AN ABSTRACT OF THE DISSERTATION OF

Chenxi Zhu for the degree of Doctor of Philosophy in Chemistry presented on November 26, 2019.

Title: Drug Discovery from Diverse Bacteria – Bioactivity-guided Isolation of Known and New Metabolites

Abstract approved: _____
Sandra Loesgen

Exploring bioactive natural products has contributed largely to clinically approved drugs we have been using over the last 100 years. Especially among the anti-infective drugs, around 70% of currently used antibiotics were discovered or derived from microbial secondary metabolites, among them compounds like vancomycin, chloramphenicol, and erythromycin. Facing the unavoidable fact of microbial drug resistance and low cure rate of cancers, exploring new drug leads is essential and urgent. Drug discovery from microbial sources has just scratched the surface: recent surveys have shown that bacterial genomes are filled with genes encoding for secondary metabolites, that have not been seen in the laboratory, and that unique environments and underexplored biodiverse niches can yield new bacterial species with unique chemotypes. Bioactivity-guided isolation with dereplication is still an efficient method used in the laboratory to discover new bioactive compounds.

This thesis includes the details on isolation of bacterial strains from diverse environments, bioactivity-guided fractionation, and dereplication/characterization of isolated metabolites. Bacterial strain library, consisting ~400 bacteria, was established in Loesgen Lab. The protocols of bacterial isolation from terrestrial and marine sources, the

workflow of methods for chemical and bioactivity screening, and dereplication methods are presented. Approximately 50% of the bacterial strains have been extracted, fractionated and tested for cytotoxicity against a colon cancer cell line. Projects were prioritized based on the chemical and bioactivity screening results. An investigation of 19 bacterial strains from Oregonian soils yielded twelve known metabolites and two new natural products, a new tetrapeptide from *Streptomyces* sp. and a new chromone from *Paraburkholderia* sp. Their absolute configuration was established via advanced Marfey's analysis and X-ray crystallography. Besides, an unusual cytotoxic diterpenoid was discovered from *Streptomyces flaveolus*, featuring five chiral centers and two double bond geometries within a fused bicyclo[8.4.0]tetradecane macrocycle. The metabolite existed in two distinct ring-flipped conformers in solution and its absolute configuration was determined by Mosher ester analysis, *J*-based coupling analysis, and DFT modeling.

Drug Discovery from Diverse Bacteria – Bioactivity-guided Isolation of Known and New Metabolites

by
Chenxi Zhu

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented November 26, 2019
Commencement June 2020

Doctor of Philosophy dissertation of Chenxi Zhu presented on November 26, 2019

APPROVED:

Major Professor, representing Chemistry

Head of the Department of Chemistry

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Chenxi Zhu, Author

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my research supervisor, Prof. Sandra Loesgen, who expertly guided me through my six years of graduate education. Her unwavering enthusiasm for chemistry deeply impacted me. Importantly, her kindness, patience, and encouragement made me go through the ups and downs of my PhD tenure. I am grateful for her never giving me up in my worst times.

My appreciation also extends to my laboratory colleagues: Dr. Donovon Adpressa, Dr. Ross Overacker, George Neuhaus, Paige Mandelare and Elizabeth Kaweesa. I want to thank them for support in my research, and especially for reviewing parts of my dissertation. I also enjoyed working with all recent and previous members of the Loesgen Lab, including Dr. Birte Plitzko, Gisela González-Montiel, Mahsa Khoshbakht, and many undergraduate students.

I am grateful to all past and present committee members for taking time to the meetings and giving me feedback, guidance, and scientific advice.

Furthermore, I also want to thank my friends, Dr. Xin Li, Dr. Jinming Li and Zhou Fang for their friendship, support, and company in Corvallis.

At last, but not least, I want to thank my parents for their never-ending support and love through life.

CONTRIBUTION OF AUTHORS

Prof. Sandra Loesgen was the major advisor on all projects and supported the design and writing of all chapters contained herein. In chapter two, Paige Mandelare assisted with the isolation of marine bacteria from Pacific fish and Arctic Ocean sediment and water samples. Elizabeth Kaweesa assisted in performing MTT-based cell viability assays. In chapter three, undergraduate student Cassandra Lew assisted in the isolation of bacteria and extract preparation. She isolated furaquinocin C and characterized its structure by NMR. Dr. Donovon Adpressa assisted the structure elucidation and determination of the absolute configuration of the new tetrapeptide from *Streptomyces resistomycificus* (CL12-5). George Neuhaus supported the structure elucidation of compound MY12-62A with computationally derived ECD spectra. Dr. Birte Plitzko and Elizabeth Kaweesa performed MTT-based cell viability assays for chapter three. In chapter four, Cassandra Lew isolated *Streptomyces flaveolus* (CL12-4) from a soil sample. Dr. Donovon Adpressa assisted with the structure elucidation of the diterpenoid by NMR analysis and computations of spectral properties, he also contributed to writing the chapter.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX FIGURES (Continued)

LIST OF APPENDIX FIGURES (Continued)

LIST OF APPENDIX FIGURES (Continued)

LIST OF APPENDIX FIGURES (Continued)

Figure Page

LIST OF APPENDIX TABLES

Chapter One: General Introduction

Chenxi Zhu

1.1 History of natural products

Natural products could be defined broadly as any molecules found in nature. Within the field of organic chemistry, natural products are described as small organic molecules (molecular weight < 1500 Da) generated by conditional metabolic pathways in bacteria, fungi, plants and animals [1]. The history of the application of natural products can be traced back over 4500 years ago in Egyptian and Sumerian written records [1]. Opium has been used as a pain reliever for thousands of years [2]. The medicinal herb Qinghao, historically recorded in medical manuscript "zhou hou bei ji fang", was utilized as a treatment of malaria as far back as 341 A.D. [3, 4]. The antipyretic effects of the willow bark were recognized for more than 200 years [5]. Starting in the 19th century, chemists began to isolate pure active compounds from bioactive plants. Morphine was first isolated from Opium between 1803 and 1805 by Friedrich Sertürner, which is generally believed to be the first isolation of an active ingredient from a plant [2]. Artemisinin was discovered from Qinghao in 1972 by Youyou Tu, who was co-awarded the 2015 Nobel Prize in Medicine [3]. Salicin, the glycoside of salicyl alcohol, was first isolated from the willow bark in 1828, and salicylic acid, was acetylated and introduced into pharmaceutical markets with the name "Aspirin" by the Bayer Company in 1898 [5].

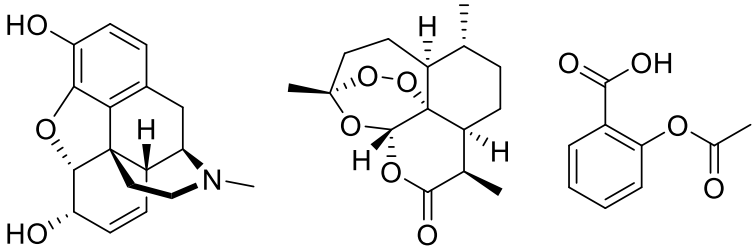


Figure 1.1 Early natural products and derivatives used as drugs to aid human suffering

The discovery of penicillin from *Penicillium notatum* by Alexander Fleming in 1928 marked a significant shift in medical treatments derived from plants to microorganisms [6]. Since then, microbes have been found to produce a variety of bioactive agents including antibiotics such as erythromycin, tetracycline, vancomycin, and chloramphenicol; antifungal agents including amphotericin B, and ieodoglucomide; anticancer drugs like doxorubicin, bleomycin, actinomycin D; or various immunosuppressive agents like rapamycin. In the modern era, over fifty percent of FDA approved drugs since 1981 have been sourced from or inspired by natural products [7, 8]. In the realm of antibacterial agents, more than 73% of the approved agents were unaltered natural products or natural product derivatives [7].

1.2 Antibiotics and anticancer agents from bacteria

Beginning with the discovery of actinomycin in 1940 by Waksman and Woodruff, a multitude of bioactive metabolites, especially antibiotics, were continuously isolated from Gram-positive bacteria of the phylum actinobacteria, leading to the 20 yearlong "golden age of natural antibiotic discovery" [9]. A review published in 2012 revealed that 45% (10,100 out of 22,500) of reported bioactive secondary metabolites were produced by actinomycetes. Of those, 7,630 were from Streptomyces and 2,479 were from rare actinomycetes [10]. One important class of natural antibiotics from *Streptomyces* are the tetracyclines, e.g. chlorotetracycline and oxytetracycline were first isolated in late 1940s as potent inhibitors of both Gram-positive and Gram-negative bacteria [11, 12]. Streptomycin, produced by *Streptomyces griseus* was the first aminoglycoside antibiotic discovered in 1944, and effective against pulmonary tuberculosis. Continuously, other aminoglycosides including kanamycin, gentamicin, sisomicin, and lividomycin were

discovered [13-15]. Aminoglycosides act by binding to the rRNA subunit of the 30S bacterial ribosome and inhibit protein synthesis [13]. Chloramphenicol, harboring a broad spectrum of activity against both Gram-positive and Gram-negative pathogens, was first discovered in 1947 and the structure was identified in 1949 [16]. Erythromycin A, a broad-spectrum antibiotic produced by *Streptomyces erythraea*, was first discovered in 1952. It is prescribed to treat a wide variety of bacterial infections, such as respiratory and gastrointestinal infections, whooping cough, syphilis, and acne, especially in patients who have adverse reactions against penicillin [17]. Vancomycin, isolated in the 1950s, is potent against Gram-positive pathogens, including methicillin-resistant *Staphylococcus aureus* (MRSA). The producer of vancomycin is *Amycolatopsis orientalis*, isolated from a soil sample collected in the jungle in Borneo [18]. After the late 1950s, in a 40-year period between 1960 and early 2000, no new classes of antibiotics were clinically approved [19]. The only 'new' drugs entities during this 40 years' time period, were close relatives of existing antibiotic scaffolds [19]. Finally in 2003, the new cyclic lipopeptide antibiotic daptomycin was marketed in the United States [20]. It first discovered from *Streptomyces reseosporus* in 1980s [21], and has rapid, concentration-dependent bactericidal activity against most clinically significant Gram-positive pathogens, including vancomycin-resistant *enterococci*, methicillin-resistant *Staphylococcus aureus*, and vancomycin-intermediate and -resistant *Staphylococcus aureus* [21, 22].

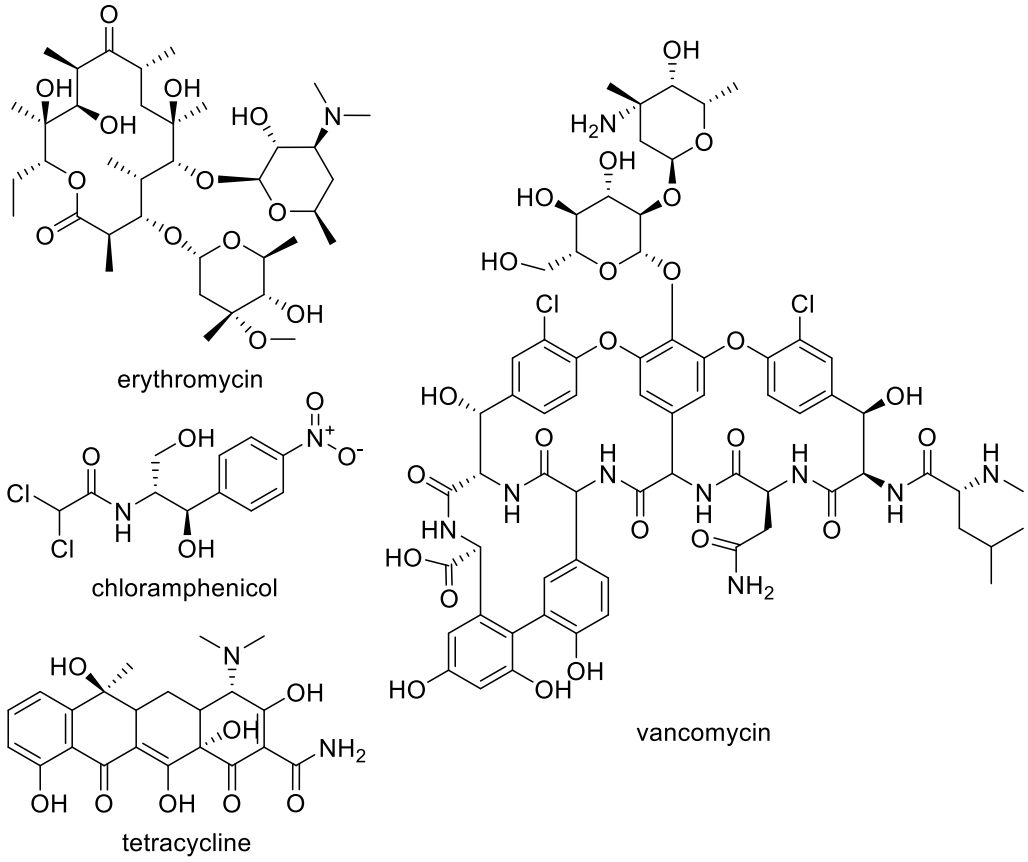


Figure 1.2 Antibiotics isolated from actinomycetes

Besides antibiotics, Gram-positive bacteria, especially *Streptomyces*, are known to produce anticancer agents that have been used in laboratories and chemotherapy. Dactinomycin, also named actinomycin D, was first isolated from *Streptomyces parvullus* in 1940 [23]. As an inhibitor of DNA-dependent RNA synthesis [23, 24], it was approved for medical use in the United States in 1964 [23]. Mitomycin was discovered in the 1950s by Japanese scientists in cultures of the microorganism *Streptomyces caespitosus* [25]. It was used to treat upper gastro-intestinal cancers (e.g. esophageal carcinoma), anal cancers, and breast cancers, as well as for superficial bladder tumors. Bleomycin, discovered in 1962 in *Streptomyces verticillus* [26], is used to treat Hodgkin's lymphoma, non-Hodgkin's lymphoma, testicular cancer, ovarian cancer, and cervical cancer [27].

Bleomycin causes DNA strand scission through formation of an intermediate metal complex requiring a metal ion cofactor such as copper or iron. This action results in inhibition of DNA synthesis, and to a lesser degree, in inhibition of RNA and protein synthesis [28]. Doxorubicin, produced by *Streptomyces peucetius* [29], was approved for medical use in the United States in 1974. It is used to treat breast cancer, bladder cancer, Kaposi's sarcoma, lymphoma, and acute lymphocytic leukemia [30]. The anthracycline structure of doxorubicin leads the intercalation with DNA, and inhibits the progression of the enzyme topoisomerase II, breaking the replication of DNA chain [31].

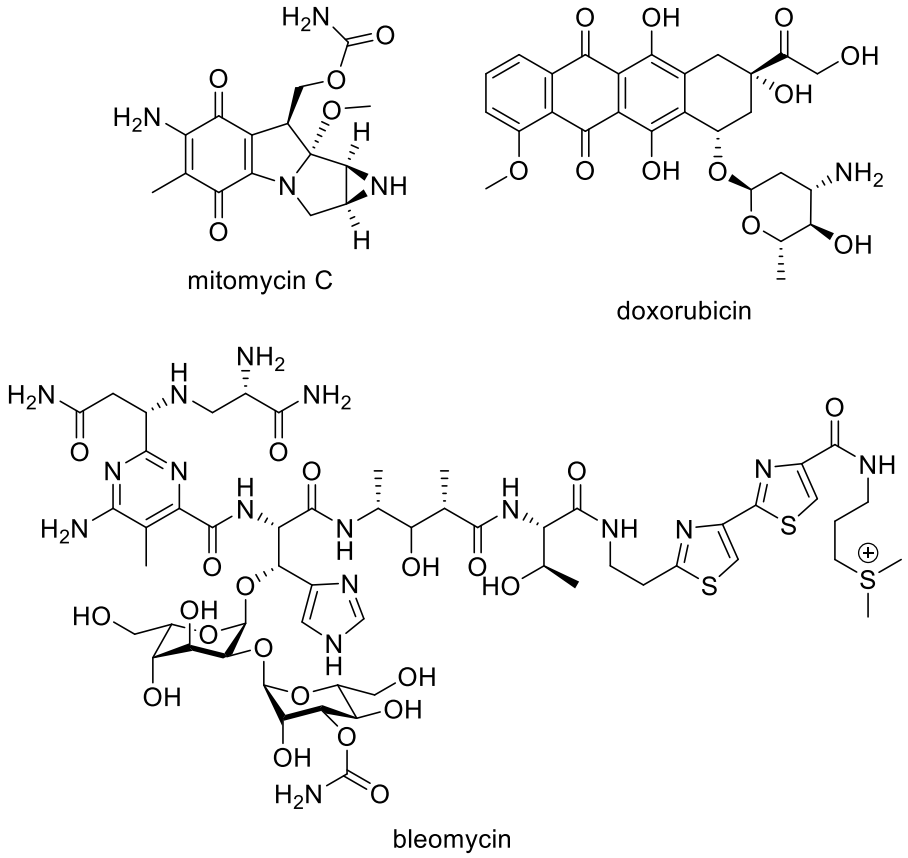


Figure 1.3 Anticancer agents from actinomycetes

1.3 New drugs are needed

Although current antibiotics are still performing well in some applications, by 2050, nearly 10 million death are predicted to be attributed to antimicrobial resistance every year [32]. Antibiotic resistance is a global crisis [33], and at our current rate of novel compound discovery we are approaching "a post-antibiotic era" [34]. Nonetheless, many of the large pharmaceutical companies closed their discovery research efforts in the antibacterial therapeutics field, during the 1990s and early 2000s due to cost constraints [19]. As a result, over a four decade span between 1960 and 2000, no new classes of antibiotics were approved [7]. Bacterial antibiotic resistance is a natural occurance and it is not a matter of if, but only of when [19]. From January 2011 to December 2014, nine small-molecule antibiotics have been approved [7], including the semisynthetic cephalosporin "ceftraoline" on October 29th , 2010, with activity against methicillin-resistant *Staphylococcus aureus*. Already in 2011, the first ceftaroline-resistant *Staphylococcus* was identified 2011 [35]. Because of unavoidable microbial resistance, the need for new drug development is constant and urgent.

Novel drugs with broad bioactivity can be found from bacterial strains isolated from unexplored environments. A chemical-biogeographic survey by Charlop-Powers concluded that arid soils showed the richest diversity of biosynthetic potential [36]. Furthermore, activating silent biosynthesis gene clusters of actinomycetes can unearth novel chemistry from previous not expressed gene clusters. The typical actinomycete genomes contain approximately 20 or more gene clusters responsible for the production of secondary metabolites, however, only a small number of them are expressed under standard laboratory conditions [37]. Co-culture techniques, ribosome engineering, and

elicitor screens [38] and heterologous expression in prioritized host organisms are reported approaches successful in activating otherwise silent biosynthetic gene clusters to obtained new structures [39, 40]. Additionally, prospecting underexplored resources like marine bacteria and complex microbiomes remains promising. A recent review from Anthony Carroll reported 242 new natural products discovered from marine microbial source in 2017, in which 137 (over 50%) are from marine actinomycete genus *Streptomyces* [41]. With advanced techniques and exploration to unknown areas, chemical diversity of microbial natural products will lead more new drugs with antibiotic, anticancer effects, and continue to supply and inspire clinic treatments.

1.4 Chapter overview

My thesis work is focused on the discovery of new natural products from diverse bacteria. My research included the isolation of bacteria from various sources to create a bacterial strain library, extracting metabolites from growing cultures, isolating new/bioactive natural products with chemical screening or bioactivity-guided fractionation, and determining the structures of new natural products.

Chapter two is an overview on the creation of a bacterial strain library. I detailed procedures of bacterial isolation from soils, sediments, water samples, and fish mucus. In total 230 bacterial strains were isolated from Oregonian soils, the Arctic Ocean, and Pacific fish. I created a workflow for the chemical and bioactivity screening of microbial extracts. Thirty out of 194 strains exhibited cytotoxic activity and five known compounds were dereplicated by UV spectra and low-resolution MS analysis. Moreover, my work set the stage to prioritize and identify projects on new and bioactive compounds described in

chapters 3 and 4 and noteworthy, several other bacteria identified by me are still actively explored in the Loesgen Lab.

Chapter three is an investigation on a group of nineteen bacteria isolated from an arid, high desert area in Oregon. Taxonomic analysis of these strains and LCMS-based PCA of the organic extracts revealed the biological and chemical diversity. Extract bioactivities were tested against five human cancer cell lines and four human bacterial pathogens. In total, 14 antibacterial and/or cytotoxic compounds were isolated via bioactivity-guided fractionation, including two new natural products - a new tetrapeptide (**3.2**) and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**).

Chapter four presents the isolation and structure elucidation of a new cytotoxic diterpenoid (**4.3**) from Stre*ptomyces flaveolus* (CL12-4). The structures of the diterpenoid contains five chiral centers and two double bond geometries within a fused bicyclo[8.4.0] tetradecane macrocycle. Interestingly the terpene consists of a mixture of ring-flipped conformers in solution. Multiple analytical techniques were used to determine its absolute configuration, including 1D and 2D NMR spectroscopic analysis, Mosher ester analysis, and DFT-based computational techniques and the absolute configuration was assuredly assigned as 1R,4S,4aR,5E,7S,9Z,12aR.

The final chapter presents a brief conclusion of the works presented in the thesis, summarizing the methods applied to discover new natural products from bacteria isolated from diverse sources.

1.5 References

1. Walsh, C.T., Yi,, *Natural product biosynthesis (1st).* Royal Society of Chemistry, Cambridge, 2017.
2. Norn, S., P.R. Kruse, and E. Kruse, *[History of opium poppy and morphine]*. Dansk medicinhistorisk arbog, 2005. **33**: p. 171-184.
3. White, N.J., T.T. Hien, and F.H. Nosten, *A Brief History of Qinghaosu.* Trends in parasitology, 2015. **31**(12): p. 607-610.
4. Klayman, D., *Qinghaosu (artemisinin): an antimalarial drug from China.* Science, 1985. **228**(4703): p. 1049-1055.
5. Awtry, E.H. and J. Loscalzo, *Aspirin.* Circulation, 2000. **101**(10): p. 1206-1218.
6. Pham, J.V., et al., *A Review of the Microbial Production of Bioactive Natural Products and Biologics.* Frontiers in Microbiology, 2019. **10**(1404).
7. Newman, D.J. and G.M. Cragg, *Natural Products as Sources of New Drugs from 1981 to 2014.* Journal of Natural Products, 2016. **79**(3): p. 629-661.
8. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the 30 years from 1981 to 2010.* J Nat Prod, 2012. **75**(3): p. 311-35.
9. Gould, K., *Antibiotics: from prehistory to the present day.* Journal of Antimicrobial Chemotherapy, 2016. **71**(3): p. 572-575.
10. Mahajan, G.B. and L. Balachandran, *Antibacterial agents from actinomycetes - a review.* Front Biosci (Elite Ed), 2012. **4**: p. 240-53.
11. Chopra, I. and M. Roberts, *Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance.* Microbiol Mol Biol Rev, 2001. **65**(2): p. 232-60 ; second page, table of contents.
12. Chopra, I. and M. Roberts, *Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance.* Microbiology and molecular biology reviews : MMBR, 2001. **65**(2): p. 232-260.
13. Wright, G.D., A.M. Berghuis, and S. Mobashery, *Aminoglycoside Antibiotics*, in *Resolving the Antibiotic Paradox: Progress in Understanding Drug Resistance and Development of New Antibiotics*, B.P. Rosen and S. Mobashery, Editors. 1998, Springer US: Boston, MA. p. 27-69.
14. Becker, B. and M.A. Cooper, *Aminoglycoside antibiotics in the 21st century.* ACS Chem Biol, 2013. **8**(1): p. 105-15.
15. Takahashi, Y. and M. Igarashi, *Destination of aminoglycoside antibiotics in the 'post-antibiotic era'.* J Antibiot (Tokyo), 2017.
16. Sills, M.R. and D. Boenning, *Chloramphenicol.* Pediatrics in Review, 1999. **20**(10): p. 357-358.
17. Cobb, R.E., et al., *Chapter 10 - Drug Discovery and Development via Synthetic Biology*, in *Synthetic Biology*, H. Zhao, Editor. 2013, Academic Press: Boston. p. 183-206.
18. Griffith, R.S., *Introduction to Vancomycin.* Reviews of Infectious Diseases, 1981. **3**(Supplement_2): p. S200-S204.
19. McCarter, Y.S., *Antibiotics: Challenges, Mechanisms, Opportunities Written by Christopher Walsh, PhD and Timothy Wencewicz, PhD.* Laboratory Medicine, 2017. **48**(3): p. e42-e42.

20. Weis, F., A. Beiras-Fernandez, and G. Schelling, *Daptomycin, a lipopeptide antibiotic in clinical practice.* Current opinion in investigational drugs (London, England : 2000), 2008. **9**(8): p. 879-884.
21. Lewis, K., *Platforms for antibiotic discovery.* Nature Reviews Drug Discovery, 2013. **12**: p. 371.
22. Eisenstein, B.I., *Lipopeptides, focusing on daptomycin, for the treatment of Gram-positive infections.* Expert Opinion on Investigational Drugs, 2004. **13**(9): p. 1159-1169.
23. Connor, H., *Drug Discovery—A History.* Journal of the Royal Society of Medicine, 2005. **98**(11): p. 517-518.
24. Sobell, H.M., *Actinomycin and DNA transcription.* Proceedings of the National Academy of Sciences, 1985. **82**(16): p. 5328-5331.
25. Tomasz, M., *Mitomycin C: small, fast and deadly (but very selective).* Chemistry & Biology, 1995. **2**(9): p. 575-579.
26. Umezawa, H., et al., *New antibiotics, bleomycin A and B.* J Antibiot (Tokyo), 1966. **19**(5): p. 200-9.
27. Soloway, M.S., *Overview of treatment of superficial bladder cancer.* Urology, 1985. **26**(4 Suppl): p. 18-26.
28. Perry, M.C., D.C. Doll, and C.E. Freter, *Chemotherapy source book.* 2012.
29. Di Marco, A., M. Gaetani, and B. Scarpinato, *Adriamycin (NSC-123,127): a new antibiotic with antitumor activity.* Cancer Chemother Rep, 1969. **53**(1): p. 33-7.
30. Chatterjee, K., et al., *Doxorubicin Cardiomyopathy.* Cardiology, 2010. **115**(2): p. 155-162.
31. Gewirtz, D., *A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin.* Biochemical Pharmacology, 1999. **57**(7): p. 727-741.
32. O'Neill, J., *Tackling drug-resistant infections globally: final report and recommendations.* The review on antimicrobial resistance 2016.
33. Aslam, B., et al., *Antibiotic resistance: a rundown of a global crisis.* Infection and drug resistance, 2018. **11**: p. 1645-1658.
34. Bragg, R.R., et al. *Potential Treatment Options in a Post-antibiotic Era*. in *Infectious Diseases and Nanomedicine III*. 2018. Singapore: Springer Singapore.
35. Ventola, C.L., *The antibiotic resistance crisis: part 1: causes and threats.* P & T : a peer-reviewed journal for formulary management, 2015. **40**(4): p. 277-283.
36. Charlop-Powers, Z., et al., *Chemical-biogeographic survey of secondary metabolism in soil.* Proc Natl Acad Sci U S A, 2014. **111**(10): p. 3757-62.
37. Scherlach, K. and C. Hertweck, *Triggering cryptic natural product biosynthesis in microorganisms.* Organic & Biomolecular Chemistry, 2009. **7**(9): p. 1753-1760.
38. Mao, D., et al., *Recent advances in activating silent biosynthetic gene clusters in bacteria.* Current opinion in microbiology, 2018. **45**: p. 156-163.
39. Ochi, K., *Insights into microbial cryptic gene activation and strain improvement: principle, application and technical aspects.* J Antibiot (Tokyo), 2017. **70**(1): p. 25-40.

40. Ochi, K. and T. Hosaka, *New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters.* Appl Microbiol Biotechnol, 2013. **97**(1): p. 87-98.
41. Carroll, A.R., et al., *Marine natural products.* Natural Product Reports, 2019. **36**(1): p. 122-173.

Chapter Two: Creation of a bacterial strain library for drug discovery

Chenxi Zhu, Paige E. Mandelare, Sandra Loesgen

2.1 Abstract

In this chapter, an overview is presented on my efforts to create a bacterial strain library for the Loesgen lab. Since 2014, I was involved with isolating bacteria from various sources, including terrestrial and marine samples, from various collection trips and collaborative projects. In total, 230 bacterial strains were isolated, purified to single colonies, and cryopreserved. Isolates came from soils, sediments, water samples, or were derived from fish mucus. Additionally, 164 bacterial strains were provided by BioViotica GmbH, Germany, for chemical screening. I was in charge of the selection of new environmental samples, isolation of bacterial strains, library maintenance, and preparation of extracts from around 400 strains. Different soil and water sample treatments as well as media conditions were tested for the isolation of bacteria, with a focus on chemically talented actinomycetes. Once in the strain collection, bacteria were cultured using my optimized and standardized protocol, extracted, fractionated, and tested in cell viability assays. In total, the screening of 194 strains has been completed, 30 extracts showed moderate cytotoxicity against human colon cancer cell line (HCT-116) with less than 50% cell survival at a sample concertation of 10 µg/mL. Five known cytotoxic antibiotics were dereplicated from extracts and fractions just by UV and molecular weight comparison. Currently, the Loesgen Lab library contains ~400 bacterial strains, expanded from the here established protocols and techniques.

2.2 Introduction

Natural products from bacterial sources have provided a great number of FDA approved drugs. For example, the antibiotics vancomycin and erythromycin are used to

treat infections with Gram-positive pathogens [1]; doxorubicin and dactinomycin, as cytotoxic agents, are used to treat breast cancer and bladder cancer [2]; the immunosuppressant agent rapamycin is used to prevent the rejection of kidney transplants [3]. But the discovery of new chemical entities was not limited to terrestrial bacteria, bioactive compounds were also discovered by exploration of marine environments [4-6]. The obligate marine genus *Salinispora* was identified in 2005 within the ever growing phylum of actinobacteria containing two distinct new species, one of which produced the potent anti-tumor agent salinosporamide A [7]. Thus, investigating bacterial products from unexplored environments and biodiverse niches is promising for the discovery of new natural products and drug leads.

In the Loesgen lab, environmental samples and bacterial strains came from multiple sources. In 2014, we started a collaboration with Prof. Giuliana Panieri at the Arctic University in Trømso, Norway. From two different arctic cruises in 2016 and 2017, we received sediment and water samples collected from the Arctic Ocean. The swabs of juvenile Pacific Ocean fish were provided by Dr. Misty Paig-Tran (California State University, Fullerton) and her students as part of our collaboration to study the fish microbiome as source for new antibiotics. We were given soil samples collected from Marys Peak by Prof. James White (Oregon State University). We collected soils and sediments in Oregon including the OSU campus, Sulphur Springs in OSU's McDonald research forest, and locations close to lava lands and lava tubes near the Deschutes National Forest. The isolates represent about 60% of the bacterial library (2014-2018) and other bacterial strains were provided by BioViotica. Herein, I present details on the isolation of bacteria from different sources, and describe the workflow of bacterial

selection, culturing, extraction, chromatographic separation, and initial chemical and bioactivity screening. From the active fractions and extracts, five known metabolites were easily dereplicated by LR-MS analysis, UV spectrum, and ^{1}H NMR experiments by comparing data with in-house and commercial databases. The protocols established herein on how to collect and process environmental samples, isolate bacteria, implement strains into the collection, and screening the extracts have expanded the strain collection and fueled the work of several graduate students. It also helped to prioritize and identify the projects described in Chapters 3 and 4 of this thesis, in which new and biologically active metabolites were characterized.

2.3 Results and discussion

In my work, I initiated the Loesgen lab bacterial collection with 394 strains, isolated from various environments. The collection included 85 strains from Arctic Ocean sediment, 62 from Oregonian soils, 83 from fish mucus, and 164 strains from the BioViotica. The BioViotica strain collection was derived from global soil sampling, all with existing permits, including sampling sites in Germany, Mexico, Malta, Africa, etc. (Figure 2.1)

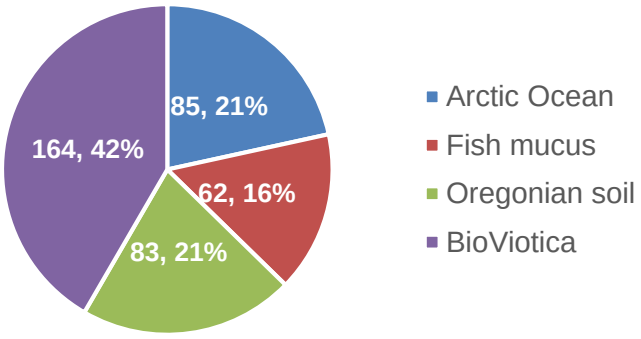


Figure 2.1 Collection of the bacterial strain library of Loesgen Lab.

This chapter is focused on procedures on the isolation/selection of bacteria and the screening of extracts and fractions thereof. Three different types of agar media were used for bacterial isolation from soil samples: nutrient-rich malt-based M2 agar, nutrient poor R2A agar, and commercial actinomycetes isolation agar. Agar plates were supplemented with antifungal agents to allow only bacteria growth. For bacterial isolation from marine environments, we used starch-yeast-peptone (SYP) medium instead of M2 media, as well as R2A agar and actinomycetes isolation agar, but all media were supplemented with 33.3 g/L instant ocean sea salt to mimic sea water (-SW) conditions. In the Loesgen lab, rich M2 media has been used successfully for terrestrial bacteria and SYP-SW media for marine bacterial strains [8].

Collected soil samples were dried in a laminar-flow bio-cabinet to enrich for hardy actinobacteria [9]. As shown in Figure 2.2., 0.5-1.0 g of dried soil, or sediment, was added to 10 mL of autoclaved ultra-pure water. Vigorous shaking for 20 seconds, 1.0 mL solution was further serial diluted into 9.0 mL sterile, ultrapure water tubes respectively to give six samples. 0.1 mL of dilutions #3, #4, #5, and #6 were inoculated into three different agar plates. Inoculated plates were incubated at 28 °C for up to 12 weeks. Any well-separated bacterial colonies, as observed by eye, were picked from the original isolation plates and sub-cultured on M2 or suitable agar plates. Most of the Gram-positive strains possessed morphological features characteristic of the class actinomycetes [10]. All strains were grouped according to colony color, morphology, and pigment production, and some subjected to taxonomic analysis via DNA fingerprinting [11]. We named the selected strains after the location of sampling (Marys Peak “MP”, lava land “LL”, lava tube “LT”), or supporting student’s initials (Cassandra Lew “CL”).

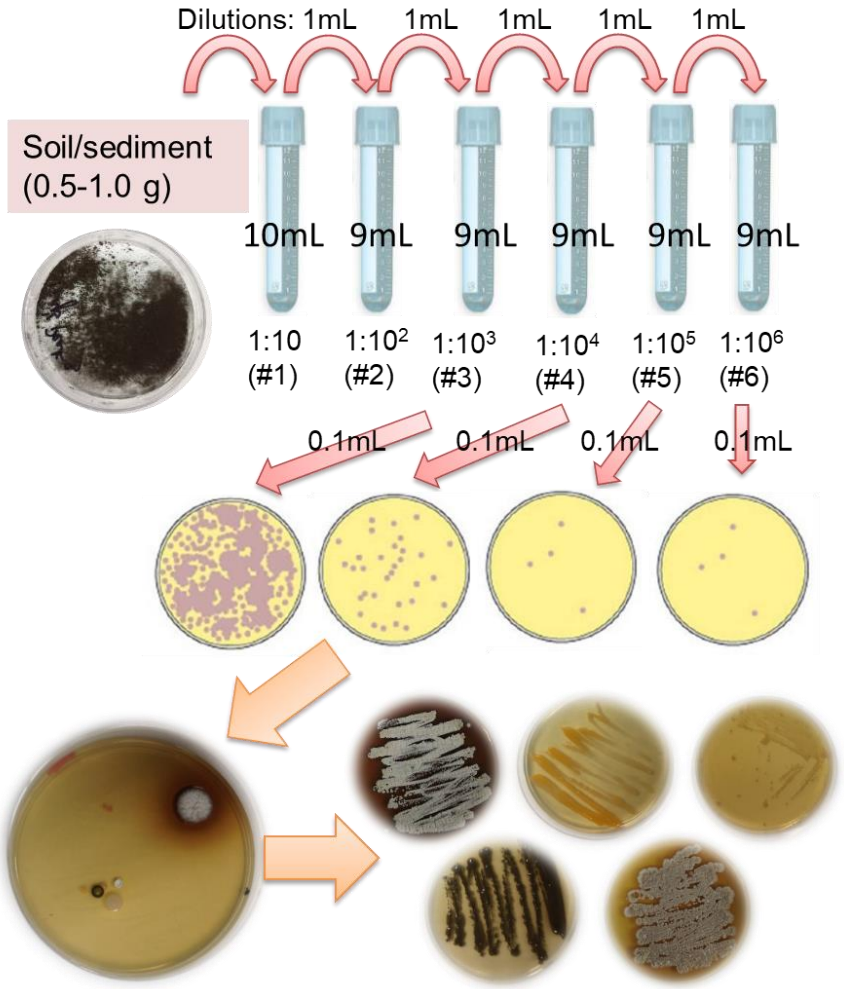


Figure 2.2 Workflow for bacterial isolation from died soil samples or ocean sediments

Collected water samples were shaken for 5 s, then 0.1 mL was directly inoculated onto SYP-SW and actinomycetes isolation-SW agar plates, and then was spread using a sterile applicator loop. Inoculated plates were incubated at 28 °C for up to 12 weeks. Any well-separated bacterial colonies, as observed by eye, were removed from the original isolation plates and sub-cultured on SYP-SW agar plates. Most of the isolated marine strains possessed creamy colored, slimy morphologies, characteristics of Gram-negative bacteria [10].

The mucous swabs of juvenile Pacific Ocean fish were streaked onto SYP-SW agar plates (with antifungal agents) and grown for 5-7 days before colonies were sub-cultured. We named these strains with “PF” as internal IDs.

Based on the goal to find new bioactive natural products, quickly identifying known chemotypes is known as ‘dereplication, and is used in many laboratories to improve discovery efficiency. I designed a protocol of dereplication for known natural products from organic extracts or fractions. All bacterial strains were grown in 500 mL of liquid culture (malt-based M2 medium for terrestrial bacteria, starch-yeast-peptone-seawater (SYP-SW) medium for marine and fish bacteria) for 2-4 week. Both used media are nutrient-rich and can be sued for all bacterial species. 500 mL of liquid culture usually yields around 20-50 mg of extract, which is a suitable amount for one round of fractionation, analysis by LC/MS, and bioassays. The culture broth was sonicated with equal volume of ethyl acetate for 20 min and extracted by separating funnel. The aqueous layer was washed with ethyl acetate twice. Organic layers were combined and then dried in vacuum. Organic extracts were fractionated into seven fractions by vacuum liquid column chromatography (VLCC), using silica gel with mobile phase from 100% dichloromethane (DCM) to 100% methanol. The extract and fractions were prepared at a concentration of 10 mg/mL in DMSO as stock solutions for bioactivity screening. An MTT-based cell viability assay using the human colon cancer cell line (HCT-116) was used to estimate the cytotoxic properties present in organic extracts and fractions. Colon carcinoma cell line HCT 116 is a robust model for initial testing as it is easy to culture, has a fast growth rate, and is sensitive to most anticancer agents [12]. Additionally, initial chemical screening for project prioritization was carried out by a low-resolution LCMS

system using UV absorbance at 210, 254, and 360 nm wavelengths and m/z values. Dereplication of known natural products was achieved by comparison of UV spectra and retention times to our in-house database, which has ~3500 UV spectra and mass information of known natural products, in addition to comparison of spectroscopic data to the commercial AntiBase [13] and AntiMarin Databases [14]. Detailed procedures were illustrated in section 2.4 materials and methods.

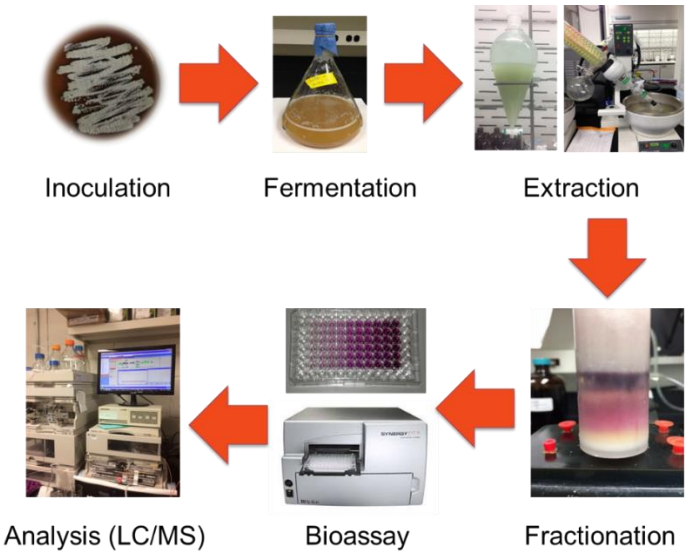


Figure 2.3 Workflow on dereplication/characterization of bioactive metabolites

In total 194 bacterial strains, approximately 50% of the bacterial library collection at that time, have been extracted and tested against human colon carcinoma (HCT-116). Out of the 194 extracts, four extracts exhibited strong activity with less than 20% cell survival against HCT-116 (yellow highlight), and 30 extracts exhibited its growth with less than 50% cell survival (bold), illustrated in Table 2.1.

Table 2.1. Results of single dose MTT-based cell viability assay against human colon cancer cell line (HCT-116). Organic extracts were tested at 10 µg/mL. 250 µM of etoposide was used as positive control.

Strain Name	Cell survival ± SD (%)	Strain Name	Cell survival ± SD (%)	Strain Name	Cell survival ± SD (%)
AO-01	100.6 ± 1.7	AO-70	103.3 ± 3.3	K14/11N	107.1 ± 13.5
AO-02	**38.7 ± 3.9**	AO-71	85.1 ± 1.5	K14/12N	106.7 ± 0.4
AO-03	133.0 ± 22.3	AO-72	72.7 ± 4.0	K14/13	90.9 ± 3.1
AO-05	84.4 ± 0	AO-73	118.5 ± 11.6	K14/5	102.0 ± 10.7
AO-06	145.3 ± 7.2	AO-74	55.8 ± 7.6	K14/7	108.3 ± 5.8
AO-07	108.9 ± 13.1	AO-75	115.4 ± 11.4	K17/4	86.6 ± 17.5
AO-08	98.9 ± 2.1	AO-76	75.2 ± 2.7	K18/4N	100.1 ± 14.8
AO-09	115.4 ± 24.7	**AO-78**	**46.6 ± 1.0**	**K18/6**	**38.5 ± 4.7**
AO-10	102.8 ± 15.1	AO-79	67.0 ± 10.0	K9/12	69.2 ± 6.6
AO-11	134.8 ± 28.2	AO-80	93.9 ± 18.4	**K9/13**	**29.9 ± 0.3**
AO-12	67.2 ± 9.6	AO-81	59.7 ± 11.9	K9/14	74.2 ± 1.5
AO-13	102.3 ± 3.7	AO-82	92.4 ± 2.8	K9/15	117.2 ± 7.7
AO-14	78.4 ± 13.9	AO-83	86.2 ± 3.5	K9/17	91.9 ± 15.7
AO-15	83.3 ± 15.9	AO-84	95.6 ± 7.2	K9/19	88.3 ± 0.7
AO-16	99.9 ± 6.4	AO-85	51.2 ± 10.5	K9/2	87.5 ± 5.8
AO-17	90.4 ± 14.4	AO-86	77.5 ± 6.6	K9/20	98.0 ± 6.5
AO-18	107.2 ± 4.3	MP1#3AB	129.8 ± 8.9	**K9/3**	**27.7 ± 1.7**
AO-19	110.1 ± 13.4	MP1#3M	52.1 ± 5.3	K9/8	77.7 ± 1.5
AO-20	130.6 ± 17.2	**MP1#3U**	**41.5 ± 5.4**	K9/9	55.9 ± 2.3
AO-21	53.7 ± 0.5	MP1#4D	59.8 ± 5.6	Kuba10	90.6 ± 10.6
AO-22	121.9 ± 5.7	**MP1#4F**	**23.2 ± 5.4**	**Lu9455N**	**24.2 ± 2.9**
AO-24	88.3 ± 6.5	MP1#4T	108.7 ± 3.7	**Lu9560**	**34.2 ± 2.8**
AO-25	102.1 ± 6.2	MP1#5AA	80.7 ± 21.6	M4	66.7 ± 2.2
AO-26	110.0 ± 0.8	MP1#6A	67.5 ± 13.5	O VI /4	83.2 ± 1.2
AO-27	118.0 ± 24.3	MP1#6AE	73.2 ± 11.0	P4	76.7 ± 30.7
AO-28	118.3 ± 6.8	MP1#6B	105.0 ± 4.7	P5	74.5 ± 10.5
AO-29	64.2 ± 8.6	MP1#6V	106.2 ± 50.9	**SP160**	**20.4 ± 8.2**
AO-30	**37.0 ±0.1**	MP1#Q	81.5 ± 1.2	**PF-1D**	**28.0 ± 0.5**
AO-31	89.7 ± 6.7	MP3#3K	65.2 ± 2.8	PF-24C	123.9 ± 5.9
AO-32	120.2 ± 18.8	MP3#3T	76.3 ± 2.3	PF-24D	93.6 ± 8.6
AO-33	105.9 ± 9.5	**MP3#4N**	**31.6 ± 3.6**	PF-24E	106.7 ± 7.2
AO-34	107.7 ± 17.6	MP3#5Z	75.8 ± 3.9	PF-32A	92.0 ± 5.4
AO-35	105.5 ± 29.5	MP3#6A	17.3 ± 0.4	PF-33A	102.9 ± 6.5
AO-36	101.1 ± 17.2	MP3#6B	79.4 ± 10.9	PF-33B	109.8 ± 14.3
AO-37	112.7 ± 15.8	MP3#6D	83.1 ± 12.2	PF-36A	109.0 ± 11.7
AO-38	88.8 ± 4.9	MP3#6G	92.2 ± 18.2	**PF-37A**	**32.0 ± 5.1**
AO-39	**47.6 ± 0.3**	Spr1sup2A	99.0 ± 20.4	PF-43A	89.5 ± 8.8
AO-40	127.2 ± 37.7	Spr1sup2C	101.7 ± 31.2	**PF-62-A**	**25.3 ± 9.1**
AO-41	**36.6 ± 2.5**	Spr1sed4	90.7 ± 13.8	PF-70A	88.8 ± 4.0
AO-42	57.6 ± 3.5	Spr2sed5	96.0 ± 16.6	PF-70B	100.6 ± 7.3

AO-43	68.0 ± 1.6	**Spr2sup5**	**42.7 ± 4.0**	PF-70C	81.3 ± 7.9
AO-44	77.5 ± 20.2	SprSS	55.6 ± 2.1	PF-71A	97.4 ± 7.3
AO-45	75.7 ± 12.6	**CL12-4**	**28.4 ± 0.3**	PF-72A	82.6 ± 0.9
AO-46	105.9 ± 20.0	CL12-5	74.4 ± 0.9	PF-72B	76.1 ± 0.1
AO-47	137.7 ± 4.7	CL13-6A	80.0 ± 2.8	PF-73A	92.8 ± 0.3
AO-48	70.8 ± 4.7	CL13-6B	99.1 ± 1.2	PF-73B	91.5 ± 2.6
AO-49	94.6 ± 16.7	**CL16-4**	**29.2 ± 1.4**	**PF-74A**	**22.6 ± 2.2**
AO-50	125.2 ± 5.3	CL16-5A	95.0 ± 1.6	PF-75A1	96.9 ± 9.1
AO-51	112.0 ± 24.4	**CL16-5B**	**26.6 ± 0.7**	PF-75A2	84.4 ± 13.5
AO-52	111.8 ± 11.6	CL16-5C	103.7 ± 0.6	PF-75B1	71.7 ± 2.6
AO-53	116.1 ± 17.8	CL17-3	99.1 ± 5.8	PF-75B2	84.3 ± 10.3
AO-54	95.7 ± 0.8	CL18-3A	57.2 ±0 .9	PF-76A	82.0 ± 1.9
AO-55	101.0 ± 8.1	**CL18-4D**	**39.3 ± 0.5**	PF-78A	84.8 ± 4.6
AO-56	104.9 ± 17.4	LL2-3	91.9 ± 6.5	PF-83A	82.5 ± 1.5
AO-57	88.0 ± 1.9	LL2-4A	68.3 ± 5.1	PF-83B	75.0 ± 3.3
AO-58	139.0 ± 19.4	LL2-4B	90.6 ± 6.0	PF-83C	85.5 ± 2.8
AO-59	95.0 ± 14.2	LL2-4C-Y	98.0 ± 8.0	**PF-85A**	**16.4 ± 3.5**
AO-60	70.3 ± 8.6	LL2-4C-B	81.5 ± 11.6	PF-87A	88.6 ± 4.3
AO-61	144.2 ± 14.8	LL2-4E	85.1 ± 6.4	PF-94A	82.9 ± 1.5
AO-63	80.3 ± 3.1	LL2-5A	83.6 ± 5.5	PF-95A	88.2 ± 3.1
AO-64	**38.7 ± 1.4**	**LL2-5B**	**8.0 ± 1.8**	PF-95B1	93.6 ± 1.4
AO-65	91.3 ± 4.5	**LT2-5**	**26.8 ± 15.4**	PF-95B2	86.9 ± 0.7
AO-66	101.9 ± 21.7	**Bay II**	**37.0 ± 1.0**	Etoposide	6.3 ± 0.9
AO-67	81.6 ± 20.1	Dra11N	92.8 ± 13.0		
AO-68	93.5 ± 5.2	**FORM1**	**0.0 ± 0.8**		
AO-70	112.8 ± 23.5	K14/10	83.0 ± 1.4		

Five cytotoxic compounds were dereplicated from four different bacterial strains from the library collection by liquid-chromatography coupled mass spectrometry and UV data. Only if needed, chemical structures of known metabolites were confirmed by ^{1}H NMR data (Figure 2.3). Noteworthy, the dereplication process presented here is useful to quickly identify known, bioactive metabolites and is meant to prioritize bioactive strains with potentially new metabolites for more detailed chemical analysis during my thesis work, this is by no means intended to be a complete structural assignment.

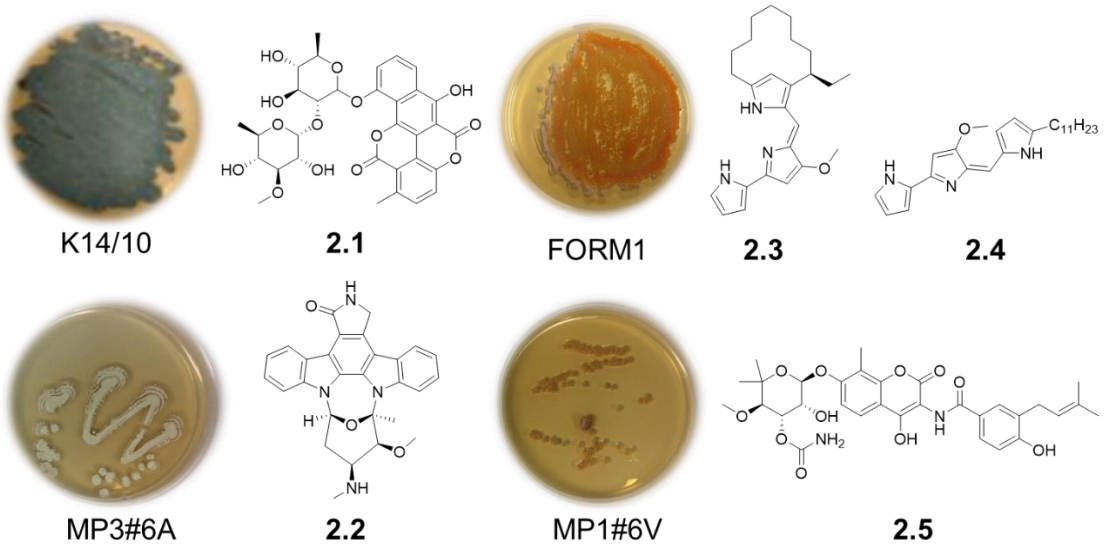


Figure 2.3 Structures of five known cytotoxic compounds were dereplicated from bioactive bacteria. Structure assignment based on UV spectrum, MS data, ^{1}H NMR data only. No stereo configuration was obtained. These are published absolute configurations for **2.1**-**2.3** and **2.5**.

Chartreusin **2.1** was isolated as a bright yellow solid from a *Streptomyces* sp. (K14/10, isolated from Mount Kilimanjaro soil, provided by BioViotica). The UV spectrum of **2.1** displayed maximum absorbances at 242, 274, 334, 380, 402 and 424 nm, matching the chartreusin UV pattern from in-house library and the literature value [15]. Low-resolution mass spectrometry (LRESIMS) gave an m/z values of 663.1 ($[M + Na]^+$) and 1303.3 ($[2M + Na]^+$). The ^{1}H- and ^{13}C- NMR spectroscopic data of isolated **2.1** confirmed the structure as chartreusin by comparison with published NMR data [15]. Chartreusin was first discovered in 1953 and isolated from *Streptomyces chartreusis* [16], it is a well-studied anticancer agent, which binds to GC-rich DNA and therefore inhibits RNA synthesis.

Staurosporine **2.2** was found in the VLCC fractions (F5 and F6) derived from extract from *Streptomyces* sp. (MP3#3A), isolated from a soil sample collected at Marys Peak,

Oregon. The characteristic UV spectrum of **2.2** displayed absorbances at 194, 206, 244, 292, 334, 356 and 370 nm, matching the UV pattern of staurosporine [17] from our in-house UV database. The LRESIMS analysis gave m/z values of 467.2 ($[M+H]^+$) and 933.2 ($[2M+H]^+$),supporting the molecular weight 466.2 for staurosporin ($C_{28}H_{26}N_4O_3$). The known anticancer agent staurosporine was first discovered in 1977 [17] and its absolute configuration established by X-ray analysis in 1994 [18].

Metacyloprodigiosin (**2.3**), also named as streptorubin A, and undecylprodigiosin (**2.4**), were isolated as red pigments from *Streptomyces* sp. FORM1 (isolated from Mexican soil, provided by Bioviotica). Both, the UV spectra of **2.3** and **2.4** displayed characteristic absorbances at 220, 228, and 452 nm wavelengths, matching the UV pattern of streptorubin B (a prodigiosin) from our in-house UV spectra database. LRESIMS gave m/z values of 392.3 for **2.3**, and 394.3 for **2.4**. By searching in this molecular weight range of known compounds within AntiBase [13], five prodigiosins (streptorubin A, streptorubin B, vitamycin A, ethylcyclononylprodiogiosin and cyclomethyldecylprodiogiosin) were found, all with a molecular formulae of $C_{25}H_{33}N_3O$ (MW: 391.56), and one (undecylprodiogiosin, $C_{25}H_{35}N_3O$, MW: 393.58) were found. By comparison with the reported ^1H NMR spectroscopic data [19, 20], compounds **2.3** and **2.4** can be tentatively assigned as metacyloprodigiosin and undecylprodigiosin, respectively. Metacyloprodigiosin (**2.3**) was reported to exhibit potent anticancer activities against multiple cancer cell lines and might be responsible for the strong bioactivity observed in Form1 extracts [21].

Novobiocin **2.5** was found in the organic extract of *Streptomyces* sp. (MP3#6V), isolated from a soil sample collected in Marys Peak, Oregon. The UV spectrum of **2.5**

displayed maximum absorbances at 204 and 334 nm wavelength, matching the UV pattern of novobiocin from our in-house UV database and literature values [22]. LRESIMS gave m/z values of 613.2 ($[M + Na]^+$) and 1247.1 ($[2M + Na]^+$), supportive of novobiocin with a molecular formulae of $C_{31}H_{36}N_2O_{11}$. Novobiocin is a coumarin-derived antibiotic, first obtained from *Streptomyces niveus* in the 1955 [22], and it binds to DNA gyrase and therefore blocks cell energy production. Here, the potent activity observed in the single-dose cell viability assay can be traced to novobiocin, found in all active fractions.

2.4 Materials and methods

2.4.1 General experimental procedures

NMR spectra were acquired on a Bruker Avance III 500 MHz or Bruker Avance III 700 MHz spectrometer, equipped with a 5 mm TXI probe or DCH cryoprobe respectively, with the residual solvent used as internal standards: acetone-*d*$_6$ (δ_H 2.05; δ_C 29.82), $CDCl_3$ (δ_H 7.26; δ_C 77.06), DMSO-*d*$_6$ (δ_H 2.50; δ_C 39.50), and methanol-*d*$_4$ (δ_H 3.31; δ_C 49.00). Low-resolution ESI-MS mass spectra were recorded on Agilent 1100 series LC with MSD 1946. Phenomenx Kinetex C-18 reverse phase column (5 µm × 150 mm × 4.6 mm) was used with a gradient elution from 10% ultra-pure H_2O/ 90% MeCN to 100% MeCN (with 0.05% formic acid in each solvent) at a flow rate of 0.8 mL/min. Dereplication of known natural products was achieved by comparison of UV spectra and retention times to our in-house database in addition to comparison of spectroscopic data to the commercial AntiBase [13] and AntiMarin Database [14]. Semi-preparative HPLC was performed using an Agilent 1100 HPLC on Phenomenx Kinetex C-18 (5 µm × 150 mm × 10 mm) with a gradient elution with ultra-pure H_2O and MeCN at a flow rate of

4.0 mL/min. Preparative HPLC was performed using an Agilent 1260 HPLC system of ultra-pure H_2O and MeCN with Phenomenex Luna C-18 (5 µm × 250 mm × 21 mm). All samples were filtered through a 0.45 µm nylon filter before LCMS and HPLC analysis. Vacuum liquid column chromatography (VLCC) was performed using silica gel 60 (70-230 mesh, Silicycle), with seven mobile phase gradients of dichloromethane-methanol (v/v 99:1, 30:1, 15:1, 9:1, 3:1, 1:1, 0:1). General reagents were purchased from Sigma-Aldrich Corp., Fisher Scientific, and VWR International.

2.4.2 Media conditions

M2 media: malt extract (10 g/L), yeast extract (4 g/L), and glycose (4 g/L). Adjust pH 7.0 prior to sterilization. For making solid agar, 15 g/L agar added before sterilization.

SYP media: soluble starch (10 g/L), yeast extract (4 g/L), and peptone (2 g/L). Adjust pH 7.0 prior to sterilization. 15 g/L agar added before sterilization for preparing solid agar. SYP-SW: SYP medium supplemented with instant ocean sea salt (33.3 g/L)

Actinomycetes isolation agar (medium powder purchased from Thomas Scientific): sodium caseinate (2.0 g/L), L-asparagine (0.1 g/L), sodium propionate (4.0 g/L), dipotassium phosphate (0.5 g/L), magnesium sulphate (0.1 g/L), ferrous sulphate (0.001 g/L), and agar (15.0 g/L). Suspend 21.70 g in 1 L DI water containing 5 ml glycerol. Adjust pH to 8.1 before sterilization.

R2A agar (medium powder purchased from BD Difco™): yeast extract (0.5 g/L), proteose peptone (0.5 g/L), casamino acids (0.5 g/L), glucose (0.5 g/L), soluble starch (0.5 g/L), Na-pyruvate (0.3 g/L), K_2HPO_4 (0.3 g/L), $MgSO_4$ x 7 H_2O (0.05 g/L), and agar (15.0 g/L).

For bacterial isolation, solid agar media were prepared with antifungal drugs by addition of sterile solutions of nystatin (25 mg/L) and cycloheximide (50 mg/L) added post-sterilization. For bacterial isolation from marine sediments and water samples, 33.3 g/L of instant ocean sea salt was added to mimic the sea water condition.

2.4.3 Soil and marine samples collection

Soil samples were collected from six sites in Oregon: OSU campus, lava lands and lava caves in the Deschutes National Forest near Bend, Sulphur springs in Douglas County, and Marys Peak. Soil samples (0.5-1.0 g) were placed in sterile 50 mL plastic Whirl-Pak bags and kept cool until processing, within 72 hours.

From our collaborators at the Arctic University in Trømso, Norway, we received sediment samples and water samples from their 2016 cruise (CAGE 16-5) and 2017 cruise (AMGG) with stops at Prins Karls Forland, Storfjordrenna Pingos and Bjørnøyrenna Craters. On site sampling were carried out with a remotely operated vehicle (ROV), complimented by conventional sampling technologies (Figure 2.4), including Van Veen grab sampler, gravity cores, push cores, box core, blade cores for sediment collection, and 5L Niskin bottles for water sampling. sediment samples (0.5-1.0 g) were placed in 15 mL sterile falcon tube and kept at 0-4 °C. Ocean water (5-10 mL) from Niskin bottles was placed in 15 mL falcon tubes.

