

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Investigations into Bioactive Secondary Metabolites Produced by Marine Cyanobacteria /

### Permalink

<https://escholarship.org/uc/item/4px5m8q7>

### Author

Mervers, Emily

### Publication Date

2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Investigations into Bioactive Secondary Metabolites Produced by Marine Cyanobacteria

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

in

Chemistry

by

Emily Mevers

Committee in charge:

Professor William H. Gerwick, Chair  
Professor Lihini Aluwihare  
Professor Timothy Baker  
Professor Michael Burkart  
Professor Thomas Hermann

2014

Copyright

Emily Mevers, 2014

All rights reserved

The Dissertation of Emily Mevers is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2014

## DEDICATION

This dissertation is dedicated to my wonderful parents, Sandy and Steve Mevers, my fabulous sisters, Kim and Jen, and all of my amazing friends! For without your love and support over the years I would have never survived these past five years, let alone been able to achieve my dreams. Thank you all for the support, and I am especially blessed to have so many remarkable people in my life!

## EPIGRAPH

“Live as if you were to die tomorrow. Learn as if you were to live forever.”

*~Mahatma Gandhi*

## TABLE OF CONTENTS

Signature Page .....	iii
Dedication .....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Abbreviations .....	x
List of Figures.....	xiii
List of Tables .....	xiv
Acknowledgements.....	xv
Vita.....	xviii
Abstract of the Dissertation .....	xx
Chapter 1: Introduction.....	1
1.1 Terrestrial Natural Products .....	1
1.1.1 Historical Uses of Natural Products.....	1
1.1.2 Early Investigations into Terrestrial Natural Products.....	3
1.1.3 Impact of Natural Products on the Pharmaceutical Industry.....	5
1.2 Marine Natural Products .....	8
1.2.1 Early Investigations into Marine Natural Products.....	8
1.2.2 Modern Day Investigations into Marine Natural Products .....	12
1.2.3 Marine Derived Drugs .....	16
1.2.4 Investigations into Marine Cyanobacteria .....	18
1.2.5 Therapeutic Potential of Secondary Metabolites from Marine Cyanobacteria .....	21
1.2.5.1 Cancer Cell Cytotoxicity.....	21
1.2.5.2 Other Biologically Important Activity Observed from Cyanobacterial Metabolites .....	25
1.3 Dissertation Contents .....	30
1.4 References.....	33
Chapter 2: Cytotoxic Veraguamides, Alkynyl Bromide-Containing Cyclic Depsipeptides from the Marine Cyanobacterium cf <i>Oscillatoria margaritifera</i> .....	41
2.0.1 Abstract.....	41
2.1 Introduction.....	41
2.2 Results and Discussion .....	44
2.2.1 Veraguamide A .....	44
2.2.2 Veraguamide B-C, and K-L.....	49
2.2.3 Bioassay Results .....	53
2.2.4 Taxonomy of Producing Organism.....	53
2.3 Conclusion .....	54

2.4 Experimental Methods .....	56
2.4.1 General Experimental Procedures .....	56
2.4.2 Cyanobacterial Collections and Morphological Identification .....	56
2.4.3 Extraction and Isolation .....	59
2.4.4 Hydrogenation, Acid Hydrolysis and Marfey's Analysis .....	57
2.4.5 Preparation and GCMS Analysis of 2-Hydroxy-3-methylpentanoic Acid (H3mpa) .....	60
2.4.6 Preparation and GCMS Analysis of Methyl 3-Hydroxy-2-Methyloctanoate (Hmoaa) .....	61
2.4.7 Tandem Mass Spectrometry Data Acquisition and Preprocessing .....	62
2.4.8 Cytotoxicity Assay .....	62
2.4.9 DNA Extraction, Amplification, and Sequencing .....	63
2.4.10 Phylogenetic Inference .....	64
2.5 Acknowledgements .....	65
2.6 Appendix .....	66
2.7 References .....	85
Chapter 3: Lyngbyabellin N, A Cytotoxic Secondary Metabolite from a Palmyra Atoll Collection of the Marine Cyanobacterium <i>Moorea bouillonii</i> .....	88
3.0.1 Abstract .....	88
3.1 Introduction .....	88
3.2 Results and Discussion .....	89
3.2.1 Lyngbyabellin N Collection and Isolation .....	89
3.2.2 Lyngbyabellin N Planar Structure Determination .....	91
3.2.3 Stereochemical Analysis of Lyngbyabellin N .....	95
3.2.4 Bioassay Results .....	98
3.3 Conclusion .....	99
3.4 Experimental Methods .....	100
3.4.1 General Experimental Procedures .....	100
3.4.2 Cyanobacterial Collection and Taxonomic Identification .....	100
3.4.3 Isolation of Lyngbyabellin N .....	100
3.4.4 Ozonolysis and Acid Hydrolysis of Lyngbyabellin N .....	101
3.4.5 Modified Marfey's Analysis of the Statine Unit .....	101
3.4.6 Preparation and GCMS Analysis of Hmba .....	102
3.4.7 Preparation and GCMS Analysis of <i>N,N</i> -Dimethylvaline .....	103
3.4.8 Cytotoxicity Assay .....	103
3.5 Acknowledgements .....	104
3.6 Appendix .....	105
3.7 References .....	112
Chapter 4: Lipopeptides from the Tropical Marine Cyanobacterium <i>Symploca</i> sp .....	114
4.0.1 Abstract .....	114
4.1 Introduction .....	114
4.2 Results and Discussion .....	116
4.2.1 Tasiamides C-E Collection and Isolation .....	116
4.2.2 Tasiamide C Absolute Structure Determination .....	116
4.2.3 Tasiamide D Absolute Structure Elucidation .....	121
4.2.4 Tasiamide E Absolute Structure Determination .....	124
4.2.5 Structural Similarities of Tasiamides C-E to Known Metabolites .....	129

4.2.6 Bioassay Results .....	130
4.3 Conclusion .....	132
4.4 Experimental Methods .....	132
4.4.1 General Experimental Procedures .....	132
4.4.2 Extraction and Isolation .....	133
4.4.3 Acid Hydrolysis and Marfey's Analysis of Tasiamides C-E .....	133
4.4.4 Preparation and GCMS Analysis of Isoleucine (Ile) in Tasiamides C-E .....	135
4.4.5 Base Hydrolysis of Hmba Unites in Tasiamides C and D .....	136
4.4.6 GCMS Analysis of H4mpa in Tasiamide E .....	136
4.4.7 H-460 Cytotoxicity Assay .....	137
4.4.8 HCT-116 Cytotoxicity Assay .....	137
4.5 Acknowledgements .....	138
4.6 Appendix .....	139
4.7 References .....	165
Chapter 5: Parguerene, Precarribowmide, and Mooreamide: New Lipopeptides from the Marine Cyanobacterium <i>Moorea</i> sp. ....	167
5.0.1 Abstract .....	167
5.1 Introduction .....	167
5.2 Results and Discussion .....	169
5.2.1 Collection and Isolation .....	169
5.2.2 Structure Elucidation of Parguerene .....	170
5.2.3 Structure Elucidation of Precarribowmide .....	174
5.2.4 Structure Elucidation of Mooreamide .....	179
5.2.5 Biosynthetic Origin .....	183
5.3 Conclusion .....	184
5.4 Experimental Methods .....	184
5.4.1 General Experimental Procedures .....	184
5.4.2 Morphological Identification .....	185
5.4.3 Collection, Extraction, and Isolation .....	185
5.4.4 Oxidation, Acid Hydrolysis, and Marfey's Analysis of Parguerene .....	186
5.4.5 Acid Hydrolysis and Marfey's Analysis of Precarribowmide .....	187
5.4.6 Polymerase Chain Reaction (PCR) and Cloning for PRM-25/Mar/11-2 .....	188
5.4.7 Phylogentic Interferences for PRM-25/Mar/11-2 .....	188
5.4.8 H-460 Cytotoxicity Assay .....	189
5.5 Acknowledgements .....	190
5.6 Appendix .....	191
5.7 References .....	205
Chapter 6: Structure-Activity Relationships of the Lyngbyamide Class of Marine Natural Product .....	208
6.0.1 Abstract .....	208
6.1 Introduction .....	208
6.2 Results and Discussion .....	210
6.2.1 Synthetic Approach to Synthesizing Lyngbyamide Analogs .....	210
6.2.2 Biological Evaluation of the First Round of Analogs .....	214
6.2.3 The Synthesis and Activity of Rounds 2A and 2B .....	216
6.2.4 Analysis of the Cathepsin L Stabilization, Brine Shrimp Toxicity, and CMC data .....	220

6.2.5 Significance of this Study .....	224
6.3 Conclusion .....	226
6.4 Experimental Methods .....	227
6.4.1 General Experimental Procedures.....	227
6.4.2 Amide Bond Formation .....	227
6.4.3 Olefin Metathesis .....	228
6.4.4 Base Hydrolysis .....	228
6.4.5 Cyclopropanation.....	228
6.4.6 Analytical Data for Synthetic Analogs .....	229
6.4.7 Brine Shrimp Toxicity Assay.....	240
6.4.8 Cathepsin L Assay .....	240
6.4.9 Critical Micelle Concentration.....	241
6.4.10 Statistical Analysis.....	241
6.5 Acknowledgements.....	242
6.6 Appendix.....	243
6.7 References.....	291
 Chapter 7: Conclusion & Future Work.....	 294
7.1 Summary of the Research Presented in the Dissertation and Future Work .....	294
7.2 Expected Future Direction of Marine Natural Products Chemistry .....	301
7.3 References.....	305

## LIST OF ABBREVIATIONS

- 2-AG – 2-arachidonoyl glycerol
- AEA – Anadamide
- AMC – 7-amino-4-methylcoumarin
- CD – Circular Dichroism
- CHCl<sub>3</sub> – Chloroform
- CH<sub>2</sub>Cl<sub>2</sub> – Dichloromethane
- CH<sub>2</sub>N<sub>2</sub> - Diazomethane
- CH<sub>3</sub>CN – Acetonitrile
- CMC – Critical Micelle Concentration
- CNP – Coiba National Park
- CNS – Central Nervous System
- COSY – Correlation Spectroscopy
- DCE – 1,2-Dichloroethane
- DIPC – *N,N*-Diisopropylcarbodiimide
- DMSO – Dimethyl sulfoxide
- ECS – Endocannabinoid System
- ESI – Electrospray Ionization
- Et<sub>2</sub>O – Diethyl Ether
- EtOAc – Ethyl Acetate
- EtOH – Ethanol
- FDAA – 1-fluoro-2,4-dinitrophenyl-5-alanine amide
- FT – Fourier Transform
- GCMS – Gas Chromatography Mass Spectrometry
- H2BC – Heteronuclear 2 Bond Correlation

HMBC – Heteronuclear Multiple Bond Correlation

HOBt – Hydroxybenzotriazole

HPLC – High Pressure Liquid Chromatography

HR – High Resolution

HSQC – Heteronuclear Single Quantum Correlation

ICBG – International Cooperative Biodiversity Groups

IR – Infrared Spectroscopy

LCMS – Liquid Chromatography Mass Spectrometry

LR – Low Resolution

MeOH – Methanol

MS – Mass spectrometry

MTPA – Methoxy(trifluoromethyl)phenylacetic acid

NCE – New Chemical Entity

NMR – Nuclear Magnetic Resonance

NOE – Nuclear Overhauser Effect

NP – Normal Phase

NRPS – Non-Ribosomal Peptide Synthetase

PKS – Polyketide Synthase

PrepTLC – Preparatory Thin Layer Chromatography

ROESY – Rotating-Frame Nuclear Overhauser Effect Correlation Spectroscopy

RP – Reverse Phase

SAR – Structure-Activity Relationship

SCUBA – Self Contained Underwater Breathing Apparatus

SPE – Solid Phase Extraction

TOCSY – Total Correlation Spectroscopy

TOF – Time of Flight

UV – Ultraviolet

VLC – Vacuum Liquid Chromatography

## LIST OF FIGURES

Figure 1.1: Plant alkaloids .....	2
Figure 1.2: Terrestrial microbial metabolites.....	4
Figure 1.3: A portion of the NCE approved for use between 1981-2010 .....	8
Figure 1.4: Early marine natural products .....	11
Figure 1.5: Modern marine secondary metabolites.....	16
Figure 1.6: Marine derived drugs.....	18
Figure 1.7: Cyanobacteria natural products .....	21
Figure 1.8: Bioactive cyanobacteria metabolites (part 1.....	24
Figure 1.9: Bioactive cyanobacteria metabolites (part 2.....	29
Figure 2.1: Veraguamides A-C, and K-L.....	43
Figure 2.2: Select 2D NMR data for veraguamide A .....	46
Figure 2.3: Sequencing by ESIMS/MS fragmentations.....	51
Figure 2.4: Differences in <sup>13</sup> C NMR chemical shifts between the natural product and synthetic versions .....	55
Figure 3.1: Lyngbyabellin N along with other related metabolites .....	90
Figure 3.2: Select 2D NMR Data for Lyngbyabellin N.....	91
Figure 3.3: <sup>13</sup> C NMR comparison between lyngbyabellin N and both H and K.....	96
Figure 3.4: Comparison of CD spectrums of lyngbyabellin K and N.....	96
Figure 3.5: Synthetic scheme for the synthesis of the four statine residues.....	97
Figure 3.6: Structure similarity of lyngbyabellin N to coibamide A, dolastatin 10, and lyngbyabellin A.....	99
Figure 4.1: Tasiamide C-E and grassystatin A-B .....	115
Figure 4.2: Select 2D NMR data for tasiamide C-E .....	121
Figure 4.3: Select low resolution MS fragmentation cleavages for tasiamide C-E .....	122
Figure 5.1: Structures of the new metabolites parguerene, mooreamide, and precarriewowmide, along with structurally related metabolites .....	169
Figure 5.2: Four partial structures and select 2D NMR data for parguerene .....	174
Figure 5.3: Select 2D NMR data for precarriewowminde .....	179
Figure 5.4: Three partial structures and select 2D NMR data for mooreamide .....	181
Figure 6.1: Alkyl amides from marine cyanobacteria.....	210
Figure 6.2: Round 1 analogs of lyngbyamide A .....	212
Figure 6.3: General synthetic scheme.....	213
Figure 6.4: Cathepsin L and brine shrimp activity for round 1 analogs .....	215
Figure 6.5: Round 2a and 2b of lyngbyamide analogs .....	216
Figure 6.6: Initial velocity of the analogs compared to the enzyme blank (E.....	218
Figure 6.7: Stabilization of cathepsin L .....	218
Figure 6.8: Surface tension readings for compound 33d .....	219
Figure 6.9: Comparison of cathepsin L dose response and surface tension readings over the same logarithmic concentration scale .....	221
Figure 7.1: Cyanobacterial natural products that were discussed in the previous research Chapter.....	295

## LIST OF TABLES

Table 2.1: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for veraguamide A in $\text{CDCl}_3$ .....	48
Table 3.1: NMR spectral data for lyngbyabellin N in $d_6$ -DMSO.....	94
Table 4.1: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for tasiamide C in $\text{CDCl}_3$ .....	119
Table 4.2: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for tasiamide D in $\text{CDCl}_3$ .....	123
Table 4.3: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for tasiamide E in $\text{CDCl}_3$ .....	127
Table 4.4: Residue sequence comparisons of the tasiamide/grassystatin/symplocin superfamily of metabolites .....	131
Table 5.1: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for parguerene in $\text{CDCl}_3$ .....	173
Table 5.2: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for precarriebowmide in $\text{CD}_3\text{OD}$ .....	177
Table 5.3: $^1\text{H}$ and $^{13}\text{C}$ NMR assignments for mooreamide in $\text{CDCl}_3$ .....	182
Table 6.1: Biological data for the lyngbyamide A analogs.....	222
Table 6.2: Comparing the CMC and MW of commercially available biologically important surfactants and two analogs .....	226

## ACKNOWLEDGEMENTS

I would like to acknowledge my entire committee for their time, feedback and support of my research but especially my advisor Professor William H. Gerwick and co-advisor Professor Michael Burkart. Professor William H. Gerwick has always been a huge supporter of my research and my future as a scientist as he had a significant role in many of my achievements, including being awarded the Teddy Traylor award, placement on the Cancer Mammals Training grant for the past 3 years, and most importantly acquiring a postdoctoral position in Jon Clardy's laboratory at the Harvard Medical School. I greatly appreciate his mentorship, support, and guidance over the past 4.5 years.

I would also like to acknowledge Bill Baker at the University of South Florida for igniting my passion of marine natural products chemistry. He gave me a coveted spot in his laboratory working with a brilliant graduate student, Alan Maschek, on secondary metabolites from nudibranchs, and I appreciate his continued mentorship, support, and guidance in my future.

I would also like to acknowledge my entire family, especially my parents and sisters for their love, support, and encouragement throughout my educational years, as I would never have made it this far without all of them. I especially appreciate all of those fishing trips in Southwest Florida with my father, trips to Hawaii and Colorado with my sisters, and those many Scrabble games with my mother.

I would also like to acknowledge all of my amazing friends, especially Alec Saitman, Paul Boudreau, Kevin Daniel, Emily Monroe, Matt Bertin, Karin Kleigrew, Sam Mascuch, Adam Bravard, Christopher DeBona, Josh Weaver, and Angi Littlejohn. I would like to thank them for lending an ear, telling a joke, responding to a text, and going for beer or a hike when I needed it the most. They have all been vital to my sanity and I know they will be friends for life!

I would like to also acknowledge my collaborators, Tara Byrum, Samantha Mascuch, Paul Boudreau, Alec Saitman, Emily Monroe, Changlun Shao, Bailey Miller, Matthew Bertin, Karin

Kleigrewe, Amanda Fenner, Cameron Coates, Emily Trentacoste, and all of the past and present Gerwick lab members. It has been a true pleasure to be able to work alongside each of them and I appreciate all of their advice and intriguing discussions.

I would like to also acknowledge the SIO dive facilities, especially Rich Walsh and Christian McDonald, for all of their help putting together collection trips. The UCSD Mass Spectrometry and NMR facilities, especially Anthony Mrse and Brendan Duggan, for their immense knowledge on acquiring 2D NMR data.

Chapter 2, in essence, is a reprint as it appears in the *Journal of Natural Products* 2011, 74, 928-936, with the following authors Emily Mevers, Wei-Ting Liu, Niclas Engene, Hosein Mohimani, Tara Byrum, Pavel A. Pevzner, Pieter C. Dorrestein, Carmenza Spadafora, and William H. Gerwick. The dissertation author was the primary investigator and first author of this paper.

Chapter 3, in part, includes a reprint as it appears in the *European Journal of Organic Chemistry* 2012, 5141-5150, with the following authors Hyukjae Choi, Emily Mevers, Tara Byrum, Frederick A. Valeriote, and William H. Gerwick. The dissertation author was a primary investigator and co-first author of this paper.

Chapter 4, in essence, is an accepted manuscript in the *Journal of Natural Products* 2014, with the following authors, Emily Mevers, F. P. Jake Haeckl, Paul D. Boudreau, Tara Byrum, Pieter C. Dorrestein, Frederick A. Valeriote, and William H. Gerwick. The dissertation author was the primary investigator and is the first author of this paper.

Chapter 5, in part, includes a reprint as it appears in the *Journal of Natural Products* 2013, 76, 1810-1814, with the following authors Emily Mevers, Tara Byrum, and William H. Gerwick. Also included in chapter 5, in part, a prepared manuscript for submission to the journal *Lipids*, with the following authors, Emily Mevers, Vincenzo Di Marzo, and William H. Gerwick. The dissertation author was the primary investigator and the first author on each of these manuscripts.

Chapter 6, in essence, is currently being prepared for submission in 2014, with the following authors Emily Mevers, Bailey Miller, Matthew Bertin, Vivian Hook, and William H. Gerwick. The dissertation author was the primary investigator and will be the first author of this material.

## VITA

### EDUCATION AND FIELDS OF STUDY

University of California, San Diego, La Jolla, CA Doctor of Philosophy in Chemistry <i>Advisors: William H. Gerwick and Michael Burkart</i>	2009-2014
University of California, San Diego, La Jolla, CA Master of Science in Chemistry	2009-2011
University of South Florida, Tampa, FL Bachelor of Science in Chemistry, minor in Mathematics	2005-2009

### FELLOWSHIP AND AWARDS

NIH Biochemistry of Growth Regulation and Oncogenesis	2011-2014
Teddy Traylor Award, University of California, San Diego	7/2013
David Carew Travel Award, American Society of Pharmacognosy	7/2012

### PUBLICATIONS

Mevers, E., Haeckl, F., Boudreau, P., Byrum, T., Gerwick, W. Three Lipopeptides from the Tropical Marine Cyanobacteria *Symploca* sp. *Journal of Natural Products* **2014**, accepted.

Mevers, E., Byrum, T., Gerwick, W. Parguerene and Precarriebowmide, Biosynthetically Intriguing Metabolites from the Marine Cyanobacterium *Moorea producens* *Journal of Natural Products* **2013**, 76, 1810-1814.

Maschek, J., Mevers, E., Diyabalanage, T., Chen, L., Ren, Yuan, McClintock, J., Amsler, C., Wu, J., Baker, B. Palmadorin Chemodiversity from the Antarctic nudibranch *Austrodoris kerguelenensis* and inhibition of Jak2/STAT5-dependent HEL Leukemia Cells *Tetrahedron* **2012**, 68, 9095-9104.

\*Choi, H., \*Mevers, E., Byrum, T., Gerwick, W. Lyngbyabellins K-N from Two Palmyra Atoll Collections of the Marine Cyanobacterium *Moorea bouillonii* *European Journal of Organic Chemistry* **2012**, 5141-5150.

\*Authors contributed equally to work presented in manuscript

Mevers, E., Liu, W.T., Engene, N., Mohimani, H., Byrum, T., Pevzner, P., Dorrestein, P., Gerwick, W. Cytotoxic Veraguamides, Alkynyl Bromide Containing Cyclic Depsipeptides from the marine Cyanobacterium cf. *Oscillatoria margaritifera* *Journal of Natural Products* **2011**, 74, 928-936.

Nunnery, J., Mevers, E., Gerwick, W. Biologically active secondary metabolites from marine cyanobacteria *Current Opinions in Biotechnology* **2010**, 6, 797-793.

## PUBLICATIONS IN PREPARATION

Mevers, E., Di Marzo, V., Gerwick, W. Mooreamide A: A Cannabinomimetic Lipid from the Marine Cyanobacterium *Moorea bouillonii* *Lipids* **2014**, in prep.

Mevers, E.; Miller, B.; Hook, V.; Gerwick, W. H. Structure Activity Relationships of the Lyngbyamide Class of Marine Natural Product, **2014**, in prep.

