

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Venomics and Functional Analysis of Venom From the Emerald Jewel Wasp, *Ampulex compressa*

Permalink

<https://escholarship.org/uc/item/1ww2p6mk>

Author

Arvidson, Ryan Scott

Publication Date

2016

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Venomics and Functional Analysis of Venom From the Emerald Jewel Wasp,
Ampulex compressa

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry and Molecular Biology

by

Ryan Scott Arvidson

June 2016

Dissertation Committee:

Dr. Michael E. Adams, Chairperson

Dr. Jason Stajich

Dr. Anandasankar Ray

Copyright by
Ryan Scott Arvidson
2016

The Dissertation of Ryan Scott Arvidson is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

Given the adage that “Science does not happen in a vacuum”, I would like to first acknowledge those that contributed to the data presented in this dissertation and who are contributing authors in manuscripts in which this data is to be published.

Victor Landa spent considerable time photographic and filming *A. compressa* in the lab which led to beautiful pictures, as is used in the introductory chapter, and an informative video that is available on the lab’s website (ampulex.ucr.edu). Victor also organized and obtained head capsule size data, and dissected mandibles from *A. compressa* larva and prepared them for electron microscopy. I would also like to thank Victor for his maintenance of the wasp colony. Sarah Frankenberg dissected and imaged cockroaches containing pupated wasps demonstrating that *A. compressa* larva is selective in which organs it consumes before pupating. Maayan Kaiser, who at the time of writing this dissertation was a Ph.D. student in the laboratory of Fredric Libersat, at the Ben-Gurion University, Beersheba, Israel, contributed her venom proteomics data in collaboration to generate the *A. compressa* venom presented in this dissertation. Haroun Mohammad contributed countless hours maintaining cell culture and developing and running assays attempting to characterize the effect of ampulexins on calcium dynamics *in vitro*, in particular on store-operated calcium entry via calcium add-back experiments. Jean Paul Urdena (J.P.) developed the cockroach brain injection protocol for the lab, and with great patience analyzed the effect of the tachykinins and ampulexins *in vivo*. J.P. also milked wasps for venom and helped maintain the insect and cell cultures. Chris Dail

assisted in milking of venom, dissection of venom glands and cockroach brains, injection of cockroach brains for *in vivo* studies and maintenance of the insect colonies. Elizabeth Duong assisted in cloning the cockroach tachykinin receptor, and the pUAST transformation plasmid for ampulexin 2 in drosophila, milked wasps for venom, and helped maintain the insect and cell cultures. Alexander Nguyen helped in cloning the cockroach tachykinin receptor. Alexander also milked wasps for venom and helped maintain the insect and cell cultures. Song Soo Lee, and Matt Mieselman for performing the behavioral assays of ampulexin 2 expression in drosophila, and James Hall for assistance in molecular modeling of ampulexin 2 dimer. Special thanks to Miguel Escalona for the ampulexin 2 dimer reaction scheme. I would like to thank Vladimir Kokoza and Alex Raikhel for the use of their beveling apparatus instrumental in successful cockroach cephalic ganglia injection experiments, and thank Peter Arensberger for his instruction in bioinformatics and for the initial assembly of the *Ampulex compressa* venom gland transcriptome.

I would like to thank SongQin Pan, Academic Coordinator, Proteomics at the Institute for Integrative Genome Biology who consistently and continuously helped in obtaining venom proteomics data, which are invaluable in writing this dissertation.

The image of the venom apparatus was taken and adapted for this dissertation with permission (Elsivier 2016) from “Venom and Dufour's glands of the emerald cockroach wasp *Ampulex compressa* (Insecta, Hymenoptera, Sphecidae): Structural and biochemical aspects”, Werner Gnatzy *et al.* 2016.

I would like to thank the UC Riverside department of Biochemistry, in particular Professors Miriam Ziegler and Thomas Baldwin, and Professor Paul Larsen for their encouragement and support in obtaining teaching assistantships. I would like to thank the department of Biochemistry for the Graduate Research Assistantship, Fall 2015. I would like to thank UC Riverside Graduate Division for the Graduate Research Mentorship Fellowship and the Distinguished Teaching Award.

I would also like to personally acknowledge those whose support made this dissertation, and my success as a graduate student possible. First, I cannot thank Professor Michael E. Adams enough for giving me the opportunity to do my thesis work in his laboratory. I would like to express my deep gratitude to my committee members: Anand Ray for his insights and support in my final year, and Jason Stajich, who greatly facilitated my success in bioinformatics analysis. Past and present Adams lab members – Robert Hice and Jason Higa, who taught me molecular biology, Sonali Deshpande, Do-Hyoung Kim, Yike Ding, and members of my cohort, Jackie Hubbard, Kosia McGovern, and Sephanie Dingwall - thank you all for your friendship and comradery.

No words can express my gratitude toward my mother Joanee Arvidson for a lifetime of unconditional love, support and encouragement, and my father Aron Arvidson for inspiration to become a Scientist and Philosopher. Leah Frankenberg for her support and hospitality and Lynn May Giovanni for inspiration and encouragement.

Above all, I would like to acknowledge and thank Sarah Frankenberg, the love and light of my life, you not only made my successes possible, but also meaningful.

Dedication

This Dissertation is a Dedication to the Love of Wisdom.

-- quod est inferius est sicut quod est superius --

ABSTRACT OF THE DISSERTATION

Venomomics and Functional Analysis of Venom From the Emerald Jewel Wasp,
Ampulex compressa

by

Ryan Scott Arvidson

Doctor of Philosophy, Graduate Program in Biochemistry and Molecular Biology
University of California, Riverside, June 2016
Dr. Michael E. Adams, Chairperson

My research involves biochemical analysis of venom from a fascinating parasitoid jewel wasp *Ampulex compressa*. Most parasitoid wasps envenomate the host by stinging into the body cavity to cause paralysis and developmental arrest, prior to deposition of eggs externally or within the body cavity. *A. compressa* instead uses a different subjugation strategy by injecting venom directly into the central nervous system, eliciting a behavioral sequence culminating in hypokinesia, a 7-10 day lethargy advantageous to wasp reproduction. Hypokinesia is a specific, venom-induced behavioral state characterized by suppression of the escape response and reduced spontaneous walking, leaving other motor functions unaffected. This specificity of action is particularly unique among venoms and interestingly, effects of the venom on the escape response are reversible as the cockroach may recover after 7-10 days if not consumed by the wasp larvae. Venom-induced hypokinesia raises an interesting biological question: How can

such a potent biochemical cocktail cause such long-lasting, specific, yet reversible effects on behavior? I approached this question in two ways: objective one - bioinformatic analysis of the venom and venom gland tissue to determine what the venom is made of, and objective two - functional analysis of key venom components to determine how the venom works. To address objective 1, I used advanced bioinformatics techniques to generate transcriptomes of the venom tissue and proteomes of the venom and venom tissue. Next generation sequencing of venom gland RNA has yielded full-length coding sequences and quantification of venom transcript levels, while mass spectroscopy based protein analysis has validated the presence of venom proteins. These analyses will allow construction of a comprehensive *A. compressa* “venome” that will help inform functional analyses of the venom and its role in hypokinesia induction. For objective two, I focused on the characterization of the most abundant peptide in the venom, tentatively named Ampulexin 1, and pharmacological analysis of an interesting venom peptide neurotransmitter, called tachykinin. Analysis of venom tachykinin action on cockroach brain receptors may reveal an interesting case of the evolution of a neurotransmitter from one animal to target the nervous system of another.

Table of Contents

Acknowledgements.....	iv
Dedication.....	vii
Abstract of the Dissertation.....	viii
List of Figures.....	xiii
List of Tables.....	xv
Forward.....	xvi

Chapter I. Exordium

Part One. Introduction to the Dissertation.....	2
Part Two. Literature Review of <i>Ampulex compressa</i>	13
Part Three. Venomics: Development and Current Perspectives.....	17
References Cited.....	23

Chapter II. Life History of *Ampulex compressa*

Abstract.....	34
Introduction.....	35
Materials and Methods.....	38
Results.....	41
Discussion.....	44
References Cited.....	47

**Chapter III. Transcriptomic and Proteomic Analysis of *Ampulex compressa*
Venom**

Abstract.....	59
Introduction.....	61
Materials and Methods.....	65
Results.....	70
Discussion.....	80
References Cited.....	87

**Chapter IV. Characterization and Functional Analysis of *Ampulex compressa*
Venom Tachykinin**

Abstract.....	107
Introduction.....	108
Materials and Methods.....	111
Results.....	115
Discussion.....	118
References Cited.....	122

Chapter V. Structural and Functional Analysis of Ampulexin Family Peptides

Introduction.....	134
Materials and Methods.....	138
Results and Discussion.....	140
References Cited.....	142

**Chapter VI. Differential Gene Expression Analysis of Cockroach Cerebral Ganglia,
Following Envenomation by *Ampulex compressa***

Introduction.....	148
Materials and Methods.....	152
Results and Discussion.....	153
References Cited.....	157

Chapter VII. Conclusions and Future Directions

Value of Venome Assembly.....	166
Ampulexin Peptide Family.....	167
Venom Tachykinins.....	168
Approaches to Functional Analysis of Venom Components.....	170
Perspectives on Venom Induced Hypokinesia.....	173
References Cited.....	177

List of Figures

<u>Figure</u>		<u>Page</u>
1-1	<i>A. compressa</i> attacks its prey	31
1-2	The venom apparatus of <i>A. compressa</i>	32
2-1	<i>A. compressa</i> development time is sexually dimorphic	51
2-2	<i>A. compressa</i> pupal and adult size is sexually dimorphic	52
2-3	<i>A. compressa</i> larva develops through three instars	53
2-4	Comparative morphological comparison of the mandibles of the three instars of <i>A. compressa</i>	54
2-5	Viscera of <i>P. americana</i> before and after parasitization by <i>A. compressa</i>	55
2-6	Life Cycle of <i>A. compressa</i>	56
2-7	Estimation of pupal size by modeling as a parabola	57
3-1	The venom apparatus of <i>A. compressa</i> is contractile	93
3-2	Experimental Design	94
3-3	Differential expression analysis between tissue subtypes of the venom apparatus	95
3-4	Bioinformatic analysis of the venom proteome	96
3-5	3D ‘gel’ analysis of the venom proteome	97
3-6	3D ‘gel’ analysis of the venom proteome by gland expression	98
3-7	Ampulexin peptide counts and signal cleavage	100
3-8	Gene ontology terms associated with venom proteins	101
3-9	Comparative genomic analysis of <i>A. compressa</i> venom proteins	102
3-10	Small molecule related genes in the venom apparatus	103

List of Figures (continued)

<u>Figure</u>	<u>Page</u>
3-11 Venom sac contents are acidic	104
3-12 A large, secreted, novel venom protein is completely disordered	105
4-1 <i>A. compressa</i> venom contains tachykinin precursor	127
4-2 <i>P. americana</i> cerebral ganglia express tachykinin and tachykinin receptor	128
4-3 Tachykinin receptor alignment	129
4-4 <i>P. americana</i> tachykinin receptor transmembrane domain and hydrophathy plot prediction	130
4-5 <i>A. compressa</i> venom tachykinins activate the cockroach brain tachykinin receptor <i>in vitro</i>	131
4-6 Venom tachykinin inhibits escape response <i>in vivo</i>	132
5-1 Ampulexins are amphipathic alpha-helices	144
5-2 Synthesis of ampulexin 2 dimer from monomer	145
5-3 Ampulexin 1 injected in the subesophageal ganglion effects short-term escape threshold	146
6-1 RNA sequencing data analysis	161
6-2 Differential expression in the cerebral ganglion post-sting	162

List of Tables

Table		<u>Page</u>
2-1	Size of larval body and head capsule by instar	50
4-1	Primers used in cloning <i>P. americana</i> tachykinin receptor.....	126
6-1	<i>P. americana</i> cerebral ganglia transcriptome assembly statistics.....	160
6-2	Synapse protein expression in stung and unstung cerebral ganglia	163
6-3	Differentially expressed genes common to both cerebral ganglia after the sting.....	164

Forward

This dissertation's first chapter is an introduction to the thesis, divided into three parts: the first describes the rationale and context for the scientific questions assessed herein, the second is a literature review of the jewel wasp *Ampulex compressa* and its interactions with its host *Periplaneta americana*, the third is focused on the application of next-generation nucleotide sequencing technology and advanced mass spectroscopy based proteomics in venom analysis, establishing "venomics" as a neologism in the scientific community. Chapters two, three, and four are formatted in academic manuscript format, where each contains an abstract, introduction, materials and methods, results, conclusion, and references. Chapters five and six are data chapters relevant to this study, but not intended as a manuscript *per se*; each with an introduction, materials and methods, and a combined results and discussion, with references.

Chapter I

Exordium

Part One. Introduction to the Dissertation

The endoparasitoid wasp *Ampulex compressa* injects venom directly into the central nervous system (CNS) of its host, the American cockroach, *Periplaneta americana*, to induce a long-term behavioral state known as hypokinesia (Haspel et al. 2003), (Libersat 2003). This renders the much larger host compliant, allowing the wasp to maneuver it physically into its burrow. When *A. compressa* first encounters its victim, it promptly and aggressively attacks, first stinging into the prothorax to cause short-term flaccid paralysis of the prothoracic legs (Moore et al. 2006). This facilitates subsequent stings into cephalic ganglia, i.e. brain and subesophageal ganglion (Figure 1-1) (Haspel et al. 2003), (Gal et al. 2010). The latter sting induces vigorous grooming behavior in the cockroach for approximately twenty minutes while the wasp readies its burrow (Weisel et al. 1999). The wasp then uses its large mandibles to grasp cockroach antennae and lead it into the burrow, whereupon a single egg is deposited externally on the leg of the host. Upon hatching, the wasp larva initially feeds on host hemolymph before entering its body cavity, whereupon it selectively consumes a subset of tissues prior to pupation (Arvidson et al.). Metamorphosis ensues over the next 30 days and the adult wasp emerges to complete the life cycle (Haspel et al. 2005). This work is focused on elucidating the biochemical and possible genetic basis for “zombie-like” hypokinesia induction (Gal et al. 2010).

Hypokinesia is a specific, venom-induced behavioral state characterized by suppression of the escape response and reduced spontaneous walking, leaving other motor functions unaffected. Stung, hypokinesic cockroaches remain able to walk, if

pulled, and fly in a wind tunnel (Gal et al. 2010). This specificity of action is particularly unique among venoms (Asgari et al. 2010). Interestingly, effects of the venom on the escape response are reversible. If the newly deposited egg is removed prior to hatching, the cockroach escape response slowly recovers, reaching normal levels within five to seven days. Initial envenomation of cephalic ganglia appears to down-modulate descending signals to the thoracic ganglia, suppressing selectively host escape behavior and leaving other motor circuits intact (Rosenberg et al. 2005). Furthermore, the venom lacks necrotic or lethal effects, instead preserving the host as a living, fresh food source for the wasp larva (Haspel et al. 2005).

The biochemical bases for some behavioral sequelae post-envenomation are described. For example, short-term paralysis of prothoracic legs, which occurs after the initial sting into the prothoracic ganglion, appears to be attributable to venom components GABA and GABA_A receptor agonists β -alanine and taurine (Moore et al. 2006). The vigorous grooming response induced by stings into cephalic ganglia may result from dopamine, another venom component; indeed dopamine injection into naïve cockroaches induces vigorous grooming (Weisel et al. 1999), (Banks 2010).

On the other hand, causes of venom-induced hypokinesia remain obscure. The present work is aimed at explaining how envenomation leads to this fascinating behavioral change. One possible clue lies in the observation that, although dopamine injection induces grooming in naïve animals, hypokinesic animals are resistant to dopamine-induced grooming (Weisel-Eichler et al. 2002). Significantly, recovery of resistance to dopamine-induced grooming follows the same time course of recovery from

hypokinesia (Weisel-Eichler et al. 2002). This suggests that the venom may inhibit dopamine signaling in some way. To address this possibility, the D2-like dopamine receptor from the cockroach CNS cloned and expressed with the intention of using it as a bioassay for screening the venom for receptor antagonists (Banks 2010). Previous to this, the most abundant peptide venom components were characterized using HPLC fractionation and MALDI-TOF, followed by Edman degradation and RACE (rapid amplification of cDNA ends). These analyses revealed the presence of several small peptides ranging in mass from 1-10 kDa (Moore 2003). One of the most abundant venom peptides, ampulexin 1, shows high sequence similarity to another venom peptide, ampulexin 2, which appears to exist in the venom as a dimer. Both ampulexin 1 and ampulexin 2 synthesized and bioassayed against cells expressing the cockroach D2-like dopamine receptor in luminescence assays (Button et al. 1993), (Dupriez et al. 2002). These components inhibit dopamine-induced calcium mobilization in a dose dependent way (Banks 2010). However further work showed equal efficacy of these peptides against endogenously expressed purinergic and muscarinic receptors, implying they may act on downstream signal transduction steps involved in calcium mobilization. Our current hypothesis is that these peptides may interfere with store-operated calcium entry.

Venom-induced hypokinesia raises an interesting biological question: How can the venom cocktail cause such long-lasting, specific, yet reversible effects on behavior? Generation of testable hypotheses to address the mode of venom action is facilitated by the characterization of the composition of the venom itself. Confounding elucidation of venom action is the fact that venom of *A. compressa* is a complex mixture of hundreds of

proteins and peptides ranging from 2 to 100 kDa, any of which might be critical or superfluous *a priori*. Additionally, attempts to reproduce sting-induced hypokinesia through injection of milked venom or venom gland extract into cerebral ganglia of cockroaches has failed.

The advent of next-generation nucleotide sequencing, e.g. Illumina, and advanced mass spectrometry based proteomics, along with an ever increasing catalog of bioinformatics analysis tools and software has allowed construction of a comprehensive *A. compressa* venome (Chapter 3), which will inform future functional analyses of the venom and its role in hypokinesia induction. Since no genome exists for *A. compressa*, a transcriptome of the venom apparatus was created to provide a database of open reading frames that served as a foundation for proteomics analysis of milked venom. Bottom-up proteomics typically yields small peptide fragments that are limited in information and can often make any identification of the parent protein ambiguous, whereas transcripts created via *de novo* transcriptome assembly are hypothetical reconstructs of RNA sequencing data that may not have a cognate sequence in the actual transcriptome. I found that construction of a comprehensive proteome of the venom apparatus using a bottom-up, shotgun approach from extracts of both venom apparatus and whole milked venom was greatly enhanced when analyzed against the transcriptome database. Combining these methods by assignment of mass spectrometry identified peptide onto hypothetical transcript database validates the sequence and presence of the transcript to which it aligned. These are further validated by having some depth in the RNA sequencing data and by PCR, when needed. This approach applied to *A. compressa*

venom has yielded a comprehensive analysis of the full length coding sequences of hundreds of venom proteins, enzymes and other components that contribute to its actions.

In addition to elucidating the venom proteome, combined transcriptomic and proteomic data can be used to assess the role of the individual glands of the venom apparatus in the manufacture of venom. The *A. compressa* venom apparatus is unusual in that the venom reservoir or sac is distinct from the venom gland, yet connected to the venom duct. The venom sac in this species is large and voluminous and could serve as a reservoir for the secreted fluid from the venom gland. The venom gland itself has a low volume, but a very large surface area with tubes branching throughout the abdomen, making it ideal for secretion, but not for storage. Venom generated by this gland therefore appears to be transferred to the venom sac for maintenance, processing, and perhaps concentration and acidification, so that the wasp is equipped to attack whenever a cockroach is encountered. This would require however, that the venom gland secretions must enter the common duct and travel up and into the venom sac. Analysis of both proteomes and transcriptomes of both the venom gland and venom sac reveal that the venom gland expresses venom proteins more highly than the venom sac for most, but not all of the venom proteins. This led me to hypothesize that venom is largely made in the venom gland and translocated into the venom sac, where it is supplemented with additional components from the venom sac epithelium. The venom sac is contractile appears capable of expelling the venom once the stinger reaches the appropriate target, such as the cockroach brain. Alternatively, the venom may be used, if rarely, for defensive purposes.

In the interest of defining the biochemical basis for hypokinesia, milked venom and its components have been tested in functional assays to evaluate actions at multiple levels from cell to behavior. Abundant venom peptides, ampulexin 1 and ampulexin 2, and several venom tachykinins were synthesized for functional analysis. Actions of ampulexin 1 and ampulexin 2 on cell culture have been tested previously, and results indicate they can inhibit GPCR-mediated calcium mobilization in cultured cells (Banks 2010). Although we observed this effect initially on the D2-like dopamine receptor, subsequent work revealed similar actions on purinergic and muscarinic receptors endogenous to CHO-K1 cells, leading us to hypothesize that a downstream step in receptor-induced calcium mobilization may be involved. Similar experiments in ovarian CHO, epithelial HEK293, and neuroblastoma/glioma NG108 cells reinforce this hypothesis, as these peptides show similar efficacies on these cell types. Results from calcium add-back experiments suggest these peptides may inhibit store-operated calcium entry. When calcium is liberated from internal stores by adding a neurotransmitter (e.g., ATP or carbachol), store-operated calcium (SOC) entry across the plasma membrane is triggered, as evidenced by a large luminescence signal when calcium is added back to the bathing medium. Pre-incubation of cells with ampulexin 1 and ampulexin 2 can block SOC entry. These results provide the first indication of putative SOC inhibitors identified in a natural venom, and point to a possible mechanism underlying hypokinesia. Unfortunately, at the time of writing, these results are not consistently reproducible, and so are presented here as an observation, rather than a data chapter.

One of the more exciting discoveries from the venomomics analysis described herein is the presence of a tachykinin precursor in milked venom. Sequences of tachykinin peptides present in this precursor show high sequence similarity to those previously described in two related cockroach species (Predel et al. 2005). I therefore undertook an investigation into the possible functional significance of this peptide precursor in generation of hypokinesia. The wasp venom preprotachykinin sequence contains nine tachykinin-related peptides with four distinct sequences, some of which are repeated. As the cockroach tachykinin receptor is the presumed target of these peptides, the cockroach tachykinin receptor was cloned, and its sensitivity to *A. compressa* venom tachykinins was assessed using established luminescence receptor assays. The tachykinin precursor *per se* does not activate the receptor with high affinity. Interestingly, the venom contains proteases, one of which, an endothelin converting enzyme-like endopeptidase, a proenzyme convertase, is known to process peptides from precursors, and furin protease is known to cleave proteins at dibasic residues, which flank the tachykinins. These proteases could be involved in cleaving venom tachykinins from the precursor. Alternatively, enzymes in the cockroach brain may process this precursor following envenomation. This precursor therefore could act as a time-release particle that gradually releases tachykinins to interfere with normal cockroach CNS function. The putative venom protease may affect cockroach brain chemistry through disruption of endogenous cockroach brain neuropeptides. In addition to cleavage of tachykinin peptides from the precursor, tachykinins require C-terminal amidation for activity, requiring a peptidyl-glycine alpha-amidating monooxygenase –like activity in either the venom, or in the

cerebral ganglia. Furthermore, direct injection of tachykinin peptides cockroach SEG causes short term increase in escape threshold. Venom tachykinin is a compelling target for functional analysis as it is a well characterized neurotransmitter, have been found in other animal venoms, and tachykinins are implicated in a broad range of behaviors, including locomotion, and recently in social aggression in *Drosophila* (Winther et al. 2006), (Asahina et al. 2014).

One of the striking features of hypokinesia is its five to seven day-long duration. This persistent feature of venom action could involve time-release of neuropeptides, as hypothesized above with respect to discovery of a tachykinin precursor in the venom. It also could involve enzymatic action of venom proteases acting to degrade endogenous signaling molecules or extracellular matrix in host cephalic ganglia. In addition, it is important to consider possible changes in gene expression patterns in cockroach cephalic ganglia due to envenomation. RNA sequencing of cockroach cephalic ganglia prior to envenomation and 24 hours post envenomation served to generate a more comprehensive transcriptome assembly, particularly for cephalic ganglia specific transcripts, and allow quantification and differential expression of genes. By comparing gene expression patterns of stung versus both unstung and recovered animals, we may be able to identify genes specifically associated with hypokinesia and its reversibility. Although many of the genes identified as differentially expressed are novel, several stand out as potentially interesting. One is bruchpilot, a synaptic protein involved in neurotransmitter release (Paul et al. 2015). Others include response to oxidative stress proteins, and proteins involved in metabolism. No whole genome assembly is yet available for *P. americana*,

and only a few transcript sequences can be downloaded from public databases, though there is one unannotated whole body transcriptome available from the one-thousand-insect-transcriptome initiative (1KITE). Additionally, this database of transcripts with quantification and differential expression is useful for proteomic studies of changes in cockroach CNS following envenomation.

One approach to functional analysis of venom proteins involves direct injection of whole venom and venom components into the cockroach subesophageal ganglion for assessment of behavioral changes. Ganglion injections have been previously used to search for *A. compressa* venom components responsible for transient paralysis of the prothoracic legs (Moore et al. 2006). These experiments led to discovery of GABA and GABA_A receptor agonists taurine and beta-alanine in the venom that contribute to the initial temporary paralysis of the prothoracic legs following the initial sting (Moore et al. 2006). I will describe behavioral analysis using a circular arena and a pulse generator to quantify spontaneous locomotory and escape responses (Chapter 4 and 5). Synthetic venom components, or other pharmacological agents can be injected into the subesophageal ganglion using fine micropipette needles attached to a hydraulic injector. The SEG is readily accessible in the ventral neck region by penetrating soft cuticle connecting the head to the prothoracic segment. Comparing the effect on behavior between venom- and saline-injected cockroaches will provide information on the components of the venom that may be sufficient or responsible for hypokinesia.

Artificial or synthetically induced hypokinesia, lasting 24 hours or longer, has not been reported. This includes attempts to inject milked venom and venom gland extracts

into both cerebral ganglia of the cockroach. This aspect is simultaneously the most compelling and frustrating in the functional biochemical analysis of *A. compressa* venom.

This raises fundamental questions as to the envenomation mechanism; is the venom itself sufficient to cause hypokinesia, but any attempts to harvest and process it for artificial injection render it inert? If this is the case, it is possible that oxidation or protease activity would be responsible for the degradation of venom components. During the milking process, venom is diluted when absorbed into the milking solution, and the change in ionic strength or pH may cause the venom to become inactive. Nonetheless, there is still the outside possibility that there some critical component outside of the milked venom or venom apparatus that has not yet been identified that is responsible for the observed effect of the venom.

I hypothesize that hypokinesia is caused by concerted action of both peptide toxins and venom enzymes to interfere with normal brain function that disrupts synaptic integration of information, abolishing the cockroach escape response. This action would target two distinct command centers for locomotion. The first and most obvious is the brain. Venom components that target the brain are multitudinous, and target brain activity at different time frames. For example, the GABA receptor agonists in the venom, and perhaps other components in the small molecular weight fraction, may disrupt synaptic transmission briefly, after which a second wave of activity causes a short to intermediate term hypokinesia. This latter action could involve release of now-mature tachykinins activating inhibitory neurons, or activity of phospholipase A2, releasing fatty-acid paracrine signals causing synaptic blockade, similar to what has been reported in snake

venoms (Paoli et al. 2009). Long term hypokinesia may be caused by disruption of extracellular matrix proteins required for synaptic integrity. This would require days to repair, during which interruption of sensory integration prevents normal locomotory activity (Kochlamazashvili et al. 2010).

Part Two. Literature Review of *Ampulex compressa*

Ampulex compressa has been the object of study for over half a century owing to its fascinating hunting behavior. Early reports its cockroach hunting prowess describe a pugnacious insect from New Caledonia, whereby the female attacks large cockroach species such as *Periplaneta americana*, and *Periplaneta australasiae*, (American and Australian cockroaches respectively) to exploit them as a living food supply for her offspring (Williams 1942). *P. americana* has been described as the exclusive host for *A. compressa* in the literature. Despite *P. americana*'s size and strength advantage, *A. compressa* is well-suited as a parasitoid, and was viewed as an ideal biological control strategy in Hawaii, since as *P. Americana* was an invasive pest. Consequently, *A. compressa* was introduced to Kauai, Maui and Oahu in 1941 (Williams 1942). Even though *A. compressa* is well suited to the pacific island ecosystem, its slow life cycle and relatively low fecundity made it ultimately unsuitable as a biological control for *P. americana* in Hawaii.

The unusual effect of the venom on *P. americana* is compelling. A successful hunt results in modified behavior of the host, but not paralysis. This makes venom of *A. compressa* a tempting target of neuropharmacological research. One of the first critical accounts of *A. compressa* venom action described block of synaptic transmission (Piek et al. 1989). This block was explained by presence of GABA and GABA analogs in the venom. However, this does not explain the gradual onset and prolonged condition of hypokinesia, indicating that venom action is complex (Piek et al. 1989), (Moore et al. 2006). The venom apparatus was described as being highly branched and bifurcated, with

no venom reservoir (Piek 1986), (Piek et al. 1989). The lack of a described venom reservoir led to description of a large sac-like organ attached to the *ductus venatus*, or common duct as a Dufour's gland. Preliminary analysis of venom and venom apparatus protein extracts by polyacrylamide gel electrophoresis shows that the venom sac contains a similar protein profile with the milked venom, especially in the small molecular weight fraction, and many common bands with the venom gland in the large molecular weight fraction (Moore 2003). The chemical similarity of the venom sac and milked venom was recently corroborated, and further, the Dufour's gland was identified as a small diminutive organ attached to the body wall with chemical characteristics distinct from the venom (Gnatzy et al. 2015) (Figure 1-2). These observations and data in the present work establish that the venom apparatus of *A. compressa* is composed of two connected but distinct venom-producing organs: the venom sac or reservoir, and the venom gland. I sought to perform proteomic analysis of each independently to provide a comprehensive view of the venom *per se* and the mechanism of envenomation.

When female *A. compressa* encounters a cockroach, she flanks the cockroach swiftly, grabbing the it by the pronotum and delivering a swift sting to the thorax, temporarily paralyzing the front legs (Williams 1942), (Haspel et al. 2003), (Moore et al. 2006). Once the cockroach is prone, stings are directed precisely into the head capsule through the soft membrane on the neck; these additional stings are responsible for induction of hypokinesia (Haspel et al. 2003), (Gal et al. 2010).

During the first moments after the sting the cockroach recovers from the encounter, but instead of escaping, it engages in vigorous, protracted grooming. This bout of

grooming has been attributed to dopamine in the venom (Weisel et al. 1999), (Banks 2010). The full effect of the venom can be observed about 60-90 min post-sting, at which time the wasp can handily lead the hypokinesic cockroach into a nearby burrow for egg deposition.

The original description of *A. compressa* venom action was described as a permanent decrease in activity of the subesophageal ganglion (Piek 1986). In contrast, current literature and present observations actually describe hypokinesia as reversible. Within 5-7 days post-envenomation, cockroaches again exhibit spontaneous walking, while escape threshold returns to pre-stung levels. Significantly, this interval coincides with time to pupation of the wasp larva. Naive wasps, i.e wasps that have yet to sting a cockroach, are less successful in inducing long lasting hypokinesia as experienced wasps (unpublished observations). However, the hunting behavior of *A. compressa* is innate. Envenomation of the head capsule targets both the head ganglia - subesophageal ganglion and brain - with little leakage into surrounding tissue (Haspel et al. 2003). Apparently, *A. compressa* identifies target ganglia as opposed to surrounding tissue by sensing differences in density through mechanoreceptors in the stinger (Gal et al. 2014). In a real sense, the wasp may feel and even taste the cockroach CNS during envenomation.

The central complex has been shown to control locomotory behavior in cockroaches, making it a suspected target for the wasp's venom (Guo et al. 2013). Recent examination of CNS targeting necessary or sufficient for hypokinesia induction revealed that either the brain central complex or the subesophageal ganglion is sufficient to cause hypokinesia

(Kaiser et al. 2015). This suggests that envenomation of both cerebral ganglia by the wasp is redundant.

The wasp larva develops outside the cockroach during the first two instars, while the third instar and pupa develop inside its body cavity. During development, *A. compressa* larva secrete large doses of antibiotics - mostly γ -lactones and isocoumarins - into the cockroach to preserve the food supply (Herzner et al. 2013). Cockroaches are prone to infections from bacteria and protozoans as well as nematodes, which could be deleterious to the developing wasp larva (Faulhaber et al. 1992), (Sasmal et al. 2008), (Morton et al. 2013). This sanitization is sophisticated in that secreted antibiotics are broad spectrum, persistent until wasp eclosion, and the anti-microbial (r)-(-)-mellein is volatile, effectively “fumigating” against opportunistic infections (Weiss et al. 2014).

A. compressa venom not only incapacitates the host, but confers advantages to its developing progeny. For example, host envenomation slows its metabolism so as to retain body mass, thus providing additional nutrients to the developing wasp larva, and even protecting the cockroach against starvation and desiccation during the larva developmental period (Haspel et al. 2005).

Part Three. Venomics: Development and Current Perspectives

The action of *A. compressa* venom is contrasting to common ideas of wasp venoms. Most hymenopteran-human encounters occur with social wasps (Vespoidea), bees (Apoidea), and ants (Formicidae). These social hymenoptera have developed venoms for defense, causing pain and inflammation in the target animal. Hymenoptera venoms are responsible for significant morbidity and mortality, therefore a great deal of venom research has gone to understanding the effects of venom in this context (Vetter et al. 1998), (Fitzgerald et al. 2006), (Matysiak et al. 2015). Indeed, proteomics analyses of hymenoptera venoms typically focus on allergy and inflammation in stung humans, implicating immunoglobulin E in mediation of allergic response to venom (dos Santos Pinto et al. 2012), (Resende et al. 2013), (Sookrung et al. 2014), (Van Vaerenbergh et al. 2015). Allergens identified in venoms seem to be ubiquitous in proteomics analysis and the current work describes several known allergens, including the aptly named venom allergen 3, phospholipase A2, metallopeptidase, and hyaluronidase (Hoffman 1995), (Justo Jacomini et al. 2013), (Palm et al. 2013). Though *A. compressa* rarely stings for defense, it is unclear if any of these conserved venom components are defensive in nature. It is thought that venom of solitary wasps would necessarily differ in content from social wasps since the former is used to capture and incapacitate prey with components such as kinins, bradykinin, mastoparan, and cationic, amphipathic, alpha-helical peptides, similar in description to ampulexin 1 (Konno et al. 2016), (Lee et al. 2016).

Another important focus in proteomics analysis of venoms is drug discovery. The modern approach to venom analysis has yielded an unprecedented amount of new gene sequence and expression information pertaining to venom proteins and peptides. However, functional analysis of venom components is required before therapeutic applications can be envisioned. Some examples of drugs derived from venoms include ziconotide, a N-type calcium channel blocker derived *Conus magus* venom useful for treatment of chronic pain and chlorotoxin from scorpion venom that targets glioblastoma cells (Miljanich 2004), (Mamelak et al. 2007), (Escoubas et al. 2009). Additional characterization of novel venom proteins, peptides, and small molecules will likely uncover new therapeutics and/or biologically relevant tools, allowing more sophisticated analysis of basic biological questions (Danneels et al. 2010).

High throughput mass spectrometry and RNA sequencing have revolutionized biochemical analysis of venoms. Prior to the advent of these technologies, venom proteomics yielded identification of relatively small numbers of proteins, usually less than one hundred. A major obstacle has been the diversity of low-abundance components, making fractionation and analysis difficult. Because of this, only the most highly abundant venom components were characterized. This is less of a problem for analysis of larger venomous animals such as snakes, though application of modern high-throughput techniques have greatly facilitated the profiling of venomous reptile proteomes as well (Vejayan et al. 2014).

Since proteomes, including venoms, are complex mixtures of hundreds of proteins, advances in protein identification technology were necessary. Coupling liquid

chromatography in-line with tandem mass spectroscopy (LC-MS/MS) allowed for separation of complex mixtures before introduction into the mass ionizer and subsequently the detector. Since analysis of large intact proteins is limited by inability of larger mass proteins to ionize, it is first necessary to proteolytically digest the sample into small peptides, typically with trypsin, a procedure referred to as trypsinizing (Shen et al. 2002), (Karpievitch et al. 2010), (Walther et al. 2010). These peptides are then separated by chromatography before passing into the ionizer, commonly electrospray ionization, or ESI, where an aerosol of the samples is propelled in a vacuum to the mass analyzer, and as the solvent evaporates, the peptide is imparted a charge (Shen et al. 2002). Introduction of an additional dimension of chromatography, where usually both dimensions have a reversed-phase stationary phase, then an acidic mobile phase for the first dimension and a basic mobile phase second dimension, or vis versa, can further separate peptides before introduction into the mass analyzer, allowing better resolution and therefore better identification of individual proteins (Zhang et al. 2010). Increases in the scanning speed and sampling rate of mass spectrometers, such as in the hybrid LTQ-Oribitrap, also greatly improves resolution; combining the two greatly facilitates discovery of low abundant proteins in complex mixtures. Once the mass spectra of a tryptic peptide have been obtained, the peptide must be assigned to its parent precursor. This is accomplished through search engine algorithms such as SEAQUEST and MASCOT, where experimental spectra are compared to theoretical spectra derived from a database of protein sequences. When a user-determined statistically significant match is obtained, the parent protein is considered identified (Eng et al. 1994), (Perkins et al.