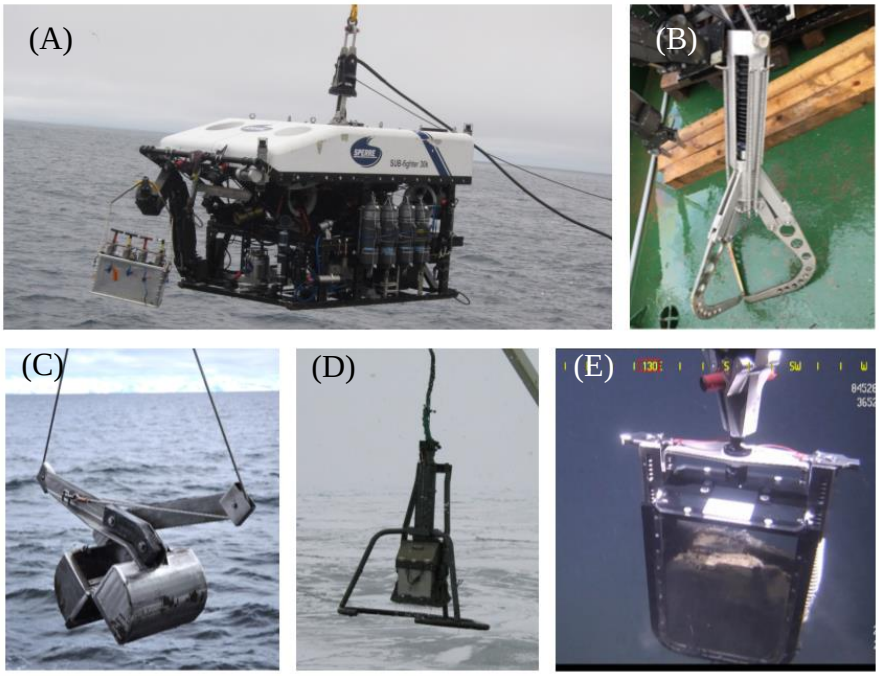


Figure 2.4 (A) ROV deployed at the with benthic sampling gear; (B) Van Veen grab in open mode from RV Helmer Hanssen; (C) Van Veen grab in open mode from RV Helmer Hanssen; (D) A box core in action from RV Helmer Hanssen; (E) A blade corer sampling a bacterial mat at the Storfjørdrenna pingo site.

2.4.4 Bacterial strain isolation and cryopreservation

Collected soil samples were dried overnight in a laminar-flow bio-cabinet to enrich for hardy actinobacteria [9]. Next, 0.5-1.0 g of soil was added to 10 mL of autoclaved ultra-pure water. After vigorous shaking for 20 seconds, 1.0 mL solution was further serial diluted into 9.0 mL sterile, ultrapure water tubes respectively to give six samples. 0.1 mL of dilutions #3, #4, #5, and #6 (dilution of 1: 1,000, 1: 10,000, 1: 100,000, and 1:1,000,000, respectively) were used to inoculate media containing agar plates (M2, R2A, and actinomycetes isolation agar, with antifungal agents) using a sterile applicator loop. Inoculated plates were incubated at 28 °C for up to 12 weeks. Any well-separated bacterial colonies, as observed by eye, were picked from the original isolation plates and sub-cultured on M2 or suitable agar plates. Most of the Gram-positive strains possessed

morphological features characteristic of the class actinobacteria. All strains were grouped according to colony color, morphology, and pigment production, and some subjected to taxonomic analysis via DNA fingerprinting [11].

In terms of marine samples, collected sediments were treated similarly as terrestrial soils without drying in the bio-cabinet. 0.1 mL of dilutions #3, #4, #5 were used to inoculate SYP agar plates and actinomycetes isolation agar plates, both supplied with instant ocean sea salt. Collected water samples were shaken for 5 s, then 0.1 mL was directly inoculated onto SYP and actinomycetes isolation agar plates supplied with instant ocean sea salt and spread using a sterile applicator loop. Inoculated plates were incubated at 28 °C for up to 12 weeks. Any well-separated bacterial colonies, as observed by eye, were removed from the original isolation plates and sub-cultured on SYP-SW agar plates. Most of the isolated marine strains possessed creamy colored, slimy morphologies, characteristics of Gram-negative bacteria.

Fish mucus bacteria (internal designation “PF”) were isolated from the swabs of juvenile Pacific Ocean fish collected by Prof. Misty Paig-Tran (California State University, Fullerton). The mucous swabs were streaked onto antifungal SYP-SW agar and grown for 5-7 days before colonies were sub-cultured. 83 different bacterial colonies were isolated and added into the strain collection. For extract preparation, each bacterium was grown in 100 mL of liquid SYP-SW medium for 7 days before the culture was extracted three times with equal volumes of organic solvent (ethyl acetate).

All isolated strains were cryopreserved by adding 15 Vol% of glycerol solution to bacterial culture broth and store the vials at -80°C.

2.4.5 Extract preparation

Pure colonies were picked and streaked on freshly prepared M2 agar plates for soil bacterial strains or SYP-SW agar plates for marine or fish mucus-derived strains and grown for 5 – 7 days. One 2 cm^2 piece of agar with cells was used to inoculate 50 mL of M2 media as seed culture and grown for 3-5 days, this culture was then used to inoculate 500 mL of M2, or SYP-SW, liquid culture. All bacterial strains were cultivated at 28 ˚C on an orbital shaker at 120 rpm for 2 - 4 weeks. Prior to extraction, all liquid cultures were checked for purity by streaking a small volume onto agar plates and allowing the bacteria to grow and develop phenotypic features for 2-3 days. Bacterial cultures were extracted by first adjusting the pH of cultures to 5.0-5.5 with diluted hydrochloric acid (1N HCl). The organic extracts of each strain were prepared by extraction with an equal volume of ethyl acetate to growth broth. The organic layer was collected, dried over magnesium sulfate, and concentrated under vacuum.

2.4.6 Single dose MTT-based cell viability assay

Inhibition of mammalian cell growth by bacterial extracts and fractions was evaluated against the human colorectal carcinoma model (HCT-116, ATCC CCL-247) [23, 24]. HCT-116 cells were maintained in MEM growth media supplemented with 10% (v/v) fetal bovine serum, penicillin (100 U/mL) and streptomycin (100 µg/mL). The cell lines were incubated at 37 °C in 5% CO_2.

Cellular cytotoxicity was determined measuring the reduction of the tetrazolium salt MTT (3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide) by metabolically active cells. Organic extracts and fractions from bacteria were prepared at 10 mg/mL in DMSO for single dose cellular cytotoxicity analysis. Cells were plated into 96-well plates

with 7,000 cells/well. The cells were maintained overnight before treatment with the addition of organic extract to each well at dose of 10 µg/mL. After 48 h, MTT reagent (5 mg/ml in PBS) was added to each well at a final concentration of 0.5 mg/ml. The plates were incubated for 2 h at 37°C. The growth media was removed, and then purple formazan product solubilized by the addition of 50 µL DMSO. Absorbance was measured at 550 nm using a Biotek Synergy 96-well plate reader. Metabolic activity of vehicle-treated cells (0.1% DMSO) was defined as 100% cell growth. Etoposide (250 µM) was used as a positive control.

2.5 Conclusion

In summary, my efforts were instrumental to expand the bacterial collection in the Loesgen lab to include 394 bacterial strains, extracts and fractions from different sources. Protocols to isolate bacteria from soils and sediments have been tested and established that are now implemented in all ongoing isolation efforts.

From these strains, 194 bacteria have been cultured, extracted, and tested for cytotoxicity in a single dose cell viability assay against a colon cancer cell line (HCT-116). 30 (out of 194) organic extracts showed less than 50% cell survival. From UV and mass spectrometry data, five known anticancer compounds were dereplicated and confirmed by ^{1}H NMR analysis. The chemical and bioactivity screening has prioritized my graduate work and yielded in the new metabolites described in Chapters three and four of this thesis. Additionally, several projects were initiated from my screening efforts, including another graduate thesis project and three undergraduate thesis projects. The microbial library in the Loesgen Lab has grown to over 4000 strains (in

2019) and my protocols have been implemented ever since. In summary, protocols to isolate bacteria, culture, and screen strains from various sources have been established.

2.6 Acknowledgments

This work was supported by OSU start-up funds (SL). We wish to thank Bioviotica (Prof. A. Zeeck and H.-P. Kroll) for providing microbial strains. We wish to thank Prof. Giuliana Panieri and Haoyi Yao (The Arctic University of Norway) and Scott Klasek for collecting sediments and water samples from the Arctic Ocean. We wish to thank Dr. Misty Paig-Tran and her team (California State University, Fullerton) for providing swabs of juvenile Pacific Ocean fish. We also wish to thank Prof. James White (Oregon State University) for collecting soil samples from Marys Peak. Meanwhile, we also wish to thank OSU undergraduate students Cassandra Lew, Kathryn Chen, Molly Austin, and Anabel Mendoza for supporting bacterial isolation, cultivation, and metabolites extraction, and Dr. Birte Plitzko and Elizabeth Kaweesa for performing MTT-based cytotoxicity assays.

2.7 Reference

1. Mahajan, G.B. and L. Balachandran, *Antibacterial agents from actinomycetes - a review.* Front Biosci (Elite Ed), 2012. **4**: p. 240-53.
2. Sobell, H.M., *Actinomycin and DNA transcription.* Proceedings of the National Academy of Sciences, 1985. **82**(16): p. 5328-5331.
3. Pham, J.V., et al., *A Review of the Microbial Production of Bioactive Natural Products and Biologics.* Frontiers in Microbiology, 2019. **10**(1404).
4. Malve, H., *Exploring the ocean for new drug developments: Marine pharmacology.* Journal of pharmacy & bioallied sciences, 2016. **8**(2): p. 83-91.
5. Carroll, A.R., et al., *Marine natural products.* Natural Product Reports, 2019. **36**(1): p. 122-173.
6. Lindequist, U., *Marine-Derived Pharmaceuticals - Challenges and Opportunities.* Biomolecules & therapeutics, 2016. **24**(6): p. 561-571.
7. Fenical, W., et al., *Discovery and development of the anticancer agent salinosporamide A (NPI-0052).* Bioorganic & medicinal chemistry, 2009. **17**(6): p. 2175-2180.
8. Haste, N.M., et al., *Activity of the thiopeptide antibiotic nosiheptide against contemporary strains of methicillin-resistant Staphylococcus aureus.* J Antibiot (Tokyo), 2012. **65**(12): p. 593-8.
9. Jensen, P.R., et al., *Culturable marine actinomycete diversity from tropical Pacific Ocean sediments.* Environ Microbiol, 2005. **7**(7): p. 1039-48.
10. Tortora, G.J., B.R. Funke, and C.L. Case, *Microbiology : an introduction*. 2016.
11. Klindworth, A., et al., *Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies.* Nucleic Acids Res, 2013. **41**(1): p. e1.
12. Skehan, P., et al., *New Colorimetric Cytotoxicity Assay for Anticancer-Drug Screening.* JNCI: Journal of the National Cancer Institute, 1990. **82**(13): p. 1107-1112.
13. Laatsch, H., *AntiBase 2014: The Natural Compound Identifier.* 2014.
14. J.W. Blunt, M.H.G.M., H. Laatsch, *AntiMarin Database.* AntiMarin Database, 2006.
15. Leach, B.E., et al., *Chartreusin, a New Antibiotic Produced by Streptomyces chartreusis, a New Species.* Journal of the American Chemical Society, 1953. **75**(16): p. 4011-4012.
16. Portugal, J., *Chartreusin, elsamicin A and related anti-cancer antibiotics.* Curr Med Chem Anticancer Agents, 2003. **3**(6): p. 411-20.
17. Omura, S., et al., *A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization.* J Antibiot (Tokyo), 1977. **30**(4): p. 275-82.
18. Funato, N., et al., *Absolute Configuration of Staurosporine By X-Ray Analysis.* Tetrahedron Letters, 1994. **35**(8): p. 1251-1254.
19. Wasserman, H.H., G.C. Rodgers, and D.D. Keith, *Structure and Synthesis of Undecylprodigiosin . A Prodigiosin Analogue from Streptomyces.* Chemical Communications, 1966(22): p. 825-+.

20. Wasserman, H.H., G.C. Rodgers, and D.D. Keith, *Undecylprodigiosin.* Tetrahedron, 1976. **32**(15): p. 1851-1854.
21. Darshan, N. and H.K. Manonmani, *Prodigiosin and its potential applications.* Journal of food science and technology, 2015. **52**(9): p. 5393-5407.
22. Hoeksema, H., J.L. Johnson, and J.W. Hinman, *STRUCTURAL STUDIES ON STREPTONIVICIN,1 A NEW ANTIBIOTIC.* Journal of the American Chemical Society, 1955. **77**(24): p. 6710-6711.
23. Kmail, A., et al., *In vitro Assessments of Cytotoxic and Cytostatic Effects of Asparagus aphyllus, Crataegus aronia, and Ephedra alata in Monocultures and Co-Cultures of Hepg2 and THP-1-Derived Macrophages.* Pharmacognosy Communications, 2015. **5**(3): p. 165-172.
24. Mosmann, T., *Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays.* J Immunol Methods, 1983. **65**(1-2): p. 55-63.

Chapter Three: Biodiversity and bioactivity of high desert derived Oregonian soil bacteria

Chenxi Zhu, Cassandra I. Lew, Donovon A. Adpressa, Elizabeth N. Kaweesa, George F. Neuhaus, Birte Plitzko, Lev N. Zakharov, and Sandra Loesgen

3.1 Abstract

In this study, 19 strains of bacteria have been isolated from arid, high desert soil samples collected near Bend, Oregon. Species identification was accomplished using 16S rRNA gene sequencing. Organic extracts were prepared from malt-based liquid cultures of each bacterium and then tested for their respective anti-bacterial and anti-proliferative activities. Antibacterial testing included four human pathogens: Gram-negative *Escherichia coli* and *Pseudomonas aeruginosa*, Gram-positive *Enterococcus faecium* and methicillin-resistant *Staphylococcus aureus*, tested by the microbroth dilution method. Anti-proliferative screens included five cancer cell line models: human colon cancer adenocarcinoma (HCT-116), breast adenocarcinoma (MCF-7), lung carcinoma (A549), malignant melanoma (SK-Mel-5), and prostate adenocarcinoma (PC-3) tested via MTT based cell viability procedures. Noteworthy, six extracts (30%) showed strong inhibition with less than 50% cell proliferation at 10 μg/mL in more than one cell line. Principal component analysis (PCA) of LCMS data revealed drastic differences in the metabolic profiles found in the organic extracts of this panel of high desert derived bacteria. In total, fourteen potent antibacterial and/or cytotoxic metabolites were isolated via bioactivity-guided fractionation, including two new natural products: a new tetrapeptide and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one, as well as twelve known compounds: (-)-furaquinocine C, bafilomycin C1 and D, FD-594, oligomycin A, chloramphenicol, MY12-62A, (4S/4R)-sclerone, isosclerone, tunicamycin VII, tunicamycin VIII, and anthrabenzoxocinone (6*S*, 16*S*)1.264-C.

3.2 Introduction

A complex microbiota can be found in soils globally and their overall contribution to nutrient cycling, the health of soils, and impact on human society has become more evident in recent years [1]. Soil-dwelling bacteria have given us a remarkable diversity and unprecedented amount of bioactive small molecules. A large amount of currently used antibiotics and many cytotoxic reagents that improved treatment of cancers have been inspired by secondary metabolite drug leads from soil bacteria [2, 3]. Many of these structures were found to be produced by filamentous, Gram-positive actinobacteria, with the genus *Streptomyces* being the most prolific and chemically studied member [1]. Inspired by recent studies that showed high abundance of unique biosynthetic gene clusters found in bacteria isolated at high altitude from arid areas in the US [4-7], we decided to explore the species biodiversity of local Oregonian soil bacteria as well as their metabolite and bioactivity potential. We report the isolation of 19 soil-derived bacterial strains, accessed their identity by 16S rRNA fingerprinting techniques, and purified 14 secondary metabolites via bioactivity-guided or chemical-guided fractionation. Among them, two new natural products, a tetrapeptide and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one, were isolated and characterized. Moreover, potent antibiotic and/or cytotoxic compounds, including bafilomycin C1 and D, (-)-furaquinocine C, chloramphenicol, anthrabenzoxocinone, FD-594, tunicamycin VII, tunicamycin VIII and oligomycin A were isolated via bioactivity-guided fractionation. Additionally, the absolute configuration of the quinolinedione natural product MY12-62A was determined by DFT-based ECD computation. This case study supports the notion that the Oregonian high desert is a prolific place for secondary metabolite exploration, besides twelve known antibiotics and cytotoxic compounds, two new metabolites were found and characterized.

3.3 Results and discussion

3.3.1 Species diversity and phylogenetic analysis

Presented here are 19 bacterial strains (Figure 3.1) that were isolated from four soil collection sites close to Bend, Oregon. All bacteria were grown on malt-based (M2) agar plates, with visible colony formation in about 5 -7 days. Only *Actinophytocola oryzae* (LL2-4E) took 6-8 weeks to establish visible growth. Each strain presented a unique phenotype and some noteworthy mentions are *Amycolatopsis saalfeldensis* (LL2-4B) which presented white spores and released a dark pigment into the agar; *Paraburkholderia fungorum* (LL2-4C-B), a non-sporulating, Gram negative bacterium with slimy, black-brown colonies; the spores of *Streptomyces sparsogenes* (LT2-5) which changed color from white to pale yellow to finally black-brown within two weeks. 18 out of 19 strains were identified using DNA fingerprinting techniques by 16S rRNA gene sequencing. Phylogenetic relationship of these bacteria were determined by comparing the partial sequences of their 16S rRNA gene to published sequences using the BLAST algorithm [8] and/or EZTaxon [9]. Next, a phylogenetic tree was generated by combining the 18 sequences with their closest relatives to visualize the relationships among the isolated bacteria by constructing neighbor-joining tree [10] (Figure 3.2). Of the isolated bacteria, 15 belong to the phylum actinobacteria, with a majority of *Streptomyces* spp. (58%). *Arthrobacter agilis*, *Mycolicibacterium hodleri*, and *Amycolatopsis saalfeldensis* were also isolated. Three Gram-negative, proteobacteria were isolated as well (*Pseudomonas trivialis*, *Paraburkholderia fungorum*, and *Mesorhizobium huakuii*).

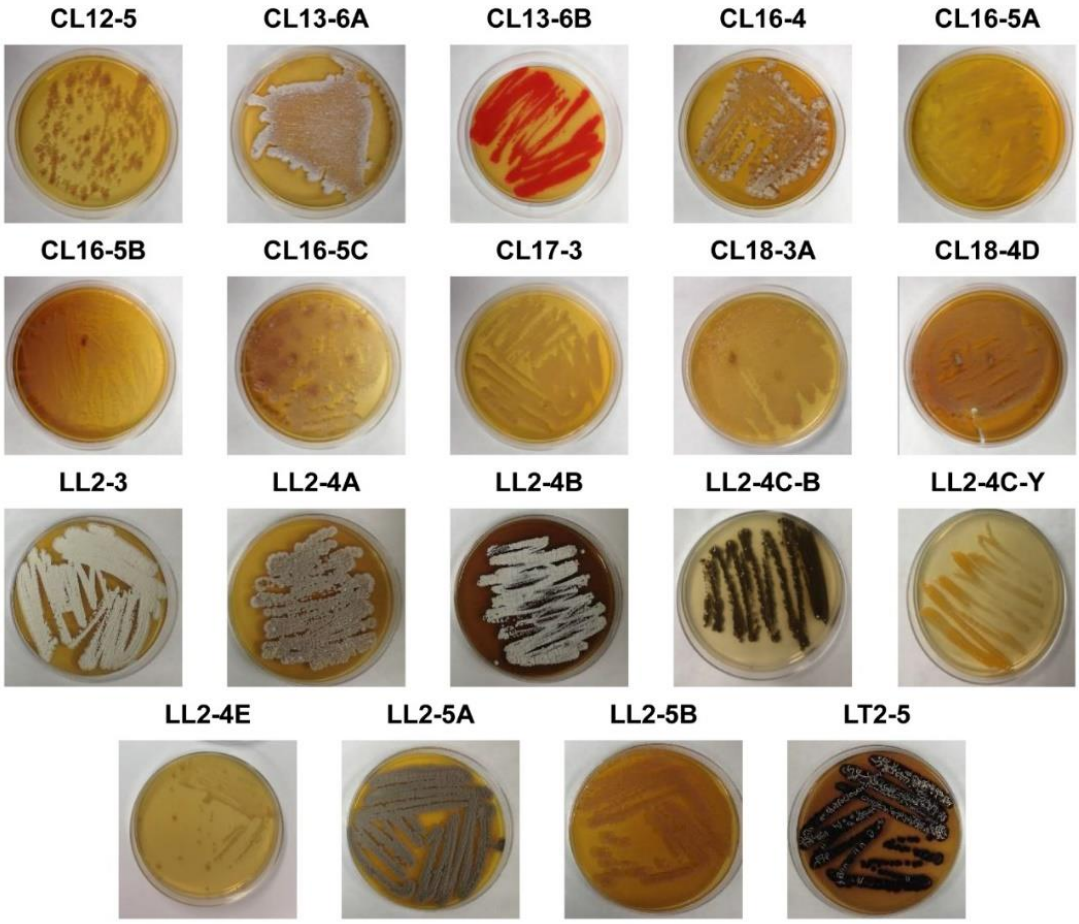


Figure 3.1 Malt agar plates of isolated bacteria, all grown at 28 °C for 2-3 weeks, except bacterium *Actinophytocola oryzae* (LL2-4E) which was grown for 8 weeks.

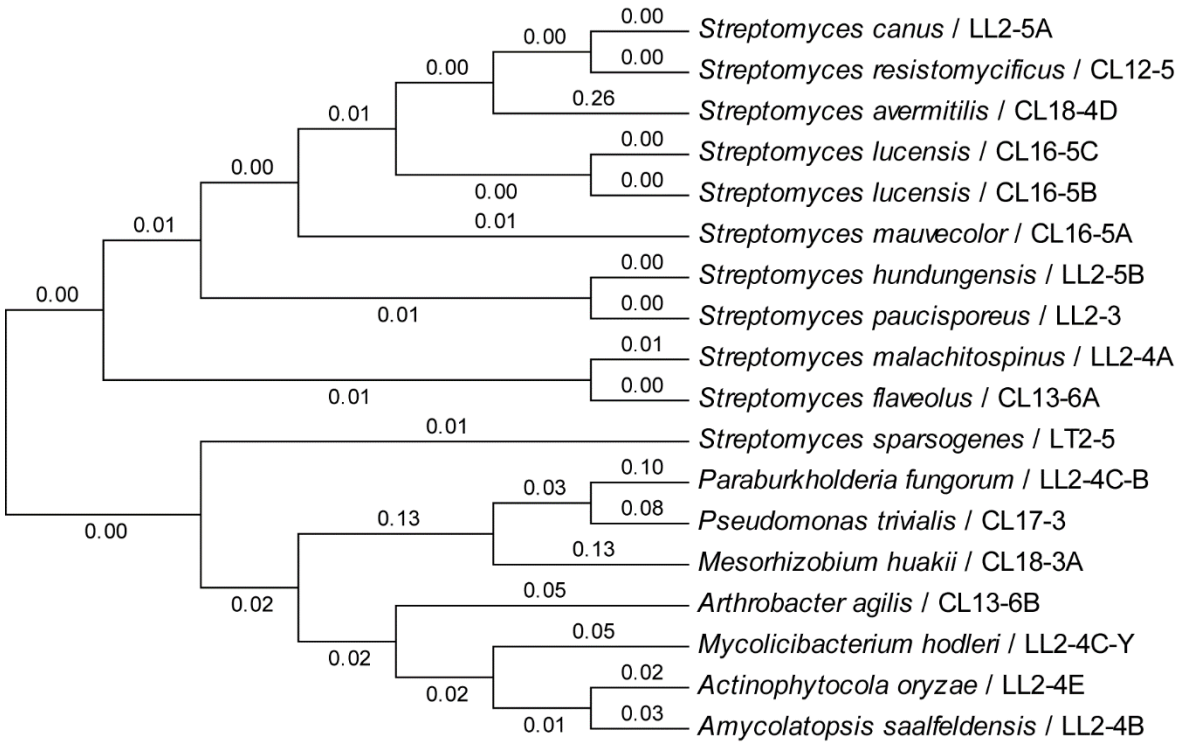


Figure 3.2 Taxonomic analysis of isolated Bend bacteria via 16S rRNA gene sequence amplification, generated with MEGA7 [10]. The evolutionary history was inferred using the Neighbor-Joining method [11].

In collaboration with Prof Brian Murphy (UIC), bacterial colonies were subjected to their IDBac pipeline [12, 13]. The hierarchical clustering based on MALDI-TOF MS protein spectra present a different approach to assess the taxonomic diversity bacterial species [14]. 11 strains were picked as a test case and subjected to MALDI-TOF MS analysis via the IDBac software which revealed similar taxonomic relationships. *Streptomyces sparsogenes* (LT2-5), with presented the unusual spore color changing phenotype, is the most distant to the other *Streptomyces*. Its closest relative is *Streptomyces paucisporeus* (LL2-3) and both share a similar chemotype with the production of anthrabenzoxocinone metabolites. The clustering of *Streptomyces hundungensis* (LL2-5B), *Streptomyces malachitospinus* (LL2-4A), and *Streptomyces resistomycificus* (CL12-5) also agreed with our in-house analysis using 16S rRNA DNA

fingerprinting. In summary, it was exciting to see that the IDBAc pipeline with its MALDI-TOF MS methodology produced comparable taxonomic information for a set of 11 strains in a fraction of the time.

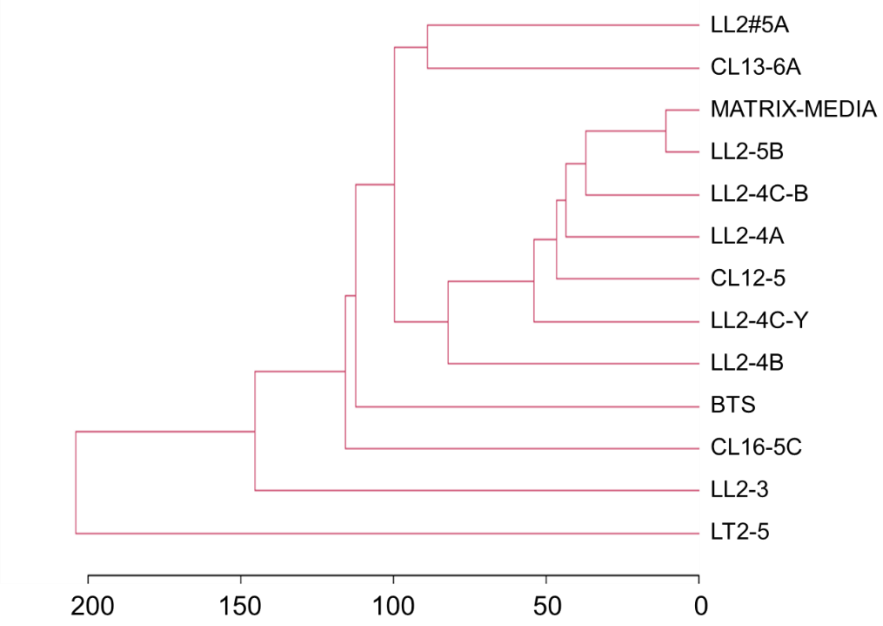


Figure 3.3 Hierarchical clustering of MALDI-TOF MS protein data, generated with the IDBac software [12, 13].

3.3.2 Bioactivity survey

Organic extracts of each bacterial culture were prepared at a concentration of 10 mg/mL in DMSO as stock solution for bioactivity assays. As seen in Figure 3.3A, cytotoxic activities were evaluated at 10 µg/mL against five cancer cell lines: human colorectal carcinoma model (HCT-116, ATCC CCL-247), human breast adenocarcinoma cell line (MCF-7, ATCC HTB-22), melanoma (SK-Mel-5, ATCC HTB70), prostate cancer (PC3, ATCC CRL-1435), and lung cancer (A-549, ATCC CCL-185). Six extracts (30% of all samples) showed strong inhibition with more than 50% cell growth inhibition in at least one cancer cell line. Both, *Streptomyces sp.* (CL16-4) and *Streptomyces*

lucensis (CL16-5B) inhibited cell proliferation by 50% in all five cancer cell lines. *Streptomyces resistomycificus* (CL12-5) showed inhibition against both melanoma (SK-Mel-5; 41% cell survival) and prostate carcinoma (PC3; 44% cell survival). *Streptomyces hundungensis* (LL2-5B) and *Streptomyces sparsogenes* (LT2-5) showed good activity against SK-Mel-5 (34% cell survival) and was highly potent against HCT-116 with 8% cell survival. *Streptomyces avermitilis* (CL18-4D) showed selective inhibition against melanoma (SK-Mel-5, 2% cell survival), and moderate inhibition against colon cancer (HCT-116, 40% cell survival).

In our survey for antibiotic activity (Figure 3.3B), 42% of the strains (8 out of 19) illustrated strong inhibition against Gram-positive human pathogens in their extracts tested at 125 µg/mL. Among these 8 strains, *Streptomyces resistomycificus* (CL12-5), *Streptomyces lucensis* (CL16-5B), *Streptomyces sparsogenes* (LT2-5), *Streptomyces paucisporeus* (LL2-3), and unidentified bacterium (CL16-4) all produce cytotoxins. Only three strains, *Streptomyces malachitospinus* (LL2-4A), *Amycolatopsis saalfeldensis* (LL2-4B), and *Paraburkholderia fungorum* (LL2-4C-B) exhibited antibacterial property without accompanged cytotoxicity. Noteworthy, except for *Paraburkholderia fungorum* (LL2-4C-B), all of the antibiotic producers belong to genus *Streptomyces* and only *Streptomyces malachitospinus* (LL2-4A) showed dual inhibition against Gram-negative and Gram-positive pathogens.

Bioactivity-guided fractionation and isolation of metabolites was based on the results of these single dose cytotoxicity and antibacterial assays.

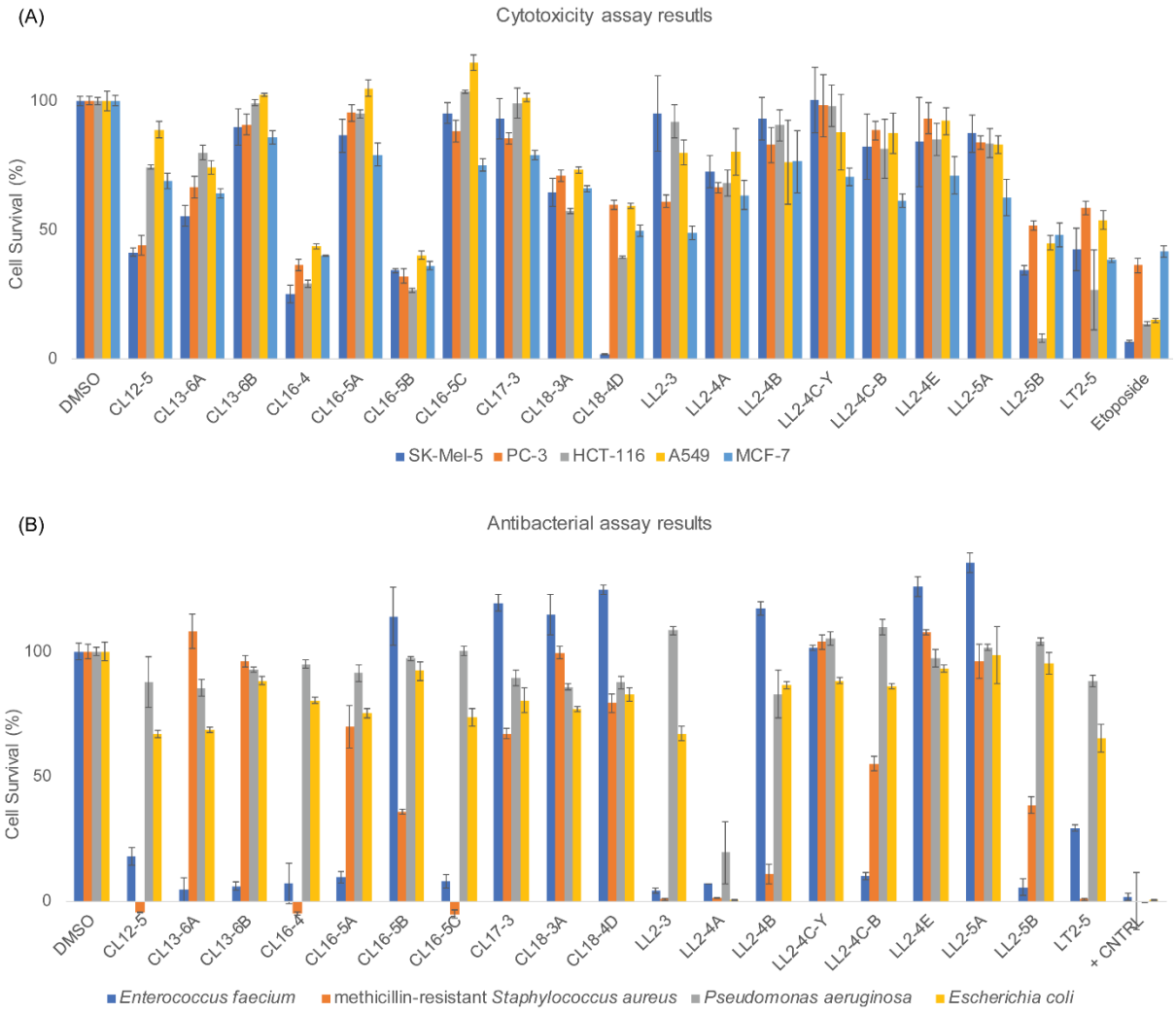


Figure 3.4 Results of single dose bioactivity tests for organic extracts. (A) Cytotoxicity assay results. (B) Antibacterial microbroth assay results. The detailed information on assay results can be found in the supporting information, Table B1 and Table B2.

3.3.3 Dereplication of known bioactive metabolites

Two to eight L liquid culture for each bioactive strain were prepared to enable compound isolation. Antibacterial/cytotoxic screening was used in-between steps of chromatographic separation to guide the isolation efforts. Twelve known metabolites and two new natural products were isolated from bioactive extracts. The structures are listed in Figure 3.5.

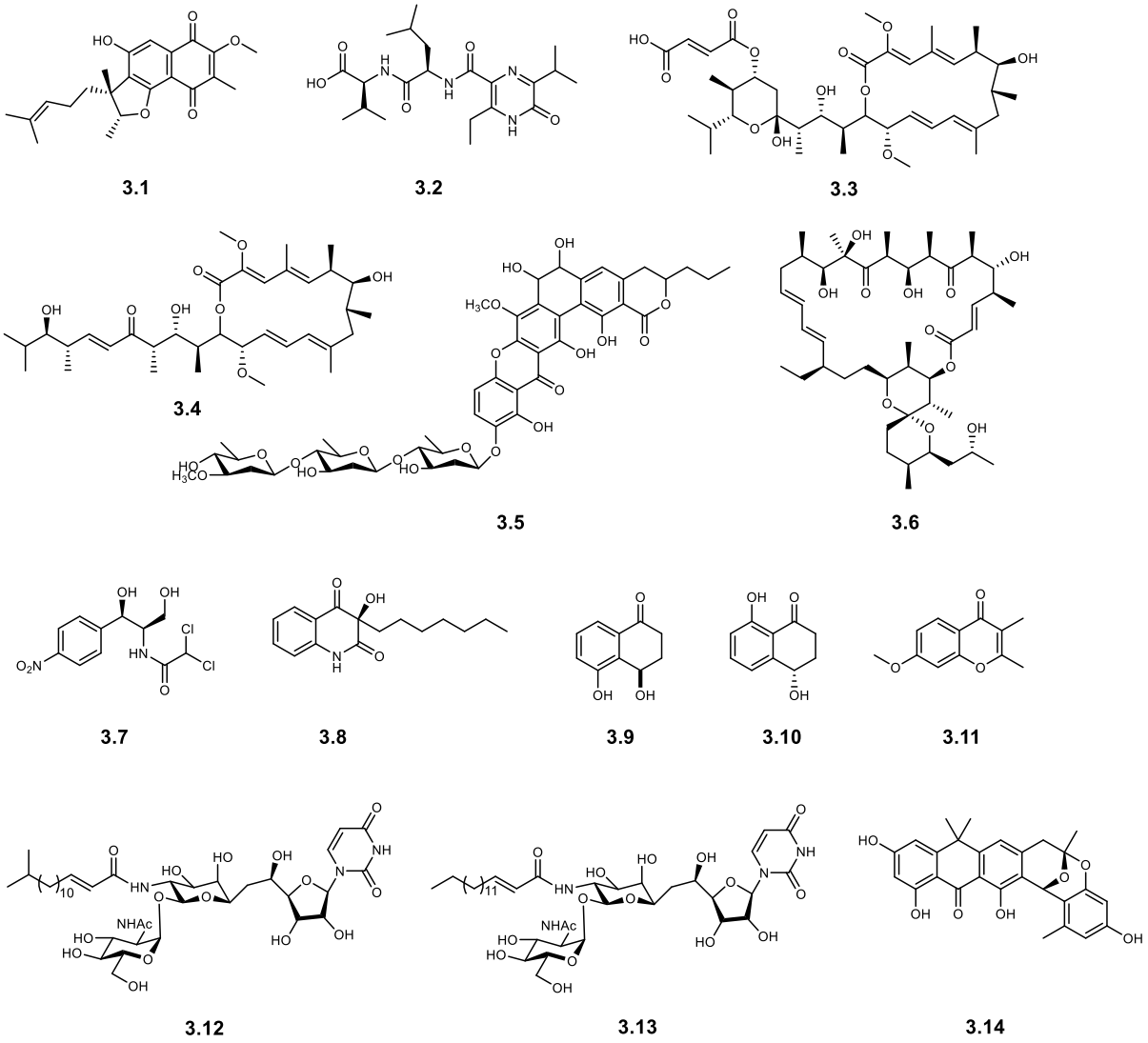


Figure 3.5 Structures of isolated metabolites from 19 bacterial strains from Bend, Oregon.

(-)-Furaquinocin C (**3.1**) was isolated as a yellow solid from *Streptomyces resistomycificus* (CL12-5), with a yield of 1.2 mg/L. The molecular formula was assigned to $C_{22}H_{26}O_5$, based on the HRESIMS result m/z 393.1670 ($[M + Na]^+$; calc. for $C_{22}H_{26}O_5Na^+$, 393.1672; Δppm = 0.6). The optical rotation, UV absorbance, and NMR spectroscopic data of **3.1** agreed with the literature values published for furaquinocin C

by Ishibashi M. et al. [15] in 1991, and synthesized (-)-C-furaquinocin by Smith A.B. et al. in 1995. [16]

Bafilomycin C1 (**3.3**) and bafilomycin D (**3.4**) were isolated from an unidentified bacterium CL16-4, which showed potent inhibition in both cytotoxicity assay and antibacterial assay. The yields were 2.4 mg/L and 1.6 mg/L, respectively. Bafilomycin C1 and D were originally isolated from *Streptomyces griseus* Tü 2599 (DSM 2610) [17, 18]. Bafilomycins are potent inhibitors of ATPases and therefore important tools to study autophagy and other cellular processess [19]. The HRMS data, NMR and ECD spectroscopic data of **3.3** and **3.4** agreed with literature values [17, 20, 21].

FD-594 (**3.5**) was obtained as a yellow solid from two *Streptomyces lucensis* (CL16-5B and CL16-5C), in small yields of 0.2 mg/L. It was identified based on UV, HRMS, and NMR spectroscopic data with literature values for FD-595 [22, 23]. The structure of pyrano[4',3':6,7]naphtho[1,2-b]xanthene antibiotic FD-594 (**3.5**) was originally isolated from *Streptomyces baldaccii* TA-0256 by Qiao Y.F. in 1998 [22, 23], as an anticancer agent inhibiting the growth of HL-60 cells and an antibacterial agent against Gram-positive bacteria.

Oligomycin A (**3.6**) was isolated as a white solid from *Streptomyces avermitilis* (CL18-4D), with yield of 0.5 mg/L. The structure of **3.6** was identified based on UV, HRMS, and NMR spectroscopic data from a recent published paper [24]. Oligomycin A is a potent inhibitor of ATP synthase and strong cytotoxic to most human cells, a recent study found inhibition pf K-Ras PM localization in colorectal carcinoma cell lines (SW620 and SW620 Ad 300) [24].

Chloramphenicol (**3.7**), a widely used broad-spectrum antibiotic, was obtained as a white solid in high yields of 5 mg/L from *Streptomyces malachitospinus* (LL2-4A). The HRMS analysis and NMR spectroscopic data in methanol-d_4 agreed well with literature values [25].

(*R*)-3-heptyl-3-hydroxyquinoline-2,4(*1H,3H*)-dione (**3.8**) was obtained as a white solid from *Amycolatopsis saalfeldensis* (LL2-4B). This metabolite was previously isolated from *Pseudomonas* as a weak 5-lipoxygenase inhibitor and named MY12-62A by Kitamura S. in 1986 [26]. The compound was also reported from a host sponge *Suberea creba* by Debitus C. et al. in 1997 [27]. The HRMS data and NMR spectroscopic data of **3.8** agreed with the literature value for quinolinedione MY12-62A [26, 27]. However, no absolute configuration of MY12-62A has been reported thus far, and the published optical rotation data $[\alpha]_D^{20}$ +2° (0.11 *c*, MeOH) did not agree with our experimental result $[\alpha]_D^{20}$ -34° (0.25 *c*, MeOH). To establish the absolute configuration, we used DFT-based computational approaches to compute the ECD spectrum of the quinolinedione and compared it to the experimental spectrum. To reduce the computational cost, the eight-carbon chain was truncated to a methyl group, this reduced the required number of conformers. Briefly, gas phase geometry optimization was performed in Gaussian [28] followed by solvated geometry optimization. Next, the optimized conformers were submitted to TDDFT computations at multiple levels of theory [29]. All levels, including CAM-B3LYP/TZVP, ωB97X/def2-TZVP, and M062X/def2-TZVP, were in agreement with the *R* enantiomer matching the experimental data (Figure 3.6).

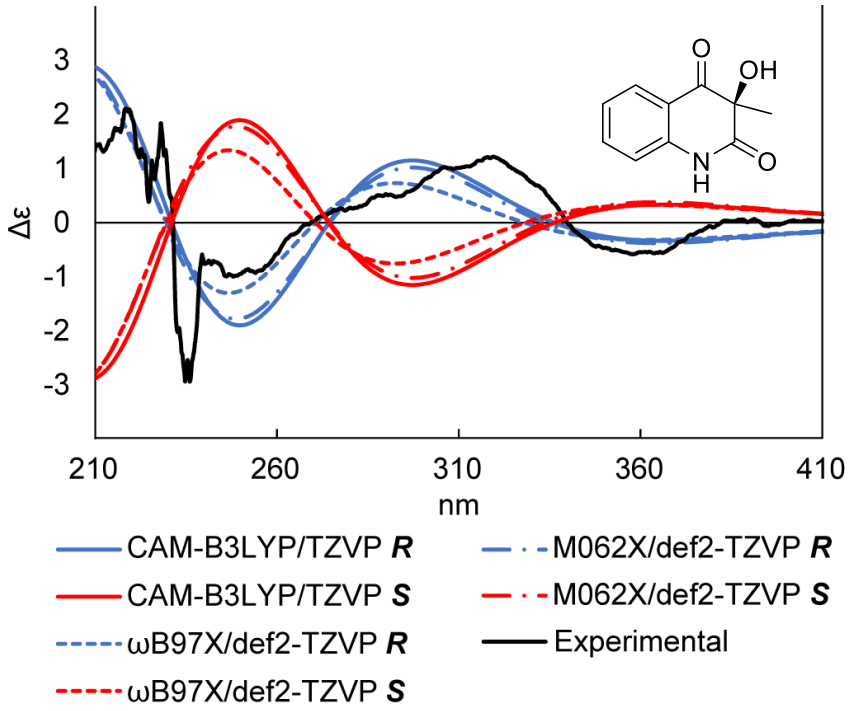


Figure 3.6. Experimental ECD spectra of MY12-62A in methanol with computed spectra of both enantiomers. Red solid line: CAM-B3LYP/TZVP ***S*** (shift: 11 nm; σ = 0.66 eV), blue solid line: CAM-B3LYP/TZVP ***R*** (shift: 11 nm; σ = 0.66 eV), red dotted line: ωB97X/def2-TZVP ***S*** (shift: 14 nm; σ = 0.7 eV), blue dotted line: ωB97X/def2-TZVP ***R*** (shift: 14 nm; σ = 0.7 eV), red dashed line: M062X/def2-TZVP ***S*** (shift: 9 nm; σ = 0.62 eV), blue dashed line: M062X/def2-TZVP ***R*** (shift: 9 nm; σ = 0.62 eV)

(4*S*/4*R*)-sclerone (**3.9**), isosclerone (**3.10**), and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) were isolated from *Paraburkholderia fungorum* (LL2-4C-B), with yield of 0.15 mg/L, 0.3 mg/L, and 1 mg/L, respectively. **3.9** and **3.10** were previously isolated from the Fruit of *Juglans* trees [26-29]; while **3.11** was isolated as a natural product for first time. The structural characterization of **3.11** is discussed in section 3.4.6.

The HRMS data and NMR spectroscopic data of **3.9** agreed with the literature value of sclerone [30]. With an optical rotation of almost zero, indicative of either an achiral compound or a racemic mixture, we moved forward and tested if we were able to separate a possible racemic mixture of **3.9** using chiral column chromatography

(supporting information Figure B28). Indeed, we were able to separate two enantiomers, the conformer with a retention time of 16 min exhibited an optical rotation value of $[\alpha]_D^{20}$ +15° (0.1 *c*, MeOH) which matched the optical rotation value of the published *S* conformer [31], and the conformer eluted at 30 min with an optical rotation value of $[\alpha]_D^{20}$ -13° (0.1 *c*, MeOH) matched data published for the *R* conformer [32]. Additionally, the ECD spectra of both isolated enantiomers matched the published data [32].

The structure of compound **3.10** was confirmed as isosclerone by comparing the HRMS, NMR spectroscopic data [33], and experimental ECD results with literature values [34].

Tunicamycin VII (**3.12**) and tunicamycin VIII (**3.13**) were isolated as white powder from *Streptomyces hundungensis* (LL2-5B). HRESIMS gave **3.12** m/z 867.4253 ([M + Na]$^+$, calc. for $C_{39}H_{64}N_4O_{16}Na^+$ 867.4210, Δppm = 5.0), and gave **3.13** m/z 867.4219 ([M + Na]$^+$; calc. for $C_{39}H_{64}N_4O_{16}Na^+$ 867.4210, Δppm = 1.1), thus $C_{39}H_{64}N_4O_{16}$ as chemical formula was assigned for both **3.12** and **3.13**. The structure of **3.13** was determined by comparing the NMR spectroscopic data of reported synthesized tunicamycin V [35], which has the same structure except for one less methylene on the aliphatic side chain. The structure of **3.12** was determined by comparing the NMR data of reported synthesized tunicamycin V and isolated tunicamycin **3.13**.

Anthrabenzoxocinone (6*S*, 16*S*)1.264-C (**3.14**) was obtained as a yellow solid from *Streptomyces sparsogenes* (LT2-5). The UV spectrum and NMR spectroscopic data agreed with the literature values reported for anthrabenzoxocinone. The ECD spectrum

indicates the absolute configuration is matching to the one reported in the literature for anthrabenzoxocinone (6*S*, 16*S*)1.264-C [36-38].

3.3.4 Structure elucidation of new natural products

Tetrapeptide (**3.2**) was obtained as an amorphous white powder from *Streptomyces resistomycificus* CL12-5. It was purified from VLCC fraction F6 using preparative HPLC with isocratic elution of 43% MeCN/ 57% water. The chemical formula was assigned as $C_{21}H_{34}N_4O_5$ based on HRESIMS ([M - H]$^-$ m/z 421.2462, calc. for $C_{21}H_{33}N_4O_5^-$ 421.2456, Δppm = 1.3) and NMR spectroscopic data. The ^{1}H (700 MHz, DMSO-d_6) and ^{13}C NMR (176 MHz, DMSO-d_6) data are displayed in Table 3.1. The IR absorption bands at 3364, 1653 cm^{-1} indicated the presence of hydroxy, amide, and carbonyl functional groups. The peptidic nature of the compound was identified by ^{13}C shift and DEPT-HSQC correlations for C-1': δ_C 171.5, C-1": 173.1, and C-2': $\delta_{C/H}$ 50.8/4.64, C-2": 57.8/4.04, respectively. Detailed structural information was obtained from the ^{1}H-^{1}H COSY, DEPT-HSQC, and HMBC spectra of **3.2**, as seen in figure 3.7. The ^{1}H-^{1}H COSY spectrum was used to establish the linear sequence from amino proton 2'-NH (δ_H 8.23) to a doublet of triplets signal of methine proton H-2' (δ_H 4.64, δ_C 50.8) and methylene protons H_2-3'(δ_H 1.53, δ_C 42.6), which correlated to a septet methine proton H-4' (δ_H 1.58, δ_C 24.6) connecting to two doublet methyl protons H_3-5' (δ_H 0.91, δ_C 22.3) and H_3-6' (δ_H 0.89, δ_C 23.1). Long range coupling of both H-2' (δ_H 4.64) and H-3' (δ_H 1.53) to carbonyl carbon C-1' (δ_C 171.5) was seen in the HMBC spectrum. Thus, the presence of a leucine residue was supported. Similarly, for a valine residue, the correlation of isopropyl group from H-2" (δ_H 4.04, δ_C 57.8) to H-3" (δ_H 2.07, δ_C 30.3) and also to the two methyl groups H_3-4" (δ_H 0.82) and H_3-5" (δ_H 0.82) was observed in ^{1}H-^{1}H COSY spectrum. The

weak COSY correlation from H-2" (δH 4.04, δC 57.8) to the amide 2"-NH indicated a Leu-Val moiety in **3.2**.

Table 3.1. ^{1}H (700 MHz, DMSO-*d6*) and ^{13}C- NMR (176 MHz, DMSO-*d6*) data of tetrapeptide (**3.2**) vs. the reported natural product JBIR-56

	CL12-5-peptide			JBIR-56 [39]	
Position	δ_C	δ_H (*J* in Hz), type	Position	δ_C	δ_H (*J* in Hz), type
2	155.8		2	155.9	
3	158.4		3	158.7	
5	120.7		5	120.3	
6	145.7		6	145.9	
7	29.3	3.22 sept (6.86), 1H	7	29.5	3.22 q (6.6), 1H
8	19.8	1.16 d (6.79), 3H	8	19.9	1.15 d (6.6), 3H
9	19.7	1.15 d (6.79), 3H	9	19.8	1.15 d (6.6), 3H
10	163.1		10	163.3	
11	22.9	3.01 dq (12.39, 7.28), 1H; 2.93 dq (12.39, 7.28), 1H	11	23.2	2.99 dq (12.6, 7.2), 1H; 2.93 dq (12.6, 7.2), 1H
12	13.8	1.13 t (7.35), 3H	12	14.0	1.13 t (7.2), 3H
Leu			Leu		
1'	171.5		1'	171.4	
2'	50.8	4.64 dt (14.00, 6.79), 1H	2'	50.7	4.52 dd (13.8, 8.4), 1H
3'	42.6	1.53 m, 2H	3'	42.6	1.52 dd (13.8, 6.0), 2H
4'	24.6	1.58 sept (6.65), 1H	4'	24.7	1.56 q (6.0), 1H
5'	22.3	0.91 d (6.37), 3H	5'	23.1	0.89 d (6.0), 3H
6'	23.1	0.89 d (6.44), 3H	6'	22.4	0.88 d (6.0), 3H
2'-NH		8.23 d (8.68), 1H	2'-NH		8.18 d (8.4), 3H
Val			Ala		
1"	173.1		1"	174.0	
2"	57.8	4.04 s, 1H	2"	48.1	4.12 dq (7.2, 6.6), 1H
3"	30.3	2.07 m, 1H	3"	17.7	1.23 d (7.2), 1H
4"	18.0	0.82 d (3.99), 3H	2"-NH		8.34 br s, 1H
5"	19.5	0.82 d (3.99), 3H			
2"-NH		8.14 br s, 1H			

The amino proton 2'-NH showed a correlation to carbonyl carbon C-10 (δ_C 163.1) in HMBC spectrum. Additionally, the ^{1}H-^{1}H COSY spectrum exhibited an ethyl group by coupling between diastereotopic protons H_2-11 (δ_H 3.01, 2.93, δ_C 22.9) and methyl group H_3-12 (δ_H 1.13, δ_C 13.8), and an isopropyl group with methine H-7 (δ_H 3.22, δ_C 29.3), methyl H_3-8 (δ_H 1.16, δ_C 19.8), and methyl H_3-9 (δ_H 1.15, δ_C 19.7). The ethyl group had HMBC correlations to C-6 (δ_C 145.7) and C-5 (δ_C 120.7); while the methyl protons H-12 also showed a correlation to carbon C-6. Both methyl protons H_3-8 and H_3-9 of the isopropyl group had HMBC correlation to carbon C-3 (δ_C 158.4), and the methine proton H-7 connected to C-2 (δ_C 155.8) (Figure 3.6). The remaining NMR resonances indicated a 3,5,6-trisubstitued 2(1*H*)-pyrazinone and the plane structure of **3.2** was constructed. A similar natural product, JBIR-56 from Streptomycetes was reported by Motohashi et al. in 2011 [39]. However, JBIR-56 features an alanine residue instead of the valine residue found in **3.2**. In Table 3.1 we list the NMR data for compound **3.2** and JBIR-56.

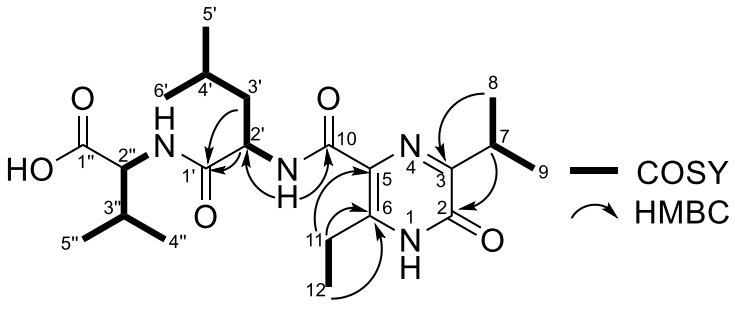


Figure 3.7 Key ^{1}H-^{1}H COSY and HMBC correlations of compound **3.2** (bold lines show ^{1}H-^{1}H COSY, and arrows show HMBC correlations)

To determine the absolute configuration of the new tetrapeptide, advanced Marfey's analysis [40, 41] was undertaken (Figure 3.8A). As seen in the LCMS chromatogram shown in Figure 3.8B, only *L*-valine was detected and confirmed in the hydrolyzed product when compared with the retention times and masses of authentic standards of

D/*L*-Val and *D*/*L*-Leu. Next, we employed DFT-based computations of chiroptical properties of compound **3.2** to determine the absolute configuration as we were limited by the amount of **3.2** isolated. ECD spectra were recorded experimentally and computed to DFT-computed spectra which clearly indicate an *L*-Val-*D*-Leu configuration for the new tetrapeptide.

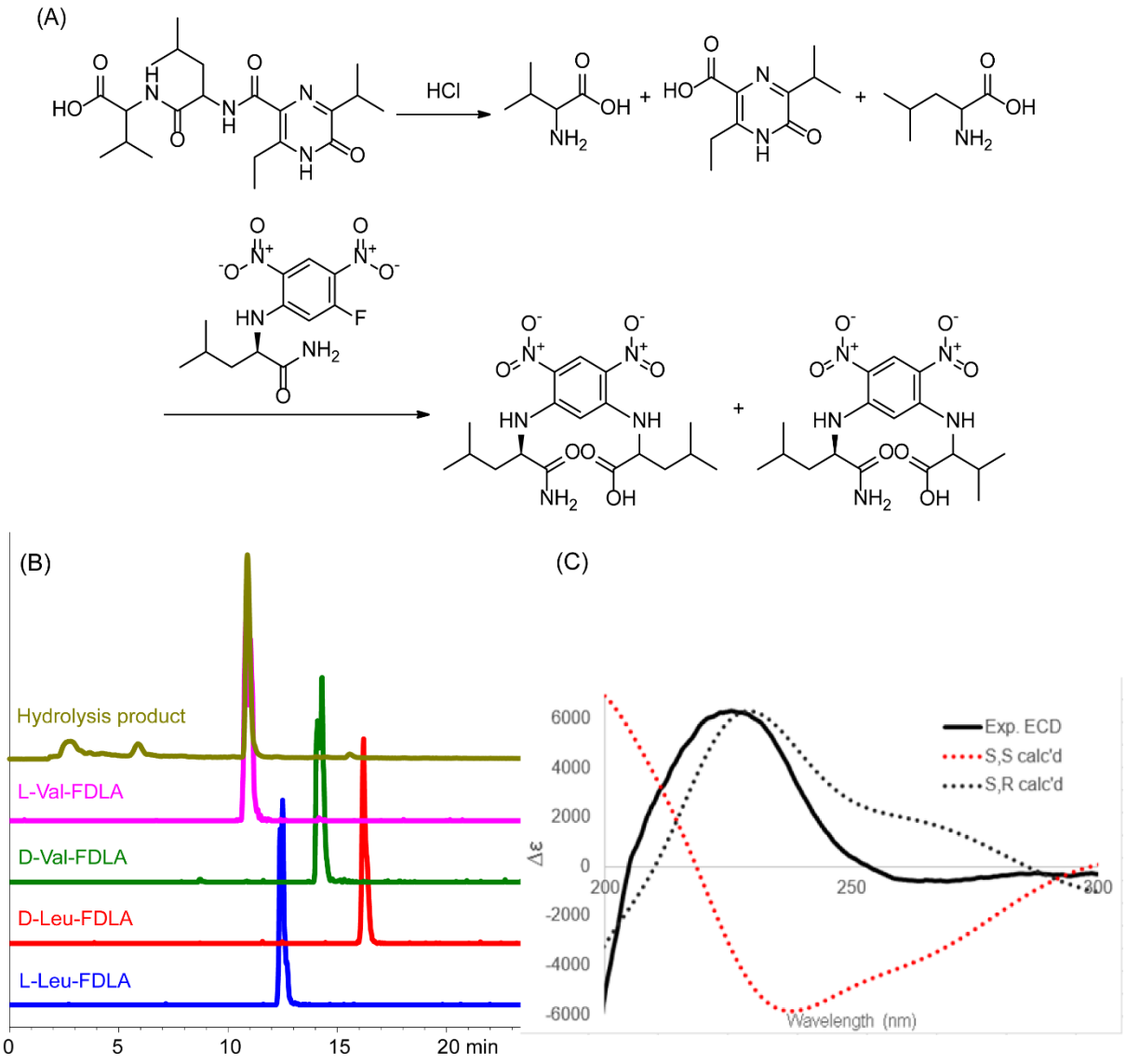


Figure 3.8 The determination of absolute configuration of tetrapeptide (**3.2**). (A) Reaction scheme for advanced Marfey's analysis. (B) LCMS data indicated an *L*-Val moiety. (C) Comparison of computed and experimental ECD spectra indicate an *L*-Val, *D*-Leu configuration for tetrapeptide from *Streptomyces resistomycifics* (CL12-5)

7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) was obtained as yellow needles. The chemical formula was assigned to be $C_{12}H_{12}O_3$ based on the 1D and 2D NMR spectroscopic data and HRESIMS ($[M + H]^+$ m/z 205.0859, calc. for $C_{12}H_{13}O_3^+$ 205.0859, Δppm = 0.1). The ^1H (700 MHz, $CDCl_3$) and ^{13}C NMR (176 MHz, $CDCl_3$) data is shown in Table 3.2. The ^1H-NMR spectrum and DEPT-HSQC revealed three aromatic protons (H-5: δ_H 8.09, δ_C 127.5; H-6: δ_H 6.92, δ_C 114.1; H-8: δ_H 6.77, δ_C 99.9), a methoxy group (7-OCH_3: δ_H 3.88, δ_C 55.9), and two methyl groups (2-CH_3: δ_H 2.38, δ_C 18.7; 3-CH_3: δ_H 2.04, δ_C 10.2). The ^{13}C NMR spectrum indicated a carbonyl carbon (δ_C 177.68) and 7 aromatic or olefin carbons (δ_C 161.5 to 114.1). The structure was elucidated based on the analysis of correlations from ^1H-^1H COSY, DEPT-HSQC, and HMBC spectra with key correlations shown in Figure 3.10. The correlation of H-5 (δ_H 8.09) to C-6 (δ_C 114.1) and C-7 (δ_C 163.7), H-6 (δ_H 6.92) to C-8 (δ_C 99.9), H-8 (δ_H 6.77) to C-7 (δ_C 163.7), C-8 (δ_C 99.9) and C-4a (δ_C 116.8) in HMBC spectra in addition to the correlation of H-5 (δ_H 8.09) to H-6 (δ_H 6.92) in ^1H-^1H COSY proved the substituted benzene. The HMBC spectrum also exhibited a strong correlation from the methoxy group 7-OCH_3(δ_H 3.88) to C-7 (δ_C 163.7). Moreover, the aromatic proton H-5 (δ_H 8.09) correlate to C-4 (δ_C 177.7), which connects to the right part of the molecule. Correlations from the methyl protons 3-CH_3 (δ_H 2.04) to C-2 (δ_C 161.5), C-3 (δ_C 116.8) and carbonyl C-4 (δ_C 177.7), and the correlations of the methyl protons 2-CH_3 (δ_H 2.38) to C-2 (δ_C 161.5) and C-3 (δ_C 116.8) completed the spin system. The high-field chemical shift of C-2 (δ_C 161.5) and C-8a (δ_C 157.8) indicated their connection to oxygen and are part of a chromone ring system. Thus, the structure of compound **3.11** was assigned as 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one. Additionally, compound **3.11** formed X-ray diffraction

analysis suitable crystals after being dissolved in methanol and then placed in a secondary container containing ultrapure water. The X-ray crystallographic analysis supported the proposed structure for compound **3.11** (Figure 3.9B).