ABSTRACT OF THE DISSERTATION

Investigations into Bioactive Secondary Metabolites Produced by Marine Cyanobacteria

by

Emily Mevers

Doctor of Philosophy in Chemistry

University of California, San Diego, 2014

Professor William Gerwick, Chair

Marine cyanobacteria are prolific producers of structurally intriguing and biologically important secondary metabolites, many of which are of mixed NRPS/PKS biosynthetic origins, and have a broad range of biological activity, including ion channel modulation, cancer cell toxicity, anti-parasitic, anti-bacterial, anti-inflammatory, brine shrimp toxicity, and molluscicidal. Presently, there is one clinically approved drug that is

an analog of the cyanobacterial natural product, dolastatin 10, while there are several agents in clinical trial, including soblidotin and synthadotin, which are analogs of dolastatin 15 and 10, respectively. Additionally, others are currently undergoing preclinical evaluation as anti-cancer agents, including apratoxin F, curacin A, desmethoxymajusculamide C (DMMC) and somacystinamide. The primary research objective of the research herein was to isolate and elucidate the structures of biologically active secondary metabolites from tropical marine cyanobacteria. In total, fifteen novel compounds from either *Oscillatoria* or *Moorea* were isolated and characterized. These include thirteen highly modified peptides (veraguamides A-C and H-L, precarriebowmide, tasiamides C-E, and lyngbyabellin N) and two alkyl amides (parguerene and mooreamide). The planar structure elucidation of each of these metabolites involved the use of 2D NMR spectroscopy and mass spectrometry techniques, including a mass spectrometry based dereplication algorithm to deduce the planar structure of several of the modified peptides. Absolute stereochemical analysis involved many techniques, such as Marfey's analysis, semi-synthesis,  $^3J$  coupling constant analysis, circular dichroism,  $^{13}\text{C}$  NMR comparisons, NOE correlations, and chiral GCMS analysis. Many of these compounds were biologically evaluated with veraguamide A and lyngbyabellin N exhibiting cancer cell cytotoxicity [ $\text{IC}_{50} = 141 \text{ nM}$  (H-460) and  $\text{IC}_{50} = 40.9$  (HCT-116), respectively], and mooreamide exhibiting cannabinoid receptor binding activity ( $K_i = 0.47 \text{ }\mu\text{M}$ ). A secondary research objective has been the structure-activity relationship (SAR) study to investigate the active pharmacophore in the lyngbyamide family of compounds, which consist of a cyclopropyl fatty acid (tail) and an amide head group. In total, 50 analogs were synthesized, designed to probe the importance of several structural characteristics of the lyngbyamides. These compounds

were tested in a wide array of biological assays, and a subset were found to possess strong activity in the stabilization of cathepsin L-mediated proteolysis, brine shrimp toxicity, and surface tension suppression.

Chapter 1:  
Introduction

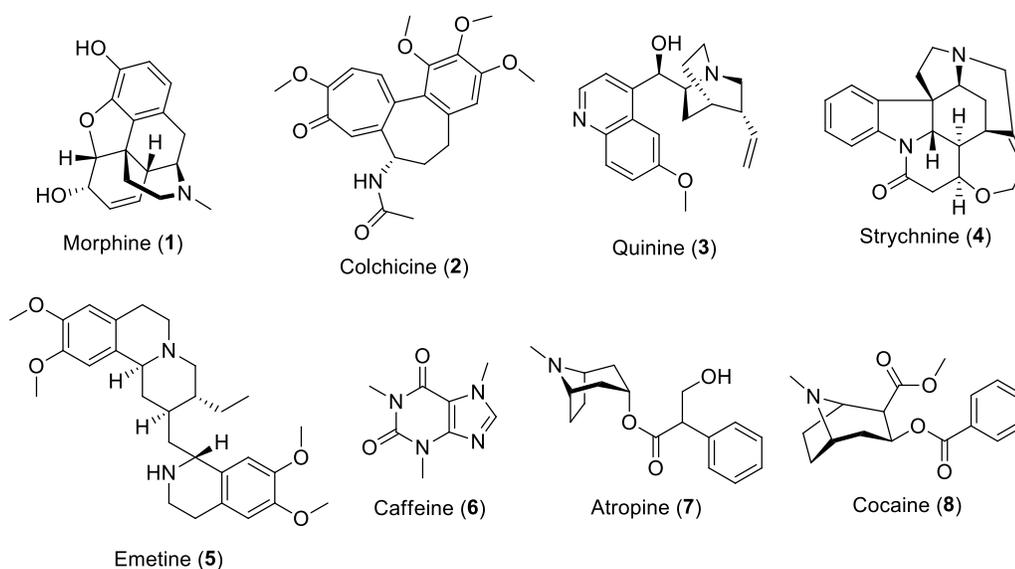
1.1 Terrestrial Natural Products

1.1.1 Historical Uses of Natural Products

Mankind has relied on nature for many essential things such as food, shelter, clothing, transportation, fertilizers, flavors, and fragrances for survival, but they also developed a significant amount of knowledge on how to make medicines from living organisms.<sup>1</sup> It has even been suggested that over 60,000 years ago, the Neanderthals may have used plants to treat common ailments, as palaeoanthropological studies near Kurdistan, Iraq revealed pollen deposits in graves.<sup>2</sup> However, the oldest written record of the use of plants or other extracts as medicines comes from Mesopotamia around 2600 B.C. and were written on clay tablets in cuneiform.<sup>3</sup> They describe over 1000 plant-derived substances, such as oils from both *Cedrus* sp. (cedar) and *Cupressus sempervirens* (cypress), *Glycyrrhiza glabra* (licorice), *Commiphora* sp. (myrrh), and *Papaver somniferum* (poppy juice), which were used to treat anything from a cold to parasitic infections, with the active ingredients from all of the aforementioned plants still used today.<sup>2,3</sup>

At around the same time period, other populations around the world were also learning to harness the potential of plants and animals as pharmaceutical agents. For example, the Egyptians writing of Elbers Papyrus (Egyptian pharmaceutical record) in 1500 B.C. consisted of over 700 drugs from mostly plants, which included agents from *Aloe vera* (aloe), *Boswellia carteri* (frankincense), and *Ricinus communis* (castor).<sup>4</sup> The Chinese had the writing *We Shi Er Bing Fang* (Prescriptions for fifty-two diseases, 1100 B.C.), which included 52 prescriptions.<sup>5</sup> This was followed up with several more thorough writings, the *Shennong Herbal* (100 B.C.) and *Tang Herbal* (659 A.D.), with the latter consisting of 850 drugs, along with information regarding their properties, efficacy, and synergies. With many of the therapeutic effects of these 850 drugs

subsequently confirmed, such as *Coptis chinensis* (anti-diarrhea), *Ephedra sinica* (anti-asthmatic), and *Melia azedarach* (anti-helminthic).<sup>6,7</sup> The people of the Indian sub-continent wrote the Charaka Samhita around 900 B.C. which contains 341 plant derived drugs,<sup>8</sup> while Hippocrates of Cos (460-377 B.C.), a Greek, wrote Corpus Hippocraticum after collecting more than 400 natural agents, including *Atropa belladonna* (anesthetic), and *Ornithogalum caudatum* (laxative).<sup>9</sup> Finally, Pedanius Dioscorides (40-90 A.D.) compiled the De Materia Medica, which consists of the dosage and efficacy of over 600 plant derived medicines and was the foundation of European pharmacology.<sup>9</sup> Interestingly, it appears that there was a convergent evolution of these different medicinal systems, where both the western and oriental populations were developing similar drugs to treat related diseases, with limited communications between them, such as Hippocrates using *Veratrum album* (white hellebore) and the Chinese using *Veratrum nigrum* (Black hellebore) both used as emetic.<sup>10</sup>



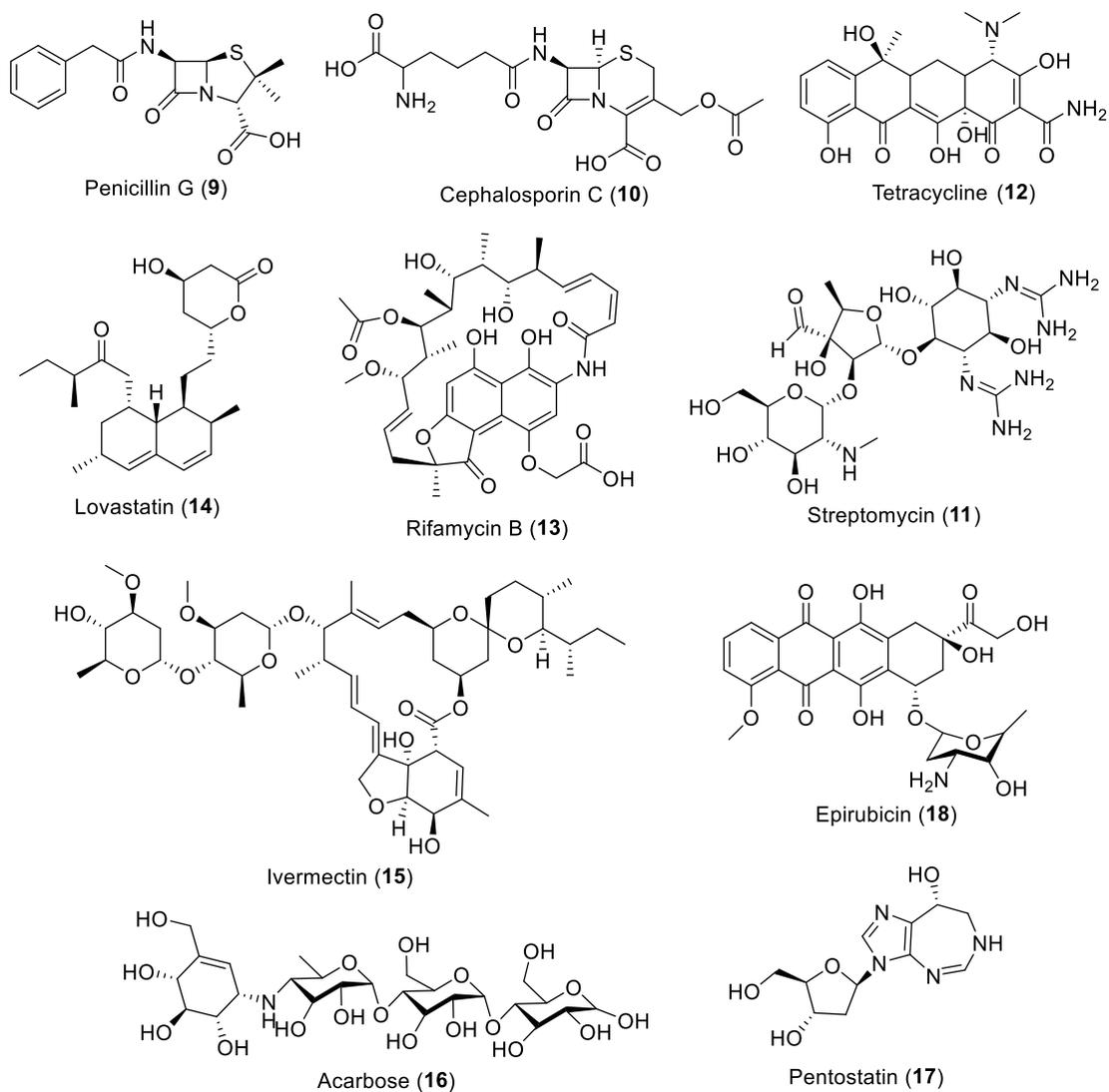
**Figure 1.1:** Plant alkaloids

### 1.1.2 Early Investigations into Terrestrial Natural Products

Although many civilizations were using whole or parts of plants to treat a host of diseases for thousands of years, it was not until the early 19<sup>th</sup> century before the active ingredients were being purified and investigated. This began the field of pharmacognosy, which is defined as the study of the physical, chemical, biochemical and biological properties of drugs with natural origins.<sup>11</sup> The first purification of a natural product occurred in 1803, when the alkaloid morphine (**1**) was purified from poppy seedpods (*Papaver somniferum*) by Friedrich Sertürner and was subsequently commercialized in 1826 by E. Merck.<sup>12</sup> This was followed by the isolations of colchicine (**2**) from meadow saffron (*Colchicum autumnale*), quinine (**3**) from the bark of the cinchona tree, strychnine (**4**) from Saint Ignatius bean (*Strychnos nux vomica*), and emetine (**5**) from ipecacuanha, all by Pierre-Joseph Pelletier and colleagues between 1817-1821.<sup>11</sup> Additionally, in 1820 caffeine (**6**) was purified by Friedlieb Ferdinand, followed by the purification of atropine (**7**) (1831) by Heinrich Friedrich Georg Mein, and cocaine (**8**) in 1860 by Albert Nieman.<sup>13</sup> These alkaloids have a range of activity, including as anti-cholinergics, analgesics, stimulants, anti-inflammatories, anti-amoebics, and toxins.

Up until 1940 plants were the primary organisms being chemically investigated, however this all changed with the isolation of the antibiotic, penicillin (**9**), from a fungus by Alexander Fleming.<sup>14</sup> Penicillin was the first drug to effectively target infections caused by gram-positive bacteria, such as *Staphylococci* and *Streptococci*, and its discovery revolutionized the field of medicine, while also opening the door for investigations into numerous other micro-organisms.<sup>14</sup> Since then, micro-organisms have been shown to produce other antibiotics, such as the cephalosporins (**10**) (*Acremonium*), aminoglycosides (**11**) (*Streptomyces*), tetracyclines (**12**) (*Streptomyces*), and polyketides [rifamycins (*Amycolatopsis mediterranei*)].<sup>15</sup> Additional notable natural products includes immunosuppressive agents [e.g. rapamycins (**13**)], cholesterol

lowering agents [e.g. lovastatin (**14**)], anti-hemintic agents [e.g. ivermectin (**15**)], anti-diabetic agents [e.g. acarbose (**16**)], and anti-cancer agents [e.g. pentostatin (**17**), and epirubicin (**18**)].<sup>16,17</sup>



**Figure 1.2:** Terrestrial microbial metabolites

To this day, the terrestrial environment is still being investigated for bioactive metabolites, with a renewed effort in following up on organisms used in traditional Chinese medicine. As of 2000, this had led to at least 122 active compounds, of which 80% are being used for the same (or related) ethnomedical purpose and are derived from only 94 plant species.<sup>18</sup>

Some scientists believe higher order plant species (angiosperms and gymnosperms) are perfect organisms to investigate because their uses are well documented, and thus it is unlikely to find metabolites with acute toxicity.<sup>18,19</sup> Furthermore, with the vast diversity of plants available, their chemical diversity should be equal or superior to that found in synthetic combinatorial chemical libraries.<sup>18</sup> However, other researchers estimate that the macroscopic terrestrial environment is running out of new chemical structural diversity, and it is therefore less likely to yield new drugs, especially from the higher plants.<sup>20,21</sup>

### 1.1.3 Impact of Natural Products on the Pharmaceutical Industry

Natural products have had a tremendous role in the development of new drugs since the establishment of the pharmaceutical industry, targeting a broad range of diseases. In the beginning many of the commercialized drugs had been described as ethnomedicines for hundreds of years, with their efficacies well documented. Many of these are hallmark drugs and are still widely used on a regular basis, like penicillin (**9**) (antibiotic), morphine (**1**) (analgesic), acetylsalicylic acid (**19**) (analgesic and anti-inflammatory), and ephedrine (**20**) (decongestant). Today, natural products still have a huge impact on drug development as lead compounds themselves, or as synthetic starting points, and this has been outlined nicely in a number of reviews by Newman and Cragg over the past ten years.<sup>16,22-24</sup>

Before discussing the impact of natural products on drug development, one must first define what the major categories of sources are, as there are many ways to classify the clinically used drugs. The categories described here closely follow those previously defined by Newman and Cragg and consist of two major sub-categories, small molecules and biologics. The small molecule category contains drugs that are either unmodified natural products (N), derived from a natural product (ND), completely of synthetic origin (S), or of synthetic origin but mimics the pharmacophore of a natural product (S/NM), with natural products involved in all but those of

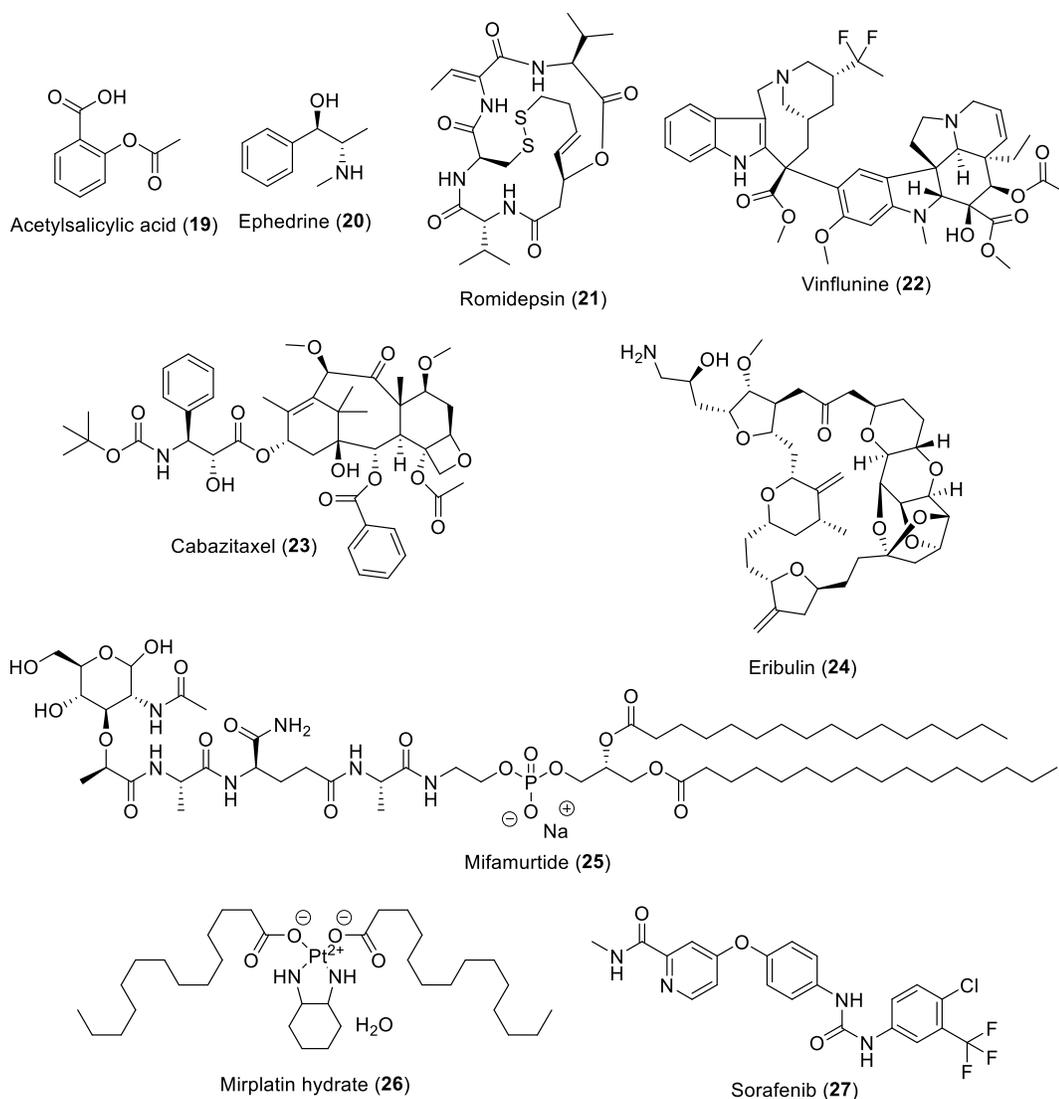
synthetic origin.<sup>22</sup> The biologics contain only large peptides (>45 residues), proteins, and vaccines, and are omitted in most of the following statistics in order to understand the impact natural products have on small molecule drug development.

Over a thirty year period (1981-2010), 1355 new chemical entities (NCE) were approved for use covering all diseases, countries, and sources of which 79.8% were considered small molecules.<sup>22</sup> Of these small molecules, only 29% were of synthetic origin, and thus natural products had a role in producing over 70% of these NCEs.<sup>22</sup> However, the importance of natural products varies across diseases with natural products playing key roles treating infectious diseases (microbial, parasitic, and viral), cancer, hypertension, and inflammation, but with no known metabolites effective as antihistamines, diuretics, or hypnotics. During this same thirty year time period, there were 104 small molecule NCEs developed as anti-bacterial (25% were synthetic), 28 agents as anti-fungal (90% were synthetic), 48 agents as anti-viral (43% were synthetic), 99 agents targeting cancer (20.2% were synthetic), and many more to treat a variety of other ailments.<sup>22</sup>

Looking back even further, 206 NCEs targeting cancer have been approved since the early 1940s, of which 175 are considered small molecules and only 25% are of synthetic origins.<sup>22</sup> This includes seven approved anti-cancer agents in 2010, consisting of two unmodified natural products [romidepsin (**21**), and polyphenon E], four derivatized natural products [vinflunine (**22**), cabazitaxel (**23**), eribulin (**24**), and mifamurtide (**25**)], and only one truly synthetic metabolite [miriplatin hydrate (**26**)]. Eribulin (Halaven), a natural product inspired compound, is likely the most complex drug that is completely produced by total chemical synthesis.<sup>25</sup>

Alarming, the number of small molecule NCEs entering the clinic have steadily declined after averaging about 40 per year between 1981-2000, to just over 20 between 2001-2010 (exception of 2002 and 2004). In 2010, only 20 small molecules NCEs were approved, the

second fewest in the 30 year time period; however half (10) were directly derived from a natural product.<sup>22</sup> This reduction in NCEs approval is thought to be a direct correlation to the shift in focus by big pharmaceutical companies away from drug discovery programs based around natural product leads and more in favor of combinatorial synthesis. This shift in focus occurred because of what were believed to be insurmountable hurdles in natural product based drug discovery research, such as the limited supply and structural complexity, as well as to the belief that because combinatorial synthesis yielded significantly more compounds, it would yield more new drugs.<sup>26</sup> Unfortunately this has failed miserably as almost 25 years of combinatorial synthesis has yielded only one approved drug, sorafenib (**27**), which is used to treat primary kidney cancer.<sup>22</sup> With the failure of combinatorial synthesis and significant improvements in the technology involved in the dereplication of natural products, there has been a significant return to natural products based drug discovery, with specific interests in small, structurally diverse, natural product libraries for high-throughput screening.<sup>26</sup>



**Figure 1.3:** A portion of the NCE approved between 1981-2010

## 1.2 Marine Natural Products

### 1.2.1 Early Investigations into Marine Natural Products

The world's oceans cover approximately 70% of Earth's surface, and certain ecosystems like the deep sea thermal vents and coral reef communities are estimated to have equal or higher species diversity than exist in the tropical rain forests.<sup>27</sup> Furthermore, organisms living in the marine environment have to deal with a number of unique stressors, like high salinity, low

nutrients, increased pressure, and aqueous surroundings. A large percentage of these organisms are relatively sessile and soft-bodied, for example tunicates, sponges, cyanobacteria, and macroalgae, thus making them extremely vulnerable to predation. However, as a result of the intense competition and the unique stressors, many of these marine organisms have evolved the capacity to protect themselves in the form of secondary metabolites.<sup>28</sup>

Unlike the terrestrial environment, chemical investigation into the marine realm is still a relatively young field. This is because of the lack of accessibility and the preconceived notion that the oceans lacked species diversity. However, in 1942 Jacques Cousteau and Emile Gagnan co-invented the modern self-contained underwater breathing apparatus (SCUBA), which gave divers access to a significant portion of the marine environment that was previously inaccessible.<sup>29</sup> SCUBA allowed scientists to easily make collections in waters up to 120 feet deep, which includes many interesting habitats and a large diversity of organisms.

