC	MKLTVCVMIVAVLFLTAWTFVTADDSKNGLENHFVKARDEMNREASKLDDK <b>EACYAPGTFCGIKPGLCCSALCLPAVCIDG</b>
	PVID (P2b)	MKLTVCVIVAVLFLTAQLITADDSRRTQKHRALRSTTKRARSNR <b>PCKKSGRKCFFHQKDCCGRACIITICP</b>
	PVIE	MKLTCVLIIAVLFLTAQLITAGYSRDKQVYRAVRLGDKMLR <b>VGEFRGCAHINQACNPPQCCRGYTCQSSYIPSCQL</b>
	PVIG	MKLTVCVIVAVLFLTAQLITADDSRRTQKHRALRSTTK <b>GATSNRPKIPGRKCFPHQKDCCGRACIITICP</b>
	PVIIA	MKLTVCVIVAVLFLTAQLITADDSRRTQKHRALRSTTKLSLSTR <b>CRIPNQKCFQHLDCCSRKCNRFNKCV</b>
	MKLTVCVIVAVLFLTAQLITADDSRRTQKHRALRSTTKLSLSTR <b>CRIPNQKCFQHLDCCSRKCNRFNKCV</b>	
O2	Contryphan-P	MGKLTILLVAAVLLSTQVMVQGDGDQPAYRNAAPRDDNP <b>GGAIGKFMNVLRRSGCPWDPWCG</b>
O3	PIIA	MSRFGIMVLTFLLLVSMATSHRYARGKQATRNRNAINIRRRSTPKTEACEEVECELEEKH <b>CCCIRSDGPKCSRKCLSSFFC</b>
S	PVIIIA	MMSKMGAMFVLLLLFTLASSQEGDVQARKRLTRDFYRTL <b>VPVSTRGCSGSPCFKNTCRDECICGGLSNCWCGYGGSRGCKCTCRE</b>
		MMSKMGAMFVLLLLFTLASSQEGDVQARKRLTRDFYRTL <b>VPVSTRGCSGSPCFKNTCRDECICGGLSNCWCGYGGSRGCKCTCRE</b>
T	PVA	MRCLPVFVILLLLIASAPSVDAHPKTKDDMPLASFHDNAKRTLQRLWKKR <b>GCCPKQMRCCTLG</b>
	PVB	MHCLPVFVILLLLIASAPSVDAHPKTKDDMPLASFHDNAKRTLQRFWKKR <b>DCCPEKMWCPLG</b>
		MRCLPVFVILLLLIASAPSVDAHPKTKDDMPLASFHDNAKRTLQRFWKKR <b>DCCPEKMWCPLG</b>
Con-ikot-ikot	P21b	MNMSMTLSMFVMVVAATV <b>TGFELLPSQDRSCCIRKTLLEENYPGQESQRAHYCQQDATTNCPDITYDFGCCPGYATCMSINAQNNVRPAHDTICINRLCFDPGF</b>

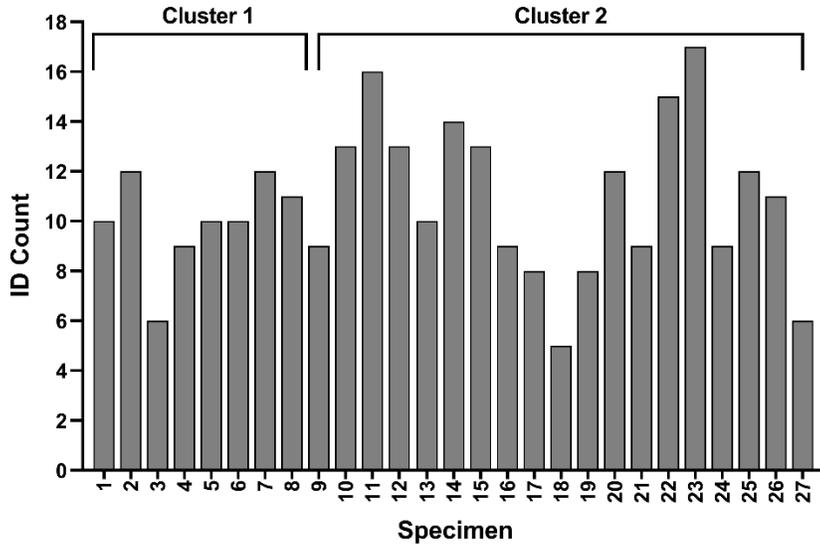


Figure 7- Conopeptide IDs for 27 *C. purpurascens* injected venom samples.

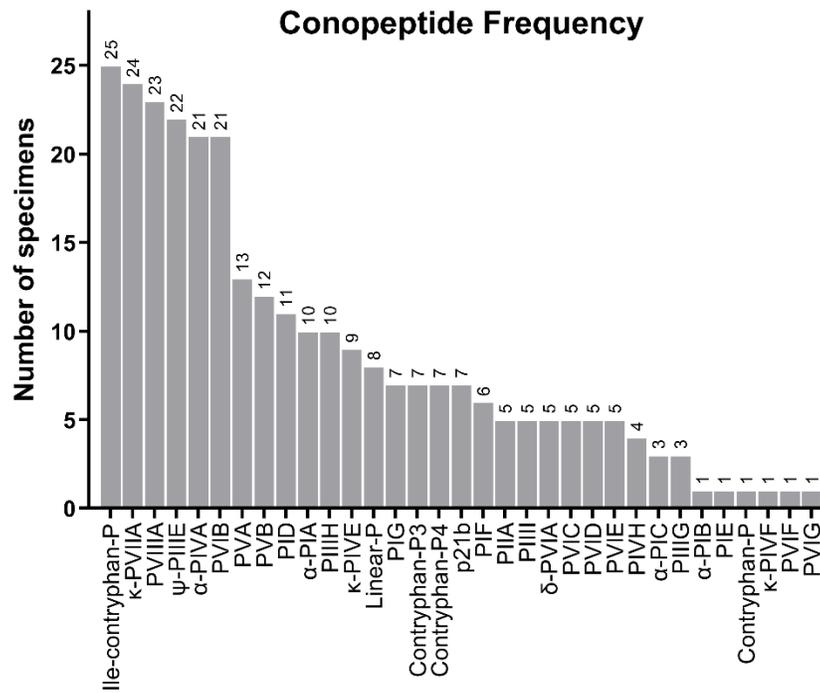


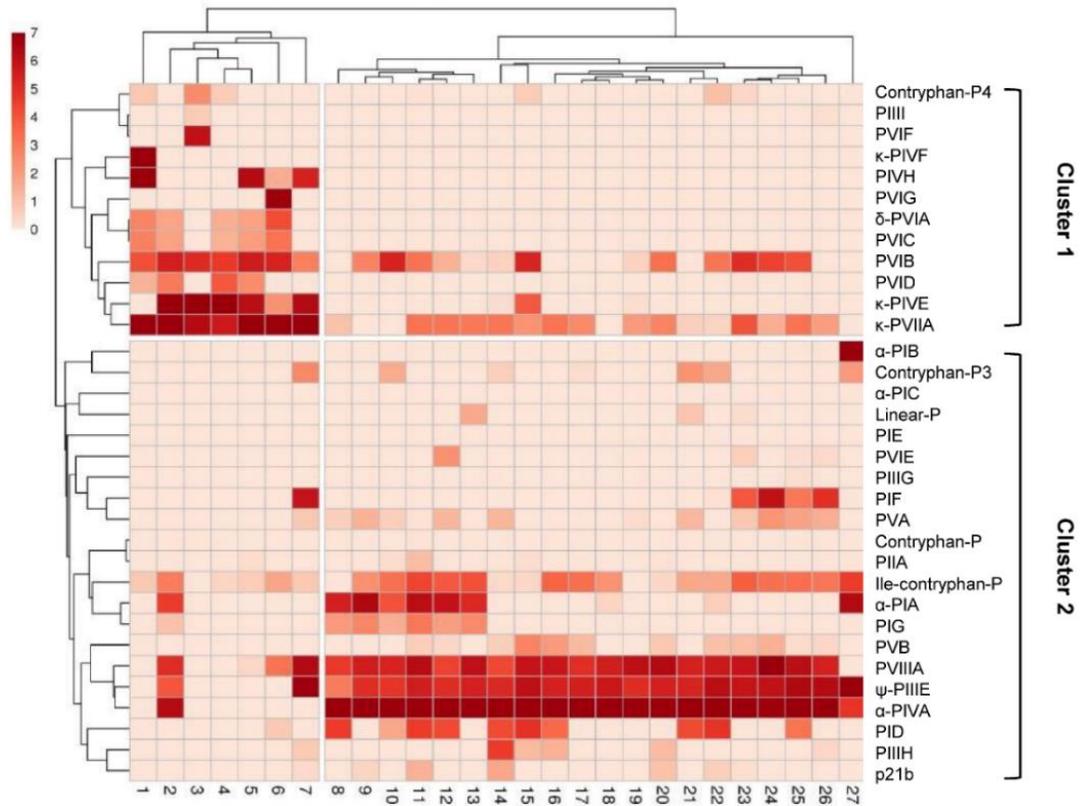
Figure 8- Conopeptide frequency in injected venom of *C. purpurascens*. N= 27

### 3.2.2 Intraspecific venom comparison

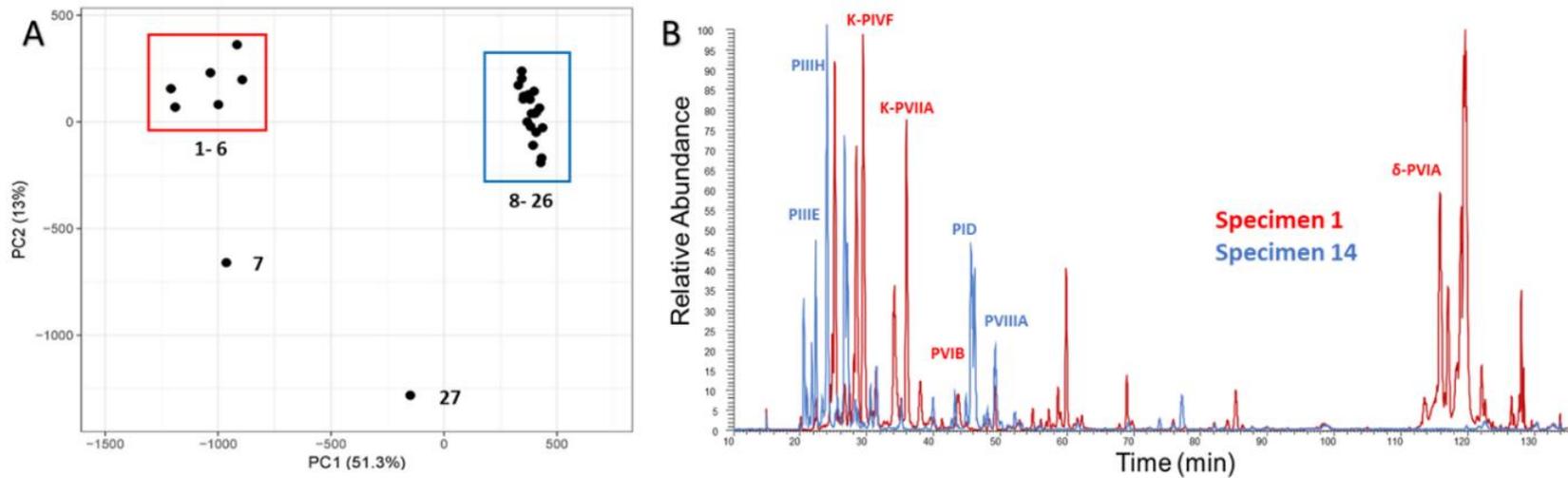
Hierarchical cluster analysis was employed to compare the venom profiles based on the total ion intensity of each base peptide (Methods Section 2.6). Total ion chromatograms (TIC) of injected venom (Appendix B) and conopeptide profiles varied among the 27 samples. Two groups were distinguished from cluster analysis of the 33 base conopeptides, specimens 1-7, and 8-27 (Figure 9). Clustering along the y-axis distinguished two groups of conopeptides that correlate to different venom compositions. The first cluster (snails 1-7) is mainly comprised of  $\delta$ - and  $\kappa$ -conotoxins that target the sodium and potassium channels, respectively. These conotoxins make up the 'lightning strike' cabal that rapidly immobilizes prey by acting on ion channels. The second cluster (snails 8-27) contains  $\psi$ - and  $\alpha$ - conotoxins that both act on nAChRs and make up the 'motor cabal'. Principal Component Analysis supported this dual expression pattern in the venom and clustered samples into two distinct groups of specimens 1-6 and specimens 8-26, with specimens 7 and 27 as outliers (Figure 10A). An overlay of chromatograms from specimen 5 from cluster 1 (blue) with specimen 14 from cluster 2 (red) emphasizes the distinction in venom profile components between the two clusters (Figure 10B).

The conopeptide identifications were made from venom gland transcriptome databases of two *C. purpurascens* specimens (transcriptomes A and B). These specimens correspond to specimen venom samples 1 (snail sacrificed for transcriptome A) and 14 (snail sacrificed for transcriptome B). To assess the coverage of the milked venom sample by the corresponding transcriptome, we compared conopeptide expression between these two specimens (Table 7). Our comparison examines the expression of each peptide between the two specimens at both transcriptomic (TPM) and proteomic (relative intensity) levels. Conopeptides expressed in the venom gland but not identified in the injected venom sample are shown in grey. Conopeptides identified in injected venom sample, but not expressed in venom gland transcriptome are shown in blue. Overall, we see a

differential expression pattern between the two specimens and between transcriptomic and proteomic expression within the same specimen.



**Figure 9- Conopeptide profiles from 27 *C. purpurascens* specimens.** Ion intensities were normalized to the highest value for each specimen and  $\ln(x + 1)$ -transformed. Clusters were determined by hierarchical cluster analysis using correlation distance and average linkage.



**Figure 10- *C. purpurascens* expresses two distinct venom profiles.** A) PCA analysis of normalized ion intensity for all conopeptide IDs (n= 27). B) Total Ion Chromatogram (TIC) overlay of Specimen 14 from cluster 1 (blue) and specimen 1 from cluster 2 (red).

**Table 7- Comparison of transcriptome expression and MS identification from two specimens of *C. purpurascens*** (Specimen 1-Transcriptome A, Specimen 14- Transcriptome B). Differential patterns of expression are shown by shading. Conopeptides expressed in the venom duct but not identified in the injected venom sample are shown in grey. Conopeptides identified in injected venom sample, but not

Superfamily	Conopeptide	Specimen 1			Specimen 14		
		Transcriptome A (TPM)	MS Intensity	Relative	Transcriptome B (TPM)	MS Intensity	Relative
A	α-PIA	0.00	0.00		0.00	0.00	
A	α-PIB	0.00	0.00		0.00	0.00	
A	α-PIC	6657.49	0.00		1214.88	0.00	
A	PID	139.43	0.00		0.00	67.16	
A	PIE	25.81	0.00		50.15	0.00	
A	PIF	0.00	0.00		0.00	0.00	
A	PIG	0.00	0.00		0.00	0.00	
A	α-PIVA	8002.60	0.00		7115.58	1000.00	
A	κ-PIVE	0.00	0.00		1043.68	0.22	
A	κ-PIVF	0.00	989.76		0.00	0.00	
A	PIVH	3962.08	1000.00		0.00	0.00	
B2	Linear-P	7081.46	0.00		2802.38	0.00	
M	Ile-Contryphan-P	1985.81	1.23		0.00	0.38	
M	Contryphan-P3	0.00	0.00		512.22	0.87	
?	Contryphan-P4	0.00	1.29		0.00	0.00	
M	ψ-PIIE	1926.10	0.00		770.99	186.94	
M	PIIG	6.84	0.00		0.00	0.00	
M	PIIH	0.00	0.00		2988.25	97.70	
M	PIII	0.00	0.02		8.32	0.00	
O1	δ-PVIA	2642.86	13.74		3651.60	0.00	
O1	PVIB	0.00	58.13		257.39	0.75	
O1	PVIC	632.78	14.82		0.00	0.00	
O1	PVID	363.92	3.02		0.00	0.00	
O1	PVIE	0.00	0.00		152.04	0.01	
?	PVIF	0.00	0.00		0.00	0.00	
O1	PVIG	0.00	0.00		569.10	0.00	
O1	κ-PVIIA	934.23	935.37		1387.09	21.09	
O2	Contryphan-P	57382.30	0.00		4647.95	0.00	
O3	PIIA	22.33	0.00		0.00	0.00	
S	PVIIIA	4082.19	0.00		1193.34	74.22	
T	PVA	340.10	0.00		0.00	2.68	
T	PVB	73.58	0.00		403.84	0.83	
Con-ikot-ikot	P21b	0.00	0.00		71.93	4.46	

### 3.2.3 Novel S-superfamily conotoxin- PVIIIA

PVIIIA is one of the 21 newly identified conopeptides and is the first member of the S-superfamily found in injected venom. The peptide has five disulfide bonds and exhibits cysteine framework VIII (C-C-C-CX<sub>aa</sub>C-CX<sub>aa</sub>C-CX<sub>aa</sub>CX<sub>aa</sub>C). It was expressed in high frequency and abundance within this *C. purpurascens* population. It was identified in 23 of the 27 venom samples (Figure 8). When venom profiles were compared, PVIIIA expression clustered closely with  $\alpha$ -PIVA and  $\psi$ -PIIE, which both target nicotinic receptors as part of the motor cabal (Figure 9). Alignment with functionally characterized S-superfamily conotoxins, known to target serotonin ( $\sigma$ -GVIIIA) and nicotinic ( $\alpha$ -GVIIIB,  $\alpha$ -RVIIIA) receptors, exhibits very little sequence homology aside from the conserved cysteine framework (Figure 11). PVIIIA is 41 residues in length and has 5 sites of modification, as determined by MS/MS spectral matching. We mapped all identified PTMs for the 33 toxiforms of PVIIIA (Appendix A). The following sites of modification were determined: oxidated Pro(6), carboxylated Glu(16), deamidated Asn(10, 24), and truncations on both N- and C- terminals. These modifications occur in most possible combinations, significantly expanding the molecular diversity of the PVIIIA base peptide. We also compared toxiform expression among the 27 specimen samples (Figure 12). The heatmap shows two clusters of peptides, which correlate to high abundance (top cluster) and lower abundance (bottom cluster). The six toxiforms in the top cluster exhibit the highest expression within the samples, as shown by color, and also within the population

PVIIIA	GCSGSPCFKN--KTCRDECI CGGLSNCWCGYG-GSR--GCKCTCRE-
$\alpha$ -GVIIIA	GCTR-TCGGP---KCTGTCTCTNS SKCGCRYNVHPSGWGCGACSG-
$\alpha$ -GVIIIB	SGSTCTCFTS--TNCQGSCECL SPPGCYCSNN-GIRQPGCSTCPGT
$\alpha$ -RVIIIA	KCNFDKCKGTGVYNCGE SCSCEGLHSRCRTYNI GSMKSGCACICTYY

**Figure 11- Alignment of PVIIIA with characterized S-superfamily conotoxins  $\alpha$ -GVIIIA,  $\alpha$ -RVIIIA and  $\alpha$ -GVIIIB.** Cysteines are highlighted in yellow to emphasize a conserved cysteine framework.



are quite significant, as the venom is a valuable source of bioactive peptides that can be used as neuronal probes and developed as novel therapeutic agents. Several conopeptides have reached clinical trials, including the approval of Prialt™, among the most powerful painkillers known [18].

Analysis of the intrinsic complexity of cone snail venom has been significantly advanced with the advent of NGS transcriptomic data that provides thousands of novel putative conopeptide sequences-- a trend that will continue to expand. It is critical to probe venom using proteomic approaches, as transcriptomic data on its own can only provide putative sequences. Large scale top-down proteomics/peptidomics is the best way to assess *de facto* PTMs and cleavage sites to generate mature conopeptides. We sought to maximize venom coverage through conopeptide identifications; however, practical aspects of these workflows, such as the number of available transcriptomes, size of the conopeptides suitable for “top-down”/enzyme-free methods, and unforeseen PTMs, may have an effect on the final coverage of components obtained. While recognizing these limitations for complete venom coverage, we were able to increase component identification by including sequences discovered through *de novo* methods and sequences previously reported for *C. purpurascens* to our search database. Regardless of the total coverage obtained, our results reveal a clear picture of the venom profiles and envenomation strategies employed by *C. purpurascens*.

We show through a functional proteogenomic comparison between specimens 1 and 14 that transcriptomic data from the venom gland does not provide complete coverage of the venom components. We identified conopeptides in the injected venom that were not represented at the transcript level, demonstrating the lack of homogeneity between the venom gland transcriptome and the injected venom. Of the 17 conopeptides reported in UniProt for *C. purpurascens*, 7 were not found in either venom gland transcriptome ( $\alpha$ -PIA,  $\alpha$ -PIB,  $\mu$ -PIIIA,  $\psi$ -PIIIF,  $\kappa$ -PIVF, p21a, conantokin-P, and Leu-contryphan-P). By

combining transcriptomes from two specimens into a search database, we were able to increase our total proteome coverage of the venom. However, these results emphasize that a transcriptome is a snapshot of gene expression at the precise moment the animal was sacrificed for mRNA extraction and cannot be used alone to fully describe the dynamics of venom expression. Other limitations include missing toxin transcripts during the transcriptome assembly process, as *de novo* assemblers can face difficulties when attempting to process large numbers of closely related transcript isoforms [115].

We aimed to achieve high-confidence peptide identifications to help describe the molecular mechanisms of predations utilized by this population of *C. purpurascens*. Our venomics approaches led to the identification of 543 conopeptides, which are the result of 33 base sequences and their corresponding toxiforms, significantly expanding the current inventory of *C. purpurascens* conopeptides. As expected, these are only a fraction of the putative conopeptide base sequences predicted by transcriptomic expression or by the number of unique masses deconvoluted at the MS1 level [89]. We were able to ascertain numerous toxiforms from the 33 identified base conopeptide sequences. Cone snails have the remarkable ability to engineer their venom peptides through hyper-modification, a molecular adaptation to hunting strategy [22, 103, 104]. These PTMs may have important implications for the development and molecular engineering of novel peptide-based therapeutics [116-118]. Using spectral matching we were able to detect sites of differential hydroxylation and carboxylation, which could not be deciphered through mass matching alone.

Our results emphasize the importance of identifying venom components from the injected venom, the actual brew delivered into prey. This is in striking contrast to intraspecific studies that utilized dissected venom [85, 119], which neglect venom processing and delivery at several levels. This is the first study using high-resolution mass spectrometry, transcriptomic data and *de novo* approaches on the injected venom of a

large group of individuals of the same species for the global identification of components, assessment of venom dynamics, and evaluation of synergistic interactions between conopeptides and their potential pharmacology.

The conopeptide composition of the predatory injected venom arsenal of *C. purpurascens* consists of cysteine-constrained peptides that range from 1 disulfide bond (contryphans) to 5 disulfide bonds (PVIIIA and p21b). The outlier is the linear peptide (Linear-P) belonging to the B2-superfamily. The molecular masses ranged from 938 Da (Contryphan-P3) to 4960 Da (PVIIIA), indicating a wide spread of molecular features of these venom components. These venomes are covered by conotoxin frameworks I-VIII, X, and 21. Except for frameworks II, VIII, and 21, 3D structural information exists to help assign disulfide bonding and folding patterns [120-123] to these newly discovered *C. purpurascens* conotoxins. This is particularly true for the well-studied  $\alpha$ -conotoxins (framework I) and  $\kappa$ -,  $\delta$ -conotoxins (frameworks VI, VII). Structural assignments of the more complex frameworks, such as those found in PVIIIA and p21b (5 disulfide bonds), remain a challenge. While the structural and even functional features of novel base sequences can be predicted by homologies, such as PID, PIE, PIF (which are homologs of other well-characterized  $\alpha$ -conotoxins), others such as PIIA, PIIG-I, PVIE, and PVIIIA have no significant homology to functionally characterized conotoxins; and therefore their activity and role in the envenomation strategy will require further investigation.

Hierarchical cluster analysis of the venom profiles of 27 specimens enabled us to ascertain strong linkages and possible synergisms between specific conopeptides through co-expression patterns. We found two distinct clustering patterns indicating that two different venom cabal combinations can be employed by *C. purpurascens* as a hunting strategy. Cluster 1 contains classical members of the lightning strike cabal, affecting neuronal transmission by disrupting the propagation of action potentials ( $\delta$ -PVIA,  $\kappa$ -PVIIA,  $\kappa$ -PIVE), but not apparent members of the motor cabal, comprising paralytic toxins acting

primarily on nicotinic muscular targets ( $\alpha$  or  $\alpha A$  conotoxins). These findings provide a significant revision to the original venom cabal configurations for *C. purpurascens*. The original cabal concept was introduced by the synergy of conotoxins  $\kappa$ -PVIIA and  $\delta$ -PVIA (the lockjaw peptide) found in the pooled venom from several individuals of *C. purpurascens* collected in the Gulf of California [111, 113]. However, when using pooled venom, the lightning strike cabal would be complemented with members of the motor cabal that includes several inhibitors of nAChRs such as a  $\alpha A$  (PIVE, PIVF, PIVG) and  $\psi M$  (PIII-I) conotoxins, which is not the case for individuals within cluster 1 (non-pooled samples). Since conotoxins PIVE,  $\kappa$ -PVIIA,  $\delta$ -PVIA, and their respective toxiforms, and novel conotoxins, PIVH, PVIB, PVIC, PVID (and toxiforms), are part of cluster 1, the latter appear to complement the lightning strike cabal within those *C. purpurascens* individuals.

Cluster 2 contains several inhibitors of nAChRs such as  $\alpha$ -PIA-F,  $\alpha A$ -PIVA, and  $\psi$ -PIII-E conotoxins in addition to components of the lightning strike cabal,  $\delta$ -PVIB and  $\kappa$ -PVIIA (also present in cluster 1). This is an indication of the use of multiple cabals as the primary arsenal of this population of *C. purpurascens*. The role of PVIIIA is intriguing, as it is highly expressed in cluster 2, but given the abundance of nicotinic inhibitors already present there, it would be unlikely that another more complex nicotinic inhibitor is necessary to complete the motor cabal. Another curious finding within cluster 1 was the presence of mini-M conotoxins PIIIG, PIIIH, and PIII-I. While these conotoxins are prevalent in worm and mollusk-hunting *Conus* species [124, 125], they have not been found in the injected venom of fish hunting species until now. The significance of this finding is under investigation.

We use PVIIIA as an example to demonstrate toxiform variability among the 27 venom samples. A heatmap of PVIIIA toxiforms shows two clusters of peptides (Fig. 6), high abundance (top cluster) and lower abundance (bottom cluster). While the toxiform

comparison does not provide insight into its role in the venom, it can help distinguish which toxiforms are most abundant within the population and provide leads for downstream bioactivity assays.

Populations of cone snails in different habitats and geographical locations can show different venom phenotypes, as seen in *C. purpurascens* venom studies carried out on animals from the Gulf of California [111], The Clipperton atoll [126], Ecuador [127], Panama [83] and Costa Rica [90] showing profound differences in venom profiles. For example, p21a, a conotoxin with the putative ability to modulate AMPA receptors, was found in a *C. purpurascens* specimen from Ecuador [127], but not in the animals from Costa Rica here studied. However, the homologous conotoxin p21b was found as part of cluster 2, but not cluster 1. Given the differences in cabals between clusters 1 and 2, it is likely that p21b participates in the lightning strike cabal within cluster 2 instead of PVIA, which is absent in this cluster. Habitat is critical to these slow-moving creatures as they must adapt to very localized areas. Part of this adaptation process will be venom production to capture prey that is prevalent to these microhabitats. Accordingly, venom profiles that we found might be a product of such an adaptation. This adaptation appears to be imprinted over their development in the wild, as upon captivity, the venom remains invariant as these animals were fed and kept under identical conditions.

Despite extensive studies on *C. purpurascens* through decades, using HR-MS/MS spectral matching, we have revealed a deeper coverage of the components of the injected venom from 27 specimens of *C. purpurascens*. Furthermore, we have shown the dramatic venom variations from specimen to specimen and the dynamic interaction of components as revealed by two patterns of synergism. These findings further develop the cabal concept in several ways. 1) The expanded reach of components due to the hypermodification to generate a plethora of toxiforms, 2) novel components belonging to distinct cabals, and 3) the possibility of multiple cabals operating independently within the same

geographical group of individuals. In addition to providing the strongest evidence of venom cabals to date, these findings will allow us to predict molecular targets of uncharacterized conopeptides based on global expression patterns. These analyses will tremendously aid the convoluted process of developing conotoxins/conopeptides into valuable molecular probes or therapeutics.

CHAPTER 4:

INSULIN-LIKE PEPTIDE FROM THE INJECTED VENOM OF

*CONUS PURPURASCENS*

## 4.1 Introduction

Insulin and insulin-like peptides (ILPs) belong to a superfamily of 6-9 kDa hormone peptides that are involved in growth and metabolism. ILP gene sequences encode signal peptides followed consecutively by B-chain, C-chain, and A-chain peptides. The protein is proteolytically processed into its active form containing A-and B-chain peptides connected by two disulfides [128]. Human insulin has provided critical information on the structure and function of the insulin scaffold, including the key amino acid residues needed for insulin dimerization and receptor binding [129-131]. These findings enabled the production of human recombinant insulin therapy, Humulin [132], and its fast-acting analog, Humalog, or insulin lispro [133]. The commercialization of insulin has significantly prolonged the lifespan and improved the quality of life for hundreds of millions of diabetics worldwide [134].

The insulin superfamily is a diverse group of hormone peptides that are found throughout animal phylogenies, from unicellular organisms to the most complex vertebrates [135]. In vertebrates, the superfamily includes insulin, insulin-like growth factors (IGF-1 and IGF-2), and relaxin, which are involved in glucose metabolism, growth, and pregnancy, respectively. Invertebrates have greatly expanded the function of the insulin superfamily through a diversity of ILPs. Unlike their vertebrate counterparts, invertebrate ILPs are the products of multi-gene families. This gene expansion is reflected through tissue heterogeneity and the multi-faceted physiological role of ILPs within invertebrate systems [136-139]. Examples of well-studied invertebrate ILPs include the insect bombyxins [140], molluscan insulin-like peptides (MIPs) [141, 142], and *Drosophila* insulin-like peptides (DILPs) [143]. Invertebrate ILPs are involved not only in carbohydrate metabolism and growth, but also in reproduction, diapause, aging, and immunity [144].

Gene diversification has led to the incorporation of endogenous peptide hormones into animal venom as an evolutionary tactic to disrupt the normal endocrine function in the prey [145, 146]. Because venom peptide hormones mimic endogenous hormones, they are promising candidates for drug development. For example, the glucagon-like peptide, exenatide, from the Gila monster saliva, mimics the exogenous hormone incretin that helps increase insulin release after a meal and it was developed into Byetta™, a drug used for the management of Type II-diabetes [147].

The first record of an insulin-like peptide (ILPs) from animal venom came from *Conus geographus* (Con-Ins G1), a fish-hunting cone snail species from the Indo-Pacific region [148]. Cone snail venom is a complex mixture of peptides, proteins, and small molecules that contains several classes of hormone-like peptides, such as conopressins (oxytocin/vasopressin analogs) [149], conoCAPS (crustacean cardioactive peptide-like peptides) [150], and RFamides [151]. Proteomic analysis of *C. geographus* venom revealed a peptide resembling fish-like insulin that when synthesized and tested, decreased blood glucose in feeder fish, causing rapid immobilization. Since then, transcriptomic data from venom ducts have revealed a diversity of ILPs from across the *Conus* genus. Some *Conus* species use 'weaponized-insulin' in their venom that more closely resemble vertebrate insulin than MIPs [152]. These vertebrate-like insulins lack an extra cysteine pair that is found in MIPs. Con-Ins G1 shows modest activity against the human insulin receptor, despite low homology to human insulin [153]. Conoinsulins from other fish-hunting species, *C. tulipa* and *C. kinoshitai*, also bind and activate the human insulin receptor [154]. Conoinsulins also contain unique post-translational modifications (i.e., carboxylated glutamic acid) that may enhance their ability to bind and activate the insulin receptor, as suggested by increased receptor binding activity of Con-Ins G1 as compared to the PTM-free peptide [153].

The functionality of conoinsulins at the human insulin receptor has opened the door for screening these venom ILPs as pharmacological agents. Recently, Con-Ins G1 has been used as a scaffold for developing a minimized human insulin peptides (mini-Ins) [155]. Mini-Ins is a truncated monomeric insulin peptide that binds and activates the insulin receptor with comparable potency to human insulin. Using Con-Ins G1 as a model, alternative binding mechanisms were determined that allowed for engineering minimized yet fully functional human insulin peptides.

Here, we describe a new conoinsulin (Con-Ins P1) found in the injected venom of *C. purpurascens*, a fish-hunting cone snail that inhabits the tropical Eastern Pacific region. *C. purpurascens* venom is well-studied and is known to contain conotoxins targeting sodium channels [113, 156], potassium channels [157], nicotinic acetylcholine receptors [158-161], as well as enzymes such as hyaluronidases [162], ACE, ECE [163] and conodipines [114]. Con-Ins P1 differs considerably in sequence and arrangement from other conoinsulins as it has a not truncated B chain. It is the first discovery of a hormone-like peptide from *C. purpurascens* venom, and the first direct evidence of conoinsulins being used in injected venom for prey capture.

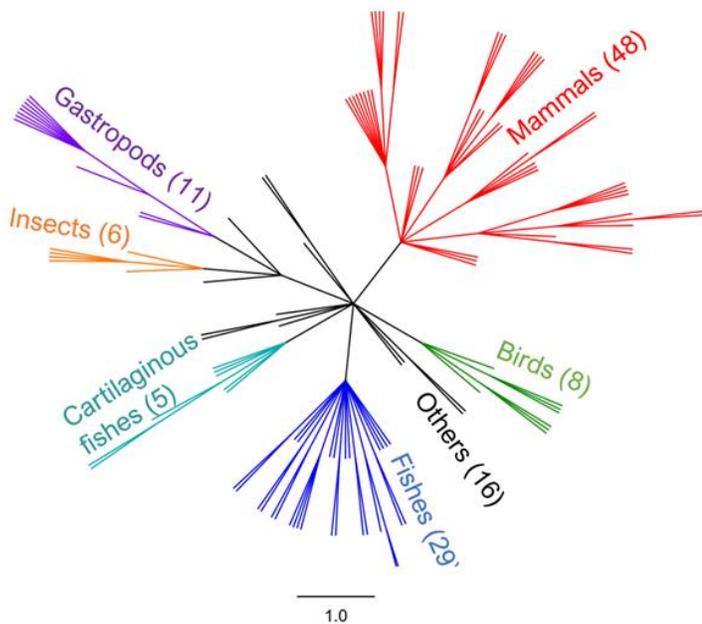
## **4.2 Results**

### *4.2.1 Phylogenetic analysis of ILP expression*

Insulin superfamily proteins were compiled from the InterPro database (IPR036438). The InterPro database contained a total of 5,000 entries for the insulin superfamily (258 reviewed), which included 942 unique taxa (123 reviewed). When considering reviewed entries alone, 85% of them are from chordates, and ~50% of these are from mammals (Figure 13). Invertebrate taxa include gastropods, insects, echinoderms (starfish), and nematode species. Gastropod taxa comprise ~10% of the InterPro insulin superfamily database; 65 entries from 31 *Conus* species and 2 entries from the venomous turrid sea snail, *Unedogemmula bisaya* (Tax ID: 746885). The reviewed proteins include 28 entries

from gastropods; 22 entries from cone snails and the remainder from model species *Aplysia californica* (Tax ID: 6500) and *Lymnaea stagnalis* (Tax ID: 6523) (Appendix D).

Within the unreviewed *Conus* entries, 13 sequences exhibit the cysteine framework of vertebrate-like insulins rather than MIPs (Figure 14). These vertebrate-like insulins come from *C. kinoshitai* (2), *C. geographus* (6), *C. tulipa* (4), and *C. lenavati* (1). All the above are fish-hunting species except for *C. lenavati*, which is a worm-hunter.



**Figure 13- Phylogenetic tree of the insulin superfamily.** Entries are from Interpro database (IPR036438, reviewed).

Identity 

```

tr|A0A3S9V8L7|Con-Ins K1 (C. kinos... MTTSSYFLLVALGLLLLYVQSSFGSPHTSDSGTTLVRRRLCGSELVT-YLGELCLG--NR-----
tr|A0A3S9V8K6|Con-Ins K2 (C. kinos... MTTSSYFLLVALGLLLLYVQSSFGNPHTRDSGTTTPDRDHS CGGELVD-RLVKLCPN--NR-----
sp|A0A0B5AC86|Con-Ins G3b (C. ge... MTTSSYFLLVALGLLLLYVQSSFGNQHTRNSDTP---KHRCGSELADQYV-QLCHG--KRNDAGK
tr|X5IWU3|(C. geographus) MTTSSYFLLVALGLLLLYVQSSFGNQHTRNSDTP---KHRCGSELADQYV-QLCHG--KRNDAGK
sp|A0A0B5A8P4|Con-Ins G3 (C. geo... MTTSSYFLLVALGLLLLYVQSSFGNQHTRNSDTP---KHRCGSELADQYV-QLCHG--KRNDAGK
sp|A0A0B5A7P2|Con-Ins G1c (C. ge... MTTSSYFLLMALGLLLLYVQSSFGNQHTRTFDTP---KHRCGSEITNSYM-DLCYR--KRNDAGK
sp|A0A0B5A8Q2|Con-Ins G1b (C. ge... MTTSSYFLLMALGLLLLYVQSSFGNQHTRTFDTP---KHRCGSEITNSYM-DLCYR--KRNDAGE
sp|A0A0B5AC95|Con-Ins G1a (C. ge... MTTSSYFLLMALGLLLLYVQSSFGNQHTRTFDTP---KHRCGSEITNSYM-DLCYR--KRNDAGE
tr|A0A3S9V8L9|(C. tulipa) MTTSSYFLLVALGLLLLYVQSSFGNQHTRNSDTP---WNRCGSQITDSYR-ELCPH--KRNDAGK
sp|A0A0B5ADU4|Con-Ins T1 (C. tuli... MTTSSYFLLMALGLLLLYVQSSFGNQHTRNSDTP---KYRCGSEIPNSYI-DLCFR--KRNDAGK
sp|A0A0B5ABD5|Con-Ins T3 (C. tuli... MTTSSYFLLMALGLLLLYVQSSFGNQHTRNSDTP---KYRCGSDIPNSYM-DLCFR--KRNDAGK
sp|A0A0B5AC90|Con-Ins T2 (C. tulipa) MTTSSYFLLMALGLLLLYVQSSFGNQHTRNSDTP---KYRCGSDIPNSYM-DLCFR--KRNDAGK
Con-Ins P1 (C. purpurascens) MTTSSYFLLVALGLLLLYVQSSFAGH-----CGSDLVERYKT--CPGVSGTNDAGK
tr|A0A0K8TU45|(C. lenavati) MTTSSYFLLVALGLLLLHVQSSFGSQHTRG-----IKCGPELTE-YLLTLC LG--KRNDAGK

```

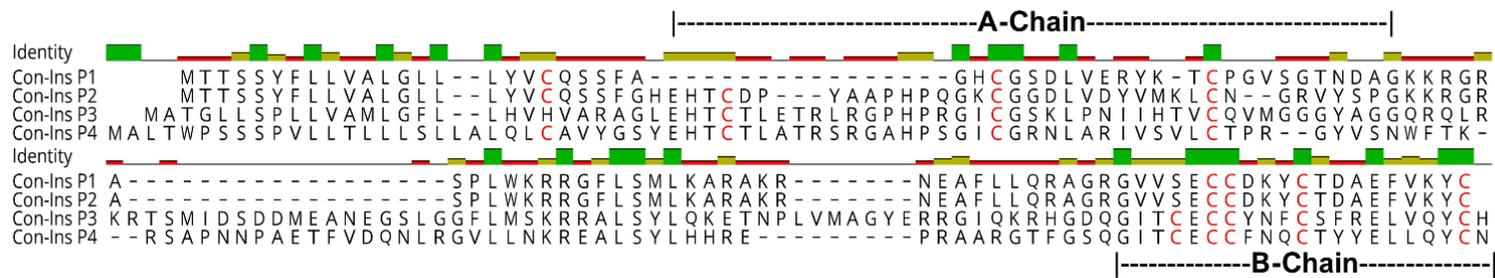
Identity 

```

tr|A0A3S9V8L7|Con-Ins K1 (C. kinos... -----KRRGFP SMLKARAKRNEAFL LQRDGRG I VEDCCYNDCTDEKLEKEYCHTLQG
tr|A0A3S9V8K6|Con-Ins K2 (C. kinos... -----KRRGFP SMLKARAKRNEAFL LQRDGRV I VGDCCDNYCTDERLKGYCASLLGL
sp|A0A0B5AC86|Con-Ins G3b (C. ge... RGRASPLWQRQGF L SMLKA--KRNEAFL LQRDGRG I VEVCCDNPCTVATLMTFCH
tr|X5IWU3|(C. geographus) RGRASPLWQRQGF L SMLKA--KRNEAFL LHRDGRG I VEVCCDNPCTVATLRTFCH
sp|A0A0B5A8P4|Con-Ins G3 (C. geo... RGRASPLWQRQGF L SMLKA--KRNEAFL LQRDGRG I VEVCCDNPCTVATLRTFCH
sp|A0A0B5A7P2|Con-Ins G1c (C. ge... RGRASPLWQRGGS L SQLKARAKRNGAFHL PRDGRGVVEHCCHRP CSNAEFKKYCS
sp|A0A0B5A8Q2|Con-Ins G1b (C. ge... RGRASPLWQRGGS L SKL KARAKRNGAFHL PRDGRGVVEHCCHRP CSNAEFKKYCG
sp|A0A0B5AC95|Con-Ins G1a (C. ge... RGRASPLWQRGGS L SKL KARAKRNGAFHL PRDGRGVVEHCCHRP CSNAEFKKYCG
tr|A0A3S9V8L9|(C. tulipa) RGQASPLWQRGGS L SMLKARAKRNEAFL LQRAHRGVVEHCCKRAC SNAEFMQFCG
sp|A0A0B5ADU4|Con-Ins T1 (C. tuli... RGRASPLWQRGGS L SMLKARAKRNEAFL LQRAHRGVVEHCCHRP CSNAEFKKF CG
sp|A0A0B5ABD5|Con-Ins T3 (C. tuli... RGQASPLWQRGGS L SMLKARAKRNEAFL LQRAHRGVVEHCCKRAC SNAEFMQFCGNS
sp|A0A0B5AC90|Con-Ins T2 (C. tulipa) RGQASPLWQRGGS L SMLKARAKRNEAFL LQRAHRGVVEHC CYRPSNAEFKKF CG
Con-Ins P1 (C. purpurascens) RGRASPLWKRQGF L SMLKARAKRNEAFL LQRAGRGVVSECCDKYCTDAEFVKYC
tr|A0A0K8TU45|(C. lenavati) RGRASPLWKRQGF L SMLKAGAKRNEAFL LQRA R P G I VEECCHKTCTDDEARKYCS SLM

```

**Figure 14- Conoinsulins that exhibit a vertebrate-like insulin cysteine framework** Sequences are from InterPro insulin superfamily database (IPR036438, unreviewed entries).