1999), (Paulo 2013). This technique also can be applied to non-proteolytically treated samples to enrich and identify small peptides, where proteolysis would make fragments too small for detection or analysis. The combination of these advances in chromatography, mass spectrometry, and systems biology is referred to as MudPIT, for multiple-dimension protein identification technology (Schirmer et al. 2003), (Graumann et al. 2004), (Chen et al. 2006). MudPIT is a sensitive, and reproducible method for analysis of whole proteomes, often in one experiment, resolving several thousand proteins in a mixture down to the attomole level (Comunian et al. 2011), (Roth et al. 2011).

A critical caveat to the power of the MudPIT approach is the requirement for a database containing sequences of every protein in the sample to be analyzed (Washburn 2004). This problem is mitigated in model organisms whose genome has been sequenced. However, the utility of MudPIT is severely compromised if sequences of the target proteins are unknown. Along with proteomics, nucleotide sequencing has undergone a revolution with the introduction of cost-effective next generation sequencing. Illumina sequencing has become the standard approach from *de novo* sequencing of whole transcriptomes, especially for non-model organisms (Surget-Groba et al. 2010), (Feldmeyer et al. 2011), (Li et al. 2014). Illumina platforms take as input fragmented cDNA, which is sequenced in short reads of 150 base-pair maximal, but 50 to 100 base pairs more commonly, in massively parallel clusters, overall yielding 200 to 400 million reads. These data, usually in files of tens to hundreds of gigabytes must be analyzed on computational clusters, by skilled bioinformaticians, using bioinformatics pipelines that

are still being vetted (Gongora-Castillo et al. 2013). Despite the difficulties, next generation sequencing data has been used successfully to generate many *de novo* transcriptome assemblies in a wide range of organisms.

Several programs have been created to assemble short read sequencing data into putative transcripts, including Trinity, Trans-ABYSS, SOAPdenovo-Trans, and Velvet/Oases (Robertson et al. 2010), (Schulz et al. 2012), (Haas et al. 2013), (Xie et al. 2014). However, publicly available short read assemblers are prone to error, generating false positives and incomplete or chimeric transcripts. Much effort has gone into optimization and benchmarking these pipelines to determine the best program for production of a universal algorithm to produce a high quality transcriptome from any RNA-seq dataset (Surget-Groba et al. 2010), (Amin et al. 2014), (Rana et al. 2016). Each assembler has strengths and weaknesses compared to its siblings and some are more amenable to certain raw datasets than others. One creative approach has been to leverage the strengths of many assemblers, and filter out redundant sequences. This approach has been made available in the EvidentialGene - tr2aacds pipeline, which recommends merging assemblies of multiple programs to generate a highly redundant sequence file of millions to tens of millions of sequences, then filters out all redundant sequences, with a focus on longest coding sequences (Gilbert 2013).

An important consideration when quality of assessing a *de novo* assembly is to determine its completeness, without knowing transcript number and sequence *a priori*. An approach to answer this is to assess how many genes that are expected to be present are extant and complete. The programs CEGMA and BUSCO approach this by

interrogating the assembly for ultra-conserved genes that are expected to be expressed in any metazoan (Parra et al. 2007), (Simao et al. 2015). The existence and completeness of these ultra-conserved genes is then a proxy for existence and completeness of all sequences in the transcriptome. A high quality transcriptome can serve as a database for MudPIT in non-model systems, greatly facilitating the discovery of new proteins.

Taken together, transcriptomic analysis can act as a surrogate genome and as a method for quantifying expression of transcripts. Proteomics can confirm existence of a transcript and, when applied to a complex mixture such as a venom, can distinguish a venom transcript from a house-keeping transcript. Since it is now possible to fully profile the protein composition of a venom, this focus of study can rightly be called venomics. Usage of venomics as a term is not without criticism, since it was considered a badomics word (Eisen 2012). I would argue that even according to the five-point formalization laid out in this criticism, venomics is suitably comprehensive and unique in that a venom is different from a venom gland proteome, transcriptome, or genome, and is not used tongue-in-cheek. Although limited in scope as a subset of the total proteome, it is no different than a tissue-specific transcriptome in that respect. Additionally, it would appear that use of the term venomics has become accepted by the scientific community at large, since it has been used in hundreds of published articles describing analysis of venom proteomes.

References Cited

- Amin, S., et al. (2014). "Assembly and annotation of a non-model gastropod (*Nerita melanotragus*) transcriptome: a comparison of de novo assemblers." BMC Res Notes **7**: 488.
- Arvidson, R., et al. "Life history of the emerald jewel wasp, *Ampulex compressa*." in preparation.
- Asahina, K., et al. (2014). "Tachykinin-expressing neurons control male-specific aggressive arousal in *Drosophila*." Cell **156**(1-2): 221-235.
- Asgari, S. and D. B. Rivers (2010). "Venom proteins from endoparasitoid wasps and their role in host-parasite interactions." Annual Review of Entomology **56**(1): 313-335.
- Banks, C. N. (2010). The Roles of Biogenic Amines and Dopamine Receptors in Envenomation by the Parasitoid Wasp *Ampulex compressa*. Ph.D., University of California.
- Button, D. and M. Brownstein (1993). "Aequorin-expressing mammalian cell lines used to report Ca²⁺ mobilization." Cell Calcium **14**(9): 663-671.
- Chen, E. I., et al. (2006). "Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT)." Mol Cell Proteomics **5**(1): 53-56.
- Comunian, C., et al. (2011). "A comparative MudPIT analysis identifies different expression profiles in heart compartments." Proteomics **11**(11): 2320-2328.
- Danneels, E. L., et al. (2010). "Venom proteins of the parasitoid wasp *Nasonia vitripennis*: recent discovery of an untapped pharmacopee." Toxins (Basel) **2**(4): 494-516.
- dos Santos Pinto, J. R., et al. (2012). "Proteomic view of the venom from the fire ant *Solenopsis invicta* Buren." J Proteome Res **11**(9): 4643-4653.

- Dupriez, V. J., et al. (2002). "Aequorin-based functional assays for G-protein-coupled receptors, ion channels, and tyrosine kinase receptors." Receptors Channels **8**(5-6): 319-330.
- Eisen, J. A. (2012). "Badomics words and the power and peril of the ome-meme." Gigascience **1**(1): 6.
- Eng, J. K., et al. (1994). "An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database." J Am Soc Mass Spectrom **5**(11): 976-989.
- Escoubas, P. and G. F. King (2009). "Venomics as a drug discovery platform." Expert Rev Proteomics **6**(3): 221-224.
- Faulhaber, L. M. and R. D. Karp (1992). "A diphasic immune response against bacteria in the American cockroach." Immunology **75**(2): 378-381.
- Feldmeyer, B., et al. (2011). "Short read Illumina data for the de novo assembly of a non-model snail species transcriptome (*Radix balthica*, Basommatophora, Pulmonata), and a comparison of assembler performance." BMC Genomics **12**: 317.
- Fitzgerald, K. T. and A. A. Flood (2006). "Hymenoptera stings." Clin Tech Small Anim Pract **21**(4): 194-204.
- Gal, R., et al. (2014). "Sensory arsenal on the stinger of the parasitoid jewel wasp and its possible role in identifying cockroach brains." PLoS One **9**(2): e89683.
- Gal, R. and F. Libersat (2010). "On predatory wasps and zombie cockroaches: Investigations of "free will" and spontaneous behavior in insects." Communicative & Integrative Biology **3**(5): 458-461.
- Gal, R. and F. Libersat (2010). "A wasp manipulates neuronal activity in the sub-oesophageal ganglion to decrease the drive for walking in its cockroach prey." PLoS ONE **5**(4): e10019.
- Gilbert, D. (2013). Gene-omes built from mRNA seq not genome DNA. 7th annual arthropod genomics symposium, Notre Dame.

- Gnatzy, W., et al. (2015). "Venom and Dufour's glands of the emerald cockroach wasp *Ampulex compressa* (Insecta, Hymenoptera, Sphecidae): structural and biochemical aspects." Arthropod Struct Dev **44**(5): 491-507.
- Gongora-Castillo, E. and C. R. Buell (2013). "Bioinformatics challenges in de novo transcriptome assembly using short read sequences in the absence of a reference genome sequence." Nat Prod Rep **30**(4): 490-500.
- Graumann, J., et al. (2004). "Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast." Mol Cell Proteomics **3**(3): 226-237.
- Haas, B. J., et al. (2013). "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nat Protoc **8**(8): 1494-1512.
- Haspel, G., et al. (2005). "Parasitoid wasp affects metabolism of cockroach host to favor food preservation for its offspring." Journal of Comparative Physiology A **191**(6): 529-534.
- Haspel, G., et al. (2003). "Direct injection of venom by a predatory wasp into cockroach brain." Journal of neurobiology **56**(3): 287-292.
- Haspel, G., et al. (2003). "Direct injection of venom by a predatory wasp into cockroach brain." **56**: 287-292.
- Herzner, G., et al. (2013). "Larvae of the parasitoid wasp *Ampulex compressa* sanitize their host, the American cockroach, with a blend of antimicrobials." Proc Natl Acad Sci U S A **110**(4): 1369-1374.
- Hoffman, D. R. (1995). "Fire ant venom allergy." Allergy **50**(7): 535-544.
- Justo Jacomini, D. L., et al. (2013). "Hyaluronidase from the venom of the social wasp *Polybia paulista* (Hymenoptera, Vespidae): Cloning, structural modeling, purification, and immunological analysis." Toxicon **64**: 70-80.
- Kaiser, M. and F. Libersat (2015). "The role of the cerebral ganglia in the venom-induced behavioral manipulation of cockroaches stung by the parasitoid jewel wasp." J Exp Biol **218**(Pt 7): 1022-1027.

- Karpievitch, Y. V., et al. (2010). "Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects." Ann Appl Stat **4**(4): 1797-1823.
- Kochlamazashvili, G., et al. (2010). "The extracellular matrix molecule hyaluronic acid regulates hippocampal synaptic plasticity by modulating postsynaptic L-type Ca(2+) channels." Neuron **67**(1): 116-128.
- Konno, K., et al. (2016). "Peptide Toxins in Solitary Wasp Venoms." Toxins (Basel) **8**(4).
- Lee, S. H., et al. (2016). "Differential Properties of Venom Peptides and Proteins in Solitary vs. Social Hunting Wasps." Toxins (Basel) **8**(2).
- Li, B., et al. (2014). "Evaluation of de novo transcriptome assemblies from RNA-Seq data." Genome Biol **15**(12): 553.
- Libersat, F. (2003). "Wasp uses venom cocktail to manipulate the behavior of its cockroach prey." Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology **189**(7): 497-508.
- Mamelak, A. N. and D. B. Jacoby (2007). "Targeted delivery of antitumoral therapy to glioma and other malignancies with synthetic chlorotoxin (TM-601)." Expert Opin Drug Deliv **4**(2): 175-186.
- Matysiak, J., et al. (2015). "Influence of honeybee sting on peptidome profile in human serum." Toxins (Basel) **7**(5): 1808-1820.
- Miljanich, G. P. (2004). "Ziconotide: neuronal calcium channel blocker for treating severe chronic pain." Curr Med Chem **11**(23): 3029-3040.
- Moore, E. L. (2003). A Biochemical and Molecular Analysis of Venom with Distinct Physiological Actions from Two Arthropod Sources: The Parasitoid Jewel Wasp, *Amplex compressa*, of the Insect Order Hymenoptera and the Obligate Entomophagous Assassin Bug *Platyeris biguttata*, of the Insect Order Hemiptera. Ph.D., UC Riverside.

Moore, E. L., et al. (2006). "Parasitoid wasp sting: a cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." Journal of neurobiology **66**(8): 811-820.

Moore, E. L., et al. (2006). "Parasitoid wasp sting: A cocktail of GABA, taurine, and β -alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." **66**: 811-820.

Morton, A. and F. Garcia-del-Pino (2013). "Sex-related differences in the susceptibility of *Periplaneta americana* and *Capnodis tenebrionis* to the entomopathogenic nematode *Steinernema carpocapsae*." J Invertebr Pathol **112**(3): 203-207.

Palm, N. W., et al. (2013). "Bee venom phospholipase A2 induces a primary type 2 response that is dependent on the receptor ST2 and confers protective immunity." Immunity **39**(5): 976-985.

Paoli, M., et al. (2009). "Mass spectrometry analysis of the phospholipase A(2) activity of snake pre-synaptic neurotoxins in cultured neurons." J Neurochem **111**(3): 737-744.

Parra, G., et al. (2007). "CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes." Bioinformatics **23**(9): 1061-1067.

Paul, M. M., et al. (2015). "Bruchpilot and Synaptotagmin collaborate to drive rapid glutamate release and active zone differentiation." Front Cell Neurosci **9**: 29.

Paulo, J. A. (2013). "Practical and Efficient Searching in Proteomics: A Cross Engine Comparison." Webmedcentral **4**(10).

Perkins, D. N., et al. (1999). "Probability-based protein identification by searching sequence databases using mass spectrometry data." Electrophoresis **20**(18): 3551-3567.

Piek, T. (1986). Venoms of the Hymenoptera: Biochemical, Pharmacological and Behavioural Aspects.

Piek, T., et al. (1989). "The venom of *Ampulex compressa*--effects on behaviour and synaptic transmission of cockroaches." Comp Biochem Physiol C **92**(2): 175-183.

- Predel, R., et al. (2005). "Tachykinin-related peptide precursors in two cockroach species." FEBS Journal **272**(13): 3365-3375.
- Rana, S. B., et al. (2016). "Comparison of De Novo Transcriptome Assemblers and k-mer Strategies Using the Killifish, *Fundulus heteroclitus*." PLoS One **11**(4): e0153104.
- Resende, V. M., et al. (2013). "Proteome and phosphoproteome of Africanized and European honeybee venoms." Proteomics **13**(17): 2638-2648.
- Robertson, G., et al. (2010). "De novo assembly and analysis of RNA-seq data." Nat Methods **7**(11): 909-912.
- Rosenberg, L. A., et al. (2005). "Wasp venom injected into the prey's brain modulates thoracic identified monoaminergic neurons." Journal of neurobiology **66**(2): 155-168.
- Roth, M. J., et al. (2011). "Sensitive and reproducible intact mass analysis of complex protein mixtures with superficially porous capillary reversed-phase liquid chromatography mass spectrometry." Anal Chem **83**(24): 9586-9592.
- Sasmal, N. K., et al. (2008). "Experimental infection of the cockroach *Periplaneta americana* with *Toxocara canis* and the establishment of patent infections in pups." J Helminthol **82**(2): 97-100.
- Schirmer, E. C., et al. (2003). "MudPIT: A powerful proteomics tool for discovery." Discov Med **3**(18): 38-39.
- Schulz, M. H., et al. (2012). "Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels." Bioinformatics **28**(8): 1086-1092.
- Shen, Y., et al. (2002). "High-Efficiency Nanoscale Liquid Chromatography Coupled On-Line with Mass Spectrometry Using Nanoelectrospray Ionization for Proteomics." Analytical Chemistry **74**(16): 4235-4249.
- Simao, F. A., et al. (2015). "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs." Bioinformatics **31**(19): 3210-3212.
- Sookrung, N., et al. (2014). "Proteome and allergenome of Asian wasp, *Vespa affinis*, venom and IgE reactivity of the venom components." J Proteome Res **13**(3): 1336-1344.

- Surget-Groba, Y. and J. I. Montoya-Burgos (2010). "Optimization of de novo transcriptome assembly from next-generation sequencing data." Genome Res **20**(10): 1432-1440.
- Van Vaerenbergh, M., et al. (2015). "IgE recognition of chimeric isoforms of the honeybee (*Apis mellifera*) venom allergen Api m 10 evaluated by protein array technology." Mol Immunol **63**(2): 449-455.
- Vejayan, J., et al. (2014). "Comparative analysis of the venom proteome of four important Malaysian snake species." J Venom Anim Toxins Incl Trop Dis **20**(1): 6.
- Vetter, R. S. and P. K. Visscher (1998). "Bites and stings of medically important venomous arthropods." Int J Dermatol **37**(7): 481-496.
- Walther, T. C. and M. Mann (2010). "Mass spectrometry-based proteomics in cell biology." J Cell Biol **190**(4): 491-500.
- Washburn, M. P. (2004). "Utilisation of proteomics datasets generated via multidimensional protein identification technology (MudPIT)." Brief Funct Genomic Proteomic **3**(3): 280-286.
- Weisel-Eichler, A. and F. Libersat (2002). "Are monoaminergic systems involved in the lethargy induced by a parasitoid wasp in the cockroach prey?" Journal of Comparative Physiology A **188**(4): 315-324.
- Weisel, E., et al. (1999). "Venom of a parasitoid wasp induces prolonged grooming in the cockroach." The Journal of experimental biology **202 (Pt 8)**: 957-964.
- Weiss, K., et al. (2014). "Multifaceted defense against antagonistic microbes in developing offspring of the parasitoid wasp *Ampulex compressa* (Hymenoptera, Ampulicidae)." PLoS One **9**(6): e98784.
- Williams, F. X. (1942). "Ampulex Compressa (Fabr.), A Cockroach-Hunting Wasp Introduced from New Caledonia Into Hawaii." Hawaiian Entomological Society **11**(2): 221-233.

Winther, A. M., et al. (2006). "Tachykinin-related peptides modulate odor perception and locomotor activity in *Drosophila*." Mol Cell Neurosci **31**(3): 399-406.

Xie, Y., et al. (2014). "SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads." Bioinformatics **30**(12): 1660-1666.

Zhang, X., et al. (2010). "Multi-dimensional liquid chromatography in proteomics--a review." Anal Chim Acta **664**(2): 101-113.



Figure 1-1. *A. compressa* attacks its prey

A. compressa observed during envenomation of cockroach cephalic ganglia. The wasp grips the cockroach pronotum with its mandibles, while maneuvering its abdomen forward to sting the cockroach directly into the brain. Photo by Victor Landa (2014)

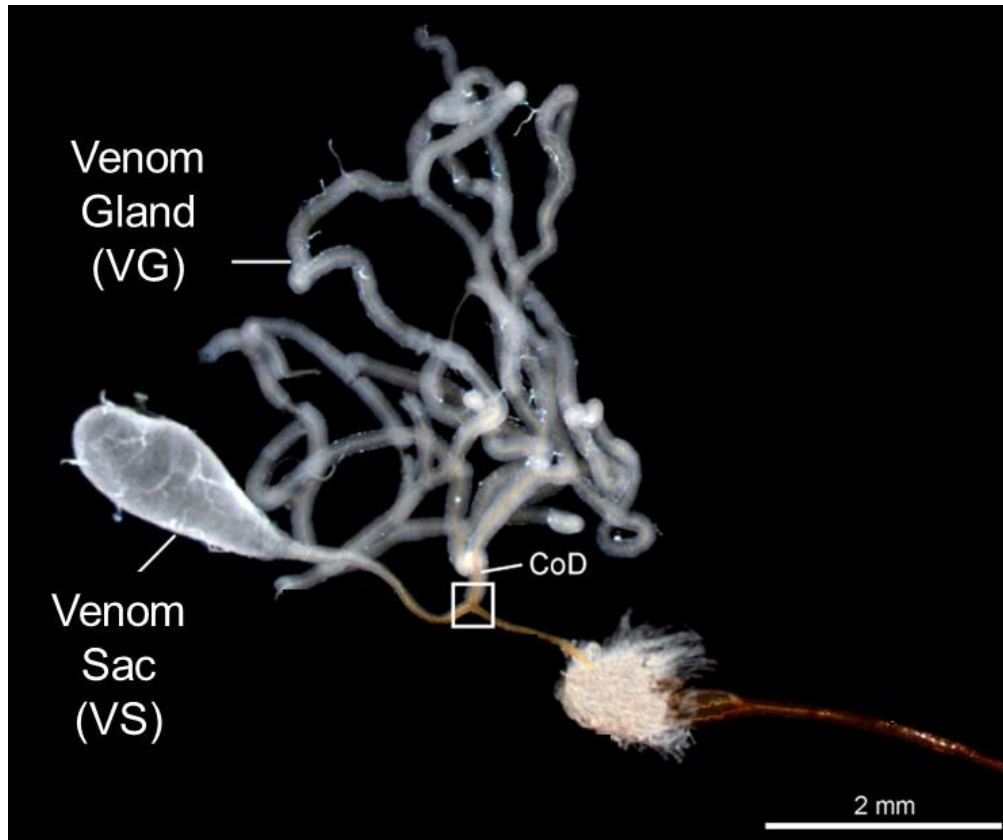


Figure 1-2. The venom apparatus of *A. compressa*

The venom apparatus of *A. compressa* is composed of two distinct glandular organs, the venom gland (VG), a long highly branched tubular structure, and a venom sac or reservoir (VS). Both are connected together and to the stinger via a common duct, or ductus venatus (CoD). Adapted with permission from (Gnatzy, 2016). Copyright Elsevier 2016.

Chapter II.

Life History of *Ampulex compressa*

Abstract

Ampulex compressa is an endoparasitoid of the American cockroach *Periplaneta americana*. Its host subjugation strategy is unusual in that envenomation is directed into the host central nervous system, eliciting a long-term behavior modification termed hypokinesia. Hypokinesic cockroaches become a living food supply for wasp offspring. *A. compressa* manipulates hypokinesic cockroaches into a burrow, followed by oviposition of a single egg on the leg. Larvae hatch three days later and pierce the cuticle of the cockroach to feed on hemolymph for the first two instars. During ecdysis to the third instar, the larva enters the body cavity of the cockroach, consuming internal tissues including fat body and skeletal muscle, sparing the gut and nervous system. Each larval instar has a distinct, specialized mandible morphology. Herein we describe the life history of *A. compressa* with emphasis on larval morphology and developmental timing.

Introduction

The Emerald Jewel Wasp *Ampulex compressa* (Fabricius, 1781), a solitary endoparasitoid wasp employs a unique strategy to subdue and exploit its host, the American cockroach *Periplaneta americana*.

Upon encountering the cockroach, *A. compressa* hunts and aggressively attacks its prey (Keasar et al. 2006). The attack is characterized by two stings: a swift sting into the thorax that leads to a transient paralysis of the prothoracic legs, and one precise sting to the head, targeting the head ganglia (Haspel et al. 2003). The head sting is remarkably precise, as the wasp is able to sense the location of the brain through mechanosensory inputs from the stinger (Gal et al. 2014). Envenomation the head ganglia induces a long term (7 – 10 day) behavioral change characterized by increased escape threshold, decreased escape distance, and decreased spontaneous walking. This modification in behavior, known as hypokinesia, facilitates the parasitization process by rendering the cockroach submissive to manipulation by the wasp (Fouad et al. 1996). After the cockroach is subdued, *A. compressa* precisely clips the antenna and ingests hemolymph from the wound (Piek et al. 1984). The wasp then leads the cockroach by grasping the truncated antennae, and pulling it into a nearby burrow (Veltman et al. 1991). She then proceeds to lay a single egg on the coxa of the host's mesothroacic leg and entombs the host by sealing the borrow entrance with miscellaneous debris (Williams 1942). This provides the young wasp hatchling with a fresh food supply during its development. As the larva develops, it eats the cockroach alive, avoiding the nervous and digestive systems, which has the added benefit of keeping the host alive as long as possible. More

recently it has been shown that *A. compressa* larvae secrete antimicrobial compounds into the cockroach as it is being consumed to sanitize and perhaps preserve the host (Herzner et al. 2013), (Weiss et al. 2014).

Most parasitoid wasps oviposit upon or within a host where the larva(e) will feed and develop (Quicke 1997). The female wasp may also inject various proteins and small molecule venom components that can suppress the host's immune system and metabolism for the benefit of her offspring. When the larvae is ready to pupate it will either chew its way out, killing the host and develop into adulthood outside of the host, or pupate inside of the host killing it or rendering it moribund, depending on the species (Heimpel et al. 1996). *A. compressa* affects the host by modifying both its behavior as well as its metabolism (Piek et al. 1984), (Haspel et al. 2005). Once most of the edible tissue of the cockroach has been devoured, the larva pupates. A few weeks later, the mature adult will eclose and chew its way out of the cockroach carcass and emerge as an adult wasp. The adult males of the species neither hunt nor have envenomation apparatus or stingers. Adult females are larger than the males, and are avid cockroach hunters. As adults, parasitoid wasps generally feed on nectar and pollen, and sometimes host hemolymph or tissues (Patt et al. 1997).

A. compressa larva hatch adhered to the cuticle of the host approximately three days after being deposited on the cockroach coxa (Paterson Fox et al. 2009). The first instar must pierce the cockroach cuticle to start the flow of hemolymph that the larva will consume as sustenance. The larva remains adhered to the cockroach through the first two stages of development, continually obtaining nourishment from host hemolymph (Fox

2006). The cockroach remains alive throughout the first two larval instars and even several hours into the third instar as it selectively eats the cockroach organs. Once ready to pupate, the larva will spin a cocoon with silk from its mandibles that will form in two layers, a hard shell that encases the pupa and a thick, woven silk outer covering that surrounds the shell (Williams 1942). Once pupated, the development time into adulthood varies with sex and will take four to five weeks after which an adult will eclose from the cocoon still inside the desiccated husk of the cockroach and the life cycle is complete.

Herein we describe three instars of larval development, according to Dyar's rule (Dyar 1890). Dyar's Rule asserts that heavily sclerotized body segments, such the head capsule and mandibles, grow in predictable quantum jumps between instars. Dyar's Rule has been applied to many hymenoptera and the number of parasitoid wasp instars can range from two to five (Hansell 1982), (Odebiyi et al. 1986), (Whitfield et al. 1987), (Giannotti 1997), (Solis et al. 2010), (Kim et al. 2011). We further characterize the unique mandibular structures for each instar, as well as developmental time course (Williams 1942), (Paterson Fox et al. 2009). The organs that the larva consumes, as well as those that it avoids during its final stage of development are described.

Materials and Methods

Animal Husbandry

A. compressa wasps were reared in 2 ft³ plexiglass cages on a 16:8 light:dark cycle with honey and water ad libitum at 28° C and 40% humidity. One female and two to three males were maintained in each cage. Cockroaches were reared in 55 gallon plastic containers with a 18 volt electric barrier along the perimeter of the top of the container on a 16:8 light:dark cycle with dog food and water ad libitum at 28 C and at 40% humidity. Adult female cockroaches were provided to single female *A. compressa* for parasitization 5 times a week but no more than one in 24 hours. Wasp development was completed in French vials in a 50-60% humidified incubator. Eclosed wasps were placed into gender-specific holding cages.

Larval and pupal development duration

Differences in development times between male and female *A. compressa* were determined relative to the time a cockroach host was introduced into *A. compressa* cages, the time between parasitization and pupation, and the time between pupation and eclosion. Sex of the wasp was determined upon emergence. Seven breeding cages were maintained, each containing a single female wasp and one to three males. Cockroaches were introduced once a day, five days a week for approximately a year. When a female wasp would die or otherwise lose fecundity, she was replaced with a 10-day staged, naive female. Over this period, 1896 cockroaches introduced to female *A. compressa*, 1656 were parasitized, and 1033 eggs progressed to pupation. From those 1033 pupa, 962

emerged as adult wasps, and of those 962, 135 (14%) were female and 818 (80%) were male.

Pupal volume and adult mass

Since pupae of male and female *A. compressa* are visually similar, we distinguished gender prior to eclosion by measuring volume. To accomplish this, 73 pupae were selected and measurements of length and the width were made (Figure 2-1). Since the cocoon is a prolate spheroid, similar to a parabola revolving about its axis, the following equations were derived to obtain volume as a function of length and width, where l is the long axis of the cocoon, L is length of the cocoon at the longest point and W is the corresponding width.

$$\text{Eq. 1} \quad \text{Volume} \sim \pi \int_0^L [(-W/2)/(L/2)^2(l - L/2)^2 + W/2]^2 dl$$

Head capsule measurements

Head capsule measurements were taken from 24 larva varying in age over a 10-day period. The age of the larva was referenced to time of oviposition. Larval body length and width was measured using a calibrated ruler. Head capsule width was taken as distance between the eyes and length from mouth to head apex. The product of length and width yielded head capsule area. Head capsule area was used instead of length or width to emphasize differences between instars.

Electron microscopy

For each *A. compressa* larva measured for head capsule size, mandibles were dissected and imaged under a Hitachi TM-1000 tabletop scanning electron microscope at the

Microscopy Core at the Institute of Integrative Biology at the University of California,
Riverside.

Results

Larval development time in *A. compressa*, defined as the interval between oviposition and larval pupation, takes $\sim 7.5 \pm 1$ days. We observed no sexual dimorphism in development time of larvae. In contrast, pupal duration, defined as the interval between pupation and adult eclosion, is significantly different between sexes. Average pupation duration of male *A. compressa* is 29.5 ± 1.7 days ($n = 200$), whereas average female pupation duration is 34.5 ± 2.3 days ($n = 135$) and both are similar to previously reported values (Figure 2-1). The longer pupation duration of female *A. compressa* correlates with larger size. Average pupal volume for females is 0.26 ± 0.02 cm³, whereas average pupal volume for males is 0.180 ± 0.01 cm³. The mass of newly eclosed adult wasps also differs significantly between males and females. Female mass is 142.8 ± 7.3 mg compared to male mass of 100.6 ± 3.4 mg (Figure 2-2).

We observed that *A. compressa* develops through three larval instars distinguished by quantum changes in head capsule size over the developmental period of 8 days (Figure 2-3). Average head capsule length and width by instar are listed in Table 2-1. The first instar hatches within 3 days of oviposition with body length averaging 3.53 ± 0.77 mm, and width at widest point averaging 1.07 ± 0.18 mm. The second instar appears during the 4th day of development, with body length measuring 10.02 ± 2.88 mm and width at 4.00 ± 0.99 mm. The third and final instar appears on day 6 with head capsule about double the size of the first instar. Third instar body length is about 3 times larger than the second instar and 9 times larger than the first instar, measuring 28.36 ± 7.27 mm with a width of 8.79 ± 2.06 mm.

The first and second instar larvae were observed in all cases to develop outside of the cockroach, while third instar larvae were observed exclusively inside of the host. We therefore conclude that the ecdysis between second to third instar is facilitated by the entry of the larva into the cockroach body cavity. This is corroborated by appearance of larval cuticle on the outside of the cockroach near the entry wound. We were not able to observe shedding of cuticle resulting from ecdysis of the first instar, however it has been observed that layers of cuticle appear between larva and host (Williams 1942).

We observed three distinct morphological changes in mandibles of *A. compressa* during larval development (Figure 2-4). The mandibles of the first instar are about 100 μm in length. These exist when the larva is in its early stages and appear to be designed for piercing the cockroach host to gain access to hemolymph. Mandibles of the second instar are about 310 μm long and have a serrated edge suitable for cutting into the intersegmental membrane. This structure appears when it is time for the larva to cut through the exoskeleton of the cockroach and begin to carve its way inside. Third instar mandibles are considerably larger at 360 μm and are relatively blunt, suitable for macerating internal organs such as muscle and fat body.

Priority tissues for consumption clearly are muscle tissue and the fat body, which were first observed to be missing following entry of the larva into the host. Next consumed, were the tracheal system and reproductive organs. The final “meal” was the central nervous system, though a significant portion of the ventral nerve cord remained following wasp pupation. The gut also remains intact following pupation (Figure 2-5). This may help to keep the cockroach alive as long as possible to prevent decay and

spoilage of the remaining food source. During pupation the larva insulates itself from the host by spinning a layer of silk between the gut and the pupal case. Eventually the host desiccates without rotting and the entombed wasp develops in its surrogate womb.

Discussion

Ampulex compressa is a remarkable wasp renowned for its unique host subjugation strategy, in essence zombifying the cockroach that then becomes a complacent, living source of nourishment for the single offspring (Gal et al. 2010). *A. compressa* and *P. americana* coevolved so that *A. compressa* instinctively knows its host from birth. We now see that this instinctual familiarity goes beyond just adult behavior as the appetite of the larva is selective in consuming cockroach tissues. *A. compressa* larvae seem to avoid consuming the gastrointestinal tract of the host, focusing instead on high nutrition tissues like fat body, muscle and in the case of female cockroaches, ovaries. The larvae do an extremely thorough job of clearing out the muscle tissues as well as most of the fat body, reproductive organs as well as some of the ventral nerve cord. However, the entire digestive tract is left completely intact. In fact, the larva takes extra measures to distance itself from this part of the host anatomy by laying down a bed of silk over the entire gut from esophagus to anus, flattening it and separating itself. The reason behind this behavior may be to protect itself from harmful contaminants that might reside within the gut of the cockroach.

The wasp larva develops through 3 instars, the first two exclusively outside of the cockroach and the third exclusively inside. This is indicated by corresponding distinctive head capsule sizes with three distinct mandibular morphologies. Morphology of the mandibles appears well-suited to the needs of each instar. The first instar must pierce the intersegmental membrane in order to begin to draw hemolymph from its host. The vampiric first instar has a circular labium that may be ideal for sucking and ingesting

hemolymph. The second instar mandibles are larger, more complex, and serrated, similar to a saw. This corresponds to the need for this instar to cut a hole into the cuticle of the cockroach so that the larva may enter. The third and final instar mandibles are larger still and have a blunt, crushing surface suitable for maceration and consumption of the softer internal organs. Once inside, the third instar larva consumes all fat body and muscle within its reach, even removing muscle into the coxa and femur of the forelegs. The wasp spins a cocoon and pupates inside the cockroach, where it develops into an adult after several weeks and ready to mate and hunt, completing the life cycle (Figure 2-6).

While larval development duration is not sexually dimorphic, pupal duration is longer in females. Further, the pupal volume is on average significantly larger than males. Since the cocoon is a prolate spheroid, it can be modeled as a rotated parabola to estimate volume (Figure 2-7). This may allow prediction of sex weeks before adult eclosion. Since *A. compressa* are haplodiploid, females may or may not fertilize their eggs, where fertilized eggs generate females, and non-fertilized eggs generate males. This results in different numbers of male and female offspring, presumably at the whim of the maternal female. Only female *A. compressa* envenomate cockroaches and pharmacological research of the envenomation mechanism relies on consistent generation of females. It is then useful to predict which pupa will yield female adults, and measurement of pupal size constitutes a practical way to predict the likelihood that the pupa will develop into a female.

I have observed that *A. compressa* is very territorial and aggressive toward others of its species, especially between females in the presence of cockroaches, and I speculate

that female *A. compressa* are less likely to fertilize eggs in the presence of other females. My recent observations have shown that percentage of female offspring has increased from 20% to 40% after isolating breeding females and obscuring their view of other females.

These results provide evidence in support of three instars during development of *A. compressa*, each with characteristic differences in morphology. We also describe the differences between male and female *A. compressa* in size and mass. We have determined that even though the time from egg deposition to pupation is the same between males and females, the time from pupation into adulthood is significantly longer for females than males, though this might be expected as the size of the female wasp is significantly larger from that of the male.

References Cited

- Buys, S. C. (2007). "Morphological studies on the last instar larva of *Ampulex compressa* (Fabricius) from Brazil (Insecta, Hymenoptera, Ampulicidae)." Spixiana **30**(1): 33-37.
- Dyar, H. G. (1890). "The Number of Molts of Lepidopterous Larvae." Psyche **5**(175-176): 420-422.
- Fouad, K., et al. (1996). "Neuromodulation of the escape behavior of the cockroach *Periplaneta americana* by the venom of the parasitic wasp *Ampulex compressa*." Journal of Comparative Physiology A **178**(1): 91-100.
- Fox, E. G. B., S.C. Mallet, R. S. Bressan-Nascimento S. (2006). "On the morphology of the juvenile stages of *Ampulex compressa* (Fabricius, 1781) (Hymenoptera, Ampulicidae)." Zootaxa **1279**: 43-51.
- Gal, R., et al. (2014). "Sensory arsenal on the stinger of the parasitoid jewel wasp and its possible role in identifying cockroach brains." PLoS One **9**(2): e89683.
- Gal, R. and F. Libersat (2010). "On predatory wasps and zombie cockroaches: Investigations of "free will" and spontaneous behavior in insects." Commun Integr Biol **3**(5): 458-461.
- Giannotti, E. (1997). "Biology of the wasp *Polistes (Epicnemius) cinerascens* Saussure (Hymenoptera: Vespidae)." Anais da Sociedade Entomológica do Brasil **26**: 61-67.
- Godfray, H. C. J. (1994). Parasitoids : behavioral and evolutionary ecology. Princeton, N.J., Princeton University Press.
- Hansell, M. H. (1982). "Brood development in the subsocial wasp *Parischnogaster mellyi* (Saussure) (Stenogastrinae, Hymenoptera)." Insectes Sociaux **29**(1): 3-14.
- Haspel, G., et al. (2005). "Parasitoid wasp affects metabolism of cockroach host to favor food preservation for its offspring." Journal of Comparative Physiology A **191**(6): 529-534.
- Haspel, G., et al. (2003). "Direct injection of venom by a predatory wasp into cockroach brain." Journal of neurobiology **56**(3): 287-292.
- Heimpel, G. E. and T. R. Collier (1996). "THE EVOLUTION OF HOST-FEEDING BEHAVIOUR IN INSECT PARASITOIDS." Biological Reviews **71**(3): 373-400.

