STRUCTURAL CHARACTERIZATION OF LIPOPEPTIDES ISOLATED FROM *BACILLUS globigii* SPORES

by

Bruce Howard Williams

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2001

Advisory Committee:

Professor Catherine Fenselau, Chairman/Advisor
Associate Professor Dorothy Beckett
Associate Professor Douglas Julin
Associate Professor David Jollie

ABSTRACT

Title of Thesis: STRUCTURAL CHARACTERIZATION OF LIPOPEPTIDES ISOLATED FROM *BACILLUS globigii* SPORES

Degree candidate: Bruce Howard Williams

Degree and year: Master of Science, 2001

Thesis directed by: Professor Catherine Fenselau, Department of Chemistry and Biochemistry

Several species of *Bacillus* spores produce surface-active antibiotic and antifungal lipopeptides, phospholipids, and polypeptides. Some of these are detected as biomarkers when intact spores are analyzed by Matrix Assisted Laser Desorption/Ionization time-of-flight (MALDI-TOF) mass spectrometry. *Bacillus globigii* spores, grown in new sporulation media (NSM), were suspended and then analyzed using a MALDI-TOF mass spectrometer to screen for biomarkers with 4-methoxycinnamic acid as matrix. Spores, vegetative cells, and the culture supernatant show a biomarker cluster composed of six molecules and centered at a molecular mass of 1478 Da. The biomarker was isolated from the cell-free culture supernatant by solid phase extraction to yield a crude product. Its components were concentrated, separated by reverse-phase high-pressure liquid

chromatography, and then eluted as a cluster of three peaks. The three peaks were collected separately and preliminary analysis was conducted with a MALDI-TOF. A Fourier Transform Mass Spectrometer (MALDI source) instrument was used to determine the monoisotopic protonated masses to be 1463.8, 1477.8, and 1505.8 Da, respectively to their order of elution. The mass differences of 14 and 28 Da, suggest that they are homologous molecules. Acid hydrolysis of each molecule showed it to contain an ester linkage. Methyl esterification, using methanolic HCl, of the open form of the molecule at m/z 1477.8 resulted in a product 56 Da higher, indicating the presence of four carboxylic acid moieties. A quadripole-time-of-flight instrument with nanospray source (Q-star) was used to sequence the reaction products. The purified compounds have identical amino acid sequences to previously identified lipopeptides in *Bacillus subtilis*.

ACKNOWLEDGEMENTS

I would like to express my appreciation to Professor Catherine Fenselau for her guidance and encouragement throughout this research.

I am deeply indebted to Dr. Yetrib Hathout for his guidance, patience, and enthusiasm, at all phases in conducting this research.

I would like to thank Dr. Xudong Yao, Dr. Carlos Afonso, Dr. Javier Ramirez, and Kristy Reynolds for their assistance in conducting the tandem mass spectrometry and peptide sequencing work. Their suggestions, interest, and friendship were invaluable to my progress.

I would like to thank Amy Freas who grew the microorganism and offered her expertise in microbiology throughout this research. Additionally, I would like to show appreciation to the members of the Fenselau research group: Dr. Plamen Demeriv, Kevin Schefcheck, Amir Rahbar, Jamie Lu, Qoung Ping Fu, and Bijan Iliasi.

Most importantly, I would like to thank my wife Patrice, whose love and support made this all possible.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Background

Over a century ago, the possible role of bacteria in disease had been suggested and widely discussed. The early studies that led to the discovery of bacterial spores also led to the final proof that microorganisms caused infectious disease [1]. New pathogens have emerged, and some have spread worldwide. Many, including *Salmonella, E. coli 0157*, and *Campylobactor,* have reservoirs in healthy food animals, from which they spread to an increasing variety of foods. These pathogens cause millions of cases of sporadic illness and chronic complications, as well as large and challenging outbreaks over many states and nations [2]. Contamination with foodborne microorganisms often eludes traditional food inspection, which relies on visual identification of foodborne hazards. The foods contaminated with emerging pathogens often look, smell, and taste normal. The pathogen frequently survives traditional preparation techniques: *E. coli 0157* in meat can survive the mild heating that a rare hamburger gets; *Salmonella* in eggs survives in an omelette; and Norwalk virus in oysters survives gentle steaming [2]. Following traditional recipes and preparation methods has caused illness and highly publicized outbreaks.

Other highly publicized incidents, including the bombing of the federal building in Oklahoma City and the 1996 Olympics in Atlanta, have highlighted the rising incidence of terrorist acts. Once thought of as a problem relegated to trouble spots in far-away nations, people of all countries are now aware that a

terrorist can attack any time, anywhere. The attraction of biological weapons in war and for terrorist attacks is attributed to easy access to a wide range of disease-producing biological agents, to their low production costs, and to their ease of transport. The issue of biological terrorism is particularly frightening because biological weapons are silent, invisible to the naked eye, and easy to conceal from routine security systems, but can be as deadly as a nuclear bomb. The fall of the Soviet Union and the defeat of Iraq uncovered extensive biological weapons programs of surprising sophistication and diversity. Major biological agents from the Cold War included the causative agents of hemorrhagic fever, plague, smallpox, and viral encephalitis and intelligence reports suggested that Iraq was preparing the disease causing agents of anthrax and botulism as weapons during the 1991 Gulf War [3]. The threat from biological weapons was no longer an issue of debate and military laboratories were tasked with the challenge of developing timely and specific methods for detecting these biological agents. Their goal was to detect and correctly identify agents so that the appropriate medical treatment could be given to those exposed [3].

Whether a medical emergency results from a new foodborne disease or from an act of terrorism, one critical question is always asked: What is a simple way to rapidly and easily identify the pathogen and diagnose the disease? The identification and characterization of microorganisms has become an important issue in diagnosing disease, detecting biological warfare agents, and monitoring the potential contamination of water and stored foods. In screening for the presence of microorganisms, distinguishing them as bacteria, viruses or fungi is

often the initial level of analysis. In more complex cases, where biological or environmental hazards are concerned, species or even strains must be identified.

Traditional Techniques for Bacterial Identification

Traditional approaches to the identification of the bacteria in biological mixtures relied heavily on certain morphological and biochemical characteristics. Commonly, these methods relied on microscopic examination of an isolated organism taken from homogeneous colonies cultured on specific growth media [4]. Heterogeneous populations, representative of field conditions, caused problems and gave inconsistent results. These culture-based methods of identification are inherently slow because of the requirement to first purify and then grow microorganisms over a period of several days. In instances where these methods failed to isolate the organism, the direct analysis of samples was occasionally used for tentative but not definitive identification. Microbiologists often made presumptive identifications based on observation of discrete masses of colonies grown on the surface of solid culture media [4]. The colonies of certain bacterial species are characteristic with regard to size and shape, outline, elevation, and translucency. Bacteria can exist as spheres, rods, or spirals. Whereas the length of the cell may vary substantially due to different physicochemical conditions of growth, the width of cells is a more constant feature [4]. For example, when grown on nutrient agar, only a few *Bacillus* species, such as *B. cereus, B. anthracis, and B. thuringiensis,* have a width less than one micrometer. Often small differences in colony structure, when

compared to known strains resulted in the creation of new species. The color of the colonies (white, pigmented, or colorless) and their changes in the medium were additional criteria useful in the tentative identification of bacteria [1]. However, simple observation of colonial morphology was insufficient to differentiate one species from another. Different species of *Salmonella* produce colonies that are indistinguishable on ordinary media from those of other intestinal bacteria. Colonies of *Staphylococci* on solid media may appear white, yellow, or orange and may consequently be misidentified by color [4]. Whereas strains of a number of *Bacillus* species form rather characteristic colonies on a certain medium, strains of other species, like *B. subtilis*, may show extreme variation. Additionally, even pure cultured colonies have been observed to show different degrees of translucence or opaqueness. We now know, that colony appearance may vary considerably with such environmental factors as composition or moisture content of the medium or temperature or atmosphere of incubation [1].

The capacity of some bacteria to differentiate into heat resistant spores has traditionally been considered a taxonomic feature of crucial importance. This transformation of vegetative cells, or sporangia, to spores, predominates in the *Bacillus* species [1]. On nutrient agar, *Bacillus* species form endospores best detected under a phase contrast microscope by their brightness [5]. They may be cylindrical, oval, or round. During the formation of endospores, the sporangium may remain unchanged in shape or may swell. However, the use of spore formation as a main criterion to classify bacteria is complicated by the fact

that sporulation often depends on the conditions under which a culture is tested [4]. This scheme is still routinely used, but is limited in its application to strains that sporulate poorly on commercially available media or sporulate only under more narrowly defined conditions.

Correlating with the different reactions to the Gram stain, examination of electron micrographs shows that Gram-positive and Gram-negative organisms have very different cell envelopes. A cytoplasmic membrane bounded by a thick layer of peptidoglycan surrounds the Gram-positive bacterium . Gram-negative bacteria possess a double-layered envelope comprising the cytoplasmic membrane, a thin peptidoglycan within a hydrophilic compartment, or periplasm, and an outer membrane of lipopolysaccharide (LPS) containing the major antigenic sites that control pathogenic interactions in mammalian hosts [6]. Structural variation in these components can be used to classify and identify bacteria.

The most common biochemical tests involved the observation of whether or not a culture of the bacterium in liquid nutrient medium would ferment particular sugars such as glucose or lactose [4]. Acid or gas may be produced, which when detected would cause a change in color of an indicator dye present in the medium. Other tests determine whether the bacterium produces particular products when grown in suitable culture media. In addition, the activity of specific enzymes can be measured to aid in the identification of bacteria [4]. Procedures that rely on biochemical changes are useful in the detection and

identification of a broad range of organisms, but fail in the identification of specific groups of organisms.

Even within the last decade, differentiation and identification of most microorganisms has relied on the use of traditional procedures [7]. Although a bacterium could be identified presumptively by a battery of certain physical characteristics and biochemical assays, further specific tests were often performed to establish definitive identity of the organisms. Today cell morphology is often considered less important and small but typical differences between species merely aid in the identification of bacteria. The increasing number of recognized species and the occurrence of disease has made routine identification problematic and necessitated the need for more rapid and sensitive methods.

New Methods of Detection and Identification of Bacteria

Traditional procedures used to identify bacteria yield the correct answer in most instances, but require at least a 24-hour period for identification. Rapid diagnosis of life threatening infections due to the presence of bacteria or one or more of their products from tissue samples is of great importance. In order to help improve reporting time in diagnosis, many new methods of detection and identification of bacteria have been developed. Some traditional methods based on morphology and staining characteristics were simply modified while other new technologies were introduced. Automation of sample preparation and monitoring of several biochemical processes simultaneously have aided our ability to

characterize microorganisms [7]. These improvements were needed for the purposes of expediting identification, increasing sensitivity of detection, and increasing the reliability and accuracy of bacterial identification methods.

Bacteria are susceptible to infection by bacteriophages that are highly host specific. This method has been used to test the susceptibility of bacteria colonies to lysis by specific bacteriophages in public health investigations. Bacterial viruses are of great value in determining the presence of specific types of organisms by susceptibility patterns, and can prove two strains to be identical or different [7]. The technique involves the use of virulent bacteriophages, which are placed onto freshly cultured bacterial lawns. After incubation for 18 – 48 hours, a positive result is recorded by the presence of zones of clearing in the otherwise uniform layer of bacterial growth [4]. The reactions between bacteriophage and bacteria are extremely specific and have proven useful for the identification of many bacterial pathogens such as *Staphylococcus aureus* [7]

Applications of molecular biology, particularly nucleic acid hybridization techniques, have revolutionized microbial identification through the development of nucleic acid probes. Hybridization probes can make side-by-side comparisons or compare patterns with reference libraries. This scheme is suited primarily to situations where the approximate identity of the organism is recognized and identification of a species is required. The major application of probes is in the simultaneous detection and identification of microorganisms [4]. These two processes are intimately entwined in probe technology, since successful detection of a microbe in a sample using hybridization means that

complementary sequences are present and some indication of the identification of the gene or organism is provided. The hybridizations are normally done as immobilized assays on nitrocellulose or nylon membranes [5] [7]. The reactions are usually semi-quantitative, but can be made quantitative by careful attention to procedures [4]. The target DNA, usually chromosomal DNA, is immobilized using a vacuum filtration manifold that allows accurate allotment of DNA samples to a membrane in a set pattern. Dot-blots use an array of circular spots on the membrane, whereas slot-blots use an array of oval shaped spots. One of the most frequently used target nucleic acids is in the form of colony hybridizations [7]. Colonies on a plate are transferred to a membrane and then lysed and processed to remove protein and leave the denatured nucleic acids as single strands bound to the support ready for hybridization to the probe. A common problem with this procedure is high background levels due to contaminating materials. One application of this technique has been the detection and identification of microorganisms directly from the environment.

Improved methods based on monoclonal antibodies and DNA amplification by the polymerase chain reaction (PCR) have reduced the detection of microbial populations and the diagnoses of infections from days to hours [7]. These approaches have been successfully used to rapidly detect and identify the bacteria in biological fluids and other environments. However, these methods have limited applicability and sensitivity in the routine identification of microorganisms. Additionally, these methods cannot search for a large set of possible identities in a single analysis. Given the importance of the problem,

however, additional methods that apply to broader, mixed populations, offer greater speed, and offer complimentary sensitivity are needed.

Microbial Chemical Markers

The ability to unambiguously identify species and strains in complex mixtures and in remote field environments remains an area of active research. One approach to this problem has been to classify based on analysis of chemical components as unique biomarkers, commonly referred to as chemotaxonomy [1]. In general, bacteria lack the complex intracellular structures found in plant and animal cells, but still contain a diverse set of chemical markers for taxonomic discrimination. Chemotaxonomy is especially suited for those microorganisms in which there are few useful morphological and physiological characteristics. These markers can vary with nutrition and life cycle of the organism, but are generally stable enough to be characteristic.

Nucleic Acids

Of the various chemical components used for identification, only chromosomal DNA is unaffected by changes in the growth conditions. DNA is therefore the most individually characteristic class of chemical markers in identifying microorganisms. It is now generally accepted that DNA base composition (mol% G + C) is a constant and therefore a taxonomically important characteristic of bacterial species. Strains belonging to the same species do not differ in DNA base composition by more than 1-2 % [1]. Species belonging to the

same genus usually have a mol% G+C within a range not broader than 10%. Whereas closely related species have similar mole percentage G+C values, it is important to recognize that two organisms with similar DNA base compositions are not necessarily closely related since the deoxyribonucleotides in the DNA may not be in the same linear order. The DNA base composition of species of the *Bacillus* has been found to be in the range of 32- 69 mole percentage G+C, showing that the genus is genetically heterogeneous [1]. At the species level, *Bacillus* strains have been shown to differ in their DNA base composition far more than expected, which reflects the fact that descriptions of species are generated on the basis of more than phenotypic characteristics [8]. The current appreciation of the species diversity of *Bacillus* species makes other-than traditional identification procedures more appropriate.

