UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MINING MICROBIAL COMMUNITIES: EXPLORING SYMBIOTIC BACTERIA AS A SOURCE OF NEW NATURAL PRODUCTS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CHRISTINE M. THEODORE
Norman, Oklahoma
2013

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/6388

MINING MICROBIAL COMMUNITIES: EXPLORING SYMBIOTIC BACTERIA AS A SOURCE OF NEW NATURAL PRODUCTS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

BY

_____
Dr. Robert H. Cichewicz, Chair

_____
Dr. Kenneth M. Nicholas

_____
Dr. Joseph M. Suflita

_____
Dr. Ann H. West

_____
Dr. Helen I. Zgurskaya

© Copyright by CHRISTINE M. THEODORE 2013
All Rights Reserved.

For my parents

Acknowledgements

To thank everyone as much as they deserve would require an infinite amount of time and space. But I would briefly like to thank my advisor, Dr. Cichewicz. His dedication and enthusiasm towards our field, and science as a whole, is not matched by many. I have appreciated his mentorship throughout my time at OU and I will carry the knowledge and experiences gained to every new challenge I encounter. I would like to thank my committee members for their advice and guidance during my PhD career. A very special thank you to Drs. Foster, Powell and Nimmo. Without their hard work and expertise my dissertation would not be possible. I am extremely grateful to all past and present members of the Natural Products Discovery Group. Every person I have worked with has brought a unique set of knowledge and strengths, and I have benefited from working with each of them. I wish to thank the faculty members of the Department of Chemistry and Biochemistry, who have worked diligently to foster an environment of collaboration and innovation.

I would like to acknowledge my family and friends, near and far, old and new, for all their encouragement along the way. Finally, I am who I am today due to the love and support from my amazing husband, Tom. His confidence in me has never wavered, especially when I had little confidence in myself. Without him my PhD would still be a wish instead of a reality.

Table of Contents

Acknowledgements iv

List of Tables viii

List of Figures ix

Abstract x

Chapter 1: Bacteria as a Source of Bioactive Natural Products 1

1.1 The importance of natural products in modern pharmaceuticals 1

1.2 Bacteria as a source of natural products 2

1.3 The biosynthetic machinery of secondary metabolite production 3

1.4 Symbiotic bacteria: Uncovering a surprising source of natural products 5

1.5 Mammalian Microbiomes: A Potential Solution to Some Field Specific Problems 7

Chapter 2: Hypothesis and chapter overviews 12

2.1 Hypothesis 12

2.2 Chapter 3. Production of cytotoxic glidobactins/luminmycins by *Photorhabdus asymbiotica* in liquid media and live crickets 12

2.3 Chapter 4. Genomic and metabolomic characterization of natural product biosynthetic diversity in a strain of *Brevibacillus laterosporus* isolated from a feral hog 13

2.4 Chapter 5. Development of an efficient and robust screening method for the detection of biosynthetically talented organisms from mammalian microbiome 13

Chapter 3. Production of cytotoxic glidobactin/luminmycins in liquic media and live crickets 15
3.1 Introduction and discussion 15
3.2 Experimental 25
3.2.1 General Experimental 25
3.2.2 Microorganisms and culture conditions 25
3.2.3 Extraction and isolation 26
3.2.3 Productin of glidobactins and luminmycins in an insect host 26
3.2.4 Cell cytotoxicity assays 27
3.2.5 20S proteasome inhibition assay 27
Chapter 4. Genomic and Metabolomic Insights into the Natural Product Biosynthetic Diversity of a Feral Hog Associated *Brevibacillus laterosporus* Strain 29
4.1 Introduction 29
4.2 Results and Discussions 30
4.3 Materials and Methods 47
4.3.1 Collection and bacteria isolation 47
4.3.2 Overlay assay 48
4.3.3 DNA extraction, PCR amplification and sequence 48
4.3.4 Culture conditions and extractions 49
4.3.5 Compound isolation 49
4.3.6 Marfey's analysis of auriporcine 51
4.3.7 X-ray experimental information 51
4.3.8 Compound characterization 53

4.4 General experimental conditions 53
Chapter 5. A high-throughput screening procedure of mammalian microbiome isolates for secondary metabolite production, antimicrobial activity and cancer cell cytotoxicity 54
5.1 Introduction 54
5.1.1 Accessing new biodiversity in the search for new natural products 54
5.1.2 Screening approach for biosynthetically talented bacteria 55
5.1.3 Purification of organisms 57
5.1.4 Selection of animals for sampling 57
5.1.5 Interpretation of LCMS data 59
5.1.6 Controls 61
5.2 Materials 61
5.2.1 Reagents 61
5.2.3 Equipment 63
5.2.4 Reagent setup 64
5.2.5 Cell preparation for biological assays 67
5.2.6 Equipment setup 67
5.3 Protocol 69
5.4 Anticipated results 72
References 73
Appendix 84
Appendix table of contents 84

List of Tables

Table 1. ^{1}H and ^{13}C NMR data. (500 Mhz, DMSO-d_6, 25 ^{o}C) for luminmycin D (3) ... 21

Table 2. Biosynthetic gene clusters in PE36 identified by antiSMASH analysis. NRPS= nonribosomal peptide synthase, PKS = polyketide synthase, T1 = type I, T2 = type II, T3 = type III..... 34

Table 3. NMR data for 8..... 37

Table 4. NMR data for 9, showing the data for the different conformations 40

Table 5. MS^{n} fragment information for 9. Pho = phenylalaninol. MHP = methylhydroxyl pentanone 41

Table 6. Gradient scheme for LCMS separation of crude extracts 68

List of Figures

Figure 1. Examples of clinically relevant natural products (producing bacterium in parenthesis) 3

Figure 2. Examples of the major structural classes of bacteria produced natural products 5

Figure 3. Examples of structurally related compounds isolated from very different sources (in parenthesis). Their biosynthetic origins are bacterial symbionts. 7

Figure 4. Examples of compounds produced by symbiotic bacteria (host listed in parenthesis). 11

Figure 5. *Photorhabdus* sp. and nematode life cycle. Phase I stages are indicated by blue boxes, phase II by red. 16

Figure 6. LC-MS traces (254 nm) of P. asymbiotica in TSB (red, top trace) and defined medium (black, lower trace). 19

Figure 7. Structures of isolated compounds 20

Figure 8. LCMS single ion trace profiles of various extracts. 23

Figure 9. *In vitro* cytoxicity and proteasome inhibition 24

Figure 10. Overlay experiment of organisms isolated from a wild hog ear. 31

Figure 11. Phylogeny of *Brevibacillus laterosporous* strain PE36 and close relatives.. 33

Figure 12. Isolated compounds. 36

Figure 13. PyMol representations of 9. 42

Figure 15. Automated liquid handling robot from Aurora Biomed 56

Figure 16. A skunk and armadillow sampled roadside. 59

Figure 17. Illustration of generated LCMS data.. 60

Abstract

As ubiquitous inhabitants of Earth, bacteria represent a diverse resource in the search for new natural products. In order for the field of natural products to progress, new and exciting sources of biodiversity must be continually accessed. Microbiomes, the community of commensal and pathogenic microorganisms that live in and on other organisms, present a unique opportunity to access new biological diversity. These symbiotic bacteria have proven to be prolific producers of natural products and represent relatively untapped resource of new secondary metabolites.

Photorhabdus asymbiotica, a bacterial symbiont of an entomopathogenic nematode is an emerging health concern for humans. It engages in a bi-phasic lifestyle that switches between symbiotic and pathogenic phases. Its genome had been previously sequenced and our analysis of the secondary metabolite pathways indicated it had the genetic machinery necessary to produce multiple natural products. Extensive media screening studies were employed to induce the production of the encoded for natural products. Two known and one new proteasome inhibitors were isolated from the optimized culture conditions. The theory that *P. asymbiotica*'s lifestyle switch is governed in part by environmental cues was explored by injecting the bacterium into live crickets. Recovery of the bioactive natural products from the cricket corpus supported the belief that the nutrient composition of the insect gut plays a role in the life cycle switch of *P. asymbiotica*.

The microbiomes of mammals are a rich and diverse source of bacteria. While these microbiomes are well studied in terms of populations, little is known about their ability to produce natural products. The biosynthetic potential of the microbiome of a feral hog

was examined. One isolated organism, identified as a new strain of *Brevibacillus latersporus*, was chosen for further genomic and metabolomic characterization. Thirty two biosynthetic gene clusters were identified in the organism's genome. Three known bioactive compounds were isolated along with two new natural products. A portion of the known and new natural products were matched up with their biosynthetic gene clusters. This work showed that symbiotic bacteria from small mammals are worthy of further investigation by natural product investigators.

Finally, a method was developed to sample, isolate, and screen members of microbiome communities for biosynthetically talented and culturable organisms. By employing the use of liquid handling robots and specialized instrumentation, an efficient method for quickly evaluating organisms for secondary metabolites was created. By combining an efficient screening method metabolomic investigations, a snap shot of the chemical products generated by microbiome members has been achieved.

Chapter 1: Bacteria as a Source of Bioactive Natural Products

1.1 The importance of natural products in modern pharmaceuticals

For thousands of years, chemical substances from plants, animals and microbes have played a substantial role in the treatment of human disease.[1] These small molecules, called natural products or secondary metabolites, have become an important source of novel pharmaceuticals, most notably for the treatment of infections and cancer.[2] While their native biological roles are often unknown,[3] these substances have played a significant role in the modern pharmaceutical industry. Since 1981, approximately 50% of all approved small-molecule drugs have been derived from or inspired by natural products.[4] From 1981-2004, more than a third of all drug sales were based on natural products or natural product inspired compounds.[5] Despite their prominent role in medicine, modern pharmaceutical companies had begun to shy away from natural products, believing the natural products pipeline to be exhausted and preferring to use a more synthesis dependent approach.[6] By 2010, the total number of approved small molecules reached its second lowest level in the past 30 years however, fully half of those were natural product based.[4] This lull in molecule approval has come at a troubling time in modern medicine, as the number of infections by antimicrobial resistance organisms is increasing around the globe.[7] The increased need for novel pharmaceuticals and decreasing molecule approval rate indicates that natural products remain an important area of investigation.

1.2 Bacteria as a source of natural products

Bacteria are an important part of the Earth's biota; they are ubiquitous and account for a greater biomass than both plants and animals.[8] As such, bacteria represent an extensive source of natural products. Unlike plant secondary metabolites, whose use in traditional medicines guided researchers towards bioactive compounds, bacterial based pharmaceuticals were the result of a more targeted screening approach. [9] Early application of these screening methods were so efficient that members of all the main clinically relevant families of antibiotics were discovered by the conclusion of the 1950s.[9] This effort in the early years of microbial natural product isolation research, through to the present day, has resulted in pharmaceuticals of many drug classes. These pharmaceutical classes include antibiotics, antitumor agents, immunosuppressive agents, enzyme inhibitors, insecticides, and antiparasitic agents.[9] Some common clinically relevant examples of compounds produced by bacteria are shown in Figure 1. Streptomycin[10] and tetracycline[11] represent examples of early antibiotics. Mitomycin[12] is an antitumor agent approved for use by the FDA in 1974.[13] Tacrolimus, discovered in 1984[14] and early usages begun in the early 90s[15], is used as an immunosuppressant agent.

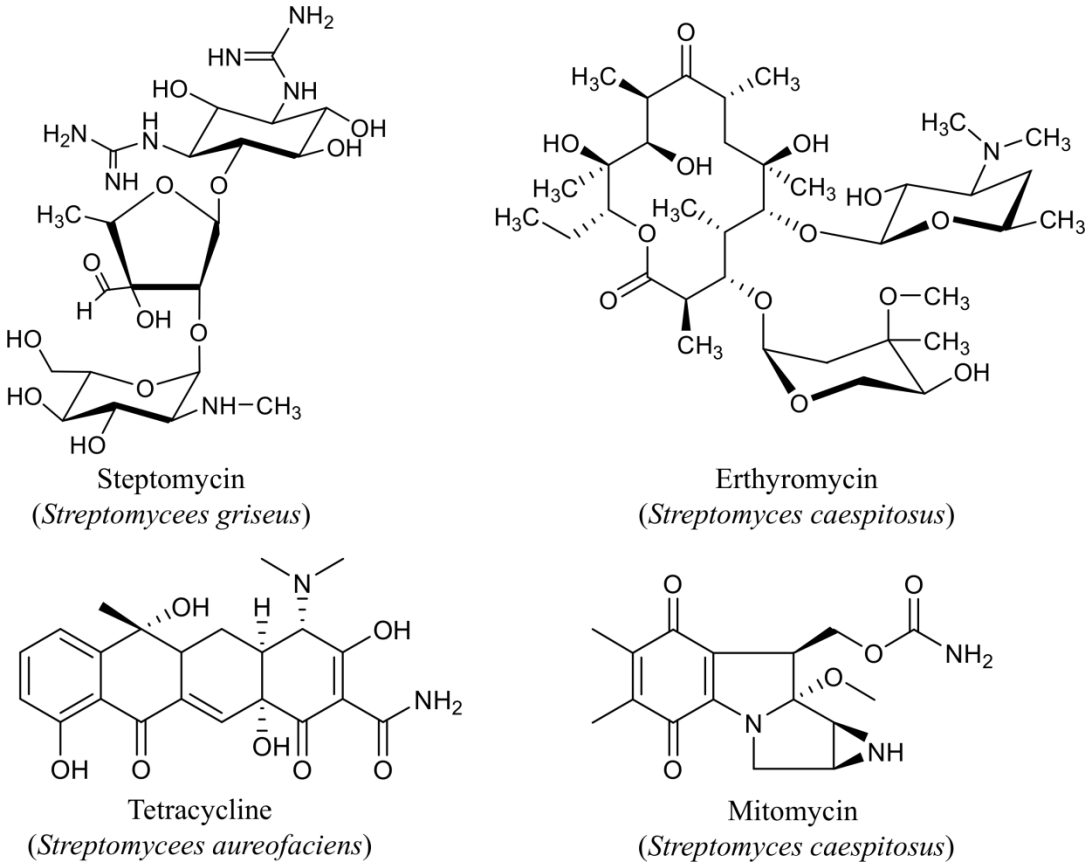


Figure 1. Examples of clinically relevant natural products (producing bacterium in parenthesis)

1.3 The biosynthetic machinery of secondary metabolite production

Bacteria have demonstrated that they are talented producers of natural products. Through the use of a few key biosynthetic strategies,[16] bacteria are able to assemble compounds spanning a range of structural classes. There is a wide variety of secondary metabolite classes but the vast majority of bacterial secondary metabolites are produced by two types of gene clusters, polyketide synthases (PKS) and non-ribosomal peptide synthases (NRPS). While they produce structurally different compounds, the logic of

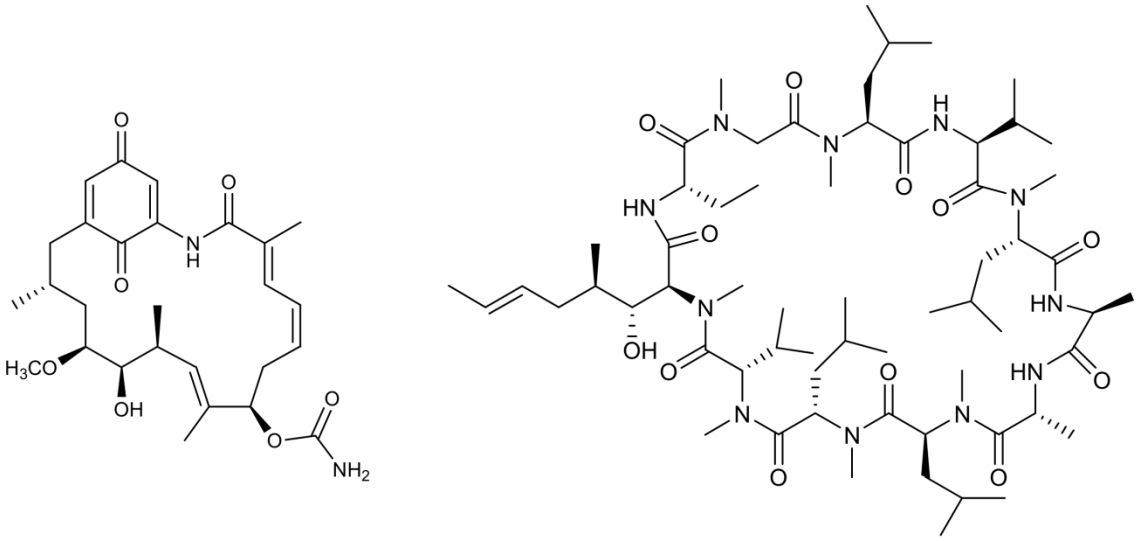
their mechanistic assembly is very similar. A brief outline of these biosynthetic assembly lines follows with examples of these structural classes shown in Figure 2.

The PKS biosynthetic system, involves large proteins, divided into modules, which each preform an enzymatic activity. Not every domain, however, is present in every system. Through the addition of simple molecular building blocks, such as malonyl-CoA and methylmalonly-CoA, the PKS domains create secondary metabolites in a process that is very similar to fatty acid biosynthesis. Both processes involve a repetitive decarboxylation Claisen thioester condensation of coenzyme A, which forms the anchor of the growing molecular chain.[17] Additional enzymatic processing produces highly functionally-substituted fatty acids.[18] The enzymatic architecture, molecular building blocks and mode of action can be used to divide PKSs into three general types, type I, II and III, to produce a heterogenous group of complex compounds.

NRPSs are large multifunctional enzymes that use proteogenic and/or nonproteogenic amino acids as substrates to produce secondary metabolites.[19] Like PKSs, they are divided into modules that perform specific tasks. The three most important subunits are the adenylation domain, the peptidyl carrier protein and the condensation domain. They are responsible for substrate recognition and activation, the transport of substrates and elongation of intermediates, and peptide bond formation, respectively. [20]

While the list of bacterial pharmaceutical examples is extensive, the potential well of natural products from bacterial sources has not been tapped dry. There is still a wealth of compounds waiting to be isolated. As an example, it is estimated that only 1-3% of all antibiotics produced by the bacteria *Streptomyces* have been discovered.[21] By

combining conventional culturing methods with novel culturing environments, genomic analysis and heterologous expression, natural products from bacterial sources are able to receive a renewed look by investigators.[22]



Geldanamycin
(Polyketide)

Cyclosprone
(Nonribosomal Peptide)

Figure 2. Examples of the major structural classes of bacteria produced natural products

1.4 Symbiotic bacteria: Uncovering a surprising source of natural products

Since the machinery necessary to produce secondary metabolites is encoded for in an organism's genome, it follows that compounds with similar scaffolds are produced by homologous gene clusters. A phenomenon in natural product discovery is bacteria natural products with structural similarity to invertebrate-derived compounds.[23] The structural similarity of natural products from these widely different organisms lead to the theory that the metabolites were being produced by associated bacteria. [24] As

prokaryotic and eukaryotic genomes exhibit distinct architectural differences, examining biosynthetic gene clusters identified through metagenomic analysis can help determine whether secondary metabolites are produced by bacteria or the animal host.[25] Figure 3 shows an example of this work by looking at the similarities between compounds of the pederin family of metabolites isolated from a beetle (*Paederus* sp.) and a sponge (*Psammocinia* sp.). Pederin, originally isolated from a beetle,[26] is structurally similar to compounds isolated from sponges. Some examples are onnamide a[27] and psymberin.[28] The Piel group illustrated in a series of publications using a genetic approach of exploring the biosynthetic gene clusters that the pederin family is produced by a symbiotic *Pseudomonas* species and not the sponge or beetle host.[25, 29] Another example of symbiotic bacteria being responsible for previously isolated compounds comes from a recently published paper showing bengamide E, an anti-tumor and immunosuppressant compound, is in actuality produced by an associated culturable myxobacteria, *Myxococcis virescens*.[30] The work by these groups, and others, show the exciting potential symbiotic bacteria have to produce secondary metabolites.

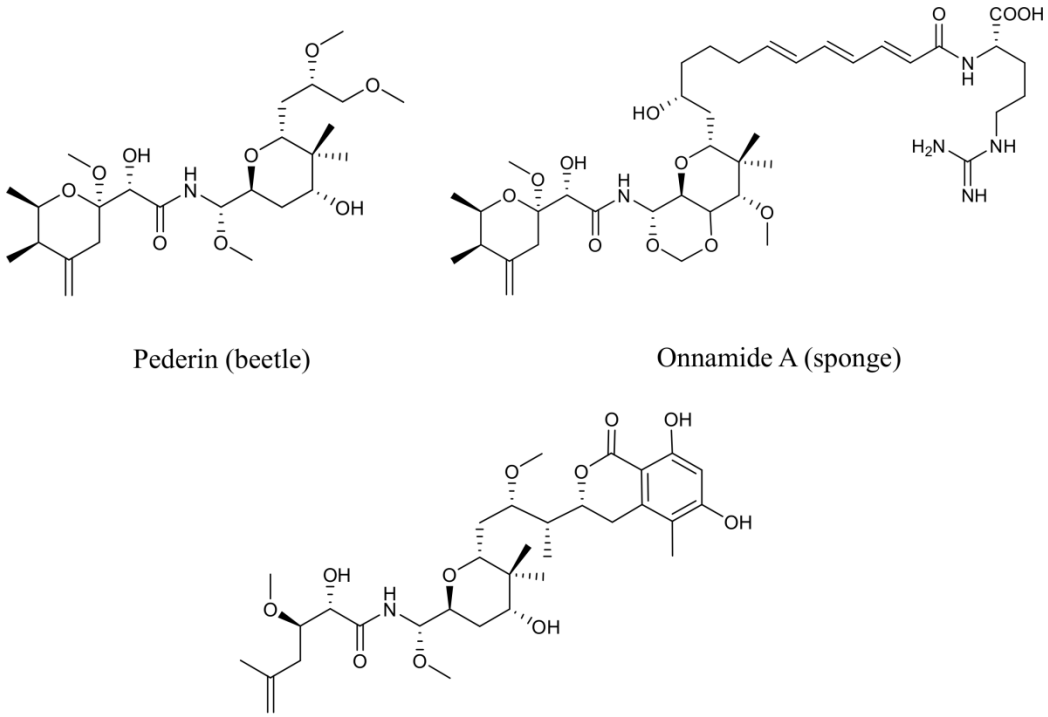


Figure 3. Examples of structurally related compounds isolated from very different sources (in parenthesis). Their biosynthetic origins are bacterial symbionts.

1.5 Mammalian microbiomes: a potential solution to some field specific problems

Natural products are widely accepted as the future of drug discovery. [31] If the field of natural products wishes to continue to be successful and move forward, it is essential to begin to solve some of the field specific challenges and reduce the bottle necks in the drug discovery process. As with any field, there are a number of areas that can be improved upon and multiple ways of achieving similar goals. The current state of natural products research, and the inherent problems associated with it, can be distilled into two broad areas: accessing new biodiversity and addressing the lingering doubts

from pharmaceutical companies. Mammalian bacterial microbiomes could be one way of solving these problems.

Originally defined by Lederberg as "[the] ecological community of commensal, symbiotic and pathogenic microorganisms that literally share our body space", [32] this term has expanded to include microbial communities of other organisms. The mammalian microbiome consists of the bacteria and fungi, both symbiotic and pathogenic, which live in association with the mammal. They are located throughout the animal's body; for example in the aural, nasal, oral, anal and vaginal cavities. The most well studied microbiomes of higher order animals, such as farm animals and humans, are the intestinal microbiomes. The gut is home to a rich and extremely diverse microbial community that lives and thrives in a symbiotic relationship with their hosts.

Gaining access to new biodiversity in the form of previously unknown or uncultured bacteria is essential to unlocking a new source of novel natural products. Historically, access to an untapped resource of biologics has resulted in an outpouring of new secondary metabolites. [33] Traditionally, sample collection by natural products researchers in remote areas is met with a variety of difficulties: multinational treaties on biodiversity[34], and concerns by local governments[6] complicate the collection process. Environmental changes due to seasonal variations or extinction due to natural or anthropogenic make return trips to collect additional samples difficult.

It is estimated that less than 0.1% of Earth's bacteria has been isolated and cultured,[35] indicating that bacteria is a largely untapped resource in the search for natural products. Gaining access to new biodiversity in the form of previously unknown

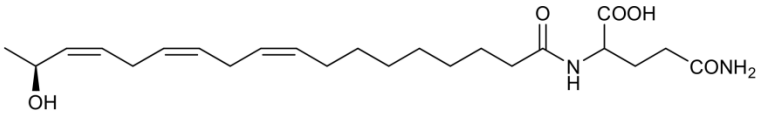
and uncultured bacteria could open the doors to a vast new resource of secondary metabolites.

Bacteria participate symbiotic relationships with other bacteria, plants, fungi, nematodes, insects, sponges and other organisms, producing a wide variety of natural products.[36] [36b] Figure 4 shows a sampling of compounds produced by symbiotic bacteria. Volicitin, an N-acyl amino acid present in the saliva of caterpillars and produced by gut associated bacteria,[37] elicits a plant to release volatile compounds that attract natural predators of the plant.[38] Swinholide, a cytotoxic compound isolated originally from a sponge,[39] was later shown to originate from associated cyanobacteria.[40] Bacilibactin, a sidereophore, is produced by *Bacillus* species associated with plants.[41]

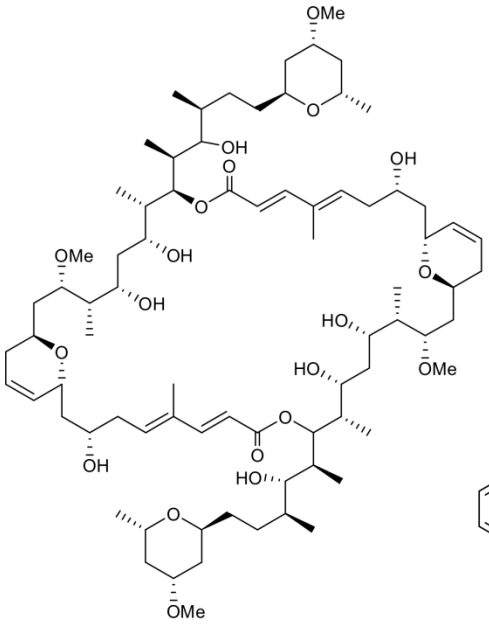
There is, however, little information on the microbiota of mammals and their natural products. Vertebrate microbiome members have evolved together over time, often persisting for the life of the mammal and are host specific.[42] However, their total community population is fluid, and relative populations can change based on diet or environment. The mammalian microbiome, although thoroughly studied in how it relates to health and disease of certain mammals, has not been explored as a source of natural products. Exploring the mammalian microbiome can address some of the pitfalls of natural product sample collection.

Finally, in order for natural products drug discovery to reach its full potential, the paradigm of the large pharmaceutical industry much change. Record sales and high stock holder expectation resulted in an industry culture of "blockbuster drugs", wherein the bulk of a company's profits are accounted for by a few high profit drugs.[6, 43] In

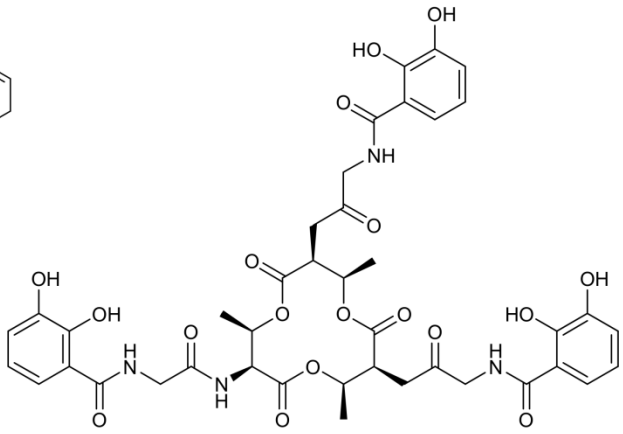
response, pharmaceutical companies have adopted a high throughput screening (HTS) method of synthetic compound libraries against specific targets.[21] While this method is quick and efficient, the "hit rate" (defined as compounds that are eventually commercialized) is extremely low, <0.001%.[44] Clearly, although fast and easy to screen an extensive amount of compounds, this method is far less effective than the hit rate associated with natural product screenings. Only looking at polyketide metabolites, the hit rate is 0.3%[44], a 300% increase. The challenge, however, is shifting the prevailing attitudes of the pharmaceutical industry. Although the nature of industry is to follow the path of least resistance to higher profits, academic and small biotechnology companies have picked up the bulk of natural products research.[31] By accessing new and exciting biodiversity, and in turn novel natural products, non-industrial labs can provide new insights into natural product research. The mammalian microbiome, with its ease of sampling and high potential for bioactive natural products, provides a gateway back toward industrial screening of secondary metabolites. With an easily accessible and reliable source of microorganisms, coupled with efficient screening methods, mammalian microbiomes have the capacity to usher in a new area of natural products research.



Volicitin (Insect)



Swinholide (Sponge)



Bacillibactin (Plants)

Figure 4. Examples of compounds produced by symbiotic bacteria (host listed in parenthesis).

Chapter 2: Hypothesis and chapter overviews

2.1 Hypothesis

Bacteria have traditionally been a rich source of natural products. At the core of discovering new natural products is the idea that accessing new biodiversity is essential for isolating new chemistry. **Symbiotic bacteria are a unique source of biodiversity and new natural products.** This hypothesis was tested via the following specific aims:

1. Isolation of new glidobactin derivatives from *Photorhabdus asymbiotica*, a nematode symbiont and entomopathogenic bacteria.
2. Isolation of a new bacteria strain from a feral hog microbiome, characterization of the secondary metabolite clusters and isolation of new natural products.
3. Development of a screening method for isolated microorganisms from mammalian microbiomes.

2.2 Chapter 3. Production of cytotoxic glidobactins/luminmycins by *Photorhabdus asymbiotica* in liquid media and live crickets

In this project, we looked at *Photorhabdus asymbiotica*, a unique bacterium that engages in a two part life cycle that requires adaption to both a symbiotic and pathogenic phases. The genome of *P. asymbiotica* contains several gene clusters that are predicted to be involved in the biosynthesis of unique secondary metabolites. However, recent reports on *Photorhabdus* sp. secondary metabolite production have indicated that many of its genes are silent under laboratory conditions. Using a panel of media and fermentation conditions, we have successfully achieved the production of a series of

new and known glidobactin/luminmycin compounds. These compounds were also obtained upon infection of live crickets with the bacterium.

2.3 Chapter 4. Genomic and metabolomic characterization of natural product biosynthetic diversity in a strain of *Brevibacillus laterosporus* isolated from a feral hog

In this project, a bacterium was isolated from a wild hog. Genomic analysis indicated that the organism was a new strain of *Brevibacillus latersporus*. From this organism basiliskamides A and B, known antifungal compounds, were isolated. New and known indole-pyrazine derivatives were isolated along with a new hybrid peptide. The structure of peptide was determined using NMR, MS/MS and x-ray crystallography. Finally, the isolated compounds were matched with their biosynthetic gene clusters.

2.4 Chapter 5. Development of an efficient and robust screening method for the detection of biosynthetically talented organisms from mammalian microbiome

This chapter focuses on the method development of collecting and purifying organisms from the mammalian biome and the development of a screening protocol for the identification of biosynthetically talent organisms. We developed a standard operating procedure for collecting samples from the mouth, ear, nasal, vaginal, and anal cavities as well as the small and large intestines of recently deceased small mammals native to central Oklahoma. These samples were purified to obtained pure cultures of each organism. Each bacterium was then grown in three different media types,

representing one nutrient rich environment and two minimal nutrient environments. These cultures were extracted using an automated liquid handling system and analyzed by LC-ESI-MS. The total ion trace was analyzed for each extract to identify prolific metabolite producers. The bacteria were regrown, extracted and each extract tested for antimicrobial and antitumor activity.

Chapter 3. Production of cytotoxic glidobactin/luminmycins in liquid media and live crickets

Adapted from a previously published article in the *Journal of Natural Products*.[45]

3.1 Introduction and discussion

Photorhabdus spp. live in association with a variety of soil-dwelling entomopathogenic *Heterorhabditis* nematodes.[46] After the nematode infects an insect, the bacterial symbiont is released, which kills the insect and provides the nematode a rich food source for growth and reproduction.[47] Once the insect food resources are exhausted, the bacteria recolonize the nematode, which reemerges from the insect carcass to begin its search for new prey .[46] These distinct environments, the nematode gut and the insect hemolymph, induce the differentiation of *Photorhabdus* spp. into what has been described as two phase variants. [48] Phase I (the pathogenic phase) occurs in the insect hemolymph and is characterized by the production of a variety of proteases, lipases and small molecules. Phase II (mutualistic phase) occurs in the nematode gut and is categorized by a general lack of production of most phase I extracellular digestive enzymes and other organic substances.[49,50]

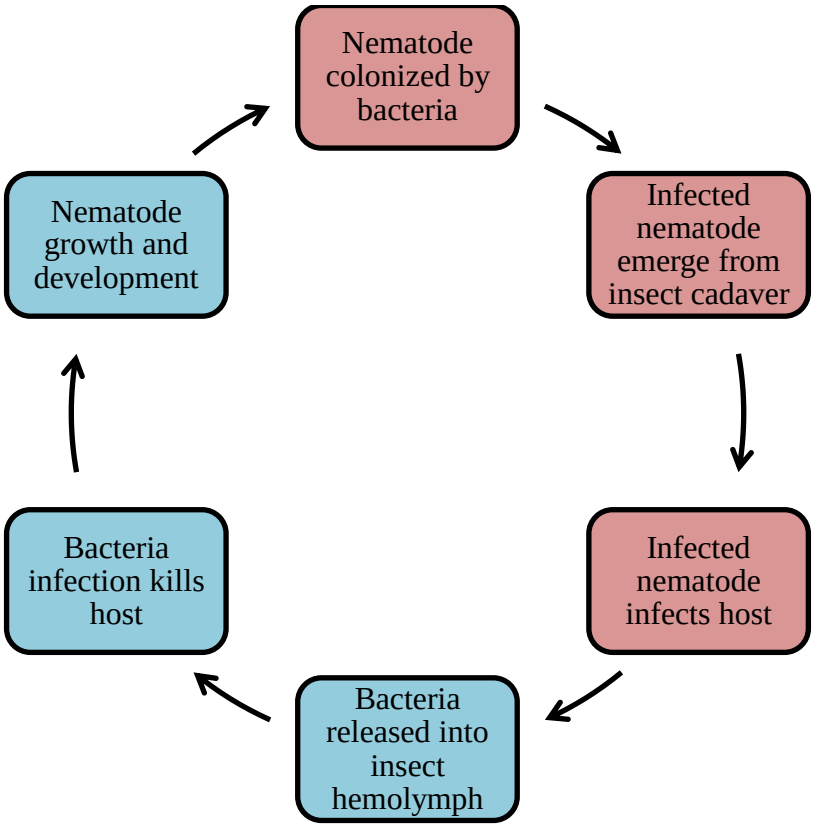


Figure 5. *Photorhabdus* sp. and nematode life cycle. Phase I stages are indicated by blue boxes, phase II by red.

Among the many described *Photorhabdus* spp., *Photorhabdus asymbiotica* has received less attention from chemists and biologists compared to its better known sister species *Photorhabdus luminescens* and *Photorhabdus temperata*, as well as closely related species in the genus *Xenorhabdus*.[48] While it is believed that *P. asymbiotica* engages in a life cycle similar to other *Photorhabdus* and *Xenorhabdus*, it is apparent that *P. asymbiotica* has an additional confirmed role as a human pathogen.[51,52] It has been reported that the genome of *P. asymbiotica*, contains a large number of virulence loci and secondary metabolite biosynthetic gene clusters.[53] The combination of an intriguing lifecycle and apparent capacity to generate a variety of secondary metabolites

makes *P. asymbiotica* an attractive organism to investigate due to its ecological impact and human health concerns.

The genome of *P. asymbiotica* ATCC43949 has been previously sequenced and annotated[54]. Our reexamination of the genome using the secondary-metabolite gene-cluster search tool antiSMASH[55] revealed 19 biosynthetic gene clusters (Table 1), which included a group of genes that was decidedly similar to the gene cluster attributable to glidobactin production [56] (Supporting Information, Figure S1). A recent publication had supported the hypothesis that a closely related *Photorhabdus* sp. was capable of producing glidobactin-like compounds.[57] Moreover, the same group recently verified the presence of a glidobactin gene cluster in *P. luminescens via* heterologous expression in *Escherichia coli* because this gene cluster was considered to be a "silent" biosynthetic pathway when the bacterium was cultured in the laboratory.[58] Knowing that the glidobactins are potent proteasome inhibitors[59], we used a combination of LC-MS and bioassay analysis to test for the presence of these compounds.

However, our initial tests with tryptic soy broth (TSB), a medium in which *P. asymbiotica* grows remarkably well, resulted in no LC-MS evidence for glidobactins and produced extracts that tested negative for cytotoxicity. Expanding on the basic TSB recipe, a variety of supplements and changes in the medium formulation were tested including: altering the amount of dextrose and peptone in the medium; the addition of soil extract, ground invertebrates, nematodes and sterile blood; and alternative carbon sources (sucrose, glucose and ribose). However, the ethyl acetate extracts of these cultures also proved negative for glidobactins and cytotoxicity. Consequently, several other media formulations were tested including Luria broth, potato dextrose broth, and a

defined medium used for the induction of secondary metabolites in *Streptomyces*.[60] In each case, the ethyl acetate extracts generated for cultures grown in these media demonstrated cytotoxicity. In addition, the culture prepared in the defined medium exhibited a golden-red hue that was quite distinctive from the colors of the all the cultures (pale yellow to cream colored) prepared using other media formulations. Analysis of the culture grown in defined medium by LC-MS led to the observation of several new peaks (Figure 6), including four substances that were indicative of glidobactins (based on their mass spectrometric and UV spectroscopic data). Note the appearance of glidobactin/luminmycin peaks appearing between 8-9 minutes (indicated by a box) in the defined medium and their absence in TSB. Refer to Figure 2 for an expansion of this region showing the retention times for each glidobactin/luminmycin metabolite. Subsequent bioassay-guided isolation led to purification and dereplication of two known compounds, glidobactin A (**1**)[61] and luminmycin A(**2**).[58] In addition, two presumably new (based on mass spectrometric analysis) glidobactin/luminmycin derivatives were also purified resulting in the chemical characterization of a new glidobactin analogue, luminmycin D (**3**). Unfortunately the other new metabolite degraded rapidly upon purification before its structure was determined.

