2008

Matrix-assisted laser desorption/ionization mass spectrometry for the analysis of collected bioaerosols

Alton John Dugas
Louisiana State University and Agricultural and Mechanical College

MATRIX-ASSISTED LASER DESORPTION/IONIZATION MASS SPECTROMETRY FOR THE ANALYSIS OF COLLECTED BIOAEROSOLS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

In

The Department of Chemistry

By
Alton John Dugas Jr.
B.S. Louisiana State University, 1995
M.S. Louisiana State University, 1998
May 2008

DEDICATION

This work is dedicated to all the men and women of the US Armed Forces who made the ultimate sacrifice while serving in the defense of our Nation.

ACKNOWLEDGEMENTS

This dissertation could not have been possible without the assistance and encouragement of a great deal of people, especially God the Almighty, who has given me the strength to persevere.

To my dearest wife Kimberly, my daughter Grace and son Trae, I am so blessed that you are all in my life. I cherish each moment we share and could not imagine my life without the joy you bring. To my Mom and Dad, you have always believed in me. I could not have asked for better parents.

To my Advisor, Dr. Murray, I thank you for your counsel and faith in me throughout my studies at Louisiana State University. You are greatly appreciated. I owe you my sincerest gratitude. I would like to thank the United State Air Force for giving me the opportunity to pursue my Doctorate and granting me the time needed to complete this degree.

To my committee members, Dr. Grover Waldrop, Dr. Doug Gilman, Dr. Jayne Garno, and Dr. Michael Tolocka, I thank you all for your gifts of wisdom and guidance. Special thanks to Dr. Bill Wischusen and Anne Jolissaint for their assistance in microbiology and the loan of equipment. Dr. Azeem Hasan for allowing me the freedom to use the mass spectrometry facility equipment. Without this freedom, none of this work would have been possible. Dr. Yohannes Rezenom for his friendship and direction. I want to thank the Socolosky Microscopy Facility for providing microscopy services and Louisiana State University for being a huge part of my life

To the Murray Research Group both past and present: Mark, Damien, Jae-kuk, Jerrell, Xing, Jeonghoon, Fan, Lancia, Jaye, Jianan, Udaya, and Nicole. I will miss you all very much. Good luck to you in all your endeavors.

Finally, I want to thank all my family and friends. It was good to be home in Louisiana.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Name
ACN	Acetonitrile
AI	Andersen Impactor
APS	Aerosol Particle Sizer
Ar	Argon
BAMS	Bioaerosol Mass Spectrometry
BEC	Bioaerosol Exposure Chamber
CA	Caffeic Acid
CE	Collision Energy
CHCA	α-cyano-4-hydroxycinnamic acid
CI	Cyclone Impactor
CID	Collision Induced Dissociation
CNBr	Cyanogen Bromide
Da	Dalton
DE	Delayed Extraction
DESI	Desorption Electrospray Ionization
DHB	2,5-dihydroxybenzoic acid
DNA	Deoxyribonucleic Acid
EBI	European Bioinformatics Institute
ELISA	Enzyme-Linked ImmunoSorbent Assay
EMBL	European Molecular Biology Laboratory
ESI	Electrospray Ionization

eV	Electron Volt
ExPASy	Expert Protein Analysis System
FA	Ferulic Acid or 4-hydroxy-3-methoxycinnamic acid
FAME	Fatty Acid Methyl Esters
FDA	Food and Drug Administration
FS	Filter System
GC	Gas Chromatography
HABA	2[4-hydroxyphenlyazo]benzoic acid
HPA	3-hydroxypicolinic acid
IR	Infrared
LADESI	Laser Assisted Desorption Electrospray Ionization
LAL	*Limulus* amoebocyte lysate
LAMMA	Laser Microtarget mass analysis
LC	Liquid Chromatography
LDI	Laser Desorption/Ionization
LNNL	Lawrence Livermore National Laboratory
lpm	Liters Per Minute
LPS	Lipopolysaccharide
KE	Kinetic Energy
Hz	Hertz
MALDI	Matrix-Assisted Laser Desorption/Ionization
MCP	Microchannel Plate
MMD	Mean Median Diameter

MS	Mass Spectrometry
NCBI	National Center for Biotechnology Information
oMALDI	Orthogonal MALDI
PCR	Polymerase Chain Reaction
PG	Peptidoglycan
PMF	Peptide Mass Fingerprinting
PRISM	Progressive Reductive Iterative Search Mapping
Py	Pyroylsis
QTOF	Quadrupole-Time-of-Flight
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
SA	Sinapic Acid or 3,5-dimethoxy-4-hydroxycinnamic acid
SIB	Swiss Institute of Bioinformatics
SRS	Sequence Retrieval System
Stk	Stokes Number
TDC	Time-to-digital converter
TFA	Trifluoroacetic acid
TOF	Time-of-Flight
TrEMBL	Translated European Molecular Biology Laboratory Database
UV	Ultraviolet

ABSTRACT

The goal of this dissertation was to demonstrate collection, detection and identification of microorganisms from bioaerosols using offline matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) for the analysis of proteins. MALDI-MS intact bacteria techniques were adapted for use with an orthogonal MALDI quadrupole-time-of-flight mass spectrometer. Results indicate the instrument is capable of analyzing intact whole-cells. The first phase was to evaluate three bioaerosol samplers, an Andersen N6 single-stage impactor (AI), a cyclone impactor, and a vacuum filter system. The samplers collected test bioaerosols using a constructed bioaerosol exposure chamber (BEC). The BEC allowed all three samplers to operate in parallel. Each sampler demonstrated the ability to successfully collect and detect the test bioaerosol by offline MALDI-MS. Using the TOF-MS spectra from impacted bacteria, the Expert Protein Analysis System's (ExPASy) sequence retrieval system (SRS) was used to search the SWISS-PROT database. A total of 19 unique proteins were identified for *E. coli*, 8 for *B. Thuringiensis*, and 6 for *B. subtilis*. Subsequently, cytochrome *c* and *E. coli* samples were proteolyzed *in situ* using trypsin and CNBr. The digestions were done using removable mini-wells. The mini-wells were placed on top of collected spots on the MALDI target and served as a mini chemical reactor for digestion. Using the TOF-MS spectra of the digested samples, peptide mass mapping was done using the MASCOT search engine. A progressive reductive iterative search mapping (PRISM) technique was used in order to assist in optimizing protein matches from *E. coli*. In this approach, four of seven iterations produced protein matches. To determine the suitability of MS/MS techniques for use with *in situ* digests, selected fragments from the cytochrome *c* and *E. coli* digests was done. MS/MS was successful for cytochrome c, but was unable to produce spectra for *E. coli*.

CHAPTER 1. INTRODUCTION

Biological warfare and terrorism are of great concern as they have the ability to inflict mass casualties onto a population. Therefore, it is imperative that rapid and accurate detection and identification of these events occur. Since bioaerosols are the primary dissemination mechanism of such events, they are the main focus of monitoring. Successful monitoring is the key element for safeguarding the general public as it allows public health and safety officials to coordinate emergency management actions. It is the goal of this research to advance understanding and improve current bioaerosol monitoring efforts through the development and enhancement of current analytical techniques, specifically those involving the use of mass spectrometry.

1.1 Aerosols and Bioaerosols

Aerosols are colloidal systems of dispersed solids or liquids in a gas. The term aerosol includes both the particle and suspending media. Their sizes can range from 2 mm to 100 μm and they can have lifetimes ranging from a few seconds to over one year.[1] Aerosols are characterized by their size, shape, and density with size being the primary characteristic for determining their physical behavior and atmospheric transport. Aerosol concentration is also important in determining environmental exposure and safety limits and is usually expressed in terms of the number of particles per unit volume of gaseous medium.

Table 1 lists some common aerosols and their respective sizes.

Aerosols are generated from both natural and anthropogenic processes. There are three modes of formation for atmospheric aerosols: nucleation, accumulation and coarse particle formation. Nucleation consists of gas-to-particle conversion which is commonly the result of particles emitted during combustion. As a result, these particles are concentrated near their

Table 1. Common aerosols and their particle sizes.[1]

Aerosol Particle	Size (μm)
Fog, Mist	0.1 – 200
Tobacco Smoke	0.08 – 1.5
Diesel Smoke	0.03 – 1.0
Cloud Droplets	2.0 – 80
Paint Spray	1.5 – 200
Cement Dust	3.0 – 100

source and are usually found in high concentrations that can lead to coagulation with other nearby particles including each other, thus causing accumulation. Accumulation mode is a result of particle growth due to weak removal processes and includes nuclei particles that have coagulated. The combination of nuclei particles and accumulation particles are termed *fine* particles and have a size range between 0.004 to 2.5 μm.[1] The last mode, coarse particle formation, is a result of mechanical process such as wind, wave action, and agriculture, and usually results in particles greater than 1 μm.

Aerosol particle motion is quite complex. However, simple models can be used to predict their persistence in the atmosphere and assist in the collection of particles of a specific size distribution.[1] Aerosol particle size is the primary factor in determining which motion regime, Newton or Stokes, a particle will follow. A good predictor of these regimes of particle motion is the particle's Reynolds number, which is the ratio of inertial forces to viscous forces felt by the particle. For Reynolds numbers greater than 1000, Newtonian motion predominates whereas Reynolds numbers less than1.0 will result in Stoke's Law application. Aerosols of interest are typically between 0.5 and 5 μm diameter, which corresponds to a Reynolds numbers less than 1.0.[1] Aerosol particles within this size range can penetrate deep into the alveolar tissue, the air sacs in the lungs in which gas exchange occurs, and enter the bloodstream and are therefore more physiologically relevant.[2]

Bioaerosols are aerosols composed of whole or parts of biological or biologically active components such as bacteria, viruses, fungi, toxins, algae, and protozoa.[3] They are ubiquitous throughout the troposphere and are extremely diverse in composition and concentration as evidenced through findings of pollen deposits at both the North and South poles.[4] While atmospheric transport of bioaerosols can be beneficial, e.g., pollination of plants, it can also have deleterious affects such as the transport of pathogenic species as in the transport of the organism *Puccinia graminis f. sp. tritici*, stem rust disease. Indeed, *P. graminis* has been found away from its endemic area of the Mississippi Valley to as far north as Canada.[4]

The size of bioaerosols range from 10 nm for viral particles to 100 μm for pollen.[5] Table 2 lists some bioaerosols and their respective sizes. Bioaerosols have been shown to cause allergies, infections, and in the worst case, death as in the case of such natural infections such as pneumonia, measles and diphtheria or biological warfare.[6, 7] Table 3 lists several pathogens and their respective bioaerosol transmitted disease.[8] Pathogenic bioaerosol are

Table 2. Common bioaerosols and their particle sizes.[1]

Bioaerosol Particle	**Size (μm)**
Viruses	0.01 – 0.25
Bacteria	0.25 – 15
Fungal Spores	0.5 – 100
Pollen	10 – 100
Thoracic Particles	0.001 – 20
Respirable Particles	.001 – 10

Table 3. List of some bioaerosol pathogens.

Organism Name	**Organism Type**	**Disease**
Rubella	Virus	Measles
Influenza	Virus	Flu
Mycobacterium tuberculosis	Bacteria	Tuberculosis
Legionella pneumophila	Bacteria	Legionellosis
Aspergillus fumigates	Fungi	Aspergillosis

spread through the atmosphere and are transported based on their physical characteristics. However, during transport, the biological activity of pathogens can be lost or diminished based on their ability to maintain viability during the transport process. The organism's survivability depends on its ability to tolerate environmental stress such as changes in humidity and temperature, exposure to electromagnetic radiation such as UV radiation, and damage from oxidative species such as O_2, O_3, SO_2, SO_2, NO, and NO_2. Therefore, depending on the analytical detection method, viability may be of great concern and will ultimately determine which collection techniques are feasible.

1.2 Bioaerosol Collection

When collecting bioaerosols, sample viability may be a concern, depending on the detection method employed. If sample viability is required, special care must be taken to ensure the microorganisms are not damaged during the collection process. Therefore, it may be necessary to limit collection times and conditions. However, for pathogenic organisms, it may be more suitable to ensure the microorganisms are non-viable to protect the health of individuals collecting and handling the sample catch. In this research, viability is not a requirement and therefore, no special collection methods to maintain viability was used.

For bioaerosol collection, there are two categories of samplers, inertial and non-inertial. Inertial samplers such as impactors and impingers are able to sort or select specific size ranges of particle for collection. Impactors are devices which force a stream of air to intercept a physical surface causing particles in the airstream to hit the object while impingers force air into a liquid medium causing the particles to collect in it. In most cases, they exploit the motion of the particles, which is based on the particle's momentum and hence it's mass, to trap particles of a given size. Conversely, non-inertial samplers such as filters and precipitators, are typically

unable to select particles based on size. Filters are porous materials in which air is drawn through. The path through the filter is irregular forcing the particles to change direction. As the particles traverse the filter, they physically contact the filter material and are entrapped. Precipitators are devices that operate using physical properties such as electrostatic and thermal gradients to force particles to change their trajectory. In the case of electrostatic precipitation, the particles are charged by passing the airstream through a high voltage or corona discharge. Immediately after charging, the particles are collected by an oppositely charged plate. For thermal precipitation, the airstream is passed through an area with a generated thermal gradient. As the particles move across this area, the particles will move towards the area of decreasing temperature. Despite the utility of non-inertial samplers, they often require more power and higher air flow to operate, which can reduce microorganism viability.

1.3 Current Bioaerosol Analytical Methods

There are many analytical methods for detecting material from collected bioaerosols. Despite the many methods, there are two detection schemes, primary (direct) and secondary (indirect). As the name implies, primary schemes involve the direct detection of the analyte such as in the detection of proteins using mass spectrometry, whereas secondary schemes involve indirect detection through the use of additional reagents such as fluorescent tags as in the case of enzyme-linked immunosorbent assays (ELISA).[9] Regardless of the detection scheme, microbial detection assays are broken into two categories, culture and non-culture based as listed in Table 4. In addition, the assays can be typed as either requiring intact microorganisms or lysates. Culture based assays involve counting the number of microbes present in order to determine limits for exposure or determining the presence of infectious microbes to minimize infectious disease outbreak such as bacterial Meningitis.

Table 4. Analytical methods for microorganisms indicating culture and intact cell requirements. (✓ = Required, X = Must not be, – = Not required but can be used)

Analytical Method	Cultured Required	Intact Cells Required
Plate counting	✓	✓
Tissue culturing	✓	✓
Infectivity Assays	✓	✓
Microscopy		
Light microscopy	–	✓
Electron microscopy	–	✓
Antibody based		
Immunofluoresence	–	–
Enzyme-linked	–	–
Immunoelectroblotting	–	X
Nucleic acid based		
DNA/RNA hybridization	–	X
Polymerase Chain Reaction	–	X
Enzyme/Substrate	–	–
GC/LC of metabolites	–	–
Bio–/Chemiluminesce	–	–
Limulus	–	–
IR/Raman Spectroscopy	–	–
Mass Spectroscopy	–	–

The "gold standard" for unambiguous identification of any organism is the comparison of genetic material to that of a sequenced genome.[10] However, routine DNA sequencing for microorganism identification is not common practice because the techniques involved are labor intensive, time consuming and costly.[10] Conversely, genome sequencing of microorganisms is increasing and the database of genomic data is constantly growing. With a sequenced genome, there are several possible approaches for identification, including protein database construction for use in proteomic analysis (which is the study of protein expression in cells) and direct genetic evaluation. Even if the genome is not completely sequenced, limited genomic information can be useful in creating primers for nucleic acid amplification by polymerase chain reaction (PCR). Using this information, PCR primers can be constructed to be either generic for a broad range of organisms, i.e., the 16S rRNA gene conserved region, or tailored specifically for a given genus,

i.e., the invA gene of *Salmonella*.[11-14] Furthermore, species specific primers can be used when a known organism, e.g., *Bacillus anthracis,* is the chosen target.[15, 16]

Immunoassays such as ELISA and Western blots, employ antibodies to recognize specific epitopes, the recognition region on an antigen, from target analytes. Consequently, antibody based detection assays are highly specific for particular antigens. Therefore, it is necessary to create antibodies for each additional target in order to broaden the analyte list. In addition, immunoassays are a secondary detection system that requires coupling of the antibody to a fluorophore, radiolabel, or enzyme. Despite this, these assays are extremely sensitive and have been demonstrated to detect femtomolar concentration of material.[17]

Microscopy is also a powerful tool in microorganism detection. While microscopy covers a broad range of instruments including light, scanning and transmission electron, fluorescence, and confocal microscopes, the easiest and cheapest to employ is the light microscope. With the light microscope, visual inspection of samples is possible to identify microorganisms down to a size 0.2 μm. However, it is often necessary to use either specialized techniques such as dark field, phase contrast or staining for viewing. While these techniques are widely used in a histology, they can only determine visible morphological characteristics such as cell wall structure and shape.

Flow cytometry is another useful technique especially if cell sorting is desired.[18] In flow cytometry, the material is hydrodynamically focused into a stream of single particles or cells. The particles are then detected by measuring sample conductivity, light scattering or fluorescence. Once detected, cells can be sorted based on user defined variables such as cell size or the presence of a chemical label such as a fluorescence antibody tag. In a typical device, the sample is extruded from the flow cell as single-cell charged droplets. These droplets are then

sorted electrostatically using selection rules defined by the user based on the detection criteria for the desired components. Through this technique, samples can either be directed into specific containers later for single cell analysis or accumulated for sample concentrating.

Another technique for the detection of cells, including microorganisms, is the luminol reaction. Luminol is a chemical that chemiluminesces in the presence of strong base and a catalyst such as iron.[19] Incubation of cells at high pH results in cell lysis and release of proteins into solution. Heme containing proteins such heme oxygenase in bacteria are then available to catalyze the chemiluminescent reaction of luminol into 3-aminophthalate producing a photon emission near 425nm.[20] While this technique is a reliable method for cell detection in forensics, it cannot be used to distinguish among cell types including the difference between bacterial cells and blood cells. However, the intensity of photoemission observed can be correlated to cell concentration. This can give a rough estimation of cell number, but it is not a definitive assay for cell counting because heme-iron content varies between cell types and species. Moreover, since this is a non-specific test, caution must be taken to prevent sample contamination.

A more specific cell identification test is the *Limulus* amoebocyte lysate (LAL) assay.[21] The LAL assay detects a Lipopolysaccharide (LPS) found in gram-negative bacteria, which is a bacteria endotoxin. Gram negative bacteria are bacteria which stain negative when treated with gram reagents. These bacteria contain two cell membranes, an inner and outer, with a thin layer of peptidoglycan, or cell wall, in the periplasmic space, which is the space between the two cell membranes. The origin of this assay dates back to the 1950s and was developed as a result of the observation that amoebocyte cells from the horse shoe crab, *Limulus polyphemus*, had the ability to stimulate clotting in the presence of gram-negative bacteria.[21] In 1977, this assay was approved by the Food and Drug Administration (FDA) to replace their standard endotoxin assay,

the rabbit pyrogen test. By 1987, the FDA had established the LAL test as standard method for testing all intravenous drugs for the presence of endotoxin contamination. The LAL method is a turbidity-based optical technique. When a sample containing gram-negative endotoxins is exposed to the LAL mixture, coagulation results, causing the optical density of the solution to increase.[22] The optical density change can be detected and used to assist in the determination of the presence and relative quantity of gram-negative bacteria.[23, 24] Despite the utility of this technique, the LAL assay cannot be used as a general bioaerosol detection assay due to its specificity for gram-negative bacteria endotoxins.

Fourier transform infrared and Raman spectroscopies are powerful techniques for detecting and identifying microorganisms. [25-27] These techniques analyze differences between sample spectra and spectra generated from pure samples. Reference spectra from pure samples are generated and cataloged into a spectral fingerprint database for use in sample identification. These techniques have demonstrated the ability to detect bacteria, yeast and fungi.[25-27] Identification is based on the principle that microorganism's produce unique spectra fingerprints due to their distinct composition and quantity of components such as lipids, protein and nucleic acids. In general, spectral differences are most notable between 1800 – 900 cm^{-1} corresponding to the key regions for well defined chemical species found in biological cells.[28] Table 5 lists some of the chemical species and their approximate wavenumber. While it usually requires approximately 10^3 cells for detection, recent studies demonstrated the ability to detect a single microorganism.[29] Despite this sensitivity, IR techniques suffer from the inability to unambiguously identify molecular species and the variability of spectra resulting from culture and experimental conditions.[30]

Table 5. IR spectroscopic bacteria chemical signature.[31, 32]

Chemical Bond Origin	Approximate Wavenumber (cm^{-1})
Amide I	1650
Amide II	1540
Phosphodiester bonds from nucleic acid backbone	1242 1080
Polysaccharide compounds	1050-950
CH3	1455
CH2	1398
Dipicolinic Acid	1447-1439

The last method to be discussed is mass spectrometry (MS). With mass spectrometry, identification down to the strain level, a subset of the species with minor differences such as the presence of a mutation, is possible.[33] This is achieved through two identification techniques: mass spectral fingerprinting and database searching. MS fingerprinting is analogous to IR and Raman fingerprinting in that mass spectra for pure isolates are recorded and stored in a database for comparison to unknown spectra. However, MS has an advantage over IR and Raman in that biological macromolecules detected in MS can be uniquely identified based on mass from either a public user protein database such as SWISS-PROT or a theoretical database based on genomic data such as TrEMBL. SWISS-PROT is a manually curated database of protein sequences and is a collaborative effort between the Swiss Institute of Bioinformatics (SIB) and the European Bioinformatics Institute (EIB). TrEMBL is the translated nucleotide sequence database of the European Molecular Biology Laboratory (EMBL) and is a supplement of SWISS-PROT in that it contains information not yet integrated into the SWISS-PROT database. Another advantage of MS for detection and identification of microorganisms is the ability of this analytical technique to detect all three major classes of biological macromolecules such as lipids, proteins, and nucleic acids. In addition, the use of tandem MS instruments, instruments with more than one mass analyzer, can provide MS^n capabilities and give specific information regarding chemical

composition such as protein and DNA/RNA sequences. The sequence information can then be used to refine search algorithms for matching against databases of genomic and proteomic data.

1.4 Mass Spectrometry Methods

Identification of microorganisms by mass spectrometry is on its way to becoming a routine analysis for biological samples. It is capable of analyzing for all three classes of biological macromolecules and has the ability to utilize spectra fingerprinting and database searching. The following sections give some background into the different mass spectrometry methods for detecting and identifying microorganisms.

1.4.1 Pyrolysis Mass Spectrometry

The first mass spectrometry technique to be discussed is pyrolysis MS. While other analytical techniques can be coupled to pyrolysis such as gas chromatography(GC)-flame ionization detection and GC-ion mobility spectroscopy,[34, 35] mass spectrometry is often the chosen technique due to its speed, sensitivity and ability to couple to separation methods such as chromatography. Pyrolysis is the thermal degradation of complex organic mixtures such as whole organisms or lysates, in an inert atmosphere. Pyrolysis mass spectrometry consists of four steps, pyrolysis, ionization, mass separation and detection. Ionization of the pyrolysate is usually accomplished by a low energy electron beam (<30 eV) which produces singly-charged positive molecular and fragment ions. Following ion formation, the ions are introduced into the mass analyzer by applying a positive voltage onto a metal plate called a repeller plate within the ionization chamber. Typically, after ionization, pyrolysis products are less than 500 Daltons (Da); however, for biological material, this range is usually less than 200.[36, 37]

Application of pyrolysis MS improved greatly through the analysis of lipid and bacterial cell wall products such as peptidoglycan (PG) material. For PG analysis, a field portable Py-GC-

IMS biodetector has demonstrated the ability to provide taxonomic data equivalent to traditional gram staining.[34] For lipid analysis, extraction and derivatization of cellular lipids is usually done prior to pyrolysis. A widespread and accepted technique for identification of bacteria by pyrolysis is analysis of cellular phospholipids through the generation of fatty acid methyl esters (FAME).[38-40] In FAMEs, the sample is treated with a sodium hydroxide or other suitable alkali followed by the production free fatty acids. Subsequently, the carboxylic acid group is methylated through acid treatment followed by the addition of a methylating solution of boron trichloride and methanol. The resultant FAME material can then be extracted with a two phase non-polar organic/ether mixture such as hexane/*tert*-butyl ether.[41] The FAMEs can then either be chromatographically separated or analyzed directly by mass spectrometry for identification.[42] While this technique is useful, it lacks the ability to identify microorganisms down to the species level, which is required for homeland security and public health safeguarding.

1.4.2 Electrospray Ionization Mass Spectrometry

ESI was first introduced by Malcom Dole in 1968[43] and refined in the 1980s by John Fenn[44] for the analysis of large biomolecules including phospholipids, glycolipids, nucleic acids, carbohydrates, and proteins.[45-56] Despite the wide range of molecules, ESI is limited in its ability to handle complex biological samples. This is a result of the complexity of mass spectra obtained. ESI often results in analytes having a wide range of charge states requiring deconvolution to obtain the parent mass. Consequently, samples with multiple analytes increase the difficulty of this deconvolution and, if the sample is sufficiently complex such as intact cells, the task becomes nearly impossible. Moreover, ESI has a low tolerance for impurities including salts, buffers and detergents which are often utilized in biological sample preparations.

Nevertheless, ESI is a crucial linchpin in proteomics applications when used with separation methods.

A crucial proponent of biological sampling by ESI is in the ease of coupling this ionization technique to analytical separations such as capillary electrophoresis (CE) and liquid chromatography (LC). Coupling of ESI to these separation techniques separates the analytes in time creating a much simpler series of mass spectra. Additionally, these separations can serve to remove incompatible sample impurities. In the realm of biological sampling, ESI analysis of whole intact-cells is limited due to the enormous degree of sample complexity. Consequently, almost all methods currently established involve cell lysing followed by either an extraction or separation. Despite these shortcomings, ESI has been demonstrated to confirm the presence of bacteria and even differentiate between bacteria species through the analysis of muramic acid, phospholipid/glycolipid composition, protein profiles and PCR product analysis[45-56]. However, these analyses are lengthy and tedious.

1.4.3 Desorption Electrospray Ionization

Desorption electrospray ionization (DESI) is a new method for desorbing and ionizing material from a surface. [57, 58] In DESI, an electrospray source is directed onto the sample surface at an angle between 0 and 90 degrees with the skimmer cone of the mass spectrometer placed at a similar angle in the same plane as the ESI source. The electrospray source is then turned on and electrospray droplets are allowed to hit the sample. Analyte molecules are detected after they are desorbed and ionized by the electrospray droplets. DESI is capable of detecting samples of proteins, explosives, chemical warfare agent simulants and bacteria biomarkers.[59, 60] In general, mass spectra produced from DESI experiments resemble those of ESI in that distributions of analyte ions with high charge states are typically observed.

1.4.4 Laser Desorption/Ionization Mass Spectrometry

Laser desorption techniques are diverse in mass spectrometry applications and include laser microtarget mass analysis (LAMMA), laser desorption ionization (LDI), laser-assisted desorption electrospray ionization (LADESI), and matrix-assisted laser desorption /ionization (MALDI). One of the first attempts to identify bacteria by LDI was done Heller et al.[61] In this work, bacteria cell lysates were analyzed by a time-of-flight mass spectrometer equipped with an IR laser to produce mass spectra of phospholipids. Later work by a group at Lawrence Livermore National Laboratory (LLNL) used both IR and UV lasers to analyze four species of *Bacillus* bacteria spores from lysozyme lysed cells.[62] In this work, the IR (λ = 3.05 - 3.8 μm) laser and the UV laser (λ = 337 nm) both were able to generate spectra, however the results concluded that IR lasers were better suited for use in the matrix-free sample analysis. In addition, they found that a wavelength of 3.05 μm produced the best spectra from whole spores of several *Bacillus* species and they were able to obtain peaks up to 19 kDa.[62]

An online detection system was also developed by LLNL called bioaerosol mass spectrometry (BAMS).[63] BAMS utilizes an aerosol time-of-flight instrument based on the design introduced by Gard et al.[64] BAMS is equipped with an UV laser for ionization of bioaerosol particles directly introduced into the instrument.[65] While this setup does not constitute desorption in a traditional sense, i.e., off a surface, they were able to record both positive and negative ions simultaneously for *m/z* values less than 400 for pure samples[66] and under 200 for intact cells.[63] Previous techniques have demonstrated a higher *m/z* range with detected peaks up to 20 kDa, however the use of a matrix was required.[67] More recently, a laser desorption post–ionization technique involving two lasers was used to detect derivatized peptides from biofilms containing *B. subtilis*.[68] The two laser setup consists of a N_2 (λ = 337

nm) laser for desorption and a F_2 (λ = 157 nm) laser for post–ionization. The F_2 laser is positioned parallel to the sample surface for postionization.[68] Despite the emergence of laser desorption/ionization techniques in biomolecule detection, the use of a matrix for the analysis of biological samples does provide additional enhancements such as increased *m/z* range.[69]

1.4.5 Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry

With regard to analysis of whole-cell microorganisms, MALDI has been used extensively due to its high tolerance of impurities and ease of sample preparation. Generally speaking, MALDI analysis of whole-cell microorganisms is the fastest offline analysis technique for obtaining data without the need for extensive sample preparations such as culturing, cell lysis or analytical separations. The speed of MALDI-MS analysis for bioaerosols lies in the ability of a MALDI target to be used as the impaction surface inside bioaerosols impactors. 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.[70] MALDI-MS techniques have been developed for the rapid analysis of biological samples for the detection and identification of microorganisms such as bacteria, viruses and fungus.[71-73] However, MALDI-MS is not limited to whole-cell microorganism analysis and is often used in the analysis of prepared biological samples to include cell lysates and LC fractions as shown in Figure 1. Moreover, MALDI-MS can be used to analyze all three major classes of biological macromolecules: lipids, nucleic acids and proteins.[74-80] For intact-cell microorganism analysis, target analytes are typically phospholipids and proteins due to their location within the cell.

1.5 MALDI-MS Bacteria Identification Techniques

There are two techniques for identifying bacteria from MALDI-MS: protein database searching and mass spectral fingerprinting (Figure 2).[78, 81] For mass spectral fingerprinting, a

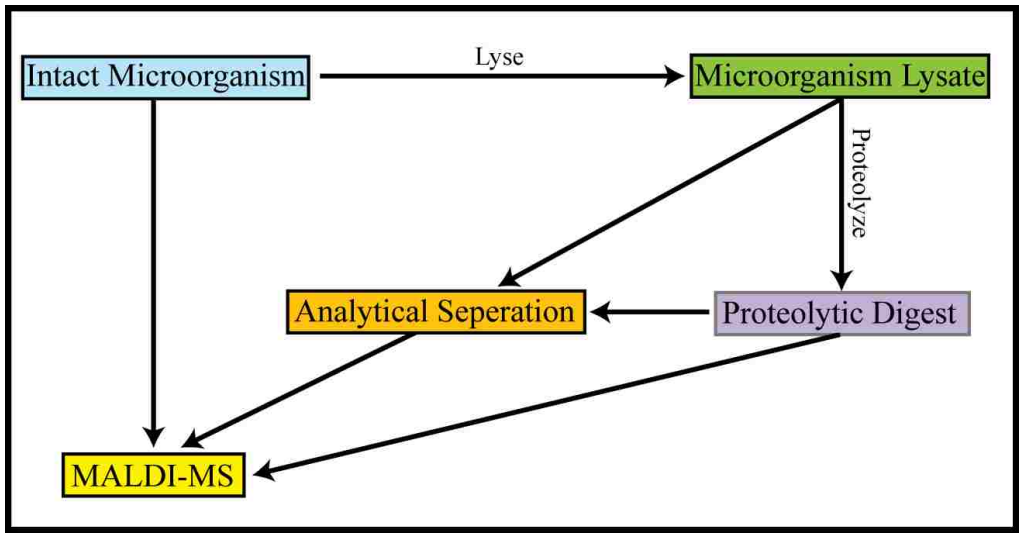


Figure 1. MALDI-MS microorganism analysis workflow.

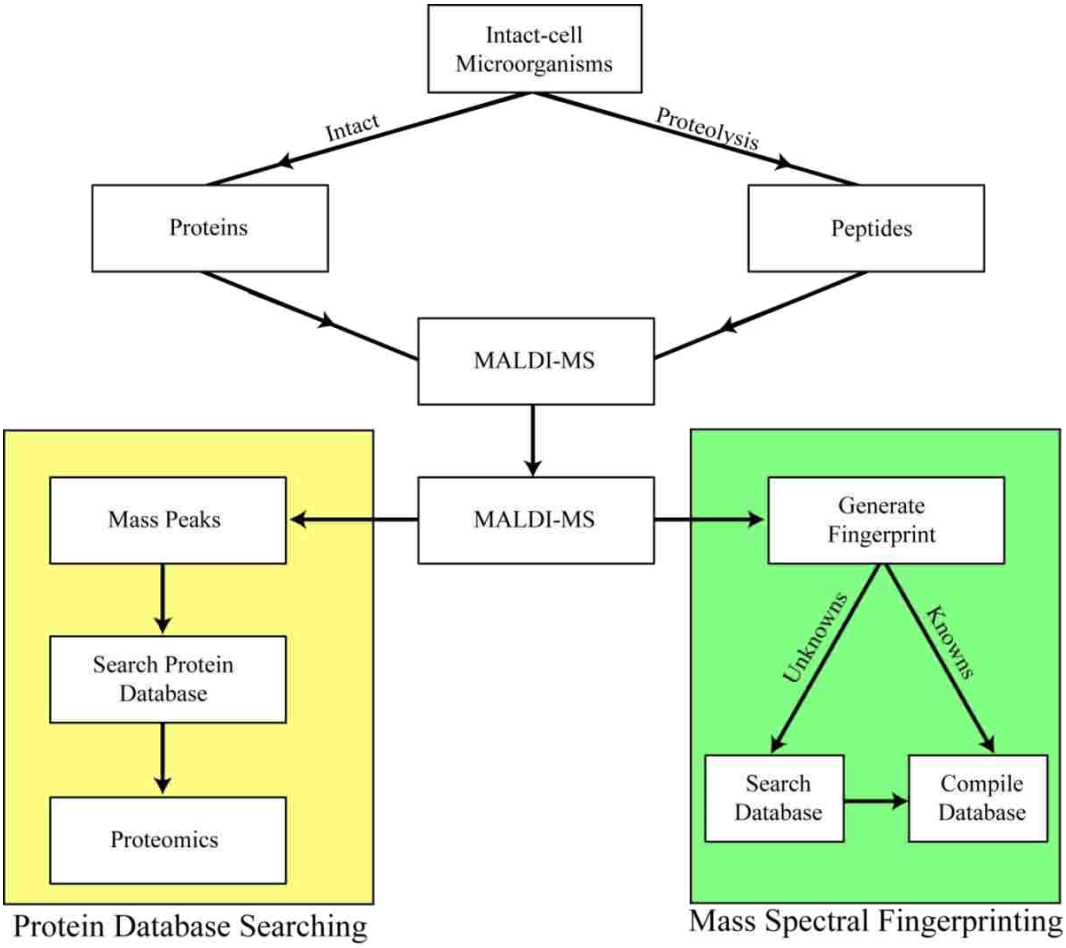


Figure 2. MALDI-MS bacteria analysis and identification schemes.

database is constructed from target bacteria containing peak location (*m/z*), peak intensity, and frequency of occurrence. [78, 82-84] With this information, an 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.[78, 85] Despite this utility, the data subjected to the matching algorithm must be smoothed prior to analysis. Also, the data must be in identical formats. Specifically, the data recording intervals must be identical or additional data manipulation is required in order to compare with stored fingerprint mass spectra. However, a more versatile technique, protein database searching, is available. In protein database searching, mass spectra peaks are searched against protein databases such as SWISS-PROT and TrEMBL for mass matches, which can then be assigned to a specific organism to create a match.[81, 86-88] This process is done for all peaks in the spectra and a list is then compiled to assist in identification.[87] Both of these methods have been demonstrated with pure bacteria samples as well as simple mixtures of up to 4 bacteria species.[85]

Bacteria identification by MALDI-MS remains an emerging technique. Current developments are focused on increasing the dependability of identification. One such procedure is proteolysis. Through proteolysis, peptide fragments are generated and detected to produce peptide mass maps. Peptide mass mapping is a powerful tool for identifying proteins based on cleavage product and parent masses. Proteolytic treatment of intact-cell microorganisms has been demonstrated and used to increase identification reliability. [88] At present, proteolysis is more suitable for protein database searching. However, an algorithm could potentially be made for mass spectra fingerprinting that includes proteolytic fragments. Another developing approach is the use of tandem mass spectrometer (MS/MS) instruments. MS/MS is a mass spectrometry technique that produces mass spectra from fragmentation of ionized molecules within the mass

spectrometer.[89] The resultant mass spectra can provide structure information and be used in molecular structure determination. Selective bond cleavage, such as peptide backbone cleavage of polypeptides, can be achieved by adjusting the amount of energy used for fragmentation. If mainly peptide backbone cleavages are found, then the MS/MS spectra can be used to determine the amino acid sequence, which can then be used to increase the confidence of protein identification.

Although both techniques providing similar identification capabilities, mass spectral fingerprinting has a major disadvantage in that it is suitable only for target organisms in which mass spectra fingerprints are already cataloged. In general, the mass spectra fingerprint database will only contain organisms that generate significant interest such as biological warfare agents, and will be unsuitable for generic microorganism analysis. The mass spectra fingerprint database can be expanded, but it will always only be capable of identifying microorganisms included in its database. Conversely, protein database searching can be used for wider range of microorganism identification. Similarly, the limitation is database construction. Protein databases are constantly being updated to reflect new proteins or translated proteins from genomic data. This continues to increase the robustness of protein database searching.

1.6 Research Objectives

The main focus of this research is to adapt MALDI-MS techniques for the identification of bacteria with specific focus on coupling bioaerosol collections to offline MALDI-MS analysis. First, several bioaerosol collection systems were tested for compatibility with MALDI-MS analysis. These included an Andersen N6 single-stage impactor, a cyclone impactor and a vacuum filter system. Each sampler system required minor adaptations for MALDI-MS analysis. Second, intact whole-cell analysis were adapted to a quadrupole-time-of-flight mass

spectrometer for MS/MS. The spectra collected were compared to those from a MALDI-TOF instrument and used for searching online protein databases such as SWISS-PROT and TrEMBL for the identification of proteins from the collected bioaerosols of proteins and bacteria. Finally, proteomic digestions using trypsin and CNBr of intact whole-cells from impacted bioaerosols were done. All digestions were done *in situ* using removable mini-wells directly on the target surface followed by analysis by MALDI-MS for use in peptide mass mapping.

CHAPTER 2. EXPERIMENTAL

2.1 Bioaerosol Generation

All bioaerosols under study were generated by compressed-air nebulization. In this work, a Collison 6-jet nebulizer (Figure 3) was used for bioaerosol generation. Nebulization operates similar to atomization where a gas is used to aspirate the liquid.[90] In atomization, compressed air stream is directed through a liquid stream where the energy from the compressed air breaks the liquid into particles. In nebulization, the sample liquid is drawn into the gas stream and the gas/liquid jet exits the nozzle creating a large size range of particles. Under most circumstances, a narrow and stable size range of particles is desired. A collison nebulizer is designed to cause the gas/liquid jet to impact onto the side of the sample jar to control the particle size distribution. This process removes the larger particles (greater than 10 μm) from the airflow inside the nebulizer sampler jar. Consequently, the smaller particles (less than 10 μm) are swept up towards the nebulizer outlet. Before exiting the nebulizer, an additional size selection is achieved with a curved tube near the exit. This tube limits the size range of particles exiting to less than 10 μm. The collison nebulizer is well characterized as shown in Appendix A where the mass median diameter (MMD), the median diameter size as calculated from the minimum and maximum particle sizes produced, and the volume consumption and flow rates are correlated to the pressure used in nebulization.

All samples were suspended in 10 mL of water and sprayed until the nebulizer was unable to generate particles from the remaining liquid. The nebulizer sample jar was filled and the T-stem was placed approximately 3/8” below the surface of the liquid to ensure proper operation. A pressure of 30 psi was used for nebulization. For proteins, 5–10 mg samples were

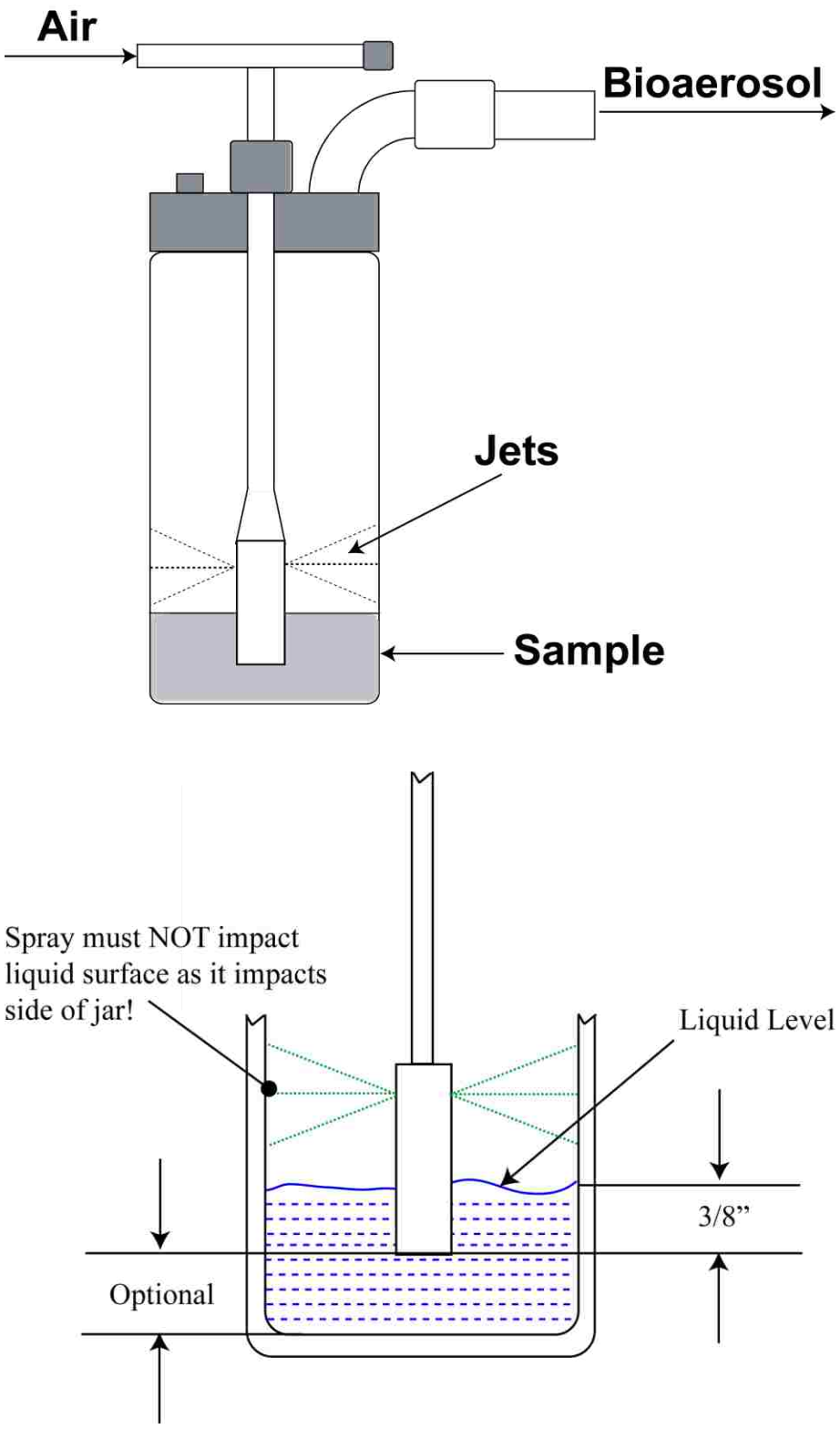


Figure 3. Collison Nebulizer Diagram.

used. For bacteria, 20–40 mg quantities were used. Nebulization was complete in less than 1 hour of continuous operation. The bioaerosol output from the nebulizer was directed into a bioaerosol exposure chamber (BEC) to ensure operator and lab personnel safety.

2.2 Bioaerosol Exposure Chamber (BEC)

The BEC was constructed from ¼ inch plexiglass (Figure 4) in a rectangular box with dimensions of 24”x18”x18” for the length, width and height, respectively. The BEC has five

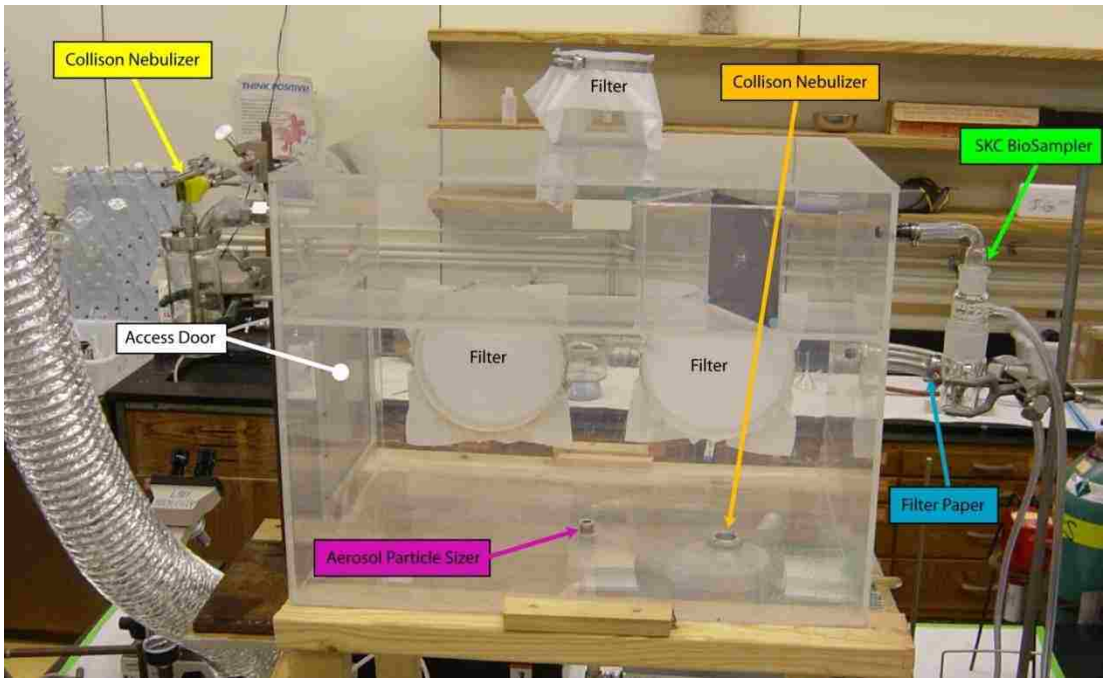


Figure 4. Bioaerosol exposure chamber.

accessible ports. In this work, four of the ports were used for sampling and one was used for bioaerosol introduction. In addition, the BEC has one sealed access port for cleaning between bioaerosol experiments. All remaining non-sampling port openings were covered with submicron nylon monofilament filters (Great Lakes Filters, Hillsdale, MI) and attached with hose clamps onto the ports outfitted with ¼” polyfoam weather seal to prevent tearing of the filter on the edges. The filter allowed gas exchange and prevented the generation of a pressure differential

during nebulization and sampling. All samplers could be operated. The top filtered port is designed for easy hookup to a hood exhaust.

2.3 Generated Bioaerosol Characterization

A Model 3321 aerosol particle sizer (APS) from TSI, Inc. was used to monitor and characterize bioaerosols generated inside the BEC by the collison nebulizer. The APS is capable of sizing particles between 0.5 and 20 μm. A schematic of the APS is shown in Figure 5. Aerosol samples are drawn into the APS at 1 L/min and combined with a sheath flow at 4 L/min before being accelerated through a orifice nozzle. The particle acceleration is directly related to its aerodynamic particle size, which is related to the particle's mass. The larger the size, the slower the particle is accelerated through

the detection system. Consequently, the particle's velocity can then be used to determine the aerodynamic particle size.

The APS utilizes a double-crest optical system for single particle detection (Figure 6). The double crested signal is the result of the particle passing through two partially overlapping laser beams, resulting in two detection events or crests for each single particle. The smallest time resolution possible for the instrument is 4 ns, which limits the instrument's ability to accurately measure particle sizes less than 0.5 μm although particle detection is possible down to 0.3 μm. The instrument's particle size limit is due to the wavelength of light used (655 nm). Detection of particles below 0.3 μm is not efficient. As single particles pass through each beam, the two detection events occur with some Δt between the signal maxima. The Δt is used to calculate the particle's aerodynamic size. Detection events that do not result in two peaks do occur and result in different outcomes for data logging and grouping as shown in Table 6.