Proteins, Lipids, Carbohydrates, and Whole Cells

The use of chemical markers and the chemical variation in bacteria have proven to be of value in bacterial classification, and its chemical properties are an essential part in the description of genera and species. Proteins, lipids, cell wall constituents, and whole cells all provide a diverse set of such markers for taxonomic discrimination [1].

Measurement of relationships between organisms using proteins has focused on either comparisons of individual molecules or large evaluation of total cellular protein by comparisons of their electrophoretic mobilities. The bacterial genome encodes between 200 and 6000 proteins, which function enzymatically

or structurally [8]. When a bacterium is grown under carefully controlled conditions, this protein complement is essentially invariant. Polyacrylamide gel electrophoresis of the total cellular proteins provides a partial separation in which individual bands mostly represent several proteins. This complex pattern is reproducible and represents a fingerprint of the strain that can be used for comparative purposes [8]. Despite the system being rapid and accurate to differentiate species, inter-laboratory comparisons are not always reliable. Additionally, it has been of little value in identifying distantly related species and performing strain differentiation [7]. Improved automation is expected to improve the use of proteins and their enzymatic digests as species-specific biomarkers that are unique to various non-pathogens, distinguishing them from other pathogenic bacteria.

The backbone of cell walls, peptidoglycan, is present in almost all bacteria, but not in eukaryotic cells. This polysaccharide backbone contains repeating sugars, such as muramic acid, which are unique to bacteria and are definitive markers for the organism. Covalently attached to muramic acid are tetra- and pentapeptides composed of repeating L- and D-amino acids crosslinked by peptide bridges. Although variations in the sequence of peptides occur in side chains and crossbridges, D-amino acids are distinctive chemical markers from this region. D-amino acids are a major class of structural components that are not synthesized by mammals and therefore, are useful in bacterial typing. The fundamental polymer common to the walls of most bacteria is the peptidoglycan murien [1], [6]. It is a heteropolymer made up of glycan

strands crosslinked through short peptides. The glycan strands consist of *N*-acetylglucosamine and *N*-acetylmuramic acid. The latter provides a site for peptide chain attachment. Adjacent peptide chains in the peptidoglycan may be crosslinked by peptide bonding. The peptide chains always consist of alternating L- and D-amino acids and the most common sequence is L-alanine/D-glutamic acid/L-diamino acid [6]. In gram negative bacteria the chemical composition and structure of the peptidoglycan is rather uniform and therefore generally of little value as a character for classification, however gram-positive bacteria reveal great variation in the composition and structure of the peptide chain. About 100 different peptidoglycan types have been described [6]. Since the structure and composition of the peptidoglycans formed by different bacteria are generally stable and do not depend on cultivation conditions, the cell wall composition in many gram-positive bacteria is often a very useful chemotaxonomic marker.

There are a number of additional components that are closely associated with Gram-positive cell walls and affect its properties. Capsules, for example, are usually large polysaccharide molecules that envelop the cell wall. They are often of vital importance to the pathogenic potential of the organism [6]. If the capsule is lost, the organisms become sensitive to killing by phagocytes and are no longer virulent. Capsular polysaccharides are not covalently linked to the wall, but remain closely associated with the cells when harvested and washed. Although most of the gram-positive capsular polysaccharides are made of repeating units from more than one sugar residue, some bacteria possess homopolysaccharides. An example is dextran that assists *Streptococcus mutans*

in adhering to teeth, resulting in the development of cavities from the formation of dental plaque [6]. The microorganism responsible for anthrax, *Bacillus anthracis*, produces a capsule composed entirely of D-glutamic acid units [6]. In addition to helping organisms adhere to surfaces and protecting them from body defense mechanisms, capsules seem to prevent cells from desiccation in environments prone to periodic drying. Over 85 distinct capsular polysaccharides have been distinguished and can be used as a basis for typing of strains [6].

Lipids provide a wealth of taxonomic information that may be used for identification. Generally, lipids have been analyzed from whole cells providing a rapid means of characterization. Fatty acids have served as useful lipid biomarkers due to their abundance and ease of extraction from cells. Long chain fatty acids have commonly been released from the LPS of Gram-negative bacteria and analyzed by Gas Chromatography [9]. Hydroxylated fatty acids are common to almost all Gram-negative species while fatty acids with branched alkyl chains predominate in Gram-positive bacteria [6]. Cylcohexane and cyclobutane fatty acids are rare, specific markers for some acidophilic and thermophilic spore forming bacteria. Other groups differ in the quantity of unsaturated fatty acids formed [1].

Chemical analyses have become a major influence in bacterial taxonomy as new methods and techniques are developed. They have proved particularly successful for identifying organisms in which there are few distinguishing morphological and physiological characteristics.

Mass Spectrometry Applications

Approaches based on the application of instrument-based chemotaxonomic identification of microorganisms can provide differentiation, identification, and trace detection of microorganisms not otherwise easily obtained by traditional methods. Mass spectrometry was first applied more than 25 years ago as an analytical tool for the typing of microorganisms. It analyzes chemical markers for a broad set of organisms and from mixed populations. Although sample preparation and culturing are still required, an instrumental measurement is made in minutes.

Several ionization techniques have been applied directly to whole or lysed cells with mass spectrometric analysis. Initially, bacteria were grown in standardized conditions and thermally degraded in an inert atmosphere (pyrolysed, Py) to produce characteristic fingerprints of the cell [4]. Analysis of chemical profiles of the small pyrolysis products by mass spectrometry was accomplished in minutes. This process was automated and more reproducible than other instrumental methods (e.g. gas chromatography). Bacteria tended to produce the same peaks and quantitative data represented by peak heights were used to successfully identify various bacterial species.

In the 1980s, Fast Atom Bombardment mass spectrometry (FAB-MS) was used to analyze polar lipids for bacterial chemotaxonomy. It effectively separated mixtures of compounds without the need for any prior separative method such as gas chromatography. Members of the phospholipid series

phosphatidylserine, phosphatidylglycerol, and phosphatidylethanolamine were well seen as their anions [10]. Members of the enterobacteria have polar lipid fingerprints that distinguish them from each other and from bacteria of other genera. FAB-MS was able to differentiate between many species and compare closely related strains with similar polar lipid profiles.

Recently, Matrix-Assisted Laser Desorption/Ionization (MALDI) has been applied for the rapid typing of microorganisms. This technique is simple, allows sample analysis in seconds, and avoids time-consuming sample preparation. Most MALDI taxonomic studies have been based on the detection of proteins in the mass range up to two kilodaltons [10]. Whole cell suspensions [11] have been analyzed by MALDI-TOFMS to effectively identify intact Gram-negative and Gram positive microorganisms [12, 13], detect food-borne pathogens [14] , discriminate between pathogenic and non-pathogenic strains of bacteria [15], and detect human disease by analysis of tissues and bodily fluids. While a multitude of information exists on characterization of whole bacteria by mass spectrometry, little work has been done on bacterial spores.

Taxonomy and General Characteristics of *Bacillus* Spores

Bacteria of the genus *Bacillus* are aerobic, spore forming, gram-positive rod-shaped organisms. The genus is one of the most diverse and commercially useful groups of microorganisms [16]. Representatives of this group are widely distributed in soil, air, and water. Nutrient deprivation of *Bacillus* triggers several cellular responses that attempt to prevent starvation or to prepare the cell for a

long period when nutrients are not available [8]. If the cell fails to provide itself with the required nutrients, the cell then initiates another sequence of events resulting in the internal formation of a dormant spore. The spores are refractive and form within the vegetative cell before it lyses giving them a common name of endospore [1]. These structures are capable of remaining dormant for many years, but can germinate within minutes when suspended in a suitable rich environment. Although sporulation is a dramatic and morphologically identifiable stage in the life cycle of the *Bacillus*, it is considered a last resort response to a stressful situation of nutritional deprivation. Other bacteria respond to starvation by growing flagella for chemotaxis and derepressing genes for catabolic functions. These bacteria tend to live in less hostile environments than *Bacillus* which live in poor environments such as soil and dry grass. During this phase of the life cycle, spores are highly resistant to environmentally hazardous conditions such as solvents, extreme temperatures, and high pressure that kill non-sporulated bacteria [17]. Many species of the genus have been recognized; e.g. *B. subtilis, B. cereus*, the causative agent in food poisoning, and *B. anthracis,* which produces the anthrax toxin.

While the isolation of these spores is relatively easy, the identification of the *Bacillus* species in this stage can be difficult. Originally, microbiologists used the generic name *Bacillus* for several unrelated bacterial groups which were rod-shaped resulting in a profusion of species [16]. It was only from about 1920 that the diagnostic importance of bacterial endospore formation was generally accepted as a character for classification of bacteria. From then on, the genus

Bacillus was defined to include all endospore forming, rod-shaped, straight or nearly straight bacteria capable of growing aerobically [1]. Since the general recognition of the genus *Bacillus*, many new species have been described. However, these descriptions were often poor and strain variations were not fully appreciated, it was not until the 1940s that comparative studies were attempted to rightfully place an organism into an existing genus or species [1]. By studying morphological, physiological, and biochemical properties of over a thousand strains with nearly two hundred species names, most of the strains were assigned to just nineteen species.

Current Research

Spores contain distinct markers consisting of both external and internal components such as cell wall lipids, carbohydrates, nucleic acids, proteins, and secondary metabolites. Previously, Fenselau and coworkers used MALDI-Time of Flight Mass Spectrometry (MALDI-TOF MS) to obtain chemotaxonomic information of both intact and lysed *Bacillus* spores [18] [19]. These investigations focused on protein biomarkers between closely related *Bacillus* species, including pathogens such as *B. anthracis* and *B. cereus*. A group of external biomarkers can be detected without any treatment, while a larger set of high-mass biomarkers, located inside the spore, are released by chemical or mechanical treatment and show masses characteristic of proteins from the *Bacillus* proteome [20].

To date, a considerable amount of research has been completed on the application of MALDI-TOF MS to identify protein biomarkers in bacteria [21]. However, few studies have applied this technique to the analysis of secondary metabolite biomarkers unique to the various *Bacillus* species. These characteristic molecules, such as cyclic antimicrobial peptides and lipopeptides, generally have molecular masses from 800 to 5000 Da and have been used to distinguish between several *B. subtilis* strains [22]. Lipopeptides are heterogeneous molecules synthesized by several *Bacillus* species, are generally cyclic, and exist as mixtures of homologous series that differ in the length of their fatty acid chains. Surfactins, for example consist of a β-hydroxyl fatty acid that is linked to the N-terminal amino acid of a heptapeptide by an amide bond. The carboxyl end of the peptide is further linked to the β-hydroxyl group of the fatty acid via a lactone bond. Iturins have similar structures, but contain a β-amino fatty acid instead of a β-hydroxy fatty acid, and contain different amino acids [23] [24]. Recently, Hathout applied mass spectrometric techniques to discover a novel lipopeptide biomarker secreted by *B. thuringiensis kurstaki* [25]. Other cyclic lipopeptides have been characterized with different peptide ring sizes and fatty acid chain lengths . These compounds are being exploited for their potential as antibiotics, as biosurfactants in oil recovery, and as emulsifiers in detergents.

In this work, the identification and structural characterization of a lipopeptide characteristic of *B. globigii* is reported. It is visible in the mass spectra of vegetative cells and spores of *B. globigii (Bacillus subtilis var. niger)*. This microorganism is non-pathogenic and is commonly used as a simulant in

biological studies focused on the detection and identification of pathogenic bacteria. The biological warfare (BW) agents of greatest concern are those that are deployable as aerosols. Experts generally agree that *B. anthracis* spores, which are highly pathogenic and require medical treatment within 48 hours, are one of a growing list of infectious agents requiring rapid detection. *B. globigii* spores have principally been used in the development of techniques for rapid microbial identification. Similarly, this microorganism was chosen for this study as a suitable stand-in for other pathogenic bacteria.

EXPERIMENTAL

Microorganisms and Culture Conditions

Lyophilized *Bacillus globigii (Bacillus subtilis var. niger)* cells were received from U.S. Army Laboratories at Dugway Proving Ground, Utah, USA and were grown in-house according to standard procedures [16]. Colonies were cultured overnight on sheep's blood agar plates and then grown in New Sporulation Medium (NSM) agar. Plates were incubated at 37°C for 48 hours followed by 24 hours incubation at room temperature. Plates were checked under a phase contrast microscope to verify the presence of spores within the vegetative cells. The cells were harvested from the NSM plates by scraping, using a spreader, and then combined and placed in 3 – 5 mL of filtered distilled water. Vegetative cells were permitted to lyse at room temperature for the next 24 to 72 hours and were monitored to ensure their disappearance. The remaining spores were suspended in 40 mL of sterile filtered water and then recovered by centrifuging three to four times at 10000*g* for 10 minutes. The pellets were resuspended in 10 mL of sterile distilled water, lyophilized, and stored at -20° C. An aliquot of suspended spores was checked under a phase contrast microscope to ensure vegetative cells and cellular debris were successfully removed during centrifugation. The cell-free supernatants (typically 40 -50 mL of material) from each washing were aspirated, pooled, and stored at 20°C for analysis.

Germinated *B. globigii* spores were also prepared. 5 mg of lyophilized spores were suspended in 5 mL of NSM media and left to germinate at room temperature. Aliquots of the suspension were taken every 30 minutes and checked under a phase contrast microscope for loss of refraction, indicating germination was occurring. Maximum germination was achieved between two and three hours.

A single colony of *B. globigii* was cultured on a Brain Heart Infusion (BHI) agar plate overnight at 37°C for analysis of vegetative cells.

Mass Spectrometry

In general, molecular masses of crude samples and purified fractions were routinely checked for the presence of the compound of interest.

All samples were analyzed in the positive ion mode using a Kompact MALDI 4 time-of-flight (TOF) instrument (Kratos Analytical Instruments, Chestnut Ridge, NY) equipped with pulsed extraction and a nitrogen laser operating at 337 nm. A 50 mM solution of 4-methoxycinnamic acid (4-MCA), in acetonitrile/water/trifluoroacetic acid (TFA) (70:30:0.1, v/v/v), was used as matrix. The analyte solution (0.4 μL) was deposited on a stainless steel sample slide, covered with an equal volume of matrix solution, and allowed to air dry. Mass spectra were accumulated over 50 individual laser shots and were obtained in the reflectron mode at an initial accelerating voltage of 20 kV. Cytochrome C (0.1 μL) was used as an internal calibrant.

Accurate mass measurements were obtained using MALDI on a Fourier Transform mass spectrometer (FTMS) (IonSpec, Irvine, CA) equipped with a 4.7 Tesla superconducting magnet and a 337 nm laser. The nitrogen laser fluence was estimated to be 40 mJ/cm^2. The mass accuracy was better than 20 ppm. A 50 mM matrix solution of 4-MCA in acetonitrile/water/trifluoroacetic acid (TFA) (70:30:0.1, v/v/v) was used as a matrix. One microliter of the analyte solution was deposited on a stainless steel sample probe between two equal volumes of matrix solution (0.5 μL) and allowed to air dry. Mass spectra were averaged over 20 individual laser shots.

