Mass Spectral Analysis of Crustacean Signaling Peptides

Using a Multi-dimensional Strategy

By

Nicole Adelita Woodards

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of the final oral examination: 07/25/2013

The dissertation is approved by the following members of the Final Oral Committee:
Lingjun Li, Professor, Analytical Chemistry
Antony Stretton, Professor, Neurobiology
Lloyd Smith, Professor, Analytical Chemistry
John Wright, Professor, Analytical Chemistry
Richard Peterson, Professor, Pharmaceutical Science

For My Son

Do as I say and as I do,

And always be true to you.

Remember my lil boy,

This is the key to obtaining true joy.

Shine your light amidst the gloom and pain,

For you are the embodiment of your name.

Huey, I prophesized you my expected unexpected,

My center, my focus, my joy, my eclectic.

My light, my hope, my love, my dare,

My muse, in my accomplishments we both shall share.

Acknowledgements

To my advisor Dr. Lingjun Li, I truly appreciate the freedom you allow in the pursuit of knowledge. Lord knows, I have benefited greatly from your understanding in the challenges surrounding method development and application. Thank you for allowing me six years of research, in which I have matured both personally and professionally.

To Dr. Antony Stretton, Dr. Lloyd Smith, Dr. John Wright, and Dr. Richard Peterson, thank you all for agreeing to serve on my committee and the individual encouragement and acknowledgements you have given me throughout the years.

To my lab mates, both past and present, thank you for either your professional and/or personal contributions you have made to this work. I also had the privilege of mentoring three students, for which I am grateful. A special thanks to Gary Girdaukas and Tom O'Connor, you guys are awesome. Thank you also to Cameron Scarlet and Molly Hahn in the AIC.

Personally, I want to thank my friends and family that have encouraged me to continue my academic pursuit. Thank you Ms. Julianna Matz, you told me to not stop and continue till I was done, and that has resulted in 25 years of continuous education for which I am several thousands of dollars in debt… but I wouldn't have it any other way. Ruth Bacon, you know I did all this because you promised me a KitchenAid. I have to also thank all the support I have had as a single mother, and there was a lot, and without that I would not be able to be here.

The funding I received from AOF and NIDDK Supplement is also greatly appreciated because it allowed me to focus on my research. There was more funding after and I am grateful.

My brain is old, and foggy, so if you feel jipped I didn't thank you personally, let's meet. You can buy me a cup of coffee and we can quantify your over-emotional feelings. Thank you.

Table of Contents

Mass Spectral Analysis of Crustacean Signaling Peptides

Using a Multi-dimensional Strategy

Nicole Adelita Woodards

Under the supervision of Professor Lingjun Li

At the University of Wisconsin-Madison

Abstract

By definition, neuropeptides are a class of signaling molecules either secreted by or that have an effect on neurons. They are present at minute concentrations in complex biological matrices. Their expression levels vary both temporally and spatially as they regulate several biological functions. Our limited knowledge of their regulatory processes has resulted in the creation or understanding of several therapeutic treatments with more advancements hinging on future neuropeptidomic breakthroughs.

A multi-dimensional strategy has been implemented for the analysis of neuropeptides in the crustacean model system. Due to their dynamic nature, biological mass spectrometry is the instrumental platform used for neuropeptide analysis. It is coupled with analytical methodologies primarily focused on improving neuropeptide detection, identification, and quantitation.

Simple protocols for rapid, robust, and reproducible analysis have been developed by modifying previous techniques for novel application on newer technologies. Several protocols were developed by acknowledging both the benefits and limitations of the techniques implemented. While this work utilizes both soft ionization techniques, more emphasis is given to the development of MALDI analysis.

Relative quantitative analysis by 4-plex N,N-dimethyl leucine (DiLeu) labels was applied to study the neuropeptide expression level changes in relation to feeding. A neuropeptide database was compiled for automated analysis of large datasets using popular protein identification and quantitation software.

Improving mass spectrometry *de novo* sequencing was achieved by modifying a simple acid-driven deamidation protocol to determine the presence and location of amides. Also, the formation of multiply charged peptides by MALDI ionization was explored to improve fragmentation spectra.

Finally, to further analysis by MALDI ionization, proof-of-principle projects were initiated. The first was established to improve detection of hemolymph samples on newer MALDI instruments to determine expression level changes of circulating neuropeptides in relation to the circadian rhythm cycle. The second was the application of isotopic DiLeu (iDiLeu) labels to generate a five point standard curve for absolute quantitation. Primary interests of these labels are in their application for imaging experiments. In conclusion, all projects were developed by implementing a multi-dimensional strategy to strengthen the mass spectrometric platform for advancing neuropeptidomic research.

Chapter 1: Introduction and Overall Summary

1.1 Abstract

A multi-dimensional approach was applied to adequately analyze neuropeptides in the crustacean model system. The dimensions incorporated into this work are (1) considering the dynamics of neuropeptide expression; (2) complementing and building upon previous knowledge; (3) improving analytical methodologies with particular emphasis on detection, identification, and quantitation; (4) improving matrix assisted laser desorption/ionization (MALDI) mass spectrometry analysis; (5) implementing novel technology and techniques to advance current research; (6) creating or modifying simple sample analysis protocols for rapid, robust, and reproducible application; (7) implementing a biological mass spectrometric platform for neuropeptide analysis; and (8) acknowledging both the advantages and disadvantages of these techniques for optimal application and improvement.

All dimensions, save MALDI analysis, were incorporated into the six projects discussed in this work. The projects are (1) a neuropeptide database compilation; (2) relative quantitative analysis of neuropeptides in relation to feeding by 4-plex N,N dimethyl leucine (DiLeu) labeling; (3) the analysis of multiply charged ions in MALDI; (4) deamidation of peptides for rapid screening of amides within a sequence or sample; (5) isotopic DiLeu (iDiLeu) for absolute quantitation by a five point standard curve generated in MALDI; and (6) improving detection of circulating neuropeptides in hemolymph samples for circadian rhythm analysis.

1.2 Introduction

Neuropeptides are an important class of chemical messengers that affect numerous biological processes such as mating, feeding, breathing, pain sensing, learning, and memory [1]. They are present at trace-level concentrations in complex biological systems and their expression

varies both spatially and temporally for regulation of biological processes [2, 3]. Due to their complexity, understanding their exact function in the nervous system has remained elusive [4]. Biological mass spectrometry is capable of analyzing complex biological mixtures with high speed, sensitivity, and specificity making it an ideal tool for peptide analysis [5, 6].

Biological mass spectrometry uses soft ionization techniques to ionize peptides then separates them by their mass-to-charge (m/z) ratio for detection [5]. The amino acid sequence, size, and length all affect the ionization efficiency of the peptide affecting its ability to be detected and subsequently identified by tandem mass spectrometry fragmentation [5, 7-10]. Analytical methodologies are employed to improve detection, identification, and quantitation of neuropeptides. Primary application of these analytical methods is to improve MALDI based analysis, although electrospray ionization (ESI) is also used in this work.

Neuropeptides and their regulatory pathways are conserved across species [11]. The relatively simple crustacean neural network has been extensively studied and its physiology is well defined [11-13]. Therefore crustaceans provide an excellent biological model system creating the unique opportunity of identifying novel peptides with the ability to verify their biological functions at the circuitry level [2, 14]. This information continues to bridge the gap between previously established knowledge and current discoveries [14]. It also provides a significant platform for exploratory method development that advances neuropeptidomic research when coupled to state-of-the-art technologies [15].

1.3 Chapter Summary

Overall, this work contains six major sections; (1) an introductory section; (2) background information; (3) automated neuropeptide identification and quantitation; (4) improving tandem mass spectrometry de novo sequencing; (5) developing MALDI based

analysis; and finally (6) a conclusions and supplemental information section. A brief background introduction and review for the accumulated work is included in Chapter 2. It describes the benefits of using the crustacean model system and discusses the need for automated identification and quantitation software for neuropeptidomic analysis. It also provides a very brief overview of neuropeptide biosynthesis and identifies the crustacean neuropeptide families listed in our in-house crustacean database.

Section 3 is the compilation and application of a neuropeptide database to automate neuropeptide identification and quantitation. The first project discussed, in Chapter 3, is the compilation of a neuropeptide database. Our previous in-house neuropeptide database was incomplete; the 245 entries were missing key information (i.e. sequence, name, family, and/or mass) and also the list lacked several of our recently discovered peptides. Utilizing several recent publications the neuropeptide database was updated but still several entries were incomplete. Broadening the literature search to incorporate earlier research completed on the crustacean model, we were able to complete and feasibly expand the neuropeptide database to 692 entries. This database was then incorporated into a larger database containing the peptide sequences of homologous species and uploaded onto three types of protein identification and quantitation software for automated peptide analysis.

The second project discussed, in Chapter 4, benefited immensely by the automated peptide analysis achieved from the database compilation. The project was developed to analyze the expression of neuropeptides in crustaceans at several stages of satiation utilizing the recently developed 4-plex N,N-dimethyl leucine (DiLeu) quantitation labels. A labeling matrix was created to test the robustness of the tags by utilizing several technical, experimental, and biological replicates. The large dataset resulting from the labeling matrix was extensively, and

initially incorrectly, manually analyzed. When newer software versions were released that allowed the user to customize the reporter ions in the quantitation labeling scheme, a month's worth of manual analysis was easily performed in minutes. The top scoring peptides from three types of commercially available software (Mascot, PEAKS, and ProteoIQ) were determined and their quantitative results compared to their raw data and a previous study employing isotopic formaldehyde labeling for relative quantitative analysis.

Section 4 is the improvement of tandem mass spectrometry *de novo* sequencing. Chapter 5 discusses the production and tandem mass spectrometry analysis of multiply charged peptide ions generated by MALDI ionization. Acknowledging the disadvantage of singly charged ion fragmentation patterns generated by the MALDI process, several simple plating and matrix manipulation techniques were implemented with the goal of producing abundant multiply charged peptide ions by MALDI ionization [16, 17]. These processes were analyzed on several newer MALDI instruments and utilized techniques from MALDI's infancy [18-21]. Results are presented that demonstrate the feasibility of this approach.

In Chapter 6, a simple acid-driven deamidation protocol was established to rapidly screen for the presence of amides in a sequence. This protocol is modified from a previously established deamidation protocol and is coupled to MALDI analysis revealing the presence and number of amides within the sample by each 1Da mass shift [22]. When coupled to liquid chromatography (LC) ESI analysis, the results indicate whether an amide was located in the main sequence of the peptide or on the C-terminus by an increase in retention time resulting from the conversion of the sequence amino acids to their deamidated acid counterparts.

Section 5 is a proof-of-principle method development section containing two projects for the improvement of MALDI based analysis. Chapter 7 discusses the use of iDiLeu labels for

generation of a five point standard curve to calculate absolute quantitation of a peptide with MALDI based analysis [23]. While this technique may be used for absolute, and possibly relative, quantification of any MALDI analysis, particular interest of its application to MALDI imaging experiments and our hemolymph analysis studies are emphasized. Chapter 8 discusses modifications of the hemolymph protocol for analysis on the newer MALDI TOF/TOF instruments. Simple sample preparation modifications are investigated to improve identification of circulating neuropeptides for circadian rhythm analysis.

In the final section, conclusions and future directions for this work are highlighted and discussed. This section also contains the appendices for this work. Appendix A contains a short description of the instruments used, the animal care, dissection, and sample collection and preparation protocols. Appendix B contains the 692 entry neuropeptide database. Appendix C contains the synthesis and labeling for DiLeu and iDiLeu. Appendix D contains the introduction packets I have created during my time in the Li laboratory to aid in information sharing and ensure proper training. Appendix E highlights my contributions to a recent collaboration that resulted in a publication [24].

1.4 Conclusion

This work embodies six years of research focused on crustacean neuropeptidomic analysis. It utilizes newer technology to build upon and strengthen previous research to provide a powerful and robust mass spectrometry platform for future analysis.

1.5 Works Cited

1. Klavdieva, M.M., *The history of neuropeptides I.* Front Neuroendocrinol, 1995. **16**(4): p. 293-321.
2. Christie, A.E., E.A. Stemmler, and P.S. Dickinson, *Crustacean neuropeptides.* Cellular and molecular life sciences : CMLS, 2010. **67**(24): p. 4135-69.
3. Strand, F.L., *Neuropeptides : regulators of physiological processes*. Cellular and molecular neuroscience. 1999, Cambridge, Mass.: MIT Press. 658 p.
4. Pozo, D. and M. Delgado, *The many faces of VIP in neuroimmunology: a cytokine rather a neuropeptide?* FASEB J, 2004. **18**(12): p. 1325-34.
5. Nilsson, C.L., et al., *Mass spectrometry of peptides in neuroscience.* Peptides, 1998. **19**(4): p. 781-9.
6. Jakubowski, J.A. and J.V. Sweedler, *Sequencing and mass profiling highly modified conotoxins using global reduction/alkylation followed by mass spectrometry.* Anal Chem, 2004. **76**(22): p. 6541-7.
7. Morrison, J., *Studies of Ionization Efficiency. I. The Determination of Molecular Appearance Potentials Using the Mass Spectrometer.* The Journal of Chemical Physics, 1951. **19**: p. 1305.
8. Morrison, J. and A. Nicholson, *Studies of ionization efficiency. Part II. The ionization potentials of some organic molecules.* The Journal of Chemical Physics, 1952. **20**: p. 1021.
9. Morrison, J., *Studies of ionization efficiency. Part III. The detection and interpretation of fine structure.* The Journal of Chemical Physics, 1953. **21**(10): p. 1767.

10. Morrison, J., *Studies of Ionization Efficiency. Part IV. The Electron Spectra of Some Simple Molecules.* The Journal of Chemical Physics, 1954. **22**: p. 1219.

11. Kwok, R., S.M. Chan, and S.S. Tobe, *Chapter 34 - Crustacean Bioactive Peptides*, in *Handbook of Biologically Active Peptides*, J.K. Abba, Editor. 2006, Academic Press: Burlington. p. 221-228.

12. Nusbaum, M.P. and M.P. Beenhakker, *A small-systems approach to motor pattern generation.* Nature, 2002. **417**(6886): p. 343-50.

13. Marder, E. and D. Bucher, *Central pattern generators and the control of rhythmic movements.* Curr Biol, 2001. **11**(23): p. R986-96.

14. Keller, R., *Crustacean neuropeptides: structures, functions and comparative aspects.* Experientia, 1992. **48**(5): p. 439-48.

15. McAlister, G.C., et al., *A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-orbitrap mass spectrometer.* J Proteome Res, 2008. **7**(8): p. 3127-36.

16. He, M., J.F. Emory, and S.A. McLuckey, *Reagent anions for charge inversion of polypeptide/protein cations in the gas phase.* Anal Chem, 2005. **77**(10): p. 3173-82.

17. He, M. and S.A. McLuckey, *Two ion/ion charge inversion steps to form a doubly protonated peptide from a singly protonated peptide in the gas phase.* J Am Chem Soc, 2003. **125**(26): p. 7756-7.

18. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Anal Chem, 1988. **60**(20): p. 2299-301.

19. Vorm, O., P. Roepstorff, and M. Mann, *Improved Resolution and Very High Sensitivity in MALDI TOF of Matrix Surfaces Made by Fast Evaporation.* Analytical Chemistry, 1994. **66**(19): p. 3281-3287.

20. Strupat, K., J. Kampmeier, and V. Horneffer, *Investigations of 2,5-DHB and succinic acid as matrices for UV and IR MALDI. Part II: Crystallographic and mass spectrometric analysis.* International Journal of Mass Spectrometry and Ion Processes, 1997. **169–170**(0): p. 43-50.

21. Kampmeier, J., et al., *Investigations of 2,5-DHB and succinic acid as matrices for IR and UV MALDI. Part: I UV and IR laser ablation in the MALDI process.* International Journal of Mass Spectrometry and Ion Processes, 1997. **169–170**(0): p. 31-41.

22. Miller, H.K. and H. Waelsch, *Utilization of glutamine and asparagine and their peptides by micro-organisms.* Nature, 1952. **169**(4288): p. 30-1.

23. Xiang, F., *Developing novel labeling tools for neuropeptide quantitation and identification by mass spectrometry*. 2011, The University of Wisconsin - Madison: Ann Arbor. p. 251.

24. Sturm, R.M., et al., *Mass spectrometric evaluation of neuropeptidomic profiles upon heat stabilization treatment of neuroendocrine tissues in crustaceans.* J Proteome Res, 2013. **12**(2): p. 743-52.

Chapter 2: An Introduction to the Discovery and Sequence Identification of Crustacean Neuropeptides in Biological Mass Spectrometry

2.1 Abstract

This research focuses on the improvement of crustacean neuropeptidomic research by incorporating analytical techniques with biological mass spectrometry analysis. The following chapter is presented as a brief overview of the concepts developed by this research. It analyzes the importance of the crustacean animal model in scientific research, presents a brief review of neuropeptides, and reviews the biological mass spectrometric techniques employed in neuropeptide research with primary emphasis on automated neuropeptide identification and database formation.

2.2 Introduction

Neuropeptidomics is a term coined by researchers to describe the comprehensive analysis of neuropeptides under specific physiological conditions [1]. Neuropeptides are a type of signaling molecule involved in regulating behaviors and functions associated with the nervous system, such as feeding, sleeping, learning, and movement [1-3]. While there are several methods for neuropeptide analysis, this work utilizes a biological mass spectrometry based analysis coupled with analytical methodologies for analysis of neuropeptides in the crustacean model system.

2.3 The Crustacean Model System

Despite the world of knowledge in the crustacean field, it is a common occurrence to justify the use of a crustacean model system to non-crab users, especially in a community dominated with rodent research. Crustaceans are not as readily accepted by the scientific

community as a viable model system yet, despite the increasing focus on the regulations, or lack thereof, placed on their research (http://www.aaalac.org) [4, 5]. The following section is dedicated to explaining the use of crustaceans as model systems with emphasis on their housing and other general information.

The crustacean model system has a long, rich history in research. The first scientific studies using crustacean model systems were performed over a century ago and range from anatomical to physiological and behavioral [6-9]. Crustacean research led to the elucidation of the sequence of the first invertebrate neuropeptide; the red pigment concentrating hormone (RPCH) [6, 9]. The entire process, from identification to sequencing of the neuropeptide, took 44 years spanning from 1928 to 1972 [6, 9]. Corresponding to the sequencing of RPCH were numerous physiological studies on the neurons involved with motor function in the crab.

Physiologically complexity lies within their simple neural network. Numerous signaling molecules regulate the functional output of the simple structure of the stomatogastric nervous system (STNS) neural circuits [10, 11]. These signaling molecules and their pathways that regulate complex behaviors are conserved across species [12, 13]. Studying the simpler neural circuitry of the crustacean nervous system instead of the complex mammalian neural network facilitates detection of these molecules and pathways [5, 14]. Studies on the crustacean neural system have generated understanding in the human neural circuit operation at a structural and regulatory level [6, 9, 14, 15].

Crustacean models are also ideal model systems for development of analytical techniques. The technique's results are validated by the model system's long history of study while simultaneously expanding upon previous knowledge [16]. Simply stated, the crustacean

nervous system has been extensively studied and its physiology is well defined, which helps validate results of exploratory neuropeptide research [16].

Crustacean Species Used in this Study

This research explores the neuropeptide content of two species of crustacean. The first is *Cancer borealis*, more commonly known by lobstermen as the Jonah crab. The common name was formulated because lobsters and crabs do not cohabitate and initially these crabs were not used for food, therefore lobstermen would release the seemingly useless animal unharmed like the biblical Jonah from the clutches of the whale (http://www.seafoodsource.com/seafoodhandbook.aspx?id=10737418967). The second species is *Callinetus sapidus* or the blue crab, thus named for the color of its shell.

Open Circulatory System

Crabs have an open circulatory system and their circulating fluid is called hemolymph [7]. The crab has a heart near the pericardial organs and an auxiliary heart near the anterior end of the dorsal median artery to help circulation in the system [17]. The hemolymph contains hemocyanin responsible for the transport of oxygen in the system [18]. Circulating neuropeptides are transported in the hemolymph to act on various distant neurons and regulate several physiological processes [18].

Stomatogastric Nervous System

The crustacean stomatogastric nervous system (STNS) is responsible for generating movement of the foregut and is composed of four connected ganglia [13, 14]. There are two commissural ganglion (CoG) containing approximately 500 neurons each, an esophageal ganglion (OG) with approximately 18 neurons, and a stomatogastric ganglion (STG) with about

30 motor neurons [13]. The actual number of neurons within each ganglion varies between species. The STG, which generates movement in the gastric mill and pylorus, receives hormonal signals from many sources like the pericardial organ (PO) and is controlled by many modulatory neurons from the CoG and OG [13]. The STNS of the crustacean is well-defined and characterized by physiological studies [13, 14, 19-21]. The STNS is displayed in Figure 2.1.

Identification of Crabs: Gender, Size, and Anatomy

Gender of a crab is determined by the shape of the tail. Males have a "lighthouse" shaped tail while females are more rounded. It may be difficult to differentiate between an unmated female and male crab, but mated female crabs are easy to identify. In blue crabs, an additional indicator of sex is the color of the claws. Females have red "painted" claws while males are all blue. The size of the crab is determined by measuring across the carapace of the crab. General crab anatomy is given in a PowerPoint reference entitled *Crustacean Anatomy and Stuff: A Quick Guide to the Anatomy of our Decapod Friends* (Appendix D).

Higher-Order Animal Model and Regulation

Crabs, lobsters, and crayfish are members of the phylum Arthopoda that are classified as a higher-order animal and are therefore perceived to feel pain or suffering and also capable of displaying complex behaviors [4, 5, 22]. Classified as such, they fall under animal regulatory guidelines as illustrated by Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC), the *Guide for the Care and Use of Laboratory Animals* (GUIDE) and the Animal Welfare Act of 1999 [23]. However, realistically, as invertebrates their regulation is minimal [5].

Invertebrate Regulation in Research

Typically invertebrates are seen as a viable replacement for vertebrate research in the reduction, replacement, refinement animal research model [5, 24]. Their research has expanded and supplemented knowledge of vertebrate systems [5]. However, sentiments about limited regulation regarding invertebrate research are changing. Recently, the European Parliament included cephalopods in its animal use directive 2010/63/EU, with decapods included in the initial draft but not the final directive [25]. Although invertebrate research regulation is not present in the United States, many oversight organizations, such as AAALAC, Institutional Animal Care and Use Committees (IACUC), and the Institute for Laboratory Animal Research (ILAR) are emphasizing the importance of proper regulation in higher order invertebrates in research (http://www.aaalac.org/accreditation/faq_landing.cfm#A2) [4, 25].

University of Wisconsin-Madison Regulations

The Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) is an oversight organization that provides accreditation to facilities recognizing their commitment to proper treatment of animals in research. The University of Wisconsin-Madison has AAALAC accreditation and such accreditation validates that our animal research promotes ethical treatment of animals in research which complements the regulatory mandate of animal research by many funding agencies. In October of 2011, the University underwent review for AAALAC re-accreditation. Given our crustacean facility location in the School of Pharmacy animal facility, that mainly houses vertebrates, and the initiation of invertebrates in European Parliament directives, we revamped our policies and updated the Standard Operating Procedures (SOPs) for our facility. The SOPs are all included in Appendix A.

Crustacean User Animal Training

Given our proximity to vertebrates in the School of Pharmacy animal facility, all crustacean users in our lab undergo online animal training as outlined in the *Li Lab Introduction Packet* (Appendix D). They are presented with the SOPs of the laboratory for general reference (Appendix A) and a PowerPoint reference entitled *Crustacean Anatomy and Stuff: A Quick Guide to the Anatomy of our Decapod Friends* (Appendix D). This coupled with personal training in tank cleaning and maintenance, crab health monitoring, crab shipping and acclimation, dissections and other experimental procedures are safeguards in ensuring that we maintain the safety of the animals and the researcher while furthering our research scope and insuring our research integrity.

2.4 Crustacean Handling and Housing

Aquarium and Facility Conditions

Crabs are housed in cold-saltwater aquariums ranging in 10-15°C depending on the species. Specific aquarium conditions for each species are documented in the SOPs in Appendix A. To maintain adequate housing conditions, the aquariums are cleaned every 8-12 days and water quality checks are performed. Refer to Appendix A for the SOP regarding this procedure.

The animal facility is also kept on an artificial light cycle. A 12 hour light and dark cycle is maintained by high beam white fluorescent and low beam red lights respectively. Since crustaceans are nocturnal, the dark cycle is maintained from 10 am to 10pm daily so that behavioral and feeding studies may be conducted during normal business hours [26].

Obtaining Crabs and Acclimation

Obtaining the crabs used in this study is done in one of two ways. Either they are ordered online and shipped to our facility or we hand pick the healthiest subjects from local groceries.

Shipped crabs were obtained from The Fresh Lobster Company (Gloucester, MA), The Crab Place (Crisfield, MD), James Hook & Co (Boston, MA), or Ocean Resources, Inc. (Sedgwick, ME) depending on availability and survival rates of the shipped crabs. They are normally shipped overnight and packaged with ice packs, artificial seaweed and/or paper towels. Once the subjects have been acquired, they are kept in cold saltwater aquariums until they are studied.

We define a good shipment of crabs as one that has greater than an 80% survival rate from shipment to full acclimation. Since crabs are packed on ice during shipment the only way to determine if they survived the shipment is if they keep their appendages tucked under them while being held upside down by the back of the carapace. Crabs have normally recovered from the cold effects of shipping within the first 3 hours of introduction to the water. Constant monitoring of the newly shipped crabs is recommended since fast removal of dead crabs is necessary to maintain tank water quality and the health of the other crabs occupying the same tank or tanks connected to the same filtration system. One dead crab left in the system for 24 hours can kill the entire tank of crabs.

When introducing a crab to the aquarium it must be burped to remove the air from its body. Burping is performed by either manually holding or physically propping the crab face up underwater to remove the air. When burping you witness the air escaping from the crab in the form of bubbles. Positioning the crabs is favored over holding each individual crab when one is stocking the tank alone and has many crabs to introduce to the system.

Acclimation to the aquariums and their new environment after obtaining the crustaceans is recommended for at least a week (7 days), especially for feeding or behavioral studies. Since the crabs are wild caught, their last feeding is unknown, therefore keeping the crabs a week

without food ensures that they will eat when presented with food. For other studies, an acclimation of three days may be used but a week is highly recommended and preferred.

Seasonal Dependence and Shipping Conditions on Survival

Although ordering is performed year round, the quality and crab availability relies heavy upon the season. Healthy crabs are easily obtained from late spring to early fall months; during the good fishing months. During this time viable crabs are shipped immediately and in healthier conditions than crabs obtained in the winter months.

The survival of a shipment of crabs seems to depend on their availability, the species and quantity packaged in a single shipment. There is a higher level of success with survival of a shipment when no more than 16 crabs are packaged together. Normally the survival rate of shipment is 80% and anything below can warrant a replacement shipment from many of shipping companies. The exception to this rule is the blue crab. Their aggressive behavior typically puts their optimal survival rate at 75% so their survival during shipment is never guaranteed, in fact, there is normally a disclaimer. Therefore the blue crabs are normal handpicked from local groceries to increase survival. Shipping is only done if their availability from the grocers is scarce.

Cannibalism: An Indication of Overcrowding and Starvation

It is well documented that up to 11% of a crab's diet is crab, so precautions must be made to prevent cannibalism from occurring [27]. Cannibalism among crabs is used as an indicator that their environment and feeding are not optimum. It suggests either overcrowding or starvation among the crabs occupying the aquarium [27].

To prevent cannibalism, different crustacean species are not mixed together in one aquarium. This is because different species have different temperaments; as demonstrated by their reaction when being prodded with a cylindrical rod. If more than one species must occupy a tank, a divider is used to prevent the crabs from interacting.

Another precaution is preventing overcrowding. Over-crowding negatively affects water quality and causes fighting among crabs. Depending on the size and species, 8-10 crabs are kept in a 55 gallon aquarium at one time. Feeding is also used to prevent cannibalism among the crabs. A fed crab is less likely to attack another crab. It is believed that crab meat is a crab's diet when other food sources are not available; this is conducive to observations made by this lab when housing crabs [28]. In an overcrowded aquarium, crabs may still fight and kill but will refrain from eating each other if already fed. Fighting due to overcrowding is indicated by crab death or the appearance of loose limbs in the aquarium [29].

Feeding

Crabs are fed raw fish or shrimp depending on availability. To prevent fighting during feedings, the food is deposited right in front of the crab with a cylindrical rod. Enough food is prepared to feed all the crabs within the tank immediately. Initial consumption of the food is verified by the motion of the mouth and the disappearance of food from sight. If the food is eaten within a 5 minute time frame, more food is given to the crab to prevent fighting. After 30 minutes to an hour, any uneaten food is removed from the tank to prevent clogging of the filtration system. Also, uneaten food will decompose in the water negatively affecting water quality and subsequently the health of the crabs.

Euthanasia / Experimental Endpoint defined

Given that crabs are a higher-order animal model and are therefore perceived to experience pain, it is crucial to identify what may be considered pain or suffering in this animal model. Since crabs are not vocal, visual cues are used to indicate that the crab is experiencing pain. Appendix A contains an endpoint criterion for the nanoparticle exposure protocol but any indication of suffering or sickness in an animal during any procedure is immediately addressed to insure the health and wellbeing of the animal and experimental accuracy, validity, and integrity. The following criteria were used in the hemolymph studies described in this dissertation.

A suffering/sick animal shows the following behavior:

- Lethargy, non-movement even when prompted to move by poking or tapping
- Uncontrolled locomotive functions, the legs twitch involuntarily
- Uncontrolled mouth movements, the mouth may hang open involuntarily or unevenly (one side open)
- Its tail is hanging away from the body; normally a crab's tail is tucked under the body

A suffering/sick animal must be isolated from the other crabs otherwise it could be attacked, killed, and possibly eaten. The animal may recover on its own or it may be euthanized by being placed on ice for 1 hour to induce hibernation and consequential death. An animal is recorded as dead when the following criteria are met:

- No movement from the animal even when prompted to move by poking or tapping
- When held upside down by the back of the carapace, the appendages hang away from the carapace. In contrast, a live crab will fold its claws close to its body when held by the back of the carapace.

2.5 Neuropeptides: an Overview

Currently defined, neuropeptides are signaling molecules that are either secreted by or have an effect on the neurons in the nervous system [3, 6, 30, 31]. Neuropeptides are present in all biological systems and are the largest and most diverse class of signaling molecules [6, 32-34]. They are amide linked polypeptides typically less than 50 α-amino acids in length [6, 35]. Neuropeptides are involved in the regulation of many physiological processes including feeding, respiration, mating, and development and are present at femtomolar and picomolar concentrations *in vivo* [36].

Biosynthesis

Neuropeptides, like other peptides, are encoded into the genome and synthesized in the endoplasmic reticulum [2, 6]. After transcription and translation, the resulting precursor protein, known as a preprohormone if it contains more than one neuropeptide or a prohormone if it only contains one neuropeptide, is packaged in secretory vesicles for transport in the Golgi apparatus [2, 6]. During transport, the precursor protein is cleaved by processing enzymes called convertases or cleavases and the resulting neuropeptide(s) undergo post-translation modifications (PTMs) such as C-terminal amidation, cyclization of N-terminal glutamine or glutamic acid, formation of disulfide bonds, sulfonation of tyrosine, oxidation of methionine, etc [2, 3, 6]. Many of these PTMs are responsible for making the neuropeptide bioactive before its release to affect its target [2, 3, 33, 34]. Once released, neuropeptides are enzymatically degraded by peptidases in the extracellular space [2, 35].

Mode of Action

Neuropeptides generate a response in the receptor of their target neuron that either promotes, by excitation, or prevents, by inhibition, a response by the targeted neuron. Targeted

neurons are classified as local or distant depending on their proximity to the neuropeptide secreting cell body [6, 37]. Local targets consist of the autocrine functioning (a self-targeting neuronal cell that synthesizes the neuropeptide and is targeted as a byproduct of the neuropeptide's release) and paracrine actions (targeting the neuronal receptors in close proximity to the releasing cell) [6]. The distant target is reached via circulating hormonal delivery, which occurs when the neuropeptide is released into the circulatory system to exert an effect on a distant neuron [6, 37].

Neuropeptides may act as neurotransmitters, neuromodulators, or neurohormones [3]. Neurotransmitters produce fast paracrine actions [3, 38]. A neuromodulator acts slower as a secondary reaction either enhancing or inhibiting the initial effects of a neurotransmitter or neurohormone [38]. Neurohormones are the neuropeptides that effect distant targets via circulating hormonal delivery [38].

Expression

Neuropeptides vary both temporally and spatially in expression [1]. Temporal diversity of neuropeptide expression is in response to the organism's environment and also dependent upon the organism's developmental stage, gender, age and health [2, 3]. Spatial diversity is the result of organ function within the organism [12]. Given that different organs perform different functions within the system, the expression of neuropeptides mimics that diversity.

Although neuropeptide expression may vary considerably it may be categorized into three types [2]. Under normal circumstances within a system, Type 1 is present at relatively high levels, Type 2 is expressed at low or undetectable levels and will up-regulate in response to stimulation such as stressors, and Type 3 is only detectable during developmental stages (i.e.

juvenile or prenatal) [2, 39]. Type 2 and Type 3 are typically connected given that Type 3 may be "reactivated" in adulthood thereby becoming Type 2 [2, 31, 39].

A neuropeptide may be expressed at different stages and therefore be categorized into all three types [2, 12]. Distinction of the types is therefore independent of the neuropeptide itself and instead dependent on its function in the system; what cell body synthesizes the neuropeptide, releases it, and what it subsequently targets and its function [2].

Families

Currently defined, neuropeptide families are groups of neuropeptides with conserved sequence homology [3, 12]. Previously, families were defined as neuropeptides that shared similar functions [3]. As more neuropeptides were studied and plasticity of neuropeptide expression and functionality discovered, the current definition was postulated and accepted [40]. Therefore, although neuropeptides may share a similar sequence and be considered part of a family, they may have different functions.

Biological Function

Although the biological function of a neuropeptide within a system is diverse, known functions of crustacean neuropeptides typically fall under the following categories: cardioactive and myotropic, chromatophorotropins, regulatory, and developmental [12]. Cardioactive and myotropic neuropeptides affect the crustacean heart and muscles. They oversee the pyloric and hindgut functions. Chromatophorotropins are neuropeptides that are involved with the expression of pigment or color. Regulatory neuropeptides are responsible for maintaining normal operations of the crustacean such as glucose levels and respiration. Developmental neuropeptides regulate sex, maturity, and molting.

Given the simplistic structure of the neural network, an entire family or single peptide within the family may have diverse functionality [12, 40]. Also, despite sequence conservation across species, functionality may not be conserved. For example, Allatostatin inhibits juvenile hormone production in insects while it modulates heart rate and pyloric motor output in crabs and lobsters [12, 41]. Therefore the functional categories listed above are not rigid and many neuropeptides are multifunctional.

The neuropeptide families represented in the 692 entry in-house neuropeptide crustacean database are shown in Table 2.1. The table contains 30 peptide families; Allatostatin A, B, and C-types are all considered subfamilies of a single family but they are listed individually in the table because of their differing functions and sequence motifs. The families are grouped by known biological function types, listed above, and their known biological functions in crustaceans as previously published [3, 6, 9, 18, 37, 42, 43].

2.6 Neuropeptidomics: Discovery and Automated Sequence Identification

Mass Spectrometry Platform for Peptide Analysis

Shotgun peptidomics is normally employed for neuropeptide discovery and identification in a complex biological sample because no prior knowledge regarding the peptide sequence is necessary [1]. The sample is analyzed using liquid chromatography coupled to electrospray ionization (LC-ESI) and followed by tandem mass spectrometry (MS/MS) for rapid separation and identification of a peptides in the complex biological sample thereby increasing both sample throughput and analysis [44-47]. Proteomic and peptidomic exploration is also performed on matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) instruments [48]. However, ESI is preferred over MALDI since, traditionally, the multiple charges created in ESI

produce better fragmentation patterns for identification than the singly charges in MALDI [49]. Identifications based on fragmentation from tandem mass spectra in ESI are more reliable than mass matching in MALDI because of interference from isobaric species. Also, peptides may cleave as a result of the ionization process or sample preparation, thereby limiting mass matching [18, 50, 51]. Finally, sequence identification by tandem mass spectrometry allows for identification of novel peptides that might have otherwise been mislabeled as an isobaric peptide or noise [1, 52, 53].

Neuropeptide discovery involves *de novo* sequencing of the tandem mass spectrum and matching the best fit amino acid sequence to known peptide sequences [1]. If literature provides no direct or homologous match to the identified peptide then the peptide in question may be synthesized and tested for biologically relevance [19].

While the LC-ESI MS/MS platform allows for improved detection and identification of peptides in the sample, peptide sequence discovery and identification by mass spectrometric data analysis is dependent upon having decent mass spectra, reliable software for spectral analysis and database mining, and a decent database. Manual *de novo* sequencing of the spectra produced is a time consuming task and is susceptible to human error leading to lower reproducibility and questionable results. Automation of the sequencing analysis allows for rapid, precise, and reproducible identification of neuropeptides in a complex biological sample which benefits larger datasets [54-56].

Automated Sequence Identification

Although numerous search engines are commercially available, they may be grouped into three different categories by their method of identification. The three categories are database

searching, *de novo* sequencing, and sequence tagging [12, 57]. As the name implies, the database search method relies on an algorithm for searching a database and matching fragmentation patterns of the sequences in the database to the tandem MS spectra of the actual data. Mascot [57] and SEQUEST [58] are popular database searching software.

For the *de novo* sequencing method, the MS spectra are *de novo* sequenced, according to an algorithm, and the possible sequences determined for the spectrum are ranked by the probability of them actually being the peptide in question. PEAKS [59] and Lutefisk [60] are popular *de novo* sequencing software.

Sequence tagging is a hybrid method of *de novo* sequencing and database searching [12, 61]. Within a fragmentation spectrum, there are sequence tags, or clearly identifiable fragment motifs of the peptide. These tags are first identified in the tandem mass spectrum by *de novo* sequencing and then searched against the database for possible peptide matches or homologous peptide matches for novel peptide identification or post-translational modification identification in a peptide. GutenTag [62] and the newer Peaks DB [63] software follow this hybrid model of peptide identification.

Data Preprocessing

Automated data processing requires preprocessing of the mass spectra before identification [53, 64]. Peptide identification software use algorithms to determine amino acid sequences based on distances between singly charged fragment ion peaks and peptide mass fingerprinting [53, 57, 59, 64]. Data preprocessing parameters adjust mass accuracy based on the mass analyzer, calibration methods, and the fragmentation process used. It involves smoothing the baseline and background subtraction for noise reduction to identify peaks, deisotoping and

centroiding peaks for accurate peptide mass matching, and deconvolution of multiply charged species to singly charged ions to reduce the number of charge states for each ion [57, 59].

Deconvolution is important for identification of fragment ions by masses [59]. Highly charged states within a sample present more fragment peaks. Since algorithms do not discrimination between multiply or singly charged fragment peaks, their presence changes the mass difference between peaks as defined by the fragmentation ion series (a,b,c,x,y,z) [57]. Therefore, deconvolution of the spectrum leaves only singly charged peaks and simplifies the dataset allowing for correct and confident peptide identification. Noise filtering also prevents noise peaks from becoming misidentified as a fragment ion peak [59]. Baseline adjustment may be required given that some algorithms used rely on more abundant peaks for identification [53, 57]. Also, peak centering may help the mass accuracy of the dataset possibly correcting for mass shifts [57].

Once the data is preprocessed it is normally searched against a database of known peptides for peptide identification. Depending on the software used, the peptide may also be matched to a protein. Quantification is possible for identified peptides, but not for unidentified peaks when using commercially available software (Mascot [57] and PEAKS [59]) thereby further emphasizing the importance of the database.

2.7 Database Construction

Peptide databases are normally compiled by genomic, transcriptomic, or peptidomic information from the organism itself or a homologous species [12]. These methods maybe classified as either top down methods, which predicts peptides from genomic or transcriptomic data [65-67] or the bottom up method, which involves the compilation of identified peptides

from shotgun peptidomic experiments [51, 68]. When an organism is missing an annotated genome, transcriptomic and peptidomic methods are commonly employed or the genomic, transcriptomic, or peptidomic information from a homologous species may be used to construct a database.

Top Down Database Compilation

Given the ease of obtaining genomic data, or using an animal model that has a completed genomic database, peptide databases may be formulated by predicting cleavages in the annotated sequences [69]. Genomic data will provide an array of information for all possible neuropeptides formed by the system independent of neuropeptide plasticity. However, despite the popularity surrounding genome sequencing, not many species have their genome readily sequenced and there are challenges surrounding genome annotation, that ironically, mostly stem from the sequencing methods themselves [69-71].

With transcriptomic database compilation, expressed mRNA or ESTs are used to formulate a neuropeptide database. This methodology is especially useful when an annotated genome is missing for an organism. However, the cDNA synthesized from the mRNA only represents a portion of the actual DNA present in the system [72]. Therefore the resulting database is limited by the condition of the organism during mRNA extraction and dependent on neuropeptide plasticity.

For both top down methods, the progression from DNA to peptide prediction presents further complications due to the limited information on proteases that cleave the pre-prohormones for neuropeptide formation [73]. Attempting to bypass the limited knowledge on proteases by conducting a “no enzyme” automated search in relation to the original genome

requires an excessive amount of time and computer processing power because every possible cleavage is evaluated [74]. The software itself might even incorporate a basic enzymatic digest, despite the designation of a “no enzyme” search, to bypass the processing problem [65]. When known proteases are used to create a neuropeptide databank for analysis, it must be emphasized that the created database is incomplete [75]. However, the availability of an annotated genomic sequence improves the identification of peptides in a sample when conducting bottom up shotgun peptidomics.

Bottom Up Database Compilation

Another methodology for formulating neuropeptide databases is by shotgun peptidomics [12]. Peptides are extracted from the system either by harvesting tissues or fluids. The sample is than processed and undergoes biological mass spectrometry for analysis. The resulting potential neuropeptide peaks are then sequenced and the potential neuropeptides are authenticated by physiological studies or searching against a homologous species for identification.

While bottom up databases compiled from literature and peptidomic methods are more beneficial for studying model systems lacking an annotated genome, it presents its own set of problems. These databases are also incomplete because of insufficient data resulting from the inherent nature of neuropeptides and mass spectrometry limitations [1].

Neuropeptides are dynamically expressed within an individual and between individuals in a population. Due to the advances in the sensitivity of mass spectrometry, it has become ever more popular to decrease the sample size to a single cell or single animal in MS analysis especially with imaging [76-79]. This ideology is also supported by the 3R’s in animal research (replacement, reduction, and refinement) [80]. While single animal experiments help create and

maintain a congruent background for quantitative analysis, these results should not be applied to the entire population due to intra-species diversity. The reverse is also true, while results obtained from a population do provide information on the behavior of the individual; the results of the population are not unvarying for the individual. Therefore, the variance that exists between populations and individuals needs to be acknowledged and established in both single cell and single animal studies.

The second reason bottom up peptide databases are incomplete are that the peptide sequences themselves may be incomplete as a result of mass spectrometric limitations [81]. A region of the peptide amino acid sequence may not be sequenced due to *de novo* sequencing ambiguity or the inability to discriminate between isobaric residues. While mass spectrometry has aided in ground breaking innovations in the peptidomic field, the drawbacks to the technique need to be noted. Mass spectrometry tends to identify readily abundant or easily ionized peptides. Peptides that are easily ionized tend to contain basic amino acids like arginine, lysine and histidine. In contrast, peptides lacking these basic amino acid residues or that are less abundant in the spectrum may not be readily identified, thereby causing reduced peptidome coverage [82].

Another drawback to mass spectrometry deals with the isolation window for *de novo* sequencing. During tandem mass spectrometry, co-fragmentation of two peptides may occur as a result of peptide masking [51]. Peptide masking is caused by two peptides co-eluting at the same time or the peptide isolation window for fragmentation may be large enough to include a nearby peptide less than 3 m/z away. The co-fragmentation negatively affects *de novo* sequencing resulting in the less than optimal sequencing of peptides.

Another emerging complication arising from MS is the sequence scrambling that occurs due to cyclization of the peptide [83]. Cyclization arises when the nucleophile N-terminus attacks the oxalone C-terminal at b_n position. The cyclized peptide can then linearize by breaking at any b_n position in the cycle creating a new amino acid sequence. Although current studies debate the actual extent to which peptide sequence scrambling is a problem, it must still be considered and has been noted in peptidomic results by other groups [83-86].

There are several techniques that may be used to fill in the missing amino acid sequence of a neuropeptide. Advances in mass spectrometry instrumentation have led to the development of more sensitive, more accurate instruments such as the Orbitrap [87]. The combination of MALDI and ESI mass spectrometry techniques can also result in better sequence coverage of a peptide [88]. Alternatively, complementary fragmentation methods such as CID, ECD, and HCD may result in more complete sequence coverage for a peptide during *de novo* sequencing [89]. Also, comparison to known peptide sequences of homologous species may be used to elucidate the sections of missing sequence or to discriminate between isobaric residues in peptidomic research [65, 90].

Homologous Species Database Searches

Given that neuropeptide sequences are conserved across homologous species, it has become common practice to search the dataset generated by one organism against a database formulated by a homologous species for potential peptide matches [37, 91]. The genomic, transcriptomic, or peptidomic information of the homologous species may be used for analysis. Homologous searches are met with different levels of success, dependent upon the degree of neuropeptide plasticity represented in the database in relation to the dataset and the quality of the

dataset. However when homologous searches are successful, the neuropeptide database of the organism in question is expanded. Due to diversity in neuropeptide function across species, physiological studies are still necessary for determining the function of peptides within the organism of interest.

Internal Curating: Database Sharing

In light of database complexity, it has become increasingly popular for laboratories to overcome these disadvantages by creating in-house databases for their research. While this allows the laboratory to maintain scientific advantages in face of limited database information and to update their database rapidly when their research results in the discovery of new peptides, it creates obstacles for advancing neuropeptidomic research. The first obstacle is database sharing among the scientific community. Laboratories are either unwilling to share to maintain competitive advantage or unable to share due to lack of proper resources for sharing these often large data files rapidly. However, many publications are now requiring that these databases be available to the general scientific community. But since there are a number of methods for both database construction and entry formatting, information regarding the formatting language (Parse rules) and the software optimal for editing or sharing the database are needed for their successful application.

2.3 Respective Contributions Statement and Acknowledges

Portions of this chapter were reviewed and edited by Claire Schmerberg for their adaptation to Schmerberg, C.M., Woodards, N.A., and Li, L. *Mass Spectrom. Rev.* **2013** [in preparation].

2.4 Works Cited

1. Yin, P., et al., *Neuropeptidomics: mass spectrometry-based qualitative and quantitative analysis.* Methods Mol Biol, 2011. **789**: p. 223-36.
2. Hokfelt, T., et al., *Neuropeptides- an overview.* Neuropharmacology, 2000. **39**(8): p. 1337-56.
3. Strand, F.L., *Neuropeptides : regulators of physiological processes*. Cellular and molecular neuroscience. 1999, Cambridge, Mass.: MIT Press. 658 p.
4. Harvey-Clark, C., *IACUC Challenges in Invertebrate Research.* ILAR J, 2011. **52**(2): p. 213-20.
5. Horvath, K., et al., *Invertebrate welfare: an overlooked issue.* Ann Ist Super Sanita, 2013. **49**(1): p. 9-17.
6. Christie, A.E., E.A. Stemmler, and P.S. Dickinson, *Crustacean neuropeptides.* Cellular and molecular life sciences : CMLS, 2010. **67**(24): p. 4135-69.
7. Halliburton, W.D., *On the Blood of Decapod Crustacea.* The Journal of physiology, 1885. **6**(6): p. 300-463 4.
8. Zeleny, C., *Compensatory regulation.* Journal of Experimental Zoology, 1905. **2**(1): p. 1-102.
9. Keller, R., *Crustacean neuropeptides: structures, functions and comparative aspects.* Experientia, 1992. **48**(5): p. 439-48.
10. Wang, J., et al., *Combining capillary electrophoresis matrix-assisted laser desorption/ionization mass spectrometry and stable isotopic labeling techniques for comparative crustacean peptidomics.* J Chromatogr A, 2010. **1217**(26): p. 4463-70.

11. Chang, E.S., S.A. Chang, and E.P. Mulder, *Hormones in the Lives of Crustaceans: An Overview.* American Zoologist, 2001. **41**(5): p. 1090-1097.

12. Kwok, R., S.M. Chan, and S.S. Tobe, *Chapter 34 - Crustacean Bioactive Peptides*, in *Handbook of Biologically Active Peptides*, J.K. Abba, Editor. 2006, Academic Press: Burlington. p. 221-228.

13. Skiebe, P., *Neuropeptides are ubiquitous chemical mediators: Using the stomatogastric nervous system as a model system.* J Exp Biol, 2001. **204**(Pt 12): p. 2035-48.

14. Nusbaum, M.P. and M.P. Beenhakker, *A small-systems approach to motor pattern generation.* Nature, 2002. **417**(6886): p. 343-50.

15. Hsu, Y.W.A., et al., *Members of the crustacean hyperglycemic hormone (CHH) peptide family are differentially distributed both between and within the neuroendocrine organs of Cancer crabs: implications for differential release and pleiotropic function.* Journal of Experimental Biology, 2006. **209**(16): p. 3241-3256.

16. Ma, M., et al., *Expanding the Crustacean neuropeptidome using a multifaceted mass spectrometric approach.* Journal of Proteome Research, 2009. **8**(5): p. 2426-37.

17. Steinacker, A., *The anatomy of the decapod crustacean auxiliary heart.* Biol Bull, 1978. **154**(3): p. 497-507.

18. Chen, R., et al., *Measurement of neuropeptides in crustacean hemolymph via MALDI mass spectrometry.* J Am Soc Mass Spectrom, 2009. **20**(4): p. 708-18.

19. Ma, M., et al., *Mass spectrometric characterization and physiological actions of novel crustacean C-type allatostatins.* Peptides, 2009. **30**(9): p. 1660-8.

20. Marder, E. and D. Bucher, *Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs.* Annu Rev Physiol, 2007. **69**: p. 291-316.

21. Marder, E. and D. Bucher, *Central pattern generators and the control of rhythmic movements.* Curr Biol, 2001. **11**(23): p. R986-96.

22. Elwood, R.W., S. Barr, and L. Patterson, *Pain and stress in crustaceans?* Applied Animal Behaviour Science, 2009. **118**(3–4): p. 128-136.

23. National Research Council (U.S.). Committee for the Update of the Guide for the Care and Use of Laboratory Animals. and Institute for Laboratory Animal Research (U.S.), *Guide for the care and use of laboratory animals*. 8th ed. 2011, Washington, D.C.: National Academies Press. xxv, 220 p.

24. Flecknell, P., *Replacement, reduction and refinement.* ALTEX, 2002. **19**(2): p. 73-8.

25. Andrews, P.L., *Introduction: laboratory invertebrates: only spineless, or spineless and painless?* ILAR J, 2011. **52**(2): p. 121-5.

26. Luschen, W., et al., *Isolation, sequence analysis, and physiological properties of enkephalins in the nervous tissue of the shore crab Carcinus maenas L.* Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8671-5.

27. Fernandez, M., *Cannibalism in Dungeness crab Cancer magister: effects of predator-prey size ratio, density, and habitat type.* Marine Ecology Progress Series, 1999. **182**: p. 221-230.

28. Stevens, B.G., D.A. Armstrong, and R. Cusimano, *Feeding habits of the dungeness crab Cancer magister as determined by the index of relative importance.* Marine Biology, 1982. **72**(2): p. 135-145.

29. Smith, L.D., *Effects of limb autotomy and tethering on juvenile blue crab survival from cannibalism.* Marine ecology progress series. Oldendorf, 1995. **116**(1): p. 65-74.

30. Kastin, A.J., *What is a neuropeptide?* Trends in Neurosciences, 2000. **23**(3): p. 113.

31. de Wied, D., *Peptide hormones and neuropeptides: birds of a feather.* Trends in Neurosciences, 2000. **23**(3): p. 113.

32. Fu, Q., et al., *Hormone complement of the Cancer productus sinus gland and pericardial organ: an anatomical and mass spectrometric investigation.* The Journal of comparative neurology, 2005. **493**(4): p. 607-26.

33. Duve, H., et al., *Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.* European journal of biochemistry / FEBS, 1997. **250**(3): p. 727-34.

34. Fu, Q., M.F. Goy, and L. Li, *Identification of neuropeptides from the decapod crustacean sinus glands using nanoscale liquid chromatography tandem mass spectrometry.* Biochemical and biophysical research communications, 2005. **337**(3): p. 765-78.

35. Behrens, H.L., R. Chen, and L. Li, *Combining microdialysis, NanoLC-MS, and MALDI-TOF/TOF to detect neuropeptides secreted in the crab, Cancer borealis.* Analytical chemistry, 2008. **80**(18): p. 6949-58.

36. Masler, E.P., T.J. Kelly, and J.J. Menn, *Insect neuropeptides: discovery and application in insect management.* Archives of insect biochemistry and physiology, 1993. **22**(1-2): p. 87-111.

37. Ma, M., et al., *Characterization of the Carcinus maenas neuropeptidome by mass spectrometry and functional genomics.* General and comparative endocrinology, 2009. **161**(3): p. 320-34.

38. MalcolmBurrows, *Neurotransmitters, neuromodulators and neurohormones*. 'Oxford University Press'.

39. Tohyama, M., *An overview of the ontogeny of neurotransmitters and neuromodulators in the central nervous system*, in *Handbook of Chemical Neuroanatomy, Vol 10: Ontogeny of Transmitters and Peptides in the CNS*, A. Bjorklund, Hokfelt, T., Tohyama, M., Editor. 1992, Elsevier: Amsterdam. p. 647-650.

40. Duve, H., et al., *Allatostatins of the tiger prawn, Penaeus monodon (Crustacea : Penaeidea).* Peptides, 2002. **23**(6): p. 1039-1051.

41. Tobe, S.S. and W.G. Bendena, *Chapter 29 - Allatostatins*, in *Handbook of Biologically Active Peptides (Second Edition)*, K. Abba, Editor. 2013, Academic Press: Boston. p. 191-196.

42. Keller, R., *Crustacean Neuropeptides - Structures, Functions and Comparative Aspects.* Experientia, 1992. **48**(5): p. 439-448.

43. Ma, M.M., et al., *Mass spectral characterization of peptide transmitters/hormones in the nervous system and neuroendocrine organs of the American lobster Homarus americanus.* General and comparative endocrinology, 2008. **156**(2): p. 395-409.

44. Boonen, K., et al., *Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis.* J Sep Sci, 2008. **31**(3): p. 427-45.

45. Desiere, F., et al., *Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry.* Genome Biol, 2005. **6**(1): p. R9.

46. Wei, X. and L. Li, *Mass spectrometry-based proteomics and peptidomics for biomarker discovery in neurodegenerative diseases.* Int J Clin Exp Pathol, 2009. **2**(2): p. 132-48.

47. Petruzziello, F., et al., *Extensive characterization of Tupaia belangeri neuropeptidome using an integrated mass spectrometric approach.* J Proteome Res, 2012. **11**(2): p. 886-96.

48. Kuster, B., et al., *Mass spectrometry allows direct identification of proteins in large genomes.* Proteomics, 2001. **1**(5): p. 641-50.

49. El-Aneed, A., A. Cohen, and J. Banoub, *Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers.* Applied Spectroscopy Reviews, 2009. **44**(3): p. 210-230.

50. Beaudry, F., *Stability comparison between sample preparation procedures for mass spectrometry-based targeted or shotgun peptidomic analysis.* Analytical Biochemistry, 2010. **407**(2): p. 290-292.

51. Sturm, R.M., et al., *Mass Spectrometric Evaluation of Neuropeptidomic Profiles upon Heat Stabilization Treatment of Neuroendocrine Tissues in Crustaceans.* Journal of Proteome Research, 2012. **12**(2): p. 743-752.

52. Xu, H. and M.A. Freitas, *A dynamic noise level algorithm for spectral screening of peptide MS/MS spectra.* BMC bioinformatics, 2010. **11**: p. 436.

53. Mujezinovic, N., et al., *Reducing the haystack to find the needle: improved protein identification after fast elimination of non-interpretable peptide MS/MS spectra and noise reduction.* BMC Genomics, 2010. **11 Suppl 1**: p. S13.

54. Nesvizhskii, A.I., O. Vitek, and R. Aebersold, *Analysis and validation of proteomic data generated by tandem mass spectrometry.* Nat Methods, 2007. **4**(10): p. 787-97.

55. Venable, J.D., et al., *Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra.* Nat Methods, 2004. **1**(1): p. 39-45.

56. Ma, B. and R. Johnson, *De novo sequencing and homology searching.* Mol Cell Proteomics, 2012. **11**(2): p. O111 014902.

57. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.

58. Jiang, X., et al., *Optimization of filtering criterion for SEQUEST database searching to improve proteome coverage in shotgun proteomics.* BMC Bioinformatics, 2007. **8**: p. 323.

59. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry.* Rapid Commun Mass Spectrom, 2003. **17**(20): p. 2337-42.

60. Johnson, R.S. and J.A. Taylor, *Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry.* Mol Biotechnol, 2002. **22**(3): p. 301-15.

61. Hernandez, P., M. Muller, and R.D. Appel, *Automated protein identification by tandem mass spectrometry: issues and strategies.* Mass Spectrom Rev, 2006. **25**(2): p. 235-54.

62. Tabb, D.L., A. Saraf, and J.R. Yates, 3rd, *GutenTag: high-throughput sequence tagging via an empirically derived fragmentation model.* Anal Chem, 2003. **75**(23): p. 6415-21.

63. Zhang, J., et al., *PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification.* Mol Cell Proteomics, 2012. **11**(4): p. M111 010587.

64. Krishnan, S., et al., *Pre-processing liquid chromatography/high-resolution mass spectrometry data: extracting pure mass spectra by deconvolution from the invariance of isotopic distribution.* Rapid Communications in Mass Spectrometry, 2013. **27**(9): p. 917-923.

65. Menschaert, G., et al., *A hybrid, de novo based, genome-wide database search approach applied to the sea urchin neuropeptidome.* J Proteome Res, 2010. **9**(2): p. 990-6.

66. Li, B., et al., *Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle Tribolium castaneum.* Genome Res, 2008. **18**(1): p. 113-22.

67. Southey, B.R., S.L. Rodriguez-Zas, and J.V. Sweedler, *Characterization of the prohormone complement in cattle using genomic libraries and cleavage prediction approaches.* BMC Genomics, 2009. **10**: p. 228.

68. Pampanin, D.M., et al., *Detection of small bioactive peptides from Atlantic herring (Clupea harengus L.).* Peptides, 2012. **34**(2): p. 423-6.

69. Yandell, M. and D. Ence, *A beginner's guide to eukaryotic genome annotation.* Nat Rev Genet, 2012. **13**(5): p. 329-42.

70. Loman, N.J., et al., *High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity.* Nat Rev Microbiol, 2012. **10**(9): p. 599-606.

71. Reed, J.L., et al., *Towards multidimensional genome annotation.* Nat Rev Genet, 2006. **7**(2): p. 130-41.

72. Rudd, S., *Expressed sequence tags: alternative or complement to whole genome sequences?* Trends Plant Sci, 2003. **8**(7): p. 321-9.

73. Sabido, E., T. Tarrago, and E. Giralt, *Towards the identification of unknown neuropeptide precursor-processing enzymes: Design and synthesis of a new family of dipeptidyl phosphonate activity probes for substrate-based protease identification.* Bioorg Med Chem, 2010. **18**(23): p. 8350-5.

74. Cottrell, J.S., *Protein identification using MS/MS data.* Journal of Proteomics, 2011. **74**(10): p. 1842-1851.

75. Fouillen, L., et al., *Neuropeptide alterations in the tree shrew hypothalamus during volatile anesthesia.* J Proteomics, 2013. **80**: p. 311-9.

76. Rubakhin, S.S., et al., *Profiling metabolites and peptides in single cells.* Nat Methods, 2011. **8**(4 Suppl): p. S20-9.

77. Svatos, A., *Single-cell metabolomics comes of age: new developments in mass spectrometry profiling and imaging.* Anal Chem, 2011. **83**(13): p. 5037-44.

78. Lanni, E.J., S.S. Rubakhin, and J.V. Sweedler, *Mass spectrometry imaging and profiling of single cells.* J Proteomics, 2012. **75**(16): p. 5036-51.

79. Berman, E.S., et al., *Preparation of single cells for imaging/profiling mass spectrometry.* J Am Soc Mass Spectrom, 2008. **19**(8): p. 1230-6.

80. *Replacement, Reduction and Refinement of Animal Experiments in the Development and Control of Biological Products. Proceedings of a conference. Langen, Germany, November 2-4, 1994.* Dev Biol Stand, 1996. **86**: p. 1-368.

81. Bogdanov, B. and R.D. Smith, *Proteomics by FTICR mass spectrometry: top down and bottom up.* Mass Spectrom Rev, 2005. **24**(2): p. 168-200.

82. Wen, X., S. Dagan, and V.H. Wysocki, *Small-molecule analysis with silicon-nanoparticle-assisted laser desorption/ionization mass spectrometry.* Anal Chem, 2007. **79**(2): p. 434-44.

83. Saminathan, I.S., et al., *The extent and effects of peptide sequence scrambling via formation of macrocyclic B ions in model proteins.* J Am Soc Mass Spectrom, 2010. **21**(12): p. 2085-94.

84. Harrison, A.G., *Peptide sequence scrambling through cyclization of b(5) ions.* J Am Soc Mass Spectrom, 2008. **19**(12): p. 1776-80.

85. Goloborodko, A.A., et al., *Sequence scrambling in shotgun proteomics is negligible.* J Am Soc Mass Spectrom, 2011. **22**(7): p. 1121-4.

86. Li, X., et al., *Structural heterogeneity of doubly-charged peptide b-ions.* J Am Soc Mass Spectrom, 2011. **22**(2): p. 245-54.

87. McAlister, G.C., et al., *A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-orbitrap mass spectrometer.* J Proteome Res, 2008. **7**(8): p. 3127-36.

88. Samgina, T.Y., et al., *N-terminal tagging strategy for de novo sequencing of short peptides by ESI-MS/MS and MALDI-MS/MS.* J Am Soc Mass Spectrom, 2010. **21**(1): p. 104-11.

89. Frese, C.K., et al., *Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry.* Anal Chem, 2012. **84**(22): p. 9668-73.

90. Ranc, V., et al., *Broad characterization of endogenous peptides in the tree shrew visual system.* J Proteomics, 2012. **75**(9): p. 2526-35.

91. Ma, M., et al., *Combining in silico transcriptome mining and biological mass spectrometry for neuropeptide discovery in the Pacific white shrimp Litopenaeus vannamei.* Peptides, 2010. **31**(1): p. 27-43.

2.8 Figures and Tables

Figure 2.1 Crustacean Nervous System

This figure displays the location of the major ganglion, pericardial organs, brain, and sinus gland in the stomatogastric nervous system within the crustacean nervous system. In the feeding study outlined in Chapter 4, the pericardial organs were harvested.

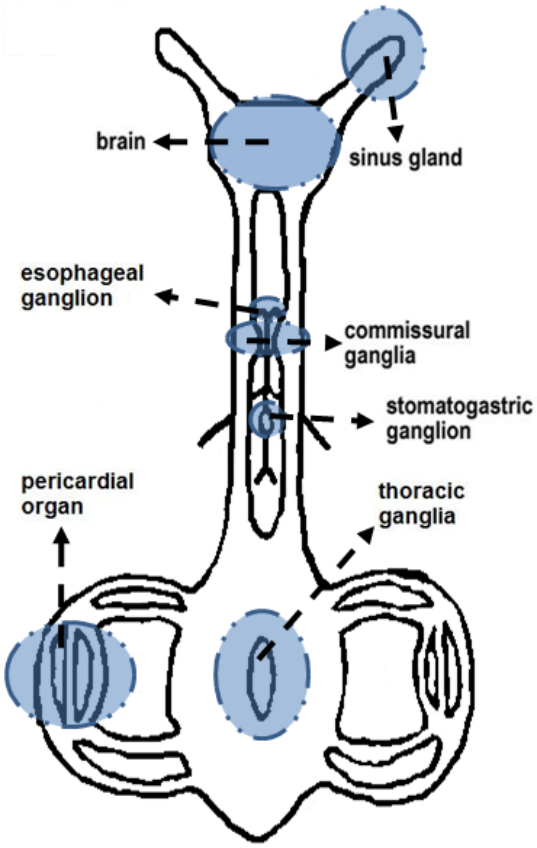


Table 2.1

This table displays the known functions of the 30 families of the updated 692 entry neuropeptide list. They are grouped by functional type. Given that Allatostatin A, B, and C type have various functions, they are listed separated despite being subgroups of the same Allatostatin family. Blank functional entries indicate that exact function of the families in the crustacean system is unknown.

Functional Type (ref[12])	Family	Function (ref[3, 6, 9, 18, 37, 42, 43])
Chromatophorotropins (pigment/color)	RPCH (Red Pigment concentrating hormone)	concentrates pigment in retinal pigment cells
	PDH(Pigment-dispersing hormone)	RPCH antagonist; disperses pigment
Cardioactive/myotropic (muscle function)	Proctolin	hindgut and cardiac
	RFamide	
	Actin	cardioactive protein
	Orcokinin	hindgut
	Others (Enkephalins)	Enkephalins: stimulate release of chromatophorotropins; inhibit ovarian maturation; contor carbohydrate metabolism and locomotory activity
	Allatostatin A-type	inhibits hindgut muscle contraction, inhibits motor output effects on STG and cardiac ganglion
	Allatostatin B-type	inhibits pyloric motor output
	Allatostatin C-type	inhibits cycle frequency of pyloric motor output by STG, modulate heart rate
	Kinins	stimulatory effect on pyloric rhythm
	myosuppressin	heart and cardiac ganglion
	Orcomyotropin	hindgut
	Pyrokinin	gastric mill excitation
	SIFamide	modulates pyloric neural network
	Tachykinin	pyloric and gastric mill activity
	Calcitonin-like	
	Allatotropin	digestive motor programs

Regulatory	CHH (Crustacean hyperglycemic hormone)	upregulates hemolymph, glucose and lipid levels; heptapancreatic amylase secretion; ion transport in gills; water uptake in molting; inhibits ecdysteroidogenesis in Y organs; MF synthesis in mandibular organs; AGH and protein synthesis in androgenic glands
Developmental	Corazonin	retraction of erythrophore pigment granules; dispersion of leucophorepigment granules; inhibits spermatogenesis
	Ecdysis triggering hormone	molting triggering hormone
	Eclosion hormone	
	bursicon	Chromatophorotropin inducing pigment dispersion and reducing RPCH
	CCAP	excites heart; excites other muscles
	Cryptocyanin	mouting protein related to hemocyanain but doesn't transport oxygen
unknown	Neuroparsin	
	CPRP (CHH precursor-related peptide)	
	Allatostatin Combos	
	Orcokinin/Orcomyotropin-related	
	RYamide	
5 functional types	**30 families**	**22 families with known functions**

Chapter 3: Crustacean Neuropeptide Database Compilation

3.1 Abstract

With the emergence of new techniques, software, and instruments for biological mass spectrometry analysis, we found it necessary to automate our neuropeptide identification and quantitation process. However, a quick review of our master in-house neuropeptide list revealed that it needed to be completed and updated for compatibility to many bioinformatics software. Our master list contained 245 neuropeptides representing 12 families and only 233 of the 245 entries contained sequences. After updating and a feasible expansion, the new in-house crustacean neuropeptide list contains 692 entries (690 with complete sequences, 1 partial sequence, and 691 with masses) from 30 families. The master neuropeptide database was expanded to 692 entries by compiling all available in-house neuropeptide lists and performing an extensive literature search (48 papers were used including 11 published by our lab and 37 by various other groups). Of the 449 new sequences added to our master list, 207 of them were identified by previous lab members and 242 were added as a result of literature searches. Of the 692 entries, only Penaeustatin 40 and 41 are missing information pertaining to their sequence and mass [1]. This chapter details the process of compiling the 692 entry crustacean neuropeptide list used throughout this research for data analysis. The entire process was completed in Microsoft Excel.

3.2 Introduction

Automated data processing software, bioinformatics, requires a database for rapid identification of the neuropeptides present in a biological system. The software uses the database in one of two ways; either by creating a virtual list of potential peptides for database

searching or to validate entries in a virtual list created by *de novo* sequencing [2]. Software that uses database derived virtual lists are Mascot [3] and SEQUEST [4]. Software that creates a virtual list from *de novo* sequenced candidates are PEAKS [5] and Lutefisk [6].

Given that software employs probability based results to determine the likelihood that a match is correct, the size of a database, the length of the database entries, and redundant entries are factors that affect the peptide score and false discovery rate [3, 5, 7, 8]. A small database will run the risk of not identifying a present peptide or, when coupled with a large dataset which is common in shotgun peptidomic analysis, it will falsely identify a peptide [9]. Alternatively, as the database increases in size the confidence in the peptide identification decreases [7, 9-11].

The accuracy and completeness of a database is not only dependent upon the method by which it was compiled but determines how accurately it identifies neuropeptides in a sample; this is of course assuming that the dataset being analyzed contains quality tandem mass spectra since dataset quality also affects identification. A database may be compiled by genomic, transcriptomic, or peptidomic methods [12]. Genomic and transcriptomic methods may be considered top down methods while peptidomic compilation is a bottom up method. Transcriptomic and peptidomic methods are commonly employed when an organism is missing an annotated genome. Since neuropeptides are commonly conserved across homologous species, another method employed in lieu of a missing genome, is searching a database constructed from a homologous species.

Crustacean Database Formation

Brachyura have no sequenced genome. In fact the first crustacean arthropoda genome was only sequenced in 2011 despite the numerous studies conducted on crustacean [13]. There

are partial EST libraries for three other crustacean models including *Carcinus maenas*; a species previously studied in our lab [14, 15]. Peptidomic strategies have also been employed by our lab for database construction [16].

Our in-house crustacean database was recently expanded from 245 neuropeptides to 692 using literature searches. The 692 entries were then combined to the HomoNP database containing 2132 entries from 75 various invertebrate and vertebrate species, most of which are homologous to crustaceans [17]. After redundant entries were eliminated, a database with 1958 entries remained and the entries were formatted according to both NCBI and SwissProt Parse rules for compatibility with Mascot, PEAKS, and ProteoIQ. The Parse rules indicate the way in which the information must formatted to meet the language requirements of the computing software so the information may be detected and analyzed.

3.3 Database Compilation Procedure

Materials and Methods

Literature and in-house databases were used with Microsoft Excel. The following outlines the method and literature criteria for composing the new database.

Literature Criteria

All literature used in this process had to contain a crustacean neuropeptide list. The list itself must contain information on the neuropeptide's family, sequence, and mass. Although lists were highly preferred for easy comparison, if the publication did not have the information listed, then name and sequence with mass had to be identifiable in the text.

In-house Database List Compilation

All available crustacean neuropeptide lists were gathered from the laboratory's shared storage drive (T:\drive) and lab members' personal lists. A total of 8 lists, including the master list, were analyzed. All lists collected, including the master list, were incomplete. They were classified as incomplete since the majority of their entries were missing at least one of the following classifications: family, name (where applicable), sequence, or monoisotopic mass.

All listed neuropeptide entries from the 8 collected lists were assembled into a single list for processing and comparison. The compiled list resulted in 1590 entries. All entries were checked for redundancy and accuracy against eleven previously published articles by our group that contained lists of identified neuropeptides [15, 16, 18-26].

Narrowing Down the 1590 Entries

The initial process of narrowing down the 1590 entries was performed by two people, independent of one another, and their results compared for accuracy. Of the 1590 entries, 455 of them were unique entries. The following sections explain how the entries were processed.

Mass Matching Unknown Sequences

A quick sort of the 1590 entries by sequence revealed that 27 entries had no sequence. These 27 masses were checked against the masses of the other 1563 entries containing sequences by use of the "countif" function. Of the 27 unknown sequences, the masses of 9 unknown sequences matched exactly to other masses whose entries contained sequences. These 9 non-sequenced matches were considered redundant entries and therefore deleted leaving 1581 entries.

Automated Redundant Sequence Elimination

Next, obvious redundant sequence entries were immediately eliminated narrowing the 1581 entries to 503 entries. This was done by using the remove replicates function in the data

tab of Excel. Before this function was applied, the sequences of the 1581 entries were formatted to allow the software to readily and properly identify redundant entries.

First all unknown sequences were assigned their mass as a sequence to prevent their loss. Then all C-terminally amidated neuropeptides were uniformly formatted with an "amide" ending instead of "a" to prevent the software from (1) misidentifying the amide with an alanine "A" and resulting in deletion of non-redundant entries and (2) redundant sequence entries ending in "amide" and "a" occurring in the final product.

Unified formatting of C-terminal amide entries was completed using the Excel Replace function with the match case mode selected. First all "amide" entries were replaced with "a" and then all "a" entries replaced with "amide" to prevent "amidemide" entries from occurring in the sequences. The match case function prevents an alanine conversion to an amide.

Bookmarking Unknown Sequences

The 18 masses with unknown sequences that did not match any of the other entries were removed from the list. The 18 masses were saved for comparison against published works but were eliminated from the following procedures for simplicity. This bookmarking action resulted in 485 entries.

Semi-automated Redundant Mass Elimination

The resulting 485 entries were sorted by mass and exactly identical mass entries, identified by using the "countif" formula, were manually checked for redundancy. The 19 identified redundant masses correlated to 39 entries that were manually checked. Manual checking was employed to differentiate between isobaric amino acids and isobaric peptide sequences (i.e. Penaeustatin 29 and Penaeustatin 16 both have the same masses but different

sequences) [1]. Entries with the similar masses as the result of isobaric amino acids (I/L and Q/K) were deleted resulting in 476 entries.

Li Lab Published Literature Validation

Manual searching of the 19 redundant mass entries also revealed potentially redundant sequences as a result of (1) poorly resolved N-terminal pyroglutamate identification ("pQ/pE/pyro/p" entries) and (2) possible sequence scrambling (XY versus YX entry in lieu of an otherwise similar sequence and mass). The eleven lab publications were searched to validate those entries.

Validating N-terminal Pyroglutamate

Pyroglutamate is formed by the enzymatic cyclization of an N-terminal Q or E or is created as an artifact during the extraction process [27, 28]. Given that crustaceans lack a genome, we thought it necessary to elucidate whether the pyroglutatmate was created from Q or E instead of assigning a universal pQ or pE entry format. This was decided to aid future genomic transcription studies or searches. The 476 entries contained a mixture of pyroglutamate formatting alternatively listing the amino acid as pQ, pE, p, or pyro. All 20 pyroglutamate entries were checked against the eleven published papers from our lab to resolve whether the peptide contained a pQ or pE so that redundant entries may be removed [15, 16, 18-26]. This resulted in narrowing the list to 472 entries.

Sequence Scrambling Clarification

Sequencing scrambling may be the result of a poor tandem mass spectrum to elucidate the structure, human error resulting in improper sequence entry, or even N-terminal cyclization of the sequence during the gas phrase of the ionization process [29]. All 13 potential sequence

scrambled entries were checked against the Li lab published literature for validation. This resulted in narrowing the list to 465 entries.

Final Literature List Validation

All 465 entries were checked for accuracy against the Li lab published literature. This resulted in 455 verified sequence entries (2 of which are not completely sequenced; Penaeustatin 40 and 41) [1]. The other 10 entries all belonged to the CHH (crustacean hyperglycemic hormone) precursor-relate peptide (CPRP) family and were only identified by name and mass.

Expanding the Database Through an Extensive Literature Search

It was our intent to compile a comprehensive list of all plausible crustacean neuropeptides within our lab. However, the resulting list of 483 entries (455 validated sequence entries and 28 unknown sequences) were still lacking imperative information pertaining to family, name, sequence, and mass that was not available in our literature. Given the century long history of crustacean research, literature was searched to complete these entries. This also resulted in the plausible expansion of the database to 692 crustacean neuropeptide entries.

The database compilation of the eight in-house neuropeptide lists and eleven publications, resulting in narrowing 1590 entries to 483, also revealed significant clues for discovering the missing information. It contained entries with (1) missing name numbers, (2) multiple sequences listed under a single name, (3) unidentified CPRP entries, and (4) missing Li lab literature entries. A criterion was formulated based on these four observations and this criterion used to expand the literature search to 37 other publications [1, 30-65]. The following explains the database observations for which the criterion was founded for both identification of unknown entries and feasible expansion of the database.

Missing Numbered Name Entries

If an entry contained a number in the name, (i.e. Penaeustatin 7 or Carcinustatin 10) then literature was searched to insure that all numbered entries were present in the database. The assumption being postulated was that if a single name was given to one peptide for easy identification and a name contained a number, then other peptides existed that had the same first name but different and sequential numbers. Basically, numbers within a name were used to differentiate different isoforms belonging to the same neuropeptide family in that species. Therefore the total number of neuropeptides within a group needed to be identified and all neuropeptide entries needed to be included.