Herzner, G., et al. (2013). "Larvae of the parasitoid wasp *Ampulex compressa* sanitize their host, the American cockroach, with a blend of antimicrobials." Proceedings of the National Academy of Sciences **110**(4): 1369-1374.

Keasar, T., et al. (2006). "Host-Handling Behavior: An Innate Component of Foraging Behavior in the Parasitoid Wasp *Ampulex compressa*." Ethology **112**(7): 699-706.

Kim, I.-K., et al. (2011). "Prolonged embryonic stage and synchronized life-history of *Platygaster robiniae* (Hymenoptera: Platygasteridae), a parasitoid of *Obolodiplosis robiniae* (Diptera: Cecidomyiidae)." Biological Control **57**(1): 24-30.

Odebiyi, J. A. and A. H. Bokonon-Ganta (1986). "Biology of *Epidinocarsis* [*Apoanagyrus*] *lopezi* [Hymenoptera: Encyrtidae] an exotic parasite of cassava mealybug, *Phenacoccus manihoti* [Homoptera: Pseudococcidae] in Nigeria." Entomophaga **31**(3): 251-260.

Paterson Fox, E. G., et al. (2009). "Notes on the Biology and Behaviour of the Jewel Wasp, *Ampulex compressa* (Fabricius, 1781) (Hymenoptera; Ampulicidae), in the Laboratory, Including First Record of Gregarious Reproduction." Entomological News **120**(4): 430-437.

Patt, J. M., et al. (1997). "Foraging success of parasitoid wasps on flowers: interplay of insect morphology, floral architecture and searching behavior." Entomologia Experimentalis et Applicata **83**(1): 21-30.

Piek, T., et al. (1984). "Change in behaviour of the cockroach, *Periplaneta americana*, after being stung by the sphecid wasp *Ampulex compressa*." Entomologia Experimentalis et Applicata **35**(2): 195-203.

Quicke, D. L. J. (1997). Parasitic wasps. London ; New York, Chapman & Hall.

Solis, D. R., et al. (2010). "Description of the Immatures of Workers of the Ant *Camponotus vittatus* (Hymenoptera: Formicidae)." Florida Entomologist **93**(2): 265-276.

Veltman, J. and W. Wilhelm (1991). "Husbandry and display of the Jewel wasp: *Ampulex compressa* and its potential value in destroying cockroaches." International Zoo Yearbook **30**(1): 118-126.

Weiss, K., et al. (2014). "Multifaceted defense against antagonistic microbes in developing offspring of the parasitoid wasp *Ampulex compressa* (Hymenoptera, Ampulicidae)." PLoS One **9**(6): e98784.

Whitfield, G. H., et al. (1987). "NUMBER OF INSTARS OF LARVAE OF THE ALFALFA LEAFCUTTER BEE, MEGACHILE ROTUNDATA (F.)

(HYMENOPTERA: MEGACHILIDAE)." The Canadian Entomologist **119**(10): 859-865.

Williams, F. X. (1942). "Ampulex Compressa (Fabr.), A Cockroach-Hunting Wasp Introduced from New Caledonia Into Hawaii " Proc Hawaiian Entomol Soc **11**(2): 221-233.

Instar	Larva		Head Capsule	
	Length (mm)	Width (mm)	Length (μm)	Width (μm)
1	3.53 \pm 0.77	1.07 \pm 0.18	514 \pm 36	561 \pm 128
2	10.0 \pm 2.88	4.00 \pm 0.99	885 \pm 88	780 \pm 72
3	28.4 \pm 7.27	8.79 \pm 2.06	1181 \pm 104	885 \pm 64

Table 2-1. Size of larval body and head capsule by instar

n = 6 for instar 1, n = 5 for instar 2, and n = 13 for instar 3.

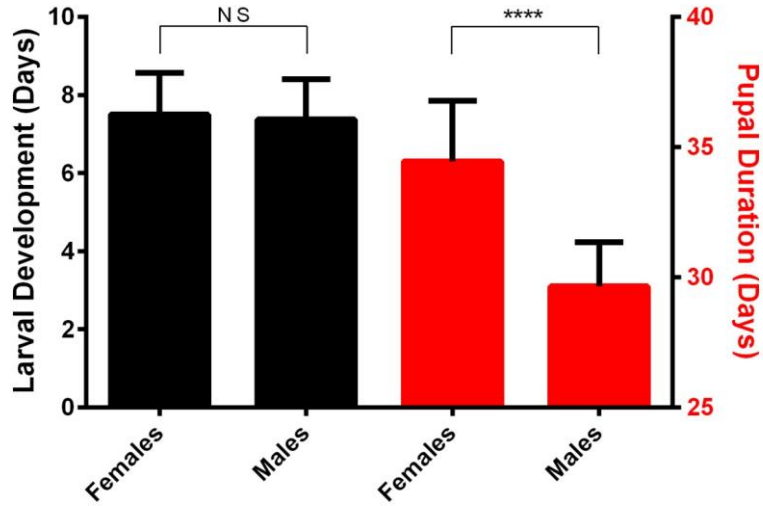


Figure 2-1. *A. compressa* development time is sexually dimorphic

Larval development as defined by the duration between egg laying and pupation shows no difference between males and females (black, $p = 0.26$). The average time to pupation for males was 7.4 days whereas the average time to pupation for females was 7.5 days. The time between pupation and eclosion is significantly different between males and females (red, $p < 2.48 \times 10^{-52}$). The average time from pupation to eclosion for males is 29.6 days whereas the average time from pupation to eclosion for females is 34.5 days. (Statistical difference determined by Student's T test, NS = not significant, **** = $p < 0.0001$, males $n = 200$, females $n = 135$)

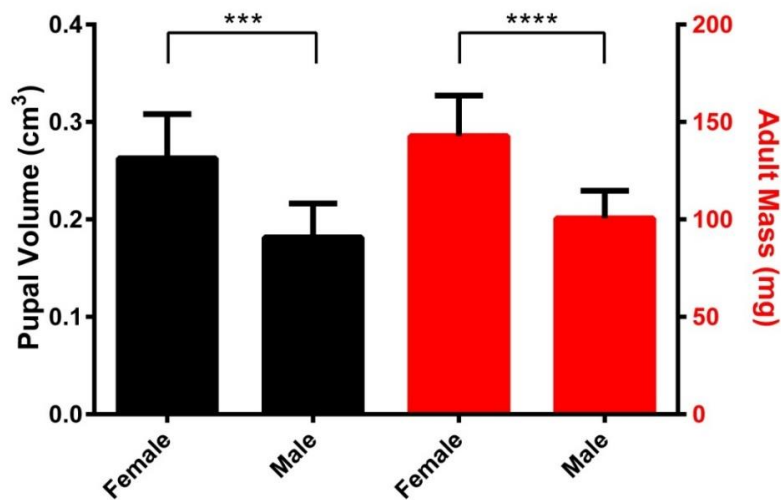


Figure 2-2. *A. compressa* pupal and adult size is sexually dimorphic

There is a significant difference in the size of pupa and adult mass between male and female *A. compressa*. The pupal volume (black) of the female *A. compressa* is significantly larger than that of the male *A. compressa*. The average pupal volume of females is 0.26 cm³, whereas the pupal volume of males is 0.182 cm³ (Female n = 8, male n=21). The adult mass (red), after eclosion is also significantly larger of females than males. The average post-eclosion mass of females is 142.8 mg, whereas the post-eclosion mass of males is 100.6 mg (Female n = 8, male n=17). (***) = $p < 0.001$, (****) = $p < 0.0001$, by Student's T test).

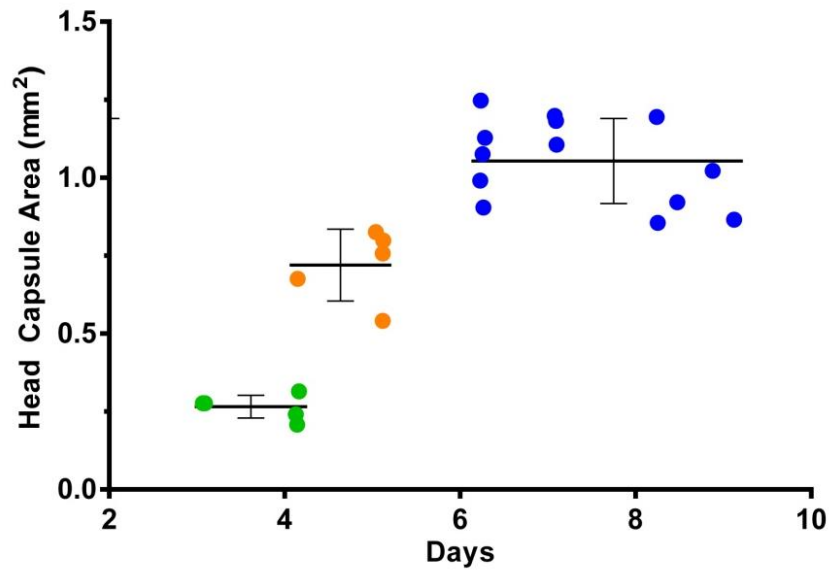


Figure 2-3. *A. compressa* larva develops through three instars

Three instars of the *A. compressa* were determined by taking representative larva from different time intervals and plotting the product of the length and width of the head capsule (area) against the age of the larva. The data plots as three distinct groups indicative of three instars. The green points represent first instar, orange represents second instar and blue represent the third instar. The horizontal bars are average size of the head capsule per instar and the vertical error bar is standard deviation.

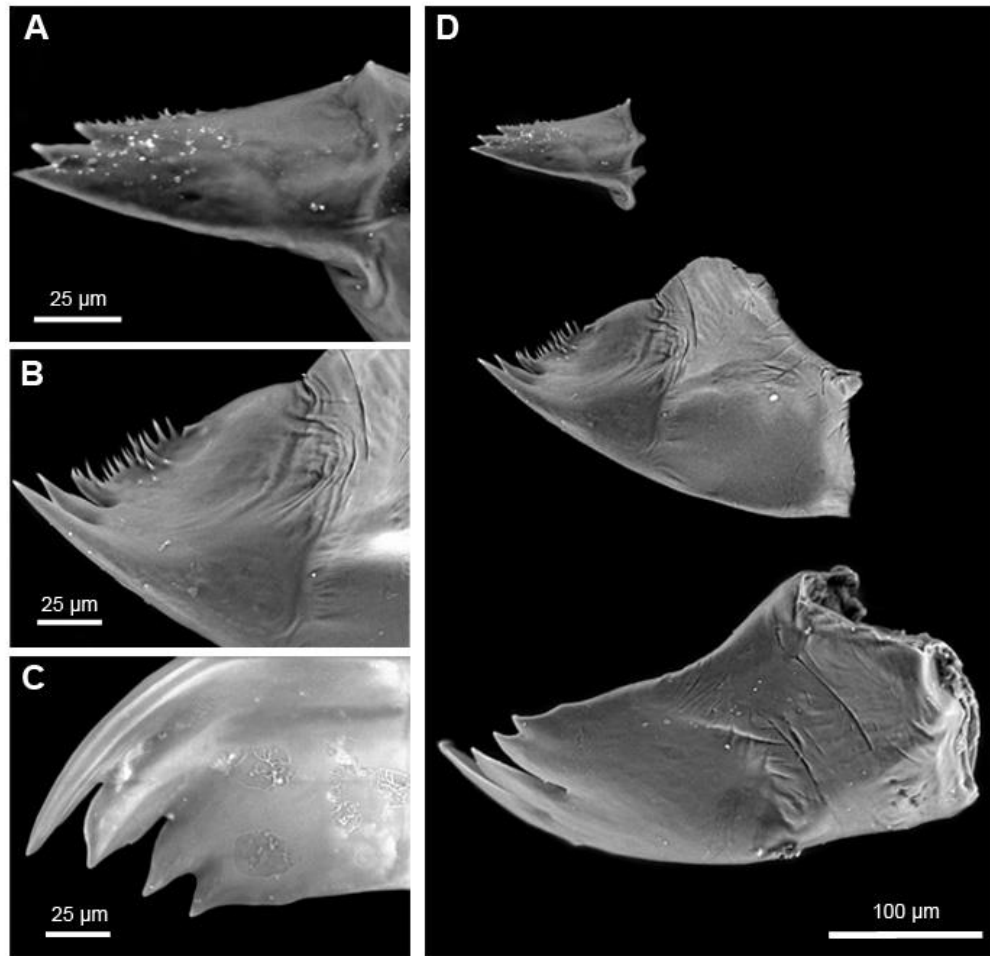


Figure 2-4. Comparative morphological comparison of the mandibles of the three instars of *A. compressa* taken by a scanning electron microscope

Mandible morphology differs between the instars and suits the needs of the instar. **A.** First instar mandibles appear suitable for piercing the cockroach cuticle and facilitating a steady flow of hemolymph without serious injury to the cockroach. **B.** The second instar mandible is larger and contains a serrated edge suitable for cutting into the cockroach to facilitate entry into the host. **C.** Third and last instar mandibles of *A. compressa*. These appear after the larva has entered the body cavity of the cockroach to consume fat body and muscle. These mandibles appear to be suited for crushing and macerating the internal soft tissues of the cockroach. **D.** All three mandible types together to scale for comparison.

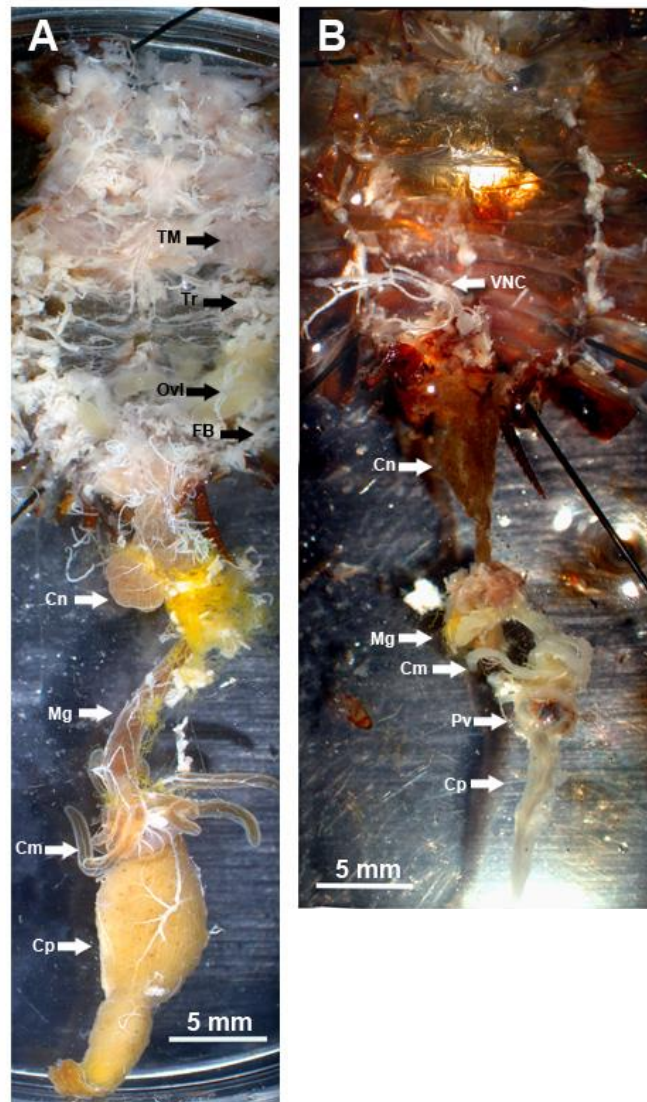


Figure 2-5. Viscera of *P. americana* before and after parasitization by *A. compressa*

A. The entire digestive system has been isolated to more thoroughly illustrate all parts of the gut as well as tissues previously obstructed such as trachea, ovaries, thoracic muscle, fat body etc. **B.** Isolated gut from after completion of *A. compressa* larval development. Abbreviations: TM = thoractic muscle, Tr =Trachea, Ovl = ovarioles of ovary, FB = fat body, Cn = colon, Mg = midgut, Cm = Caeca, Cp = crop, VNC = ventral nerve cord, Pv = proventriculus.

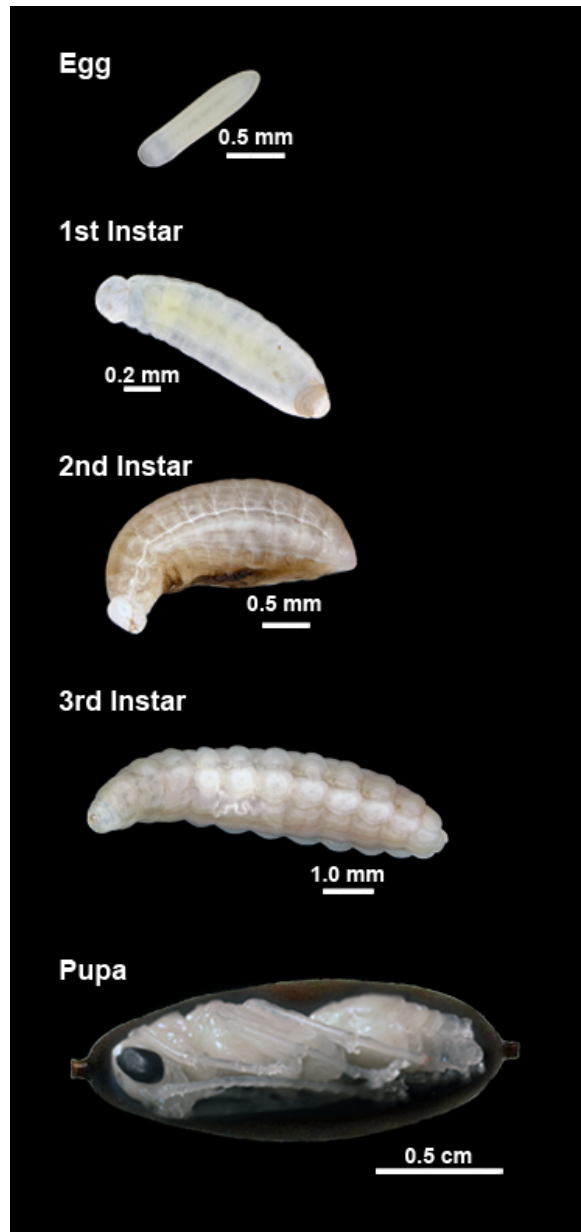


Figure 2-6. Life Cycle of *Ampulex compressa*

The life cycle of *A. compressa* from egg to adult. The different developmental stages of the *A. compressa* includes egg, first instar, second instar, third instar, and pupa.

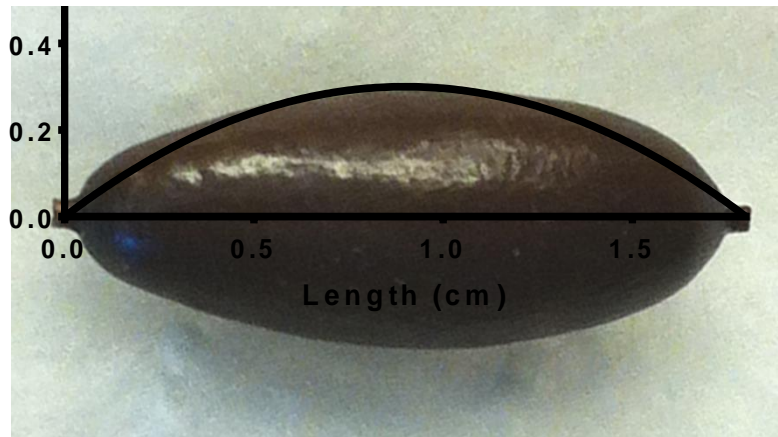


Figure 2-7. Estimation of pupal size by modeling as a parabola

The shape of the cocoon of *Ampulex compressa* pupa is a prolate spheroid, thus the volume of the cocoon can be estimated by modeling it as parabola, rotated about the longitudinal axis, using only length and width measurements.

Chapter III.

Transcriptomic and Proteomic Analysis of *Ampulex compressa* Venom

Abstract

The parasitoid jewel wasp *Ampulex compressa* exploits its host, the American cockroach *Periplaneta americana*, as a food source for its young. The host is subdued through direct envenomation into the central nervous system, inducing a days-long “zombie-like state” referred to as hypokinesia. Hypokinesia is a specific, venom-induced behavioral state characterized by suppression of the escape response and reduced spontaneous walking, leaving other motor functions unaffected. The condition is reversible, provided that egg deposition is prevented. To facilitate elucidation of the biochemical mechanisms underlying venom-induced hypokinesia, we developed a comprehensive inventory of the venom by combining: 1) next-generation RNA sequencing, transcriptomic analysis of the venom gland apparatus, and 2) proteomics of venom apparatus extracts and whole milked venom via mass spectroscopy-based Multi-Dimensional Protein Identification Technology (MudPIT). A total of 267 proteins and peptides were identified in the venom, 103 with identifiable Pfam domains, 189 similar to other proteins in Uniprot databases; 78 are novel. *A. compressa* venom proteins include members of the M13 family of proteases, phospholipase A2 isoforms, along with hyaluronidase, icarapin, venom acid phosphatase, serpins, and signaling peptides including tachykinin and corazonin. Differential expression analysis and proteomics of the two major tissues of the venom apparatus (venom gland and venom sac), identified tissue-specific expression of key venom components. venom of *A. compressa* is a rich collection of both well-known and novel, pharmacologically diverse compounds. My findings provide new and

comprehensive information on *A. compressa* venom contents that inform hypotheses regarding its mode of action.

Introduction

The endoparasitoid wasp *Ampulex compressa* injects venom directly into the central nervous system of its host, the American cockroach (*Periplaneta americana*) to induce a long-term behavioral state known as hypokinesia (Haspel et al. 2003), (Libersat 2003). Envenomation results in dramatic, long term behavioral changes rendering the much larger host compliant, allowing the wasp to physically manipulate and maneuver it into its burrow. When *A. compressa* first encounters its victim, it promptly and aggressively attacks, first stinging into the prothorax to cause short-term flaccid paralysis of the prothoracic legs (Haspel et al. 2003). This facilitates subsequent stings into cephalic ganglia (brain and subesophageal ganglion). Once cephalic stings are completed, the cockroach engages in vigorous grooming behavior for about twenty minutes while the wasp readies its burrow (Weisel-Eichler et al. 1999). The wasp then uses its large mandibles to grasp the cockroach antennae and lead it into the burrow, whereupon a single egg is deposited on the host's leg. Within three days, the wasp larva hatches, whereupon it feeds on hemolymph of the cockroach. Upon completion of the first two instars, the larva enters the body cavity of the host and consumes its internal tissues of the living host during the third instar; pupation occurs around day eight. After four weeks, the adult wasp emerges and the life cycle is completed (Haspel et al. 2005).

Hypokinesia is a specific, venom-induced behavioral state characterized by increased threshold for the escape response and reduced spontaneous walking; other motor functions such as the righting response remain unaffected. Stung, hypokinesic cockroaches walk if pulled, swim if submerged in water and fly in a wind tunnel (Gal et

al. 2010). Thus hypokinesia does not constitute paralysis observed in most other host-parasitoid interactions. This specificity of behavior modification is particularly interesting and unique among host-parasitoid interactions (Piek 1990), (Asgari et al. 2011). Interestingly, effects of the venom on the escape response are reversible. If the newly deposited egg is removed from an envenomated cockroach, the cockroach escape response slowly recovers, returning to normal levels within ~7-10 days. The venom appears to modulate descending signals from cephalic ganglia, resulting in suppression of host escape behavior without affecting other motor circuits (Gal et al. 2010). Furthermore, the venom lacks necrotic or lethal effects, so that the placid host remains in good condition as a food source for the wasp larva (Gal et al. 2005).

Venom-induced hypokinesia raises an interesting biological question: How can such a potent biochemical cocktail cause such long-lasting, specific, yet reversible effects on behavior? To address this question, we have generated comprehensive *A. compressa* venom proteome and venom apparatus transcriptomes and proteomes to inform our functional analyses of the venom and its role in hypokinesia induction. The venom apparatus of *A. compressa* is composed of two distinct glands: venom gland and venom sac (Figure 3-1). In other hymenoptera the venom gland is a secretory tube integrated into the venom sac or reservoir where the venom is stored before injection (Piek 1986). *A. compressa*'s venom gland is bifurcated and highly branched; it is larger than most hymenoptera with respect to body size and does not contain a reservoir in itself. The venom gland is distinct and separated from the venom sac, which is a prominent bulbous glandular sac, situated between the left and right major branches of the venom gland

(Gnatzy et al. 2015). Both venom gland and venom sac in *A. compressa* connect to the ductus venatus, or venom duct, which then enters the stinger (Piek et al. 1989), (Gnatzy et al. 2015). Previous work showed that major peptide components of the venom are present in the venom sac, but not the venom gland (Moore 2003). Therefore, both components of the venom apparatus appear to have distinct roles in venom production and require closer scrutiny in how each may participate in envenomation. Transcriptomes of the venom sac and venom gland were generated *de novo*, using Illumina short read sequencing and the Trinity pipeline (Grabherr et al. 2011), (Haas et al. 2013). Protein coding sequences extracted from the transcriptome then serve as a database for mass spectrometry based proteomics, allowing for protein discovery in non-model organisms (Evans et al. 2012). Using HPLC to introduce tryptic digestion of a complex protein mixture into the mass analyzer increases resolution and sensitivity of analysis. Furthermore, analyzing experimentally determined tandem mass spectra against a theoretical spectra derived from a protein database allows assignment of small tryptic peptides to larger coding sequences that would otherwise not be detectable by mass spectrometry. This approach, coined MudPIT (Multiple dimension Protein Identification Technology), has been used to profile complex proteomes, including venoms (Batista et al. 2004), (dos Santos et al. 2009), (Haney et al. 2014). New nucleotide sequencing and mass spectrometry technologies have greatly facilitated protein discovery in non-model systems and have advanced the field of venomomics (Escoubas et al. 2008), (de Graaf et al. 2009).

The biochemical bases for some behavioral sequelae post-enuenomation are described. For example, short-term paralysis of prothoracic legs, which occurs after the initial sting into the prothoracic ganglion, appears to be attributable to venom components GABA and GABA_A receptor agonists β -alanine and taurine (Moore et al. 2006). The vigorous grooming response induced by stings into cephalic ganglia may result from dopamine, another venom component, as dopamine injection into the cephalic ganglia of naïve cockroaches induces grooming (Weisel et al. 1999). On the other hand, causes of venom-induced hypokinesia remain obscure. Elucidation of the venom proteome allows for the generation of hypotheses on the mechanism of hypokinesia induction and a platform for future functional analysis and inquiry into the effect *A. compressa* venom may have on the cockroach brain.

Materials and Methods

Animal Husbandry

A. compressa and *P. americana* were reared as previously described. In brief: Single female *A. compressa* were housed with three males in 1 ft.² x 2 ft. tall plexiglass cages with the view of other female wasps occluded. Water and honey were provided *ad libitum*. Individual female cockroaches were introduced to the wasp 5 times per week for parasitization. *P. americana* were reared in 55-gallon trash cans with water and dog food *ad libitum*. All animals were reared at 28°C and 30 - 40% humidity on a 16-8 light-dark cycle.

RNA extraction and Sequencing

Venom sacs and venom glands were dissected from nine wasps and pooled into two biological replicates of each tissue type. RNA was extracted from each tissue using the Trizol method (Invitrogen), and quality was assessed on a Agilent 2100 Bioanalyzer. Sequencing libraries were generated and multiplexed using the Illumina TruSeq RNA Library Preparation Kit, according to manufacturer's instructions. All four libraries were combined and sequenced on the Illumina HiSeq 2000 platform at the Institute for Integrative Genome Biology, UC Riverside (IIGB). Sequencing data from each sample were combined and assembled using the Trinity software suite with the CuffFly and extended lock options, and a k-mer overlap of 2, to minimize spurious isoforms. RSEM and Deseq2 plugins for Trinity were used to quantify transcripts and calculate differential expression between tissue types, respectively (Li et al. 2011), (Love et al. 2014). The

Transdecoder plugin for Trinity was used to extract putative ORFs with a minimum length of 30 amino acids (Haas et al. 2013). The ORF database (896984 sequences) generated with Transdecoder was used for MudPIT. All computational analyses were performed on the IIGB Linux Cluster.

Mass spectrometry sample preparation

Venom was milked from healthy adult female *A. compressa* as described previously (Moore et. al. 2006). In brief: CO₂ anesthetized wasps were placed into a modified P1000 tip with the abdomen protruding from the tip, covered with parafilm, and allowed to recover. Wasps were aggravated to sting through the parafilm and venom drops were absorbed into 5 µl of deionized water, frozen on dry ice, and stored at -80°C until processed. Proteins were extracted from ten combined venom sacs or venom glands in extraction buffer (50 mM HEPES, 250 mM NaCl, 1X Protease Arrest (G-Biosciences) with 0.5 mM EDTA and 0.1% Nonidet P-40) by gentle homogenization with a pestle. Tissue homogenate was centrifuged at 10,000 rcf for 10 min at 4°C. The supernatant was purified by C18 SepPak (Waters) and concentrated by SpeedVac (Savant).

Proteins were separated by TRIS-Tricine SDS-PAGE on a 16.5% gel (BioRad) with 20 µg protein in each lane at a constant 50 volts and stained with AcquaStain Protein Gel Stain (Bulldog Bio). Precision Plus Protein Dual Xtra Prestained Protein Standards were used as a reference (BioRad).

For analysis by mass spectrometry, 1200 sting equivalents of SepPak-purified milked venom protein were split into two samples: one was separated by SDS-PAGE and the other analyzed directly. The sample to be analyzed directly was split into two samples,

one of which was trypsinized before analysis and the other was to be analyzed without protease treatment. For SDS-PAGE separated venom proteins and venom gland protein extracts, the gel was cut into fractions where indicated (Figure 3-2, A-F), and each fraction was in-gel trypsin digested and analyzed by mass spectrometry.

MudPIT Nano-UPLC-MS/MS analysis and Protein Identification

Trypsin-digested samples were dried to pellets with a SpeedVac concentrator and resuspended in 20 μ l of 0.1% formic acid solution. MudPIT LC/MS was employed to analyze these final peptide samples using two-dimensional nanoAcquity UPLC (Waters) and Orbitrap Fusion mass spectrometer (Thermo Fisher). Two-dimensional nano-UPLC fractionation and separation gradient, and the MS survey scan using data-dependent acquisition (DDA) was described previously (Drakakaki et al. 2012) (Hebert et al. 2014).

All raw MS data were processed with Proteome Discoverer version 1.4 (Thermo Fishers, San Jose, CA) to generate mgf files that were used in Mascot searches (version 2.5) against a custom ORF database. All searches were performed with the following settings: peptide mass tolerance: \pm 10 ppm, fragment mass tolerance: \pm 0.3 Da, Variable modifications: Acetyl (N-term), Amidated (C-term), Formyl (N-term), Gln-> pyro-Glu (N-term Q), Glu-> pyro-Glu (N-term E), Oxidation (M), with one max missed trypsin cleavages. Spectra were accepted for the four venom samples if the MASCOT score of the identified protein was greater than the MASCOT score that corresponds to a false discovery rate (FDR) of 1% against a reversed-decoy database. Spectra were accepted for the venom sac and venom gland samples if the MASCOT score of the identified protein was greater than the MASCOT score that corresponds to an FDR of 5%.

Protein annotation

ORFs identified as venom proteins via MudPIT were assessed for predicted secretory signals by SignalP 4.1. Molecular mass and isoelectric points for ORFs were calculated by ExPASy Compute pI/Mw tool (web.expasy.org/compute_pi/) with secretory signals removed from those sequences for which they were predicted. ORFs were searched against NCBI-nr, Uniprot and PfamA databases using standalone BLAST 2.2.30+, hmmscan or phmmer (Hmmer 3.0) where indicated.

Comparative genomic analysis

Protein sequences from genomes of *N. vitripennis*, *S. invicta*, *P. barbatus*, *L. humile*, *H. saltator*, *A. echinatoi*, *C. obscurior*, *A. cephalotes*, *B. impatiens*, *A. mellifera*, were obtained from hymenoptera base; *D. melanogaster* sequences from Flybase, *T. castaneum* sequences from Beetlebase; *O. abietinus*, *L. reclusa*, *L. Hesperus*, *C. exilicauda*, *S. maritima*, from Baylor College of Medicine Human Genome Sequencing Center; *M. musculus* and *O. Hannah* sequences from NCBI. Each genome protein set was interrogated with *A. compressa* venom ORFs using phmmer (Hmmer 3.0), with an expect cutoff of 10^{-5} . Species key: Jewel wasp, *Nasonia vitripennis* (Werren et al. 2010); Wood Wasp, *Orussus abietinus* (Poelchau et al. 2015); Fire Ant, *Solenopsis invicta* (Wurm et al. 2011); Harvester Ant, *Pogonomyrmex barbatus* (Smith et al. 2011); Argentine Ant, *Linepithema humile* (Smith et al. 2011); Jumping Ant, *Harpegnathos saltator* (Bonasio et al. 2010); Leaf-Cutter Ant, *Atta cephalotes* (Suen et al. 2011); Tramp Ant, *Cardiocondyla obscurior* (Elsik et al. 2016); Leaf-Cutter Ant, *Acromyrmex echinatoi* (Nygaard et al. 2011); Bumble Bee, *Bombus impatiens* (Sadd et al. 2015); Honey Bee,

Apis mellifera (Honeybee Genome Sequencing 2006), (Elsik et al. 2014); Fruit Fly, *Drosophila melanogaster* (Attrill et al. 2016); Flour Beetle, *Tribolium castaneum* (Tribolium Genome Sequencing et al. 2008); Brown Recluse Spider, *Loxosceles reclusa* (Poelchau et al. 2015); Black Widow Spider, *Latrodectus hesperus* (Poelchau et al. 2015); Bark Scorpion, *Centruroides exilicauda* (Poelchau et al. 2015); Centipede, *Strigamia maritima* (Poelchau et al. 2015); Mouse, *Mus musculus* (Mouse Genome Sequencing et al. 2002), (Church et al. 2009); King Cobra, *Ophiophagus Hannah* (Vonk et al. 2013).

Results

Transcriptome Assembly

RNA sequencing yielded 36M reads for venom gland replicate 1 (VG1), 78M reads for venom gland replicate 2 (VG2), 17M reads for venom sac replicate 1 (VS1), and 26M reads for venom sac replicate 2 (VS2). These reads were assembled *de novo* using the Trinity pipeline into 50667 components containing 69,009 transcripts. Completeness of assembly was assessed using CEGMA, which found complete sequences with homology to 244 out of 248 ultra-conserved, core eukaryotic genes (CEGs) (98.39%), and partial sequences for 247 out of 248 tested CEGs (99.60%) in the assembly. A total of 724 CEGs were identified with an average of 2.93 orthologs per CEG. The assembly also was queried for presence of previously identified venom proteins, ampulexins, via tBLASTn. CEGMA and BLAST results suggest that the assembly reconstructed venom transcripts efficiently.

Differential Expression

Individual sequencing libraries were mapped back to the transcriptome using RSEM to quantify tissue specific transcript abundance, and transcript levels were compared between tissue types using DEseq2. The venom gland and venom sac share 52% of assembled transcripts, with the venom gland expressing 1,535 shared transcripts more highly than the venom sac and the venom sac expressing 249 transcripts more highly than the venom gland ($p < 0.001$, fold change > 4) (Figure 3-3A). The venom gland expressed 472 unique transcripts that were differentially expressed, whereas the venom sac

expressed 132 unique transcripts that were differentially expressed; all other transcripts unique to each gland did not pass the $p < 0.001$, and 4-fold cutoff for determining differential expression. Of those differentially expressed transcripts, some are exceedingly so with false discovery rate (FDR) less than 10^{-4} , into 10^{-280} , indicating that even though the venom sac and venom gland share many transcripts in common, they have distinct expression profiles (Figure 3-3B). Of all transcripts assembled, 68,003 (98.5%) were returned with abundance information by RSEM. Abundance estimation for all transcripts ranged seven orders of magnitude. The mean venom-specific transcript counts were two orders of magnitude greater than the venom gland or venom sac total transcript counts (756, 5.21, and 4.61 respectively), indicating that the venom apparatus tissues specialize in producing venom specific transcripts (Figure 3-3C). Of all quantifiable transcripts, 19457 transcripts were unique to the venom gland, 12889 were unique to the venom sac, and 35,657 were shared between the tissue types (Figure 3-3D). Assembly transcript length ranged from 200 bp (set minimum) to 44,751, with the largest likely to be artifacts of the assembly and chimeric transcripts. The mean length of venom transcripts (2,715 bp) is twice that of the complete assembly (1,301 bp) (Figure 3-3E).