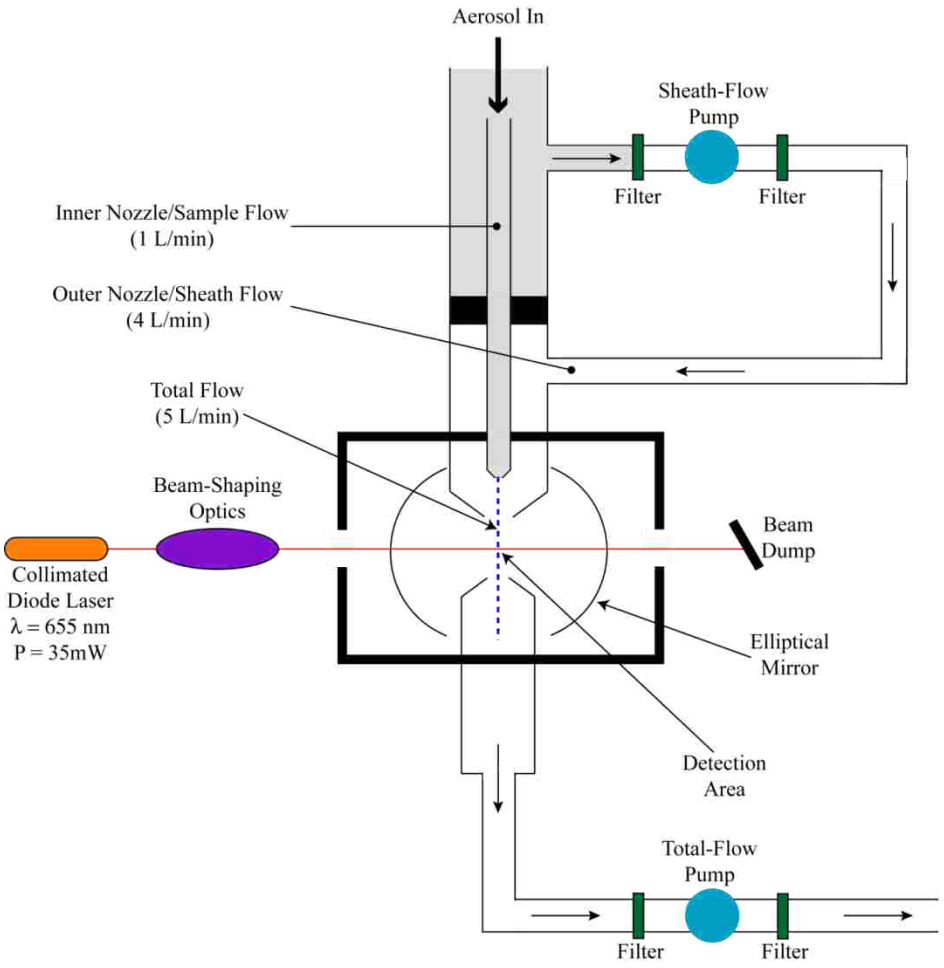


Figure 5. Aerosol particle sizer schematic.

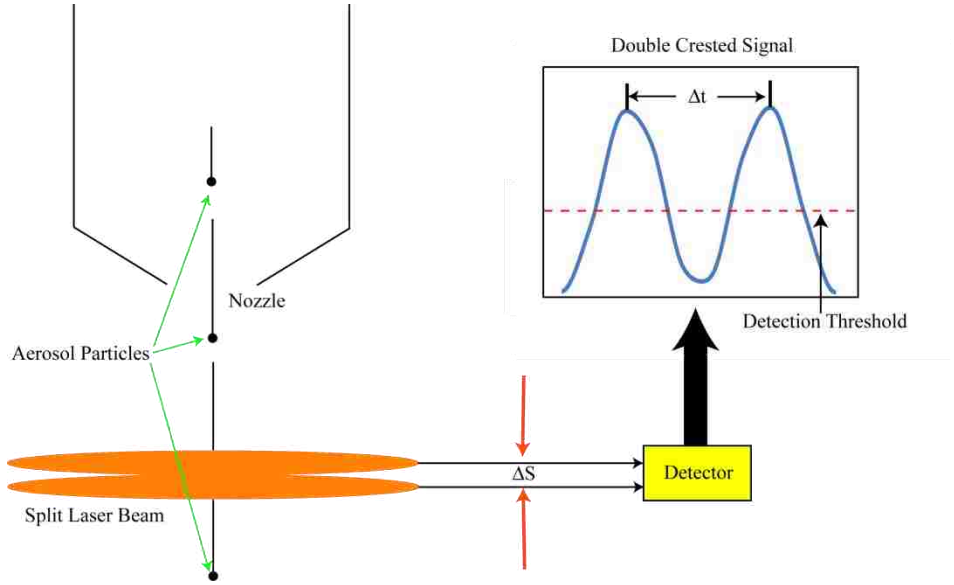


Figure 6. Double beam 2-crest process.

Table 6. APS Time-of-Flight Measurement Results.

Event Type	Valid	Outcome of Data
1 sample crest detected	No – Particle size Yes- Concentration	Logged for concentration purposes in the <0.523 μm cat.
3 sample crests detected	No	Logged but not used
2 sample crests detected with no minima below detection threshold	Yes	Logged for concentration and TOF
2 sample crests detected with minima below detection threshold	Yes	Logged for concentration and TOF
2 sample crests detected outside TOF timer window	No	Logged but not used

2.4 Bioaerosol Collection

For bioaerosol collections, three different samplers were used, an Andersen N6 single-stage impactor, an SKC BioSampler (cyclone impactor), and a filter vacuum assembly. Each of these samplers were used simultaneously for collecting the generated bioaerosols from the BEC. All samplers were run for the duration of nebulization which was typically less than 1 hour. Post-collection sample processing was minimized to streamline data collection and analysis. All collector's samples were analyzed by MALDI mass spectrometry.

2.4.1 Collection by Impaction

Impaction as the name implies is the phenomenon whereby particles physically impact a surface. Upon impaction, the particles either adhere to the surface or bounce off. Particle bounce is the single most important phenomenon for impaction samplers. To minimize particle bounce, samplers should be operated at the lowest flow to decrease impact velocity and detachment of adhered particles during collection. In addition, surface coatings can be used to enhance the adhesive forces for impacted particles with the surface. In this work, two impactors are used, a cyclone impactor and an Andersen N6 single-stage impactor.

The Andersen N6 single-stage impactor (AI) is used extensively in the collection of viable microorganisms including bacteria and fungi (Figure 7). Typical operation utilizes 90 to

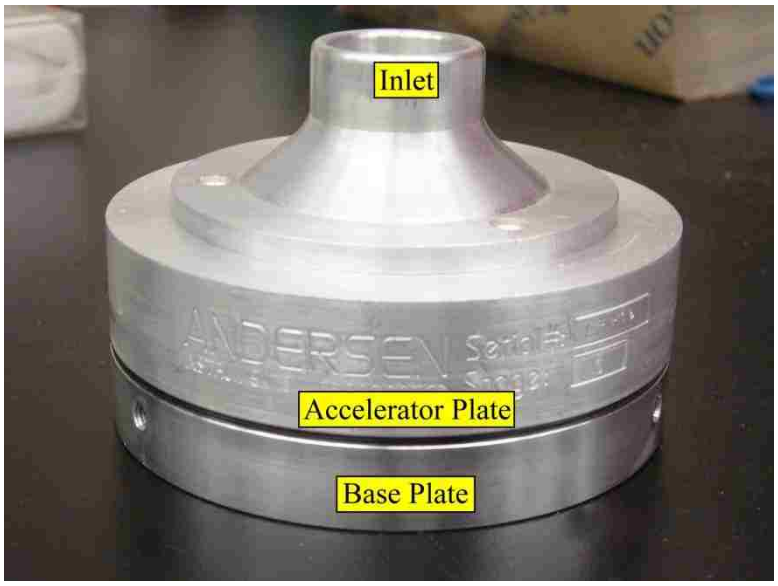


Figure 7. Image of an Assembled Andersen N6- Single-stage impactor.

100 mm plastic Petri plates filled with agar, which is typically a 1 to 2% solution of polygalactose, and nutrient growth media. The agar plate is placed on top of the base of the impactor to serve as the impaction surface. For our studies, a bare stainless steel MALDI target and an uncoated microscope glass slide were used for the impaction surfaces. The impactor is equipped with a 400 hole accelerator plate containing 0.26 mm diameter holes (Figure 8). Each hole serves as a separate impaction mini-jet. Increasing the number of mini-jets decreases the chance of particle coincidence in which two particles are sampled at the same location. However, in this work, particle coincidence is not a concern since this only increases the concentration of particles for analysis. The AI was operated at 28.3 L/min when used individually and at 14.5 L/min when used in conjunction with other samplers. The AI uses a rotary vane pump specifically designed and calibrated for this impactor.

The theory of inertial impaction for the collection of bioaerosols is simple and straightforward. A vacuum pump is used to sample a stream of air which is directed against a surface. The direction of the air stream is typically perpendicular to the impaction surface

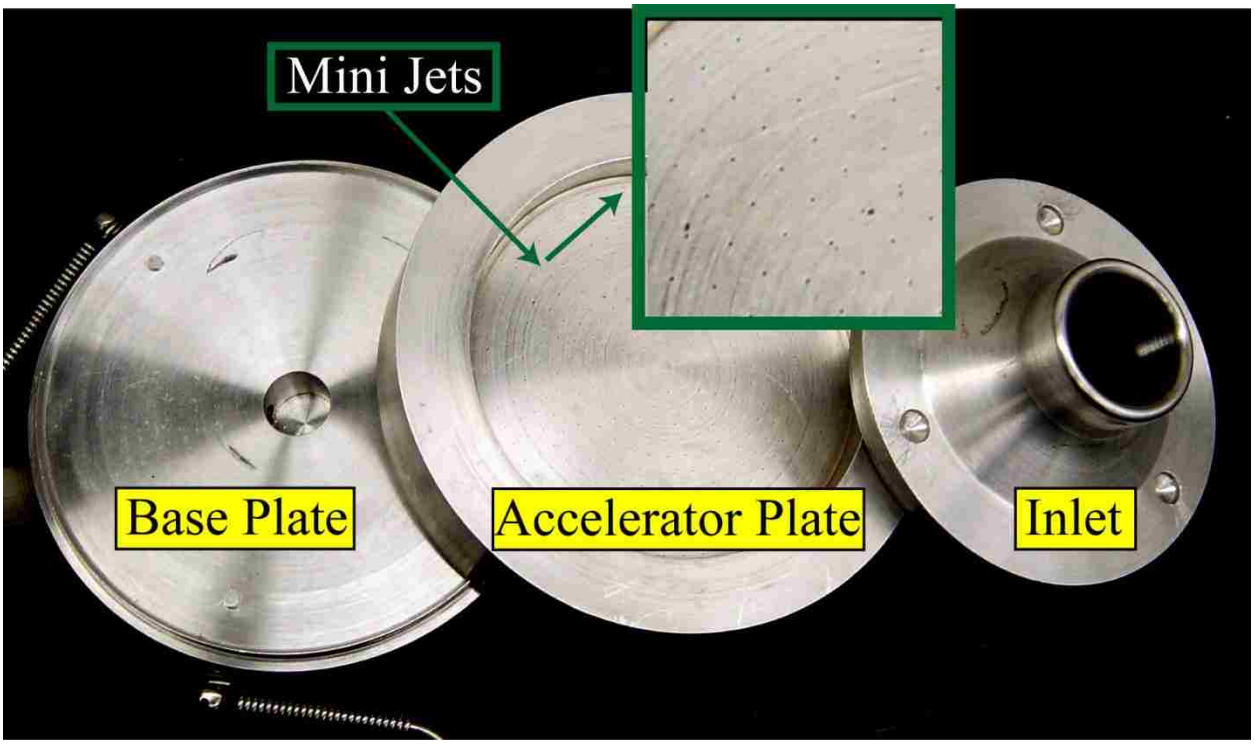


Figure 8. Disassembled Andersen N6 Single-stage Impactor with zoom in of mini-jets on Accelerator plate.

(Figure 9). In this case, the particles initially follow the flow of the airstream. As the airstream approaches the impaction surface, the flow changes directions. A particle's ability to change directions to follow the airstream is related to its inertia, which in turn is based on the particle's mass. To simplify this discussion, an example on the theory of operation of inertial impactors will be made using a single mini-jet environment of the AI used in this work.

For a particle to impact the surface during collection, it must move a certain distance away from its original streamline and towards the surface. Each particle entering the mini-jet orifice has a position within the airstream with respect to the centerline as shown in Figure 10. For simplicity, the streamline path is assumed to be a quarter circle resulting in curvilinear motion. Each particle experiences a radial velocity (V_r) during this motion as the particles follow the path of the changing airflow. Due to this radial velocity, the particle moves away from

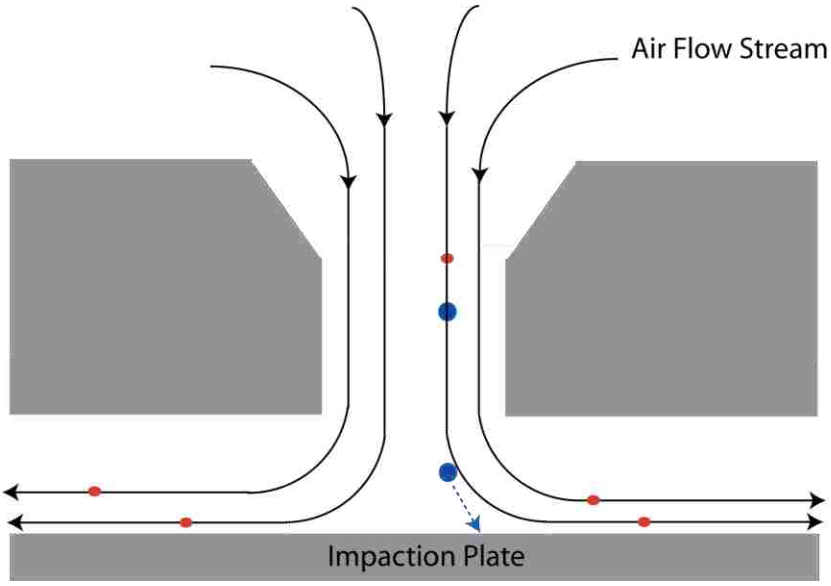


Figure 9. Impaction particle collection based on curvilinear motion.

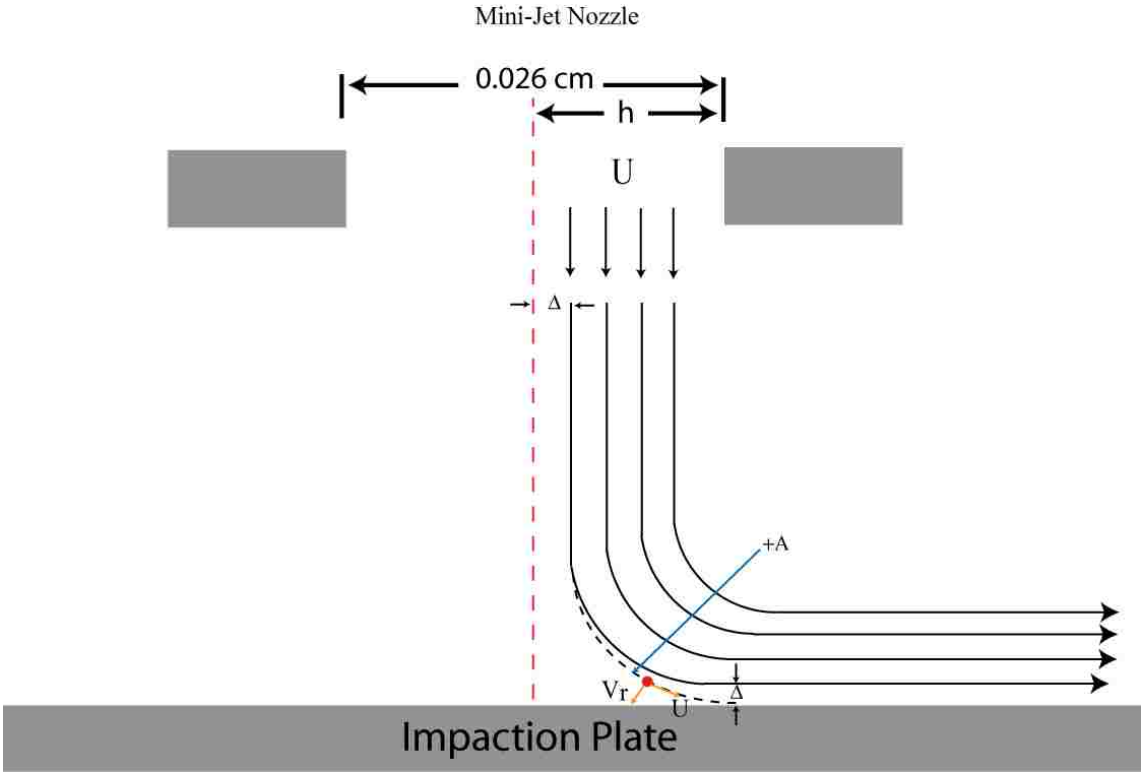


Figure 10. Impactor Theory.

its original streamline and toward surface until the particle's streamlines are no longer curved. During this time, if a particle impacts the surface, the distance between its original streamline and the surface is given by Δ. For all identical particles within a distance Δ from the mini-jet centerline, impaction will occur. This information can be used to calculate the size of the particles that will impact the surface.[1]

The Andersen N6 single stage impactor has a sharp cutoff size or diameter (d_{50}) for particles. The cutoff diameter is the size in which particles larger will be collected and particles smaller will not. To calculate this, an explanation of the stokes number, Stk, is required. The Stk is a dimensionless number that relates a particle's stopping distance to that of an obstacle's diameter (Equation 1). [1] For our application of using the AI, a slight derivation is required to change the obstacle diameter, D_c, to the mini-jet radius. This change is shown in Equation 2. Equation 2 can be transformed into Equation 3 by substitution of the appropriate parameters for stopping distance.[1] Using equation 3, the cutoff size for the AI with a 0.26 mm diameter mini-jet

$$\mathrm{Stk} = \frac{\tau U_0}{d_c} \qquad \textbf{Equation 1}$$

$$\mathrm{Stk} = \frac{\tau U}{D_j/2} \qquad \textbf{Equation 2}$$

$$d_{50}\sqrt{C_c} = \left[\frac{9\eta D_j^3 (Stk_{50})}{4\rho Q}\right]^{1/2} \qquad \textbf{Equation 3}$$

can be calculated with a few assumptions. In this equation, the viscosity of air is 1.81×10^{-5} Pa·s under standard conditions (293K and 101kPa). In addition, the particle density chosen is that for a unit spherical water droplet (1000 kg/m^3) and the Stk_{50}, the stokes number that gives 50 percent collection efficiency, for the N6 stage of an Andersen impactor with a circular jet is 0.24.[1] Finally, since the slip correction factor is a function of d_{50}, Equation 4 can be used to calculate

$$d_{50} = d_{50}\sqrt{C_c} - 0.078\mu m$$ **Equation 4**

the value of d_{50} to within 2% at pressures between 91 and 101 kPa.[1] Therefore, d_{50} for the AI operated at 28.3 L/min is equal to approximately 0.6 μm.

The cyclone impactor (CI) is a unique bioaerosol impactor in that the impaction surface is not perpendicular to the bioaerosol jet stream. The CI has three sonic jet nozzles of 0.25 mm diameters configured to produce a net circular flow of air within the collection vessel. This configuration is designed to minimize particle bounce and allow for the use of solvents in the collection reservoir to enhance collection efficiencies. The technical specification for the angles and distances from the nozzle to the cylinder surface is shown in Figure 11. The collection efficiency of this configuration for 0.5 μm particles is greater than 80 percent.[91, 92] Figure 12 shows an image and partial schematic of the CI depicting the airflow through the sampler. The solvent used in all of these studies was water. The sampler was operated at 14.5 L/min using a rotary vane pump.

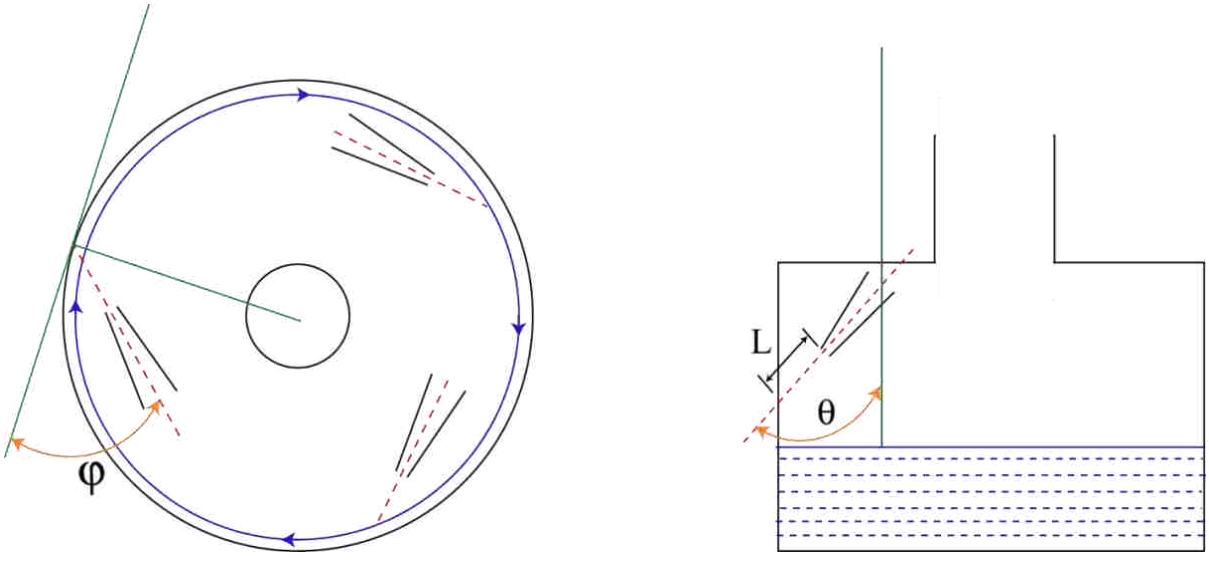


Figure 11. Cyclone impactor technical specifications [φ = 30 degrees, θ = 60 degrees].

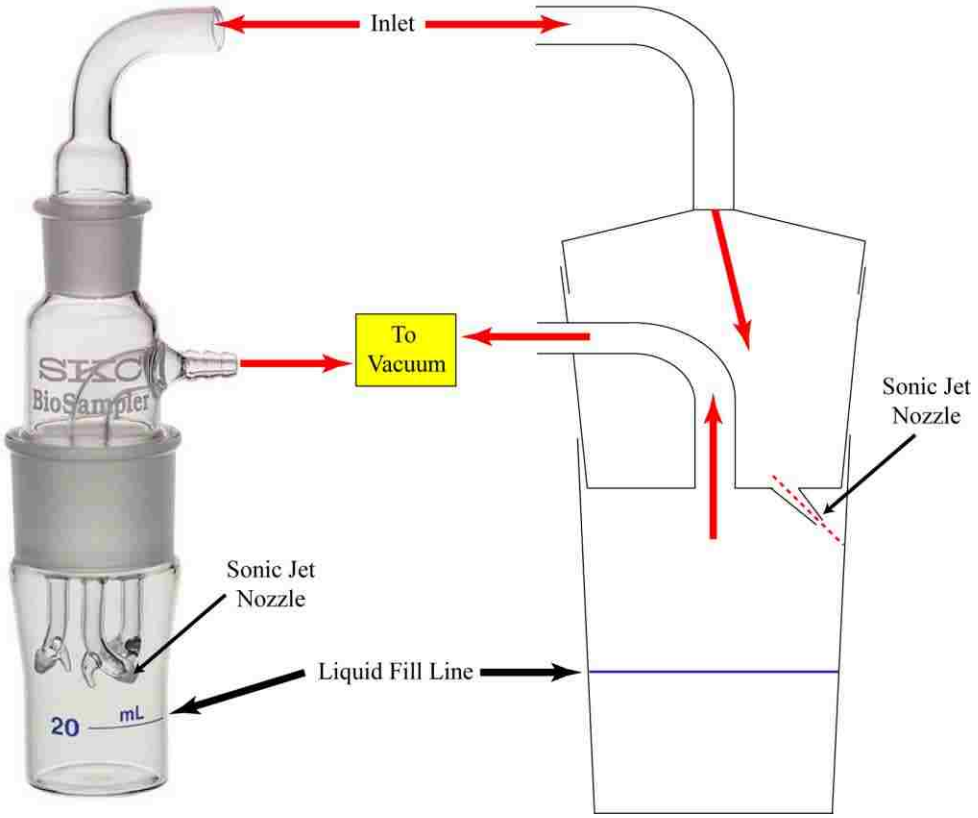


Figure 12. SCK BioSampler schematic.

2.4.2 Collection by Filtration

The last collection method used in this study was filtration. Two types of fibrous filters were used: Whatman GF/F and Whatman #1. The GF/F filters (ca. 0.42 mm thickness) are made from borosilicate glass microfibers and the #1 filters (ca. 0.18 mm thickness) are made from cellulose fibers. The liquid particle retention sizes are 11.0 μm and 0.7 μm for #1 and GF/F, respectively. However, experimental evidence indicates the aerosol particle retention sizes are smaller than those indicated for liquids. For the setup, the filter paper was sandwiched between two ISO–KF half nipple flanges such that the filter is located before the o-ring/centering ring in the air stream. The filter, flanges, and o-ring are then secured by a hinged clamp as shown in Figure 13. A rotary vane vacuum pump was then used to draw air through the filter setup at 14.5 L/min.

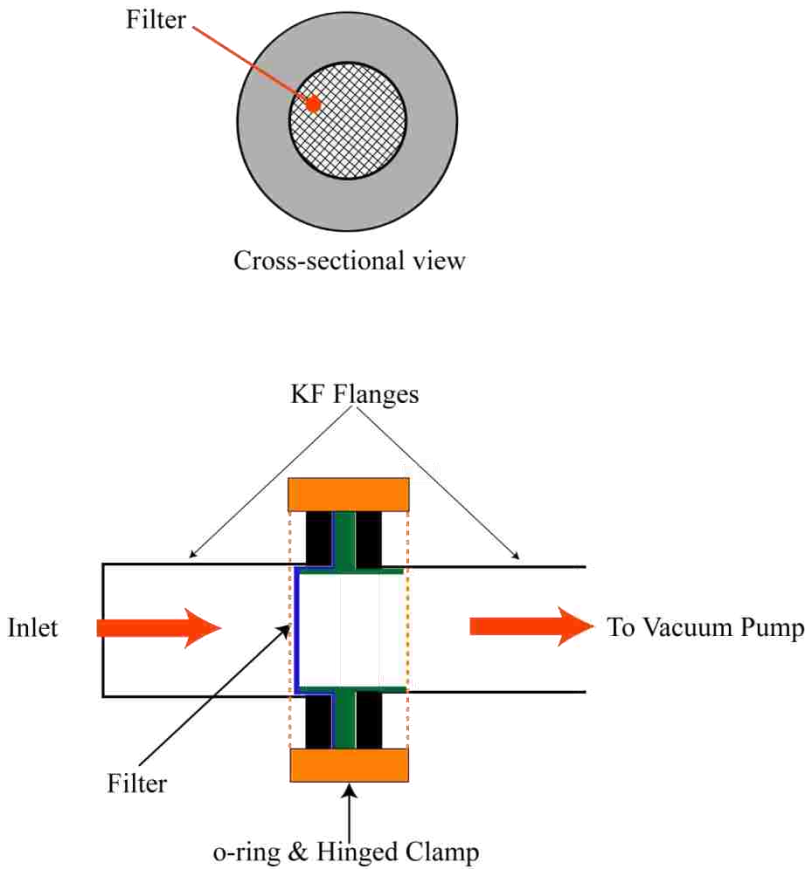


Figure 13. Filter Paper sampler.

For filtration sampling, sample viability is usually lost due to exposure to air and mechanical forces during collection. Since our analysis does not require viability, this was not a concern. Fibrous filters all have some common characteristics. The first characteristic is the velocity of the air stream through the filter sampler. At the face of the filter, the velocity is termed U_0 and can be calculated from Equation 5.[1] As the air flows through the filter, the volume available for air flow decreases due to the presence of the filter fibers. Since Q is constant across the filter, the velocity of the air flow must increase and can be calculated from Equation 6 where α is equal to the volume fraction of fibers or packing density.[1] In addition, the increase in air velocity results in a decrease in pressure as a result of the Bernoulli principle as shown in Equation 7.[1] For the Bernoulli principle's application to a system, P is equal to the static

pressure, ρ is equal to the air density, v is equal to velocity, g is equal to gravitational acceleration, and h is equal to the height change of the system. Since the airflow is horizontal through the filter in our experimental setup, the gravitational term reduces to zero. Therefore, an increase in air velocity will result in a decrease in pressure which is termed ΔP or pressure drop. For all practical purposes, ΔP is directly proportional to the thickness of the filter, which means thicker filters are more efficient than thinner ones of the same material.

$$U_0 = \frac{Q}{A} \qquad \textbf{Equation 5}$$

$$U_0 = \frac{Q}{A(1-\alpha)} \qquad \textbf{Equation 6}$$

$$P + \frac{1}{2}\rho v^2 + \rho g h = Constant \qquad \textbf{Equation 7}$$

Collection of particles by filtration occurs through five distinct mechanisms: interception, inertial impaction, diffusion, gravitational settling, and electrostatic attraction. The last mechanism is extremely difficult to quantify, but usually only contributes if charging of the particles and filter was done. Nevertheless, each mechanism can be described separately with respect to collection efficiency and summed to give a simplified expression for overall filter collection efficiency. Typically, collection efficiency increases initially after the start of collection because the volume within the filter decreases as a result of particle loading on the filter fibers. The resulting increase in collection efficiency and decrease in ΔP will continue as particle loading increases; however, the filter will eventually reach a clog point where efficiency drops dramatically.

Along with collection efficiencies, the analytical application dictates the type of filter based on compatibility with solvents and chemical interferences. In this work, analysis of the samples from the filter was done using three different techniques: solvent extraction, adhesive

removal and direct analysis. For the solvent extraction of particles, two approaches were used. In the first approach, approximately 1 to 2 µl of either solvent or matrix was deposited onto a bare stainless steel MALDI target. Before the droplet dried, the filter was placed face down onto the liquid for approximately 10 seconds. After this time, the filter was removed and matrix solution was added to the MALDI target. In the second approach, the filter was placed face down on the MALDI target and equivalent solvent or matrix was added to the backside of the filter. Once the filter was removed, additional matrix was added to the MALDI target. For the adhesive removal, double sided tape was adhered to the front side of the filter. The tape was then removed and adhered to a MALDI target. Matrix was then added prior to analysis. For the last technique, direct analysis, double sided tape was adhered to a MALDI target and the filter was placed back side down onto the tape. Once attached, matrix was added. All samples were dried before introduction into the mass spectrometer for analysis

2.5 Mass Spectrometric Analysis

Mass spectrometry is a versatile and robust analytical technique that is capable of providing information regarding the chemical composition and mass of a sample. A mass spectrometer consists of three main parts as shown in Figure 14. In mass spectrometry, gaseous

Figure 14. Components of a Mass Spectrometer.

ions are separated in space or time resulting in spectra with peaks based on the on mass-to-charge ratio (*m/z*) of the ion. In order to perform mass spectrometry, the analyte must be in the gas phase. For large biological compounds, this presents a unique challenge. In addition, since

most ionization techniques such as electron ionization, impart a large amount of energy (greater than 70 eV) to the sample, the molecule's covalent bonds break, creating fragment ions from the analyte. Since most biological molecules are polymers of repeating units, this type of fragmentation would produce spectra with limited information for analyte identification. In 2002, John B. Fenn and Koichi Tanaka were awarded the Nobel prize in chemistry for their work on the *soft ionization* techniques, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), respectively. [93, 94] These techniques make it possible to analyze large intact biomolecules such as oligonucleotides, peptides/proteins and lipids with limited to no covalent bond fragmentation.[93, 94]

In this work, two mass spectrometers, an axial MALDI (MALDI) time-of-flight (TOF) (Bruker OmniFlex) and an orthogonal MALDI (oMALDI) quadrupole-time-of-flight (QTOF) mass spectrometer (ABI QSTAR) were used. Both of these instruments are commercial and are equipped with nitrogen lasers (λ= 337nm) for desorption/ionization of analytes. A description of each instrument is provided in the subsequent sections. All analyses were done offline from the bioaerosol collections.

2.5.1 Matrix-Assisted Laser Desorption/Ionization

MALDI mass spectrometry is a *soft ionization* technique designed to produce gas phase ions of biological macromolecules including lipids, peptides/proteins, and oligonucleotides.[95-97] In MALDI, a sample is deposited onto a solid target sample target such as stainless steel in combination with an excess of a molecule termed the matrix. The matrix and analyte then crystallize on the surface for analysis. The matrix assists in the ionization and desorption of the biomolecule into the gas phase.[98] For a molecule to serve as a matrix, it must have several characteristics. First, the matrix must absorb light at the wavelength of the laser. Second, the

matrix must be soluble in a solvent that is compatible with the analyte and promote co-crystallization. Lastly, the matrix must be able to influence the co-desorption and ionization of the analyte upon absorbing laser light. Furthermore, if operating conditions require low pressures as in vacuum MALDI, the matrix must be vacuum stable. Table 7 lists some common matricies and their respective analytes.

Table 7. Some common MALDI matricies[99]

Matrix Name	**Abb.**	**Laser**	**Applications**
α-cyano-4hydroxucinnamic acid	CHCA	337/355	Peptides, lipids, nucleotides
4-hydroxy-3-methoxycinnamic acid	FA	266/337/356	Proteins
2,5-dihydroxybenzoic acid	DHB	266/337/355	Oligos, peptides, nucleotides
3,5-dimethoxy-4-hydroxycinnamic acid	SA	266/337/355	Lipids, peptides, proteins
3,4-dihydroxycinnamic acid	CA	337/355	Peptides, proteins, lipids
2[4-hydroxyphenlyazo]benzoic acid	HABA	266/337	Proteins, Lipids
3-hydroxypicolinic acid	HPA	337/355	oligonucleotides

Once the analyte is co-crystallized with the matrix on the MALDI target (Figure 15), the target is then inserted into the MALDI mass spectrometer. Ionization is accomplished by

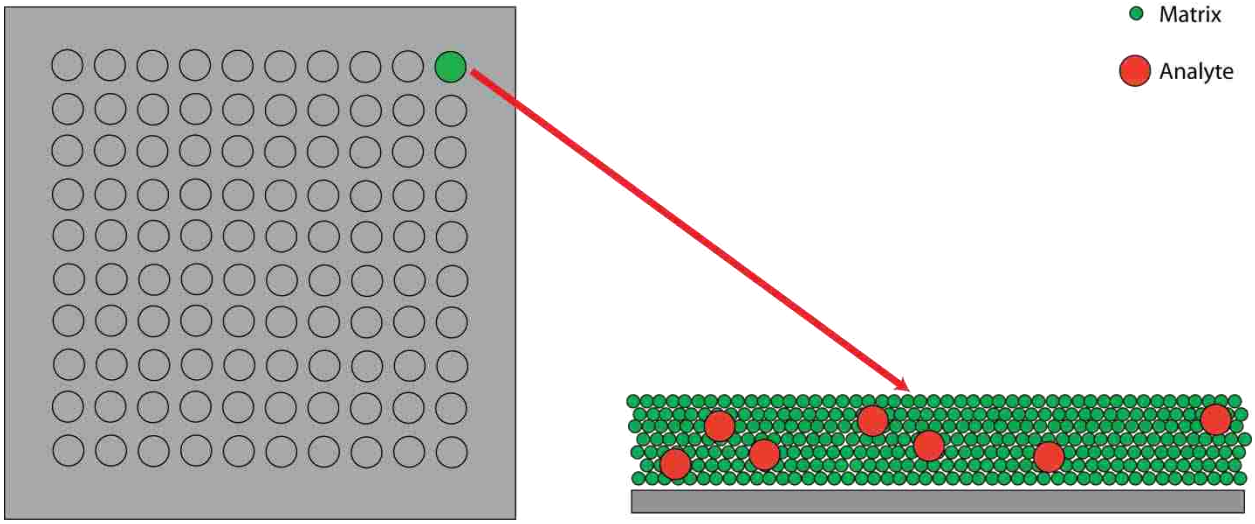


Figure 15. Matrix Deposition illustration on a 10 x 10 target.

directing a pulsed a laser (0.5 to 25 ns) onto the target sample location as shown in Figure 16. The absorbed light energy is absorbed by the matrix causing desorption. The amount of energy per unit area, fluence (H), for the MALDI process is in the range of 20 to 200 J m^{-2}.[100] Since the

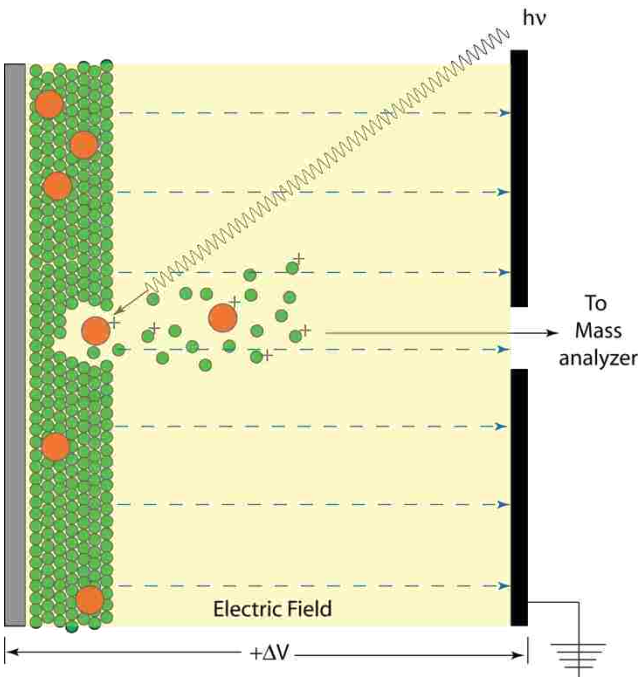


Figure 16. MALDI Ionization/Desorption Illustration.

analyte is embedded within the matrix during co-crystallization, the analyte molecules are desorbed and are entrained in the gas-phase matrix plume. Ionization is thought to occur within the plume by gas-phase proton transfer reactions without significant fragmentation of the biological molecules.[101] The generated ions are then guided into the mass analyzer such as time-of-flight (TOF), where they are separated. Resolution enhancing techniques such as delayed extraction and reflectrons, are used to improve mass resolution by MALDI as ionization often occurs in different temporal and spatial locations.

Two types of MALDI ionization sources were used in this work, axial and orthogonal. For the axial configuration, the MALDI target is perpendicular to the mass analyzer and ions are extracted directly into the flight tube. This configuration is designed to be used in either linear or reflectron time-of-flight modes. For the orthogonal MALDI ionization source, the MALDI target is perpendicular to the initial mass analyzer or ion guides. However, the flight tube is positioned

orthogonal to the original ion beam trajectory. Both of these instrument setups are described in the subsequent sections.

2.5.2 Mass Analyzers

The two types of mass analyzers used in this work were a time-of-flight mass spectrometer and a tandem quadrupole-time-of-flight mass spectrometer. The MALDI-TOF is capable of operating in both linear and reflectron modes while the oMALDI-QTOF instrument operates in reflectron mode. In the subsequent paragraphs, a brief description of quadrupoles, time-of-flight and reflectrons will be given before proceeding to the instrument descriptions.

A quadrupole mass analyzer consists of four parallel rods (2 pairs) whose long axis is parallel to the ion beam trajectory from the ion source. Ions are accelerated out of the source and into the quadrupole mass analyzer where a direct current (DC=U) and a radio frequency (rf=Vcos(ωt)) is applied to each of the two pairs of rods as indicated in Table 8. As the ions

Table 8. Quadrupole rod settings.

Rod Pair	Quadrupole Settings
1	+[U+ Vcos(ωt)]
2	-[U+ Vcos(ωt)]*

** rf is 180 degrees out of phase with respect to that of rod 1.*

travel through the quadrupole, they are subjected to the quadrupole field. The quadrupole field can be altered to control operation by either controlling the potentials (U/V) or rf frequency. Traditionally, the potentials are varied while the frequency is kept constant. The ion path within the field is either stable and ions exit the quadrupole region to reach the detector, or unstable and ions will not reach the detector.

Within the quadrupole, the potential, Φ, at any point within the quadrupole field can be calculated from Equation 8 where r_0 is the inscribed radius between the rods and x and y are the distances from the center of the field. Within this field, the motion of an ion is described using

the Mathieu equation as shown in Equation 9. In this equation, u represents the transverse displacement in the x and y direction from the center of the field, ξ is equal to ωt divided by 2 and the variables a and q are dimensionless numbers defined by Equations 10 and 11, respectively.[99] A generalized stability diagram with respect to a and q corresponding to the first

$$\Phi_{(x,y)} = \Phi\left[\frac{x^2 - y^2}{r_0^2}\right] \qquad \textbf{Equation 8}$$

$$\frac{d^2u}{d^{2\xi}} + [a_u - 2q_u cos(2\xi)]u = 0 \qquad \textbf{Equation 9}$$

$$a_u = \frac{8eU}{m\omega^2 r_0^2} \qquad \textbf{Equation 10}$$

$$q_u = \frac{4eV_0}{m\omega^2 r_0^2} \qquad \textbf{Equation 11}$$

stability region is shown in Figure 17. In this figure, the mass scan lines are generated by adjusting the ratios of the DC potential and AC amplitude while keeping the *rf* frequency

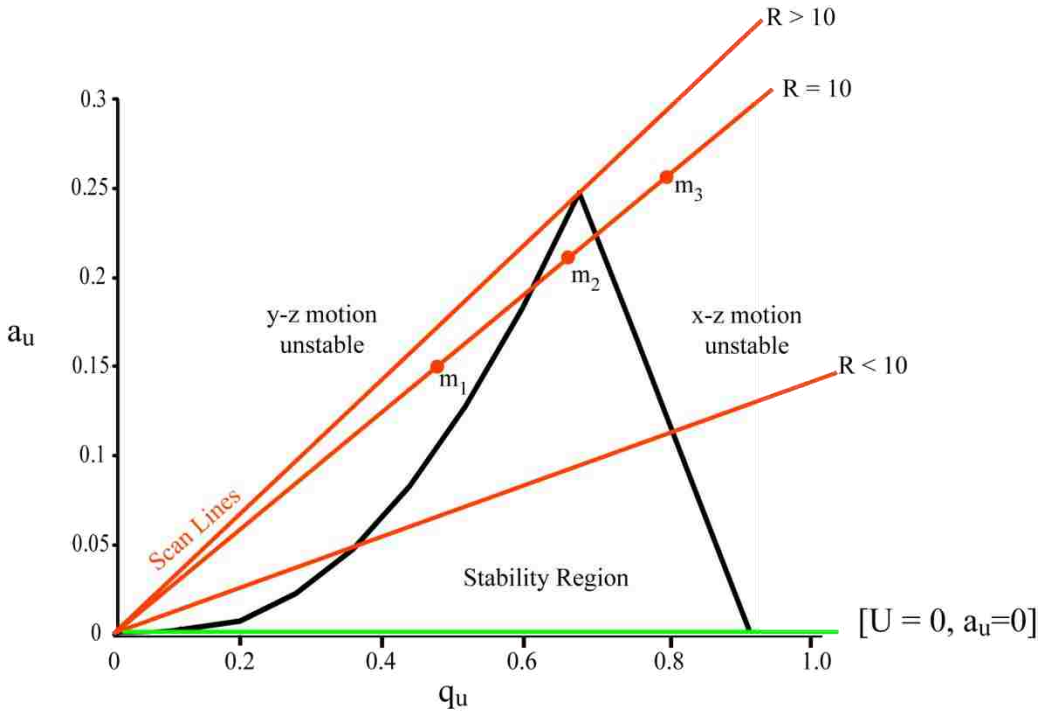
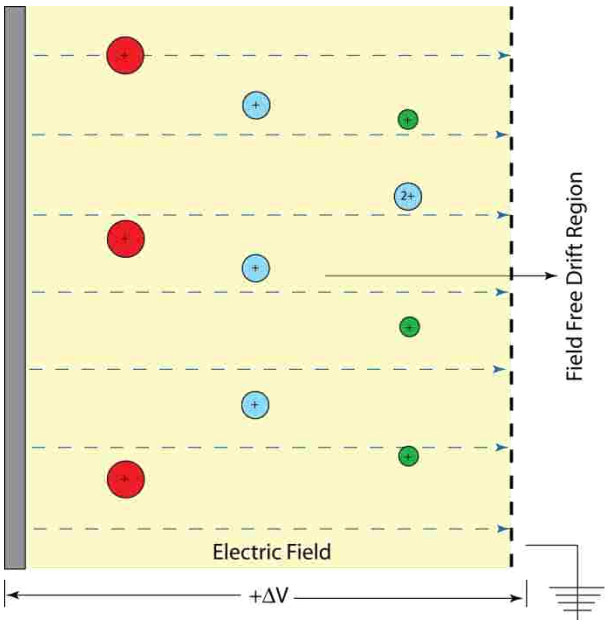


Figure 17. First stability region of a quadrupole mass spectrometer.

constant. As shown, increasing the slope of the scan line, results in a decrease in the *m/z* range falling within the stability region. For the scan line of R equal to 10, masses m_1 and m_3 have unstable trajectories and not reach the detector whereas mass m_2 will. In essence, this allows the quadrupole to serve as a narrow *m/z* bandpass filter.

On the other hand, if the quadrupole is in tandem to another mass analyzer, it is often desired that the largest number of ions pass through the quadrupole. In this mode of operation, the quadrupoles serve as an ion guide. This operation can be achieved while operating the quadrupoles in *rf* mode only. In *rf* mode only, the DC component, U, is zero. When U is equal to zero, a_u also is zero resulting in a horizontal mass scan line. This means that the scan line runs along the q_u axis resulting in a large *m/z* range falling within the stability region of the quadrupole field. When talking about tandem mass spectrometers that incorporate quadrupoles, there is specific terminology to indicate which mode of operation the quadrupole operates. A q is used to represent a quadrupole operating in *RF* mode only such as ion guides, and Q is used to represent a quadrupole operating as a mass filter. The QTOF mass spectrometer used in this work is actually a qQqTOF instrument, however, for simplicity, the instrument will be called a QTOF. For the QTOF, the mass selecting quadrupole can easily be switched between ion guiding for TOFMS mode and mass filtering for MS/MS mode.

The other type of mass analyzer used in this work was a time-of-flight mass analyzer. In TOF, ions are generated in the source and then pulsed out into the TOF flight tube. This pulse out is accomplished by exposing the ions to a uniform electric field as shown in Figure 18. Inside the electric field, all the ions experience a force in either the same or opposite direction of the electric field depending on their charge. This force causes the ions to accelerate until they reach



$$v = \sqrt{\frac{2qV}{m}}$$

q = 1.60x10^{-19} C
m = M.W. x 1.66x10^{-27} kg
V = Voltage (volts)

For a 1 kDa positive ion with +1 charge and a 3 KV potential, v = ~24 Km/s

Figure 18. Ion extraction for TOF mass analysis.

the field free drift region. Upon exiting the electric field, each ion has been accelerated to a final velocity based on its initial spatial position and velocity as well as its mass and charge. If initial spatial positions and velocities are neglected, all ions with the same *m/z* have the same kinetic energy (KE) and same velocity upon exiting the ion source. Once the ions enter the field free drift region, they no longer experience any forces or undergo any collisions since the mean free path is greater than the length of the flight tube. When the ions hit the detector, their flight time is recorded. As shown in Figure 19, if the flight time is known, then the time can be used to calculate the *m/z* ratio of the ion. From this, mass spectra of time versus ion intensity can be constructed. If the TOF is calibrated, the times can be converted into *m/z* values to generate mass spectra. TOF mass spectrometers provide good resolution and a theoretical unlimited *m/z* range. However, there are a few factors that effect the TOF resolving power, which is the ratio of the mass of a peak divided by its width.

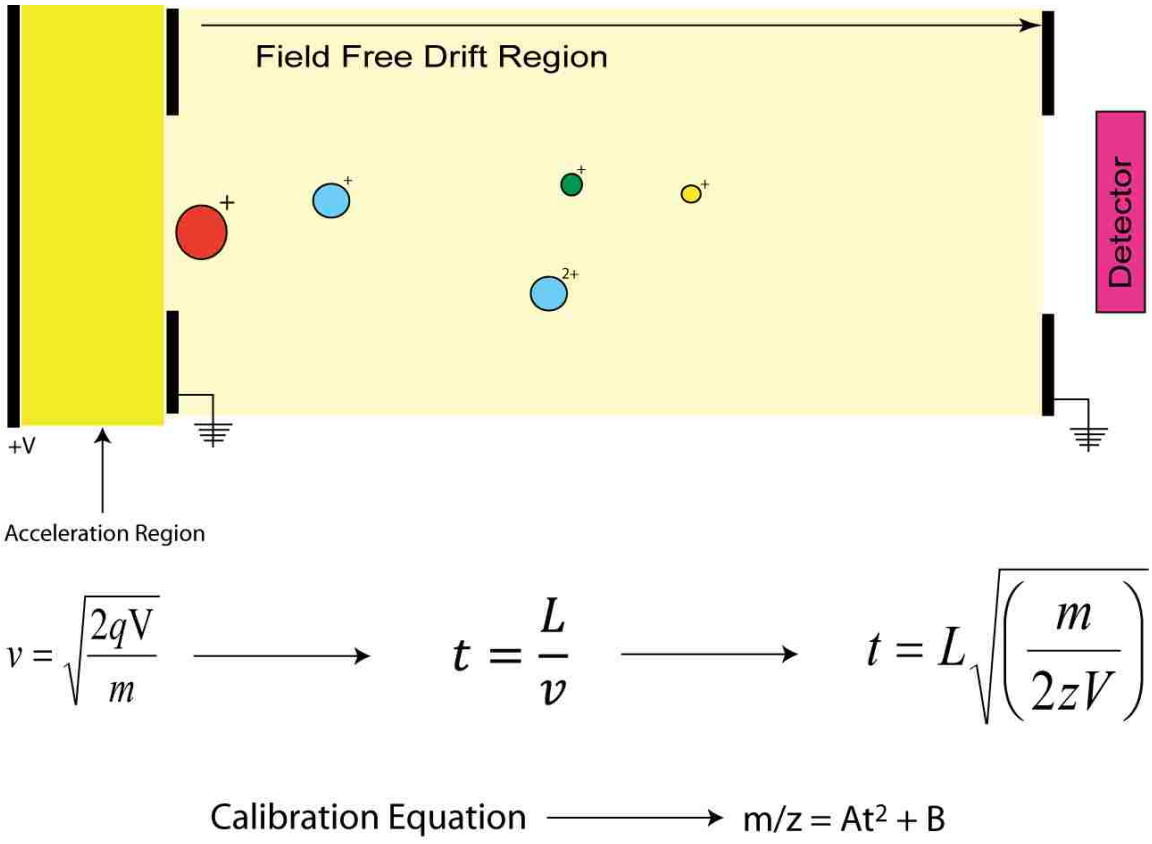


Figure 19. TOF theory and drift time calculations.