For example both Penaeustatin and Carcinustatin were listed in the 483 entries (see Table 3.1). Penaeustatin contained numbers ranging from 1 to 41 but 40 was missing a completed sequence and mass and 41 was missing a sequence. Similarly, Carcinustatin had 18 entries numbered from 2 to 20. Both families had one overlapping sequence. According to our hypothesis, then (1) Penaeustatin contained at least 41 neuropeptides, hopefully all with completed sequences and masses and (2) at least 20 neuropeptides belonged to the Carcinustatin group so two were possibly missing.

Literature searches revealed that Penaeustatin did only have 41 neuropeptides and that 40 was missing a completed sequence and mass and 41 missing a sequence [1] and that Carcinustatin contained 20 neuropeptides, three of which overlapped with Penaeustatin [1, 65]. Literature searches also revealed that Carcinustatin has overlapping sequences with three Procastatin neuropeptides [43]. Procastatin contains 27 neuropeptides, five of those (the three overlapping the Carcinustatin entries and two unnamed additional ones) were present in the 483

entries and the other 22 were added [43]. All are Allatostatin A type neuropeptides and resulted in expanding our database by 22 new entries and properly identifying the family and name of the other 60 entries the database already contained.

Multiple Entries Under One Name

It was speculated that one sequence would contain one name for rapid and easy identification of neuropeptides with names. This viewpoint gained more validity during the missing name numbers searches. Therefore any entries containing the same name required further validation of the entry.

For instance, the 28 entries with unknown sequences posed a challenge. Luckily, 12 of them had a name and family identification; 11 from CPRP and 1 identified solely by the name Lepidopteran peptide. A quick search of the other 455 known sequences revealed the existence of 14 more entries identified solely as Lepidopteran peptides thereby breaking the one peptide, one name rule. A literature search revealed the missing sequence of the one Lepidopteran and identified the species that the neuropeptides were identified from, helocostatins and cydiastatins [59]. Also, 12 more Lepidoteran peptides were added [60]. The 26 Lepidoteran peptides range across 5 separate families (Allatostatin (A, B, and C Types), Allatotropin, RFamide, kinins, and CCAP).

CPRP Entries

Of the 227 neuropeptides added to the database, 113 are from CPRP, 96 truncated CPRPs and 14 full CPRPs from various crustacean species. Literature searches for identification of the 10 invalid sequence entries lead to the discovery and appreciation of the diversity in expression

of CPRP in crustaceans [18]. And given that CPRPs are derived from the same preprohormone as CHH, CHHs were also added to the database [26, 31, 35, 36, 42, 44, 45, 55, 57, 64].

Missing Literature Entries

In a few instances, when checking the accuracy of the 485 entries against the eleven Li lab publications, it became apparent that some of the published sequences were not represented by any of the 485 entries. For example, buriscon β fragment was identified in *Cancer meaenus* by our group and published but not present in any of the databases collected from the lab [15]. Bursicon contains an α and β, both are identified as major crustacean neuropeptides and were added [30]. In total 6 neuropeptides from five families (Bursicon, Calcitonin-like, Ecdysis triggering hormone, Eclosion hormone, and neuroparsin) were added to the database. These entries represent five of the six new families added to the database.

Criteria Validation of Neuropeptide Additions

Given that our objective was to formulate a complete and updated database for the group, and being quite aware of the expansive literature from which we could search and use to expand, criteria were formulated to prevent the addition of otherwise frivolous information to our database. It was our intent to add only neuropeptides whose plausibility of existence in our crustacean models was highly probable.

Also, only neuropeptides whose sequences were supported by literature were added to our database. Again, the addition of the neuropeptide had to be logical, in the sense that it was highly probable that it could exist in our models. We validated this by the existence of other family members already identified in our database. Therefore all added neuropeptides were linked by family, name, sequence, or mass to those neuropeptides already existing in the

database. The family criterion was only implemented in the case of CHHs in relation to CPRPs being present; all others were added based on name, sequence or mass.

Automated Database Completed

Although the 692 entry crustacean neuropeptide list identified by this extensive literature search is currently being implemented by our lab in crustacean research, it was not large enough to validate a database for automated searching by Mascot or PEAKS. We therefore incorporated this list into a database previously compiled by a former lab member for use on the software he developed. The database we collected from him contained 2132 entries.

Combined, the database entries totaled 2824 entries. Redundant sequence entries were identified by the "countif" function, sorted from largest redundant to smallest, and then manually searched to remove redundant entries. Manual searching was required to insure that all Li lab entries were retained in the final list. The final list contained 1958 entries.

This crustacean database with 1958 entries was formatted according to both NCBI and SwissProt Parse rules for compatibility with Mascot, Peaks, and ProteoIQ software for analysis.

3.4 Results

This research was initiated with the goal of completing the 245 entry master in-house crustacean neuropeptide list and formulating a database for automated neuropeptide identification. Quickly, it became evident that the list required updating to include recent discoveries by our laboratory resulting in 455 entries. However, the crustacean list was still missing information; therefore literature searches expanding beyond our published work were searched. This resulted in 692 entries with completed information.

The neuropeptide families represented in the 692 entry in-house neuropeptide crustacean database contains 30 peptide families; Allatostatin A, B, and C-types are all considered subfamilies of a single family but they are listed individually in the Table 2.1. The families are grouped by known biological function types explained in Chapter 2 and their known biological functions in crustacean decapod models are listed [15, 23, 30, 48, 66-68]. Table 3.2 shows the distribution of neuropeptides within each family. It shows the frequency of entries for the original 455 entries in comparison to the 692 and how many were added by outside literature searches. The completed 692 entry list is included in Appendix B of this work while the resulting 1958 database is found in Appendix B of Dr. Claire Schmerberg's dissertation [69].

3.5 Future Work and Conclusions

Since we lack a completed genome for any of the crustacean decapod models used by our laboratory, we developed a neuropeptide database to assist in their automated sequence identification. The database was uploaded on Mascot, PEAKS, and ProteoIQ for easy crustacean neuropeptide identification in future and current research by all lab members. The fasta file is also available on the T:/ drive so that the database may be uploaded to other software as needed.

With the advancement of crustacean neuropeptidome research, the database will require updating in the future. But for now, it is complete regarding both currently identified neuropeptides and neuropeptides with a highly probable chance of existence in the crustacean system. The database was formatted to be easily updated by the use of Microsoft Excel and easy addition to the fasta list by copy/pasting once the Parse rules of formatting are observed.

This research highlighted the gaps in knowledge regarding neuropeptide research despite a century's worth of analysis. It highlighted two neuropeptides that still require sequence

identification (Penaeustatin 40 and 41) which might be rectified with current mass spectrometry instrumentation [1]. It also revealed that many neuropeptides still require physiological studies to identify their functionality in the crustacean model systems. In conclusion, despite the wealth of knowledge on the crustacean neuropeptidome, it is far from complete and hopefully this database will be used to expand and complete that knowledge.

3.6 Respective Contributions Statement and Acknowledgments

Claire Schmerberg contributed greatly to this work by running parallel entry identification and formatting the fasta list according to Parse rules. We used Weifeng Cao's HomoNP database to expand our own work. I would also like to thank the cooperation of the Li lab in handing over their personal databases which resulted in the compilation of seven additional neuropeptide lists for analysis.

3.7 Works Cited

1. Duve, H., et al., *Allatostatins of the tiger prawn, Penaeus monodon (Crustacea : Penaeidea).* Peptides, 2002. **23**(6): p. 1039-1051.
2. Boonen, K., et al., *Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis.* J Sep Sci, 2008. **31**(3): p. 427-45.
3. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.
4. Eng, J.K., A.L. McCormack, and J.R. Yates Iii, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* Journal of the American Society for Mass Spectrometry, 1994. **5**(11): p. 976-989.
5. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry.* Rapid Commun Mass Spectrom, 2003. **17**(20): p. 2337-42.
6. Taylor, J.A. and R.S. Johnson, *Sequence database searches via de novo peptide sequencing by tandem mass spectrometry.* Rapid Commun Mass Spectrom, 1997. **11**(9): p. 1067-75.
7. Reidegeld, K.A., et al., *An easy-to-use Decoy Database Builder software tool, implementing different decoy strategies for false discovery rate calculation in automated MS/MS protein identifications.* Proteomics, 2008. **8**(6): p. 1129-37.
8. Zhang, J., et al., *PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification.* Mol Cell Proteomics, 2012. **11**(4): p. M111 010587.
9. MacCoss, M.J., *Computational analysis of shotgun proteomics data.* Curr Opin Chem Biol, 2005. **9**(1): p. 88-94.

10. Yates, J.R., 3rd, *Database searching using mass spectrometry data.* Electrophoresis, 1998. **19**(6): p. 893-900.

11. Searle, B.C., *Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies.* Proteomics, 2010. **10**(6): p. 1265-9.

12. Kwok, R., S.M. Chan, and S.S. Tobe, *Chapter 34 - Crustacean Bioactive Peptides*, in *Handbook of Biologically Active Peptides*, J.K. Abba, Editor. 2006, Academic Press: Burlington. p. 221-228.

13. Colbourne, J.K., et al., *The ecoresponsive genome of Daphnia pulex.* Science, 2011. **331**(6017): p. 555-61.

14. Ma, M., et al., *Combining in silico transcriptome mining and biological mass spectrometry for neuropeptide discovery in the Pacific white shrimp Litopenaeus vannamei.* Peptides, 2010. **31**(1): p. 27-43.

15. Ma, M., et al., *Characterization of the Carcinus maenas neuropeptidome by mass spectrometry and functional genomics.* General and comparative endocrinology, 2009. **161**(3): p. 320-34.

16. Cape, S.S., et al., *Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster, Homarus americanus.* Journal of Neurochemistry, 2008. **105**(3): p. 690-702.

17. Cao, W., *MS Informatics: Using Bioinformatic Tools to Enhance MS-based Neuropeptidomics and Proteomics*. 2012, The University of Wisconsin - Madison: Ann Arbor. p. 439.

18. Hui, L., et al., *Discovery and characterization of the Crustacean hyperglycemic hormone precursor related peptides (CPRP) and orcokinin neuropeptides in the sinus glands of the blue crab Callinectes sapidus using multiple tandem mass spectrometry techniques.* Journal of Proteome Research, 2011. **10**(9): p. 4219-29.

19. Fu, Q., M.F. Goy, and L.J. Li, *Identification of neuropeptides from the decapod crustacean sinus glands using nanoscale liquid chromatography tandem mass spectrometry.* Biochemical and Biophysical Research Communications, 2005. **337**(3): p. 765-778.

20. Szabo, T.M., et al., *Distribution and Physiological Effects of B-Type Allatostatins (Myoinhibitory Peptides, MIPs) in the Stomatogastric Nervous System of the Crab Cancer borealis.* Journal of Comparative Neurology, 2011. **519**(13): p. 2658-2676.

21. Behrens, H.L., R.B. Chen, and L.J. Li, *Combining microdialysis, nanoLC-MS, and MALDI-TOF/TOF to detect neuropeptides secreted in the crab, Cancer borealis.* Analytical chemistry, 2008. **80**(18): p. 6949-6958.

22. Hui, L.M., et al., *Discovery and Functional Study of a Novel Crustacean Tachykinin Neuropeptide.* ACS chemical neuroscience, 2011. **2**(12): p. 711-722.

23. Ma, M.M., et al., *Mass spectral characterization of peptide transmitters/hormones in the nervous system and neuroendocrine organs of the American lobster Homarus americanus.* General and comparative endocrinology, 2008. **156**(2): p. 395-409.

24. Fu, Q., et al., *Hormone complement of the Cancer productus sinus gland and pericardial organ: An anatomical and mass spectrometric investigation.* Journal of Comparative Neurology, 2005. **493**(4): p. 607-626.

25. Ma, M., et al., *Expanding the Crustacean neuropeptidome using a multifaceted mass spectrometric approach.* Journal of Proteome Research, 2009. **8**(5): p. 2426-37.

26. Ma, M., et al., *Combining bottom-up and top-down mass spectrometric strategies for de novo sequencing of the crustacean hyperglycemic hormone from Cancer borealis.* Analytical chemistry, 2009. **81**(1): p. 240-7.

27. Kumar, A. and A.K. Bachhawat, *Pyroglutamic acid: throwing light on a lightly studied metabolite.* Current Science, 2012. **102**(2): p. 288.

28. Predel, R., et al., *Post-translational modifications of the insect sulfakinins: sulfation, pyroglutamate-formation and O-methylation of glutamic acid.* European journal of biochemistry/FEBS, 1999. **263**(2): p. 552.

29. Bleiholder, C., et al., *Sequence-Scrambling Fragmentation Pathways of Protonated Peptides.* Journal of the American Chemical Society, 2008. **130**(52): p. 17774-17789.

30. Christie, A.E., E.A. Stemmler, and P.S. Dickinson, *Crustacean neuropeptides.* Cellular and molecular life sciences : CMLS, 2010. **67**(24): p. 4135-69.

31. Chan, S.M., et al., *Crustacean neuropeptide genes of the CHH/MIH/GIH family: implications from molecular studies.* General and comparative endocrinology, 2003. **134**(3): p. 214-9.

32. Chang, E.S., M.J. Bruce, and S.L. Tamone, *Regulation of Crustacean Molting - a Multi-Hormonal System.* American Zoologist, 1993. **33**(3): p. 324-329.

33. Duve, H., et al., *Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.* European Journal of Biochemistry, 1997. **250**(3): p. 727-734.

34. Dickinson, P.S., et al., *Identification of SYWKQCAFNAVSCFamide: a broadly conserved crustacean C-type allatostatin-like peptide with both neuromodulatory and cardioactive properties.* Journal of Experimental Biology, 2009. **212**(8): p. 1140-1152.

35. Fanjul-Moles, M.L., *Biochemical and functional aspects of crustacean hyperglycemic hormone in decapod crustaceans: Review and update.* Comparative Biochemistry and Physiology C-Toxicology & Pharmacology, 2006. **142**(3-4): p. 390-400.

36. de Kleijn, D.P., et al., *Cloning and expression of two mRNAs encoding structurally different crustacean hyperglycemic hormone precursors in the lobster Homarus americanus.* Biochimica et biophysica acta, 1995. **1260**(1): p. 62-6.

37. Sithigorngul, P., et al., *Four novel PYFs: members of NPY/PP peptide superfamily from the eyestalk of the giant tiger prawn Penaeus monodon.* Peptides, 2002. **23**(11): p. 1895-906.

38. de Kleijn, D.P., et al., *Cloning and expression of mRNA encoding prepro-gonad-inhibiting hormone (GIH) in the lobster Homarus americanus.* FEBS letters, 1994. **353**(3): p. 255-8.

39. Chang, E.S., G.D. Prestwich, and M.J. Bruce, *Amino-Acid-Sequence of a Peptide with Both Molt-Inhibiting and Hyperglycemic Activities in the Lobster, Homarus-Americanus.* Biochemical and Biophysical Research Communications, 1990. **171**(2): p. 818-826.

40. Tensen, C.P., et al., *Comparative Characterization of Hyperglycemic Neuropeptides from the Lobster Homarus-Americanus.* Peptides, 1991. **12**(2): p. 241-249.

41. Sithigorngul, P., et al., *Seven novel FMRFamide-like neuropeptide sequences from the eyestalk of the giant tiger prawn Penaeus monodon.* Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 2002. **131**(3): p. 325-337.

42. Tensen, C.P., et al., *Isolation and Amino-Acid-Sequence of Crustacean Hyperglycemic Hormone Precursor-Related Peptides.* Peptides, 1991. **12**(4): p. 673-681.

43. Yasuda-Kamatani, Y. and A. Yasuda, *Characteristic expression patterns of allatostatin-like peptide, FMRFamide-related peptide, orcokinin, tachykinin-related peptide, and SIFamide in the olfactory system of crayfish Procambarus clarkii.* Journal of Comparative Neurology, 2006. **496**(1): p. 135-147.

44. Hsu, Y.W.A., et al., *Members of the crustacean hyperglycemic hormone (CHH) peptide family are differentially distributed both between and within the neuroendocrine organs of Cancer crabs: implications for differential release and pleiotropic function.* Journal of Experimental Biology, 2006. **209**(16): p. 3241-3256.

45. Tensen, C.P., D.P. De Kleijn, and F. Van Herp, *Cloning and sequence analysis of cDNA encoding two crustacean hyperglycemic hormones from the lobster Homarus americanus.* European journal of biochemistry / FEBS, 1991. **200**(1): p. 103-6.

46. Stemmler, E.A., et al., *Molecular and mass spectral identification of the broadly conserved decapod crustacean neuropeptide pQIRYHQCYFNPISCF: The first PISCF-allatostatin (Manduca sexta- or C-type allatostatin) from a non-insect.* General and comparative endocrinology, 2010. **165**(1): p. 1-10.

47. Terwilliger, N.B., L. Dangott, and M. Ryan, *Cryptocyanin, a crustacean molting protein: Evolutionary link with arthropod hemocyanins and insect hexamerins.* Proceedings of the

National Academy of Sciences of the United States of America, 1999. **96**(5): p. 2013-2018.

48. Keller, R., *Crustacean Neuropeptides - Structures, Functions and Comparative Aspects.* Experientia, 1992. **48**(5): p. 439-448.

49. Sithigorngul, P., et al., *Three more novel FMRFamide-like neuropeptide sequences from the eyestalk of the giant freshwater prawn Macrobrachium rosenbergii.* Peptides, 2001. **22**(2): p. 191-197.

50. Sithigorngul, P., et al., *Novel FMRFamide-like neuropeptides from the eyestalk of the giant freshwater prawn Macrobrachium rosenbergii.* Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 1998. **120**(3): p. 587-595.

51. Trimmer, B.A., L.A. Kobierski, and E.A. Kravitz, *Purification and characterization of FMRFamidelike immunoreactive substances from the lobster nervous system: isolation and sequence analysis of two closely related peptides.* The Journal of comparative neurology, 1987. **266**(1): p. 16-26.

52. Mercier, A.J., et al., *Isolation of two FMRFamide-related peptides from crayfish pericardial organs.* Peptides, 1993. **14**(2): p. 137-43.

53. Luschen, W., et al., *Isolation, Sequence-Analysis, and Physiological-Properties of Enkephalins in the Nervous-Tissue of the Shore Crab Carcinus-Maenas L.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(19): p. 8671-8675.

54. Quackenbush, L.S., *Crustacean Endocrinology, a Review.* Canadian Journal of Fisheries and Aquatic Sciences, 1986. **43**(11): p. 2271-2282.

55. Soyez, D., *Occurrence and diversity of neuropeptides from the crustacean hyperglycemic hormone family in arthropods - A short review.* Neuropeptides in Development and Aging, 1997. **814**: p. 319-323.

56. DeKleijn, D.P.V. and F. VanHerp, *Molecular biology of neurohormone precursors in the eyestalk of Crustacea.* Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 1995. **112**(4): p. 573-579.

57. Chang, E.S., *Crustacean hyperglycemic hormone family: Old paradigms and new perspectives.* American Zoologist, 2001. **41**(3): p. 380-388.

58. Stay, B. and S.S. Tobe, *The role of allatostatins in juvenile hormone synthesis in insects and crustaceans.* Annual review of entomology, 2007. **52**: p. 277-99.

59. Duve, H., et al., *Lepidopteran peptides of the allatostatin superfamily.* Peptides, 1997. **18**(9): p. 1301-9.

60. Audsley, N. and R.J. Weaver, *Identification of neuropeptides from brains of larval Manduca sexta and Lacanobia oleracea using MALDI-TOF mass spectrometry and post-source decay.* Peptides, 2003. **24**(10): p. 1465-74.

61. Davis, N.T., et al., *Allatostatin-like-immunoreactive neurons of the tobacco hornworm, Manduca sexta, and isolation and identification of a new neuropeptide related to cockroach allatostatins.* The Journal of comparative neurology, 1997. **385**(2): p. 265-84.

62. Torfs, P., et al., *The kinin peptide family in invertebrates.* Annals of the New York Academy of Sciences, 1999. **897**: p. 361-73.

63. Gaus, G., et al., *The Sequences of Five Neuropeptides Isolated from Limulus using Antisera to FMRFamide.* The Biological Bulletin, 1993. **184**(3): p. 322-329.

64. Keller, R., et al., *Biological effects of neurohormones of the CHH/MIH/GIH peptide family in crustaceans*, in *Recent Developments in Comparative Endocrinology and Neurobiology*, E. Roubos, et al., Editors. 1999, Shaker Publishing BV: Nijmegen. p. 209–212.

65. Duve, H., et al., *Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab Carcinus maenas.* European journal of biochemistry / FEBS, 1997. **250**(3): p. 727-34.

66. Keller, R., *Crustacean neuropeptides: structures, functions and comparative aspects.* Experientia, 1992. **48**(5): p. 439-48.

67. Chen, R., et al., *Measurement of neuropeptides in crustacean hemolymph via MALDI mass spectrometry*. J Am Soc Mass Spectrom, 2009. **20**(4): p. 708-18.

68. Strand, F.L., *Neuropeptides : regulators of physiological processes*. Cellular and molecular neuroscience. 1999, Cambridge, Mass.: MIT Press. 658 p.

69. Schmerberg, C.M., *Functional Neuropeptidomics in the Decapod Crustacean: Method Development and Application to Behavioral Neuroscience Research*. 2012, The University of Wisconsin - Madison: Ann Arbor. p. 494.

3.8 Figures and Tables

Table 3.1

A list of the Carcinustatin, Penaeustatin, and Procastatin entries present in the final 692 entry crustacean neuropeptide database; all added information is **bolded**, information contained in the original 455 entries is not bolded. The missing names (Carcinustatin 1 and Carcinustatin 14) and missing sequences and masses (Penaeustatin 40 and 41) inspired literature searches that resulted in the identification of names for two unnamed sequences and the addition of 22 neuropeptides. All are Allatostatin A-type neuropeptides.

Name	(M+H)	Sequence
Carcinustatin 1, Penaeustatin 13	569.30821	YAFGLa
Carcinustatin 2	769.38790	EAYAFGLa
Carcinustatin 3	795.40357	EPYAFGLa
Carcinustatin 4	781.38792	DPYAFGLa
Carcinustatin 5	780.40390	NPYAFGLa
Carcinustatin 6	753.39300	SPYAFGLa
Carcinustatin 7	824.43012	ASPYAFGLa
Carcinustatin 8, Penaeustatin 6, **Procastatin 17**	794.41955	AGPYAFGLa
Carcinustatin 9	780.40390	GGPYAFGLa
Carcinustatin 10	962.50943	APQPYAFGLa
Carcinustatin 11	926.4730	ATGQYAFGLa
Carcinustatin 12	912.42840	PDMYAFGLa
Carcinustatin 13	2224.0743	EYDDMYTEKRPKVYAFGLa
Carcinustatin 14, Penaeustatin 35	585.30312	YSFGLa
Carcinustatin 15, **Procastatin 27**	810.41450	AGPYSFGLa
Carcinustatin 16	812.39370	GGPYSYGLa
Carcinustatin 17, **Procastatin 10**	857.41520	SGQYSFGLa
Carcinustatin 18	918.4026	SDMYSFGLa
Carcinustatin 19	1100.5081	APTDMYSFGLa

Carcinustatin 20	3150.5279	GYEDEDEDRPFYALGLGKRPRTYSFGLa
Penaeustatin 1	1568.7227	ANEDEDAASLFAFGLa
Penaeustatin 2	1979.9934	PDAEESNKRDRLYAFGLa
Penaeustatin 3	953.5203	DRLYAFGLa
Penaeustatin 4	881.45160	TGGPYAFGLa
Penaeustatin 5	881.45160	SAGPYAFGLa
Penaeustatin 7	850.42060	SGHYAFGLa
Penaeustatin 8	882.44680	ANQYAFGLa
Penaeustatin 9	825.42540	AGQYAFGLa
Penaeustatin 10	854.44068	TPSYAFGLa
Penaeustatin 11	951.4683	pQRDYAFGLa
Penaeustatin 12	771.36720	SDYAFGLa
Penaeustatin 14	912.4574	ANQYTFGLa
Penaeustatin 15	885.44650	ASQYTFGLa
Penaeustatin 16	814.40938	SQYTFGLa
Penaeustatin 17	599.31880	YTFGLa
Penaeustatin 18	893.42640	SGHYNFGLa
Penaeustatin 19	806.39440	GHYNFGLa
Penaeustatin 20	852.42500	AGPYEFGLa
Penaeustatin 21	838.40940	GGPYEFGLa
Penaeustatin 22	866.44070	AAPYEFGLa
Penaeustatin 23	781.38790	GPYEFGLa
Penaeustatin 24	811.39850	SPYEFGLa
Penaeustatin 25	838.40940	NPYEFGLa
Penaeustatin 26	1980.8934	NEVPDPETERNSYDFGLa
Penaeustatin 27	1866.8504	EVPDPETERNSYDFGLa
Penaeustatin 28	1426.6597	PETERNSYDFGLa
Penaeustatin 29	814.37300	NSYDFGLa
Penaeustatin 30	613.29800	YDFGLa
Penaeustatin 31	850.42060	AGHYSFGLa
Penaeustatin 32	957.4789	DRTYSFGLa
Penaeustatin 33	840.42500	PSAYSFGLa
Penaeustatin 34	941.4186	pQNMYSFGLa
Penaeustatin 36	3113.5651	DARGALDLDQSPAYASDLGKRIGSAYSFGLa
Penaeustatin 37	3099.5858	TARGALDLDQSPAYASDLGKRIGSAYSFGLa
Penaeustatin 38	812.43010	SVAYGFGLa
Penaeustatin 39	826.44580	TVAYGFGLa
Penaeustatin 40		(X)GIYGFGLa
Penaeustatin 41	2385.60000	

Procastatin 1	**911.43700**	**QNNYGFGLa**
Procastatin 2	**881.45160**	**TPNYAFGLa**
Procastatin 3	**901.42360**	**QGMYSFGLa**
Procastatin 4	**928.42330**	**PDMYSFGLa**
Procastatin 5	**910.46690**	**PDLYSFGLa**
Procastatin 6	**902.40770**	**ADMYSFGLa**
Procastatin 7	**884.45120**	**ADLYSFGLa**
Procastatin 8	**870.41040**	**SGNYNFGLa**
Procastatin 9	**956.49480**	**SRQYSFGLa**
Procastatin 11	**936.50500**	**PRNYAFGLa**
Procastatin 12	**2052.78290**	**TSDEEDDEDDQYYPYGLa**
Procastatin 13	**907.51480**	**PRVYGFGLa**
Procastatin 14	**1856.84830**	**ADSYGLAFGNGGDALEMGLa**
Procastatin 15	**700.33010**	**SYDFGLa**
Procastatin 16	**895.46720**	**TAGPYAFGLa**
Procastatin 18	810.41450	SGPYAFGLa
Procastatin 19	**824.43010**	**TGPYAFGLa**
Procastatin 20	**852.42500**	**ADPYAFGLa**
Procastatin 21	**877.45670**	**PNPYAFGLa**
Procastatin 22	**888.39200**	**DGMYSFGLa**
Procastatin 23	**841.42030**	**AGQYSFGLa**
Procastatin 24	826.40938	SGPYSFGLa
Procastatin 25	**1320.52260**	**EDYDSSDQYSLa**
Procastatin 26	**800.39370**	**SGAYSFGLa**

Table 3.2

The 30 families identified in the final 692 entry list and their respective number of entries (their frequency; freq) in the final 692 entry and 455 entry lists. The final column displays the number of added entries per family; all added entries were the result of the feasible expansion of the list by the external literature search and not previously identified by our laboratory.

Families	**692 freq**	**455 freq**	**Added**
Actin	2	1	1
Allatostatin A-type	149	125	24
Allatostatin B-type	34	32	2
Allatostatin Combos	7	7	0
Allatostatin C-type	3	0	3
Allatotropin	1	1	0
Bursicon	2	0	2
Calcitonin-like	1	0	1
CCAP	3	2	1
CHH (Crustacean hyperglycemic hormone)	27	0	27
Corazonin	1	1	0
CPRP (CHH precursor-related peptide)	164	41	123
Cryptocyanin	11	11	0
Ecdysis triggering hormone	1	0	1
Eclosion hormone	1	0	1
Kinins	10	2	8
Myosuppressin	1	1	0
Neuroparsin	1	0	1
Orcokinin	42	41	1
Orcokinin/Orcomyotropin-related	1	1	0
Orcomyotropin	4	4	0
Others	8	3	5
PDH(Pigment-dispersing hormone)	17	3	14
Proctolin	1	1	0
Pyrokinin	6	6	0
RFamide	142	131	11
RPCH	10	2	8
RYamide	23	22	1
SIFamide	4	4	0
Tachykinin	15	13	2
30	**692**	**455**	**237**

Chapter 4: *Cancer borealis* Feeding Time Course Study on Neuropeptides using DiLeu Quantitation: A Comparative Analysis of Identification and Quantitation Software

4.1 Abstract

The analysis of temporal changes in neuropeptide expression as it relates to feeding behavior was conducted using N, N-dimethyl leucine (DiLeu) labels for relative quantitation. The robustness of the tags was tested by preforming a DiLeu labeling matrix formulated from several biological, technical, and experimental replicates. In response to the large dataset generated, an automated data analysis approach was developed and applied. This chapter focuses on the comparison of three types of commercially available software for the automated identification and quantitation of DiLeu labeled crustacean neuropeptides as they temporally vary in relation to feeding.

4.2 Introduction

A fundamental aspect of energy homeostasis and animal survival is feeding behavior [1, 2]. Feeding behavior is regulated by a wide array of chemical messengers including neuropeptides [2, 3]. Allatostatins (ASTs) and RFamides are two peptide families that have been previously associated with feeding behavior through their regulation of the crustacean gastric mill and pyloric system [4, 5]. However each family contains numerous individual peptides, therefore identification of the exact peptides within a family that contribute to regulating feeding functions is beneficial [6, 7]. We studied the dynamics of crustacean neuropeptide expression as it relates to its various states of satiation by utilizing the 4-plex isobaric DiLeu labels.

Quantitation is important for tracking the differential expression of neuropeptides in a biological sample. Two categories for mass spectral quantitative analysis exist. The first is label

free quantitation that utilizes spectral counting while the other is labeled quantitation methods or tagging. Tagging is either isotopic or isobaric for analysis in the MS or tandem MS modes respectively. This research focuses on tagged quantitation.

Isotopic tags, analyzed in MS mode, may be used with MALDI or ESI ionization techniques. Isobaric tags are normally used with ESI ionization given the need for tandem mass analysis for quantitation, but with the development of newer mass analyzers, isobaric tags may gain popularity with MALDI ionization as well.

Formaldehyde labeling is a fast, simple, and stable isotopic labeling reaction for comparison of typically two samples but may also analyze three samples [8, 9]. The two-plex method involves labeling the two samples of interest with either H_2- or D_2- formaldehyde creating a 4Da mass difference between the light and heavy labels [9]. Previous studies from the Li lab demonstrate accurate concentration determination from peak areas for samples within a 10 fold concentration differences [8]. These tags have also been applied to a previous feeding study comparing the neuropeptide expression between fed and unfed Jonah crabs [10].

The Li lab has developed N, N-dimethyl leucine (DiLeu) reagents that serve as attractive alternatives to the commercially available isobaric tags for relative and absolute quantitation (iTRAQ) reagents due to their cost effectiveness, synthetic simplicity, improved stability in the aqueous phase, labeling efficiency, and improved fragmentation efficiency [11]. Similar to iTRAQ, DiLeu tags contain an amine reactive group, a balance group, and a reporter ion group [12-14]. The triazine ester reactive group labels the peptide at the N-terminus and ε-amino group of the lysine side-chain. A mass shift of 146 Da is observed for each DiLeu tag incorporated onto the peptide. Intense a_1 reporter ions at *m/z* 115.1, 116.1, 117.1, and 118.1 are

observed in the fragmentation spectrum for quantitation. DiLeu and iTRAQ reagents show comparable protein sequence coverage (~43%) and quantitation accuracy (<15%) for peptides resulting from a tryptic digest [11].

To test the robust nature of the labels, a DiLeu labeling matrix was created that compared both aliquot samples from the same time point and time point analysis across the four time points analyzed (unfed and 1, 3, and 6 hours after feeding). Also, technical, biological, and experimental replicates were performed. This analysis resulted in the generation of a very large dataset.

Advances in bioinformatics have emerged with identification and quantitation software for fast, accurate and reproducible analysis of the larger datasets created by shotgun proteomics. Automated data analysis first requires that the dataset be preprocessed before analysis. After preprocessing, the dataset is analyzed against a database for peptide identification followed by protein identification. Furthermore, quantitation may be performed on identified peptides and proteins [15]. Newer software versions allow the user to enter their own reporter ion labels for quantitation and upload their own database for searching.

For automated peptide identification and quantitation analysis I used three separate protein identification and quantitation software: Mascot, PEAKS, and ProteoIQ. To ensure that the data being analyzed by the software was as identical as possible, I utilized PLGS for preprocessing. This was necessary given that PEAKS had preprocessing capabilities and Mascot did not. The processing capability of the software and overall workflow for the dataset is illustrated in Figure 4.1.

Mascot is a probability based scoring database searching software [16]. Mascot is developed from MOWSE (molecular weight search) fragment database that uses the molecular weights inferred by the protease peptide fragments for protein identification. For MS/MS ion searches, first peptide fragment ions from the database are calculated and then the tandem fragment data is matched to the calculated values.

PEAKS identification approach utilizes a de novo sequencing algorithm based on dynamic programming for peptide sequencing, rescoring of the 10000 candidate sequences based on a stricter mass tolerance, and computation of a confidence score for the overall success of the sequencing [17]. Recently, PEAKS has evolved to allow the list of candidate sequences acquired by de novo sequencing to be searched against a database for identification [18].

ProteoIQ analyzes the results from other software (www.nusep.com). It is compatible with the database search results from Mascot, SEQUEST and X!Tandem software. It is capable of statistical validation of protein identifications by two methods. The first uses ProValT to arrange results by false discovery rate [15]. The second method is by protein probability. ProteoIQ also preforms quantitation by calculating expression tag ratios. Finally the software is capable of comparative proteomics; identifying unique entries between biological samples.

4.3 Materials and Methods

Crabs and Feeding Procedures

Refer to Appendix A for complete housing and feeding protocol. The Jonah crabs used for this experiment were all males roughly the same size (5.5-6 inches) and shipped from The Fresh Lobster Company. They were acclimated to their tanks for 72 hours without feeding. After the 72 hour equilibration time period, unfed crabs underwent tissue extraction and fed

crabs were given thawed pieces of frozen shrimp and viewed to ensure consumption. Consumption is characterized by the shrimp disappearing into the crab's mouth and was validated by the contents of the crab's stomach during dissection. Any fed crab that lack proper stomach content was removed from further analysis.

Tissue Collection

Refer to Appendix A for general tissue collection protocols and Appendix C for those pertaining to DiLeu labeling. Although both the pericardial organs (PO) and brains were both collected for analysis, this study only focuses on the datasets obtained by PO analysis.

DiLeu Tissue Neuropeptide Extraction

Refer to Appendix C.

DiLeu Labeling and Synthesis

Refer to Appendix C. Analysis was performed on the Waters QTOF Micro, which features a nanoACQUITY ultra-performance liquid chromatography front end coupled to ESI.

DiLeu Labeling Matrix

A four by four DiLeu labeling matrix was created and applied to individual crabs and aliquot samples (Figure 4.2). Initially, twenty crabs were collected for analysis to conduct analysis and an individual crab vs. aliquot sample analysis was also performed (Figure 4.3). The crabs were separated into five groups of four individuals. Two groups remained unfed while the other three were dissected 1, 3, and 6 hours after feeding. Neuropeptides were individually extracted from the PO of all twenty crabs. Four of the neuropeptide extracts from unfed crabs were pooled into a single sample and then divided back into four equal samples, creating aliquots and technical replicates. Each aliquot of the pooled sample was labeled with a separate DiLeu

tag. The other sixteen neuropeptide extracts were labeled with DiLeu tags and analyzed according to the matrix in Figure 4.2 where 1, 2, 3, and 4 arbitrarily represents the individual crabs collected within each time point. Analysis by rows gives four biological replicates of satiated vs. unfed crabs, they are considered time point analyses. Analysis by columns gives data comparing the neuropeptide content of four individual crabs within the same group, they are considered aliquot analyses. Therefore running the matrix analysis by both row and column resulted in eight chromatograms.

Sixteen crabs were collected to apply the DiLeu labeling matrix to pooled aliquot samples. They were separated into four groups of four crabs each. One group remained unfed while the other three were dissected 1, 3, and 6 hours after feeding. Neuropeptides were individually extracted from the PO of all sixteen crabs. The four neuropeptide extracts from the crabs in same feeding group were pooled into a single sample and then divided back into four equal samples, creating aliquots and technical replicates. Each aliquot of the pooled sample was labeled with a separate DiLeu tag. They were analyzed according to the matrix in Figure 4.2 where 1, 2, 3, and 4 arbitrarily represents the aliquot from the pooled analysis collected within each time point. These sixteen crabs represent one biological replicate.

Since the samples are distributed into aliquots, analysis by rows no longer gives four biological replicates of satiated vs. unfed crabs but four technical replicates of satiated vs. unfed crabs. Analysis by columns gives data comparing the neuropeptide content of technical replicates within the same feeding group. Since only four crabs were pooled for each time group to make aliquots, three biological replicates were achieved by dissecting a total of forty-eight crabs. Eight chromatograms for three biological replicates are twenty four chromatograms. The

QTOF settings were first ran to promote the identification of +2 charges and then reran with settings to promote the identification of the +1 charges. Analysis on each setting was performed three times for each sample to negate instrumental variability, resulting in the collection of one hundred and forty four chromatograms (Figure 4.4).

Automated Search Parameters

Only 72 chromatographs, pertaining to the +2 charges of the aliquot analysis, underwent automated data processing. The sequence of analysis described in Figure 4.1 was followed.

PLGS

The dataset underwent minor preprocessing using smooth only parameters in PLGS for quantitative purposes [19].

Mascot

Mascot Daemon was utilized for analysis of the PLGS pkl files. Between aliquots, the biological and technical replicates for the same time group were combined using the merge MS/MS files function. Between time points, the biological and technical replicates with the same tagging sequence were combined.

The Mascot parameters were analyzed with seven variable modifications (C-terminal amidation, dehydrogenate cysteine for disulfide bonds, both oxidation and dioxidation on methionine, sulfonate tyrosine, and N-terminus cyclization for glutamine and glutamate) and a fixed DiLeu modification. The analysis was initially run using the default Deamon parameters of a 1.2Da peptide tolerance and 0.6Da tandem MS tolerance, which are default standards for the default instrument settings. Then analysis was performed with an established ESI-QUAD-TOF instrument setting for 1.2Da, 200ppm, 100ppm, and 50ppm peptide tolerances for comparison.

All the Mascot results were converted to dat files for further analysis on ProteoIQ.

PEAKS

The PLGS pkl files were loaded onto PEAKS and batch analyzed according to labeling scheme. The data was analyzed using the batching scheme and with similar parameters to Mascot. The PEAKS parameters slightly differed from Mascot parameters because only five of the seven variable modifications were available (C-terminal amidation, dehydrogenate cysteine for disulfide bonds, methionine oxidation, sulfonate tyrosine, and N-terminus cyclization). Also a fixed DiLeu modification was calculated. A peptide tolerance of 200ppm, 100ppm, and 50pm was run with the software. All other parameters were not changed.

ProteoIQ

ProteoIQ analysis was conducted on the .dat files from the 200ppm, 100ppm, and 50ppm Mascot results. Reporter ion quantitation for custom isobaric tags was performed individually on the 24 Mascot results. The same Mascot PTMs were used. No correction factors were assigned to the DiLeu tags. For Peptide Parsing Options, default parameters were used expect the minimum peptide length was change from 6 amino acids to 5. All results are reported for the initial protein set, no changes to minimum peptide probability were made.

Manual Analysis

Manual analysis requires *de novo* processing of each peak in the chromatogram individually [20]. ProteinLynx DDA finder is applied to the chromatograph to detect the peaks. The reporter ion intensities corresponding to all detected peaks were recorded using MassLynx.

Repetitive peaks analysis was performed on the peaks corresponding to the +2 aliquot time point analysis with reporter ion intensities above 50, which is our preset minimum reporter

ion intensity limit for quantitation. Repetitive peaks were identified by technical, then experiment, and finally biological replicates based on retention time and m/z. A standard deviation cutoff of 0.5 was applied to both the retention time and m/z to determine repetitive peaks between each replicate analysis.

The coefficient of variance (CV) was calculated for the reporter ions of every detected peak.

4.4 Results and Discussions

Individual Crab Analysis and Animal to Animal Variability

An experiment was designed to determine the robustness of the labeling using the DiLeu tags developed by the Li lab for analysis of neuropeptide content of satiated vs. unfed crabs and for comparison of labeled neuropeptide content of a pooled sample versus individual crabs. This analysis resulted in nine chromatographs; eight from the matrix and one from the aliquot sample. Manual analysis of the peaks detected by the ProteinLynx DDA search function revealed five doubly charged and labeled neuropeptides that were present in both the aliquot and individual unfed crab datasets. They were chosen for comparison of the tags' relative peak heights to demonstrate the DiLeu tags' reproducibility of individual crabs (Figure 4.2). Analysis of the aliquot data obtained reveled that a 1:1:1:1 ratio of the four DiLeu tags demonstrated roughly a 1:1:1:1 peak ratio with 20% error, which is typical for commercially available iTRAQ reagents (Figure 4.5A) [21]. However, about two-fold expression level changes were observed for the individual crabs in the same time point group instead of an equal ratio (1:1:1:1) for our initial experiments (Figure 4.5B). Since the pooled aliquot samples demonstrates the robustness of the

tags, the different heights of the tags within the individual unfed crabs demonstrates that the animal to animal variability is great and individual crab analysis is not favorable.

The variability may be attributed to actual differences in the neuropeptide expression from animal to animal or it may be the result of dissection error. Removal of the entire PO requires precision in extraction that is not always reproducible from animal to animal even if performed by the same person. To minimize the variability, pooled tissues from four crabs were collected and aliquoted for analysis in the time course feeding comparison. This initial data displayed the feasibility of using the DiLeu reagents for peptide quantitation of complex tissue extracts in the crustacean model.

Aliquot PO Labeling Matrix

The experiment was repeated, following the DiLeu labeling matrix in Figure 4.2, but instead of using individual crab PO neuropeptide extractions, the four crab PO neuropeptide extractions at the same feeding time point were pooled and divided into aliquots before labeling. The relative quantitative analysis of the pooled PO is shown for the five doubly charged and labeled neuropeptides originally identified in the individual crab study (Figure 4.6). The ratios and coefficient of variance (CV) for the biological replicates (arbitrarily labeled DAGJ, EBHK, and FCIL refer to Figure 4.4) are the averages of twelve technical replicates (four set of experimental replicates ran three times each). The CV is also displayed for experimental group comparing unfed, 1, 3, and 6 hours after feeding. All experimental CVs are above 20%, the threshold for the aliquots, therefore indicating the statistical significance of the difference in the peak heights for the reporter ions of each feeding group.

The identification of the five doubly charged and labeled neuropeptides is given in Figure 4.5. Identification was initially achieved by mass matching to an in-house list of crustacean neuropeptides. The mass matching results were verified by *de novo* sequencing as described in Dr. Mingming Ma's dissertation using MassLynx [20]. Notice that peak 595 m/z has no sequence, it is a novel peptide that is not yet identified and has yet to be sequenced, even by our updated database. Identification of the peptide between runs was accomplished by comparing its tandem mass spectra for similar fragmentation patterns.

Manual Analysis

When ProteinLynx DDA finder was applied to the 36 chromatographs obtained by the +2 time point analysis, a total of 5274 peaks where identified. For manual quantitation, the peak list for each chromatograph generated by ProteinLynz DDA finder was transferred to an Excel file and their respective reporter ion intensities manually recorded by viewing the tandem mass fragmentation pattern of each peak using MassLynx. For identification, further manual analysis was required for *de novo* processing of each peak individually [20]. This extensive manual dataset analysis was necessary for two reasons. The first was due to the lack of a database for automated searching, which was eventually addressed in Chapter 3. The second was because our DiLeu tags were novel and the software available at the time did not allow the user to customize reporter ion intensities or enter tag masses for proper identification and quantitation of the dataset. Eventually, with the increase popularity of developing novel tags for quantitation, bioinformatics finally caught up to the demand for software capable of identification and quantitating novel tags. The automated data analysis is analyzed later in this chapter.

Given the overwhelming data acquired by technical, biological, and experiment replication for quantitative validation it was determined that repetitive peaks should first be determined to narrow the list of candidates for *de novo* analysis. Of the 5274 peaks, 4543 were hand analyzed first for technical replicates for repeat peaks (as determined by retention time and m/z), then by experimental replicates for repeat peaks, and finally by biological replicates for repeat peaks. The reason for narrowing the initial peaks to 4543 was initially to avoid analyzing peaks with a reporter ion signal intensity lower than 50; our preset limit for quantitation.

Technical replicates were created as a result of running each sample three times on the QTOF. This was performed to validate the identification of the peaks present in the analysis and remove the inconsistency of analysis in data-dependent acquisition (DDA) mode, resulting from switching between MS and MS/MS [22]. Only the peaks present in all three runs were selected for further analysis. This stringent criterion was selected so that a peak had three spectra for future quantitative comparison and three spectra to choose from to validate its *de novo* sequence identification; it was also helpful in narrowing down the list of candidates for further analysis. Of the 4543 peaks, 937 of them were determined to be present in all three QTOF runs. Statistically, 61% of the peaks identified were present in all three runs (Table 4.1).

Experimental replicates are the result of the tagging variability produced by labeling an aliquot sample with a different tag. Experimental replicates were performed to validate that quantitation was not dependent upon the location of the tag. Only the peaks present in all four replicates were chosen for further analysis. Of the 937 peaks, 109 were present in all four replicates. Statistically, the location of the tags was not significant in reference to number of peaks detected (Table 4.2).

Biological replicates were produced as a result of repeating the experiments on three sets of crabs. Biological replicates were performed for statistical significance of the quantitation method performed. Of the 109 peaks analyzed for repetitive peaks across biological samples, only 14 peaks were present in all three biological samples. The differences in biological replicates were not deemed significant in relation to the number of identified peaks (Table 4.3). The resulting 14 peaks were identified by mass matching. Of the 14, 10 matched to our list of neuropeptides (Table 4.4).

Of the 14 neuropeptides, 5 were previously identified in the aliquot analysis initially performed for the animal to animal variability experiments. One of these peptide identifications was questionable given a pE at the N-terminus and lack of lysine in the sequence so it was removed from the following quantitative analysis.

The CV calculation is not influenced by tag location and is an easy way to cover analysis on all four reporter tags. Given that the samples analyzed are not technical replicates but experimental analyses of the different feeding time points, a CV greater than 20% will indicate significant differences among the reporter ions thereby identifying peptides that are significantly up or down regulated. However, the CV calculated from the CVs of the same technical replicate, will determine the validity of the tags' quantitation. Since the technical replicates are the same sample measured three times on the QTOF, their CVs should be the same for each run. This technical replicate CV should be under 20% to indicate a decent measure. Unfortunately, only 6 of 24 CVs were under 20%.

Review of the data indicated that not only were the parameters for determining replicate peaks too lenient but the initial process of removing the peaks with reporter intensities below 50

was either done poorly or there was data loss between transfer of the dataset. The parameters used were based on standard deviation calculations of the peaks. It is a relatively easy calculation in Excel and was utilized to overcome the large dataset. Since a standard deviation cutoff of 0.5 was applied to both the retention time and m/z this resulted in an allowed error of 0.7 min and 1320 ppm. The enormous ppm allowance was unknown at the time, it is now realized that a standard deviation of 0.076 would be needed to correlate to a more appropriate 200 ppm cut-off. In retrospect, standard deviation was not an accurate measure for determining m/z repetition and ppm should have been utilized from the start. Since a year's worth of data analysis resulted in failure, the need for automating identification and quantitation was greatly emphasized.

Automated Data processing

The month it would take to completely analyze a single chromatogram manually was completed in one minute, after parameters were set. However, since the software used was initially designed for protein identification, the results still required manual analysis.

Mascot

The PLGS files were batch analyzed using Mascot Deamon according to labeling scheme. The initial default parameters for Deamon resulted in 18 top scoring neuropeptides. Comparison to the raw data corresponding to the top 18 neuropeptides resulted in similarity of only 11 of the 37 ratios (Table 4.5). Reanalysis was performed using an established ESI-QUAD-TOF instrument setting for 1.2Da, 200 ppm, 100 ppm, and 50 ppm peptide tolerances. The top hits, as determined by Mascot score, of all five settings were compared; the majority of hits were found with 200 ppm peptide tolerance.

In total, 39 peptides were identified from the analysis. Of the 39 peptides, 25 were from our 692 entry crustacean database and the other 14 were from various species including fly and worm (Table 4.6). The 25 crustacean neuropeptides were spread across 9 families; 6 families and 10 neuropeptides were previously identified in a feeding analysis experiment using formaldehyde labeling on extracts from the brain and PO tissues of Jonah crabs at unfed and one hour after feeding [10].

Of 39 peptides identified, 7 peptides were found in all parameters used and 9 were found across all time points, 5 of which overlapped each other, resulting in 11 neuropeptides (Table 4.7). Of the five most prevalent neuropeptides only three underwent further quantitation analysis. For quantitation, Mascot has a rigid ratio structure that assigns the lowest m/z as the control. For DiLeu, this was m/z 115, so all ratios were compared to m/z 115. Because of the staggered labeling scheme used, it prevented direct analysis of the ratios between time point analyses. Inverse ratios and comparisons between experimental replicates were required for statistical analysis of the quantitation data (Figure 4.7). Also the batch feature on Mascot Daemon pulls all processed data into a single file which represented 9 QTOF runs (3 technical replicates of 3 biological replicates using the same labeling scheme) so a N>3 was applied for quantitative comparison. Only two of the three peptides were also identified by the previous formaldehyde labeling study and only one peptide was similar to those results.

PEAKS

Peaks identified 47 neuropeptides from our 692 entry list. Of those 47, 18 matched to the ones previously identified by Mascot, and 11 matched to ones previously identified in the formaldehyde labeling study (Table 4.8).

ProteoIQ

ProteoIQ analysis was conducted on the .dat files from the 200ppm, 100ppm, and 50ppm Mascot results. Only 43 peptides were identified in total and only 8 of those were unique and only 7 of those 8 matched the 692 entry database. However, ProteoIQ allows the user to determine which reporter ion represents the control. This was beneficial considering the labeling scheme applied, allowing for direct comparison between experimental replicates. Of the 7 neuropeptides, half were previously identified by the formaldehyde study (Table 4.9).

Comparison of the software

The raw data used for the ratio comparison to the initial 18 top scoring Mascot hits was compared to the 3 peptides for Mascot quantitation analysis, 11 overlapping peptides identified in PEAKS, and the 7 neuropeptides identified by ProteoIQ. This resulted in 23 neuropeptides from the four types of analysis. Given that Mascot quantification is the limiting factor, only 3 of those 23 peptides were identified in more than two of the four analyses (raw data, Mascot, PEAKS, and ProteoIQ) and only 2 were previously identified in the formaldehyde labeling study (Figure 4.8).

The ratios of the 4 analyses were compared to that of the formaldehyde study with an average theoretical error well over 100%. However, statistically there is no difference between the ratios found by three types of software and the formaldehyde labeling study.

A Venn diagram displaying the overlap of the identified peptides for each of the software is shown in Figure 4.9. The results from Mascot and PEAKS analysis appear to be complementary for identification of neuropeptides. Since ProteoIQ is based on Mascot results, it is not surprising that the neuropeptides identified by ProteoIQ completely overlap with Mascot.

Disadvantages with the Peptidomic Application

Both PEAKS and Mascot are originally created for protein identification by first identifying the peptide fragments produced by enzymatic digestion of the protein and then matching those peptide fragments to their respective proteins. Despite its primary use for protein identification, the software may be applied to neuropeptide identification. In order to do this, important factors must be taken under consideration. These include the software forcing incorrect protein identification and missed identifications of smaller neuropeptides.

Given the sequence homology of neuropeptide families, the software will group homologous peptides together to identify a non-existent protein. This is actually the result of hierarchical clustering of peptide hits [23]. This phenomenon is especially apparent within the Orcokinin family where several of the members differ by the addition of one amino acid to the C-terminus. The longest Orcokinin neuropeptide will be identified as the protein and the other peptides will be considered hits for identification of the single protein (Figure 4.10). For this reason, identified proteins and their scores are irrelevant and only the peptides identified for the "protein" are considered.

Another problematic feature of the peptide database is the length of the peptides in the database. Mascot does not identify sequences smaller than five amino acids. Due to the small probability of matching a short peptide sequence, peptides shorter than seven amino acids in length normally receive very low scores [16]. Given that scoring is important in the algorithms, the lower scoring peptides are either "lost" in the filtering process or not shown in the dataset [16-18, 23]. A more valid method of determining the accuracy of peptide identification is the

sequence coverage. Sequence coverage also compensates for the hierarchical clustering mentioned previously.

4.5 Future Work

First, this work represents sample collection on the PO, although the brain samples were also collected they never underwent neuropeptide extraction or labeling. Therefore, analysis on the brain and other neural tissues should be considered. Also the data collected was collected by runs using instrument parameters to detect both multiply and singly charged species, but only the multiply charged run are analyzed here, the singly charged runs should also be reviewed.

Analysis of the data produced from the automated data processing could be expanded to include the peptides not receiving a top score. This would be a feasible expansion considering that the software is originally meant for protein analysis and is therefore not adequate for identification of smaller peptides less than 6 amino acids in length and also since the score is based on the size of our database (1958), when a normal protein database is at least a few thousand entries.

Since quantitation is only performed on identified peptides, manual analysis is not negated. As demonstrated above, although present in multiple runs, 4 of the 14 prevalent neuropeptides identified in the by hand analysis were not identified by our database. More attention should also be focused on the high scoring neuropeptides that are not on our 692 entry list but present in other homologous species. It should be determined if these peptides are novel crustacean peptides present in our sample or false database identifications.

Although more analysis may be applied to this dataset, reanalysis of the isobaric labeled samples on the Synapt G2 QTOF or Q-Extractive orbitrap mass spectrometers may also be

explored in the future for better quality tandem mass spectra to improve software based identification.

4.6 Respective Contributions and Acknowledgements

The experiment was conducted in collaboration with Feng Xiang who synthesized the 4-plex DiLeu labeling reagents and designed the data collection matrix. Dr. Claire Schmerberg added substantially in the formation of the database, as indicated in Chapter 3, Micromass training for data acquisition and with data analysis discussion. Chenxi Yang and Jingxin Wang aided with ProteoIQ training, and Robby Cunningham and Rob Sturm aided with PEAKS training.

4.7 Works Cited

1. Hillebrand, J.J.G., D. de Wied, and R.A.H. Adan, *Neuropeptides, food intake and body weight regulation: a hypothalamic focus.* Peptides, 2002. **23**(12): p. 2283-2306.
2. Barbano, M.F. and M. Cador, *Various aspects of feeding behavior can be partially dissociated in the rat by the incentive properties of food and the physiological state.* Behav Neurosci, 2005. **119**(5): p. 1244-53.
3. Klavdieva, M.M., *The history of neuropeptides I.* Front Neuroendocrinol, 1995. **16**(4): p. 293-321.
4. Bechtold, D.A. and S.M. Luckman, *The role of RFamide peptides in feeding.* The Journal of endocrinology, 2007. **192**(1): p. 3-15.
5. Fu, Q., et al., *Mass spectrometric characterization and physiological actions of VPNDWAHFRGSWamide, a novel B type allatostatin in the crab, Cancer borealis.* Journal of neurochemistry, 2007. **101**(4): p. 1099-107.
6. Ma, M., et al., *Mass spectrometric characterization and physiological actions of novel crustacean C-type allatostatins.* Peptides, 2009. **30**(9): p. 1660-8.
7. Kwok, R., S.M. Chan, and S.S. Tobe, *Chapter 34 - Crustacean Bioactive Peptides*, in *Handbook of Biologically Active Peptides*, J.K. Abba, Editor. 2006, Academic Press: Burlington. p. 221-228.
8. DeKeyser, S.S. and L. Li, *Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry quantitation via in cell combination.* Analyst, 2006. **131**(2): p. 281-90.
9. Hsu, J.L., et al., *Stable-isotope dimethyl labeling for quantitative proteomics.* Anal Chem, 2003. **75**(24): p. 6843-52.

10. Chen, R., et al., *Comparative Neuropeptidomic Analysis of Food Intake via a Multi-faceted Mass Spectrometric Approach.* ACS Chem Neurosci, 2010. **1**(3): p. 204-214.

11. Xiang, F., et al., *N,N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics.* Anal Chem, 2010. **82**(7): p. 2817-25.

12. Wiese, S., et al., *Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research.* Proteomics, 2007. **7**(3): p. 340-50.

13. Ow, S.Y., et al., *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.

14. Pottiez, G., et al., *Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples.* J Proteome Res, 2012. **11**(7): p. 3774-81.

15. Weatherly, D.B., et al., *A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results.* Mol Cell Proteomics, 2005. **4**(6): p. 762-72.

16. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.

17. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry.* Rapid Commun Mass Spectrom, 2003. **17**(20): p. 2337-42.

18. Zhang, J., et al., *PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification.* Mol Cell Proteomics, 2012. **11**(4): p. M111 010587.

19. Sun, W., et al., *Characterization of the liver tissue interstitial fluid (TIF) proteome indicates potential for application in liver disease biomarker discovery.* J Proteome Res, 2010. **9**(2): p. 1020-31.

20. Ma, M., *Exploring neuropeptidomes via novel mass spectral techniques*. 2008, The University of Wisconsin - Madison: Ann Arbor. p. 289.

21. Gan, C.S., et al., *Technical, Experimental, and Biological Variations in Isobaric Tags for Relative and Absolute Quantitation (iTRAQ).* Journal of Proteome Research, 2007. **6**(2): p. 821-827.

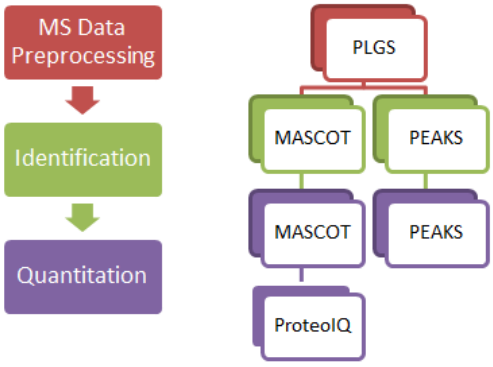
22. Schmerberg, C.M., *Functional Neuropeptidomics in the Decapod Crustacean: Method Development and Application to Behavioral Neuroscience Research*. 2012, The University of Wisconsin - Madison: Ann Arbor. p. 494.

23. Koskinen, V.R., et al., *Hierarchical clustering of shotgun proteomics data.* Mol Cell Proteomics, 2011. **10**(6): p. M110 003822.

4.8 Figures and Tables

Figure 4.1 Automated Data Processing

For automated data processing the dataset underwent three steps for analysis. First preprocessing was performed by PLGS, the identification and quantitation performed by Mascot and PEAKS independently, and finally the Mascot results were analyzed by ProteoIQ. The software capabilities are listed in the table. A plus sign indicates the software is capable of performing the task, a negative indicates it is not, a plus and minus indicates that the software is normally capable of preforming the task but only on preset parameters that do not match our requirements.



	MASCOT	PEAKS	ProteoIQ	PLGS
MS Data Preprocessing	-	+	-	+
Protein Identification	+	+	-	+/-
Protein Quantitation	+	+	+	+/-

Figure 4.2 DiLeu Labeling Matrix

The four by four experimental labeling matrix used for analyzing the PO samples collected. This matrix was applied to both individual crab samples and the pooled aliquot samples of four crabs. The description resulting from analysis by row or column is given in relation to whether the matrix was applied to single crabs or aliquot samples.

	1	2	3	4
Unfed	115	116	117	118
1h	116	117	118	115
3h	117	118	115	116
6h	118	115	116	117

Matrix Analysis	Single Crab	Aliquots (4 crabs pooled)
By Row	4 Biological Replicates Animal to animal variability	4 Technical Replicates Tag to tag variability
By Column	4 sets of experiments run: Same feeding setup Different biological replicates Different tags	4 experimental replicates: Same feeding setup Same biological replicate Different tags

Figure 4.3 Individual Crab vs. Aliquot Analysis

This figure represents the analysis performed on the individual and aliquot samples demonstrating how aliquots were created.

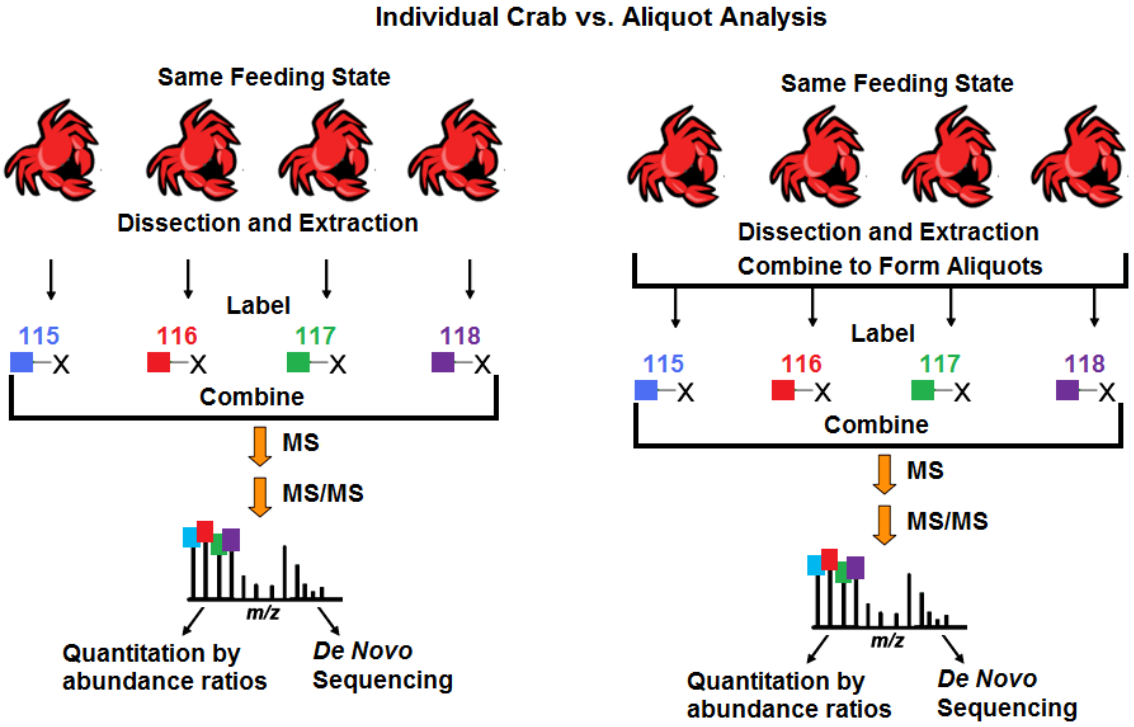
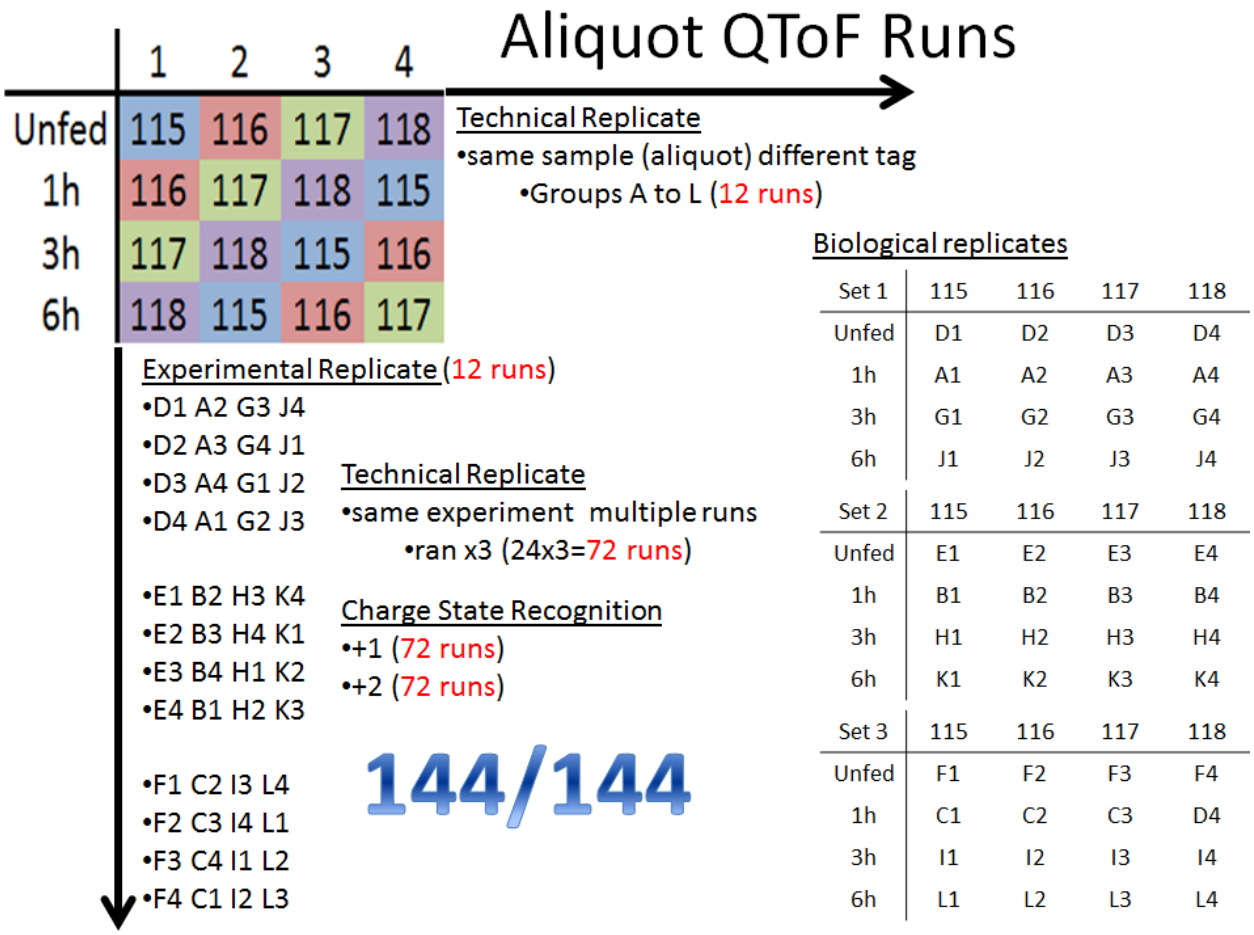


Figure 4.4 Aliquot QTOF Runs

This figure represents the application of the aliquots with the DiLeu labeling matrix resulting in 144 chromatograms collected by the analysis. The labels given to the samples are arbitrary. Letter A to L represent the three biological replicates of each of the four time points while 1-4 indicates the aliquot within each group.



Set 1	115	116	117	118
Unfed	D1	D2	D3	D4
1h	A1	A2	A3	A4
3h	G1	G2	G3	G4
6h	J1	J2	J3	J4

Set 2	115	116	117	118
Unfed	E1	E2	E3	E4
1h	B1	B2	B3	B4
3h	H1	H2	H3	H4
6h	K1	K2	K3	K4

Set 3	115	116	117	118
Unfed	F1	F2	F3	F4
1h	C1	C2	C3	D4
3h	I1	I2	I3	I4
6h	L1	L2	L3	L4

Figure 4.5 Identification of the Five Doubly Charged and Labeled Neuropeptides

Identification of the neuropeptides was initially calculated by peak matching and confirmed by tandem mass *de novo* sequencing. The relative peak heights of the DiLeu labels for five labeled neuropeptides in the (A) aliquot and (B) individually extracted unfed crab groups. The error bars represent 20% error for the DiLeu tags.

m/z	labeled M+H	M+H	Peptide Sequence	Peptide Family
595.30	1189.60	1044.60		
555.80	1110.60	965.54	NRNFLRFa	RFamide
535.82	1070.64	925.49	SRPYSFGLa	A-Type Allatostatin
543.30	1085.60	940.5	QRAYSFGLa	A-Type Allatostatin
719.86	1438.72	1293.63	STNWSSLRSAWa	B-Type Allatostatin

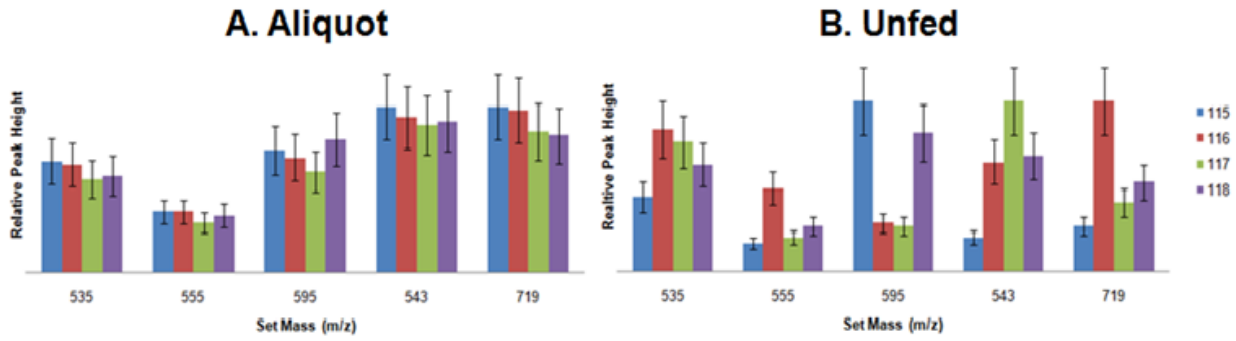


Figure 4.6 Relative-quantitative Analysis of Crustacean Feeding

The ratio of unfed verses fed peaks for the five labeled neuropeptides for the three biological replicates arbitrarily labeled DAGJ, EBHK, and FCIL. CV is coefficient of variance for the DiLeu tags.

595	CV	1hr/unfed	3hr/unfed	5hr/unfed
DAGJ	40.38	1.8	1.54	1.78
EBHK	45.91	1.17	0.98	1.32
FCIL	34.49	1.13	1.3	1.54
average	40.27	1.39	1.29	1.56

535	CV	1hr/unfed	3hr/unfed	5hr/unfed
DAGJ	28.78	1.84	1.38	1.49
EBHK	41.57	0.79	0.96	0.77
FCIL	40.28	0.63	0.6	0.6
average	37.04	1.08	0.98	0.94

555	CV	1hr/unfed	3hr/unfed	5hr/unfed
DAGJ	28.12	1.59	1.35	1.43
EBHK	55.83	2.23	1.85	2.77
FCIL	28.28	1.25	1.03	1.14
average	37.95	1.7	1.42	1.63

543	CV	1hr/unfed	3hr/unfed	5hr/unfed
DAGJ	37.11	2.43	1.98	2.06
EBHK	24.68	1.33	1.22	1.37
FCIL	18.51	1.32	1.34	1.25
average	27.18	1.72	1.53	1.58

719	CV	1hr/unfed	3hr/unfed	5hr/unfed
DAGJ	35.71	2.19	1.94	1.59
EBHK	55.66	1.94	1.33	1.91
FCIL	29.35	1.44	1.46	1.39
average	40.9	1.85	1.56	1.64

Figure 4.7 Mascot Label Comparisons

The labeling for a biological set and the resulting Mascot quantitation results are demonstrated. Since there is a rigid assignment of the m/z 115 as the control peak, all for experimental replicates must be present for ratio comparison. The inverse ratios are used for comparison resulting in comparison of two instead of four experimental replicates.

DAGJ	115	116	117	118	116/115	117/115	118/115
D1A2G3J4	D	A	G	J	A/D	G/D	J/D
D2A3G4J1	J	D	A	G	D/J	A/J	G/J
D3A4G1J2	G	J	D	A	J/G	D/G	A/G
D4A1G2J3	A	G	J	D	G/A	J/A	D/A

Figure 4.8 Ratio of Fed to Unfed Comparison

A comparison of the ratios of fed to unfed for three software to the formaldehyde labeling results (FH2 Study) and the raw data (RAW) for the two peaks identified in all five methods.

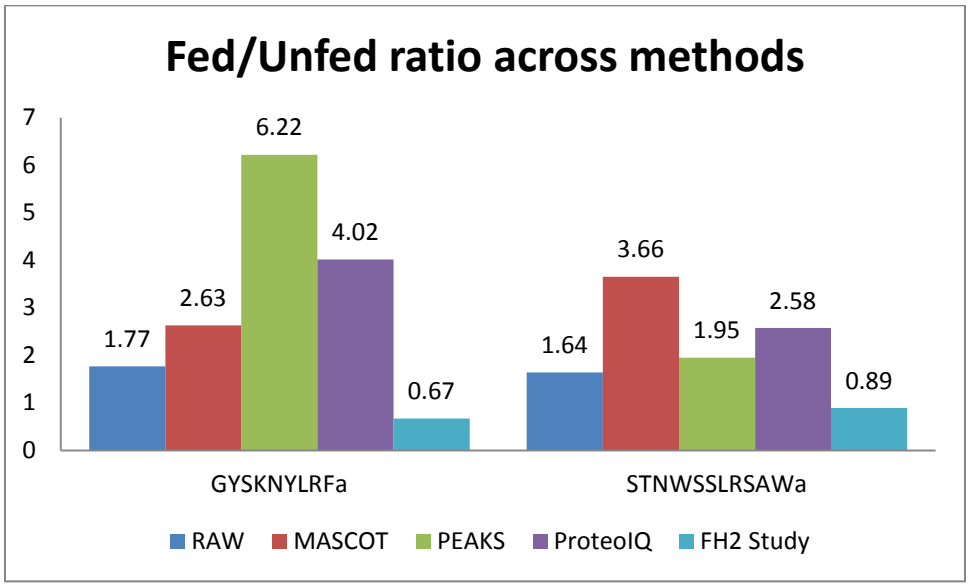


Figure 4.9 Software Comparisons of Identified Neuropeptide

The overlap of the identified neuropeptides for the three software is demonstrated. Mascot identified a total of 25 neuropeptides, PEAKS identified 47 neuropeptides and ProteoIQ identified 7.

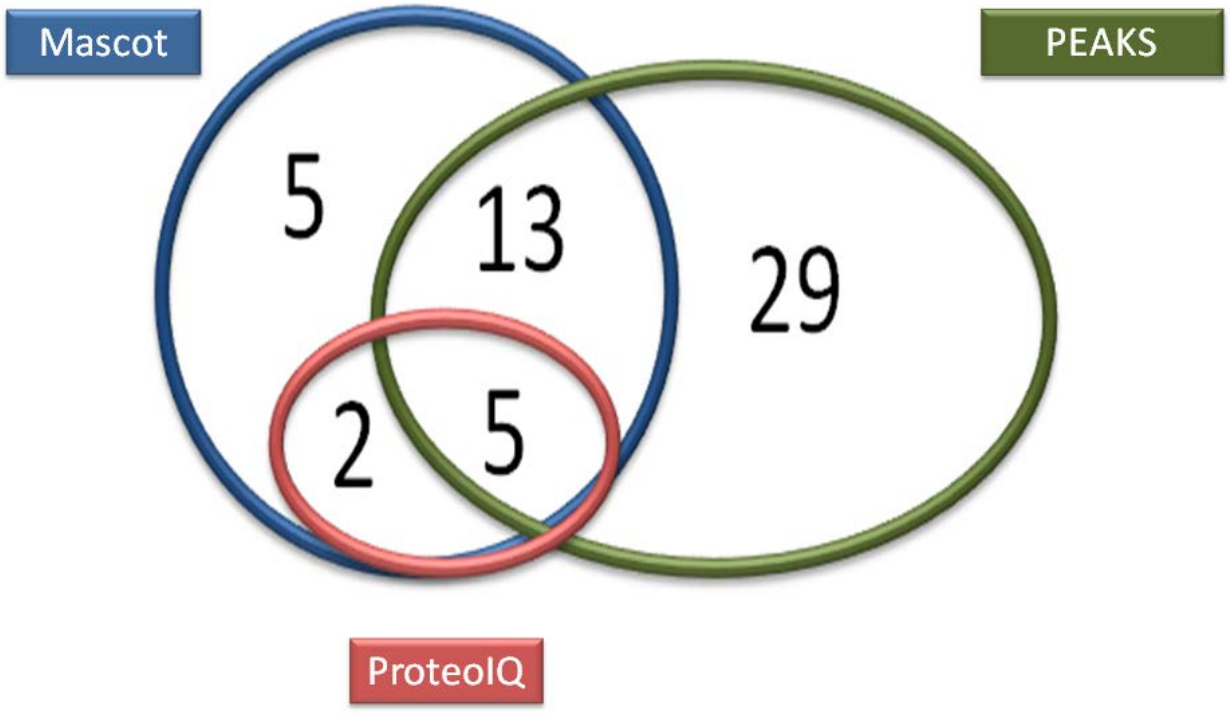


Figure 4.10 Forced Protein Matches

Six members of the Orcokinin family are displayed below. Given their sequence homology, if all six are identified are in a sample, hierarchical clustering of peptide hits will result in the identification of [Ala13]-orcokinin as a "protein" and the other five will be considered its peptide hits. The sequence coverage of each peptide in relation to the protein is also calculated.