**Figure 15- Conoinsulin sequences expressed in *C. purpurascens* tissue.**

#### 4.2.2 ILP expression in *C. purpurascens* tissues

Transcriptomes from six different *C. purpurascens* tissue types were analyzed for ILP expression: venom duct, venom bulb, proboscis, eye, liver, and foot. In total, 4 transcripts encoding ILPs were identified (Figure 15). The transcripts exhibit heterologous expression among the different tissue types (Table 9). Con-Ins P1 is the most ubiquitously expressed in all tissue types, aside from the liver or foot where no ILP expression was found. Con-Ins P1 has the highest expression in the venom duct (866 TPM) and the venom bulb (143 TPM). P1 expression is significantly lower in the eye (12 TPM) and proboscis (9 TPM). Con-Ins P2 was also expressed in the venom duct (89 TPM); however, this protein was not identified through MS/MS analysis of the milked venom. Con-Ins P3 was expressed in low quantities (1-3 TPM) in the venom duct, bulb, and eye, and Con-Ins P4 was expressed only in the eye (2 TPM).

Con-Ins P1 exhibits a cysteine framework more like vertebrate insulins than molluscan-type insulins. When comparing the A and B chain peptides between the four transcripts, Con-Ins P1 has one less cysteine residue in each, resulting in one less inter-peptide bond. Con-Ins P1 shares a cysteine framework with other venom insulins from fish-hunting cone snail species (Figure 14). Con-Ins P3 and P4 are homologous to other MIPs, exhibiting an 8-cysteine framework. Interestingly, Con-Ins P2, also expressed in the venom duct, exhibits a hybrid sequence. It has an identical A chain peptide to Con-Ins P1, but its B chain is more like molluscan-type insulins Con-Ins P3 and P4.

**Table 8- RNA expression of conoinsulins in *C. purpurascens* tissues.**

RNA expression (TPM)	Con-Ins P1	Con-Ins P2	Con-Ins P3	Con-Ins P4
Venom Duct	866.51	89.02	3.36	0
Venom Bulb	143.13	0	2.51	0
Eye	12.34	0	1.40	2.07
Proboscis	9.12	0	0	0
Liver	0	0	0	0
Foot	0	0	0	0

#### 4.2.3 Con-Ins P1, new ILP from injected venom

Con-Ins P1 was identified from the LC-MS/MS analysis of the injected venom from *C. purpurascens*. Of the 27 specimens examined, Con-Ins P1 was only identified in the venom of two snails (Figure 16). Two unique peptides were identified from the RNA transcript for Con-Ins P1. The A-chain contains 4 cysteines and the B-chain contains 2 cysteines, which form two disulfide bonds to connect the peptides and one intra-peptide bond on the A chain (Figure 17). Both peptides have multiple PTMs (Figure 17). The A-chain peptide contained either 1 or 2 carboxylated glutamates at E5, E15, or both. The B-chain was identified with or without hydroxylated proline (P15), deamidated asparagine (N21), and C-terminal amidation. C-terminal truncation occurred on the B chain resulting in peptides ending in either amidated A23 or G24.

When compared to other vertebrate-like conoinsulins that have demonstrated activity at the human insulin receptor (Con-Ins K1, Con-Ins G1, Con-Ins-T1), Con-Ins P1 shows the most sequence homology to Con-Ins K1, and very little homology to G1 and T1, aside from a conserved insulin cysteine residues (Figure 18). All four conoinsulins maintain a carboxylated glutamate (E) residue; however, this is located at E5 in Con-Ins P1 and E4 in the other three peptides. Con-Ins P1 has a second site of carboxylation at E15 and lacks the amidated C-terminal. Vertebrate insulins from human and zebrafish also contain 2 glutamate acid residues in the A-chain, but carboxylation is only found in cone snail venom ILPs.

The Con-Ins P1 B-chain peptide demonstrates even more sequence diversity between other *Conus* species and vertebrate homologs. It lacks an N-terminal tail but maintains an extended C-terminus that is lost in the other venom ILPs. The C-terminal extension contains an oxidated proline at P15, a deamidated asparagine at N21, and an amidated C-terminus. Interestingly, the oxidated residue at P15 occurs at the predicted site for insulin receptor binding, based on previous Con-Ins modeling studies [154]. Con-Ins P1

is unique from the other characterized conoinsulins in that it contains an extended C-terminus, similar to human and zebrafish insulin, however, the P1 B-chain tail lacks the functional aromatic triplet, FFY, res possible for receptor recognition and dimerization.

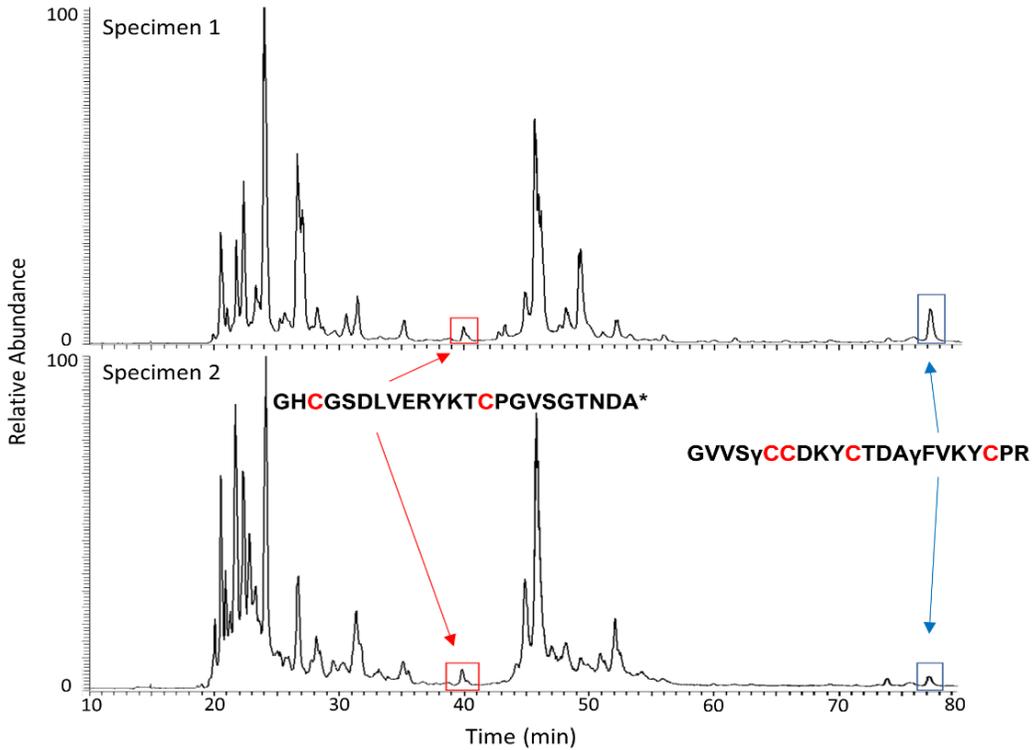


Figure 16- Con-Ins P1 from the injected venom of two *C. purpurascens* specimens.



Figure 17- Con-Ins P1 modifications identified by MS/MS. PTM symbols:  $\gamma$ = carboxylated glutamic acid, O= hydroxylated proline, N= deamidated asparagine, \* = amidated C-terminal.

	<b>A-chain</b>	<b>B-chain</b>
Human	GIVEQ <u>CC</u> TSI <u>C</u> SLYQLENY <u>C</u> N	FVNQHLCGSHLVEALYLVCGERGFFYTP <u>PK</u> T
Human lispro	GIVEQ <u>CC</u> TSI <u>C</u> SLYQLENY <u>C</u> N	FVNQHLCGSHLVEALYLVCGERGFFYTP <u>PK</u> T
Zebrafish	GIVEQ <u>CC</u> HKP <u>C</u> SIFELQNY <u>C</u> N	NPGTPQHLCGSHLVDALYLVCPTGFFY <u>NP</u>
Con-Ins P1	GVS <u>y</u> CCDKYCTDA <u>y</u> FVK <u>y</u> CP <u>R</u>	GH <u>C</u> GS <u>D</u> LVERYKT- <u>CO</u> GVSGT <u>NDA</u> * SDSGTTPDRD <u>HS</u> CG <u>Sy</u> LVTYL <u>GEL</u> CLGN
Con-Ins K1	GIV <u>y</u> D <u>CC</u> YNDCTDERLKEY <u>C</u> HTLQ*	TFDTOK <u>HR</u> CG <u>Sy</u> ITNSYMDL <u>CYR</u>
Con-Ins G1	GVV <u>y</u> H <u>CC</u> HR <u>PC</u> SNAEFKK <u>yC</u> *	NSDTOK <u>YR</u> CG <u>Sy</u> IPNSYIDL <u>CF</u>
Con-Ins T1	GVV <u>y</u> H <u>CC</u> HR <u>PC</u> SNAEFKK <u>yC</u> *	

**Figure 18- Con-Ins P1 displays cysteine framework homology to vertebrate insulins.** Sequence comparison between Con-Ins P1, human and zebrafish insulin, and other previously described conoinsulins. Residues involved in dimerization are underlined. Homolgy to human insulin shown in blue. PTM symbols: **y**= carboxylated glutamic acid, **O**= hydroxylated proline, **N**= deamidated asparagine, \* = amidated C-terminal.

### 4.3 Discussion

Con-Ins P1 from *C. purpurascens* is the first conoinsulin identified directly from injected venom. *C. purpurascens* hunts using a hook-and-pull strategy that allows for direct collection of the venom without dissecting the venom duct. Until now, proteomic evidence of conoinsulins in venom has been limited to Con-Ins G1 and Con-Ins G3 from the venom gland of *C. geographus*. In contrast to the hook-and-line strategy employed by *C. purpurascens*, *C. geographus* uses a net-hunting strategy in which the snail expel venom into the water through their expanded rostrum or 'net' to immobilize the fish, and then rapidly engulf their prey [112]. Because of the challenge of collecting venom through the net-capture strategy, the conoinsulins identified from *C. geographus* came from venom extruded from the dissected venom gland. Dissected venom is inherently more complex; it includes many protein-processing enzymes and extra components that do not ultimately end up in the true injected venom that is utilized by the snail against prey. Other vertebrate-like conoinsulin sequences from fish-hunting species *C. tulipa* and *C. kinoshitai* were

identified from RNAseq data and lacked proteomic evidence of their composition in the venom. For these reasons, the indication of 'weaponized' venom insulins remained unsubstantiated, despite evidence of their activity on human insulin receptors [154]. Our identification of Con-Ins P1 provides the first solid evidence that conoinsulin is actually injected into the prey, indicating that it plays a physiological role in prey capture.

The identification of conoinsulins at the proteomic level has been limited and most of them have been found at the transcript level within the venom ducts. When considering the unreviewed InterPro insulin superfamily database, there are 13 (of 67) conoinsulin transcripts that exhibit a cysteine framework resembling vertebrate insulin peptides rather than MIPs and other invertebrate insulins. These 13 vertebrate-like conoinsulins come from fish-hunting species, *C. geographus*, *C. kinoshitai*, and *C. tulipa*, and worm-hunting species *C. lenavati*. The presence of vertebrate-like insulin in the venom duct transcriptome of a worm-hunter contradicts the theory that these insulins have been 'weaponized' by piscivorous cone snails. If the theory of weaponized insulin holds, *C. lenavati* may be capable of a mixed hunting strategy, switching between fish and worms depending on what is readily available, but this is speculation. At this point, there is no proteomic evidence to confirm the presence of conoinsulin in the injected venom of worm-hunting snails. Proteomic analysis of the worm and mollusk-hunting cone snails is an area that is not well-studied and must be investigated to better understand the function of insulin in the venom/duct.

Another striking observation is the absence of vertebrate-like insulin transcripts in the venom ducts of other fish-hunting species (i.e. *C. bullatus*, *C. magus*, *C. striatus*, and *C. ermineus*). It may be that only some lineages of fish-hunters gained this alternative function of ILPs by incorporating it in their venom arsenal, or that others have lost this evolved trait. The absence of vertebrate-like conoinsulin in *C. ermineus* is especially interesting, being that it is a close relative to *C. purpurascens* and that the two species

exhibit high sequence homology and even overlap in their venom toxins (Grandal et al. 2020, in review). From the venom mRNA data currently available, we cannot conclude that ‘weaponized’ insulin is unique to fish-hunting species because it is present *C. levitani*, nor can we state that it is ubiquitous in fish-hunting cone snails. Furthermore, our current understanding of venom conoinsulins is based on transcriptomic data of a limited sample size – in most cases, one or two specimens per species. In this situation, genomic, transcriptomic, and proteomic data are all critical to deduce how and why these venom conoinsulins evolved.

Venom conoinsulins are likely the result of gene duplication and diversification. Such is the case with lactrodectins, neuropeptide hormone homologs recruited into the venom by various arthropod taxa [145, 146], and glucagon-like peptides from lizard venom [164]. Evidence for this in cone snails lies in the heterogeneous expression of different ILPs in different tissue types. This has been shown previously in the net-hunting species *C. geographus* [152], and here in the hook-and-pull hunter *C. purpurascens*. In both cases, venom conoinsulins have higher expression rates than endogenous MIPs, but this high expression was limited to venom duct and venom bulb tissue. Con-Ins P1 (venom insulin) was also expressed in the proboscis and eye tissue where its function remains in question. The other two conoinsulins identified from *C. purpurascens* (Con-Ins P3 and P4) more closely resemble MIPs rather than vertebrate insulins. We found that Con-Ins P3 and P4 had much lower expression than Con-Ins P1, but the expression patterns between the two differed. Con-Ins P3 was expressed in the venom duct, bulb, and eye, while Con-Ins P4 expression was limited to the eye. Con-Ins P2 resembles a hybrid ILP, its signal peptide and A-chain sequence show 100% homology to Con-Ins P1, while its B-chain is more similar to Con-Ins P3, P4, and other MIPs. Con-Ins P2 may be key to our understanding of conoinsulin diversification and the evolution of venom insulins from endogenous MIPs. *Lymnaea stagnalis*’ MIPs are produced and secreted by neuroendocrine glands and are

important for growth, metabolism, and other processes of neurodevelopment [137, 165]; however, their role in *Conus* venom is unclear. They may be part of endogenous signaling mechanisms, or cone snails may utilize these molluscan-type insulins in their venom to interrupt the cellular processes of their molluscan prey. Although we don't know why some species express these vertebrate-like venom insulin peptides and others don't, the new evidence of conoinsulins in injected venom supports the model of insulin weaponization through gene diversification.

Con-Ins P1 is unique from all previously identified conoinsulins in that we have identified differential modes of PTMs determined by MS/MS spectral matching. Previously, the synthetic modified version of Con-Ins G1 showed more activity at the human insulin receptor than the PTM-free peptide [153]. Since Con-Ins P1 contains different modifications at different positions than G1, it will be important to assess the effect of its different modified forms for their ability to bind and activate the insulin receptor. It will be especially important to assess the effect of hydroxylated P15 residue on the C-terminus of Con-Ins P1 because it is located in the predicted functional site for Con-Ins activity based on previous molecular docking studies [154].

Con-Ins P1 is also unique from other vertebrate-like conoinsulins in that it lacks the N-terminal tail on the B-chain but maintains an elongated C-terminus, also found on human and zebrafish insulins. The C-terminus of the B-chain in human insulin includes the aromatic triplet, known to be critical to the binding mechanism [166]. A recently developed minimized insulin analog, based on the framework of Con-Ins G1, mini-Ins, maintains binding activity despite lacking the C-terminus region of the B-chain [155]. This, along with the proven activity of conoinsulins, is evidence that the aromatic triplet is not essential for conoinsulin function. However, since the B-chain C-terminal has significant structural implications for binding, Con-Ins P1 may provide insight into new mechanisms for ligand-receptor interaction

Molecular docking studies for vertebrate-like conoinsulins reported six conserved residues within the bioactive conoinsulins: Glu4, Lys/Arg9, and Ser12 in the A-chain and Ser9, Glu/Asp10, and Glu/Asp17 in the B-chain [154]. Con-Ins P1 only has one of these six conserved residues, emphasizing its uniqueness among the vertebrate-like venom insulins. The one conserved residue is aspartate at site B10, consistent with both Con-Ins T2 and T3, while the others have glutamate, which is predicted to be carboxylated. Human insulin has a histidine at site B10 that is thought to be involved in receptor recognition and activation. Amino acid substitution studies have shown that replacing this histidine with an aspartate increases the peptide's activity by 4-5 fold, presumably due to the negatively charged residue [167]. This mutated peptide was developed into a rapid-acting insulin analog but it was discontinued due to mitogenicity [168, 169]. All vertebrate-like venom insulins identified thus far contain a negatively charged residue (i.e. histidine) at position B10. Con-Ins P1, however, also maintains the following three residues (LeuB11, ValB12, and GluB13) that are identical to human insulin. For these reasons Con-Ins P1 makes an interesting case study for molecular dynamics, using Con-Ins P1 as a model.

The discovery of vertebrate-like conoinsulins from cone snail venom has opened the door to the development of novel insulin analogs, like mini-Ins. The uniqueness of Con-Ins P1 compounded by its presence in injected venom makes this peptide a prime candidate for drug development. The first step is to use model the molecular dynamics using advanced cryo-EM structures of the insulin receptor [170-172]. We can use this information to deduce critical residues for binding and assess the effects of individual PTMs, as well as single amino acid mutations. These molecular modeling studies can advance our current understanding of human insulin-receptor interaction and provide grounds for developing novel, high-efficiency insulin analogs as complementary therapies for hyperglycemic metabolic conditions.

CHAPTER 5:

FUNCTIONAL SCREENING OF NOVEL ALPHA-CONOTOXIN ACTIVITY ON  
NICOTINIC ACETYLCHOLINE RECEPTOR SUBTYPES

## 5.1 Introduction

Alpha-conotoxins are inhibitors of the nAChR and are the most ubiquitous venom components across the *Conus* genus; all species analyzed express one or more  $\alpha$ -conotoxins in their venom [100]. Unique from other nAChR ligands,  $\alpha$ -conotoxins exhibit subtype selectivity, specifically among neuronal receptor subtypes [52, 55]. Selective ligands of nAChR subtypes are clinically important probes for studying the pathophysiology behind diseases like addiction, cognitive disorders, neurodegenerative diseases, and pain [48, 173, 174]. Subtype specificity makes  $\alpha$ -conotoxins valuable molecular probes for basic neuroscience research and drug design.

The framework I  $\alpha$ -conotoxins belong to the A Superfamily, as defined by a conserved gene signal sequence. This is the largest group of characterized  $\alpha$ -conotoxins with the greatest diversity in subtype selectivity. The post-translational modifications commonly found on these conotoxins (C-terminal amides and hydroxyprolines) are important for peptide stability and bioactivity [51]. Within the framework I  $\alpha$ -conotoxins, there is significant diversity in amino acid composition and the size of the inter-cysteine loops. The size of the  $\alpha$ -conotoxin inter-cysteine loops has been shown to affect affinity toward either muscular or neuronal nAChR subtypes (Figure 3) [53, 54], and can therefore be used to predict the target of uncharacterized toxins based on homology. In general,  $\alpha$ -conotoxins exhibiting a 3/5 inter-cysteine loop pattern are inhibitors of neuromuscular junction subtypes, and 4/3 and 4/7 toxins inhibit neuronal subtypes [55].

Neuronal subtypes are expressed heterogeneously throughout the central nervous system [56]. The most widely expressed subtype in the mammalian brain,  $\alpha 4\beta 2$ , is heavily involved in the dopaminergic pathway and nicotine addiction. This receptor can exhibit different subtype stoichiometry;  $(\alpha 4)_2(\beta 2)_3$  is more sensitive to nicotine than  $(\alpha 4)_3(\beta 2)_2$  [57]. Because of its role in nicotine addiction,  $\alpha 4\beta 2$  is the clinical target for smoking cessation therapies, including the partial agonist, varenicline (Chantix, Pfizer, Inc.). Very few  $\alpha$ -

conotoxins inhibit  $\alpha 4\beta 2$  but not exclusively,  $\alpha$ -GID being the most potent inhibitor [58]. The  $\alpha 3\beta 4$  subtype is the predominant nAChR in the sensory and autonomic ganglia neurons and is expressed in the mesolimbic dopamine circuitry of the midbrain where it modulates addiction to nicotine and potentially other drugs of abuse [61, 62]. Only  $\alpha 4/6$ - conotoxins  $\alpha$ -AulB and  $\alpha$ -TxID have demonstrated specificity for this subtype [63, 175].

The  $\alpha 6$ -containing subtypes are also expressed abundantly in the midbrain dopaminergic neurons and are mediators of the addiction reward pathway [65, 66, 176, 177]. Ligands with selectivity for  $\alpha 6$ -subtype nAChRs are important molecular probes to study the pathophysiology of addiction and other dopamine-related disorders, such as Parkinson's disease. However,  $\alpha 6$  selectivity is rare because of its high homology with the  $\alpha 3$  subunit. There is one  $\alpha 6$ -biased ligand known thus far from *C. purpurascens*,  $\alpha$ -PIA, which preferentially inhibits  $\alpha 6$ -containing receptors with 75-fold greater affinity than  $\alpha 3$  receptors [67, 68].

Homomeric  $\alpha 7$  nAChRs are a unique subtype expressed throughout the brain, including the hippocampus and cerebral cortex involved in learning and memory [71], as well as in non-neuronal tissues, like immune cells [70]. Along with  $\alpha 6$  receptors,  $\alpha 7$  receptors are involved in nicotine reward pathways and present molecular targets for smoking cessation therapeutics [178-180]. The  $\alpha 7$  subtype is also involved in pain and inflammation pathways [174, 181]. Ric-3 is an important assembly protein found to increase the assembly and expression of  $\alpha 7$  receptors [182] and is therefore co-expressed with the  $\alpha 7$  subunit in this study for functional assays. Several structure-activity studies with  $\alpha 4/7$ -conotoxins have identified critical residues for  $\alpha 7$  activity [183-185].

The ligand-binding properties of nAChRs have been studied using the X-ray crystal structure of the soluble acetylcholine binding protein (AChBP) of mollusks *Lymnaea stagnalis* and *Aplysia californica* [43, 186]. AChBPs are not functional ion channels, however, they form stable homopentamers that preserve features of the ligand-binding

domain of nAChRs, and thereby serve as useful binding models for  $\alpha$ -conotoxins [186, 187]. The AChBP is particularly useful for the homology modeling of the homopentameric  $\alpha 7$  subtypes [160, 188].

There are several previously characterized  $\alpha$ -conotoxins from the venom of *C. purpurascens*. Framework IV toxin,  $\alpha$ -PIVA, is a paralytic nAChR antagonist selective for muscle-subtype receptors. It is present in the venom in differentially modified forms, including multiple sites of proline hydroxylation, which affects bioactivity [189]. The characterized framework I toxins from *C. purpurascens* include  $\alpha$ -PIA,  $\alpha$ -PIB, and  $\alpha$ -PIC. Both  $\alpha$ -PIB and  $\alpha$ -PIC target muscle receptor subtypes [160, 190], while  $\alpha$ -PIA is selective for  $\alpha 6$  containing neuronal receptors [191].

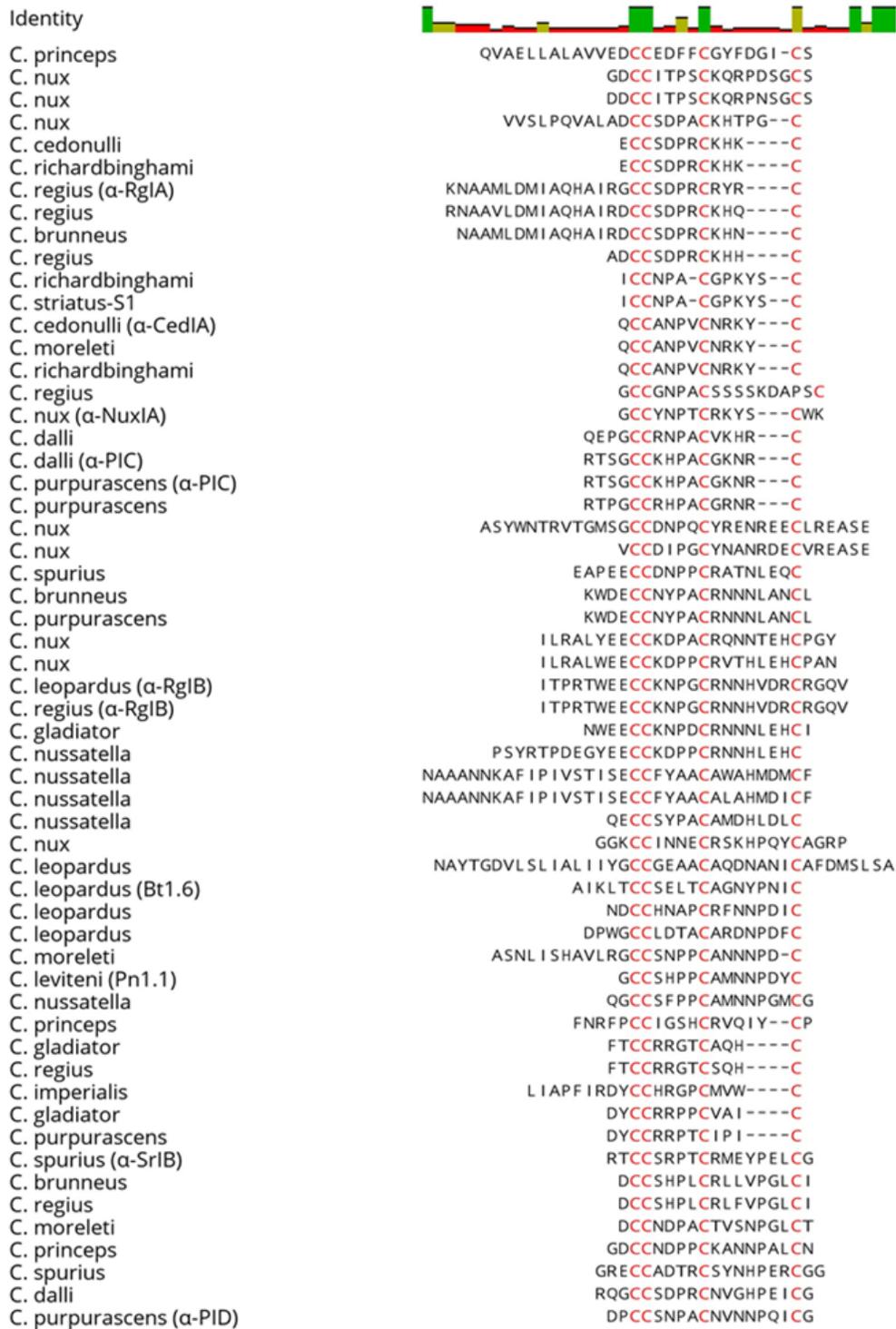
Here, we will focus specifically on the framework I  $\alpha$ -conotoxins because they 1) have a conserved two-disulfide bonding pattern (1-3, 2-4), and 2) are relatively small peptides (<22 residues). These characteristics simplify solid-state peptide synthesis for functional assays. First, we considered  $\alpha$ -conotoxins identified from the injected venom of *C. purpurascens*. These included  $\alpha 6$ -selective conotoxin,  $\alpha$ -PIA, and newly identified  $\alpha$ -PID (Grandal et al. 2020, in review). Both peptides exhibit a 4/7 inter-cysteine loop size and significant sequence homology. We also evaluated venom duct transcriptomic data from other *Conus* species for expression of  $\alpha$ -conotoxins. Two  $\alpha 4/4$ - conotoxins,  $\alpha$ -NuxI from *Conus nux* and  $\alpha$ -CedI from *Conus cedonulli*, were also synthesized for functional screening. Here, we propose to test the inhibitory activity of three novel  $\alpha$ -conotoxins,  $\alpha$ -PID,  $\alpha$ -NuxIA, and  $\alpha$ -CedIA, as well as previously described  $\alpha$ -PIA. We predict they will be selective for neuronal subtypes over neuromuscular nAChRs based on their inter-cysteine loop size.

## 5.2 Results:

### 5.2.1 Bioinformatic approach to identifying novel $\alpha$ -conotoxins

A conserved A superfamily signal sequence was used to identify putative  $\alpha$ -conotoxin transcripts from the transcriptomes of 17 *Conus* species. The BLAST search returned a total of 57 transcripts, from which we predicted the mature  $\alpha$ -conotoxins sequences (Figure 19). The search included some previously characterized  $\alpha$ -conotoxins, including  $\alpha$ -SI (*C. striatus*),  $\alpha$ -SrlB (*C. spurius*),  $\alpha$ -RglA and  $\alpha$ -RglB (*C. regius*), and  $\alpha$ -PIC (*C. purpurascens*).

All the sequences share framework I cysteine pattern, CC-(X)<sub>m</sub>-C-(X)<sub>n</sub>-C. In these sequences, the first inter-cysteine loop (m) contained either 3 or 4 residues, while the second loop (n) contained either 3, 4, 5, 6, 7, or 9 residues, resulting in the following loop patterns: 3/5, 4/3, 4/4, 4/5, 4/6, 4/7, or 4/9. There is high sequence homology among  $\alpha$ -conotoxins from different species; in many cases only 1 or 2 residues are different. Six sequences are expressed in multiple species, including  $\alpha$ -PIC from *C. purpurascens*, which was also found in *C. dalli*.



**Figure 19- Sequences of α-conotoxins extracted from Conus venom duct RNA-seq transcripts.**

### 5.2.2 Functional screening of novel $\alpha$ -conotoxins on nAChR subtypes

We chose three  $\alpha$ -conotoxins from the 57 identified sequences to have synthesized for functional assays:  $\alpha$ -PID from *C. purpurascens*,  $\alpha$ -NuxIA from *C. nux*, and  $\alpha$ -CedIA from *C. cedonulli*. We also obtained previously characterized  $\alpha$ -PIA from a commercial source as a positive control for screening assays. Both  $\alpha$ -CedIA and  $\alpha$ -NuxIA exhibit a 4/4/ loop pattern, while  $\alpha$ -PIA and  $\alpha$ -PID exhibit a 4/7 loop pattern (Figure 20).  $\alpha$ -PIA and  $\alpha$ -PID have highly homologous sequences, with a one residue difference in the first loop and a two residue difference in the second loop (82% similarity). Each contains 18 residues, however,  $\alpha$ -PID has a one residue N-terminal deletion and C-terminal elongation compared to  $\alpha$ -PIA.  $\alpha$ -NuxIA and  $\alpha$ -CedIA contain 15 and 13 residues, respectively, and very little sequence homology aside from asparagine and proline residues in the first loop that are consistent across all 4 sequences (46% similarity).

Functional screenings were performed on nAChR subtypes expressed in *Xenopus* oocyte vectors. Inhibition of ACh-induced current was measured for the following nAChR subtypes: human neuronal receptors  $\beta_3\alpha_4\beta_2\alpha_6\beta_2$ ,  $\alpha_7$ ,  $\alpha_3\beta_4$ ,  $\alpha_4(2)\beta_2(3)$ ,  $\alpha_4(3)\beta_2(2)$ ,  $\alpha_4\alpha_5\beta_2$ , and mouse muscle receptor  $\alpha_1\beta_1\delta\epsilon$  (Table 4). Screening assays measured  $\alpha$ -CNTX (1  $\mu$ M) inhibition of ACh-induced current (Figure 21).  $\alpha$ -CNTX -elicited responses were normalized to ACh control currents so that inhibition values represent % control.

All four  $\alpha$ -conotoxins inhibit the muscle subtype  $\alpha_1\beta_1\delta\epsilon$  (Table 10); the most potent inhibitor at 1  $\mu$ M was  $\alpha$ -NuxIA ( $0.15 \pm 0.01$ ,  $p= 0.008$ ) with IC<sub>50</sub> of 47.4 nM (Figure 22). The other three elicited ~50% inhibition at 1  $\mu$ M (Figure 21).

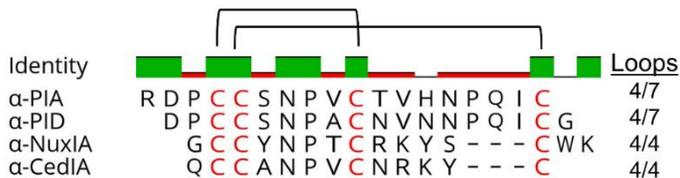
$\alpha$ -PID significantly inhibits neuronal subtypes  $\alpha_3\beta_4$  ( $0.73 \pm 0.03$ ,  $p= 0.001$ ) and  $\alpha_4(2)\beta_2(3)$  ( $0.84 \pm 0.02$ ,  $p= 0.006$ ).  $\alpha$ -CedIA also significantly inhibits  $\alpha_4(2)\beta_2(3)$  ( $0.90 \pm 0.02$ ,  $p= 0.005$ ), while  $\alpha$ -NuxIA significantly inhibits  $\alpha_4(3)\beta_2(4)$  with alternative stoichiometry ( $0.78 \pm 0.04$ ,  $p= 0.002$ ). No significant inhibition was measured on receptor  $\alpha_4\alpha_5\beta_2$ .

Both  $\alpha$ -PID ( $0.14 \pm 0.02$ ,  $p= 0.0003$ ) and  $\alpha$ -PIA ( $0.06 \pm 0.03$ ,  $p= 0.055$ ) strongly inhibited the  $\beta_3\alpha_4\beta_2\alpha_6\beta_2$  receptor concatemer. These values corresponded to an 86% inhibition of ACh-induced current by  $\alpha$ -PID and 94% inhibition by  $\alpha$ -PIA. An inhibitory dose-response curve (30 nM, 100 nM, 300 nM, and 1  $\mu$ M) calculated an IC<sub>50</sub> of 119.7 nM for  $\alpha$ -PID and 329.9 nM for  $\alpha$ -PIA (Figure 23A).  $\alpha$ -PIA exhibited a greater response at higher concentrations (1  $\mu$ M and 300 nM), while  $\alpha$ -PID exhibited a greater response at lower concentrations (100 nM and 30 nM) (Table 11). The  $\beta_3\alpha_4\beta_2\alpha_6\beta_2$  receptor was also significantly inhibited by 1  $\mu$ M  $\alpha$ -CedIA ( $0.74 \pm 0.02$ ,  $p= 0.031$ ), although to a lesser extent than  $\alpha$ -PIA and  $\alpha$ -PID.  $\alpha$ -NuxIA also demonstrated minor inhibition on  $\beta_3\alpha_4\beta_2\alpha_6\beta_2$ , although not significant ( $0.78 \pm 0.07$ ,  $p= 0.294$ ).

Alpha-PID also significantly inhibited  $\alpha_7$  receptors at 1  $\mu$ M ( $0.35 \pm 0.05$ ,  $p= 0.057$ ) and 300 nM ( $0.54 \pm 0.05$ ,  $p= 0.029$ ) (Table 11, Figure 23B).

**Table 9- Inhibitory activity of Framework I  $\alpha$ -conotoxins from *C. purpurascens*.**

$\alpha$ -CNTX	Sequence	Reference	nAChR selectivity
$\alpha$ -PIA	-RDPCCSNPVCTVHNPQIC	This work	$h\beta_3\alpha_4\beta_2\alpha_6\beta_2^* \gg ma1\beta_1\delta\epsilon \gg ha_{4(3)}\beta_{2(2)} > ha_{4(2)}\beta_{2(3)} > ha_{4\alpha}5\beta_2 \approx ha3\beta_4 > ha_7$
		Dowell (2002)	$ra6/\alpha_3\beta_2\beta_3 > ra6/\alpha_3\beta_4 > ra_3\beta_2 > ra_3\beta_4 \gg ra_4\beta_2$
$\alpha$ -PIB	QSPGCCWNPACVKNR--C	Lopez-Vera (2007)	$ha6/\alpha_3\beta_2\beta_3 > ha6/\alpha_3\beta_4 \gg ha1\beta_1\delta\epsilon$ $ma1\beta_1\delta\epsilon \approx ma1\beta_1\delta\gamma$
$\alpha$ -PIC	--SGCCKHPACGKNR--C	Hoggard (2017)	$ra1\beta_1\delta\epsilon > ra1\beta_1\delta\gamma \approx ha_3\beta_2 \gg ha_7$
$\alpha$ -PID	--DPCCSNPACNVNPNQICG	this work	$h\beta_3\alpha_4\beta_2\alpha_6\beta_2^* > ha_7 > ma1\beta_1\delta\epsilon > ha_3\beta_4 > ha_{4(2)}\beta_{2(3)} > ha_{4(3)}\beta_{2(2)} \approx ha_{4\alpha}5\beta_2$



**Figure 20- Sequences of the  $\alpha$ -conotoxins tested for inhibitory activity.**

### Inhibition of ACh-induced current

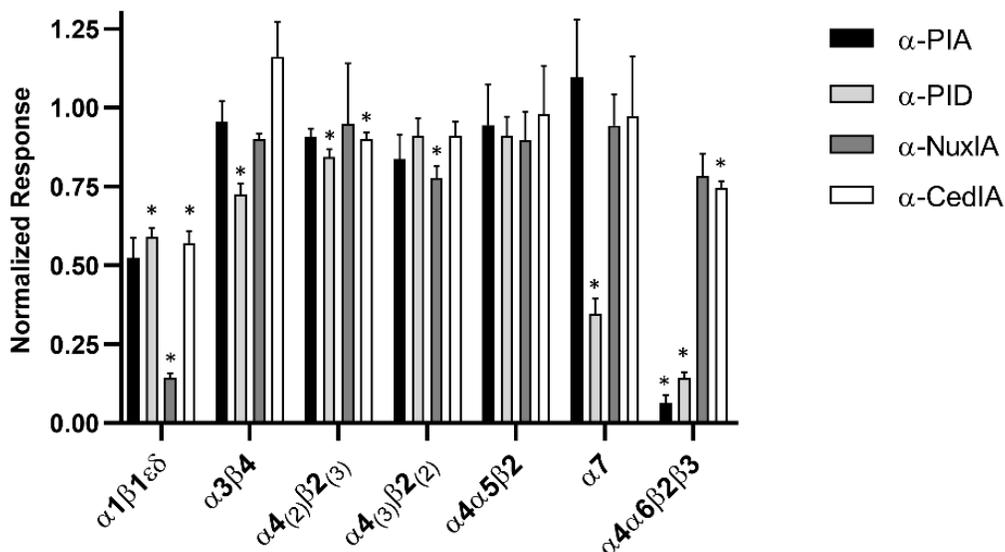
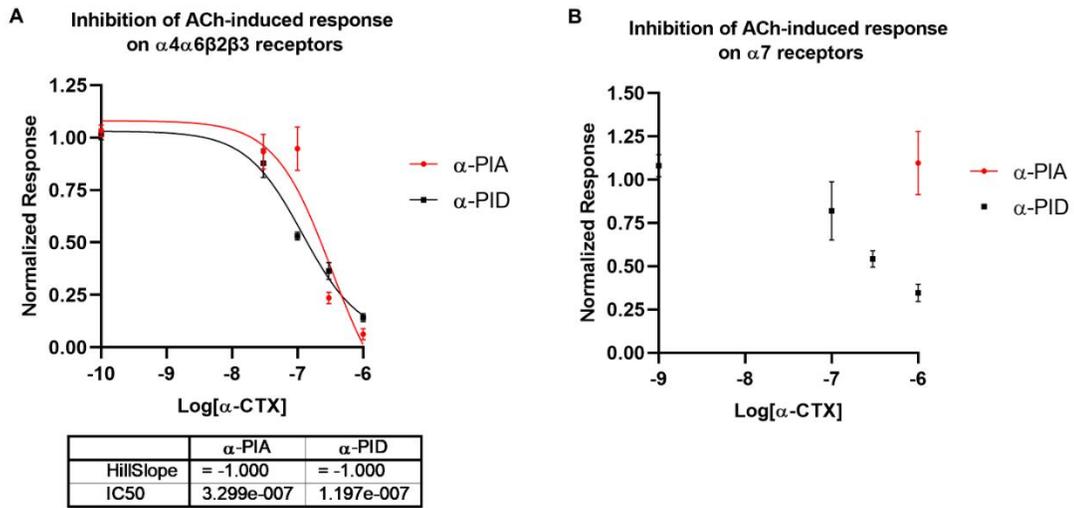


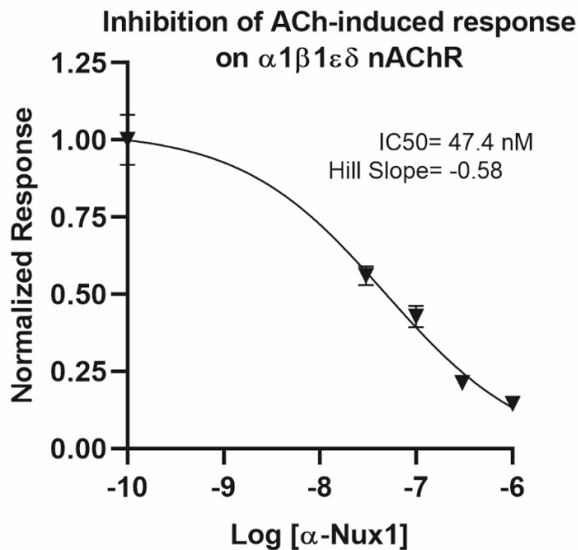
Figure 21- Inhibitory activity of  $\alpha$ -PIA,  $\alpha$ -PID,  $\alpha$ -NuxIA, and  $\alpha$ -CedIA on nAChR subtypes. \*P-value  $\leq$  0.05

Table 10- Inhibitory activity of  $\alpha$ -PIA,  $\alpha$ -PID,  $\alpha$ -NuxIA, and  $\alpha$ -CedIA on nAChR subtypes. Values represent % ACh- elicited response  $\pm$ SEM (n). p-values calculated by a paired t-test.