With TOF, the resolving power is based on the simultaneous detection of ions of the same *m/z*. Since TOF is based on ion velocities, a small spread in kinetic energy is essential in achieving good resolving power. The two factors that cause differences in kinetic energies are the initial spatial position and velocity of the ions. These two factors are extremely important in MALDI considering the irregular surface of the matrix crystal creating a non-uniform spatial distribution and the uneven distribution of initial velocities as a result of collisions during desorption. To correct the resolution problems, two techniques are used, delayed extraction (DE) and the reflectron.

In DE, the application of the electric field is time delayed from ion formation.[102] In MALDI, the laser pulse desorbs and ionizes analyte molecules. Each ion has a mass-independent velocity, v_0, away from the crystal surface as illustrated by t_0 in Figure 20. This creates a

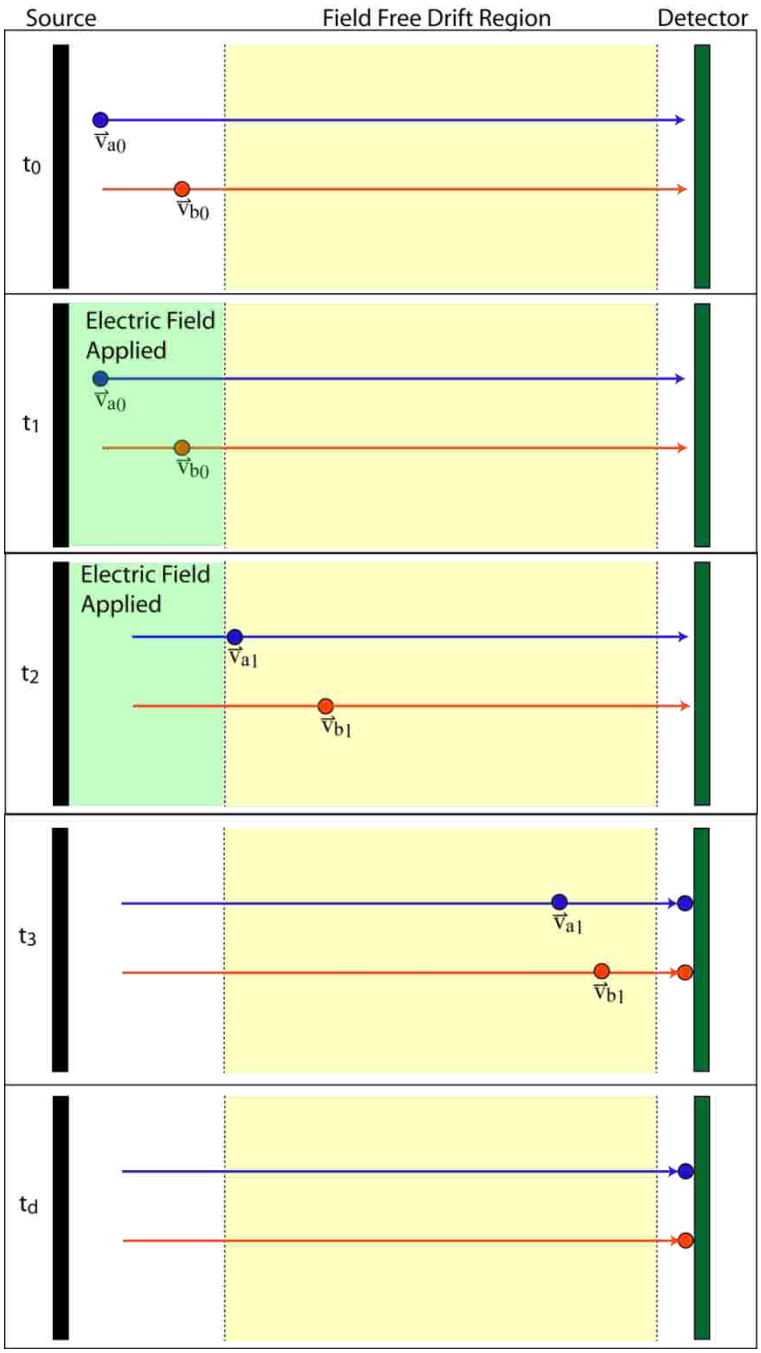


Figure 20. Delayed Extraction.

distribution of velocities among ions with the same *m/z* ratios. In the absence of a delay, the kinetic energies of the ions with same *m/z* values are different, resulting in low resolving power. By delaying the application of the electric field for ion extraction, the ions with greater v_0 (illustrated by the ion with velocity v_{b0} in Figure 20) is farther away from the surface than those

with smaller velocities,v_0, as shown by v_{a0}. As a result, when the extraction field is applied, the ions with greater v_0 spend less time in the electric field and has less kinetic energy when entering the field free drift region. Consequently, ions with smaller v_0 spend more time in the electric field and have more kinetic energy when entering the field free drift region. Initially, v_{a0} is less than v_{b0}. However, after entering the field free drift region, the relative velocities is reversed, i.e., v_{a1} is greater than v_{b1}. Therefore, at some point in the flight tube, ion a and ion b will become *space focused* such that the travel time to a given distance is equal. If the detector is placed at this distance, then each *m/z* ion, reaches the detector at the same time regardless of its initial energy. Since this effect is mass dependent, the delay time can be varied and adjusted to increase the resolution for a particular mass range.

The other resolution enhancing technique is ion reflection.[103] While there are three types of reflectrons generally available, a single-stage, a dual-stage, and a curved field, only single stage reflectrons will be discussed to illustrate the general principle of ion reflection for resolution enhancement. The reflectron, or ion mirror, is composed of a series of ring electrodes to which voltage gradient is applied. For single-stage reflectrons, the voltage gradient is constant thereby creating a uniform electric field near the center of the electrode rings. Ions with the same *m/z* and different kinetic energies within the flight tube must be refocused such that they are able to reach the detector simultaneously as shown in Figure 21. In this figure, the identical *m/z* ions, A and B, have different kinetic energies upon ion extraction. In Figure 21, the KE of ion A is greater than ion B, therefore, v_a is greater than v_b. If not focused, the resolution suffers because the ions do not reach the detector simultaneously. A reflectron can be used to compensate for these differences and enhance resolution. As the ions exit the field free drift region and enter the reflectron region, those with higher kinetic energies (A) travel deeper into the reflectron field

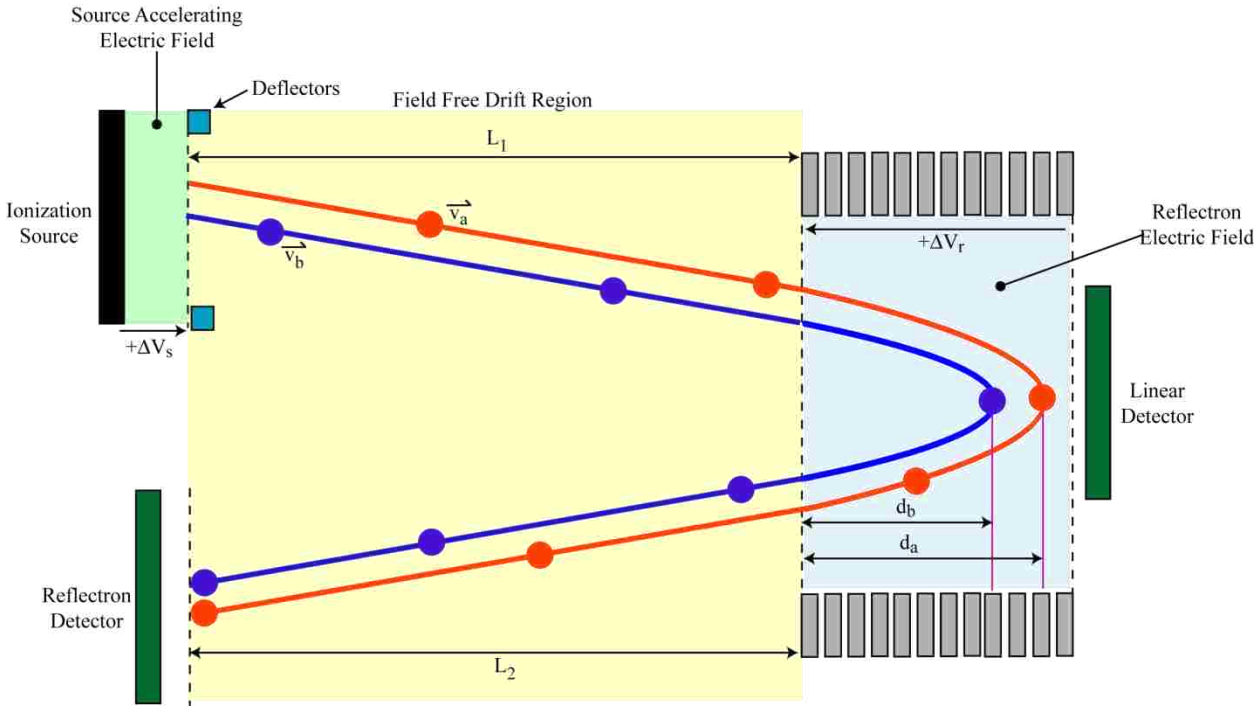


Figure 21. Reflectron TOF configuration.

before they are reflected back out. At the same time, the ions with less kinetic energy (B) travel a shorter path within the reflectron field before being reflected. As a result, the total path length in the flight tube for each ion is different. The flight path for Ions A and B are given by Equations 12 and 13, respectively.[99] These equations can be used to calculate the total flight time for an ion in a reflectron TOF as shown by Equation 14. Therefore, the differences in the total path length and KE of identical ions are corrected resulting in the refocusing identical *m/z* ions at the detector and an increase in the resolution. In addition, a reflectron can also increase the effective length of the TOF flight tube which also increases resolution.

$$S = (L_1 + L_2) + 4d_a \qquad \textbf{Equation 12}$$

$$S = (L_1 + L_2) + 4d_b \qquad \textbf{Equation 13}$$

$$t = \left(\frac{m}{2zV}\right)^{1/2} (L_1 + L_2 + 4d) \qquad \textbf{Equation 14}$$

2.5.3 Mass Spectrometer Instrumentation

The first mass spectrometer used was a Bruker OmniFlex, an axial MALDI-TOF (MALDI). The MALDI-TOF is capable of both linear and reflectron modes of operation. It is equipped with a nitrogen laser (λ =337 nm, 3 ns pulse width) operated at 1 Hz. For ion extraction, an accelerating voltage of 19 kV was used and operated with delayed extraction to increase resolution. In all experiments, signal suppression was used and consists of a pair of deflection plates turned on and off to deflect the desired mass range from the detector. In this work, three layouts of stainless steel targets were used: 7x7, 10x10, and a special modified target for use with filter paper. On average, 25 laser shots were collected to generate mass spectra from samples. The laser energy was adjusted to about 10% above threshold for the chosen matrix as determined through the analysis of peptide standards.

The other mass spectrometer used was an Applied Biosystems QSTAR XL, which is a qQqTOF (QTOF). The QTOF mass spectrometer is equipped with a N_2 laser (λ = 337 nm, 3 ns pulse width) operated between 1 and 40 Hz. This instrument is equipped with a single-stage reflectron. While the QTOF was designed to be an electrospray instrument, the instrument is equipped with removable ionization sources including an orthogonal MALDI (oMALDI) source.[104] A general diagram of the QTOF is shown in Figure 22. This tandem mass spectrometer is capable of performing TOF-MS as well as MS/MS experiments. In TOF-MS modes, all three quadrupoles are operated in *RF* mode only, whereas in MS/MS mode, Q_1 is operated as a mass filter. In addition, q_2 is located within the collision chamber and guides all ions, precursor and products, into the TOF pulser. As the ions exit, they are reaccelerated and focused by ion optics into a narrow continuous beam which is pulsed out orthogonally from the original ion beam direction into the TOF tube. The orthogonal pulser operates at a specified

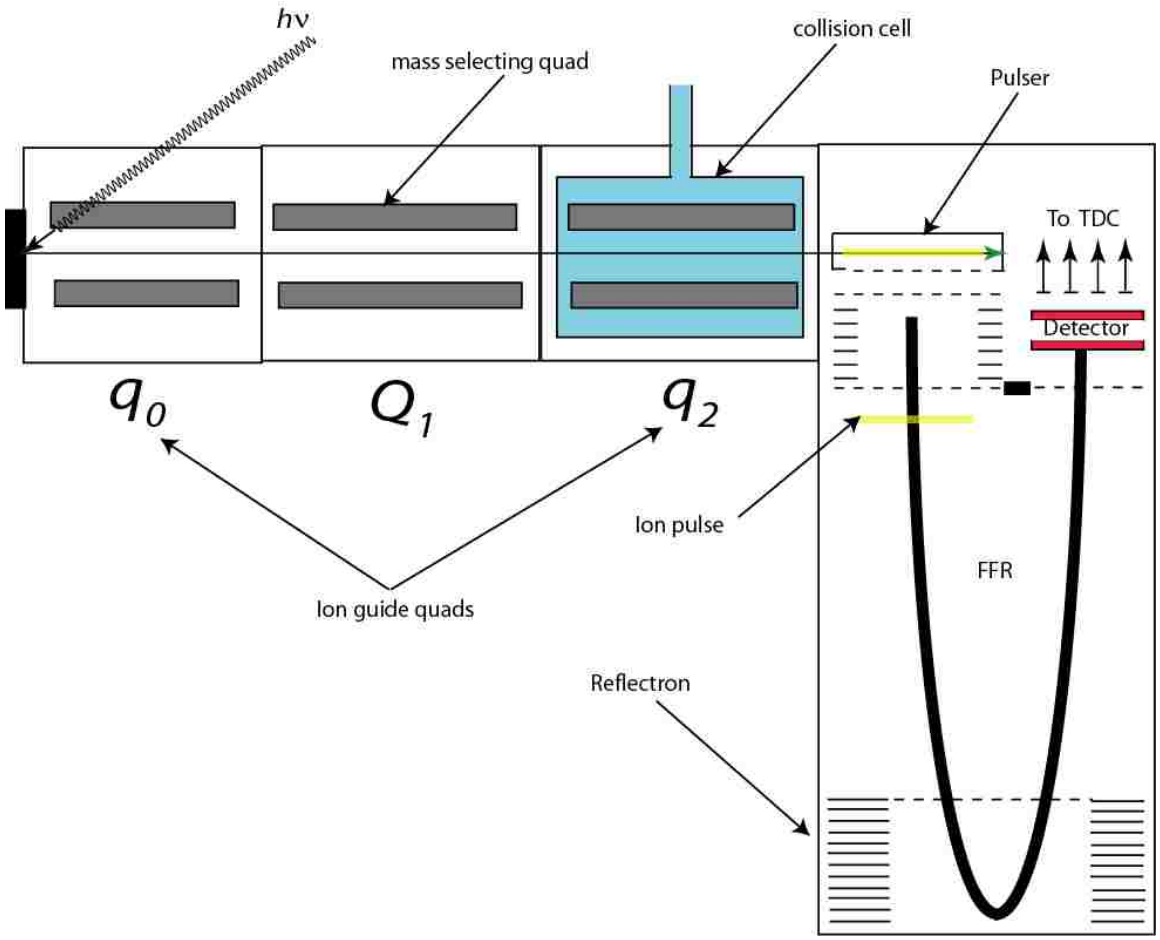


Figure 22. QTOF mass spectrometer diagram.

repetition rate as determined by either the *m/z* range set for Q_1 for MS/MS experiments or the maximum *m/z* range set for TOF-MS experiments. The ions are pulsed out using an electric field of 10 kV at a repetition rate of several kHz. The repetition rate is determined by the software to allow the highest *m/z* ion to reach the detector before a subsequent pulse is initiated. For the QTOF, the upper *m/z* limit is 40 kDa for TOF-MS mode and 6 kDa for MS/MS mode.

Both mass spectrometers used are equipped with a channel electron multiplier array or what is more commonly called a microchannel plate (MCP). The MCP detectors are actually a pair of MCPs constructed in a chevron configuration to prevent ions from transiting the channel. Each MCP has a gain of approximately 10^4 illustrated by Figure 23. The signal is received by anodes mounted after the MCP. For the QTOF, there are four discrete anodes. Each anode is

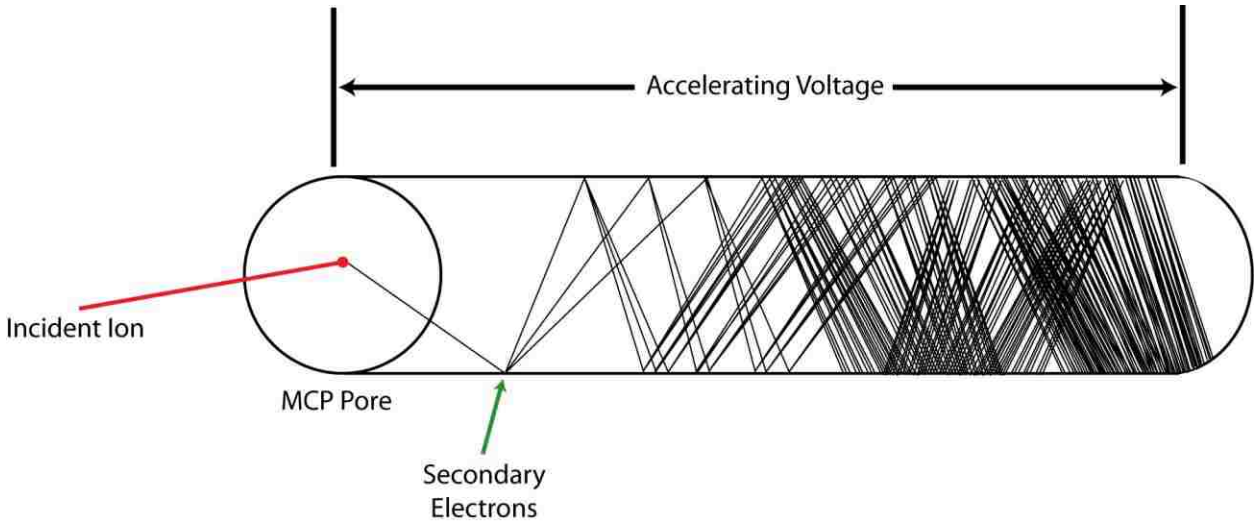


Figure 23. Illustration of electron multiplication inside an MCP.

connected to a separate pre-amplifier and discriminator which increases the dynamic range of the detector. A time-to-digital converter converts each single ion's pulse into digital value. Since MCPs can only detect one ion event per multiplication signal, the detector has to recover before a second event can be detected. This is called the dead-time. For the QTOF, the dead time is minimized by dividing the MCP into four sections with each section having a separate anode. Each anode has its own channel in the TDC providing up to four simultaneous ion detections per dead time unit. Since each TOF pulse results in a separate mass spectrum, creating a complete mass spectrum is done by summing each of the collected spectra from the entire set of TOF pulse cycles.

The greatest benefit of using the QTOF is the ability to perform MS/MS experiments. For MS/MS mode, the maximum *m/z* possible for selection is 6 kDa. To conduct MS/MS experiments, an *m/z* value is first selected for filtering. Next, the *m/z* filter width is selected as either low resolution (the entire precursor isotope series), high resolution (only the precursor and limited isotope series i.e., within ± 1 *m/z* units), or unit resolution (only allowing the precursor's exact *m/z* through). Once this is determined, the selected *m/z* range passes through Q_1 to enter the

collision cell, q_2 which is being operated in *RF* mode only. While traversing the collision cell, the analyte ions collide with an inert gas such as Ar (MALDI) or N_2 (ESI). When the collisions produce fragmentation, it is called collision induced dissociation (CID).[105] The resulting fragments produced by CID are then guided by q_2 into the orthogonal pulser for extraction into the TOF mass analyzer.

2.5.4 Collision Induced Dissociation

Collision induced dissociation is the term used to describe fragmentation resulting from inelastic collisions between analyte ions and neutral inert gases. In this inelastic collision, some of the Kinetic Energy (KE) from the collision is transferred into the internal energy of the colliding molecules. The transferred energy is converted into vibrational and electronic modes within the excited ion causing the ions to fragment. Since this is an energy driven process, the easiest bonds, the bonds with the lowest energy, would break first followed by the next weakest and so forth until the molecule is no longer energized. This process continues for all collisions occurring as the ions move through the collision cell including product ions. Therefore, as ions are fragmented, the initial fragments collide and produce fragments as illustrated in Figure 24.

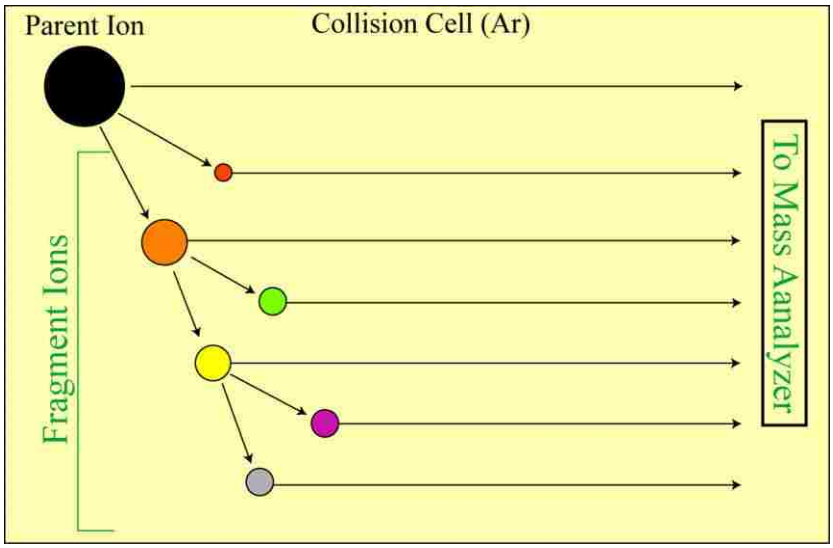


Figure 24. Illustration of Collision Induced Dissociation with Ar gas.

The intensities of the parent and product fragments are based on several factors all related to collision frequency and energy and include the ease of fragmentation for both precursor and products, length of time in collision cell (related to initial ion KE), concentration of collision gas and the overall concentration of the precursor ion.

For CID, there are two energy regimes useful in mass spectrometry, low energy CID (collision energy < 200 eV) and high energy CID (collision energy > 1 keV).[106, 107] In this work, low energy CID was used with Ar as the collision gas to fragment biological samples of polypeptides. For polypeptides, there are three different equivalent bonds across the chain backbone as shown in Figure 25 (excluding N-and C-terminus). When any of these bonds are

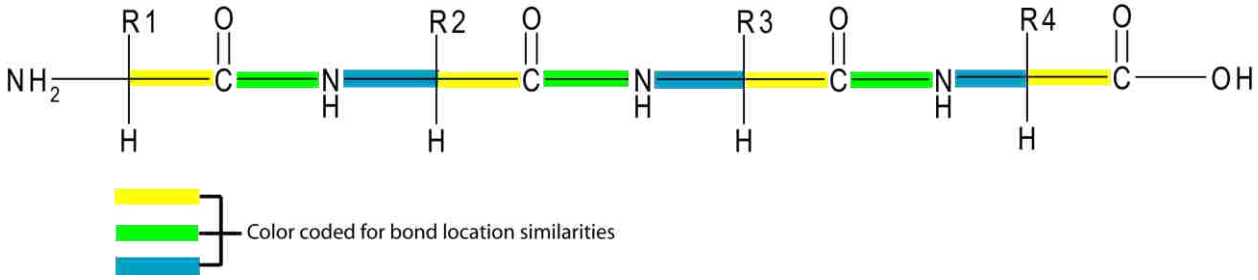


Figure 25. Peptide Bond Backbone Similarities Across a Polypeptide.

cleaved during CID, only one of the two fragments retains the ion's initial charge. When the charge is retained by the N-terminus, the ions are labeled, a, b and *c* ions and when they are retained by the C-terminus, the ions are labeled x, y, and z.[108] Internal fragmentation is also possible. Here, two peptide bonds are broken with the internal peptide fragment retaining the charge. In this case, the nomenclature combines the two cleavages, one N- terminal and one C-terminal type. To indicate which bonds are broken, subscripts are used in the following format $(b_3x_4)_2$. In this format, the subscripts inside the parentheses indicate the distance from terminus and the subscript outside indicates the length of the internal fragment. Figure 26 illustrates the fragmentation nomenclature used for polypeptide analysis by CID. By using these fragmentation patterns, the primary sequence of the peptide (the amino acid residue sequence from N- to C-

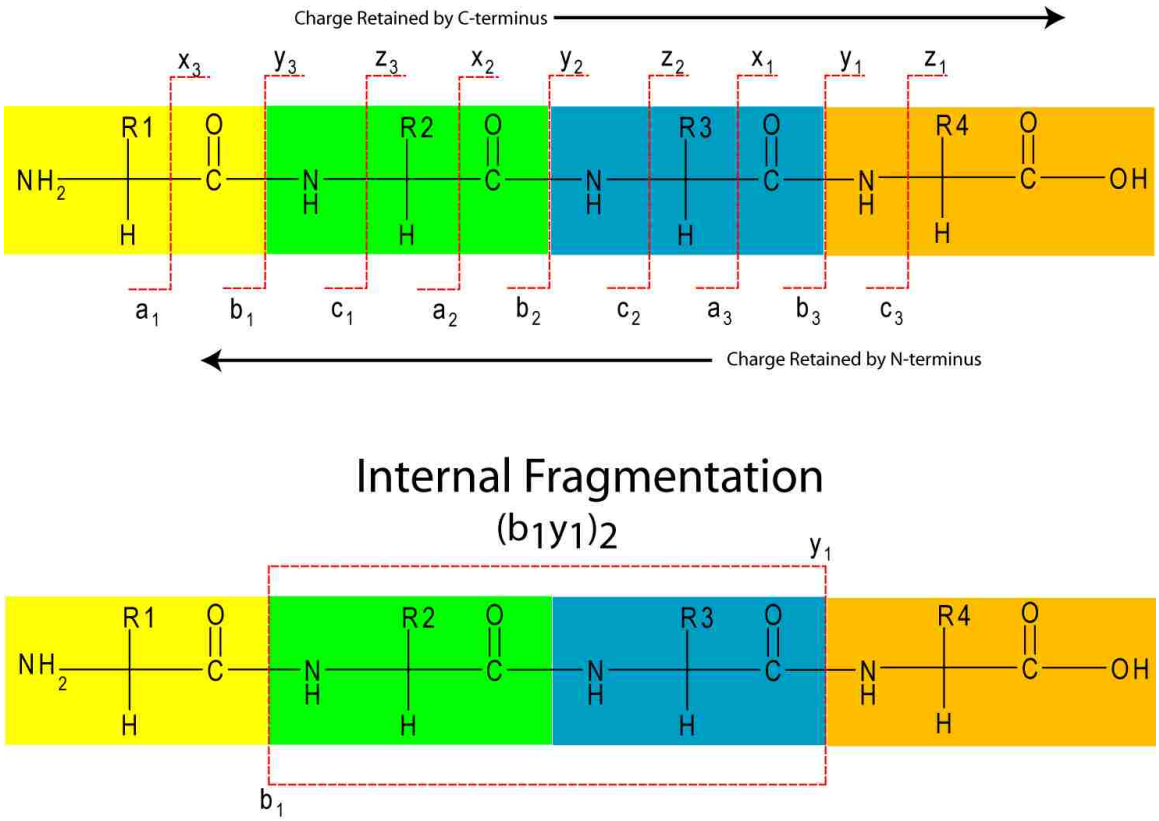


Figure 26. Polypeptide CID Bond Fragmentation Terminology.

terminus) or a sequence tag (a stretch of amino acid residues with missing terminal mass values) can be determined from tandem MS.

2.6 Intact Whole-Cell Bacteria Analysis

For intact whole-cell analysis, solutions of *E. coli*, *B. subtilis*, and *B. thuringiensis* were prepared by adding 20 mg of the lyophilized bacteria in 1 mL of water. The bacteria solutions were then sonicated for 1 minute without heat followed by a 5 minute sonication without heat. Several MALDI sample preparation techniques were tried. However, with the motivation of analyzing bioaerosol samples, the dried-layer method was used. In this method, 0.5 – 1.0 μL of sample was deposited onto the MALDI target and allowed to dry. Once dried, 0.5 μL of a 1% TFA solution was added and allowed to dry. Once the solvent was dry, 0.5 to 1.0 μL of the appropriate matrix solution was added and allowed to dry before analysis. Whole-cell analysis

was performed on both the MALDI-TOF and oMALDI-QTOF mass spectrometers. An equivalent analysis was done on intact proteins to simulate the presence of bacteriotoxins in bioaerosols. For the protein analysis, a 1 mM stock of the protein was created, and from this, 0.5 μL of the sample was added to the target and prepared using the dried-layer method described above.

2.7 Proteomic Analysis of Bioaerosols

MALDI-MS can be used to detect all four major classes of biological macromolecules: proteins, nucleic acids, lipids and saccharides. While this work focuses on protein analysis from intact whole-cell microorganisms, molecule types from mass spectral peaks are initially not determinated except for the fact that the MALDI matricies chosen demonstrate a propensity for desorption and ionization of proteins and peptides.

In the work described below, the method used to identify microorganisms from bioaerosols is proteomics. Proteomics is the study of protein structure, function and expression for an organism.[109] All proteins are encoded by an organism's genome or DNA. While an organism's genome is constant (except for random mutations), its proteome is not. As an example, in humans, the proteome of a liver cell is different than that of a nerve cell despite each cell having identical genomes. The same is true for bacteria, although not as pronounced. For bacteria, their proteome can change in response to environmental factors, growth requirements and metabolic needs.[110, 111] Proteins are also responsible for diverse set of biological functions such as immunity, metabolism, and cell signaling, and are found localized throughout the a cell, and in the case of eukaryotes, even inside the organelles (a compartment within a cell containing a separate membrane) within the cell. In addition, some proteins are so vital that they were conserved from some distant evolutionary ancestor which creates the probability of significant

sequence homology (amino acid sequences with a high degree residue similarities) between organisms from different species. Despite this complexity, every organism, at any given time, expresses a distinct proteome.

2.7.1 Generation of E. coli Protein Database

The National Center for Biotechnology Information (NCBI) contains a database for identified proteins cataloged by taxa. In addition, NCBI contains genomic sequences, partial and complete. There are several *E. coli* fully sequenced genomes available. From these sequenced genomes, translated databases (protein databases produced by *in silico* translation of genomic data into proteins using standard open reading frames) are constructed. The translated protein sequences are then stored in FASTA format, which is a standardized text format for representing DNA and protein sequences. A FASTA sequence extractor (FASE) program (Appendix C) was written to parse the data from the translated databases for population into an in-house protein database. The FASE program populates the database while calculating molecular masses from the primary sequences. The major limitation of FASE or any protein database populated using FASTA data is the inability to extrapolate post-translational modifications.

2.7.2 Protein Database Searching and Mass Spectral Fingerprinting

The analysis of proteins by mass spectrometry can be used to characterize the proteome of an organism. This can be done on either intact whole-cells or cell lysates. The detected proteins can be used to create a mass spectral library for the organism. Once the library is populated by enough data from different organisms, peaks unique to a specific organism can be identified and labeled as biomarker for the organism. Mass spectra from unknown samples can then be searched for these biomarkers to identify a match within the mass spectral library using a peak based probability algorithm.[112] This technique is called mass spectral fingerprinting (MSF)

due to its similarity to forensic fingerprinting. Another caveat of MSF is the ability to store entire mass spectra including non-biomarker peaks. Comparisons between unknown mass spectra and the library will reveal similarities. The simplest comparison can be done by calculating a cross-correlation coefficient score based on Equation 15. Despite the simplicity,

$$r = \frac{\sum_i (x_i - \bar{x})\left((y_i - \bar{y})\right)}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \qquad \textbf{Equation 15}$$

this approach requires x-axis data points to be identical between the compared spectra and is susceptible to instrument and background noise.[112] Since the MSF approach relies on a well populated database, a different method was used due to the limited database.

Protein database searching is the other technique used to identify microorganisms from MALDI-MS data. In protein database searching, each peak found in the mass spectrum was used to search the Swiss–Prot and TrEMBL databases using the sequence retrieval system (SRS) found on the Expert Protein Analysis System (ExPASy) web server hosted by the SIB.[113] The SRS searches were done by inputting the peak *m/z* values as protein molecular weights. The search results containing hits from the analyzed organism were stored and counted as a tentative protein identification. The same peak *m/z* values were used to search the in-house *E. coli* database constructed using the FASE program. In this work, protein database searching using the SRS was done for three organisms: *E. coli*, *B. subtilis*, and *B. thuringiensis*.

2.7.3 Proteolysis of Impacted Bioaerosols

Since there is always a possibility of organisms from different species or subspecies containing peaks with identical masses, a different approach is needed to elucidate these differences. One such approach is the proteolytic digestion of whole-cell microorganisms. Proteolysis results in smaller peptide fragments from the enzymatic of chemical cleavage of the peptide bonds. As a result, smaller peptides are produced, which are more easily ionized. In

addition, there are two potentially useful bioinformatic approaches available for the identification of proteolytic peptide fragments. One approach is termed peptide mass mapping or fingerprinting (PMF). For PMF, the peptide fragment *m/z* values from the detected peaks can be used to search a database in which *in silico* proteolysis was done on all proteins contained within the database. These generated masses are then matched against those obtained from the mass spectra. This can be done using a search engine such as MASCOT.[114] The other approach is sequence or sequence tag analysis. In this approach, MS/MS experiments of the peptide mass fragments below an *m/z* of 6 kDa can be done. The sequence can be used to search the same *in silico* proteolytic database. However, the confidence of a match is greater due to additional peptide primary structure information.

Proteolysis experiments were done for peptide mass mapping to assist in the identification of microorganisms from collection bioaerosols. Since the impacted bioaerosol samples were collected directly onto a MALDI target, all digestions were accomplished directly on the flat surface of the MALDI target. To accomplish this, a mini-well technique is developed. The mini-well is a small diameter metal ferule or spacer with smooth edges for enhancing the contact with the MALDI target. Once the bioaerosol was collected, the mini-well was placed on top of a spot on the target. Once placed on top, the proteolysis was carried out by placing the reagents inside the mini-well. Two proteolytic agents were used, trypsin and cyanogen bromide (CNBr). Trypsin cleaves on the C-terminal side of lysine and arginine residues while CNBr cleaves on the C-terminal side of methionine residues as shown in Figure 27. Also, CNBr cleavage can result in the formation of homoserine lactone from the methionine residue on the C-terminal side of the N-terminal cleavage product as shown in Figure 28. The CNBr digestion was carried out at room temperature in a fume hood due to its toxic nature and byproducts. For

trypsin, the digestion was carried out in an incubator at 37° C. For both proteolysis, the reaction was allowed to proceed until all solvent was evaporated. Once the solvent evaporates, the mini-well was removed and matrix was applied to the target on the proteolytic digested spot.

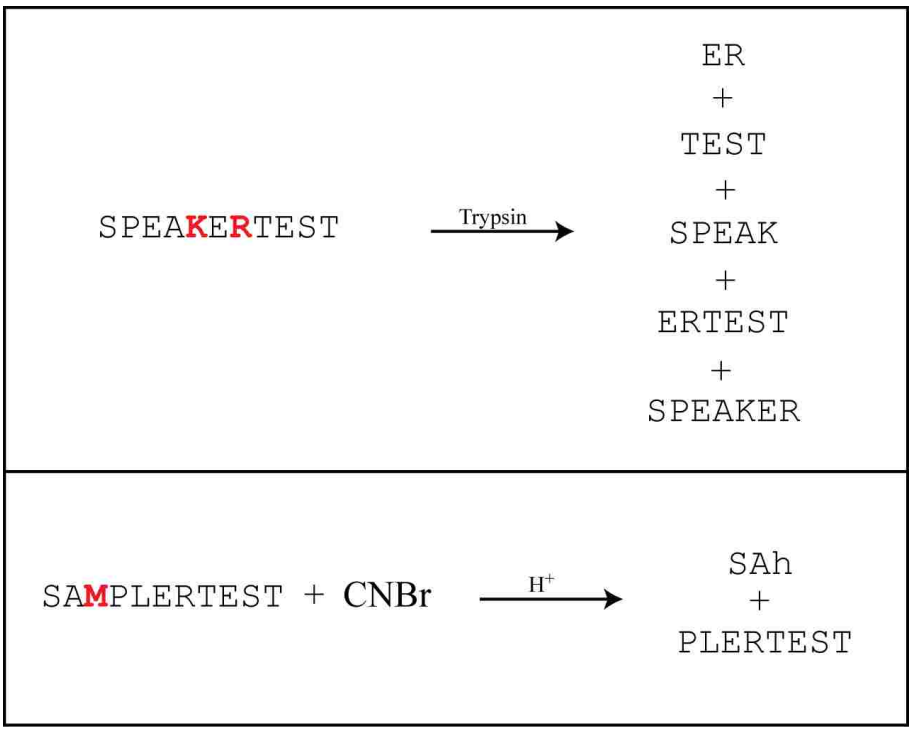


Figure 27. Proteolytic agent digestion products for two hypothetical peptides using (A) trypsin and (B) CNBr

For CNBr digestions, a 1 M solution was made in acetonitrile and stored in the freezer at -20°C until needed. The reaction mixture for the CNBr digestion was as follows: 2 µL of CNBr solution, 5 µL of acid solution (5% TFA, 50% Formic acid or 0.1 M HCl), and 5 µL of water. The acid solution was varied to maximize digestion performance. For trypsin digestion, a 100 µM solution was made in water fresh each time. The reaction solution was as follows: 2 µL of trypsin (100µM), 5 µL of NH_4HCO_3 (50mM) and 5 µL of water. The MALDI target was placed inside a petri dish and covered before placing it into the incubator. In addition to the mini-well proteolysis, *in vitro* digestions were done for relative digestion efficiency determinations.

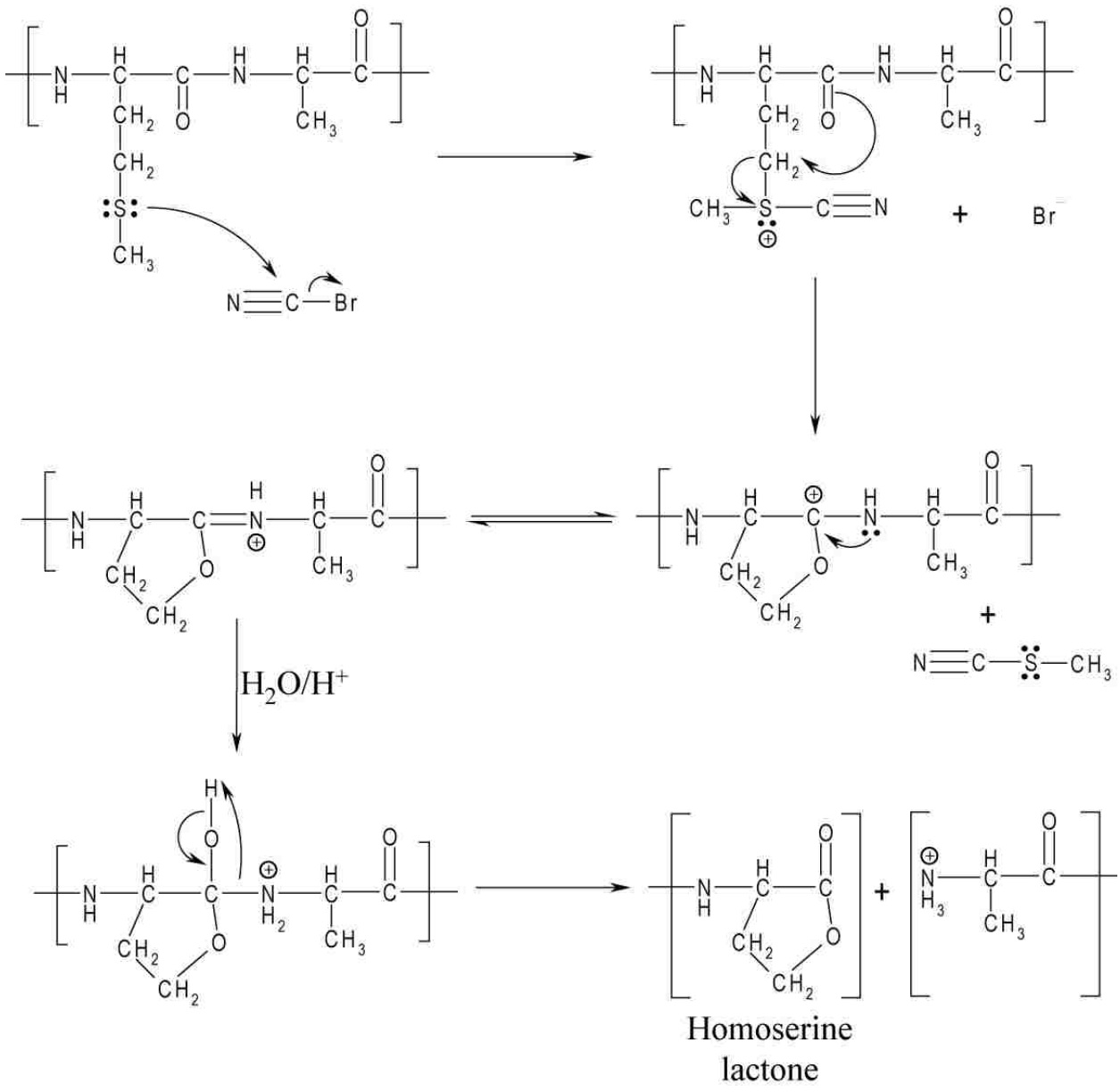


Figure 28. CNBr proteolytic reaction resulting in the conversion of methionine into homoserine lactone.

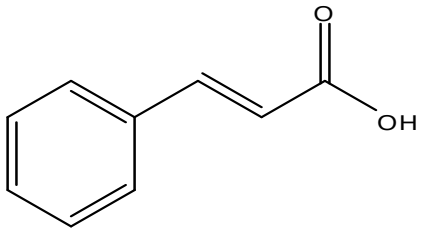
2.8 Reagents and Chemicals

Solvents used in this work include: acetonitrile (ACN), acetone, methanol, ethanol, and isopropanol from Fisher Scientific. Other chemicals used in this work were: acetic acid, trifluoroacetic acid (TFA), formic acid and hydrochloric acid from Fisher Scientific, cyanogen

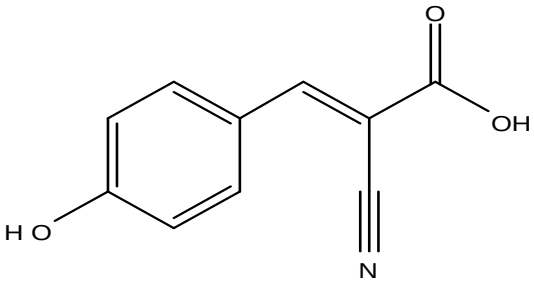
bromide (CNBr), ammonium bicarbonate, Iodine, and ammonium oxalate monohydrate were obtained from Sigma-Aldrich, and safranin T and crystal violet were obtained from Fluka.

The MALDI matricies used in this work were all derivatives of cinnamic acid[115]: α-cyano-4-hydroxycinnamic acid (CHCA, Sigma-Aldrich), 3,5-dimethoxy-4-hydroxycinnamic acid, (sinapic acid, Fluka), 4-hydroxy-3-methoxycinnamic acid (ferulic acid, Fluka), and 3,4-dihydroxycinnamic acid (caffeic acid, Fluka). Structures of these matricies are in Figure 29. Typical matrix solutions consisted of 20 to 40 mg per 1 mL of solvent. In most instances the solvent used was a 50/50 ACN/water containing 0.1% TFA solution. The ACN/water ratios can be adjusted to enhance MALDI crystal formation and signal quality.

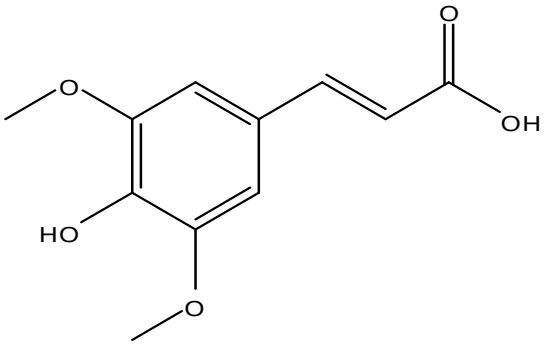
All the biological samples used for analysis were purchased from Sigma-Aldrich: *Escherichia coli* w strain (ATCC 9637), *Escherichia coli* b strain (ATCC 11303), *Bacillus subtilis* (ATCC 6633), DPCC treated bovine pancreas trypsin (E.C. 3.4.21.4), bovine heart cytochrome c, bovine pancreas insulin, horse heart myoglobin, ferrous stabilized human hemoglobin subunit Ax, chicken egg white lysozyme (E.C. 3.2.1.17), bovine serum albumin (BSA), and chicken egg white albumin (ovalbumin).



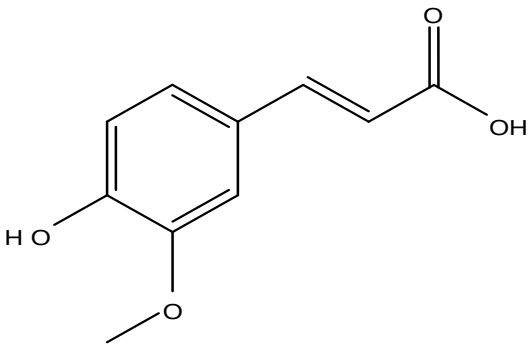
Cinnamic Acid



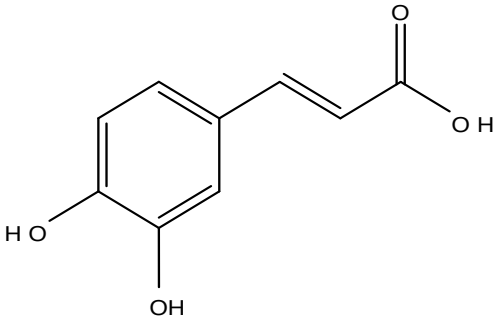
α-cyano-4-hydroxycinnamic acid
(CHCA)



3,5-dimethoxy-4-hydroxycinnamic acid
(Sinapic Acid)



4-hydroxy-3-methoxycinnamic acid
(Ferulic Acid)



3,4-dihydroxycinnamic acid
(Caffeic Acid)

Figure 29. Structures of MALDI matricies used in this work.

CHAPTER 3. ANALYSIS OF INTACT WHOLE-CELL BACTERIA USING A QUADRUPOLE-TIME-OF-FLIGHT MASS SPECTROMETER COMBINED WITH PROTEIN DATABASE SEARCHING

3.1 Introduction

In this Chapter, a study of the analysis of intact cell bacteria using a tandem quadrupole-time-of-flight mass spectrometer is described. This study was done in order to further develop this technique for use on collected bioaerosols. Three bacteria species, *Escherchia coli*, *Bacillus subtilis* and *Bacillus thuringiensis*, were used in this work. After analysis of the intact cells by oMALDI-QTOF mass spectrometry, the mass spectral peaks were searched against an online protein database to identify tentative protein matches. Mass spectral peaks that were unique to a given microorganism were identified as potential biomarkers for use in identification protocols. A benefit of using a tandem instrument for this application is the ability MS/MS experiments.

3.2 Experimental

Samples of proteins and intact cell bacteria were analyzed by the MALDI-TOF and oMALDI-QTOF. In this work, biological samples were prepared using both a dried droplet and an overlay method for MALDI preparation. Sample utilization for MALDI preparations is usually expressed in terms of the amount of material deposited on the target. For proteins, this amount was in the picomolar range and for bacteria was in the amount of micrograms. All stock samples were prepared in water and were vortexed and sonicated without heat prior to analysis. All protein samples were completely dissolved at 1mM concentrations except for insulin due to its low solubility at neutral pH. Suspensions of bacteria were prepared using the same protocol as for proteins and were thoroughly mixed prior to removing material for sample analysis.