Family	Name	[M+H]	Sequence	**Sequence Coverage**
Orcokinin	[Ala13]-orcokinin	1474.65973	NFDEIDRSGFGFA	
	Orcokinin[1-12]	1403.62262	NFDEIDRSGFGF	92.31%
		1360.61680	FDEIDRSGFGFA	92.31%
		1213.54839	DEIDRSGFGFA	84.615%
		1098.52145	EIDRSGFGFA	76.92%
		1027.48430	EIDRSGFGF	69.23%

Table 4.1 Repetitive Technical Replicate Peaks

The total number of peaks determined as repetitive by the number of QTOF runs that are present for each experimental replicate performed. Single identifies peaks only seen in one run. Double are the peaks seen in two runs. Triple are peaks seen in all three QTOF runs. The total is the number of unique (non-repetitive peaks) for each experimental replicate for the time point analysis. The final three columns are the percentages.

TP Analysis	Single	Double	Triple	Total	Single	Double	Triple
D1A2G3J4	112	56	207	375	29.87	14.93	55.20
D2A3G4J1	68	74	186	328	20.73	22.56	56.71
D3A4G1J2	54	86	249	389	13.88	22.11	64.01
D4A1G2J3	71	106	231	408	17.40	25.98	56.62
E1B2H3K4	74	62	189	325	22.77	19.08	58.15
E2B3H4K1	53	78	273	404	13.12	19.31	67.57
E3B4H1K2	62	108	234	404	15.35	26.73	57.92
E4B1H2K3	55	80	270	405	13.58	19.75	66.67
F1C2I3L4	58	84	261	403	14.39	20.84	64.76
F2C3I4L1	48	64	282	394	12.18	16.24	71.57
F3C4I1L2	60	72	171	303	19.80	23.76	56.44
F4C1I2L3	69	78	258	405	17.04	19.26	63.70
	784	948	2811	4543	17.26	20.87	61.88

Table 4.2 Repetitive Experimental Replicate Peaks

The total number of peaks determined as repetitive by the four experimental replicates performed. The first triple corresponds to Table 4.1 indicating the number of peaks detected in all three QTOF runs. Single identifies peaks only seen in one experimental replicate. Double are peaks seen in two experimental replicates. Triple are peaks seen in three experimental replicates. Quadruple are peaks seen in all four experimental replicates. The total is the number of unique (non-repetitive peaks) for each experimental replicate for the time point analysis. The two tailed student t-test results are displayed on the right for each biological replicate by the location of the unfed control label. The location of the tag doesn't affect the number of peaks detected.

	Triple	Single	Double	Triple	Quadruple	Total
D1A2G3J4	207	3	8	24	34	69
D2A3G4J1	186	4	5	19	34	62
D3A4G1J2	249	9	15	25	34	83
D4A1G2J3	231	9	12	22	34	77
E1B2H3K4	189	5	5	21	33	63
E2B3H4K1	273	14	14	31	33	91
E3B4H1K2	234	4	15	27	33	78
E4B1H2K3	270	9	14	35	33	90
F1C2I3L4	261	8	15	22	42	87
F2C3I4L1	282	14	13	25	42	94
F3C4I1L2	171	0	8	7	42	57
F4C1I2L3	258	8	12	24	42	86
	2811	87	136	282	432	937

D	115	116	117	118
115				
116	0.2933			
117	0.1402	0.0837		
118	0.3534	0.0868	0.1816901	

E	115	116	117	118
115				
116	0.0584			
117	0.2441	0.2838		
118	0.1128	0.9007	0.2522155	

F	115	116	117	118
115				
116	0.391			
117	0.0922	0.1099		
118	0.824	0.2362	0.1402346	

Table 4.3 Repetitive Biological Replicate Peaks

The total number of peaks determined as repetitive by the three biological replicates performed. Quadruple are peaks previously identified in Table 4.2 indicating the number of peaks detected in all four experimental replicates. Single identifies peaks only seen in one biological replicate. Double are the peaks seen in two biological replicates. Triple are peaks seen in all three biological replicates. The standard deviation and average of the peaks detected per biological replicate are given. The two tailed student t-test results are displayed for each biological replicate by the location of the unfed control label (right) and between the biological replicates (below).

Tab	Quad	Single	Double	Triple	STDev	average
DAGJ	34	5	15	14	5.507	11.333
EBHK	32	5	14	14	5.196	11
FCIL	42	13	15	14	1	14
	108	23	44	42		

	D	E	F
D			
E	0.942894		
F	0.4913745	0.4235	

115	D	E	F
D			
E	0.9034		
F	0.6759	0.5864	

116	D	E	F
D			
E	0.428		
F	0.4436	0.9552	

117	D	E	F
D			
E	0.9101		
F	0.5782	0.6491	

118	D	E	F
D			
E	0.701		
F	0.8212	0.9055	

All	D	E	F
D			
E	0.5983949		
F	0.6603372	0.9787309	

Table 4.4 Raw Data 10 Match

The top 10 peaks identified from the top 14 peaks detected by manual results. The peaks were initially mass matched to the 692 entry database, and the 4 sequences that were previously identified in the individual crab analysis were verified by tandem MS *de novo* sequencing and are highlighted in blue. The orange identifies the peptide identified from updating the database and the red identifies either an incorrect match given the pQ N-terminus or abundant but unlabeled match.

Family	Name	(M+H)	Sequence
Allatostatin A-type	Lepidopteran peptide cydiastatin 3,helicostatin 3	925.4890	SRPYSFGLa
	Penaeustatin 32	957.4789	DRTYSFGLa
Allatostatin B-type	CbAST-B8	1293.63346	STNWSSLRSAWa
CPRP		989.46870	GFLSQDVHS
RFamide		1005.57409	GPRNFLRFa
		1019.58974	APRNFLRFa
		1022.56462	GNRNFLRFa
	NPY/PP peptide pem-PYF 3	1049.60031	YAIAGRPRFa
RYamide		976.46354	SGFYANRYa
		1030.47411	pQGFYSQRYa

Table 4.8 Mascot Top 18

The top 18 neuropeptides detected by MASCOT search using the Deamon default settings. Theses peptides were compared to the raw data. The peptides identified in all four experimental replicates are indicated by a yellow box. The peptides that are also identified in the formaldehyde study are indicated in blue and red.

			Time Points					
			1234	2341	3412	4123		
Allatostatin B-type	LL160	TSWGKFQGSWa		1			1	No comparison
	LL164	GNWNKFQGSWa		1			1	Same ratios
	LL166	NNWSGAFKGSWa		1		1	2	Different ratios
	LL170	SGKWSNLRGAWa	1				1	50:50same:differ
	LL174	STNWSSLRSAWa	1	1	1	1	4	≤1/4 differ
	LL175	STDWSSLRSAWa	1	1		1	3	
CCAP	LL200	PFCNAFTGCa	1				1	In all 4 TPs
Orcokinin	LL447	NFDEIDRSGFGFA	1	1	1	1	4	
Orcomyotropin	LL465	FDAFTTGFGHS	1	1			2	FH2 Study brain
RFamide	LL543	APRNFLRFa	1	1	1	1	4	FH2Study PO
	LL561	TNYGGFLRFa			1		1	
	LL567	GAHKNYLRFa		1		1	2	
	LL588	GYSKNYLRFa	1	1	1	1	4	
	LL591	APQRNFLRFa	1				1	
	LL600	AYNRSFLRFa				1	1	
	LL605	SENRNFLRFa		1			1	
RYamide	LL659	SGFYANRYa	1				1	
Tachykinin	WF_NPDB0417	APSGFLGMRFa		1	1	1	3	37
7 families	18 neuropeptides		10	12	6	9	37	
						37		

Table 4.6 Mascot Top 39 Peptides

The top proteins identified by Mascot parameter comparisons are listed below. The peptides that were previous identified by the formaldehyde study are also indicated. For the formaldehyde study, neuropeptides identified in both PO and brain samples are highlighted in purple.

Family		Sequence
Allatostatin A-type	LL095	pQRAYSFGLa
	LL097	SRPYSFGLa
Allatostatin B-type	LL155	AGWSSMRGAWa
	LL160	TSWGKFQGSWa
	LL164	GNWNKFQGSWa
	LL166	NNWSGAFKGSWa
	LL170	SGKWSNLRGAWa
	LL174	STNWSSLRSAWa
	LL175	STDWSSLRSAWa
	LL176	NNNWSKFQGSWa
CCAP	LL200	PFCNAFTGCa
CHH	LL203	
Orcokinin	LL447	NFDEIDRSGFGFA
Orcomyotropin	LL465	FDAFTTGFGHS
RFamide	LL532	pQRNFLRFa
	LL543	APRNFLRFa
	LL544	GNRNFLRFa
	LL561	TNYGGFLRFa
	LL567	GAHKNYLRFa
	LL588	GYSKNYLRFa
	LL591	APQRNFLRFa
	LL600	AYNRSFLRFa
	LL605	SENRNFLRFa
RYamide	LL659	SGFYANRYa
Tachykinin	WF_NPDB0417	APSGFLGMRFa

MASCOT 39 NPs

FH2 Study brain
FH2 Study PO
Both brain and PO

14 other species (fly, worm)
25 crustacean

14 Questionable Matches	
WF_NPDB0817	Procambarus Clarkii, (QB14) actin 1000513
WF_NPDB0968	Aplysia californica, (TK17) Luqin-C
WF_NPDB1037	Aplysia californica, (TK86) atrial amino-terminal peptide
WF_NPDB1165	Ascaris suum , (WC15) Rfamide
WF_NPDB1166	Ascaris suum , (WC16) Rfamide
WF_NPDB1420	Drosophila, (YB48) PTTH
WF_NPDB1518	Anopheles gambiae, (YJ22) EH
WF_NPDB1639	Cydia pomonella, (YR01) actin-3
WF_NPDB1642	Cydia pomonella, (YR01) actin
WF_NPDB1643	,
WF_NPDB1674	Pachygrapsus marmoratus, (OC02) CHH A OX
WF_NPDB1906	AST, (AD26) Schistocerca gregaria-PR-1
WF_NPDB1955	AST , (AD28) ALLP_HELAM Helicostatins-PR-5
WF_NPDB2113	Macrobrachium rosenbergii, (QH05) SG peptide A precursor-2

Table 4.7 Mascot Most Prevalent Peptides

The eleven peptides detected in both the time point (TP) and parameter (para) comparisons for the top Mascot hits for all parameters analyzed are identified below. The five neuropeptides found in both comparisons are indicated. Peptides identified in the former formaldehyde study are highlighted.

Family		Sequence	TPs	para
Allatostatin B-type	LL155	AGWSSMRGAWa	0	1
	LL164	GNWNKFQGSWa	1	0
	LL174	STNWSSLRSAWa	1	1
	LL160	TSWGKFQGSWa	1	1
Orcokinin	LL447	NFDEIDRSGFGFA	1	1
Orcomyotropin	LL465	FDAFTTGFGHS	0	1
RFamide	LL543	APRNFLRFa	1	0
	LL567	GAHKNYLRFa	1	1
	LL588	GYSKNYLRFa	1	1
	WF_NPDB1166	SDMPGVLRFa	1	0
Tachykinin	WF_NPDB0817	APSGFLGMRFa	1	0
			9	7

FH2 Study brain
FH2 StudyPO

Overlapping NPs

Table 4.8 PEAKS Top hits matched to Mascot top hits

The 18 peptides identified by PEAKS that overlap with Mascot identification. Those quantified by Mascot are indicated by orange. Those peptides that match the previous formaldehyde labeling results for the brain (red), PO (blue), or both (purple) are identified.

Family	Name	(M+H)	Sequence
AST A-type	Lepidopteran peptide cydiastatin 3,helicostatin 3	925.4890	SRPYSFGLa
AST B-type		1107.51530	AGWSSMRGAWa
	CbAST-B5	1182.56907	TSWGKFQGSWa
	CbAST-B6	1222.57522	GNWNKFQGSWa
		1252.58580	NNWSGAFKGSWa
	CbAST-B8	1293.63346	STNWSSLRSAWa
	CbAST-B9	1366.62871	NNNWSKFQGSWa
Orcokinin	[Ala13]-orcokinin	1474.65973	NFDEIDRSGFGFA
Orcomyotropin		1186.51636	FDAFTTGFGHS
RFamide		1019.58974	APRNFLRFa
		1022.56462	GNRNFLRFa
		1073.55270	TNYGGFLRFa
		1104.60612	GAHKNYLRFa
		1146.60545	GYSKNYLRFa
	PrcFaRP 6	1147.64832	APQRNFLRFa
	[ala1]-FaRP	1172.63234	AYNRSFLRFa
		1181.61742	SENRNFLRFa
RYamide		976.46354	SGFYANRYa
6 families			18NPs

Table 4.9 Top ProteoIQ Matches

The top seven neuropeptides identified by ProteoIQ across all three parameters (200ppm, 100ppm, and 50ppm). The peptides previously quantified in Mascot are indicated in orange. Those peptides that match the previous formaldehyde labeling results for the brain (red), PO (blue), or both (purple) are identified.

Family	Name	(M+H)	Sequence
Allatostatin B-type	CbAST-B8	1107.51530	AGWSSMRGAWa
		1293.63346	STNWSSLRSAWa
		1294.61750	STDWSSLRSAWa
Orcokinin	[Ala13]-orcokinin	1474.65973	NFDEIDRSGFGFA
Orcomyotropin		1186.51636	FDAFTTGFGHS
RFamide		1146.60545	GYSKNYLRFa
Tachykinin		950.48765	APSGFLGM(O)Ra

Chapter 5 Generation of Multiply Charged Ions in MALDI Based Analysis

5.1 Abstract

The matrix-assisted laser desorption/ionization (MALDI) process is currently known for primarily producing singly charged ions. However, multiply charged ions are capable of providing enhanced fragmentation efficiency and better tandem mass spectral data for identification. This chapter describes the pursuit of creating multiply charged ions in MALDI by simple sample preparation and plating techniques and their resulting fragmentation spectra.

5.2 Introduction

In biological mass spectrometry, soft ionization techniques have created a platform for analyzing complex biological mixtures with high speed, high sensitivity and great specificity making it an ideal tool for peptide analysis [1, 2]. Electrospray ionization (ESI) coupled to liquid chromatography (LC) is commonly used for shotgun proteomics experiments for discovery and identification because of the fast separation and reproducible results. Also, the quality tandem mass spectra produced by multiple charged ions are primarily found with ESI thereby adding to its appeal [3]. The other soft ionization technique, matrix-assisted laser desorption/ionization (MALDI), is finding popularity in histological studies with its imaging and direct tissue capabilities [4-6]. However, the predominantly singly charged molecules produced in the MALDI ionization process limit the identification of peptides primarily to mass matching [3, 7-10]. It is therefore of great interest to manipulate the charge state of the ions produced in MALDI to improve peptide fragmentation and subsequent sequencing for identification [11].

The predominant charge species generated in MALDI analysis is the singly charged ion [7]. However, the first MALDI experiments not only reported the existence of multiply charged

ions but used them as validation that the protocol used was acceptable for large molecule analysis [4]. It is unknown when the prevalent thought of associating MALDI analysis solely with the production of singly charged ions became popular [7].

Since the actual mechanism for ion formation in MALDI is still unknown, there are several prevailing theories on why singly charged species dominate the spectra [12]. It is speculated that the laser desorption ionization process initially releases multiply charged species from the sample that are reduced in plume before even reaching analyzer [10]. Charge reduction of the multiply charged species may occur by matrix clusters, free electrons from the metal MALDI plate, or as charge transfers just to name a few theories [8, 10, 11, 13].

A simple method for creating multiply charged ion states of a protein or peptide in MALDI employed matrix manipulation [14]. The matrix choice, matrix solution, matrix to analyte ratios, crystallization conditions, and sample deposition method are all contributing factors in creating these multiply charged ions [15].

The detection of multiply charged ions in MALDI has been primarily coupled with time of flight (TOF) mass spectrometers [15, 16]. Manipulation of the MALDI instrument, in terms of breaking the vacuum (AP-MALDI) or incorporating other ESI-like features into the system have successfully resulted in the production of multiply charged ions [11, 17]. By developing a methodology that creates multiply charged ions using MALDI with no instrument modifications allows the method to be easily incorporated into future research designs involving any MALDI instrument. Also, such a design will further current efforts in understanding the exact mechanism behind ion formation in the MALDI process; elucidating matrix-analyte interactions

upon initial ionization by the laser and secondary reactions that occur in the MALDI plume resulting in predominately singly charged ions [7, 13].

The goal is to obtain a mass spectrum where multiply charged ions are detected with great abundance by the MALDI process. A relative abundance of 100 for intensity would be ideal. It is believed that once multiply charged ions are detected with great abundance, fragmentation of the multiply charged ions will produce enhanced tandem mass spectrum for improved *de novo* sequencing and thus improved peptide identification.

5.3 Materials and Methods

Matrices

2',4',6'-trihydroxyacetophenone monohydrate (THAP), α-cyano-4-hydroxycinnamic acid (CHCA), sinapic acid (SA), 2,5-dihydroxybenzoic acid (DHB), 3-nitrobenzonitrile (NBN), nicotinic acid (NA), 2-nitrophloroglucinol (NPG), 2,5-dihydroxyacetophenone (DHAP), benzoic acid (BA), and hydroquinone (H).

Solvents

Deionized water, methanol (MeOH), formic acid (FA), 2-propanol (IPA), acetonitrile (ACN), trifluoroacetic acid (TFA).

Peptides

All peptide standards were obtained from American Peptide Company. Bradykinin was predominately used for analysis. The other peptides were used with the Orbitrap.

Instruments

Five MALDI instruments were utilized in this experiment. One FTICR, three MALDI TOF/TOF instruments and the MALDI orbitrap were used. The Micromass QTOF was also used

to obtain an ESI spectrum of bradykinin by direct infusion. All instruments parameters are listed in Appendix A.

Methods

Dry Droplet Plating Method

Equal volume standard and matrix were premixed before plating 0.5-1μL of the mixture. This was done to ensure even mixture of the analyte and matrix for analysis. All spots were left to dry, crystalizing, at room temperature before analysis.

Ultra-thin Layer Method

A modified ultra-thin layer (UTL) experiment was conducted where a dilute sample of the matrix was plated onto the metal MALDI plate as the 1st layer and allowed to dry before plating the analyte and matrix at either a 1:1 or 1:4 by volume mixture [15]. The 1st layer of diluted matrix is prepared as follows: The regular matrix is diluted 4-fold in the organic solute used in its original matrix solution. For example the matrix solution used for CHCA was 1:2 water: acetonitrile so it was diluted in acetonitrile and the matrix solution used for DHB is 50% methanol so it is diluted in methanol. Then it is further diluted in a 1:4 or 1:9 mix with 25% isopropanol in water. A 0.2μL aliquot of 1st layer is spotted on the MALDI plate and allowed to dry. The top layer is a 0.5μL aliquot mixture of analyte and matrix generated by the dry droplet method.

DHB Crystal Formation

Three saturated DHB (saDHB) (150mg DHB in 50% methanol) mixtures were made. One was placed in the 4°C refrigerator (dark cold temperature conditions), another in a bench top

drawer away from light (dark room temperature conditions), and the last left on the bench top (light room temperature conditions) for 42 days to allow crystals to grow in solution.

Matrix Solution Analysis

Initially, THAP, CHCA, and SA were freshly prepared in one of the following four solutions: (1) water/methanol at 2:1 ratio by volume (2) formic acid/water/2-propanol at 1:3:2 ratio by volume (3) water/acetonitrile at 1:1 ratio by volume and (4) trifluoroacetic acid/water/acetonitrile at 1:4.5:4.5 ratio by volume. They were analyzed using the dry droplet method with a 1μg/μL concentration sample of bradykinin. Later, DHB was analyzed using those four solutions as well.

NA was analyzed in five solutions. A 10^{-3}M NA solution in water, 50% methanol, solutions 2 and 4 listed above, and 0.1% FA in 50% ACN with water solution.

Binary Matrix Analysis

Binary matrix analysis was performed in two ways. The first was by UTL analysis. The matrix used as the 1st layer is different from the analyte matrix layer. The second was the addition of NBN directly to either the saDHB matrix or the bradykinin standard before plating using the dry droplet method.

Electron Free Plating Material Analysis

Samples were plated on metal MALDI plates for analysis. A modified electron free protocol was performed by wrapping the metal plate with Parafilm M before spotting using the dry droplet method [13]. Also, glass slipcovers were taped above the Parafilm M for electron free analysis in the FTICR as well.

5.4 Results and Discussion

All data is collected for the bradykinin peptide. It was chosen because of its availability and its m/z at 1060 which is the middle mass range of typical peptide analysis. It was theorized that if multiple charges could be created with bradykinin then they could be created with the other peptide masses as well.

All initial results were obtained on the MALDI FTICR. It was my goal to produce multiply charged ions on this instrument since the majority of multiply charged ion work focused on the MALDI TOF. Given that we normally used saDHB with our MALDI analysis, I designated saDHB as the baseline. A protocol is considered successful if it produced better results than those obtained by saDHB analysis (Figure 5.1).

With FTICR, the dominant multiply charged species was the +3 peak (Figure 5.1). The deconvolution setting on the FTICR had to be turned off to see multiple charges. Later, due to instrument problems, analysis was performed on the MALDI TOF instruments. With the TOF, the dominant multiply charged species detected is the +2 peak. ESI is considered the golden scale for production of multiple charges. The spectrum obtained by ESI is generally considered the best representation of multiply charged species that exist in a sample. For bradykinin, the dominant peak in the ESI spectrum was the +2 peak (data not shown but performed as direct infusion on the Micromass).

I explored different matrices, solvent combinations, varying matrix to analyte ratios, different electron free plating materials, and various matrix manipulations with the goal of producing multiply charged ions. To begin optimization of these conditions I modified procedures from two papers that report detection of multiply charged protein and peptide ions by MALDI TOF analysis [14, 15]. Three matrices, THAP, CHCA, and SA were freshly prepared in

one of four matrix solutions and compared to saDHB using the FTICR. The results are shown in Table 5.1. Use of the matrices SA and CHCA resulted in highly fragmented mass spectra because they are classified as hot matrices [18].

Also displayed in Table 5.1, are the results for the electron free plating material analysis. All four matrices, saDHB and the three listed above in the four matrix solutions, were plated on both the Parafilm M and the glass covers. Only spots whose spectra displayed a signal were analyzed and are shown on the table. All were analyzed on the FTICR. The Bruker MALDI TOF/TOF instruments have a glass slide plate holder normally utilized for MALDI imaging. On a glass slide, saDHB was plated and analyzed on the AutoFlex but there was no signal detected. For both AutFlex and FTICR, plating on the glass proved difficult. The sample diffused across the plate before drying.

Given the bad spectrum acquired by plate alteration, alternative methods for plating were researched and the UTL method discovered [15]. The UTL method utilizes a thin layer of dilute matrix as a buffer between the sample and the plate. It also produces more evenly distributed crystals for analysis, which aids shot to shot reproducibility. Binary matrix application was also explored with the UTL method using different matrices for the 1st layer and matrix analyte mix.

The single matrix UTL method was performed with DHB, SA, THAP, and CHCA (see Figures 5.2-5.4). The UTL for DHB produced the best +3 peak abundances. The binary matrix UTL method was performed with DHB, DHAP, SA and NPG (Figures 5.5-5.8). DHAP and NPG were incorporated into the UTL studies to determine how successful they are in producing multiply charged ions in MALDI FTICR [16]. Again, the best +3 peak abundances were found with DHB analysis.

When analyzed independently, NPG produced the fragmentation peak m/z 904 that emerged as the dominant peak. This fragment is thought to occur from the cleavage of an arginine. NPG may be considered a hot matrix like CHCA and SA; however their spectra never resulted in a dominant fragmentation peak. Why this fragmentation is the dominant peak is unknown, however arginine is a highly basic residue that sequesters the proton preventing the charge from transferring [19, 20]. It was also determined that the ratio of the +1/+2 charges for the fragmented peptide was similar to that of the full length peptide. The potential location of the second protonation site is one of the three imino acid prolines [20].

Originally assigned as the baseline, saDHB became the gold standard for producing multiply charged ions. To determine why, a structural and matrix solution analysis was performed. First the DHB was made in the first 4 solutions and compared to normal saDHB (Table 5.2). Normal saDHB proved the best technique. Next the concentration of methanol was tested (Table 5.3). The 75% methanol solution was the best. A structural analysis resulted in adding BA and H as matrices. But plating with these two matrices proved difficult and results not obtained until a tertiary UTL matrix analysis was performed (Figure 5.10). Again, regular saDHB provided the best results.

Normally, fresh DHB matrix is prepared on the day of analysis and the spot is analyzed within 24 hours of plating. In the midst of instrument complications, a plated sample was analyzed both day of and eight days after plating. This delay in analysis resulted in better spectra. The age of the DHB was analyzed to determine optimum multiple peak analysis. This analysis remains inconclusive.

Review of older literature revealed that for DHB experiments crystals were grown and plated for MALDI analysis[21] . Optimum crystal growth was explored in dark cold, dark room temperature, and light room temperature conditions. After 42 days, crystals only formed in the dark cold conditions. All solutions and crystals were plated for analysis. The most notable results came from binary matrix plating of the DHB crystal with the NA in water matrix. The NA matrix was added after it was revealed by literature search to be one of the first MALDI matrices, and it use resulted in the production of multiply charged ions [4]. The resulting experiment peak ratios were the best reported and the relative abundance of the +2 peak was by far the best of all experiments combined (Figure 5.11). As more shots were obtained per analysis, the ratio did not change but the relative abundance grew with each shot (Figure 5.11).

As instruments changed, and peak ratios remained relatively the same, the primary question became what is the required abundance for the +2 peak to produce adequate tandem MS fragmentation. Bradykinin was plated with saDHB and the +2 peak isolated for tandem fragmentation on the UltraFlexTreme. Although the peak was selected for fragmentation, nothing above m/z 530 was displayed on the spectrum (Figure 5.12). There are two theories for this anomaly. In the first there is a timing issue with the analyzer while the second is a deflection issue with the reflectron. Both the analyzer timing and deflector energies are calculated assuming that the species being isolated is singly charged. The +2 species is actually larger and has a higher energy than the +1 species, therefore they are slower (not reaching the detector before the analyzer turns off) and heavier (not deflecting in the reflectron) than the +1 species.

A previous study reported the improved fragmentation pattern generated from the multiply charged species of large proteins by MALDI TOF/TOF [15]. However, none of the

spectra shown in the publication were above the m/z for the multiply charged species. Also, in this study an Applied Biosystems instrument was used, and no mention was made of this cut-off phenomenon.

Given the recent data, in relation to the TOF/TOF cut-off and the FTICR deconvolution, it was apparent that MALDI software calculations based on singly charged ion formation were a limiting factor for analysis. With newer generation mass spectrometers having an interchangeable MALDI or ESI ionization source option, I hypothesized that such software assumptions may not be made on the Orbitrap and that fragmentation above the m/z 530 would be detected. Analysis of bradykinin plated with saDHB was performed on the Orbitrap and complete fragmentation of bradykinin was seen with both CID and HCD (Figure 5.12-5.13). Granted, this is also a comparison between an LTQ Orbitrap and a TOF/TOF, but again, I believe that the ability to see the complete fragmentation spectrum is the result of the interchangeable interface on the Orbitrap.

In conclusion, multiple charges are detected when a single peptide containing two or more basic sites is analyzed with a cold matrix in MALDI. Although not abundant, the multiply charged peaks are present. Hotter matrices produce fragmentation peaks in the normal mass spectrum without requiring tandem MS fragmentation analysis. Also, depending on the MALDI instrument, tandem mass spectrometry may be performed on the multiply charged peaks.

5.5 Future Focus

Future work should focus on the development of highly abundant multiply charged ions with emphasis on spot to spot and shot to shot reproducibility. A glass plate could be etched to aid in preventing the sample from spreading. Also coating with the saDHB matrix by TMsprayer

or growing saDHB crystals directly on the glass slide should be analyzed. A binary matrix application for both methods could be utilized with NA.

Since they have the added benefit of generating fragments in the MS mode, a more detailed analysis of the fragmentation patterns for hot matrices may result in a simpler alternative to improving peptide identification then pursing the formation of multiply charged ions.

5.6 Respective Contributions and Acknowledgements

I would like to thank Chris Lietz for his fruitful discussion and help with the MALDI Orbitrap analysis. Chris and I are both working on the production of multiply charged peptide ions in MALDI. He, however, is attacking the problem from a top down approach by first analyzing proteins and I directly and solely focused on peptides. Also Paul Kowalski from Bruker helped formulate the theories regarding the TOF/TOF cut-off.

5.7 Works Cited

1. Nilsson, C.L., et al., *Mass spectrometry of peptides in neuroscience.* Peptides, 1998. **19**(4): p. 781-9.
2. Jakubowski, J.A. and J.V. Sweedler, *Sequencing and mass profiling highly modified conotoxins using global reduction/alkylation followed by mass spectrometry.* Anal Chem, 2004. **76**(22): p. 6541-7.
3. He, M., J.F. Emory, and S.A. McLuckey, *Reagent anions for charge inversion of polypeptide/protein cations in the gas phase.* Anal Chem, 2005. **77**(10): p. 3173-82.
4. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Anal Chem, 1988. **60**(20): p. 2299-301.
5. Schwartz, S.A., M.L. Reyzer, and R.M. Caprioli, *Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation.* Journal of Mass Spectrometry, 2003. **38**(7): p. 699-708.
6. Stoeckli, M., et al., *Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues.* Nat Med, 2001. **7**(4): p. 493-6.
7. Karas, M., M. Gluckmann, and J. Schafer, *Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors.* J Mass Spectrom, 2000. **35**(1): p. 1-12.
8. Knochenmuss, R., et al., *Secondary ion-molecule reactions in matrix-assisted laser desorption/ionization.* J Mass Spectrom, 2000. **35**(11): p. 1237-45.

9. Burkitt, W.I., et al., *Discrimination Effects in MALDI-MS of Mixtures of Peptides—Analysis of the Proteome.* Australian Journal of Chemistry, 2003. **56**(5): p. 369-377.

10. Zenobi, R. and R. Knochenmuss, *Ion formation in MALDI mass spectrometry.* Mass Spectrometry Reviews, 1998. **17**(5): p. 337-366.

11. He, M. and S.A. McLuckey, *Two ion/ion charge inversion steps to form a doubly protonated peptide from a singly protonated peptide in the gas phase.* J Am Chem Soc, 2003. **125**(26): p. 7756-7.

12. Lewis, J.K., J. Wei, and G. Siuzdak, *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis*, in *Encyclopedia of Analytical Chemistry*. 2006, John Wiley & Sons, Ltd.

13. Frankevich, V., et al., *Production and fragmentation of multiply charged ions in 'electron-free' matrix-assisted laser desorption/ionization.* Rapid Commun Mass Spectrom, 2003. **17**(20): p. 2343-8.

14. Cohen, S.L. and B.T. Chait, *Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins.* Anal Chem, 1996. **68**(1): p. 31-7.

15. Liu, Z. and K.L. Schey, *Fragmentation of multiply-charged intact protein ions using MALDI TOF-TOF mass spectrometry.* J Am Soc Mass Spectrom, 2008. **19**(2): p. 231-8.

16. Trimpin, S., et al., *Extending the Laserspray Ionization Concept to Produce Highly Charged Ions at High Vacuum on a Time-of-Flight Mass Analyzer.* Analytical chemistry, 2011. **83**(14): p. 5469-75.

17. McEwen, C.N., B.S. Larsen, and S. Trimpin, *Laserspray ionization on a commercial atmospheric pressure-MALDI mass spectrometer ion source: selecting singly or multiply charged ions.* Anal Chem, 2010. **82**(12): p. 4998-5001.

18. Shahgholi, M., et al., *Sugar additives for MALDI matrices improve signal allowing the smallest nucleotide change (A:T) in a DNA sequence to be resolved.* Nucleic Acids Res, 2001. **29**(19): p. E91.

19. Paizs, B. and S. Suhai, *Fragmentation pathways of protonated peptides.* Mass Spectrom Rev, 2005. **24**(4): p. 508-48.

20. Ewing, N.P., et al., *Gas-phase basicities for ions from bradykinin and its des-arginine analogues.* J Mass Spectrom, 2001. **36**(8): p. 875-81.

21. Westman, A., et al., *Sample morphology effects in matrix-assisted laser desorption/ionization mass spectrometry of proteins.* Journal of Mass Spectrometry, 1995. **30**(1): p. 206-211.

5.8 **Figures and Tables**

Figure 5.1Bradykinin Plated with saDHB

The baseline spectrum produced by dry droplet analysis of bradykinin with saDHB on the MALDI FTICR.

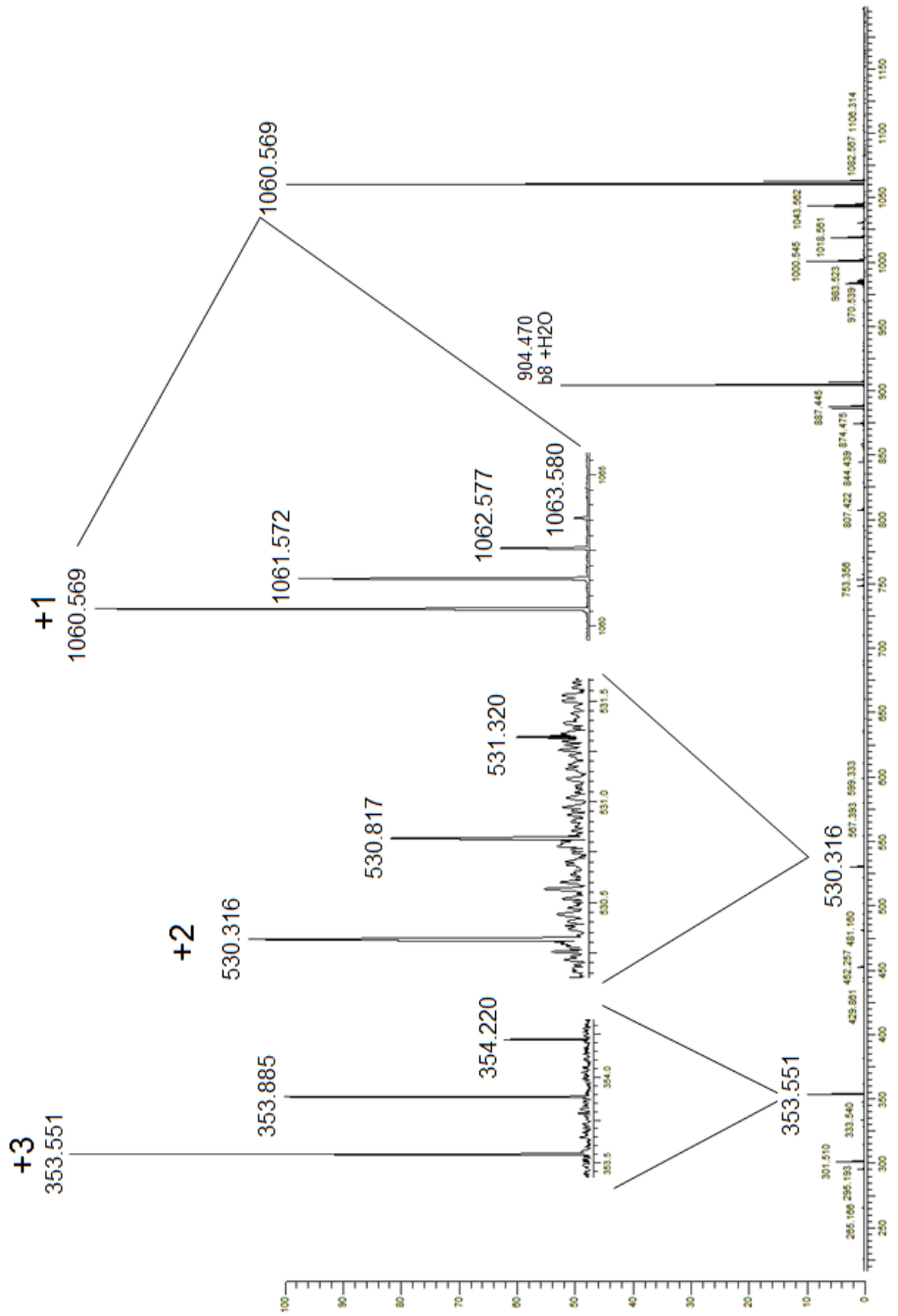
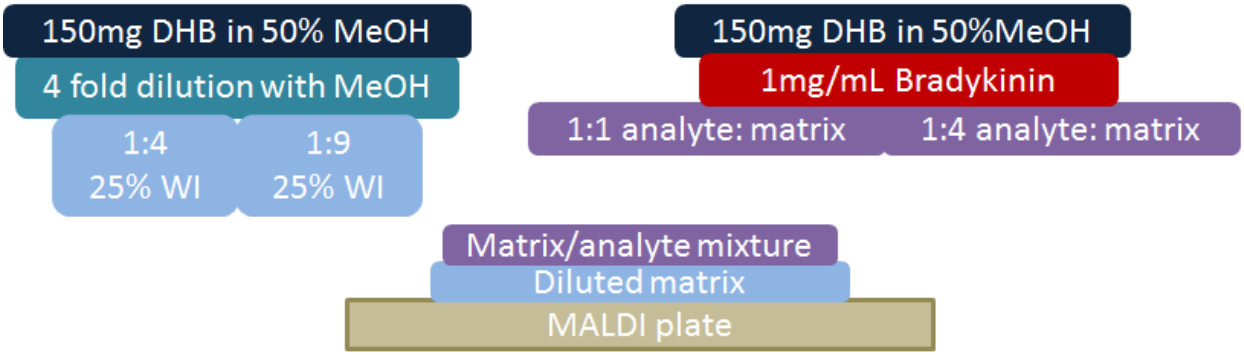


Figure 5.2 UTL results for DHB

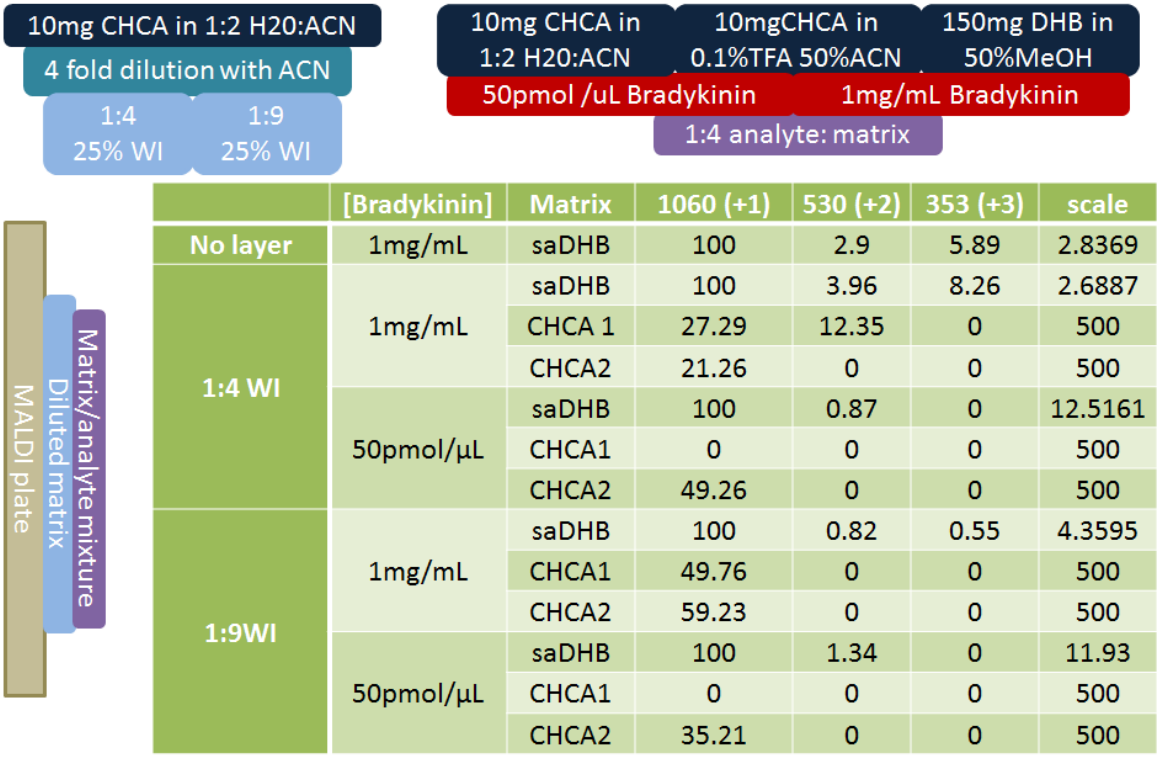
The formation of the 1st layer and the matrix and analyte mixes are displayed with the relative abundance for each peak and scale for each spectrum.



	Brady: saDHB	1060 (+1)	530 (+2)	353 (+3)	scale
No layer	1:1	100	3.002	6.06	3.0331
1:4 WI	1:1	100	4.39	9.96	2.549
	1:4	100	4.74	11.05	2.4852
1:9 WI	1:1	100	1.8	3.42	3.5169
	1:4	100	5.27	13.05	2.351

Figure 5.3 UTL results for CHCA.

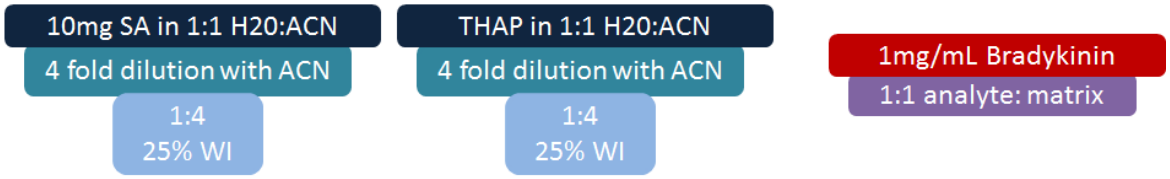
The formation of the 1st layer and the three matrices used for the matrix analyte mixes for the two bradykinin concentrations are displayed with the relative abundance for each peak and scale for each spectrum.



	[Bradykinin]	Matrix	1060 (+1)	530 (+2)	353 (+3)	scale
No layer	1mg/mL	saDHB	100	2.9	5.89	2.8369
1:4 WI	1mg/mL	saDHB	100	3.96	8.26	2.6887
		CHCA 1	27.29	12.35	0	500
		CHCA2	21.26	0	0	500
	50pmol/µL	saDHB	100	0.87	0	12.5161
		CHCA1	0	0	0	500
		CHCA2	49.26	0	0	500
1:9WI	1mg/mL	saDHB	100	0.82	0.55	4.3595
		CHCA1	49.76	0	0	500
		CHCA2	59.23	0	0	500
	50pmol/µL	saDHB	100	1.34	0	11.93
		CHCA1	0	0	0	500
		CHCA2	35.21	0	0	500

Figure 5.4 UTL results for THAP and SA.

The formation of the 1st layer and the matrix and analyte mix are displayed with the relative abundance for each peak and scale for each spectrum.

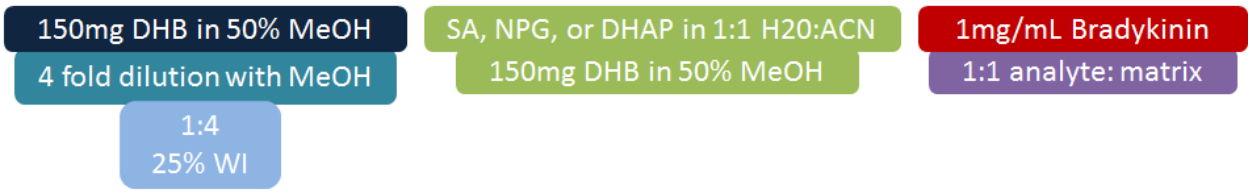


MALDI plate / Diluted matrix / Matrix/analyte mixture

	[Bradykinin]	1060 (+1)	530 (+2)	353 (+3)	scale
THAP	No layer	100	4.946	8.147	2.5747
		100	2.886	4.883	3.0169
		100	3.445	4.779	2.6651
	UTL	100	2.783	5.790	3.2390
		100	1.955	7.129	4.0717
		100	2.657	6.477	3.5987
SA	No layer	53.725	1.321	5.543	45.1183
		72.625	1.698	7.126	15.0565
		51.941	1.331	5.380	18.0743
	UTL	64.119	1.542	6.452	30.9257
		70.427	1.639	6.580	14.9701
		57.761	1.472	5.580	18.4403

Figure 5.5 Binary UTL results for DHB as 1st layer

The binary matrix application using the UTL method with DHB as the 1st layer and DHB, SA, NPG, or DHAP for the matrix analyte mixture is displayed. The formation of the 1st layer and the four matrices used for the matrix analyte mixes for the bradykinin concentration are displayed with the relative abundance for each peak and scale for each spectrum. The most abundant +3 peak is circled in red.

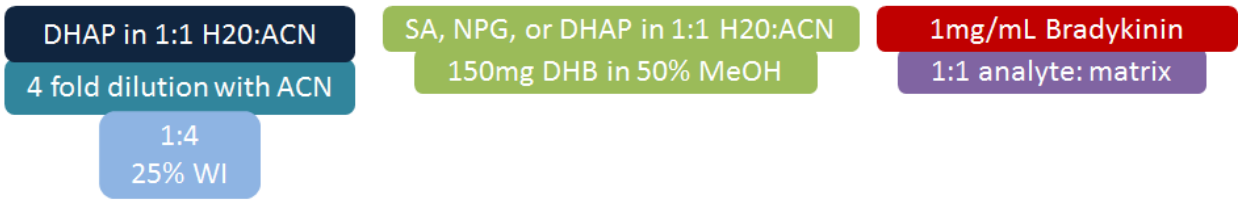


MALDI plate / Diluted matrix / Matrix/analyte mixture

Matrix	ULT	1060(+1)	530(+2)	353(+3)	Scale
DHB	None	100	2.647	9.9914	5.3518
		100	4.1147	7.1564	2.9805
	DHB	100	2.766	9.2849	4.2574
		100	2.628	10.366	5.6239
		100	2.6589	9.793	5.3996
SA		100	1.9344	8.0146	5.5775
		47.716	9.618	4.4908	32.1965
DHAP		74.9592	2.0966	7.6349	126.965
		100	2.4715	9.8678	15.4101
NPG		100	1.3909	10.1437	25.5974
		76.8395	1.9442	7.9064	27.178

Figure 5.6 Binary UTL results for DHAP as 1st layer

The binary matrix application using the UTL method with DHAP as the 1st layer and DHB, SA, NPG, or DHAP for the matrix analyte mixture is displayed. The formation of the 1st layer and the four matrices used for the matrix analyte mixes for the bradykinin concentration are displayed with the relative abundance for each peak and scale for each spectrum. The peaks circled in red have higher relative intensity than the UTL DHB (Figure 5.5).

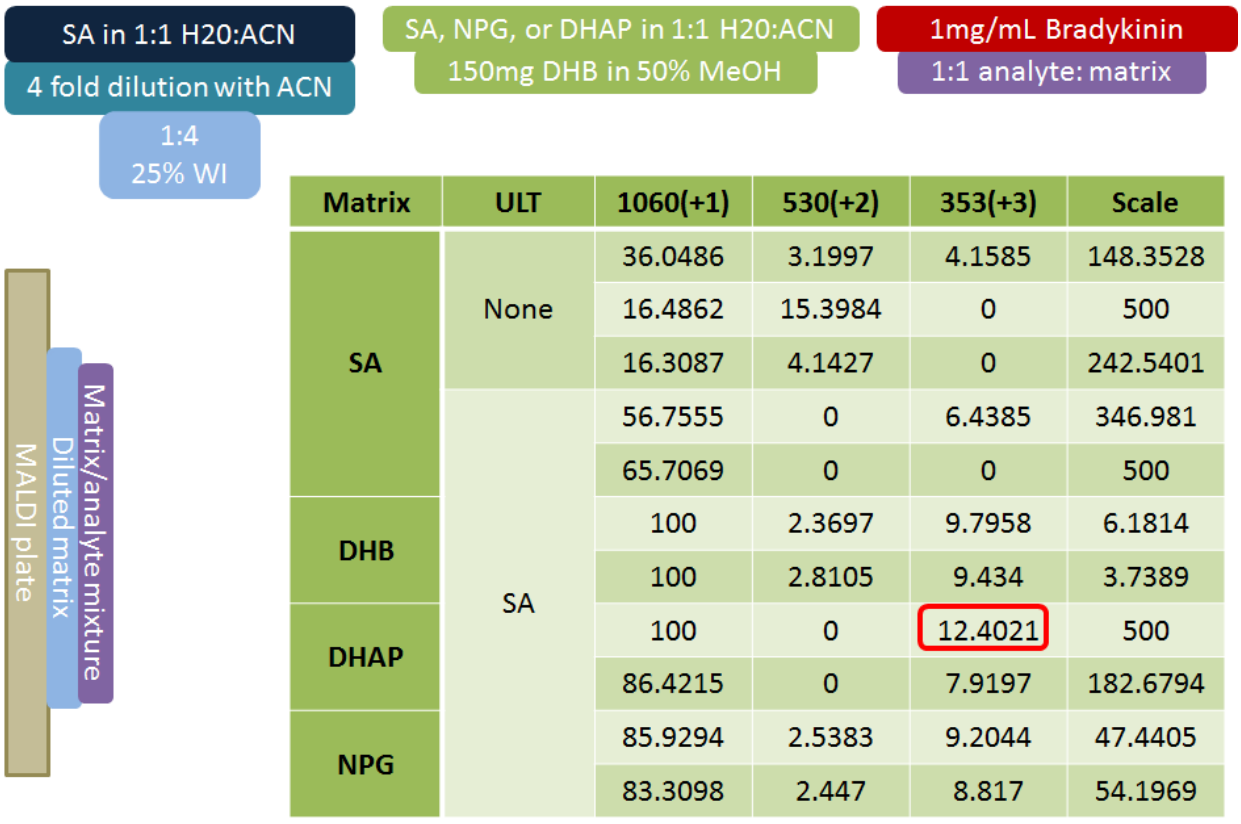


MALDI plate / Diluted matrix / Matrix/analyte mixture

Matrix	ULT	1060(+1)	530(+2)	353(+3)	Scale
DHAP	None	100	3.4857	9.7559	183.6978
		100	2.5339	10.0749	18.3339
		100	2.2997	9.4709	11.5161
	DHAP	100	3.107	10.6151	86.1902
		100	2.5977	10.1147	20.8454
		100	2.6638	10.3949	29.4233
DHB		100	3.6924	7.5094	3.1975
		100	3.2489	10.3755	40.0781
SA		58.5808	0	0	500
		82.0642	0	7.0958	407.2894
NPG		87.5442	2.5354	9.335	44.5211
		89.9413	2.2526	8.9176	46.4918

Figure 5.7 Binary UTL results for SA as 1st layer

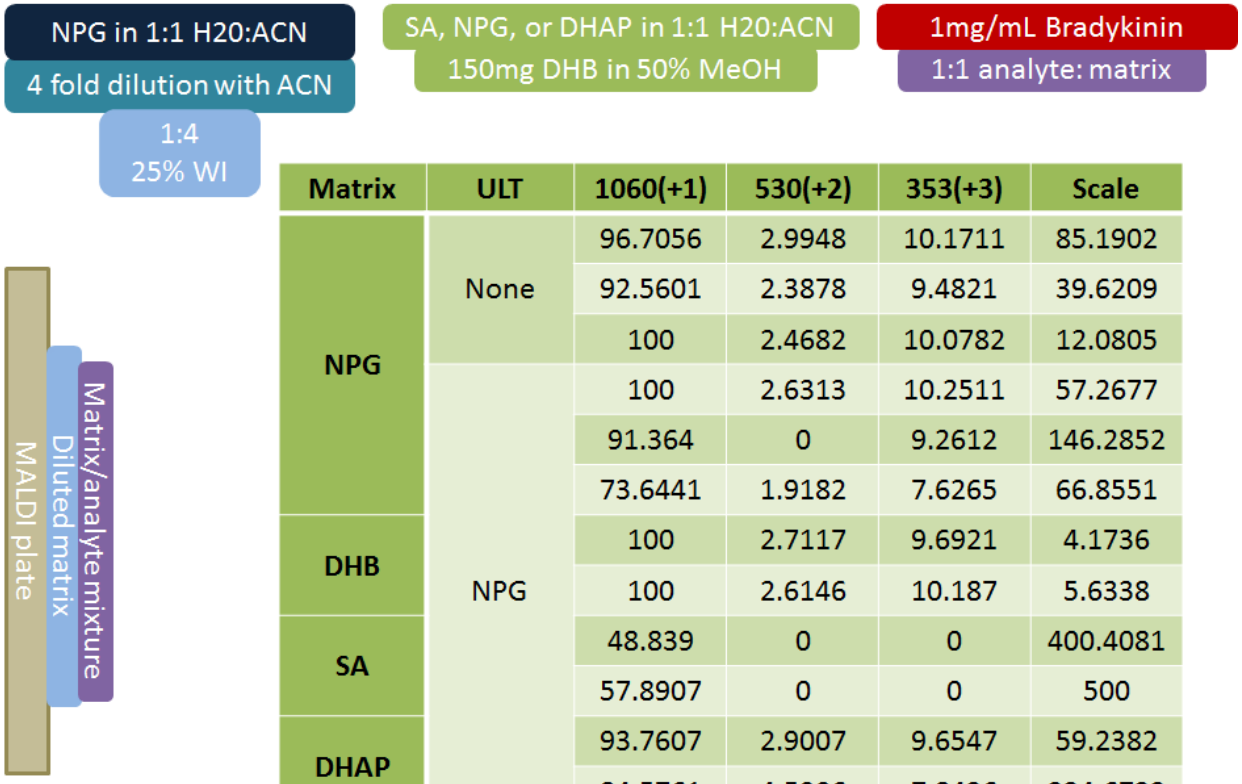
The binary matrix application using the UTL method with SA as the 1st layer and DHB, SA, NPG, or DHAP for the matrix analyte mixture is displayed. The formation of the 1st layer and the four matrices used for the matrix analyte mixes for the bradykinin concentration are displayed with the relative abundance for each peak and scale for each spectrum. The peak circled in red has a higher relative intensity than the UTL DHB (Figure 5.5).



Matrix	ULT	1060(+1)	530(+2)	353(+3)	Scale
SA	None	36.0486	3.1997	4.1585	148.3528
		16.4862	15.3984	0	500
		16.3087	4.1427	0	242.5401
	SA	56.7555	0	6.4385	346.981
		65.7069	0	0	500
DHB		100	2.3697	9.7958	6.1814
		100	2.8105	9.434	3.7389
DHAP		100	0	12.4021	500
		86.4215	0	7.9197	182.6794
NPG		85.9294	2.5383	9.2044	47.4405
		83.3098	2.447	8.817	54.1969

Figure 5.8 Binary UTL results for NPG as 1st layer

The binary matrix application using the UTL method with SA as the 1st layer and DHB, SA, NPG, or DHAP for the matrix analyte mixture is displayed. The formation of the 1st layer and the four matrices used for the matrix analyte mixes for the bradykinin concentration are displayed with the relative abundance for each peak and scale for each spectrum. No peaks were detected above the DHB UTL abundance (Figure 5.5).



Matrix	ULT	1060(+1)	530(+2)	353(+3)	Scale
NPG	None	96.7056	2.9948	10.1711	85.1902
		92.5601	2.3878	9.4821	39.6209
		100	2.4682	10.0782	12.0805
	NPG	100	2.6313	10.2511	57.2677
		91.364	0	9.2612	146.2852
		73.6441	1.9182	7.6265	66.8551
DHB		100	2.7117	9.6921	4.1736
		100	2.6146	10.187	5.6338
SA		48.839	0	0	400.4081
		57.8907	0	0	500
DHAP		93.7607	2.9007	9.6547	59.2382
		84.5761	4.5906	7.8436	234.6722

Figure 5.9 Bradykinin Fragmentation

The spectrum obtained by plating NPG with bradykinin. The m/z 904 peak is dominant, resulting from the cleavage of the first arginine. Sites of protonation for bradykinin are displayed in red. Multiply charged species were detected for both the the fragment peptide and the full length peptide and their relative abundances displayed.

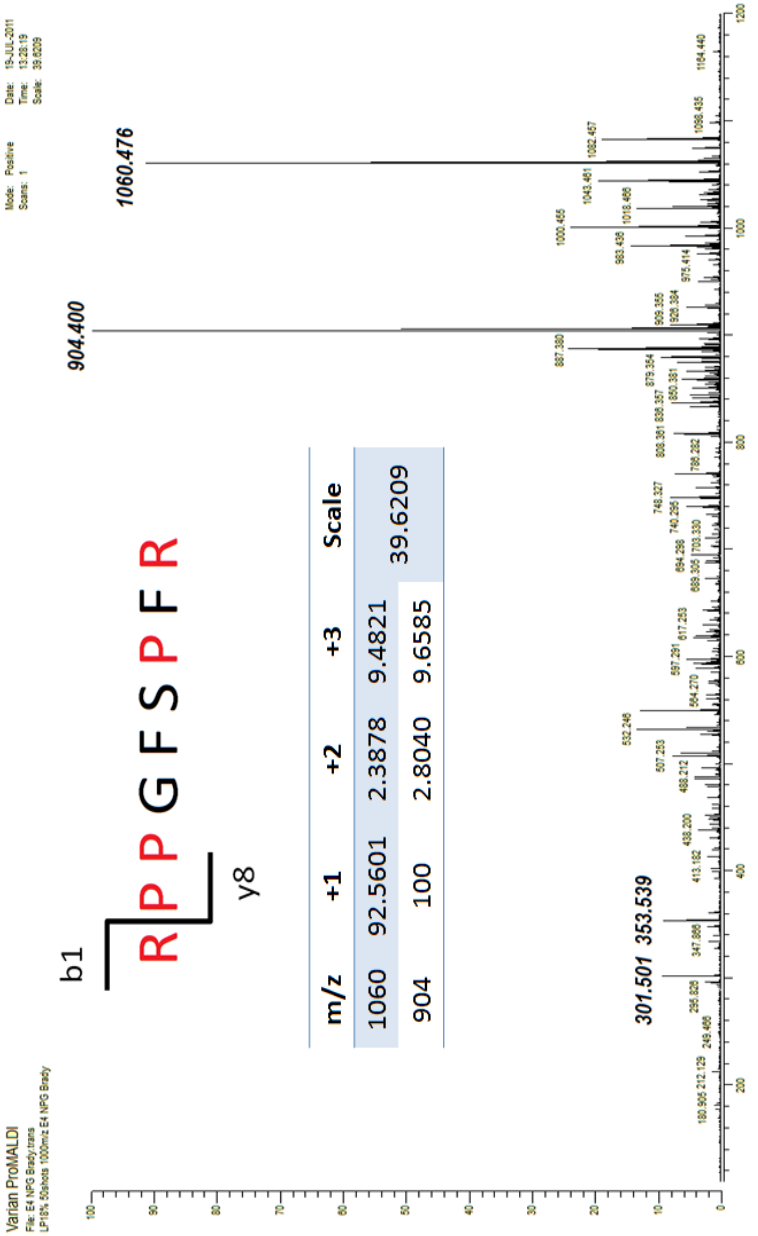
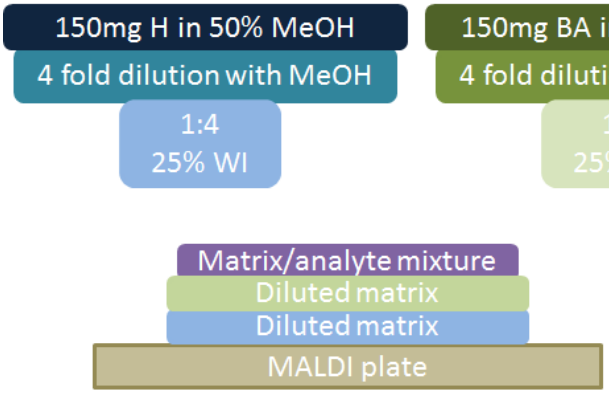


Figure 5.10 DHB Structural Analysis

An UTL tertiary matrix scheme is demonstrated below. A 1st and 2nd layer are created by diluting H and BA in methanol and isopropanol. The top layer is saDHB at 1:1 with bradyknin. The spots were analyzed on the TOF/TOF and the ratio of +2 to the +1 are demonstrated and compared to a dry droplet saDHB sample.



Sample	Ratio (+2/+1)
saDHB	0.003869
UTL H50 UTL BA100 Top DHB50	0.0040877

Figure 5.11 Binary NA and DHB Crystal

A DHB crystal formed in solution in cold, dark conditions was plated and NA mixed at a 1:1 ration with bradykinin was deposited on top of the crystal by the dry droplet method. The instrument was set to collect 50,000 shots but analysis was halted at 30,099 shots and the results are shown. The blue brackets indicate the location of the abundance for the number of indicated shots during the 50,000 shot analysis. The data collected is compared both to a freshly prepared saDHB dry droplet sample and a regular 500 shot analysis collected on the binary crystal and NA sample.

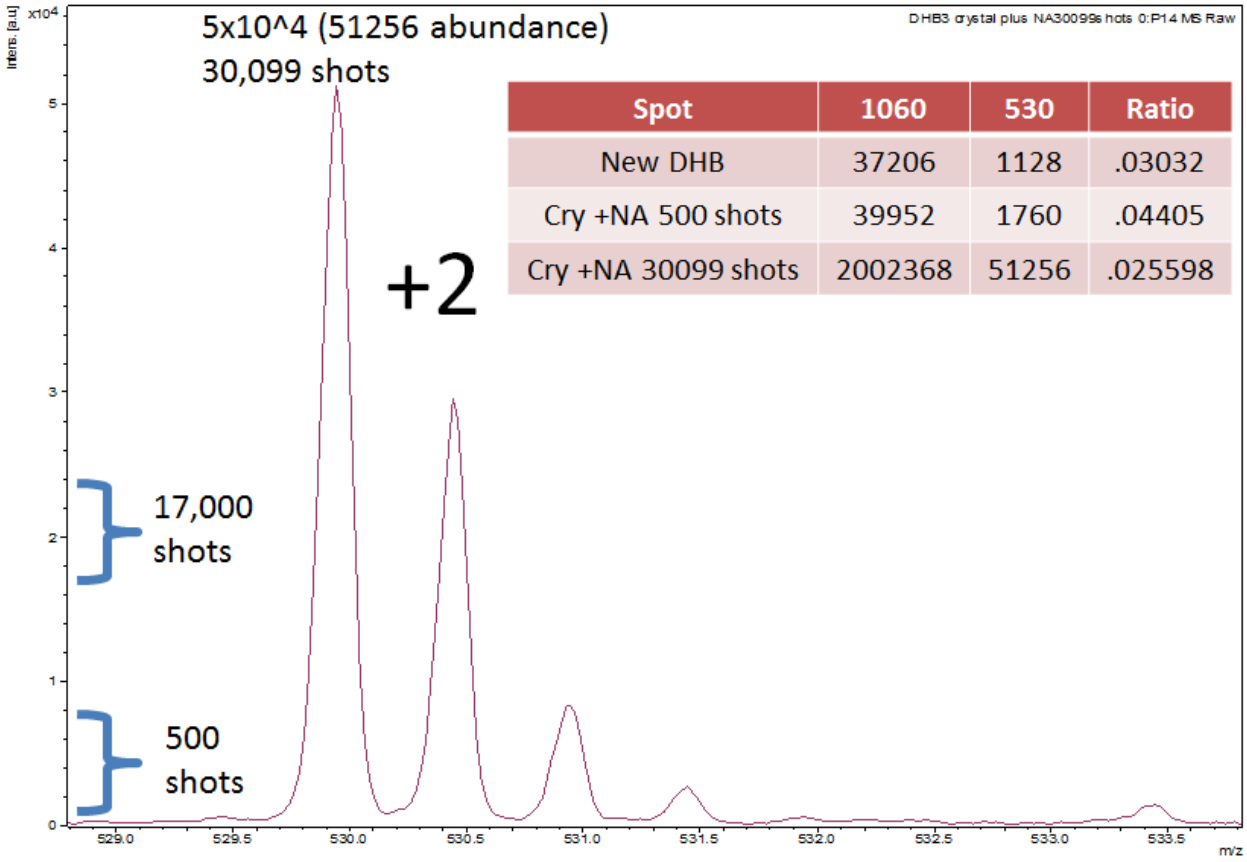


Figure 5.12 TOF/TOF Cut-off

Nothing above the isolated m/z is detected by the TOF/TOF for the +2 species of bradykinin. The fragments detected are both internal fragmentation fragments and fragment ions which are listed in Table 5.4.

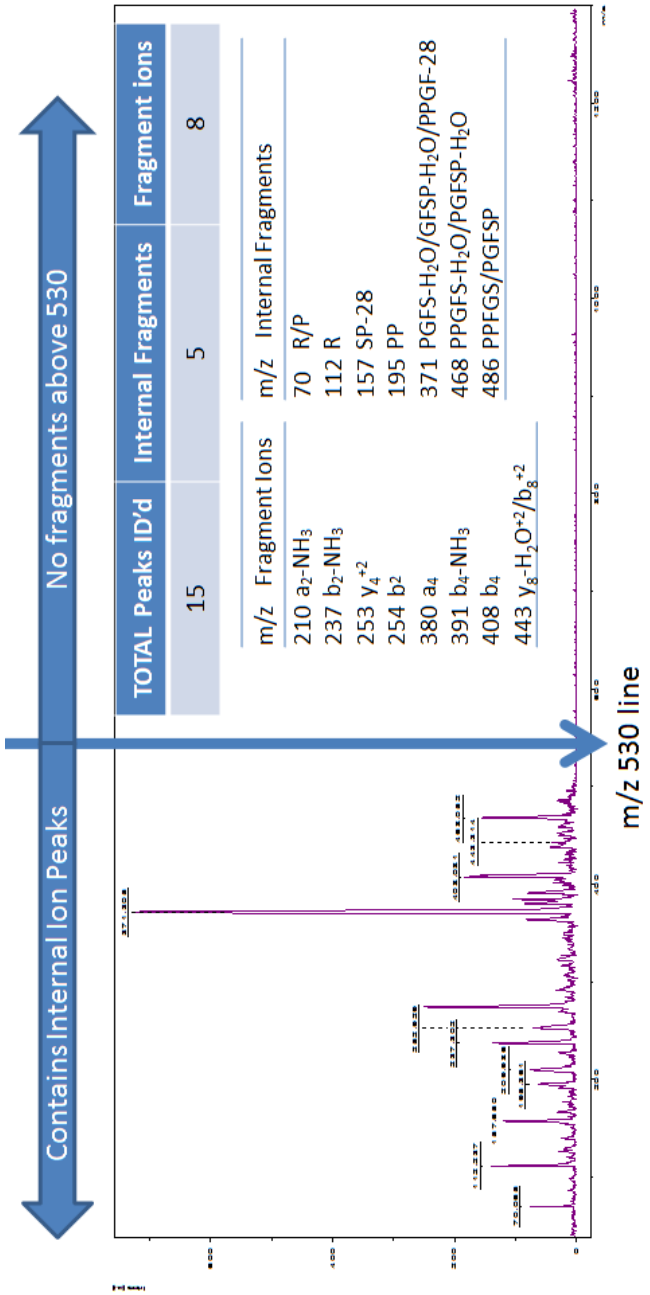


Figure 5.13 Orbitrap CID Fragmentation

The tandem MS spectrum for the doubly charged bradykinin peak generated by CID fragmentation on the Orbitrap.

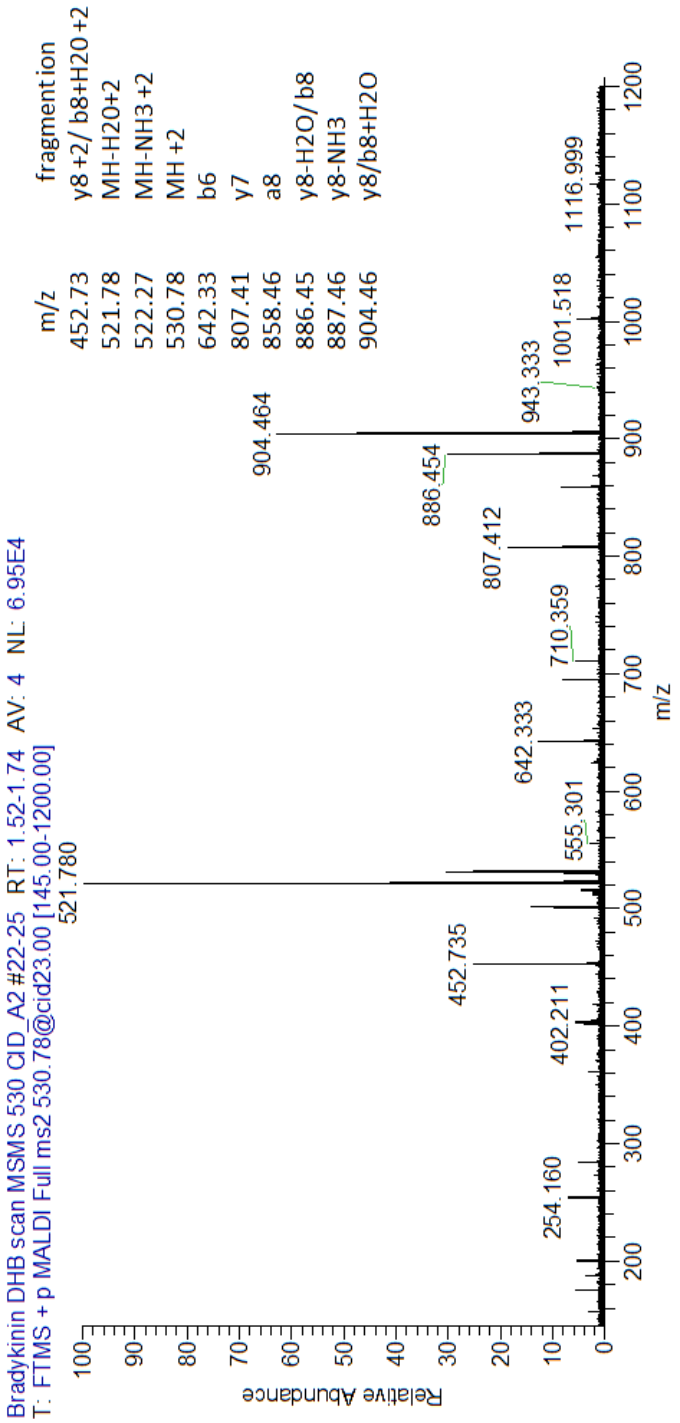


Figure 5.13 Orbitrap HCD Fragmentation

The tandem MS spectrum for the doubly charged bradykinin peak generated by HCD fragmentation on the Orbitrap.

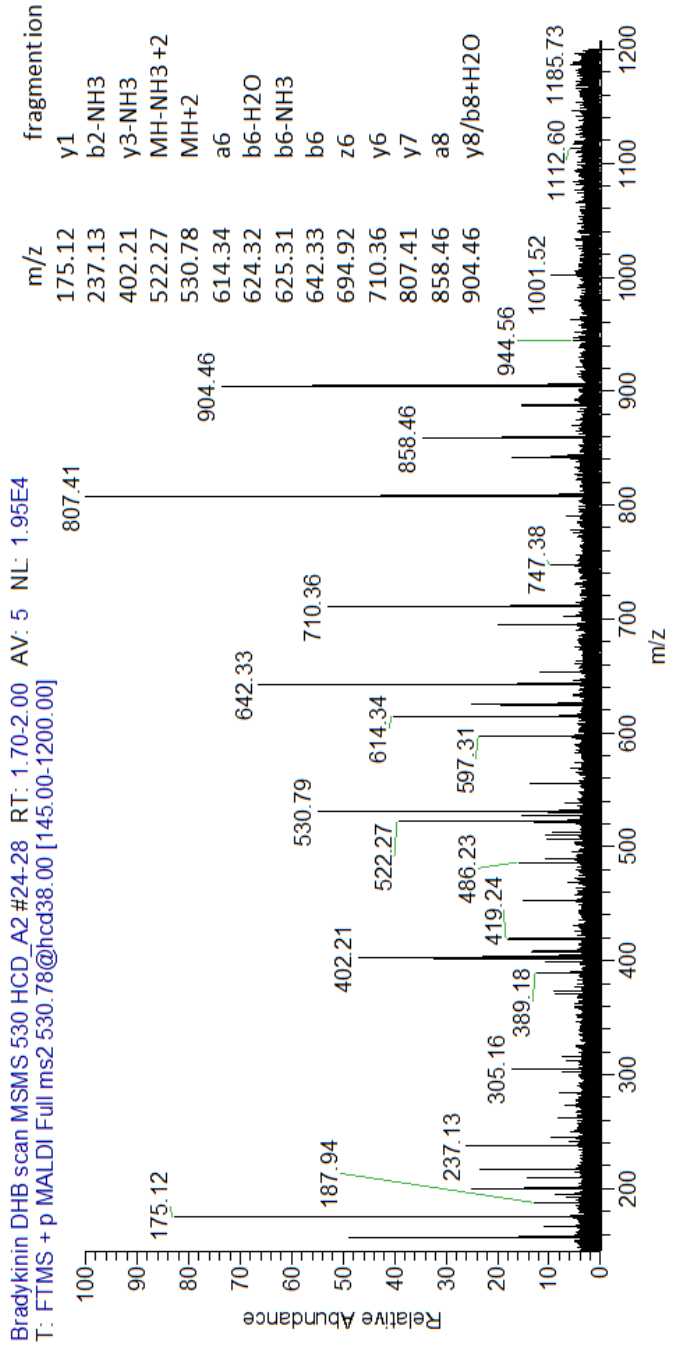


Table 5.1

Multiply charged MALDI data obtained by varying plate surface, matrix, and matrix solution. The solutions 1-4 are (1) water/methanol at 2:1 ratio by volume (2) formic acid/water/2-propanol at 1:3:2 ratio by volume (3) water/acetonitrile at 1:1 ratio by volume and (4) trifluoroacetic acid/water/acetonitrile at 1:4.5:4.5 ratio by volume.

Plate Surface	Matrix	Matrix Solution	Scale	Relative Abundance		
				1060 (+1)	530 (+2)	353 (+3)
Regular Metal	DHB		23.2844	100	2.7	10.1
			8.5889	100	2.6	8.4
			8.4121	100	4.5	10.6
	CHCA	1	103.3712	100	2.9	10.2
			271.4234	100	0	9.7
			174.5093	79.4	0	5.9
			79.9033	45	1.6	4.3
		2	96.4235	60.4	0	5.6
			250.149	51.5	0	0
		3	314.8849	54	0	0
			192.3343	41.9	0	0
		4	250.8654	38.9	0	0
			120.9609	67.2	0	7.4
	THAP	1	41.7169	100	3.4	10.1
		2	268.7758	100	0	7.32
		3	60.1325	100	3.7	10.2
		4	74.3871	100	3.7	10.2
	SA	1	453.7859	72.3	10.6	0
			247.2592	100	7.6	10.4
		2	170.4244	100	5.2	8.3
		3	95.3973	100	4.6	11.1
		4	178.9976	100	3.8	9.7
Parafilm Covered	DHB		23.9721	100	2.9	10.2
			122.7604	100	2.6	7.7
Parafilm and Glass Slip Covers	DHB		41.5890	79.9	2.3	7.1
	THAP	1	500	100	0	0
		2	500	100	0	0
		3	78.7101	100	3.18 (1)	8.5

Table 5.2 Matrix Solution

DHB made in different matrix solutions and compared to saDHB using bradykinin as the peptide. The solutions 1-4 are (1) water/methanol at 2:1 ratio by volume (2) formic acid/water/2-propanol at 1:3:2 ratio by volume (3) water/acetonitrile at 1:1 ratio by volume and (4) trifluoroacetic acid/water/acetonitrile at 1:4.5:4.5 ratio by volume.

Matrix Solution	1060(+1)	530(+2)	353(+3)	Scale
saDHB	100	3.142	9.146	3.8543
1	93.936	2.449	8.809	5.8328
2	63.267	1.708	6.663	9.4941
3	73.063	2.368	7.882	27.8049
4	76.27	1.986	4.577	18.0771

Table 5.3 Methanol Concentrations

DHB is made in varying concentrations of methanol to water and analyzed with bradykinin using the dry droplet method on the TOF/TOF. The number following DHB indicates the percentage of methanol in the solution.

[MeOH]	Ratio (+2/+1)
DHB 100	0.00259
DHB 75	0.004566
DHB 50	0.003869
DHB 25	0.0012209

Chapter 6: Deamidation for Identification of Amides in a Peptide Sequence

6.1 Abstract

Modification of Miller and Waelsch's acid driven deamidation protocol resulted in the formulation of a simple, non-enzymatic method for determining the number and location of amides in a peptide sequence by mass spectrometric analysis. When analyzed by MALDI mass spectrometry the protocol reveals the total number of amides within a peptide. This is identifiable by the 1Da mass shift that results from the conversion of each amide to a carboxylic acid. When analyzed by nanoLC-ESI mass spectrometry the protocol reveals the location of the amide. This is determined because deamidation of asparagine and glutamine creates a retention time shift while deamidation of the C-terminal amide does not. Such information is pertinent in *de novo* mass spectrometric sequencing since it is a quick technique that will facilitate elucidating sequence ambiguity.

