UNIVERSITY OF CALIFORNIA SAN DIEGO

Isolation and Characterization of Natural Products from Marine Cyanobacteria Aided by Enhanced NMR Methodologies

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Chemistry

by

Kelsey Leigh Alexander

Committee in charge:

Professor William Gerwick, Chair
Professor Michael Burkart, Co-Chair
Professor Elizabeth Komives
Professor Clifford Kubiak
Professor Alina Schimpf
Professor Dionicio Siegel

2021

The dissertation of Kelsey Leigh Alexander is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

This dissertation is dedicated to my parents, Leslie and Paul Alexander. I wouldn't be here without your continual support and encouragement in my life. And to Dana Humphrey, who always showed up at the right moments with a treat and the perfect words, thank you. And to Rylie Kern I am forever thankful to be able to do life with you and the fact that you always celebrate the little things. And to my sister Kendall Cavaletto, thank you for always being there and making me laugh. Thank you to all of you for being my biggest cheerleaders, your support has meant the world.

EPIGRAPH

"Just keep swimming"
-Dory, Finding Nemo

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor William Gerwick for his support as the chair of my committee.

Chapter 2, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Duggan, Brendan M.; Gerwick, William H. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Naman, C. Benjamin; Iwasaki, Arihiro; Leao, Tiago; Reher, Raphael; Kim, Hyun Woo; Ternon, Eva, Caro-Diaz, Eduardo J.E.; Glukhov, Evgenia; Duggan, Brendan M.; Gerwick, Lena; Gerwick, William H. The dissertation author was the primary investigator and author of this material.

Chapter 4, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Reher, Raphael; Naman, Benjamin C.; Caro-Diaz Eduardo J.E.; Glukhov, Evgenia; Gerwick, Lena; Gerwick, William H. The dissertation author was the primary investigator and author of this material.

The Appendix, in full, is a reprint of the material as it appears in Chemistry- A European Journal 2021. Proetto, Maria T.; Alexander, Kelsey; Melaimi, Mohand; Bertrand, Guy; Gianneschi, Nathan C. vol. 27, 2021." The dissertation author was the secondary investigator and author of this paper.

VITA

2015 B.S. in Biology/Chemistry, Point Loma Nazarene University

2017 M.S. in Chemistry, University of California San Diego

2021 Ph.D. in Chemistry, University of California San Diego

PUBLICATIONS

Alexander, K.L., Reher, R., Naman, C., Caro-Diaz, E., Glukhov,E., Gerwick, L. and Gerwick, W. Curacin A analogues from a Puerto Rican cyanobacterium. *Manuscript in preparation*.

Alexander, K.L., Naman, C., Iwasaki, A., Leao, T., Reher, R., Kim, H., Ternon, E., Caro-Diaz, E., Glukhov,E., Duggan, B., Gerwick, L., and Gerwick, W. Fatuamide A, a hybrid PKS/NRPS ionophore from the American Samoan cyanobacterium Leptolyngbya sp., *Manuscript in preparation.*

L. Tiago, Wang, M., Moss, N., Silva, R., Sanders, J., Nurk, S., Gurevich, A., Humphery, G., Reher, R., Zhu,Q., Belda-Ferre, P., Glukhov,S., Whitner,S., Alexander, K.L., Rex, R., Pevzner, P., Dorrestien, P., Knight,R., Bandeira,N., Gerwick, W.H., Gerwick,L. A Multi-Omics Characterization of the Natural Product Potential of Tropical Filamentous Marine Cyanobacteria. Mar. Drugs. 2021, 19,20.

Proetto, M.; Alexander, K.L.; Melaimi, M.; Bertrand, G.; Gianneschi, N. Cyclic (Alkyl)(Amino)Carbene (CAAC) Gold(I) complexes as Chemotherapeutic Agents. Chem. Eur. J. 2021, 27, 3772-3778.

Li, Y. Naman, C.B., Alexander, K.L., Guan, H., Gerwick, W.H. The Chemistry, Biochemistry, and Pharmacology of Marine Natural Products from Leptolyngbya, a Chemically Endowed Genus of Cyanobacteria., Marine Drugs. 2020, 18(10), 508.

Ndukwe, I., Wang, X., Lam, N., Ermanis, K., Alexander, K.L., Bertin, M., Martin, G., Muir, G., Paterson, I.,Britton, R., Goodman, J.M., Helfrich, E.J., Piel, J., Gerwick, W.H. and Williamson, T. Synergism of anisotropicand computational NMR methods reveals the likely configuration of phormidolide A. Chem. Commun. 2020,56, 7565-7568.

Reher, R., Kim, H., Zhang, C., Mao, H., Wang, M., Nothias-Scaglia, L., Caraballo, M., Glukhov, E., Teke, B., Leao,T., Alexander, K.L., Duggan, B., Everbroeck, E., Dorrestein, P., Cottrell, G., and Gerwick, W. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. JACS.2020, 142, 9, 4114-4120.

Mohimani, H.; Gurevich, A.; Alexander, K.L.; Naman, C.B.; Leao, T.; Glukhov, E.; Moss, N.A.; Knaan, T.L.;Vargas, F.; Nothias, L.; Singh, N.K.; Sanders, J.G.; Benitez, R.A.S.; Thompson, L.R.; Hamid, M.N.; Morton, J.T.; Mikheenko, A.; Shlemov, A.; Korobeynikov, A.; Friedberg, I.; Knight, R.; Venkateswaran, K.V.; Gerwick, W.; Gerwick, L.; Dorrestien, P.C.;

Pevzner, P.A. MetaMiner: A Peptidogenomics Approach for the Discovery of Ribosomally Synthesized and Post-translationally Modified Peptides from microbial communities. Cell Systems. 2019, 9,600-608.

Shellhamer, D.F.; Alexander, K. L.; Beavis, Z. J.; Bucardo, M. S.; Elwin, S. L.; Gomez, L.; Licata, C. J.; Van Horne, S. and Perry, M. C. Comparing neat vs. solution reactions of chlorosulfonyl isocyanate with alkenes. Trends in Organic Chemistry **2017,** 18, 15-20.

Shellhamer, D. F.; Alexander, K. L.**;** Bunting, S. A.; Elwin, S. L.; Licata, C. J.; Milligan, J. C.; Robinson, R. D.; Shipowick, D. E.; Smith, L. B.; Perry, M. C. and Rouffet, M. Improved Synthetic Utility of a Sluggish Electrophile: Reaction of Chlorosulfonyl Isocyanate with Unreactive and Reactive Alkenes. Synthesis **2015,** 47, 1994-1950.

ABSTRACT OF THE DISSERTATION

Isolation and Characterization of Natural Products from Marine Cyanobacteria Aided by Enhanced NMR Methodologies

by

Kelsey Leigh Alexander

Doctor of Philosophy in Chemistry

University of California San Diego, 2021

Professor William Gerwick, Chair
Professor Michael Burkart, Co-Chair

The natural world is an extraordinary source of diverse organisms and natural product compounds. The use of natural products and their derivatives has been a recurring theme in the discovery and development of new therapeutics. The marine world is vast and contains incredible biological and chemical diversity. Therefore, the oceans have the potential for new discoveries of compounds with medicinal applications. One group of marine organism that has been shown to be a prolific producer of bioactive compounds is the cyanobacteria. This work

focuses on the study of new compounds from marine cyanobacteria for their chemical diversity and biological activity.

The work in this dissertation focuses on the discovery and characterization of compounds from two different species of cyanobacteria, *Leptolyngbya sp.* and *Moorena producens*. These were collected from American Samoa and Puerto Rico, respectively. The characterization of these compounds was achieved through the integration of a variety of techniques including mass spectrometry, advanced NMR technology and genomic information. This analysis led to the characterization of a new ionophore from the *Leptolygbya sp.* The collection of *Moorena producens* afforded new analogues of the bioactive compound, curacin A. 2D NMR was an essential aspect in these studies. Therefore, the optimization of different NMR techniques for faster data acquisition was studied in this dissertation for its applications on lower concentrations of natural products.

This work highlights the novelty of compounds that can be discovered by the application of the multiple techniques. Additionally, it demonstrates the reduction in time that is possible with new techniques for the experimental acquisition of NMR data. Finally, this dissertation adds to the body of knowledge of the chemistry that originates from the world's oceans.

Chapter One: Marine Natural Product Discovery

1.1 Natural Products as medicine

Natural products contribute to a wide variety of medicines that are currently prescribed by physicians for a range of ailments. Paclitaxel (brand name Taxol®) was isolated from the bark of the Pacific yew and is used for the treatment of breast cancer, ovarian cancer, malignant brain tumors, and non-small cell lung cancers (*figure 1.1*).[1] Of all new and approved drugs from January 1981 to September 2019, 3.8% were from an unaltered natural product. Additionally this study showed that there were a variety of other classifications of drugs that were based on natural products which included biological macromolecules, botanical drugs, natural product derivatives, synthetic drugs based on a NP pharmacophore, and synthetic drugs that are mimics of a natural product. These natural product "like" categories total 85.2% of the new drugs from the last 38 years.[2] Therefore, natural products offer a promising source for novel therapeutics.

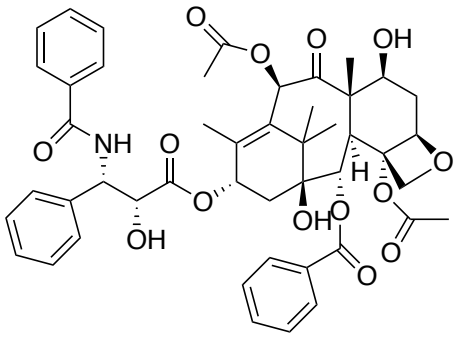


Figure 1.1: Structure of paclitaxel.

1.1.1 Approved Marine Drugs

Table 1.1: Approved marine natural product drugs.

Name	Drug	Drug Class	Treats	Source organism
Fludarabine Phosphate	Fludara®	cancer	lymphomatic leukemia	sponge
Nelarabine	Arranon®	cancer	T-cell acute lymphoblastic leukemia	sponge
Cytarabine	Cytosar-U®	cancer	leukemia and lymphomas	sponge
Vidarabine	Vira-A®	antiviral	ocular	sponge
Ziconotide	Prialt®	pain	pain	cone sail
Plitidepsin	Aplidin®	cancer	multiple melanoma	tunicate
Omega-3-Acid Ethyl Ester	Lovaza®, Omtryg®	cardiovascular	hyperlipidemia	fish
Icosapent ethyl	Vascepa®	cardiovascular	hyperlipidemia	fish
omega-3-carboxylic acids	Epanova®	cardiovascular	hyperlipidemia	fish
Propylene glycol alginate Sulfate Sodium	PSS	cardiovascular	hyperlipidemia	algae
Iota-carrageenan	Carragelose®	nasal spray	viral cold	seaweed
Trabectedin	Yondelis®	cancer	sarcomas	tunicate
Brentuximab vedotin	Adcetris®	cancer	lymphoma	cyanobacteria
Eribulin	Halaven®	cancer	breast cancer	tunicate
Polatuzumab vedotin	Polivy®	cancer	B-cell lymphoma	cyanobacteria
Protamine sulfate	Protamine, NovoMix 30	Antidote, Diabetes	heparin overdose, diabetes	fish
Hemocyanin (KLH)	KLH	vaccine	immune stimulation	keyhole limpet
Belantamab mafodotin-blmf	BLENREP®	cancer	multiple mylenoma	cyanobacteria
Lurbinectedin	Zepzelca®	cancer	metastic small cell lung cancer	sea squirt
Enfortumab vedotin	Padcev®	cancer	bladder cancer	cyanobacteria

There are 21 approved marine natural product drugs to date that are prescribed for a variety of treatments including pain, cancer, antiviral, and cardiovascular diseases (*table 1.1*). These therapeutics have sources from a variety of different marine organisms, including algae, cone snails, cyanobacteria, fish, keyhole limpets, seaweed, tunicates, and sponges. These treatments cover a variety of different drug classes as well, including antiviral, cancer, vaccine, pain, cardiovascular disease, antidote, and nasal sprays. Additionally, there are many other marine natural products that are currently undergoing clinical trials.[3]

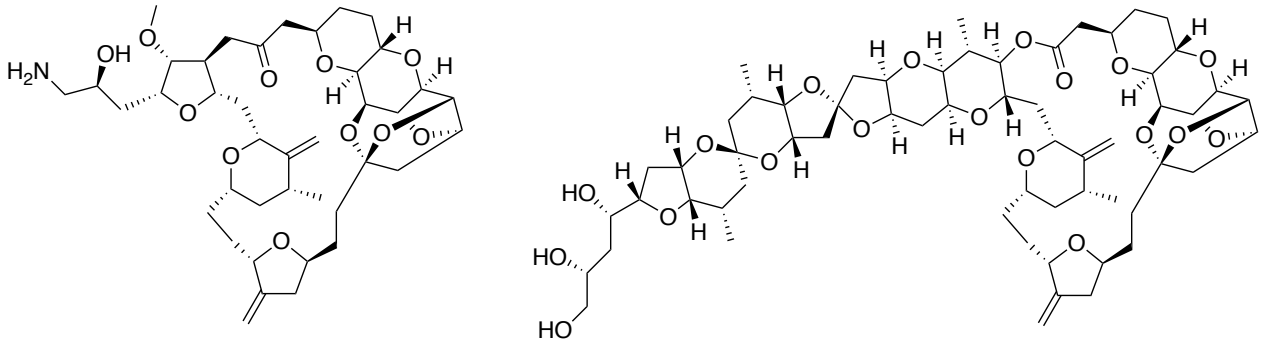


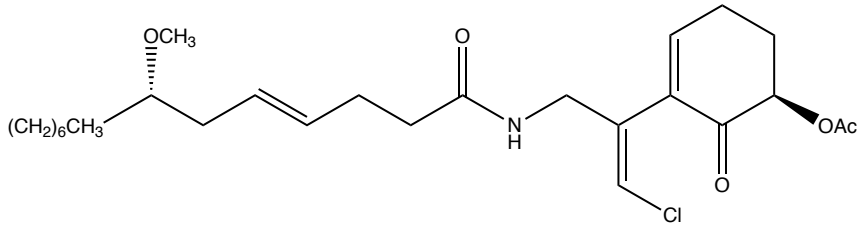
Figure 1.2: a. Structure of eribulin b. structure of halichondrin B.

A clinically available marine drug that is prescribed as an anticancer agent is Eribulin (*figure 1.2a*). Eribulin is a natural product derivative of halichondrin B (*figure 1.2b*). Halichondrin B was isolated from the marine sponge *Halichondria Okadai Katodata*.[4] The total synthesis of halichondrin B was initially reported in 1992 by the Kishi group and required 90 chemical steps that utilized key reactions to assemble smaller building blocks into the complete halichondrin B molecule.[5] The synthetic scheme of halichondrin B was further developed and through testing of its derivatives, and it was found that the macrocyclic lactone portion of the molecule gave rise to its cytotoxicity.[6] Further studies on the structure and optimization of the active portion of the molecule was done at the Eisai Research Institute and led to the drug known

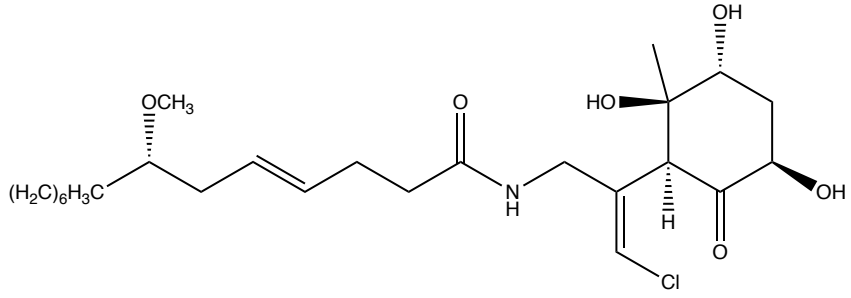
as Eribulin.[7] Eribulin is prescribed for metastatic breast cancer and it has been shown to increase the overall survival of the patients compared to treatment of physician's choice.[8]

1.1.2 Natural products as anti-inflammatory agents, neurotoxins, proteasome inhibitors and cytotoxins

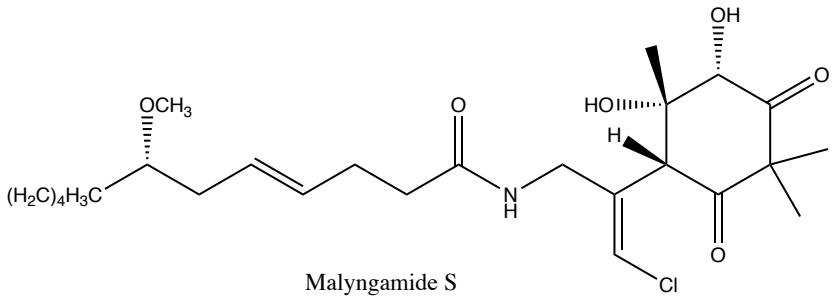
Moore *et al* reported finding bioactive compounds from cyanobacteria since 1981.[9] It has been hypothesized that cyanobacteria produce these bioactive secondary metabolites to fend off predators in their natural environment.[10] It has also been suggested that the method of extraction and isolation of these metabolites has led to the discovery of bioactive natural products, that include anti-inflammatory agents, neurotoxins and proteasome inhibitors.[11]



Malyngamide F acetate



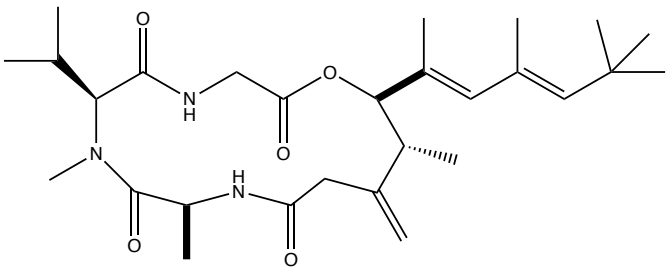
Malyngamide 2



Malyngamide S

Figure 1.3: Structures of anti-inflammatory agents malygamide F acetate, malyngamide 2, and malyngamide S.

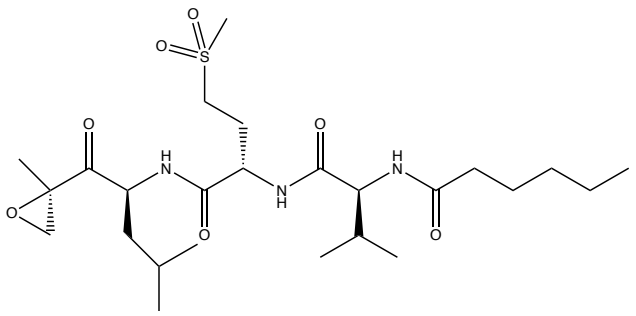
Some anti-inflammatory agents that have been discovered from cyanobacteria are the malyngamide F series. Malyngamide F acetate (*figure 1.3*) was found to have IC_{50} of 7.1 μM in the NO assay for anti-inflammatory activity. Malyngamide F acetate was found to selectively inhibit the MyD88-dependent pathway.[12] Additionally, malyngamide 2 (*figure 1.3*) was also found to have anti-inflammatory activity to LPS-induced RAW macrophage cells with an IC_{50} of 8.0 uM.[13] Malyngamide S (*figure 1.3*) was found to inhibit superoxide production in an activated human peripheral blood neutrophil assay.[14]



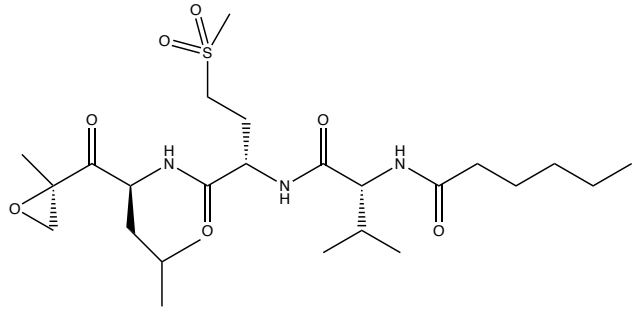
Antillatoxin

Figure 1.4: Structure of the neurotoxic agent antillatoxin.

Antillatoxin, is a lipopeptide that was found to have ichthyotoxic effects (*figure 1.4).*[15] Further studies showed antillatoxin was neurotoxic to rat cerebellar granule neurons and had an acute neurotoxicity with an LC_{50} of 20.1 ± 6.4 nM.[16] Antillatoxin has a neurotoxicity similar to brevetoxin (a neurotoxin from a dinoflagellate) in that it is a sodium channel activator.[17,18] Antillatoxin was found to have activity in cells that heterologously express rNAv1.2, rNav1.4, and rNav1.5 alpha subunits and had a unique efficacy in comparison to veratridine and brevetoxin-2.[11,18]



Carmaphycin B



Analog 18

Figure 1.5: Structures of proteasome inhibitors caramphycin B and analog 18.

Proteasomes are important in the regulation of many different cellular processes. Therefore, inhibition of the proteasome in pathogens is a relevant target for therapeutics.[19] carmaphycin B (*figure 1.5a*), which features an α,β-epoxyketone warhead, inhibits the β5 subunit of the 20S proteasome in yeast at the nanomolar range.[20] Further modifications on the carmaphycin B scaffold were made to develop more potent analogues that are proteasome inhibitors with antimalarial activity. The lead compound, analog 18 (*figure 1.5b*), from these modifications has a 100-fold wider therapeutic window than carmaphycin B.[21]

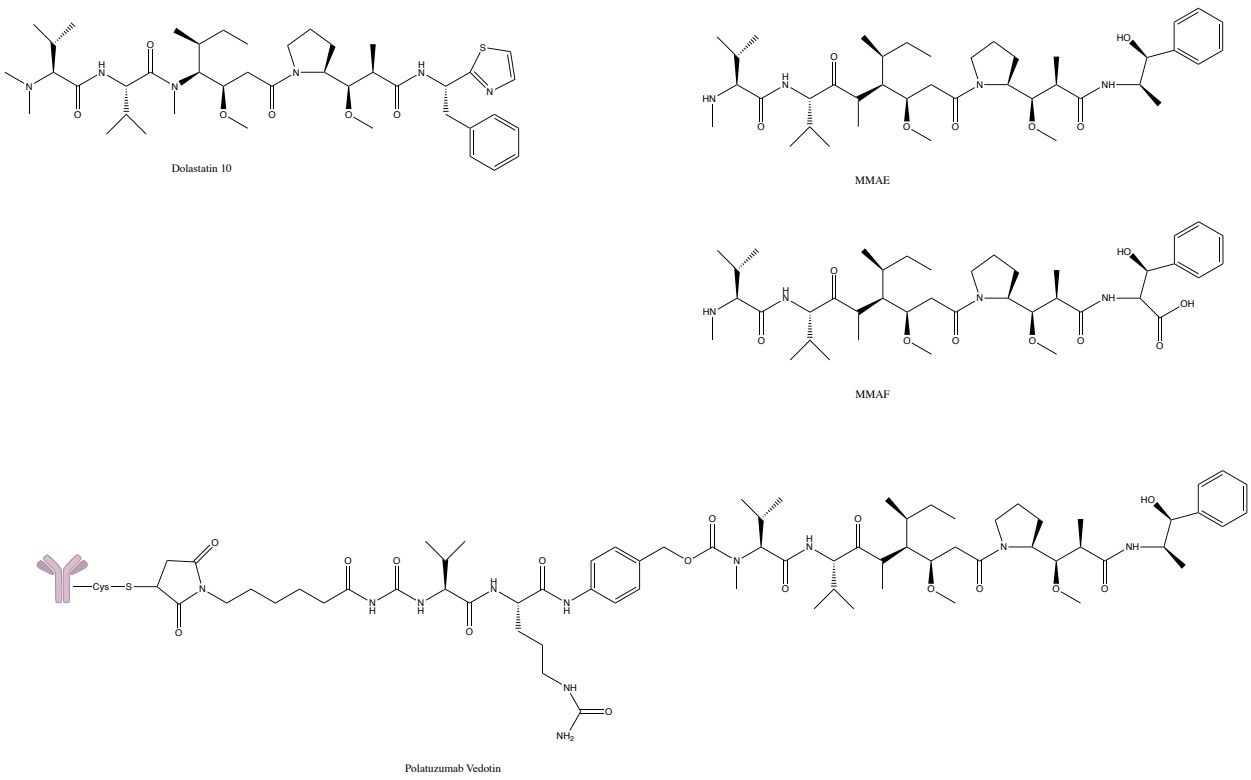


Figure 1.6: Structures of dolastatin 10, MMAE, MMAF, and polatuzumab vedotin.

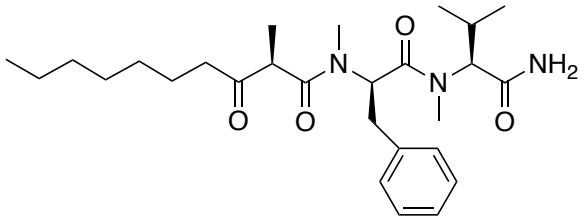
Dolastatin 10 (*figure 1.6*) was initially isolated from the sea hare *Dolabella auriculara*; however, it was later discovered to originate from a marine cyanobacterium.[22,23] Dolastatin 10 showed potency in the 0.1 nM range in the NCI cancer cell cytotoxicity primary screen with the brain, ovary, and colon tumors being the most sensitive.[23] Dolastatin 10 and derivatives, auristatin PE and tasidotin, were studied in multiple phase I and phase II trials; however none of these made it past phase II trials.[24] In a phase II clinical trials for advanced breast cancer, dolastatin 10 showed minimal activity.[25] Further SAR studies were done with dolastatin 10[26], which included the analogue auristatin E. A new analogue monomethyl auristatin E (MMAE) was published by Seattle Genetics Scientist (*figure 1.6*).[27] This new analogue was linked to a monoclonal antibody mAb and was then taken to clinical trials and led to the development of brentuziman vedotin which is an FDA approved drug.[25] The MMAE warhead is also a part of the FDA approved drug polatuzumab vedotin (*figure 1.6*) which is linked to the IgC1-kappa anti-

CD22 monoclonal antibody which is used to target B-Cell Lymphoma.[25,28] Additionally, this warhead, MMAE, is currently under multiple clinical trials with a variety of different antibody conjugates, allowing it to target different cancers.[25] Another warhead (*figure 1.6*), based on dolastatin 10, is MMAF (monomethylauristatin F-linked). This compound has also been linked to an anti-CD38 monoclonal antibody to give rise to the drug Belantamab mafodotin-blmf, which recently obtained approval by the FDA for the treatment of multiple melanoma.[29,30,31] MMAF has been incorporated into multiple ADCs that are in clinical trial. One such MMAF-based drug is depatuxizumab mafodotin that is in Phase III clinical trials for glioblastoma multiforme.[32]

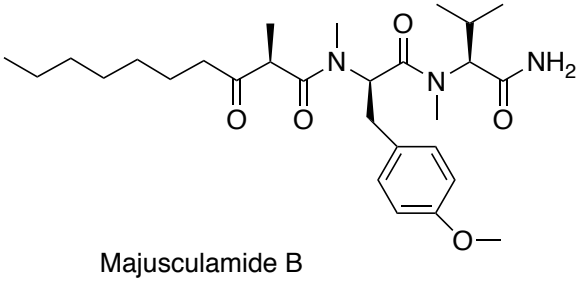
1.2. Secondary Metabolites from marine cyanobacteria

1.2.1 Cyanobacteria

Cyanobacteria or 'blue-green algae' have the capability to produce bioactive compounds. Cyanobacteria are gram-negative prokaryotes that are able to carry out oxygenic photosynthesis.[33] Benthic tropical cyanobacteria can grow in large quantities in the environment allowing for collection by skin and SCUBA diving.[34] The discovery of majusculamides A and B (*figure 1.7*) by Richard E. Moore began the pursuit of structurally diverse compounds produced by marine cyanobacteria.[35] Cyanobacteria are noted for their rich color and contain many pigments including chlorophyll a, phycobiliproteins and carotenoids.[36] These pigments have different biotechnological applications as well. Phycoerythrin and phycocyanin have been used for antibody labeling and immunofluorescence techniques.[36,37,38]



Majusculamide A



Majusculamide B

Figure 1.7: Structures of cyanobacteria metabolites majusculamide A and B.

Filamentous cyanobacteria can be kept in culture in the laboratory with specific growth media.[39] Changes in nitrogen levels, temperature, pH, and symbionts can impact the production of the expressed secondary metabolites.[40] Production of some cyanobacteria metabolites can be enhanced by modifying the constituents of the media. For example, UIC 10036 exhibited a higher production of tolytoxin with high phosphate conditions.[41]

1.2.1.1 Leptolyngbya sp.

Leptolyngbya is a genus of non-heterocystous cyanobacteria that is a part of the order *Pseudanabaenales*. [40] *Leptolyngbya* are thin filamentous cyanobacteria that range from 0.5 to 3.5 μm in width.[42] A diverse range of metabolites have been isolated from *Leptolyngbya sp.* that have been collected all over the world from both marine and fresh water environments; these metabolites include polypeptides, simple esters, macrolides, pyrones, polyaromatics, oxazolines, toxins, phenolic compounds, and pigments.[40] One bioactive secondary metabolite that was isolated from a Panamanian *Leptolyngbya sp.* is coibamide A (*figure 1.8*), reported to have antiproliferative activity and had nM cytotoxicity against multiple cell lines.[43]

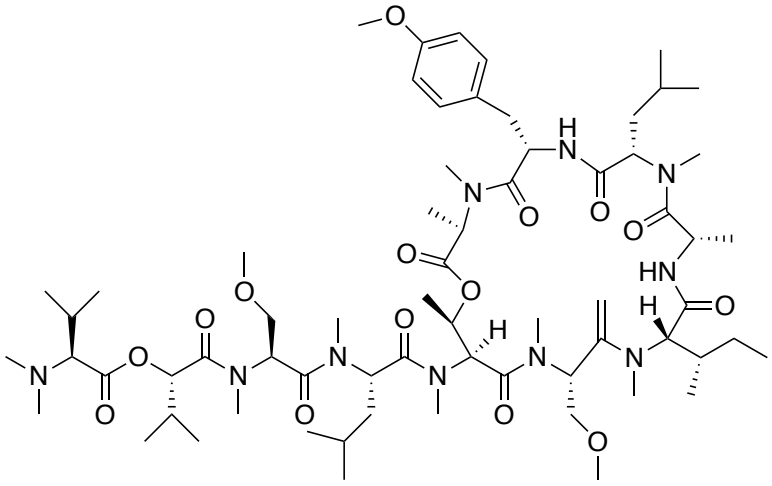
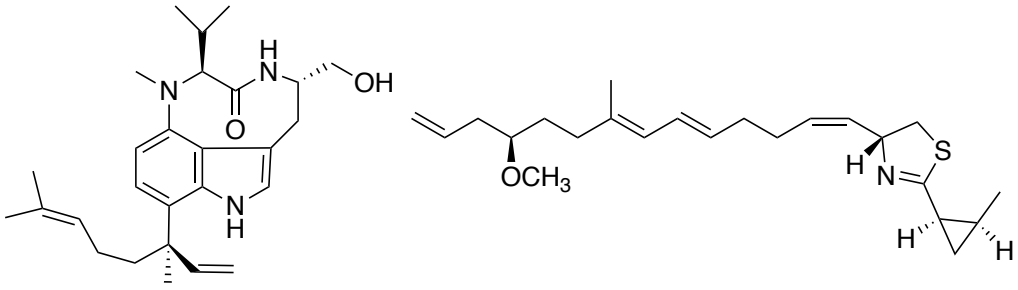


Figure 1.8: Structure of antiproliferative agent coibamide A.

1.2.1.2 Moorena producens

More than 40% of the reported marine cyanobacteria natural products have come from the genus Moorena *(previously Moorea)*.[44,45] Some strains of *Moorena producens* (formerly identified as *Lyngbya majuscula)* were noted for their ability to cause skin irritation, and this characteristic has been attributed to the toxins produced by these organisms which includes lyngbyatoxin A (*figure 1.9)* .[46,47] Since the discovery of lyngbyatoxin A, more bioactive compounds have been discovered including the curacins. Curacin A (*figure 7*), originally isolated from *Lyngbya majuscula* from the island of Curaçao, was shown to be a novel antimitotic agent.[48]



Lyngybatoxin A Curacin A

Figure 1.9: Structures of lyngbyatoxin A and curacin A.

1.2.2 Siderophores and Ionophores

Metals play an important role in the human body and in the environment. Iron (III) is insoluble in aerobic environments at a physiological pH. Iron is essential for the growth of bacteria; however, it is only found at concentrations of 0.01- 2 nM in surface waters of the ocean.[49] Bacteria use siderophores to acquire this limited element. Siderophores are low molecular weight organic compounds that scavenge Iron from the environment and have a higher affinity for iron (III) than iron (II).[50] Siderophores typically bind to Iron (III) through bidentate ligands that include such functional groups as catechols, hydroxamates, phenolates and carboxylates.[51] Enterobactin (*figure 1.10)* is an example of a tris-catechol cyclic siderophore, which was isolated from multiple enteric and pathogenic bacteria.[49] Marine siderophores tend to be amphiphiles with fatty acids that have iron(III) binding head groups; one example of this is marinobactin (*figure 1.10*).[49] Citric acid, schizokinen, and synechobactin A-C (*figure 1.10)* are examples of siderophores that have been isolated from cyanobacteria.[52]

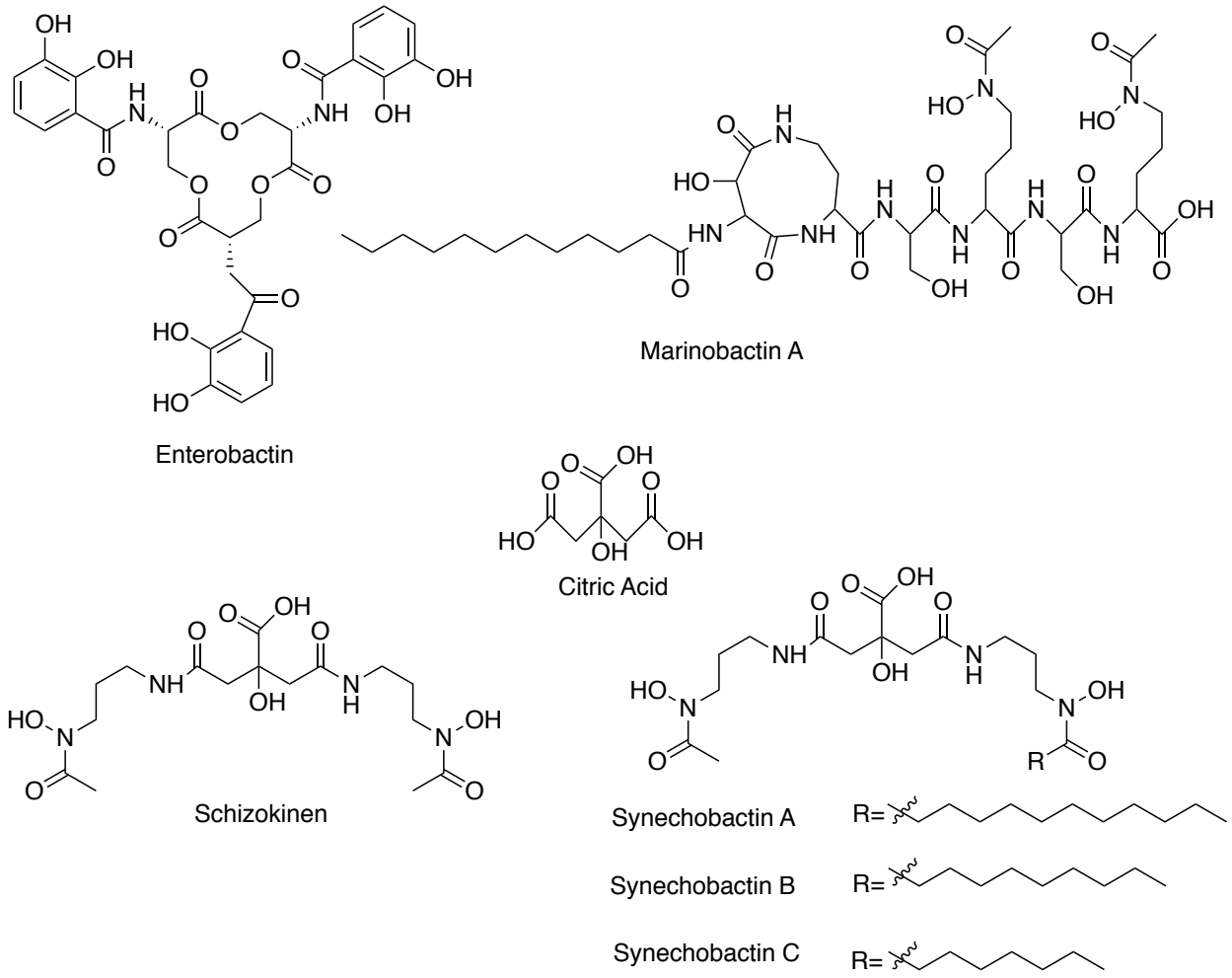
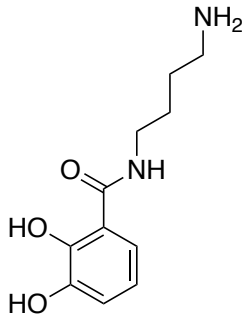


Figure 1.10: Structures of the siderophores enterobactin, marinobactin A, citric acid, schizokinen, synechobactin A-C.

Ionophores are compounds that sequester other metals from the environment and are named after the metal they acquire; some examples include chalkophores (Cr), molybdophores (Mo), and zincophores (Zn).[53] Methanobactins (Mbns) are a class of chalkophore that derive from methane oxidizing bacteria.[54] Mbns have potential as drug therapeutics for Wilson's disease as well as treatment of hemochromatosis with desferrioxamine.[54] Nitrogen fixing organisms utilize Mo in their nitrogenase enzyme, and a molybdophore such as aminochelin (*figure 1.11*) are secreted to acquire Mo.[55] Fungi have the ability to transport Zn in their metabolism and utilize zincophores to acquire zinc.[56] A siderophore inspired chelator was

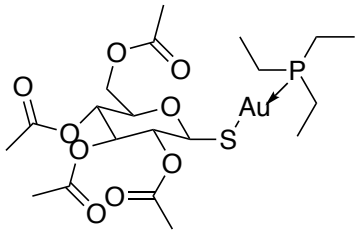
developed to bind to Uranium; this took advantage of the fact that the uranyl ion reacts similarly to Fe(III) to hydroxamate ligands.[57]



Aminochelin

Figure 1.11: The ionophore aminochelin.

Metal complexes have been used throughout history as therapeutics for a variety of ailments. The drug Auranofin (*figure 1.12*) contains gold and is used to treat Rheumatoid Arthritis.[58] Cisplatin is a platinum drug prescribed for the treatment of multiple cancers including testicular, ovarian, lung, and bladder.[59] Metal complexes also have the potential for use as antimicrobial drugs. Analysis of the Community for Open Antimicrobial Drug Discovery (CO-ADD) revealed that complexes containing a metal atom had a higher hit rate (9.9%) as opposed to simple organic compounds (0.87%).[60] A preclinical study utilized siderophores as a conjugate to carry antibiotics through the bacterial membrane and this had activity against Gram-negative pathogens.[61]



Auranofin

Figure 1.12: Structure of Auranofin.

1.3 Isolation Strategies and Structural Analysis

1.3.1 Bioassay Guided

One traditional natural products discovery approach for natural products has been 'bioassay guided'. This involves a process by which extracts and fractions are screened against different bioassays in pursuit of an active compound. Different extraction and fractionation procedures can be utilized depending on the organism and compound of interest. In the extraction of secondary metabolites of cyanobacteria different techniques such as solid phase extraction, microwave treatment, sonication, liquid extraction, pulverization, and lyophilization, are used based on the type of compound that is being isolated.[62] For example, hoiamide D was isolated through a bioassay guided fractionation approach and was found to have activity against p53/MDM2 interaction.[63]

1.3.2 Genetic Information Guided

In recent years, genetic sequencing has become more efficient and cheaper, allowing for the sequencing of organisms to be accessible and has allowed for genetic approaches to be implemented in natural products discovery. These approaches have been referred to as a "bottom-up" isolation approach, as opposed to the traditional "top down" bioassay guided isolation approach. This "bottom up" isolation utilizes bioinformatic tools as well as native and heterologous hosts.[64] Secondary metabolites are produced by the enzymes encoded by biosynthetic gene clusters (BGCs), and these can be found through the process of genome mining.[65] AntiSMASH is a bioinformatics tool that detects BGCs from the queried genome.[66] While genome mining is a powerful tool to investigate biosynthetic gene clusters for novel natural products, many of these gene clusters are silent in that they don't produce the natural product of interest under normal conditions. Therefore, manipulation of the native host is required for expression of the metabolites from these silent genes.[65] Genetic manipulation such

as gene deletions and promoter exchanges can allow for the awakening of these silent gene clusters.[67] Additionally, changes in culturing conditions can upregulate metabolites from "silent" gene clusters.[68] Heterologous expression is a technique that can be used to express a gene or multiple genes in an alternative host, allowing for the production of secondary metabolites that are either not produced or are produced only in low quantity.[64] This technique has been used in the cyanobacteria *Anabaena sp.* to produce lyngbyatoxin A, pendolmycin teleocidin B-4, and cyrptomaldamide.[69,70]

1.3.3 Mass Spectrometry Guided

Mass spectrometry can be used for structure elucidation as well as to guide isolation of secondary metabolites. Tools such as GNPS (Global Natural Products Social Molecular Networking), utilizes the tandem mass spectrometry data to group together chemically similar compounds.[71] Molecular networking is a useful tool that can be used to dereplicate known natural products, allowing for faster isolation of novel secondary metabolites.[72] Molecular networking can also be used to identify bioactive molecules, called bioactive molecular networking. This strategy was used on a *E. dendroides* extract, which led to the identification of antiviral diterpenoids that were previously missed by bioassay fractionation.[73] Microcolins E-M were isolated through a bioassay guided and LC MS/MS molecular networking approach. The fractions showed cytotoxicity to H-460 cancer cells and the molecular network clustered previously isolated microcolins with 9 new microcolins.[74]

1.3.4 Advances in NMR Spectroscopy

1.3.4.1 2D/3D techniques

Structure elucidation is a key aspect of drug discovery and typically utilizes a variety of NMR (nuclear magnetic resonance) experiments to solve the structure of an unknown candidate

therapeutic. NMR spectroscopy allows for the visualization of chemically unique proton and carbon atoms. Whereas 1D NMR experiments afford valuable information about the different types of atoms, 2D experiments show the connectivity between different atoms and therefore are very powerful in structure elucidation. Common 2D experiments used in structure elucidation included ^{1}H-^{13}C HSQC, ^{1}H-^{13}C HMBC, ^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H ROESY and ^{1}H-^{1}H NOESY.[75] Each 2D experiment utilizes a different pulse sequence, which probes the molecule of interest in a different way, giving high specialized information. For example, the ^{1}H-^{13}C HSQC experiment gives information about the one-bond connectivities between proton and carbon atoms in a molecule. These data, in combination with other experiments, can be used to deduce the overall structure of the compound.

3D NMR experiments have also started to gain broader utilization. Labeled peptides have been characterized by 3D experiments that have the dimensions of ^{13}C, ^{15}N, and ^{1}H.[76] 3D diffusion experiments have been developed by combining a typical 2D pulse sequence with a diffusion pulse sequence.[77] This concept has led to the rise of many different experiments including DOSY-HMQC, which was used to separate a mixture of quinine, camphene, and geraniol in the diffusion dimension, with each compound on a different HMQC plane.[78] An additional 3D experiment that was successful in separating mixtures was TOCSY-iDOSY.[79]

1.3.4.2 FAST NMR methods

NMR spectroscopy is integral for drug development, however, it can be costly and time consuming. Therefore, acceleration of two-dimensional NMR experiments has been an important area of research. One way that faster acquisition times can be achieved is through the shortening of the pulse sequence of the experiment. Techniques such as NUS (nonuniform sampling) can be used to shorten the acquisition time of a HSQC experiment through sampling only a fraction of

the points in the indirect dimension.[80,81] The points which are sampled in a non-uniform manner in these experiments can be selected through different sampling schemes such as radial, Poisson, random and EMS (envelope matched sampling).[82] The Poisson sampling scheme is often chosen because it has fewer artifacts but maintains sensitivity.[82,83] The points that are not sampled in a NUS scheme get reconstructed after the completion of the experiment and before Fourier transformation using different algorithms, e.g. compressed sensing and Iterative Soft Thresholding (IST) .[84,85,86] ASAP (Acceleration by Sharing Adjacent polarization) is a technique that speeds up the acquisition time by reducing the relaxation delay in the pulse sequence.[87] Linear prediction is used to extend out the missing points from the experiment and complete the FID.[88] SMART (SMAll Recovery Times) is a similar technique that reduces the length of the experiment by reducing the recovery delay, however it requires 3-dimensional field gradients.[89,90] The Earnst angle first introduced the reduction of recovery delay and other techniques have been developed that utilize this delay including SOFAST-HMQC, SMART, ASAP and the BEST approach.[91,88]

1.3.4.3 SMART

The NMR based tool Small Molecule Accurate Recognition Technology 2.0 (SMART 2.0) utilizes machine learning to place a query HSQC spectra into a 180D space of over 27,000 spectra at the time of publication, and now the database includes 104,349 spectra for version SMART 2.1 (https://smart.ucsd.edu/classic).[92] This is an update to the original tool, SMART 1.0, which used 2,048 spectra and a deep conventional network with a Siamese architecture.[93] SMART 2.0 allows for the rapid dereplication of compounds based on their NMR spectra. Additionally, based on the location of the compound in 180 dimensional space, its relationship to other compounds in the database allows for information to be obtained about the structure class

of the compound in question, as well as gives insight into other aspects of its structure. The discovery of several new sesquiterpene lactones was guided through the use of the tool SMART 2.0. Through inputting HSQC data of the chromatography fractions, this tool enabled the targeted isolation of 10 new compounds.[94]

1.4 Dissertation Chapters

1.4.1 Chapter 1: Introduction

The research presented here highlights the isolation and characterization of secondary metabolites from marine cyanobacteria for their potential use as therapeutics to a wide variety of diseases. Additionally, the use of NMR is critical to elucidate the structure and techniques to reduce acquisition time are also presented. Natural products have played an important role in drug development and design. Specifically, marine natural products offer a genetically diverse source of novel therapeutics. Approaches to isolate and characterize these bioactive compounds have included a variety of techniques including NMR, mass spectrometry and genome-guided approaches.