The two instruments used in this work are both equipped with N_2 (λ= 337 nm) lasers. While the lasers on the two instruments are comparable, they operate at considerably different

repetition rates. The MALDI-TOF laser operates at 1 Hz while the oMALDI-QTOF laser can operate between 1 and 40 Hz. For this work, a repetition rate of 40 Hz was used unless otherwise indicated. MALDI sample preparations were done on stainless steel targets designed specifically for each instrument. Each plate was cleaned and polished prior to analysis. Typical sample volumes were less than 1 μL to prevent spot spreading and promote faster drying times. Several matricies were used in this study; however, best results were obtained with CHCA, SA, CA, and FA. Table 9 provides the matrix sample masses used with solvent composition and indicate whether the solution was saturated. MALDI samples were prepared with equal

Table 9. Preparation methods for MALDI matricies used.

Matrix	**Mass Used (g)**	**Solvent (1mL volumes)**	**Saturated**
CHCA	0.0250	50/50 ACN/H_2O + 0.1% TFA 70/30 ACN/H_2O + 0.1% TFA	Yes Yes
SA	0.0250	50/50 ACN/H_2O + 0.1% TFA	Yes
CA	0.0300	50/50 ACN/H_2O + 0.1% TFA	No
FA	0.0350	50/50 ACN/H_2O + 0.1% TFA	No

Volumes of matrix to analyte solution. For protein samples, no additional sample preparation was needed. For bacteria samples, application of a dilute TFA solution was used to enhance signal intensity and spectra quality. For most samples, a gentle cold air stream was used to assist in solvent evaporation and matrix crystallization. However, this is not a necessary step and only serves to reduce the amount of wait time between analyses.

After bacteria samples were analyzed, mass spectral peaks were searched using two approaches, the ExPASy sequence retrieval system (SRS)[113] and an in-house generated bacteria protein database generated from translated sequenced genomic information from NCBInr. In addition, each peak was recorded in a separate database for the initial generation of a bacteria mass spectral fingerprint protocol.

3.3 Results and Discussion

3.3.1 Comparison of an Axial MALDI-TOF to an Orthogonal MALDI-QTOF

Mass spectra from cytochrome *c* and *E. coli* from the two mass spectrometers used are shown in Figure 30 and 31. Figure 30a and 31a were obtained on the MALDI-TOF in positive reflectron mode using a 350 ns delay, 2 ns acquisition time and 1 Hz laser repetition rate. Ion suppression below *m/z* 1000 was used with the ion deflector operating at 2 kV. Figure 30b and 31b were obtained on the oMALDI-QTOF in positive mode with a mass range of 1000 – 16000 *m/z* selected and a 40 Hz laser repetition rate. Unlike the MALDI-TOF, the QTOF-TOF only records the data for the selected mass range. The mass spectra obtained from each instrument are similar. In fact, for MALDI, these spectra would be indistinguishable since MALDI mass spectra inconsistencies are usually the result of varying peak intensities and background noise. One of the largest differences between the two instruments is that of mass resolution and dynamic range. While the oMALDI-QTOF has better resolution, the dynamic range is limited due to the quadrupole mass analyzer.[104] However from these results, it is apparent that spectra from intact cells can be obtained using both instruments and are sufficiently similar to be used interchangeably if required.

3.3.2 MALDI Matrix Evaluation

Selecting a suitable MALDI matrix is critical for obtaining optimal results. In general, a given matrix is better at analyzing specific classes of molecules.[100] Since this work involves biomolecules, specifically proteins, peptide and protein matricies were chosen. Of all the matricies tried, the four with best results for intact cell bacteria were CHCA, CA, SA, and FA. Of the four, FA demonstrated the largest tolerance for sample impurities and concentration variations. Figure 32 shows mass spectra of *E. coli* obtained using each matrix using the

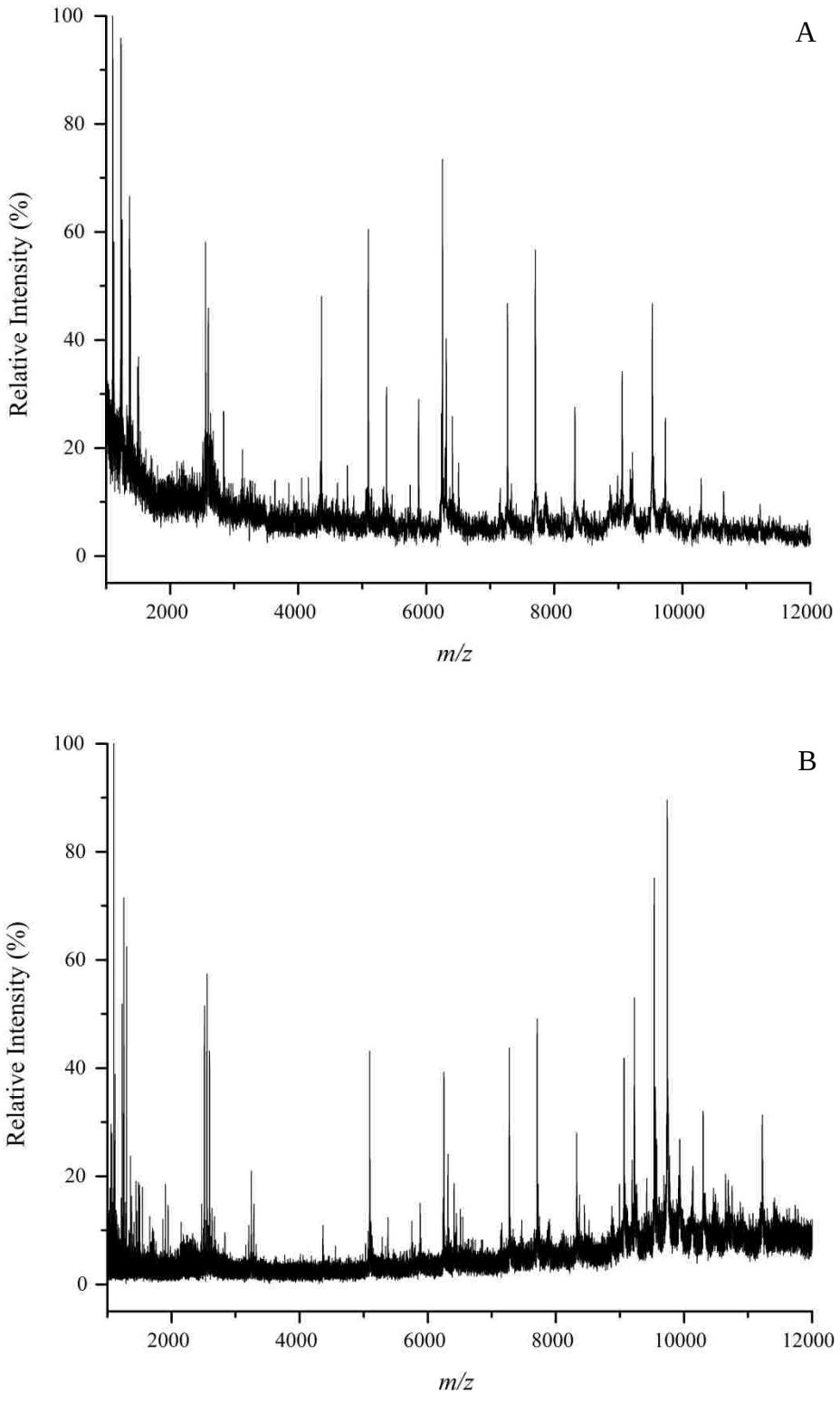


Figure 30. MALDI mass spectra of whole intact cell *E. coli* in ferulic acid matrix from (A) MALDI-TOF and (B) oMADLI-QTOF.

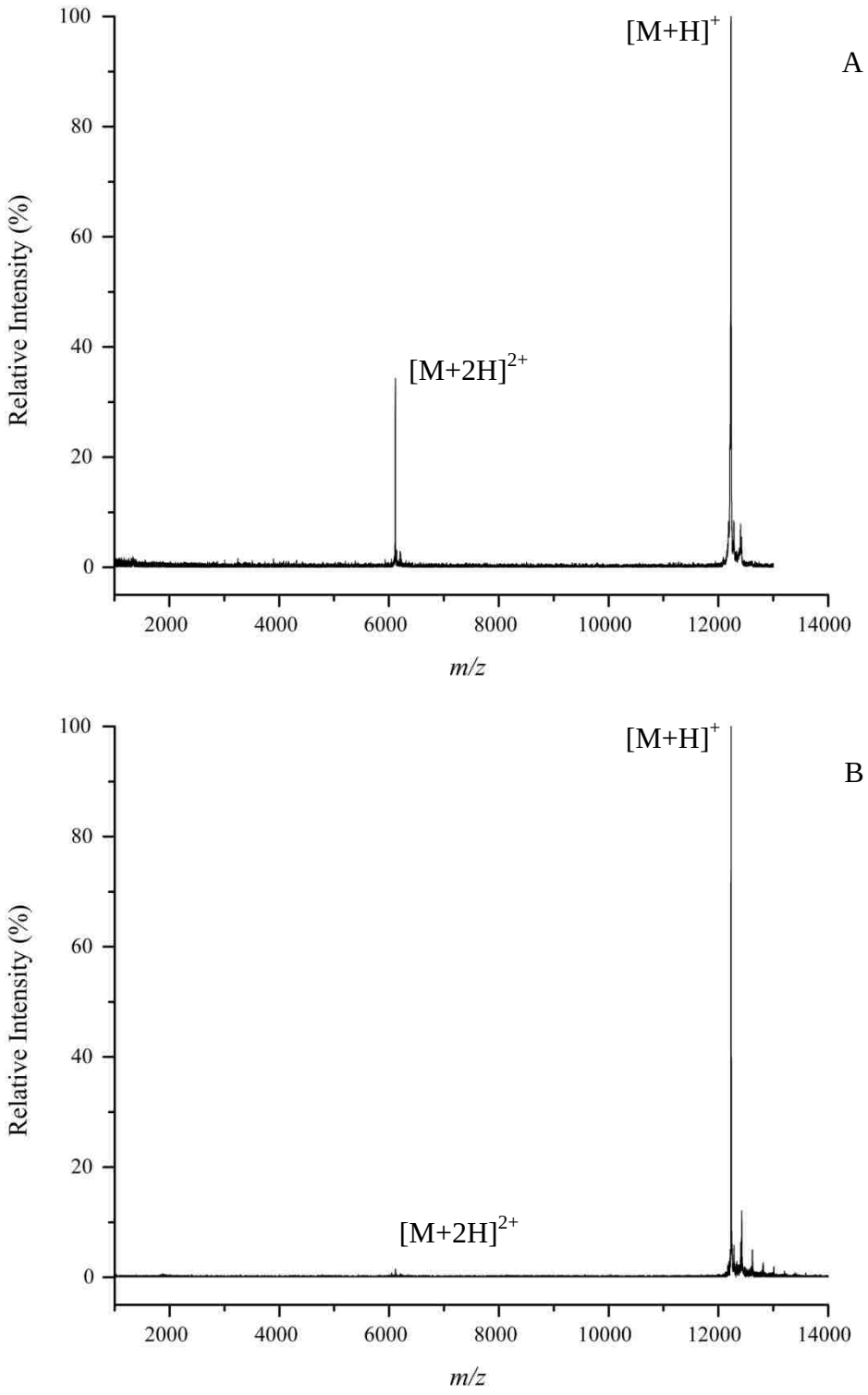


Figure 31. MALDI mass spectra of cytochrome *c* in ferulic acid matrix from (A) MALDI-TOF and (B) oMALDI-QTOF.

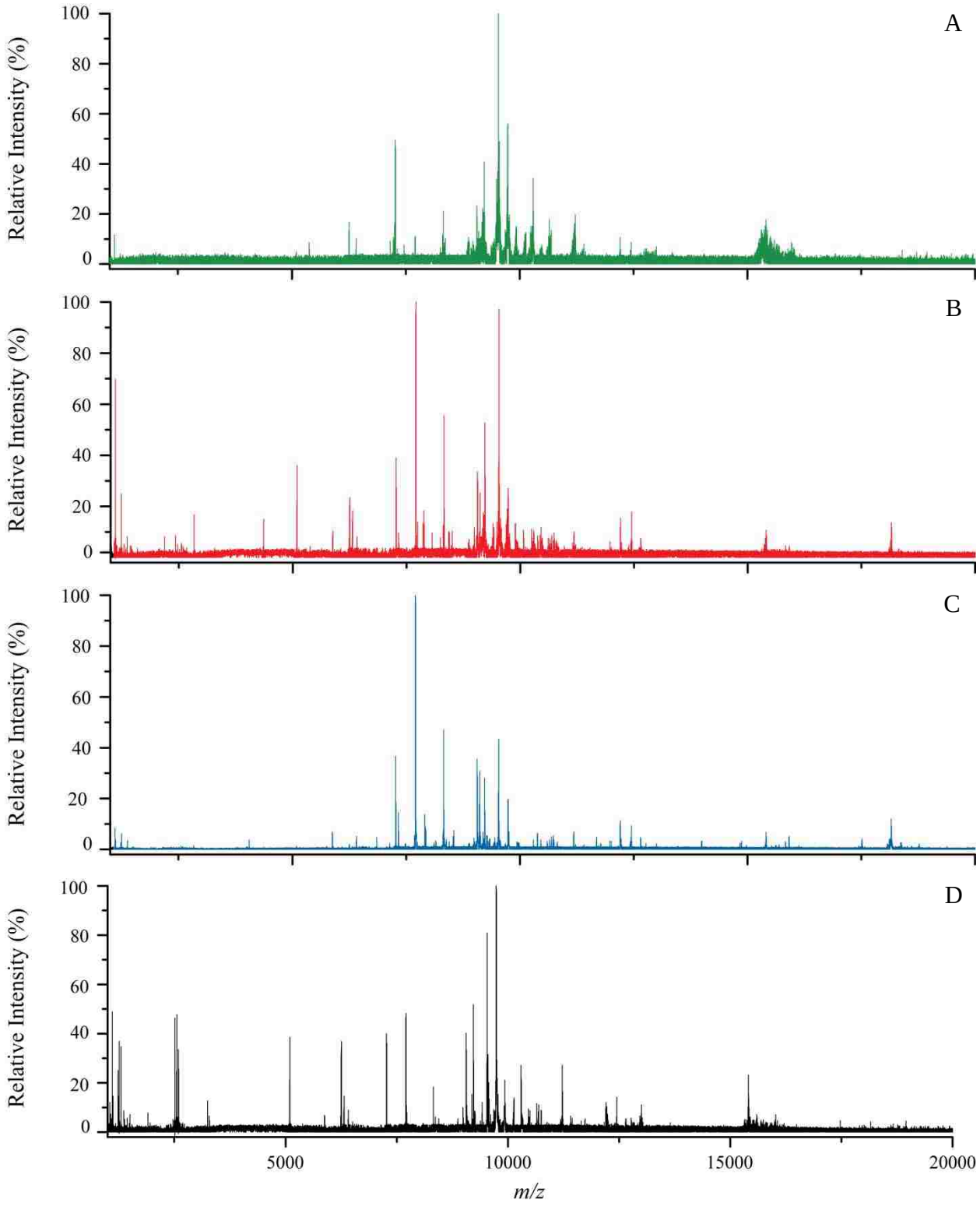


Figure 32. oMALDI-QTOF mass spectra of whole intact cell *E. coli* using (A) CHCA (B) CA (C) SA and (D) FA.

oMALDI-QTOF. As shown, each of these four matricies can be used for analysis of whole cells successfully. However, FA demonstrated the broadest mass range such that it was possible to produce low mass peaks (< 3000) and high mass peaks (> 8000). Despite this, most of the initial work involved CHCA due to its ease of ionization and low laser energy requirements.

3.3.3 Database Searching

Database searching was done using ExPASy's sequence retrieval system (SRS) against the SWISS-PROT and TrEMBL databases. For ease of data handling, three initial searches were done using the organism as the query parameter, which produced a table of proteins for the selected organism only. The resulting reports were copied into a spreadsheet for quick offline analysis. In addition, an in-house protein database was generated from the genomes of five *E. coli* strains. The in-house database was populated with protein sequences in FASTA format from translated NCBInr protein databases. A small visual basic application was written to calculate molecular weights from these FASTA sequences. The in-house database and the SRS results were compared for the number of matches and protein identifications.

Figures 33 and 34 are from a single *E. coli* spectrum acquired using the oMALDI-QTOF mass spectrometer. The spectrum was generated using a FA matrix and was collected from 25 TOF cycles at 45% laser energy set at a repetition rate of 40 Hz. Figure 33a shows the entire mass spectrum while Figure 33b, 34a and 34b show expanded regions of the smoothed spectrum with peak labels that are recorded in Tables 10 and 11. For this analysis, peak *m/z* values were searched using the ExPASy's sequence retrieval system (SRS) against the SWISS-PROT and TrEMBL databases. The searches were carried out using the masses as determined from the *m/z* values with a +1 charge.

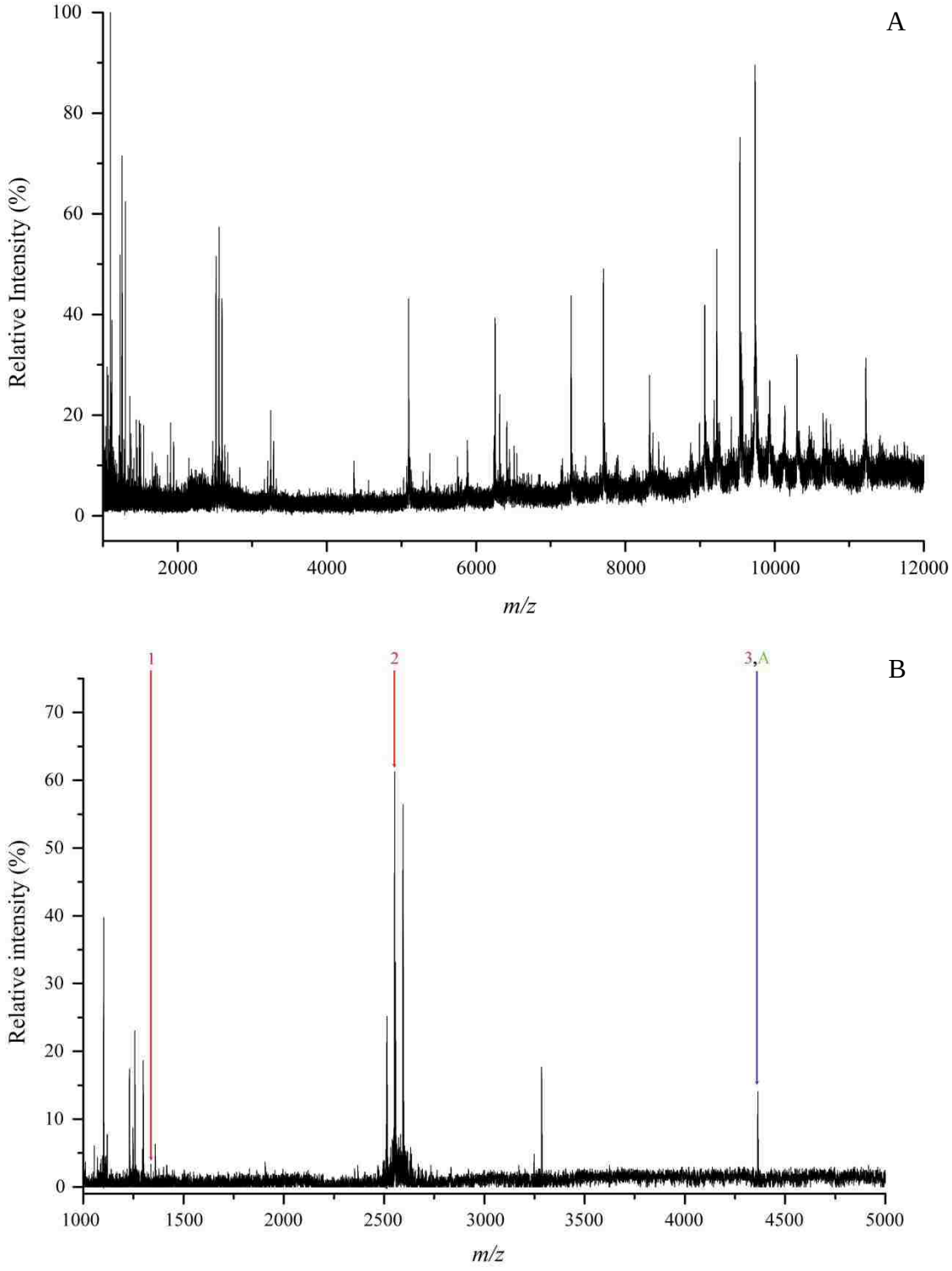


Figure 33. oMALDI-QTOF mass spectra of intact whole cell *E. coli* using FA (A) full spectrum and (B) expanded mass region (1000-5000 *m/z*) with labels of identified proteins.

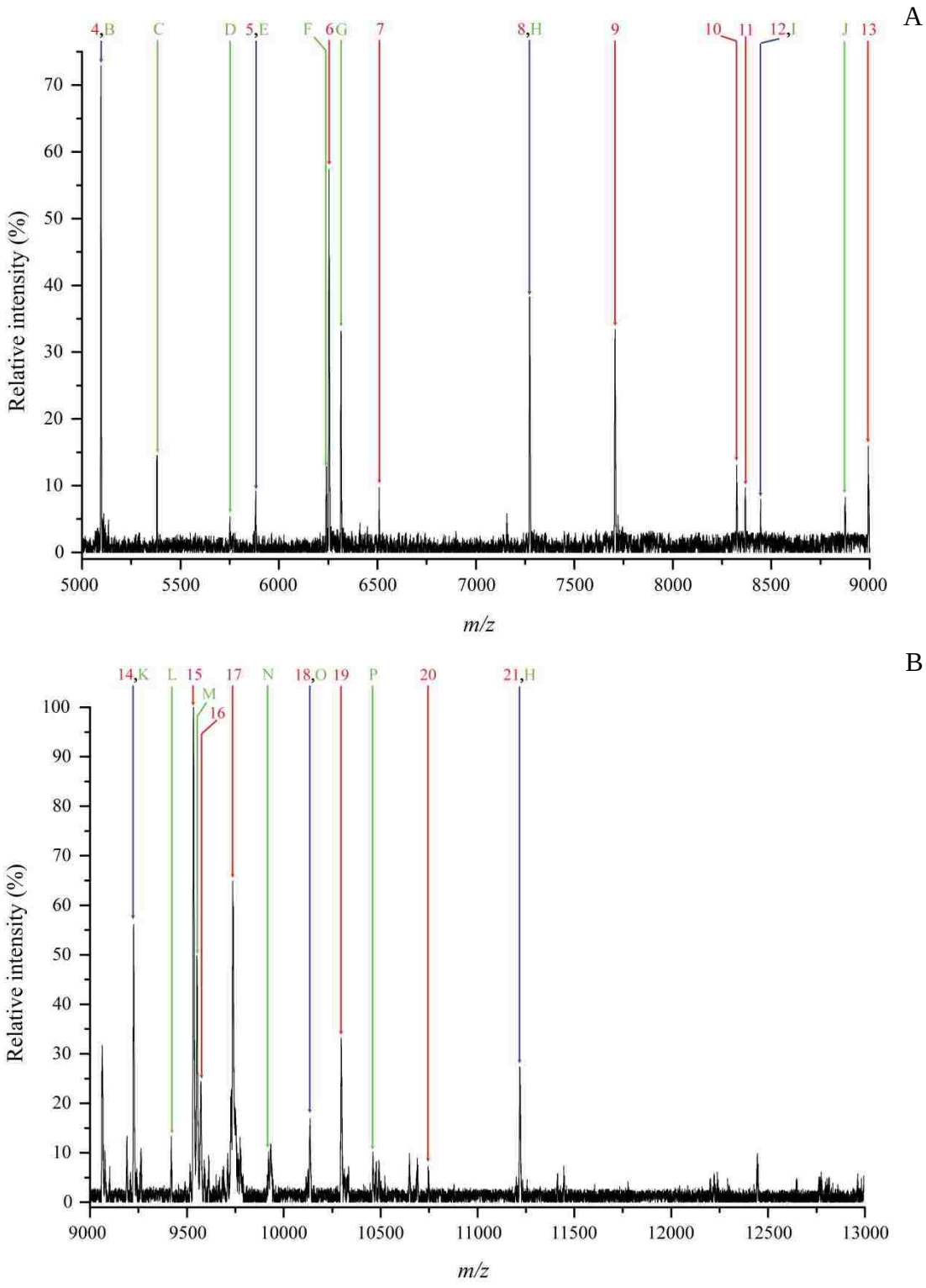


Figure 34. oMALDI-QTOF mass spectra of intact whole cell *E. coli* using FA with expanded mass regions (A) 5000-9000 *m/z* and (B) 9000 – 13000 *m/z* and labeled proteins.

Table 10. SRS search results of identified proteins from *E. coli*. (* = hits unique to *E. coli*)

Peak	Peak *m/z*	Description
1	1337*	*E. coli* [*] Dihydropteroate synthase 1 (Fragment)
2	2552*	*E. coli* [*] PhoA protein (Fragment)
3	4364*	*E. coli* [O157:H7][O6:K15:H31 (strain 536/UPEC)][(strain UTI89/UPEC)][Strain K12] 50S ribosomal protein L36 *E. coli*[*] UreG protein (Fragment)
4	5096*	*E. coli* [O157:H7][O6:K15:H31(strain 536/UPEC)][O6][Strain K12][strain UT189/UPEC] Stationary-phase-induced ribosome-associated protein (SRA) (30S|ribosomal protein S22) *E. coli* [strain B] 30S ribosomal subunit S22
5	5881*	*E. coli* [*][O6:K15:H31 (strain 536/UPEC)][(strain UTI89/UPEC)] Putative uncharacterized protein
6	6254*	*E. coli* [O6] Putative uncharacterized protein
7	6508*	*E. coli* [O6:K15:H31 (strain 536/UPEC)] Putative uncharacterized protein *E. coli* [strain UTI89 / UPEC][O6] Putative conserved protein *E. coli* O157:H7 Putative lipoprotein Rz1
8	7272*	*E. coli* [O157:H7][O6][strain K12][strain B][strain UTI89/UPEC] [O6:K15:H31 (strain 536/UPEC)] 50S ribosomal protein L29
9	7706*	*E. coli* [*] Class I integrase (Fragment)
10	8323*	*E. coli* [O157:H7][strain K12][O6] Major outer membrane lipoprotein precursor (Murein-lipoprotein) *E. coli* [O6:K15:H31 (strain 536 / UPEC)] Major outer membrane lipoprotein *E. coli* [(strain UTI89 / UPEC)] Murein lipoprotein
11	8368*	*E. coli* [strain B] IS3 element protein InsF
12	8446*	*E. coli* [*] Dihydrofolate reductase
13	8993*	*E. col i*[*] Putative uncharacterized protein.
14	9225*	*E. coli* [O157:H7][O6][strain K12][O1:K1 (APEC)][strain UTI89/UPEC][O6:K15:H31 (strain 536/UPEC)] DNA-binding protein HU-beta (NS1) (HU-1). *E. coli* [strain B] Histone family protein DNA-binding protein. *E. coli* [O157:H7] Putative uncharacterized protein ydaQ (Putative uncharacterized|protein ECs1930)
15	9535*	*E. coli* [O6:K15:H31 (strain 536 / UPEC)] Putative uncharacterized protein.
16	9572*	*E. coli* [O6][strain K12] Acetolactate synthase isozyme 2 small subunit (EC 2.2.1.6) (AHAS-II)|(Acetohydroxy-acid synthase II small subunit) (ALS-II)
17	9739	*E. coli* [*] Dr hemagglutinin AFA-III operon regulatory protein afaF. *E. coli* [O1:K1 / APEC] Putative uncharacterized protein
18	10137*	*E. coli* [strain K12] Uncharacterized protein ydfK.
19	10299*	*E. coli* [*] Putative uncharacterized protein (Fragment).
20	10750	*E. coli* [O157:H7][O6][strain 12] [O6:K15:H31 (strain 536/UPEC)][strain UTI89/UPEC][O1:K1/APEC] Probable sigma(54) modulation protein (ORF3) (ORF95) *E. coli* [strain B] Sigma 54 modulation protein/ribosomal protein S30EA *E. coli* [*] Hibernation promoting factor
21	11217*	*E. coli* [strain B] Cyanide hydratase.

Table 11. In-house search results of identified proteins of *E. coli*.

Peak	Peak *m/z*	Description
A	4364	*E. coli* [UT189] Hypothetical protein
B	5096	*E. coli* [0157_H7 Sakai] 50S ribosomal protein L36
C	5380	*E. coli* [0157_H7 EDL933] unknown protein associated with Rhs element *E. coli* [UT189] kil protein of bacteriophage HK97
D	5750	*E. coli* [536] Hypothetical protein
E	5881	*E. coli* [0157_H7 EDL933] unknown protein encoded within prophage CP-933O
F	6239	*E. coli* [UT189] Hypothetical protein *E. coli* NinF
G	6447	*E. coli* [0157_H7 EDL933] 50S ribosomal protein L30
H	8323	*E. coli* [UT189] Hypothetical protein *E. coli*[536] hypothetical protein YdhZ
I	8446	*E. coli* [0157_H7 EDL933] unknown protein encoded within prophage CP-933R
J	8873	*E. coli* [UT189] [536] Hypothetical protein
K	9225	*E. coli* [0157_H7 EDL933] Hypothetical protein
L	9419	*E. coli* [UT189][536 hypothetical protein YcgZ
M	9552	*E. coli* [UT189] Hypothetical protein
N	9923	*E. coli* [0157_H7 Sakai] Hypothetical protein ECs5537
O	10137	*E. coli* [536] hypothetical protein ECP_4624
P	10461	*E. coli* [0157_H7 Sakai] putative transcription antitermination protein
Q	11217	*E. coli* [0157_H7 Sakai] hypothetical protein ECs4308

For the SRS searches, 21 of the 47 peaks searched produced a hit for the expected microorganism, *E. coli,* in the database. Within the 21 hits, 14 of the matches produced a hit with a corresponding protein identification while the remaining seven were labeled as putative orhypothetical. For the in-house database search, 17 of the 47 peaks produced hits. However, only 5 of these peaks provided a corresponding protein identification while the remaining 12 were labeled as hypothetical or unknown protein. From these results, the SRS had an overall hit percentage of 44.7% and the in-house database search of 36.2%. While these percentages are not significantly different, the major difference between these two techniques is in the number of corresponding protein identifications. For the SRS, the protein identification was approximately 30 % while the in-house database was only 17 %. This indicates the SRS technique is almost twice as successful as the in-house database searching. This in part could be due to the lack of data common between the in-house database and the SRS databases, SWISS-PROT and TrEMBL or from false positives. When analyzing both search techniques for similarities, only 8 of the peaks were found in both lists. Of these eight cross referenced matches, only the peak with an *m/z* 5096 gave results of meaningful similarity. This peak is typically labeled as ribosomal in nature and is consistent with the results from both searches despite not being identified as the identical protein from both searches.

In database searching, multiple entries for a given protein mass are typically found. Table 12 is an SRS search for the mass 5096. As shown, there are 64 entries of which 6 of them are *E. coli*. A brief look at the entire table yields a broad spectrum of organisms including bacteria, mouse, human, and virus proteins. In this work, there was on average 80.3 hits per mass value searched using the SRS. Since this search technique accesses database information from multiple organisms, results are typically plentiful with matches for many organisms. Of the 80.3 average

Table 12. SRS search results of molecular weight 5096.

Description	Organism	MW
Heme exporter protein D (Cytochrome c-type biogenesis protein ccmD)	*Paracoccus denitrificans*	5096
50S ribosomal protein L36	*Sphingopyxis alaskensis*	5096
Stationary-phase-induced ribosome-associated protein (SRA) (30S|ribosomal protein S22)	*Escherichia coli* [O157:H7][O6:K15:H31]	5096
Stationary-phase-induced ribosome-associated protein (SRA) (Protein D)|(30S ribosomal protein S22)	*Escherichia coli* (strain K12)	5096
Stationary-phase-induced ribosome-associated protein (SRA) (30S|ribosomal protein S22)	*Escherichia coli* (strain UTI89)	5096
Stationary-phase-induced ribosome-associated protein (SRA) (30S|ribosomal protein S22)	*Shigella boydii* serotype 4 (strain Sb227)	5096
Stationary-phase-induced ribosome-associated protein (SRA) (30S|ribosomal protein S22)	*Shigella sonnei* (strain Ss046)	5096
Uncharacterized protein YHL048C-A	*Saccharomyces cerevisiae*	5096
Putative uncharacterized protein	*Haloarcula marismortui*	5096
Predicted protein	*Botryotinia fuckeliana* B0510	5096
Calmodulin (Fragment)	*Sporothrix schenckii*	5096
Calmodulin (Fragment)	*Scedosporium prolificans*	5096
Calmodulin (Fragment)	*Scedosporium aurantiacum*	5096
Calmodulin (Fragment)	*Pseudallescheria africana*	5096
Calmodulin (Fragment)	*Pseudallescheria minutispora*	5096
Calmodulin (Fragment)	*Pseudallescheria angusta*	5096
Calmodulin (Fragment)	*Pseudallescheria fusoidea*	5096
Calmodulin (Fragment)	*Pseudallescheria ellipsoidea*	5096
Calmodulin (Fragment)	*Pseudallescheria boydii*	5096
Calmodulin (Fragment)	*Duddingtonia flagrans*	5096
Calmodulin (Fragment)	*Arthrobotrys oligospora*	5096
3-Hydroxybutyryl-CoA dehydrogenase (Frag)	*Phanerochaete chrysosporium*	5096
Adenosine A2b receptor (Frag)	*Homo sapiens*	5096
Putative uncharacterized protein	*Trichomonas vaginalis* G3	5096
Putative uncharacterized protein (Frag)	*Plasmodium berghei*	5096
Spinocerebellar ataxia type 1 protein (Frag)	*Macrotus californicus*	5096
RstA2 (Frag)	*Vibrio* phage CTX	5096
Chromosome chr5 scaffold_64, whole genome shotgun sequence	*Vitis vinifera*	5096
Putative uncharacterized protein	*Oryza sativa* subsp Japonica	5096
Putative uncharacterized protein	*Vibrio cholerae* 2740-80	5096
Putative uncharacterized protein	*Methylibium petroleiphilum* (strain PM1)	5096
Putative inner membrane protein translocase component YidC	*Polaribacter dokdonensis* MED152	5096
30S ribosomal subunit S22	*Escherichia coli* B	5096
Putative uncharacterized protein	*Sulfitobacter* sp NAS-141	5096
Putative uncharacterized protein	*Janibacter* sp HTCC2649	5096
Putative uncharacterized protein	*Marinomonas* sp MED121	5096
50S ribosomal protein L34	*Marinomonas* sp MED121	5096
50S ribosomal protein L36	*Roseobacter* sp CCS2	5096
Degenerate transposase (Orf1)	*Streptococcus pneumoniae* SP23-BS72	5096
50S ribosomal protein L36	*Sphingomonas wittichii* RW1	5096
Putative uncharacterized protein	*Eubacterium ventriosum* ATCC 27560	5096
Putative uncharacterized protein	*Bifidobacterium adolescentis* L2-32	5096
Putative uncharacterized protein	*Beggiatoa* sp PS	5096
Putative uncharacterized protein	*Campylobacter hominis*	5096
Putative uncharacterized protein	*Bacteroides ovatus* ATCC 8483	5096
Putative uncharacterized protein	*Synechococcus* sp BL107	5096
DNA topoisomerase IV subunit beta (Frag)	uncultured bacterium	5096
Putative uncharacterized protein	*Candidatus Kuenenia* stuttgartiensis	5096
Nicotinate phosphoribosyltransferase (EC 24211)	*Vibrio angustum* S14	5096
Putative uncharacterized protein	uncultured bacterium 582	5096
Reverse transcriptase	*Ralstonia eutropha*	5096
Putative 4-hydroxy-2-oxovalerate aldolase (Frag)	*Sphingomonas* sp P2	5096
KaiB (Frag)	*Nostoc linckia*	5096
Putative uncharacterized protein	*Rickettsia felis*	5096
Putative uncharacterized protein	*Streptomyces avermitilis*	5096
Putative uncharacterized protein	*Pseudomonas syringae* pv tomato	5096
Sepiapterin reductase	*Mus musculus*	5096
NS5a protein (Frag)	Hepatitis C virus subtype 1b	5096
Chromosome undetermined SCAF18766, whole genome shotgun sequence|(Frag)	*Tetraodon nigroviridis*	5096
Chromo-helicase-DNA binding protein (Frag)	*Nisaetus nipalensis*	5096
Collagen alpha1(III) (Frag)	*Xenopus laevis*	5096
Gag protein (Frag)	Human immunodeficiency virus 1	5096

hits per mass searched using the SRS, only 2.3 of those hits were from *E. coli*. Therefore, on average, there were over 75 non *E. coli* hits per mass searched. While this presents a unique bioinformatics problem, the lists are easily sorted using a spreadsheet program for manual data analysis. While both search techniques can be employed successfully, the SRS technique is better suited because this database is continually updated with the latest protein information.

To further test this technique, SRS searches were performed using the organism as the query parameter and the resultant information stored in a spreadsheet. This was done for *E. coli*, *B. subtilis* and *B. thuringiensis*. Each microorganism's mass spectra produced peaks which were used to search the data from the SRS. Figure 35 contains the spectra from each of the organisms in this work. From these spectra, differences are easily noted. Tables 13 and 14 contain the search results for the *Bacillus* microorganisms. For *B. subtilis*, 19 searchable peaks were found. Of the 19 peaks, only 6 produced hits from the SRS dataset as shown in Figure 36. For *B. thuringiensis*, 24 searchable peaks were found. Of the 24 searchable peaks, 14 produced hits from the SRS dataset as shown in Figures 37a and 37b. Upon examining the dataset for each microorganism, a larger dataset appears to provide for greater matches as should be the case from a strictly statistical perspective. Table 15 outlines the SWISS-PROT and TrEMBL results for each microorganism. While *B. thuringiensis* contains the greatest number of strains and/or subspecies, *E. coli* has the greatest number of unique proteins in the database. Alternatively, while *B. subtilis* only has a single species in the database, it contains the greatest percentage of non-putative protein descriptors and unique molecular masses at over 90 and 95 percent, respectively.

Biomarker determination is essential for successful microorganism identification. While the microorganism dataset in this work is limited to three, similar approaches could be used for

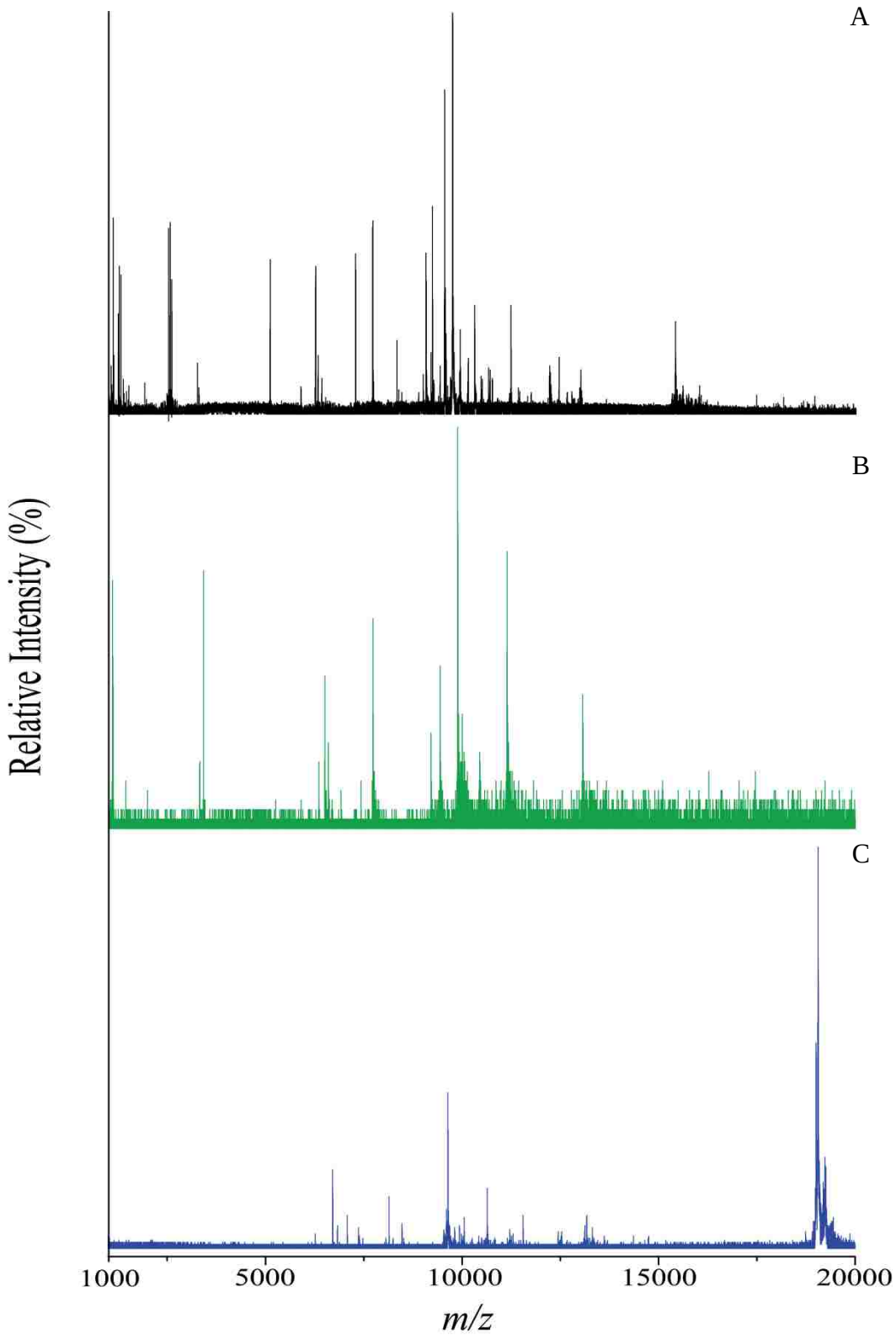


Figure 35. oMALDI-QTOF mass spectra (smoothed) of whole intact cells of (A) *E. coli* (B) *B. subtilis* and (C) *B. thuringiensis* in FA.

Table 13. SRS search results of identified proteins from *B. subtilis*. (* = hits unique to *B. subtilis*)

Peak	*m/z*	Description
1	3319*	*Bacillus subtilis* Methylglyoxal synthase (Fragment)
2	7728	*Bacillus subtilis* Uncharacterized protein ypmT
3	9209*	*Bacillus subtilis* Uncharacterized protein yscA
4	9884	*Bacillus subtilis* DNA-binding protein HU 1 (DNA-binding protein II) (HB) Uncharacterized protein ypbS
5	11142*	*Bacillus subtilis* 50S ribosomal protein L24 (BL23) (12 kDa DNA-binding protein) (HPB12).
6	13066	*Bacillus subtilis* Putative uncharacterized protein yhaH (YhaH protein).

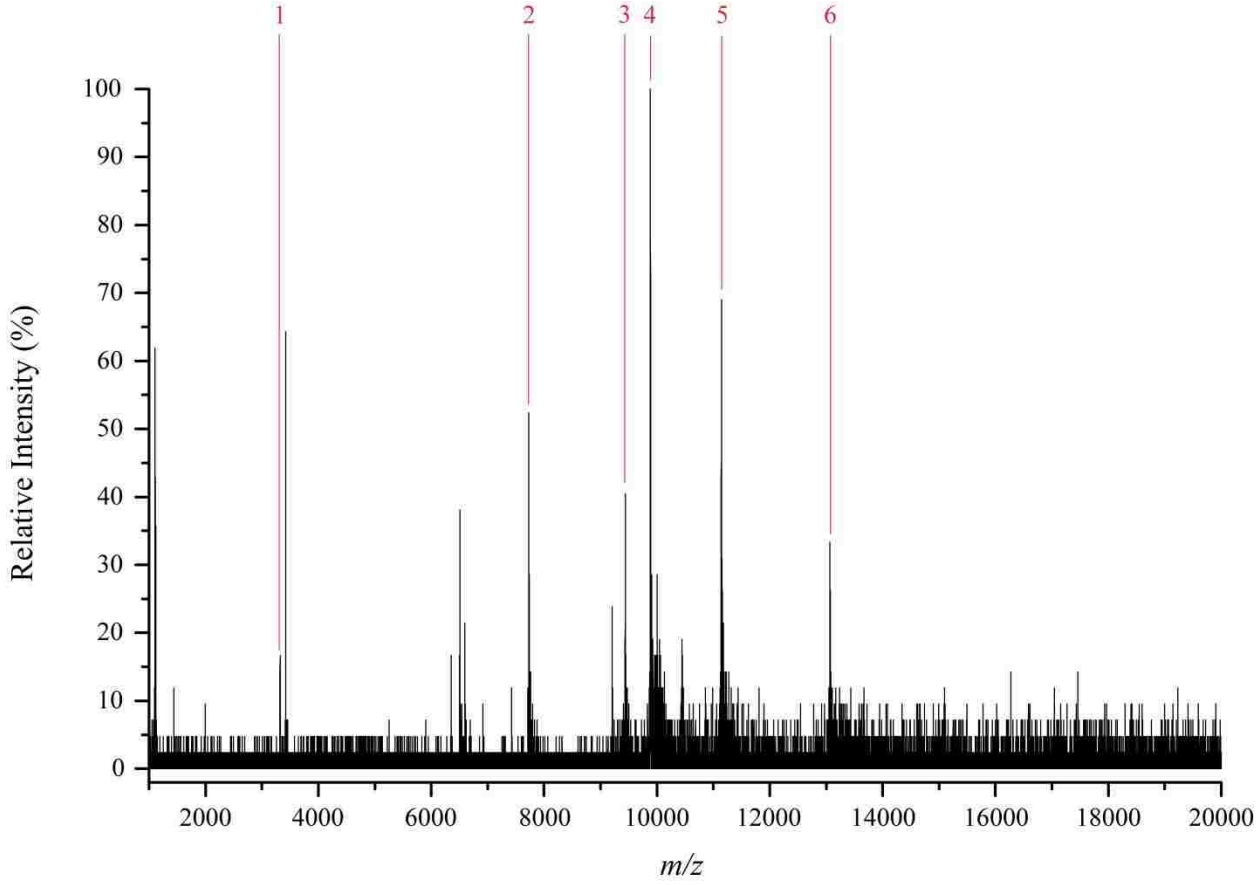


Figure 36. oMALDI-QTOF mass spectra of intact whole cell *B. subtilis* using FA with protein label for SRS identified peaks.