Tandem mass spectrometry (MS/MS) experiments were performed on an Applied Biosystems (Foster City, CA) Q-Star with a Protana A/s (Odense, Denmark) nanospray source. All experiments were performed using nanoelectrospray disposable needles (Protana). The sample was dried and dissolved in a solvent solution of methanol/water/acetic acid, 50:50:2, v/v/v). Four microliters of this mixture was loaded into the needle for analysis. The voltage applied to this needle was 0.9 kV. The curtain gas flow rate was 1.3 L/min.

Screening for Lipopeptide-Producing Microorganisms

All colonies of *B. globigii* (germinated and dormant spores grown on NSM plates and vegetative cells grown in BHI) were harvested using a sterile loop and suspended in 100 μL of acetonitrile/water/trifluoroacetic acid (TFA) (70:30:0.1, v/v/v). Each suspension (0.4 μL) was deposited in a well of a MALDI sample

slide and covered with an equal volume of 4-MCA. Germinated spores and vegetative cells were analyzed immediately after preparation, while dormant spores were kept lyophilized and at –20°C until analyzed. The cell-free culture supernatant (0.3 μL) was deposited on a sample slide and covered with 4-MCA as above.

Isolation and Purification of Lipopeptides

Three methods proved viable for extraction of lipopeptides with little noticeable difference in the resulting purified samples. These involved solid phase extraction, Folch extraction, and surface washing. Typically, both solid phase extraction and Folch extraction methods were used together to maximize the total amount of product recovered from each culture.

Solid Phase Extraction

Two to three milliliters of culture supernatant was applied to a disposable Bakerbond spe cartridge column (J.T. Baker) packed with 500 mg of C_{18} silica. The cartridge was equilibrated successively with 5 mL of methanol and 5 mL of 0.1% trifluoroacetic acid (TFA). The cartridge was washed with 10 mL of water, 20 mL of 45% aqueous methanol, and finally with 10 mL of methanol [26]. The eluate was evaporated in a vacuum centrifuge and the residue was dissolved in 1 mL of methanol. The sample (0.4 μL) was analyzed directly by MALDI-TOF MS to confirm the elution of the lipopeptide. This purification scheme is shown in Figure 1.

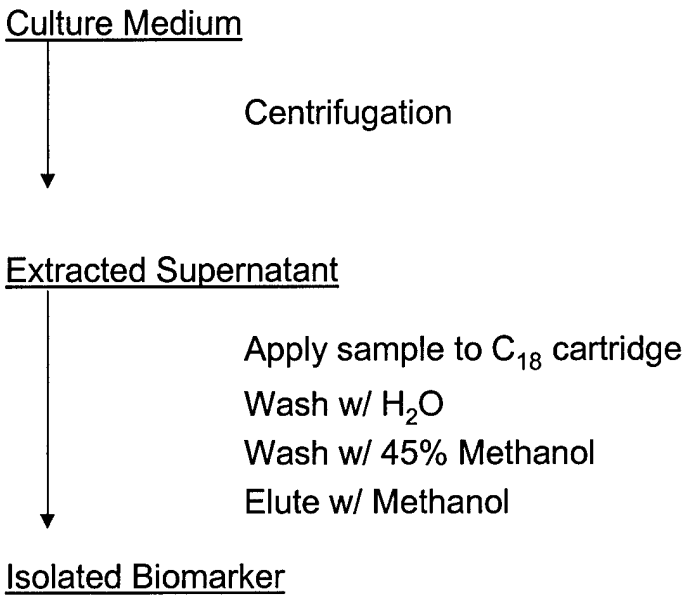


Figure 1. Solid Phase Extraction Purification Scheme

Folch Extraction

Lyophilized spores (50 – 100 mg) were added to a 2:1 chloroform-methanol (v/v) mixture in a glass test tube. The crude extract was mixed with 20% its volume of water, thoroughly vortexed, and allowed to separate into two phases by standing followed by mild centrifugation at 3000*g* for five minutes [27]. The phases were separated, collected in glass vials, and 0.4 μL of each was analyzed directly by MALDI-TOF MS to determine its components.

Solvent Washing

Dry spores (50 to 100 mg) were suspended in acetonitrile/water (50:50, v/v) and vigorously mixed using a vortex, followed by mild centrifugation at 3000*g*

for five minutes. The supernatant, containing components washed from the spore surface, was extracted and dried using a vacuum centrifuge. The dry extract was dissolved in 5 mL of acetonitrile/water (20:80, v/v) and applied to a cartridge column packed with 500 mg of C_{18} silica. The cartridge was rinsed and the retained components were eluted as described above. The eluate was evaporated in a vacuum centrifuge, dissolved in methanol, and analyzed by MALDI-TOF MS.

HPLC Purification

Instrumentation and Chemicals

A Shimadzu (Columbia, MD) HPLC system, consisting of two LC-600 delivery pumps, an SPD-6A UV Spectrophotometric detector, and a C-R6 A Chromatopac recorder were used in all experiments. Acetonitrile was HPLC grade from Fischer Scientific (Fair Lawn, NJ) and trifluoroacetic acid (TFA) was from Aldrich Chemicals (Milwaukee, WI).

Reversed-Phase Chromatography

Separations were performed on an analytical Aquapore 250 x 4.6 mm RP-300 C_8 column (Applied Biosystems) with a 10 μm particle size. The mobile phase components were (A) 0.1% TFA in water and (B) acetonitrile/water/trifluoroacetic acid (TFA) (80:20:0.1, v/v/v). Lipopeptides were detected at 215 nm and eluted by a linear gradient of solvent B, developed from

30% - 100% in 32 minutes at a flow rate of 0.8 mL/minute. Peaks with different retention times were collected, analyzed by MALDI-TOF MS, and freeze-dried.

Structural Analysis of Lipopeptides

A series of chemical reactions were performed with the purified lipopeptides. The reaction products were characterized by MALDI and nanospray tandem mass spectrometry (MS/MS) analysis.

Determination of the Lactone Bond

To determine the lactone linkage, the lipopeptide was dissolved in 1 M potassium hydroxide (KOH) and reacted overnight at room temperature. The excess KOH was removed by passing the solution through a gel-loader tip packed with Poros 50 R2 (PerSeptive Biosystems, Framingham, MA) [28]. The trapped sample was washed twice with 20 μL of 0.1% TFA, then eluted from the cartridge with 20 μL of methanol. The eluate was partially evaporated from the sample in a vacuum centrifuge. To the sample, 20 μL of filtered distilled water and acetic acid up to a concentration of 2% were added, and then subjected to MS/MS analysis.

Esterification

The open form of the purified lipopeptide was subjected to esterification of its carboxylic acid groups. After the sample was dissolved in 1 M KOH, hydrolyzed, and cleaned as described above, it was freeze dried. Methanolic

HCl was made by the dropwise addition of 160 μL of acetyl chloride to 1 mL of methanol while stirring. One hundred microliters of this solution was added and the reaction carried out for one hour at room temperature. The sample was dried in a vacuum centrifuge and suspended in acetonitrile/water/trifluoroacetic acid (TFA) (70:30:0.1, v/v/v). An aliquot of the reaction product was analyzed by MALDI-TOF MS.

Fatty Acid Analysis

About 10 mg of the purified lipopeptide was hydrolyzed with 200 μL of 6 M HCl at 150°C for 18 hours. After the solution was cooled, 200 μL of filtered distilled water was added and the fatty acids were extracted with three 1 mL additions of ether. The ether extracts were combined and evaporated to dryness under a stream of nitrogen. The acid was converted into its acyl chloride by reaction with 100 μL of thionyl chloride (Acros Chemicals, NJ) for one min at room temperature. The thionyl chloride was evaporated with a stream of nitrogen and the residue was reacted for one minute at room temperature with 20 μL of a 10% solution of 3-pyridylcarbinol (Aldrich Chemical Co.) in acetonitrile. Analysis of the picolinyl esters was performed on a Shimadzu gas chromatography/mass spectrometer (Columbia, MD). About 4 μL of sample was injected into the gas chromatograph equipped with a 25 m X 0.25 mm DB5 capillary column (J & W Scientific, Folsom, CA). The temperature of the injector was set at 250°C and the helium carrier gas flow was set at 40 mL/min. The column temperature was programmed from 180°C to 280°C at 10°C per minute.

RESULTS

Microorganisms and Culture Conditions

The progress of *B. globigii* cells through sporulation was marked by several characteristic morphological and physical changes. Observation of the suspension under a phase contrast microscope distinguished some of these changes [5]. Early in sporulation, the rod-shaped vegetative cells lost motility and began to darken due to the internal development of a protoplast. This prespore became phase dark indicating that sporulation was occurring. The spores finally became phase bright with no surrounding mother cells, due to their lysis.

Isolation of Compounds from the Culture Supernatant of *B. globigii* Cultures

In the present study, *B. globigii* was found to produce a biomarker centered at a peak with a molecular mass of 1478 Da, when analyzed by MALDI-TOF MS. Significantly, the biomarker was present in samples obtained from vegetative cells, dormant spores, and germinating spores. Analysis of the cell-free supernatant, extracted from the culture medium, also showed the presence of the biomarker. Thus it is unique to the *B. globigii* strain in all stages of its life cycle. The MALDI-TOF MS spectrum of each sample is shown in Figures 2 - 5.