6.2 Introduction

In *de novo* sequencing, when a peptide sequence cannot be elucidated by tandem mass spectrometric fragmentation, it is essential to pursue alternative means for sequence identification. Chemical manipulation of the peptide allows for improved fragmentation thus improving detection of the amino acid residues present. Previously, the Li lab used methyl esterification to determine the exact number and location of aspartic and glutamic acids in Orcokinins and other peptides [1]. The advantage of using deamidation is that it can also be coupled with mass spectrometry to identify the number of amides in the sequence.

Deamidation is the process of converting an amide (CONH2) into a carboxylic acid (COOH). In peptides, deamidation converts an asparagine (N) to aspartic acid (D), a glutamine

(Q) to a glutamic acid (E), or an amidated C-terminus to its counterpart carboxylic acid. The conversion process produces a 1Da mass shift for each amide present in the peptide.

Deamidation of the N and Q occurs naturally in peptides and is believed to be a method for regulating the timing of biological processes [2]. The half-life of this process varies greatly, on a scale ranging from minutes to years, depending on the size of the peptide and the number and location of the N and Q within the peptide sequence [2]. Deamidation is also induced chemically by enzymes or nonezymatic processes. Acid driven deamidation is favored because it is inexpensive and has been proven to work in conjunction with methyl esterification [1].

Acid induced deamidation is dependent upon the reaction time, temperature, and the strength and pH of the acid [3-7]. All four factors have been varied until the optimum conditions for inducing deamidation were found. This resulted in the acid driven deamidation of Q and N in relation to the C-terminal amide for rapid screening of the presence of these amino acids to improve the specificity for peptide characterization.

6.3 Materials and Methods

Most standards were purchased from American Peptide Company (Table 6.1). Orcokinin was synthesized by the Biotechnology Center of the University of Illinois at Urbana-Champaign. Trifluoroacetic acid (TFA), formic acid (FA), methanol (MeOH), trichloroacetic acid (TCA), acetic acid (AA), and acidified methanol (AcMeOH; 90:9:1 MeOH:H_2O:Acetic acid) were also used.

Deamidation Protocol

A 10μL sample was mixed at 1:1 by volume ratio with 40% TFA in water (10μL); the final concentration of the acid for the reaction is 20%. The mixture was placed in boiling water

for 30 minutes then placed on ice to cool for 5 minutes. Then 1mL of 50% methanol in water was added and the sample was placed in the speed vacuum and completely dried before being reconstituted in 0.1% FA for mass spectrometric analysis.

MALDI TOF Analysis

Two MALDI instruments, MALDI FTICR and the ABI MALDI TOF/TOF, were used for analysis as determined by instrument availability. See Appendix A for instrument description.

Amazon Analysis

See Appendix A.

6.4 Results and Discussion

A previously published procedure reported the deamidation of single and di-amino acid chains of Q and N in 10% TCA for one hour at 70°C [8]. Initially, this protocol was modified by acid, temperature, and incubation time. Either 20% FA or 20% TFA was mixed at a 1:1 ratio by volume with a single peptide standard and incubated at one of eight temperatures (30, 40, 50, 60, 70, 80, 90, or 100°C) maintained by a water bath for one or three hours. De-ionized water was added to quench the reaction and reduce the acid concentration before completely drying the sample by speed vacuum and reconstitution in 0.1% FA for analysis by MALDI FTMS.

Peptide standards were chosen for their different combinations and locations of Q and N in relation to the C-terminal amide. Angiotensin and bactenecin are experimental standards to ensure that mass shifts are the result of deamidation and not adverse side reactions to other amino acids present. Bactenecin has the added advantage of containing a disulfide bond to

ensure that the shift changes are the result of deamidation and not the reduction of the disulfide bond.

Since the mass shift caused by the deamidation process is only 1Da for each amide in the sequence, partial deamidation of the sample resulted in peak interference between the original and deamidated peptide isoforms. When the deamidation peak is the dominant peak in the spectrum, it is assumed that the deamidation reaction was driven mostly toward completion. This is indicated by a whole number in the results. The value of the whole number represents the mass shift between the original and dominant deamidated peak. However, if the original peptide peak dominates the signal then either the deamidation reaction did not occur or was not favored. Any disruption in the normal isotopic distribution pattern of the original peptide is used to indicate that deamidation occurred but was not favored. A disfavored deamidation reaction, regardless of the mass shift between the deamidated and original peak, was arbitrarily assigned a value of 0.5 in the results. Table 6.2 displays the results for one hour incubation while Table 6.3 displays the results for the three hour incubation using this numbering scheme. Figures 6.1, 6.2, and 6.3 demonstrate peak interference patterns resulting in how the assignment of 0.5, 1, and 3 values in the results were determined, respectively.

These initial experiments quickly revealed that analysis on all the standards simultaneously was time consuming and the rate determining step for adequate progress in these experiments (Table 6.4). To speed up the process for determining an optimal deamidation protocol, all additional protocol optimization experiments were only conducted on Orcokinin. Orcokinin was chosen since it was the original peptide used in the methyl esterification experiments previously performed [1]. For additional parameter simplification, a fixed one hour

incubation time and 70°C temperature were chosen based on the original deamidation protocol [8]. Once the procedure was optimized for this single peptide, the protocol was applied to the initial eight peptides then to a mix of the eight peptides for validation.

However, there was a complication with determining what acid to use. Recent studies suggest that methyl esterification could result from the acidified methanol extraction protocol [9, 10]. Acetic acid (AA) was therefore added to the analysis. TCA was also added because it was used in the initial study [8]. Acidified methanol (AcMeOH; 90:9:1 MeOH: H_2O: acetic acid) was tested to determine if it had any effects on deamidation (Figure 6.5). The acid comparison studies indicated that TFA provided the best deamidation results (Table 6.5).

Since peak interference was still present, to attempt to drive the reaction to completion, the effects of time and acid concentration on deamidation were analyzed. First, the incubation time was increased, incrementally to 24 hours. An equilibrium mixture was determined at the seven hour incubation time and anything beyond negatively affected the ratio of the deamidation peak to the original peak (Figure 6.4). Increasing the heat of the reaction was investigated to speed up the reaction time. The ratio of deamidated to original at 7 hours in 70°C water bath was matched by 30 minutes in boiling water (98°C) (Figure 6.5). A control for the effects of heat on the peptide was conducted with each experiment where water, instead of acid, was added to the peptide before incubation. Heat was determined to have no effect on the peptide (results not shown).

In pursuit of complete deamidation, higher concentrations of acid were analyzed (10%, 20% and 50% final concentrations). Although 50% final acid concentration was producing

better deamidation ratios, they were only slightly better than 20% (Figure 2.6). Higher concentrations of acids are not compatible with speed vacuum usage.

An acid concentration below 1% is required for speed vacuum use because of the negative effects the acid has on the pump. Initially water was used to decrease the concentration of the acid. However, an increase in water volume requires a longer drying time therefore 50% methanol was used decrease the time required to complete the drying step since organics dry faster in the speed vacuum.

The protocol was applied to the eight peptides individually and as a mixture. The peptides were analyzed by tandem mass spectrometry to validate that deamidation was occurring (Table 6.6). Deamidation was occurring, but determining whether or not it was occurring within the peptide or at the C-terminal was not possible using the established method analyzed with MALDI ionization. Deciding to abandon this aspect of the research and move forward with the results to perform the deamidation protocol on a biological sample proved beneficial. A peptide standard mix was spiked into the neuropeptide extract from blue crab pericardial organs and divided into three aliquot amounts to test the deamidation protocol.

Since the biological sample was more complex in neuropeptide content then the peptide mixture, analysis of the deamidation protocol on the biological samples was performed by nanoLC-ESI. This resulted in rapid sample separation coupled to tandem mass spectrometry for confirmation of the peptides detected. Ironically, coupling the protocol to LC- ESI enabled the desired determination of amide location. A shorter retention time was observed for a peptide containing an N or Q then its D and E counterparts. This observation is in agreement with previous publications which detect deamidation by LC methods [5, 11-14].

Another interesting result obtained from the biological sample suggested pyroglutamate (pQ) rings were open under the deamidation protocol due to a 20 Da mass shift for the pQGFYSQRYamide peptide. To test this theory, both corazonin and LH-RH peptide standards were analyzed for pQ opening and deamidation with the deamidation protocol but the results were inconclusive.

In conclusion, a simple acid driven deamidation protocol was formulated and analyzed with mass spectrometry. MALDI analysis determines the number of amides within the peptide while nanoLC-ESI analysis determines whether it is an N/Q or C-terminal amide. This protocol also has potential application with pQ opening for paired labeling quantitation.

6.5 Future Work

Application of the deamidation protocol is expected to help elucidate the presence of amide-containing peptides with sequence ambiguity, such as many of the novel Orcokinins, CHHs, and CPRPs identified in crustaceans.

Further development of pQ opening for labeling quantitation is highly desirable. The opening of the pQ ring has great potential for label quantitation since tags are the popular method of quantitation. Since many chemical tags label peptides on the N-terminus or lysine in a peptide, a pQ modification would prevent labeling at the N-terminus and subsequent analysis which is primarily used for quantitation and identification of potential biomarkers for disease. For our own studies, over 4% of the peptides listed in our crustacean database are both N-terminally blocked and lack a lysine for labeling. N-terminally blocked peptides are also determined to be biological significant in studying Down syndrome and Alzheimer's [15-17].

6.6 Respective Contributions and Acknowledges

I would like to thank Robby Cunningham for running my samples on the AmaZon and helping with the initial data analysis of those AmaZon samples. I would also like to thank his undergraduate, Dan Wellner, for performing the neuropeptide extraction on the PO samples Robby and I collected.

6.7 Works Cited

1. Ma, M., K.K. Kutz-Naber, and L. Li, *Methyl esterification assisted MALDI FTMS characterization of the orcokinin neuropeptide family.* Anal Chem, 2007. **79**(2): p. 673-81.
2. Robinson, N.E. and A.B. Robinson, *Amide molecular clocks in drosophila proteins: potential regulators of aging and other processes.* Mech Ageing Dev, 2004. **125**(4): p. 259-67.
3. Karty, J.A. and J.P. Reilly, *Deamidation as a consequence of beta-elimination of phosphopeptides.* Anal Chem, 2005. **77**(14): p. 4673-6.
4. Wilson, H. and R.K. Cannan, *THE GLUTAMIC ACID-PYRROLIDONECARBOXYLIC ACID SYSTEM.* Journal of Biological Chemistry, 1937. **119**(1): p. 309-331.
5. Capasso, S. and P. Di Cerbo, *Kinetic and thermodynamic control of the relative yield of the deamidation of asparagine and isomerization of aspartic acid residues.* J Pept Res, 2000. **56**(6): p. 382-7.
6. Dekker, C.A., D. Stone, and J.S. Fruton, *A peptide from a marine alga.* J Biol Chem, 1949. **181**(2): p. 719-29.
7. Flatmark, T. and K. Sletten, *Multiple forms of cytochrome c in the rat. Precursor-product relationship between the main component Cy I and the minor components Cy II and Cy 3 in vivo.* J Biol Chem, 1968. **243**(7): p. 1623-9.
8. Miller, H.K. and H. Waelsch, *Utilization of glutamine and asparagine and their peptides by micro-organisms.* Nature, 1952. **169**(4288): p. 30-1.

9. Stemmler, E.A., et al., *C-terminal methylation of truncated neuropeptides: An enzyme-assisted extraction artifact involving methanol.* Peptides, 2013. **46**(0): p. 108-125.

10. Kutz, K.K., J.J. Schmidt, and L. Li, *In Situ Tissue Analysis of Neuropeptides by MALDI FTMS In-Cell Accumulation.* Analytical Chemistry, 2004. **76**(19): p. 5630-5640.

11. Bischoff, R. and H.V. Kolbe, *Deamidation of asparagine and glutamine residues in proteins and peptides: structural determinants and analytical methodology.* J Chromatogr B Biomed Appl, 1994. **662**(2): p. 261-78.

12. Tyler-Cross, R. and V. Schirch, *Effects of amino acid sequence, buffers, and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides.* J Biol Chem, 1991. **266**(33): p. 22549-56.

13. Lindner, H. and W. Helliger, *Age-dependent deamidation of asparagine residues in proteins.* Exp Gerontol, 2001. **36**(9): p. 1551-63.

14. Sarioglu, H., et al., *Deamidation as a widespread phenomenon in two-dimensional polyacrylamide gel electrophoresis of human blood plasma proteins.* Electrophoresis, 2000. **21**(11): p. 2209-18.

15. Schilling, S., et al., *Glutaminyl cyclase inhibition attenuates pyroglutamate Abeta and Alzheimer's disease-like pathology.* Nat Med, 2008. **14**(10): p. 1106-11.

16. Jawhar, S., O. Wirths, and T.A. Bayer, *Pyroglutamate amyloid-beta (Abeta): a hatchet man in Alzheimer disease.* J Biol Chem, 2011. **286**(45): p. 38825-32.

17. Marcello, A., et al., *Reduced levels of IgM autoantibodies against N-truncated pyroglutamate Abeta in plasma of patients with Alzheimer's disease.* Neurobiol Aging, 2011. **32**(8): p. 1379-87.

2.8 Figures and Tables

Figure 6.1 Defining the 0.5 mixture

A 0.5 mixture is demonstrated. The normal isotopic distribution of Vasopressin is shown on the left. The new peak distribution on the right indicates deamidation partially occurred since the normal distribution is disrupted but the deamidation peak is not the dominant peak.

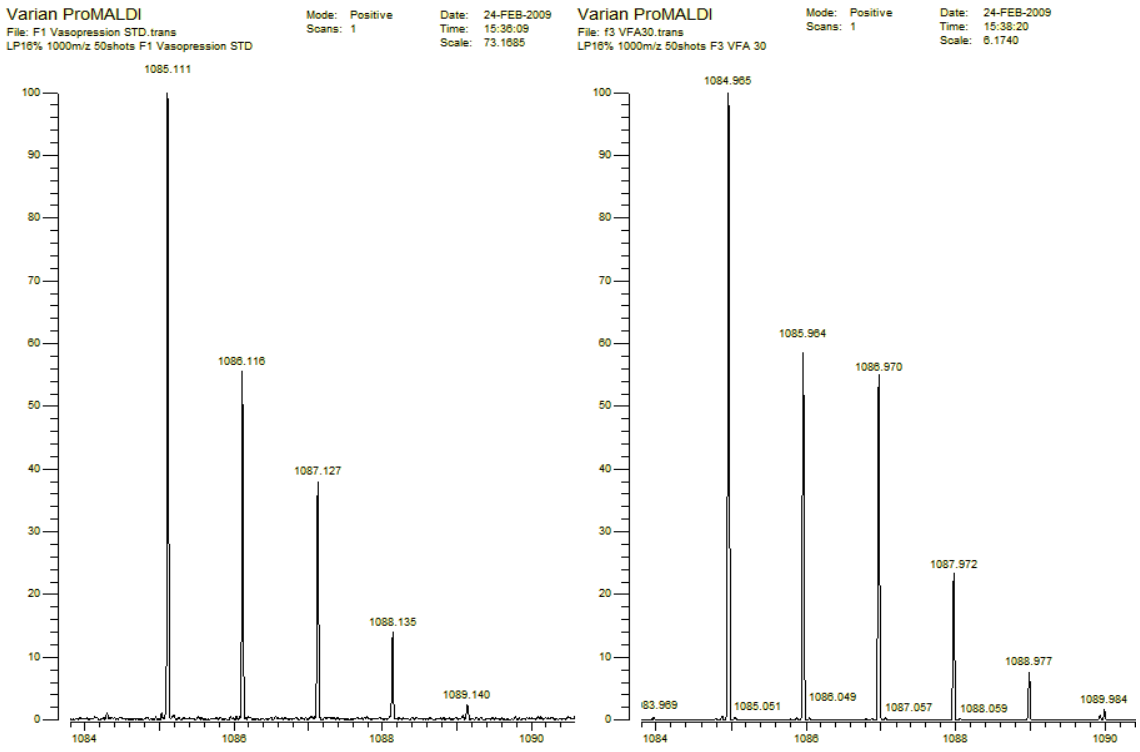


Figure 6.2 Defining the 1 mixture

A 1 mixture is defined. The normal isotopic distribution of Somatostatin is shown on the left. The new peak distribution on the right indicates a single deamidation has occurred. Due to partial deamidation of the peptide, peak interference is occurring between the original and deamidated peptide isoforms. Since the dominant peak only 1 Da away from the original peptide peak, a 1 is used to indicate the mass shift.

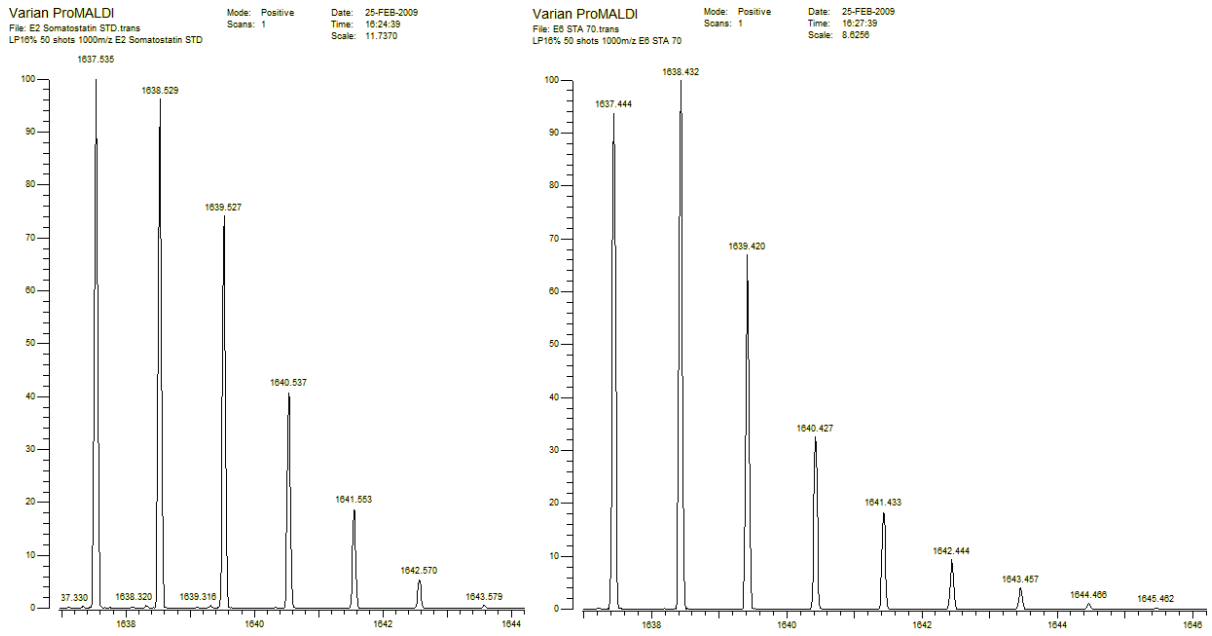


Figure 6.3 Defining the 3 mixture

A 3 mixture is defined. The normal isotopic distribution of Vasopressin is shown on the left. Due to partial deamidation of the peptide, peak interference is occurring between the original and deamidated peptide isoforms. Since the dominant peak is 3 Da away from the original peptide peak, a 3 is used to indicate the mass shift.

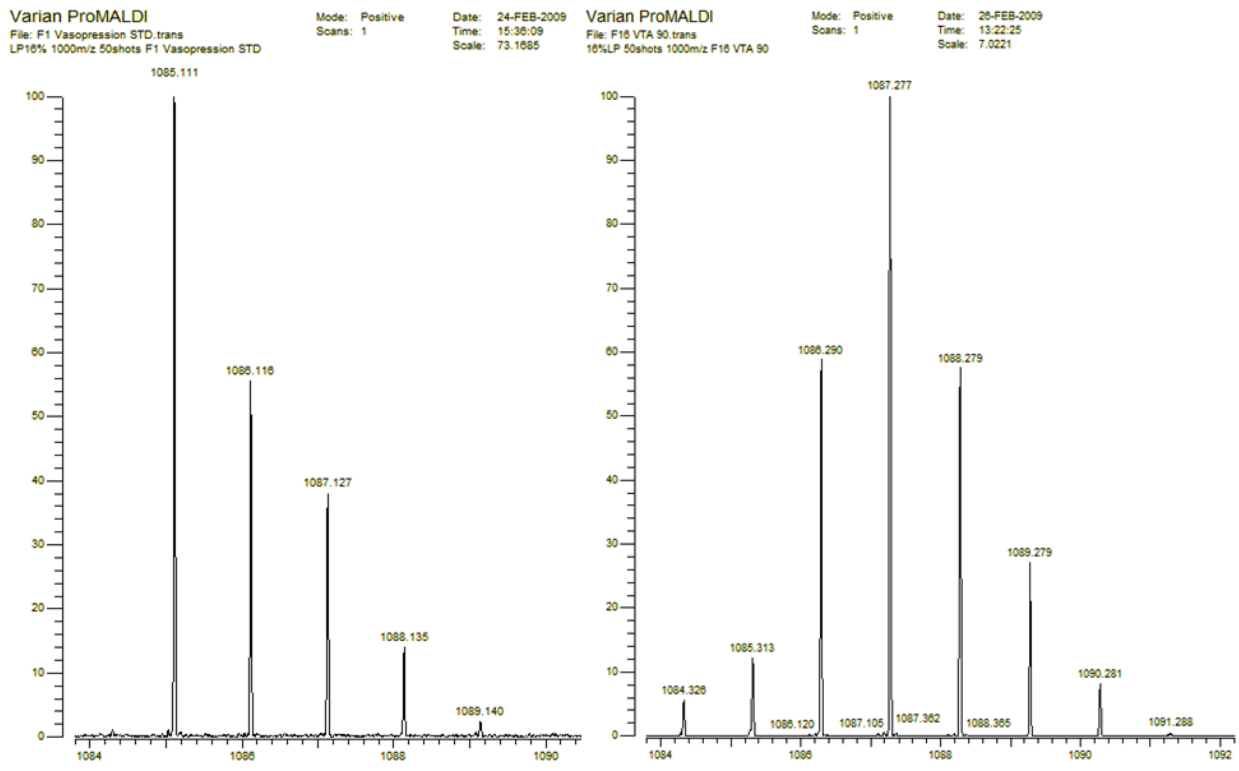


Figure 6.4 TFA incubation time analyses at 70°C

Deamidation results for Orcokinin peptide at various incubation times at 70°C. The ratio is the result of averaging the deamidation peak to the regular peak.

	Ratio	n	STDev
STD	0.801339	6	0.177986
TFA 1	0.902453	1	
TFA 3	1.069789	4	0.11471
TFA 5	1.249877	3	0.102445
TFA 7	1.595857	3	0.059013
TFA 10	1.303092	4	0.083962
TFA 12	1.235843	4	0.037355
TFA 24	1.393659	3	0.038234

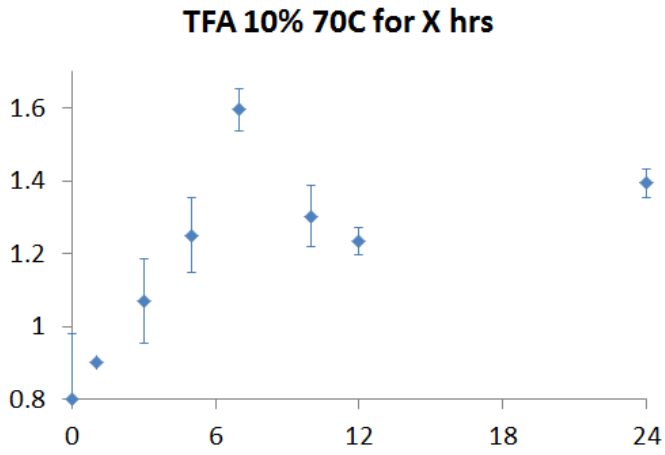


Figure 6.5 TFA incubation time analyses in boiling water

Deamidation results for Orcokinin peptide at various incubation times at 98°C (in boiling water). The ratio is the result of averaging the deamidation peak to the regular peak.

Time	Ratio	n	STDev
5 min	0.880926	4	0.009394
10 min	0.915672	4	0.035078
15 min	1.02399	4	0.041287
20 min	1.056528	4	0.039724
30 min	1.574855	3	0.110126

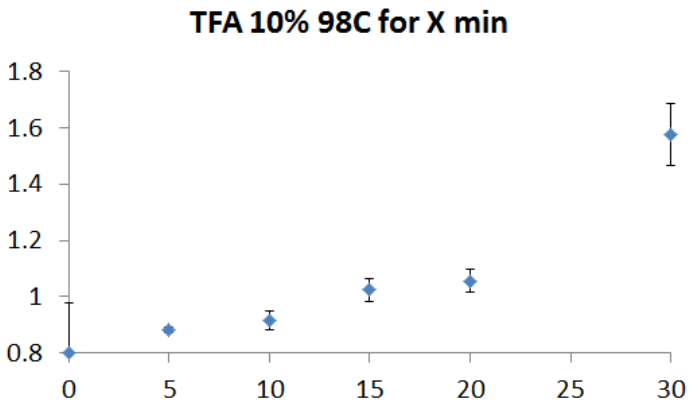
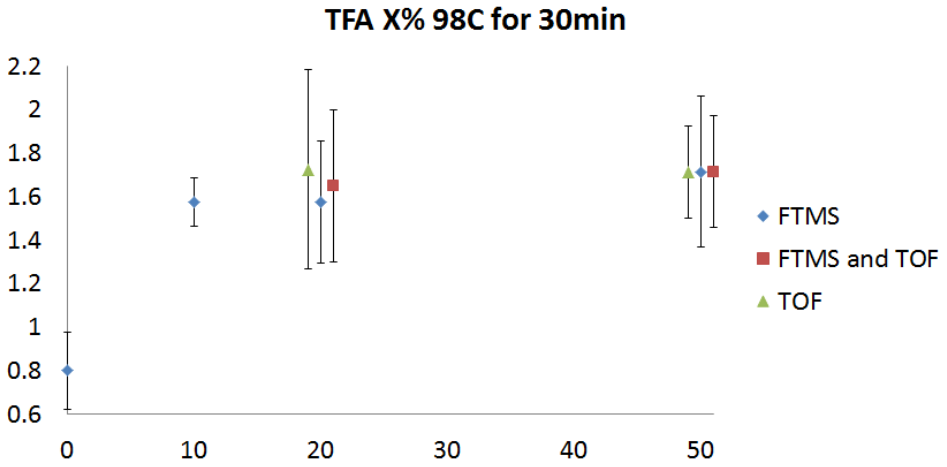


Figure 6.6 TFA Concentration analyses

Deamidation results for the Orcokinin peptide at various contractions of TFA. The concentration listed is the final reaction concentration of the acid at a 1:1 incubation with the peptide standard. Data was analyzed by MALDI FTMS and TOF/TOF as determined by instrument availability.



	FTMS			TOFTOF			Both		
	Ratio	n	STDev	Ratio	n	STDev	Ratio	n	STDev
STD	0.801339	6	0.177986	0.801339					
TFA 10	1.574855	3	0.110126						
TFA 20	1.576491	3	0.281222	1.725314	3	0.456031	1.650903	6	0.348518
TFA 50	1.715341	3	0.345971	1.712697	3	0.211088	1.714019	6	0.256328

Table 6.1 Eight standards list

The eight standards analyzed in this experiment are listed. Their sites of deamidation are bolded red in the sequence. Location of disulfide bonds are indicated in blue. Bactenecin was added after the initial experiments to validate that deamidation and not disulfide bond reduction was occurring.

Standard	M+H	Sequence	Concentration
Angiotensin I, Human	1296.5	DRVYIHPFHL	$5*10^{-5}$M
Bactenecine, bovine	1483.9	RLCRIVVIRVCR	10^{-5}M
Crustacean Cardioactive Peptide CCAP	956.1	PFC**N**AFTGC**a**	10^{-5}M
[Ala 13] Orcokinin	1474.70	**N**FDEIDRSGFGA	10^{-5}M
Peptide K	1104.6	KHK**N**YLRF**a**	$4*10^{-6}$M
Somatostatin	1637.74	AGCK**N**FFWKTFTSC	$6*10^{-5}$M
Substance P	1347.74	RPKP**QQ**FFGLM**a**	$7.42*10^{-6}$M
Vasopressin AVP	1084.3	CYF**QN**CPRG**a**	$5.5*10^{-5}$M

Table 6.2 One hour incubation for two acids

Mass shifts determined from the MALDI FTMS analysis for the deamidation protocol modification experiments. A 0.5 indicates that deamidation was not favored but did occur (Figure 6.1).

	30°C		40°C		50°C		60°C		70°C		80°C		90°C		100°C	
	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA
Angiotensin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Orcokinin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Peptide K	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5
Somatostatin	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
Substance P	0	0.5	0	0.5	0	0.5	0	0.5	0	0	0.5	0.5	0	0.5	0.5	0.5
Vasopressin	0.5	0.5	0	0.5	0.5	0	0	0	0	0.5	0	0.5	0	3	0	0.5

Table 6.3 Three hour incubation for two acids

Mass shifts determined by the MALDI FTMS analysis for the deamidation protocol modification experiments (N=3). A dash indicates that the sample was not detected by MALDI analysis. Slashes indicate that both mass shifts were observed.

	30°C		40°C		50°C		60°C		70°C		80°C		90°C		100°C	
	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA	FA	TFA
Angiotensin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bactenecin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CCAP	0	-	-	-	0	-	0	1	-	-	1	-	1	-	-	-
Orcokinin	0	-	-	0	-	0	0	0	0	-	0	-	-	-	-	-
Peptide K	2	2	-	-	2	0	2	-	2	-	2	1/2	-	1/2	-	-
Somatostatin	0	-	-	-	0	-	0	0	0/2	0	0	0	0	-	-	-
Substance P	0/3	-	0	1	0	1	0	-	0	-	0	-	0	-	0	-
Vasopressin	0/3	0/5	0	1	0/5	1/5	0/5	1/2	0/1/3	3	0/3	3	0	3	1	3

Table 6.4 Experimental Simplification Factors

Experimental simplification was performed to minimize the number of spots analyzed. This was necessary to rapidly determine if the protocol modifications were effect for deamidation.

Factor	1hr	3 hr	New
Standards	6	8	1
Acids	2	2	4
Temperature	8	8	1
Time	1	1	1
Total Spots	96	128	4

Table 6.5 Four acid comparisons

The MALDI TOF/TOF results for analysis on the 4 acids and AcMeOH are demonstrated below. The initial concentration of the acid is given before the 1:1 by volume incubation with the standard. The m/z of the dominant peak for Orcokinin treated with each acid for each incubation time is reported. The number below the m/z indicates the intensity of the peak. The bolded numbers under the intensities are the ratios of the deamidation peak to the normal monoisotopic peak.

Acid	**Orcokinin standard**	**1 hour**	**3 hour**	**5 hour**
20% FA	1474.2 77317	1474.1 31250	1474.0 25169	1474.0 18550
20% TFA	1474.2 77317	1474.1 20309	1475.0 13617 **1.227332**	1475.1 15115 **1.367997**
20% AA	1474.2 77317	1474.2 24446	1474.2 36889	1474.1 34848
20% TCA	1474.8 79955	1474.6 6007	1474.6 6180	1475.6 6831 **1.066829**
AcMeOH	1474.2 77317	1474.1 50849	1474.1 44112	1474.1 22053

Table 6.6 Established Deamidation Protocol on Eight Standards

The results for deamidation analysis of the peptide standards as individuals and as a mix are given (N=3). Whether or not the deamidation was MS/MS proven is also indicated. The locations of the amides within the peptide sequence are indicated by red and the total expected mass shift for the peptide for complete deamidation is given. In the mixture, for the four peptides that differed in sequence only by the C-terminal amide, it was difficult to determine where the deamidation was occurring so the results were applied to both as indicated by the slash.

Standard	Peptide Sequence	Expected Mass Shift	Single Analysis Shift	MS/MS proven	Mix shifts
Angiotensin I, Human	DRVYIHPFHL	0	0	Yes	0
Bradykinin	RPPGFSPFR	0	-	-	0
CCAP	PFcNAFTGc	1	+1/0	Yes	+1/0
CCAPamide	PFcNAFTGcamide	2	+1	No	+1/0
Orcokinin	NFDEIDRSGFGFA	1	+1	Yes	+1/+2
Orcokinin-amide	NFDEIDRSGFGFAamide	2	+2	No	+1/+2
Peptide K	KHKNYLRFamide	2	-	-	+2
Somatostatin	AGcKNFFWKTFTSc	2	+2	No	+1
Substance P	RPKPQQFFGLMamide	3	+3/+4	No	+3

Chapter 7: iDiLeu Standard Curve Generation for Absolute Quantitation by MALDI Based Analysis

7.1 Abstract

Standard curve calibration is an important approach in MS-based quantitation. Originally developed for LC/MS, a set of five isotopic N, N-dimethyl leucine (iDiLeu) mass tags have been applied to MALDI TOF analysis for absolute quantitation. The features of the tags are: (1) the tags developed are structurally identical with different numbers of deuterium incorporated in them, (2) five mass tags introduce five unique mass shifts (*m/z* 141, 145, 148, 151 and 155) to the analyte standard after labeling, (3) labeled standards can be combined and detected in a single LC/MS run to generate a standard curve. A standard curve generated in MALDI TOF has significant implications for absolute quantitation in several MALDI based applications including tissue imaging and neuropeptide hemolymph analysis.

7.2 Introduction

Due to the complexity of the sample, a liquid chromatography coupled to mass spectrometry (LC/MS) method is often used for neuropeptidomic research. With the advancement of labeling methods, absolute quantitation has become an attractive option for the small quantities of neuropeptides within a sample [1-3]. Standard curve calibration is normally implemented for absolute quantitation in LC/MS since it offers accurate quantitation and the wide dynamic range necessary to quantify the trace amounts of neuropeptide [4].

While standard curve generation is the key to successful quantitation, it was initially a very time consuming task requiring multiple LC/MS runs [5]. For a five point standard curve

with standard deviation error bars, 15 LC/MS runs are needed resulting in an experimental time frame of 300-450 minutes for a typical 30-45 minute LC acquisition [5]. Originally the iDiLeu tags were formulated as a high throughput LC/MS standard curve generation method. The iDiLeu tags are structurally identical with different numbers of deuterium atoms incorporated in them (0, 4, 7, 10, and 14) [5]. Structures, molecular weights, and numbers of deuterium atoms of the synthesized tags are shown in Figure 7.1. However, with the advancement of multiplexing in labeled quantitation, an 8-plex labeling system can generate the five point standard curve while simultaneously analyzing three samples in a single LC MS/MS run [6, 7].

A 1-2 minute shorter retention time was recorded as a result of the large number of deuterium atoms in the two heavier tags (d10 and d14) [5]. This deuterium effect, although a minor inconvenience in generating the LC/MS standard curve, is not a foreseeable problem with either using the multiplex tags or MALDI MS analysis. MALDI analysis of the iDiLeu labeled Allatostatins (ASTs) verified the correlation between peak intensities and concentration and the probability of the tags' application in MALDI analysis (Figure 7.2) [5].

Since hemolymph analysis and tissue imaging experiments are primarily conducted on a MALDI TOF instrument, the use of the iDiLeu tags for absolute quantitation in MALDI became appealing. The mass shifts generated by the tags are detectable in the MS survey mode which adds to their appeal for MALDI quantitation. The following chapter outlines the proof of principle application of the iDiLeu tags for absolution quantitation by standard curve generation in a MALDI analysis.

7.3 Materials and Methods

Materials

A mixture containing 6 peptide standards; the list of the peptides used in this study is displayed in Table 7.1 and were obtained from American Peptide Company (Sunnyvale, CA).

Refer to Appendix C for iDiLeu materials.

Synthesis of 5-plex i-DiLeu Reagents

Refer to Appendix C.

Synthesis of i-DiLeu Triazine Ester

Refer to Appendix C.

iDiLeu Peptide Labeling

Refer to Appendix C.

MALDI Labeling Verification Experiment

A 1:1:1:1:1 mixture of iDiLeu labeled peptide mix was performed on an UltraflexTreme MALDI TOF/TOF analyzer (Bruker, Framingham, MA) equipped with a Smartbeam. For sample spotting, equal volumes of 1 µL sample solution and matrix solution (5mg α-cyano-4-hydroxycinnamic acid (CHCA) in 80% acetonitrile 0.1%FA or 150mg 2,5-dihydroxybenzoic acid (DHB) in 1mL of 1:1 methanol: water) were premixed and 1 µL of the mixture spotted and allowed to dry prior to analysis.

7.4 Results and Discussion

Heavy Labels

Initial analysis of a 1:1:1:1:1 mixture of the iDiLeu labeled peptide mix revealed problems with analysis between the lighter (d0, d4, and d7) and heavier (d10 and d14) tags. The lighter tags provided a 1:1:1 ratio with respect to each other, and the heavier tags also

produced a 1:1 ratio with respect to one another, but this ratio was not maintained across the five tags. Two reasons were hypothesized. The first questioned the labeling protocol and the second the synthesis of the heavier tags. Possibly a higher excess of the heavier tags was required for proper labeling or possibly, given that the initial synthesis of the tags resulted in a brown product, they were inadvertently reduced and required a new synthesis.

Two additional experiments were performed in response to this problem. The first was relabeling the heavier tags in excess while the later involved replacing the heavier tags with the 2-plex formaldehyde labels. Formaldehyde was selected over other isotopic labels because it was used in the synthesis of the iDiLeu labels. The first experiment improved the ratio between the heavy and light labels, but they were still not at a 1:1. The results of the second experiment emphasized the necessity of the labels to have a similar structure for ionization.

Previously reported was the improved ionization efficiencies of the labeled peptides [5]. This was achieved since the iDiLeu coupling to the N-terminus of an peptide changed the overall structure of the amine from a primary to a tertiary, which enhances both the proton affinity and the resulting ionization efficiency [5]. While formaldehyde labeling also converts the labeled peptide to a tertiary amine, the iDiLeu tags generally resulted in a 10-fold increase in signal intensity over the formaldehyde labels. These results emphasis that structure similarity between the tags is important for adequate quantitation.

Shot to shot and spot to spot variability was also dominant in analysis. The saDHB original used for iDiLeu analysis proved inefficient for our analysis [5]. CHCA had to be

utilized to obtain a sample signal from the Ultraflex. CHCA provided a more evenly disbursed crystal formation which aids in quantitative analysis.

Since the ratio between the heavier tags was improved with relabeling in excess, and the color of the heavier tags called the synthesis into question, the heavier tags were resynthesized. A clear solution resulted in the new batch of resynthesized heavier labels and subsequent 1:1:1:1:1 analysis was conducted with the newer labels. This resulted in the desired 1:1 ratio across all the labels with a coefficient of variance less than 11% (Table 7.2).

7.5 Future Work

Since this work was only on generating a 1:1:1:1:1 ratio among the tags, testing the tags at different ratios is necessary to generate a standard curve. Typically, a 1:5:1:5:10 ratio is normally tested next, but for the purposes of generating a standard curve five increasing amounts are needed. Therefore, the limit of detection and limit of quantitation needs to be established for the tags for each peptide used because the limit of detection depends on the ionization efficiency of the peptide and each peptide will have a different ionization efficiency [5]. Once the dynamic range of the tags is established in relation to the range of peptide concentrations the tags can adequately calculate, then the increasing ratios for the tags may be determined to generate a standard curve for absolute quantitation. In the previous study a 1:2:4:8:16 was used for AST-I (Figure 7.2) [5].

For the tags to applicable to hemolymph analysis and tissue imaging experiments it is necessary to determine how effective the tags are for determining quantitation among peptides in family; where sequences are conserved. Whether or not the standard curve of one peptide may be used to determine the concentration of another homologous peptide sequence,

for peptides in the same family, needs to be tested and validated. If such homologous peptide standard curve generation is possible, then a single peptide standard may be used to determine the concentration of the majority of peptides within a family instead of each peptide requiring their own standard curve for quantitation.

Also methods of improving matrix application to facilitate accurate quantitation must also be explored. For reliable and reproducible quantitative analysis, the shot to shot and spot to spot variability must be addressed. Changing the matrix from saDHB to CHCA was met with some success however the shot to shot variability is still unsatisfactory. Possible spraying or sublimation techniques to produce evenly dispersed crystals may be utilized to address this issue. Coupling of the matrix application to the sprayer or sublimation may find more relevance in the method becoming more compatible for MALDI imaging experiments.

Another option is performing analysis on the MALDI orbitrap. While the sensitivity of the instrument will help with quantitation this may emphasize the importance of an evenly distributed matrix application.

7.6 Acknowledgments and Respective contributions

Dr. Feng Xiang is responsible for development of the original application of the iDiLeu tags for which this project is based. All synthesis and labeling of the iDiLeu tags was performed by Tyler Greer. His undergraduate, Anatoliy Nechyporenko, assisted with a few analyses.

7. 7 Works Cited

1. Barr, J., et al., *Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I.* Clin Chem, 1996. **42**(10): p. 1676-1682.
2. Barnidge, D.R., et al., *Absolute Quantification of the G Protein-Coupled Receptor Rhodopsin by LC/MS/MS Using Proteolysis Product Peptides and Synthetic Peptide Standards.* Analytical Chemistry, 2003. **75**(3): p. 445-451.
3. Kirkpatrick, D.S., S.A. Gerber, and S.P. Gygi, *The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications.* Methods, 2005. **35**(3): p. 265-273.
4. Lange, V., et al., *Selected reaction monitoring for quantitative proteomics: a tutorial.* Mol Syst Biol, 2008. **4**.
5. Xiang, F., *Developing novel labeling tools for neuropeptide quantitation and identification by mass spectrometry*. 2011, The University of Wisconsin - Madison: Ann Arbor. p. 251.
6. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat Protoc, 2009. **4**(4): p. 484-94.
7. Pottiez, G., et al., *Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples.* J Proteome Res, 2012. **11**(7): p. 3774-81.

7. 8 Figures and Tables

Figure 7.1

a) Deuterium atom combinations of reduction dimethylation reaction; b) structures, number of locations of deuterium atoms and molecular weights of the 5-plex iDiLeu reagents [5].

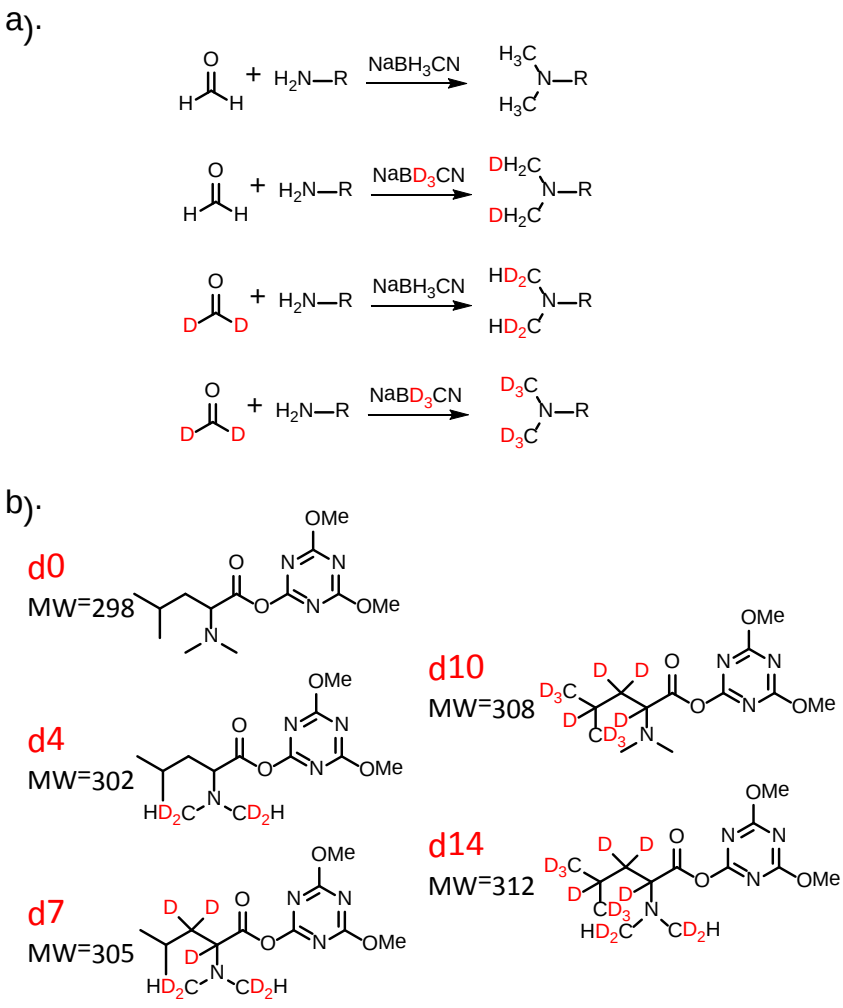


Figure 7.2

The correlation between peak intensities and labeled AST-I concentrations by LC/MS (insert) and MALDI TOF analysis [5]. The concentrations of the peptide to generate this curve are 1.5×10^{-6} M, 7.48×10^{-7} M, 3.74×10^{-7} M, 1.87×10^{-7} M and 9.35×10^{-8} M (a 1:2:4:8:16 ratio).

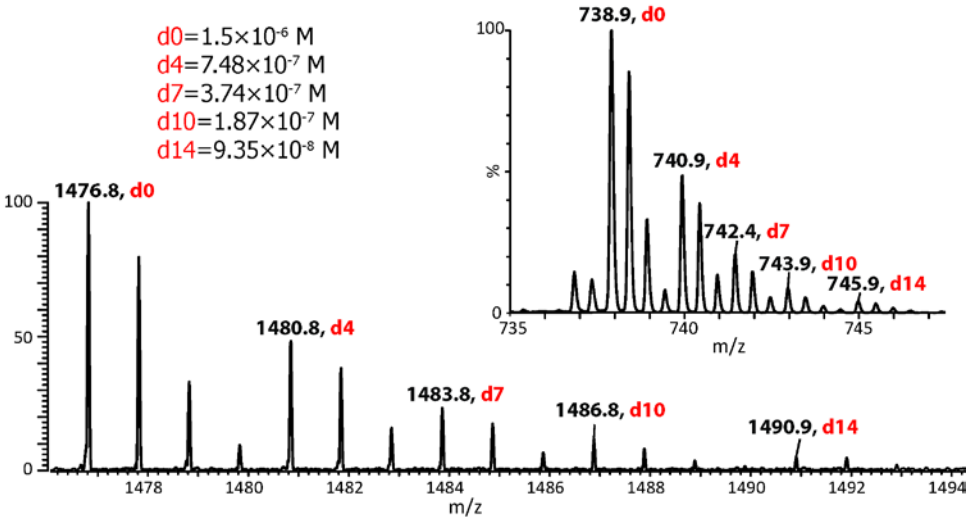


Table 7.1

The 6 peptide standards used for labeling analysis are listed below. The mass for the labeled peptides are also given for each tag. The respective mass shift generated by labeling is given in parenthesizes for each tag. Since Somatostatin 14 also contains two lysines, which are labeled by these tags, the masses for the partial (1 and 2 label) and fully labeled (3 label) peptide are all listed.

Standard	MW	Peptide Sequence	FH2 (28)	FD2 (32)	D0 (141)	D4 (145)	D7 (148)	D10 (151)	D14 (155)
FMRF-amide	598.8	FMRFa	626.8	630.8	739.8	743.8	746.8	749.8	753.8
Allatostatin III	899	GGSLYSFGLa	927	931	1040	1044	1047	1050	1054
FMRF amide-like peptide I, lobster	1053.3	SDRNFLRFa	1081.3	1085.3	1194.3	1198.3	1201.3	1204.3	1208.3
Allatostatin I	1335.5	APSGAQRLYGFGLa	1363.5	1367.5	1476.5	1480.5	1483.5	1486.5	1490.5
Bactenecin, bovine	1483.9	RLCRIVVIRVCR	1511.9	1515.9	1624.9	1628.9	1631.9	1634.9	1638.9
Somatostatin 14 (1 label)	1637.74	AGCKNFFWKTFTSC	1665.74	1669.74	1778.74	1782.74	1785.74	1788.74	1792.74
Somatostatin 14 (2 label)	1637.74	AGCKNFFWKTFTSC	1693.74	1701.74	1919.74	1927.74	1933.74	1939.74	1947.74
Somatostatin 14 (3 label)	1637.74	AGCKNFFWKTFTSC	1721.74	1733.74	2060.74	2072.74	2081.74	2090.74	2102.74

Table 7.2 The 1:1:1:1:1 Ratios

The ratios obtained by labeling with the resynthesized heavier tags are calculated below from the intensity of the tags, d0 is assigned as control for ratio calculation.

FMRF-amide						
	d0	d4	d7	d10	d14	Average
Intensity	116227	106026	93502	90379	99778	101182.4
Ratios	1	0.912232098	0.804477	0.777608	0.858475	std dev
Ratio to add	1.00	1.10	1.24	1.29	1.16	10337.73
Error from average	0.148687914	0.047869985	0.075906	0.106772	0.01388	

FMRF amide-like peptide I						
	d0	d4	d7	d10	d14	Average
Intensity	30502	31602	28841	27487	26701	29026.6
Ratios	1	1.036063209	0.945545	0.901154	0.875385	std dev
Ratio to add	1.00	0.97	1.06	1.11	1.14	2038.55
Error from average	0.050829239	0.687672955	0.71496	0.728342	0.73611	

Allatostatin I						
	d0	d4	d7	d10	d14	Average
Intensity	32316	36551	40623	37294	41566	37670
Ratios	1	1.131049635	1.257055	1.154041	1.286236	std dev
Ratio to add	1.00	0.88	0.80	0.87	0.78	3672.80
Error from average	0.142129015	0.029705336	0.078391	0.009981	0.103424	

Chapter 8: Sample Collection Methods for Improving Detection of Circulating Neuropeptides by MALDI Based Analysis for Application in a Circadian Rhythm Study

8.1 Abstract

In 2007 and 2009 Bruker released the AutoFlex III Smartbeam and the UltrafleXtreme MALDI –TOF/TOF mass spectrometers, respectively. Normally, the application of established analytical techniques on newer instruments improves analysis. Sometimes however, the analysis by the newer instrument may result in poorer or no data. Transferring the analysis of our hemolymph protocol from the 4800 MALDI TOF/TOF to the newly acquired Bruker MALDI TOF/TOF instruments was met with limited success. This research was initiated to improve the detection of neuropeptides from a hemolymph sample on the Bruker MALDI TOF/TOF instruments for application in a circadian rhythm analysis.

8.2 Introduction

Hemolymph analysis allows for multiple samples to be collected from a single animal; which negates animal to animal variability in different animals and follows the reduction rule of animal models [1]. It also has a higher probability of collecting signaling molecules then tissue analysis [2]. Despite these advantages, neuropeptide analysis from hemolymph samples is challenging.

Hemolymph flows through an open-circulatory system and has a high salt concentration [2]. Also present in hemolymph are large hemocyanin and cryptocyanin proteins that degrade rapidly when hemolymph is extracted; the resulting peptide fragments mask peptide identification [2, 3]. Hemolymph also clots rapidly after removal and this clotting has been demonstrated to trap peptide content [4]. To address these issues, both chemical and physical

means (i.e. non-clotting reagents, protease inhibitors, and high velocity centrifugation) have been previously employed to maximum neuropeptide recovery from hemolymph samples [2-7].

Circadian rhythm effects, in reference to neuropeptide expression, have been previously reported for the crustacean model [3, 8]. An increase in the concentration of crustacean hyperglycemic hormone (CHH) has been previously reported in the first hour of darkness while its levels remain relatively low during light hours [9]. Also, since crustaceans are nocturnal, more myotrophic neuropeptides are theorized to increase in expression as a result of the crab's increased locomotive activity [5].

Previously, our hemolymph analysis was performed on the 4800 MALDI TOF/TOF located in the Biotech Center (Figure 8.1 A) [2]. However, we were fortunate enough to obtain a demo instrument, Bruker AutoFlex III Smartbeam, for several months before opting to purchase the Bruker UltrafleXtreme mass spectrometer. While we had the AutoFlex, we noticed that the data analysis of any hemolymph sample was not possible despite the results detected from the 4800 analysis (Figure 8.1 B). In this chapter, I attempt to improve detection of neuropeptides extracted from hemolymph samples by the Bruker MALDI TOF/TOF mass spectrometer using basic techniques involving sample collection modifications. The primary goal is the application of the revised protocol to investigate changes in neuropeptide expression in response to day and night cycles. Such a study will aid in improving our knowledge on neuropeptide dynamic changes in response to circadian rhythm.

8.3 Materials and Methods

Refer to Appendix A for SOPs on standard handling procedures. The Jonah crabs for this experiment were obtained from Ocean Resources, Inc. They were allowed to acclimate to their

tank for a month before circadian rhythm analysis and three days before regular protocol modification analysis. The standard protocol workflow for hemolymph analysis is described in Appendix A. Described below are modifications made to the hemolymph collection.

Non-clotting Reagents and Hemolymph Volume Concentration Analysis

A modified version of previous non-clotting experiment is conducted to determine the effects of two non-clotting solutions on neuropeptide extraction (Figure 8.2) [4]. A volume of 5 mL hemolymph is withdrawn from a crab and immediately separated into the following 4 volumes for analysis (2-0.5 mL, 1 mL, and 2 mL). One 0.5 mL sample will serve as the control while the other 3 volumes are immediately added to tubes already containing equal volumes of acidified methanol or sodium citrate solution (0.11 M Na citrate, 0.1 M NaCl) [4].

The samples remained on ice for 20 minutes. The hemolymph extraction protocol as described in Appendix A is then followed. The desalting step was performed twice on the sodium citrate solutions. In additional experiments, the desalting step was only performed once and a published modified molecular weight cut-off (MWCO) protocol applied [10].

Circadian Rhythm Experimental Design

Hemolymph was collected from the crab one hour before (light cycle) and one hour after (dark cycle) lights dimmed in the animal room; the low beam red cycle began. Hemolymph from the 3 light and 3 dark cycles were collected over a course of two weeks. Collection of a light cycle was paired to collection of a dark cycle to ensure a full set was collected in case of crab death. Which cycle was collected first was alternated for each of the three experimental replicates. No saline replacement was performed despite 0.5 mL collections being performed for four days, but feeding occurred after each set of light/dark cycles was complete.

8.4 Results and Discussion

Determining Hemolymph Extraction Volume

It was hypothesized that the concentration of neuropeptides in the hemolymph sample was too low to be detected on the AutoFlex. Two areas of focus were analyzed to increase the neuropeptide concentration during initial hemolymph collection to improve detection and subsequent analysis. For the first focus, it was hypothesized that more hemolymph should contain more neuropeptides; therefore different volumes of hemolymph were collected from the crab for analysis.

For the second focus, it had previously been determined that the clotting behavior of hemolymph traps neuropeptides preventing their extraction and analysis [4]. Previously, our lab had implemented a Protease inhibitor to prevent this from occurring, however it was our goal to utilize a methodology that was more mass spectrometry friendly given that the Protease inhibitor contained various polymers [2]. We therefore utilized simple non-clotting procedures that just required immediate incubation of the hemolymph with the non-clotting solution [2, 4].

Since hemocyanin degradation products could mask neuropeptide signal, we also initially looked into methods for simple hemocyanin removal before neuropeptide extraction. The simplest method found required centrifugation at 170,000g for 1 hour [11]. Hemocyanin removal was abandoned and instead we focused on preventing the degradation of the hemocyanin by using non-clotting solutions and keeping the sample on ice.

In the initial experiments testing the non-clotting reagents and hemolymph volumes, hemolymph was withdrawn from the 3rd appendage of the crab and aliquoted into different volumes. These studies indicated that more neuropeptides were identified by the 0.5 mL

samples than the larger volumes. However, when the experiments were repeated by individually removing 0.5 mL, 1 mL and 3 mL volumes instead of creating aliquots from a 5 mL sample, nothing was seen in the 0.5 mL spectrum (Table 8.1 A). It was speculated that despite the crab having an open circulatory system, neuropeptides were not evenly distributed among the hemolymph. To test this theory, both 1mL and 3mL volumes of hemolymph were withdrawn and a 0.5 mL aliquot used for analysis. It was confirmed that a larger hemolymph volume resulted in better neuropeptide content (Table 8.1 A).

To minimize the amount of stress this would have on the crab for long term analysis, a new location was chosen to facilitate more neuropeptides being removed in the collection. Collection was moved to the tissue between the 3rd and 4th appendage. The number of neuropeptides detected was greater in this area (Table 8.1 B).

Although the change in location increased the detection of neuropeptides in lower collection volumes, higher volumes were still required to obtain robust detection (Figure 8.3). It was determined, by literature, that crustaceans have approximately 22 mL of hemolymph [6, 12]. Following rodent protocols for 10% analysis, 2mL of hemolymph was initially assigned to begin circadian rhythm analysis. However, after the second extraction the experiment was terminated for humane purposes; the crabs became overly aggressive resulting in the death and cannibalism of 1/3 of the study population overnight.

The experiment was redesigned and repeated with a 0.5 mL volume of hemolymph extracted for analysis and scheduled feedings after every other extraction to prevent aggression among the population. Analysis of the three sets of day and night hemolymph resulted in very poor data. The samples underwent formaldehyde labeling (Appendix A) in hopes of improving

ionization and detection of the neuropeptides. No peak pairs were detected above m/z 556 which correlates to the m/z for the smallest neuropeptide in our 692 entry database but several peaks were detected below it.

Although analysis of neuropeptides with this procedure was not possible, it must be stated that metabolites were seen in the sample (Table 8.2). And they were subsequently labeled and seen in the mass range m/z 450-2000 used for the formaldehyde labeled circadian rhythm analysis.

8.5 Future Focus

It is speculated that neuropeptides are being lost in the MWCO step due to the clogging of the spin column. The MWCO step is performed to (1) remove higher mass proteins from the sample which helps remove proteins responsible for clotting and degradation and (2) aid in desalting, otherwise the sample is too viscous. This theory is strengthened by the fact that metabolites are prevalent in analysis. Before any further hemolymph analysis is pursued, it is necessary to determine where the loss of neuropeptides is occurring, if at all. This analysis may be completed by (1) running the samples on the 4800 MALDI TOF/TOF or MALDI Orbitrap for comparison and (2) spiking standards into one of four aliquot hemolymph samples before each of the three steps (NP extraction, MWCO, and desalting); leaving the four aliquot as a control.

It would also be beneficial to explore alternative desalting methods for processing a high sample volume. Once the cause of the sample loss is determined further steps maybe considered, but the overall goal is to apply the technique to the analysis of neuropeptide secretion changes in relation to day-night cycle on the crab.

Since metabolites were detected, it might also prove interesting to continue analysis on the metabolites identified by this study.

8.6 Respective Contributions and Acknowledgement

The initial concentration and non-clotting solutions studies were performed by Santhoshini Ramani on the AutoFlex. Stephanie Rawson performed the extraction protocols to determine the optimum location for hemolymph withdrawal on the AutoFlex and also concentration analysis on the UltrafleXtreme. Ayrel Lyra also performed neuropeptide extraction for analysis on the UltrafleXtreme. This project was inspired by Claire Schmerberg because of the difficulties she was having with analysis of her microdialysis samples. Claire also aided in the initial structure of this project and aided with fruitful discussion.

8.7 Works Cited

1. *Replacement, Reduction and Refinement of Animal Experiments in the Development and Control of Biological Products. Proceedings of a conference. Langen, Germany, November 2-4, 1994.* Dev Biol Stand, 1996. **86**: p. 1-368.
2. Chen, R., et al., *Measurement of neuropeptides in crustacean hemolymph via MALDI mass spectrometry.* J Am Soc Mass Spectrom, 2009. **20**(4): p. 708-18.
3. Terwilliger, N. and K. Dumler, *Ontogeny of decapod crustacean hemocyanin: effects of temperature and nutrition.* J Exp Biol, 2001. **204**(Pt 5): p. 1013-20.
4. Kwok, R. and S.S. Tobe, *Hemolymph clotting in crustaceans: Implications for neuropeptide extraction from invertebrate hemolymph.* Peptides, 2006. **27**(3): p. 590-596.
5. Luschen, W., et al., *Isolation, sequence analysis, and physiological properties of enkephalins in the nervous tissue of the shore crab Carcinus maenas L.* Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8671-5.
6. Martin, J.-L., A. Van Wormhoudt, and H. Ceccaldi, *Zinc-hemocyanin binding in the hemolymph of carcinus maenas (crustacea, decapoda).* Comparative Biochemistry and Physiology, 1977. **58**: p. 193-195.
7. Cytrynska, M., et al., *Purification and characterization of eight peptides from Galleria mellonella immune hemolymph.* Peptides, 2007. **28**(3): p. 533-46.
8. Strand, F.L., *Neuropeptides : regulators of physiological processes*. Cellular and molecular neuroscience. 1999, Cambridge, Mass.: MIT Press. 658 p.

9. Fanjul-Moles, M.L., *Biochemical and functional aspects of crustacean hyperglycemic hormone in decapod crustaceans: Review and update.* Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 2006. **142**(3–4): p. 390-400.

10. Cunningham, R., et al., *Investigation and reduction of sub-microgram peptide loss using molecular weight cut-off fractionation prior to mass spectrometric analysis.* J Mass Spectrom, 2012. **47**(10): p. 1327-32.

11. Harris, J.R. and J. Markl, *Electron microscopy and biochemical characterization of a 350-kDa annular hemolymph protein from the keyhole limpet Megathura crenulata.* European Journal of Biochemistry, 1994. **225**(2): p. 521-528.

12. Gleeson, R.A. and L.Z. Paul, *The determination of hemolymph volume in the blue crab, Callinectes sapidus, utilizing 14C-thiocyanate.* Comparative Biochemistry and Physiology Part A: Physiology, 1977. **56**(3): p. 411-413.

8.8 Figures and Tables

Figure 8.1 MALDI TOF/TOF Comparison

The top spectrum (A) is representative of the results collected with analysis of the standard hemolymph extraction on the 4800 MALDI TOF/TOF mass spectrometer [2]. The bottom spectrum (B) is representative of the results collected with analysis of the standard hemolymph extraction on the Bruker AutoFlex III Smartbeam mass spectrometer.

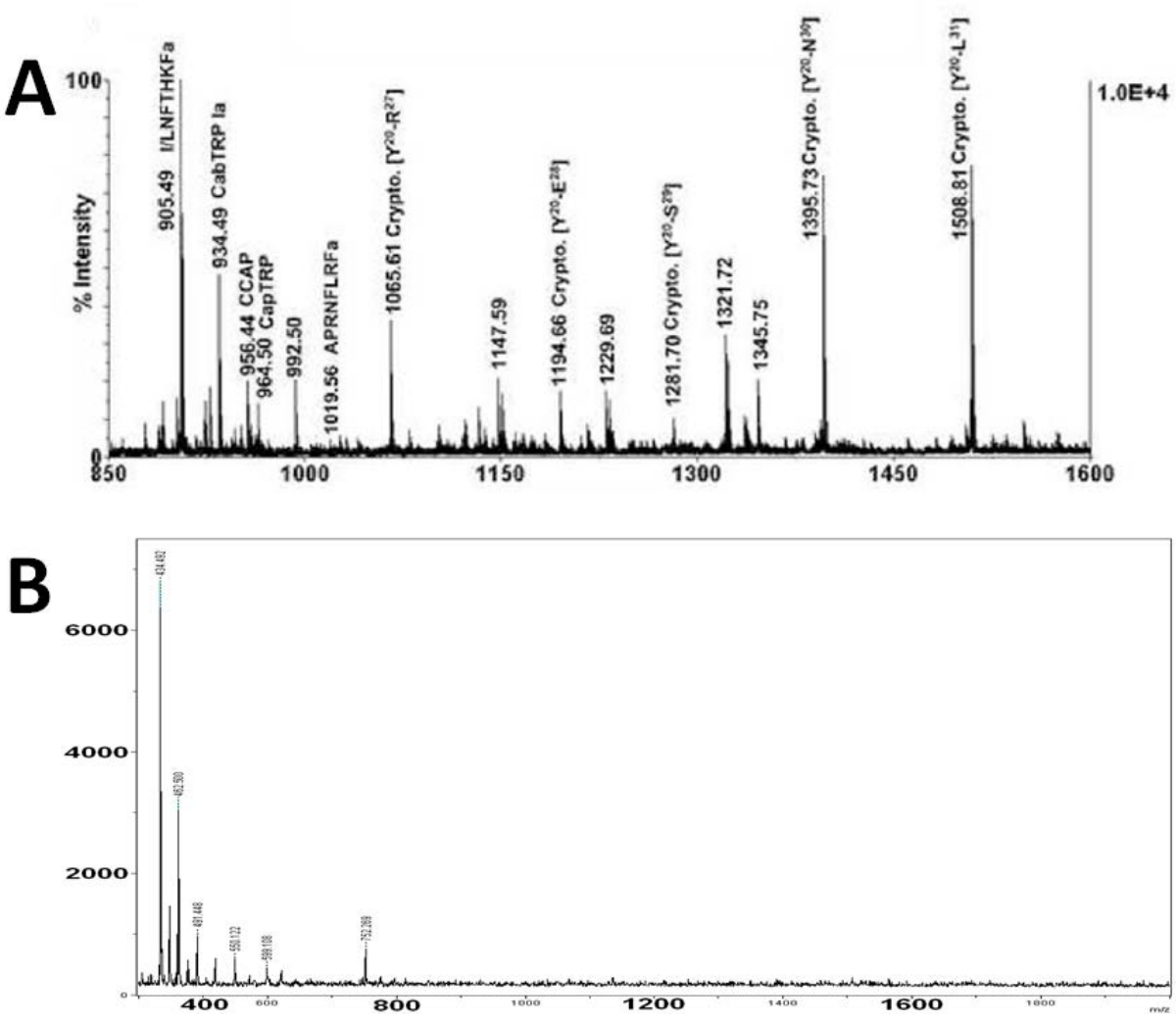


Figure 8.2 Non-clotting and concentration factor experiments

A figure outlining the experiments conducted to improve the detection of neuropeptides extracted from a hemolymph sample by non-clotting solutions and hemolymph concentration.

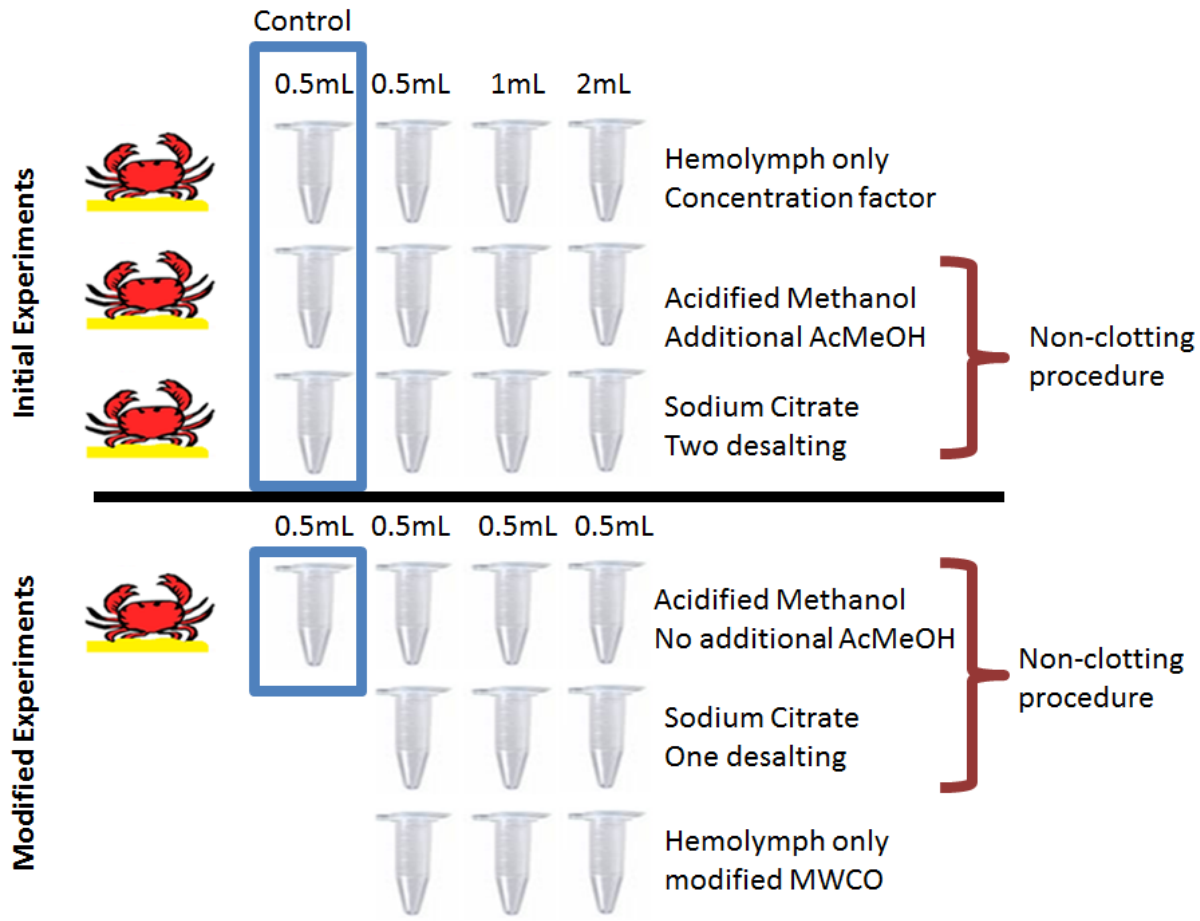


Figure 8.3 Volume Analysis at new location

The mass spectra for 1mL and 3mL (A and B respectively) drawn from between the 3rd and 4th appendages. Identification of the neuropeptides is given in the blue table of Table 8.1.

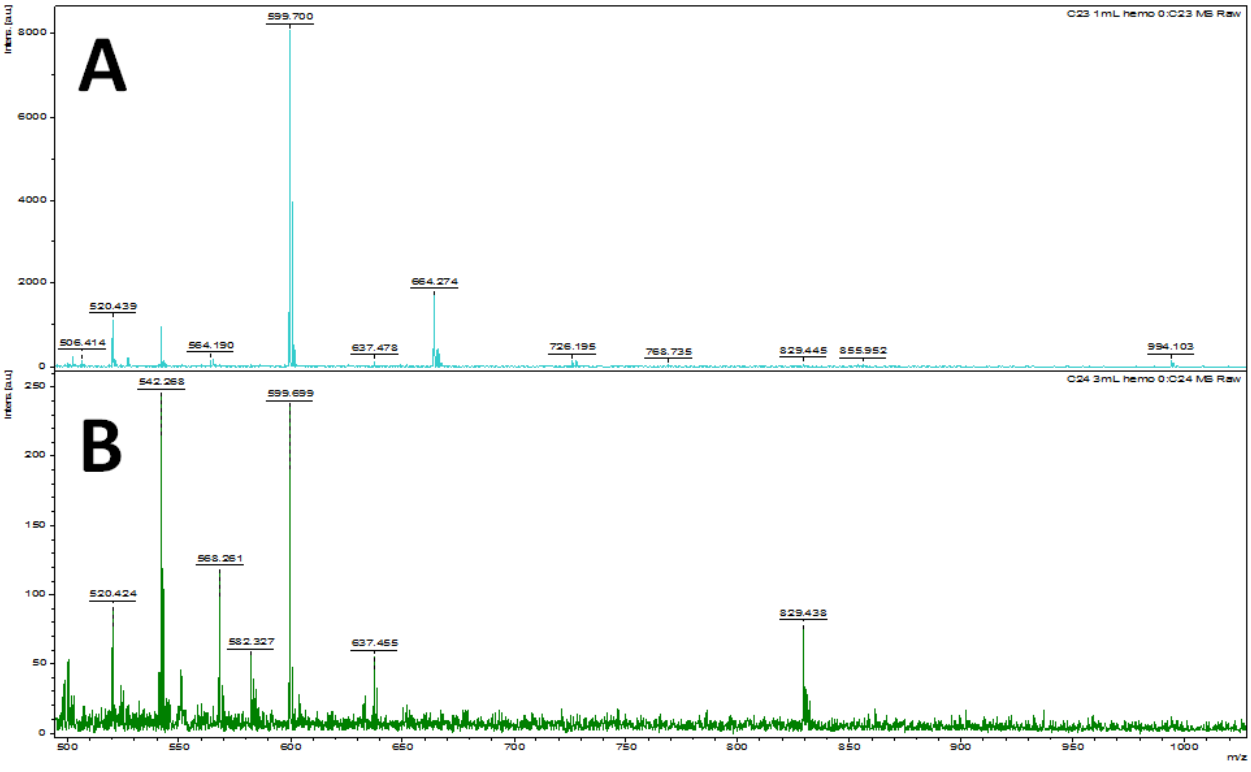


Table 8.1 Improved Neuropeptide Detection by Changing Location

The identification for the neuropeptides collected either at the leg joint (grey table) or between the 3rd and 4th appendages (blue table). Only neuropeptides present in more than one technical replicate were identified. Also given is the frequency of their detection and in which volume of hemolymph they were detected.

Family	Name	M+H	Sequence	0.5mL	1mL	2mL
Allatostatin A-type	Penaeustatin 30	599.31880	YTFGLamide		2	5
Orcokinin		1393.67940	ERDMHRFSFGLamide		2	
			2		4	5

Family	Name	(M+H)	Sequence	0.5mL	1mL	2mL
	Penaeustatin 17	599.31880	YTFGLa	3	3	2
AST A-type	Carcinustatin 1, Penaeustatin 13	569.30821	YAFGLa		2	
		1052.58870	SPRLTYFGLa	1	1	
CPRP		1415.79000	RSAEGLGRMGRLL	2		1
	truncated Hoa CPRP-C 20-33	1531.72100	SMPLGFLSQDHSVN	1	1	
Cryptocyanin	cryptocyanin fragment CC1 20-29	1281.68376	YKIFEPLRES	1		1
Tachykinin		950.48765	APSGFLGM(O)Ra	2		
			7	6	4	3

Table 8.2 Metabolites Identified

The metabolites identified from a single mass spectrum collected from 1 mL and mass matched from a list of compiled from the **M**ouse **M**ultiple tissue **M**etabolome **D**ata**B**ase (MMMDB). The mass of the peaks highlighted in purple also correspond to matrix peaks.

Metabolite Name	m/z	#	3mL	1mL	0.5mL
1-Methyl-2-pyrrolidone	100	3	1	1	1
Isonicotinamide/nicotinamide	123	1	1	0	0
o-Hydroxybenzoate/Hypoxanthine/1-Methylnicotinamide	137	3	1	1	1
His	156	2	1	1	0
2-Oxoadipate/Allantoin/Pimelate	159	2	0	1	1
Terephthalate/Lumazine	165	1	0	1	0
N-Formylmethionine	176	3	1	1	1
Serotonin	177	3	1	1	1
DOPA	198	1	1	0	0
Erythrose 4-phosphate	199	3	1	1	1
Cytidine	244	2	1	1	0
Uridine	245	2	1	1	0
Pyridoxal 5-phosphate	246	2	1	1	0
2'-Deoxyguanosine	268	1	1	0	0
Galacturonate 1-phosphate	273	3	1	1	1
Ophthalmate	290	2	1	1	0
cAMP	328	3	1	1	1
GMP	362	1	0	1	0
NADH	664	1	0	1	0
	19	39	15	16	8

Chapter 9: Conclusions and Future Focus

9.1 Conclusions

In this work, several projects were conducted with the overall focus of improving crustacean neuropeptide analysis. Mass spectrometry was chosen because of its capability of analyzing low concentration analytes in complex biological mixtures. Ironically, several of the projects reported were initiated because of the disadvantages of mass spectrometric techniques with the goal of improving analysis.

Chapter 2 provides a brief summary of the limitations of bioinformatics with emphasis on database compilation. This background chapter also explains the use of crustaceans as model systems in our research.

Chapter 3 describes the compilation of our database that is currently being used for automated peptide identification in our lab. Originally containing 245 entries, the database was upgraded and feasibly expanded to 692 entries. It was also incorporated into a larger homologous database, resulting in 1958 entries, for software compatibility. This 1958 entry database was used to compare three commercially available protein identification and quantitation software using the large dataset collected in Chapter 4.

Chapter 4 describes the semi-quantitative analysis using the 4-plex N,N dimethyl leucine (DiLeu) labels to determine neuropeptide expression level changes in relation to feeding. The application of a complex labeling matrix determined the robust labeling nature of the tags and validated the quantitation results for future quantitative application. The analysis of the dataset by three quantitative software demonstrated that PEAKS and Mascot are complementary software. The Mascot results also revealed several potentially novel neuropeptides in the Jonah

crab. Also, several additional neuropeptides were identified in relation to feeding by DiLeu labeling then the previous formaldehyde labeling study performed by our lab.

Chapter 5 explores the production of multiple charges in the matrix assisted laser desorption/ionization (MALDI) process. The chapter demonstrates that production of multiple charges in peptides with MALDI analysis is possible by simple sample preparation techniques and that fragmentation of these multiply charged ions is possible. The best protocol for producing multiply charged bradykinin ions was when the peptide was mixed with nicotinic acid matrix solution and plated on top of a crystal grown in saDHB solution.

Chapter 6 develops a simple acid-driven deamidation protocol that determines the presence and location of amides in peptide sequence. This technique is also shown to be potentially useful in opening pyroglutamate rings for quantitative labeling analysis.