1.4.2 Chapter 2: 2D NMR experiments for Accelerating Natural Product Discovery

Chapter 2 concentrates on the acceleration of two-dimensional NMR experiments. ^{1H}-^{13}C HSQC experiments are one of the standard 2D experiments that are utilized in structure elucidation of natural product medicines. However, one common problem that plagues natural products investigations is the limited quantities of material. Higher quantities of material typically allow for NMR experiments that are quicker and have better signal-to-noise than when doing the same experiment on smaller quantities. When there are limited quantities of material, an HSQC experiment can be modified to record additional scans so as to obtain a better signal-

to-noise. However, this is both costly and time consuming. In this chapter, the focus was on two techniques, ASAP and NUS, to speed up the rate of data accumulation for an HSQC experiment. We tested this suite of experiments on a standard compound and on the novel compound fatuamide A, the structure elucidation of which will be presented in the following chapter.

1.4.3 Chapter 3: Characterization of Fatuamide A

Chapter 3, focuses on a culture of cyanobacteria (ASX22JUL!4-2) that was originally collected by SCUBA diving in coastal waters of American Samoa and has been cultured in laboratory at UC San Diego. This culture has been identified as a *Leptolyngbya* sp. after DNA extraction and sequencing by Nanopore minION®. After extraction and fractionation, the most cytotoxic fraction to NCI-H460 human lung cells afforded the novel compound fatuamide A. This chapter focuses on the structure elucidation and characterization of this novel compound. The final structure was determined through a multitude of NMR experiments, mass spectrometry, and genomic data, and suggested a hybrid PKS/NRPS cyclic peptide. The organism was genetically sequenced and a putative biosynthetic gene cluster has been identified. Furthermore, analysis of the gene cluster suggests that fatuamide A may have the ability to chelate to metals. Through analysis of the genomic data from the cultured cyanobacterium, a biosynthetic proposal is presented.

1.4.4 Chapter 4: Structure Determination of Curacin Analogs

Chapter 4 describes the characterization of natural analogs of the antimitotic agent curacin A. These analogues were isolated from a collection of cyanobacteria that was made near Culebra, Puerto Rico through SCUBA diving. The cyanobacteria was initially identified as *Moorena producens*. After extraction and fractionation, multiple fractions were noted to have cytotoxicity to H460 human lung carcinoma cells *in vitro.* An LCMS/MS network was made of

the different fractions and curacin A was identified in the sample. This led to the targeted isolation and structure elucidation of compounds in the same MS/MS molecular network cluster as curacin A, giving rise to several new cuarcin A analogues.

1.4.5 Chapter 5: Conclusions and future work

Chapter 5 focuses on the summary of all the research findings of the previous chapters. It highlights the novel secondary metabolites and the techniques utilized to isolate and characterize them. It also discusses the future directions of each of these research projects that have been outlined in the chapters previous.

1.4.6 Appendix: Cylic (Alkyl)(Amino) Carbene (CAAC) Gold(I) Complexes as Chemotherapuetic Agents

This appendix chapter depicts novel therapeutics that possess a gold complex with a CAAC ligand. These gold complexes were modeled after the gold-containing drug Auranofin (used for the treatment of rheumatoid arthritis), but were designed for their potential to be more potent and selective inhibitors in the treatment of various forms of cancer. This appendix chapter will highlight the cytotoxicity of these complexes across different cancer cell lines. Auranofin's mechanism of action is through inhibition of the enzyme thioredoxin reductase. Therefore, the CAAC complexes were tested for their inhibition of this enzyme as well. The selectivity of the binding of the CAAC complexes to the target was analyzed through investigating their off target binding to serum albumin.

1.5 References:

(1) Marupudi, N. I.; Han, J. E.; Li, K. W.; Renard, V. M.; Tyler, B. M.; Brem, H. Paclitaxel: A Review of Adverse Toxicities and Novel Delivery Strategies. *Expert Opin. Drug Saf.* **2007**, *6* (5), 609–621. https://doi.org/10.1517/14740338.6.5.609.

(2) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly

Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(3) Mayer, A. M. S. Marine Pharmacology.

(4) Bai, R.; Paull, K. D.; Herald, C. L.; Malspeis, L.; Pettit, G. R.; Hamel, E. Halichondrin B and Homohalichondrin B, Marine Natural Products Binding in the Vinca Domain of Tubulin: Discovery of Tubulin-Based Mechanism of Action by Analysis of Differential Cytotoxicity Data. *J. Biol. Chem.* **1991**, *266* (24), 15882–15889.

(5) Aicher, T. D.; Buszek, K. R.; Fang, F. G.; Forsyth, C. J.; Jung, S. H.; Kishi, Y.; Matelich, M. C.; Scola, P. M.; Spero, D. M.; Yoon, S. K. Total Synthesis of Halichondrin B and Norhalichondrin B. *Journal of the American Chemical Society*. American Chemical Society April 1, 1992, pp 3162–3164. https://doi.org/10.1021/ja00034a086.

(6) Towle, M. J.; Salvato, K. A.; Budrow, J.; Wels, B. F.; Kuznetsov, G.; Aalfs, K. K.; Welsh, S.; Zheng, W.; Seletsky, B. M.; Palme, M. H.; Habgood, G. J.; Singer, L. A.; Dipietro, L. V.; Wang, Y.; Chen, J. J.; Quincy, D. A.; Davis, A.; Yoshimatsu, K.; Kishi, Y.; Yu, M. J.; Littlefield, B. A. In Vitro and in Vivo Anticancer Activities of Synthetic Macrocyclic Ketone Analogues of Halichondrin B. *Cancer Res.* **2001**, *61* (3), 1013–1021.

(7) Bauer, A. Story of Eribulin Mesylate: Development of the Longest Drug Synthesis; 2016; pp 209–270. https://doi.org/10.1007/7081_2016_201.

(8) Cortes, J.; O'Shaughnessy, J.; Loesch, D.; Blum, J. L.; Vahdat, L. T.; Petrakova, K.; Chollet, P.; Manikas, A.; Diéras, V.; Delozier, T.; Vladimirov, V.; Cardoso, F.; Koh, H.; Bougnoux, P.; Dutcus, C. E.; Seegobin, S.; Mir, D.; Meneses, N.; Wanders, J.; Twelves, C. Eribulin Monotherapy versus Treatment of Physician's Choice in Patients with Metastatic Breast Cancer (EMBRACE): A Phase 3 Open-Label Randomised Study. *Lancet* **2011**, *377* (9769), 914–923. https://doi.org/10.1016/S0140-6736(11)60070-6.

(9) Patterson, G. M. L.; Larsen, L. K.; Moore, R. E. Bioactive Natural Products from Blue-Green Algae. *J. Appl. Phycol.* **1994**, *6* (2), 151–157. https://doi.org/10.1007/BF02186069.

(10) Nagle, D. G.; Paul, V. J. Production of Secondary Metabolites by Filamentous Tropical Marine Cyanobacteria: Ecological Functions of the Compounds. *J. Phycol.* **1999**, *35* (6 SUPPL.), 1412–1421. https://doi.org/10.1046/j.1529-8817.1999.3561412.x.

(11) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. Biologically Active Secondary Metabolites from Marine Cyanobacteria. *Curr. Opin. Biotechnol.* **2010**, *21* (6), 787–793. https://doi.org/10.1016/j.copbio.2010.09.019.

(12) Villa, F. A.; Lieske, K.; Gerwick, L. Selective MyD88-Dependent Pathway Inhibition by the Cyanobacterial Natural Product Malyngamide F Acetate. *Eur. J. Pharmacol.* **2010**, *629* (1–3), 140–146. https://doi.org/10.1016/j.ejphar.2009.12.002.

(13) Malloy, K. L.; Villa, F. A.; Engene, N.; Matainaho, T.; Gerwick, L.; Gerwick, W. H. Malyngamide 2, an Oxidized Lipopeptide with Nitric Oxide Inhibiting Activity from a

Papua New Guinea Marine Cyanobacterium. *J. Nat. Prod.* **2011**, *74* (1), 95–98. https://doi.org/10.1021/np1005407.

(14) Appleton, D. R.; Sewell, M. A.; Berridge, M. V.; Copp, B. R. A New Biologically Active Malyngamide from a New Zealand Collection of the Sea Hare Bursatella Leachii. *J. Nat. Prod.* **2002**, *65* (4), 630–631. https://doi.org/10.1021/np010511e.

(15) Orjala, J.; Nagle, D. G.; Hsu, V. L.; Gerwick, W. H. Antillatoxin: An Exceptionally Ichthyotoxic Cyclic Lipopeptide from the Tropical Cyanobacterium Lyngbya Majuscula. *J. Am. Chem. Soc.* **1995**, *117* (31), 8281–8282. https://doi.org/10.1021/ja00136a031.

(16) Berman, F. W.; Gerwick, W. H.; Murray, T. F. Antillatoxin and Kalkitoxin, Ichthyotoxins from the Tropical Cyanobacterium Lyngbya Majuscula, Induce Distinct Temporal Patterns of NMDA Receptor-Mediated Neurotoxicity. *Toxicon* **1999**, *37* (11), 1645–1648. https://doi.org/10.1016/S0041-0101(99)00108-7.

(17) Nakanishi, K. The Chemistry of Brevetoxins: A Review. *Toxicon* **1985**, *23* (3), 473–479. https://doi.org/10.1016/0041-0101(85)90031-5.

(18) Cao, Z.; Gerwick, W. H.; Murray, T. F. Antillatoxin Is a Sodium Channel Activator That Displays Unique Efficacy in Heterologously Expressed RNav1.2, RNav1.4 and RNav1.5 Alpha Subunits. *BMC Neurosci.* **2010**, *11*, 1–13. https://doi.org/10.1186/1471-2202-11-154.

(19) Dybdal-Hargreaves, N. F.; Risinger, A. L.; Mooberry, S. L. Eribulin Mesylate: Mechanism of Action of a Unique Microtubule Targeting Agent. *Clin. Cancer Res.* **2015**, *21* (11), 2445–2452. https://doi.org/10.1016/j.physbeh.2017.03.040.

(20) Pereira, A. R.; Kale, A. J.; Fenley, A. T.; Byrum, T.; Debonsi, H. M.; Gilson, M. K.; Valeriote, F. A.; Moore, B. S.; Gerwick, W. H. The Carmaphycins: New Proteasome Inhibitors Exhibiting an α,β-Epoxyketone Warhead from a Marine Cyanobacterium. *ChemBioChem* **2012**, *13* (6), 810–817. https://doi.org/10.1002/cbic.201200007.

(21) LaMonte, G. M.; Almaliti, J.; Bibo-Verdugo, B.; Keller, L.; Zou, B. Y.; Yang, J.; Antonova-Koch, Y.; Orjuela-Sanchez, P.; Boyle, C. A.; Vigil, E.; Wang, L.; Goldgof, G. M.; Gerwick, L.; O'Donoghue, A. J.; Winzeler, E. A.; Gerwick, W. H.; Ottilie, S. Development of a Potent Inhibitor of the Plasmodium Proteasome with Reduced Mammalian Toxicity. *J. Med. Chem.* **2017**, *60* (15), 6721–6732. https://doi.org/10.1021/acs.jmedchem.7b00671.

(22) Luesch, H.; Moore, R. E.; Paul, V. P.; Mooberry, S. L.; Corbett, T. H. Isolation of Dolastatin 10 from the Marine Cyanobacterium Symploca Species VP642 and Total Stereochemistry and Biological Evaluation of Its Analogue Symplostatin 1. *J. Nat. Prod.* **2001**, *64* (7), 907–910. https://doi.org/10.1021/np010049y.

(23) Pettit, G. R.; Kamano, Y.; Herald, C. L.; Fujii, Y.; Kizu, H.; Boyd, M. R.; Boettner, F. E.; Doubek, D. L.; Schmidt, J. M.; Chapuis, J. C.; Michel, C. Isolation of Dolastatins 10-15 from the Marine Mollusc Dolabella Auricularia. *Tetrahedron* **1993**, *49* (41), 9151–9170.

https://doi.org/10.1016/0040-4020(93)80003-C.

(24) Amador, M. L.; Jimeno, J.; Paz-Ares, L.; Cortes-Funes, H.; Hidalgo, M. Progress in the Development and Acquisition of Anticancer Agents from Marine Sources. *Ann. Oncol.* **2003**, *14* (11), 1607–1615. https://doi.org/10.1093/annonc/mdg443.

(25) Newman, D. J.; Cragg, G. M. Current Status of Marine-Derived Compounds as Warheads in Anti-Tumor Drug Candidates. *Mar. Drugs* **2017**, *15* (4). https://doi.org/10.3390/md15040099.

(26) Pettit, G. R.; Hogan, F.; Toms, S. Antineoplastic Agents. 592. Highly Effective Cancer Cell Growth Inhibitory Structural Modifications of Dolastatin 10. *J. Nat. Prod.* **2011**, *74* (5), 962–968. https://doi.org/10.1021/np1007334.

(27) Doronina, S.; Senter, P. D.; Toki, B. E. Pentapeptide Compounds and Uses Related Thereto. WO Patent 200288172, 2002.

(28) Wang, E.; Sorolla, M. A.; Krishnan, P. D. G.; Sorolla, A. From Seabed to Bedside: A Review on Promising Marine Anticancer Compounds. *Biomolecules* **2020**, *10* (2). https://doi.org/10.3390/biom10020248.

(29) Yu, B.; Jiang, T.; Liu, D. BCMA-Targeted Immunotherapy for Multiple Myeloma. *J. Hematol. Oncol.* **2020**, *13* (1), 1–24. https://doi.org/10.1186/s13045-020-00962-7.

(30) Lonial, S.; Lee, H. C.; Badros, A.; Trudel, S.; Nooka, A. K.; Chari, A.; Abdallah, A. O.; Callander, N.; Lendvai, N.; Sborov, D.; Suvannasankha, A.; Weisel, K.; Karlin, L.; Libby, E.; Arnulf, B.; Facon, T.; Hulin, C.; Kortüm, K. M.; Rodríguez-Otero, P.; Usmani, S. Z.; Hari, P.; Baz, R.; Quach, H.; Moreau, P.; Voorhees, P. M.; Gupta, I.; Hoos, A.; Zhi, E.; Baron, J.; Piontek, T.; Lewis, E.; Jewell, R. C.; Dettman, E. J.; Popat, R.; Esposti, S. D.; Opalinska, J.; Richardson, P.; Cohen, A. D. Belantamab Mafodotin for Relapsed or Refractory Multiple Myeloma (DREAMM-2): A Two-Arm, Randomised, Open-Label, Phase 2 Study. *Lancet Oncol.* **2020**, *21* (2), 207–221. https://doi.org/10.1016/S1470-2045(19)30788-0.

(31) Administration, U. S. F. and D. FDA Granted Accelerated Approval to Belantamab Mafodotin-Blmf for Multiple Myeloma. **2020**, 1–2.

(32) Newman, D. J. The " Utility " of Highly Toxic Marine-Sourced Compounds. **2019**, No. April.

(33) Singh, R. K.; Tiwari, S. P.; Rai, A. K.; Mohapatra, T. M. Cyanobacteria: An Emerging Source for Drug Discovery. *J. Antibiot. (Tokyo).* **2011**, *64* (6), 401–412. https://doi.org/10.1038/ja.2011.21.

(34) Leão, P. N.; Ramos, V.; Gonçalves, P. B.; Viana, F.; Lage, O. M.; Gerwick, W. H.; Vasconcelos, V. M. Chemoecological Screening Reveals High Bioactivity in Diverse Culturable Portuguese Marine Cyanobacteria. *Mar. Drugs* **2013**, *11* (4), 1316–1335. https://doi.org/10.3390/md11041316.

(35) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. The Natural Products Chemistry of Cyanobacteria. *Compr. Nat. Prod. II Chem. Biol.* **2010**, *2*, 141–188. https://doi.org/10.1016/b978-008045382-8.00041-1.

(36) Saini, D. K.; Pabbi, S.; Shukla, P. Cyanobacterial Pigments: Perspectives and Biotechnological Approaches. *Food Chem. Toxicol.* **2018**, *120* (June), 616–624. https://doi.org/10.1016/j.fct.2018.08.002.

(37) Coates, R. C.; Trentacoste, E.; Gerwick, W. H. Bioactive and Novel Chemicals from Microalgae. *Handb. Microalgal Cult. Appl. Phycol. Biotechnol. Second Ed.* **2013**, 504–531. https://doi.org/10.1002/9781118567166.ch26.

(38) Becker, W. E. Microalgae for Human and Animal Nutrition. In *Handbook of Microalgal Culture*; 2013; pp 461–503. https://doi.org/10.1002/9781118567166.

(39) Moss, N. A.; Leao, T.; Glukhov, E.; Gerwick, L.; Gerwick, W. H. *Collection, Culturing, and Genome Analyses of Tropical Marine Filamentous Benthic Cyanobacteria*, 1st ed.; Elsevier Inc., 2018; Vol. 604. https://doi.org/10.1016/bs.mie.2018.02.014.

(40) Li, Y.; Naman, C. B.; Alexander, K. L.; Guan, H.; Gerwick, W. H. The Chemistry, Biochemistry and Pharmacology of Marine Natural Products from Leptolyngbya, a Chemically Endowed Genus of Cyanobacteria. *Mar. Drugs* **2020**, *18* (10). https://doi.org/10.3390/md18100508.

(41) Crnkovic, C. M.; May, D. S.; Orjala, J. The Impact of Culture Conditions on Growth and Metabolomic Profiles of Freshwater Cyanobacteria. **2019**, *30* (1), 375–384. https://doi.org/10.1007/s10811-017-1275-3.The.

(42) Y. Hirose, T. Fijusawa, Y. Ohtsubo, M. Katayama, N. Misawa, S. Wakazuki, Y. Shimura, Y. Nakamura, M. Kawachi, H. Yoshikawa, T. Eki, Y. K. Complete Genome Sequence of Cyanobacterium. *Am. Soc. Microbiol.* **2016**, *4* (2), 2–3. https://doi.org/10.1128/genomeA.00090-16.Copyright.

(43) Medina, R. A.; Goeger, D. E.; Hills, P.; Mooberry, S. L.; Huang, N.; Romero, L. I.; Ortega-Barría, E.; Gerwick, W. H.; McPhail, K. L. Coibamide A, a Potent Antiproliferative Cyclic Depsipeptide from the Panamanian Marine Cyanobacterium Leptolyngbya Sp. *J. Am. Chem. Soc.* **2008**, *130* (20), 6324–6325. https://doi.org/10.1021/ja801383f.

(44) Leao, T.; Castelão, G.; Korobeynikov, A.; Monroe, E. A.; Podell, S.; Glukhov, E.; Allen, E. E.; Gerwick, W. H.; Gerwick, L. Comparative Genomics Uncovers the Prolific and Distinctive Metabolic Potential of the Cyanobacterial Genus Moorea. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (12), 3198–3203. https://doi.org/10.1073/pnas.1618556114.

(45) Tronholm, A.; Engene, N. *Moorena Gen. Nov.*, a Valid Name for "*Moorea* Engene & *Al.*" *Nom. Inval.* (*Oscillatoriaceae, Cyanobacteria*). *Not. algarum* **2019**, *122* (20 December 2019), 122.

(46) Jiang, W.; Zhou, W.; Uchida, H.; Kikumori, M.; Irie, K.; Watanabe, R.; Suzuki, T.; Sakamoto, B.; Kamio, M.; Nagai, H. A New Lyngbyatoxin from the Hawaiian Cyanobacterium Moorea Producens. *Mar. Drugs* **2014**, *12* (5), 2748–2759. https://doi.org/10.3390/md12052748.

(47) Cardellina II, J. H.; Marner, F.-J.; Moore, R. E. Seaweed Dermatitis : Structure of Lyngbyatoxin A. *Science (80-.).* **2013**, *204* (4389), 193–195.

(48) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. Structure of Curacin A, a Novel Antimitotic, Antiproliferative and Brine Shrimp Toxic Natural Product from the Marine Cyanobacterium Lyngbya Majuscula. *J. Org. Chem.* **1994**, *59* (6), 1243–1245. https://doi.org/10.1021/jo00085a006.

(49) Sandy, M.; Butler, A. Microbial Iron Acquisition: Marine and Terrestrial Siderophores. *Chem. Rev.* **2009**, *109* (10), 4580–4595. https://doi.org/10.1021/cr9002787.

(50) Hider, R. C.; Kong, X. Chemistry and Biology of Siderophores. *Nat. Prod. Rep.* **2010**, *27* (5), 637–657. https://doi.org/10.1039/b906679a.

(51) Miethke, M.; Marahiel, M. A. Siderophore-Based Iron Acquisition and Pathogen Control. *Microbiol. Mol. Biol. Rev.* **2007**, *71* (3), 413–451. https://doi.org/10.1128/MMBR.00012-07.

(52) Gademann, K.; Portmann, C. Secondary Metabolites from Cyanobacteria: Complex Structures and Powerful Bioactivities. *Curr. Org. Chem.* **2008**, *12* (4), 326–341. https://doi.org/10.2174/138527208783743750.

(53) Hofmann, M.; Retamal-Morales, G.; Tischler, D. Metal Binding Ability of Microbial Natural Metal Chelators and Potential Applications. *Nat. Prod. Rep.* **2020**. https://doi.org/10.1039/c9np00058e.

(54) Kenney, G. E.; Rosenzweig, A. C. Chalkophores. *Annu. Rev Biochem* **2018**, *87*, 645–676. https://doi.org/10.1146/annurev-biochem-062917-012300.Chalkophores.

(55) Liermann, L. J.; Guynn, R. L.; Anbar, A.; Brantley, S. L. Production of a Molybdophore during Metal-Targeted Dissolution of Silicates by Soil Bacteria. *Chem. Geol.* **2005**, *220* (3–4), 285–302. https://doi.org/10.1016/j.chemgeo.2005.04.013.

(56) Łoboda, D.; Rowińska-Zyrek, M. Candida Albicans Zincophore and Zinc Transporter Interactions with Zn(II) and Ni(II). *Dalt. Trans.* **2018**, *47* (8), 2646–2654. https://doi.org/10.1039/c7dt04403h.

(57) Ivanov, A. S.; Parker, B. F.; Zhang, Z.; Aguila, B.; Sun, Q.; Ma, S.; Jansone-Popova, S.; Arnold, J.; Mayes, R. T.; Dai, S.; Bryantsev, V. S.; Rao, L.; Popovs, I. Siderophore-Inspired Chelator Hijacks Uranium from Aqueous Medium. *Nat. Commun.* **2019**, *10* (1), 1–7. https://doi.org/10.1038/s41467-019-08758-1.

(58) Minigh, J. Auranofin. *xPharm Compr. Pharmacol. Ref.* **2007**, No. I, 1–6.

https://doi.org/10.1016/B978-008055232-3.61268-1.

(59) Dasari, S.; Bernard Tchounwou, P. Cisplatin in Cancer Therapy: Molecular Mechanisms of Action. *Eur. J. Pharmacol.* **2014**, *740*, 364–378. https://doi.org/10.1016/j.ejphar.2014.07.025.

(60) Frei, A.; Zuegg, J.; Elliott, A. G.; Baker, M.; Braese, S.; Brown, C.; Chen, F.; G. Dowson, C.; Dujardin, G.; Jung, N.; King, A. P.; Mansour, A. M.; Massi, M.; Moat, J.; Mohamed, H. A.; Renfrew, A. K.; Rutledge, P. J.; Sadler, P. J.; Todd, M. H.; Willans, C. E.; Wilson, J. J.; Cooper, M. A.; Blaskovich, M. A. T. Metal Complexes as a Promising Source for New Antibiotics. *Chem. Sci.* **2020**, *11* (10), 2627–2639. https://doi.org/10.1039/c9sc06460e.

(61) Möllmann, U.; Heinisch, L.; Bauernfeind, A.; Köhler, T.; Ankel-Fuchs, D. Siderophores as Drug Delivery Agents: Application of the "Trojan Horse" Strategy. *BioMetals* **2009**, *22* (4), 615–624. https://doi.org/10.1007/s10534-009-9219-2.

(62) Haque, F.; Banayan, S.; Yee, J.; Chiang, Y. W. Extraction and Applications of Cyanotoxins and Other Cyanobacterial Secondary Metabolites. *Chemosphere* **2017**, *183*, 164–175. https://doi.org/10.1016/j.chemosphere.2017.05.106.

(63) Bergendal, E. Hoiamide D, a Marine Cyanobacteria-Derived Inhibitor of P53/ MDM2 Interaction Karla. *Bone* **2008**, *23* (1), 1–7. https://doi.org/10.1016/j.bmcl.2011.10.054.Hoiamide.

(64) Luo, Y.; Cobb, R. E.; Zhao, H. Recent Advances in Natural Product Discovery. *Curr. Opin. Biotechnol.* **2014**, *30*, 230–237. https://doi.org/10.1016/j.copbio.2014.09.002.

(65) Ziemert, N.; Weber, T.; Medema, M. H. *Genome Mining Approaches to Bacterial Natural Product Discovery*, 3rd ed.; Elsevier Ltd., 2020. https://doi.org/10.1016/b978-0-12-409547-2.14627-x.

(66) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(67) Walker, J. M. *Fungal Secondary Metabolism - Methods and Protocols*; 2009; Vol. 531.

(68) Tomm, H. A.; Ucciferri, L.; Ross, A. C. Advances in Microbial Culturing Conditions to Activate Silent Biosynthetic Gene Clusters for Novel Metabolite Production. *J. Ind. Microbiol. Biotechnol.* **2019**, *46* (9–10), 1381–1400. https://doi.org/10.1007/s10295-019-02198-y.

(69) Videau, P.; Wells, K. N.; Singh, A. J.; Eiting, J.; Proteau, P. J.; Philmus, B. Expanding the Natural Products Heterologous Expression Repertoire in the Model Cyanobacterium Anabaena Sp. Strain PCC 7120: Production of Pendolmycin and Teleocidin B-4. *ACS Synth. Biol.* **2020**, *9* (1), 63–75. https://doi.org/10.1021/acssynbio.9b00334.

(70) Taton, A.; Ecker, A.; Diaz, B.; Moss, N. A.; Anderson, B.; Reher, R.; Simkovsky, R.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H.; Lea, T. F.; Golden, J. W. Heterologous Expression of Cryptomaldamide in a Cyanobacterial Host. **2020**. https://doi.org/10.1021/acssynbio.0c00431.

(71) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L.A.; Quinn, R.A; Duncan, K.R.; Hsu, C.; Floros, D.J.; Gavilan, R.G.; Kleigrew,K.; Northen, T.; Dutton, R.J.; Parrot, D.; Carlosn, E.E.; Aigle, B., Michelsen, C.F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B.T.; Gerwick, L.; Liaw, C.; Yang, Y.; Humpf, H.; Maansson, M., Keyzers, R.A.; Sims, A.C.; Johnson, A.R.; Sidebottom, A.M.; Sedio, B.E.; Klitgaard, A.; Larson, C.B; Boya P, C.A.; Torres-Mendoza, D.; Gonzalez, D.J.; Silva, D.B.; Marques, L.M.; Demarque, D.P.; Pociute, E.; O'Neill, E.C.; Briand, E.; Helfrich, E.J.N.; Granatosky, E.A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J.J; Zeng, Y.; Vorholt, J.; Kurita, K.; Charusanti, P.; McPhail, K.L.; Nielsen, K.F.; Vuong, L.; Elfeki, M.; Traxler, M.F.; Engene, N.; Koyama, N.; Vining, O.B.; Baric, R.; Silva, R.R.; Mascuch, S.J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P.G.; Dai, J.; Neupane, R.; Gurr, J.; Rodrígues, A.M.C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B.M.; Almaliti, J.; Allard, P.; Phapale, P.; Nothias, L.; Alexandrov, T.; Litaudon, M.; Wolfender, J.; Kyle, J.E.; O Metz, T.; Peryea, T.; Nguyen, D.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K.M; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P.R.; O Palsson, B.; Pogliano, K.; Linington, R.G.; Gutiérrez, M.; Lopes, N.P.; Gerwick, W.H.; Moore, B.S.; Dorrestein, P.C.; Bandeira, N.; . Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837. https://doi.org/10.1038/nbt.3597.

(72) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonsi, H. M.; Gerwick, W. H.; Dorrestein, P. C. Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **2013**, *76* (9), 1686–1699. https://doi.org/10.1021/np400413s.

(73) Nothias, L. F.; Nothias-Esposito, M.; Da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J. Nat. Prod.* **2018**, *81* (4), 758–767. https://doi.org/10.1021/acs.jnatprod.7b00737.

(74) Yu, H. B.; Glukhov, E.; Li, Y.; Iwasaki, A.; Gerwick, L.; Dorrestein, P. C.; Jiao, B. H.; Gerwick, W. H. Cytotoxic Microcolin Lipopeptides from the Marine Cyanobacterium Moorea Producens. *J. Nat. Prod.* **2019**, *82* (9), 2608–2619. https://doi.org/10.1021/acs.jnatprod.9b00549.

(75) Reynolds, W. F. *Natural Product Structure Elucidation by NMR Spectroscopy*; Elsevier

Inc., 2017. https://doi.org/10.1016/B978-0-12-802104-0.00029-9.

(76) Zhang, F.; Adnani, N.; Vazquez-Rivera, E.; Braun, D. R.; Tonelli, M.; Andes, D. R.; Bugni, T. S. Application of 3D NMR for Structure Determination of Peptide Natural Products. *J. Org. Chem.* **2015**, *80* (17), 8713–8719. https://doi.org/10.1021/acs.joc.5b01486.

(77) Dal Poggetto, G.; CastanÌar, L.; Foroozandeh, M.; Kiraly, P.; Adams, R. W.; Morris, G. A.; Nilsson, M. Unexploited Dimension: New Software for Mixture Analysis by 3D Diffusion-Ordered NMR Spectroscopy. *Anal. Chem.* **2018**, *90* (22), 13695–13701. https://doi.org/10.1021/acs.analchem.8b04093.

(78) Barjat, H.; Morris, G. A.; Swanson, A. G. A Three-Dimensional DOSY-HMQC Experiment for the High-Resolution Analysis of Complex Mixtures. *J. Magn. Reson.* **1998**, *131* (1), 131–138. https://doi.org/10.1006/jmre.1997.1332.

(79) Birlirakis, N.; Guittet, E. A New Approach in the Use of Gradients for Size-Resolved 2D-NMR Experiments. *J. Am. Chem. Soc.* **1996**, *118* (51), 13083–13084. https://doi.org/10.1021/ja9627151.

(80) Sidebottom, P. J. A New Approach to the Optimisation of Non-Uniform Sampling Schedules for Use in the Rapid Acquisition of 2D NMR Spectra of Small Molecules. *Magn. Reson. Chem.* **2016**, No. April, 689–694. https://doi.org/10.1002/mrc.4444.

(81) Kuwata, K.; Brooks, D.; Yang, H.; Schleich, T. Exponential Sampling, an Alternative Method for Sampling in Two-Dimensional NMR Expirements. *J. Magn. Reson. Ser. B* **1994**, *104* (1), 11–25. https://doi.org/10.1006/jmrb.1994.1049.

(82) Mobil, Mehdi; Hoch, J. Nonuniform Sampling and Non-Fourier Signal Processing Methods in Multidimensional NMR Mehdi. *Prog Nucl Magn Reson Spectrosc.* **2014**, *83*, 21–41. https://doi.org/10.1016/j.pnmrs.2014.09.002.Nonuniform.

(83) Hyberts, S. G.; Takeuchi, K.; Wagner, G. Poisson-Gap Sampling and Forward Maximum Entropy Reconstruction for Enhancing the Resolution and Sensitivity of Protein NMR Data. *J. Am. Chem. Soc.* **2010**, *132* (7), 2145–2147. https://doi.org/10.1021/ja908004w.

(84) Kazimierczuk, K.; Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chemie - Int. Ed.* **2011**, *50* (24), 5556–5559. https://doi.org/10.1002/anie.201100370.

(85) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52* (4), 315–327. https://doi.org/10.1007/s10858-012-9611-z.

(86) Gołowicz, D.; Kasprzak, P.; Orekhov, V.; Kazimierczuk, K. Fast Time-Resolved NMR with Non-Uniform Sampling. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 40–55. https://doi.org/10.1016/j.pnmrs.2019.09.003.

(87) Kupce, Eriks, Freeman, R. Fast Multidimensional NMR by Polarization Sharing. *J. Med. Pharm. Chem.* **2001**, *2* (1), 1941–1944. https://doi.org/10.1002/mrc.

(88) Becker, J.; Luy, B. CLIP-ASAP-HSQC for Fast and Accurate Extraction of One-Bond Couplings from Isotropic and Partially Aligned Molecules. *Magn. Reson. Chem.* **2015**, *53* (11), 878–885. https://doi.org/10.1002/mrc.4276.

(89) Vitorge, B.; Bodenhausen, G.; Pelupessy, P. Speeding up Nuclear Magnetic Resonance Spectroscopy by the Use of SMAll Recovery Times - SMART NMR. *J. Magn. Reson.* **2010**, *207* (1), 149–152. https://doi.org/10.1016/j.jmr.2010.07.017.

(90) Ndukwe, I. E.; Shchukina, A.; Kazimierczuk, K.; Butts, C. P. Rapid and Safe ASAP Acquisition with EXACT NMR. *Chem. Commun.* **2016**, *52* (86), 12769–12772. https://doi.org/10.1039/c6cc07140f.

(91) Schulze-Sünninghausen, D.; Becker, J.; Luy, B. Rapid Heteronuclear Single Quantum Correlation NMR Spectra at Natural Abundance. *J. Am. Chem. Soc.* **2014**, *136* (4), 1242–1245. https://doi.org/10.1021/ja411588d.

(92) Reher, R.; Kim, H. W.; Zhang, C.; Mao, H. H.; Wang, M.; Nothias, L. F.; Caraballo-Rodriguez, A. M.; Glukhov, E.; Teke, B.; Leao, T.; Alexander, K. L.; Duggan, B. M.; Van Everbroeck, E. L.; Dorrestein, P. C.; Cottrell, G. W.; Gerwick, W. H. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *J. Am. Chem. Soc.* **2020**, *142* (9), 4114–4120. https://doi.org/10.1021/jacs.9b13786.

(93) Zhang, C.; Idelbayev, Y.; Roberts, N.; Tao, Y.; Nannapaneni, Y.; Duggan, B. M.; Min, J.; Lin, E. C.; Gerwick, E. C.; Cottrell, G. W.; Gerwick, W. H. Small Molecule Accurate Recognition Technology (SMART) to Enhance Natural Products Research. *Sci. Rep.* **2017**, *7* (1), 1–17. https://doi.org/10.1038/s41598-017-13923-x.

(94) Lee, J.; Park, J.; Kim, J.; Jeong, B.; Choi, S. Y.; Jang, H. S.; Yang, H. Targeted Isolation of Cytotoxic Sesquiterpene Lactones from Eupatorium Fortunei by the NMR Annotation Tool, SMART 2.0. *ACS Omega* **2020**, *5* (37), 23989–23995. https://doi.org/10.1021/acsomega.0c03270.

Chapter Two: Optimization of FAST 2D NMR experiments for Accelerating Natural Product Characterization

2.1 Abstract

NUS (Non-Uniform Sampling) and ASAP (Acceleration by Sharing Adjacent Polarization) are NMR techniques that allow for the rapid accumulation of data for two-dimensional NMR experiments. Fast techniques are integral for a cost and time effective approach to potential therapeutic compound discovery and development. This chapter highlights the application of NUS and ASAP protocols for HSQC experiments to increase the speed of data acquisition for natural products at different concentrations. These techniques work by shortening the acquisition time of the experiment and are processed using different algorithms such as IST (Iterative Shrinkage Threshold) and Compressed Sensing. It was found that an HSQC of a 12 mM solution of a standard compound can be completed in 3 min and 51 seconds using a combination of ASAP with 50% Non-Uniform Sampling, compared to the same number of scans took 47 min to complete without these techniques. These techniques were then tested on lower concentrations of the standard, to replicate lower concentrations found in natural product drug discovery, and on a novel natural product, fatuamide A. More rapid dereplication and structure elucidation can be achieved with the faster data collection that is explored in these studies.

2.2 Introduction

NMR is an essential tool in all fields of chemistry research. 2D NMR is crucial for its ability to observe the connection between atoms. While it is a critical technique for establishing and identifying structures, it can be very time intensive, and even more so when done on samples present in small quantity. Therefore, reducing the acquisition time of these experiments has been

of interest. Traditionally, NMR experiments involved sampling all of the points in the Nyquist grid for complete data acquisition. Developments with nonuniform sampling have shown a complete spectrum can be produced without sampling all of the points.[1]

Nonuniform sampling collects only a fraction of the data points in the indirect dimension. The amount of points that are sampled correlates to the time that the experiment takes. Nonuniform sampling takes advantage of the fact that only a certain amount of points in a NMR experiment contain the vital information to produce a complete spectrum, and the rest is baseline.[1] The optimally sampled spectrum will have the minimum amount of points, but still retains all the signals of a fully sampled spectrum. Algorithms are applied to a partially sampled spectrum to fill in the missing data and a full spectrum is thereby produced. The points that are sampled are important in terms of the reconstruction of the data.[2] Different sampling schemes such as radial, poisson, random and EMS (envelope matched sampling), can be applied such that a different pattern of points are sampled.[3] The poisson sampling scheme has been shown to have fewer artifacts but retain sensitivity according to literature reports, and for this reason, this was the sampling scheme used in this study.[3,4]

ASAP reduces the relaxation time between each scan; typically the relaxation time is decreased from 1-1.5 second to 60 ms, which necessarily makes the experiment significantly shorter. There have been a variety of techniques that have utilized shortening the relaxation time to decrease the experiment length, such as SO-FAST-HMQC, BEST approach, and the SMART approach.[5,6] There can be problems with the instrument overheating with too short of a delay, and this is a important consideration in planning an ASAP protocol.[7]

The selection of optimal processing protocols is an essential aspect for these experiments. The Nonuniform sampled spectra need to be reconstructed before Fourier Transformation. This

can be achieved through a variety of algorithms, including Iterative Soft Thresholding and Compressed Sensing.[1,8,9] Linear prediction is also used in processing for ASAP to extend out points that are missing from the experiment and complete the FID.[10]

In this study, I focused on combining these two methods, ASAP and NUS, within the same experiment, so as to minimize data collection time while not sacrificing spectral quality. Additionally, the goal was to apply this combination of techniques on relatively low concentrations of natural products, as the isolation of small quantities of natural products can be a common problem.

2.3 Results and Discussion

2.3.1 Selection of a Standard Compound

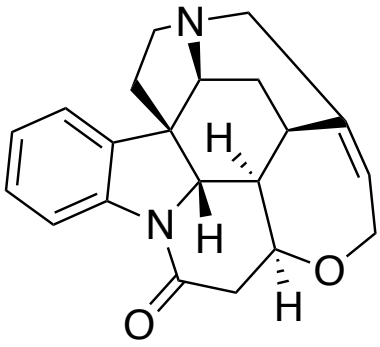


Figure 2.1: Structure of strychnine.

Strychnine (*figure 2.1*) is a natural product often utilized in NMR studies for its variety of chemical shifts, and thus was chosen a standard compound for these studies. *Figure 2.2* shows the results of a standard ^{1}H-^{13}C HSQC experiment without NUS or ASAP that was recorded in 47 min. *Figure 2.3* shows the results of an ^{1}H-^{13}C HSQC spectra that was accumulated with 50% Nonuniform Sampling and the ASAP protocol; it took only 3 min and 51 seconds to complete. This is a nearly 12 fold difference in the amount of time it took compared to the standard method (*figure 2.2*). However, there are limitations to how low of level of sampling can occur and still produce high quality data. The appearance of artifacts as well as the disappearance of peaks are

prevalent when too low a level of sampling is applied. This is evident when comparing 100% sampling and 3.125% NUS in *figure 2.4.* There are both missing peaks and artifacts in the 3.125% NUS.

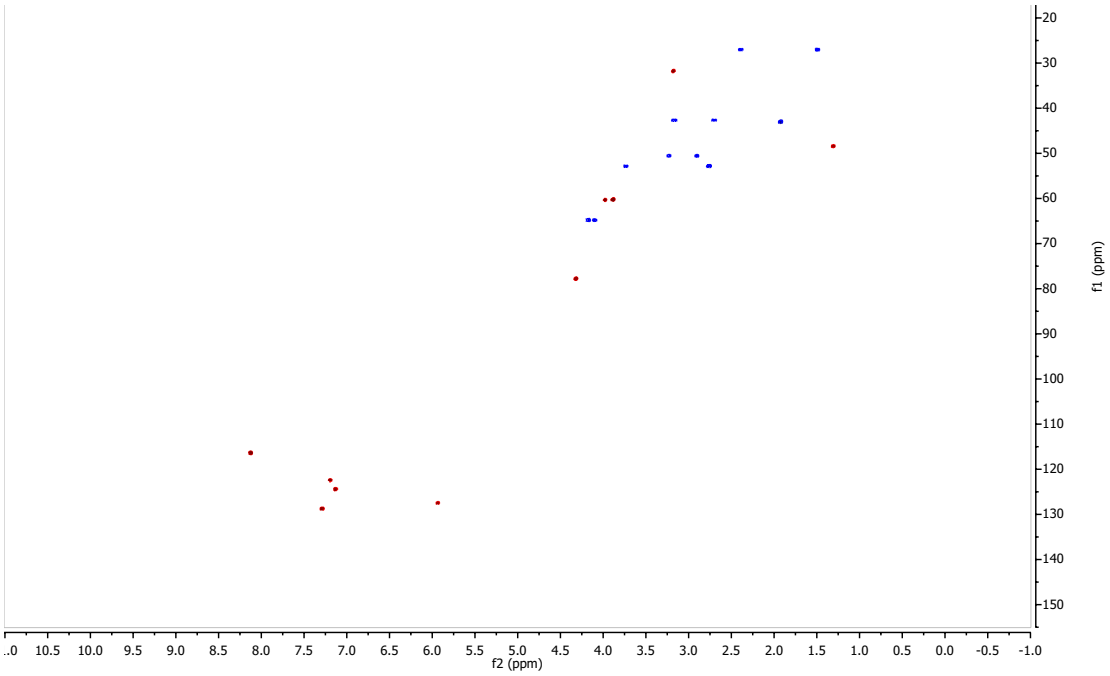


Figure 2.2: 100% Sampling of a multiplicity edited ^{1}H-^{13}C HSQC spectra with positive CH and CH_3 peaks (red) and negative CH_2 peaks (blue) of strychnine (12 mM, 600 MHz).

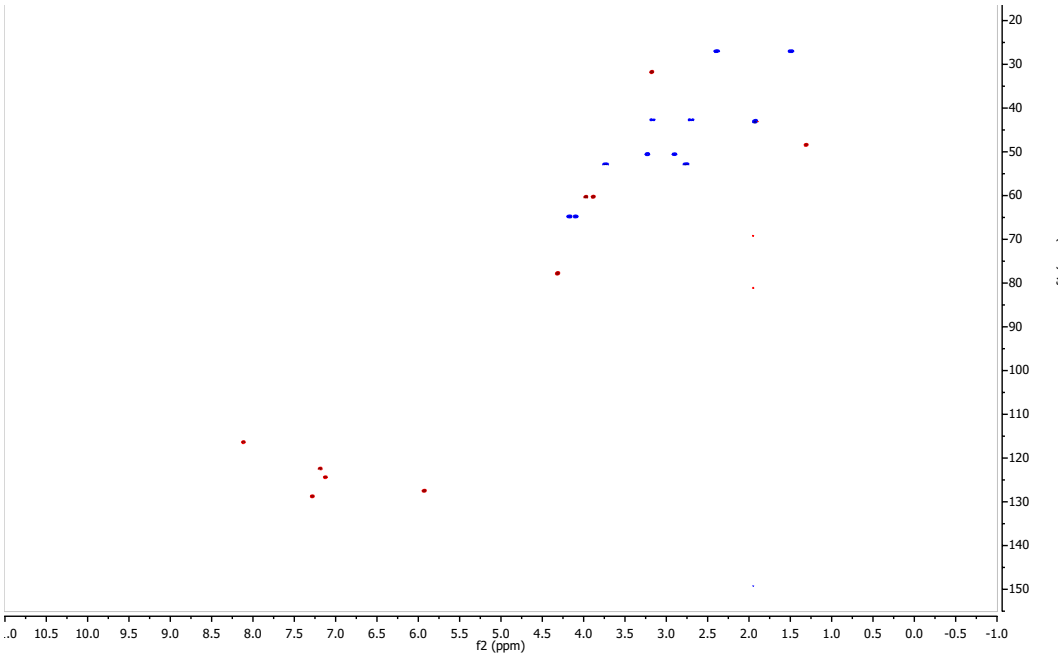


Figure 2.3: 50% Nonuniform Sampling (NUS) with application of the ASAP protocol to obtain a multiplicity edited ^{1}H-^{13}C HSQC spectra with positive CH and CH_3 peaks (red) and negative CH_2 peaks (blue) of strychnine (12 mM, 600 MHz). Note that this spectrum shows all of the same peaks as the 100% sampled spectra (figure 2.2).

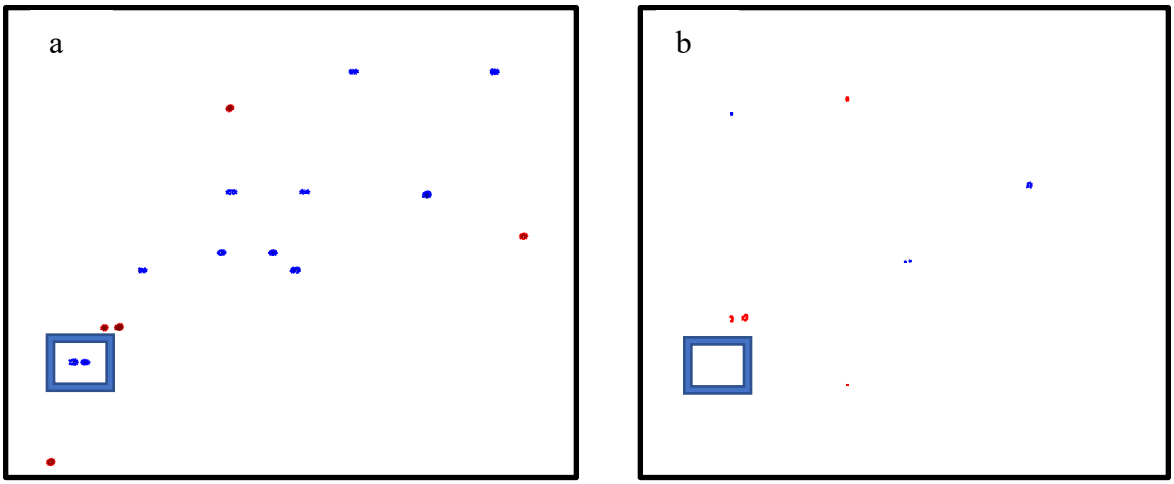


Figure 2.4: a. Zoomed in view of 100% sampled ^{1}H-^{13}C HSQC spectra with positive CH and CH_3 peaks (red) and negative CH_2 peaks (blue), and b. zoomed in view of 3.125% NUS spectra with positive CH and CH_3 peaks (red) and negative CH_2 peaks (blue) of strychnine (12 mM, 600 MHz). Blue box highlights two peaks that disappear at 3.125%.