Table 14. SRS search results of identified proteins from B *thuringiensis*. (* = hits unique to *B. thuringiensis*)

Peak	*m/z*	Description
1	6708*	*Bacillus thuringiensis* [serovar israelensis] ClpB protein
2	7081*	*Bacillus thuringiensis* Putative uncharacterized protein *Bacillus thuringiensis* [serovar israelensis] Hypothetical membrane spanning protein
3	7367	*Bacillus thuringiensis* [strain Al Hakam][subsp. Konkukian] Cold shock protein *Bacillus thuringiensis* [serovar israelensis] [subsp. Konkukian] Putative uncharacterized protein
4	8142	*Bacillus thuringiensis* [serovar israelensis][(strain Al Hakam][subsp. Konkukian] Putative uncharacterized protein
5	8473*	*Bacillus thuringiensis* [strain Al Hakam][subsp. Konkukian] Putative uncharacterized protein
6	9605*	*Bacillus thuringiensis* [serovar israelensis] DNA-binding protein HU *Bacillus thuringiensis* Hypothetical secreted protein
7	9642*	*Bacillus thuringiensis* [strain Al Hakam][subsp. Konkukian] Putative uncharacterized protein *Bacillus thuringiensis* [subsp. Konkukian] DNA-binding protein HU
8	10057*	*Bacillus thuringiensis* [strain Al Hakam][subsp. Konkukian] Putative uncharacterized protein
9	10643	*Bacillus thuringiensis* Bt SpoIIID
10	11552*	*Bacillus thuringiensis* [serovar israelensis] Microbial collagenase (EC 3.4.24.3) *Bacillus thuringiensis* [serovar konkukian str. 97-27] Putative uncharacterized protein
11	12446*	*Bacillus thuringiensis* [strain Al Hakam] Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96). *Bacillus thuringiensis* [subsp. Konkukian] 4a-hydroxytetrahydrobiopterin dehydratase (Pterin-4-alpha-|carbinolamine dehydratase) (EC 4.2.1.96)
12	13171	*Bacillus thuringiensis* [serovar israelensis][subsp. Konkukian] Putative uncharacterized protein *Bacillus thuringiensis* [subsp. Sotto] Putative glycerol uptake facilitator protein (Fragment).
13	19011	*Bacillus thuringiensis* [subsp. Konkukian] Acyl-CoA hydrolase (Cytosolic long-chain acyl-CoA thioester hydrolase)|(EC 3.1.2.-)
14	19052	*Bacillus thuringiensis* [subsp. Konkukian] Superoxide dismutase [Cu-Zn] (EC 1.15.1.1)

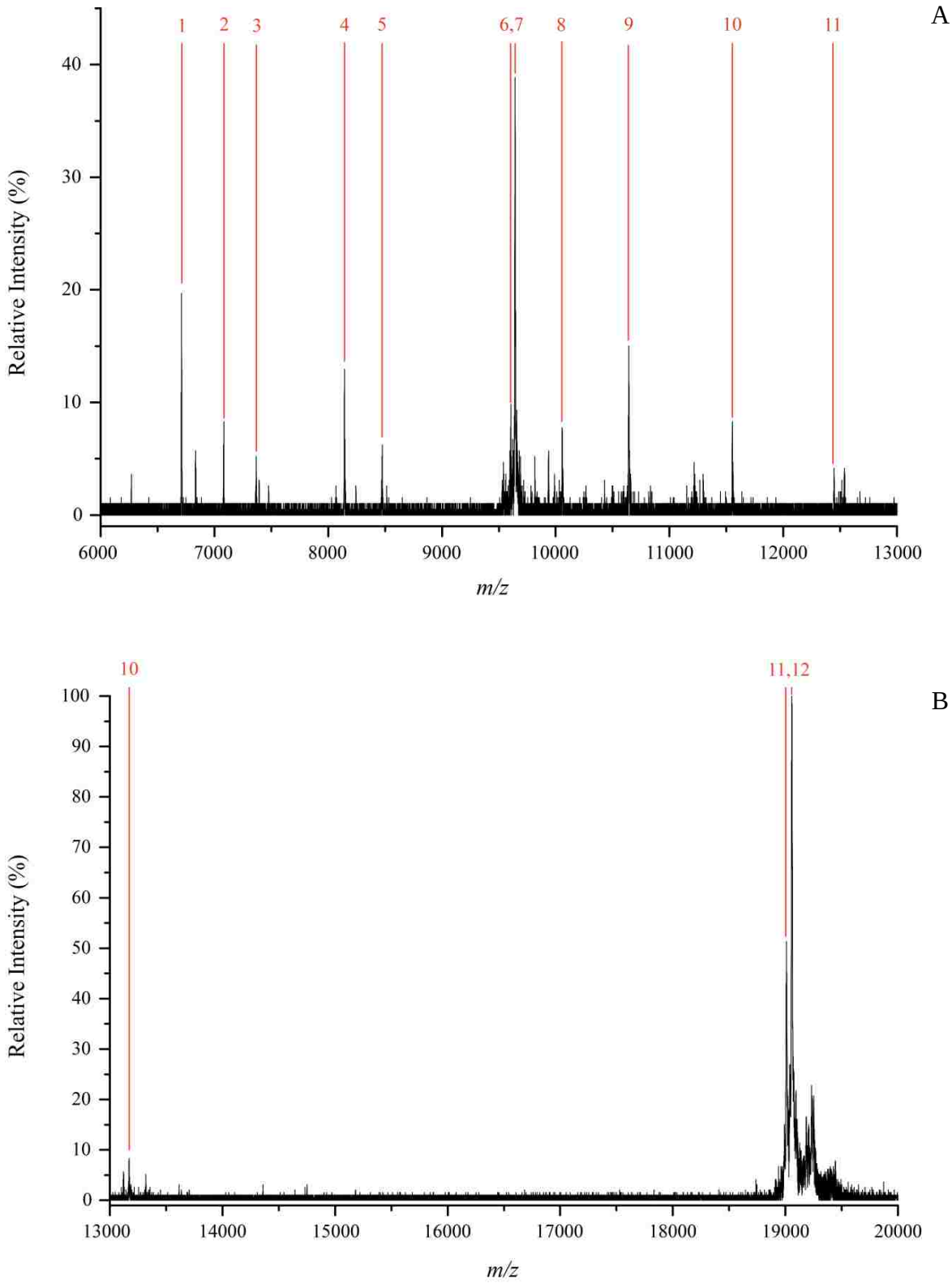


Figure 37. oMALDI-QTOF mass spectra of intact whole cell *B. thuringiensis* using FA with expanded mass regions (A) 6000-13000 *m/z* and (B) 13000 – 20000 *m/z* and labeled proteins

Table 15. SWISS-PROT and TrEMBL database statistics for microorganisms under analysis.

Microorganism	Strains or subspecies	Total Entries (Protein)	Unique Protein Entries (MW)	Unique Protein Entries (Name)
Escherichia coli	61	45,325	25,217	21,175
Bacillus subtilis	1	5,214	4,937	4,791
Bacillus thuringiensis	99	17,886	14,419	5,643

larger datasets and easily applied to any unknown microorganism with a nearly complete genome. By analyzing the peak lists from Tables 10, 13 and 14, unique peaks can be determined and labeled as biomarkers for the given microorganism. For *E. coli*, 19 of the 21 hits from the peak list were unique. For *B. subtilis*, 3 of the 6 hits from the peak list were unique, and for *B. thuringiensis*, 8 of the 14 hits were unique. The identified biomarkers are labeled with asterisks in their corresponding peak list tables.

3.4 Summary

In the work described in this Chapter, three microorganisms were subjected to intact whole cell bacteria MALDI-MS. While work of this nature is traditionally done on an MALDI-TOF mass spectrometer, it is possible to use a tandem mass spectrometer for the analysis of intact whole cells. This work demonstrates the successful analysis of three intact cell microorganisms using an oMALDI-QTOF mass spectrometer. The oMALDI-QTOF mass spectrometer was capable of analyzing intact cell bacteria using a variety of MALDI matricies with good reproducibility. By using an oMALDI-QTOF mass spectrometer, MS/MS experiments are possible. However, after analyzing current data, most detected peaks were greater than 6000 and thus are unable to be analyzed by MS/MS. Despite this, the oMALDI-QTOF mass spectrometer is suitable for intact cell MALDI-MS analysis. In addition, biomarker peaks were identified by database searching of protein databases both online and in-house. While the online databases provided more useful and meaningful data, in-house databases could be enhanced by

the addition of more information from more organisms and theoretically be as successful as the online database searches.

CHAPTER 4. FORENSIC EVALUATION OF COLLECTED BIOAEROSOLS FROM IMPACTION AND FILTRATION SAMPLES USING MATRIX-ASSISTED LASER DESORPTION/IONIZATION MASS SPECTROMETRY

4.1 Introduction

In the work described in this chapter, three bioaerosol samplers, an Andersen N6 single-stage impactor (AI), a cyclone impactor (CI) and a vacuum filter system (FS), were used to collect test bioaerosols for offline MALDI-MS analysis. All three samplers were evaluated in parallel through the simultaneous collection from a single bioaerosol source, a collision nebulizer. This was made possible through the use of a gas permeable sealed chamber called the bioaerosol exposure chamber (BEC). As the bioaerosols were sprayed inside the BEC, the three sampling systems and an aerosol particle sizer were attached and used to collect the bioaerosols for analysis. Each sampler system was evaluated for compatibility with MALDI-MS analysis. As such, all post collection sample handling, such as sample extracting and concentrating, was minimized to decrease the amount of time required for analysis. By analyzing these different sampling techniques and their compatibility with MALDI, their potential use for the detection and identification of biological material from bioaerosols was assessed.

4.2 Experimental

Three samplers, an Andersen N6 single-stage impactor, a cyclone impactor and a vacuum filter were used to collect bioaerosols from the BEC. The bioaerosols were generated using a collison 6-jet nebulizer operated at 30 psi. During nebulization, the biological samples were suspended in 10 mL of water inside the collison reservoir. Nebulization was done until the level of the sample liquid was below the collison nebulizer jet inlet. This takes around 45 to 60 minutes to complete. All samplers were operated for the duration of nebulization and operated at 14.5 lpm. After collection, each sample catch was prepared for MALDI analysis using minimal

post collection sample handling techniques. However, since each sampler has different collection medium, sample handling protocols were varied accordingly.

For the FS collection, two analysis methods were tried, direct and indirect. For direct analysis, the filters were attached to a MALDI target using double sided tape. Once attached, MALDI matrix was applied to the filter and allowed to dry. Once dried, the MALDI target was inserted into the mass spectrometer and analyzed. For the indirect method, two techniques were tried: reverse deposit and tape pulls. For the reverse deposit method, both wet and dry surfaces were used. In each case, the filter paper was placed face down onto the MALDI target. After 15 to 30 seconds, the filter paper was removed. Once removed, MALDI matrix was deposited to the target surface and allowed to dry. For the tape pull extraction, double sided tape was adhered to a MALDI target. Once attached, the filter paper was placed face down on top of the tape. The filter paper was then peeled off the tape surface. Following removal of the filter paper, MALDI matrix was applied to the area of the tape where the filter sample was deposited. Two different filter types were investigated: cellulose and a borosilicate glass.

For the cyclone impactor, water was chosen as the collection medium due to its compatibility with MALDI analysis and its non-toxic nature. A total volume of 20 ml of collection medium was used during sampler operation as outlined in the operation manual. Following sample collection, the liquid volume was reduced to less than 1 mL to concentrate the sample. Sample concentration was done by centrifugation or lyophilization. Sample concentration is required due to the large collection medium volume. Once concentrated, the samples was diluted as needed for MALDI analysis.

For the AI, a glass microscope slide and a MALDI target were placed inside on the base plate during collection. These served as the impaction surface for collection. Following

collection, the glass microscope slide was gram stained as outlined in Appendix B. The gram staining was done to allow visual identification of the samples collected. For the MALDI target impacted samples, MALDI matrix was applied after an application of 1 μL of 1.0 % TFA and allowed to dry. Once dried, the samples were analyzed by MALDI-MS.

4.3 Results and Discussion

4.3.1 Characterization of the BEC

The BEC provides a sealed environment for the generation of biosafety level 1 bioaerosols for simultaneous sampling from up to four samplers. From Appendix A, the operating performance characteristics of the collision nebulizer provides the particle size distribution, liquid consumption rates and air flow rates for a given operating pressure. Therefore it is possible to adjust the pressure to increase the flow rates and thus particle concentrations. For this work, it is critical to ensure there is no unexpected pressure change during nebulization or sampling. At 30 psi, the collision nebulizer produces an air flow rate of 15.9 lpm. Since each sampler was operated at flow rate of 14.5 lpm, it is critical that a counter-balance of air volume be added to the BEC to prevent a pressure gradient from developing. This is achieved through the use of gas permeable nylon air filters. In addition, it is critical that the concentration and particle size distribution inside of the BEC is unchanged while sampling. This ensures a homogenous sampling environment and prevents any one sampler from experiencing particle bias as a result of uneven bioaerosol distributions within the BEC. To demonstrate this, an *E. coli* and cytochrome *c* bioaerosol was generated at 30 psi and characterized by the APS with and without samplers operating. As shown in Figure 38, the median and mean particle size

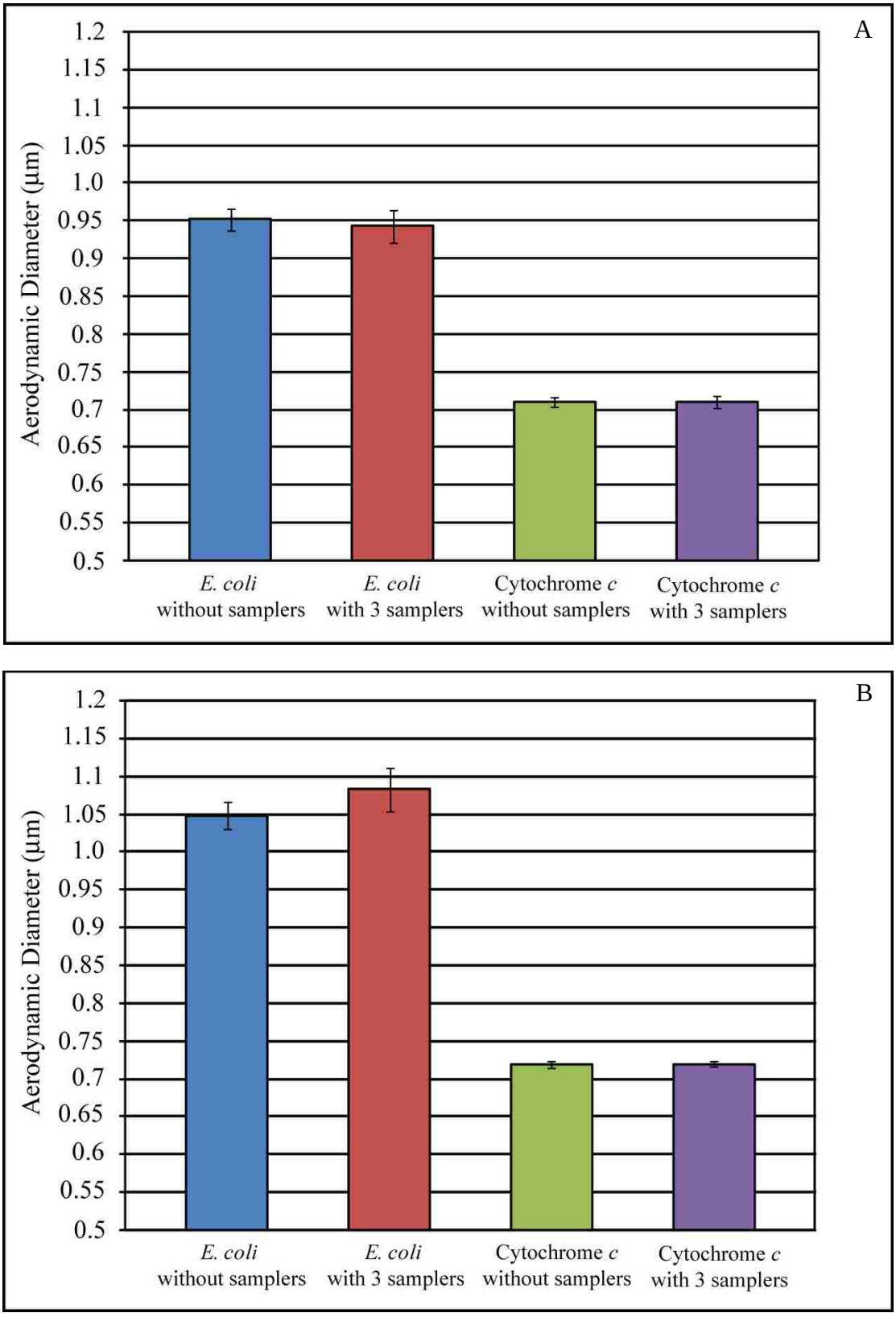


Figure 38. Bioaerosol (A) median and (B) mean particle sizes for *E. coli* and cytochrome *c* generated bioaerosols within the BEC with and without the samplers operating.

distributions are not significantly different between the sampler on and off events. The distribution of particle sizes is stable providing reliable system for sampler comparisons within the lab. In addition, the observed particle median and mean sizes are approximately 30% larger for *E. coli* than cytochrome *c* as expected due to the significant size difference between the two samples. Therefore, the BEC provides a safe and consistent sampling environment for bioaerosols.

4.3.2 Filter Collection

Filtration is a technique that captures particulates within fibrous material. In the case of bioaerosols, the particulates are biological material, specifically bacteria and proteins. While filters are capable of higher flow rates compared to non-filter samplers, a flow of 14.5 lpm was used in this work. This was done for two reasons. First, operating at this flow rate provided a more meaningful comparison with the other samplers used. Second, when using filter samplers in the future, a 14.5 lpm flow rate should reduce the amount of microorganism damage during sampling and allow viability studies to be conducted.

In this work, two bioaerosols, an *E. coli* and a cytochrome *c* bioaerosol, were generated separately and collected using both filter types. Figures 39 – 41 are MALDI-TOF mass spectra obtained from the cellulose based filter. For the cellulose based filter, both direct and indirect analysis methods were used for cytochrome *c*. However, it was not possible to produce results through the direct analysis of *E. coli*. This could be due to the nature of MALDI preparation and the complexity of the sample on the filter. For cytochrome c, the sample is composed of one analyte. Application of the matrix to this filter paper results in the integration of the entrained cytochrome *c* within the MALDI matrix crystal. Alternatively, the *E. coli* sample is more complex with varying amounts of protein with different concentrations. The initial results

suggest that the matrix to analyte ratio is not sufficient to provide good MALDI-MS signal. As a result, more matrix was added, but again, no positive results were seen. The application of additional matrix and a larger volume to the *E. coli* collected filter effectively diluted the sample or unintentionally extracted the sample from the filter. Similar findings with cytochrome *c* were observed when large volumes of matrix was used during matrix application to the cytochrome *c* collected filters, thus supporting this hypothesis. In addition, the tape pull method provides better quality mass spectra for both cases in that the baseline drift and noise is less than that observed for the reverse deposition technique.

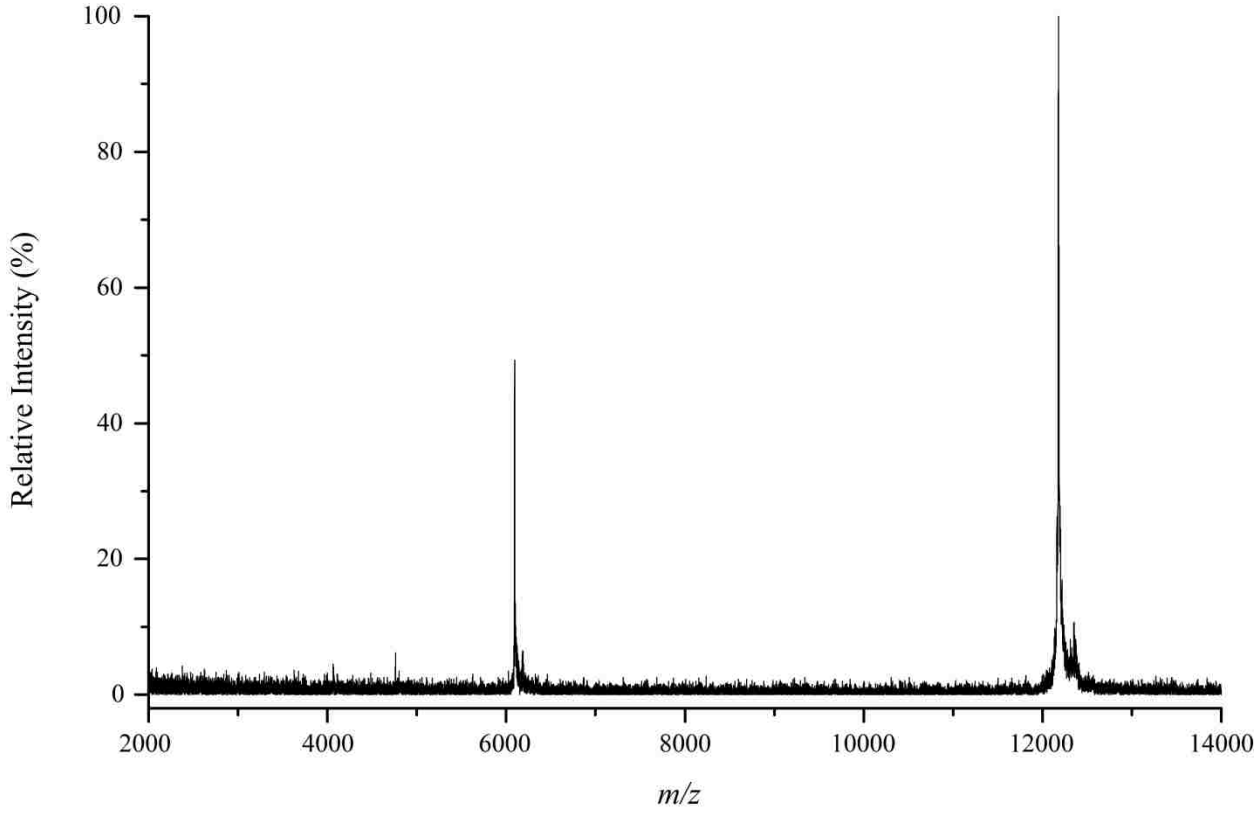


Figure 39. MALDI-TOF mass spectrum with SA for direct analysis from cellulose based filters of collected cytochrome c.

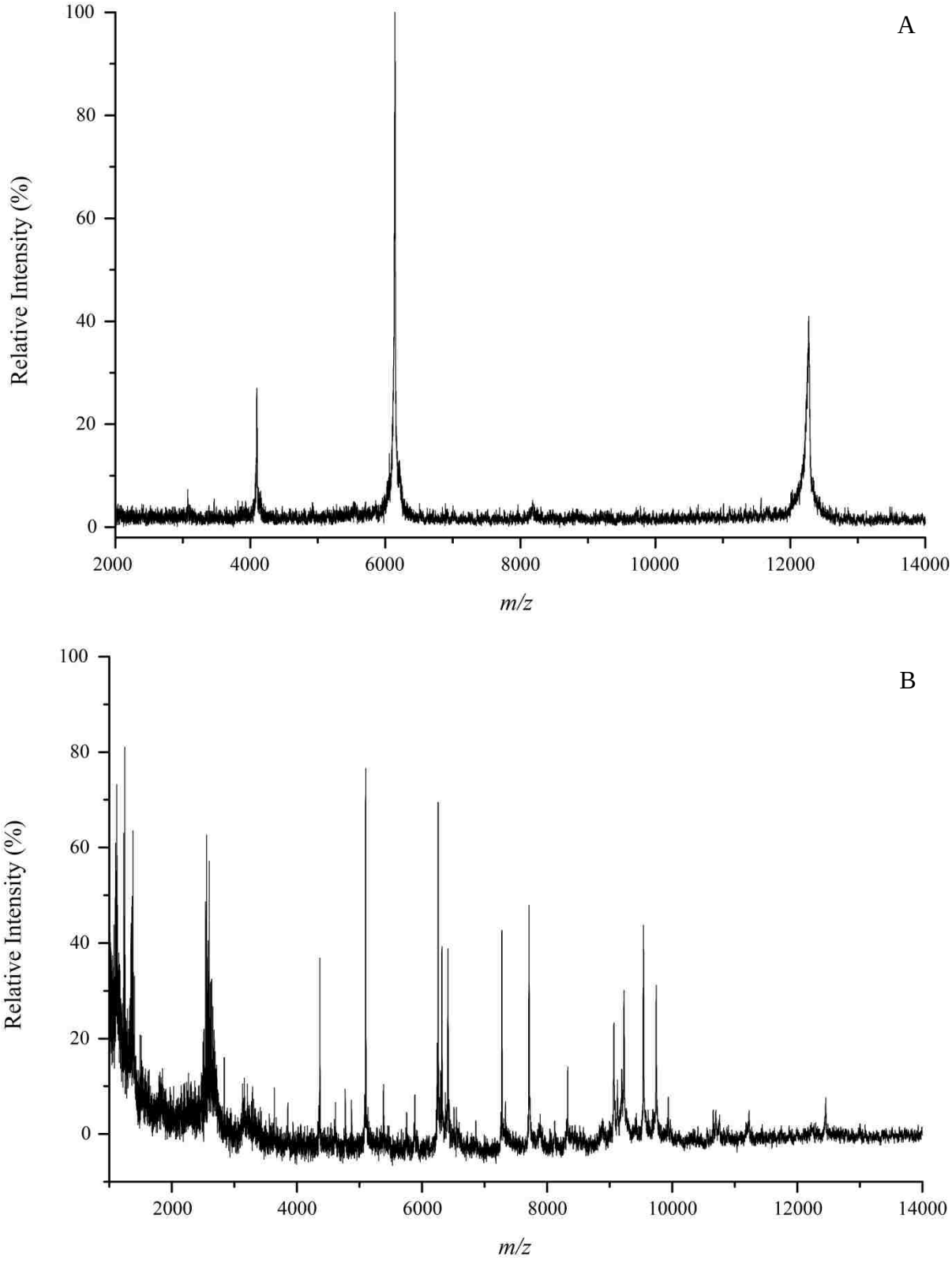


Figure 40. MALDI-TOF mass spectra with SA for reverse deposited samples from cellulose based filters of collected (A) cytochrome *c* and (B) *E. coli*.

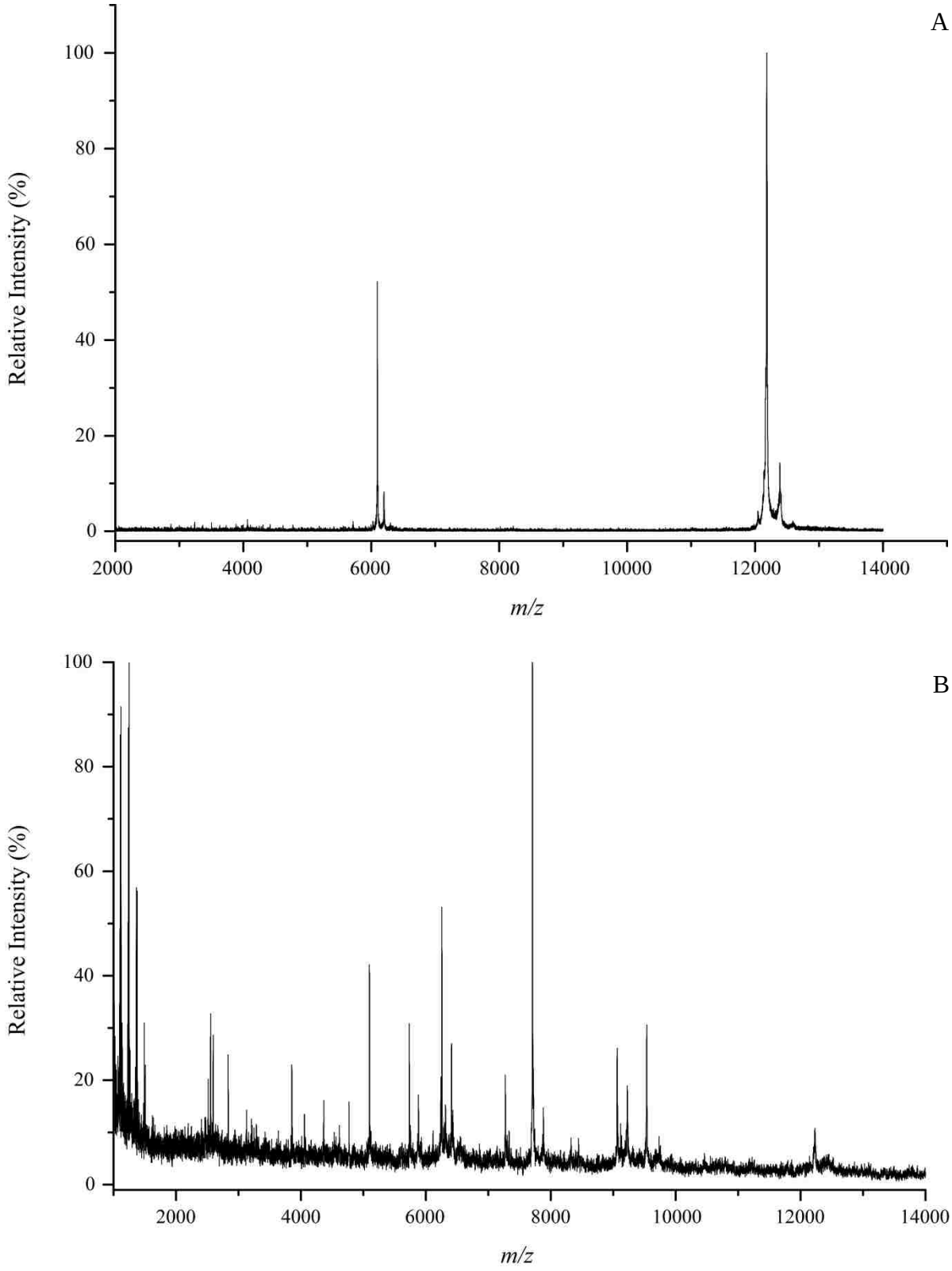


Figure 41. MALDI-TOF mass spectra with SA for tape pulled samples from cellulose based filters of collected (A) cytochrome *c* and (B) *E. coli*.

For the borosilicate filter, tape pull deposition was not possible because fibers from the paper were pulled apart during adhesion to the tape. However, reverse deposition (the process of removing material directly from one surface to another) onto a bare MALDI target and direct analysis was possible for cytochrome *c*. Figures 42 and43 are mass spectra from an MALDI-TOF mass spectrometer of collected cytochrome c and *E. coli* by direct analysis and reverse deposition. While cytochrome *c* was detected by direct analysis, *E. coli* was not. Additionally, the reverse deposition spectrum for cytochrome *c* provided better quality spectra over direct analysis as noted above in the preceding paragraph for the cellulose filter paper. While both filter types do provide mass spectra from bioaerosol collected samples, the cellulose based filter is easier to handle post collection as evidenced by the paper integrity and durability. Overall, vacuum filter collection of bioaerosols shows promise and is suitable for MALDI-MS analysis.

4.3.3 Impaction Samplers

The two impactors used in this work were an Andersen N6 single stage 400-hole impactor (AI) and a cyclone impactor (CI). The AI utilizes solid surfaces for particle impaction while the CI utilizes either solid (collection vessel wall) or liquid surfaces. The media used with the CI was water instead of the typical mineral oil. During operation, the water swirls due to the cyclonic air flow directed against the collection vessel wall. As the liquid swirls, the sample is swept off the walls and into the collection medium liquid reservoir. Once collection is complete, the liquid medium can be analyzed or concentrated prior to analysis. For the cytochrome *c* collection, the collection liquid developed a very slight pink tint indicative of a dilute cytochrome *c* solution. Due to the relatively low viscosity of water, the cyclonic action failed to produce a steady uniform wet-film across the impaction area on the vessel wall. As a result, cytochrome *c* deposits formed on the vessel wall.

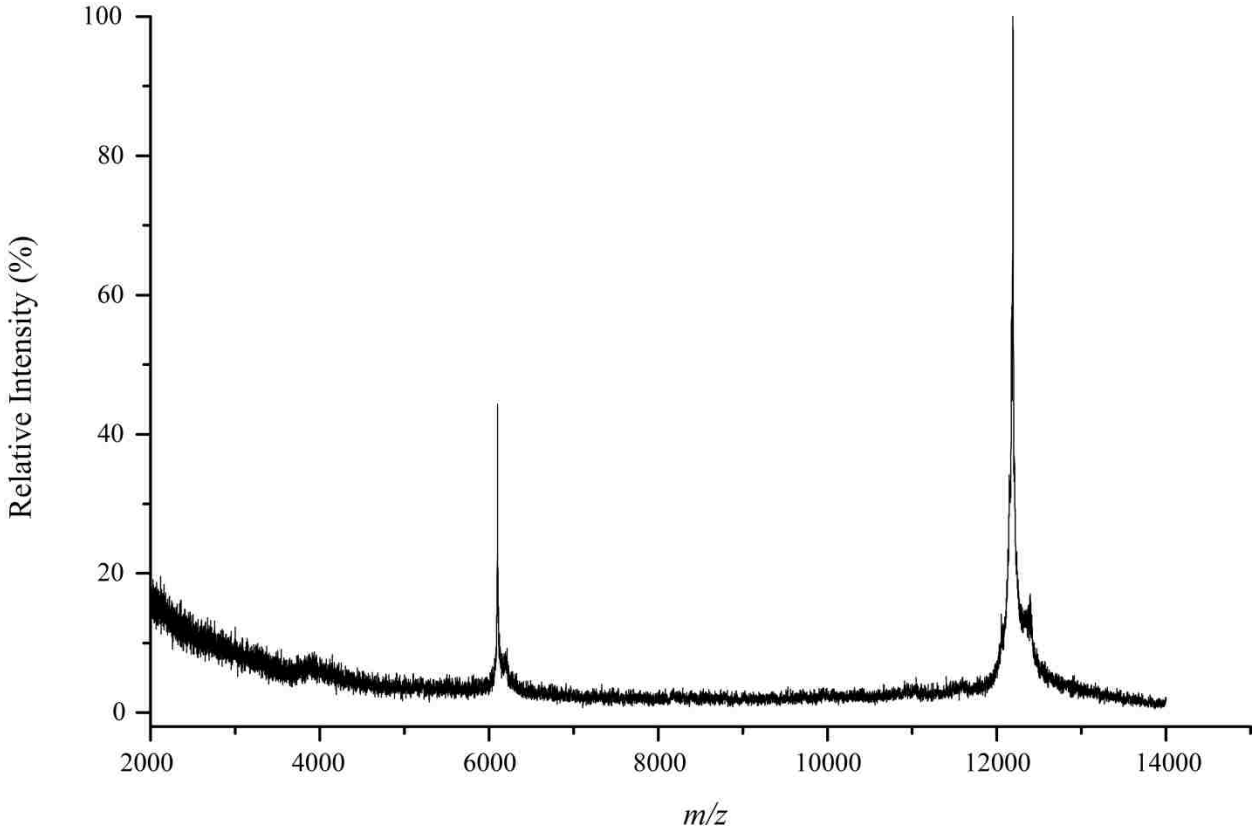


Figure 42. MALDI-TOF mass spectra with SA for direct analysis from borosilicate glass filters of collected cytochrome *c*.

Due to the unexpected deposition of cytochrome *c* on the collection vessel wall, two concentrating steps were tried: sample centrifugation and vessel wall washing. Since evaporating a large volume of water could take over 24 hours in a rotary evaporator or lyopholizer and cytochrome c is unable to form a pellet during centrifugation, the vessel wall washing proved to be the best method for obtaining an appropriately concentrated solution for analysis. Therefore, all subsequent collections were handled by first removing the liquid medium from the collection vessel followed by centrifugation. Subsequently, the collection vessel wall was washed with no greater than 1 mL of water and collected in a microcentrifuge tube. For cytochrome *c*, only the wall wash was able to produce good signal. However, for *E. coli* the centrifugation step was adequate as a result of the ease of pellet formation of bacteria. Once the pellet is formed, the

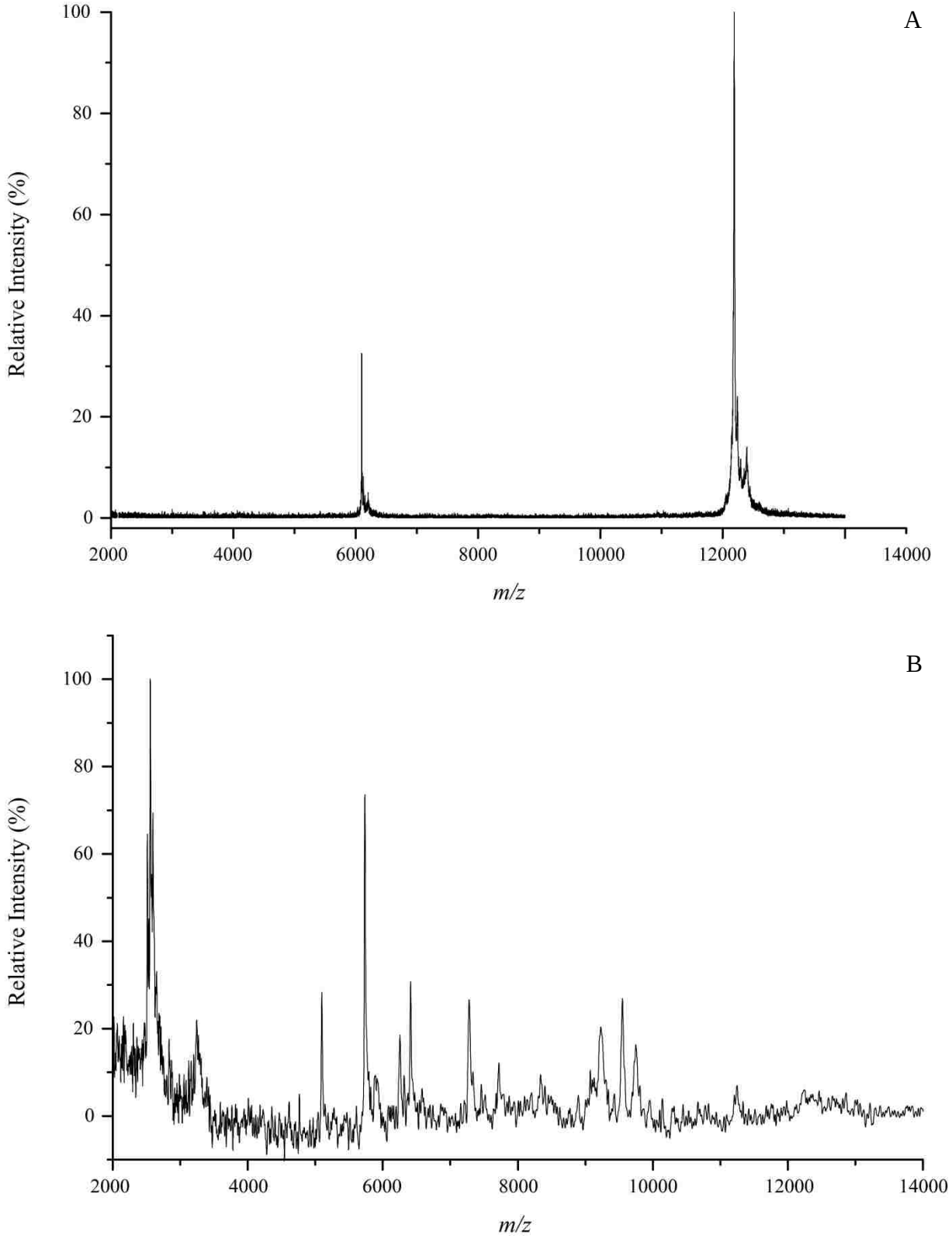


Figure 43. MALDI-TOF mass spectra with SA for reverse deposited samples from borosilcate glass based filters of collected (A) cytochrome *c* and (B) *E. coli*.

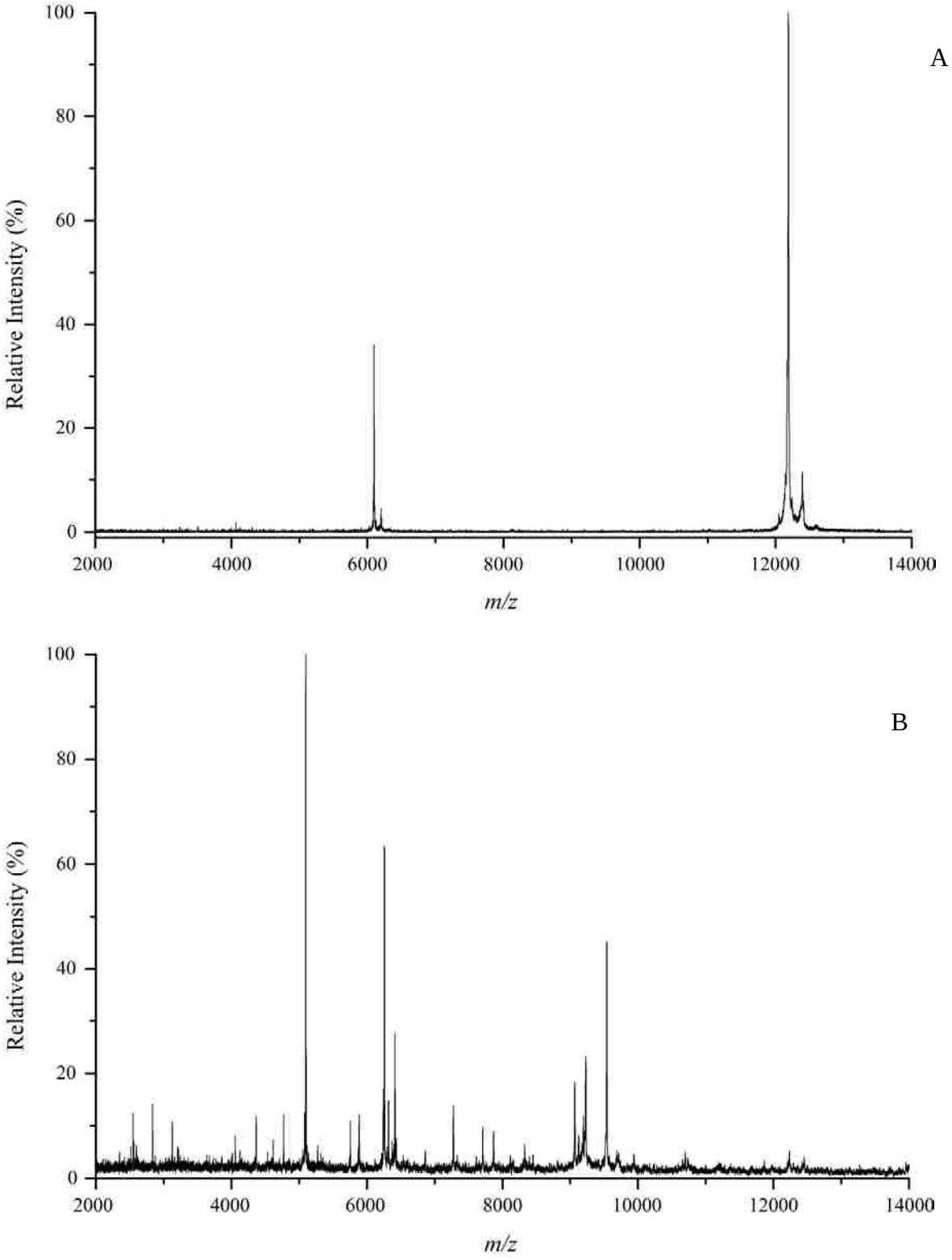


Figure 44. oMALDI-QTOF mass spectra of CI collected samples of (A) cytochrome *c* and (B) *E. coli* in FA.

supernatant was decanted and the pellet is resuspended in no greater than 1 mL of water. For unknown samples, both steps are necessary due to the unknown nature of the material, i.e., proteins or cells.

The benefit of using the CI is the variety of sample handling techniques possible. Once collected, the sample can be treated as any normal liquid sample. The sample can be used for culturing or any other technique for detection of biological material. Figure 44 shows the oMALDI-QTOF mass spectra of the collected CI samples. As evidenced by the spectra, CI collected samples greatly resemble those of the solution based samples from Chapter 3. While the CI sample does provide sample handling versatility and good quality spectra from collected bioaerosols, the concentrating step is critical. Therefore, both concentrating techniques are necessary in order to handle unknown samples. Specifically, if only protein toxins are present without any microorganisms as in the case of ricin, centrifugation will not be sufficient for sample analysis. Additionally, if no visible deposits are noted on the vessel wall in the case of a colorless material, the operator will be unable to determine which technique to use for concentrating. Therefore, a 1 mL wash down and resultant centrifugation of the remaining 19 mL volume must be done to determine whether a bacteria pellet forms thus increasing the overall time for analysis.

For the AI, a glass microscope slide and a stainless steel MALDI target were used for collection surfaces. After collection, the surfaces were removed from the AI for analysis. Upon initial inspection of the MALDI target following collection, spots were observed coinciding with each of the 400 jets of the AI accelerator plate. For analysis, an impacted spot was chosen for addition of solvents and MALDI matrix. On the impacted MALDI target, a 1μL volume of 1.0% TFA is first added to the impaction spot and allowed to dry. After the solvent dried, 1 μL of

MALDI matrix was then and allowed to dry. Once dried, the sample was analyzed by MALDI-MS. Figure 45 shows the spectra obtained from the analysis of the MALDI target impacted bioaerosols. To demonstrate the significance of the visible impacted spots, locations devoid of visible spots on the MALDI target were also analyzed. With the exception of insulin, only visible impaction spots produced signal. For insulin, the signal of the location devoid of an impaction spot was only 5% of the signal of insulin from an impaction spot. Since insulin was the smallest protein analyzed, this finding is consistent with inertial impactor theory which predicts the greatest deflection from the center of the impactor jet for the smallest particles. As a result, only the visible impacted spots should be analyzed while the locations devoid of visible spots can be used for peptide standard placement.

4.4 Summary

In this chapter, three samplers were investigated for offline MALDI compatibility. Each sampler can be used with appropriate post collection sample handling techniques. For filter collections, cellulose based paper appears to be better suited due to its robustness during sample collection. While direct analysis of bacteria was not accomplished, direct analysis of proteins from the filter was done. Additionally, sample extraction by adhesive tape removal or reverse deposition can be done for both proteins and bacteria. For the impactors, the CI is capable of collecting both protein and bacteria bioaerosols, however, proteins provide an additional challenge during concentration due to difficulties in centrifugation to pellet protein material. Despite this, the CI provides the advantage of being more suitable for post collection sample treatment such as proteolysis. For the AI, sample preparation is simple and fast. The AI provides the quickest time to analysis of all the samplers examined and is capable of collecting both bacteria and protein bioaerosols. Additionally, AI samples are handled the similarly

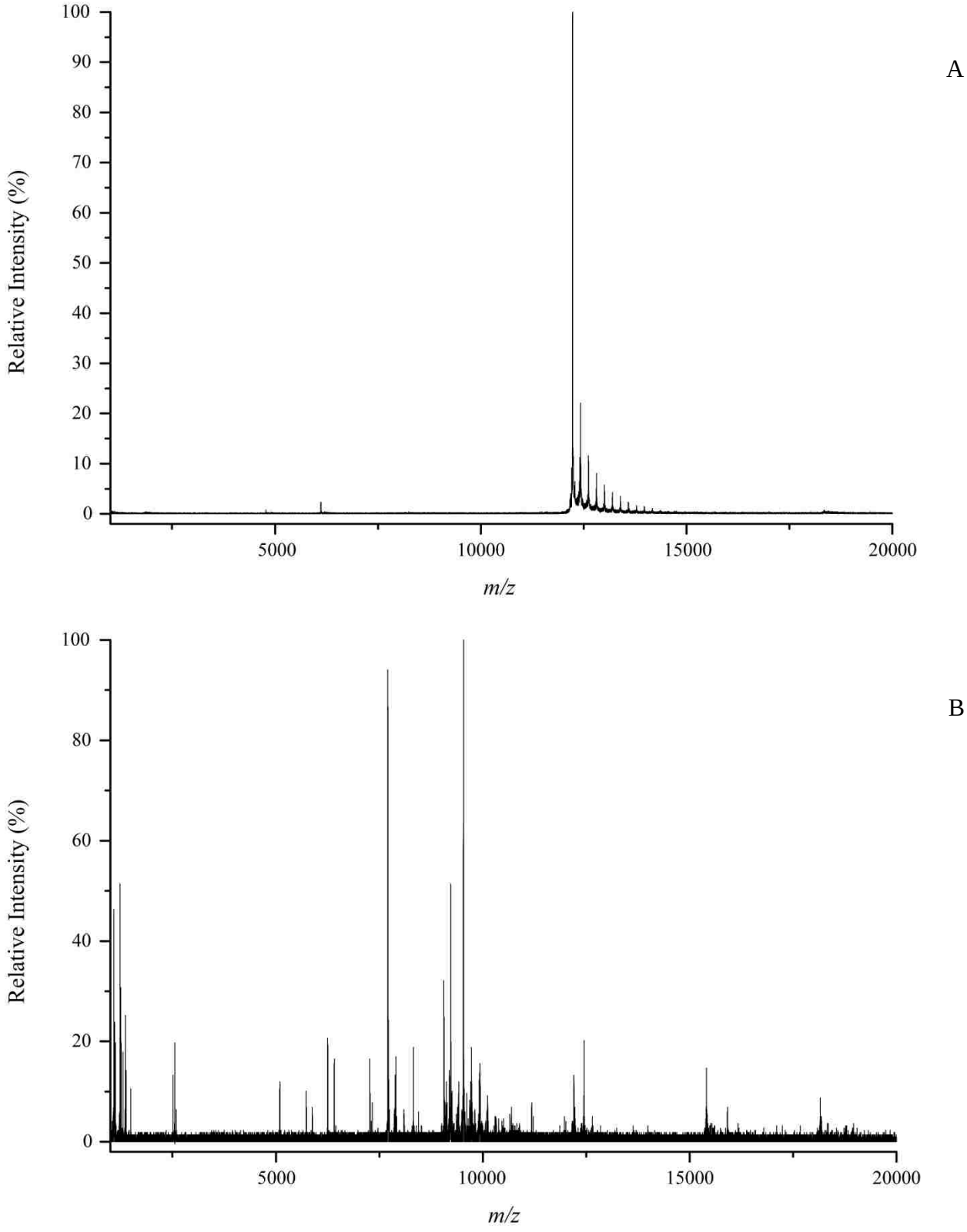


Figure 45. oMALDI-QTOF mass spectra of AI collected samples of (A) cytochrome *c* and (B) *E. coli* in FA.

unlike the CI sample catches which differ for proteins and bacteria. Even though the AI is the simplest and quickest, it has one major disadvantage, AI samples are difficult to process using post collection sample chemistry in contrast to CI samples.

While all three sampling techniques were useful for analysis of collected bioaerosols, the AI and CI techniques produce the largest signal with minimal noise and no baseline drift. However, with respect to dried droplet intact whole-cell bacteria MALDI-MS spectra, which produces on average 55 ± 5 peaks, the number of peaks found for the AI and CI spectra are fewer (39 ± 6 and 35± 7, respectively). For the filter papers, the cellulose paper performed to a higher level than the borosilicate based paper as observed by the greater number of peaks and overall spectra quality (low noise). The cellulose paper performed to nearly the same level as the AI and CI in that on average 34 ± 11 and 37 ± 8 peaks were observed from the reverse deposit and tape pull methods, respectively. Conversely, the borosilicate based filter paper was the worse sampler investigated and only produced on average 17 ± 6 peaks and typically was plagued by high noise and low signal intensity. Despite this, filter paper collections are a feasible approach for collecting bioaerosols for analysis by MADLI-MS.

CHAPTER 5. ON-TARGET PROTEOLUSIS OF IMPACTED BIOAEROSOLS USING MINI-WELLS

5.1 Introduction

Analysis of bioaerosols by MALDI-MS provides mass spectral peaks which can be used to search protein databases for tentative matches.[81] Often, these searches produce matches against a broad spectrum of organisms.[81, 86, 87] Therefore, any additional information obtained can result in an increase in the confidence of identification. Since most of the peaks observed in bioaerosol analysis by MALDI-MS of intact cells and cell products are proteins, one way to produce additional information is through proteolysis. In this chapter, enzymatic and chemical proteolysis of impacted bioaerosols using trypsin and cyanogen bromide is described. Since the primary bioaerosol sampler used in this work is an Andersen impactor (AI), the collected sample is deposited directly onto a stainless steel MALDI target as detailed in Chapter 4.