Table 3.2. ^{1}H (700 MHz, $CDCl_3$) and ^{13}C-NMR (700 MHz, $CDCl_3$) data of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**)

Position	δ_C	δ_H (*J* in Hz), type
1		
2	161.5	
2-CH_3	18.7	2.38 s, 3H
3	116.8	
3-CH_3	10.2	2.04 s, 3H
4	177.7	
4a	116.8	
5	127.5	8.09 d (8.9), 1H
6	114.1	6.92 dd (8.9, 2.4), 1H
7	163.7	
7-OCH_3	55.9	3.88 s, 3H
8	99.9	6.77 d (2.4), 1H
8a	157.8	

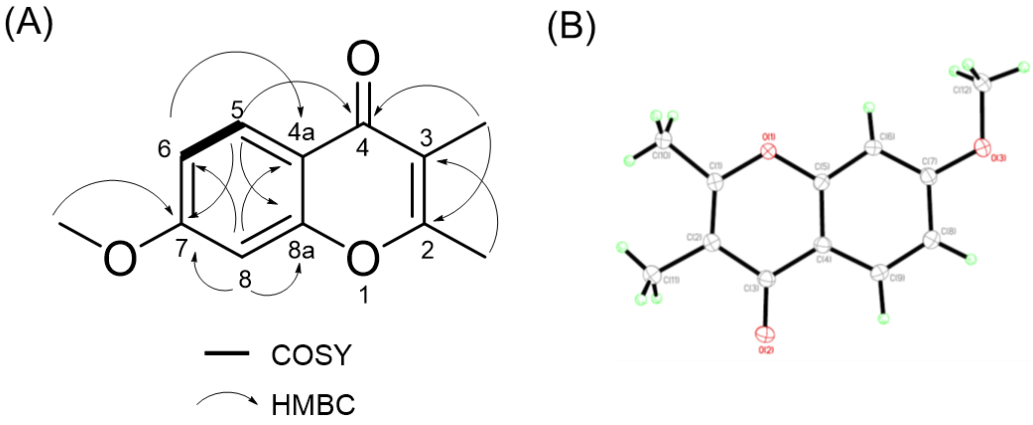


Figure 3.9 (A) Key ^{1}H-^{1}H COSY and HMBC correlations of compound **3.11** (bold lines show ^{1}H-^{1}H COSY, and arrows show HMBC correlations). (B) ORTEP plot of **3.11**.

3.3.5 LCMS-based principal component analysis (PCA) of extracted metabolites

Liquid chromatography mass spectrometry data derived from all bacterial organic extracts was subjected to principal component analysis (PCA). The data demonstrated great chemical diversity is found in the isolated bacteria which aided in strain prioritization and compound isolation efforts. The PCA plot of bacteria isolated in 2014 (Figure 3.10A) showed both *Streptomyces lucensis* CL16-5B and CL16-5C clustering together and both are producing FD-594. The unidentified bacterium (CL16-4) and *Streptomyces avermitilis* (CL18#4D) grouped closely together and both produced macrolide antibiotics bafilomycin C1/D and oligomycin A. Interestingly, *Streptomyces resistomycificus* (CL12-5) which produced furaquinocin C and the new tetrapeptide, were the most distant in its chemotype compared to the other strains in the PCA analysis. The bacteria isolated in 2016 were analyzed at a later timepoint (Figure 3.10B). Here, no overlapping patterns were observed, indicative of a somewhat unique chemical profile in each extract. For example, only *Streptomyces hundungensis* (LL2-5B) produced the glycosylated nucleoside antibiotics tunicamycins and is seen distant from other strains in the PCA, while the two anthrabenzoxocinone producing strains, *Streptomyces sparsogenes* (LT2-5) and *Streptomyces paucisporeus* (LL2-3), clustered closer together. Noteworthy, we found that extracts containing active metabolites were placed far from the center in the PCA plot which could indicate that there is a set of commonly shared metabolites present in all extracts, but specialized secondary metabolites are more species specific.

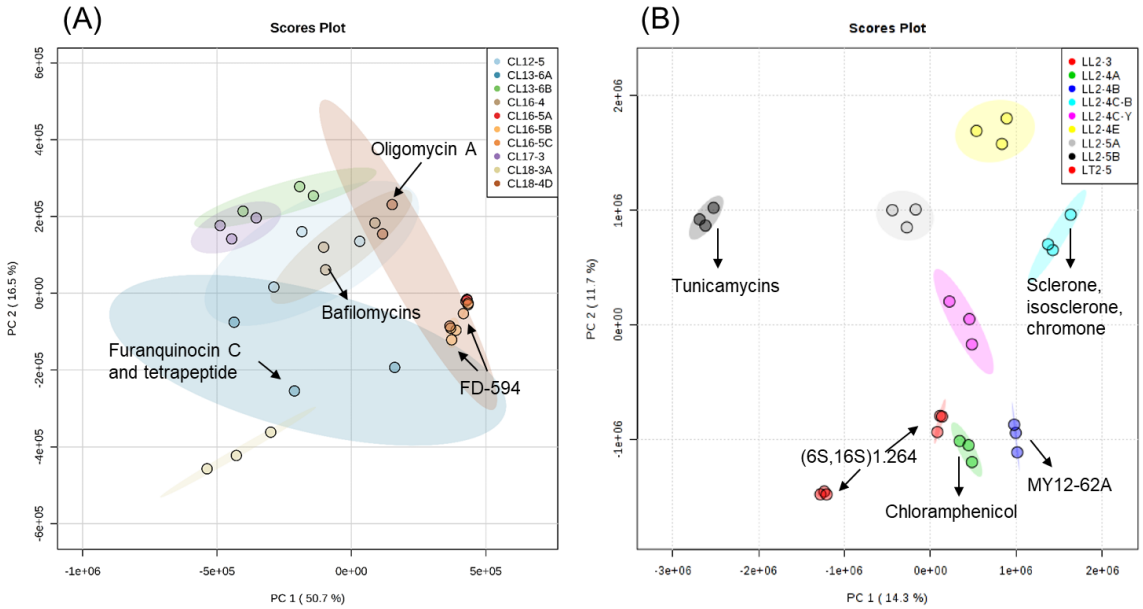


Figure 3.10 PCA plots derived from LCMS data of bacterial extracts from strains collected in 2014 (A) and 2016 (B).

3.4 Materials and methods

3.4.1 General experimental equipment

Optical rotations were measured on a JASCO P-1010 polarimeter. Circular dichroism (CD) and UV measurements were recorded using a JASCO J-815 spectropolarimeter. Infrared (IR) spectra were recorded on a Thermo Scientific Nicolet IR100 FTIR spectrometer. NMR spectra were acquired on a Bruker Avance III 500 MHz or Bruker Avance III 700 MHz spectrometer, equipped with a 5 mm TXI probe or ^{13}C cryoprobe respectively, with the residual solvent used as internal standards: acetone-*d6* (δ_H 2.05; δ_C 29.82), CDCl3 (δ_H 7.26; δ_C 77.06), DMSO-*d6* (δ_H 2.50; δ_C 39.50), and methanol-*d4* (δ_H 3.31; δ_C 49.00) [42]. Low-resolution ESI-MS mass spectra were recorded on Agilent 1100 series LC with MSD 1946. High-resolution ESI-MS spectra were recorded on an Agilent 6545 LC/Q-TOF MS. Dereplication of known natural products was achieved by comparison of UV spectra and retention times to our in-house database in addition to

comparison of spectroscopic data to the commercially available databases AntiBase [43] and AntiMarin Database [44]. A Teledyne ISCO Combi-Flash Companion system was used for automated flash chromatography. Analytical and semipreparative HPLC was performed using an Agilent 1100 HPLC. Three columns were used with this system: column 1—Phenomenx kinetex C-18, 5 µm × 150 mm × 4.6 mm (a gradient elution with ultra-pure H_2O and MeCN with 0.05% formic acid in each solvent at a flow rate of 0.8 mL/min). Column 2 – Chiralcel® OD-H column 10 µm × 250 mm × 4.6mm (at a flow rate of 0.7 mL/min). Column 3 – Phenomenx kinetex C-18, 5 µm × 150 mm × 10 mm (at a flow rate of 4.0 mL/min). Preparative HPLC was performed using an Agilent 1260 HPLC system of ultra-pure H_2O and MeCN. Column 4 Phenomenex Luna C-18, 5 µm × 250 mm × 21 mm). All samples were filtered through a 0.45 µm nylon filter before LCMS and HPLC analysis. Vacuum liquid column chromatography (VLCC) was performed using silica gel 60 (70-230 mesh, Silicycle), with seven mobile phase gradients of dichloromethane-methanol (v/v 99:1, 30:1, 15:1, 9:1, 3:1, 1:1, 0:1). Analytical thin layer chromatography (TLC) was performed on a pre-coated silica gel 60 F_{254} plates (Eppendorf). TLC plates were visualized by UV (254 and 360 nm) and by spraying with anisaldehyde solution followed by heating at 80 °C. General reagents were purchased from Sigma-Aldrich Corp., Fisher Scientific, and VWR International.

3.4.2 Media conditions and isolation of soil bacteria

All cultures were fermented in either M2 media (malt extract (10 g/L), yeast extract (4 g/L), glucose (4 g/L), pH 7.0 prior to sterilization) or R2A (purchased from BD Difco™). Media with antifungal drugs was prepared by the addition of sterile solutions of nystatin

(25 mg/L) and cycloheximide (50 mg/L) added post-sterilization. For solid growth conditions agar was added at 15 g agar per liter to culture media prior to sterilization.

Four soil samples were collected from two different sites close to Bend, OR on June 2014 (43°59'14"N, 121°20'11"W) and June 2016 (44°18'41"N, 121°25'30"W; 43°59'14"N, 121°22'12"W). Soil samples were placed in sterile 50 mL plastic Whirl-Pak bags and kept cool until processing within 72 hours. Collected soil samples were dried overnight in a laminar-flow biocabinet [45]. 0.5-1.0 g of soil was added to 10 mL of autoclaved ultra-pure water. After vigorous shaking for 20 seconds, 1.0 mL solution was further serial diluted into 9.0 mL sterile, ultrapure water respectively to give six samples. 0.1 mL of dilutions #3, #4, #5, and #6 (dilution of 1: 1,000, 1: 10,000, 1: 100,000, and 1:1,000,000) was used to inoculate high and low nutrient media agar plates (medium M2 and R2A) using a sterile applicator loop (VWR).

Inoculated plates were incubated at 28°C for up to 12 weeks. Any well-separated bacterial colonies, as observed by eye, were removed from the original isolation plates and sub-cultured on M2 agar plates. Most of the Gram-positive strains possessed morphological features characteristic of the class actinobacteria [46]. All strains were grouped according to colony color, morphology, and pigment production, and subjected to phylogenetic analysis. All strains were cryopreserved with 15% of glycerol in liquid media at -80°C.

3.4.3 Extract preparation

Pure colonies were picked and streaked on freshly prepared M2 agar plates and grown for 5 – 7 days. One 2 cm^2 piece of agar with mycelium was used to inoculate 50 mL of M2 media as seed culture and grown for 4-7 days, this culture was then used to inoculate

500 mL of M2 liquid culture. All bacterial strains were cultivated at 28 ˚C on an orbital shaker at 120 rpm for 2 - 4 weeks. Prior to extraction, all liquid cultures were checked for purity by streaking a small volume on M2 agar plates and allowing the bacteria to grow and develop phenotypic features for 2-3 days.

Bacterial cultures were extracted by first adjusting the pH of cultures to 5.0-5.5 with diluted hydrochloric acid (1N HCl). The organic extracts of each strain were prepared by extraction with an equal volume of ethyl acetate to growth broth. The organic layer was collected, dried over magnesium sulfate, and concentrated under vacuum. For large scale extraction, the mycelia were filtered off by vacuum filtration after pH adjustment, and the supernatant combined with XAD-7 resin over 24 h. The extract was recovered from the resin using acetone and methanol washes, respectively, after rinsing with an equal volume of DI water to growth broth. The separated cells/spores were sonicated in acetone for 20 minutes and this extract combined with the extract derived from XAD resin. The combined organic fractions were evaporated to dryness using a rotary evaporator.

3.4.4 Species identification of bacteria

For taxonomic analysis, each strain was grown, and cells were collected by centrifugation. The cell pellet was frozen and crushed in liquid nitrogen using a micro pellet pestle. The crushed cells were incubated with lysozyme solution at 37°C for 30 minutes and then incubated with 3 µl of proteinase K and 10% (w/v) sodium dodecyl sulfate (SDS) at 55°C for 30 min. DNA was extracted from a phenol-chloroform solution (v/v 1:1) by centrifugation at 12,000 rpm for 5 minutes. The 16S rRNA loci were amplified from genomic DNA using PCR protocols using the following primers: 8f (5’-AGAGTTTGATCMTGGCTCAG) and 1513r (5’- ACGGCTACCTTGTTACGACTT)

[47], (both purchased from Integrated DNA Technologies, IDT). The 50 μL PCR mixture contained 20 to 40 ng of DNA template, 10 pmol of each primer, 25 μL of MangoMix™ (purchased from Bioline, UK) containing Mango*Taq*™ DNA polymerase, dNTPs, red and orange reference dyes and Mg^{2+}. The PCR program consisted of 30 cycles at 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 90 seconds followed by a final extension step at 72°C for 10 minutes. Amplified DNA products were examined and purified by agarose gel electrophoresis using the PureLink™ quick gel extraction kit (purchased from Invitrogen by life technologies, Germany) according to the manufacturer's suggested protocol. Sequencing reactions were carried out using an ABI 3730 capillary sequence machine at OSU's CGRB Core Facilities with primers 8f and 1513r.

All nucleotide sequences were analyzed and compared to sequences within the NCBI database (http://www.ncbi.nlm.nih.gov/) [8] using the Basic Local Alignment Search Tool (BLAST) [8] or EzBiocloud 16S database [9]. Partial 16S rRNA gene sequences can be found in the supporting information. A phylogenetic tree was created with MEGA7 [10]. The evolutionary history was inferred using the Neighbor-Joining method [11]. The optimal tree with the sum of branch length = 1.01839152 is shown (next to the branches). The evolutionary distances were computed using the Maximum Composite Likelihood method [11].

3.4.5 IDBac

In collaboration with Prof Brian Murphy (UIC), bacterial colonies were subjected to their IDBac pipeline [12]. For matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis, the bacteria were streaked on M2 agar

plates individually and grown for 12 days. Single bacterial colonies were picked for in situ analysis by MALDI-TOF MS for secondary metabolites present and taxonomic protein markers [14]. Converted data results was imported and analyzed in the freely-available IDBac software, following published guidelines[13]. IDBac provides the hierarchical clustering of protein spectra and places bacterial isolates into putative genus- and species-level groups. Isolates within each group were further discriminated based on their differences in specialized metabolites production through analysis of Metabolites Association Networks (MANs).

3.4.6 Principal component analysis (PCA) of extract LC/MS data

Liquid chromatography mass spectrometry data derived from all bacterial organic extracts was subjected to principal component analysis (PCA). The data demonstrated great chemical diversity present in isolated bacteria and aided in compound isolation efforts. Organic extracts, prepared at 10 mg/mL in MeCN, were tested on analytical LC/MS system using a gradient elution (10-100% MeCN with 0.05% formic acid) and samples were tested in triplicates. Peak lists were modified and exported from MZmine [48], and PCA plots were generated with MetaboAnalyst 4.0 [49]. Bacteria isolated from the 2014 collection and the 2016 collection were analyzed separately.

3.4.7 Single dose cell viability assay

Inhibition of cell viability of bacterial extracts were evaluated against five cancer cell line models: human colorectal carcinoma model (HCT-116, ATCC CCL-247), human breast adenocarcinoma cell line (MCF-7, ATCC HTB-22), melanoma (SK-Mel-5, ATCC HTB70), prostate cancer (PC3, ATCC CRL-1435), and lung cancer (A-549, ATCC CCL-185). HCT-116 and SK-Mel-5 cells were maintained in MEM growth media (VWR),

A549, MCF-7 and PC-3 in RPMI 1640 media, each supplemented with 10% (v/v) fetal bovine serum, penicillin (100 U/mL) and streptomycin (100 µg/mL). The cell lines were incubated at 37 °C in 5% CO_2.

Cellular cytotoxicity was determined measuring the reduction of the tetrazolium salt MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) by metabolically active cells. Organic extracts and fractions from bacteria were prepared at 10 mg/mL in DMSO for single dose cellular cytotoxicity analysis. Cells were plated into 96-well plates (HCT-116: 7,000 cells/well; SK-Mel-5: 5,000 cells/well; A549: 7,000 cells/well; PC-3: 5,000 cells/well; MCF-7: 9,000 cells/well). The cells were maintained overnight before treatment with the addition of 10 µg/ml organic extract to each well. After 48 h, MTT reagent (5 mg/ml in PBS) was added to each well at a final concentration of 0.5 mg/ml. The plates were incubated for 2 h at 37°C. The growth media was removed, and then purple formazan product solubilized by the addition of 50 µL DMSO. Absorbance was measured at 550 nm using a Biotek Synergy 96-well plate reader. Metabolic activity of vehicle-treated cells (0.1% DMSO) was defined as 100% cell growth. Etoposide (250 µM) was used as a positive control.

3.4.8 Single dose antibacterial assay

All extracts were tested for antibacterial activity in cell-based assays following established protocols [50]. The antibacterial activity was evaluated against two Gram-positive pathogens including *Enterococcus faecium* (ATCC 49032) and methicillin-resistant *Staphylococcus aureus* (ATCC BAA-41) as well as two Gram-negative pathogens: *Pseudomonas aeruginosa* (ATCC 15442) and *Escherichia coli* (ATCC 8739). Vancomycin and chloramphenicol were used as positive controls in Gram-positive and

Gram-negative pathogens, respectively. Tested extracts and antibiotic controls were dosed at 125 µg/mL. Bacterial growth rates were measured by absorbance at 620 nm using Bioteck Synergy 96-well plate reader.

3.4.9 Advanced Marfey's analysis

Approximately 0.2 mg of tetrapeptide **3.2** was hydrolyzed with 6N HCl at 110°C for 24 hours, then evaporated to dryness and reconstituted in 100 µL of MeOH-H_2O (v/v 1:1) solution. 100 µL of FDLA solution in acetone (0.1% w/v) and 20 µL of 1 M $NaHCO_3$ were added to the hydrolysate and heated to 40°C for 1 hour. The solution was cooled down to room temperature, and then neutralized with 2 N HCl and evaporated to dryness. The residue was resuspended in 100 µL of MeCN-H_2O (v/v 1:1) solution, and then analyzed with LCMS system (column 1 with a gradient method from 35% to 50% MeCN).

3.4.10 X-ray crystallography

Diffraction intensities for compound **3.11** were collected at 173 K on a Bruker Apex2 DUO CCD diffractometer using CuKα radiation, λ= 1.54178 Å. Absorption correction was applied by SADAB [51]. Space group was determined based on intensity statistics. Structure was solved by direct methods and Fourier techniques and refined on F^2 using full matrix least-squares procedures. All non-H atoms were refined with anisotropic thermal parameters. H atoms in the aromatic rings were found on the residual density map and refined with isotropic thermal parameters. Methyl H atoms were refined in calculated positions without restrictions on rotation around the C—C bonds, HFIX 138 in SHELXL [52]. All calculations were performed by the Bruker SHELXL-2014/7 package [52].

Crystallographic Data for 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**): $C_{12}H_{12}O_3$, M = 204.22, 0.10 x 0.05 x 0.04 mm, T = 173(2) K, Triclinic, space group P-1, a = 6.9631(3) Å, b = 10.9375(4) Å, c = 13.2925(5) Å, α = 89.743(2) Å, β = 81.014(2) Å, γ = 80.834(2) Å, V = 986.93(7) Å3, Z = 4, Z'=2, Dc = 1.374 Mg/m3, μ(Cu) = 0.810 mm-1, F(000) = 432, 2θmax = 133.15°, 12378 reflections, 3466 independent reflections [Rint = 0.0532], R1 = 0.0447, wR2 = 0.1167 and GOF = 1.024 for 3466 reflections (307 parameters) with I>2s(I), R1 = 0.0573, wR2 = 0.1267 and GOF = 1.024 for all reflections, max/min residual electron density +0.265/-0.267 eÅ-3.

3.4.11 Computational details

ECD Computational details for My12-62A (**3.8**). Initial conformational analysis was performed in Spartan [53] using the MMFF Merck molecular force field and random rotor conformational search with 100 conformers for the *R* enantiomer. Density functional theory calculations were performed using the Gaussian 09 package.[54] The generated conformer sets was subjected to gas phase QM geometry optimizations and frequency calculations at B3LYP/6-31+G before solvated geometry optimizations at M062X/TZVP with IEFPCM in methanol. ECD spectra were calculated by the TDDFT methodology using the following basis sets: ωB97X/def2-TZVP (UV shift of 14 nm and a σ value of 0.70 eV) utilizing IEFPCM in methanol, CAM-B3LYP/TZVP with IEFPCM in methanol (UV shift of 11 nm and a σ value of 0.66 eV) and M062X/def2-TZVP (UV shift of 9 nm and a σ value of 0.62 eV) utilizing IEFPCM in methanol. ECD spectra were simulated using SpecDis 1.71 [55]. ECD spectra of *S* enantiomer were generated by reflection of spectra of *R* enantiomer.

ECD Computational details for tetrapeptide (**3.2**). Conformers were generated for computation of ECD spectra by a MCMM search at the MMFF level in Spartan 14 [53] for diastereomers of 3.2. Conformers below 5 kcal/mol were carried forward into a QM geometry optimization at the PCM-B97D/def2-TZVP level (acetonitrile) in ORCA. Conformers accounting for greater than 2% of the Boltzmann distribution were subjected to TDDFT calculations at the PCM-ωB97xD/def2-TZVP level. Computationally generated UV spectra were matched to the experimental spectra using SpecDis [55] and the σ and wavelength correction were applied to the computational derived ECD spectra.

3.5 Conclusion

In this study, the biodiversity and bioactivity of high desert derived Oregonian soil strains was explored by phylogenetic analysis by 16S rRNA sequencing, and phenotypic assays for cell viability and antibiotic activity. Prioritization of strains for chemical analysis was made easy by the IDBac pipeline to analyze MALDI-based TOFMS data, and in-house LCMS-based metabolomics approaches. Four bacterial pathogens and five human cancer cell lines were used to screen the organic extracts. Over 50% of extracts (10 out of 19 strains) showed potent antibacterial activity at dose of 125 μg/mL and/or cytotoxic property at dose of 10 μg/mL. Three of them exhibited antibacterial properties without cytotoxicity. Known antibacterial or cytotoxic agents, including the broad antibiotic chloramphenicol, macrolide antibiotics bafilomycin C1/D and oligomycin A, glycosylated natural products tunicamycin VII/VIII, FD-594 and the terpenoid (-)-furaquinocine C were isolated from Oregonian bacteria. Bioactivity- and chemical-guided isolation yielded in the discovery of two new natural products tetrapeptide (**3.2**) and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**), along with twelve known antibiotic

and/or cytotoxic compounds. For compound MY12-62A (**3.8**), we established the absolute configuration for the first time by comparative analysis of experimental and DT-DFT derived spectral data. The new tetrapeptide from *Streptomyces resistomycificus* (CL12-5) (**3.2**) did not show bioactivity in either antibacterial and cytotoxicity assay, while 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) from *Paraburkholderia fungorum* (LL2-4C-B) was slightly active against methicillin-resistant *Staphylococcus aureus* (MRSA) at 10 μM with 65% cell survival. Overall, we were able to isolate 19 strains from only four soil samples. Especially the isolated Streptomycetes dominated the bioactivity survey, no surprise as they are known as the most chemical prolific genus within the actinobacteria.

3.6 Acknowledgements

This work was supported by OSU start-up funds (SL). We wish to thank Prof. Murphy and his team for carrying out MALDI-TOFMS experiments and their support in data conversion and analysis with IDBac. We thank Prof. Kerry McPhail and Dr. David Gallegos for assistance with Marfey's analysis.

3.7 References

1. Traxler, M.F. and R. Kolter, *Natural products in soil microbe interactions and evolution.* Nat Prod Rep, 2015. **32**(7): p. 956-70.
2. Giddings, L.A. and D.J. Newman, *Microbial natural products: molecular blueprints for antitumor drugs.* J Ind Microbiol Biotechnol, 2013. **40**(11): p. 1181-210.
3. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the 30 years from 1981 to 2010.* J Nat Prod, 2012. **75**(3): p. 311-35.
4. Charlop-Powers, Z., et al., *Global biogeographic sampling of bacterial secondary metabolism.* Elife, 2015. **4**: p. e05048.
5. Milshteyn, A., J.S. Schneider, and S.F. Brady, *Mining the metabiome: identifying novel natural products from microbial communities.* Chem Biol, 2014. **21**(9): p. 1211-23.
6. Charlop-Powers, Z., et al., *Chemical-biogeographic survey of secondary metabolism in soil.* Proc Natl Acad Sci U S A, 2014. **111**(10): p. 3757-62.
7. Reddy, B.V., et al., *Natural product biosynthetic gene diversity in geographically distinct soil microbiomes.* Appl Environ Microbiol, 2012. **78**(10): p. 3744-52.
8. Johnson, M., et al., *NCBI BLAST: a better web interface.* Nucleic Acids Research, 2008. **36**(suppl_2): p. W5-W9.
9. Kim, O.S., et al., *Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species.* Int J Syst Evol Microbiol, 2012. **62**(Pt 3): p. 716-21.
10. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.* Mol Biol Evol, 2016. **33**(7): p. 1870-4.
11. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Mol Biol Evol, 1987. **4**(4): p. 406-25.
12. Clark, C.M., et al., *Coupling MALDI-TOF mass spectrometry protein and specialized metabolite analyses to rapidly discriminate bacterial function.* Proc Natl Acad Sci U S A, 2018. **115**(19): p. 4981-4986.
13. Clark, C.M., et al., *Using the Open-Source MALDI TOF-MS IDBac Pipeline for Analysis of Microbial Protein and Specialized Metabolite Data.* J Vis Exp, 2019(147).
14. Silva, R., N.P. Lopes, and D.B. Silva, *Application of MALDI Mass Spectrometry in Natural Products Analysis.* Planta Med, 2016. **82**(8): p. 671-89.
15. Ishibashi, M., et al., *Novel antibiotics, furaquinocins C, D, E, F, G and H.* J Antibiot (Tokyo), 1991. **44**(4): p. 390-5.
16. Smith, A.B., J.P. Sestelo, and P.G. Dormer, *Total Synthesis of (-)-Furaquinocin C.* Journal of the American Chemical Society, 1995. **117**(43): p. 10755-10756.
17. Kretschmer, A., et al., *The Structures of Novel Insecticidal Macrolides - Bafilomycin-D and Bafilomycin-E, and Oxohygrolidin.* Agricultural and Biological Chemistry, 1985. **49**(8): p. 2509-2511.
18. Werner, G., et al., *Metabolic products of microorganisms. 224. Bafilomycins, a new group of macrolide antibiotics. Production, isolation, chemical structure and biological activity.* J Antibiot (Tokyo), 1984. **37**(2): p. 110-7.

19. Bowman, E.J., A. Siebers, and K. Altendorf, *Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 7972-6.
20. Kretschmer, A., et al., *The Structures of Novel Insecticidal Macrolides: Bafilomycins D and E, and Oxohygrolidin.* Agricultural and Biological Chemistry, 2014. **49**(8): p. 2509-2511.
21. Pan, H.Q., et al., *Identification and characterization of the antifungal substances of a novel Streptomyces cavourensis NA4.* J Microbiol Biotechnol, 2015. **25**(3): p. 353-7.
22. Kondo, K., et al., *Structure and biosynthesis of FD-594; a new antitumor antibiotic.* J Antibiot (Tokyo), 1998. **51**(3): p. 288-95.
23. Qiao, Y.F., et al., *Isolation and characterization of a new pyrano[4',3':6,7]naphtho[1,2-b]xanthene antibiotic FD-594.* J Antibiot (Tokyo), 1998. **51**(3): p. 282-7.
24. Salim, A.A., et al., *Oligomycins as inhibitors of K-Ras plasma membrane localisation.* Org Biomol Chem, 2016. **14**(2): p. 711-715.
25. Alshaibani, M.M., et al., *Isolation and characterization of cyclo-(tryptophanyl-prolyl) and chloramphenicol from Streptomyces sp. SUK 25 with antimethicillin-resistant Staphylococcus aureus activity.* Drug Des Devel Ther, 2016. **10**: p. 1817-27.
26. Kitamura, S., et al., *Studies on lipoxygenase inhibitors. II. KF8940 (2-n-heptyl-4-hydroxyquinoline-N-oxide), a potent and selective inhibitor of 5-lipoxygenase, produced by Pseudomonas methanica.* J Antibiot (Tokyo), 1986. **39**(8): p. 1160-6.
27. Debitus, C., et al., *Quinolones from a bacterium and tyrosine metabolites from its host sponge, Suberea creba from the Coral Sea.* J Mar Biotechnol, 1998. **6**(3): p. 136-41.
28. M. J. Frisch, G.W.T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, , *Gaussian 09.* Gaussian 09 (Gaussian, Inc., Wallingford CT, 2009), 2009.
29. Pescitelli, G. and T. Bruhn, *Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra.* Chirality, 2016. **28**(6): p. 466-74.
30. Yang, Q., et al., *Antimicrobial and cytotoxic juglones from the immature exocarps of Juglans mandshurica.* Nat Prod Res, 2019. **33**(22): p. 3203-3209.

31. Liu, L., et al., *New alpha-tetralonyl glucosides from the fruit of Juglans mandshurica.* Chem Pharm Bull (Tokyo), 2004. **52**(5): p. 566-9.
32. Machida, K., et al., *Studies on the constituents of Juglans species. I. Structural determination of (4S)- and (4R)-4-hydroxy-alpha-tetralone derivatives from the fruit of Juglans mandshurica MAXIM. var. sieboldiana MAKINO.* Chem Pharm Bull (Tokyo), 2005. **53**(8): p. 934-7.
33. Zhou, Y., et al., *Studies on Cytotoxic Activity against HepG-2 Cells of Naphthoquinones from Green Walnut Husks of Juglans mandshurica Maxim.* Molecules, 2015. **20**(9): p. 15572-88.
34. Evidente, A., et al., *Regiolone and Isosclerone, Two Enantiomeric Phytotoxic Naphthalenone Pentaketides: Computational Assignment of Absolute Configuration and Its Relationship with Phytotoxic Activity.* European Journal of Organic Chemistry, 2011. **2011**(28): p. 5564-5570.
35. Yamamoto, K., et al., *Total Synthesis of Tunicamycin V.* Org Lett, 2018. **20**(1): p. 256-259.
36. Lam, Y.K.T., et al., *L-755,805, a New Polyketide Endothelin Binding Inhibitor from an Actinomycete.* Tetrahedron Letters, 1995. **36**(12): p. 2013-2016.
37. Herath, K.B., et al., *Anthrabenzoxocinones from Streptomyces sp. as liver X receptor ligands and antibacterial agents.* J Nat Prod, 2005. **68**(9): p. 1437-40.
38. Chen, H., et al., *Isolation of an anthrabenzoxocinone 1.264-C from Streptomyces sp. FXJ1.264 and absolute configuration determination of the anthrabenzoxocinones.* Tetrahedron: Asymmetry, 2014. **25**(2): p. 113-116.
39. Motohashi, K., et al., *JBIR-56 and JBIR-57, 2(1H)-pyrazinones from a marine sponge-derived Streptomyces sp. SpD081030SC-03.* J Nat Prod, 2011. **74**(7): p. 1630-5.
40. Bhushan, R. and H. Bruckner, *Marfey's reagent for chiral amino acid analysis: a review.* Amino Acids, 2004. **27**(3-4): p. 231-47.
41. Bhushan, R. and H. Bruckner, *Use of Marfey's reagent and analogs for chiral amino acid analysis: assessment and applications to natural products and biological systems.* J Chromatogr B Analyt Technol Biomed Life Sci, 2011. **879**(29): p. 3148-61.
42. Babij, N.R., et al., *NMR Chemical Shifts of Trace Impurities: Industrially Preferred Solvents Used in Process and Green Chemistry.* Organic Process Research & Development, 2016. **20**(3): p. 661-667.
43. Laatsch, H., *AntiBase 2014: The Natural Compound Identifier.* 2014.
44. J.W. Blunt, M.H.G.M., H. Laatsch, *AntiMarin Database.* AntiMarin Database, 2006.
45. Mincer, T.J., et al., *Widespread and Persistent Populations of a Major New Marine Actinomycete Taxon in Ocean Sediments.* Applied and Environmental Microbiology, 2002. **68**(10): p. 5005-5011.
46. Barka, E.A., et al., *Taxonomy, Physiology, and Natural Products of Actinobacteria*. Microbiology and Molecular Biology Reviews, 2016. **80**(1): p. 1-43.

47. Klindworth, A., et al., *Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies.* Nucleic Acids Res, 2013. **41**(1): p. e1.
48. Pluskal, T., et al., *MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data.* BMC Bioinformatics, 2010. **11**(1): p. 395.
49. Chong, J., et al., *MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis.* Nucleic acids research, 2018. **46**(W1): p. W486-W494.
50. Wiegand, I., K. Hilpert, and R.E. Hancock, *Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances.* Nat Protoc, 2008. **3**(2): p. 163-75.
51. Sheldrick, G.M., *Bruker/Siemens Area Detector Absorption Correction Program.* Bruker AXS, 1998.
52. Sheldrick, G., *SHELXT - Integrated space-group and crystal-structure determination.* Acta Crystallographica Section A, 2015. **71**(1): p. 3-8.
53. Hanwell, M.D., et al., *Avogadro: an advanced semantic chemical editor, visualization, and analysis platform.* J Cheminform, 2012. **4**(1): p. 17.
54. Frisch, M.J., et al., *Gaussian 09*. 2009, Gaussian, Inc.: Wallingford, CT, USA.
55. Bruhn, T., et al., *SpecDis: quantifying the comparison of calculated and experimental electronic circular dichroism spectra.* Chirality, 2013. **25**(4): p. 243-9.

Chapter Four: A new diterpenoid from soil bacterium *Streptomyces flaveolus* CL12-4

Chenxi Zhu, Donovon A. Adpressa, Cassandra I. Lew, Sandra Loesgen

4.1 Abstract

A new cyclic diterpenoid, along with two known porphyrins, was isolated from soil-dwelling *Streptomyces flaveolus* (CL12-4), found in arid, high desert soil near Bend, Oregon. The chemical formula and molecular structure of the terpene was determined by HRESIMS, 1D and 2D NMR spectroscopic analysis, including COSY, NOESY, ROESY, HSQC and HMBC. Interestingly, the compound exists in two ring-flipped macrocyclic conformations in solution. The new terpene contains five stereo centers and two double bond geometries within a fused bicyclo[8.4.0] tetradecane macrocycle. To determine its three-dimensional structure, advanced Mosher ester analysis, *J*-based coupling analysis, and density functional theory (DFT)-based computational techniques including ForceGen 3D structure generation and electronic circular dichroism (ECD) calculations were used. The absolute configuration of the new diterpenoid was assigned as 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*S*,12a*S*. The compound was found to inhibit cell growths of human colon carcinoma (HCT-116) with 24 % cell survival after treatment with a single dose of 10 µg/mL, comparable to the positive control etoposide (10 µg/mL).

4.2 Introduction

Terpenes are an enormous large class of natural products, with more than 50,000 members that have been isolated and characterized thus far [1, 2]. Terpenoids are mostly isolated from plants and fungi, with an historically smaller amount found in bacteria [3]. The first described odoriferous terpenes from actinomycetes were geosmin [4] and 2-exo-hydroxy-2-methylisoborne [5] reported in the late 1960s. Recent studies revealed that terpene synthases, the enzymes responsible for the biosynthesis of terpenoid chemodiversity, are widely distributed in bacteria, particularly in actinomycetes [6-9]. Although many of these presumptive terpene synthase encoding genes appear to be silent in laboratory studies, more and more terpenes were recently identified using engineered terpenoid biosynthetic gene clusters and heterologous expression in model microorganisms [8, 10]. Terpenoids have been found to exhibit a wide range of biological activity, including antimicrobial, antifungal, antiparasitic, antiviral, anti-allergenic, antispasmodic, antihyperglycemic, anti-inflammatory, and immunomodulatory properties, only few have been shown to exhibit cytotoxic properties against a variety of tumors cells and cancers [11, 12].

Herein, we report the isolation of a newly discovered diterpenoid and two known porphyrins from the cytotoxic organic extract of an Oregonian soil *Streptomyces* sp., internal designation CL12-4. Due to the slow chemical exchange from two ring-flipped conformers in solution, structure elucidation by NMR spectroscopy posed to be a difficult challenge. Mosher ester analysis, 2D NMR experiments including *J*-based analysis, and DFT-based modeling were employed to determine its absolute configuration.

4.3 Results and discussion

The bacterium CL12-4 was isolated from an arid, high desert soil collected near Bend, Oregon. The species was identified by partial amplification and sequencing of the 16S rRNA gene (1394 bp), with universal primers: 8f (5'-AGAGTTTGATCMTGGCTCAG) and 1513r (5'- ACGGCTACCTTGTTACGACTT) [13]. The DNA finger printing technique revealed that CL12-4 showed 99.93% similarity to *Streptomyces flaveolus* NMRC 3715 [14] (supporting information Figure C2)

The ethyl acetate extract from liquid, malt media cultures of *Streptomyces* CL12-4 showed potent inhibition of cancer cell growths with on average of less than 50% cell survival at a dose of 10 µg/ml against five human cancer cell lines: human colorectal carcinoma model (HCT-116), human breast adenocarcinoma cell line (MCF-7), melanoma (SK-Mel-5), prostate cancer (PC3), and lung cancer (A-549). Coproporphyrin III (**4.1**), zincphyrin (**4.2**), and diterpenoid (**4.3**) were isolated from 420 mg extract derived from 4.5 L shaking culture of *Streptomyces flaveolus*.

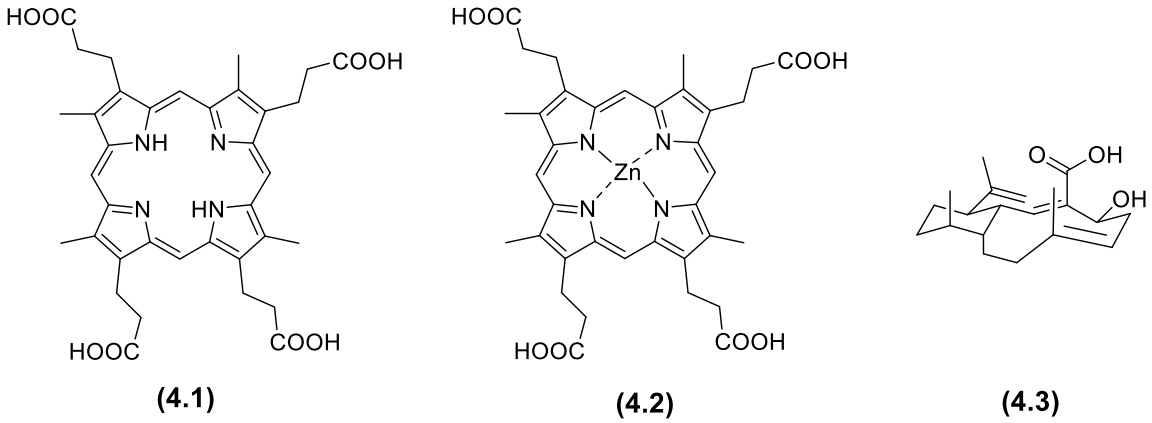


Figure 4.1 Structures of metabolites isolated from *Streptomyces flaveolus* (CL12-4)

Coproporphyrin III (**4.1**) was isolated as a dark brown solid. The HRESIMS gave an m/z 655.2764 ($[M + H]^+$), providing a molecular formula of $C_{36}H_{38}O_8N_4$ (Δppm = 0.2; calcd for $C_{36}H_{39}O_8N_4^+$, 655.2768) with 20 degrees of unsaturation. The UV spectrum of **4.1** (in MeCN) displayed maximum absorbances at 392, 496, 528 and 562 nm, indicating a large conjugated chromophore, which matched the literature values from our in-house database and AntiBase [15] for coproporphyrin III. The ^{1}H NMR spectrum of isolated **4.1** matched previously published spectral data [16].

Zincphyrin (**4.2**) was isolated as a pink solid. The HRESIMS gave an m/z 717.1888 ($[M + H]^+$), providing a molecular formula of $C_{36}H_{36}O_8N_4Zn$ (Δppm = 1.3; calcd for $C_{36}H_{37}O_8N_4Zn^+$, 717.1897). The UV spectrum of **4.2** displayed maximum absorbances at 332, 406, 536 and 572 nm, indicating again a large conjugated chromophore, which matched the literature value from our in-house database and AntiBase. Zincphyrin was first isolated from *Streptomyces* sp. in 1993 by a Japanese chemist Toriya [17]. The ^{1}H NMR spectrum of isolated **4.2** was confirmed as zincphyrin by comparison with published NMR data [16, 17].

Diterpenoid (**4.3**) was obtained as an amorphous white powder. The chemical formula was assigned as $C_{20}H_{30}O_3$ based on HRESIMS m/z 319.2261 (Δppm = 2.1; calcd for $C_{20}H_{31}O_3^+$ 319.2268). The UV spectrum of **4.3** (recorded in MeCN) displayed maximum absorbances at 210 and 228 nm. The major IR bands include a broad band for a carboxy/hydroxyl group at 3446 cm^{-1}, a sharp strong stretch at 2922 cm^{-1} for a terminal alkene, and a carbonyl stretch at 1697 cm^{-1} for a carboxylic acid. ^{1}H-, ^{13}C- and 2D NMR experiments, including COSY, NOESY, ROESY, TOCSY, HSQC and HMBC were utilized to define the structure of diterpenoid **4.3**.

The ^{1}H NMR spectrum of **4.3** exhibited one methyl doublet (δ_H 1.13), two methyl singlets (δ_H 1.60, 1.62), one methine signal (δ_H 3.55) with a doublet of doublets of doublets pattern, four broad signals in the olefinic region (δ_H 5.8-4.9) with integration of 0.4 or 0.6, and two broad methine signal (δ_H 4.26, 4.15) also with integration of 0.4 or 0.6. The ^{13}C NMR spectrum of **4.3** displayed 31 signals, providing 11 carbons more than the number from the predicted chemical formula. These 31 carbon signals consist of one carbonyl carbon at δ_C 170.6 ppm, one olefinic carbon of a terminal alkene moiety at δ_C 111.3 ppm, 11 signals for aromatic or olefinic carbons, and the rest aliphatic carbon signals. From the 2D NMR data, it became evident that some of the carbons and the protons have multiple assignments, as Table 4.1 listed the ^{1}H (700 MHz, methanol-d_4) and ^{13}C NMR (176 MHz, methanol-d_4) data. The complexity observed in the 1H spectra with broad signals and partial doubling of signals exhibiting smaller integration values as well as the additional carbon signals found in ^{13}C NMR spectrum together with the broad, overlapping signals in the ROESY/NOESY spectra, indicate that this compound is undergoing slow chemical exchange. Two possibilities for chemical exchange come to mind on the NMR timescale, either the partial doubling of signals could arise from slow sampling of two tautomeric forms, or slow sampling of two macrocyclic conformations.

Table 4.1 ^{1}H and ^{13}C NMR spectroscopic data for **4.3**

Position	δ_C, type	δ_H (*J* in Hz), integration
1*	35.6, CH	1.69 m, 0.6H
1*	37.9, CH	1.70 m, 0.4H
1-CH_3	19.2, CH_3	1.13 d (7.20), 3H
2	27.8, CH_2	1.74 m, 1.38 m, 2H
3	26.9, CH_2	1.54 m, 1.36 m, 2H
4*	47.6, CH	2.06, 0.4H
4*	47.9, CH	1.99, 0.6H

4a	36.8, CH	3.55 ddd (11.00, 10.30, 4.20), 1H
5*	148.4, CH	5.38 br s, 0.4H
5*	153.7, CH	5.63 br s, 0.6H
6*	131.0, C	
6*	133.5, C	
6-COOH	170.6, C	
7*	77.1, CH	4.26 br s, 0.4H
7*	79.1, CH	4.15 br s, 0.6H
8*	36.2, CH_2	2.72, 0.4H; 2.66, 0.4H
8*	36.3, CH_3	2.35, 0.6H; 2.31, 0.6H
9*	124.8, CH	5.09, 0.4H
9*	126.0, CH	4.96, 0.6H
10*	138.0, C	
10*	140.2, C	
10-CH_3	23.0, C	1.60 s, 3H
11*	35.8, CH_2	2.25, 1.59
11*	42.6, CH_2	2.09, 1.83
12*	34.2, CH_2	1.75, 1.37
12*	30.8, CH_2	1.72, 1.29
12a*	45.6, CH	1.57 m, 0.6H
12a*	49.6, CH	1.48 m, 0.4H
1'	150.5, C	
1'-CH_3	19.1, CH_3	1.62, 3H
2'	111.3, CH_2	8a: 4.60 d; 8b: 4.72 s, 2H

* multiple assignments for the same carbon/proton caused by two ring-flipped macrocycle conformers

COSY, NOESY, and HMBC correlations constructed the plane structure of **4.3** (Figure 4.2). Starting from the terminal alkene, protons H-2'a (δ_H 4.60) and H-2'b (δ_H 4.72) both showed correlations to methine carbon C-6 (δ_C 47.6/47.9) and methyl group 10-CH_3 (δ_C 19.1). The methyl protons 1'-CH_3 (δ_H 1.62) connected to the olefinic carbon C-1' (δ_C 150.5) and methine carbon C-4 (δ_C 47.6/47.9). For the six-member ring, three spin systems were constructed with correlations from the ^1H-^1H COSY spectrum between H-4 (δ_H 2.06/1.99) and methine proton H-4a (δ_H 3.55) and methine proton H-12a (δ_H 1.57/1.48), from methylene protons H_2-3 (δ_H 1.54, 1.36) to methylene protons H_2-2 (δ_H

1.74, 1.38), and from methyl protons 1-CH_3 (δ_H 1.13) to H-1 (δ_H 1.69/1.70). All three spin systems were connected using HMBC correlations. The connection between C-3 and C-4 from cross correlations from H-4a (δ_H 3.55) to C-4 (δ_C 47.6/47.9) and C-3 (δ_C 26.9), and the correlations from methyl protons 1-CH_3 (δ_H 1.13) to methylene carbon C-2 (δ_C 27.8) and methine carbon C-12a (δ_C 45.6/49.6) established the full six-member ring. Due to the coupling between the proton at δ_H 4.60 ppm and methyl protons 1'-CH_3 (δ_H 1.62) and the coupling between the proton at δ_H 4.72 ppm and H-4 observed in the NOESY spectrum, the *cis* proton H-2'a was assigned the chemical shift δ_H 4.60 ppm, and the *trans* proton H-2'b the chemical shift δ_H 4.72 ppm.

The structure of the ten-member ring was also established by COSY and HMBC correlations. The ^{1}H-^{1}H coupling between H-4a (δ_H 3.55) to H-5 (δ_H 5.38/5.63) and the HMBC correlation from H-4a (δ_H 3.55) to C-6 (δ_C 131.0/133.5) and C-12 (δ_C 34.2) indicated that the methine carbons C-4a (δ_C 36.8) and C-12a (δ_C 45.6/49.6) were shared in two rings. The HMBC correlation from H-5 (δ_H 5.38/5.63) to C-6 (δ_C 131.0/133.5) supported the connection to the spin system H-7 (δ_H 4.26/4.15) to H-8 (δ_H 2.72, 2.35/2.66, 2.31) to H-9 (δ_H 5.09/4.96); similarly, the HMBC correlation from H-4a (δ_H 3.55) to C-12 (δ_C 34.2) supported the extension to the spin system of H_2-12 (δ_H 1.75, 1.37) to H_2-11 (δ_H 2.25, 1.59). The HMBC correlations from the methyl protons 10-CH_3 (δ_H 1.60) to the quaternary C-6 (δ_C 138.0/142.2), methine C-5 (δ_C 124.8/127.0) and methylene C-11 (δ_C 35.8), finalized the assignment of the ten-membered ring. Two double bonds C-5 (δ_C 148.4/153.7) to C-6 (δ_C 131.0/133.5) and C-9 (δ_C 124.8/127.0) to C-10 (δ_C 138.0/142.2) were assigned due to their down field chemical shifts. Additionally, the carboxylic acid was placed on C-6 (δ_C 131.0/133.5) due to the HMBC correlation

from H-5 (δ_H 5.38/5.63) to the carbonyl carbon at δ_C 170.6. The hydroxyl group was placed on C-7 due to its down field chemical shift (δ_C 47.6/47.9).

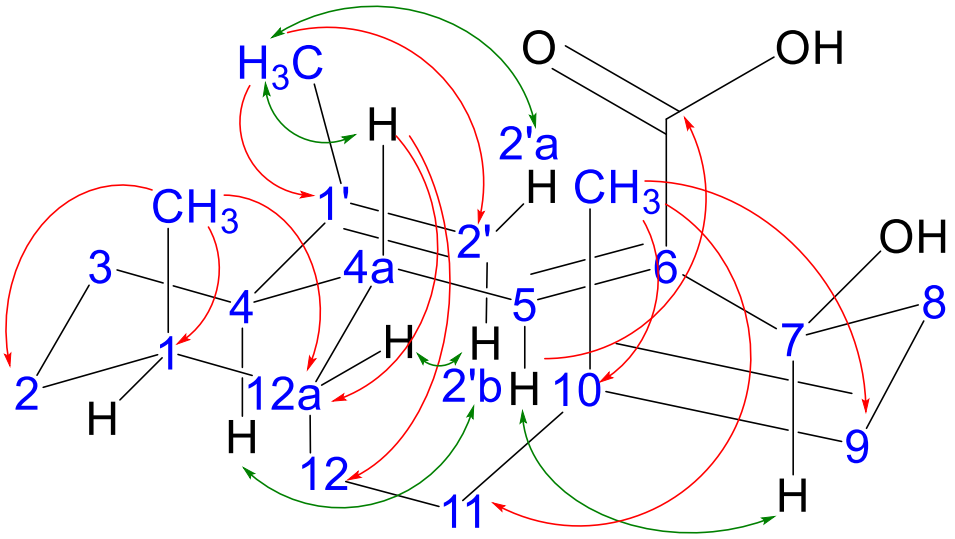
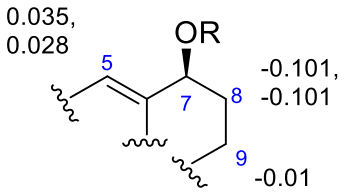


Figure 4.2 Selected COSY (bold line), NOESY (green), and HMBC (red) correlations found in **4.3**

In order to determinate the absolute configuration of diterpenoid **4.3**, with its five asymmetric carbon centers and flexible macrocyclic ring system, Mosher ester analysis [18], NOESY spatial analysis, *J*-based conformational analysis [19] and DFT-based computations were used. The secondary alcohol was accessible to Mosher ester analysis by individually reacting **4.3** with both *R* and *S* α-methoxy-α-trifluoromethylphenylacetic (MTPA) acid chlorides to give the **4.3*S*** and **4.3*R*** MTPA esters, respectively. Comparative analysis of the ^{1}H NMR spectra of both MTPA esters assigned unequivocally C-7 as *S* configured (Figure 4.3 and supporting information Figure C26).

4.3*S* R = (*S*)-MTPA
4.3*R* R = (*R*)-MTPA



Position	δ_H **4.3*S*** (ppm)	δ_H **4.3*R*** (ppm)	$\Delta\delta_{S-R}$ (ppm)
5	5.594, 5.502	5.559, 5.474	0.035, 0.028 > 0
8', 8"	3.011, 2.504	3.112, 2.605	-0.101, -0.101 < 0
9	5.756	5.768	-0.012 < 0

Figure 4.3 $\Delta\delta_{S-R}$ values for MTPA esters of **4.3**

With the absolute configuration of the secondary alcohol defined, NOESY and *J*-based conformational analysis was used to determine the remaining four stereocenters. The multiplicity of H-4a exhibits a doublet of doublets of doublets (ddd) pattern with two large coupling constants (11.00 and 10.30 Hz) and a small one (4.30 Hz), which indicates two trans and one cis relationship with regards to the decalin ring system. Based on the DQF-COSY data extraction, the $^3J_{(H5\text{-}H4a)}$ coupling constant is 10.00 Hz in both conformers which supports the orientation to be trans and accounting for one of the two large coupling constants within the ddd multiplet of H-4a. The NOESY spectrum showed a strong NOE correlation from H-4a to methyl 1-CH_3 placing them cis to one another and counting for the small coupling in the ddd signal for H-4a; while a very weak NOE correlation from H-4a to H-4 allowed to place these signals trans to each other, accounting for the other large coupling constant of the H-4a multiplet (supporting information Figure C23). Additionally, a strong NOE correlation between H-5 and H-7 was found in both conformers in the NOESY spectrum, which set the geometry of the C-5/C-6 double bond as *Z*. With the absolute configuration of C-7 determined to be *S* by Mosher ester analysis, consequently the stereocenters at C-4a, C-12a, and C-4 can be

assigned as *S*, *R*, and *R* configured, based on the mentioned NOEs and the *J*-based analysis above.

IPAP-HMBC NMR data displayed the $^{3}J_{(Me10\text{-}H9)}$ coupling to be approximately 7 Hz in both ring conformers which indicated the geometry of the C-9/C-10 double bond to be *E* configured (supporting information Figure C24). A very small coupling constant $^{4}J_{(C4a\text{-}Me1)}$, approximately 1 Hz, would suggest a cis position between H-4a to CH_3-1, inferring *S* configuration for C-1 (supporting information Figure C25). These values assignments were corroborated by DFT-based calculations of coupling constants at the B3LYP/6-31G* level.

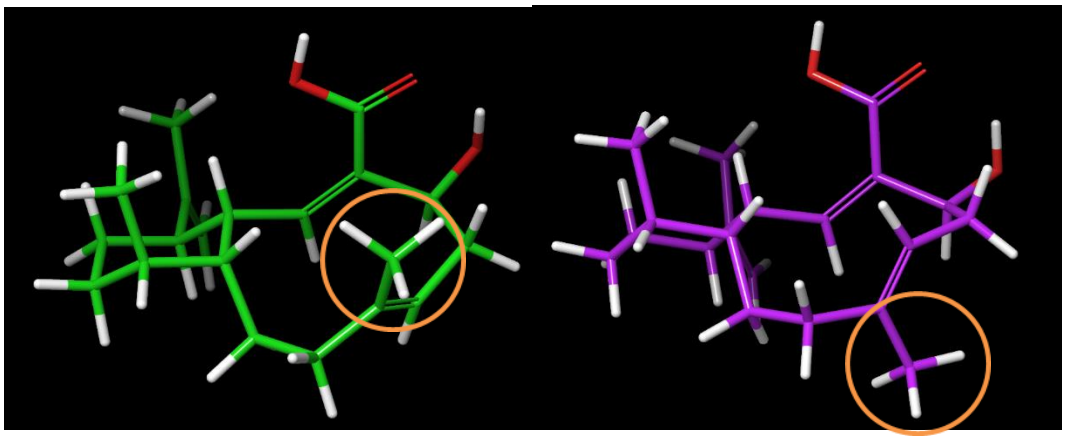


Figure 4.4 The two lowest energy conformers calculated by DFT. Orange circles point to the ring-flipped methyl groups between the two conformers in green color or purple color.

Next, we employed 3D structural computation of conformers to support the assignments of the configuration. Here, the conformational space of diterpenoid **4.3** was calculated using ForceGen applying minimal distance and torsion restraints derived from NOE and *J* values, namely the observed NOE between H-5 and H-7, and between H-4a and 1-CH_3, and the $^{3}J_{HH}$ between H-4a and H-5, to guide accurate conformational sampling [20]. From this conformational ensemble the Boltzmann population was

computed by DFT at the MPW1PW91/cc-PVTZ level. The two lowest energy conformers are shown in Figure 4.4. The macrocycle has two distinct low energy forms (methyl up in green color and methyl down in purple color) with nearly identical energies but with a significant rotational barrier between them. This observation explains the chemical exchange shown in the NOESY/ROESY spectra. The Boltzmann populations are approximately 3:2, which matched the integration ratios observed in the ^{1}H NMR experiment (Table 4.1). Computationally derived UV spectrum and circular dichroism (CD) spectrum at the ωB97XD/def2TZVP//MPW1PW91/cc-PVTZ level compared with experimental spectra further support the stereochemical assignment of diterpenoid **4.3** (supporting information Figure C27). Thus, the absolute configuration of **4.3** was assigned as 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*S*,12a*S*.

The closest related structures we could identify are odyverdine A and B, both produced by heterologous expression in *Streptomyces avermitilis* [10]. Odyverdine A exhibits a 6-8-4 ring system (a fused tricyclo[8.4.0.0^{3,6}] structure), and odyverdine B has a 6-7-5 ring system (a fused tricyclo[8.4.0.0^{2,6}] structure), different from the bicyclo[8.4.0] structure of **4.3.** Their planar structures were determined by NMR.

4.4 Materials and methods

4.4.1 General experimental procedure

Circular dichroism (CD) and UV measurements were recorded using a JASCO J-815 spectropolarimeter. Infrared (IR) spectra were recorded on a Thermo Scientific Nicolet IR100 FTIR spectrometer. NMR spectra were acquired on a Bruker Avance III 700 MHz spectrometer equipped ^{13}C cryoprobe, with the residual solvent used as internal standards: methanol-*d*$_4$ (δ_H 3.31; δ_C 49.00) and chlorform-*d*$_1$ (δ_H 7.26; δ_C 77.06) [21].

Low-resolution ESIMS spectra were recorded on Agilent 1100 series LC with MSD 1946. High-resolution ESIMS spectra were recorded on an Agilent 6545 LC/Q-TOF MS. Dereplication of known natural products was achieved by comparison with of UV maxima and retention times to an internal lab database while commercially available databases including AntiBase [15] and AntiMarin Database [22], were used to compare spectroscopic data. Analytical HPLC was performed using an Agilent 1100 HPLC with a reverse phase Phenomenx kinetex C18 column (5 µm × 150 mm × 4.6 mm) with a standard gradient method was employed with ultrapure H_2O (A) and MeCN (B) with 0.05% formic acid (FA) in each solvent (10% B to 100% B in 35 min) at a flow rate of 0.8 mL/min. Preparative HPLC was performed using an Agilent 1260 HPLC system of ultra-pure H_2O (A) and MeCN (B) with a reverse phase Phenomenex Luna C-18 column (5 µm × 150 mm × 21 mm) with isocratic methods at a flow rate of 20 mL/min. All samples were filtered through a 0.45 µm nylon filter or centrifuged at 14,000 rpm for 5 min before MS and HPLC analysis. Analytical thin layer chromatography (TLC) was performed on a pre-coated silica gel 60 F254 plates (Eppendorf). TLC plates were visualized by UV (254 and 360 nm) and by spraying with anisaldehyde solution followed by heating at 80 °C. General reagents were purchased from Sigma-Aldrich Corp., Fisher Scientific, and VWR International.

4.4.2 Media condition and bacterial isolation

Culture media. M2: malt extract (10 g/L), yeast extract (4 g/L), glucose (4 g/L), pH 7.0 prior to sterilization. R2A (purchased from BD Difco™). Agar plate: addition with 15 g agar per liter, for agar with antifungal drugs, nystatin (25 mg/L) and cycloheximide (50 mg/L) were added.

The bacterium *Streptomyces flaveolus* (CL12-4) was isolated from an arid, high desert soil collected near Bend, Oregon on June 14th, 2014 (43°59'14"N, 121°20'11"W). The soil sample was placed in a sterile 50 mL plastic Whirl-Pak bag, and was kept cool until processing within 72 h.

After drying overnight in a laminar-flow bio-cabinet, 1.0 g of soil was transferred into a 10 mL of autoclaved ultrapure water. 0.1 mL of dilution #3, #4, #5, and #6 (dilution of 1: 1,000, 1: 10,000, 1: 100,000, and 1:1,000,000) was inoculated onto high and low nutrient media agar plates (M2 and R2A) and was spread with a sterile loop. *Streptomyces flaveolus* CL12-4 emerged from dilution #4. It possessed morphological features, including production of white, filamentous hyphae and earthy smell, often characteristics for the genus Streptomyces. Clean streaked on M2 agar plate and incubated at 28 °C for 5-7 days, pure bacterial colonies were observed by eye. One piece of 1 cm^2 agar was used to inoculate 50 mL of M2 liquid broth as seed culture for DNA extraction, glycerol cryopreservation, and large-scale cultivation.