	$\alpha$ -PID	$\alpha$ -PIA	$\alpha$ -NuxIA	$\alpha$ -CedIA
$\alpha 1\beta 1\delta\epsilon$	0.59 $\pm$ 0.03 (8) p= 0.025	0.52 $\pm$ 0.06 (5) p= 0.093	0.15 $\pm$ 0.01 (4) p= 0.008	0.57 $\pm$ 0.04 (4) p= 0.046
$\alpha 3\beta 4$	0.73 $\pm$ 0.03 (4) p= 0.001	0.96 $\pm$ 0.07 (3) p= 0.266	0.90 $\pm$ 0.02 (3) p= 0.066	1.16 $\pm$ 0.11 (3) p= 0.145
$\alpha 4_{(2)}\beta 2_{(3)}$	0.84 $\pm$ 0.02 (8) p= 0.006	0.91 $\pm$ 0.02 (4) p= 0.802	0.95 $\pm$ 0.19 (8) p= 0.298	0.90 $\pm$ 0.02 (8) p= 0.005
$\alpha 4_{(3)}\beta 2_{(2)}$	0.91 $\pm$ 0.06 (4) p= 0.129	0.84 $\pm$ 0.08 (4) p= 0.120	0.78 $\pm$ 0.04 (4) p= 0.002*	0.91 $\pm$ 0.04 (4) p= 0.137
$\alpha 4\alpha 5\beta 2$	0.91 $\pm$ 0.06 (4) p= 0.112	0.94 $\pm$ 0.13 (3) p= 0.316	0.90 $\pm$ 0.09 (4) p= 0.201	0.98 $\pm$ 0.15 (4) p= 0.234
$\alpha 7$	0.35 $\pm$ 0.05 (4) p= 0.057	1.09 $\pm$ 0.18 (4) p= 0.380	0.94 $\pm$ 0.10 (4) p= 0.257	0.97 $\pm$ 0.19 (4) p= 0.336
$\alpha 4\alpha 6\beta 2\beta 3$	0.14 $\pm$ 0.02 (19) p= 0.0003	0.06 $\pm$ 0.03 (4) p= 0.055	0.78 $\pm$ 0.07 (3) p= 0.294	0.74 $\pm$ 0.02 (4) p= 0.031



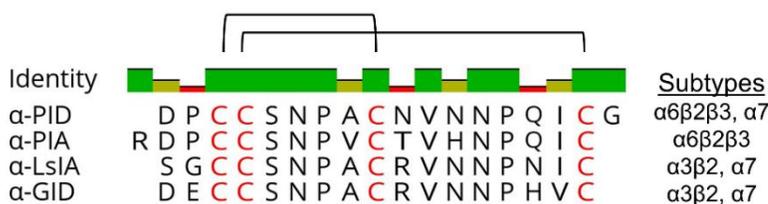
**Figure 23-** Dose response curve for  $\alpha$ -PIA and  $\alpha$ -PID on  $\alpha 4\alpha 6\beta 2\beta 3$  and  $\alpha 7$  nAChRs. Error bars represent  $\pm$  SEM



**Figure 22-** Dose response curve for  $\alpha$ -Nux1A on  $\alpha 1\beta 2\delta\epsilon$  nAChRs. Error bars represent  $\pm$  SEM

**Table 11- Dose-response of  $\alpha$ -PIA and  $\alpha$ -PIDc on  $\alpha 4\alpha 6\beta 2\beta 3$  and  $\alpha 7$  nAChRs.**  
 Values represent % ACh- elicited response  $\pm$ SEM. p-values calculated by a paired

[ $\alpha$ -CNTX]	$\alpha 4\alpha 6\beta 2\beta 3$		$\alpha 7$	
	$\alpha$ -PID	$\alpha$ -PIA	$\alpha$ -PID	$\alpha$ -PIA
1 $\mu$ M	0.14 $\pm$ 0.02 p= 0.0003	0.06 $\pm$ 0.03 p= 0.055	0.35 $\pm$ 0.05 p= 0.057	1.09 $\pm$ 0.18 p= 0.380
300 nM	0.36 $\pm$ 0.04 p= 0.002	0.24 $\pm$ 0.03 p= 0.006	0.54 $\pm$ 0.05 p= 0.029	0.97 $\pm$ 0.30 p= 0.491
100 nM	0.53 $\pm$ 0.02 p= 0.025	0.95 $\pm$ 0.10 p= 0.439	0.82 $\pm$ 0.17 p= 0.215	1.25 $\pm$ 0.34 p= 0.317
30 nM	0.88 $\pm$ 0.07 p= 0.166	0.93 $\pm$ 0.08 p= 0.199	----	----

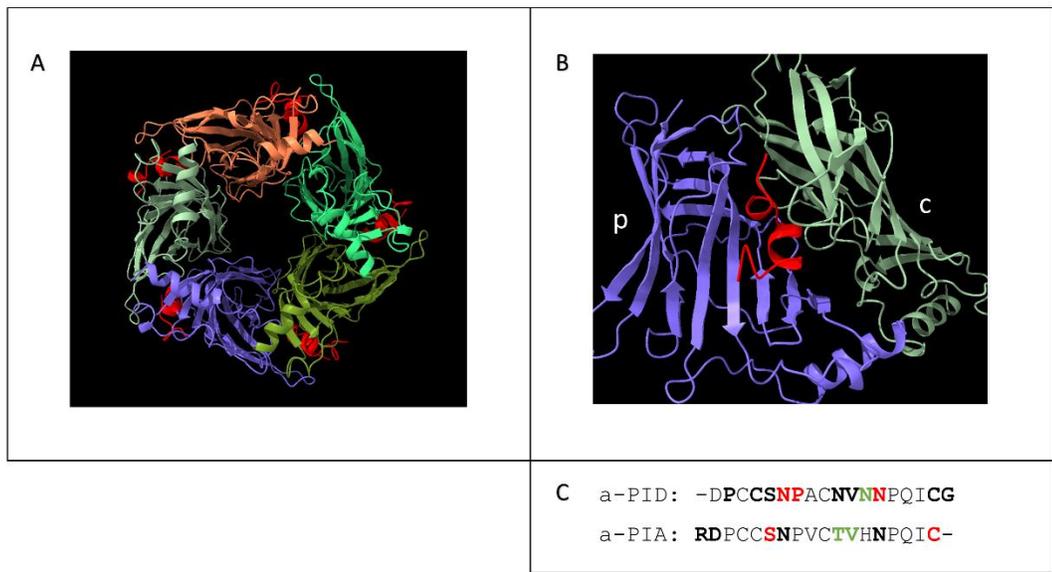


**Figure 24- Alignment of  $\alpha$ -PID and  $\alpha$ -PIA with conotoxins that inhibit  $\alpha 7$  receptors.**

### 5.2.3 Modeling $\alpha$ -PID and $\alpha$ -PIA binding to the $\alpha 7$ nAChR

Homology modeling was used to model the extracellular domain of the human  $\alpha 7$  receptor based on the structure of *Aplysia* AChBP bound to  $\alpha$ -LvIA (PDB: 5XGL) [192]. By replacing  $\alpha$ -LvIA with  $\alpha$ -PIA and  $\alpha$ -PID, we were able to model and compare their binding dynamics to the  $\alpha 7$  homopentamer. The models contain five  $\alpha$ -CNTX molecules bound between each adjacent subunit (Figure 25A). Each of the five bound  $\alpha$ -CNTX was analyzed for contacts, in the form of hydrogen bonds, with the principal (p) and the complementary (c) receptor subunits (Figure 25B). All possible sites of contacts combined from the five bound  $\alpha$ -CNTXs are summarized in Figure 25C.  $\alpha$ -PID had seven residues in contact with the c subunit (P2, C4, S5, N10, V11, C17, and G18), three residues in contact with the p subunit (N6, P7, N13), and one residue with contact to both (N12).  $\alpha$ -

PIA had four residues in contact with the c subunit (R1, D2, N7, N14), two residues in contact with the p subunit (S6, C18), and two residues in contact with both (T11, V12). Overall,  $\alpha$ -PID showed more possible interactions (10 residues) compared to  $\alpha$ -PIA (8 residues).



**Figure 25- Model of  $\alpha$ -PID bound to human  $\alpha$ 7 receptor.** A)  $\alpha$ -PID (red) bound at all five binding sites and B)  $\alpha$ -PID (red) in the binding pocket between two adjacent  $\alpha$ 7 subunits, the principal (p) subunit (purple) and the complementary (c) subunit (green). This model used the structure of  $\alpha$ -LvIA bound to the AChBP (PDB: 5XGL) as a template. C)  $\alpha$ -PID and  $\alpha$ -PIA residue contacts with the  $\alpha$ 7 receptor determined by molecular energy and distance calculations. Contacts are color coded by the following: **black**= c subunit, **red**= p subunit, **green**= both c and p subunit contacts.

### 5.3 Discussion

Alpha-conotoxins are venom peptides that selectively bind and inhibit nAChRs. Their role in cone snail venom is analogous to that of  $\alpha$ -bungarotoxins in snake venom, to block neurotransmission at the postsynaptic membrane in the neuromuscular junction, thereby immobilizing prey. What sets  $\alpha$ -conotoxins apart from other nAChR ligands is their unique affinity for neuronal nAChR subtypes, which are not involved in muscle movement, but rather are highly involved in neurotransmitter signaling processes in the brain. It is not known whether cone snails have evolved to target receptors homologous to vertebrate

neuronal receptors in their invertebrate prey [193], or if the neuronal subtype specificity is an anomaly that happens to be to our benefit, as these neuronal subtypes contribute to the pathophysiology of many devastating human diseases [181].

Neuronal nAChR subtypes include all subunit arrangements other than the muscle subtype,  $\alpha 1\beta 1(\delta/\gamma/\epsilon)$ . In this study we screened a subset of human neuronal receptor subtypes:  $\beta 3\alpha 4\beta 2\alpha 6\beta 2$ ,  $\alpha 7$ ,  $\alpha 3\beta 4$ ,  $\alpha 4(2)\beta 2(3)$ ,  $\alpha 4(3)\beta 2(2)$ , and  $\alpha 4\alpha 5\beta 2$ . Alpha-conotoxin selectivity has been broadly attributed to the inter-cysteine loop sizes of these small disulfide-constrained peptides. Muscle-subtype selectivity is commonly found in  $\alpha 3/5$  conotoxins (3 and 5 residues in the first and second inter-cysteine loops), while  $\alpha 4/3$ ,  $\alpha 4/4$ , and  $\alpha 4/7$  conotoxins tend to inhibit neuronal subtypes. Here, we screened two new  $\alpha 4/4$  conotoxins ( $\alpha$ -NuxIA and  $\alpha$ -CedIA) identified from RNA sequencing, and two  $\alpha 4/7$  conotoxins ( $\alpha$ -PIA and  $\alpha$ -PID) identified in the injected venom of *C. purpurascens* (Grandal et al. 2020, in review).  $\alpha$ -PIA has been previously described as an  $\alpha 6$ -selective conotoxin with the unique ability to distinguish between the similar  $\alpha 3$  and  $\alpha 6$  subunit, (Table 9) [67]. We aimed to validate these previous findings and expand the functional screening to include  $\alpha 7$  receptors. We predicted that all four  $\alpha$ -conotoxins would show preferential inhibition toward neuronal subtypes over neuromuscular receptors, based on their inter-cysteine loop sizes.

The results from the screening assays showed that all four  $\alpha$ -conotoxins did not exhibit selectivity for neuronal subtypes. The first notable finding from this body of work is that all four  $\alpha$ -conotoxins showed inhibition of the muscle subtype,  $m\alpha 1\beta 1\delta\epsilon$ , by at least 40%. It is well supported that  $\alpha 3/5$  conotoxins have a high affinity for the muscle subtype of the nicotinic receptor [50], but we did not expect to see  $\alpha 1\beta 1$  inhibition by  $\alpha 4/7$  conotoxins, based on previously characterized  $\alpha$ -conotoxins [53, 54]. This is also in stark contrast to previous functional assays with  $\alpha$ -PIA on  $h\alpha 1\beta 1\delta\epsilon$  receptors, where 10 $\mu$ M (10x our screening concentration)  $\alpha$ -PIA did not affect inhibition. This could be due to the

discrepancy in the species cDNA, as previous studies used human receptors [67], and here we used mouse muscle receptors. The strongest inhibition of  $\alpha 1\beta 1\delta \epsilon$  muscle receptors detected in this study was demonstrated by  $\alpha 4/4$  conotoxin,  $\alpha$ -NuxIA, (IC<sub>50</sub>= 47 nM). Existing data on  $\alpha 4/4$  conotoxin subtype specificity is limited; however, two  $\alpha 4/4$  conotoxins have been identified from *C. purpurascens* ( $\alpha$ -PIB and  $\alpha$ -PIC). Both conotoxins also showed high affinity for muscle  $\alpha 1\beta 1\delta \epsilon$  receptors over neuronal subtypes (Table 9) [160, 190]. Although  $\alpha$ -NuxIA shows little sequence homology to  $\alpha$ -PIB or  $\alpha$ -PIC, the combined evidence suggests that the  $\alpha 4/4$  framework conotoxins, like the  $\alpha 3/5$  conotoxins, are selective for neuromuscular nicotinic receptors.

The second strongest inhibition was on the  $\alpha 6$ -containing concatemer,  $h\beta 3\alpha 4\beta 2\alpha 6\beta 2$  by both  $\alpha$ -PIA (329 nM) and  $\alpha$ -PID (119 nM). At the highest concentration (1  $\mu$ M)  $\alpha$ -PIA shows greater inhibition than  $\alpha$ -PID (96% vs 84%), but at lower concentrations (100 nM and 30 nM)  $\alpha$ -PID has a stronger effect, resulting in a lower IC<sub>50</sub> value than  $\alpha$ -PIA. Previous studies report a much lower IC<sub>50</sub> (1.7 nM) for  $\alpha$ -PIA on  $h\alpha 6/\alpha 3\beta 2\beta 3$  receptors [67]. However, we cannot make a direct comparison to these studies because we used an  $\alpha 6$  concatemer in this screening study that contained  $\alpha 6$  and  $\alpha 4$  subunits ( $\beta 3\alpha 4\beta 2\alpha 6\beta 2$ ), while previous studies used an  $\alpha 6/\alpha 3$  chimeric receptor. Because  $\alpha 6$  and  $\alpha 3$  are highly homologous structures, this likely explains the stronger inhibition by  $\alpha$ -PIA, as it is much more selective for  $\alpha 6/\alpha 3$  than for  $\alpha 4$ . Interestingly, both  $\alpha$ -NuxIA and  $\alpha$ -CedIA also inhibited the  $h\beta 3\alpha 4\beta 2\alpha 6\beta 2$  receptor (~25% inhibition), although to a much lesser extent than  $\alpha$ -PIA and  $\alpha$ -PID.

The finding that both  $\alpha$ -PIA and  $\alpha$ -PID conotoxins inhibited  $\alpha 6$  receptors is not altogether surprising since they share 84% homology. More surprising was their drastic difference in affinity for the homomeric neuronal  $\alpha 7$  receptor.  $\alpha$ -PIA did not affect the  $h\alpha 7$  receptors, while  $\alpha$ -PID exhibited strong inhibition (65% at 1  $\mu$ M). When comparing to other  $\alpha 4/7$  conotoxins with  $\alpha 7$  receptor activity (Figure 24), there is significant sequence

homology within the loops. Because the two share 82% identity, they present an excellent case for comparing binding dynamics on the  $\alpha 7$  receptor. Therefore, we employed homology modeling to construct a human  $\alpha 7$  model and predict molecular interactions between  $\alpha$ -PIA/  $\alpha$ -PID in the binding domain.

The h $\alpha 7$  extracellular binding domain bound to  $\alpha$ -PIA/ $\alpha$ -PID was modeled based on the structure of the AChBP bound to the conotoxin  $\alpha$ -LvIA. The resulting models provided information on the molecular interactions of  $\alpha$ -PIA and  $\alpha$ -PID with the  $\alpha 7$  receptor. Despite their sequence homology, the two conotoxins exhibit very different binding dynamics. Previous AChBP binding studies suggest that  $\alpha$ -CNTX inhibitory activity is a result of interaction with the C loop located on the principal subunit. This was the case with the potent  $\alpha 7$  blocker,  $\alpha$ -PnIA, which upon binding locked the C loop in a resting state conformation rendering it unable to be activated [186]. Both the  $\alpha$ -PIA and  $\alpha$ -PID models had more interactions between the complementary subunit than the principal subunit. However,  $\alpha$ -PID has three residue contacts with the principal subunit that are not present in the  $\alpha$ -PIA model. These hydrogen bonds occur at positions N6, P7, and N13, and may be involved in  $\alpha$ -PID inhibition of h $\alpha 7$ . It is important to note that while  $\alpha$ -CNTX interactions with the principal subunit are assumed critical for nAChR inhibition, interactions with the complementary subunit may also play a role. Going forward, amino acid substitution experiments with  $\alpha$ -PIA and  $\alpha$ -PID could be used to confirm the critical residues for  $\alpha 7$  activity.

Alpha conotoxins have been studied for decades to examine their selective binding to nAChRs. These studies, however, have been primarily screens of a single  $\alpha$ -conotoxin after isolation from the venom or through cDNA libraries. Here we demonstrate how advances in sequencing technology aid large-scale prediction of novel  $\alpha$ -conotoxin sequences, allowing us to choose sequences based on amino acid composition and loop size to further dissect the underpinnings for nAChR subtype selectivity.

CHAPTER 6:

CONCLUSIONS AND FUTURE DIRECTIONS

## 6. Conclusions and Future Directions

The therapeutic application of venom peptides has been known for centuries [194]. The use of venom for medicinal purposes is far from a new notion, however, modern advances in molecular technologies have revolutionized venom studies. Natural product discovery, specifically natural peptide discovery, has evolved tremendously due to next-generation sequencing. Genomic data 1) enables large-scale venom protein/peptide discovery and 2) allows us to study the evolution of venom genes across venomous phylogenies. In this body of work, we have capitalized on venomics methodologies to identify cone snail venom peptides with therapeutic potential. Our venomics approach merged transcriptomic and proteomic workflows to interrogate the venom of 17 *Conus* species, focusing on the fish-hunting cone snail, *C. purpurascens*.

We used top-down, high-resolution MS/MS analysis and venom duct transcriptome databases to analyze the injected venom of *C. purpurascens*. We described 33 conopeptides and all of their associated toxiforms. Twenty-one of these venom peptides (64% of the identified components) were identified here for the first time. Of these new conopeptides was PVIIIA, a five disulfide-bonded conotoxin that was abundant in 85% of the injected venom samples, and therefore likely plays an important role in the venom. Comparison of the venom profiles across 27 specimens showed that the venom peptides clustered into two possible profiles with distinct molecular targets. The results provide insight into the probable pharmacological targets of newly identified venom peptides.

We also described a new conoinsulin from the venom of *C. purpurascens*, Con-Ins P1. This is the first ILP identified directly from injected venom, supporting previous evidence of ILP incorporation into the venom. Con-Ins P1 is unique from other conoinsulins in that it displays a different PTM profile and varies drastically in the C-chain sequence/ structure. We hypothesize that the sequence of the B chain C-terminal will

allow it to maintain activity at the insulin receptor but will not lend to self-association of Con-Ins P1 peptides, as self-association of human insulin decreases its bioactivity. Going forward, we will delve into the molecular dynamics and function of Con-Ins P1. First, we will model the Con-Ins P1 peptide using Con-Ins G1 as a template [153], then we will model the dimerization of Con-Ins P1 to assess the dimer stability. We will also model binding dynamics at the human insulin receptor to test our hypothesis. Based on modeling studies, we plan to synthesize the peptide in differentially modified forms for functional assays.

The venom approach employed in this study resulted in the identification of a diverse range of conopeptides with different cysteine frameworks and potential pharmacological targets (Table 1, Table 5). We decided to focus on the  $\alpha$ -conotoxins for functional characterization due to their relative ease of synthesis to obtain large quantities. The  $\alpha$ -conotoxin family of peptides has well-defined cysteine connectivity, and the ones we had synthesized for this study were small ( $\alpha$ -CedIA-13 residues,  $\alpha$ -NuxIA-15 residues,  $\alpha$ -PID- 18 residues). We screened these toxins, along with commercially purchased  $\alpha$ -PIA, on nAChR subtypes using a *Xenopus* oocyte expression system and two-electrode voltage clamp to measure inhibition of ACh-induced current. We found that  $\alpha$ -NuxIA was a potent inhibitor of muscle subtype nAChRs ( $\alpha$ 1 $\beta$ 1 $\delta$  $\epsilon$ , IC<sub>50</sub>= 47 nM). Our results also supported previous work that  $\alpha$ -PIA selectively inhibits  $\alpha$ 6 receptors.  $\alpha$ -PID shares 82% identity with  $\alpha$ -PIA and is a potent inhibitor of both  $\alpha$ 6 and  $\alpha$ 7 receptors. Neuronal  $\alpha$ 6 and homopentameric  $\alpha$ 7 play important roles in neurological and neurodegenerative diseases, as well as addiction [179, 181]. Selective ligands for specific nicotinic receptors are critical for dissecting the pathophysiology of individual subtypes. Using molecular homology modeling we set out to find the discriminating factor for  $\alpha$ 7 selectivity.  $\alpha$ -PIA and  $\alpha$ -PID binding to the h $\alpha$ 7 receptor was modeled based on the AChBP- $\alpha$ -LvIA structure, allowing analysis of molecular interactions between bound  $\alpha$ -CNTX and receptor. This analysis

revealed three  $\alpha$ -PID residues in contact with the principal receptor subunit that may explain increased  $\alpha$ -PIA selectivity toward h $\alpha$ 7 over  $\alpha$ -PIA. Further amino acid substitution functional assays are required to test the critical nature of these residue interactions.

Scientists have historically, and will continue to look towards nature for the answer to medical questions. This body of work demonstrates how venomous methods (transcriptomics, proteomics) advance traditional natural product discovery techniques. In the case of cone snails, millions of years of evolution have engineered libraries of bioactive peptides with high selectivity for clinically important molecular targets. We have sequenced, assembled, and mined their genetic information to harness their venom evolution for our medical advantage. The tight disulfide constrained nature of many conopeptides, such as the cysteine knot peptides, has been shown to increase thermal and proteolytic stability, both important for developing successful drugs. However, there are obstacles and limitations when developing peptides as drugs.

The largest obstacle we face in developing venom proteins/peptides as pharmaceuticals is their inherent inability to cross biological membranes, such as intestinal walls or the blood-brain barrier. In the case of Prialt®, which requires an intrathecal pump for administration, we see how delivery techniques can limit the indication of a drug and can cause potential secondary health effects. Peptide engineering methods for increasing peptide permeation through biological membranes include cyclization and glycosylation [11, 12]. Some conopeptides are naturally glycosylated, such as the analgesic neurotensin analog, contulakin-G [195]. Based on this, cono-glycopeptides may make good drug candidates, however, these peptides are commonly linear, leaving them susceptible to proteolysis. Some peptide therapeutics can be successfully administered by systemic injection, such is the case for insulin and the venom derived diabetes drug, exenatide [196]. This has positive indications for future conoinsulin analogs which are currently being developed [155].

Taken together, it could be argued that the value of venom research in drug discovery is in studying the molecular interactions and binding dynamics of the naturally engineered, highly selective ligands with clinical receptor targets. The large libraries of bioactive molecules provided by venomomics approaches allow us to distinguish the critical residues and receptor interactions necessary for maximal response. Venom peptides thereby provide the tools to develop specific and effective mimetic drugs for a range of neuropathologies (i.e. pain, addiction, neurodegenerative diseases, neuroendocrine disorders). As we continue to understand the value of peptides and biologics as therapeutics, we will need to promote engineering efforts toward improved drug delivery mechanisms.

**APPENDIX A: Toxiforms and sites of modification for each conopeptide identified from the injected venom of *C. purpurascens***

**Table 12-  $\alpha$ -PIA toxiforms**

$\alpha$ -PIA	R	R	D	P	C	C	S	N	P	V	C	T	V	H	N	P	Q	I	C	G
				0					0							D	0	E	*	*

Annotated Sequence	MH+ [Da]
CCSNPVCTVHNPQIC*	1844.77
PCCSNPVCTVHNPQIC*	1941.82
PCCSNPVCTVHNPQICG*	1998.84
DPCCSNPVCTVHNPQIC*	2056.84
DPCCSNPVCTVHNPQIC	2057.83
DPCCSNPVCTVHNOQIC*	2072.84
DPCCSNOVCTVHNPQIC*	2072.84
DPCCSNPVCTVHNPQICG*	2113.87
RDPCCSNPVCTVHNPQIC*	2212.95
RDPCCSNPVCTVHNPQIC	2213.93
RDPCCSNPVCTVHNOQIC*	2228.94
RDPCCSNOVCTVHNPQIC*	2228.94
RDPCCSNOVCTVHNPQIC	2229.92
RDPCCSNPVCTVHNOQIC	2229.93
RDPCCSNPVCTVHDOQIC*	2229.94
RDPCCSNPVCTVHNOqIC*	2229.94
RDPCCSNOVCTVHNOQIC*	2244.94
RDOCCSNOVCTVHNPQIC*	2244.94
RDPCCSNPVCTVHNPQICG*	2269.97
RDPCCSNPVCTVHNPQICG	2270.96
RDPCCSNOVCTVHNPQICG*	2285.96
RRDPCCSNPVCTVHNPQIC*	2369.06

**Table 13-  $\alpha$ -PIB toxiforms**

$\alpha$ -PIB	Q	S	P	G	C	C	W	N	P	A	C	V	K	N	R	C
	Z		0						0						D	*

Annotated Sequence	MH+ [Da]
GCCWNPACVKNRC*	1680.70
OGCCWNPACVKNRC*	1793.74
ZSPGCCWNPACVKNRC*	1975.82
ZSOGCCWNPACVKNRC*	1991.81

ZSOGCCWNPACVKDRC*	1992.79
ZSOGCCWNOACVKNRC*	2007.80

**Table 14- α-PIC toxiforms**

α-PIC	S G C C K P A C G K N R C
-------	---------------------------

Annotated Sequence	Thyo. MH+ [Da]
SGCCKHPACGKNRC	1691.70

**Table 15- α-PID toxiforms**

PID	R D P C C S N P A C N V N N P Q I C G
	0 0 0 E *

Annotated Sequence	MH+ [Da]
CCSNPACNVNNPQIC*	1806.71
PCCSNPACNVNNPQIC*	1903.77
DPCCSNPACNVNNPQIC*	2018.79
RDPCCSNPACNVNNPQIC*	2174.90
RDPCCSNPACNVNNPQIC	2175.88
RDPCCSNPACNVNNOQIC*	2190.89
RDPCCSNOACNVNNPQIC*	2190.89
RDOCCSNOACNVNNPQIC*	2206.88
RDPCCSNOACNVNNOQIC*	2206.89
RDPCCSNPACNVNNPQICG	2232.90
RDPCCSNPACNVNNOEICGRRCSRTLTKTyCOLPD	4348.78
RDPCCSNPACNVNNPEICGRRCSRTLTKTyCOLOD	4348.79

**Table 16- PIE toxiforms**

N A A A K A F D L T A P T A G E G C C F N P A C A N N P N I C
0 0

Annotated Sequence	MH+ [Da]
NAAAKAFDLTAPTAGEGCCFNOCVNNPNIC	3412.48
NAAAKAFDLTAOTAGEGCCFNPAVNNPNIC	3412.48

**Table 17- PIF toxiforms**

PIF	Q	E	P	G	C	C	R	N	P	A	C	V	K	H	R	C
	Z	Y	O				D	O								*

Annotated Sequence	MH+ [Da]
GCCRNPACVKHRC*	1673.73
PGCCRNPACVKHRC*	1770.79
PGCCRNOACVKHRC*	1786.78
OGCCRNPACVKHRC*	1786.79
EPGCCRNPACVKHRC*	1899.83
ZEPGCCRNPACVKHRC*	2010.86
ZEPGCCRDPAVKHRC*	2011.84
ZEPGCCRNPAVKHRC	2011.85
ZEPGCCRNOACVKHRC*	2026.86
ZEOGCCRNPAVKHRC*	2026.86
ZEOGCCRNOACVKHRC*	2042.85
ZyPGCCRNPACVKHRC*	2054.85
ZyOGCCRNPACVKHRC*	2070.85
ZyPGCCRNOACVKHRC*	2070.85

**Table 18- PIG toxiforms**

PIG	P	C	C	S	N	P	V	C	T	V	H	G	G	P	Q	L	C
																	*

Annotated Sequence	MH+ [Da]
CCSNPVC TVHGGPQLC*	1844.77
PCCSNPVC TVHGGPQLC*	1941.82

**Table 19- α-PIVA toxiforms**

	G	C	C	G	S	Y	P	N	A	A	C	H	P	C	S	C	K	D	R	P	S	Y	C	G	Q	G	
				B	O	D						O							O		*	*	E*	*			

Annotated Sequence	MH+ [Da]
GCCGSYPNAA CHP CSCKDRPSYC	2764.03
GCCGSYPDAACHP CSCKDRPSYC	2765.03
GCCGSYPNAA CHP CSCKDROSYC*	2779.03
GCCGSYPNAA CHP CSCKDROSYC	2780.02
GCCGSYPNAA CHO CSCKDRPSYC	2780.02
GCCGSYPDAACHO CSCKDRPSYC	2781.00
GCCGSYPDAACHP CSCKDROSYC	2781.02

GCCGSYONAA <b>CHOC</b> SCCKDRPSYC*	2795.03
GCCGSYONAA <b>CHOC</b> SCCKDRPSYC	2796.01
GCCGSYPNAA <b>CHOC</b> SCCKDR <b>OS</b> YC	2796.01
GCCGSYODAA <b>CHOC</b> SCCKDRPSYC	2797.00
GCCGSYONAA <b>CHOC</b> SCCKDR <b>OS</b> YC	2812.01
GCCGSYODAA <b>CHOC</b> SCCKDR <b>OS</b> YC	2813.01
GCCGSYPNAA <b>CHP</b> SCCKDRPSYCG	2821.05
GCCGSYPNAA <b>CHP</b> SCCKDR <b>OS</b> YCG	2837.04
GCCGSYPNAA <b>CHOC</b> SCCKDRPSYCG	2837.04
GCCGSYPDA <b>CHP</b> SCCKDR <b>OS</b> YCG	2838.04
GCCGSYPDA <b>CHOC</b> SCCKDRPSYCG	2838.04
GCCGSYPNAA <b>CHOC</b> SCCKDR <b>OS</b> YCG	2853.03
GCCGSYONAA <b>CHOC</b> SCCKDRPSYCG	2853.04
GCCGSYONAA <b>CHOC</b> SCCKDR <b>OS</b> YCG	2869.03
GCCGSYODAA <b>CHOC</b> SCCKDR <b>OS</b> YCG	2870.03
CCGSYPNAA <b>CHP</b> SCCKDRPSYCGQ**	2891.10
CCGSYPNAA <b>CHOC</b> SCCKDRPSYCGQ*	2907.09
CCGSYONAA <b>CHP</b> SCCKDRPSYCGQ*	2907.09
CCGSYPNAA <b>CHP</b> SCCKDR <b>OS</b> YCGQ*	2907.10
CCGSYPDA <b>CHP</b> SCCKDR <b>OS</b> YCGQ*	2908.09
CCGSYONAA <b>CHOC</b> SCCKDRPSYCGQ*	2923.09
CCGSYPNAA <b>CHOC</b> SCCKDR <b>OS</b> YCGQ*	2923.09
CCGSYODAA <b>CHOC</b> SCCKDRPSYCGQ*	2924.08
CCGSYPDA <b>CHOC</b> SCCKDR <b>OS</b> YCGQ*	2924.09
CCGSYONAA <b>CHOC</b> SCCKDR <b>OS</b> YCGQ*	2939.08
GCCGSYPNAA <b>CHP</b> SCCKDRPSYCGQ*	2948.12
GCCGSYPNAA <b>CHP</b> SCCKDRPSYCGQ	2949.11
GCCGSYPNAA <b>CHP</b> SCCKDRPSYCGQ*	2950.10
GCCGSYPDA <b>CHP</b> SCCKDRPSYCGQ	2950.11
CCGSYPNAA <b>CHP</b> SCCKDR <b>OS</b> YCGQG*	2964.11
GCCGSYPNAA <b>CHOC</b> SCCKDRPSYCGQ*	2964.11
GCCGSYPNAA <b>CHP</b> SCCKDR <b>OS</b> YCGQ*	2964.12
GCCGSYONAA <b>CHP</b> SCCKDRPSYCGQ*	2964.13
GCCGSYODAA <b>CHP</b> SCCKDRPSYCGQ*	2965.10
GCCGSYPDA <b>CHOC</b> SCCKDRPSYCGQ*	2965.10
GCCGSYPNAA <b>CHP</b> SCCKDR <b>OS</b> YCGQ	2965.11
GCCGSYONAA <b>CHP</b> SCCKDRPSYCGQ	2965.11
GCCGSYPNAA <b>CHOC</b> SCCKDRPSYCGQ	2965.11
GCCGSYPDA <b>CHP</b> SCCKDR <b>OS</b> YCGQ*	2965.12
GCCGSYONAA <b>CHP</b> SCCKDRPSYCGQ*	2966.10
GCCGSYODAA <b>CHP</b> SCCKDRPSYCGQ*	2967.09
CCGSYPNAA <b>CHOC</b> SCCKDR <b>OS</b> YCGQG*	2980.08
CCGSYONAA <b>CHOC</b> SCCKDRPSYCGQG*	2980.11

GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGQ*	2980.11
GCCGSYPNA <b>AACHOC</b> SC <b>KDR</b> OSYCGQ*	2980.11
GCCGSYPDA <b>AACHOC</b> SC <b>KDR</b> OSYCGQ*	2981.10
GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGQ	2981.10
GCCGSYPNA <b>AACHOC</b> SC <b>KDR</b> OSYCGQ	2981.10
GCCGSYODAA <b>CHOC</b> SC <b>KDRPSY</b> CGQ*	2981.11
GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGQ*	2982.10
GCCGSYONAA <b>CHOC</b> SC <b>KDR</b> OSYCGQ*	2996.11
GCCGSYODAA <b>CHOC</b> SC <b>KDR</b> OSYCGQ*	2997.09
GCCGSYONAA <b>CHOC</b> SC <b>KDR</b> OSYCGQ	2997.10
GCCGSYPNA <b>AACHP</b> CS <b>KDRPSY</b> CGQG*	3005.14
GCCGSYPNA <b>AACHP</b> CS <b>KDRPSY</b> CGQG	3006.14
GCCGSYPDA <b>AACHP</b> CS <b>KDRPSY</b> CGEG	3008.10
GCCGSYPNA <b>AACHP</b> CS <b>KDR</b> OSYCGQG*	3021.14
GCCGSYPNA <b>AACHOC</b> SC <b>KDRPSY</b> CGQG*	3021.14
GCCGSYONAA <b>AACHP</b> CS <b>KDRPSY</b> CGQG*	3021.14
GCCGSYPNA <b>AACHOC</b> SC <b>KDRPSY</b> CGQG	3022.12
GCCGSYPNA <b>AACHOC</b> SC <b>KDRPSY</b> CGEG*	3022.12
GCCGSYPNA <b>AACHP</b> CS <b>KDR</b> OSYCGQG	3022.13
GCCGSYPNA <b>AACHP</b> CS <b>KDR</b> OSYCGEG*	3022.13
GCCGSYPDA <b>AACHP</b> CS <b>KDR</b> OSYCGQG*	3022.13
GCCGSYPDA <b>AACHOC</b> SC <b>KDRPSY</b> CGQG*	3022.13
GCCGSYPNA <b>AACHP</b> CS <b>KDR</b> OSYCGEG	3023.12
GCCGSYPNA <b>AACHOC</b> SC <b>KDRPSY</b> CGEG	3023.12
GCCGSYPDA <b>AACHOC</b> SC <b>KDRPSY</b> CGEG*	3023.12
GCCGSYPDA <b>AACHOC</b> SC <b>KDRPSY</b> CGQG	3023.12
GCCGSYPDA <b>AACHP</b> CS <b>KDR</b> OSYCGEG*	3023.12
GCCGSYPDA <b>AACHP</b> CS <b>KDR</b> OSYCGQG	3023.12
GCCGSYONAA <b>AACHP</b> CS <b>KDRPSY</b> CGEG	3023.12
GCCGSYODAA <b>AACHP</b> CS <b>KDRPSY</b> CGEG	3024.10
GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGQG*	3037.13
GCCGSYPNA <b>AACHOC</b> SC <b>KDR</b> OSYCGQG*	3037.13
GCCGSYPNA <b>AACHOC</b> SC <b>KDR</b> OSYCGQG	3038.11
GCCGSYPNA <b>AACHOC</b> SC <b>KDR</b> OSYCGEG*	3038.11
GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGEG*	3038.12
GCCGSYONAA <b>CHOC</b> SC <b>KDRPSY</b> CGQG	3038.12
GCCGSYODAA <b>CHOC</b> SC <b>KDRPSY</b> CGQG*	3038.12
GCCGSYPDA <b>AACHOC</b> SC <b>KDR</b> OSYCGQG*	3038.12
GCCGSYONAA <b>CHOC</b> SC <b>KDR</b> OSYCGQG*	3053.13
GCCGSYODAA <b>CHOC</b> SC <b>KDR</b> OSYCGQG*	3054.11
GCCGSYONAA <b>CHOC</b> SC <b>KDR</b> OSYCGEG*	3054.12
GCCGSYONAA <b>CHOC</b> SC <b>KDR</b> OSYCGQG	3054.12
CCGSBONAA <b>CHOC</b> SC <b>KDR</b> OSYCGQG*	3074.04

RGCCGSYPNAA <b>CHP</b> CSCKDR <b>OSY</b> CGQ*	3120.22
RGCCGSYPDA <b>A</b> CHP <b>CS</b> CKDR <b>OSY</b> CGQ*	3121.17
RGCCGSYPDA <b>A</b> CHP <b>CS</b> CKDR <b>OSY</b> CGQ*	3178.19
RGCCGSYPNAA <b>CHP</b> CSCKDR <b>OSY</b> CGEG*	3178.19
RGCCGSYPNAA <b>CHP</b> CSCKDR <b>OSY</b> CGQG	3178.19
GCCGSYPNAA <b>CHO</b> CSCKDR <b>PSY</b> CGEGR	3179.20
GCCGSYPNAA <b>CHP</b> CSCKDR <b>OSY</b> CGEGR	3179.20
SHVVRG <b>CC</b> GSYPDA <b>A</b> CHP <b>CS</b> CKDR <b>OSY</b> CGQ*	3543.45
SHVVRG <b>CC</b> GSY <b>ONAA</b> CHO <b>CS</b> CKDR <b>PSY</b> CGQ*	3558.44
SHVVRG <b>CC</b> GSYPNAA <b>CHO</b> CSCKDR <b>OSY</b> CGQ*	3558.44
LLSHVVRG <b>CC</b> GSY <b>ONAA</b> CHO <b>CS</b> CKDR <b>PSY</b> CGQ*	3784.60