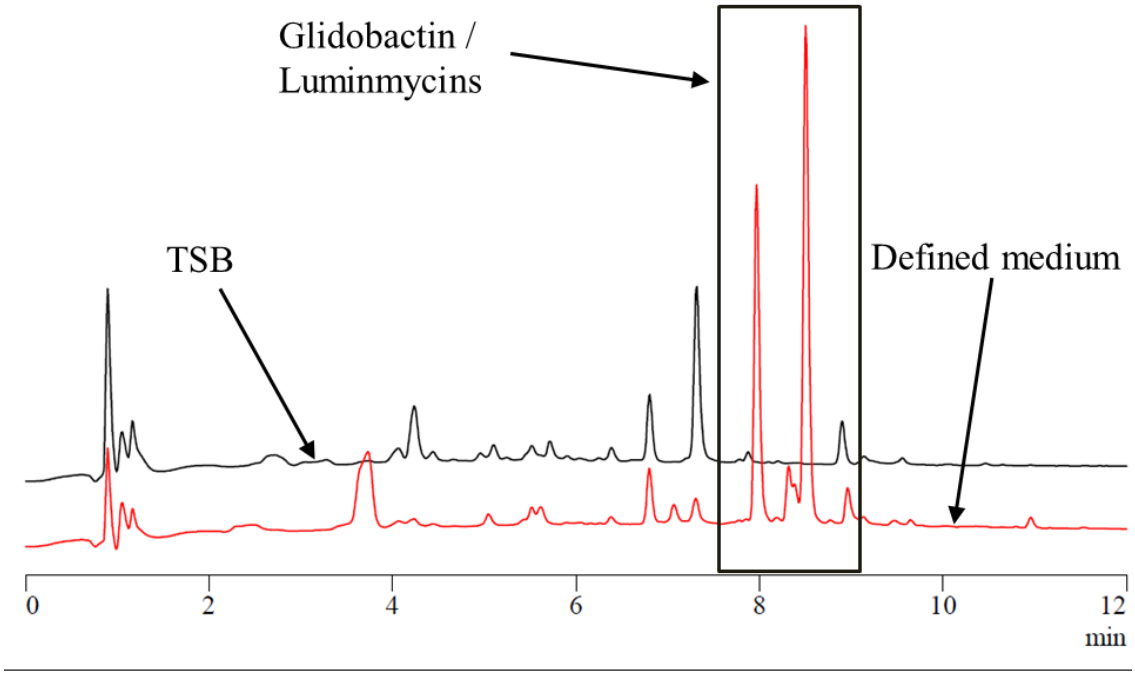
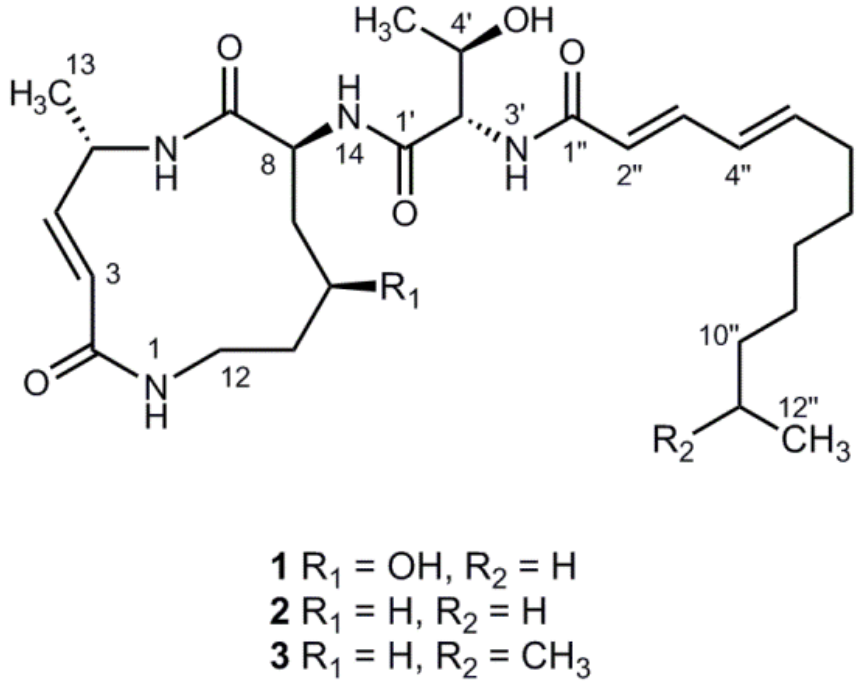


Figure 6. LC-MS traces (254 nm) of P. asymbiotica in TSB (red, top trace) and defined medium (black, lower trace).

Compound **3** was obtained as a white solid. High resolution electrospray ionization mass spectrometry (HRESIMS) analysis provided a $[M+H]^+$ peak at *m/z* 519.3543, which indicated that the metabolite possessed the molecular formula $C_{28}H_{47}N_4O_5$. The ^1H and ^{13}C NMR data were similar to the data collected for **2** with the majority of the carbon and proton chemical shifts throughout the 12-membered macrocycle and much of the polyketide tail remaining the same. However, integration of the ^1H NMR spectrum revealed a new methyl signal that was not present in **2**. In addition, the existing H-11" methyl protons were shifted downfield from 1.25 ppm to

1.50 ppm, while the ^{13}C NMR signal of its associated carbon atom was shifted from 22.1 ppm to 27.0 ppm. Inspection of the 2D NMR data for **3** helped unambiguously determine that the new methyl group was part of a terminal isopropyl unit. Since the specific rotation data for **1** and **2** were identical to those values previously[58, 62], and we presume that compound **3** shares the same biogenic origins, the absolute configurations is assumed to be the same as **2**. Double bond configurations were confirmed by coupling constants. Metabolite **3** has been given the trivial name luminmycin D. Structures are shown in Figure 6, NMR data of luminmycin D (**3**) in table 1.



1 R_1 = OH, R_2 = H
2 R_1 = H, R_2 = H
3 R_1 = H, R_2 = CH_3

Figure 7. Structures of isolated compounds, glidobactin A (**1**), luminmycin A (**2**), and luminmycin D (**3**).

Table 1. ^{1}H and ^{13}C NMR data. (500 Mhz, DMSO-d_6, 25 ^{o}C) for luminmycin D (**3**)

Luminmycin D (**3**)		
Position	δ_C, type[a]	δ_H (*J* in Hz)
2	166.5, CH	
3	118.2, CH	6.23, m
4	146.4, CH	6.78, dd (15.4,4.6)
5	45.3, CH	4.41, m
7	171.0, C	
8	51.1, CH	4.51, m
9	29.5, CH_2	2.07, m
10	29.4, CH_2	1.69, m
11	29.5, CH_2	1.44, m
12a	37.7, CH_2	3.20, m
12b		2.94,m
13	18.0, CH_3	1.20, d (7.0)
1'	169.3, C	
2'	57.9, CH	4.32, dd (8.8,3.9)
4'	66.2, CH	3.99, m
5'	19.4, CH	1.02, d (6.4)
1''	165.9, C	
2''	122.7, CH	6.17, m
3''	139.4, CH	7.00, dd (15.2, 11.2)
4''	128.3, CH	6.17, dd (15.2)
5''	141.6, CH	6.09, m
6''	31.9, CH_2	2.13, q (7.0,7.0,7.0)
7''	28.1, CH_2	1.37, m
8''	28.6, CH_2	1.28-1.25, m
9''	28.6, CH_2	1.28-1.25, m
10''	31.0, CH_2	1.28-1.25, m
11''	27.8, CH	1.5, m
12''	23.0, CH_3	0.85, d (6.8)
13''	23.0, CH_3	0.85, d (6.8)
NH (1)		7.34, dd (6.6)
NH (6)		8.45, d (7.8)
NH (14)		7.62, d (7.8)
NH (3')		8.01, d (8.8)
OH		4.94, d (4.4)

The distinct lifecycle variations experienced by *P. asymbiotica* during its symbiotic and pathogenic periods indicates that tight regulation of secondary metabolite production must be essential to the bacterium and its host nematode's survival.[48] Our observations pertaining to the marked variation in the *P. asymbiotica* metabolome in different media types appear to be consistent with the idea that the bacterium requires specific environmental cues to induce secondary metabolite production. Paralleling our studies, a recent report from the Müller group had asserted that the biosynthetic genes for glidobactins/luminmycins in *P. luminescens* were "silent," which led the authors to pursue a heterologous expression-based approach for their production.[58] The challenges of producing *Photorhabdus* metabolites highlighted in this report, coupled with our successful production of glidobactins/luminmycins under liquid-state fermentation conditions, led us to propose that an insect might provide an appropriate trigger for inducing the production of metabolites by *P. asymbiotica*. Accordingly, *P. asymbiotica* cells were grown in a non-glidobactin/luminmycin-inducing TSB medium and injected into live crickets. After 24 hours post-inoculation, we observed that all of the *P. asymbiotica*-infected crickets were dead. The insect carcasses were ground and extracted with ethyl acetate. Examination of the extracts by LC-MS confirmed that all four of the glidobactin/luminmycin metabolites produced by *P. asymbiotica* in the defined medium were also present in the cricket corpus (Figure 8). These data further reinforce the theory that insects can provide the requisite environment for stimulating glidobactin/luminmycin production as part of *P. asymbiotica*'s phase I growth strategy.

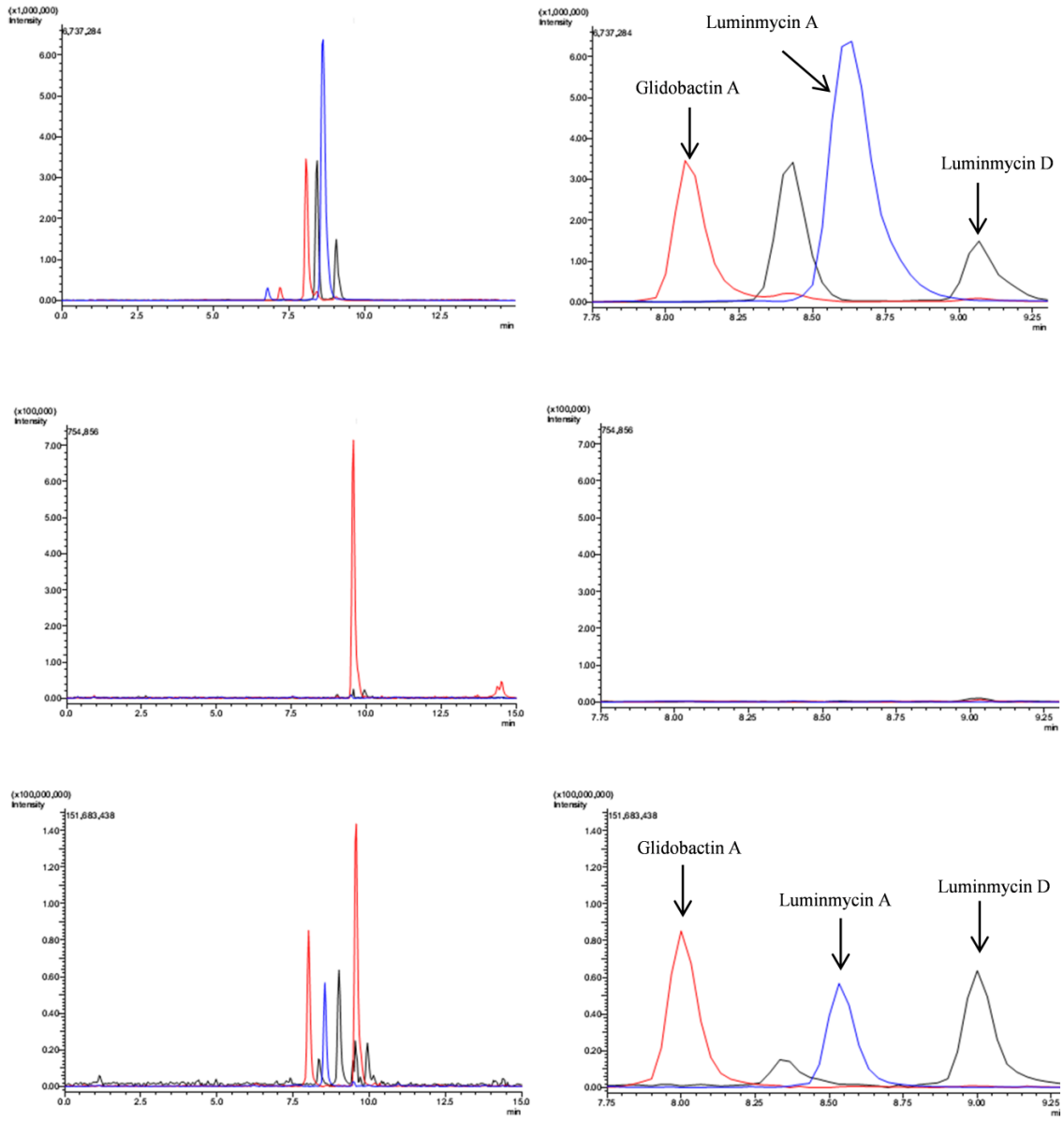


Figure 8. LCMS single ion trace profiles of various extracts. m/z 521 (glidobactin A) in red, m/z 519 (luminmycin D and unknown derivative) in black and m/z 505 (luminmycin A) in blue. (A) defined media (C) PBS injected cricket blank (E) crickets inoculated with *P. asymbiotica*. (B) (D) and (F) zoomed in to glidobactin/luminmycin fingerprint area of the adjacent trace, between 7.75 and 9.29 mins.

Metabolites **1**-**3** were evaluated for toxicity against human pancreatic cancer cells to assess their relative potencies (Figure 9 A). Compounds **1** and **2** exhibited similar activity profiles, although **1** was more potent with an IC_{50} value of 6.99 nM (4.35-11.2 95% confidence interval) compared to an IC_{50} value 17.9 nM (11.4-28.2 95% confidence interval) for **2**. In contrast, compound **3** was surprisingly much less active with an IC_{50} value of 105 nM (65.9-166 95% confidence interval). The compounds exhibited relatively less toxicity against mouse fibroblasts (Supplemental, Figure 9). Since **1** is known to inhibit the proteasomes,[59] all of the compounds were analyzed in a proteasome inhibition assay (Figure 9B). The results of these experiments showed a similar pattern of inhibition compared to the cancer cell cytotoxicity tests with **1** being slightly more active than **2** and much more active than **3** (IC_{50} values of 68.6, 87.6, and 375 nM, and 95% confidence intervals of 24.2-194, 58.5-131, and 322-437, respectively)

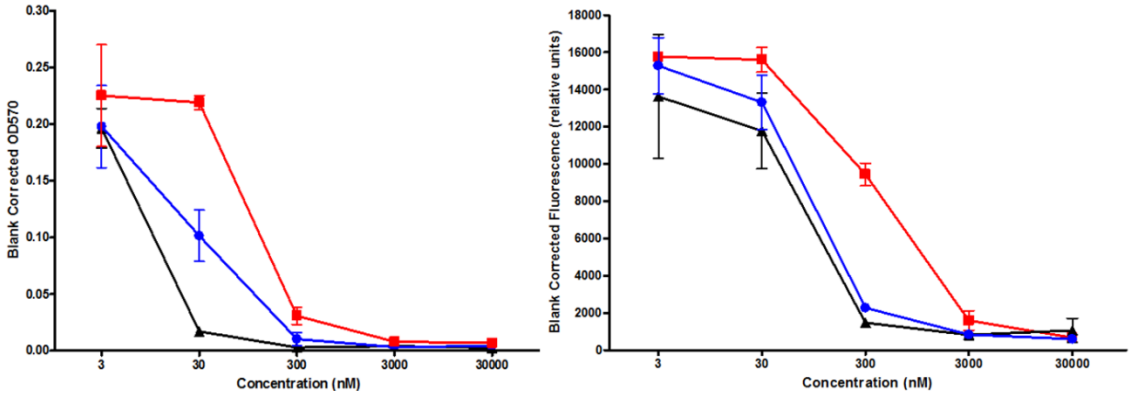


Figure 9. *In vitro* cytoxicity (panel A) and proteasome inhibition (panel B) of 1, 2, and 3 in MIA-PaCa-2. The data are shown as an average of triplicates with error bars representing the mean plus the standard deviation.

3.2 Experimental

3.2.1 General Experimental

UV data were collected on a Hewlett Packard 8452A diode array spectrophotometer. IR data were collected on a Shimadzu IRAffinity FTIR. Optical rotation data was determined on an AutopolIII automatic polarimeter. NMR data were collected on Varian VNMR 400 MHz NMR and Varian 500 MHz instruments. Accurate mass data were collected on an Agilent 6538 HRESI QTOF MS coupled with an Agilent 1290 HPLC. Metabolite profiles were obtained on a Shimadzu LCMS 2020 system (ESI quadrupole) coupled to a photodiode array detector. LC-MS samples were separated using a Phenomenex Kintex column (2.6 µm C_{18} column, 100 Å, 75 x 3.0 mm).

3.2.2 Microorganisms and culture conditions

Photorhabdus asymbiotica ATCC 43949 was purchased from ATCC (originally deposited as *Xenorhabdus luminescens*). For the media panel tests, starter cultures of the bacterium were prepared in TSB at room temperature on a rotary shaker at 120 rpm for 12 h. Flasks (1 L Erlenmeyer) containing 300 mL of autoclaved medium (refer to Supplemental Information for a list of the media tested and their formulations) were seeded with 1 mL of inoculum and allowed to grow for 1 week at 25 ºC on a rotary shaker at 120 rpm. For the scale up metabolite production, a starter culture of the bacterium was prepared in TSB at room temperature on a rotary shaker at 120 rpm. After 12 hours, the starter culture was used to inoculate a 20 L bioreactor containing the defined medium [60] [in $g \cdot L^{-1}$: KH_2PO_4 (2.0), NH_4Cl (1.5), $MgSO_4 \cdot 7H_2O$ (0.5), glycerol (10), myo-insitol (0.4), monosodium L-glutamate monohydrate (5.0), NaF (0.084),

$FeSO_4 \cdot 7H_2O$ (0.025), $ZnSO_4 \cdot 7H_2O$ (0.01), $CoCl_2 \cdot 6H_2O$ (0.01), $CaCO_3$ (0.25), *p*-aminobenzoate (0.001). No pH adjustments were performed. The culture was stirred at 20 rpm at 25 ºC for 1 week.

3.2.3 Extraction and isolation

A bioactivity guided fractionation procedure was followed. The scale-up culture was removed from the bioreactor and extracted three times with ethyl acetate in a 1:1 (vol.:vol.). The organic layers were removed, combined and dried under reduced pressure on a rotary evaporator. The reduced crude extract (3.46g) was absorbed onto silica and subjected to fractionation by flash chromatography using a hexane-CH_2Cl_2-MeOH gradient. The bioactive fractions (340 mg) were further fractionated by C_{18} HPLC. A MeOH-H_20 gradient (20% for 5 minutes, 20-100% over 50 minutes, 100% for 10 minutes) at flow rate of 10 mL/min was used to elute the glidobactin/luminmycin metabolites . This procedure provided an active fraction that was further purified by reverse phase semi-preparative HPLC (60% isocratic MeCN) to provide **1** (3.08 mg), **2** (2.6 mg), and **3** (0.6 mg).

Luminmycin D (3): white solid; ; $[\alpha]_D$ (c 0.06, MeOH) -14.5; UV(MeOH) λ_{max} 256nm (ε 20511); IR (film) ν_{max} 3421, 1643cm^{-1}, ^{1}H and ^{13}C NMR, see Table 1; HRESIMS $[M+H]^+$ 519.3543 (calcd for $C_{28}H_{46}N_4O_5$, 519.3541).

3.2.3 Productin of glidobactins and luminmycins in an insect host

Live crickets (*Acheta domestica*) were purchased from Timberline Fisheries through a local reseller. Inoculations of live crickets was performed based on modifications to a

published method.[63] The *P. asymbiotica* inoculum was grown in a non-glidobactin inducing TSB medium. Bacteria were obtained by centrifugation (8000 ×g), washed with PBS, and diluted to a standardized concentration so that we could deliver 10^8 cells in a 2 µL injection.[64] Crickets were briefly chilled on ice until they appeared lethargic and were immediately injected between the third and fourth abdominal segments with *P. asymbiotica*. Control crickets were injected with sterile PBS. The treated crickets were given access to water and food (corn meal) and left at 25 ºC in the light. After 24 h, crickets were flash frozen in liquid nitrogen, crushed with a mortar and pestle, and extracted with EtOAc. The EtOAc was dried under vacuum and resuspended in 9:1 MeOH-H_2O at a concentration of 10 mg/mL for LC-MS analysis.

3.2.4 Cell cytotoxicity assays

Mammalian cell cytotoxicity assays were performed on cancer cells (MIA PaCa-2) and mouse fibroblast cells (NIH/3T3) by adding 2,000 cells per well into 96-well plates and the cells were allowed to attach overnight at 37 ^oC in a humidified incubator (5% CO_2 atmosphere). Test compounds were diluted in DMSO and added to the wells so that the final concentration of DMSO per well did not exceed 1% by volume. The plates of treated cells were incubated for 72 hours and cell viability was determined by MTT assay.[65]

3.2.5 20S proteasome inhibition assay

The proteasome assay was performed using a colorometric commercial kit purchased from Cayman Chemical (item number 10008041). Briefly, cells were seeded

at a density of 50,000 cells/well and incubated overnight at 37 ^{o}C. Cells are treated with compounds for 90 minutes, the plate centrifuged, assay buffers were added, and the results read by a plate reader for fluorescent intensity (360 nm and 480 nm excitation and emission wavelengths, respectively). The assays were performed cancer cells (MIA PaCa-2) and mouse fibroblast cells (NIH/3T3).

Chapter 4. Genomic and Metabolomic Insights into the Natural Product Biosynthetic Diversity of a Feral Hog Associated *Brevibacillus laterosporus* Strain

4.1 Introduction

For decades, nature has served as a valuable source of bioactive compounds with many natural products (secondary metabolites) having entered into clinical use [4]. One of the primary reasons the supply of new and inspiring secondary metabolites has continued after such a long period of intense research is linked to the tremendous biological diversity of organisms from which these compounds are obtained [4]. Bioactive substances with unique chemical features have been discovered from plants, microbes, and animals originating from terrestrial and marine environments. In order to sustain the rich pipeline of new chemical entities from nature sources, researchers need to maintain a dedicated focus on investigating new biological resources [33].

Initial efforts to characterize animal microbiomes, especially those from vertebrate hosts, has uncovered a virtually untapped source of bacterial and archaeal diversity. These microorganisms fulfill a range of critical functions through their participation in commensal, symbiotic and pathogenic relationships with the animal hosts [32]. The spectrum of habitats afforded by a body's many distinct microenvironments substantially increases the number of microbial species that can inhabit a single animal. Bacteria associated with other microbes, plants, nematodes, insects and sponges produce an intriguing variety of secondary metabolites [36]. In contrast, relatively little is known about the natural products generated by the microbial symbionts of mammalian hosts. Larger mammals, such as humans, have a diverse

microbial population with some members capable of producing bioactive small molecules [66]. Secondary metabolites isolated from human microbial symbionts with antifungal [67], antibiotic [68], cytotoxic [69], anti-biofilm [70], and anti-tumor [71] properties demonstrate the biosynthetic potential of mammalian microbiome members. It is reasonably expected that the unexplored bacterial symbionts of smaller mammals will harbor a comparable assortment of biosynthetic gene clusters encoding for secondary metabolite production and that these compounds will exhibit a correspondingly wide range of promising biological activities.

In this report, we describe the use of an opportunistic sampling method to access secondary metabolites produced by a bacterium derived from a mammalian host. This work describes a new strain of *Brevibacillus laterosporus* acquired from the ear canal of a feral hog originating from southwestern Oklahoma, USA. The natural product biosynthetic potential of the bacterial isolate was revealed using a combination of bioassays, LC-MS, and genomic sequencing data. These efforts provided several compounds including a new and unusual peptidic metabolite auriporcine (**6**), a new pyrazine auripyrzine (**5**), and the previously described antifungal metabolites basiliskamides A and B (**1** and **2**, respectively). Our research highlights how the integrated application of metagenomic and metabolomic approaches offers exciting opportunities for mining mammalian microbiomes for new natural products.

4.2 Results and Discussions

The oral cavity, ear canal, and nasal cavity of a feral hog taken by a hunter in southwestern Oklahoma were swabbed for microbial inhabitants within 24 h of being

bagged. The samples were spread onto agar media and microbes purified by subculturing until bacterial isolates exhibited stable, uniform morphologies. The collection of isolates was transferred to fresh media and agar overlays seeded with *Staphylococcus aureus* were applied. Several isolates produced clear zones with no *S. aurerues* growth (Figure 10). One of the bioactive isolates derived from a swab sample taken from of the hog's ear (PE36) was selected for further investigation.

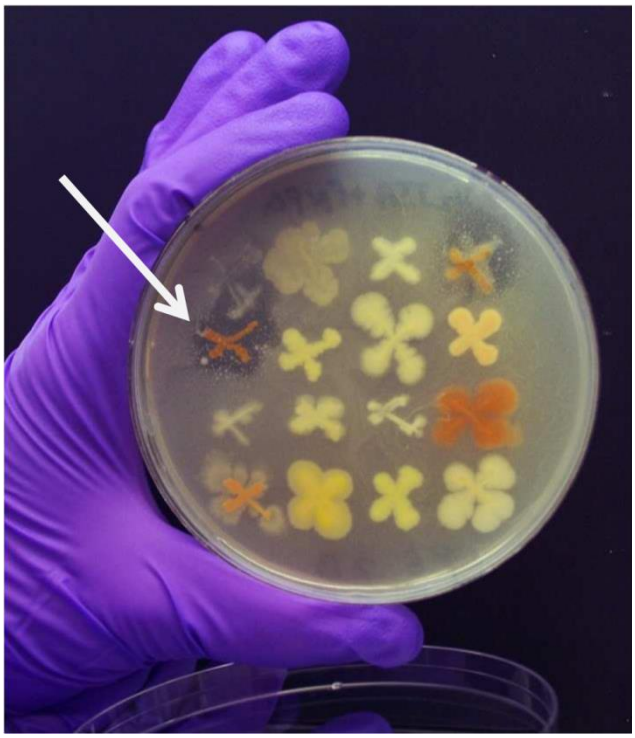


Figure 10. Overlay experiment of organisms isolated from a wild hog ear. Isolate PE36 indicated by an arrow

Genomic analysis of the isolate yielded a total of 16.75 million reads or 2.49 Gbp with 99.8 percent of reads passing quality filtering (>Q30 across a sliding window of 50bp). A total of 63 scaffolds with a genome size of 5.14 Mbp were obtained upon

assembly. These scaffolds had an N50 of 155.81 Kbp and a maximal scaffold length of 457.91 Kbp producing a near-complete genome assembly (489-fold coverage). After annotation, 4791 coding sequences were split between 430 subsystems along with 100 RNAs. The complete 16S small subunit rRNA sequence was used to identify isolate PE36 as a member of the family Paenibacillaceae and the genus *Brevibacillus* (Figure 11A). A sequence based comparison of the PE36 genome and those of related sequenced strains (i.e., *Brevibacillus laterosporus* strains LMG15441, GI-9, and DSM 25, as well as *Brevibacillus* phR) warranted the designation of PE36 as a separate strain of *Brevibacillus laterosporus* (Figure 11B). The new *Brevibacillus laterosporus* strain PE36 genome was submitted to the secondary metabolite analysis pipeline antiSMASH [72] revealing 32 putative biosynthetic gene clusters. These included 11 nonribosomal peptide synthase (NRPS) clusters, two polyketide synthase (PKS) clusters, and four hybrid pathways (Table 2).

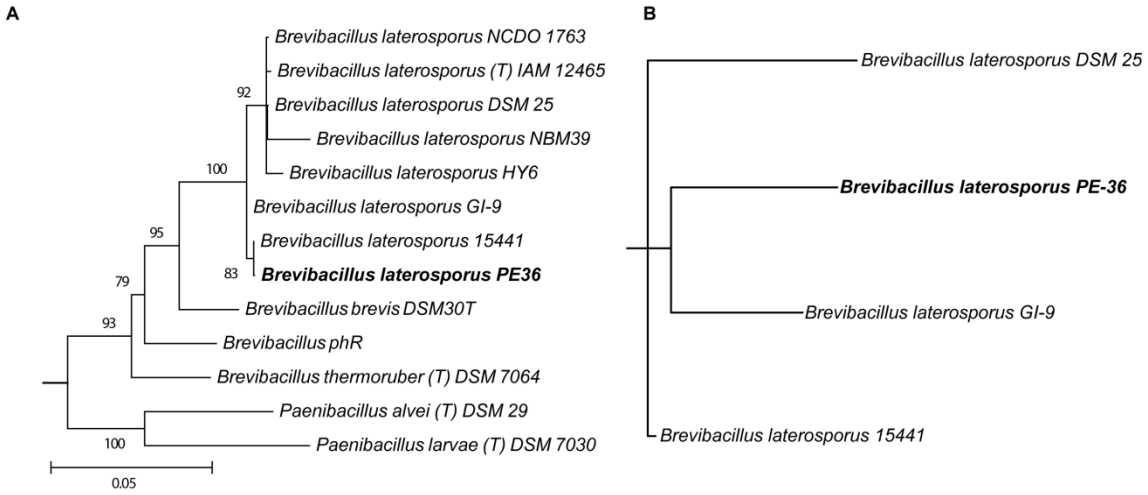


Figure 11. Phylogeny of *Brevibacillus laterosporous* strain PE36 and close relatives depicted by A) maximum likelihood of 16S rRNA gene sequences and B) a consensus network based whole genome comparisons. Bootstrap values >70% are listed for nodes in the maximum likelihood tree, with the scale bar representing 0.05 substitutions per position. Branch lengths for the genome consensus network are proportional to the number of likely inversion events of least collinear blocks (LCBs) inferred by Mauve.

Table 2**.** Biosynthetic gene clusters in PE36 identified by antiSMASH analysis. NRPS= nonribosomal peptide synthase, PKS = polyketide synthase, T1 = type I, T2 = type II, T3 = type III

Cluster #	**Type**
1	Bacteriocin
2	NRPS
3	NRPS/T1PKS/terpene
4	NRPS
5	NRPS
6	Putative biosynthetic gene cluster
7	Putative biosynthetic gene cluster
8	Putative biosynthetic gene cluster
9	Putative biosynthetic gene cluster
10	Siderophore
11	Bacteriocin/T2PKS/NRSP/*trans*-ATPKS
12	*Trans* AT-PKS
13	Other
14	NRPS
15	NRPS/*trans*-ATPKS
16	Putative
17	NRPS
18	Bacteriocin
19	Putative biosynthetic gene cluster
20	Putative biosynthetic gene cluster
21	Putative biosynthetic gene cluster
22	NRPS
23	NRPS
24	NRPS
25	NRPS
26	NRPS/*trans*-ATPKS
27	T3PKS
28	NRPS
29	NRPS
30	Putative biosynthetic gene cluster
31	Phosphonate
32	Putative biosynthetic gene cluster

Metabolites produced by strain PE36 were isolated and characterized from scale-up cultures for structure characterization. Silica flash chromatography, C_{18} VLC, and preparative and semipreparative C_{18} RP-HPLC were used to purify the metabolites (Figure 3) that were responsible for the extract's biological activities. One of the active fractions was determined to contain basiliskamides A (**4**) and B (**5**) [73] the structures of which were confirmed by ESIMS, UV and ^{1}H NMR data and ^{1}H-^{13}C-HSQC. The basiliskamides exhibit potent antifungal activities and modest antibacterial properties [74]. These compounds had been previously reported from a sample taken from a bacterium associated with a tubeworm collected in Papua New Guinea [73]. Another active (antibacterial) fraction contained 12-methyltetradecanoic acid (**6**), which was dereplicated by GC-MS and ^{1}H NMR.

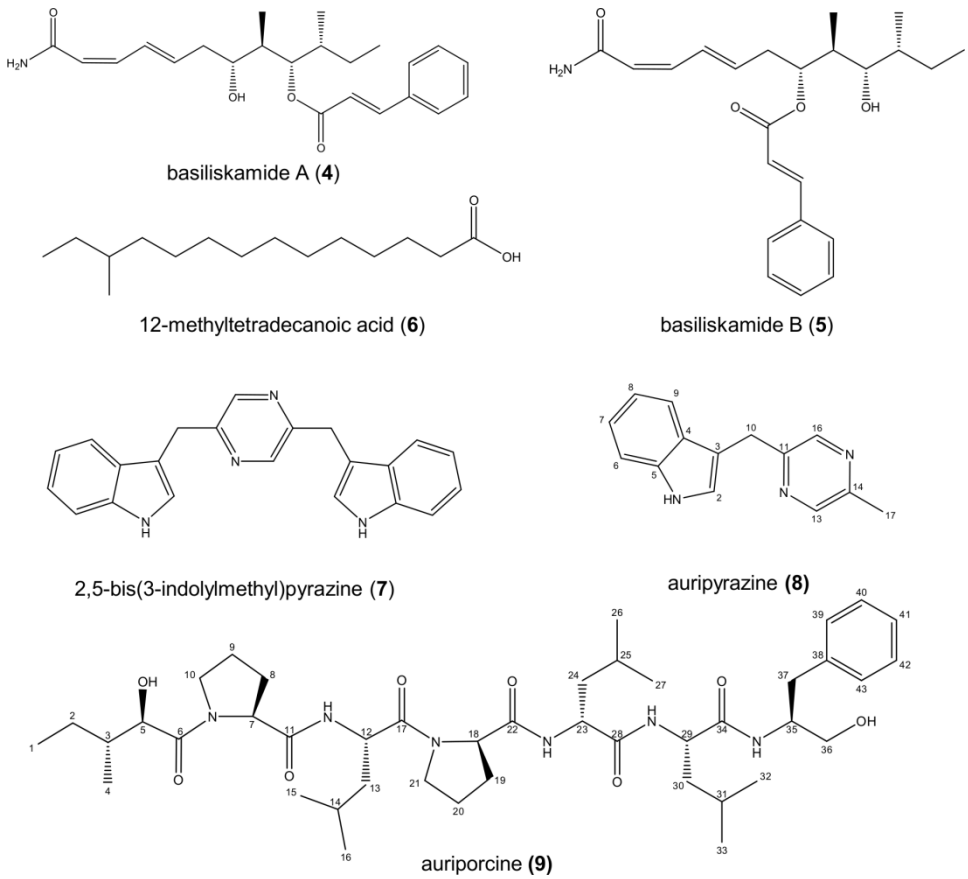


Figure 12. Isolated compounds.

To further access the chemical diversity of the PE36 metabolome, two additional fractions were selected for chemical analysis. LC-MS profiling of the first fraction revealed the presence of two compounds possessing similar UV adsorption spectra (λ_{max} ~220). One of the compounds was identified as 2,5-bis(3-indolylmethyl)pyrazine (**7**), whereas the second compound is reported here for the first time as a natural product. HRESIMS analysis of **5** yielded a molecular ion at *m/z* 222.1039 $[M-H]^-$, which substantiated a molecular formula of $C_{14}H_{13}N_3$. The ^1H and ^{13}C NMR data for **8** were remarkably similar to those obtained for **7** (Table 3). Analysis of the ^1H NMR data revealed a broad exchangeable singlet at δ_H 10.9, two downfield singlets at δ_H 8.32 and 8.36, five aromatic protons (δ_H 7.22 – 6.94), a singlet integrating for two hydrogens at

δ_H 4.17, and a methyl singlet at δ_H 2.45. These features could be accounted for if one of the two indoles attached to the pyrazine in **4** was replaced by a methyl group in compound **5**. The proposed structural change was subsequently confirmed by 2D NMR to give the structure of auripyrazine (**5**).

Table 3. NMR data for **8**

Position	**^{13}C**	**^{1}H (*J* in Hz)**
2	123.7, CH	7.22, m
3	111.1, C	-
4	125.7, C	-
5	136.1, C	-
6	111.2, CH	7.33, d (8.2)
7	119.0, CH	6.94, m
8	119.0, CH	7.48, m
9	121.5, CH	7.06, m
10	31.3, CH_2	4.17, s
11	155.4, C	-
12-N	-	-
13	142.3, CH	8.32, s
14	152.3, C	-
15-N	-	-
16	141.6, CH	8.35, s
17	21.8, CH_3	2.45, s
1-N*H*	-	10.92, brs

HRESIMS analysis of **9** provided a molecular ion with *m/z* 799.5330 [M+H], which indicated that the compound had a molecular formula of $C_{43}H_{70}N_6O_8$. An initial inspection of the ^{1}H NMR data revealed that **9** was likely peptidic in nature with signals characteristic of amide doublet protons, as well as amino-acid α-hydrogen spins. Further inspection of the 1D and 2D NMR data revealed the occurrence of many more signals than could be accounted for by the proposed molecular formula; however, additional attempts at purification of the compound by analytical HPLC proved futile.

We also observed by ^{1}H NMR that the relative distribution of these spins was not appreciably altered by switching to different solvent systems (DMSO-d_6, acetone- d_6, $CDCl_3$, MeOH-d_4, pyridine-d_5) or upon changes to the NMR probe temperature. Therefore, we conjectured that **9** might exist in two dominant conformational states that were relatively insensitive to the influence of the surrounding solvent.

Analysis of the ^{1}H-^{1}H TCOSY, ^{1}H-^{13}C-gHMBC, and ^{1}H-^{1}H NOESY data provided evidence for several discrete spin systems attributable to two prolines and three leucines that were linked to form a Pro-Leu-Pro-Leu-Leu peptidic fragment. Another spin set was probed in greater detail with ^{1}H-^{13}C-gHMBC leading to the generation of a six carbon fragment in which two of the carbon atoms were attached to oxygen atoms (δ_C 172.0 and 72.5, amide carbonyl and a hydroxyl-group-bearing methine, respectively) (supplemental Figure 1). Examination of the ^{1}H-^{13}C-gHMBC data helped linked this fragment to the peptide portion of the new metabolite (supplemental Figure 1). The placement of this six carbon fragment at the *N*-terminus of the peptide was further supported by an^{1}H-^{15}N HSQC experiment from which we found no evidence for primary amide hydrogens. The remaining carbon atoms in **9** were determined to comprise a phenylalaniol residue (phenylalaninol – Pho) at the *C*-terminus of the peptide (supplemental Figure 1). The reduction of the phenylalanine carboxyl group to a primary alcohol was confirmed by the presence of a new carbon chemical shift at δ_C 62.4 for the atom bearing the primary alcohol. A series of MS^n experiments (Table 5) were performed to provide additional substantiation for the proposed planar structure of compound **9**. A majority of the fragment ions exhibited a loss of water and/or putative intramolecular cyclization as reported for similar structures

[75]. Several of the fragments were generated through cleavage of the amide bonds, thus confirming the proposed sequence of amino acid residues in **9**. The assigned NMR shifts, parsed out between the two major conformations is shown in Table 4.