To deduce the optimal level of sampling of NUS at varying concentrations, we did a series of experiments at different percentages of sampling (*table 2.1*). The table depicts the

length of the ^{1}H-^{13}C HSQC experiment, and above the redline shows by visual inspection the level of NUS that is acceptable without significant loss of data. At the lowest concentration tested (0.5 mM), a ^{1}H-^{13}C HSQC experiment could be recorded in 6 hours and 12 minutes (50% NUS); at a higher concentration (12 mM) a ^{1}H-^{13}C HSQC experiment could be recorded in 23 minutes and 30 seconds. To further reduce the acquisition time the nonuniform sampling was combined with ASAP (acceleration by Sharing Adjacent Polarization) (*table 2*). For the 12 mM sample with 50% NUS and no ASAP, the experiment took 23 minutes and 30 seconds (*table 1*); however, it only took 3 minutes and 51 seconds for the same level of NUS (50%) with the ASAP protocol applied (*table 2*). Each concentration had the same number scans per experiment. The number of scans was determined by first taking a ^{1}H NMR for an optimal signal-to-noise ratio of approximately 20:1. Therefore, combining the techniques of ASAP and NUS reduces acquisition time, and can be achieved for ^{1}H-^{13}C HSQC experiments without data loss.

Table 2.1: Strychnine at different concentrations (constant number of scans per column based on the number of scans needed to obtain an optimal S/N of approximately 20:1 in a trial ^{1}H NMR) with different levels of NUS sampling to obtain ^{1}H-^{13}C HSQC spectra. Below the red line indicates that artifacts or missing peaks were observed compared to the spectrum obtained without NUS or ASAP (Figure 2.2).

% Sampled	12 mM	1 mM	0.5 mM
100%	0:46:49	3:06:21	12:24:31
50%	0:23:30	1:33:15	6:12:04
25%	0:11:52	0:46:42	3:06:06
12.50%	0:06:03	0:23:29	1:33:08
10%	0:04:48	0:18:23	1:12:48
8%	0:03:42	0:14:02	0:55:24
6.25%	0:03:09	0:11:52	0:46:41
3.13%	0:01:43	0:06:04	0:23:28

2.3.2 Combination of NUS and ASAP techniques

Table 2.2 Standard strychnine at different concentrations (constant number of scans per column) and different levels of sampling with ASAP to obtain ^{1}H-^{13}C HSQC spectra. Below the red line indicates that artifacts or missing peaks were observed compared to the spectrum obtained without NUS or ASAP (Figure 2.2).

%Sampled	12 mM	1 mM	0.5 mM
100%	0:07:40	0:30:29	2:01:45
50%	0:03:51	0:15:12	1:00:37
25%	0:01:57	0:07:36	00:30:15
12.50%	0:01:00	0:03:49	0:15:07
10%	0:00:47	0:03:00	0:11:48
8%	0:00:43	0:02:18	0:09:00
6.25%	0:00:33	0:01:56	0:07:35
3.13%	0:00:18	0:01:00	0:03:49

2.3.3 Application of NUS and ASAP to the Analysis of a Natural Product

2.3.3.1 Nonuniform Sampling of Fatuamide A

Additionally, I evaluated the level of NUS sampling necessary for high quality ^{1}H-^{13}C HSQC spectra using the new natural product fatuamide A. Its isolation and characterization are further discussed in Chapter 3. Table 2.3 shows the amount of time each experiment took for each percent sampled. With application of these data acquisition techniques I was able to significantly reduce the acquisition time for this new natural product.

Table 2.3: Fatuamide A at different concentrations (constant number of scans per column) and different levels of sampling. Below the red line indicates the presence of artifacts or missing peaks from the ^{1}H-^{13}C HSQC spectra.

% Sampled	2.2 mM
100%	1:33:25
50%	0:46:45
25%	0:23:30
12.5%	0:11:53
10%	0:09:21
8%	0:07:11
6.25%	0:06:05
3.13%	0:03:11

2.3.4 Quantification of Signal-to-Noise Ratios in ^{1}H-^{13}C HSQC Spectra Accumulated with NUS and ASAP Protocols

One problem that arises from using too low of a sampling scheme is that some peaks are missing or additional non-real peaks are observed. These phenomena were observed in spectra below the red lines in Tables 2.1, 2.2, and 2.3. To further examine this problem, the signal-to-noise ratio was studied in each NMR experiment (*figure 2.5*). An interpretable spectrum needs an acceptable signal-to-noise ratio such that the peaks due to the compound of interest can be distinguished from noise. On the instrument used a 200:1 signal-to-noise provides a quality spectrum, and values even lower than that could also produce quality spectrum. The signal-to-noise ratio decreases as the percentage of nonuniform sampling decreases. Additionally, for each level of NUS, the entire spectrum was subtracted from the reference spectra (100% sampled) and the difference of the spectra was calculated as the RMS (root mean square). Higher RMS values show that the intensities and positions of the peaks are not well reconstructed. Lower sampling levels showed a greater difference between the nonuniform sampling spectra and the reference spectra and had a higher RMS value.

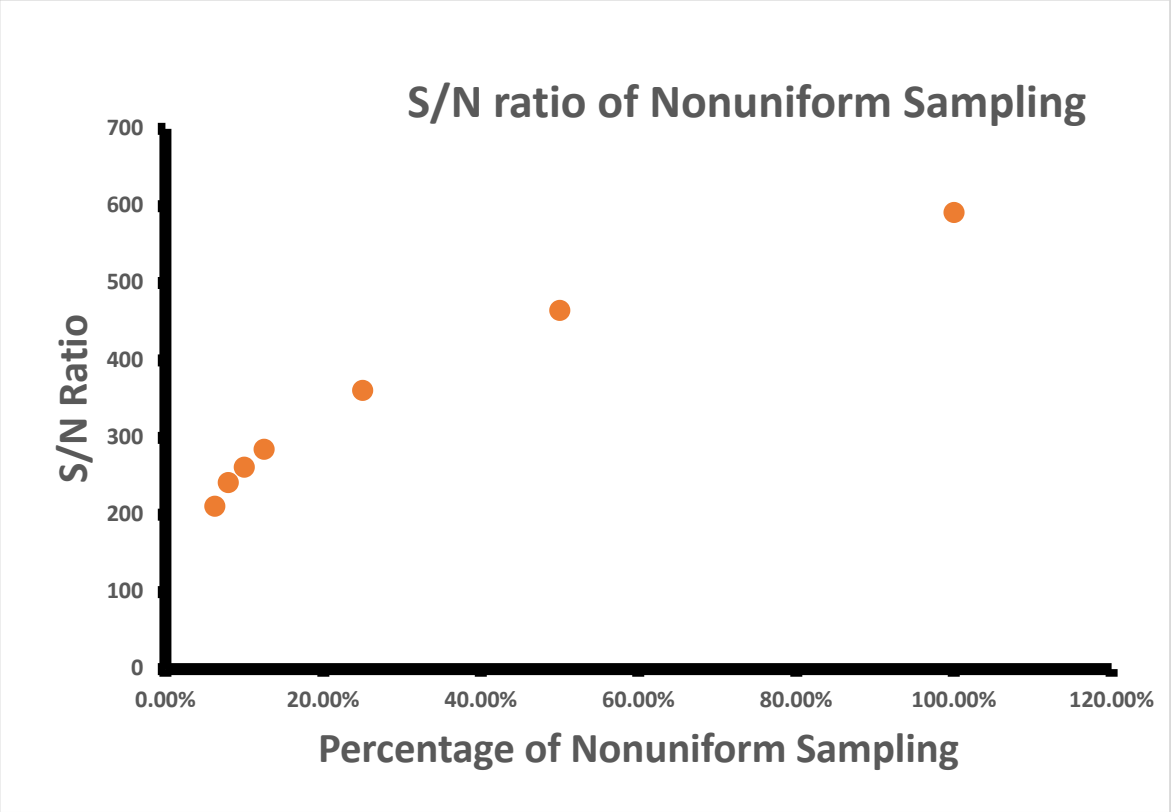


Figure 2.5: The signal-to-noise ratio of standard strychnine (12 mM) at different percentages of nonuniform sampling (NUS). Higher S/N ratio allows for easier and more reliable interpretation of data.

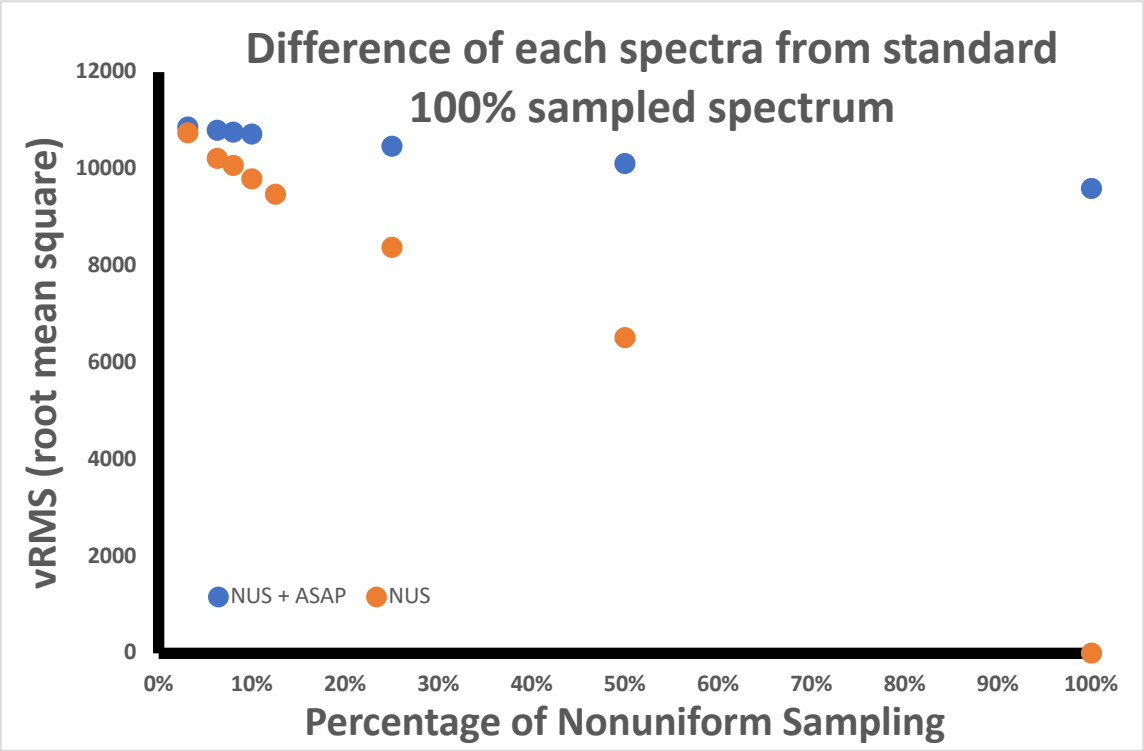


Figure 2.6: Difference between each spectra and the 100% sampled spectrum of strychnine (12 mM) shown as vRMS (root mean square). Lower RMS values indicate better reconstruction.

2.4 Conclusions

NMR is a vital tool in structure elucidation research, and thus is integral to drug development and design. Techniques that allow for rapid data accumulation can shorten the length of the experiment, saving critical time and money. Combination of techniques such as Nonuniform Sampling (NUS) and ASAP (acceleration by sharing adjacent polarization) can significantly decrease the length of an experiment. This is important for using other automatic structure annotation tools such as SMART[11], which requires HSQC data. Acquiring high quality NMR data quickly is also beneficial for dereplication of known compounds as well as structure elucidation of new ones.

2.5 Experimental Section

2.5.1 Sample Preparation

The samples of strychnine were prepared to 12 mM, 1 mM, and 0.5 mM concentrations in deuterated chloroform. The smaller concentrations required careful preparation to exclude water. The NMR tube and syringes were lyophilized overnight. The sample of fatuamide was prepared at a concentration of 2.2 mM in deuterated methanol.

2.5.2 Instrumentation and Experiments

The spectra were either recorded on a 1.7 mm dual tune TCI cryoprobe on a Bruker AVANCE III 600 MHz with standard Bruker pulse sequences at 298 ^{o}K or a 5 mm inverse detection triple resonance cryoprobe with z-gradients on a Bruker AVANCE III 600 MHz. HSQC experiments were obtained with the pulse sequence hsqcedetgpsisp2.3.nus.bmd for NUS experiments and hsqcedetgpsisp2.3.asap.nus.bmd for NUS/ASAP experiments. The program hmsIST was used to generate Poisson-gap schedules for the NUS experiments.[8]

2.5.3 Processing programs

The tool NMRbox was used to process all of the NMR data. The program nus2pipe was used to process the NMR Bruker data conversion script. NMR Pipe and SMILE were used for iterative soft thresholding. MDD (multi-dimensional decomposition method) was used for compressed sensing. The conversion script was edited in the fidSP.com. The -yN and the -yT were set to the length of the vclist. -xP0 was set to the zero order phase correction and -xP1 was set to the first order phase correction and the command nmrPipe -fn POLY -auto -xn 5.0ppm -ord 1 was changed to nmrPipe -fn POLY -auto -ord. A recFT.com was created for MDD processing. A proc.SH was created to include the number of lines in the vclist and the number of points in full sampling. A file nls.in was created with the following parameters (NDIM 2, SPARSE y seed 431 sptype poisson gap, MINHOLE 0, f180 nn, CT_SP nn, CEXP yn, NIMIN 0

0, SW 24145.536 7211.538, T2 0.02 1, Jsp 0 0. The files were made executable and run. In the ASAP experiments, Linear prediction was used (nmrPipe -fn LP -f -ord 9 -pred 3392).

2.5.4 Quantification

The quantification of the NMR data was achieved by subtracting each spectrum from the 100% sampled spectrum in nmrbox. A new spectrum was generated from this subtraction protocol and the vRMS value was taken from the statistics generated and plotted for each level of sampling.

2.6 Acknowledgements

I acknowledge Brendan Duggan for his help and expertise with the NMR instrumentation.

Chapter 2, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Duggan, Brendan M.; Gerwick, William H. The dissertation author was the primary investigator and author of this material.

2.7 References

(1) Kazimierczuk, K.; Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chemie - Int. Ed.* **2011**, *50* (24), 5556–5559. https://doi.org/10.1002/anie.201100370.

(2) Sidebottom, P. J. A New Approach to the Optimisation of Non-Uniform Sampling Schedules for Use in the Rapid Acquisition of 2D NMR Spectra of Small Molecules. *Magn. Reson. Chem.* **2016**, No. April, 689–694. https://doi.org/10.1002/mrc.4444.

(3) Mobil, Mehdi; Hoch, J. Nonuniform Sampling and Non-Fourier Signal Processing Methods in Multidimensional NMR Mehdi. *Prog Nucl Magn Reson Spectrosc.* **2014**, *83*, 21–41. https://doi.org/10.1016/j.pnmrs.2014.09.002.Nonuniform.

(4) Hyberts, S. G.; Takeuchi, K.; Wagner, G. Poisson-Gap Sampling and Forward Maximum Entropy Reconstruction for Enhancing the Resolution and Sensitivity of Protein NMR Data. *J. Am. Chem. Soc.* **2010**, *132* (7), 2145–2147. https://doi.org/10.1021/ja908004w.

(5) Ndukwe, I. E.; Shchukina, A.; Kazimierczuk, K.; Butts, C. P. Rapid and Safe ASAP Acquisition with EXACT NMR. *Chem. Commun.* **2016**, *52* (86), 12769–12772. https://doi.org/10.1039/c6cc07140f.

(6) Schulze-Sünninghausen, D.; Becker, J.; Luy, B. Rapid Heteronuclear Single Quantum Correlation NMR Spectra at Natural Abundance. *J. Am. Chem. Soc.* **2014**, *136* (4), 1242–1245. https://doi.org/10.1021/ja411588d.

(7) Kupce, Eriks, Freeman, R. Fast Multidimensional NMR by Polarization Sharing. *J. Med. Pharm. Chem.* **2001**, *2* (1), 1941–1944. https://doi.org/10.1002/mrc.

(8) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52* (4), 315–327. https://doi.org/10.1007/s10858-012-9611-z.

(9) Gołowicz, D.; Kasprzak, P.; Orekhov, V.; Kazimierczuk, K. Fast Time-Resolved NMR with Non-Uniform Sampling. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 40–55. https://doi.org/10.1016/j.pnmrs.2019.09.003.

(10) Becker, J.; Luy, B. CLIP-ASAP-HSQC for Fast and Accurate Extraction of One-Bond Couplings from Isotropic and Partially Aligned Molecules. *Magn. Reson. Chem.* **2015**, *53* (11), 878–885. https://doi.org/10.1002/mrc.4276.

(11) Reher, R.; Kim, H. W.; Zhang, C.; Mao, H. H.; Wang, M.; Nothias, L. F.; Caraballo-Rodriguez, A. M.; Glukhov, E.; Teke, B.; Leao, T.; Alexander, K. L.; Duggan, B. M.; Van Everbroeck, E. L.; Dorrestein, P. C.; Cottrell, G. W.; Gerwick, W. H. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *J. Am. Chem. Soc.* **2020**, *142* (9), 4114–4120. https://doi.org/10.1021/jacs.9b13786.

Chapter Three: Characterization of Fatuamide A, a Novel Cyanobacterial Ionophore

3.1 Abstract:

A novel compound, fatuamide A, was discovered from the marine cyanobacterium *Leptolyngbya* sp. that was cultured in the laboratory from a collection made in Faga'itua Bay, American Samoa. A bioassay-guided approach with NCI lung H460 cells was utilized to guide the isolation of fatuamide A, which was obtained from the most cytotoxic fraction. Fatuamide A has a unique structure based on a hybrid PKS/NRPS pathway. The DNA of the cultured cyanobacterium was sequenced and a putative biosynthetic gene cluster responsible for the production of fatuamide A was identified. This biosynthetic gene cluster had elements that suggested the potential of fatuamide A to bind to metals, and this property was later shown by LC-MS/MS analysis. Fatuamide A is a novel ionophore and therefore has potential medicinal and environmental applications.

3.2 Introduction:

Secondary metabolites from marine cyanobacteria are a rich source of diverse natural products, many of which have therapeutic potential.[1] The second most major cause of death worldwide is cancer. Lung cancer is the most common cancer and has the highest number of cancer deaths worldwide.[2] Natural products and their derivatives make up almost 65% of the currently employed anticancer drugs.[3] Genomic analysis has shown that marine cyanobacteria are an underexplored area for natural product therapeutics.[4] Members of the genus of marine cyanobacteria*, Leptolygnbya* sp., have been a rich source of novel secondary metabolites including coibamide A, honaucin a, and phormidolide.[5] Secondary metabolites that have been isolated from *Leptolygnbya sp* have been shown to possess different biological activities

including antibacterial, cytotoxicity to cancer cells, and anti-inflammatory activity.[5] American Samoa is a potential site for the discovery of new and chemically diverse natural products as it has been little explored for natural products from its rich marine biota. A collection was made of a live *Leptolygnbya sp*. in 2014 and brought to the laboratory and cultured. Culturing cyanobacteria is a technique that is used to obtain larger quantities of biomass than is available from field collections. Large scale production of algal cultures is a technique that has been used in industrial settings, and can be time and space intensive.[6] Fatuamide A was produced in relatively low quantities, so a scale up of *Leptolygnbya* culture in large carboys was employed.

Siderophores are naturally produced compounds that are used by an organism to acquire metals from its surrounding environment. Such metal complexes can have unique applications in medicine. For example, siderophores have been used to chelate and thus remove more toxic metals such as aluminum and vanadium, as well as for the treatment of different diseases that result in an excess of iron.[7] Siderophores can also be utilized as a drug delivery system in a siderophore-antibiotic conjugate.[8] Some siderophores have the ability to bind metals such as copper and zinc with higher affinity than to iron. For example, methanobactin is an ionophore that has higher selectivity for binding to Cu^{2+} due to its heterocyclic rings.[9]

Most of the reported cyanobacterial siderophores have been obtained from terrestrial sources.[10] The types of siderophores reported to date from cyanobacteria include 19 hydroxamates, 6 catecholates, and another 5 that are uncategorized.[10] Three cyanobacterial siderophores have been more fully characterized are synechobactin, schizokinen and anachelin.[10] Schizokinen, a hydroxamate siderophore, was first reported from the freshwater cyanobacterium *Anabaena* sp.[11] Synechobactin A-C are amphiphilic hydroxamate siderophore that were obtained from a marine *Synechococcus sp.*[12] Anachelin is a catechol siderophore that was obtained from

the freshwater species *Anabaena cylindrica.*[13] Fatuamide A has also shown ability to bind metals, and is a new example of marine cyanobacteria phenolate siderophore (*figure 3.1*).

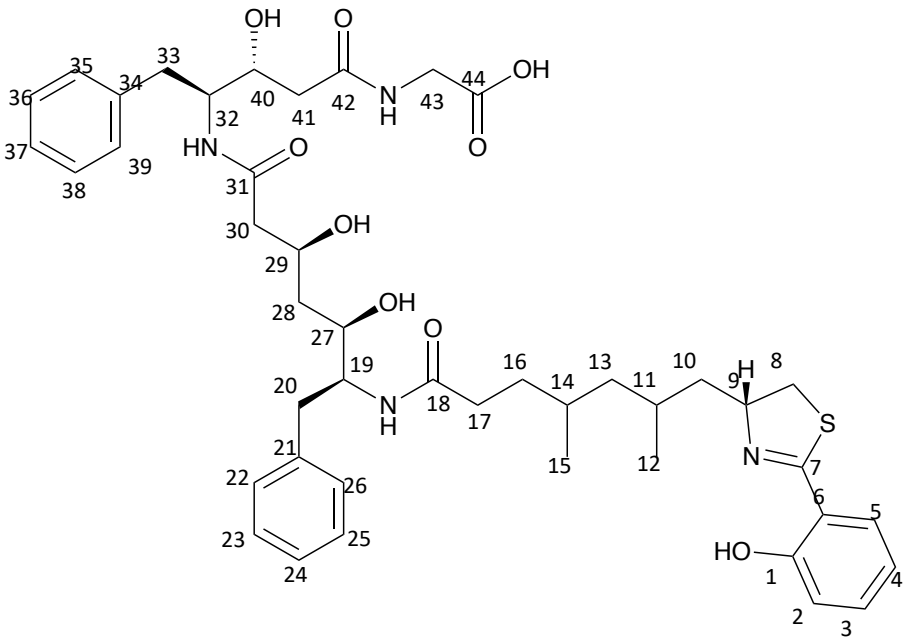


Figure 3.1: Structure of fatuamide A.

3.3 Results and Discussion

3.3.1 Collection, Culture, and Scale Up

The sample ASX22JUL14-2 was collected in American Samoa, isolated as a pure strain, and then maintained in culture in the laboratory in SWBG11 media (*figure 3.2a*). The sample was identified as *Leptolygnbya* sp. on the basis of morphology and genetic analysis *(figure 3.2b).*[14] *Leptolyngbya sp.* have been reported to grow slowly[5], and this was observed with the *Leptolyngbya sp.* ASX22JUL14-2 culture. To acquire more material, a number of large carboys were used and aerated using aquarium pumps (*figure 3.2c*). Each glass carboy contained 10-13L of SWBG11 media, aerated, and grown for several months before harvest.

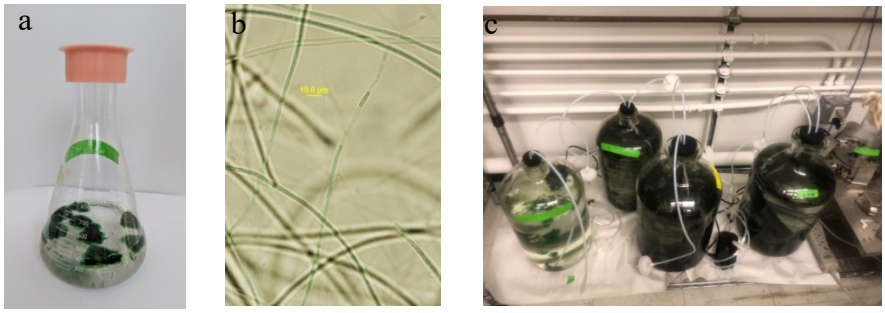


Figure 3.2: a. Culture of *Leptolyngbya* sp. ASX22JUL14-2 **b**. Microscopy of cultured ASX22JUL14-2 **c**. Set up for culture scale up of ASX22JUL14-2.

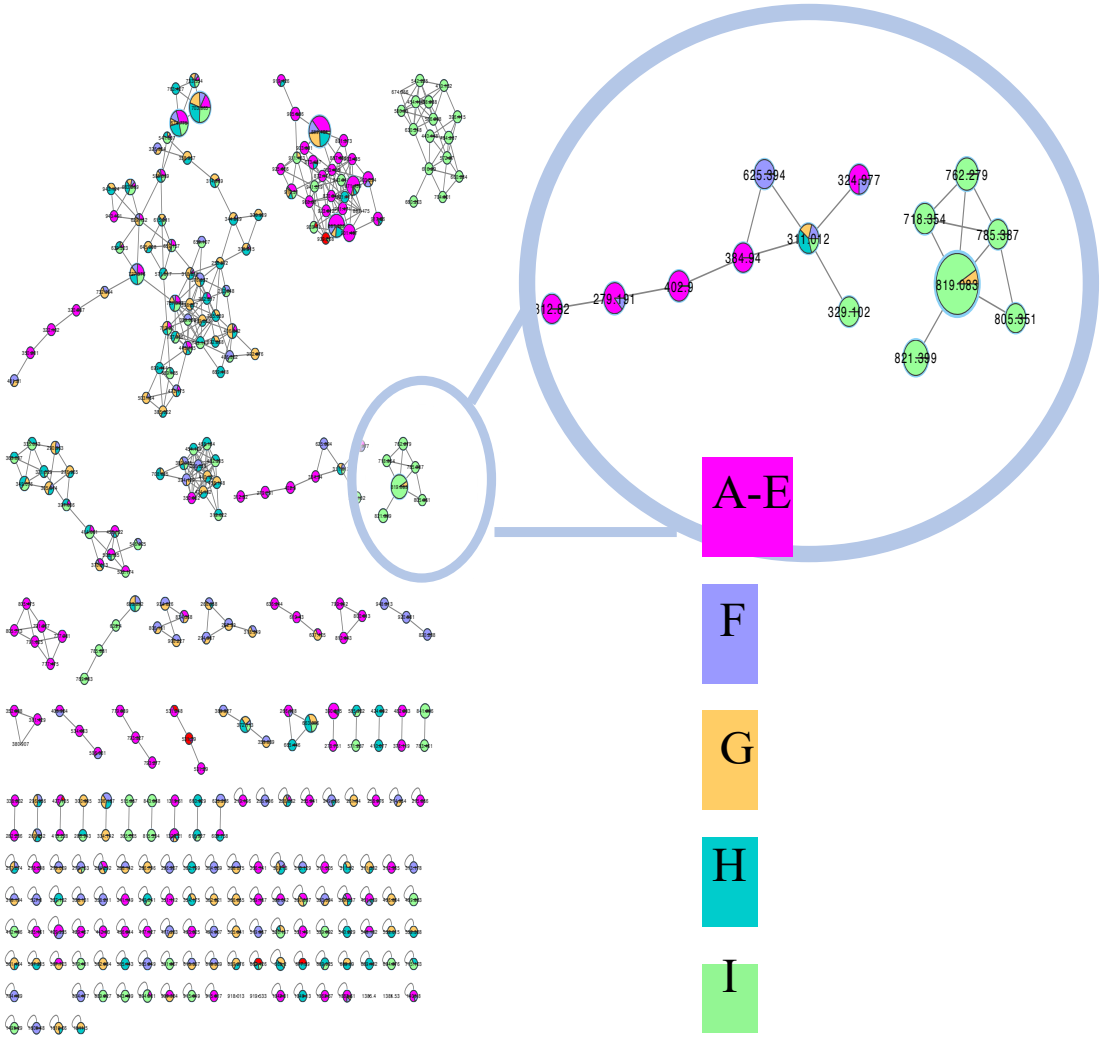


Figure 3.3: MS/MS-Based Molecular Network of ASX22JUL14-2 (using GNPS). Blue circle shows cluster from the most cytotoxic fraction I.

3.3.2 Extraction and Vacuum Liquid Chromatography

The sample was extracted after multiple harvests and then extracted with CH_2Cl_2:MeOH (2:1). Following extraction, it was fractionated into nine sub-fractions by vacuum liquid chromatography. The eluted material in fraction I (the most polar fraction, 100% MeOH) on normal phase silica column was found to cause 74.8% growth inhibition to NCI H460 lung cancer cells *in vitro* at 10 µg/mL. An LC-MS/MS based molecular network (figure 3.3) was made from this extract to further probe and annotate the compounds present in this sample. The molecular network contained a cluster comprised totally of compounds from the cytotoxic fraction I. This cluster possessed a compound with an *m/z* 819 that was subsequently targeted for isolation.

3.3.3 Cytotoxicity and Anti-inflammatory Assays

The chromatography fractions were evaluated for cytotoxicity to H460 human lung carcinoma cells (*figure 3.4*). The I fraction showed moderate cytotoxicity at the 10 µg/mL concentration, and therefore, this fraction was further pursued. The fractions were also evaluated for anti-inflammatory activity (*figure 3.5*) at concentrations of 30 and 10 ug/mL. Fraction I only showed a slight anti-inflammatory effect at 30 µg/mL. Fatuamide A was initially isolated from a fraction that was cytotoxic to lung cancer cells. Therefore, the pure compound fatuamide A was tested against lung cancer H460 cells. However, at the highest concentration tested, 60µM, it did not show any cytotoxicity. Another component, or mixture of components, must have been responsible for the activity associated with the parent fraction.

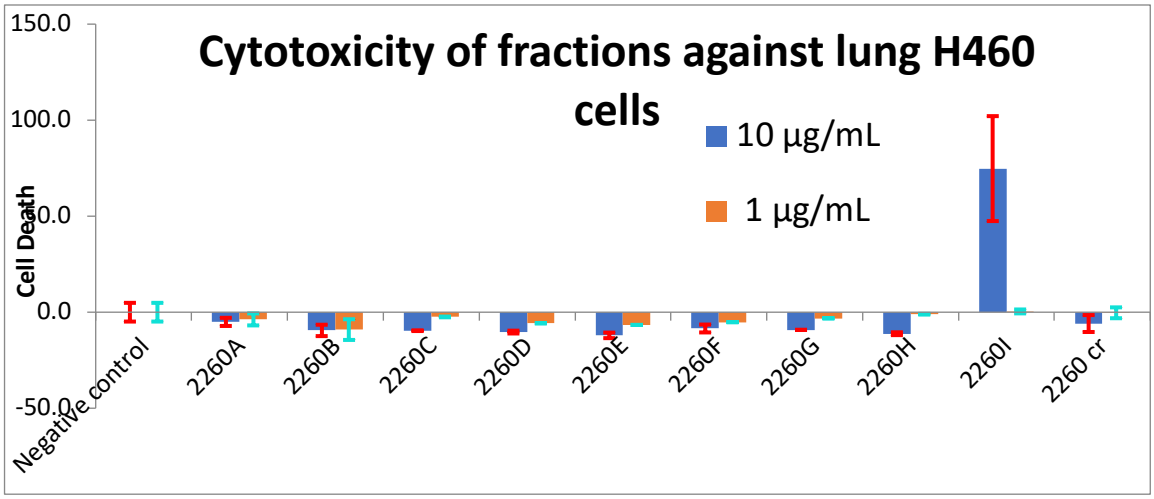
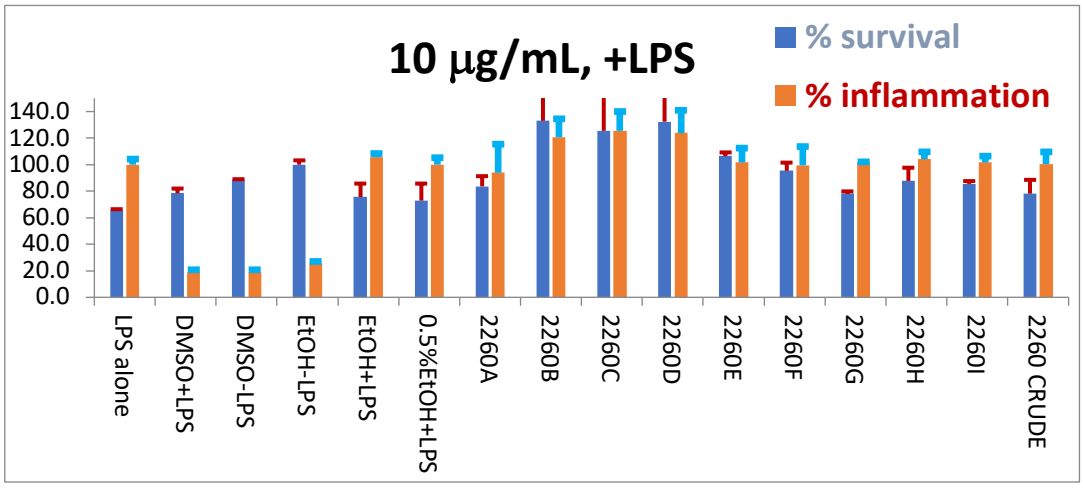


Figure 3.4: Fractions were tested against NCI H460 human lung carcinoma cells.



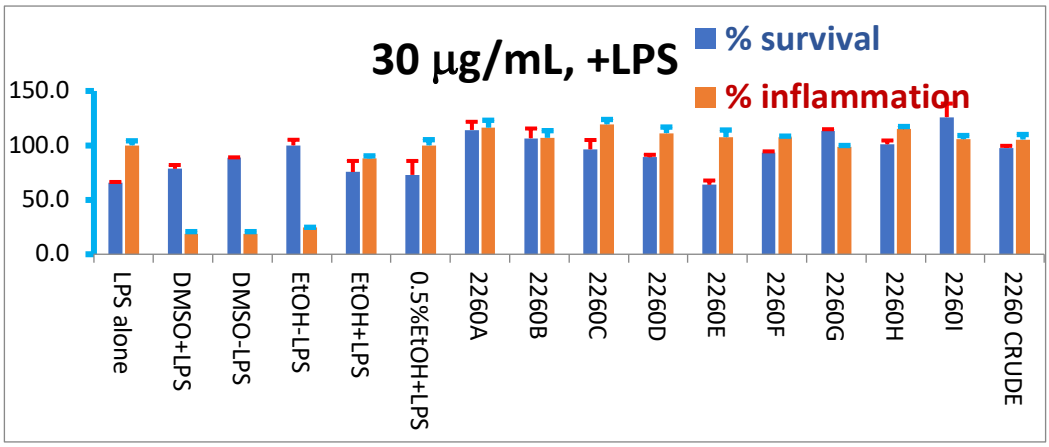


Figure 3.5: Fractions were tested for anti-inflammation activity at a. 10µg/mL and b. 30µg/mL.

3.3.4 Structure elucidation

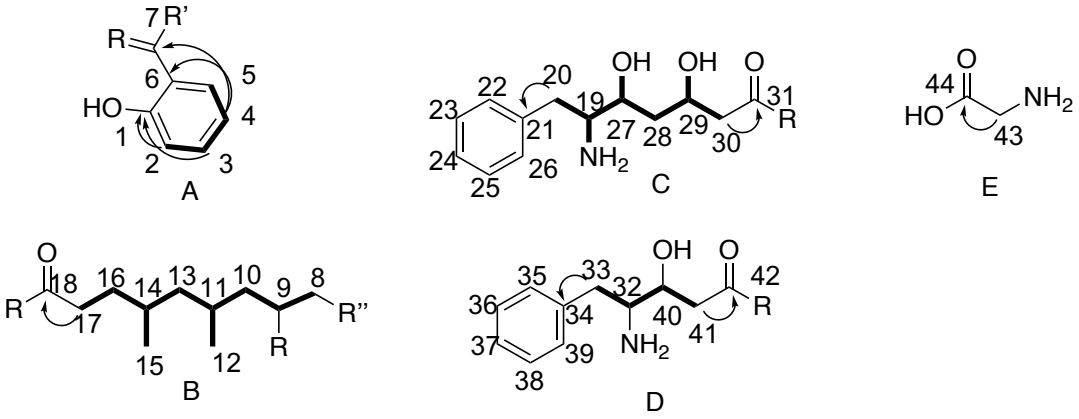


Figure 3.6: Partial structures determined by NMR for fatuamide A.

Fatuamide A was isolated as a white solid and a molecular formula of $C_{44}H_{58}N_4O_9S$ was deduced from observation of the $[M + H]^+$ ion at *m/z* 819.4000 by HR-ESI-TOFMS (*figure 3.18*). This molecular formula has 18 degrees of unsaturation. There were 4 ester/amide-type carbonyls at δ_C 173-176. The ^1H NMR spectrum had a methylene group at δ_H 3.02 and 3.52 which were associated with δ_C 37.58 by HSQC, and could be assigned to the methylene group of a thiazoline ring, which was confirmed with HMBC correlations. This accounted for 2 additional degrees of unsaturation. One phenol ring accounted for 4 degrees of unsaturation. This was shown through the four protons δ_H 6.89, 6.93,7.35, and 7.42. Finally there were two phenyl rings accounting for the final 8 degrees of unsaturation. This was apparent through 10 protons at δ_H 7.15-7.22. The inclusion of a sulfur atom in the molecular formula was first predicted by the molecular formula from the high-resolution spectrometry. The formula $C_{44}H_{58}N_4O_9S$ had only a 0.3 ppm difference to the measured molecular ion. This formula was also predicted using SIRIUS 4.0 software, which uses isotope pattern analysis(*figure 3.7*).[15] Additionally, the IR spectra (*figure 3.31*) of fatuamide showed a band at 701.49 cm^{-1}; it is reported that a thioether or thiol has a C-S stretch of 710-685 cm^{-1}.[16]

Figure 3.7: Analysis of the MS/MS spectrum of fatuamide A by SIRIUS 4.0 software. The blue highlighted formula was determined to be the molecular formula.

Table 3.1: ^{1}H and ^{13}C NMR spectroscopic data of fatumide A in MeOH-d_4.

Position	δ_C, type	δ_H (J in Hz)	Selected HMBC (H-> C)
1	160.2, C	--	
2	117.8, CH	6.93, m	1,2,4,6
3	134.0, CH	7.35, td (8.6, 8.1,1.5)	1,5
4	120.0, CH	6.89, m	2,6
5	131.6, CH	7.42, dd (7.91, 1.47)	1,3,7
6	117.6, C	--	
8a	37.58, CH_2	3.02, m	9,**10**
8b		3.52, dd (10.9,8.0)	10
7	171.9, C	--	
9	75.5, CH	4.76, p (8.1, 8.0)	7
10a	44.3, CH_2	1.59, m	9,10,11,12,13a
10b		1.69, m	8a, 9,10,11,12,13a
11	29.3, CH	1.78, ddt (11.3,9.4,5.8)	
12	20.3, CH_3	0.94, d (6.6)	10,11,13
13a	45.1, CH_2	1.11, ddd (13.6,9.3,4.5)	14,17,18,
13b		1.25, m	
14	31.1, CH	1.47, m	
15	19.4, CH_3	0.84, d (6.3)	13,14,16,17
16	35.0, CH_2	2.07, t (7.88)	
17a	35.0, CH_2	1.20, m	14
17b		1.41, m	
18	176.2, C	--	
19	56.9, CH	4.01, ddt (15.87,6.33,3.46)	
20a	36.91, CH_2	2.61, m	21,22,19
20b		3.05, m	
21	139.9, C	--	
22	130.4, CH	7.22[a], m	
23	129.3, CH	7.22[a], m	
24	127.3, CH	7.15[a], m	22
25	129.2, CH	7.22[a], m	
26	130.4, CH	7.22[a], m	
27	71.3, CH	3.75, m	
28a	41.8, CH_2	1.44, m	
28b		1.56, m	
29	66.7, CH	4.13, m	
30a	45.3, CH_2	2.18, dd (14.2,4.2)	31
30b		2.26, dd (14.2,8.5)	29,31
31	173.9, C	--	
32	56.5, CH	4.11, m	
33a	37.28, CH_2	3.13, dd (13.9, 3.6)	
33b		2.65, m	35,32
34	140.3, C	--	
35	130.4, CH	7.22[a], m	
36	130.4, CH	7.22[a], m	
37	129.3, CH	7.22[a], m	
38	127.2, CH	7.15[a], m	36
39	129.2, CH	7.22[a], m	
40	71.9, CH	3.95	
41a	41.4, CH_2	2.40, dd (14.7, 9.18)	42
41b		2.58, m	42
42	174.5, C	--	
43a	42.2, CH_2	3.90, s	
43b		3.94 d (4.75)	44
44	173.5, C	--	

[a]Signals may be interchanged due to overlap

Table 3.1 provides a tabulation of the shifts from the ^{1}H and ^{13}C NMR spectra. *Figure 3.8* shows the structures of fatuamide A determined from COSY, HSQC, H2BC, HSQC TOCSY and HMBC data. The final structure of fatuamide (*figure 3.1*) was determined through a combination of NMR, mass spectrometry analysis, and a detailed investigation of the biosynthetic gene cluster. The partial structures A-D were determined through COSY, HSQC, and HMBC (*figure 3.6*). Overlapped signals were distinguished using H2BC and band selective HSQC experiments.

From the aromatic region, it was evident that there were 3 different phenyl rings, two of which corresponded to two phenylalanine derived subunits and one from a salicylic acid derived unit. Starting with the 4 aromatic protons, δ_H 6.89, 6.93,7.35, and 7.42, which were connected by COSY correlations that corresponded to the phenol ring which corresponds to C2-C5 . H2 and H3 had an HMBC correlation to the quaternary carbon C1 (δ_c 160.2) suggesting the hydroxy group at this position in the phenol ring. H4 and H2 had HMBC correlations to the quaternary carbon C6 (δ_c 117.6) completing fragment A. Fragment B had two strong three proton doublets at δ_H 0.94 and 0.84, which corresponded to two methyl groups. These two methyl groups were connected to carbon chain C8-C17 that were sequentially connected by COSY correlations. The methyl groups (C12 and C15) were attached at C11 and C14 respectively through COSY correlations giving a 1,3 dimethyl group. C9 had a shift at δ_c 75.5 suggesting a connection to a heteroatom. H17 had an HMBC correlation to the carbonyl C18, completing partial structure B. Fragment C consisted of a carbon chain connected by COSY correlations comprised of C27-C30. H27 and H29 were methine protons with shifts of δ_H 3.75 and 4.13, respectively, suggesting two hydroxy groups at those positions. C27 was connected by COSY correlations to C19 and C20. C19 had a shift of δ_c 56.9 and a methine proton, suggesting a connection to an amide at that position. H20 had HMBC correlations to the aromatic protons corresponding to one of the

phenyl rings, completing fragment C. Fragment D was comprised of a carbon chain involving C33, C32, C40, and C41. H40 was a methine proton with a shift of δ_H 3.95, suggesting a hydroxy group at that position. C32 had a shift of δ_c 56.5 and has a single attached proton by HSQC, suggesting a connection to an amide at that position. H32 had HMBC correlations to aromatic protons corresponding to the second of the two phenyl rings. H41 had an HMBC correlation to the carbonyl C42, thus completing fragment D. Finally fragment E was comprised of a methylene group at C43. H43 had an HMBC correlation to a carbonyl at δ_c 173.5 (C44). H9 from fragment B and H4 from fragment A both had an HMBC correlation to the same quaternary carbon (C7) at δ_c 171.9; this combined with the chemical shifts of C8 and C9, established that a thiazoline ring connected the A and B fragments. NMR could not connect the other fragments, and therefore, we used mass spectrometry for the connections between these partial structures (*figure 3.8*).

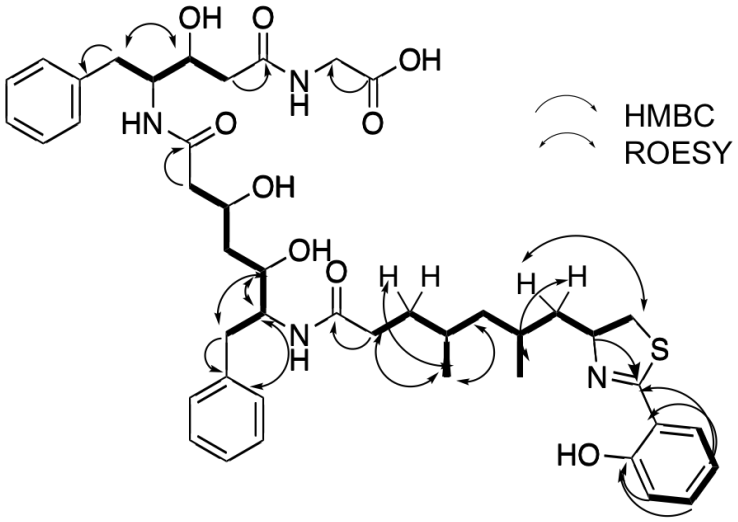


Figure 3.8: COSY (bolded lines), ROESY and HMBC correlations of fatuamide A.

A salicylate subunit connected to a thiazoline or oxazoline ring is a moiety seen in several other siderophore and ionophores, such as in yersiniabactin, pyochelin, and amychelin.[17,18,19] The thiazoline-salicylic unit in yersiniabactin has similar chemical shifts to what was observed for

fatuamide A. In the two isomers of yersiniabactin, the methylene group has shifts of δ_c 33.5 and 35.4 in DMF, while fatuamide has a shift of δ_c 37.6 MeOH-d_4. C9 of the thiazoline ring in yersiniabactin has shifts of δ_c 79.7 and 83.8 in DMF, while fatuamide has a shift of δ_c 75.5 MeOH-d_4. C7 of the thiazoline ring in yersiniabactin has δ_c 172.5 in DMF, while fatuamide has a shift of δ_c 171.9 MeOH-d_4.