Table 20- κ-PIVE toxiforms

R	D	C	C	G	V	K	L	E	M	C	H	P	C	L	C	D	N	S	C	K	N	Y	G	K	G	
								y	O			O					D					D	*	*	*	*

Annotated Sequence	MH+ [Da]
DCCGVKLEM <b>CHP</b> CLDNSCKN	2657.02
DCCGVKLEM <b>CHP</b> CLDD <b>S</b> CKN	2658.00
DCCGVKLEM <b>CHP</b> CLDNSCKD	2658.00
DCCGVKLE <b>m</b> CHPCLDNSCKN	2673.01
DCCGVKLEM <b>CHO</b> CLDNSCKN	2673.01
DCCGVKLE <b>m</b> CHOCLDNSCKN	2689.01
DCCGVKLY <b>M</b> CHPCLDNSCKN	2701.02
<b>CC</b> GVKLEM <b>CHP</b> CLDNSCKNY	2705.06
<b>CC</b> GVKLEM <b>CHP</b> CLDNSCKNYG	2762.07
DCCGVKLEM <b>CHP</b> CLDNSCKNY*	2819.10
DCCGVKLEM <b>CHP</b> CLDNSCKNY	2820.08
DCCGVKLEM <b>CHP</b> CLDD <b>S</b> CKDY*	2821.07
DCCGVKLEM <b>CHP</b> CLDD <b>S</b> CKNY	2821.07
DCCGVKLE <b>m</b> CHPCLDNSCKNY	2836.08
DCCGVKLEM <b>CHO</b> CLDNSCKDY*	2836.08
DCCGVKLEM <b>CHO</b> CLDNSCKNY	2836.08
DCCGVKLE <b>m</b> CHPCLDNSCKDY*	2836.09
DCCGVKLE <b>m</b> CHPCLDD <b>S</b> CKNY*	2836.09
DCCGVKLY <b>M</b> CHPCLDNSCKNY	2864.09
DCCGVKLY <b>M</b> CHPCLDNSCKDY*	2864.09
DCCGVKLEM <b>CHP</b> CLDNSCKNYG	2877.10
DCCGVKLEM <b>CHP</b> CLDD <b>S</b> CKNYG	2878.09
DCCGVKLEM <b>CHP</b> CLDD <b>S</b> CKDYG*	2878.09
DCCGVKLEM <b>CHP</b> CLDNSCKDYG	2878.09
<b>CC</b> GVKLEM <b>CHP</b> CLDNSCKNYGK*	2889.19

CCGVKLEMCHPCLDNSCKNYGK	2890.19
DCCGVKLEmCHPCLDNSCKDYG*	2893.10
DCCGVKLEmCHPCLDNSCKNYG	2893.10
DCCGVKLEMCHOCLDNSCKNYG	2893.10
DCCGVKLEMCHOCLDNSCKDYG*	2893.10
DCCGVKLyMCHPCLDNSCKNYG*	2920.11
DCCGVKLyMCHPCLDNSCKDYG*	2921.09
DCCGVKLyMCHPCLDNSCKNYG	2921.11
CCGVKLyMCHPCLDNSCKNYGK*	2933.19
DCCGVKLEMCHPCLDNSCKNYGK*	3004.21
DCCGVKLEMCHPCLDNSCKNYGK	3005.20
DCCGVKLEMCHPCLDDSCCKDYG*	3006.20
DCCGVKLEmCHPCLDNSCKNYGK*	3020.20
DCCGVKLEMCHOCLDNSCKNYGK*	3020.20
DCCGVKLEMCHOCLDNSCKDYGK*	3021.20
DCCGVKLEmCHPCLDNSCKDYGK*	3021.20
DCCGVKLEmCHPCLDDSCCKNYGK*	3021.21
DCCGVKLEMCHOCLDDSCCKNYGK*	3021.21
DCCGVKLEmCHOCLDNSCKNYGK*	3036.20
DCCGVKLEmCHOCLDDSCCKNYGK*	3037.21
DCCGVKLEmCHOCLDDSCCKDYGK*	3038.20
DCCGVKLyMCHPCLDNSCKNYGK*	3048.20
DCCGVKLyMCHPCLDNSCKDYGK*	3049.21
DCCGVKLyMCHPCLDDSCCKNYGK*	3049.21
DCCGVKLEMCHPCLDNSCKNYGKG*	3061.23
DCCGVKLEMCHPCLDNSCKNYGKG	3062.22
DCCGVKLyMCHPCLDNSCKNYGK*	3064.22
DCCGVKLyMCHOCLDNSCKNYGK*	3064.22
DCCGVKLEmCHPCLDNSCKNYGKG*	3077.23
DCCGVKLyMCHOCLDNSCKNYGK*	3080.21
DCCGVKLyMCHPCLDNSCKNYGKG*	3105.24
RDCCGVKLEMCHPCLDDSCCKDYGK*	3162.25
RDCCGVKLEmCHPCLDNSCKNYGK*	3176.28
RDCCGVKLEMCHOCLDNSCKNYGK*	3176.28
RDCCGVKLEMCHOCLDDSCCKNYGK*	3177.27
RDCCGVKLEmCHPCLDDSCCKNYGK*	3177.28
RDCCGVKLEmCHOCLDNSCKNYGK*	3192.27
RDCCGVKLEmCHOCLDDSCCKNYGK*	3193.26
RDCCGVKLyMCHPCLDNSCKNYGK*	3220.28
DCCGVKLEMCHPCLDNSCKNYGKGKKEY*	3609.53
DCCGVKLEMCHPCLDNSCKNYGKGKKEY	3610.53
DCCGVKLEMCHPCLDDSCCKDYGKGKKEY*	3611.53
DCCGVKLyMCHPCLDNSCKNYGKGKKEY*	3653.54

NAAVNDKASHLIDNVIRDCCGVKLEMCHPCLCDNSCKNYGK*	4835.18
--------------------------------------------	---------

**Table 21- κ-PIVF toxiforms**

D	C	C	G	V	K	L	E	M	C	H	P	C	L	C	D	N	S	C	K	K	S	G	K
								Y	O		O				D				*	*	*	*	

Annotated Sequence	MH+ [Da]
DCCGVKLEMCHPCLCDNSCKK	2671.07
DCCGVKLEMCHPCLCDDSCKK	2672.05
DCCGVKLEmCHPCLCDNSCKK	2687.06
DCCGVKLEmCHPCLCDDSCKK*	2687.06
DCCGVKLEmCHPCLCDDSCKK	2688.07
DCCGVKLEMCHOCLCDDSCKK	2688.07
DCCGVKLYMCHPCLCDNSCKK	2715.06
DCCGVKLYMCHPCLCDDSCKK	2716.06
DCCGVKLYmCHPCLCDNSCKK	2731.07
DCCGVKLEMCHPCLCDNSCKKS*	2757.12
DCCGVKLEMCHPCLCDNSCKKS	2758.10
DCCGVKLEmCHPCLCDNSCKKS*	2773.11
DCCGVKLYMCHPCLCDNSCKKS*	2801.12
DCCGVKLYmCHPCLCDNSCKKS	2802.11
DCCGVKLEMCHPCLCDNSCKKSG*	2814.13
DCCGVKLEMCHPCLCDNSCKKSG	2815.12
DCCGVKLEMCHPCLCDDSCKKSG	2816.12
CCGVKLEMCHPCLCDNSCKKSGK*	2827.20
CCGVKLEMCHPCLCDNSCKKSGK	2828.21
DCCGVKLYMCHPCLCDNSCKKSG*	2858.14
DCCGVKLEMCHPCLCDNSCKKSGK*	2942.23
DCCGVKLEMCHPCLCDNSCKKSGK	2943.24
DCCGVKLEMCHOCLCDNSCKKSGK*	2958.22
DCCGVKLEmCHPCLCDNSCKKSGK*	2958.23
DCCGVKLEMCHOCLCDDSCKKSGK*	2959.23
DCCGVKLEmCHPCLCDDSCKKSGK*	2959.23
DCCGVKLEmCHOCLCDNSCKKSGK*	2974.22
DCCGVKLEmCHOCLCDDSCKKSGK*	2975.22
DCCGVKLYMCHPCLCDNSCKKSGK*	2986.23
DCCGVKLYMCHPCLCDDSCKKSGK*	2987.23
DCCGVKLYmCHPCLCDNSCKKSGK*	3002.23
DCCGVKLYmCHOCLCDNSCKKSGK*	3002.24

**Table 22- PIVH toxiforms**

PIVH	R	D	C	C	G	V	V	M	E	E	C	H	K	C	L	C	N	Q	T	C	K	K	K	G
								O	Y	Y								D	E	*	*	*	*	*

Annotated Sequence	MH+ [Da]
DCCGVMEECHKCLCNQTC	2459.90
DCCGVMEECHKCLCNETC	2460.90
DCCGVMEECHKCLCDQTC	2460.90
DCCGVmEECHKCLCNQTC	2475.89
DCCGVVMγECHKCLCNQTC*	2502.90
DCCGVMEγCHKCLCNQTC*	2502.90
DCCGVMEγCHKCLCNQTC	2503.91
DCCGVVMγECHKCLCNQTC	2503.91
DCCGVMEECHKCLCNQTCK	2587.99
DCCGVMEECHKCLCDQTCK	2589.00
DCCGVMEECHKCLCNETCK	2589.00
DCCGVVMγECHKCLCNQTCK*	2631.00
DCCGVMEγCHKCLCNQTCK*	2631.00
DCCGVMEγCHKCLCNQTCK	2632.00
DCCGVVMγECHKCLCNQTCK	2632.00
DCCGVMEECHKCLCNQTCKK*	2715.11
DCCGVMEECHKCLCNQTCKK	2716.09
DCCGVMEECHKCLCNETCKK	2717.09
DCCGVMEECHKCLCDQTCKK	2717.09
CCGVMEECHKCLCNQTCKKK*	2728.17
CCGVMEECHKCLCNQTCKKK	2729.18
DCCGVMEγCHKCLCNQTCKK*	2759.10
DCCGVVMγECHKCLCNQTCKK*	2759.10
DCCGVVMγECHKCLCNQTCKK	2760.10
CCGVmγγCHKCLCNQTCKKK*	2832.18
DCCGVMEECHKCLCNQTCKKK*	2843.20
DCCGVMEECHKCLCNQTCKKK	2844.20
DCCGVmEECHKCLCNQTCKKK*	2859.19
DCCGVmEECHKCLCNETCKKK*	2860.20
DCCGVmEECHKCLCDQTCKKK*	2860.20
DCCGVmEECHKCLCDETCKKK*	2861.18
DCCGVVMγECHKCLCNQTCKKK*	2887.19
DCCGVMEγCHKCLCNQTCKKK*	2887.19
DCCGVMEγCHKCLCDQTCKKK*	2888.20
DCCGVVMγECHKCLCDQTCKKK*	2888.20
DCCGVMEECHKCLCNQTCKKKG*	2900.22
DCCGVMEECHKCLCNQTCKKKG	2901.23
DCCGVMEECHKCLCDETCKKKG*	2902.20

DCCGVMEECHKCLCDQTCCKKG	2902.20
DCCGVMEECHKCLCNETCCKKG	2902.20
DCCGVMEECHKCLCNQTCCKKG*	2916.22
DCCGVMEECHKCLCDQTCCKKG	2918.20
DCCGVMEECHKCLCNETCCKKG	2918.20
DCCGVMEECHKCLCDETCCKKG*	2918.20
DCCGVMEYCHKCLCNQTCCKKG*	2944.22
DCCGVMEYCHKCLCNQTCCKKG*	2944.22
RDCCGVMEECHKCLCNQTCCKKG*	3015.27

**Table 23- Linear-P toxiforms**

LINEAR-P	F	Q	P	S	A	E	N	E	E	G	K	F	R	F	F	D	K	Q	Q
		Z	O															E	*

Annotated Sequence	MH+ [Da]
ZPSAENEEGKFRFFDKQ	2039.94
ZOSAENEEGKFRFFDKQ	2055.94
ZPSAENEEGKFRFFDKQQ	2168.00
ZPSAENEEGKFRFFDKEQ*	2168.00
FQPSAENEEGKFRFFDKQ	2204.04
FQPSAENEEGKFRFFDKQQ	2332.10

**Table 24- Ile-Contryphan-P toxiforms**

Ile-Contryphan-P	A	T	S	L	G	C	V	I	W	P	W	C
									O			*

Annotated Sequence	MH+ [Da]
CVIWPWC	1020.44
CVIWOWC	1036.44
GCVIWPWC*	1076.48
GCVIWPWC	1077.46
GCVIWOWC	1093.46
LGCVIWPWC	1190.55
SLGCVIWPWC	1277.58
ATSLGCVIWPWC	1449.66

**Table 25- Contryphan-P3 toxiforms**

Contryphan-P3	A	T	S	L	A	C	A	I	W	T	K	C

Annotated Sequence	MH+ [Da]
CAIWT <b>C</b>	938.42
ACAIWT <b>C</b>	1009.46
ATSLACAIWT <b>C</b>	1381.66

**Table 26- Contryphan-P4 toxiforms**

Contryphan-P4	<b>C</b> V Y W R K <b>C</b>
---------------	-----------------------------

Annotated Sequence	MH+ [Da]
<b>CVYWRK</b> <b>C</b>	1071.49

**Table 27-  $\psi$ -PIIE toxiforms**

R	H	P	P	<b>C</b>	<b>C</b>	L	Y	G	K	<b>C</b>	R	R	Y	P	G	<b>C</b>	S	S	A	S	<b>C</b>	<b>C</b>	Q	R	
		0	0												0									E	*

Annotated Sequence	MH+ [Da]
<b>OC</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQ</b>	2659.04
<b>C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	2701.11
<b>H00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCC</b>	2781.09
<b>P</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	2782.17
<b>P</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	2798.16
<b>OC</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	2798.16
<b>OC</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	2814.16
<b>OC</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCER*</b>	2815.16
<b>OC</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR</b>	2815.16
<b>H00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQ</b>	2909.16
<b>00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	2927.20
<b>00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCER*</b>	2928.20
<b>00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR</b>	2928.20
<b>HPP</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	3016.28
<b>HPO</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	3032.27
<b>HOP</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	3032.28
<b>HPP</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	3032.28
<b>H00C</b> CLY <b>GK</b> <b>CRRYPG</b> <b>CSSASCCQR*</b>	3048.27
<b>HPO</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	3048.27
<b>H00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR*</b>	3064.26
<b>H00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCER*</b>	3065.26
<b>H00C</b> CLY <b>GK</b> <b>CRRYOG</b> <b>CSSASCCQR</b>	3065.26

RHPOCCLYGKCRRYOGCSSASCCQR*	3204.37
RH00CCLYGKCRRYPGCSASCCQR*	3204.37
RH00CCLYGKCRRYOGCSSASCCQR*	3220.38
RH00CCLYGKCRRYOGCSSASCCER*	3221.36
RH00CCLYGKCRRYOGCSSASCCQR	3221.36

**Table 28- PIIIG toxiforms**

PIIIG	Q W G C C P V N A C R S C H C C
	Z D *

Annotated Sequence	MH+ [Da]
ZWGCCPVNACRSCHCC*	2093.75
ZWGCCPVDACRSCHCC*	2094.74

**Table 29- PIIIH toxiforms**

E I I L P A L G Q R K C C P L T A C K L G S G C K C C E
Z O

Annotated Sequence	MH+ [Da]
KCCPLTACKLGSGCKCC	2059.87
CCPLTACKLGSGCKCCE	2060.81
KCCPLTACKLGSGCKCCE	2188.91
KCCOLTACKLGSGCKCCE	2204.90
LGQRKCCPLTACKLGSGCKCCE	2643.17
ZIILPALGQRKCCPLTACKLGSGCKCCE	3261.55
EIILPALGQRKCCPLTACKLGSGCKCCE	3279.56

**Table 30- PIIII toxiforms**

PIIII	C C Q A Y C S R Y H C L P C C
-------	-------------------------------

Annotated Sequence	MH+ [Da]
CCQAYCSRYHCLPCC	2094.75

**Table 31- δ-PVIA toxiforms**

E A C Y A P G T F C G I K P G L C C S E F C L P G V C F G
O O *

Annotated Sequence	MH+ [Da]
EAC <b>Y</b> A <b>O</b> G <b>T</b> F <b>C</b> G <b>I</b> K <b>O</b> G <b>L</b> <b>C</b> C <b>S</b> E <b>F</b> <b>C</b> L <b>P</b> G <b>V</b> <b>C</b>	3142.29
<b>C</b> <b>Y</b> A <b>O</b> G <b>T</b> F <b>C</b> G <b>I</b> K <b>O</b> G <b>L</b> <b>C</b> C <b>S</b> E <b>F</b> <b>C</b> L <b>P</b> G <b>V</b> <b>C</b> F <b>G</b> *	3145.31
EASKLDKKEA <b>C</b> <b>Y</b> A <b>O</b> G <b>T</b> F <b>C</b> G <b>I</b> K <b>O</b> G <b>L</b> <b>C</b> C <b>S</b> E <b>F</b> <b>C</b> L <b>P</b> G <b>V</b> <b>C</b> F <b>G</b> *	4244.90

**Table 32- PVIB toxiforms**

Q	C	T	P	Y	G	G	S	C	G	V	D	S	T	C	C	G	R	C	N	V	P	R	N	K	C	E	L
Z			0																D		O		D			Y	*

Annotated Sequence	MH+ [Da]
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> *	3033.20
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b>	3033.20
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>N</b> K <b>C</b>	3034.20
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>D</b> K <b>C</b> *	3034.20
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> *	3049.19
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b>	3049.19
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>D</b> K <b>C</b> *	3050.19
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>N</b> K <b>C</b>	3050.19
<b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3051.21
<b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3067.20
<b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b> *	3067.20
<b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> OR <b>N</b> K <b>C</b> <b>E</b>	3083.20
<b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> OR <b>D</b> K <b>C</b> <b>E</b> *	3083.20
<b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> OR <b>N</b> K <b>C</b> <b>E</b> *	3083.20
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b> *	3162.24
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b> *	3162.24
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3162.25
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b> *	3163.24
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3163.24
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b>	3163.24
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b>	3164.23
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3178.23
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b> *	3178.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> OR <b>D</b> K <b>C</b> <b>E</b> *	3178.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> OR <b>N</b> K <b>C</b> <b>E</b>	3178.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> OR <b>N</b> K <b>C</b> <b>E</b> *	3178.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> N <b>V</b> OR <b>D</b> K <b>C</b> <b>E</b>	3179.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> OR <b>N</b> K <b>C</b> <b>E</b>	3179.23
Z <b>C</b> T <b>P</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> OR <b>D</b> K <b>C</b> <b>E</b> *	3179.23
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>N</b> K <b>C</b> <b>E</b>	3179.24
Z <b>C</b> T <b>O</b> Y <b>G</b> G <b>S</b> <b>C</b> G <b>V</b> D <b>S</b> T <b>C</b> C <b>G</b> R <b>C</b> D <b>V</b> PR <b>D</b> K <b>C</b> <b>E</b> *	3179.24

ZCTOYGGSCGVDSTCCGRCNVPRDKCE	3179.24
QCTPYGGSCGVDSTCCGRCNVPRNKCE	3179.24
ZCTOYGGSCGVDSTCCGRCDVPRDKCE	3180.22
QCTPYGGSCGVDSTCCGRCNVPRDKCE	3180.22
ZCTPYGGSCGVDSTCCGRCDVORDKCE	3180.22
QCTPYGGSCGVDSTCCGRCDVPRDKCE*	3180.24
QCTPYGGSCGVDSTCCGRCDVPRNKCE	3180.24
ZCTPYGGSCGVDSTCCGRCNVPRNKCE	3180.24
ZCTPYGGSCGVDSTCCGRCNVPRDKCE*	3180.24
QCTPYGGSCGVDSTCCGRCDVPRDKCE	3181.21
ZCTOYGGSCGVDSTCCGRCNVORNKCE*	3193.25
ZCTOYGGSCGVDSTCCGRCNVORNKCE	3194.23
ZCTOYGGSCGVDSTCCGRCNVORDKCE*	3194.23
ZCTOYGGSCGVDSTCCGRCDVORNKCE*	3194.24
ZCTOYGGSCGVDSTCCGRCNVORDKCE	3195.22
ZCTOYGGSCGVDSTCCGRCDVORNKCE	3195.23
QCTPYGGSCGVDSTCCGRCNVORDKCE*	3195.23
QCTPYGGSCGVDSTCCGRCNVORNKCE	3195.23
ZCTOYGGSCGVDSTCCGRCNVPRNKCy*	3222.22
ZCTPYGGSCGVDSTCCGRCNVORNKCy	3222.25
ZCTOYGGSCGVDSTCCGRCDVPRNKCy	3223.22
ZCTPYGGSCGVDSTCCGRCNVPRDKCEL	3276.28
ZCTPYGGSCGVDSTCCGRCDVPRNKCEL	3276.28
ZCTPYGGSCGVDSTCCGRCDVPRDKCEL*	3276.28
ZCTPYGGSCGVDSTCCGRCDVPRDKCEL	3277.29
ZCTOYGGSCGVDSTCCGRCNVPRDKCEL	3292.27
ZCTOYGGSCGVDSTCCGRCDVPRDKCEL*	3292.27
ZCTOYGGSCGVDSTCCGRCDVPRNKCEL	3292.27
ZCTPYGGSCGVDSTCCGRCDVORNKCEL	3292.27
ZCTPYGGSCGVDSTCCGRCDVORDKCEL*	3292.27
ZCTOYGGSCGVDSTCCGRCDVPRDKCEL	3293.27
ZCTOYGGSCGVDSTCCGRCNVORDKCEL	3308.27
ZCTOYGGSCGVDSTCCGRCDVORNKCEL	3308.27
ZCTOYGGSCGVDSTCCGRCDVORDKCEL	3309.27
ZCTPYGGSCGVDSTCCGRCNVORNKCyL	3335.31
ZCTPYGGSCGVDSTCCGRCNVORDKCyL*	3335.31
ZCTPYGGSCGVDSTCCGRCDVORNKCyL*	3335.31
ZCTOYGGSCGVDSTCCGRCDVPRDKCyL	3337.27

**Table 33- PVIC sites of modification** \*\*There were no spectral matches to the full PVIC mature peptide. Sites of modification were determined by matches to peptide fragments

E A C Y A P G T F C G I K P G L C C S A L C L P A V C I D

Z 0 0 \*

**Table 34- PVID toxiforms**

S N R P C K K S G R K C F P H Q K D C C G R A C I I T I C P  
 D O O E \*

Annotated Sequence	MH+ [Da]
SNROCCKSGRKCFOHQKDCCGRACIITICP	3723.74
SNROCCKSGRKCFOHEKDCCGRACIITICP*	3723.74
SDROCCKSGRKCFOHQKDCCGRACIITICP*	3723.74

**Table 35- PVIE toxiforms**

V G E F R G C A H I N Q A C N P P Q C C R G Y T C Q S S Y I P S C Q L  
 D E D O O E O E \*

Annotated Sequence	MH+ [Da]
GCAHINQACN00QCCRGYTCQSSYIOSCQL	3635.47
GCAHINQACN00QCCRGYTCQSSYIOSCEL*	3635.47
VGEFRGCAHINQACNPPQCCRGYTCQSSYIOSCQL	4191.78
VGEFRGCAHINQACNPPQCCRGYTCQSSYIOSCEL*	4191.78
VGEFRGCAHINQACNPOQCCRGYTCQSSYIOSCQL	4207.77
VGEFRGCAHINQACNPOQCCRGYTCQSSYIOSCEL*	4207.77
VGEFRGCAHINQACNPOECCRGYTCQSSYIOSCQL*	4207.77
VGEFRGCAHINQACDPOQCCRGYTCQSSYIOSCQL*	4207.77
VGEFRGCAHINQACN00QCCRGYTCQSSYIPSCQL	4207.77
VGEFRGCAHINQACN00QCCRGYTCQSSYIPSCCEL*	4207.77
VGEFRGCAHINQACN00QCCRGYTCQSSYIOSCQL	4223.76
VGEFRGCAHINQACN00QCCRGYTCQSSYIOSCEL*	4223.76
VGEFRGCAHINEACN00QCCRGYTCQSSYIOSCQL*	4223.77
VGEFRGCAHINQACD00QCCRGYTCQSSYIOSCQL*	4223.77
VGEFRGCAHIDQACN00QCCRGYTCQSSYIOSCQL*	4223.77
VGEFRGCAHINQACN00ECCRGYTCQSSYIOSCQL*	4223.77

**Table 36- PVIF toxiforms**

A T S N R P C K K T G R K C F P H Q K D C C G R A C I I T I C P  
 D O O E \*

Annotated Sequence	MH+ [Da]
ATSNROCCKTGRKCFHEKDCCGRACIITICP*	3909.83

ATSDROCKKTGRKCFQHqKDCCGRACIITICP*	3909.83
ATSNROCKKTGRKCFQHqKDCCGRACIITICP	3909.83

**Table 37- PVIG toxiforms**

S	T	T	K	G	A	T	S	N	R	P	C	K	I	P	G	R	K	C	F	P	H	Q	K	D
								D	O			O							O	E				

C	C	G	R	A	C	I	I	T	I	C	P
											*

Annotated Sequence	MH+ [Da]
GATSNROCKIogrKCFPHqKDCCGRACIITICP*	3947.85
GATSNRPCKIogrKCFOHqKDCCGRACIITICP*	3947.85
GATSNROCKIogrKCFPHQKDCCGRACIITICP	3947.85
GATSNRPCKIogrKCFOHQKDCCGRACIITICP	3947.85
GATSNROCKIogrKCFPHqKDCCGRACIITICP	3948.84
GATSNRPCKIogrKCFOHqKDCCGRACIITICP	3948.84
GATSDROCKIogrKCFOHQKDCCGRACIITICP*	3963.83
GATSNROCKIogrKCFOHqKDCCGRACIITICP*	3963.85
GATSNROCKIogrKCFOHQKDCCGRACIITICP	3963.85
GATSDROCKIogrKCFOHqKDCCGRACIITICP	3965.82
STTKGATSNROCKIogrKCFPHqKDCCGRACIITICP*	4365.07
STTKGATSDROCKIogrKCFPHQKDCCGRACIITICP*	4365.07
STTKGATSDRPCKIogrKCFOHQKDCCGRACIITICP*	4365.07
STTKGATSNROCKIogrKCFPHQKDCCGRACIITICP	4365.07
STTKGATSNROCKIogrKCFOHqKDCCGRACIITICP*	4381.06
STTKGATSDROCKIogrKCFOHQKDCCGRACIITICP*	4381.07

**Table 38- κ-PVIIA toxiforms**

C	R	I	P	N	Q	K	C	F	Q	H	L	D	D	C	C	S	R	K	C	N	R	F	N	K	C	V
			O	D	E				E											D		D				*

Annotated Sequence	MH+ [Da]
CRIPNQKCFQHLDDCCSRKCNRFNKC	3600.58
CRIPNQKCFQHLDDCCSRKCDRFNKC	3601.57
CRIPNQKCFQHLDDCCSRKCNRFDKC	3601.57
CRIPNQKCFQHLDDCCSRKCDRFDKC*	3601.57
CRIPNQKCFEHLDDCCSRKCNRFNKC	3601.58

CRIPDQKCFQHLDDCCSRKCDRFNKCV	3602.57
CRIONQKCFQHLDDCCSRKCDRFNKCV*	3616.60
CRIONQKCFQHLDDCCSRKCNRFDKCV*	3616.60
CRIONQKCFQHLDDCCSRKCNRFNKCV	3616.60
CRIODQKCFQHLDDCCSRKCDRFNKCV*	3617.59
CRIODQKCFQHLDDCCSRKCNRFNKCV	3617.59
CRIONQKCFQHLDDCCSRKCDRFDKCV*	3617.60
CRIONQKCFQHLDDCCSRKCDRFNKCV	3617.60
CRIONQKCFQHLDDCCSRKCNRFDKCV	3617.60
CRIONEKCFQHLDDCCSRKCNRFNKCV	3617.60
CRIONQKCFEHLDDCCSRKCNRFNKCV	3617.61
CRIONEKCFQHLDDCCSRKCNRFDKCV*	3617.61
CRIONQKCFQHLDDCCSRKCDRFDKCV	3618.59

**Table 39- Contryphan-P toxiforms**

Contryphan-P	G C P W D P W C
	O *

Annotated Sequence	Thyo. MH+ [Da]
GCOWDPWC*	1092.403

**Table 40- PIIA toxiforms**

PIIA	C C C I R S D G P K C S R K C L S S F F C

Annotated Sequence	MH+ [Da]
CCIRSDGPKCSRKCLSSFFC	2525.10
CCCIRSDGPKCSRKCLSSFFC	2685.13

**Table 41- PVIIIA toxiforms**

	G C S G S P C F K N K T C R D E C I C G G L S N C W
	O D Y D

	G Y G G S R C G C K C T C R E
--	-------------------------------

Annotated Sequence	MH+ [Da]
CSGSO <del>C</del> FKNKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4857.92
CSGSO <del>C</del> FKNKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4901.90
CSGSP <del>C</del> FKNKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4841.91
CSGSP <del>C</del> FKDKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4842.92
CSGSP <del>C</del> FKNKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4842.89
CSGSO <del>C</del> FKNKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4858.90
CSGSO <del>C</del> FKDKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4858.90
CSGSO <del>C</del> FKDKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4859.92
CSGSO <del>C</del> FKDKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4902.93
CSGSO <del>C</del> FKNKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4902.93
CSGSO <del>C</del> FKDKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4903.91
CSGSP <del>C</del> FKNKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4885.87
CSGSP <del>C</del> FKDKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4843.91
GCSGSO <del>C</del> FKNKTCRDEICGGLSDCWCGYGGSRGCKCTC	4630.80
GCSGSO <del>C</del> FKDKTCRDEICGGLSNCWCGYGGSRGCKCTC	4630.78
GCSGSO <del>C</del> FKDKTCRDyCICGGLSDCWCGYGGSRGCKCTC	4675.78
GCSGSO <del>C</del> FKDKTCRDEICGGLSDCWCGYGGSRGCKCTC	4631.79
GCSGSP <del>C</del> FKNKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4898.92
GCSGSP <del>C</del> FKDKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4900.92
GCSGSO <del>C</del> FKNKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4914.93
GCSGSO <del>C</del> FKNKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4915.94
GCSGSO <del>C</del> FKDKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4915.93
GCSGSP <del>C</del> FKNKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4899.93
GCSGSO <del>C</del> FKDKTCRDEICGGLSDCWCGYGGSRGCKCTCRE	4916.93
GCSGSP <del>C</del> FKNKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4942.91
GCSGSO <del>C</del> FKDKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4959.92
GCSGSO <del>C</del> FKNKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4959.91
GCSGSO <del>C</del> FKNKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4958.93
GCSGSP <del>C</del> FKDKTCRDEICGGLSNCWCGYGGSRGCKCTCRE	4899.93
GCSGSO <del>C</del> FKDKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4960.91
GCSGSP <del>C</del> FKDKTCRDyCICGGLSNCWCGYGGSRGCKCTCRE	4943.91
GCSGSP <del>C</del> FKDKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4944.92
GCSGSP <del>C</del> FKNKTCRDyCICGGLSDCWCGYGGSRGCKCTCRE	4943.92

**Table 42- PVA toxiforms**

PVA	G	C	C	P	K	Q	M	R	C	C	T	L
												*

Annotated Sequence	MH+ [Da]
GCCPKQMRCTL	1570.64
GCCPKQMRCTL*	1569.66

**Table 43- PVB toxiforms**

PVB	R	D	C	C	P	E	K	M	W	C	C	P	L
					0			0				0	*

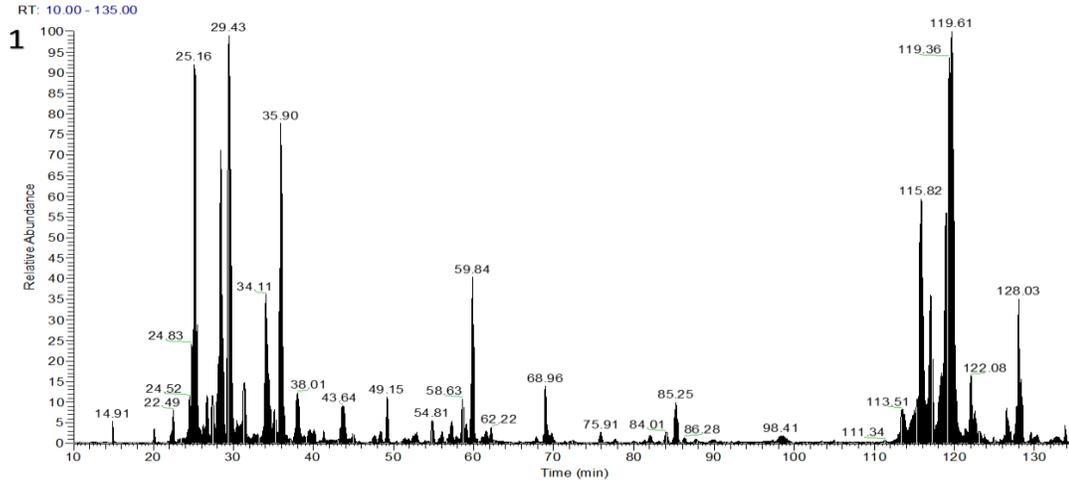
Annotated Sequence	MH+ [Da]
DCCPEKMWCCP	1542.53
DCCOEKmwCCP	1574.52
DCCOEKMWCCP	1558.52
DCCPEKmwCCP	1558.52
DCCOEKMWCCPL*	1670.62
DCCPEKmwCCPL*	1670.62
DCCPEKMWCCPL*	1654.63
DCCOEKMWCCOL*	1686.62
DCCOEKmwCCPL*	1686.62
RDCCPEKMWCCP	1698.63
RDCCPEKMWCCPL*	1810.73

**Table 44- p21b sites of modification on the two identified peptides.** \*\*There were no spectral matches to the full PVIC mature peptide. Sites of modification were determined by matches to peptide fragments

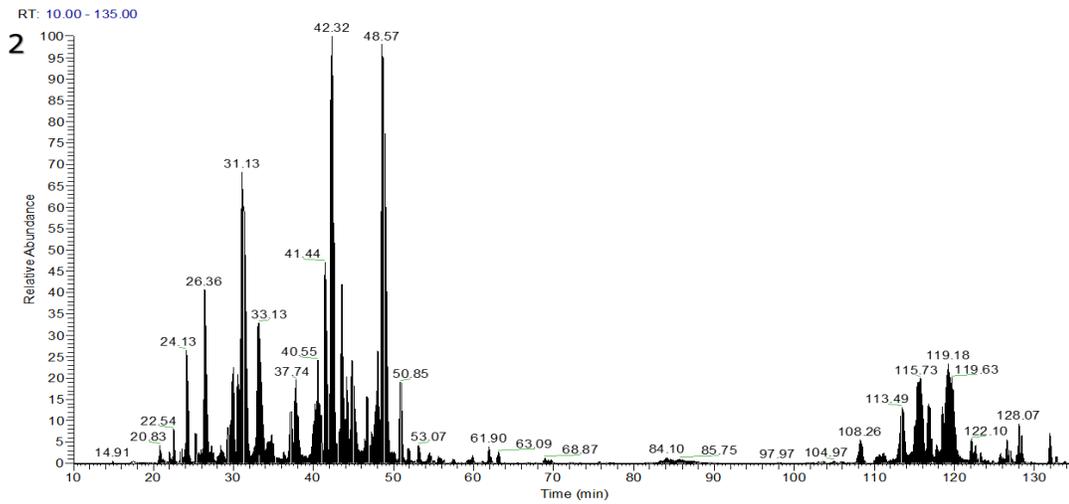
A	F	E	L	L	P	S	Q	D	R	S	C	C	I	R	K	T	L	E	C	L	E	N	Y	P	G	Q	E	S	Q	R	A	H	Y
						0		E																0				E					*

B	S	I	N	A	Q	N	N	V	R	P	A	H	D	T	C	I	N	R	L	C	F	D	P	G	F	
															0											

**APPENDIX B: Chromatograms (TIC) of injected venom samples from *C. purpurascens* specimens.**



**Figure 26- Total ion chromatogram from venom sample of specimen 1**



**Figure 27- Total ion chromatogram from venom sample of specimen 2**

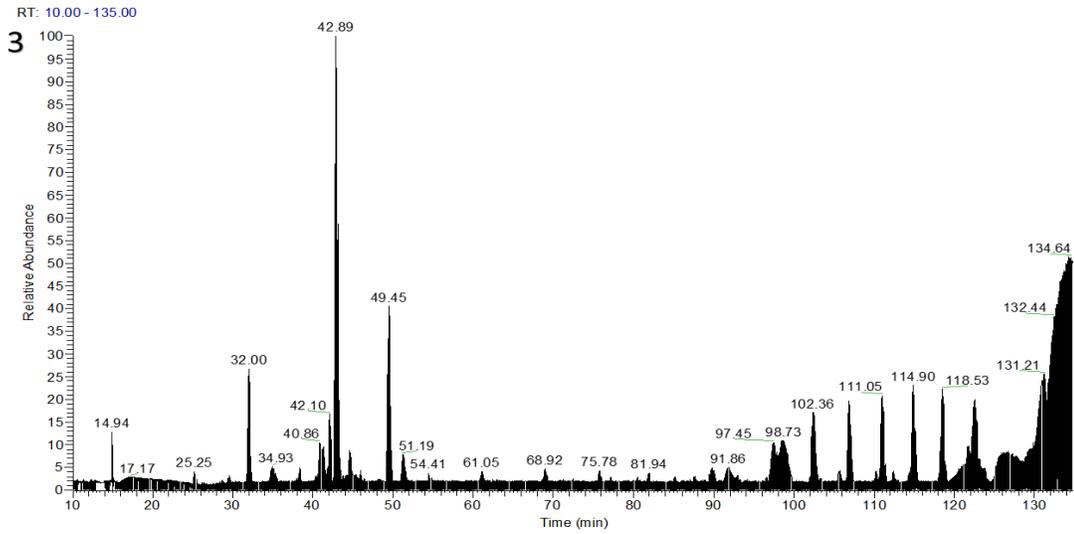


Figure 28- Total ion chromatogram from venom sample of specimen 3

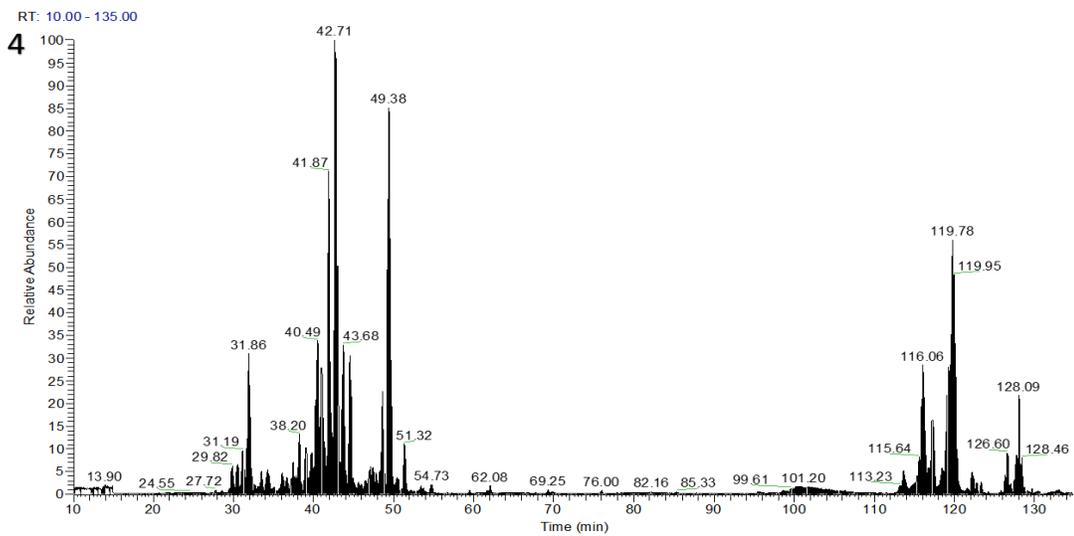


Figure 29- Total ion chromatogram from venom sample of specimen 4

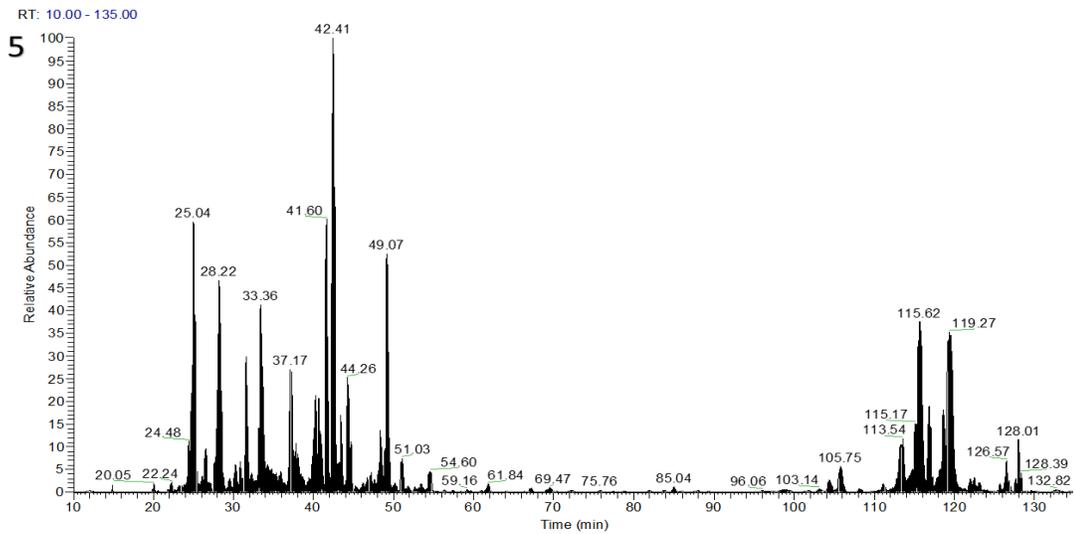


Figure 30- Total ion chromatogram from venom sample of specimen 5

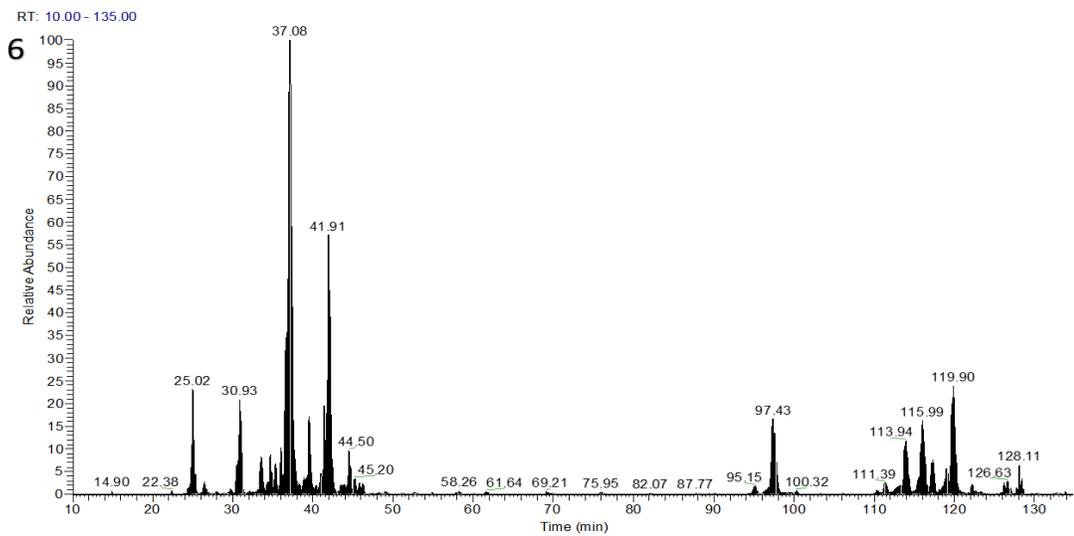
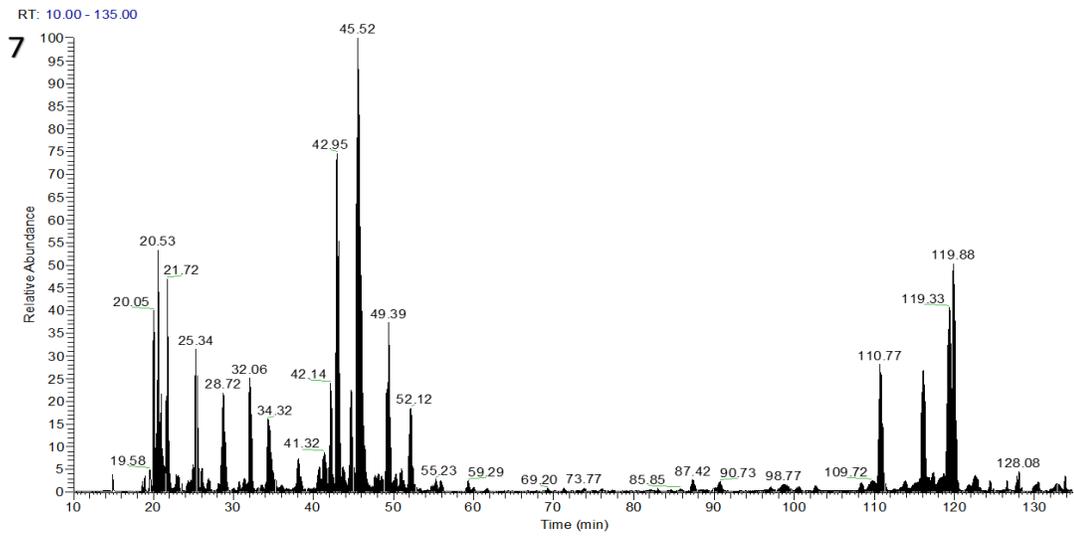
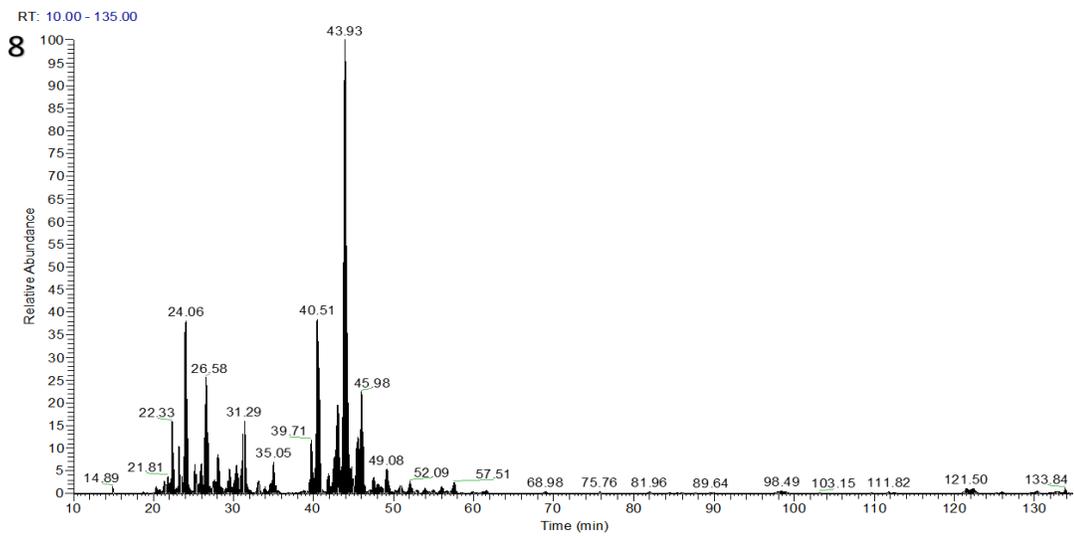