Table 4. NMR data for **9**, showing the data for the different conformations

Residue					
		^{13}C	^{1}H	^{13}C	^{1}H
MHP	1	11.1, CH_3	0.82	11.1, CH_3	0.82
	2	22.8, CH_2	1.52, m; 1.06, m	22.8, CH_2	1.51, m; 1.08, m
	3	37.1, CH	1.66, m	37.1, CH	1.67, m
	4	15.0, CH_3	0.86, m	15.0, CH_3	0.86, m
	5	72.5, CH	3.91, m	72.5, CH	3.94, m
	6	172.6, C	-	172.6, C	-
Pro	7	59.1, CH	4.84, m	59.1, CH	4.82, m
	8	31.5, CH_2	2.09, m	31.5, CH_2	2.09, m
	9	28.5, CH_2	1.75, m	28.5, CH_2	1.75, m
	10	46.2, CH_2	3.41, m	46.2, CH_2	3.41, m
	11	171.9, C	-	171.9, C	-
Leu	NH	-		-	
	12	51.3, CH	4.35, m	51.3, CH	4.36, m
	13	39.9, CH_2	1.45, m	39.6, CH_2	1.25, m
	14	23.9, CH	1.53, m	23.7, CH	1.47, m
	15	20.8-23.0, CH_3	0.86, m	20.8-23.0, CH_3	0.76, m
	16	20.8-23.0, CH_3	0.76, m	20.5, CH_3	0.66, m
	17	170.9-172.2	-	-	
Pro	18	60.0, CH	4.25, m	59.7, CH	4.34, m
	19	28.9, CH_2	2.05, m	28.9, CH	2.05, m
	20	24.1, CH_2	1.78, m; 1.91, m	24.1, CH_2	1.78, m; 1.91, m
	21	46.4, CH_2	3.67, m; 3.48, m	46.4, CH	3.65, m; 3.50, m
	22	171.6	-	171.6	-
Leu	NH	-	7.83, m	-	7.84, m
	23	48.6, CH	4.53, m	51.5, CH	4.17, m
	24	39.6, CH_2	1.45, m	39.6, CH_2	1.48, m
	25	23.6, CH	1.56, m	23.6, CH	1.51, m
	26	20.8-23.0, CH_3	0.87, m	20.8-23.0, CH_3	0.82, m
	27	20.8-23.0, CH_3	0.87, m	20.8-23.0, CH_3	0.87, m
	28	170.3, C	-	170.9-172.2, C	-
Leu	NH	-	8.00, d (8.3)	-	8.15, m
	29	51.6, CH	4.15, m	50.6, CH	4.21, m
	30	40.9, CH_2	1.41, m	40.0, CH	1.38 m
	31	23.7, CH	1.50, m	23.7, CH	1.45, m
	32	20.8-23.0, CH_3	0.84, m	20.8-23.0, CH_3	0.86, m
	33	20.8-23.0, CH_3	0.78, m	20.8-23.0, CH_3	0.79, m
	34	170.9-172.2, C	-	170.9-172.2, C	-
Pho	NH	-	7.57, d (8.8)	-	7.67, d (8.3)
	35	52.6, CH	3.86, m	52.6, CH	3.86, m
	36	62.4, CH	3.30, m; 3.25,m	62.36, CH	3.29, m; 3.21, m
	37	36.5, CH	2.86, m; 2.67, m	36.5, CH	2.85, m; 2.65, m
	38	139.0, C	-	139.0, C	-
	39	128.6, CH	7.21, m	128.6, CH	7.21, m
	40	127.7, CH	7.25, m	127.7, CH	7.25, m
	41	125.5, CH	7.17, m	125.5, CH	7.14, m
	42	127.7, CH	7.25, m	127.7, CH	7.25, m
	43	128.6, CH	7.21, m	128.6, CH	7.21, m

Table 5. MS^n fragment information for **9**. Pho = phenylalaninol. MHP = methylhydroxyl pentanone

Fragment	***m/z***
$[MHP\text{-}Pro\text{-}Leu\text{-}Pro\text{-}Leu]^+$ +	535
$[MHP\text{-}Pro\text{-}Leu\text{-}Pro]^+$	422
$[Pro\text{-}Leu\text{-}Leu\text{-}Pho]^+ + 2H$	457
$[Leu\text{-}Leu\text{-}Pho]^+ + 2H$	360
$[Leu\text{-}Pho]^+ + 2H$	247

The relative configuration of **9** was determined by X-ray diffraction experiment on a single crystal prepared in a vapor diffusion chamber (acetone and ether). Compound **9** approximated a helix-like conformation in the crystalline state with the leucine residue side chains projecting outward from the compound in a relatively disordered state (Figure 4). We observed a 3*S**,5*S**,7*S**,12*R**,18*S**,23*R**,29*S**,35*S** relative configuration for **9**, which indicated that some of the incorporated amino acid residues possessed a D-configuration. Marfey's analysis was subsequently carried out demonstrating that both of the proline residues were L-configured (Supplemental Figure 2). In contrast, we detected a ~1:2 mixture of the Marfey's-derivatized D and L leucines, which was in agreement with the proposed 12R*, 23R*, 29S* assignments for the three leucine residues detected in the crystal structure of 6 (Supplemental Figure 2). The Marfey's analysis also provided definitive evidence for the presence of L-phenylalaninol (Supplemental Figure 2). Thus, the absolute configuration of compound **9** was determined to be 3*S*,5*S*,7*S*,12*R*,18*S*,23*R*,29*S*,35*S*.

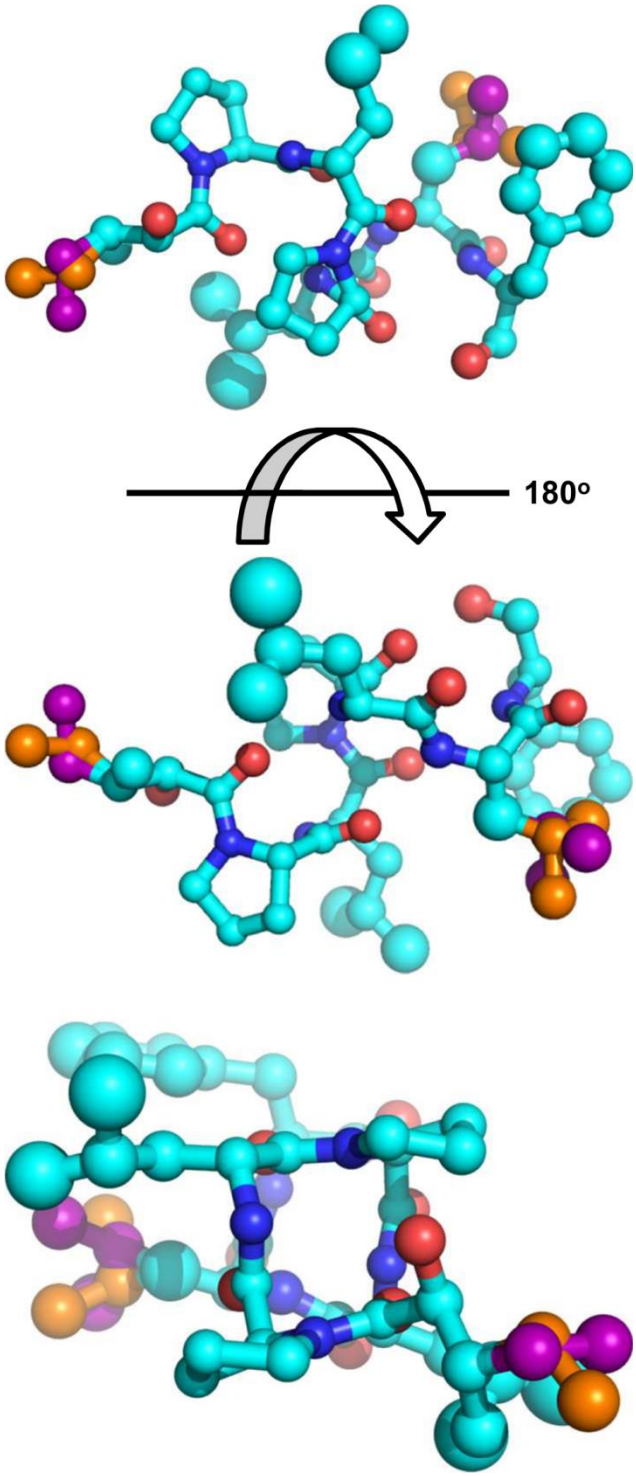


Figure 13. PyMol representations of **9**. The top two views are rotated 180°. The bottom is a view of the helical wheel, looking down the center of the helix from the N-terminal.

The genome of *B. laterosporus* strain PE36 was reexamined to provide a link between the secondary metabolites made by this isolate and their respective biosynthetic genes. Of the 11 NRPS gene clusters identified in PE36 (Table 2), cluster 23 (Figure 5a) exhibited several distinctive features making it the likely source of metabolite **9**. Namely, cluster 23 contains an initial PKS-like initiation module that we postulate is responsible for loading the non-amino acid starter unit at the *N*-terminus of the **9**. This is followed by NRPS modules containing adenylation domains that are predicted to sequentially incorporate proline → leucine → proline → leucine → isoleucine → leucine → tyrosine → leucine → isoleucine → threonine. The observed order of amino acids in **9**, L-proline → D-leucine → L-proline → D-leucine → L-leucine → L-phenylalaniol, fits fairly well with this prediction. Upon closer inspection, epimerization domains are associated with each of the predicted leucine incorporation steps. This explains the observed D-configurations of the first two leucines; however, the L-leucine origins are not clear. We hypothesize that either the epimerization domain associated with the incorporation of the third leucine is not active or the third leucine incorporation module is skipped with the downstream isoleucine incorporation module, which lacks an epimerization domain, serving to install the third leucine. The termination of the NRPS chain at the tyrosine/phenylalanine residue could occur via several scenarios including 1) phenylalanine incorporation followed by termination and post-production reductive tailoring , 2) tyrosine incorporation followed by termination, dehydroxylation, and reduction, or 3) full incorporation of all the predicted amino acids followed proteolysis of a tyrosine/phenylalanine → leucine bond.

Of the two predicted primarily PKS gene clusters (Table 2), biosynthetic cluster 12 is the most likely candidate for the production of **4** and **5**. Cluster 12 is a predicted *trans*-AT PKS that also contains an NRPS-like domain (based on antiSMASH prediction). *Trans*-AT PKSs are unique in a few key architectural and biosynthetic ways. Most notably, each module lacks an AT (acetyltransferase) domain; instead the module receives its carbon building blocks from a shared AT, in this case a fused tandem AT domain [76]. Figure 5B illustrates the proposed biosynthetic gene cluster and plausible biosynthetic pathway responsible for the generation of **4** and **5**. The starter united is loaded by a predicted acetyl-loading AT of the GCN5-related superfamily (GNAT). Each ketosynthase domain (KS) uses malonyl-CoA to add to carbon units to the growing polyketide chain. Ketoreductases (KR), enolreductases (ER) and methyl transferases (MT) derivatize the chain to add hydroxyl and methyl groups. The fourth module is split exhibiting a domain configuration of KS – KR – ACP – KS – DH – ACP – KR that is indicative of type B dehydrating bimodules. These types of modules are capable of producing double bonds with either the *E* or *Z* configurations [76].

The non-polyketide portions of **4** and **5** are likely introduced from three additional genes co-located with the gene cluster. One of these genes was identified as a putative aminotransferase that could have converted a carboxylate into an amid [77]. The cinnamic acid portion of the molecule is most likely added by an NRPS-like adenylation domain. Adenylation domains have previously been shown to incorporate non-amino acid moieties into compounds via ester linkages [78].

Mammalian microbiomes provide a unique opportunity to explore novel microbial biodiversity. As we have shown in this report, members of these microbial

communities contain the genetic machinery necessary to produce exciting natural products. Genomic characterization of a new strain of *B. laterosporus*, PE36, revealed a number of proposed secondary metabolite gene clusters. This information was used in conjunction with standard compound isolation and elucidation techniques to more fully explore PE36's biosynthetic capabilities. As the field of natural products advances, and the need to access new biodiversity increases, mammalian microbiomes will be a vital resource in the search for new natural products. We have demonstrated that these microbial communities are an exciting source of new secondary metabolites. With continued attention and experimentation, mammalian microbiomes will be a rich and varied source of new natural products.

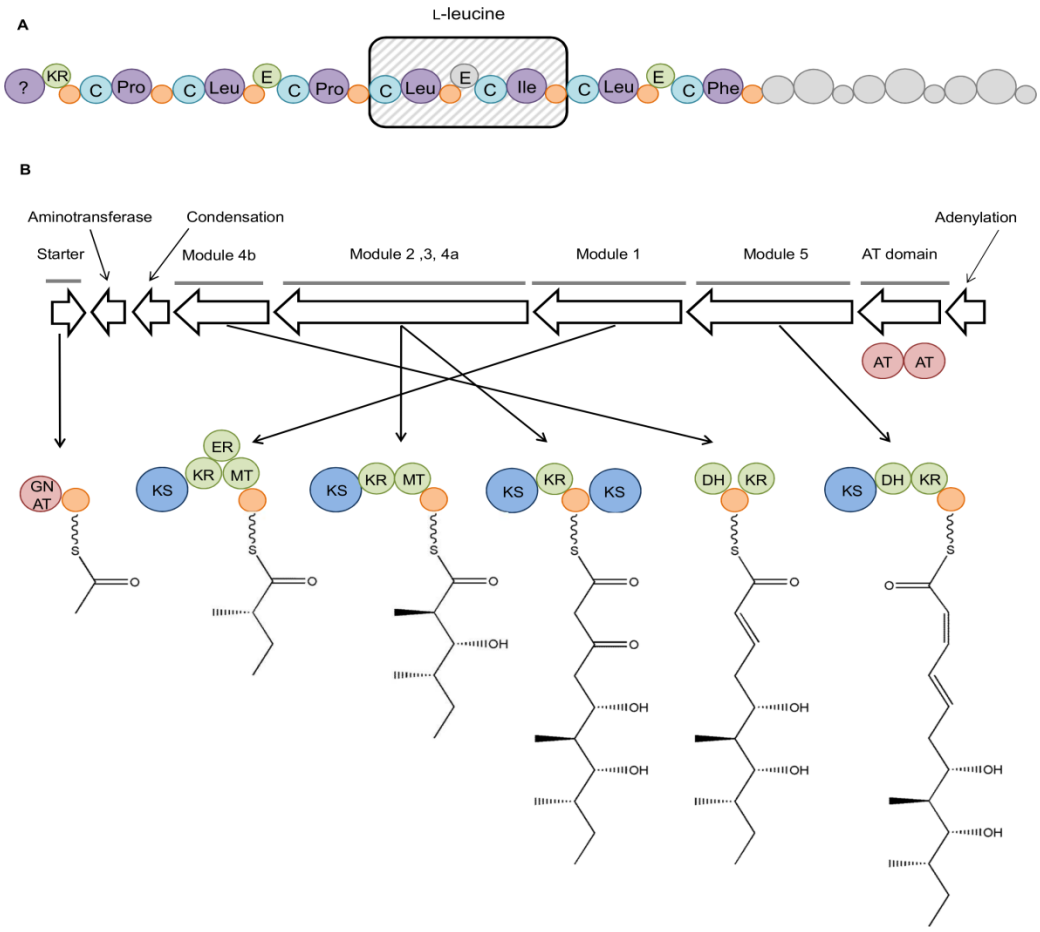


Figure 14. Graphical interpretation of antiSMASH results of auriprocine (**9**) and proposed biosynthesis of the polyketide portion of the basiliskamides A (**4**) and B (**5**). Panel A illustrates the biosynthetic gene cluster, cluster 12, responsible for the production 6. Adenylation domains are labeled with the amino acid they are predicted to contribute, with the N-terminal 6 carbon fragment's unknown origin indicated with a ? in the corresponding adenylation domain. The two domains leading to the uncertainty of the L-leucine addition are contained in the indicated box and the two proposed hypothesis illustrated. Either the epimerization domain is inactive and the second domain skipped, or the first domain is skipped and the predicted isoleucine adenylation domain is instad adding a leucine. The remaining encoded domains that are not believed

to have contributed to the peptide are shown in grey. Panel B shows the organization of the proposed basiliskamide gene cluster. The fused, tandem AT domains are shown separate from the rest of the modules. The dehydrating bimodule is shown split, as the gene that encodes the two KS domains are separated from the dehydrogenase and ketoreductase. The genes encoding the predicted aminotransferase, condensation, and adenylation domains (hypothesized to add the amide and cinnamic acid moieties to the polyketide chain) are labeled. A: adenylation C: condensation, KR: ketoreductase, E: epimerization, GNAT: GCN5-related acetyltransferase, KS: ketosynthase, ER: enolreductase, MT: methyltransferase, DH: dehydrogenase, AT: acetyltransferase

4.3 Materials and Methods

4.3.1 Collection and bacteria isolation

The oral, nasal and aural cavities of a recently deceased wild pig, shot by a hunter in southwest Oklahoma, were swabbed with a sterile cotton swab when the animal was brought in for processing at a local osteological processor in Oklahoma City, Oklahoma. The tip of each individual swab was removed and placed in a sterile 15mL falcon tube containing 0.9% NaCl in water (w/v) solution. The tube was vortexed and diluted 1:10 with the salt solution. 50 uL was spread on half strength TSA. The streaked plates were incubated at 37°C and 5% CO_2 for approximately two weeks. Individual colonies were picked based on color and morphological differences and re-streaked on ½ strength TSA plates until pure colonies were obtained.

4.3.2 Overlay assay

Bacterial isolates were arrayed onto half strength TSA plates and incubated for up to one week. A layer of half strength TSA seeded with *Staphylococcus aureus* was genetly applied to the surface of each plate of isolations. The plates were incubated at 37 $^{\circ}$C under a 5% CO_2 atmosphere for five days and periodically examined for zones of inhibited *S. aureus* growth.

4.3.3 DNA extraction, PCR amplification and sequence

Genomic DNA was extracted using the PowerBiofilm DNA Isolation kit (MoBio Laboratories) by spinning down 2.0mL of turbid culture and following the manufacturer instructions. The purity of gDNA was confirmed by spectrophotometery (Implen) and submitted to the Oklahoma Medical Research Foundation (OMRF) to be sequenced on a Illumina MiSeq using TruSeq LT 2x150bp chemistry. Reads were assembled using the CLC Genomics Workbench suite *de novo* assembly algorithm. Contigs smaller than 800bp were discarded. After assembly, scaffolded contigs were submitted to the RAST server for total-genome annotation [79]. These scaffolds were also uploaded to antiSMASH [72] to identify putative pathways associated with the production of secondary metabolites. Pathway annotations identified by antiSMASH were amended to the RAST annotation prior to submission to GenBank. Annotated gDNA and 16S rRNA sequences were were deposited with GenBank (submission is currently in review).

4.3.4 Culture conditions and extractions

A starter culture of the bacterium was prepared in tryptic soy broth (TSB) and shaken on a rotary shaker at 130 rpm for 12hr. One liter flasks (Erlenmeyer) containing 300 mL of autoclaved TSB were inoculated with 1 mL of starter culture. The flasks were shaken at 130 rpm at room temperature for 1 week. The culture was transferred to a glass extraction vessel for partitioning. The culture was partitioned three times against EtOAc (1:1 vol/vol) and the solvent removed from the organic layer under reduced pressure on a rotary evaporator.

4.3.5 Compound isolation

The crude extract (approximately 4g from 30L culture) was absorbed onto silica and subjected to fractionation by flash chromatography using a hexane-CH_2Cl_2-MeOH gradient. Isolation of each compound was as follows:

The crude extract (~4g from 30 L culture) was absorbed onto silica gel and subjected to flash chromatography fractionation using a hexane-CH_2Cl_2-MeOH gradient. The final purification of each compound was achieved as follows:

*12-methyltetradecanoic acid (**6**)*: A silica flash column was run (Hexane:DCM:MeOH) and the fraction that eluded with 100% DCM was further separated using C_{18} HPLC. A MeOH-H_2O gradient (20% MeOH for 5 minutes, 20-100% MeOH gradient over 50 minutes, 100% MeOH for 10 minutes at a 10 mL/min flow rate) resulted in a single active fraction. This fraction was further purified by semi-preparative HPLC (60%

ACN for 10 minutes, 60-100% ACN gradient over 45 minutes, 100 % ACN for 10 minutes at 2 ml/min) to provide **6**. The structure of the metabolite was confirmed by ^{1}H NMR and GC-MS analysis.

*Basiliskamides A (**4**) and B (**5**)*: C_{18} VLC (75:25, 50:50, 25:75, 100% Water:MeOH) was run and the fraction that eluded with 100% MeOH was further separated using C_{18} HPLC. A MeOH-H_2O gradient (20% MeOH for 5 minutes, 20-100% MeOH gradient over 50 minutes, 100% MeOH for 10 minutes at a 10 mL/min flow rate) resulted in a single active fraction. This fraction was further purified by semi-preparative C_{18} HPLC (isocaratic, 80% MeOH) to yield **4** and **5**. The compounds were identified based on dereplication of their published MS, UV and NMR data [73].

*2,5-bis(3-indolylmethyl)pyrazine (**7**)*: An HPLC fraction that eluted following **4** and **5** was further purified by semi-preparative C_{18} HPLC (60% MeOH for 5 minutes, 60-100% MeOH over 50 minutes, 100% MeOH for 10 minutes at a 10 mL/min flow rate) to provide **7**. The structure of **7** was confirmed based on dereplication of its MS and NMR data [80]

*Auripyrazine (**8**)*: A fraction eluting from HPLC prior to **4** and **5** was further purified by preparative C_{18} HPLC (20% MeOH for 5 minutes, 20-100% MeOH gradient over 50 minutes, 100% MeOH for 10 minutes at a 10 mL/min flow rate). Further purification by semi-preparative C_{18} HPLC (isocratic, 40% acetonitrile with 0.1% formic acid) provided purified **5.**

Auriporcine (9): A very late eluting fraction from the same HPLC purification step that yielded **4** and **5** was pursued for further analysis. Semi-preparative C_{18} HPLC with acetonitrile and H_2O treated with 0.1% TFA (10% acetonitrile for 5 minutes, 10-100% acetonitrile gradient over 55 minutes) provided purified **9**.

4.3.6 Marfey's analysis of auriporcine

Marfey's analysis was performed as described in [81], with slight modifications. Briefly, approximately 1 mg of **9** was dissolved in 6M HCl and heated at 100°C for approximately 18 h. After heating, the sample was dried under reduced pressure, redissolved in 50 µL of water and transferred to a 1.5 mL centrifuge tube. 50 µL of 50 mM $_{D}$ and $_{L}$-enantiomenrs of each amino acid of interest was aliquoted into 1.5 mL centrifuge tubes. To each standard and sample, 100 µL of 1% FDAA and 20 µL of 1M $NaHCO_3$ was added. Tubes were capped and heated in a 40°C water bath, with periodic mixing, for 1 h. After cooling to room temperature, 10 µL of 2M HCl was added to each tube. The mixtures were dried under reduced pressure and redissolved in 200 µL of 90:10 MeOH:Water. Samples were diluted 1:10 with 90:10 MeOH:water for LCMS analysis.

4.3.7 X-ray experimental information

A colorless prism-shaped crystal of dimensions 0.440 x 0.160 x 0.120 mm was selected for structural analysis. Intensity data for this compound were collected using a diffractometer with a Bruker APEX ccd area detector and graphite-monochromated Mo

Kα radiation (λ = 0.71073 Å). The sample was cooled to 100 K. Cell parameters were determined from a non-linear least squares fit of 6210 peaks in the range 2.41 < θ < 26.04°. A total of 40661 data were measured in the range 1.554 < θ < 28.344° using ϕ and ω oscillation frames. The data were corrected for absorption by the empirical method giving minimum and maximum transmission factors of 0.964 and 0.990. The data were merged to form a set of 10800 independent data with R(int) = 0.0586 and a coverage of 99.9 %.

The orthorhombic space group $P2_12_12_1$ was determined by systematic absences and statistical tests and verified by subsequent refinement. The structure was solved by direct methods and refined by full-matrix least-squares methods on F^2. The positions of hydrogens bonded to carbons were initially determined by geometry and were refined using a riding model. Hydrogens bonded to nitrogens and oxygens were located on a difference map, and their positions were refined independently. Non-hydrogen atoms were refined with anisotropic displacement parameters. Hydrogen atom displacement parameters were set to 1.2 (1.5 for methyl) times the isotropic equivalent displacement parameters of the bonded atoms. A total of 573 parameters were refined against 118 restraints and 10800 data to give wR(F^2) = 0.1925 and S = 1.011 for weights of w = 1/[σ2 (F^2) + (0.0900 P)^2 + 5.0000 P], where P = [F_o^2 + 2F_c^2] / 3. The final R(F) was 0.0707 for the 9029 observed, [F > 4σ(F)], data. The largest shift/s.u. was 0.012 in the final refinement cycle. The final difference map had maxima and minima of 0.315 and -0.374 e/Å^3, respectively. The absolute structure could not be determined by refinement of the Flack parameter.

4.3.8 Compound characterization

Auripyrazine (**4**): yellow solid; UV (MeOH) λ_{max} 222 (log ε 3.49); ^{1}H and ^{13}C data see Table 2; HRESIMS $[M-H]^-$ *m/z* 222.1039 (calculated for $C_{14}H_{12}N_3$, 222.1031).

Auriporcine (**5**): white solid; UV (MeOH) λ_{max} 206 (log ε 4.19), $[\alpha]_D$ (*c* 0.065) -27.7; ^{1}H and ^{13}C data see Table 3; HRESIMS $[M+H]^+$ *m/z* 799.5330 (calculated for $C_{43}H_{71}N_6O_8$, 799.5333).

4.4 General experimental conditions

UV data were collected on a Hewlett Packard 8452A diode array spectrophotometer. Optical rotation data was determined on an AutopolIII automatic polarimeter. UV–CD spectra were measured on an AVIV circular dichroism spectrometer model 202-01. NMR data were obtained on Varian VNMR spectrometer (500 mHz). Accurate mass data were collected on an Agilent 6538 HRESI QTOF MS coupled with an Agilent 1290 HPLC. LCMS analysis were performed on a Shimadzu LCMS 2020 system (ESI quadrupole) coupled to a photodiode array detector, samples were separated using a Phenomenex Kintex column (2.6 µm C_{18} column, 100 Å, 75 x 3.0 mm). The HPLC system utilized SCL-10A VP pumps and system controller with a Gemini 5 µm C18 column (110 Å, 250 × 21.2 mm, flow rates of 1 to 10 mL/min). X-ray data were collected using a diffractometer with a Bruker APEX ccd area detector and graphite-monochromated Mo Kα radiation (λ = 0.71073 Å)

Chapter 5. A high-throughput screening procedure of mammalian microbiome isolates for secondary metabolite production, antimicrobial activity and cancer cell cytotoxicity

5.1 Introduction

5.1.1 Accessing new biodiversity in the search for new natural products

Natural products are widely considered the future of drug discovery.[33] These small molecules, isolated primarily from marine and terrestrial derived plants, fungi and bacteria, produce and inspire a staggering number of current pharmaceutical products. [4] In order to continue to have such a strong impact on pharmaceutical research, and entice pharmaceutical companies away from a more synthetically driven approach,[4] there exists a continual need to access new biodiversity. Despite being a vital part of discovering natural product chemistry, [35] access to new biodiversity is often a challenge.

This need to access new biodiversity, and overcome some common issues in natural products research, has led us to consider exploring the mammalian microbiome as a source of novel microorganisms. Animal microbiomes are a virtually unexplored source of bacterial and archaeal diversity. Bacteria associated with other bacteria, plants, fungi, nematodes, insects and sponges produce an array of interesting and biologically active secondary metabolites. [36] Bacterial symbionts of larger mammals, such as humans, have been shown to produce a variety of natural products,[66] and we can expect the microbial communities of smaller mammals to be similarly biosynthetically talented. In this case, investigating the mammalian microbiome as a source of new natural products is an attractive solution to the inherent problems of sample collection.

The challenge is effectively isolating and screening these organism for secondary metabolite production.

5.1.2 Screening approach for biosynthetically talented bacteria

Screening of crude extracts in our in-house assay procedure has proven to be an effective method for identifying biosynthetically talented organisms. LCMS traces of crude extracts were considered in tandem with *in vitro* activity against methicillin-resistant *Staphylococcus aureus* (MRSA), *Candida albicans*, and a human pancreatic cancer cell line. Together, this information created a comprehensive data set to determine whether to pursue an organism for further work. The organizational goals for the screening project were to establish a method that quickly and efficiently identified organisms worth pursuing further and minimized hands-on time for a researcher. These goals were accomplished by biasing the method towards the most productive organisms. This was accomplished by limiting the liquid media to only three types: one nutrient rich medium (tryptic soy broth) and two defined media with minimal nutrients. The minimal media were chosen to stress the organism by providing a minimal amount of nutrients necessary for it to survive and induce the production of secondary metabolites.

The downside to this approach is the potential to miss interesting secondary metabolites by selecting for the organisms that are easiest to grow in liquid culture. However, new isolates are cheap and comparatively easy to obtain. It is more efficient to collect additional isolates that easily produce metabolites in the lab than it is to tease secondary metabolites out of each organism that is isolated. By limiting the media and

growth conditions, the focus is shifted towards organisms with a greater potential to produce more metabolites in lab conditions.

Finally, efficiency was also addressed through the use of a liquid handling robot to perform the extraction procedure on the liquid bacteria cultures, as shown in Figure 15. When coupled with automated solvent drying systems, the hands on time required by a researcher is reduced from hours to minutes. The liquid handling robot is programed to add EtOAc to each culture tube and allow each sample culture to extract for one hour. A conductivity sensing feature is used to remove only the organic layer and deposit the solvent into a clean 96 well plate. By using the liquid handling robot during the extraction phase of this protocol, a group could extract 300-400 cultures in a typical 8 hour work day.

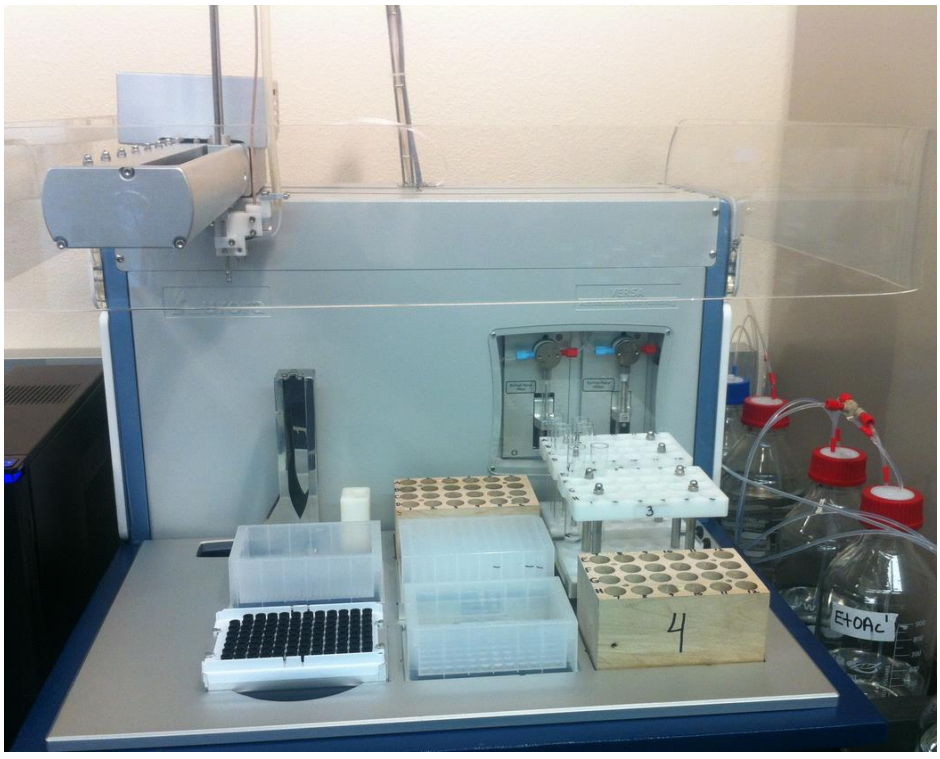


Figure 15. Automated liquid handling robot from Aurora Biomed

5.1.3 Purification of organisms

While the mammalian microbiome has proven to be a rich and diverse source of microorganisms, there will inevitably be some organisms that grow faster than others. These organisms tend to spread out quickly, monopolize the space on a petri plate and decrease the diversity of organisms on any given plate. Ultimately, the isolation conditions were chosen to inhibit some of these faster growing organisms. By using hypoxic conditions during the initial isolation phase, a greater diversity of organism types was seen. By choosing one growth condition, we improved efficiency but potentially limited our final results. From our experiments, the benefits of this initial plate diversity outweighed the chance of missing a select few organisms and improved overall efficiency.

Two plate types were used in this protocol for the initial plating of microbiome samples, 10% strength TSA and BHI, with the salt concentrated adjusted to the levels of a full strength mixture. Individual colonies were chosen for further purification based on unusual characteristics such as color, growth patterns, and incubation period before growth is observed. Isolates were restreaked onto the similar plate types for further purification. The media used for an organism's isolation was noted and used for later grow ups for DNA extraction and cryogenic storage.

5.1.4 Selection of animals for sampling

It should be noted that selection of the mammals to study is extremely important. Since the animals being sampled were already deceased by other causes,

there existed the potential of sampling opportunistic environmental bacteria instead of the natural microbial community of the animal. For animals that have died due to being hit by vehicles, it is best to sample in the hours just after sunrise. Nocturnal animals were most likely hit no more than a few hours prior. In warmer months, early morning collection ensures microbial colonization by other organisms is not sped up by ambient heat.

Visual inspection of the animal was necessary, and carcasses with bloating, insect infestation or removal of flesh from scavengers were not chosen for sampling. Additionally, a manner of death could sometimes be deduced by looking at the corpus. Animals that expired due to substantial internal injuries to the gut were rejected as candidates for intestinal sampling. The ideal carcass for microbiome sampling would have all the visual cues of a recent death and minimal damage to the main body cavity. Excessively bloodied orifices were not sampled. Figure 16 demonstrates ideal sampling conditions. These steps ensured that the samples collected were representative of the native microbial community and not of opportunistic environmental organisms colonizing the animal after its death.

Finally, the legality of collecting deceased roadside animals was investigated for our sampling area. Collection of samples required a permit from the town government. Other sampling locations may have similar requirements.



Figure 16. A skunk (top panel) is sampled roadside. An armadillo (bottom panel) was opened to access the large and small intestines. Both carcasses have all visual signs of a recent death. The main body cavities are intact and there is minimal bleeding from sampled offices.

5.1.5 Interpretation of LCMS data

Evaluation of the LCMS data required the largest amount of hands on involvement. Accurate interpretation of the generated spectra is vital to the success of

this protocol. The PDA and mass spectrum data was evaluated in tandem to determine whether an organism is biosynthetically talented. Typically, spectra that are dramatically different than the baseline chromatograms of the media control are easily identified. As previously described, the objective of this method is to quickly identify the most prolific natural product producers, and not be concerned with missing less productive organisms. As shown in Figure 17, the most talented organisms will produce crude extracts with chromatographic profiles that are clearly different than controls and other less talented organisms.

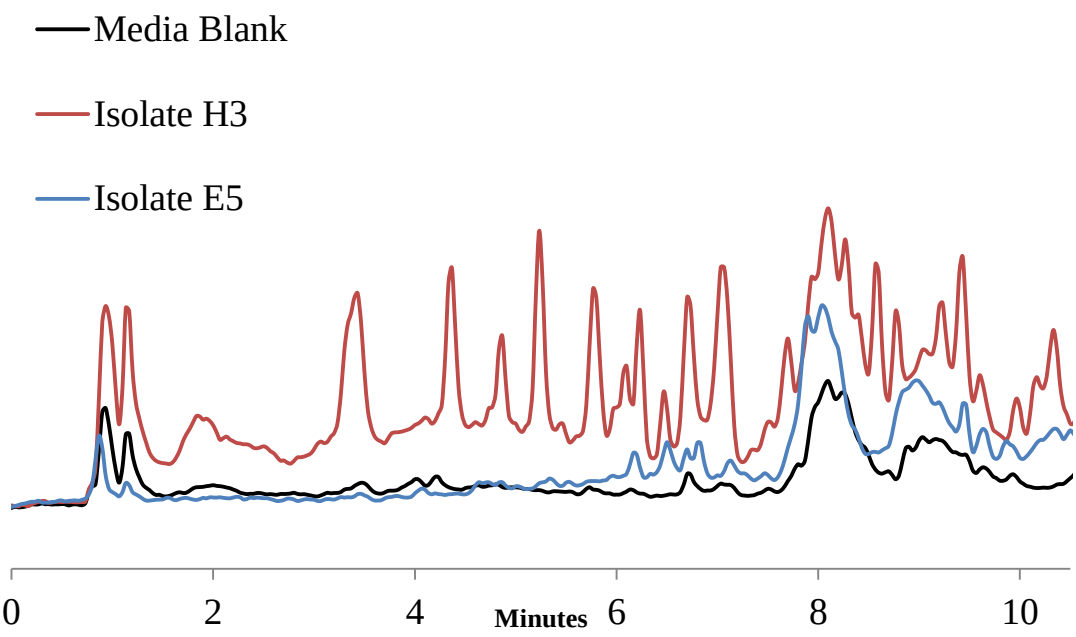


Figure 17. Illustration of generated LCMS data. A biosynthetically talented strain (red trace) compared to a less talented strain (blue trace) and the media control (black trace). Chromatogram shown 0-10 minutes for ease of viewing.

5.1.6 Controls

Controls are relatively straightforward for this procedure. As space permitted, 1-3 tubes of each liquid media type were left un-inoculated. These tubes were placed on the shaker with the inoculated broths. The culture blanks remained clear with no signs of microbial growth, indicating contamination of the samples by environmental organisms was unlikely. The culture blanks were than extracted along with the inoculated cultures and used as a comparison during chromatographic analysis.

During bioactivity assays, chloramphenicol and fluconazole (100µg/mL each in DMSO) were used as positive controls to qualitatively determine the level of bioactivity seen in a crude extract. DMSO, the solvent used to dissolve each crude extract and the antimicrobial controls, was used as a negative control.

5.2 Materials

5.2.1 Reagents

All reagents are cell culture grade unless otherwise stated.