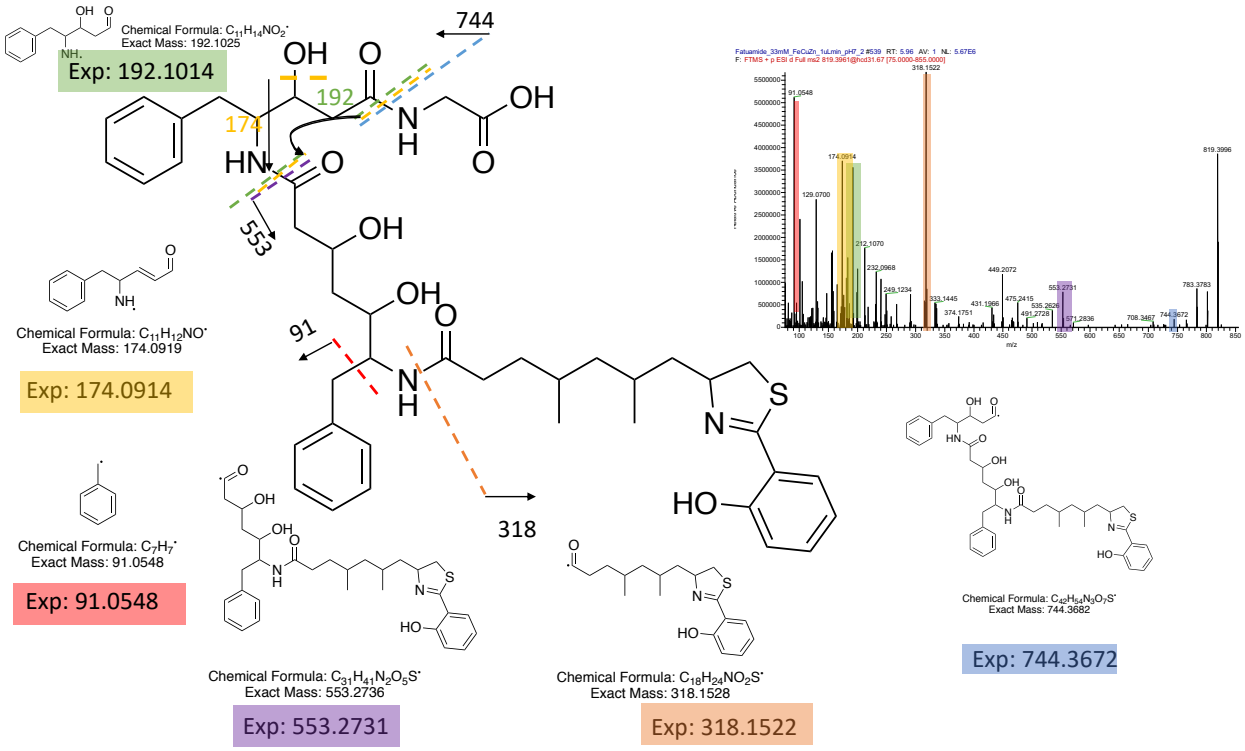


Figure 3.9: Fatuamide A fragments observed by ESI MS. Fragments are highlighted and color coded in the MS/MS spectrum, and the corresponding fragments are drawn.

The MS/MS (figure 3.9) showed a fragment at *m/z* 318.1522, which corresponds to fragments A and B. The *m/z* 553 fragment likely corresponds to the combination of fragment A,B, and C, and therefore connecting fragment B and C. The *m/z* 744 fragment corresponds to fragments A, B, C, and D, therefore connecting fragment C to fragment D. *Figure 3.10* shows the ms/ms of one of the fragments (*m/z* 553) of fatuamide A. The thiazoline salicylic acid unit fragments in the thiazoline ring, giving rise to the *m/z* 434 fragment (*figure 3.10*). This

fragmentation is consistent to what is seen for this moiety in yersiniabactin.[17] Additionally, it shows the *m/z* 318 fragment is a part of the larger *m/z* 553 fragment. Therefore, the partial structures A,B and C must be adjacent to each other.

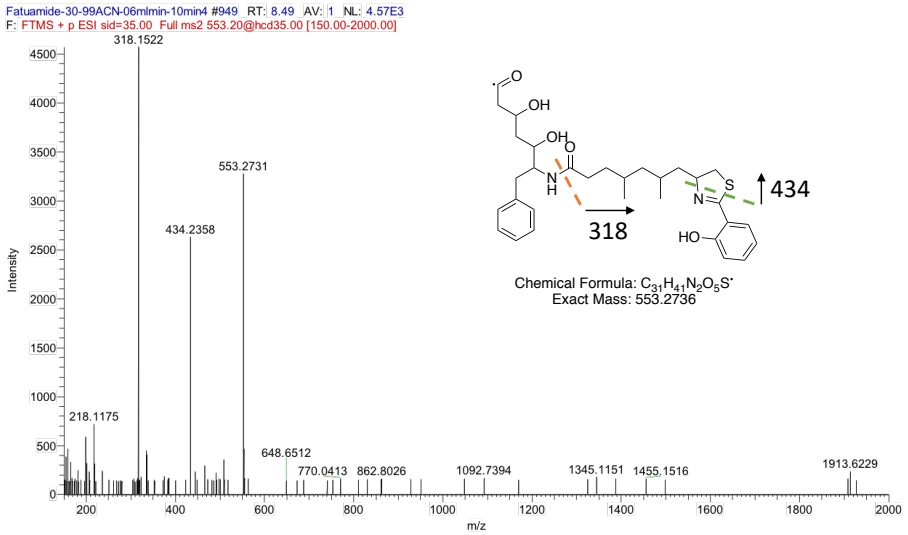


Figure 3.10: MS/MS of *m/z* 553 fragment of fatuamide A.

3.3.5 Putative Gene Cluster

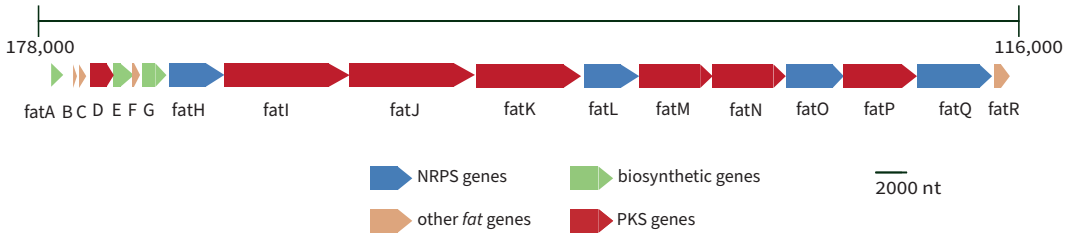


Figure 3.11: Fatuamide gene cluster adapted from AntiSMASH output, highlighting the genes important for fatuamide A biosynthesis.

The putative biosynthetic gene cluster was identified through different genetics hooks from the NMR data (*figure 3.11*). The culture that produced fatuamide was sequenced and assembled and previously published in literature.[14] The antiSMASH[20] output of the genome

showed two regions that possessed mixed PKS/NRPS biosynthetic gene clusters (BGC). The putative fatuamide BGC had adenylation domains that were specific for phenylalanine incorporation and had a cMT domain. The putative gene cluster for futamide A is 122,552nt located 72,780-195,331 nt.

FatE is an is a salicylate synthase homologue and likely catalyzes the conversion of isochorismate into salicylate, as seen in other compounds such as attinimicin and amychelin.[21,19] FatA is a thioesterase that could be responsible for release of the natural product. Blasting the thioesterase showed similarity to the gramidicin dehydrogenase LgrE (67% identity) and Gramicidin S biosynthesis protein GrsT (63% identity) and thioesterase MycT (63% identity). FatH through fatQ are made up of a mixture of NRPS and PKS genes (*table 3.2*). Fat R is a TauD/TfdA family dioxygenase. Additionally, the Fat BGC contains an MFS (major facilitator superfamily) transporter, ABC transporter binding protein, and a Ton-B dependent receptor. *Anabena sp.* also contains an MFS protein, and this is one of the proteins that is used for the secretion of the siderophore schizokinen.[10] Siderophores with bound iron are imported by Ton-B dependent transporters and the ABC transporters.[10] The presence of these various accessory genes supports the prediction that fatuamide A is a metal transporting siderophore, as discussed further below.

Table 3.2: Predicted functions of the fatuamide Biosynthetic Gene Cluster (BGC)

Protein	Size (nt)	Predicted function	GenBank Accesion number
fatA	768	thioesterase	F6J95_025050
fatB	249	acyl carrier protien	F6J95_025045
fatC	459	4'-phosphopantetheinyl transferase superfamily protein	F6J95_025040
fatD	1539	benzoate-CoA ligase family protein	F6J95_025035
fatE	1323	Salicylate Synthase	F6J95_025030
fatF	468	holo-ACP synthase	F6J95_025025
fatG	1524	3-oxoacyl-(acyl carrier protein) synthase III	F6J95_025020
fatH	3474	C (heterocyclization), A- Cys, P	F6J95_025015
fatI	7944	KS, AT, cMT, KR, DH, ER, PCP	F6J95_025010
fatJ	7944	KS, AT, cMT, KR, DH, ER, PCP	F6J95_025005
fatK	6654	KS, AT, KR, DH, ER, PCP	F6J95_025000
fatL	3447	C, A-Phe, P	F6J95_024995
fatM	4650	KS, AT, KR, PCP	F6J95_024990
fatN	4638	KS, AT, KR, PCP	F6J95_024985
fatO	3678	C, A-Phe, P	F6J95_024980
fatP	4644	KS, AT, KR, PCP	F6J95_024975
fatQ	4755	C, A-gly, P, C	F6J95_024970
fatR	1005	TauD/TfdA family dioxygenase	F6J95_024965

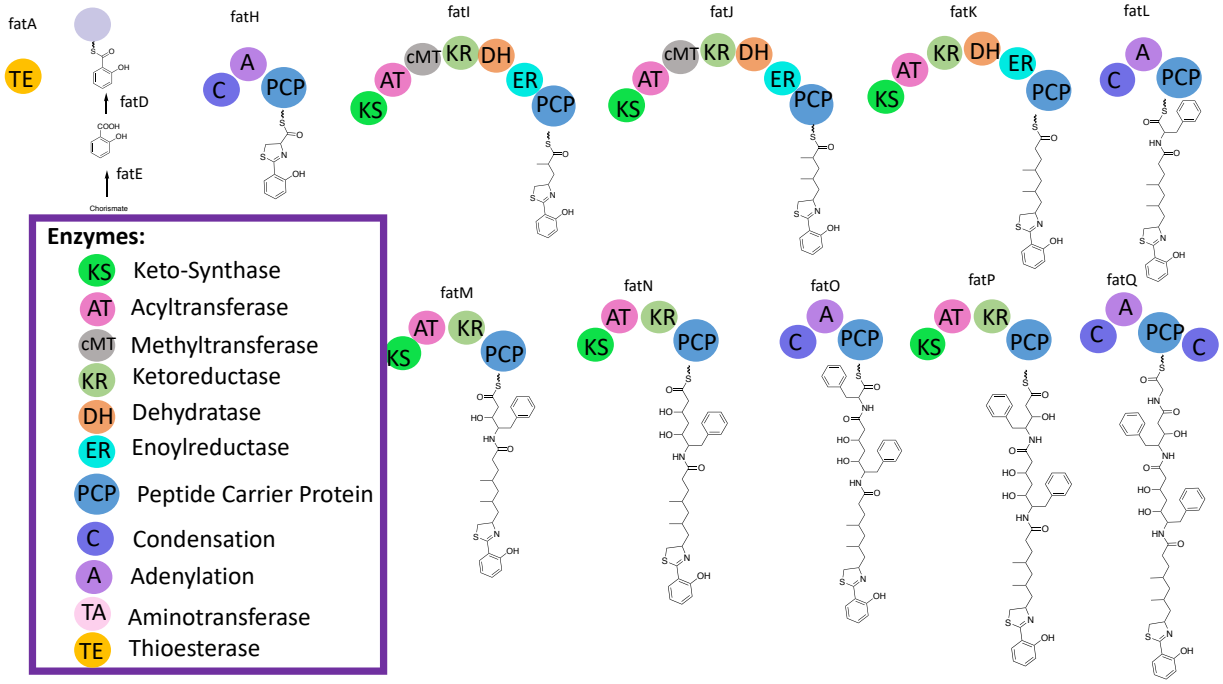


Figure 3.12: Proposed biosynthesis of fatuamide A from the putative biosynthetic gene cluster.

In the biosynthesis of amychelin, the 2-hydroxybenzyl moiety is produced by a salicylate synthase that converts chorismate into salicylate, which is incorporation into the biosynthetic

pathway by a hydroxybenzoyl AMP ligase.[19] Fat E is a salicylate synthetase and FatD shows 44.9% identity to a 4-hydroxybenzoate ligase in a Candidatus Rokubacteria bacterium. The thiazole-salicylic unit in yersinabactin is produced by a NRPS module that combines an activated salicylate on an aryl carrier protein with a cysteine unit that is attached through an adenylation domain; the latter is subsequently cyclized to a thiazoline ring.[22,23,24]. Similarly, fatH is an NRPS module that adds a cysteine through an adenylation domain that is subsequently cyclized to form a thiazoline-salicylic unit. This is followed by two PKS extensions (fatI and fatJ) in which the PKS additions are fully reduced and methylated by cMTs. Next, fatK is responsible for another PKS extension that is fully reduced. An addition of a phenylalanine by fatL is followed by two PKS extensions (fatM and fatN) that are partially reduced to secondary alcohols in each case. A second phenylalanine is added by fatO and undergoes a PKS extension that is partially reduced by fatP to a secondary alcohol. Finally, a glycine residue is added by fatQ.

A thioesterase domain, fatA, is at the upstream edge of the cluster that could be responsible for the release of fatuamide and a BLAST search that shows that the domain is an alpha/beta hydrolase with 68% identity. Due to its placement in the cluster we are classifying it at as a type II thioesterase (TE).[25] Type II thioesterases have many different functions in secondary metabolite pathways.[25] Most biosynthetic pathways have an in-line thioesterase at the terminus which is responsible for product release, and an additional TE II that is used to edit or reprime units.[25,26] However, there are few ionophores that use a TEII to release the final product, such as nanchangmycin, monensin, and nigericin/abierixin.[27] Alignment of the amino acids from the thioesterase in fatuamide A with the thioesterase from nanchagmycin, monesin, and nigericin showed 17.9%, 41.4%, and 18.0% pairwise identity, respectively. Additionally, the alignment with the thioesterase from yersiniabactin showed 29.2% identity. Alternatively, the release of

fatuamide A could be done through an additional condensation unit that is present after the addition of glycine in the biosynthetic gene cluster. Release of the natural product with a condensation domain has been seen previously with FK520, apratoxin A which had an additional condensation and adenylation domain, and aeruginoside which had an additional condensation and PCP.[28] Alignment of the amino acids from the additional condensation in fatuamide A with the additional condensation domains listed above showed 28.9% identity with Fk520, 41.7% identity with apratoxin, and 40.9% identity with aeruginoside.

3.3.5 Stereochemistry of Fatuamide A

There are eight stereocenters in fatuamide A. The first one in the chain (C9) is the thiazoline ring derived from a cysteine residue; due to the lack of an epimerase domain in the NRPS module responsible for its incorporation, the configuration is proposed as deriving from the natural L-stereoisomer. Additionally, there are two stereocenters derived from the addition of two phenylalanine residues (C19 and C32); these also are predicted to incorporate L-phenylalanine without epimerization according to the biosynthetic gene cluster, and thus to both be of *S* configuration.

The stereochemistry of the three hydroxy groups at C27, C29, and C40 were annotated by antiSMASH[20] to have L-configuration. This deduction of configuration of the hydroxy groups can be deduced from the KR domain which is responsible for their formation.[29] These three KR domains were aligned with other cyanobacterial domains to gain further insight into the likely configuration of the resulting hydroxy group (*figure 3.13*). In all cases, the KR domains lacked the characteristic LDD from the LDD loop in B-type KR domains. It has been shown that the second D of this sequence is highly conserved in B-type KRs, but this residue was absent in all three fatuamide KR domains, suggesting that all three KRs were A-type. However, we also

found that all three hydroxy-producing KR domains of the fatuamide BGC lacked the highly conserved W residue that is consistently present in A-type KR domains. Therefore, it appears that fatuamide A has a mix of A-type and B-type KR domains because they lack both the pivotal W residue and the LDD motif.

In a revised study of the stereochemistry of phormidolide, another marine cyanobacterial metabolite, it was determined that they possessed the crucial D of the LDD loop, identifying them as B-type KRs; the L-configuration was confirmed by synthesis and spectroscopic studies.[30,31] In another study that analyzed KR domains from different organisms, all that contained the second D in the LDD loop went on to form the D-product.[32] Trans-AT KR domains have been shown to produce the L or D hydroxy groups based on the presence of the D in the LDD loop.[33] Therefore, because fatuamide A lacks the diagnostic D of the LDD loop for the three hydroxy-producing KRs, we conclude that these KR domains form L-products in each case, consistent with the AntiSMASH[20] results.

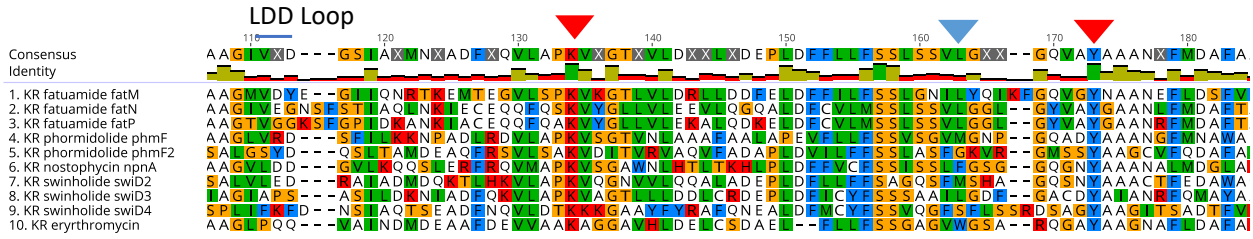


Figure 3.13: Alignment of different KR domains from cyanobacterial natural products made with *Geneious version 2019.2 created by Biomatters.* Position of the conserved W for A-type KRs denoted with a blue triangle. The LDD Loop is denoted by a blue line. Red triangles denote catalytic sites.

The stereochemistry of the methyl groups at C11 and C14 are even more difficult to characterize using bioinformatics. The KS domains were analyzed on NaPDOS2, to evaluate if there were any correlations between their sequences and the stereochemistry of the resulting methyl groups; however, this analysis was inconclusive. The cMT in cisAT pathways is not well

studied in terms of the stereochemistry of its products.[34] It was found that only 1.7% of cisAT modules contain cMT domains.[35] Nevertheless, homology was found in the sequences of cMTs that produce similar products, such as in yersiniabactin and pyochelin.[36] There have also been some recent studies on the stereochemistry of products of cMT enzymes present in transAT polyketide synthases.[37] To determine if certain sequences of cMTs are associated with the stereochemistry of the corresponding methyl groups they introduce, a dendrogram using the neighbor-joining tree building method with 1,000 bootstraps on Geneious was created. The input data were sequences of cMT from cyanobacterial compounds with known methyl group stereochemistry., Unfortunately, there was no distinguishable stereochemical outcome deduced from the clusters established in the dendrogram. Further investigation of the amino acid sequences for C-methyl transferases that produce cyanobacterial compounds with methyl groups of known stereochemistry was performed in Figure 3.14. Lack of a Y residue at the position for stereocontrol is denoted by a green triangle. This analysis tentatively suggests that both fatuamide methyl groups are 2R; however, this is quite speculative and needs experimental verification.

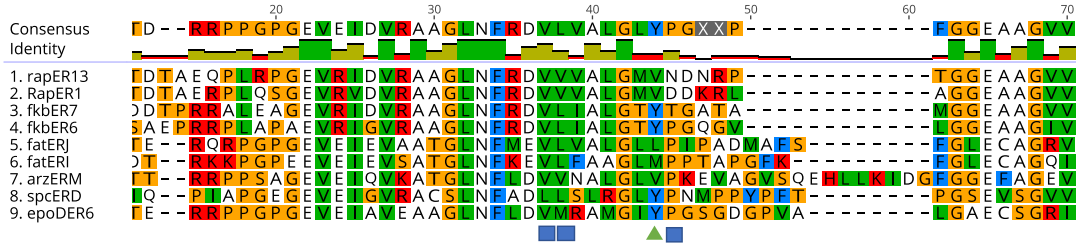


Figure 3.14: Alignment of cyanobacterial ER domains on Geneious. The green triangle represents the site for stereocontrol according to Kwan *et al*.[38] A Y residue at this position gives a 2S configuration for the resulting secondary methyl group. A V,A, or F residue at this position gives 2R configuration. The blue boxes denote additional sites that are involved in stereocontrol of the resulting methyl group.

The relative configuration of 1,3-n-methyl branched can also be determined through the chemical shift difference of the geminal methylene protons.[39] The chemical shift difference of the geminal methylene protons (H13a and H13b) on fatuamide A is 0.14. Values less than 0.1 have been shown to be *anti*-configured, and those more than 0.4 have been shown to be *syn*-configured. Intermediate values between 0.1 and 0.4 could be either *syn*- or *anti*-, and require further analysis and comparison to literature data.[39] In the case of fatuamide A, considering the nature of adjacent functionalities and literature examples, (4S,6S,8R) – 9-tert-butoxy-4,6,8-trimethyl-9 oxononanoic acid and (2S,4S)-5-acetoxy-2,4-dimethylpentanoic acid, we propose an *anti*- configuration.[40,41]

3.3.6 Fatuamide A in SMART

The NMR tool SMART (www.smart.ucsd.edu) was used to guide for dereplication and structure elucidation of fatuamide A. The SMART tool compares HSQC spectra with an extensive library of compounds, and the top 10 results (the compounds most similar) to fatuamide A are shown in *figure 3.15*. The relatively lower cosine scores reflect the novelty of fatuamide A. Additionally there are some motifs that are highlighted in the top hit compounds that are seen in fatuamide A, such as phenol groups. Fatuamide A contains a salicylic acid, consistently with this deduction of the SMART tool. Additionally, many of the top hits include phenylalanine, and as noted, fatuamide A has two PKS extended phenylalanine residues.

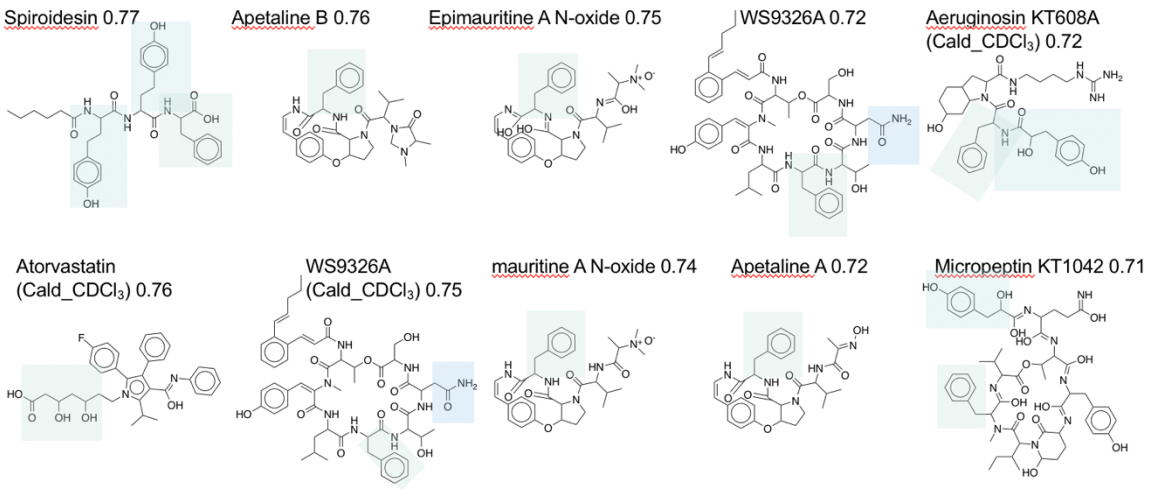


Figure 3.15: Fatuamide A top 10 SMART 2.0 results based on cosine score. The colored boxes highlight the substructures that have similar motifs to those found in fatuamide A.

3.3.7 Metal Infusion Experiment

To evaluate the metal binding capabilities of fatuamide A, the compound was injected to a native spray MS with a metal infusion post LC (*figure 3.16*). This method was established by Aron *et al.* for metabolomics studies of metal binding compounds.[42] There is a peak corresponding to an apo form of fatuamide A and Cu^{2+} bound form of fatuamide A having an *m/z* delta= 61.92. There were also other metal adducts in smaller quantities, showing that fatuamide A can bind to Zn^{2+} (*m/z* delta = 62.92077) and Fe^{3+} (*m/z* delta = 52.9115) as well.

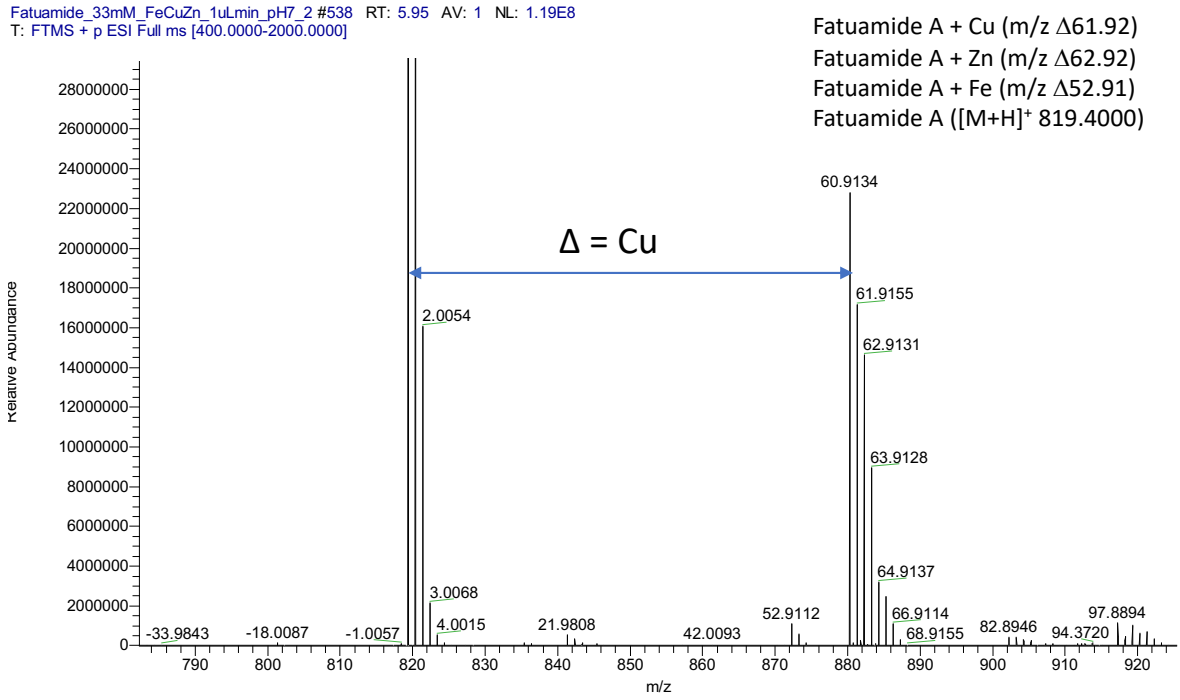


Figure 3.16: LC-MS/MS trace of native electron spray mass spectrometry with post column metal infusion. Mass shifts indicate proposed binding of Cu, Zn, and Fe to fatuamide A.

3.3.8 Siderophore CAS Assay

The production of siderophores by other bacteria have been detected in the literature through a siderophore CAS (chrome azurol S)binding assay.[10] Change in the absorption and color of solution indicates the release of the metal from the chrome azurol CAS solution and therefore the presence of a siderophore.[43] An established protocol for liquid media was followed.[44] It was found that the ASX22JUL14-2 culture that produces fatuamide A caused a change in color of the CAS solution, indicating the presence of a siderophore (figure 3.17). The absorbance was measured at 655 nm, and a decrease in absorbance was seen, indicative of the presence of a siderophore (table 3.3)

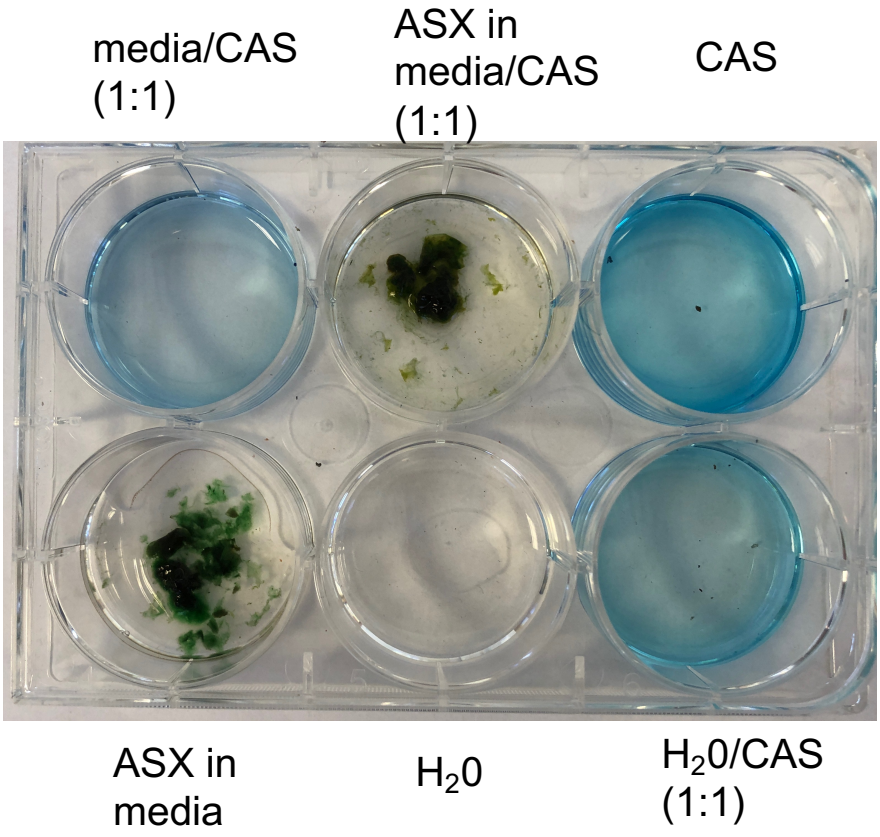


Figure 3.17: CAS (chrome azurol S) assay with fatuamide A producer ASX22JUL14-2. Loss of blue color in the ASX culture in media/CAS (1:1, top middle) indicates a positive test for siderophore activity.

3.4 Conclusions

The natural product fatuamide A was isolated from an American Samoan *Leptolyngbya sp.* The structure was determined through different NMR experiments, mass spectrometry and genomic analysis. It was initially prioritized for isolation because fractions containing this compound possessed strong cytotoxicity against NCI H460 lung cancer cells. Through analysis of the biosynthetic gene cluster and LCMS/MS studies, it is shown that fatuamide is an ionophore and has a highly unique structure. Many ionophores and siderophores have biomedical and environmental relevance in terms of removing metals from the body as well as the

environment; therefore, in this aspect further studies on fatuamide A and its biological properties are warranted.

3.5 Experimental Section

3.5.1 General Experimental Procedures

The LC-MS analysis was done with a Thermo Finnigan Surveyor HPLC System with a Thermo-Finnigan LCQ Advantage Max Mass Spectrometer with a Phenomenex Kinetex 5 μm C18 100 x 4.6mm column. A linear gradient that had a flow rate of 0.6 mL/min with the solvents A) H_2O + 0.1% Formic Acid and (B) CH_3CN + 0.1% Formic Acid. A 5 min isocratic step of 30% B was followed by an increase to 99% B over 17 min. It was held at 99% B for 5 min and then decreased to 30% B in 1 min and then held for 4 min at 30% B. The mass spectra were gathered with an ESI source (*m/z* 200-2000). HR ESI MS data was collected by the UCSD Chemistry and Biochemistry Mass Spectrometry Facility on an Agilent 6230 Accurate-Mass TOFMS in positive ion mode. The ^{13}C data was recorded at 298 ^{o}K with standard pulse sequences on a Varian Vx 500 NMR with a cold probe and z-gradients. The ROESY, HSQC-TOCSY,TOCSY, H2MBC, and band selective HMBC NMR spectra were recorded in a 1.7 mm dual tune TCI cryoprobe on a Bruker AVANCE III 600 MHz with standard Bruker pulse sequences at 298 ^{o}K. ^{1}H, HSQC, HMBC, and COSY were recorded on a JEOL 500 MHz NMR spectrometer at 298 ^{o}K. The NMR data were recorded in deuterated methanol and calibrated with residual solvent peaks (δ3.31, δ49.00). A CombiFlash Rf+ Lumen flash chromatography TELEDYNE ISCO system was used for chromatography. The IR data was collected on a ThermoScientifc Nicolet 6700 FT-IR instrument. The metal infusion method was done by methods reported in literature using a Vanquish UPLC system coupled to a Thermo Fisher Scientific Q-Exactive orbitrap mass spectrometer.[42]

3.5.2 Biological Material Collection and Identification

The sample ASX22JUL14-2 was collected in Faga'itua Bay, American Samoa on July 22nd, 2014. A chemistry collection, RNA later, and culture was collected at 1-2 meter depth. It was grown for approximately 120 days in a 16 hour light / 8 hour dark routine in SWBG11 media at 27.2-27.3°C.

3.5.3 Culture Techniques

The culture of ASX22JUL14-2 was scaled up in 13L and 9L glass carboys with a rubber stopper. For aeration, an air pump with a HEPA filter was connected to an autoclavable tube, and run through the stopper and into the SWBG11 media. A shorter second tube was in the head space inside the carboy. This shorter second tube went through the stopper and connected to a HEPA filter. The whole system (minus the air pump) was autoclaved on G20 cycle before the addition of the cyanobacteria and SWBG11 media. After autoclaving, 10L of media were added and ASX22JUL14-2 culture from a 2L culture were added in a biosafety cabinet. The cultures were harvested after 3-5 months of growth.

3.5.4 Extraction and isolation

The culture was extracted after multiple rounds of harvest in DCM:MeOH (2:1) for 30 minutes with sonication at a temp <30°C. The organic layers were combined and dried for a crude extract of 1.156g. The crude was further fractionated by Vacuum Liquid Chromatography afforded 1.011 g of fraction I (100% MeOH) which contained fatuamide A. This fraction was loaded on a Combiflash by solid phase in celite with a linear gradient with a C_{18} 5.5 g Gold column (A) H_2O and (B) MeOH and (C) ACN at a flow rate 18 ml/min and monitored at wavelengths of 214 nm and 254 nm. It was initiated with under isocratic conditions of 30% (B) and 70% (A) for 6 minutes followed by a gradient to 100% (B) until minute 21, held for 1 min, at

minute 22 50% (B) and 50% (C) to 25 minutes, held for a minute and then decreased to 20% (B) and 80% (C) and held for one minute, followed by 100% (C) until 33.8 minutes. Under these conditions, fatuamide A eluted at 16 minutes, and a total of 1 mg was obtained .

3.5.5 Cytotoxicity/ Anti-inflammatory Assay

Human lung carcinoma H460 cells were grown in a flask in monolayers to near confluence and seeded at 6.66 x 10^3 cells/ml in wells (180 μL each) containing RPMI medium with FBS. They were incubated for 24 h at 37°C and 5% CO_2. The samples were prepared dissolving in DSMO and diluted in RPMI medium so the DMSO was less than 1%. A 20 μL aliquot of these solutions were added per well and the final concentrations of the samples were 10 and 1 μg/mL; these were tested in duplicate. Plates were incubated for 48 h and then stained for 25 min with MTT. The plates were analyzed by optical density measurements at 570 nm and 630 nm. Cell survival rates were calculated by negative controls comprised of RPMI medium.

The anti-inflammatory activity was tested using RAW 264.7 ATCC murine macrophage cells in Dulbecco's Modified Eagle Medium (DMEM) with 10% endotoxin-low fetal bovin serium at 37°C with 5% CO_2. The cells were seeded in a 96-well plate (5x104 cells/well) in triplicate and were subjected for 24 h with 3 μg/mL of E. coli lipopolysaccharide (LPS), with the tested fractions at concentrations of 10 and 30 μg/mL. The accumulation of NO in the supernatant of the cell cultures was analyzed by the quantification of nitrite by the Griess reaction. A 50 μL aliquot of each supernatant was added to 96-well plates with 50 μL of 0.1 % sulfanilamide and 50 μL 0.1% N-(1-napthyl)-ethylenediamine (NED). Absorbance analysis of the mixtures was quantitatively measured at 570 nm to calculate the release of NO based on a standard curve from a nitrate standard in DMEM (0-100 μM). LPS was used as the negative

control as an indication of 0% anti-inflammation and 1% DMSO was used as the positive anti-inflammation control.

3.5.6 Bioinformatics

The DNA was extracted and reported in "A Multi-Omics Characterization of the Natural Product Potential of Tropical Filamentous Marine Cyanobacteria" and can be found on NCBI with the accession JAAHFU000000000.[14] Alignments of the KR domains were done with *Geneious version 2019.2 created by Biomatters* with a Global Alignment (Needleman-Wunsch).

3.5.7 CAS chrome azurol Assay

A chrome azural S assay was used that was modified for seawater samples.[44,45] A solution that contained fatuamide was mixed with the CAS assay solution in equal quantities. The CAS assay solution consisted of CAS (chrome azural S) ($2x10^{-4}$ M), $FeCl_3$ ($2x10^{-5}$M), HDTMA (hexadecyltrimethylammonium bromide) ($1.6x10^{-1}$ M), and perazindiethanesulfonic acid (PIPES) at pH 5.8 ($1.1x10^{-1}$M). After 50 hours of incubation the absorbance was measured at 655 nm. The concentrations of Fe(III) complexes were calculated from the absorbance. One negative control was the solution containing the ASX22JUL14-2 culture in its media without CAS reagent. A second negative control was the CAS assay solution by itself (no cultures added). A third negative control was comprised of CAS assay solution in pure H_20 (1:1). Additionally, a solution of media, containing only 50% of the metal concentration compared to standard media preparation, that had previously had ASX22JUL14-2 growing it for three weeks and was then removed by filtration, and then was tested with the CAS solution (1:1). The lack of reaction may have been due to all of the fatuamide siderophore being complexed during the one day between filtration and CAS assay. The final assay condition was comprised of ASX22JUL14-2 in full strength metal mix media and directly assayed with the CAS reagent, a

condition which gave a positive reaction (See *figure 3.17*). All samples were analyzed for absorption at 655 nm.

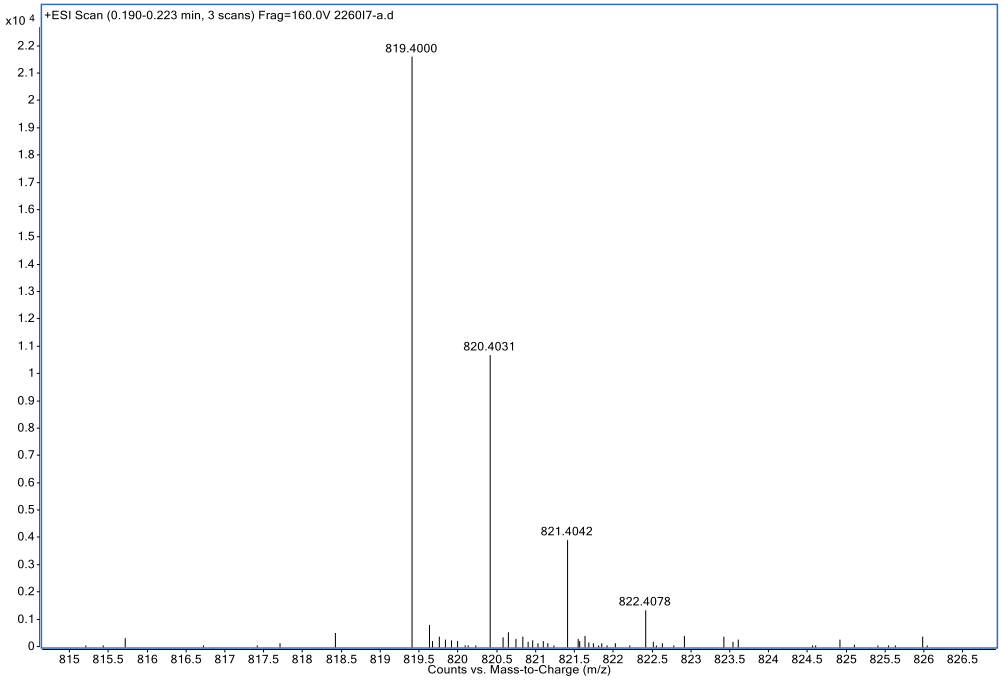


Figure 3.18: HR-ESI-TOFMS of fatuamide A $[M+H]^+$ 819.4000.

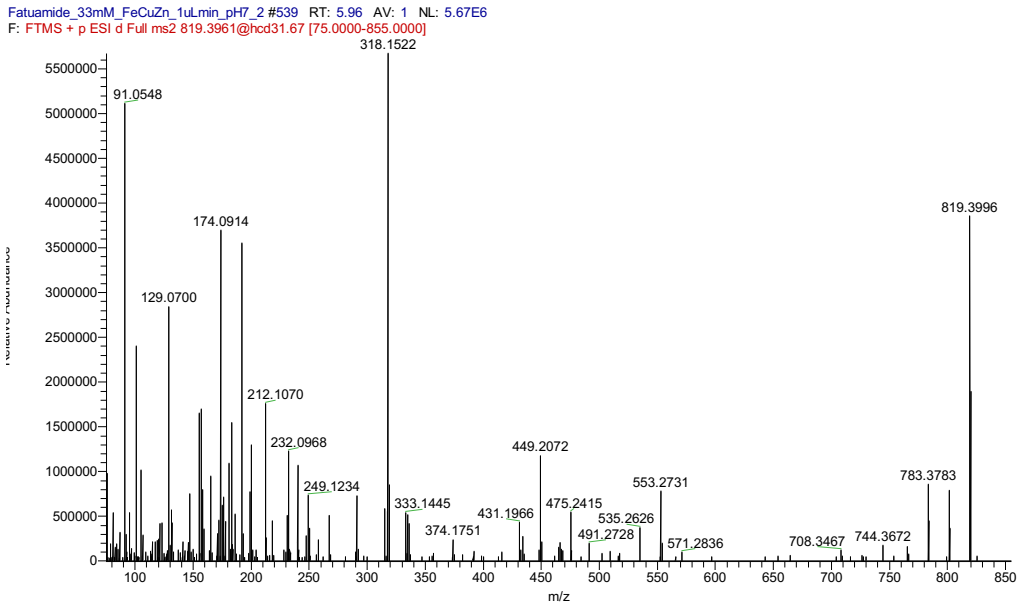


Figure 3.19: HR-ESI-TOFMS MS/MS of fatuamide A.

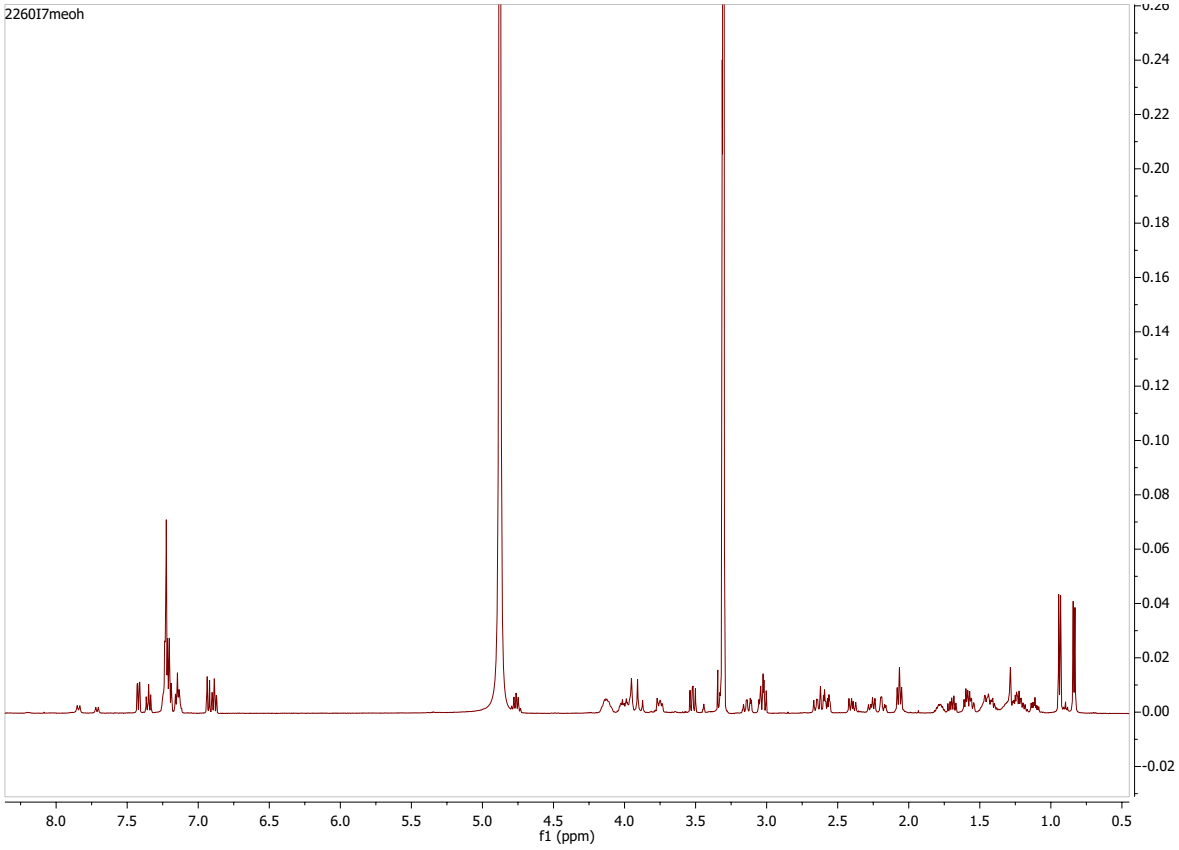


Figure 3.20: ^{1}H NMR spectrum of fatuamide A in MeOH- d_4 on 500 MHz JEOL.