Proteomics

The venom proteome was generated by combining MuDPIT results from seven sample preparations (Figure 3-2). The combined results, filtered by a $p < 0.01$ significance threshold, totaled 267 identified proteins. Of the 267 proteins identified, 196 were differentially expressed beyond the $p_{adj} < 0.001$ and fold change > 4 cutoff (Figure 3-4A). The venom sac over-expresses 6 of the 196 differentially expressed proteins,

including the two members of a previously identified family of novel toxins coined ampulexins and icarapin, a hymenopteran venom allergen. Venom sac expresses several additional venom proteins highly, including hyaluronidase, calreticulin, and venom allergen 3, however these are not statistically over-expressed from the venom gland.

Phmmer search against NCBI-nr and Uniprot databases returned 57 results with homology to characterized proteins, 132 results to putative/uncharacterized proteins and 78 with no homology to any proteins in the databases. Hmmscan identified 103 domain hits from PfamA database. The species associated with the top BLAST hit for venom proteins against NCBI-nr and Uniprot databases were overwhelmingly hymenopteran, with the majority being ants and bees. Custom protein databases were generated from NCBI-nr to include all hymenoptera proteins, all insect proteins excluding hymenoptera, all arthropod proteins excluding insects and all metazoan proteins excluding arthropods. The venom proteome, venom gland and venom sac proteomes, and 1000 random ORFs from the assembly were analyzed against these databases via Stand-alone BLASTp (Figure 3-4B). Combined MudPIT analysis of the venom gland yielded 1,008 unique proteins with 792 identified in the high molecular weight fraction (Figure 3-2B), and 331 identified in the low molecular weight fraction (Figure 3-2E). Combined MudPIT analysis of the venom sac yielded 108 unique proteins; 13 were identified in the high molecular weight fraction (Figure 3-2C), and 102 identified in the low molecular weight fraction (7 were found in both fractions) (Figure 3-2F). Of the total proteins identified in the venom gland, 139 also were present in the venom proteome, indicating that the venom gland contributes more to the total number of venom proteins than the venom sac.

The venom sac proteome contains only 3 proteins also found in the milked venom, but not in the venom gland proteomes. However, 38 proteins are shared between the venom sac, venom gland and milked venom proteomes, including venom sac enriched proteins such as the ampulexins, tachykinin, and serpin. There were 59 venom proteins that were not detected in either the venom gland or venom sac proteomes (Figure 3-4C). No apparent linear functional relationship exists between transcript read counts and estimated protein abundance (Pearson's $p = 0.69$), but there is a positive correlation between transcript abundance and protein abundance (Spearman's $p < 0.0001$) (Figure 3-4D), indicating that the abundance of protein cannot be predicted from transcript counts. Nevertheless, if there is an increase in transcript counts, an increase in protein abundance is usually observed.

The venom proteome contains hundreds of proteins, many with multiple isoforms, paralogs, or representatives of the same enzyme family, arguing for a high degree of functional redundancy in venom components. The M13 peptidase family is well represented in the venom, and one isoform of endothelin-converting enzyme 1 stands out in both transcript counts and protein abundance, suggesting that the dominant M13 action is from this enzyme. Hyaluronidase is an exception, as it is a highly represented venom protein with only one isoform (Figure 3-5). Every protein identified in the venom proteome had non-zero transcript levels in either the venom gland or the venom sac and 96% had non-zero transcript levels in both the venom gland and the venom sac. A significant number (27%) of venom protein transcripts are not differentially expressed at the $p < 0.001$ significance level and show a similar protein abundance between tissue

types (Figure 3-6, A and B). The venom, as is evidenced by SDS-PAGE (Figure 3-2), is comprised of proteins ranging from below 2 kDa to over 100 kDa.

The venom gland proteome is enriched in large molecular weight proteins (>15 kDa), and the venom sac is enriched in low molecular weight (< 12 kDa) peptides. The SDS-PAGE banding pattern of the venom sac protein extract closely resembles that of the milked venom below 25 kDa. On the other hand, the banding pattern of the milked venom has analogous bands in the venom gland above 25 kDa, though the banding match between milked venom and venom gland samples is not as stark as the banding match between venom sac and milked venom in the sub-25 kDa range. This suggests that milked venom contains a combination of venom gland and venom sac specific proteins, with the venom gland contributing to the larger enzyme fraction of the venom and the venom sac contributing the smaller peptide toxins. This trend is reflected in the RNAseq counts as well, with the venom gland read counts higher for larger molecular weight proteins, in particular the M13 peptidase family members endothelin-converting enzyme and neprilysin, and phospholipase A2 (Figure 3-6A). In contrast, venom sac read counts for venom proteins are much lower overall than the venom gland except for the low molecular weight peptide toxins, in particular the ampulexins, where ampulexin 1 is the second-most expressed venom protein in the venom sac with low expression in the venom gland (Figure 3-6B). The most abundant venom toxin expressed in the venom sac is a novel, glycosylated peptide, that has been previously identified in the venom (VS5842) (Moore 2003). However, there are several secreted, novel peptides with molecular mass between 2.4 and 10.3 kDa in the venom proteome that have higher

transcript counts in the venom gland, though much lower peptide abundance (emPAI) than venom sac peptides (Ishihama et al. 2005).

Only 65% of *A. compressa* venom proteins are predicted to have a secretory signal, which might be counter-intuitive assuming all venom proteins are to be secreted. However, this trend is consistent with other hymenoptera-venom proteins deposited in NCBI-nr, as 56% are predicted to have secretory signals, while 45% of all venom proteins in NCBI-nr contain secretory signals. Proteomics analysis of black widow venom found that 67% of venom gland specific transcripts have secretory signals (Haney et al. 2014). In contrast only 2% of 1000 randomly selected *A. compressa* ORFs contained secretory signals, while 10% of all hymenopteran proteins from NCBI-nr contain secretory signals, demonstrating enrichment of secretory signals in venom-specific genes. The mechanism by which the non-signal bearing proteins are secreted into the venom is unknown. Mass spectrometry based proteomics can facilitate understanding of signal cleavage in small peptides.

The ampulexin family of peptides are well represented in both the venom apparatus transcriptome and venom proteome. These peptides are predicted to have secretory signals that are cleaved as the peptides are processed into the venom. The majority species of these peptides, as determined by spectral count, corresponds to peaks previously seen in MALDI-TOF analysis of the venom for ampulexin 1 and 3. The signal is predicted to be two amino acids N-terminal of the major ion species for ampulexin 1, and three amino acids C-terminal and of the major ion species for ampulexin 3. Ampulexin 2 is largely found as a dimer of the peptide identified one amino acid N-

terminal of the predicted cleavage site. Ampulexin 4 cleavage is divided between two major species; a peak corresponding to the mass of the smaller fragment has been observed in MALDI-TOF spectra, but the larger peak was not resolved due to noise at its predicted mass (Moore 2003) (Figure 3-7).

The majority of enzymes in *A. compressa* venom are predicted to be proteases, which are a common component of animal venoms in general (King et al. 2000), (Matsui et al. 2000) (Figure 3-8). The target of these proteases could be in the host, targeting and modulating immune response, or disrupting endogenous cerebral processes. On the other hand, or the protease activity could process venom proteins into active conformations, including other proteases, in an activation cascade. Proteases in bee venom have multiple functions, including targeting and activating the phenoloxidase cascade in arthropod targets, but also targeting fibrinogen, affecting blood clotting in mammals (Choo et al. 2010). *A. compressa* venom proteases fall in multiple families, including serine, cysteine and zinc containing metalloproteases, and are expected to have diverse substrates and activities. Second to proteases, *A. compressa* venom contains carbohydrate targeting enzymes (CTE), including hyaluronidase, trehalase, carbohydrate sulfotransferase, and glucose dehydrogenase. Presumably these enzymes target extracellular proteoglycan domains in the cockroach brain. Extracellular polysaccharides are involved in many critical functions in the central nervous system, including axonal growth and synapse formation (Frischknecht et al. 2008), (Zimmermann et al. 2008). Therefore, glycoproteins and proteoglycans are not unreasonable targets for *A. compressa* venom.

Comparative genomics

To assess how the composition of *A. compressa* venom compares to other animal venoms, 201 identified venom proteins were analyzed against sixteen venomous, and three non-venomous animal genomes. Approximately 50% of 201 identified *A. compressa* venom proteins are shared with other venomous animals, with the highest proportion of positive hits from ants and bees (115-122) (Figure 3-9). The proportion of positive hits is lowest in the mouse (24) and in the non-venomous insects, fruit fly and flour beetle (51 and 75 respectively). Interestingly the number of positive hits is reduced in the jewel wasp *Nasonia vitripennis* (78) compared to other hymenoptera and other venomous animals. *N. vitripennis* is an ectoparasitoid whose venom targets pupae of Diptera, whereas other venomous animals examined target the CNS or other body tissues of the organism, either for parasitism or defensive purposes. The difference in target may account for the difference in composition. *N. vitripennis* venom has been studied in detail and despite the differences, contains many protein classes in common with *A. compressa*, including metalloprotease, serine protease and serine protease inhibitors, chitinase and trehalase, phosphatases, and lipases. Also interestingly there was high representation of *A. compressa* venom homologs in the king cobra genome (91), similar to the black widow (91) and brown recluse (93) spiders. The bark scorpion (103) and centipede (105) had fewer hits, but were similar to that of the ants and bees.

These results suggest that conservation of certain venom proteins, in particular the protease and lipase families, extends beyond the hymenoptera clade to include venomous

animals in general. On the other hand, almost half of identified *A. compressa* venom proteins remain uncharacterized or are novel. *A. compressa* proteins in common with other venomous animals are generally confined to specific protein families. The M13 protease family is represented in all genomes examined; it is preserved in venomous animals, with a more limited representation in the non-venomous animals and *N. vitripennis*. The serpin family and cysteine-rich secretory family of proteins are present in all animals examined. The phospholipase A2 family, a ubiquitously identified venom component, has good representation in all animals examined except mouse, and to a lesser extent in *N. vitripennis* and the non-venomous insects.

Small molecule synthesis enzymes

A. compressa venom contains the GABA_A receptor agonists GABA, β -alanine and taurine (Moore et al. 2006). GABA and β -alanine are synthesized by decarboxylation of glutamate and aspartate respectively, and taurine is synthesized from cysteine through cysteine sulfinic acid. Dopamine also has been detected in the venom (Banks 2010). Transcriptomes for the venom apparatus were interrogated for enzymes that are part of the biosynthetic pathway for these compounds. A query for glutamate decarboxylase, the enzyme that synthesizes GABA, reciprocally BLASTs as both acidic amino acid decarboxylase and glutamate decarboxylase with an e-value of 0, and 6e-173 respectively. RNA-seq read counts are high in the venom sac, indicating it is the predominant locus of synthesis for GABA, β -alanine, and taurine. Dopamine synthesis enzymes, tyrosine decarboxylase and tyrosine hydroxylase, are highly expressed in the venom gland, indicating that the venom gland is the predominant locus of dopamine

synthesis (Figure 3-10). Dopamine β -hydroxylase and tyramine β -hydroxylase were not found in the transcriptome assembly, suggesting that octopamine is not synthesized in the venom apparatus.

Discussion

Venom-induced hypokinesia is remarkable for its specificity, reversibility and duration. To understand the biochemical basis for this unique behavioral alteration, a thorough analysis of the venom is required. The present work is an iterative study beginning with characterization of the venom apparatus on the transcriptomic level to build a database of transcripts and read counts. Combining transcriptomics with proteomics of milked venom identifies venom-specific transcript ORFs, which greatly facilitates comprehensive identification and characterization, as the entire ORF can be assayed against global annotated databases instead of peptide fragments. Further, this approach facilitates discovery of novel venom proteins, since novel peptide fragments would remain unidentified with proteomics alone. While transcriptome analysis yields quantitative information regarding transcript number, it does not guarantee direct correlation with translated protein levels. Conversely, proteomics alone often does not provide complete sequence information, given that sequences of non-conserved peptides may not be present in existing hymenopteran genome databases. Combining the transcriptome data with the proteomics data allows global identification of venom components and gene expression profiling of the venom apparatus.

Transcriptomics of both the venom gland and venom sac has provided insights into the compartmentalization and tissue specific expression of identified venom components. I have found that the venom gland and venom sac differ greatly in expression levels of certain venom transcripts, though each have at least some level of expression for the great majority of venom proteins, suggesting that each tissue type shares a common ancestry.

Therefore, transcriptomics reveals an interesting functional morphology of the venom apparatus as two related yet distinct glandular structures that jointly contribute to overall venom composition. In other words, neither structure alone can completely account for the protein repertoire of the venom. This precludes the idea that the venom sac serves only as a passive venom reservoir, wherein the venom gland expresses and excretes the entire complement venom components. Indeed, the ampulexins are among the most abundant venom components and are products of the venom sac, affirming that the venom sac expresses and secretes a venom gland-independent contribution to the venom.

The venom has also been shown to be acidic, and venom sac contents are acidic with respect to the venom gland (Gal et al. 2014) (Figure 3-11). It is then reasonable to conjecture that the venom is secreted from the venom gland into the venom sac, where it is supplemented with additional proteins and peptides, acidified and maintained, ready to contract and expel venom under the appropriate cue.

A. compressa venom collection is as novel as it is well-annotated. It contains a large representation of M13 proteases, especially of the neprilysin and endothelin-converting enzyme family. Hmmscan of the venom against Swiss-Prot and PfamA databases shows that 10% of the venom proteins contain M13 protease domains. These proteases are reported to typically be anchored to the extra-cellular surface of expressing cells and process neurotransmitter signaling peptides, either deactivating mature peptides or processing peptides from precursors. These proteases are the most well represented proteins in the venom, in terms of both peptide spectral counts and RNA expression levels in the venom gland. Yet only four of the seventeen proteins with M13 domains are

predicted to have secretory signals, and the only one predicted to have a secretory signal that is high expressed is neprylisin-2, an enzyme known to inactivate tachykinin peptides in mammals, interesting considering the presence of tachykinin in the venom.

The venom also contains phospholipase A2-like proteins as a major component, which are ubiquitous in venomous animals. In honeybees, it has cytolytic activity, especially in the presence of melittin, though *A. compressa* venom is not lytic (Banks 2010). Hyaluronidase, present at relatively high spectral count and expression level in *A. compressa* venom, is also found in other venoms and is thought to target the extracellular matrix (Girish et al. 2002), (King et al. 2011). Hyaluronan is a major component of the extra-cellular matrix and is important in maintaining synapse connectivity (Pyka et al. 2011), (Bikbaev et al. 2015). Disrupting synapse connections in the SEG or the central complex of the brain would disrupt integration of sensory data and impair motor function, and could also be reversible, as is the effect of the venom. The presence of phospholipase A2 and hyaluronidase enzyme activity in other venoms has been thought to be a spreading factor, in which the enzymes “loosen” the extracellular space to allow the venom to penetrate deeper into the tissue (Tu et al. 1983), (Girish et al. 2004), (Kemperaju et al. 2006), (Bordon et al. 2015). It is interesting to consider what the effect of “loosening” the cellular connectivity of a brain, without killing the cells, would have on synaptic transmission. *A. compressa* venom also contains cysteine-rich secretory proteins. Homologous proteins were found to block cyclic nucleotide-gated ion channels in snake venom, however, only the small molecular weight fraction of *A. compressa* was shown to block synaptic transmission (Yamazaki et al. 2004), (Moore 2006). The venom

contains serpins, serine protease inhibitors that act as immunosuppressants in other parasitoid wasp venom by preventing the phenoloxidase cascade (Bordon et al. 2015). Presence of serpins in the venom may allow other venom components to continue their action in the cockroach brain without interference from endogenous proteases.

The presence of 28 novel, highly expressed (> 1000 FPKM) proteins and peptides in *A. compressa* venom compel the imagination as to whether they are critical to behavioral changes resulting from envenomation. A secreted, 68 kDa novel venom protein is predicted to be completely disordered by IUPRED with 17 possible binding domains by ANCHOR (Dosztanyi et al. 2005), (Dosztanyi et al. 2009) (Figure 3-12). Intrinsically disordered proteins may serve as binding partners for a wide range of proteins, or to act as chaperones to protect protein activity. Additional possible protective or antioxidant proteins in the venom include glutathione peroxidase, perhaps to act as a preservative, protecting the venom from oxidation.

Comparison of *A. compressa* venom proteins to other venomous animals highlights those functions that are conserved in envenomation and those that may be unique to *A. compressa*. A significant portion of *A. compressa* venom proteins have some homology to other venomous animals. This is perhaps surprising considering its unique target location, the cockroach central nervous system, and the specific behavioral modification caused by the venom. The large molecular weight fraction contains proteins that are most likely homologous to those in other venomous animals, whereas the small molecular weight fraction peptides are likely to be novel. Included in the more conserved venom set are known common venom allergens such as the phospholipase A2, icarapin, and venom

acid phosphatases. The specialized ability of animal venoms to block or modify ion channel gating in the target nervous system, can often be conferred by small peptides (Yoshikami et al. 1989), (Adams 2004), (Dutertre et al. 2010), (Pringos et al. 2011). So far, *A. compressa* venom peptides have not shown this type of activity, though its small molecule fraction activates GABA_A receptors in the cockroach central nervous system (Moore 2003), (Moore et al. 2006).

This analysis is a comprehensive survey of proteins that are contained in the venom of *A. compressa*. We can infer from this list what proteins might be sufficient to induce hypokinesia, though we cannot conclude what proteins are necessary. The nature of this approach and the sensitivity of LTQ-Orbitrap Fusion mass spectrometer to detect as many proteins as possible. This will most certainly include spurious proteins not necessary for the venom's pathology, though nevertheless found in the biological fluid by "accident" through simple diffusion or errant trafficking. The dilemma then is how to identify proteins as being true venom components, i.e. is necessary in the pathology of the protein, from what may be proteins found in the venom fluid with no role in the venom action that confers fitness to the wasp, or negligible effect on the cockroach cerebral ganglia.

A practical approach to determining what venom components are sufficient to induce hypokinesia would be a systematic, functional analysis of synthetic or recombinant venom components or venom fractions. To date there has been no report of a synthetic, reversible, and long term (>24 hours) hypokinesia induction with venom or venom gland extract. This is a consequence either of the wasp's unparalleled sophistication in targeting

specific brain regions, or of activity loss due to sample handling and preparation methods.

The wasp itself can be used as tool to determine what proteins are necessary for hypokinesia induction through RNA interference. RNA interference has been implemented in wasps and bees and has even been shown to knock down venom transcripts in a *Drosophila* parasitoid, *Leptopilina boulardi* (Li-Byarlay et al. 2013), (Colinet et al. 2014). However there are limitations in how many genes may be targeted at a time, as each dsRNA species will compete for the RISC machinery and reduce efficiency of knockdown for each (Miller et al. 2012). Attempting to render *A. compressa* incompetent to induce hypokinesia by RNAi knockdown of critical venom transcripts would require that there are dsRNA-sensitive, non-redundant venom constituents, without which the venom would be unable to induce long-term hypokinesia.

Until hypokinesia can be induced in the laboratory, we must be cautious as to what we can exclude from the current venom list as spurious so as to not dismiss a critical component because of its apparent benignity or scarcity. Additionally, there is evidence of glycosylation of many of these proteins, and the possibility of other post translational modifications exists, such as phosphorylation or precursor processing, that are not revealed by MudPIT, and yet may be critical for venom action (Moore 2003), (Gnatzy et al. 2015). Whatever its limitations, this analysis will undoubtedly inform future functional analysis of the interaction of the venom and the cockroach brain, largely by identifying the full coding sequence for all proteinaceous venom components for recombinant expression or synthesis, and secondly by allowing speculations and

generation of hypotheses of venom action on the cockroach brain based on identification herein.

References Cited

- Adams, M. E. (2004). "Agatoxins: ion channel specific toxins from the American funnel web spider, *Agelenopsis aperta*." Toxicon **43**(5): 509-525.
- Asgari, S. and D. B. Rivers (2011). "Venom Proteins from Endoparasitoid Wasps and Their Role in Host-Parasite Interactions." Annu Rev Entomol **56**(1): 313-335.
- Attrill, H., et al. (2016). "FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*." Nucleic Acids Res **44**(D1): D786-792.
- Banks, C. N. (2010). The Roles of Biogenic Amines and Dopamine Receptors in Envenomation by the Parasitoid Wasp *Ampulex compressa*. Ph.D., University of California.
- Batista, C. V., et al. (2004). "Proteomics of the venom from the Amazonian scorpion *Tityus cambridgei* and the role of prolines on mass spectrometry analysis of toxins." J Chromatogr B Analyt Technol Biomed Life Sci **803**(1): 55-66.
- Bikbaev, A., et al. (2015). "Brain extracellular matrix retains connectivity in neuronal networks." Sci Rep **5**: 14527.
- Bonasio, R., et al. (2010). "Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*." Science **329**(5995): 1068-1071.
- Bordon, K. C., et al. (2015). "Arthropod venom Hyaluronidases: biochemical properties and potential applications in medicine and biotechnology." J Venom Anim Toxins Incl Trop Dis **21**: 43.
- Choo, Y. M., et al. (2010). "Dual function of a bee venom serine protease: prophenoloxidase-activating factor in arthropods and fibrin(ogen)olytic enzyme in mammals." PLoS One **5**(5): e10393.
- Church, D. M., et al. (2009). "Lineage-specific biology revealed by a finished genome assembly of the mouse." PLoS Biol **7**(5): e1000112.
- Colinet, D., et al. (2014). "Development of RNAi in a *Drosophila* endoparasitoid wasp and demonstration of its efficiency in impairing venom protein production." J Insect Physiol **63**: 56-61.

- de Graaf, D. C., et al. (2009). "Bee, wasp and ant venomics pave the way for a component-resolved diagnosis of sting allergy." J Proteomics **72**(2): 145-154.
- dos Santos, L. D., et al. (2009). "Brown recluse spider venom: proteomic analysis and proposal of a putative mechanism of action." Protein Pept Lett **16**(8): 933-943.
- Dosztanyi, Z., et al. (2005). "IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content." Bioinformatics **21**(16): 3433-3434.
- Dosztanyi, Z., et al. (2009). "ANCHOR: web server for predicting protein binding regions in disordered proteins." Bioinformatics **25**(20): 2745-2746.
- Drakakaki, G., et al. (2012). "Isolation and proteomic analysis of the SYP61 compartment reveal its role in exocytic trafficking in Arabidopsis." Cell Res **22**(2): 413-424.
- Dutertre, S. and R. J. Lewis (2010). "Use of venom peptides to probe ion channel structure and function." J Biol Chem **285**(18): 13315-13320.
- Elsik, C. G., et al. (2016). "Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine." Nucleic Acids Res **44**(D1): D793-800.
- Elsik, C. G., et al. (2014). "Finding the missing honey bee genes: lessons learned from a genome upgrade." BMC Genomics **15**: 86.
- Escoubas, P., et al. (2008). "Venomics: unravelling the complexity of animal venoms with mass spectrometry." J Mass Spectrom **43**(3): 279-295.
- Evans, V. C., et al. (2012). "De novo derivation of proteomes from transcriptomes for transcript and protein identification." Nat Methods **9**(12): 1207-1211.
- Frischknecht, R. and C. I. Seidenbecher (2008). "The crosstalk of hyaluronan-based extracellular matrix and synapses." Neuron Glia Biol **4**(3): 249-257.
- Gal, R., et al. (2014). "Sensory arsenal on the stinger of the parasitoid jewel wasp and its possible role in identifying cockroach brains." PLoS ONE **9**(2): e89683.
- Gal, R. and F. Libersat (2010). "A wasp manipulates neuronal activity in the sub-esophageal ganglion to decrease the drive for walking in its cockroach prey." PLoS One **5**(4): e10019.

- Gal, R., et al. (2005). "Parasitoid wasp uses a venom cocktail injected into the brain to manipulate the behavior and metabolism of its cockroach prey." Arch Insect Biochem Physiol **60**(4): 198-208.
- Girish, K. S., et al. (2002). "Snake venom hyaluronidase: an evidence for isoforms and extracellular matrix degradation." Mol Cell Biochem **240**(1-2): 105-110.
- Girish, K. S., et al. (2004). "Isolation and characterization of hyaluronidase a "spreading factor" from Indian cobra (*Naja naja*) venom." Biochimie **86**(3): 193-202.
- Gnatzy, W., et al. (2015). "Venom and Dufour's glands of the emerald cockroach wasp *Ampulex compressa* (Insecta, Hymenoptera, Sphecidae): structural and biochemical aspects." Arthropod Struct Dev **44**(5): 491-507.
- Grabherr, M. G., et al. (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.
- Haas, B. J., et al. (2013). "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nat Protoc **8**(8): 1494-1512.
- Haney, R. A., et al. (2014). "Dramatic expansion of the black widow toxin arsenal uncovered by multi-tissue transcriptomics and venom proteomics." BMC Genomics **15**: 366.
- Haspel, G., et al. (2005). "Parasitoid wasp affects metabolism of cockroach host to favor food preservation for its offspring." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **191**(6): 529-534.
- Haspel, G. and F. Libersat (2003). "Wasp venom blocks central cholinergic synapses to induce transient paralysis in cockroach prey." J Neurobiol **54**(4): 628-637.
- Haspel, G., et al. (2003). "Direct injection of venom by a predatory wasp into cockroach brain." J Neurobiol **56**(3): 287-292.
- Hebert, A. S., et al. (2014). "The one hour yeast proteome." Mol Cell Proteomics **13**(1): 339-347.
- Honeybee Genome Sequencing, C. (2006). "Insights into social insects from the genome of the honeybee *Apis mellifera*." Nature **443**(7114): 931-949.

- Ishihama, Y., et al. (2005). "Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein." Mol Cell Proteomics **4**(9): 1265-1272.
- Kemparaju, K. and K. S. Girish (2006). "Snake venom hyaluronidase: a therapeutic target." Cell Biochem Funct **24**(1): 7-12.
- King, T. P. and M. D. Spangfort (2000). "Structure and biology of stinging insect venom allergens." Int Arch Allergy Immunol **123**(2): 99-106.
- King, T. P. and K. M. Wittkowski (2011). "Hyaluronidase and hyaluronan in insect venom allergy." Int Arch Allergy Immunol **156**(2): 205-211.
- Li-Byarlay, H., et al. (2013). "RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee." Proc Natl Acad Sci U S A **110**(31): 12750-12755.
- Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC Bioinformatics **12**: 323.
- Libersat, F. (2003). "Wasp uses venom cocktail to manipulate the behavior of its cockroach prey." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **189**(7): 497-508.
- Love, M. I., et al. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biol **15**(12): 550.
- Matsui, T., et al. (2000). "Snake venom proteases affecting hemostasis and thrombosis." Biochim Biophys Acta **1477**(1-2): 146-156.
- Miller, S. C., et al. (2012). "Dissecting systemic RNA interference in the red flour beetle *Tribolium castaneum*: parameters affecting the efficiency of RNAi." PLoS One **7**(10): e47431.
- Moore, E. L. (2003). A Biochemical and Molecular Analysis of Venom with Distinct Physiological Actions from Two Arthropod Sources: The Parasitoid Jewel Wasp, *Amplex compressa*, of the Insect Order Hymenoptera and the Obligate Entomophagous Assassin Bug *Platymeris biguttata*, of the Insect Order Hemiptera. Ph.D., UC Riverside.

- Moore, E. L., et al. (2006). "Parasitoid wasp sting: a cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." Journal of neurobiology **66**(8): 811-820.
- Mouse Genome Sequencing, C., et al. (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-562.
- Nygaard, S., et al. (2011). "The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming." Genome Res **21**(8): 1339-1348.
- Piek, T. (1986). Venoms of the Hymenoptera: Biochemical, Pharmacological and Behavioural Aspects.
- Piek, T. (1990). "Neurotoxins from venoms of the Hymenoptera--twenty-five years of research in Amsterdam." Comp Biochem Physiol C **96**(2): 223-233.
- Piek, T., et al. (1989). "The venom of *Ampulex compressa*--effects on behaviour and synaptic transmission of cockroaches." Comp Biochem Physiol C **92**(2): 175-183.
- Poelchau, M., et al. (2015). "The i5k Workspace@NAL--enabling genomic data access, visualization and curation of arthropod genomes." Nucleic Acids Res **43**(Database issue): D714-719.
- Pringos, E., et al. (2011). "Peptide neurotoxins that affect voltage-gated calcium channels: a close-up on omega-agatoxins." Toxins (Basel) **3**(1): 17-42.
- Pyka, M., et al. (2011). "Chondroitin sulfate proteoglycans regulate astrocyte-dependent synaptogenesis and modulate synaptic activity in primary embryonic hippocampal neurons." Eur J Neurosci **33**(12): 2187-2202.
- Sadd, B. M., et al. (2015). "The genomes of two key bumblebee species with primitive eusocial organization." Genome Biol **16**: 76.
- Smith, C. D., et al. (2011). "Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*)." Proc Natl Acad Sci U S A **108**(14): 5673-5678.
- Smith, C. R., et al. (2011). "Draft genome of the red harvester ant *Pogonomyrmex barbatus*." Proc Natl Acad Sci U S A **108**(14): 5667-5672.

- Suen, G., et al. (2011). "The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle." PLoS Genet **7**(2): e1002007.
- Tribolium Genome Sequencing, C., et al. (2008). "The genome of the model beetle and pest *Tribolium castaneum*." Nature **452**(7190): 949-955.
- Tu, A. T. and R. R. Hendon (1983). "Characterization of lizard venom hyaluronidase and evidence for its action as a spreading factor." Comp Biochem Physiol B **76**(2): 377-383.
- Vonk, F. J., et al. (2013). "The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system." Proc Natl Acad Sci U S A **110**(51): 20651-20656.
- Weisel-Eichler, A., et al. (1999). "Venom of a parasitoid wasp induces prolonged grooming in the cockroach." J Exp Biol **202** (Pt 8): 957-964.
- Weisel, E., et al. (1999). "Venom of a parasitoid wasp induces prolonged grooming in the cockroach." The Journal of experimental biology **202** (Pt 8): 957-964.
- Werren, J. H., et al. (2010). "Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species." Science **327**(5963): 343-348.
- Wurm, Y., et al. (2011). "The genome of the fire ant *Solenopsis invicta*." Proc Natl Acad Sci U S A **108**(14): 5679-5684.
- Yamazaki, Y. and T. Morita (2004). "Structure and function of snake venom cysteine-rich secretory proteins." Toxicon **44**(3): 227-231.
- Yoshikami, D., et al. (1989). "The inhibitory effects of omega-conotoxins on Ca channels and synapses." Ann N Y Acad Sci **560**: 230-248.
- Zimmermann, D. R. and M. T. Dours-Zimmermann (2008). "Extracellular matrix of the central nervous system: from neglect to challenge." Histochem Cell Biol **130**(4): 635-653.

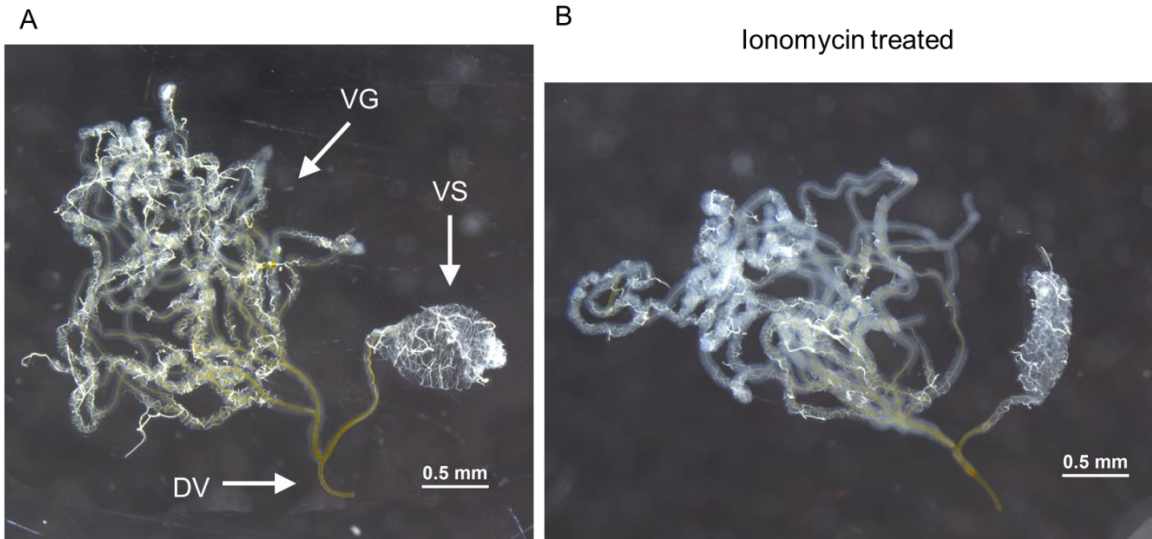


Figure 3-1. The venom apparatus of *A. compressa* is contractile

A. The venom apparatus of *A. compressa* is composed of two distinct, glandular organs: the tubular venom gland, and the venom sac which is distinct from, but connected to the venom gland at the common duct. **B.** the venom sac is contractile, as evidenced by application of ionomycin forcing contraction of the venom sac.

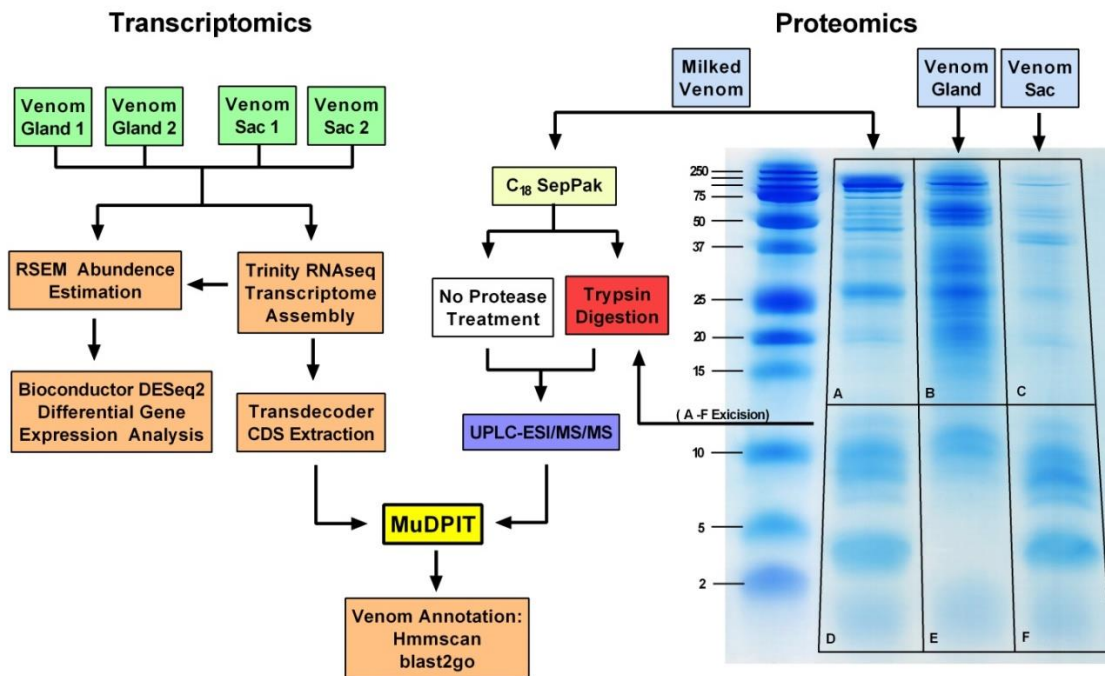


Figure 3-2. Experimental Design

Transcriptomics: RNA was sequenced from two distinct components (the venom gland and the venom sac) of the venom apparatus in biological and technical replicates. RNA sequencing reads were combined and assembled into a transcriptome using the Trinity pipeline. Transcript levels for each replicate were estimated using RSEM, and differential gene expression analysis between tissue types was examined using DESeq2 plugins for Trinity. Probable coding sequences (ORFs) were extracted from the transcriptome by Transdecoder plugin for Trinity. Proteomics: Protein was extracted from the venom gland, venom sac and milked venom and separated by tris-tricine SDS-PAGE. Fractions of the gel were excised as indicated, trypsinized and analyzed by ultra-performance liquid chromatography in-line electrospray ionization tandem mass spectrometry (UPLC-ESI/MS/MS). Additional milked venom protein was purified and concentrated by C18 SepPak and speed-vac, and analyzed by mass spectrometry, either by trypsinizing the sample, or direct analysis for better small peptide coverage. Venom and venom apparatus proteins were identified by multiple dimension protein identification technology (MudPIT) analysis using MASCOT interrogation of the Transdecoder database. Functional annotation of the identified proteins was performed with hmmscan, phmmer and SignalP 4.1, against the SwissProt and PfamA databases.

