2013

Matrix-assisted laser desorption ionization mass spectrometry for identification of shrimp

Venkatesulu Salla
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the Chemistry Commons

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_theses/853

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
MATRIX-ASSISTED LASER DESORPTION IONIZATION MASS SPECTROMETRY FOR IDENTIFICATION OF SHRIMP

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirement for the degree of Master of Science

in

The Department of Chemistry

by

Venkatesulu Salla
B.S., Jawaharlal Nehru Technological University, 2006
M.S., McNeese State University, 2010
August 2013
This dissertation is dedicated to my parents: Munaswami Reddy and Chandramma and my wife, Srilakshmi, for their endless love and support throughout my life.
ACKNOWLEDGMENTS

This thesis would not have been possible without the help and support of my advisor Prof. Kermit K. Murray. I would like to express my sincere gratitude to my advisor for his invaluable guidance, leadership and patience during this program. I would also like to thank my committee members Dr. Isaiah Warner, Dr. Bin Chen, for your valuable time. To Dr. Azeem Hasan, thank you for allowing me to work in the mass spectrometry facility and giving me help. Thanks to Dr. Julie A. Anderson for assisting me in collecting shrimp samples. I am also grateful to the Gulf of Mexico Research Initiative and the Department of Chemistry, Louisiana State University for financial support.

To the Murray Research Group both past and present: Damien, Fan, Suman, Jaye, Lancia, Thabiso, Sung-gun, Jon, Pratap, Yonathan, and Chinthaka thank you for all your support and encouragement. I had a wonderful time working with you. Thanks to my friends Pratap and Dinesh for the great times we had together.

Finally, an immeasurable amount of gratitude to my dear wife, Srilakshmi, for her endless love, patience, and constant support during this journey. I am so blessed that you are in my life. I am greatly thankful to my wonderful parents for their love and support throughout my life and for always believing in me.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ iii

LIST OF TABLES .................................................................................................................. v

LIST OF FIGURES .............................................................................................................. vi

ABSTRACT .......................................................................................................................... vii

CHAPTER 1. INTRODUCTION ............................................................................................... 1
  1.1 Seafood Analysis for Quality Control ........................................................................... 3
  1.2 Mass Spectrometry for Biomolecule Based Seafood Analysis ........................................ 10
  1.3 Shrimp .......................................................................................................................... 16
  1.4 Research Objectives ................................................................................................... 18

CHAPTER 2. EXPERIMENTAL ............................................................................................. 19
  2.1 Bruker UltrafleXtreme Tandem Time-of-Flight Mass Spectrometer ............................... 19
  2.2 Biotyper ...................................................................................................................... 20
  2.3 Sample Collection ...................................................................................................... 21
  2.4 Sample Preparation and MALDI MS Analysis .............................................................. 22
  2.5 Construction of Reference Database and Identification of Unknown Shrimp ............ 23
  2.6 Reagents and Standards ............................................................................................ 24

CHAPTER 3. MATRIX-ASSISTED LASER DESORPTION IONIZATION MASS SPECTROMETRY FOR IDENTIFICATION OF SHRIMP ........................................... 27
  3.1 Introduction ................................................................................................................ 27
  3.2 Experimental .............................................................................................................. 27
  3.3 Results ....................................................................................................................... 28

CHAPTER 4. CONCLUSION AND FUTURE DIRECTION .................................................. 37

REFERENCES ...................................................................................................................... 38

VITA .................................................................................................................................. 48
LIST OF TABLES

Table 2-1. Description of identification scores ................................................................. 21
Table 2-2. Shrimp species used to establish the reference database for MALDI-TOF MS based species identification .......................................................... 22
Table 2-3. List of peptide calibration standard II peptide masses ........................................ 25
Table 3-1. Species identification results for the unknown shrimp samples ......................... 32
Table 3-2. MALDI mass spectra of individual shrimp matched against the mass spectrum library. ........................................................................................................... 33
Table 3-3. Identification test for Litopenaeus setiferus and Farfantepenaeus azticus from different regions, different sizes and fresh/frozen, F. azticus (fresh vs. frozen, small vs. jumbo) ........................................................................................................... 35
Table 3-4. Species identification results for the imported shrimp samples ......................... 36
LIST OF FIGURES

Figure 2-1. Diagram of the Bruker Ultraflextreme MALDI TOF/TOF mass spectrometer ........ 20

Figure 2-2. MALDI mass spectrum of Peptide Calibration Standard II.................................. 26

Figure 3-1. MALDI mass spectra of white shrimp (*Litopenaeus setiferus*) a) without desalting and b) with desalting using a C$_{18}$ reverse phase pipette tip. ..................................................... 28

Figure 3-2. MALDI-TOF mass spectra obtained from the analysis of (a) *Litopenaeus setiferus*, (b) *Pandalus platyceros*, (c) *Pandalopsis dispar*, (d) *Pleoticus robustus*, (e) *Sicyonia brevirostris*, and *Farfantepenaeus azticus* (f) showing the differences among the different shrimp species. .................................................................................................................... 29

Figure 3-3. Graphical output from a *P. robustus* reference spectrum ................................... 30

Figure 3-4. Dendrogram of six shrimp MSPs.................................................................................. 31

Figure 3-5. The match of individual species to the species library ................................................. 34
ABSTRACT

Matrix-assisted laser desorption ionization (MALDI) time of flight mass spectrometry was used to identify shrimp at the species level using commercial mass spectral fingerprint matching software (Bruker Biotyper). In the first step, a mass spectrum reference database was constructed from the analysis of six shrimp species that are commercially important in the United States: *L. setiferus, F. azticus, S. brevirostris, P. robustus, P. dispar and P. platyceros*. In the second step, the reference database was tested using 74 unknown shrimp samples from these six species. Correct identification was achieved for 72 of the 74 samples (97%): 72 samples were identified at the species level and 2 samples were identified at the genus level using the manufacturer’s log score specifications. The MALDI fingerprinting method for the identification of shrimp species was found to be reproducible and accurate with rapid analysis.
CHAPTER 1. INTRODUCTION

The central objective of the work described in this thesis was to develop a matrix-assisted laser desorption ionization mass spectrometry (MALDI MS) fingerprinting approach for the identification of shrimp at the species level.

The quality and safety of food is essential for the food industry and public health because food consumers are susceptible to any form of contamination that may occur in any phase starting from food production to until food reaches to the consumers. The United States is as a high seafood consuming nation, particularly shrimp, which is a widely consumed seafood item, authentication of shrimp is important for the U.S. seafood industry. As a consequence of recent BP Oil Spill in the Gulf of Mexico, seafood consumers are more concerned about food safety.\(^1\)\(^-\)\(^2\) Moreover, with increased demands for higher quality seafood products, adulterated seafood products where willful or unintentional substitution of higher value of species with less valuable species have become more widespread.\(^3\) Food authentication is a major concern not only for the elimination of commercial fraud, but also for the assessment of safety risks caused by the introduction of any food ingredient that might be harmful to human health.\(^4\)\(^-\)\(^6\)