Chapter 7 demonstrates the feasibility of using isotopic DiLeu (iDiLeu) labels to generate a five point standard curve for absolute quantitation with emphasis on MALDI imaging techniques. Chapter 8 demonstrates the modification of the current hemolymph protocol for MALDI TOF/TOF analysis on the Bruker AutoFlex and UltrafleXtreme with reference to circadian rhythm analysis. It demonstrates that the current protocol is better for detection of metabolites than neuropeptides.

9.2 Future Focus

The work presented in this dissertation provides a solid foundation for future mass spectrometry based analysis. However, several aspects of this work still require varying degrees of focus to develop a complete analysis of crustacean neuropeptides. The top results from the DiLeu data suggest the existence of several peptides from homologous species could exist within

our crustacean model. This needs to be explored given that a third of the hits were from the homologous species database. Since protein identification and quantitation software was utilized, the sequence homology among peptide families resulted in the data still requiring manual analysis for validation of the hits. Therefore more information may be acquired from a more, in-depth analysis of the datasets instead of solely focusing on the tops hits that matched our in-house database. Also, despite being up-to-date, we must still acknowledge that our database is incomplete. It would be greatly beneficial to obtain a completed genome for our model systems.

The acid driven deamidation study could still benefit from more liquid chromatography (LC) electrospray ionization (ESI) analysis; however it is ready for application. This has great potential for elucidating the sequences of crustacean hypoglycemic hormone (CHH) and CHH-protein related peptides (CPRP) families. In our database alone, there are 187 peptide entries under these two families and they contain a total of 403 amides from either the presence of asparagine or glutamine. Also the acid-driven cyclization opening of pyroglutamate needs to be optimized for use with labeling quantitation. This will affect quantitation on 8% of the peptides in our 692 entry database with potential application for Down syndrome and Alzheimer's quantitation research.

The multiply charged MALDI research still requires that multiple charged ions be represented as the dominant species in the spectrum. Once this is feasible with single peptide analysis, ways to expand this technique for production of abundant multiple charges in peptide mixtures and biological samples need to be pursued.

While proof of a 1:1:1:1:1 ratio was shown with the iDiLeu labels, future studies need to analyze standard curve generation, dynamic range of the curve with MALDI analysis through limit of detection and limit of quantitation, and the implication of quantification on homologous species.

The hemolymph analysis protocol still requires modification in the sample preparation steps for peptide identification. Currently however, it could be applied to metabolite analysis. It might also prove beneficial to ensure that the presence of the signals in the lower mass range is actually from metabolites and not peptide degradation.

These projects are not presented as a completed analysis of crustacean signaling peptides, but instead as a foundation for further development. This work strengthens several aspects of the mass spectrometric platform by building upon previous knowledge while simultaneously pushing the forefront of research. Several simple protocols are established for easy application, rapid analysis, and reproducibility to be implemented into future studies.

Appendix A: Standard Operating Procedures for Animal Care, Sample Collection and Preparation, and Instrumental Analysis

A.1 Introduction

The following appendix is a compilation of protocols used for the research presented in this dissertation. It includes the Stand Operating Procedures for housing and dissection of the crustaceans, the general protocol used for neuropeptide analysis on tissue and hemolymph samples. For the protocol regarding DiLeu and iDiLeu, please refer to Appendix C.

A.2 Standard Operating Procedures

The standard operating procedures (SOPs) are attached at the end of this appendix. They were written and revised during the preparation for the AAALAC reaccreditation visit in October 2011. Claire Schmerberg formatted and primarily authored the documents. Chuanzi OuYang recently completed the SOP regarding nanoparticle exposure research.

The SOPs only include those for Jonah and blue crabs since those are the species in primary use by our lab, however our lab has worked with other crustacean species. As new species are investigated by our lab, the SOPs will require updating. Also, SOPs should be formulated for research conducted that exceeds the scope of the current SOPs on file in the animal facility.

A.3 Hemolymph Analysis

The following describes the collection and sample preparation required for neuropeptide analysis of a hemolymph sample in the circadian rhythm analysis described in Chapter 8. It is adapted from both the SOPs attached at the end of this appendix and Claire's thesis [1]. In

general, hemolymph withdrawal is sequentially followed by neuropeptide extraction, molecular weight cutoff (MWCO), desalting and, when required, formaldehyde labeling.

Hemolymph Withdrawal

Hemolymph was obtained from Jonah crabs via the SOP for Marian Crustacean Dissection with the following modifications.

Materials

Needle (23gauge, 1" long, Beckton Dickinson and Co. (BD), Franklin Lakes, NJ), plastic 1 mL syringe (BD), crab saline (440 mM NaCl, 26 mM $MgCl_2$, 13 mM $CaCl_2$, 11.2 mM Trizma base, 11 mM KCl, 5.1 mM maleic acid, pH adjusted to 7.4), acidified methanol (90:9:1 MeOH:H_2O:Acetic acid)

Instructions

1. Prepare the needle and syringe. If obtaining more than 0.5mL or removing 0.5ml from the crab over multiple consecutive days, it is necessary to replace the volume of hemolymph taken with crab saline. Prepare this needle with enough saline to equal the volume of hemolymph being removed. Loosen the cap on the needle without completely uncapping the needle since the rest of the procedure will be done one handed. Preparing everything beforehand (both the saline replacement needle and hemolymph withdrawal needle) will quicken the procedure causing less stress to the crab and preventing injury to the researcher. Smaller gauge needles have a slight tendency to clog, but minimalize the damage to the animal while larger needles do not clog as easily but are more likely to cause internal injury.
2. Remove the crab from the tank and place it on its carapace with the ventral side facing up in a metal dissecting pan, holding the side of the claw where it bends to immobilize the crab and

prevent it from pinching you. A crab cannot reach directly across from itself and will move in response to being poked with a needle.

3. Uncap the needle and insert at one of the soft joints. While the optimal location for a hemolymph withdrawal varies depending on species, amount of hemolymph needed, and experimenter preference my experience proved that the best location is between the 3rd and 4th legs where they join the body (Figure A.1). This location closer to the area where the heart, pericardial organs, and thoracic ganglion are located; which is an area richer in neuropeptide content despite an open circulatory system. Regardless of placement, insert the needle into the soft tissue, and withdraw the plunger on the syringe to create a vacuum. With a larger needle you must be mindful of the internal organs of the crab, especially the pericardial ridges. Ease the needle in and switch angles if you hit a something hard. DO NOT punch through any obstacles. This will create internal damage lessen the survival of the crab during and after the procedure and create a more stressful response in the crab that will lead to questionable results especially in quantitative analysis. Gently change the orientation of the needle inside the body until hemolymph is observed to enter the syringe. Hemolymph is clear, slightly greyish in color. Once a good location is obtained, the syringe and needle are held still until either the desired amount of hemolymph is collected or the flow of hemolymph into the syringe ceases, even with additional pressure applied via the plunger. Multiple locations may need to be used to obtain enough hemolymph.
4. Add hemolymph to equal volume acidified methanol and keep on ice. The resulting mix of hemolymph and acidified methanol will create a cloudy white solution. The sooner this step

is performed the better, acidified methanol is both used to extract neuropeptide content and prevent clotting [2].

5. Inject the replacement saline solution into the crab at this point. Removing 0.5 mL of hemolymph or more each day will result in the crab's death after 3-4 days, so if this volume or more is taken, it is recommended to inject an equal volume of crab saline into the animal to replace lost volume. Repeated withdrawals of large volumes (more than 2mL) of hemolymph will result in the crabs' death even if the volume is replaced (Table A.1). The estimated total hemolymph volume in a crab is 22 mL [3].

Hemolymph Neuropeptide Extraction

The following protocol was slightly modified from Claire's thesis [1]. Neuropeptide extraction, MWCO and desalting were performed on hemolymph samples followed by formaldehyde labeling when necessary.

Materials for Hemolymph Extraction

Materials: Hemolymph, acidified methanol (90:9:1 MeOH:H_2O: acetic acid), solution for resuspension, microtubes, plastic pestles (optional), sonicator, microcentrifuge, speedvac, pipettes and tips

Note: Keep samples on ice whenever possible.

Neuropeptide Extraction Steps

1. Collect hemolymph add equal volume of acidified methanol and spin down (5 min x 16,100g).
2. Remove supernatant and set aside in a new tube.

3. Add a 0.5 mL acidified methanol to the pellet and reconstitute; vortexing and/or a plastic pestle can be used to aid in breaking up the pellet.
4. Sonicate the tube for 5 min. Do not use heat.
5. Centrifuge 5 min x 16,100 g to separate liquid and solid portions.
6. Repeat steps 2-5 twice (for a total of 3 extractions) and combine the supernatants. The pelleted solids can be saved, but it has been determined that the first and second extractions of most tissues contain the most neuropeptides. Subsequent extractions contain primarily lipids and/or no compounds of interest.
7. Evaporate the solvent using the speedvac. Medium or Low heat is recommended.
8. Reconstitute in 0.5mL of 70/30 H_2O/MeOH prior to storage at -20°C. It is important to reconstitute prior to freezing, because it has been found that hemolymph samples frozen before resuspension will not achieve full dissolution.

Molecular Weight Cutoff (MWCO)

Hemolymph samples were usually subjected to molecular weight cutoff (MWCO) following a combination of the manufacturer's recommended procedure and a published modification [4]. This step is conducted on the crude extracts prior to desalting or other steps. If MWCO was not conducted, the hemolymph samples were typically too thick for efficient desalting.

MWCO Materials

Amicon Utracel 10 kDa MWCO 0.5mL size, 70/30 H_2O/MeOH, 50/50 H_2O/MeOH, 0.1M NaOH, tubes, centrifuge, speedvac, pipette and tips

MWCO Procedure

1. Rinse MWCO device with 0.2mL 0.1M NaOH: add solution to tube and spin 4 min x 14,000g. Ensure that the solution flows through the membrane before completing the next step.
2. Rinse MWCO with 0.5 mL of 50/50 H_2O/MeOH: spin 8 min x 14,000g. Ensure that the solution flows through the membrane before completing the next step. Steps 1 and 2 are collected in the same collection tube. A new collection is required for the following step.
3. Load sample, dissolved in 0.5mL 30/70 MeOH/H_2O (sonicate if necessary to reconstitute), into device, and run through the membrane. This will take 25 min x 14,000g.
4. Rinse sample tube with 0.1 mL of 30/70 MeOH/H_2O, and run through the membrane. This will take 5 min x 14,000g. This rinse is combined with the flow-through in step 3.
5. Both the flow-through and concentrate were saved, but only the flow-through was used for neuropeptide analysis.
6. Evaporate the solvent for the flow-through using a speedvac (Low/Med heat).
7. Reconstitute the sample in a minimal volume of 0.1% FA in water, up to 60 µL for desalting. Sonicate 5 min if necessary. The sample may be frozen at this step for desalting at a later time.

Desalting

Hemolymph extracts were subjected to a desalting step, especially prior to analysis via MALDI ionization or formaldehyde labeling.

Desalting Materials

C18 Zip Tip pipette tips, 10 µL size (Millipore), acetonitrile, 0.1% FA in H_2O, 50/50/0.1 ACN/H_2O/FA, tubes, speedvac, pipettes and tips.

Desalting Procedure

1. Wet the C18 bed in the tips with 15 µL of acetonitrile (ACN). (Draw up and expel to waste)
2. Equilibrate the tips with 15 µL of 0.1% FA in water. (Draw up and expel to waste)
3. Bind sample to the tip by drawing it up and expelling it back into the original tube 12 times, with 15 µL moved through the C18 bed each time.
4. Rinse the C18 bed with 15 µL of 0.1% FA in water 4 times to remove salts and highly hydrophilic compounds. For some samples, the first two rinses can be added back to the original sample in case re-analysis will be required and/or the sample is very precious. Other rinses were expelled to waste.
5. Elute hydrophobic compounds, including peptides, from the C18 material by pulling up and dispensing a solution of 50/50/0.1 ACN/H_2O/FA in 15 µL amounts 6 times.
6. Evaporate the sample till dry using a speedvac (Low/Med heat).
7. Reconstitute the sample in ~10 µL of water. Sonicate to ensure full dissolution. Formaldehyde labeling is performed next. The sample may be frozen at this point.

Isotopic Formaldehyde Labeling

Two-plex [5] formaldehyde labeling schemes, as described by Claire were used [1]. The two-plex uses formaldehyde (OCH_3) and deuterium formaldehyde (OCD_3) as dimethylation agents. A light label incorporates 2(CH_2) per primary amine while a heavy label incorporates 2(CD_2) per primary amine. All peptide N-termini (except those blocked by pyroglutamation or acetylation) and lysine side chains will be labeled.

Formaldehyde Labeling Materials

Formaldehyde, 1% solution in water (v/v) (Sigma), D-formaldehyde, 1% solution in water (v/v) (Isotech), borane pyridine, solution in isopropanol (Sigma), water, desalted sample of neuropeptides from 1-3 organs, 100 mM ammonium bicarbonate solution, water bath at 37°C, speedvac, pipette

Formaldehyde Labeling Instructions

1. Sonicate the dissolved sample to ensure full dissolution.
2. Prepare formaldehyde, deuterium formaldehyde, and borane pyridine solutions in a chemical fume hood. Solutions a, b, and c must be made fresh the day of the experiment, solution d may be made in advance and diluted day of experiment.

    a. borane pyridine (3uL in 1mL isopropanol)
    b. formaldehyde-H2 (4% in water, v/v; 20uL of 37% FH2 in 740µL water)
    c. formaldehyde-D2 (4% in water, v/v; 40uL of 20% FD2 in 800µL water),
    d. ammonium bicarbonate (0.1 M; 15.8mg in lmL water=200mM stock; dilute 1:1 in water for 100mM when needed).

3. Add 10 µL of formaldehyde solution to samples for light labeling or 10µL of deuterium formaldehyde solutions to samples for heavy labeling.
4. Add 10 µL of borane pyridine solution to each sample.
5. Incubate at 37°C for 15 min in the water bath.
6. Take samples out of the water bath and add 10 µL of the 100 mM ammonium bicarbonate solution.

7. Dry samples in speedvac. Samples will be fully dry when they do not smell of ammonia. A sample may look dry, but still contain ammonia. Highly basic conditions can cause peptide breakdown, so it is necessary to remove all ammonia.

A.4 Tissue Sample Preparation Protocols

Tissue Extraction

Pericardial organs (POs) were obtained from crabs via the SOP for Marian Crustacean Dissection with the following modifications. For DiLeu analysis the tissues were stored in DMF while for all other analysis the tissue was stored in acidified methanol. See Appendix C for DiLeu and iDiLeu protocols.

Tissue Neuropeptide Extraction Procedure

1. Homogenize, grind, in a manual glass tissue homogenizer using 200 µL of acidified methanol.
2. Centrifuge the homogenate at 5 min x 16,100 g.
3. Remove supernatant and set aside in a new tube.
4. Reconstitute pellet in 200 µL acidified methanol.
5. Repeat steps 2-4 twice (for a total of 3 extractions) and combine the supernatants. The pelleted solids can be saved, but it has been determined that the first and second extractions of most tissues contain the most neuropeptides. Subsequent extractions contain primarily lipids and/or no compounds of interest.
6. Evaporate the solvent using the speedvac. Medium or Low heat is recommended.

A.5 Instrumental Analysis

The majority of the data analyses for the hemolymph samples were performed on the MALDI TOF/TOF instruments. The tissue extracts were analyzed on LC-ESI instruments. Robby Cunningham performed data acquisition of the deamidation samples in Chapter 6 using the amaZon ETD (Bruker Daltonics, Bremen, Germany). DiLeu labeled tissues were analyzed on the Waters micromass LC-ESI (prefer to Appendix C). Multiply Charged MALDI experiments were performed on both MALDI TOF/TOF and MALDI FTICR instruments. The iDiLeu analysis was performed on the MALDI TOF (refer to Appendix C). A complete list of instruments is supplied in Table A.2.

A.5.1. MALDI Mass Spectrometry

MALDI FTICR

The Varian/IonSpec Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (Lake Forest, CA) is equipped with a 7.0 T actively shielded superconducting magnet and a 355nm Nd:YAG laser (Laser Science, Inc., Franklin, MA). All of the mass spectra were collected in the positive ion mode. 50 laser shots were employed and all the mass spectra were recorded from the mass range of *m/z* 200-2000.

MALDI TOF/TOF

There are three MALDI TOF/TOF instruments. The two Bruker (Bruker Daltonics, Bremen, Germany) MALDI TOF/TOF instruments used for analysis are the AutoFlex III Smartbeam and UltraflexTreme. Both are equipped with a Smartbeam laser which is a frequency-tripled Nd:YAG laser, at 355 nm, that uses the structured focus profile,'Nd:YAG (structured A) and Nd:YAG (structured B)'[6].

The 4800 MALDI-TOF/TOF analyzer (Applied Biosystems, Framingham, MA) is equipped with a 200 Hz, 355 nm Nd:YAG laser (spot diameter of 75 μm). The full MS spectral analysis was operated in the positive ion reflectron mode. Instrument parameters were set using the 4000 Series Explorer Software (Applied Biosystems). Each mass spectrum was generated by averaging 500 laser shots over the mass range *m/z* 500-2000. Mass spectra were externally calibrated using peptide standards applied onto the MALDI sample plate.

The AutoFlex III MALDI-TOF/TOF analyzer (Bruker Daltonics, Bremen, Germany) is equipped with 200 Hz Smartbeam II laser. Mass spectra were acquired in a positive ion reflectron mode. 500 shots were employed and all the mass spectra were recorded from *m/z* 100 to *m/z* 2000. External calibration was performed, using a standard peptide mixture. The results were processed with the FlexAnalysis software from Bruker Daltonics.

The UltrafleXtreme MALDI TOF/TOF analyzer (Bruker, Framingham, MA) is equipped with a Smartbeam. Acquisitions were performed in positive ion reflectron mode. Mass spectra were obtained by averaging 500 laser shots covering the mass range *m/z* 400-2000. External calibration was performed, using a standard peptide mixture. The results were processed with the FlexAnalysis software from Bruker Daltonics.

MALDI Orbitrap

The MALDI LTQ Orbitrap XL instrument (Thermo Scientific, Bremen, Germany) is equipped with a commercial 60 Hz N2 laser at 337 nm (LTB Lasertechnik Berlin GmbH, Berlin, Germany). All experiments were performed with automatic gain control (AGC) turned on. The Orbitrap analyzer is fully tuned and calibrated weekly with the aid of two calibration peptide

mixtures (MSCal4, Sigma Aldrich, St. Louis, MO) for optimization in two mass ranges: *m/z* 50-1100 for small molecules and *m/z* 900-3000 for peptides.

A.5.2 ESI mass spectrometry

microMass

A Waters Micromass QTOF mass spectrometer (Waters Corp., Milford, MA, USA) is coupled to a Waters nanoAcquity UPLC system. Primarily used in DiLeu analysis, refer to Appendix C for LC operation. The micromass was used for direct infusion analysis of the bradykinin peptide in chapter 5 to determine the best multiply charged spectrum that could be expected or achieved in MALDI analysis.

AmaZon ETD

Robby Cunningham described his acquisition techniques in his thesis [7]. Briefly, the amazon ETD (Bruker Daltonics, Bremen, Germany) is an ion trap mass spectrometer equipped with an on-line nanospray source. A Hystar (Version 3.2, Bruker Daltonics, Bremen, Germany) controlled the Eksigent nanoLC software (Dublin, CA; Version 3.0 Beta, Build 080715) for MS acquisition of all experiments.

A.6 Works Cited

1. Schmerberg, C.M., *Functional Neuropeptidomics in the Decapod Crustacean: Method Development and Application to Behavioral Neuroscience Research*. 2012, The University of Wisconsin - Madison: Ann Arbor. p. 494.
2. Kwok, R. and S.S. Tobe, *Hemolymph clotting in crustaceans: Implications for neuropeptide extraction from invertebrate hemolymph.* Peptides, 2006. **27**(3): p. 590-596.
3. Martin, J.-L., A.V. Wormhoudt, and H. Ceccaldi, *Zinc-hemocyanin binding in the hemolymph of< i> Carcinus maenas</i>(Crustacea, Decapoda).* Comparative Biochemistry and Physiology Part A: Physiology, 1977. **58**(2): p. 193-195.
4. Cunningham, R., et al., *Investigation and reduction of sub-microgram peptide loss using molecular weight cut-off fractionation prior to mass spectrometric analysis*. J Mass Spectrom, 2012. **47**(10): p. 1327-32.
5. Chen, R., et al., *Comparative Neuropeptidomic Analysis of Food Intake via a Multi-faceted Mass Spectrometric Approach.* ACS Chem Neurosci, 2010. **1**(3): p. 204-214.
6. Holle, A., et al., *Optimizing UV laser focus profiles for improved MALDI performance.* Journal of Mass Spectrometry, 2006. **41**(6): p. 705-716.
7. Cunningham, R.S., *Mass Spectrometry Applications for Comparative Proteomics and Peptidomic Discovery*. 2012, The University of Wisconsin - Madison: Ann Arbor. p. 242.

A.7 Figures and Tables

Figure A.1

Optimal locations for hemolymph removal are demonstrated by the black arrows. While success is found on the 3rd walk leg, and larger volumes of hemolymph must be withdrawn from this area to obtain decent neuropeptide content. The optimal location is between the 3rd and 4th walking legs, closer in proximity to the heart, PO and TG. This area is rich in neuropeptide content. The final location is the tail, which also provides access to the heart, PO and TG area.

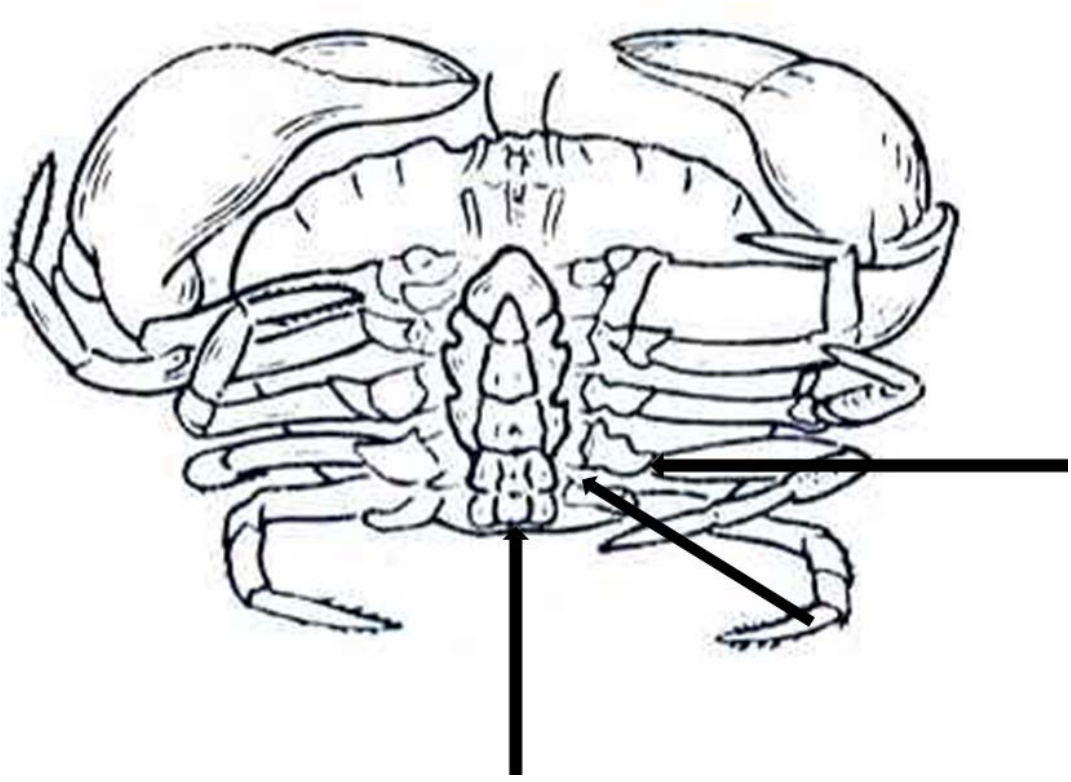


Table A.1 Hemolymph Withdrawal Template

The following table is based on personal observations. It demonstrates the survival rate of crabs when a certain amount of hemolymph is withdrawn from the crab; under optimal experimental conditions (i.e. no internal injuries). Following the notes will increase the likelihood of survival in the crab, but as a general rule the more hemolymph withdrawn (either all at once of over several days) the more likely the crab will not survive.

Volume withdrawn	**#withdrawals per wk**	**Morbidity Expectation**	**Notes**
0.5mL	1	Low	Replace hemolymph and feed every other day if conducting more than 3/wk
	5	High	
1mL	1	Low	Replace hemolymph and feed every other day for more than 2/wk
	5	High	
2mL	1	High	Replace hemolymph and feed at least 5 days before and within 2 days after
3mL	1	High	Replace hemolymph and feed at least 5 days before and within 1 day after
5mL	1	High	Replace hemolymph and feed at least 5 days before and within 1 day after

Table A.2

A list of the instruments used in this thesis for analysis. Instruments used for each experiment is indicated by an X. Under each experiment is the chapter (indicated by C4-8) where further explanation of the experiments conducted may be found.

Ionization	Analyzer	Company Name	Experiments				
			DiLeu C4	MCM C5	D C6	iDiLeu C7	Hemo C8
MALDI	FTICR	Varian/IonSpec FTICR		X	X		
	TOF/TOF	AB/MDS SCIEX 4800 MALDI TOF/TOF		X	X		
		Bruker AutoFlex II Smartbeam		X	X		X
		Bruker UltraflexTreme		X	X	X	X
	LTQ Orbitrap	Thermo Scientific MALDI LTQ Oribtrap		X			
ESI	QTOF	Waters micromass QTOF	X	X			
	IT	Bruker amaZon ETD			X		

Abbreviations used: MALDI=matrix assisted laser desorption/ionization; ESI=electrospray ionization; FTICR= Fourier transform ion cyclotron resonance; TOF=time of flight; LTQ=linear trap quadrupole; QTOF=quadrupole time of flight; IT=ion trap; AB=Applied Biosystems; ETD=electron transfer dissociation; DiLeu= N, N-dimethyl leucine; MCM= multiply charged MALDI; D=Deamidation; iDiLeu=isotopic DiLeu; Hemo=Hemolymph analysis

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER: 321-R **EFFECTIVE DATE: 10-12-2011**

TITLE: Revision to Routine Jonah Crab (*Cancer borealis*) Husbandry

Revisions are denoted by underlining.

1. General Animal Holding Room Procedures: See Policy/Procedure No. 300.

2. Temperature:

2.1. Room temperature shall be maintained between 64-66 F (18-19C)

2.2. Tank water temperature shall be maintained between 50-57 F (10-14C)

3. Relative Air Pressure: negative

4. Relative Humidity: 30%-70% **RH**

5. Lighting: Light timer controls shall normally be set for a 12:12 (light: dark) lighting cycle.

6. Personal Protective Equipment:

6.1. To enter room: lab coat or dedicated uniform

6.2. To work with crabs: gloves

7. Caging:

7.1. Jonah crabs shall normally be housed in 55 gal saltwater aquaria with re-circulating pumps.

7.2. The maximum density/number of crabs per tank shall be: 10 crabs/55 gal of saltwater.

7.3. Tank water shall be either reverse osmosis (R/O) water or distilled water with 0.67-1.1 grams of instant ocean salts per 5 gallons of water.

7.4. Aquarium pH level: Optimal pH is 8.3. It is important to maintain pH anywhere between 8.2 and 8.4 so that the good bacteria that live on the tank's biological filter and chew up the bad waste products don't go into hibernation when the pH is too low. In order to bring up the pH add some e buffer which is purchased from an aquarium store.

7.5. Water Quality Control:

7.5.1. Check tank temperature and specific gravity every day- adjusted as needed.

7.5.2. It is necessary to perform a 10%-20% water change prior to the arrival of a shipment of crabs. In our case, we change water almost every week. In addition, the mesh filter that the water coming out of the tanks flow into and the protein skimmer need to be changed, too.

7.5.3. Check water over cooling coils every day- always keep coils completely covered, but don't let the water level go over the filter bags (This applies only to the square aquarium (Tank #3)).

7.5.4. Make sure that the flow rate through the filters is as high as possible (= more filtration).

7.5.5. Don't put more than 10 crabs in the tank when they first come in- if the tank is overcrowded, then the crabs may fight and kill each other.

7.5.6. Remove dead animals (and animal pieces) as quickly as possible.

7.5.7. Clean the tank when it starts to get cloudy.

7.5.8. If possible, newly clean tank should be pre-condition for 24 hours before introducing new

animals.

8. Feeding:

8.1. Crabs shall be fed as needed. Blue crabs can survive for several weeks without being fed.. They will be fed frozen fish of any white flesh (cod, tilapia, whiting, etc) or frozen cooked shrimp.

8.2. The procedure for feeding the crabs is as follows:

1. Thaw a portion of fish fillet in the microwave, or thaw the shrimp by running warm water over them.
2. Put enough pieces of fillet (about 1 inch square) or enough shrimp into the tank that every crab has one. Add 2-3 extra pieces/shrimp. Each crab will eat 2-3 pieces of shrimp.
3. Watch to make sure that no fights break out.
4. The crabs may be left unattended to eat for 1-2 hours if every crab is eating and not bothering the others.
5. Return in an hour or two and remove any pieces of food or food debris that remain in the tank as they may foul the water. Do not try to take food that a crab is holding on to.

9. Cage Cleaning/Sanitizing Schedule: Re-circulating water with suitable baffles installed in the tank to slow down water circulation to promote effective settlement of suspended particles. When it is necessary to flush out bottom sludge the pump circulating the water is switched off, the baffles are raised slightly. Alternatively, the sludge can be siphoned or pumped from the settlement tank. A separate protocol for tank cleaning is provided.

9.1. Tanks shall be sanitized at least every 6 months or when empty.

10. Approval

I have read and approved this revised procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER: **EFFECTIVE DATE: 10-12-2011**

TITLE: Routine Blue Crab (*Callinectes sapidus*) Husbandry

1. General Animal Holding Room Procedures: See Policy/Procedure No. 300.

2. Temperature:

2.1. Room temperature shall be maintained between 64-66 F (18-19C)

2.2. Tank water temperature shall be maintained between 50-57 F (10-14C)

3. Relative Air Pressure: negative

4. Relative Humidity: 30%-70% **RH**

5. Lighting: Light timer controls shall normally be set for a 12:12 (light: dark) lighting cycle.

6. Personal Protective Equipment:

6.1. To enter room: lab coat or dedicated uniform

6.2. To work with crabs: gloves

7. Caging:

7.1. Blue crabs shall normally be housed in 55 gal saltwater aquaria with re-circulating pumps.

7.2. The maximum density/number of crabs per tank shall be: 10 crabs/55 gal of saltwater.

7.3. Tank water shall be either reverse osmosis (R/O) water or distilled water with 0.67-1.1 grams of instant ocean salts per 5 gal of water.

7.4. Aquarium pH level: Optimal pH is 8.3. It is important to maintain pH anywhere between 8.2 and 8.4 so that the good bacteria that live on the tank's biological filter and chew up the bad waste products don't go into hibernation when the pH is too low. In order to bring up the pH add some buffer which is purchased from an aquarium store.

7.5. Water Quality Control:

7.5.1. Check tank temperature and specific gravity every day- adjusted as needed.

7.5.2. It is necessary to perform a 10%-20% water change prior to the arrival of a shipment of crabs. In our case, we change water almost every week. In addition, the mesh filter that the water coming out of the tanks flow into and the protein skimmer need to be changed, too.

7.5.3. Check water over cooling coils every day- always keep coils completely covered, but don't let the water level go over the filter bags (This applies only to the square aquarium (Tank #3)).

7.5.4. Make sure that the flow rate through the filters is as high as possible (= more filtration).

7.5.5. Don't put more than 10 crabs in the tank when they first come in- if the tank is overcrowded, then the crabs may fight and kill each other.

7.5.6. Remove dead animals (and animal pieces) as quickly as possible.

7.5.7. Clean the tank when it starts to get cloudy.

7.5.8. If possible, newly clean tank should be pre-condition for 24 hours before introducing new animals.

Number: **Effective Date**: 10-12-2011
Title: Routine Blue Crab Husbandry

8. Feeding:

8.1. Crabs shall be fed as needed. Blue crabs can survive for several weeks without being fed.. They will be fed frozen fish of any white flesh (cod, tilapia, whiting, etc) or frozen cooked shrimp.

8.2. The procedure for feeding the crabs is as follows:

1. Thaw a portion of fish fillet in the microwave, or thaw the shrimp by running warm water over them.
2. Put enough pieces of fillet (about 1 inch square) or enough shrimp into the tank that every crab has one. Add 2-3 extra pieces/shrimp. Each crab will eat 2-3 pieces of shrimp.
3. Watch to make sure that no fights break out.
4. The crabs may be left unattended to eat for 1-2 hours if every crab is eating and not bothering the others.
5. Return in an hour or two and remove any pieces of food or food debris that remain in the tank as they may foul the water. Do not try to take food that a crab is holding on to.

9. Cage Cleaning/Sanitizing Schedule: Re-circulating water with suitable baffles installed in the tank to slow down water circulation to promote effective settlement of suspended particles. When it is necessary to flush out bottom sludge the pump circulating the water is switched off, the baffles are raised slightly. Alternatively, the sludge can be siphoned or pumped from the settlement tank. A separate protocol for tank cleaning is provided.

9.1. Tanks shall be sanitized at least every 6 months or when empty.

10. Approval

I have read and approved this procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER: **EFFECTIVE DATE:** October 12, 2011
TITLE: Marine Crustacean Dissection (Procedure edited from the thesis of Joshua Schmidt: Figures found on T:/drive)

I. Dissection Introduction

The process of dissection involves two major procedures, a macrodissection and a microdissection. The macrodissection is the rough dissection without a microscope to remove the areas of the crab that contain the neurosecretory organs. Following macrodissection, microdissection under a microscope is performed to remove the nervous system organs from their surrounding tissue. As stated, the macrodissection is performed first followed by microdissection. The steps for performing a full brachyuran crab dissection are listed below with pictures outlining the method.

To begin dissection, you must first decide what organs/tissues you want to obtain. This must be established initially as it will determine how much how little of this manual you will need to follow. To make this easier to follow, the instructions have been split into sections in order to make it easier for the reader to find the section needed to perform the dissection they are interested in.

II. Macrodissection

1. Drawing Hemolymph

To perform the most successful hemolymph draw, the crab must not have been cooled on ice and thus should be the first thing you do before any other dissection. The hemolymph must also be drawn first so that none is lost when legs are removed later for dissection. If only hemolymph is needed and no other tissue, not icing the crab greatly improves the odds of survival of the crab after drawing. So the first step for drawing hemolymph is to get everything ready. You will need an aluminum dissecting pan, a syringe that will hold at least 5 mL (even if you only want 1 mL), a large needle (20 gauge or lower, for the needle), a spatula, and a tube to receive the hemolymph (this means having whatever extraction buffer you are using in the tube already)(Figure 2.2). Get these ready before beginning because once you start you will not want to be rushing around trying to get the other supplies you need. Once you have these, you are ready to draw hemolymph.

1. You may draw hemolymph from the tail or a leg of the crab.
   i. To draw hemolymph from the crab's tail, you will want to lean the crab's top side against the back of the pan (so you are looking at the tail) and then hold the claws of the crab back so they do not pinch you (Figure 2.3). Most people do not have large enough hands to hold both claws with one hand and may require someone to assist them in doing this. Initially, the crabs generally will cooperate and you can hold them fairly easily, but as soon as you carry out the next step, you definitely need help holding the crab's claws.

   After you have the crabs claws controlled, use the spatula to pry the tail of the crab away from his body. At this point, you will probably want another person to help you control the crab. Generally, as soon as you pry the tail up, the crab begins

squirming and will do almost anything to stop you from doing what you are trying to do. If you have someone helping you, it is possible to hold the tail back and go on to the next step, but if alone, it is not necessary to continue holding the tail away from the body. Once you release the tail, it will fold against the body, but not so tight that you will not be able to easily pull it back again when you need to.

Once the tail has been pulled away from the body, you are ready to insert the needle to draw hemolymph (hopefully you have the needle on the syringe already). Insert the needle into the crab right at the base of the tail, where it meets the body (Figure 2.5). Angle the needle into the crab towards the ventral side, into the cardial chamber of the crab.

ii. To draw hemolymph from a leg, insert the needle at the base of one of the legs (Figure 2.7).

2. With the needle in the crab, draw back on the syringe plunger to begin drawing hemolymph. Draw back on the plunger enough that you create a vacuum inside the syringe to assist in the drawing (Figure 2.6). You will also probably need to move the syringe in and out of the crab to find the spot that provides the most hemolymph. This will be evident when hemolymph begins rapidly filling up the syringe. Try not to move the needle around too much inside the crab as you will be damaging the organs inside the crab and will reduce the survivability of the crab.
3. Once you have drawn as much hemolymph as you need, remove the needle. For crab survival, the smaller the amount of hemolymph that is drawn, the better the chances are. Generally, 2 mls can be drawn without a serious detrimental effect on the animal, and as much as 10-15 mls can be drawn if the crab is to be sacrificed following hemolymph removal.
4. Once you have removed the needle, let the crab rest and/or place it back into the tank, making sure to allow all the air to burp out of the crab before letting it settle to the bottom.
5. Next, take the syringe and place the appropriate amount of drawn hemolymph into the tube you have prepared. Place this tube on ice and continue with dissection. Once you have done this, you are finished and can go on with dissection or whatever you need to do for your experimentation.

2. Stomatogastric Nervous System (STNS) and Eyestalk Removal

Macrodissection for the removal of the following organs:

Brain
Stomatogastric Ganglia (STG)
Commissural Ganglia (CoG)
Oesophagial Ganglia (OG)
Anterior Cardial Plexus (ACP)
Sinus Gland (SG)

Items needed:

Aluminum dissecting pan
Side cutters or rongeurs
Small scissors
Toothed forceps or locking hemostat

Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Spatula
Rubber lined dissecting dish

1. Before performing this dissection, you need to place the crab on ice for 15-30 minutes. This anesthetizes the crab as well as makes it easier to deal with initially. This is done by burying a crab in ice, in a bucket. Choose the crab you want, capture it and place it in the bucket of ice, making sure it is completely covered, for 15-30 minutes. If you need to draw hemolyph, follow the directions above, and then place the crab on ice. While waiting for the crab to ice, prepare all the tools you will need for this dissection (Figure 2.8).
2. Once the crab has been chilled for 15-30 minutes, remove it and place it in the dissection pan. Then, use the side cutters to remove the legs and claws at their base where they meet the body of the crab (Figure 2.9). Generally, by removing the claws first, the whole process is made easier.
3. Next, turn the crab over so the ventral side is up and use the side cutters to crack away the outer edge of the crabs shell. Crack into the crab roughly 0.5 inch or at least enough to provide space to slip in a spatula (Figure 2.10 A). Crack around the crabs shell from the rear edge of the crab, up around the crab almost to the eyestalks (Figure 2.10 B).
4. After you have the edge of the shell cracked away, slip the more rounded end of the spatula into the shell and separate the hypodermis of the crab from the crab shell (Figure 2.11). The hypodermis is the thin skin that is present just under the shell of the crab. Be careful when doing this as you do not want to break the skin (although breaking the skin does not mean you can not continue, it does make the dissection more difficult later).
5. Separate an area of the hypodermis from the shell large enough that you can crack away more a sufficient amount of shell that you can continue using the spatula to separate the hypodermis all the way across the crab. Be especially careful when separating the hypodermis under the middle section of the crab as this is where the stomatogastric nervous system lies and you can damage it if not careful. You will essentially be going back and forth between cracking the shell away and separating the hypodermis. Eventually, you want to be left with most of the front half of the crab's shell removed, except for a small bridge (maybe 1 inch wide) from the eyestalk area to the rear of the crab. The hypodermis should be separated completely from the front portion of the crab (Figure 2.12).
6. Once you have the top of the crab prepared, with the appropriate amount of shell removed and the hypodermis separated, turn the crab over and begin removing the mouthparts of the crab. These can be removed by pulling them away from the body of the crab, and twisting until they crack off (Figure 2.13 A). There are a good number of mouthparts, some of which you will need to pull away from the crab's body with a spatula (Figure 2.13 B). Finally, you will get down to the main pair of mandibles over lips and the entrance to the oesophagus. These can be removed by using the side cutters and cutting the base of the mandibles and levering the up and away from the crab (Figure 2.13 C).

7. Next, flip the crab back over and crack along the front of the crab, around the eyestalk sockets. Make sure to crack a good amount along the very front of the crab (Figure 2.14 A). Once you have the front of the shell cracked away, use the spatula to make sure the hypodermis is separated from the shell as far forward as is possible without damaging it. After this is done, use the cutters to crack the shell all the way across the middle of the crab. Once the shell is cracked all the way across, carefully lift from the rear the remaining middle/front portion of the shell (Figure 2.14 B). If you have cracked enough shell away along the front, this piece should lift away fairly easily with no hypodermis or other tissue still attached to it. If hypodermis or other tissue is still stuck to this portion of the shell, either crack more of the shell on the front of the crab, or carefully use the small scissors to cut away the connected tissue as close to the shell as is possible.
8. Once the front portion of the crab has been removed, you should now be able to remove the eyestalks (for SG). They may either be on the piece of shell you just removed, or remain on the crab. Wherever they are, use the small pair of scissors to cut the softer section at the base of the eyestalk to remove it from the rest of the shell (Figure 2.15). Place the eyestalks into saline solution and place on ice until carrying out microdissection (in next section).
9. After removing the eyestalks, use the spatula to separate the hypodermis from lower shell on either side of the eyestalks (Figure 2.16 A). Then, use the small scissors to cut the tissue along the front of the crab. If the shell was cracked correctly, there should be two small muscles that remain attached to the front portion of the crab's lower shell. Cut these, and then continue to cut down and forward along the inside edge of the crab's lower shell. Make sure to keep your scissors as close to the shell as possible, making small cuts as you progress (Figure 2.16 B). Use the scissors or the spatula to occasionally scrape the cut tissue down out of the way. Continue doing this to the lower part of the shell where you will eventually reach an arch structure in the shell with a small bump in the shell protruding into the crab's body (Figure 2.16 C).
10. Once you have reached the arch in the shell over the oesophagus, use the small scissors to carefully cut around the oesophagus and lips of the crab (Figure 2.17 A). Once the lips and oesophagus are free, use the cutters to cut the shell on either side of where the eyestalks were down towards the arch in the shell over the oesophagus (Figure 2.17 B). This should release the whole piece of shell over the oesophagus and will allow you to remove the stomach.
11. Now that you have the oesophagus revealed, you will be removing the stomach. This is probably one of the more difficult parts of the dissection and generally takes practice to become proficient at. In any case, you will need to first turn the crab so that you are facing it head on. Also, prop the front of the crab on the front of the dissecting tray so the crab is leaning up toward you. Now, take the toothed forceps or the locking hemostats in your left hand and grab the lips. Carefully lift the lips a little and using the small scissors, cut at the base of the oesophagus (Figure 2.18 A). Continue carefully lifting the lips as the oesophagus is cut away, making sure not to tear the lips but to use them to control the oesophagus. As you lift, continue cutting the tissue under the oesophagus, towards the rear of the crab. If everything has gone

well so far, you should find two small ossicles just below the oesophagus that you can use to guide your cutting. Keep cutting above these ossicles but not into the oesophagus and stomach. Also, try to avoid cutting into any white tissues as these will make the whole preparation much messier to deal with. Continue lifting the oesophagus with the forceps and cutting back into the crab until you see two whitish yellow balls on the underside of the stomach (Figure 2.18 B). Once you have reached these, you can cut the hypodermis on either side of the stomach angling back and away from the stomach (Figure 2.18 C). Extend this cut towards the back of the stomach, and then cut behind the stomach and lift the whole stomach away from the crab, cutting any other tissue holding the stomach in place (Figure 2.18 D). Keep on holding the stomach by the lips with the forceps and go on to the next step.

12. The next step is the removal of the teeth from the stomach and can also be challenging to perform the first several times attempted. To do this, take the stomach that you have just removed and flip it over, turning the forceps in the process. Then, drape the stomach the index finger of your left hand and then hold onto the forceps with your left hand as well. Once you have the stomach positioned well, use the small scissors to cut straight down the bottom of the oesophagus (that should now be up because the stomach is flipped over) (Figure 2.19 A). Cut straight back, between the two whitish yellow balls and down on the back side of the stomach so that there is a cut running from front to back on the bottom of the stomach. Then take the scissors and at the widening of the oesophagus into the stomach, make cuts angling towards the back side corners of the stomach (Figure 2.19 B). Once these cuts are made, the stomach should almost fall open, revealing the inside of the stomach. Inside the stomach, you may find the remains of the crab's last meal. Wash this out with saline if necessary. You will also find three dark green, harder parts in the stomach. These are the teeth the crab uses to grind its food inside the stomach. To remove the teeth, cut them out with the small scissors; do this while you have the stomach draped across your left hand while you hold the lips with the forceps (Figure 2.19 C).
13. Once you are done removing the teeth, flip the stomach back over and place it into a rubberized dissecting dish with enough saline to cover it completely (Figure 2.20). Set this aside on ice or in a fridge until you are ready to perform the microdissection to remove the stomatogastric nervous system organs.
14. Now that you have removed the eyestalks and the stomach, you are ready to go on to either perform macrodissection to remove the pericardial organs and thoracic ganglia, or to carry out the microdissection of the sections you have just dissected.

3. Cardial Ridge and Thoracic Ganglia Removal

Macrodissection for removal of the following organs:

Heart
Pericardial Organ (PO)
Thoracic Ganglia (TG)

Items needed:

All items for macrodissection of stomatogastric ganglia
Large scissors (curved or straight)
Forceps, fine and paddle (NOT microdissecting forceps)

Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Tube prepared for heart and another for TG (if necessary)

1. To perform this dissection, you must have performed dissection of the stomach before beginning. Once you have the eyestalks and stomach in saline and placed on ice, you can continue with this dissection in preparation to remove the heart (if necessary), the cardial ridges, and the thoracic ganglia. Make sure you have all the supplies prepared before beginning (Figure 2.21).
2. First, pick up the crab in your left hand and take the large scissors in your right hand. Using the scissors cut the shell straight back from 1 inch on either side of the midline all the way back to where the top shell meets the bottom shell along the back of the crab (Figure 2.22 A). When you cut all the way through on one side, you will be able to break off that side of the shell. Once the shell is removed, use the scissors to cut away the tissue that remains there along with the gills along the cardial ridges. Make sure to remove the tissue on both sides of the crab so that all you have left are the tail section of the crab, the cardial ridges, and the shell over the top of that (Figure 2.22 B).
3. Once you have both sides of the crab removed, turn the scissors on their side, open them and slide them under the upper part of the shell with the blades on either side of the top of the cardial ridge (Figure 2.23 A and B). Cut the connection between both cardial ridges and the upper portion of the shell remaining. Once these connections are cut, lift the shell away to reveal the heart and the inside of the cardial ridges.
4. If you need to remove the heart, use a clean pair of forceps to carefully tug on the heart until it pulls out of the crab. If necessary, use the small scissors to cut any connections to the surrounding tissue (Figure 2.24). Be careful when removing the heart not to damage the inside of the cardial ridges as this is where the PO's reside. Blot the heart on a Kimwipe to remove any excess hemolymph and place the heart into the tube you have prepared.
5. Once you have removed the heart, lift the tail and pull it away from the remaining portion of the crab (Figure 2.25 A). Then use a pair of forceps (I use the paddle ended forceps) to remove all the tissue from between the cardial ridges (Figure 2.25 B). This includes removing all the yellow tissue (hepatopancreas) that you can, as well as removing the long white tubes which run between the cardial ridges. Again, make sure not to damage the inside of the cardial ridges where the PO's are.
6. After the inside of the cardial ridges is fairly clean, use the large scissors to cut the cardial ridges away from the remaining crab at their bases (Figure 2.26). Make sure to get the entire cardial ridge, and place it in saline solution as soon as it can be removed from the body of the crab.
7. Once the cardial ridges have been placed in saline, you can go on to work on removing the TG. To begin this process, use the small scissors to cut away the ossicles around the center of the remaining piece of the crab (Figure 2.27 A). Be careful as the TG lies directly in the center of the crab. Once you have cut most of the ossicles, use the forceps to remove as much of the ossicles as you can (Figure 2.27 B).

Number:

Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

8. After removing the ossicles, you should be able to see the TG in the bottom of the crab piece (Figure 2.28 A). It looks like a wagon wheel, with ganglia projecting out to the legs and rest of the crab's body. To remove it, use the small scissors to cut around the TG (Figure 2.28 B), and using a fine (clean) forceps, carefully lift the TG. As you pull it up, use the scissors to cut any remaining connections it has to the surrounding tissue (Figure 2.28 C).
9. Once you have the TG removed from the crab, use a Kimwipe to blot the TG and to remove any excess tissue from it. After the TG is clean, place it into the tube you have prepared for it.
10. After removing the TG from the crab, you are finished with the macrodissection of the crab. Now, you should have the eyestalks in saline, the stomach section in saline in a dissecting dish, and the PO in saline, all of these chilled on ice or in a fridge (Figure 2.29). Additionally, you should have the drawn hemolymph, heart tissue, and TG, all in separate tubes that each has the appropriate extraction buffer in them. These tubes should also be placed on ice so as to keep them cool while you perform microdissection to obtain all the other tissues. The final step of the macrodissection process is to clean up all the tools you have used in the process. All the crab parts need to be placed in a dead animal bag and placed in the dead animal freezer provided by the animal care facility. All the tools should be rinsed and all of the "gunk" scrubbed off of them. After rinsing in water, rinse them in a solution of 70% EtOH before lying out to dry. Once this is all done, you are all finished with the macrodissection and can continue onto the microdissection.

III. Microdissection

1. Sinus gland (SG) dissection from eyestalks

Microdissection for removal of the following organs:

Sinus glands (SGs)

Items needed:

Dissecting microscope
Rubberized dissecting dish
Large spring scissors
Two pairs of the least and most sturdy, fine micro- forceps
Fine micro- forceps
Small spring scissors
Tube ready for SGs

1. After you have cleaned up from macrodissection, you are ready to begin the microdissection using the microscope to complete the gathering of nervous system organ from the surrounding crab tissue. Prepare by pouring saline into the dissecting dish and getting all your tools ready and at hand (Figure 2.30).
2. Next, take the eyestalks out of the saline. Using the large spring scissors, slide one blade along the inside edge of the eyestalk and cut first the concave side, and then turn the eyestalk over and cut the convex side (Figure 2.31). Do this for both eyestalks and then place them in the dissecting dish.
3. While looking through the microscope, use the two pairs of the least fine, sturdy forceps and carefully pull the two halves of the eyestalk apart (Figure 2.32 A). You

want to use these forceps because if you use the fine forceps, you will bend them and ruin them. Never use the fine forceps when there is hard shell present. In any case, while prying the two halves apart, watch for a bright white, almost iridescent, spot that seems to be embedded in the tissue on the concave side of the eyestalk; this is the SG (Figure 2.32 B). Once you have identified the SG, use one of the forceps to pull the whole piece of tissue out of the eyestalk shell. Do this for both eyestalks. Once you have removed the tissue from the shell, remove the shell from the dissecting dish and place it aside to throw away later.

4. Once you have just the soft tissue, you can grab the fine forceps and the small spring scissors. For microdissection, you should always hold the forceps in your left hand and the scissors in your right hand. This allows you the best control with both tools. Once you have the tools, use them to carefully cut away the tissue surrounding the SG (Figure 2.33 A). It is almost impossible to get rid of all the surrounding tissue, but you should be able to clean the organ up fairly well. One trick is to grab the tissue with the forceps, then place the open scissors over the very tip of the forceps (above the tissue) and carefully pull the tissue through the blades of the scissors up to the point you want to cut (Figure 2.33 B). By doing this multiple times, it is generally possible to clean most of the extraneous tissue from the SGs.
5. After removing the SGs from the eyestalks, grab them with the forceps and remove them from the saline. Before placing them in the tube, make sure to blot the forceps with a Kimwipe to remove any excess saline solution that is on them. To prevent the loss of SGs by sticking to the Kimwipe, blot just the upper part of the forceps, above where you are holding the SGs. Finally, place the SGs in the tube you have prepared with the appropriate buffer in it. You have now completed the SG dissection and can continue on with your microdissection.

2. Pericardial organ (PO) dissection from cardial ridges

Microdissection for removal of the following organs:

Pericardial organs (POs)

Items needed:

Dissecting microscope
Rubberized dissecting dish with two pins
Small scissors
Fine micro- forceps
Small spring scissors
Tube ready for POs

1. To begin this dissection, you again want to fill a dissecting dish with saline solution and get all your tools and materials ready (Figure 2.34).
2. Once you are ready to start, you can remove the cardial ridges from the saline you have them stored in from the macrodissection. Take each cardial ridge and identify which side was inside (towards the heart) and which side was out. If you imagine the cardial ridge as having a triangular shape, the inner side makes up one side of the triangle and the outer side makes up one of the angles of the triangle. Because the POs are removed from the inside side of the cardial ridge, you normally would have to balance the cardial ridge on the point of the outer side, making the dissection

fairly difficult. To simplify this, use the small scissors to cut the point of the outer side off to create a flat side on which to lay the cardial ridge in the dissecting dish.

3. Once you have cut the outer side of the cardial ridge, place it in the dissecting dish with the inner side up and use the pins to stabilize it from moving. This is best done by placing one pin in the bottom (or ventral) side of the cardial ridge, and placing another on the top (dorsal) side, making sure to not to place the top pin through the PO itself (Figure 2.36). If you are not sure about what is PO, just keep the pin as close to the top ossicle as possible when placing it. These pins should also be placed at an angle so that you can work on the cardial ridge between the pins. By placing these pins, you will stabilize the cardial ridge so that you can dissect the PO from it. Because you have two cardial ridges, it is best to pin one down, dissect the PO from it, discard the one you just dissected from, pin the other one and finish by dissecting the PO out of the second cardial ridge.
4. After pinning the cardial ridge so that it is stable and will not shift when you start dissecting you can use the microscope to identify the PO. If macrodissection was done correctly, there should not be any tissue over the area containing the POs, but there very well may be some. If there is extraneous tissue, carefully use the forceps and the small spring scissors to cut it away and remove it. To identify the POs, use the microscope and look for bright white or iridescent strings that run across the main portion of the cardial ridge (Figure 2.37). When looking down on the inside of the cardial ridge, you should see a set of ossicles that create a triangular structure, with another ossicle projecting off one of the triangles vertexes (this side is the forward or rostral side). Over the ossicles that project out of the triangle, there are two commissures that run along this ossicle and connect with the main portion of the PO inside the triangular ossicle structure. In the triangular structure, the PO has junctions on the inside of two of the vertexes of the triangle. Between these junctions, there are two to three commissures that connect them. These can be difficult to see sometimes so it is often easier to identify the junctions first.
5. Once the POs (junctions or commissures) are identified, look through the microscope and use the forceps to gently lift them away from the cardial ridge. It is generally easier to begin by lifting one of the junctions and then using the spring scissor to cut any of the smaller connections to the cardial ridge (Figure 2.38 A). If possible, do not cut the large commissures that cross through the triangular ossicle. Once a junction is free, carefully lift it and as you do, the commissures will begin pulling away as well. Because the commissures are generally under a layer of tissue, it is necessary to cut them out. This can be done by either cutting alongside the commissures, or carefully making small cuts at the spot where the commissure is held in the tissue (Figure 2.38 B). As there are two to three commissures connecting the two main junctions, complete removal requires this process to be carried out multiple times. Overall, this is the most tedious part of the PO dissection as well as being most at risk for cutting the PO. If cut, continue from the other junction and try to obtain as much of the PO as possible. You have the entire PO when you have two junctions, with all the commissures connecting them as well as a bit of PO that

projected from one of the junctions, along the ossicle that projected towards the front end of the crab.

6. After removing the POs, pull them out of the saline solution and carefully blot the forceps with a Kimwipe to remove any remaining saline. Because the POs will easily stick to the Kimwipe, blot the forceps behind the tips where you are holding the POs. Doing this reduces the amount of saline transferred to the tubes with buffer in them. Once blotted, place the POs in the tube, making sure that the POs actually get into the buffer and do not just stick to the walls of the tube.
7. Now that you have dissected one set of POs, discard the cardial ridge you were just working on and start work on the other one. After completely finishing, make sure to clean up the dish you have been working in as well as any tools that were dirtied in the process.

3. Stomatogastric nervous system (STNS) dissection

Microdissection for removal of the following organs:

Brain
Commissural Ganglia (CoG)
Oesophageal Ganglia (OG)
Anterior Cardial Plexus (ACP)
Stomatogastric Ganglia (STG)

Items needed:

Dissecting microscope
Rubberized dissecting dish with at least seven pins
Fine micro- forceps
Small spring scissors
Five tubes with buffer for each type of organ

1. After dissecting the SGs and the POs, we now turn to the STNS dissection. This can be performed before or after the SG and PO dissection and is up to personal preference. To begin this dissection, first make sure everything is ready (Figure 2.39). If ready, remove the dissecting dish with the stomach section in it and place it under the microscope. You should have saline on the stomach from the macrodissection, but make sure that it is fairly clean and is covering the stomach completely. If dirty, make sure the lips are pinned down first and empty out the dirty saline and replace with clean saline. You want enough saline so that the whole stomach is completely submerged. This makes the dissection much easier as tissues can free float in the saline and will not lie down on top of each other. Additionally, refer often to the schematic of the STNS (Figure 2.40) as it will assist you greatly in identifying the location of many of the organs you will be dissecting.
2. Now that you have everything ready, you need to pin out the stomach. This is necessary as it stretches everything out so that it is easily accessible for dissection. There should already be a pin in the lips holding them in place. Next place a pin in each of the lower corners of the stomach, slightly stretching it taught between them and the lips (Figure 2.41 A). After these pins are placed, put a pin on either side of the stomach about halfway between the lips and the lower pins (Figure 2.41 B). When placing these pins, you do not want to just pin the hypodermis, but actually get some of the stomach tissue

underneath and slightly stretch these to either side. Finally, you will place a pin on each side of the oesophagus, just below the lips. Look for where the oesophagus was cut open during macrodissection just below the lips, and place the pins there, pulling up slightly to stretch these up and apart (Figure 2.41 C). Be careful when placing these last pins as it is easy to stick a pin through the CoG. If you are unsure, wait until you have located the CoGs before placing these pins.

3. After the stomach has been pinned out, you can begin the actual dissection. Initially, you want to identify the brain. If macrodissection was done correctly, there should be tissue protecting the brain that will need to be removed. When looking through the microscope, you should see hypodermis from the top side of the crab (a mottled red brown tissue) and hypodermis from the bottom side of the crab (a whitish yellow tissue). Under where these two tissues meet is roughly where the brain should be located (Figure 2.42). If it is not revealed, carefully use your forceps and spring scissors to cut hypodermis out of the way until you see the brain. Do not be afraid to physically remove any tissue that is getting in the way; if it is in the way now, it will probably continue to stay in the way. The brain looks like a white iridescent blob with many commissures projecting from it. It should be fairly easy to spot as it is one of the larger structures on top of the stomach. Once you have identified the brain DO NOT remove it, but carefully clear away some of the tissue surrounding it.
4. Removing the CoGs: With the brain identified, you will now dissect out the CoGs. To do this start at the brain and identify one of the two large commissures that project out of the brain at an angle up towards the lips, but away from the midline. Follow these commissures, cutting away tissue as you go, until you reach a junction (Figure 2.43 A). The junction will look iridescent white and will have another large commissure projecting from it as well as several smaller commissures (Figure 2.43 B). This junction is the commissural ganglia (CoG). Carefully, cut the commissure between the brain and the CoG and use this to pull the CoG up and out of the surrounding tissue. Because there are other commissures connecting it to the rest of the STNS, you will have to make several more cuts to completely remove it. Once you have it removed, move it to a clear area of the dissecting dish and shorten the projecting commissures so that only the CoG remains (Figure 2.43 C). When trimmed down, remove it from the saline and blot the forceps with a Kimwipe to remove the excess saline, again being careful not to lose the CoG in the process. After blotting, place the CoG into the appropriate tube and make sure it gets into the buffer. Finally, repeat this process on the other side to remove the other CoG. One thing to note is that you may have to rotate the dish in order to get a good angle to cut from when working on dissecting the other CoG. Just remember that you can rotate the dish any way you find necessary, as long as you keep your relative sense of where the lips are.
5. Removing the brain: After removing the CoG, you can now return to the brain and remove it. This is most easily done by grasping the right commissure that projected towards the CoG and lifting slowly. As you lift, you will need to cut the other commissures that are projecting out of the brain (Figure 2.44 A). Eventually, you will also have to cut directly underneath the brain to free it from the surrounding tissue (Figure 2.44 B). Be careful not to cut too deeply under the brain as other STNS organs

lie there and will need to be removed later. Continue to lift with the forceps and cutting anything that holds the brain in place. Eventually, you should be able to completely remove the brain. Like the CoG, move the brain to a clear area of the dissecting dish and trim short all the projecting commissures. Once trimmed, remove the brain from the saline, blot it as with the other organs, and place it in the appropriate tube.

6. Once you have dissected out the brain, you are ready to carry on and get the rest of the STNS organs. This next set of steps is probably the most difficult part of the whole dissection, so do not expect to be able to perform it the first, second, or even tenth time you attempt it. It just takes practice and perseverance. Eventually you will get good enough to do it quickly and efficiently.
   In any case, for the next step of dissection, you will begin cutting straight down from where the brain was (Figure 2.45 A). You only will need to cut down maybe half a centimeter or so. You want to pay attention and stop cutting down when you reach an artery that runs along the midline of the crab. This artery will look like an opening in the tissue. It will be deceiving though as there are two muscles that are inside of the artery (Figure 2.45 B). Watch for these two muscles as well as the artery wall to indicate when you have reached the artery. Once you have reached the artery, do not cut any further or you make your dissection very difficult.
7. After finding the artery and the two muscles inside it, look down between the two muscles to find a commissure running inside the artery along the midline, this is the stomatogastric nerve (stn) (Figure 2.46 A). Quite often the stn is difficult to see between the muscles, so carefully cut their connection to the inside wall of the artery on the bottom (caudal) end (Figure 2.46 B). Again, look for the stn lying along the bottom of the artery. It may be necessary to use the forceps blindly to grab anything that might be lying inside the artery. Generally, if you have not seen the stn by now, blindly grabbing for it will reveal its location. If you do grab it, do so gently and do not pull hard as you can snap the stn and make this dissection much more difficult. In any case, once you have identified the stn, grab the muscles that you just cut and lift them up. As you lift them, you may have to cut the top of the artery open to continue pulling up on the muscles. As you reach the end of the muscles, continue pulling on them and cutting towards the lips, all the while paying attention to where the stn is and making sure not to cut it (Figure 2.46 C). This may seem confusing, but will make sense once you begin performing the dissection. After you have cut almost all the way o the lips, cut the piece free and discard it. If you have done this correctly, you should have just revealed the majority of the forward section of the STNS.
8. To help identify all of the necessary organs, refer often to the STNS schematic (Figure 2.40). Once you have removed the main portion of the artery, find the stn again. Next, follow the stn forward to a thickening of the nerve. Although you may not be able to see them, there should be two projecting nerves (anterior cardial nerve, acn) coming off the stn at this thickening. Most likely, you will be unable to see the acns at this point, but take note of where this thickening is as we will be returning here shortly to remove the ACPs. In any case, as you move anterior, you will find a major intersection where two relatively large commissures project off the stn towards where the CoGs were (the superior oesophageal nerve, son). You may find it necessary to clear away any excess

tissue that has remained in the way, just make sure when doing so that you do not cut the stn in the process. From this intersection, the stn should continue forward to a final branch point. At this final branching, there will be many projecting nerves, including two that curve forward and then back towards where the CoGs were (the inferior oesophageal nerve, ion), and another that projects straight up towards you (the inferior ventricular nerve, ivn). Just posterior to this junction is where the oesophageal ganglia (OG) is located (Figure 2.47).

9. Removing the OG: Once you have identified the location of the OG, use the spring scissors to cut the ions that project out of this junction. Then using the forceps, grab either the ivn or the ion and carefully lift the OG up away from the stomach tissue. As you lift it away, you will need to cut the stn below the OG but above the junction of the stn and the sons (Figure 2.48). Once you cut the stn, the OG should be free and you can move it to a clear place in the dissecting dish. Do so and as you did with the CoGs and the brain, trim down the projecting nerves before removing the OG from the saline, blotting the forceps to remove excess saline, and placing into the appropriate tube.
10. Removing the ACPs: After removing the OG, follow the stn posterior from where the OG was to the thickening in the stn that you identified earlier, this is where the acns (that contain the ACPs) branch off. If you do not see a thickening, do not worry; you can usually also find the acns by carefully grabbing the stn and lifting it gently. By doing so you should see two nerves projecting out the stn and winding their way to either side and into the surrounding tissue. Because the ACPs location on the acn is vaguely known, it is necessary to take a length of the acn to guarantee removal of the ACP. To do this, first grab the acn close to the stn, but not right at their junction. Then slide the spring scissors alongside the stn and cut the acn off of the stn while still holding onto the acn with the forceps (Figure 2.49). Finally, lift the acn and carefully cut away tissue that is holding it in place until you have a section of acn roughly 0.5 centimeters long. Cut this segment out, remove it from saline, blot it and place it in the appropriate tube. Do the same on the other side of the acn and you have obtained the ACPs.
11. Removing the STG: At this point, you are almost finished. If everything has gone well, these last steps should be fairly easy to carry out. To find the STG, continue following the stn posterior from where you found the acns. When you have reached the point that you first cut into the artery, you should see the stn starting to thicken into the STG. Occasionally, the STG is further posterior than where you cut into the artery and you will be required to cut open the artery a little more to have uninhibited access to the STG. The crab STG is fairly distinctive as it has two nerves projecting directly out of the sides (the anterior lateral nerves, aln) and another that projects posterior (the dorsal ventricular nerve, dvn) with a nearby junction where two more nerves angle posterior and to either side (the median ventricular nerves, mvn). The STG is also marked by a more opaque spot at its center that is the neuropil, or meeting of all the incoming neurons (Figure 2.50 A). In any case, once you have identified the STG, use the forceps to grab the stn just anterior of the STG and carefully pull it towards the anterior. Once there is a little tension on the STG, cut the dvn, the mvns, the alns, and finally the stn anterior to where you are holding it (Figure 2.50 B and C). After doing all this, the STG should be free and you can move it to a clear area of the dish. As with other organs, trim the

projecting neurons, remove it from the saline, blot the forceps and place it into the appropriate tube.

12. Now, if you were able to do all that, you have just carried out a complete dissection of the brachyuran crab. Make sure to discard the remaining stomach tissue, clean out the dish, and clean all the tools you have used. As with the macrodissection, rinse and scrub everything with water first, and then rinse with 70% EtOH. Also, make sure that your tubes are chilled and that you place them in the freezer as soon as possible.

IV. Selected Figures

Figure 2.5: Insertion of the needle at the base of the crab's tail for collecting hemolymph.

Figure 2.7: Drawing hemolymph from the base of a leg.

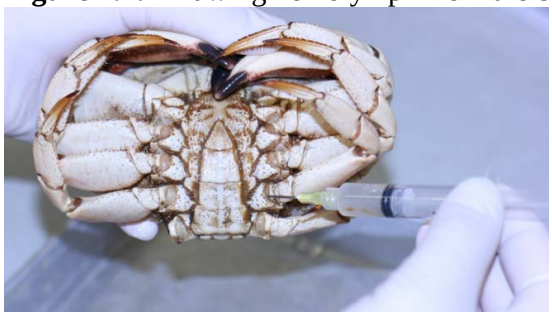
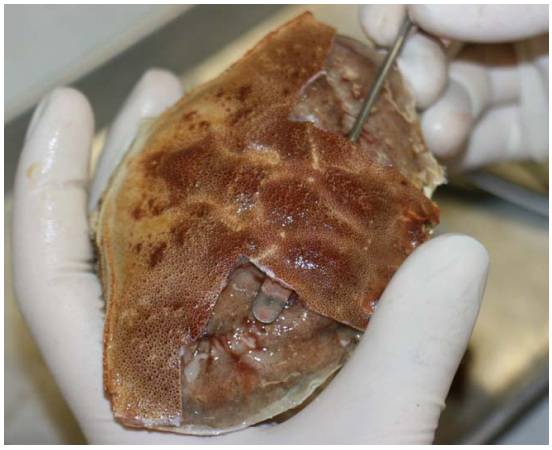


Figure 2.12: Shell should be cracked away as illustrated and hypodermis separated all the way across crab.



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.14: A. Crack the shell away from the front of the crab and around the eyestalks. **B.** After cracking it loose, carefully lift the middle of the crab shell away from the crab.

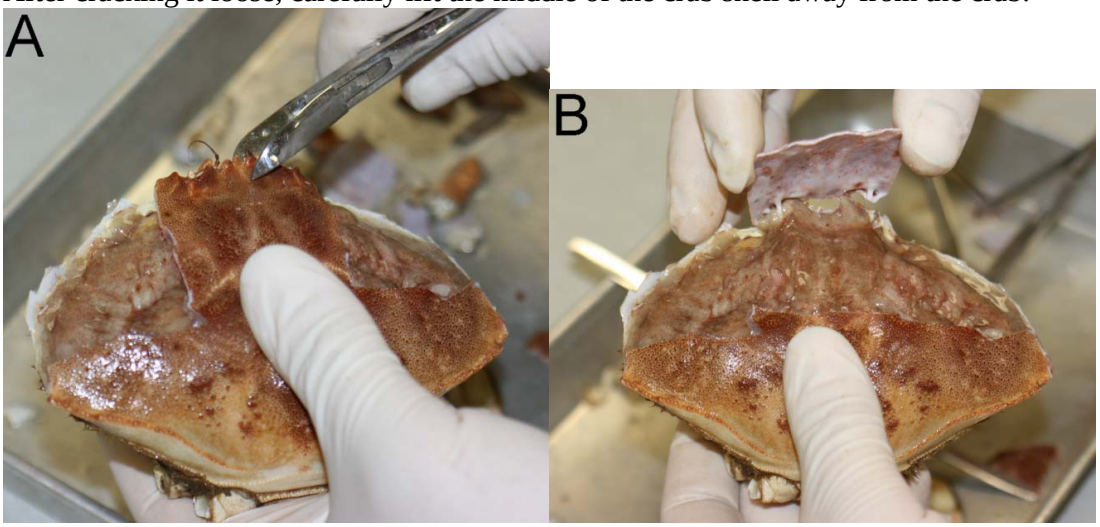


Figure 2.15: Removal of the eyestalks for later SG removal.

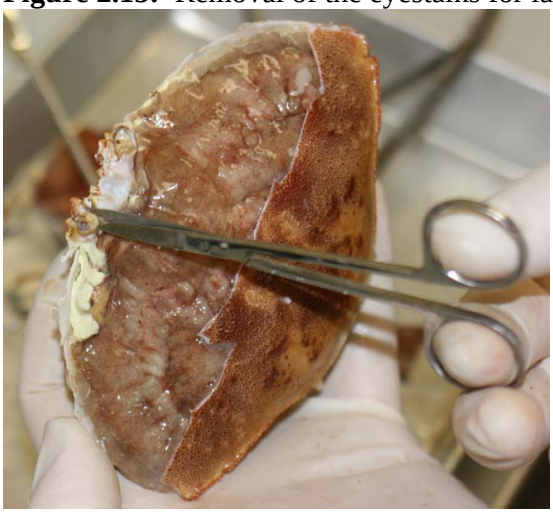


Figure 2.16: A. Use the spatula to separate the hypodermis from the lower shell. **B.** Use the scissors against the inside front of the crab to cut tissue away. **C.** Cut and scrape tissue away until a small bump on the inside of the shell and an arch structure created by the shell are revealed.

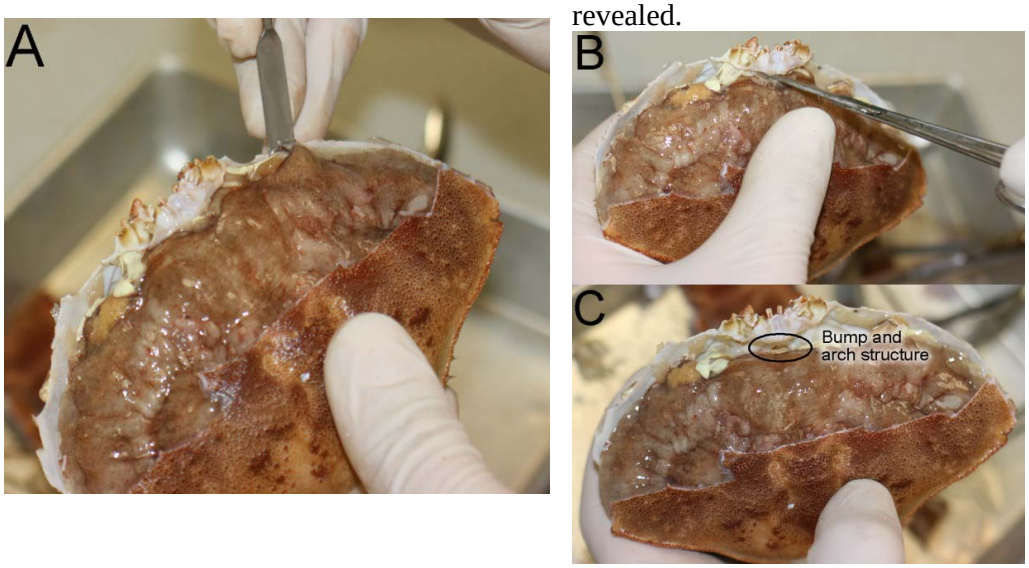
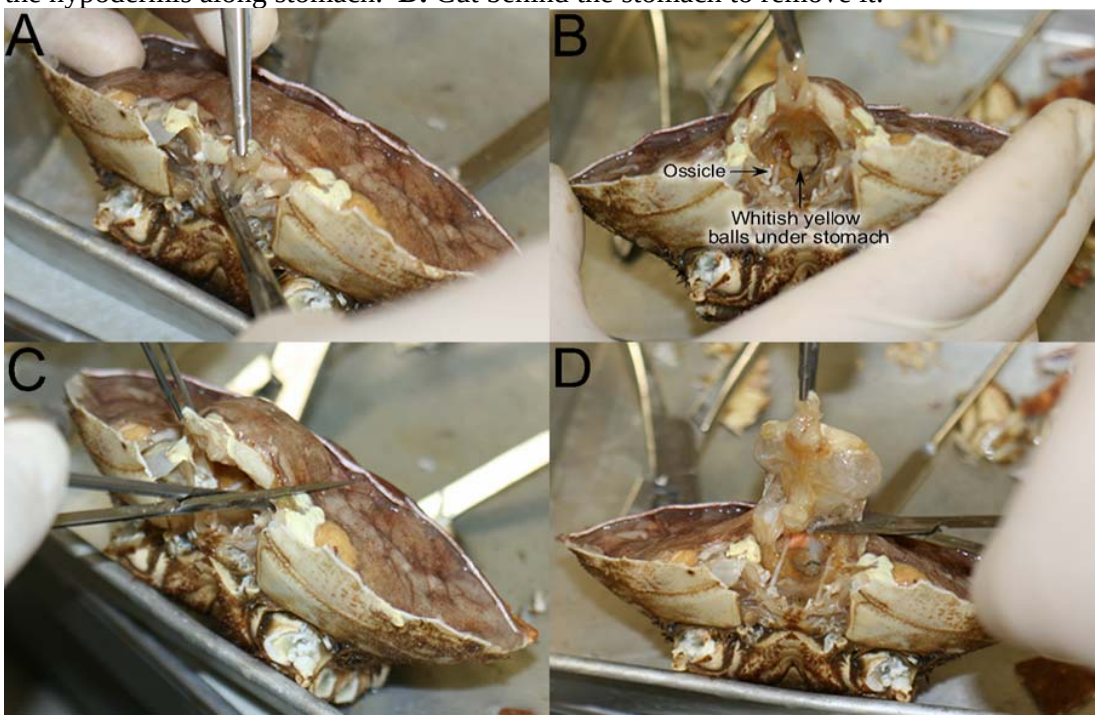
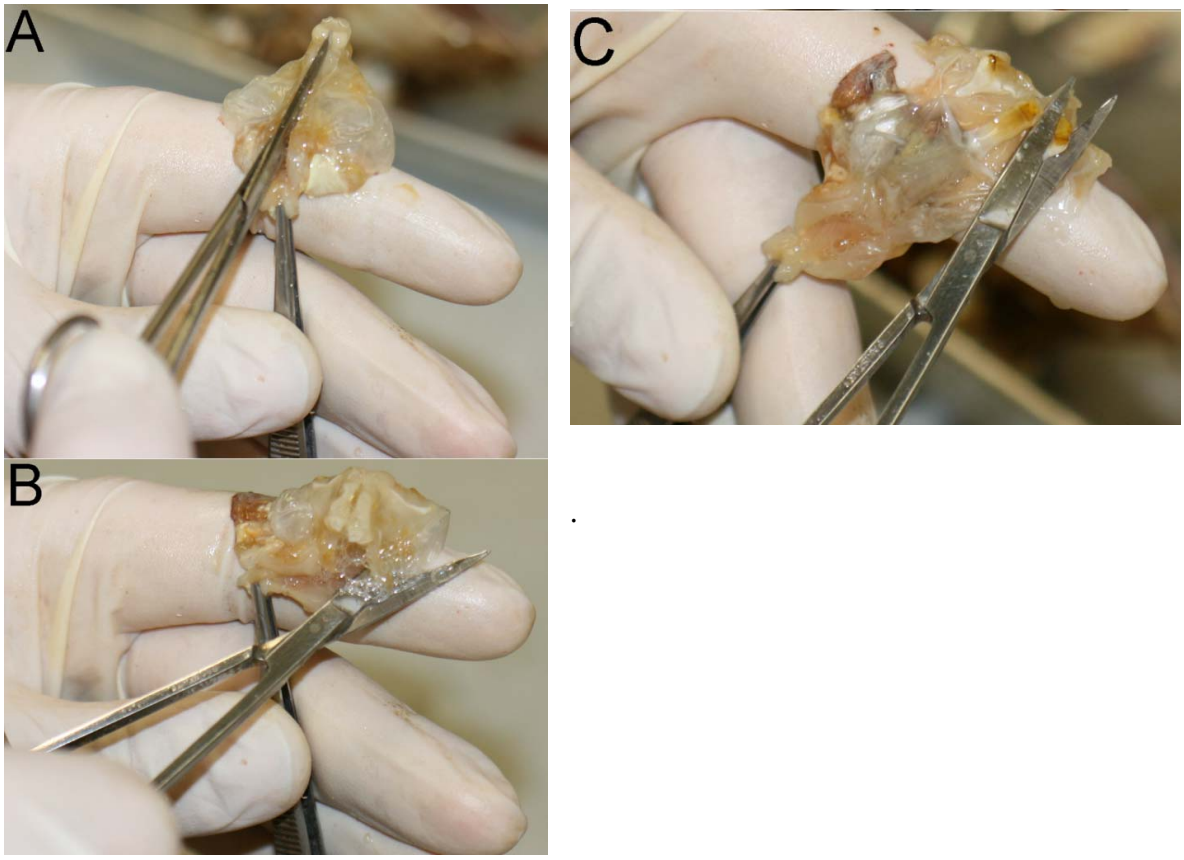


Figure 2.18: A. Grasp the lips with the toothed forceps and cut at the base of the oesophagus. **B.** Cut back, into crab, above ossicles, till you see the pylorus (two whitish yellow balls). **C.** Cut the hypodermis along stomach. **D.** Cut behind the stomach to remove it.