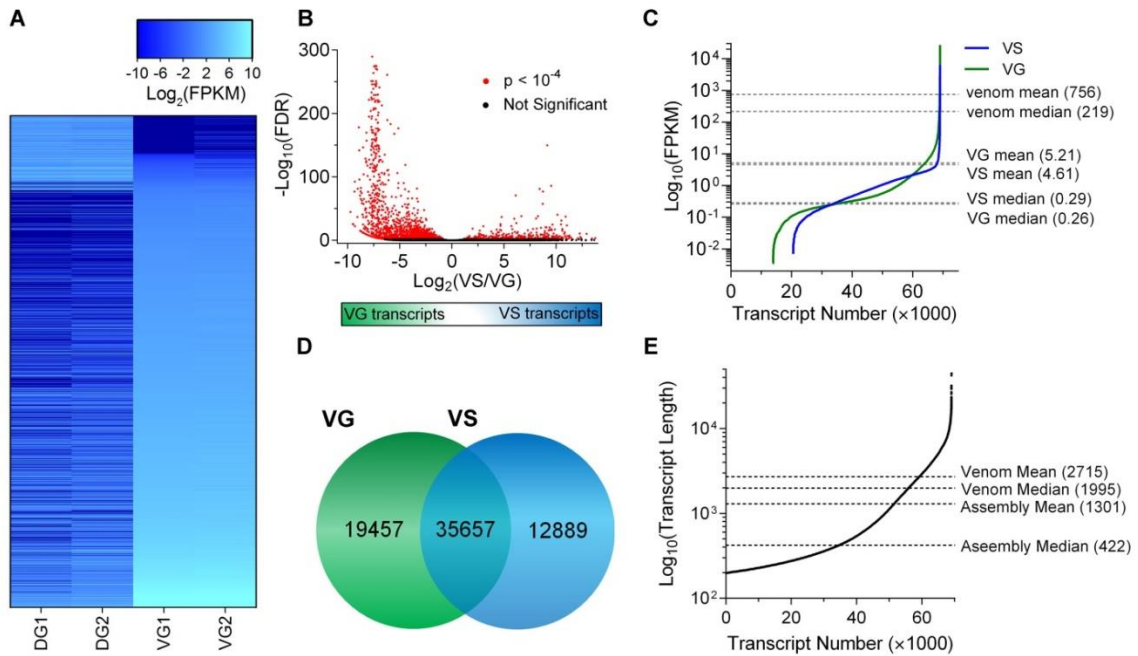


Figure 3-3. Differential expression analysis between tissue subtypes of the venom apparatus

A. Heat map representing Log_2 fold changes in expression of 2,222 differentially expressed ($p < 0.001$, > 4 -fold change) transcripts common to both the venom gland (VG) and the venom sac (VS), sorted by VG1form low to high. Columns are replicate RNA-Seq data and rows are differentially expressed transcripts between tissue types. In general, the VG has higher expression of most shared transcriptome components, though there is a population of transcripts more highly expressed in the VS. **B.** Volcano plot of fold change in transcript levels between VS and VG, and significance as false discovery rate (FDR). Positive x-values represent transcripts more highly expressed in the VS, whereas negative x-values represent transcripts more highly expressed in the VS. **C.** Transcript abundance estimation as fragments per kilobase per million mapped reads (FPKM) by transcript number. The average transcript abundance for identified venom transcripts (mean = 756) is two orders of magnitude higher than the average for all transcripts (VS mean = 4.61, VG mean = 5.21) **D.** Venn diagram of quantifiable transcripts found only in the VG, only in the VS, and transcripts common to both. Of the 35,657 transcripts shared between tissue types only 2,222 are differentially expressed. **E.** Transcript length in base-pairs by transcript number. Transcripts range in length from 200 (assembly minimum) to 44,751 bp. Venom-specific transcripts tend to be longer on average than the mean assembled transcript length.

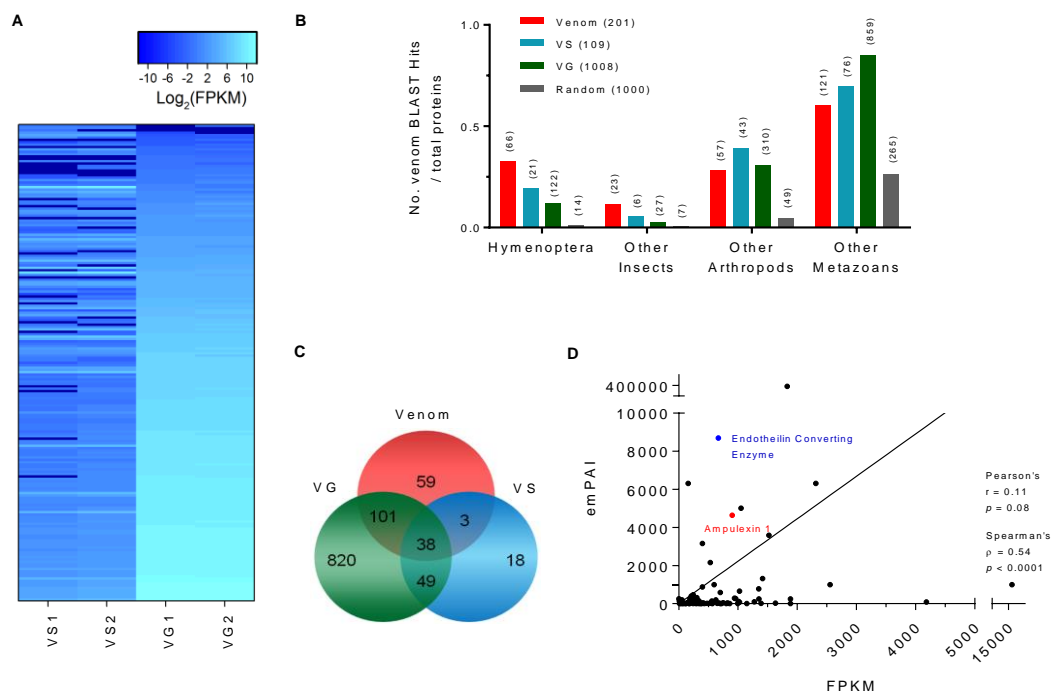


Figure 3-4. Bioinformatic analysis of the venom proteome

A. Heat map representing Log_2 fold change in transcript abundance of 201 mass spectrometry identified venom transcripts between venom gland (VG) and venom sac (VS) sorted by VG1form low to high. **B.** Milked venom, venom sac, and venom gland proteomes, and a random selection of 1000 ORFs taken from the assembly, were searched against taxonomic subsets of the NCBI-nr database. The boxplot represents the number of BLAST hits with an E-value $< 10^{-5}$ divided by the total number of proteins in the search. **C.** Venn diagram of proteins identified in milked venom, VS and VG proteomes. While there is overlap between the proteins discovered in each proteome, 59 venom proteins were not found in either venom sac or venom gland proteomes (all identified venom proteins have transcript representation in either the venom sac or venom gland or both). **D.** Correlation between transcript counts given as fragments per kilobase per million mapped reads (FPKM) and protein abundance estimation as exponentially modified protein abundance index (emPAI). Protein abundance does not functionally correlate with transcript levels (Pearson's $r = 0.11$, $p = 0.08$); however, there is a significant and positive Spearman's correlation (Spearman's $\rho = 0.54$, $p < 0.001$) between protein abundance and transcript levels.

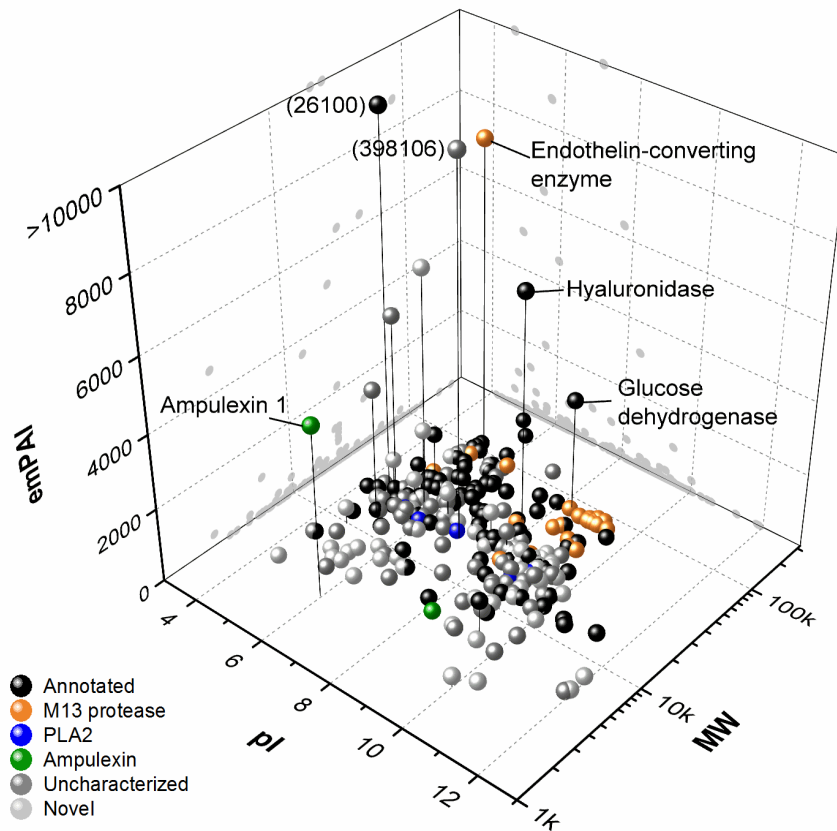


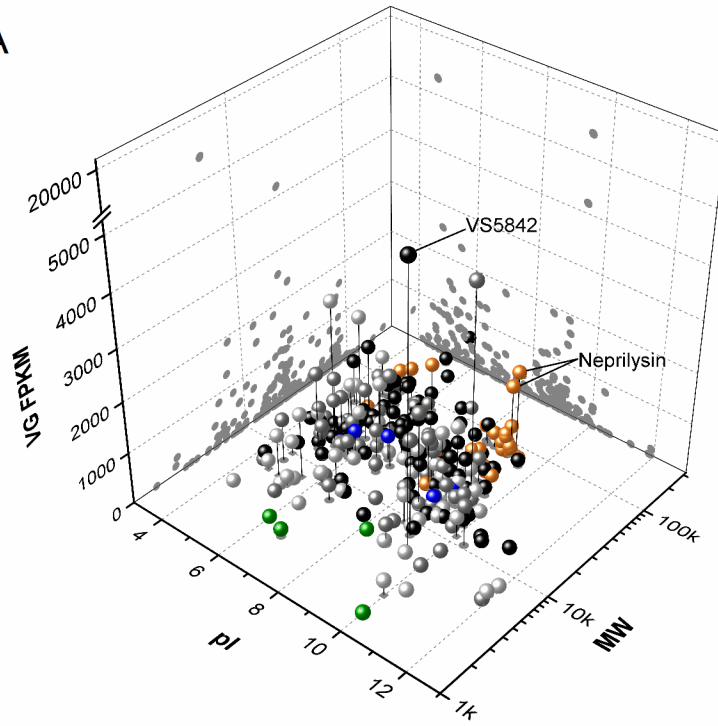
Figure 3-5. 3D ‘gel’ analysis of the venom proteome

Theoretical values of isoelectric point (pI) and molecular weight (MW) of venom proteins plotted against exponentially modified Protein Abundance Index (emPAI). The pI and MW were extrapolated from entire coding sequences identified by MudPIT. Values in parentheses represent the emPAI of the two proteins that were above 10,000. The M13 family peptidases (orange) and phospholipase A2 (blue) fall onto similar mass ranges but different pI's. The most abundant venom proteins are highlighted, notably endothelin-converting enzyme, hyaluronidase, and ampulexin 1. All other proteins identified in PfamA and Swiss-Prot databases by Hmmscan and BLAST are in black (Annotated). Uncharacterized proteins (dark gray) are differentiated from novel proteins (light gray) in that uncharacterized proteins were found represented in Uniprot database as “putative” or “uncharacterized”, whereas novel proteins did not return any significant hits ($E\text{-value} < 10^{-5}$) from Uniprot, or PfamA databases.

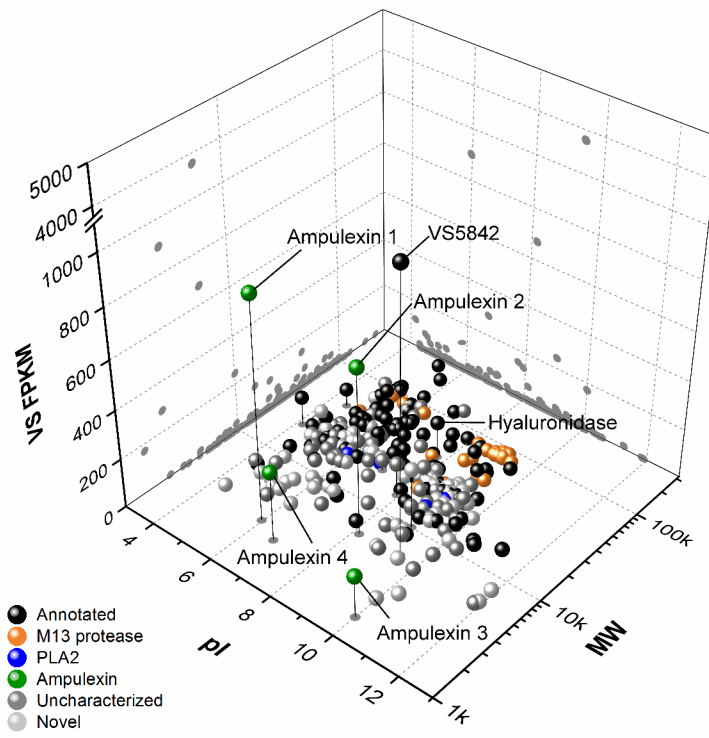
Figure 3-6. 3D ‘gel’ analysis of the venom proteome by gland expression

Theoretical values of isoelectric point (pI) and molecular weight (MW) of venom proteins plotted against relative abundance of transcripts (FPKM, fragments per kilobase per million mapped reads). **A.** Venom proteins plotted against venom gland (VG) transcript abundance. **B.** Venom proteins plotted against venom sac (VS) transcript abundance. The theoretical gel results correlate with the 1D SDS-PAGE (Figure 3-2) with higher molecular weight protein expression in the venom gland and higher expression of low molecular peptides in the venom sac. The M13 family peptidases (red) and phospholipase A2 (blue) fall onto similar mass ranges but different pI's. The most expressed venom proteins in each gland are highlighted; notably M13 peptidases and chitinase in the venom gland and hyaluronidase, serpin, and the ampulexin peptides in the venom sac. All other proteins identified by BLAST are in black. Uncharacterized proteins (grey) are differentiated from novel proteins (white) in that uncharacterized proteins were found represented in NCBI-nr database as “putative” or “uncharacterized”, whereas novel proteins did not return any significant hits ($E\text{-value} < 10^{-5}$) from Uniprot, or PfamA databases.

A



B



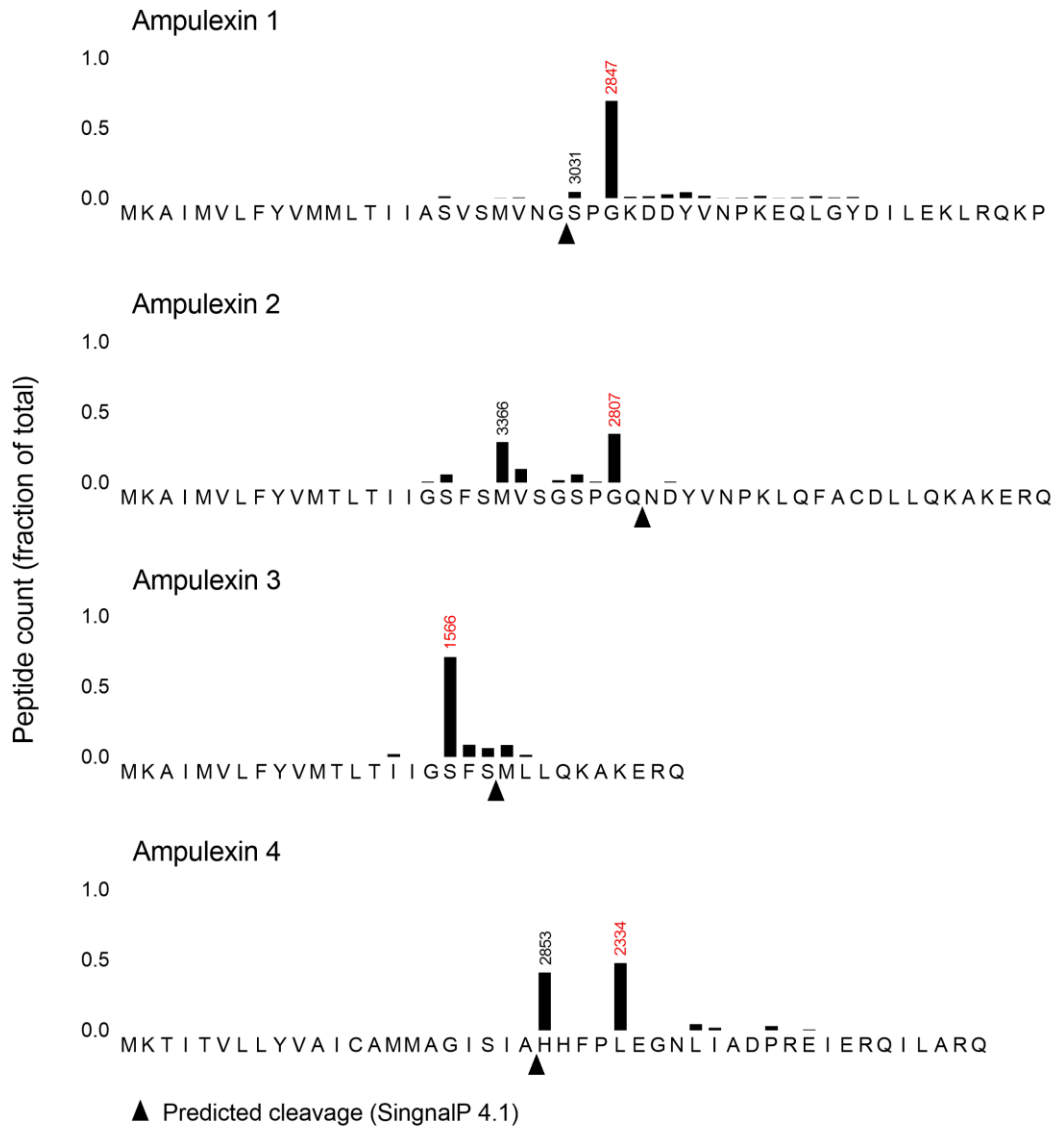


Figure 3-7. Ampulexin peptide counts and signal cleavage

The ampulexin family peptides cleavage sites as identified by mass spectrometry. The relative abundance of a peptide's spectral count is plotted as a histogram above the most N-terminal residue of the detected peptide. Predicted cleavage sites are indicated with an arrow. The molecular mass of the peptide is above the histogram, molecular mass in red has been previously detected by MALDI.

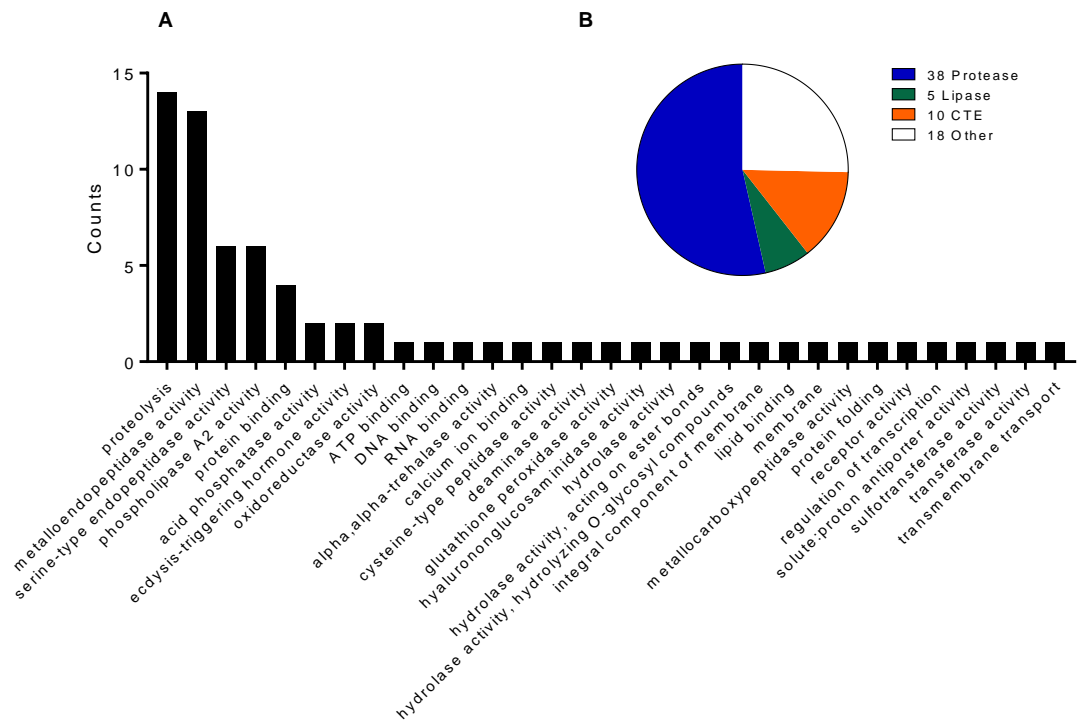


Figure 3-8. Gene ontology terms associated with venom proteins

A. Hierarchical distribution of GO terms associated with identified venom proteins. The venom proteome contains representatives of many protein activities, proteases are a dominant fraction. **B** Venom proteins with enzymatic activity by broad enzymatic classification. Proteases in the venom include serine, cysteine, and metalloproteases. A significant fraction of enzyme activity target carbohydrates or glycans as substrates and are referred to as carbohydrate-targeting-enzymes (CTE).

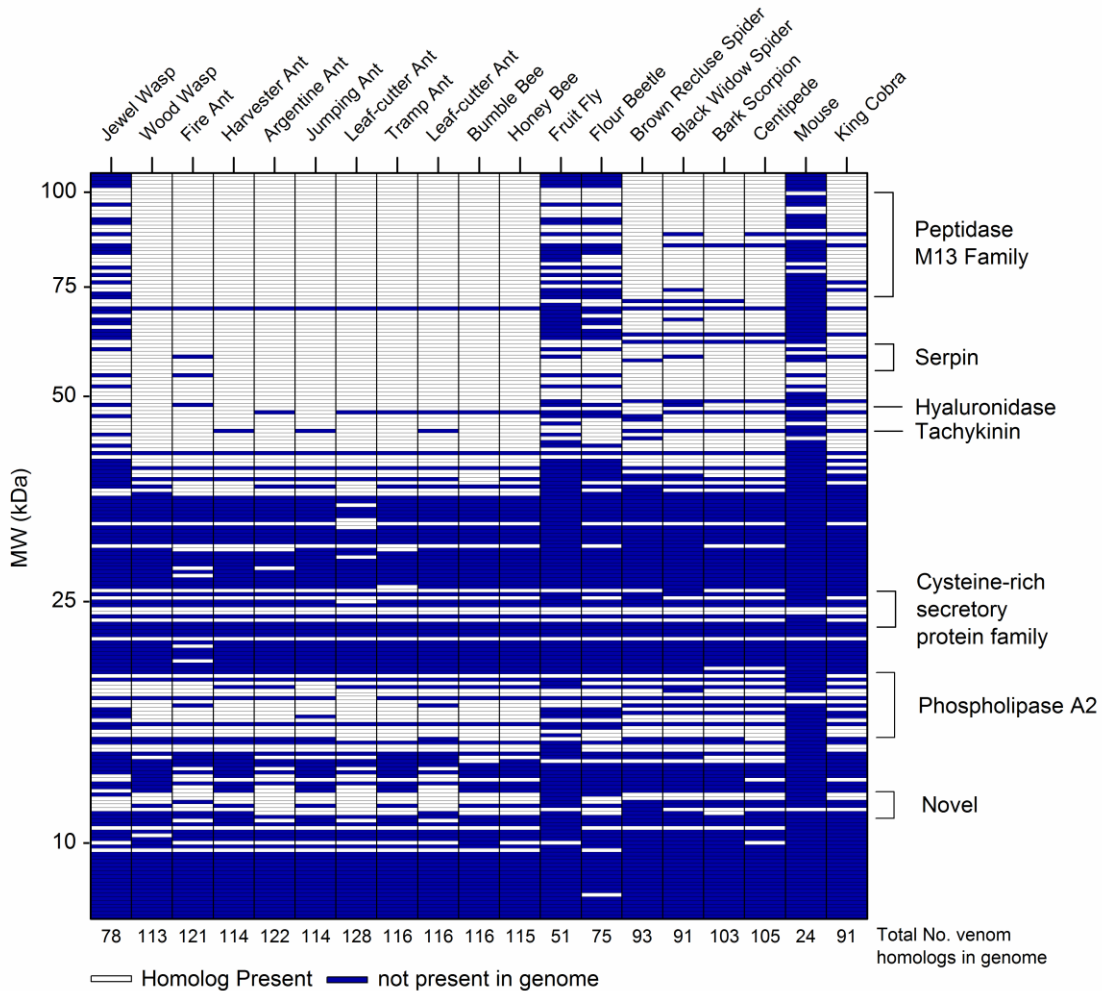


Figure 3-9. Comparative genomic analysis of *A. compressa* venom proteins

Venom proteins were searched for homologs against the respective genome using phmmer. Rows are venom proteins by theoretical molecular mass. Columns indicate whether at least one homolog was identified with an E-value cutoff $< 10^{-5}$. The total number of positive hits is summed at the bottom of the column. Common protein families across species are identified at right. Scientific names for each species are given in Methods.

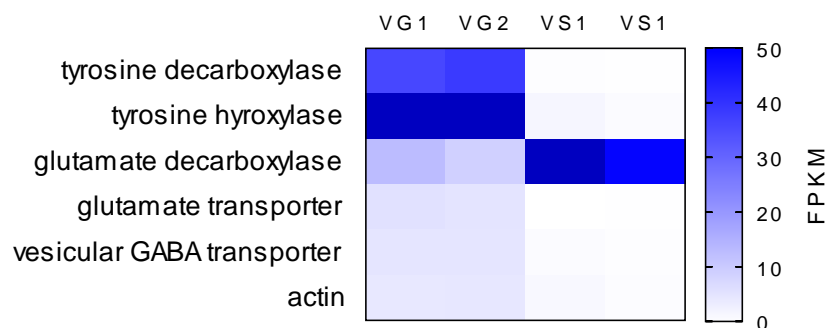


Figure 3-10. Small molecule related genes in the venom apparatus

Enzymes associated with small molecules in *A. compressa* transcriptome and associated read counts in FPKM. Glutamate decarboxylase, which synthesizes a venom component GABA is highly expressed in the venom sac. Glutamate transporter is expressed more highly in the venom gland. Synthesis enzymes for dopamine, tyrosine decarboxylase and tyrosine hydroxylase, are highly expressed in the venom gland.

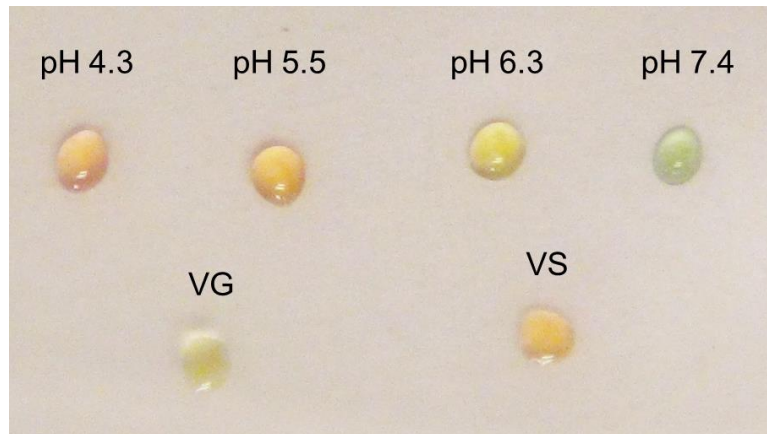


Figure 3-11. Venom sac contents are acidic

Phosphate buffered saline at varying pH was used as reference standards with a universal pH indicator solution (top row). Venom sac (VS) contents are acidic with respect to venom gland contents (VG).

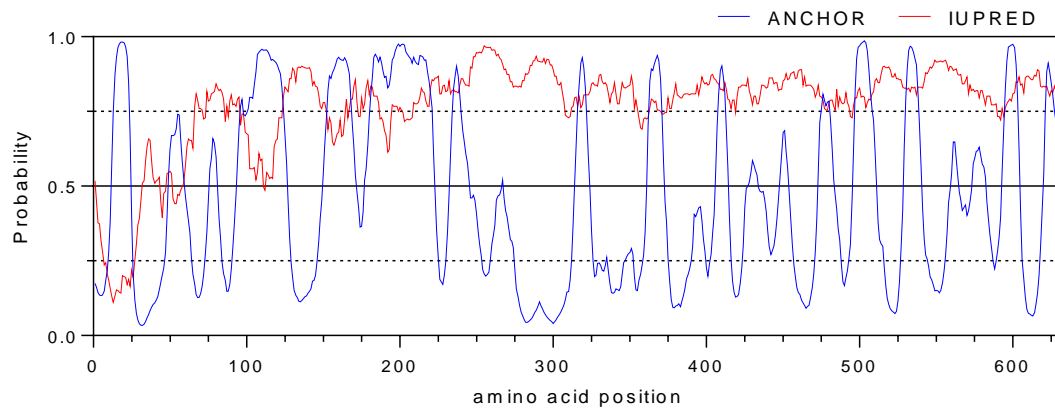


Figure 3-12. A large, secreted, novel venom protein is completely disordered

The potential for disorder as predicted by IUPRED (red trace) plotted over amino acid position suggests that the protein is completely disordered past its secretory signal (probability > 0.5 is considered disordered). The protein contains 17 putative protein binding domains as predicted ANCHOR (blue trace).

Chapter IV.

Characterization and Functional Analysis of *A. compressa* Venom Tachykinin

Abstract

The emerald jewel wasp, *Ampulex compressa*, is a parasitoid of the American cockroach *Periplaneta americana*. *A. compressa* subdues its host by envenomation of the cerebral ganglia, resulting in long-term behavioral modification characterized by an increased escape response threshold and decreased spontaneous walking. Transcriptomic and proteomic analysis of *A. compressa* venom reveals a complex mixture of enzymes and peptides, including the precursor of a known neurotransmitter, tachykinin. Tachykinin, Substance-P like peptide, has been implicated in regulation diverse behaviors, including courtship, aggression, olfaction, and locomotion. The tachykinin receptor was cloned from the cockroach subesophageal ganglion for assay of venom tachykinin activity. *A. compressa* venom tachykinins (AcVTk) activate the cockroach brain tachykinin receptor with high affinity (~5 nM) *in vitro*. Injection of synthetic AcVTk into the subesophageal ganglion, a known target of the wasp venom, increases escape response threshold for one-hour post treatment. These findings indicate that tachykinin may play an important role in the sequelae resulting from envenomation of the cockroach cerebral ganglia.

Introduction

The parasitoid jewel wasp *Ampulex compressa* is renowned for its unique ability to induce a long term behavioral modification in its host, the American cockroach. *A. compressa* subdues its host by injecting venom directly into the central nervous system of cockroach. The first sting is directed into the thorax, causing a short-term paralysis of the forelegs, allowing the wasp to subsequently deliver a second sting precisely into the cephalic ganglia (Haspel et al. 2003). Over the course of an hour, the stung cockroach engages in vigorous grooming, while the escape responses diminishes. Within two hours, the cockroach stops grooming and falls into a hypokinetic state, where the escape response is severely compromised and the cockroach remains stationary, not moving more than a few centimeters over the course of several days, unless provoked. Venom-induced hypokinesia compromises the host escape response, but leaves major motor functions, such as righting, flying and swimming intact (Gal et al. 2010). This renders the cockroach compliant to manipulation, allowing *A. compressa* to lead it into its burrow. There, the wasp lays a single egg on the coxa of the cockroach and meticulously barricades the burrow to seal the cockroach inside. (Piek et al. 1989). Over the next eight days, the cockroach remains protected and complacent as a fresh, living food supply for the developing wasp larva.

A. compressa venom is a rich pharmacopeia of enzymes, peptide toxins and neurotransmitters (Moore et al. 2006), (Banks 2010), including a ubiquitous peptide neurotransmitter, tachykinin. Tachykinins are a large family of Substance-P related peptides that are active in both the peripheral and central nervous systems of mammals

and insects (Maggio 1988), (Holzer-Petsche et al. 1995), (Vanden Broeck et al. 1999), (Siviter et al. 2000). Tachykinins are involved in many physiological and behavioral processes in insects, including nociception, gut contraction, locomotion, pheromone detection, and aggressive social behavior (Winther et al. 1998), (Siviter et al. 2000), (Winther et al. 2006), (Asahina et al. 2014), (Im et al. 2015), (Shankar et al. 2015).

Insect tachykinins have a $\text{FX}_1\text{GX}_2\text{R}$ -amide C-terminus, where X_1 and X_2 are any amino acid and activate tachykinin receptors in the nanomolar range (Torfs et al. 2000), (Poels et al. 2007), (Jiang et al. 2016). The tachykinin receptor by some accounts is $\text{G}\alpha\text{q}$ -coupled, although evidence also exists for pertussis toxin sensitive signal transduction (Li et al. 1991), (Quartara et al. 1997). Tachykinin peptides have been characterized for in two cockroach species, including *P. americana* (Johard et al. 2001), (Predel et al. 2005), (Jung et al. 2013). While, to my knowledge, there have been no reports of tachykinins in insect venoms, tachykinins have been found in octopus and stonefish venom (Hopkins et al. 1996), (Ruder et al. 2013).

Next-Generation sequencing and mass spectrometry-based proteomics revealed presence of a pre-protachykinin-like protein expressed in the *A. compressa* venom apparatus. This precursor four distinct sequences, some of which are repeated in the precursor. Interestingly, no evidence for amidation of the peptides was found in mass spectrometry data, and sequence fragments fail to show evidence for expected processing at cleavage sites known to result in active tachykinins. Since mass spectrometry analysis of venom provides evidence for presence of tachykinin precursor,

but no processed peptides, my findings therefore suggest that wasp injects tachykinin precursor into the cockroach CNS.

To establish that *A. compressa* venom tachykinins deduced from the precursor sequence would be active if processed at putative cleavage sites, I cloned the cockroach tachykinin receptor and expressed it in cell culture for aequorin based luminescence assays. Dose response curves for three unique *A. compressa* venom tachykinins, and four known cockroach tachykinins for control and comparison, were generated. To assess whether exogenously applied tachykinins mimic the effects of venom *in vivo*, synthetic tachykinin precursor and fully processed tachykinins were injected into the cockroach subesophageal ganglion and the cockroach was assayed for increased escape threshold.

Functional analysis of tachykinin venom peptides, possibly acting on cockroach tachykinin receptors could reveal a novel strategy for manipulation of host behavior and yield evolutionary insights into interspecific actions of neuropeptide signaling molecules.

Materials and Methods

Cloning of the cockroach tachykinin receptor

Five cockroach SEGs were dissected and total RNA was extracted by TRIzol method according to manufacturer's instructions (Invitrogen). cDNA was synthesized using the SuperScript III First-Strand Synthesis kit (Invitrogen) and an anchored oligo-dT primer.

Alignments of several arthropod tachykinin receptors and a partial sequence of the cockroach *Rhyarobia maderae* were aligned and degenerate primers were designed based on regions of strong homology. Amplicons were cloned into pJet1.2 and sequenced, and gene specific primers were designed for 5' and 3' RACE (rapid amplification of cDNA ends) Both 5' and 3' RACE were performed using the ExactSTART Eukaryotic mRNA 5'- & 3' RACE Kit according to manufacturer's instructions. For 5' RACE, a gene specific primer was used to generate cDNA, and a nested primer was used for PCR. For 3' RACE, cDNA was generated with a adapted, anchored oligo-dT, and nested PCR was performed with two gene-specific forward primers. The 5' and 3' amplicons were cloned into pJet1.2 and sequenced. The full coding sequence was amplified and inserted into pcDNA3.1 for expression in WTA11 cells. All PCR reaction performed with Q5 High-Fidelity DNA Polymerase (NEB). Primers are listed in Table 4-1.