It is difficult to identify shrimp species based on their morphological characteristics due to their phenotypic similarities, especially when the external carapace is removed in processing. Identification based on the biomolecule signatures of the different species can be a more reliable option and methods based on DNA and protein analysis have been developed to achieve this goal.\(^7\)\(^-\)\(^8\) Analytical electrophoretic and immunological methods have been adapted to the identification of seafood species as have techniques based on DNA amplification and hybridization. DNA based studies have focused not only on species identification, but also on population structures, phylogeography and phylogenetic relationships.\(^7\)
Recently, proteomics tools based on mass spectrometry methods have been adapted to food analysis and to the area of seafood identification and assessment.\(^9\)-\(^{10}\) Among these proteomics tools, MALDI MS is a well-established technique for the identification of bacteria, yeasts, and fungi in food quality control.\(^{11}\)-\(^{12}\) MALDI analysis can provide a molecular fingerprint of individual species of interest with reliability that is comparable to established genetic methods.\(^{13}\)-\(^{14}\) MALDI MS has been used for the identification of microorganisms in two approaches: mass fingerprinting and proteomic-based identification.\(^{15}\)-\(^{16}\) The mass fingerprinting approach has the advantage of fast and minimal sample preparation compared to proteomic-based identification.\(^{17}\)

In our research work, six shrimp species that are important to the U.S. commercial seafood industry were analyzed. Due to the fact that seafood species contain high levels of salts, an extraction protocol was developed for salt removal using the pipet tips containing reverse phase C\(_{18}\) stationary phase inside them. The desalted sample was mixed with MALDI matrix (2,5-dihydroxy benzoic acid) and then spotted onto a target for MALDI MS analysis. The MALDI-TOF mass spectra were then analyzed using fingerprinting software (Bruker Biotyper) that was originally developed for microorganism identification.

In the database, a library of reference spectra of each species was created using an average of 20 mass spectra from each species. To measure the accuracy of the shrimp reference library, mass spectra from unknown shrimp samples (from the six library species) were used for the identification. Out of 74 shrimp samples analyzed for species identification, 72 samples were correctly identified at the species level, and two samples were identified at the genus level. In addition, shrimp samples of different size (small, jumbo) and state (fresh, frozen) were also
identified. Three imported shrimp species analyzed using the database and did not show a false positive.

1.1 **Seafood Analysis for Quality Control**

Seafood is a rich source of nutrients for a healthy diet and its consumption has been linked to a wide range of health benefits.\(^\text{18}\) Consequently, seafood consumption has dramatically increased in recent decades and reached new high with average Americans consuming 15.1 pounds per year in 2011.\(^\text{19}\) Although a major share of this seafood is imported from different countries around the globe, the U.S. also maintains a large seafood production sector. In this context, food quality, including safety and authentication are major concerns facing the food industry today. There is increasing consumer concern for food quality and safety and, at the same time, there has been a significant market increment in high value products consumption. Therefore, it is very important that both domestically processed and imported seafood is safe, free from adulteration, and properly labeled.

Food quality refers to the quality characteristics of food that is acceptable to consumers such as intrinsic qualities (e.g. food safety, nutrition, taste, and texture) and extrinsic qualities (e.g. certification, labeling, and price). Important intrinsic and extrinsic quality attributes are foodborne pathogens, heavy metals and toxins, pesticide or drug residues, soil and water contaminants, place of origin, authentication, food substitution, and labeling.\(^\text{20}\) Food quality can be affected in any phase including location and conditions of catch, processing, and handling throughout the supply chain.

Food safety risks are highly important as they are directly related to human health. In recent years, consumers have become increasingly concerned about consuming wild-caught seafood due to water contamination in coastal regions.\(^\text{21}\) The majority of the U.S. domestic
seafood supply comes from wild population (wild-caught); contamination of water in fisheries such as the Gulf of Mexico, South Atlantic, and Pacific Coast could potentially cause seafood contamination. For instance, during the BP Deepwater Horizon oil spill in 2010, over 200 million gallons of oil poured into the Gulf, followed by 1.8 million gallons of dispersants intended to break down the oil into droplets.\textsuperscript{22-23} This oil spill had a tremendous impact on the U.S. seafood industry. Oil contains numerous toxic compounds that are released into the water as it breaks down.\textsuperscript{24-25} The most toxic oil compounds such as polycyclic aromatic hydrocarbons (PAHs) and poly chlorinated biphenyls (PCBs) are known carcinogens and neurotoxins.\textsuperscript{1-2} These toxic compounds are bioavailable to marine species because their solubility in water.\textsuperscript{26} PAHs have the potential to accumulate in seafood species and cause health risks to humans via ingestion of contaminated seafood.\textsuperscript{27-28}

It is also important to assess potential food safety risks associated with imported seafood.\textsuperscript{29} With increased international seafood trade, consumers benefit from year-round supplies, lower prices, and the introduction of food variety. However, globalization of seafood supply could introduce new food safety risks and spread contaminated seafood.\textsuperscript{30-31} Moreover, food safety risks associated with imported seafood are much higher when compared to domestic seafood due since three fourths of the seafood consumed in the U.S. are imported from different countries. Most of the seafood such as shrimp and fish imported from these developing countries are labeled as farm-raised, but in reality, they often are produced under unsafe and unhealthy conditions.\textsuperscript{32} The presence of a high number of developing country exporters that may have limited infrastructure increases the risk of contaminants in exports, which may be exacerbated by toxins and bacteria that are native to the tropical areas that produce much of the supply of seafood.\textsuperscript{33} Potential seafood hazards that can affect human health include bacteria, virus,
parasites, chemical residues, and antibiotics. Upon entry into the U.S., shrimp account for the highest number of FDA detentions among all seafood imported products. Most of these detentions are due to presence of salmonella bacteria and harmful antibiotics such as oxytetracycline, nitrofurazone, chloramphenicol, and other chemicals in the imported seafood.

Another important concern that is affecting the U.S. seafood industry is commercial fraud or adulteration. With increased seafood imports and decreased monitoring, fraud and deception in seafood marketing is becoming more widespread. Because of intra-species phenotypic similarities, unintentional or willful adulteration of seafood products can occur. Moreover, the intra species observable characteristics are often removed during processing and the flesh of intra species is similar in taste and texture; therefore, it can be relatively easy to create adulterated products. Most seafood adulteration is based on supplying the consumer with something different from and inferior to the product expected. Mislabeled and substituted species are the most commonly found adulterated products in the seafood industry. Numerous mislabeled seafood products have been found in the U.S. seafood markets; the majority of these mislabeled seafood products are produced in other countries such as Mexico and China and are labeled as U.S. wild caught seafood. Since there is a high demand for certain seafood species, substituting an inexpensive species for one of higher value is becoming more common.

There are many analytical methods that have been developed to ensure the quality and the safety of food. Most of these analytical methods are designed for the seafood species authentication or identification, the detection of hazardous chemicals and harmful microorganisms in the seafood products.
DNA based methods are widely used for the seafood analysis, primarily for food authentication purposes. DNA based methods for the authentication of seafood species are based on the genetic variations among the species of interest. To detect species-specific genetic variations, DNA is extracted from the target species using appropriate extraction buffers. After DNA extraction, a specific genetic sequence (target DNA fragments) is selected from either nuclear DNA or mitochondrial DNA (mt DNA) for the amplification. Among these DNA sequences, mt DNA is widely utilized because it is appropriate for phylogenetic analysis and universal primers are readily available for the amplification of mt DNA sequences. The commonly employed mt DNA sequences in seafood species authentication are mt cytochrome b (mt cyt b), 12S ribosomal RNA gene, 16S ribosomal RNA gene, and 5S ribosomal RNA.