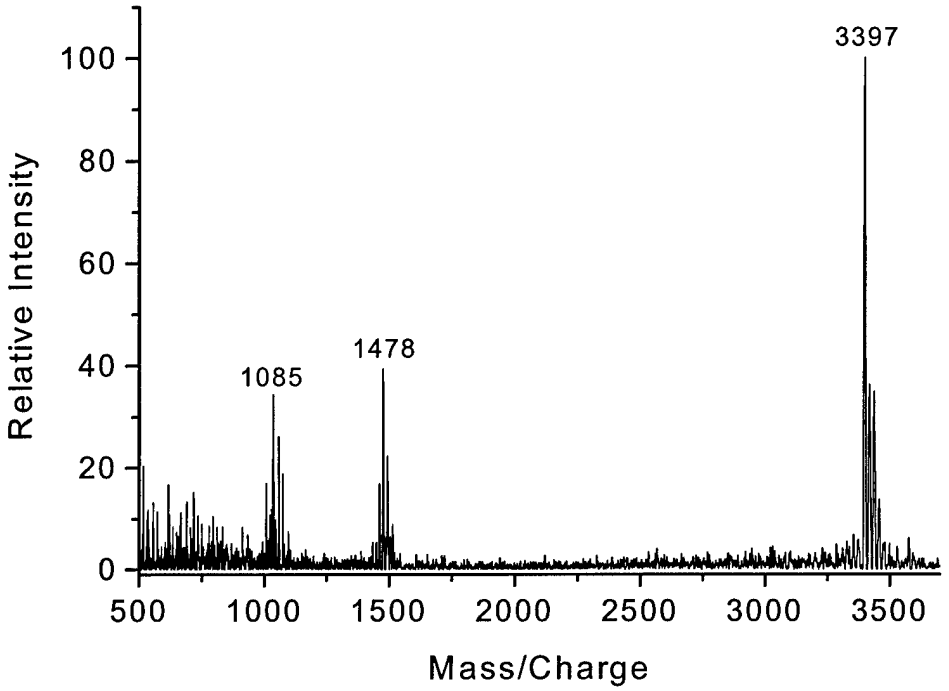


Figure 2. MALDI-TOF MS of Vegetative Cells from *B. globigii*

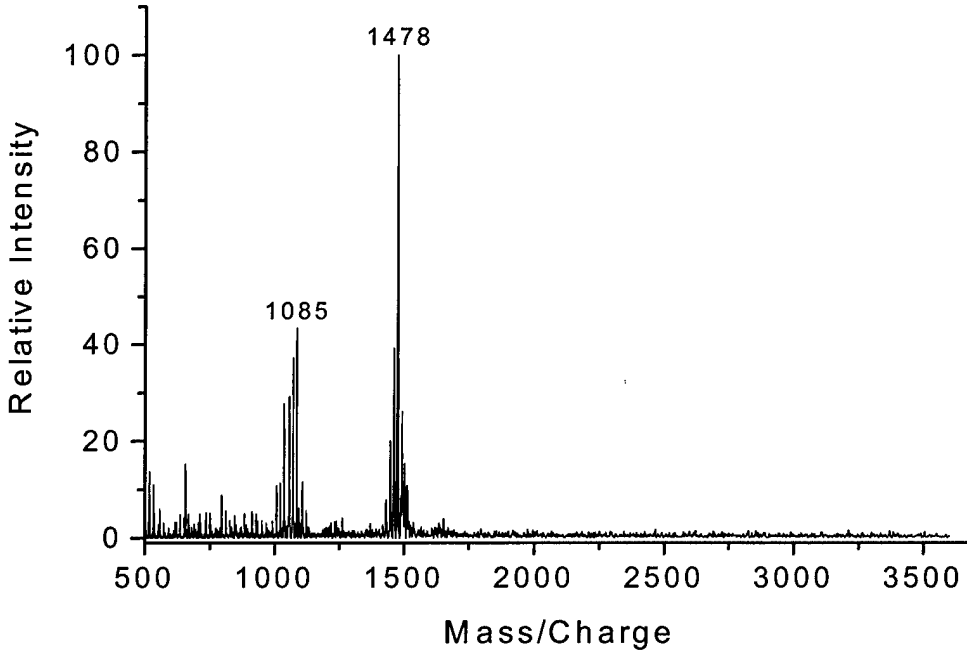


Figure 3. MALDI-TOF MS of Whole Spores from *B. globigii*

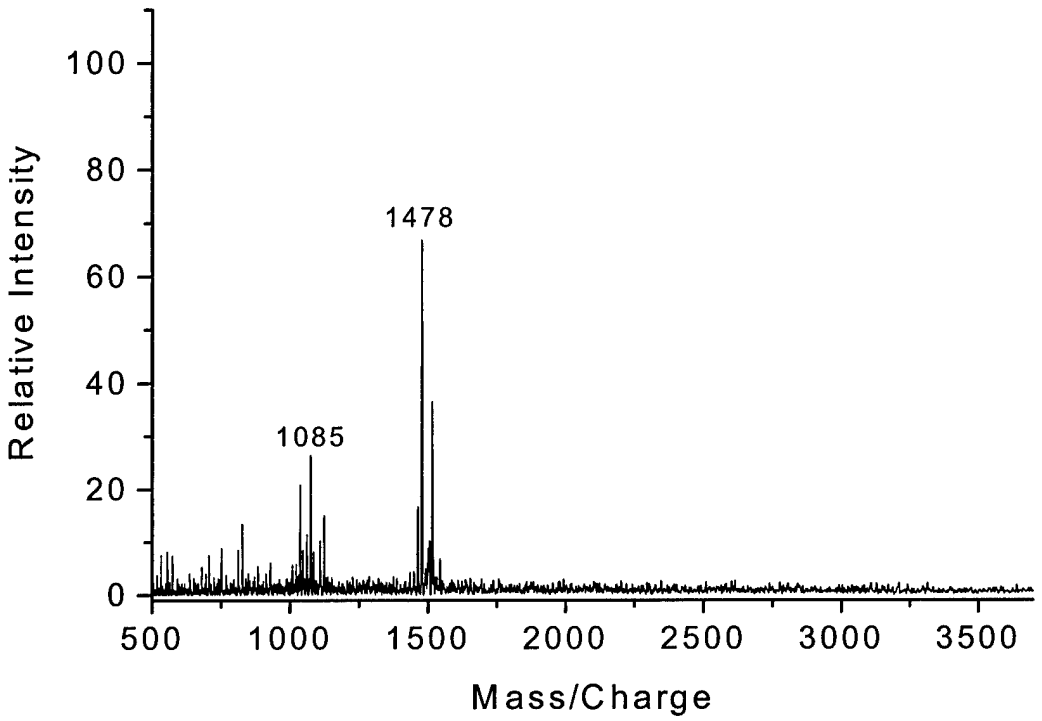


Figure 4. MALDI-TOF MS of Germinated Spores from *B. globigii*

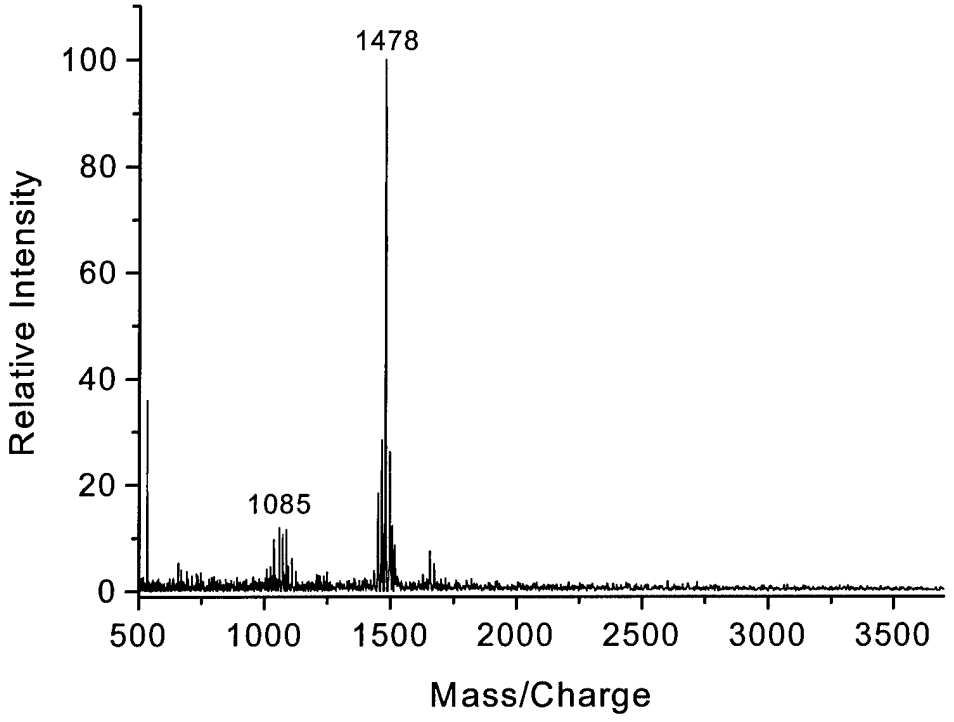


Figure 5. MALDI-TOF MS of Cell-Free Culture Supernatant

This biomarker was also produced in different growth media and showed the highest intensity in spores when grown in NSM. The biomarker was isolated from the cell-free culture supernatant of continuously growing cells as described in the experimental chapter. The stepwise purification of the raw material to the crude product, using Bakerbond C_{18} solid phase extraction cartridges, resulted in enrichment of the compounds. Mass spectral analysis of the water and the aqueous methanol fractions did not show the presence of the biomarkers. The biomarker cluster began to elute when the cartridge was washed with aqueous methanol of concentrations greater than 50% and completely eluted with methanol. Methanol eluted fractions were dried and analyzed by MALDI –TOF MS confirming the presence of two biomarker clusters, centered at m/z 1085 and m/z 1478. The mass spectrum of a methanol eluted fraction is shown in Figure 6.

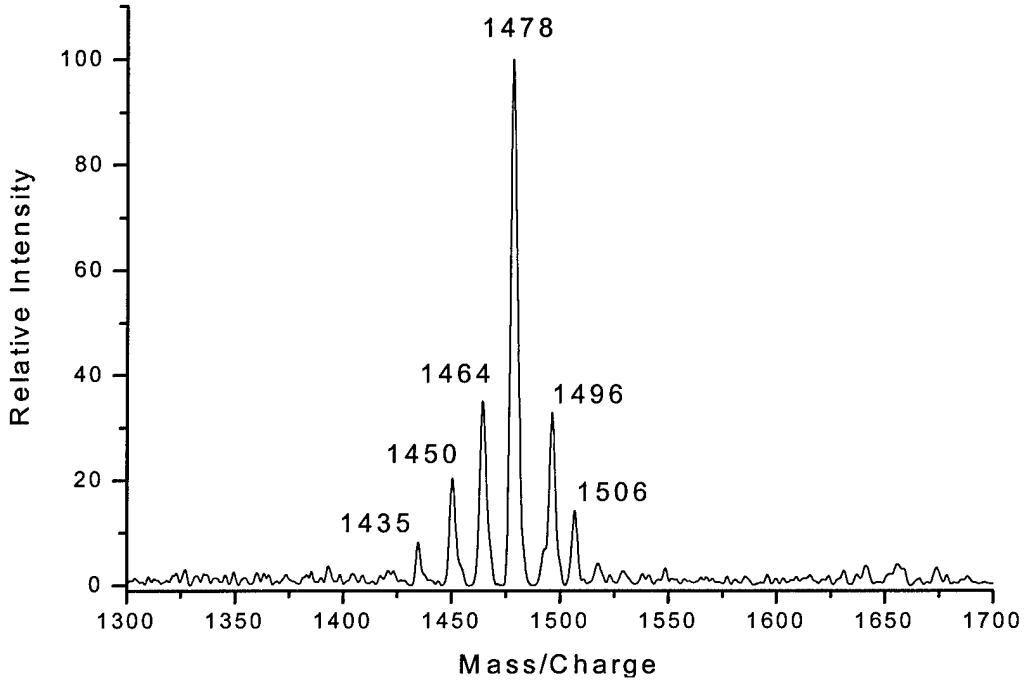


Figure 6. MALDI-TOF MS of Methanol Eluted Fraction from C_{18} cartridge

The mass spectrum shows a cluster containing six molecules centered around an intense peak at m/z 1478. Three peaks, with molecular masses of 1450, 1464, and 1478 Da, differed by 14 Da, suggesting a series of homologous molecules. The molecule at m/z 1506 is 28 Da above the last peak in this series and is possibly a fourth homologous molecule. The peak at m/z 1435 was consistently 15 Da from the peak at m/z 1450. A sixth peak at m/z 1496 is 18 Da higher than the peak at 1478 Da. The accurate mass of the cluster was measured by a Fourier Transform Mass Spectrometer (FTMS). The monoisotopic protonated ($[M+H]^+$) masses were determined to be 1449.8, 1463.8, 1477.8, 1495.8, and 1505.8 m/z. The FTMS spectrum of the molecule weighing 1477.8 is shown in Figure 7.

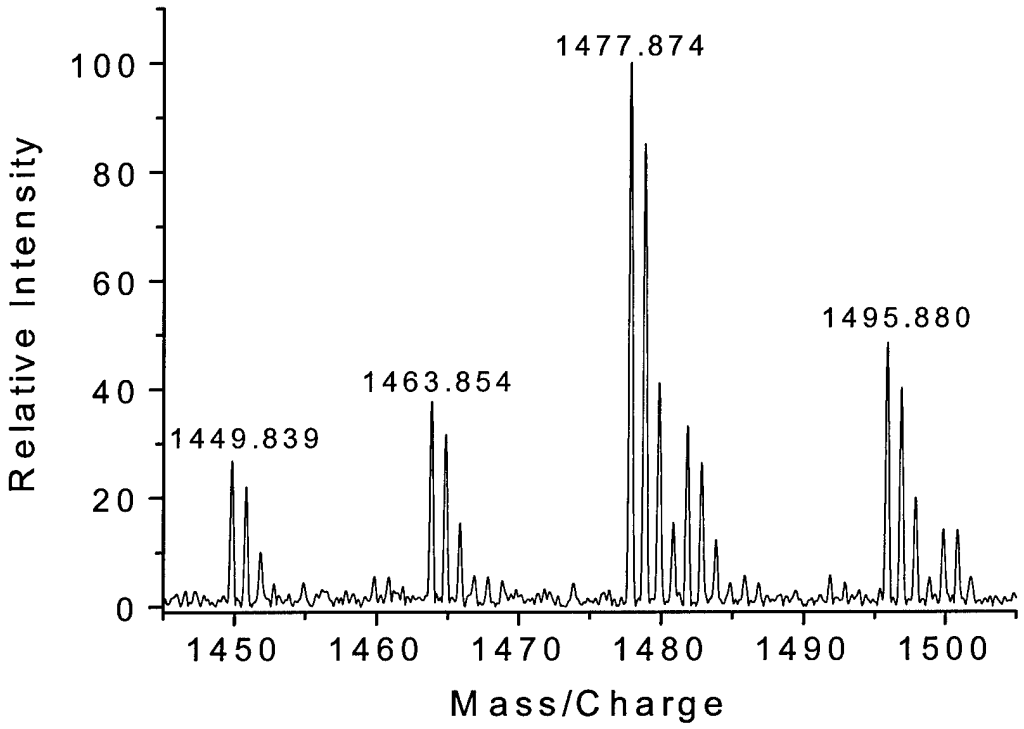


Figure 7. FTMS Spectrum of Methanol Eluted Fraction from C_{18} cartridge

The FTMS spectrum confirmed the initial analysis of the biomarker using MALDI-TOF MS. Based on the mass difference pattern and amphiphilic nature of the biomarker, it is suggested to be a family of closely related lipopeptides that vary in their fatty acid chain length.

After the biomarker was extracted from the culture supernatant of *B. globigii,* the individual lipopeptides were purified and separated by reverse-phase HPLC, shown in Figure 8. HPLC purification provided purified samples for further analysis. About ten peaks were collected representing individual lipopeptides from the two observed biomarkers clusters. The lipopeptides, from the cluster of interest, eluted between 65 and 70% organic solvent B. All peaks were collected and the molecular mass of each was determined by MALDI-TOF MS. Five of the six peaks observed in the C_{18} pre-purified sample were detected by MALDI analysis of the HPLC fractions. Fraction one contained a mixture of three compounds that showed peaks at m/z 1450, m/z 1472, and m/z 1482. Fractions two through five contained lipopeptides, that eluted as single peaks, at m/z 1464, 1478, 1496, and 1506 respectively. Fraction three ($[M+H]^+$ at m/z 1478), the main component of the crude product and of greatest relative intensity in the mass spectrum, was the most abundant of all the eluted fractions. The MALDI mass spectra of these purified fractions are shown in Figures 9, 10 and 11. The cluster observed at m/z 1085 Da, eluted as separate components prior to the fractions noted above.

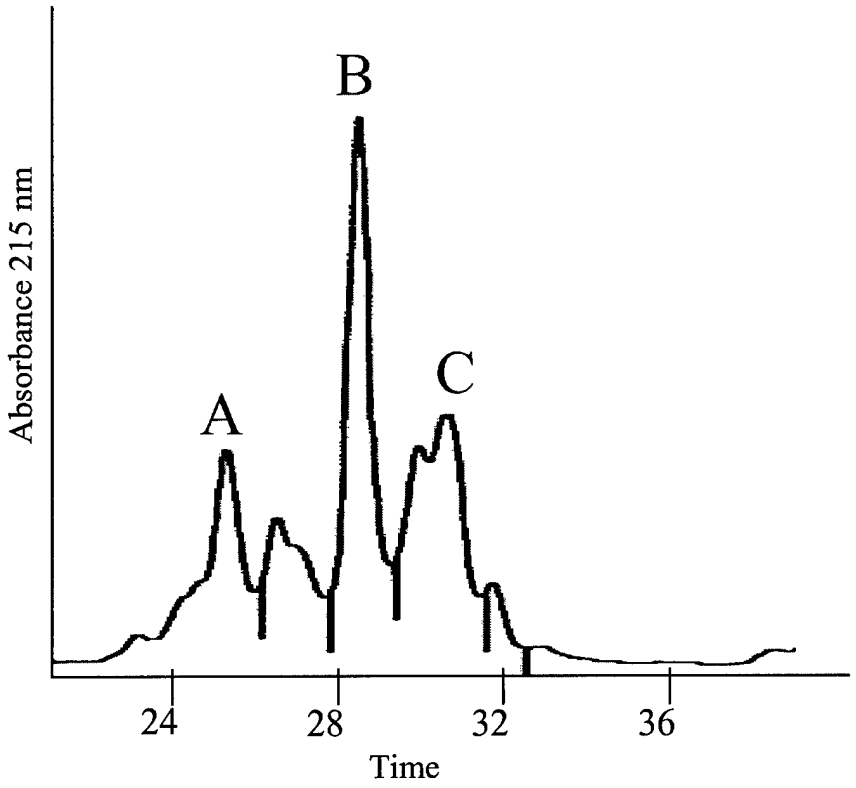


Figure 8. Purification of Lipopeptides by RP-HPLC.