Tachykinin receptor characterization

The *P. americana* tachykinin receptor was aligned to other insect tachykinin receptors with Clustal Omega (Sievers et al. 2011), (Li et al. 2015). Hydropathy plots were

predicted by on the ExPASy bioinformatics resource portal, using the Kyte and Doolittle hydrophathy scale with a sliding window of 19 (Kyte et al. 1982), (Gasteiger et al. 2003). Transmembrane domains were predicted using the TMHMM webserver (Krogh et al. 2001).

Luminescence Assays

Tachykinin receptor activity was assayed in WTA11 cells via an aequorin-based luminescence assay as previously described (Torfs et al. 2002), (Park et al. 2003), (Vleugels et al. 2013). WTA11 cells are Chinese-hamster ovary (CHO) cell clones that stably express the luminescent calcium reporter aequorin, along with the promiscuous G-protein $G\alpha_{16}$. WTA11 cells were maintained in DMEM:F12 media (Gibco) supplemented with 10% FBS (Sigma), 1X antibiotic/antimycotic, (Gibco) and 250 $\mu\text{g}/\text{ml}$ Zeocin. Cells were transfected with cockroach tachykinin receptor in the mammalian expression vector pcDNA3.1 with X-tremeGENE 9 transfection reagent (Roche). Cells were harvested with enzyme-free cell dissociation buffer and treated with coelenterazine F (Nanolight Technology) for 3 hours in suspension, protected from light. Serial dilutions of tachykinins in a 96 well plate were injected with an equal volume of cell suspension at a density of approximately 50,000 cells per well and luminescence was recorded for 20 seconds post injection on a LUMIstar Omega Microplate Reader (BMG). All peptides were synthesized by China Peptides, except AcVTk 1 and PaTk 12, which were synthesized by Peptides 2.0. All peptides were delivered at $\geq 95\%$ purity. Dose response curves were generated with GraphPad Prism's four-parameter nonlinear fit function

normalized to the response of the highest dose (2 μ M) (100% relative luminescence; % RLU).

Cockroach Brain RNA sequencing

Each of two replicates containing either two brains or two SEGs were prepared for RNA sequencing as follows: RNA was extracted by the TRIzol method, total RNA was assessed for quality and quantity by Bioanalyzer (Agilent) and cDNA libraries were constructed using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB). Sequencing was performed on an Illumina NextSeq 500. RNA preparation and sequencing was performed at the Institute for Integrative Genome Biology at the University of California, Riverside. Transcripts were assembled *de novo* using the Trinity pipeline (Grabherr et al. 2011), (Haas et al. 2013). *De novo* sequences were concatenated with a whole body transcriptome and filtered for non-redundant sequences by the EvidentialGene tr2aacds pipeline (Gilbert 2013), (Nakasugi et al. 2014), (Chen et al. 2015). Transcripts were quantified by RSEM (Li et al. 2011), (Chen et al. 2015). The transcriptome was interrogated for tachykinin and tachykinin receptor sequences by Standalone BLAST (Altschul et al. 1990).

RT-PCR of tachykinin and tachykinin receptor

RNA was extracted from cockroach subesophageal ganglia using the TRIzol method as above. cDNA was synthesized using the SuperScript III First-Strand Synthesis System (Invitrogen) using an anchored oligo-dT primer, according to manufactures instructions. Primers for protachykinin were designed based on the previously published sequence

(Predel et al. 2005). A 30-cycle PCR was performed as follows: 95 for 30 sec; 55 for 30 sec; 72 for 90 sec, using AccuPower PCR PreMix (Bioneer).

Tachykinin injection into cockroach subesophageal ganglion

Synthetic peptides were dissolved in an injection vehicle of cockroach saline (Moore et al. 2006) with 0.1 percent Janus Green B as a tracer. Injections were performed with a Drummond Nanoject II and microcapillaries beveled to 30 degrees. Cockroaches were cold-anesthetized on ice for 5 to 10 min prior to injection. Cockroaches were then placed ventral side up on a Peltier cold-plate set to 4° C to prevent the cockroach from waking during surgery, with the head pinned through the labrum to expose the neck. The submentum was cut laterally, and the anterior portion was flipped back toward the mandibles. A small platinum spoon mounted on a micro manipulator was used to move and clamp soft tissue covering the subesophageal ganglion. Tachykinins or saline controls were injected in a volume of 210 nL of solution. The submentum was replaced, and cockroaches were allowed to recover ventral side up at room temperature for 15 min.

To assess changes in the cockroach's escape response to aversive stimuli, foot shocks were administered to standing cockroaches with a Grass SD9 stimulator, with each lead connected to metal tape strip in the middle of a 30 cm radius circular arena modified from (Gal et al. 2008), (Gavra et al. 2011). Cockroaches were positioned across metal strips and stimulated with voltage pulses of 200 msec duration at 3 Hz for 3 seconds or until the animal attempted escape. The minimum voltage required to elicit an escape response was averaged over three consecutive trials, to a maximum of 20 volts, to avoid injury.

Results

Identification of protachykinin in venom

Proteomic analysis detected presence of tachykinin in milked *A. compressa* venom and in protein extracts of the venom sac (see chapter 3 for details). Close examination of peptide fragments from mass spectrometry data revealed that un-trypsinized samples did not contain mature tachykinins, yet pieces of the precursor were detected (Figure 4-1, black underlines). In trypsinized samples, each tachykinin identified in the precursor was resolved, though without amidation (Figure 4-1, red overline). This is expected, since trypsin cleaves C-terminally to basic residues and each tachykinin is flanked by a dibasic cleavage site. These data indicate that the venom contains unprocessed tachykinin precursor.

Cockroach cerebral ganglia express both tachykinin and tachykinin receptor

A. compressa venom targets cockroach cerebral ganglia, specifically the central complex of the brain and the SEG (Haspel et al. 2003). Therefore, it would seem likely that the presumed target of venom tachykinin, the tachykinin receptor, would be expressed in these regions, if venom tachykinin were to have an effect on the cockroach. It has been previously reported that several regions in the brain express tachykinin and its receptor, including the central complex (Winther et al. 2006). RT-PCR and RNA-Seq demonstrate that the SEG also expresses both tachykinin, and tachykinin receptor (Figure 4-2).

A. compressa* venom tachykinins activate the cockroach tachykinin receptor *in vitro

The *P. americana* tachykinin receptor was cloned for the purpose of assaying both endogenous and venom-derived tachykinins in cell culture. The full *P. americana* tachykinin receptor sequence is homologous to that of other insects across diverse orders, which are highly conserved in general (Figure 4-3). Simulated hydropathy plots and transmembrane domain predictions correlate well and predict seven transmembrane domains (Figure 4-4). The cockroach tachykinin receptor is predicted to couple to Gαq by PRED-COUPLE 2 with a normalized score of 0.99 (Sgourakis et al. 2005), (Sgourakis et al. 2005).

Luminescence assays show *A. compressa* venom tachykinins activate the cockroach tachykinin receptor with affinities comparable to endogenous peptides *in vitro*, with EC₅₀ values in the low nanomolar range, with an activity range over three orders of magnitude. The venom does not contain mature tachykinins, but rather the precursor protachykinin. A fragment of protachykinin that contains AcVTk 1 (proAcVTk) exhibits detectable activity, albeit in the micromolar range. The venom itself does not show activity against the receptor in cell assays at 50 stings venom per well (data not shown). This, taken together with the lack of mature tachykinins in mass spectra, indicates that the venom itself does not activate the tachykinin receptor *per se*, but may have the potential to be processed under the appropriate conditions (Figure 4-5).

A. compressa venom tachykinins are similar in sequence to each other with one variable amino acid second from the N-terminus in AcVTk 1-3; AcVTk 4 has a glycine

instead of the N-terminal alanine, but otherwise identical to AcVTk 1. The cockroach tachykinin precursor is more complex than that found in the wasp venom, containing 13 unique peptides, with no repeats in the precursor.

It does not appear that *A. compressa* tachykinin is unique to the venom, as primers designed from the venom tachykinin sequence were also able to amplify transcripts from the wasp central nervous system. These transcripts were identical to venom tachykinin by sequence analysis (data not shown). This does not preclude however, that there are additional, unidentified tachykinins in the wasp central nervous system.

A. compressa* venom tachykinins modulate escape response *in vivo

The effect of envenomation is characterized as an attenuation in escape response, by increasing escape threshold. To address whether venom tachykinins can recapitulate the effect of the venom, a synthetic venom tachykinin, AcVTk1 was injected into cockroach subesophageal ganglia and the effect on escape threshold was monitored.

Injection of synthetic *A. compressa* tachykinins causes a significant increase in escape threshold in cockroaches shortly after injection, comparable to the maximum increase induced by the wasp. However, this effect is temporary returning to non-significant levels after 2 hours. This indicates that targeting the tachykinin signaling system in the SEG can be effective in modulating escape, at least in the short term (Figure 4-6). Injection of proAcVTk did not affect escape response (data not shown).

Discussion

I report for the first time the presence of tachykinin in an insect venom, and the functional characterization of the putative target, the cockroach tachykinin receptor. The role of tachykinin in induction of hypokinesia is supported by *in vivo* injection. Venom induced hypokinesia is most likely caused by the concerted action of many elements in the venom in which tachykinin and its processing may play an interesting and critical part. I show that mature *A. compressa* tachykinins can activate the cockroach tachykinin receptor with comparable affinities to endogenous tachykinins *in vitro*. These data further support the role of tachykinin in modulating locomotion, and establish that tachykinin may modulate escape threshold in the subesophageal ganglion.

Tachykinin deficiency has been associated with hyperactivity in flies, suggesting they may be involved in suppressing locomotory activity (Winther et al. 2006), (Nassel et al. 2010). The sub-esophageal ganglion, a target of the wasp venom, regulates locomotion in cockroaches (Gal et al. 2010), (Kaiser et al. 2015), and injection of tachykinin into the subesophageal ganglion of cockroaches causes a reversible effect on its escape response. Tachykinin has been implicated in affecting presynaptic inhibition in crayfish amacrine neurons, and inhibits responses in cockroach olfactory receptor neurons (Glantz et al. 2000), (Jung et al. 2013). Given that tachykinin has been implicated in synaptic inhibition, and that lack of tachykinin signaling leads to hyperactivity, I speculate that injection of venom tachykinin may lead to synaptic inhibition in the cockroach SEG, leading to a decrease in locomotion. The wasp also targets the cockroach brain, near the

central complex, a region known to regulate locomotion. This area of the brain is reported to express tachykinin receptors (Vitzthum et al. 1998), (Johard et al. 2001).

It is compelling that tachykinin exists in the venom as an unprocessed precursor requiring processing to become bioactive. If it is processed, it must happen after injection into the cockroach brain, since the venom must be made, stored and maintained, and functionally competent to induce hypokinesia without tachykinin being processed. There are enzymes detected in the venom proteome that may be able to participate in tachykinin precursor processing: furin protease, known to target dibasic cleavage sites, such as those flanking each tachykinin, and peptidylglycine alpha-amidating monooxygenase, known to C-terminally amidate peptides. While these enzymes do not appear active in the venom sac, they may become active once injected into the cockroach brain. This may be due to significant pH change in the venom upon injection into the cephalic ganglia, among other possible chemical changes that increase activity. There is evidence that the venom is acidic, and I show that the contents of the venom sac is acidic with respect to the cockroach hemolymph. Once active in the brain, these enzymes may process the tachykinin precursor, releasing bioactive peptides. In this way the precursor-enzyme mixture may act as a time-release particle that gradually releases tachykinins to interfere with normal cockroach CNS function. Future work will test this hypothesis.

It remains possible that the mature tachykinin peptides are present in the venom at a concentration below the detection of the limit of the LTQ-Orbitrap, though mature tachykinin standards were detected when spiked into a venom sample, and milked venom does not activate cells expressing the tachykinin receptor. To my knowledge, this is the

first report of a precursor of a toxin has been found in a venom. Future experiments to reveal if this precursor can indeed be processed in the cockroach brain *in vivo* will require proteomic analysis of the cockroach brain post-sting. So far, any attempts to find wasp tachykinins in the cockroach brain have failed, however these attempts have recovered endogenous tachykinins. In fact, the only wasp venom product identified this way was the abundant ampulexin 1.

An interesting model for the onset of long term hypokinesia, defined as an increase in escape threshold that lasts longer than 24 hours, is that different venom components act over different time frames, per their specific activity. The first effects of the venom are seen immediately with the short term paralysis caused by GABA_A receptor agonist activity in the venom. This effect wanes quickly; however, due to the presence of dopamine in the venom, as the cockroach recovers from the sting, it is more concerned with grooming than fleeing, even though at this time the cockroach still has an escape threshold similar to pre-sting levels. By the time the dopamine-induced grooming bout is finished the cockroach has lost its escape response, and the wasp returns to claim its prize. I hypothesize that the time-release nature of the venom tachykinin precursor processing can activate inhibitory tachykinin receptor neurons in the subesophageal ganglion for several hours post-sting. This allows sufficient time for venom enzymes, such as hyaluronidase and matrix-metalloprotease to disrupt synapse connectivity which ultimately results in long term hypokinesia.

While tachykinin receptor neurons have been shown to be inhibitory in antennal lobe neurons, it remains to be shown that they are inhibitory in the subesophageal ganglion.

Despite the remaining uncertainty in precursor processing and exact mechanism of inhibition that, in the subesophageal ganglion, tachykinin can play a role in modulating escape response.

References Cited

- Altschul, S. F., et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Asahina, K., et al. (2014). "Tachykinin-expressing neurons control male-specific aggressive arousal in *Drosophila*." Cell **156**(1-2): 221-235.
- Banks, C. N. (2010). The Roles of Biogenic Amines and Dopamine Receptors in Envenomation by the Parasitoid Wasp *Ampulex compressa*. Ph.D., University of California.
- Chen, S., et al. (2015). "Optimizing transcriptome assemblies for leaf and seedling by combining multiple assemblies from three de novo assemblers." The Plant Genome **8**(1).
- Gal, R. and F. Libersat (2008). "A parasitoid wasp manipulates the drive for walking of its cockroach prey." Curr Biol **18**(12): 877-882.
- Gal, R. and F. Libersat (2010). "A wasp manipulates neuronal activity in the sub-esophageal ganglion to decrease the drive for walking in its cockroach prey." PLoS One **5**(4): e10019.
- Gasteiger, E., et al. (2003). "ExPASy: The proteomics server for in-depth protein knowledge and analysis." Nucleic Acids Res **31**(13): 3784-3788.
- Gavra, T. and F. Libersat (2011). "Involvement of the opioid system in the hypokinetic state induced in cockroaches by a parasitoid wasp." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **197**(3): 279-291.
- Gilbert, D. (2013). Gene-omes built from mRNA seq not genome DNA. 7th annual arthropod genomics symposium, Notre Dame.
- Glantz, R. M., et al. (2000). "Tachykinin-related peptide and GABA-mediated presynaptic inhibition of crayfish photoreceptors." J Neurosci **20**(5): 1780-1790.
- Grabherr, M. G., et al. (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.
- Haas, B. J., et al. (2013). "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nat Protoc **8**(8): 1494-1512.

- Haspel, G., et al. (2003). "Direct injection of venom by a predatory wasp into cockroach brain." Journal of neurobiology **56**(3): 287-292.
- Holzer-Petsche, U. and T. Rordorf-Nikolic (1995). "Central versus peripheral site of action of the tachykinin NK1-antagonist RP 67580 in inhibiting chemonociception." Br J Pharmacol **115**(3): 486-490.
- Hopkins, B. J., et al. (1996). "Evidence for adrenergic and tachykinin activity in venom of the stonefish (*Synanceja trachynis*)." Toxicon **34**(5): 541-554.
- Im, S. H., et al. (2015). "Tachykinin acts upstream of autocrine Hedgehog signaling during nociceptive sensitization in *Drosophila*." Elife **4**: e10735.
- Jiang, H., et al. (2016). "Ligand selectivity in tachykinin and natalisin neuropeptidergic systems of the honey bee parasitic mite *Varroa destructor*." Sci Rep **6**: 19547.
- Johard, H. A., et al. (2001). "A putative tachykinin receptor in the cockroach brain: molecular cloning and analysis of expression by means of antisera to portions of the receptor protein." Brain Res **919**(1): 94-105.
- Jung, J. W., et al. (2013). "Neuromodulation of olfactory sensitivity in the peripheral olfactory organs of the American cockroach, *Periplaneta americana*." PLoS One **8**(11): e81361.
- Kaiser, M. and F. Libersat (2015). "The role of the cerebral ganglia in the venom-induced behavioral manipulation of cockroaches stung by the parasitoid jewel wasp." J Exp Biol **218**(Pt 7): 1022-1027.
- Krogh, A., et al. (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." J Mol Biol **305**(3): 567-580.
- Kyte, J. and R. F. Doolittle (1982). "A simple method for displaying the hydropathic character of a protein." J Mol Biol **157**(1): 105-132.
- Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC Bioinformatics **12**: 323.
- Li, W., et al. (2015). "The EMBL-EBI bioinformatics web and programmatic tools framework." Nucleic Acids Res **43**(W1): W580-584.
- Li, X. J., et al. (1991). "Cloning, heterologous expression and developmental regulation of a *Drosophila* receptor for tachykinin-like peptides." The EMBO journal **10**(11): 3221-3229.

- Maggio, J. E. (1988). "Tachykinins." Annu Rev Neurosci **11**: 13-28.
- Moore, E. L., et al. (2006). "Parasitoid wasp sting: a cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." Journal of neurobiology **66**(8): 811-820.
- Nakasugi, K., et al. (2014). "Combining transcriptome assemblies from multiple de novo assemblers in the allo-tetraploid plant *Nicotiana benthamiana*." PLoS One **9**(3): e91776.
- Nassel, D. R. and A. M. Winther (2010). "Drosophila neuropeptides in regulation of physiology and behavior." Prog Neurobiol **92**(1): 42-104.
- Park, Y., et al. (2003). "Two subtypes of ecdysis-triggering hormone receptor in *Drosophila melanogaster*." J Biol Chem **278**(20): 17710-17715.
- Piek, T., et al. (1989). "The venom of *Ampulex compressa*--effects on behaviour and synaptic transmission of cockroaches." Comp Biochem Physiol C **92**(2): 175-183.
- Poels, J., et al. (2007). "Functional comparison of two evolutionary conserved insect neurokinin-like receptors." Peptides **28**(1): 103-108.
- Predel, R., et al. (2005). "Tachykinin-related peptide precursors in two cockroach species." FEBS J **272**(13): 3365-3375.
- Quartara, L. and C. A. Maggi (1997). "The tachykinin NK1 receptor. Part I: ligands and mechanisms of cellular activation." Neuropeptides **31**(6): 537-563.
- Ruder, T., et al. (2013). "Functional characterization on invertebrate and vertebrate tissues of tachykinin peptides from octopus venoms." Peptides **47**: 71-76.
- Sgourakis, N. G., et al. (2005). "Prediction of the coupling specificity of GPCRs to four families of G-proteins using hidden Markov models and artificial neural networks." Bioinformatics **21**(22): 4101-4106.
- Sgourakis, N. G., et al. (2005). "A method for the prediction of GPCRs coupling specificity to G-proteins using refined profile Hidden Markov Models." BMC Bioinformatics **6**: 104.
- Shankar, S., et al. (2015). "The neuropeptide tachykinin is essential for pheromone detection in a gustatory neural circuit." Elife **4**: e06914.
- Sievers, F., et al. (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." Mol Syst Biol **7**: 539.

Siviter, R. J., et al. (2000). "Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A." J Biol Chem **275**(30): 23273-23280.

Torfs, H., et al. (2002). "Analysis of C-terminally substituted tachykinin-like peptide agonists by means of aequorin-based luminescent assays for human and insect neurokinin receptors." Biochem Pharmacol **63**(9): 1675-1682.

Torfs, H., et al. (2000). "Characterization of a receptor for insect tachykinin-like peptide agonists by functional expression in a stable *Drosophila* Schneider 2 cell line." J Neurochem **74**(5): 2182-2189.

Vanden Broeck, J., et al. (1999). "Tachykinin-like peptides and their receptors. A review." Ann N Y Acad Sci **897**: 374-387.

Vitzthum, H. and U. Homberg (1998). "Immunocytochemical demonstration of locustatachykinin-related peptides in the central complex of the locust brain." J Comp Neurol **390**(4): 455-469.

Vleugels, R., et al. (2013). "Pharmacological characterization of a 5-HT₁-type serotonin receptor in the red flour beetle, *Tribolium castaneum*." PLoS One **8**(5): e65052.

Winther, A. M., et al. (2006). "Tachykinin-related peptides modulate odor perception and locomotor activity in *Drosophila*." Mol Cell Neurosci **31**(3): 399-406.

Winther, A. M., et al. (1998). "Characterization of actions of *Leucophaea* tachykinin-related peptides (LemTRPs) and proctolin on cockroach hindgut contractions." Peptides **19**(3): 445-458.

Designation	Sequence
Degenerate For	TGGCTKCCITWYCASMKITWCTTCAT
Degenerate Rev	TTCATCCARCAGTADATRATNGGATTG
5' RACE gene specific primer	CCAGTAAATGGCGAGATACACC
5' RACE PCR Rev	GATACACCTCCTGGATGTACTTAGTTGTGG
PeaTkR 3' RACE For 1	AGCGGTACTIONTCATCGTGACG
PeaTkR 3' RACE For 2	CACCACAATAAGTACATCCAGG
PeaTkR Nhe1 For	AGTATGCTAGCATGAACTGTAGCACGGGA
PeaTkR Kpn1 Rev	AGAAGGTACCAGACCCGACGATCTCC

Table 4-1. Primers used in cloning *P. americana* tachykinin receptor

Degenerate primers are coded according to IUPAC nucleotide code; I is inosine.

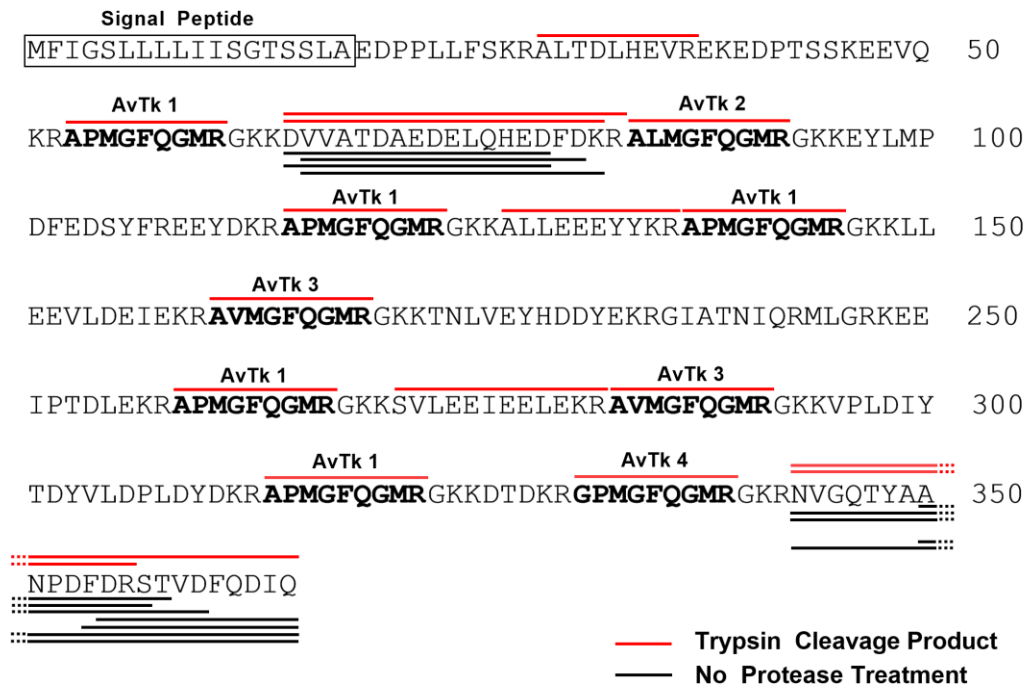


Figure 4-1. *A. compressa* venom contains tachykinin precursor

A. compressa venom preprotachykinin sequence. *A. compressa* venom tachykinins (AcVTk) are in bold and labeled, predicted secretory signal is boxed. Peptide fragments detected by mass spectrometry in a trypsinized sample are overlined in red, whereas peptide fragments in a sample not treated with protease before analysis are underlined in black.


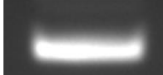
	SEG (RT PCR)	Brain (RNA Seq)	SEG (RNA Seq)
Tachykinin		1573	2590
Tachykinin Receptor		15	1.5

Figure 4-2. *P. americana* cerebral ganglia express tachykinin and tachykinin receptor

Tachykinin and tachykinin receptor transcripts are detectable in both cerebral ganglia (brain and SEG). RT-PCR analysis of the cockroach SEG yields amplicons for both tachykinin and tachykinin receptor, visualized by ethidium bromide. RNA-Seq data from transcriptomic analysis of both cerebral ganglia show expression of both tachykinin and tachykinin receptor; values are read counts in FPKM.

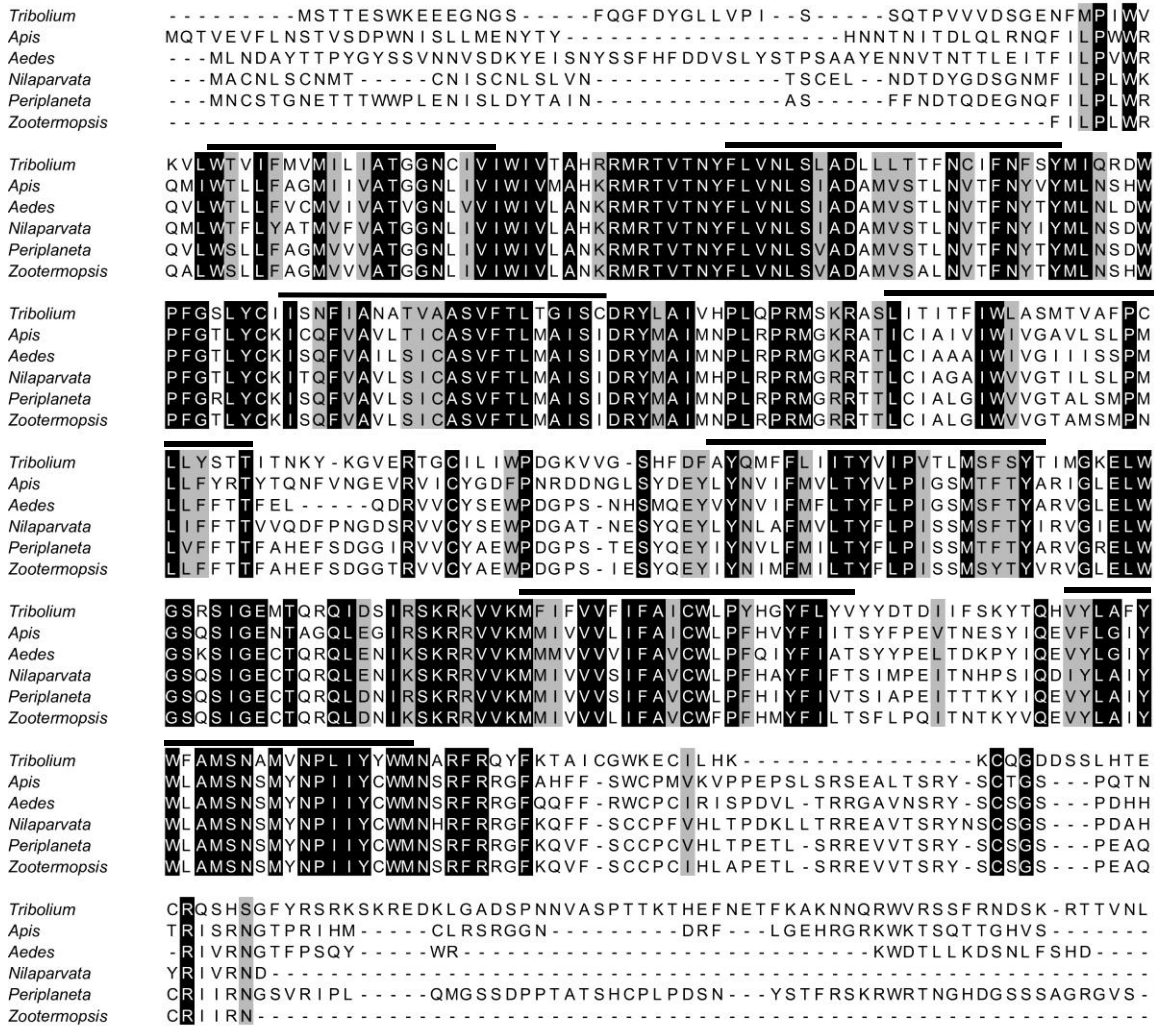


Figure 4-3. Tachykinin receptor alignment

Alignment of tachykinin receptors from diverse insect orders shows a high degree of conservation. Predicted transmembrane domains are overlined in black.

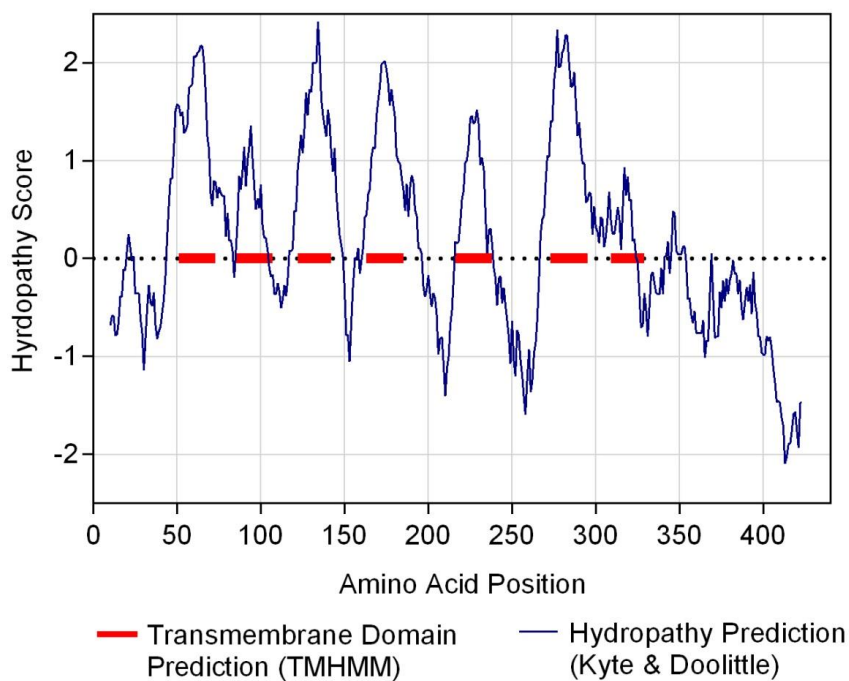
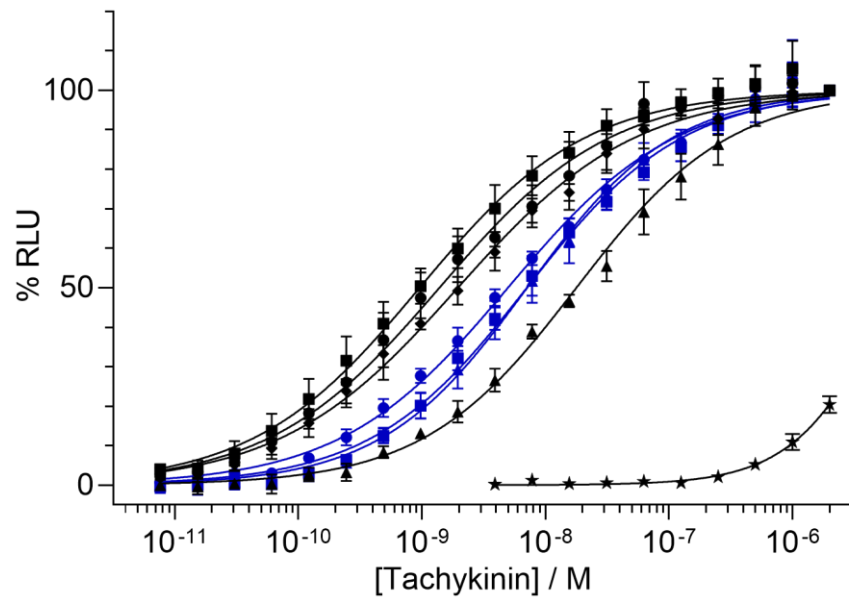


Figure 4-4. *P. americana* tachykinin receptor transmembrane domain and hydropathy plot prediction

P. americana tachykinin receptor sequence is predicted to contain seven transmembrane segments, consistent with GPCRs. Simulated hydropathy and transmembrane prediction are correlative, and predict seven transmembrane domains.



Tachykinin	EC ₅₀	Sequence
● AcVtk1	5.0 ± 0.2	APMGFQGMRa
■ AcVtk2	7.1 ± 0.3	AVMGFQGMRa
▲ AcVtk3	7.1 ± 0.4	ALMGFQGMRa
● PaTk6	1.4 ± 0.1	APASGFFGMRa
■ PaTk9	1.0 ± 0.1	APSLGFQGMRa
▲ PaTk11	17.8 ± 0.8	MGFMGMRa
◆ PaTk14	2.0 ± 0.1	APSAGFHGMRa
★ proAcVtk1		LEEEYYKR APMGFQGM RGKKLLEEV

Figure 4-5. *A. compressa* venom tachykinins activate the cockroach brain tachykinin receptor *in vitro*

Endogenous and *A. compressa* venom tachykinins were applied to WTA11 cells expressing the cockroach tachykinin receptor. Tachykinin-induced, percent-relative luminescence (% RLU) is plotted as a function of concentration. Traces of endogenous tachykinins are in black (PaTk), *A. compressa* venom tachykinin traces are in blue (AcVtk). A fragment of the tachykinin precursor (proAcVtk) was assayed at the same concentrations as mature peptides, but no EC₅₀ was calculated, internal tachykinin is in bold. EC₅₀ and sequence for each tachykinin are at right.

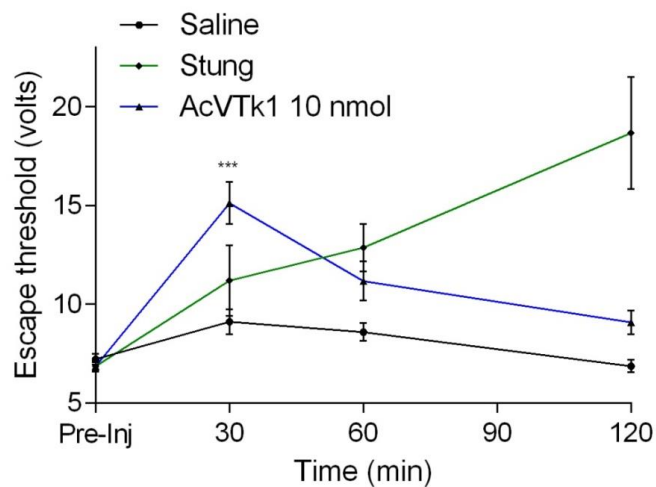


Figure 4-6. Venom tachykinin inhibits cockroach escape response *in vivo*

Injection of *A. compressa* venom tachykinin 1 (AcVTk1) into cockroach SEG increases threshold of escape significantly up to an hour after treatment (Kruskal-Wallis test, $p < 0.001$ at 30 min., $p < 0.06$ at 60 min.). Escape threshold was assayed as the minimum voltage applied to the tarsi of standing cockroaches necessary to illicit an escape response.

Chapter V.

Characterization and Functional Analysis of the Ampulexin Peptide Family

Introduction

Ampulex compressa venom contains many novel proteins and peptides. Among these, the ampulexin peptides are the most abundant. These peptides were first discovered by HPLC and MALDI-TOF analysis (Moore 2003). In these analyses, ampulexin 1 stands out as a major peak in both the chromatograms and mass spectra. Sequence analysis of this peptide revealed a novel venom toxin with a theoretical mass of 2847 Da matching the mass spectral peak precisely, indicating it was a free acid with no post-translational modifications. Similar analysis of additional major mass spectra peaks revealed a sequence with striking homology to ampulexin 1. However the theoretical mass, 2807 Da was one half that of the experimental mass of 5613 Da, indicating that it was a homodimer. This peptide, ampulexin 2, is the second most abundant transcript in the venom gland transcriptome. A third peptide, ampulexin 3 was fortuitously amplified with primers designed for ampulexin 2 with an almost identical sequence to ampulexin 2 with an internal deletion. This peptide has a theoretical mass of 1566 Da also matching a peak in mass spectra. While all three of these peptides were identified in milked venom, corresponding peaks were found almost exclusively in MALDI-TOF mass spectra of the venom sac, with very little or no respective peak in venom gland extracts. This was corroborated in RNA-Seq analysis which showed that while both the venom sac and venom gland express all the ampulexins, venom sac expression is orders of magnitude greater than that of the venom gland.