Figure 31- Total ion chromatogram from venom sample of specimen 6



**Figure 32- Total ion chromatogram from venom sample of specimen 7**



**Figure 33- Total ion chromatogram from venom sample of specimen 8**

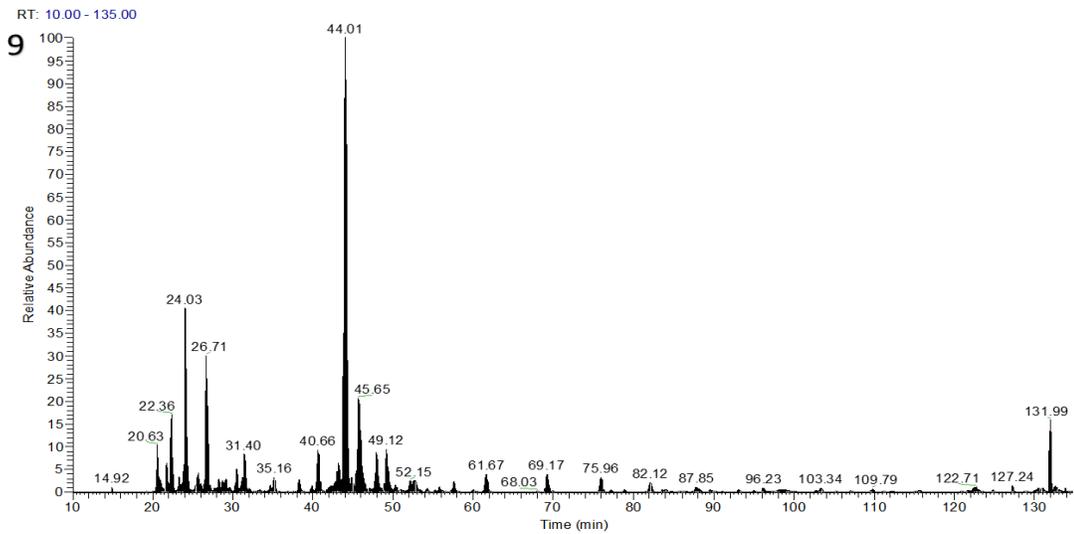


Figure 34- Total ion chromatogram from venom sample of specimen 9

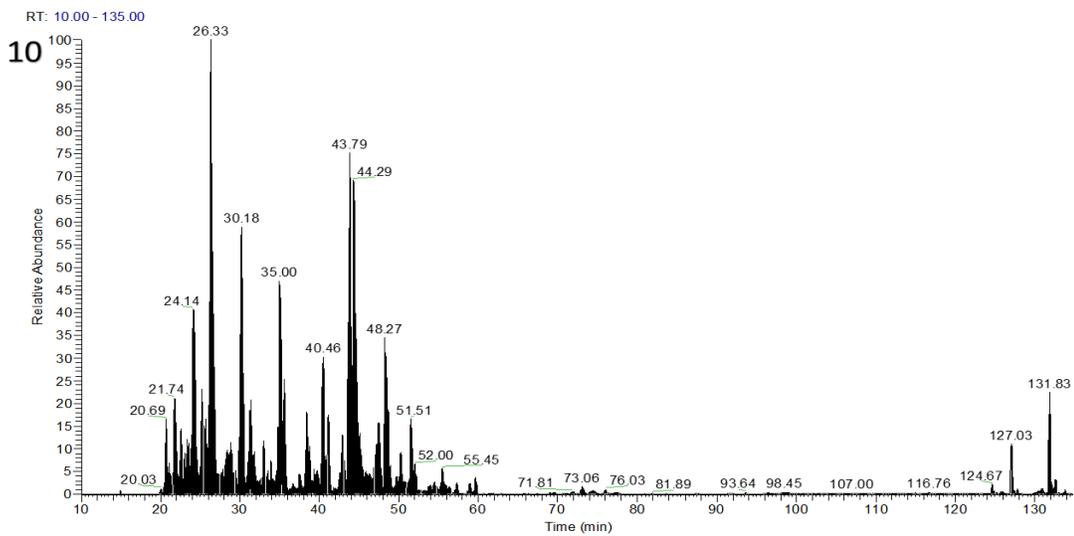
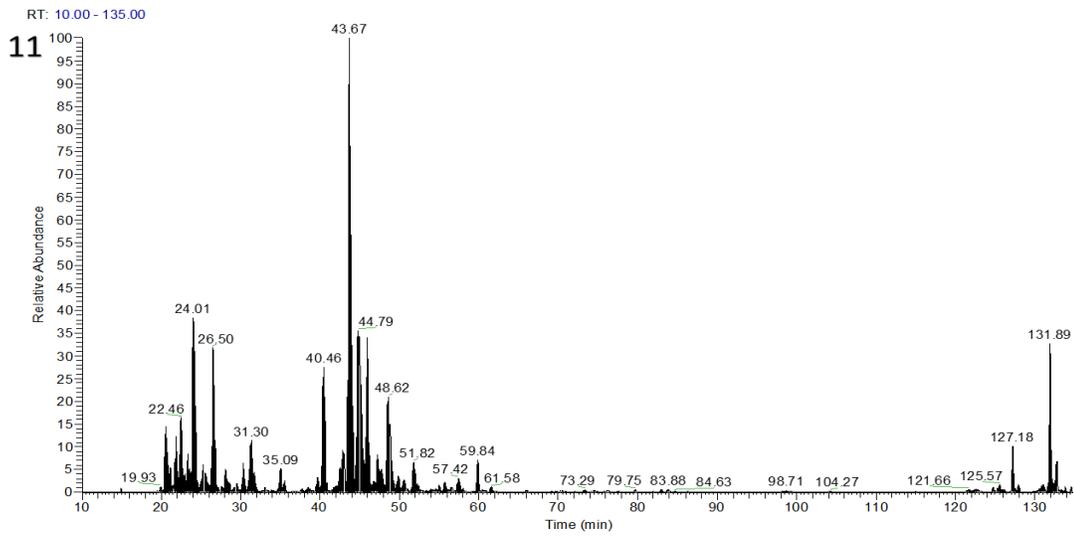
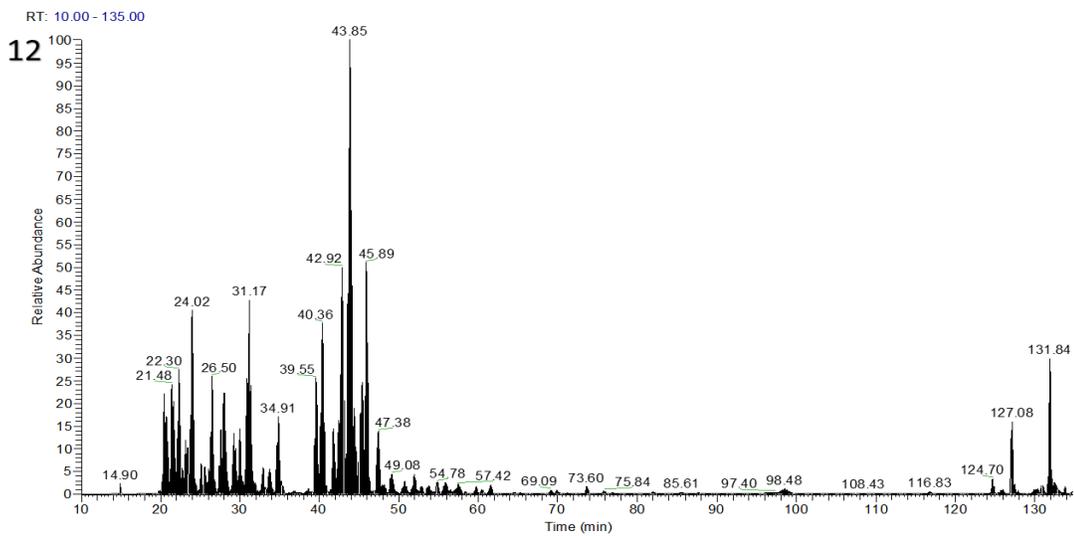


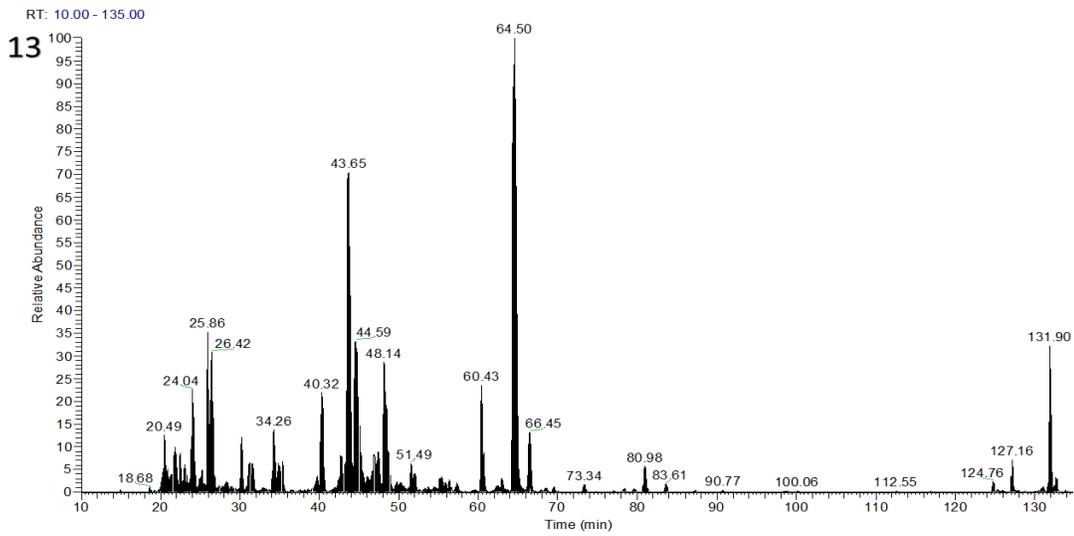
Figure 35- Total ion chromatogram from venom sample of specimen 10



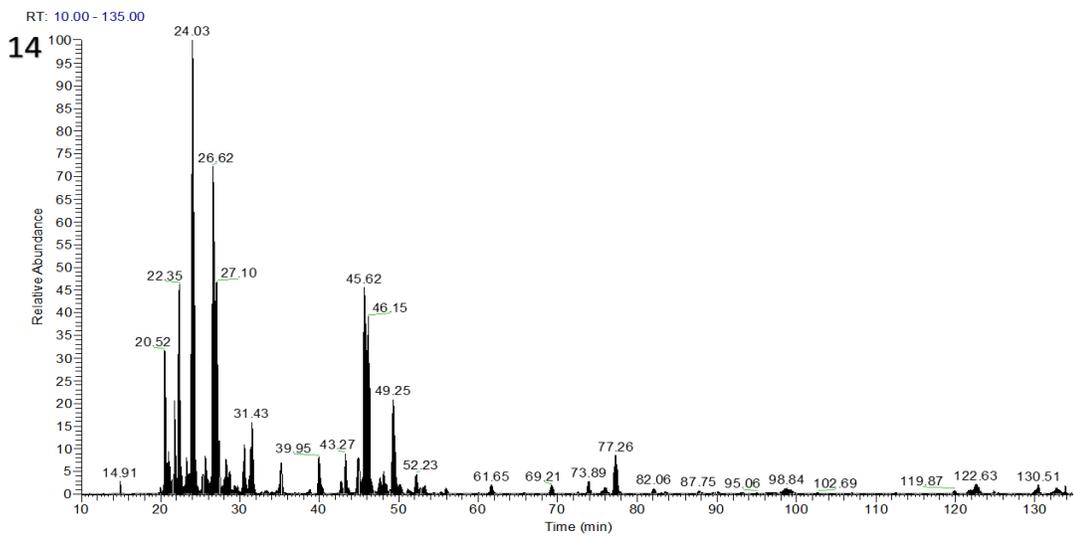
**Figure 36-** Total ion chromatogram from venom sample of specimen 11



**Figure 37-** Total ion chromatogram from venom sample of specimen 12



**Figure 38-** Total ion chromatogram from venom sample of specimen 13



**Figure 39-** Total ion chromatogram from venom sample of specimen 14

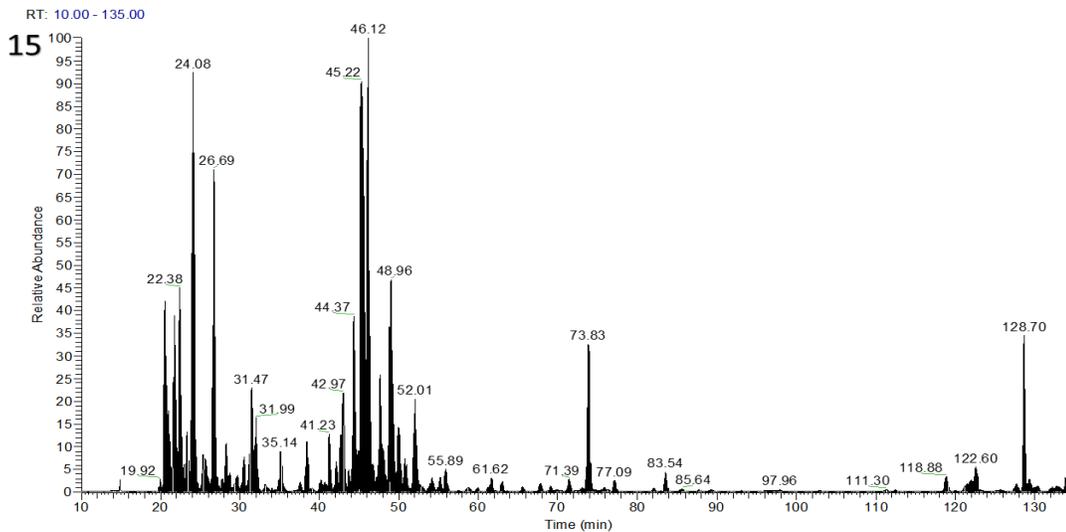


Figure 40- Total ion chromatogram from venom sample of specimen 15

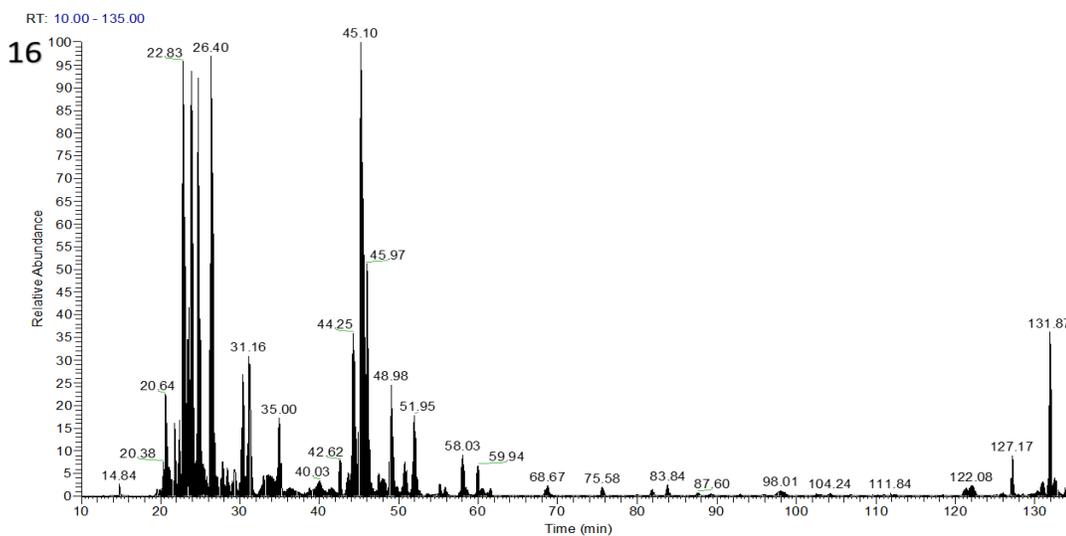
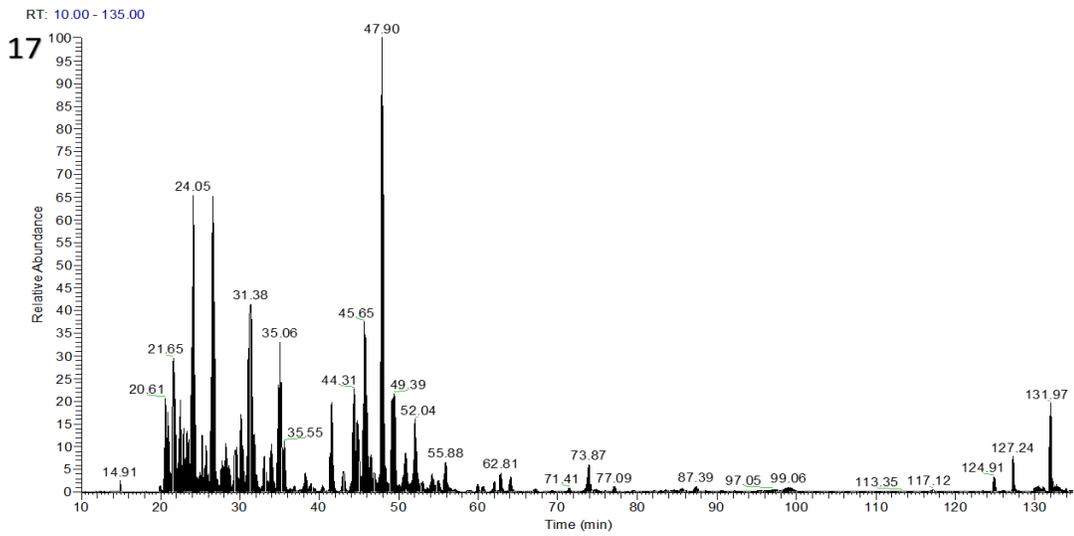
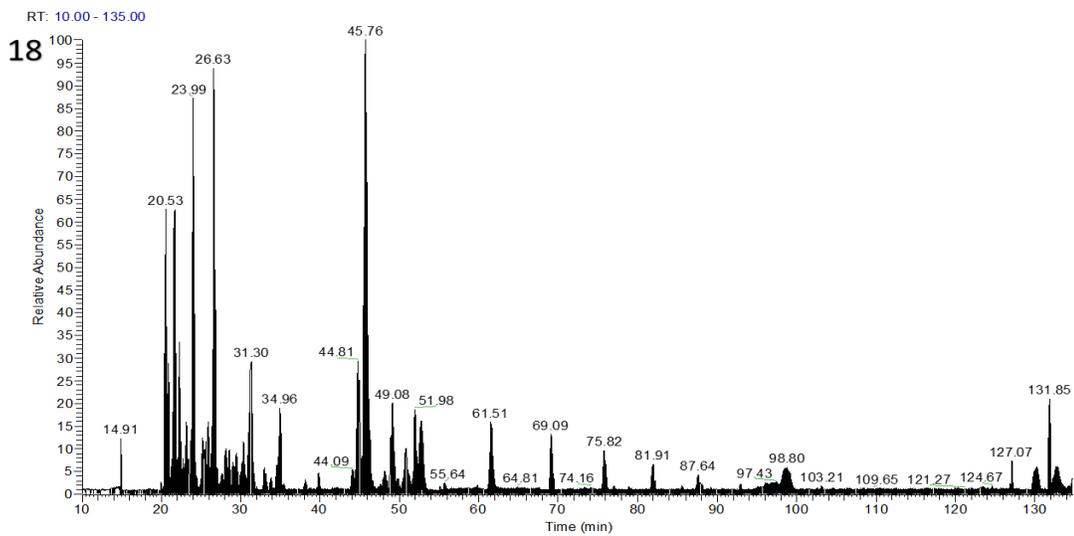


Figure 41- Total ion chromatogram from venom sample of specimen 1



**Figure 42- Total ion chromatogram from venom sample of specimen 17**



**Figure 43- Total ion chromatogram from venom sample of specimen 18**

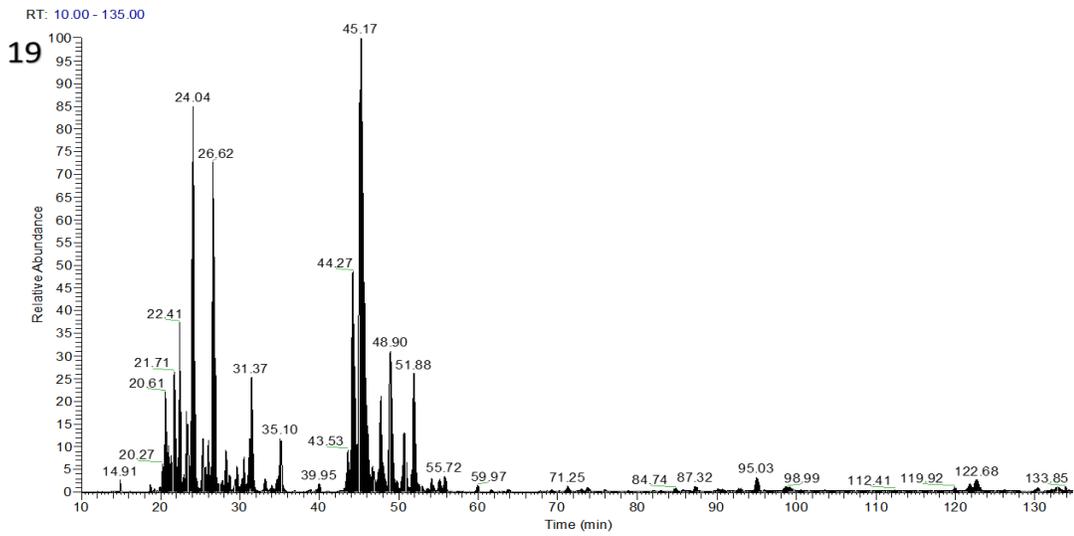


Figure 44- Total ion chromatogram from venom sample of specimen 19

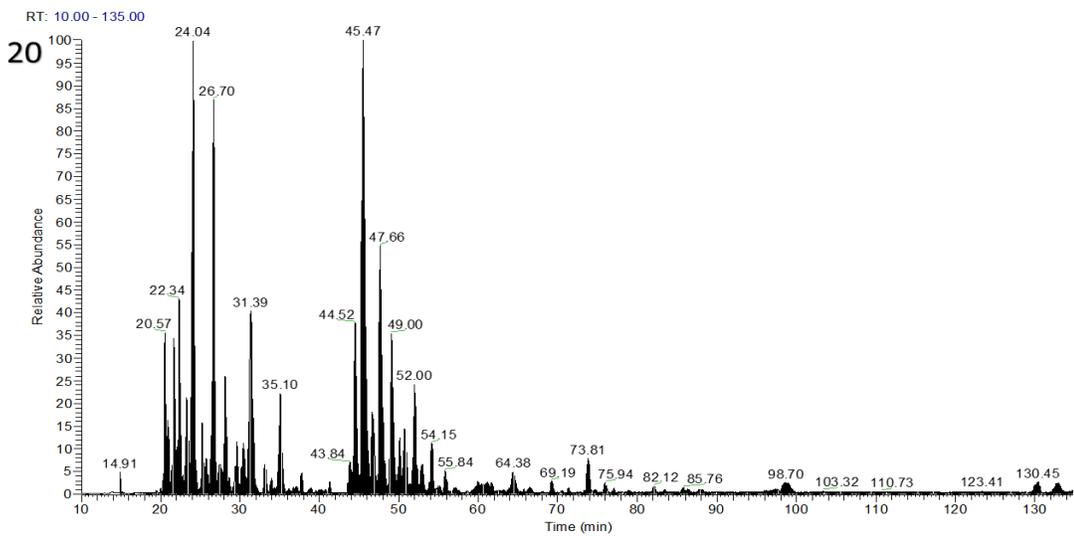


Figure 45- Total ion chromatogram from venom sample of specimen 20

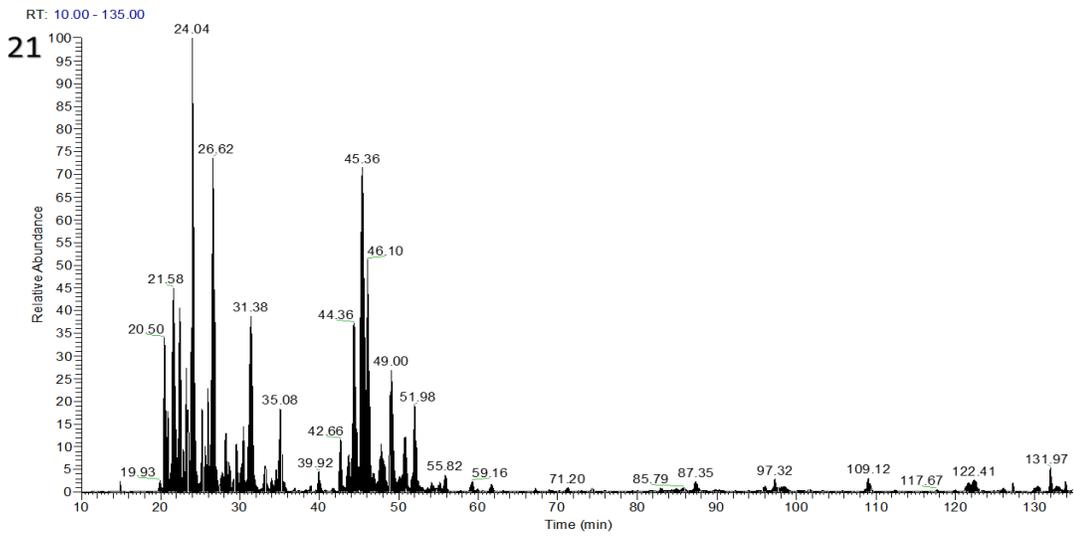


Figure 46- Total ion chromatogram from venom sample of specimen 21

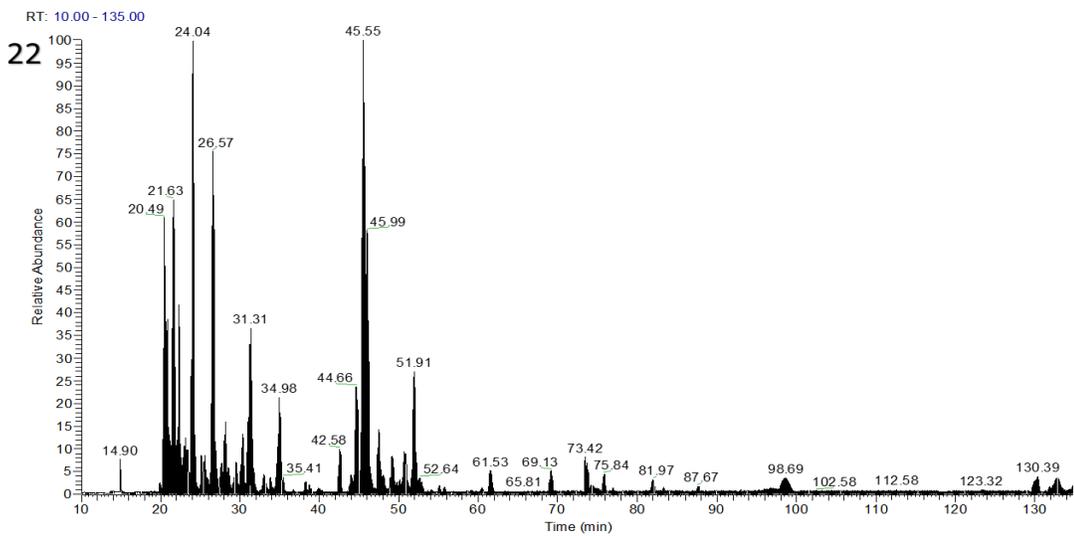
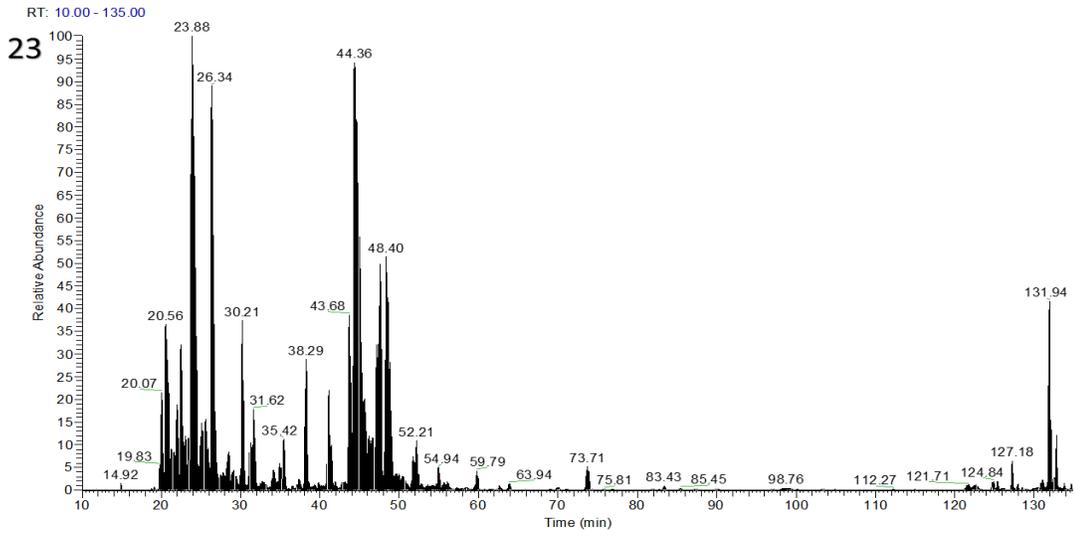
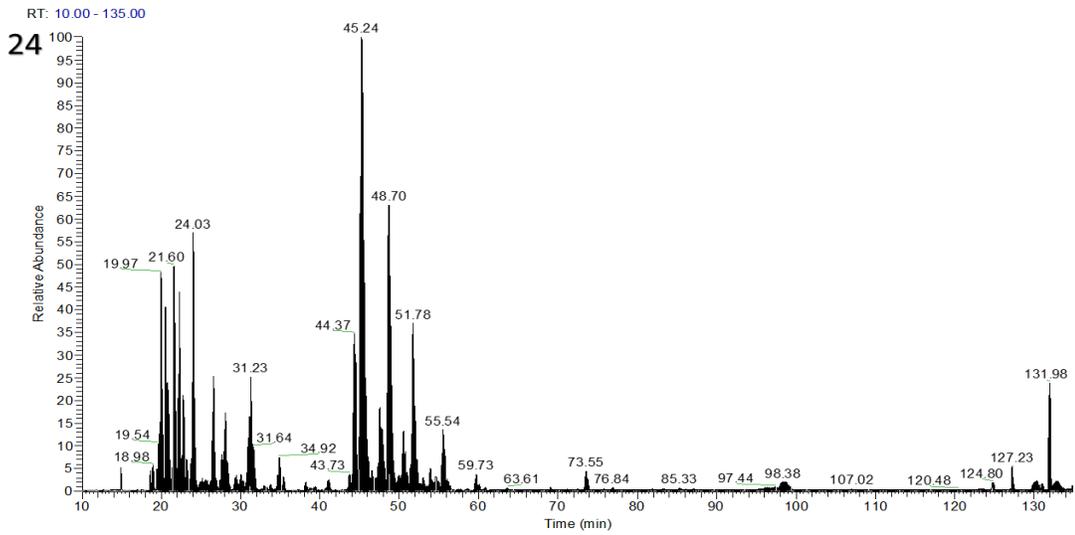


Figure 47- Total ion chromatogram from venom sample of specimen 22



**Figure 48-** Total ion chromatogram from venom sample of specimen 23



**Figure 49-** Total ion chromatogram from venom sample of specimen 24

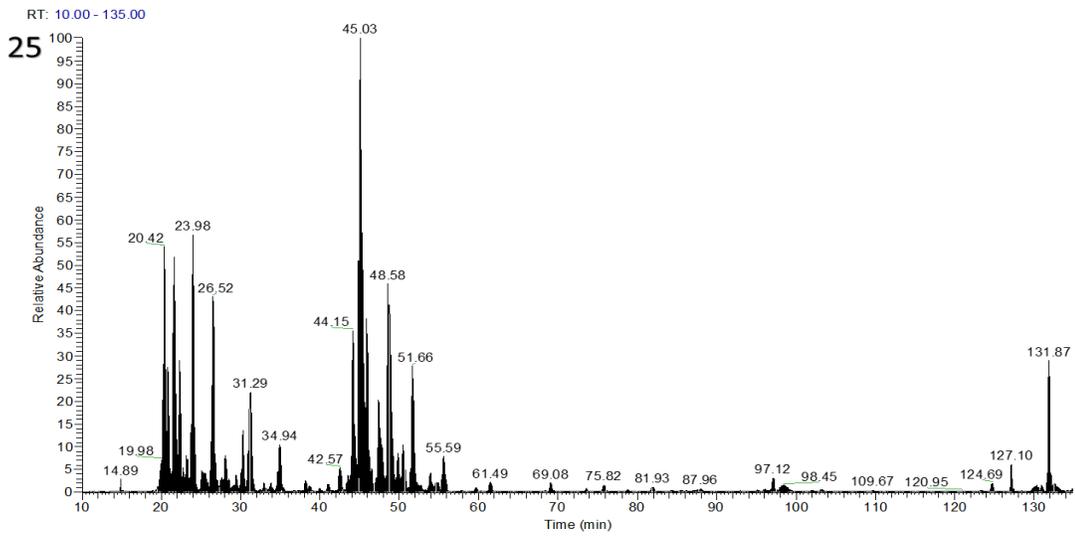


Figure 50- Total ion chromatogram from venom sample of specimen 25

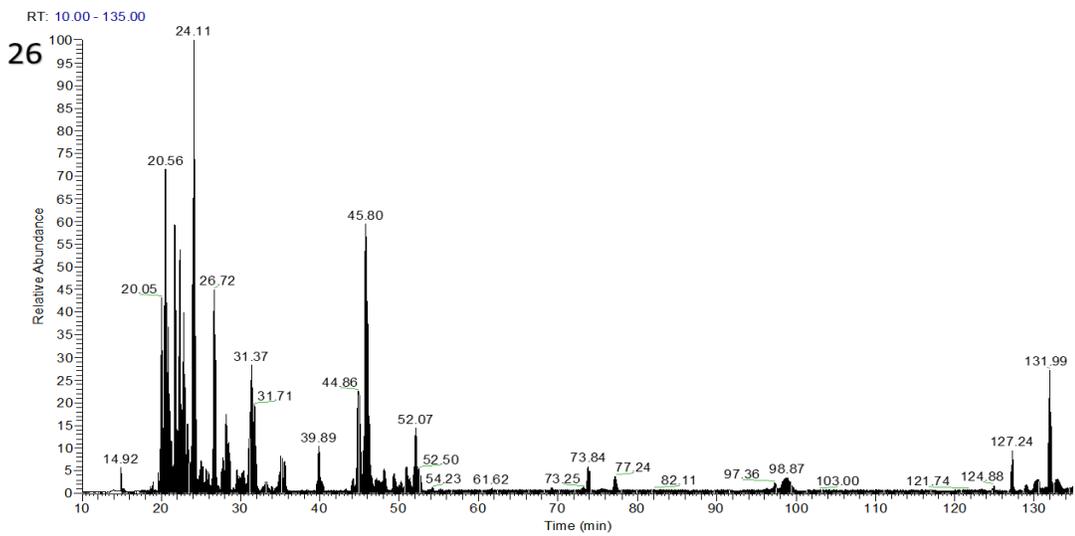


Figure 51- Total ion chromatogram from venom sample of specimen 26

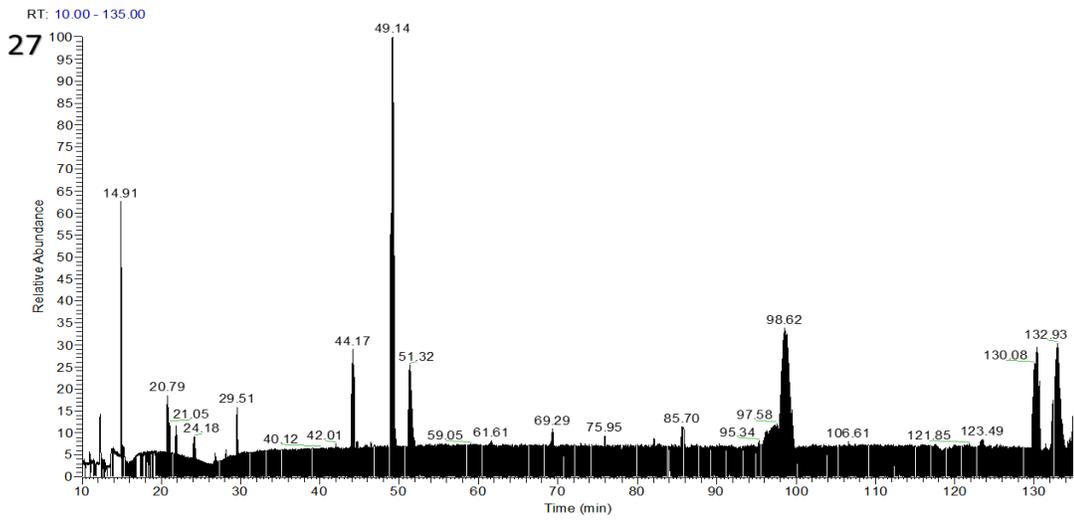


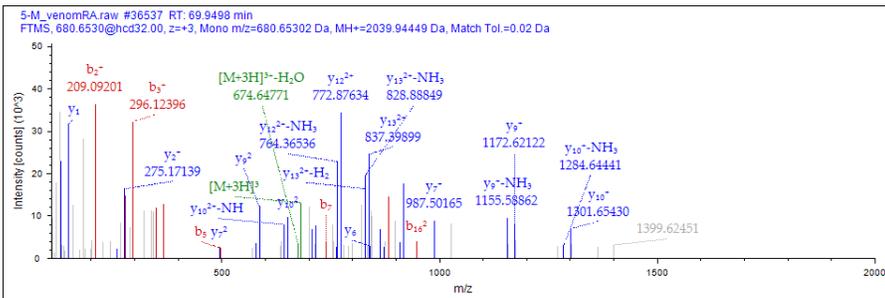
Figure 52- Total ion chromatogram from venom sample of specimen 27

## APPENDIX C: Features of the New Conopeptides in *Conus purpurascens*

Here we describe the main features of newly discovered conopeptides in the injected venom of *C. purpurascens*. We include the sequence of the precursor proteins from the transcriptome and highlight the corresponding signal and mature sequences. We also show the annotated MS/MS spectra. When possible, we compare the novel *C. purpurascens* sequences to known conopeptides, which can confer putative structural and functional characteristics to these newly described peptides.