Proteolysis is a technique used to cleave proteins along the polypeptide backbone resulting in a systematic production of polypeptide fragments. Traditionally, this process involves denaturation, reduction, and alkylation in order to increase proteolysis efficiency. However, any of these procedures could potentially interfere with MALDI. Therefore, development of proteolysis methods without these steps is preferable since will minimize the potential interferences. Furthermore, performing proteolysis on impacted bioaerosols produces an additional complication because proteolysis, especially enzymatic proteolysis, requires more time than can be achieved with a liquid droplet on a flat surface before evaporation of the solvent. Several approaches to combat this problem are addresses in this work. Once proteolysis is achieved, peptide mass mapping is done using the proteomics search engine, MASCOT.[114] In addition, MS/MS experiments were done with selected fragments from the digested samples.

5.2 Experimental

Proteolysis protocols were developed for trypsin and CNBr without any denaturation, reduction or alkylation. Two proteolysis protocols were developed: *in vitro* and *in situ*. For proteolysis, trypsin reactions were carried out at 37 °C while CNBr reactions were performed at room temperature in a hood for safety. Reaction times varied from 2 to 24 hours before analysis. All reaction volumes were kept to a minimum with no reaction volume exceeding 100 μL. All samples were analyzed by MALDI-TOF or an oMALDI-QTOF mass spectrometry. Since proteolysis reactions results in the generation of lower mass polypeptides, CHCA and FA were the two MALDI matricies used in this work.

In vitro protocol development for proteolysis of native proteins and intact whole-cell bacteria was done using 0.5 mL microcentrifuge tubes as the reaction vessels. For trypsin digestions, a stock solution of 1mM trypsin in water was made fresh every 4 days. A concentration study of the digestion of cytochrome *c* by trypsin was done. In this work a 50 mM solution of cytochrome *c* was digested with varying amounts of trypsin. These reactions were carried out in equal volumes by varying the amount of stock trypsin added while adjusting to the final volume using water. The final amounts of trypsin in the reaction vessels were between 2 and 20 nanomoles. Reactions were initiated by transferring the appropriate stock trypsin volume into the reaction vessel and insertion of the reaction vessel into 37 °C incubator. For CNBr digestions, a stock solution of 2 M CNBr in acetonitrile was made fresh every week and stored in the freezer until needed. All reactions were carried out in a total of 25 μL: 10 μL of acid solution, 10 μL of CNBr solution and 5 μL of protein solution. The acid solution was varied and contained various concentrations of TFA, formic acid, or HCl. The CNBr was added last, capped immediately and shook vigorously for 30 seconds.

For the *in situ* protocols, two methods were investigated, direct application and mini-wells (Figure 46). In the direct application method, each proteolytic reagent is added to a single

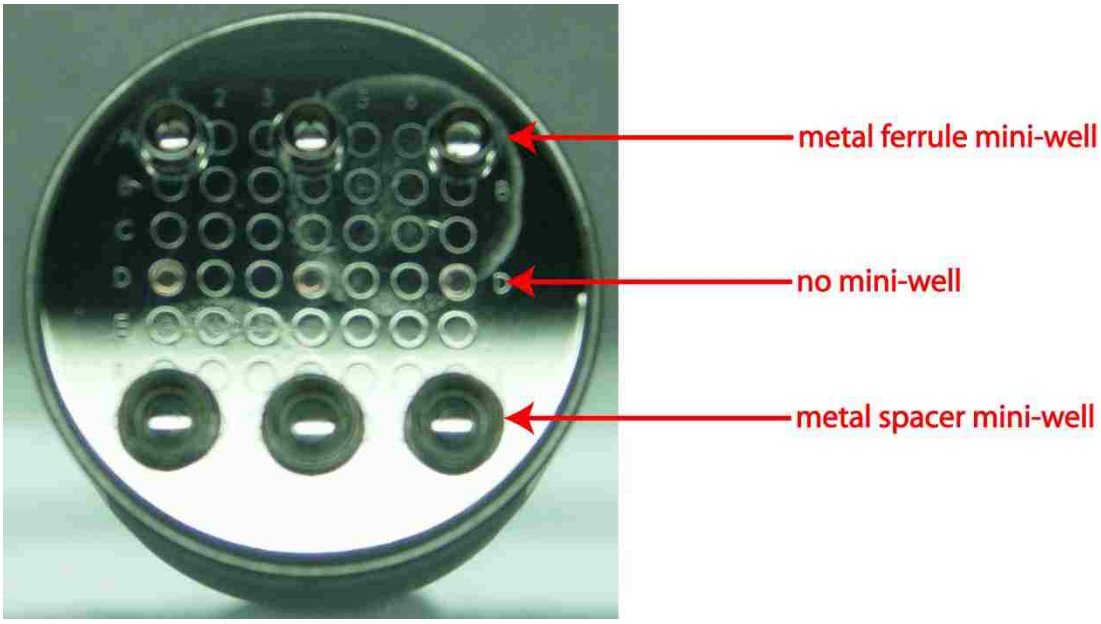


Figure 46. *In situ* mini-wells on an 7x7 OmniFlex MALDI target.

impacted spot on the MALDI target using a total volume on target no greater than 4 μL. The order of reagent addition was critical to make certain the solvents did not spread to other impacted spots. The first reagent added must contain the greatest amount of water to make sure the surface tension is adequate to limit spreading. For the trypsin reagents, this is not a problem since all the reagents are in water. This is not the case for the CNBr reagents. For CNBr, 1 μL of water was first added to the impacted spot followed by the addition of the acid solution. The last component added was the CNBr because it was dissolved in acetonitrile. In this method, the reagents are usually dry within 10 minutes. The only way to increase the reaction time is to rewet the spot or to use alternate methods such as mini-wells, which serves to increase the total volume applied to the impacted spot and decrease the evaporation rate.

The mini-wells used in this work were small diameter metal ferrules around 5 times the diameter of one impacted spot. This provided room for the insertion of the mini-well onto the

MALDI target. Once the mini-well was in place, the proteolytic reagents were added. Again, the reagent containing the most water was added first since the mini-well is not sealed to the MALDI target. The mini-wells are capable of holding up to 15 μL. After the reagents are added to the mini-well, the MALDI plate is placed in an empty petri dish and covered. When trypsin was used, the entire petri dish was placed in the heater set at 37 °C. For the CNBr, the petri dish was left at room temperature in a hood. The complete evaporation of the solvents typically took 2 hours and 1 hour for the trypsin and CNBr reactions, respectively. When the solvents were completely evaporated, the MALDI target was removed from the petri dish and the mini-well detached from the target with forceps. Once the mini-well was removed, matrix was added onto the digested spot and allowed to dry.

MALDI-MS analysis was done on all digested samples by oMALDI-QTOF-MS. Each full scan TOF spectrum was analyzed for proteolytic peptide fragments. From the full scan, peaks were selected and MS/MS experiments were performed using low energy CID. This was done by setting Q_1 to the desired parent mass with a low mass resolution to allow transmission of the parent peaks isotopes. Argon was used as the collision gas in q_2 and the collision energy was adjusted according to Equation 16 based on the precursor mass. Equation 16 only serves as a guide and collision energies can be adjusted based on the fragmentation efficiency.

$$CE = \frac{[Peptide\ m/z]}{20} \qquad \textbf{Equation 16}$$

From the combined information of the full spectrum and MS/MS, two MASCOT searches were performed: peptide mass fingerprinting and sequence tag searching. For the MASCOT searches, a peptide mass tolerance of 1.2 and 0.6 Da was used for the MS and MS/MS tolerance, respectively. Additional search parameters included the use of no variable or fixed modifications, up to 3 missed cleavages and the monoisotopic masses as input data. The search

results indicated the closest protein match and the match confidence using the MOWSE peptide fragment database and scoring algorithms.[116]

In order to compare the results from the MS and MS/MS experiments, *in silico* proteolysis was done using both trypsin and CNBr using the Protein Prospector version 4.0.4.[117] All protein sequences used were obtained from the SWISS-PROT database in FASTA format. All *in silico* digestions were set to provide peptide fragments for MALDI-MS with masses between 1000 and 6000 and a maximum of three missed cleavages. In addition, no post-translational modifications were selected as these were unknown.

5.3 Results and Discussion

5.3.1 *In Vitro* Proteolysis of Cytochrome c

Before conducting *in situ* proteolysis, protocols were developed for *in vitro* digestions. Since the two priorities are to limit possible MALDI interferents and reduce time for proteolysis, steps were taken to reduce the number of reagents while minimizing the time required for digestion. For trypsin proteolysis, a slightly alkaline pH increases enzymatic efficiency. However, this requires the introduction of a buffer. Therefore, initial trials were done without a buffer. As shown in Figure 47, the trypsin was inactive for up to 24 hours without a buffer. After the 24 hour period, trypsin was made active by the addition of 10 μL from a 100 mM NH_4HCO_3 (pH 7.8) solution. Two hours following the buffer addition, five peptide fragments were observed for cytochrome *c* and the complete disappearance of the intact cytochrome *c* peak was observed. These masses were then compared to the peptide fragments from an *in silico* trypsin digestion. Tables 16 and 17 contain the experimental and *in silico* tryptic peptide masses, respectively.

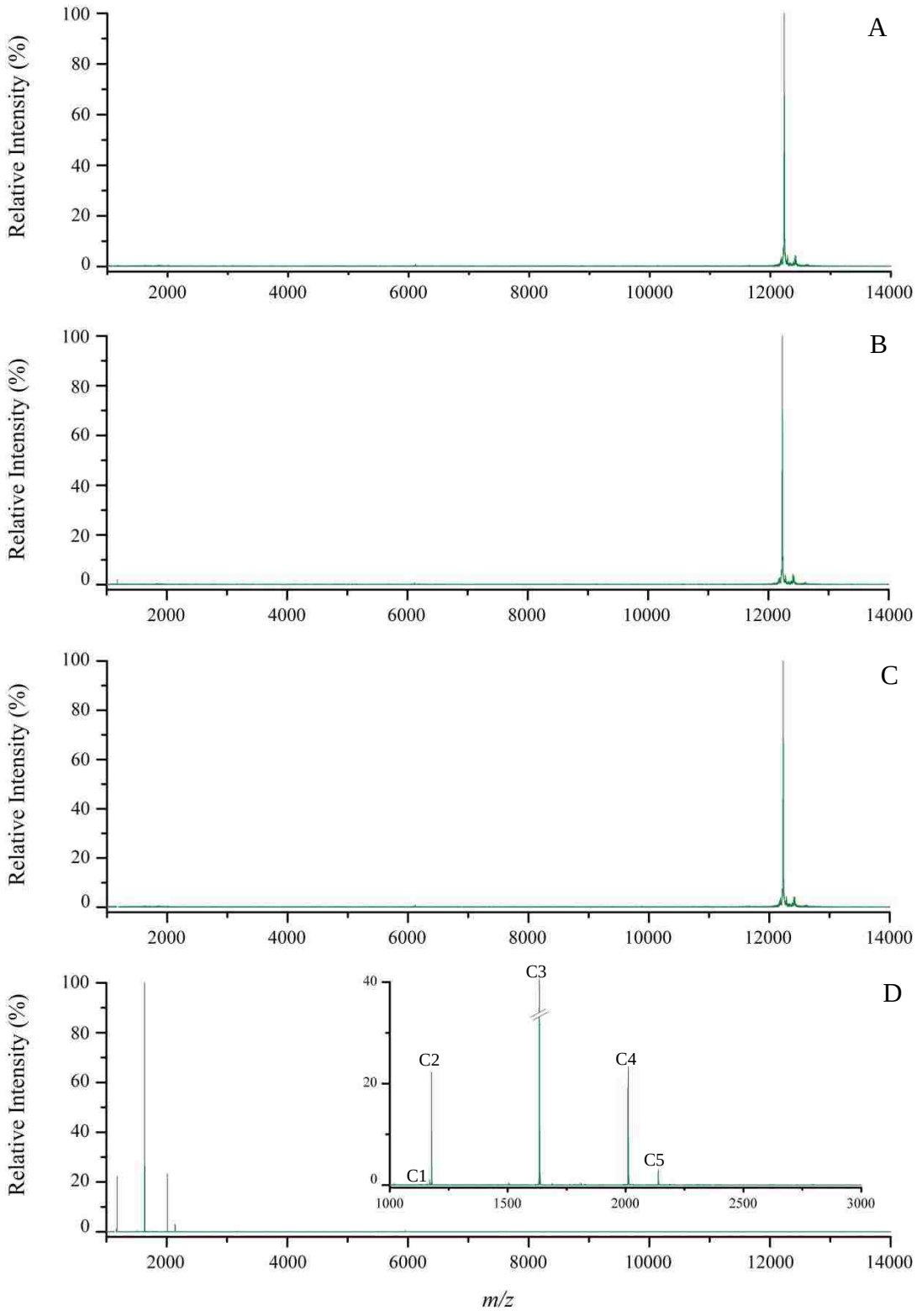


Figure 47. oMALDI-QTOF-MS of cytochrome *c* in FA (A) control (B) 2 hours post trypsin (C) 24 hours post trypsin (D) 26 hours post trypsin and 2 hours post buffer addition. Peak labels for peptide fragments are provided in Table 17.

Table 16. Experimental masses of trypsin digested fragments of cytochrome *c* obtained by *in vitro* proteolysis and oMALDI-QTOF-MS in FA.

Peak Label	Experimental Mass
C1	1168
C2	1176
C3	1633
C4	2009
C5	2138

Table 17. Protein Prospector *in silico* trypsin digestion of cytochrome *c* with one missed cleavage. Asterisks mark the matched peaks as found in Table 16.

MGDVEKGKKI FVQKCAQCHT VEKGGKHKTG PNLHGLFGRK TGQAPGFSYT
DANKNKGITW GEETLMEYLE NPKKYIPGTK MIFAGIKKKG EREDLIAYLK
KATNE

Mono-isotopic mass	Start	Stop	# of Missed Cleavages	Sequence
964.5355	93	100	0	(R)EDLIAYLK(K)
1018.445	15	23	0	(K)CAQCHTVEK(G)
1168.6227*	29	39	0	(K)TGPNLHGLFGR(K)
1456.6708	41	54	0	(K)TGQAPGFSYTDANK(N)
2009.953*	57	73	0	(K)GITWGEETLMEYLENPK(K)
806.4776	74	80	1	(K)KYIPGTK(M)
863.4297	1	8	1	(-)MGDVEKGK(K)
907.5439	81	88	1	(K)MIFAGIKK(K)
1092.6305	93	101	1	(R)EDLIAYLKK(A)
1260.5829	15	26	1	(K)CAQCHTVEKGGK(H)
1296.7177	29	40	1	(K)TGPNLHGLFGRK(T)
1306.7007	90	100	1	(K)GEREDLIAYLK(K)
1433.7766	27	39	1	(K)HKTGPNLHGLFGR(K)
1438.8132	75	87	1	(K)YIPGTKMIFAGIK(K)
1584.7658	40	54	1	(R)KTGQAPGFSYTDANK(N)
1633.8194*	10	23	1	(K)IFVQKCAQCHTVEK(G)
1698.8087	41	56	1	(K)TGQAPGFSYTDANKNK(G)
2138.048*	57	74	1	(K)GITWGEETLMEYLENPKK(Y)
2252.0909	55	73	1	(K)NKGITWGEETLMEYLENPK(K)

In vitro CNBr proteolysis was also done as shown in Figure 48. For this, the experimental fragment masses and the *in silico* digestion masses are listed in table 18 and 19, respectively. Exact mass matches between the *in silico* digest and *in vitro* digest were limited to those fragments that did not contain cysteines 15 and 18. These cysteine residues are covalently linked to the prosthetic heme group in cytochrome c. As such, the mass of the heme is not accounted for in the *in silico* digest. However, for peaks 1 and 2, the masses correspond to the *in silico* masses due to the absence of the cysteine and thus the heme prosthetic group. In addition, decarboxylation of the heme carboxyl groups as shown in Figure 49 seems to be present as both a single and double decarboxylation event resulting in a further spectral complexity. Regardless, two MASCOT searches were done, one using the experimental masses found and the other using the adjusted experimental masses to remove the prosthetic group mass information.

5.3.2 *In Vitro* Proteolysis of E. coli

E. coli was used to optimize the proteolysis and MALDI protocols for intact bacteria,. Traditionally, this would include cell lysing and centrifugation in order to release the proteins and remove cellular debris. However, these steps were not included so that the protocols could be adapted to on target proteolysis. Both trypsin and CNBr proteolysis were investigated. Since these protocols were developed for intact whole cells, no *in silico* digestion is possible due to the complex nature of bacterial proteomes. Therefore, all proteolytic sample mass spectra were first compared to the mass spectrum of the undigested *E. coli* as shown in Figures 50 and 51. All peaks observed in the proteolysis samples and not in the control sample were recorded as tentative unique peptides. An additional uniqueness check is needed for trypsin digested samples. For these samples, the peak *m/z* values were compared to the *in silico* digest products of trypsin autolysis products. Matched values were excluded from the list of unique peptide masses.

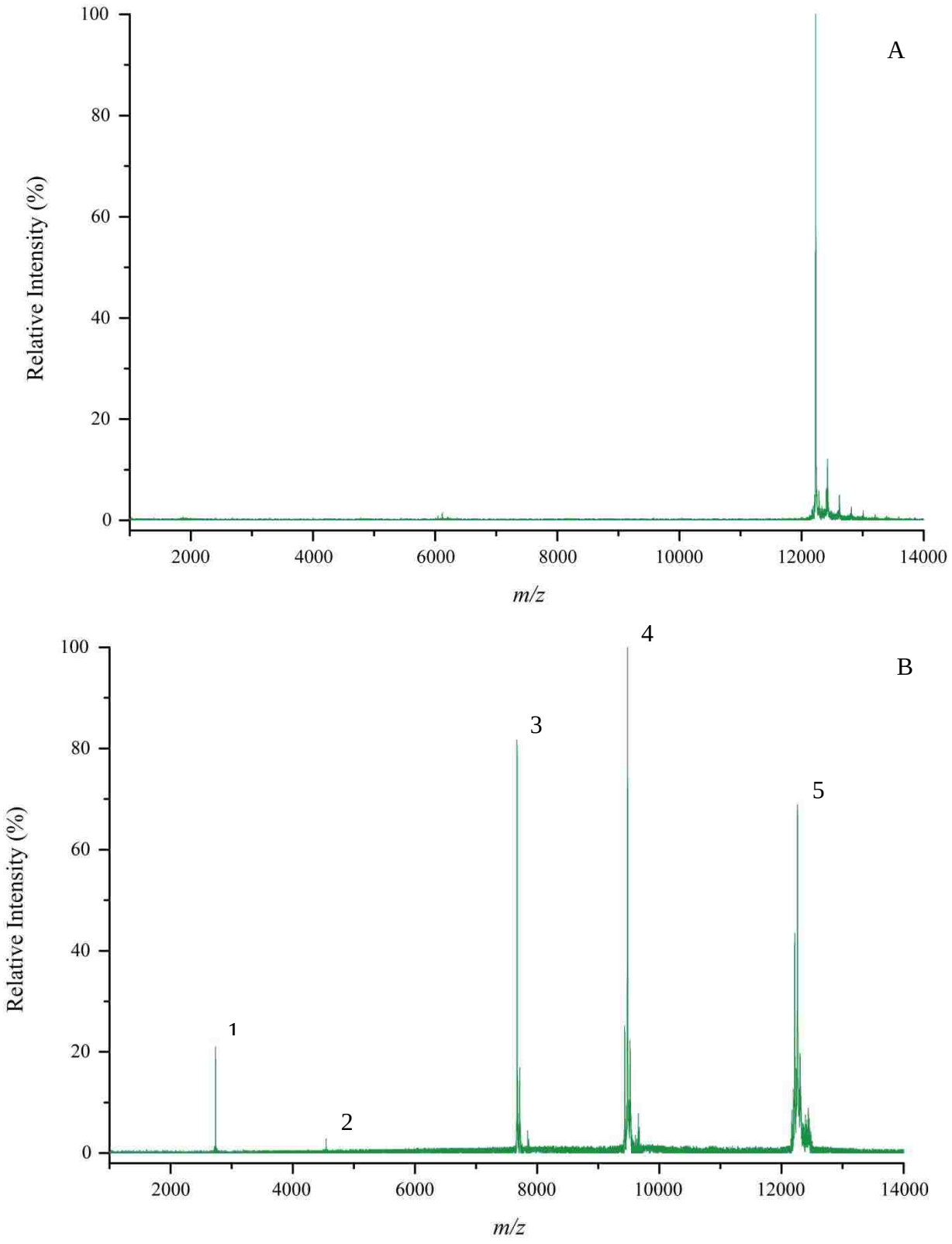


Figure 48. oMALDI-QTOF-MS of cytochrome *c* in FA (A) undigested (B) *in vitro* CNBr digested

Table 18. Experimental masses of CNBr digested fragments of cytochrome *c* obtained by *in vitro* proteolysis and analysis on an oMALDI-QTOF-MS in FA.

Peak Label	Experimental Mass	Mass Difference from *in silico*
1	2737	N/A
2	4527	N/A
3	7668	$[M+H]^+$ {M = fragment 1–66h + heme – 2[CO_2]}
4	9478	$[M+H]^+$ {M = fragment 1–81h + heme – 2[CO_2] + H_2O}
5	12,216	$[M+H]^+$ { M = intact cytochrome *c* + heme – 2[CO_2]}

Table 19. *In silico* digestion of cytochrome *c* with up to three missed cleavages. Asterisks mark the matched peaks as found in Table 20. Mass values may not match due to prosthetic group addition and/or fragmentation for peptides containing cysteines 15 and 18.

MGDVEKGKKI FVQKCAQCHT VEKGGKHKTG PNLHGLFGRK TGQAPGFSYT
DANKNKGITW GEETLMEYLE NPKKYIPGTK MIFAGIKKKG EREDLIAYLK
KATNE

Mono-isotopic mass	Start	Stop	# of Missed Cleavages	Sequence
1761.90	67	81	0	(h)EYLENPKKYIPGTKh
2735.55*	82	105	0	(h)IFAGIKKKGEREDLIAYLKKATNE
7008.53	2	66	0	(h)GDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLh
4527.47*	67	105	1	(h)EYLENPKKYIPGTKMIFAGIKKKGEREDLIAYLKKATNE
7139.59*	1	66	1	(–)MGDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLh
8818.46	2	81	1	(h)GDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLMEYLENPKKYIPGTKh
8931.50*	1	81	2	(–)MGDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLMEYLENPKKYIPGTKh
11566.03	2	105	2	(h)GDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLMEYLENPKKYIPGTKMIFAGIKKKGEREDLIAYLKKATNE
11697.07*	1	105	(intact) 3	(–)MGDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGFSYTDANKNKGITWGEETLMEYLENPKKYIPGTKMIFAGIKKKGEREDLIAYLKKATNE

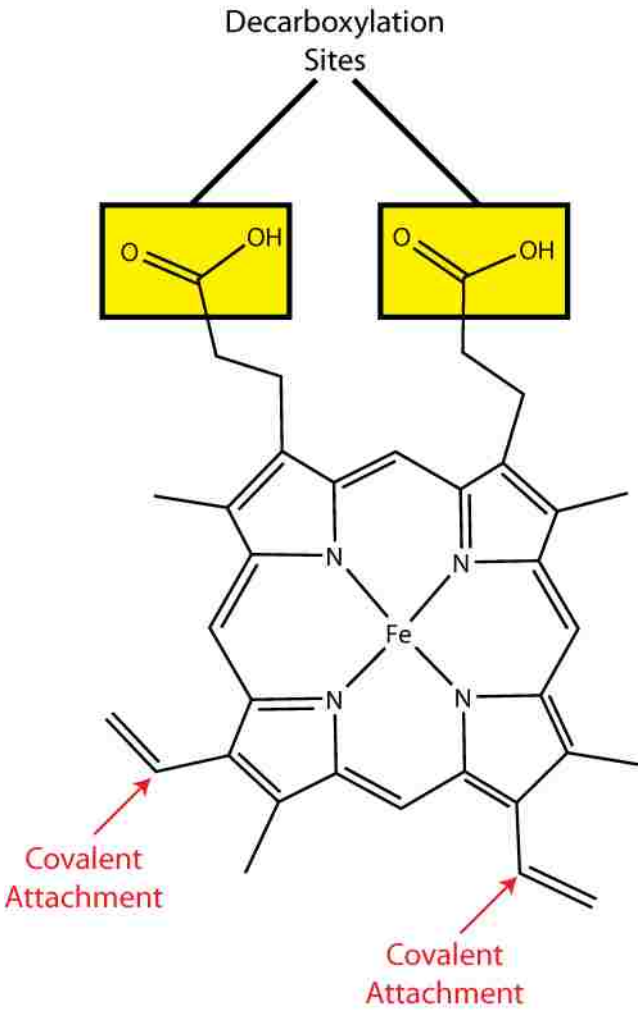


Figure 49. Heme *c* structure from cytochrome *c*.

Finally, an SRS database search was done on all peaks labeled as tentative to determine whether the new peaks were due to proteolysis or were new proteins resulting from new sample conditions. Tables 20 and 21 list the tentative unique peaks observed for proteolysis using trypsin and CNBr, respectively.

For the trypsin digestion, a total of 14 peaks were observed under 6000 *m/z* and none matched the peaks observed from the control *E. coli* in the same mass region. Of the 14 peaks, only one peak (peak 2162) matched a trypsin autolysis fragment. For the CNBr digestion, 16 peaks were observed in mass region under 10,000 *m/z*. Out of the 16 peaks, 12 were identified as

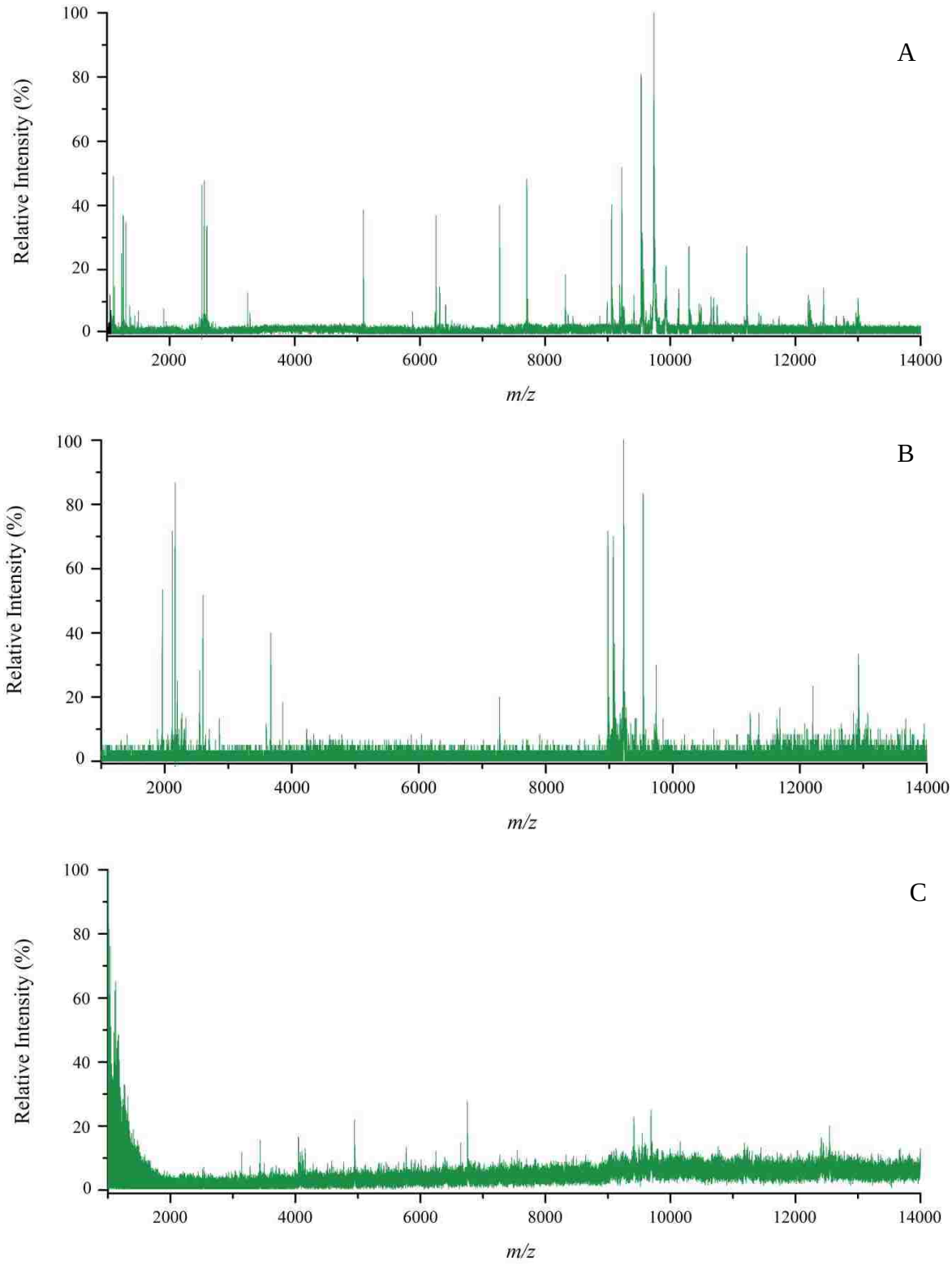


Figure 50. oMALDI-QTOF-MS of *E. coli* in FA (A) control (B) Trypsin digested (C) CNBr digested.

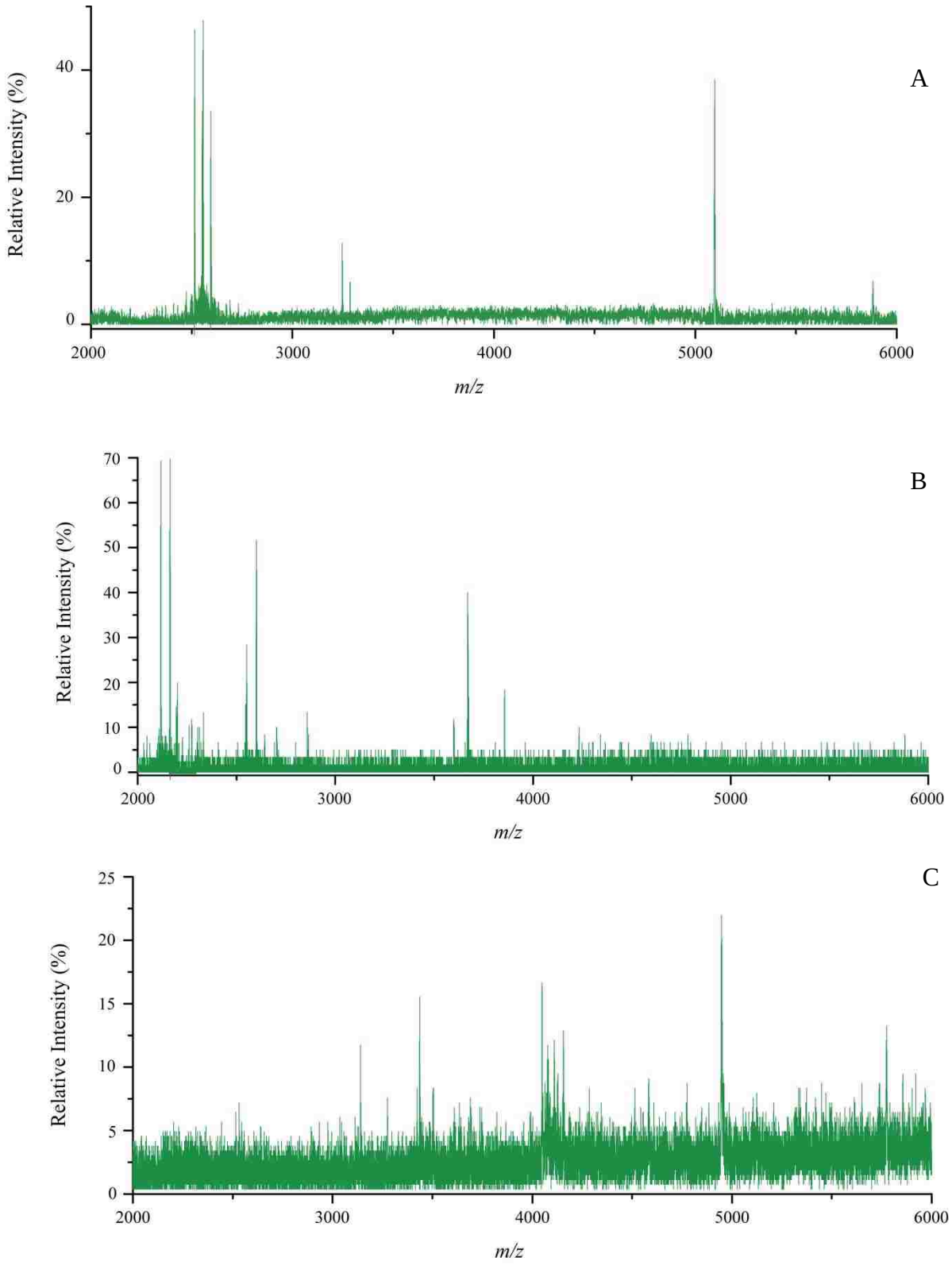


Figure 51. oMALDI-QTOF-MS expanded low mass region (2000–6000) of *E. coli* in FA (A) control (B) Trypsin digested (C) CNBr digested.

Table 20. Unique trypsin peptides from *in vitro* intact whole cell proteolysis of *E. coli*.

Peak *m/z*	SRS Hit	Description	Iteration Number
1961	No	N/A	2
2116	No	N/A	3
2194	No	N/A	3
2200	No	N/A	4
2259	No	N/A	1
2272	No	N/A	4
2545	No	N/A	3
2551	No	N/A	2
2601	No	N/A	2
2858	No	N/A	2
3599	No	N/A	1
3669	No	N/A	1
3855	No	N/A	1

Table 21. Unique CNBr peptides from *in vitro* intact whole cell proteolysis of *E. coli*.

Peak *m/z*	SRS Hit	SRS Description	Iteration Number
2297	No	N/A	
2640	No	N/A	
3140	No	N/A	
3435	No	N/A	
4155	No	N/A	
4583	No	N/A	
4944	No	N/A	
5771	YES	Putative uncharacterized protein ECs1334	
6206	No	N/A	
6752	No	N/A	
7269	YES	Putative uncharacterized protein nine Bacteriophage lambda nin 60-like protein Putative uncharacterized protein ECs2980	
9690	No	N/A	

unique peaks. Of the four non-unique, two matched peaks from the control and two were found in a SRS search as an *E. coli* protein. Therefore, those peaks were not treated as proteolytic peptide fragments.

5.3.3 *In Situ* Proteolysis of Collected Bioaerosols Using Mini-Wells

Collection of bioaerosols directly on a stainless steel MALDI target is the fastest way to sample and analyze bioaerosols by mass spectrometry. An adaptation of the proteolytic technique described in the previous section increases the analytical power for identifying the biological components from bioaerosols. One of the major difficulties in performing proteolysis on a MALDI target is the lack of container vessel, which limits the volume of reagents used. One approach to overcome this disadvantage was the use of immobilized trypsin.[87] By immobilizing the trypsin, the autolysis fragments were limited and this allowed a larger concentration of the protease to be used. Despite this, the sample volume was still limited and a humidifier chamber was necessary to maintain the reaction volume. In this work, a new technique utilizing free trypsin for the proteolysis of impacted bioaerosols without the use of a humidifier chamber was investigated. In order to accomplish this, individual removable mini-wells were placed directly over an impacted bioaerosol spots on a MALDI target. The proteolytic reagents were then be added to the mini-well. Once the reaction solvent evaporated, the mini-well was removed and the target was analyzed after MALDI sample preparation for impacted bioaerosols.

Impacted bioaerosols of cytochrome *c* and *E. coli* were proteolyzed using trypsin and CNBr protocols with mini-wells. The mini-wells hold a maximum volume of 15 μL, therefore, all reaction volumes were scaled down to a volume of 12 μL. All digestions were finished within 2 hours after adding the proteolytic reagents. Figures 52 through 55 show the mass spectra from the mini-well digestions. The cytochrome *c* trypsin digestions produced results which were consistent with the *in vitro* digestion except for the absence of peaks 1168 and 1176 and the appearance of two new peaks at *m/z* 3154 and 11645 corresponding to partially digested protein. The disappearance of the low mass peaks and the appearance of peaks from two new partially

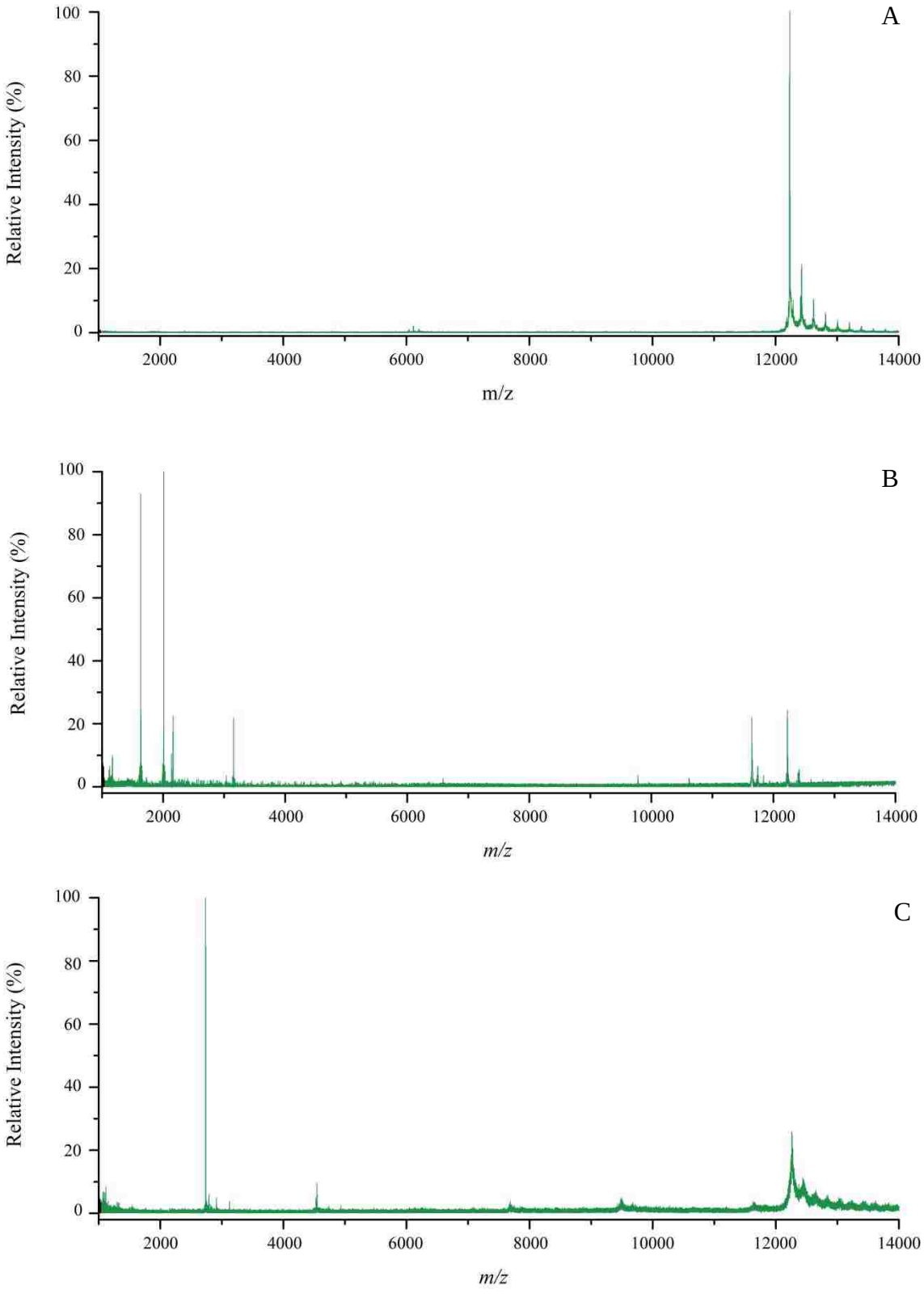


Figure 52. oMALDI-QTOF-MS of Cytochrome *c* from impacted *in situ* mini-wells proteolysis in FA (A) control (B) Trypsin (C) CNBr.

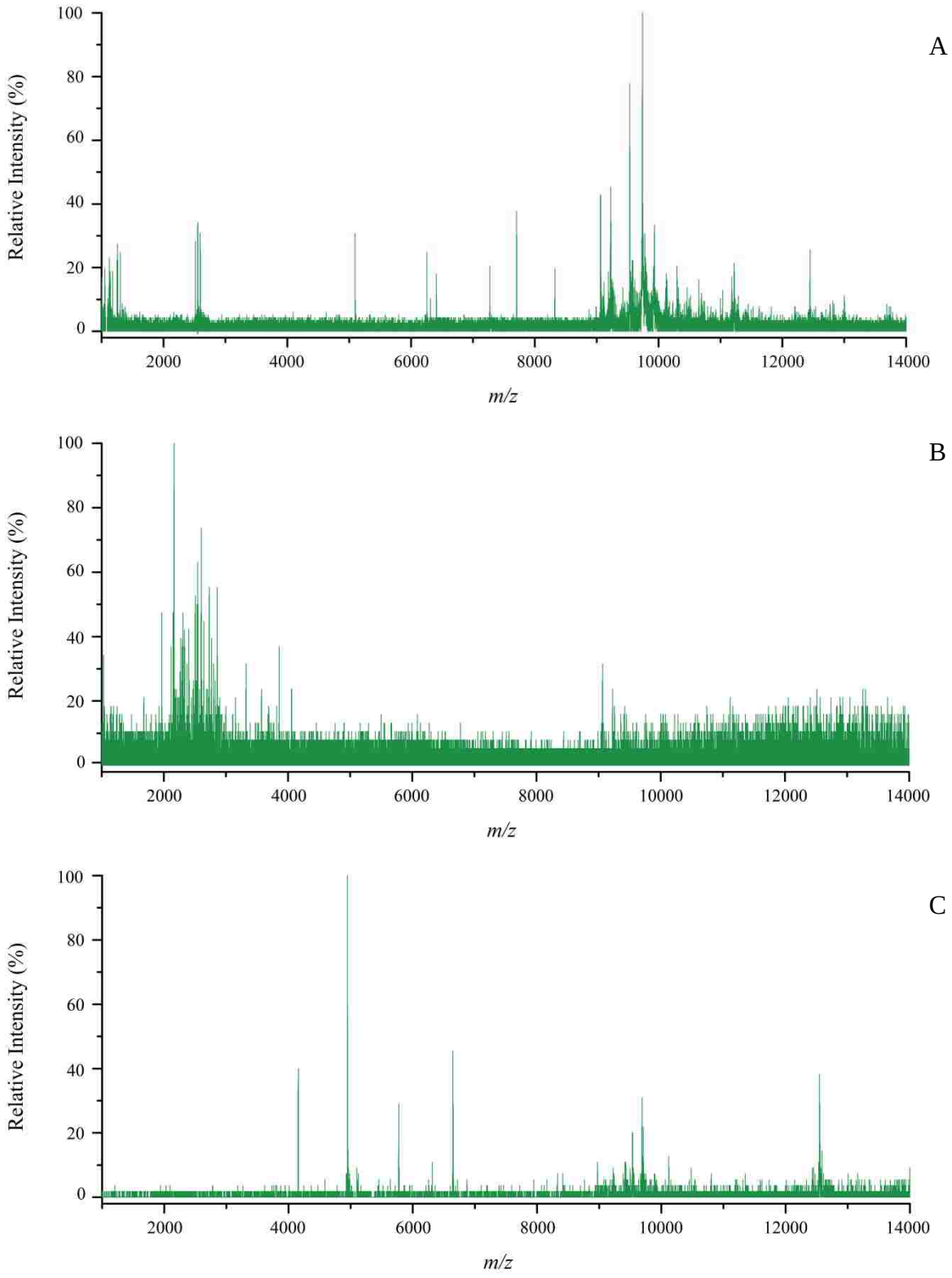


Figure 53. oMALDI-QTOF-MS of *E. coli* from impacted *in situ* mini-wells proteolysis in FA (A) control (B) Trypsin (C) CNBr.

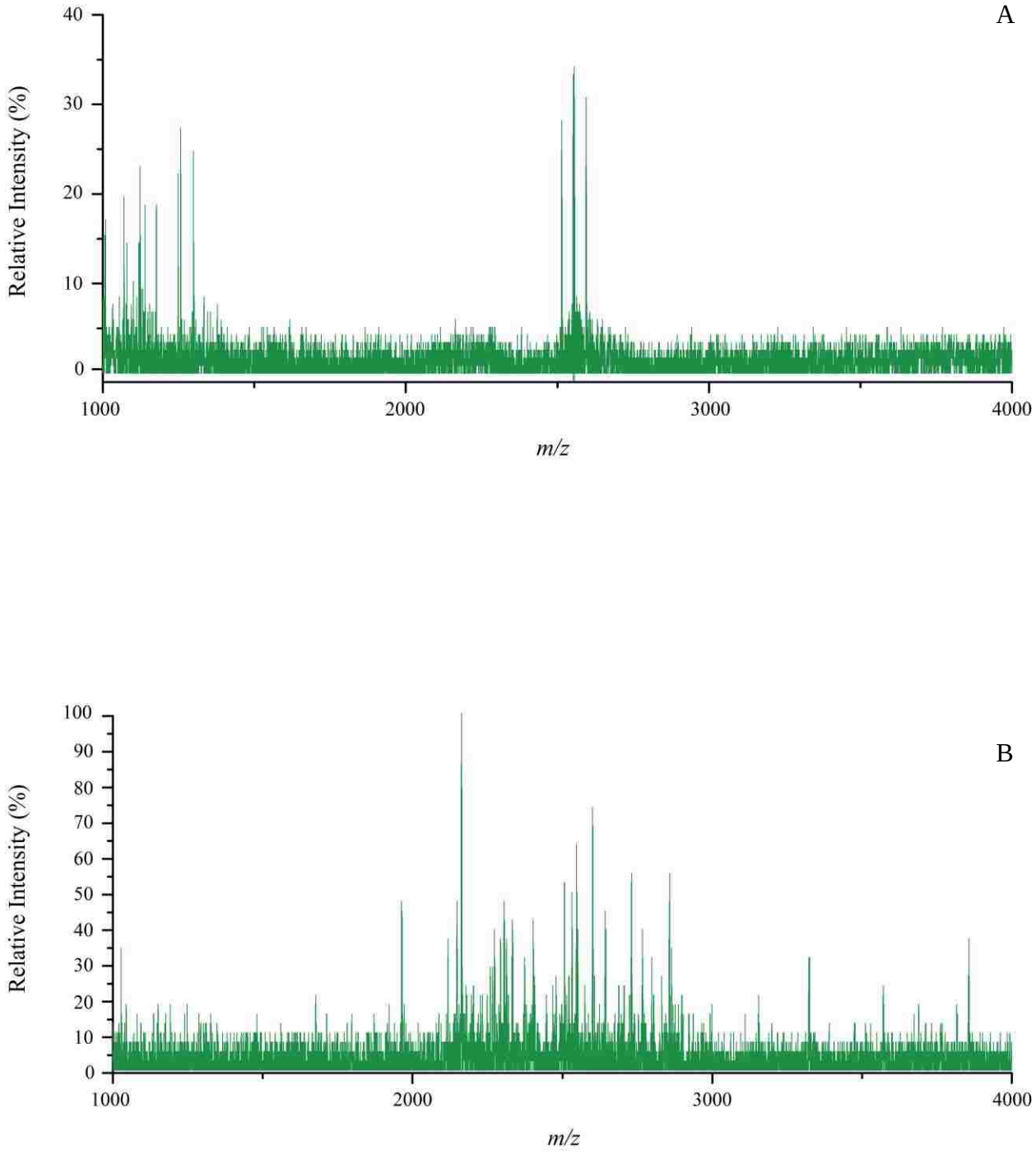


Figure 54. Expanded low mass region of oMALDI-QTOF-MS of *E. coli* from impacted *in situ* mini-wells proteolysis in FA (A) control (B) Trypsin

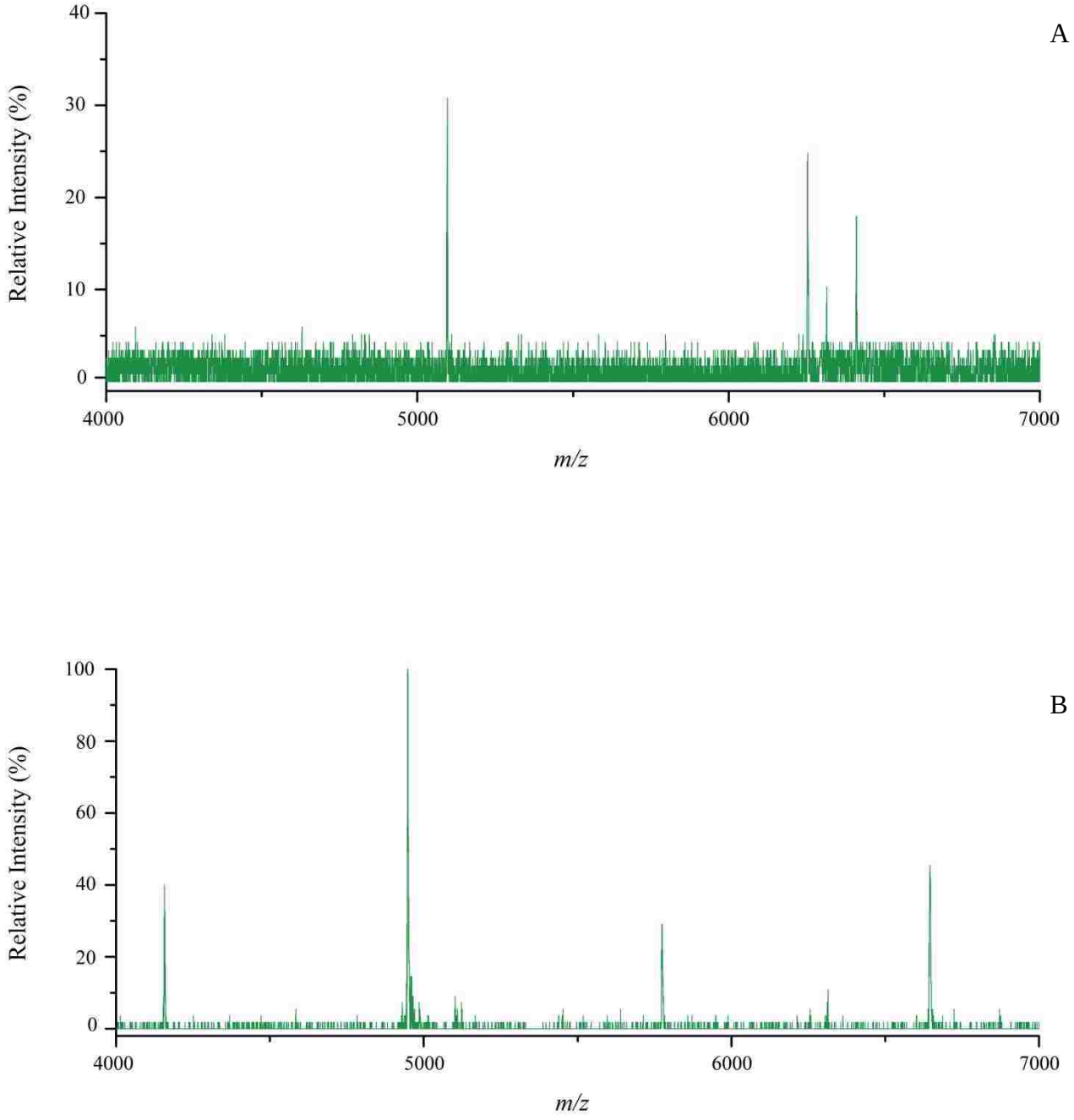


Figure 55. Expanded low mass region of oMALDI-QTOF-MS of *E. coli* from impacted *in situ* mini-wells proteolysis in FA (A) control (B) CNBr.

digested peptides is an indication the digestion did not have as much time for proteolysis as the *in vitro* digestions. This could be due to the effects of solvent evaporation or lack of agitation of the reaction vessel which is possible for *in vitro* reactions but not the mini-wells. However, when comparing the trypsin digested *E. coli* between the *in vitro* and the mini-wells, the mini-wells appear to have performed more efficiently as indicated by the greater number of peptide fragments observed. Table 22 lists the peptide fragments from the *in situ* trypsin proteolysis. The higher efficiency of the mini-wells compared to the *in vitro* proteolysis is a result of the *E. coli* cells settling in the reaction vessel (despite agitation). As the *E. coli* cells settle they become unavailable for proteolysis and this limits the access of trypsin to exposed proteins. For the mini-wells, settling is not as much an issue since the concentration of bacteria used is less than in the *in vitro* trials. In addition, the reaction volume is less inside the mini-wells while the surface area is equivalent in to the surface area at the bottom of the microcentrifuge tube.