The venom apparatus transcriptome yielded four sequences in the ampulexin gene cluster. The fourth ampulexin was confirmed via MuDPIT proteomics. Its theoretical mass has a cognate and previously unidentified peak in MALDI-TOF mass spectra at 2333 Da. Ampulexin 4 has a lower expression level than the other ampulexins and is equally expressed in both the venom sac and venom gland. Predictions of the secondary structure of ampulexin 1 and 2 suggested that these peptides are alpha helical and amphipathic. To confirm its alpha-helical character, circular dichroism was used to determine secondary structure. Circular dichroism is a qualitative method for estimating the overall secondary structure of a polypeptide (Greenfield 2006). Ampulexin 1 and 2 are each composed of 24 amino acids, and in an alpha-helical conformation, they are about the width of the plasma membrane. Given that these peptides are amphipathic, it is a reasonable hypothesis that these peptides may be transmembrane, or cell penetrating. These peptides, therefore, have general features shared by cell penetrating peptides, such as a poly-basic C-terminus, and amphipathicity (Stalmans et al. 2013). The abundance and novelty of these peptides makes them intriguing targets for functional analysis.

Amphipathic alpha-helical peptides are common in hymenopteran venoms, and venoms in general (Nakajima et al. 1986). Although ubiquitous, such peptides have diverse functions and properties. A major wasp venom peptide is mastoparan, which has an overall noxious effect when injected into mammals. Mastoparan is chemotactic, attracting immune cells, and causes inflammation by releasing histamine by accelerating release of GTP from G-proteins in a GPCR independent way (Todokoro et al. 2006). This is thought to occur by mimicking the structure of a GPCR in the lipid bilayer

(Higashijima et al. 1988). Mastoparan has been compared to melittin, a 26 amino acid alpha-helical peptide, which is the principal toxic component of bee venom (Terwilliger et al. 1982). Melittin intercalates into the plasma membrane lipid bilayer, where it enhances the activity of phospholipase A2, another major hymenopteran venom protein, to cause cell lysis. Due to its ability to penetrate and destroy cells, melittin has been used in cancer therapy by conjugating it to ligands of receptors expressed in tumor cells (Orsolic 2012). Melittin infused nanoparticles also reduce the rate of HIV infection, while leaving normal cells unaffected (Hood et al. 2013). Mastoparan and melittin are also antimicrobial, having the ability to destroy a broad range of bacteria (Asthana et al. 2004), (Leandro et al. 2015). The ampulexins, despite sharing structural features with melittin and mastoparan, do not demonstrate antimicrobial or otherwise cytotoxic activity (Banks 2010). However, lack of lytic activity does not preclude the ampulexins from intercalating into, if not penetrating the plasma membrane. In doing so, the ampulexins could disrupt cellular signaling and membrane trafficking. Analysis of ampulexin 1 and ampulexin 2 in cell culture has shown that these peptides, at relatively high concentrations (100 μ M) can disrupt GPCR mediated calcium flux after thirty minutes of incubation or more (Banks 2010). Further analysis shows that these peptides can interfere with store-operated calcium entry at similar efficacies and incubation times. However, at the time of this writing, these results are not consistently reproducible, and these assays require further refinement.

It is tempting to expect that the target of ampulexins must be in the cockroach brain. They may interfere with some cellular process, or perhaps work synergistically with A.

compressa phospholipase A2, as is the case for bee melittin. However, it is also possible that ampulexins work to modulate actions of other venom components, and not the cockroach brain directly. Ampulexins might directly inhibit or allosterically modulate venom enzymes, delaying their action to extend the effect of venom, or to protect the venom from protease activity or other degradation. It is also possible that ampulexins have a beneficial or neuroprotective effect on the cockroach brain. Cell-penetrating peptides have neuroprotective effects. For example, viper venom small peptides inhibit mitochondrial swelling, and a scorpion venom peptide protects against a Parkinson's model of neural degeneration in rats (Martins et al. 2010), (Yin et al. 2014), (Meloni et al. 2015). The ampulexins represent a novel and still not well understood family of peptides whose abundance in the venom suggests an important role in the physiology of venom-induced hypokinesia.

Materials and Methods

Circular Dichroism

Circular dichroism was performed at the Analytical Chemistry Instrumentation Facility at the University of California, Riverside, on a Jasco J-815 CD spectrophotometer. Ampulexin 1 solutions were prepared at 0.5 mg/ml in 18.1 M Ω Milli-Q water and 30% aqueous 1-propanol (Fisher) where indicated. Data was acquired at 100 nm/min, 2 sec response, 1 nm bandwidth, and 8 scans, from 260 nm to 190 nm, under continuous nitrogen in a quartz, 0.1 cm cuvette.

Molecular Modeling

Structure prediction of axn1 and axn2 was performed on the I-TASSER server, Zhong laboratory, University of Michigan (Zhang 2008), (Roy et al. 2010), (Yang et al. 2015). The models with the highest c-score were chosen for visualization. The I-TASSER PDB coordinates for each model was visualized on Visual Molecular Dynamics (Humphrey et al. 1996). The axn2 dimer model was created in VMD by loading the axn2 PDB coordinates twice in anti-parallel and forcing the S-S bond between cysteines.

Synthesis of Axn2 dimer from monomer

Ampulexin 2 monomer (synthesized as previously described) and oxidized dithiothreitol (trans-4,5-dihydroxy-1,2-dithiane, Sigma) in 10-fold molar excess were added to a 50 mM phosphate buffer solution, pH 9. The reaction was incubated at room temperature on the bench for 30 min, after which it was desalted and DTT removed by SepPak, according to manufacturer's instructions.

Ampulexin effect on behavior *in vitro*

Ampulexin 1 was dissolved in cockroach saline containing 0.1 percent Janus Green B as a tracer (Moore et al. 2006). Injections were performed with a Drummond Nanoject II and microcapillaries that were beveled to a 30 degree angle. Cockroaches were cold-anesthetized on ice for 5 to 10 min prior to injection. Cockroaches were then placed ventral side up on a Peltier cold-plate set to 4°C to maintain anesthesia, with the head pinned back through the labrum to expose the neck. The submentum was cut laterally, and the anterior portion was flipped back toward the mandibles. A small platinum spoon mounted on a micromanipulator was used to move and clamp soft tissue covering the SEG. The exposed SEG was injected with 210 nL of solution. The submentum was replaced, and the cockroaches were allowed to recover ventral side up at room temperature for 15 min. Controls were injected with cockroach saline.

To assess changes in escape response to aversive stimuli, foot shocks were administered to standing cockroaches with a Grass SD9 stimulator. Opposite poles of the stimulator were connected to metal tape strips in the middle of a 30 cm radius circular arena, modified from (Gal et al. 2008), (Gavra et al. 2011). Cockroaches were positioned to be standing across the metal strips while pulses of increasing voltage (200 msec at 3 Hz) for 3 seconds or until the cockroach attempted escape. The minimum voltage required to elicit an escape was recorded for three consecutive trials and averaged.

Results and Discussion

Conformational analysis

Circular dichroism spectra of ampulexin 1 in water indicate that its mostly disordered in aqueous solutions. These spectra are similar in shape to disordered spectra, but where disordered spectra have a negative minimum at 195, ampulexin 1 in water has a large negative minimum at 201 nm. These spectra are also negative at 212 nm, where completely disordered spectra have a small positive, indicating some degree of secondary structure (Greenfield 2006). In 30% propanol, circular dichroism spectra of ampulexin 1 is stereotypically alpha helical, with negative minima at 221 nm and 209 nm and a positive maxima at 193 nm (Kelly et al. 2005) (Figure 5-1A). Molecular models of ampulexin 1 ampulexin 2 generated on the I-TASSER server were alpha-helical (Figure 5-1B and 5-1C). The models c-score and predicted accuracy are as follows: axn1 model: C-score = -1.19, TM-score = 0.57 ± 0.15 , RMSD = 3.5 ± 2.4 Å; axn2 model: C-score = -1.23, TM-score = 0.56 ± 0.15 , RMSD = 3.6 ± 2.5 Å. The alpha-helical conformation of these peptides is amphipathic with the hydrophobic residues concentrated in the middle of the peptide and charged residues toward the termini with a C-terminus that is particularly basic.

Synthesis of axn2 dimer

Oxidized dithiothreitol is a convenient and effective reagent for oxidation of two monomers of ampulexin 2 to one ampulexin 2 dimer. The reaction was followed by tris-tricine SDS-PAGE, which is also a cost-effective and convenient method for small

peptide analysis. The dimer reaction gave a strong band where the dimer was predicted to appear, with no detectable monomer. The reaction was re-reduced with β -mercaptoethanol yielding two bands. This demonstrated an incomplete reaction, though fortuitously showing that the monomer and dimer were present (Figure 5-2A). The reaction first forms an adduct between the monomer and oxidized DTT, allowing attack of a second monomer and release of reduced DTT (Figure 5-2B). According to SDS-PAGE analysis, the majority of the monomer was reacted. Simple purification with a reversed-phase SepPak would remove the reaction conditions and desalt, leaving whatever unreacted monomer behind. If necessary, further purification with a size exclusion column is recommended.

Effect of ampulexin 1 on behavior *in vivo*

To assess the possible role ampulexins may play in venom-induced hypokinesia, we assayed synthetic ampulexin 1 on escape behavior *in vivo*. Injection of ampulexin 1 into the cockroach subesophageal ganglion causes an increase in short-term escape response for up to an hour post injection (Figure 5-3A). Unlike the wasp sting, the escape threshold returns to control levels within 180 min post-injection. Interestingly, the increase in escape threshold for 50 nmol axn1 injected (estimated natural dose) mimics the wasp sting at 30 min post injection (Figure 5-3A). The effect of axn1 on escape response is dose dependent and increases escape threshold from 9 volts for vehicle injected to maximum of about 13 volts (Figure 5-3B). This is still less than the maximal wasp-sting-induced increase in escape threshold to 17 volts (Figure 5-3A).

References Cited

- Asthana, N., et al. (2004). "Dissection of antibacterial and toxic activity of melittin: a leucine zipper motif plays a crucial role in determining its hemolytic activity but not antibacterial activity." J Biol Chem **279**(53): 55042-55050.
- Banks, C. N. (2010). The Roles of Biogenic Amines and Dopamine Receptors in Envenomation by the Parasitoid Wasp *Ampulex compressa*. Ph.D., University of California.
- Gal, R. and F. Libersat (2008). "A parasitoid wasp manipulates the drive for walking of its cockroach prey." Curr Biol **18**(12): 877-882.
- Gavra, T. and F. Libersat (2011). "Involvement of the opioid system in the hypokinetic state induced in cockroaches by a parasitoid wasp." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **197**(3): 279-291.
- Greenfield, N. J. (2006). "Using circular dichroism spectra to estimate protein secondary structure." Nat Protoc **1**(6): 2876-2890.
- Higashijima, T., et al. (1988). "Mastoparan, a peptide toxin from wasp venom, mimics receptors by activating GTP-binding regulatory proteins (G proteins)." J Biol Chem **263**(14): 6491-6494.
- Hood, J. L., et al. (2013). "Cytolytic nanoparticles attenuate HIV-1 infectivity." Antivir Ther **18**(1): 95-103.
- Humphrey, W., et al. (1996). "VMD: visual molecular dynamics." J Mol Graph **14**(1): 33-38, 27-38.
- Kelly, S. M., et al. (2005). "How to study proteins by circular dichroism." Biochim Biophys Acta **1751**(2): 119-139.
- Leandro, L. F., et al. (2015). "Antimicrobial activity of apitoxin, melittin and phospholipase A(2) of honey bee (*Apis mellifera*) venom against oral pathogens." An Acad Bras Cienc **87**(1): 147-155.
- Martins, N. M., et al. (2010). "Low-molecular-mass peptides from the venom of the Amazonian viper *Bothrops atrox* protect against brain mitochondrial swelling in rat: potential for neuroprotection." Toxicon **56**(1): 86-92.

- Meloni, B. P., et al. (2015). "Neuroprotective peptides fused to arginine-rich cell penetrating peptides: Neuroprotective mechanism likely mediated by peptide endocytic properties." Pharmacol Ther **153**: 36-54.
- Moore, E. L. (2003). A Biochemical and Molecular Analysis of Venom with Distinct Physiological Actions from Two Arthropod Sources: The Parasitoid Jewel Wasp, *Amplex compressa*, of the Insect Order Hymenoptera and the Obligate Entomophagous Assassin Bug *Platyeris biguttata*, of the Insect Order Hemiptera. Ph.D., UC Riverside.
- Moore, E. L., et al. (2006). "Parasitoid wasp sting: a cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." J Neurobiol **66**(8): 811-820.
- Nakajima, T., et al. (1986). "Amphiphilic peptides in wasp venom." Biopolymers **25 Suppl**: S115-121.
- Orsolic, N. (2012). "Bee venom in cancer therapy." Cancer Metastasis Rev **31**(1-2): 173-194.
- Roy, A., et al. (2010). "I-TASSER: a unified platform for automated protein structure and function prediction." Nat Protoc **5**(4): 725-738.
- Stalmans, S., et al. (2013). "Chemical-functional diversity in cell-penetrating peptides." PLoS One **8**(8): e71752.
- Terwilliger, T. C. and D. Eisenberg (1982). "The structure of melittin. II. Interpretation of the structure." J Biol Chem **257**(11): 6016-6022.
- Todokoro, Y., et al. (2006). "Structure of tightly membrane-bound mastoparan-X, a G-protein-activating peptide, determined by solid-state NMR." Biophys J **91**(4): 1368-1379.
- Yang, J., et al. (2015). "The I-TASSER Suite: protein structure and function prediction." Nat Methods **12**(1): 7-8.
- Yin, S. M., et al. (2014). "Neuroprotection by scorpion venom heat resistant peptide in 6-hydroxydopamine rat model of early-stage Parkinson's disease." Sheng Li Xue Bao **66**(6): 658-666.
- Zhang, Y. (2008). "I-TASSER server for protein 3D structure prediction." BMC Bioinformatics **9**: 40.

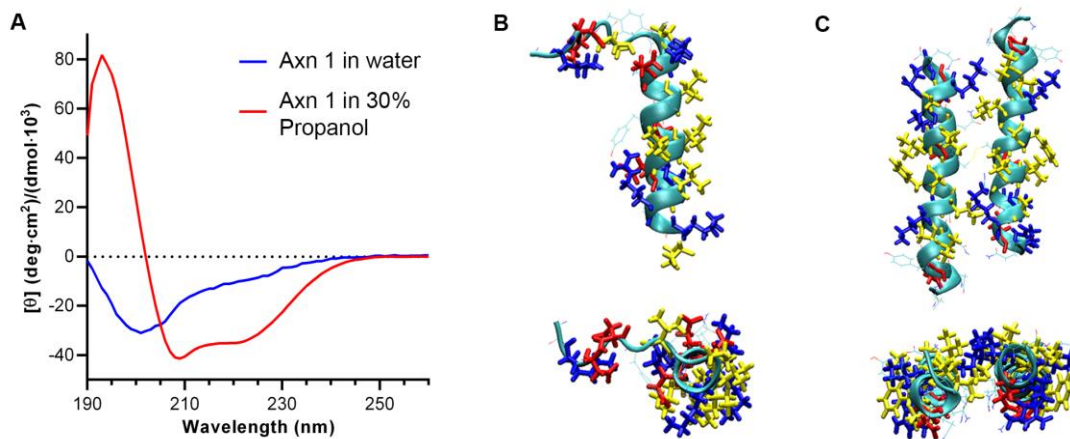


Figure 5-1. Ampulexins are amphipathic alpha-helices

A. Circular Dichroism of ampulexin 1 (axn1) in water is disordered (blue trace), and shows distinct alpha-helical spectrum in 30% propanol (red trace). **B.** Predicted 3-dimensional structure of axn1 and, **C.** axn2 dimer. Hydrophobic residues are drawn in yellow, acidic residues in red, and basic residues in blue.

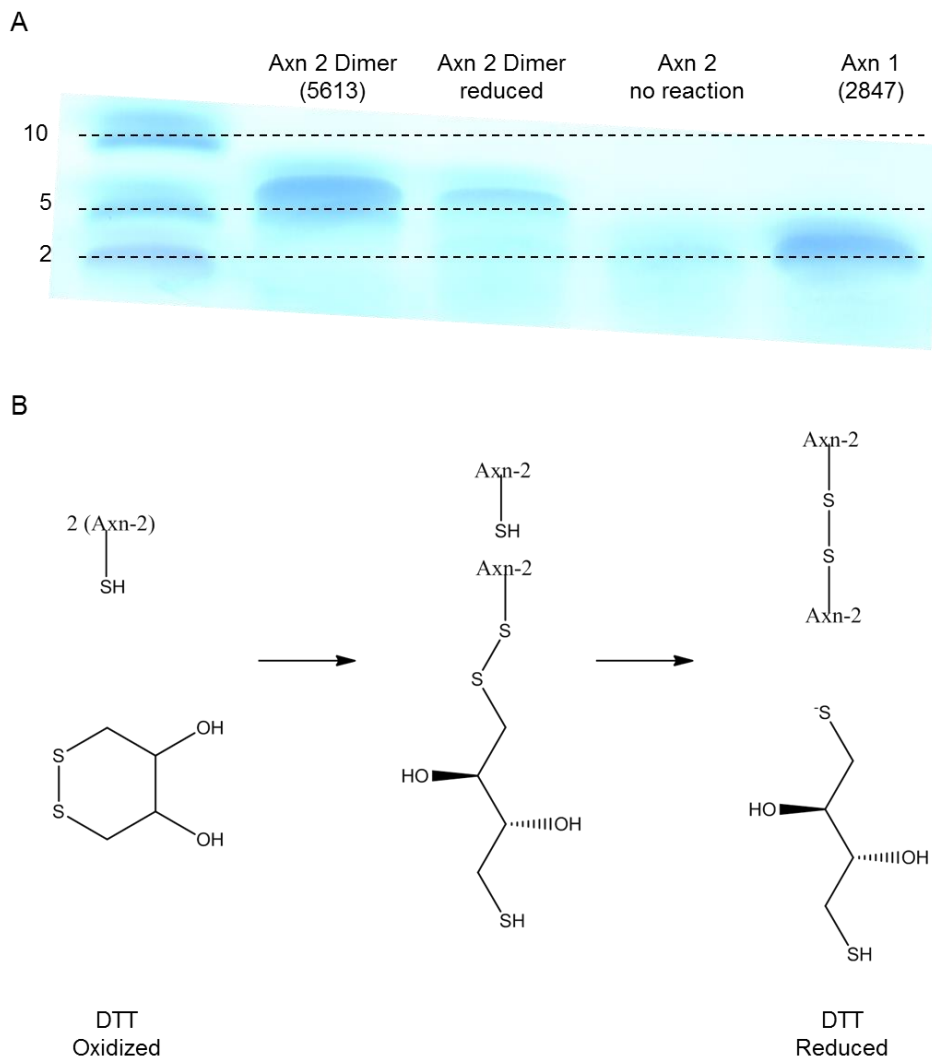


Figure 5-2. Synthesis of ampulexin 2 dimer from monomer

A cost effective and convenient way to synthesize ampulexin 2 dimer from monomer. **A.** Tris-tricine SDS PAGE gel analysis of the monomer reaction, with molecular weight standards marked at right. The second lane, Axn2 dimer, shows a band above 5 kDa, indicating the presence of a 5613 Da primer. The third lane, Axn2 dimer, reduced, shows two bands, one at the dimer mass, 5613, and one at the monomer mass, 2807. The fourth lane shows the monomer band, and the fifth lane shows axn1 for comparison and control. **B.** Schematic for the dimer reaction. Oxidized dithiothreitol (DTT) was used as an oxidizing agent, which forms a disulfide adduct with the first monomer. The second monomer attacks the first releasing reduced DTT.

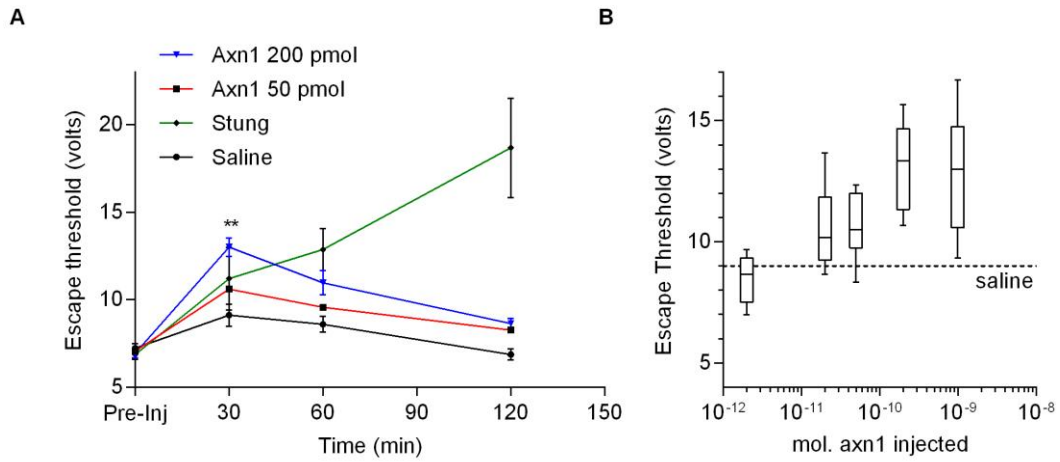


Figure 5-3. Ampulexin 1 injected in the subesophageal ganglion effects short-term escape threshold

A. Injection of ampulexin 1 (Axn1) increases cockroaches escape response to electric shock applied to the tarsi of standing cockroaches. Effect peaks at 30 min. post-injection and escape response returns to pre-injection levels after 2 hours. **B.** Dose response of ampulexin 1 injections at 30 min. post injection.

Chapter VI.

Differential Gene Expression Analysis of Cockroach Cerebral Ganglia, Following Envenomation by *Ampulex compressa*

Introduction

Envenomation of the cerebral ganglia, specifically the central complex and subesophageal ganglion, the main motor centers in the cockroach central nervous system, results in an approximately 5-7 day long hypokinesia. The duration of this condition leads to the hypothesis that changes in gene expression in these target ganglion are associated with long-term behavioral change. This was tested by quantifying global gene expression in the brain and subesophageal ganglia of control and stung cockroaches, and assessing differential gene expression patterns in both cerebral ganglia post-sting.

Next-generation RNA sequencing has been established as a powerful technique for the analysis of global gene expression in non-model organisms, as *a priori* knowledge of genetic sequences is not required (Ekblom et al. 2011). The sheer number of short sequences generated by the next-generation sequencing platforms, often referred to as “depth”, gives this technique its power, in that even relatively scarce transcripts will be sampled multiple times. The number of times a transcript fragment is sequenced is proportional to its abundance in the sample, making this technique quantitative (Kulahoglu et al. 2014). This allows comparison of gene expression levels relative to each other, or to themselves in another tissue.

Despite its power, reconstruction of transcripts is a computationally laborious process and is prone to false positives, chimeric transcripts, and truncated transcripts (Janes et al. 2015). To help mitigate these problems, a publically available whole body transcriptome of *P. americana* from the 1KITE initiative was used as a reference set of transcripts (Misof et al. 2014). Instead of using these sequences simply as a scaffold to align RNA

sequencing reads, this assembly was augmented, especially with respect to central nervous system specific transcripts, by combining with *de novo* generated transcripts from central nervous system-specific RNA sequencing data and filtering out redundant transcripts.

Quantification of differential gene expression in libraries derived from cockroach central nervous system tissues required that many central nervous system specific transcript sequences be assembled *de novo*. Methods for *de novo* assembly of transcriptomes from RNAseq data have been developed and used to analyze a number of arthropod systems (Crawford et al. 2010), (Wang et al. 2010), (Sze et al. 2012). The pipeline from RNA sequencing to differential expression analysis can be described broadly in three steps. First comes assembly of short read data into representative nucleotide sequence present in cockroach tissues. This is accomplished with programs such a Trinity and Velvet/Oases (Schulz et al. 2012), (Haas et al. 2013). Second, quantification of these transcripts is accomplished by realigning the short read data back onto the transcriptome, and counting the “depth”, or number of times a transcript has aligned a short read. This data is compiled by short read aligner programs such as RSEM or Bowtie (Langmead 2010), (Li et al. 2011). Third, biological function of identified transcripts is identified by searching global databases, such as PfamA or SwissProt for homologous sequences via search engines such as phmmer or BLAST (Altschul et al. 1990), (Bairoch et al. 1996), (Eddy 2011), (Finn et al. 2011), (Finn et al. 2014). Advances in RNA sequencing technology, a prime example being the Illumina NextSeq platform, which generates upwards of 400 million reads, allows multiplexing of multiple samples

in one sequencing experiment. Since a complete transcriptome assembly requires approximately 30 million reads, with a upper limit of 60 million reads, any additional depth does not yield new sequence discovery (Francis et al. 2013). Therefore, several relatively complete transcriptomes can be generated in parallel.

An assessment of a *de novo* assembly's completeness is an interesting challenge, since the process, by definition, recreates unknown sequences. To address this, the transcriptome is interrogated for sequences that are expected to be present in all metazoan or arthropod systems. This set of “benchmarking universal single-copy orthologs”, or BUSCOs are available in a software package from OrthoDB (Waterhouse et al. 2013), (Simao et al. 2015). Implicitly, an assembly that contains more complete BUSCOs is more complete in general than one that contains less. Furthermore, an assembly that contains multiple copies of a BUSCO is more redundant than an assembly that contains less, or one. To address issues of redundancy and completeness, the “EvidentialGene, tr2aacds” pipeline combines large datasets of multiply redundant transcriptomes, usually through multiple assembly methods, to assure completeness. It then filters for redundancy via the Cd-hit program in an attempt to yield a maximally complete, yet minimally redundant transcriptome set (Li et al. 2006), (Gilbert 2013). The *P. americana* transcriptome used in this study was a combination of the 1KITE whole body transcriptome concatenated to a *de novo* assembly using central nervous system specific reads and filtered by EvidentialGene, tr2aacds.

Tissue specific gene expression was estimated by aligning short RNA sequencing reads onto the reference transcriptome using the program RSEM (Li et al. 2011). The

number of mapped reads per transcript can then be compared between replicates and samples to determine differential expression using DESeq2 (Love et al. 2014). Those transcripts that were differentially expressed are analyzed for biological function in an attempt to describe the physiological effect of envenomation on cerebral ganglia. Additionally, gene products involved in synaptic transmission will be examined to address hypotheses that envenomation does not simply chemically interfere with synaptic function, but also genetically interferes with synaptic phenomena by effecting transcription of synaptic genes. This possibility is supported by discovery of nucleases and putative cell-penetrating peptides in the venom proteome. Lastly, in lieu of a cockroach genome, a complete transcriptome of the central nervous system, the target of *A. compressa* envenomation, is valuable for not only gene level studies, but also as a database for proteomics analysis of envenomated cerebral ganglia.

Materials and Methods

A total of 8 libraries were independently prepared and sequenced: two replicate libraries of control tissue, and two replicates of stung tissue, for both the brain and subesophageal ganglion, with 2 brains or subesophageal ganglia per sample. Cockroaches were 10-day post-eclosion, adult male, *Periplaneta americana*. Total RNA was extracted by the TRIzol method, according to manufacturer's instructions (Invitrogen). RNA quantity and quality were assessed via Bioanalyzer (Agilent). Sequencing libraries were generated using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB). Paired-end, 150 bp sequencing was performed on the Illumina NextSeq 500. Sample preparation and sequencing were performed at the Institute for Integrative Genome Biology, in the genomics core facility. Sequencing results were concatenated and assembled using the Trinity software suite with and without the "no path merging" switch (Grabherr et al. 2011). The *de novo* assembly was combined with a previously published whole body transcriptome, from the 1KITE initiative (Misof et al. 2014) and filtered using the EvidentialGene tr2aacds pipeline (Gilbert 2013), (Nakasugi et al. 2014), (Chen et al. 2015). The combined transcriptome was assessed for completeness using BUSCO (Simao et al. 2015). Transcripts were quantified using RSEM, and differential expression was determined using DESeq2 as part of the Trinity software suite (Li et al. 2011), (Haas et al. 2013), (Love et al. 2014). Synapse specific gene products were first extracted from the termite genome, and termite sequences were used as queries using phmmer and BLAST against the cockroach assembly (Terrapon et al. 2014).

Results and Discussion

The combined 1KITE and *de novo* assembly contained the most BUSCOs and therefore is considered the best assembly (Table 6-1). The *de novo* assembly itself was incomplete, owing most likely to the fact that the RNA came from a very specialized source, the central nervous system. Addition of the *de novo* sequence, however, improved the whole body 1KITE assembly enriching in central nervous system specific sequences. Another benchmark for completeness of assembly is the length of the longest sequences. The upper limit of protein length is 1800 amino acids in arthropods (Gilbert 2013). The longest 1000 protein coding sequences in a complete assembly should approach that number. The combined assembly contains the longest set of 1000 longest proteins with an average length of 1672 amino acids; however, the average of all proteins is the smallest at 203 amino acids. This indicates that the EvidentialGene tr2aacds pipeline selected the longest representative transcript of all redundant transcripts. This improves the assembly, since it is assumed that the longest assembled transcript of a gene cluster is the correct one.

Summing the number of short reads from a unique sample that align to a specific transcript is a way to quantify the relative abundance of that transcript in the sample. Quantification, especially with replicates, allows comparison between a control and an experimental tissue, in this case unstung and envenomated cerebral ganglia. The agreement between replicates shows significant increase in scatter in both the subesophageal ganglion and brain post-sting (Figure 6-1, A and B). The more divergent the agreement between replicates, the less significant any difference in transcripts

between tissue types becomes, making more subtle differences between stung and control tissue difficult to discern (Rapaport et al. 2013). Despite the scatter in the replicate data, differential expression was detected in both the subesophageal ganglion and the brain post-envenomation (Figure 6-1, C and D). Differentially expressed genes in the subesophageal ganglion are more prevalent than in the brain. However, the subesophageal ganglion is small with respect to the brain, and most of the tissue appears effected by the venom. The central complex is only a part of the larger brain, and any differences in expression levels in the central complex may be obscured by remaining brain tissue not affect by the venom (Gal et al. 2010), (Kaiser et al. 2015).

Of the set of differentially expressed genes in the subesophageal ganglion, 90 are overexpressed and 131 are underexpressed; in the brain, 90 are overexpressed and 62 are underexpressed (Figure 6-2, A and B). Interestingly, of those differentially expressed genes common to both cerebral ganglia, all but one are either upregulated or downregulated in both (Table 6-2). It is unlikely that all differentially expressed genes common to both ganglia would all altered in the same direction randomly, suggesting that these genes in particular are differentially expressed due to envenomation. However, whatever link they have to hypokinesia remains enigmatic. Synapse specific genes were analyzed in particular to show any general or specific effect on expression of synaptic proteins. If the venom disrupts synaptic structure, it might be expected that changes in synapse specific genes would be global. This was not seen in these data. Rather, specific genes were seen as differentially expressed, while others were not (Table 6-3). This could indicate a specific effect of the venom on certain synapse genes in particular, such as

synapsin, CAST1 and bruchpilot, rather than a global change in synapse gene expression. Interestingly the only synapse gene analyzed that was differentially expressed in the brain was a down regulated L-type calcium channel.

While it is expected that there would be significant alteration of gene expression in the cerebral ganglia after the physical trauma of being pierced by a stinger, it will be a challenge to determine what, if any, expression level changes are caused by venom action or what expression level changes contribute to venom-induced hypokinesia. Since hypokinesia is reversible, effects of the venom eventually subside, and gene expression changes are expected to return to a pre-stung state. This may not be the case however, and future experiments profiling recovery from hypokinesia should involve monitoring those differentially expressed genes.

Of these genes, bruchpilot stands out as most interesting. Changes in bruchpilot expression, has been shown to change according to a circadian rhythm in synapse active zones in *Drosophila* (Gorska-Andrzejak et al. 2013). Changes in bruchpilot expression in synapse active zones also has been implicated in sleep state in *Drosophila* (Gilestro et al. 2009). The effect of the venom on bruchpilot expression in the cockroach brain has been confirmed on the protein level, and bruchpilot actions in sleep and escape response are actively being researched (Maayan Kaiser, personal communication). Changes in gene expression may play a role in hypokinesia, however, there is still a significant amount of hypothesis testing and functional analysis remaining to confirm any particular role. An interesting approach would be injection of double stranded RNA targeting bruchpilot into the cockroach cerebral ganglia, or other differentially expressed genes, and monitor any

effects on behavior. Gene knockdown studies using RNAi have been effective in cockroaches previously (Maestro et al. 2006), (Irls et al. 2013). Such experiments lay a foundation for future examination into the possible genetic role of venom induced hypokinesia, and provide a promising approach for identifying target genes in the central nervous system in general.

References Cited

- Altschul, S. F., et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Bairoch, A. and R. Apweiler (1996). "The SWISS-PROT protein sequence data bank and its new supplement TREMBL." Nucleic Acids Res **24**(1): 21-25.
- Chen, S., et al. (2015). "Optimizing transcriptome assemblies for leaf and seedling by combining multiple assemblies from three de novo assemblers." The Plant Genome **8**(1).
- Crawford, J. E., et al. (2010). "De novo transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology." PLoS One **5**(12): e14202.
- Eddy, S. R. (2011). "Accelerated Profile HMM Searches." PLoS Comput Biol **7**(10): e1002195.
- Eklom, R. and J. Galindo (2011). "Applications of next generation sequencing in molecular ecology of non-model organisms." Heredity (Edinb) **107**(1): 1-15.
- Finn, R. D., et al. (2014). "Pfam: the protein families database." Nucleic Acids Res **42**(Database issue): D222-230.
- Finn, R. D., et al. (2011). "HMMER web server: interactive sequence similarity searching." Nucleic Acids Res **39**(Web Server issue): W29-37.
- Francis, W. R., et al. (2013). "A comparison across non-model animals suggests an optimal sequencing depth for de novo transcriptome assembly." BMC Genomics **14**: 167.
- Gal, R. and F. Libersat (2010). "A wasp manipulates neuronal activity in the sub-esophageal ganglion to decrease the drive for walking in its cockroach prey." PLoS One **5**(4): e10019.
- Gilbert, D. (2013). Gene-omes built from mRNA seq not genome DNA. 7th annual arthropod genomics symposium, Notre Dame.
- Gilestro, G. F., et al. (2009). "Widespread changes in synaptic markers as a function of sleep and wakefulness in *Drosophila*." Science **324**(5923): 109-112.
- Gorska-Andrzejak, J., et al. (2013). "Circadian expression of the presynaptic active zone protein Bruchpilot in the lamina of *Drosophila melanogaster*." Dev Neurobiol **73**(1): 14-26.

- Grabherr, M. G., et al. (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.
- Haas, B. J., et al. (2013). "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nat Protoc **8**(8): 1494-1512.
- Irles, P., et al. (2013). "RNAi reveals the key role of Nervana 1 in cockroach oogenesis and embryo development." Insect Biochem Mol Biol **43**(2): 178-188.
- Janes, J., et al. (2015). "A comparative study of RNA-seq analysis strategies." Brief Bioinform **16**(6): 932-940.
- Kaiser, M. and F. Libersat (2015). "The role of the cerebral ganglia in the venom-induced behavioral manipulation of cockroaches stung by the parasitoid jewel wasp." J Exp Biol **218**(Pt 7): 1022-1027.
- Kulahoglu, C. and A. Brautigam (2014). "Quantitative transcriptome analysis using RNA-seq." Methods Mol Biol **1158**: 71-91.
- Langmead, B. (2010). "Aligning short sequencing reads with Bowtie." Curr Protoc Bioinformatics **Chapter 11**: Unit 11 17.
- Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC Bioinformatics **12**: 323.
- Li, W. and A. Godzik (2006). "Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences." Bioinformatics **22**(13): 1658-1659.
- Love, M. I., et al. (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biol **15**(12): 550.
- Maestro, J. L. and X. Belles (2006). "Silencing allatostatin expression using double-stranded RNA targeted to preproallatostatin mRNA in the German cockroach." Arch Insect Biochem Physiol **62**(2): 73-79.
- Misof, B., et al. (2014). "Phylogenomics resolves the timing and pattern of insect evolution." Science **346**(6210): 763-767.
- Nakasugi, K., et al. (2014). "Combining transcriptome assemblies from multiple de novo assemblers in the allo-tetraploid plant *Nicotiana benthamiana*." PLoS One **9**(3): e91776.
- Rapaport, F., et al. (2013). "Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data." Genome Biol **14**(9): R95.