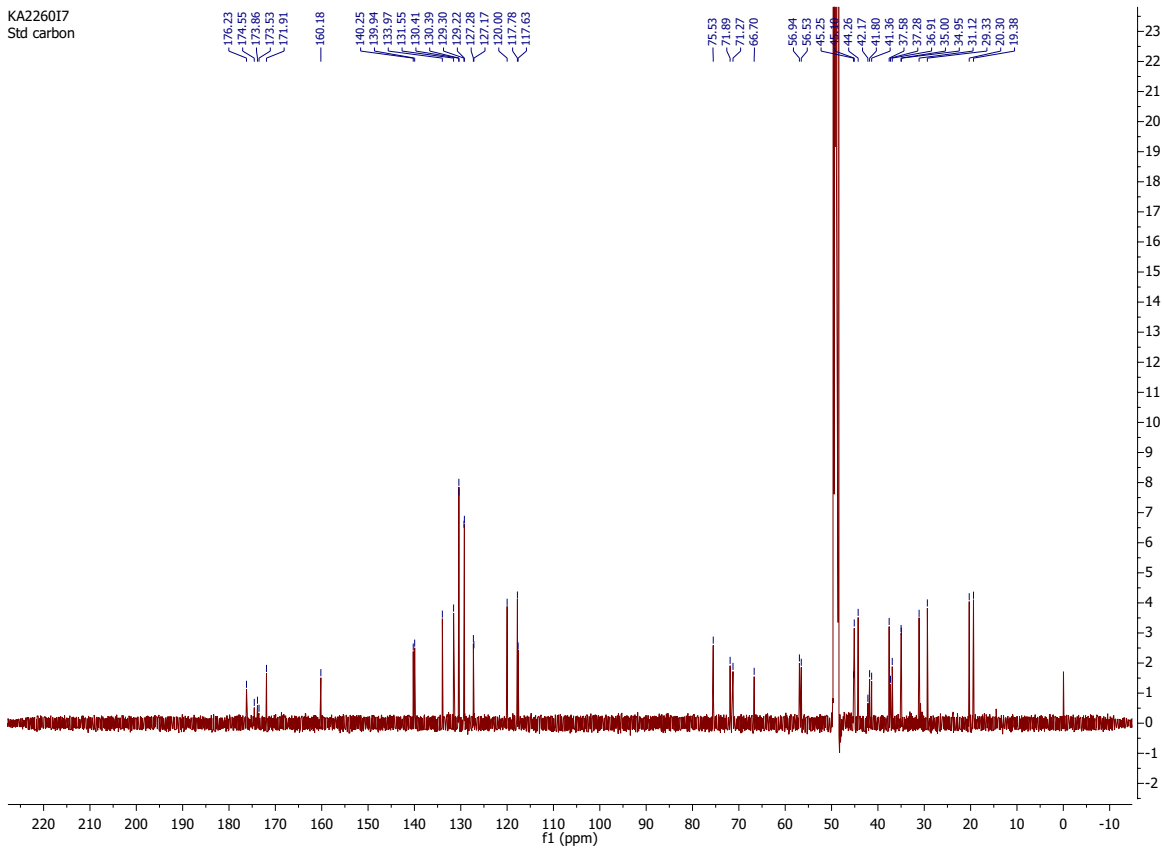


Figure 3.21: ^{13}C spectrum of fatuamide A in MeOH-d_4 on Varian Vx 500 MHz NMR

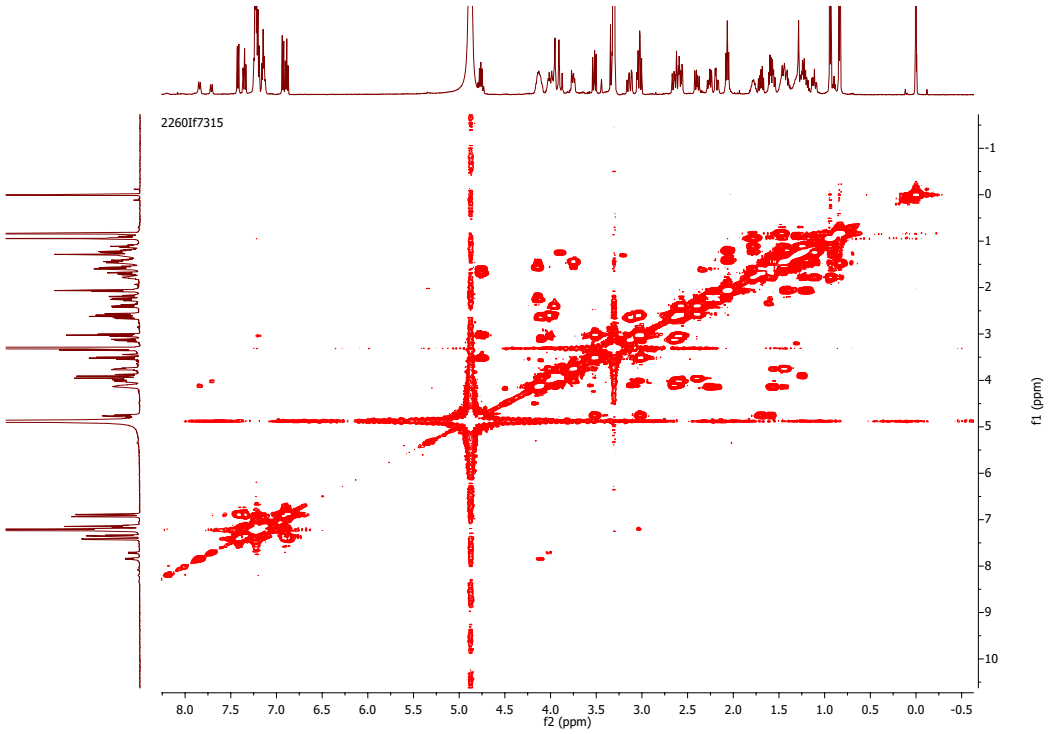


Figure 3.22: ^{1}H-^{1}H COSY NMR spectra of fatuamide A in MeOH-d_4 (500 MHz).

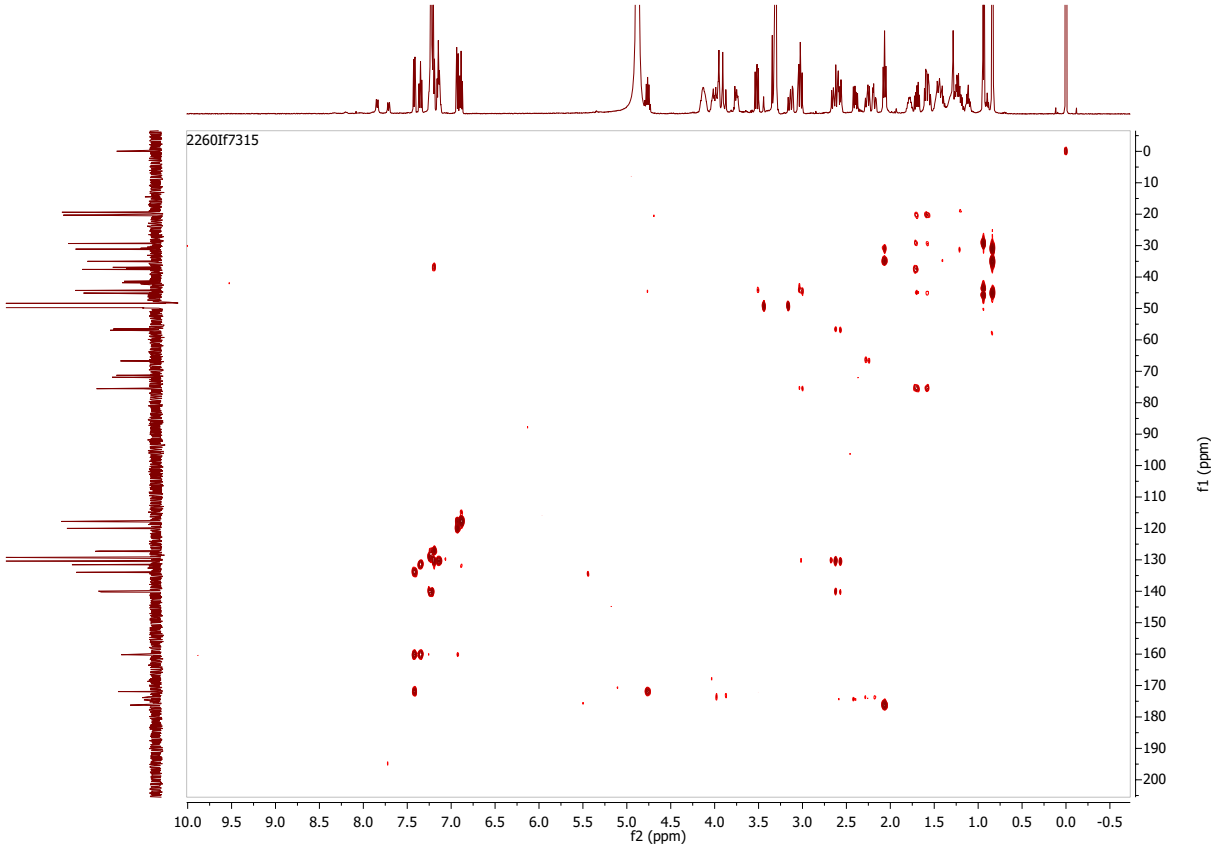


Figure 3.23: ^{1}H- ^{13}C HMBC spectra of fatuamide A in MeOH-d_4 (500 MHz).

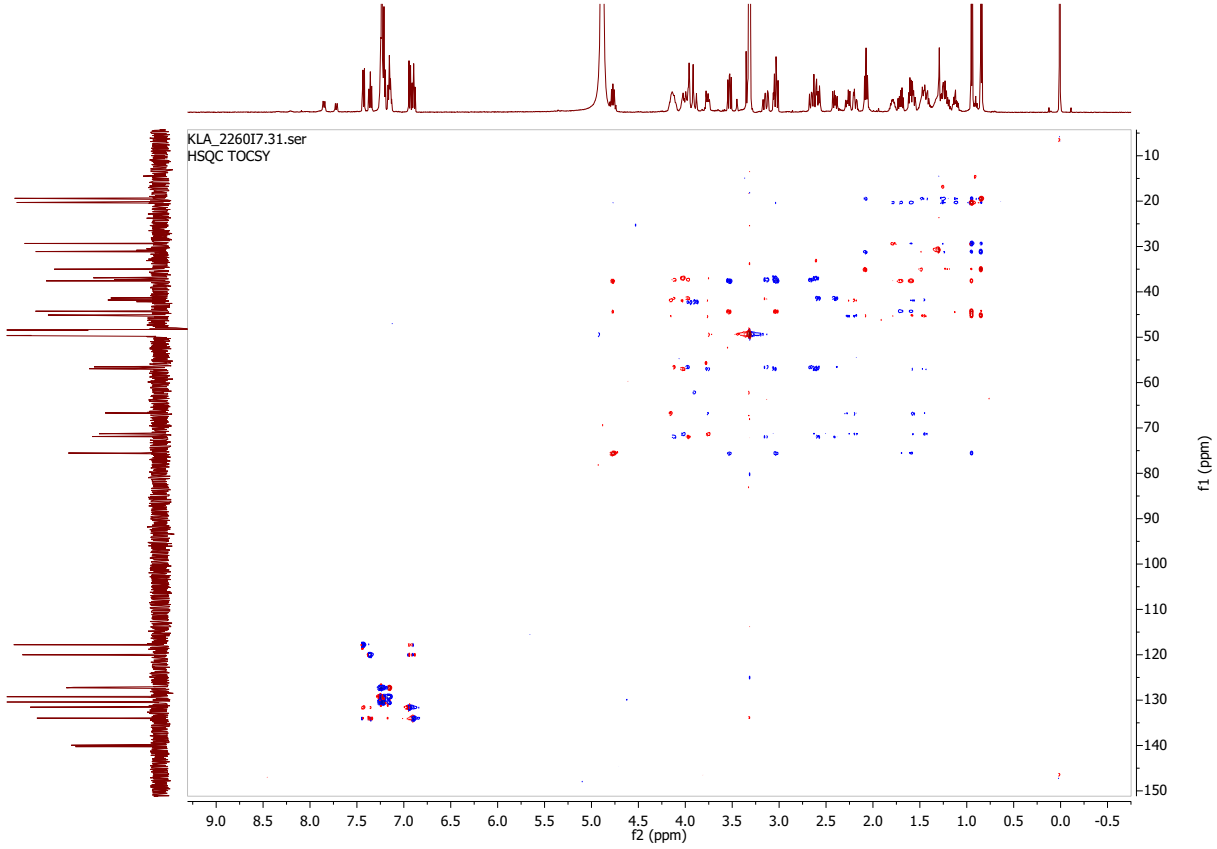


Figure 3.24: ^{1}H-^{13}C HSQC TOCSY spectra of fatuamide A in MeOH-d_4 (600 MHz).

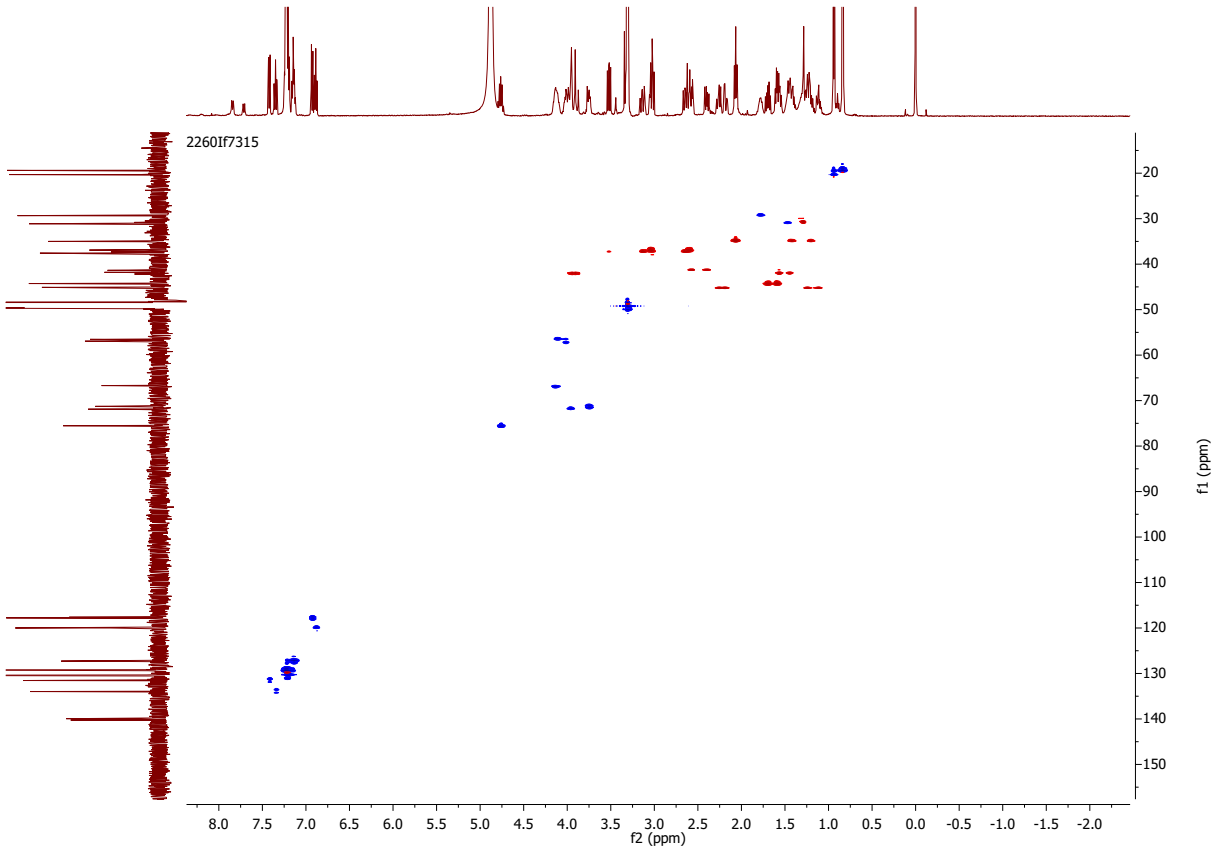


Figure 3.25: ^{1}H-^{13}C HSQC spectra of fatuamide A in MeOH-d_4 (500 MHz).

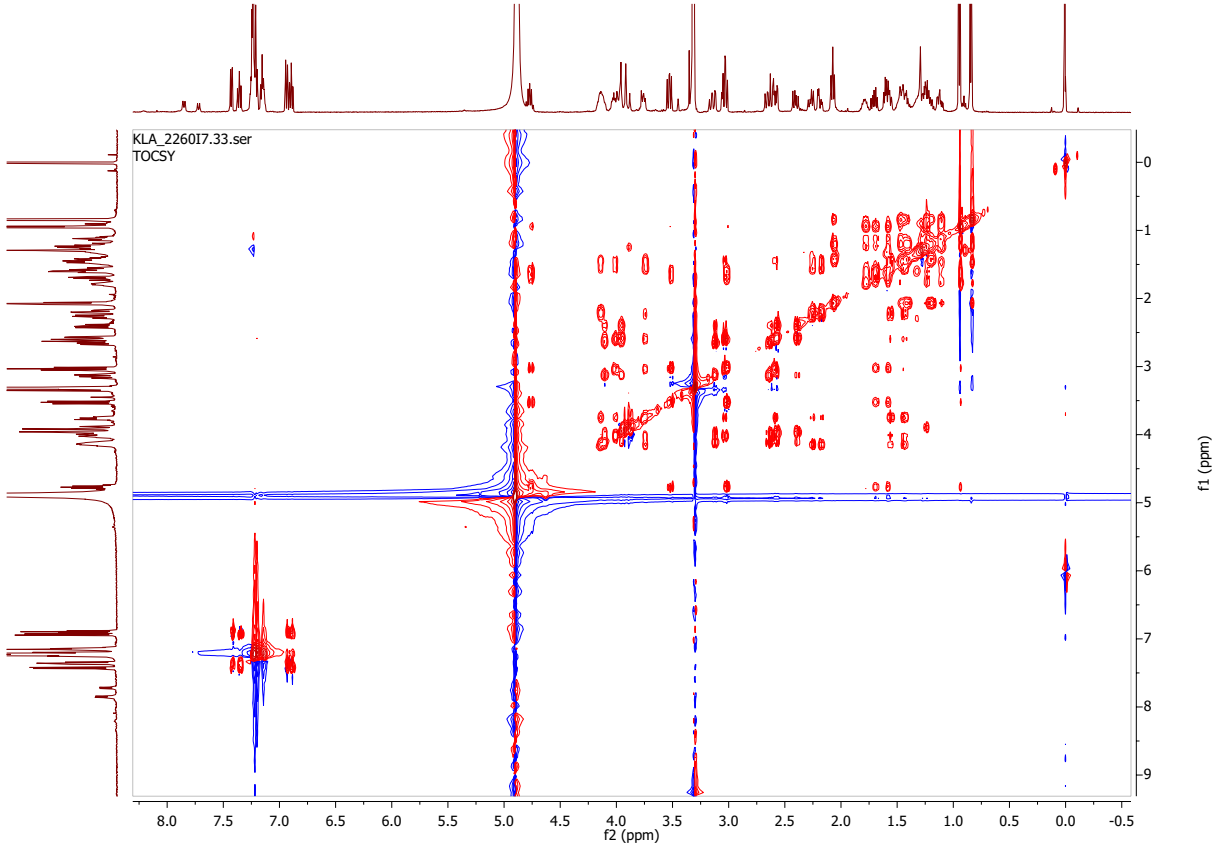


Figure 3.26: ^{1}H-^{1}H TOCSY spectra of fatuamide A in MeOH-d_4 (600 MHz).

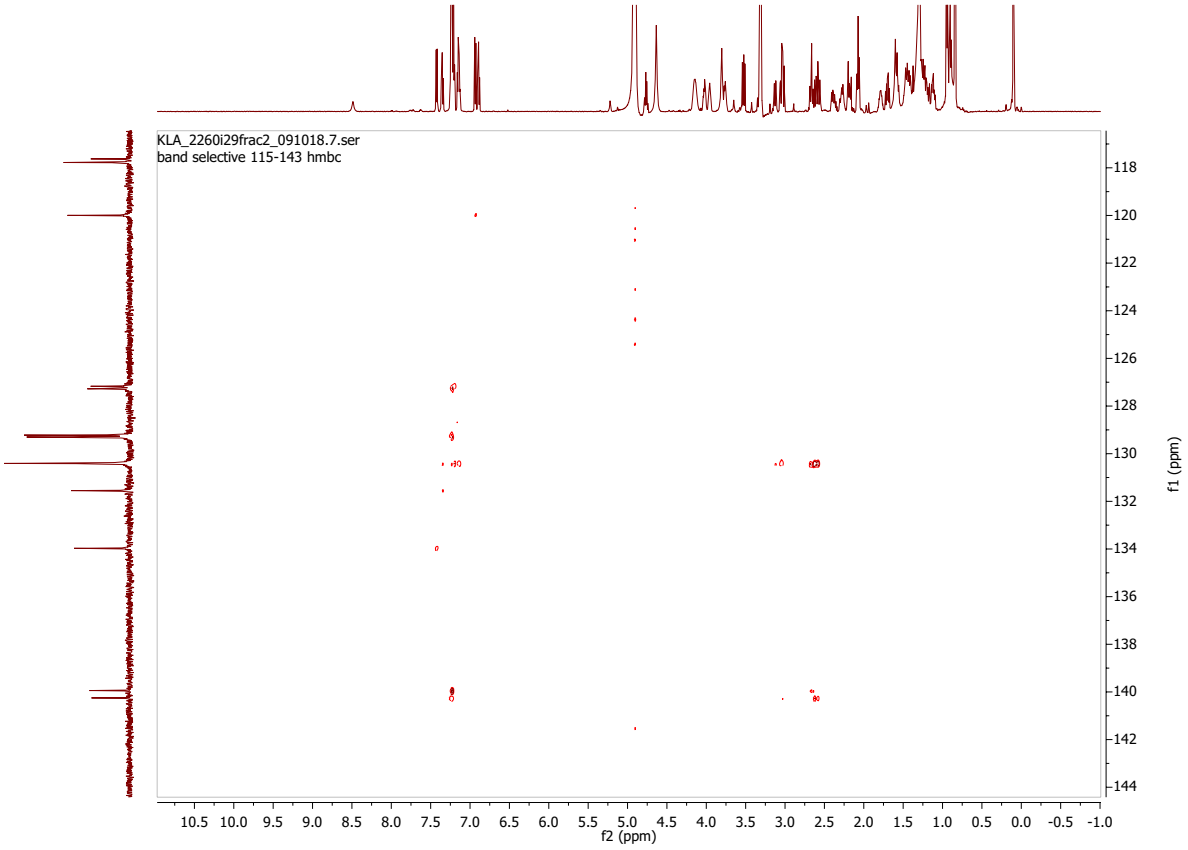


Figure 3.27: ^{1}H-^{13}C HMBC selective for δ_{13C} 115-143 ppm of fatuamide A in MeOH-d_4 (600 MHz).

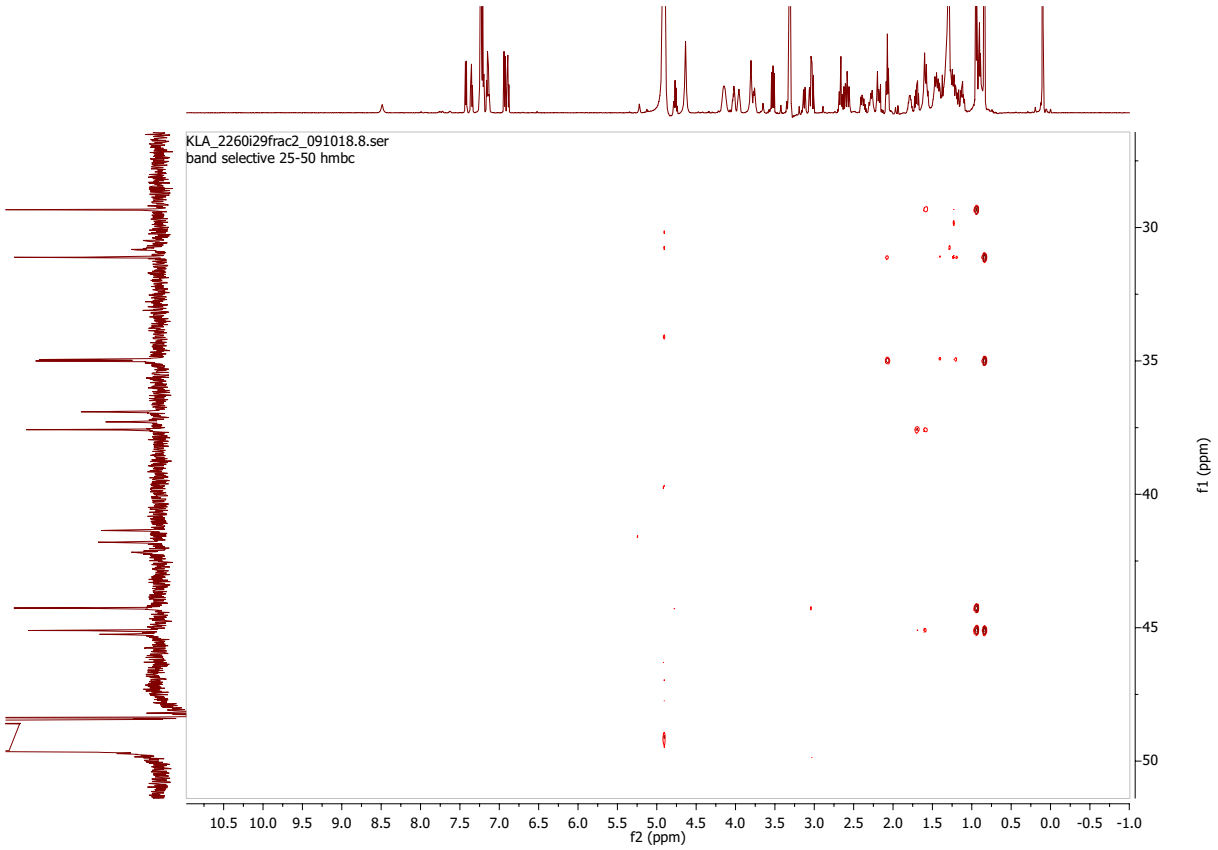


Figure 3.28: ^{1}H-^{13}C HMBC selective for δ_{13C} 25-50 ppm of fatuamide A in MeOH-d_4 (600 MHz).

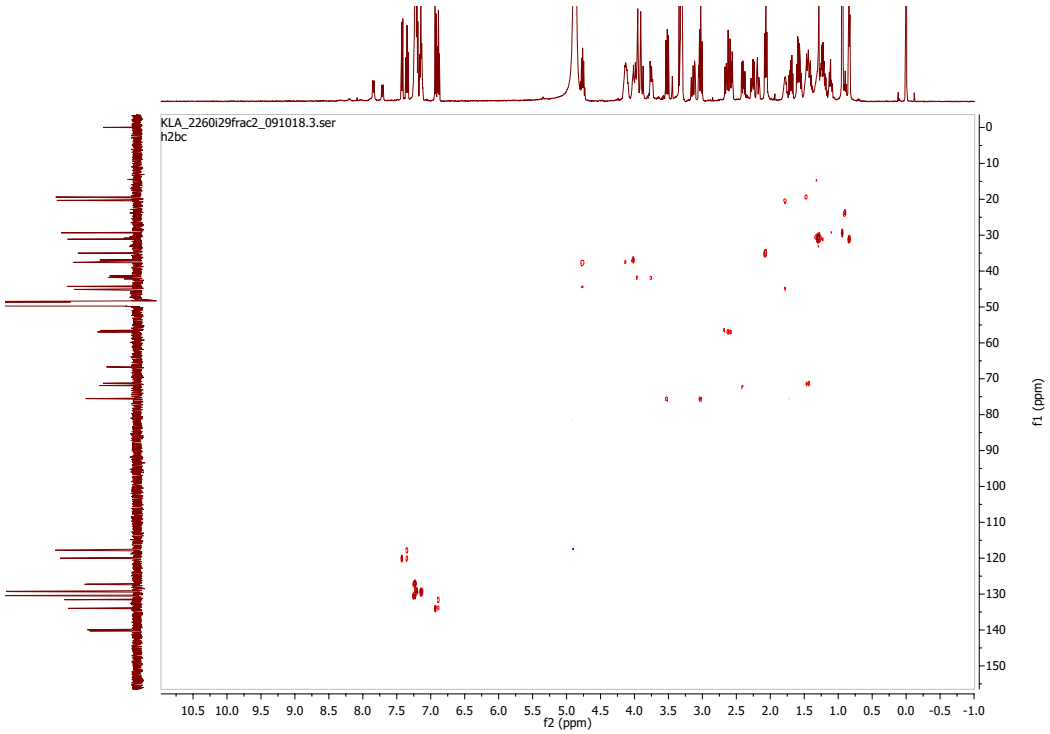


Figure 3.29: ^{1}H-^{13}C H2BC spectra of fatuamide A in MeOH-d_4(600 MHz).

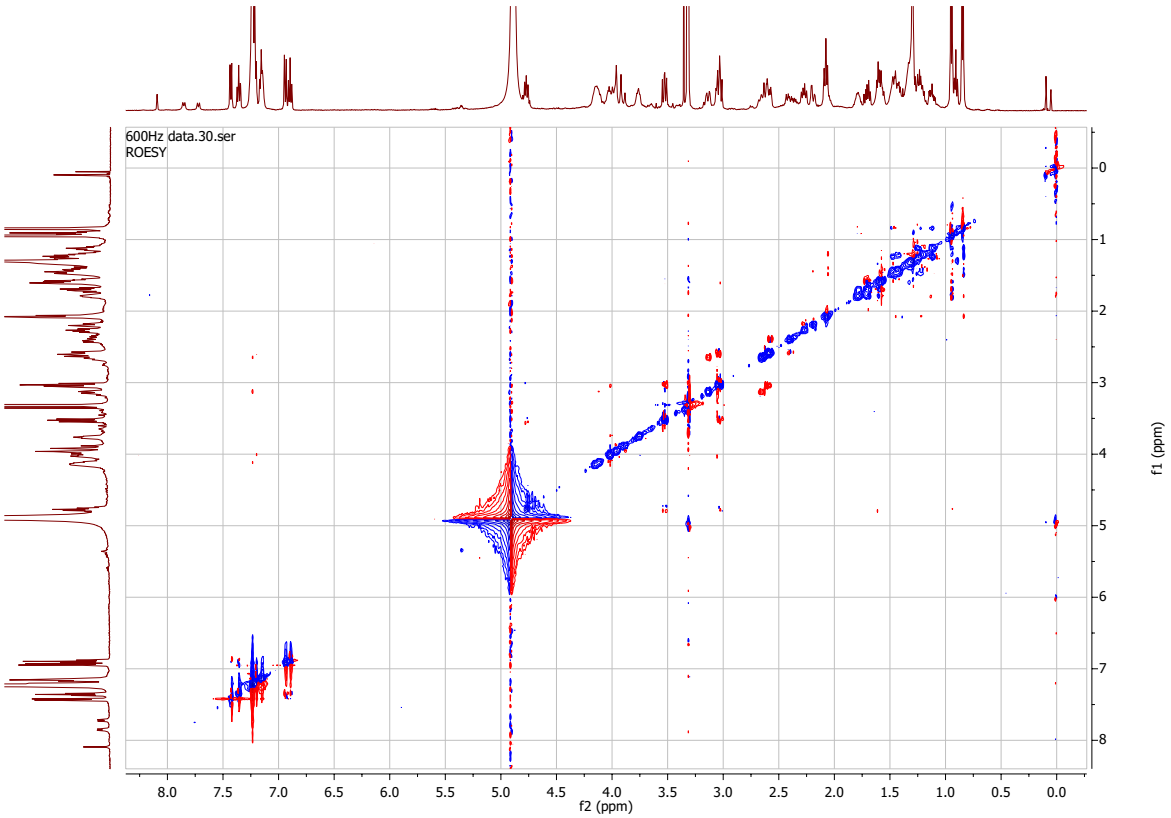


Figure 3.30: ^{1}H-^{1}H ROESY spectra of fatuamide A in MeOH-d_4 (600 MHz).

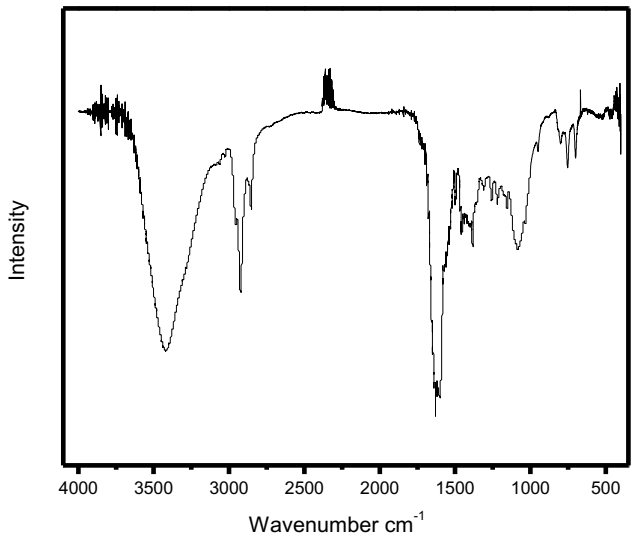


Figure 3.31: FT-IR spectrum of fatuamide A.

Table 3.3: Absorbance of the samples from the CAS assay. Decrease of absorbance value indicates a positive siderophore activity.

Sample ID	Abs @ 655.0
ASX22	0.0642
ASX22+CAS	0.3448
CAS + H2O	3.0376
media+CAS	1.9166
CAS	3.7858

3.6 Acknowledgements

We would like to thank the UCSD Chemistry and Biochemistry Mass Spectrometry facility and Dr. Su and Dr. Gross for the HREISIMS mass spec data. We would also like to thank Eva Ternon and the ECAL facility for use of the Orbitrap Elite hybrid MS. We thank Dr. Brendan Duggan for his help with the NMR experiments. We would also like to thank Sol Bloomfield, Andrew Ecker, and Syrena Whitner for help with culturing. We would also like to thank Tiago Leao for help with the bioinformatics. We would also like to thank Daniel Pietras and Raphael Reher for their help with the metal binding assay. We would also like to thank Bailey Miller and Nathan Moss for the collection of the culture in American Samoa.

Chapter 3, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Naman, C. Benjamin; Iwasaki, Arihiro; Leao, Tiago; Reher, Raphael; Kim, Hyun Woo; Ternon, Eva, Caro-Diaz, Eduardo J.E.; Glukhov, Evgenia; Duggan, Brendan M.; Gerwick, Lena; Gerwick, William H. The dissertation author was the primary investigator and author of this material

3.7 References

(1) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. Biologically Active Secondary Metabolites from Marine Cyanobacteria. *Curr. Opin. Biotechnol.* **2010**, *21* (6), 787–793. https://doi.org/10.1016/j.copbio.2010.09.019.

(2) Cancer.World Health Organization.

(3) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(4) Leao, T.; Castelão, G.; Korobeynikov, A.; Monroe, E. A.; Podell, S.; Glukhov, E.; Allen, E. E.; Gerwick, W. H.; Gerwick, L. Comparative Genomics Uncovers the Prolific and Distinctive Metabolic Potential of the Cyanobacterial Genus Moorea. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (12), 3198–3203. https://doi.org/10.1073/pnas.1618556114.

(5) Li, Y.; Naman, C. B.; Alexander, K. L.; Guan, H.; Gerwick, W. H. The Chemistry, Biochemistry and Pharmacology of Marine Natural Products from Leptolyngbya, a Chemically Endowed Genus of Cyanobacteria. *Mar. Drugs* **2020**, *18* (10). https://doi.org/10.3390/md18100508.

(6) Borowitzka, M. A.; Vonshak, A. Scaling up Microalgal Cultures to Commercial Scale. *Eur. J. Phycol.* **2017**, *52* (4), 407–418. https://doi.org/10.1080/09670262.2017.1365177.

(7) Nagoba, B.; Vedpathak, D. Medical Applications of Siderophores. *Eur. J. Gen. Med.* **2011**, *8* (3), 229–235. https://doi.org/10.29333/ejgm/82743.

(8) Song, W. Y.; Jeong, D.; Kim, J.; Lee, M. W.; Oh, M. H.; Kim, H. J. Key Structural Elements for Cellular Uptake of Acinetobactin, a Major Siderophore of Acinetobacter Baumannii. **2017**. https://doi.org/10.1021/acs.orglett.6b03671.

(9) DiSpirito, A. A.; Semrau, J. D.; Murrell, J. C.; Gallagher, W. H.; Dennison, C.; Vuilleumier, S. Methanobactin and the Link between Copper and Bacterial Methane Oxidation. *Microbiol. Mol. Biol. Rev.* **2016**, *80* (2), 387–409. https://doi.org/10.1128/mmbr.00058-15.

(10) Kubota, T.; Sakuma, Y.; Tsuda, M. Cyanobacterial Siderophores-Physiology, Structure, Biosynthesis and Applications. *Mar. Drugs* **2019**, *17* (5), 281.

(11) Simpson, F. B.; Neilands, J. B. Siderochromes in Cyanophycea: Isolation and Characterization of Schizokinen from Anabena Sp. *J. Phycol.* **1976**, *12*, 44–48.

(12) Ito, Y.; Butler, A. Structure of Synechobactins, New Siderophores of the Marine Cyanobacterium Synechococcus Sp. PCC 7002. *Limnol. Oceanogr.* **2005**, *50* (6), 1918–1923. https://doi.org/10.4319/lo.2005.50.6.1918.

(13) Beiderbeck, H.; Taraz, K.; Budzikiewicz, H.; Walsby, A. E. Anachelin, the Siderophore of the Cyanobacterium Anabaena Cylindrica CCAP 1403/2A. *Zeitschrift fur Naturforsch. - Sect. C J. Biosci.* **2000**, *55* (9–10), 681–687. https://doi.org/10.1515/znc-2000-9-1002.

(14) Leão, T.; Wang, M.; Moss, N.; da Silva, R.; Sanders, J.; Nurk, S.; Gurevich, A.; Humphrey, G.; Reher, R.; Zhu, Q.; Belda-Ferre, P.; Glukhov, E.; Whitner, S.; Alexander, K. L.; Rex, R.; Pevzner, P.; Dorrestein, P. C.; Knight, R.; Bandeira, N.; Gerwick, W. H.; Gerwick, L. A Multi-Omics Characterization of the Natural Product Potential of Tropical Filamentous Marine Cyanobacteria. *Mar. Drugs* **2021**, *19* (1).

https://doi.org/10.3390/md19010020.

(15) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299–302. https://doi.org/10.1038/s41592-019-0344-8.

(16) Nandiyanto, A. B. D.; Oktiani, R.; Ragadhita, R. How to Read and Interpret Ftir Spectroscope of Organic Material. *Indones. J. Sci. Technol.* **2019**, *4* (1), 97–118. https://doi.org/10.17509/ijost.v4i1.15806.

(17) Drechsel, H.; Stephan, H.; Lotz, R.; Haag, H.; Zähner, H.; Hantke, K.; Jung, G. Structure Elucidation of Yersiniabactin, a Siderophore from Highly Virulent Yersinia Strains. *Liebigs Ann.* **1995**, *1995* (10), 1727–1733. https://doi.org/10.1002/jlac.1995199510243.

(18) Cox, C. D.; Rinehart, K. L.; Moore, M. L.; Cook, J. C. Pyochelin: Novel Structure of an Iron-Chelating Growth Promoter for Pseudomonas Aeruginosa. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78* (7), 4256–4260. https://doi.org/10.1073/pnas.78.7.4256.

(19) Seyedsayamdost, M. R.; Traxler, M. F.; Zheng, S. L.; Kolter, R.; Clardy, J. Structure and Biosynthesis of Amychelin, an Unusual Mixed-Ligand Siderophore from Amycolatopsis Sp. AA4. *J. Am. Chem. Soc.* **2011**, *133* (30), 11434–11437. https://doi.org/10.1021/ja203577e.

(20) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(21) Fukuda, T. T. H.; Helfrich, E. J. N.; Mevers, E.; Melo, W. G. P.; Van Arnam, E. B.; Andes, D. R.; Currie, C. R.; Pupo, M. T.; Clardy, J. Specialized Metabolites Reveal Evolutionary History and Geographic Dispersion of a Multilateral Symbiosis. *ACS Cent. Sci.* **2021**, acscentsci.0c00978. https://doi.org/10.1021/acscentsci.0c00978.

(22) Pfeifer, B. A.; Wang, C. C. C.; Walsh, C. T.; Khosla, C. Biosynthesis of Yersiniabactin, a Complex Polyketide-Nonribosomal Peptide, Using Escherichia Coli as a Heterologous Host. *Appl. Environ. Microbiol.* **2003**, *69* (11), 6698–6702. https://doi.org/10.1128/AEM.69.11.6698-6702.2003.

(23) Miller, D. A.; Luo, L.; Hillson, N.; Keating, T. A.; Walsh, C. T. Yersiniabactin Synthetase: A Four-Protein Assembly Line Producing the Nonribosomal Peptide/Polyketide Hybrid Siderophore of Yersinia Pestis. *Chem. Biol.* **2002**, *9* (3), 333–344. https://doi.org/10.1016/S1074-5521(02)00115-1.

(24) Gehring, A. M.; Mori, I.; Perry, R. D.; Walsh, C. T. The Nonribosomal Peptide Synthetase HMWP2 Forms a Thiazoline Ring during Biogenesis of Yersiniabactin, an Iron-Chelating Virulence Factor of Yersinia Pestis. *Biochemistry* **1998**, *37* (33), 11637–11650. https://doi.org/10.1021/bi9812571.

(25) Kotowska, M.; Pawlik, K. Roles of Type II Thioesterases and Their Application for Secondary Metabolite Yield Improvement. *Appl. Microbiol. Biotechnol.* **2014**, *98* (18), 7735–7746. https://doi.org/10.1007/s00253-014-5952-8.

(26) Heathcote, M. L.; Staunton, J.; Leadlay, P. F. Role of Type II Thioesterases: Evidence for Removal of Short Acyl Chains Produced by Aberrant Decarboxylation of Chain Extender Units. *Chem. Biol.* **2001**, *8* (2), 207–220. https://doi.org/10.1016/S1074-5521(01)00002-3.

(27) Liu, T.; Cane, D. E.; Deng, Z. Chapter 9 The Enzymology of Polyether Biosynthesis. *Methods Enzymol.* **2009**, *459* (B), 187–214. https://doi.org/10.1016/S0076-6879(09)04609-6.

(28) Grindberg, R. V.; Ishoey, T.; Brinza, D.; Esquenazi, E.; Coates, R. C.; Liu, W. ting; Gerwick, L.; Dorrestein, P. C.; Pevzner, P.; Lasken, R.; Gerwick, W. H. Single Cell Genome Amplification Accelerates Identification of the Apratoxin Biosynthetic Pathway from a Complex Microbial Assemblage. *PLoS One* **2011**, *6* (4). https://doi.org/10.1371/journal.pone.0018565.

(29) Keatinge-clay, A. T.; Keatinge-clay, A. T. Stereocontrol within Polyketide Assembly Lines. **2016**, *33* (2). https://doi.org/10.1039/c5np00092k.

(30) Bertin, M. J.; Vulpanovici, A.; Monroe, E. A.; Korobeynikov, A.; Sherman, D. H.; Gerwick, L.; Gerwick, W. H. The Phormidolide Biosynthetic Gene Cluster: A Trans-AT PKS Pathway Encoding a Toxic Macrocyclic Polyketide. *ChemBioChem* **2016**, *17* (2), 164–173. https://doi.org/10.1002/cbic.201500467.

(31) Ndukwe, I. E.; Wang, X.; Lam, N. Y. S.; Ermanis, K.; Alexander, K. L.; Bertin, M. J.; Martin, G. E.; Muir, G.; Paterson, I.; Britton, R.; Goodman, J. M.; Helfrich, E. J. N.; Piel, J.; Gerwick, W. H.; Williamson, R. T. Synergism of Anisotropic and Computational NMR Methods Reveals the Likely Configuration of Phormidolide A. *Chem. Commun.* **2020**, *56* (55), 7565–7568. https://doi.org/10.1039/d0cc03055d.

(32) Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D. V.; Richard Hutchinson, C.; McDaniel, R. A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases. *Biochemistry* **2003**, *42* (1), 72–79. https://doi.org/10.1021/bi0268706.

(33) Piasecki, S. K.; Zheng, J.; Axelrod, A. J.; Detelich, M. E.; Keatinge-Clay, A. T. Structural and Functional Studies of a Trans-Acyltransferase Polyketide Assembly Line Enzyme That Catalyzes Stereoselective α- and β-Ketoreduction. *Proteins Struct. Funct. Bioinforma.* **2014**, *82* (9), 2067–2077. https://doi.org/10.1002/prot.24561.

(34) D. Cole Stevens, Drew T. Wagner, Hannah R. Manion, Bradley K. Alexander, and A. T. K.-C. Methyltransferases Excised from Trans-AT Polyketide Synthases Operate on N-Acetylcysteamine-Bound Substrates. *Physiol. Behav.* **2016**, *69* (7), 567–570. https://doi.org/10.1007/s13679-019-00335-3.Metformin.

(35) Adrian T. Keatinge-Clay. The Uncommon Enzymology of Cis-AcylTransferase Assembly

Lines. *Physiol. Behav.* **2017**, *117* (8), 5334–5366. https://doi.org/10.1021/acs.chemrev.6b00683.The.

(36) Ansari, M. Z.; Sharma, J.; Gokhale, R. S.; Mohanty, D. In Silico Analysis of Methyltransferase Domains Involved in Biosynthesis of Secondary Metabolites. *BMC Bioinformatics* **2008**, *9*, 1–21. https://doi.org/10.1186/1471-2105-9-454.

(37) Xie, X.; Khosla, C.; Cane, D. E. Elucidation of the Stereospecificity of C-Methyltransferases from Trans-AT Polyketide Synthases. *J. Am. Chem. Soc.* **2017**, *139* (17), 6102–6105. https://doi.org/10.1021/jacs.7b02911.

(38) Kwan, D. H.; Leadlay, P. F. Mutagenesis of a Modular Polyketide Synthase Enoylreductase Domain Reveals Insights into Catalysis and Stereospecificity. *ACS Chem. Biol.* **2010**, *5* (9), 829–838. https://doi.org/10.1021/cb100175a.

(39) Schmidt, Y.; Breit, B. Direct Assignment of the Relative Configuration in 1,3, n-Methyl-Branched Carbon Chains by 1H NMR Spectroscopy. *Org. Lett.* **2010**, *12* (10), 2218–2221. https://doi.org/10.1021/ol1005399.

(40) Novak, T.; Tan, Z.; Liang, B.; Negishi, E. I. All-Catalytic, Efficient, and Asymmetric Synthesis of α,ω- Diheterofunctional Reduced Polypropionates via "One-Pot" Zr-Catalyzed Asymmetric Carboalumination - Pd-Catalyzed Cross-Coupling Tandem Process. *J. Am. Chem. Soc.* **2005**, *127* (9), 2838–2839. https://doi.org/10.1021/ja043534z.

(41) Brand, G. J.; Studte, C.; Breit, B. Iterative Synthesis of (Oligo) Deoxypropionates via Cross-Coupling. *Org. Lett.* **2009**, *3* (7), 9–11.

(42) Title: Native Electrospray-Based Metabolomics Enables the Detection of Metal-Binding Compounds. **2019**.

(43) Neilands, B. Universal Chemical Assay for the Detection Determination of Siderophores'. **1987**, *56*, 47–56.