Polymerase chain reaction (PCR) is used for the amplification of a specific DNA sequence (target DNA) in DNA-based seafood analysis. This technique is conducted by allowing a reaction mixture containing cell lysate, DNA polymerase, nucleotides, and species specific primers (strands of nucleic acid that initiate DNA synthesis) to undergo a number of heating and cooling cycles. In each cycle of synthesis, the number of target DNA molecules used and the number of new molecules produced is doubled. Typically, about 20-50 cycles are required to obtain millions of copies of target DNA fragments. Following DNA extraction and PCR amplification, the resulting DNA fragments (amplicons) are analyzed to verify the presence or absence of species-specific genetic markers. For this purpose, a variety of post-PCR analysis methods such as DNA sequencing, multiplex PCR, restriction fragment length polymorphism (RFLP), forensically informative nucleotide sequencing (FINS), single-stranded conformational polymorphism (SSCP), random amplified polymorphic DNA (RAPD), and amplified fragment length polymorphism (AFLP) techniques have been widely used for the authentication or
identification of seafood species including shrimp, fish, flatfish, mollusks, salmonids, gadoids, scombroids, and eels.7,48-51

DNA barcoding, the identification of species using a standard DNA segment, is being used increasingly for the identification of a broad range of organisms.52-53 In DNA barcoding, a short section of DNA from a standardized region of the genome is extracted, replicated using PCR amplification and sequenced to produce reference sequences or DNA barcodes that act as molecular identification codes for each species profiled.43,54 Typically, a fragment of mitochondrial cytochrome c oxidase I (COI) is used as target gene region for the identification of most of the animal groups. The barcode sequences are saved in The Barcode of Life Data System (BOLD) database which is a reference library of DNA barcodes.55 Fish Barcode of Life Initiative (FISH-BOL) was launched in 2005 with the goal of establishing DNA barcodes for fish species and is also being developed for the identification of crustaceans.56-58

DNA-based studies have focused not only on species identification, but also on population structures, phylogeography and phylogenetic relationships.59-62 In addition to time consuming and labor intensiveness of these methods, some processed seafood products do not contain sufficient variations for interspecies differentiation.63-64

Classical electrophoretic methods such as sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and isoelectric focusing (IEF) are utilized to detect and authenticate seafood species.65 In SDS-PAGE, SDS unfolds the protein structures and coats the protein with a uniform negative charge and then, in the presence of electric field, negatively charged proteins migrate towards the positive electrode through a sieving matrix (polyacrylamide gel) at different rates depending on their molecular weight. SDS-PAGE was used for the discrimination of the pink (Farfantepenaeus duorarum), white (Litopenaeus
setiferus), and rock shrimp (Sicyonia brevirostris) by analyzing the species specific protein patterns.\textsuperscript{66} SDS-PAGE has also been applied to the identification of several commercially important fish species through the separation patterns of proteins with molecular weight lower than 30 000 Da.\textsuperscript{67} Although SDS-PAGE is sensitive and requires little protein sample, it is less effective identifying processed seafood products due to the lack of stability of polypeptide targets in the processed foods.

Isoelectric focusing (IEF) is another electrophoretic method that is based on the separation of proteins on polyacrylamide gel by the use of a pH gradient and the subsequent staining of the species-specific proteins. In a pH gradient and under the influence of an electric field, proteins move to the position in the gradient where its net charge is zero (isoelectric point). The IEF technique is a commonly used classical method for seafood species identification purposes. In IEF studies of shrimp species identification, sarcoplasmic proteins are separated in a narrow acidic pH range and the isoelectric point of the resulting bands is measured for species-specific band patterns.\textsuperscript{68-70} Several commercial fish species have also been identified with IEF, where species are characterized by the specific protein bands.\textsuperscript{67, 71}

In addition to electrophoretic methods for the identification of shrimp and other seafood species, immunological methods such as enzyme-linked immune-sorbent assay (ELISA) and eastern blot have also been applied for the species authentication, specifically for the identification of rock shrimp (S. brevirostris).\textsuperscript{72} These immunological methods are based on antibodies that specifically bind to a target protein. However, immunoassays can be ineffective at differentiating closely related species and require the development of an antibody against the specific protein of interest.
Food consumers are concerned about unwanted chemical substances such as heavy metals in seafood products because of their toxicity and accumulation in biological tissues.\textsuperscript{73-74} Inductively coupled plasma mass spectrometry (ICP-MS) is a well-established and widely utilized technique for the detection of heavy metals such as Hg, As, Cd and Pb in tuna, salmon, shrimp, and other seafood products.\textsuperscript{75-78} In this method, seafood samples are digested in nitric acid and hydrogen peroxide followed by the determination of heavy metals concentration using ICP-MS.

Chromatographic techniques coupled with mass spectrometry are extensively used for the analysis of pesticides and other drug residues in seafood products. Liquid chromatography-mass spectrometry (LC-MS) has been widely employed for the detection of different classes of drug residues such as oxytetracycline, sulfonamides, quinolones, cationic dyes, and toltrazuril sulfone in seafood.\textsuperscript{79-82} In this method, the shrimp meat is extracted with acid and then analytes are separated in the chromatography column prior to analysis the mass spectrometer.\textsuperscript{83} Gas chromatography-mass spectrometry (GC-MS) has been widely used in the analysis of organic pollutants or crude oil components in marine species such as crab, finfish, oysters, and shrimp.\textsuperscript{84-85}

In the past 20 years, the scientific community has seen great advances in different fields due to the development of high throughput methods and analytical instruments. As a result, researchers in food science and nutrition are moving from classical methodologies to more advanced strategies, typically borrowing methods that are well established in medical, pharmacological or biotechnology research.\textsuperscript{86} In this context, foodomics has been defined as a new discipline that studies the food and nutrition domains through the application of advanced omics technologies to improve consumer’s well-being, health, and confidence.\textsuperscript{87-88} The main
concepts of foodomics include genomic, transcriptomic, proteomic, and metabolomic study of foods for compound profiling, authenticity, and biomarkers analysis related to food quality and safety. Due to the large amount of data obtained from these omics approaches, it has been necessary to develop strategies to convert the complex raw data obtained into useful information. Thus, bioinformatics has become also a crucial tool in foodomics.

In proteomics and metabolomics, MS-based techniques play a crucial role in the detection and characterization of proteins, peptides, and metabolites in biological samples. Proteomics is one of the popular omics techniques that is extensively used for the seafood quality and safety. The field of proteomics is defined as the large-scale analysis of proteins expressed by an organism. The entire collection of an organism’s proteins, called its proteome, is expressed from the genes of the organism. Mass spectrometry (MS) has become the primary analysis method in the study of the proteome because of its high throughput, sensitivity and high mass accuracy. In addition, MS techniques have low detection limits and require no prior knowledge of the proteins or peptides to be identified and thus MS represents a powerful tool in proteomics. The results of these MS techniques are spectra with a number of different mass peaks that can be attributed to biomolecules. Identification of the molecules or species of interest can be performed by the bioinformatics treatment of the data obtained from MS analysis. A review of MS methods for the analysis of seafood is discussed below.

1.2 Mass Spectrometry for Biomolecule Based Seafood Analysis

Mass spectrometry (MS) is one of the most powerful techniques in chemical analysis with applications in wide range of scientific disciplines. MS has been incorporated into almost every realm of science since its development in the early 1900s with commercialization of these instruments beginning in the 1950s. Over the years, MS has become a common tool in areas
such as chemistry, biochemistry, physics, environmental science, food science, forensic science, pharmaceutical, and petroleum industries, encompassing a broad scope of target analytes. The versatility and extensive applicability of mass spectrometry make it a valuable tool for obtaining qualitative or quantitative chemical information.