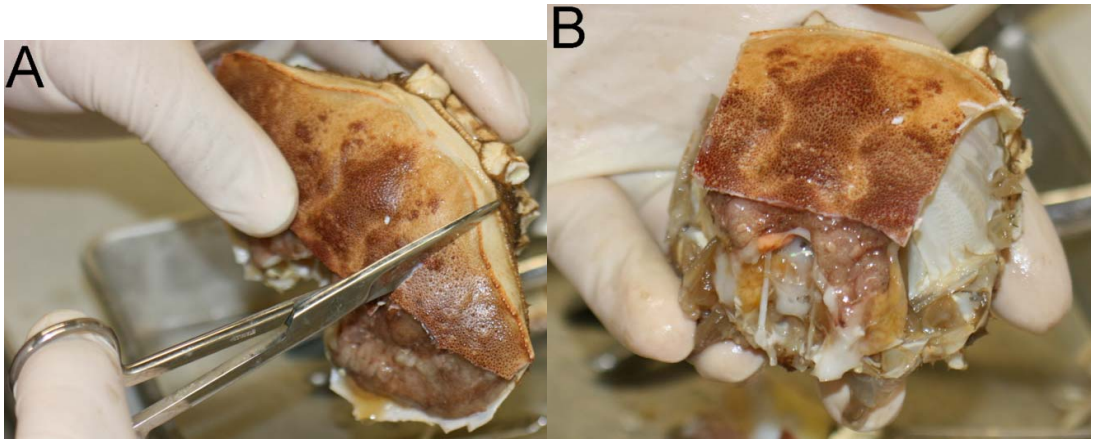
Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.19: A. Cut the oesophagus open straight back to the rear of the stomach. **B.** Make two cuts on either side of the stomach angling back to the rear corners of the stomach. **C.** Cut out all three teeth.



.

Figure 2.22: A. Cut the shell on both sides of the midline. **B.** Remove the tissue from both sides.



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.25: A. Remove the tail. **B.** Use forceps to clean out the tissue between the cardial ridges.

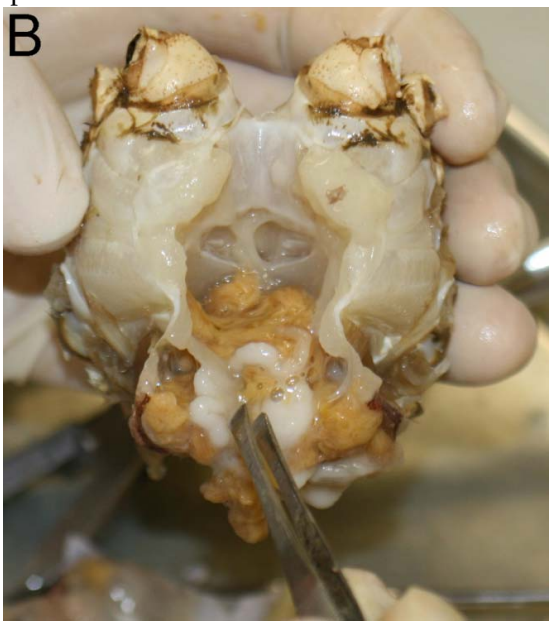
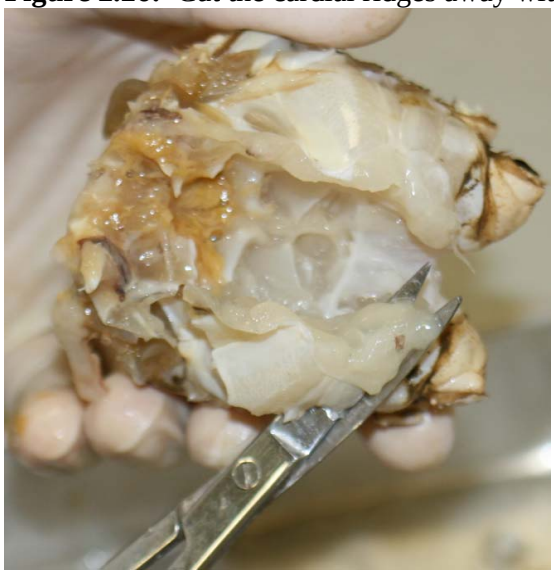


Figure 2.26: Cut the cardial ridges away with the large scissors and place them in saline.



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.28: A. The thoracic ganglia looks like a wagon wheel with ganglia projecting out of it. **B.** Cut around the TG to free it. **C.** Grab it with clean forceps and lift it out, cutting any remaining tissue as you do so.

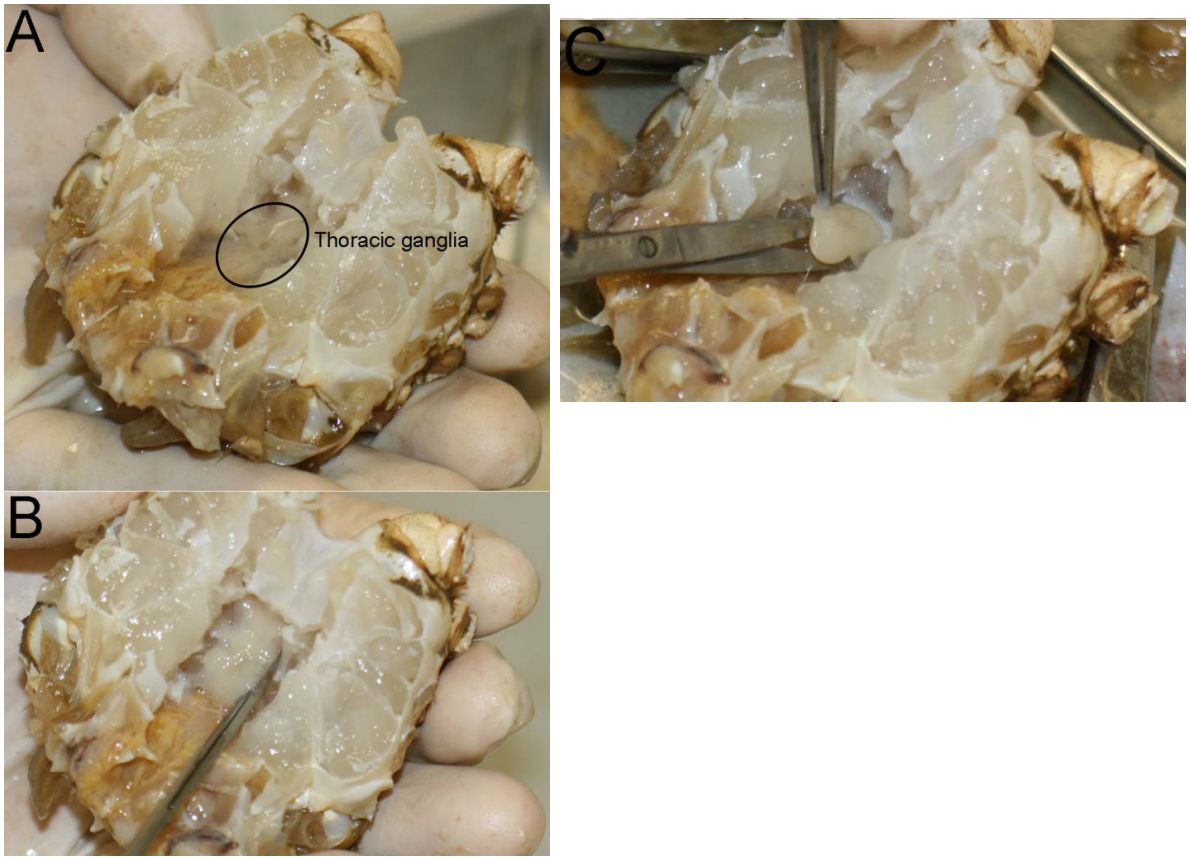
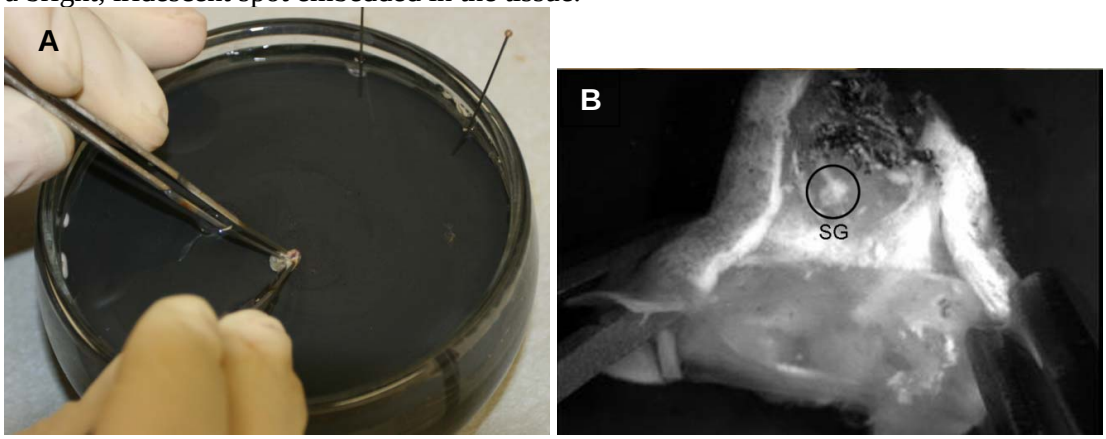


Figure 2.32: A. Using the sturdy, fine forceps, separate both sides of the eyestalk. **B.** Watch for a bright, iridescent spot embedded in the tissue.



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.37: The PO can be identified as highly iridescent ganglia spanning the inside of the cardial ridge.

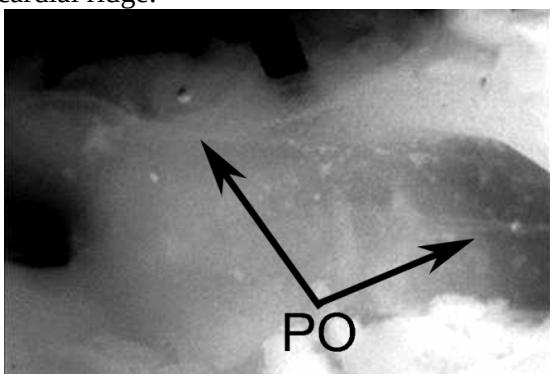
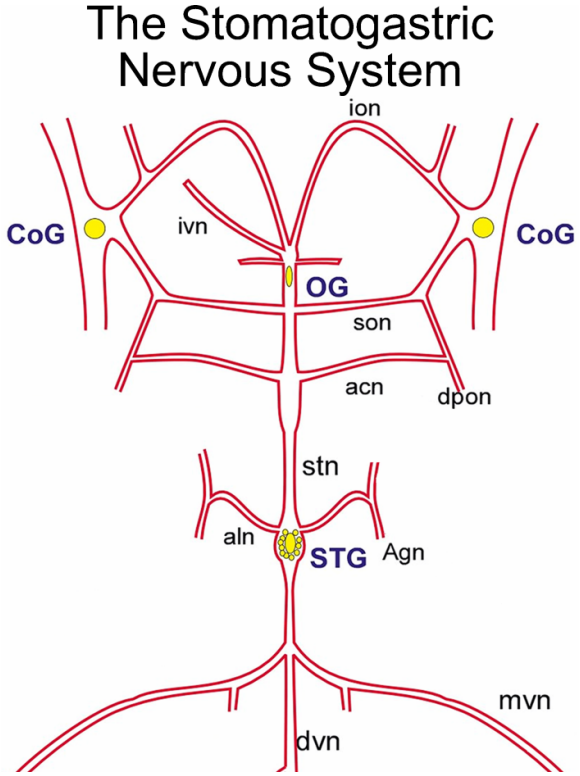


Figure 2.40: Schematic of the STNS. Refer to this during your dissection. Image taken from Farzan Nadim's Lab website at Rutgers University. (http://cancer.rutgers.edu/stg_lab/protocols.htm)

CoG	--	Commissural Ganglia	OG	--	Oesophageal Ganglia
STG	--	Stomatogastric Ganglia	ACP	--	Anterior Cardial Plexus
ion	--	inferior oesophageal nerve	*ivn*	--	inferior ventricular nerve
son	--	superior oesophageal nerve	*acn*	--	anterior cardial nerve
stn	--	stomatogastric nerve	*aln*	--	anterior lateral nerve
dvn	--	dorsal ventricular nerve	*mvn*	--	median ventricular nerve



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.41: **A.** Pin the bottom of the stomach, pulling it slightly tight. **B.** Next, pin the sides of the stomach, again slightly tight. **C.** Finally, pin either side of the oesophagus, just below the lips.

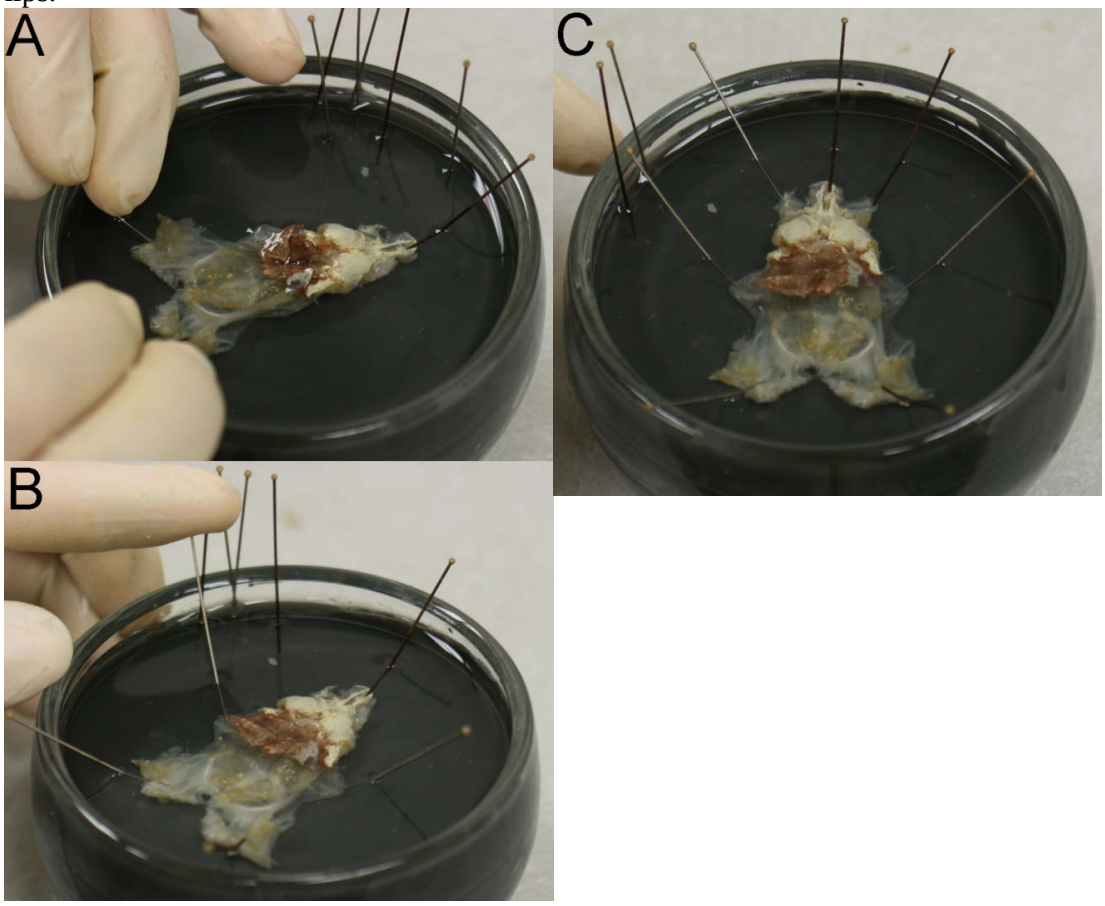


Figure 2.42: The brain (circled) embedded in the tissue above the stomach.

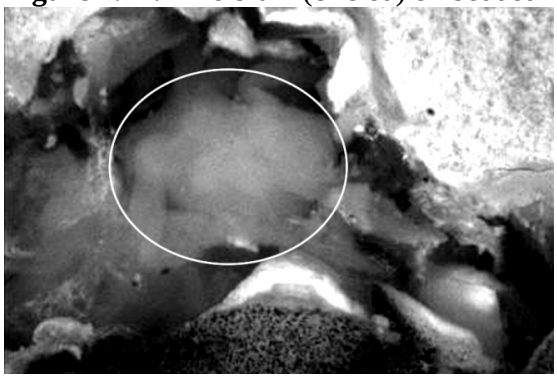


Figure 2.46: A. Look down between the muscles for the *stn* (circled). **B.** Carefully cut out the muscles. **C.** Continue cutting forward under the muscles, making sure not to cut the *stn*.

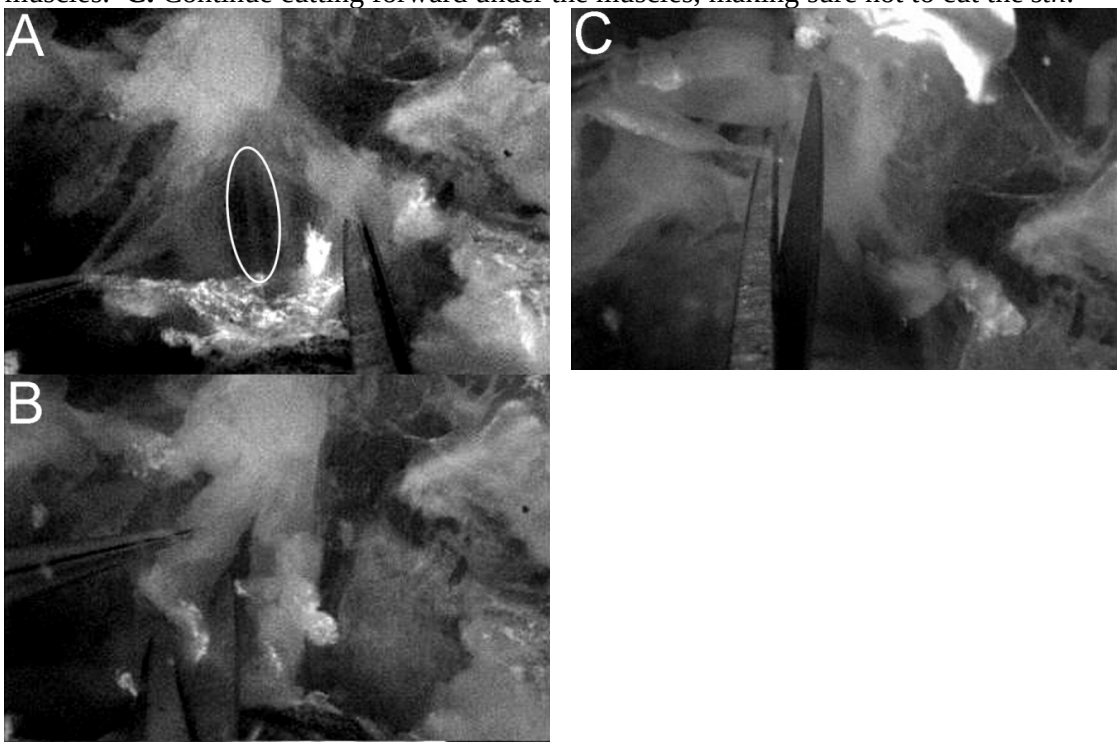
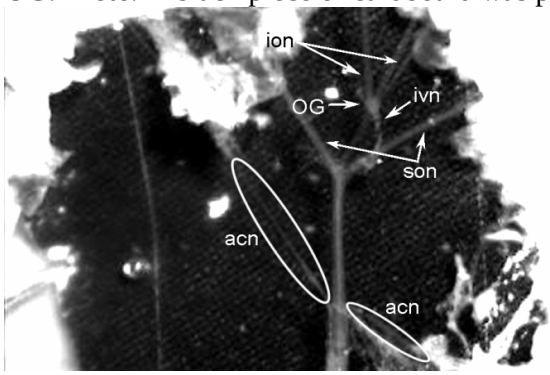
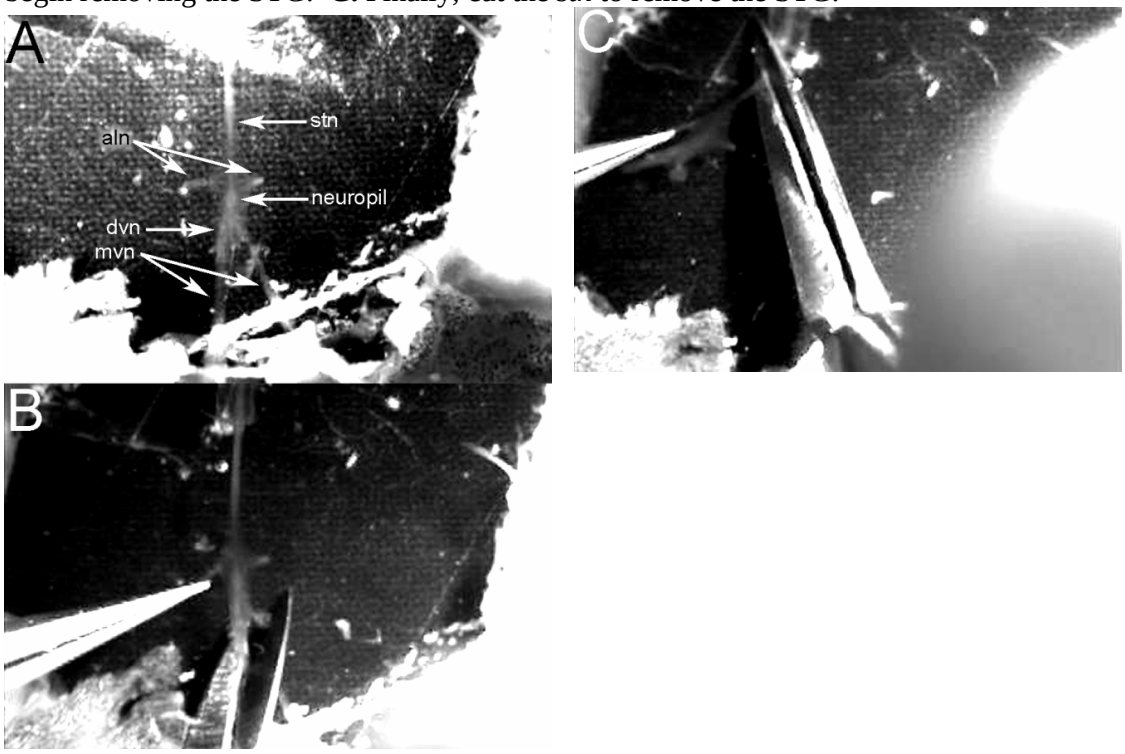


Figure 2.47: The forward (rostral) portion of the STNS showing the *acn*, *son*, *ion*, *ivn*, and the OG. Note: A black piece of cardboard was placed behind the STNS for better contrast.



Number:
Title: Marine Crustacean Dissection **Effective Date:** 10-12-2011

Figure 2.50: A. STG with the neuropil and the projecting commissures. **B.** Cut the *mvn*s to begin removing the STG. **C.** Finally, cut the *stn* to remove the STG.



Approval
I have read and approved this procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER:n/a **EFFECTIVE DATE: 02-18-2013**

TITLE: Procedures for nanoparticle exposure with live crabs

APPLICABLE SPECIES: *Cancer borealis, Callinectes sapidus*

1. Materials for Nanoparticle Exposure

1. Crab saline, composition in mM: NaCl, 440; KCl, 11; $MgCl_2$, 26; $CaCl_2$, 13; Trizma base, 11; maleic acid, 5; pH 7.45. Store at 4ºC.
2. 1 mL plastic syringe (BD Biosciences).
3. Small gauge needle, 26 ga.
4. 0.6 mL low retention microcentrifuge tube (Fisher).
5. Dissecting tools: side cutters, spatula, small scissors.
6. In-house synthesized nanoparticles.
7. Artificial sea water. Temperature is maintained between 50-57F (10-14ºC)
8. Exposure buckets.

2. Procedures for Nanoparticle Exposure in a Live Crab

1. Put on gloves. Measure desired volume of nanoparticle solution and/or crab saline using 1mL plastic syringe with 26 ga needle.
2. Set up two buckets with 1 gallon artificial sea water (10-14ºC) separately.
3. Remove crab from the tank and place it with dorsal side down. Inject the measured nanoparticle solution into the animal from the bottom leg joint and put it into the exposure bucket for one hour. The reference animal should be injected with same volume crab saline and left in the other exposure bucket for the same amount of time.
4. After desired exposure time, take animals out of the buckets and place them in ice to anesthetize the animals.
5. Follow the Marine Crustacean Dissection Protocol to dissect the crabs and obtain desired tissues.

3. Humane Endpoint Criteria after Nanoparticle Exposure in Crustacean

After nanoparticle injection, the presence of one or more of these criteria below may be signs to indicate morbidity in animals. Animals with these signs would be put into ice to slow down their body temperature hence the metabolism and ease the distress. After one hour in ice, the animal should be dead.

1. Tail away from the body
2. Constant uncontrolled movement in limbs
3. Mouth falling open

4. Approval

I have read and approved this procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER: **EFFECTIVE DATE: 10/12/2011**
TITLE: Procedures for *in vivo* microdialysis with live crabs
APPLICABLE SPECIES: *Cancer borealis, Callinectes sapidus*

1. Materials for *In Vivo* Microdialysis

1. Crab saline, composition in mM: NaCl, 440; KCl, 11; $MgCl_2$, 26; $CaCl_2$, 13; Trizma base, 11; maleic acid, 5; pH 7.45. Store at 4ºC.
2. 1 or 3 mL plastic syringe (BD Biosciences).
3. Needle from CMA 20 or 20 Elite package with the sharp tip cut off.
4. Syringe pump capable of handling microliter flow rates with adjustable syringe diameters.
5. Fluorinated ethylene propylene (FEP) tubing connectors (CMA). Stored in 70% ethanol at room temperature.
6. Microdialysis probe: CMA 20 or 20 Elite with 4 mm polyarylethersulfone membrane. It is preferable that the probe be rinsed with crab saline *in vitro* prior to use in vivo to remove residual glycerol used for storage of the probe.
7. FEP tubing, 0.12 mm inner diameter (CMA), or PEEK tubing, 0.12 mm inner diameter (Upchurch Scientific).
8. . 1.5 mL microcentrifuge tube.
9. 1000 mL pipette tip
10. Scissors
11. Hot glue and glue gun
12. 30% bleach solution in water
13. 95% ethanol.
14. Kimwipes or equivalent.
15. 1-2 sheets of plexiglass, 3/8" thick, sized to fit your salt water tank, or separate small tank for the animal to be isolated in.
16. Rotary tool, such as Dremel 7.2 V MultiPro Cordless (Dremel, Racine, WI)
17. Rotary tool chuck accepting 1/32" to 1/8" shafts (Dremel)
18. Rotary tool grinding wheel. (Dremel)
19. Plumber's epoxy (Mighty Putty) and something to cut it with (plastic knife works well)
20. Rotary tool drill bit, 1/32" (0.8 mm) (Dremel)
21. Small gauge needle, 25 or 26 ga.
22. Super glue gel or other cyanoacrylate glue.
23. Corrugated plastic covers for electrical cords
24. Collection bucket, i.e. something to hold collection tubes on ice, or a commercial refrigerated fraction collector.
25. Collection tubes.
25. Green food dye.
26. Dissecting tools: side cutters or rongeurs, spatula, small scissors
27. Camera

2. Procedure for *In Vivo* Microdialysis of a Live Crab

1. Cut desired lengths of tubing to extend the inlet and outlet tubing of the probe (tubing should be cut with a sharp razor blade or special tubing cutter). Attach to the inlet and outlet tubing with FEP connectors. Attach a needle to the inlet tubing with the same connectors.
2. Cut the tip off the 1000 mL pipette tip and place the top part in the 1.5 mL microcentrifuge tube. Place the tip of the microdialysis probe into the tube and use the pipette tip to hold the flexible top of the probe in place. Fill the tube with crab saline.
3. Fill a syringe with crab saline and attach to the needle. Ensure that the entire setup has no leaks by pushing crab saline through.
4. With the hot glue gun, apply a small ball of hot glue to the probe shaft about 1 cm from the probe tip, being careful not to get any glue on the membrane (See Fig. 1). This will prevent insertion of the probe too far into the crab. Let dry. Place probe tip back in the liquid in the microcentrifuge tube.
5. Set up the tank where the crab will be housed, either with plexiglass sheets in the tank to section off one corner or with a small separate tank.
6. Put on gloves. Remove the crab from the tank and place it in a bucket filled with ice dorsal side up. Note the time when the crab was first placed on ice. Leave the caudal area of the dorsal shell (above the pericardial sinus) free from ice. Wash this area first with 30% bleach and then 95% ethanol. Vigorous cleaning with the bleach should remove some waxy substance from the shell. Dry the shell. Leave the crab on ice for at least 15 min.
7. While the crab is on ice, set up the surgery area in a cold room, preferably in the same room as the crab tank. Set out a dissection pan and fill it with a thin layer of ice.
8. Cut approx. 2 cm of plumber's epoxy and mix it by kneading the two colors together. Once a uniform color is achieved, form the epoxy into a tube shape about 4-5 cm long. It is often helpful to change gloves after this, as the epoxy leaves a residue.
9. Once the crab is anesthetized (approx. 15 min.), remove the crab from the bucket and place it dorsal-side-up in the dissection pan for surgery.
10. Put on safety glasses. Use the rotary tool with attached grinding wheel to score several lines in the shell above the pericardial sinus in a cross-hatching pattern to increase the surface roughness of the shell. This will improve the glue's hold by providing a cleaner surface and more surface area. The location of the pericardial sinus can be determined by looking for a pentagonal area with lines surrounding it in the caudal area of the top shell (See Fig. 2). Clean this area with 95% ethanol to remove dust and provide an aseptic surface.
11. Put the drill bit in the rotary tool. Drill a hole in the shell in the middle of the pericardial sinus, as indicated by the external features of the shell. This hole should be drilled carefully. Ideally, the epidermis of the crab under the shell will not be punctured by the drill bit. This is achieved by applying minimal downward force on the rotary tool and stopping as soon as dust stops coming out of the drill hole. If the hypodermis is punctured, hemolymph will come out of the hole in pulses corresponding to the crab's heartbeat. If it is not, gently puncture it with the small gauge needle. Place a Kimwipe over the hole to absorb hemolymph and dry the shell.
12. Place the tip of the probe inside the hole, stopping when the glue ball meets the shell. The probe tip should be aimed in a rostro-ventral direction, and putting the probe in should stop the flow of hemolymph out of the hole in the shell.
13. Dry off the shell surface.
14. Put cyanoacrylate glue gel in a circle around the hole (see Fig. 3A). Make sure not to get the glue too close to the hole, as it will then go into the crab and/or damage the probe membrane.

15. Put the epoxy tube around the probe and form it into a conical shape securing the probe shaft to the shell surface (See Fig. 3B). Put a dab of cyanoacrylate glue gel at the top of this cone. Test the fluidic integrity of the system by pushing crab saline through the probe. If the probe tip was damaged during implantation, liquid will go in but not come out the other end of the tube. If the tubing was damaged, you will not be able to push liquid into the tubing.
16. Take a 20-30 cm length of corrugated plastic electrical cord cover and place around the tubing coming out of the probe. Put one end flush with the epoxy cone on the crab's shell. Push it gently into the epoxy. The corrugated plastic prevents the crab from snipping or crushing the probe tubing with its claws.
17. Use hot glue on the outside of the plastic cord cover where it meets the epoxy to provide a secure hold. Try not to get the hot glue on the crab, as it is uncomfortable for them (see Fig. 4).
18. Wait for the epoxy to harden. This normally takes around 15 min. The crab should not be kept on ice for more than 45 min if possible. Beyond 45 min, it is less likely to recover from anesthesia. Check for leaks again.
19. Take the crab in the pan to the tank. Place the pan in the tank, ice and all. Flood the pan with tank water. Hold onto the crab by its claws, with the stomach and claws pointing up. Expel gas from the stomach by gently shaking the crab.
20. If the tank is shallow, simply let go of the crab and allow it to fall to the bottom. If the tank is deep, put the crab in a net and use that to gently lower it to the substrate. The crab should start to awaken and move. Check the fluidic integrity of the system.
21. Put the lid on the tank and set up the syringe pump and collection device. A microdialysis swivel can also be used. Check that liquid can flow through the system.
22. Set the pump at the desired flow rate. Conduct the experiment as planned. At least 30 min should be allowed for equilibration of the probe following any changes in flow rate. It is our experience that the crab should be allowed to recover from the surgery for 24 hours prior to conducting experiments on it, unless measuring surgical stress is the aim of the experiment. The crab usually will not eat within 24 hours after surgery, even if in an unfed state. Most post-surgical complications occur within 24 hours after surgery as well.

Post-surgical complications include: the crab bleeding out from the drilled hole due to a poor seal with the glues. This is fatal and can be avoided by careful technique, especially in puncturing the epidermis and creating the seal with the glue. The probe may become dislodged—sometimes it can be reattached but often this damages the probe beyond reuse.
23. Once the experiment is done, stop the syringe pump and put the crab in a bucket with ice.
24. Fill the syringe with 0.5 mL of green food dye. The dye can be diluted somewhat, and alternatively, all tubing except those directly connected to the probe can be removed to keep them from getting dye in them. Push the dye through the system either by hand or with the pump at a high flow rate (5-10 µL/min). Once the dye appears at the outlet, you can stop and disconnect the syringe.
24. While the crab is on ice, set out the dissection pan, spatula, side cutters, and small scissors.
25. Once the crab is anesthetized (after at least 15 min on ice), place the crab in the dissecting pan and begin by using the side cutters to remove the claws and legs at the base where they meet the body.
26. Use the side cutters to cut around the outer rim of the crab shell. Then use the spatula to reach in between the top and bottom shells and separate the hypodermis from the upper shell.
27. Use the side cutters to remove the upper shell back to the pericardial region, avoiding the area behind the pericardial ridges. Use the scissors to cut the connective tissue between the

upper shell and the pericardial region so that you can pull the shell off. The stomach and brain can be removed as in the typical dissection procedure.
28. Place one hand on the top half of the shell, with the thumb hooked slightly into the area between the pericardial ridges. Use the other hand to secure the bottom half. If the front of the crab has been dissected away, the other thumb can be placed immediately ventral to the first thumb. Pull the top half of the shell away. The hypodermis in the pericardial sinus should stay attached to the shell. Upon removing the shell, you should be able to see where the probe tip was located. Look for a small amount of green dye under the probe tip, probably on the surface of the heart. A photo can be taken for documentation of proper probe placement.
29. Finish dissecting the crab and dispose of it according to the procedures outlined by the animal care facility. Dispose of sharps as indicated. No other hazardous materials are used.

3. Figures

Figure 1: Placement of a small amount of hot glue near the probe tip to act as a guard to prevent insertion of the probe too far into the crab.

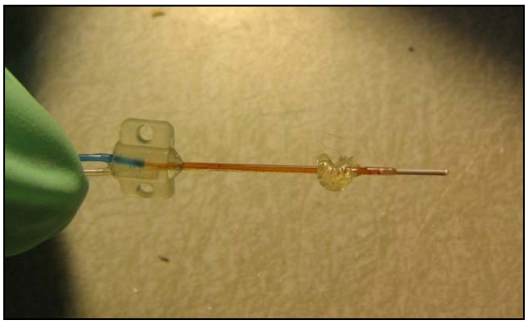


Figure 2: Location of the pericardial sinus in A) *Cancer borealis* and B) *Callinectes sapidus*. Dotted lines indicate the pentagonal marks on the outside of the shell that define the pericardial sinus. C) Preparation of the surface in *C. borealis*, showing the rotary tool with grinding wheel.

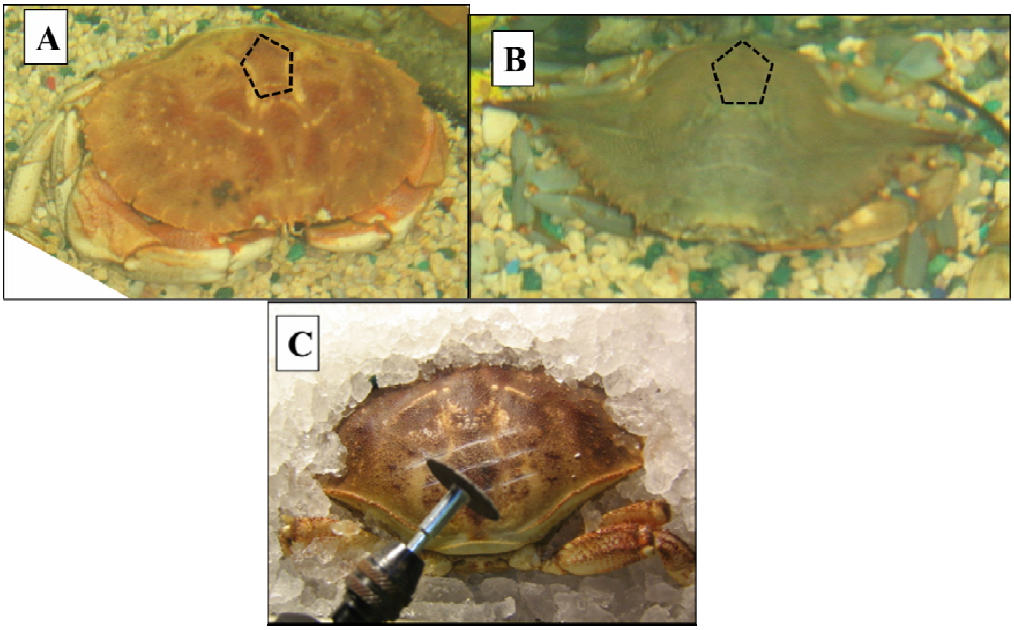


Figure 3: Placement of A) cyanoacrylate gel around hole and B) epoxy putty to secure probe.

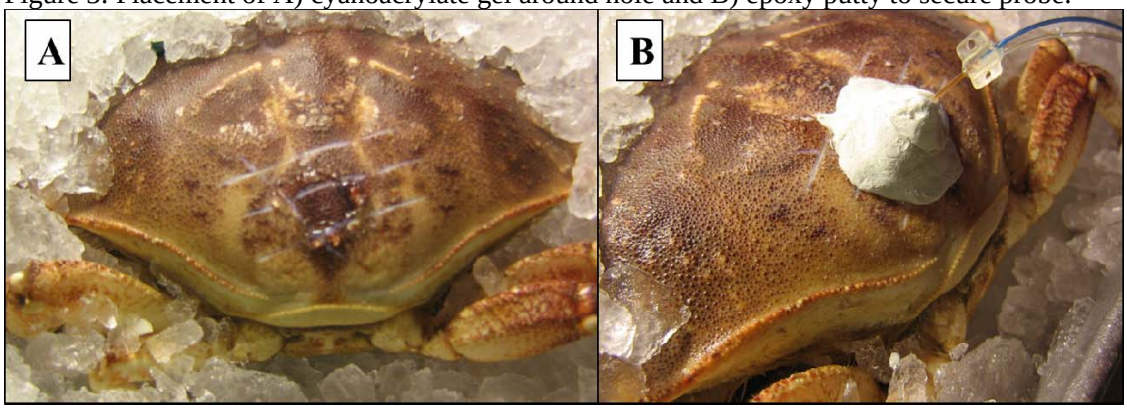
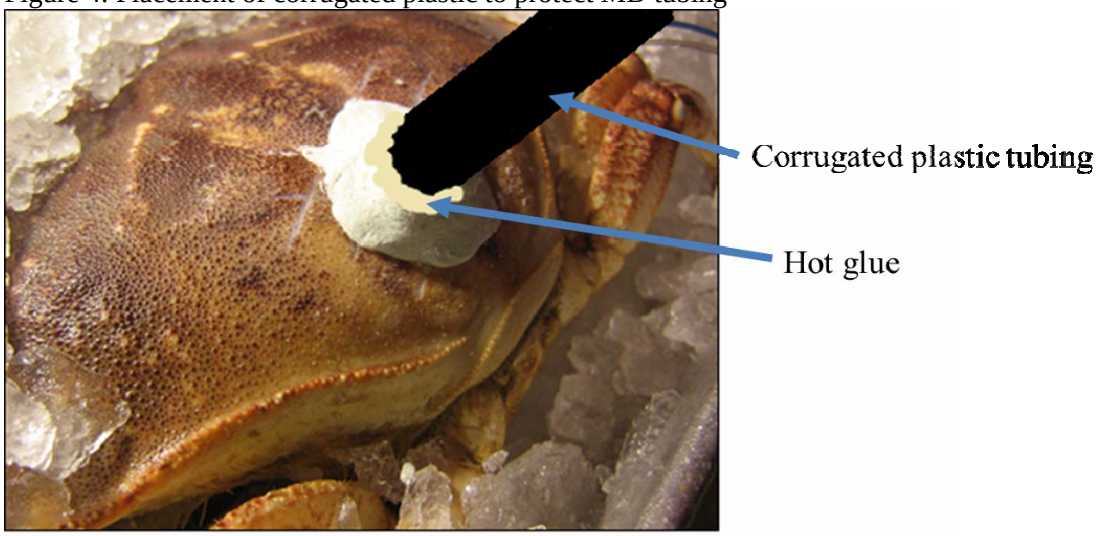


Figure 4: Placement of corrugated plastic to protect MD tubing



4. Feeding MD Experiment

While liquid flows through the probe and is collected, the animal is given food. Raw fish, crayfish, or shrimp are preferred by the crabs. Sample tubes are changed approximately every 30 min. Liquid is collected for at least 2 hours post-feeding. Food is removed with a net after the experiment. Crabs will usually only eat every 48 hrs., but occasionally will eat daily.

5. Modification by adding affinity agent to the perfusate

Occasionally, solid affinity agent may be added to the liquid flowing through the probe. This solid affinity agent is typically a solid particle ranging from 100 nm-2 µm in diameter, so it will not pass out of the probe into the animal. The surface of the solid particle is modified by attaching different functional groups or antibodies that have affinity for compounds of interest. Although these affinity agents increase the amount of analyte recovered, they do not remove enough to cause perturbation to the animal. The same liquid is used: crab physiological saline. These particles present a particular challenge in keeping them in solution and preventing clogging of the tubing and probe. The typical solution is to put a small metal ball bearing inside the syringe. The syringe pump is then placed on a rocking platform at an angle of approximately

45º to the axis of rotation. The motion of the platform thus causes the ball bearing to roll up and down the syringe, agitating the particles to keep them in solution.

Occasionally, clogs of particles will occur in the tubing. This is prevented by making sure the cut ends of the tubing are flush to each other within the FEP connectors. It can be remedied by using a high flow rate to push the particles out, but occasionally sections of tubing will have to be cut out and replaced. It is even possible for something that cannot be accessed to become clogged (i.e. the probe itself), which causes the experiment to be discontinued.

6. Approval

I have read and approved this procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

Number: **Effective Date**: 10-12-2011

Title: Tank Cleaning Procedures

All water quality tests and tank cleaning are the responsibility of Li lab personnel.

Part 1: Routine Tank Cleaning Instructions

General instructions:

- Water changes are done every 10±2 days. Water quality is also checked at this time.
- Tanks may be cleaned with animals in them. However, the danger of being pinched by crabs while cleaning a tank is high, so it is recommended to move the crabs to another tank while cleaning with the gravel vacuum, and then move them back once this is done.
- There are two ways to get the gravel vacuum to work.
  - I prefer to put the end of the tubing in the drain on the floor and create a vacuum with it by filling the cup with water, then letting the water in the cup drain. While it is draining from the cup, place the cup back under the water. I can demonstrate how to do this.
  - We also have an apparatus for creating vacuum with the sink, but it doesn't work very well.
- Always mix up the salt water in a bucket and check its salinity with the hydrometer before adding it to the tank, especially if there are already animals in the system. Double-check salinity of the tank as a whole after adding the new water to the tank.
- Water to use is RO water from the hose on the wall. This is turned on with the valve in the hallway. Stepstools to reach the valve are found in the storage room where all the clean cages, mouse food, etc are kept.
- 10-day water changes are approximately 10-20% of the water.
- It is important to make sure the water levels are high enough for filters and chillers to work properly.
- Scrape algae off the walls of the tank with the scrubber weekly.
- Do not ever let any pump run without water in it—running the pumps while they are "dry" can cause damage.

Tanks 1, 2, and 6:

These 3 tanks are connected to the same filter and chiller apparatus, so you have to be a bit careful about how you clean them.

1. Leave the filtration apparatus and chiller on.
2. On one tank, close the valves that drain water and supply water to the tank. This isolates this tank from the rest of the system.
3. Use the gravel vacuum to clean the rocks and remove 10-20% of the water. This can be estimated by noting the tank water level and looking for a 10-20% decrease in this level.
4. Clean the glass sides of the aquarium by using the algae scrubber to remove any built algae from the sides of the tanks.
5. To refill, add seawater prepared as mentioned previously to the tank. Add new water to the tank you have just cleaned.
6. Close the valves leading to and from the next tank you are going to clean.

Number:
Title: Tank cleaning procedures **Effective Date:** 10-12-2011

a. In a 2-tank system, now open the valves leading to and from the tank that has already been cleaned.
b. In a 3-tank system, leave the just-cleaned tank's valves in the OFF position. Once two of the three tanks have been cleaned, close off the third tank and then re-open the valves on the two clean tanks.
c. This is to prevent mixing of the water from the tank(s) that have not been cleaned yet with the tank(s) that have been cleaned.
d. Repeat steps 1-5 on all tanks.
7. Change the white filter bag on the filter apparatus by replacing the dirty one with a clean one from the shelf. Rinse a bag from the shelf before putting it into the filter. Put the dirty one on the dirty equipment cart in the hall outside the cage cleaning room to be cleaned.
8. Remove the top of the protein skimmer and clean it in the sink. Wipe out the bottom chamber as well. Replace.
9. Do a final adjustment on the water level. The best way to tell if enough has been added is to check the water level in the tanks AND the filter. There is a piece of tape on the body of the filter that represents where the water should be. Add water to the system until the water level in the filter is near that tape line.
10. Adjust the protein skimmer so that only a few bubbles reach the top chamber. There is a valve on the tubing leading to the skimmer that you can open and close to adjust the amount that flows into it. The liquid that collects in the top chamber should be a dark brown.

Tank 3

1. Unplug the tank (2 plugs) while cleaning, or press the "TEST" button on the plug on the wall.
2. Clean the rocks with the gravel vacuum—remove 10-20% of the water. Judge by monitoring the water level.
3. Also clean the sides of the tank with an algae scrubber to remove built algae on the glass.
4. Add water back to tank, prepared as above. In this tank, only 10 gallons need to be replaced.
5. The water level is correct in this tank when it reaches at least halfway up the gravel bags on the back of the filter. The water should not be higher than the top of the bags. The black coils of the chiller must be covered in water.
6. Rinse the cotton pieces from the filter out in the sink.
7. Plug it all back in, or press the "RESET" button on the plug on the wall.

Tanks 4 and 5

1. Unplug the filter on the tank while cleaning. They are plugged into the power strip on the wall.
2. Clean the rocks with the gravel vacuum to remove 10-20% of the water. Judge by monitoring the water level.
3. Clean the sides of the tanks removing built up algae with an algae scrubber.
4. Add water back to the tank, prepared as above. This tank only needs 1-2 gallons to be replaced.
5. The water level in this tank is correct when it reaches the bottom of the black plastic on the tank.
6. Rinse the sponge, carbon, and biological medium from inside the filter in the sink and replace.
7. Refill the filter with water from the tank before plugging it back in. These filters will be damaged greatly if they run dry, so always fill up the filter before plugging them back in.

Number:
Title: Tank cleaning procedures **Effective Date:** 10-12-2011

Water Quality Tests

Water quality should be checked prior to changing the water.

Salinity: checked with a hydrometer. For the floating hydrometer/thermometers, the water level should be at the green line on the hydrometer. For the benchtop hydrometer, immerse deeply in the water and allow to fill with water. Remove from the tank and let the indicator arm settle. It should indicate a density around 1.02. Acceptable range is 1.01-1.03.

Ammonia/nitrate/nitrite/pH: commercial kits for testing these substances are purchased from an aquarium supplier. Follow the instructions on the kit. Acceptable ranges are ammonia <0.025, nitrate <30ppm, nitrite <0.5ppm, pH: 8.2-8.4. If ammonia, nitrate, or nitrite is too high, a commercial aquarium product designed for the purpose of lowering these compounds is used according to its instructions. If pH is incorrect, a commercial buffer powder is added according to its instructions.

Dissolved oxygen is measured with a meter. The meter is first calibrated with the probe unplugged and the meter in "CAL" mode. The indicated screw is turned until the value is 0.00. The probe is then plugged in and the cap removed. The probe is allowed to equilibrate to the air. The indicated screw is the turned until the value shown is 20.9. The probe is then switched to "DO" mode, and can be put into the tanks to measure dissolved oxygen. Usually, it is at its maximum, which is 9-12 ppm for water at the temperatures we keep our crabs at. If it is below 5 ppm, it is too low and can cause stress to the animals. The flow of water through the filter is increased to add aeration, and if that is not sufficient, an air stone connected to an air pump is used. This is often necessary in the smaller tanks.

Regular cleanings are documented by initialing the sheet on the inside of the crab room door.

Part 2: Annual Cleaning

All tanks are drained for a thorough cleaning annually. This cleaning requires that the tanks be completely empty of animals. It will also take several days to a week for the tanks to re-establish the proper biological components of filtration. It is also important to annually check for metal corrosion and take preventative measures.

Tanks 1, 2, and 6

1. Start draining water using the gravel vacuum.
2. Unplug the pumps and chiller.
3. Remove water from the filter box once the water level in the tanks is below the intake tubes and clean out algae in the filter box with sponges.
4. Remove the remaining water with the gravel vacuum.
5. Clean accumulated gunk out of the protein skimmer.
6. Clean algae off the aquarium walls and plumbing.
7. Take the green sponge out of the filter and rinse it several times in salt water. This sponge houses the beneficial microbes that detoxify many of the animals' waste products, so we do not want to kill them by rinsing in fresh water. Squeeze the sponge while it sits in a clean bucket of salt water to do this. Replace the salt water when it becomes dirty. Continue to do this until the water coming out of the sponge is no longer colored.
8. Replace the filtration bag.
9. Check the bag containing biobeads. If they are brown in color, bleach them overnight in 30% bleach solution. After bleaching, rinse very well and allow to sit overnight in fresh water with a commercial chlorine detoxifying product (most aquarium chemicals fit this requirement).

Number:
Title: Tank cleaning procedures **Effective Date:** 10-12-2011

10. Refill the aquaria. They may be filled with fresh water and have the salt water added later, since no animals will be in the tanks. There is no need to pre-mix the water.
11. Add the appropriate amount of salt, measured by the hydrometer. Approximately 1 g of salt will be required per 5 gal of water. More detailed instructions on proportions are found on the salt box. Different manufacturers vary.
12. Replace the green sponge and allow the filter box to fill with water to the line indicated with tape.
13. Turn the pumps and chiller back on.
14. Reprogram the temperature controller if necessary, referring to the manual.
15. Wipe down the outside of the tanks and filter box to remove any salt.
16. Check for rust. Remove with CLR cleaner if necessary.
17. Allow several days for the tank's microbial community to cycle before putting animals in the tanks. Put only a few in at first.

Tank 3

1. Unplug and turn off. ON/OFF switch is to the left of the temperature knob.
2. If the tank is far away from the sink, pull/push tank closer to the sink. It may help to take out a few buckets of water first. The top and bottom halves of the tank are not attached, so be careful pushing from the top.
3. Remove filter bags, white cottony material, and the gravel. Place them in sink.
4. Remove coarse and fine netting below the filters, paying attention to not drop rocks/ sand down into the coil area. Place in sink
5. Remove Plexiglas above coils. Place in sink
6. DO NOT push/pull on cooling coils. Avoid applying any pressure on them at all.
7. Take hose with threaded fitting at one end and attach it to the valve in the front of the tank. Take care to not get any water on the pump – this will be harmful to it. Take the other end of the hose and attach it to the deionized water outlet using the spigot.
8. Start siphoning out water from the front compartment: fill up a hose with water and close one end of it with a finger, submerge that end in the water. Gravity will do the rest. Try to pick up as much rock/sand/stuff out with the hose.
9. Open valve, turn water on. This will cause water to start running into the front compartment, flowing up through the tubing of the tank.
10. Using this extra water coming up, continue picking up stuff from the front. The water will also flow in from the side compartments in the front. Rapid switching of the siphoning hose between compartments should allow you to prevent water from flowing into the back compartment.
11. When all stuff has been siphoned out, pull out the little netting pieces on top of the holes in the bottom of the front compartment right in the middle of the coils. Place in sink.
12. Stop the water flow into the front compartment by turning the water off. Detach that end of the hose and let the remaining water in the front compartment drain out through the tubing and the hose.
13. Switch the siphon to the back compartment.
14. Sometimes the tanks start growing algae. Scrub the sides and base of the back compartment and siphon out the solid stuff floating around.
15. Set the siphon on the tank and let it go by itself, making sure the hose will not whip out of the tank and get everything wet.

Number:
Title: Tank cleaning procedures **Effective Date:** 10-12-2011

16. While the back compartment is being emptied, clean the filters. Squeeze/massage white cottony material until all the stuff gets out of it. If hopeless, throw away and get new. The cotton is stored in the main lab space. Rinse out gravel and make sure the netting/bag does not have any holes in it. If it does, there are additional bags in the storage containers.
17. If the netting has started growing stuff or has acted as a filter and has stuff stuck on it, scrub gently to get rid of it.
18. Clean out small netting pieces, and the long rectangular coarse and fine netting that was under the graver bags.
19. Clean Plexiglas covering of the coils.
20. Replace small netting over the holes at the bottom of the front compartment.
21. Replace Plexiglas, the two nettings and all of the filters.
22. Close the valve at the front of the tank, unscrew the hose and remove.
23. Reattach other end of that hose to the deionized water outlet and start filling up the tank. Fill up both the front and back compartments.
24. Add about 4kgs of salt to the tank while it is filling up.
25. Mix up the salt while filling
26. Once the water level in the back compartment is about 2 inches from the top of the tank, stop filling.
27. Move tank back to its place.
28. Plug in and turn on. Reprogram the temperature controller if necessary. Refer to the manual for this.
29. Check back in about an hour to see how the water level and salinity are doing. The tank may require more salt and/or water. The temperature will take longer, check that the next day.

Tanks 4 and 5

1. Unplug the filters and air pump(s).
2. Clean the filters in the sink. They can be disassembled and the inner parts, including the impeller, cleaned. Refer to the manuals for details. Remove any scummy buildup.
3. If necessary, replace the carbon, sponge, and/or biological substrate components of the filter.
4. Remove all the water from the tanks using the gravel vacuum.
5. If the substrate is still not clean, the rocks can be removed and rinsed in a bucket.
6. Refill the tanks as above.
7. Put the filters and air stones back. Fill the filters with water before plugging them in to ensure that the pumps do not run while dry.

Corrosion Monitoring/Prevention

- Look for rust spots (iron oxide) or white spots (other metals' oxide) on the metal parts of the tanks and tables. Clean these off with CLR and SOS pads if necessary. Rinse well before re-filling, as these products can negatively affect water quality.
- If this is insufficient, more drastic measures will need to be taken to remove the rust. These include commercial products for rust removal, such as Naval Jelly and muritatic acid. If these are unsuccessful, mechanical means will need to be used, such as rust stripping pads for an electric drill. Rust will need to be removed with these products down to the bare metal. These will require the tank to be taken out of commission for a while. Cover the bare metal with a rust-preventing paint.

Number:
Title: Tank cleaning procedures **Effective Date:** 10-12-2011

- White spots will also show up on the plastic and glass parts of the tanks. This is lime and calcium buildup due to hard water. These can be removed with CLR. In extreme cases, they can also be soaked in vinegar (5% acetic acid). Rinse well before adding water. Sometimes a razor blade can also be used to scrape these off.
- Although we use rust-preventing paint, it is not sufficient due to the extreme humidity and salt water of the crab room. We need to check for corrosion annually.
- A rust-preventing spray can also be applied to stainless steel surfaces, and this should be done annually. The surfaces must be clean, dry, and free from rust. The product comes in an aerosol spray can. Spray on the surface and wipe this oily substance around to get full coverage.

Approval
I have read and approved this procedure.

Signature	Printed Name and Credentials	Date
DVM Signature	DVM Printed Name	Date

UNIVERSITY OF WISCONSIN-MADISON
SCHOOL OF MEDICINE AND PUBLIC HEALTH
LABORATORY ANIMAL RESOURCES

STANDARD OPERATING PROCEDURE

NUMBER: 300 **EFFECTIVE DATE:** February 16, 2011

TITLE: Standard Animal Room Procedures

1. **Check and Record Room Environmental Conditions:**
   - **1.1.** Room air temperature, relative humidity, relative room air pressure, and light cycles shall be maintained in accordance with *The Guide*[1] unless otherwise approved by the IACUC and posted. The minimum and maximum room air temperature and relative humidity will be recorded daily on the Daily Environmental Log.
   - **1.2.** See applicable 300 series routine husbandry SOPs for the required room environmental conditions for each species.
   - **1.3.** Check the room logbook for any special instructions. (See 7.1.4.)
2. **Observe Animals:** Every animal shall be observed at least daily by the LAR staff for health status, food and water consumption, urine and feces production, behavior and general appearance. Abnormal or unexpected conditions shall be reported to the veterinary staff in accordance with LAR SOPs 201 or 210.
3. **Feed/Water Animals:** See routine husbandry SOPs for each species as applicable.
4. **Clean/Sanitize Primary Enclosure:** See routine husbandry SOPs for each species as applicable.
5. **Clean/Sanitize Animal Room:**
   - **5.1.** For rooms housing small animals (e.g., mice and rats), floors are swept and mopped at least once a week using a veterinary approved detergent. This should be done additionally as needed.
   - **5.2.** For rooms housing large animals (e.g., dogs, pigs, cats, rabbits, NHPs), floors are swept or sprayed down on a daily basis. At least once a week the floors are sprayed down and/or mopped using a veterinary approved detergent. This should be done additionally as needed.
   - **5.3.** Sinks are cleaned at least weekly or as needed.
   - **5.4.** Paper towel and soap or hand sanitizer supply is checked daily and replenished as needed.
   - **5.5.** Pre-exhaust filters are checked weekly and changed as needed.
     - **5.5.1.** Microbial Sciences Vivarium BSL3 and Select Agent suites: pre-filters will be supplied for each suite on a regular basis by the LAR staff but changed and documented by the research staff.
   - **5.6.** Barrels (e.g., food, bedding, litter, trash) housed in animal rooms must be changed out and cleaned monthly.
   - **5.7.** Totes (e.g., food, water bottle, enrichment device) kept in animal rooms must be changed out and cleaned as they are emptied or monthly, whichever comes first.

[1] *Guide for the Care and Use of Laboratory Animals*, 8th edition.

5.8. Doors, including doorjambs and knobs, are cleaned inside and outside as needed or at least monthly.

5.9. Walls are cleaned as needed to keep them dust and spatter free.

5.10. Nothing is to be taped to the walls.

5.11. Ceilings, exposed pipes, and light fixtures must be kept clean and dust free.

5.12. Brooms and dustpans must be hung up when not in use.

5.13. Animal rooms shall be kept free of research equipment and supplies except for PPE and a minimal number of exchange cages. Exceptions to this policy may be made by getting prior approval from the facility supervisor in consultation with the veterinary staff.

5.14. Racks shall not be stored in animal rooms containing animals.

6. Animal Room Log Books:

6.1. Every animal room shall have a log book that contains:

6.1.1. A copy of LAR SOP No. 300 and the applicable species-specific 300 Series SOP

6.1.2. Monthly Room Activity Log(s)

6.1.3. Where applicable, a monthly Room Census Sheet for each investigator/requisition.

6.1.4. A Special Husbandry Log with any special instructions; these should be cleared through the area supervisor or lead, contain start and end dates, and be initialed by the Principal Investigator Staff.

6.2. All animal husbandry activities shall be recorded in the animal room activity log on a daily basis and initialed by the person making the entry.

6.3. The animal research technician assigned to the animal room is responsible for maintaining the animal room logbook.

6.4. Supervisors and/or lead technicians shall check each logbook at least once a month to insure that they are current and properly annotated.

6.5. Where applicable, changes in the number of animals and/or cages are recorded every Thursday or as they occur on the monthly room census sheet for the respective investigator/requisition.

6.6. The monthly room activity log and census sheets shall be turned into the supervisor's office on the first of every month and kept for 3 years.

Drew Jefcoat, DVM
Attending Veterinarian, SMPH

James Shull, PhD
Director, McArdle Laboratories

Appendix B: New Master Neuropeptide 692 Entry List

Family	Name	[M+H]	Sequence
Actin	Actin fragment Hoa cardiac muscle actin 95-105	1209.68376	LRVAPEESPVL
Actin	Actin fragment Limpoly actin3&actin11 95-106	1372.79470	LRVAPEEHPVLL
Allatostatin A-type	Carcinustatin 1, Penaeustatin 13	569.30821	YAFGLamide
Allatostatin A-type	Carcinustatin 14, Penaeustatin 35	585.30312	YSFGLamide
Allatostatin A-type	Penaeustatin 17	599.31880	YTFGLamide
Allatostatin A-type	Penaeustatin 30	613.29800	YDFGLamide
Allatostatin A-type		656.34020	SYAFGLamide
Allatostatin A-type	Procastatin 15	700.33010	SYDFGLamide
Allatostatin A-type		739.37740	GPYSFGLamide
Allatostatin A-type	Carcinustatin 6	753.39300	SPYAFGLamide
Allatostatin A-type		754.38825	GQYAFGLamide
Allatostatin A-type	Carcinustatin 2	769.38790	EAYAFGLamide
Allatostatin A-type		770.38317	GGAYSFGLamide
Allatostatin A-type	Penaeustatin 12	771.36720	SDYAFGLamide
Allatostatin A-type	Carcinustatin 5	780.40390	NPYAFGLamide
Allatostatin A-type	Carcinustatin 9	780.40390	GGPYAFGLamide
Allatostatin A-type	Penaeustatin 23	781.38790	GPYEFGLamide
Allatostatin A-type	Carcinustatin 4	781.38792	DPYAFGLamide
Allatostatin A-type		793.39910	YAHSFGLamide
Allatostatin A-type	Carcinustatin 8, Penaeustatin 6, Procastatin 17	794.41955	AGPYAFGLamide

Allatostatin A-type	Carcinustatin 3	795.40357	EPYAFGLamide
Allatostatin A-type		796.39880	NPYSFGLamide
Allatostatin A-type		798.42570	GRYSFGLamide
Allatostatin A-type	Procastatin 26	800.39370	SGAYSFGLamide
Allatostatin A-type	Penaeustatin 19	806.39440	GHYNFGLamide
Allatostatin A-type		808.43520	AAPYAFGLamide
Allatostatin A-type	Procastatin 18	810.41450	SGPYAFGLamide
Allatostatin A-type	Carcinustatin 15, Procastatin 27	810.41450	AGPYSFGLamide
Allatostatin A-type	Penaeustatin 24	811.39850	SPYEFGLamide
Allatostatin A-type	Carcinustatin 16	812.39370	GGPYSYGLamide
Allatostatin A-type	Penaeustatin 38	812.43010	SVAYGFGLamide
Allatostatin A-type	Penaeustatin 29	814.37300	NSYDFGLamide
Allatostatin A-type	Penaeustatin 16	814.40938	SQYTFGLamide
Allatostatin A-type		822.45085	VGPYAFGLamide
Allatostatin A-type	Procastatin 19	824.43010	TGPYAFGLamide
Allatostatin A-type	Carcinustatin 7	824.43012	ASPYAFGLamide
Allatostatin A-type	Penaeustatin 9	825.42540	AGQYAFGLamide
Allatostatin A-type	Procastatin 24	826.40938	SGPYSFGLamide
Allatostatin A-type	Penaeustatin 39	826.44580	TVAYGFGLamide
Allatostatin A-type		838.40938	GDPYAFGLamide
Allatostatin A-type	Penaeustatin 21	838.40940	GGPYEFGLamide
Allatostatin A-type	Penaeustatin 25	838.40940	NPYEFGLamide
Allatostatin A-type		838.44580	TAPYAFGLamide
Allatostatin A-type	Penaeustatin 33	840.42500	PSAYSFGLamide
Allatostatin A-type		841.42028	AGGAYSFGLamide
Allatostatin A-type	Procastatin 23	841.42030	AGQYSFGLamide
Allatostatin A-type		842.44070	AGLYSYGLamide
Allatostatin A-type	Penaeustatin 31	850.42060	AGHYSFGLamide
Allatostatin A-type	Penaeustatin 7	850.42060	SGHYAFGLamide

Allatostatin A-type		851.47740	GKPYAFGLamide
Allatostatin A-type	Penaeustatin 20	852.42500	AGPYEFGLamide
Allatostatin A-type	Procastatin 20	852.42500	ADPYAFGLamide
Allatostatin A-type		853.40900	EPYEFGLamide
Allatostatin A-type		854.40430	DGPYSFGLamide
Allatostatin A-type	Penaeustatin 10	854.44068	TPSYAFGLamide
Allatostatin A-type	Carcinustatin 17, Procastatin 10	857.41520	SGQYSFGLamide
Allatostatin A-type	Penaeustatin 22	866.44070	AAPYEFGLamide
Allatostatin A-type		870.39270	FNGCNFGLamide
Allatostatin A-type	Procastatin 8	870.41040	SGNYNFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 7	872.43350	KMYDFGLamide
Allatostatin A-type	Procastatin 21	877.45670	PNPYAFGLamide
Allatostatin A-type		879.48350	RGPYAFGLamide
Allatostatin A-type	Penaeustatin 4	881.45160	TGGPYAFGLamide
Allatostatin A-type	Penaeustatin 5	881.45160	SAGPYAFGLamide
Allatostatin A-type	Procastatin 2	881.45160	TPNYAFGLamide
Allatostatin A-type	Penaeustatin 8	882.44680	ANQYAFGLamide
Allatostatin A-type		883.43080	SNPYSFGLamide
Allatostatin A-type		884.43349	PSMYAFGLamide
Allatostatin A-type	Procastatin 7	884.45120	ADLYSFGLamide
Allatostatin A-type	Penaeustatin 15	885.44650	ASQYTFGLamide
Allatostatin A-type	Procastatin 22	888.39200	DGMYSFGLamide
Allatostatin A-type		892.46760	SGHYIFGLamide
Allatostatin A-type	Penaeustatin 18	893.42640	SGHYNFGLamide
Allatostatin A-type		893.49920	ARPYAFGLamide
Allatostatin A-type	Procastatin 16	895.46720	TAGPYAFGLamide

Allatostatin A-type		897.44650	FSGASPYGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 5	897.45770	ARGYDFGLamide
Allatostatin A-type		898.41275	PDMYGFGLamide
Allatostatin A-type		898.44174	GSGQYAFGLamide
Allatostatin A-type	Lepidopteran peptide Lepidostatin I	898.47810	AKSYNFGLamide
Allatostatin A-type	AST-3	899.46210	GGSLYSFGLamide
Allatostatin A-type	Procastatin 3	901.42360	QGMYSFGLamide
Allatostatin A-type	Procastatin 6	902.40770	ADMYSFGLamide
Allatostatin A-type	Procastatin 13	907.51480	PRVYGFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 4, helicostatin 4	909.4941	ARPYSFGLamide
Allatostatin A-type	Procastatin 5	910.46690	PDLYSFGLamide
Allatostatin A-type	Procastatin 1	911.43700	QNNYGFGLamide
Allatostatin A-type		911.46210	AASPYSFGLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 5	911.4734	ARAYDFGLamide
Allatostatin A-type	Carcinustatin 12	912.42840	PDMYAFGLamide
Allatostatin A-type	Penaeustatin 14	912.4574	ANQYTFGLamide
Allatostatin A-type		914.40770	PDM(O)YGFGLamide
Allatostatin A-type	Carcinustatin 18	918.4026	SDMYSFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 2, helicostatin 2 [11-18]	921.51926	LPVYNFGLamide

Allatostatin A-type		923.47338	pQRAYSFGLamide
Allatostatin A-type		923.53490	KLPYSFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 3,helicostatin 3	925.4890	SRPYSFGLamide
Allatostatin A-type		925.51420	LKAYDFGLamide
Allatostatin A-type		925.55060	LVKYSFGLamide
Allatostatin A-type	Carcinustatin 11	926.4730	ATGQYAFGLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 7	926.4843	ARSYNFGLamide
Allatostatin A-type	Procastatin 4	928.42330	PDMYSFGLamide
Allatostatin A-type		928.45231	SSGQYAFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 1	933.4577	SPHYNFGLamide
Allatostatin A-type		934.39750	SDM(O)YSFGLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 1	934.4417	SPHYDFGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 6	935.5349	LPLYNFGLamide
Allatostatin A-type	Procastatin 11	936.50500	PRNYAFGLamide
Allatostatin A-type		937.48903	PRDYAFGLamide
Allatostatin A-type		937.52540	PRVYSFGLamide
Allatostatin A-type		939.50470	TRPYSFGLamide
Allatostatin A-type		940.49993	QRAYSFGLamide
Allatostatin A-type	Penaeustatin 34	941.4186	pQNMYSFGLamide
Allatostatin A-type		949.48903	pQRPYSFGLamide

Allatostatin A-type	Penaeustatin 11	951.4683	pQRDYAFGLamide
Allatostatin A-type		953.48390	DRPYSFGLamide
Allatostatin A-type		953.48394	pQRTYSFGLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 6	953.49134	LPMYNFGLamide
Allatostatin A-type	Penaeustatin 3	953.5203	DRLYAFGLamide
Allatostatin A-type	Procastatin 9	956.49480	SRQYSFGLamide
Allatostatin A-type	Penaeustatin 32	957.4789	DRTYSFGLamide
Allatostatin A-type	Carcinustatin 10	962.50943	APQPYAFGLamide
Allatostatin A-type		967.46320	pQRDYSFGLamide
Allatostatin A-type		967.49959	ERPYSFGLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 8	974.5094	YSKFNFGLamide
Allatostatin A-type		984.51490	SKSPYSFGLamide
Allatostatin A-type		1004.48360	FSGTYNFGLamide
Allatostatin A-type		1021.51020	TPHTYSFGLamide
Allatostatin A-type		1023.51457	PADLYEFGLamide
Allatostatin A-type		1052.58870	SPRLTYFGLamide
Allatostatin A-type		1066.55677	PATDLYAFGLamide
Allatostatin A-type	Carcinustatin 19	1100.5081	APTDMYSFGLamide
Allatostatin A-type		1173.62000	HGTEGPYPFGLamide
Allatostatin A-type		1266.57490	HSPSSASYDFGLamide
Allatostatin A-type	Procastatin 25	1320.52260	EDYDSSDQYSLamide
Allatostatin A-type	Lepidopteran peptide helicostatin 9	1393.67940	ERDMHRFSFGLamide
Allatostatin A-type	Penaeustatin 28	1426.6597	PETERNSYDFGLamide
Allatostatin A-type	Penaeustatin 1	1568.7227	ANEDEDAASLFAFGLamide
Allatostatin A-type	Procastatin 14	1856.84830	ADSYGLAFGNGGDALEMGLamide

Allatostatin A-type	Penaeustatin 27	1866.8504	EVPDPETERNSYDFGLamide
Allatostatin A-type	Penaeustatin 2	1979.9934	PDAEESNKRDRLYAFGLamide
Allatostatin A-type	Penaeustatin 26	1980.8934	NEVPDPETERNSYDFGLamide
Allatostatin A-type	Procastatin 12	2052.78290	TSDEEDDEDDQYYPYGLamide
Allatostatin A-type	Lepidopteran peptide cydiastatin 2, helicostatin 2	2168.1175	AYSYVSEYKRLPVYNFGLamide
Allatostatin A-type		2173.12290	DPYAFGLRHTSFVLYAFGLamide
Allatostatin A-type	Carcinustatin 13	2224.0743	EYDDMYTEKRPKVYAFGLamide
Allatostatin A-type	Penaeustatin 41	2385.60000	
Allatostatin A-type	Penaeustatin 37	3099.5858	TARGALDLDQSPAYASDLGKRIGSAYSFGLamide
Allatostatin A-type	Penaeustatin 36	3113.5651	DARGALDLDQSPAYASDLGKRIGSAYSFGLamide
Allatostatin A-type	Carcinustatin 20	3150.5279	GYEDEDEDRPFYALGLGKRPRTYSFGLamide
Allatostatin A-type	Penaeustatin 40		(X)GIYGFGLamide
Allatostatin B-type	Lepidopteran peptide MIP IV	997.50030	AWSALHGAWamide
Allatostatin B-type		1031.50570	AWSNLGQAWamide
Allatostatin B-type	CbAST-B2	1107.51526	QWSSMRGAWamide
Allatostatin B-type		1107.51530	AGWSSMRGAWamide
Allatostatin B-type		1107.53300	AGWSSTSRAWamide
Allatostatin B-type		1123.51020	AGWSSM(O)RGAWamide
Allatostatin B-type	CbAST-B4	1165.55375	NWNKFQGSWamide
Allatostatin B-type		1179.56940	AGWNKFQGSWamide
Allatostatin B-type	CbAST-B5	1182.56907	TSWGKFQGSWamide
Allatostatin B-type		1194.60550	VTWGKFQGSWamide
Allatostatin B-type		1209.58000	TGWNKFQGSWamide
Allatostatin B-type		1220.58070	SGDWSSLRGAWamide
Allatostatin B-type	CbAST-B6	1222.57522	GNWNKFQGSWamide
Allatostatin B-type		1246.60760	GSNWSNLRGAWamide
Allatostatin B-type		1252.58580	NNWSGAFKGSWamide
Allatostatin B-type	CbAST-B7	1252.58580	NNWSKFQGSWamide
Allatostatin B-type		1253.56980	NDWSKFGQSWamide