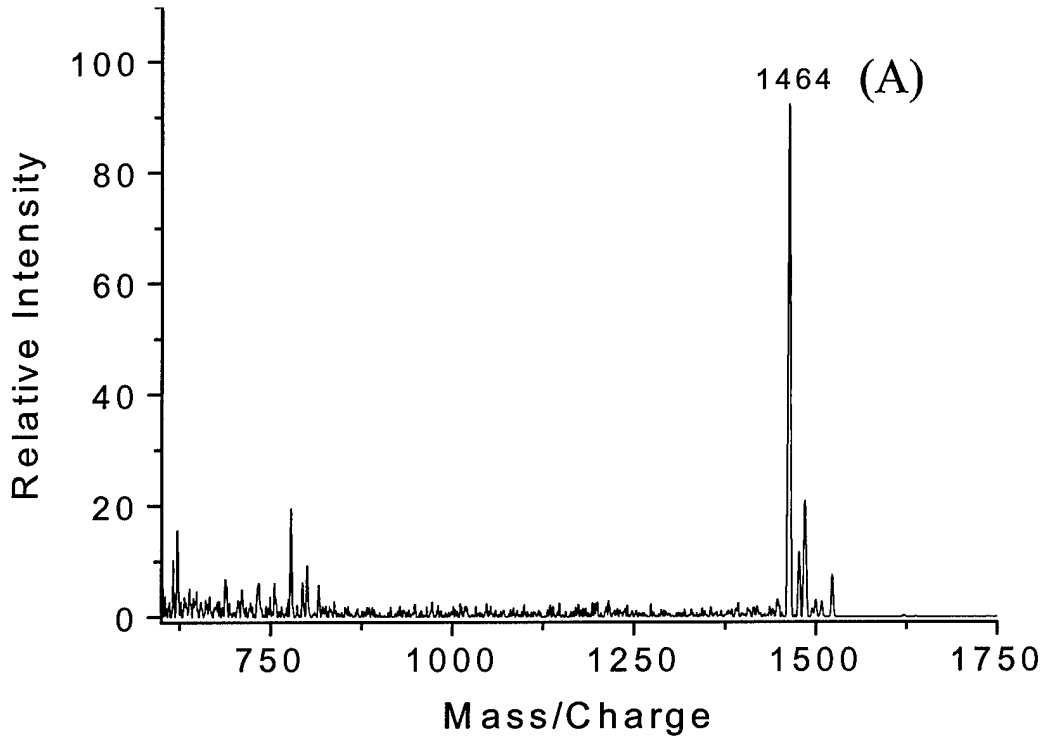


Figure 9. MALDI-TOF MS of the Purified Lipopeptide $[M+H]^+$ 1464

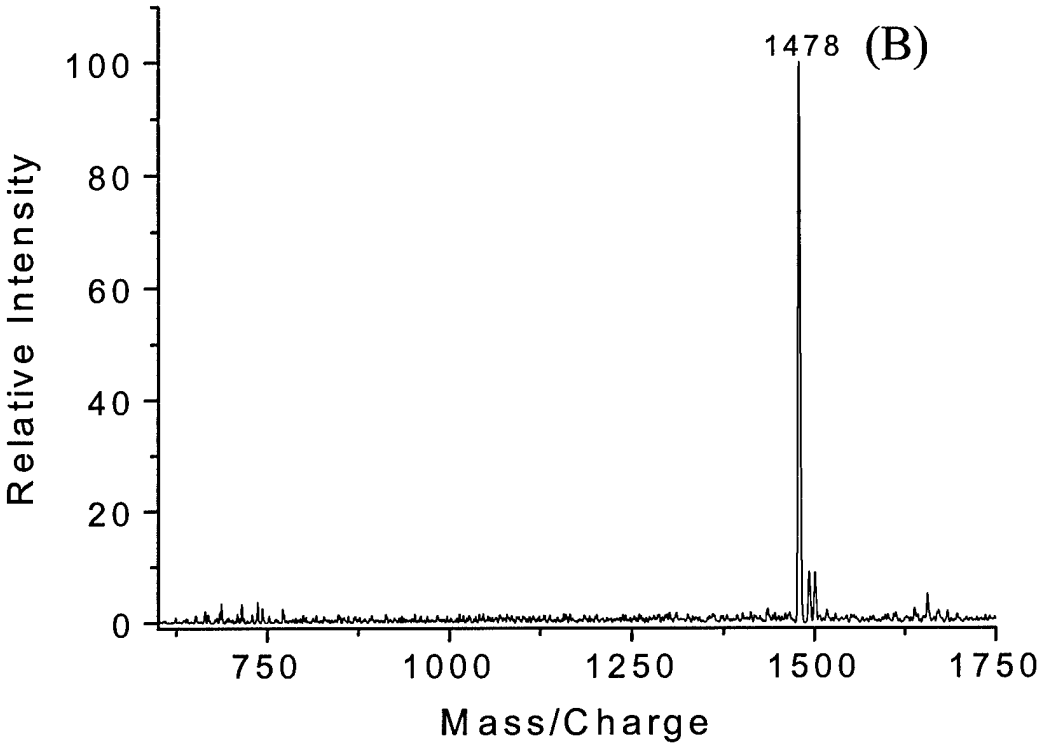


Figure 10. MALDI-TOF MS of the Purified Lipopeptide $[M+H]^+$ 1478

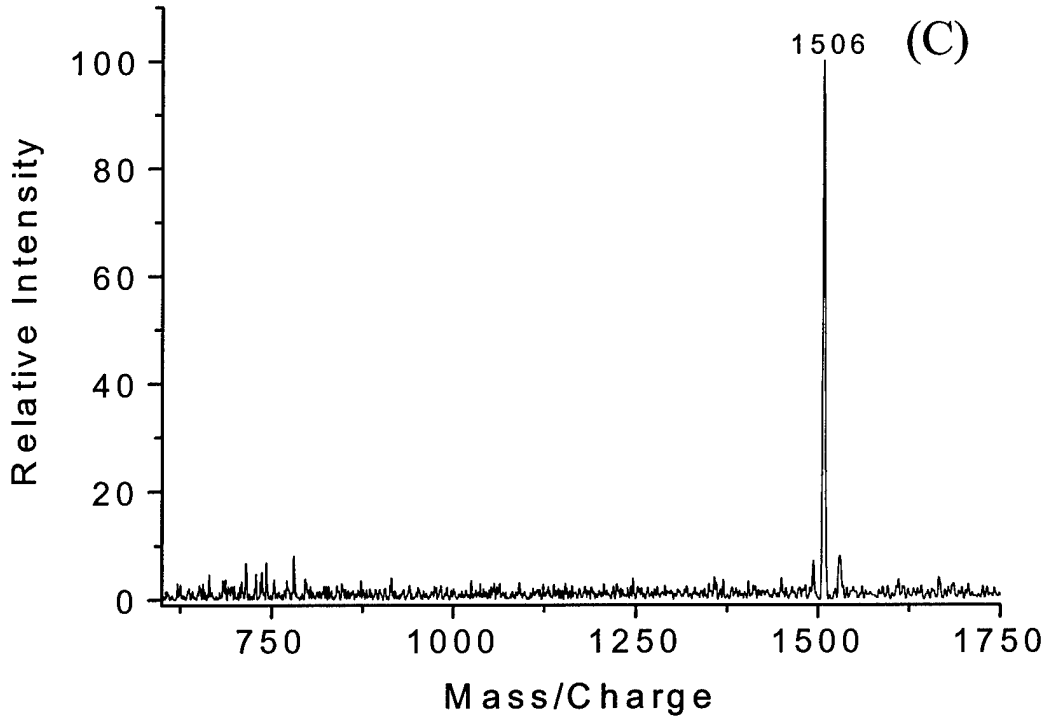


Figure 11. MALDI-TOF MS of the Purified Lipopeptide $[M+H]^+$ 1506

The purified lipopeptides weighing 1464, 1478, and 1506 Da were subjected to a mild alkaline hydrolysis that does not cleave amide bonds, using 1 M KOH for 1 hour at room temperature. Each gave new products with $[M+H]^+$ masses of 1482, 1496, and 1524 Da respectfully. A mass gain of 18 Da was observed for each molecule and was tentatively assigned to hydrolysis of a lactone. The presence of a lactone confirmed that the lipopeptides contained a cyclic peptide domain analogous to previously characterized microbial lipopeptides. Spectra of two of the opened lipopeptides are shown in Figures 12 and 13.

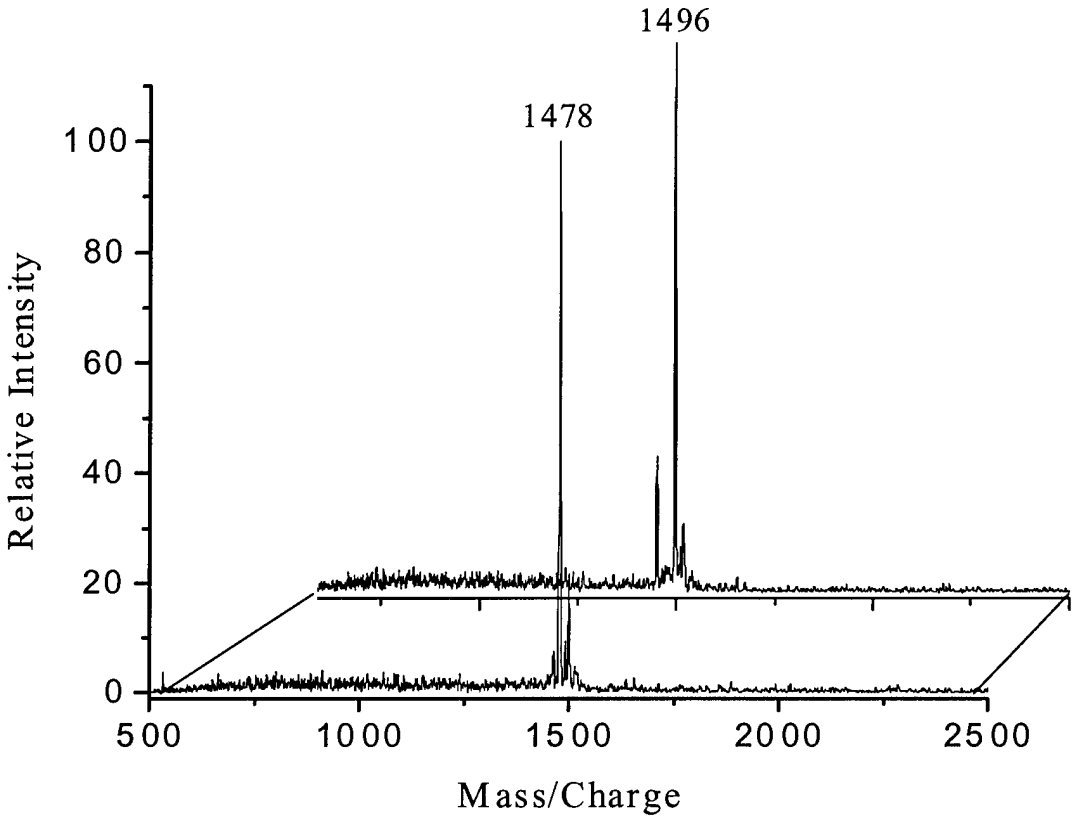


Figure 12. Hydrolyzed Product from Purified Lipopeptide $[M+H]^+$ 1478

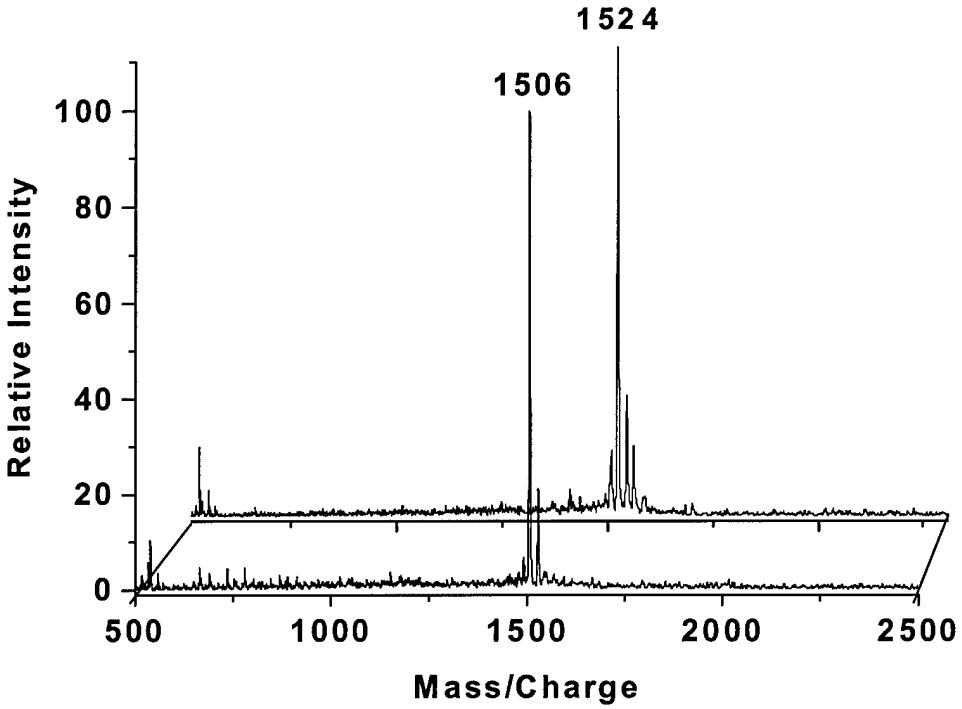


Figure 13. Hydrolyzed Product from Purified Lipopeptide $[M+H]^+$ 1506

The open form of the lipopeptide ($[M+H]^+$at m/z 1496) was esterified as described in the Experimental Section and its mass checked by MALDI-TOF MS. The resulting product was found to have an $[M+H]^+$ ion peak at m/z 1552, 56 mass units higher than the starting material. This mass gain identifies the presence of four carboxyl groups in the peptide domain that were converted to methyl esters. One carboxyl group was assigned to the C-terminal position of the peptide. The remaining carboxyl groups were assigned to amino acids possessing a carboxyl side chain, suggesting the presence of three glutamic acid or aspartic acid residues. The spectrum of a hydrolyzed lipopeptide with its converted methyl esters is shown in Figure 14.

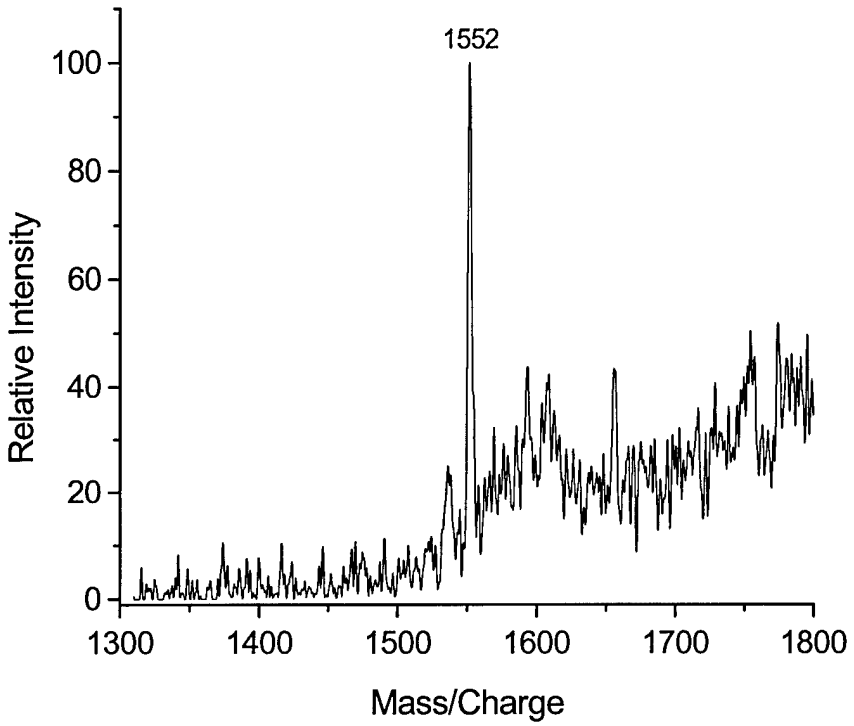


Figure 14. MALDI-TOF MS of the Esterified Lipopeptide $[M+H]^+$ 1552

To determine the amino acid sequence, a collision-induced dissociation (CID) method was applied to the hydrolyzed lipopeptides weighing 1482, 1496, and 1524 Da. A CID mass spectrum of the hydrolyzed lipopeptide, $[M+H]^+$ 1482.8, showed a doubly charged $[M+2H]^{2+}$ precursor ion at 741.9. This ion was selected for CID to yield the fragment ions indicated in Figure 15. The fragment ion peak at m/z 1351.6 is explained by the loss of 131.2 Da, isoleucine or leucine and water, from the $[M+H]^+$ ion peak at m/z 1482.8. The fragment ion peak at m/z 1188.6 is formed by the loss of 163 Da from the peak at m/z 1351.6. This loss corresponds to tyrosine. The fragment ion peak at m/z 962.5 is formed by the loss of 226.1 Da from the peak at m/z 1188.6, corresponding to glutamic acid

and proline. Neither the loss of glutamic acid or proline alone was observed. The fragment ion peak at m/z 891.6 is formed by the loss of 71 Da from the peak at m/z 962.6, corresponding to alanine. The fragment ion peak at m/z 762.4 is formed by the loss of 129.2 Da from the peak at m/z 891.6, corresponding to glutamic acid. The fragment ion peak at m/z 661.4 is formed by the loss of 101 Da from the peak at m/z 762.4, corresponding to threonine. The fragment ion peak at m/z 498.3 if formed by the loss of 163.1 Da from the peak at m/z 661.4, corresponding to tyrosine. The peak at m/z 384.3 reflects the loss of 114 Da from the peak at m/z 498.3, corresponding to ornithine.