MS is a versatile and robust analytical technique that is capable of providing information regarding the chemical composition and the masses of individual molecules in the sample. Typically, the mass spectrometer consists of three fundamental components: an ion source, a mass analyzer, and a detector. In the ion source, the analyte compounds are converted into gas-phase ions. These ions are then separated according to their mass-to-charge ratio ($m/z$) in mass analyzer and are detected. The resulting mass spectrum is a plot of ion abundance versus its $m/z$ ratio.

Mass spectrometry (MS) based techniques have been extensively utilized in food analysis. The application of MS to large biomolecules has been revolutionized in the past decade with the development of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) techniques. Use of ESI and MALDI for biological mass spectrometry is widespread and the biological applications are continuously growing. The two techniques are complementary; MALDI has a higher tolerance for impurities and the singly charged ions produced in the MALDI process give less complex mass spectra. In the following sections, mass spectrometry techniques that are used for the seafood authentication are thoroughly discussed.

Electrospray ionization (ESI) is a soft ionization technique used in mass spectrometry to produce ions using liquid samples. Briefly, ESI produces charged droplets from an analytic solution at the end of capillary at high voltage. Solvent evaporation causes droplets to shrink to a
point where the repelling coulombic forces are large enough to exceed the surface tension and subsequently droplets breakdown into highly charged smaller droplets.\textsuperscript{98} This process repeats itself several times until highly charged solute molecules remain. Depending on the size of the molecule and its ionizable sites, ESI often results in analytes having wide range of charge states.

Because ESI uses liquid sample introduction and ionization at atmospheric pressure, it is well suited for coupling to liquid separations and thus it is routinely used for MS detection of liquid chromatography (LC)\textsuperscript{99} and capillary electrophoresis (CE).\textsuperscript{100} Moreover, coupling of ESI to these techniques separates the analytes in time and results a much simpler series of mass spectra. Additionally, these separations remove sample impurities.

Electrospray ionization mass spectrometry has been applied in various seafood analysis studies such as species authentication and detection of chemical residues in seafood products. In the ESI MS based proteomic method of seafood species authentication, sarcoplasmic proteins are considered as the target analytes.\textsuperscript{101} Sarcoplasmic proteins are abundant proteins in marine species and majority of these proteins are enzymes participating in cell metabolism. The composition of the sarcoplasmic proteins changes among the species and thus these proteins are extensively used as biomarkers in marine species identification. An ESI MS based approach has been used for shrimp identification studies and it has been proposed that arginine kinase (AK) be used a potential biomarker for shrimp species identification because of the interspecific variability of AK and to its high concentration in shrimp muscle.\textsuperscript{101-102} In this study, protein extracts from shrimp samples were analyzed using 2-dimentional gel electrophoresis and arginine kinase spots at 40 kDa were localized. These arginine kinase spots were excised and digested with trypsin to produce peptides. Further identification and characterization of the differential peptides from the AK in different shrimp species was performed by electrospray
ionization ion trap tandem mass spectrometer, followed by database searching. In addition, sequence analysis was performed to obtain the primary genetic sequence of a species of interest by de novo sequencing.

Commercial fish species such as hakes and grenadiers were also identified by the characterization of different isoforms of nucleoside diphosphate kinase B (NDK B) using LC-MS/MS and ESI-MS/MS. In addition to the ESI MS food authentication studies, ESI MS is also widely used for the detection of antibiotic drug residues in seafood products. Although several commercial shrimp and other seafood species were identified, the ESI MS-based method of species identification requires time consuming steps such as extraction, separation, and digestion of proteins.

Matrix-assisted laser desorption ionization (MALDI) is another soft ionization technique that is used in mass spectrometry to produce gas phase ions from solid samples. In MALDI, the analyte is embedded in a matrix, often a small organic acid with strong absorption at the laser wavelength. A pulsed laser fired on the co-crystallized matrix/analyte mixture; the matrix is responsible for absorbing energy from the laser, isolating analyte molecules, and providing protons for analyte ionization without significant fragmentation of the biological molecules. This process generally produces singly protonated ions, yielding easily interpretable mass spectra. Typically, the molecular masses of the gas-phase ions are determined with a time-of-flight mass analyzer. Sample preparation and selection of matrix compound is crucial in MALDI. The required characteristics for a molecule to be considered as MALDI matrix include: a strong light absorption at the wavelength of the laser, soluble in a solvent that is compatible with the analyte and promote co-crystallization, be vacuum stable, and influence the co-desorption and ionization of the analyte upon absorbing laser light. There are several MALDI
matrices available commercially, but the selection of particular matrices is depends on various factors.95

The development of MALDI allows mass spectrometry detection of intact biomolecules and because of its high tolerance of impurities and ease of sample preparation; MALDI MS has been used extensively in identification and characterization of proteins, synthetic polymers, oligonucleotides, polysaccharides and lipids. Based on the type of study, various MALDI MS approaches such as protein profiling, peptide mass fingerprinting, and peptide sequence tags have been developed for species identification purpose. Among the MALDI MS methodologies, peptide mass fingerprinting (PMF) is a commonly used MS-based proteomic approach for the discrimination of seafood species.14,107 In PMF based identification of shrimp authentication studies, 2-DE analysis of the sarcoplasmic proteome shows interspecific variability in the isoelectric point (pI) of arginine kinase (AK) proteins and thus AK spots are selected as a potential molecular marker and subjected to tryptic digestion to cleave this protein into peptides. Then the digested sample is analyzed with MALDI MS to get the masses of peptides. These masses are then compared to a database containing AK protein sequences.

Commercial fish species such as hake and grenadier were also identified by analyzing the species-specific parvalbumin protein through a MALDI MS-based proteomic approach.13,108 In another fish identification study, a MALDI-TOF MS molecular profiling approach was used, in which species specific mass spectrometric profiles from 25 different fish species were observed and these specific proteins were then further identified with a PMF strategy.14

In addition to the time-consuming sample preparation steps such as proteins separation and digestion, PMF based methods of identification also require isolation of single proteins. If
the protein mixture contains more than one protein, it typically requires the additional use of MS/MS based protein identification to achieve species identification.

Another popular MALDI MS approach is the mass fingerprinting method, which is a well-established analytical technique for the identification of microorganisms. With regard to analysis of whole-cell microorganisms, MALDI has been extensively used for the identification and classification of microorganisms such as bacteria, virus, fungus, and yeasts in food quality control and taxonomical research.\textsuperscript{11, 109} MALDI MS identification of whole-cell microorganisms is a fast and reliable technique for obtaining data without the need for extensive sample preparations such as digestion and analytical separations.\textsuperscript{15} Once the sample is deposited directly on the MALDI target, the sample can be analyzed following the addition of MALDI matrix to the collected sample. Mass spectra generated from microbial samples are easy to interpret and are typically attribute to proteins. For mass spectra fingerprinting, a database is constructed (reference mass spectral library) from target bacteria containing peak location ($m/z$), peak intensity, and frequency of occurrence.\textsuperscript{110-111} With this information, the algorithm is then applied to mass spectra from unknown samples for comparison. In the end, a degree of association is calculated to provide information for a match.\textsuperscript{110}

Among the MALDI MS methods that are currently used for the species identification, the MALDI MS fingerprinting approach has the advantage that it can analyze whole-cells and requires minimal sample preparation. Moreover, the mass spectral databases can be expanded to other species by developing reference spectra of species of interest and thus subsequent identification of species is possible. Although several studies have used MALDI MS based methods for shrimp and other seafood species identification, the MALDI MS fingerprinting approach has seen limitedly used for marine species identification.
Matrix-assisted laser desorption ionization mass spectrometry has the capacity to provide valuable information in biomarker discovery, biomolecule characterization, and in identification of wide variety of species. The generation of large volumes of data requires bioinformatics tools for processing the information into easily understandable results. Bioinformatics is the utilization of computation and analysis to interpret biological data. Bioinformatic tools, such as databases, are used to manage large data sets for bioanalysis in molecule identification as well as in structure functions relationships. There are different types of databases commercially available for the identification of biomolecules or microorganisms. For example, protein databases contain sequence information for biomolecules and fingerprint databases contain reference mass spectra of different microorganisms. In this work, we have used a fingerprint database for the identification of shrimp species.