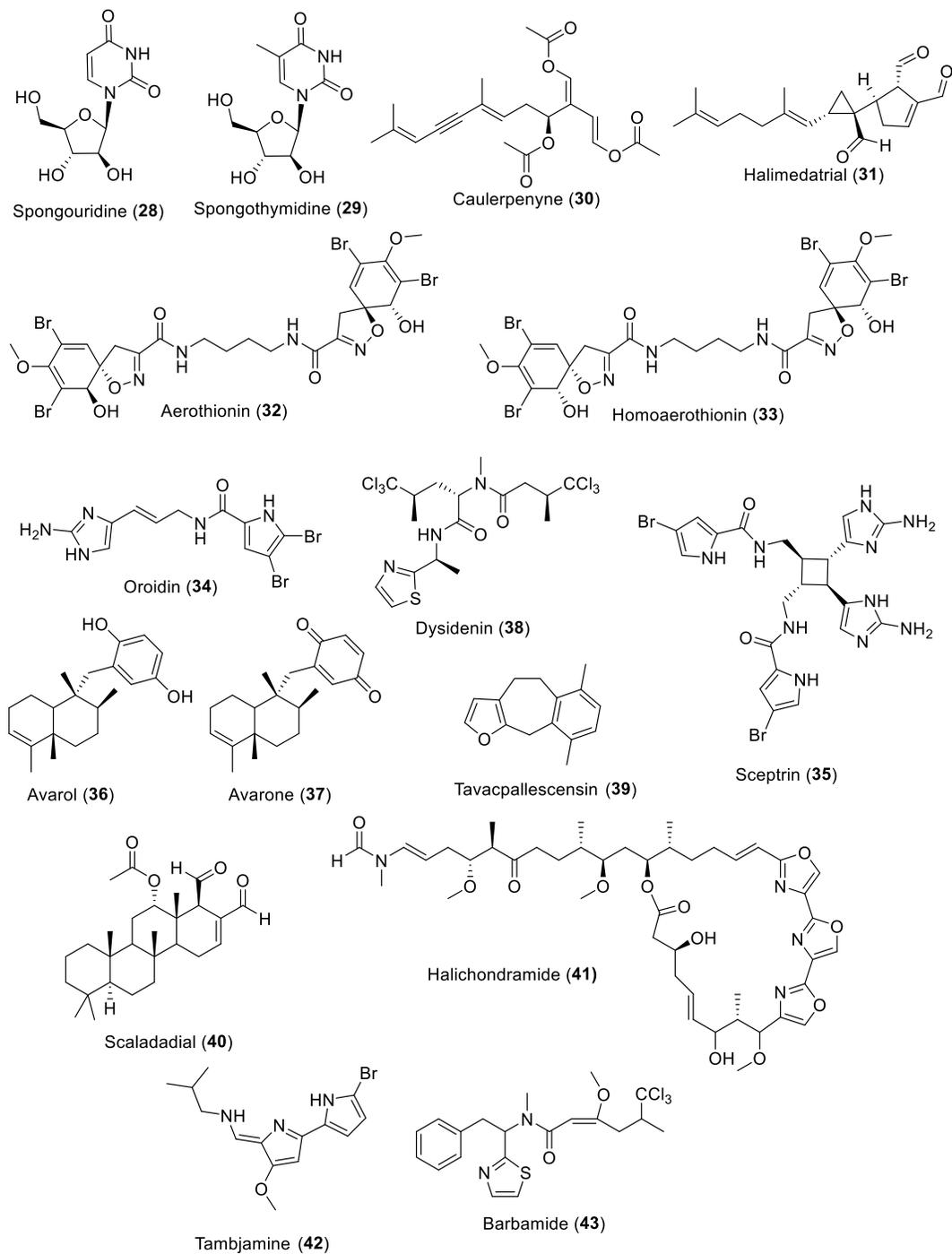
Some of the first documented metabolites that were discovered from a marine organism were reported in 1951 by Bergmann *et al.* after they discovered several unusual arabino-nucleosides [e.g. spongouridine (**28**), and spongothymidine (**29**)] from the marine sponge *Cryptotheca crypta* collected off the Florida coast.<sup>30-32</sup> Two analogs of these new metabolites, cytarabine, and vidarabine, have been used for years to clinically treat cancer and viral infections, respectively.<sup>33,34</sup> This example superbly demonstrates the extraordinary potential marine organisms possess to biosynthesize structurally intriguing and biologically active secondary metabolites, and ignited the field of marine natural products drug discovery.

After the discovery of the arabino-nucleosides, chemical investigations into other organisms such as tunicates, mollusks, soft coral, and macroalgae began, as these organisms grow prolifically in shallow tropical waters and are easily collected. Over the years a multitude of compounds have been isolated from these organisms with the identification of unique chemical species belonging to specific families of organisms. For example, hundreds of halogenated

terpenes and C<sub>15</sub>-acetogenins have been characterized from red algae, in particular the genus *Laurencia*.<sup>35</sup> Sesquiterpenes featuring a di-aldehyde moiety [e.g. caulerpenyne<sup>36</sup> (**30**) and halimedatrial<sup>37</sup> (**31**)] are hallmark metabolites from green algae,<sup>35</sup> while biologically important prostaglandins are found in the soft coral gorgonian *Plexaura homomalla*.<sup>38</sup>

Although, investigations into sponges, tunicates, and mollusks have shown these organisms to be extraordinarily prolific sources of secondary metabolites, there appears to be no chemical conformity in their observed metabolome. For example, early compounds isolated from sponges include several bromotyrosine-derived molecules [e.g. aerothionin<sup>39</sup> (**32**) and homoaerothionin<sup>40</sup> (**33**)], bromopyrroles [e.g. oroidin<sup>41</sup> (**34**) and sceptrin<sup>42</sup> (**35**)], hybrid isoprenoids [e.g. avarol<sup>43</sup> (**36**) and avarone<sup>44</sup> (**37**)], and trichloropeptides [e.g. dysidenin<sup>45</sup> (**38**)]. As for mollusks, degraded terpenoids [e.g. tavacpallescensin<sup>46</sup> (**39**)], sesterterpenes [e.g. scaladadial<sup>47</sup> (**40**)], polyketides [e.g. halichondramide<sup>48</sup> (**41**)], and bromopyrroles [e.g. tambjamine<sup>49</sup> (**42**)] have all been isolated. Over time it has become apparent that many, of these compounds are likely produced by other sources such as algae, bacteria, or cyanobacteria. Evidence of this is seen by the structural similarity of these compounds to other known metabolites produced by these microbes. For example, the dysidinins were originally isolated from the sponge *Dysidea herbacea*,<sup>45</sup> but subsequently a related metabolite, barbamide (**43**),<sup>50</sup> was found in a collection of the cyanobacterium *Moorea producens*, and the isolation of scaladadial, a 1,4-di-aldehyde, from the nudibranch *Glossodoris pallida*,<sup>47</sup> is structurally reminiscent of numerous green algal metabolites.<sup>35</sup> Conceptually, this overlap in secondary metabolites is reasonable, as both tunicates and sponges acquire their food by filtering seawater; thus, they come into contact with a multitude of different microorganisms and are also known to live symbiotically with cyanobacteria and bacteria. Similarly, mollusks graze on a variety of algae and other organisms, and seem to have the ability to sequester secondary metabolites from their diet.<sup>51,52</sup> This key observation initiated a change in the way marine natural product

researchers selected organisms to investigate, as it is always the goal to focus attention on organisms that are prolific producers of secondary metabolites, in order to continue to isolate new and interesting compounds, as well as facilitate ensuing studies on biosynthesis.



**Figure 1.4:** Early marine natural products

### 1.2.2 Modern Day Investigations into Marine Natural Products

From around the early 1980's, investigators have commented on the structural similarities of metabolites isolated from both invertebrates and microbes, and in some cases between organisms that inhabit very different ecosystems such as a marine sponge and a terrestrial beetle.<sup>53</sup> These observations led to the hypothesis that many of the true producers of secondary metabolites in the marine environment were microbes (bacteria, cyanobacteria, and fungi), that either lived symbiotically or were grazed upon by invertebrates. Recognition that microbes were responsible for producing thousands of structurally distinct metabolites, many with potent biological properties, shifted the focus of marine natural product chemists away from large invertebrates to study instead their microbial associates.

This ideological shift was made possible in part because of improvements in technologies that have significantly improved the isolation and characterization of novel metabolites, some of which are present in very small quantities. Although, some of these microbes, such as filamentous marine cyanobacteria, are able to be collected in the field, only relatively small quantities are generally found in any given location. Even worse, studying either symbiotic or free-living unicellular bacteria requires laboratory cultivation, which has only been successful for a small percentage of isolates. Thus, researchers have found it necessary to learn how to work with significantly smaller quantities of the natural products.

Some of the critical technological advancements that have been important to this development in natural products were actually improvements to existing instrumentation. Examples include improving the sensitivity of mass spectrometers (MS) and nuclear magnetic resonance (NMR) spectrometers. Today, the magnets in the NMRs have become so powerful that in 2009, Dalisay and Molinski reported the complete structure elucidation of hemi-phorboxazole A (**44**), a large cyclic macrolide, with only 16.5  $\mu\text{g}$  of the natural product.<sup>54</sup> Along with the increase in magnet strengths, probes have been designed to both reduce the amount of solvent

from approximately 500  $\mu\text{L}$  (5 mL NMR tube) to less than 40  $\mu\text{L}$  (1.7 mm NMR tube) and cooled with liquid nitrogen to reduce electronic noise (CryoProbe and Cold Probe), which greatly reduces the amount of sample needed.<sup>55-57</sup> Additionally, numerous experiments have been developed to aid in the structure elucidation process, such as heteronuclear 2 bond correlation (H2BC)<sup>58</sup> and HETLOC.<sup>59</sup>

As for MS systems, they too have become more sensitive, but more important has been the development of algorithms to analyze large datasets acquired over either single or multiple runs. An example is the nonribosomal peptide dereplication and sequencing algorithm, which uses both  $\text{MS}^2$  and  $\text{MS}^3$  fragmentation of a target peptide to compare with compounds in a fragment library. The algorithm then identifies the location and molecular weight of the modification using the top-scoring peptide in the Norine database.<sup>60</sup> This works very well for new metabolites that possess only minor structural differences from a known family of compounds; in addition, because it is MS-based, it only requires nanogram quantities of an impure mixture. Another useful MS program is called ‘molecular networking’, as it uses  $\text{MS}^1$  and  $\text{MS}^2$  data to compare all metabolites in a sample or several samples and clusters them based on relatedness in their fragmentation patterns.<sup>61</sup> This allows visualization of metabolites in extracts and quickly identifies analogs or biosynthetic precursors of a known metabolite. These improvements in technology have allowed for quick dereplication of fractions and simplified the task of deducing planar structures of new metabolites.

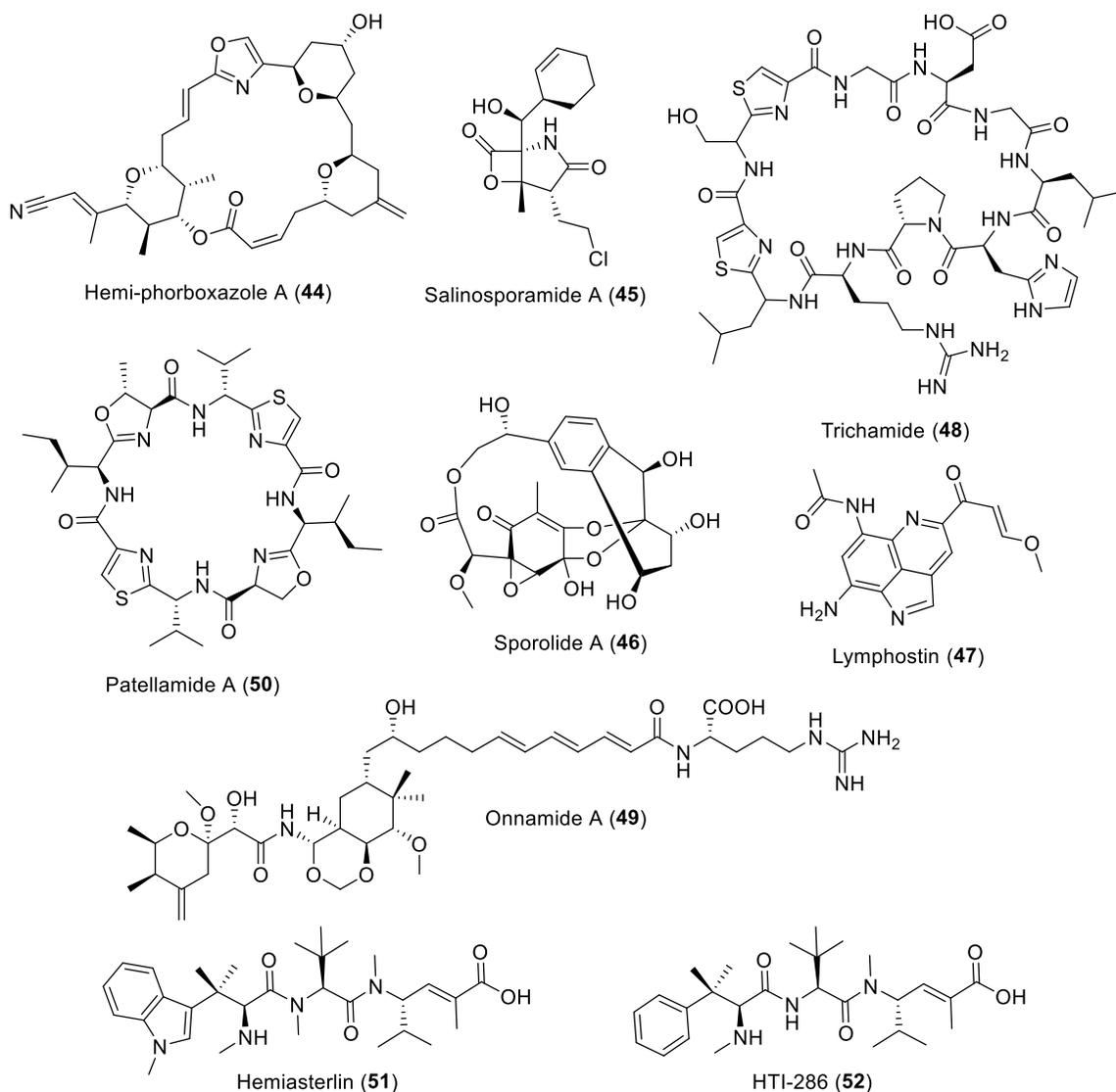
In the last 10 years, the emergence of faster and cheaper DNA sequencing protocols has enabled the sequencing of both individual genomes and meta-genomes, thus yielding a better overview of the biosynthetic capacity a particular microbe has to produce secondary metabolites. In 2007, the first genome of a marine actinomycete, *Salinospora tropica*, was completed and subsequently confirmed *S. tropica*'s exceptional biosynthetic ability. Approximately 10% of its genome encodes for the production of secondary metabolites, with at least 17 distinct pathways.<sup>62</sup>

Knowledge gained from the genome has facilitated molecular cloning of salinosporamide A (**45**),<sup>63</sup> sporolide (**46**),<sup>64</sup> and lymphostin (**47**).<sup>65</sup> Furthermore, genomic information on organisms with compounds yet to be discovered has created a new protocol called ‘genome mining’, which is the targeted isolation of predicted metabolites from the genome.<sup>66</sup> Good examples are the prochlorosins from *Prochlorococcus* MIT9313<sup>67</sup> and the ribosomal peptides trichamide (**48**) from *Trichodesmium erythraeum* ISM101.<sup>68</sup> Still, a significant portion of all bacteria have not been cultured and many live symbiotically with invertebrates or other macroorganisms, thus complicating the process to obtain their genomic material. However, this obstacle can be overcome by metagenomics, which is based on the relatively unbiased sequencing of the total environmental DNA, which in this case may involve a marine organism and its associated microflora.<sup>69</sup> This method can also unequivocally prove the identity of the true producing organism of a secondary metabolite that was isolated from an assortment of organisms. Good examples of this are Jorn Piel’s work on onnamide (**49**) from a sponge<sup>70</sup> and Eric Schmidt’s work on the cyanobactin (**50**) family of metabolites from ascidians.<sup>71</sup>

The development and improvement of chemical synthesis has also had an important role in natural products chemistry as it aids in stereochemical analysis, provides a reliable supply of material for biological assays, and it opens an avenue to improve the druggability of a new metabolite.<sup>72</sup> Commonly, fragments of natural products are synthesized in order to use as authentic standards for LCMS or GCMS analysis to assist in determining absolute stereochemistry of portions of a new metabolite. These fragments generally involve only a couple of linear synthetic steps, such as making unusual hydroxy acids or short polyketide fragments. Much more important to the pharmaceutical industry is the total synthesis of bioactive metabolites. In this case, it is critically important that the number of linear steps is minimized, and the overall yield is maximized, in order to provide an economical and reliable supply of the drug for clinical evaluation.<sup>73</sup> Some of the early total synthesis involved extraordinarily complex

natural products, and although their syntheses were often quite elegant in design, they tended to involve a significant number of steps with extremely low overall yields.<sup>74</sup>

More recently though, there seems to be an understanding of the importance of an intelligently designed synthesis of a bioactive metabolite has on it reaching clinical trials.<sup>73</sup> An example involves hemiasterlin (**51**), a tri-peptide that was isolated by the Andersen lab back in 1994 from the sponge *Hemiasterella minor*. Hemiasterlin was exquisitely potent against murine leukemia P388 cells with an IC<sub>50</sub> of just 87 pM; however, it was only isolated in very small quantities, and thus a total synthesis was needed to further evaluate its cytotoxic activity.<sup>75</sup> The Andersen group designed a relatively short convergent synthesis that yielded hemiasterlin as well as numerous other analogs for developing an understanding of structure activity relationships. From the material obtained via total synthesis, further biological evaluations were conducted, and revealed that hemiasterlin inhibits the spindle microtubule dynamics at mitosis.<sup>76</sup> Also, the SAR study identified key structural elements that are critical for the potent activity and yielded one analog [HTI-286 (**51**)] which had increased potency along with a slightly easier synthetic protocol.<sup>77</sup> With a reliable source of HTI-286 from total synthesis, it progressed into clinical development with Wyeth and eventually made it to Phase II clinical trials.<sup>72</sup>

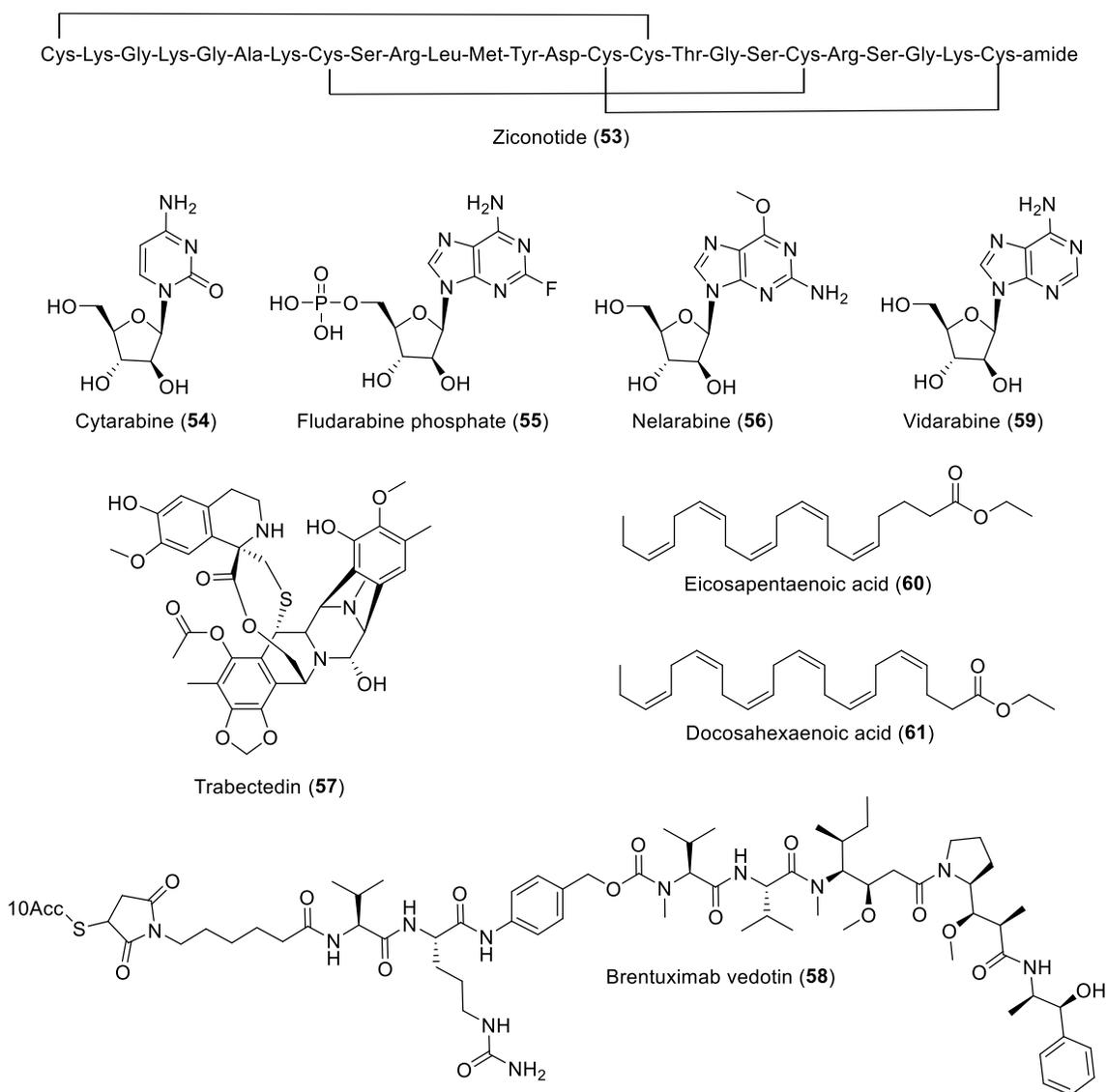


**Figure 1.5:** Modern marine secondary metabolites

### 1.2.3 Marine Derived Drugs

Although the study of marine natural products is a relatively young field, it has already yielded eleven FDA approved drugs. These drugs treat a host of diseases such as pain [ziconotide (**53**)], cancer [cytarabine (**54**), fludarabine phosphate (**55**), nelarabine (**5**), eribulin (**24**), trabectedin (**57**), brentuximab vedotin (**58**), and hemocyanin (KLH)], viral infections [vidarabine (**59**)], hypertriglyceridemia [omega-3-acid ethyl esters (**60,61**)], and coagulants (protamine

sulfate) and were isolated from a diverse group of organisms including sponges, a cone snail, fish, tunicates, and a mollusk, although many are predicted to be of microbial biosynthetic origin.<sup>78</sup> As of early 2012, there was one marine-derived drug in phase III, six in phase II, and seven in phase I clinical trials.<sup>78</sup> The vast majority of these drugs are being evaluated for the treatment of cancer; however two drugs are being evaluated against schizophrenia and for wound healing.<sup>78</sup> Of the fifteen drugs that are either approved or in clinical trials targeting cancer, they have eight different molecular targets, including microtubules, 20S proteasome, protein kinase C, and DNA binding.<sup>78</sup> The pharmaceutical success that marine natural products have obtained in such a short period is approximately 1.7- to 3.3-fold better than the industry average, with approximately one drug per 3,140 described marine natural products versus one drug to every 5,000-10,000 compounds screened, respectively.<sup>78,79</sup>