- Water (deionized)
- EtOAc
- DMSO
- Tryptic Soy Broth (TSB), premade powder
- Brain heart infusion (BHI), premade powder
- Agar
- KCl
- Na_2HPO_4

- KH_2PO_4
- NH_4Cl
- $MgSO_4 \cdot 7H_2O$
- Glycerol
- *Myo*-inositol
- Monosodium $_L$-glutamate monohydrate
- NaF
- $FeSO_4 \cdot 7H_2O$
- $ZnSO_4 \cdot 7H_2O$
- $CoCl_2 \cdot 6H_2O$
- $CaCO_3$
- *P*-aminobenzoate
- $_L$-Glutamine
- K_2HPO_4
- Chloramphenicol
- Fluconazole
- Acetonitrile (LCMS grade)
- Water with 0.1% formic acid (LCMS grade)
- Methanol (LCMS grade)
- *C. albicans* cell line ATCC 38245, frozen in 20% glycerol
- *S. aureus* cell line ATCC 700787, frozen in 20% glycerol
- MIA PaCa-2 cells ATCC CRL-1420
- YM media

- MTT solution and solvent[74]

5.2.3 Equipment

- Autoclave
- Balance
- Stir plate
- Magnetic stir bar
- pH meter
- Sterile swaps
- Large scalpel
- Small scalpel, sterile
- 15 mL polypropylene tubes, sterile
- Vortexer
- Pipettes (capable of pipetting 200, 100 and 50 µL)
- Appropriate pipette tips, sterile
- Hypoxic chamber
- Petri dishes, 100x15 mm, sterile
- Culture tubes, 16x100 mm, sterile
- Caps for 16x100 mm tubes
- Rotary Shaker, with rack capable of holding 16x100 mm culture tubes
- Deep well 96 well plates
- Evaporator capable of drying 96 well plates under reduced pressure (Genevac, EZ-2 series)

- Cryogenic tubes
- Centrifuge, capable of centrifuging at 10,000 x *g*
- Refridgerator, 4°C
- Freezer, -80°C
- Sonicator
- 250 mL Erlenmeyer flask
- 250 mL separatory funnel
- 250 mL round bottom flask
- Rototary evaporator
- Scintillation vials
- Speedvac evaporator
- LCMS system (Shimadzu 2020, ESI quadrupole, photodiode array detector, autosapler with 96-well plate tray, Phenomenex Kinetex column, 2.6 µm C_{18}, 100Å, 75 x 3.0 mm)
- Incubator (37°C, 5% CO_2, humidified)
- Plate reader (capable of determining OD_{600})
- 96-well plate for biological assay, capable of being used in plate reader

5.2.4 Reagent setup

Phosphate buffer solution (PBS)

Combine the following with 1L of water, all eights in grams: NaC (8), KCl (0.2), Na_2HPO_4 (1.78), KH_2PO_4 (0.27). Stir on a magnetic stir plate with a magnetic stir

bar until dissolved. Adjust pH to 7.4. Dispense 5 mL into 15 mL conical tubes and autoclave. Allow to cool to room temperature.

TSB Broth

Prepare full strength TSB according to the manufacturer's instructions by mixing 30g of product and 1L of distilled water. Mix on a magnetic stir plate with a magnetic stir bar until fully dissolved. For small scale analysis, dispense 2 mL into 16x100 mm culture tubes and cap. For biological assay, put 50 mL in a 250 mL Erlenmeyer flask and cover with foil. Autoclave and allow to cool to room temperature before use. This solution is stable indefinitely at room temperature if firmly sealed.

Minimal Media A

Media adapted from Rateb, *et al.*[60] Combine the following with 1L of distilled water (all weights in grams): KH_2PO_4 (2.0), NH_4Cl (1.5), $MgSO_4 \cdot 7H_2O$ (0.5), glycerol (10), *myo*-inositol (0.4), monosodium $_L$-glutamate monohydrate (5.0), NaF (0.084), $FeSO_4 \cdot 7H_2O$ (0.025), $ZnSO_4 \cdot 7H_2O$ (0.01), $CoCl_2 \cdot 6H_2O$ (0.01), $CaCO_3$ (0.25), *p*-aminobenzoate (0.001). Mix using a magnetic stir plate and magnetic stir bar. Adjust pH to 7.0. For small scale analysis, dispense 2 mL into 16 x 100 mm culture tubes and cap. For biological assay, put 50 mL in a 250 mL Erlenmeyer flask and cover with foil. Autoclave and allow to cool to room temperature. This solution is stable indefinitely at room temperature if firmly sealed.

Minimal Media B

Combine the following with 1L of distilled water (all weights in grams): Glycerol (15 mL), $_{L}$-glutamine (5), K_2HPO_4 (1.5), $MgSO_4$ (0.2). Mix on a stir plate with a magnetic stir bar until completely dissolved. Adjust the pH to 7.0. For small scale analysis, dispense into 16x100 mm culture tubes and cap. For biological assay, put 50 mL in a 250 mL Erlenmeyer flask and cover with foil. Autoclave and allow to cool to room temperature. This solution is stable at room temperature indefinitely if firmly sealed.

TSB Agar

Mix 3g of TSB powder with 15g of agar and 4.5g of NaCl with 1L of distilled water. Mix on a magnetic stir plate with a magnetic stir bar. Autoclave in an appropriate container and allow to cool enough to handle. Pour into sterile petri dishes. Allow to cool completely. Store at 4°C indefinitely.

BHI Agar

Mix 3.7g of BHI powder with 15g of agar and 4.5g of NaCl with 1L of distilled water. Mix on a magnetic stir plate with a magnetic stir bar until dissolved. Autoclave in an appropriate container and allow enough to handle. Pour into sterile petri dishes. Allow to cool completely. Store at 4°C indefinitely.

YM media

Mix 21g of YM powder with 1L of distilled water. Mix on a magnetic stir plate with a magnetic stir bar until dissolved. Autoclave in an appropriate container and allow to cool to room temperature before use.

Positive control stocks for bioassay

Prepare 10 mg/ml stock solution in DMSO of chloramphenicol and fluconazole.

5.2.5 Cell preparation for biological assays

Thaw fungal or bacteria cell lines and dilute with YM media (fungi) or TSB (bacteria), 20 µL of cells with 11 mL of media. Dispense stocks into an appropriate number of wells in a 96 well plate, 100 µL per well.

For cancer cells, add 10,000 MIA PaCa-2 cells in an appropriate number of wells in a 96-well plate. Allow cells to attach overnight in a 37°C incubator with a 5% CO_2 atmosphere.

5.2.6 Equipment setup

Aurora Biomed liquid handling robot

The Aurora Biomed liquid handling robot is used for dispensing organic solvent (EtOAc) into each culture. After a brief extraction period, the robot uses conductivity sensing to remove the organic layer from the top of each culture and transfers it to a 96 well plate. Program the instrument for your specific arrangement of plate and rack

holders. Dispense 1.5 mL of EtOAc into each culture broth and set a 1 h wait time before removal to a 96 well plate.

Shimadzu LCMS: ESI single quad MS, PDA detector, autosampler outfitted with 96well plate track holder

The Shimadzu LC2020 series LCMS is used for generating PDA and mass chromatograms of crude bacteria culture extracts. Calibrate your particular autosampler to aspirate 10 µL from each well. Use the software to program a 17 minute analysis. For the mass spectrometry analysis, set up two events, one for positive ion detection and one for negative ion detection. Set both events to scan *m/z* 200-1500. Program the PDA to scan the entire available wavelength, from 190-800 nm. Set up the gradient as per Table 5, which allows for a relatively complete separation of all components and incorporates a column equilibration step. Solvent A is water with 0.1% formic acid, Solvent B is ACN (see Reagents section)

Table 6. Gradient scheme for LCMS separation of crude extracts

Time	B concentration
1 min	10
13 min	100
15 min	100
15.01 min	10
17 min	10

5.3 Protocol

Sampling of deceased mammals (timing: 1-2 hours)

CAUTION: Use appropriate personal protective equipment while sampling animal carcasses

1. Locate carcasses that meet the criteria described in 5.1.4. Use a sterile swab and sample the oral, nasal, aural, anal and vaginal (if present) cavities.
2. Use a sharp scalpel to open the body cavity and locate the large and small intestines. Using a sterile scalpel, make a small incision in the intestines and swipe inside with a sterile swab.
3. Put swabs in a tube with 5 mL sterile PBS, breaking off the stick if necessary to fit in the tube.
4. Store at 25°C until plating.

Plating of samples (timing: ~3 days)

5. Vortex each tube for approximately 15 seconds.
6. From each sample tube pipette 200, 100, and 50 µL aliquots onto separate plates of each media type. (If using two plate types, there will be 6 plates per sample swab)
7. Grow plates in a hypoxic chamber (N_2:O_2:CO_2 1%:1%:98% (vol/vol)) for 36 hours.
8. Visually identify individual colonies and restreak onto new plates. Repeat the process as necessary to obtain pure cultures.

9. Once pure, remove plates from hypoxic chamber and allow to grow aerobically at 25°C for 36 hours.

Glycerol stocks (timing: ~ 3 days)

10. After aerobic growth is observed, inoculate each pure bacterium into 5mL of sterile broth, corresponding to the solid medium used for its isolation.
11. Shake at 250 rpm for 36 hours.
12. After the growth period, dilute the remaining culture with 8.5:10 with glycerol.
13. Transfer a small aliquot to a cryogenic tube and free at -80°C for storage.

Small scale cultures (~3 hours per plate)

14. Scrape a small amount of purified organism off of each plate and inoculate a tube containing 2 mL of autoclaved medium. Cap securely.
15. Shake at 250 rpm for 36 hours.
16. Record qualitative observations after 36 hours on hardiness of growth, color, or other visual observations.
17. Use the Aurora liquid handler to extract cultures using EtOAc. Cultures should extract for at least one hour before allowing the robot to remove the EtOAc and distribute it into a clean 96-well plate.
18. Dry the plate containing the extract under reduced pressure.

LCMS analysis of crude extracts (17 minutes per sample, ~27 hours per plate)

19. Add 100 µL of 90:10 MeOH:water to each well.

20. Sonicate briefly to ensure extract is fully dissolved
21. Load the 96-well plates into LCMS autosampler if available or transfer samples into the appropriately sized vials for your autosampler.
22. Analyze samples using the gradient and instrument settings described in section 5.2.4.
23. Analyze samples as described in section 5.1.5. Identify with organisms are prolific natural product producers.

Preparation of samples for biological assay (~3.5 days)

24. Innoculate 50 mL of the appropriate broth with any organisms identified in the screening procedure as being biosynthetically talented.
25. Shake on a rotary shaker at 250 rpm for 3 days.
26. Add 50 mL of EtOAc and let sit for 1 hour.
27. Remove organic layer using a separatory funnel. Put in a round bottom flask and dry on a rotary evaporator.
28. Transfer extract to a tared scintillation vial and dry using a speedvac evaporator.
29. Create a stock solution for analysis, 10 mg/mL for antimicrobial assays, 1 mg/ml for cancer call cytotoxicity.

Antimicrobial bioassay (~24-48 hours)

30. Dissolve crude extracts to a concentration of 10 mg/mL in DMSO for antimicrobial assays, 2.5 mg/mL for cancer cell cytotoxicity

31. 1 µL of crude extract solution or positive control stock solution is added to each well (concentration of 100 µg/mL for antimicrobial assays, 25 µg/mL for cancer cell cytotoxicity). Add 1 µL of 100% DMSO to 3 wells to use for blank correction. Plates are shaken orbitally for 5 seconds.
32. Initial OD_{600} readings are obtained for antimicrobial assays.
33. Fungal strains are incubated 48 hours in a humidified incubator at 37°C and 5% CO_2. Bacteria strains and cancer cells are incubated for 24 hours at the same conditions.
34. At the conclusion of the incubation period, bacterial and fungal plates are shaken orbitally prior to obtaining OD_{600} readings. The final OD_{600} is substracted from the initial OD_{600} to obtain the change in OD_{600} reading.
35. After incubation, cancer cell viability is determined by MTT assay.[74]

5.4 Anticipated results

The majority of organisms sampled will not be biosynthetically talented organisms under protocol conditions, and thus most LCMS traces observed will be relatively unexciting compared to the media controls. In our experience, the hit rate by LCMS analysis averages approximately 10%, with screenings sometimes greatly deviating from the norm in both directions. However, of these organisms designated as biosynthetically talented, the hit rate in biological assays is approximately 75% or greater. This implies that an organism with an interesting metabolic profile by LCMS has a strong likelihood of producing biologically active compounds. Consequently, this is a sensitive method for isolating and screening organisms from the mammalian

microbiome for prolific natural product producers. The method is applicable to any deceased small mammal that is available for sampling and can be easily adjusted to include more specialized and selective media types. As written, this is an efficient method for screening large amounts of symbiotic bacteria that requires little hands on time for researchers. It makes use of robotics and instrumentation to increase the number of organisms that can be screened at one time beyond what an individual research can perform in any given day.

References

1. Koehn, F. E.; Carter, G. T.; The evolving role of natural products in drug discovery. *Nat Res* **2005**, *4*, 206-220.

2. Butler, M. S., The Role of Natural Product Chemistry in Drug Discovery†. *J nat Prod* **2004**, *67* (12), 2141-2153.

3. Sarker, S. D. L., Zahid; Gray, Alexander L., *Natural Products Isolation*. 2nd ed.; Humana Press: **2006**.

4. Newman, D. J.; Cragg, G. M., Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J Nat Prod* **2012**, *75* (3), 311-335.

5. von Nussbaum, F.; Brands, M.; Hinzen, B.; Weigand, S.; Häbich, D., Antibacterial Natural Products in Medicinal Chemistry—Exodus or Revival? *Angew Chem Int Edit* **2006**, *45* (31), 5072-5129.

6. Li, J. W.-H.; Vederas, J. C., Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325* (5937), 161-165.

7. Palmer, A. C.; Kishony, R., Understanding, predicting and manipulatinf the genotypic evolution of antibiotic resistance. *Nat Rev Genet* **2013**, *14*, 243-248.

8. Whitman, W. B.; Coleman, D. C.; Wiebe, W. J., Prokaryotes: The unseen majority. *P Natl Acad Sci USA* **1998**, *95* (12), 6578-6583.

9. Lancini, G. D., Arnold L., Bacteria Pharmaceutical Products. In *The Prokaryotes*, Rosenberg, E. D., Edward F.; Stakebrandt, Erko; Thomson, Fabiano. , Ed. Springer-Verlag: Berlin, **2013**; pp 257-280.

10. Schatz, A.; Bugle, E.; Waksman, S. A., Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.∗†. *Exp Biol Med* **1944**, *55* (1), 66-69.

11. Antibiotics annual, 1953-1954: Proceedings of the symposium on antibiotics, october 28, 29, and 30, 1953, Washington D.C. Chairman: Henry Welch, Ph.D., Director, Division of Antibiotics, Food and Drug Administration. *J Amer Med Assoc* **1954**, *155* (11), 1019-1019.

12. Tomasz, M., Mitomycin C: small, fast and deadly (but very selective). *Chem Biol* **1995**, *2* (9), 575-579.

13. Doll, D. C.; Weiss, R. B.; Issell, B. F., Mitomycin: ten years after approval for marketing. *J Clin Oncol* **1985**, *3* (2), 276-86.

14. Kino, T. H., Hiroshi; Hashimoto, Michisane; Nishitama, Michihisa; Goto, Toshio; Okuhara, Masakuni; Kohsaka, Masanobu; Aoki, Hatsuo; Imanaka, Hiroshi, FK-506, a novel immunosuppressant isolated from a Streptomyces I. Fermentation, isolation and physio-chemical and biological characteristics. *J Antibiot* **1987**, *40* (9), 1249-1255.

15. Bowman, L. J.; Brennan, D. C., The role of tacrolimus in renal transplantation. *Expert opin Pharmaco* **2008**, *9* (4), 635-43.

16. Secondary Metabolism in Bacteria: Antibiotic Pathways, Regulation, and Function. In *Biology of the Prokaryotes*, Blackwell Science Ltd: **2009**; pp 627-651.

17. Hertweck, C., The Biosynthetic Logic of Polyketide Diversity. *Angew Chem Int Edit* **2009**, *48* (26), 4688-4716.

18. Robinson, J. A., Polyketide Synthase Complexes: Their Structure and Function in Antibiotic Biosynthesis. *Philso T B* **1991**, *332* (1263), 107-114.

19. Finking, R.; Marahiel, M. A., Biosynthesis of Nonribosomal Peptides. *Annu Rev Microbiol* **2004**, *58* (1), 453-488.

20. Schwarzer, D.; Finking, R.; Marahiel, M. A., Nonribosomal peptides: from genes to products. *Nat Prod Rep* **2003**, *20* (3), 275-287.

21. Baltz, R., Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J Ind Microbiol Biotechnol* **2006**, *33* (7), 507-513.

22. Clardy, J.; Fishback, M. A.; Wals, C. T., New Antibiotics from bacterial natural products. *Nat Biotechnol* **2006**, *24*, 1541-1550.

23. Uria, A.; Piel, J., Cultivation-independent approaches to investigate the chemistry of marine symbiotic bacteria. *Phytochem Rev* **2009**, *8* (2), 401-414.

24. Clardy, J., Using genomics to deliver natural products from symbiotic bacteria. *Genome biol* **2005**, *6* (9), 232.

25. (a) Piel, J.; Butzke, D.; Fusetani, N.; Hui, D.; Platzer, M.; Wen, G.; Matsunaga, S., Exploring the Chemistry of Uncultivated Bacterial Symbionts: Antitumor

Polyketides of the Pederin Family. *J Nat Prod* **2005**, *68* (3), 472-479; (b) Piel, J.; Hui, D.; Wen, G.; Butzke, D.; Platzer, M.; Fusetani, N.; Matsunaga, S., Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge Theonella swinhoei. *P Natl Acad Sci USA* **2004**, *101* (46), 16222-16227.

26. Pavan, M. B., G. , Pederin, the toxic principle obtained in the crystalline state from blister beetles (Paederus fuscipes). *Phys Comp Oecol* **1953**, *3*, 307-312.

27. Sakemi, S.; Ichiba, T.; Kohmoto, S.; Saucy, G.; Higa, T., Isolation and structure elucidation of onnamide A, a new bioactive metabolite of a marine sponge, Theonella sp. *J Am Chem Soc* **1988**, *110* (14), 4851-4853.

28. Cichewicz, R. H.; Valeriote, F. A.; Crews, P., Psymberin, A Potent Sponge-Derived Cytotoxin from Psammocinia Distantly Related to the Pederin Family. *Org Let* **2004**, *6* (12), 1951-1954.

29. Piel, J., A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles. *P natl Acad Sci USA* **2002**, *99* (22), 14002-14007.

30. Johnson, T. A.; Sohn, J.; Vaske, Y. M.; White, K. N.; Cohen, T. L.; Vervoort, H. C.; Tenney, K.; Valeriote, F. A.; Bjeldanes, L. F.; Crews, P., Myxobacteria versus sponge-derived alkaloids: The bengamide family identified as potent immune modulating agents by scrutiny of LC–MS/ELSD libraries. *Bioorg Med Chem* **2012**, *20* (14), 4348-4355.

31. Future Challenges. In *Drug Discovery from Natural Products*, The Royal Society of Chemistry: **2012**; pp 391-393.

32. Lederberg, J. M., Alexa, 'Ome Sweet 'Omics-- A genealogical Treasury of Words. *The Scientist* **2001**, *17* (7).

33. Clardy, J.; Walsh, C., Lessons from natural molecules. *Nature* 2004, *432* (7019), 829-37.

34. Kirsop, B. E., The convention on Biological Diversity: Some implications for microbiology and microbial culture collections. *J Inf Microbiol Biot* **1996**, *17* (5-6), 505-511.

35. Pace, N. R., A Molecular View of Microbial Diversity and the Biosphere. *Science* 1997, *276* (5313), 734-740.

36. (a) Piel, J., Metabolites from symbiotic bacteria. *Nat Prod Rep* **2004**, *21* (4), 519-538; (b) Piel, J., Metabolites from symbiotic bacteria. *Nat Prod Rep* **2009**, *26* (3), 338-362.

37. Zwickl, P.; Voges, D.; Baumeister, W., The proteasome: a macromolecular assembly designed for controlled proteolysis. *Philos T Roy Soc B* **1999**, *354* (1389), 1501-11.

38. Szókán, G.; Mezö, G.; Hudecz, F., Application of marfey's reagent in racemization studies of amino acids and peptides. *J Chrom A* **1988**, *444* (0), 115-122.

39. Pevzner, Y.; Metcalf, R.; Kantor, M.; Sagaro, D.; Daniel, K., Recent advances in proteasome inhibitor discovery. *Expert Opin Drug Dis* **2013**, *8* (5), 537-68.

40. DeMartino, G. N.; Slaughter, C. A., The proteasome, a novel protease regulated by multiple mechanisms. *J Biol Chem* **1999**, *274* (32), 22123-6.

41. Fuchs, S. A.; Berger, R.; Klomp, L. W. J.; de Koning, T. J., d-Amino acids in the central nervous system in health and disease. *Mol Genet Metab* **2005**, *85* (3), 168-180.

42. Blaser, M. J.; Falkow, S., What are the consequences of the disappearing human microbiota? *Nat Rev Microbiol* **2009**, *7* (12), 887-94.

43. Malik, N. N., Drug discovery: past, present and future. *Drug Discov Today* **2008**, *13* (21–22), 909-912.

44. Weissman, K. J.; Leadlay, P. F., Combinatorial biosynthesis of reduced polyketides. *Nat Rev Microbiol* **2005**, *3* (12), 925-36.

45. Theodore, C. M.; King, J. B.; You, J.; Cichewicz, R. H., Production of Cytotoxic Glidobactins/Luminmycins by Photorhabdus asymbiotica in Liquid Media and Live Crickets. *J nat Prod* **2012**, *75* (11), 2007-2011.

46. Goodrich-Blair, H.; Clarke, D. J., Mutualism and pathogenesis in Xenorhabdus and Photorhabdus: two roads to the same destination. *Mol Microbiol* **2007**, *64* (2), 260-8.

47. Herbert, E. E.; Goodrich-Blair, H., Friend and foe: the two faces of Xenorhabdus nematophila. *Nat Rev Microbiol* **2007**, *5* (8), 634-646.

48. Forst, S.; Dowds, B.; Boemare, N.; Stackebrandt, E., Xenorhabdus and Photorhabdus spp.: bugs that kill bugs. *Annu Rev Microbiol* **1997**, *51*, 47-72.

49. Boemare, N., Biological and physiological characterizations of colony forms variants in *Xenorhabdus* spp. (*Enterobacteriaceae*). *J Gen Microbiol* **1988**, *134*.

50. Somvanshi, V. S.; Sloup, R. E.; Crawford, J. M.; Martin, A. R.; Heidt, A. J.; Kim, K.-s.; Clardy, J.; Ciche, T. A., A Single Promoter Inversion Switches

Photorhabdus Between Pathogenic and Mutualistic States. *Science* **2012**, *337* (6090), 88-93.

51. Gerrard, J. G.; Joyce, S. A.; Clarke, D. J.; ffrench-Constant, R. H.; Nimmo, G. R.; Looke, D. F.; Feil, E. J.; Pearce, L.; Waterfield, N. R., Nematode symbiont for Photorhabdus asymbiotica. *Emerging Infec Dis* **2006**, *12* (10), 1562-4.

52. Gerrard, J.; Waterfield, N.; Vohra, R.; ffrench-Constant, R., Human infection with Photorhabdus asymbiotica: an emerging bacterial pathogen. *Microbes Infect* **2004**, *6* (2), 229-237.

53. Waterfield, N. R.; Sanchez-Contreras, M.; Eleftherianos, I.; Dowling, A.; Yang, G.; Wilkinson, P.; Parkhill, J.; Thomson, N.; Reynolds, S. E.; Bode, H. B.; Dorus, S.; ffrench-Constant, R. H., Rapid Virulence Annotation (RVA): Identification of virulence factors using a bacterial genome library and multiple invertebrate hosts. *P Natl Acad Sci USA* **2008**, *105* (41), 15967-15972.

54. Wilkinson, P.; Waterfield, N.; Crossman, L.; Corton, C.; Sanchez-Contreras, M.; Vlisidou, I.; Barron, A.; Bignell, A.; Clark, L.; Ormond, D.; Mayho, M.; Bason, N.; Smith, F.; Simmonds, M.; Churcher, C.; Harris, D.; Thompson, N.; Quail, M.; Parkhill, J.; ffrench-Constant, R., Comparative genomics of the emerging human pathogen Photorhabdus asymbiotica with the insect pathogen Photorhabdus luminescens. *BMC Genomics* **2009**, *10* (1), 302.

55. Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R., antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene

clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **2011**, *39* (suppl 2), W339-W346.

56. Schellenberg, B.; Bigler, L.; Dudler, R., Identification of genes involved in the biosynthesis of the cytotoxic compound glidobactin from a soil bacterium. *Environ Microbiol* **2007**, *9* (7), 1640-50.

57. Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Muller, R.; Stewart, A. F.; Zhang, Y., Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat Biotech* **2012**, *30* (5), 440-446.

58. Bian, X.; Plaza, A.; Zhang, Y.; Müller, R., Luminmycins A–C, Cryptic Natural Products from Photorhabdus luminescens Identified by Heterologous Expression in Escherichia coli. *J Nat Prod* **2012** *75*,(9), 1652-1655 .

59. Groll, M.; Schellenberg, B.; Bachmann, A. S.; Archer, C. R.; Huber, R.; Powell, T. K.; Lindow, S.; Kaiser, M.; Dudler, R., A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. *Nature* **2008**, *452* (7188), 755-758.

60. Rateb, M. E.; Houssen, W. E.; Harrison, W. T. A.; Deng, H.; Okoro, C. K.; Asenjo, J. A.; Andrews, B. A.; Bull, A. T.; Goodfellow, M.; Ebel, R.; Jaspars, M., Diverse Metabolic Profiles of a Streptomyces Strain Isolated from a Hyper-arid Environment. *J Nat Prod* **2011**, *74* (9), 1965-1971.

61. Oka, M.E.A., Glidobactins a, b, and c, new antitumor antibiotics ii. structure elucidation. *J Antibiot* 1988, *xli*.

62. Oka, M. E. A, Glidobactins A, B, and C, new antitumor antibiotics I. production, isolation, chemical properties and biological activity. *J Antibiot* **1988**, *xli* (10).

63. Park, Y.; Stanley, D., The entomopathogenic bacterium, Xenorhabdus nematophila, impairs hemocytic immunity by inhibition of eicosanoid biosynthesis in adult crickets, Gryllus firmus. *Biol Control* **2006**, *38* (2), 247-253.

64. Sutton, S., Measurement of Microbial Cells by Optical Density. *Journal of Val Tech* **2011**, *17* (1).

65. Hansen, M. B.; Nielsen, S. E.; Berg, K., Re-examination and further development of a precise and rapid dye method for measuring cell growth/cell kill. *Journal of Immunol Methods* **1989**, *119* (2), 203-210.

66. Yang, J. Y.; Karr, J. R.; Watrous, J. D.; Dorrestein, P. C., Integrating '-omics' and natural product discovery platforms to investigate metabolic exchange in microbiomes. *Curr Opin Chem Biol* **2011**, *15* (1), 79-87.

67. Al-Lahham, S. H.; Peppelenbosch, M. P.; Roelofsen, H.; Vonk, R. J.; Venema, K., Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. *Biochim Biophys Acta* **2010**, *1801* (11), 1175-83.

68. Asensio, C.; Pérez-Díaz, J. C.; Martínez, M. C.; Baquero, F., A new family of low molecular weight antibiotics from enterobacteria. *Biochem and Bioph Res Co* **1976**, *69* (1), 7-14.

69. Ruas-Madiedo, P.; Medrano, M.; Salazar, N.; De Los Reyes-Gavilan, C. G.; Perez, P. F.; Abraham, A. G., Exopolysaccharides produced by Lactobacillus and Bifidobacterium strains abrogate in vitro the cytotoxic effect of bacterial toxins on eukaryotic cells. *J App Microbiol* **2010**, *109* (6), 2079-86.

70. (a) Joyner, P. M.; Liu, J.; Zhang, Z.; Merritt, J.; Qi, F.; Cichewicz, R. H., Mutanobactin A from the human oral pathogen Streptococcus mutans is a cross-

kingdom regulator of the yeast-mycelium transition. *Organic & Biomolecular Chemistry* 2010, *8* (24), 5486-5489; (b) Wang, X.; Du, L.; You, J.; King, J. B.; Cichewicz, R. H., Fungal biofilm inhibitors from a human oral microbiome-derived bacterium. *Org Biomol Chem* **2012**, *10* (10), 2044-2050.

71. Wei, M. Q.; Mengesha, A.; Good, D.; Anné, J., Bacterial targeted tumour therapy-dawn of a new era. *Cancer lett* **2008**, *259* (1), 16-27.

72. Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T., antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nuc Acids Res* **2013**, *41* (W1), W204-W212.

73. Barsby, T.; Kelly, M. T.; Andersen, R. J., Tupuseleiamides and basiliskamides, new acyldipeptides and antifungal polyketides produced in culture by a Bacillus laterosporus isolate obtained from a tropical marine habitat. *J Nat Prod* **2002**, *65* (10), 1447-51.

74. Kelly, M. T. S.; Anderson, R. J. V.; Barsby, T. A. V. Antifungal and antimycobacterial basiliskamides. US 2005/0277779 A1, 2005.

75. Bythell, B. J.; Maître, P.; Paizs, B., Cyclization and Rearrangement Reactions of an Fragment Ions of Protonated Peptides. *J Am Chem Soc* **2010**, *132* (42), 14766-14779.

76. Piel, J., Biosynthesis of polyketides by trans-AT polyketide synthases. *nar Prod Rep* **2010**, *27* (7), 996-1047.

77. Moore, B. S.; Hertweck, C., Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat Prod Rep* **2002**, *19* (1), 70-99.

78. (a) Zaleta-Rivera, K.; Xu, C.; Yu, F.; Butchko, R. A. E.; Proctor, R. H.; Hidalgo-Lara, M. E.; Raza, A.; Dussault, P. H.; Du, L., A Bidomain Nonribosomal

Peptide Synthetase Encoded by FUM14 Catalyzes the Formation of Tricarballylic Esters in the Biosynthesis of Fumonisins. *Biochem* **2006**, *45* (8), 2561-2569; (b) Lin, S.; Huang, T.; Horsman, G. P.; Huang, S. X.; Guo, X.; Shen, B., Specificity of the ester bond forming condensation enzyme SgcC5 in C-1027 biosynthesis. *Org Lett* **2012**, *14* (9), 2300-3.

79. Bhushan, R.; Brückner, H., Marfey's reagent for chiral amino acid analysis: A review. *Amino Acids* **2004**, *27* (3-4), 231-247.

80. Shaaban, M.; Maskey, R. P.; Wagner-Dobler, I.; Laatsch, H., Pharacine, a natural p-cyclophane and other indole derivatives from Cytophaga sp. strain AM13.1. *J Nat Prod* **2002**, *65* (11), 1660-3.

81. Kochhar, S.; Christen, P., Amino Acid Analysis by Precolumn Derivatization with 1-Fluoro-2,4-Dinitrophenyl-5-I-Alanine Amide (Marfey's Reagent). *Protein Protocol Handbook*. Humana Press: Totowa, N.J. **2002** pp 995-1000.

Appendix

Appendix table of contents

Discussion A1. Explanation of Marfey's analysis……..…….…….....….87

Figure M1. L and D-amino acids…….....87

Figure M2. Marfey's reagent and derivatized amino acid diastereomers……...88

Figure M3. Single ion traces of derivatized amino acids……89

Discussion A2. Explanation of Proteasome inhibition…….……...90

Figure P1. Structure of SUC-LLVY-AMC……91

Table A1. Summary of antiSMASH analysis of the *P. asymbiotica* ATCC43696 genome……92

Table A2. Summary of media types tested to induce the production of glidobactin/luminmycins…….....93

Table A3. Crystal data and structure refinement for **9**……95

Table A4. Atomic coordinates and equivalent isotropic displacement parameters for **9**……...96

Table A5. Bond lengths and angles for **9**……98

Table A6. Anisotropic displacement paraments for **9**……...103

Table A7. Hydrogen coordinates and isotropic displacement parameters for **9**……105

Table A8. Torsion angles for **9**……..108

Table A9. Hydrogen bonds for **9**……...111

Figure A1. Screenshot of antiSMASH results for secondary metabolite gene cluster 15……...112

Figure A2. ^{1}H –NMR (500 MHz, DMSO-*d*6, 25°C) spectrum of compound **3**……....113

Figure A3. ^{1}H-^{13}C HSQC (500 MHz, DMSO-*d*6, 25°C) spectrum of compound **3**…..114

Figure A4. ^{1}H-^{13}C *g*HMBC (500 MHz, DMSO-*d*6, 25°C) spectrum of compound **3**..115

Figure A5. ^{1}H-^{1}H *g*COSY (500 MHz, DMSO-*d*6, 25°C) spectrum of compound **3**….116

Figure A6. ^{1}H-^{1}H tCOSY (500 MHz, DMSO-*d*6, 25°C) spectrum of compound **3**…..117

Figure A7. IR spectrum (thin film, room temperature) of compound **3**……118

Figure A8. HRESIMS (positive mode) data for compound **3**……...119

Figure A9. *In vitro* cytoxicity (upper panel) and proteasome inhibition (lower panel) for **1**,**2**, and **3** in normal mouse fibroblast cells……...120

Figure A10. ^{1}H-NMR (500 MHz, DMSO-*d*6, 25°C) spectrum of **8**…….....121

Figure A11. ^{1}H-^{13}C HSQC (500 MHz, DMSO-*d*6, 25°C) spectrum of **8**…….122

Figure A12. ^{1}H-^{13}C gHMBC (500 MHz, DMSO-*d*6, 25°C) spectrum of **8**……..123

Figure A13. HRESIMS (negative mode) data for **8**……..124

Figure A14. UV spectrum of **8**……..125

Figure A15. ^{1}H-NMR (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**…….126

Figure A16. ^{1}H-^{13}C HSQC (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**…….127

Figure A17. ^{1}H-^{13}C gHMBC (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**……..128

Figure A18. ^{1}H-^{1}H tCOSY (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**…….129

Figure A19. ^{1}H-^{1}H gCOSY (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**…….130

Figure A20. ^{1}H-^{1}H NOESY (500 MHz, DMSO-*d*6, 25°C) spectrum of **9**……131

Figure A21. ^{15}N-^{1}H HSQC (500 MHz, DMSO-*d*6, 25°C) overlay spectrum of **9**……132

Figure A22. Thermal ellipsoid plot of **9**……133

Figure A23. Packing diagram of **9**……134

Figure A24. Chromatographic analysis of standards for Marfey's analysis…….135

Figure A25. Marfey's analysis of **9**……...136

Figure A26. HRESIMS (positive mode) data for **9**……...137

Figure A27. UV data for **9**……138

Figure A28. CD data for **9**……139

Discussion A1. Explanation of Marfey's analysis

Peptides are comprised standard and non-standard amino acids. These amino acids consist of an amine group, an α-carbon with a pendant group and carboxylic acid. Amino acids can exist in one of two isomer forms, L or D (figure M1) L-amino acids are almost exclusively used in protein production in animals and humans, although there are a few exceptions for D-amino acids playing biological roles in humans and other animals.[41] D-amino acids, can frequently be found in bacteria, especifically in bacterial cell walls. [41]

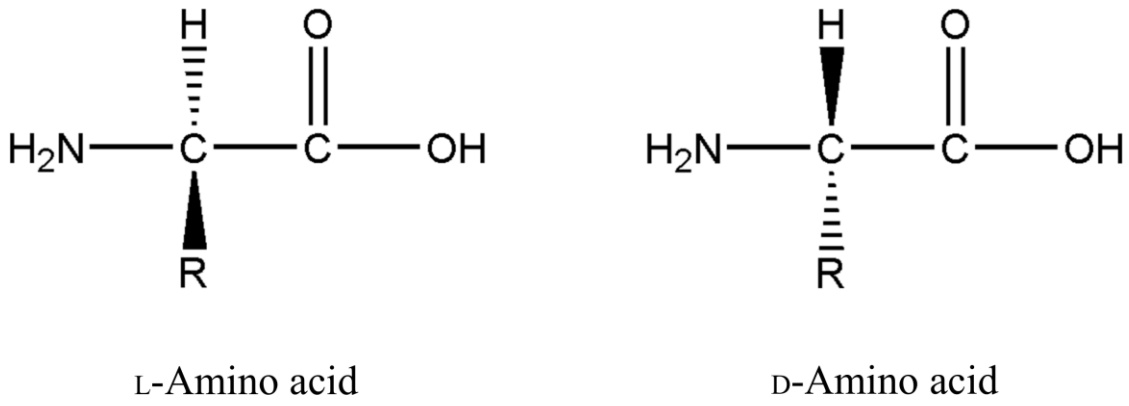
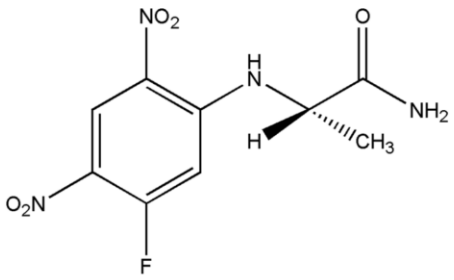


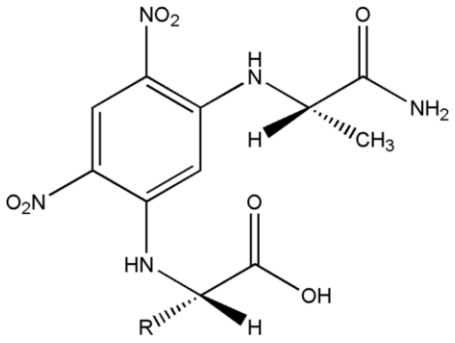
Figure M1. L and D-amino acids.

A common method to determine the chirality of amino acids is through chromatographic separation of their diastereomers. Amino acids are derivatized with an enatiomerically pure reactant and then separated by HPLC chromatography. The reactant used typically has a good chromophore, to enhance detection by HPLC UV detectors and is relatively "bulky" to enhance resolution. In chapter 4, the absolute configuration of auriprocine was determined through a method known as Marfey's analysis. Using this method, the peptide of interest was hydrolyzed to yield its

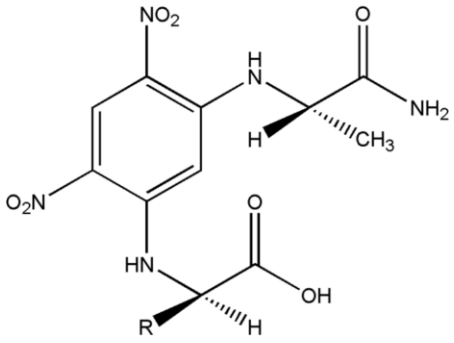
individual amino acid components and each amino acid is then derivatized with 1,5-difluro-2,4-dinitrophenyl-5-L-alanineamide (FDAA), most commonly referred to as Marfey's reagent. The result of this reaction, shown in figure M2, is either the LL or LD diastereomer of the precursor amino acid. Under the correct chromatographic conditions, these derivatized amino acids will have different retention times. By comparing the retention times of appropriate amino acid standards, the configuration of each amino acid in the can be determined.



Marfey's reagent



L,D diastereomer



L,L diastereomer

Figure M2. Marfey's reagent, and derivatized amino acid diastereomers

The separation of diasteromeric pairs is generally preferred to more direct methods (i.e., enantiomer separation), as it is often simpler to optimized chromatographic conditions, and does not require columns with chiral stationary phases.

However, there are a few disadvantages to this method. Amino acids such as tyrosine or histidine or those that contain two amino groups, can form both mono and disubstiuted derivatives. This is generally overcome through the use of excess Marfey's reagent, minimizing the amount of mono-substituted compound. Additionally, depending on the conditions used during hydrolysis, some racemization can occur. By running the full hydrolysis reaction of the standard amino acids as well as a sample can aid in determining if epimerization is occurring. Finally, depending on chromatographic conditions, peaks can sometimes overlap in the UV profile, making analysis more time consuming as chromatographic conditions must be manipulated. However, as was the case with auriprocine in chapter 4, this can be avoided by using an LCMS. By extracting the mass data for each derivatized amino acid, overlapping peaks in the UV will often resolve to allow the determination of retention times.