1.3 Shrimp

Seafood products include a wide variety of species with a significant importance for the food industry in the United States.¹⁹ Among the most valuable of all seafood species are shrimp, which are crustaceans belonging to the order Decapoda and suborder Dendrobranchiata.¹¹² Of the Dendrobranchiata, the Penaeidae family are one of the most important economic resource in U.S. fisheries.¹¹³ Shrimp typically contains full carapace or head shield, five pairs of walking legs, and a tail.¹¹³ Like other anthropods, shrimp have an exoskeleton. Although some species show significant difference in their morphological characteristics, many of the species have similar external features.¹¹⁴ Most of the commercially important species are closely related and share similar external features.

Shrimp are produced using two methods: wild-caught and farm-raised. The wild-caught process is considered more natural because shrimp are harvested from their natural habitats and,
unlike the farm-raised process, no chemicals are used.\textsuperscript{115} The farm-raised process involves producing shrimp in large man-made ponds using aquaculture. Most of the shrimp produced in the U.S. are wild-caught shrimp. The shrimp fisheries in the U.S. are the most valuable and fisheries provides thousands of jobs and broader economy.\textsuperscript{116} United States maintains a large wild shrimp fishery and they fall into two groups: warm water shrimp and cold water shrimp.\textsuperscript{117}

All of the U.S. Gulf Coast and several South Atlantic states have warm water shrimp fisheries.\textsuperscript{118} The commercially important shrimp species fished in the Gulf of Mexico and South Atlantic are brown shrimp (\textit{Farfantepeneaus aztecus}); white shrimp (\textit{Litopenaeus setiferus}); and pink shrimp (\textit{Farfantepeneaus duorarum}).\textsuperscript{114} Shrimp that are belongs to the genus Penaeus are also called as Penaeid shrimp and are the most valuable and commercially dominant shrimp in warm water. Together, the brown and white Penaeids comprise more than 95\% of shrimp landings in the Gulf of Mexico and the majority of shrimp landings in the South Atlantic.\textsuperscript{119} The other warm water shrimp are rock shrimp (\textit{Sicyonia brevirostris}), royal red shrimp (\textit{Pleoticus robustus}), and seabob shrimp (\textit{Xiphopenaeus kroyeri}) that are also fished for commercial purpose in the U.S. Gulf and South Atlantic coast.\textsuperscript{120} Although U.S. fisherman catch warm water shrimp from both the Gulf of Mexico and South Atlantic coast, more than 90\% of shrimp landings come from the Gulf of Mexico. In the Gulf region, Louisiana and Texas account for more than 80\% of the total U.S. wild caught shrimp production.\textsuperscript{121}

Besides the warm water shrimp, cold-water shrimp are also commercially important. The dominant shrimp species in cold water are from the Pandalid family: spot shrimp (\textit{Pandalus platyceros}); sidestripe shrimp (\textit{Pandalopsis disparr}); coonstriped shrimp (\textit{Pandalus hypsinotus}); Northern shrimp (\textit{Pandalus borealis}); and humpy shrimp (\textit{Pandalus goniurus}).\textsuperscript{122} Among the Pandalids, spot shrimp and side stripe shrimp are the commercially dominant species fished on
the Pacific Coast of Alaska. Although, the majority of the U.S. shrimp landings come from warm waters, cold-water shrimp account for more than 10% of total shrimp landings.

United States has a strong domestic shrimp production history but as a high seafood-consuming nation, the domestic landings are not sufficient for the needs of consumers. As a result, 86% of the shrimp consumed in the U.S. is imported.\textsuperscript{123} The leading suppliers of shrimp to the U.S. market are Thailand, Mexico, Ecuador, Indonesia, and China. The most commonly imported shrimp species from these countries are Penaeid shrimp: \textit{Penaeus monodon} and \textit{Penaeus vannamei}.\textsuperscript{124} While the majority of domestic shrimp are wild caught, most of these imported shrimp are farm-raised.

1.4 Research Objectives

The overall objective of this research was to develop a protocol for the identification of shrimp using MALDI MS fingerprinting. To achieve this goal, six commercially important domestic shrimp species of U.S. were selected and prepared for MALDI MS analysis. The resulting spectra were uploaded into commercial mass spectrometry software for the creation of a mass spectrum reference database for the six species analyzed. Finally, the reference database was tested using unknown shrimp samples from the six species. A detailed descriptions of experimental setups used in this research are discussed in Chapter 2. The results of this research are discussed in Chapter 3.
CHAPTER 2. EXPERIMENTAL

The goal of the research described in this dissertation was to develop a matrix-assisted laser desorption ionization mass spectrometry fingerprinting method for the species identification of shrimp. In this chapter, a detailed description of instrument and experimental setup including sample preparation and MALDI MS analysis used in this study is presented. A description of the shrimp mass spectrum database and is also presented.

2.1 Bruker UltrafleXtreme Tandem Time-of-Flight Mass Spectrometer

The mass spectrometer used was a Bruker UltrafleXtreme MALDI-TOF/TOF mass spectrometer. This instrument is capable of both linear and reflectron modes of operation. This system has a 4 GHz digitizer, a resolution of 40,000 from 700 to 5000 \( m/z \), and a 1 ppm and 5 ppm mass accuracy for internal and external calibration, respectively. It is equipped with frequency tripled Nd:YAG 355 nm solid-state laser with a homogenized modulated beam (Smartbeam II) laser operating at 1Hz. The MALDI ion source accepts several targets or adapter plates of the same size and shape. The MALDI target most commonly used with the MALDI-TOF/TOF has 384 possible deposition spots for high-throughput analysis. The MALDI-TOF/TOF operates at a maximum acceleration potential of 25 kV, as shown in Figure 2-1. All the ions have the same kinetic energy and the flight time depends on the mass of the analyte. This instrument is equipped with time-of-flight mass analyzer, which measures the ions mass to charge ratio via time measurement.
2.2 Biotyper

MALDI BioTyper is software that utilizes a fingerprint matching method for the identification of microorganisms. This microbial database has a total of 3290 spectra corresponding to a number of microorganisms including yeasts, fungi, viruses, and bacteria. Based on the mass spectra of individual species, this database has a unique average reference spectrum, called the main spectrum profile (MSP), for each microorganism. The MSPs contain a list of mass peaks and their abundance for each microorganism. Each MSP is composed of an average of 20 independent peak lists that contain peak position ($m/z$), peak intensities, and peak occurrence among the set of 20 spectra for the most prevalent peaks. Detailed information on creating a MSP for the species of interest is discussed in Chapter 3. After experimental mass spectra are imported into the database and before searching for identification, the database uses an alignment function to calibrate the unknown spectrum against the reference spectrum to reduce the number of mass deviations. Then using an integrated pattern-matching algorithm, which considers $m/z$ values, peak intensity, and how often each peak appears compared to
reference spectrum, the experimental spectrum is compared to the reference spectrum. Finally, a score table of matched reference spectra is displayed with score (log) values ranging from 0 to 3.