4.4.3 Bacterial species identification

Phylogenetic identification of the bacterium was achieved by PCR amplification and sequencing of the 16S rRNA gene sequence. *Streptomyces flaveolus* CL12-4 was grown in liquid culture for 5 days, and then cells were collected by centrifugation at 10,000 rpm for 10 min. The cell pellet was frozen and crushed with liquid nitrogen using a micro pellet pestle. The crushed cells were incubated with lysozyme at 37°C for 30 minutes and then incubated with 3 µl of proteinase K and 10% (w/v) sodium dodedyl sulfate (SDS) at 55°C for 30 min. DNA was extracted in a phenol-chloroform solution (v/v 1:1) by centrifugation at 12,000 rpm for 5 minutes. The 16S rRNA gene was amplified from

genomic DNA using PCR amplification with the primer: 8f (5'-AGAGTTTGATCMTGGCTCAG) and 1513r (5'- ACGGCTACCTTGTTACGACTT) [13] (both purchased from Integrated DNA Technologies, IDT). The 50 μL PCR mixture contained 20 to 40 ng of DNA template, 10 pmol of each primer, 25 μL of MangoMix™ (purchased from Bioline, UK) containing Mango*Taq*™ DNA polymerase, dNTPs, red and orange reference dyes and Mg^{2+}. The PCR program consisted of 30 cycles at 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 90 seconds followed by a final extension step at 72°C for 10 minutes. Amplified DNA products were examined and purified by agarose gel electrophoresis using the PureLink™ quick gel extraction kit (purchased from Invitrogen by life technologies, Germany) according to the manufacturer's suggested protocol. PCR products were sequenced on an ABI Prism 3730 genetic analyzer (Applied Biosystems) using a Big Dye Terminator v 3.1 cycle sequencing kit at OSU's CGRB Core Facilities. The nucleotide sequences were analyzed and compared to sequences within the NCBI database (http://www.ncbi.nlm.nih.gov/) using the Basic Local Alignment Search Tool (BLAST) or EzBiocloud 16S database [14]. The phylogenetic analysis was generated by MEGA7 by applying Neighbor-joining tree [23-25].

4.4.4 Extraction and isolation

Streptomyces flaveolus CL12-4 was inoculated into 4.5 L of M2 broth from 50 mL seed culture. The culture grew constant shaking at 110 rpm for 4 weeks at 28 ˚C. Cultures were clean streaked on M2 agar plates prior to extraction to test for culture purity. The culture was extracted with equal parts ethyl acetate three times after adjusting pH from

6.5-7.0 in original culture to 5.0 with 1 N HCl. The organic phase was then died over $MgSO_4$ and concentrated under reduce pressure.

The organic extract (approximately 420 mg) was separated into seven fractions (Fraction 1 to Fraction 7) by normal-phase vacuum-liquid chromatography (VLC) eluted with a dichloromethane (DCM) to methanol gradient (v/v 99:1, 30:1, 15:1, 9:1, 3:1, 1:1, 0:1). Altogether, 5.0 mg of coproporphyrin III (**4.1**) and 2.4 mg of zincphyrin (**4.2**) were isolated and purified from Fraction 6 and Fraction 7 via preparative HPLC separation using an isocratic elution of 46% MeCN resulting in elution at 8 min and 16 min, respectively. Diterpenoid (**4.3**) was isolated and purified from Fraction 4 with isocratic elution of 65% MeCN via preparative HPLC system.

4.4.5 Cytotoxicity assay

Cytotoxic activities of organic extracts were evaluated against five cancer cell lines [26, 27]: human colorectal carcinoma (HCT-116, ATCC CCL-247), human breast adenocarcinoma (MCF-7, ATCC HTB-22), melanoma (SK-Mel-5, ATCC HTB70), prostate cancer (PC3, ATCC CRL-1435), and lung cancer (A-549, ATCC CCL-185). HCT-116 and SK-Mel-5 cells were maintained in MEM growth media, A549, MCF-7 and PC-3 in RPMI 1640 media, each supplemented with 10% (v/v) fetal bovine serum, penicillin (100 Units/mL) and streptomycin (100 µg/mL). The cell lines were maintained in a humidified chamber at 37 °C in 5% CO_2.

Cell viability was determined by measuring the reduction of the tetrazolium salt MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) by metabolically active cells. Cells were plated into 96-well plates (HCT-116: 7,000 cells/well; SK-Mel-5: 5,000 cells/well; A549: 7,000 cells/well; PC-3: 5,000 cells/well; MCF-7: 9,000 cells/well) and

maintained overnight before treatment. Organic extracts and fractions from bacterial cultures were prepared at 10 mg/mL in DMSO for single dose treatment giving a final concentration of 10 µg/ml in each well. After 48 h, MTT reagent (5 mg/ml in PBS) was added to each well at a final concentration of 0.5 mg/ml. The plates were incubated for 2 h at 37°C. The growth media was removed, and then purple formazan product solubilized by the addition of 50 µL DMSO. Absorbance was measured at 550 nm using a Biotek Synergy 96-well plate reader. Metabolic activity of vehicle-treated cells (0.1% DMSO) was defined as 100% cell growth. Etoposide (250 µM) was used as a positive control.

4.4.6 Antibacterial assay

All extracts were tested for antibacterial activity in cell-based assays following established protocols [28]. The antibacterial activity was evaluated against two Gram-positive pathogens including *Enterococcus faecium* (ATCC 49032) and methicillin-resistant *Staphylococcus aureus* (ATCC BAA-41) as well as two Gram-negative pathogens: *Pseudomonas aeruginosa* (ATCC 15442) and *Escherichia coli* (ATCC 8739). Vancomycin and chloramphenicol were used as positive controls in Gram-positive and Gram-negative pathogens, respectively. Tested extracts and antibiotic controls were dosed at 125 µg/mL. Bacterial growth rates were measured by absorbance at 620 nm using Bioteck Synergy 96-well plate reader. All human pathogens used in the study were acquired from the American Type Culture Collection (ATCC, Manassas, VA, USA).

4.4.7 Mosher ester analysis

Mosher ester analysis [18] was used to identify the stereochemistry of the secondary alcohol of diterpenoid **4.3**. 0.64 mg of **4.3** (2.0 µmol) was dissolved in 100 µL of DCM. To this was added anhydrous pyridine (3.0 µL, 37.4 µmol) dissolved in 100 µL of dry

DCM, followed by (*S*)-(+)-α-methoxy-α-trifluoromethylphenylacetyl chloride (*S*-(+)-MTPA-Cl), dissolved in 100 µL of dry DCM. The reaction was monitored by TLC. After 24 h, the reaction was quenched with H_2O, and then extracted with ethyl acetate. The organic layers were combined, dried with anhydrous $MgSO_4$, and concentrated to dryness under reduced pressure. The desired product, the (R)-MTPA ester of compound 4.3, purified by preparative HPLC, and the its structure was determined by NMR spectroscopic analysis. In an analogous manner, compound 4.3 (0.64 mg, 2.0 µmol) was converted with (*R*)-(-)-α-methoxy-α-trifluoromethylphenylacetyl chloride (*R*-(-)-MTPA-Cl) to yield the (*S*)-MTPA ester. The absolute configuration of the secondary alcohol can be solved by comparison of ^{1}H-NMR chemical shifts.

4.4.8 Conformational modeling for 4.3

The relative configuration of the cyclo-hexyl ring and double-bond geometry was set as described in section 4.3 based on *J*-based configurational analysis and NOE correlations. Two absolute configurations, 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*E*,12a*S* or 1*R*,4*S*,4a*R*,5*Z*,7*S*,9*E*,12a*R*, exhibiting the observed relative configuration about the cyclohexyl portion of the molecule, and the absolute configuration at position 7 determined by Mosher ester analysis, were subjected to conformer generation in ForceGen with two NMR derived distance restraints and one NMR derived torsional restraint set to guide the conformational search. The NOE correlation of diastereotopic methylene protons H-12’ and H-12’’ were used as a reference distance for the calculation of NOE distance restraints according to $\sigma_{IS} = kr_{IS}^{-6}$ [29]. Accounting for errors in distance due to relaxation and error of measurement, a maximum distance of 2.8 Å was set for H-5 to H-7 and 3.6 Å for Me-1 to H-4a. Based on the 10 Hz coupling constant between H-4a

and H-5, a torsional restraint in the range of 150°-210° was also used to restrict the conformational search. Of the two configurations calculated, only the 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*E*,12a*S* configuration produced conformers with no restraint violations (649 of 963 total conformers generated), while all conformers (545 of 545) for 1*R*,4*S*,4a*R*,5*Z*,7*S*,9*E*,12a*R* exhibited greater than 0.5 Å distance violations. Both configurations were carried forward for further DFT computations.

4.4.9 Computational details for 4.3

Conformers from the ForceGen conformational search described above were carried forward for geometry optimization, NMR coupling constant calculations, as well as TDDFT calculation of chiroptical spectra. A preliminary geometry optimization and frequency calculation at the B3LYP/6-31+G* level to eliminate conformations exhibiting negative frequencies and relative energies greater than 10 kcal/mol was initiated. 1*R*,4*S*,4a*R*,5*Z*,7*S*,9*E*,12a*R* produced a single conformer accounting for 83% of the Boltzmann population with all additional low energy conformations varying only in minor torsional flexions or methyl rotations. 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*E*,12a*S* exhibited 12 low energy conformers which exhibited a macrocyclic flip accounting for the chemical exchange observed in the ROESY spectra. Geometries for these 12 conformers were optimized and Boltzmann populations were calculated at the MPW1PW91/cc-PVTZ level with the two lowest energy conformers exhibiting relative energies of 0.0 and 0.06 kcal/mol with a ‘methyl up’ and ‘methyl down’ macrocyclic conformation, respectively. NMR coupling constants were calculated at the B3LYP/6-31G* level. NOE intensities, and available experiment nJXH coupling constants were all in good agreement with the conformational ensemble generated. TDDFT was performed at the ωB97XD/def2TZVP

level with SMD implicit solvent modeling in acetonitrile. UV and ECD spectra were extracted utilizing SpecDis [30] and compared with experimental spectra and were in good agreement.

4.5 Conclusion

In summary, three metabolites coproporphyrin III (**4.1**), zincphyrin (**4.2**), and a new, cytotoxic diterpenoid **(4.3)** were isolated from *Streptomyces flaveolus* (CL12-4). The diterpenoid featured five stereocenters and two double bond geometries in a fused bicyclo[8.4.0] tetradecane macrocycle. Noteworthy, the diterpenoid existed in two distinct ring-flipped conformers in solution. The absolute configuration of **4.3** was assesses using multiple spectroscopic techniques, supported by DFT-based 3D space modeling and computational ECD spectra to be 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*S*,12a*S*. **4.3** exhibited ability to strongly inhibit the growth of human colon cancer cell line (HCT-116) at 10 µg/mL by 76%, comparable to the positive control used in the assays, etoposide (10 µg/mL). Currently, IC_{50} values for the new diterpenoid are being tested against multiple cancer cell lines.

4.6 Acknowledgements

This work was supported by OSU start-up funds (SL). We wish to thank Dr. Birte Plitzko for assisting with cell viability assays and thank Cassandra Lew for support in extracting cultures.

4.7 References

1. Lange, B.M., *The Evolution of Plant Secretory Structures and Emergence of Terpenoid Chemical Diversity.* Annual Review of Plant Biology, 2015. **66**(1): p. 139-159.
2. Dickschat, J.S., *Bacterial terpene cyclases.* Nat Prod Rep, 2016. **33**(1): p. 87-110.
3. Walsh, C.T., Yi,, *Natural product biosynthesis (1st).* Royal Society of Chemistry, Cambridge, 2017.
4. Gerber, N.N., *Geosmin, an earthy-smelling substance isolated from actinomycetes.* Biotechnology and Bioengineering, 1967. **9**(3): p. 321-327.
5. Medsker, L.L., et al., *Odorous compounds in natural waters. 2-exo-Hydroxy-2-methylbornane, the major odorous compound produced by several actinomycetes.* Environmental Science & Technology, 1969. **3**(5): p. 476-477.
6. Citron, C.A., et al., *Terpenoids are widespread in actinomycetes: a correlation of secondary metabolism and genome data.* Chembiochem, 2012. **13**(2): p. 202-14.
7. Yamada, Y., D.E. Cane, and H. Ikeda, *Chapter Seven - Diversity and Analysis of Bacterial Terpene Synthases*, in *Methods in Enzymology*, D.A. Hopwood, Editor. 2012, Academic Press. p. 123-162.
8. Yamada, Y., et al., *Terpene synthases are widely distributed in bacteria.* Proc Natl Acad Sci U S A, 2015. **112**(3): p. 857-62.
9. Hanson, J.R., *Diterpenoids of terrestrial origin.* Nat Prod Rep, 2017. **34**(10): p. 1233-1243.
10. Yamada, Y., et al., *Novel terpenes generated by heterologous expression of bacterial terpene synthase genes in an engineered Streptomyces host.* J Antibiot (Tokyo), 2015. **68**(6): p. 385-94.
11. Thoppil, R.J. and A. Bishayee, *Terpenoids as potential chemopreventive and therapeutic agents in liver cancer.* World journal of hepatology, 2011. **3**(9): p. 228-249.
12. Huang, M., et al., *Terpenoids: natural products for cancer therapy.* Expert Opinion on Investigational Drugs, 2012. **21**(12): p. 1801-1818.
13. Fischer, J.P. and H. Cypionka, *Analysis of aerotactic band formation by Desulfovibrio desulfuricans in a stopped-flow diffusion chamber.* FEMS Microbiology Ecology, 2006. **55**(2): p. 186-194.
14. Kim, O.S., et al., *Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species.* Int J Syst Evol Microbiol, 2012. **62**(Pt 3): p. 716-21.
15. Laatsch, H., *AntiBase 2014: The Natural Compound Identifier.* 2014.
16. Guo, H., et al., *Natural products and morphogenic activity of gamma-Proteobacteria associated with the marine hydroid polyp Hydractinia echinata.* Bioorg Med Chem, 2017. **25**(22): p. 6088-6097.
17. Toriya, M., et al., *Zincphyrin, a novel coproporphyrin III with zinc from Streptomyces sp.* J Antibiot (Tokyo), 1993. **46**(1): p. 196-200.
18. Hoye, T.R., C.S. Jeffrey, and F. Shao, *Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons.* Nat Protoc, 2007. **2**(10): p. 2451-8.

19. Matsumori, N., et al., *Stereochemical Determination of Acyclic Structures Based on Carbon−Proton Spin-Coupling Constants. A Method of Configuration Analysis for Natural Products.* The Journal of Organic Chemistry, 1999. **64**(3): p. 866-876.
20. Jain, A.N., et al., *Complex macrocycle exploration: parallel, heuristic, and constraint-based conformer generation using ForceGen.* J Comput Aided Mol Des, 2019. **33**(6): p. 531-558.
21. Babij, N.R., et al., *NMR Chemical Shifts of Trace Impurities: Industrially Preferred Solvents Used in Process and Green Chemistry.* Organic Process Research & Development, 2016. **20**(3): p. 661-667.
22. J.W. Blunt, M.H.G.M., H. Laatsch, *AntiMarin Database.* AntiMarin Database, 2006.
23. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.* Mol Biol Evol, 2016. **33**(7): p. 1870-4.
24. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Mol Biol Evol, 1987. **4**(4): p. 406-25.
25. Tamura, K., M. Nei, and S. Kumar, *Prospects for inferring very large phylogenies by using the neighbor-joining method.* Proc Natl Acad Sci U S A, 2004. **101**(30): p. 11030-5.
26. Kmail, A., et al., *In vitro Assessments of Cytotoxic and Cytostatic Effects of Asparagus aphyllus, Crataegus aronia, and Ephedra alata in Monocultures and Co-Cultures of Hepg2 and THP-1-Derived Macrophages.* Pharmacognosy Communications, 2015. **5**(3): p. 165-172.
27. Mosmann, T., *Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays.* J Immunol Methods, 1983. **65**(1-2): p. 55-63.
28. Wiegand, I., K. Hilpert, and R.E. Hancock, *Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances.* Nat Protoc, 2008. **3**(2): p. 163-75.
29. Jones, C.R., C.P. Butts, and J.N. Harvey, *Accuracy in determining interproton distances using Nuclear Overhauser Effect data from a flexible molecule.* Beilstein journal of organic chemistry, 2011. **7**: p. 145-150.
30. Bruhn, T., et al., *SpecDis: quantifying the comparison of calculated and experimental electronic circular dichroism spectra.* Chirality, 2013. **25**(4): p. 243-9.

Chapter Five: General Conclusions

Chenxi Zhu

5.1 General Conclusion

According to the CDC's 2019 antibiotic resistance threats report, more than 2.8 million antibiotic-resistant infections occur in the U.S. each year, and more than 35,000 people die as a result [1]. Obviously, discovery of new drugs is essential and necessary. Many of the FDA proved antibiotics are produced by Gram-positive actinobacteria [2, 3] and they have been utilized in clinic treatments over the last 50 years, such as kanamycin [4, 5], vancomycin [6], and chloramphenicol [7]. The continued discovery of effective natural products from microbes, especially antibiotics or anticancer agents, is feasible – we have just scratched the surface! Recent work on bacterial genomes displayed an abundance of genes responsible for unseen, small molecule diversity and activation of silent biosynthesis gene cluster in actinomycetes has yielded new, bioactive metabolites [8-11]. Another approach is to investigate the metabolites of newly-isolated bacteria from unexplored areas and various microbiomes [12, 13], which is the focus in my research.

From 2014-2018, I was responsible for bacteria isolation from diverse sources in order to create a bacterial strain collection for the Loesgen lab. As described in Chapter two, different media conditions and isolation procedures were utilized for collected samples from different resources. Protocols were established that are still in used for bacterial isolation in the lab. In total, 400 bacterial strains, of which 230 were isolated from Oregonian soils, the Arctic Ocean, and Pacific fish, were screened. I performed cell viability assays with mammalian cell culture and 30 strains out of 194 inhibited cancer cell growths. Five metabolites were de-replicated by UV and MS

data and most importantly, the screening laid the foundation for projects that yielded new metabolites described in later Chapters.

Chapter three and Chapter four were focused on isolation of new metabolites guided by chemical features or bioactivity. In total, seventeen metabolites were isolated and characterized, including three new natural products: a new tetrapeptide (**3.2**), 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**), and a cyclic diterpenoid (**4.3**). A combination of various spectroscopic techniques was used in the structure elucidation of all three compounds. Advanced Marfey's analysis and DFT-based ECD computations assisted in the assignment of the *L*-Val-*D*-Leu configuration of **3.2**. X-ray crystallographic analysis confirmed the chromenone structure of **3.11**. The diterpenoid (**4.3**) contained five chiral centers and two double bond geometries within a fused bicyclo[8.4.0] tetradecane macrocycle. The terpene existed in two ring-flipped conformers in solution. Techniques, involving Mosher ester analysis, *J*-based coupling analysis, and DFT-based 3D space calculation and ECD computation, were used to in assignment of the absolute configuration as **4.3** as 1*S*,4*R*,4a*S*,5*Z*,7*S*,9*E*,12a*R*. In single-dose cell viabilities assays, **4.3** strongly inhibited colon carcinoma, comparable to the positive control etoposide used in the assay. Future work will expand the testing to our in-house cancer cell line panel.

In summary, my work used traditional approaches of prospecting bacteria for the discovery of bioactive metabolites, new and known. Bacterial strain isolation efforts focused on the isolation of chemically rich actinobacteria, but also expanded into marine microbiomes where Gram negative bacteria were mostly found. In cell growths inhibition assays, actinobacteria derived extracts dominated the panel with

their activity (63% of active strains, 10% of all tested strains). Thus, exploring the metabolites of the actinobacteria in culture collections, combined with recent efforts of gene activation, remains a promising route to identify bacterial producers of known and new metabolites.

5.2 References

1. CDC., *Antibiotic Resistance Threats in the United States, 2019.* Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2019., 2019.
2. Genilloud, O., *Actinomycetes: still a source of novel antibiotics.* Natural Product Reports, 2017. **34**(10): p. 1203-1232.
3. Mahajan, G.B. and L. Balachandran, *Antibacterial agents from actinomycetes - a review.* Front Biosci (Elite Ed), 2012. **4**: p. 240-53.
4. Wright, G.D., A.M. Berghuis, and S. Mobashery, *Aminoglycoside Antibiotics*, in *Resolving the Antibiotic Paradox: Progress in Understanding Drug Resistance and Development of New Antibiotics*, B.P. Rosen and S. Mobashery, Editors. 1998, Springer US: Boston, MA. p. 27-69.
5. Becker, B. and M.A. Cooper, *Aminoglycoside antibiotics in the 21st century.* ACS Chem Biol, 2013. **8**(1): p. 105-15.
6. Connor, H., *Drug Discovery—A History.* Journal of the Royal Society of Medicine, 2005. **98**(11): p. 517-518.
7. Sills, M.R. and D. Boenning, *Chloramphenicol.* Pediatrics in Review, 1999. **20**(10): p. 357-358.
8. Ochi, K., *Insights into microbial cryptic gene activation and strain improvement: principle, application and technical aspects.* J Antibiot (Tokyo), 2017. **70**(1): p. 25-40.
9. Ochi, K. and T. Hosaka, *New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters.* Appl Microbiol Biotechnol, 2013. **97**(1): p. 87-98.
10. Mao, D., et al., *Recent advances in activating silent biosynthetic gene clusters in bacteria.* Current opinion in microbiology, 2018. **45**: p. 156-163.
11. Scherlach, K. and C. Hertweck, *Triggering cryptic natural product biosynthesis in microorganisms.* Organic & Biomolecular Chemistry, 2009. **7**(9): p. 1753-1760.
12. Jensen, P.R., et al., *Culturable marine actinomycete diversity from tropical Pacific Ocean sediments.* Environ Microbiol, 2005. **7**(7): p. 1039-48.
13. Charlop-Powers, Z., et al., *Global biogeographic sampling of bacterial secondary metabolism.* Elife, 2015. **4**: p. e05048.

Appendix Section

Appendix A: Supporting Information for Chapter Two

Figure A1. Map of the three sampling locations for cruise CAGE 16-5: Prins Karls Forland, Storfjordrenna and the CraterArea in Bjørnøyrenna.

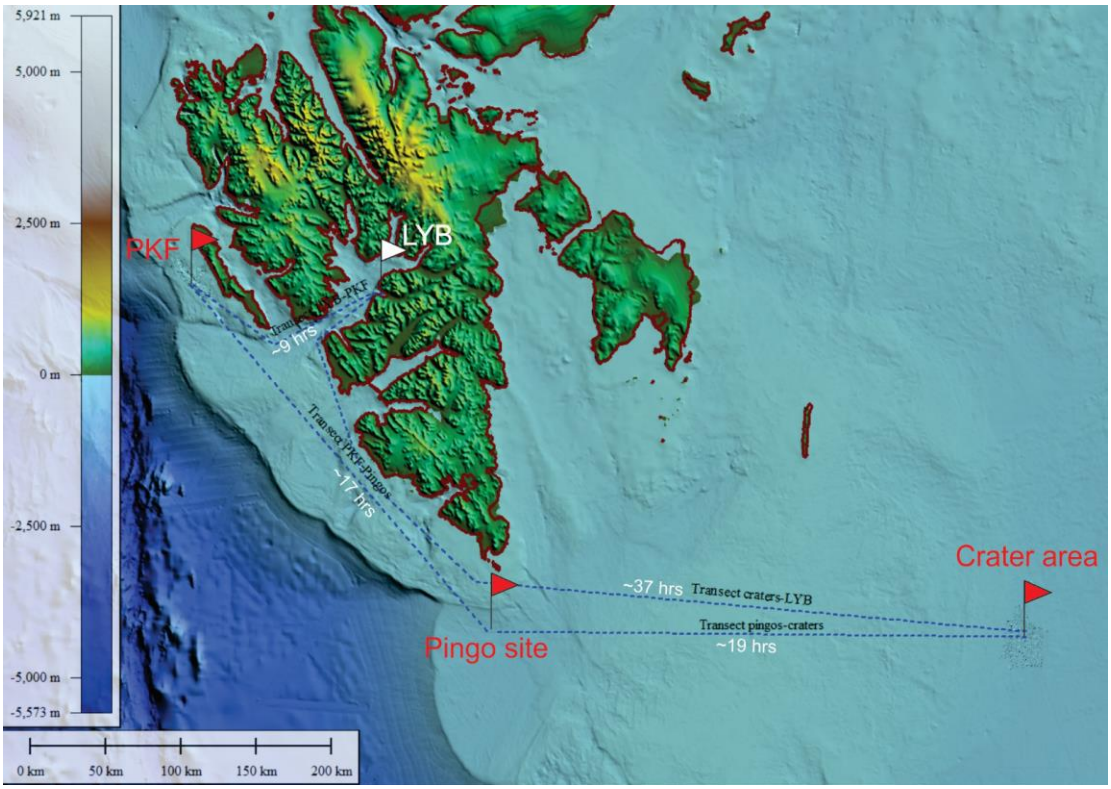


Figure A2. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of chartreusin **2.1**

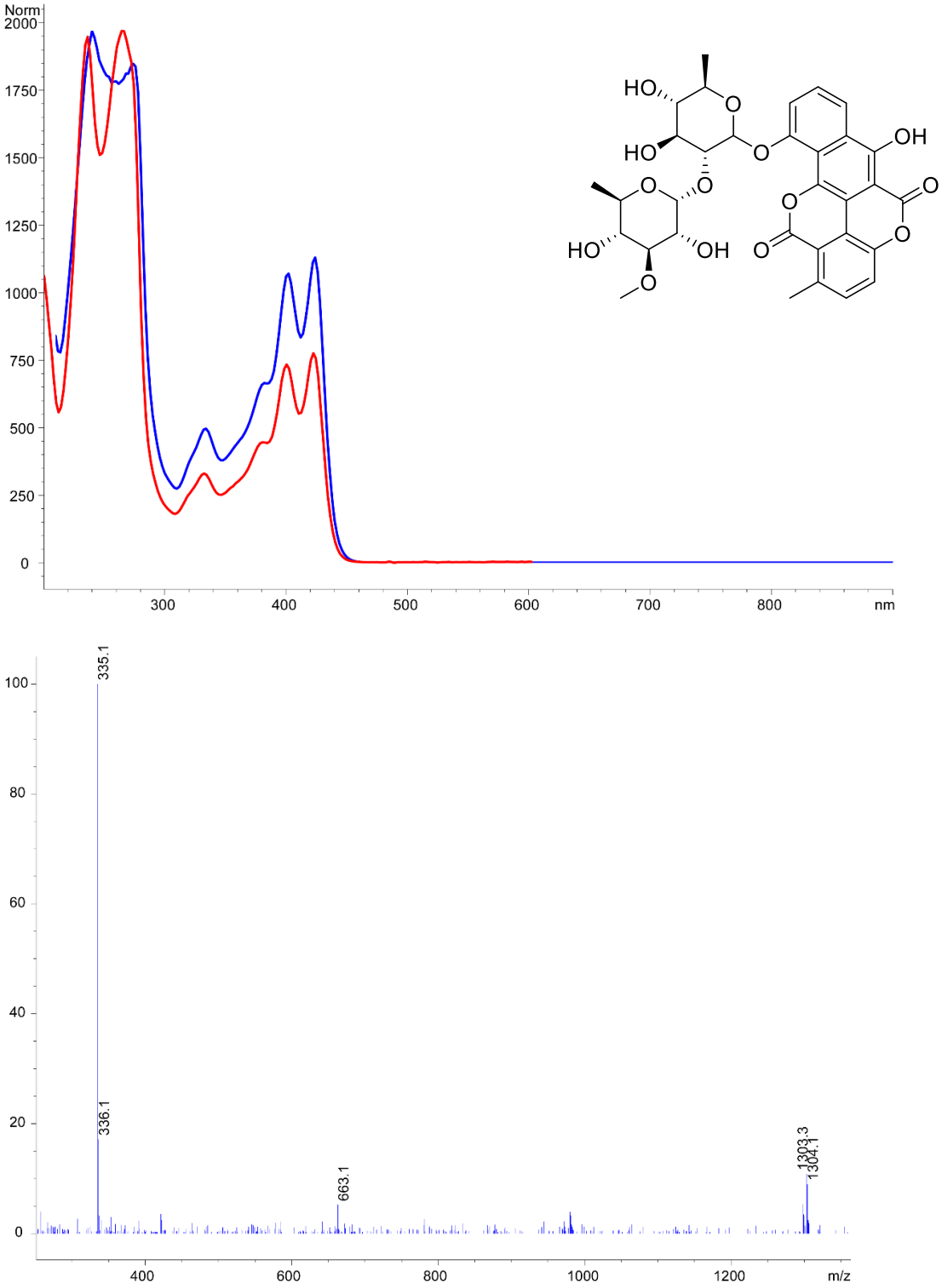


Figure A3. ^{1}H NMR spectrum of chartreusin **2.1**

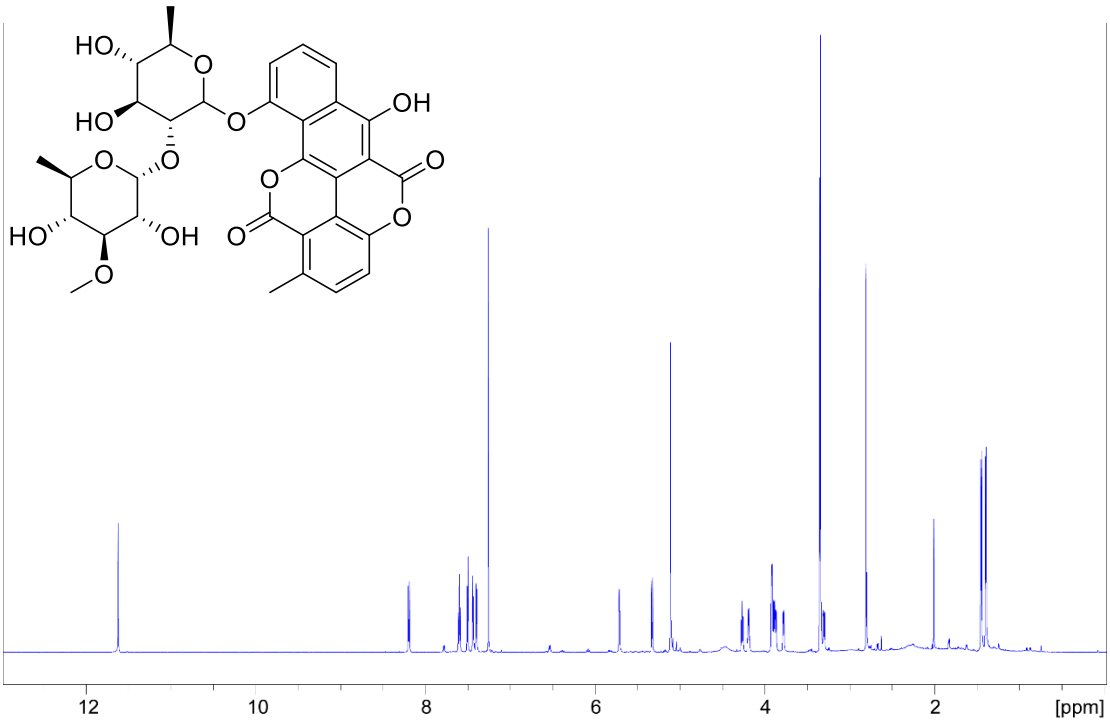


Figure A4. ^{13}C NMR spectrum of chartreusin **2.1**

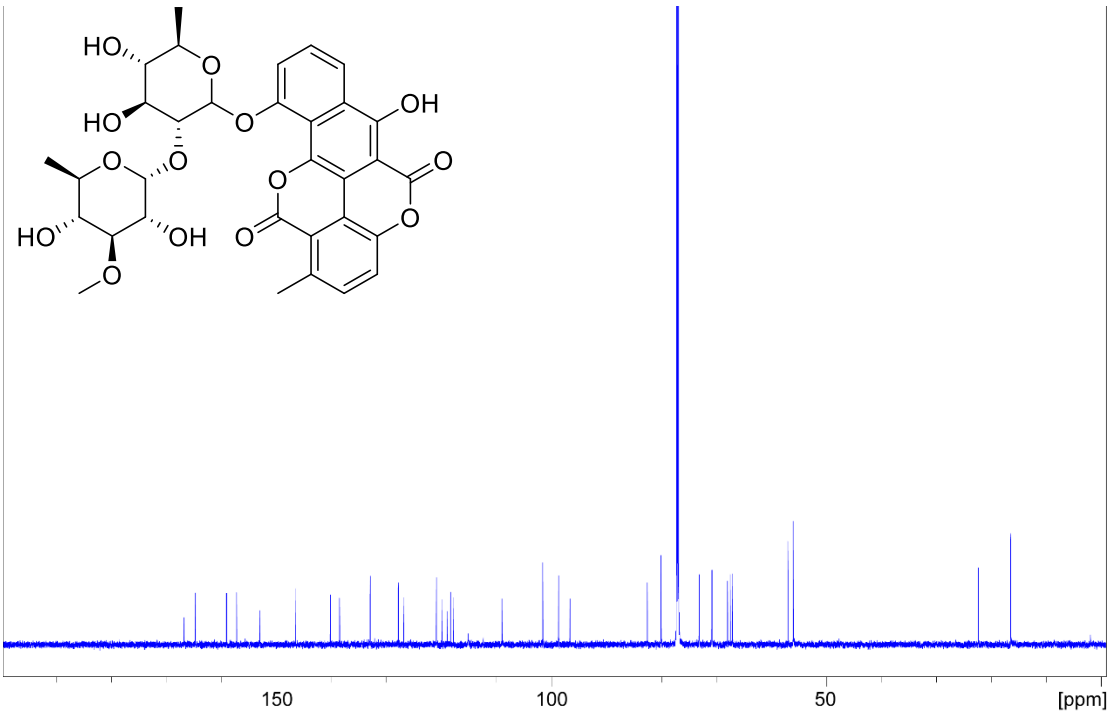
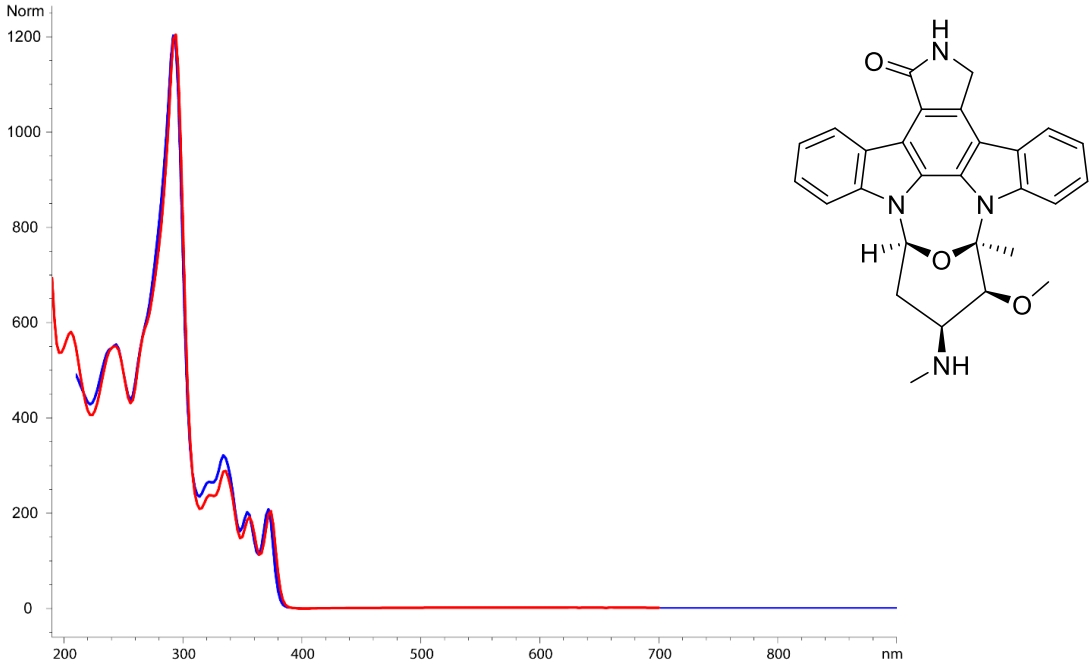


Figure A5. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of staurosporine **2.2**



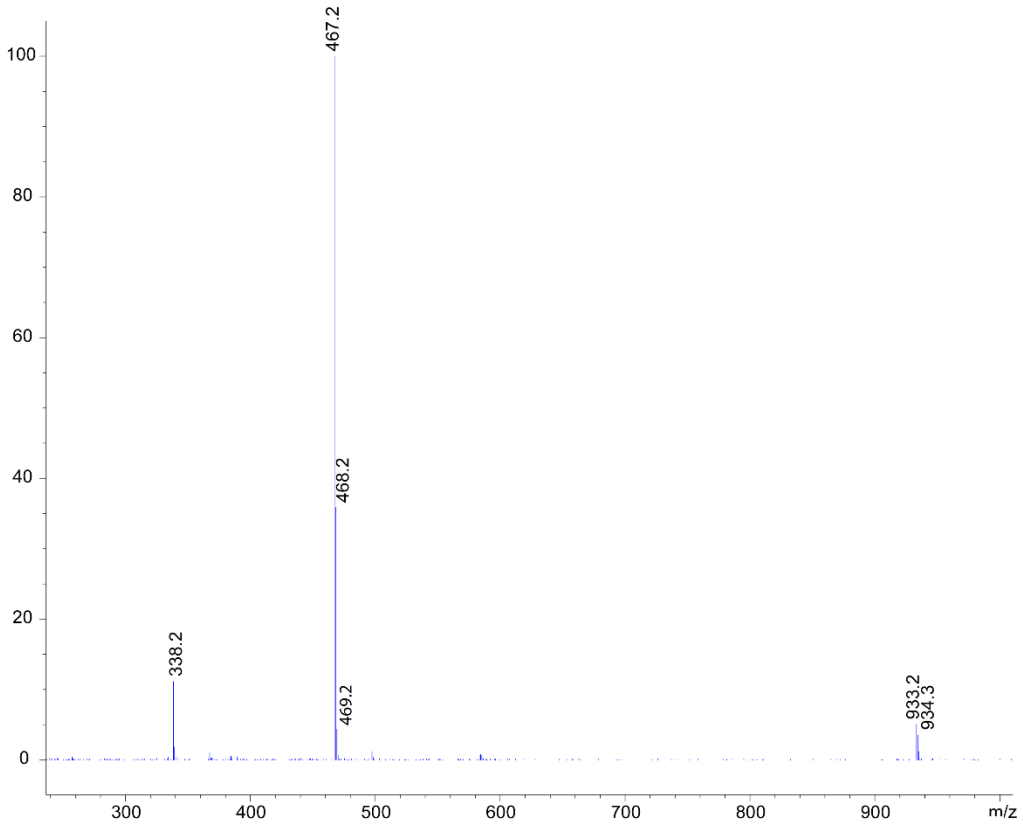
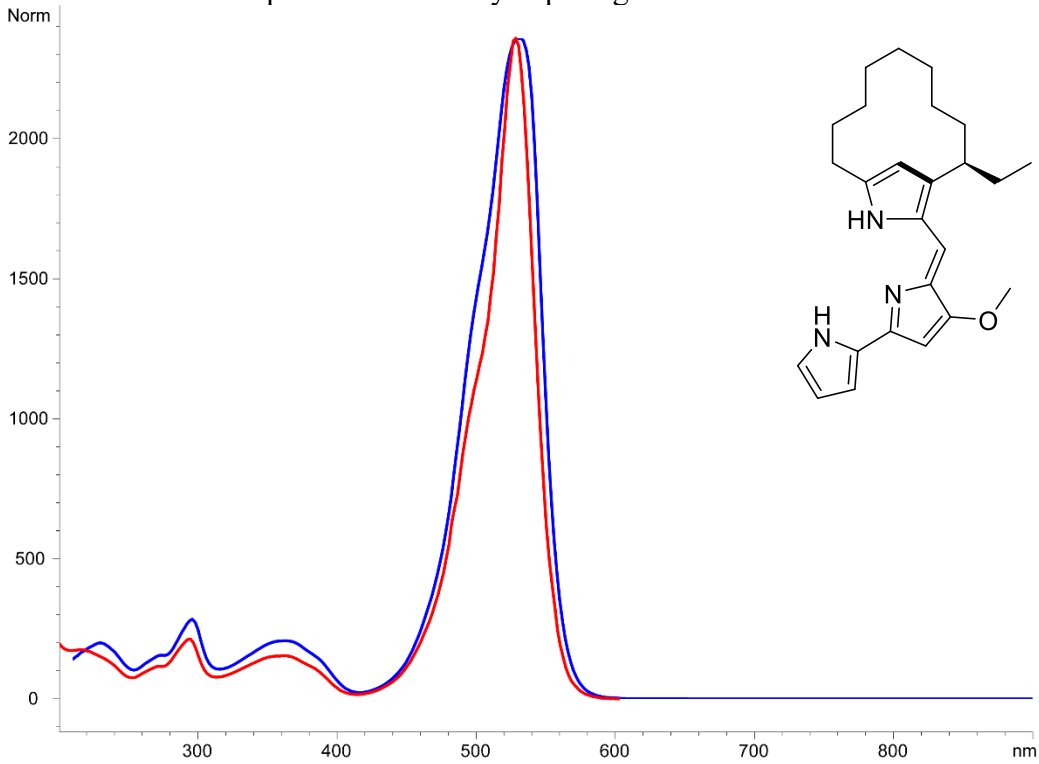


Figure A6. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of metacycloprodigiosin **2.3**



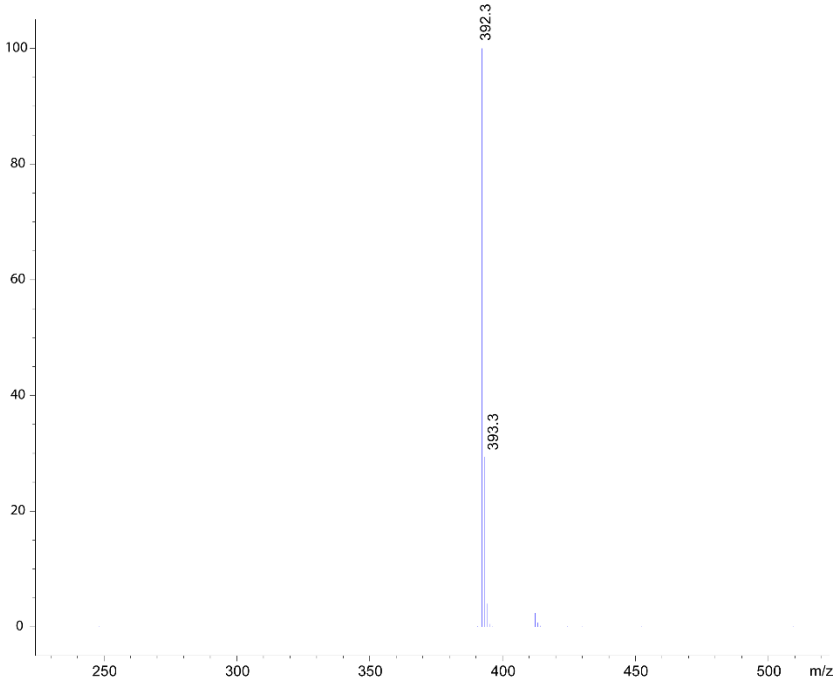


Figure A7. ^{1}H NMR spectrum of metacycloprodigiosin **2.3**

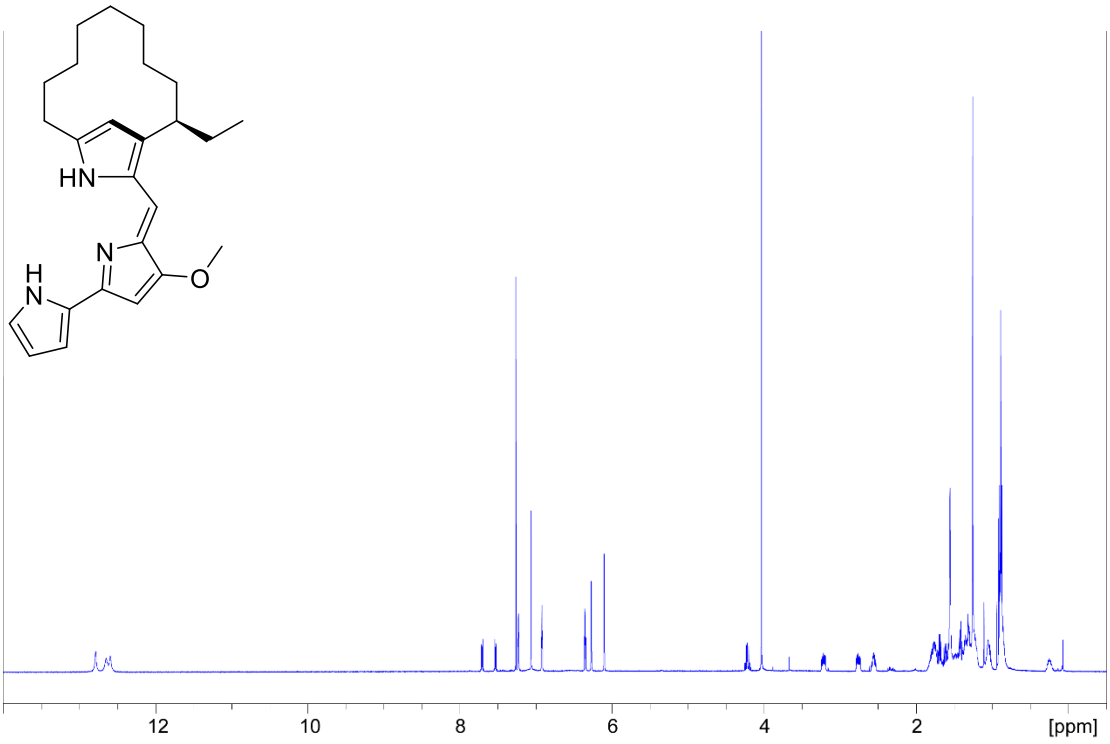


Figure A8. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of undecylprodigiosin **2.4**

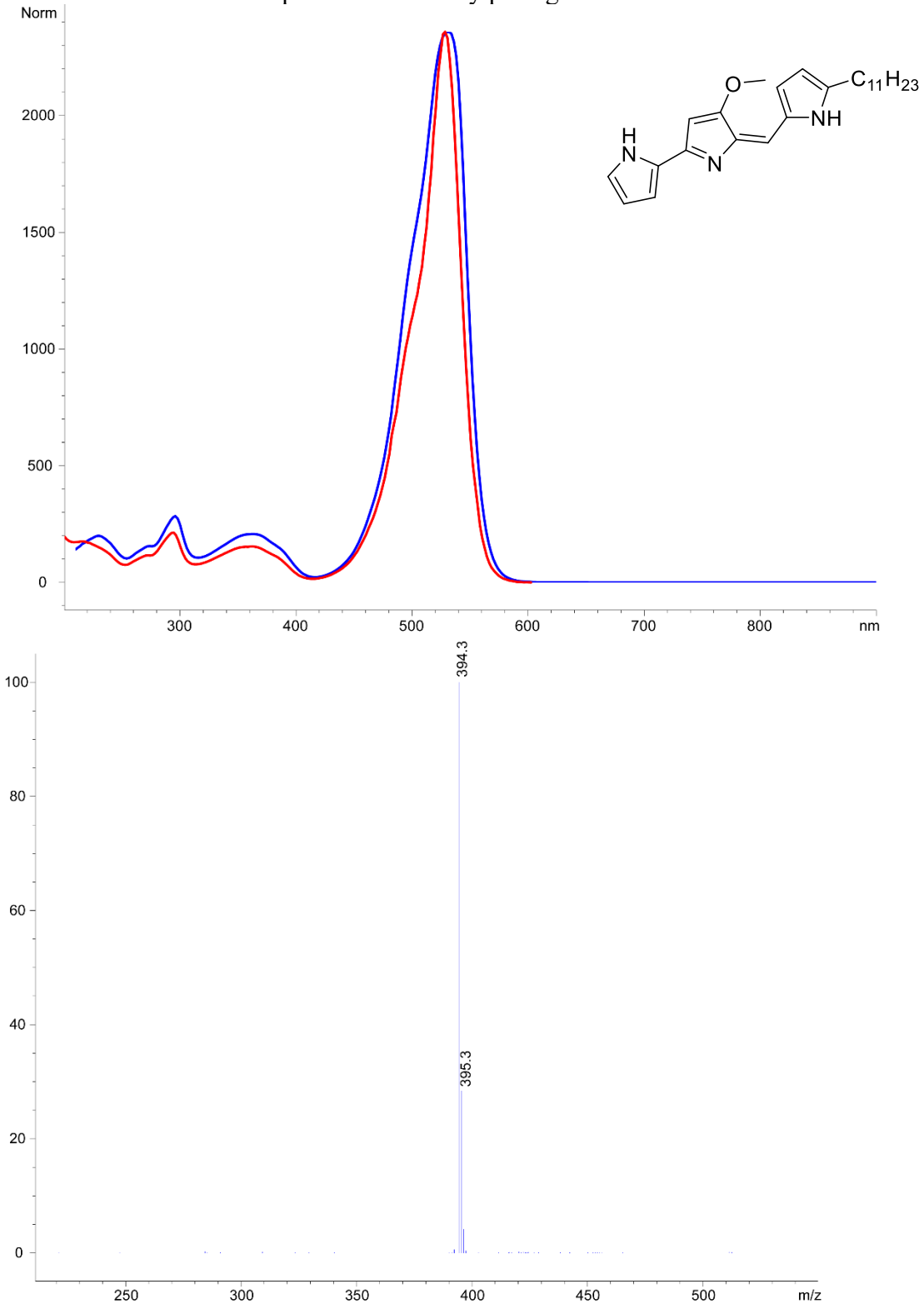


Figure A9. ^{1}H NMR spectrum of undecylprodigiosin **2.4**

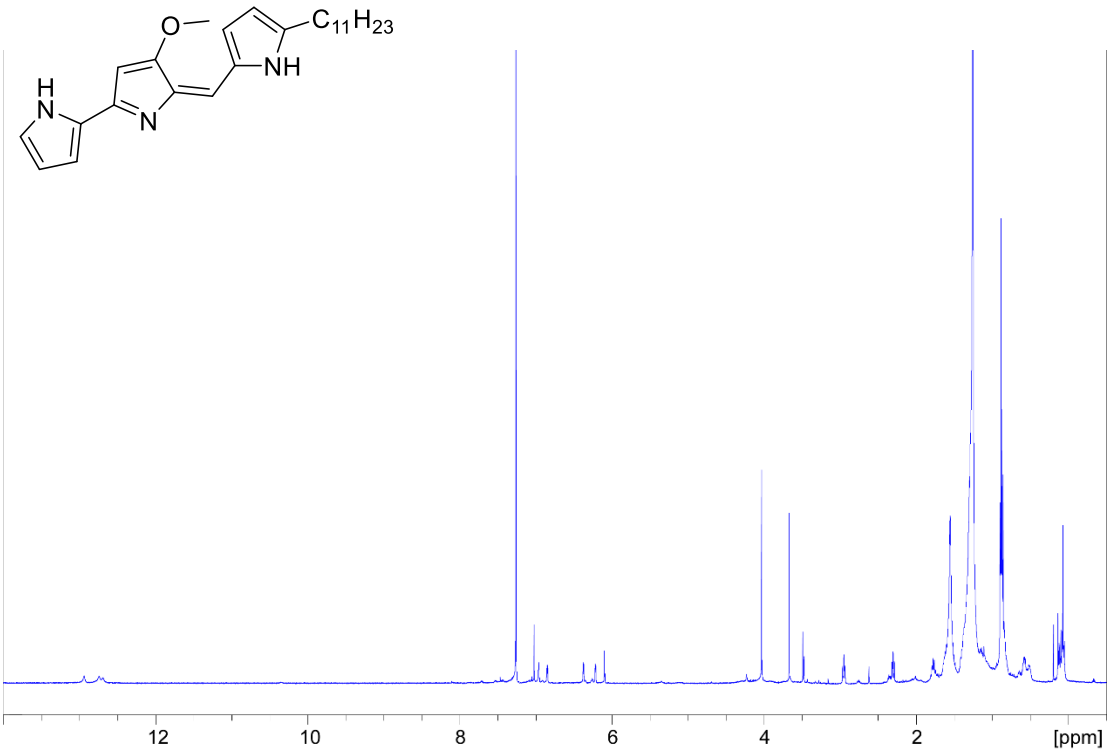
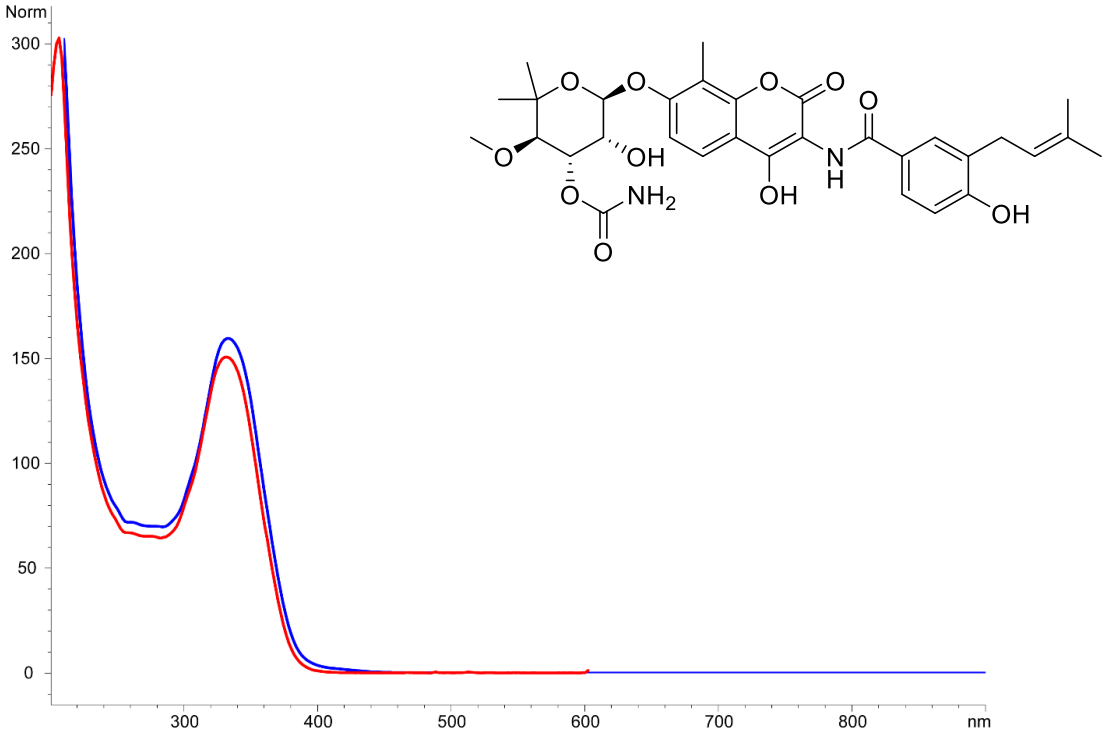


Figure A10. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of novobiocin **2.5**



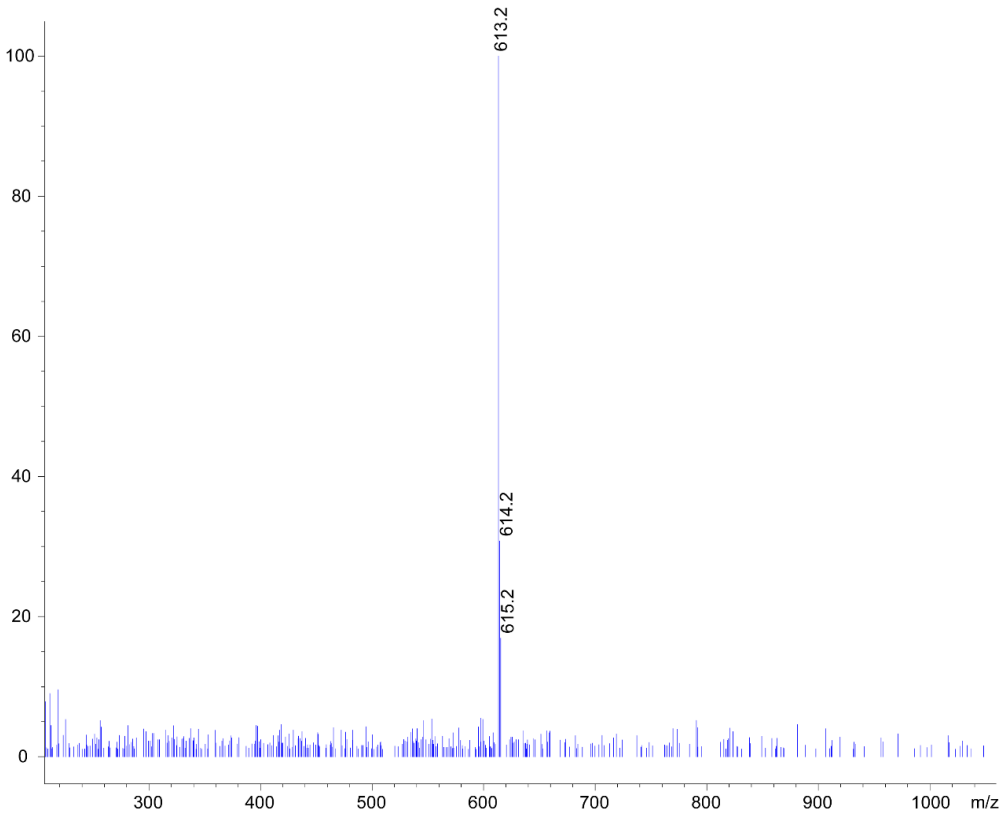


Figure A11. Results of single dose MTT-based cell viability assay for organic extracts of the Arcitic Ocean (AO) strains. The detailed information on assay results can be found in Chapter two, Table 2.1

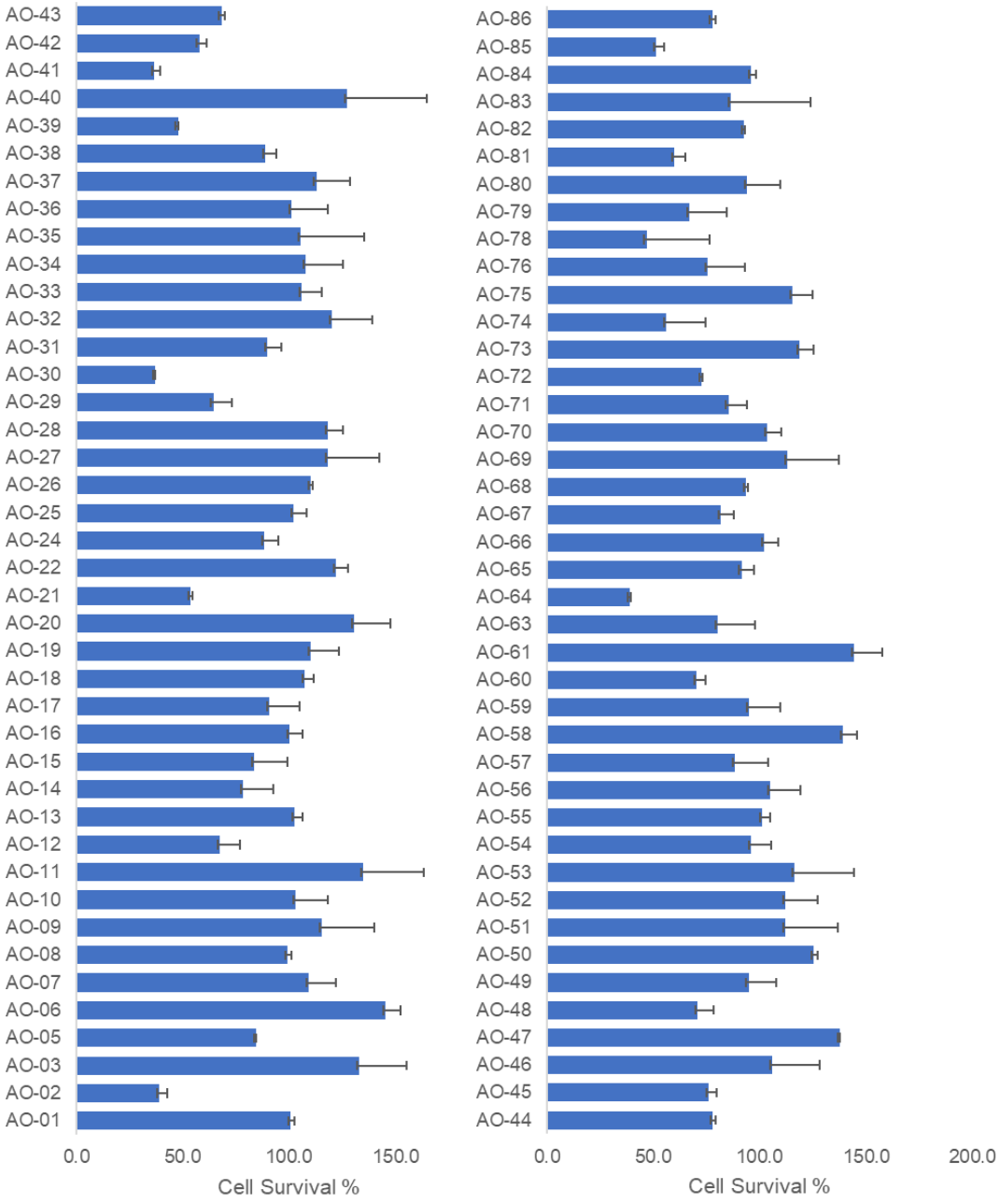


Figure A12 Results of single dose MTT-based cell viability assay for organic extracts of the Pacific Fish (PF) strains. The detailed information on assay results can be found in Chapter two, Table 2.1

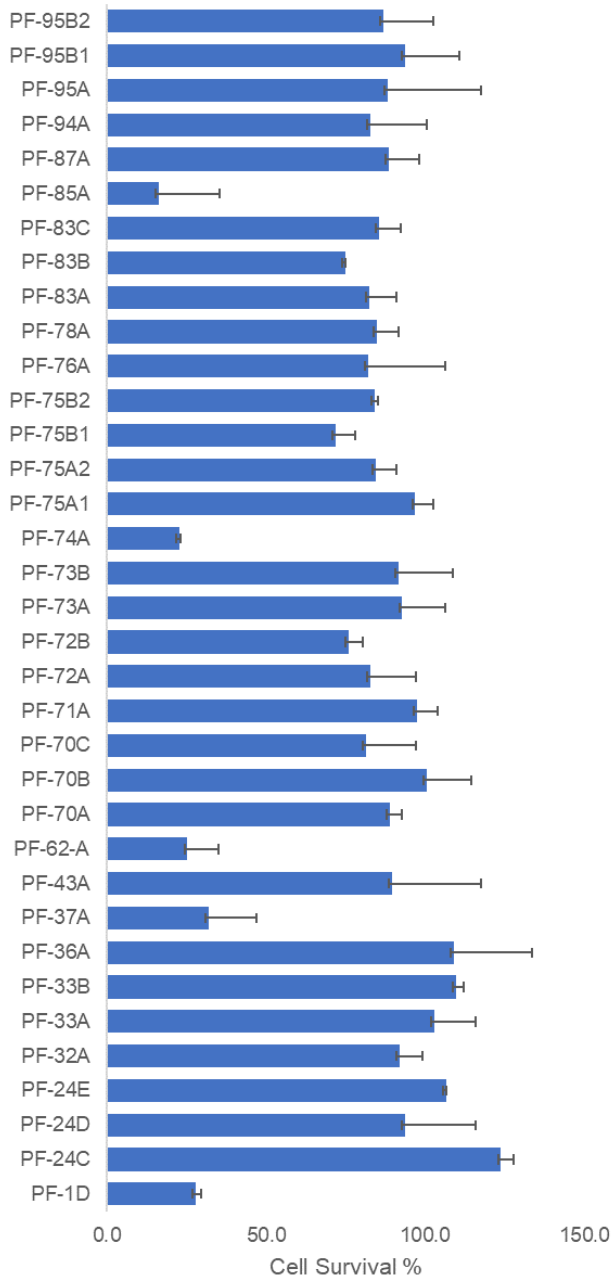
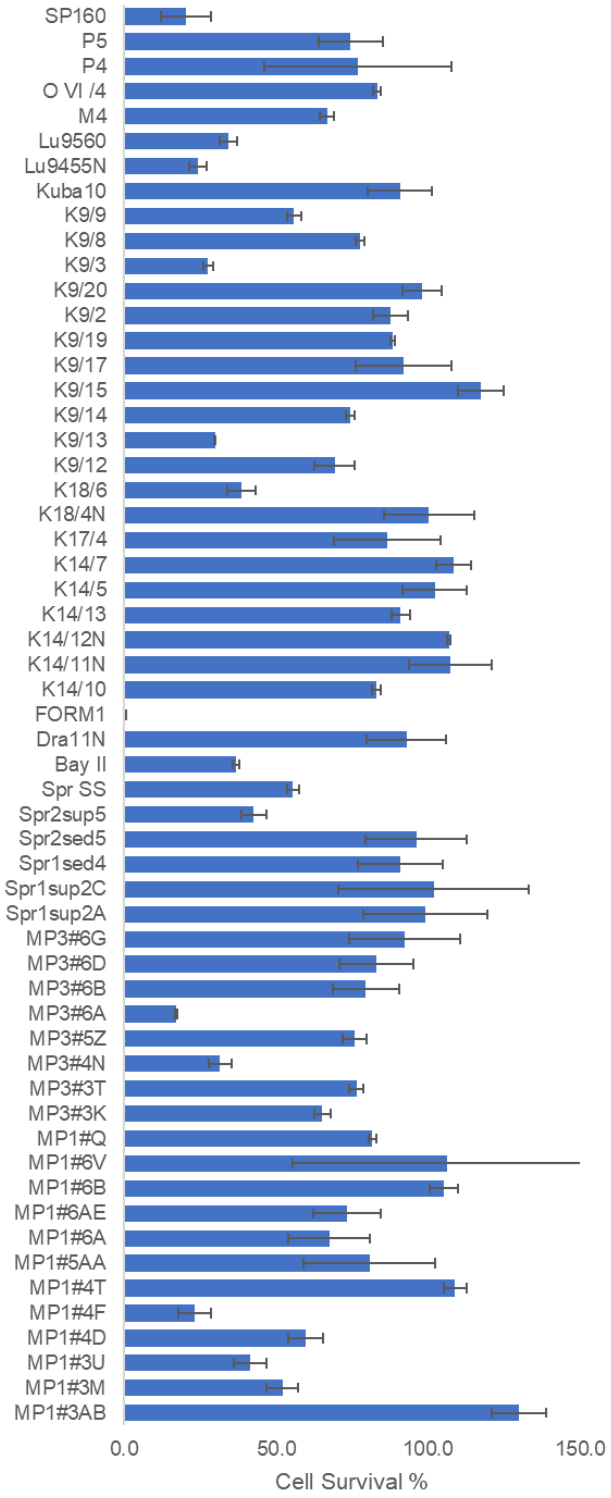


Figure A13. Results of single dose MTT-based cell viability assay for organic extracts from the other strains (excluded AO and PF strains). The detailed information on assay results can be found in Chapter two, Table 2.1



Appendix B: Supporting Information for Chapter Three

(-)-Furaquinocin C (**3.1**) Yellow solid; $[\alpha]_D^{20}$ -58° (0.25 *c*, $CHCl_3$); UV (MeCN) λmax 414, 300, 264, 222; HRESIMS m/z 393.1670 ($[M + Na]^+$; calc. for $C_{22}H_{26}O_5Na^+$, 393.1672; Δppm = 0.6); ^1H NMR (500 MHz, $CDCl_3$) δ_H 7.15 (s, 1H), 5.07 (m, 1H), 4.85 (q, *J* = 6.50 Hz, 1H), 3.99 (s, 3H), 2.06 (s, 3H), 1.97 (m, 1H), 1.95 (m, 1H), 1.84 (m, 1H), 1.65 (m, 1H), 1.64 (s, 3H), 1.53 (s, 3H), 1.45 (d, *J* = 6.60 Hz ,3H), 1.26 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ_C 184.0, 181.2, 161.2, 156.8, 134.0, 133.3, 132.1, 127.9, 123.8, 109.6, 108.9, 88.0, 60.8, 46.8, 37.7, 25.7, 23.7, 19.7, 17.6, 15.4, 9.5.