A CID mass spectrum of the hydrolyzed lipopeptide, $[M+H]^+$ 1496.8, showed a doubly charged $[M+2H]^{2+}$ precursor ion at m/z 748.9. This ion was selected for CID to yield the fragment ions indicated in Figure 16. The fragment ion at m/z 1365.5 is explained by the loss of 131.3 Da, isoleucine or leucine and water, from the $[M+H]^+$ ion peak at m/z 1496.8. The fragment ion peak at m/z 1202.6 is formed by the loss of 162.9 Da from the peak at m/z 1365.5, corresponding to tyrosine. The fragment ion peak at m/z 976.6 is formed by the loss of 226 Da from the peak at m/z 1202.6. This loss corresponds to glutamic acid and proline residues. Again, the loss of either a glutamic acid or a proline fragment alone was not observed. The fragment ion peak at m/z 905.6 is formed by the loss of 71 Da from the peak at m/z 976.6, corresponding to alanine. The fragment ion peak at m/z 776.4 is formed by the loss of 129.2 Da from the peak at m/z 905.6, corresponding to glutamic acid. The fragment ion peak at m/z 675.4 results from the loss of 101.1 Da from the peak at m/z 776.5, corresponding to

threonine. The fragment ion peak at m/z 512.4 is formed by the loss of 163 Da from the peak at m/z 675.4 and this loss corresponds to tyrosine. The fragment ion peak at m/z 398.3 is formed by the loss of 114.1 Da from the peak at m/z 512.4, corresponding to ornithine.

Comparison of the two spectra showed that the lipopeptides weighing 1464 and 1478 Da contained the same amino acids in the same sequence. Neither spectra revealed any fragments below 384.3 and 398.3 Da respectively. To summarize, the partial amino acid sequence is composed of nine residues; isoleucine or leucine, two tyrosine residues, proline, two glutamic acid residues, alanine, threonine, and ornithine. These fragmentation patterns due not reflect the stereochemistry of the individual amino acids in the peptide.

The CID method shows sequence b_n and y_n series ions from the N and C termini of the molecule respectively. Figures 17 and 18 show the assignment of b_n and y_n series ions to the CID spectra of $[M+2H]^{2+}$ 741.9 and 748.9. The amino acid residues corresponding to the b_n and yn series ions are indicated in Figure 19. A typical fragmentation pattern, of b and y series ions is explained below in the Discussion Section.

The hydrolyzed lipopeptide with a mass ($[M+H]^{+}$) of 1524 Da was also selected for CID. Based on its mass it appeared to be another homologue of the above sequenced lipopeptides. The CID mass spectrum of this lipopeptide contained a doubly charged $[M+2H]^{2+}$ precursor ion at 762.4. This ion was selected for CID to yield the fragment ions indicated in Figure 20. The fragment ion peak at m/z 1392.6 is explained by the loss of 131.1 Da, corresponding to