Allatostatin B-type		1258.64400	GVNWSNLRGAWamide
Allatostatin B-type	CbAST-B3	1260.65962	SGKWSNLRGAWamide
Allatostatin B-type		1266.60140	TNWNKFQGSWamide
Allatostatin B-type		1266.63782	SKWNKFQGSWamide
Allatostatin B-type		1272.65960	LGNWSNLRGAWamide
Allatostatin B-type		1278.63780	LNWNKFQGSWamide
Allatostatin B-type	CbAST-B8	1293.63346	STNWSSLRSAWamide
Allatostatin B-type		1294.61750	STDWSSLRSAWamide
Allatostatin B-type		1335.65930	LGNWNKFQGSWamide
Allatostatin B-type	CbAST-B9	1366.62871	NNNWSKFQGSWamide
Allatostatin B-type		1371.63410	NPDWAHFRGSWamide
Allatostatin B-type		1380.64440	NNNWTKFQGSWamide
Allatostatin B-type		1381.66480	TQNWTKFQGSWamide
Allatostatin B-type	CbAST-B1	1470.70254	VPNDWAHFRGSWamide
Allatostatin B-type		1479.71280	LNNNWSKFQGSWamide
Allatostatin B-type		1586.84130	MFAPLAWPKGGARWamide
Allatostatin B-type		1699.92530	LMFAPLAWPKGGARWamide
Allatostatin Combos		1519.7904	DPYAFGLGKRPADL
Allatostatin Combos		1634.8537	EPYAFGLGKRPATDL
Allatostatin Combos		2002.9844	DPYAFGLGKRPDMYGFGLamide
Allatostatin Combos		2017.0000	DPYAFGLGKRPDMYAFGLamide
Allatostatin Combos		2035.0396	GSGQYAFGLGKKAGGAYSFGLamide
Allatostatin Combos		2128.0862	DPYAFGLGKRPADLYEFGLamide
Allatostatin Combos		2185.1441	EPYAFGLGKRPATDLYAFGLamide
Allatostatin C-type	Lepidopteran peptide Manse-AS	1888.86220	pEVRFRQCYFNPISCF
Allatostatin C-type		1650.71920	SYWKQCAFNAVSCFamide
Allatostatin C-type	PISCF-AST	1899.83050	pQIRYHQcYFNPIScF
Allatotropin	Lepidopteran peptide Manse-AT	1486.72935	GFKNVEMMTARGFamide

bursicon	bursicon beta	12618.98120	RSYGECETLPSTIHISKEEYDDTGRLVRVCEEDVAVNKCEGACVSKVQPSVNTP SGFLKDCRCCREVHLRARDITLTHCYDGDGARLSGAKATQHVKLREPADCQCF KCGDSTR
bursicon	bursicon alpha	13257.45060	DECSLRPVIHILSYPGCTSKPIPSFACQGRCTSYVQVSGSKLWQTERSMCCQESG EREAAITLNCPKPRPGEPKEKKVLTRAPIDCMCRPCTDVEEGTVLAQKIANFIQD SMPDSVPFLK
Calcitonin-like	diuretic hormone	3272.67460	GLDLGLGRGFSGSQAAKHLMGLAAANFANFAGGPamide
CCAP	CCAP	956.37532	PFCNAFTGCamide
CCAP	Lepidopteran peptide CAP2b	975.54110	pELYAFPRVamide
CCAP	CCAP precursor related peptide	1074.5313	DIGDLLEGKD
CHH	MetenCHHA	8255.8778	SLFDPSCSGVFDRELLGRLNRVCDDCYNVFRDPKVAMECKSNCFLNPAFIQCLE YLLPEDLHEEYQSHQVVamide
CHH	Orl-CHH	8402.1102	pEVFDQACKGIYDRAIFKKLDRVCEDCYNLYRKPYVATTCRQNCYANSVFRQC LDDLLLIDVLDEYISGVQTVamide
CHH	Capa-CHH II	8419.8194	pEIYDTSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQC MEELLLMDEFDKYARAVQIV
CHH	CanpaCHH-II	8428.9102	pEIYDTSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQV MEELLLMEEFDKYARAVQIVamide
CHH	CanproCHH-II	8431.9211	QIYDSSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQV MEELLLMEEFDKYARAVQIVamide
CHH	Capr-CHH II	8436.8459	QIYDSSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQC MEELLLMEEFDKYARAVQIV
CHH	Capa-CHH I	8436.8459	QIYDTSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQC MEELLLMDEFDKYARAVQIV
CHH	Canpa/proCHH-I	8445.9368	QIYDTSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQV MEELLLMEEFDKYARAVQIVamide
CHH	Capr-CHH I	8450.8616	QIYDTSCKGVYDRGLFSDLEHVCDDCYNLYRNSYVASACRSNCYSNVVFRQC MEELLLMEEFDKYARAVQIV
CHH	DappuITPL	8471.1357	SFFDIQCKGNYDKSIFARLDRICEDCYNLFREPQLHSLCRSDFKSPYFKGCLQAL LLIDEEEKFNQMVEILamide
CHH	Hoa-MIH (pMIH[pQ61-M131])	8479.0099	pEVFDQACKGVYDRNLFKKLDRVCEDCYNLYRKPFVATTCRENCYSNWVFRQ CLDDLLLSNVIDEYVSNVQM
CHH	CarmaCHH	8525.7853	pQIYDTSCKGVYDRALFNDLEHVCDDCYNLYRTSYVASACRSNCYSNLVFRQC MDDLLMMDEFDQYARKVQMVamide

CHH	Hoa-CHH-A (pCHH-A[pQ61-V132amide])	8578.0783	pEVFDQACKGVYDRNLFKKLDRVCEDCYNLYRKPFVATTCRENCYSNWVFRQCLDDLLLSDVIDEYVSNVQMVamide
CHH	CarmaPOCHH	8631.8672	pQIYDTSCKGVYDRALFNDLEHVCDDCYNLYRTSYVASACRNNCFENEVFDVCVYQLYFPNHEEYLRSRDGLKG
CHH	Hoa-CHH-B (pCHH-B[pQ61-V132amide])	8633.1715	pEVFDQACKGVYDRNLFKKLNRVCEDCYNLYRKPFIVTTCRENCYSNRVFRQCLDDLLMIDVIDEYVSNVQMVamide
CHH	Prc-MIH	8646.13836	RYVFEECPGVMGNRAVHGKVTRVCEDCYNVFRDTDVLAGCRKGCFSSEMFKLCLLAMERVEEFPDFKRWIGLLNAamide
CHH	CanboCHH	8660.0144	pQIYDTSCKGVYDRALFSDLEHVCDDCYNLYRSSYVASECRRNCYSNVVFRQCMEELLLLMEEFDKYARAVQIVamide
CHH	Mee-MIH	8981.37306	SYIENTCRGVMGNRDIYKKVVRVCEDCTNIFRLPGLDGMCRDRCFNNEWFLVCLKAANRDDELDKFKVWISILNPGL
CHH	Pej-MIH	9016.41146	SFIDNRCRGVMGNRDIYKKVVRVCEDCTNIFRLGLDGMCRNRCFYNEWFLICLKAANREDEIEKFRVWISILNAGQ
CHH	HomamVIH	9136.2264	ASAWFTNDECPGVMGNRDLYEKVAWVCNDCANIFRNNDVGVMCKKDCFHTMDFLWCVYATERHGEIDQFRKWVSILRAamide
CHH	HoaVIH	9137.2257	ASAWFTNDECPGVMGNRDLYEKVAWVCNDCANIFRNNDVGVMCKKDCFHTMWFLWCVYATERHGEIDQFRKWVSILR
CHH	CanpaMIH	9154.3147	RVINDECPNLIGNRDLYKKVEWICEDCSNIFRKTGMASLCRRNCFFNEDFVWCVHATERSEELEDLEEWVGILGAGRD
CHH	CapaMIH	9195.3599	RVINDDCPNLIGNRDLYKKVEWICEDCSNIFRNTGMATLCRKNCFFNEDFLWCVYATERTEEMSQLRQWVGILGAGRE
CHH	CarmaMIH	9223.3661	RVINDDCPNLIGNRDLYKRVEWICEDCSNIFRNTGMATLCRKNCFFNEDFLWCVYATERTEEMSQLRQWVGILGAGRE
CHH	CanpaMOIH2	9236.2964	RRINNDCQNFIGNRAMYEKVDWICKDCANIFRQDGLLNNCRSNCFYNTEFLWCIDATENTRNKEQLEQWAAILGAGWN
CHH	CanpaMOIH1	9236.3328	RRINNDCQNFIGNRAMYEKVDWICKDCANIFRKDGLLNNCRSNCFYNTEFLWCIDATENTRNKEQLEQWAAILGAGWN
CHH	DappuITPL	9491.7107	SFFEDINCKGLYDKSIFARLDRICQDCYSLYREPELHTLCRKNCFTTNYFKGCLDALLINDEKDIQRVMKDISIIHQIPI
Corazonin	Corazonin	1369.6284	pQTFQYSRGWTNamide
CPRP	Truncated Capr CPRP I 30-38	904.48870	PGGLVHPVE
CPRP	truncated Hoa CPRP-A (SN-16)	932.44720	FLSQDHSV

	26-33		
CPRP		975.51530	RSAEGLGRMamide
CPRP		989.46870	GFLSQDVHS
CPRP		1033.52070	RSAEGLGRMG
CPRP	truncated Hoa CPRP-A (SN-16) 26-34	1046.49010	FLSQDHSVN
CPRP	truncated Hoa CPRP-A (SN-16) 23-32, B&C 22-31	1100.53710	PLGFLSQDHS
CPRP	truncated Hoa CPRP-A (SN-16) 13-25	1103.51160	GFLSQDHSVN
CPRP	truncated Hoa CPRP-B&C 1-10	1149.56810	RSVEGVSRME
CPRP	truncated Hoa CPRP-A (SN-16) 23-33, B&C 22-32	1199.60550	PLGFLSQDHSV
CPRP		1200.53920	RGFEGETGHPN
CPRP	truncated Hoa CPRP-A (SN-16) 24-34, B&C 23-33	1216.59570	LGFLSQDHSVN
CPRP	truncated Casap CPRP 15-25	1216.68960	SLKSDTVTPLR
CPRP	truncated Casap CPRP 18-28	1221.61100	SDTVTPLRGFE
CPRP		1230.69400	SLKSDTVTPLLG
CPRP	truncated Cb CPRP I,II&IV 1-12	1232.65280	RSAQGLGKMER
CPRP	truncated Hoa CPRP-A (SN-16) 13-25	1233.59570	LSSSNSPSSTPLG

CPRP	truncated Hoa CPRP-B 2-12	1234.64600	SVEGVSRMEKL
CPRP		1244.68450	DLKSDTVTPLR
CPRP		1269.75250	ASLKSPTVTPLR
CPRP	truncated Casap CPRP 14-25	1287.72670	ASLKSDTVTPLR
CPRP	truncated Hoa CPRP-A21-32, B20-31	1288.61680	STPLGFLSQDHS
CPRP	truncated Hoa CPRP-A (SN-16) 12-24	1289.65830	LLSSSNSPSSTPL
CPRP	truncated Hoa CPRP-A (SN-16) 22-33, B 21-32	1300.65320	TPLGFLSQDHSV
CPRP		1302.70590	RSAEGLGRMGRL
CPRP	truncated Hoa CPRP-A (SN-16) 11-23	1304.66920	KLLSSSNSPSSTP
CPRP	truncated Hoa CPRP-A (SN-16) 23-34, B&C 22-33	1313.64840	PLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 11-23	1329.76240	KLLSSISPSSTPL
CPRP	truncated Capr CPRP I 1-12	1344.6511	RSAQGMGKMEHL
CPRP		1361.73180	RSAQGLGKM(O)ERL
CPRP	Truncated Hoa CPRP-A (SN-16, I-16) 1-12	1362.71580	RSVEGASRMEKL
CPRP	truncated Capa CPRP 1-12, Capr CPRP II 1-12	1363.69330	RSAQGMGKMERL
CPRP	truncated Hoa CPRP-A20-32,	1375.64880	SSTPLGFLSQDHS

	B19-31		
CPRP	truncated Cb CPRP II,III&IV15-38	1377.70090	PLGDLSGSLGHPVE
CPRP		1383.81790	RSAEGLGRVGRLL
CPRP	truncated Hoa CPRP-B&C 1-12	1390.74710	RSVEGVSRMEKL
CPRP	truncated Casap CPRP 13-25	1400.81080	LASLKSDTVTPLR
CPRP	truncated Hoa CPRP-A (SN-16) 22-34, B 21-33	1414.69610	TPLGFLSQDHSVN
CPRP		1415.79000	RSAEGLGRMGRLL
CPRP	truncated Hoa CPRP-C 11-23	1416.77670	KLLSSISPSSMPLG
CPRP	truncated Hoa CPRP-A (SN-16) 11-24	1417.75330	KLLSSSNSPSSTPL
CPRP	truncated Capr CPRP I 1-13	1457.7352	RSAQGMGKMEHLL
CPRP	truncated Cb CPRP I,II&IV 1-13	1458.82090	RSAQGLGKMERLL
CPRP	truncated Hoa CPRP-A (SN-16) 11-25	1474.77480	KLLSSSNSPSSTPLG
CPRP	Truncated Hoa CPRP-A (SN-16, I-16) 1-13	1475.79990	RSVEGASRMEKLL
CPRP	truncated Capa CPRP 24-38, CbCPRP II,III&IV 24-38	1478.74850	TPLGDLSGSLGHPVE
CPRP	truncated Cama CPRP I 1-13	1493.75290	RSTQGYGRMDPIL
CPRP	truncated Hoa	1501.72810	STPLGFLSQDHSVN

	CPRP-A21-34, B20-33		
CPRP	truncated (PRO-4) 1-13	1521.79550	RSTPGYGRMDRIL
CPRP	truncated Capr CPRP I 1-14	1528.77230	RSAQGMGKMEHLLA
CPRP	truncated Cb CPRP I&II 1-14	1529.85810	RSAQGLGKMERLLA
CPRP	truncated Hoa CPRP-C 20-33	1531.72100	SMPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 11-25	1533.85230	KLLSSISPSSTPLGF
CPRP	Truncated Hoa CPRP-A (SN-16, I-16) 1-14	1562.83190	RSVEGASRMEKLLS
CPRP		1576.84760	RSVEGASRMEKLLT
CPRP	truncated Hoa CPRP-A20-34, B19-33	1588.76020	SSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B&C 1-14	1590.86320	RSVEGVSRMEKLLS
CPRP	truncated Cb CPRP II,III&IV 23-38	1592.79150	NTPLGDLSGSLGHPVE
CPRP		1604.87890	RSVEGVSRMEKLLT
CPRP	truncated Cb CPRP I&II 1-15	1616.89010	RSAQGLGKMERLLAS
CPRP	truncated Hoa CPRP-A (SN-16) 11-26	1621.84320	KLLSSSNSPSSTPLGF
CPRP		1629.82310	PSAALAVEHGTTHPLE
CPRP		1632.88500	RSAQGLGKM(O)ERLLAS
CPRP	truncated Casap CPRP 22-36	1639.80750	TPLRGFEGETGHPLE
CPRP	truncated Hoa CPRP-B 11-26	1646.93630	KLLSSISPSSTPLGFL

CPRP	Truncated Hoa CPRP-A (SN-16, I-16) 1-15	1649.86390	RSVEGASRMEKLLSS
CPRP	truncated (PRO-4) 1-15	1663.86970	RSTPGYGRMDRILAA
CPRP	truncated Hoa CPRP-A (SN-16) 17-32	1673.77660	NSPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-A (SN-16) 19-34, B 18-33	1685.81290	PSSTPLGFLSQDHSVN
CPRP	truncated Cama CPRP I 1-15	1694.87550	RSTQGYGRMDRILAA
CPRP	truncated Hoa CPRP-A (SN-16) 11-27	1734.92720	KLLSSSNSPSSTPLGFL
CPRP	truncated Hoa CPRP-B 13-31	1759.84970	SISPSSTPLGFLSQDHS
CPRP		1760.91120	RSAQGLGKMEHLLASY
CPRP	truncated Cb CPRP I&II 1-16	1779.95340	RSAQGLGKMERLLASY
CPRP		1796.01770	RSAQGLGKYLRLLASY
CPRP	truncated Hoa CPRP-C 16-32	1801.87890	ISPSSMPLGFLSQDHSV
CPRP	truncated Cb CPRP IV 1-16	1807.98470	RSAQGLGKMERLLVSY
CPRP	truncated Hoa CPRP-B 14-31	1846.88170	SSISPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-A (SN-16) 15-32	1847.84060	SSNSPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-B 13-32	1858.98100	SISPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-A (SN-16) 16-33	1859.87700	SNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa	1862.02700	KLLSSISPSSTPLGFLSQ

	CPRP-B 11-28		
CPRP	truncated Hoa CPRP-A (SN-16) 17-34	1886.88970	NSPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-C 16-33	1915.92180	ISPSSMPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 14-32	1945.95020	SSISPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-A (SN-16) 15-33	1946.90900	SSNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-B 13-33	1972.96110	SISPSSTPLGFLSQDHSVN
CPRP	truncated Cb CPRP II&III 19-38	2003.00800	ALEPNTPLGDLSGSLGHPVE
CPRP	truncated Hoa CPRP-A (SN-16) 14-33	2033.94110	SSSNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-A (SN-16) 13-32	2047.95670	LSSSNSPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-B 13-32	2059.03420	LSSISPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-B 14-33	2059.99310	SSISPSSTPLGFLSQDHSVN
CPRP	truncated Cb CPRP II 18-38	2060.02950	GALEPNTPLGDLSGSLGHPVE
CPRP	truncated Hoa CPRP-A (SN-16) 15-34	2060.95200	SSNSPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B&C 1-19	2062.09610	RSVEGVSRMEKLLSSISPS
CPRP	truncated (PRO-4) 19-38	2073.99100	SPMEPSAALAVEHGTTHPLE
CPRP	Truncated Hoa CPRP-A (SN-16)	2122.05570	RSVEGASRMEKLLSSSNSPS

	1-20		
CPRP	truncated Hoa CPRP-A (SN-16) 13-33	2147.02510	LSSSNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-A (SN-16) 14-34	2147.98400	SSSNSPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 12-32	2172.11830	LLSSISPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-B 13-33	2173.07720	LSSISPSSTPLGFLSQDHSVN
CPRP	truncated (PRO-4) 18-38	2175.03870	TSPMEPSAALAVEHGTTHPLE
CPRP	truncated Cb CPRP II&III 16-38	2216.13060	RGALEPNTPLGDLSGSLGHPVE
CPRP	truncated Hoa CPRP-A (SN-16) 12-33	2260.10920	LLSSSNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-A (SN-16) 13-34	2261.06800	LSSSNSPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 12-33	2286.16120	LLSSISPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-A (SN-16) 11-32	2289.13570	KLLSSSNSPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-B 11-32	2300.21330	KLLSSISPSSTPLGFLSQDHSV
CPRP		2354.27210	RASQGLGKMERLLASRGALEPN
CPRP	truncated Hoa CPRP-A (SN-16) 12-34	2374.15210	LLSSSNSPSSTPLGFLSQDHSVN
CPRP	truncated Cb CPRP II&III 17-38	2379.19390	YRGALEPNTPLGDLSGSLGHPVE

CPRP	truncated Hoa CPRP-A (SN-16) 11-33	2388.20410	KLLSSSNSPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-B 11-33	2414.25620	KLLSSISPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-C 11-33	2444.17000	KLLSSISPSSMPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-C 1-23	2490.24900	RSVEGVSRMEKLLSSISPSSMPL
CPRP	truncated Hoa CPRP-A (SN-16) 11-34	2502.24710	KLLSSSNSPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-B 1-24	2517.33410	RSVEGVSRMEKLLSSISPSSTPLG
CPRP	truncated Hoa CPRP-B 9-33	2543.29880	EKLLSSISPSSTPLGFLSQDHSVN
CPRP	truncated Hoa CPRP-C 1-24	2547.32690	RSVEGVSRMEKLLSSISPSSMPLG
CPRP	Truncated Hoa CPRP-A (SN-16) 1-25	2577.29370	RSVEGASRMEKLLSSSNSPSSTPLG
CPRP	truncated Hoa CPRP-A (SN-16) 9-33	2648.28720	MEKLLSSSNSPSSTPLGFLSQDHSV
CPRP	truncated Cb CPRP II&III 13-38	2650.34710	LASYRGALEPNTPLGDLSGSLGHPVE
CPRP	Truncated Capr CPRP II [13-38]	2658.38860	LASYRGAVEPNTPLGDLPGGLVHPVE
CPRP	Truncated Hoa CPRP-A (SN-16) 1-26	2724.36210	RSVEGASRMEKLLSSSNSPSSTPLGF
CPRP	truncated Hoa CPRP-A (SN-16) 9-34	2762.33020	MEKLLSSSNSPSSTPLGFLSQDHSVN
CPRP	truncated Cb CPRP II&III 12-	2763.43120	LLASYRGALEPNTPLGDLSGSLGHPVE

	38		
CPRP	truncated Hoa CPRP-B 1-26	2777.48660	RSVEGVSRMEKLLSSISPSSTPLGFL
CPRP	Truncated Hoa CPRP-A (SN-16) 1-27	2837.44620	RSVEGASRMEKLLSSSNSPSSTPLGFL
CPRP	truncated Hoa CPRP-B 1-28	2992.57720	RSVEGVSRMEKLLSSISPSSTPLGFLSQ
CPRP	Truncated Hoa CPRP-A (SN-16) 1-29	3052.53680	RSVEGASRMEKLLSSSNSPSSTPLGFLSQ
CPRP	Mros CPRP	3286.73900	WSVDGLARIEKLLSTSSSASAASPTRGQALNL
CPRP	Mlan CPRP	3300.75460	WSLDGLARIEKLLSTSSSASAASPTRGQALNL
CPRP	truncated Hoa CPRP-B 2-33	3388.70530	SVEGVSRMEKLLSSISPSSTPLGFLSQDHSVN
CPRP	Truncated Hoa CPRP-A (SN-16) 1-32	3391.65470	RSVEGASRMEKLLSSSNSPSSTPLGFLSQDHS
CPRP	truncated Hoa CPRP-B 1-32	3430.76350	RSVEGVSRMEKLLSSISPSSTPLGFLSQDHSV
CPRP	truncated Hoa CPRP-C 1-32	3460.75630	RSVEGVSRMEKLLSSISPSSMPLGFLSQDHSV
CPRP	Truncated Hoa CPRP-A (SN-16) 1-33	3490.72310	RSVEGASRMEKLLSSSNSPSSTPLGFLSQDHSV
CPRP	Olim CPRP A*	3516.68710	RSVEGSSRMERLLSSGSSSSEPLSFLSQDQSVS
CPRP	Hoa CPRP A (I-16)	3516.77510	RSVEGASRMEKLLSSISPSSTPLGFLSQDHSVN
CPRP	Olim CPRP A	3525.68740	RSVEGSSRMERLLSSGSSSSEPLSFLSQDHSVS
CPRP	Nnor CPRP B	3530.79080	RSVEGASRMEKLLSSISPSSTPLGFLSQEHSVN
CPRP	Olim CPRP B, Pcla CPRP	3543.69800	RSVEGSSRMERLLSSGSSSSEPLSFLSQDQSVN
CPRP	Hoa CPRP-B	3544.80640	RSVEGVSRMEKLLSSISPSSTPLGFLSQDHSVN
CPRP	Hoa CPRP C	3574.79920	RSVEGVSRMEKLLSSISPSSMPLGFLSQDHSVN
CPRP	Hoa CPRP-A	3604.76600	RSVEGASRMEKLLSSSNSPSSTPLGFLSQDHSVN

	(SN-16)		
CPRP	Nnor CPRP A	3618.78170	RSVEGASRMEKLLSSSNSPSSTPLGFLSQEHSVN
CPRP	truncated Cb CPRP II 3-39	3733.9330	AQGLGKMERLLASYRGALEPNTPLGDLSGSLGHPVE
CPRP	truncated Cb CPRP II 2-38	3820.96500	SAQGLGKMERLLASYRGALEPNTPLGDLSGSLGHPVE
CPRP	Sserr CPRP	3824.95010	RSAEGFGRMGRLLASLKADSLGPVQDFGVEGAAHPVE
CPRP	Casap CPRP	3838.00280	RSAEGLGRMGRLLASLKSDTVTPLRGFEGETGHPLE
CPRP	Cb CPRP III	3958.02400	RSAQGLGKMEHLLASYRGALEPNTPLGDLSGSLGHPVE
CPRP	Cb CPRP I	3963.05050	RSAQGLGKMERLLASYRGALEPNTPLGDLSGSVGHPVE
CPRP	Capa CPRP	3968.01170	RSAQGMGKMERLLASYRGALEPSTPLGDLSGSLGHPVE
CPRP	Capr CPRP I	3975.98040	RSAQGMGKMEHLLASYRGALESNTPTGDLPGGLVHPVE
CPRP	Cb CPRP II	3977.06620	RSAQGLGKMERLLASYRGALEPNTPLGDLSGSLGHPVE
CPRP	Capr CPRP III	3988.01680	RSAQGMGKMEHLLASYRGALESNTPLGDLPGGLVHPVE
CPRP	Cb CPRP IV	3991.08180	RSAQGLGKMERLLVSYRGAVEPNTPLGDLSGSLGHPVE
CPRP	Capr CPRP II	4003.06410	RSAQGMGKMERLLASYRGAVEPNTPLGDLPGGLVHPVE
CPRP	Capr CPRP IV	4017.07970	RSAQGMGKMERLLASYRAAVEPNTPLGDLPGGLVHPVE
CPRP	(PRO-4)	4061.06950	RSTPGYGRMDRILAALKTSPMEPSAALAVEHGTTHPLE
CPRP	Cama CPRP I	4069.05940	RSTQGYGRMDRILAALKTSPMEPSAALAVENGTTHPLE
CPRP	Cama CPRP II	4091.09130	RSTQGYGRMDRILAALKTSPMEPSAALAVQHGTTHPLE
CPRP	(GLN-4)	4092.07530	RSTQGYGRMDRILAALKTSPMEPSAALAVEHGTTHPLE
CPRP	Bthe CPRP	4150.03050	RSAEGFGRMERLLASIRGGADSMGHLGELTGAGEGAGHPLE
Cryptocyanin	cryptocyanin fragment CC1 21-27 & CC2 40-46	902.54580	KIFEPLR
Cryptocyanin	cryptocyanin fragment CC1 20-26	909.50803	YKIFEPL
Cryptocyanin	cryptocyanin related fragment	916.55020	KIFEPLVA
Cryptocyanin	cryptocyanin fragment CC1	1065.60914	YKIFEPLR

	20-27		
Cryptocyanin	cryptocyanin fragment CC1 20-28	1194.65170	YKIFEPLRE
Cryptocyanin	cryptocyanin related fragment	1259.71060	KIFEPLRDKN
Cryptocyanin	cryptocyanin fragment CC1 20-29	1281.68376	YKIFEPLRES
Cryptocyanin	cryptocyanin related fragment	1372.79470	KIFEPLRDKNL
Cryptocyanin	cryptocyanin fragment CC2 40-50	1373.74234	KIFEPLREDNL
Cryptocyanin	cryptocyanin fragment CC1 20-30	1395.72669	YKIFEPLRESN
Cryptocyanin	cryptocyanin fragment CC1 20-31	1508.81076	YKIFEPLRESNL
Ecdysis triggering hormone	ETH	2584.35310	DPSPEPFNPNYNRFRQKIPRIamide
Eclosion hormone		6932.38870	AVAANRKVSICIKNCGQCKKMYTDYFNGGLCGDFCLQTEGRFIPDCNRPDILIP FFLQRLE
Kinins	pev-Kinin 6	677.34060	AFSPWAamide
Kinins	pev-Kinin 2	695.31475	DFSAWAamide
Kinins	pev-Kinin 1	750.35700	ASFSPWGamide
Kinins	Lepidopteran peptide helicokinin I	755.35110	YFSPWGamide
Kinins	pev-Kinin 3	760.37770	PAFSPWGamide
Kinins	pev-Kinin 4	762.39334	VAFSPWGamide
Kinins	pev-Kinin 5	788.37260	pEAFSPWAamide
Kinins	Lepidopteran peptide helicokinin II	847.45730	VRFSPWGamide

Kinins	Lepidopteran peptide helicokinin III	921.53050	KVKFSAWGamide
Kinins	cholecyctokinin cck8	1063.40120	DYMGWMDFamide
myosuppressin		1125.56870	pQDLDHVFLR
Neuroparsin		7708.21040	APRCDRHDEEAPKNCKYGTTQDWCKNGVCAKGPGETCGGYRWSEGKCGEGT PCSCGICGGCSPFDGKCGPTSIC
Orcokinin		1027.48430	EIDRSGFGF
Orcokinin		1066.48000	NFDEIDRSA
Orcokinin		1098.52145	EIDRSGFGFA
Orcokinin		1142.51130	FDEIDRSGFG
Orcokinin		1156.52690	FDEIDRSGFA
Orcokinin		1171.53783	EIDRSSFGFN
Orcokinin		1186.53750	FDEIDRSSFA
Orcokinin		1198.54873	NFDEIDRSGFamide
Orcokinin		1199.53274	NFDEIDRSGF
Orcokinin	Hoa-Orcokinin	1213.50414	SSEDMDRLGFG
Orcokinin		1213.54839	DEIDRSGFGFA
Orcokinin		1227.53100	SSEDMDRLGFA
Orcokinin		1228.55930	NFDEIDRSSFamide
Orcokinin	Orcokinin[1-11]	1256.55421	NFDEIDRSGFG
Orcokinin		1257.53820	DFDEIDRSGFG
Orcokinin		1270.56986	NFDEIDRSGFA
Orcokinin		1271.55390	DFDEIDRSGFA
Orcokinin		1280.66336	VYGPRDIANLY
Orcokinin		1286.56477	NFDEIDRSSFG
Orcokinin		1287.54880	DFDEIDRSSFG
Orcokinin	Orcokinin-like peptides	1300.58042	NFDEIDRSSFA
Orcokinin		1301.56440	DFDEIDRSSFA
Orcokinin		1360.61680	FDEIDRSGFGFA
Orcokinin	Orcokinin[1-12]	1403.62262	NFDEIDRSGFGF

Orcokinin		1433.63320	NFDEIDRSSFGF
Orcokinin	[Ala8-Ala13]-Orcokinin	1458.66482	NFDEIDRAGFGFA
Orcokinin	Hoa-Orcokinin	1474.61549	SSEDMPSSLGFGFN
Orcokinin	[Ala13]-orcokinin	1474.65973	NFDEIDRSGFGFA
Orcokinin		1475.64370	DFDEIDRSGFGFA
Orcokinin		1490.61040	SSEDM(o)PSSLGFGFN
Orcokinin	[val13]-OrcoKinin	1502.69103	NFDEIDRSGFGFV
Orcokinin		1503.67505	DFDEIDRSGFGFV
Orcokinin		1504.67030	NFDEIDRSSFGFA
Orcokinin		1505.65430	DFDEIDRSSFGFA
Orcokinin	[Asn13]-OrcoKinin	1517.66555	NFDEIDRSGFGFN
Orcokinin		1517.69070	DFEDIERSGFGFV
Orcokinin	[Ser9-val13]-Orcokinin	1532.70160	NFDEIDRSSFGFV
Orcokinin		1533.68560	DFDEIDRSSFGFV
Orcokinin		1540.68153	NFDEIDRSGFGFH
Orcokinin	[Ser9]-Orcokinin	1547.67611	NFDEIDRSSFGFN
Orcokinin		1548.66010	DFDEIDRSSFGFN
Orcokinin	[Thr8-His13]-Orcokinin	1554.69718	NFDEIDRTGFGFH
Orcokinin/ Orcomyotropin-related		1061.57380	TPRDIANLYamide
Orcomyotropin		904.41994	FDAFTTGFamide
Orcomyotropin		1168.54220	FPAFTTGFGSH
Orcomyotropin		1186.51636	FDAFTTGFGHS
Orcomyotropin		1213.52726	FDAFTTGFGHN
Others	[Leu]-enkephalin	556.27658	YGGFL
Others	[Met]-enkephalin	574.23300	YGGFM
Others		844.47880	HL/IGSL/IYRamide
Others		858.54070	KPKTEKK

Others		1099.63580	DLPKVDTALK
Others		1103.55920	GPSGGFNGALAR
Others		1254.81440	AVLLPKKTEKK
Others		1363.67400	EVEEPEAPAPPAK
PDH		1190.58000	pQELHVPEREA
PDH	Pandalus borealis PDH	1901.99360	NSGMINSILGIPRVMTEAamide
PDH	Penaeus aztecus PDH	1927.03170	NSELINSLLGIPKVMNDAamide
PDH	Uca pugilator, Cancer magister PDH	1927.03172	NSELINSILGLPKVMNDAamide
PDH	Procambarus clarkii PDH	1941.04740	NSELINSILGLPKVMNEAamide
PDH		1973.04844	NSELINSLLGISRLMNEAamide
PDH	PDH-Cancer-Iia	5685.07310	QEVNVSEREAVATLAAHILKVVCAPLEGAGGLPHKRNSELINSILGLPKVMNEA amide
PDH	PDH-Callinectes-II	5712.10310	QELHVPEREAVANLAARILKIVHAPHDAAGVPHKRNSELINSLLGISALMNEAa mide
PDH	PDH-Marsupenaeus-II	5792.12150	QREPTASKCQAATELAIQILQAVKGAHTGVAAGPHKRNSELINSLLGLPKFMID Aamide
PDH	PDH-Cancer-I	5829.08030	QDLKYQAREMVAELAQQIYRVAQAPQAGAVGPHKRNSELINSILGLPKVMND Aamide
PDH	PDH-Marsupenaeus-I	5851.13270	DSSLKYFEREVVSELAAQILRVAQGPSAFVAGPHKRNSELINSLLGIPKVMTDAa mide
PDH	PDH-orconectes	5853.13850	QELKYPEREVVAELAAQIYRVAQAPWAGAVGPHKRNSELINSILGLPKVMNEA amide
PDH	PDH-Carcinus	5945.10650	QDLKYQEREMVAELAQQIYRVAQAPWAGAVGPHKRNSELINSILGLPKVMND Aamide
PDH	PDH-Marsupenaeus-III	5947.16510	QEDLKYFEREVVSELAAQILRVAQGPSAFVAGPHKRNSELINSLLGIPKVMNDA amide
PDH	PDH-Callinectes-I	5973.13780	QELKYQEREMVAELAQQIYRVAQAPWAAAVGPHKRNSELINSILGLPKVMND Aamide
PDH	PDH-armadillidium	6199.21790	QDLNPTEKEVLSNMLDFLQRHSRTTYMFPLLSESKRNSELINSLLGAPRVLNNA amide
PDH	PDH-eurydice	6250.38160	QSRDFSISEREIVASLAKQLLRVARMGYVPEGDLPRKRNAELINSLLGVPRVMS

			DAamide
Proctolin	Proctolin	649.36679	RYLPT
Pyrokinin		618.37221	FSPRLamide
Pyrokinin		878.52470	LYFAPRLamide
Pyrokinin		1024.55740	TSFAFSPRLamide
Pyrokinin		1037.55269	SGGFAFSPRLamide
Pyrokinin	Cb Pyrokinin I	1051.56834	TNFAFSPRLamide
Pyrokinin		1109.57380	TDGFAFSPRLamide
RFamide		695.39876	NFLRFamide
RFamide		710.39840	EFLRFamide
RFamide		735.43010	GPFLRFamide
RFamide	pem-PYF1	801.49545	RARPRFamide
RFamide		817.48320	HVFLRFamide
RFamide		817.51552	LRNLRFamide
RFamide		820.44240	ASNNLRFamide
RFamide		824.48897	RSFLRFamide
RFamide		825.44400	MPYLRFamide
RFamide	Lepidopteran peptide FLRFamide II	839.45220	GNSFLRFamide
RFamide		851.49987	RNFLRFamide
RFamide	Lepidopteran peptide FLRFamide III	853.40900	DPSFLEFamide
RFamide		863.45220	pQGNFLRFamide
RFamide		865.51550	RQFLRFamide
RFamide		887.55738	PSLRLRFamide
RFamide		901.57300	LTRPLRFamide
RFamide	S	905.49920	I/LNFTHKFamide
RFamide		905.51380	PSMRLRFamide
RFamide		905.53558	LNPFLRFamide
RFamide		908.52133	GRNFLRFamide
RFamide		916.51520	THPFLRFamide

RFamide		921.50870	PSM(O)RLRFamide
RFamide		925.48900	ETNFLRFamide
RFamide		926.52066	SKNYLRFamide
RFamide		937.52540	ELNFLRFamide
RFamide		938.53190	NRSFLRFamide
RFamide	PrcFaRP 3	938.53190	SRNFLRFamide
RFamide	Mar-FLP 4, sNPF (short neuropeptide F)	942.59958	APALRLRFamide
RFamide		948.54140	PGVNFLRFamide
RFamide		952.54755	TRNFLRFamide
RFamide		954.52680	SRNYLRFamide
RFamide		958.47810	FDDFLRFamide
RFamide		962.53190	pQRNFLRFamide
RFamide	NF 1, PrcFaRP 2	965.54279	NRNFLRFamide
RFamide	DF 2, Mar-FLP 1, PrcFaRP 5	966.52681	DRNFLRFamide
RFamide		977.56790	GPKNFLRFamide
RFamide		978.59960	LRQFLRFamide
RFamide		993.52650	NPSNFLRFamide
RFamide		994.51050	NPSDFLRFamide
RFamide		995.55336	SGRNFLRFamide
RFamide		1005.57409	GPRNFLRFamide
RFamide		1007.57850	PKSNFLRFamide
RFamide	Mar-FLP 2	1009.5578	ADKNFLRFamide
RFamide		1019.58974	APRNFLRFamide
RFamide		1022.56462	GNRNFLRFamide
RFamide	pem-FLP 1	1023.54827	GDRNFLRFamide
RFamide		1029.56290	DHVPFLRFamide
RFamide		1029.59920	LTHPFLRFamide
RFamide		1031.58970	AHKNFLRFamide
RFamide		1034.55300	NQPNFLRFamide

RFamide		1035.58470	SPRNFLRFamide
RFamide		1038.57310	LETNFLRFamide
RFamide		1041.63163	GPPSLRLRFamide
RFamide		1045.58020	GHRNFLRFamide
RFamide		1048.56870	APQGNFLRFamide
RFamide	NPY/PP peptide pem-PYF 3	1049.60031	YAIAGRPRFamide
RFamide		1050.60950	LELNFLRFamide
RFamide		1051.61600	LNRSFLRFamide
RFamide	F1	1053.55884	SDRNFLRFamide
RFamide		1059.59590	AHRNFLRFamide
RFamide		1061.62550	LPGVNFLRFamide
RFamide		1065.59520	DRNFVLRFamide
RFamide	F2	1066.59047	TNRNFLRFamide
RFamide		1069.55375	SDRNYLRFamide
RFamide		1071.56220	LFDDFLRFamide
RFamide		1073.55270	TNYGGFLRFamide
RFamide		1078.62690	LNRNFLRFamide
RFamide		1079.61090	LDRNFLRFamide
RFamide	Mar-FLP 6	1080.56974	DGGRNFLRFamide
RFamide		1094.58540	ENRNFLRFamide
RFamide		1104.60612	GAHKNYLRFamide
RFamide	pem-FLP 3	1104.60949	AQPSMRLRFamide
RFamide		1105.62650	APNKNFLRFamide
RFamide	pem-FLP 5	1105.62989	SMPSLRLRFamide
RFamide		1106.61050	LNPSNFLRFamide
RFamide		1116.54730	YEQDFLRFamide
RFamide		1118.65820	LGRPNFLRFamide
RFamide		1119.64550	SMPTLRLRFamide
RFamide	pem-FLP 4	1120.60441	SQPSMRLRFamide
RFamide		1121.62481	SM(O)PSLRLRFamide

RFamide		1121.65782	AYPSLRLRFamide
RFamide		1122.62792	RDRNFLRFamide
RFamide		1123.61190	TGNRNFLRFamide
RFamide		1124.63230	GLSRNYLRFamide
RFamide		1131.63090	DGPLAPFLRFamide
RFamide	Mar-FLP 8	1132.60104	VSHNNFLRFamide
RFamide		1135.58540	SQGLNSDLRFamide
RFamide		1135.58940	EFVDPNLRFamide
RFamide		1136.63230	LGDRNFLRFamide
RFamide		1137.59120	DGNRNFLRFamide
RFamide	NPY/PP peptide pem-PYF2	1138.61160	YSQVSRPRFamide
RFamide		1146.60545	GYSKNYLRFamide
RFamide		1147.63710	LNQPNFLRFamide
RFamide		1147.64050	EMPSLRLRFamide
RFamide	PrcFaRP 6	1147.64832	APQRNFLRFamide
RFamide		1148.66870	LSPRNFLRFamide
RFamide	PrcFaRP 4	1150.64800	ALDRNFLRFamide
RFamide		1158.61669	GYNRSFLRFamide
RFamide		1158.62790	YGNRPHLRFamide
RFamide		1159.61930	QYFMRLFRamide
RFamide		1163.63540	EM(O)PSLRLRFamide
RFamide		1163.6684	FEPSLRLRFamide
RFamide	NPY/PP peptide pem-PYF4	1164.67487	YSLRARPRFamide
RFamide	[ala1]-FaRP	1172.63234	AYNRSFLRFamide
RFamide		1174.63680	FTSKNYLRFamide
RFamide		1175.69090	ARPRNFLRFamide
RFamide		1179.67450	LTNRNFLRFamide
RFamide		1181.61742	SENRNFLRFamide
RFamide		1186.63680	LTNYGGFLRFamide
RFamide		1208.62831	DQNRNFLRFamide

RFamide		1209.61230	DENRNFLRFamide
RFamide	Mar-FLP 3	1215.62692	NYDKNFLRFamide
RFamide		1229.63130	LYEQDFLRFamide
RFamide	Lepidopteran peptide FLRFamide I	1229.64260	pEDVVHSFLRFamide
RFamide		1236.69600	LTGNRNFLRFamide
RFamide		1238.66400	SQPSKNYLRFamide
RFamide	Mar-FLP 7	1243.63307	GYGDRNFLRFamide
RFamide	Mar-FLP 5	1243.73820	DRTPALRLRFamide
RFamide		1244.71500	LDGPLAPFLRFamide
RFamide	Lepidopteran peptide FLRFamide I (Q1)	1246.66910	QDVVHSFLRFamide
RFamide		1253.65380	HDSPHVFLRFamide
RFamide	pem-FLP 2	1259.65313	AYSNLNYLRFamide
RFamide		1260.73240	LAPQRNFLRFamide
RFamide	[glu2-leu3]-SchistoFLRFa	1271.65310	pQDLDHVFLRFamide
RFamide		1271.68950	EPLDHVFLRFamide
RFamide		1272.61200	pQDNDHVFLRFamide
RFamide		1272.72120	HDLVQVFLRFamide
RFamide		1272.72120	HPLSFVSALRFamide
RFamide		1273.64360	YGSDRNFLRFamide
RFamide		1275.71680	DTSTPALRLRFamide
RFamide		1285.71640	LAYNRSFLRFamide
RFamide	PrcFaRP 1 FLRFamide	1287.65930	AYSDRNFLRFamide
RFamide		1287.68440	GYSVGLNYLRFamide
RFamide		1288.67968	QDLDHVFLRFamide
RFamide		1289.63855	GYSDRNYLRFamide

RFamide	pem-FLP 6	1300.75966	DGRTPALRLRFamide
RFamide	PrcFaRP 1 YLRFamide	1303.65420	AYSDRNYLRFamide
RFamide		1314.77530	DARTAPLRLRFamide
RFamide	[val2] pem-FLP 6	1342.80661	DVRTPALRLRFamide
RFamide	sulfakinin	1391.51430	pQFDEY(sulf)GHMRFamide
RFamide		1484.65332	GGGEYDDYGHLRFamide
RFamide	NPF (neuropeptide F)	3913.99050	KPDPSQLANMAEALKYLEQELDKYYSQVSRPRFamide
RPCH	Manto-CC	916.43120	pEVNFSPGWamide
RPCH	Nepci-AKH	920.42610	pELNFSSGWamide
RPCH	Grybi-Akh	920.42610	pEVNFSTGWamide
RPCH	Panbo-RPCH	930.44680	pELNFSPGWamide
RPCH	Ile2-Panbo-RPCH	930.44680	pEINFSPGWamide
RPCH	RPCH	930.4468	pQLNFSPGWamide
RPCH	Schgr-AKH-III	934.44170	pELNFSTGWamide
RPCH	Anaim-AKH	946.44170	pEVNFSPSWamide
RPCH	Dappu-RPCH	950.43670	pEVNFSTSWamide
RPCH	Corpu-AKH	960.45740	pELNFSPSWamide
RYamide		784.41005	FVGGSRYamide
RYamide		784.41005	FVNSRYamide
RYamide		832.41005	FYANRYamide
RYamide		841.43150	GFVSNRYamide
RYamide		862.42061	FYSQRYamide
RYamide		897.49411	L/IFVGGSRYamide
RYamide		940.51116	RFVGGSRYamide
RYamide		954.51560	LGFVSNRYamide
RYamide		959.47340	SGFYAPRYamide
RYamide		975.50470	LFYSQRYamide
RYamide		976.46354	SGFYANRYamide
RYamide		977.44760	SGFYADRYamide

RYamide		988.49990	VGFYANRYamide
RYamide		988.51120	RFYANRYamide
RYamide		1027.54320	SRFVGGSRYamide
RYamide		1030.47410	EWYSQRYamide
RYamide		1030.47411	pQGFYSQRYamide
RYamide		1030.51050	PAFYSQRYamide
RYamide		1089.54760	LSGFYANRYamide
RYamide		1114.57522	SSRFVGGSRYamide
RYamide		1143.55820	LEWYSQRYamide
RYamide		1227.65930	LSSRFVGGSRYamide
RYamide		2091.76000	(X)YANRYamide
SIFamide		1161.65274	RKPPFNGSIFamide
SIFamide	pem-FLP 7	1381.73753	GYRKPPFNGSIFamide
SIFamide		1395.75318	AYRKPPFNGSIFamide
SIFamide		1423.78448	VYRKPPFNGSIFamide
Tachykinin	TRP	605.30420	APSFGQamide
Tachykinin	TRP	679.37080	GFLGMRamide
Tachykinin	CabTRP lb	766.40285	SGFLGMRamide
Tachykinin		779.37560	APSGFLGM
Tachykinin		863.45560	PSGFLGMRamide
Tachykinin		917.46230	APSGFLGMRamide-NH3-loss
Tachykinin	CabTRP la	934.49273	APSGFLGMRamide
Tachykinin		935.47670	APSGFLGMR
Tachykinin		950.48765	APSGFLGM(O)Ramide
Tachykinin	CabTRP II	964.50330	TPSGFLGMRamide
Tachykinin		980.49821	TPSGFLGM(O)Ramide
Tachykinin		992.49820	APSGFLGMRG
Tachykinin	CalsTRP	1026.51890	YPSGFLGMRamide
Tachykinin		1027.50300	YPSGFLGMR
Tachykinin		1042.51830	YPSGFLGM(O)Ramide

CHH (Crustacean hyperglycemic hormone), CPRP (CHH precursor-related peptide), PDH(Pigment-dispersing hormone).

Appendix C: DiLeu and iDiLeu Synthesis and Labeling Protocols

C.1 Abstract

The protocols for N, N-dimethyl leucine (DiLeu) and isotopic DiLeu (iDiLeu) synthesis and labeling were performed exactly as written by Dr. Feng Xiang in his thesis [1]. For easy accessibility the following appendix contains the slightly modified passages from Chapters 2, 4, and 6 of his thesis that pertain to the synthesis of the tags and labeling used in this dissertation. The DiLeu tags follow the following overall protocol: tag synthesis, ^{18}O exchange, triazine ester synthesis (tag activation), and labeling. The iDiLeu tags follow the following protocol: tag synthesis, triazine ester synthesis (tag activation), and labeling.

C.2 Introduction

Isobaric MS^2 tagging approaches have been successfully used in MS-based quantitative proteomics. However, their application as a routine tool for quantitative MS studies is limited by high cost. The high cost of commercial TMTs and iTRAQ comes from the challenge of synthesizing these compounds; multiple steps involved in synthesis lead to moderate to low yield. A set of 6-plex deuterium-labeled DiART reagents was reported very recently with reduced cost of isobaric labeling. However, seven steps were still required to synthesize these compounds with 30%-40% overall yield [2]. A new type of isobaric MS^2 tags with fewer steps involved in synthesis is desirable to further reduce experimental cost while taking full technical advantages of isobaric MS^2 tagging approach. Formaldehyde dimethylation represents one of the most affordable approaches among all isotopic chemical derivatization techniques used for MS-based peptide and protein quantitation [3-18]. However, isotopic formaldehyde labeling is a mass-difference labeling approach and thus

lacks the advantages offered by isobaric labeling approach. We have used formaldehyde labeling technique for improved peptide fragmentation, enhanced neuropeptide *de novo* sequencing and quantitation [19-23]. A notable feature of this labeling approach is the production of intense immonium a_1 ions when dimethylated neuropeptides undergo MS^2 dissociation.[19, 24] The formation of the dimethylated a_1 ion is shown in Figure C.1A [19, 24]. The structural similarity of iTRAQ reporter ion and dimethylated amino acid a_1 ion inspired the design of novel dimethyl leucine isobaric MS^2 tags (DiLeu).

C.3 Materials for DiLeu and iDiLeu

All isotopic reagents for the synthesis of labels including: leucines (L-leucine and L-leucine-1-^{13}C, ^{15}N), heavy formaldehyde (CD_2O), sodium cyanoborodeuteride ($NaBD_3CN$), ^{18}O water 97% ($H_2{}^{18}O$) and deuterium water (D_2O) were purchased from ISOTEC Inc (Miamisburg, OH). Sodium cyanoborohydride ($NaBH_3CN$) was purchased from Sigma. N-Methylmorpholine (NMM) was purchased from TCI America (Tokyo, Japan). Chromasolv water, acetonitrile, and formic acid (FA) for UPLC were purchased from Fluka (Büchs, Switzerland). ACS grade methanol (MeOH), dichloromethane (CH_2Cl_2), acetonitrile (ACN), *N*, *N*-Dimethylformamide (DMF), and 4-(4, 6-Dimethoxy-1, 3, 5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM) were purchased from Fisher Scientific (Pittsburgh, PA). Ethanol was purchased from Pharmco-AAPER (Brookfield, CT). Ammonium formate, formaldehyde (CH_2O), tris-(2-carboxyethyl)phosphine (TCEP) 1M pH = 8.5, iodoacetamide (IAA), triethylammonium bicarbonate (TEAB), sodium dodecyl sulfate (SDS, ≥98%), α-cyano-4-hydroxy-cinnamic acid (CHCA) and bovine serum albumin (BSA) were purchased from Sigma (St. Louis, MO).

C.4 N, N-Dimethylated Leucine Protocol

Synthesis of N, N-Dimethylated Leucine (DiLeu)

A 120 mg portion of sodium cyanoborohydride ($NaBH_3CN$) was dissolved in 125 μL of H_2O or D_2O. A 100 mg portion of leucine or isotopic leucine was suspended in the mixture, the vial was sealed and mixture was kept in an ice-water bath for 30 min to cool down. 285 μL of light formaldehyde CH_2O (37% w/w) or a 530 μL heavy formaldehyde CD_2O (20% w/w) was then added to the mixture dropwise. The mixture was stirred in an ice-water bath for 30 min. The reaction was monitored by ninhydrin staining on a thin-layer chromatography (TLC) plate. The target molecule was purified by flash column chromatography ($MeOH/CH_2Cl_2$). Synthetic routes are shown in Figure C.2A. Both formaldehyde and sodium cyanoborohydride are very toxic by inhalation, in contact with skin, or if swallowed and may cause cancer and heritable genetic damage. These chemicals and reaction should be handled in a fume hood.

^{18}O Exchange of the DiLeu tag

^{18}O exchange is required for both the 115 and 116 labels to have the same masses as labels 117 and 118. ^{18}O exchange was carried out according to a procedure previously reported.[25] Briefly, 25 mg leucine or isotopic leucine was dissolved in 1N HCl $H_2{}^{18}O$ solution (pH 1) and stirred on a hot plate at 65˚C for 24 h, followed by reductive N, N-dimethylation. Synthetic routes are shown in Figure C.2A.

Synthesis of DiLeu Triazine Ester: Label Activation.

1 mg DiLeu in 50 μL DMF was combined with 1.86 mg DMTMM and 0.74 μL NMM in a 1.6 mL eppendorf vial. Mixing was performed at room temperature for 1h and

stored at -20˚C until needed for future labeling. The general synthetic route is shown in Figure C.2B.

General Procedure DiLeu Quntitation Preparation

The general structure, labeling, and quantitation scheme for DiLeu is demonstrated in Figure C.3. Crustacean tissue is dissected out according to the dissection protocols in Appendix A and then immediately transferred to 20 µL cold DMF. The tissue collected from a single crab is individually and manually homogenized in a glass tissue homogenizer using 120 µL of dimethylformamide (DMF). The homogenate from the single crab's neuropeptide extraction was centrifuged at 13200 rpm for 10min at room temperature in an Eppendorf 5415 D microcentrifuge (Brinkmann Instruments Inc., Westbury, NY). The supernatant was transferred to a clean 0.6 mL eppendorf vial and were dried under speedvac and stored at -80˚C for future labeling.

Dried tissue extracts were reconstituted in 10 µL of EtOH then labeled with 10 µL of freshly prepared 115,116,117 and 118 labeling stock solutions respectively for 2 h at room temperature. Labeled samples were dried under speedvac and reconstituted in 25 µL of 0.1% formic acid (FA) water solution. 4 µL of each labeled sample in the same subgroup were transferred in a Waters Maximum Recovery sample vial containing 5 µL of 0.1% FA to make 25 µL of sample solution for nanoLC ESI MS/MS analysis. 2 µL of prepared sample solution was used for analysis. Labeled samples across time points were also prepared and analyzed the same way described above.

C.5 Isotopic DiLeu Protocol

Synthesis of 5-plex iDiLeu Reagents

A 120 mg portion of sodium cyanoborohydride ($NaBH_3CN$) was dissolved in 125 μL of H_2O or D_2O. A 100 mg portion of leucine or isotopic leucine was suspended in the mixture, the vial was sealed and mixture was kept in an ice-water bath for 30 min to cool down. Light formaldehyde (CH_2O, 285 μL, 37% w/w) or heavy formaldehyde (CD_2O, 530 μL, 20% w/w) was then added to the mixture drop-wise. The mixture was stirred in an ice-water bath for 30 min. The reaction was monitored by ninhydrin staining on a thin-layer chromatography (TLC) plate. The target molecule was purified by flash column chromatography ($MeOH/CH_2Cl_2$). Synthetic routes are shown in Figure C.4.

Caution: Both formaldehyde and sodium cyanoborohydride are very toxic by inhalation, in contact with skin, or if swallowed and may cause cancer and heritable genetic damage. These chemicals and reactions should be handled in a fume hood.

Synthesis of iDiLeu Triazine Ester: Label Activation (same as DiLeu)

1 mg iDiLeu in 50 μL DMF was combined with 1.86 mg DMTMM and 0.74 μL NMM in a 1.6 mL eppendorf vial. Mixing was performed at room temperature for 1h and stored at -20˚C until needed for future labeling.

General Procedure of Standard Curve Mixture Preparation

Structures, molecular weights, and numbers of deuterium atoms of the synthesized iDilEu tags are demonstrated in Figure C.5. Five aliquots of neuropeptide standard EtOH solution (5 μg/μL) were prepared. 10 μL of each aliquot was labeled by 10 μL of freshly prepared i-DiLeu labeling solution (20 mg/mL) respectively. Labeled aliquots were diluted into the same concentration (10^{-5}-10^{-6} M). Reconstituting the standard samples into 50 μL of 0.1% FA water solution. 1 μL of the prepared standard curve mixture was analyzed by

MALDI TOF/TOF instrument. A standard curve was generated by averaging the peak intensities of the labeled neuropeptide standards after three technical replicates. A standard curve equation was then derived from the linear fitting of the standard curve.

C.6 Instrument Analysis

Reversed Phase NanoLC ESI MS/MS for DiLeu analysis

The DiLeu labeled neuropeptide samples were analyzed using a Waters nanoAcquity UPLC system coupled online to a Waters Micromass QTOF mass spectrometer (Waters Corp., Milford, MA, USA). The samples were dissolved in 0.1% formic $acid_{(aq)}$ (FA), injected and trapped onto a C_{18} trap column for 10 minutes (Zorbax 300SB-C_{18} Nano trapping column, Agilent Technologies, Santa Clare, CA, USA), and eluted onto a home-made C_{18} column (75 µm × 100 mm, 3 µm, 100Å) using a linear gradient (300 nl/min) from 95% buffer A (0.1% FA in H_2O) to 45% buffer B (0.1% FA in ACN) over 60 min. Survey scans were acquired from *m/z* 400-1800 with up to two precursors selected for MS^2 from *m/z* 50-1800 with 3 min dynamic exclusion.

MALDI TOF/TOF MS iDiLeu Analysis

An UltraflexTreme MALDI TOF/TOF analyzer (Bruker, Framingham, MA) equipped with a Smartbeam was used. Acquisitions were performed in positive ion reflectron mode. Mass spectra were obtained by averaging 500 laser shots covering the mass range *m/z* 400-2000. For sample spotting, equal volumes of 1 µL sample solution and matrix solution (150mg 2, 5 dihydroxybenzoic acid (DHB) in 50% methanol) were premixed and 1 µL of the mixture spotted and allowed to dry prior to analysis.

C.7 Works Cited

1. Xiang, F., *Developing novel labeling tools for neuropeptide quantitation and identification by mass spectrometry*. 2011, The University of Wisconsin - Madison: Ann Arbor. p. 251.
2. Zeng, D. and S. Li, *Revival of deuterium-labeled reagents for protein quantitation.* Chem. Commun., 2009(23): p. 3369-3371.
3. Ji, C., N. Guo, and L. Li, *Differential Dimethyl Labeling of N-Termini of Peptides after Guanidination for Proteome Analysis.* Journal of Proteome Research, 2005. **4**(6): p. 2099-2108.
4. Ji, C., et al., *Identification and Quantification of Differentially Expressed Proteins in E-Cadherin Deficient SCC9 Cells and SCC9 Transfectants Expressing E-Cadherin by Dimethyl Isotope Labeling, LC 鉴抏 ALDI MS and MS/MS.* Journal of Proteome Research, 2005. **4**(4): p. 1419-1426.
5. Ji, C. and L. Li, *Quantitative Proteome Analysis Using Differential Stable Isotopic Labeling and Microbore LC 鉴抏 ALDI MS and MS/MS.* Journal of Proteome Research, 2005. **4**(3): p. 734-742.
6. Huang, S.-Y., et al., *Quantitation of protein phosphorylation in pregnant rat uteri using stable isotope dimethyl labeling coupled with IMAC.* PROTEOMICS, 2006. **6**(6): p. 1722-1734.
7. Ji, C., et al., *Effect of 2MEGA Labeling on Membrane Proteome Analysis Using LC-ESI QTOF MS.* Journal of Proteome Research, 2006. **5**(10): p. 2567-2576.

8. Ji, C., et al., *A study of reproducibility of guanidination-dimethylation labeling and liquid chromatography matrix-assisted laser desorption ionization mass spectrometry for relative proteome quantification.* Analytica Chimica Acta, 2007. **585**(2): p. 219-226.

9. Guo, K., C. Ji, and L. Li, *Stable-Isotope Dimethylation Labeling Combined with LC 鐠杖SI MS for Quantification of Amine-Containing Metabolites in Biological Samples.* Analytical Chemistry, 2007. **79**(22): p. 8631-8638.

10. Wang, P., et al., *Targeted Quantitative Mass Spectrometric Identification of Differentially Expressed Proteins between Bax-Expressing and Deficient Colorectal Carcinoma Cells.* Journal of Proteome Research, 2009. **8**(7): p. 3403-3414.

11. Boersema, P.J., et al., *Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates.* PROTEOMICS, 2008. **8**(22): p. 4624-4632.

12. Raijmakers, R., et al., *Automated Online Sequential Isotope Labeling for Protein Quantitation Applied to Proteasome Tissue-specific Diversity.* Mol Cell Proteomics, 2008. **7**(9): p. 1755-1762.

13. Synowsky, S.A., et al., *Comparative Multiplexed Mass Spectrometric Analyses of Endogenously Expressed Yeast Nuclear and Cytoplasmic Exosomes.* Journal of Molecular Biology, 2009. **385**(4): p. 1300-1313.

14. Lemeer, S., et al., *Comparative Phosphoproteomics of Zebrafish Fyn/Yes Morpholino Knockdown Embryos.* Mol Cell Proteomics, 2008. **7**(11): p. 2176-2187.

15. Khidekel, N., et al., *Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics.* Nat Chem Biol, 2007. **3**(6): p. 339-348.

16. Rogers, L.D. and L.J. Foster, *The dynamic phagosomal proteome and the contribution of the endoplasmic reticulum.* Proceedings of the National Academy of Sciences, 2007. **104**(47): p. 18520-18525.

17. Aye, T.T., et al., *Selectivity in Enrichment of cAMP-dependent Protein Kinase Regulatory Subunits Type I and Type II and Their Interactors Using Modified cAMP Affinity Resins.* Mol Cell Proteomics, 2009. **8**(5): p. 1016-1028.

18. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat. Protocols, 2009. **4**(4): p. 484-494.

19. Fu, Q. and L. Li, *De Novo Sequencing of Neuropeptides Using Reductive Isotopic Methylation and Investigation of ESI QTOF MS/MS Fragmentation Pattern of Neuropeptides with N-Terminal Dimethylation.* Analytical Chemistry, 2005. **77**(23): p. 7783-7795.

20. Fu, Q. and L. Li, *Fragmentation of Peptides with N-terminal Dimethylation and Imine/Methylol Adduction at the Tryptophan Side-Chain.* Journal of the American Society for Mass Spectrometry, 2006. **17**(6): p. 859-866.

21. Fu, Q. and L. Li, *Investigation of several unique tandem mass spectrometric fragmentation patterns of NFDEIDR, an orcokinin analog, and its N-terminal dimethylated form.* Rapid Communications in Mass Spectrometry, 2006. **20**(4): p. 553-562.

22. DeKeyser, S.S. and L. Li, *Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry quantitation via in cell combination.* Analyst, 2006(131): p. 281-290.

23. Xiang, F., Q. Fu, and L. Li, *N, N-Dimethyl Amino Acids as iTRAQ Reagent for Improved Peptdiomics and Proteomics* 56th ASMS Conference Poster, 2008.

24. Hsu, J.L., et al., *Beyond Quantitative Proteomics: Signal Enhancement of the a1 Ion as a Mass Tag for Peptide Sequencing Using Dimethyl Labeling.* Journal of Proteome Research, 2005. **4**(1): p. 101-108.

25. Murphy, R.C. and K.L. Clay, *Preparation of Labeled Molecules by Exchange with Oxygen-18 Water.* Methods in Enzymology, 1990. **193**: p. 338-348.

C.8 Figures and Tables

Figure C.1

Formation of (A) dimethyl amino acid a_1 ion and (B) iTRAQ reporter ion [1].

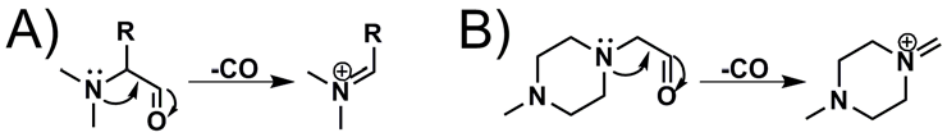


Figure C.2

(A) Synthesis of isobaric labels and (B) activation using 4-(4, 6-Dimethoxy-1, 3, 5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM)/ N-Methylmorpholine (NMM) [1].

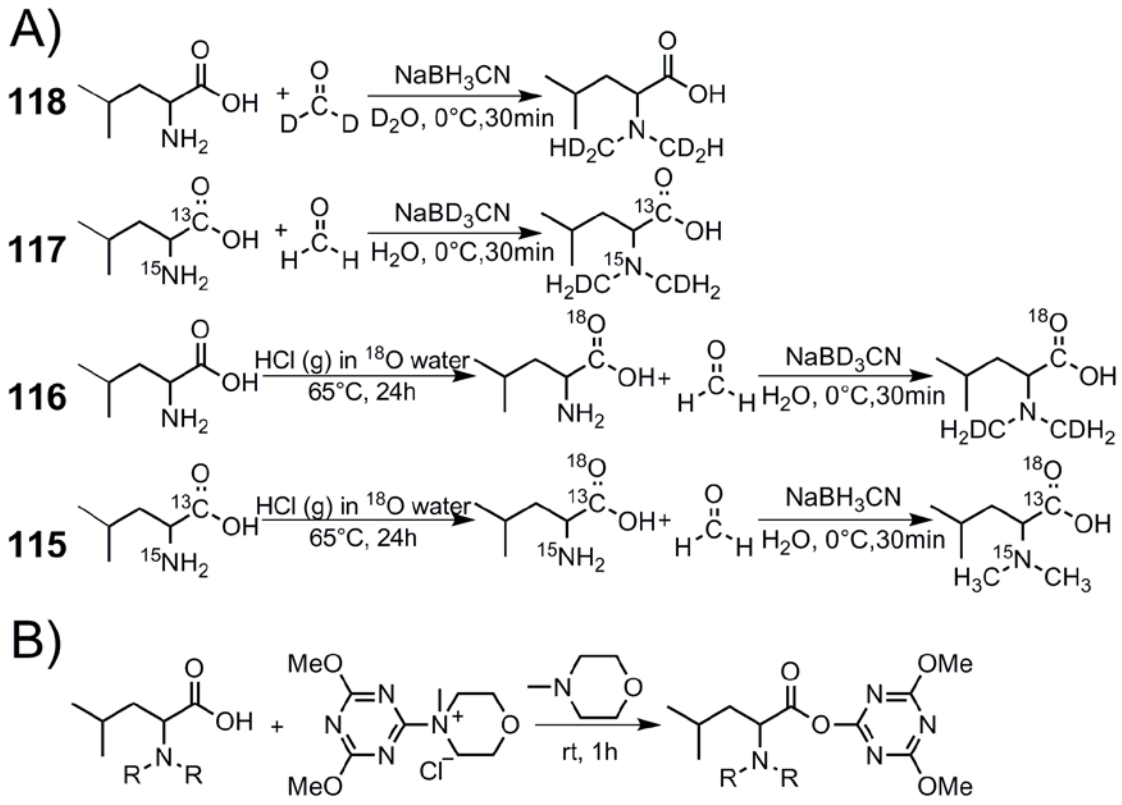


Figure C.3

General structure of dimethyl leucine isobaric tags is shown in (A). Formation of new peptide bond at N-terminus or ε-amino group of the lysine side-chain and isotope combination of isobaric tags are illustrated in (B). Quantitation of 4-plex isobarically labeled peptide is illustrated in (C) [1].

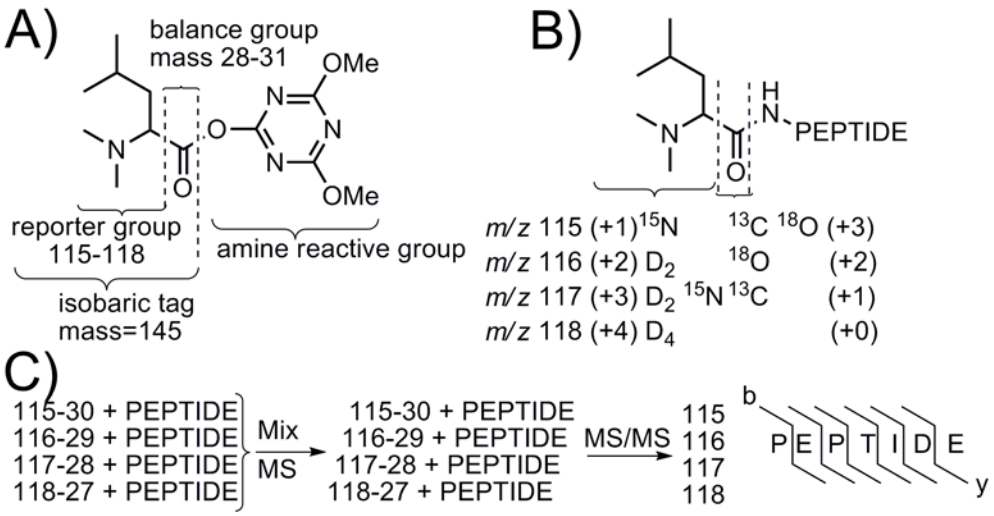


Figure C.4

Synthesis of 5-plex i-DiLeu reagents.

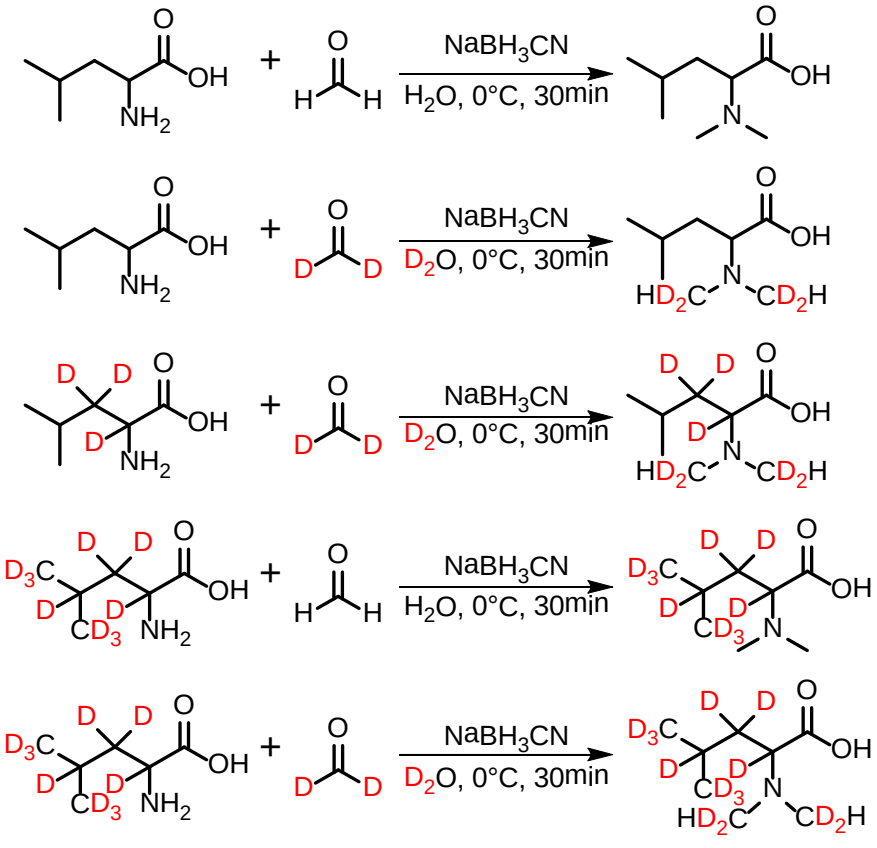
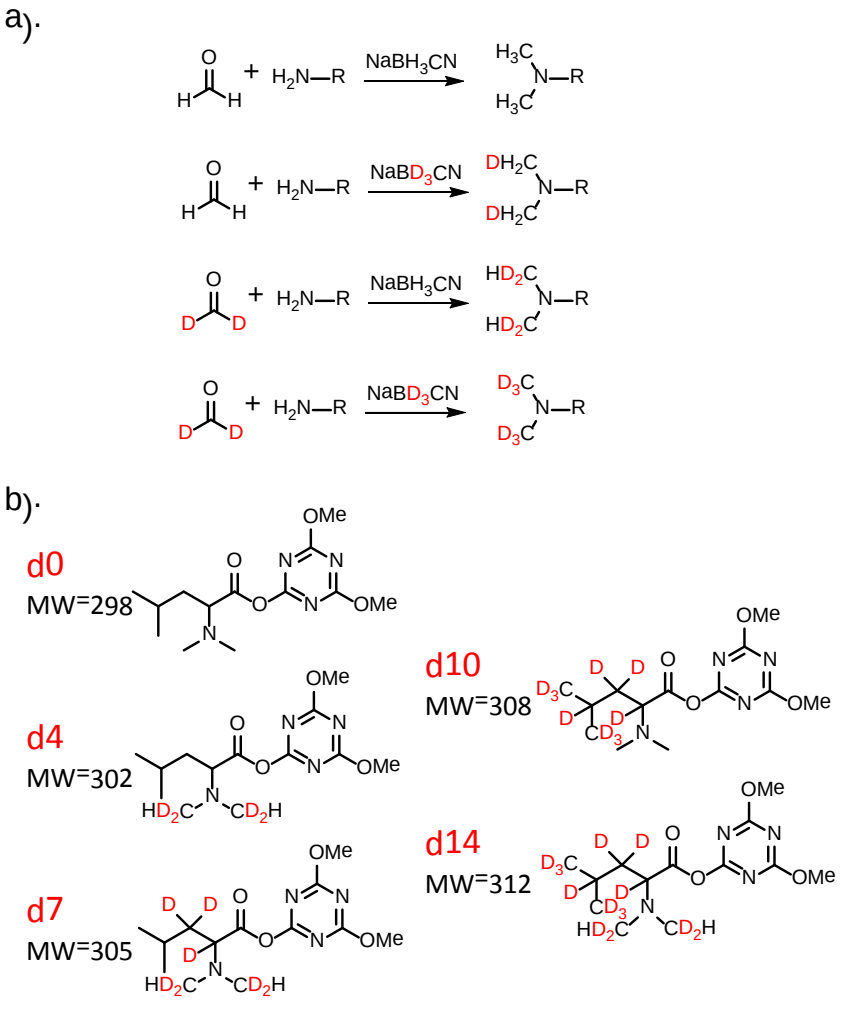


Figure C.5

a) Deuterium atom combinations of reduction dimethylation reaction; b) structures, number of locations of deuterium atoms and molecular weights of the 5-plex iDiLeu reagents) [1].



Crustacean Anatomy and Stuff

A quick guide to the anatomy of our decapod friends.

A few decapod species we have used in our lab.

Cancer irroratus Rock Crab

Cancer pagurus Edible Crab

Cancer borealis Jonah Crab

Homarus americanus
American Lobster

Carcinus maenas
Green Crab

Panulirus interruptus
California Spiny Lobster

Callinectes sapidus
Blue Crab

Taxonomy breakdown to Cancer

› cellular organisms

Superkingdom › Eukaryota

› Fungi/Metazoa group

Kingdom › Metazoa adult body composed of numerous cells differentiated into tissues and organs and usually a digestive cavity lined with specialized cells

› Eumetazoa multicellular animals with a digestive tract

› Bilateria all animals having a bilateral symmetry; having a front and a back end, as well as an upside and downside

› Coelomata body cavity between the body wall and intestine that is lined with a mesodermal epithelium

› Protostomia lower invertebrate whose mouth appears before the anus during development, cleavage is spiral and determinate, and the coelom forms as a splitting of the mesoderm

› Panarthropoda large monophyletic group including arthropods, tardigrades (water bears), and onycophorans (velvet worms)

Phylum › Arthropoda invertebrates having a segmented body and jointed appendages, usually a shell of chitin molted at intervals, and an anterior brain dorsal to the alimentary canal and connected with a ventral chain of ganglia

› Mandibulata arthropods possessing mandibles (lateral jaws) and antennae

› Pancrustacea proposed taxon comprising all crustaceans and hexapods

Subphylum › Crustacea aquatic arthropods characteristically having a segmented body, a chitinous exoskeleton, and paired, jointed limbs

Class › Malacostraca characterized by having a maximum of 19 pairs of appendages and trunk limbs which are sharply differentiated into thoracic and abdominal series

Subclass › Eumalacostraca crustaceans having eight thoracic segments, six abdominal segments, and a telson (a projection of the last body segment)

Superorder › Eucarida

Order › Decapoda crustaceans with five pairs of thoracic appendages one or more of which are modified into pincers, with stalked eyes, and with the head and thorax fused into a cephalothorax and covered by a carapace

Suborder › Pleocyemata

infraorder › Brachyura decapod Crustacea characterized by a small and short abdomen, which is bent up beneath the large cephalo-thorax.

Section › Eubrachyura

› Heterotremata/Thoracotremata group these two groups depend on where the gential open for the male species is (either the sternum or coxal (legs)); the gential opening for the female is always at the sternum.

Subsection › Heterotremata the gential opening for the female is sternum and that of the male is at the legs

Superfamily › Cancroidea

Family › Cancridae crabs

Genus › Cancer

Two Crab Families belonging to Heterotremata

CANCRIDAE p. 337

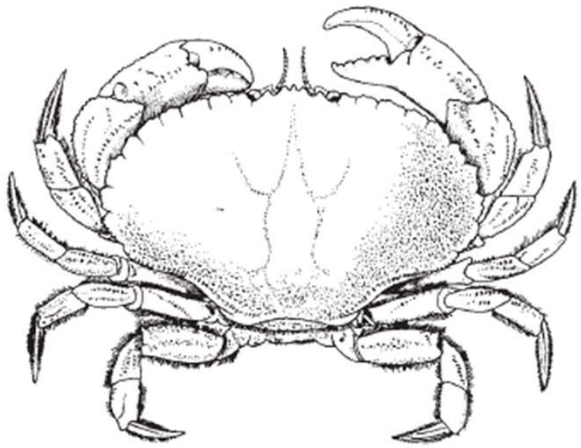
Rock crabs

Carapace broadly oval or hexagonal; front not produced in form of a rostrum but having a central tooth; anterolateral margins toothed (9 quadrangular or pentagonal teeth in species listed herein); lateral spines not strongly developed; antennules folding lengthwise. Found only in northern part of area. This family comprises 1 genus, *Cancer* Linnaeus, 1758, and 4 living species in the Atlantic ocean, 2 of which are eastern Atlantic in distribution (*Cancer bellianus* Johnson, 1861, and *Cancer pagurus* Linnaeus, 1758) and 2 western Atlantic (*Cancer borealis* Stimpson, 1859, and *Cancer irroratus* Say, 1817). Nations' (1979) proposition of dividing the genus *Cancer* into subgenera is followed here.

Two species of interest to fisheries marginally in the area:

Cancer (*Cancer*) *irroratur* Say, 1817.