For the CNBr digestions, the *in vitro* and *in situ* results for cytochrome *c* produced identical peaks albeit with different intensity ratios. This is likely a result of the decrease in reaction time due to solvent evaporation. Solvent evaporation is more critical for CNBr digestions compared to trypsin owing to the higher ratio of volatile solvents. Similarly, the *E. coli in vitro* digestion did produce more peaks with a lower *m/z* distribution compared to the mini-well digestions. Once more, this is a result of the reduced reaction time as a result of solvent evaporation. Moreover, since CNBr digestions are non-enzymatic, higher levels of CNBr were required as compared to trypsin to make certain detectable levels of fragments were produced.

For enzymatic digestions, the concentration of the protease is kept to a minimum in order to limit autolysis products. In order to increase fragment production, enzymatic reactions are

Table 22. Table of *in situ E. coli* peptide fragments.

Peak *m/z*	*In vitro*	*In situ*	SRS	Description
1043	–	X	–	
1079	–	X	–	
1597	–	X	–	
1657	–	X	–	
1675	–	X	–	
1728	–	X	–	
1804	–	X	–	
1961	X	X	–	
2116	X	X	–	
2146	–	X	–	
2162	X	X	X	Trypsin autolysis peak Somatotropin
2194	X	–	–	
2200	X	–	–	
2257	X	X	–	
2272	X	X	–	
2291	–	X	–	
2303	–	X	–	
2311	–	X	–	
2331	–	X	–	
2373	–	X	–	
2401	–	X	–	
2441	–	X	–	
2469	–	X	–	
2505	–	X	–	
2520	–	X	X	Putative Uncharacterized Protein
2544	X	X	–	
2551	X	–	–	
2599	X	X	–	
2640	–	X	–	
2671	–	X	–	
2686	–	X	–	
2728	–	X	–	
2765	–	X	–	
2797	–	X	–	
2855	X	X	–	
3323	–	X	–	
3599	X	–	–	
3669	X	–	–	
3858	X	X	–	

allowed to run longer. In the case of CNBr, this requires the addition of more proteolytic agent. While MALDI is tolerant of low levels of impurities, it is not unaffected by them.[118] Overall, both proteolytic techniques, trypsin and CNBr, proved successful with trypsin being more suitable for mini-well proteolysis and overall effectiveness with intact whole-cell bacteria.

5.3.4 Peptide Mass Mapping of Trypsin Digested Impacted Bioaerosols

Peptide mass maps of the *in situ* trypsin digested collected cytochrome *c* and *E. coli* bioaerosols were done. For cytochrome c, two peptide mass fingerprint searches were conducted using the MASCOT search engine. One search used the four *in silico* matched peptides found in Table 17 while the other used all five peptide masses from the mass spectrum as listed in Table 16. Only peptide masses, experimental mass minus prosthetic group (non-protein component) mass, e.g., heme, were used in the MASCOT searches since it is unable to account for this protein modification. The MASCOT results are shown in Figures 56 and 57 with reports for top hit protein matches shown in Tables 23 and 24. As indicated, the use of the fifth mass resulted in a decrease in MOWSE below the significance threshold value score of 67. Moreover, the top protein score resulted in non-cytochrome *c* protein. Despite this, the remaining protein matches, albeit insignificant, are a correct match. This demonstrates the importance of unmatched peptide fragments in the search algorithm. For purified proteins such as cytochrome *c*, this does not present a problem; however, for more complex samples such as intact whole-cell *E. coli*, a modified peptide mass mapping technique was needed to compensate for unmatched peptide fragments.

Peptide mass mapping of intact whole-cell *in situ* digested collected *E. coli* was done using the mass spectrum shown in Figure 53. From Figure 53, 34 tryptic fragments were observed, however, two peaks were found to have a hit in an SRS database search. Therefore,

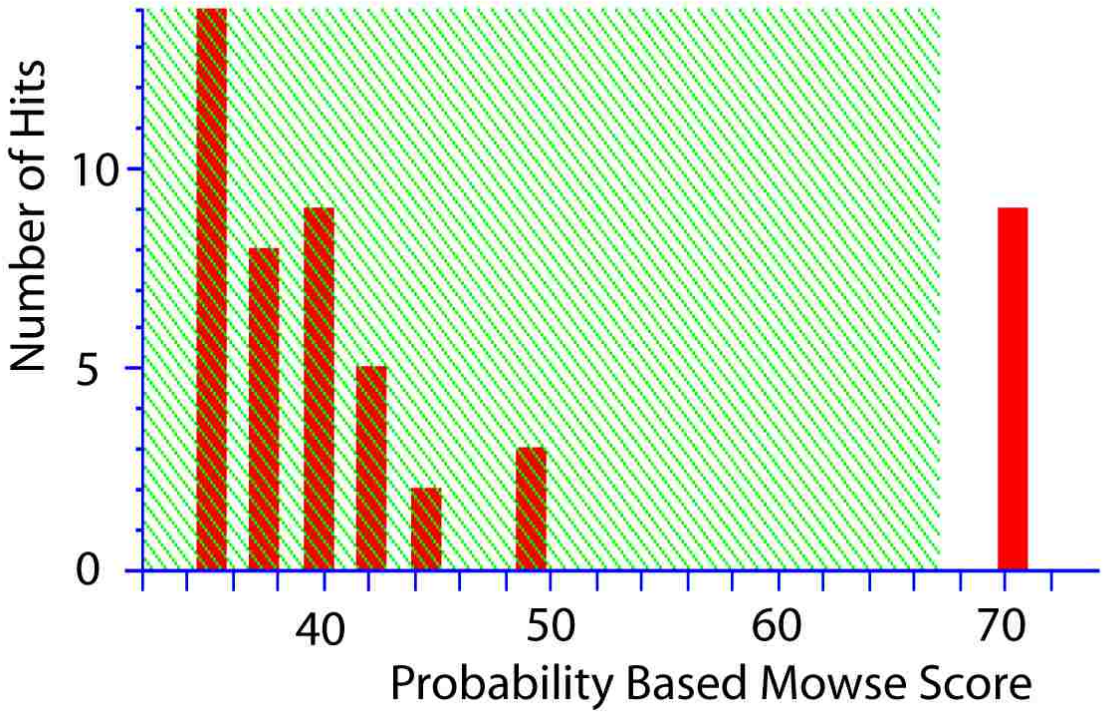


Figure 56. MASCOT graphical results for the 4 cross-matched peaks between the *in silico* and experimental peptide mass fragments.

Table 23. MASCOT results for the 4 cross-matched peaks between the *in silico* and experimental peptide mass fragments.

Match	Accession	Mass	Score	Description	Organism
1	CYC_BOVIN	11696	70	Cytochrome *c*	*Bos taurus*
2	CYC_CAMDR	11640	70	Cytochrome *c*	*Camelus dromedarius*
3	CYC_CANFA	11625	70	Cytochrome *c*	*Canis familiaris*
4	CYC_ESCGI	11640	70	Cytochrome *c*	*Eschrichtius gibbosus*
5	CYC_LAMGU	11640	70	Cytochrome *c*	*Lama guanicoe*
6	CYC_PIG	11696	70	Cytochrome *c*	*Sus scrofa*
7	CYC_SHEEP	11696	70	Cytochrome *c*	*Ovis aries*
8	CYC_HIPAM	11654	70	Cytochrome *c*	*Hippopotamus amphibius*
9	CYC_MIRLE	11610	70	Cytochrome *c*	*Mirounga leonina*
10	Y3487_RHOPA	21678	51	UPF0341 protein RPA3487	*Rhodopseudomonas palustris*

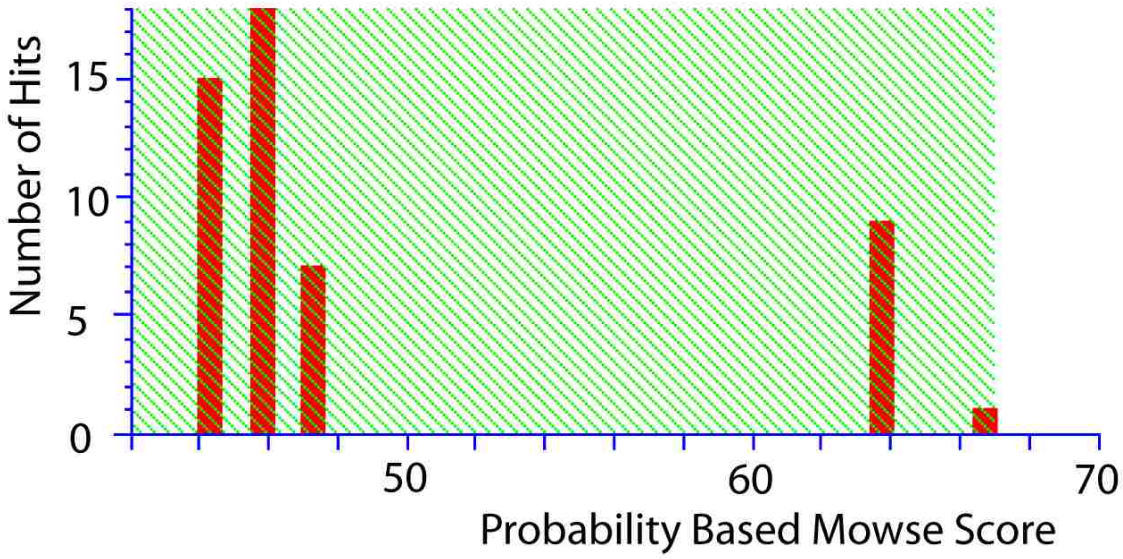


Figure 57. MASCOT graphical results for all 5 experimental peptide mass fragments.

Table 24. MASCOT results for all 5 experimental peptide mass fragments.

Match	Accession	Mass	Score	Description	Organism
1	RF3_HAEIN	59137	66	Peptide chain release factor 3	*Haemophilus influenzae*
2	CYC_BOVIN	11696	64	Cytochrome *c*	*Bos taurus*
3	CYC_CAMDR	11640	64	Cytochrome *c*	*Camelus dromedarius*
4	CYC_CANFA	11625	64	Cytochrome *c*	*Canis familiaris*
5	CYC_ESCGI	11640	64	Cytochrome *c*	*Eschrichtius gibbosus*
6	CYC_LAMGU	11640	64	Cytochrome *c*	*Lama guanicoe*
7	CYC_PIG	11696	64	Cytochrome *c*	*Sus scrofa*
8	CYC_SHEEP	11696	64	Cytochrome *c*	*Ovis aries*
9	CYC_HIPAM	11654	63	Cytochrome *c*	*Hippopotamus amphibius*
10	CYC_MIRLE	11610	63	Cytochrome *c*	*Mirounga leonina*

only the remaining 32 peaks were used in the MASCOT searches. The *in situ* digestion of the *E. coli* sample can produce fragments from any protein found inside the cell; it is an extremely complex protein mixture. Samples of this nature are typically separated by LC or CE prior to mass spectrometry analysis. However, since the sample being analyzed is on the surface of a MALDI target, this is not feasible. Therefore, any MASCOT search results in peptide fragments from different proteins. As shown with the cytochrome *c* search, this lowers the MOWSE score below the significance threshold. As a consequence, a different approach is needed.

For this work, a progressive reduction iterative search mapping (PRISM) technique was applied to the set of peak values. Since this is a known sample of *E. coli*, the searches were initially done using the organism specific SWISS-PROT database entries. In PRISM, all the unique masses from the *in situ* peaks from Table 22 were used in the initial search. After this initial mass mapping search, the peptide matches from the most significant protein hit is recorded and removed from the peak list for the next search. Afterwards, unmatched peaks from the initialsearch are searched again. The PRISM process is shown in Figure 58 and is repeated until

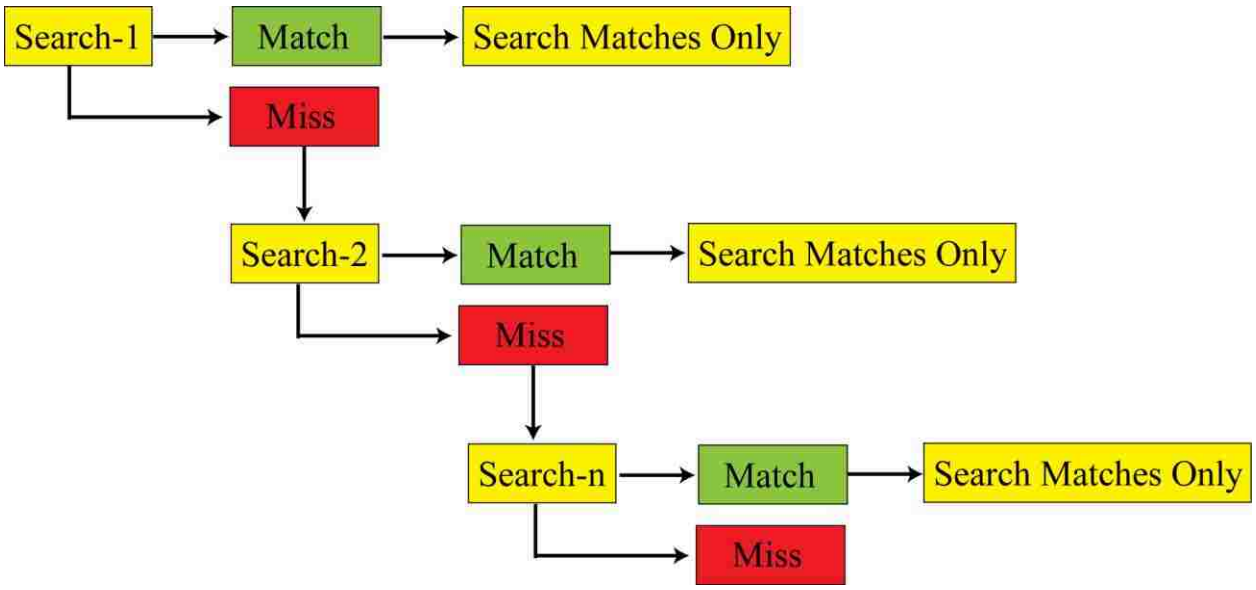


Figure 58. PRISM process for MASCOT searches of Bacteria.

all the peaks from Table 22 have been fit. Following the PRISM searches, the peptide matches from the most significant hit during each iteration were used for an individual search. These serves to increase the MOWSE score by removing unmatched peaks without modifying the significance threshold. During the individual searches, both the *E. coli* database subset was used as well as the entire database. Regardless of which dataset is used, the MOWSE score is unchanged. However, the significance threshold is increased when using the entire database, which could lead to an insignificant match.

For the PRISM MASCOT searches of the *in situ E. coli* digests, a total of seven iterations were necessary to exhaust all the masses used. Of the seven iterations, only four produced protein matches with significant MOWSE scores. After conducting all the searches with the *E. coli* subset of the database, the entire database was used. When utilizing the entire database after the first iteration, only three protein matches had MOWSE scores above the increased significance threshold values. Tables 25 – 28 and Figures 59 – 62 are the results of the searches done with the *E. coli* subset from the database. While the searches have identified four proteins from the *in situ* digest using the *E. coli* subset for the initial iteration, a similar result was not possible when starting with the entire database for iteration number 1. While analyzing unknown samples for the purpose of identification remains difficult, *in situ* proteolysis of collected bioaerosols provides valuable information when coupled to the SRS searches for unproteolyzed samples.

5.3.5 MS/MS of Proteolytic Fragments from *In situ* Proteolysis of Impacted Bioaerosols

While protein database searching and peptide mass mapping are powerful tools, the use of other mass spectrometric techniques can add to the information obtained from collected bioaerosols. Since the oMALDI-QTOF-MS is a tandem instrument capable of performing

Table 25. Iteration 1 from PRISM of trypsin proteolysis fragments from *E. coli*.

Description	MOWSE	Significance P (<0.5)	Matched	Sequence Coverage	Expect
Glutamate 5-kinase	107	52	8	35%	$1.8x10^{-7}$

MSDSQTLVVK LGTSVLTGGS RRLNRAHIVE LVR**QCAQLHA AGHRIVIVTS**
GAIAAGREHL GYPELPATIA SKQLLAAVGQ SRLIQLWEQL FSIYGIHVGQ
MLLTRADMED RERFLNARDT LRALLDNNIV PVINENDAVA TAEIKVGDND
NLSALAAILA GADK**LLLLTD QKGLYTADPR SNPQAELIKD VYGIDDALR**A
IAGDSVSGLG TGGMSTKLQA ADVACRAGID TIIAAGSKPG VIGDVMEGIS
VGTLFHAQAT PLENRK**RWIF GAPPAGEITV DEGATAAILE RGSSLLPK**GI
KSVTGNFSRG EVIRICNLEG R**DIAHGVSRY NSDALR**RIAG HHSQEIDAIL
GYEYGPVAVH RDDMITR

Peak *m/z*	Start-Stop	Missed Cleavage
2401	34 – 57	1
3858	45 – 82	2
1043	73 – 82	0
2797	165 – 189	2
2116	181 – 199	1
2640	267 – 291	1
3323	267 – 298	2
1675	322 – 336	1

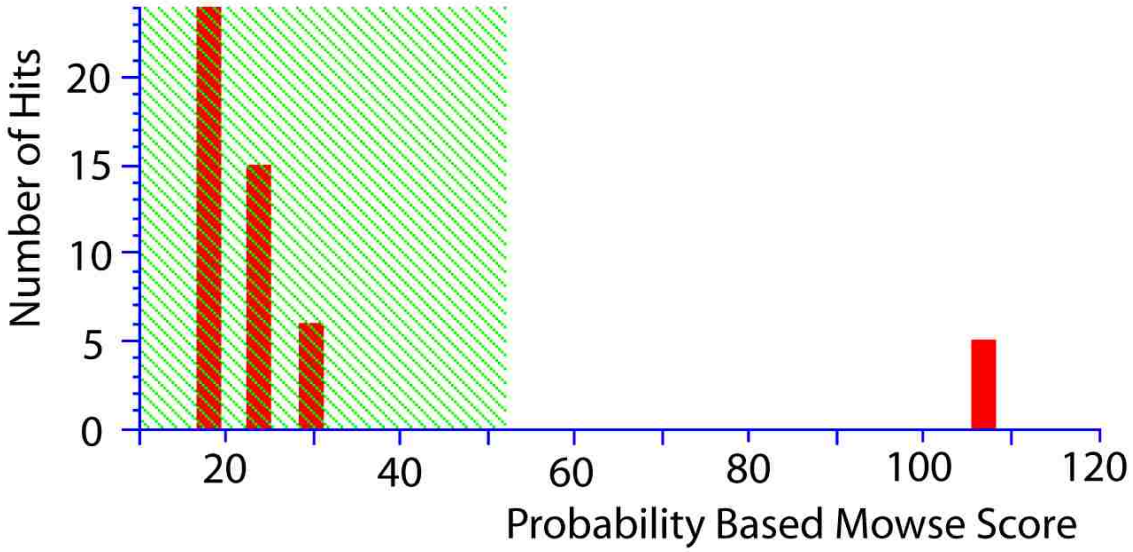


Figure 59. MASCOT search results of Iteration 1 from *E. coli* proteolytic peptides.

Table 26. Iteration 2 from PRISM of trypsin proteolysis fragments from *E. coli*.

Description	MOWSE	Significance P (<0.5)	Matched	Sequence Coverage	Expect
Uncharacterized ferrodoxin-like protein ydhX	73	67	5	18%	0.00045

MSWIGWTVAA TALGDNQMSF TRRKFVLGMG TVIFFTGSAS SLLANTRQEK
EVRYAMIHDE SRCNGCNICA RACRKTNHAP AQGSRLSIAH IPVTDNDNET
QYHFFRQSCQ HCEDAPCIDV CPTGASWRDE QGIVRVEKSQ CIGCSYCIGA
CPYQVRYLNP VTKVADKCDF CAESRLAKGF PPICVSACPE HALIFGREDS
PEIQAWLQDN KYYQYQLPGA GKPHLYRRFG QHLIKKENV

Peak *m/z*	Start-Stop	Missed Cleavage
2599	1 – 23	1
2728	1 – 24	2
2311	2 – 22	0
2469	2 – 23	1
2441	51 – 71	2

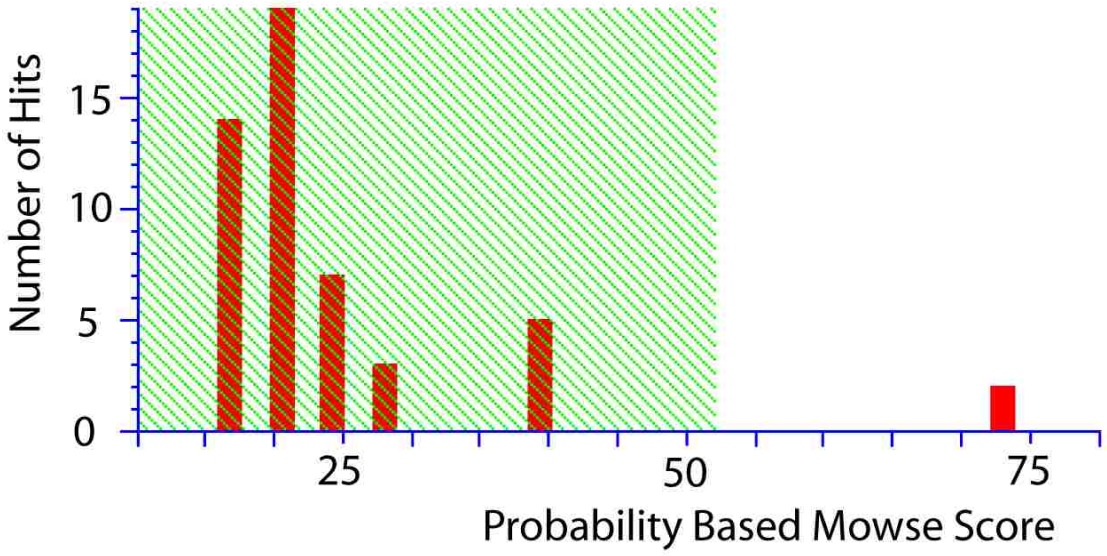


Figure 60. MASCOT search results of Iteration 2 from *E. coli* proteolytic peptides.

Table 27. Iteration 3 from PRISM of trypsin proteolysis fragments from *E. coli*.

Description	MOWSE	Significance P (<0.5)	Matched	Sequence Coverage	Expect
Crotonobetainyl-CoA carnitine CoA-transferase	70	52	6	20 %	0.00082

MDHLPMPKFG PLAGLRVVFS GIEIAGPFAG QMFAEWGAEV IWIENVAWAD
TIRVQPNYPQ LSRR**NLHALS LNIFKDEGRE AFLKLMETTD IFIEASKGPA**
FARRGITDEV LWQHNPKLVI AHLSGFGQYG TEEYTNLPAY NTIAQAFSGY
LIQNGDVDQP MPAFPYTADY FSGLTATTAA LAALHKVRET GKGESIDIAM
YEVMLRMGQY FMMDYFNGGE MCPRMTKGKD PYYAGCGLYK **CADGYIVMEL**
VGITQIAECF KDIGLAHLLG TPEIPEGTQL IHRIECPYGP LVEEKLDAWL
AAHTIAEVKE R**FAELNIACA K**VLTVPELES NPQYVARESI TQWQTMDGRT
CKGPNIMPKF KNNPGQIWRG MPSHGMDTAA ILKNIGYSEN DIQELVSKGL
401 AKVED

Peak *m/z*	Start-Stop	Missed Cleavage
1728	6 – 79	1
2544	76 – 97	2
2686	80 – 103	2
2291	98 – 117	2
2303	241 – 261	0
1079	312 – 321	0

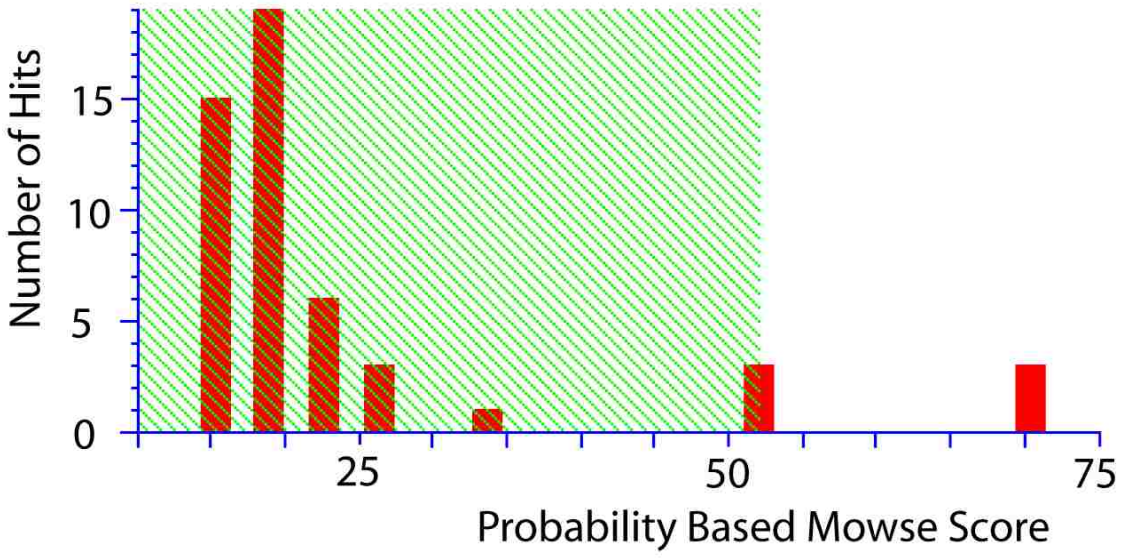


Figure 61. MASCOT search results of Iteration 3 from *E. coli* proteolytic peptides.

Table 28. Iteration 4 from PRISM of trypsin proteolysis fragments from *E. coli*.

Description	MOWSE	Significance P (<0.5)	Matched	Sequence Coverage	Expect
Lysine-arginine-ornithine-binding periplasmic protein precursor	63	52	5	28 %	0.0045

MKKSILALSL LVGLSTAASS YAALPETVRI GTDTTYAPFS SKDAKGDFVG
FDIDLGNEMC KRMQVKCTWV ASDFDALIPS LKAKKIDAII SSLSITDKRQ
QEIAFSDKLY AADSRLIAAK GSPIQPTLDS LK**GKHVGVLQ GSTQEAYANE**
TWRSKGVDVV AYANQDLVYS DLAAGR**LDAA LQDEVAASEG FLKQPAGK**DF
AFAGSSVKDK KYFGDGTGVG LR**KDDAELTA AFNKALGELR** QDGTYDK**MAK**
KYFDFNVYGD

Peak *m/z*	Start-Stop	Missed Cleavage
2331	133 – 153	1
2146	135 – 153	0
2258	177 – 198	1
1961	223 – 240	2
1597	248 – 260	2

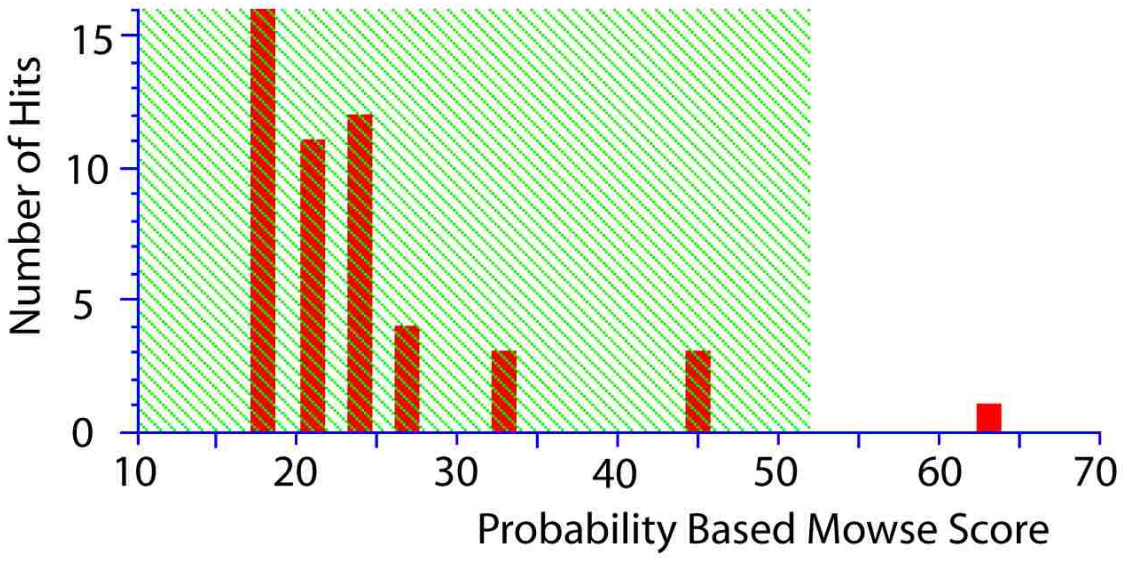


Figure 62. MASCOT search results of Iteration 4 from *E. coli* proteolytic peptides.

MS/MS experiments on *m/z* values less than 6 kDa, MS/MS experiments were done on selected masses from the *in situ* digestion of the collected bioaerosols. Figures 63 and 64 are the MS/MS spectra for the peak 1168 from the cytochrome *c in situ* trypsin digestion. Peak 1168 contains residues 29 – 39 with no missed cleavages. For this peptide, an incomplete fragment sequence is observed which is typical for low energy CID. For peak 1168 fragmentation, a b-series was seen from $b_3 – b_8$ and a y-series from $y_1 – y_6$. All other peptide backbone cleavage ion types were observed with a-type ions being the most abundant. In addition, internal fragmentation ions were present. For the internal fragments, nearly 50 % were found with an a-type cleavage resulting in *m/z* values decreased by 28 Da which corresponds to the loss of CO from the C-terminus side of the peptide. Overall, all amino acid residues for this peptide could be determined. However, the presence of the internal fragment ions with mixed cleavage types prevented the instrument software from determining the peptide's primary sequence automatically. For the CNBr *in situ* digest, peak 2735 was selected (Figure 65 and 66). In this MS/MS spectrum, eight b-series ions and five y-series ions were observed. However, the longest sequence stretch was four residues of the a-series. This is consistent with the internal fragmentation pattern resulting in internal fragments ions displaying a loss of 28 Da which is indicative of a-type cleavage corresponding to a loss of CO. For peak 2735, 33 % of the internal fragments ions displayed the loss of 28 Da.

The same MS/MS experiments were done for the *E. coli in situ* digested samples. However, *E. coli* digested samples did not produce MS/MS spectra of proteolytic peptide fragments. This could be due to the complexity of the *E. coli* sample as well as the overall intensity and ion abundances of each analyte during ionization and mass selection by Q_1. For *E. coli* samples, the total ion count seen during acquisition was substantially higher than indicated by the observed mass spectra. The total ion count could be responsible for the decrease in ion

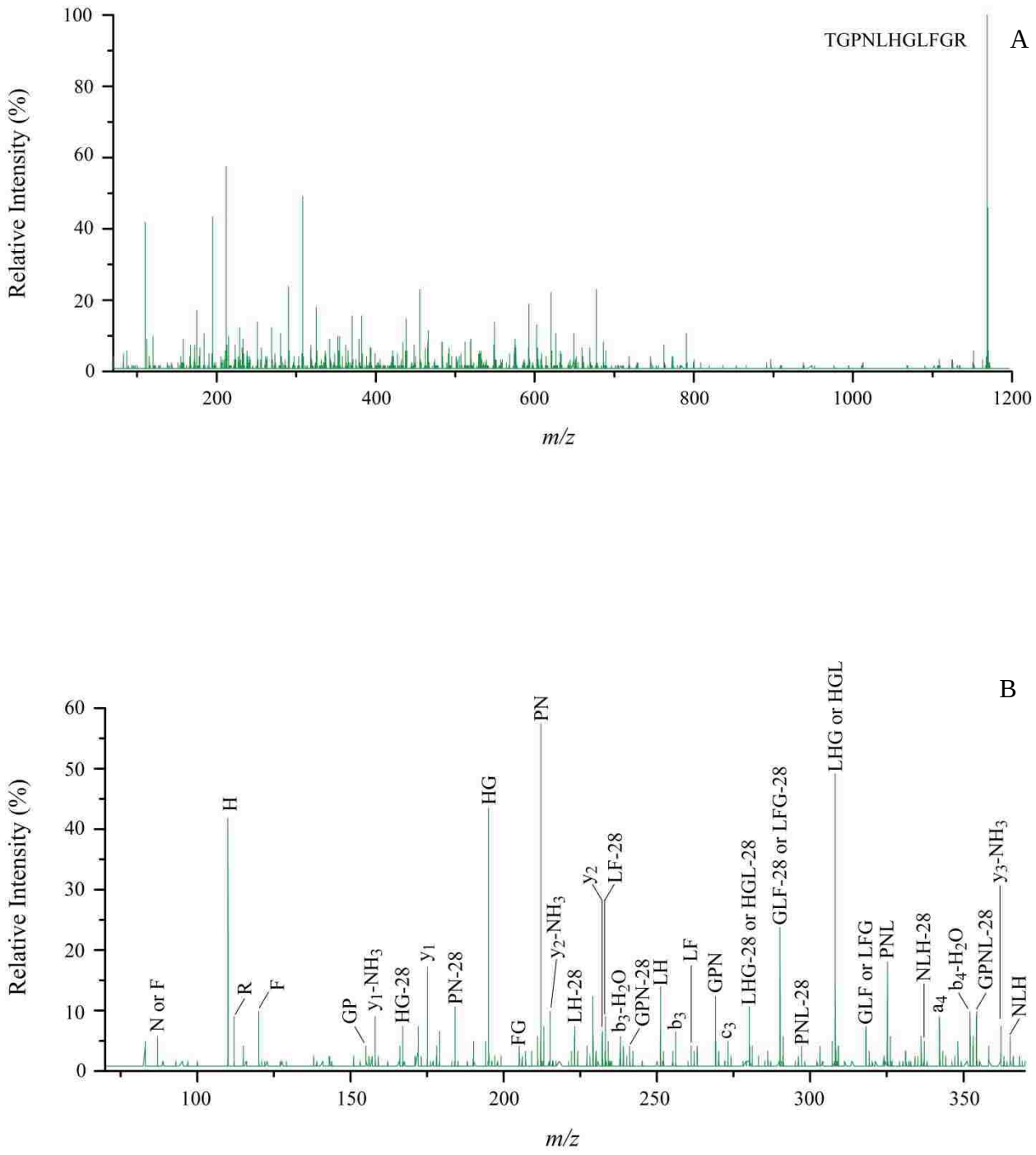


Figure 63. MS/MS experiments of *in situ* trypsin digested cytochrome *c* peak 1168 (A) full *m/z* range and (B) expanded range from 70 to 370 *m/z*.

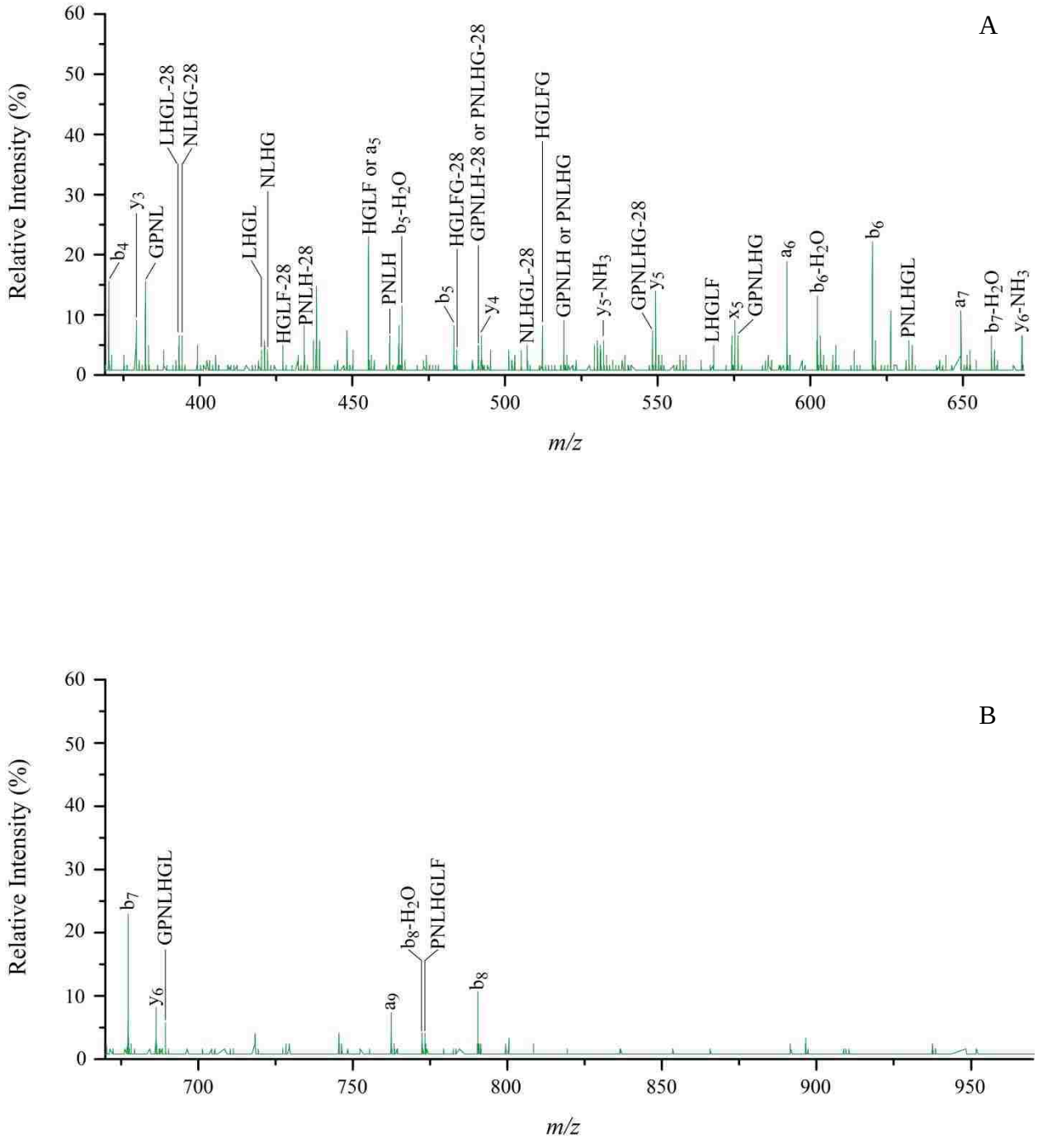


Figure 64. MS/MS experiments of *in situ* trypsin digested cytochrome *c* peak 1168 (A) expanded range from 370 to 670 *m/z* and (B) expanded range from 670 to 970 *m/z*.

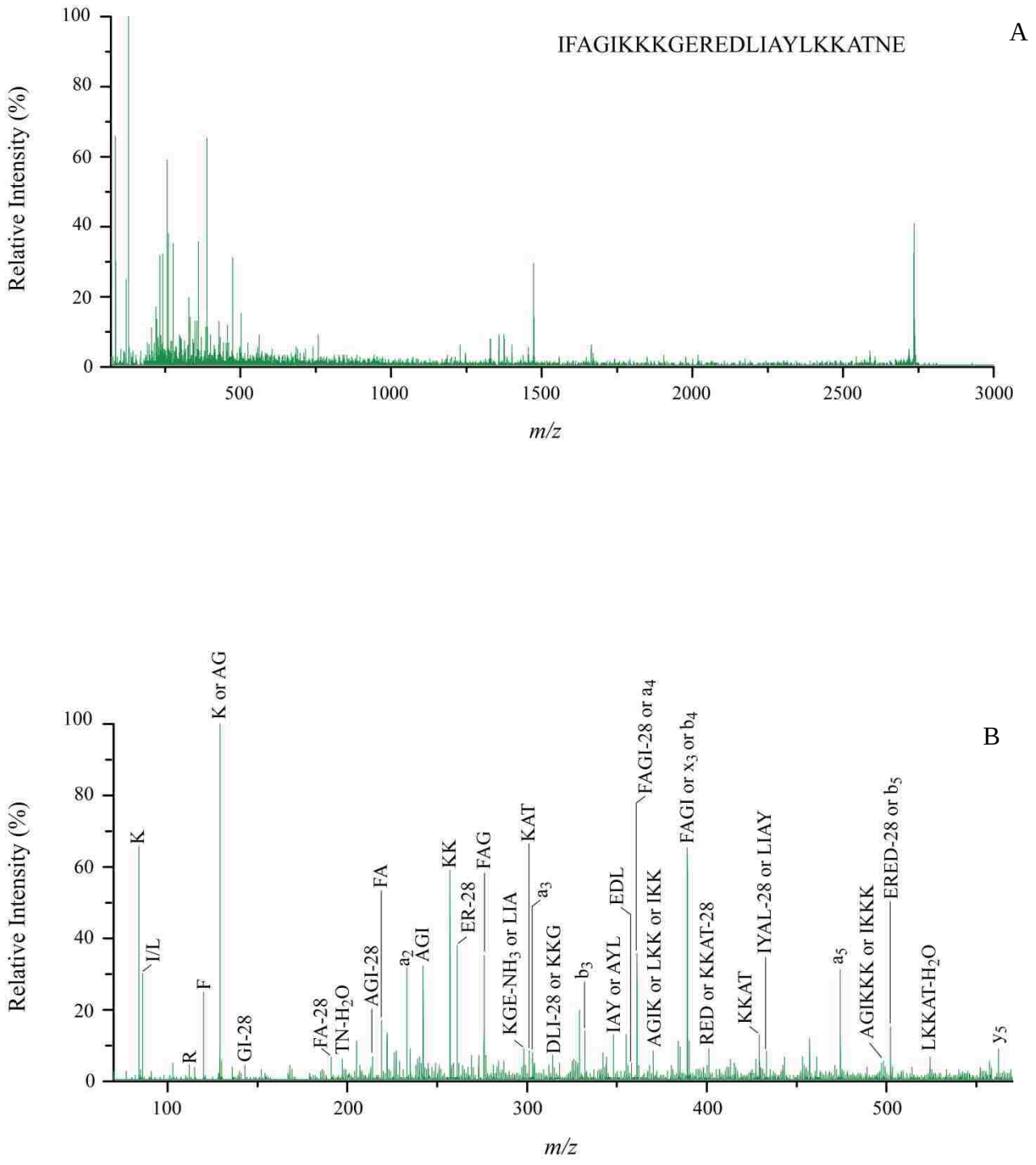


Figure 65. MS/MS experiments of *in situ* CNBr digested cytochrome *c* peak 2735 (A) full *m/z* range and (B) expanded range from 70 to 570 *m/z*.

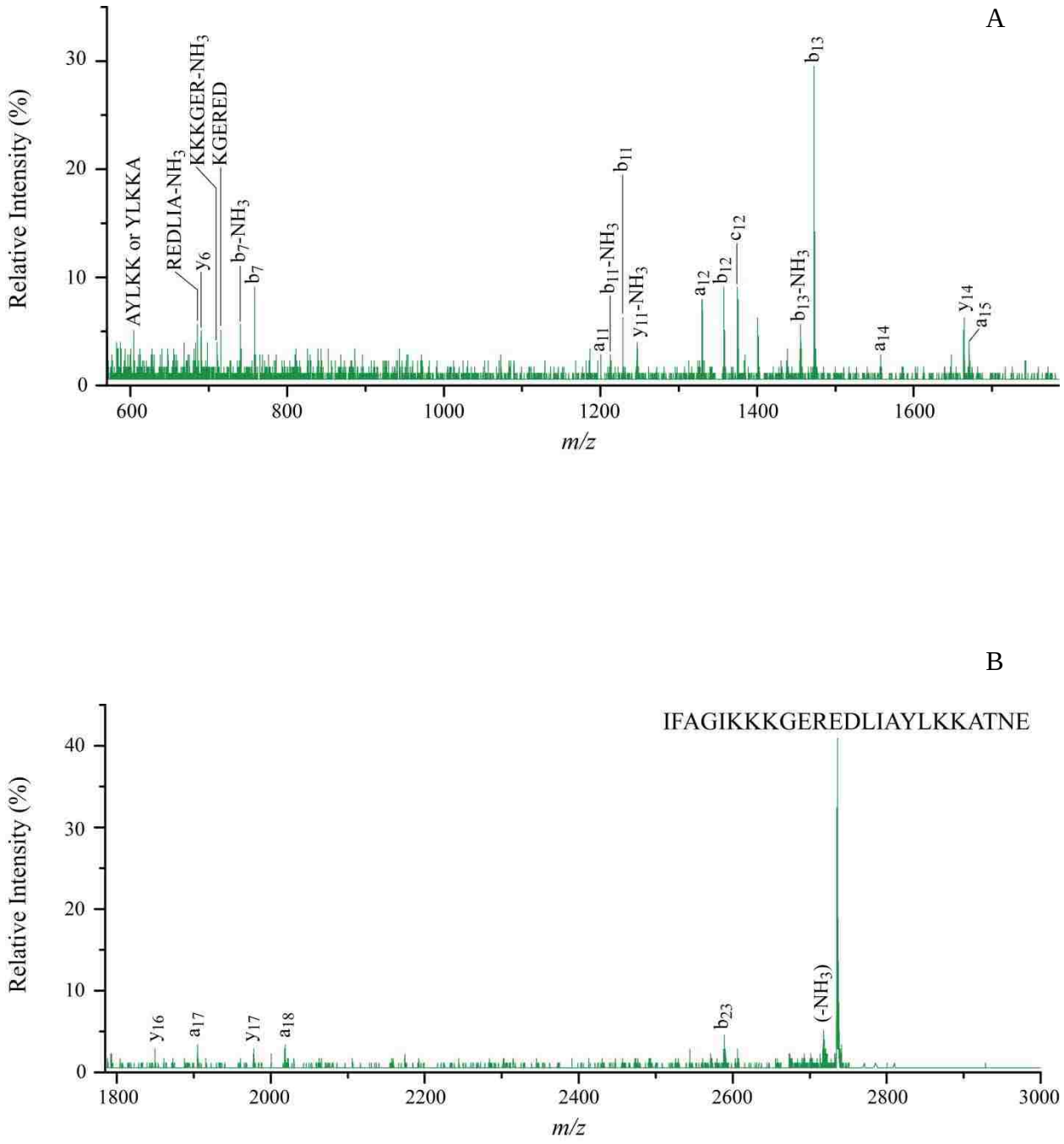


Figure 66. MS/MS experiments of *in situ* CNBr digested cytochrome *c* peak 2735 (A) expanded range from 570 to 1785 *m/z* and (B) expanded range from 1785 to 3000 *m/z*.

transmission of the selected mass through Q_1 for fragmentation. As a result, the precursor mass was not seen and therefore no fragmentation was possible.