(44) Hasegawa, H.; Maki, T.; Asano, K.; Ueda, K.; Ueda, K. Detection of Iron(III)-Binding Ligands Originating from Marine Phytoplankton Using Cathodic Stripping Voltammetry. *Anal. Sci.* **2004**, *20* (1), 89–93. https://doi.org/10.2116/analsci.20.89.

(45) Asegawa, H. H.; Atsui, M. M.; Uzuki, M. S.; Aito, K. N.; Eda, K. U.; Ohrin, Y. S. The Possibility of Regulating the Species Composition of Marine Phytoplankton Using Organically Complexed Iron. **2001**, *17* (January), 209–211.

Chapter Four: Structure Determination of Curacin Analogs

4.1 Abstract:

New compounds curacin F and G were isolated from a collection of *Moorena producens* obtained from Culebra, Puerto Rico. These new analogues are structurally similar to the antimitotic agents curacin A-D (from *Lyngbya Majuscula*) and were targeted for isolation after clustering with curacin A in an LC-MS/MS molecular network. The planar structures were determined through NMR and MS/MS analysis.

4.2 Introduction:

Curacin A ***(1)*** was initially isolated by Gerwick and coworkers from a collection of what was called at the time *Lyngbya majuscula* (now known as *Moorena producens*) that was made in Curaçao and was noted for its antimitotic and antiproliferative activities.[1] Further studies have shown that curacin A competitively inhibits the binding of [^{3}H]colchicine to tubulin.[2] Its unique structure, which contains a thiazoline and cyclopropyl ring, was found to be derived from a hybrid PKS and NRPS system.[3]

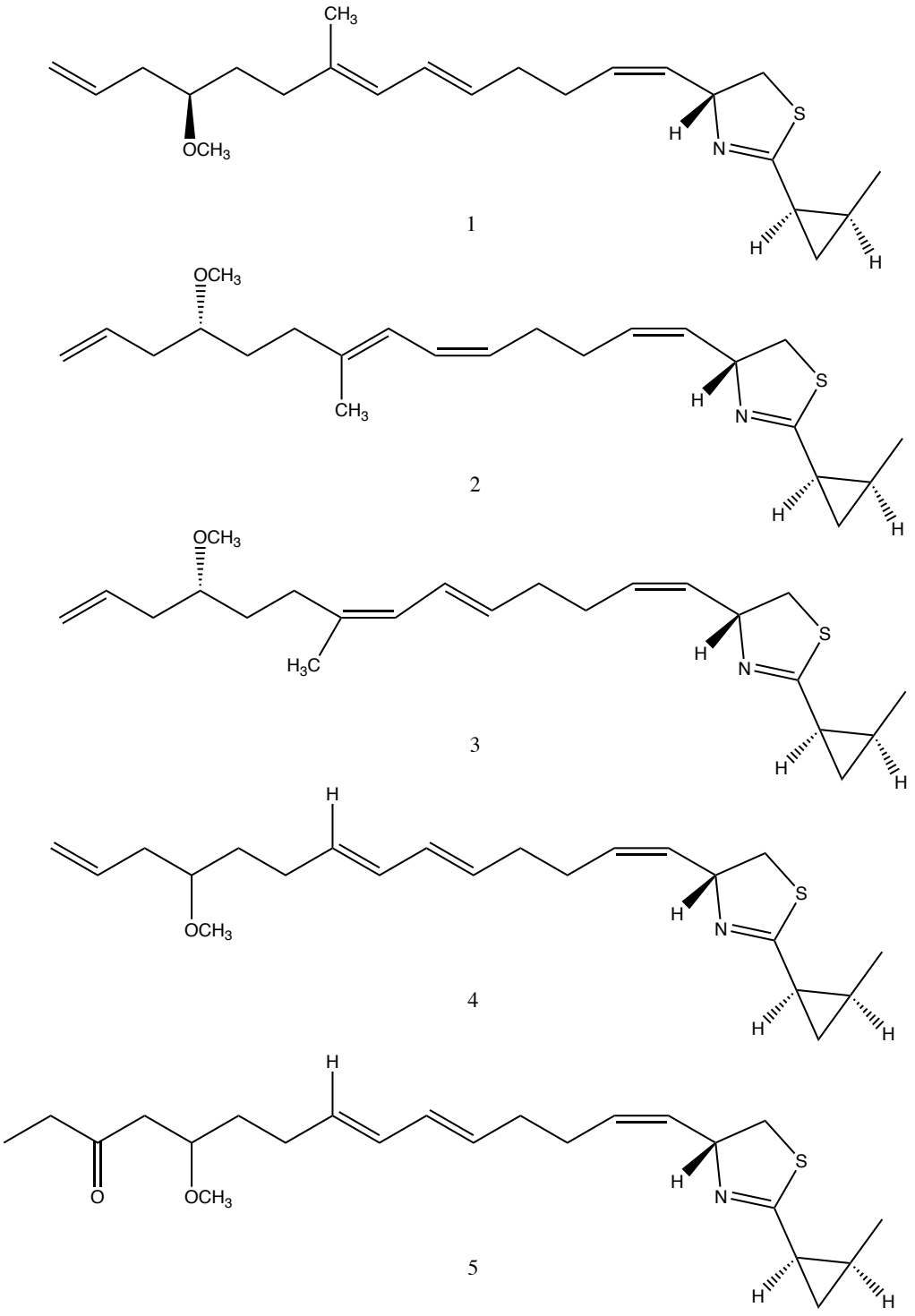


Figure 4.1: Curacin A (**1**) and known analogues B (**2**), C (**3**), D (**4**), and E (**5**).

Since the isolation of curacin A, several other analogues have been reported. Curacin B (**2**) and curacin C (**3**), which differ in their double bond geometry compared to curacin A, have

been found to also possess antimitotic activity.[4] Curacin D (**4**), also an antimitotic agent, has a proton that replaces the methyl at C-10.[5] Remarkably, an extract from the brittle star species *Ophiocoma scolopendrina* afforded curacin E (**5**), which has a unique carbonyl function at C-15.[6] Curacin E showed some cytotoxicity to P388 murine leukemia cells.[6]

The total synthesis of curacin A was first reported by White and coworkers.[7] Since then, additional analogues of curacin A that have been synthesized,[8–10] one of which was found to be a more potent inhibitor of tubulin assembly.[8]

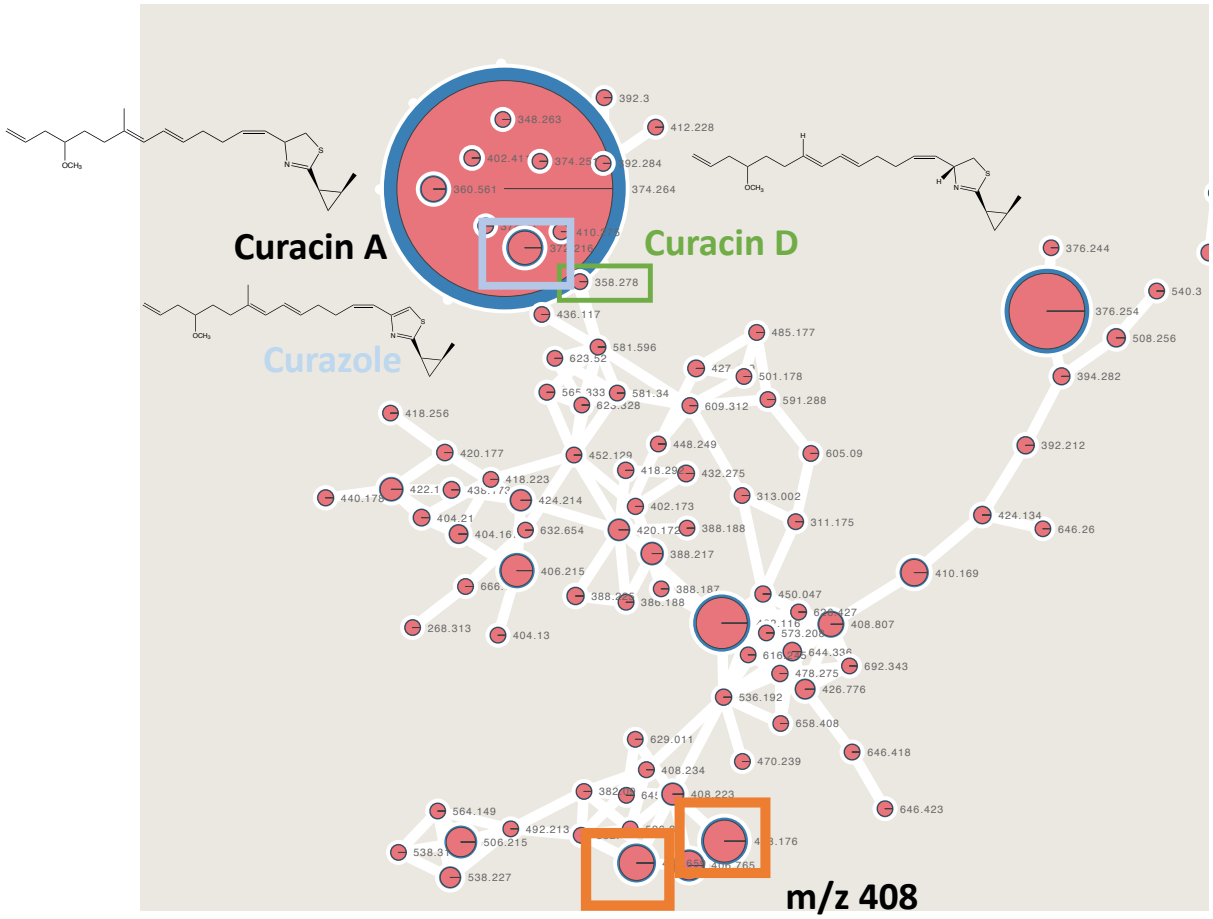


Figure 4.2: GNPS cluster that contains curacin A and analogues (curacin D and curazole) from LC-MS/MS molecular network. The orange boxes highlight new analogues of interest.

A collection of *Moorena producens (*formerly named *Moorea producens*) was made in Culebra, Puerto Rico. *Moorena producens* are a species of cyanobacteria that have been known

to produce products that can cause inflammation and blistering.[11] The *Moorena* genus is a prolific producer of natural products and has been reported to encompass more than 40% of all reported marine cyanobacteria natural products.[12]

After an initial analysis by LC-MS/MS molecular networking, the extract was found to contain curacin A. Since curacin A has shown potential as an anticancer therapeutic, there was interest to isolate and characterize natural analogs. This targeted isolation of curacin A analogs was initiated by investigation of the cluster containing curacin A in the LC-MS/MS molecular network (*figure 4.2*). As a result, we report herein the isolation and characterization of the new analogues curacin F and curacin G. They differ from previously isolated curacin anologues by the inclusion of hydroxy groups at C-10 and C-7, and instead of a diene between C7-C10. Furthermore, curacin F and G have a single double bond between C-8 to C-9.

4.3 Results and Discussion:

4.3.1 Collection, Extraction, and Isolation

A collection of reddish colored filamentous cyanobacterium (CUM3APR19-1) was made by snorkel and SCUBA diving at Punta de Maguey in Culebra, Puerto Rico on April 3, 2019. Based upon its morphology it was identified in the field to be a *Moorena producens.* The sample was extracted in DCM:MeOH (2:1) and a portion was fractionated by Vacuum Liquid Chromatography (VLC) to afford 9 fractions (A-I). Additional portions were fractionated by Combiflash column chromatography. Curacin A was collected with a retention time of 7 min. Purification of 1.5 mg of curacin F was achieved through RP-HPLC on a Kinetex C18 semipreparative column with ACN/H_2O after the initial fractionation by Combiflash column chromatography using hexanes and ethyl acetate. Purification of 3.4 mg of curacin G was achieved through RP-HPLC on a Synergi 4u Hydro-RP column with ACN/H_2O.

4.3.2 Structure Elucidation of Curacin F and G

Table 4.1: ^{1}H and ^{13}C NMR spectroscopic data of curacin F in C_6D_6.

	Curacin F		
carbon #	**δ_C, type**	**δ_H (J in Hz)**	**Selected HSQC (H->C)**
1a	39.24, CH_2	2.9, dd (10.9, 8.3)	
1b	39.24, CH_2	2.62, m	
2	72.84, CH	4.97, m	18
3	130.08, CH	5.4, dddd (10.6, 7.4, 3.2, 1.6)	2
4	133.36, CH	5.34, m	
5a	24.34, CH_2	1.93, dp (13.0,4.5)	
5b	24.34, CH_2	2.68, m	
6a	37.28, CH_2	1.48, m	
6b	37.28, CH_2	1.7, m	
7	68.18, CH	4.24, m	
8	131.67, CH	5.9, m	7
9	136.5, CH	5.9, m	10
10	72.1, C	-	
11a	38.63, CH_2	1.61,m	
11b	38.63, CH_2	1.65, m	10
12a	27.89, CH_2	1.61, m	11
12b	27.89, CH_2	1.65, m	11
13	81.00, CH	3.06, m	
14a	37.88, CH_2	2.17, m	13
14b	37.88, CH_2	2.25, m	
15	135.29, CH	5.84, m	
16	116.68, CH	5.04, m	
17	28.98, CH_3	1.25, s	9,10,11,12
18	170.9, C	-	
19	20.11, CH	1.57, m	
20a	14.36, CH_2	0.7, tdd (8.3, 4.8, 1.0)	
20b	14.36, CH_2	1.26	18
21	16.04, CH	0.9, ddddd (10.3, 8.1, 6.4, 5.1, 1.6)	
22	12.39, CH_3	1.13, dd (6.3,1.7)	
23	56.23, CH_3	3.13 d (5.7)	13

The molecular formula of curacin F was determined to be $C_{23}H_{37}NO_3$ (*m/z* 408) and has 6 degrees of unsaturation. The ^{1}H NMR spectrum had a methylene group at δ_H 2.90 and 2.62 which was associated with a carbon atom at δ_C 39.24 by HSQC, and were therefore assigned to the methylene group of a thiazoline ring. This was also established by HMBC correlations to

carbon atoms at 72.84 and 170.9, similar to curacin A. The thiazoline ring accounted for two degrees of unsaturation. An additional degree of unsaturation was accounted for from a second ring structure. This was shown by COSY correlations between two methine protons and one methylene group (δ_H 1.57, 0.90, 0.70 and 1.26), forming a cyclopropyl group, and again in close similarity to the data for curacin A. The final three degrees of unsaturation were due to three double bonds. This was evident from six carbon atoms at δ_C 130.1, 133.6, 131.7, 136.5, 135.3, 116.7 for C3-C-4, C8-C9 and C15-C16, respectively.

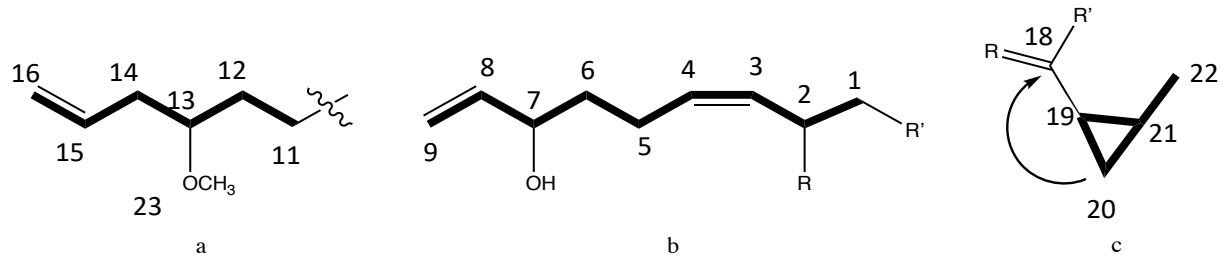


Figure 4.3: Partial structures of curacin F.

The following partial structures (*figure 4.3*) were determined from analysis of the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HSQC data. Starting with the protons δ_H 5.04 and 5.84, which corresponded to the terminal double bond at C16 and C15, respectively, were connected by COSY to the carbon chain C14 to C11. C13 (δ_c 81.12) was shown by HSQC to contain a single proton, and by chemical shift, suggested a methoxy group at this position, identical to curacin A. This was confirmed by an HMBC correlation between a methoxy methyl group (C23, δ_c 56.2, δ_H 3.13) and C13, completing fragment A. Fragment B consisted of another carbon chain, C1 to C9, that was connected by COSY correlations. C3 and C4 (δ_c 130.1 and 133.4) corresponded to a double bond as well as C8 and C9 (δ_c 131.7 and 136.5). A methine proton (δ_H 4.24) was present at C7 (δ_c 72.8), suggesting a hydroxy group at this position. C2 had a methine proton at (δ_H 4.97) suggesting connection to a heteroatom as in curacin A, and completing fragment B. Fragment C

was comprised of a cyclopropyl ring consisting of C19-C21, which was deduced from COSY correlations. C21, a methine proton (δ_H 0.90) , was correlated by COSY to a methyl group at δ_H 1.13 (C22). The methylene proton at δ_H 1.26 at C20 showed an HMBC correlation to the quaternary carbon at C18 (δ_c 170.9), and thus completed fragment C. Left unassigned at this point was a quaternary carbon atom and a methyl group, as these were not connected by COSY to any of these three partial structures.

The connections between the three partial structures of curacin F were made through HMBC correlations (*figure 4.4*). The proton H2 in fragment B showed a correlation to the quaternary carbon C18 of fragment C. This, with the chemical shifts of C1 and C2 established a thiazoline ring that connected fragments B and C, just as in curacin A. The H11 proton had an HMBC correlation to the remaining quaternary carbon C10 (δ_c 72.1). The chemical shift of C10 indicated that it had an oxygen substituent, likely as a hydroxy group. The remaining methyl group, C23, showed HMBC correlations to both H11 and H9, suggesting its position to also be at the quaternary carbon (C10). Therefore, the quaternary carbon (C10) was deduced to possess both a methyl group (C23) and a hydroxy group, and connected fragment A to B.

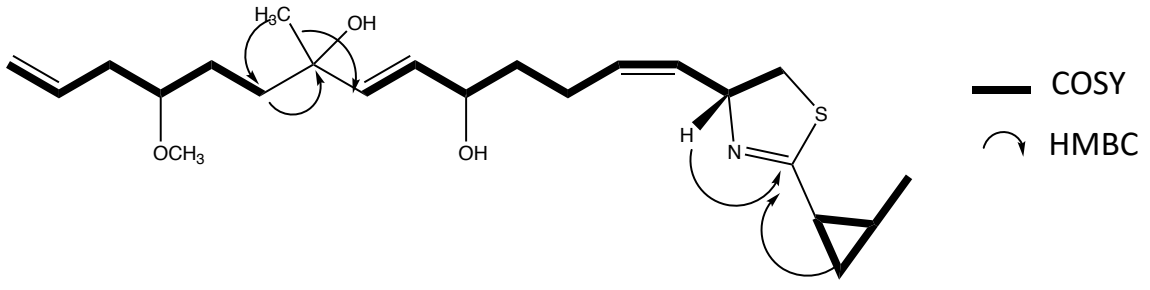


Figure 4.4: Structure of curacin F with COSY and HMBC correlations.

The structure of curacin F shows striking similarities to curacin A; the latter compound was a major constituent in the LC-MS/MS molecular network (*figure 4.2*). Curacin A possesses a diene between carbons C7- C8 and C9- C10 whereas curacin F has one double bond in that

region (C8 to C9). Curacin F also has two additional hydroxy groups not present in curacin A at positions C7 and C10. These differences are apparent in the MS/MS of curacin F which shows two sequential losses of 18 amu for the neutral ion loss of two water molecules (*figure 4.5*).

Table 4.2: ^{1}H and ^{13}C NMR spectroscopic data of curacin G in C_6D_6.

	Curacin G		
carbon #	**δ_C, type**	**δ_H (J in Hz)**	**Selected HSQC (H->C)**
1a	39.37, CH_2	2.71 td (10.8, 1.4)	2
1b	39.37, CH_2	3.01 dd (10.7, 8.4)	18
2	73.89, CH	4.98, m	18
3	129.77, CH	5.5, t (8,8)	
4	131.85, CH	5.39 t (8.3)	2
5a	24.16, CH_2	2.15, m	7
5b	24.16, CH_2	2.21, m	7
6a	36.94, CH_2	1.51, m	
6b	36.94, CH_2	1.57, m	
7	70.68, CH	4.1, m	
8	130.7, CH	5.76, m	10
9	137.25, CH	5.76, m	10
10	72, C	-	
11a	38.05, CH	1.57, m	10
11b	38.05	1.67, m	
12a	27.74, CH_2	1.59, m	
12b	27.74, CH_2	1.65, m	
13	80.66, CH	3.06, m	
14a	38.05, CH_2	2.15, m	13
14b	38.05, CH_2	2.23, m	13
15	135.12, CH	5.82, m	
16	116.49, CH	5.05, m	
17	27.81, CH_3	1.23, d (2.9)	9,10
18	170.3, C	-	
19	19.64, CH	1.67,m	
20a	14.83, CH_2	0.73, td (8.0,4.9)	
20b	14.83, CH_2	1.22 d (2.9)	18
21	19.64, CH	0.95	
22	12.06, CH_3	1.14 d (6.2)	
23	55.92, CH_3	3.12 d (4.7)	

Unfortunately, at this point in the structure elucidation process, curacin F degraded; chemical lability has been reported with other curacin anologues.[1] Therefore, isolation of additional amounts of the compound *m/z* 408 was attempted. Table 2 shows the NMR shifts of this isolation of the *m/z* 408 compound; comparison with the shifts of curacin F revealed that it was different. The NMR data gave the same partial structures as deduced for curacin F (*figure 4.3*). Additionally, curacin G, showed the same correlations between partial structures as curacin F (*figure 4.4*), indicating that curacin F and G have the same planar structure. However, there were some differences in the chemicals shifts between F and G (Table 3). These differences in chemical shifts suggested that they were stereoisomers of one another.

Table 4.3: ^{1}H and ^{13}C NMR spectroscopic data of curacin F and G in C_6D_6.

Curacin F		Curacin G	
δ_C, type	δ_H (J in Hz)	δ_C, type	δ_H (J in Hz)
39.24, CH_2	2.9, dd (10.9, 8.3)	39.37, CH_2	2.71 td (10.8, 1.4)
39.24, CH_2	2.62, m	39.37, CH_2	3.01 dd (10.7, 8.4)
72.84, CH	4.97, m	73.89, CH	4.98, m
130.08, CH	5.4, dddd (10.6, 7.4, 3.2, 1.6)	129.77, CH	5.5, t (8,8)
133.36, CH	5.34, m	131.85, CH	5.39 t (8.3)
24.34, CH_2	1.93, 2.68, dp (13.0,4.5) , m	24.16, CH_2	2.15,2.21, m
37.28, CH_2	1.7, 1.48, m	36.94, CH_2	1.57, 1.51, m
68.18, CH	4.24, m	70.68, CH	4.1, m
131.67, CH	5.9, m	130.7, CH	5.76, m
136.5, CH	5.9, m	137.25, CH	5.76, m
72.1, C	-	72.0, C	-
38.63, CH_2	1.65, 1.61, m	38.05, CH	1.67,1.57, m
27.89, CH_2	1.65, 1.61, m	27.74, CH_2	1.65,1.59, m
81.00, CH	3.06, m	80.66, CH	3.06, m
37.88, CH_2	2.17, 2.25, m	38.05, CH_2	2.23, 2.15, m
135.29, CH	5.84, m	135.12, CH	5.82, m
116.68, CH	5.04, m	116.49, CH	5.05, m
28.98, CH_3	1.25, s	27.81, CH_3	1.23, d (2.9)
170.9, C	-	170.3, C	-
20.11, CH	1.57, m	19.64, CH	1.67,m
14.36, CH_2	0.7, tdd (8.3, 4.8, 1.0)	14.83, CH_2	0.73, td (8.0,4.9)
14.36, CH_2	1.26	14.83, CH_2	1.22 d (2.9)
16.04, CH	0.9, ddddd (10.3, 8.1, 6.4, 5.1, 1.6)	19.64, CH	0.95
12.39, CH_3	1.13, dd (6.3,1.7)	12.06, CH_3	1.14 d (6.2)
56.23, CH_3	3.13 d (5.7)	55.92, CH_3	3.12 d (4.7)

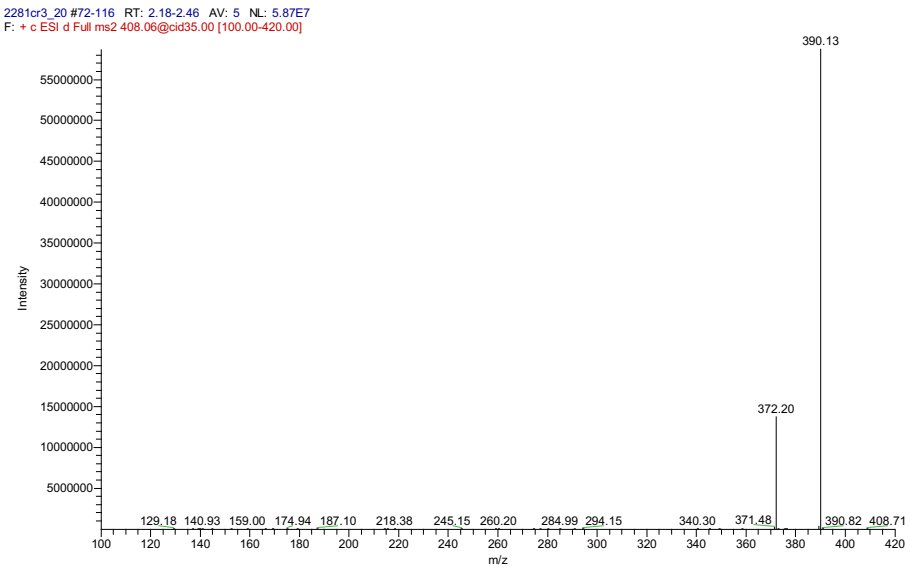


Figure 4.5: MS/MS spectrum of curacin F.

The inclusion of the 1,4 diol in Curacin F and Curacin G, and the fact that they are epimeric to one another, could suggest that they are isolation artifacts of curacin A rather than new natural products. However, these compounds were observed by LCMS from the very beginning analysis of the crude extract, suggesting that they are both natural products.

The LC-MS/MS molecular network showed additional analogues with *m/z* 422. These analogues were isolated, and had peaks by ^{1}H NMR that corresponded to methoxy groups. To investigate if these groups were an artifact of isolation, an experiment was done with a micro extraction of fresh tissue with DCM:CD_3OD (2:1) and DCM:CH_3OH (2:1) and the chemical profiles were compared from each extraction by LCMS (*figure 4.6).* This experiment showed that the *m/z* 422 methoxy analogues incorporated deuterium from the extraction with CD_3OD, which was apparent by the mass shift of *m/z* 422 to *m/z* 425, indicating that they were extraction artifacts. Therefore, the *m/z* 408 analogues were prioritized instead. It is interesting to speculate on the nature of the precursor to these methoxy analogs; it is conceivable that an epoxide containing precursor is the true natural product which reacts with methanol to form these analogs (*figure*

4.7).

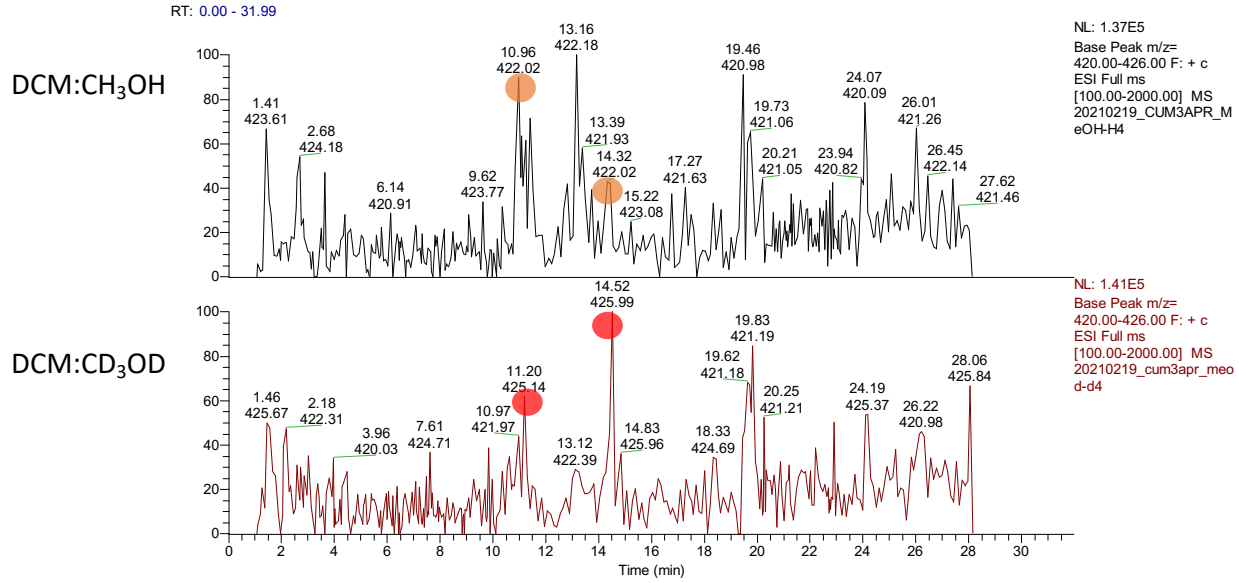


Figure 4.6: LCMS traces of *m/z* 380-430 for extracts of the culture CUM3APR19-1 using either DCM:CH_3OH (2:1, top) or DCM:CD_3OD (2:1, bottom). The methoxy-containing curacin analogues (*m/z* 422) are highlighted by orange circles in the top trace. The incorporation of deuterium into these methoxy-containing curacin analogues (*m/z* 425) are highlighted with red circles in the bottom trace.

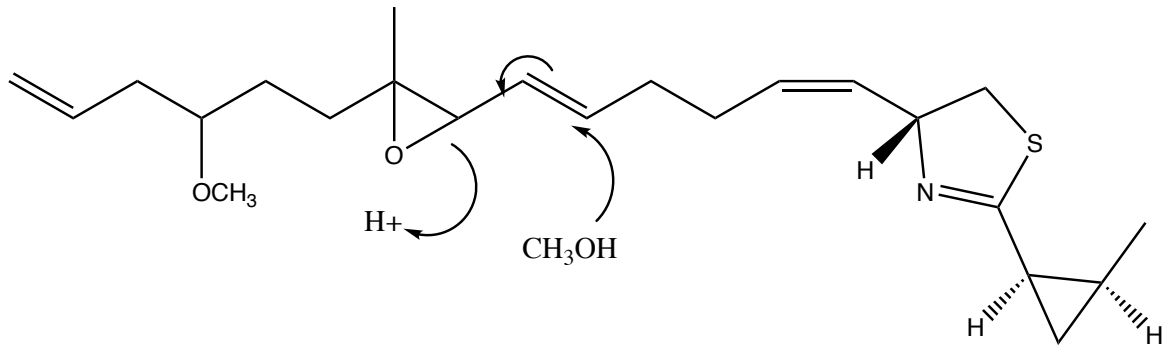


Figure 4.7: Reaction of possible epoxide precursor with methanol to form *m/z* 422 compound.

The MS/MS spectra of curacin F showed a prominent fragment at *m/z* 390. A compound with this molecular ion was also seen in small quantities in the cyanobacterial extract. By diode array detection of this eluting material, an ultraviolet absorption of 240 nm was observed, indicating that this compound possessed a conjugated diene moiety. The ms/ms of compound

with *m/z* 390 (*figure 4.8*) had fragments that suggested it may be a precursor to curacin F in that it possesses a diene with one hydroxy group. This is depicted in Scheme 1 in figure 4.9, in which the diene precursor is hydrolyzed to afford curacin F. However, curacin F could be the precursor to the diene as proposed in scheme 2 of figure 4.9. In this case, an enzyme removes the hydroxy group after the biosynthesis of curacin F and a double bond is formed, giving the diene compound.

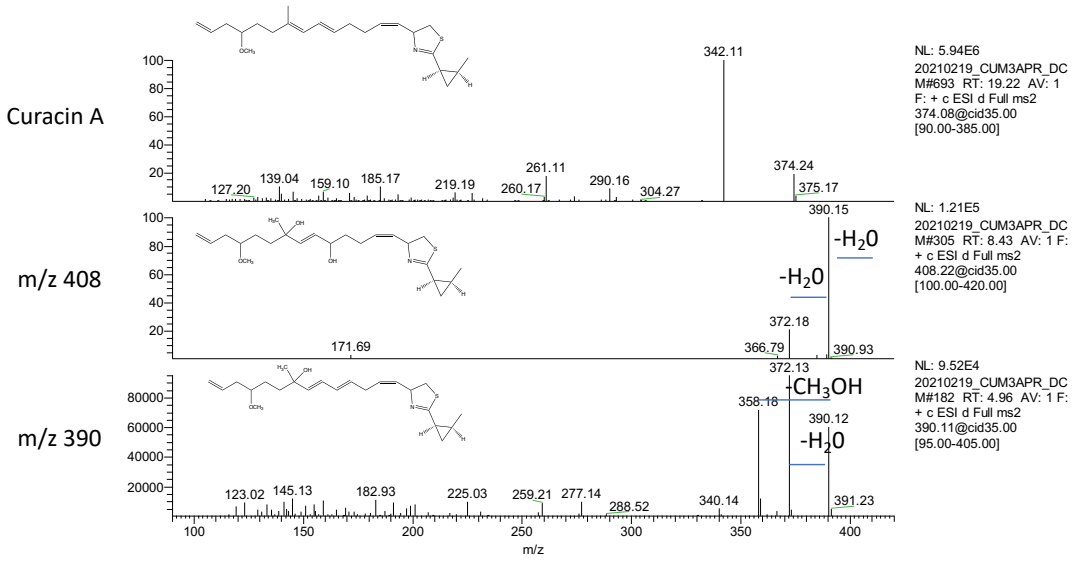


Figure 4.8: Comparison of MS/MS of curacin A (*m/z* 374), curacin F (*m/z* 408) and *m/z* 390

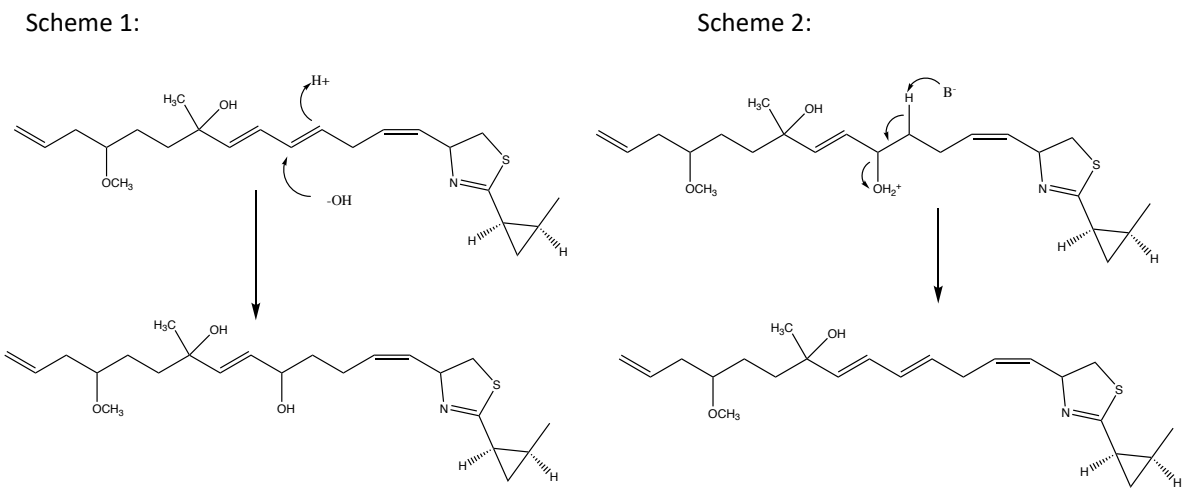


Figure 4.9: Different possible schemes relating curacin F to the diene anologue (*m/z* 390). In Scheme 1, the diene is the precursor and leads to the formation of curacin F, whereas in scheme 2, curacin F is the precursor and base-catalyzed loss of water results in the diene.

4.3.3 Cytotoxicity Testing

The fractions produced from vacuum liquid chromatography (VLC) were tested against H460 human lung carcinoma cells for cytotoxicity (*figure 4.10*). The crude extract killed 90.8% of the H460 cells at 1µg/mL. The resulting VLC fractions A, C, D, E, F, G, and H all showed cytotoxicity to H460 cell line. Curacins F and G were present in fractions E-H. Curacin G was pursued for isolation and structure determination because it was similar to the bioactive compound curacin A. The pure compound curacin G was therefore tested against lung cancer H460 cells. However, it only showed slight cytotoxicity at the highest concentration tested, 302 µM, and thus is essentially an inactive compound. However, upon checking the sample by LCMS after cytotoxicity testing, curacin G appeared to have degraded. Therefore, it remains a possibility that curacin G has cytotoxic properties.

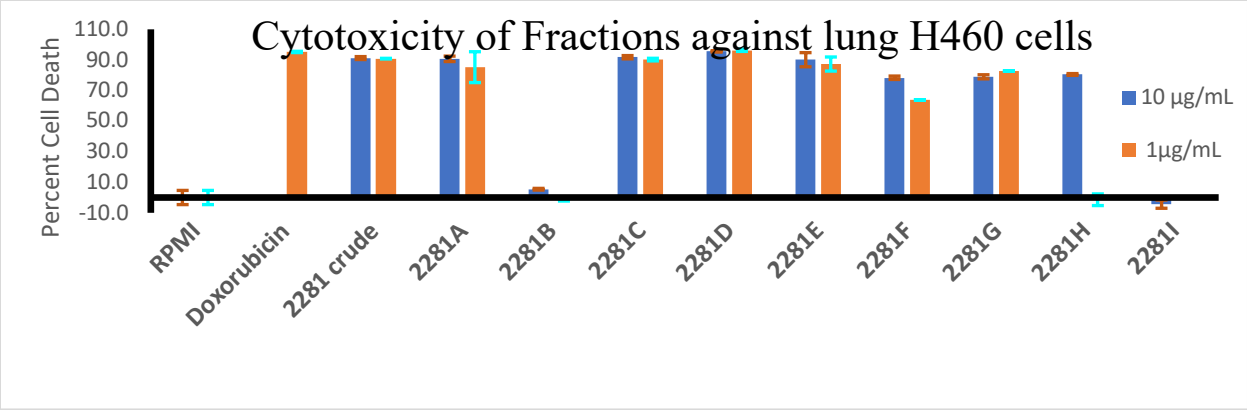


Figure 4.10: Cytotoxicity of extract and fractions from CUM3APR19-1 to human lung H460 cancer cells.

4.4 Conclusion

A collection *of Moorena producens* from Puerto Rico gave rise to two new curacin analogs. Curacin A was originally isolated from a reported *Lyngbya majuscula* (subsequently renamed to *M. producens*) obtained from Curaçao.[1] This study highlights the presence of curacin A and various analogs can be obtained from collections made in a different geographical locations. These anologs further illustrate the flexibility of the curacin A biosynthetic gene cluster, with the presence of new hydroxy groups at position C7 and C10. Since curacin A is a potent microtubule inhibitor, these anologs could have therapeutic potential as anticancer lead compounds.

4.5 Experimental Section:

4.5.1 Collection Identification and Cultivation

The sample was collected at 3-3.5 meters from the side of a rock shelf at 18°17'34.8"N 65°17'52.8"W at Punta de Maguey in Culebra, Puerto Rico on April 3, 2019 by snorkel and SCUBA diving. A sample was preserved in RNA later and a collection of CUM3APR19-1 was preserved in 1:1 seawater/2-proponol solution and frozen at -20°C until extraction. It was field ID as *Moorena producens* based on its morphology.

4.5.2 Extraction and Isolation

The collection CUM3APR19-1 was filtered by cheesecloth and extracted using DCM/MeOH (2:1) and warming (<30°C) with sonication for 30 minutes to give 12 grams of crude extract. This crude extract was further fractionated by drying a DCM/MeOH solution of the extract on celite under vacuum, and then Silica gel column chromatography (120 gram gold) on a NP Combiflash system, monitoring elution at wavelengths 254 nm and 214 nm. A linear gradient with (A) EtOAc and (B) hexanes were used for a 36 minute method. The gradient started with a 3 min isocratic step with 5% A followed by an increase to 20% A over 10 min, followed by an increase to 100% A over 18 min, and then held at 100% for 5 min and then moved to 50 % A in 1 min, and then held for 3 min.

The fractions containing curacin F from the NP Combiflash, as deduced by LCMS and ^{1}H NMR, were combined and purified through RP-HPLC on a Kinetex C18 semipreparative column (150 x 10.0 mm x 5µm Phenomenex), eluting at RT16.621-17.190. The flow rate was 2.500 mL/min with (A) H_2O and (B) acetonitrile and was monitored at a wavelength of 232 nm. It was initiated under isocratic conditions of 30% (B) for 5 minutes and then increased to 99% (B) until minute 25, it was held for 5 min at 99% (B) and then decreased to 30% (B) in 1 minute and then held at 30% (B) for one minute.

The fractions containing curacin G, as deduced by LCMS and ^{1}H NMR, were combined and further purified using RP-HPLC on Synergi 4µ hydro-RP column, 250 x 10.00 mm. The same method that was used for curacin F above was performed, and curacin G eluded at RT 19.986-20.571.

A small scale extraction was carried out on three samples of the CUM3APR19-1 biomass. A total of 654 mg of biomass in DCM:CD_3OD (2:1), 402 mg of biomass in DCM:CH_3OH (2:1), and 593 mg of biomass in DCM, were each extracted for 35 min followed

by 30 minutes of sonication. The mixtures were dried and afforded 19.0 mg, 41.1 mg, and 2.9 mg of crude extract ,respectively. LCMS analysis was performed for each of these extracts with an injection quantity of 10 microliters of a 1 mg/mL solution.

4.5.3 Cytotoxicity Assay

H460 human lung carcinoma cells were grown to confluence in monolayers and seeded at 6.66x103 cells/mL in wells with RPMI medium and FBS (180 μL each). They were incubated at 37°C and 5% CO for 24 hours. Sample preparation was done by dissolving in DMSO and diluted to final concentrations with less than 1% DMSO in RPMI medium. Aliquots of 20 μL of these solutions were added to the wells in duplicate such that the concentration of the sample was either 10 or 1 μg/mL. The plates were incubated for 48 hours and then stained with MTT for 25 minutes. The optical density was measured at 570 nm and 630 nm. The cell survival rates were calculated through the negative control of the RPMI medium.

4.5.4 General Experimental Procedures:

NMR data was acquired for curacin G using a Bruker Avance III 600 NMR with 1.7 mm dual tune TCI cyroprobe. Curacin F NMR data was acquired using a JEOL ECZ 500 NMR spectrometer with a 3 mm inverse detection probe. LC/MS data was recorded using a Thermo Finnigan LCQ Advantage Plus mass spectrometer with an Autosampler-Plus/LC-Pump-Plus/PDA-Plus system using a Phenomenex Kinetics C18 analytical column (100 x 4.6 mm x 5 μm). Reverse Phase HPLC was done with a Kinetex C18 semipreparative column (150 x 10.0 mm x 5 μm Phenomenex) or a Synergi 4μ Hydro-RP column 80A(250 x 10.00 mm) on a Thermo Fisher Scientific HPLC system with a Thermo Dionex UltiMate 3000 pump, RS autosampler, RS diode array detector, and a fraction collector.

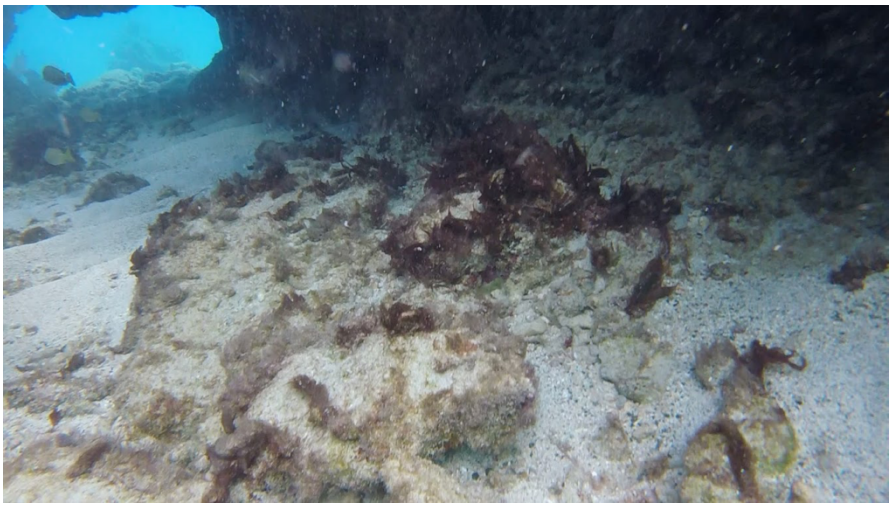


Figure 4.11: Underwater photograph of CUM3APR19-1 in Punta de Maguey in Culebra, Puerto Rico.

4.5.5 NMR Spectra

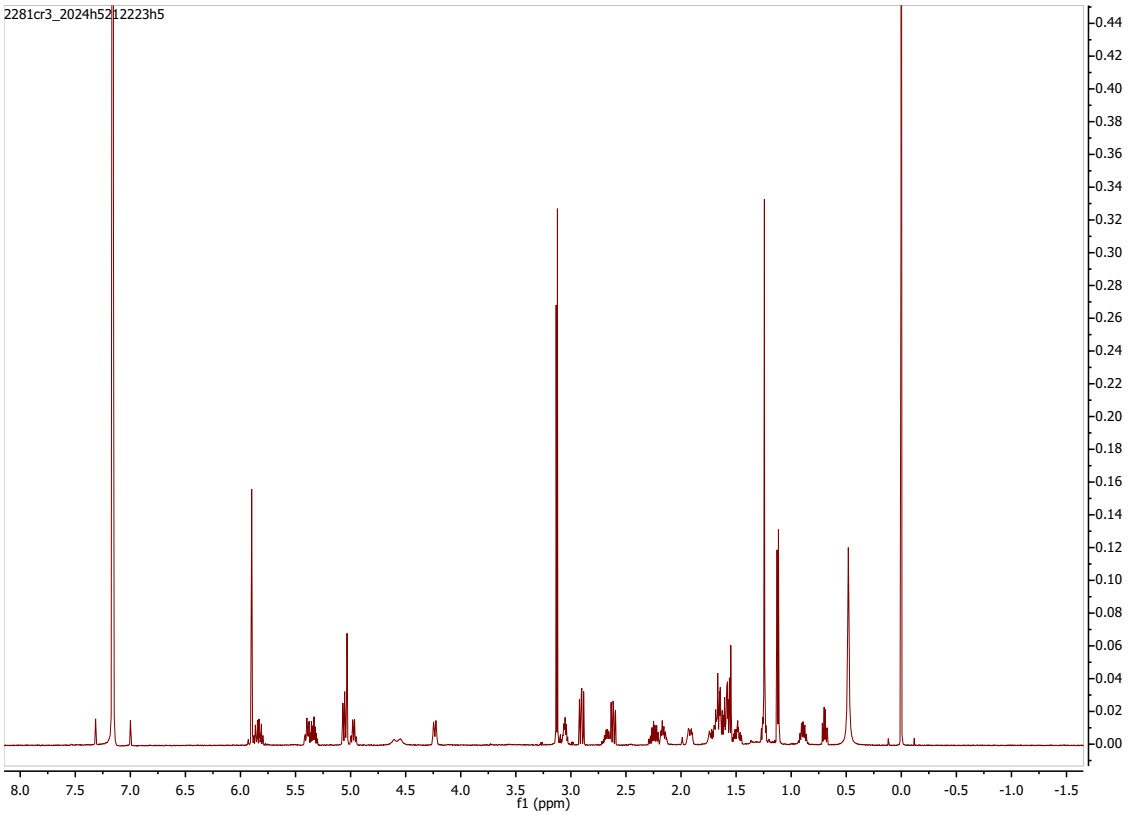


Figure 4.12: ^{1}H NMR of curacin F in C_6D_6 (500 MHz).