**Figure 1.6:** Marine derived drugs

#### 1.2.4 Investigations into Marine Cyanobacteria

Cyanobacteria are amongst the oldest life forms on earth and are ubiquitously distributed among all ecosystems. In the marine environment, cyanobacteria are known to play an important ecological role as both a carbon source via photosynthesis and as nitrogen fixers.<sup>80</sup> A particular subset, the marine filamentous cyanobacteria, have been shown to be prolific producers of natural products, especially those containing nitrogen atoms.<sup>81</sup> Filamentous marine cyanobacteria grow

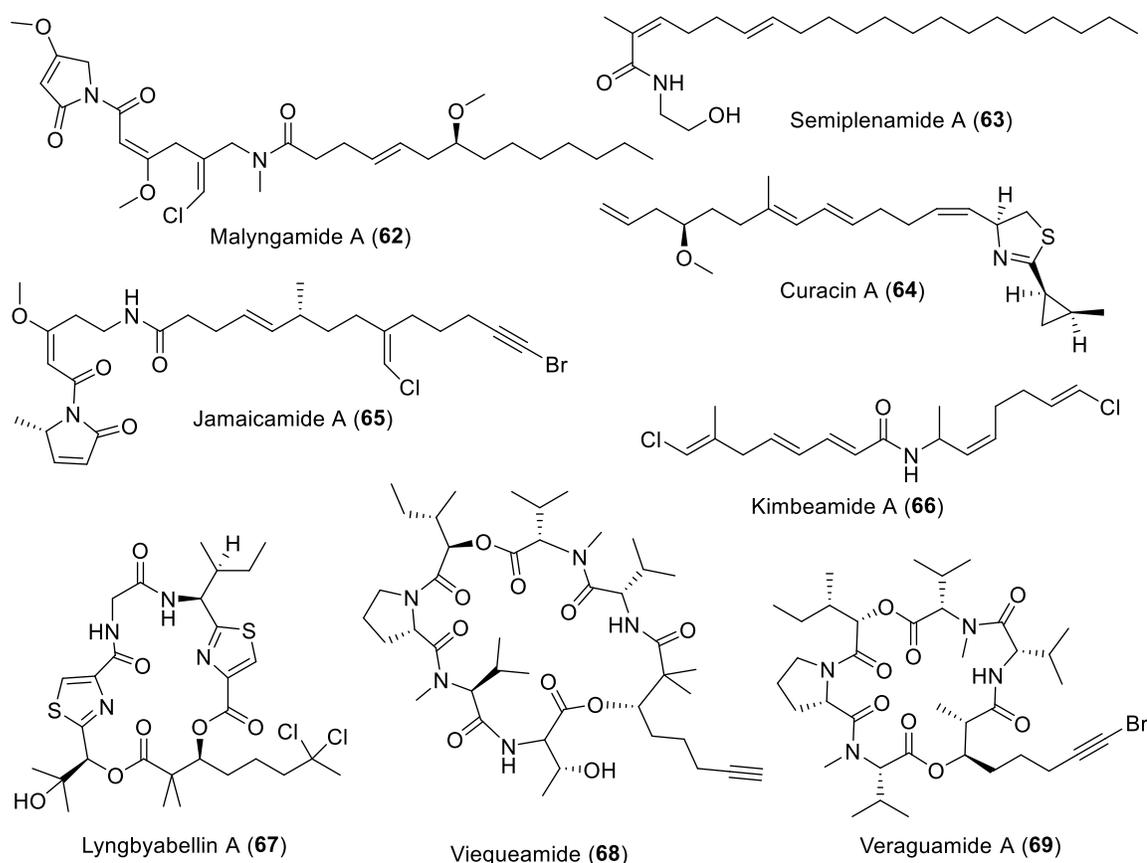
abundantly in shallow tropical waters around the world, and can grow attached to nearly any surface including, algae, rocks, sand, reefs, and mangrove roots.

Over 800 compounds have been isolated from filamentous marine cyanobacteria, with the majority coming from the genera *Moorea* (formerly *Lyngbya*),<sup>82</sup> *Oscillatoria*, and *Symploca*.<sup>83</sup> The predominant theme in these metabolites is the incorporation of nitrogen and that they are produced by either the polyketide synthase (PKS), the nonribosomal peptide synthetase (NRPS), or a mixture of the two, PKS/NRPS, biosynthetic pathways.<sup>84</sup> The metabolites of mixed biosynthesis belong to two different subfamilies of lipopeptides, the alkyl amides and the modified peptides. Alkyl amides generally consist of an acetate-derived fatty acid portion that is coupled through an amide bond to a variety of amines in linear fashion, as seen in the malyngamides (**62**),<sup>85</sup> semiplenamides (**63**),<sup>86</sup> curacin A (**64**),<sup>87</sup> jamaicamides (**65**),<sup>88</sup> and kimbeamides (**66**).<sup>89</sup> Common post-translational modifications observed in this class of metabolites are C- and N- methylation, halogenation, and cyclization.<sup>83,84</sup>

The NRPS/PKS derived peptides consist primarily of amino acids, however they also incorporate at least one acetate containing residue. Neutral L-amino acids (Val, Ala, Phe, Ile, Tyr, Pro, Cys, Ser, Leu, and Gly) are predominantly incorporated; however, it is not uncommon to find a D-form amino acid as well, which indicates that an epimerization has occurred during biosynthesis.<sup>81</sup> Other common modifications to the peptide portion of these molecules are the incorporation of hydroxy acids, N- and S- methylation, hydroxylation, cyclization, and oxidation.<sup>81</sup> As for the acetate-derived portion, one of the more common residues consist of a  $\beta$ -hydroxy/amino unit ranging in carbon chain length from 4-carbons (3-amino-2-methyl-butanoic acid moiety) to 12-carbons (3-amino-2,5-dihydroxy-dodecanoic acid moiety).<sup>90,91</sup> A unique feature of this moiety in cyanobacteria is the oxidation of the tail to either an alkene or alkyne, with the latter being subsequently brominated in some cases. Furthermore, this moiety is commonly methylated, with either mono- or di-methylated at the  $\alpha$ -position [lyngbyabellin (**67**),<sup>92</sup>

and viequeamide<sup>93</sup> (**68**) and halogenated [lyngbyabellin,<sup>92</sup> and veraguamides<sup>93,94</sup> (**69**)]. The incorporation of a number of modified amino acids and various forms of the polyketide portion yields the large structural diversity observed within this class of metabolites.

Most of the efforts to discover new metabolites from cyanobacteria have had a primary focus to identify metabolites with interesting biological activity. This has largely been accomplished by the early incorporation of biological screening of semi-crude fractions in a variety of assays, including cancer cell toxicity, anti-inflammatory, and molluscicidal. Once a fraction exhibits some type of activity, then that sample is further fractionated following a bioassay-guided fractionation scheme, until the pure active component has been identified. In this regard, numerous compounds have been discovered that have potentially useful activities, such as anti-cancer, anti-inflammation, parasitic, modulation of ion channels, receptor binding capabilities, and many more.<sup>83,95</sup>



**Figure 1.7:** Cyanobacteria natural products

## 1.2.5 Therapeutic Potential of Secondary Metabolites from Marine Cyanobacteria

### 1.2.5.1 Cancer Cell Cytotoxicity

In the U.S., cancer is the second leading cause of death behind only heart disease, and affects over thirteen million people, with prostate and breast cancer as the two most prevalent cancers afflicting men and women, respectively.<sup>96</sup> Between 1981 and 2010, 128 new anti-cancer agents were approved for use worldwide, and greater than 84% of these drugs were either natural products, natural product botanicals, natural product-derived compounds, vaccines, botanicals, or synthetics but mimicking of a natural product.<sup>97</sup> These drugs have only a handful of mechanism of actions, including, as alkylating agents {attach an alkyl group to *N*-7 of guanine [e.g. streptozocin<sup>98</sup> (**70**)]}, anti-metabolites {mimic natural building blocks of DNA [e.g. cytarabine<sup>99</sup>

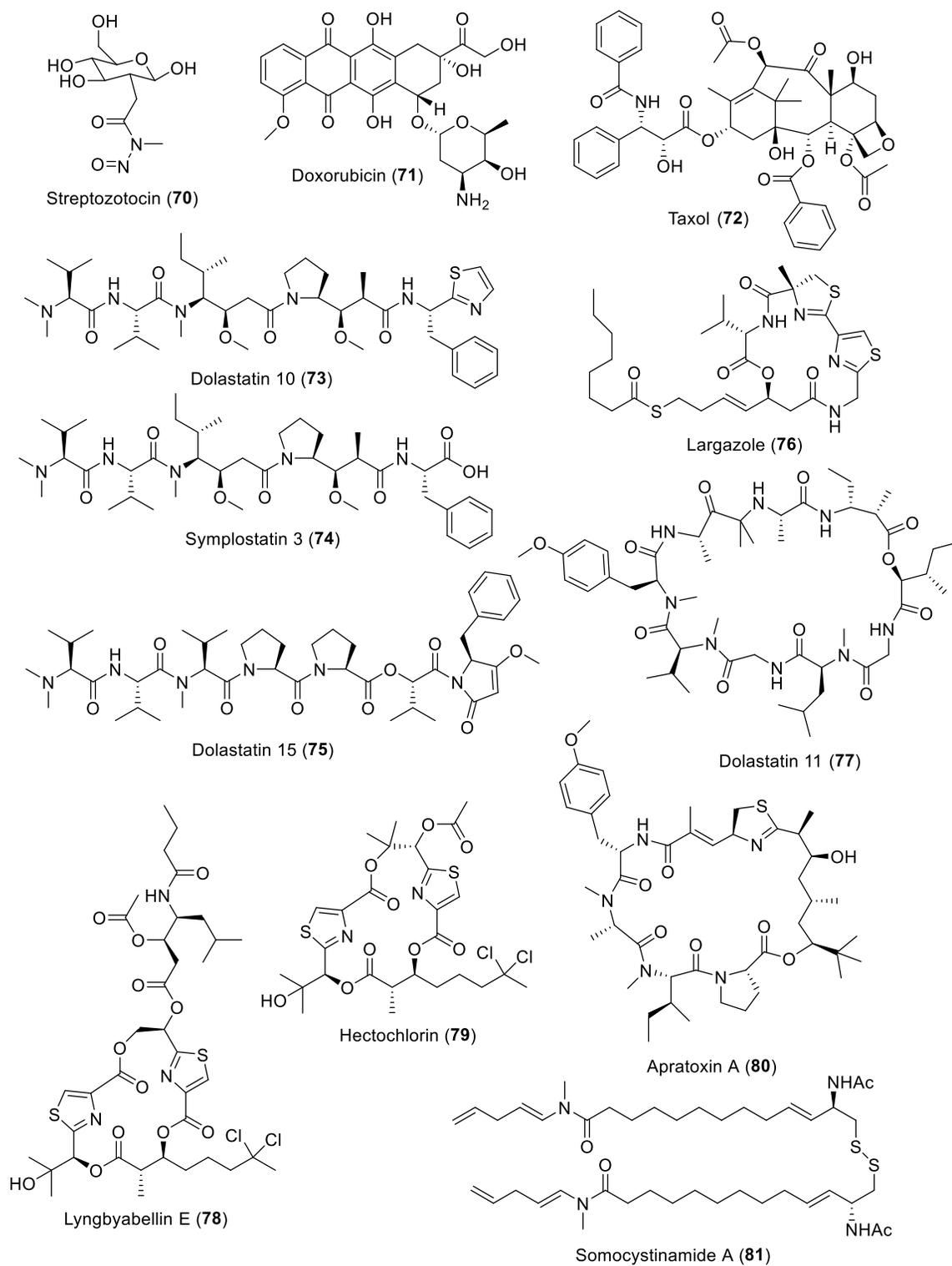
(**54**]], DNA intercalaters {reversible inclusion between DNA base pairs [e.g. doxorubicin<sup>100</sup> (**71**)]}, and as anti-mitotics {disrupt normal microtubules formation [e.g. taxol<sup>101</sup> (**72**)]}.<sup>102</sup>

Bretuximab vedotin (**56**) is the only cyanobacterial-derived drug with current FDA approval, and it is used to treat both Hodgkin's lymphoma and systemic anaplastic large cell lymphoma.<sup>103</sup> The natural product portion is derived from dolastatin 10 (**73**),<sup>104</sup> a linear lipopeptide of mixed biosynthetic origin, which has been linked to an antibody forming an antibody drug conjugate (ADC). There are other cyanobacterial-derived drugs in the clinical pipeline, with one in phase II (glembatumumab vedotin), two in phase I (SGN-75 and ASG-5ME)), and several in preclinical trials, most of which are of mixed NRPS/PKS biosynthetic origin.<sup>78</sup>

Two common mechanisms of action (MOA) of cyanobacterial natural products are the disruption of microtubules and actin filaments, both of which are involved in mitosis and thus blocking cell division.<sup>83</sup> There are several well-known microtubule inhibiting cyanobacterial metabolites that exhibit potent cytotoxicity including curacin A (**64**) [IC<sub>50</sub> = 9 nM (L1210)],<sup>87</sup> symplostatins 3 (**74**) [IC<sub>50</sub> = 3.9 nM (KB)],<sup>105</sup> dolastatin 10 (**73**) [IC<sub>50</sub> = 0.059 nM (P388)],<sup>104</sup> dolastatin 15 (**75**) [IC<sub>50</sub> = 0.13 nM (HT)],<sup>106</sup> and largazole (**76**) [IC<sub>50</sub> = 7.7 nM (MDA-MB-231)]<sup>107</sup>. Similarly, there are numerous compounds that exhibit potent cytotoxicity by inhibiting the formation of actin filaments, such as dolastatin 11 (**77**) [IC<sub>50</sub> = 47 nM (PtK1)],<sup>108</sup> lyngbyabellin E (**78**) [IC<sub>50</sub> = 400 nM (H460)],<sup>109</sup> and hectochlorin (**79**) [IC<sub>50</sub> = 20 nM (CA46)]<sup>110</sup>. Interestingly, most of these metabolites that target microtubules and actin are of mixed NRPS/PKS biosynthetic origin; however, there is a key structural distinction between these two groups of compounds, as metabolites that inhibit microtubules are primarily linear lipopeptides and those targeting actin are cyclic depsipeptides.

There are other cyanobacterial metabolites that also exhibit potent activity against a range of cancer cell lines and have a variety of mechanisms of action, including the apratoxins

[apratoxin A (**80**):  $IC_{50} = 0.52$  nM (KB)], which are highly cytotoxic, modified cyclic depsipeptides, with two polyketide sections among several amino acids.<sup>111</sup> They appear to potentially have a multitude of targets including the Heat Shock Protein (HSP),<sup>112</sup> as well as a secretory pathway.<sup>113</sup> Somocystinamide A (**81**), a disulfide dimer of an alkyl amide, selectively induces apoptosis in cancer cell lines that express caspase 8 [ $IC_{50} = 14$  nM (CEM)].<sup>114</sup> Coibamide A (**82**), a large cyclic depsipeptide containing eleven residues with significant *O*- and *N*- methylation, exhibits exceptional anti-proliferative activity against several cancer cell lines [ $IC_{50} = 2.8$  nM (MDA-MB-231)], but with an unknown biochemical target.<sup>115</sup>



**Figure 1.8:** Bioactive cyanobacteria metabolites (part 1)

### 1.2.5.2 Other Biologically Important Activity Observed from Cyanobacterial Metabolites

Cyanobacteria secondary metabolites have also been shown to exhibit a broad range of pharmaceutically pertinent activities, including as anti-inflammatory, anti-infective, neurotoxic, or neuro-receptor binding agents.<sup>95</sup> Discovery of these different types of activities came about by primarily two different isolation protocols, one of which involves thoroughly screening cyanobacterial crude extracts in as many assays as possible to obtain a 'hit' and then following up on that 'hit' with bioassay guided fractionation to yield a pure active metabolite. The other involves testing a pure compound in an assay based on structural similarities between the new metabolite and that of a known drug or endogenous ligand.

In recent years, there have been a number of metabolites that have shown potent and mechanistically intriguing anti-inflammatory activities, such as several of the malyngamides, particularly ones in the F series. In a nitric oxide (NO) inhibition assay using a mouse RAW macrophage cell line, malyngamides F (**83**) depressed Interleukin 1 and 6 (IL 1 and 6) and enhanced Tumor Necrosis Factor  $\alpha$  (TNF $\alpha$ ), thus working through a MyD88-independent pathway.<sup>116</sup> Another family of anti-inflammatory metabolites are the honaucins, which were isolated from a bloom-forming *Leptolyngbya crossbyana* off the coast of Hawaii (the big island).<sup>117</sup> These metabolites are uncharacteristically low in molecular weight and are entirely of PKS origin with the major metabolite, honaucin A (**84**), consisting of two fragments, (S)-3-hydroxy- $\gamma$ -butyrolactone and a 4-chlorocrotonic acid. From a rather thorough structure activity relationship (SAR) study on the honaucins, it appears that most modifications diminish the anti-inflammatory activity except for the replacement of the allylic chlorine with more electron rich bromine or iodine atoms, these substitutions increased the activity [iodo-honaucin A: IC<sub>50</sub> = 0.9  $\mu$ M; bromo-honaucin A: IC<sub>50</sub> = 1.5  $\mu$ M; honaucin A: IC<sub>50</sub> = 4.0  $\mu$ M].<sup>117</sup>

Another emerging theme among cyanobacterial secondary metabolites is the production of neurotoxic substances, with many appearing to target the Voltage Gated Sodium Channel

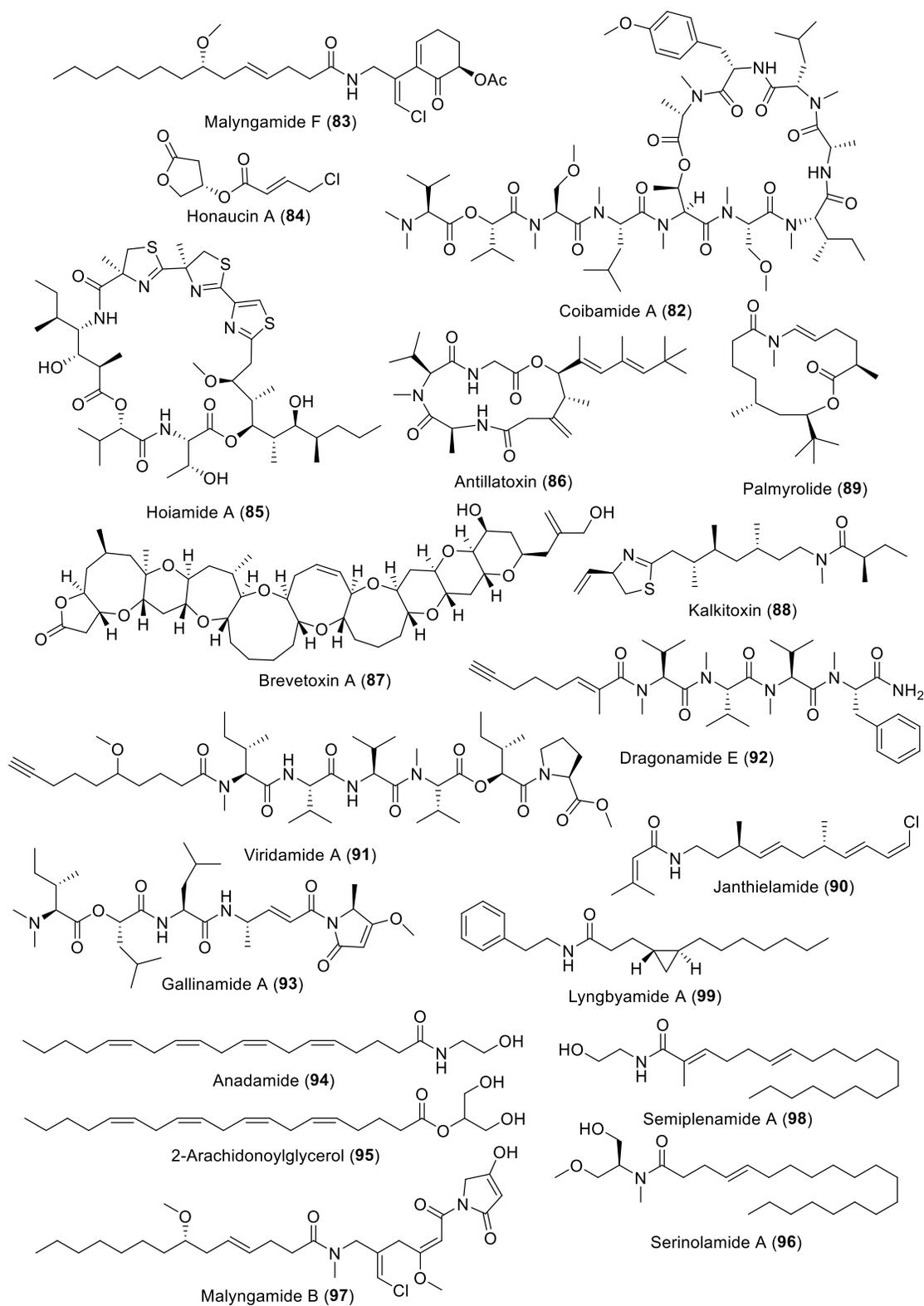
(VGSC). Metabolites that target the VGSC are thought to have a number of potential therapeutic effects, such as antiarrhythmic, enhancement of neurite outgrowth, treatment of cystic fibrosis, and as epilepsy treatments.<sup>95</sup> One such family of metabolites is the hoiamides, which are large NRPS/PKS hybrid peptides consisting of eight residues, including a 15-carbon-long linear polyketide and a tri-heterocyclic ring system.<sup>118</sup> Hoiamide A (**85**) and B both stimulate sodium flux ( $IC_{50} = 1.7 \mu\text{M}$  and  $3.9 \mu\text{M}$ , respectively) and potently suppress spontaneous calcium oscillations ( $EC_{50} = 45.6 \text{ nM}$  and  $79.8 \text{ nM}$ , respectively) in mouse neocortical neurons.<sup>119</sup> Antillatoxin (**86**), another hybrid peptide containing four amino acid residues, including a PKS residue that has seven methyl groups, is one of the more potent activators of the VGSC ( $EC_{50} = 20.1 \text{ nM}$ ).<sup>120</sup> Furthermore, it appears to interact at a distinctly different site than that of the marine dinoflagellate toxin, brevetoxin B (**87**).<sup>121</sup> On the other hand, the cyanobacterial metabolite kalkitoxin (**88**), which is a linear lipopeptide featuring four secondary methyl groups and a thiazoline ring, exhibits extraordinarily low VGSC blocking activity ( $EC_{50} = 1 \text{ nM}$ ).<sup>122</sup> Other neurotoxic metabolites include palmyrolide A (**89**),<sup>123</sup> janthielamide (**90**), and kimbeamides A-C (**66**).<sup>89</sup>