### Linear

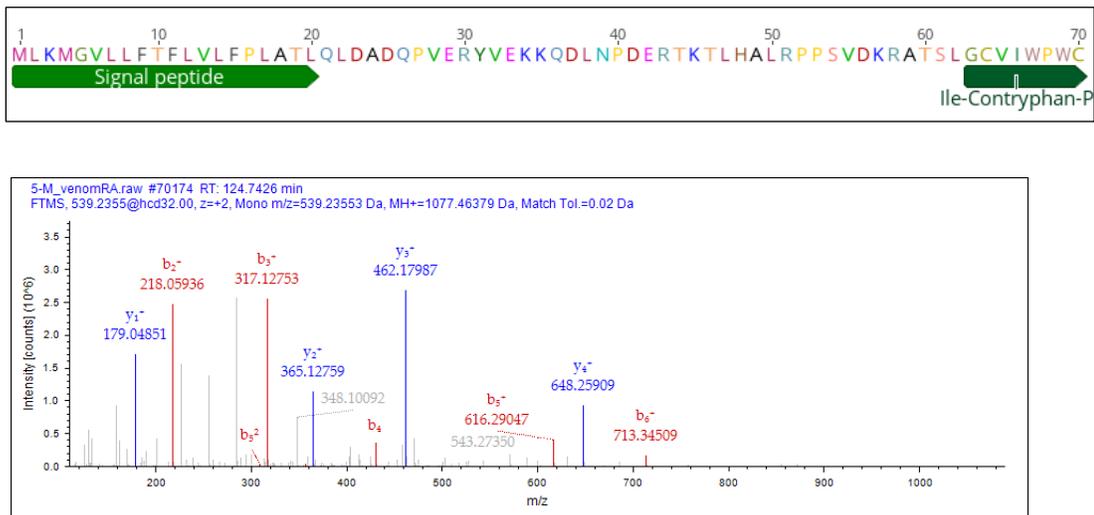
**Linear-P** belongs to the B2 Superfamily. This family of linear peptides is expressed in other *Conus* species based on sequences deposited in NCBI. A similar sequence, differing in a single residue, is expressed by *C. ermineus* (Sequence ID: AXL95472) [197], a close relative to *C. purpurascens*. This is the first evidence of these B2 linear peptides in injected venom.



**Figure 53- Supplementary information for Linear-P.** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

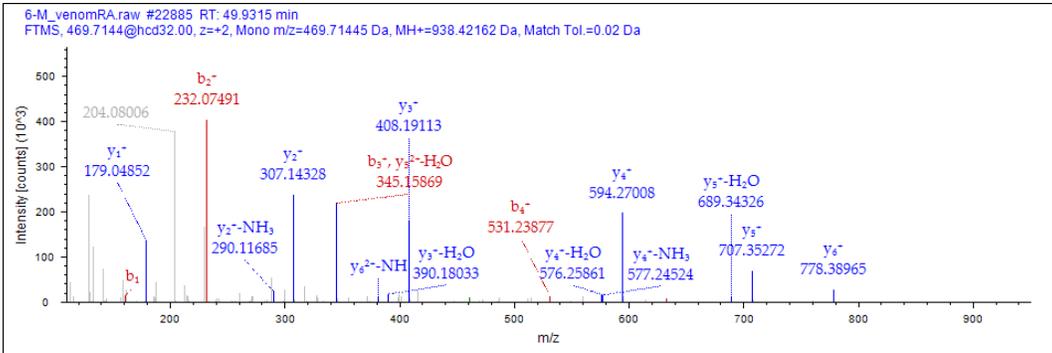
## One disulfide (C-C)

**Ile-contryphan-P** belongs to the M superfamily. Ile-contryphan-P exhibits sequence homology to previously described leu-contryphan-P, except for a switch from L5 to W5. In general, contryphans are characterized by a conserved motif containing *D*-tryptophan or leucine and a single disulfide bond. Contryphans typically classify as part of the O2 superfamily based on their signal sequence, however Ile-contryphan-P does not follow this trend. The molecular target of contryphans remain unclear. Here, Ile-contryphan-P was identified in 25 of the 27 venom samples, suggesting it has an important function that needs to be discerned.



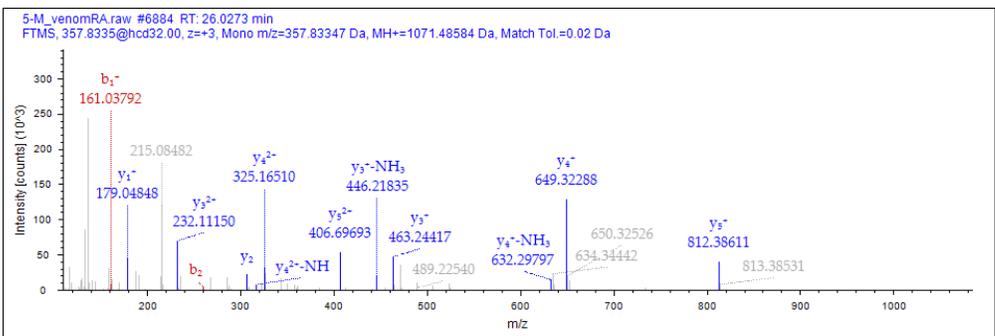
**Figure 54- Supplementary information for Ile-Contryphan-P.** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**Contryphan-P3** belongs to the M superfamily. Based on precursor analysis, these peptides form a new group of one disulfide peptides within the M superfamily [98]. Contryphan-P3 is also expressed by *C. ermineus* (Sequence ID: AXL95407) [197], but this is the first instance in venom.



**Figure 54- Supplementary information for contryphan-P 1)** full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

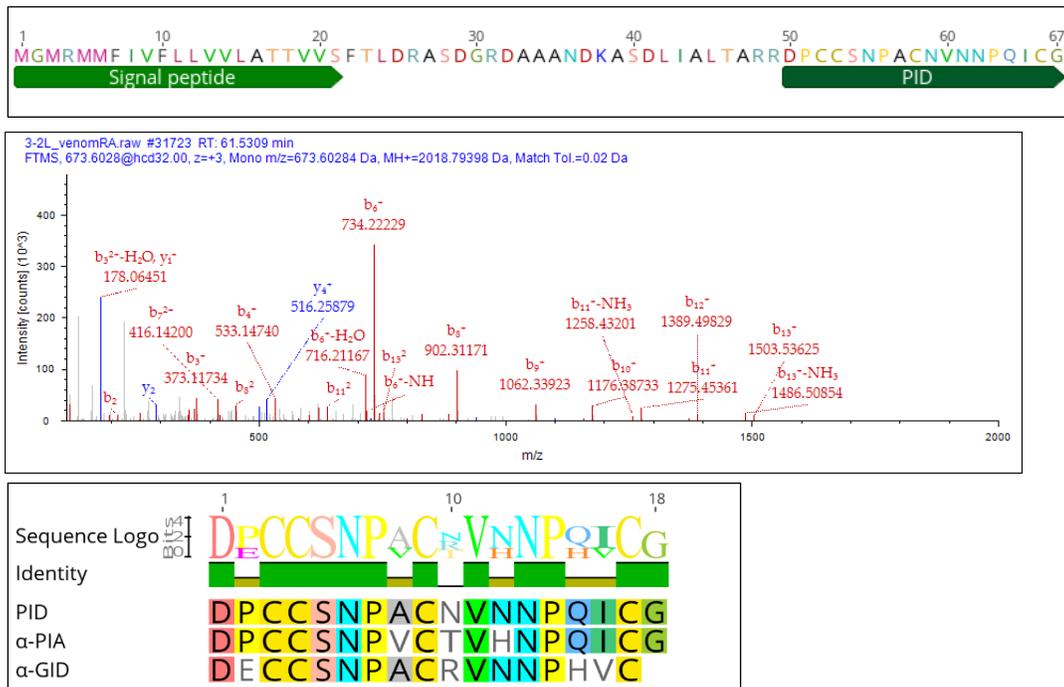
**Contryphan-P4** is not expressed in the venom duct, it was first discovered by *de novo* sequencing using PEAKS software (Unpublished data from our lab). Like Contryphan P3, the same sequence is found in *C. ermineus* venom duct transcriptome and belongs to the M superfamily (Sequence ID: AXL95569). Contryphan-P3 and P4 have different expression patterns and cluster into separate cabals.



**Figure 54- Supplementary information for Contryphan-P3 1)** mature peptide sequence as determined by *de novo* sequencing and 2) annotated MSMS spectra

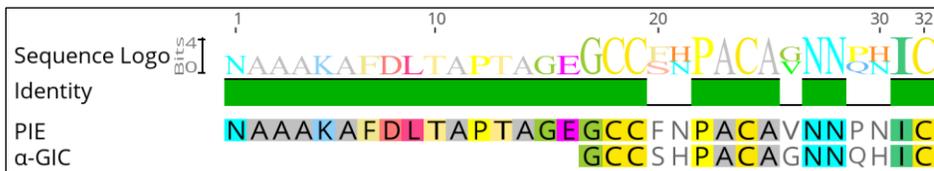
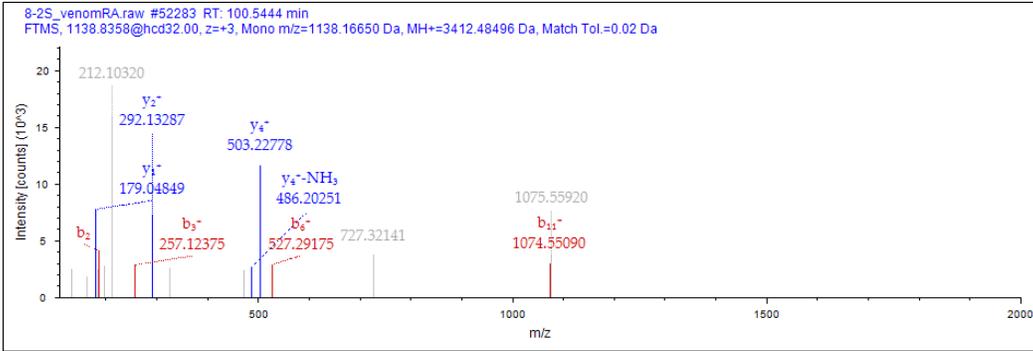
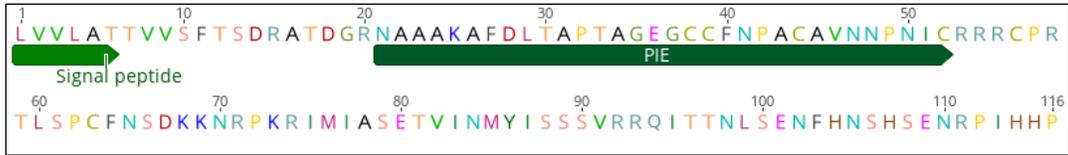
## Framework I (CC-C-C)

**PID** belongs to superfamily A. The sequence of PID has been previously reported as a nucleic acid sequence from *C. ermineus*, E1.1 precursor (P03002, Conoserver). Here we provide the first evidence for the mature peptide in milked venom from *C. purpurascens*. The mature peptide exhibits homology to  $\alpha$ -PIA (*C. purpurascens*) and to  $\alpha$ -GID (*C. geographus*, P60274), suggesting PID will inhibit the nAChR [161, 198].



**Figure 55- Supplementary information for PID** 1) full transcript with annotated signal sequence and mature peptide regions, 2) annotated MS/MS spectra and 3) sequence alignment with  $\alpha$ -PID and  $\alpha$ -GID.

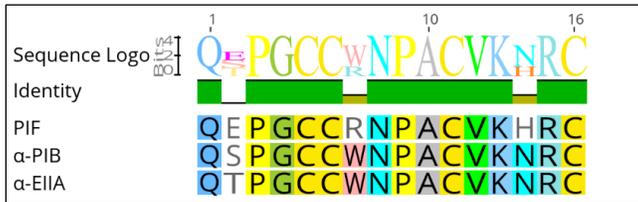
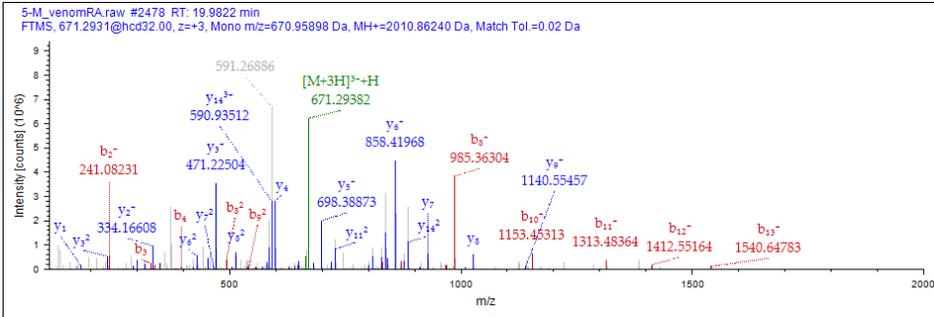
**PIE** belongs to superfamily A. Homology to  $\alpha$ -conotoxin precursor Bt1.8 from *C. betulinus* (Sequence ID: A0A068B6Q6) and characterized  $\alpha$ -GIC (Sequence ID: Q86RB2) that inhibits  $\alpha 3\beta 2$  receptors [199]. Interestingly, MS/MS data from this study identified PIE with an extended N-terminal tail lacked by characterized  $\alpha$ -conotoxins.



**Figure 56- Supplementary information for PIE** 1) full transcript with annotated signal sequence and mature peptide regions, 2) annotated MSMS spectra, and 3) sequence alignment with  $\alpha$ -GIC

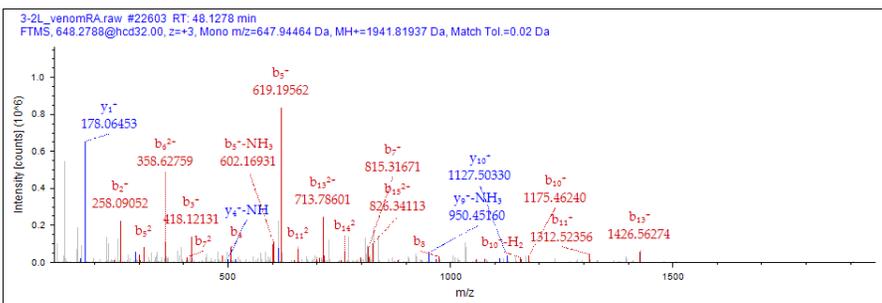
**PIF** is a new framework I conotoxin that was first sequenced *de novo* using PEAKS (Unpublished data). Because it was sequenced *de novo*, the superfamily is unable to be assigned. The mature peptide shows homology to EIIA from *C. ermineus* (Sequence ID: D4HRK4) which inhibits muscle subtype nAChRs [200]. PIF is also similar to  $\alpha$ -PIB, a muscle subtype-selective conotoxin from *C. purpurascens* [201].

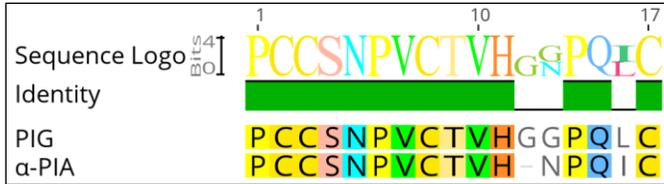




**Figure 57- Supplementary information for PIF** 1) mature peptide sequence as determined by *de novo* sequencing, 2) annotated MSMS spectra, and 3) sequence alignment with α-PIF and α-EIIA

**PIG** is a framework I conotoxin that was sequenced *de novo* using PEAKS software (Unpublished data). We lack signal sequence information, however the mature peptide sequence shows high homology (82%) to α-PIA that inhibits α6 nAChRs [161].

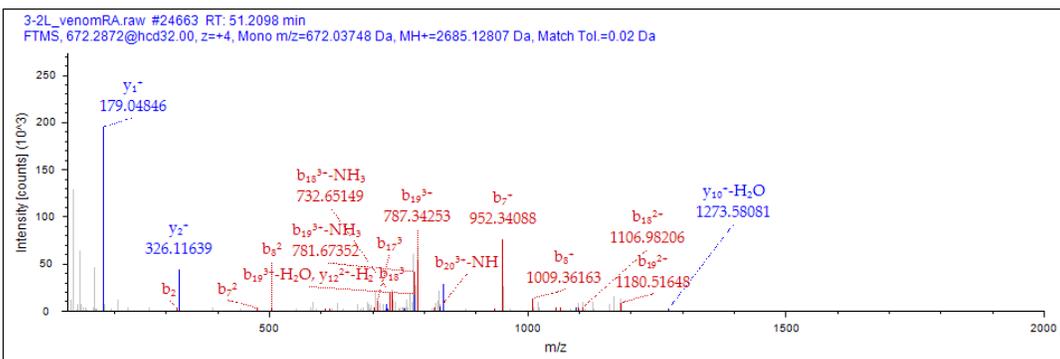




**Figure 58- Supplementary information for PIG** 1) mature peptide sequence as determined by *de novo* sequencing, 2) annotated MSMS spectra, and 3) alignment with  $\alpha$ -PIA

### Framework II (CCC-C-C-C)

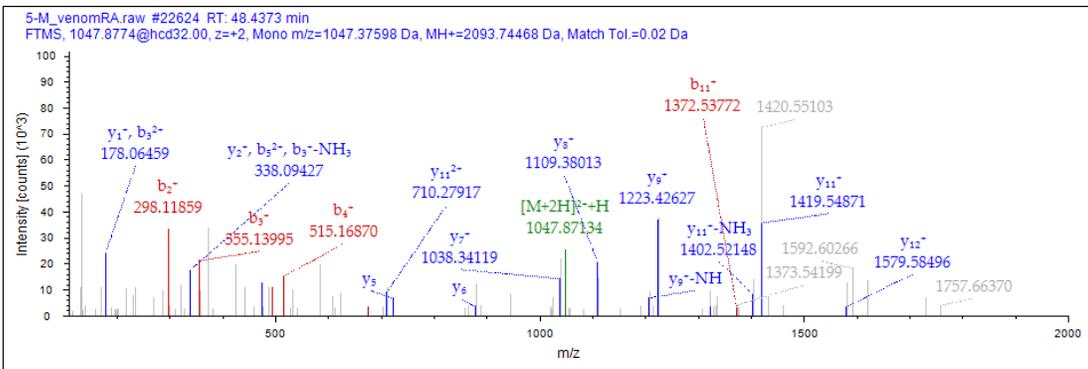
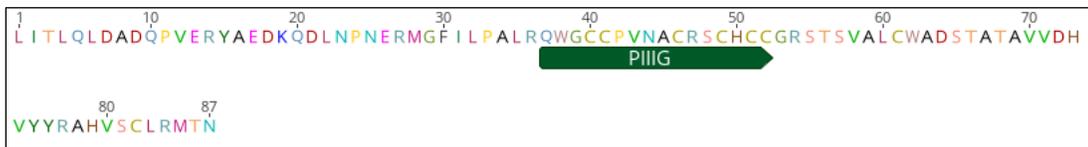
**PIIA** exhibits cysteine framework II and belongs to the O3 Superfamily. A very similar sequence is found in the *C. ermineus* venom duct transcriptome (Sequence ID: AXL95373) [197]. There are no similar characterized conotoxins from which we can infer activity. It is important to note that there is an extra cysteine pair upstream from the identified N-terminus, and it is possible that the mature peptide identified through our approaches was in fact a truncated version. If this extra cysteine pair is part of the mature peptide, it would form a novel four-disulfide peptide with a new cysteine framework (C-C-CCC-C-C-C).



**Figure 59- Supplementary information for PIIA** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**Framework III (CC-C-C-CC)**

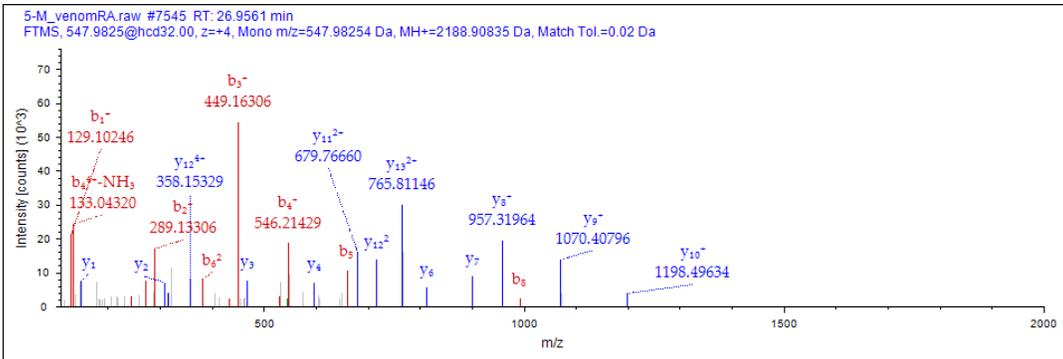
**PIIIG** belongs to the M superfamily that clusters with the motor cabal. It is an M1 mini-M with loops sizes 4/2/1 according to the number of residues between cysteine residues [125]. The mature peptide has very little sequence homology to other mini-Ms. The pharmacology of the mini-Ms remains unknown, despite their prevalence across *Conus* species [124].



**Figure 60- Supplementary information for PIIIG** 1) full transcript with annotated mature peptide region and 2) annotated MSMS spectra

**PIIIH** is an M superfamily conotoxin that clusters with the motor cabal. It is an M1 mini-M with loop sizes 4/5/1. PIIIH clusters with motor cabal toxins. It exhibits sequence homology

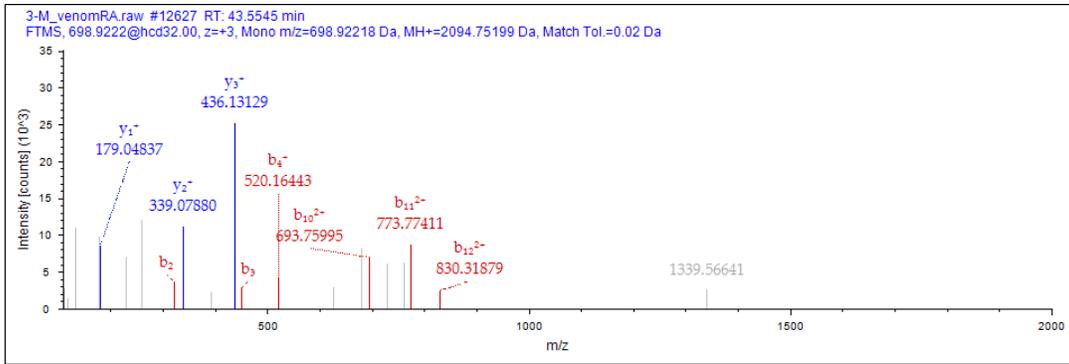
to a peptide expressed in the venom duct of fish-hunting cone snail *C. magus* (Sequence ID: QFQ61044), however there is no evidence of this peptide in the injected venom [202].



**Figure 61- Supplementary information for PIIIIH** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**PIIII** is a mini-M (M2) from the M Superfamily. It is the only one of the 3 identified venom mini-Ms that clustered with lightning-strike cabal peptides. Similar transcripts are found in *Turriconus* species (ATF27414, ATF27651) [203], and in *C. regius* (P85021) [124]. These are all worm-hunting species, supporting previous evidence that *C. purpurascens* may employ a mixed-mode feeding strategy (Unpublished data).

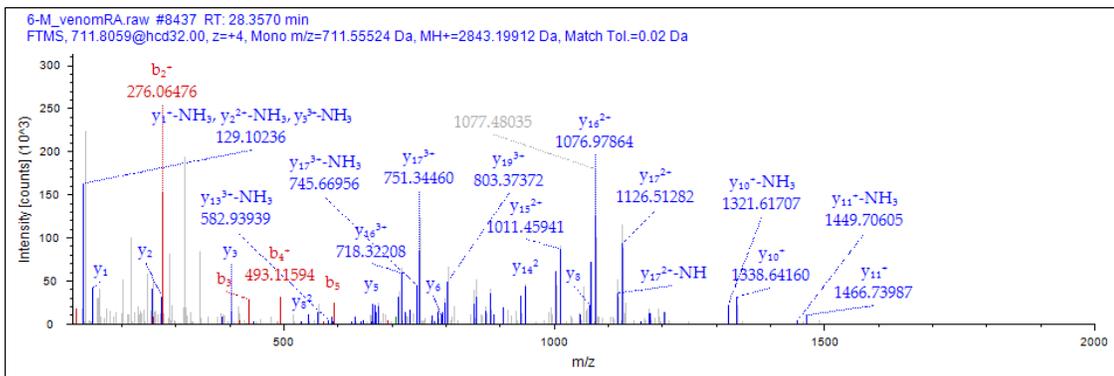


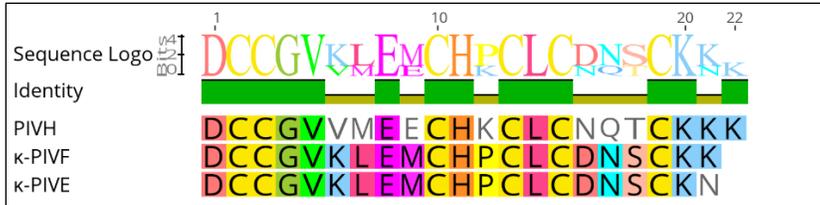


**Figure 62- Supplementary information for PIII 1)** full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**Framework IV (CC-C-C-C)**

**PIVH** belongs to the A superfamily. Its expression in the milked venom clusters closely with  $\kappa$ -PIVF and other lightning strike cabal toxins. PIVH shares ~60% homology with  $\kappa$ -PIVE and  $\kappa$ -PIVF, strongly indicating it also targets the potassium channel [204].

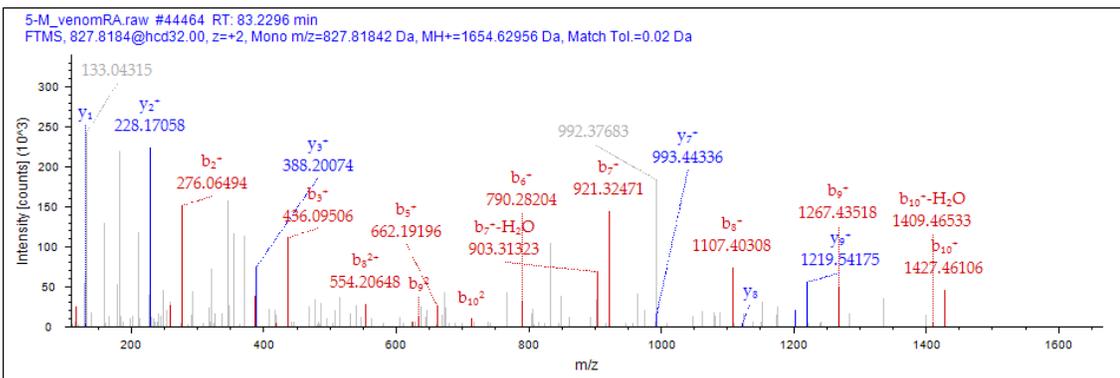
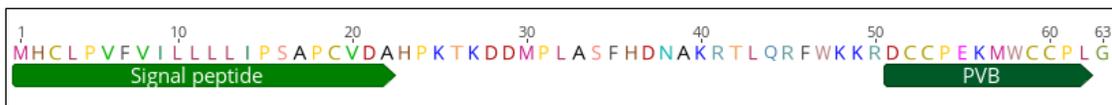




**Figure 63- Supplementary information for PIVH** 1) full transcript with annotated signal sequence and mature peptide regions 2) annotated MSMS spectra, and 3) alignment with  $\kappa$ -PIVF and  $\kappa$ -PIVE

### Framework V (CC-CC)

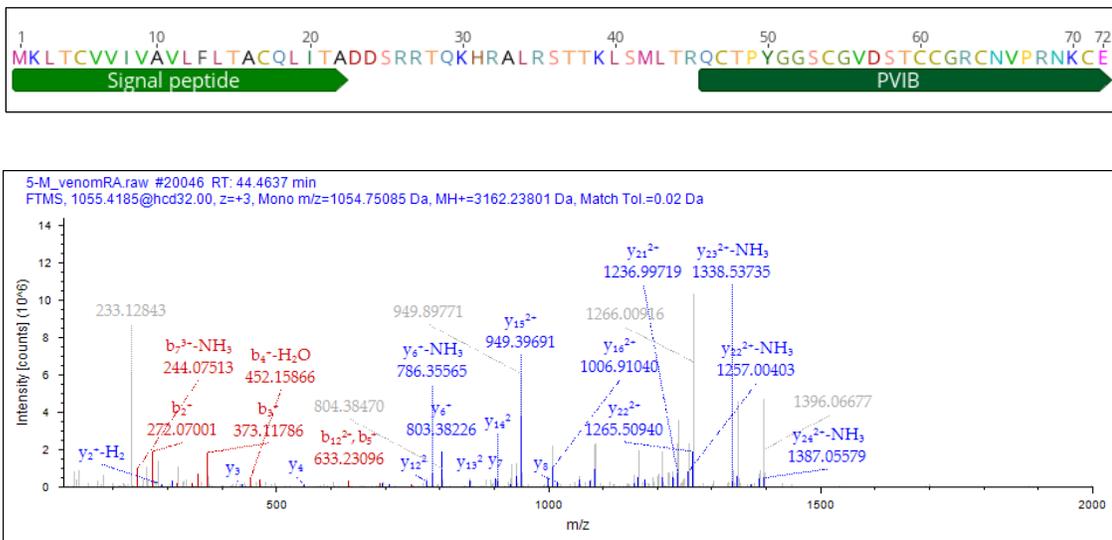
**PVB** is a T Superfamily conotoxin expressed in both transcriptomes. Its expression in the milked venom clustered closely with other motor cabal peptides. PVB is the second T Superfamily peptide identified from *C. purpurascens* venom but shows limited sequence homology to PVA aside from the conserved cysteine framework. The same sequence is found in the transcriptome of *C. ermineus* (Sequence ID: AXL95476) [197]. The pharmacology of T superfamily conotoxins is not well defined, however, two framework V conotoxins from the T Superfamily are known to target the somatostatin-3 receptor [205].



**Figure 64- Supplementary information for PVB** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**Framework VI/VII (C-C-CC-C-C)**

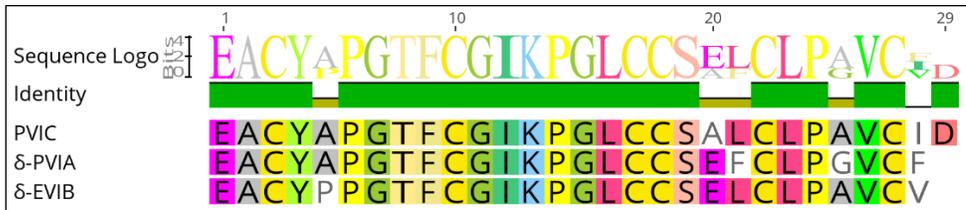
**PVIB** is an O1 Superfamily conotoxin. PVIB was identified in 21 of the 27 venom samples suggesting it plays an important role in prey capture. It falls within cluster 1 (lightning-strike cabal) and is expressed in the venom of all 7 specimens that form this cluster. PVIB has high homology to a sequence from the venom duct transcriptome of *C. ermineus* (Sequence ID: AXL95467) [197].



**Figure 64- Supplementary information for PVIB** 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**PVIC** belongs to the O1 Superfamily. It shows high homology (85%) to  $\delta$ -PVIA, and the two share similar patterns of expression in the venom. Its sequence is also similar to  $\delta$ -

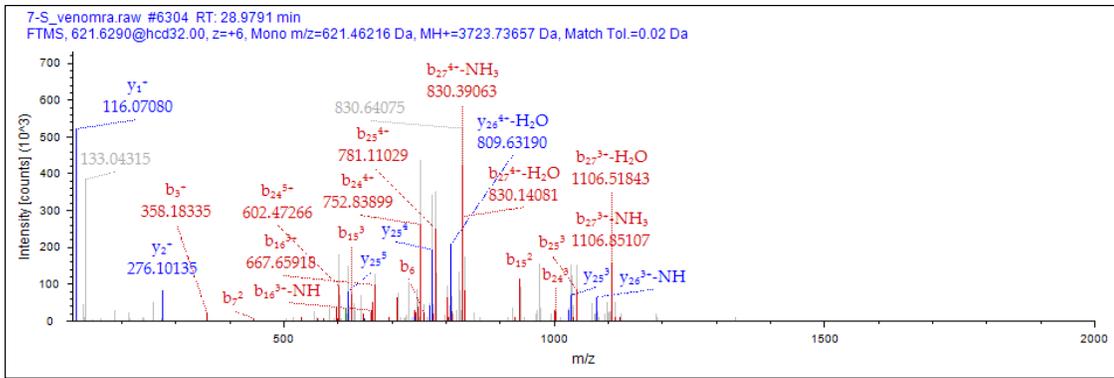
EVIB (Sequence ID: P69752). For this reason, it is likely PVIC will also target sodium channels [113] as part of the lightning-strike cabal.



**Figure 65- Supplementary information for PVIC** A) full transcript with annotated signal sequence and mature peptide regions and B) alignment with  $\delta$ -PVIA and  $\delta$ -EVIB

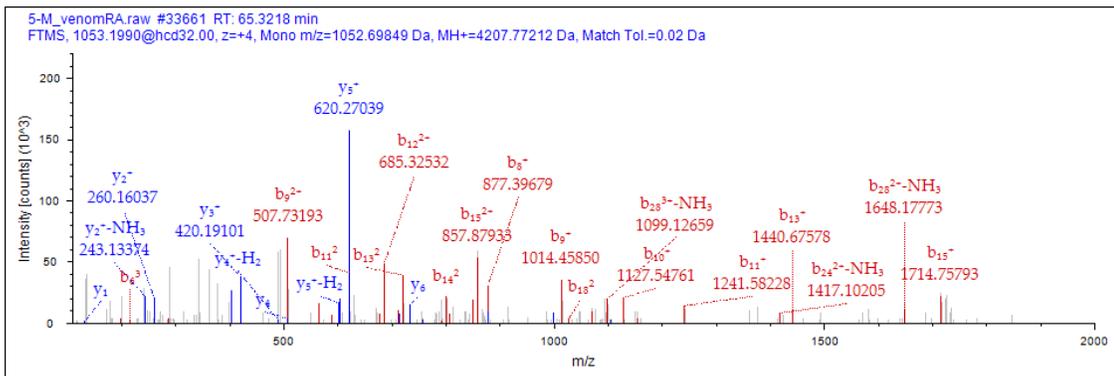
**PVID** belongs to the O1 Superfamily. It was previously reported as a nucleic acid sequence, P2b (Sequence ID: AAQ05866) [1], but this is the first time reported in the venom. It clusters within the lightning strike cabal. It closely resembles *C. purpurascens* nucleic acid sequence p2a (Sequence ID: AAQ05865), and new peptides PVIF (p2c, AAQ05867), and PVIG [1]. It does not share homology to any peptides with known activity. It does not share significant homology to any peptides with known activity, however it has the same cysteine framework and minimal homology to  $\kappa$ -PVIIA.





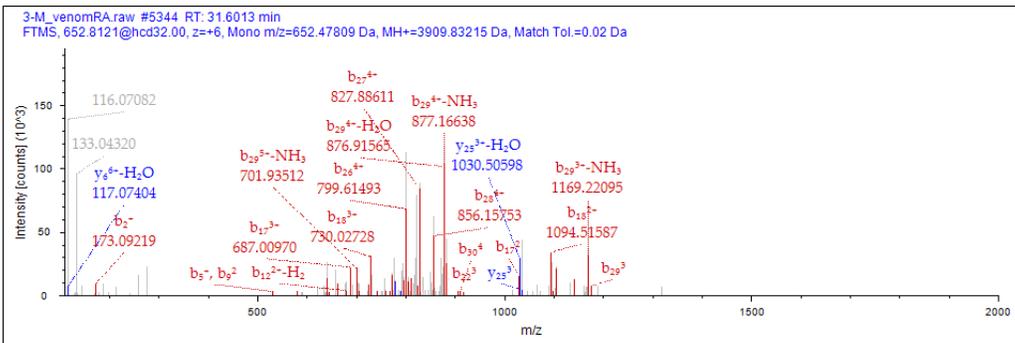
**Figure 66- Supplementary information for PVID 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra**

PVIE belongs to the O1 Superfamily. The mature peptide has little sequence homology to any characterized conotoxin but has high homology to a nucleic acid sequence from *C. ermineus* (Sequence ID: AXL95668) [197]. Interestingly, it is the only framework VI/ O1 Superfamily toxin that is expressed within cluster 2 (motor cabal).



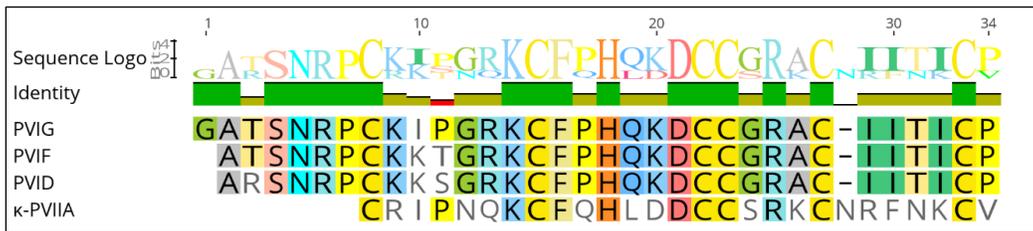
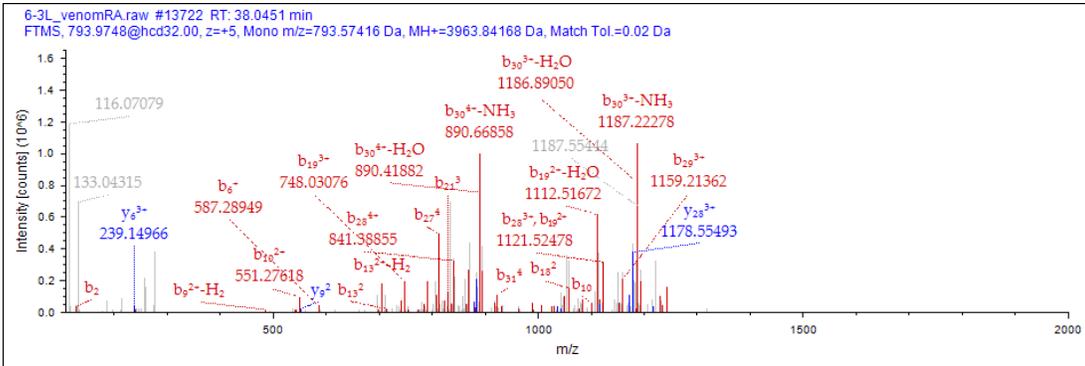
**Figure 67- Supplementary information for PVIE 1) full transcript with annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra**

**PVIF** belongs to the O1 Superfamily. It was previously reported as a nucleic acid sequence, P2c (Sequence ID: AAQ05867) [1], but this is the first time reported in the venom. PVIF was only identified in the venom of one specimen. It closely resembles *C. purpurascens* nucleic acid sequences p2a (Sequence ID: AAQ05865), PVID (p2b, AAQ05866) and PVIG [1]. It clusters within the lightning strike cabal. It does not share significant homology to any peptides with known activity but shares minimal homology with  $\kappa$ -PVIIA.