5.4 Summary

The proteolysis of proteins and intact whole-cell bacteria is possible using both trypsin and CNBr as the proteolytic agents. For cytochrome c, there was no difference between the *in vitro* and *in situ* mini-well digestions. However, for *E. coli*, the mini-well digestion of the impacted bioaerosol produced more fragments than the *in vitro* proteolysis. This is not true for the CNBr proteolysis. For *E. coli*, the CNBr digestion produced more peptide fragments for the *in vitro* proteolysis compared to that of the mini-wells. Despite this, CNBr digests inside the mini-wells is possible. Overall, the *in situ* proteolysis through the use of the mini-wells can be performed in just under 2 hours from the time of collection. Once data is obtained for the proteolyzed collected bioaerosols, the information can be used conduct peptide mass mapping searches using the MASCOT search engine. For the case of complex samples done *in situ*, such as *E. coli*, a modified search technique called PRISM was used to identify up to four proteins.

CHAPTER 6. CONCLUSIONS AND FUTURE DIRECTIONS

In this dissertation, matrix-assisted laser desorption/ionization mass spectrometry was described in which collected bioaerosols were analyzed for the detection of polypeptides and identification through online database searching. Intact whole-cell bacteria were analyzed for the first time using an oMALDI-QTOF mass spectrometer. A MALDI-TOF mass spectrometer was also used for comparison. Three bioaerosol samplers, an inertial impactor, a cyclone impactor and a vacuum filter system, were used to determine their MALDI compatibility. The significance of this work lies in the ability of MALDI to serve as a fast and reliable analytical technique for identifying microorganisms.

In Chapter 3, an orthogonal MALDI source was used on a tandem quadrupole-time-of-flight mass spectrometer and compared to an axial MALDI time of flight mass spectrometer. While the overall signal is less on the QTOF, the mass resolution was considerably greater providing improved mass accuracy for protein searching in online protein databases. Despite the upper mass limit of 40,000 *m/z* for the QTOF, the instrument was capable of analyzing the complex samples of intact whole cell bacteria samples of *E. coli*, *B subtilis*, and *B. thuringiensis*.

An evaluation of bioaerosol samplers was done in order to determine the compatibility of those samplers for use with MALDI-MS. MALDI has proven to be a fast and reliable analytical method for analyzing intact whole-cell bacteria. Bacteria analysis by MALDI can characterize bacteria down to the strain level.[119] However, most of these analyses are done from cultured cells. While it is possible to culture cells from a collected bioaerosol using the techniques described herein, analyzing collected cells without culturing provides the fastest analysis of any offline mass spectrometric method. To evaluate possible collection methods for use in an offline

bioaerosol identification technique, three bioaerosol samplers were used and assessed for MALDI compatibility in terms of speed, quality and reproducibility.

Of the three biosamplers studied, the two impactors, the Andersen and the cyclone impactor proved to be provide the fastest analysis times and best spectra quality as indicated by the lack of sample background noise. Between these two bioaerosol samplers, the Andersen impactor is superior to the cyclone sampler in terms of post collection sample handling in that there is no additional concentration or isolation step required. In addition, the cyclone sampler is constructed from glass and is susceptible to breakage as well as requires additional reagents for operation whereas the Andersen impactor does not. However, when post collection sample processing or storage is required, the CI is ideal due to its ability to use of a liquid collection medium such as water allowing for rapid storage through deep freezing.

While the AI and CI MALDI-MS analyzed samplers are similar, they differ in their ability to analyze collected proteins such as bacteriotoxins. In this case, the AI has an advantage over the CI due to difficulties in concentrating proteins post collection. As observed for the cytochrome *c* collection studies, the CI requires considerable time to concentrate had the sample not deposited on the surface vessel wall. In fact, removing approximately 20 mL of water would have exceeded 24 hours. Despite this, the cyclone impactor can be used with some modified post collection strategies such as protein precipitation or even capillary electrophoresis separation of the collected liquid. For the later, lab on a chip technology could be used, which opens an entirely new set of processing options such as on chip cell lysing and proteolytic digestion. An even simpler approach is to use a more volatile solvent to facilitate evaporation. Regardless, the overall time for analysis is shortest for analysis of AI bioaerosol collections.

The non-inertial sampler used in this study was a vacuum filter system. In this system, any filter paper can potentially be used depending on need. Unfortunately, the results of this study indicate that direct analysis, while possible, does not produce the best results. The best results for the FS were when the collected samples were removed from the filter by reverse deposition on a wet, dry or adhesive coated surface. Once the sample was removed, the analysis is identical to that of dried droplet processing and timescale. Since integration of the sample within the MALDI matrix crystal is vital for analysis, steps could be taken to assist in this process. For instance, filter paper coated with excess matrix can be used in which the post collection processing would include the addition of solvent. The application of solvent to matrix coated filter paper could improve the analyte integration in to the MALDI matrix crystal and enhance direct analysis. However, the probable best course would be to extract the sample from the filter paper through addition of solvent and agitation. More to the point, the analysis of collected bioaerosols by filtration merits further study.

Since the AI post collection times are shorter than all other collection techniques, some proteomic practices were investigated. In Chapter 5, collected bioaerosols were proteolyzed *in situ* using a removable mini-well. The mini-well serves as a temporary container allowing the addition of proteolytic reagents. Previously, immobilized trypsin was used to digest samples on a MALDI target.[120, 121] In the earliest work, the bacteria were lysed prior to being deposited onto an immobilized trypsin–coated MALDI target.[87] More recently, intact whole cell bacteria in a methanol and water solvent were deposited and digested using immobilized trypsin on a MALDI target. While these techniques were successful, none were done on collected samples and the intactness of the bacteria in the later work could be questionable due to the methanol/water solvent used to create the bacteria suspension. The addition of solvents other than water to the

bacteria suspension could alter the distribution of proteins due to extraction, or in the case of gram negative bacteria, could even cause the outer membrane to become more porous resulting in the release of periplasmic proteins. In this work, in order to ensure cells remained intact before proteolysis, only water was used as the solvent for bacteria suspensions.

The *in situ* digestions using the mini-wells provided the best results for trypsin proteolysis. While CNBr did work, the results were not as good as the *in vitro* results. This is most probably the result bacteria exposure surface and the time allowed for proteolysis. In the case of trypsin, the reaction was completed in 2 hours and the CNBr reactions were done in less than 30 minutes. From the results of the proteolysis and intact whole-cell mass spectra, peptide mass mapping and protein database searching was done. For peptide mass mapping, the PRISM technique does demonstrate promise as shown in the identification of proteins from the *in situ* proteolyzed *E. coli*. As for the protein database searching, the results obtained could be filtered to determine which matches produce a more significant overall contribution while removing those which are not significant. While this was manually accomplished in this work, future developments of search and filtering algorithms could reduce the tedious nature of this analysis. Despite this, the successful collection and analysis of bioaerosols including proteins and bacteria was accomplished. Moreover, *in situ* proteolysis of impacted bioaerosols was fast and easy leading the way for the use of this technique in future studies using different proteolytic agents.

REFERENCES

1. Hinds, W. C., *Aerosol Technology: Properties, Behavior and Measurement of Airborne Particles*. 2nd ed.; Wiley Interscience: New York, 1999.

2. Heyder, J., Deposition of Inhaled Particles in the Human Respiratory Tract and Consequences for Regional Targeting in Respiratory Drug Delivery. *Proc Am Thorac Soc* 2004, 1, 315-320.

3. Douwes, J.; Thorne, P.; Pearce, N.; Heederik, D., Bioaerosol health effects and exposure assessment: progress and prospects. *Ann Occup Hyg* 2003, 47, (3), 187-200.

4. Cox, C. S., Stability of Airborne Microbes and Allergens. In *Bioaerosols Handbook*, Cox, C. S.; Wathes, C. M., Eds. Lewis Publishers: 1995; pp 77-99.

5. Agranovski, I. E.; Safatov, A. S.; Borodulin, A. I.; Pyankov, O. V.; Petrishchenko, V. A.; Sergeev, A. N.; Agafonov, A. P.; Ignatiev, G. M.; Sergeev, A. A.; Agranovski, V., Inactivation of viruses in bubbling processes utilized for personal bioaerosol monitoring. *Appl Environ Microbiol* 2004, 70, (12), 6963-7.

6. Riedel, S., Biological warfare and bioterrorism: a historical review *Baylor University Medical Center Proceedings* 2004, 17, (4), 400-406.

7. Herr, C. E.; Nieden Az, A.; Stilianakis, N. I.; Eikmann, T. F., Health effects associated with exposure to residential organic dust. *Am J Ind Med* 2004, 46, (4), 381-5.

8. Burge, H. A., Bioaerosols in the Residential Environment. In *Bioaerosols Handbook*, Cox, C. S.; Wathes, C. M., Eds. Lewis Publishers: 1995; pp 579-597.

9. Van Weemen, B. K.; Schuurs, A. H., Immunoassay using antigen-enzyme conjugates. *FEBS Lett* 1971, 15, (3), 232-236.

10. Janda, J. M.; Abbott, S. L., 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 2007, 45, (9), 2761-4.

11. Lin, C. L.; Chiu, C. H.; Chu, C.; Huang, Y. C.; Lin, T. Y.; Ou, J. T., A multiplex polymerase chain reaction method for rapid identification of Citrobacter freundii and Salmonella species, including Salmonella Typhi. *J Microbiol Immunol Infect* 2007, 40, (3), 222-6.

12. Widjojoatmodjo, M. N.; Fluit, A. C.; Verhoef, J., Molecular identification of bacteria by fluorescence-based PCR-single-strand conformation polymorphism analysis of the 16S rRNA gene. *J Clin Microbiol* 1995, 33, (10), 2601-6.

13. Jensen, M. A.; Webster, J. A.; Straus, N., Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl Environ Microbiol* 1993, 59, (4), 945-52.

14. Imaoka, K.; Kimura, M.; Suzuki, M.; Kamiyama, T.; Yamada, A., Simultaneous detection of the genus Brucella by combinatorial PCR. *Jpn J Infect Dis* 2007, 60, (2-3), 137-9.

15. Burton, J. E.; Oshota, O. J.; Silman, N. J., Differential identification of Bacillus anthracis from environmental Bacillus species using microarray analysis. *J Appl Microbiol* 2006, 101, (4), 754-63.

16. Hong, G. E.; Kim, D. G.; Bae, J. Y.; Ahn, S. H.; Bai, S. C.; Kong, I. S., Species-specific PCR detection of the fish pathogen, Vibrio anguillarum, using the amiB gene, which encodes N-acetylmuramoyl-L-alanine amidase. *FEMS Microbiol Lett* 2007, 269, (2), 201-6.

17. Hayashi, Y.; Matsuda, R.; Maitani, T.; Imai, K.; Nishimura, W.; Ito, K.; Maeda, M., Precision Limit of Detection and Range of Quantitation in Competitive ELISA. *Anal Chem* 2004, 76, (5), 1295-1301.

18. Hernlem, B. J.; Ravva, S. V., Application of flow cytometry and cell sorting to the bacterial analysis of environmental aerosol samples. *J Environ Monit* 2007, 9, (12), 1317-22.

19. Huntress, E.; Stanley, L.; Parker, A., The Preparation of 3-Aminophthalhydrazide for Use in the Demonstration of Chemiluminescence *J A. Chem Soc* 1934, 56, (1), 241-242.

20. Ewetz, L.; Thore, A., Correlation Between Hemin Content and the Chemiluminescent Luminol Reaction with Bacteria. *Appl Environ Microbiol* 1978, 36, (6), 790-793.

21. Levin, J.; Bang, F. B., The Role of Endotoxin in the Extracellular Coagulation of Limulus Blood. *Bull Johns Hopkins Hosp* 1964, 115, 265-74.

22. Prior, R. B.; Spagna, V. A., Application of Limulus Test Device in Rapid Evaluation of Gonococcal and Nongonococcal Urethritis in Males. *J Clin Microbiol* 1981, 14, (3), 256-260.

23. Zhang, G.; Baek, L.; Koch, C., New Microarray for Quantitation of Endotoxin Using Limulus Amebocyte Lysate Combined with Enzyme-Linked Immunosorbent Assay. *J Clin Microbiol* 1988, 26, (8), 1464-1470.

24. Novitsky, T. J.; Roslansky, P. F.; Siber, G. R.; Warren, H. S., Turbidimetric Method for Quantifying Serum Inhibition of Limulus Amoebocyte Lysate. *J Clin Microbiol* 1985, 20, (2), 211-216.

25. Essendoubi, M.; Toubas, D.; Bouzaggou, M.; Pinon, J. M.; Manfait, M.; Sockalingum, G. D., Rapid identification of Candida species by FT-IR microspectroscopy. *Biochim Biophys Acta* 2005, 1724, (3), 239-47.

26. Al-Qadiri, H. M.; Al-Holy, M. A.; Lin, M.; Alami, N. I.; Cavinato, A. G.; Rasco, B. A., Rapid detection and identification of Pseudomonas aeruginosa and Escherichia coli as pure and mixed cultures in bottled drinking water using fourier transform infrared spectroscopy and multivariate analysis. *J Agric Food Chem* 2006, 54, (16), 5749-54.

27. Rebuffo-Scheer, C. A.; Kirschner, C.; Staemmler, M.; Naumann, D., Rapid species and strain differentiation of non-tubercoulous mycobacteria by Fourier-Transform Infrared microspectroscopy. *J Microbiol Methods* 2007, 68, (2), 282-90.

28. Naumann, D.; Keller, S.; Helm, D.; Schultz, C.; Schrader, B., FT-IR Spectroscopy and FT-Raman Spectroscopy Are Powerful Analytical Tools for the Non-invasive Characterization of Intact Microbial Cells. *J Mol Struct* 1995, 347, 399-406.

29. Rosch, P.; Harz, M.; Schmitt, M.; Peschke, K. D.; Ronneberger, O.; Burkhardt, H.; Motzkus, H. W.; Lankers, M.; Hofer, S.; Thiele, H.; Popp, J., Chemotaxonomic identification of single bacteria by micro-Raman spectroscopy: application to clean-room-relevant biological contaminations. *Appl Environ Microbiol* 2005, 71, (3), 1626-37.

30. Manoharan, R.; Ghiamati, E.; Chadha, S.; Nelson, W.; Sperry, J., Effect of Cultural Conditions on Deep UV Resonance Raman Spectra of Bacteria. *Appl Spectrosc* 1993, 47, 2145-2150.

31. Ghiamati, E.; Manoharan, R.; Nelson, W.; Sprerry, J., UV Resonance Raman Spectra of *Bacillus* Spores. *Appl Spectrosc* 1992, 46, 357-364.

32. Goodacre, R.; Shann, B.; Gilbert, R. J.; Timmins, E. M.; McGovern, A. C.; Alsberg, B. K.; Kell, D. B.; Logan, N. A., Detection of the dipicolinic acid biomarker in Bacillus spores using Curie-point pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. *Anal Chem* 2000, 72, (1), 119-27.

33. Arnold, R. J.; Karty, J. A.; Reilly, J. P., Bacterial Strain Differentiation by Mass Spectrometry. In *Identification of Microorganisms by Mass Spectrometry*, Wilkins, C.; Lay, J. O., Jr., Eds. John Wiley & Sons, Inc.: 2006; pp 181-201.

34. Snyder, A. P.; Dworzanski, J. P.; Tripatha, A.; Maswadeh, W. M.; Wick, C. H., Correlation of Mass Spectrometry Identified Bacterial Biomarkers from a Fielded-Pyrolysis-Gas Chromatography-Ion Mobility Spectrometry Biodetector with the Microbiologiucal Gram Stain Classification Scheme. *Anal Chem* 2004, 76, 6492-6499.

35. Donohue, J. P.; Welsh, W. J., Speciation of Listeria via pyrolysis/gas chromatography. *J Anal Appl Pyrolysis* 2004, 72, 221-228.

36. Islas, C. A.; Suelves, I.; Carter, J. F.; Herod, A. A.; Kandiyoti, R., Pyrolysis-gas chromatography/mass spectrometry of coal extrac and its fractions seperated by planar chromatrography: correlation of structure features with molecular mass. *Rapid Commun Mass Spectrom* 2000, 14, 1766-1782.

37. Timmins, E. M.; Goodacre, R., Rapid and quantitative analysis of binary mixtures of *Escherichia coli* strains using pyrolysis mass spectrometry with multivariate calibration and artifical neural networks. *J Appl Microbiol* 1997, 82, 208-218.

38. Kaneda, T., Biosynthesis of Branched Chain Fatty Acids. *J Biol Chem* 1962, 238, (4), 1222-1228.

39. Abel, K.; DeSchmertzing, H.; Peterson, J. I., Classification of Microorganisms by Analysis of Chemical Composition. *J Bacteriol* 1962, 1039-1044.

40. Whilttaker, P.; Fry, F. S.; Curtis, S. K.; Al-Khadli, S. F.; Mossoba, M. M.; Yurawecz, M. P.; Dunkel, V., C., Use of Fatty Acid Profiles to Identify Food-Borne Bacterial Pathogens and Aerobic Endospore-Forming Bacilli. *J Agric Food Chem* 25005, 53, 3735-3742.

41. Dees, S. B.; Moss, C. W., Cellular fatty acids of Alcaligenes and Pseudomonas species isolated from clinical specimens. *J Clin Microbiol* 1975, 1, (5), 414-9.

42. Basile, F.; Berverly, M. B.; Abbas-Hawks, C.; Mowry, C. D.; Voorhees, K. J., Direct Mass Spectrometric Analysis of *in Situ* Thermally Hydrolyzed and Methylated Lipids from Whole Bacterial Cells. *Anal Chem* 1998, 70, 1555-1562.

43. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., Molecular beams of macroions. *J Chem Phys* 1968, 49, (5), 2240-2249.

44. Yahashita, M.; Fenn, J. B., Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J Phys Chem* 1984, 88, 4451-4459.

45. Goodacre, R.; Heald, J. K.; Kell, D. B., Characterization of intact microorganisms using electrospray ionization mass spectrometry. *FEMS Microbiology Letters* 1999, 176, 17-24.

46. Vaidyanathan, S.; Rowland, J. J.; Kell, D. B.; Goodacre, R., Discrimination of Aerobic Endospore-forming Bacteria via Electorspray-Ionization Mass Spectrometry of Whole Cell Suspensions. *Anal Chem* 2001, 73, 4134-4144.

47. Black, G. E.; Fox, A.; Fox, K.; Snyder, A. P.; Smith, P. B., Electrospray tandem mass spectrometry for analysis of native muramic acid in whole bacterial cell hydrolysates. *Anal Chem* 1994, 66, (23), 4171-6.

48. Fox, A.; Rogers, J. C.; Gilbart, J.; Morgan, S.; Davis, C. H.; Knight, S.; Wyrick, P. B., Muramic acid is not detectable in Chlamydia psittaci or Chlamydia trachomatis by gas chromatography-mass spectrometry. *Infect Immun* 1990, 58, (3), 835-7.

49. Krishnamurthy, T.; Davis, M. T.; Stahl, D. C.; Lee, T. D., Liquid chromatography/microspray mass spectrometry for bacterial investigations. *Rapid Commun Mass Spectrom* 1999, 13, (1), 39-49.

50. Wunschel, D.; Fox, K. F.; Fox, A.; Nagpal, M. L.; Kim, K.; Stewart, G. C.; Shahgholi, M., Quantitative analysis of neutral and acidic sugars in whole bacterial cell hydrolysates using high-performance anion-exchange liquid chromatography-electrospray ionization tandem mass spectrometry. *J Chromatogr A* 1997, 776, (2), 205-19.

51. Wang, W.; Liu, Z.; Ma, L.; Hao, C.; Liu, S.; Voinov, V. G.; Kalinovskaya, N. I., Electrospray ionization multiple-stage tandem mass spectrometric analysis of diglycosyldiacylglycerol glycolipids from the bacteria Bacillus pumilus. *Rapid Commun Mass Spectrom* 1999, 13, (12), 1189-1196.

52. Smith, P. B.; Snyder, A. P.; Harden, C. S., Characterization of bacterial phospholipids by electrospray ionization tandem mass spectrometry. *Anal Chem* 1995, 67, (11), 1824-30.

53. Gibson, B. W.; Melaugh, W.; Phillips, N. J.; Apicella, M. A.; Campagnari, A. A.; Griffiss, J. M., Investigation of the structural heterogeneity of lipooligosaccharides from pathogenic Haemophilus and Neisseria species and of R-type lipopolysaccharides from Salmonella typhimurium by electrospray mass spectrometry. *J Bacteriol* 1993, 175, (9), 2702-12.

54. Krahmer, M. T.; Johnson, Y. A.; Walters, J. J.; Fox, K. F.; Fox, A.; Nagpal, M., Electrospray quadrupole mass spectrometry analysis of model oligonucleotides and polymerase chain reaction products: determination of base substitutions, nucleotide additions/deletions, and chemical modifications. *Anal Chem* 1999, 71, (14), 2893-900.

55. Liu, C.; Hofstadler, S. A.; Bresson, J. A.; Udseth, H. R.; Tsukuda, T.; Smith, R. D.; Synder, A. P., On-line dual microdialysis with ESI-MS for direct analysis of complex biological samples and microorganism lysates. *Anal Chem* 1998, 70, (9), 1797-801.

56. Fang, J.; Barcelona, M. J.; Semrau, J. D., Characterization of methanotrophic bacteria on the basis of intact phospholipid profiles. *FEMS Microbiol Lett* 2000, 189, (1), 67-72.

57. Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G., Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* 2004, 306, (5695), 471-3.

58. Takats, Z.; Wiseman, J. M.; Cooks, R. G., Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *J Mass Spectrom* 2005, 40, (10), 1261-75.

59. Song, Y.; Talaty, N.; Tao, W. A.; Pan, Z.; Cooks, R. G., Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. *Chem Commun (Camb)* 2007, (1), 61-3.

60. Cooks, R. G.; Ouyang, Z.; Takats, Z.; Wiseman, J. M., Detection Technologies. Ambient mass spectrometry. *Science* 2006, 311, (5767), 1566-70.

61. Heller, D. N.; Fenselau, C.; Cotter, R. J.; Demirev, P.; Olthoff, J. K.; Honovich, J.; Uy, M.; Tanaka, T.; Kishimoto, Y., Mass spectral analysis of complex lipids desorbed directly from lyophilized membranes and cells. *Biochem Biophys Res Commun* 1987, 142, (1), 194-9.

62. Ullom, J. N.; Frank, M.; Gard, E. E.; Horn, J. M.; Labov, S. E.; Langry, K.; Magnotta, F.; Stanion, K. A.; Hack, C. A.; Benner, W. H., Discrimination between bacterial spore types using time-of-flight mass spectrometry and matrix-free infrared laser desorption and ionization. *Anal Chem* 2001, 73, (10), 2331-7.

63. Fergenson, D. P.; Pitesky, M. E.; Tobias, H. J.; Steele, P. T.; Czerwieniec, G. A.; Russell, S. C.; Lebrilla, C. B.; Horn, J. M.; Coffee, K. R.; Srivastava, A.; Pillai, S. P.; Shih, M. T.; Hall, H. L.; Ramponi, A. J.; Chang, J. T.; Langlois, R. G.; Estacio, P. L.; Hadley, R. T.; Frank, M.; Gard, E. E., Reagentless detection and classification of individual bioaerosol particles in seconds. *Anal Chem* 2004, 76, (2), 373-8.

64. Gard, E.; Mayer, J. E.; Morrical, B. D.; Dienes, T.; Fergenson, D. P.; Prather, K. A., Real-Time Analysis of Individual Atmospheric Aerosol Particles: Design and Performance of a Portable ATOFMS. *Anal Chem* 1997, 69, (20), 4083-4091.

65. Steele, P. T.; Tobias, H. J.; Fergenson, D. P.; Pitesky, M. E.; Horn, J. M.; Czerwieniec, G. A.; Russell, S. C.; Lebrilla, C. B.; Gard, E. E.; Frank, M., Laser power dependence of mass spectral signatures from individual bacterial spores in bioaerosol mass spectrometry. *Anal Chem* 2003, 75, (20), 5480-7.

66. Russell, S. C.; Czerwieniec, G.; Lebrilla, C.; Tobias, H.; Fergenson, D. P.; Steele, P.; Pitesky, M.; Horn, J.; Srivastava, A.; Frank, M.; Gard, E. E., Toward understanding the ionization of biomarkers from micrometer particles by bio-aerosol mass spectrometry. *J Am Soc Mass Spectrom* 2004, 15, (6), 900-9.

67. Murray, K. K.; Russell, D. H., Aerosol Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *J Am Soc Mass Spectrom* 1993, 5, 1-9.

68. Edirisinghe, P. D.; Moore, J. F.; Skinner-Nemec, K. A.; Lindberg, C.; Giometti, C. S.; Veryovkin, I. V.; Hunt, J. E.; Pellin, M. J.; Hanley, L., Detection of *in situ* derivatized peptides in microbial biofilms by laser desorption 7.87 eV postionizaton mass spectrometry. *Anal Chem* 2007, 79, (2), 508-14.

69. Sampson, J. S.; Hawkridge, A. M.; Muddiman, D. C., Direct characterization of intact polypeptides by matrix-assisted laser desorption electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun Mass Spectrom* 2007, 21, (7), 1150-4.

70. Jackson, S. N.; Murray, K. K., Matrix addition by condensation for matrix-assisted laser desorption/ionization of collected aerosol particles. *Anal Chem* 2002, 74, (18), 4841-4.

71. Valentine, N. B.; Wahl, J. H.; Kingsley, M. T.; Wahl, K. L., Direct surface analysis of fungal species by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* 2002, 16, (14), 1352-7.

72. Lay, J. O., Jr.; Holland, R. D., Rapid identification of bacteria based on spectral patterns using MALDI-TOFMS. *Methods Mol Biol* 2000, 146, 461-87.

73. Yao, Z. P.; Demirev, P. A.; Fenselau, C., Mass spectrometry-based proteolytic mapping for rapid virus identification. *Anal Chem* 2002, 74, (11), 2529-34.

74. Lay, J. O., Jr., MALDI-TOF mass spectrometry of bacteria. *Mass Spectrom Rev* 2001, 20, (4), 172-94.

75. Honisch, C.; Chen, Y.; Mortimer, C.; Arnold, C.; Schmidt, O.; van den Boom, D.; Cantor, C. R.; Shah, H. N.; Gharbia, S. E., Automated comparative sequence analysis by base-specific cleavage and mass spectrometry for nucleic acid-based microbial typing. *Proc Natl Acad Sci U S A* 2007, 104, (25), 10649-54.

76. Jones, J. J.; Batoy, S. M.; Wilkins, C. L., A comprehensive and comparative analysis for MALDI FTMS lipid and phospholipid profiles from biological samples. *Comput Biol Chem* 2005, 29, (4), 294-302.

77. Tao, L.; Yu, X.; Snyder, A. P.; Li, L., Bacterial identification by protein mass mapping combined with an experimentally derived protein mass database. *Anal Chem* 2004, 76, (22), 6609-17.

78. Jarman, K. H.; Cebula, S. T.; Saenz, A. J.; Petersen, C. E.; Valentine, N. B.; Kingsley, M. T.; Wahl, K. L., An algorithm for automated bacterial identification using matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 2000, 72, (6), 1217-23.

79. Rose, N. L.; Sporns, P.; McMullen, L. M., Detection of bacteriocins by matrix-assisted laser Desorption/Ionization time-of-flight mass spectrometry. *Appl Environ Microbiol* 1999, 65, (5), 2238-42.

80. Ryzhov, V.; Hathout, Y.; Fenselau, C., Rapid characterization of spores of Bacillus cereus group bacteria by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry. *Appl Environ Microbiol* 2000, 66, (9), 3828-34.

81. Demirev, P. A.; Ho, Y. P.; Ryzhov, V.; Fenselau, C., Microorganism identification by mass spectrometry and protein database searches. *Anal Chem* 1999, 71, (14), 2732-8.

82. Jarman, K. H.; Daly, D. S.; Petersen, C. E.; Saenz, A. J.; Valentine, N. B.; Wahl, K. L., Extracting and visualizing matrix-assisted laser desorption/ionization time-of-flight mass spectral fingerprints. *Rapid Commun Mass Spectrom* 1999, 13, (15), 1586-94.

83. Dworzanski, J. P.; Snyder, A. P., Classification and identification of bacteria using mass spectrometry-based proteomics. *Expert Rev Proteomics* 2005, 2, (6), 863-78.

84. Dworzanski, J. P.; Snyder, A. P.; Chen, R.; Zhang, H.; Wishart, D.; Li, L., Identification of bacteria using tandem mass spectrometry combined with a proteome database and statistical scoring. *Anal Chem* 2004, 76, (8), 2355-66.

85. Wahl, K. L.; Wunschel, S. C.; Jarman, K. H.; Valentine, N. B.; Petersen, C. E.; Kingsley, M. T.; Zartolas, K. A.; Saenz, A. J., Analysis of microbial mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal Chem* 2002, 74, (24), 6191-9.

86. Demirev, P. A.; Lin, J. S.; Pineda, F. J.; Fenselaut, C., Bioinformatics and mass spectrometry for microorganism identification: proteome-wide post-translational modifications and database search algorithms for characterization of intact H. pylori. *Anal Chem* 2001, 73, (19), 4566-73.

87. Warscheid, B.; Fenselau, C., A targeted proteomics approach to the rapid identification of bacterial cell mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics* 2004, 4, (10), 2877-92.

88. Pribil, P.; Fenselau, C., Characterization of enterobacteria using MALDI-TOF mass spectrometry. *Anal Chem* 2005, 77, (18), 6092-5.

89. McLuckey, S. A., Principles of Collisional Activation in Analytical Mass Spectrometry. *J Am Soc Mass Spectrom* 1992, 3, 599-614.

90. Mitchell, J. P., Aerosol Generation for Instrument Calibration. In *Bioaerosols Handbook*, Cox, C. S.; Wathes, C. M., Eds. Lewis Publishers: 1995; pp 101-175.

91. Lin, X.; Reponen, T.; Willeke, K.; Wang, Z.; Grinshpun, S. A.; Trunov, M., Survival of Airborne Microorganisms During Swirling Aerosol Collection. *Aerosol Sci Technol* 2000, 32, (3), 184-196.

92. Willeke, K.; Lin, X.; Grinshpun, S. A., Improved Aerosol Collection by Combined Impaction and Centrifugal Motion. *Aerosol Sci Technol* 1998, 28, (5), 439-456.

93. Fenn, J. B., Electrospray Wings for Molecular Elephants. *Nobel Lectures* 2002, 154-184.

94. Tanaka, K., The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew Chem Int Ed Engl* 2003, 42, (33), 3860-70.

95. Kirpekar, F.; Berkenkamp, S.; Hillenkamp, F., Detection of double-stranded DNA by IR- and UV-MALDI mass spectrometry. *Anal Chem* 1999, 71, (13), 2334-9.

96. Jackson, S. N.; Wang, H. Y.; Woods, A. S., Direct profiling of lipid distribution in brain tissue using MALDI-TOFMS. *Anal Chem* 2005, 77, (14), 4523-7.

97. Nielsen, M. L.; Bennett, K. L.; Larsen, B.; Moniatte, M.; Mann, M., Peptide end sequencing by orthogonal MALDI tandem mass spectrometry. *J Proteome Res* 2002, 1, (1), 63-71.

98. Dreisewerd, K., The desorption process in MALDI. *Chem Rev* 2003, 103, (2), 395-426.

99. Dass, C., *Principles and Practice of Biological Mass Spectrometry*. John Wiley & Sons, Inc.: 2001.

100. Hillenkamp, F.; Karas, M., The MALDI Process and Method. In *MALDI MS: A Pratcial Guide to Instrumentation, Methods and Applications*, Hillenkamp, F.; Peter-Katalinic, J., Eds. Wiley-VHC: 2007; pp 1-28.

101. Karas, M.; Gluckmann, M.; Schafer, J., Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom* 2000, 35, (1), 1-12.

102. Wiley, W. C.; McLaren, I. H., Time-of-flight mass spectrometer with improved resolution. *Rev Sci Instrum* 1955, 26, (12), 1150-1157.

103. Mamyrin, B. A.; J., K. V.; Shmikk, D. V.; Zagulin, V. A., The mass reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Sov Phys JETP* 1973, 37, 45-48.

104. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A., An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom* 2001, 36, (8), 849-65.

105. Jennings, K. R., Collision-induced decompositions of aromatic molecular ions. *Int J Mass Spectrom Ion Phys* 1968, 1, (3), 227-235.

106. Hayes, R. N.; Gross, M. L.; James, A. M., [10] Collision-induced dissociation. In *Methods in Enzymology*, Academic Press: 1990; Vol. Volume 193, pp 237-263.

107. Despeyroux, D.; Jennings, K. R., *Collision-induced dissociation*. Wiley: New York, 1994.

108. Roepstorff, P; Fohlman, J., Letter to the editors. *J Biomed Mass Spectrom* 1984, 11, (11), 601.

109. Wilkins, M. R.; Sanchez, J. C.; Gooley, A. A.; Appel, R. D.; Humphery-Smith, I.; Hochstrasser, D. F.; Williams, K. L., Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 1996, 13, 19-50.

110. Behr, J.; Israel, L.; Ganzle, M. G.; Vogel, R. F., Proteomic Approach for Characterization of Hop-Inducible Proteins in Lactobacillus brevis. *Appl Environ Microbiol* 2007, 73, (10), 3300-3306.

111. Pruteanu, M.; Neher, S. B.; Baker, T. A., Ligand-Controlled Proteolysis of the Escherichia coli Transcriptional Regulator ZntR. *J Bacteriol* 2007, 189, (8), 3017-3025.

112. Jarman, K. H.; Wahl, K. L., *Development of spectral pattern-matching approaches to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for bacterial identification*. John Wiley and Sons, Inc.: Hoboken, 2006.

113. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A., ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003, 31, (13), 3784-8.

114. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, 20, (18), 3551-67.

115. Beavis, R. C.; Chait, B. T., Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid Commun Mass Spectrom* 1989, 3, (12), 432-5.

116. Pappin, D. J.; Hojrup, P.; Bleasby, A. J., Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* 1993, 3, (6), 327-32.

117. Clauser, K. R.; Baker, P.; Burlingame, A. L., Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* 1999, 71, (14), 2871-82.

118. Bajuk, A.; Gluch, K.; Michalak, L., Effect of impurities on the matrix-assisted laser desorption/ionization mass spectra of insulin. *Rapid Commun Mass Spectrom* 2001, 15, (24), 2383-2386.

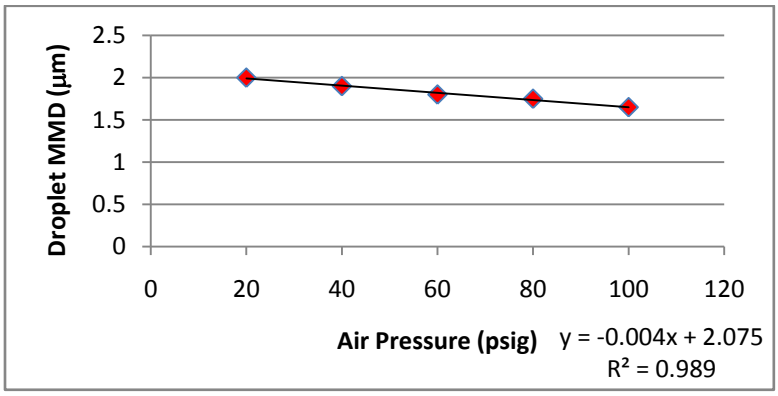
119. Arnold, R. J.; Reilly, J. P., Fingerprint matching of *E. coli* strains with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of whole cells using a modified correlation approach. *Rapid Commun Mass Spectrom* 1998, 12, (10), 630-636.

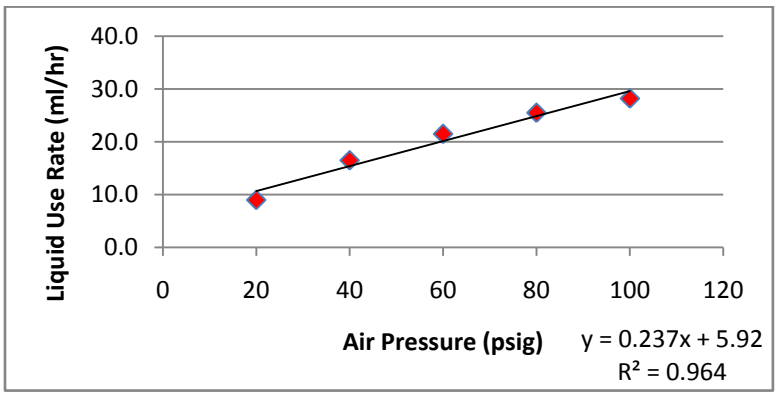
120. Yao, Z. P.; Afonso, C.; Fenselau, C., Rapid microorganism identification with on-slide proteolytic digestion followed by matrix-assisted laser desorption/ionization tandem mass spectrometry and database searching. *Rapid Commun Mass Spectrom* 2002, 16, (20), 1953-6.

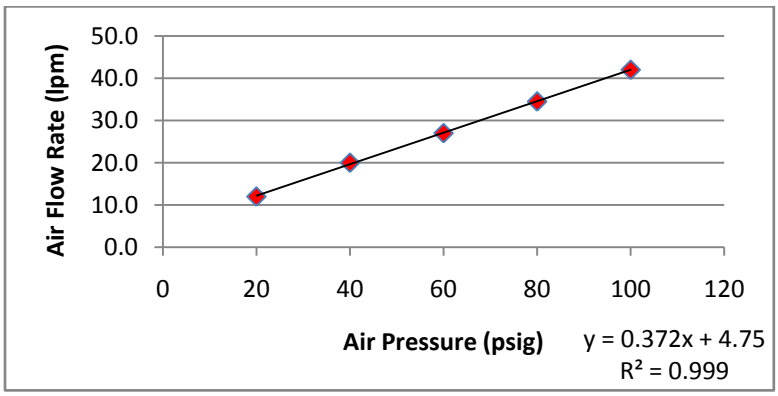
121. Harris, W. A.; Reilly, J. P., On-probe digestion of bacterial proteins for MALDI-MS. *Anal Chem* 2002, 74, (17), 4410-6.

APPENDIX A: COLLISON NEBULIZER CHARACTERISTICS[a]

Operational characteristics (particle size, liquid sample volume consumption, and air flow rate) of the 6-jet collison nebulizer as controlled by pressure.







[a] May, K. R., The collison nebulizer: Description, performance and application. *Journal of Aerosol Science* **1973,** 4, (3), 235-238.

APPENDIX B. GRAM STAINING PROCEDURE FOR BACTERIA

A tradiational gram staining procedure for the classical gram-stain for determining bacteria gram-type.

PROCEDURES:

1. Slide preparation.
    a. Place a wet droplet of the sample onto a microscope slide.
    b. Heat the slide for few seconds over a Bunsen flame until it becomes hot to the touch (This mounts/fixes the bacteria firmly to the slide).
    c. Let the sample cool.
2. Add the primary stain (crystal violet) and let stand for 1 minute.
3. Rinse off the stain with water and use a wipe to wick the moisture off the slide.
4. Add Gram's iodine (the mordant) and let stand for 30 seconds.
5. Wash with the decolorizer until all excess dye is removed.
6. Add the secondary stain (safranin O) and let stand for 1 minute.
7. Wash with water until the excess stain is removed.

RESULTS:

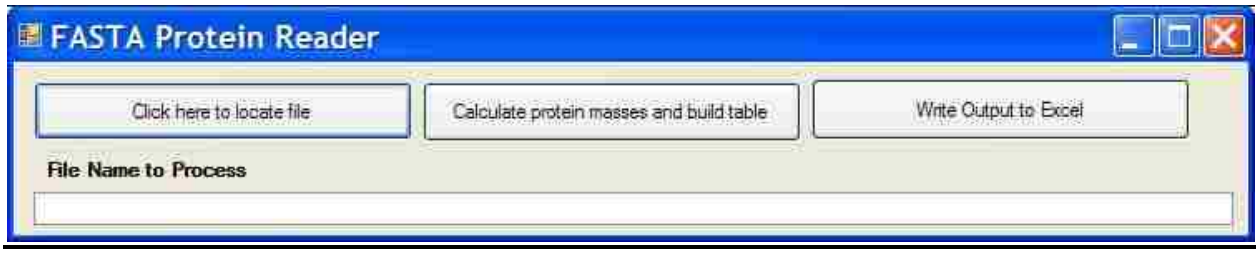
- Gram-negative bacteria will lose the primary stain, take secondary stain, and will appear red-pink.
- Gram-positive bacteria will keep the primary stain and will appear violet.

REAGENTS:

- Primary stain = 20 g of crystal violet dissolved in 200 ml of 95% ethyl alcohol.
- Secondary stain = 2.5 g of safranin O in 100 ml of 95 % ethyl alcohol
- Gram's iodine = 1 g of iodine and 2 g of potassium iodide dissolved in 300 ml of dH_2O.
- Decolorizer = 50% ethyl alcohol/ 50% acetone

APPENDIX C. FASTA SEQUENCE EXTRACTOR

Visual basic code for the program used to calculate molecule weights of all proteins from FASTA databases for organisms and incorporation into an in-house microorganism database.



MainForm.vb 1

```
1    Public Class MainForm
2
3         Private Sub Button1_Click(ByVal sender As
     System.Object, ByVal e             As System.EventArgs)
     Handles Button1.Click
4         OpenFileDialog1.ShowDialog()
5
6
7    End Sub
8
9    Private Sub OpenFileDialog1_FileOk(ByVal sender As
     System.Object, ByVal
      e As System.ComponentModel.CancelEventArgs) Handles
      OpenFileDialog1.FileOk
10        Dim vartext As String
11        vartext = OpenFileDialog1.FileName
12        TextBox1.Text = vartext
13        Call CountSequences()
14
15   End Sub
16
17   Private Sub Button2_Click(ByVal sender As System.Object,
     ByVal e As
      System.EventArgs) Handles Button2.Click
18        Dim fileReader As String
19        fileReader =
     My.Computer.FileSystem.ReadAllText(TextBox1.Text)
20        Call ProteinSequences(fileReader)
21   End Sub
```

```
22
23  Sub ProteinSequences(ByVal fileString As String)
24       Dim count1, count2, count3, count4 As Integer
25       Dim PositionCounter As Integer
26       Dim parsedText As String
27       Dim mass2 As Double
28
29
30       count1 = 0
31       count2 = 0
32       count4 = 0
33
34       PositionCounter = 1
35       count3 = Strings.InStr(PositionCounter, fileString,
    Chr(13))
36       count4 = Strings.InStr(count3 + 1, fileString, Chr(13))
37       parsedText = Mid(fileString, count3 + 2, count4 -
    count3 - 2)
38       TextBox1.Text = count3
39       TextBox2.Text = count4
40       TextBox3.Text = parsedText
41       Call CalculateMass(parsedText, mass2)
42
43       Dim objExcel As New
    Microsoft.Office.Interop.Excel.Application
44       Dim ExportFileName As String
45       Dim Prompt, Title, DefaultResponse As String
46       Dim XPos, YPos As Integer
47       XPos = 10
48       YPos = 10
49       Prompt = "Please enter a name for the export file."
50       Title = "EXPORT FILENAME"
51       DefaultResponse = ""
52       ExportFileName = Interaction.InputBox(Prompt, Title,
          DefaultResponse, XPos,YPos)
53       With objExcel
54            .Visible = True
55            .Workbooks.Add()
56            .Cells(1, 1) = TextBox6.Text
57            .SaveWorkspace("C:\My Programs\PHAT\Data\" &
    ExportFileName
               & ".xls")
58       End With
59
60  End Sub
61
62
```

```
63   Private Sub MainForm_Load(ByVal sender As System.Object,
     ByVal e As
      System.C:\Documents and Settings\Alton Dugas\My
      Documents\ProteinFASTAreader\MainForm.vb 2EventArgs)
     Handles
      MyBase.Load
64
65        Call AminoAcidMasses()
66   End Sub
67
68   Sub AminoAcidMasses()
69        Dim count As Integer
70        Dim filereader As IO.StreamReader = New
           IO.StreamReader("C:\Documents and Settings\Alton
     Dugas\My
           Documents\My
           LSU\AminoAcidMasses.txt")
71
72        count = 1
73
74        Do While count < 21
75             AminoAcid(count) = filereader.ReadLine()
76             ListBox1.Items.Add(AminoAcid(count))
77             count = count + 1
78        Loop
79
80
81   End Sub
82
83   Sub CalculateMass(ByVal Sequence As String, ByVal mass As
     Double)
84        Dim SeqLen As Integer
85        Dim count, AAcount As Integer
86        Dim AA As String
87        AAcount = 0
88        count = 1
89        mass = 0
90
91        SeqLen = Sequence.Length
92        TextBox4.Text = SeqLen
93
94        Do Until SeqLen <= 0
95             AA = Mid(Sequence, count, 1)
96             Select Case AA
97             Case "A"
98                  AAcount = 1
99             Case "C"
```

```
100                 AAcount = 2
101             Case "D"
102                 AAcount = 3
103             Case "E"
104                 AAcount = 4
105             Case "F"
106                 AAcount = 5
107             Case "G"
108                 AAcount = 6
109             Case "H"
110                 AAcount = 7
111             Case "I"
112                 AAcount = 8
113             Case "K"
114                 AAcount = 9
115             Case "L"
116                 AAcount = 10
117             Case "M"
118                 AAcount = 11
119             Case "N"
120                 AAcount = 12
121             Case "P"
122                 AAcount = 13
123             Case "Q"
124                 AAcount = 14
125             Case "R"
126                 AAcount = 15
127             Case "S"
128                 AAcount = 16
129             Case "T"
130                 AAcount = 17
131             Case "V"
132                 AAcount = 18
133             Case "W"
134                 AAcount = 19
135             Case "Y"
136                 AAcount = 20
137             Case Else
138                 MsgBox("wrong AA")
139             End Select
140             mass = mass + AminoAcid(AAcount)
141             SeqLen = SeqLen - 1
142             count = count + 1
143         Loop
144         mass = mass + 18.01056469
145         TextBox5.Text = mass
146
```

```
147 End Sub
148
149 Sub CountSequences()
150      Dim CountFileReader As String
151      Dim count, SumCount, count2 As Integer
152      count = 1
153      SumCount = 0
154      count2 = 0
155
156      CountFileReader =My.Computer.FileSystem.ReadAllText
          (TextBox1.Text)
157      count2 = Strings.InStr(count, CountFileReader, ">")
158      count = count2
159      Do Until count2 = 0
160           count = count + 1
161           SumCount = SumCount + 1
162           count2 = Strings.InStr(count, CountFileReader,
    ">")
163           count = count2
164      Loop
165      TextBox6.Text = SumCount
166 End Sub
167
168 Private Sub Button3_Click(ByVal sender As System.Object,
    ByVal e As
     System.EventArgs) Handles Button3.Click
169
170
171 End Sub
172 End Class
```

Module1.vb

```
1   Module Module1
2        Public SequenceArray(10000) As String
3        Public AminoAcid(20) As Double
4
5   End Module
```

VITA

Alton John Dugas, Jr., was born in Donaldsonville, Louisiana. He graduated with a Bachelor of Science degree in 1995 and a Master of Science degree in 1998 in biochemistry from Louisiana State University. After receiving his M.S. degree, Alton was accepted into the United States Air Force's Officer Training School where he received a commission and the rank of Second Lieutenant on September 30, 1999. He then went on to serve an assignment as a chemical research officer in support of counterproliferation efforts by acting as the field test director for the chemical technologies branch of the materials directorate. During this assignment, Alton attended the Air Force's Aerospace Basic Course. In 2001, he was promoted to the rank of First Lieutenant. He was reassigned to the United States Air Force Academy's Preparatory School as a science instructor in 2002. In 2003, Alton was promoted to the rank of Captain and attended the Air Force's Squadron Officer School. To date, Alton has earned the following awards and decorations while serving in the United States Air Force: the Air Force Commendation Medal with one oak leaf cluster, the Air Force Organizational Excellence Award, the National Defense Service Medal, the Global War on Terrorism Service Medal, the Air Force Longevity Medal with one oak leaf cluster and the Air Force Training Ribbon. In 2004, Alton was selected to pursue a doctorate degree in chemistry through the Air Force Institute of Technology's civilian school program. For his doctoral studies, Alton chose to return to Louisiana State University and selected Dr. Kermit K. Murray as his doctoral advisor. Alton is a member of the American Chemical Society and the American Society of Mass Spectrometry. During his doctoral program, Alton was selected to present research at four conferences. Captain Dugas is currently a candidate for the degree of Doctor of Philosophy in chemistry, which will be awarded at the May 2008 Commencement.