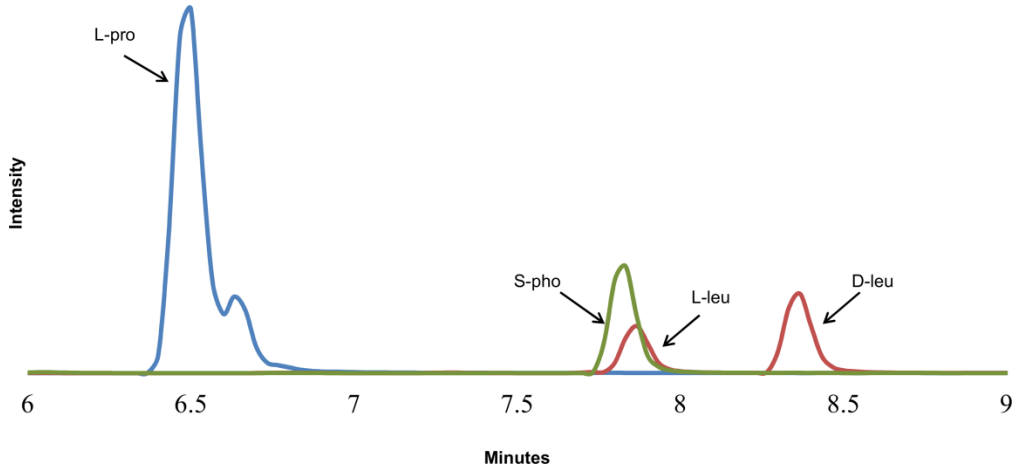


Figure M3. Single ion traces of derivatized amino acids, illustrating the resolution between overlapping peaks.

Discussion M2. Proteasome inhibition assay

The proteasome is a large (700 kDa) cylinder-shaped proteinase complex that is involved in the degradation of intercellular proteins.[40] The proteasome is made up of 35 different subunits, with the proteolytic core complex residing in the 20s proteasome.[37] The proteasome is responsible for degrading a range of intercellular proteins, involved in cell cycle control, singal transduction, transcription factors, oncogenes and oncogene products. [40] Although protein degradation is vital to the survival of a cell, proteasome inhibitors exhibit anti-inflammatory and antiproliferative effects, making the search for proteasome inhibitors an exciting area of research.

In chapter 3, the glidobactin and luminmycin family of compounds were shown to be proteasome inhbitors, the effect of which was studied through the use of a kit purchased from Cayman Chemical (kit number 10008041). The kit employs a 20s subunit substrate, *N*-Succinyl-Leu-Leu-Val-Tyr-7-amino-4-methylcoumarin (SUC-LLVY-AMC), which when cleaved by the proteasome releases a highly fluorescent product. If the proteasome is active, the SUC-LLVY-AMC will be cleaved, and the fluorescence can be measured at 360 and 480 nm, excitation and emission wavelengths, respectively. Proteasome inhibitors will show a decrease in fluorescence dependent on the concentration of the inhibitor present

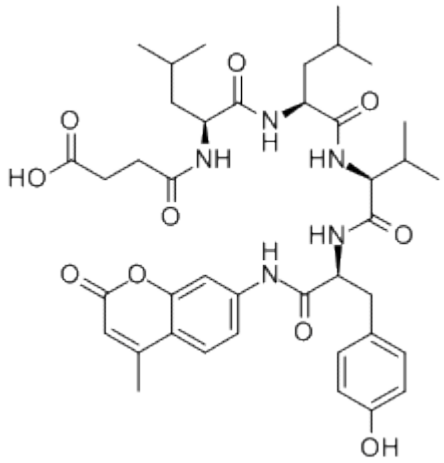


Figure P1. Structure of SUC-LLVY-AMC

Table A1. Summary of antiSMASH results of the *P. asymbiotica* ATCC43949 genome. Putative gene clusters for secondary metabolites and their genomic locatins are shown. Cluster 15 was identified as homologous to the genes responsible for glidobactin biosynthesis. NRPS = non-ribosomal peptide synthase, T1PKS = type I polyketide synthase

Gene Cluster	Type of Cluster
1	NRPS/T1PKS
2	NRPS/nucleoside
3	NRPS
4	T1PKS
5	T1PKS
6	Phosphoglycolipid
7	NRPS
8	NRPS
9	NRPS
10	NRPS
11	NRPS/T1PKS
12	NRPS/T1PKS
13	Other
14	NRPS
15	T1PKS/NRPS
16	NRPS
17	NRPS
18	NRPS
19	Butyrolacetone

Table A2. Summary of media types explored to induce the production of glidobactins/luminmycins.

Name	Formulation	Compounds induced
TSB	Commercially available	-
TSB plus chitin	Commercially available mix at 75% strength with 7.5gL^{-1} chitin	-
TSB plus soil	Commercially available mix with 10% v/v soil extract added prior to autoclaving	-
TSB plus blood	Commercially available mix with 5% v/v sterile sheep blood added after autoclaving	-
TSB plus bugs	Commercially available mix with 3 ground bugs collected in Norman, OK added to 300 mL of broth prior to autoclaving	-
TSB mimic, less sugar increased tryptone	In gL^{-1}: dextrose (1.7), tryptone (27), NaCl (5), K_2HPO_4(2.5)	-
TSB mimic, less tryptone, increased sugar	In gL^{-1}: dextrose (1.7), tryptone (27), NaCl (5), K_2HPO_4(2.5)	-
TSB mimic, glucose	In gL^{-1}: glucose (2.5), tryptone (20), NaCl (5), K_2HPO_4(2.5)	-
TSB mimic, ribose	In gL^{-1}: ribose (2.5), tryptone (20), NaCl (5), K_2HPO_4(2.5)	-
TSB mimic, sucrose	In gL^{-1}: sucrose (2.5), tryptone (20), NaCl (5), K_2HPO_4(2.5)	-
Actinomyces broth	Commercially available	-
Potato dextrose	Commercially available	**1**
Luria broth	Commercially available	**1**
Luria broth plus homogenized bugs	Commercially available mix with approximately 10 bugs collected in Norman, OK added to 300 mL of broth and homogenized with mechanical homogenizer prior to autoclaving	**1**
Defined medium[a]	In gL^{-1}: KH_2PO_4 (2.0), NH_4Cl (1.5), $MgSO_4{\cdot}7H_2O$ (0.5), glycerol (10), myo-insitol (0.4), monosodium L-glutamate monohydrate (5.0), NaF (0.084), $FeSO_4{\cdot}7H_2O$ (0.025), $ZnSO_4{\cdot}7H_2O$ (0.01), $CoCl_2{\cdot}6H_2O$ (0.01), $CaCO_3$ (0.25), p-aminobenzoate (0.001)	**1**, **2**, **3**
Defined medium, honey as carbon source	Same as defined medium but with 10gL^{-1} of honey in place of glycerol	**1,** (trace amounts of **2** and **3)**
Defined medium, sucrose as carbon source	Same as defined medium but with 10gL^{-1} of sucrose in place of glycerol	**1,** (trace amounts of **2** and **3)**
Defined medium, glucose as carbon source	Same as defined medium but with 10gL^{-1} of glucose in place of glycerol	**1**, **2**, **3**
Defined medium, 10% v/v sheep blood	Same as defined medium, with the addition of 10% v/v of sterile sheep blood added after autoclaving	**1**, **2**, **3**

0.25x seawater	In gL^{-1}: NH_4Cl (1.5), glycerol (10), monosodium L-glutamate monohydrate (5), instant ocean (8.8)	-
0.5x seawater	In gL^{-1}: NH_4Cl (1.5), glycerol (10), monosodium L-glutamate monohydrate (5), instant ocean (17.5)	-
1x seawater	In gL^{-1}: NH_4Cl (1.5), glycerol (10), monosodium L-glutamate monohydrate (5), instant ocean (35.0)	-
Dulbecco's modified Eagle Medium (DMEM)	Sterile, filtered, commercially available	**1**
DMEM plus 5% fetal bovine serum (FBS)	Sterile, filtered, commercially available DMEM and FBS added at 5% v/v	-
DMEM plus glycerol	In gL^{-1}: glycerol (10). Autoclaved and mixed with sterile, filtered, commercially available DMEM	**1**
DMEM plus glycerol and 5% FBS	In gL^{-1}: glycerol (10). Autoclaved and mixed with sterile, filtered, commercially available DMEM and 5% v/v FBS	-
Plant food plus glycerol	In gL^{-1}: plant food, commercially available water soluble (1.6), glycerol (10)	**1**
Plant food plus sucrose	In gL^{-1}: plant food, commercially available water soluble (1.6), sucrose (10)	**1**
Plant food plus glucose	In gL^{-1}: plant food, commercially available water soluble (1.6), glucose (10)	**1**
Plant food plus honey	In gL^{-1}: plant food, commercially available water soluble (1.6), honey (10)	**1**
Plant food plus cholesterol	In gL^{-1}: plant food, commercially available water soluble (1.6), cholesterol (10)	-
Skim milk	In gL^{-1}: skim milk powder (10)	-
Skim milk plus glycerol	In gL^{-1}: skim milk powder (10), glycerol (10)	-

Compound **1** : Glidobactin A; Compound **2**: Luminmycin A; Compound **3**: Luminmycin D;

[a] Defined medium described by Rateb *et al*. *J. Nat. Prod*. **2011**, *74*, 1965-1971.

Table A3. Crystal data and structure refinement for auriprocine, **9**

Emperical formula	$C_{43}H_{70}N_6O_8$
Formula weight	799.05
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Unit cell dimensions	a = 11.1021(10) Å $\alpha = 90^o$
	b = 14.9391(14) Å $\beta = 90^o$
	c = 26.211(2) Å $\gamma = 90^o$
Volume	4347.2(7) $Å^3$
Z,Z′	4, 1
Density (calculated)	1.221 Mg/m^3
Wavelength	0.71073 Å
Temperature	100(2) K
F(000)	1736
Absorption coefficient	0.084 mm^{-1}
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.990 and 0.964
Theta range for data collection	1.554 to 28.344^o
Reflections collected	40661
Independent reflections	10800 [R(int) = 0.0586]
Data / restraints / parameters	10800 / 118 / 573
wR(F^2 all data)	wR2 = 0.1925
R(F obsd data)	R1 = 0.0707
Goodness-of-fit on F^2	1.011
Observed data [$I>2\sigma(I)$]	9029
Absolute structure parameter	1.2(6)
Largest and mean shift / s.u.	0.012 and 0.000
Largest diff. peak and hole	0.315 and -0.374

$wR2 = \{\Sigma [w (F_o^2-F_c^2)^2] / \Sigma [w (F_o^2)^2]\}^{1/2}$

$R1 = \Sigma ||F_o| - |F_c|| / \Sigma |F_o|$

Table A4. Atomic coordinates and equivalent isotropic displacement parameters for **9**. U(eq) is defined as one third of the trace of the orthogonalized U_{ij} tensor.

	x	y	z	U(eq)
C(1)	-0.0302(9)	0.7182(6)	0.9049(3)	0.032(2)
C(2)	-0.0906(7)	0.6462(8)	0.8728(5)	0.031(3)
C(1A)	-0.1152(11)	0.6459(9)	0.9234(4)	0.035(3)
C(2A)	-0.0561(14)	0.6675(6)	0.8725(5)	0.031(3)
C(3)	0.0039(4)	0.5876(3)	0.84645(16)	0.0259(10)
C(4)	0.1015(5)	0.5476(4)	0.88184(18)	0.0317(11)
C(5)	0.0656(4)	0.6276(3)	0.79971(17)	0.0227(9)
O(1)	-0.0148(3)	0.6604(2)	0.76155(12)	0.0250(7)
C(6)	0.1627(4)	0.5700(3)	0.77505(15)	0.0172(8)
O(2)	0.2689(3)	0.5970(2)	0.77529(11)	0.0228(6)
N(1)	0.1361(3)	0.4930(2)	0.75030(13)	0.0182(7)
C(7)	0.2338(4)	0.4434(3)	0.72609(15)	0.0181(8)
C(8)	0.1719(4)	0.3578(3)	0.70557(16)	0.0209(8)
C(9)	0.0381(4)	0.3812(3)	0.70385(16)	0.0196(8)
C(10)	0.0220(4)	0.4432(3)	0.74968(17)	0.0203(8)
C(11)	0.2936(4)	0.4976(3)	0.68379(15)	0.0188(8)
O(3)	0.4020(3)	0.4887(2)	0.67580(11)	0.0223(6)
N(2)	0.2222(3)	0.5513(3)	0.65565(13)	0.0195(7)
C(12)	0.2687(4)	0.6031(3)	0.61235(15)	0.0215(8)
C(13)	0.3057(5)	0.5427(3)	0.56787(17)	0.0281(10)
C(14)	0.2022(5)	0.4845(3)	0.54749(18)	0.0341(12)
C(15)	0.1090(7)	0.5395(5)	0.5188(3)	0.0593(19)
C(16)	0.2493(7)	0.4096(4)	0.5146(2)	0.0502(16)
C(17)	0.3742(4)	0.6626(3)	0.62954(16)	0.0206(8)
O(4)	0.4766(3)	0.6539(2)	0.61321(12)	0.0238(7)
N(3)	0.3498(3)	0.7256(2)	0.66559(14)	0.0194(7)
C(18)	0.4495(4)	0.7804(3)	0.68477(17)	0.0201(8)
C(19)	0.3850(4)	0.8553(3)	0.71504(18)	0.0236(9)
C(20)	0.2650(4)	0.8133(3)	0.73080(18)	0.0256(9)
C(21)	0.2305(4)	0.7552(3)	0.68505(17)	0.0226(9)
C(22)	0.5365(4)	0.7264(3)	0.71704(16)	0.0176(8)
O(5)	0.6435(3)	0.7476(2)	0.72007(12)	0.0231(6)
N(4)	0.4913(3)	0.6557(2)	0.74201(13)	0.0183(7)
C(23)	0.5677(4)	0.5955(3)	0.77162(17)	0.0205(8)
C(24)	0.6078(4)	0.6357(4)	0.82206(18)	0.0318(11)
C(25)	0.5061(5)	0.6500(5)	0.8609(2)	0.0449(15)
C(26)	0.4549(6)	0.5615(6)	0.8797(2)	0.062(2)
C(27)	0.5536(8)	0.7074(8)	0.9043(3)	0.087(3)

C(28)	0.6753(4)	0.5640(3)	0.73923(17)	0.0215(8)
O(6)	0.7772(3)	0.5580(2)	0.75787(12)	0.0266(7)
N(5)	0.6505(4)	0.5398(3)	0.69141(16)	0.0246(8)
C(29)	0.7423(4)	0.5032(3)	0.65695(18)	0.0265(9)
C(30)	0.6774(5)	0.4565(4)	0.6129(2)	0.0460(15)
C(31)	0.7508(6)	0.4342(6)	0.5648(2)	0.0338(16)
C(32)	0.6699(9)	0.3991(7)	0.5208(3)	0.0380(19)
C(33)	0.8387(9)	0.3568(7)	0.5777(4)	0.045(2)
C(31A)	0.7728(8)	0.3983(5)	0.5843(3)	0.0320(18)
C(32A)	0.8213(11)	0.4517(8)	0.5381(4)	0.040(2)
C(33A)	0.7190(11)	0.3079(6)	0.5675(4)	0.038(2)
C(34)	0.8314(4)	0.5756(3)	0.63934(16)	0.0214(9)
O(7)	0.9403(3)	0.5617(2)	0.63893(12)	0.0255(7)
N(6)	0.7797(3)	0.6508(3)	0.62220(14)	0.0211(7)
C(35)	0.8503(4)	0.7245(3)	0.60118(16)	0.0235(9)
C(36)	0.8877(4)	0.7909(3)	0.64222(18)	0.0268(10)
O(8)	0.7928(3)	0.8455(2)	0.65972(13)	0.0291(7)
C(37)	0.7796(5)	0.7710(3)	0.55821(17)	0.0289(10)
C(38)	0.7549(4)	0.7102(3)	0.51347(17)	0.0270(10)
C(39)	0.6430(5)	0.6701(4)	0.50671(19)	0.0354(12)
C(40)	0.6191(6)	0.6159(4)	0.4652(2)	0.0431(14)
C(41)	0.7104(6)	0.5976(4)	0.4299(2)	0.0433(14)
C(42)	0.8219(6)	0.6361(4)	0.4365(2)	0.0408(13)
C(43)	0.8446(5)	0.6914(4)	0.47787(18)	0.0333(11)

Table A5. Bond lengths [Å] and angles [°] for **9**

C(1)-C(2)	1.522(5)	N(2)-C(12)	1.467(5)
C(1)-H(1A)	0.9800	N(2)-H(2N)	0.82(5)
C(1)-H(1B)	0.9800	C(12)-C(13)	1.530(6)
C(1)-H(1C)	0.9800	C(12)-C(17)	1.538(6)
C(2)-C(3)	1.531(5)	C(12)-H(12)	1.0000
C(2)-H(2A)	0.9900	C(13)-C(14)	1.537(7)
C(2)-H(2B)	0.9900	C(13)-H(13A)	0.9900
C(1A)-C(2A)	1.521(6)	C(13)-H(13B)	0.9900
C(1A)-H(1AA)	0.9800	C(14)-C(16)	1.506(8)
C(1A)-H(1AB)	0.9800	C(14)-C(15)	1.522(9)
C(1A)-H(1AC)	0.9800	C(14)-H(14)	1.0000
C(2A)-C(3)	1.529(5)	C(15)-H(15A)	0.9800
C(2A)-H(2AA)	0.9900	C(15)-H(15B)	0.9800
C(2A)-H(2AB)	0.9900	C(15)-H(15C)	0.9800
C(3)-C(5)	1.526(6)	C(16)-H(16A)	0.9800
C(3)-C(4)	1.546(6)	C(16)-H(16B)	0.9800
C(3)-H(3)	1.0000	C(16)-H(16C)	0.9800
C(3)-H(3A)	1.0000	C(17)-O(4)	1.222(5)
C(4)-H(4A)	0.9800	C(17)-N(3)	1.361(5)
C(4)-H(4B)	0.9800	N(3)-C(18)	1.466(5)
C(4)-H(4C)	0.9800	N(3)-C(21)	1.486(5)
C(5)-O(1)	1.428(5)	C(18)-C(22)	1.516(6)
C(5)-C(6)	1.524(6)	C(18)-C(19)	1.547(6)
C(5)-H(5)	1.0000	C(18)-H(18)	1.0000
O(1)-H(1O)	0.77(6)	C(19)-C(20)	1.529(6)
C(6)-O(2)	1.246(5)	C(19)-H(19A)	0.9900
C(6)-N(1)	1.353(5)	C(19)-H(19B)	0.9900
N(1)-C(7)	1.459(5)	C(20)-C(21)	1.530(6)
N(1)-C(10)	1.469(5)	C(20)-H(20A)	0.9900
C(7)-C(11)	1.525(5)	C(20)-H(20B)	0.9900
C(7)-C(8)	1.549(6)	C(21)-H(21A)	0.9900
C(7)-H(7)	1.0000	C(21)-H(21B)	0.9900
C(8)-C(9)	1.527(6)	C(22)-O(5)	1.232(5)
C(8)-H(8A)	0.9900	C(22)-N(4)	1.339(5)
C(8)-H(8B)	0.9900	N(4)-C(23)	1.459(5)
C(9)-C(10)	1.527(6)	N(4)-H(4N)	0.85(5)
C(9)-H(9A)	0.9900	C(23)-C(24)	1.520(6)
C(9)-H(9B)	0.9900	C(23)-C(28)	1.539(6)
C(10)-H(10A)	0.9900	C(23)-H(23)	1.0000
C(10)-H(10B)	0.9900	C(24)-C(25)	1.536(7)
C(11)-O(3)	1.229(5)	C(24)-H(24A)	0.9900
C(11)-N(2)	1.347(5)	C(24)-H(24B)	0.9900

C(25)-C(27)	1.520(10)	C(31A)-H(31A)	1.0000
C(25)-C(26)	1.521(11)	C(32A)-H(32D)	0.9800
C(25)-H(25)	1.0000	C(32A)-H(32E)	0.9800
C(26)-H(26A)	0.9800	C(32A)-H(32F)	0.9800
C(26)-H(26B)	0.9800	C(33A)-H(33D)	0.9800
C(26)-H(26C)	0.9800	C(33A)-H(33E)	0.9800
C(27)-H(27A)	0.9800	C(33A)-H(33F)	0.9800
C(27)-H(27B)	0.9800	C(34)-O(7)	1.226(5)
C(27)-H(27C)	0.9800	C(34)-N(6)	1.339(6)
C(28)-O(6)	1.236(5)	N(6)-C(35)	1.460(6)
C(28)-N(5)	1.333(6)	N(6)-H(6N)	0.75(6)
N(5)-C(29)	1.468(6)	C(35)-C(36)	1.520(6)
N(5)-H(5N)	0.78(6)	C(35)-C(37)	1.539(6)
C(29)-C(30)	1.530(6)	C(35)-H(35)	1.0000
C(29)-C(34)	1.538(6)	C(36)-O(8)	1.409(6)
C(29)-H(29)	1.0000	C(36)-H(36A)	0.9900
C(30)-C(31)	1.538(5)	C(36)-H(36B)	0.9900
C(30)-C(31A)	1.562(6)	O(8)-H(8O)	0.78(6)
C(30)-H(30A)	0.9900	C(37)-C(38)	1.509(7)
C(30)-H(30B)	0.9900	C(37)-H(37A)	0.9900
C(30)-H(30C)	0.9900	C(37)-H(37B)	0.9900
C(30)-H(30D)	0.9900	C(38)-C(39)	1.390(7)
C(31)-C(33)	1.550(6)	C(38)-C(43)	1.393(7)
C(31)-C(32)	1.553(6)	C(39)-C(40)	1.383(8)
C(31)-H(31)	1.0000	C(39)-H(39)	0.9500
C(32)-H(32A)	0.9800	C(40)-C(41)	1.398(9)
C(32)-H(32B)	0.9800	C(40)-H(40)	0.9500
C(32)-H(32C)	0.9800	C(41)-C(42)	1.375(9)
C(33)-H(33A)	0.9800	C(41)-H(41)	0.9500
C(33)-H(33B)	0.9800	C(42)-C(43)	1.386(8)
C(33)-H(33C)	0.9800	C(42)-H(42)	0.9500
C(31A)-C(33A)	1.542(6)	C(43)-H(43)	0.9500
C(31A)-C(32A)	1.548(6)		

C(2)-C(1)-H(1A)	109.5	H(2A)-C(2)-H(2B)	108.1
C(2)-C(1)-H(1B)	109.5	C(2A)-C(1A)-H(1AA)	109.5
H(1A)-C(1)-H(1B)	109.5	C(2A)-C(1A)-H(1AB)	109.5
C(2)-C(1)-H(1C)	109.5	H(1AA)-C(1A)-H(1AB)	109.5
H(1A)-C(1)-H(1C)	109.5	C(2A)-C(1A)-H(1AC)	109.4
H(1B)-C(1)-H(1C)	109.5	H(1AA)-C(1A)-H(1AC)	109.5
C(1)-C(2)-C(3)	110.6(6)	H(1AB)-C(1A)-H(1AC)	109.5
C(1)-C(2)-H(2A)	109.5	C(1A)-C(2A)-C(3)	114.4(8)
C(3)-C(2)-H(2A)	109.5	C(1A)-C(2A)-H(2AA)	108.7
C(1)-C(2)-H(2B)	109.5	C(3)-C(2A)-H(2AA)	108.7
C(3)-C(2)-H(2B)	109.5	C(1A)-C(2A)-H(2AB)	108.7

C(3)-C(2A)-H(2AB)	108.7	C(10)-C(9)-H(9A)	111.1
H(2AA)-C(2A)-H(2AB)	107.6	C(8)-C(9)-H(9A)	111.1
C(5)-C(3)-C(2A)	104.3(5)	C(10)-C(9)-H(9B)	111.1
C(5)-C(3)-C(2)	116.4(7)	C(8)-C(9)-H(9B)	111.1
C(5)-C(3)-C(4)	108.6(4)	H(9A)-C(9)-H(9B)	109.1
C(2A)-C(3)-C(4)	109.9(8)	N(1)-C(10)-C(9)	102.4(3)
C(2)-C(3)-C(4)	115.5(6)	N(1)-C(10)-H(10A)	111.3
C(5)-C(3)-H(3)	105.0	C(9)-C(10)-H(10A)	111.3
C(2)-C(3)-H(3)	105.0	N(1)-C(10)-H(10B)	111.3
C(4)-C(3)-H(3)	105.0	C(9)-C(10)-H(10B)	111.3
C(5)-C(3)-H(3A)	111.3	H(10A)-C(10)-H(10B)	109.2
C(2A)-C(3)-H(3A)	111.3	O(3)-C(11)-N(2)	123.2(4)
C(4)-C(3)-H(3A)	111.3	O(3)-C(11)-C(7)	119.5(4)
C(3)-C(4)-H(4A)	109.5	N(2)-C(11)-C(7)	117.3(4)
C(3)-C(4)-H(4B)	109.5	C(11)-N(2)-C(12)	122.0(4)
H(4A)-C(4)-H(4B)	109.5	C(11)-N(2)-H(2N)	123(4)
C(3)-C(4)-H(4C)	109.5	C(12)-N(2)-H(2N)	115(4)
H(4A)-C(4)-H(4C)	109.5	N(2)-C(12)-C(13)	111.9(4)
H(4B)-C(4)-H(4C)	109.5	N(2)-C(12)-C(17)	110.2(3)
O(1)-C(5)-C(6)	109.8(3)	C(13)-C(12)-C(17)	111.1(4)
O(1)-C(5)-C(3)	114.6(4)	N(2)-C(12)-H(12)	107.8
C(6)-C(5)-C(3)	115.9(4)	C(13)-C(12)-H(12)	107.8
O(1)-C(5)-H(5)	105.1	C(17)-C(12)-H(12)	107.8
C(6)-C(5)-H(5)	105.1	C(12)-C(13)-C(14)	113.4(4)
C(3)-C(5)-H(5)	105.1	C(12)-C(13)-H(13A)	108.9
C(5)-O(1)-H(1O)	112(4)	C(14)-C(13)-H(13A)	108.9
O(2)-C(6)-N(1)	119.0(4)	C(12)-C(13)-H(13B)	108.9
O(2)-C(6)-C(5)	119.0(4)	C(14)-C(13)-H(13B)	108.9
N(1)-C(6)-C(5)	122.0(4)	H(13A)-C(13)-H(13B)	107.7
C(6)-N(1)-C(7)	118.5(3)	C(16)-C(14)-C(15)	110.8(5)
C(6)-N(1)-C(10)	128.6(3)	C(16)-C(14)-C(13)	111.1(5)
C(7)-N(1)-C(10)	112.3(3)	C(15)-C(14)-C(13)	112.0(5)
N(1)-C(7)-C(11)	111.7(3)	C(16)-C(14)-H(14)	107.6
N(1)-C(7)-C(8)	103.9(3)	C(15)-C(14)-H(14)	107.6
C(11)-C(7)-C(8)	112.3(3)	C(13)-C(14)-H(14)	107.6
N(1)-C(7)-H(7)	109.6	C(14)-C(15)-H(15A)	109.5
C(11)-C(7)-H(7)	109.6	C(14)-C(15)-H(15B)	109.5
C(8)-C(7)-H(7)	109.6	H(15A)-C(15)-H(15B)	109.5
C(9)-C(8)-C(7)	104.6(3)	C(14)-C(15)-H(15C)	109.5
C(9)-C(8)-H(8A)	110.8	H(15A)-C(15)-H(15C)	109.5
C(7)-C(8)-H(8A)	110.8	H(15B)-C(15)-H(15C)	109.5
C(9)-C(8)-H(8B)	110.8	C(14)-C(16)-H(16A)	109.5
C(7)-C(8)-H(8B)	110.8	C(14)-C(16)-H(16B)	109.5
H(8A)-C(8)-H(8B)	108.9	H(16A)-C(16)-H(16B)	109.5
C(10)-C(9)-C(8)	103.3(3)	C(14)-C(16)-H(16C)	109.5

H(16A)-C(16)-H(16C)	109.5	C(23)-C(24)-H(24A)	108.6
H(16B)-C(16)-H(16C)	109.5	C(25)-C(24)-H(24A)	108.6
O(4)-C(17)-N(3)	120.1(4)	C(23)-C(24)-H(24B)	108.6
O(4)-C(17)-C(12)	123.0(4)	C(25)-C(24)-H(24B)	108.6
N(3)-C(17)-C(12)	116.9(4)	H(24A)-C(24)-H(24B)	107.6
C(17)-N(3)-C(18)	118.3(4)	C(27)-C(25)-C(26)	112.2(6)
C(17)-N(3)-C(21)	128.4(4)	C(27)-C(25)-C(24)	108.6(5)
C(18)-N(3)-C(21)	112.9(3)	C(26)-C(25)-C(24)	111.7(5)
N(3)-C(18)-C(22)	112.1(3)	C(27)-C(25)-H(25)	108.1
N(3)-C(18)-C(19)	103.3(3)	C(26)-C(25)-H(25)	108.1
C(22)-C(18)-C(19)	113.2(4)	C(24)-C(25)-H(25)	108.1
N(3)-C(18)-H(18)	109.4	C(25)-C(26)-H(26A)	109.5
C(22)-C(18)-H(18)	109.4	C(25)-C(26)-H(26B)	109.5
C(19)-C(18)-H(18)	109.4	H(26A)-C(26)-H(26B)	109.5
C(20)-C(19)-C(18)	104.2(3)	C(25)-C(26)-H(26C)	109.5
C(20)-C(19)-H(19A)	110.9	H(26A)-C(26)-H(26C)	109.5
C(18)-C(19)-H(19A)	110.9	H(26B)-C(26)-H(26C)	109.5
C(20)-C(19)-H(19B)	110.9	C(25)-C(27)-H(27A)	109.5
C(18)-C(19)-H(19B)	110.9	C(25)-C(27)-H(27B)	109.5
H(19A)-C(19)-H(19B)	108.9	H(27A)-C(27)-H(27B)	109.5
C(19)-C(20)-C(21)	103.8(4)	C(25)-C(27)-H(27C)	109.4
C(19)-C(20)-H(20A)	111.0	H(27A)-C(27)-H(27C)	109.5
C(21)-C(20)-H(20A)	111.0	H(27B)-C(27)-H(27C)	109.5
C(19)-C(20)-H(20B)	111.0	O(6)-C(28)-N(5)	122.8(4)
C(21)-C(20)-H(20B)	111.0	O(6)-C(28)-C(23)	121.0(4)
H(20A)-C(20)-H(20B)	109.0	N(5)-C(28)-C(23)	116.2(4)
N(3)-C(21)-C(20)	102.4(3)	C(28)-N(5)-C(29)	122.4(4)
N(3)-C(21)-H(21A)	111.3	C(28)-N(5)-H(5N)	111(4)
C(20)-C(21)-H(21A)	111.3	C(29)-N(5)-H(5N)	126(4)
N(3)-C(21)-H(21B)	111.3	N(5)-C(29)-C(30)	107.9(4)
C(20)-C(21)-H(21B)	111.3	N(5)-C(29)-C(34)	111.7(4)
H(21A)-C(21)-H(21B)	109.2	C(30)-C(29)-C(34)	113.4(4)
O(5)-C(22)-N(4)	122.2(4)	N(5)-C(29)-H(29)	107.9
O(5)-C(22)-C(18)	120.9(4)	C(30)-C(29)-H(29)	107.9
N(4)-C(22)-C(18)	117.0(4)	C(34)-C(29)-H(29)	107.9
C(22)-N(4)-C(23)	121.9(4)	C(29)-C(30)-C(31)	117.9(4)
C(22)-N(4)-H(4N)	120(4)	C(29)-C(30)-C(31A)	107.2(5)
C(23)-N(4)-H(4N)	118(4)	C(29)-C(30)-H(30A)	107.8
N(4)-C(23)-C(24)	112.9(4)	C(31)-C(30)-H(30A)	107.8
N(4)-C(23)-C(28)	110.2(3)	C(29)-C(30)-H(30B)	107.8
C(24)-C(23)-C(28)	111.9(4)	C(31)-C(30)-H(30B)	107.8
N(4)-C(23)-H(23)	107.2	H(30A)-C(30)-H(30B)	107.2
C(24)-C(23)-H(23)	107.2	C(29)-C(30)-H(30C)	110.3
C(28)-C(23)-H(23)	107.2	C(31A)-C(30)-H(30C)	110.3
C(23)-C(24)-C(25)	114.6(4)	C(29)-C(30)-H(30D)	110.3

C(31A)-C(30)-H(30D)	110.3	C(34)-N(6)-C(35)	122.0(4)
H(30C)-C(30)-H(30D)	108.5	C(34)-N(6)-H(6N)	121(4)
C(30)-C(31)-C(33)	108.4(5)	C(35)-N(6)-H(6N)	117(4)
C(30)-C(31)-C(32)	112.0(5)	N(6)-C(35)-C(36)	111.8(4)
C(33)-C(31)-C(32)	105.9(7)	N(6)-C(35)-C(37)	110.1(4)
C(30)-C(31)-H(31)	110.1	C(36)-C(35)-C(37)	111.3(4)
C(33)-C(31)-H(31)	110.1	N(6)-C(35)-H(35)	107.8
C(32)-C(31)-H(31)	110.1	C(36)-C(35)-H(35)	107.8
C(31)-C(32)-H(32A)	109.5	C(37)-C(35)-H(35)	107.8
C(31)-C(32)-H(32B)	109.4	O(8)-C(36)-C(35)	113.8(4)
H(32A)-C(32)-H(32B)	109.5	O(8)-C(36)-H(36A)	108.8
C(31)-C(32)-H(32C)	109.5	C(35)-C(36)-H(36A)	108.8
H(32A)-C(32)-H(32C)	109.5	O(8)-C(36)-H(36B)	108.8
H(32B)-C(32)-H(32C)	109.5	C(35)-C(36)-H(36B)	108.8
C(31)-C(33)-H(33A)	109.5	H(36A)-C(36)-H(36B)	107.7
C(31)-C(33)-H(33B)	109.5	C(36)-O(8)-H(8O)	97(4)
H(33A)-C(33)-H(33B)	109.5	C(38)-C(37)-C(35)	112.9(4)
C(31)-C(33)-H(33C)	109.4	C(38)-C(37)-H(37A)	109.0
H(33A)-C(33)-H(33C)	109.5	C(35)-C(37)-H(37A)	109.0
H(33B)-C(33)-H(33C)	109.5	C(38)-C(37)-H(37B)	109.0
C(33A)-C(31A)-C(32A)	111.2(8)	C(35)-C(37)-H(37B)	109.0
C(33A)-C(31A)-C(30)	111.2(6)	H(37A)-C(37)-H(37B)	107.8
C(32A)-C(31A)-C(30)	108.9(5)	C(39)-C(38)-C(43)	117.8(5)
C(33A)-C(31A)-H(31A)	108.5	C(39)-C(38)-C(37)	121.4(5)
C(32A)-C(31A)-H(31A)	108.5	C(43)-C(38)-C(37)	120.8(5)
C(30)-C(31A)-H(31A)	108.5	C(40)-C(39)-C(38)	121.6(5)
C(31A)-C(32A)-H(32D)	109.4	C(40)-C(39)-H(39)	119.2
C(31A)-C(32A)-H(32E)	109.5	C(38)-C(39)-H(39)	119.2
H(32D)-C(32A)-H(32E)	109.5	C(39)-C(40)-C(41)	119.7(5)
C(31A)-C(32A)-H(32F)	109.5	C(39)-C(40)-H(40)	120.1
H(32D)-C(32A)-H(32F)	109.5	C(41)-C(40)-H(40)	120.1
H(32E)-C(32A)-H(32F)	109.5	C(42)-C(41)-C(40)	119.2(5)
C(31A)-C(33A)-H(33D)	109.4	C(42)-C(41)-H(41)	120.4
C(31A)-C(33A)-H(33E)	109.5	C(40)-C(41)-H(41)	120.4
H(33D)-C(33A)-H(33E)	109.5	C(41)-C(42)-C(43)	120.7(5)
C(31A)-C(33A)-H(33F)	109.5	C(41)-C(42)-H(42)	119.6
H(33D)-C(33A)-H(33F)	109.5	C(43)-C(42)-H(42)	119.6
H(33E)-C(33A)-H(33F)	109.5	C(42)-C(43)-C(38)	120.9(5)
O(7)-C(34)-N(6)	124.2(4)	C(42)-C(43)-H(43)	119.5
O(7)-C(34)-C(29)	121.2(4)	C(38)-C(43)-H(43)	119.5
N(6)-C(34)-C(29)	114.5(4)		

Table A6. Ansiotropic displacement parameters (A^2 x 10^3) for **9**. The anisotropic displacement factor exponent takes the form $-2\pi^2[h^2a^{*2}U_{11}+\ldots+2h\,k\,a^*\,b^*\,U_{12}]$

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C(1)	40(5)	32(5)	24(4)	-1(4)	7(4)	3(4)
C(2)	17(6)	49(6)	26(3)	-11(4)	3(4)	9(5)
C(1A)	43(7)	40(7)	21(5)	3(5)	6(5)	6(5)
C(2A)	17(6)	49(6)	26(3)	-11(4)	3(4)	9(5)
C(3)	25(2)	34(3)	19(2)	-2(2)	1(2)	7(2)
C(4)	40(3)	35(3)	19(2)	1(2)	-1(2)	11(2)
C(5)	21(2)	25(2)	22(2)	-4(2)	-2(2)	2(2)
O(1)	26(2)	24(2)	25(2)	2(1)	-5(1)	3(1)
C(6)	21(2)	18(2)	12(2)	3(1)	0(2)	0(2)
O(2)	22(1)	26(2)	20(2)	-1(1)	-2(1)	-7(1)
N(1)	17(2)	23(2)	15(2)	0(1)	1(1)	-3(1)
C(7)	19(2)	20(2)	16(2)	1(2)	1(2)	0(2)
C(8)	24(2)	18(2)	20(2)	-1(2)	1(2)	-2(2)
C(9)	19(2)	17(2)	22(2)	0(2)	-1(2)	-1(2)
C(10)	16(2)	18(2)	26(2)	2(2)	2(2)	-1(2)
C(11)	21(2)	17(2)	18(2)	1(2)	0(2)	-4(2)
O(3)	20(1)	26(2)	21(2)	0(1)	2(1)	-1(1)
N(2)	18(2)	25(2)	16(2)	2(1)	2(1)	-1(2)
C(12)	24(2)	25(2)	15(2)	5(2)	-2(2)	-3(2)
C(13)	37(3)	29(2)	19(2)	-2(2)	6(2)	-8(2)
C(14)	51(3)	30(3)	21(2)	-4(2)	4(2)	-13(2)
C(15)	60(4)	59(4)	59(4)	-20(3)	-31(3)	6(3)
C(16)	69(4)	44(3)	37(3)	-19(3)	-1(3)	-5(3)
C(17)	27(2)	17(2)	17(2)	4(2)	0(2)	-2(2)
O(4)	21(1)	26(2)	25(2)	-2(1)	6(1)	-3(1)
N(3)	18(2)	18(2)	23(2)	0(1)	2(1)	0(1)
C(18)	18(2)	17(2)	25(2)	-1(2)	-1(2)	0(2)
C(19)	23(2)	17(2)	31(2)	-2(2)	0(2)	5(2)
C(20)	18(2)	24(2)	34(2)	-3(2)	3(2)	2(2)
C(21)	19(2)	22(2)	26(2)	1(2)	2(2)	2(2)
C(22)	18(2)	17(2)	19(2)	-2(2)	1(2)	3(2)
O(5)	18(1)	24(2)	27(2)	-3(1)	1(1)	-4(1)
N(4)	14(2)	21(2)	20(2)	2(1)	2(1)	2(1)
C(23)	18(2)	21(2)	23(2)	4(2)	1(2)	6(2)
C(24)	23(2)	49(3)	23(2)	2(2)	-5(2)	7(2)
C(25)	33(3)	79(5)	23(2)	-5(3)	-3(2)	18(3)