Another interesting pharmaceutical potential for these metabolites is as anti-infectives, particularly against several neglected diseases such as malaria, leishmania, and Chagas disease. These diseases are caused by the parasites *Plasmodium falciparum* (malaria), *Leishmania donovani* (leishmania), and *Trypanosma cruzi* (Chagas) and infect more than 2 million people worldwide each year.<sup>124</sup> Through efforts of the International Cooperative Biodiversity Group (ICBG) in Panama,<sup>125</sup> a handful of potential lead compounds have been identified for the treatment of these neglected diseases. Each of these metabolites are NRPS/PKS hybrid linear lipopeptides, ranging in size from five to seven residues and containing a high degree of post-assembly modifications. For example, viridamide A (**91**) has three *N*- and two *O*- methyl groups, along with a ten carbon PKS chain that has a terminal alkyne.<sup>126</sup> Similarly, dragonamide E (**92**) is highly methylated

(four *N*-methyls) but also has a *C*-terminal amide.<sup>127</sup> While viridamide A exhibits activity against *P. falciparum* (IC<sub>50</sub> = 5.8 μM), *L. donovani* (IC<sub>50</sub> = 1.37 μM), and *T. cruzi* (IC<sub>50</sub> = 1.0 μM), dragonamide E only exhibited activity against *L. donovani* (IC<sub>50</sub> = 5.1 μM).<sup>126,127</sup> The smallest of these anti-infective agents is gallinamide A (**93**), which was isolated from a collection of *Schizothrix* sp. in the Portobelo Marine Park on Panama's Caribbean coast.<sup>128</sup> Gallinamide A also exhibits activity against *P. falciparum* (IC<sub>50</sub> = 8.4 μM), however more recently it was shown to be a potent inhibitor of the human cysteine cathepsin L protease (IC<sub>50</sub> = 5.0 nM).<sup>129</sup> The latter activity gives insights into what the potential MOA for the observed toxicity toward *P. falciparum* could be, as cysteine proteases are responsible for converting pro-neuropeptides to active neuropeptides through nucleophilic attack by the cysteine thiol on the carbonyl groups in proteins, thus catalyzing the hydrolysis of amide bonds. A functionality that is structurally significant to gallinamide A is the α,β-unsaturated ketone, which is derived from an acetate extended alanine residue. This functionality is prone to nucleophilic addition via a Michael's reaction, and thus it is possible that a similar reaction is occurring in *P. falciparum*, as proteases are ubiquitously found in all life forms.<sup>130</sup>

Finally, a class of cyanobacterial natural products that has been increasingly reported and of growing physiological and pharmacological importance, is the alkyl amides. One rather intriguing biological target that these alkyl amides seem to target are neuro-receptors, more specifically, the cannabinoid receptors. There are two known subtypes of this G protein-coupled receptor (GPCR), CB<sub>1</sub> and CB<sub>2</sub>, which are primarily localized in the central nervous (CNS) and immune systems, respectively.<sup>131</sup> The endocannabinoid system (ECS) is known to have a host of important biological functions including appetite regulation,<sup>132</sup> development,<sup>133</sup> learning and memory,<sup>134</sup> pain management,<sup>135</sup> cancer,<sup>136</sup> and diabetes.<sup>137</sup> Two endogenous ligands for CB<sub>1</sub> are anandamide [AEA (**94**)] and 2-arachidonoyl glycerol [2-AG (**95**)]; these consist of arachidonic acid with either an ethanolamine (AEA) or glycerol (2-AG) amide linkage, respectively.<sup>130</sup>

Structural resemblance of a number of cyanobacterial metabolites to known endocannabinoids has revealed these prokaryotes to be a rich source of these bioactive lipids, such as serinolamide A (**96**) (CB<sub>1</sub>: K<sub>i</sub> = 16.4 μM; CB<sub>2</sub>: K<sub>i</sub> = 5.2 μM),<sup>138</sup> malyngamide B (**97**) (CB<sub>1</sub>: K<sub>i</sub> = 3.6 μM; CB<sub>2</sub>: K<sub>i</sub> = 2.6 μM),<sup>139</sup> semiplenamamide A (**98**) (CB<sub>1</sub>: K<sub>i</sub> = 19.5 μM),<sup>86</sup> and lyngbyamide A (**99**) (CB<sub>1</sub>: K<sub>i</sub> = 4.7 μM).<sup>140</sup>



**Figure 1.9:** Bioactive cyanobacteria metabolites (part 2)

### 1.3 Dissertation Contents

The primary focus of the research described in following chapters is the isolation and structure elucidation of secondary metabolites from marine cyanobacteria, with a secondary focus on a structure activity relationship study into the alkyl amide lyngbyamide A (**97**). Chapter 2 describes the isolation and characterization of eight new metabolites, veraguamides A-C and H-L, which were isolated from a collection of *Oscillatoria margaritifera* from Coiba National Park (CNP) off of Panama's west coast, as part of the Panama International Cooperative Biodiversity Group program. The planar structure of veraguamides A and L were fully deduced by 2D NMR spectroscopy and mass spectrometry, whereas the structures of veraguamides B, C, and H-K were mainly determined by a combination of <sup>1</sup>H NMR and MS<sup>2</sup>/MS<sup>3</sup> techniques.<sup>60</sup> These new compounds are analogous to the mollusk-derived kulomo'opunalide natural products, with two of the veraguamides (C and H) containing the same terminal alkyne moiety.<sup>141</sup> However, four veraguamides, A (**69**), B, K, and L, also feature an alkynyl bromide, a functionality that has been previously observed in only one other marine natural product, jamaicamide A (**65**).<sup>88</sup> Veraguamide A showed potent cytotoxicity to the H-460 human lung cancer cell line (LD<sub>50</sub> = 141 nM).<sup>67</sup>

Chapter 3 discusses the isolation and structure elucidation of a new lipopeptide, lyngbyabellin N, from an extract of the marine cyanobacterium, *Moorea bouillonii*, collected from Palmyra Atoll in the Central Pacific Ocean. Its planar structure and absolute configuration were elucidated by the combination of spectroscopic and chromatographic analyses as well as chemical synthesis of fragments. In addition to structural features typical of the lyngbyabellins, such as two thiazole rings and a chlorinated 2-methyloctanoate residue, this new compound possesses several interesting aspects, including an unusual *N,N*-dimethylvaline terminus and a leucine statine.<sup>86</sup> Lyngbyabellin N exhibits strong cytotoxic activity against the HCT-116 colon cancer cell line (IC<sub>50</sub> = 40.9 ± 3.3 nM).<sup>142</sup>

Chapter 4 describes the isolation and characterization of three new lipopeptides, tasiamides C-E, from a collection of the tropical marine cyanobacterium *Symploca* sp., collected near Kimbe Bay, Papua New Guinea. This collection has been particularly rich in secondary metabolites, such as kimbeamides A-C (**66**), kimbelactone A, and tasihalide C, which were previously characterized.<sup>89</sup> However, a renewed investigation into a relatively more polar cytotoxic fraction yielded three new lipopeptides.<sup>83</sup> Their planar structures were deduced by traditional 2D NMR spectroscopy and tandem mass spectrometry, and their absolute configurations were determined by a combination of Marfey's and chiral GC-MS analysis. These new metabolites are similar to several previously isolated families of metabolites, including tasiamide, the grassystatins, and symplocin A, all of which were isolated from similar marine filamentous cyanobacteria.<sup>143-145</sup>

Chapter 5 consists of the isolation and structure elucidation of three new marine cyanobacterial natural products, parguerene, precarriebowmide, and mooreamide, from two separate collections of *Moorea* sp., one obtained from Puerto Rico and the other from Papua New Guinea. The planar structures of each were deduced by 2D NMR spectroscopy and mass spectrometry. Parguerene and mooreamide are modified alkyl amides, whereas precarriebowmide is a lipopeptide and represents only a minor modification compared to two other known metabolites, carriebowmide and carriebowmide sulfone. The identification of precarriebowmide led to an investigation into whether carriebowmide and carriebowmide sulfone were true secondary metabolites or isolation artifacts.<sup>146</sup> Both parguerene and mooreamide are structurally reminiscent of the endocannabinoids anadamide (**94**) and 2-arachidonoylglycerol (**95**), and thus it was hypothesized that each would exhibit some cannabinoid receptor binding activity. Unfortunately, parguerene decomposed prior to being evaluated, but mooreamide exhibited moderate selective binding affinity towards CB<sub>1</sub> over CB<sub>2</sub> ( $K_i = 0.47 \mu\text{M}$  and  $K_i > 25 \mu\text{M}$ , respectively).

Chapter 6 examines the structure-activity relationship (SAR) of lyngbyamide A in a broad range of bioassays. The lyngbyamide family of metabolites were isolated from a collection of *M. bouillonii* obtained from Grenada in 1995, and were shown to exhibit both brine shrimp toxicity and cannabinoid receptor binding activity.<sup>140</sup> They are rather small alkyl amides consisting of a twelve carbon fatty acid tail group which is functionalized with a trans-cyclopropyl ring at the C4 position, and a head group portion which typically consists of a biogenic amine deriving from isoleucine, tyrosine, or phenylalanine. In total, 49 analogs were designed and synthesized to probe the importance of key functional groups, such as the cyclopropyl ring, the chain length (both shorter and longer), head group polarity, and number of amine substituents. These analogs were constructed in three different rounds of synthesis and were evaluated for cathepsin L activation/inhibition, brine shrimp toxicity, cannabinoid receptor binding, cancer cell cytotoxicity, nitric oxide production in RAW cells, and ion channel modulation. Interestingly, a subset of these analogs showed strong activity toward stabilization of cathepsin L and brine shrimp toxicity. It is well known that several commercially available surfactants also have the potential to stabilize cathepsin L, and by analysis with a tensiometer, it was determined that this subset of analogs were indeed surfactants.

The dissertation finishes with a conclusion and future work chapter in which I give a brief summary about each of the research chapters herein and elaborate on some interesting potential future directions for each project. Additionally, I discuss the key role natural products should have on drug development in the future, with an emphasis on those deriving from a marine source.

## 1.4 References

- (1) Newman, D. J.; Cragg, G. M.; Snader, K. M. *Nat. Prod. Rep.* **2000**, *17*, 215-234.
- (2) Solecki, R. S. *Science* **1975**, *190*, 880-881.
- (3) Anonymous *HerbalGram* **1998**, *42*, 33-47.
- (4) Zhong, G. S.; Wan, F. *Chin. J. Med. Hist.* **1999**, *29*, 178-182.
- (5) Jiao, Y. M.; Wang, F. *J. TCM* **2005**, *36*, 58-59.
- (6) Gao, X. M. *Advanced Traditional Chinese Medicine Series/Chinese Materia Medica*; People's Medical Publishing House: Beijing, China, 2004; Vol. 1
- (7) Magner, L. N. In *A History of Medicine*; Marcel Dekker: New York, NY, 1992; Vol. 1, pp. 37-62.
- (8) Dev, S. *Environ. Health Perspect.* **1999**, *107*, 783-789.
- (9) Magner, L. N. In *A History of Medicine*; Marcel Dekker: New York, NY, 1992; Vol. 1, pp. 63-98.
- (10) Kong, D. X.; Li, X. J.; Zhang, H. Y. *ChemMedChem* **2008a**, *3*, 1169-1171.
- (11) Neumeyer, J. L. In *Foye's Principles of Medicinal Chemistry*; Lueke, T. L., Williams, D. A.; Lippincott Williams & Wilkins: Baltimore, MD, 2013; Vol. 7, pp. 1-12.
- (12) Grabley, S.; Thiericke, R. *Adv. Biochem. Eng. Biotechnol.* **1999**, *64*, 101-154.
- (13) Kinghorn, D. A. In *Foye's Principles of Medicinal Chemistry*; Lueke, T. L., Williams, D. A. (Eds.); Lippincott Williams & Wilkins: Baltimore, MD, 2013; Vol. 7, pp. 13-28.
- (14) Chin, Y. W.; Baluas, M. J.; Chai, H. B.; Kinghorn, A. D. *The AAPS Journal* **2006**, *28*, E239-E253.
- (15) Dewick, P. M. In *Medicinal Natural Products: A Biosynthetic Approach*; John Wiley & Sons: Chichester, UK, 2002, Vol. 2, pp. 453-504.
- (16) Newman, D. J.; Cragg, G. M.; Snader, K. M. *J. Nat. Prod.* **2003**, *66*, 1022-1037.
- (17) Bulter, M. S. *Nat. Prod. Rep.* **2005**, *22*, 162-195.
- (18) Fabricant, D. S.; Farnsworth, N. R. *Environ. Health Persp.* **2001**, *109*, 69-75.
- (19) Feher, M.; Schmidt, J. M. *J. Chem. Info. Comput. Sci.* **2003**, *43*, 218-227.
- (20) Hoffman, T.; Mueller, S.; Nadmid, S.; Garcia, R.; Mueller, R. *J. Am. Chem. Soc.* **2013**, *135*, 16904-16911.

- (21) Muigg, P.; Josefin, R.; Bohlin, L.; Backlund, A. *Phytochem Rev.* **2013**, *12*, 449-457.
- (22) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311-335.
- (23) Cragg, G. M.; Newman, D. J.; Snader, K. M. *J. Nat. Prod.* **1997**, *60*, 52-60.
- (24) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2007**, *70*, 461-477.
- (25) Preston, J. N.; Trivedi, M. V. *Ann. Pharmacother.* **2012**, *46*, 802-811.
- (26) Carter, G. T. *Nat. Prod. Rep.* **2011**, *28*, 1783-1789.
- (27) Scheuer, P. J. *J. Chem. Ed.* **1999**, *76*, 1075-1079.
- (28) Butler, A.; Walker, J. V. *Chem. Rev.* **1993**, *93*, 1937-1944.
- (29) Wallace, R. W. *Mol. Med. Today* **1997**, *3*, 291-295,
- (30) Bergmann, W.; Feeney, R. J. *J. Org. Chem.* **1951**, *16*, 981-987.
- (31) Bergmann, W.; Burke, D. C. *J. Org. Chem.* **1956**, *22*, 226-228.
- (32) Bergmann, W.; Stempien, M. F. *J. Org. Chem.* **1957**, *22*, 1575-1557.
- (33) <http://www.cancer.gov/cancertopics/druginfo/cytarabine> (accessed January 27th, 2014)
- (34) Boon, R. *Antivir. Chem. Chemoth.* **1997**, *8*, 5-10.
- (35) Faulkner, D. J. *Marine Natural Reports* **1986**, *3*, 1-33.
- (36) Amico, V.; Oriente, G.; Piattelli, M.; Tringali, C.; Fattorusso, E.; Magno, S.; Mayol, L. *Tetrahedron Lett.* **1978**, *38*, 3593-3596.
- (37) Paul, V. J.; Fenical, W. *Science* **1983**, *221*, 747-749.
- (38) Kitagawa, I.; Kobayashi, M.; Yasuzawa, T.; Byeng Wha, S.; Yoshihara, M.; Kyogoku, Y. *Tetrahedron* **1985**, *41*, 995-1005.
- (39) Fattorusso, E.; Minale, L.; Guido, S.; Moody, K.; Thomson, R. H. *J. Chem. Soc.* **1970**, *12*, 752-753.
- (40) Moody, K.; Thomson, R. H.; Fattorusso, E.; Minale, L.; Sodano, G. *J. Chem. Soc.* **1972**, *1*, 18-24.
- (41) Forenza, S.; Minale, L.; Riccio, R.; Fattorusso, E. *J. Chem. Soc.* **1971**, *18*, 1129-1130.
- (42) Walker, R. P.; Faulkner, D. J.; Van Engen, D.; Clardy, J. *J. Am. Chem. Soc.* **1981**, *103*, 6772-6773.

- (43) Crispino, A.; de Giulio, A.; de Rosa, S.; Strazzullo, G. *J. Nat. Prod.* **1989**, *52*, 646-648.
- (44) Muller, W. E.; Maidhof, A.; Zahn, R. K.; Schroder, H. C.; Gasic, M. J.; Heidemann, D.; Bernd, A.; Kurelec, B.; Eich, E.; Seibert, G. *Cancer Res.* **1985**, *45*, 4822-4826.
- (45) Kazlauskas, R.; Lidgard, R. O.; Wells, R. J.; Vetter, W. *Tetrahedron Lett.* **1977**, *36*, 3183-3186.
- (46) Guella, G.; Mancini, I.; Guerriero, A.; Pietra, F. *Helv. Chim. Acta* **1985**, *68*, 1276-1282.
- (47) Copley, J. T. P.; Tyler, P. A.; Van Dover, C. L.; Philip, S. J. *Mar. Ecol. Prog. Ser.* **2003**, *255*, 171-181.
- (48) Kernan, M. R.; Molinski, T. F.; Faulkner, D. J. *J. Org. Chem.* **1988**, *53*, 5014-5020.
- (49) Carte, B.; Faulkner, D. J. *J. Org. Chem.* **1983**, *48*, 2314-1318.
- (50) Orjala, J.; Gerwick, W. H. *J. Nat. Prod.* **1996**, *59*, 427-430.
- (51) Imhoff, J. F.; Stöhr, R. In *Sponges*; Müller, W. E. G. (Ed.); Springer: Heidelberg, 2003; Vol. 1, pp. 35-54.
- (52) Cimino, G.; Ghiselin, M. T. In *Chemical Defense and Evolution of Opisthobranch Gastropods*; California Academy of Sciences: San Francisco, CA, 2009; Vol. 60, pp 179-187.
- (53) Perry, N. B.; Blunt, J. W.; Munro, M. H. G.; Pannell, L. K. *J. Am. Chem. Soc.* **1988**, *110*, 4850-4851.
- (54) Dalisay, D. S.; Molinski, T. F. *Org. Lett.* **2009**, *11*, 1967-1970.
- (55) <http://www.bruker.com/products/mr/nmr/probes/cryoprobes/microcryoprobe/technical-details.html> (accessed January 24th, 2014)
- (56) Molinski, T. F. *Nat Prod. Rep.* **2010**, *27*, 321-329.
- (57) Molinski, T. F. *Curr Opin. Biotech.* **2010**, *21*, 819-826.
- (58) Nyber, N. T.; Duus, J. O.; Sorensen, O. W. *J. Am. Chem. Soc.* **2005**, *127*, 6154-6155.
- (59) Kurz, M.; Schmieder, P.; Kessler, H. *Angew. Chem. Int. Ed. Eng.* **1991**, *30*, 1329-1331.
- (60) Mohimani, H.; Liu, W. T.; Mylne, J. S.; Poth, A. G.; Colgrave, M. L.; Tran, D.; Selsted, M. E.; Dorrestein, P. C.; Pevzner, P. A. *J. Proteome Res.* **2011**, *10*, 4505-4512.
- (61) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Banderia, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1743-1752.

- (62) Udworthy, D. W.; Zeigler, L.; Asolkar, R. N.; Singan, V.; Lapidus, A.; Fenical, W.; Jensen, P. R.; Moore, B. S. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10376-10381.
- (63) Eustáquio, A. S.; McGlinchey, R. P.; Liu, y.; Hazzard, C.; Beer, L. L.; Florova, G.; Alhamadsheh, N. M.; Lechner, A.; Kale, A. J.; Kobayashi, Y.; Reynolds, K. A.; Moore, B. S. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12295-12300.
- (64) McGlinchey, R. P.; Nett, M.; Moore, B. S. *J. Am. Chem. Soc.* **2008**, *130*, 2406-2407.
- (65) Miyanaga, A.; Janso, J. E.; McDonald, L.; He, M.; Liu, H.; Barbieri, L.; Eustáquio, A. S.; Fielding, E. N.; Carter, G. T.; Jensen, P. R.; Feng, X.; Leighton, M.; Koehn, F. E.; Moore, B. S. *J. Am. Chem. Soc.* **2011**, *133*, 13311-13313.
- (66) Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22-31.
- (67) Li, B.; Sher, D.; Kelly, L.; Shi, Y.; Huang, K.; Knerr, P. J.; Joewono, I.; Rusch, D.; Chisholm, S. W.; van der Donk, W. A. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10430-10435.
- (68) Sudek, S.; Haygood, M. G.; Youssef, D. T.; Schimdt, E. W. *Appl. Environ. Microbiol.* **2006**, *72*, 4382-4387.
- (69) Hugenholtz, P.; Goebel, B. M.; Pace, N. R. *J. Bacteriol.* **1998**, *180*, 4765-4774.
- (70) Piel, J.; Hui, D.; Wen, G.; Butzke, D.; Platzer, M.; Fusetani, N.; Matsunaga, S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16222-16227.
- (71) Schmidt, E. W.; Nelson, J. T.; Rasko, D. A.; Sudek, S.; Eisen, J. A.; Haygood, M. G.; Ravel, J. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7315-7320.
- (72) Andersen, R. J.; Williams, D. E.; Strangman, W. K.; Roberge, M. In *Anticancer Agents from Natural Products*; Gordon, C. M., Kingston, D. G. I., Newman, D. J. (Eds.); CRC Press: London, GBR, 2011; Vol. 2, pp. 347-364.
- (73) Wender, P. A. *Tetrahedron* **2013**, *69*, 7529-7550.
- (74) Nicolaou, K. C.; Sorensen, E. J. *Classics in Total Synthesis: Targets, Strategies, Methods*; VCH Publishers, Inc: New York, NY, 1996; Vol. 1.
- (75) Anderson, H. J.; Coleman, J. E.; Andersen, R. J.; Roberge, M. *Cancer Chemother. Pharmacol.* **1997**, *39*, 223-226.
- (76) Nieman, J. A.; Coleman, J. E.; Wallace, D. J.; Piers, E.; Lim, L. Y.; Roberge, M.; Andersen, R. J. *J. Nat. Prod.* **2003**, *66*, 183-199.
- (77) Loganzo, F.; Miscafani, C. M.; Annable, R.; Beyer, C.; Musto, S.; Hari, M.; Tan, X.; Hardy, C.; Hernandez, R.; Baxter, M.; Singanallore, T.; Khafizova, G.; Poruchynsky, M. S.; Fojo, T.; Nieman, J. A.; Ayrál-Kaloustian, S.; Zask, A.; Andersen, R. J.; Greenberger, L. M. *Cancer Res.* **2003**, *63*, 1838-1845.