Cancer (*Metacarcinus*) *borealis* Stimpson, 1859.



Also apart of the Portunidae family.

PORTUNIDAE p. 343

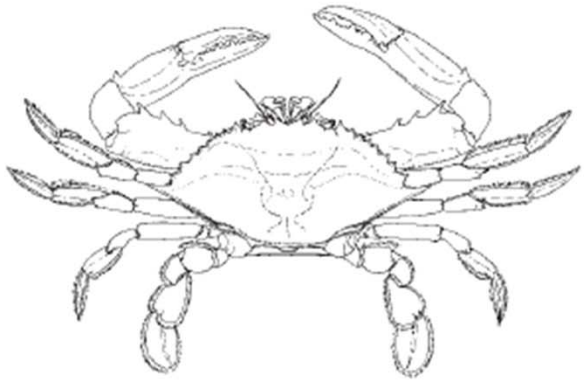
Swimming crabs

Carapace broad and flat, with 5 to 9 teeth on anterolateral margin; well-developed lateral spine at greatest width (in the species listed herein); last pair of legs usually distinctly adapted for swimming, flattened and paddle-shaped, but flattened without being paddle-shaped in a few species (not included here).

Ten species of interest to fisheries in the area:

Arenaeus cribrarius (Lamarck, 1818).

Callinectes bocourti A. Milne Edwards, 1879.
Callinectes danae Smith, 1869.
Callinectes exasperatus (Gerstaecker, 1856).
Callinectes larvatus Ordway, 1863.
Callinectes maracaboensis Taissoun, 1969.
Callinectes ornatus Ordway, 1863.
Callinectes rathbunae Contreras, 1930.
Callinectes sapidus Rathbun, 1896.
Callinectes similis Williams, 1966.



Male vs. Female

Claw Color

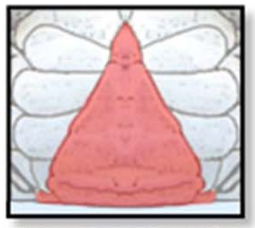
Female blue crab
Note her "painted" red claws

Male blue crab.
Note his blue tipped claws.

Abdomen

Blue Crab Abdomens

Male
"Jimmy"

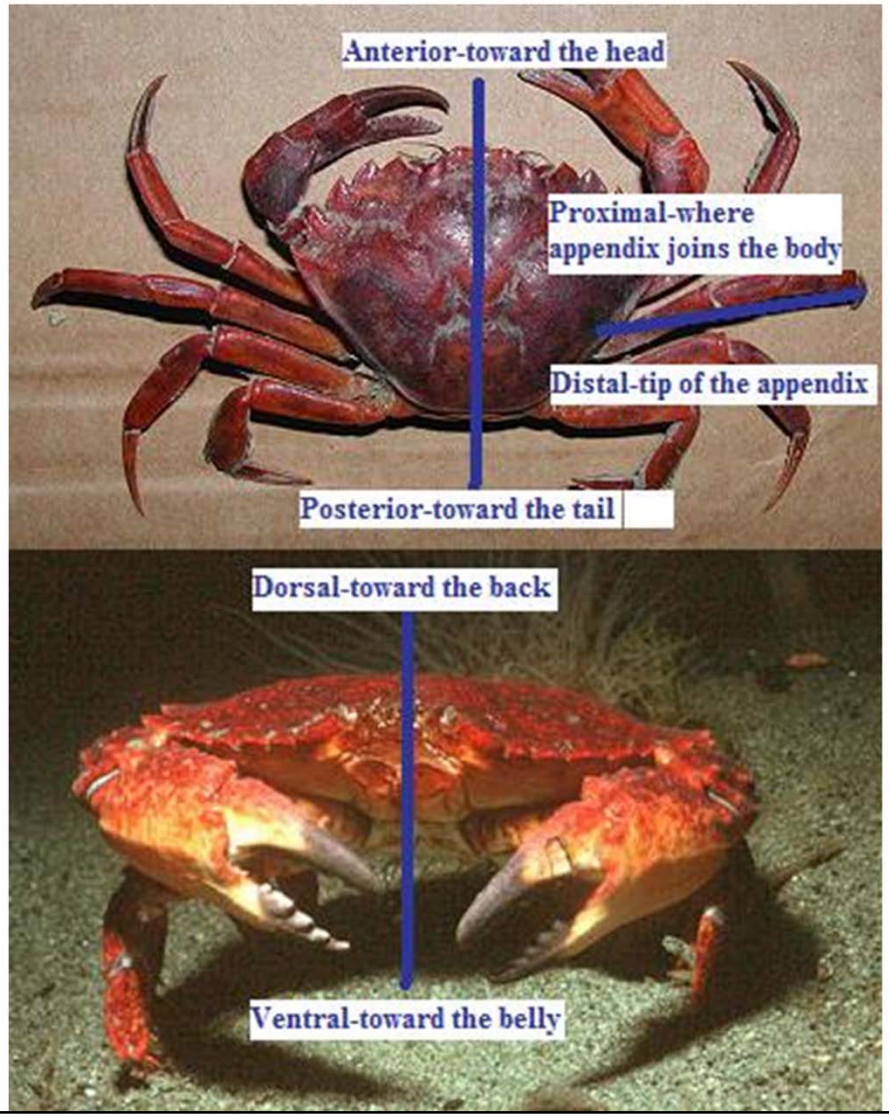


Immature Female
"Sally" or "She-Crab"

Mature Female
"Sook"

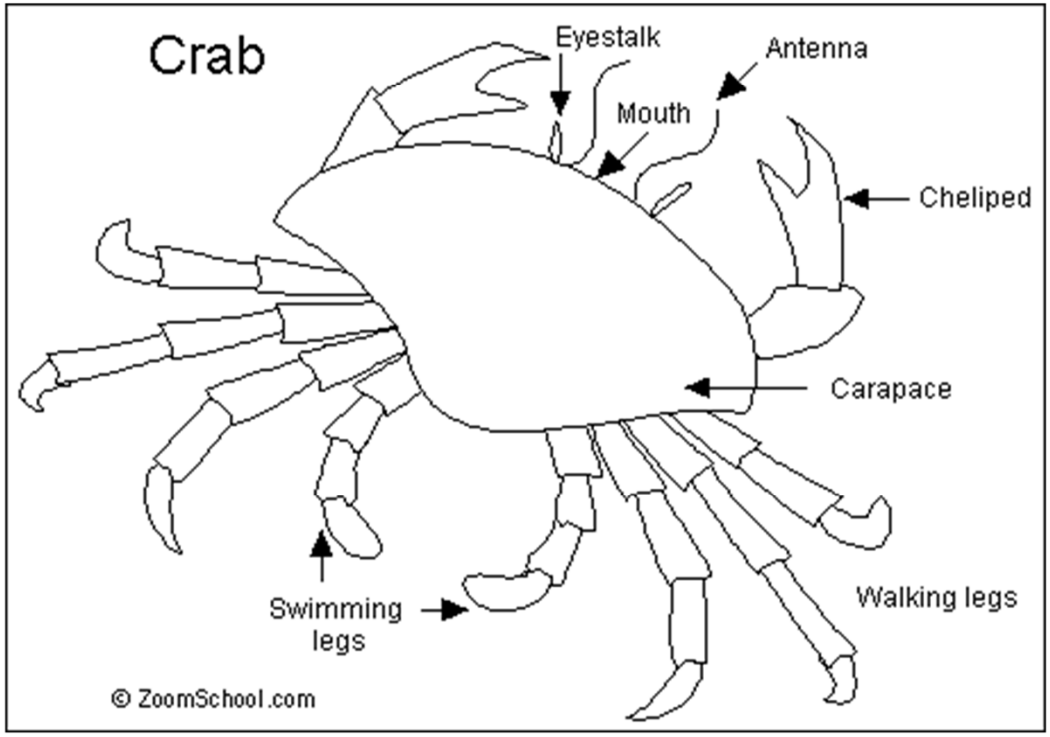
Note that the abdomen part is true for all crabs.

Anatomy Terms in Relation to the Crab

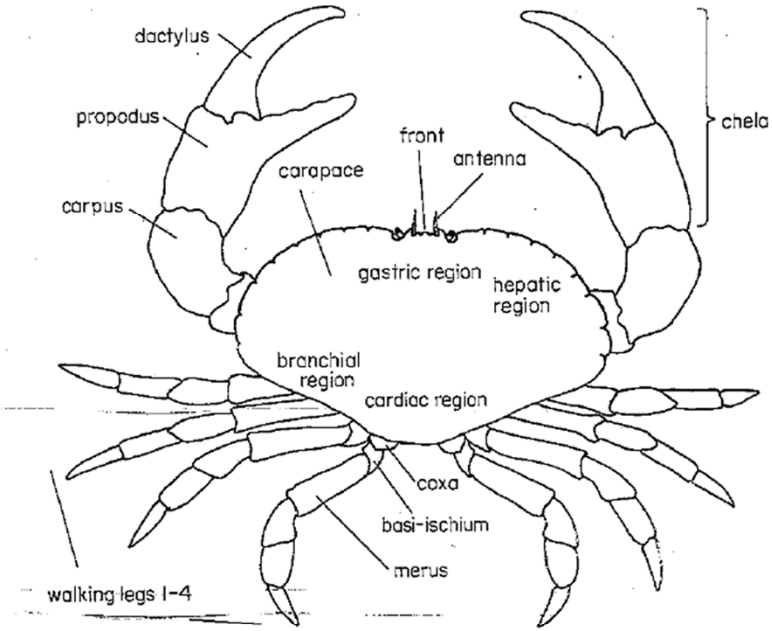


The top picture is of *C. maenas* and the bottom is of the Jonah.

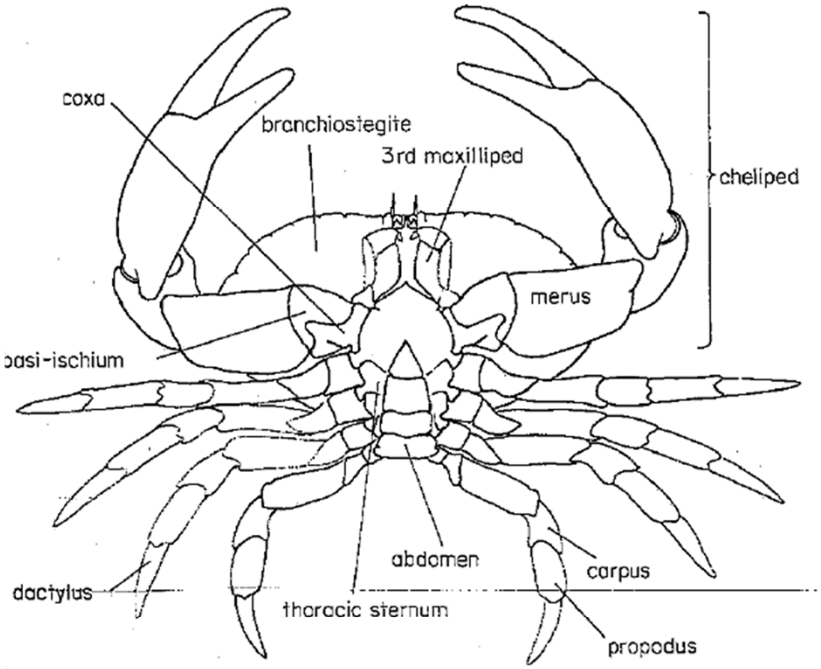
Very Basic Crab Anatomy



External parts of *Cancer pagurus* (male)



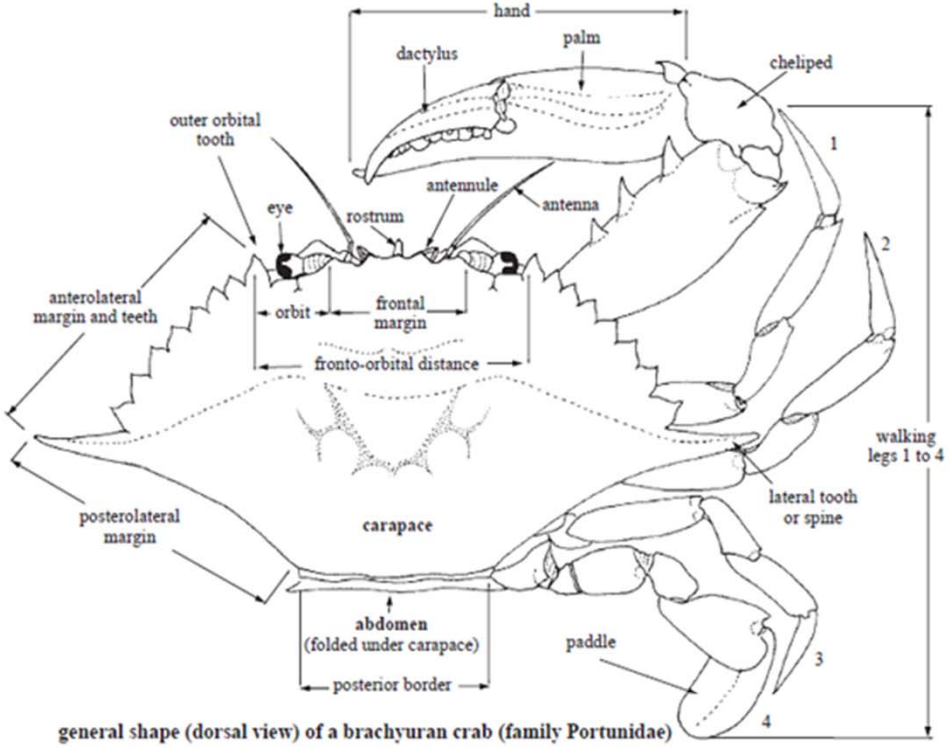
Dorsal view



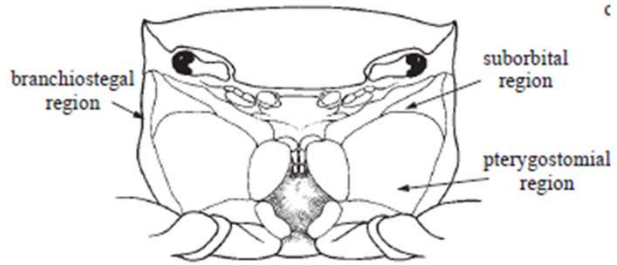
Ventral view

More Complex Crab Anatomy

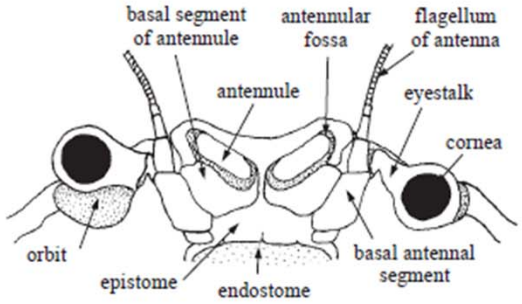
More Technical Terms and Measurements



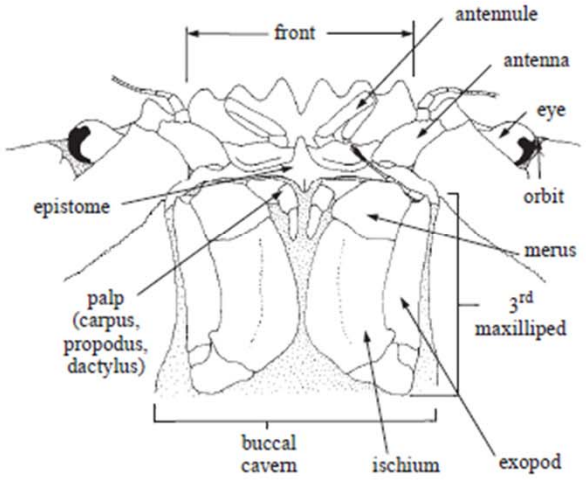
general shape (dorsal view) of a brachyuran crab (family Portunidae)



frontal view of body

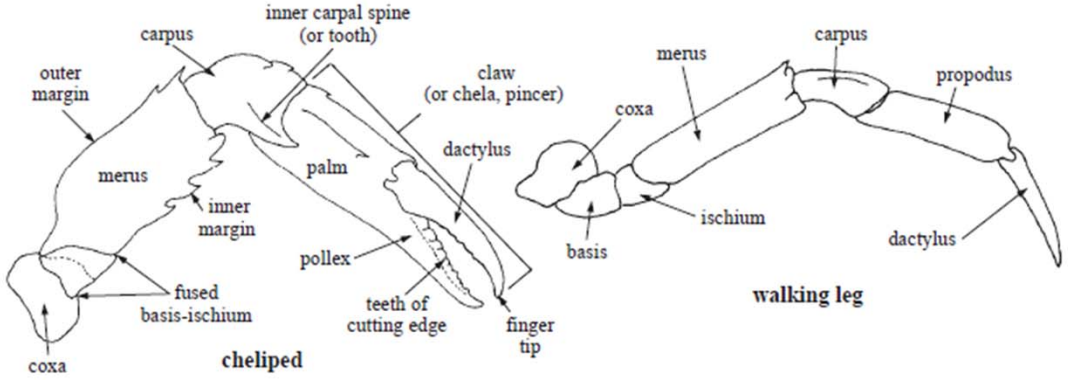


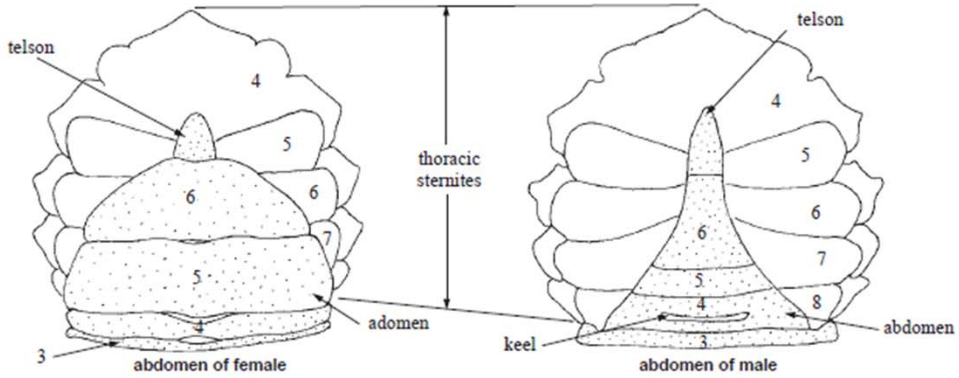
eyes, antennae, and antennules (ventral view)



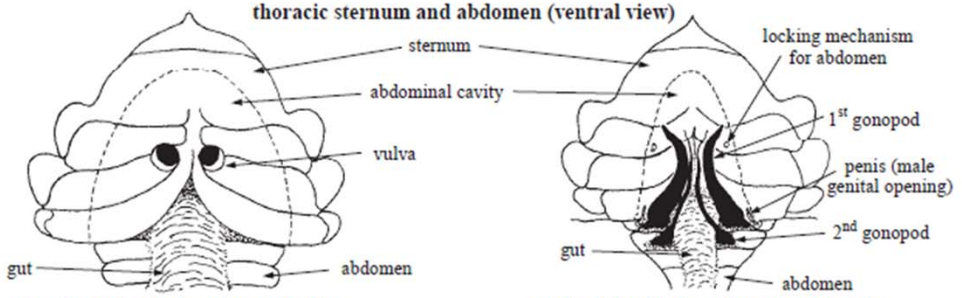
mouth field

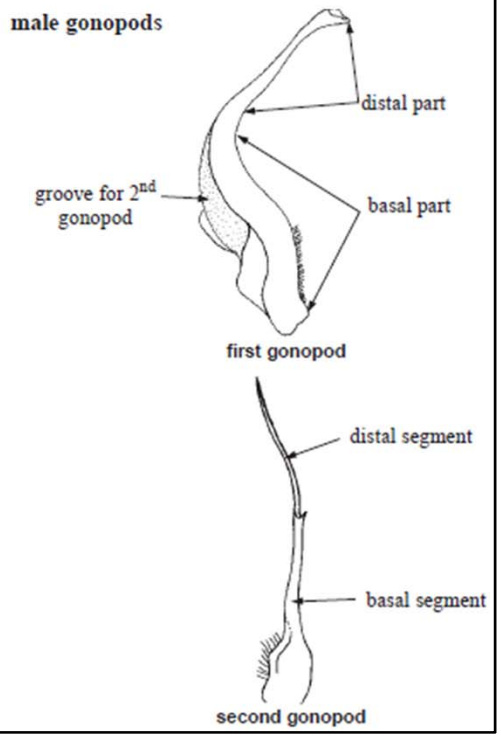
More Technical Terms



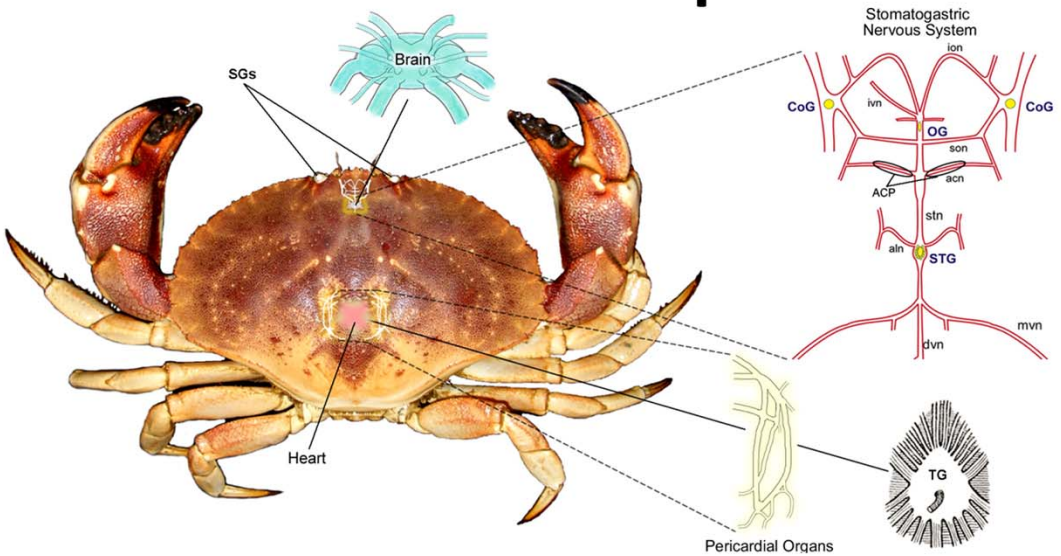


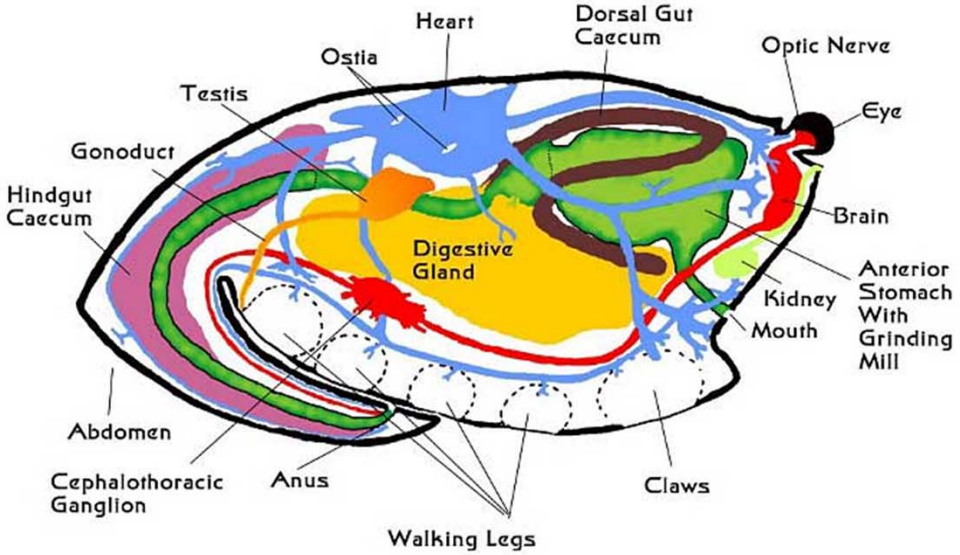
thoracic sternum and abdomen (ventral view)



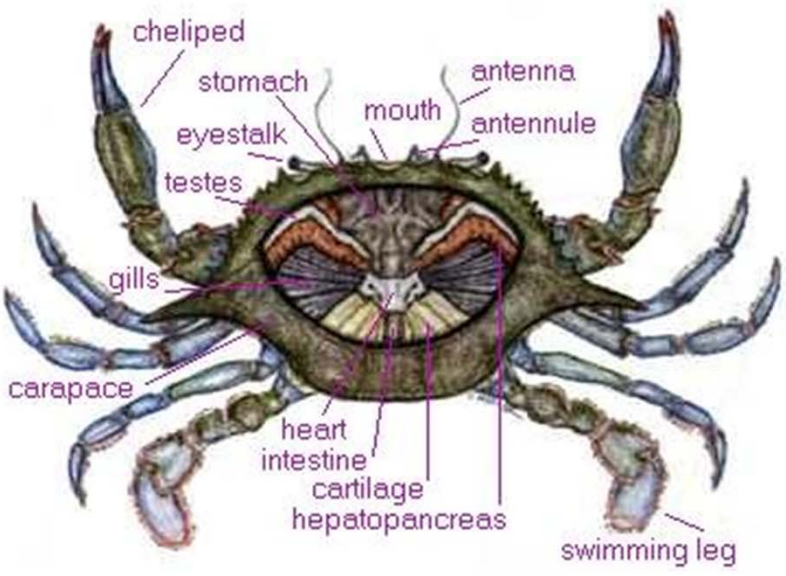


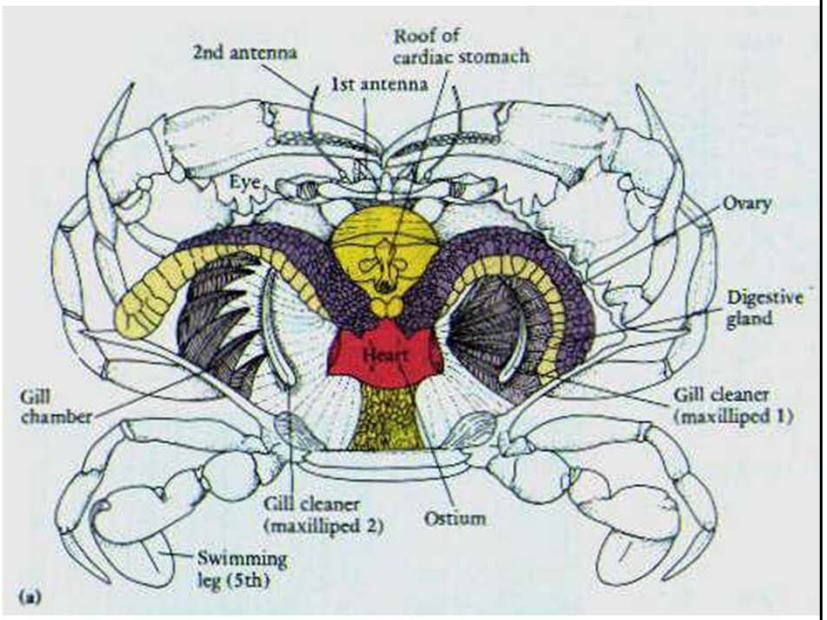
Location of important areas in crab



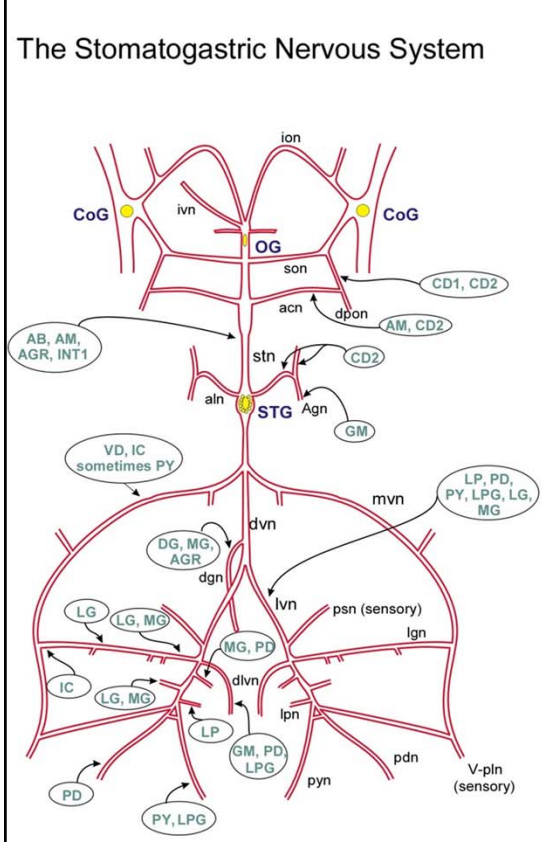


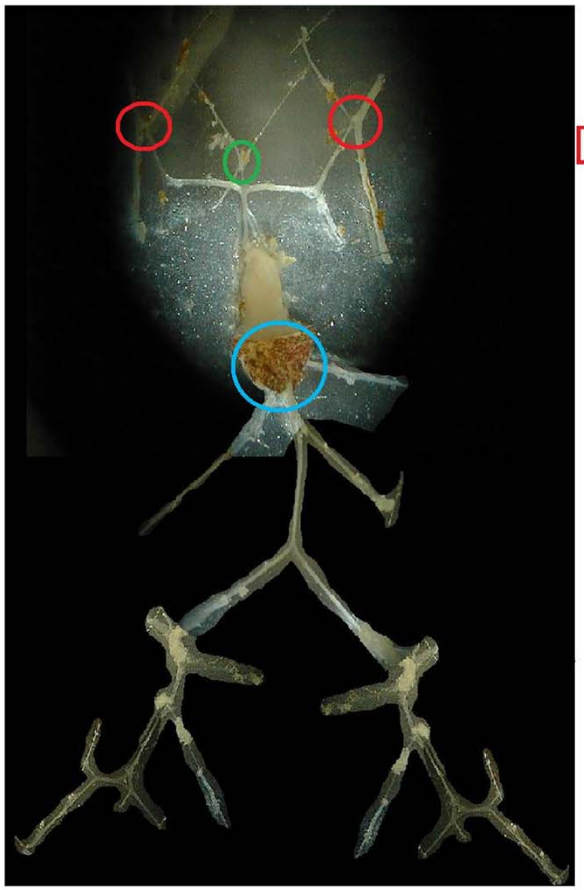
More images of the crab's insides

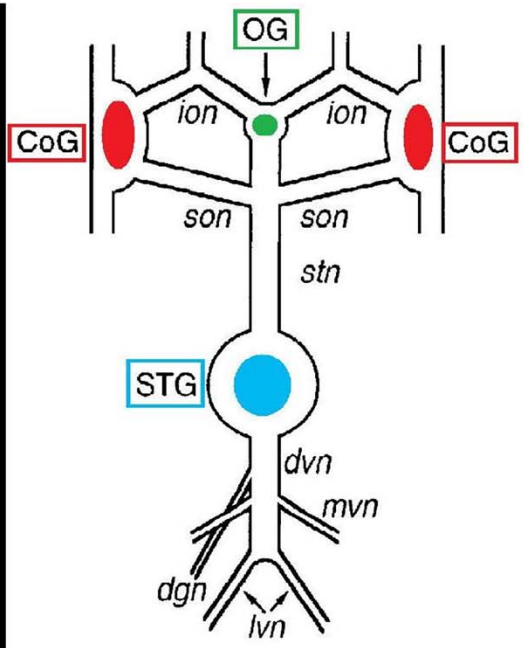




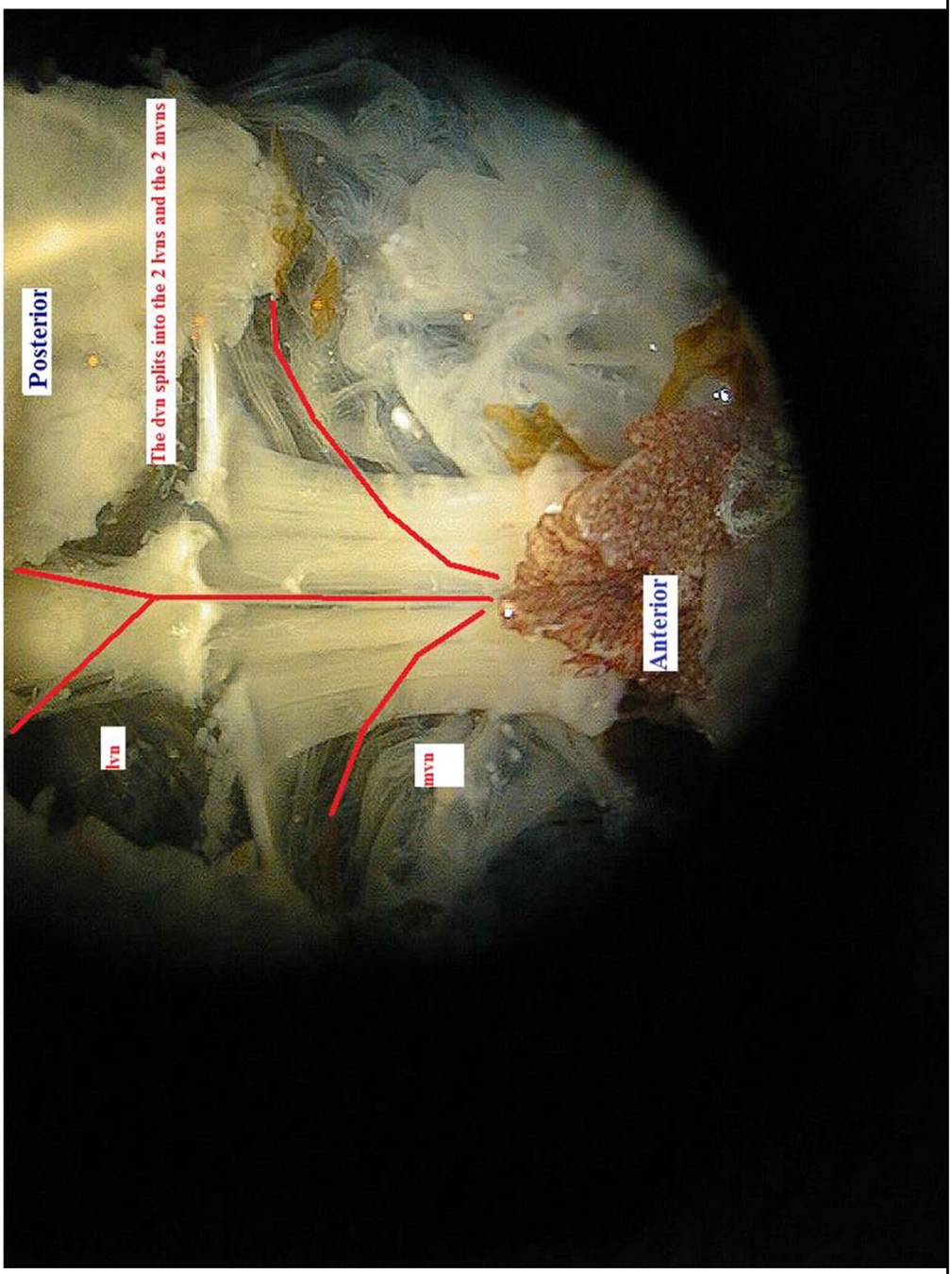
The Stomatogastric Nervous System



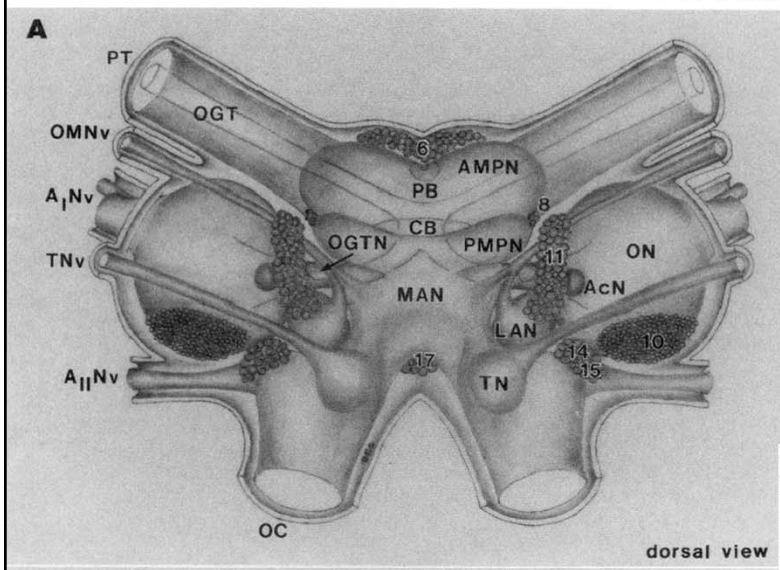


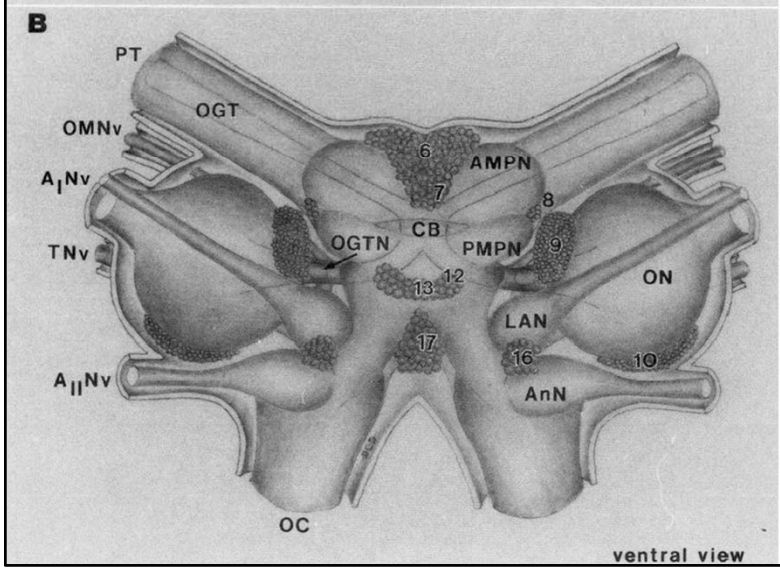


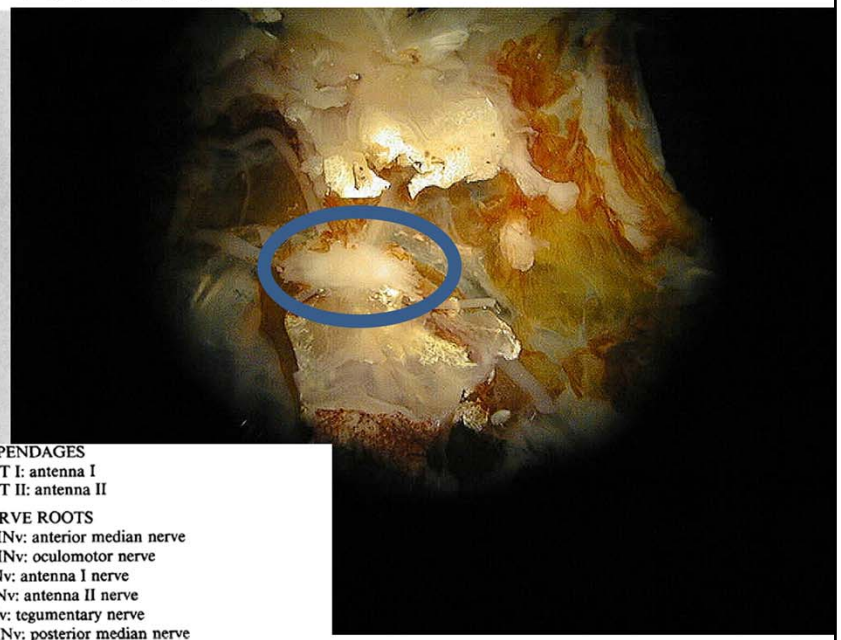
Locating the STG



Crab Brain







APPENDAGES
ANT I: antenna I
ANT II: antenna II

NERVE ROOTS
AMNv: anterior median nerve
OMNv: oculomotor nerve
A_INv: antenna I nerve
A_{II}Nv: antenna II nerve
TNv: tegumentary nerve
PMNv: posterior median nerve

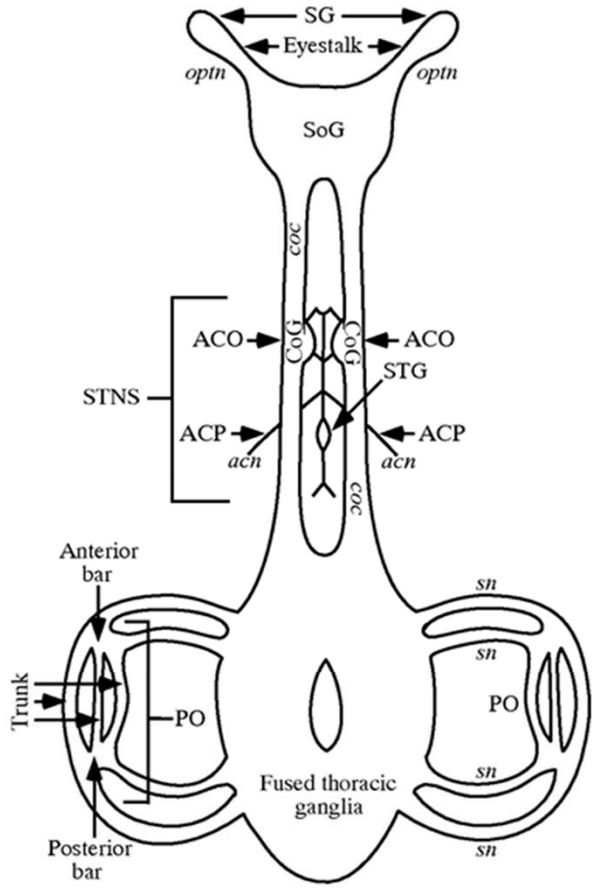
NEUROPILS
L: lamina
EM: external medulla
IM: internal medulla
TM: terminal medulla
HN: hemiellipsoid body
AMPN: anterior medial protocerebral neuropil
PMPN: posterior medial protocerebral neuropil
PB: protocerebral bridge
CB: central body
ON: olfactory lobe
LAN: lateral antenna I neuropil
MAN: median antenna I neuropil
AcN: accessory lobe
DCN: deutocerebral commissure neuropil
OGTN: olfactory globular tract neuropil
AnN: antenna II neuropil
TN: tegumentary neuropil

TRACTS AND COMMISSURES
OT: optic tract
PT: protocerebral tract
OGT: olfactory globular tract
DC: deutocerebral commissure
OC: oesophageal connectives

Brachyura (crabs; *e.g., Scylla, Carcinus, Callinectes, Hemigrapsus, Leptograpsus*). The crab *Scylla serrata* is taken as an example of a brachyuran. A dorsal view of the brain is shown in Figure 6A and a ventral view in Figure 6B. Like the crayfish, the three brain regions of the crab brain lie in one plane so that most of the neuropils can be seen from both sides.

The protocerebral tracts are large in brachyurans and contain a number of axons of a large diameter. The anterior medial protocerebral neuropils (AMPN) are also proportionately larger than in crayfish. The protocerebral bridge (PB), central body (CB), and the olfactory globular tract (OGT) are not different from those of crayfish; they occupy the same sites and are similarly proportioned.

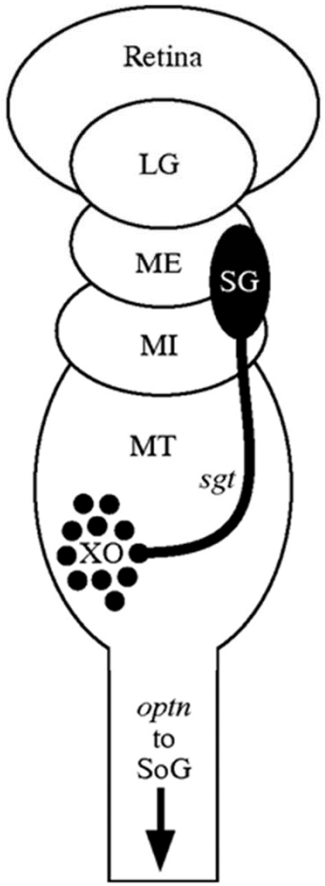
Schematic representation of the nervous system of a brachyuran crab illustrating the relative locations of the known neuroendocrine organs of Cancer species.



Hsu Y A et al. J Exp Biol 2006;209:3241-3256

The Journal of Experimental Biology

Schematic diagram of the optic ganglia of Cancer species highlighting the location and organization of the X-organ-sinus gland (XO-SG) system.



Hsu Y A et al. J Exp Biol 2006;209:3241-3256

The Journal of Experimental Biology

Dissection Resources

- http://www.jove.com/video/1207/cancer-borealis-stomatogastric-nervous-system-dissection
- The Tdrive
  - Folder Crab Dissection Procedures
  - Folder Tank and Animal Care Protocols
    - Folder Current protocols
      - File: Marine Crustacean Dissection Protocol 10122011.pdf

Works Cited

Slide #	References/notes
2	http://www.scandfish.com/ig/gallery.asp?categoryid=22 (used for all pictures except the spiny lobster.) http://www.dfg.ca.gov/marine/newsletter/1207.asp (spiny lobster picture)
3	http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?lvl=0&id=39395 Use the above website to find the taxonomy of any species. The image was modified to explain and identify the taxonomy levels.
4	http://www.mikedelaney.org/CrabPhotos.html Tavares, M., *True crabs*. The living marine resources of the Western Central Atlantic, 2002. **1**.
5	Images taken from http://www.bluecrab.info/identification.html
6	Online google photos
7	http://www.enchantedlearning.com/subjects/Crab.shtml
8	Warner, G.F., *The Biology of Crabs*. 1977: Elek Science.
9,10	Tavares, M., *True crabs*. The living marine resources of the Western Central Atlantic, 2002. **1**.
11	DeKeyser, S.S., *Development of mass spectrometric techniques for the analysis of neuropeptides: Differential display, quantitation, and imaging*. 2007, The University of Wisconsin - Madison: Ann Arbor. p. 232. http://reefkeeping.com/issues/2003-12/rs/images/image008LG.htm A side view of the internal anatomy of a typical crab, drawn as if the animal were cut open just to the side of the midline. The location of the claws and walking legs are given. The heart and major blood vessels are shown in blue, the gut is in green, and the nervous system is in red. Specific structures are labeled. Modified from McLaughlin
12	http://web.vims.edu/adv/ed/crab/guts3.html?svr=www http://library.thinkquest.org/26153/marine/arthro.htm

Slide #	References/notes (continued)
13	Image on left: http://www.bio.brandeis.edu/marderlab/figures/STNS_cancer_2.jpg Image on far right modified from Nusbaum, M.P. and M.P. Beenhakker, *A small-systems approach to motor pattern generation.* Nature, 2002. **417**(6886): p. 343-50. Image in the middle from dissection photos that use to be on the Tdrive
14	T drive image
15	Right brain image taken from Sandeman, D., et al., *Morphology of the brain of crayfish, crabs, and spiny lobsters: a common nomenclature for homologous structures.* The Biological Bulletin, 1992. **183**(2): p. 304-326.
16	Schematic representation of the nervous system of a brachyuran crab illustrating the relative locations of the known neuroendocrine organs of Cancer species. The central nervous system (CNS) of brachyurans is generally considered to consist of the supraesophageal (SoG) and fused thoracic ganglia, which are connected via the circumesophageal connectives (cocs). The optic nerves (optns) link the SoG with the ganglia of the eyestalks, the location of the neuroendocrine sinus gland (SG). Another well-known neuroendocrine site is the pericardial organ (PO), which is located in the pericardial chamber surrounding the heart. The POs consist of elaborations of the segmental nerves (sns), which project from the fused thoracic ganglia. Two additional neuroendocrine sites, the anterior cardiac plexus (ACP) and the anterior commissural organ (ACO), are contained within the stomatogastric nervous system (STNS), an offshoot from the CNS that overlies the foregut. The ACPs are located on the anterior cardiac nerves (acns) and the ACOs are located within the commissural ganglia (CoGs). For the sake of future discussion, the stomatogastric ganglion (STG) is also shown in this schematic. It should be noted that this illustration is not drawn to scale and that other portions of the nervous system have been excluded for the sake of simplicity. Hsu, Y.W.A., et al., *Members of the crustacean hyperglycemic hormone (CHH) peptide family are differentially distributed both between and within the neuroendocrine organs of Cancer crabs: implications for differential release and pleiotropic function.* Journal of Experimental Biology, 2006. **209**(16): p. 3241-3256.
17	Schematic diagram of the optic ganglia of Cancer species highlighting the location and organization of the X-organ-sinus gland (XO-SG) system. The nervous system contained within the eyestalk consists of several distinct regions, including the medulla terminalis (MT), the medulla interna (MI), the medulla externa (ME), the lamina ganglionaris (LG) and the retina. This system of ganglia is connected to the supraesophageal ganglion (SoG) via the optic nerve (optn). Located in the MT is a loosely associated collection of neurosecretory somata that are collectively termed the X-organ (XO). The release site of hormones produced by these somata is the sinus gland (SG), which is located at the junction of the MI and ME. The sinus gland tract (sgt) links the XO and SG. Hsu, Y.W.A., et al., *Members of the crustacean hyperglycemic hormone (CHH) peptide family are differentially distributed both between and within the neuroendocrine organs of Cancer crabs: implications for differential release and pleiotropic function.* Journal of Experimental Biology, 2006. **209**(16): p. 3241-3256.

Li Lab Introductory Packet

Hello and welcome to the Li Lab! Whether you are joining as a Graduate Student with the Department of Chemistry or Pharmacy, having a rotation with Dr. Li as a Pharmacy student, an Undergraduate Student here for the semester, a High School Intern, or beginning your Post-doctorate, this Introductory Packet will get you started.

Table of Contents

I. NIH Biosafety Online Courses: Biosafety 101, 104, and 201

Since the laboratory and campus receives NIH funding, we must follow certain protocols. One protocol is that all of lab members must undergo Biosafety Training. The Office of Biological Safety, here on campus has established online courses for all lab members to take to fulfill NIH and general safety requirements. Training must be completed before you began any work in the lab. You are only required to take the training once during your UW scientific career.

Biosafety 101, 104, and 201 training courses are online through the self- registration on the Learn@UW (https://learnuw.wisc.edu/) website. There is an online presentation with mini

quizzes during the presentation then a

final quiz at the end. You are allowed to take the final quiz as many times as necessary to pass. A passing score is at or above 70%.

Once you have completed the online courses, please create a pdf document of the passing scores and place in the Biosafety Training Verification Folder so that they are easily accessible to anyone as per regulations. There is also an excel sheet. Please check the excel page, add your name and place a Y once you have taken and passed (passing is a score at or above 70%) the quizzes.

If you have any further questions regarding the Biosafety tests, you may contact Erin Zimmerman or Chris Lietz. Erin and Chris are in charge of Laboratory Safety. Those of you first year Chemistry Graduates will need to contact them when you take the Laboratory Safety course offered in the Spring.

II. Li Lab Group Contact Information Sheet

Please give your name, phone number, and preferred email to Nicole Woodards to be placed on the Information Contact Sheet attached to this pack. This is how you are informed of important lab information.

III. T:/drive access, General Lab Computer Etiquette, and Virus Protection

The T:/ drive is an external drive that we keep lab information on. Each one of us has our own folders in the T:/drive and also general lab related information is also kept on the T:/ drive. All lab computers have access to the T:/drive. If you are joining the lab in a more permanent fashion, as a graduate or post-doc, then your personal computer that you will use in the lab may be connected to the T:/ drive as well. You will need to contact the DoIT help desk at help@doit.wisc.edu. You will need to say what type of antivirus software you have, identify yourself as a permanent Li Lab member, give the lab room and building (Rennebohm) you are in, and your MAC address. The MAC address can be located by clicking on the START button, typing cmd into the search bar, ENTER, and typing ipconfig /all. The physical address is the same thing as the MAC address.

General lab computer use is restricted to lab research only. You are not allowed to browse the internet if it is not research related nor perform personal tasks on laboratory computers. If you plan on using your own laptop or computer for data processing or analysis then please insure that you have current virus protection on your computer (this includes MAC books, they get virus as well).

IV. Crustacean Users: Occupational Health 101

If you think you are going to be using crustaceans in your research, and most everyone in the Li does. You may want to consider taking the Occupational Health 101: Safety for Personnel with Animal Contact also located in the self-registration section of the Learn@UW website. Although this course only mentions vertebrate animals, the information is still relevant

to us not only because we are surrounded by vertebrate animals in the vivarium but because the information contains sound safety practices that should be adopted by all animal users.

All information you need regarding the animal room is on the T:/ drive. Also, you may want to let it be known that you are interested in learning dissection procedure. Emails will go out during the semester about the availability of dead crabs to practice dissection on. If you require more help please feel free to contact Nicole Woodards.

V. Li Lab Key Card Access

To gain access to the Li laboratories, please contact Jenny Hergenrother, room 1214 in the business office. You will need your UW wiscard.

Again, if you have any questions, please ask anyone for help. The main people mentioned in this introductory packet are willing to help but anyone in the Li lab will point you the right direction. Welcome to the lab.

Appendix E: Mass Spectrometric Evaluation of Neuropeptidomic Profiles Upon Heat Stabilization Treatment of Neuroendocrine Tissues in Crustaceans

This appendix highlights my specific contribution to the published work [1]. Adapted from Robert Sturm, Tyler Greer, Nicole Woodards, Erin Gemperline, Lingjun Li. *J. Proteome Res.* **2013** 12(2): 743-52.

E.1 Abstract

To minimize biological variability between crab samples, tissue handling and extraction considerations were made and applied to determine the effectiveness of several tissue stabilization techniques. Heat stabilization using boiling and treatment with the Denator Stabilizor T1 was test against the normal extraction protocol of iced incubation with the extraction solvent to determine the degree to which post-mortem protease activity affects neuropeptidomic tissue studies in the crustacean *Callinectes sapidus* (blue crab). Neuropeptides in fixed and non-fixed tissue are extracted using acidified methanol or N,N-Dimethylformamide (DMF) and analyzed by MALDI-TOF and nanoLC-ESI-MS/MS platforms. Post-mortem fragments did not significantly affect MALDI analysis in the range *m*/*z* 650-1600, but observations in ESI MS/MS experiments suggest that putative post-mortem fragments can mask neuropeptide signal and add spectral complexity to crustacean neuropeptidomic studies. The impact of the added spectral complexity did not dramatically affect the number of detected neuropeptides between fixed and non-fixed tissues.

E.2 My Contributions

I developed a method of analysis for limiting animal to animal variability while obtaining the minimum sample required for detection thereby observing the 3Rs of animal research [2]. It

has been previously demonstrated that the pericardial organs (PO) within the crustacean nervous system are complementary in expression [3]. Two animals were required for analysis for the comparison of two techniques. To prevent the biological variability demonstrated in Chapter 4 of this dissertation, and allow for the heat stabilization techniques to be performed rapidly on the tissues, the POs of a single animal underwent separate fixing/non-fixing experimental protocols. In this respect, the animal became its own baseline for comparison. However, for detection on ESI mass spectrometer, both POs from a single crab are needed. Therefore two individual crabs underwent the experimental analysis and their POs were pooled at fixing/non-fixing was performed but before neuropeptide extraction performed. The experimental protocol is displayed in Figure E.1.

I preformed the majority of crustacean ordering, handling, and care. Macro and micro dissection and the fixing/non-fixing protocols were divided between three of us resulting in an assembly line. I preformed all micro dissections, which was timed for an average. My microdissection times were approximately four minutes for complete removal of the PO once the pericardial ridges were removed by macrodissection.

The assumption that the neuropeptide expression between right and left POs are similar required validation since left POs from two crabs were pooled for one extraction protocol and POs from the right side of two crabs were pooled for a different extraction protocol in each experiment[4, 5]. For validation of complementary neuropeptide expression, POs from two crabs were dissected out and the two left and right POs were pooled together separately. The neuropeptides were extracted from both sets of left/right PO samples with DMF and analyzed on the Synapt G2 QTOF mass spectrometer. Figure E.2 shows that the base peak ion (BPI)

chromatograms for left and right POs were similar, validating the complementary nature of neuropeptide expression between left and right organs. Therefore any differences in neuropeptide expression in the experiments results from the extraction protocols used and not the location of the POs or animal to animal variability.

The database I helped compile was utilized in this experiment for both manual and automated analysis to obtain the results represented in Tables E.1-E.2. Collecting the tandem mass spectra necessary to validate the results was equally divided among the four of us (Figure E.3).

E.3 Article Conclusions

The observations reported in the article suggest that post-mortem protein fragments do not significantly affect crustacean neuropeptidomic studies using MALDI or ESI MS. In MALDI experiments, Stabilizor T1 treated samples produce MS spectra containing peak profiles similar to those observed in samples not treated with heat fixation methods. In ESI experiments, tissues not treated with the Stabilizor T1 produced more intense and more complex mass spectra over the course of the elution gradient. The added complexity in the non-Stabilizor T1 treated tissue did not significantly affect the identification of endogenous crustacean neuropeptides through masking effects. The degree of added complexity was more prominent in DMF extracted tissues than in acMeOH extracted tissues. This suggests that acMeOH is a better protease inhibitor extraction solvent than DMF, possibly due to its low pH. Neuropeptides that were affected by coeluting chimeric species were still successfully *de novo* sequenced in four of five instances. Only one peptide (NFDEIDRSSFa) was completely masked by a chimeric species in non-treated DMF extracted tissues and not clearly observed. Although the majority of

extra ion peaks observed in non-Stabilizor T1 treated samples did not mask identified neuropeptides in blue crab PO extracts in this study, neuropeptidomic studies performed in other species and neural organs may be more affected by the high abundant, highly charged ions. We suggest that a preliminary experiment utilizing the Stabilizor T1 in the sample processing workflow should be performed to test its necessity in fully characterizing the neuropeptidome of another neural organ or another crustacean species. The benefits of decreased spectral complexity need to be balanced with slightly reduced neuropeptide signal to determine if the Stabilizor T1 or other heat deactivation step would be necessary and beneficial for the outcome of the experiment.

E.4 Acknowledgements

This work is supported in part by the National Institutes of Health grants (1R01DK071801, 1R56DK071801, and 1S10RR029531) and the National Science Foundation grant (CHE-0967784). L. Li acknowledges an H.I. Romnes Faculty Research Fellowship. R.M.S. acknowledges the NIH-supported Clinical Neuroengineering Training Program Predoctoral Fellowship (T32 EB011434). All authors contributed equally to the work.

E.5 Works Cited

1. Sturm, R.M., et al., *Mass spectrometric evaluation of neuropeptidomic profiles upon heat stabilization treatment of neuroendocrine tissues in crustaceans.* J Proteome Res, 2013. **12**(2): p. 743-52.

2. *Replacement, Reduction and Refinement of Animal Experiments in the Development and Control of Biological Products. Proceedings of a conference. Langen, Germany, November 2-4, 1994.* Dev Biol Stand, 1996. **86**: p. 1-368.

3. Fu, Q., et al., *Hormone complement of the Cancer productus sinus gland and pericardial organ: an anatomical and mass spectrometric investigation.* J Comp Neurol, 2005. **493**(4): p. 607-26.

4. DeKeyser, S.S., et al., *Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues.* J. Proteome Res., 2007. **6**(5): p. 1782-1791.

5. Pulver, S.R. and E. Marder, *Neuromodulatory complement of the pericardial organs in the embryonic lobster, Homarus americanus.* The Journal of comparative neurology, 2002. **451**(1): p. 79-90.

E.6 Figures and Tables

Figure E.1 Experimental workflow

For each experiment the pericardial organs (POs) from two blue crabs were divided in two and subjected to parallel extraction protocols. POs were either placed directly into an extraction solvent (acMeOH or DMF) or subjected to tissue fixation (Stabilizer T1 or boiling) and then placed in extraction solvent. Tissue was then homogenized, neuropeptides extracted, concentrated, and subjected to mass spectral analysis by nanoLC-ESI-MS/MS and MALDI-TOF-MS. It is expected to find *m/z* regions where no interference was observed, regions where an isotopic envelope from non-fixed samples masked neuropeptide signal, and regions where ion peaks were present only in non-fixed tissues, constituting additional non-interfering peaks [1].

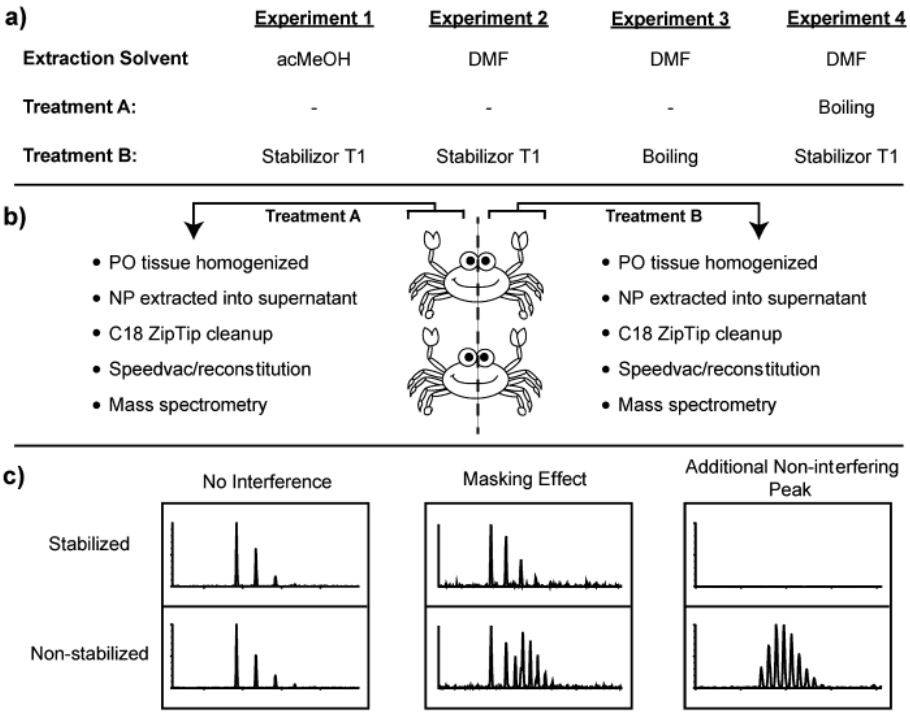


Figure E.2

Base peak ion (BPI) chromatograms of two pooled left and right PO neuropeptide extracts. Neuropeptides were extracted from the tissue using DMF. The BPI chromatograms for both the left and right extracts are similar in elution profiles and peak heights. This suggests that the neuropeptide contents of left and right POs are similar [1].

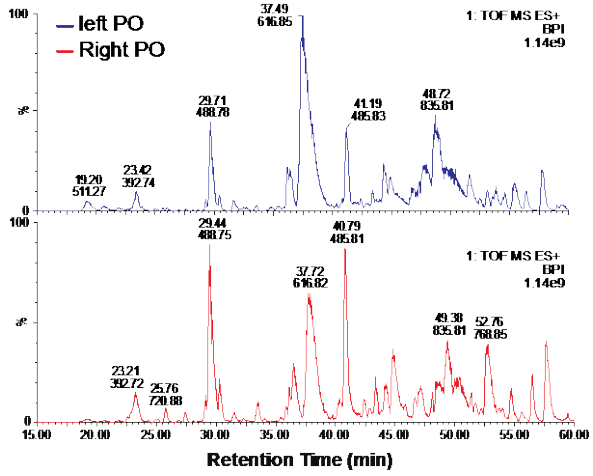


Figure E.3

MS/MS fragmentation spectra of masked orcokinin NFDEIDRSSFGFA ($[M+H]^+$ = *m/z* 1504.67) identified using Waters PepSeq algorithm. a) MS/MS of PO tissue extract treated with Stabilizor T1. b) MS/MS of PO tissue extract treated with boiling. Both tissue samples were extracted with DMF. The $[M+H]^+$ ion is correctly identified in the Stabilizor T1 treated sample with a PepSeq score of 403 whereas the $[M+H]^+$ ion is incorrectly identified in the boiled tissue extract resulting in a PepSeq score of 260.

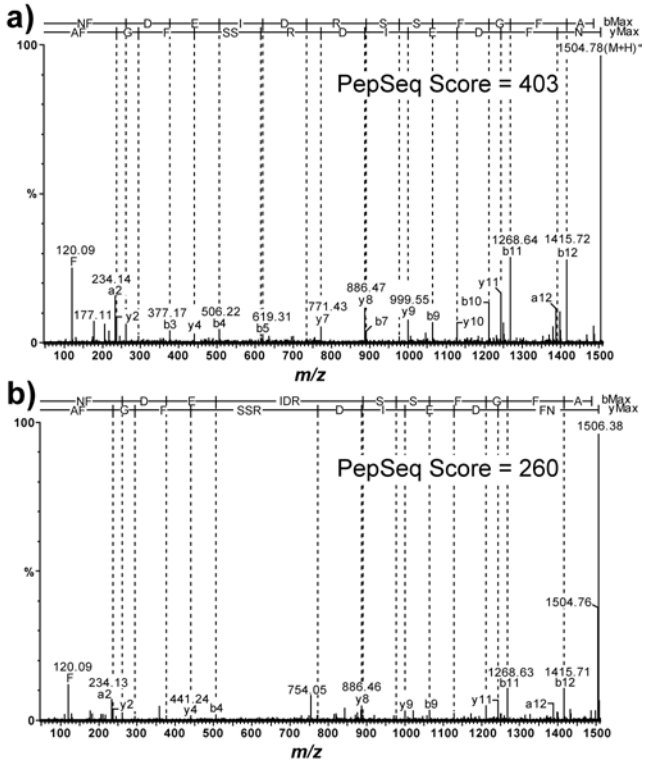


Table E.1

Neuropeptides observed in ESI mass spectra of each sample preparation experiment

Neuropeptides observed in ESI mass spectra of each sample preparation experiment[a]

			Experiment 1		Experiment 2		Experiment 3		Experiment 4	
Family	[M+H]+	Sequence	acMeOH	Stabilizor T1	DMF	Stabilizor T1	DMF	Boiling	Boiling	Stabilizor T1
AST-A	781.3879	DPYAFGLa				+				+
	854.4043	DGPYSFGLa	+	+	+	+				
	883.4308	SNPYSFGLa								+
	[b]893.4264	SGHYNFGLa	+						+	+
	897.4577	ARGYDFGLa	+	+	+	+	+	+	+	+
	909.4941	ARPYSFGLa	+	+	+	+	+	+	+	+
	911.4734	ARAYDFGLa							+	+
	928.4523	SSGQYAFGLa					+	+		+
	937.5254	PRVYSFGLa	+	+	+	+	+	+	+	+
	939.5047	TRPYSFGLa	+	+	+	+	+	+	+	+
	962.5094	APQPYAFGLa	+	+	+	+	+	+	+	
	[b]986.4876	RNMYSFGLa	+	+	+	+	+	+	+	+
	998.4942	TSPQYSFGLa			+	+				
	1004.4836	FSGTYNFGLa	+	+	+	+	+	+	+	+
	1021.5102	THPTYSFGLa	+	+	+	+	+	+	+	+
	1097.5990	AGLSALYSFGLa		+	+					
	1167.5793	AAGLQNYDFGLa	+	+	+	+	+	+	+	+
	2051.0345	GSGQYAYGLGKKAGQYSFGLa	+	+	+		+	+	+	
AST-B	1031.5057	AWSNLQGAWa	+	+	+	+		+	+	+
	1061.5527	AGWSSLKGAWa	+	+	+	+	+	+	+	+
	1107.5153	AGWSSMRGAWa	+	+	+	+	+	+	+	+
	1123.5102	AGWSSM(O)RGAWa	+	+	+	+	+	+	+	+
	[b]1165.5537	NWNKFQGSWa	+	+	+		+	+	+	
	1182.5691	TSWGKFQGSWa	+	+	+	+	+	+	+	+
	1209.5800	TGWNKFQGSWa	+	+	+	+	+	+	+	+
	1220.5807	SGDWSSLRGAWa	+	+	+	+	+	+	+	+
	1222.5752	GNWNKFQGSWa	+	+	+	+	+	+	+	+
	1252.5858	NNWSKFQGSWa	+	+	+	+	+	+	+	+
	1253.5698	NDWSKFGQSWa	+	+	+	+	+			
	1293.6335	STNWSSLRSAWa	+	+	+	+	+	+	+	+
	1380.6444	NNNWTKFQGSWa	+	+	+	+	+	+	+	+
	1470.7025	VPNDWAHFRGSWa	+	+	+	+	+	+	+	+
	1586.8413	MFAPLAWPKGGARWa	+	+	+	+	+	+	+	+
AST-C	[b]1899.8305	pQIRYHQcYFNPISdF	+	+	+	+	+	+	+	+
CCAP	956.3753	PFCNAFTGCa	+	+	+	+	+	+	+	+
FLP	[b]695.3988	NFLRFa	+	+	+	+	+	+	+	+
	[b]851.4999	RNFLRFa			+					
	[b]916.5152	THPFLRFa	+	+	+		+		+	+
	[b]925.4526	DDNFLRFa	+	+	+	+		+		
	926.5207	SKNYLRFa	+	+	+	+	+	+	+	+
	965.5428	NRNFLRFa	+	+	+	+	+	+	+	+
	966.5268	DRNFLRFa	+	+	+	+	+	+	+	+
	1022.5646	GNRNFLRFa	+	+	+	+	+	+	+	+
	1104.6061	GAHKNYLRFa	+		+	+	+		+	+
	1124.6323	GLSRNYLRFa	+	+	+	+	+		+	+
	[b]1132.6010	NVGSHGFLRFa	+	+		+	+		+	
	1146.6055	GYSKNYLRFa	+	+	+	+	+	+	+	+
	1147.6483	APQRNFLRFa	+	+	+	+	+	+	+	+
	1158.6167	YGNRSFLRFa	+	+	+	+	+	+	+	+
Myosuppressin	1125.5687	pQDLDHVFLR			+	+	+	+	+	
	1271.6531	pQDLDHVFLRFa	+	+	+	+	+	+	+	+
	1288.6797	QDLDHVFLRFa	+	+	+	+	+		+	
Orcomyotropin	1186.5164	FDAFTTGFGHS	+	+	+	+	+	+	+	+
Orcokinin	1228.5593	NFDEIDRSSFa	+	+	#	+				
	1256.5542	NFDEIDRSGFG	+	+						
	1270.5699	NFDEIDRSGFA	+	+	+	+	+	+	+	+
	1403.6226	NFDEIDRSGFGF	+	+	+	+	+	+	+	+
	1433.6332	NFDEIDRSSFGF	#	+	#	+	#	#		+
	1474.6597	NFDEIDRSGFGFA	+	+	+	+	+	+	+	+
	1502.6910	NFDEIDRSGFGFV	+	+	+	+	+	+	+	+
	1504.6703	NFDEIDRSSFGFA	#	+	#	+	#	#	#	+
	1532.7016	NFDEIDRSSFGFV	#	+	#	+	#	#	#	+
	1547.6815	NFDEIDRSSFGFN	+	+	+	+	+	+	+	+
Others	844.4788	HLGSLYRa	+	+	+	+	+	+	+	+
Proctolin	649.3668	RYLPT	+	+	+	+	+	+	+	+
RYamide	784.4101	FVGGSRYa	+	+	+	+	+	+	+	+
	832.4101	FYANRYa	+	+	+	+	+	+	+	+
	959.4734	SGFYAPRYa	+	+	+	+	+	+	+	+
	976.4635	SGFYANRYa	+	+	+	+	+	+	+	+
	1027.5432	SRFVGGSRYa					+	+	+	+
	1030.4741	pEGFYSQRYa	+	+	#	+	+	+	#	+
	1114.5752	SSRFVGGSRYa	+	+	+	+	+	+	+	+
Total Peptides			60	62	59	61	64	61	57	58
Unique Peptides			2	1	5	2	6	2	6	4

[a] AST, allatostatin; CCAP, crustacean cardioactive peptide; CPRP, crustacean hyperglycemic hormone precursor related peptide; FLP, FMRFamide like peptide. Amidation indicated by lowercase a. Pyroglutamation indicated by lowercase p. Disulfide bond indicated by lowercase c. Methionine oxidation indicated by M(O).

[b] indicates peptide had BPI > 1000, but no MS/MS event triggered

+ indicates that peptide was observed in ESI spectra

# indicates that there are additional peaks at $[M+2H]^{+2}$ that mask neuropeptide signal

Table E.2

Neuropeptides observed in MALDI mass spectra of each sample preparation experiment.

Neuropeptides observed in MALDI mass spectra of each sample preparation experiment[a]

			Experiment 1		Experiment 2		Experiment 3		Experiment 4	
Family	**[M+H]$^+$**	**Sequence**	**acMeOH**	**Stabilizor T1**	**DMF**	**Stabilizor T1**	**DMF**	**Boiling**	**Boiling**	**Stabilizor T1**
AST-A	656.34	SYAFGLa	+	+	+	+		+		
	739.38	GPYSFGLa				+				
	796.40	NPYSFGLa	+	+	+	+	+	+	+	+
	810.41	AGPYSFGLa	+	+	+	+	+		+	+
	854.40	DGPYSFGLa	+	+		+				+
	909.49	ARPYSFGLa	+	+	+	+	+	+	+	+
	937.52	PRVYSFGLa	+	+	+	+	+	+	+	+
	998.49	TSPQYSFGLa	+	+	+	+	+		+	+
	1004.49	FSGTYNFGLa			+	+	+		+	+
	1021.51	THPTYSFGLa	+	+	+	+				
	1052.59	SPRLTYFGLa	+	+	+	+	+	+	+	+
	1068.61	ALTTLYAFGLa	+	+	+	+		+	+	+
	1167.58	AAGLQNYDFGLa		+						
AST-B	1061.52	AGWSSLKGAWa			+	+	+	+	+	+
	1107.52	AGWSSMRGAWa	+	+	+	+	+	+	+	+
	1123.52	AGWSSM(O)RGAWa	+	+	+	+	+	+	+	+
	1182.57	TSWGKFQGSWa	+	+	+	+	+	+	+	+
	1209.58	TGWNKFQGSWa	+	+	+	+	+	+		
	1220.58	SGDWSSLRGAWa	+	+	+	+	+	+	+	+
	1222.58	GNWNKFQGSWa	+	+	+	+	+	+	+	+
	1252.59	NNWSKFQGSWa	+	+	+	+	+	+	+	+
	1253.62	NDWSKFGQSWa					+	+	+	+
	1293.63	STNWSSLRSAWa	+	+	+	+	+	+	+	+
	1380.64	NNNWTKFQGSWa	+	+	+	+	+	+	+	+
	1470.70	VPNDWAHFRGSWa	+	+	+	+	+	+	+	+
	1246.61	GSNWSNLRGAWa	+	+						
	1586.84	MFAPLAWPKGGARWa	+	+	+	+	+	+	+	+
AST-C	1899.83	pQIRYHQcYFNPISCF	+		+	+	+	+	+	+
CCAP	956.38	PFCNAFTGCa	+	+	+	+	+		+	+
CPRP	1046.49	FLSQDHSVN				+			+	+
	1363.69	RSAQGMGKMERL							+	+
FLP	965.54	NRNFLRFa	+	+	+	+	+	+	+	+
	966.53	DRNFLRFa	+	+			+	+		
	1022.56	GNRNFLRFa	+	+	+	+	+	+	+	+
	1104.61	GAHKNYLRFa					+	+		
	1147.65	APQRNFLRFa	+	+	+	+	+	+	+	+
	1158.62	YGNRSFLRFa	+		+	+	+	+	+	
Myosuppressin	1271.65	pQDLDHVFLRFa	+	+	+	+	+	+		
Orcomyotropin	1186.52	FDAFTTGFGHS	+	+						
Orcokinin	1171.54	EIDRSSFGFN	+	+						
	1256.55	NFDEIDRSGFG		+						
	1270.57	NFDEIDRSGFA		+						
	1403.62	NFDEIDRSGFGF	+	+	+	+	+	+	+	+
	1433.63	NFDEIDRSSFGF	+	+	+	+		+	+	+
	1474.66	NFDEIDRSGFGFA	+	+	+	+	+	+	+	+
	1502.69	NFDEIDRSGFGFV	+	+	+	+	+	+	+	+
	1504.67	NFDEIDRSSFGFA	+	+	+	+	+	+	+	+
	1532.70	NFDEIDRSSFGFV	+	+	+	+	+	+	+	+
	1547.68	NFDEIDRSSFGFN	+	+	+	+			+	+
Others	844.48	HLGSLYRa					+	+		+
RYamide	784.41	FVGGSRYa	+	+	+	+	+	+	+	+
	832.41	FYANRYa	+	+			+		+	+
	976.46	SGFYANRYa	+	+	+	+	+	+	+	+
	1030.45	pEGFYSQRYa	+	+	+	+	+	+	+	+
	1114.58	SSRFVGGSRYa	+	+	+	+	+	+	+	+
Total Peptides			44	45	40	43	40	38	40	41
Unique Peptides			2	3	0	3	5	3	1	2

[a] AST, allatostatin; CCAP, crustacean cardioactive peptide; CPRP, crustacean hyperglycemic hormone precursor related peptide; FLP, FMRFamide like peptide. Amidation indicated by lowercase a. Pyroglutamation indicated by lowercase p. Disulfide bond indicated by lowercase c. Methionine oxidation indicated by M(O).

+ indicates that peptide was observed in MALDI spectra