For identity scoring, the algorithm implemented in the database counts mass signals in experimental spectra that matched with reference spectra. Furthermore, the algorithm also correlates signal intensities of matched signals. Together, three scores obtained from such a procedure are multiplied and normalized to a value of 1000 and then converted in its common logarithm. Based upon score values, it can be determined whether identification is correct at species or genus level. A detailed description of identification scores is displayed in Table 2-1.

Table 2-1. Description of identification scores

<table>
<thead>
<tr>
<th>Range</th>
<th>Description</th>
<th>Symbols</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.30 — 3.00</td>
<td>Highly probable species identification</td>
<td>(+++)</td>
<td>green</td>
</tr>
<tr>
<td>2.00 — 2.99</td>
<td>Secure genus identification, probable species identification</td>
<td>(++)</td>
<td>green</td>
</tr>
<tr>
<td>1.70 — 1.99</td>
<td>Probable genus identification</td>
<td>(+)</td>
<td>yellow</td>
</tr>
<tr>
<td>0.00 — 1.69</td>
<td>No reliable identification</td>
<td>(-)</td>
<td>red</td>
</tr>
</tbody>
</table>

2.3 Sample Collection

Six different shrimp species (Table 2-2) were considered in this study: two penaeidae, *Litopenaeus setiferus* and *Farfantepenaeus aztecs* (from Louisiana); two pandalidae, *Pandalopsis dispar* and *Pandalus platyceros* (Alaska) one sicyoniidae, *Sicyonia brevirostris* (Florida Atlantic Coast) and one solenoceridae, *Pleoticus robustus* (Florida Atlantic Coast). Specimens were collected using extractive fishing in the Gulf of Mexico, North Pacific Coast,
and North Atlantic Coast and shipped to our location on ice or, for Louisiana shrimp, obtained locally. Imported penaeus species were Chinese white shrimp (*Fenneropenaeus chinensis*), black tiger shrimp (*Penaeus monodon*) farmed in Thailand, and whiteleg shrimp (*Litopenaeus vannamei*) farmed in Ecuador that were obtained from a local grocery store.

**Table 2-2.** Shrimp species used to establish the reference database for MALDI-TOF MS based species identification

<table>
<thead>
<tr>
<th>Common Name</th>
<th>Scientific Name</th>
<th>Place of origin (USA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern White Shrimp</td>
<td><em>Litopenaeus setiferus</em></td>
<td>LA, FL, TX</td>
</tr>
<tr>
<td>Northern Brown Shrimp</td>
<td><em>Farfantepenaeus azticus</em></td>
<td>LA, FL, TX</td>
</tr>
<tr>
<td>Royal Reds</td>
<td><em>Pleoticus robustus</em></td>
<td>FL</td>
</tr>
<tr>
<td>Rock Shrimp</td>
<td><em>Sicyonia brevirostris</em></td>
<td>FL</td>
</tr>
<tr>
<td>Side Stripe Shrimp</td>
<td><em>Pandalopsis dispar</em></td>
<td>AK</td>
</tr>
<tr>
<td>Spot Shrimp</td>
<td><em>Pandalus platyceros</em></td>
<td>AK</td>
</tr>
</tbody>
</table>

LA = Louisiana (Gulf of Mexico), FL = Florida (South Atlantic coastal), TX = Texas (Gulf of Mexico), AK = Alaska (North Pacific coast).

All of the samples were obtained unfrozen on ice and stored at −80 °C upon receipt. Samples were classified in their respective taxa based on their anatomical external features.

### 2.4 Sample Preparation and MALDI MS Analysis

Samples of 1 g of shrimp skeletal muscle were obtained by dissecting a shrimp and then homogenizing at room temperature in 2 mL of nanopure water using a mortar and pestle. The homogenate was then centrifuged at 13,000 rpm for 20 min. The supernatant was removed and further purified using desalting pipette tips (OMIX C18, 10 µl, Agilent Technologies, Santa Clara, CA, US) according to the manufacturer’s protocol. A 4 µl volume of the desalted sample
was directly pipetted into 4 µl of 30 mg/ml 2, 5-dihydroxy benzoic acid (Sigma-Aldrich, St Louis, MO) matrix solution in 1:1 (v/v) ethanol: 0.1 % TFA. A 1 µl aliquot of the analyte/matrix solution was spotted onto a stainless steel MALDI target and allowed to dry at room temperature. This sample preparation protocol was used for all experiments in this study.

MALDI-TOF spectra were recorded in positive ion reflectron mode with an accelerating voltage of 25 kV and analyzed in the mass range of 1,000 – 5,000 Da. The spectra were acquired after calibration of the instrument with a peptide standard. A minimum of 500 laser shots per sample was used to generate each mass spectrum. MALDI BioTyper 2.0 software was used for the mass spectra fingerprinting.

2.5 Construction of Reference Database and Identification of Unknown Shrimp

An average of 20 spectra is recommended to make one profile mass spectrum. For this work, a collection of 20 acceptable spectra from each species was used to generate a single library mass spectrum representative of that species. Mass resolution was greater than 10,000 for the majority of the mass spectra peaks and mass accuracy was greater than 30 ppm. Five individual shrimp were used to generate the 20 mass spectra. The database stores a list of peak positions and their relative intensities for all of the library entries.

MALDI mass spectra were obtained for the unknown shrimp using 3 shrimp per species, 4 MALDI target spots per species and one mass spectrum per spot. The mass spectra were processed in the mass range from 1000 – 5000 m/z using the fingerprinting software and the shrimp database. The software returned the top species matches in ranked order with confidence scores ranging between 0 and 3.
2.6 Reagents and Standards

Solvents used in this work include: acetonitrile (ACN), acetone, methanol, and ethanol from Fisher Scientific. The MALDI matrices used in this work were α-cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxy benzoic acid (DHB), and 3,5-dimethoxy-4-hydroxycinnamic acid (sinapic acid) and were purchased from Sigma-Aldrich, Saint Louis, MO. Typical matrix solutions consisted of 20-30 mg/mL of matrix. In most instances the solvent used was a 50/50 ethanol/water containing 0.1% TFA solution.

Calibration Standards

Mass calibration of MALDI TOF MS was achieved by using Peptide Calibration Standard II from Bruker. This peptide standard mixture contains nine peptides in the mass range between 700 and 3500 Da as shown in Table 2-3. The masses for the peptide calibration mixture can be observed as protonated monoisotopic masses. Monoisotopic mass is the mass of an ion or molecule obtained using the mass of the most abundant isotope of each element. The peptide calibration standards are dry peptides in a microcentrifuge tube. This peptide mixture was dissolved in 125 µL of 0.1% trifluoro acetic acid (TFA) in nanopure water and mixed in the micro centrifuge tube. A volume of 1 µL of this peptide standard was mixed with 1 µL of CHCA matrix solution and deposited on a MALDI target and allowed to dry at room temperature. The matrix solution was prepared by dissolving CHCA in an organic-acidic solvent mixture of 1:1 (v/v) ACN: 0.1% TFA in nanopure water up to saturation at room temperature and then centrifuged. The remaining peptides were stored at less than -20°C.
Table 2-3. List of peptide calibration standard II peptide masses