**Figure 68- Supplementary information for PVIF** 1) full transcript with an annotated signal sequence and mature peptide regions and 2) annotated MSMS spectra

**PVIG** is a new O1 Superfamily conotoxin. It closely resembles *C. purpurascens* nucleic acid sequences p2a (Sequence ID: AAQ05865), PVID (p2b, AAQ05866) and PVIF (p2c, AAQ05867) [1]. PVIG clustered with the lightning-strike cabal, although was only identified in one specimen. It does not share significant homology to any peptides with known activity, however, it has the same cysteine framework and minimal homology to  $\kappa$ -PVIIA.

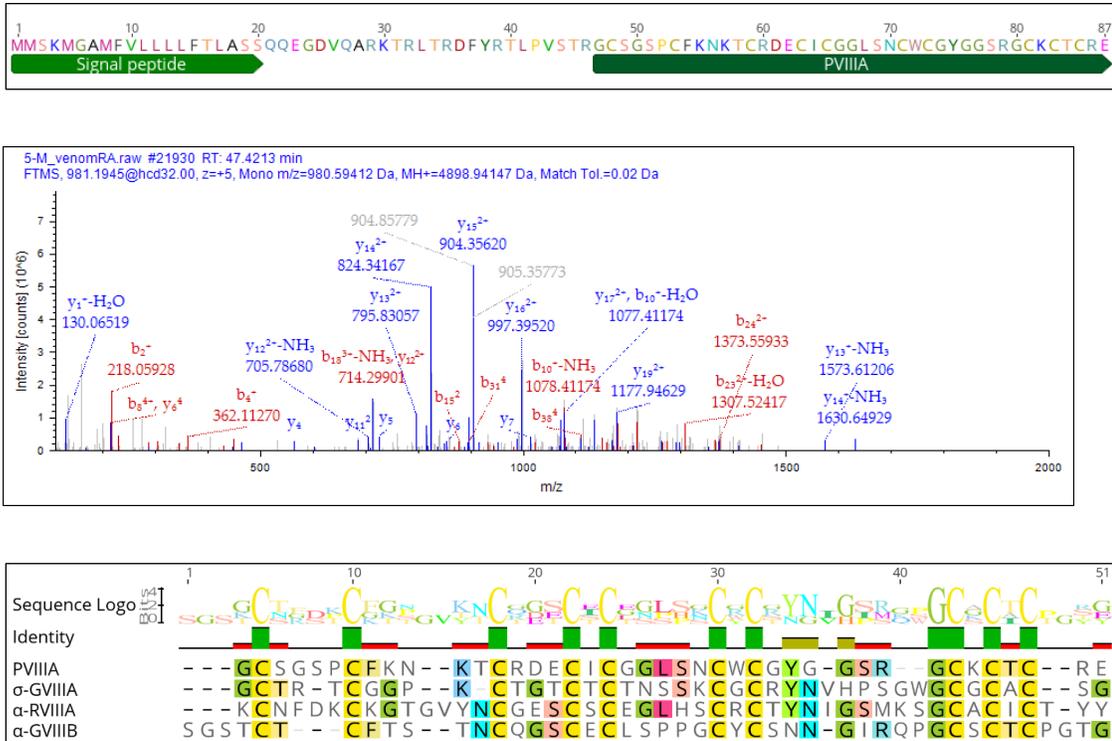


**Figure 69- Supplementary information for PVIG** 1) full transcript with annotated signal sequence and mature peptide regions, 2) annotated MSMS spectra, and 3) sequence alignment with PVIF, and PVID, and κ-PVIIA.

**Framework VIII (C-C-C-CXC-CXC-CXCXC)**

**PVIIIA** belongs to the S Superfamily of conotoxins and contains 5 disulfide bonds. It was identified in 25 of the 27 venom samples, and therefore likely has an important role in the venom that has yet to be revealed. It clusters closely with ψ-PIIIIE and α-PIVA and is a major component of the motor cabal. There are few framework VIII conotoxins with known bioactivity; two target the nAChR (α-GVIIIB, α-RVIIA) [206, 207] and one targets the serotonin receptor (σ-GVIIIA) [208]. However, PVIIIA does not display high sequence

homology to any of these characterized toxins, aside from a conserved cysteine framework.

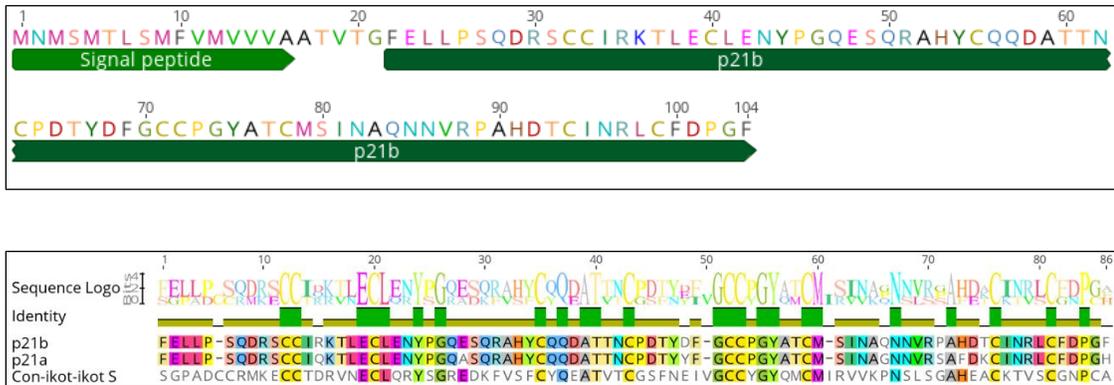


**Figure 70- Supplementary information for PVIIIA** 1) full transcript with annotated signal sequence and mature peptide regions, 2) annotated MSMS spectra, and 3) sequence alignment with  $\sigma$ -GVIIIA,  $\alpha$ -GVIIIB, and  $\alpha$ -RVIIIA

**Framework XXI (CC-C-C-C-CC-C-C-C)**

**p21b** was expressed in transcriptome B and its precursor sequence classifies it as part of the con-ikot-ikot family. It shows 91% identity to P21a, a previously described 10 cysteine, 5-disulfide conotoxin [127]. P21a was not expressed in either transcriptome and was not detected in the milked venom. However, P21b was identified in 10 of the 27 venom samples. It clusters closely to PVB and PIIIH, both newly described here. A con-ikot-ikot

isolated from *C. striatus* targets the AMPA receptor and is presumed to contribute to the lightning strike cabal [209]. In this study, P21b expression in the venom clusters with motor cabal toxins, suggesting an alternative molecular target for the con-ikot-ikot family of knottin peptides.



**Figure 71- Supplementary information for p21b** 1) full transcript with annotated signal sequence and mature peptide regions and 2) sequence alignment with p21a and con-ikot-ikot S

## APPENDIX D: Alignment of gastropod insulin superfamily proteins.

**A0A0B5ADV0 | Con-Ins Me1** -----MATSS**C**FLLVTLG--LLLHVQQ-AFLHE-HT**C**SPSEP---AAPGGI**CG**SNLAELHSFL**CE**KELEDY-----  
**A0A0B5A8P4 | Con-Ins G3** -----MTTSFYFLLVALG--LLLYV**C**QSSFGNQ-HTRNSDTP-----KHR**CG**SELADQYVQL**CH**-----GK-----  
**A0A0B5AC86 | Con-Ins G3b** -----MTTSFYFLLVALG--LLLYV**C**QSSFGNQ-HTRNSDTP-----KHR**CG**SELADQYVQL**CH**-----GK-----  
**A0A0B5A7P2 | Con-Ins G1C** -----MTTSFYFLLMALG--LLLYV**C**QSSFGNQ-HTRTFDTP-----KHR**CG**SEITNSYMDL**CY**-----RK-----  
**A0A0B5A8Q2 | Con-Ins G1b** -----MTTSFYFLLMALG--LLLYV**C**QSSFGNQ-HTRTFDTP-----KHR**CG**SEITNSYMDL**CY**-----RK-----  
**A0A0B5AC95 | Con-Ins G1a** -----MTTSSYFLLMALG--LLLYV**C**QSSFGNQ-HTRTFDTP-----KHR**CG**SEITNSYMDL**CY**-----RK-----  
**A0A0B5ADU4 | Con-Ins T1** -----MTTSFYFLLMALG--LLLYV**C**QSSFGNQ-HTRNSDTP-----KYR**CG**SEIPNSYIDL**CF**-----RK-----  
**A0A0B5ABD5 | Con-Ins T3** -----MTTSFYFLLMALG--LLLYV**C**QSSFGNQ-HTRNSDTP-----KYR**CG**SDIPNSYMDL**CF**-----RK-----  
**A0A0B5AC90 | Con-Ins T2** -----MTTSFYFLLMALG--LLLYV**C**QSSFGNQ-HTRNSDTP-----KYR**CG**SDIPNSYMDL**CF**-----RK-----  
**A0A0B5ABD9 | Con-Ins G2** -----MTTSSYFLLVALG--LLLYV**R**QSFSTHE-HT**C**QLDDP---AHPQ**GK****CG**SDLVNYHEEK**CE**EEEEARRGG-----  
**A0A0B5ADT3 | Con-Ins G2b** -----MTTSSYFLLVALG--LLLYV**R**QSFSTHE-HT**C**QLDDP---AHPQ**GK****CG**SDLVNYHEEK**CE**EEEEARRGG-----  
**A0A0B5ABE4 | Con-Ins Q1b** -----MTTSSYFLLVALG--LLLYL**C**QSSFGTE-HT**C**EPGAS---PHPQ**GK****CG**PELAEFHETM**CE**VEESLQGG-----  
**A0A0B5ABE6 | Con-Ins Q1** -----MTTSSYFLLVALG--LLLYL**C**QSSFGTE-HT**C**EPGAS---PHPQ**GK****CR**PELAEFHETM**CE**VEESLQGG-----  
**A0A0B5AC98 | Con-Ins F1** -----MTTSSYFLLVTLG--LLLYV**CR**SSFGTE-HT**C**ESDAS---PHPQ**GV****CG**SPLAEAVEAA**CE**LEEYLQGG-----  
**A0A0B5A7N5 | Con-Ins F2C** -----MTTSSYFLLVALG--LLLYV**CR**SSFGSE-HT**C**ESDAS---PHPQ**GV****CG**SPLAEAVEAA**CE**LEESLQGG-----  
**A0A0B5A7N1 | Con-Ins F2b** -----MTTSSYFLLVALG--LLLYV**CR**SSFGSE-HT**C**ESDAS---PHPQ**GV****CG**SPLAEAVEAA**CE**LEESLQGG-----  
**A0A0B5ADT9 | Con-Ins F2** -----MTTSSYFLLVALG--LLLYV**CR**SSFGSE-HT**C**ESDAS---PHPQ**GV****CG**SPLAEAVEAA**CE**LEESLQGG-----  
**A0A0B5A8Q6 | Con-Ins M1** -----MTTSSYFLLVALG--LLLYV**C**QSSFGGE-HV**CG**SNQP---NHPNG**K****CG**SKMADYLEEQ**CE**EEEEAAHGG-----  
**A0A0B5A7N8 | Con-Ins Tx1** -----MTTSSYFLLVALG--LLLYV**F**QSSFGGE-HV**C**WLGD---NHPQ**GI****CG**PQVADIVEIR**CE**EKEAEQGG-----  
**A0A0F7YYV0 | ILP-1** -----MTTSSYFLLVALG--LLLYV**C**QSSFGGE-HV**C**WLDDP---NHPE**GI****CG**PQVSDIVEIR**CE**EKEAEQGG-----  
**P91797 | MIP-7** -----MNASVES**C**LTFTFVL--VAL**CV**GLTIG--QQVNT**CT**MFSR---QHPR**GL****CG**NRLARAHANL**CF**LLRNTY**PD**IFPRK-----  
**Q9NDE7 | MIP-1** MSKFL**LQ**SHSANA**CL**LTLTLLT-LAS**NLD**ISLANFE-HS**C**NGYMR---PHPR**GL****CG**EDLHV**I**IS**NL****CS**SLGGNR-----  
**P80090 | MIP-3** MASV--HLTLTKAFM**VT**VFLT-LLLN**VS**ITRGT**TQ**-HT**CS**ILSR---PHPR**GL****CG**STLANM**VQ**WL**CS**TYTT**SS**KVK--R-  
**P07223 | MIP-1** MAGV--RLVFTKAFM**VT**VLLT-LLLN**IG**VKPAEG**Q**FSAC**N**INDR---PHRR**GV****CG**SALADL**VD**FA**CS**SS**NQ**PAMVK--R-  
**P25289 | MIP-2** MVGV--RLVFTNAF**VT**VLLT-LLLN**IG**VKPAEG**Q**FSAC**S**FSR---PHPR**GI****CG**SNLAG**FR**AFI**CS**N**Q**NS**PS**MVK--R-  
**P31241 | MIP-5** MAGV--RLVFTKAFM**VT**VLLT-LLLN**IG**VKPAEG**Q**FSAC**S**FSR---PHPR**GI****CG**SDLADL**RA**FI**CS**RR**NQ**PAMVK--R-  
**A0A0B5A7M7 | Con-Ins Im1** MA-----T**S**LLS**PL**L**V**AMLG--FLLHVH**V**ARAGLE-HT**CT**LETR**MQ**GAHP**QGI****CG**SKLPD**IV**HT**V****C**Q**VM**GRGY-----  
**A0A0B5A8P8 | Con-Ins Im2** MAL---TW**P**SS**PP**VLLTLL**L**SLLAL**QL****CA**VYGS**YE**-HT**CT**L**AT**RSRGA**HP**SGI**CG**RNLAR**IV**SVL**CT**PRG--Y-----

**A0A0B5ADV0 | Con-Ins Me1** --SGSALKKRGRPSR-----RMKRR-----D**FL**SALKTRV**KR**-----  
**A0A0B5A8P4 | Con-Ins G3** --RNDAGKKRGRASP-----L**W**QRQ-----G**FL**SML**KA**--**KR**-----  
**A0A0B5AC86 | Con-Ins G3b** --RNDAGKKRGRASP-----L**W**QRQ-----G**FL**SML**KA**--**KR**-----  
**A0A0B5A7P2 | Con-Ins G1C** --RNDAGKKRGRASP-----L**W**QRR-----G**S**L**SQL**KAR**AKR**-----  
**A0A0B5A8Q2 | Con-Ins G1b** --RNDAGEK**R**GRASP-----L**W**QRR-----G**FL**SKL**KAR**AKR-----

**A0A0B5AC95 | Con-Ins G1a** --RNDAGEKRGGRASP-----LWQRR-----GSLSKLKARAKR-----  
**A0A0B5ADU4 | Con-Ins T1** --RNDAGKKRGRASP-----LWQRG-----GSLSMLKARAKR-----  
**A0A0B5ABD5 | Con-Ins T3** --RNDAGKKRGRASP-----LWQRG-----GSLSMLKARAKR-----  
**A0A0B5AC90 | Con-Ins T2** --RNDAGKKRGRASP-----LWQRG-----GSLSMLKARAKR-----  
**A0A0B5ABD9 | Con-Ins G2** --TNDGGKRRRRASP-----LWKRR-----RFLSMLKARAKR-----  
**A0A0B5ADT3 | Con-Ins G2b** --TNDGGKRRRRASP-----LRKRR-----RFISMLKARAKR-----  
**A0A0B5ABE4 | Con-Ins Q1b** --TDDARKKRGRASL-----LRKRR-----GFLSMLKARAKR-----  
**A0A0B5ABE6 | Con-Ins Q1** --TDDARKKRGRASL-----LRKRR-----GFLSMLKARAKR-----  
**A0A0B5AC98 | Con-Ins F1** --TG---KKRGRASP-----LRKRR-----AFLSMLKARAKR-----  
**A0A0B5A7N5 | Con-Ins F2C** --TG---KKRGRASL-----LRKRR-----AFLSMLKARAKR-----  
**A0A0B5A7N1 | Con-Ins F2b** --TG---KKRGRASL-----LRKRR-----AFLSMLKARAKR-----  
**A0A0B5ADT9 | Con-Ins F2** --TG---KKRGRASL-----LRKRR-----AFLSMLKARAKR-----  
**A0A0B5A8Q6 | Con-Ins M1** --TNDARATTGRALS-----LSKRR-----GFLSMLKRRGKR-----  
**A0A0B5A7N8 | Con-Ins Tx1** --ANNARANTGRTSS-----LMKRR-----GFLSLLKKGKR-----  
**A0A0F7YYV0 | ILP-1** --ANNARAYTGRTSS-----LMKRR-----GFLSLLKKGKR-----  
**P91797 | MIP-7** RSVDNTEFEKV-YSIPLSVLAELDLSDDDWGAYVSKKDIPIRSETNGLSGANFESSAFDKQLELPAMKSTTSQLFRIKLKR  
**Q9NDE7 | MIP-1** -----RFL---AKYMKRDT-TENVNDKLRGILLNKKE-----AFSYLTKR-----  
**P80090 | MIP-3** QAE-----PDEEDDAMSKIMISKKR-----ALSYLTKR-----  
**P07223 | MIP-1** -----RNAETDLDDPLRNIKLSSES-----ALTYLTKR-----  
**P25289 | MIP-2** DAETGWLL---PETMVKRNAETDLDDPLRNIKLSSES-----ALTYLTKR-----  
**P31241 | MIP-5** DAETGWLL---PETMVKRNAQTDLDDPLRNIKLSSES-----ALTYLTKR-----  
**A0A0B5A7M7 | Con-Ins Im1** --AGGQRQLRKRTSMIDSDDMEAEGGSRGGFLMSKRR-----ALSYLQKETNPL--VMAGY  
**A0A0B5A8P8 | Con-Ins Im2** --VSNWFTK--RSAP-NKPAETFVDQNLRGVLLNKRE-----ALSYLPR-----

**A0A0B5ADV0 | Con-Ins Me1** KEGRSVKRSPTS GMS**CECC**KNS**CDAEE**ILEY**CPPL**PSS-----  
**A0A0B5A8P4 | Con-Ins G3** NEAFFLQRD-GRGIVEV**CCDN**P**CTVAT**LRT**FC**H-----  
**A0A0B5AC86 | Con-Ins G3b** NEAFFLQRD-GRGIVEV**CCDN**P**CTVAT**LMT**FC**H-----  
**A0A0B5A7P2 | Con-Ins G1C** NGAFHLPRD-GRGVVEH**CCHRP****CSNAEF**KKY**CS**-----  
**A0A0B5A8Q2 | Con-Ins G1b** NGAFHLPRD-GRGVVEH**CCHRP****CSNAEF**KKY**CG**-----  
**A0A0B5AC95 | Con-Ins G1a** NGAFHLPRD-GRGVVEH**CCHRP****CSNAEF**KKY**CG**-----  
**A0A0B5ADU4 | Con-Ins T1** NEAFHLQRA-HRGVVEH**CCHRP****CSNAEF**KK**FC**G-----  
**A0A0B5ABD5 | Con-Ins T3** NEAFHLQRA-HRGVVEH**CKRA****CSNAEF**M**QFC**GN-----  
**A0A0B5AC90 | Con-Ins T2** NEAFHLQRA-HRGVVEH**CCYRP****CSNAEF**KK**FC**G-----  
**A0A0B5ABD9 | Con-Ins G2** TG-----YKGIA**CECC**QHY**CTDQ**EFINY**CPPV**TESSSSSSAA-----  
**A0A0B5ADT3 | Con-Ins G2b** RG-----YQGIA**CECC**QHY**CTDQ**EFINY**CPPV**TESSSSSSAV-----  
**A0A0B5ABE4 | Con-Ins Q1b** NEASPLPRA-GRGIV**CECC**KNS**CTYEE**I**TEY****CPPV**TEGSG-----  
**A0A0B5ABE6 | Con-Ins Q1** NEASPLPRA-GRGIV**CECC**KNS**CTYEE**I**TEY****CPPV**TEGSG-----

```

A0A0B5AC98 | Con-Ins F1  NEASPLQRS-GRGIVCECCKNHCNIEELTEYCPPVTEGSG-----
A0A0B5A7N5 | Con-Ins F2C NEASPLQRS-GRGIVCECCKNHCNIEELTEYCPPVTEGSG-----
A0A0B5A7N1 | Con-Ins F2b NEASPLQRS-GRGIVCECCKNHCNLEELTEYCPPVTEGSG-----
A0A0B5ADT9 | Con-Ins F2  NEASPLQRS-GRGIVCECCKNHCNIEELTEYCPPVTEGSG-----
A0A0B5A8Q6 | Con-Ins M1  NEASPLQRA-GRGIVCECCKNHCTDEEFTEYCPHVTESG-----
A0A0B5A7N8 | Con-Ins Tx1  DEGSPLQRS-GRGIVCECCKHHCTKEEFTEYCH-----
A0A0F7YYV0 | ILP-1      DEGS-LQRS-GRGIVCECCKHHCTKEELTEYCH-----
P91797 | MIP-7      GSRLKREVMAEPSLVCDCCYNECSVRKLATYC-----
Q9NDE7 | MIP-1      -----EASGSITCECCFNQCRIFELAQYCRLPDHFFSRISRTGRSNSGHAQLEDNFS
P80090 | MIP-3      -----ESRPSIVCECCFNQCTVQELLAYC-----
P07223 | MIP-1      -----QGTTNIVCECCMKPCTLSELRQYCP-----
P25289 | MIP-2      -----QRTTNLVCECCFNYCTPDVVRKYCY-----
P31241 | MIP-5      -----QRTTNLVCECCYNVCTVDVFYEYCY-----
A0A0B5A7M7 | Con-Ins Im1  ERRGIQKRHGEQGITCECCYNHCSFRELVQYCN-----
A0A0B5A8P8 | Con-Ins Im2  EPRATRGTFGSQGITCECCFNQCTYYELLQYCN-----

```

Figure 72- Alignment of gastropod insulin superfamily proteins. Sequences are from InterPro database (reviewed).

## CURRICULUM VITAE

Meghan K. Grandal was born on July 14, 1989, and grew up in Cranford, NJ. She attended Cranford High School where she graduated in 2007. She received her Bachelor of Science degree in marine biology from The University of North Carolina Wilmington (UNCW) in 2011, during which time worked in the marine natural product laboratory of Dr. Dan Baden and Dr. Andrea Bourdelias. She went on to receive her Master of Science degree in marine biology at UNCW working under the direction of both Dr. Alison Taylor and Dr. Andrea Bourdelais. Her work at UNCW aimed to understand the cellular mechanisms of the algal neurotoxin brevetoxin and other non-toxic algal derivatives from the Florida red tide dinoflagellate, *Karenia brevis*. She completed her master's degree in 2014 and began a Ph.D. program at the Medical University of South Carolina in 2015. Her dissertation work was completed under the supervision of Dr. Frank Mari at the National Institute of Standards and Technology at Hollings Marine Laboratory in Charleston, SC.

## REFERENCES

1. Duda TF, Jr. and Palumbi SR (2004) Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*. *Proc Biol Sci* 271:1165-74. doi: 10.1098/rspb.2004.2708
2. Lewis RJ and Garcia ML (2003) Therapeutic potential of venom peptides. *Nature Reviews Drug Discovery* 2:790. doi: 10.1038/nrd1197
3. Pennington MW, Czerwinski A and Norton RS (2018) Peptide therapeutics from venom: Current status and potential. *Bioorganic and Medicinal Chemistry* 26:2738-2758. doi: 10.1016/j.bmc.2017.09.029
4. Veisheh M, Gabikian P, Bahrami S-B, Veisheh O, Zhang M, Hackman RC, Ravanpay AC, Stroud MR, Kusuma Y, Hansen SJ, Kwok D, Munoz NM, Sze RW, Grady WM, Greenberg NM, Ellenbogen RG and Olson JM (2007) Tumor Paint: A Chlorotoxin: Cy5.5 Bioconjugate for Intraoperative Visualization of Cancer Foci. *Cancer Research* 67:6882-6888. doi: 10.1158/0008-5472.can-06-3948
5. Tarcha EJ, Olsen CM, Probst P, Peckham D, Muñoz-Elías EJ, Kruger JG and Iadonato SP (2017) Safety and pharmacodynamics of dalazatide, a Kv1.3 channel inhibitor, in the treatment of plaque psoriasis: A randomized phase 1b trial. *PLOS ONE* 12:e0180762. doi: 10.1371/journal.pone.0180762
6. Shcherbatko A, Rossi A, Foletti D, Zhu G, Bogin O, Galindo Casas M, Rickert M, Hasa-Moreno A, Bartsevich V, Crameri A, Steiner AR, Henningsen R, Gill A, Pons J, Shelton DL, Rajpal A and Strop P (2016) Engineering Highly Potent and Selective Microproteins against Nav1.7 Sodium Channel for Treatment of Pain. *Journal of Biological Chemistry* 291:13974-13986. doi: 10.1074/jbc.M116.725978
7. Di Cesare Mannelli L, Cinci L, Micheli L, Zanardelli M, Pacini A, McIntosh JM and Ghelardini C (2014) alpha-conotoxin RglA protects against the development of nerve injury-induced chronic pain and prevents both neuronal and glial derangement. *Pain* 155:1986-95. doi: 10.1016/j.pain.2014.06.023
8. Egleton RD and Davis TP (1997) Bioavailability and Transport of Peptides and Peptide Drugs into the Brain. *Peptides* 18:1431-1439. doi: [https://doi.org/10.1016/S0196-9781\(97\)00242-8](https://doi.org/10.1016/S0196-9781(97)00242-8)
9. Clark RJ, Fischer H, Dempster L, Daly NL, Rosengren KJ, Nevin ST, Meunier FA, Adams DJ and Craik DJ (2005) Engineering stable peptide toxins by means of backbone cyclization: stabilization of the alpha-conotoxin MII. *Proc Natl Acad Sci U S A* 102:13767-72. doi: 10.1073/pnas.0504613102
10. Clark RJ, Akcan M, Kaas Q, Daly NL and Craik DJ (2012) Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* 59:446-455. doi: <https://doi.org/10.1016/j.toxicon.2010.12.003>
11. Craik DJ, Fairlie DP, Liras S and Price D (2013) The future of peptide-based drugs. *Chem Biol Drug Des* 81:136-47. doi: 10.1111/cbdd.12055

12. Moradi SV, Hussein WM, Varamini P, Simerska P and Toth I (2016) Glycosylation, an effective synthetic strategy to improve the bioavailability of therapeutic peptides. *Chem Sci* 7:2492-2500. doi: 10.1039/c5sc04392a
13. Kintzing JR and Cochran JR (2016) Engineered knottin peptides as diagnostics, therapeutics, and drug delivery vehicles. *Current Opinion in Chemical Biology* 34:143-150. doi: 10.1016/j.cbpa.2016.08.022
14. Zhu S (2003) Evolutionary origin of inhibitor cystine knot peptides. doi: 10.1096/fj.02-1044fje
15. Su M (2014) Marine Knottins with Remarkable Biological Functions Cast a Promising Outlook on Clinical Translation. *Oceanography: Open Access* 02. doi: 10.4172/2332-2632.1000e110
16. Li H, Bowling JJ, Su M, Hong J, Lee B-J, Hamann MT and Jung JH (2014) Asteropsins B–D, sponge-derived knottins with potential utility as a novel scaffold for oral peptide drugs. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1840:977-984. doi: 10.1016/j.bbagen.2013.11.001
17. Corsetti M and Tack J (2013) Linaclotide: A new drug for the treatment of chronic constipation and irritable bowel syndrome with constipation. *United European Gastroenterology Journal* 1:7-20. doi: 10.1177/2050640612474446
18. Schmidtke A, Lotsch J, Freynhagen R and Geisslinger G (2010) Ziconotide for treatment of severe chronic pain. *Lancet* 375:1569-77. doi: 10.1016/S0140-6736(10)60354-6
19. Colgrave ML and Craik DJ (2004) Thermal, Chemical, and Enzymatic Stability of the Cyclotide Kalata B1: The Importance of the Cyclic Cystine Knot†. *Biochemistry* 43:5965-5975. doi: 10.1021/bi049711q
20. Werle M, Schmitz T, Huang H-L, Wentzel A, Kolmar H and Bernkop-Schnürch A (2006) The potential of cystine-knot microproteins as novel pharmacophoric scaffolds in oral peptide drug delivery. *Journal of Drug Targeting* 14:137-146. doi: 10.1080/10611860600648254
21. Wu Y, Wang L, Zhou M, You Y, Zhu X, Qiang Y, Qin M, Luo S, Ren Z and Xu A (2013) Molecular Evolution and Diversity of Conus Peptide Toxins, as Revealed by Gene Structure and Intron Sequence Analyses. *PLOS ONE* 8:e82495. doi: 10.1371/journal.pone.0082495
22. Espiritu MJ, Cabaltega CC, Sugai CK and Bingham JP (2014) Incorporation of post-translational modified amino acids as an approach to increase both chemical and biological diversity of conotoxins and conopeptides. *Amino Acids* 46:125-51. doi: 10.1007/s00726-013-1606-x
23. Buczek O, Bulaj G and Olivera BM (2005) Conotoxins and the posttranslational modification of secreted gene products. *Cellular and Molecular Life Sciences CMLS* 62:3067-3079. doi: 10.1007/s00018-005-5283-0
24. Lewis RJ, Dutertre S, Vetter I and Christie MJ (2012) Conus Venom Peptide Pharmacology. *Pharmacological Reviews* 64:259-298. doi: 10.1124/pr.111.005322
25. Terlau H and Olivera BM (2004) Conus Venoms: A Rich Source of Novel Ion Channel-Targeted Peptides. *Physiological Reviews* 84:41-68. doi: 10.1152/physrev.00020.2003

26. Trim SA and Trim CM (2013) Venom: the sharp end of pain therapeutics. *British journal of pain* 7:179-188. doi: 10.1177/2049463713502005
27. Mathie A (2010) Ion channels as novel therapeutic targets in the treatment of pain. *J Pharm Pharmacol* 62:1089-95. doi: 10.1111/j.2042-7158.2010.01131.x
28. Ostling PS, Davidson KS, Anyama BO, Helander EM, Wyche MQ and Kaye AD (2018) America's Opioid Epidemic: a Comprehensive Review and Look into the Rising Crisis. *Curr Pain Headache Rep* 22:32. doi: 10.1007/s11916-018-0685-5
29. McGivern J (2007) McGivern JG. Ziconotide: a review of its pharmacology and use in the treatment of pain. *Neuropsychiatr Dis Treat* 3: 69-85.
30. Satkunanathan N, Livett B, Gayler K, Sandall D, Down J and Khalil Z (2005) Alpha-conotoxin Vc1.1 alleviates neuropathic pain and accelerates functional recovery of injured neurones. *Brain Research* 1059:149-158. doi: <https://doi.org/10.1016/j.brainres.2005.08.009>
31. Livett BG, Sandall DW, Keays D, Down J, Gayler KR, Satkunanathan N and Khalil Z (2006) Therapeutic applications of conotoxins that target the neuronal nicotinic acetylcholine receptor. *Toxicon* 48:810-29. doi: 10.1016/j.toxicon.2006.07.023
32. Romero HK, Christensen SB, Di Cesare Mannelli L, Gajewiak J, Ramachandra R, Elmslie KS, Vetter DE, Ghelardini C, Iadonato SP, Mercado JL, Olivera BM and McIntosh JM (2017) Inhibition of  $\alpha 9\alpha 10$  nicotinic acetylcholine receptors prevents chemotherapy-induced neuropathic pain. *Proceedings of the National Academy of Sciences* 114:E1825-E1832. doi: 10.1073/pnas.1621433114
33. Vincler M, Wittenauer S, Parker R, Ellison M, Olivera BM and McIntosh JM (2006) Molecular mechanism for analgesia involving specific antagonism of  $\alpha 9\alpha 10$  nicotinic acetylcholine receptors. *Proc Natl Acad Sci U S A* 103:17880-4. doi: 10.1073/pnas.0608715103
34. Cuny H, de Faoite A, Huynh TG, Yasuda T, Berecki G and Adams DJ (2012)  $\gamma$ -Aminobutyric Acid Type B (GABAB) Receptor Expression Is Needed for Inhibition of N-type (Cav2.2) Calcium Channels by Analgesic  $\alpha$ -Conotoxins. *Journal of Biological Chemistry* 287:23948-23957. doi: 10.1074/jbc.M112.342998
35. Carstens BB, Berecki G, Daniel JT, Lee HS, Jackson KAV, Tae H-S, Sadeghi M, Castro J, O'Donnell T, Deiteren A, Brierley SM, Craik DJ, Adams DJ and Clark RJ (2016) Structure–Activity Studies of Cysteine-Rich  $\alpha$ -Conotoxins that Inhibit High-Voltage-Activated Calcium Channels via GABAB Receptor Activation Reveal a Minimal Functional Motif. *Angewandte Chemie International Edition* 55:4692-4696. doi: doi:10.1002/anie.201600297
36. Sine SM and Engel AG (2006) Recent advances in Cys-loop receptor structure and function. *Nature* 440:448-55. doi: 10.1038/nature04708
37. Miller PS and Smart TG (2010) Binding, activation and modulation of Cys-loop receptors. *Trends in Pharmacological Sciences* 31:161-174. doi: <https://doi.org/10.1016/j.tips.2009.12.005>
38. Unwin N (1993) Nicotinic Acetylcholine Receptor an 9 Å Resolution. *Journal of Molecular Biology* 229:1101-1124. doi: <https://doi.org/10.1006/jmbi.1993.1107>
39. Papke RL (2014) Merging old and new perspectives on nicotinic acetylcholine receptors. *Biochemical pharmacology* 89:1-11. doi: 10.1016/j.bcp.2014.01.029

40. Alberola-Die A, Martinez-Pinna J, Gonzalez-Ros JM, Ivorra I and Morales A (2011) Multiple inhibitory actions of lidocaine on Torpedo nicotinic acetylcholine receptors transplanted to *Xenopus* oocytes. *J Neurochem* 117:1009-19. doi: 10.1111/j.1471-4159.2011.07271.x
41. Balass M, Katchalski-Katzir E and Fuchs S (1997) The  $\alpha$ -bungarotoxin binding site on the nicotinic acetylcholine receptor: Analysis using a phage-epitope library. *Proceedings of the National Academy of Sciences* 94:6054-6058. doi: 10.1073/pnas.94.12.6054
42. Gotti C and Clementi F (2004) Neuronal nicotinic receptors: from structure to pathology. *Progress in Neurobiology* 74:363-396. doi: <https://doi.org/10.1016/j.pneurobio.2004.09.006>
43. Hansen SB, Talley TT, Radic Z and Taylor P (2004) Structural and ligand recognition characteristics of an acetylcholine-binding protein from *Aplysia californica*. *The Journal of biological chemistry* 279:24197-24202. doi: 10.1074/jbc.M402452200
44. Mukhtasimova N, Free C and Sine SM (2005) Initial coupling of binding to gating mediated by conserved residues in the muscle nicotinic receptor. *The Journal of general physiology* 126:23-39. doi: 10.1085/jgp.200509283
45. Hansen SB, Sulzenbacher G, Huxford T, Marchot P, Taylor P and Bourne Y (2005) Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *The EMBO journal* 24:3635-3646. doi: 10.1038/sj.emboj.7600828
46. Unwin N (2005) Refined Structure of the Nicotinic Acetylcholine Receptor at 4Å Resolution. *Journal of Molecular Biology* 346:967-989. doi: 10.1016/j.jmb.2004.12.031
47. Gotti C, Clementi F, Fornari A, Gaimarri A, Guiducci S, Manfredi I, Moretti M, Pedrazzi P, Pucci L and Zoli M (2009) Structural and functional diversity of native brain neuronal nicotinic receptors. *Biochem Pharmacol* 78:703-11. doi: 10.1016/j.bcp.2009.05.024
48. Hurst R, Rollema H and Bertrand D (2013) Nicotinic acetylcholine receptors: From basic science to therapeutics. *Pharmacology & Therapeutics* 137:22-54. doi: <https://doi.org/10.1016/j.pharmthera.2012.08.012>
49. McIntosh JM, Yoshikami D, Mahe E, Nielsen DB, Rivier JE, Gray WR and Olivera BM (1994) A nicotinic acetylcholine receptor ligand of unique specificity, alpha-conotoxin Iml. *J Biol Chem* 269:16733-9.
50. McIntosh JM, Santos AD and Olivera BM (1999) Conus peptides targeted to specific nicotinic acetylcholine receptor subtypes. *Annu Rev Biochem* 68:59-88. doi: 10.1146/annurev.biochem.68.1.59
51. Kang TS, Vivekanandan S, Jois SDS and Kini RM (2005) Effect of C-Terminal Amidation on Folding and Disulfide-Pairing of  $\alpha$ -Conotoxin Iml. *Angewandte Chemie International Edition* 44:6333-6337. doi: 10.1002/anie.200502300
52. Lebbe EK, Peigneur S, Wijesekara I and Tytgat J (2014) Conotoxins targeting nicotinic acetylcholine receptors: an overview. *Mar Drugs* 12:2970-3004. doi: 10.3390/md12052970
53. Jin A-H, Daly NL, Nevin ST, Wang C-IA, Dutertre S, Lewis RJ, Adams DJ, Craik DJ and Alewood PF (2008) Molecular Engineering of Conotoxins: The Importance of Loop Size to  $\alpha$ -Conotoxin Structure and Function. *Journal of Medicinal Chemistry* 51:5575-5584. doi: 10.1021/jm800278k

54. Janes RW (2005)  $\alpha$ -Conotoxins as selective probes for nicotinic acetylcholine receptor subclasses. *Current Opinion in Pharmacology* 5:280-292. doi: <https://doi.org/10.1016/j.coph.2005.01.013>
55. Azam L and McIntosh JM (2009) Alpha-conotoxins as pharmacological probes of nicotinic acetylcholine receptors. *Acta Pharmacologica Sinica* 30:771. doi: 10.1038/aps.2009.47
56. Gotti C, Zoli M and Clementi F (2006) Brain nicotinic acetylcholine receptors: native subtypes and their relevance. *Trends in Pharmacological Sciences* 27:482-491. doi: <https://doi.org/10.1016/j.tips.2006.07.004>
57. Nelson ME, Kuryatov A, Choi CH, Zhou Y and Lindstrom J (2003) Alternate stoichiometries of alpha4beta2 nicotinic acetylcholine receptors. *Mol Pharmacol* 63:332-41. doi: 10.1124/mol.63.2.332
58. Banerjee J, Yongye AB, Chang YP, Gyanda R, Medina-Franco JL and Armishaw CJ (2014) Design and synthesis of alpha-conotoxin G1D analogues as selective alpha4beta2 nicotinic acetylcholine receptor antagonists. *Biopolymers* 102:78-87. doi: 10.1002/bip.22413
59. Luo S, Zhangsun D, Schroeder CI, Zhu X, Hu Y, Wu Y, Weltzin MM, Eberhard S, Kaas Q, Craik DJ, McIntosh JM and Whiteaker P (2014) A novel alpha4/7-conotoxin Lv1A from *Conus lividus* that selectively blocks alpha3beta2 vs. alpha6/alpha3beta2beta3 nicotinic acetylcholine receptors. *Faseb j* 28:1842-53. doi: 10.1096/fj.13-244103
60. Hogg RC, Miranda LP, Craik DJ, Lewis RJ, Alewood PF and Adams DJ (1999) Single amino acid substitutions in alpha-conotoxin Pn1A shift selectivity for subtypes of the mammalian neuronal nicotinic acetylcholine receptor. *J Biol Chem* 274:36559-64. doi: 10.1074/jbc.274.51.36559
61. Grady SR, Moretti M, Zoli M, Marks MJ, Zanardi A, Pucci L, Clementi F and Gotti C (2009) Rodent habenulo-interpeduncular pathway expresses a large variety of uncommon nAChR subtypes, but only the alpha3beta4\* and alpha3beta3beta4\* subtypes mediate acetylcholine release. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29:2272-2282. doi: 10.1523/JNEUROSCI.5121-08.2009
62. Toll L, Zaveri NT, Polgar WE, Jiang F, Khroyan TV, Zhou W, Xie XS, Stauber GB, Costello MR and Leslie FM (2012) AT-1001: a high affinity and selective  $\alpha 3\beta 4$  nicotinic acetylcholine receptor antagonist blocks nicotine self-administration in rats. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 37:1367-1376. doi: 10.1038/npp.2011.322
63. Grishin AA, Cuny H, Hung A, Clark RJ, Brust A, Akondi K, Alewood PF, Craik DJ and Adams DJ (2013) Identifying key amino acid residues that affect alpha-conotoxin AulB inhibition of alpha3beta4 nicotinic acetylcholine receptors. *J Biol Chem* 288:34428-42. doi: 10.1074/jbc.M113.512582
64. Glick SD, Maisonneuve IM, Kitchen BA and Fleck MW (2002) Antagonism of alpha 3 beta 4 nicotinic receptors as a strategy to reduce opioid and stimulant self-administration. *Eur J Pharmacol* 438:99-105. doi: 10.1016/s0014-2999(02)01284-0
65. Yang KC, Jin GZ and Wu J (2009) Mysterious alpha6-containing nAChRs: function, pharmacology, and pathophysiology. *Acta Pharmacol Sin* 30:740-51. doi: 10.1038/aps.2009.63