Bafilomycin C1 (**3.3**). White Solid; UV (MeCN) λmax 288, 248, 214; HRESIMS m/z 743.3973 ($[M + Na]^+$; calc. for $C_{39}H_{60}O_{12}Na^+$, 742.2977; Δppm = 0.5); ^1H NMR (500 MHz, methanol-d_4) δ_H 6.92 (d, *J* = 15.75 Hz, 1H), 6.70 (s, 1H), 6.60 (dd, *J* = 15.1, 10.8 Hz, 1H), 6.49 (d, *J* = 15.75 Hz, 1H), 5.89 (d, *J* = 8.7 Hz, 1H), 5.77 (d, *J* = 10.8 Hz, 1H), 5.11 (dd, *J* = 15.1, 8.75 Hz, 1H), 5.04 (dd, *J* = 7.75, 1.35 Hz, 1H), 5.00 (m dd, 1H), 4.16 (dd, *J* = 10.3, 4.7 Hz, 1H), 3.99 (dd, *J* = 8.4, 8.25 Hz, 1H), 3.65 (m, 1H), 3.62 (s, 3H), 3.59 (dd, *J* = 10.4, 2.0 Hz, 1H), 3.254 (s, 3H), 3.245 (m, 1H), 2.52 (m, 1H), 2.30 (dd, *J* = 11.95, 4.9 Hz, 1H), 2.08 (ddd, *J* = 10.85, 7.3, 1.1Hz, 1H), 1,98 (s, 3H), 1.94 (m, 1H), 1.90 (s, 3H), 1.84 (m, 1H), 1.65 (m, 1H), 1.81 (dd, *J* = 7.1, 1.5 Hz, 1H), 1.33 (d, *J* = 11.65 Hz, 1H), 1.29 (m, 1H), 1.06 (d, *J* = 7.00 Hz, 3H), 0.99 (d, *J* = 5.85 Hz, 3H), 0.93 (d, *J* = 6.75 Hz, 3H), 0.97 (d, *J* = 6.95 Hz, 3H), 0.86 (d, *J* = 6.9 Hz, 3H), 0.84 (d, *J* = 6.4 Hz, 3H), 0.83 (d, *J* = 6.75 Hz, 3H).

Bafilomycin D (**3.4**). White solid. UV (MeCN) λmax 284, 244; HRESIMS m/z 627.3856 ($[M + Na]^+$; calc. for $C_{35}H_{56}O_8Na^+$, 627.3867; Δppm = 1.8); ^1H NMR (500 MHz, methanol-d_4) δ 6.94 (dd, *J* = 15.75, 8.75 Hz, 1H), 6.69 (s, 1H), 6.57 (dd, *J* =

14.85, 11.00 Hz, 1H), 6.26 (d, *J* = 15.75 Hz, 1H) , 5.89 (d, *J* = 8.60 Hz, 1H), 5.79 (d, *J* = 11.00 Hz, 1H), 5.19 (dd, *J* = 6.76, 1.88 Hz, 1H), 5.16 (dd, *J* = 14.85, 8.35 Hz, 1H), 3.96 (dd, *J* = 7.85, 7.55 Hz, 1H), 3.81 (dd, *J* = 9.02, 3.50 Hz, 1H), 3.64 (s, 3H), 3.24 (s, 3H), 3.11 (dd, *J* = 7.25, 4.45 Hz, 1H), 3.09 (dd, *J* = 7.10, 3.55 Hz, 1H), 2.55 (m, 2H), 2.54 (m, 1H), 2.07 (m, 1H), 1.97 (s, 3H), 1.96 (m, 1H), 1.85 (s, 3H), 1.80 (m, 1H), 1.63 (m, 1H), 1.11 (d, *J* = 6.85 Hz, 3H), 1.08 (d, *J* = 6.85 Hz, 3H), 1.06 (d, *J* = 7.14 Hz, 3H), 0.99 (d, *J* = 6.86 Hz, 3H), 0.94 (d, *J* = 6.86 Hz, 3H), 0.91 (d, *J* = 6.75 Hz, 3H), 0.90 (d, *J* = 6.60 Hz, 3H).

FD-594 (**3.5**) was obtained as a yellow solid from two *Streptomyces lucensis* (CL16-5B and CL16-5C). The organic extract was subjected to VLCC fractionation and yielded 7 fractions (F1-F7). The bioactive fraction F4 was separated into 8 subfractions (SF1-SF8) using semiautomated flash column chromatography with an ISCO Combiflash on a 10 g silica column (Grace Reveleris) with a gradient elution from 100% hexane to 100 % EtOAc to 100% MeOH. Compound FD-594 was isolated from subfraction SF6 using preparative HPLC with isocratic elution (51% water/ 49% MeCN), and then purified with semi-preparative HPLC with isocratic elution (55% water/ 45% MeCN) .

FD-594 (**3.5**). Yellow solid. UV (MeCN) λmax 418, 362, 298, 280, 234, 214; HRESIMS m/z 963.3251 ($[M + Na]^+$; calc. for $C_{47}H_{56}O_{20}Na^+$ 963.3257; Δppm = 0.6);

Oligomycin A (**3.6**). White solid. UV (MeCN) λmax 264; HRESIMS m/z 813.5094 calc. for $C_{45}H_{74}O_{11}Na^+$, 813.5123, Δppm = 3.6); ^{1}H NMR (500 MHz, $CDCl_3$)

Chloramphenicol (**3.7**) was obtained as a white solid from *Streptomyces malachitospinus* (LL2-4A). The organic extract was subjected to VLCC fractionation and yielded 7 fractions (F1-F7). Compound **3.7** was isolated from fraction F5 using preparative HPLC with an isocratic elution of 27% MeCN/ 73% water. UV (MeCN) λmax 278; HRESIMS m/z 321.0050 ([M - H]$^-$; calc. for $C_{11}H_{11}Cl_2N_2O_5^-$, 321.0051; Δppm = 0.2); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 8.18 (d, *J* = 8.80 Hz, 2H), 7.64 (d, *J* = 8.80, 2H), 6.24 (s, 1H), 5.16 (d, *J* = 2.70 Hz, 1H), 4.13 (ddd, *J* = 8.80, 6.20, 2.70 Hz, 1H), 3.81 (dd, *J* = 10.90, 7.20 Hz, 1H), 3.60 (dd, *J* = 10.90, 6.10 Hz, 1H);

(R)-3-heptyl-3-hydroxyquinoline-2,4(*1H,3H*)-dione (**3.8**). The organic extract was subjected to VLCC fractionation. Compound **3.8** was isolated from fraction F4 and F5 using preparative HPLC with an isocratic elution protocol of 49% MeCN/ 51%water, and further purified on semipreparative HPLC with an isocratic elution protocol of 55% water /45% MeCN. White solid; $[\alpha]_D^{20}$ -34° (0.25 *c*, MeOH); UV (MeCN) λmax 326, 236, 216; HRESIMS m/z 274.1457 ([M - H]$^-$; calc. for $C_{16}H_{20}NO_3^-$ 274.1449, Δppm = 3.0); ^{1}H NMR (700 MHz, methanol-*d*$_4$) δ_H 7.825 (dd, *J* = 7.77, 1.26 Hz, 1H), 7.596 (td, *J* = 7.52, 1.47 Hz, 1H), 7.16 (t, *J* = 7.28 Hz, 1H), 7.05 (d, *J* = 8.05 Hz, 1H), 1.86 (ddd, *J* = 13.72, 11.83, 4.69 Hz, 1H), 1.77 (ddd, *J* = 13.79, 11.90, 4.83 Hz, 1H), 1.40 (m, 2H), 1.31 (m, 2H), 1.29 (m, 2H), 1.26 (m, 2H), 1.21 (m, 2H), 0.86 (t, *J* = 7.07 Hz, 3H); ^{13}C NMR (176 MHz, methanol-*d4*) δ_C 197.9, 175.1, 142.7, 128.3, 124.3, 120.8, 117.4, 83.9, 41.6, 32.8, 30.4, 30.1, 24.0, 23.6, 14.4.

(4*S*/4*R*)-sclerone (**3.9**), isosclerone (**3.10**), and 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) were isolated from *Paraburkholderia fungorum* (LL2-4C-B),

of Compounds **3.9**, **3.10**, and **3.11** were respectively eluted at 8 min,12 min, and 28 min, and then purified on semi-preparative HPLC.

(4S/4R)-sclerone (**3.9**). White solid. UV (MeCN) λmax 310, 260, 222; HRESIMS m/z $[M - H]^-$ 177.0552 (calc. for $C_{10}H_{10}O_3^-$ 177.0557, Δppm = 0.5); ^{1}H NMR (700 MHz, DMSO-*d6*) δ_H 7.30 (dd, *J* = 7.77, 1.12 Hz, 1H), 7.23 (dd, *J* = 7.84, 7.84 Hz, 1H), 7.07 (dd, *J* = 7.98, 0.84 Hz, 1H), 5.10 (t, *J* = 3.71 Hz, 1H), 2.88 (ddd, *J* = 17.21, 11.21, 5.91 Hz, 1H), 2.4087 (dt, *J* = 17.08, 4.34 Hz, 1H), 2.10 (m, 2H); ^{13}C NMR (176 MHz, DMSO-*d6*) δ_C 197.9, 155.7, 132.4, 131.4, 128.5, 120.5, 116.3, 60.0, 33.0 30.4

(4S)-isosclerone (**3.10**). White solid. UV (MeCN) λmax 310, 260, 214; HRESIMS m/z 177.0560 $([M + H]^+$, calc. for $C_{10}H_{11}O_3^+$ 179.0703, Δppm = 1.6); ^{1}H NMR (500 MHz, DMSO-*d_6*) δ_H 12.39 (s, 1H), 7.51 (dd, *J* = 8.15, 7.65 Hz, 1H), 7.08 (dd, *J* = 7.55, 0.9 Hz, 1H), 6.85 (d, *J* = 8.2 Hz, 1H), 5.60 (brs, 1H), 4.75 (dd, *J* = 8.9, 4.0 Hz, 1H), 2.75 (m, 2H), 2.20 (m, 1H), 1.98 (m, 1H); ^{13}C NMR (125 MHz, methanol-*d4*) δ_C 206.3, 163.7, 148.6, 137.9, 118.8, 117.7, 116.5, 68.3, 36.0, 32.6.

Tunicamycin VII (**3.12**) and tunicamycin VIII (**3.13**) were isolated *from Streptomyces hundungensis* (LL2-5B). The organic extract was subjected to VLCC fractionation. Compound **3.12** and **3.13** was eluted at 11.5 min and 12.5 min respectively on preparative HPLC system with isocratic elution of 50% MeCN/50% water from active fraction F7.

Tunicamycin VII (**3.12**). White solid. UV (MeCN) λmax 260; HRESIMS m/z 867.4253 $([M + Na]^+$; calc. for $C_{39}H_{64}N_4O_{16}Na^+$ 867.4210, Δppm = 5.0); ^{1}H NMR (500 MHz, methanol-*d_4*) δ_H 7.91 (d, *J* = 8.00 Hz, 1H), 6.82 (dt, *J* = 15.5, 6.85 Hz,

1H), 5.94 (d, *J* = 15.5 Hz, 1H), 5.92 (d, *J* = 5.65 Hz, 1H), 5.75 (d, *J* = 8.00 Hz, 1H), 4.93 (d, *J* = 3.25 Hz, 1H), 4.58 (d, *J* = 8.45 Hz, 1H), 4.21 (m, 1H), 4.19 (t, *J* = 5.80 Hz, 1H), 4.07 (t, *J* = 9.40 Hz, 1H), 4.01 (d, *J* = 10.00 Hz, 1H), 4.00 (t, *J* = 9.30, 1H), 3.87 (m, 1H), 3.85 (m, 1H), 3.81 (m, 1H), 3.77 (d, *J* = 10.20 Hz, 1H), 3.69 (m, 1H), 3.67 (d, *J* = 9.75 Hz, 1H), 3.66 (d, *J* = 9.75 Hz, 1H), 3.65 (m, 1H), 3.34 (t, *J* = 10.00 Hz, 1H), 2.11 (m, 1H), 1.93 (s, 3H), 1.54 (m, 2H), 1.37-1.24 (m, 16H), 0.88 (d, *J* = 6.60 Hz, 6H)

Tunicamycin VIII (**3.13**). White solid. UV (MeCN) λmax 260; HRESIMS m/z 867.4219 ($[M + Na]^+$; calc. for $C_{39}H_{64}N_4O_{16}Na^+$ 867.4210, Δppm = 1.1); ^{1}H NMR (500 MHz, methanol-d_4) δ_H 7.91 (d, *J* = 8.00 Hz, 1H), 6.82 (dt, *J* = 15.5, 6.85 Hz, 1H), 5.94 (d, *J* = 15.5 Hz, 1H), 5.92 (d, *J* = 5.65 Hz, 1H), 5.75 (d, *J* = 8.00 Hz, 1H), 4.93 (d, *J* = 3.25 Hz, 1H), 4.58 (d, *J* = 8.45 Hz, 1H), 4.21 (m, 1H), 4.19 (t, *J* = 5.80 Hz, 1H), 4.07 (t, *J* = 9.40 Hz, 1H), 4.01 (d, *J* = 10.00 Hz, 1H), 4.00 (t, *J* = 9.30, 1H), 3.87 (m, 1H), 3.85 (m, 1H), 3.81 (m, 1H), 3.77 (d, *J* = 10.20 Hz, 1H), 3.69 (m, 1H), 3.67 (d, *J* = 9.75 Hz, 1H), 3.66 (d, *J* = 9.75 Hz, 1H), 3.65 (m, 1H), 3.34 (t, *J* = 10.00 Hz, 1H), 2.11 (m, 1H), 1.93 (s, 3H), 1.54 (m, 1H), 1.37-1.24 (m, 20H), 0.92 (t, *J* = 6.50 Hz, 3H).

Anthrabenzoxocinone (6*S*, 16*S*)1.264-C (**3.14**) was obtained as a yellow solid from *Streptomyces sparsogenes* (LT2-5). The organic extract was subjected to VLCC fractionation and compound **3.14** was isolated from fractions F4 and F5 on preparative HPLC system with isocratic elution of 60% MeCN/40% water. HRESIMS m/z 459.1445 ($[M - H]^-$; calc. for $C_{27}H_{23}O_7^-$ 459.1449, Δppm = 0.9); ^{1}H NMR (700 MHz, acetone-d_6) δ_H 13.47 (s, 1H), 12.78 (s, 1H), 7.16 (s, 1H), 6.75 (d, *J*

= 2.24 Hz, 1H), 6.32 (d, *J* = 2.24 Hz, 1H), 6.27 (dd, *J* = 2.45, 0.56 Hz, 1H), 6.24 (s, 1H), 6.15 (d, *J* = 2.45 Hz, 1H), 3.33 (d, *J* = 18.0 Hz, 1H), 3.13 (d, *J* = 18.0 Hz, 1H), 2.44 (s, 3H), 1.68 (s, 3H), 1.64 (s, 3H), 1.60 (s, 3H); ^{13}C NMR (176 MHz, acetone-d6) δ_C 191.7, 166.6, 166.2, 158.5, 157.9, 155.8, 153.3, 151.2, 142.9, 136.9, 124.4, 118.4, 114.6, 112.0, 110.8, 108.0, 107.2, 101.8, 101.5, 98.5, 65.7, 40.7, 39.3, 33.9, 33.8, 27.7, 19.3.

Table B1. Results of single dose MTT-based cell viability assay. Organic extracts were tested at 10 µg/mL. 250 µM of etoposide was used as a positive control

Strain Name	Cell survival (%)				
	SK-Mel-5	PC-3	HCT-116	A549	MCF-7
DMSO	100.0 ± 1.9	100.0 ± 1.6	100.0 ± 1.3	100.0 ± 3.7	100.0 ± 2.1
CL12-5	41.4 ± 1.6	44.1 ± 3.8	74.4 ± 0.9	88.8 ± 3.3	68.9 ± 2.9
CL13-6A	55.4 ± 3.9	66.5 ± 4.3	80.0 ± 2.8	74.2 ± 2.6	64.1 ± 1.7
CL13-6B	89.9 ± 6.9	90.8 ± 4.0	99.1 ± 1.2	102.3 ± .6	85.8 ± 2.6
CL16-4	25.2 ± 3.3	36.4 ± 2.3	29.2 ± 1.4	43.8 ± 1.0	40.1 ± 0.1
CL16-5A	86.5 ± 6.4	95.4 ± 3.1	95.0 ± 1.6	104.9 ± 3.2	79.1 ± 4.4
CL16-5B	34.2 ± 0.9	32.1 ± 2.8	26.6 ± 0.7	40.2 ± 1.6	36.2 ± 1.5
CL16-5C	95.1 ± 4.0	88.2 ± 4.2	103.7 ± .6	114.8 ± 3.0	75.1 ± 2.5
CL17-3	93.0 ± 8.0	85.4 ± 2.3	99.1 ± 5.8	101.3 ± 1.6	79.0 ± 1.8
CL18-3A	64.5 ± 5.5	71.1 ± 2.2	57.2 ± 0.9	73.4 ± 1.2	66.0 ± 1.3
CL18-4D	1.7 ± 0.1	59.8 ± 1.7	39.3 ± 0.5	59.4 ± 1.1	49.8 ± 2.2
LL2-3	95.2 ± 14.8	61.0 ± 2.4	91.9 ± 6.5	79.9 ± 4.8	48.7 ± 2.6
LL2-4A	72.8 ± 6.2	66.4 ± 2.0	68.3 ± 5.1	80.3 ± 8.9	63.4 ± 5.7
LL2-4B	93.1 ± 8.3	82.9 ± 6.9	90.6 ± 6.0	76.4 ± 16.3	76.4 ± 12.2
LL2-4C-Y	100.4 ± 12.8	98.2 ± 12.1	98.0 ± 8.0	87.8 ± 14.6	70.6 ± 3.3
LL2-4C-B	82.2 ± 12.6	88.5 ± 3.7	81.5 ± 11.6	87.6 ± 7.9	61.3 ± 2.6
LL2-4E	84.2 ± 17.3	93.2 ± 5.9	85.1 ± 6.4	92.2 ± 5.2	71.1 ± 7.2
LL2-5A	87.4 ± 7.2	83.8 ± 2.6	83.6 ± 5.5	83.1 ± 3.4	62.6 ± 7.2
LL2-5B	34.4 ± 2.0	51.7 ± 1.8	8.0 ± 1.8	45.0 ± 2.8	48.1 ± 4.5
LT2-5	42.6 ± 8.3	58.6 ± 2.7	26.8 ± 15.4	53.8 ± 3.7	38.2 ± 1.0
Etoposide	6.7 ± 0.4	36.2 ± 2.9	13.7 ± 0.6	14.9 ± 0.7	41.6 ± 2.4

Table B2. Results of single dose microbroth antibacterial assay. Organic extracts were tested at 125 µg/mL. Vancomycin (µg/mL) and chloramphenicol (125µg/mL) were respectively used as positive controls against Gram-positive and Gram-negative pathogens.

Sample	*E. faecium*	MRSA	*P. aeruginosa*	*E. coli*
DMSO	100.0 ± 3.2	100.0 ± 3.0	100.0 ± 1.7	100.0 ± 3.8
CL12-5	18.0 ± 3.6	-4.5 ± 0.2	87.9 ± 10.3	67.0 ± 1.4
CL13-6A	4.7 ± 4.7	108.3 ± 6.9	85.5 ± 3.5	68.6 ± 1.0
CL13-6B	5.9 ± 1.7	96.1 ± 2.4	92.9 ± 1.1	88.3 ± 1.7
CL16-4	7.0 ± 8.1	-5.1 ± 0.6	95.0 ± 1.7	80.4 ± 1.2
CL16-5A	9.4 ± 2.3	69.9 ± 8.4	91.4 ± 3.2	75.3 ± 1.9
CL16-5B	114.1 ± 11.6	35.9 ± 1.1	97.2 ± 0.8	92.3 ± 3.8
CL16-5C	7.8 ± 2.7	-5.2 ± 1.6	100.2 ± 1.9	73.6 ± 3.6
CL17-3	119.5 ± 3.3	67.2 ± 2.1	89.4 ± 3.0	80.5 ± 4.9
CL18-3A	114.8 ± 8.1	99.6 ± 2.6	85.9 ± 1.4	77.1 ± 1.0
CL18-4D	124.9 ± 2.0	79.4 ± 3.8	87.6 ± 2.3	82.7 ± 2.7
LL2-3	4.2 ± 1.2	0.7 ± 0.3	108.5 ± 1.7	67.2 ± 2.8
LL2-4A	6.9 ± 0.0	1.2 ± 0.3	19.4 ± 12.6	0.5 ± 0.3
LL2-4B	117.4 ± 2.8	10.7 ± 3.9	83.0 ± 9.6	86.7 ± 1.5
LL2-4C-Y	101.6 ± 1.1	103.9 ± 2.8	105.2 ± 2.8	88.4 ± 1.1
LL2-4C-B	10.1 ± 1.5	55.1 ± 3.0	109.8 ± 3.2	86.2 ± 1.1
LL2-4E	126.2 ± 4.0	107.7 ± 1.0	97.3 ± 3.6	93.1 ± 1.5
LL2-5A	135.7 ± 3.9	96.1 ± 7.0	101.6 ± 1.3	98.6 ± 11.6
LL2-5B	5.6 ± 3.2	38.4 ± 3.4	104.0 ± 1.5	95.2 ± 4.3
LT2-5	29.2 ± 1.2	0.7 ± 0.3	88.1 ± 2.3	65.3 ± 5.6
+ CNTRL	1.6 ± 1.4	0.0 ± 11.3	-0.4 ± 0.1	0.5 ± 0.3

Figure B1. 16S rRNA sequence of bacterium CL12-5

TGCAGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGTGA
GTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACGG
GGTCTAATACCGGATAACACCTTGCCTCTCCTGAGGGAAGGTTAAAAGCT
CCGGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGTAAT
GGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCA
CACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG
AATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGA
TGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTG
ACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGG
TAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTC
GTAGGCGGCTTGTCACGTCGGGTGTGAAAGCCCGGGGCTTAACCCCGGGT
CTGCATTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGATCGGAATTCC
TGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAA
GGCGGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGGAG
CGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGAACTA
GGTGTTGGCGACATTCCACGTCGTCGGTGCCGCAGCTAACGCATTAAGTT
CCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACG
GGGGCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACGCGA
AGAACCTTACCAAGGCTTGACATACACCGGAAAGCATCAGAGATGGTGC
CCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTC
GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTTCTGTGTT
GCCAGCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACCGCCGGGGT
CAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTC
TTGGGCTGCACACGTGCTACAATGGCAGGTACAATGAGCTGCGATACCGT
GAGGTGGAGCGAATCTCAAAAAGCCTGTCTCAGTTCGGATTGGGGTCTGC
AACTCGACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTGC
TGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGA
AAGTCGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGC

Figure B2. 16S rRNA sequence of bacterium CL13-6A

TGCAAGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGTG
AGTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACG
GGGTCTAATACCGGATACTGACCGGCCTGGGCATCCAGGCGGGTCGAAA
GCTCCGGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGT
AATGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGG
CCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTG
GGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAG
GGATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAA
GTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCG
CGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAG
CTCGTAGGCGGCTTGTCACGTCGGTTGTGAAAGCCCGGGGCTTAACCCCG
GGTCTGCAGTCGATACGGGCAGGCTAGAGTTCGGTAGGGGAGATCGGAA
TTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGC
GAAGGCGGATCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGGG
GAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGC
ACTAGGTGTGGGCAACATTCCACGTTGTCCGTGCCGCAGCTAACGCATTA
AGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATT
GACGGGGGCCCGCACAAGCGGCGGAGCATGTGGCTTAATTCGACGCAAC
GCGAAGAACCTTACCAAGGCTTGACATACACCGGAAAGCATTAGAGATA
GTGCCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCCG
TGTTGCCAGCAGGCCCTTGTGGTGCTGGGGACTCACGGGAGACCGCCGGG
GTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATG
TCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACC
GCGAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCT
GCAACTCGACCCCATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCATT
GCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCAC
GAAAGTCGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGA
GCTGTCGAAG

Figure B3. 16S rRNA sequence of bacterium CL13-6B

CATGCAGTCGAACGATGAACCTCACTTGTGGGGGGATTAGTGGCGAACGG
GTGAGTAACACGTGAGTAACCTGCCCTTGACTCTGGGATAAGCCTGGGAA
ACCGGGTCTAATACTGGATACGACCTTCTGGCGCATGCCATGTTGGTGGA
AAGCTTTTGTGGTTTTGGATGGACTCGCGGCCTATCAGCTTGTTGGTGGGG
TAATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGACCG
GCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGT
GGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCGTGA
GGGATGAAGGCCTTCGGGTTGTAAACCTCTTTCAGTAGGGAAGAAGCCGG
CCTTTTGGGTTGGTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGT
GCCAGCAGCCGCGGTAATACGTAGGGCGCAAGCGTTATCCGGAATTATTG
GGCGTAAAGAGCTCGTAGGCGGTTTGTCGCGTCTGCCGTGAAAGTCCGGG
GCTTAACTCCGGATCTGCGGTGGGTACGGGCAGACTAGAGTGCAGTAGGG
GAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGA
ACACCGATGGCGAAGGCAGGTCTCTGGGCTGTAACTGACGCTGAGGAGC
GAAAGCATGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGT
AAACGTTGGGCACTAGGTGTGGGGGACATTCCACGTTTTCCGCGCCGTAG
CTAACGCATTAAGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACT
CAAAGGAATTGACGGGGGCCCGCACAAGCGGCGGAGCATGCGGATTAAT
TCGATGCAACGCGAAGAACCTTACCAAGGCTTGACATGAACCGGAATGAT
GCAGAGATGTGTCAGCCACTTGTGGCCGGTTTACAGGTGGTGCATGGTTG
TCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAA
CCCTCGTTCCATGTTGCCAGCGGGTTATGCCGGGGACTCATGGGAGACTG
CCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAATCATCATGCCC
CTTATGTCTTGGGCTTCACGCATGCTACAATGGCCGGTACAAAGGGTTGC
GATACTGTGAGGTGGAGCTAATCCCAAAAAGCCGGTCTCAGTTCGGATTG
AGGTCTGCAACTCGACCTCATGAAGTTGGAGTCGCTAGTAATCGCAGATC
AGCAACGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCA
AGTCACGAAAGTTGGTAACACCCGAAGCCGGTGGCCTAACCCCTTGTGGG
AG

Figure B4. 16S rRNA sequence of bacterium CL16-5A

AGTCCGAACGGTGAAGCCTTCGGGTGGATCAGTGGCGAACGGGTGAGTA
ACACGTGGGCAATCTGCCCTGCACTCTGGGACAAGCCCTGGAAACGGGGT
CTAATACCGGATATGACCGGCTCTCGCATGGGGGTTGGTGTAAAGCTCCG
GCGGTGCAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGGGGTGATGGC
CTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACAC
TGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
ATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGA
CGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGCAAGTGACG
GTACCTGCAGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAA
TACGTAGGGTGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTA
GGCGGCTTGTCGCGTCGGATGTGAAAGCCCGGGGCTTAACCCCGGGTCTG
CATTCGATACGGGCAGGCTAGAGTGTGGTAGGGGAGATCGGAATTCCTGG
TGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGC
GGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGA
ACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGTTGGGAACTAGGT
GTTGGCGACATTCCACGTCGTCGGTGCCGCAGCTAACGCATTAAGTTCCC
CGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGG
GCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGA
ACCTTACCAAGGCTTGACATACACCAGAAAACCGTAGAGATACGGTCCCC
CTTGTGGCTGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGA
GATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGTTCTGTGTTGCCA
GCATGCCTTTCGGGGTGATGGGGACTCACAGGAGACCGCCGGGGTCAACT
CGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGG
CTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCGTGAGGT
GGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACTC
GACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTGCTGCGG
TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTC
GGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGTCG

Figure B5. 16S rRNA sequence of bacterium CL16-5B

CATGCAGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGT
GAGTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAAC
GGGGTCTAATACCGGATATGACCATCTTGGGCATCCTTGATGGTGTAAAG
CTCCGGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGTA
ATGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGC
CACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGG
GGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGG
GATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAG
TGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGC
GGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGC
TCGTAGGCGGCTTGTCACGTCGGGTGTGAAAGCCCGGGGCTTAACCCCGG
GTCTGCATTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGATCGGAATT
CCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCG
AAGGCGGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGG
AGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGAA
CTAGGTGTTGGCGACATTCCACGTCGTCGGTGCCGCAGCTAACGCATTAA
GTTCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTG
ACGGGGGCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACG
CGAAGAACCTTACCAAGGCTTGACATACACCGGAAAGCATTAGAGATAGT
GCCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTGTG
TTGCCAGCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACCGCCGGGG
TCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGT
CTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCG
TGAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTG
CAACTCGACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTG
CTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACNNCACG
AAAGTTGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAG
CT

Figure B6. 16S rRNA sequence of bacterium CL16-5C

TGCAAGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGTG
AGTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACG
GGGTCTAATACCGGATATGACCATCTTGGGCATCCTTGATGGTGTAAAGC
TCCGGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGTAA
TGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCC
ACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGG
GAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGG
ATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGT
GACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGCG
GTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCT
CGTAGGCGGCTTGTCACGTCGGGTGTGAAAGCCCGGGGCTTAACCCCGGG
TCTGCATTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGATCGGAATTC
CTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGA
AGGCGGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGGA
GCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGAACT
AGGTGTTGGCGACATTCCACGTTGTCGGTGCCGCAGCTAACGCATTAAGT
TCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACGCG
AAGAACCTTACCAAGGCTTGACATACACCGGAAAGCATTAGAGATAGTGC
CCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTC
GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTGTGTT
GCCAGCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACCGCCGGGGT
CAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTC
TTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCGT
GAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGC
AACTCGACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTGC
TGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGA
AAGTTGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGC

Figure B7. 16S rRNA sequence of bacterium CL17-3

GTCGAGCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCGGCGGACGGGTGA
GTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACGTTCGGAAACGGA
CGCNAATACCGCATACGTCCTACGGGAGAAAGCAGGGGACCTTCGGGCC
TTGCGCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGGGGTAATG
GCTCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGATCAGTCACA
CTGGAACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
TATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGA
AGGTCTTCGGATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTTACCT
AATACGTATCTGTTTTGACGTTACCGACAGAATAAGCACCGGCTAACTCT
GTGCCAGCAGCCGCGGTAATACAGAGGGTGCAAGCGTTAATCGGAATTAC
TGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGGATGTGAAATCCCC
GGGCTCAACCTGGGAACTGCATTCAAAACTGACTGACTAGAGTATGGTAG
AAGGGTGGTGGGATNTCCTGTGTAGCGGNGAAATGCGTAGATATAGNAA
GGAACACCAGTGGNGAAGGGGACCNCCTGGACTNATACTGACACTGAGG
TGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGC
CGTNAACGATGTCAACTAGCCGTTGGGAGCCTTGAGCTNTTAGTGGCGCA
GCTAACGCATTAAGTTGNCCNCCTGGGGAGTACGGCCGCAAGGTTAAAAC
TCAAATGAATTGACGGGGGCCCGCACAAACGGTGGAGCATGTGGTTTAAT
TCGAAGCAACGCGAAGAACCTTTCCAGGCCTTGACATCCAATGAACTTTC
CAGAGATGGATTGGTGCCTTCGGGAACATTGAGACAGGTGCTGCATGGCT
GTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGCGCA
ACCCTTGTCCTTAGTTACCAGCACGTAATGGTGGGCACTCTAAGGAGACT
GCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGC
CCTTACGGCCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTG
CCAAGCCGCGAGGTGGAGCTAATCCCACAAAACCGATCGTAGTCCGGATC
GCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGCGAAT
CAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCCCCCGTCA
CACCATGGGAGTGGGTTGCACCAGAAGTAGCTAGTCTAACCTTCGGGAGG

Figure B8. 16S rRNA sequence of bacterium CL18-3A

GTCGAGCGCCCCGCAAGGGGAGCGGGGACGGGGGAGTAACGCGTGGGAA
TCTACCCATCACTACGGAACAACTCCGGGAAACTGGAGCTAATACCGTAT
ACGTCCTTCGGGAGAAAGATTTATCGGTGATGGATGAGCCCGCGTTGGAT
TAGCTAGTTGGTGGGGTAATGGCCTACCAAGGCGACGATCCATAGCTGGT
CTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCT
ACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCC
AGCCATGCCGCGTGAGTGATGAAGGCCCTAGGGTTGTAAAGCTCTTTCAA
CGGTGAAGATAATGACGGTAACCGTAGAAGAAGCCCCGGCTAACTTCGTG
CCCTCAGCCGCGGTAATACGAAGGGGGCTAGCGTTGTTCGGATTTACTGG
GCGTAAAGCGCACGTAGGCGGATTGTTAAGTTAGGGGTGAAATCCCAGG
GCTCATCCCTGGAACTGCCTTTAATACTGGCAATCTCGAGTCCGGAAGAG
GTGAGTGGAATTCCGAGTGTAGAGCTGATATTNGTAGATATTCGGAGGAA
CACACAGTGGNGAAGGCGGCTCTCTGGTCCGGTAGTGACGNTGANGTGCG
AAAGCGTGGGGAGCAAACAGGATTAGATTCCCTGGTAGTNCACGCTGTAA
ACGATGGAAGCTAGCCGTCGGCATGTTTACTTGTCGGTGGCGCAGCTAAC
GCATTAAGCTTCCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAG
GAATTGACGGGGGCCCGCCCAAGCGGTGGAGCATGTGGTTTAATTCGAAG
CAACGCGCAGAACCTTACCAGCCCTTGACATCCCGGTCGCGGCCTAGAGA
GATTTAGGCCTTCAGTTCGGCTGGACCGGTGACAGGTGCTGCATGGCTGT
CGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC
CCTCGCCCTTAGTTTCCATCATTCAGTTGGGCACTCTAAGGGGACTGCCGG
TGATAAGCCGAGAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTA
CGGGCTGGGCTACACACGTGCTACAATGGTGGTGACAGTGGGCAGCGAG
GCCGCGAGGTCGAGCTAATCTCCAAAAGCCATCTCAGTTCGGATTGCACT
CTGCAACTCGAGTGCATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCA
TGCCGCGGTGAATACGTTCCCCGGCCTTGTACCCACCGCCCGTCACACCA
AGGGAGTTGGCTTTACCCGAAGGCGCTGTCCTAACCGCAAGGAGGCAGG
CGACC

Figure B9. 16S rRNA sequence of bacterium CL18-4D

ACCATGCAGTCGAGCGATGAAGCCCTTCGGGGTGGATTAGAGGGAGAGG
GTGAGTAGCACGCGGGCTTCTGCCCTGCACTCTGGGACAAGACCTGGTGA
CGGGGTCTAATACCGGATAATACTCTCAGATGCATCTGTGAGGGTTAAAA
GCTCCGGTGATGCAGGATGAGGCCAGGGAACTATCTCCTTGTTGGTGAGA
TAGTGGCTCACCCTGGCAACAACAGGTAGCCGGCCTGNGAGGGCGACCG
GCCCCNCTGCCTACTGAGACACGGCCCAGACTCCTACTGGAGGCAGGGGG
GGGGAATATTGCACACTGGGCGAAAGACTGATGAAGGGCCTCCGGGTGA
GGGATGACGGTCTTCGGGAAGAAAACCTCTTTCACCAGGGAAGAAGCCC
CAGTGACCTTCCCTGCAGAAGACGCGCCAGTTAACTACGTGCCAGCAGCC
GTCGGAATACCTGGGGCGCAAGGGTGGTCCGGAATTATTGGGCGTAAAG
AGCTCAATCCTGGCTTGTCACTCCAGGTGTGAAAGCCCGGCTGTTAACCC
CGGGTCTGCATTCGATACGGGCTATNTAGAGTGTGGTAGGGGAGATNGGA
AATTCCTGGTGTANCGGTGAAATTAGCAGATATCATGAGAGACATTCGGT
GAGCGAAGGCGGATCTCTGGGCCATTCACTGGCCCTGATGATCGANATGA
TGCAGAGCAAACAGGATTACATACCCTGGTAGTTCACGCCGTAATCCGTG
GGATATAGGAGTTANNGACATTGCTGGTAATTGGTCTTGCAAGTAACGCA
TTAAGTTCCCTAAGCNNGGAGTACGGGGAGAAGGGTAAAAATCAAAGGA
ATCAAAGGGGGCCCGCACAAGCAGCGGAGCATGTGGCTTAATTGGTTTNA
ACGCGAAGAACCTTACCAAGGCTTNACATCCNCCGGAAAGCATTAGAGA
TAGTGCCCCCCTTGTGGTTCCCTTACAGGTGGTGCATGCAGGTCGTCAACT
CGTGTCGTGAGATGGTGGGGTAAGTCCCTCGGTTAACGCAACCCTTGTTN
TGTGTCCCCATCATTCCTTTCGGGGAGATGGGGGATCACAGGAAGCAGCC
GGGGTCAGATNGGAGGAAGGGGGGGTGGGGGTCAAGTCATCATCCCCCT
TACCTTTTGGGCTGCGCTCCTCCTGCAATGCCAGATCCGATGAAATGGGAT
NCCGCAAGGTGGACCGAATCTCAAAAAGTCGGTCTCAGTTCGGATTGGGG
TGTGCAACTCGCCCCCATCCAGTTGGAGTTGATAGTAATTGCAGATCAGC
ATTCATGCCGCGAATAAGTTCCCGGCCCTCCACACCCCGCCCCCCATCTCA
CCAAAGTAGGTGACTCCCGAAACCGGGGCCCAACCCCAAAGGGAGCAAC
CGG

Figure B10. 16S rRNA sequence of bacterium LL2-3

TCGAACGATGAAGCCCTTCGGGGTGGATTAGTGGCGAACGGGTGAGTAAC
ACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACGGGGTCT
AATACCGGATAACACTCCTCAGGGCATCTTGAGGGGTTAAAAGCTCCGGC
GGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGGGGTAATGGCCT
ACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTG
GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATAT
TGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACG
GCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGACGGT
ACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGGTAATA
CGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGG
CGGCTTGTCACGTCGGATGTGAAAGCCCGGGGCTTAACCCCGGGTCTGCA
TTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGATCGGAATTCCTGGTG
TAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGG
ATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAAC
AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGTTGGGAACTAGGTGT
TGGCGACATTCCACGTTGTCGGTGCCGCAGCTAACGCATTAAGTTCCCCG
CCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGC
CCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGAAC
CTTACCAAGGCTTGACATATACCGGAAACGGCTAGAGATAGTCGCCCCCT
TGTGGTCGGTATACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAG
ATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTTCTGTGTTGCCAGC
ATGCCCTTCGGGGTGATGGGGACTCACAGGAGACTGCCGGGGTCAACTCG
GAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGGCT
GCACACGTGCTACAATGGCCGGTACAAAGAGCTGCGATGCCGCGAGGCG
GAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACTCG
ACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTGCTGCGGT
GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTCG
GTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGA

Figure B11. 16S rRNA sequence of bacterium LL2-4A

GTCGACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGTGAGTAAC
ACGTGGGCAATCTGCCCTGCACTCTGGGACAAGCCCTGGAAACGGGGTCT
AATACCGGATACTGAACCTCGCAGGCATCTGTGAGGGTCGAAAGCTCCGG
CGGTGCAGGATGAGCCCGCGGCCTATCAGCTAGTTGGTGAGGTAATGGCT
CACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACT
GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATA
TTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGAC
GGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGACGG
TACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGGTAAT
ACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAG
GCGGCTTGTCACGTCGGTTGTGAAAGCCCGGGGCTTAACCCCGGGTCTGC
AGTCGATACGGGCAGGCTAGAGTTCGGTAGGGGAGATCGGAATTCCTGGT
GTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCG
GATCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAA
CAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGCACTAGGTG
TGGGCAACATTCCACGTTGTCCGTGCCGCAGCTAACGCATTAAGTGCCCC
GCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGG
CCCGCACAAGCGGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGAA
CCTTACCAAGGCTTGACATACACCGGAAACGGCCAGAGATGGTCGCCCCC
TTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGA
GATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCCGTGTTGCCA
GCAGGCCCTTGTGGTGCTGGGGACTCACGGGAGACCGCCGGGGTCAACTC
GGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGGC
TGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCGCAAGGTG
GAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACTCG
ACCCCATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCATTGCTGCGGT
GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTCG
GTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGC

Figure B12. 16S rRNA sequence of bacterium LL2-4B

CATGCAGTCGAACGATGAAGCCTTCGGGTGGATTAGTGGCGAACGGGTGA
GTAACACGTGGGTAATCTGCCCTGCACTCTGGGATAAGCCTTGGAAACGA
GGTCTAATACCGGATATCACAACTCTTCGCATGGGGGGTTGTTGAAAGTT
CCGGCGGTGCAGGATGAACCCGCGGCCTATCAGCTTGTTGGTGGGGTAAT
GGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGACCGGCCA
CACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG
AATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGA
TGACGGCCTTCGGGTTGTAAACCTCTTTCGCCAGGGACGAAGCGCAAGTG
ACGGTACCTGGATAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGCGG
TAATACGTAGGGTGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGAGCTC
GTAGGCGGTTTGTCGCGTCGGCTGTGAAATCTGGAGGCTTAACCTTCAGC
GTGCAGTCGATACGGGCAGACTTGAGTTCGGTAGGGGAGACTGGAATTCC
TGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAA
GGCGGGTCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGGGGAG
CGAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAACGGTGGGTGCTA
GGTGTGGGCGACATCCACGTTGTCCGTGCCGTAGCTAACGCATTAAGCAC
CCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGG
GGGCCCGCACAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAA
GAACCTTACCTGGGCTTGACATGCGCCAGACATCCCCAGAGATGGGGCTT
CCCTTGTGGTTGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGT
GAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATCCTACGTTGC
CAGCGCGTTATGGCGGGGACTCGTGGGAGACTGCCGGGGTCAACTCGGA
GGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTATGTCCAGGGCTTC
ACACATGCTACAATGGCTGGTACAGAGGGCTGCGATACCGCGAGGTGGA
GCGAATCCCTTAAAGCCGGTCTCAGTTCGGATCGCAGTCTGCAACTCGAC
TGCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGA
ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCATGAAAGTCGGT
AACACCCGAAGCCCATGGCCCAACCCGCAAGGGGGNGAGTGT

Figure B13. 16S rRNA sequence of bacterium LL2-4C-B

GACCGTCCTCCTTGCGGTTAGACTAGCCACTTCTGGTAAAACCCACTCCCA
TGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCA
TGCTGATCCGCGATTACTAGCGATTCCAGCTTCACGCACTCGAGTTGCAG
AGTGCGATCCGGACTACGATCGGTTTTCTGGGATTGGCTCCCCCTCGCGG
GTTGGCGACCCTCTGTTCCGACCATTGTATGACGTGTGAAGCCCTACCCAT
AAGGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACC
GGCAGTCTCCCTAGAGTGCTCTTGCGTAGCAACTAGGGACAAGGGTTGCG
CTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGC
CATGCAGCACCTGTGTTATGGCTCCCTTTCGGGCACTCCCACCTCTCAGCA
GGATTCCATACATGTCAAGGGTAGGTAAGGTTTTTCGCGTTGCATCGAATT
AATCCACATCATCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTT
TAATCTTGCGACCGTACTCCCCAGGCGGTCAACTTCACGCGTTAGCTACGT
TACCAAGTCAATGAAGACCCGACAACTAGTTGACATCGTTTAGGGCGTGG
ACTACCAGGGTATCTAATCCTGTTTGCTCCCCACGCTTTCGTGCATGAGCG
TCAGTATTGGGCCCAGGGGGCTGCCTTCGCCATCGGTATTCCTCCACATCT
CTACGCATTTCACTGCTACACGTGGAAATTCTACCCCCNCTGCCATACTCT
AGCCCGNCAGTCACNAATGCAGTTCCCAGGTTAAGCCCGGGGGATTTCAC
ATCTG

Figure B14. 16S rRNA sequence of bacterium LL2-4C-Y

TACACATGCAGTCGAACGGAAAGGCCCTTCGGGGTACTCGAGTGGCGAA
CGGGTGAGTAACACGTGGGTGATCTGCCCTGCACTTCGGGATAAGCCTGG
GAAACTGGGTCTAATACCGGATATGACCTTGGGATGCATGTCCTTTGGTG
GAAAGCTTTTGCGGTGTGGGATGGGCCCGCGGCCTATCAGCTTGTTGGTG
GGGTTAAGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGA
CCGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGGAGGCAGC
AGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCG
TGAGGGACGACGGCCTTCGGGTTGTAAACCTCTTTCAGCACAGACGAAGC
GCAAGTGACGGTATGTGCAGAAGAAGGACCGGCCAACTACGTGCCAGCA
GCCGCGGTAATACGTAGGGTCCGAGCGTTGTCCGGAATTACTGGGCGTAA
AGAGCTCGTAGGTGGTTTGTCGCGTTGTTCGTGAAAACTCACAGCTCAAC
TGTGGGCGTGCGGGCGATACGGGCAGACTAGAGTACTGCAGGGGAGACT
GGAATTCCTGGTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCG
GTGGCGAANGCGGGTCTCTGGGCAGTAACTGACGCTGAGGAGCGAAAGC
GTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGT
GGGTACTAGGTGTGGGTTTCCTTCCTTGGGATCCGTGCCGTAGCTAACGCA
TTAAGTACCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGAA
ATTGACGGGGGCCCGCACAAGCGGCGGAGCATGTGGATTAATTCGATGCA
ACGCGAAGAACCTTACCTGGGTTTGACATGCACAGGACGTGCCTAGAGAT
AGGTATTCCCTTGTGGCCTGTGTGCAGGTGGTGCATGGCTGTCGTCAGCTC
GTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATCTT
ATGTTGCCAGCGCGTCATGGCGGGGACTCGTAAGAGACTGCCGGGGTCAA
CTCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTATGTCCAG
GGCTTCACACATGCTACAATGGCCGGTACAAAGGGCTGCGATGCCGTAAG
GTGGAGCGAATCCTTTAAAGCCGGTCTCAGTTCGGATCGGGGTCTGCAAC
TCGACCCCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGC
GGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCATGAAAG
TCGGTAACACCCGAAGCCGGTGGCCTAACCCCTCGTGGGAGNGAGCCGTC
GAA

Figure B15. 16S rRNA sequence of bacterium LL2-4E

ATGCAAGTCGAACGATGAAGCCTTTCGGGGTGGATTAGTGGCGAACGGGT
GAGTAACACGTGGGTAACCTGCCCTTGGCTTTGGGATAAGCCCTGGAAAC
GGGGTCTAATACCGGATATGACTGACTCTCGCATGGGGGTTGGTGGAAAG
TTCCGGCGGCCAGGGATGGGCCCGCGGCCTATCAGCTTGTTGGTGGGGTG
ATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGC
CACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGG
GGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGG
GATGACGGCCTTCGGGTTGTAAACCTCTTTCGGCAGGGACGAAGCGAAAG
TGACGGTACCTGCAGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGC
GGTAATACGTAGGGTGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGAGC
TCGTAGGCGGTTTGTCGCGTCGACTGTGAAAACCTGGGGCTTAACTCTGG
GCTTGCAGTCGATACGGGCAGACTTGAGTTCGGTAGGGGAGACTGGAATT
CCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCG
AAGGCGGGTCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGGGG
AGCGAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAACGTTGGGCGC
TAGGTGTGGGGGACATTCACGTCTTCCGTGCCGTAGCTAACGCATTAAGC
GCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCG
AAGAACCTTACCTGGGTTTGACATGCACCAGATATCCCTAGAGATAGGGC
TTCCCTTGTGGTTGGTGTGCAGGTGGTGCATGGCTGTCGTCAGCTCGTGTC
GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCGTTCCATGTT
GCCAGCGCGTAATGGCGGGGACTCATGGGAGACTGCCGGGGTCAACTCG
GAGGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTATGTCCAGGGCT
TCACACATGCTACAATGGCCGGTACAGAGGGCTGCGAGACCGTGAGGTG
GAGCGAATCCCTTAAAGCCGGTCTCAGTTCGGATCGGGGTCTGCAACTCG
ACCCCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGT
GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTCG
GTAACACCCGAAGCCCATGGCCCAACCCGTAAGGGAGGGAG

Figure B16. 16S rRNA sequence of bacterium LL2-5A

AGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGTGAGT
AACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACGGGG
TCTAATACCGGATAACACTGTCACGGGCATCTGTGGTGGTTAAAAGCTCC
GGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGTAATGG
CTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACA
CTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAA
TATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATG
ACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGAC
GGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGGTA
ATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGT
AGGCGGCTTGTCACGTCGGGTGTGAAAGCCCGGGGCTTAACCCCGGGTCT
GCATTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGATCGGAATTCCTG
GTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGG
CGGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAGCGTGGGGAGCG
AACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGAACTAGG
TGTTGGCGACATTCCACGTTGTCGGTGCCGCAGCTAACGCATTAAGTTCCC
CGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGG
GCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGA
ACCTTACCAAGGCTTGACATACGCCGGAAAGCATCAGAGATGGTGCCCCC
CTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGA
GATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTGTGTTGCCA
GCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACCGCCGGGGTCAACT
CGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGG
CTGCACACGTGCTACAATGGCAGGTACAATGAGCTGCGATACCGTGAGGT
GGAGCGAATCTCAAAAAGCCTGTCTCAGTTCGGATTGGGGTCTGCAACTC
GACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGATCAGCATTGCTGCGG
TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTC
GGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGC

Figure B17. 16S rRNA sequence of bacterium LL2-5B

CTTANNACATGCAGTCGAACGATGAAGCCCTTCGGGGTGGATTAGTGGCG
AACGGGTGAGTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCC
TGGAAACGGGGTCTAATACCGGATAACACCCCCCATCGCATGGTGGGGG
GTTAAAAGCTCCGGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTG
GTGGGGTGATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGG
CGACCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC
AGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCC
GCGTGAGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGA
AGCGAAAGTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCA
GCAGCCGCGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGAGCTCGTAGGCGGCTTGTCACGTCGGATGTGAAAGCCCGGGGCTT
AACCCCGGGTCTGCATTCGATACGGGCTAGCTAGAGTGTGGTAGGGGAGA
TCGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACC
GGTGGCGAAGGCGGATCTCTGGGCCATTACTGACGCTGAGGAGCGAAAG
CGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGT
TGGGAACTAGGTGTTGGCGACATTCCACGTCGTCGGTGCCGCAGCTAACG
CATTAAGTTCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGG
AATTGACGGGGGCCCGCACAAGCAGCGGAGCATGTGGCTTAATTCGACG
CAACGCGAAGAACCTTACCAAGGCTTGACATATACCGGAAACGGCTAGA
GATAGTCGCCCCCTTGTGGTCGGTATACAGGTGGTGCATGGCTGTCGTCA
GCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTG
TTCTGTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACT
GCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCC
CCTTATGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAAAGAGCTG
CGATGCCGCGAGGCGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATT
GGGGTCTGCAACTCGACCCCATGAAGTCGGAGTTGCTAGTAATCGCAGAT
CAGCATTGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTC
ACGTCACGAAAGTCGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGG
GAGGGAGCTGTCGAAGG

Figure B18. 16S rRNA sequence of LT2-5

CATGCAAGTCGAACGATGAAGCCGCTTCGGTGGTGGATTAGTGGCGAACG
GGTGAGTAACACGTGGGCAATCTGCCCTGCACTCTGGGACAAGCCCTGGA
AACGGGGTCTAATACCGGATATTACCTGGTGAGGCATCTTACCGGGTGGA
AAGCTCCGGCGGTGCAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGGG
GTGATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACC
GGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAG
TGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTG
AGGGATGACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGC
AAGTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGC
CGCGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAG
AGCTCGTAGGCGGCTTGTCACGTCGGATGTGAAAGCCCGGGGCTTAACCC
CGGGTCTGCATTCGATACGGGCAGGCTAGAGTTCGGTAGGGGAGATCGGA
ATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGG
CGAAGGCGGATCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGG
GGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGTTGGGA
ACTAGGTGTGGGCGACATTCCACGTTGTCCGTGCCGCAGCTAACGCATTA
AGTTCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATT
GACGGGGGCCCGCACAAGCGGCGGAGCATGTGGCTTAATTCGACGCAAC
GCGAAGAACCTTACCAAGGCTTGACATACATCGGAAACATCCAGAGATG
GGTGCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTC
GTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCT
GTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACTCACAGGAGACTGCCG
GGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA
TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATA
CCGCGAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGT
CTGCAACTCGACCCCATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCA
TTGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTC
ACGAAAGTCGGTAACACCCGAAGCCGGTGGCCCAACCCCTGGGAGGGAG
CCGTCGAA

Table B3. ^{1}H-NMR (500 MHz, $CDCl_3$) and ^{13}C-NMR (125 MHz, $CDCl_3$) data of (-)-furaquinocin C (**3.1**)

position	δ_C	δ_H (*J* in Hz), type
2	88.03	4.85 q (6.50), 1H
2-CH_3	15.42	1.45 d (6.60), 3H
3	46.85	
3-CH_3	19.75	1.26 s, 3H
3a	127.91	
4	156.85	
5	108.91	7.15 s, 1H
5a	133.27	
6	184.04	
7	156.82	
7-OCH_3	60.76	3.99 s, 3H
8	134.03	
8-CH_3	9.45	2.06 s, 3H
9	181.17	
9a	109.57	
9b	161.19	
10	37.78	1.97 m, 1H; 1.65 m, 1H
11	23.71	1.95 m, 1H; 1.84 m, 1H
12	123.76	5.07 m, 1H
13	132.11	
14	17.65	1.53 s, 3H
15	25.71	1.64 s, 3H

Figure B19. ^{1}H NMR spectrum of (-)-furaquinocin C (**3.1**) in $CDCl_3$ at 500 MHz NMR.