- Schulz, M. H., et al. (2012). "Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels." Bioinformatics **28**(8): 1086-1092.
- Simao, F. A., et al. (2015). "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs." Bioinformatics **31**(19): 3210-3212.
- Sze, S. H., et al. (2012). "A de novo transcriptome assembly of *Lucilia sericata* (Diptera: Calliphoridae) with predicted alternative splices, single nucleotide polymorphisms and transcript expression estimates." Insect Mol Biol **21**(2): 205-221.
- Terrapon, N., et al. (2014). "Molecular traces of alternative social organization in a termite genome." Nat Commun **5**: 3636.
- Wang, X. W., et al. (2010). "De novo characterization of a whitefly transcriptome and analysis of its gene expression during development." BMC Genomics **11**: 400.
- Waterhouse, R. M., et al. (2013). "OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs." Nucleic Acids Res **41**(Database issue): D358-365.

	<i>De novo</i> assembly	1KITE assembly	merged assembly
Total number transcripts	145228	54937	1480226
N50	422	1148	2013
Filter by <i>Evidentialgene</i>			
Total number transcripts after filter	37107	15022	60277
N50	465	1780	1469
Average length of top 1000 proteins (aa)	400	1140	1672
Average length of all proteins (aa)	368	306	203
BUSCO: Arthropoda			
Complete Single-copy BUSCOs	259	1276	1483
Complete Duplicated BUSCOs	12	140	130
Fragmented BUSCOs	726	312	108
Missing BUSCOs	1678	947	954
Total BUSCO groups searched	2675	2675	2675
BUSCO: Metazon			
Complete Single-copy BUSCOs	113	468	21
Complete Duplicated BUSCOs	11	67	768
Fragmented BUSCOs	350	134	44
Missing BUSCOs	369	174	10
Total BUSCO groups searched	843	843	843

Table 6-1. *P. americana* cerebral ganglia transcriptome assembly statistics

Assembly statistics comparing the *de novo* transcriptome, the 1KITE whole body transcriptome, and a merged transcriptome of a *de novo* assembly, and the 1KITE assembly. Each assembly was determined for completeness using BUSCO (benchmarking universal single copy orthologs), with both an arthropod specific, and metazoan specific BUSCO sets. The merged assembly contains more complete BUSCOs than either of the component assemblies, indicating it is a more complete assembly, however with more redundancy in the metazoan set, as indicated by large number of duplicated BUSCOs.

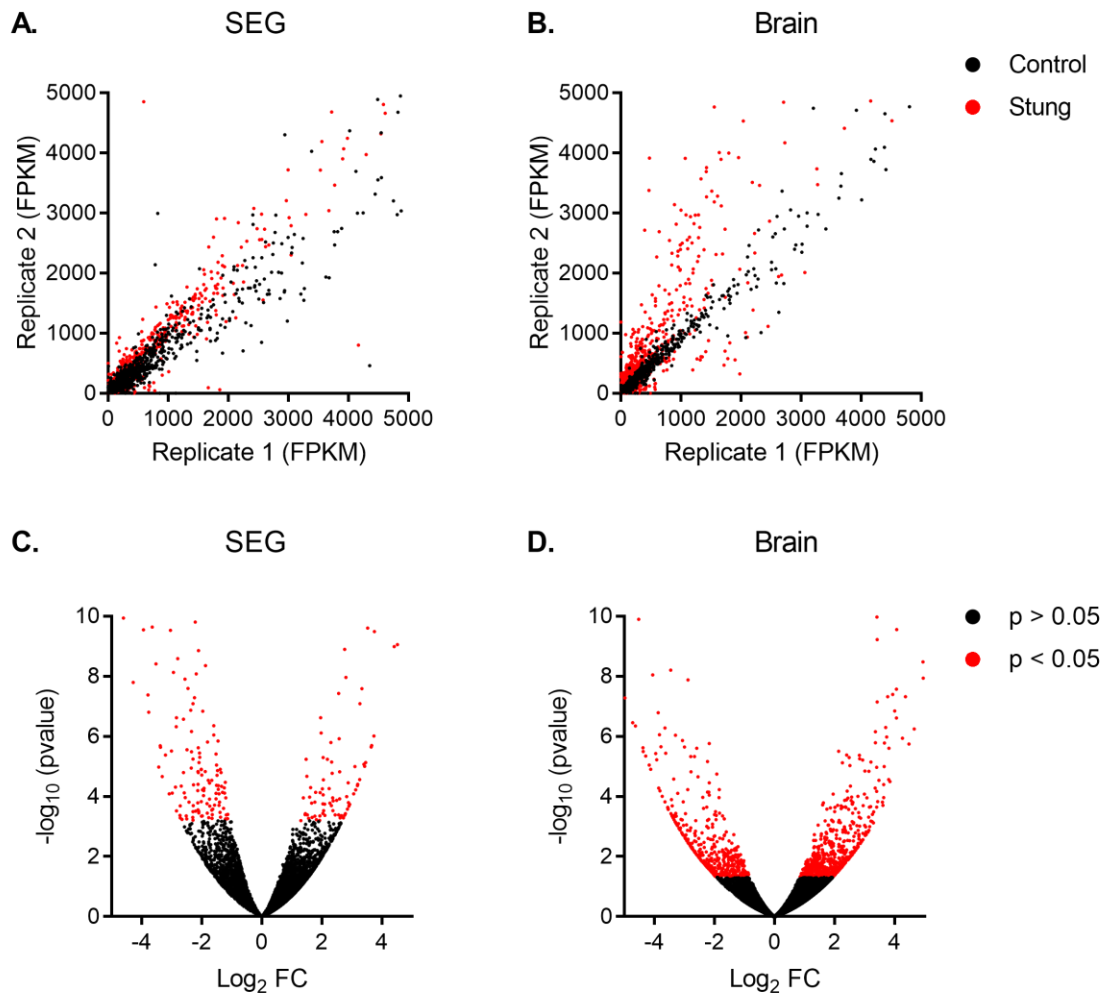


Figure 6-1. RNA sequencing data analysis

Replicate agreement in the SEG, **A**, and the brain, **B**, shows increased scatter after the sting, indicating a systematic effect of the sting on gene expression. Volcano plots of the subesophageal ganglion, **C**, and brain, **D**, show that indeed, the sting causes differential gene expression in both cerebral ganglia.

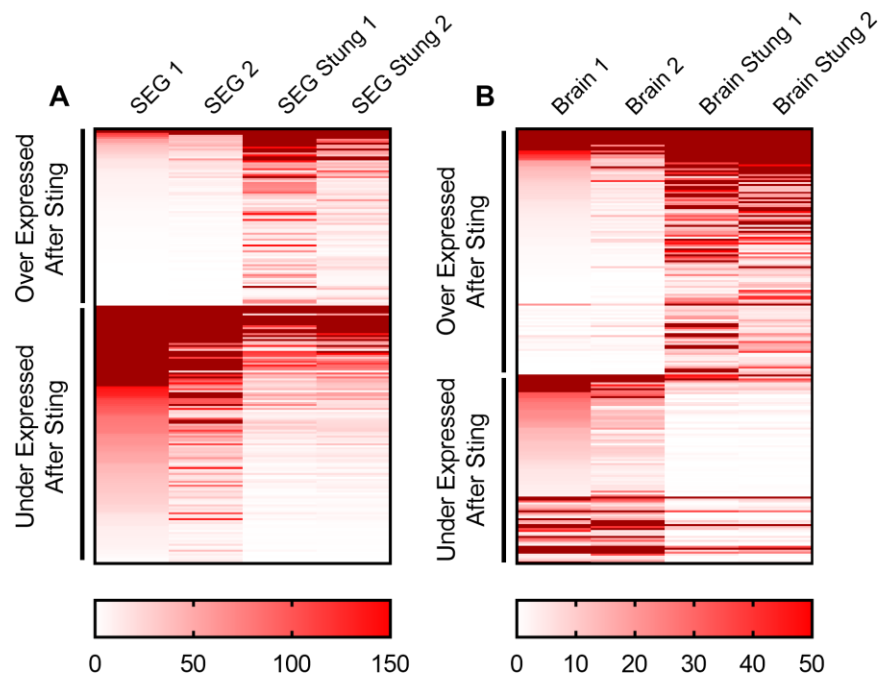


Figure 6-2. Differential expression in the cerebral ganglion post-sting

Heat maps representing differential expression in **A.** subesophageal ganglion and **B.** brain. Both the brain and subesophageal ganglion underexpress and overexpress genes post-venomation.

phmmer result	Eval	SEG	SEG Stung	p val	brain	brain stung	p val	
Putative facilitated trehalose transporter	4E-148	40	3103	7.36E-19	19	891	4.05E-17	Overexpressed
Uncharacterized protein	2E-14	1	97	1.93E-03	0	260	5.07E-09	
		280	1198	9.15E-05	79	1008	2.60E-07	
Sodium/hydrogen exchanger-like protein	5E-66	41	1567	1.70E-12	67	1672	4.94E-07	
Hemolymph lipopolysaccharide-binding protein	1E-60	426	3528	9.02E-07	239	3238	9.86E-07	
		0	201	3.75E-09	0	91	2.75E-05	
Mitochondrial sodium/hydrogen exchanger NHA2	2E-101	48	1345	6.58E-12	67	1425	3.91E-05	
Apolipoprotein D	1E-143	3725	21240	2.31E-11	4407	21288	1.22E-03	
Uncharacterized protein	4E-40	344	1357	9.85E-03	147	886	1.50E-03	
Inositol monophosphatase 2	1E-167	1217	6646	1.39E-03	1478	9717	1.68E-03	
Serine protease gd	2E-201	134	537	3.79E-02	30	279	1.80E-03	
Trypsin zeta	9E-108	0	87	1.70E-03	4	111	2.93E-03	
Putative facilitated trehalose transporter	1E-154	1	161	7.73E-07	1	74	6.98E-03	
		0	123	6.51E-04	4	108	7.30E-03	
Integrin alpha-PS3	3E-232	76	481	1.67E-03	62	435	1.09E-02	
Uncharacterized protein	1E-112	270	1445	3.43E-02	147	1004	2.17E-02	Underexpressed
		80	595	2.10E-02	35	288	3.06E-02	
Phosphoenolpyruvate carboxykinase	0	7216	21124	1.39E-03	5813	18754	3.19E-02	
Putative membrane protein	0.0027	1662	14794	5.99E-06	1185	6617	3.32E-02	
Heterogeneous nuclear ribonucleoprotein Q	7E-26	0	147	6.93E-07	85	494	3.45E-02	
Uncharacterized protein	0	245	1522	8.02E-03	108	715	3.45E-02	
Protein takeout	2E-137	206	1613	3.87E-04	188	1051	3.68E-02	
		0	63	1.94E-02	1	54	4.58E-02	
Acetyl-coenzyme A synthetase	8E-289	742	106	6.66E-06	498	4	6.17E-15	
Glutamate receptor 1	3E-70	484	12	1.95E-11	285	7	2.74E-07	
Alkaline nuclease	2E-47	948	90	2.54E-07	449	31	8.18E-06	
28S ribosomal protein S10, mitochondrial	5E-78	275	0	4.82E-15	100	0	2.47E-05	
Predicted protein	2E-29	80	0	3.81E-03	142	0	4.12E-05	
Regucalcin (Fragment)	3E-80	530	69	1.01E-04	249	22	6.71E-04	
alpha-trehalose-phosphate synthase	0	21082	6116	4.91E-03	14663	2974	7.93E-04	
Fatty-acid amide hydrolase 2-like protein	5E-68	174	4	1.95E-05	159	5	9.55E-04	
Putative 4-coumarate--CoA ligase 3	3E-188	571	52	6.90E-03	305	29	9.55E-04	
F-box/LRR-repeat protein 7	3E-109	101	0	8.37E-06	62	0	2.65E-03	
Cytochrome b5-related protein	3E-199	5199	972	2.33E-05	4305	603	5.81E-03	
Protein croquemort	7E-139	1657	355	2.54E-03	1008	192	1.55E-02	
Angiopietin-1 receptor	5E-212	596	94	6.59E-03	524	96	1.55E-02	
Uncharacterized protein	3E-109	1934	468	2.87E-03	1476	377	2.12E-02	
Uncharacterized protein (Fragment)	6E-23	156	16	1.01E-02	12	254	3.45E-02	

Table 6-2. Differentially expressed genes common to both cerebral ganglia after the sting

Genes that are differentially expressed in both the brain and subesophageal ganglia, are either overexpressed in both, or underexpressed in both, except for one uncharacterized protein. These genes were analyzed for function against PfamA database using phmmer. Eval refers to the significance of the phmmer hit, SEG, SEG stung, brain, and brain stung are normalized FPKM values, with corresponding significance of differential expression (*p* val).

Presynaptic genes

Query	Reciprocal BLAST	SEG	SEG stung	p value	Brain	Brain Stung	p value
syntaxin	Syntaxin	907	1194	0.327	703	640	0.818
Soluble NSF attachment protein	Soluble NSF attachment protein	5512	5837	0.795	4202	3546	0.591
synaptobrevin	VAMP-7/Synaptobrevin	800	729	0.742	649	577	0.741
vesicle-fusing ATPase 1	Vesicle-fusing ATPase 1	1056	1199	0.622	1062	755	0.358
complexin	not found						
tomosyn 1	Syntaxin-binding protein/Tomosyn	2211	2203	0.988	5898	3227	0.073
synaptophysin	Synaptophysin-like protein 2	353	511	0.246	620	456	0.426
Synaptic vesicle glycoprotein 2C	Synaptic vesicle glycoprotein 2B	3846	3976	0.901	3400	5659	0.189
Synapsin	Synapsin	81	211	0.041	287	193	0.329
vacuolar ATPase B subunit	Vacuolar ATP synthase subunit B	15924	17627	0.636	9231	9819	0.837
Cysteine string protein	Cysteine string protein	435	549	0.456	510	544	0.858
neurexin 1	Neurexin-3-alpha	93	62	0.514	268	257	0.914
CAST1	CAST1	54	161	0.050	248	226	0.816
L-type Ca channel subunit beta-2	L-type Ca channel subunit beta-2	133	214	0.294	1408	719	0.048
synaptotagmin 1	Synaptotagmin	11838	13795	0.492	13786	10710	0.420
CDPKII	CDPKII	8	14	0.709	55	8	0.068
myosin v	Myosin-Va	253	338	0.430	510	415	0.556
bruchpilot	bruchpilot/ERC protein 2	54	161	0.050	248	226	0.816
bruchpilot	bruchpilot/ERC protein 2	478	2094	< E-07	981	1618	0.227

Postsynaptic genes

Query	Reciprocal BLAST	SEG	SEG stung	p value	Brain	Brain Stung	p value
neuroligin	Neuroligin-3	134	172	0.589	359	273	0.484
discs-large (PSD-95)	Disks large 1	7594	7697	0.952	6355	8289	0.395
NMDA receptor	N-methyl-D-aspartate receptor	2114	2181	0.900	2950	2389	0.498
serotonin receptor	5-hydroxytryptamine receptor 2B	25	15	0.660	45	45	1.000
Dopamine receptor	Dopamine receptor	0	9	0.430	2	9	0.417
octopamine receptor	Octopamine receptor	830	1105	0.310	1423	1114	0.477

Table 6-3. Synapse protein expression in stung and unstung cerebral ganglia

Synapse specific proteins were assessed for differential expression. The pre-synaptic proteins, synapsin, CAST1, and bruchpilot were significantly differentially expressed after the sting in the subesophageal ganglion, but not in the brain. L-type calcium channel was underexpressed after the sting in the brain, but not the subesophageal ganglion. Differentially expressed genes are in bold.

Chapter VII.

Conclusions and Future Directions

Value of Venome Assembly

Assembly of the venome described in this dissertation has provided a comprehensive dataset presenting a complete picture of *Ampulex compressa* venom components. This information will likely to lead to insights into mechanisms underlying venom-induced behavioral modification. Heretofore, the molecular composition of the venom was ill-defined. Previous work had assayed venom fractions for action on synaptic transmission and found only the small molecular weight fraction exerted inhibitory activity (Moore et al. 2006). Component analysis of this fraction characterized several abundant peptides and small molecules, including ampulexins 1, 2 and 3, and GABA, β -alanine, and taurine (Moore 2003), (Moore et al. 2006). Additional small molecule analysis found evidence for dopamine as well (Banks 2010).

Venomics of *A. compressa* presented here represents a fundamental advancement in the understanding of venom composition and action. This study yielded full length coding sequences for the vast majority of proteinaceous venom compounds as well as transcript quantification and semi-quantitative protein abundance. These data allow generation of hypotheses regarding venom action on cockroach cerebral ganglia that consider the entire spectrum of venom components. Venome assembly also demonstrates how next-generation shotgun sequencing and proteomics are used to obtain a comprehensive analysis of a venom of considerable complexity, novelty, and a small sample size, informing future research in this field of venomics in general. Since action of the venom appears to be non-cytotoxic, specific and reversible, its analysis yields a rich collection of novel, pharmacologically relevant compounds. Based on the putative

effects of some venom peptides on store-operated calcium entry, it is not inconceivable that we may discover additional modes of action that, once understood, could hold potential biological research or clinical diagnostic tools.

Small molecules found in the venom thus far - dopamine and GABA - appear to play important roles in the envenomation process. A complete chemical analysis of the venom would necessarily involve identifying other small signaling molecules. Such small molecules might include local anesthetics such as polyamine derivatives known to be present in wasp and spider venoms (Adams 2004), (Stromgaard et al. 2006). Indeed, injection of the local anesthetic procaine mimics some of the symptoms associated with hypokinesia (Gal et al. 2010). Therefore, a thorough evaluation at the small molecule complement is integral to further characterization of the venom.

Ampulexin Peptide Family

Elucidation of ampulexin 1 and ampulexin 2 actions on calcium mobilization may yet reveal the first store-operated calcium entry antagonism caused by a venom. Structural and functional analysis of these unique peptides hold potential as new tools to investigate mechanisms of store-operated calcium entry as well as insight into the mechanism of hypokinesia induction. Hypothetically, if the ampulexins indeed disrupt store operated calcium entry, they appear to do so indirectly, perhaps by preventing the channel from being gated by the stromal interaction molecule, also known as STIM. However, once these channels are opened the ampulexins cannot block calcium entry.

Although it seems natural to assume the target of ampulexin action is in the cockroach brain, ampulexins might modulate venom action, perhaps through allosteric or

active site binding of venom enzymes. Ampulexins may also act synergistically with other venom components, such as how melittin enhances necrotic action of phospholipase A2 in bee venom, despite melittin being a non-competitive inhibitor of phospholipase A2 (Ownby et al. 1997), (Saini et al. 1997). Due to structure similarities with cell penetrating peptides, ampulexins or other venom peptides may be able to intercalate into plasma membranes, effecting cell membrane dynamics, or penetrating completely into the cells, perhaps facilitating translocation of venom components into the cytoplasm of brain cells (Stalmans et al. 2013). The venom proteome contains polypeptides with homology to known endoribonucleases, suggesting that if these nucleases target cockroach transcripts, they must enter the cell by some mechanism. Disrupting protein expression through destroying transcripts would be an interesting component of hypokinesia.

Venom Tachykinins

Functional analysis of tachykinin venom peptides, presumably acting on tachykinin receptors in the cockroach central nervous system, could reveal a novel strategy for manipulation of host behavior and yield evolutionary insights into interspecific actions of neuropeptide signaling molecules. Although tachykinin receptor signal transduction has been reported to be sensitive to pertussis toxin, suggesting coupling to $G_{\alpha i}$, the cockroach tachykinin receptor is predicted to couple to $G_{\alpha q}$. (Li et al. 1991), (Poels et al. 2007). Future investigations of tachykinin ligand-receptor interactions naturally will involve uncovering whether the cockroach tachykinin receptor is coupled to $G_{\alpha q}$ or $G_{\alpha i/o}$ using chimeric G proteins and established luminescence techniques. These studies implicate tachykinin signaling in modulating control of locomotion in spontaneous walking and

escape in the subesophageal ganglion. Given that injection of venom tachykinin into the subesophageal ganglion increases escape threshold, tachykinin receptor neurons may be inhibitory, as previously reported in the cockroach antennal lobe (Jung et al. 2013), (Fusca et al. 2015). Activation of inhibitory tachykinin receptor neurons could increase threshold of locomotory circuits, thereby preventing descending permissive inputs from activating central pattern generators in thoracic ganglia.

One of the more critical and fascinating aspects of venom tachykinin functional analysis is that venom tachykinin is present and maintained in a protachykinin precursor form, as evidenced from mass spectroscopy experiments. Analysis of venom, venom gland extracts, and unprocessed tachykinin precursor fragments show reduced ability to activate the tachykinin receptor *in vitro*, despite the high affinity of venom tachykinins for the cockroach tachykinin receptor. Further study is required to ascertain of how the tachykinin precursor is processed to become bioactive. Tachykinins must be cleaved from the precursor at dibasic cleavage sites flanking each peptide in the precursor (Hook et al. 2008). Additionally, each tachykinin peptide must become C-terminally amidated by oxidation of the C-terminal glycine α -carbon (Eipper et al. 1992). Proteomics analyses reveals the presence of enzymes sufficient for these activities, in particular, a furin endoprotease whose canonical cleavage site matches the C-terminal cleavage site of each tachykinin (Duckert et al. 2004). Additional enzymes implicated in processing peptides from precursors are zinc containing endoproteases such as endothelin converting enzyme (Schmidt et al. 1994). Interestingly, neprilysin, a homolog of endothelin converting enzyme and a major venom component, degrades neuropeptide signaling molecules such

as Substance P and bradykinin (Skidgel et al. 2004). Elucidating the roles of proteases in processing or possible degradation of venom components, or protein and peptide targets in the cockroach brain, promise to reveal important roles in venom induced hypokinesia. In addition, a peptidylglycine alpha-amidating mono-oxygenase, known to amidate peptides with a C-terminal glycine, was found in the venom proteome (Eipper et al. 1993), (Mueller et al. 1993). Preliminary results indicate that venom sac extract, incubated at room temperature for one hour, at pH 7, in the presence of para-methylsulfonyl fluoride at 0.1 mM, can yield fully processed tachykinins. This supports the hypothesis that the venom tachykinin is maintained as fragments of protachykinin in the acidic venom sac, then when injected into the neutral pH cockroach brain, proteolysis and alpha-amidating activity increases, and mature, bioactive tachykinins appear. This could constitute a time-release mechanism, perhaps keeping inhibitory neurons active over several hours. It remains to be demonstrated whether the tachykinin processing in the venom sac is pH sensitive. Additionally, further validation of this hypothesis requires identification of processed venom tachykinins in stung cockroach cerebral ganglia. Attempts at venom peptide identification in stung cockroach subesophageal ganglia have only detected ampulexin 1, suggesting refinement in sample preparation is necessary to enrich for venom components, or tachykinins in subesophageal ganglia extracts.

Approaches to Functional Analysis of Venom Components

Description of the venom proteome has revealed many highly abundant novel proteins and peptides. Each of these remains to be functionally analyzed and

characterized in detail. It is not necessary however for a venom component to be abundant to be important, confounding functional analysis of overall venom action, as fleeting masses of specific venom proteins in the proteinaceous sea that is a cockroach brain will be difficult to follow.

While much work has been applied to determining what venom compounds directed to which cerebral location can be sufficient to induce a hypokinetic state in cockroaches, it is a different challenge altogether to determine which compounds in the venom are necessary to induce hypokinesia, especially in the long-term. One approach is to employ RNA silencing to knock down expression of venom transcripts. RNA silencing of wasp venom toxins is effective in a *Drosophila* parasitoid, and RNA interference of other non-venomous genes is effective in bees and ants (Choi et al. 2012), (Li-Byarlay et al. 2013), (Colinet et al. 2014). This approach suffers at least three caveats. First, venom components may be redundant, whereby different isoforms of enzymes or multiple agonists or antagonists may be present, such is the case with respect to multiple GABA_A receptor agonists GABA, β -alanine, and taurine (Moore et al. 2006). Knockdown of one redundant compound therefore would not compromise overall venom potency. Second, introduction of multiple double-stranded RNA species for simultaneous knockdown of several genes, would compete for the RNA-induced silencing complex (RISC) machinery. Even a carefully dosed injection of multiple species of double stranded RNA might lose efficiency overall. Therefore, only a few venom components may be knocked down at a time. Third, RNA interference may not be efficient in *A. compressa*, especially for highly abundant peptides such as ampulexin 1, where knockdown might not lower

concentration enough to yield an effect. Nonetheless, successful knockdown of a venom component sufficient to render a wasp incompetent to induce hypokinesia would be unmitigated evidence of that component's necessity in the envenomation pathology.

An alternative to *in vivo* analysis of venom toxins is to take advantage of the UAS-GAL4 system in *Drosophila* (Duffy 2002), (Busson et al. 2007). This system allows for expression of transgenes in a tissue specific way. Transgenic *Drosophila* can be generated whose genome would contain wasp toxin coding sequences under the control of UAS promoters. When crossed with *Drosophila* expressing a tissue specific GAL4, the offspring of these *Drosophila* would express a wasp toxin in the tissue specified by the GAL4 line. A UAS-ampulexin-2 line was generated for this purpose, and in preliminary geotaxis assays, expression of ampulexin 2 pan-neuronally via the *nrv*-GAL4 driver, causes hyperactivity, suggesting that ampulexin 2 can affect locomotion in *Drosophila* (Sun et al. 1999). It is also possible to express these UAS controlled toxins conditionally using additional transgenic tools, such as GAL80 and GeneSwitch (Pilauri et al. 2005), (Nicholson et al. 2008). While further refinements are necessary in terms of spatio-temporal control of toxin expression, this approach could become a valuable tool in venom functional analysis.

It is theoretically possible to generate a multiple UAS-toxin line with a combination of venom components to assess the combinatorial effect of these peptide toxins in *Drosophila*. Another possibility is to drive a polycistronic, multiple toxin-containing transcript by separating toxins with a T2A sequence (Daniels et al. 2014). The T2A sequence forces a ribosome to “skip” a peptide bond during translation, allowing

translation of multiple polypeptides from one transcript. In this scenario several toxins can be expressed under control of one promoter. However, T2A sequences leave an epitope of several amino acids on the protein to which they are fused to. This may be a problem for functional assays, since addition of several amino acids onto a protein could disrupt its function, especially for small peptides. Despite possible logistical difficulties, this approach could be used to “induce a sting”, by moving *Drosophila* to the permissive temperature (GAL80^{ts}) or feeding a compound such as RU486 (GeneSwitch). This approach, at least hypothetically, would be an elegant way to show effects of venom toxins on locomotion.

Perspectives on Venom Induced Hypokinesia

Development of a cerebral ganglia transcriptome has additional utility besides quantifying transcript expression for differential gene expression analysis. This database contains coding sequences for the majority of expressed proteins in the cerebral ganglia which can be used for mass spectrometry based proteomics of these tissues. Future elucidation of venom action must turn to observing the venom in its ultimate destination, the cockroach brain and subesophageal ganglion. These lines of analyses may be accomplished through whole brain proteomics of stung cockroaches. Although the venom would be a miniscule amount of total protein in this sample, modern proteomics methods are very sensitive, and with proper sample preparation and fractionation, even a very small amount of a protein, even into the attomole range, can be detected. Besides the possible processing of protachykinin in the cerebral ganglia after the sting, it is not unreasonable to hypothesize that other venom proteins would undergo modifications after

the sting, and these modifications may very well be critical to the venom action, especially over the long term. Future mass spectrometry analysis of cockroach cerebral ganglia may detect these changes and be used to follow the progress of venom processing over the course of hypokinesia. Since hypokinesia is a reversible phenomenon, it is interesting to consider the fate of the venom components over the days-long effect of the venom. It is possible that the venom disrupts cellular function in a persistent, but reversible way such that even after it has been neutralized or eliminated from cockroach cerebral ganglia, the tissue remains affected and the cockroach remains hypokinesic. This is similar to the action of botulinum toxin in that, although its half-life is about four hours in rat serum, the effects can last weeks to months (Ravichandran et al. 2006), (Nigam et al. 2010), (Wheeler et al. 2013). The venom might also persist only as long as effects are seen, whereby the time course of recovery would correspond to the time course of elimination of the venom. Therefore, monitoring the half-life of venom components would be a way to monitor its action. In this scenario, it would be assumed that the venom is acting consistently and persistently to control escape. Lastly, the venom may persist well after its effects are no longer noticeable, where the cockroach may habituate, or otherwise adapt to its presence. *A. compressa* venom contains enzymes, such as hyaluronidase and matrix metallo-protease that are known to act on extracellular matrix proteins and glycans (Kochlamazashvili et al. 2010), (Kurshan et al. 2014), (Levy et al. 2014), (Bikbaev et al. 2015). These enzymes are ubiquitous in venoms and are thought to be spreading factors which allow the venom to penetrate deeper than the site of injection (Girish et al. 2004). This could very well be the case for *A. compressa* venom. However,

is worth considering what effect disrupting and/or loosening cell-cell interactions in the central nervous system would be. Hypothetically, these enzymes could disrupt structural integrity of synapses, causing the hypokinetic state until synapse connection is re-configured. This might take days to repair and could be integral to the onset and duration of venom induced hypokinesia.

With the discovery that an important synapse protein, bruchpilot, is differentially expressed after the sting, and that bruchpilot is critical to circadian rhythm of sleep-wake cycles in *Drosophila*, studies are now under way to explore the hypothesis that venom induced hypokinesia induces a prolonged narcolepsy, or sleep like state (Maayan Kaiser, Fredric Libersat, personal communication). In fact, quiescent cockroaches do not spontaneously walk and have an increased escape response threshold until roused (Fredric Libersat, personal communication, and personal observation). This hypothesis is exciting, since it correlates observed effects of venom on a biochemical level with the fascinating venom induced-behavior.

Understanding venom composition is integral to understanding venom action. The venome of *A. compressa* venom is a rich, neuroactive biochemical admixture whose neuropharmacology is potent, yet reversible, suppressing escape without paralysis. Elucidation of the venome reveals more questions than it answers, and a significant amount of investigation remains to unravel the mechanism of venom action. Each protein or peptide described herein plays some role in venom action and each warrants further investigation. Hypokinesia is a syndrome, likely caused by the concerted action of many

venom components, orchestrated temporally to usurp control of the cockroach's motility, for *A. compressa*'s maternal, yet macabre, motives.

References Cited

- Adams, M. E. (2004). "Agatoxins: ion channel specific toxins from the American funnel web spider, *Agelenopsis aperta*." Toxicon **43**(5): 509-525.
- Banks, C. N. (2010). The Roles of Biogenic Amines and Dopamine Receptors in Envenomation by the Parasitoid Wasp *Ampulex compressa*. Ph.D., University of California.
- Bikbaev, A., et al. (2015). "Brain extracellular matrix retains connectivity in neuronal networks." Sci Rep **5**: 14527.
- Busson, D. and A. M. Pret (2007). "GAL4/UAS targeted gene expression for studying *Drosophila* Hedgehog signaling." Methods Mol Biol **397**: 161-201.
- Choi, M. Y., et al. (2012). "Phenotypic impacts of PBAN RNA interference in an ant, *Solenopsis invicta*, and a moth, *Helicoverpa zea*." J Insect Physiol **58**(8): 1159-1165.
- Colinet, D., et al. (2014). "Development of RNAi in a *Drosophila* endoparasitoid wasp and demonstration of its efficiency in impairing venom protein production." J Insect Physiol **63**: 56-61.
- Daniels, R. W., et al. (2014). "Expression of multiple transgenes from a single construct using viral 2A peptides in *Drosophila*." PLoS One **9**(6): e100637.
- Duckert, P., et al. (2004). "Prediction of proprotein convertase cleavage sites." Protein Eng Des Sel **17**(1): 107-112.
- Duffy, J. B. (2002). "GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife." Genesis **34**(1-2): 1-15.
- Eipper, B. A., et al. (1993). "Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains." Protein Sci **2**(4): 489-497.
- Eipper, B. A., et al. (1992). "The biosynthesis of neuropeptides: peptide alpha-amidation." Annu Rev Neurosci **15**: 57-85.

Fusca, D., et al. (2015). "Colocalization of allatotropin and tachykinin-related peptides with classical transmitters in physiologically distinct subtypes of olfactory local interneurons in the cockroach (*Periplaneta americana*)." J Comp Neurol **523**(10): 1569-1586.

Gal, R. and F. Libersat (2010). "A wasp manipulates neuronal activity in the sub-oesophageal ganglion to decrease the drive for walking in its cockroach prey." PLoS ONE **5**(4): e10019.

Girish, K. S., et al. (2004). "Isolation and characterization of hyaluronidase a "spreading factor" from Indian cobra (*Naja naja*) venom." Biochimie **86**(3): 193-202.

Hook, V., et al. (2008). "Proteases for processing proneuropeptides into peptide neurotransmitters and hormones." Annu Rev Pharmacol Toxicol **48**: 393-423.

Jung, J. W., et al. (2013). "Neuromodulation of olfactory sensitivity in the peripheral olfactory organs of the American cockroach, *Periplaneta americana*." PLoS One **8**(11): e81361.

Kochlamazashvili, G., et al. (2010). "The extracellular matrix molecule hyaluronic acid regulates hippocampal synaptic plasticity by modulating postsynaptic L-type Ca(2+) channels." Neuron **67**(1): 116-128.

Kurshan, P. T., et al. (2014). "Regulation of synaptic extracellular matrix composition is critical for proper synapse morphology." J Neurosci **34**(38): 12678-12689.

Levy, A. D., et al. (2014). "Extracellular matrix control of dendritic spine and synapse structure and plasticity in adulthood." Front Neuroanat **8**: 116.

Li-Byarlay, H., et al. (2013). "RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee." Proc Natl Acad Sci U S A **110**(31): 12750-12755.

Li, X. J., et al. (1991). "Cloning, heterologous expression and developmental regulation of a *Drosophila* receptor for tachykinin-like peptides." EMBO J **10**(11): 3221-3229.

Moore, E. L. (2003). A Biochemical and Molecular Analysis of Venom with Distinct Physiological Actions from Two Arthropod Sources: The Parasitoid Jewel Wasp,

Amplex compressa, of the Insect Order Hymenoptera and the Obligate Entomophagous Assassin Bug Platyeris biguttata, of the Insect Order Hemiptera. Ph.D., UC Riverside.

Moore, E. L., et al. (2006). "Parasitoid wasp sting: a cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host." Journal of neurobiology **66**(8): 811-820.

Mueller, G. P., et al. (1993). "Peptide alpha-amidation and peptidylglycine alpha-hydroxylating monooxygenase: control by disulfiram." Mol Pharmacol **44**(5): 972-980.

Nicholson, L., et al. (2008). "Spatial and temporal control of gene expression in *Drosophila* using the inducible GeneSwitch GAL4 system. I. Screen for larval nervous system drivers." Genetics **178**(1): 215-234.

Nigam, P. K. and A. Nigam (2010). "Botulinum toxin." Indian J Dermatol **55**(1): 8-14.

Ownby, C. L., et al. (1997). "Melittin and phospholipase A2 from bee (*Apis mellifera*) venom cause necrosis of murine skeletal muscle in vivo." Toxicon **35**(1): 67-80.

Pilauri, V., et al. (2005). "Gal80 dimerization and the yeast GAL gene switch." Genetics **169**(4): 1903-1914.

Poels, J., et al. (2007). "Functional comparison of two evolutionary conserved insect neurokinin-like receptors." Peptides **28**(1): 103-108.

Ravichandran, E., et al. (2006). "An initial assessment of the systemic pharmacokinetics of botulinum toxin." J Pharmacol Exp Ther **318**(3): 1343-1351.

Saini, S. S., et al. (1997). "Melittin binds to secretory phospholipase A2 and inhibits its enzymatic activity." Biochem Biophys Res Commun **238**(2): 436-442.

Schmidt, M., et al. (1994). "Molecular characterization of human and bovine endothelin converting enzyme (ECE-1)." FEBS Lett **356**(2-3): 238-243.

Skidgel, R. A. and E. G. Erdos (2004). "Angiotensin converting enzyme (ACE) and neprilysin hydrolyze neuropeptides: a brief history, the beginning and follow-ups to early studies." Peptides **25**(3): 521-525.

Stalmans, S., et al. (2013). "Chemical-functional diversity in cell-penetrating peptides." PLoS One **8**(8): e71752.

Stromgaard, K., et al. (2006). "Protolytic properties of polyamine wasp toxin analogues studied by ¹³C NMR spectroscopy." Magn Reson Chem **44**(11): 1013-1022.

Sun, B., et al. (1999). "Dynamic visualization of nervous system in live *Drosophila*." Proc Natl Acad Sci U S A **96**(18): 10438-10443.

Wheeler, A. and H. S. Smith (2013). "Botulinum toxins: mechanisms of action, antinociception and clinical applications." Toxicology **306**: 124-146.