- (78) Gerwick, W. H.; Moore, B. S. *Chem. Biol.* **2012**, *19*, 85-98.
- (79) [http://www.phrma.org/sites/default/files/159/rd\\_brochure\\_020307](http://www.phrma.org/sites/default/files/159/rd_brochure_020307) (accessed Feb. 3rd, 2007)
- (80) Bothe, H. In *The Biology of Cyanobacteria*; Carr, N. G., Whitton, B. A. (Eds.); University of California Press: Berkeley, CA, 1982; Vol. 19, pp. 87-104.
- (81) Tan, L. T. *Phytochemistry* **2007**, *68*, 954-979.
- (82) Engene, N.; Rottacker, E. C.; Kaštovský, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komárek, J.; Gerwick, W. H. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1171-1178.
- (83) Tan, L. T. *J. Appl. Phycol.* **2010**, *22*, 659-676.
- (84) Gerwick, W. H.; Tan, L. T.; Sitachitta, N. *Alkaloids Chem. Biol.* **2001**, *57*, 75-184.
- (85) Cardellina II, J. H.; Marnier, F. J.; Moores, R. E. *J. Am. Chem. Soc.* **1979**, *101*, 240-242.
- (86) Han, B.; McPhail, K. L.; Ligresti, A.; Di Marzo, V.; Gerwick, W. H. *J. Nat. Prod.* **2003**, *66*, 1364-1368.
- (87) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. *J. Org. Chem.* **1994**, *59*, 1243-1245.
- (88) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (89) Nunnery, J. K.; Engene, N.; Byrum, T.; Cao, Z.; Jabba, S. V.; Pereira, A. R.; Matainaho, T.; Murray, T. F.; Gerwick, W. H. *J. Org. Chem.* **2012**, *77*, 4198-4208.
- (90) Nunnery, J. K.; Suyama, T. L.; Linington, R. G.; Gerwick, W. H. *Tetrahedron Lett.* **2011**, *52*, 2929-2932.
- (91) Boudreau, P. D.; Byrum, T.; Liu, W. T.; Dorrestein, P. C.; Gerwick, W. H. *J. Nat. Prod.* **2012**, *75*, 1560-1570.
- (92) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Mooberry, S. L. *J. Nat. Prod.* **2000**, *63*, 611-615.
- (93) Mevers, E.; Liu, W. T.; Engene, N.; Mohimani, H.; Byrum, T.; Pevzner, P. A.; Dorrestein, P. C.; Spadafora, C.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 928-936.
- (94) Salvador, L. A.; Biggs, J. S.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2011**, *74*, 917-927.
- (95) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. *Curr. Opin. Biotech.* **2010**, *21*, 1-7.
- (96) <http://www.cancer.org> (accessed January 27, 2014)

- (97) Cragg, G. M.; Newman, D. J. *Biochim. Biophys. Acta* **2013**, *1830*, 3670-3695.
- (98) Reusser, F. *J. Bacteriol.* **1971**, *105*, 580-588.
- (99) Wilmanns, W. In *Cancer Chemotherapy*; Elkerbout, F., Ed.; Leiden Unvi. Press: Leiden, Netherlands, 1972; Vol. 1, pp 301-310.
- (100) Thorn, C. F.; Oshiro, C.; Marsh, S.; Hernandez-Boussard, T.; McLeod, H.; Klein, T. E.; Altman, R. B. *Pharmacogenet. Genom.* **2011**, *21*, 440-446.
- (101) Horwitz, S. B. *Ann. Oncol.* **1994**, *5*, S3-6.
- (102) Avendaño, C.; Menéndez, J. C. In *Medicinal Chemistry of Anticancer Drugs*; Elsevier: Amsterdam, 2008; Vol. 1, 9-247.
- (103) <http://www.adcetris.com> (accessed January 27, 2014)
- (104) Luesch, H.; Moore, R. E.; Paul, V. J.; Mooberry, S. L.; Corbett, T. H. *J. Nat. Prod.* **2001**, *64*, 907-910.
- (105) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Mooberry, S. L.; Corbett, T. H. *J. Nat. Prod.* **2002**, *65*, 16-20.
- (106) Bai, R.; Friedman, S. J.; Pettit, G. R.; Hamel, E. *Biochem. Pharmacol.* **1992**, *43*, 2637-2645.
- (107) Hong, J.; Luesch, H. *Nat. Prod. Rep.* **2012**, *29*, 449-456.
- (108) Bai, R.; Verdier-Pinard, P.; Gangwar, S.; Stessman, C. C.; McClure, K. J.; Sausville, E. A.; Pettit, G. R.; Bates, R. B.; Hamel, E. *Mol. Pharmacol.* **2001**, *59*, 462-469.
- (109) Han, B.; McPhail, K. L.; Gross, H.; Goeger, D. E.; Mooberry, S. L.; Gerwick, W. H. *Tetrahedron* **2005**, *61*, 11723-11729.
- (110) Marquez, B. L.; Watts, K. S.; Yokochi, A.; Roberts, M. A.; Verdier-Pinard, P.; Jimenez, J. I.; Hamel, E.; Scheuer, P. J.; Gerwick, W. H. *J. Nat. Prod.* **2002**, *65*, 866-871.
- (111) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418-5423.
- (112) Shen, S.; Zheng, P.; Lovchik, M. A.; Li, Y.; Tang, L.; Chen, Z.; Zeng, R.; Ma, D.; Yuan, J.; Yu, Q. *J. Cell Biol.* **2009**, *185*, 629-639.
- (113) Liu, Y.; Law, B. K.; Luesch, H. *Mol Pharmacol.* **2009**, *76*, 91-104.
- (114) Wrasidlo, W.; Mielgo, A.; Torres, V. A.; Barbero, S.; Stoletov, K.; Suyama, T. L.; Klemke, R. L.; Gerwick, W. H.; Carson, D. A.; Stupack, D. G. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2313-2318.

- (115) 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. *J. Am. Chem. Soc.* **2008**, *130*, 6234-6235.
- (116) Villa, F. A.; Lieske, K.; Gerwick, L. *Eur. J. Pharmacol.* **2010**, *629*, 140-146.
- (117) Choi, H.; Mascuch, S. J.; Villa, F. A.; Byrum, T.; Teasdale, M. E.; Smith, J. E.; Preskitt, L. B.; Rowley, D. C.; Gerwick, L.; Gerwick, W. H. *Chem. Biol.* **2012**, *19*, 589-598.
- (118) Pereira, A.; Cao, Z.; Murray, T. F.; Gerwick, W. H. *Chem. Biol.* **2009**, *16*, 893-906.
- (119) Choi, H.; Pereira, A. R.; Cao, Z.; Shuman, C. F.; Engene, N.; Byrum, T.; Matainaho, T.; Murray, T. F.; Mangoni, A.; Gerwick, W. H. *J. Nat. Prod.* **2010**, *73*, 1411-1421.
- (120) Orjala, J.; Nagle, D. G.; Hsu, V.; Gerwick, W. H. *J. Am. Chem. Soc.* **1995**, *117*, 8281-8282.
- (121) Li, W. I.; Berman, F. W.; Okino, T.; Yokokawa, F.; Shioiri, T.; Gerwick, W. H.; Murray, T. F. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7599-7604.
- (122) LePage, K. T.; Goeger, D.; Yokokawa, F.; Asano, T.; Shioiri, T.; Gerwick, W. H.; Murray, T. F. *Toxicol. Lett.* **2005**, *158*, 133-139.
- (123) Pereira, A. R.; Cao, A.; Engene, N.; Soria-Mercado, I. E.; Murray, T. F.; Gerwick, W. H. *Org. Lett.* **2010**, *12*, 4490-4493.
- (124) <http://www.cdc.gov/parasites/> (accessed January 24<sup>th</sup>, 2014)
- (125) <http://www.icbg.org/groups/panama.php> (accessed January 24<sup>th</sup>, 2014)
- (126) Simmons, T. L.; Engene, N.; Urena, L. D.; Romero, L. I.; Ortega-Barría, E.; Gerwick, L.; Gerwick, W. H. *J. Nat. Prod.* **2008**, *71*, 1544-1550.
- (127) Balunas, M. J.; Linington, R. G.; Tidgewell, K.; Fenner, A. M.; Urena, L. D.; Togna, G. D.; Kyle, D. E.; Gerwick, W. H. *J. Nat. Prod.* **2010**, *73*, 60-66.
- (128) Linington, R. G.; Clark, B. R.; Trimble, E. E.; Almanza, A.; Urena, L. D.; Kyle, D. E.; Gerwick, W. H. *J. Nat. Prod.* **2009**, *72*, 14-17.
- (129) Miller, B.; Friedman, A. J.; Choi, H.; Hogan, J.; McCammon, J. A.; Hook, V.; Gerwick, W. H. *J. Nat. Prod.* **2014**, *77*, 92-99.
- (130) Stolze, S. C.; Deu, E.; Kaschani, F.; Nan, L.; Florea, B. I.; Richau, K. H.; Colby, T.; van der Hooft, R. A. L.; Overkleeft, H. S.; Bogyo, M.; Kaiser, M. *Chem. Biol.* **2012**, *19*, 1546-1555.
- (131) Elphick, M. R.; Egertová, M. *Phil. Trans. R. Soc. Lond. B.* **2001**, *356*, 381-408.
- (132) Williams, C. M.; Kirkham, T. C. *Psychopharmacology* **1999**, *143*, 315-317.

- (133) Ramos, J. A.; Gómez, M.; de Miguel, R. In *Cannabinoids*; Pertwee, R. G., Ed.; Springer: Berlin Heidelberg, 2005; Vol 168, pp 643-656.
- (134) Riedel, G.; Davies, S. N. In *Cannabinoids*; Pertwee, R. G., Ed.; Springer: Berlin Heidelberg, 2005; Vol. 168, pp 445-478.
- (135) Walker, J. M.; Hohmann, A. G. In *Cannabinoids*; Pertwee, R. G., Ed.; Springer: Berlin Heidelberg, 2005; Vol. 168, pp 509-554.
- (136) Guzmán, M. *Nat. Rev. Cancer* **2003**, *3*, 745-755.
- (137) Di Marzo, V.; Piscitelli, F.; Mechoulam, R. In *Diabetes – Perspectives in Drug Therapy*; Schwanstecher, M., Ed.; Springer: Heidelberg, 2011; Vol. 203, pp 75-104.
- (138) Guitiérrez, M.; Pereira, A. R.; Debonisi, H. M.; Ligresti, A.; Di Marzo, V.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 2313-2317.
- (139) Montaser, R.; Paul, V. J.; Luesch, H. *ChemBioChem* **2012**, *13*, 2676-2681.
- (140) Sitachitta, N.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 681-684.
- (141) Nakao, Y.; Yoshida, W. Y.; Szabo, C. M.; Baker, B. J.; Scheuer, P. J. *J. Org. Chem.* **1998**, *63*, 3272-3280.
- (142) Choi, H.; Mevers, E.; Byrum, T.; Valeriote, F. A.; Gerwick, W. H. *Eur. J. Org. Chem.* **2012**, *27*, 5141-5150.
- (143) Kwan, J. C.; Eksioğlu, E. A.; Liu, C.; Paul, V. J.; Luesch, H. *J. Med. Chem.* **2009**, *52*, 5732-5747.
- (144) Molinski, T. F.; Reynolds, K. A.; Morinaka, B. I. *J. Nat. Prod.* **2012**, *75*, 425-431.
- (145) Williams, P. G.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2003**, *66*, 1006-1009.
- (146) Mevers, E.; Byrum, T.; Gerwick, W. H. *J. Nat. Prod.* **2013**, *76*, 1810-1814.

## Chapter 2:

### Cytotoxic Veraguamides, Alkynyl Bromide-Containing Cyclic Depsipeptides from the Marine Cyanobacterium *Oscillatoria margaritifera*

#### 2.0.1 Abstract

A family of cancer cell cytotoxic cyclodepsipeptides, veraguamides A-C (**1-3**) and H-L (**4-8**), were isolated from a collection of cf. *Oscillatoria margaritifera* obtained from the Coiba National Park, Panama as part of the Panama International Cooperation Biodiversity Group (ICBG) program. The planar structure of veraguamide A (**1**) was deduced by 2D NMR spectroscopy and mass spectrometry whereas the structures of **2-8** were mainly determined by a combination of <sup>1</sup>H NMR and MS<sup>2</sup>/MS<sup>3</sup> techniques. These new compounds are analogous to the mollusk-derived kulomo'opunalide natural products, with two of the veraguamides (C and H) containing the same terminal alkyne moiety. However, four veraguamides, A, B, K and L, also feature an alkynyl bromide, a functionality that has only been previously observed in one other marine natural product, jamaicamide A. Veraguamide A showed potent cytotoxicity to the H-460 human lung cancer cell line (LD<sub>50</sub> = 141 nM).

#### 2.1 Introduction

Marine cyanobacteria are exceptionally prolific producers of structurally diverse secondary metabolites, of which many have intriguing biological properties.<sup>1-4</sup> An emerging biosynthetic theme in cyanobacterial natural products is the frequent combination of polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS) derived portions, and this results in a highly diverse suite of nitrogen-rich structural frameworks, most of which are lipid soluble.<sup>5,6</sup> A number of these cyanobacterial metabolites possess terminal alkyne functionalities in the PKS-derived sections, including carmabin A,<sup>7</sup> georgamide,<sup>8</sup> pitipeptolide A,<sup>9</sup> yanucamides,<sup>10</sup>

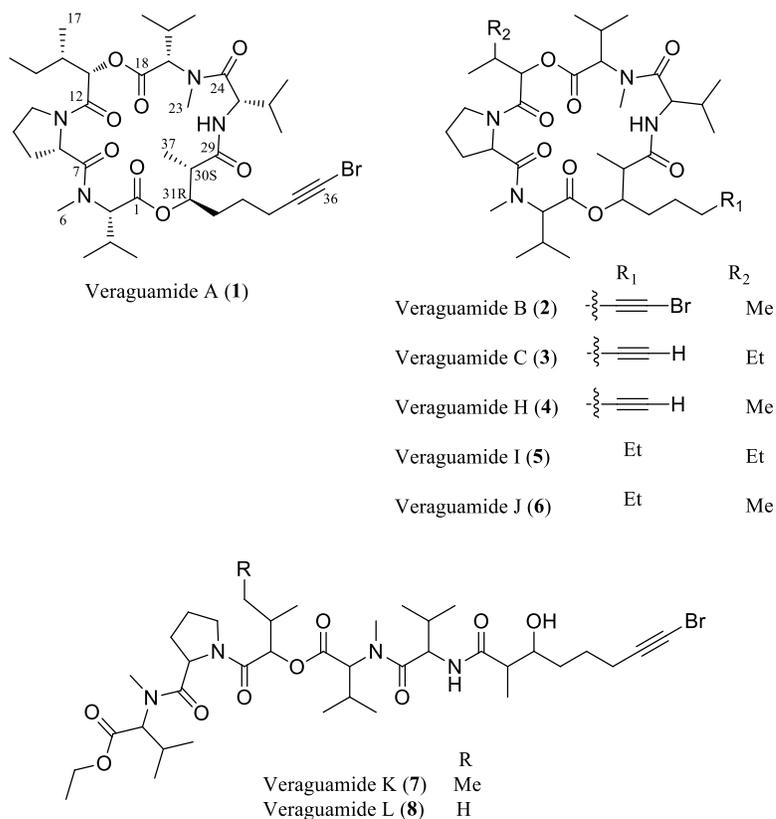
antanapeptin,<sup>11</sup> trungapeptin A,<sup>12</sup> hantupeptin,<sup>13</sup> wewakpeptins,<sup>14</sup> dragonamide,<sup>15</sup> and viridamide A.<sup>16</sup> Similar metabolites have also been obtained from several species of mollusks, namely, *Onchidium* sp. and *Dolabella auricularia*, yielding onchidins A<sup>17</sup> and B<sup>18</sup> and a family of kulolides,<sup>19</sup> respectively. Due to the strong and distinctive similarity between these secondary metabolites isolated from mollusks and those of cyanobacterial origin, it is highly likely that the mollusks obtain these compounds from their diet of cyanobacteria.

Since 1998, the International Cooperative Biodiversity Group (ICBG) in Panama, a program of the Fogarty International Center of the National Institutes of Health, has enabled unique opportunities to conduct integrated natural products investigations, biodiversity inventories and conservation, infrastructure development, and educational training.<sup>20</sup> Moreover, the country of Panama permits the study of marine cyanobacteria from two very different tropical environments, the Caribbean Sea in the Western Atlantic, and the Eastern Pacific. Some of these sites are quite pristine and of exceptional biodiversity, such as the Coiba National Park (CNP), some 15 kilometers off the Pacific coast of Panama. The CNP was formed in 2003 as a result of a developing recognition of its high number of indigenous and endemic plant, animal and microbial species, and in 2005 it was named a World Heritage Site by UNESCO.<sup>21</sup>

Several filamentous tuft-forming species of marine cyanobacteria were collected from the CNP in 2010, and their extracts evaluated in a number of biological assays. Two reduced complexity fractions from one extract, subsequently tentatively identified as *Oscillatoria margaritifera*, were found to be highly cytotoxic to H-460 human lung cancer cells in vitro (2% survival at 3 µg/mL), and these were chosen for further investigation. As a result of a bioassay-guided fractionation process, one major and several minor new cytotoxic lipopeptides were isolated and structurally defined. The major compound, named veraguamide A (**1**) (the CNP lies within the Panamanian state of Veraguas),<sup>22</sup> was highly cytotoxic to H-460 cells (LD<sub>50</sub> = 141 nM); the minor compounds were all of lesser potency to this cancer cell line. As described

below, the structure of **1** was fully characterized, including the absolute configuration at all chiral centers, whereas the planar structures of the minor compounds were largely determined by integrated  $^1\text{H}$  NMR and  $\text{MS}^2/\text{MS}^3$  analysis. Additionally, a new iteration of a recently developed computer algorithm was applied to the  $\text{MS}^2/\text{MS}^3$  data and allowed deduction of the structures of the minor metabolites.<sup>23-25</sup>

During the final stages of this project, a parallel effort in the Luesch and Paul laboratories in Florida found several of the same compounds (veraguamides A, B, and C) (**1-3**) as well as several new derivatives from an Atlantic collection, and these form the substance of a parallel report.<sup>26</sup> It is interesting and potentially insightful to the origin and evolution of the genetic pathways responsible for veraguamide biosynthesis that these same distinctive metabolites have been isolated from cyanobacteria collected from these two well-separated oceans.



**Figure 2.1:** Veraguamides A-C, and K-L

## 2.2 Results and Discussion

### 2.2.1 Veraguamide A

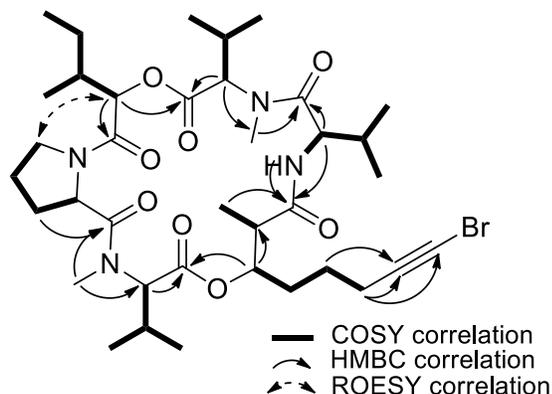
The tentatively identified cyanobacterium *O. margaritifera* was collected by hand from shallow waters (1-5 m deep) in the CNP, Panama, in February 2010. The ethanol-preserved collection was repetitively extracted (CH<sub>2</sub>Cl<sub>2</sub>-MeOH, 2:1) and fractionated using normal-phase vacuum liquid chromatography (VLC). Two fractions that eluted with 100% EtOAc and 75% EtOAc/MeOH were cytotoxic to H-460 human lung cancer cells (both exhibiting 2% survival at 3 µg/mL). Further fractionation with reversed-phase solid-phase extraction (SPE) yielded 2.3 mg of veraguamide A (**1**), a pure amorphous solid, and between 0.1 to 0.5 mg of several analogues, veraguamides B, C, and H-L (**2-8**).

HRESIMS of **1** gave a [M+H]<sup>+</sup> at *m/z* 767.3594 as well as peaks for the [M+Na]<sup>+</sup> and [M+K]<sup>+</sup> adducts at *m/z* 789.3405 and 805.3148, respectively, indicating a molecular formula of C<sub>37</sub>H<sub>59</sub>N<sub>4</sub>O<sub>8</sub>Br and requiring 10 degrees of unsaturation. IR spectroscopy suggested a peptide with a strong absorption band at 1763 cm<sup>-1</sup>, and this was supported by observation of six ester or amide type carbonyls by <sup>13</sup>C NMR analysis (δ<sub>C</sub> 173.5, 172.2, 170.9, 170.7, 169.7, and 166.0). The <sup>1</sup>H NMR spectrum also suggested a peptide with one amide (NH) proton resonating at δ<sub>H</sub> 6.28 and two *N*-methyl groups at δ<sub>H</sub> 3.01 and 2.95. The <sup>13</sup>C NMR spectrum also revealed the presence of an unusually polarized alkyne functionality (δ<sub>C</sub> 79.4 and 38.4), accounting for a further 2 degrees of unsaturation. Thus, 8 of the 10 degrees of unsaturation were explained, and indicated that veraguamide A must possess two rings.