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bradykinin 1-7</td>
<td>757.429</td>
<td>757.86</td>
</tr>
<tr>
<td>Angiotensin II</td>
<td>1046.556</td>
<td>1047.19</td>
</tr>
<tr>
<td>Angiotensin I</td>
<td>1296.703</td>
<td>1297.49</td>
</tr>
<tr>
<td>Substance P</td>
<td>1347.729</td>
<td>1348.64</td>
</tr>
<tr>
<td>Bombesin</td>
<td>1619.798</td>
<td>1620.86</td>
</tr>
<tr>
<td>Renin Substrate</td>
<td>1758.932</td>
<td>1760.03</td>
</tr>
<tr>
<td>ACTH clip 1-17</td>
<td>2093.100</td>
<td>2094.43</td>
</tr>
<tr>
<td>ACTH clip 18-39</td>
<td>2465.213</td>
<td>2466.68</td>
</tr>
<tr>
<td>Somatostatin 28</td>
<td>3147.450</td>
<td>3149.57</td>
</tr>
</tbody>
</table>

The peptide calibration masses can be observed as a protonated monoisotopic protonated molecule peaks [M+H]$^+$ or average protonated molecule peaks. A MALDI mass spectrum of the peptide calibration mixture is shown in Figure 2-2.
Figure 2-2. MALDI mass spectrum of Peptide Calibration Standard II
CHAPTER 3. MATRIX-ASSISTED LASER DESORPTION IONIZATION MASS SPECTROMETRY FOR IDENTIFICATION OF SHRIMP

3.1 Introduction

In this chapter, the identification of shrimp species using a MALDI MS fingerprinting approach is described. As mentioned in Chapter 1, shrimp is a popular seafood item in the U.S. and thus shrimp species identification is important to assure seafood quality and safety. Six commercially important shrimp species, *Litopenaeus setiferus*, *Pleoticus robustus*, *Sicyonia brevirostris*, *Pandalopsis dispar*, *Pandalus platyceros*, and *Farfantepenaeus azticus*, were used in this work. After the analysis of shrimp samples by MALDI-TOF/TOF mass spectrometry, the resulting mass spectra were imported into the fingerprinting software and a shrimp reference database was created. The unknown shrimp samples were analyzed and the results were used to test the database accuracy. Three imported shrimp species, *Fenneropenaeus chinensis*, *Penaeus monodon*, *Litopenaeus vannamei*, were used as negative controls.

3.2 Experimental

A detailed discussion of experimental setup and analysis of shrimp samples were presented in Chapter 2. Briefly, approximately 1 g of skeletal muscle of shrimp was obtained by dissecting a shrimp which was homogenized and then centrifuged to produce a supernatant. Due to the fact that seafood samples give salt interference, the samples were desalted using pipette tips, which have C_{18} reverse phase material packed in their tips. The desalted samples were then mixed with MALDI matrix (30 mg/mL 2,5-dihydroxy benzoic acid) in 1:1 ratio for the MALDI-TOF/TOF analysis. The resulting MS data were then used for the creation of the shrimp reference database. An average of 20 spectra for each shrimp species was used to create the shrimp reference spectra in the database. Shrimp samples from the six species used to generate
the library mass spectra were analyzed and used to evaluate the accuracy of the shrimp reference library.

3.3 Results

The initial step in the creation of the mass spectrum database was the development of a sample preparation protocol for the removal of salt and other impurities from the seafood samples. Because shrimp samples contain significant amounts of salt that can cause interferences in the mass spectra, purification of the sample was critical to obtaining reproducible mass spectra.

Figure 3-1. MALDI mass spectra of white shrimp (*Litopenaeus setiferus*) a) without desalting and b) with desalting using a C$_{18}$ reverse phase pipette tip.
Figure 3-1 depicts example mass spectra with and without desalting for white shrimp (*Litopenaeus setiferus*). When the supernatant of the shrimp muscle homogenate was mixed with the matrix solution and spotted in 1 μl volumes at 8 positions on stainless steel MALDI target, only two of the 8 spots produced acceptable mass spectra. After desalting using C$_{18}$ reverse phase pipette tips, all spots produced acceptable mass spectra with an increase in signal typically an order of magnitude or greater.

**Figure 3-2.** MALDI-TOF mass spectra obtained from the analysis of (a) *Litopenaeus setiferus*, (b) *Pandalus platyceros*, (c) *Pandalopsis dispar*, (d) *Pleoticus robustus*, (e) *Sicyonia brevirostris*, and *Farfantepenaeus azteicus* (f) showing the differences among the different shrimp species.
Representative mass spectra obtained from the six shrimp species are shown in Figure 3-1. Peaks indicative of the species were observed in the range from 1000 to 5000 m/z. In this range, all the six species showed significant differences in their mass spectra. Spectra from species in the same family appeared qualitatively similar to each other compared to those in different families. Our sample preparation protocol was found to be reproducible and provide enough peaks in MALDI spectra to allow us to build a reference database containing the characteristic mass fingerprinting profiles of representative shrimp species.

To create a single library mass spectrum, 20 individual mass spectra from each of the six species were averaged. A database containing the reference peak lists of individual shrimp species was created.

![Reference Spectrum: P. robustus](image_url)

**Figure 3-3.** Graphical output from a *P. robustus* reference spectrum
For example, in the graphical output of the *Pleoticus robustus* reference spectrum displayed in Figure 3-3, the blue lines represent averaged peak masses with respect to their relative intensity and the red lines show occurrence of the peak masses contributing in that MSP.

A dendrogram is a diagram representing a hierarchy of categories based on similarity between species and is typically used in biological taxonomy. The dendrogram based on a cluster analysis of the mass spectra database is shown in Figure 3-4. The x-axis represents the distance level: a lower distance level value corresponds to greater similarity. The dendrogram indicates that the mass spectra give a good indication of species similarity.

**Figure 3-4.** Dendrogram of six shrimp MSPs

For example, the members of the Penaeidae family (*Litopenaeus setiferus* and *Farfantepenaeus aztecus*), and the Pandalidae family (*Pandalopsis dispar* and *Pandalus platyceros*) have a low distance level within the family. The distance level between the species *Sicyonia brevirostris* and *Pleoticus robustus* is greater, which is consistent with both species being classified under different families.
To test the shrimp database, 74 shrimp samples were analyzed using shrimp from the six species to generate the library mass spectra. Multiple samples for each species were prepared using the desalting protocol described above. For the identification of the unknown samples, the MALDI TOF mass spectrum that was obtained after analysis was compared with the shrimp database. Three shrimp were analyzed for each species and four spots and four mass spectra were obtained for each shrimp.

Table 3-1. Species identification results for the unknown shrimp samples

<table>
<thead>
<tr>
<th>Species (no. of samples tested)</th>
<th>Number of samples with the indicated log score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 – 1.99</td>
</tr>
<tr>
<td><em>Litopenaeus setiferus</em> (12)</td>
<td></td>
</tr>
<tr>
<td><em>Farfantepenaeus azticus</em> (14)</td>
<td></td>
</tr>
<tr>
<td><em>Pleoticus robustus</em> (12)</td>
<td></td>
</tr>
<tr>
<td><em>Sicyonia brevirostris</em> (12)</td>
<td></td>
</tr>
<tr>
<td><em>Pandalopsis dispar</em> (10)</td>
<td></td>
</tr>
<tr>
<td><em>Pandalus platyceros</em> (14)</td>
<td>1</td>
</tr>
<tr>
<td>Total (74)</td>
<td>2</td>
</tr>
</tbody>
</table>

The log score values of 74 analyzed shrimp samples shown in Table 3-1. Out of 74 shrimp samples analyzed for species identification, 72 samples were correctly identified at the species level, and two samples were identified at the genus level. The *F. azticus* and *L. setiferus* are the most closely related (see Figure 3) and this led to a genus rather than species level fit for one of the 14 mass spectra for *F. azticus*. Similarly, the close relationship between the *P. platyceros* and *P. dispar* led to a genus rather than a species match for one of the 14 mass spectra for *P.
platyceros. Low signal for some of the peaks in the m/z region between 1500 and 4000 m/z limited the score for the two poorly matching mass spectra. Based on the identification results, the shrimp database shows 97% accuracy.