66. Berry JN, Engle SE, McIntosh JM and Drenan RM (2015) alpha6-Containing nicotinic acetylcholine receptors in midbrain dopamine neurons are poised to govern dopamine-mediated behaviors and synaptic plasticity. *Neuroscience* 304:161-75. doi: 10.1016/j.neuroscience.2015.07.052
67. Dowell C, Olivera BM, Garrett JE, Staheli ST, Watkins M, Kuryatov A, Yoshikami D, Lindstrom JM and McIntosh JM (2003) Alpha-conotoxin PIA is selective for alpha6 subunit-containing nicotinic acetylcholine receptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 23:8445-8452. doi: 10.1523/jneurosci.23-24-08445.2003
68. Chi SW, Lee SH, Kim DH, Kim JS, Olivera BM, McIntosh JM and Han KH (2005) Solution structure of alpha-conotoxin PIA, a novel antagonist of alpha6 subunit containing nicotinic acetylcholine receptors. *Biochem Biophys Res Commun* 338:1990-7. doi: 10.1016/j.bbrc.2005.10.176
69. McIntosh JM, Azam L, Staheli S, Dowell C, Lindstrom JM, Kuryatov A, Garrett JE, Marks MJ and Whiteaker P (2004) Analogs of  $\alpha$ -Conotoxin MII Are Selective for  $\alpha$ 6-Containing Nicotinic Acetylcholine Receptors. *Molecular Pharmacology* 65:944-952. doi: 10.1124/mol.65.4.944
70. Bertrand D, Lee C-HL, Flood D, Marger F and Donnelly-Roberts D (2015) Therapeutic Potential of  $\alpha$ 7 Nicotinic Acetylcholine Receptors. *Pharmacological Reviews* 67:1025-1073. doi: 10.1124/pr.113.008581
71. Breese CR, Adams C, Logel J, Drebing C, Rollins Y, Barnhart M, Sullivan B, Demasters BK, Freedman R and Leonard S (1997) Comparison of the regional expression of nicotinic acetylcholine receptor alpha7 mRNA and [125I]-alpha-bungarotoxin binding in human postmortem brain. *J Comp Neurol* 387:385-98. doi: 10.1002/(sici)1096-9861(19971027)387:3<385::aid-cne5>3.0.co;2-x
72. Sgard F, Charpantier E, Bertrand S, Walker N, Caput D, Graham D, Bertrand D and Besnard F (2002) A novel human nicotinic receptor subunit, alpha10, that confers functionality to the alpha9-subunit. *Mol Pharmacol* 61:150-9. doi: 10.1124/mol.61.1.150
73. McIntosh JM, Absalom N, Chebib M, Elgoyhen AB and Vincler M (2009) Alpha9 nicotinic acetylcholine receptors and the treatment of pain. *Biochem Pharmacol* 78:693-702. doi: 10.1016/j.bcp.2009.05.020
74. Dutertre S (2014) Venomics in medicinal chemistry. *Future Med Chem* 6:1609-10. doi: 10.4155/fmc.14.117
75. Prashanth JR, Lewis RJ and Dutertre S (2012) Towards an integrated venomics approach for accelerated conopeptide discovery. *Toxicon* 60:470-477. doi: <https://doi.org/10.1016/j.toxicon.2012.04.340>
76. Sunagar K, Morgenstern D, Reitzel AM and Moran Y (2016) Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. *J Proteomics* 135:62-72. doi: 10.1016/j.jprot.2015.09.015
77. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G and Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* 37:420-423. doi: 10.1038/s41587-019-0036-z

78. Tran NH, Qiao R, Xin L, Chen X, Liu C, Zhang X, Shan B, Ghodsi A and Li M (2019) Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nature Methods* 16:63-66. doi: 10.1038/s41592-018-0260-3
79. Webb B and Sali A (2016) Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics* 54. doi: 10.1002/cpbi.3
80. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH and Ferrin TE (2020) UCSF ChimeraX : Structure visualization for researchers, educators, and developers. *Protein Science*. doi: 10.1002/pro.3943
81. Holford M, Daly M, King GF and Norton RS (2018) Venoms to the rescue. *Science* 361:842-844. doi: 10.1126/science.aau7761
82. Abdel-Rahman MA, Abdel-Nabi IM, El-Naggar MS, Abbas OA and Strong PN (2011) Intraspecific variation in the venom of the vermivorous cone snail *Conus vexillum*. *Comparative Biochemistry and Physiology* 154:318-25. doi: 10.1016/j.cbpc.2011.06.019
83. Chun JB, Baker MR, Kim DH, Leroy M, Toribo P and Bingham JP (2012) Cone snail milked venom dynamics--a quantitative study of *Conus purpurascens*. *Toxicon* 60:83-94. doi: 10.1016/j.toxicon.2012.03.019
84. Davis J, Jones A and Lewis RJ (2009) Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS. *Peptides* 30:1222-7. doi: 10.1016/j.peptides.2009.03.019
85. Jakubowski JA, Kelley WP, Sweedler JV, Gilly WF and Schulz JR (2005) Intraspecific variation of venom injected by fish-hunting *Conus* snails. *Journal of Experimental Biology* 208:2873-83. doi: 10.1242/jeb.01713
86. Rivera-Ortiz JA, Cano H and Mari F (2011) Intraspecies variability and conopeptide profiling of the injected venom of *Conus ermineus*. *Peptides* 32:306-16. doi: 10.1016/j.peptides.2010.11.014
87. Romeo C, Di Francesco L, Oliverio M, Palazzo P, Massilia GR, Ascenzi P, Polticelli F and Schinina ME (2008) *Conus ventricosus* venom peptides profiling by HPLC-MS: a new insight in the intraspecific variation. *Journal of Separation Science* 31:488-98. doi: 10.1002/jssc.200700448
88. Biass D, Dutertre S, Gerbault A, Menou JL, Offord R, Favreau P and Stocklin R (2009) Comparative proteomic study of the venom of the piscivorous cone snail *Conus consors*. *Journal of Proteomics* 72:210-8. doi: 10.1016/j.jprot.2009.01.019
89. Himaya SWA, Mari F and Lewis RJ (2018) Accelerated proteomic visualization of individual predatory venoms of *Conus purpurascens* reveals separately evolved predation-evoked venom cabals. *Scientific Reports* 8:330. doi: 10.1038/s41598-017-17422-x
90. Rodriguez AM, Dutertre S, Lewis RJ and Mari F (2015) Intraspecific variations in *Conus purpurascens* injected venom using LC/MALDI-TOF-MS and LC-ESI-TripleTOF-MS. *Analytical and Bioanalytical Chemistry* 407:6105-16. doi: 10.1007/s00216-015-8787-y
91. Dutertre S, Biass D, Stocklin R and Favreau P (2010) Dramatic intraspecimen variations within the injected venom of *Conus consors*: an unsuspected contribution to venom diversity. *Toxicon* 55:1453-62. doi: 10.1016/j.toxicon.2010.02.025

92. Dutertre S, Jin AH, Alewood PF and Lewis RJ (2014) Intraspecific variations in *Conus geographus* defence-evoked venom and estimation of the human lethal dose. *Toxicon* 91:135-44. doi: 10.1016/j.toxicon.2014.09.011
93. Dutertre S, Jin AH, Vetter I, Hamilton B, Sunagar K, Lavergne V, Dutertre V, Fry BG, Antunes A, Venter DJ, Alewood PF and Lewis RJ (2014) Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nature Communications* 5:3521. doi: 10.1038/ncomms4521
94. Prashanth JR, Dutertre S and Lewis RJ (2017) Pharmacology of predatory and defensive venom peptides in cone snails. *molecular biosystems* 13:2453-2465. doi: 10.1039/c7mb00511c
95. Prator CA, Murayama KM and Schulz JR (2014) Venom variation during prey capture by the cone snail, *Conus textile*. *PLoS One* 9:e98991. doi: 10.1371/journal.pone.0098991
96. Olivera BM (1997) E.E. Just Lecture, 1996. *Conus* venom peptides, receptor and ion channel targets, and drug design: 50 million years of neuropharmacology. *Molecular Biology of the Cell* 8:2101-9. doi: 10.1091/mbc.8.11.2101
97. Olivera BM, Walker C, Cartier GE, Hooper D, Santos AD, Schoenfeld R, Shetty R, Watkins M, Bandyopadhyay P and Hillyard DR (1999) Speciation of cone snails and interspecific hyperdivergence of their venom peptides. Potential evolutionary significance of introns. *Annals of the New York Academy of Sciences* 870:223-37. doi: 10.1111/j.1749-6632.1999.tb08883.x
98. Lebbe EK and Tytgat J (2016) In the picture: disulfide-poor conopeptides, a class of pharmacologically interesting compounds. *J Venom Anim Toxins Incl Trop Dis* 22:30. doi: 10.1186/s40409-016-0083-6
99. Kaas Q, Westermann JC, Halai R, Wang CK and Craik DJ (2008) ConoServer, a database for conopeptide sequences and structures. *Bioinformatics* 24:445-6. doi: 10.1093/bioinformatics/btm596
100. Kaas Q, Yu R, Jin AH, Dutertre S and Craik DJ (2012) ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. *Nucleic Acids Research* 40:D325-30. doi: 10.1093/nar/gkr886
101. Kaas Q, Westermann JC and Craik DJ (2010) Conopeptide characterization and classifications: an analysis using ConoServer. *Toxicon* 55:1491-509. doi: 10.1016/j.toxicon.2010.03.002
102. Craig AG, Bandyopadhyay P and Olivera BM (1999) Post-translationally modified neuropeptides from *Conus* venoms. *European Journal of Biochemistry* 264:271-5. doi: 10.1046/j.1432-1327.1999.00624.x
103. Buczek O, Bulaj G and Olivera BM (2005) Conotoxins and the posttranslational modification of secreted gene products. *Cellular and Molecular Life Sciences* 62:3067-79. doi: 10.1007/s00018-005-5283-0
104. Franco A, Pisarewicz K, Moller C, Mora D, Fields GB and Mari F (2006) Hyperhydroxylation: a new strategy for neuronal targeting by venomous marine molluscs. *Progress in Molecular and Subcellular Biology* 43:83-103. doi: 10.1007/978-3-540-30880-5\_4

105. Degueldre M, Verdenaud M, Legarda G, Minambres R, Zuniga S, Leblanc M, Gilles N, Ducancel F, De Pauw E and Quinton L (2017) Diversity in sequences, post-translational modifications and expected pharmacological activities of toxins from four *Conus* species revealed by the combination of cutting-edge proteomics, transcriptomics and bioinformatics. *Toxicon* 130:116-125. doi: 10.1016/j.toxicon.2017.02.014
106. Puillandre N, Duda TF, Meyer C, Olivera BM and Bouchet P (2015) One, four or 100 genera? A new classification of the cone snails. *Journal of Molluscan Studies* 81:1-23. doi: 10.1093/mollus/eyu055
107. Jones RM and Bulaj G (2000) Conotoxins - new vistas for peptide therapeutics. *current pharmaceutical design* 6:1249-85. doi: 10.2174/1381612003399653
108. Mir R, Karim S, Kamal MA, Wilson CM and Mirza Z (2016) Conotoxins: Structure, Therapeutic Potential and Pharmacological Applications. *current pharmaceutical design* 22:582-9. doi: 10.2174/1381612822666151124234715
109. Xie B, Huang Y, Baumann K, Fry BG and Shi Q (2017) From Marine Venoms to Drugs: Efficiently Supported by a Combination of Transcriptomics and Proteomics. *Marine Drugs* 15:103. doi: 10.3390/md15040103
110. Wilson D and Daly NL (2018) Venomics: A Mini-Review. *High-Throughput* 7:19. doi: 10.3390/ht7030019
111. Terlau H, Shon KJ, Grilley M, Stocker M, Stuhmer W and Olivera BM (1996) Strategy for rapid immobilization of prey by a fish-hunting marine snail. *Nature* 381:148-51. doi: 10.1038/381148a0
112. Olivera BM, Seger J, Horvath MP and Fedosov AE (2015) Prey-Capture Strategies of Fish-Hunting Cone Snails: Behavior, Neurobiology and Evolution. *Brain, Behavior and evolution* 86:58-74. doi: 10.1159/000438449
113. Shon KJ, Grilley MM, Marsh M, Yoshikami D, Hall AR, Kurz B, Gray WR, Imperial JS, Hillyard DR and Olivera BM (1995) Purification, characterization, synthesis, and cloning of the lockjaw peptide from *Conus purpurascens* venom. *Biochemistry* 34:4913-8. doi: 10.1021/bi00015a002
114. Moller C, Davis WC, Clark E, DeCaprio A and Mari F (2019) Conodipine-P1-3, the First Phospholipases A2 Characterized from Injected Cone Snail Venom. *Molecular and Cellular Proteomics* 18:876-891. doi: 10.1074/mcp.RA118.000972
115. Macrander J, Broe M and Daly M (2015) Multi-copy venom genes hidden in *de novo* transcriptome assemblies, a cautionary tale with the snakelocks sea anemone *Anemonia sulcata* (Pennant, 1977). *Toxicon* 108:184-188. doi: 10.1016/j.toxicon.2015.09.038
116. Kapon CA, Thapa P, Cabalteja CC, Guendisch D, Collier AC and Bingham JP (2013) Conotoxin truncation as a post-translational modification to increase the pharmacological diversity within the milked venom of *Conus magus*. *Toxicon* 70:170-8. doi: 10.1016/j.toxicon.2013.04.022
117. Van Lierop BJ, Robinson SD, Kompella SN, Belgi A, McArthur JR, Hung A, MacRaild CA, Adams DJ, Norton RS and Robinson AJ (2013) Dicarba alpha-conotoxin Vc1.1 analogues with

- differential selectivity for nicotinic acetylcholine and GABAB receptors. ACS Chemical Biology 8:1815-21. doi: 10.1021/cb4002393
118. Craik DJ and Adams DJ (2007) Chemical modification of conotoxins to improve stability and activity. ACS Chemical Biology 2:457-68. doi: 10.1021/cb700091j
119. Martinez JS, Olivera BM, Gray WR, Craig AG, Groebe DR, Abramson SN and McIntosh JM (1995) alpha-Conotoxin EI, a new nicotinic acetylcholine receptor antagonist with novel selectivity. Biochemistry 34:14519-26. doi: 10.1021/bi00044a030
120. Gomez MC, Aquino AMC, Matira AR, Alvarico RAD, Valbuena RE and Tayo LL Alpha-family of Conotoxins. ACM Press,
121. Fang G-M, Chen X-X, Yang Q-Q, Zhu L-J, Li N-N, Yu H-Z and Meng X-M (2018) Discovery, structure, and chemical synthesis of disulfide-rich peptide toxins and their analogs. Chinese Chemical Letters 29:1033-1042. doi: 10.1016/j.ccl.2018.02.002
122. Heimer P, Tietze AA, Bauml CA, Resemann A, Mayer FJ, Suckau D, Ohlenschlager O, Tietze D and Imhof D (2018) Conformational mu-Conotoxin PIIIA Isomers Revisited: Impact of Cysteine Pairing on Disulfide-Bond Assignment and Structure Elucidation. Analytical Chemistry 90:3321-3327. doi: 10.1021/acs.analchem.7b04854
123. Gomez MC, Alvarico RAD, Valbuena RE, Aquino AMC, Matira AR and Tayo LL (2019) In silico Protein Structure Comparison of Conotoxins with VI/VII Cysteine Framework. Proceedings of the 2019 3rd International Conference on Computational Biology and Bioinformatics, Association for Computing Machinery, Nagoya, Japan pp. 52–57
124. Franco A, Dovell S, Moller C, Grandal M, Clark E and Mari F (2018) Structural plasticity of mini-M conotoxins - expression of all mini-M subtypes by *Conus regius*. FEBS Journal 285:887-902. doi: 10.1111/febs.14372
125. Jacob RB and McDougal OM (2010) The M-superfamily of conotoxins: a review. Cellular and Molecular Life Sciences 67:17-27. doi: 10.1007/s00018-009-0125-0
126. Jacobsen RB, Jimenez EC, De la Cruz RG, Gray WR, Cruz LJ and Olivera BM (1999) A novel D-leucine-containing Conus peptide: diverse conformational dynamics in the contryphan family. Journal of Peptide Research 54:93-9. doi: 10.1034/j.1399-3011.1999.00093.x
127. Moller C and Mari F (2011) 9.3 KDa components of the injected venom of *Conus purpurascens* define a new five-disulfide conotoxin framework. Biopolymers 96:158-65. doi: 10.1002/bip.21406
128. Steiner DF, Peterson JD, Tager H, Emdin S, Ostberg Y and Falkmer S (1973) Comparative Aspects of Proinsulin and Insulin Structure and Biosynthesis. 13:591-604. doi: 10.1093/icb/13.3.591
129. Shoelson SE, Lu ZX, Parlautan L, Lynch CS and Weiss MA (1992) Mutations at the dimer, hexamer, and receptor-binding surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions. 31:1757-1767. doi: 10.1021/bi00121a025
130. Menting JG, Whittaker J, Margetts MB, Whittaker LJ, Kong GW, Smith BJ, Watson CJ, Žáková L, Kletvíková E, Jiráček J, Chan SJ, Steiner DF, Dodson GG, Brzozowski AM, Weiss MA,

- Ward CW and Lawrence MC (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* 493:241-245. doi: 10.1038/nature11781
131. Menting JG, Yang Y, Chan SJ, Phillips NB, Smith BJ, Whittaker J, Wickramasinghe NP, Whittaker LJ, Pandeyarajan V, Wan ZL, Yadav SP, Carroll JM, Strokes N, Roberts CT, Ismail-Beigi F, Milewski W, Steiner DF, Chauhan VS, Ward CW, Weiss MA and Lawrence MC (2014) Protective hinge in insulin opens to enable its receptor engagement. *Proceedings of the National Academy of Sciences* 111:E3395-E3404. doi: 10.1073/pnas.1412897111
132. Chance RE and Frank BH (1993) Research, development, production, and safety of biosynthetic human insulin. *Diabetes Care* 16 Suppl 3:133-42. doi: 10.2337/diacare.16.3.133
133. Howey DC, Bowsher RR, Brunelle RL and Woodworth JR (1994) [Lys(B28), Pro(B29)]-Human Insulin: A Rapidly Absorbed Analogue of Human Insulin. *Diabetes* 43:396-402. doi: 10.2337/diab.43.3.396
134. Garg SK, Rewers AH and Akturk HK (2018) Ever-Increasing Insulin-Requiring Patients Globally. *Diabetes Technology & Therapeutics* 20:S2-1-S2-4. doi: 10.1089/dia.2018.0101
135. Pertseva MN and Shpakov AO (2002). *Journal of Evolutionary Biochemistry and Physiology* 38:547-561. doi: 10.1023/a:1022008932029
136. Gregoire FM, Chomiki N, Kachinskas D and Warden CH (1998) Cloning and Developmental Regulation of a Novel Member of the Insulin-like Gene Family in *Caenorhabditis elegans*. *Biochemical and Biophysical Research Communications* 249:385-390. doi: 10.1006/bbrc.1998.9164
137. Smit AB, Van Kesteren RE, Li KW, Van Minnen J, Spijker S, Van Heerikhuizen H and Geraerts WPM (1998) Towards Understanding the Role of Insulin in the Brain: Lessons from Insulin-related Signaling Systems in the Invertebrate Brain. *Progress in Neurobiology* 54:35-54. doi: 10.1016/s0301-0082(97)00063-4
138. Ebberink RHM and Joosse J (1985) Molecular properties of various snail peptides from brain and gut. *Peptides* 6:451-457. doi: 10.1016/0196-9781(85)90413-9
139. Leever SJ (2001) Growth control: Invertebrate insulin surprises! *Current Biology* 11:R209-R212. doi: 10.1016/s0960-9822(01)00107-5
140. Nagasawa H, Kataoka H, Isogai A, Tamura S, Suzuki A, Mizoguchi A, Fujiwara Y, Suzuki A, Takahashi SY and Ishizaki H (1986) Amino acid sequence of a prothoracicotropic hormone of the silkworm *Bombyx mori*. *Proceedings of the National Academy of Sciences* 83:5840-5843. doi: 10.1073/pnas.83.16.5840
141. Smit AB, Vreugdenhil E, Ebberink RHM, Geraerts WPM, Klootwijk J and Joosse J (1988) Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. 331:535-538. doi: 10.1038/331535a0
142. Smit AB, Marle Av, Elk Rv, Bogerd J, Heerikhuizen Hv and Geraerts WPM (1993) Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*. *Journal of Molecular Endocrinology* 11:103-113. doi: 10.1677/jme.0.0110103

143. Brogiolo W, Stocker H, Ikeya T, Rintelen F, Fernandez R and Hafen E (2001) An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Current Biology* 11:213-221. doi: 10.1016/s0960-9822(01)00068-9
144. Antonova Y, Arik AJ, Moore W, Riehle MA and Brown MR (2012) 2 - Insulin-Like Peptides: Structure, Signaling, and Function. In: Gilbert LI (ed) *Insect Endocrinology*, Academic Press, San Diego pp. 63-92
145. McCowan C and Garb JE (2014) Recruitment and diversification of an ecdysozoan family of neuropeptide hormones for black widow spider venom expression. *Gene* 536:366-375. doi: 10.1016/j.gene.2013.11.054
146. Undheim EA, Grimm LL, Low CF, Morgenstern D, Herzig V, Zobel-Thropp P, Pineda SS, Habib R, Dziemborowicz S, Fry BG, Nicholson GM, Binford GJ, Mobli M and King GF (2015) Weaponization of a Hormone: Convergent Recruitment of Hyperglycemic Hormone into the Venom of Arthropod Predators. *Structure* 23:1283-92. doi: 10.1016/j.str.2015.05.003
147. Schnabel CA, Wintle M and Kolterman O (2006) Metabolic effects of the incretin mimetic exenatide in the treatment of type 2 diabetes. *Vascular Health and Risk Management* 2:69-77. doi: 10.2147/vhrm.2006.2.1.69
148. Safavi-Hemami H, Gajewiak J, Karanth S, Robinson SD, Ueberheide B, Douglass AD, Schlegel A, Imperial JS, Watkins M, Bandyopadhyay PK, Yandell M, Li Q, Purcell AW, Norton RS, Ellgaard L and Olivera BM (2015) Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proceedings of the National Academy of Sciences* 112:1743-1748. doi: 10.1073/pnas.1423857112
149. Cruz LJ, de Santos V, Zafaralla GC, Ramilo CA, Zeikus R, Gray WR and Olivera BM (1987) Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from *Conus geographus* and *Conus straitus* venoms. *J Biol Chem* 262:15821-4.
150. Möller C, Melaun C, Castillo C, Díaz ME, Renzelman CM, Estrada O, Kuch U, Lokey S and Marí F (2010) Functional Hypervariability and Gene Diversity of Cardioactive Neuropeptides. *Journal of Biological Chemistry* 285:40673-40680. doi: 10.1074/jbc.m110.171397
151. Robinson SD, Safavi-Hemami H, Raghuraman S, Imperial JS, Papenfuss AT, Teichert RW, Purcell AW, Olivera BM and Norton RS (2015) Discovery by proteogenomics and characterization of an RF-amide neuropeptide from cone snail venom. *Journal of Proteomics* 114:38-47. doi: 10.1016/j.jprot.2014.11.003
152. Safavi-Hemami H, Lu A, Li Q, Fedosov AE, Biggs J, Showers Corneli P, Seger J, Yandell M and Olivera BM (2016) Venom Insulins of Cone Snails Diversify Rapidly and Track Prey Taxa. *Molecular Biology and Evolution* 33:2924-2934. doi: 10.1093/molbev/msw174
153. Menting JG, Gajewiak J, MacRaild CA, Chou DH, Disotuar MM, Smith NA, Miller C, Erchegyi J, Rivier JE, Olivera BM, Forbes BE, Smith BJ, Norton RS, Safavi-Hemami H and Lawrence MC (2016) A minimized human insulin-receptor-binding motif revealed in a *Conus geographus* venom insulin. *Nat Struct Mol Biol* 23:916-920. doi: 10.1038/nsmb.3292
154. Ahorukomeye P, Disotuar MM, Gajewiak J, Karanth S, Watkins M, Robinson SD, Flórez Salcedo P, Smith NA, Smith BJ, Schlegel A, Forbes BE, Olivera B, Hung-Chieh Chou D and Safavi-

- Hemami H (2019) Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. *eLife* 8. doi: 10.7554/elife.41574
155. Xiong X, Menting JG, Disotuar MM, Smith NA, Delaine CA, Ghabash G, Agrawal R, Wang X, He X, Fisher SJ, Macrauld CA, Norton RS, Gajewiak J, Forbes BE, Smith BJ, Safavi-Hemami H, Olivera B, Lawrence MC and Chou DH-C (2020) A structurally minimized yet fully active insulin based on cone-snail venom insulin principles. *Nature Structural & Molecular Biology*. doi: 10.1038/s41594-020-0430-8
156. Shon K-J, Olivera BM, Watkins M, Jacobsen RB, Gray WR, Floresca CZ, Cruz LJ, Hillyard DR, Brink A, Terlau H and Yoshikami D (1998)  $\mu$ -Conotoxin PIIIA, a New Peptide for Discriminating among Tetrodotoxin-Sensitive Na Channel Subtypes. *The Journal of Neuroscience* 18:4473-4481. doi: 10.1523/jneurosci.18-12-04473.1998
157. Shon K-J, Stocker M, Terlau H, Stühmer W, Jacobsen R, Walker C, Grilley M, Watkins M, Hillyard DR, Gray WR and Olivera BM (1998)  $\kappa$ -Conotoxin Pviia Is a Peptide Inhibiting the Shaker K<sup>+</sup> Channel. *Journal of Biological Chemistry* 273:33-38. doi: 10.1074/jbc.273.1.33
158. Hopkins C, Grilley M, Miller C, Shon KJ, Cruz LJ, Gray WR, Dykert J, Rivier J, Yoshikami D and Olivera BM (1995) A new family of *Conus* peptides targeted to the nicotinic acetylcholine receptor. *Journal of Biological Chemistry* 270:22361-7. doi: 10.1074/jbc.270.38.22361
159. Mitchell SS, Shon KJ, Foster MP, Davis DR, Olivera BM and Ireland CM (1998) Three-Dimensional Solution Structure of Conotoxin  $\psi$ -P iiiie , an Acetylcholine Gated Ion Channel Antagonist  $\dagger$  ,  $\ddagger$ . *Journal of Biological Chemistry* 273:1215-1220. doi: 10.1021/bi972186t
160. Hoggard MF, Rodriguez AM, Cano H, Clark E, Tae H-S, Adams DJ, Godenschwege TA and Marí F (2017) In vivo and in vitro testing of native  $\alpha$ -conotoxins from the injected venom of *Conus purpurascens*. *Neuropharmacology* 127:253-259. doi: 10.1016/j.neuropharm.2017.09.020
161. Dowell C, Olivera BM, Garrett JE, Staheli ST, Watkins M, Kuryatov A, Yoshikami D, Lindstrom JM and McIntosh JM (2003)  $\alpha$ -Conotoxin PIA Is Selective for  $\alpha 6$  Subunit-Containing Nicotinic Acetylcholine Receptors. *Journal of Neuroscience* 23:8445-8452. doi: 10.1523/jneurosci.23-24-08445.2003
162. Möller C, Clark E, Safavi-Hemami H, Decaprio A and Marí F (2017) Isolation and characterization of Conohyal-P1, a hyaluronidase from the injected venom of *Conus purpurascens*. *Journal of Proteomics*. doi: 10.1016/j.jprot.2017.05.002
163. Safavi-Hemami H, Möller C, Marí F and Purcell AW (2013) High molecular weight components of the injected venom of fish-hunting cone snails target the vascular system. *Journal of Proteomics* 91:97-105. doi: 10.1016/j.jprot.2013.07.007
164. Irwin DM (2012) Origin and convergent evolution of exendin genes. *General and Comparative Endocrinology* 175:27-33. doi: 10.1016/j.ygcen.2011.11.025
165. Geraerts WP, Smit AB, Li KW and Hordijk PL (1992) The Light Green Cells of *Lymnaea*: a neuroendocrine model system for stimulus-induced expression of multiple peptide genes in a single cell type. *Experientia* 48:464-73. doi: 10.1007/bf01928165
166. Weiss MA and Lawrence MC (2018) A thing of beauty: Structure and function of insulin's "aromatic triplet". *Diabetes Obes Metab* 20 Suppl 2:51-63. doi: 10.1111/dom.13402

167. Schwartz GP, Burke GT and Katsoyannis PG (1987) A superactive insulin: [B10-aspartic acid]insulin(human). *Proceedings of the National Academy of Sciences* 84:6408-6411. doi: 10.1073/pnas.84.18.6408
168. Brange J, Ribel U, Hansen JF, Dodson G, Hansen MT, Havelund S, Melberg SG, Norris F, Norris K, Snel L and et al. (1988) Monomeric insulins obtained by protein engineering and their medical implications. *Nature* 333:679-82. doi: 10.1038/333679a0
169. Hansen BF, Kurtzhals P, Jensen AB, Dejgaard A and Russell-Jones D (2011) Insulin X10 revisited: a super-mitogenic insulin analogue. *Diabetologia* 54:2226-2231. doi: 10.1007/s00125-011-2203-8
170. Scapin G, Dandey VP, Zhang Z, Prosis W, Hruza A, Kelly T, Mayhood T, Strickland C, Potter CS and Carragher B (2018) Structure of the insulin receptor-insulin complex by single-particle cryo-EM analysis. *Nature* 556:122-125. doi: 10.1038/nature26153
171. Uchikawa E, Choi E, Shang G, Yu H and Bai X-C (2019) Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor–ligand complex. *eLife* 8. doi: 10.7554/elife.48630
172. Gutmann T, Schäfer I, Poojari C, Brankatschk B, Vattulainen I, Strauss M and Coskun Ü (2019) Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *bioRxiv*:679233. doi: 10.1101/679233
173. Papke RL and Lindstrom JM (2020) Nicotinic acetylcholine receptors: Conventional and unconventional ligands and signaling. *Neuropharmacology* 168:108021. doi: 10.1016/j.neuropharm.2020.108021
174. Hone AJ and McIntosh JM (2018) Nicotinic acetylcholine receptors in neuropathic and inflammatory pain. *FEBS Letters* 592:1045-1062. doi: 10.1002/1873-3468.12884
175. Luo S, Zhangsun D, Zhu X, Wu Y, Hu Y, Christensen S, Harvey PJ, Akcan M, Craik DJ and McIntosh JM (2013) Characterization of a novel  $\alpha$ -conotoxin TxID from *Conus textile* that potently blocks rat  $\alpha 3\beta 4$  nicotinic acetylcholine receptors. *Journal of medicinal chemistry* 56:9655-9663. doi: 10.1021/jm401254c
176. Quik M, Perez XA and Grady SR (2011) Role of  $\alpha 6$  nicotinic receptors in CNS dopaminergic function: relevance to addiction and neurological disorders. *Biochemical Pharmacology* 82:873-882. doi: 10.1016/j.bcp.2011.06.001
177. Brunzell DH (2012) Preclinical Evidence That Activation of Mesolimbic Alpha 6 Subunit Containing Nicotinic Acetylcholine Receptors Supports Nicotine Addiction Phenotype. 14:1258-1269. doi: 10.1093/ntr/nts089
178. Brunzell DH, McIntosh JM and Papke RL (2014) Diverse strategies targeting  $\alpha 7$  homomeric and  $\alpha 6\beta 2^*$  heteromeric nicotinic acetylcholine receptors for smoking cessation. *n/a-n/a*. doi: 10.1111/nyas.12421
179. Papke RL, Brunzell DH and De Biasi M (2020) *Cholinergic Receptors and Addiction*. Springer International Publishing, pp. 123-151

180. Papke RL, Dwoskin LP and Crooks PA (2007) The pharmacological activity of nicotine and norm nicotine on nAChRs subtypes: relevance to nicotine dependence and drug discovery. *Journal of Neurochemistry* 101:160-167. doi: 10.1111/j.1471-4159.2006.04355.x
181. Bagdas D, Gurun MS, Flood P, Papke RL and Damaj MI (2018) New Insights on Neuronal Nicotinic Acetylcholine Receptors as Targets for Pain and Inflammation: A Focus on  $\alpha 7$  nAChRs. *Current neuropharmacology* 16:415-425. doi: 10.2174/1570159X15666170818102108
182. Alexander JK, Sagher D, Krivoshein AV, Criado M, Jefford G and Green WN (2010) Ric-3 Promotes  $\gamma 7$  Nicotinic Receptor Assembly and Trafficking through the ER Subcompartment of Dendrites. *J Neurosci* 30:10112-10126. doi: 10.1523/jneurosci.6344-09.2010
183. Quiram PA (2000) Pairwise Interactions between Neuronal  $\alpha 7$  Acetylcholine Receptors and  $\alpha$ -Conotoxin PnIB. *J Biol Chem* 275:4889-4896. doi: 10.1074/jbc.275.7.4889
184. Whiteaker P, Marks MJ, Christensen S, Dowell C, Collins AC and McIntosh JM (2008) Synthesis and Characterization of 125I- $\alpha$ -Conotoxin Arib[V11L;V16A], a Selective  $\gamma 7$  Nicotinic Acetylcholine Receptor Antagonist. *J Biol Chem* 283:910-919. doi: 10.1074/jbc.M80136895
185. Luo S, Nguyen TA, Cartier GE, Olivera BM, Yoshikami D and McIntosh JM (1999) Single-Residue Alteration in  $\alpha$ -Conotoxin PnIA Switches Its nAChR Subtype Selectivity<sup>†</sup>. *Biochemistry* 38:14542-14548. doi: 10.1021/bi991252j
186. Celie PH, Kasheverov IE, Mordvintsev DY, Hogg RC, van Nierop P, van Elk R, van Rossum-Fikkert SE, Zhmak MN, Bertrand D, Tsetlin V, Sixma TK and Smit AB (2005) Crystal structure of nicotinic acetylcholine receptor homolog AChBP in complex with an  $\alpha$ -conotoxin PnIA variant. *Nat Struct Mol Biol* 12:582-8. doi: 10.1038/nsmb951
187. Ulens C, Hogg RC, Celie PH, Bertrand D, Tsetlin V, Smit AB and Sixma TK (2006) Structural determinants of selective  $\gamma$ -conotoxin binding to a nicotinic acetylcholine receptor homolog AChBP. *Proceedings of the National Academy of Sciences* 103:3615-3620. doi: 10.1073/pnas.0507889103
188. Heghinian MD, Mejia M, Adams DJ, Godenschwege TA and Marí F (2015) Inhibition of cholinergic pathways in *Drosophila melanogaster* by  $\alpha$ -conotoxins. *The FASEB Journal* 29:1011-1018. doi: 10.1096/fj.14-262733
189. Hopkins C, Grilley M, Miller C, Shon K-J, Cruz LJ, Gray WR, Dykert J, Rivier J, Yoshikami D and Olivera BM (1995) A New Family of *Conus* Peptides Targeted to the Nicotinic Acetylcholine Receptor. *Journal of Biological Chemistry* 270:22361-22367. doi: 10.1074/jbc.270.38.22361
190. López-Vera E, Jacobsen RB, Ellison M, Olivera BM and Teichert RW (2007) A novel  $\alpha$  conotoxin ( $\alpha$ -PIB) isolated from *C. purpurascens* is selective for skeletal muscle nicotinic acetylcholine receptors. *Toxicon* 49:1193-1199. doi: <https://doi.org/10.1016/j.toxicon.2007.02.007>
191. Dowell C, Olivera BM, Garrett JE, Staheli ST, Watkins M, Kuryatov A, Yoshikami D, Lindstrom JM and McIntosh JM (2003)  $\alpha$ -Conotoxin PIA Is Selective for  $\alpha 6$  Subunit-Containing Nicotinic Acetylcholine Receptors. *The Journal of Neuroscience* 23:8445-8452. doi: 10.1523/jneurosci.23-24-08445.2003

192. Xu M, Zhu X, Yu J, Yu J, Luo S and Wang X (2017) The crystal structure of Ac-AChBP in complex with  $\alpha$ -conotoxin LvIA reveals the mechanism of its selectivity towards different nAChR subtypes. *Protein & Cell* 8:675-685. doi: 10.1007/s13238-017-0426-2
193. Kehoe J and McIntosh JM (1998) Two Distinct Nicotinic Receptors, One Pharmacologically Similar to the Vertebrate  $\alpha$ 7-Containing Receptor, Mediate Cl Currents in *Aplysia* Neurons. *The Journal of Neuroscience* 18:8198-8213. doi: 10.1523/jneurosci.18-20-08198.1998
194. Utkin YN (2015) Animal venom studies: Current benefits and future developments. *World Journal of Biological Chemistry* 6:28. doi: 10.4331/wjbc.v6.i2.28
195. Lee H-K, Zhang L, Smith MD, Walewska A, Vellore NA, Baron R, McIntosh JM, White HS, Olivera BM and Bulaj G (2015) A marine analgesic peptide, Contulakin-G, and neurotensin are distinct agonists for neurotensin receptors: uncovering structural determinants of desensitization properties. *Frontiers in Pharmacology* 6. doi: 10.3389/fphar.2015.00011
196. Lau JL and Dunn MK (2018) Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* 26:2700-2707. doi: <https://doi.org/10.1016/j.bmc.2017.06.052>
197. Abalde S, Tenorio MJ, Afonso CML and Zardoya R (2018) Conotoxin Diversity in *Chelyconus ermineus* (Born, 1778) and the Convergent Origin of Piscivory in the Atlantic and Indo-Pacific Cones. *Genome Biology and Evolution* 10:2643-2662. doi: 10.1093/gbe/evy150
198. Nicke A, Loughnan ML, Millard EL, Alewood PF, Adams DJ, Daly NL, Craik DJ and Lewis RJ (2003) Isolation, structure, and activity of GID, a novel alpha 4/7-conotoxin with an extended N-terminal sequence. *J Biol Chem* 278:3137-44. doi: 10.1074/jbc.M210280200
199. Chi SW, Kim DH, Olivera BM, McIntosh JM and Han KH (2004) Solution conformation of alpha-conotoxin GIC, a novel potent antagonist of alpha3beta2 nicotinic acetylcholine receptors. *Biochem J* 380:347-52. doi: 10.1042/BJ20031792
200. Quinton L, Servent D, Girard E, Molgo J, Le Caer JP, Malosse C, Haidar el A, Lecoq A, Gilles N and Chamot-Rooke J (2013) Identification and functional characterization of a novel alpha-conotoxin (EIIA) from *Conus ermineus*. *Anal Bioanal Chem* 405:5341-51. doi: 10.1007/s00216-013-6926-x
201. Lopez-Vera E, Jacobsen RB, Ellison M, Olivera BM and Teichert RW (2007) A novel alpha conotoxin (alpha-PIB) isolated from *C. purpurascens* is selective for skeletal muscle nicotinic acetylcholine receptors. *Toxicon* 49:1193-9. doi: 10.1016/j.toxicon.2007.02.007
202. Pardos-Blas JR, Irisarri I, Abalde S, Tenorio MJ and Zardoya R (2019) Conotoxin Diversity in the Venom Gland Transcriptome of the Magician's Cone, *Pionoconus magus*. *Mar Drugs* 17:553. doi: 10.3390/md17100553
203. Li Q, Barghi N, Lu A, Fedosov AE, Bandyopadhyay PK, Lluisma AO, Concepcion GP, Yandell M, Olivera BM and Safavi-Hemami H (2017) Divergence of the Venom Exogene Repertoire in Two Sister Species of *Turriconus*. *Genome Biol Evol* 9:2211-2225. doi: 10.1093/gbe/evx157

204. Teichert RW, Jacobsen R, Terlau H, Yoshikami D and Olivera BM (2007) Discovery and characterization of the short kappaA-conotoxins: a novel subfamily of excitatory conotoxins. *Toxicon* 49:318-28. doi: 10.1016/j.toxicon.2006.10.001
205. Petrel C, Hocking HG, Reynaud M, Upert G, Favreau P, Biass D, Paolini-Bertrand M, Peigneur S, Tytgat J, Gilles N, Hartley O, Boelens R, Stocklin R and Servent D (2013) Identification, structural and pharmacological characterization of tau-CnVA, a conopeptide that selectively interacts with somatostatin sst3 receptor. *Biochem Pharmacol* 85:1663-71. doi: 10.1016/j.bcp.2013.03.019
206. Teichert RW, Jimenez EC and Olivera BM (2005) Alpha S-conotoxin RVIIIA: a structurally unique conotoxin that broadly targets nicotinic acetylcholine receptors. *Biochemistry* 44:7897-902. doi: 10.1021/bi047274+
207. Christensen SB, Bandyopadhyay PK, Olivera BM and McIntosh JM (2015) alphaS-conotoxin GVIIIB potently and selectively blocks alpha9alpha10 nicotinic acetylcholine receptors. *Biochem Pharmacol* 96:349-56. doi: 10.1016/j.bcp.2015.06.007
208. England LJ, Imperial J, Jacobsen R, Craig AG, Gulyas J, Akhtar M, Rivier J, Julius D and Olivera BM (1998) Inactivation of a serotonin-gated ion channel by a polypeptide toxin from marine snails. *Science* 281:575-8. doi: 10.1126/science.281.5376.575
209. Walker CS, Jensen S, Ellison M, Matta JA, Lee WY, Imperial JS, Duclos N, Brockie PJ, Madsen DM, Isaac JT, Olivera B and Maricq AV (2009) A novel *Conus* snail polypeptide causes excitotoxicity by blocking desensitization of AMPA receptors. *Curr Biol* 19:900-8. doi: 10.1016/j.cub.2009.05.017