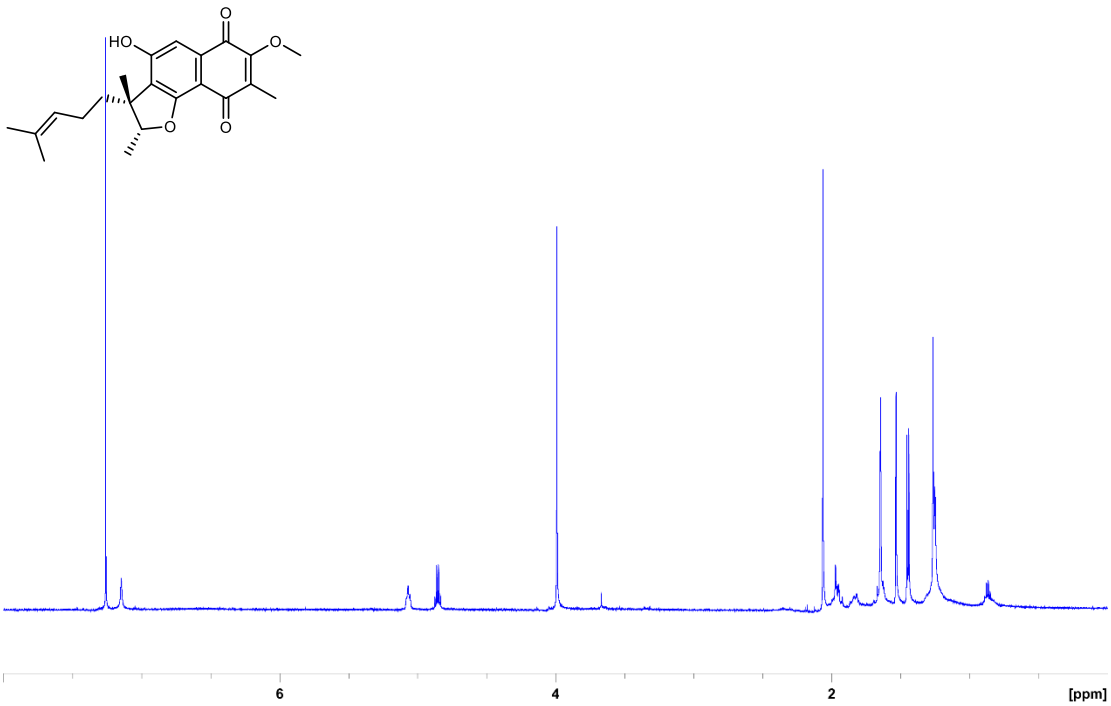


Table B4. ^{1}H-NMR (500 MHz, in methanol-*d4*) and ^{13}C-NMR (125 MHz, in methanol-*d4*) data of bafilomycin C1 (**3.3**)

position	δ_H (*J* in Hz), type
2-OCH_3	3.62 s, 3H
3	6.70 s, 1H
4	
5	5.77 d (10.80), 1H
6	2.52 m, 1H
7	3.24 m, 1H
8	1.94 m, 1H
9a	2.08 ddd (10.85, 7.30, 1.10), 1H
9b	1.81 dd (7.10, 1.50), 1H
10	
11	5.89 d (8.70), 1H
12	6.60 dd (15.10, 10.80), 1H
13	5.11 dd (15.10, 8.70), 1H
14	3.99 dd (8.40, 8.25), 1H
14-OCH_3	3.25 s, 3H
15	4.99 dd (10.40, 4.70), 1H
16	2.19 m, 1H
17	4.16 dd (10.30, 1.65), 1H
18	1.65 m, 1H
19	
20a	2.30 dd (11.95, 4.90), 1H
20b	1.33 d (11.65), 1H
21	3.65 m, 1H
22	1.29 m, 1H
23	3.59 dd (10.40, 2.00), 1H
24	1.84 m, 1H
25	0.84 d (6.40), 3H
26	2.03 s, 3H
27	1.06 d (7.00), 3H
28	0.93 d (6.75), 3H
29	1.98 s, 3H
30	0.83 d (6.75), 3H
31	0.99 d (5.85), 3H
32	0.97 d (6.95), 3H
33	0.86 d (6.90), 3H
2'	6.49 d (15.75), 1H
3'	6.92 d (15.75), 1H

Figure B20. ^{1}H NMR spectrum of bafilomycin C1 (**3.3**) in methanol-d_4 at 500 MHz

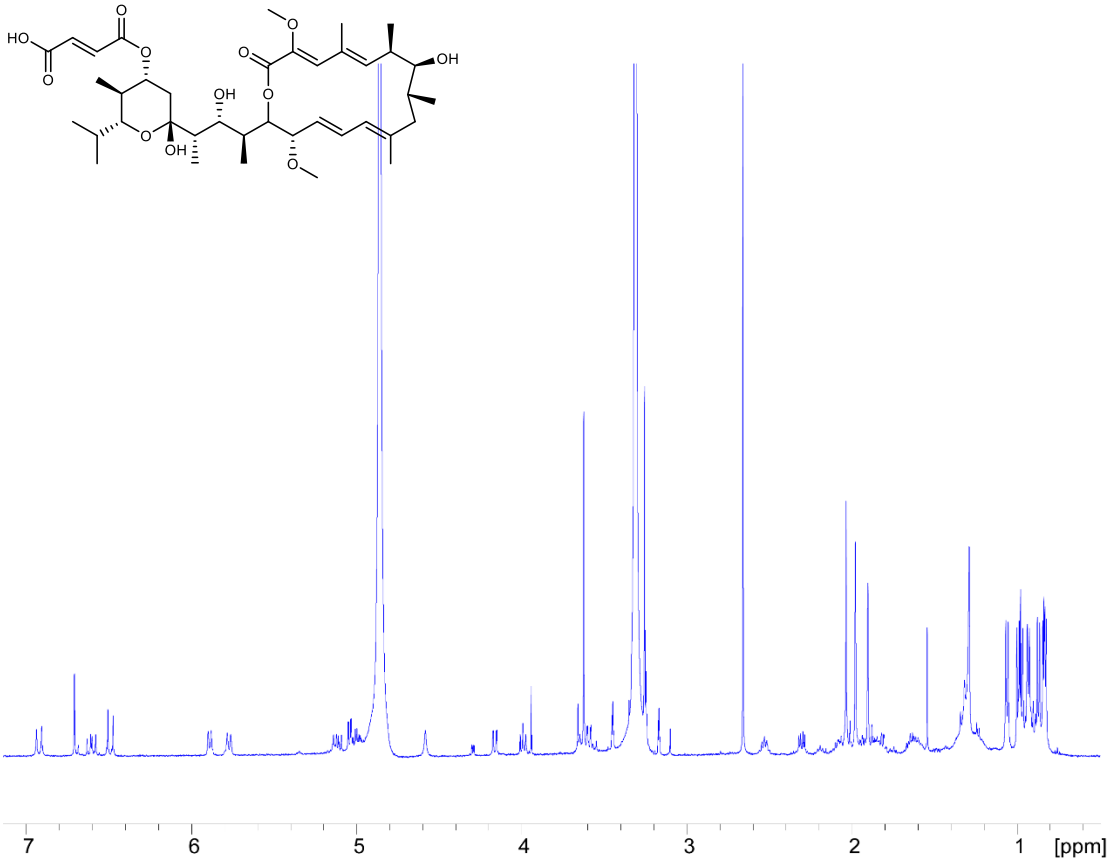


Table B4. ^{1}H-NMR (500 MHz, in methanol-d_4) of bafilomycin D (**3.4**)

position	δ_H (*J* in Hz), type
1	
2	
2-OCH_3	3.64 s, 3H
3	6.69 s, 1H
4	
4-CH_3	1.97 s, 3H
5	5.89 d (8.60),1H
6	2.54 m, 1H
6-CH_3	1.06 d (7.14), 3H
7	3.25 m, 1H
8-CH_3	0.91 d (6.75), 1H
8	1.97 m, 1H
9a	2.07 m, 1H
9b	1.80 m, 1H
10	
10-CH_3	1.85 s, 3H
11	5.79 d (11.00), 1H
12	6.57 dd (14.85, 11.00), 1H
13	5.16 dd (14.85, 8.35), 1H
14	3.96 dd (7.85, 7.55), 1H
14-OCH_3	3.24 s, 3H
15	5.19 dd (6.76, 1.88), 1H
16	2.55 m, 1H
16-CH_3	0.91 d (6.60), 3H
17	3.81 dd (9.02, 3.50), 1H
18	3.09 dd (7.10, 3.55), 1H
18-CH_3	1.11 d (6.85), 3H
19	
20	6.26 d (15.70), 1H
21	6.94 dd (15.70, 8.75), 1H
22	2.55 m, 1H
22-CH_3	1.08 d (6.85), 3H
23	3.11 dd (7.25, 4.45), 1H
24	1.63 m, 1H
24-CH_3	0.99 d (6.86), 3H
25	0.94 d (6.86), 3H

Figure B21. ^{1}H NMR spectrum of bafilomycin D (**3.4**) in methanol-*d4* at 500 MHz

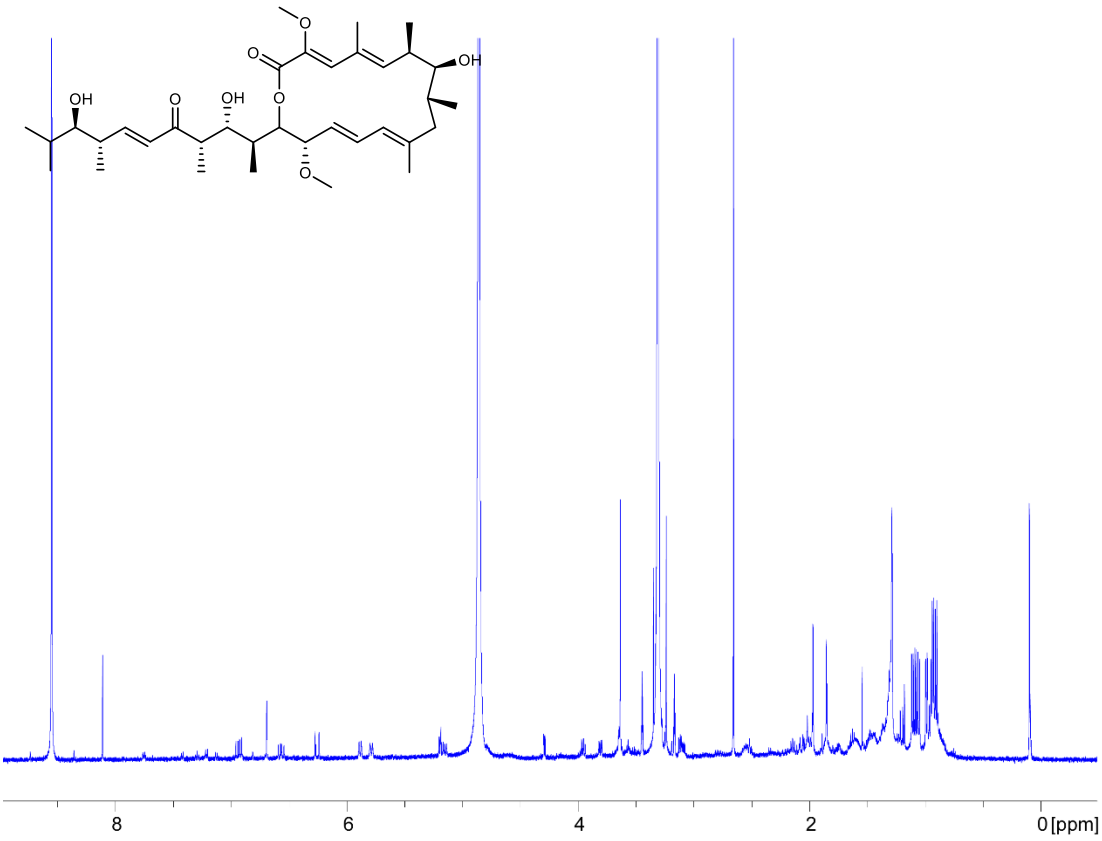


Figure B22. ^{1}H NMR spectrum of FD-594 (**3.5**) in DMSO-*d*$_4$ at 700 MHz

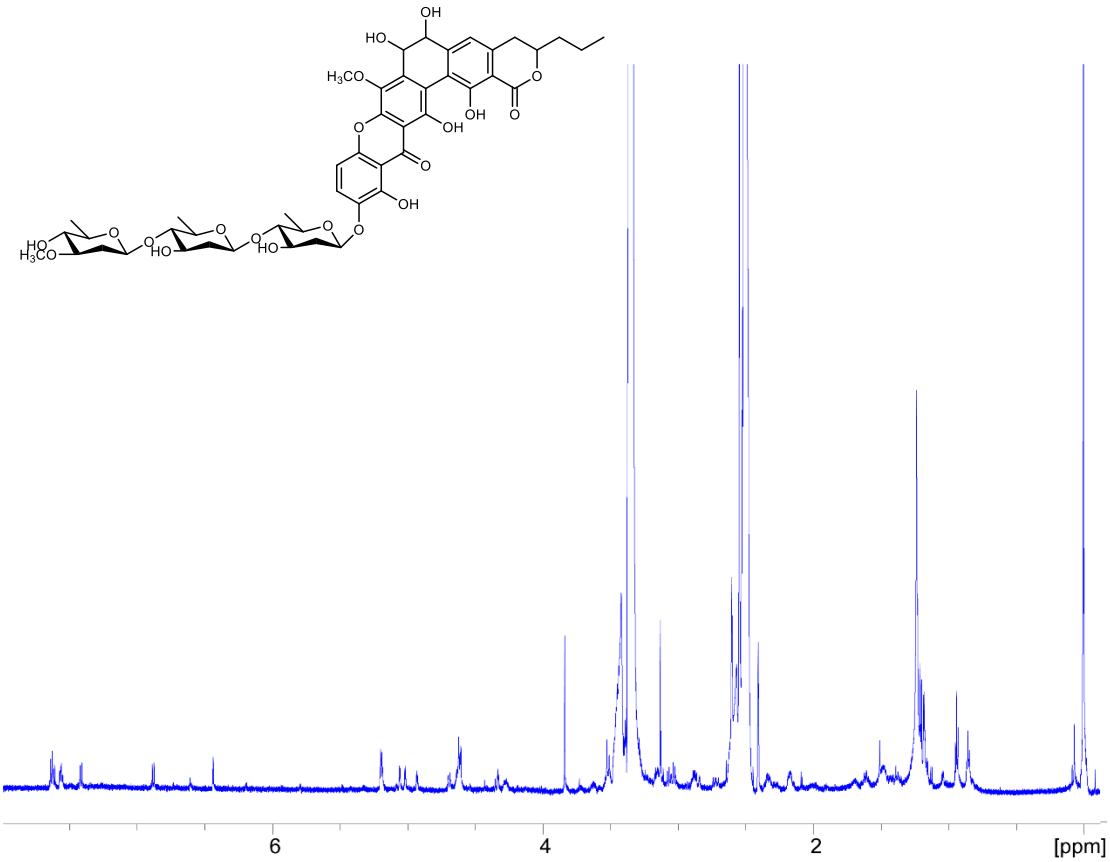


Figure B23. ^{1}H NMR spectrum of oligomycin A (**3.6**) in DMSO-d_4 at 800 MHz

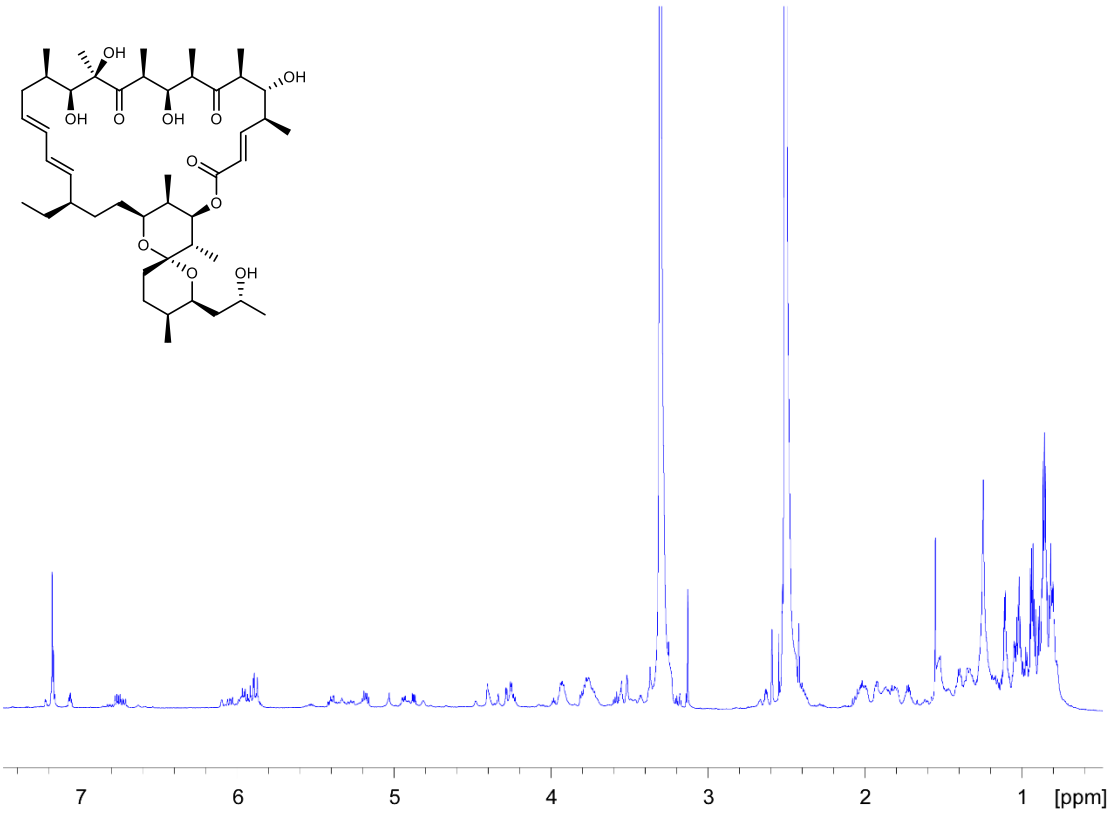


Table B5. ^{1}H-NMR (700 MHz, in methanol-d_4) of chloramphenicol (**3.7**)

Position	δ_H (*J* in Hz), type
1	
2, 6	8.18 d (8.80), 1H
3, 5	7.64 d (8.80), 1H
4	
7	5.16 d (2.70),1H
8	4.13 ddd (8.80, 6.20, 2.70), 1H
9a	3.81 dd (10.90, 7.20), 1H
9b	3.60 dd (10.90, 6.10), 1H
10	
11	6.24 s, 1H

Figure B24. ^{1}H NMR spectrum of chloramphenicol (**3.7**) in methanol-d_4 at 700 MHz

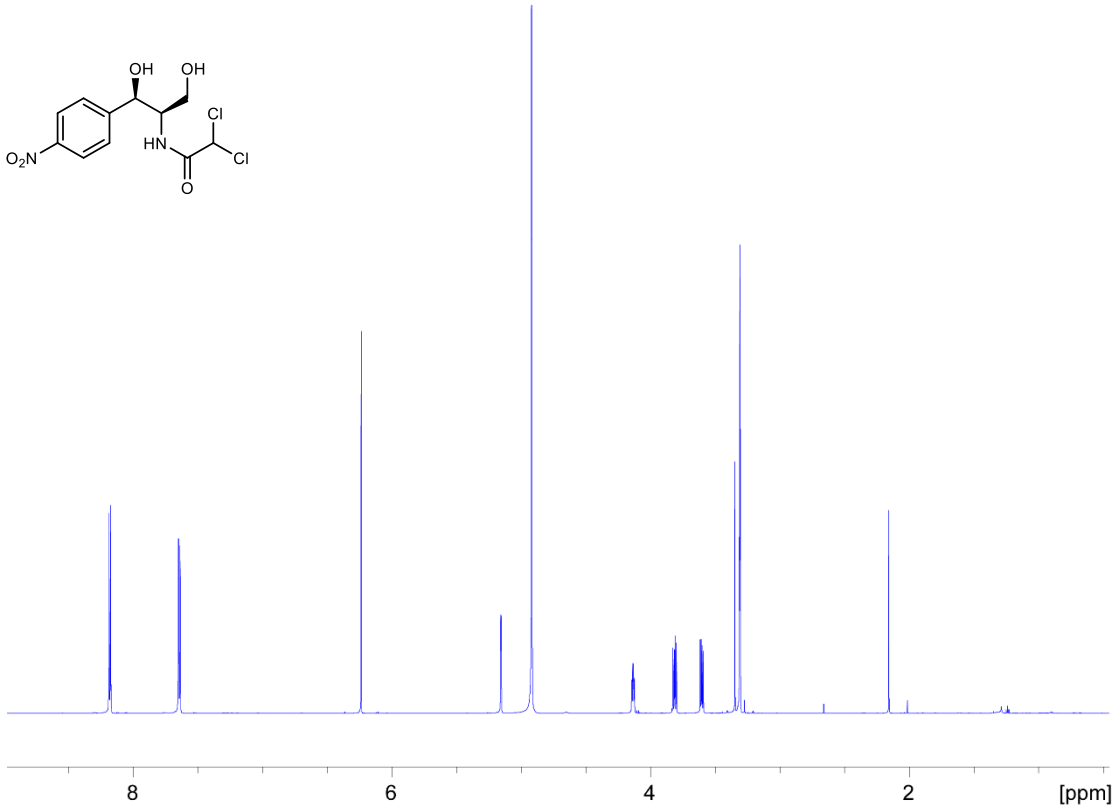


Table B6. ^{1}H-NMR (700 MHz, in methanol-d_4) and ^{13}C-NMR (176 MHz, in methanol-d_4) data of MY12-62A (**3.8**)

position	δ_C	δ_H (J in Hz), type
1		
2	175.10	
3	83.92	
4	197.87	
4a	120.83	
5	128.31	7.82 dd (7.77, 1.26), 1H
6	124.34	7.16 t (7.28), 1H
7	137.34	7.60 td (7.52, 1.47), 1H
8	117.37	7.05 d (8.05),1H
8a	142.69	
1'	41.59	1.86 ddd (13.72, 11.90, 4.69), 1H; 1.77 ddd (13.72, 11.90, 4.83) 1H
2'	23.96	1.40 m, 2H
3'	30.44	1.21 m, 2H
4'	30.12	1.31 m, 2H
5'	32.80	1.26 m, 2H
6'	23.65	1.29 m, 2H
7'	14.42	0.86 t (7.07), 3H

Figure B25. ^{13}C NMR spectrum of chloramphenicol (**3.7**) in methanol-d_4 at 176 MHz

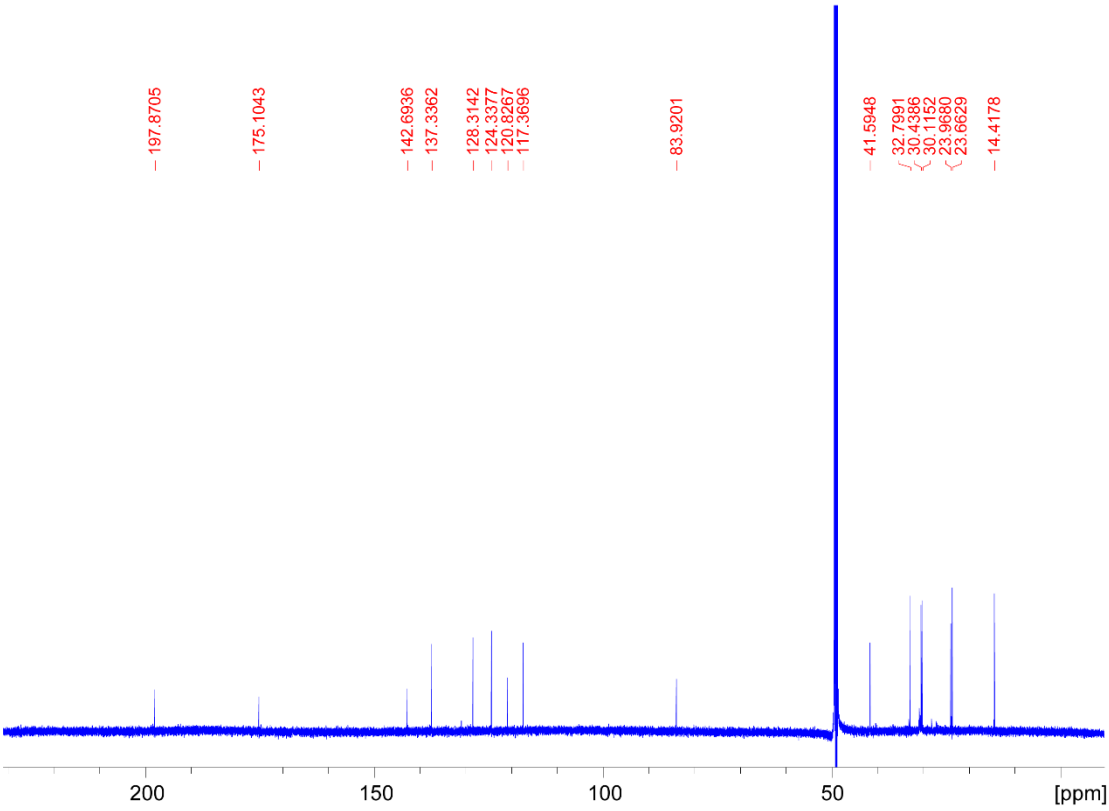


Figure B26. ^{1}H NMR spectrum of MY12-62A in methanol-d_4 at 700 MHz

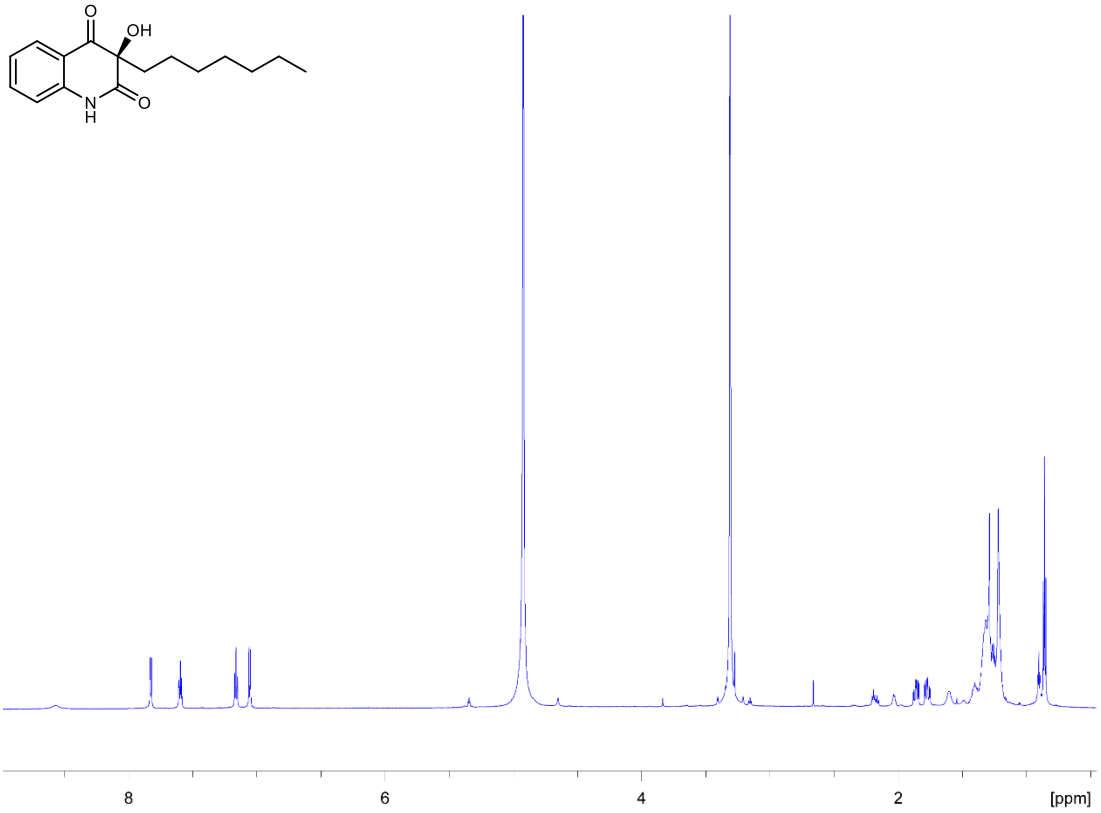


Table B7. ^{1}H-NMR (700 MHz, in methanol-d_4) and ^{13}C-NMR (176 MHz, in methanol-d_4) data of (4*S*/4*R*)-sclerone (**3.9**)

Position	δ_C	δ_H (*J* in Hz), type
1	197.9	
		2.88 ddd (17.21, 11.21, 5.91), 1H;
2	33.0	2.41 dt (17.08, 4.34),1H
3	30.5	2.10 m, 2H
4	60.0	5.10 t (3.71), 1H
4-OH		
4a	131.4	
5	155.7	
5-OH		
6	120.5	7.07 dd (7.98, 0.84), 1H
7	128.5	7.23 dd (7.84, 7.84), 1H
8	116.3	7.30 dd (7.77, 1.12), 1H
8a	132.4	

Figure B27. ^{1}H NMR spectrum of (4S/4R)-sclerone (**3.9**) in methanol-d_4 at 700 MHz

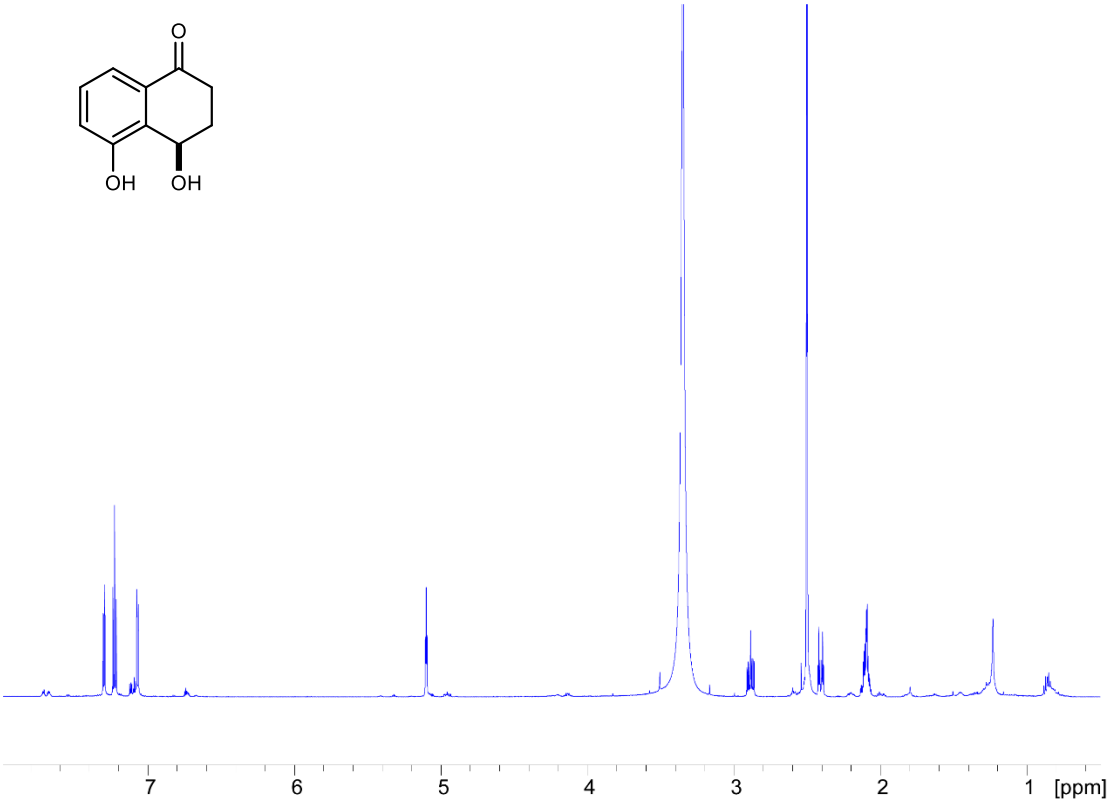
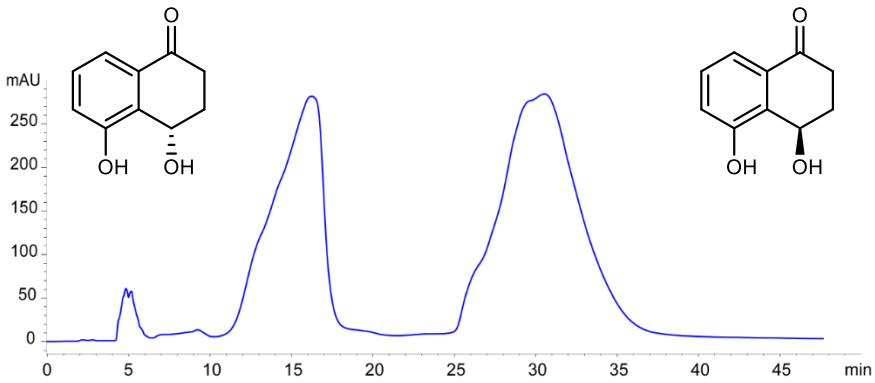


Figure B28. (A) HPLC separation of the racemic mixture of (4*S*/4*R*)-sclerone (**3.9**) on chiral column. (B) Experimental ECD spectra for (4*S*)-sclerone (blue line) and (4*R*)-sclerone (red line)

A



B

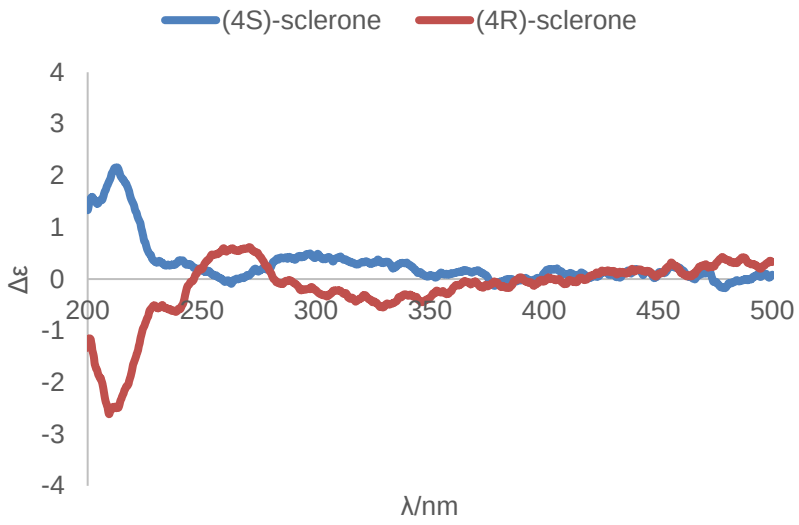


Table B8. ^{1}H-NMR (700 MHz, in methanol-d_4) and ^{13}C-NMR (176 MHz, in methanol-d_4) data of isosclerone (**3.10**)

Position	δ_C	δ_H (*J* in Hz), type
1	206.3	
2	36.0	2.75 m, 2H
3	32.6	2.20 m, 1H; 1.98 m, 1H
4	68.3	4.75 dd (8.9, 4.0), 1H
4-OH		5.60 br s, 1H
4a	148.6	
5	118.8	7.08 dd (7.55, 0.9), 1H
6	137.9	7.51 dd (8.15, 7.65), 1H
7	117.7	6.85 d (8.2), 1H
8	163.7	
8-OH		12.39 s, 1H
8a	116.6	

Figure B29. ^{1}H NMR spectrum of isosclerone (**3.10**) in methanol-d_4 at 700 MHz

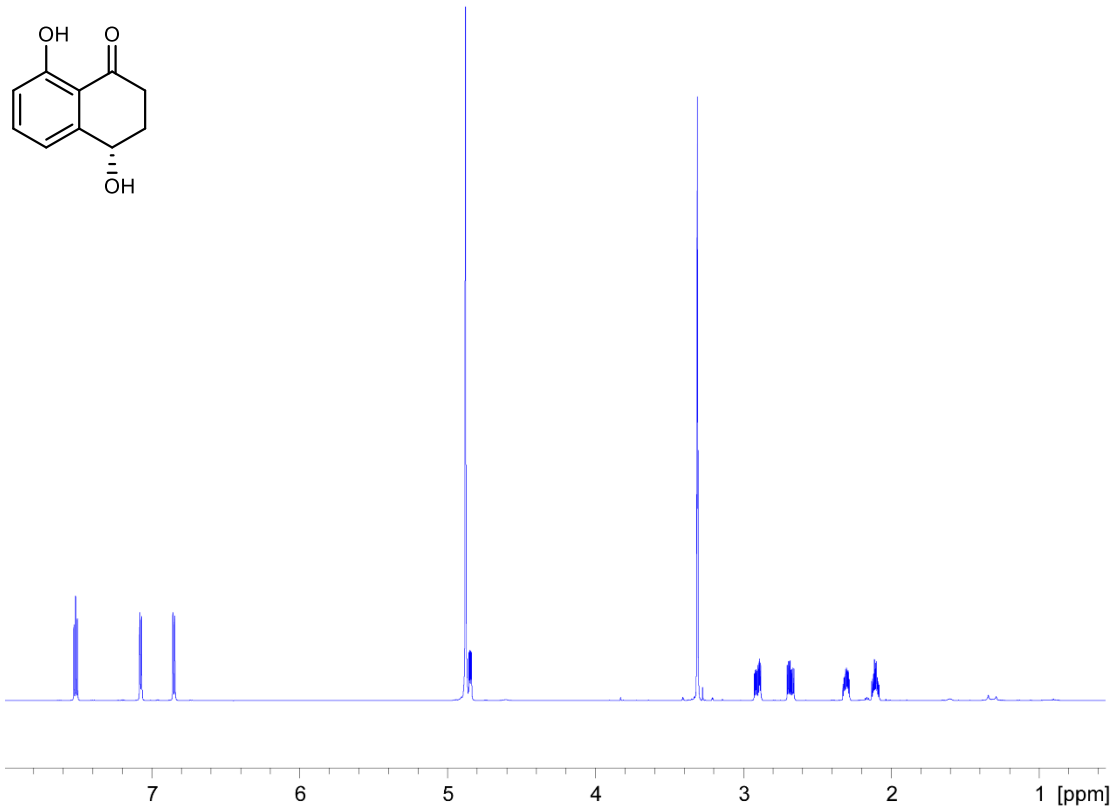


Figure B30. Experimental ECD spectra for isosclerone (**3.10**)

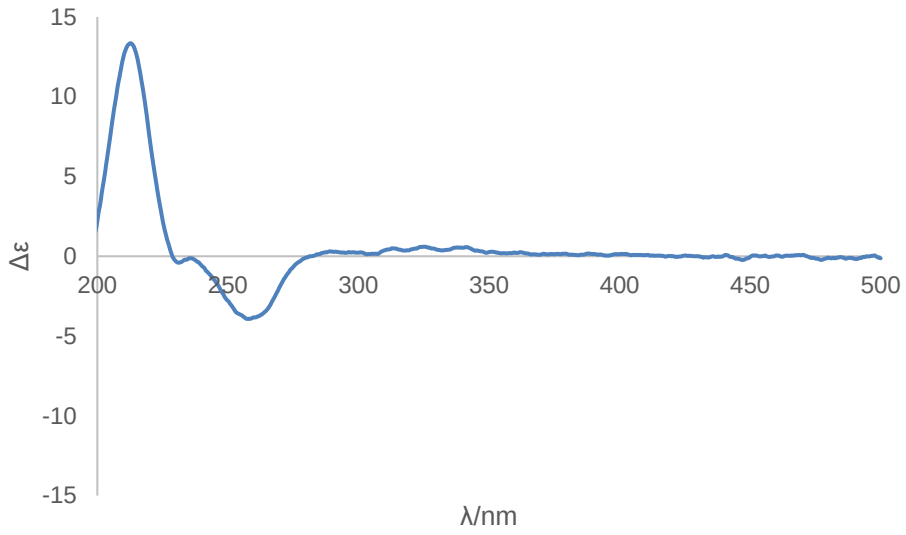


Table B9. ^{1}H-NMR (700 MHz, in methanol-d_4) and ^{13}C-NMR (176 MHz, in methanol-d_4) data of tunicamycin VII (**3.12**)

Position	δ_C	δ_H (*J* in Hz), type
2	152.7	
3		
4	166.2	
5	103.0	5.75 d (8.00), 1H
6	142.7	7.91 d (8.00), 1H
1'	89.7	5.92 d (5.65)
2'	75.4	4.19 t (5.80), 1H
3'	70.9	4.21 m, 1H
4'	88.2	3.85 m, 1H
5'	74.3	4.00 t (9.3), 1H
6'	35.9	2.11 m, 1H; 1.54, m, 1H
7'	72.5	3.77 d (10.20), 1H
8'	72.9	3.67 d (9.75), 1H
9'	73.2	3.65 d (9.75), 1H
10'	54.5	4.07 t (9.40), 1H
11'	102.0	4.58 d (8.45), 1H
1"	100.3	4.93 d (3.25), 1H
2"	54.9	3.87 m, 1H
3"	72.1	3.65 s, 1H
4"	73.3	3.34 t (10.00), 1H
5"	68.4	4.01 d (10.00), 1H
6"	63.1	3.81 m,1H; 3.69 m, 1H
2'"	169.7	
3'"	124.9	5.94 d (15.5), 1H
4'"	146.4	6.82 dt (15.5, 6.85), 1H
5'"	33.1	2.21 q (6.85), 1H
6'"-14'"	28.8-29.6	1.37-1.24 m, 16H
15'"	27.4	1.54 m, 1H
16'"-CH3	23.1	0.88 d (6.60), 6H
-COCH3	173.4	
-COCH3	23.3	1.93 s, 3H

Figure B31. ^{1}H NMR spectrum of tunicamycin VII (**3.12**) in methanol-d_4 at 700 MHz

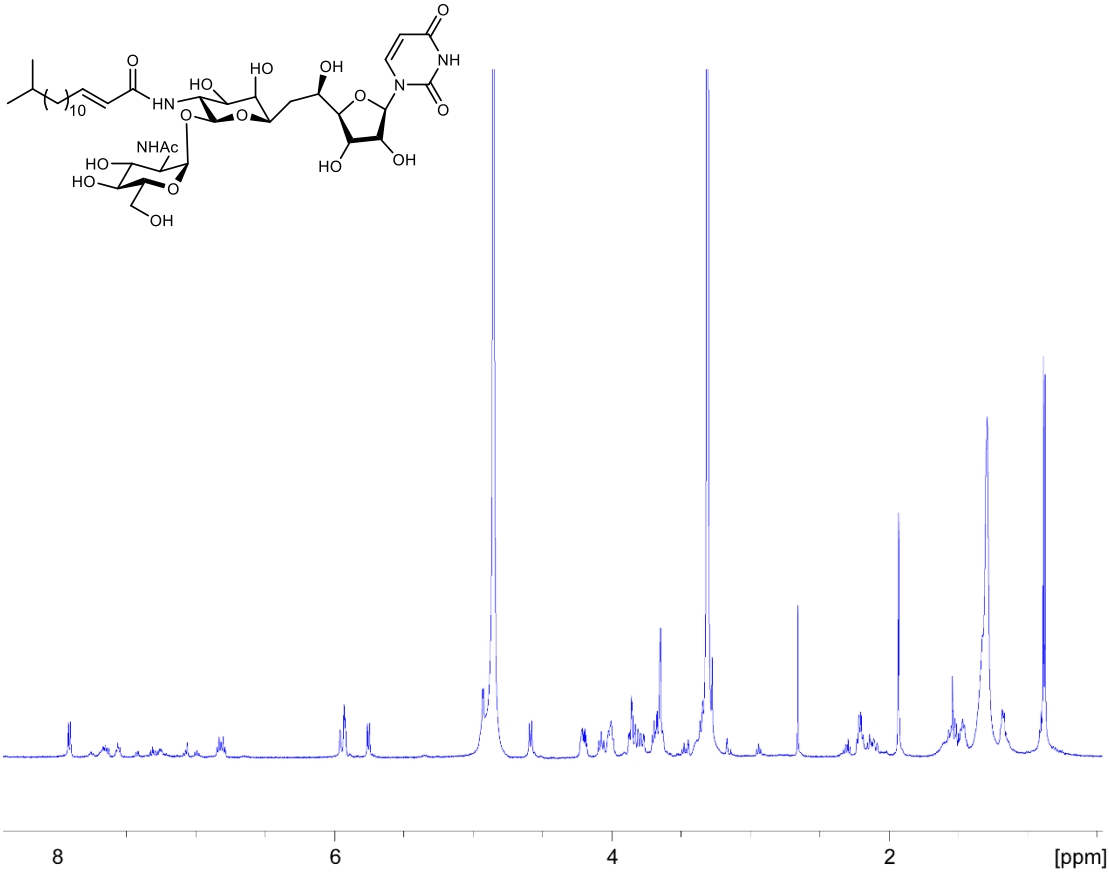


Figure B32. ^{1}H NMR spectrum of tunicamycin VIII (**3.13**) in methanol-d_4 at 700 MHz

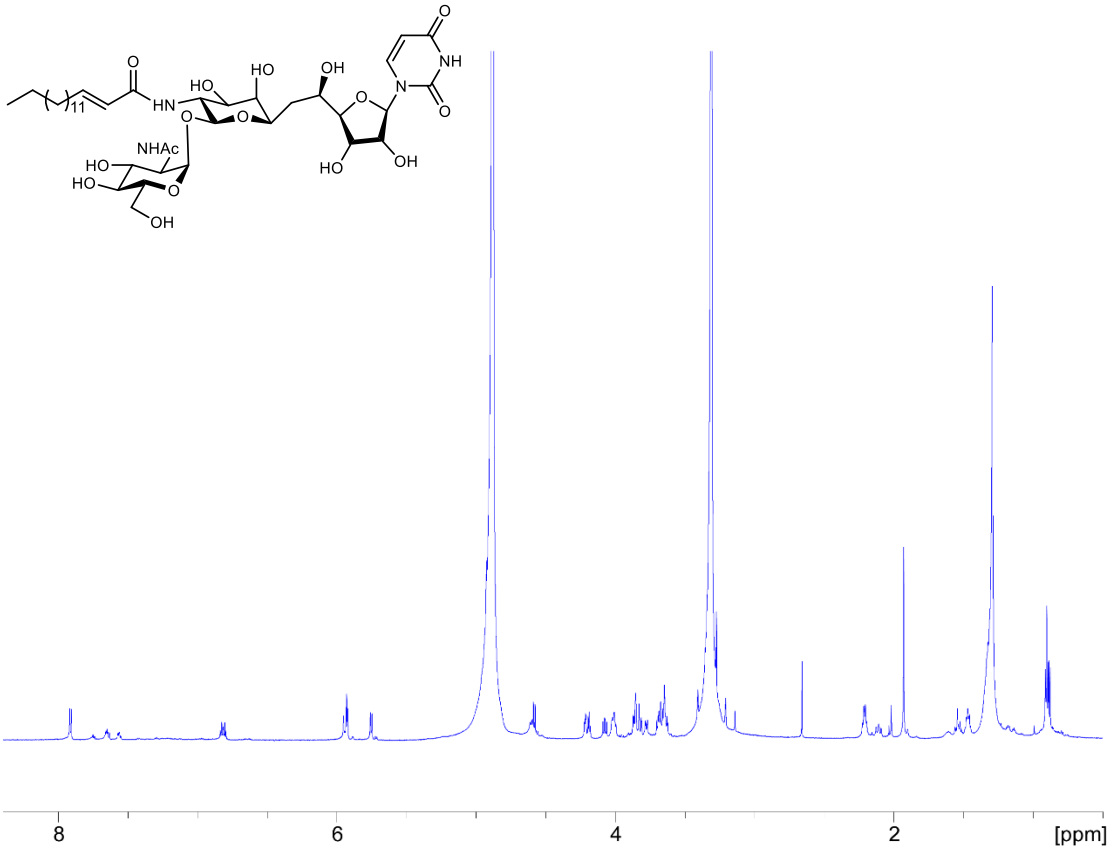


Table B10. ^{1}H-NMR (700 MHz, in methanol-*d4*) and ^{13}C-NMR (176 MHz, in methanol-*d4*) data of anthrabenzoxocinone (6*S*, 16*S*)1.264-C (**3.14**)

position	δ_C	δ_H (*J* in Hz), type
1	136.9	
2	112.0	6.27 dd (2.45, 0.56), 1H
3	157.9	
4	101.9	6.32 d (2.24), 1H
6	98.5	
7	40.7	3.33 d (18.00), 1H
		3.13 d (18.00), 1H
8	118.4	7.16 s, 1H
9	39.3	
10	107.2	6.75 d (2.24), 1H
11	166.2	
12	101.5	6.15 d (2.45), 1H
13	166.6	
14	191.7	
15	158.5	
16	65.74	6.24 s, 1H
4a	153.3	
7a	142.9	
8a	151.3	
9a	155.8	
13a	108.0	
14a	110.8	
15a	124.4	
16a	114.6	
1-CH3	19.3	2.44 s, 3H
6-CH3	27.7	1.64 s, 3H
9-2CH3	33.9	1.60 s, 3H
	33.8	1.68 s, 3H
13-OH		12.78 s, 1H
15-OH		13.47 s, 1H

Figure B33. ^{1}H NMR spectrum of anthrabenzoxocinone (6*S*, 16*S*)1.264C (**3.14**) in methanol-d_4 at 700 MHz

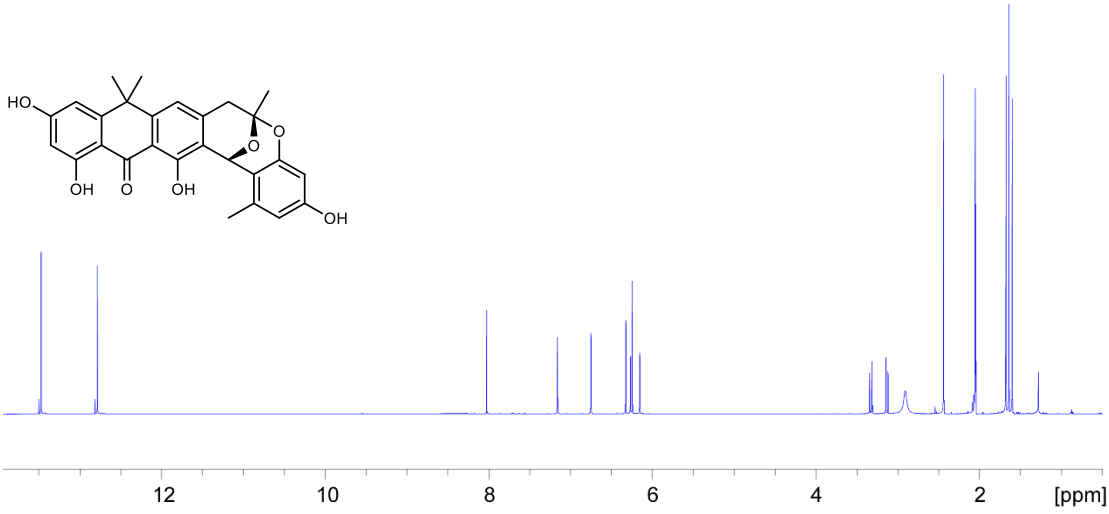


Figure B34. Experimental ECD spectrum of isolated anthrabenzoxocinone (6*S*, 16*S*) 1.264C (**3.14**)

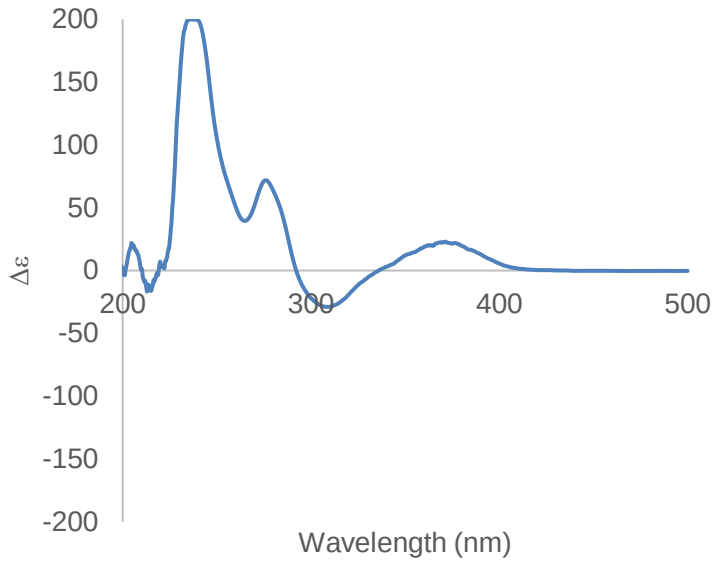


Figure B35. ^{1}H NMR spectrum of tetrapeptide (**3.2**) in methanol-d_4 at 700 MHz

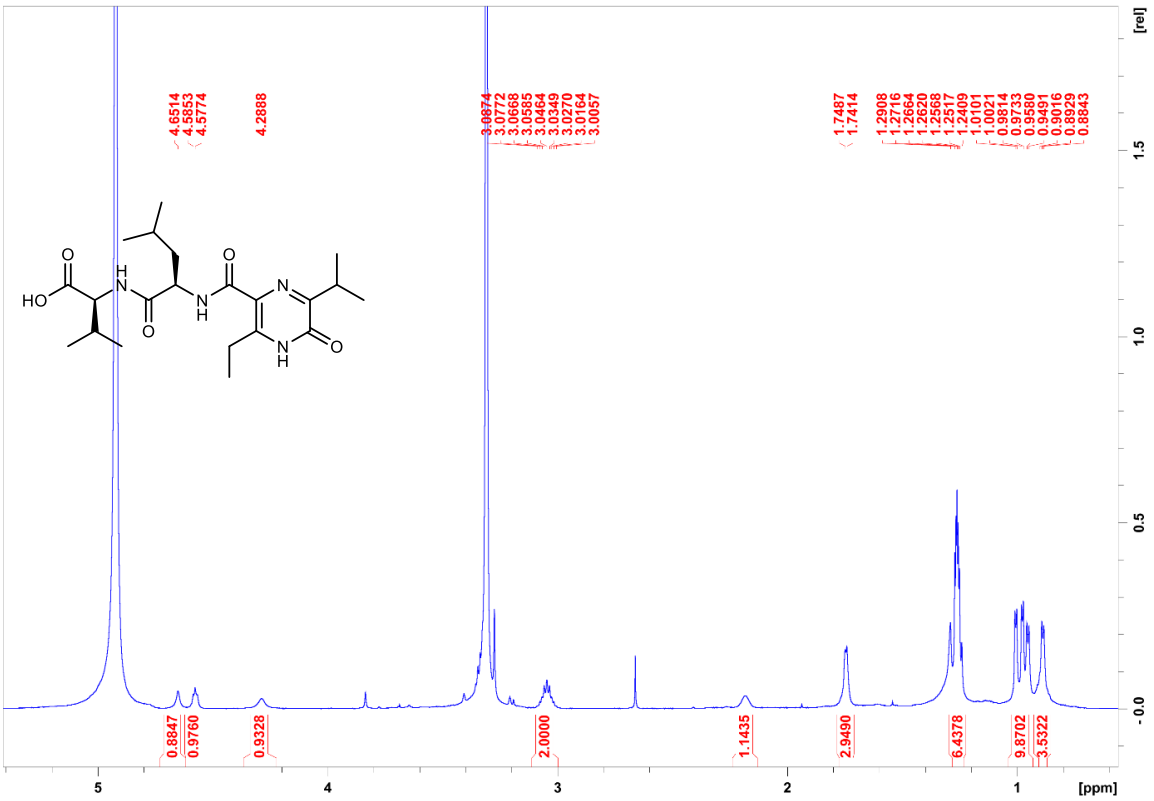


Figure B36. ^{1}H NMR spectrum of tetrapeptide (**3.2**) in DMSO-*d6* at 700 MHz

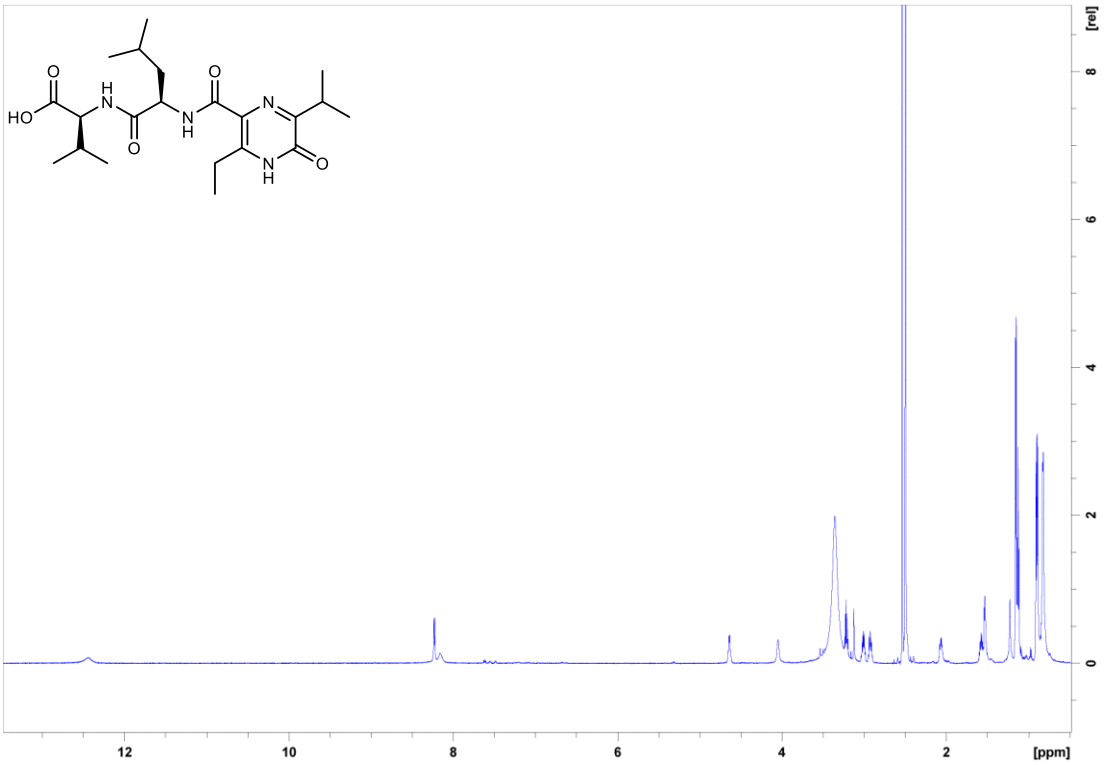


Figure B37. ^{1}H NMR spectrum, 13.5 to 3.7 ppm expansion, of tetrapeptide (**3.2**) in DMSO-d_6 at 700 MHz

Figure B38. ^{1}H NMR spectrum, 3.7 to 0.5 ppm expansion, of tetrapeptide (**3.2**) in DMSO-d_6 at 700 MHz

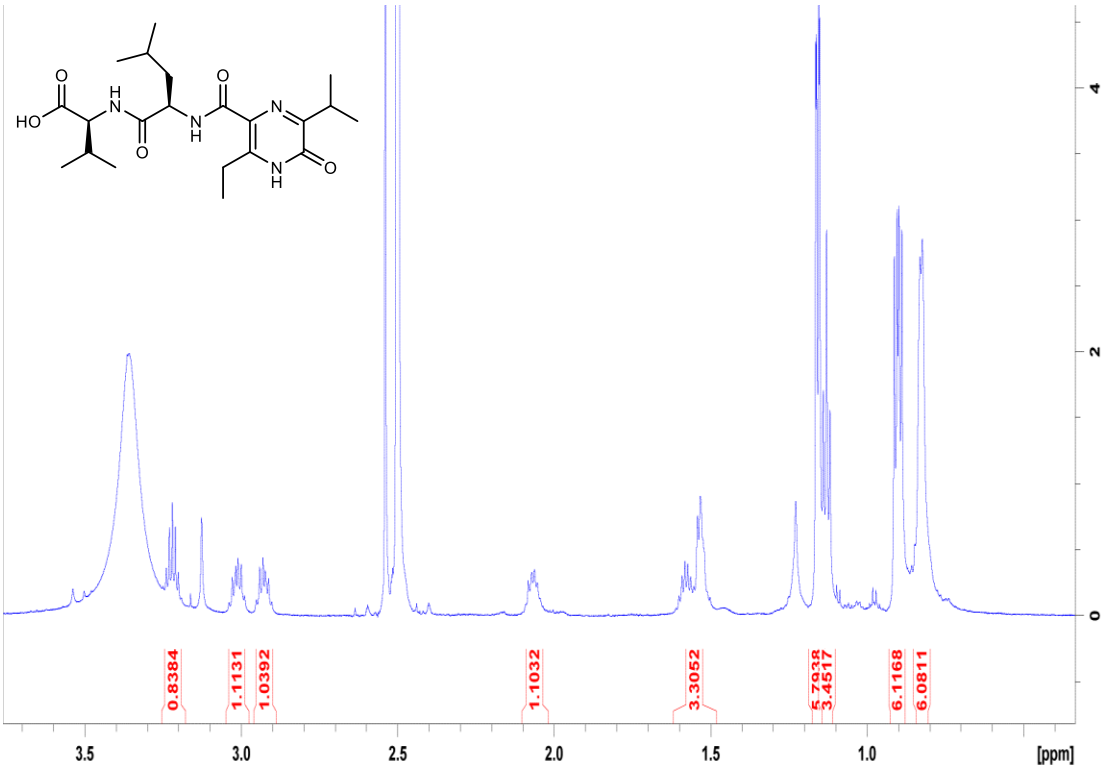


Figure B39. ^{1}H-^{1}H COSY NMR spectrum of tetrapeptide (3.2) in DMSO-d_6 at 700 MHz