The match of individual species to the species library is indicated in Table 3-2. The table contains the scores obtained when a shrimp from a particular species was matched against the database. The standard deviation is indicated for the replicates of the mass spectra that yielded the correct species match. Species that belong to same family have higher scores compared to species different families, although none give a species level match.

<table>
<thead>
<tr>
<th></th>
<th>F. azticus</th>
<th>L. setiferus</th>
<th>P. robustus</th>
<th>P. platyceros</th>
<th>P. dispar</th>
<th>S. brevirostrris</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. azticus</td>
<td>2.7(2)</td>
<td>1.8</td>
<td>0.8</td>
<td>1</td>
<td>0.9</td>
<td>0.6</td>
</tr>
<tr>
<td>L. setiferus</td>
<td>1.9</td>
<td>2.7(1)</td>
<td>1.1</td>
<td>0.8</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td>P. robustus</td>
<td>1.1</td>
<td>0.9</td>
<td>2.7(1)</td>
<td>1.2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>P. platyceros</td>
<td>0.7</td>
<td>1.1</td>
<td>1.1</td>
<td>2.6(1)</td>
<td>1.8</td>
<td>0.8</td>
</tr>
<tr>
<td>P. dispar</td>
<td>0.4</td>
<td>0.9</td>
<td>0.8</td>
<td>1.8</td>
<td>2.7(1)</td>
<td>0.6</td>
</tr>
<tr>
<td>S. brevirostrris</td>
<td>0.8</td>
<td>1.1</td>
<td>0.5</td>
<td>0.6</td>
<td>1.1</td>
<td>2.7(1)</td>
</tr>
</tbody>
</table>

The *F. azticus* and *L. setiferus* are the most closely related (see Figure 3) and this led to a genus rather than species level fit for one of the 14 mass spectra for *F. azticus*. Similarly, the close relationship between the *P. platyceros* and *P. dispar* led to a genus rather than a species match for one of the 14 mass spectra for *P. platyceros*. Low signal for some of the peaks in the m/z region between 1500 and 4000 m/z limited the score for the two poorly matching mass spectra.

A heat map was constructed (Figure 3-5) to visualize the database hit values. Scores in the Table 3-2 were color-coded respectively to generate a heat map. In the heat map, for the
same species, the diagonal shows highest matching values. Species that belong to same family shows more similarity compared to species that are from different family.

![Image of a species matching chart](image.png)

**Figure 3-5.** The match of individual species to the species library

Among six commercial shrimp species, the penaeidae *Litopenaeus setiferus* (Northern White shrimp) and *Farfantepenaeus aztecus* (Northern Brown shrimp) are widely distributed and fished commercially in a number of coastal regions. The shrimp database (which was generated using Louisiana penaeidae shrimp) was evaluated against penaeidae that were collected in Texas and Florida coastal waters. The identification results are displayed in Table 3-3. Out of 12 samples, 11 samples were correctly identified at the species level and one sample was identified at the genus level.
Table 3-3. Identification test for *Litopenaeus setiferus* and *Farfantepenaeus azticus* from different regions, different sizes and fresh/frozen, *F. azticus* (fresh vs. frozen, small vs. jumbo)

<table>
<thead>
<tr>
<th>Log score</th>
<th><em>L. setiferus</em> and <em>F. azticus</em> from FL, TX</th>
<th><em>F. azticus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.30 – 3.0</td>
<td>6/6</td>
<td>5/6</td>
</tr>
<tr>
<td>2.00 – 2.29</td>
<td>1/6</td>
<td></td>
</tr>
<tr>
<td>0 – 1.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The size and fresh/frozen state of the shrimp was evaluated using *Farfantepenaeus azticus* shrimp. Small (about 9 g) and jumbo (about 23 g) were used and both fresh and frozen (small). The results are also shown in Table 3-3. All *F. azticus* samples (4 each for fresh, frozen, small, and jumbo) were correctly identified to species level irrespective of their size and storage conditions.

To test the robustness of the database against false positive identification, we tested three imported shrimp species from Thailand, Indonesia, and China. Samples were analyzed using the protocol that was used for the domestic species and the resulting mass spectra were analyzed using the database. The identification results are shown in Table 3-4. None of the samples were in the reliable identification range (all scores <1.7), giving further confidence for correct identification and lack of false positives.
Table 3-4. Species identification results for the imported shrimp samples

<table>
<thead>
<tr>
<th>Species (no. of samples tested)</th>
<th>Number of samples with the indicated log score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 – 1.69</td>
</tr>
<tr>
<td><em>Fenneropenaeus chinensis</em> (3)</td>
<td>3</td>
</tr>
<tr>
<td><em>Penaeus monodon</em> (3)</td>
<td>3</td>
</tr>
<tr>
<td><em>Litopenaeus vannamei</em> (3)</td>
<td>3</td>
</tr>
<tr>
<td>Total samples analyzed (9)</td>
<td>9</td>
</tr>
</tbody>
</table>
CHAPTER 4. CONCLUSION AND FUTURE DIRECTION

The application of MALDI-TOF MS identification of shrimp using commercial fingerprinting software was accomplished using sample purification to remove salt from the tissue samples. After using salt removal pipet tips, the resulting spectra were more reproducible. The mass spectra show significant difference between all six species analyzed. The species from the same family have more common mass peaks than the species that belong to different families. A library of mass spectra was constructed from six shrimp species that are commercially important in the United States. The dendrogram that was generated from the constructed reference mass spectra clearly demonstrates the differences among constructed mass spectra. Except for two samples that were identified at genus level, the other 72 unknown samples were identified at the species level. The library identified shrimp to the species level with 97% accuracy. Shrimp from different geographical regions could also be identified at the species level. Negative controls using imported shrimp species were identified.

In the future, few other commercially important U.S. domestic shrimp species such as pink shrimp (*Farfantepenaeus duorarum*) can be included in the shrimp spectral library. In addition, imported shrimp species can also be added to the database. Moreover, MS/MS studies of species-specific mass peaks can provide the valuable information on potential biomarkers identification and characterization. Identifying shrimp from oil contaminated water or finding which part of the shrimp body accumulates oil dispersants can be useful for seafood safety purposes. Also, extending the shrimp library to other seafood species such as fish will be useful for the rapid seafood authentication and identification studies.
REFERENCES


3. Buck, E. H. *Seafood Marketing: Combating Fraud and Deception*; Congressional Research Service, **2010**.


VITA

Venkatesulu Salla was born in Chittoor, A.P, India. He graduated in 2006 with a Bachelor of Science degree in Pharmacy at Jawaharlal Nehru Technological University. After his bachelor’s education, in 2008 he enrolled at McNeese State University located in Lake Charles, Louisiana, where he received Master of Science degree in chemistry in 2010. In fall 2010, he enrolled at Louisiana State University, Baton Rouge, Louisiana, and expected to receive Master of Science degree in chemistry in summer 2013. His research advisor at Louisiana State University was Dr. Kermit K. Murray and submitted a research article for publication in Analytica Chimica Acta journal. During his education at Louisiana State University, Salla gave two poster presentations in American Chemical Society conferences. He was a member of American Chemical Society for Mass Spectrometry and the American Chemical Society.