C(26)	44(3)	108(6)	33(3)	30(4)	18(3)	15(4)
C(27)	66(5)	153(9)	43(4)	-43(5)	-3(4)	10(6)
C(28)	21(2)	18(2)	25(2)	4(2)	6(2)	3(2)
O(6)	16(1)	32(2)	32(2)	5(1)	1(1)	4(1)
N(5)	19(2)	23(2)	32(2)	-1(2)	5(2)	2(2)
C(29)	25(2)	20(2)	35(2)	-4(2)	6(2)	2(2)
C(30)	37(3)	43(3)	58(3)	-31(3)	26(3)	-19(2)
C(31)	39(4)	34(4)	29(3)	-7(3)	3(3)	0(3)
C(32)	46(4)	35(4)	33(4)	-5(3)	-2(3)	-4(4)
C(33)	47(4)	47(4)	40(4)	-13(4)	-4(3)	12(3)
C(31A)	36(4)	28(4)	32(4)	-10(3)	1(3)	1(3)
C(32A)	46(5)	37(4)	38(4)	-8(4)	10(3)	0(4)
C(33A)	49(5)	32(4)	35(4)	-10(3)	-3(4)	-5(4)
C(34)	26(2)	23(2)	15(2)	-2(2)	0(2)	3(2)
O(7)	19(2)	28(2)	30(2)	4(1)	2(1)	1(1)
N(6)	17(2)	23(2)	22(2)	2(1)	1(2)	1(2)
C(35)	29(2)	22(2)	20(2)	7(2)	2(2)	0(2)
C(36)	30(2)	27(2)	24(2)	0(2)	2(2)	-6(2)
O(8)	36(2)	21(2)	30(2)	-2(1)	7(2)	-4(1)
C(37)	42(3)	24(2)	20(2)	4(2)	-3(2)	4(2)
C(38)	36(3)	24(2)	21(2)	7(2)	-3(2)	3(2)
C(39)	34(3)	45(3)	28(2)	7(2)	-4(2)	1(2)
C(40)	43(3)	46(3)	41(3)	10(3)	-18(3)	-8(3)
C(41)	74(4)	33(3)	23(2)	3(2)	-12(3)	-9(3)
C(42)	57(4)	37(3)	28(3)	4(2)	9(2)	0(3)
C(43)	42(3)	33(3)	25(2)	8(2)	6(2)	-9(2)

Table A7. Hydrogen coordinates and isotropic displacement parameters for **9**

	x	y	z	U(eq)
H(1A)	-0.0917	0.7580	0.9192	0.048
H(1B)	0.0150	0.6901	0.9327	0.048
H(1C)	0.0250	0.7530	0.8835	0.048
H(2A)	-0.1425	0.6747	0.8467	0.037
H(2B)	-0.1422	0.6085	0.8948	0.037
H(1AA)	-0.1525	0.7002	0.9373	0.052
H(1AB)	-0.1770	0.6000	0.9184	0.052
H(1AC)	-0.0541	0.6238	0.9472	0.052
H(2AA)	0.0054	0.7146	0.8780	0.037
H(2AB)	-0.1179	0.6923	0.8492	0.037
H(3)	-0.0418	0.5348	0.8332	0.031
H(3A)	-0.0568	0.5413	0.8364	0.031
H(4A)	0.0629	0.5203	0.9117	0.048
H(4B)	0.1472	0.5019	0.8633	0.048
H(4C)	0.1562	0.5952	0.8930	0.048
H(5)	0.1089	0.6818	0.8127	0.027
H(1O)	-0.074(5)	0.633(4)	0.760(2)	0.030
H(7)	0.2953	0.4267	0.7523	0.022
H(8A)	0.2021	0.3426	0.6711	0.025
H(8B)	0.1865	0.3064	0.7286	0.025
H(9A)	0.0172	0.4122	0.6716	0.024
H(9B)	-0.0124	0.3270	0.7072	0.024
H(10A)	-0.0475	0.4838	0.7449	0.024
H(10B)	0.0108	0.4086	0.7816	0.024
H(2N)	0.150(5)	0.559(4)	0.662(2)	0.023
H(12)	0.2026	0.6432	0.6001	0.026
H(13A)	0.3721	0.5032	0.5792	0.034
H(13B)	0.3366	0.5805	0.5398	0.034
H(14)	0.1610	0.4568	0.5775	0.041
H(15A)	0.0442	0.5002	0.5068	0.089
H(15B)	0.0755	0.5852	0.5415	0.089
H(15C)	0.1472	0.5686	0.4894	0.089
H(16A)	0.2920	0.4347	0.4852	0.075
H(16B)	0.3048	0.3724	0.5346	0.075
H(16C)	0.1819	0.3728	0.5027	0.075
H(18)	0.4938	0.8074	0.6553	0.024
H(19A)	0.4326	0.8729	0.7454	0.028
H(19B)	0.3719	0.9087	0.6934	0.028
H(20A)	0.2035	0.8600	0.7371	0.031

H(20B)	0.2744	0.7765	0.7620	0.031
H(21A)	0.1806	0.7035	0.6957	0.027
H(21B)	0.1863	0.7902	0.6591	0.027
H(4N)	0.417(5)	0.643(3)	0.7396(19)	0.022
H(23)	0.5184	0.5413	0.7798	0.025
H(24A)	0.6468	0.6941	0.8152	0.038
H(24B)	0.6692	0.5959	0.8375	0.038
H(25)	0.4399	0.6838	0.8436	0.054
H(26A)	0.3920	0.5731	0.9052	0.092
H(26B)	0.4202	0.5286	0.8509	0.092
H(26C)	0.5193	0.5259	0.8952	0.092
H(27A)	0.6167	0.6747	0.9226	0.131
H(27B)	0.5870	0.7631	0.8905	0.131
H(27C)	0.4877	0.7215	0.9279	0.131
H(5N)	0.581(5)	0.538(4)	0.687(2)	0.030
H(29)	0.7887	0.4568	0.6762	0.032
H(30A)	0.6089	0.4948	0.6025	0.055
H(30B)	0.6430	0.3999	0.6261	0.055
H(30C)	0.6423	0.5014	0.5894	0.055
H(30D)	0.6116	0.4182	0.6261	0.055
H(31)	0.7969	0.4880	0.5533	0.041
H(32A)	0.6231	0.3476	0.5328	0.057
H(32B)	0.6150	0.4467	0.5098	0.057
H(32C)	0.7206	0.3808	0.4920	0.057
H(33A)	0.7939	0.3075	0.5935	0.067
H(33B)	0.8773	0.3355	0.5463	0.067
H(33C)	0.9004	0.3784	0.6014	0.067
H(31A)	0.8413	0.3863	0.6082	0.038
H(32D)	0.7535	0.4779	0.5193	0.060
H(32E)	0.8745	0.4996	0.5501	0.060
H(32F)	0.8663	0.4115	0.5155	0.060
H(33D)	0.6367	0.3172	0.5551	0.058
H(33E)	0.7683	0.2824	0.5401	0.058
H(33F)	0.7177	0.2666	0.5965	0.058
H(6N)	0.713(5)	0.657(4)	0.624(2)	0.025
H(35)	0.9252	0.6987	0.5860	0.028
H(36A)	0.9213	0.7574	0.6715	0.032
H(36B)	0.9524	0.8296	0.6285	0.032
H(8O)	0.759(5)	0.809(4)	0.675(2)	0.035
H(37A)	0.7020	0.7930	0.5720	0.035
H(37B)	0.8259	0.8236	0.5462	0.035
H(39)	0.5814	0.6803	0.5312	0.042
H(40)	0.5410	0.5912	0.4606	0.052
H(41)	0.6956	0.5590	0.4018	0.052
H(42)	0.8840	0.6246	0.4125	0.049

H(43)	0.9224	0.7169	0.4820	0.040

Table 8. Torsion angles [°] for **9**.

C(1A)-C(2A)-C(3)-C(5)	-174.8(11)
C(1A)-C(2A)-C(3)-C(2)	53(3)
C(1A)-C(2A)-C(3)-C(4)	-58.6(14)
C(1)-C(2)-C(3)-C(5)	-77.1(11)
C(1)-C(2)-C(3)-C(2A)	-24(3)
C(1)-C(2)-C(3)-C(4)	52.1(12)
C(2A)-C(3)-C(5)-O(1)	-67.9(8)
C(2)-C(3)-C(5)-O(1)	-52.6(6)
C(4)-C(3)-C(5)-O(1)	175.0(4)
C(2A)-C(3)-C(5)-C(6)	162.5(8)
C(2)-C(3)-C(5)-C(6)	177.8(5)
C(4)-C(3)-C(5)-C(6)	45.4(5)
O(1)-C(5)-C(6)-O(2)	112.9(4)
C(3)-C(5)-C(6)-O(2)	-115.3(4)
O(1)-C(5)-C(6)-N(1)	-63.2(5)
C(3)-C(5)-C(6)-N(1)	68.7(5)
O(2)-C(6)-N(1)-C(7)	1.7(5)
C(5)-C(6)-N(1)-C(7)	177.8(4)
O(2)-C(6)-N(1)-C(10)	172.3(4)
C(5)-C(6)-N(1)-C(10)	-11.6(6)
C(6)-N(1)-C(7)-C(11)	-62.7(5)
C(10)-N(1)-C(7)-C(11)	125.2(4)
C(6)-N(1)-C(7)-C(8)	176.1(3)
C(10)-N(1)-C(7)-C(8)	3.9(4)
N(1)-C(7)-C(8)-C(9)	19.0(4)
C(11)-C(7)-C(8)-C(9)	-101.8(4)
C(7)-C(8)-C(9)-C(10)	-34.1(4)
C(6)-N(1)-C(10)-C(9)	163.8(4)
C(7)-N(1)-C(10)-C(9)	-25.1(4)
C(8)-C(9)-C(10)-N(1)	35.7(4)
N(1)-C(7)-C(11)-O(3)	146.6(4)
C(8)-C(7)-C(11)-O(3)	-97.1(5)
N(1)-C(7)-C(11)-N(2)	-35.6(5)
C(8)-C(7)-C(11)-N(2)	80.6(5)
O(3)-C(11)-N(2)-C(12)	0.8(6)
C(7)-C(11)-N(2)-C(12)	-176.9(4)
C(11)-N(2)-C(12)-C(13)	68.4(5)
C(11)-N(2)-C(12)-C(17)	-55.7(5)
N(2)-C(12)-C(13)-C(14)	58.8(5)
C(17)-C(12)-C(13)-C(14)	-177.5(4)
C(12)-C(13)-C(14)-C(16)	-164.3(4)
C(12)-C(13)-C(14)-C(15)	71.3(6)

N(2)-C(12)-C(17)-O(4)	116.7(5)
C(13)-C(12)-C(17)-O(4)	-7.9(6)
N(2)-C(12)-C(17)-N(3)	-61.6(5)
C(13)-C(12)-C(17)-N(3)	173.8(4)
O(4)-C(17)-N(3)-C(18)	-1.5(6)
C(12)-C(17)-N(3)-C(18)	176.9(3)
O(4)-C(17)-N(3)-C(21)	170.9(4)
C(12)-C(17)-N(3)-C(21)	-10.8(6)
C(17)-N(3)-C(18)-C(22)	-69.5(5)
C(21)-N(3)-C(18)-C(22)	117.0(4)
C(17)-N(3)-C(18)-C(19)	168.4(4)
C(21)-N(3)-C(18)-C(19)	-5.1(4)
N(3)-C(18)-C(19)-C(20)	25.6(4)
C(22)-C(18)-C(19)-C(20)	-95.8(4)
C(18)-C(19)-C(20)-C(21)	-36.6(4)
C(17)-N(3)-C(21)-C(20)	170.0(4)
C(18)-N(3)-C(21)-C(20)	-17.3(4)
C(19)-C(20)-C(21)-N(3)	32.6(4)
N(3)-C(18)-C(22)-O(5)	151.8(4)
C(19)-C(18)-C(22)-O(5)	-91.8(5)
N(3)-C(18)-C(22)-N(4)	-28.6(5)
C(19)-C(18)-C(22)-N(4)	87.8(4)
O(5)-C(22)-N(4)-C(23)	-4.4(6)
C(18)-C(22)-N(4)-C(23)	176.0(4)
C(22)-N(4)-C(23)-C(24)	75.0(5)
C(22)-N(4)-C(23)-C(28)	-50.9(5)
N(4)-C(23)-C(24)-C(25)	67.4(6)
C(28)-C(23)-C(24)-C(25)	-167.6(4)
C(23)-C(24)-C(25)-C(27)	-168.4(6)
C(23)-C(24)-C(25)-C(26)	67.4(6)
N(4)-C(23)-C(28)-O(6)	139.2(4)
C(24)-C(23)-C(28)-O(6)	12.6(6)
N(4)-C(23)-C(28)-N(5)	-44.0(5)
C(24)-C(23)-C(28)-N(5)	-170.6(4)
O(6)-C(28)-N(5)-C(29)	0.6(7)
C(23)-C(28)-N(5)-C(29)	-176.2(4)
C(28)-N(5)-C(29)-C(30)	162.5(4)
C(28)-N(5)-C(29)-C(34)	-72.2(5)
N(5)-C(29)-C(30)-C(31)	164.3(5)
C(34)-C(29)-C(30)-C(31)	40.0(7)
N(5)-C(29)-C(30)-C(31A)	-166.2(4)
C(34)-C(29)-C(30)-C(31A)	69.5(5)
C(29)-C(30)-C(31)-C(33)	71.3(8)
C(31A)-C(30)-C(31)-C(33)	-3.6(8)
C(29)-C(30)-C(31)-C(32)	-172.2(7)

C(31A)-C(30)-C(31)-C(32)	113.0(11)
C(29)-C(30)-C(31A)-C(33A)	141.0(8)
C(31)-C(30)-C(31A)-C(33A)	-102.3(11)
C(29)-C(30)-C(31A)-C(32A)	-96.2(8)
C(31)-C(30)-C(31A)-C(32A)	20.6(8)
N(5)-C(29)-C(34)-O(7)	135.0(4)
C(30)-C(29)-C(34)-O(7)	-102.8(5)
N(5)-C(29)-C(34)-N(6)	-48.8(5)
C(30)-C(29)-C(34)-N(6)	73.5(5)
O(7)-C(34)-N(6)-C(35)	-0.2(7)
C(29)-C(34)-N(6)-C(35)	-176.3(4)
C(34)-N(6)-C(35)-C(36)	-89.7(5)
C(34)-N(6)-C(35)-C(37)	146.1(4)
N(6)-C(35)-C(36)-O(8)	-73.8(5)
C(37)-C(35)-C(36)-O(8)	49.7(5)
N(6)-C(35)-C(37)-C(38)	-63.0(5)
C(36)-C(35)-C(37)-C(38)	172.5(4)
C(35)-C(37)-C(38)-C(39)	101.2(5)
C(35)-C(37)-C(38)-C(43)	-78.0(6)
C(43)-C(38)-C(39)-C(40)	-2.3(7)
C(37)-C(38)-C(39)-C(40)	178.5(5)
C(38)-C(39)-C(40)-C(41)	2.6(8)
C(39)-C(40)-C(41)-C(42)	-1.8(8)
C(40)-C(41)-C(42)-C(43)	0.9(8)
C(41)-C(42)-C(43)-C(38)	-0.7(8)
C(39)-C(38)-C(43)-C(42)	1.3(7)
C(37)-C(38)-C(43)-C(42)	-179.5(5)

Table A9. Hydrogen bonds for **9** [Å and°]

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(1)-H(1O)...O(6)#1	0.77(6)	2.00(6)	2.772(5)	178(6)
C(8)-H(8B)...O(5)#2	0.99	2.48	3.272(5)	136.9
C(9)-H(9A)...O(7)#1	0.99	2.54	3.368(5)	141.0
C(9)-H(9B)...O(1)#3	0.99	2.64	3.431(5)	137.4
C(10)-H(10A)...O(6)#1	0.99	2.26	3.221(5)	162.0
N(2)-H(2N)...O(7)#1	0.82(5)	2.41(5)	3.164(5)	153(5)
C(13)-H(13A)...O(3)	0.99	2.56	3.130(6)	116.2
C(21)-H(21A)...N(2)	0.99	2.55	3.143(6)	118.7
N(4)-H(4N)...O(2)	0.85(5)	2.02(5)	2.762(5)	147(5)
C(24)-H(24A)...O(5)	0.99	2.62	3.178(6)	115.8
N(5)-H(5N)...O(3)	0.78(6)	2.14(6)	2.891(5)	161(6)
N(6)-H(6N)...O(4)	0.75(6)	2.64(6)	3.374(5)	167(6)
O(8)-H(8O)...O(5)	0.78(6)	1.96(6)	2.718(5)	162(6)

Symmetry transformations used to generate equivalent atoms:

#1 x-1, y, z #2 -x+1, y-1/2, -z+3/2 #3 -x, y-1/2, -z+3/2

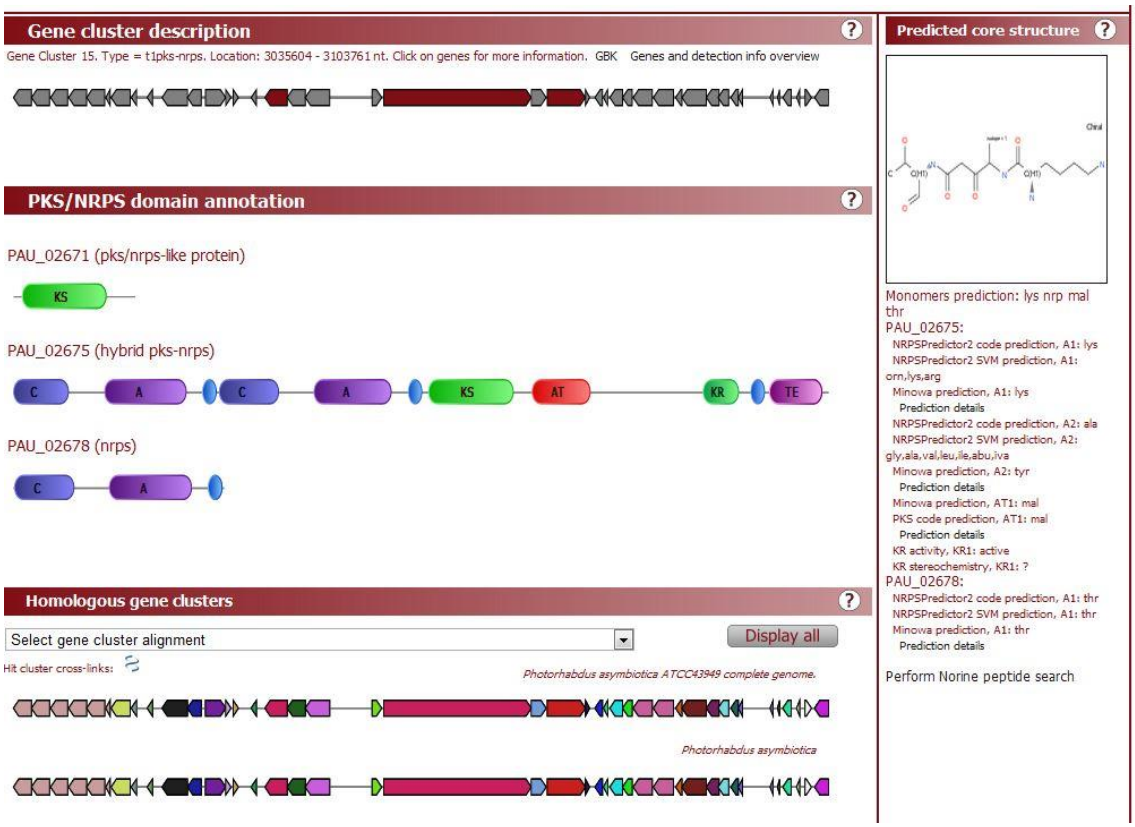


Figure A1. Screenshot of antiSMASH results for secondary metabolite gene cluster 15; the cluster predicted to produce glidobactin-like compounds. KS = ketosynthase, A = AMP binding domain, C = condensation domain, AT = acyltransferase, KR = ketoreductase, TE = thioesterase

Figure A2.1H-NMR (500 MHz, DMSO*d*6, 25°C) spectrum of compound **3**

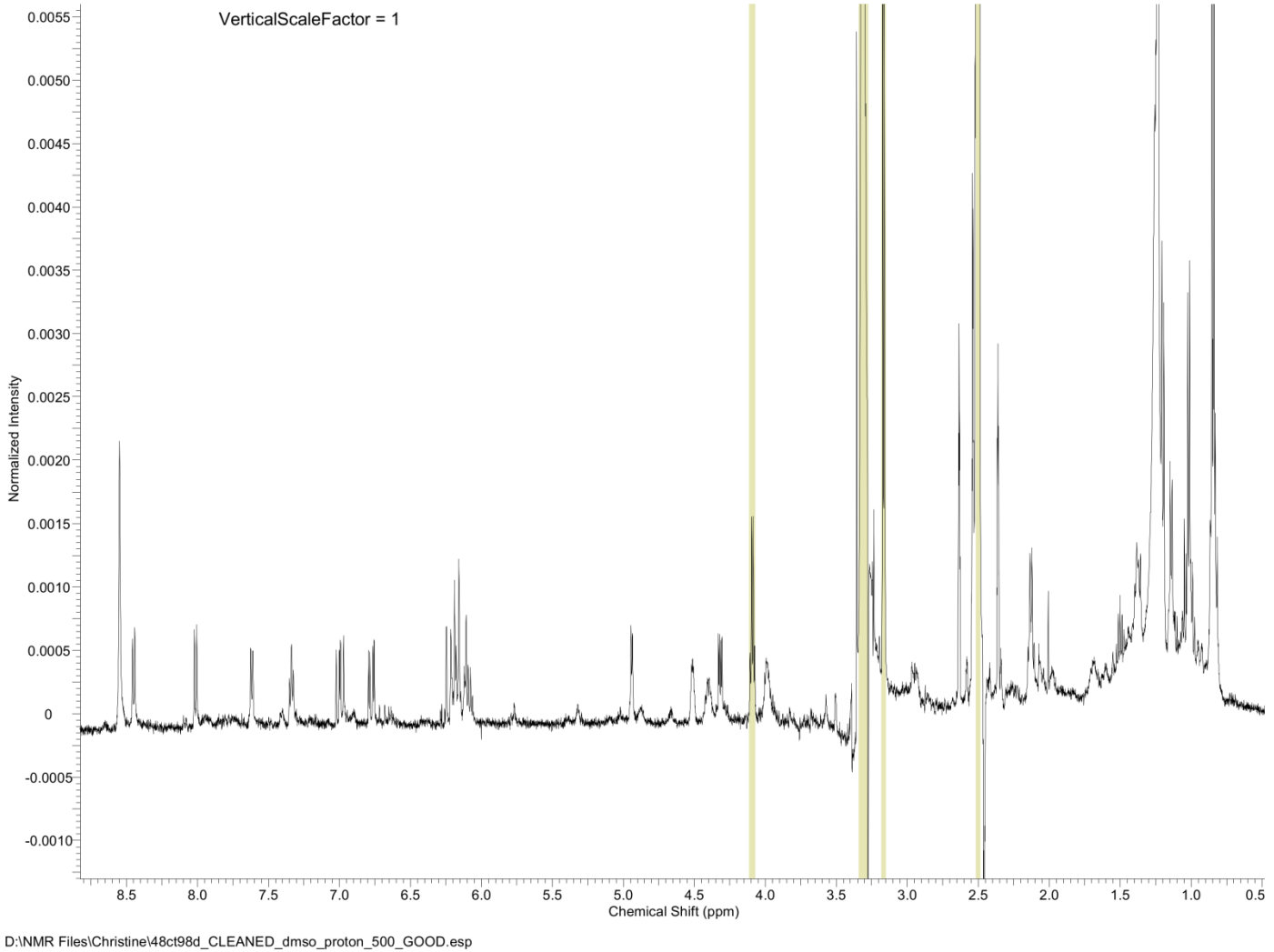


Figure A3. ^{1}H-^{13}C HSQC (500 MHz, DMSO*d*6, 25°C) spectrum of compound **3**

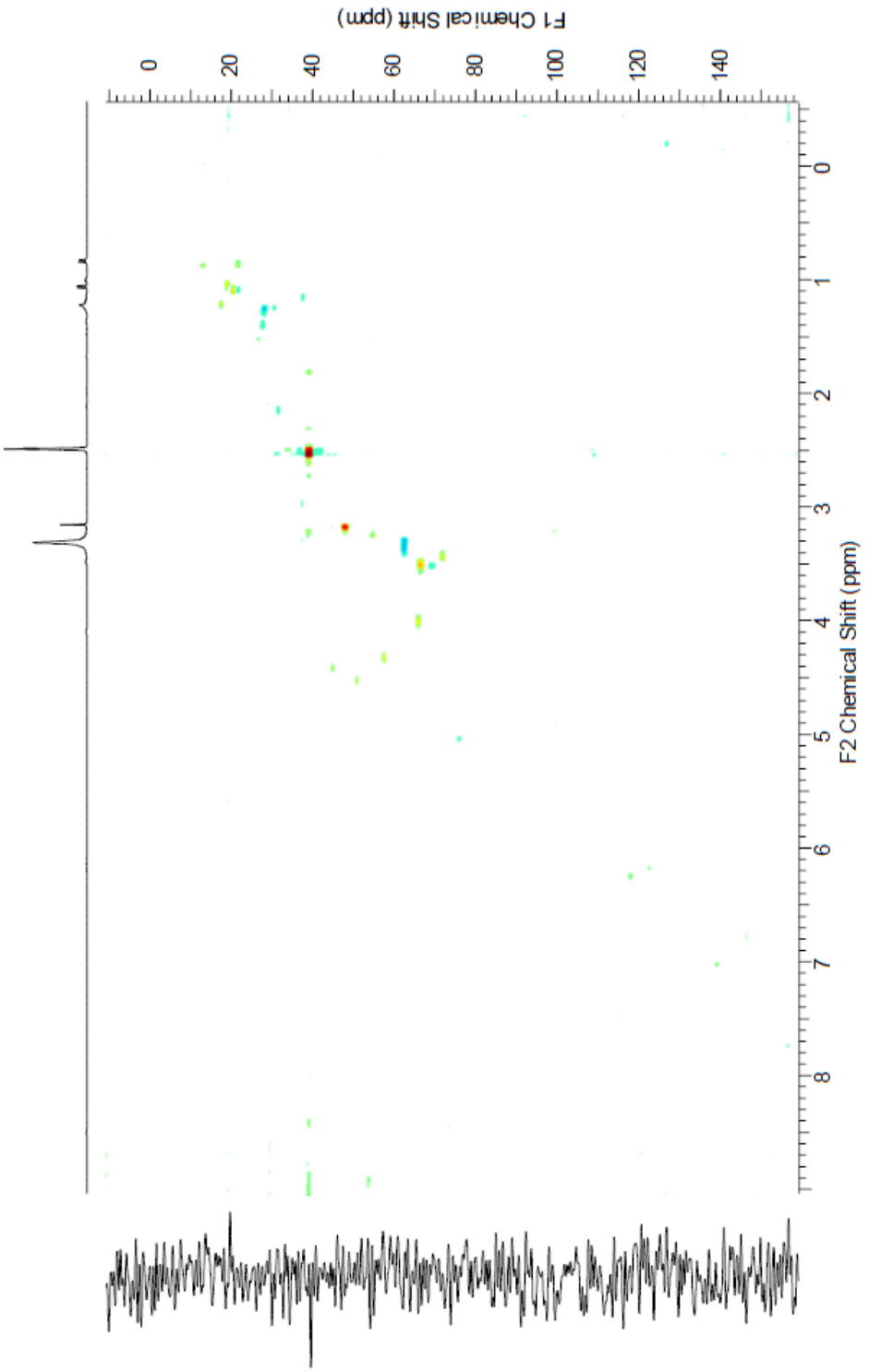
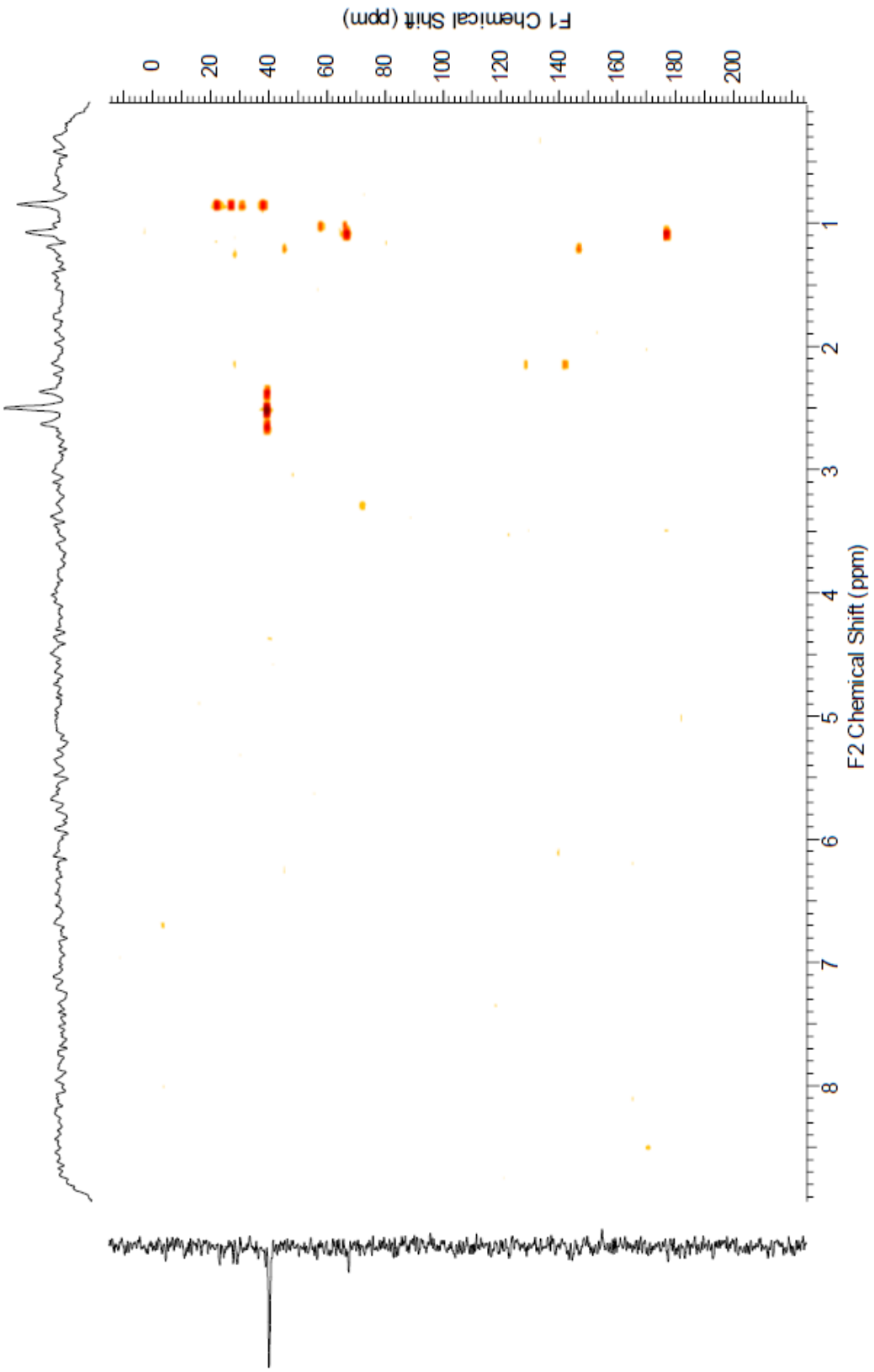


Figure A4. ^{1}H-^{13}C gHMBC (500 MHz, DMSO*d6*, 25°C) spectrum of compound **3**



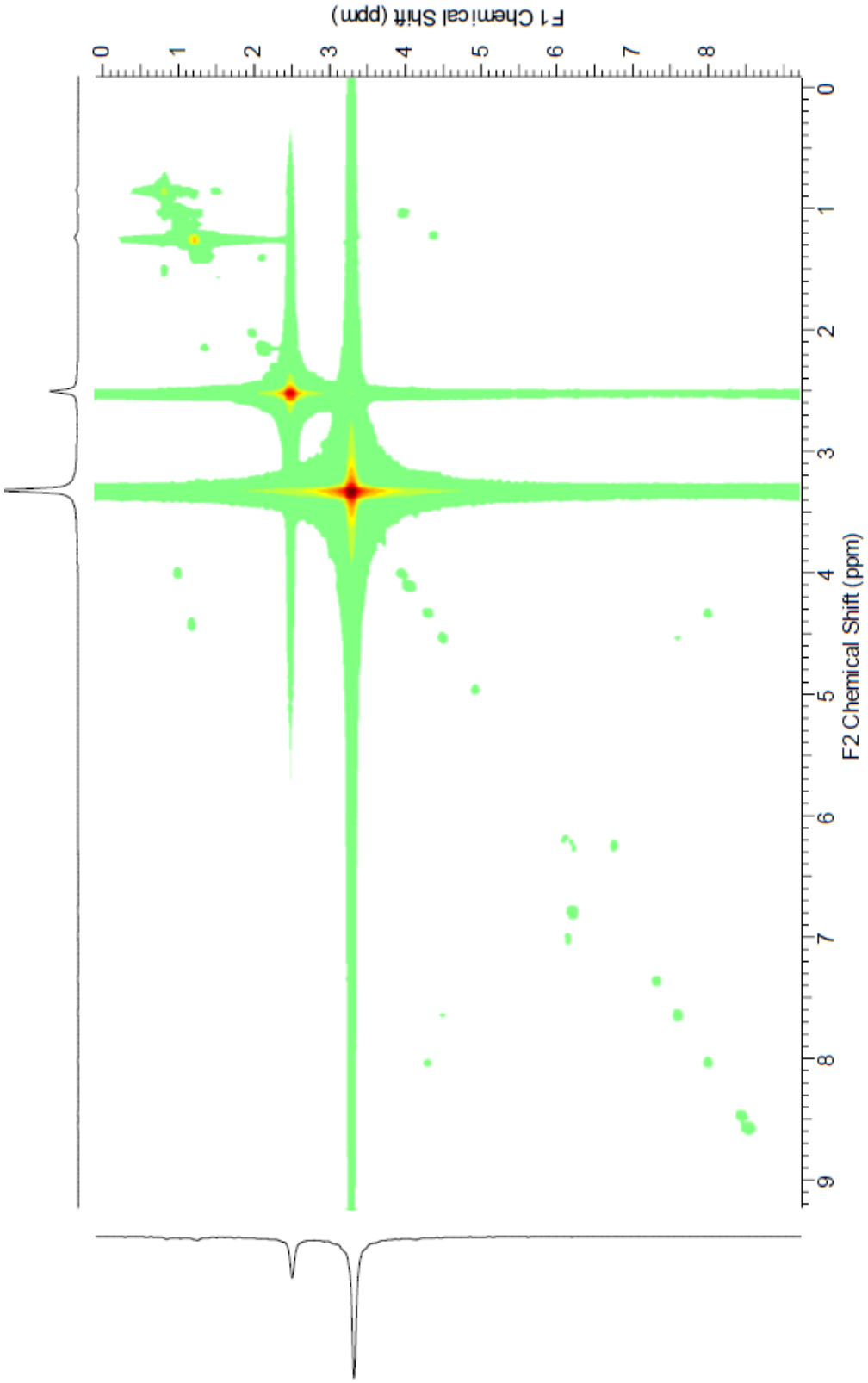


Figure A5. ^{1}H-^{1}H gCOSY (500 MHz, DMSO*d*6, 25°C) spectrum of compound **3**

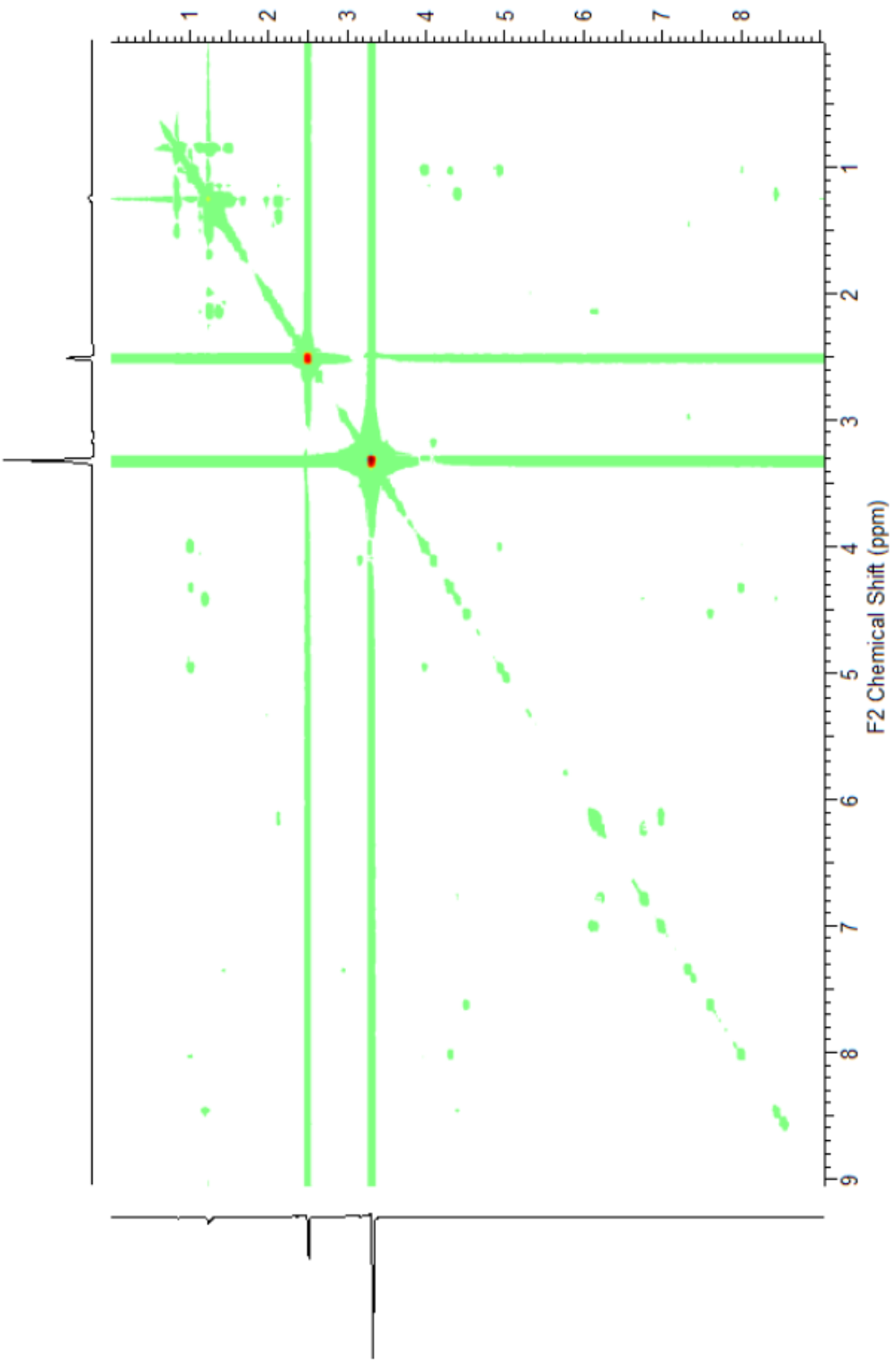


Figure A6. ^{1}H-^{1}H tCOSY (500 MHz, DMSO*d6*, 25°C) spectrum of compound **3**