Analysis of 1D and 2D NMR spectra (COSY, TOCSY, ROESY, HSQC and HMBC) led to the identifications of four amino acids [one valine (Val), two *N*-methyl-valines (*N*-MeVal) and one proline (Pro)], one hydroxy acid [2-hydroxy-3-methylpentanoic acid (H3mpa)] and one extended chain polyketide. The proton chemical shifts of the H3mpa residue were very similar to those reported for isoleucine, however, the carbon chemical shift for the α-carbon was

significantly downfield ( $\delta_c$  76.1), consistent with a hydroxy acid. The identity of the extended polyketide residue was deduced from a combination of COSY and HMBC correlations. A CH-CH<sub>3</sub> constellation formed one spin system, and a deshielded methine adjacent to three sequential methylene residues formed a second spin system. By HMBC, the two methine centers were found to be adjacent, and thus a nearly 90° angle must exist between their proton substituents. HMBC between the H-30 methine, as well as its attached secondary methyl group (H<sub>3</sub>-37), and an amide-type carbonyl at  $\delta$  170.9 completed one terminus of this residue. At the other end, HMBC cross peaks were observed between the methylene protons H-34a/H-34b and both carbons C-35 and C-36, whereas methylene protons H-33a and H-33b showed only correlations with C-36. The chemical shift of the distal carbon of the alkyne was quite unusual ( $\delta_c$  38.4), but matched quite well with that reported for the alkynyl bromide present in jamaicamide A, the only other marine natural product reported with this functionality.<sup>27</sup> Thus, this last residue in veraguamide A (**1**) was identified as a derivative of 8-bromo-3-hydroxy-2-methyloct-7-ynoic acid (Br-Hmoya).

As the proline residue accounted for one additional degree of unsaturation, the tenth and final degree of unsaturation must arise from veraguamide A (**1**) having an overall cyclic constitution; this was apparent from the residue connectivities observed by HMBC and ROESY (Table 2.1). HMBC correlations from the two *N*-Me groups and the NH to their respective adjacent carbonyls and  $\alpha$ -carbons were used to connect three of the residues in veraguamide A. A correlation from the  $\alpha$ -hydroxy proton of the Br-Hmoya residue (H-31) to the carbonyl of *N*-MeVal-1 (C-1) served to connect these two residues. Similarly, the H3mpa and *N*-MeVal-2 residues were connected by a HMBC cross peak from the  $\alpha$ -hydroxy proton of the H3mpa residue (H-13) to the C-18 carbonyl of the *N*-MeVal-2 residue. Finally, a ROESY correlation was used to make the concluding connection between the Pro and H3mpa residues. Thus, veraguamide A was deduced to have a cyclo-[*N*-MeVal – Pro – H3mpa - *N*-MeVal – Val – Br-Hmoya] structure.



**Figure 2.2:** Select 2D NMR data for veraguamide A

The absolute configuration of the four  $\alpha$ -amino acids in veraguamide A (**1**) were determined by LC-MS analysis of the acid hydrolysate appropriately derivatized with Marfey's reagent (D-FDAA). The six standards, L-Pro, D-Pro, L-Val, D-Val, L-*N*-MeVal, and D,L-*N*-MeVal were also reacted with D-FDAA and compared to the derivatized hydrolysate by LC-MS. From the retention times and co-injections it was clear that all four of the amino acids, Pro, Val and two *N*-MeVal residues, were of the L configuration.

The absolute configuration of the H3mpa residue was determined by comparing the GC-MS retention time of the methylated residue liberated by acid hydrolysis with authentic standards. The four standards, L-*allo*-H3mpa, L-H3mpa, D-*allo*-H3mpa, and D-H3mpa, were synthesized from L-*allo*-Ile, L-Ile, D-*allo*-Ile, and D-Ile, respectively, following literature procedures.<sup>28</sup> The four standards each possessed distinctly different retention times by GC-MS [44.86 (L-*allo*-H3mpa), 45.06 (D-*allo*-H3mpa), 45.26 (D-H3mpa), and 45.63 min (L-H3mpa)]. The methylated residue from the acid hydrolysate gave a single peak at 45.63 min, thus indicating its configuration as L-H3mpa.

To determine the absolute configuration of the Br-Hmoya residue, compound **1** was hydrogenated with 10% Pd/C to remove simultaneously the bromine atom and fully reduce the terminal alkyne functionality. This hydrogenation product was then hydrolyzed with 6 N HCl in

a microwave reactor to yield the free residues. An aliquot of the methylated hydrolysate was treated with the *S*-Mosher's acid chloride [*S*-(+)- $\alpha$ -methoxy- $\alpha$ -(trifluoromethyl)phenylacetyl chloride, *S*-(+)-MTPA-Cl) and compared to four synthetic standards, as described below. Two core standards, *2S,3S*-Hmoaa and *2S,3R*-Hmoaa, were synthesized using a published procedure.<sup>13b</sup> To create the four chromatographic standards, *2S,3S*-Hmoaa and *2S,3R*-Hmoaa were each separately treated with *S*-MTPA-Cl and *R*-MTPA-Cl, yielding four diastereomeric compounds. These four standards were then compared to the *S*-MTPA-Cl derivatized hydrolysate of veraguamide A (**1**). Two of the standards (*2S,3S*-Hmoaa and *2S,3R*-Hmoaa reacted with *S*-MTPA-Cl) are each identical to a possible configuration of the natural residue, whereas the other two standards (*2S,3S*-Hmoaa and *2S,3R*-Hmoaa reacted with *R*-MTPA-Cl) are enantiomeric to the other two possible configurations of the natural residue (*2R,3R*-Hmoaa and *2R,3S*-Hmoaa, respectively). A GC-MS instrument equipped with a DB5-MS column was then used to compare the retention times of the four diastereomeric standards with the derivatized hydrolysate. The retention time of the hydrolysate product (47.13 min) matched *2S,3R*-Hmoaa that was reacted with *S*-MTPA-Cl, identifying that the absolute configuration of the Hmoya residue in **1** is *30S,31R*. In summary, the above experiments established that veraguamide A (**1**) has *2S, 8S, 13S, 14S, 19S, 25S, 30S* and *31R* absolute configuration.

**Table 2.1:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for veraguamide A (**1**) in  $\text{CDCl}_3$ 

residue	position	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}} (J \text{ in Hz})^{\text{a}}$	HMBC <sup>a</sup>	ROESY <sup>a</sup>
NMeVal-1	1	170.7			
	2	65.0	3.94, d (10.7)	1, 3, 4, 7	8, 9b, 4, 5
	3	28.3	2.28, m	4	4, 6
	4	19.6	0.98, d (6.8)	2, 3, 5	3
	5	19.3	0.92, d (6.6)	2, 4	2, 3, 6
	6	29.5	3.01, s	2, 7	2, 3, 5
Pro	7	172.2			
	8	57.3	4.95, dd (8.5, 6.3)	9, 10, 11	2, 9a, 9b, 10a, 10b
	9a	28.7	2.28, m	8, 10, 11	6, 8, 9b
	9b		1.79, m	7, 8, 10, 11	8, 9a
	10a	25.0	2.03, m	8, 9, 11	8, 11b, 31
	10b		1.99, m	8, 9, 11	11b
	11a	47.3	3.85, dt (9.3, 7.1)	8, 9, 10	10a, 11b, 13
	11b		3.61, dt (9.3, 7.1)	9, 10	10b, 11a, 13
H3mpa	12	166.0			
	13	76.1	4.90, d (9.3)	12, 14, 15, 18	11a, 11b, 14, 16, 31
	14	35.7	1.98, m	17	13
	15a	24.9	1.54, m	14	15b, 16
	15b		1.13, m	14	15a
	16	20.3	1.00, d (6.8)	14, 15	14
	17	10.6	0.87, t (7.3)	13, 14, 15	15a, 16
NMeVal-2	18	169.7			
	19	66.1	4.15, d (9.8)	18, 20, 22, 23, 24	20, 21, 22, 25
	20	28.5	2.25, m	19	19, 23
	21	20.4	1.11, d (6.3)	19, 20, 22	19
	22	20.2	0.99, d (6.8)	19, 20, 21	19, 23

**Table 2.1: continued**

residue	position	$\delta_C^b$	$\delta_H$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	ROESY <sup>a</sup>
	23	30.1	2.95, s	19, 24	20, 22, 28
Val	24	173.5			
	25	52.8	4.71, dt (6.3, 8.5)	24, 26, 27, 28, 29	19, 26, 27, NH-1
	26	32.1	1.90, m	25, 27	25
	27	20.2	0.94, d (6.8)	25, 26, 28	25
	28	17.5	0.88, d (6.8)	25, 26, 27	25
	NH-1		6.28, d (8.5)	29	25, 32, 37
Br-HMOYA	29	170.9			
	30	42.4	3.12, m	29, 31, 37	31, 32
	31	76.4	4.85, d (10.5)	1	10a, 30, 32, 33a, 33b, 34a, 34b, 37
	32	29.7	1.26, m	30, 31	NH-1
	33a	24.8	1.59, m	32, 34, 35	31
	33b		1.42, m	34, 35	31
	34a	19.5	2.20, m	33, 35, 36	31
	34b		1.97, m	33, 35, 36	31
	35	79.4			
	36	38.4			
	37	13.9	1.25, m	29, 30, 31	31, NH-1

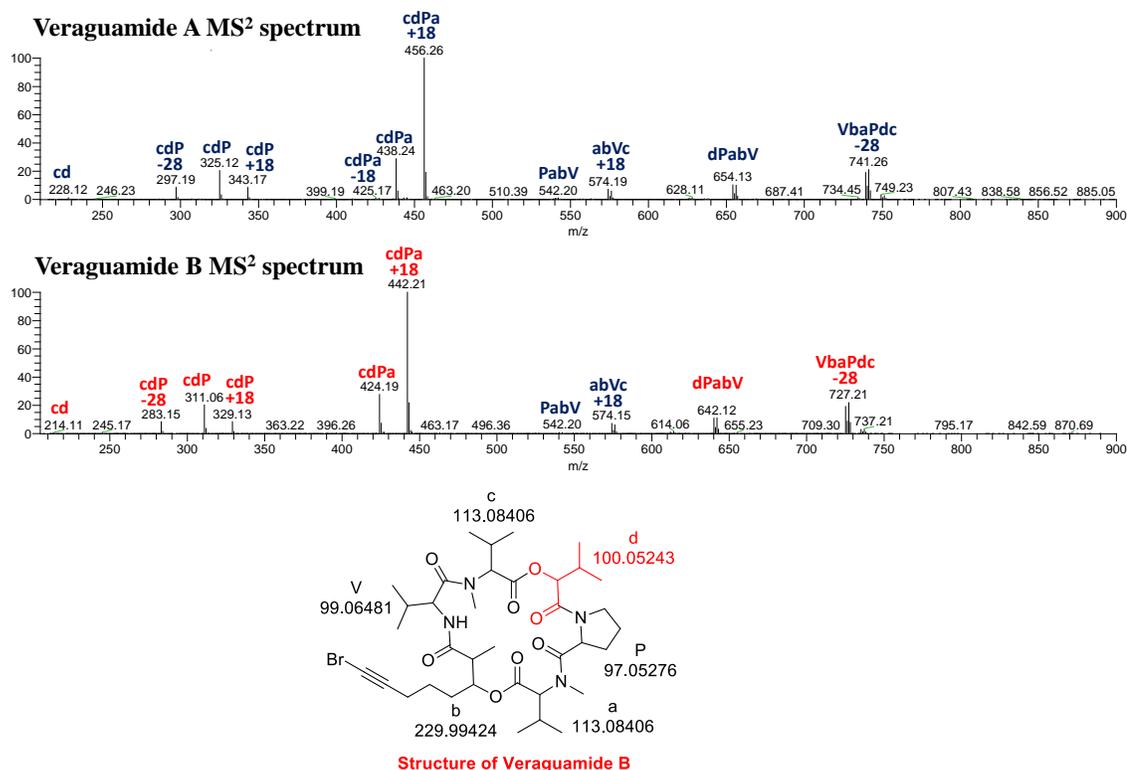
<sup>a</sup>500 MHz for <sup>1</sup>H NMR, HMBC, and ROESY; <sup>b</sup>125 MHz for <sup>13</sup>C NMR

### 2.2.2 Veraguamide B-C, and K-L

Several analogues of compound **1** were isolated from the more polar chromatographic fraction (eluted with 75% EtOAc/MeOH) of the crude extract. Because these analogues were obtained in quite small yield (0.1-0.5 mg), we were motivated to examine their structures using a newly reported computer analysis of MS<sup>2</sup>/MS<sup>3</sup> data obtained for cyclic peptides.<sup>23-25</sup> Additionally, because <sup>1</sup>H NMR analysis of several of these analogues showed them to be similar in overall

structure to veraguamide A (**1**), the position of structural modifications could be determined based on mass shifts in characteristic fragments. With the structure of **1** rigorously determined by a full spectrum of spectroscopic and chemical techniques, it was possible to use this parent structure to determine the characteristic fragmentation pattern for this family of metabolites. Thus, by both a manual comparison of MS<sup>2</sup> fragmentation pattern for each of the analogues to that of **1**, and by application of this newly developed computer algorithm for cyclic peptides, the location and nature of the structural modifications to the veraguamide A (**1**) parent structure were determined readily. In most cases, confirmatory <sup>1</sup>H NMR data were also obtained.

Compound **2** was isolated as a slightly more polar secondary metabolite in approximately 0.3 mg yield, and by HRESIMS indicated a molecular formula of C<sub>36</sub>H<sub>57</sub>N<sub>4</sub>O<sub>8</sub>Br. This mass is 14 Da less than that of veraguamide A (**1**), and thus, veraguamide B (**2**) possesses one fewer fully saturated carbon atom. Consistent with this observation, <sup>1</sup>H NMR analysis showed a nearly identical spectrum as obtained for veraguamide A with only small differences observed in the high field methyl and methylene regions. To localize this mass offset, the MS ion dataset tree for **2**, containing both MS<sup>2</sup> and a series of MS<sup>3</sup> spectra, were subjected to the comparative dereplication algorithm.<sup>23-25</sup> This algorithm compares the MS dataset to the Norine database plus any user inputted sequences (such as veraguamide A); as expected, **1** was the top hit with the 14 Da difference located to the H3mpa residue.<sup>29</sup> To verify this assignment, the MS<sup>2</sup> spectra for compounds **1** and **2** were compared manually, and this also indicated that the 14 Da structural difference was present in the H3mpa residue (Figure 2.3). Thus, the H3mpa residue in veraguamide A (**1**) was replaced by a 2-hydroxy-3-methyl-butanoic acid (Hmba) residue in veraguamide B (**2**). Due to the small amount of compound obtained, and the desire to explore the biological properties of these veraguamide A analogues (discussed below), the absolute configuration was not established for compound **2**, but we speculate that it is likely identical to that of veraguamide A (**1**).



**Figure 2.3:** Sequencing by ESIMS/MS fragmentations

In a similar fashion, the structures for compounds **3**, **4**, **5**, and **6** were also determined, with each possessing only a single modified residue in comparison with either veraguamide A (**1**) or veraguamide B (**2**). Veraguamides C (**3**) and H (**4**) were found to be analogues of compounds **1** and **2**, respectively; however, they lacked the alkynyl bromine atom but retained the alkyne functionality. Veraguamides I (**5**) and J (**6**) also proved to be analogues of compounds **1** and **2**, respectively; in this case they lack both the bromine atom as well as the alkynyl functionality in the polyketide section of the molecule. Again, due to the low yields of compounds **3-6**, their absolute configurations were not determined experimentally; it may be that they are the same as veraguamide A (**1**).

Two additional veraguamides, K (**7**) and L (**8**), were isolated from the more polar and biologically active VLC fraction; however, their structures could not be determined by the

MS<sup>2</sup>/MS<sup>3</sup> method because the algorithm is currently designed specifically for the analysis of cyclic peptides. Additional development of the algorithm is underway to expand its ability to distinguish between linear and cyclic peptides using mass spectrometry data, as this is a long-standing problem in the proteomics and peptidomics fields. Nevertheless, using 600 MHz cryoprobe NMR it was possible to obtain a nearly complete 2D NMR data set for **8** (HSQC, HMBC, and TOCSY). Additionally, HRESIMS of **8** gave a [M+Na]<sup>+</sup> peak at  $m/z$  821.3673, indicating a molecular formula of C<sub>38</sub>H<sub>63</sub>N<sub>4</sub>O<sub>9</sub>Br (9 degrees of unsaturation), differing from veraguamide B (**2**) by C<sub>2</sub>H<sub>6</sub>O and one less degree of unsaturation. <sup>13</sup>C NMR shifts were deduced by a combination of HMBC and HSQC data, and revealed the presence of six ester- or amide-type carbonyls ( $\delta_c$  176.0, 172.5, 172.5, 171.0, 170.0, 166.7) and an alkynyl bromide ( $\delta_c$  79.7 and 38.0), accounting for 8 degrees of unsaturation. As detailed below, a proline in **8** accounted for the ninth and final degree of unsaturation in veraguamide L, signifying that **8** is a linear depsipeptide.

The NMR spectra of veraguamide L (**8**) possessed similar <sup>1</sup>H and <sup>13</sup>C NMR shifts to most of the resonances present for veraguamide A (**1**). Analysis of the 1D and 2D NMR spectra led to the assignments of four amino acids [valine (Val), two *N*-methyl-valines (*N*-MeVal) and proline (Pro)] as well as one hydroxy acid [2-hydroxy-3-methylbutyric acid (Hmba)] and 8-bromo-3-hydroxy-2-methyloct-7-ynoic acid (Br-Hmoya). In addition, HMBC correlations were observed from a deshielded methylene ( $\delta_H$  4.15) to both a methyl carbon ( $\delta_c$  14.3) and a carbonyl ( $\delta_c$  171.0), features not observed for compound **1**. By TOCSY, this same deshielded methylene was directly adjacent to the new methyl group, thus defining an ethyl ester at the carboxylic acid terminus of veraguamide L (**8**). Subsequently, comparison of the MS<sup>2</sup> data for compounds **7** and **8** revealed that the only difference between these two compounds is in the hydroxy acid residue. In **7**, this residue is H3mpa (comparable to **1**) while in **8** it is Hmba (comparable to **2**). At this

point, we are uncertain if veraguamide K (**7**) and L (**8**) are artifacts of the preservation of the original sample in ethanol, or if they represent true natural products of the cyanobacterium.

### 2.2.3 Bioassay Results

Only compounds **1**, **2**, **3**, **7**, and **8** were available in sufficient quantity for evaluation in the H-460 cytotoxicity assay. Compound **1** showed potent activity ( $LD_{50} = 141$  nM), while compounds **2**, **3**, **7**, and **8** all exhibited activity in the low micromolar range, but due to insufficient quantities, no further evaluation of these analogues was possible. However, two structural analogues of veraguamide A, kulomo'opunalide-1 and -2, have similar or identical NRPS portions of the molecule but lack the alkynyl bromide in the PKS portion. These two compounds were previously tested against P388 cells, but were reported to exhibit only moderate cytotoxicity,<sup>19</sup> suggesting that the alkynyl bromide may be an essential structural feature for the potent cytotoxic activity observed for veraguamide A (**1**).

### 2.2.4 Taxonomy of Producing Organism

A taxonomic investigation of the veraguamide-producing cyanobacterium (PAC-17-FEB-10-2) showed that the morphology agreed relatively well with the current definition of *Oscillatoria margaritifera*.<sup>30</sup> *O. margaritifera* was described initially from brackish and marine environments of northern Europe,<sup>30</sup> which makes it geographically and environmentally unlikely that tropical marine PAC-17-FEB-10-2 would belong to the same taxon.<sup>31</sup> Moreover, specimens of *Oscillatoria* have overlapping morphological characters with the genus *Lyngbya*,<sup>32</sup> and phylogenetic analysis is therefore essential to delineate these morphologically similar but evolutionarily unrelated genera.<sup>33</sup> Phylogenetic inferences of the SSU (16S) rRNA gene of PAC-17-FEB-10-2 revealed that this strain nested within the *Oscillatoria* lineage with *O. sancta* PCC 7515 as the closest related reference strain.<sup>31</sup> However, the *Oscillatoria* lineage forms two distinct

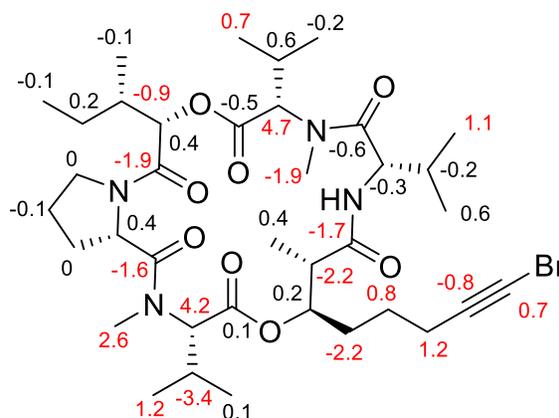
sister clades, one temperate *sensu stricto* (including PCC 7515) and one tropical marine (including PAC-17-FEB-10-2). The DNA bar-coding gap between the two clades was 4.2 (mean p-distance: inter-clade = 2.3 %; intra-clade = 0.6 %), which may support the separation of temperate and tropical marine *Oscillatoria* into two distinct genera. However, because such a revision in the taxonomy of tropical marine *Oscillatoria* has not yet occurred, at the present time the best taxonomic definition of the veraguamide-producing strain PAC-17-FEB-10-2 is cf. *Oscillatoria margaritifera*.

### 2.3 Conclusion

A chemical investigation into a potentially cytotoxic extract of *Oscillatoria margaritifera* which was collected in Coiba National Park, Panama, led to the isolation of veraguamides A-C and H-L (**1**). The planar structure of veraguamide A was determined by the combination of NMR spectroscopy and mass spectrometry techniques. However the structures of veraguamides B, C and H-L were all determined using a MS comparative dereplication algorithm, which needs only nanogram quantities of an impure natural product. Due to limited quantities of each of the analogs, the absolute configuration was only determined for veraguamide A, by utilizing both Marfey's analysis and synthetic chemistry. These new compounds are analogous to the mollusk-derived kulomo'opunalide natural products, with two of the veraguamides (C and H) containing the same terminal alkyne moiety. However, four veraguamides, A, B, K and L, also feature an alkynyl bromide, a functionality that has only been previously observed in one other marine natural product, jamaicamide A. Furthermore, veraguamide A exhibited potent cytotoxicity against H-460, human lung cancer cells, with a LD<sub>50</sub> of 141 nM.

Since publishing this research in back-to-back publications with Hendrik Luesch at the University of Florida, no additional veraguamides have been isolated; however, there is one report of a total synthesis of veraguamide A. The total synthesis was completed in 2012 by Zhang

et al. and they suggest that there is an error in the original stereochemical assignment of veraguamide A as they report significant differences in the  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra between the two natural products and the synthetic material (Figure 2.4).<sup>34</sup> Almost half (18 of 37) of the total carbons are shifted by at least 0.7 ppm from the original reported data, with several shifted as far off as 4.7 ppm, involving carbons from every residue. The carbons that are furthest shifted are the two alpha carbons on the *N*-MeVal's, and thus possibly suggesting that the synthetic version may have an issue with configuration of the tertiary amide bonds. Furthermore, the isolation and absolute configurational analysis on veraguamide A was conducted independently by both myself and Lilibeth Salvador of Hendrik Luesch's laboratory and our findings were consistent with one another. However, further investigations are needed to definitively deduce the difference between the natural products and synthetic versions.