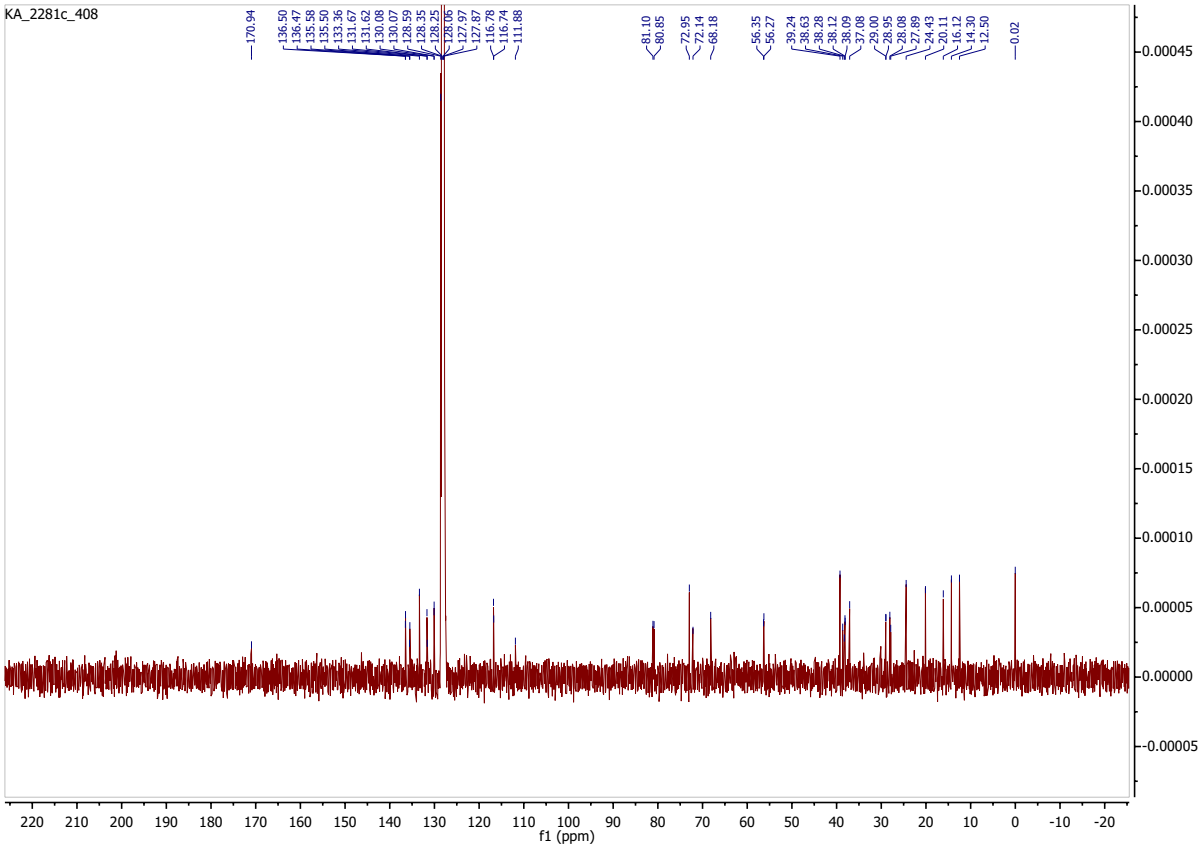


Figure 4.13: ^{13}C NMR of curacin F in C_6D_6.

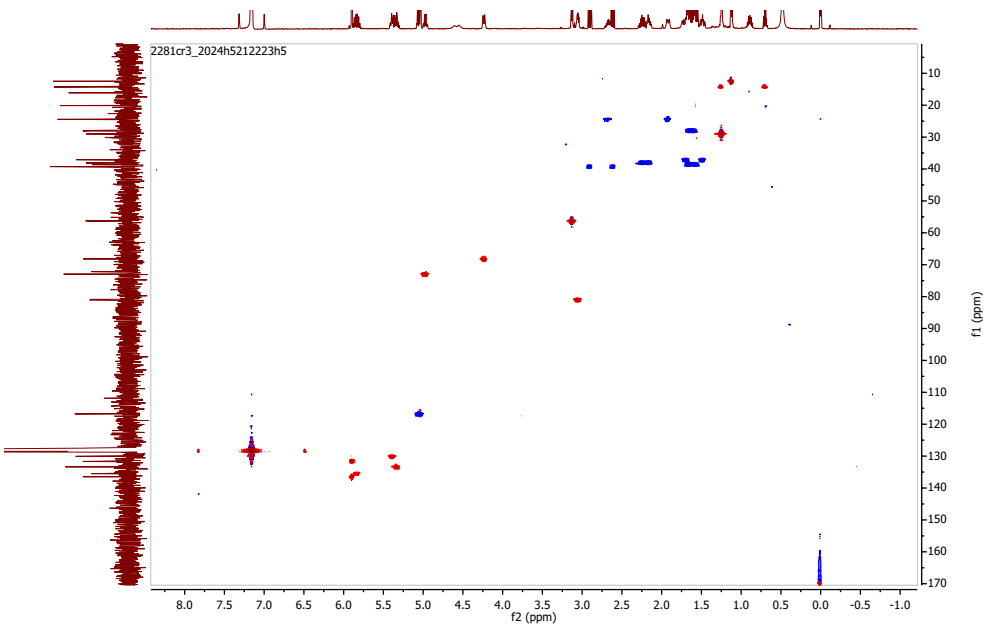


Figure 4.14: ^{1}H- ^{13}C HSQC spectrum of curacin F in C_6D_6.

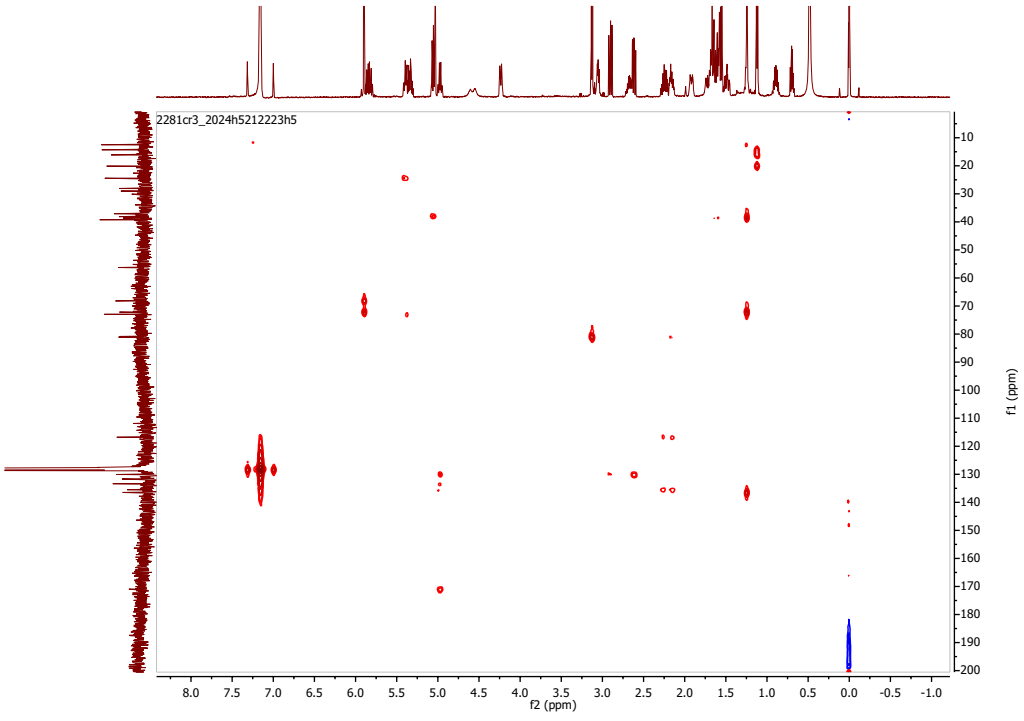


Figure 4.15: ^{1}H- ^{13}C HMBC of curacin F in C_6D_6.

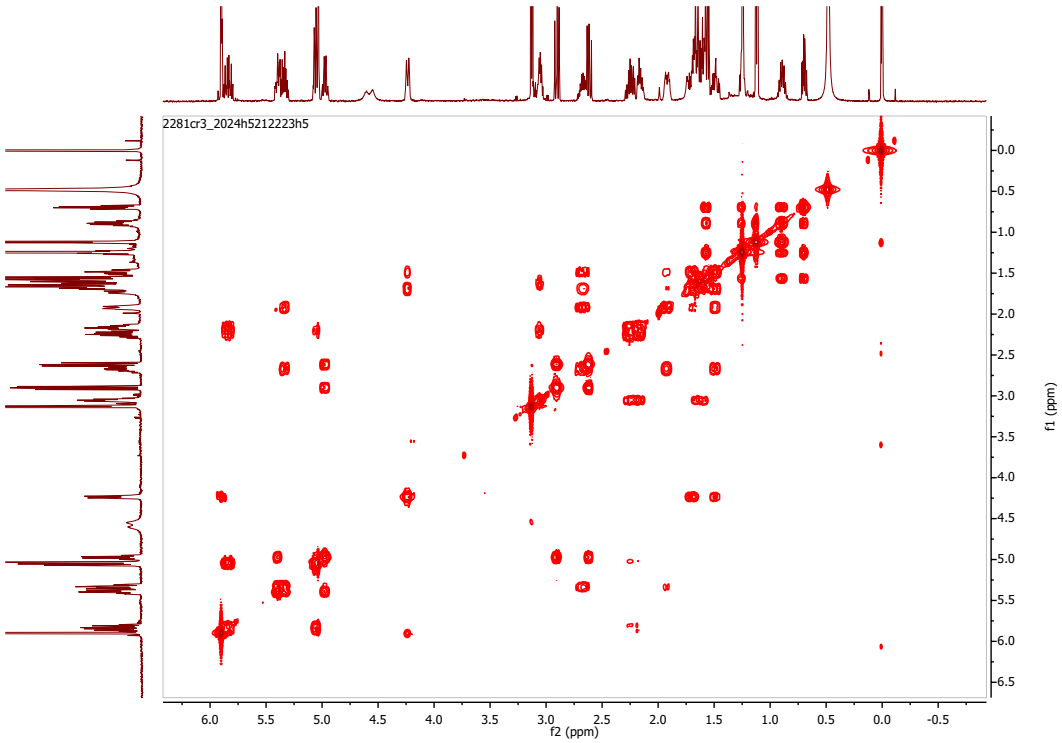


Figure 4.16: ^{1}H-^{1}H COSY of curacin F in C_6D_6.

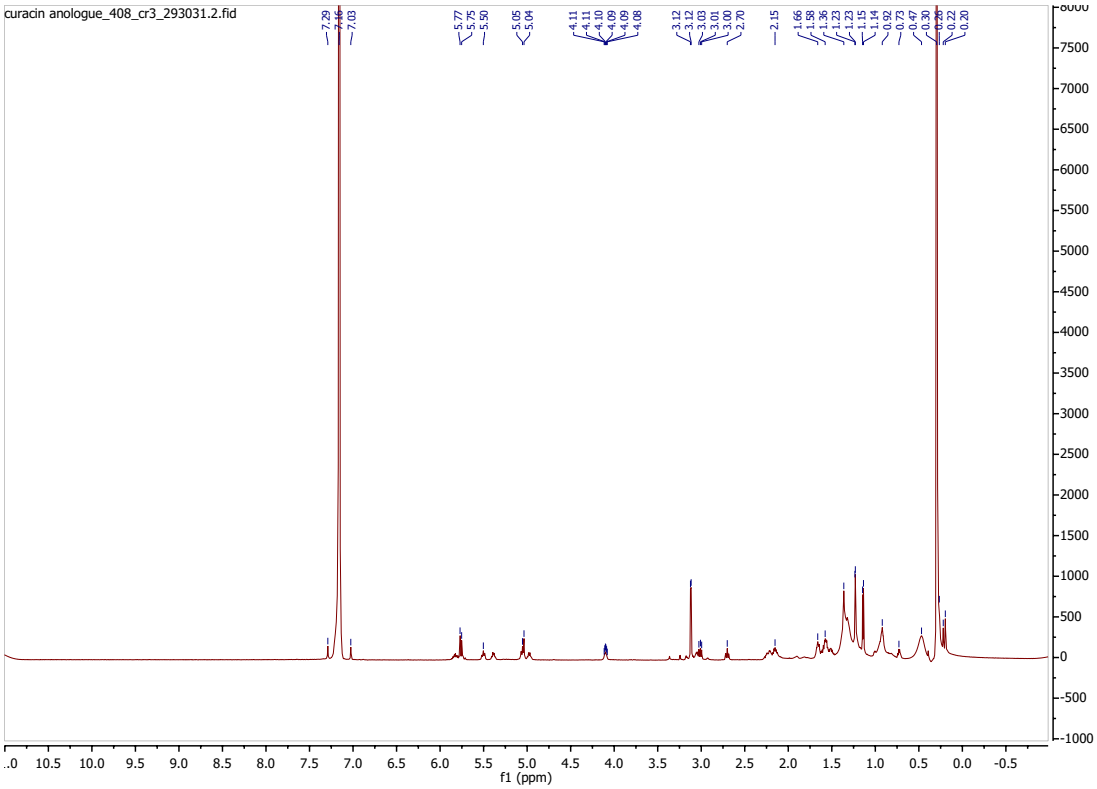


Figure 4.17: ^{1}H NMR spectrum of curacin G in C_6D_6.

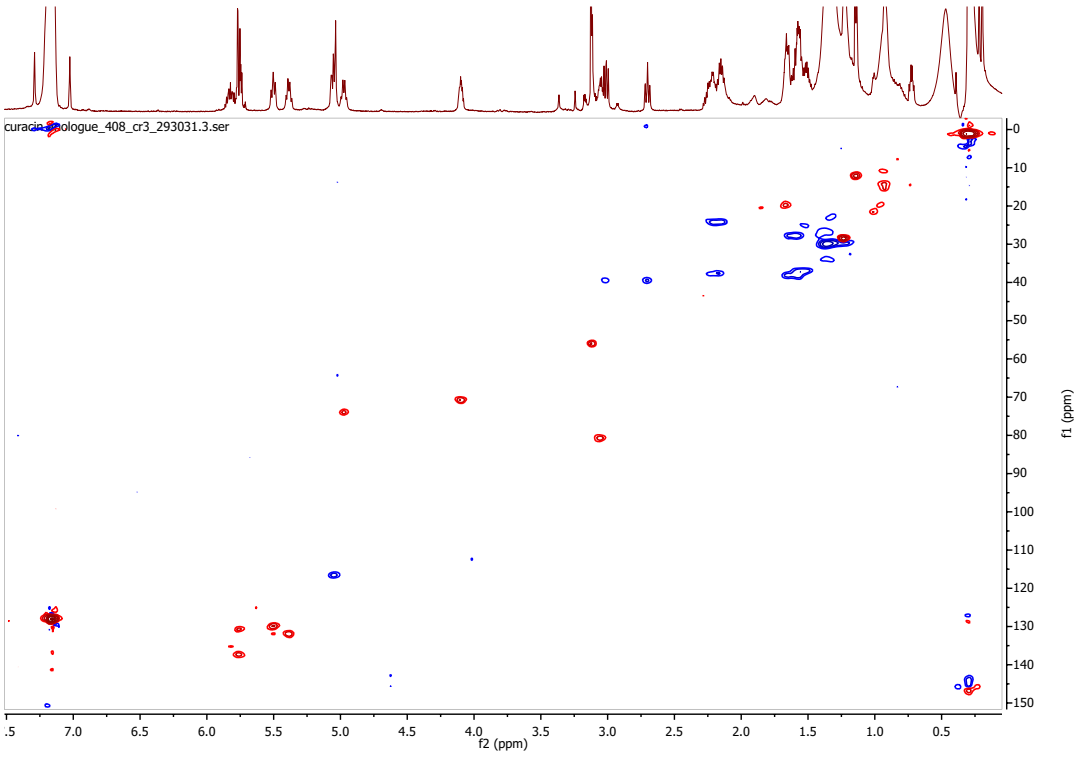


Figure 4.18: ^{1}H- ^{13}C HSQC of curacin F in C_6D_6.

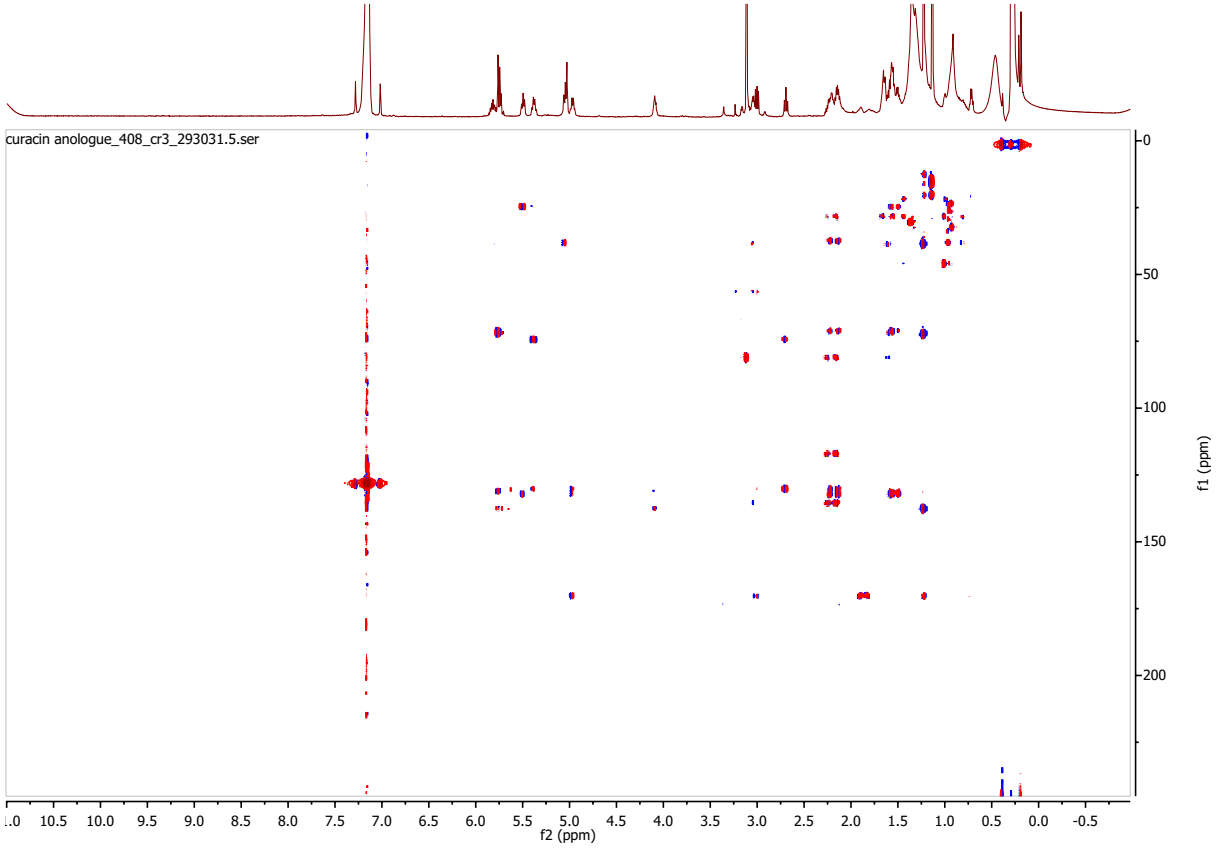


Figure 4.19: ^{1}H- ^{13}C HMBC of curacin G in C_6D_6.

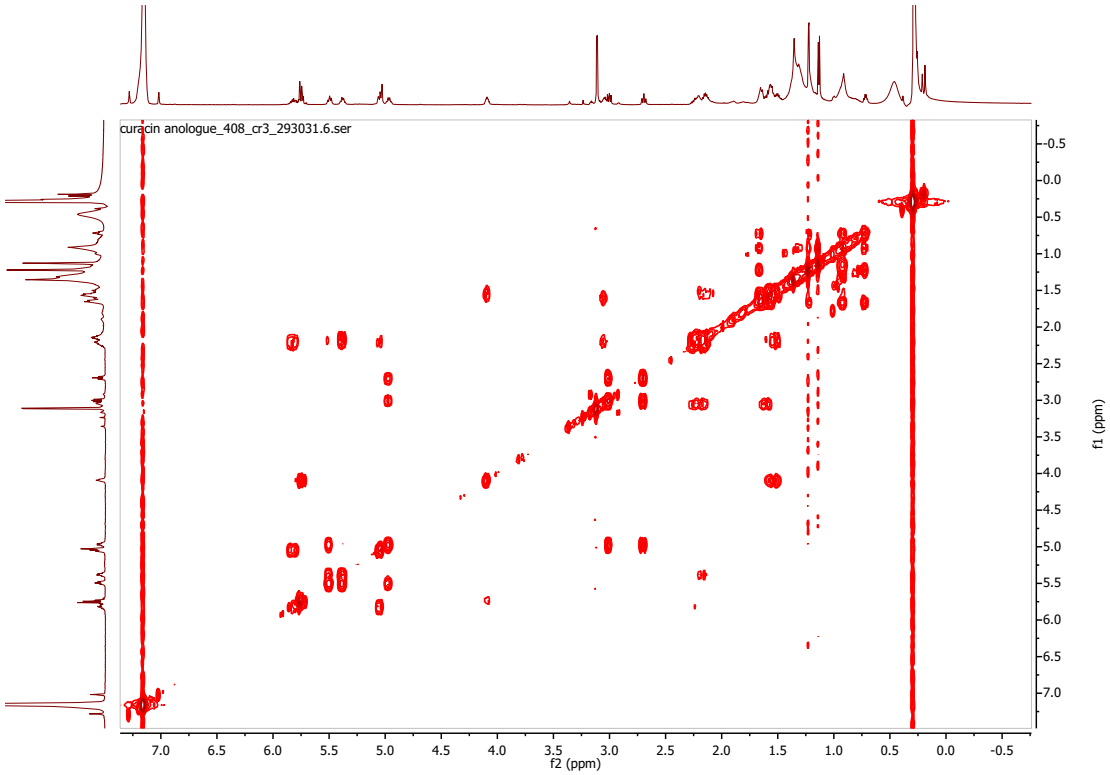


Figure 4.20: ^{1}H- ^{1}H COSY of curacin G in C_6D_6.

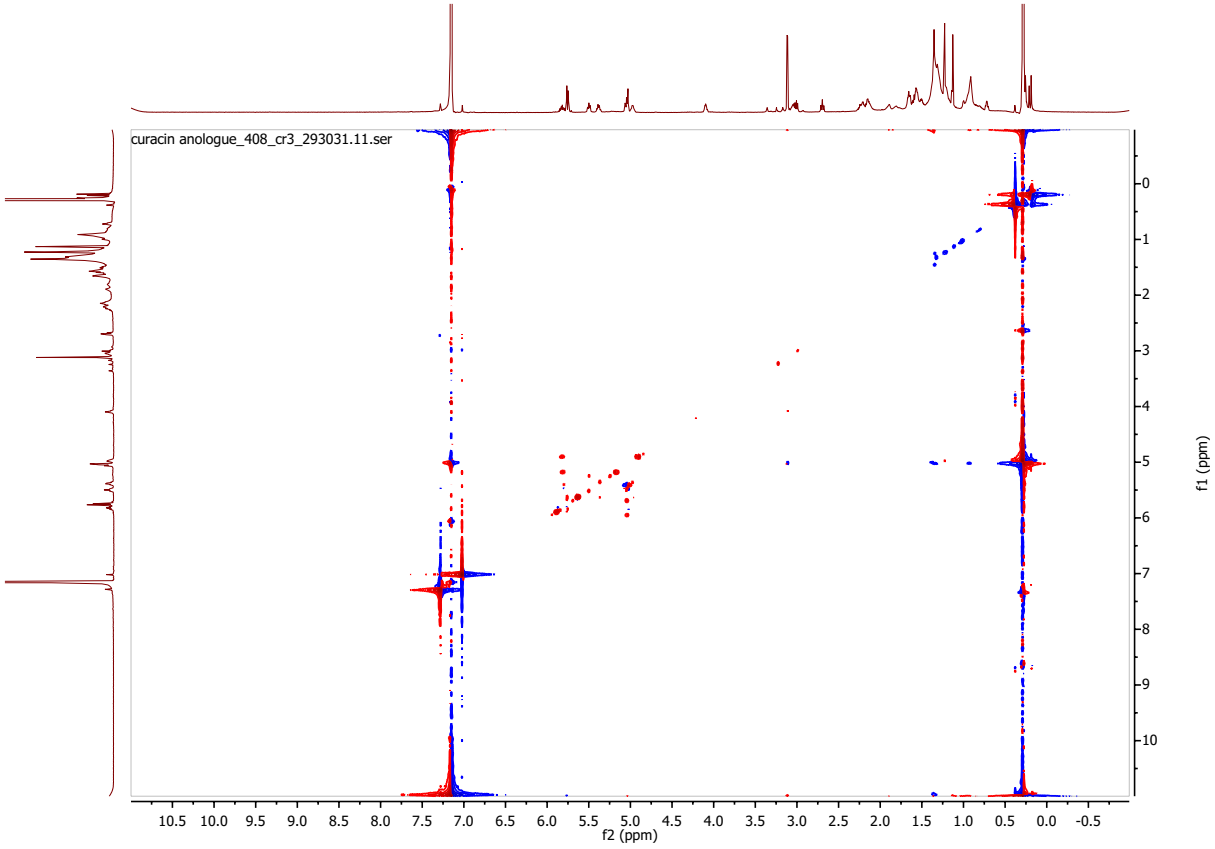


Figure 4.21: ^{1}H- ^{1}H NOESY of curacin G in C_6D_6.

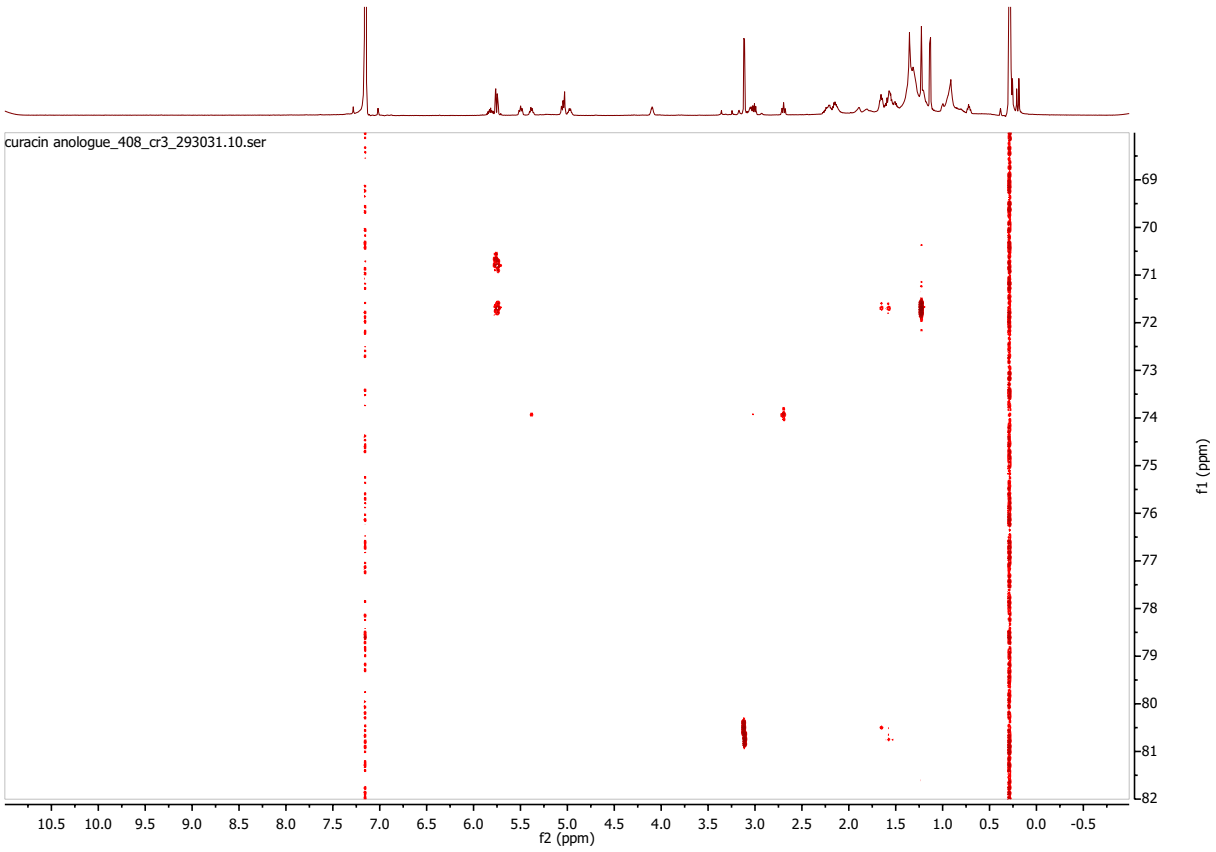


Figure 4.22: Band selective ^{1}H- ^{13}C HMBC of curacin G in C_6D_6.

4.6 Acknowledgements:

I thank Sociedad Ambiente Marino for their help in facilitating collections in Culebra, Puerto Rico.

Chapter 4, in full is currently being prepared for submission for publication of the material. Alexander, Kelsey L.; Reher, Raphael; Naman, Benjamin C.; Caro-Diaz Eduardo J.E.; Glukhov, Evgenia; Gerwick, Lena; Gerwick, William H. The dissertation author was the primary investigator and author of this material

4.7 References:

(1) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. Structure of Curacin A, a Novel Antimitotic, Antiproliferative and Brine Shrimp Toxic

Natural Product from the Marine Cyanobacterium Lyngbya Majuscula. *J. Org. Chem.* **1994**, *59* (6), 1243–1245. https://doi.org/10.1021/jo00085a006.

(2) Blokhin, Andrei V.; Yoo, Hye-Dong; Geralds, Robin S.; Nagle, Dale G.; Gerwick, William H.; Hamel, E. Characterization Cyanobacterial Site of Tubulin Analogues. *Mol. Pharmacol.* **1995**, *48*, 523–531.

(3) Chang, Z.; Sitachitta, N.; Rossi, J. V.; Roberts, M. A.; Flatt, P. M.; Jia, J.; Sherman, D. H.; Gerwick, W. H. Biosynthetic Pathway and Gene Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium Lyngbya Majuscula. *J. Nat. Prod.* **2004**, *67* (8), 1356–1367. https://doi.org/10.1021/np0499261.

(4) yoo, H. D.; Gerwick, W. H. Curacins b and c, New Antimitotic Natural Products from the Marine Cyanobacterium Lyngbya Majuscula. *J. Nat. Prod.* **1995**, *58* (12), 1961–1965. https://doi.org/10.1021/np50126a029.

(5) Márquez, B.; Verdier-Pinard, P.; Hamel, E.; Gerwick, W. H. Curacin D, an Antimitotic Agent from the Marine Cyanobacterium Lyngbya Majuscula. *Phytochemistry* **1998**, *49* (8), 2387–2389. https://doi.org/10.1016/S0031-9422(98)00365-3.

(6) Ueoka, R.; Hitora, Y.; Ito, A.; Yoshida, M.; Okada, S.; Takada, K.; Matsunaga, S. Curacin e from the Brittle Star Ophiocoma Scolopendrina. *J. Nat. Prod.* **2016**, *79* (10), 2754–2757. https://doi.org/10.1021/acs.jnatprod.6b00701.

(7) White, J. D.; Kim, T. S.; Nambu, M. Synthesis of Curacin A: A Powerful Antimitotic from the Cyanobacterium Lyngbya Majuscula. *J. Am. Chem. Soc.* **1995**, *117* (20), 5612–5613. https://doi.org/10.1021/ja00125a034.

(8) Wipf, P.; Reeves, J. T.; Balachandran, R.; Day, B. W. Synthesis and Biological Evaluation of Structurally Highly Modified Analogues of the Antimitotic Natural Product Curacin A. *J. Med. Chem.* **2002**, *45* (9), 1901–1917. https://doi.org/10.1021/jm0105171.

(9) Martin, B. K. D.; Mann, J.; Sageot, O. A. Synthesis of Analogues of the Marine Anti-Tumour Agent Curacin. *J. Chem. Soc. - Perkin Trans. 1* **1999**, *8* (17), 2455–2460. https://doi.org/10.1039/a904378k.

(10) Nishikawa, A.; Shirai, R.; Koiso, Y.; Hashimoto, Y.; Iwasaki, S. Design and Synthesis of Curacin a Analogs with Varied Side Chain Structures. *Bioorganic Med. Chem. Lett.* **1997**, *7* (20), 2657–2660. https://doi.org/10.1016/S0960-894X(97)10055-5.

(11) Curren, E.; Leong, S. C. Y. Global Phylogeography of Toxic Cyanobacteria Moorea Producens Reveals Distinct Genetic Partitioning Influenced by Proterozoic Glacial Cycles. *Harmful Algae* **2019**, *86* (May), 10–19. https://doi.org/10.1016/j.hal.2019.04.010.

(12) Leao, T.; Castelão, G.; Korobeynikov, A.; Monroe, E. A.; Podell, S.; Glukhov, E.; Allen, E. E.; Gerwick, W. H.; Gerwick, L. Comparative Genomics Uncovers the Prolific and Distinctive Metabolic Potential of the Cyanobacterial Genus Moorea. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (12), 3198–3203. https://doi.org/10.1073/pnas.1618556114.

Chapter Five: Conclusion

The oceans offer a vast and complex chemical space for the discovery of new therapeutics. Cyanobacteria are known to be unique and prolific producers of compounds with potential for bioactivity. This work has focused on the isolation and characterization of metabolites from two different species of marine cyanobacteria; *Moorena producens* and *Leptolyngbya sp.*

A culture of *Leptolyngbya sp.* cyanobacteria from American Samoa was initially studied for its cytotoxicity to lung NCI-H460 cells. Fatuamide A was isolated and characterized from the most cytotoxic fraction. It was found that it has the ability to bind to metals and therefore is an ionophore.

Curacin F and G were new compounds isolated from a collection of *Moorena producens* from Puerto Rico. They were prioritized because they initially networked by LC-MS/MS with curacin A, which is a known antimitotic agent. Their planar structures were determined through NMR and mass spectrometry, and shared many similarities with the structure of curacin A, but had two additional hydroxy groups and lacked the diene functionality.

NMR was an integral part to the characterization of these compounds. In this work, we studied the use of rapid NMR approaches for natural products. The combination of two different techniques, NUS (nonuniform sampling) and ASAP (acceleration by sharing adjacent polarization), were applied in combination to significantly shorten the acquisition time. We found the optimal leveling of sampling in these accelerated experiments as a function of concentration of a standard and the natural product fatuamide A.

This work highlights the interconnection between different tools and fields of study within natural products research. The structure of fatuamide A was determined by studying it

through multiple orthogonal approaches. The NMR studies allowed for the construction of all of the partial structures comprising fatuamide A. The MS/MS and MS^3 data enabled for the partial structures to be put in the correct order for the completion of the final structure. This structure matched with the natural product that would be produced by a corresponding biosynthetic gene cluster deduced from the draft sequenced genome. Moreover, an analysis of the gene cluster revealed biosynthetic genes that suggested that fatuamide A was a siderophore. The biosynthetic gene cluster was also used to make predictions on the stereochemistry of fatuamide A. It was a combination of all these techniques, NMR, mass spectrometry, and genomic analysis, that allowed for the final proposed structure of fatuamide A. This interdisciplinary study within natural products research illustrates the potential to unlock new discoveries and potentially new therapeutic scaffolds.

There are more analogues of curacin A that were identified in the LCMS/MS molecular network that could be further characterized. These would be of particular interest because of the antimitotic nature of curacin A. It would also be interesting to study the biosynthetic gene cluster and see if the differences between curacin A and these new analogues can be explained by differences in the encoded enzymes.

The study of the methyl group stereochemistry of fatuamide A highlights the need for more research to be done on the stereochemistry of methyl groups that are installed by C-methyl transferases (cMT) in biosynthetic pathways. There are multiple examples in the literature where these methyl groups in cyanobacterial natural products remain uncategorized. Deciphering how these enzymes control the stereochemistry of methyl group additions from S-adenosylmethionine (SAM) would help not only give absolute configurations for compounds in the future, but for compounds that have already been published as well.

Natural products research can be faced with major bottlenecks in terms of discovery and characterization of new compounds. This work highlights a few different ways in which this can be alleviated. The first involves the use of techniques in NMR spectroscopy that can reduce the acquisition time. While 2D NMR data can typically take longer to acquire, it is valuable for gaining information about the connectivity of atoms in a molecule. Being able to gain this data in rapid matter can allow for faster dereplication and structure elucidation, and therefore reduce some of the bottlenecks in natural products drug discovery. A second way that this work highlights a more rapid approach to natural products discovery is through the integration of genomic, mass spectrometric, and NMR data in combination to arrive at a final proposed structure.

Studies in marine natural products allow for us to dive into the diverse chemical space that is present in our oceans. The ocean continues to be a source of new compounds that have potential for use on land for a variety of purposes, including human health. Not only studying the different compounds produced, but studying the systems in which they are made, can potentially help solve our need for different types of therapeutics.

Natural products are of critical importance for human health and have either directly contributed or inspired a majority of the therapeutics that are on the market today. The ocean remains an incredible source of these diverse natural products. It is important that we maintain this unique biodiversity so that these ecosystems remain intact. This way we can continue to learn from the chemistry the ocean utilizes and implement it into our own development of medicine. It is important now more than ever, to take measures to help preserve these environments, so we can not only enjoy them, but study them for generations to come.

Maintaining and studying the oceans is an investment in our future health, as ocean health is intimately tied to human health.

Appendix: Cyclic (Alkyl)(Amino)Carbene (CAAC) Gold(I) Complexes as Chemotherapeutic Agents

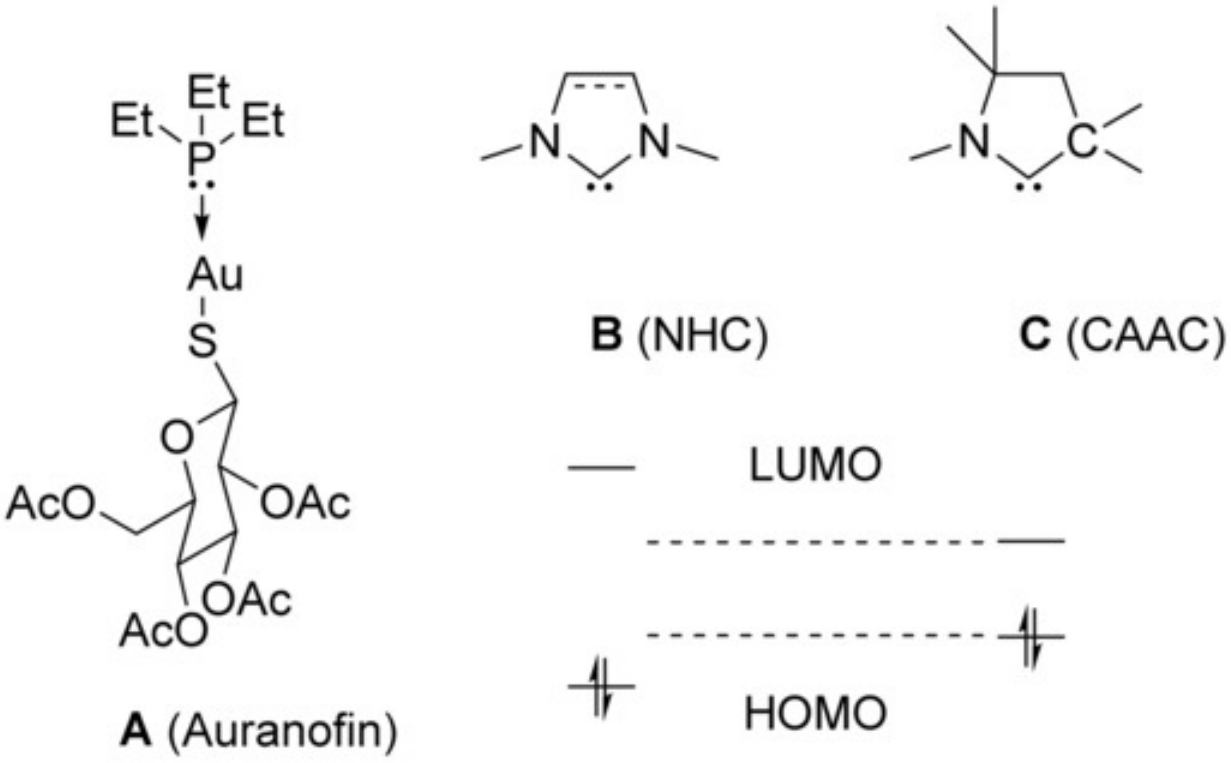
A.1: Abstract:

Cyclic (Alkyl)(Amino)Carbenes (CAACs) have become forceful ligands for gold due to their ability to form very strong ligand-metal bonds. Inspired by the success of Auranofin and other gold complexes as antitumor agents, we have studied the cytotoxicity of bis- and mono-CAAC-gold complexes on different cancer cell lines: HeLa (cervical cancer), A549 (lung cancer), HT1080 (fibrosarcoma) and Caov-3 (ovarian cancer). Further investigations aimed at elucidating their mechanism of action are described. This includes quantification of affinities for TrxR, evaluation of their bioavailability and determination of associated cell death process. Moreover, Transmission Electron Microscopy (TEM) was used to study morphological changes upon exposure. Noticeably, a significant reduction in non-specific binding to serum proteins was observed with CAAC complexes when compared to Auranofin. These results confirm the potential of CAAC-gold complexes in biological environments, which may result in more specific drug-target interactions and decreased side effects.

A.2 Introduction:

Many civilizations, starting with ancient China, employed gold containing concoctions, long touted for their healing powers in the fight against disease.[1] In the late 19th century, scientific study began with the examination of gold(I) complexes against pulmonary tuberculosis,[2] and then later against rheumatoid arthritis.[3] The latter became a successful endeavor with a clear demonstration of the beneficial effect of gold-thiolate drugs. Introduced in the late 1970s, Auranofin (**A**, Scheme **1**) proved very promising and was approved for clinical use in 1985.[3a] The antitumor activity of Auranofin was reported for the first time in 1979.[4] Since

that time, numerous gold(I) complexes supported by phosphorous-, nitrogen- or carbon-based ligands have been investigated.[5] Although it is still not fully understood, the proposed mechanism is associated with the ability of Auranofin and other gold-based drugs to bind a selenocysteine moiety located in the C-terminal active site of the Thioredoxin Reductase (TrxR) enzyme,[6] which is involved in controlling reactive oxygen species (ROS) homeostasis.[7] It should be noted that the 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-gluco-pyranosyl-1'-thiolate (Glc) moiety was selected due to the over-expression of the corresponding glucose transporters on cancerous cells, resulting in a higher uptake of the gold complex compared to versions lacking this anionic carbohydrate ligand.[8]



Scheme A.1: Auranofin structure (**A**); frontier orbitals comparison between NHCs (**B**) and CAACs (**C**).