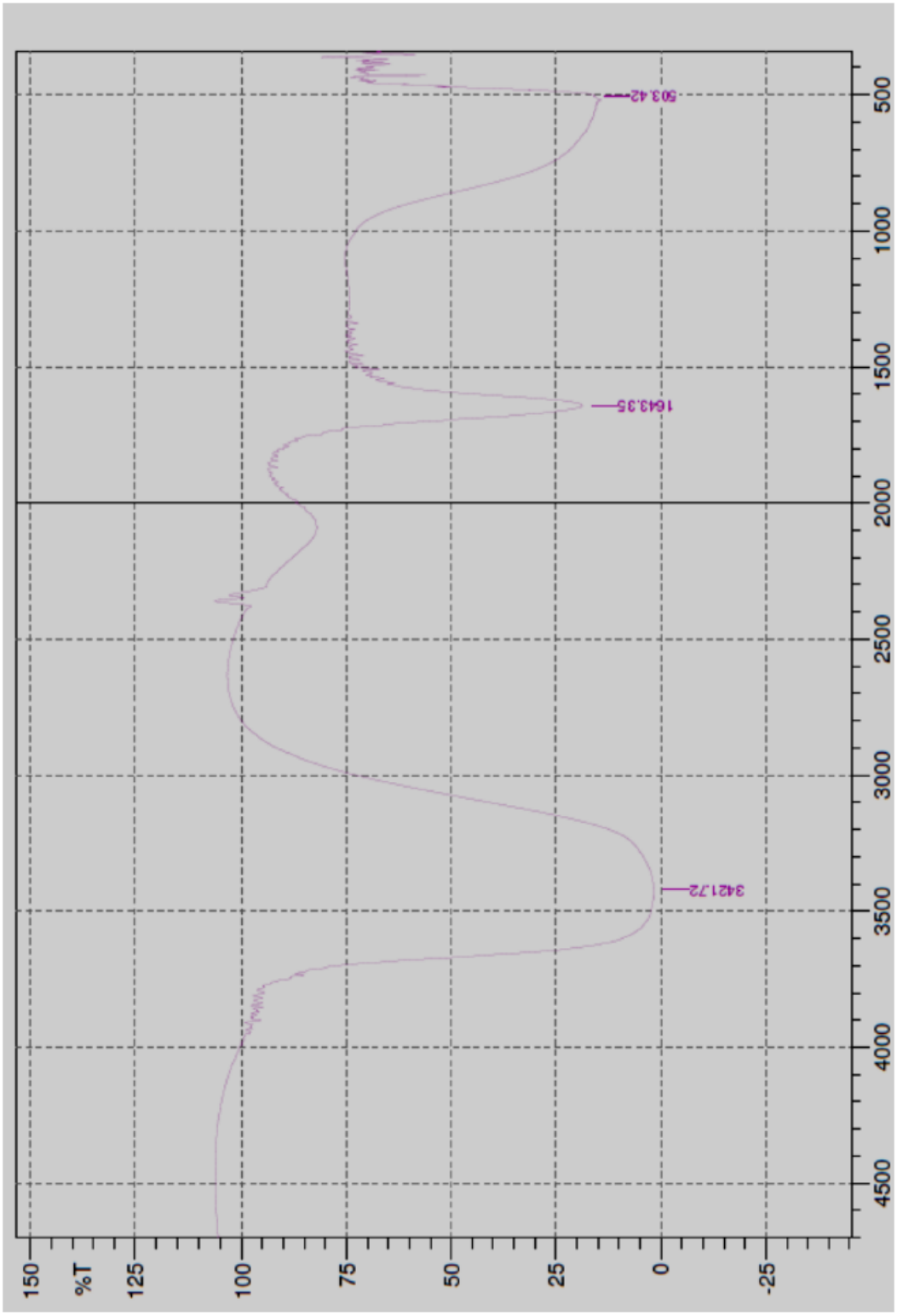
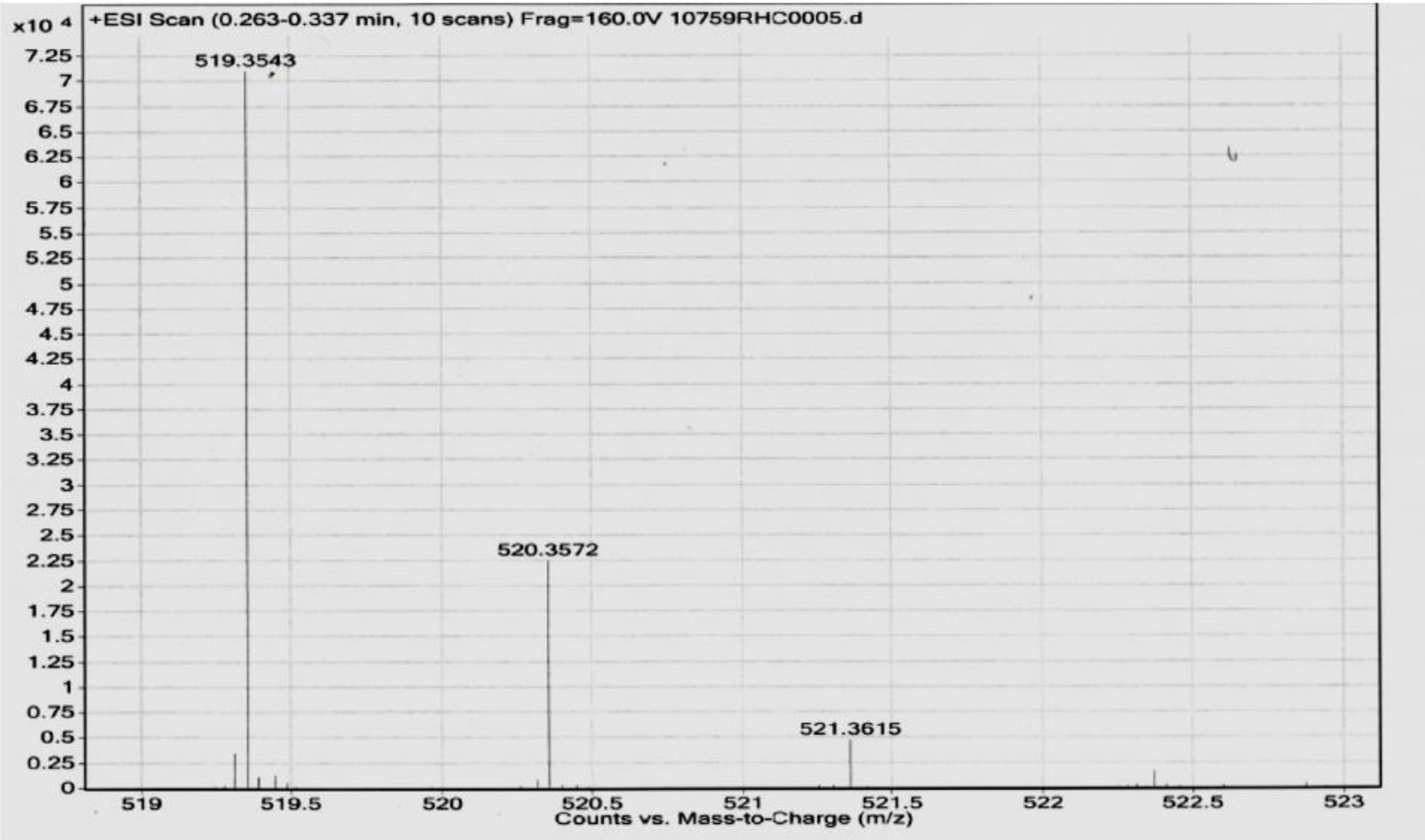


Figure A7. IR spectrum (thin film, room temperature) of compound **3**

Figure A8. HRESI MS (positive mode) for compound **3**



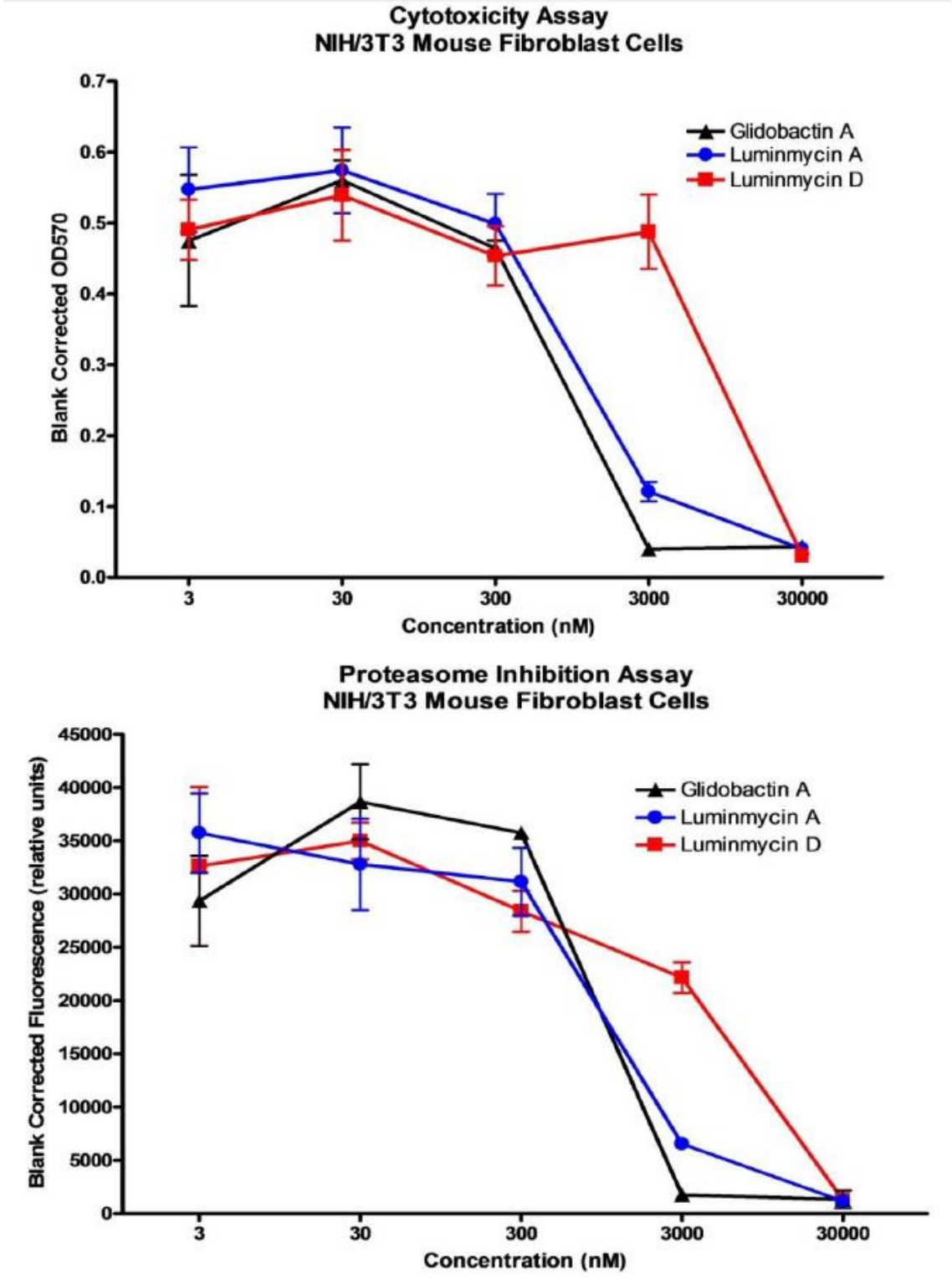
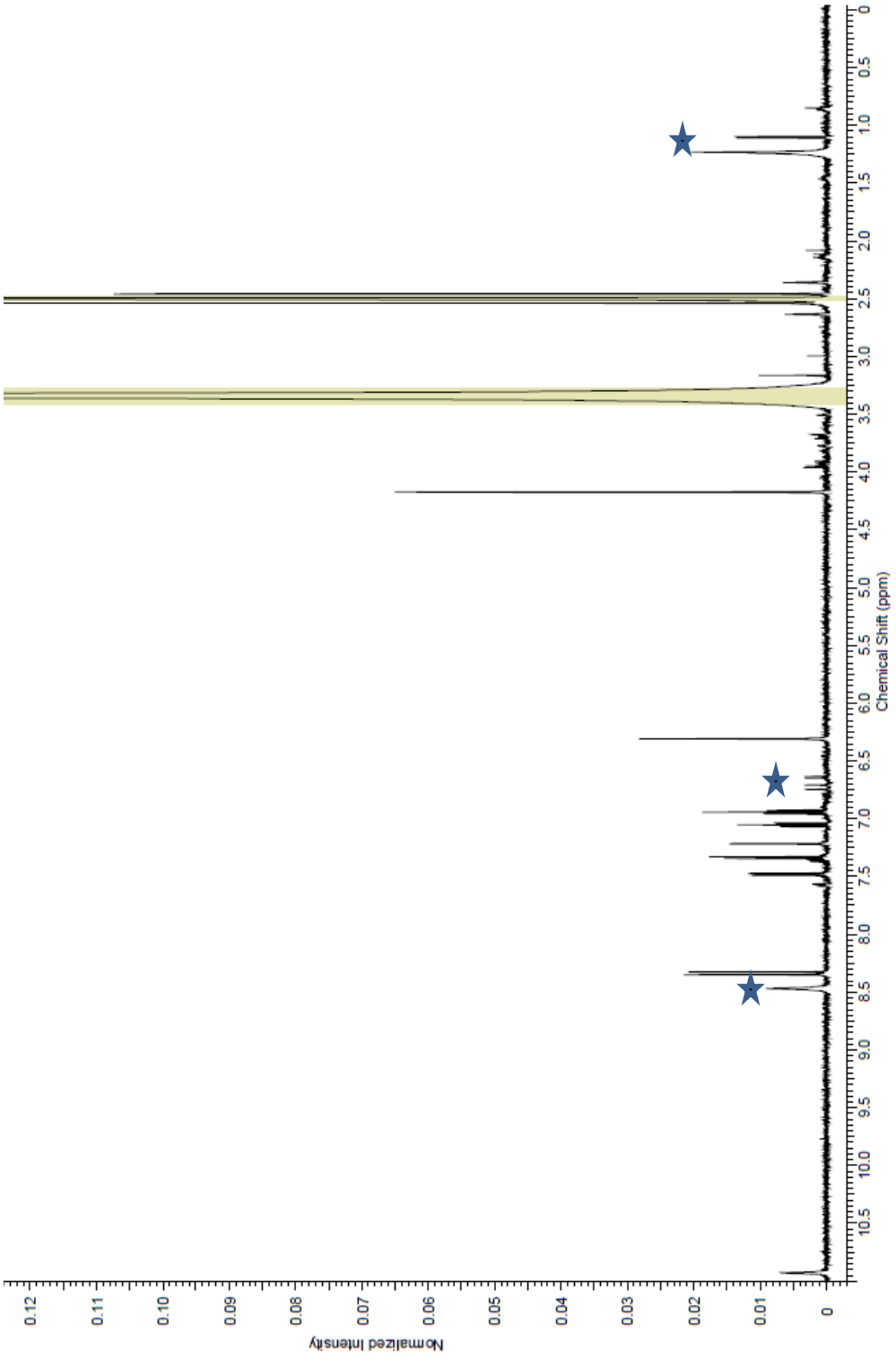


Figure A9. *In vitro* cytoxicity (upper panel) and protesome inhibition (lower panel) for **1**,**2**, and **3** in norman mouse fibroblast cells.

Figure A10. ^{1}H (500 MHz, DMSO*d*6, 25°C) spectrum of compound **8.** Unknown impurities marked with a *



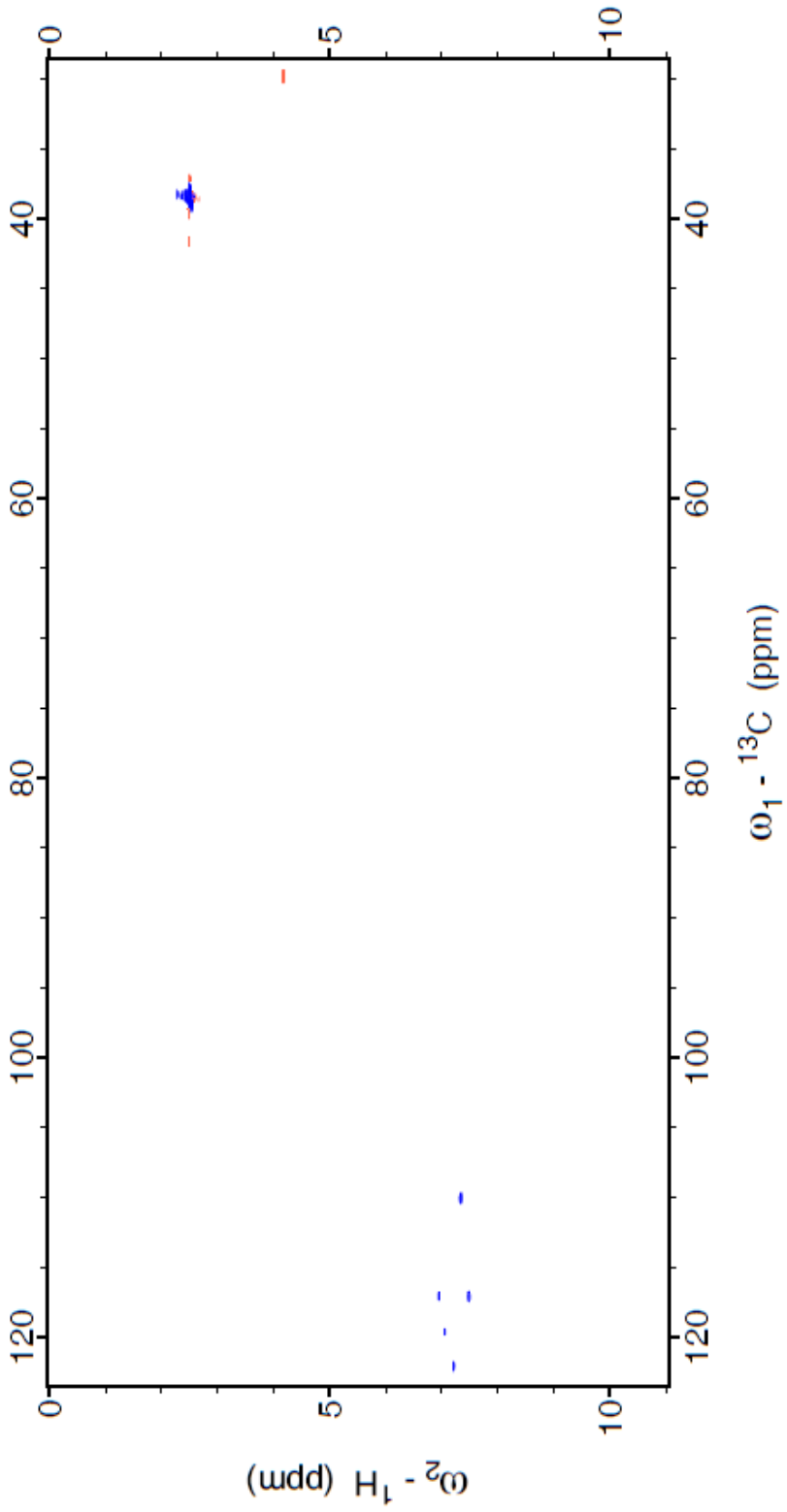


Figure A11. ^{1}H-^{13}C HSQ (500 MHz, DMSO*d6*, 25°C) spectrum of compound **8.** CH groups in blue, CH_2 in red.

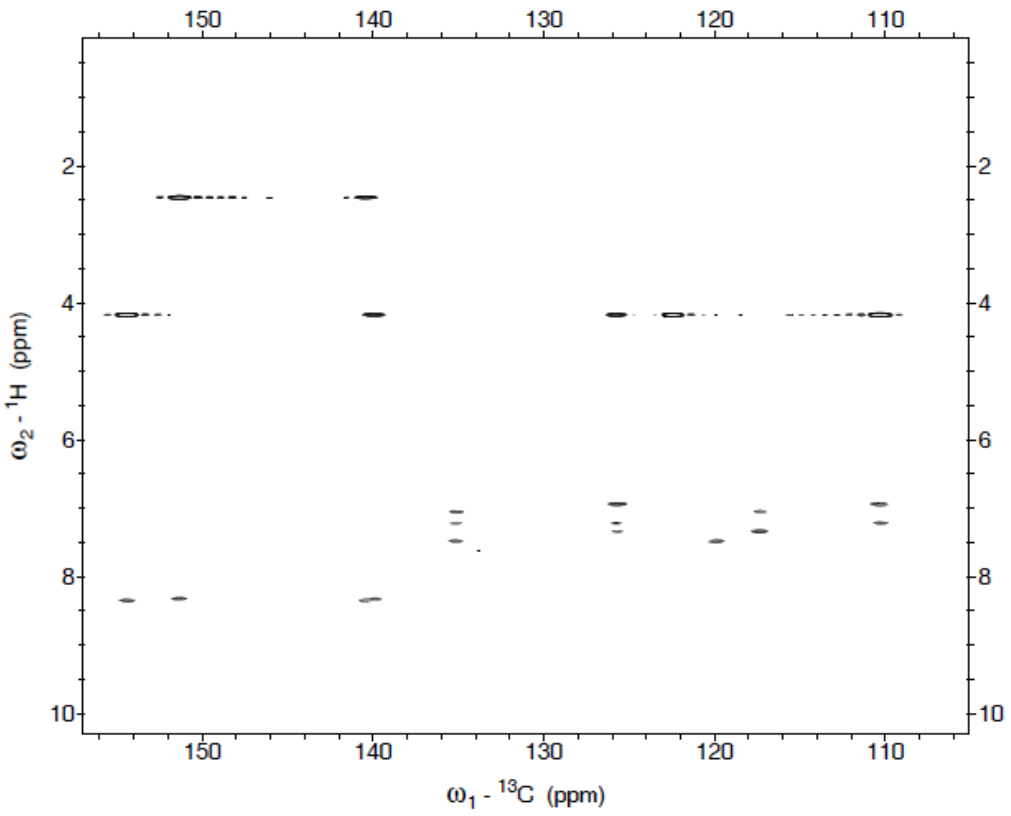
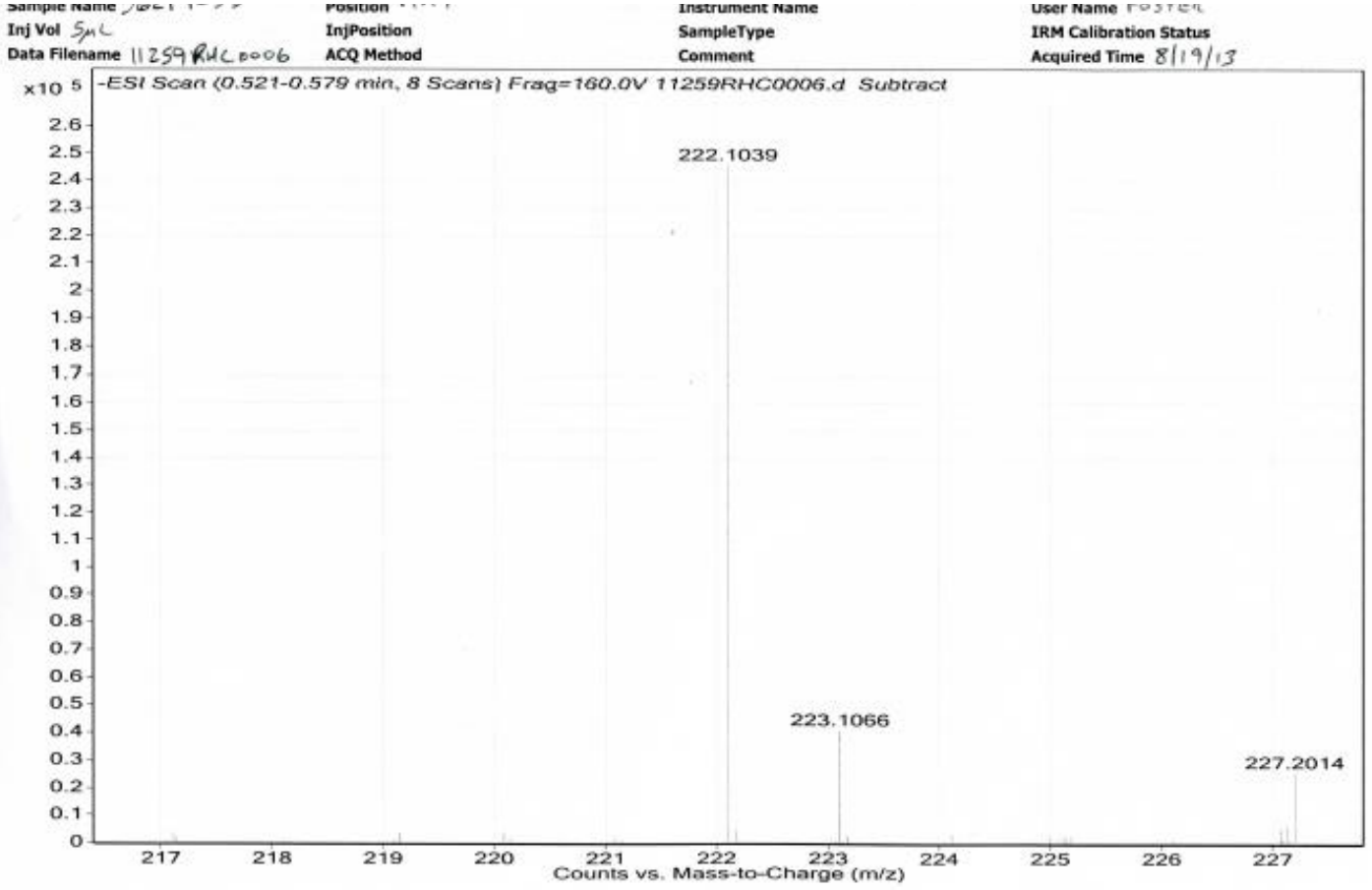


Figure A12. ^{1}H-^{13}C gHMBC (500 MHz, DMSO*d*6, 25°C) of **8**. Spectrum zoomed in to area of interest for ease of viewing.

Figure A12. HRESIMS (negative mode) for **8**



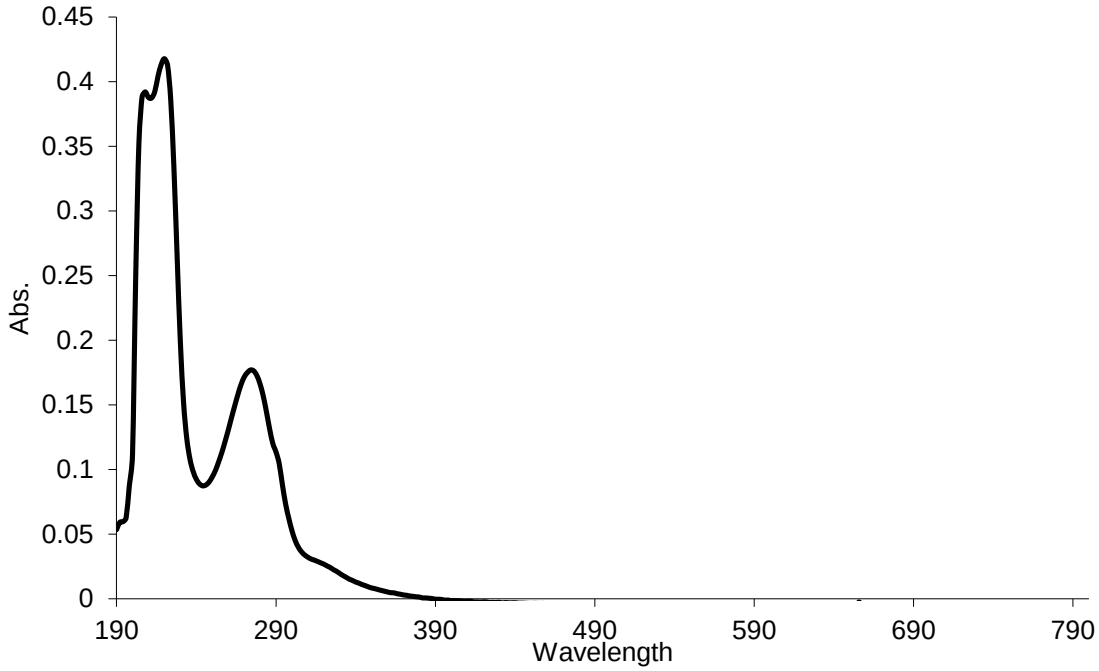
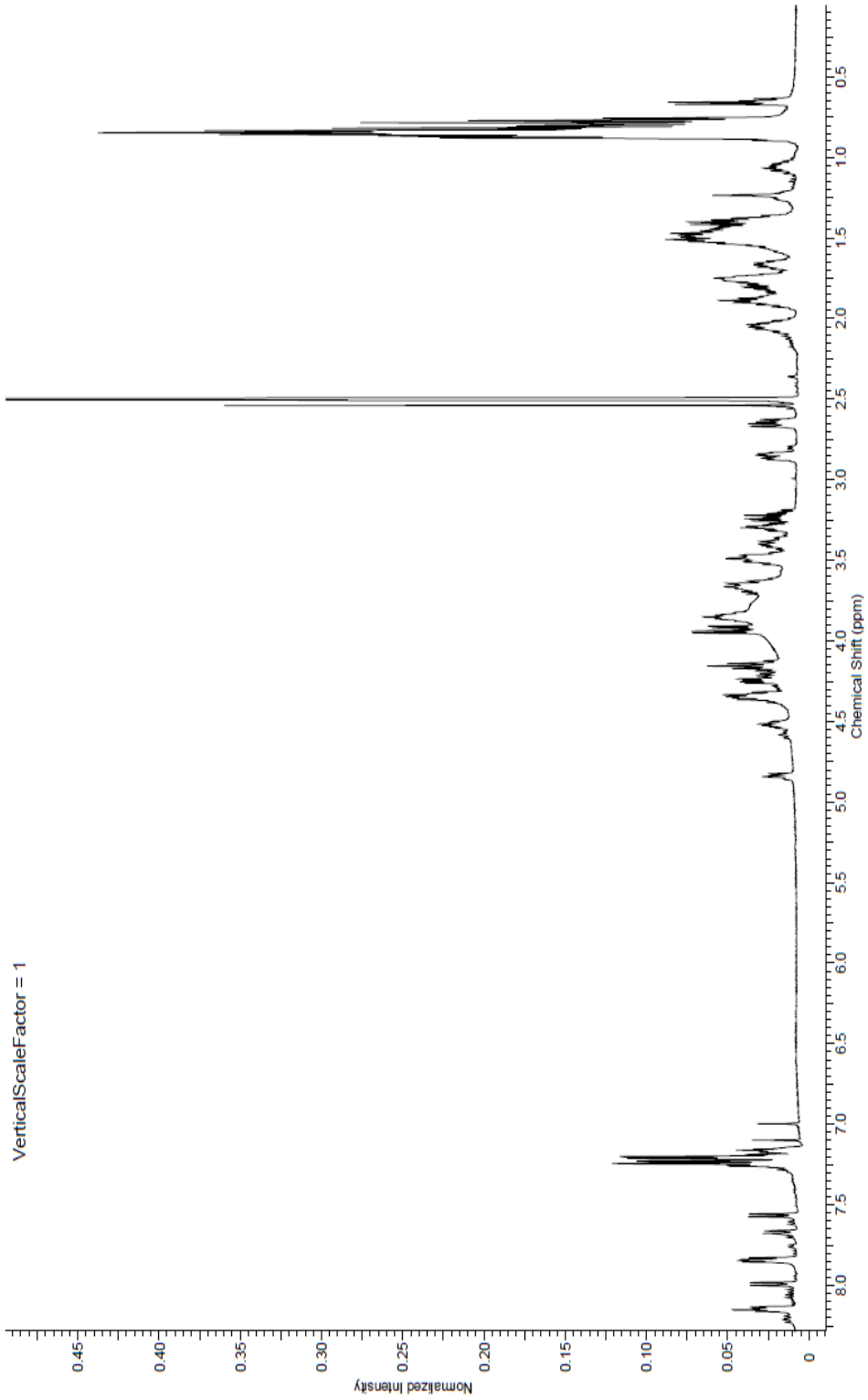


Figure A14. UV data for **8**

Figure A15. ^{1}H (500 MHz, DMSO*d6*, 25°C) spectrum of compound **9**



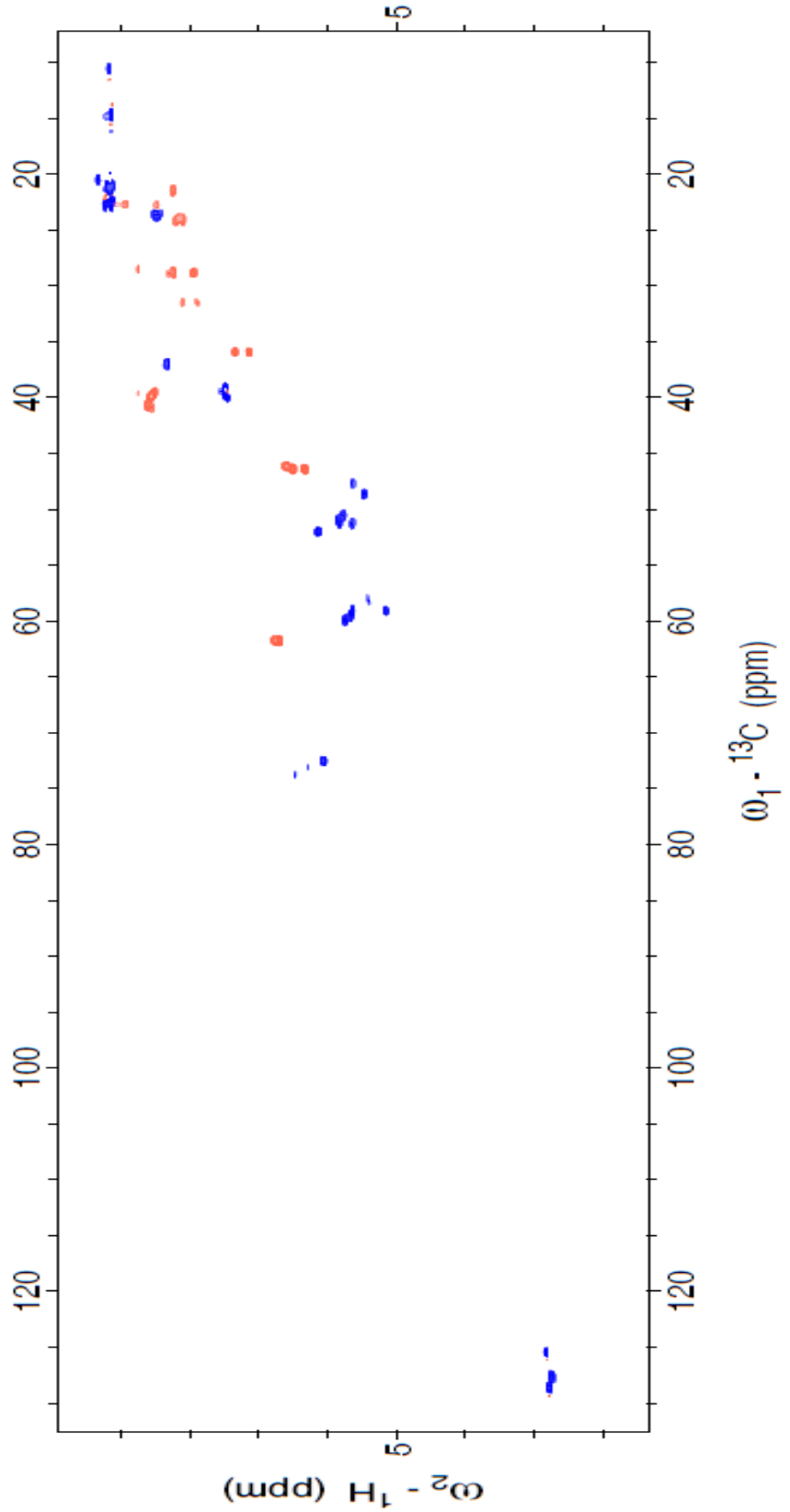


Figure A16. ^{1}H-^{13}C HSQC (500 MHz, DMSO*d*6, 25°C) spectrum of compound **9**. CH and CH_3 peaks in blue. CH_2 peaks in red.