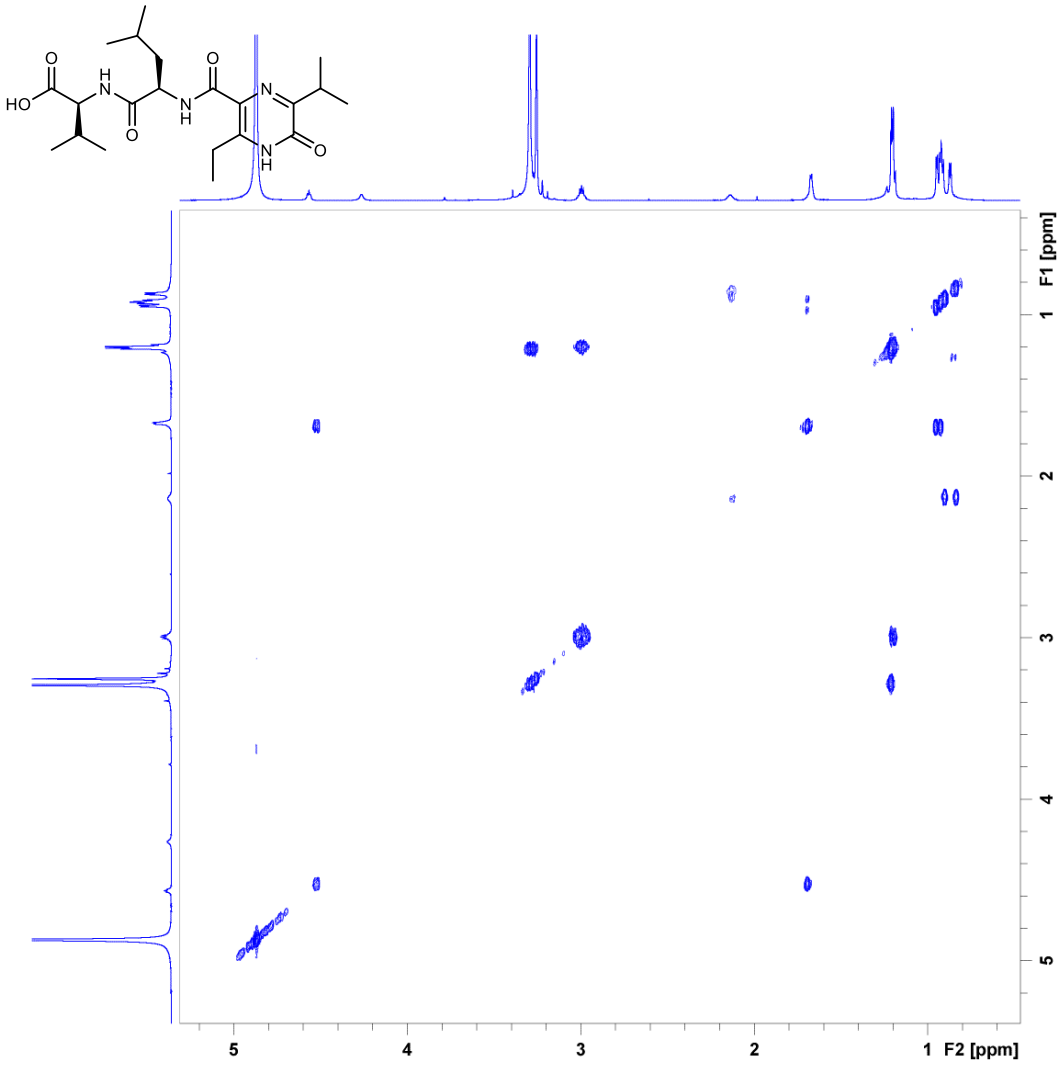


Figure B40. ^{1}H-^{13}C HSQC NMR spectrum of tetrapeptide (**3.2**) in DMSO-d_6 at 700 MHz and 176 MHz

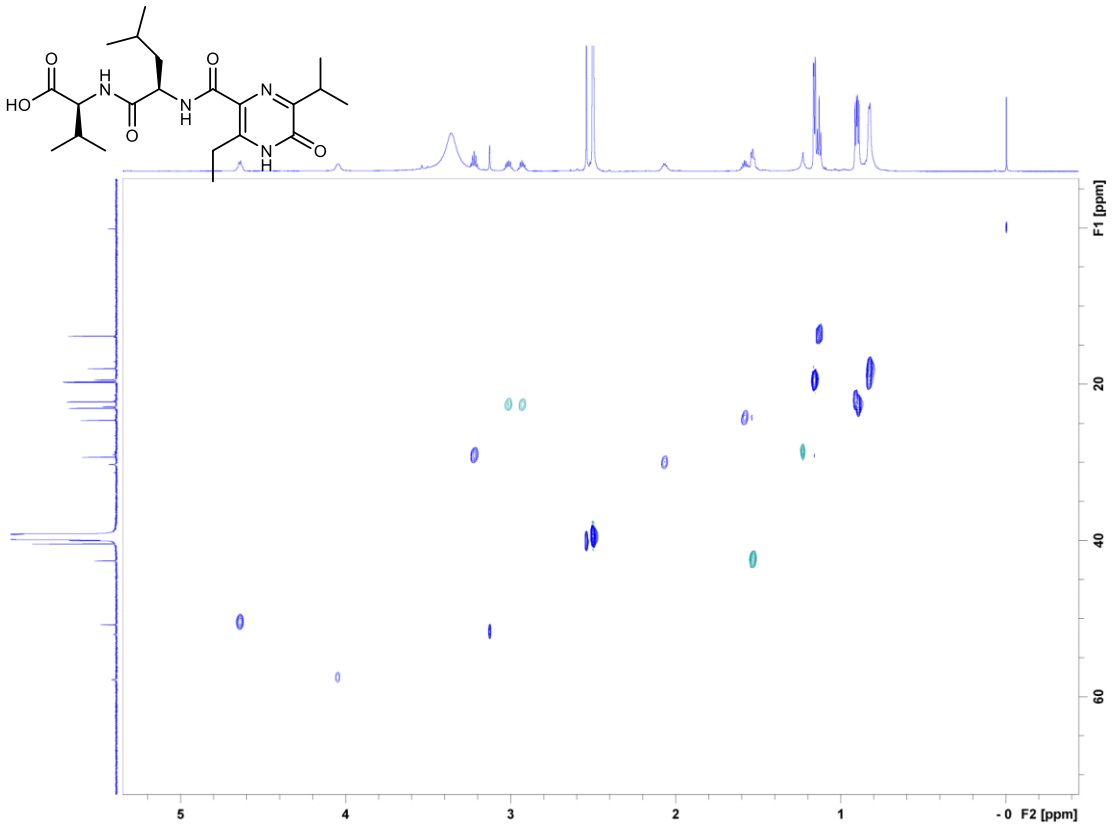


Figure B41. ^{1}H-^{13}C HMBC NMR spectrum of tetrapeptide (**3.2**) in DMSO-d_6 at 700 MHz (J_{CH} 8Hz)

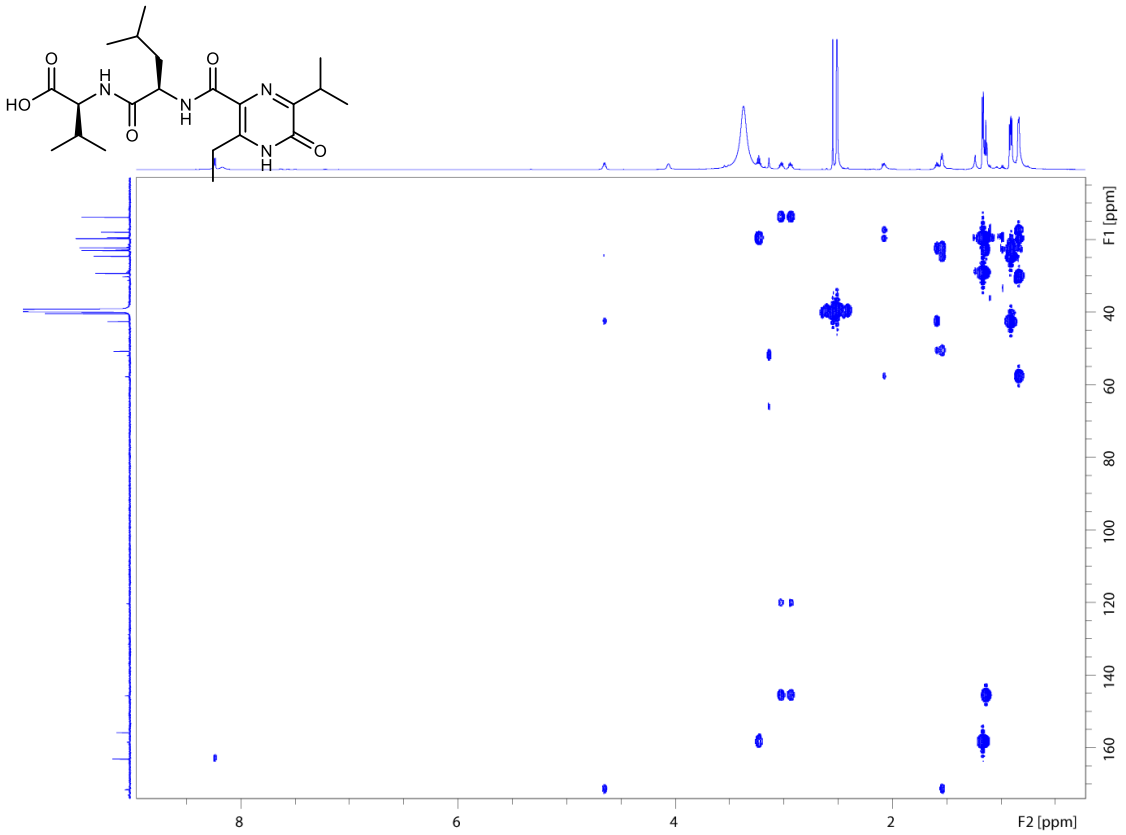


Figure B42. ^{13}C NMR spectrum of tetrapeptide (**3.2**) in DMSO-d_6 at 176MHz

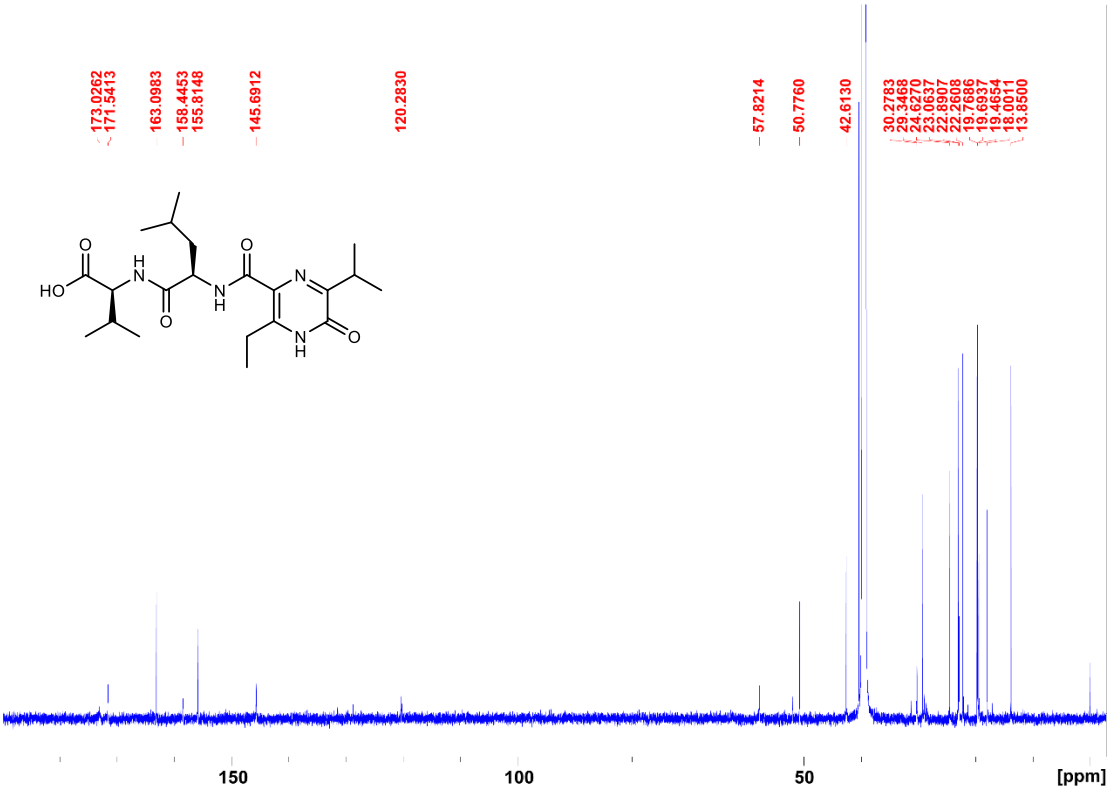


Figure B43. IR spectrum of tetrapeptide (**3.2**)

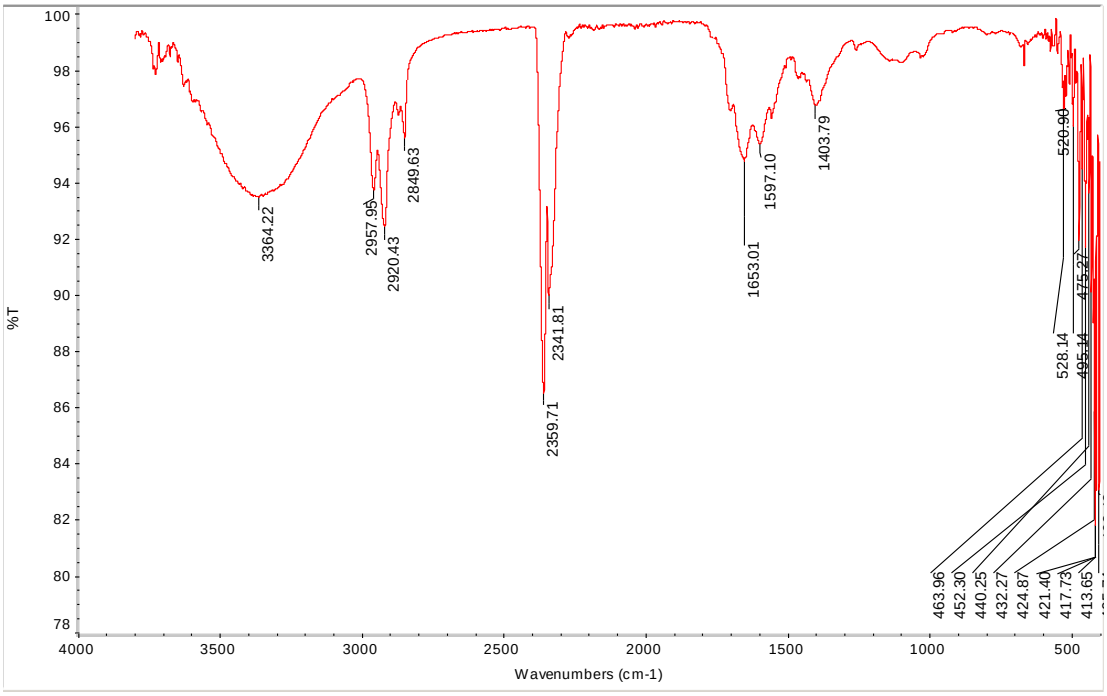


Figure B44. ^{1}H NMR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) in $CDCl_3$ at 700 MHz

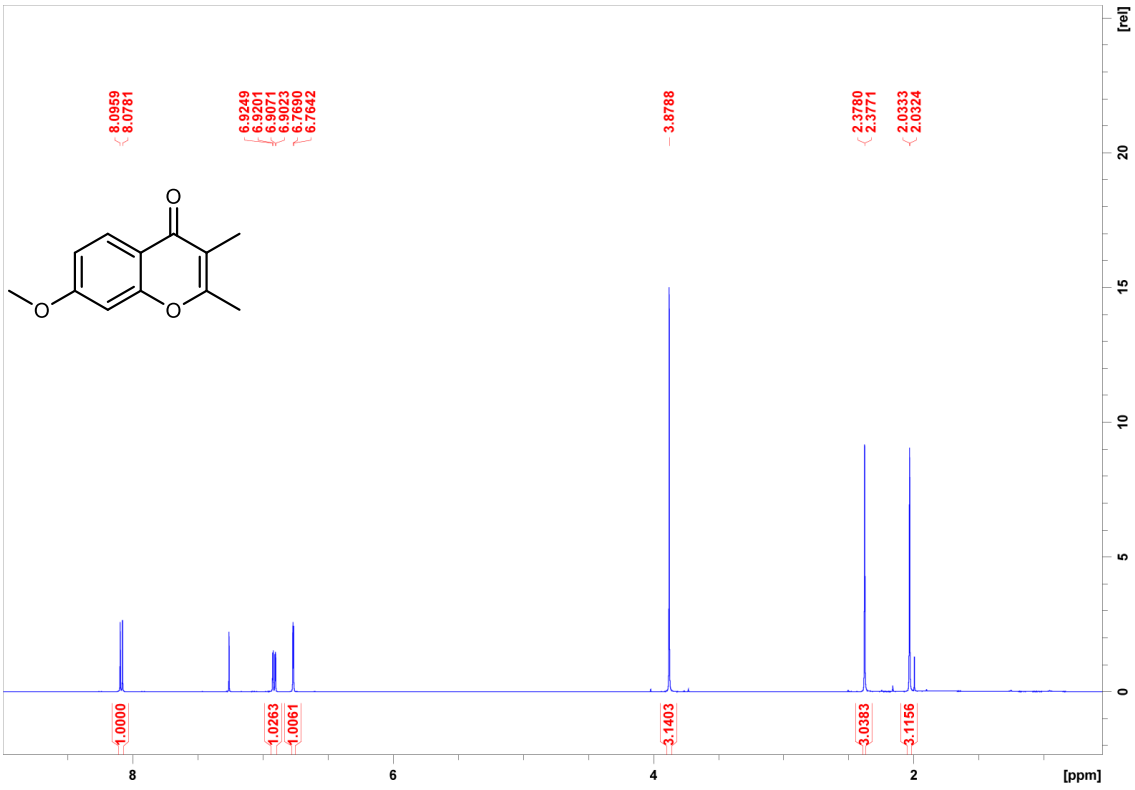


Figure B45. ^{1}H-^{1}H COSY NMR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) in $CDCl_3$ at 700 MHz

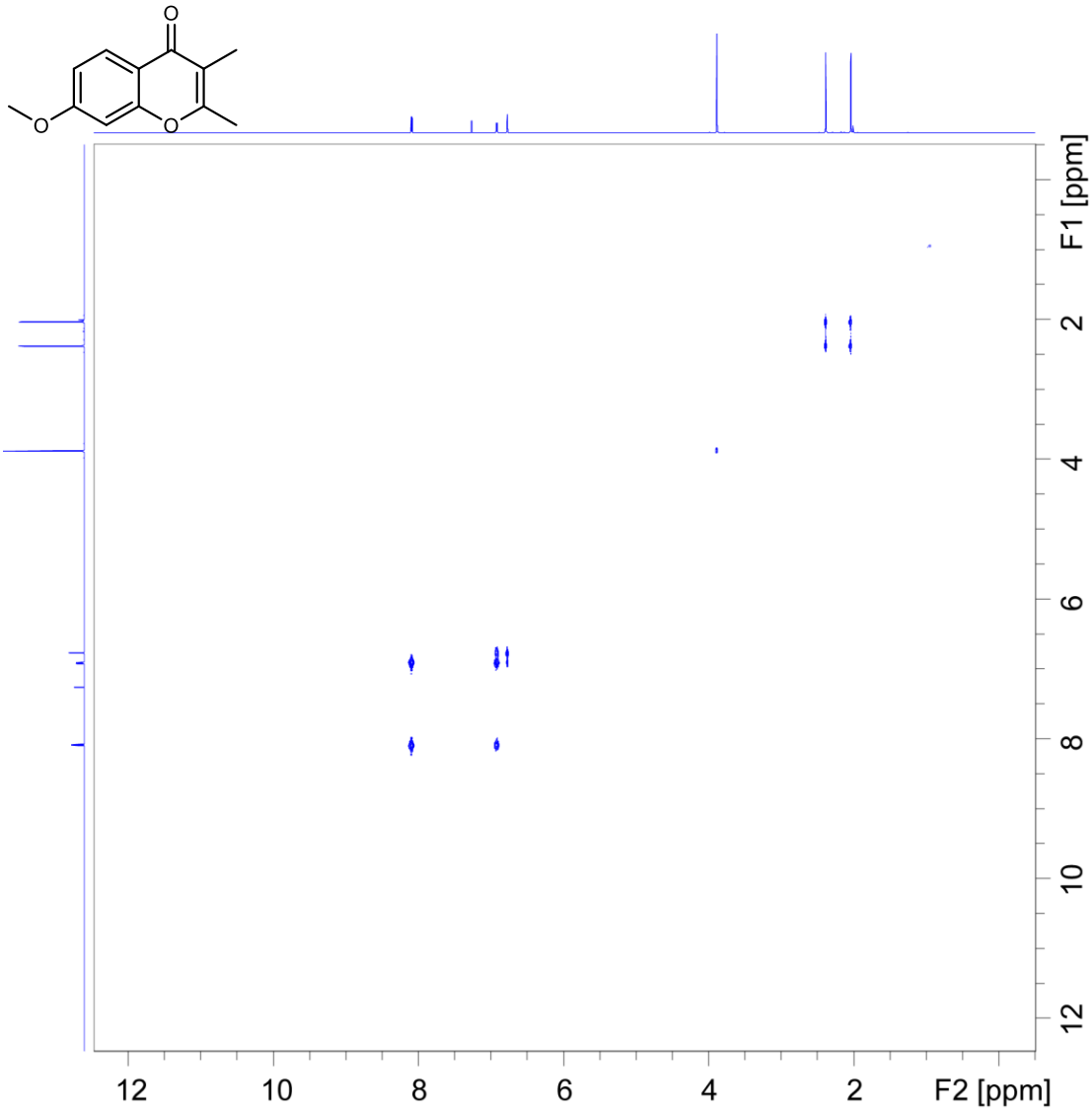


Figure B46. ^{1}H-^{13}C HSQC NMR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) in $CDCl_3$ at 700 MHz and 176 MHz

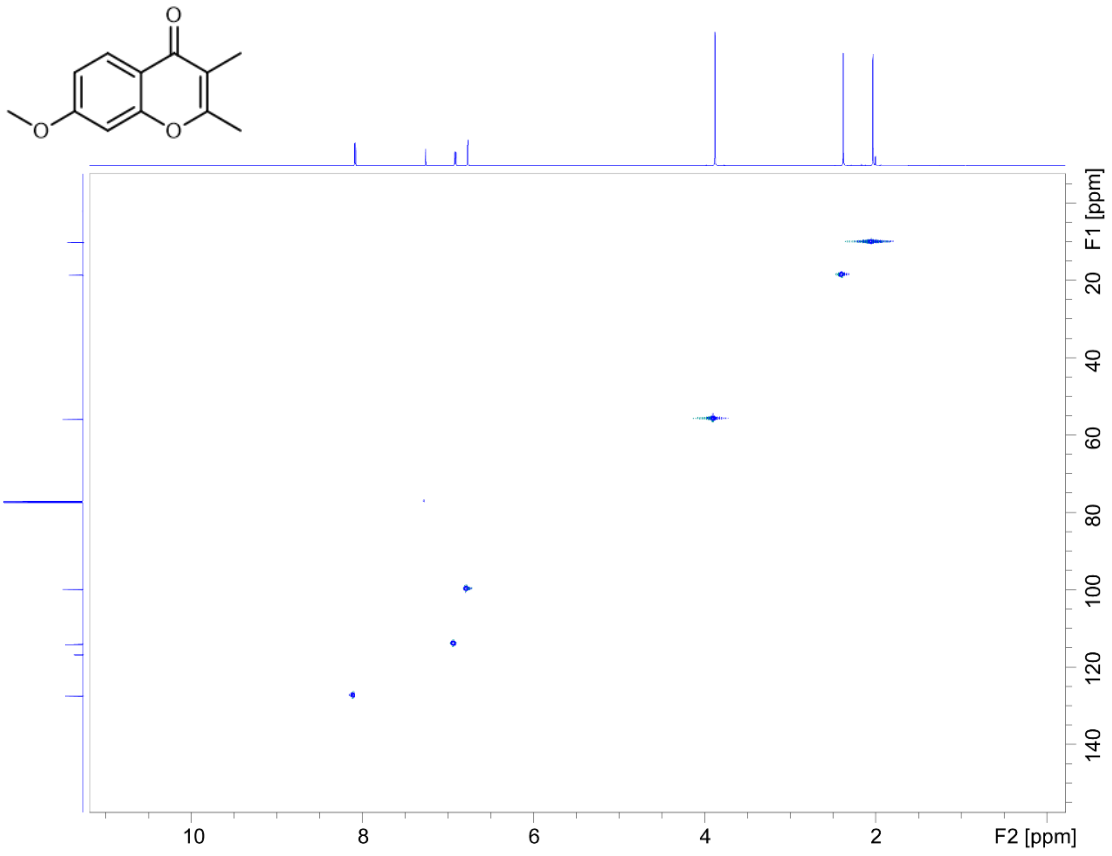


Figure B47. ^{1}H-^{13}C NMR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) in $CDCl_3$ at 700 MHz and 176 MHz

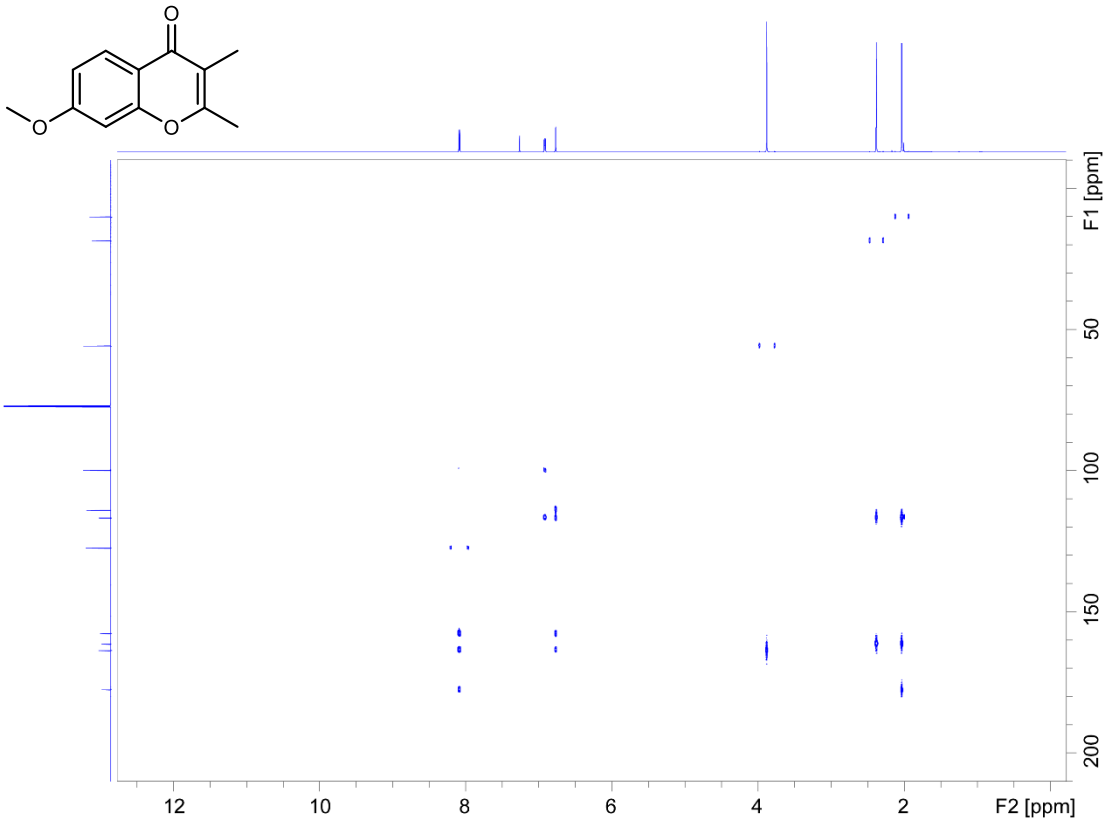


Figure B48. ^{13}C NMR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**) in $CDCl_3$ at 176 MHz

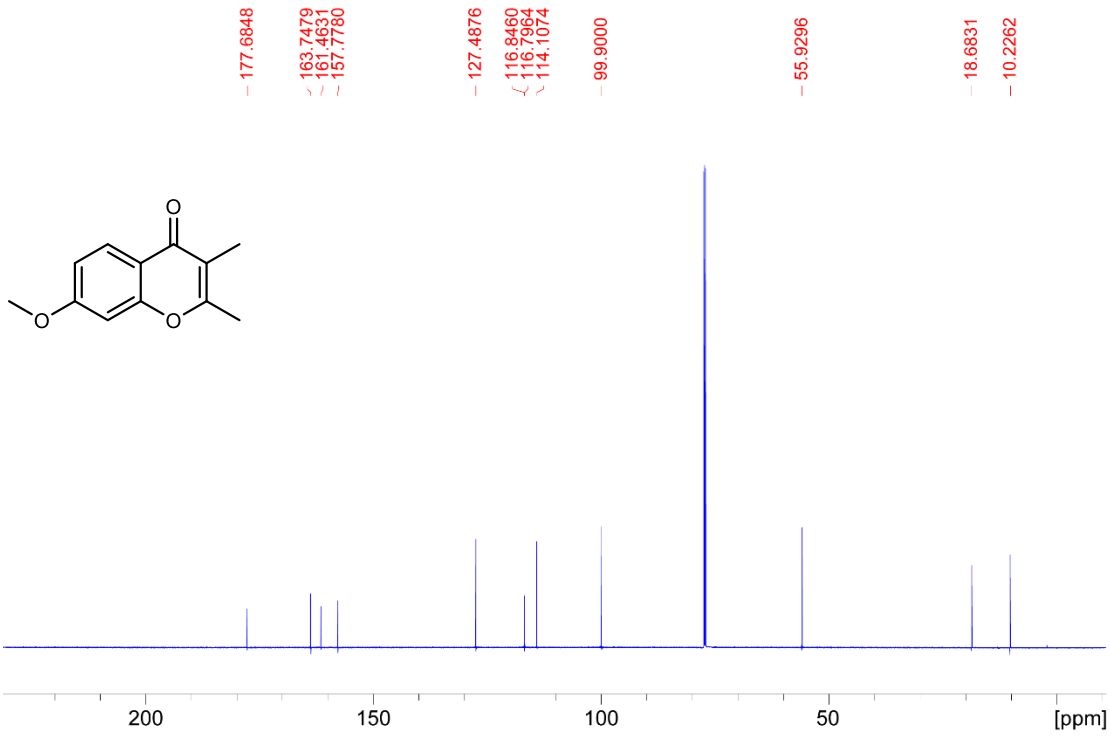
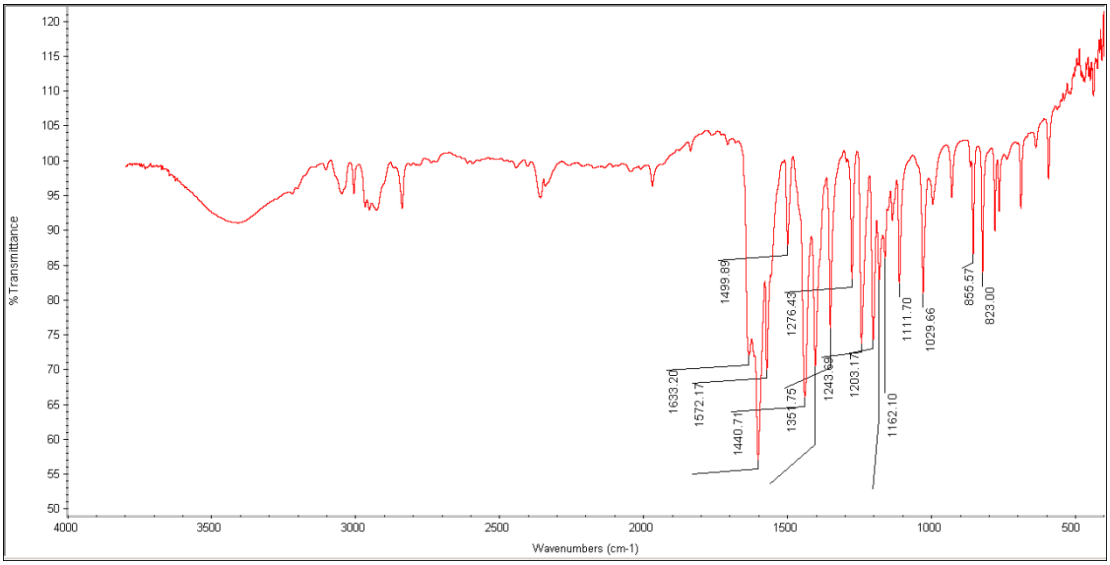


Figure B49. IR spectrum of 7-methoxy-2,3-dimethyl-4*H*-chromen-4-one (**3.11**)



Appendix C: Supporting Information for Chapter Four

Figure C1. 16S rRNA sequence of bacterium *Streptomyces sp.* (CL12-4)

ACATGCAAGTCGAACGATGAACCACTTCGGTGGGGATTAGTGGCGAACGGGT
GAGTAACACGTGGGCAATCTGCCCTTCACTCTGGGACAAGCCCTGGAAACGG
GGTCTAATACCGGATACTGACCGGCCTGGGCATCCAGGCGGGTCGAAAGCTCC
GGCGGTGAAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGTAATGGCTC
ACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTGGG
ACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCAC
AATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACGGCCTTCGG
GTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGACGGTACCTGCAGAA
GAAGCGCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGCGCAAG
CGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGCTTGTCACGTCGG
TTGTGAAAGCCCGGGGCTTAACCCCGGGTCTGCAGTCGATACGGGCAGGCTA
GAGTTCGGTAGGGGAGATCGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATA
TCAGGAGGAACACCGGTGGCGAAGGCGGATCTCTGGGCCGATACTGACGCTG
AGGAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGC
CGTAAACGGTGGGCACTAGGTGTGGGCAACATTCCACGTTGTCCGTGCCGCA
GCTAACGCATTAAGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTC
AAAGGAATTGACGGGGGCCCGCACAAGCGGCGGAGCATGTGGCTTAATTCGA
CGCAACGCGAAGAACCTTACCAAGGCTTGACATACACCGGAAAGCATTAGAG
ATAGTGCCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTC
GTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCCGTG
TTGCCAGCAGGCCCTTGTGGTGCTGGGGACTCACGGGAGACCGCCGGGGTCA
ACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGG
CTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCGCGAGGTGGA
GCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACTCGACCCC
ATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCATTGCTGCGGTGAATACGT
TCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTCGGTAACACCCG
AAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGC

Figure C2. (A) Bacterium *Streptomyces* sp. CL12-4 on M2 agar plate. (B) Phylogenetic analysis of bacterium CL12-4. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.05141921 is shown. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 19 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1349 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

(A)

(B)

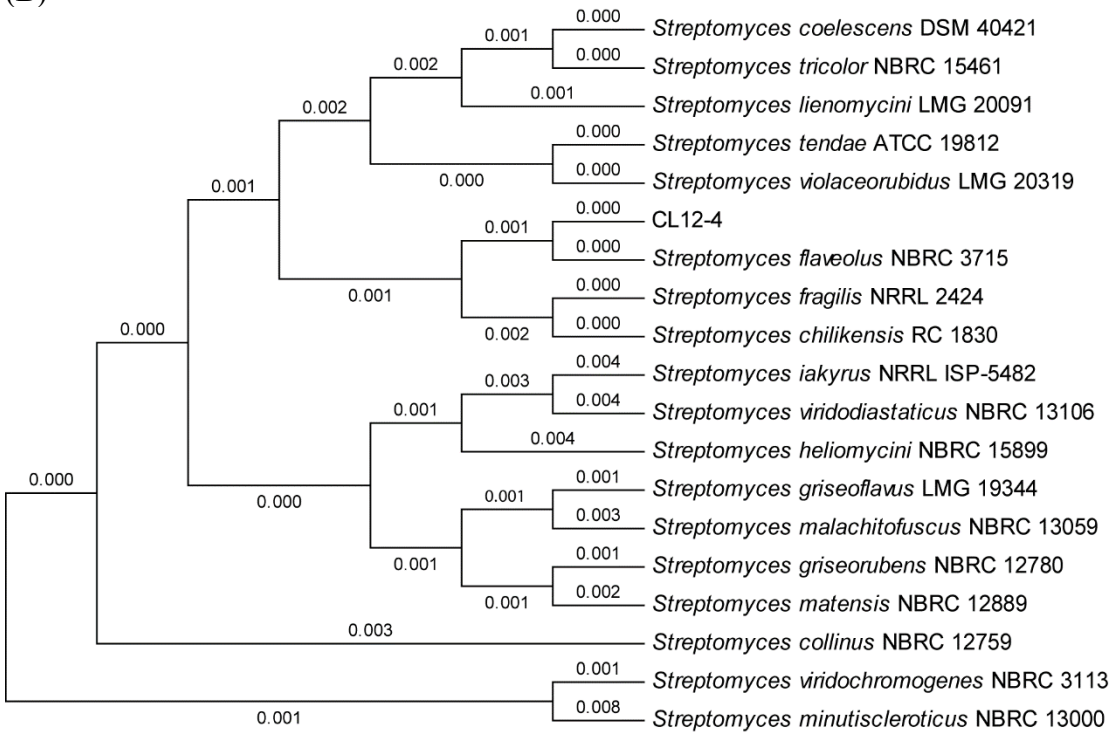


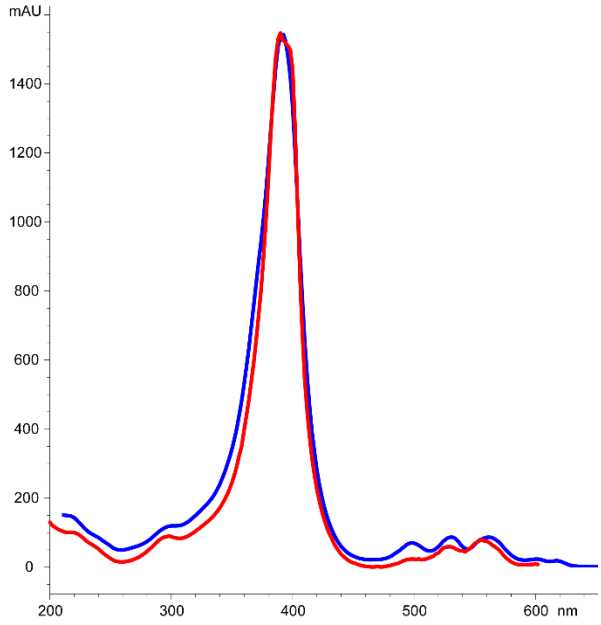
Table C1. Results of single dose antibacterial assay for organic extract of *Streptomyces* sp. CL12-4 at 125 µg/mL; vancomycin (125 µg/mL) and chloramphenicol (125 µg/mL) as positive control.

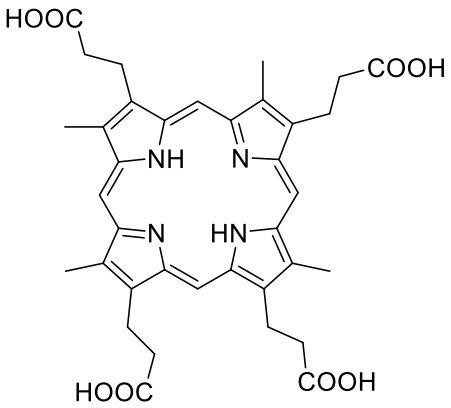
Sample	Antibacterial assay			
	E. faecium	MRSA	*P. aeruginosa*	*E. coli*
CL12-4	98.2 ± 0.1	102.0 ± 9.4	84.9 ± 1.3	89.3 ± 3.7
Vancomycin	1.8 ± 1.0	0.1 ± 0.2	-	-
Chloramphenicol	-	-	0 ± 0.1	0.5 ± 0.3

Table C2. Results of single dose MTT-based cell viability assay for organic extract of *Streptomyces* sp. CL12-4 at 10 µg/mL; etoposide (250 µM) as positive control.

Sample	Cytotoxicity assay				
	SK-Mel-5	PC-3	HCT-116	A549	MCF-7
CL12-4	7.9 ± 0.7	46.3 ± 2.3	28.4 ±0.3	37.6 ± 0.6	46.2 ± 2.0
Etoposide	6.7 ± 0.4	36.2 ± 2.9	13.7 ± 0.6	14.9 ± 0.7	41.6 ± 2.4

Figure C3. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of **4.1**





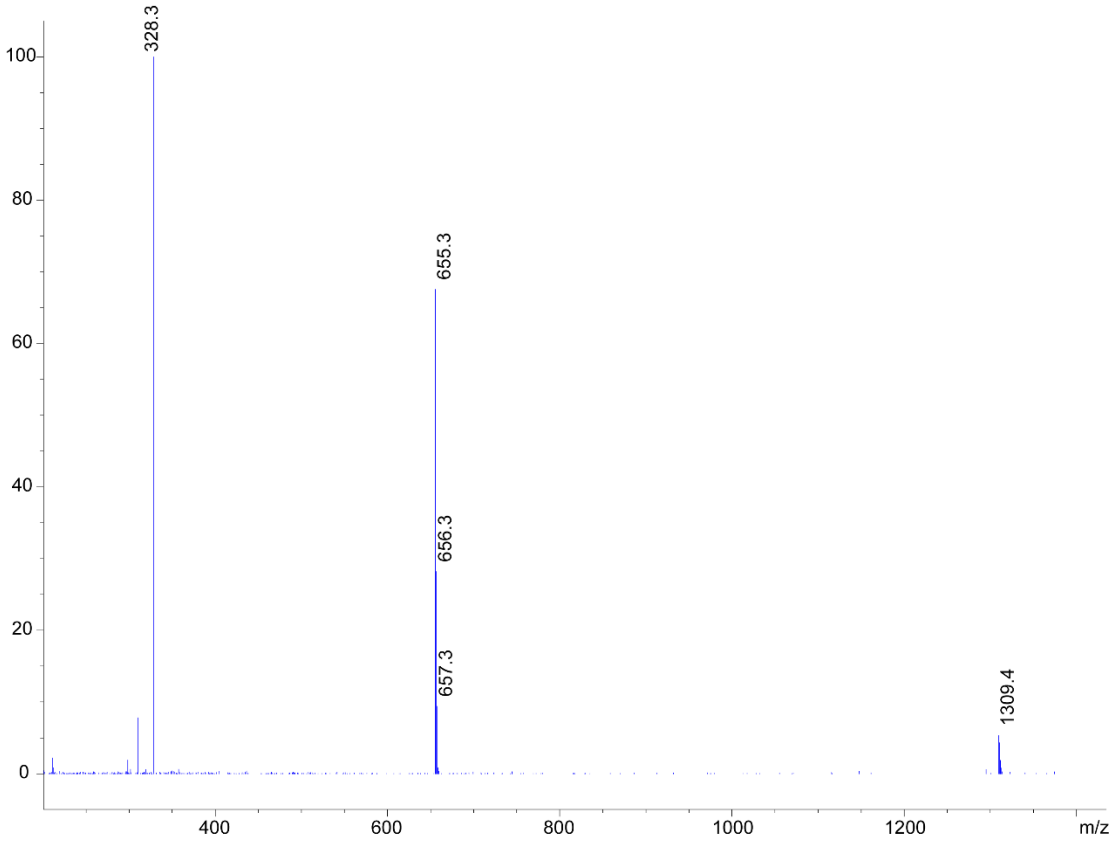
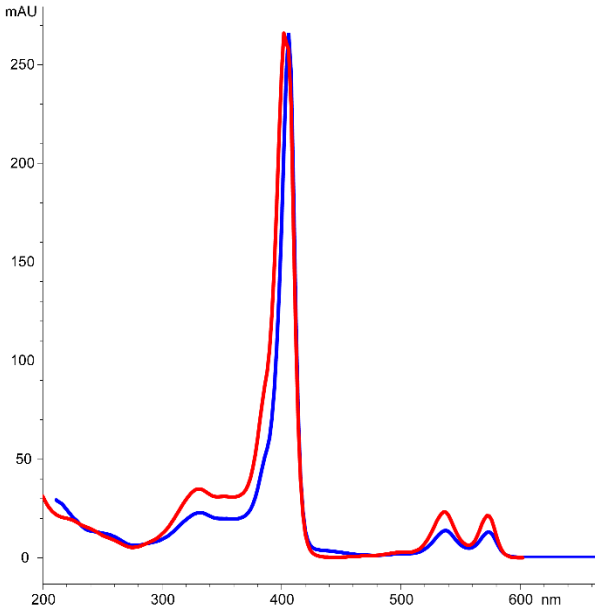
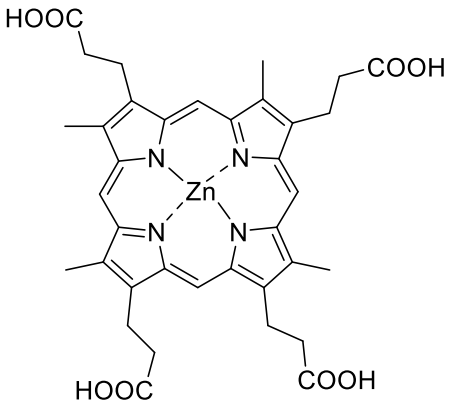


Figure C4. UV spectra (experimental result in blue, in-house UV library data in red) and low-resolution mass spectrum of **4.2**





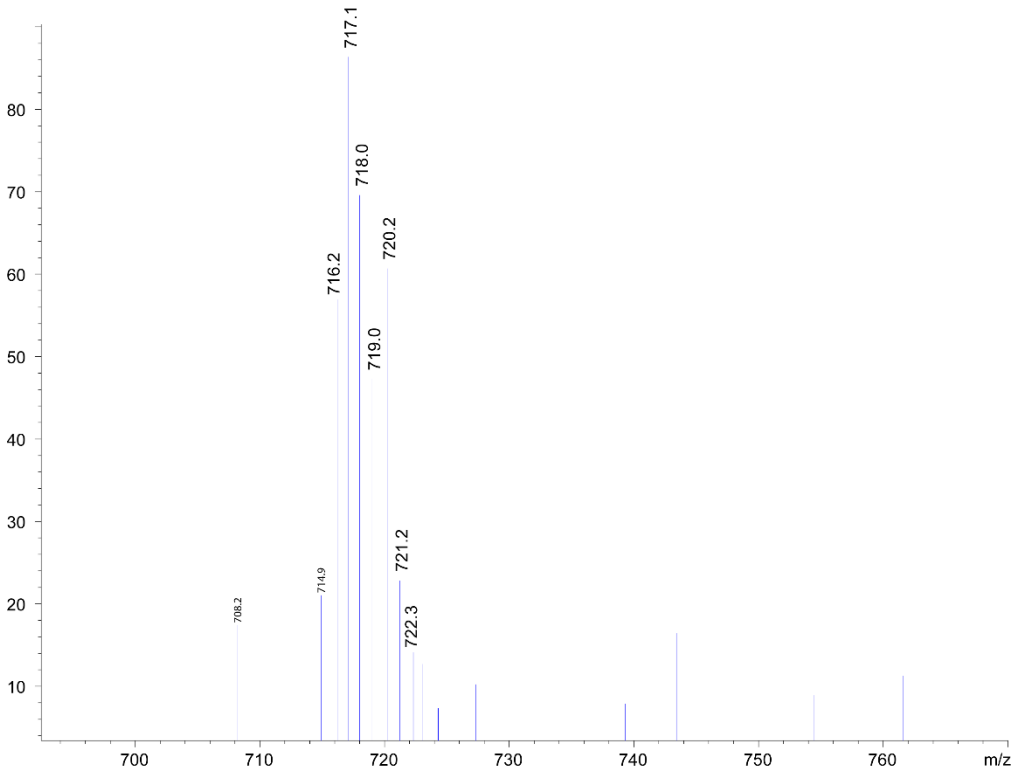


Figure C5. ^{1}H NMR spectrum of coproporphyrin (**4.1**) in methanol-d_4 at 500 MHz

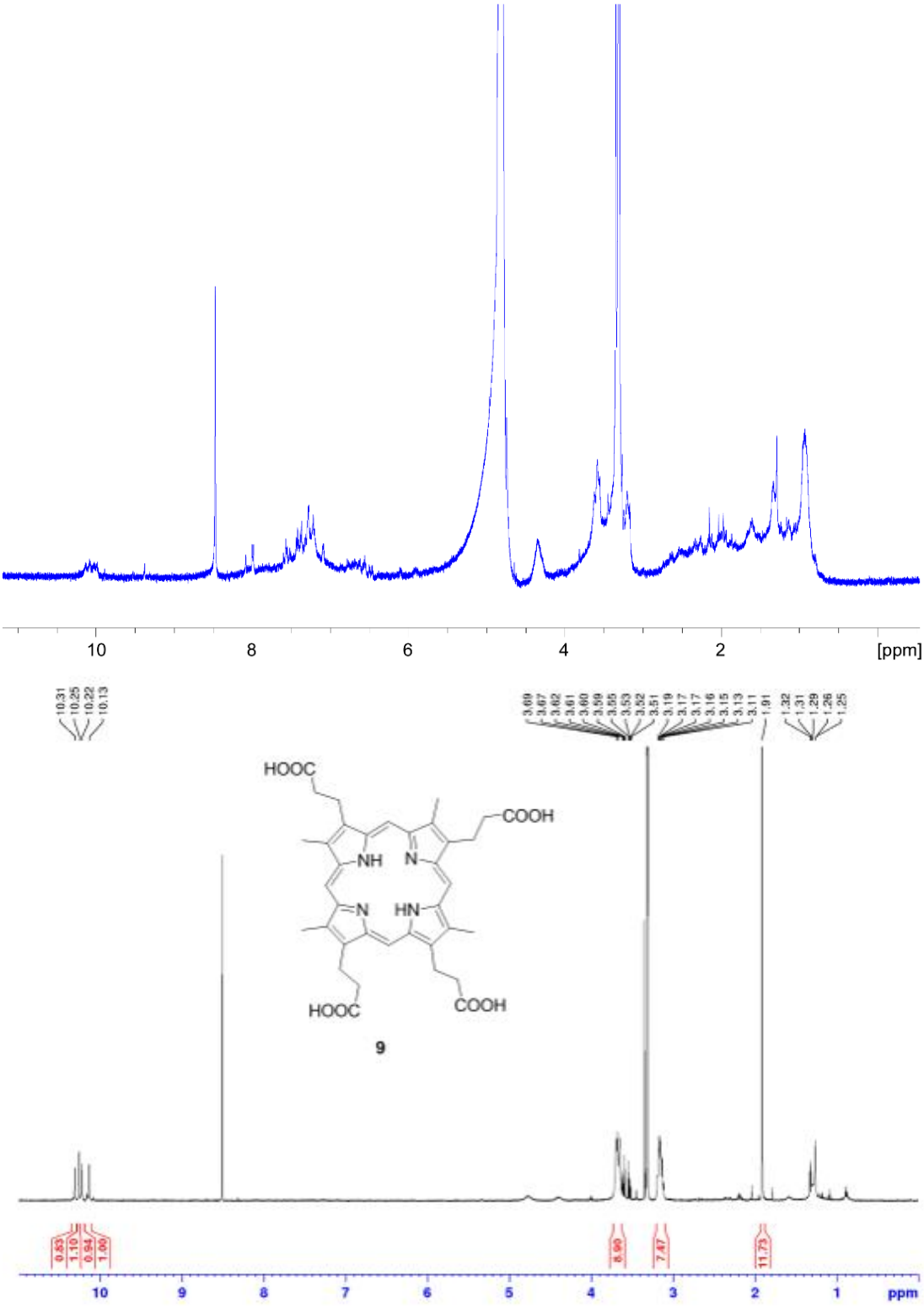


Figure C6. ^{1}H NMR spectrum of Zincphyrin (**4.2**) in methanol-d_4 at 500 MHz

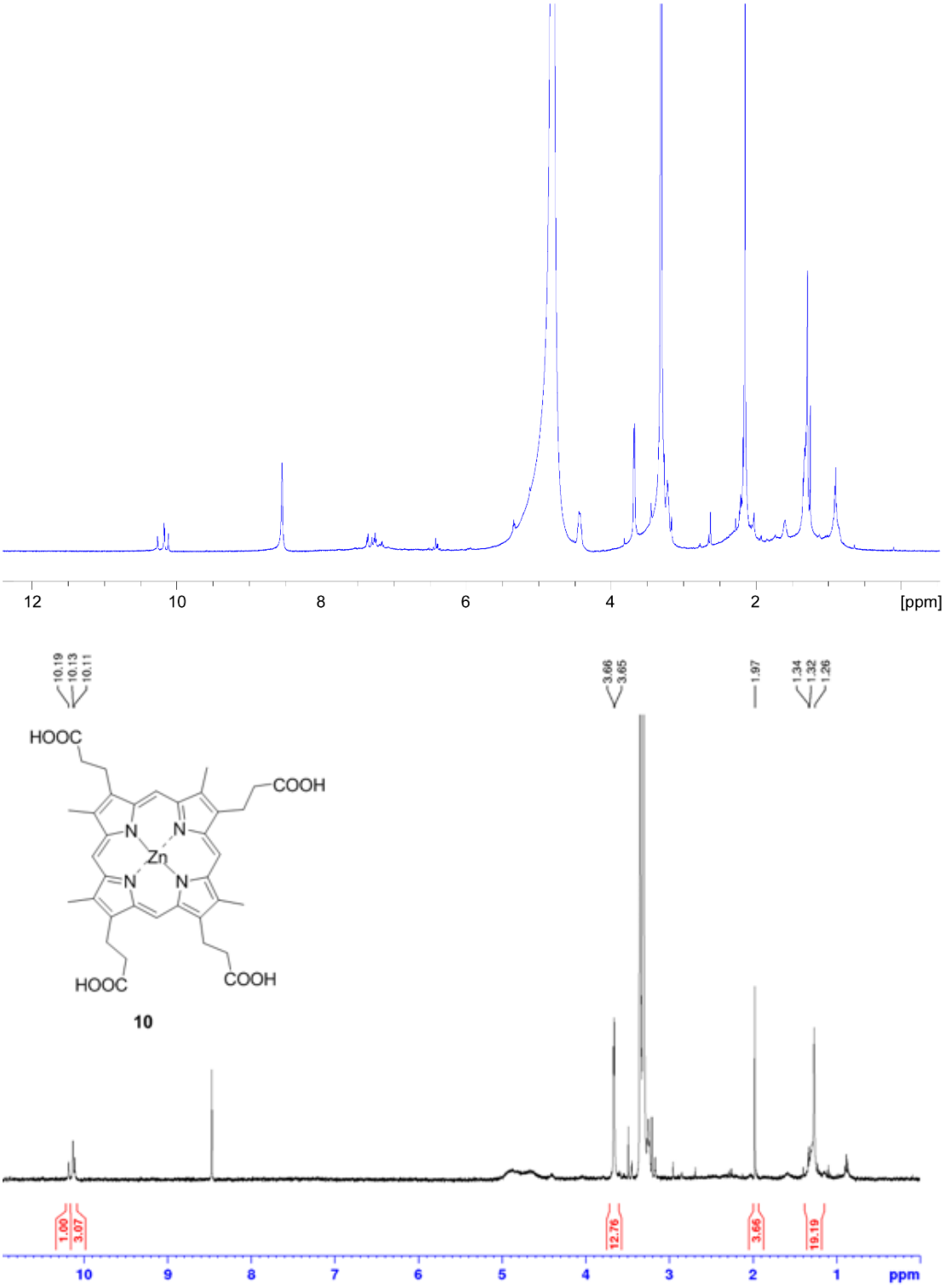


Figure C7. UV spectrum and low-resolution mass spectrum of **4.3**

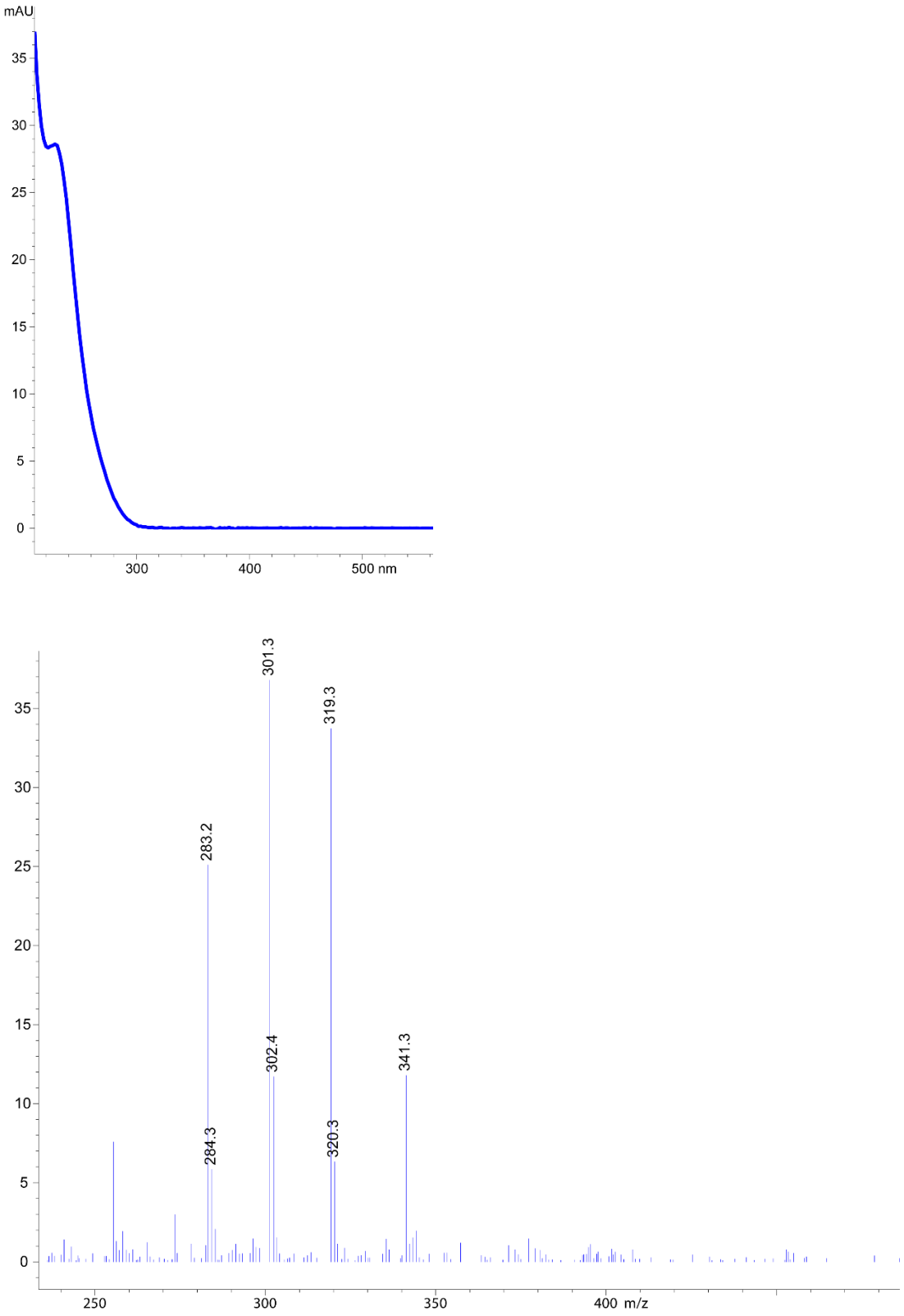


Figure C8. IR spectrum of **4.3**

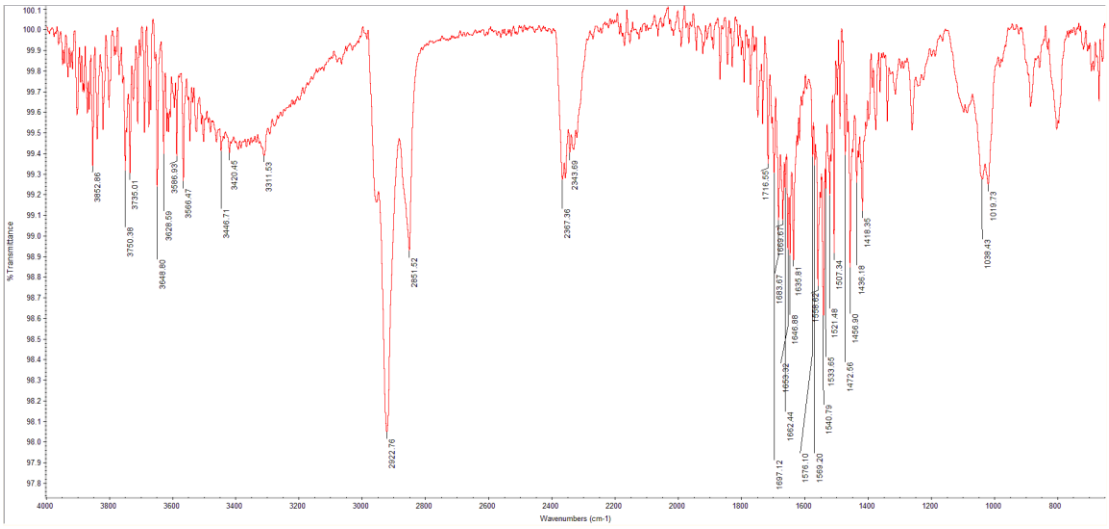


Figure C9. ^{1}H NMR spectrum of **4.3** in methanol-d_4 at 700 MHz

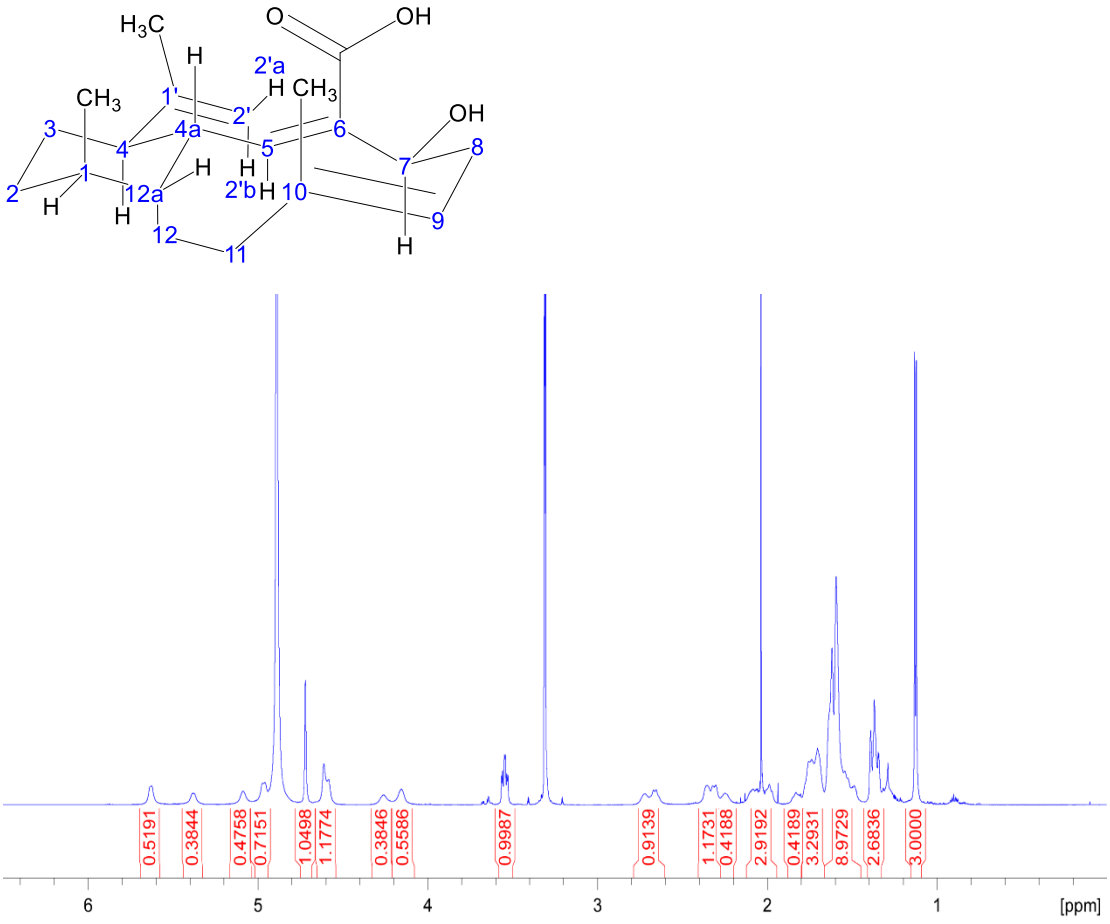


Figure C10. ^{1}H NMR spectrum of **4.3** (6.0-3.0 ppm) in methanol-d_4 at 700 MHz