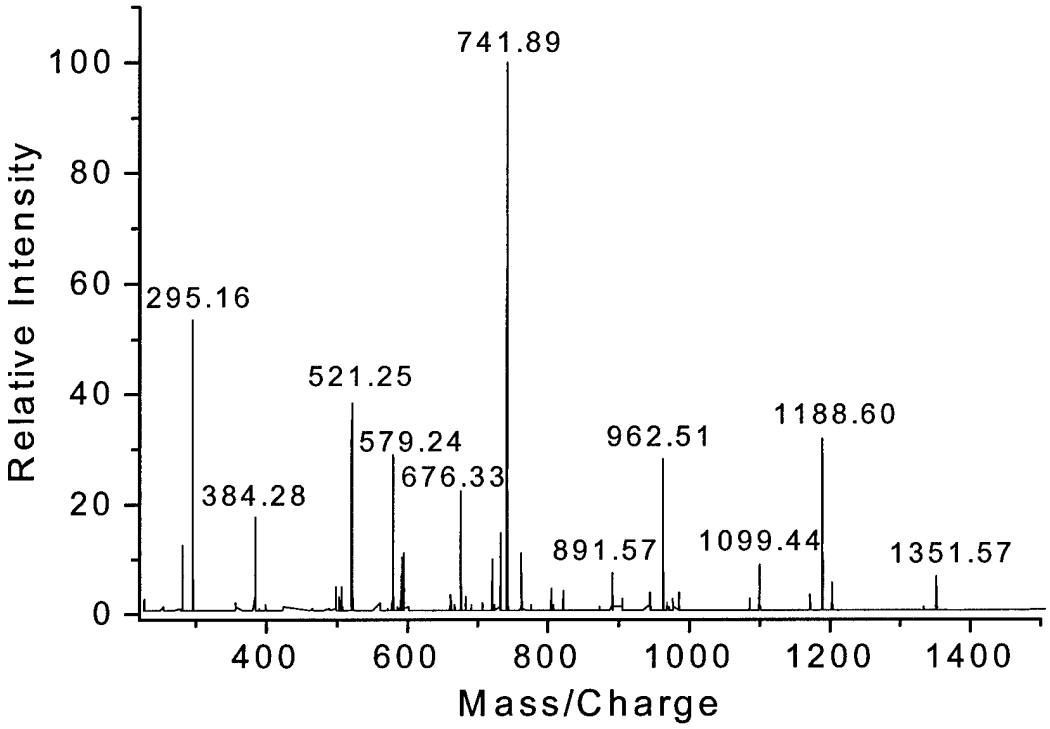


Figure 15. Product Ion Scan of $[M+2H]^{2+}$ 741.89

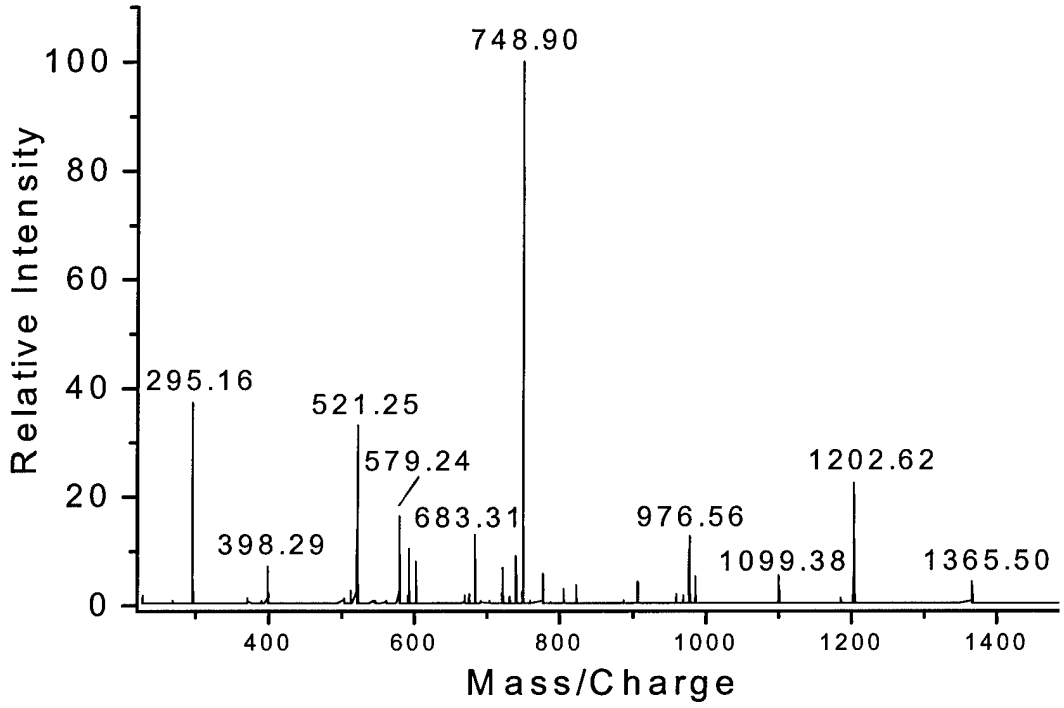


Figure 16. Product Ion Scan of $[M+2H]^{2+}$ 748.90

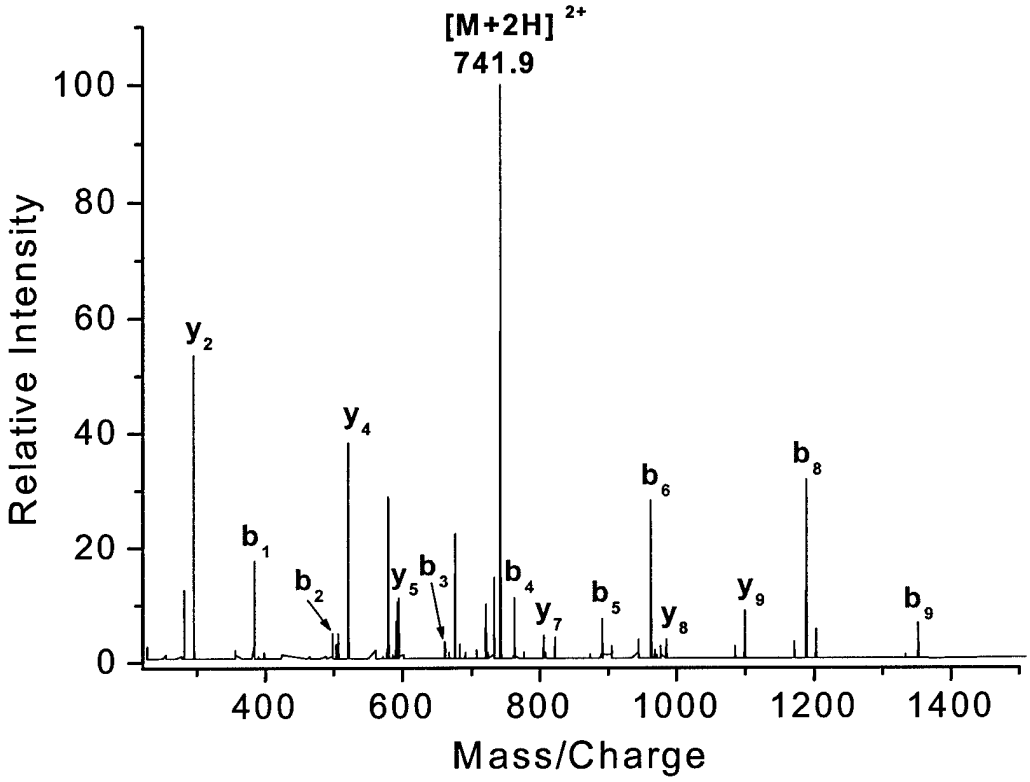


Figure 17. Assignment of b_n and y_n series ions for $[M+2H]^{2+}$ 741.9

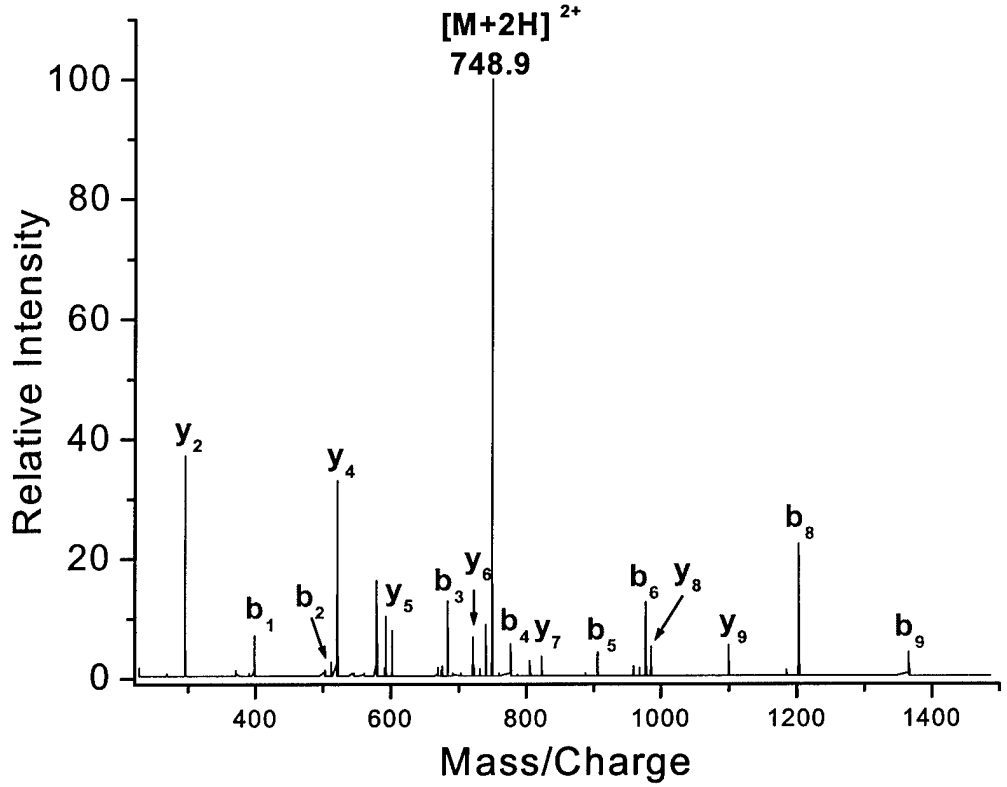


Figure 18. Assignment of b_n and y_n series ions for $[M+2H]^{2+}$ 748.9

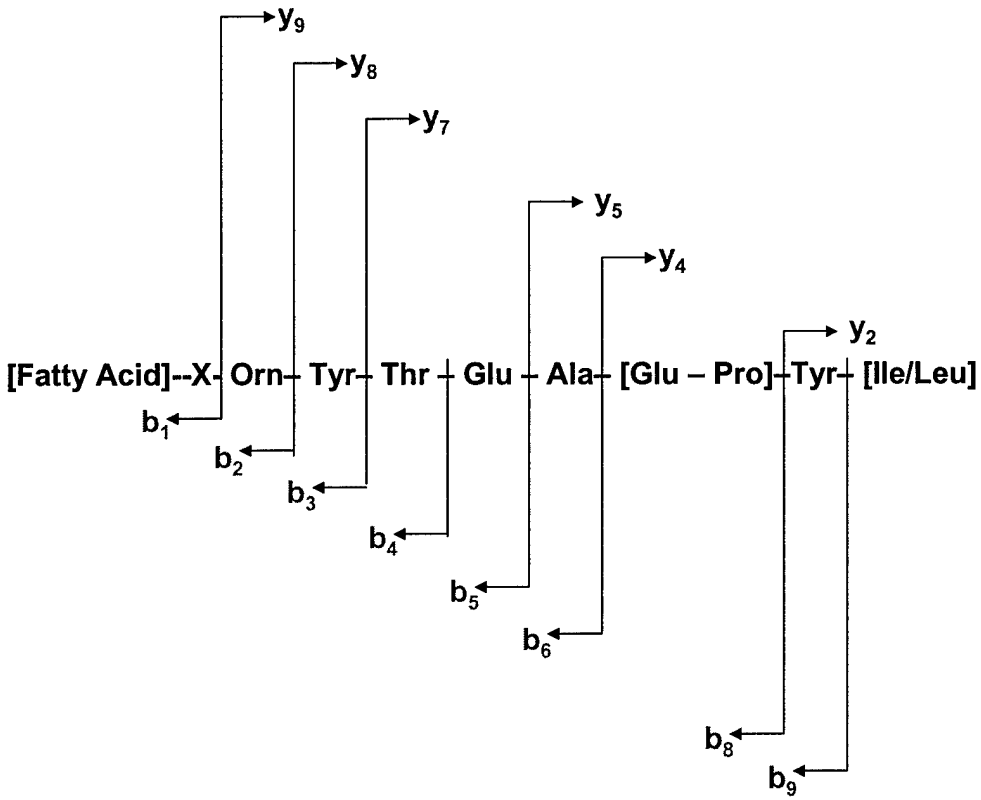


Figure 19. Assignment of Amino Acid Sequence to b_n and y_n Series Ions for Lipopeptides $[M+H]^+$ 1464 and $[M+H]^+$ 1478

isoleucine or leucine and water, from the $[M+H]^+$ ion peak at m/z 1523.7. The fragment ion peak at m/z 1229.6 is formed by the loss of 163 Da from the peak at m/z 1392.6, corresponding to tyrosine. The fragment ion peak at m/z 1004.5 is formed by the loss of 225 Da from the peak at m/z 1229.6, corresponding to glutamine and proline. Again, cleavage of the glutamate-proline or proline-glutamate bond was not observed. The fragment ion peak at m/z 905.6 is formed by the loss of 98.9 Da from the peak at m/z 1004.5, corresponding to valine. The fragment ion peak at m/z 776.4 is formed by the loss of 129.2 Da from the peak at m/z 905.6, corresponding to glutamic acid. The fragment at m/z 675.3 represents the loss of 101.1 Da from the peak at m/z 776.4, corresponding

to threonine. The fragment ion peak at m/z 512.3 is formed by the loss of 163 Da from the peak at m/z 675.3, corresponding to tyrosine. The fragment ion peak at m/z 398.3 is formed by the loss of 114 Da from the peak at m/z 512.4, corresponding to ornithine. Figure 21 is the assignment of b and y series ions to the CID spectra of $[M+2H]^{2+}$.

Significantly, this sequence shows a valine substitution for alanine at the sixth position and a glutamine substitution for glutamic acid in the fragment linked to the proline residue. To summarize, the lipopeptide weighing 1506 Da is composed of at least nine amino acids containing the following: isoleucine or leucine, two tyrosine residues, proline, glutamine, valine, threonine, glutamic acid, and ornithine. The amino acid residues corresponding to the b_n and y_n series ions are in Figure 22.

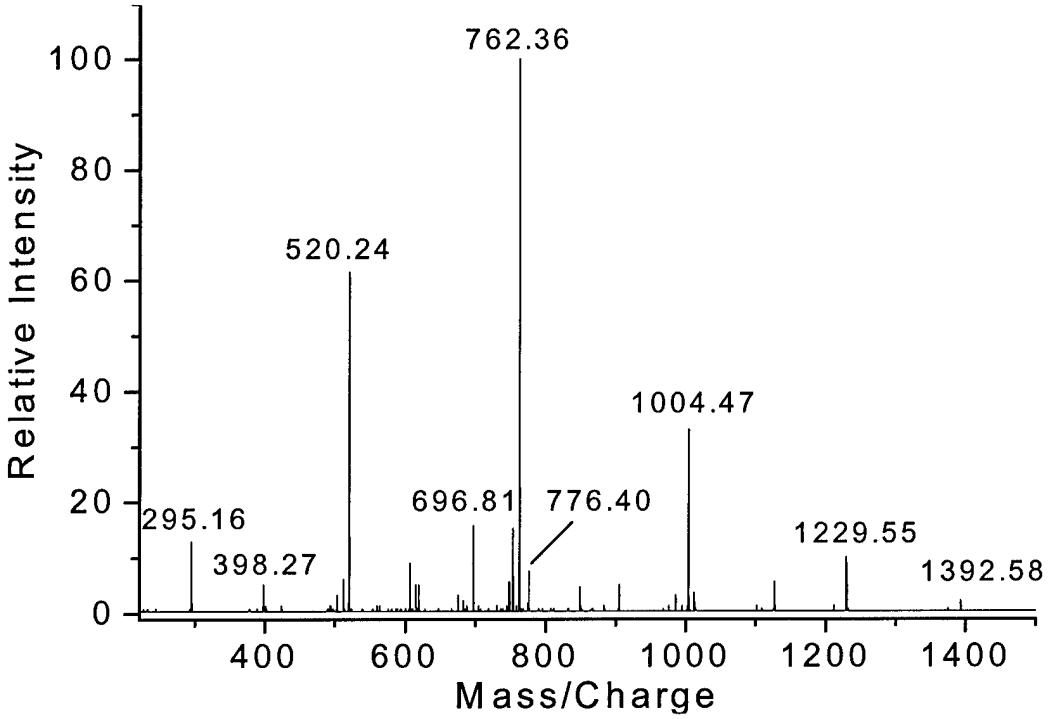


Figure 20. Product Ion Scan of $[M+2H]^{2+}$ 762.36

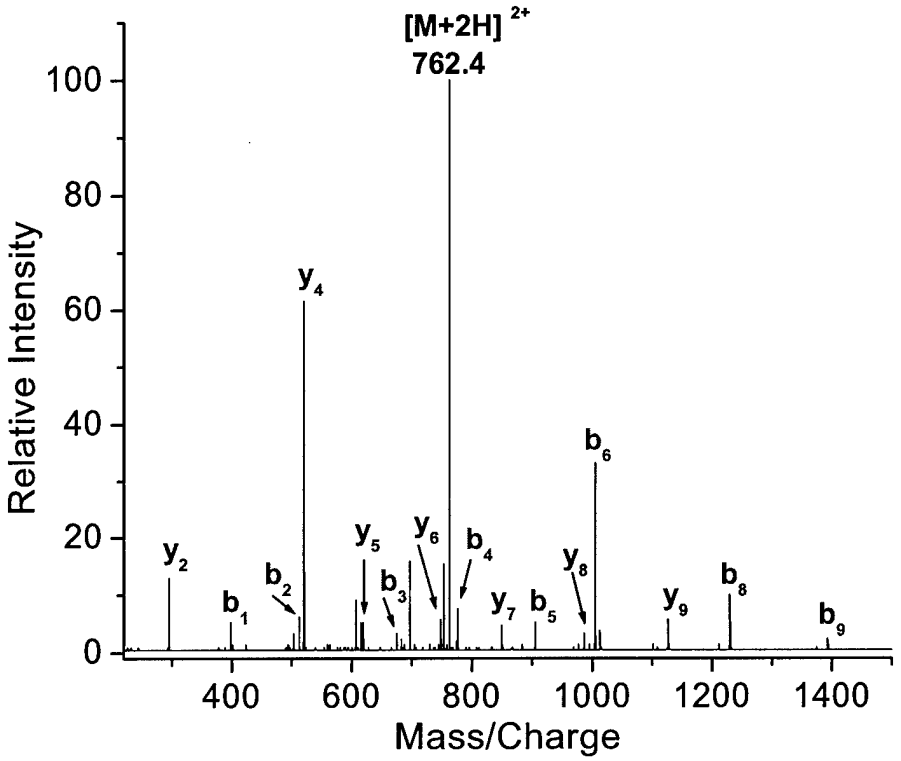


Figure 21. Assignment of b_n and y_n series Ions for $[M+2H]^{2+}$ 762.36

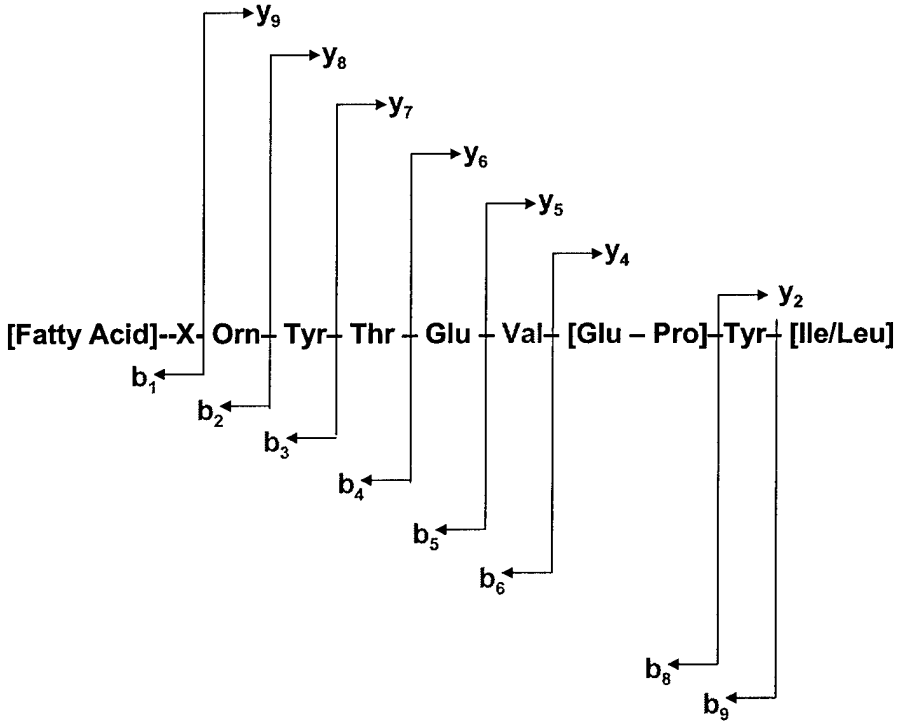


Figure 22. Assignment of Amino Acid Sequence to b_n and y_n Series Ions for Lipopeptide $[M+H]^+$ 1506

The fatty acid portion was cleaved from each of the sequenced lipopepetides and picolynyl derivatives were formed as stated in the Experimental Section. Analysis by GCMS showed peaks at m/z 53, 65, 80, 92, and 108 for each sample. The peaks at m/z 92 and 108 (data not shown) are characteristic of 3-picolynyl derivatives of fatty acids. No other peaks were detected to determine the structure of the fatty acid chain. This was possibly due to a low concentration of the derivatives.

DISCUSSION

B. globigii produced a biomarker peak at m/z 1478 that was consistently present in spores and detected with minimal treatment using standard protocols. The biomarker was determined to be a lipopeptide, structurally characterized, and shown to be a distinct marker for *B. globigii*. This biomarker was well suited for this study in that it was visible in all phases of the bacterium's life cycle and appeared to be unique to *B. globigii*. Screening of multiple samples took place in minutes and was limited only by the requirements to culture the bacteria, demonstrating the rapid analysis capabilities of MALDI-TOF mass spectrometry. While different culture conditions may change the expression of certain biomarkers, little variation was detected in the presence and relative abundance of this biomarker in *B. globigii* spores.

The lipopeptide initially showed three distinct properties: 1. It was easily washed from the spore surface indicating a loose association with the cell wall; 2. it was visible in the cell-free culture medium indicating it is excreted during sporulation; and 3. It was observed as a signal cluster of related molecules, a characteristic of microbial lipopepetides. *Bacillus subtilis* spores are noted for producing three general classes of bioactive compounds that appear as signal clusters: small hydrophilic antifungal peptides, antibacterial polypeptide lantibiotics, and amphiphilic, antifungal lipopeptide antibiotics. Therefore it was not alarming to observe similar compounds in a *B. globigii* spores. A peak at m/z 3397 was seen in the mass spectrum of B. globigii vegetative cells and was

similar in mass to the lantibiotic subtilin (at m/z 3343) from *B. subtilis*. A peak at m/z 1085 was visible in all spectra and matched previously observed mass spectra of the iturin lipopepetides (m/z at 1079-1137) from *B. subtilis*. Two papers [29], [30] have described a lipopeptide, named fengycin, from two strains of *B. subtilis,* S499 and F-29-3. Each reported observation of different peaks at m/z 1541 and 1462, respectively, but indicated higher and lower homologues identical in mass to those seen in the biomarker cluster centered at m/z 1478. Additionally, fengycin is reported to exist in two forms that differ in a variable amino acid identified at the peptide's sixth position. Therefore, the biomarker cluster, observed at m/z 1478 in *B. globigii*, was characterized to determine its relation to fengycin.

The lipopeptide purification procedure includes an extraction step exploiting a solid phase technique and reverse-phase chromatography. Classical extraction methods use acid precipitation followed by extraction with methanol or a chloroform-methanol (2:1) mixture. In comparison, the solid-phase extraction on a C_{18} cartridge column was cited as a rapid technique that avoided reactions such as hydrolysis or esterification of lipopeptide functional groups due to prolonged acid or solvent exposure. No reactions of this kind were observed during purification. However, the cluster taken from the culture medium always appeared to contain a hydrolyzed compound peak at m/z 1496. This peak was always present when using surface washing and Folch extraction of spores. The C_{18} phase was efficient at absorbing lipopepetides due to the hydrophobic, fatty acid part of these molecules. Therefore the majority of impurities such as polar

contaminants are removed by washing the cartridge with water and aqueous methanol. Lipopeptides are completely retained using less than 50% aqueous methanol as a washing agent and eluted completely with methanol for separation. The iturin lipopepetides, observed at m/z 1085, were retained by the cartridge and eluted in the methanol fraction. Reverse phase HPLC enabled separation of three homologues fractions whose MALDI-TOF molecular masses seemed to correspond to fengycin homologues. The fraction with a mass of 1506 Da matched the mass determined by FAB-MS reported by Vanittankom [30]. The fraction with a mass of 1464 Da was close to the mass (at m/z 1462) determined by FAB-MS and reported by Schneider and coworkers [29]. The biomarker's solubility and mass measurements strongly indicated the similarity of both the *B. globigii* and *B. subtilis* lipopepetides.