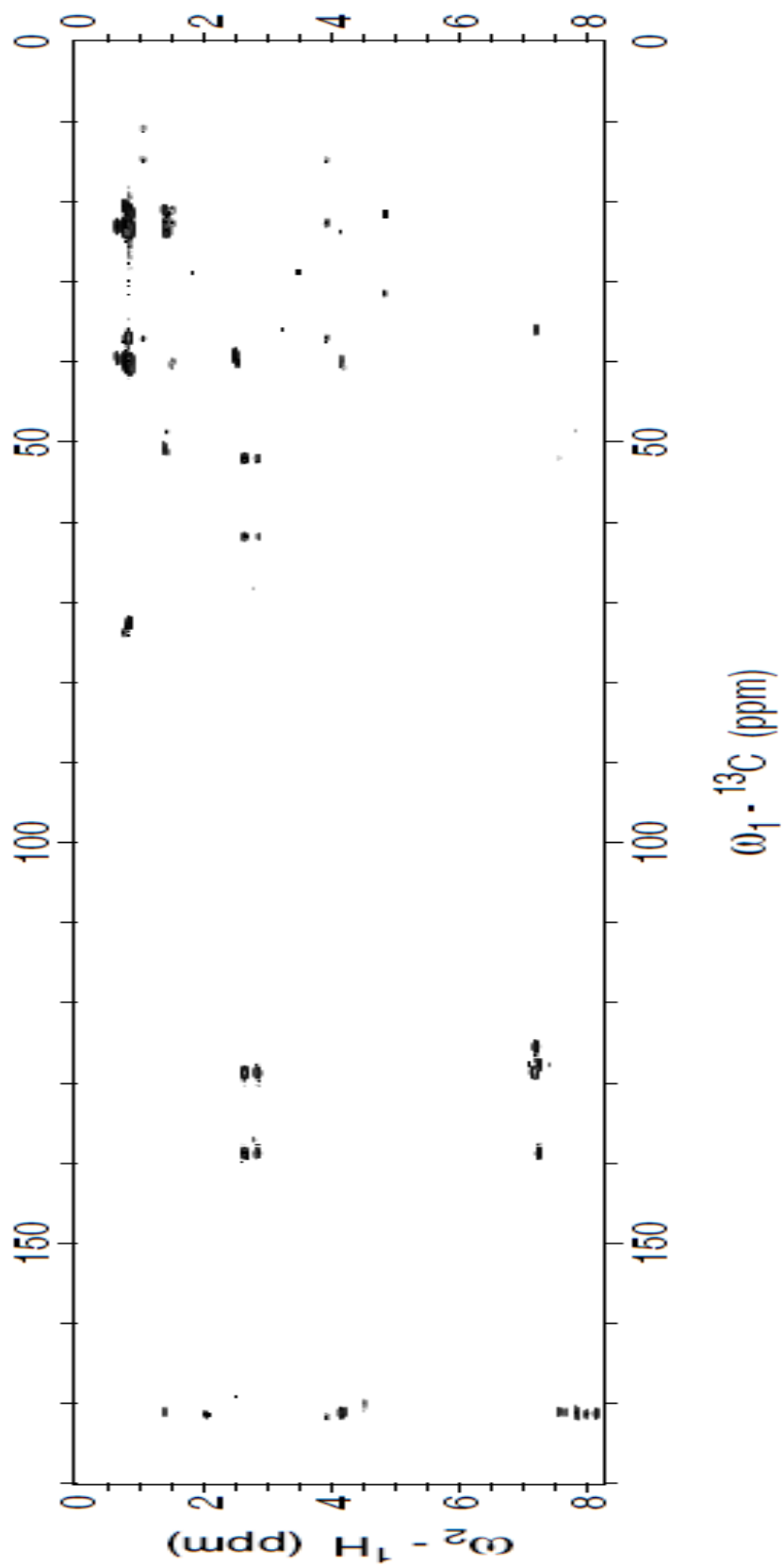


Figure A17. ^{1}H-^{13}C gHMBC (500 MHz, DMSO*d*6, 25°C) spectrum of compound **9**.

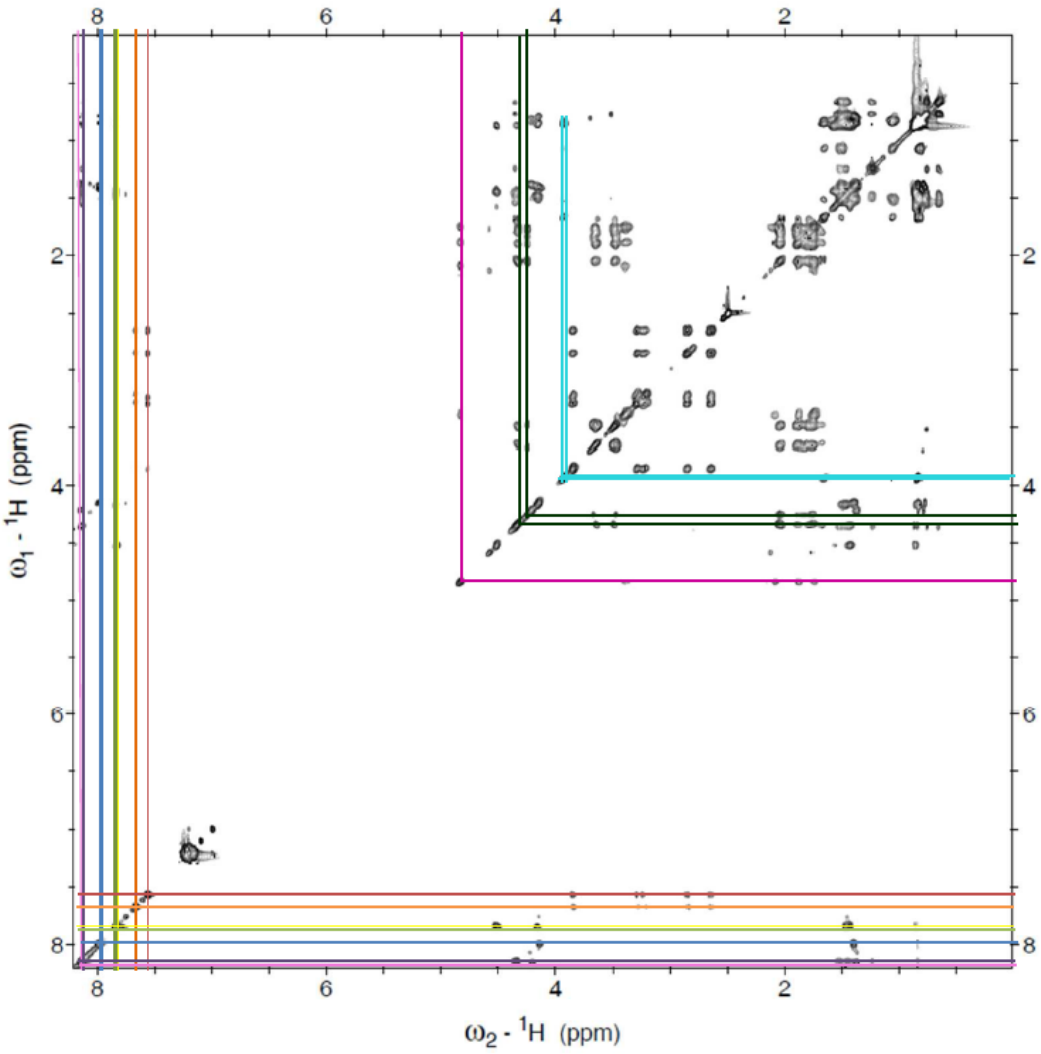


Figure A18. ^{1}H-^{1}H tCOSY (500 MHz, DMSO*d*6, 25°C) spectrum of compound **9**. Colored lines indicate spin system originating from NH groups (far left and bottom) and prolines (center).

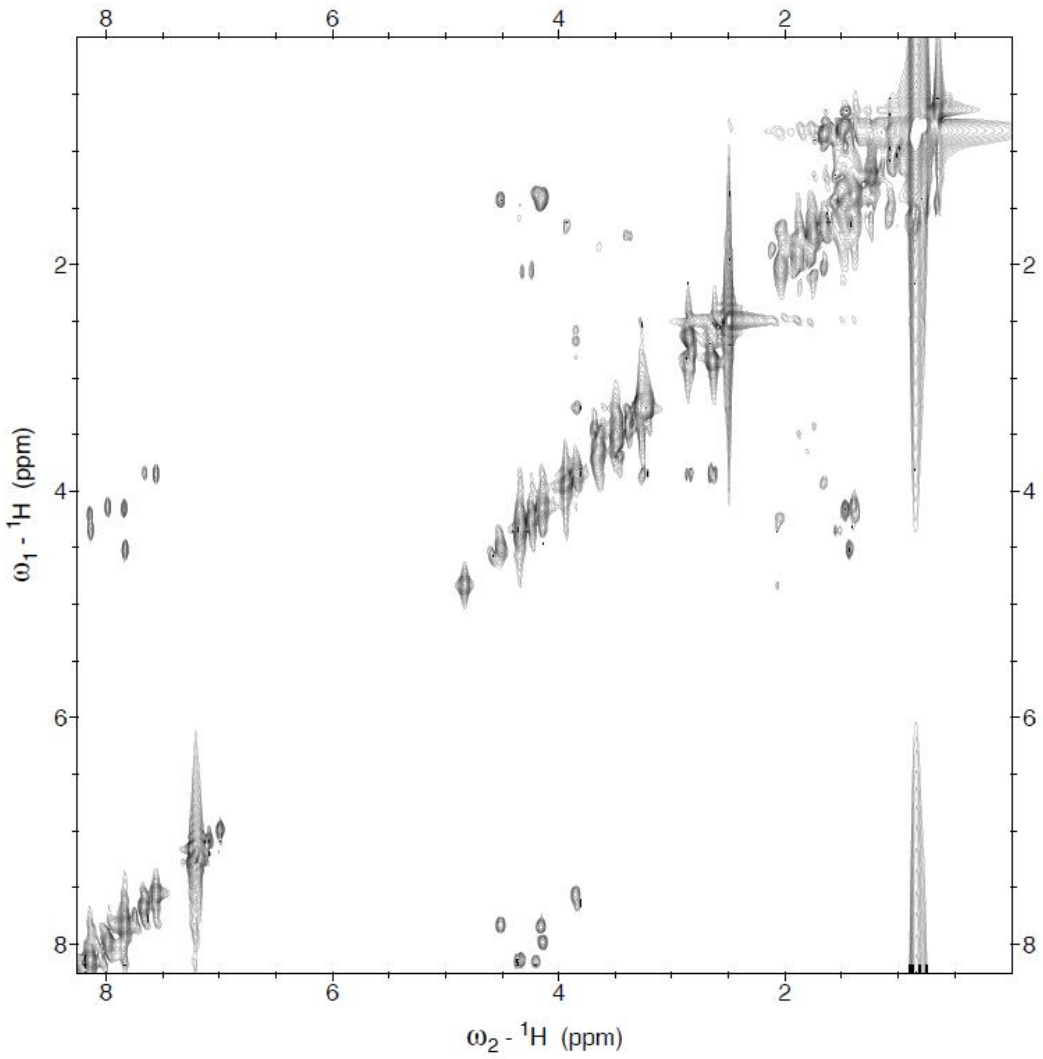


Figure A19. ^{1}H-^{1}H gCOSY (500 MHz, DMSO*d*6, 25°C) of compound **9**

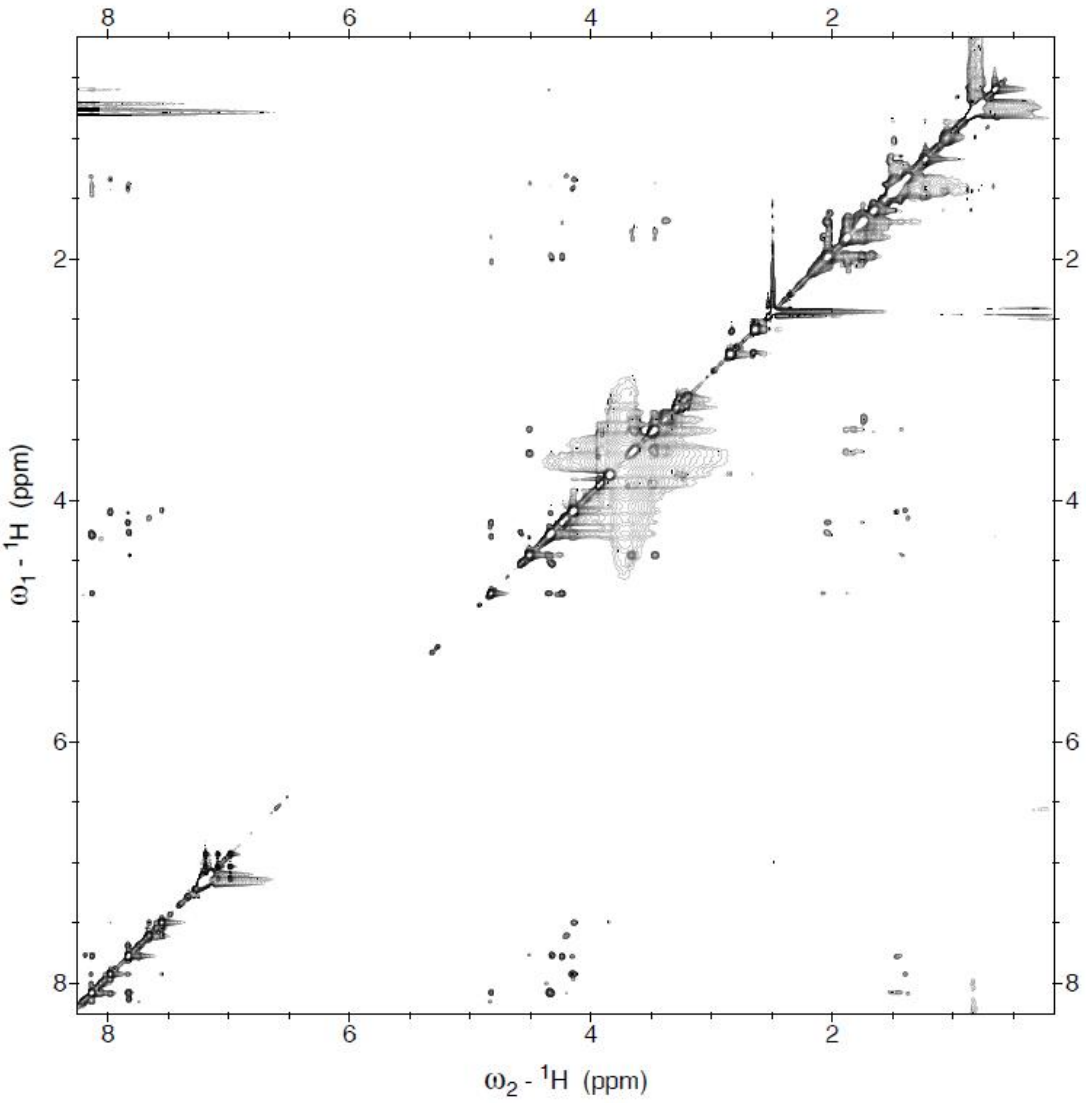


Figure A20. ^{1}H-^{1}H NOESY (500 MHz, DMSO*d*6, 25°C) of compound **9**

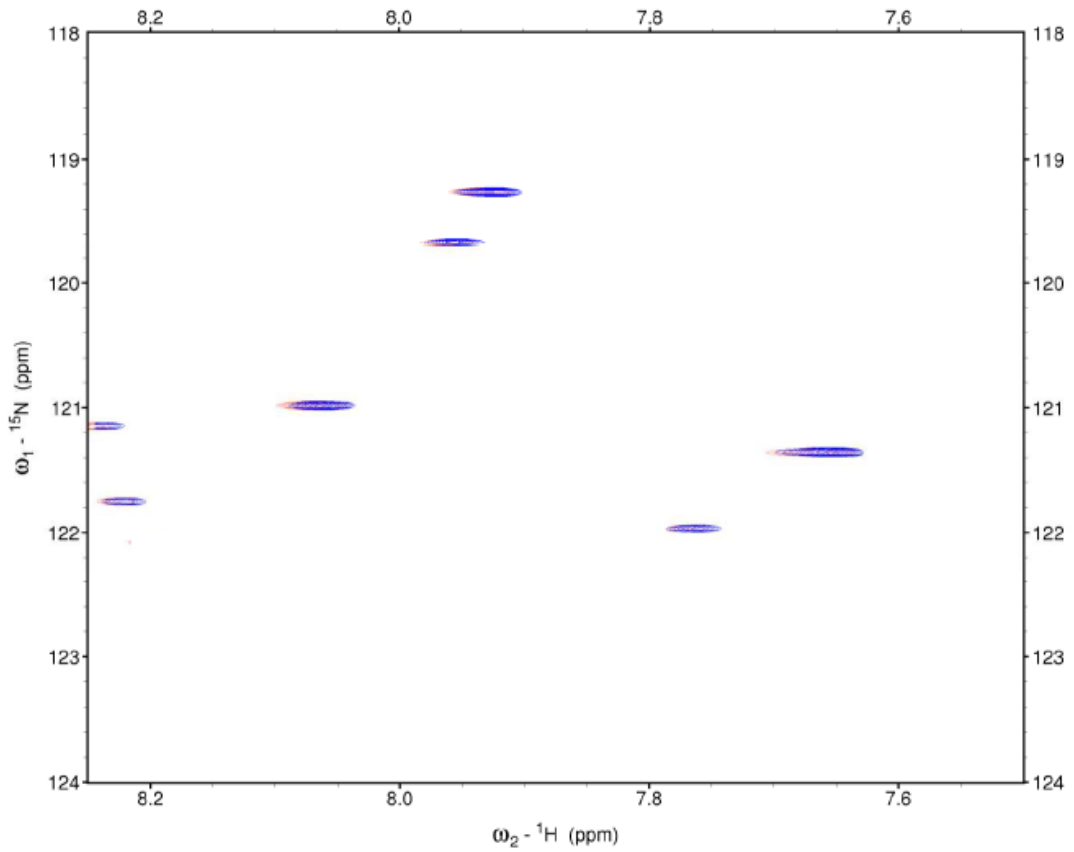


Figure A21. ^{1}H-^{15}N HSQC (500 MHz, DMSO*d*6, 25°C) of compound **9**. Figure shows overlay of two NMR experiments. The first experiment was run to show on NH signales (in red). The second showed NH and NH_2. Both experiments overlayed exactly, indicating that there were no NH_2 groups in **9**. Contaminate signal removed for ease of viewing.

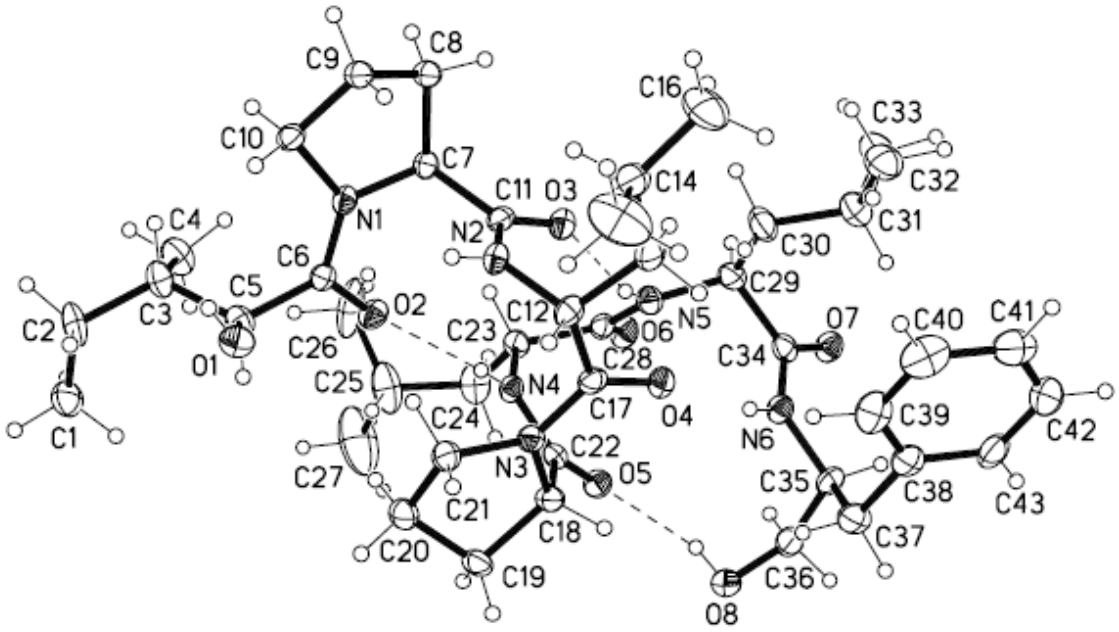


Figure A22. Thermal ellipsoid plot of **9**

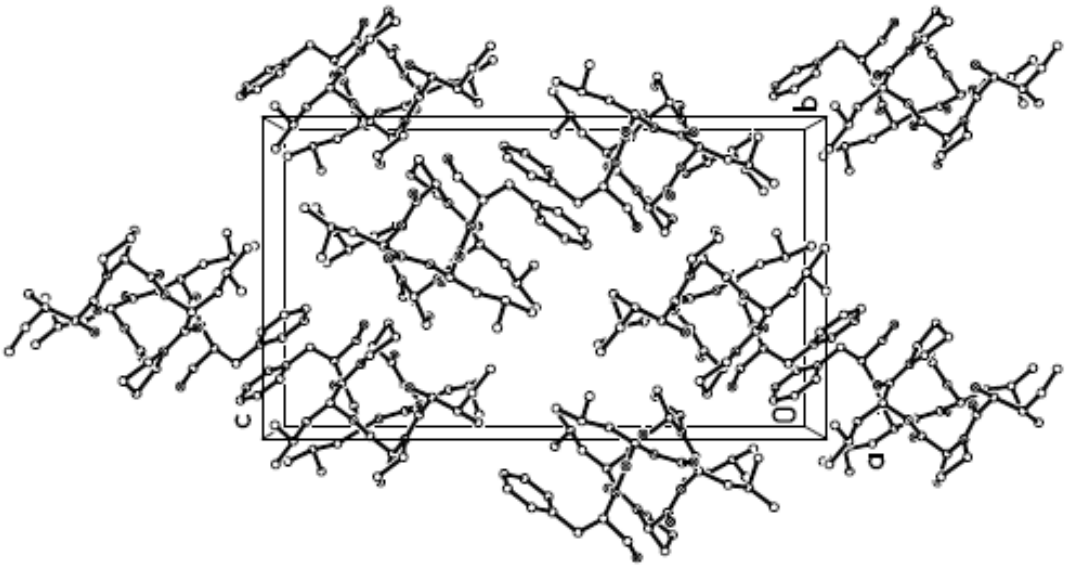


Figure A23. Packing diagram of **9**

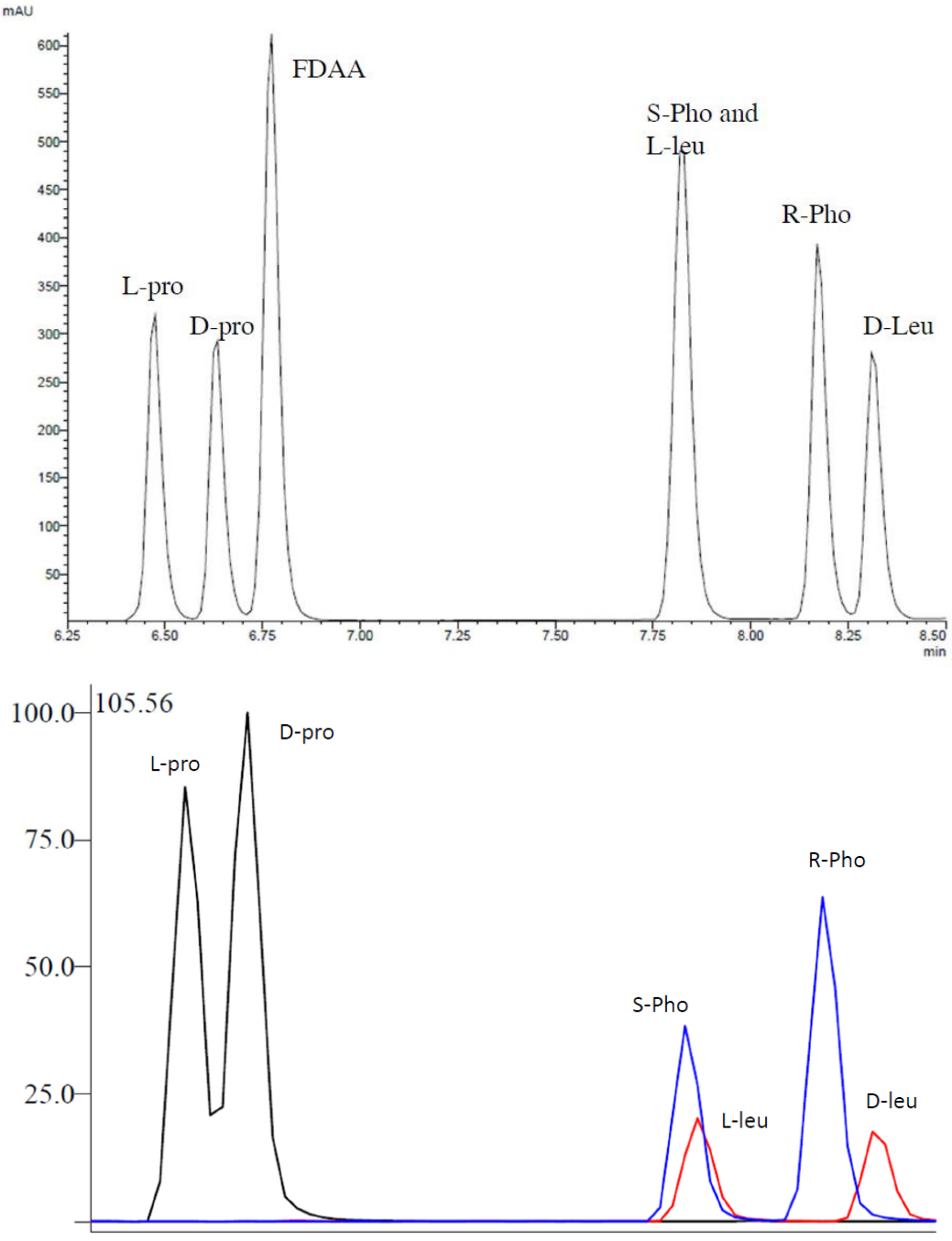


Figure A24. Chromatograms of standard $_{D}$ and $_{L}$-amino acids. PDA trace (top panel) and MS single ion trace (bottom panel) of amino acids. Mass trace shows that although they overlap in the PDA chromatogram, *S*-phenylalanine and and $_{L}$-leucine can be resolved by looking at the ion trace specific for each molecules *m/z*.

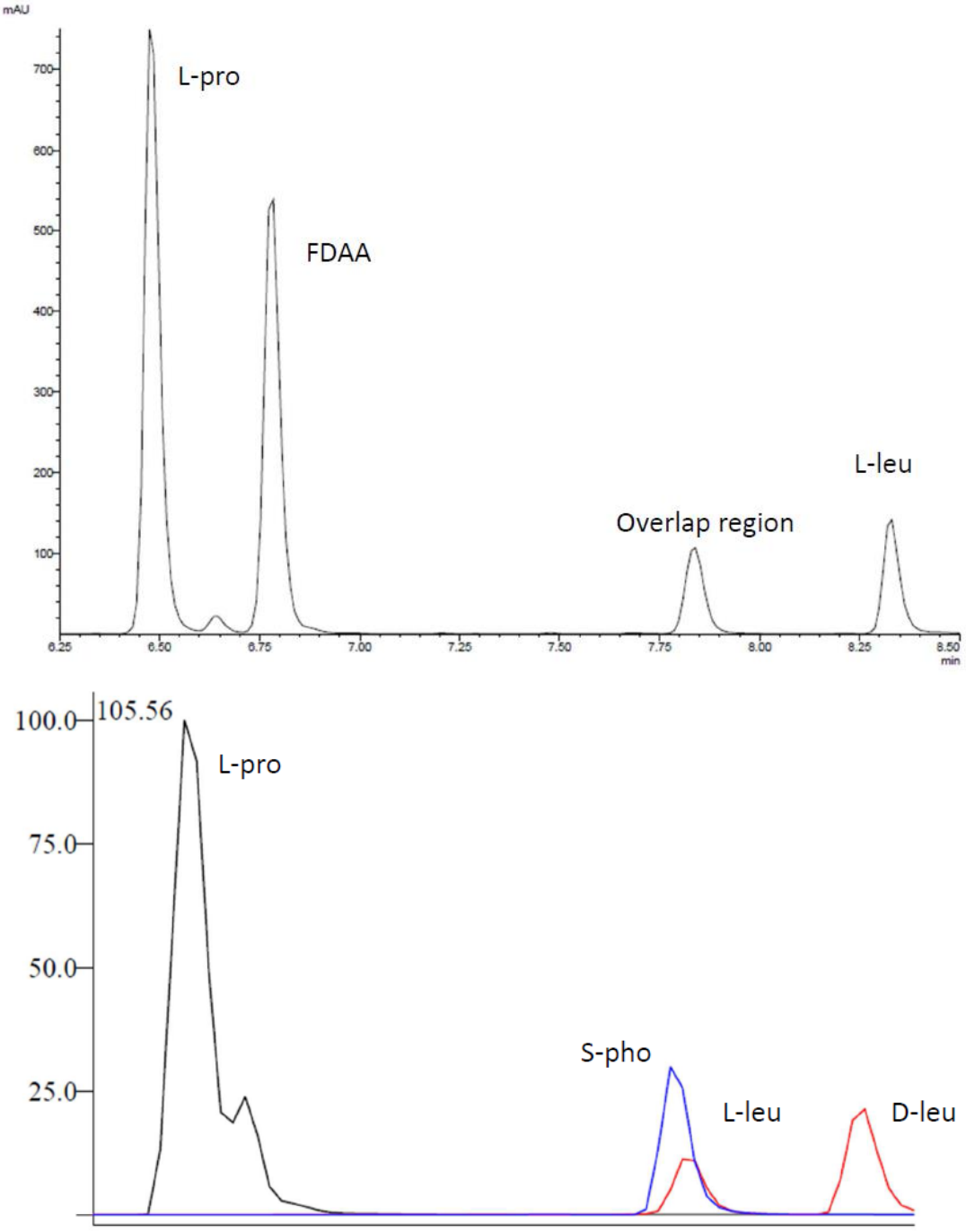
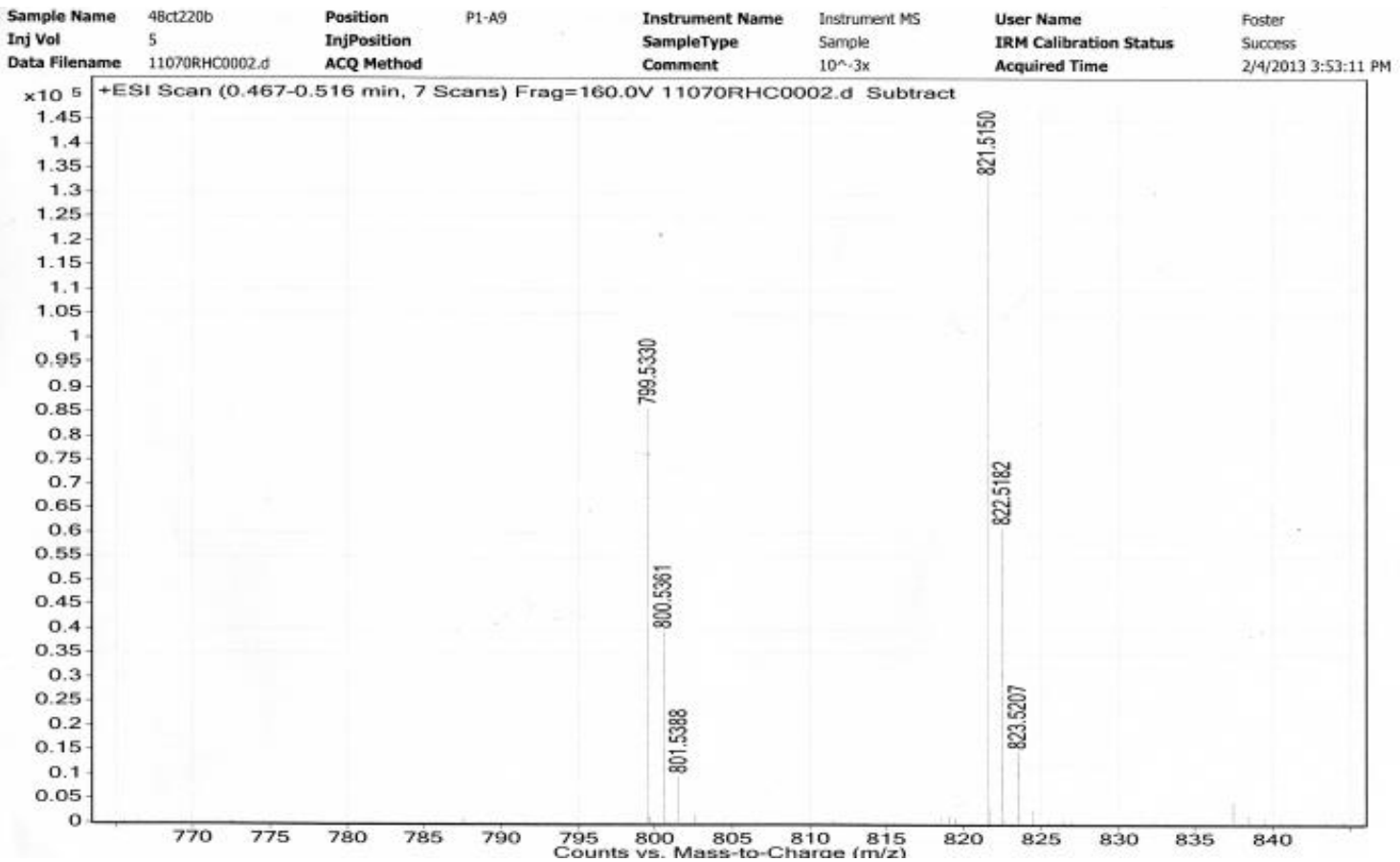


Figure A25. Marfey's analysis of **9**. PDA trace (top panel) and single ion trace (bottom panel). The Marfey's analysis confirmed the presence of *S*-phenylalanine. Which lead to the final assignment of all $_{L}$-prolines. D and L leucine exists in a roughly 1:2 ratio, which matches the assigned absolute configuration.

Figure A26. HRESIMS (positive mode) of **9**



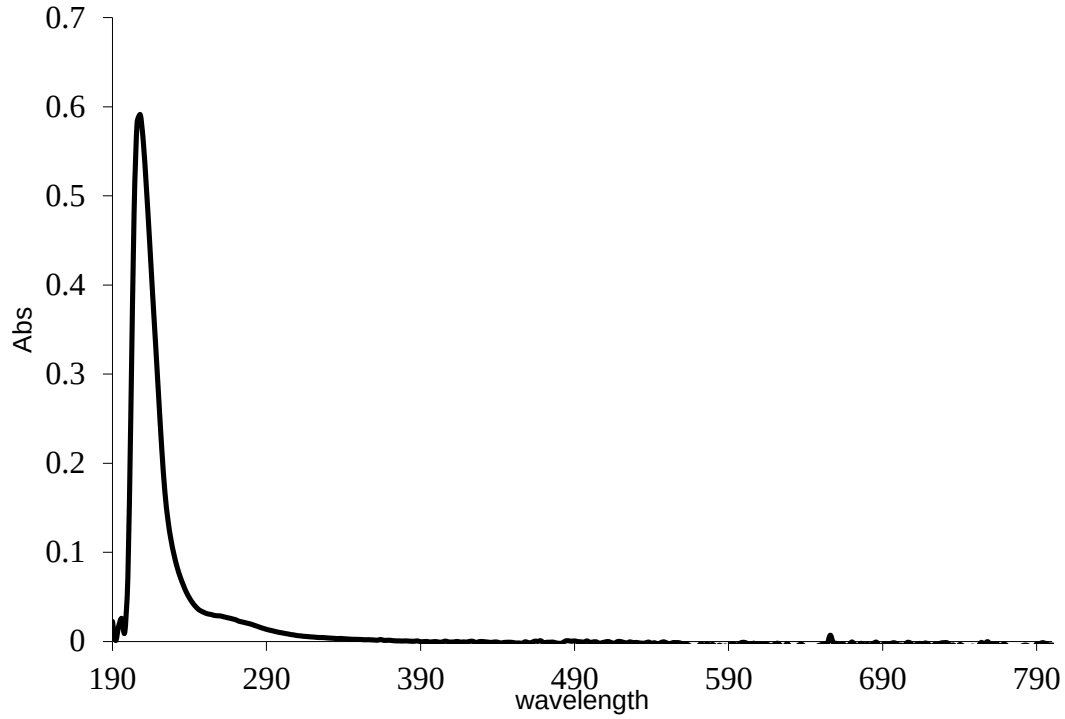


Figure A27. UV data for **9**

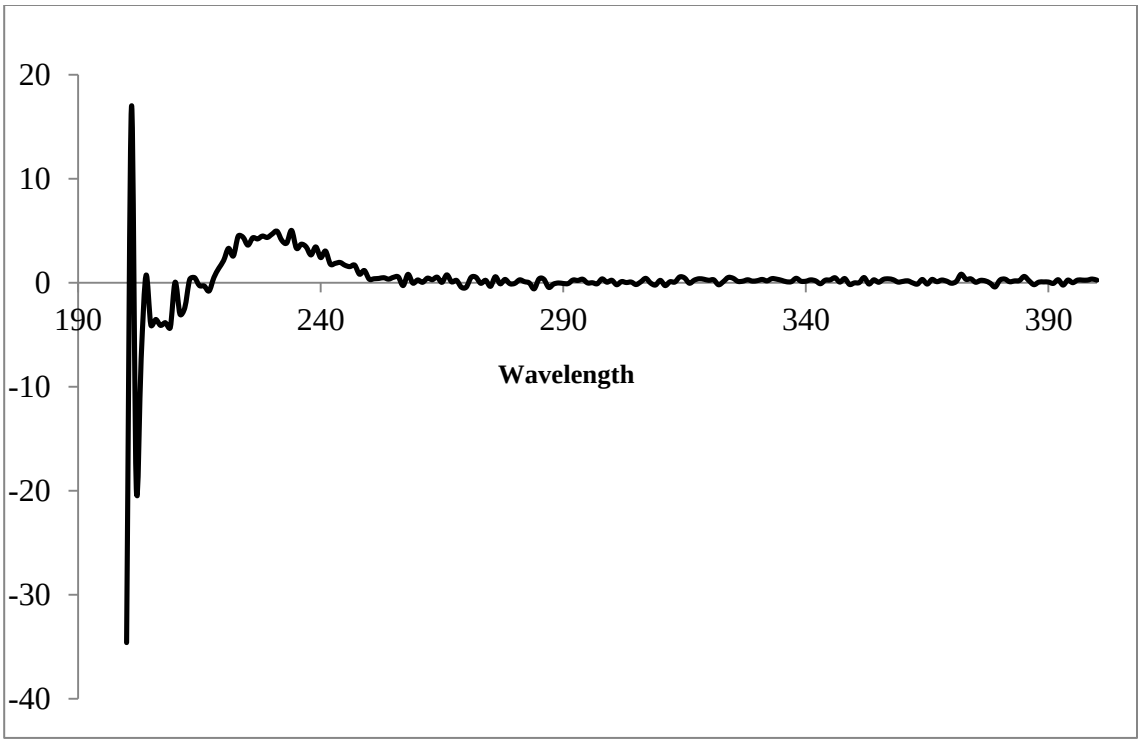


Figure A28. CD data for **9**.