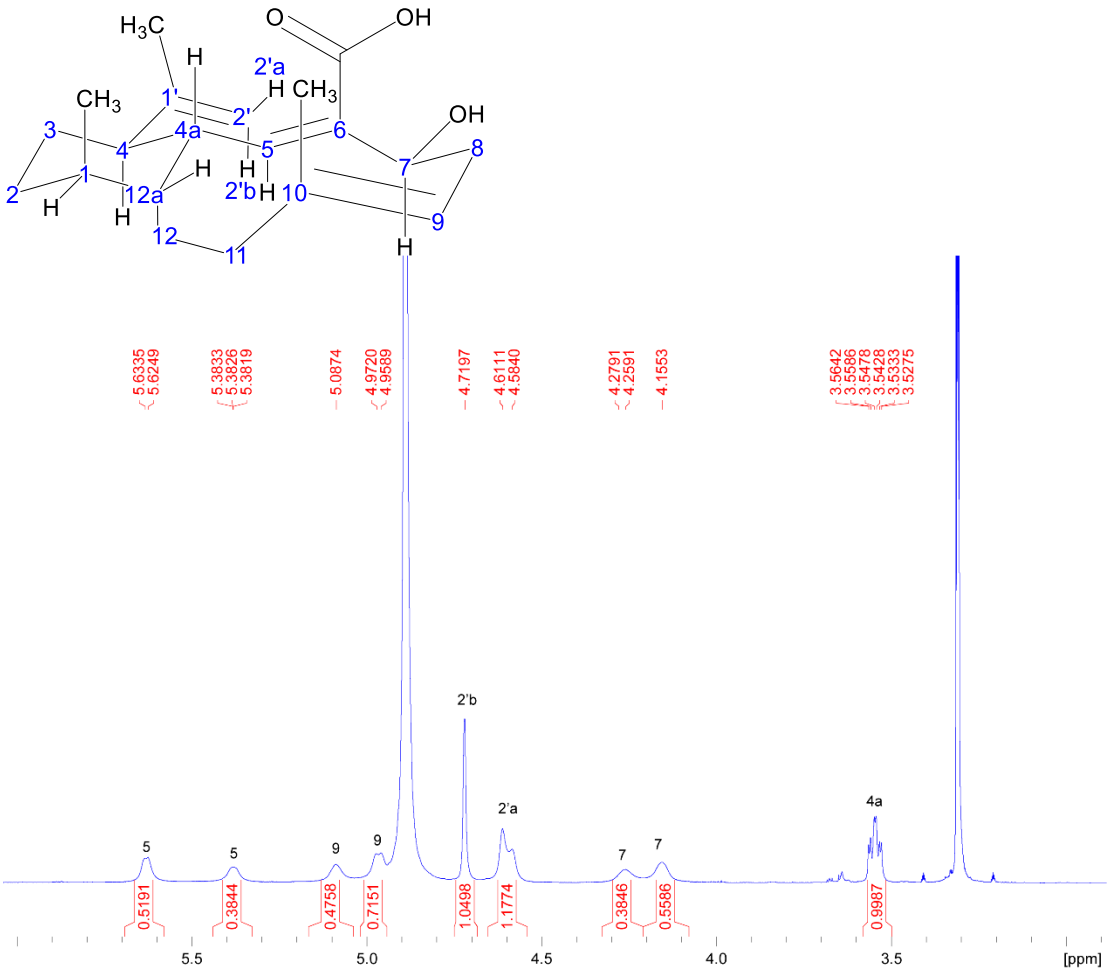


Figure C11. ^{1}H NMR spectrum of **4.3** (3.0-0.5 ppm) in methanol-d_4 at 700 MHz

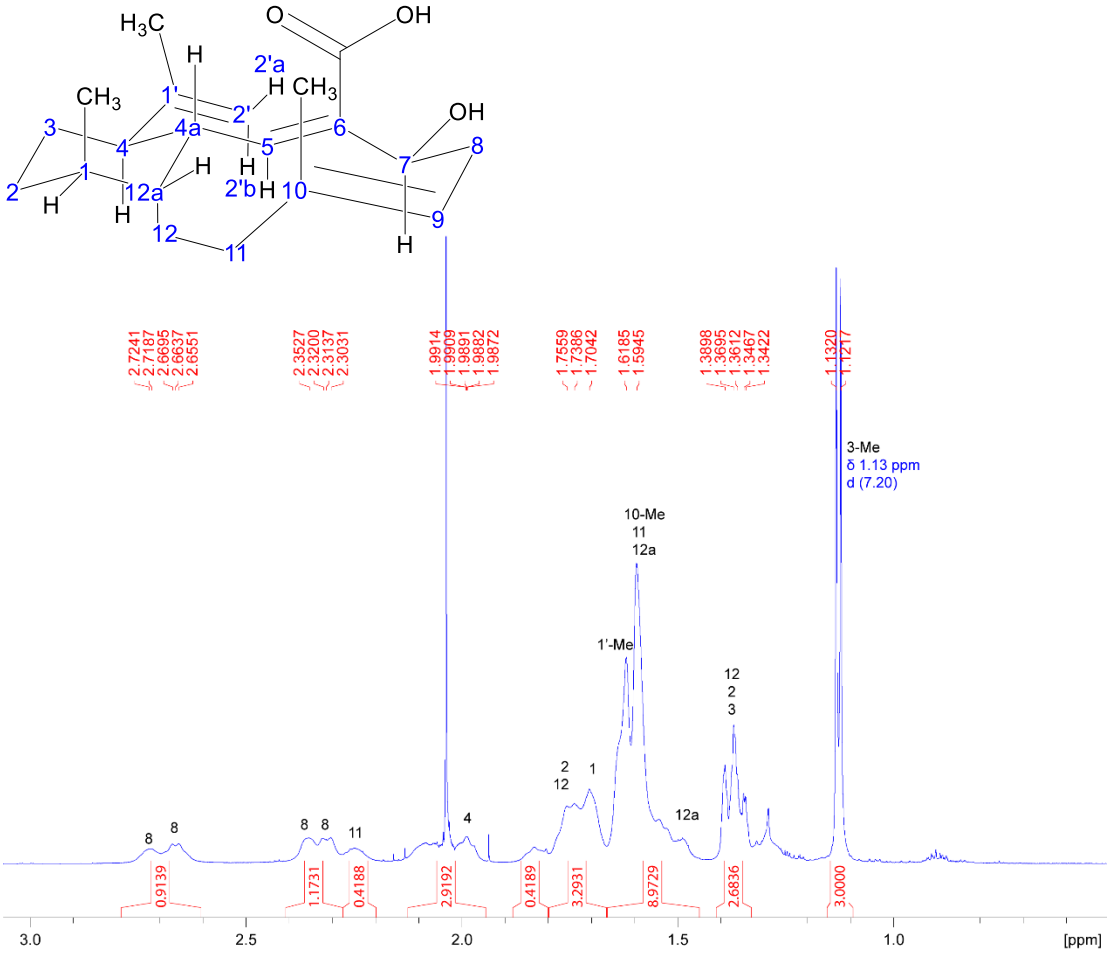


Figure C12. ^{13}C NMR spectrum of **4.3** in methanol-d_4 at 176 MHz

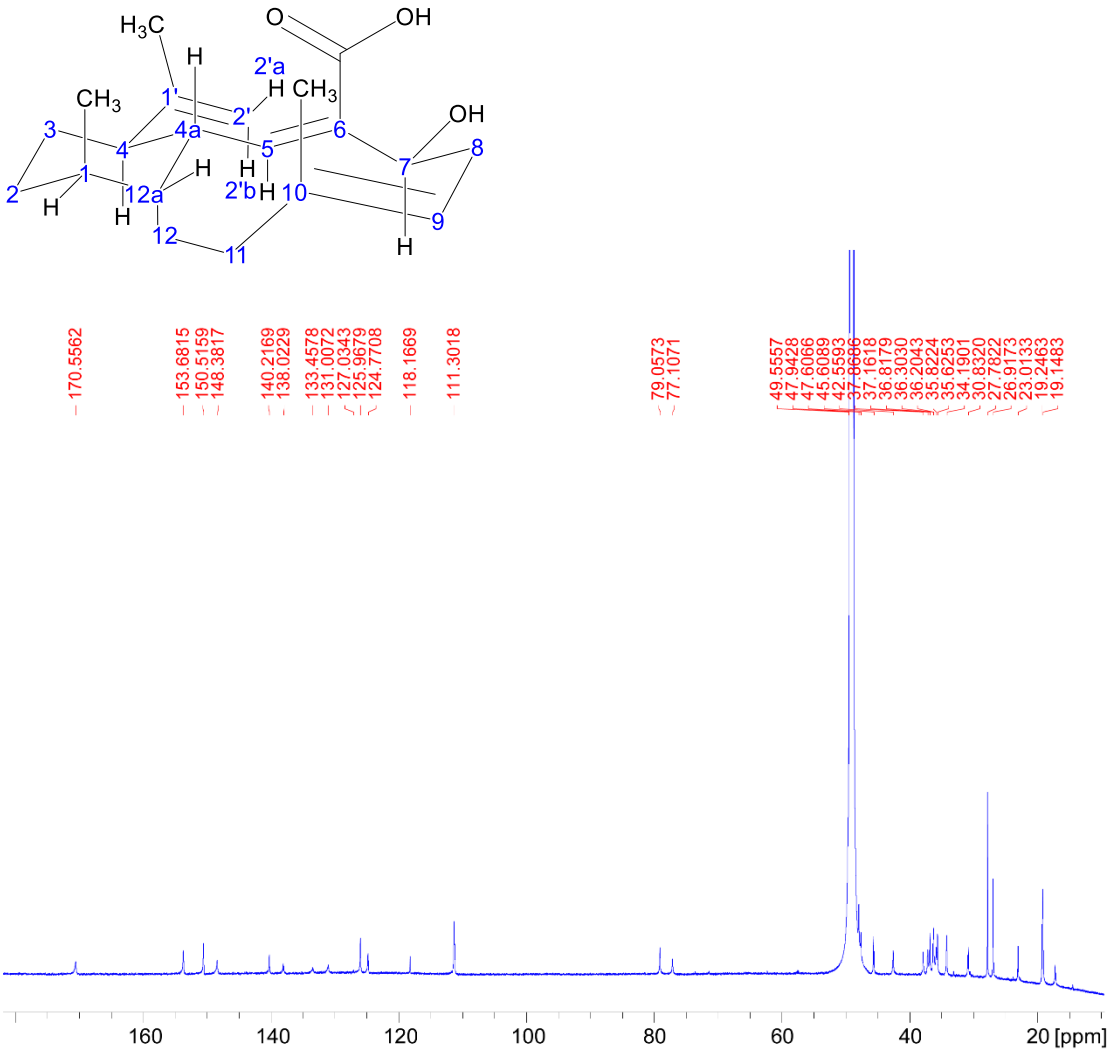


Figure C13. COSY NMR spectrum of **4.3** in methanol-d_4

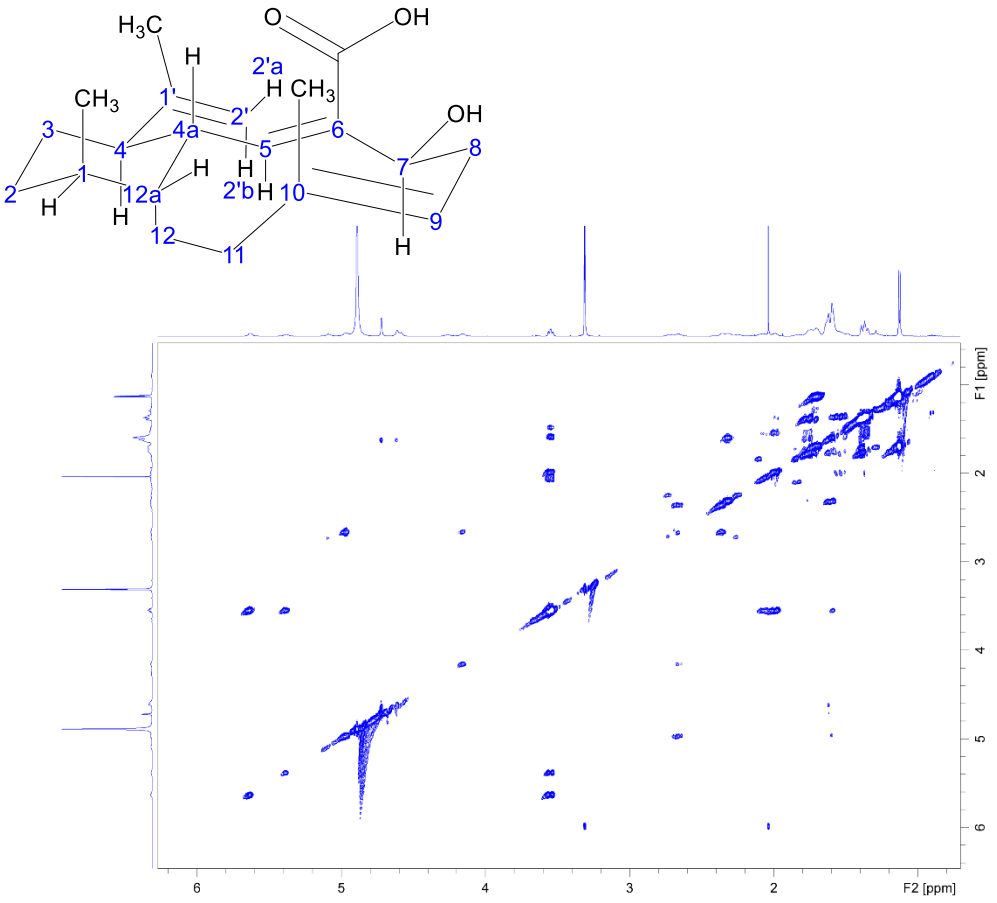


Figure C14. NOESY NMR spectrum of **4.3** in methanol-d_4. Overlapping signals (in orange circles) indicate that this compound is undergoing slow chemical exchange.

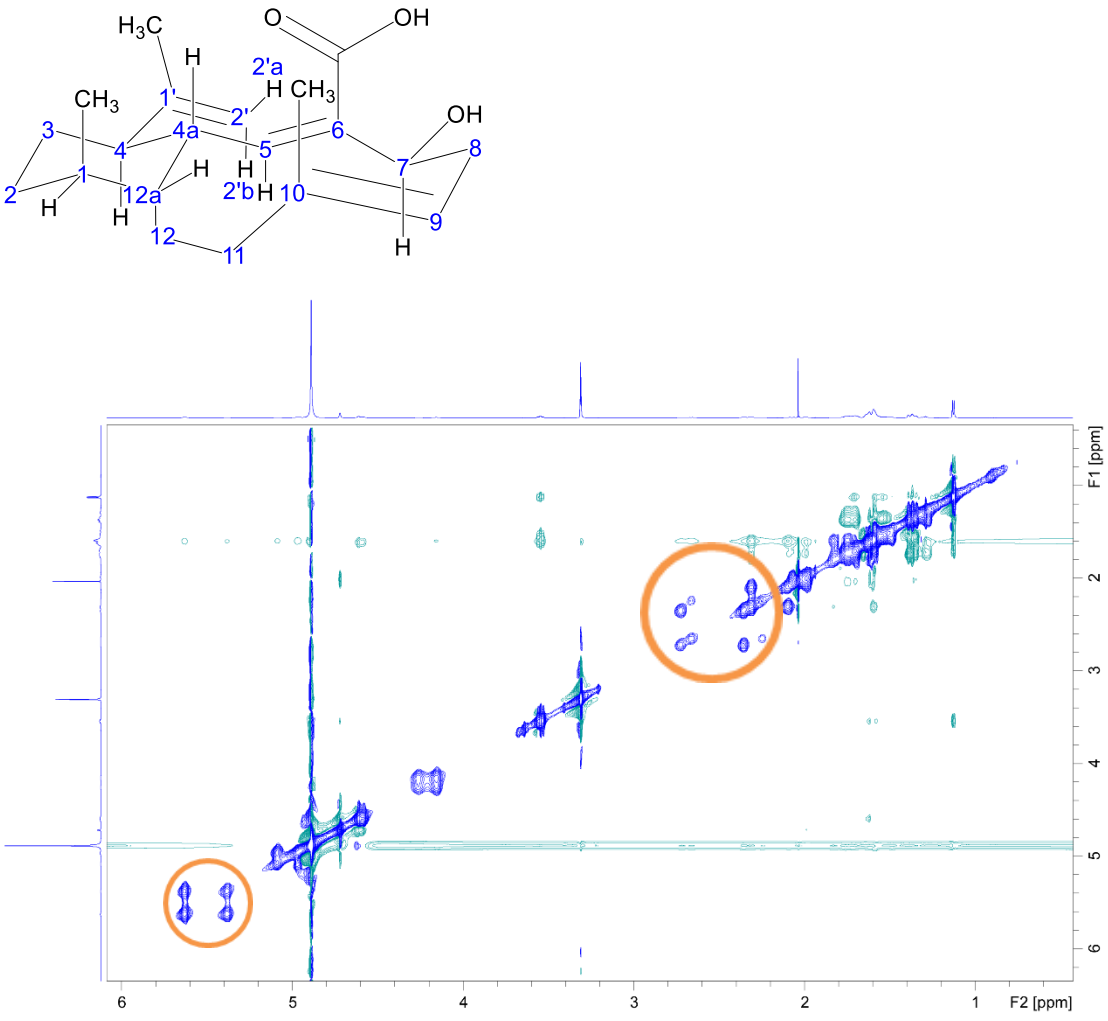


Figure C15. ROESY NMR spectrum of **4.3** in methanol-d_4. Overlapping signals (in orange circles) indicate that this compound is undergoing slow chemical exchange.

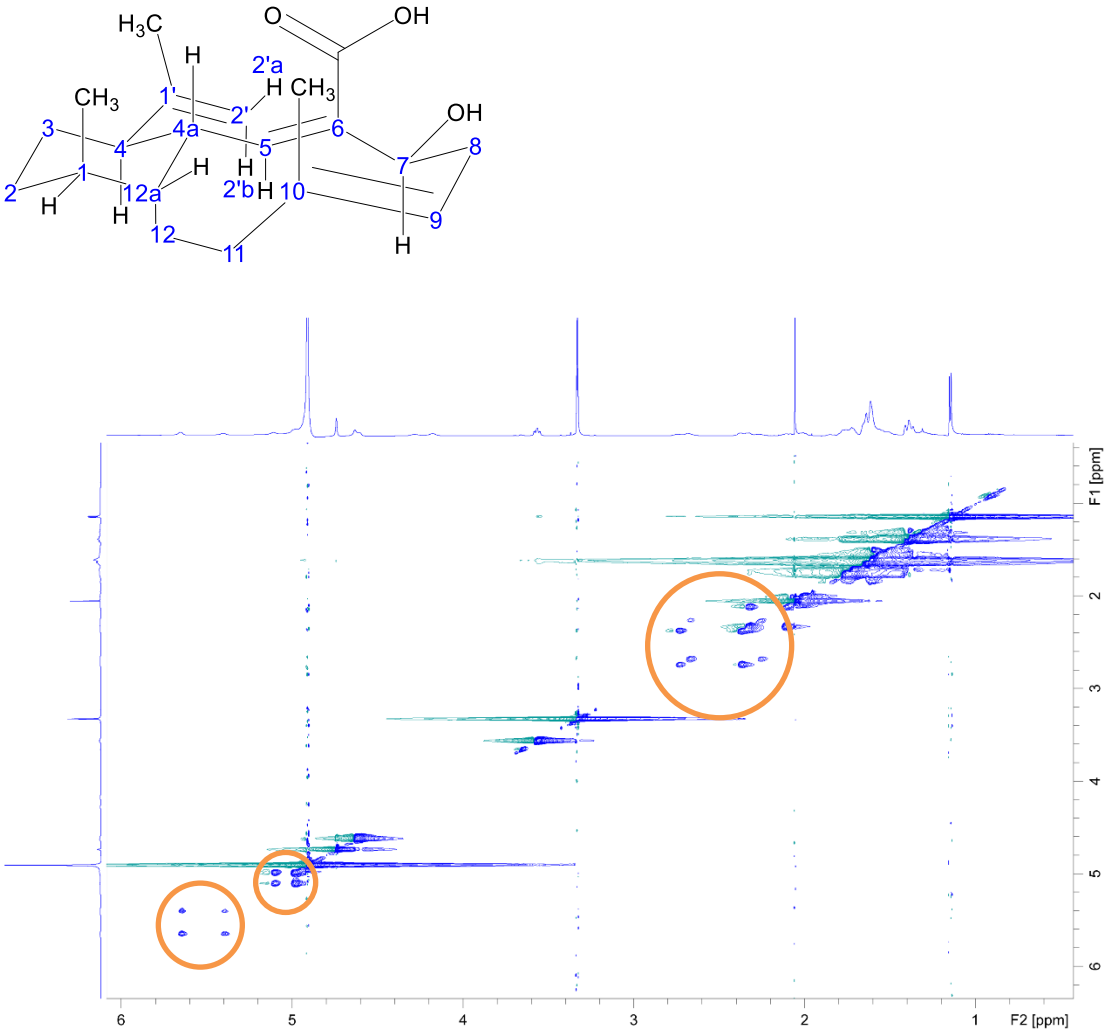


Figure C16. TCOSY NMR spectrum of **4.3** in methanol-d_4

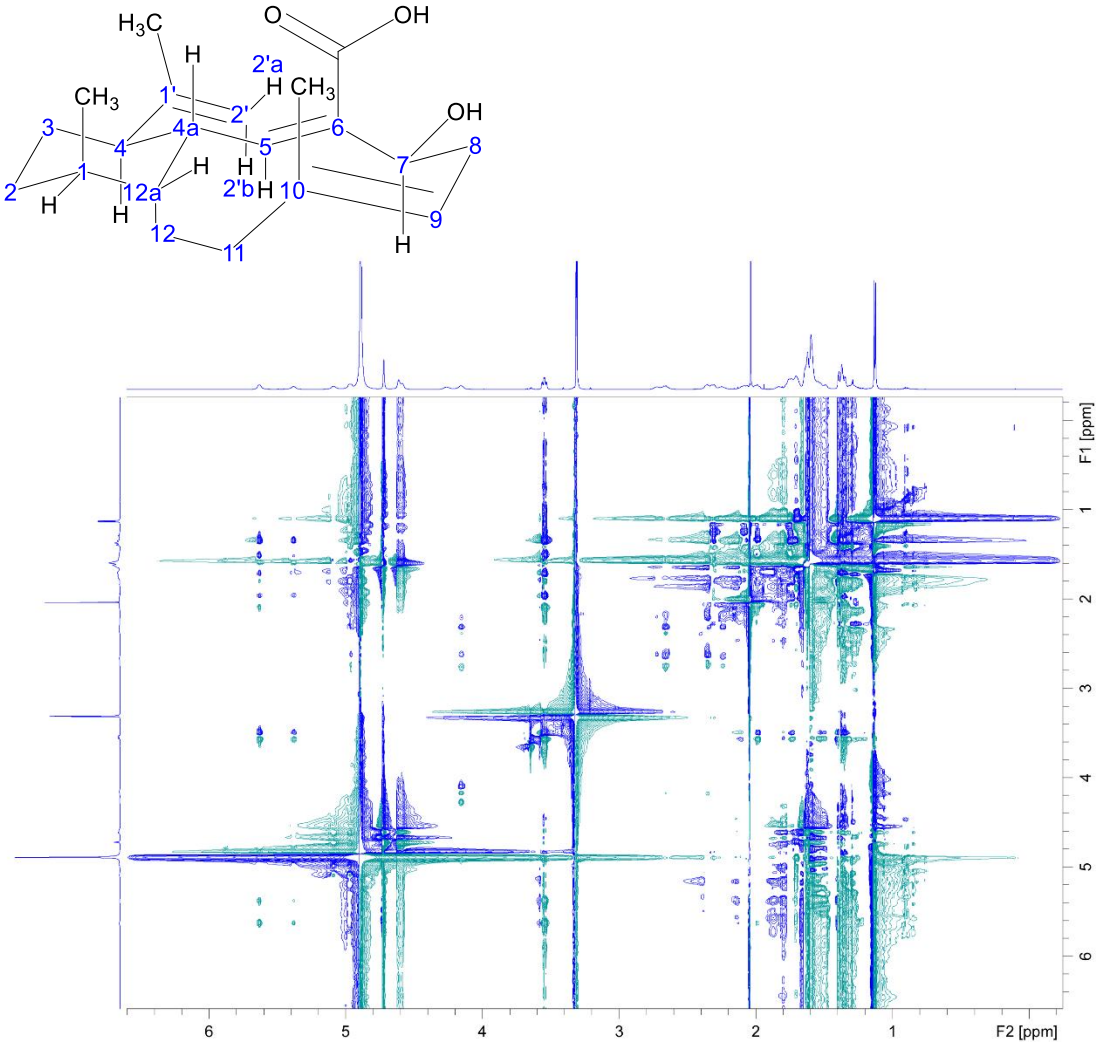


Figure C17. HSQC-DEPT NMR spectrum of **4.3** in methanol-d_4

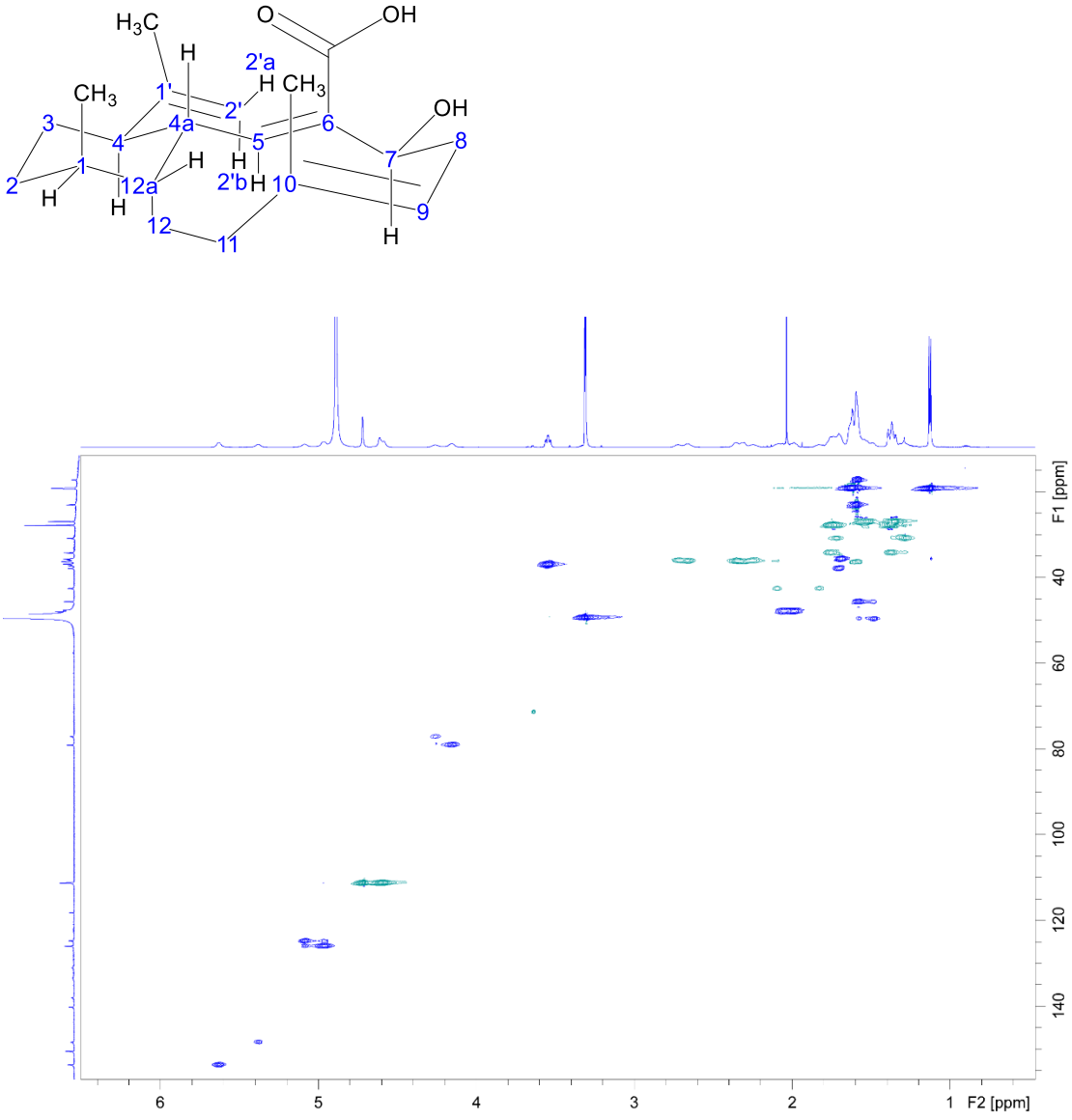


Figure C18. HMBC NMR (J_{CH} = 8 Hz) spectrum of **4.3** in methanol-d_4

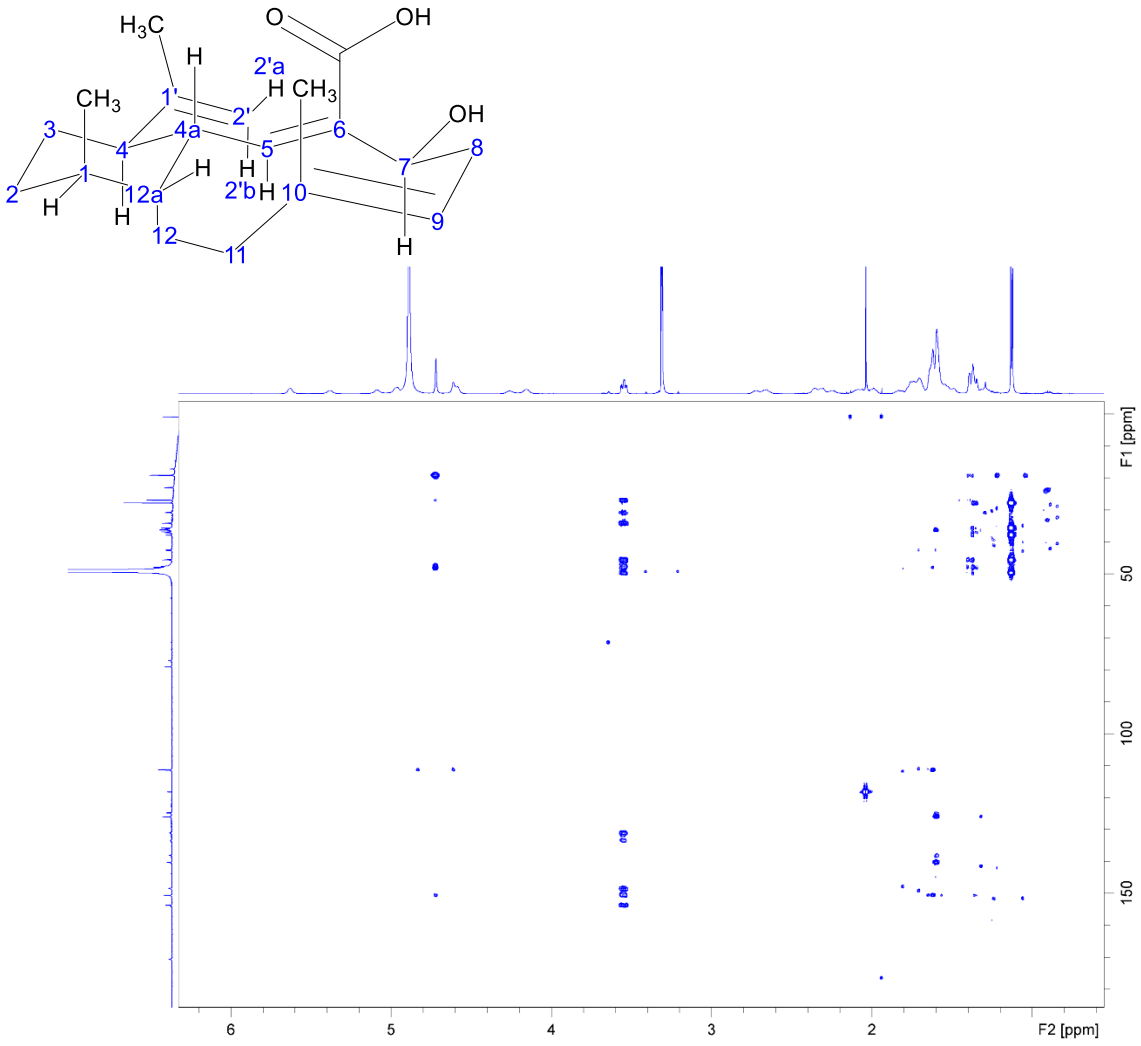


Figure C19. HMBC NMR (J_{CH} = 4 Hz) spectrum of **4.3** in methanol-d_4

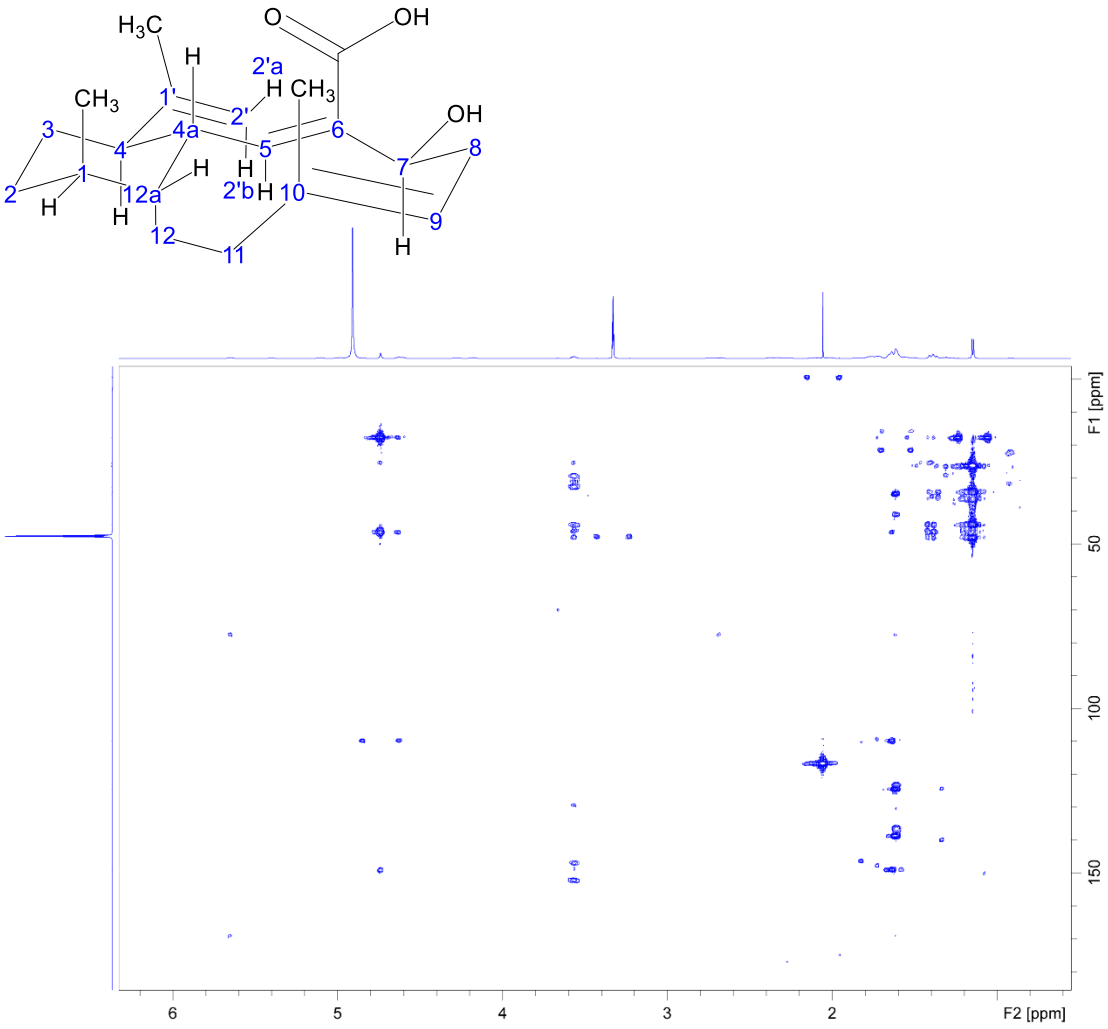


Figure C20. HMBC NMR (J_{CH} = 50 Hz) spectrum of **4.3** in methanol-d_4

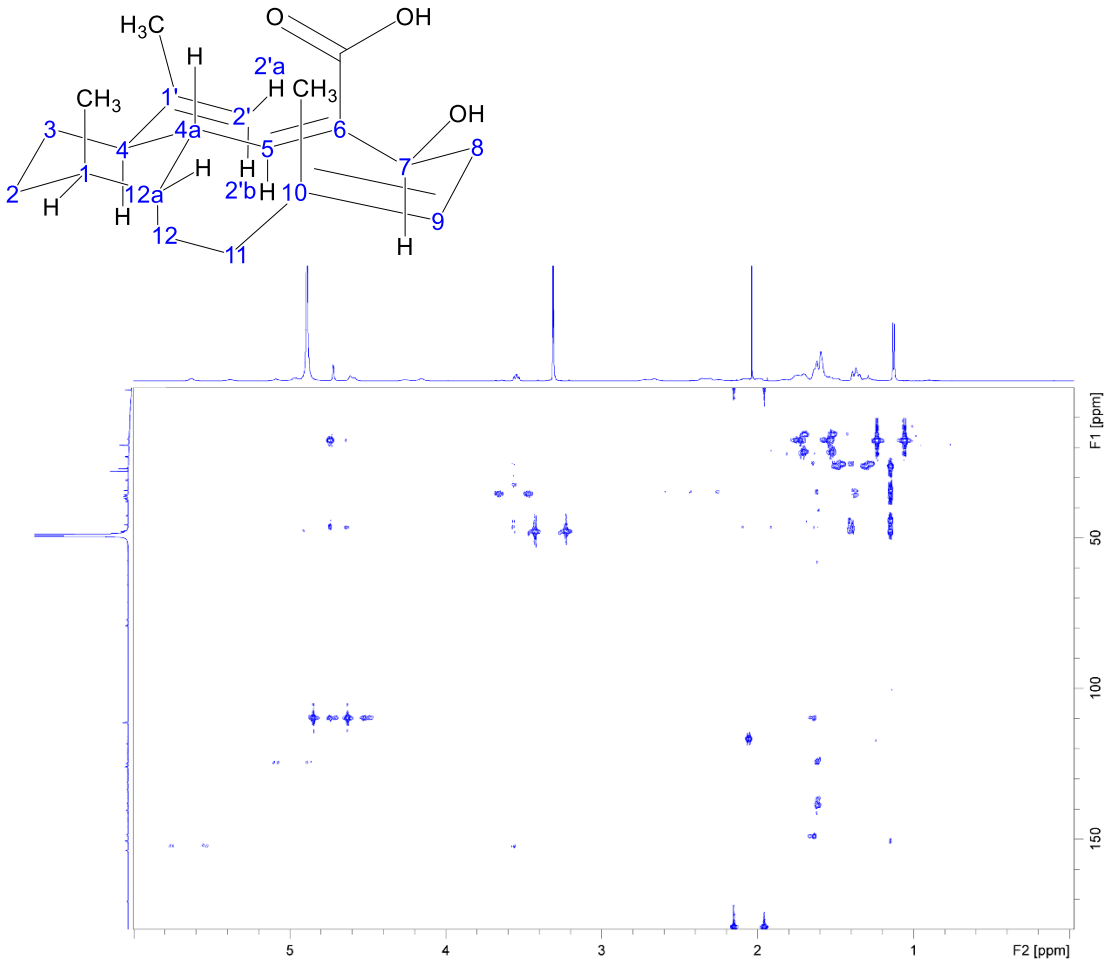


Figure C21. DQF-COSY NMR spectrum of **4.3** in methanol-d_4

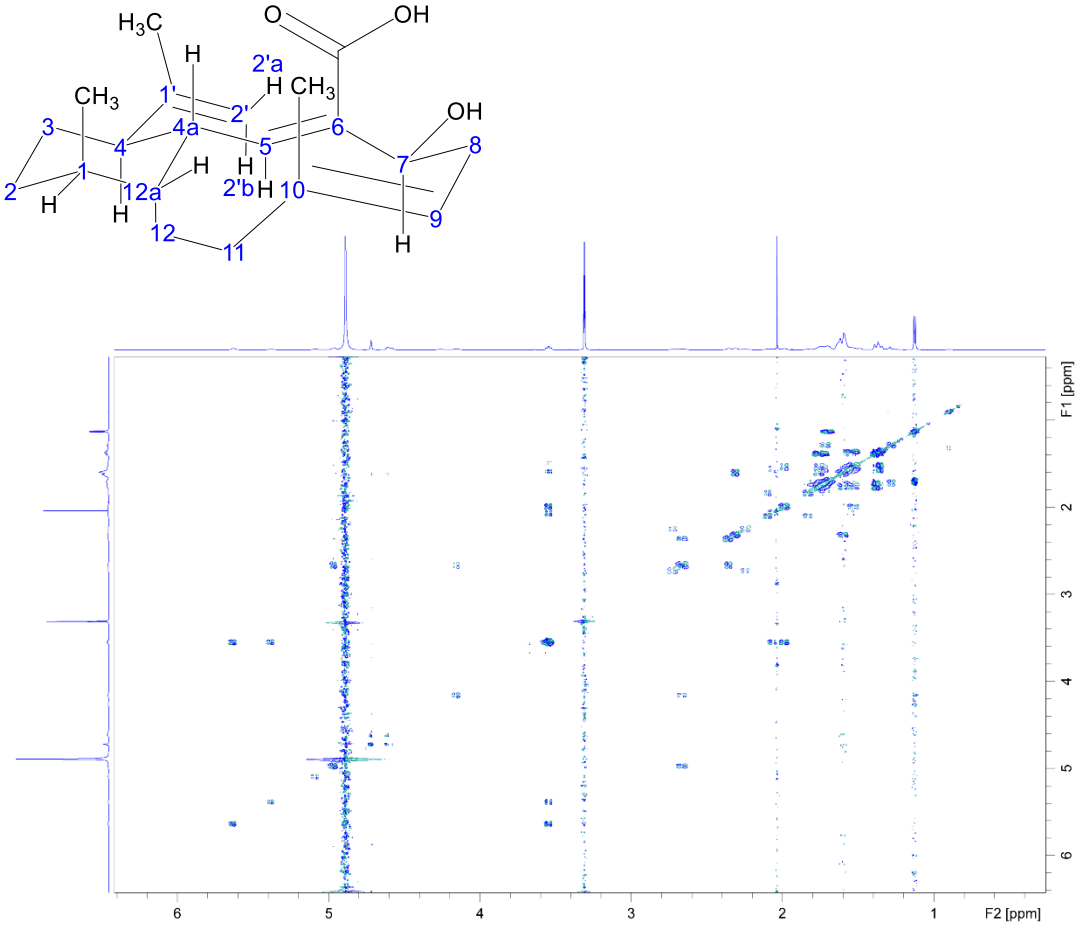


Figure C22. 1D trace of H-9 to H-1 multiplet from DQF-COSY NMR spectrum of **4.3** in methanol-d_4

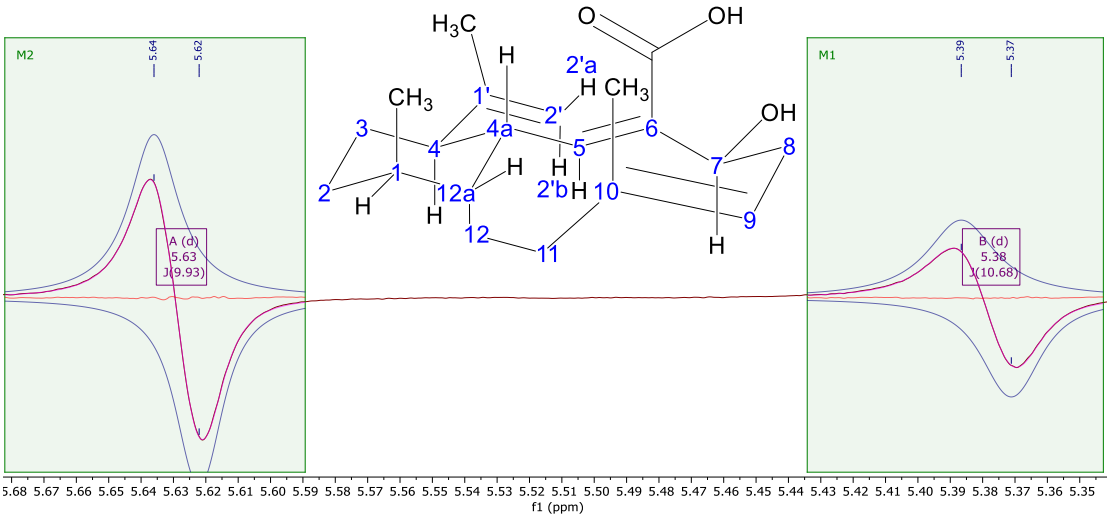


Figure C23. DQF-COSY NMR spectrum of **4.3** in methanol-d_4

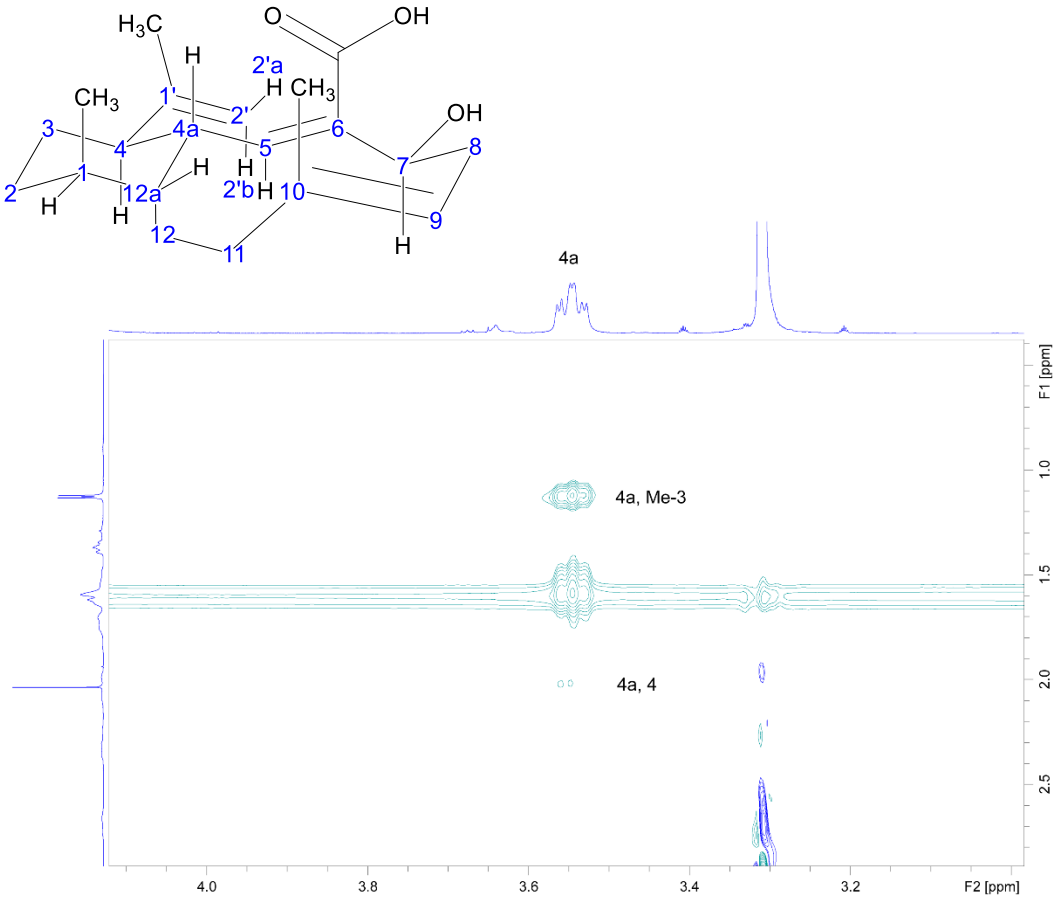


Figure C24. 1D trace of $^{3}J_{(Me10\text{-}H9)}$ coupling constant from IPAP-HMBC NMR spectrum of **4.3** in methanol-d_4

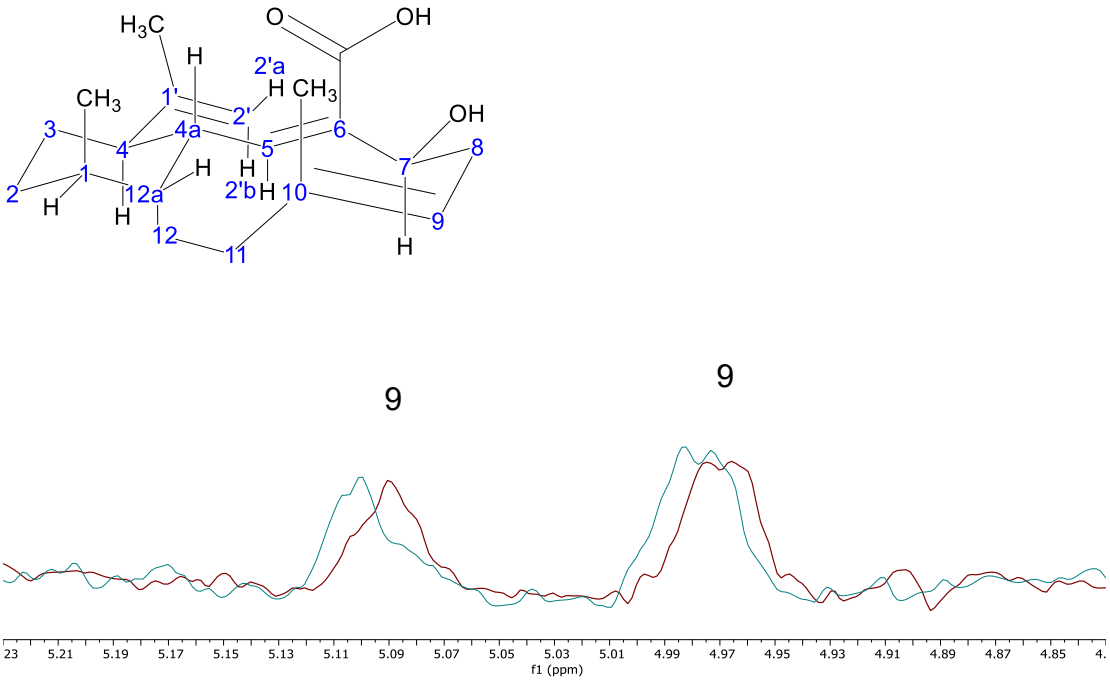


Figure C25. 1D trace of $^{4}J_{(C4\text{-}Me1)}$ ~1 Hz coupling constant from HETLOC NMR spectrum of **4.3** in methanol-d_4

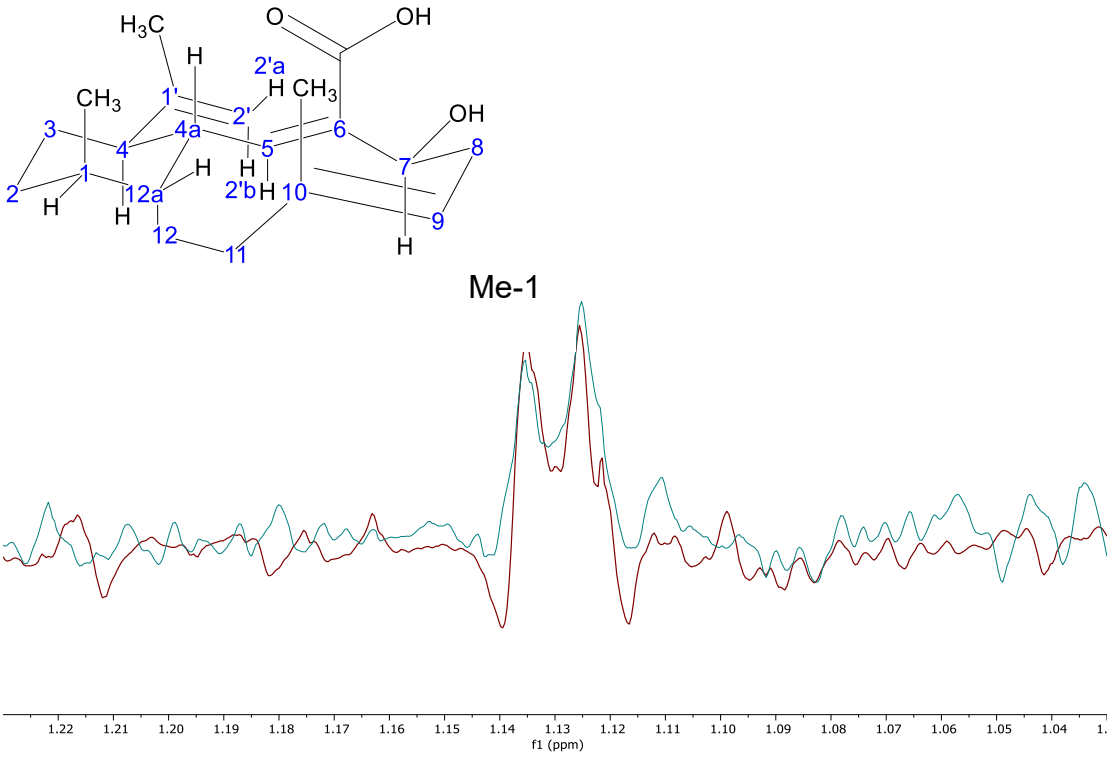


Figure C26. Key regions of ^{1}H NMR spectra of Mosher's ester analysis of **4.3** in $CDCl_3$

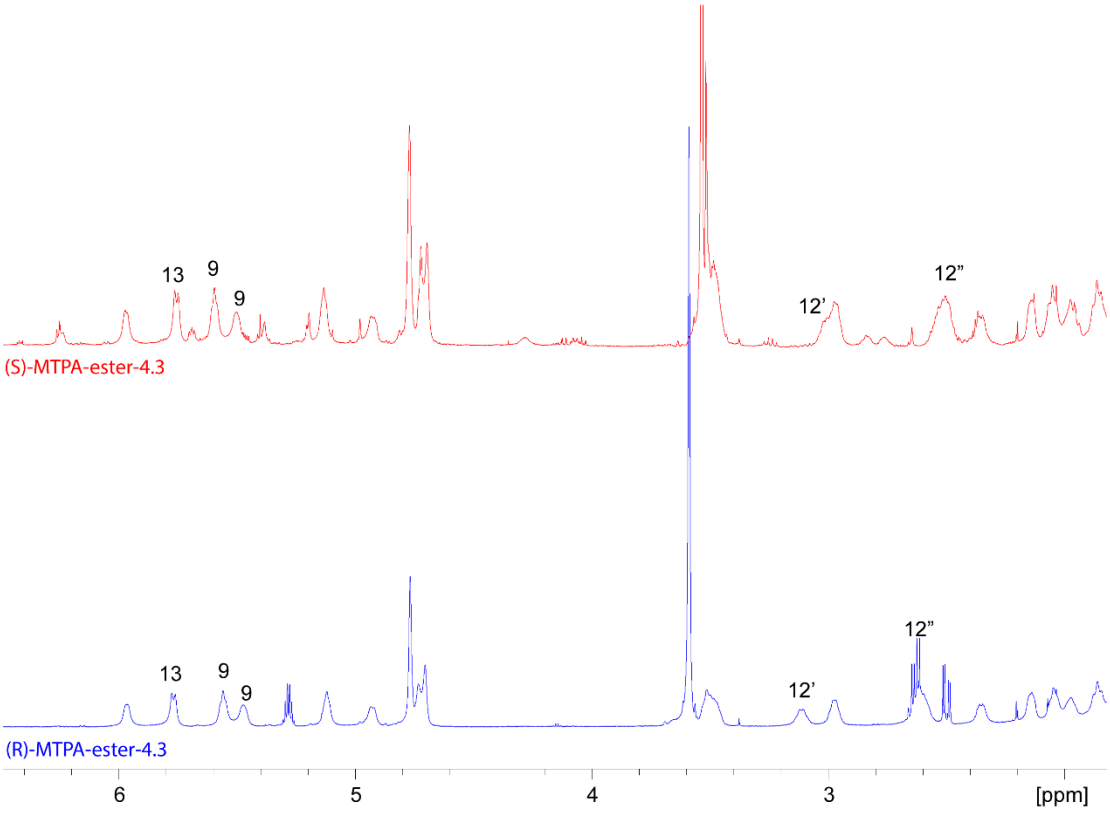
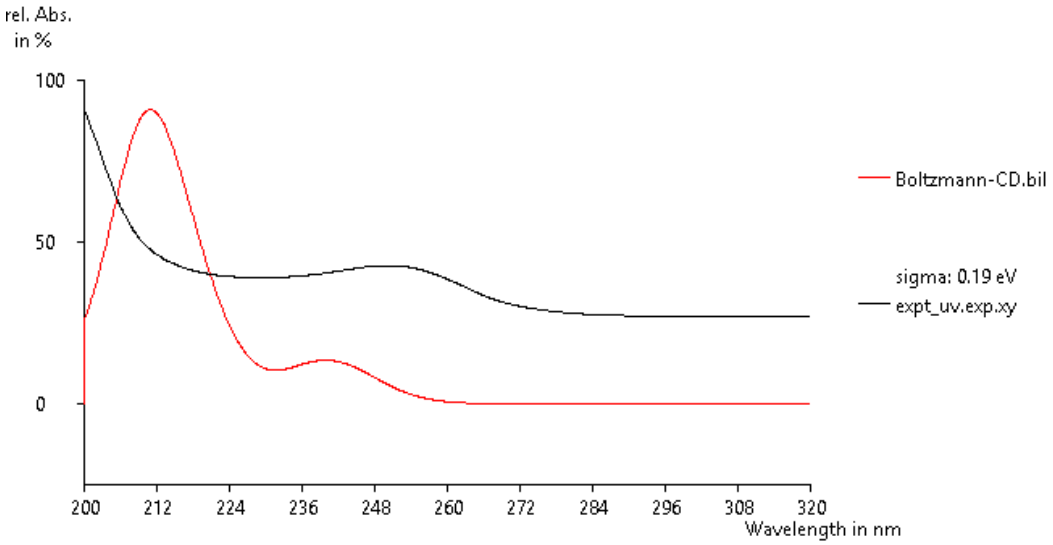


Figure C27. (A) Computational UV spectrum (in red color) vs. experimental UV spectrum (in black color); (B) computational ECD spectrum (in red color) vs. experimental ECD spectrum (in black color).

(A)



(B)