Despite its potency, Auranofin suffers from significant side effects due to its undesired binding to cysteine residues on serum proteins and intracellular proteins.[9] A step forward in the design of gold(I) drugs was achieved with the replacement of the phosphine ligand by *N*-Heterocyclic Carbenes (NHCs) **B**,[10, 11] which form stronger ligand-metal bonds. Similarly to Auranofin, neutral complexes of type Au(**B**)Cl mainly exhibit cytotoxicity to their ability to inhibit TrxR overexpressed in cancer cells. Alternatively, cationic bis-carbene complexes of type $Au(\mathbf{B})^{2+}X^{-}$ show higher stability under physiological conditions than neutral complexes and are characterized by a better ability to navigate through membrane with elevated potential, to accumulate in the mitochondria, to interact with membranal vital systems and also to induce deregulation and/or alterations of elements essential for cell life.[5d-5f]

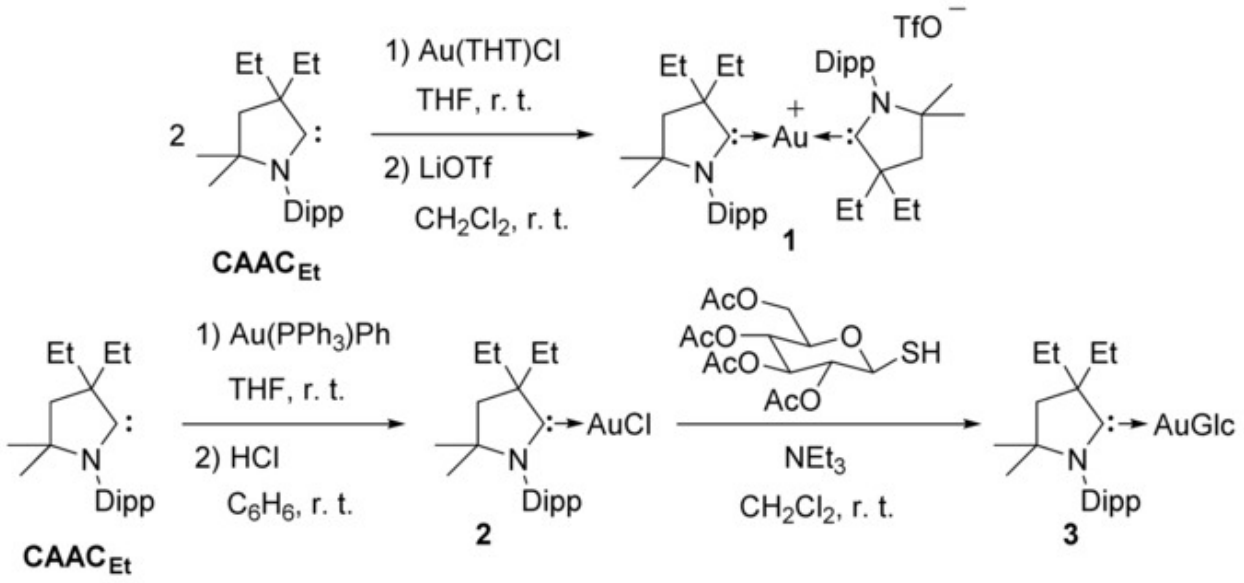
More recently, another family of carbenes, namely cyclic (alkyl)(amino) carbenes (CAACs)[12, 13] **C** have found many applications, due to their peculiar steric and electronic properties. In-deed, when compared to NHCs, the replacement of one nitrogen atom (σ-attractor and π-donor) by a quaternary carbon (solely σ-donor) results in a higher HOMO (carbene lone pair) and lower LUMO (formal empty orbital) (Scheme **1**).[14] Therefore, CAACs are simultaneously amongst the most basic and the most π-acidic singlet carbenes. The advantages of this class of carbene over NHCs are best exemplified by success in the isolation of active catalytic intermediates,[15] efficient and robust pre-catalysts,[16] and complexes with unusual configuration and/or oxidation state.[17] In all cases the stronger carbene-metal bond is characteristic of these CAAC bound complexes with respect to NHC analogs. Therefore, there is reason to believe that CAAC-bound gold complexes would exhibit lower affinity for cellular proteins compared to desired biological targets, leading to more specific drug-target interactions and decreased side effects. Recently, different CAAC-bound gold, silver and copper complexes

have been reported as potent cytotoxic agents. Further, their mechanism of action appears to be metal dependent.[18] Herein, we report a detailed investigation on the in vitro antitumoral activity of a set of gold(I) complexes (vide infra), all bearing CAAC ancillary ligands for better comparison between them, namely [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1**, Au(**CAAC$_{Et}$**)Cl **2**[17d] and Au(**CAAC$_{Et}$**)Glc **3**. We have also evaluated the ability of these complexes to inhibit TrxR, measured their affinity for bovine serum albumin (BSA) and provide evidence for the mechanism of cell death.

A.3 Results and Discussion:

First, we reacted free carbene **CAAC$_{Et}$** with a half-equivalent of Au(THT)Cl in tetrahydrofuran. Subsequent anion exchange by treating the cationic complex with lithium trifluoromethanesulfonate in methylene chloride afforded the homoleptic salt **1** in 91 % yield (Scheme 2, top). The neutral complex **2** is prepared by a two-step sequence in which **CAAC$_{Et}$** is first reacted with Au(PPh$_3$)Ph to substitute the PPh$_3$ ligand. Then protolysis of the resulting Au(**CAAC$_{Et}$**)Ph with dry HCl in benzene allowed for the isolation of the desired complex **2** in an overall yield of 87 % (Scheme 2, bottom center). Note that both compounds **1** and **2** can be synthesized by bypassing the isolation of free carbene **CAAC$_{Et}$**. In the former case, addition of one equivalent of KHMDS to the mixture of the cyclic iminium **CAAC$_{Et}$**H$^+$TfO$^-$ and Au(THT)Cl also led to complex **1**, albeit with overall lower yields of 62 %. Similarly, addition of one equivalent of KHMDS to the mixture of the cyclic iminium **CAAC$_{Et}$**H$^+$TfO$^-$ and Au(PPh$_3$)Ph and further reaction with HCl afforded complex **2** with an 82 % yield. Finally, reaction of complex **2** with the gluco-pyranosyl-1-thiol in the presence of triethylamine in methylene chloride led to the expected CAAC-bound Auranofin analog **3** in good yield (94 %, Scheme 2, bottom right). Each of these complexes were found to be indefinitely air and moisture

stable, and were fully characterized by NMR, HR-MS and X-ray diffraction of single crystals (Figure 1).



Scheme A.2: Syntheses of complexes **1** (top), **2** (bottom center) and **3** (bottom right).

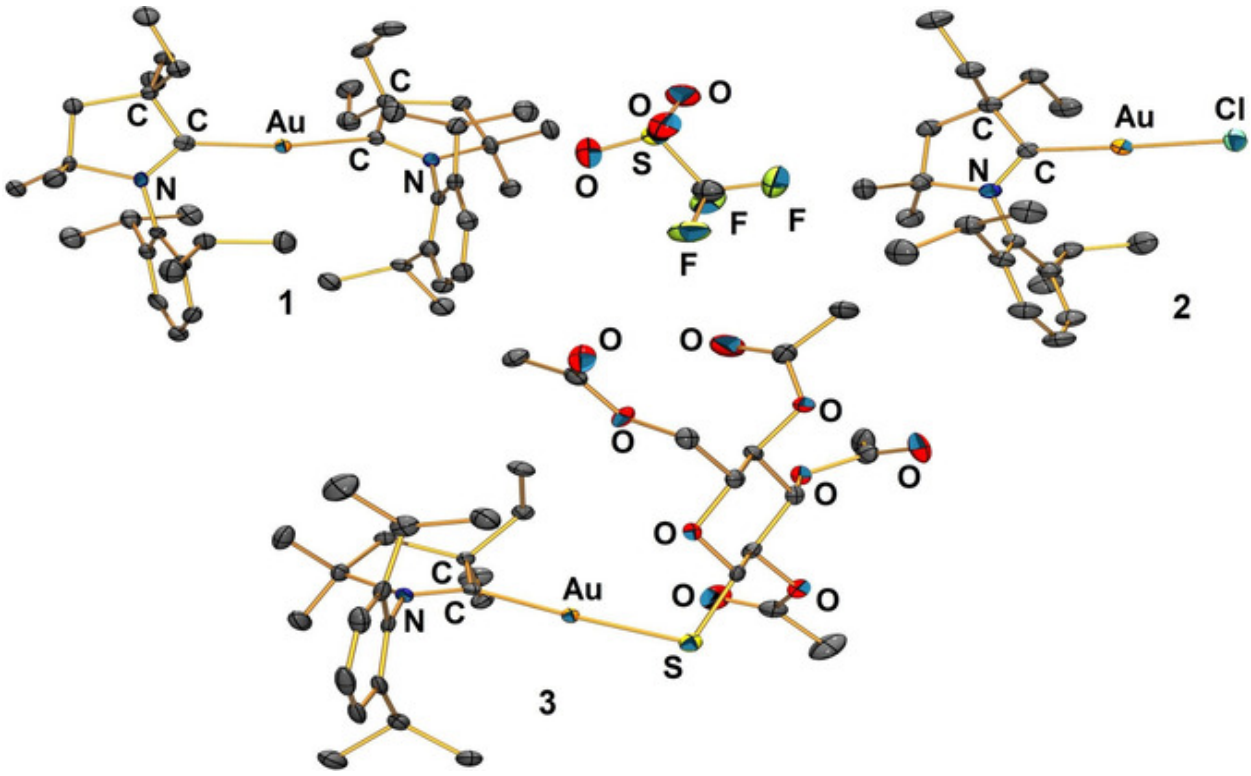


Figure A.1: Solid-state structures of complexes **1**, **2** and **3**. Ortep views are shown with ellipsoids at 50 % probability. H atoms, solvents are omitted for clarity.

As for any CAAC supported divalent gold(I) complex,[19] we expected a linear geometry at the metal center. Analysis of the solid-state structures[20] confirmed this spatial arrangement with C-Au-C/X angles of 170.56° (13), 178.05° (7) and 176.47° (10) for compounds **1**, **2** and **3**, respectively. The slightly bent geometry in the homoleptic complex **1** is unexpected with regards to previously reported $[Au(\mathbf{CAAC})_2]^+$ fragments in which the carbene-Au-carbene angle is found to be 180°. However, this bent geometry is commonly observed in bis-CAAC complexes of Ni^0,[21a] Pd^0[21b] and Pt^0,[21b] which are isolobal to Au^I, and exhibit carbene-metal-carbene angles of 169° ±4°. More importantly, in complex **1**, the trifluoromethanesulfonate anion is located more than 5 Å away from the gold atom and therefore does not engage in any interaction with the metal center. Finally, carbene-gold bond lengths remain relatively constant (**1**: 2.036 (3) Å, **2**: 1.975 (3) Å, **3**: 2.004 (4) Å). Note that the minor shortening of the carbene bond length for the gold-chloride complex has already been observed for analogous compounds.[19]

With complexes **1**–**3** and the $\mathbf{CAAC_{Et}}H^+TfO^-$ salt in hand, we tested their cytotoxicity against three different human cancer cell lines: HeLa (cervical cancer), A549 (lung cancer), HT1080 (fibrosarcoma) and Caov-3 (ovarian cancer). A cell titer blue assay was used to determine the extent of the inhibitory effect of the complexes on cancer cell growth. Cytotoxic activities were demonstrated to be compound and cell line dependent: against HeLa cells, activities decreased following the order **1**≈**2**>**A**>**3**≫$\mathbf{CAAC_{Et}}H^+TfO^-$. whereas, against A549 and HT1080, the order **1**>**A**>**3**≈**2**≫$\mathbf{CAAC_{Et}}H^+TfO^-$ was observed. Finally, the activities followed **1**>**A**>**3**>**2**≫$\mathbf{CAAC_{Et}}H^+TfO^-$ when exposed to Caov-3. Overall, Auranofin **A**, Au($\mathbf{CAAC_{Et}}$)Cl **2** and Au($\mathbf{CAAC_{Et}}$)Glc **3** triggered death on all cell lines with IC50 values in the

micromolar range (Table 1). Interestingly, [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1** proved to be more cytotoxic than the others, with IC50 values lower by 1 to 2 orders of magnitude across all cell lines.

Table A.1: Cytotoxicity (IC50) against cancer cell lines (in μM).[a]

Compound	HeLa	A549	HT1080	Caov-3
Auranofin **A**	1.7±0.2	3.7±0.6	1.4±0.9	0.9±0.3
[Au(**CAAC$_{Et}$**)$_2$]$^+$OTf$^-$ **1**	0.3±0.2	0.07±0.06	0.14±0.04	0.3±0.2
Au(**CAAC$_{Et}$**)Cl **2**	0.6±0.2	4.5±0.7	4.4±1.3	3.9±0.8
Au(**CAAC$_{Et}$**)Glc **3**	2.7±0.1	6.6±2.5	3.1±1.8	1.9±0.4
CAAC$_{Et}$H$^+$TfO$^-$	60±13	76±34	26±16	>50

[a] IC50 values represent the concentration that caused 50 % cell death reported. These values were estimated by interpolation in the sigmoidal dose response fitted curve and are the results of at least three independent experiments.

This trend has previously been observed for NHC supported gold(I) complexes and is linked to the ability of delocalized lipophilic cations to pass through hydrophobic barriers and mitochondrial membranes to reach their targets in the mitochondria.[22] Auranofin (**A**) was observed to be slightly more active than the **CAAC$_{Et}$** analog **3** Au(**CAAC$_{Et}$**)Glc in all cell lines. Remarkably, compound **2** exhibited higher cytotoxicity than Auranofin only on HeLa cells, suggesting a cell dependent mechanism of action. Furthermore, the ligand

precursor **CAAC$_{Et}$**H$^+$OTf$^-$ was found to be at least 5- to 1000-fold less toxic than the complexes against this set of cancer cell lines. The significantly lower cytotoxicity of **CAAC$_{Et}$**H$^+$OTf$^-$, clearly demonstrates the gold center involvement in the antitumor activity. In general, the gold complexes bearing **CAAC$_{Et}$** as a mono- or bis- carrier ligand were demonstrably more active than those reported with bulkier CAAC-ligands,[18] suggesting that steric parameters and the hydrophilicity/lipophilicity balance are key features impacting activity.

Thus, as part of an investigation on the physiochemical properties of the CAAC-gold complexes, we determined their lipophilicity using the octanol-water partition protocol. Based on our experimental results, the most lipophilic compound was Au(**CAAC$_{Et}$**)Cl **2**, followed by Au(**CAAC$_{Et}$**)Glc **3**. As expected, because of its positively charged nature [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1** showed the lowest lipophilicity value (Table **2**). Indeed, it has been reported that the lipophilicity of compounds can be finely tuned to increase their ability to cross cellular membranes and accumulate in mitochondria, therefore modifying their toxicity.[6, 23] To examine the impact of lipophilicity on cytotoxicity, we studied the cellular uptake of [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1**, Au(**CAAC$_{Et}$**)Cl **2** and Au(**CAAC$_{Et}$**)Glc **3** in Caov-3 at 1 μM concentration after 6 h incubation. Our results showed a 10-fold greater uptake for the cationic complex [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1** when compared to the neutral complexes Au(**CAAC$_{Et}$**)Cl **2** and Au(**CAAC$_{Et}$**)Glc **3**. Interestingly, complex Au(**CAAC$_{Et}$**)Cl **2**, which is the most lipophilic compound, showed the lowest uptake and cytotoxicity values on Caov-3 cells. Among these CAAC-complexes, lower lipophilicity is associated with higher cellular uptake and higher cytotoxicity.

Table A.2: LogP values, Cellular Uptake, TrxR Inhibition and % BSA binding.

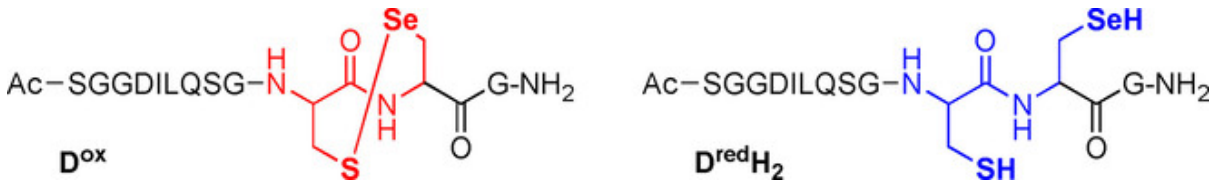
Compound	**Log*P***	**Uptake (ng Au/mg protein)[a]**	**TrxR (IC_{50}, nM)[b]**	**BSA binding (%)[c]**
Auranofin **A**	n.d.[d]	n.d.[d]	15.6±4.7	79±6
$[Au(\mathbf{CAAC_{Et}})_2]^+OTf^-$ **1**	1.4±0.1	276±115	>100 000	17±21[c]
$Au(\mathbf{CAAC_{Et}})Cl$ **2**	2.4±0.1	28±11	819±145	55±9
$Au(\mathbf{CAAC_{Et}})Glc$ **3**	2.2±0.2	36±8	2226±826	50±9
$\mathbf{CAAC_{Et}}H^+TfO^-$	n.d.[d]	n.d.[d]	>200 000	n.d.[d]

[a] Uptake in Caov-3 cells after 6 h incubation with 1 μM of the respective Au-complex. [b] IC50 values represent the concentration that caused 50 % cell death reported. [c] BSA binding was estimated as the percentage of free drug recovered in the presence and absence of protein. [d] not determined.

Based on the accepted mode of action for Auranofin A and other gold complexes, we opted to explore affinity towards TrxR, a known target for gold complexes. In parallel, we investigated non-specific interactions between gold complexes and serum proteins, using bovine serum albumin as a model. Compounds **A**, **1**, **2** and **3** were incubated with TrxR and bovine serum albumin and results are summarized in Table **2**.

Under our protocol and in accordance with previously published literature,[24] Auranofin **A** showed the highest TrxR inhibitory effect at low nanomolar concentrations. Complex **2** showed increased activity over **3**, with both exhibiting TrxR-IC50 values in the micromolar ranges. Note that previously reported, analogous NHC and CAAC complexes show TrxR inhibition at similar concentration ranges.[25] Interestingly, despite displaying the highest cytotoxicity against all cancer cell lines studied here, the cationic complex [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1** did not inhibit TrxR up to its solubility limit (100 μM). Moreover, as expected, **CAAC$_{Et}$**H$^+$TfO$^-$ bearing no gold metal center, was demonstrated to have no affinity for TrxR up to its solubility limit.

To further understand TrxR inhibition by the gold complexes, we chose to study the binding of gold complexes **1**–**3** with dodecapeptide **D^{ox}** (Ac-SGGDILQSG**C**U**U**G-NH2, Scheme **3**)[6b] which is analogous to the peptide sequence found in TrxR that is the specific target of gold complexes. Note that in its oxidized form, **D^{ox}** corresponds to the non-active form of the TrxR enzyme. Peptide **D^{ox}** was prepared according to literature procedures and was converted into **D^{red}**H$_2$ (Ac-SGGDILQSG**C$_H$U$_H$**G-NH) upon treatment with tris(2-carboxyethyl)phosphine (TCEP). Under its reduced form, **D^{red}**H$_2$, as well as analogous deprotonated forms of the reduced peptide (**D^{red}**H$^-$ and **D^{red}**.$^{2-}$), correspond to the active forms of the TrxR. To investigate the interactions between the gold complexes and peptides of type **D**, we incubated freshly reduced peptide **D^{red}**H$_2$/**D^{red}**H$^-$/**D^{red}**.$^{2-}$ with gold complexes **1**–**3** and analyzed the results by mass spectrometry. As expected, complexes **2** and **3** form anionic adducts [**D^{red}**-Au(**CAAC$_{Et}$**)]$^-$ in which the reduced di-anionic peptide **D^{red}**.$^{2-}$ is bound to the [CAAC-Au]$^+$ cationic fragment resulting from the formal substitution of the anionic ligand (X=Cl$^-$ for **2** or X=Glc$^-$ for **3**).



Scheme A.3: Dodecapeptide **D**, analogous to TrxR active sequence, in its oxidized ($\mathbf{D^{ox}}$) and reduced ($\mathbf{D^{red}H_2}$) forms.

Interestingly, incubation of the cationic complex **1** under the same conditions as above did not allow for observation of the corresponding adduct [$\mathbf{D^{red}}$-Au($\mathbf{CAAC_{Et}}$)$_2$]$^-$; thus, explaining the lack of inhibition of TrxR detailed in Table 2.

As stated, one of our goals was to overcome concerns regarding the undesired affinity of Auranofin to albumin. Here, all complexes showed reduced binding compared to Auranofin **A** (See Table 2). Moreover, compared to similar NHC supported gold complexes,[26] the use of CAAC ligands lead to decreased non-specific binding from close to 100 %[27] to 55 %. We believe the lower affinity of CAAC-supported complexes **1**, **2** and **3** to BSA compared to **A** and analogous NHC complexes is a result of the stronger ligand-metal bond, and thus the greater stability of complexes **1**, **2** and **3**.

Several important pieces of information arise from these data. First, it appears that the cytotoxicity of complex **1** results without targeting the TrxR enzyme, hence causing cell death through a mechanism different than for complexes **A**, **2** and **3**. Also, even though the lack of interaction between [Au($\mathbf{CAAC_{Et}}$)$_2$]$^+$TfO$^-$ **1** and the selenocysteine moiety is not surprising, it does suggest that the CAAC ligand is not labile under these conditions, and that the fragment [Au($\mathbf{CAAC_{Et}}$)$_2$]$^+$ retains its structural integrity. Again, this stability is clearly the result of the stronger carbene-metal bond strength compared to other supporting ligands.[28] Finally, while the Auranofin **A** inhibition of TrxR, in the nanomolar range, is much stronger than complexes

2 and **3**, the higher propensity of **A** to undergo non-selective binding to BSA[29] partially accounts for similar cytotoxicity observed for complexes **A**, **2** and **3**. Indeed, in contrast to **A**, complexes **2** and **3** inhibit TrxR in the micromolar range but show lower non-specific binding. In other words, taken together, the results from Table **1** and **2** confirm the high efficiency at which CAAC-bound complexes reach their biological target and may be reflected less by lower affinity, non-specific interactions with other, undesired, biomolecular targets. This increased selectivity might also result in improved bioavailability and bio-distribution profiles which are important parameters for experimental drugs at the developmental stage.

After establishing these intrinsic features, we decided to investigate the modes of cell death (necrosis vs. apoptosis) induced by each compound, **A**, **1**, **2** and **3**. We used a standard protocol in which HeLa cells were incubated with our set of gold compounds (10 μM) and their effects were studied in a time-dependent manner. Thus, at each selected time-point, cells were stained with FITC-Annexin V and Propidium Iodide (PI). Annexin V is a protein that binds to phosphatidylserine, a lipid constituent of plasma membrane which is normally restricted to the inner membrane surface but becomes exposed in the outer leaflet when cells undergo apoptosis. PI is used as a marker of cell membrane integrity and combined with Annexin V it becomes a tool to distinguish necrotic from apoptotic cells using flow cytometry.[30]

Analysis at various selected times, show that complexes **A**, **2** and **3** induce apoptosis at early time-points (Figure **2**), after which cells probably undergo secondary necrosis resulting in an increase of necrotic cells with respect to cells undergoing apoptosis. Among this series of compounds, the glucopyranosyl containing complex **3** stands out as the most potent cell death promoter, with almost 80 % of the total population of cells undergoing either apoptosis or necrosis. Interestingly, when exposed to $[Au(\mathbf{CAAC_{Et}})_2]^+TfO^-$ **1** for a duration of up to 6 h, a

negligible amount of cells undergo either apoptosis or necrosis. These results, as well as the previously discussed lack of TrxR inhibition exhibited by [Au(**CAAC$_{Et}$**)$_2$]$^+$TfO$^-$ **1**, supports the hypothesis that its cell death mechanism involves a different pathway than the gold complexes bearing only one CAAC ligand.

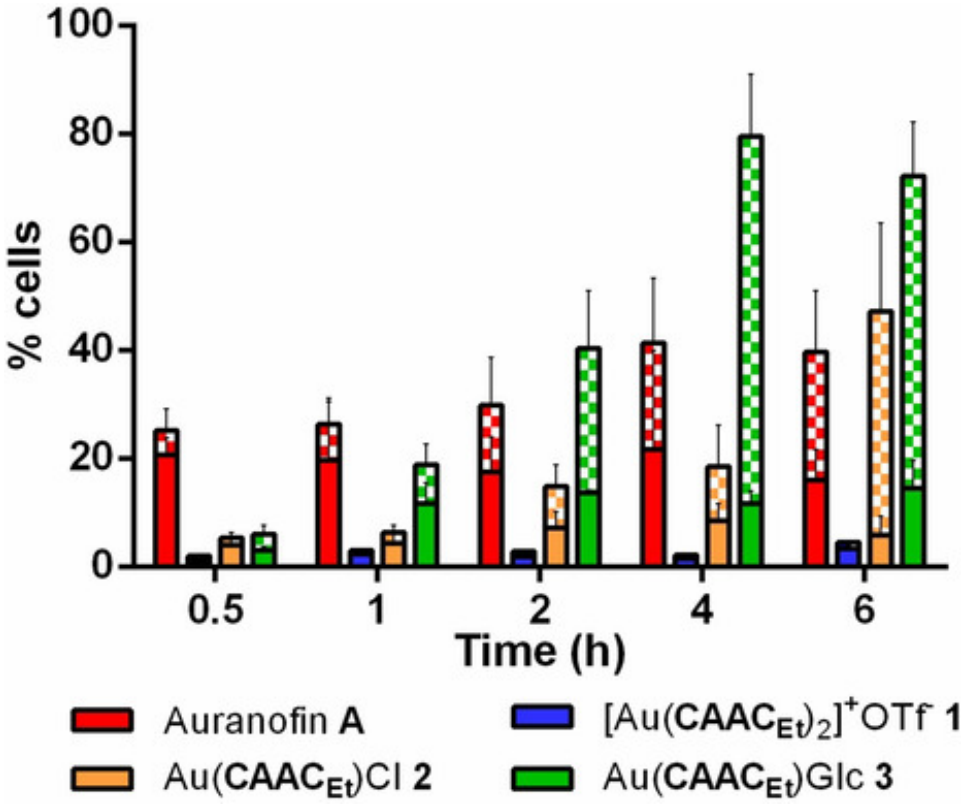


Figure A.2: Population of apoptotic vs. necrotic HeLa cells after a 30 min to 6 h exposure to complexes **A**, **1**, **2** and **3**. Full bars represent % cells undergoing apoptosis, while gridded bars represent cells under-going necrosis. The height of the bars represents the % of cells undergoing apoptosis and necrosis.

To further understand these results, we studied the ability of these compounds to induce depolarization and loss of mitochondrial *trans*-membrane potential ($\Delta\Psi m$), which is closely related to apoptotic cellular pathways.[31] To test this, we employed a lipophilic cationic dye, 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolylcarbocyanine iodide (JC-1), which is known to enter mitochondria and arrange in red fluorescent J-aggregates on cells with high

ΔΨm. Upon mitochondrial membrane depolarization, JC-1 remains in its green fluorescent monomeric form. Flow cytometry plots were obtained (FL1, green vs. FL2, red) for HeLa cells pretreated with Auranofin **A**, [Au(**CAAC$_{Et}$**)$_2$]$^+$TfO$^-$ **1**, Au(**CAAC$_{Et}$**)Cl **2**, Au(**CAAC$_{Et}$**)Glc **3** or dimethylformamide (DMF, solvent control) and subsequently stained with JC-1 (Figure S7). A healthy cell population was gated on an untreated cell sample (P1, double positive) and compared to the plot from the cells incubated with DMF or the gold complexes. Accordingly, a significant decrease of red fluorescent aggregates in cells exposed to Auranofin **A**, Au(**CAAC$_{Et}$**)Cl **2** and Au(**CAAC$_{Et}$**)Glc **3** can be observed. Once more, the latter display the largest effect, while [Au(**CAAC$_{Et}$**)^2]$^+$TfO- **1** showed only minimal decrease on J-aggregates or mitochondrial depolarization, suggesting a different mechanism of action.

Apoptosis and necrosis are the result of different biochemical events, which can be monitored as cellular morphological changes.[32] Owing to its high resolution and the possibility of introducing high-contrast staining agents to the samples, Transmission Electron Microscopy (TEM) is widely used to study cellular ultra-structures. Therefore, HeLa cells were incubated with DMF solutions of complexes Auranofin **A**, [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1**, Au(**CAAC$_{Et}$**)Cl **2** and Au(**CAAC$_{Et}$**)Glc **3** (10 μM) plus a control voided of any gold complex for 30 min and processed for TEM imaging. Representative TEM images of cells exposed to gold complexes are shown in low magnification (x2, 900, Figure **3 A**–D) and high magnification (x6, 800, Figure **3 F**–I). Lastly, TEM micro-graphs of the control experiment, with cells exposed to only DMF (Figure **3 E** and J) were obtained to ascertain the effect of each gold complex.

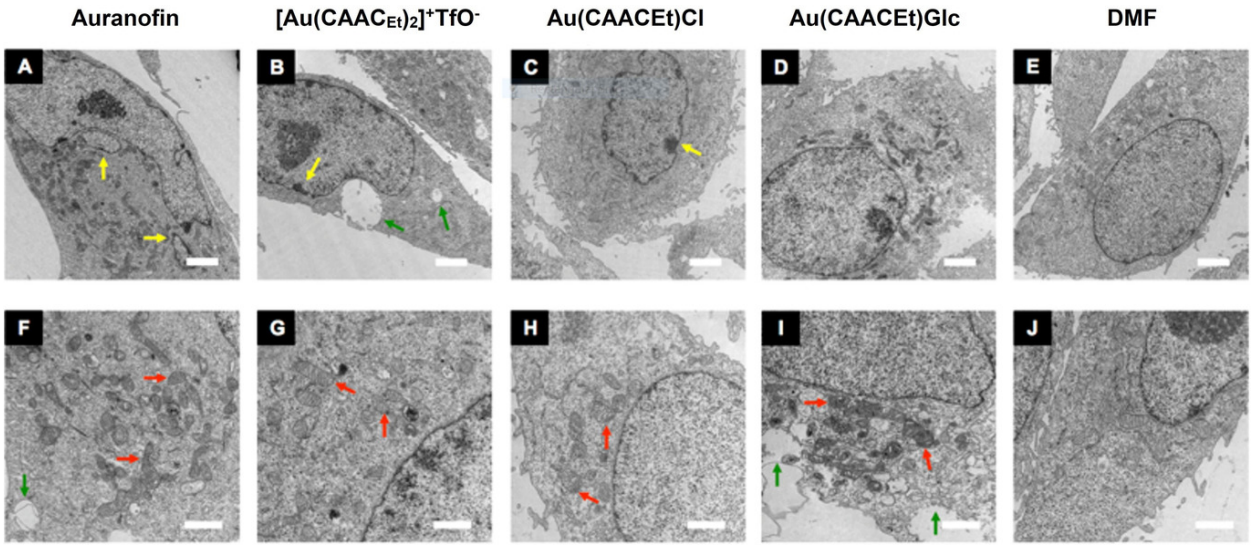


Figure A.3: TEM images of cells pretreated with Auranofin A (A, F), [Au(**CAAC$_{Et}$**)$_2$]$^+$TfO$^-$ **1** (B, G), Au(**CAAC$_{Et}$**)Cl **2** (C, H), Au(**CAAC$_{Et}$**)Glc **3** (D, I) at a 10 μM concentration and DMF as a vehicle control (E, J). Top panels show low magnification images (×2900, A-E, scale bars represent 2 μm) and bottom panels show high magnification images (×6800, F-J, scale bars represent 1 μm). Specific features of cell morphologies are shown with arrows (further discussion in the main text).

We consider normal cellular morphologies as rounded with homogenous cellular and nuclear membranes, each containing a nucleolus, chromatin, in addition to well-preserved cytoplasmic organelles. Auranofin **A** (Figure 3 A and F) is well known to promote apoptosis on a wide variety of cancer cells.[33] Under our experimental protocol, we could observe clear signs of apoptosis such as nuclear fragmentation or karyorrhexis (Figure 3 A, yellow arrows), and swollen cytoplasmic organelles and vacuoles (Figure 3 F, red and green arrows, respectively).[32b, 34] Au(**CAAC$_{Et}$**)Cl **2** (Figure 3 C and H) pretreated cells show morphological signs of early apoptosis such as condensation of chromatin at the nuclear mem-brane (Figure 3 C, yellow arrow) and preserved mitochondria (Figure 3 H, red arrows). Cells exposed to Au(**CAAC$_{Et}$**)Glc **3** (Figure 3 D and I) show the highest degree of pathological morphological features with cells transitioning from apoptosis to secondary necrosis such as swelling of mitochondria and disruption of plasma membrane (Figure 3 I, red and green arrows,

respectively). In contrast, cells exposed to [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1** (Figure **3 B** and G) show well rounded nuclei and preserved cytoplasmic organelles (Figure **3 G**, red arrows). Some degree of chromatin fragmentation and cytoplasmic vacuoles were observed (Figure **3 B**, yellow and green arrows, respectively), which may suggest a population of cells undergoing early apoptosis.

A.4 Conclusion:

In conclusion, we have demonstrated the effect of a CAAC as a gold ligand in anticancer cell studies. As previously reported with NHC supported complexes, mono- and bis-CAAC gold complexes possess different cellular targets.[7b] Both mono- and bis-CAAC supported gold centers benefit from the peculiar coordination environment and intrinsic electronic properties provided by the CAAC ligand. As a result, this new family of cytotoxic agents demonstrated significantly reduced non-specific binding toward non-tumor associated proteins such as albumin, while still retaining high affinity to the target TrxR (Au(**CAAC$_{Et}$**)Cl **2** and Au(**CAAC$_{Et}$**)Glc **3**). Moreover, morphological, and biochemical changes suggest that Au(**CAAC$_{Et}$**)X complexes **2** and **3** induce cell death through a mechanism indicative of apoptosis followed by secondary necrosis. [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1**, showed the highest cytotoxic activity with an onset of activity at longer time-points than the mono-ligated complexes and no affinity to TrxR. Therefore, other mechanisms of action need to be considered, such as the induction of Mitochondrial Membrane Permeabilization,[35] or binding to telomeric DNA G-quadruplexes.[36] Such mechanisms of action unrelated to TrxR inhibition are currently under investigation for [Au(**CAAC$_{Et}$**)^2]$^+$TfO$^-$ **1**. Altogether this investigation reports on the applicability of CAAC-supported gold complexes with application as antitumor agents and reveals their biological targets which may lead to lower undesired side-effects for gold metal-based therapeutics.

A.5: Acknowledgements

Thanks are due to the US National Science Foundation (CHE-1954380). M.T.P. thanks the UCSD Cancer Researchers in Nanotechnology (CRIN) for a postdoctoral fellowship and the mentorship of Dr A. Kummel within that program.

The Appendix, in full, is a reprint of the material as it appears in Chemistry- A European Journal 2021. Proetto, Maria T.; Alexander, Kelsey; Melaimi, Mohand; Bertrand, Guy; Gianneschi, Nathan C. vol. 27, 2021." The dissertation author was the secondary investigator and author of this paper.

A.6: References:

(1) H. Zhao, Y. Ning, Gold Bull. 2001, 34, 24-29.

(2) R. Koch, Dtsch. Med. Wochenstr. 1890, 16, 756-757.

(3) (a) W. F. Kean, F. Forestier, Y. Kassam, W. W. Buchanan, A. G. Rooney, Semin. Arthritis Rheum. 1985, 14,180-186. (b) W. F. Kean, C. J. L. Lock, H. Howard-Lock, Inflammopharmacology 1991, 1, 103-114.(c) W. F> Kean, L. Hart, W. W. Buchanan, Rheumatology 1997, 36,560-572; (d) A. G. Baldwin, D. Brough, S. Freeman, J. Med. Chem.2016, 59, 1691-1710.

(4) T. M. Simon, D. H. Kunishima, G. J. Vibert, A. Lorber, Cancer 1979, 44,1965-1975.

(5) (a) S. J. Berners-Price, A. Filipovska, Metallomics 2011, 3, 863-873. (b)C. Schmidt, L. Albrecht, R. M. Balasupramaniam, B. Karge, M. Brönstrup, A. Prokop, K. Baumann, S. Reichl, I. Ott, Metallomics 2019,11, 533-545. (c) J. L. Hickey, R. A. Ruhayel, P. J. Barnard, M. V. Baker, S. J. Berners-Price, A. Filipovska, J. Am. Chem. Soc. 2008, 130, 12570-12571.(d) W. Liu, R. Gust, Coord. Chem. Rev. 2016, 329, 191-213. (e) M. Mora, M. C. Gimeno,R. Visbal, Chem. Soc. Rev. 2019, 48,447-462. (f) C. H. G. Jakob, Bruno Dominelli, E. M. Hahn, T. O. Berghausen, T. Pinheiro, F. Marques, R. M. Reich, J. D. G. Correia, F.E. Kühn, Chem Asian J. 2020, 15, 2754-2762. (g) F. Magherini, T. Fiaschi, E. Valocchia, M. Becatti, A. Pratesi, T. Marzo, L. Massai, C. Gabbiani, I. Landini, S. Nobili, E. Mini, L. Messori, A. Modesti, T. Gamberi, Oncotarget, 2018, 9, 28042-28068

(6) (a) F. Angelucci, A. A. Sayed, D. L. Williams, G. Boumis, M. Brunori, D. Dimastrogiovanni, A. E. Miele, F. Pauly, A. Bellelli, J. Biol. Chem. 2009, 284, 28977-

28985. (b) A. Pratesi, C. Gabbiani, E. Michelucci, M.Ginanneschi, A. M. Papini, R. Rubbiani, I. Ott, L. Messori, J. Inorg.Biochem. 2014, 136, 161-169.

(7) (a) D. Trachootham, J. Alexandre, P. Huang, Nat Rev Drug Discov 2009, 8, 579-591. (b) W. Liu, K. Bensdorf, M. Proetto, A. Hagenbach, U. Abram, R. Gust, J. Med .Chem. 2012, 55, 3713-3724.

(8) (a) D. A. Chan, P. D. Sutphin, P. Nguyen, S. Turcotte, E. W. Lai, A. Banh, G. E. Reynolds, J. -T. Chi, J. Wu, D. E. Solow-Cordero, M. Bonnet, J.U. Flanagan, D.M. Bouley, E.E. Graves, W.A. Denny, M.P. Hay, A.J. Giaccia, Sci.Transl. Med. 2011, 3, 94ra70. b) A. K. Buck, S. N. Reske, J. Nucl. Med. 2004; 45, 461-463.

(9) M. S. Iqbal, S. G. Taqi, M. Arif, M. Wasim, M. Sher, Biol. Trace Elem.Res. 2009, 130, 204-209.

(10) For thematic issues and books on NHCs, see: (a) T. Rovis, S. P. Nolan, Synlett 2013, 24, 1188-1189. (b) A. J. Arduengo, G. Bertrand, Chem. Rev. 2009, 109, 3209-3210. (c) S. Díez-González, in N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools,Royal Society of Chemistry: Cambridge, 2016. (d) Nolan, S. P. in NHeterocyclic Carbenes: Effective Tools for Organometallic Synthesis;Wiley-VCH: Weinheim, 2014.

(11) (a) K. M. Hindi, M. J. Panzner, C. A. Tessier, C. L. Cannon, W. J. Youngs, Chem. Rev. 2009, 109, 3859-3884. (b) W. Liu, R. Gust, Chem.Soc. Rev. 2013, 42, 755-773.

(12) For the synthesis of CAACs, see: (a) V. Lavallo, Y. Canac, C. Präsang, B. Donnadieu, G. Bertrand, Angew. Chem. 2005, 117, 5851-5855; Angew. Chem., Int. Ed. 2005, 44, 5705-5709. (b) V. Lavallo, Y. Canac, A. DeHope, B. Donnadieu, G. Bertrand, Angew. Chem. 2005, 117, 7402-7405; Angew. Chem., Int. Ed. 2005, 44, 7236-7239; (c) R. Jazar, R. D. Dewhurst, J. B. Bourg, B. Donnadieu, Y. Canac, G. Bertrand, Angew. Chem. 2007, 119, 2957-2960; Angew. Chem., Int. Ed. 2007, 46, 2899-2902. (d) X. Zeng, G. D. Frey, R. Kinjo, B. Donnadieu, G. Bertrand J. Am. Chem. Soc. 2009, 131, 8690-8696. (e) J. Chu. D. Munz, R. Jazzar, M. Melaimi, G. Bertrand, J. Am. Chem. Soc. 2016, 138, 7884-7887.

(13) For reviews on CAACs see: (a) M. Soleilhavoup, G. Bertrand, Acc. Chem. Res. 2015, 48, 256-266. (b) M. Melaimi, M. Soleilhavoup, G. Bertrand, Angew. Chem. 2010, 122, 8992-9032; Angew. Chem., Int. Ed. 2010, 49, 8810-8849. (c) S. Roy, K. C. Mondal, H. W. Roesky, Acc. Chem. Res. 2016, 49, 357-369. (d) M. Melaimi, R. Jazzar, M. Soleilhavoup, G. Bertrand, Angew. Chem. 2017, 129, 10180-10203; Angew. Chem., Int. Ed. 2017, 56, 10046-10068. (e) R. Jazzar, M. Soleilhavoup, G. Bertrand, Chem. Rev. 2020, 120, 4141-4168.

(14) (a) V. Lavallo, Y. Canac, B. Donnadieu, W. W. Schoeller, G. Bertrand, Angew. Chem. 2006, 118, 3568-3571; Angew. Chem., Int. Ed. 2006, 45, 3488-3491. (b) O. Back, M.

Henry-Ellinger, C. D. Martin, D. Martin, G. Bertrand, Angew. Chem. 2013, 125, 3011-3015; Angew. Chem., Int. Ed. 2013, 52, 2939-2943.

(15) (a) L. Jin, D. R. Tolentino, M. Melaimi, G. Bertrand, Sci. Adv. 2015, 1, e1500304. (b) L. Jin, E. A. Romero, M. Melaimi, G. Bertrand, J. Am. Chem. Soc. 2015, 137, 15696-1569. c) R. Kinjo, B. Donnadieu, G. Bertrand, Angew. Chem. 2011, 123, 5674-5677; Angew. Chem., Int. Ed. 2011, 50, 5560-5563.

(16) (a) V. Lavallo, G. D. Frey, S. Kousar, B. Donnadieu, G. Bertrand, Proc. Natl. Acad. Sci. USA 2007, 104, 13569-13573. (b) V. M. Marx, A. H. Sullivan, M. Melaimi, S. C. Virgil, B. K. Keitz, D. S. Weinberger, G. Bertrand, R. G. Grubbs. Angew. Chem. 2015, 127, 1939-1943; Angew. Chem., Int. Ed. 2015, 54, 1919-1923.

(17) (a) D. S. Weinberger, M. Melaimi, C. E. Moore, A. L. Rheingold, G. Frenking, P. Jerabek, G. Bertrand. Angew. Chem. 2013, 125, 9134-9137; Angew. Chem., Int. Ed. 2013, 52, 8964-8967. (b) D. S.Weinberger, S. K. N. Amin, K. C. Mondal, M. Melaimi, G. Bertrand, A. C. Stückl, H. W. Roesky, B. Dittrich, S. Demeshko, B. Schwederski, W. Kaim, P. Jerabek, G. Frenking. J. Am. Chem. Soc. 2014, 136, 6235-6238. (c) L. Jin, D. S. Weinberger, M. Melaimi, C. E. Moore, A. L. Rheingold, G. Bertrand. Angew. Chem. 2014, 126, 9205-9209; Angew. Chem., Int. Ed. 2014, 53, 9059-9063. (d) L. Jin, M. Melaimi, A. Kostenko, M. Karni, Y. Apeloig, C. E. Moore, A. L. Rheingold, G. Bertrand, Chem. Sci. 2016, 7, 150-154.

(18) B. Bertrand, A. S. Romanov, M. Brooks, J. Davis, C. Schmidt, I. Ott. M. O'Connell, M. Bochmann, Dalton Trans. 2017, 46, 15875-15887. (b) M.R. M. Williams, B. Bertrand, J. Fernandez-Cestau, Z. A. E. Waller, M. AO'Connell, M. Searcey, M. Bochmann, Dalton Trans 2018, 47, 13523-13534. (c) B. Bertrand, M. R. M. Williams, M. Bochmann, Chem. Eur. J. 2018, 24, 11840-11851.

(19) G. D. Frey, R. D. Dewhurst, S. Kousar, B. Donnadieu, G. Bertrand, J. Organomet. Chem. 2008, 693, 1674-1682.

(20) CCDC 2022581 ([Au(CAACEt)2]+TfO- 1), 2022582 (Au(CAACEt)Cl 2)and 2022583 (Au(CAACEt)Glc 3) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre.

(21) (a) K. C. Mondal, P. P. Samuel, Y. Li, H. W. Roesky, S. Roy, L. Ackermann, N. S. Sidhu, G. M. Sheldrick, E. Carl, S. Demeshko, S. De,P. Parameswaran, L. Ungur, L. F. Chibotaru, D. M. A. Andrada, Eur. J. Inorg. Chem. 2014, 818–823. (b) S. Roy, K. C. Mondal, J. Meyer, B. Niepotter, C. Kohler, R. Herbst-Irmer, D. Stalke, B. Dittrich, D. M. Andrada, G. Frenking, H. W. Roesky, Chem. Eur. J. 2015, 21, 9312-9318.

(22) (a) J. S. Modica-Napolitano, J. R. Aprille, Adv. Drug Delivery Rev. 2001, 49, 63-70; (b) S. J. Berners-Price, A. Filipovska, Aust. J. Chem. 2008,61, 661-668.

(23) (a) A. Bindoli, M. P. Rigobello, G. Scutari, C. Gabbiani, A. Casini, L. Messori, Coord. Chem. Rev. 2009, 253, 1692-1707. (c) R. Rubbiani, I. Kitanovic, H. Alborzinia, S. Can, A. Kitanovic, L. A. Onambele, M.Stefanopoulou, Y. Geldmacher, W. S. Sheldrick, G. Wolber, A. Prokop, S. Wölfl, I. Ott, J. Med. Chem. 2010, 53, 8608-8618.

(24) W. Liu, K. Bensdorf, M. Proetto, U. Abram, A. Hagenbach, R. Gust, J. Med. Chem. 2011, 54, 8605-8615.

(25) W. Walther, O. Dada, C. O'Beirne, I. Ott, G. Sánchez-Sanz, C. .Schmidt, C. Werner, X. Zhu, M. Tacke, Lett. Drug Des. Discov. 2017, 14, 125-134.

(26) C. Schmidt, B. karge, R. Misgeld, A. Prokop, M. Brönstrup, I. Ott, Med. Chem. Commun. 2017, 8, 1681-1689.

(27) C. Schmidt, B. Karge, R. Misgeld, A. Prokop, R. Franke, M. Bronstrup, I. Ott, Chem. Eur. J. 2017, 23, 1869-1880.

(28) (a) H. Sivaram, J. Tan, H. V. Huynh, Organometallics 2012, 31, 5875-5883. (b) H. F. Dos Santos, M. A. Viera, G. Y. Sánchez Delgado, D.Paschoal, J. Phys. Chem. A 2016, 120, 2250-2259.

(29) Non-specific binding of Auranofin A to HSA, in the range of 60 to 80 %, is a well-studied phenomenon that results in higher required concentration to cause cell death: see reference 9

(30) L. C. Crowley, B. J. Marfell, A. P. Scott, N. J. Waterhouse, Cold Spring Harb. Protoc. 2016, 2016, pdb prot087288.

(31) J. D. Ly, D. R. Grubb, A. Lawen, Apoptosis 2003, 8, 115-128.

(32) (a) J. F. Kerr, A. H. Wyllie, A. R. Currie, Br. J. Cancer 1972, 26, 239-257; (b) U. Ziegler, News Physiol Sci 2004, 19, 124-128.

(33) (a) P. Zou, M. X. Chen, J. S. Ji, W. Q. Chen, X. Chen, S. L. Ying, J. R. Zhang, Z. H. Zhang, Z. G. Liu, S. L. Yang, G. Liang, Oncotarget 2015, 6,36505-36521. (b) N. Park, Y. J. Chun, J Toxicol Env Heal A 2014, 77, 1467-1476.

(34) (a) D. V. Krysko, T. Vanden Berghe, K. D'Herde, P. Vandenabeele, Methods 2008, 44, 205-221. (b) S. Elmore, Toxicol Pathol 2007, 35,495-516.

(35) M. V. Baker, P. J. Barnard, S. J. Berners-Price, S. K. Brayshaw, J. L. Hickey, B. W. Skelton, A. H. White, Dalton Trans. 2006, 3708-3715.

(36) C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri, A. Casini, Angew. Chem. 2016, 128, 4328-4331; Angew. Chem. Int. Ed. 2016, 55, 4256-4259