The general structure of microbial lipopepetides is commonly composed of a β-hydroxy fatty acid integrated into a cyclic peptide, normally of seven amino acids. Fengycin has an ester bond (lactone) formed between the C-terminal amino acid and the third amino acid in the sequence. Except for the ester bond, only peptide bonds are present in the peptide domain. The lactone bond was confirmed chemically by mild alkaline hydrolysis that does not cleave amide bonds. This feature proved useful for obtaining structure specific fragment ions by collision with a gas (CID).

With the Q-Star mass spectrometer, ions of a single mass can be selected for activation. In this way amino acid sequences can be determined. The most

important fragments are formed by the cleavage of the amide bonds and can be separated into two classes. Fragments that retain the charge on the N-terminus and cleave on the C-terminus are designated b_n series fragments, with n representing the number of amino acids contained. Fragments generated from the N-terminal retains the charge on the C-terminal and are designated y_n series fragments. This fragmentation pattern [34] is shown in Figure 23.

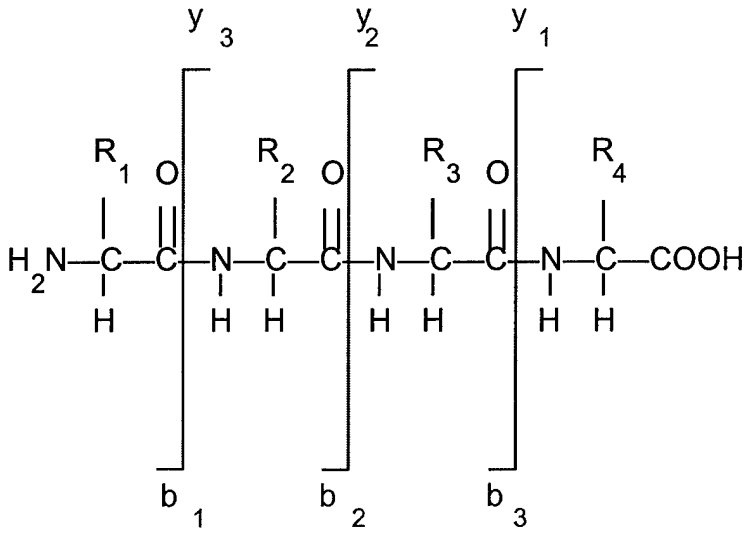


Figure 23. Peptide Fragmentation Pattern

Decomposition of the cyclic peptide may open any peptide bond and cause complex mass spectrum of fragmentation patterns. Therefore, the ring was opened and the peptide was subjected to CID to determine its amino acid composition sequence.

Series b_n ions were identified by comparing the noted fragments from the product ion scans of m/z 741.9, 748.9, and 762.4. A variable region of each molecule of the lipopepetides contained a fatty acid that differed from other homologues by units of 14 Da. Fragments produced in the mass spectra of the

three scans which varied by 14 Da were labeled as b ions. Fragments which were visible in all three spectra and did not differ, were labeled y ions. Upon analysis of the b series ions, a partial amino acid sequence was determined. Both molecules weighing 1464 and 1478 Da had identical sequences. When this partial sequence was compared to the sequence of fengycin, the amino acids differed from version IX with a glutamic acid located at the third position. The open form amino acid sequence comparison is shown in Figure 24.

[Fatty Acid]– Glu -- Orn – Tyr– Thr – Glu – Ala– Gln – Pro–Tyr– Ile

[Fatty Acid]-- X-- Orn – Tyr– Thr – Glu – Ala–[Glu – Pro]–Tyr– [Ile/Leu]

Figure 24. Amino acid sequences of Fengycin IX (top) and *B. globigii* lipopeptides $[M+H]^+$ 1464 and $[M+H]^+$ 1478

The methyl esterification reaction identified four sites at which methyl groups were attached to produce a total mass shift of 56 Da. One site is the carboxy terminus. Only two glutamic acids were reported in the earlier study of fengycin IX. The increase in molecular mass and the fourth methyl esterification site are considered to arise from the longer base hydrolysis reaction. The glutamine residue at position 8 was converted to glutamic acid during this reaction.

Analysis of the product ion scan at m/z 762.4 showed a valine, instead of alanine, at the sixth position of the peptide. This partial sequence matches the structure reported for fengycin XII [29]. The amino acid sequence are compared in Figure 25.

[Fatty Acid]– Glu -- Orn – Tyr– Thr – Glu – Val – Gln – Pro–Tyr– Ile

[Fatty Acid]-- X– Orn – Tyr– Thr – Glu – Val–[Glu – Pro]–Tyr– [Ile/Leu]

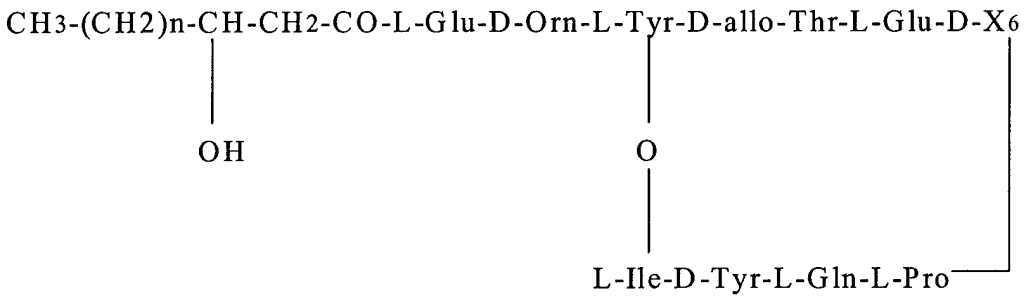
Figure 25. Amino Acid Sequence of Fengycin XII (top) and *B. globigii* lipopeptide $[M+H]^+$ 1506

The mass spectra of the derivatized fatty acids were similar and contained ions characteristic of side chain structure but offered no real structural information. Fengycin IX is reported to have C_{16} to C_{18}-3-hydroxy fatty acids while fengycin XII is reported to have C_{16} and C_{17} fatty acids [29], [30]. These structures were established by NMR and GCMS.

Based on these results, the compounds with protonated molecular masses of 1464 and 1478 are proposed to be fengycin IX and that with a protonated mass of 1506 is proposed to be fengycin XII. The complete structures of each, as solved by earlier studies [32] are shown in Figure 26.

Although, I did not test the functional role of fengycin in *B. globigii*, isolates from *B. subtilis* possess antimicrobial properties. When applied to fungi and

other bacteria, fengycin showed potential as an antibiotic against fungi. Like extracellular enzymes of *Bacillus*, synthesis of lipopeptide antibiotics generally



Fengycin	3-OH-Fatty acid	D-X6
IX	C16	Ala
XII	C17	Val

Figure 26. Fengycin Structures from *B. subtilis*

occurs as the culture enters the early stages of sporulation [32]. The event can be used as a phenotypic marker for the early biochemical events of spore formation. Production of fengycin in B. globigii followed this pattern and was evident upon analysis of the culture supernatant. There have been many attempts to assign functional roles in sporulation to these antibiotics but the isolation of antibiotic deficient mutants that sporulate has excluded any essential role in sporulation [33], [34]. It seems therefore that antibiotic synthesis, like the production of extracellular enzymes, is a response to the same environmental conditions that signal sporulation and may serve other functions. Possibly, it is produced to kill or inhibit other microorganisms, thereby providing a competitive advantage to the producing bacteria.

Interestingly, both forms of fengycin (IX and XII) were observed in spores and the culture supernatant of *B. globigii* and varied in the amino acid present at the sixth position of the molecule as well as in fatty acid length. Other homologues are suspected at masses of 1450 and 1492 Da, but were only seen in mass spectra of crude samples. Additionally, analogous lipopepetides have been discovered in the pathogen *B. cereus* and differ only in the stereochemistry of their tyrosine residues [29]. Further studies to isolate and characterize these molecules and their antifungal properties could prove useful to understanding the role of microbial lipopepetides for all members of the genera *Bacillus*.

Principally, our efforts have focused on the use of mass spectrometry to detect biomarkers in intact spores for differentiation of microorganisms. *Bacillus globigii spores* (*B. subtilis var. niger*) have been safely used as a stand-in for pathogenic bacteria in field tests and in this work, the organism served as a model organism for the rapid use of mass spectrometry to detect and characterize specific biomarkers in pathogenic bacteria not distinguishable by other methods. With the application of instrument-based analytical techniques to biochemistry and molecular biology, the reliance on traditional microbiological tests for gathering phenotypic data has decreased. Instead a wealth of information has been obtained by using biomarkers to identify microorganisms. The pairing of this chemotaxonomic approach with mass spectrometry has proven successful for detecting and differentiating many microorganisms. This is particularly relevant in identifying related Gram-positive bacteria and pathogens. While this approach does not necessarily mirror the genome more than

morphological traits, it does provide a more powerful and realistic tool that supplements traditional approaches. In many cases, such as cell wall components and proteins, the chemotaxonomic identifications are consistent with classifications derived from nucleic acid based methods. Use of mass spectrometry will continue to develop as a strategy in microorganism identification. It is well suited for many approaches to current and future problems such as the screening for the presence of contamination or infection, identification of a specific genus or species, and in the structure elucidation and structure function studies of biomolecules in microorganisms.

REFERENCES

1. Harwood, C., ed. *Bacillus*. Biotechnology Handbooks, ed. T.E.a.R. Sherwood. Vol. 2. 1989, Plenum Press: Salisbury, England. 414.
2. Buzby, J.C., *Bacterial Foodborne Disease: Medical Costs and Productivity Loses*, . 1996, U. S. Department of Agriculture: Washington, D.C. p. 80.
3. Walt, D., *Biowarfare Detection.* Analytical Chemistry, 2000: p. 739A - 746A.
4. Mitruka, B., *Methods of Detection and Identification of Bacteria*. 1976: CRC Press. 256.
5. Doi, R., ed. *Biology of Bacilli: Applications to Industry*. Biotechnology Series, ed. J. Davies. 1992, Reed: Paris. 370.
6. S. Hammond, P.L., and A. Rycroft, *The Bacterial Cell Surface*. 1983, Washington, D.C.: Kapitan Szabo Publishers. 226.
7. Nelson, W., ed. *Modern Techniques for Rapid Microbial Analysis*. Vol. 1. 1991: Kingston, Rhode Island. 263.
8. Holland, R.S., *Protein Fingerprinting and microbial identification*, in *Biotechnology Information*, R. Wakeford, Editor. 1986, IRL Press: Oxford, England.
9. Shaw, N., *Lipid Composition as a guide to the Classification of Bacteria*, in *Chemotaxonomy of Bacterial Lipids*. 1975.
10. Fenselau, C., ed. *Mass Spectrometry for the Characterization of Microorganisms*. ACS Symposium Series 541, ed. M. Comstock. 1994, American Chemical Society: Washington, DC. 240.

11. Holland R., *Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns using Matrix-Assisted Laser Desorption/Ionization with Time-of-Flight Mass Spectrometry.* Rapid Comminications in Mass Spectrometry, 1996.

12. Claydon, M., *The Rapid Identification of Intact Microorganisms Using Mass Spectrometry.* Nature Biotechnology, 1996. **14**(November): p. 1584-1586.

13. Welham, K., *The Characterization of Micro-organisms by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry.* Rapid Communications in Mass Spectrometry, 1998. **12**: p. 176-180.

14. Baar, B., *Characterization of bacteria by matrix-assissted laser desorption/ionization and elestrospray mass spectrometry.* FEMS Microbiology Reviews, 2000. **24**: p. 193-219.

15. Krishnamurthy, T., *Detection of Pathogenic and Non-Pathogenic Bacteria by Matrix-assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry.* Rapid Communications in Mass Spectrometry, 1996. **10**: p. 883-888.

16. Harwood, C., ed. *Molecular Biological Methods for Bacillus.* Modern Microbiological Methods, ed. M. Goodfellow. 1990, John Wiley & sons: Tyne. 564.

17. Russell A., *The Destruction of Bacterial Spores.* 1982, New York: Academic Press Inc. 333.

18. Ryzhov, V., *Rapid Characterization of Spores of Bacillus cereus Group Bacteria by Matrix-Assisted Laser Desorption-Ionization Time-of-Flight Mass Spectrometry.* Applied and Environmental Microbiology, 2000. **66**(9): p. 3828-3834.

19. Hathout, Y., *Identification of Bacillus Spores by Matrix-Assisted Laser Desorption Ionization-Mass Spectrometry.* Applied and Environmental Microbiology, 1999. **65**(10): p. 4313-4319.

20. Demirev, P., *Microorganism Identification by Mass Spectrometry and Protein Database Searches.* Analytical Chemistry, 1999. **71**(14): p. 2732-2738.

21. Lay, J.O., *MALDI-TOF mass spectrometry bacterial taxonomy.* Trends in Analytical Chemistry, 2000. **19**(8): p. 507-516.

22. Leenders, F., *Rapid Typing of Bacillus subtilis Strains by their Secondary Metabolites Using Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Intact Cells.* Rapid Communications in Mass Spectrometry, 1999. **13**: p. 943-949.

23. Peypoux, F., *Iturins A and B.* Tetrahedron, 1973. **29**: p. 3455.

24. Michel, F.B., *Isolation and Charcterization of New Iturins: Iturin D and Iturin E.* The Journal of Antibiotics, 1987. **40**(4): p. 437-442.

25. Hathout, Y., *Kurstakins: A New Class of Lipopeptides Isolated from Bacillus thuringiensis.* Journal of Natural Products, 2000.

26. Razafindralambo, H., *Purification of antifungal lipopeptides by reversed-phase high performance liquid chromatography.* Journal of Chromatography, 1993. **639**: p. 81-85.

27. Folch, J., and Stainley, G., *A simple method for the isolation and purification of total lipids from animal tissues.* Journal of Biological Chemistry, 1957. **226**: p. 497-509.

28. Yao, X., Personal Comminication. 2000.

29. Sonenshein, ed. *Bacillus subtilis and Other Gram-Positive Bacteria*. . 1993, American Society for Microbiology: Washington, D.C. 987.

30. Desai, J., *Microbial Production of Surfactants and Their Commercial Potential.* Micribiology and Molecular Biology Reviews, 1997. **61**(1): p. 47-64.

31. Cooper, D., *Biosurfactantas.* Microbiological Sciences, 1986. **3**(5): p. 145-149.

32. Schneider, J., *The Structure of Two Fengycins from Bacillus subtilis S499.* Z. Naturforsch, 1999. **54c**: p. 859-866.

33. Vanittannakom, N., *Fengycin - A novel Antifungal Lipopeptide Antibiotic Produced by Bacillus subtilis F-29-3.* The Journal of Antibiotics, 1986. **39**(7): p. 888-901.

34. Siuzdak, G., *Mass Spectrometry for biotechnology*. 1996, San Diego: Academic Press. 161.