**Figure 2.4:** Differences in  $^{13}\text{C}$  NMR chemical shifts between the natural product and synthetic version

## 2.4 Experimental Methods

### 2.4.1 General Experimental Procedures

Optical rotations were measured on a JASCO P-2000 polarimeter, UV spectra on a Beckman Coulter DU-800 spectrophotometer, and IR spectra were obtained using a Nicolet IR-100 FT-IR spectrophotometer using KBr plates. NMR spectra were recorded with chloroform as internal standard ( $\delta_C$  77.0,  $\delta_H$  7.26) on a Varian Unity 500 MHz spectrometer (500 and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively), on a Varian VNMRs (Varian NMR System) 500 MHz spectrometer equipped with a Cold Probe (500 and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively). Also used were a Bruker 600 MHz spectrometer equipped with a 1.7 mm MicroCryoProbe (600 and 150 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively) and a JOEL 500 MHz spectrometer (500 and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively). LR- and HRESIMS were obtained on a ThermoFinnigan LCQ Advantage Max mass detector and Thermo Scientific LTQ Orbitrap-XL mass spectrometer, respectively.  $\text{MS}^2/\text{MS}^3$  spectra were obtained on Biversa Nanomate with nanoelectrospray ionization on a ThermoFinnigan LTQ-MS which utilized Tune Plus software version 1.0. HPLC was carried out using a Waters 515 pump system with a Waters 996 PDA detector. All solvents were either distilled or of HPLC quality. Acid hydrolysis was performed using a Biotage (Initiator) microwave reactor equipped with high pressure vessels.

### 2.4.2 Cyanobacterial Collections and Morphological Identification

The veraguamide-producing cyanobacterium PAC-17-FEB-10-2 was collected by hand using snorkel gear in shallow water off Isla Canales de Afuera on the Pacific coast of Panama ( $7^\circ 41.617'N$ ,  $81^\circ 38.379'E$ ). Morphological characterization was performed using an Olympus IX51 epifluorescent microscope (1000X) equipped with an Olympus U-CMAD3 camera. Morphological comparison and putative taxonomic identification of the cyanobacterial specimen was performed in accordance with modern classification systems.<sup>32,35</sup>

### 2.4.3 Extraction and Isolation

The cyanobacterial biomass (9.75 g, dry wt) was extracted with 2:1 CH<sub>2</sub>Cl<sub>2</sub>-MeOH to afford 1.8 g of dried extract. A portion of the extract was fractionated by silica gel VLC using a stepwise gradient solvent system of increasing polarity starting from 100% hexanes to 100% MeOH (nine fractions, A-I). The two fractions eluting with 100% EtOAc (fraction G) and 75% EtOAc in MeOH (fraction H) were separated further using RP SPE [500 mg SPE, stepwise gradient solvent system of decreasing polarity starting with 20% CH<sub>3</sub>CN in H<sub>2</sub>O to 100% CH<sub>2</sub>Cl<sub>2</sub>, to produce four fractions (1-4) each] to yield pure veraguamide A (**1**). Further fractionation by RP HPLC using a Phenomenex 4 μm Synergi Fusion analytical column, with a gradient from 50% CH<sub>3</sub>CN/H<sub>2</sub>O to 100% CH<sub>3</sub>CN over 30 min, yielded pure veraguamides B, C and K-L (**2-8**).

**Veraguamide A (1):** amorphous solid;  $[\alpha]_D^{22}$  -14.7 (*c* 0.33, CH<sub>2</sub>Cl<sub>2</sub>); UV (MeCN)  $\lambda_{\max}$  (log  $\epsilon$ ) 204 (4.00), 266 (2.83) nm; IR (neat)  $\nu_{\max}$  3327, 2964, 2930, 1734, 1700, 1456, 1272, 1194, 1128 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (500 MHz, CDCl<sub>3</sub>), see Table 1; ESIMS/MS *m/z* 741.26 (C<sub>36</sub>H<sub>61</sub>N<sub>4</sub>O<sub>7</sub><sup>80</sup>Br), 654.13 (C<sub>31</sub>H<sub>50</sub>N<sub>3</sub>O<sub>7</sub><sup>78</sup>Br), 574.13 (C<sub>26</sub>H<sub>44</sub>N<sub>3</sub>O<sub>6</sub><sup>80</sup>Br), 542.20 (C<sub>25</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub><sup>80</sup>Br), 463.20 (C<sub>20</sub>H<sub>35</sub>N<sub>2</sub>O<sub>5</sub><sup>80</sup>Br), 456.26 (C<sub>23</sub>H<sub>42</sub>N<sub>3</sub>O<sub>6</sub>), 438.24 (C<sub>23</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 343.17 (C<sub>17</sub>H<sub>31</sub>N<sub>2</sub>O<sub>5</sub>), 325.12 (C<sub>17</sub>H<sub>29</sub>N<sub>2</sub>O<sub>4</sub>), 297.19 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>3</sub>), 228.12 (C<sub>12</sub>H<sub>22</sub>NO<sub>3</sub>); HRESIMS [M+H]<sup>+</sup> *m/z* 767.3594 (calcd for C<sub>37</sub>H<sub>61</sub>N<sub>4</sub>O<sub>8</sub><sup>78</sup>Br 767.3594).

**Veraguamide B (2):** amorphous solid;  $[\alpha]_D^{23}$  -13.1 (*c* 0.25, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  0.90 (d, *J* = 6.7, 3 H), 0.94 (d, *J* = 7.3, 3 H), 0.95 (d, *J* = 6.7, 3 H), 0.96 (d, *J* = 5.8, 3 H), 1.00 (d, *J* = 7.6, 3 H), 1.01 (d, *J* = 7.3, 3 H), 1.04 (d, *J* = 6.7, 3 H), 1.12 (d, *J* = 6.7, 3 H), 1.27 (d, *J* = 4.7, 3 H), 1.27 (m, 1 H), 1.45 (m, 1 H), 1.81 (m, 2 H), 1.97-2.12 (m, 4 H), 2.14-2.39 (m, 5 H), 2.96 (s, 3 H), 3.02 (s, 3 H), 3.14 (m, 1 H), 3.62 (q, *J* = 7.4, 1 H), 3.81 (q, *J* = 7.4, 1 H), 3.95 (d, *J* = 10.3, 1 H), 4.16 (d, *J* = 9.3, 1 H), 4.73 (t, *J* = 6.2, 1 H), 4.86 (d, *J* = 8.1, 2 H), 4.96 (t, *J* = 6.1, 1 H), 6.27 (d, *J* = 8.2, 1 H); ESIMS/MS *m/z* 727.21 (C<sub>35</sub>H<sub>59</sub>N<sub>4</sub>O<sub>7</sub><sup>80</sup>Br), 642.12 (C<sub>30</sub>H<sub>48</sub>N<sub>3</sub>O<sub>7</sub><sup>80</sup>Br), 574.13 (C<sub>26</sub>H<sub>44</sub>N<sub>3</sub>O<sub>6</sub><sup>80</sup>Br), 542.20 (C<sub>25</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub><sup>80</sup>Br), 463.20 (C<sub>20</sub>H<sub>35</sub>N<sub>2</sub>O<sub>5</sub><sup>80</sup>Br),

442.21 (C<sub>22</sub>H<sub>40</sub>N<sub>3</sub>O<sub>6</sub>), 424.19 (C<sub>22</sub>H<sub>38</sub>N<sub>3</sub>O<sub>5</sub>), 329.13 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>5</sub>), 311.06 (C<sub>16</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>), 283.15 (C<sub>15</sub>H<sub>27</sub>N<sub>2</sub>O<sub>3</sub>), 214.11 (C<sub>11</sub>H<sub>20</sub>NO<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 775.3257 (calcd for C<sub>36</sub>H<sub>57</sub>N<sub>4</sub>O<sub>8</sub><sup>78</sup>BrNa 775.3252).

**Veraguamide C (3):** amorphous solid; [ $\alpha$ ]<sub>D</sub><sup>23</sup> -13.0 (*c* 0.17, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  0.86-0.89 (m, 6 H), 0.94 (d, *J* = 6.4, 3 H), 0.96 (d, *J* = 6.6, 3 H), 1.00 (d, *J* = 6.4, 3 H), 1.01 (d, *J* = 7.1, 3 H), 1.03 (d, *J* = 7.0, 3 H), 1.12 (d, *J* = 6.6, 3 H), 1.26 (s, 3 H), 1.81 (m, 1 H), 1.93-2.12 (m, 7 H), 2.15-2.40 (m, 4 H), 2.95 (s, 3 H), 3.02 (s, 3 H), 3.12 (m, 1 H), 3.63 (m, 1 H), 3.86 (m, 1 H), 3.95 (d, *J* = 10.7, 1 H), 4.15 (d, *J* = 10.8, 1 H), 4.71 (m, 1 H), 4.88 (m, 1 H), 4.90 (d, *J* = 9.2, 1 H), 4.96 (m, 1 H), 6.27 (m, 1 H); ESIMS/MS *m/z* 661.35 (C<sub>36</sub>H<sub>61</sub>N<sub>4</sub>O<sub>7</sub>), 576.23 (C<sub>31</sub>H<sub>50</sub>N<sub>3</sub>O<sub>7</sub>), 496.25 (C<sub>26</sub>H<sub>46</sub>N<sub>3</sub>O<sub>6</sub>), 462.30 (C<sub>25</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 456.25 (C<sub>23</sub>H<sub>42</sub>N<sub>3</sub>O<sub>6</sub>), 438.22 (C<sub>23</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 383.26 (C<sub>20</sub>H<sub>35</sub>N<sub>2</sub>O<sub>4</sub>) 343.17 (C<sub>17</sub>H<sub>31</sub>N<sub>2</sub>O<sub>5</sub>), 325.12 (C<sub>17</sub>H<sub>29</sub>N<sub>2</sub>O<sub>4</sub>), 297.19 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 711.4302 (calcd for C<sub>37</sub>H<sub>60</sub>N<sub>4</sub>O<sub>8</sub>Na 711.4303).

**Veraguamide H (4):** amorphous solid; ESIMS/MS *m/z* 647.33 (C<sub>35</sub>H<sub>59</sub>N<sub>4</sub>O<sub>7</sub>), 562.21 (C<sub>30</sub>H<sub>48</sub>N<sub>3</sub>O<sub>7</sub>), 496.25 (C<sub>26</sub>H<sub>46</sub>N<sub>3</sub>O<sub>6</sub>), 462.30 (C<sub>25</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 442.23 (C<sub>22</sub>H<sub>40</sub>N<sub>3</sub>O<sub>6</sub>), 424.20 (C<sub>22</sub>H<sub>38</sub>N<sub>3</sub>O<sub>5</sub>), 365.24 (C<sub>19</sub>H<sub>33</sub>N<sub>2</sub>O<sub>4</sub>) 329.14 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>5</sub>), 311.07 (C<sub>16</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>), 283.18 (C<sub>15</sub>H<sub>27</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 697.4141 (calcd for C<sub>36</sub>H<sub>58</sub>N<sub>4</sub>O<sub>8</sub>Na 697.4147).

**Veraguamide I (5):** amorphous solid; ESIMS/MS *m/z* 665.37 (C<sub>36</sub>H<sub>65</sub>N<sub>4</sub>O<sub>7</sub>), 580.28 (C<sub>31</sub>H<sub>54</sub>N<sub>3</sub>O<sub>7</sub>), 500.28 (C<sub>26</sub>H<sub>50</sub>N<sub>3</sub>O<sub>6</sub>), 466.34 (C<sub>25</sub>H<sub>44</sub>N<sub>3</sub>O<sub>5</sub>), 456.25 (C<sub>23</sub>H<sub>42</sub>N<sub>3</sub>O<sub>6</sub>), 438.22 (C<sub>23</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 383.26 (C<sub>20</sub>H<sub>35</sub>N<sub>2</sub>O<sub>4</sub>) 343.17 (C<sub>17</sub>H<sub>31</sub>N<sub>2</sub>O<sub>5</sub>), 325.12 (C<sub>17</sub>H<sub>29</sub>N<sub>2</sub>O<sub>4</sub>), 297.19 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 715.4619 (calcd for C<sub>37</sub>H<sub>64</sub>N<sub>4</sub>O<sub>8</sub>Na 715.4616).

**Veraguamide J (6):** amorphous solid; ESIMS/MS *m/z* 651.35 (C<sub>35</sub>H<sub>63</sub>N<sub>4</sub>O<sub>7</sub>), 566.20 (C<sub>30</sub>H<sub>52</sub>N<sub>3</sub>O<sub>7</sub>), 500.25 (C<sub>26</sub>H<sub>50</sub>N<sub>3</sub>O<sub>6</sub>), 467.05 (C<sub>25</sub>H<sub>45</sub>N<sub>3</sub>O<sub>5</sub>), 442.23 (C<sub>22</sub>H<sub>40</sub>N<sub>3</sub>O<sub>6</sub>), 424.20 (C<sub>22</sub>H<sub>38</sub>N<sub>3</sub>O<sub>5</sub>), 365.24 (C<sub>19</sub>H<sub>33</sub>N<sub>2</sub>O<sub>4</sub>) 329.14 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>5</sub>), 311.07 (C<sub>16</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>), 283.18 (C<sub>15</sub>H<sub>27</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 699.4298 (calcd for C<sub>36</sub>H<sub>62</sub>N<sub>4</sub>O<sub>8</sub>Na 699.4303).

**Veraguamide K (7):** amorphous solid;  $[\alpha]^{23}_D$  -21.4 (*c* 0.33, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  0.85 (d, *J* = 6.9, 3 H), 0.88 (d, *J* = 7.7, 3 H), 0.90 (t, *J* = 6.5, 3 H), 0.91 (d, *J* = 6.5, 3 H), 0.99 (d, *J* = 6.5, 3 H), 1.00 (d, *J* = 6.9, 6 H), 1.04 (d, *J* = 6.5, 3 H), 1.16 (m, 1 H), 1.20 (d, *J* = 6.9, 3 H), 1.25 (t, *J* = 7.0, 3 H), 1.47 (m, 1 H), 1.47 (m, 2 H), 1.74 (m, 1 H), 1.89 (m, 1 H), 2.13-2.30 (m, 7 H), 2.41 (m, 1 H), 2.93 (d, *J* = 5.5, 1 H), 3.10 (s, 3 H), 3.13 (s, 3 H), 3.68 (m, 1 H), 3.79 (t, *J* = 6.5, 1 H), 3.89 (m, 1 H), 4.15 (m, 1 H), 4.17 (m, 1 H), 4.81 (m, 1 H), 4.86 (t, *J* = 7.0, 1 H), 4.88 (d, *J* = 10.8, 1 H), 4.90 (m, 1 H), 6.36 (d, *J* = 8.6, 1 H); ESIMS/MS *m/z* 769.24 (C<sub>37</sub>H<sub>60</sub>N<sub>4</sub>O<sub>8</sub><sup>80</sup>Br), 656.21 (C<sub>31</sub>H<sub>49</sub>N<sub>3</sub>O<sub>7</sub><sup>80</sup>Br), 559.14 (C<sub>26</sub>H<sub>42</sub>N<sub>2</sub>O<sub>6</sub><sup>80</sup>Br), 484.22 (C<sub>25</sub>H<sub>46</sub>N<sub>3</sub>O<sub>6</sub>), 443.10 (C<sub>20</sub>H<sub>33</sub>N<sub>2</sub>O<sub>4</sub><sup>78</sup>Br), 438.22 (C<sub>23</sub>H<sub>40</sub>N<sub>3</sub>O<sub>5</sub>), 371.06 (C<sub>19</sub>H<sub>35</sub>N<sub>2</sub>O<sub>5</sub>), 325.11 (C<sub>17</sub>H<sub>29</sub>N<sub>2</sub>O<sub>4</sub>), 297.20 (C<sub>16</sub>H<sub>29</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 835.3831 (calcd for C<sub>38</sub>H<sub>65</sub>N<sub>4</sub>O<sub>9</sub><sup>78</sup>BrNa 835.3827).

**Veraguamide L (8):** amorphous solid;  $[\alpha]^{22}_D$  -27.9 (*c* 0.50, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>)  $\delta$  0.85 (d, *J* = 6.7, 3 H), 0.88 (d, *J* = 6.7, 3 H), 0.89 (d, *J* = 6.6, 3 H), 0.95 (d, *J* = 6.6, 3 H), 0.97 (d, *J* = 6.5, 3 H), 0.98 (d, *J* = 6.2, 3 H), 1.00 (d, *J* = 6.9, 3 H), 1.04 (d, *J* = 1.04, 3 H), 1.17 (d, *J* = 7.1, 3 H), 1.23 (t, *J* = 7.12, 3 H), 1.46 (dt, *J* = 7.1, 6.9, 2 H), 1.51 (m, 1 H), 1.72 (m, 1 H), 1.87 (m, 1 H), 2.00 (m, 1 H), 2.06 (m, 1 H), 2.15 (m, 1 H), 2.19 (m, 1 H), 2.20 (m, 1 H), 2.22 (m, 1 H), 2.24 (m, 1 H), 2.26 (m, 1 H), 2.39 (m, 1 H), 2.93 (d, *J* = 3.81, 1 H), 3.09 (s, 3 H), 3.13 (s, 3 H), 3.67 (dt, *J* = 7.7, 7.5, 1 H), 3.78 (t, *J* = 6.6, 1 H), 3.85 (dt, *J* = 7.7, 7.5, 1 H), 4.14 (m, 1 H), 4.17 (m, 1 H), 4.81 (dt, *J* = 6.2, 7.8, 1 H), 4.82 (d, *J* = 8.6, 1 H), 4.85 (d, *J* = 10.5, 1 H), 4.87 (d, *J* = 10.3, 1 H), 4.90 (dd, *J* = 8.5, 6.3, 1 H), 6.37 (d, *J* = 8.8, 1 H); ESIMS/MS *m/z* 755.24 (C<sub>36</sub>H<sub>58</sub>N<sub>4</sub>O<sub>8</sub><sup>80</sup>Br), 642.21 (C<sub>30</sub>H<sub>47</sub>N<sub>3</sub>O<sub>7</sub><sup>80</sup>Br), 545.14 (C<sub>25</sub>H<sub>40</sub>N<sub>2</sub>O<sub>6</sub><sup>80</sup>Br), 470.21 (C<sub>24</sub>H<sub>44</sub>N<sub>3</sub>O<sub>6</sub>), 445.11 (C<sub>20</sub>H<sub>33</sub>N<sub>2</sub>O<sub>4</sub><sup>80</sup>Br), 424.21 (C<sub>22</sub>H<sub>38</sub>N<sub>3</sub>O<sub>5</sub>), 357.06 (C<sub>18</sub>H<sub>33</sub>N<sub>2</sub>O<sub>5</sub>), 311.10 (C<sub>16</sub>H<sub>27</sub>N<sub>2</sub>O<sub>4</sub>), 283.20 (C<sub>15</sub>H<sub>27</sub>N<sub>2</sub>O<sub>3</sub>); HRESIMS [M+Na]<sup>+</sup> *m/z* 821.3673 (calcd for C<sub>38</sub>H<sub>63</sub>N<sub>4</sub>O<sub>9</sub><sup>78</sup>BrNa 821.3671).

#### 2.4.4 Hydrogenation, Acid Hydrolysis, and Marfey's Analysis

Veraguamide A (**1**, 1 mg) was dissolved in 1 mL of EtOH and treated with a small amount of 10% Pd/C and then placed under an atmosphere of H<sub>2</sub> (g) for 5 h. The reaction product was treated with 1.5 mL of 6 N HCl in a microwave reactor at 160 °C for 5 min. An aliquot (~300 µg) of the hydrolysate was dissolved in 300 µL of 1 M sodium bicarbonate, and then 16 µL of 1% D-FDAA (1-fluoro-2,4-dinitrophenyl-5-D-alanine amide) was added in acetone. The solution was maintained at 40 °C for 90 min at which time the reaction was quenched by the addition of 50 µL of 6 N HCl. The reaction mixture was diluted with 200 µL of CH<sub>3</sub>CN and 10 µL of the solution was analyzed by LC-ESIMS.

The Marfey's derivatives of the hydrolysate and standards were analyzed by RP HPLC using a Phenomenex Luna 5 µm C<sub>18</sub> column (4.6 x 250 mm). The HPLC conditions began with 10% CH<sub>3</sub>CN/90% H<sub>2</sub>O acidified with 0.1% formic acid (FA) followed by a gradient profile to 50% CH<sub>3</sub>CN/ 50% H<sub>2</sub>O acidified with 0.1% FA over 85 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. The retention times of authentic acid D-FDAA derivatives were D-Pro (66.49), L-Pro (69.30), D-Val (78.45), D-N-Me-Val (86.61), L-Val (88.00), and L-N-Me-Val (91.66); the hydrolysate product gave peaks with retention times of 69.49, 88.07, and 91.74 min, according to L-Pro, L-Val and L-N-Me-Val, respectively.

#### 2.4.5 Preparation and GCMS Analysis of 2-Hydroxy-3-methylpentanoic Acid (H3mpa)

Veraguamide A (**1**, 1 mg) was dissolved in 1 mL of ethanol and treated with a small amount of 10% Pd/C and H<sub>2</sub> (g). The reaction product was then treated with 1.5 mL of 6 N HCl at 110 °C for 16 hrs. The reaction product was dried under N<sub>2</sub> (g) then dissolved in 0.5 mL of MeOH and Et<sub>2</sub>O and treated with diazomethane. L-Ile (20 mg) was dissolved in 5 mL of cold (0 °C) 0.2 N HClO<sub>4</sub>, and then 2 mL of NaNO<sub>2</sub> (aq) was added with rapid stirring. The reaction mixture was stored at room temperature for 1 h. The solution was boiled for 3 min, cooled to room temperature, and then saturated with NaCl. The mixture was extracted three times with

Et<sub>2</sub>O, and the Et<sub>2</sub>O layer was then dried under N<sub>2</sub> (g) to yield the oily 2*S*,3*S*-H3mpa. An aliquot was dissolved in 1.5 mL of MeOH and Et<sub>2</sub>O and treated with diazomethane. The product was then dried under N<sub>2</sub> (g). Correspondingly, 2*R*,3*R*-H3mpa, 2*S*,3*R*-H3mpa, and 2*R*,3*S*-H3mpa were synthesized with the same procedure from D-Ile, L-*allo*-Ile, and D-*allo*-Ile, respectively.

Each authentic stereoisomer of H3mpa was dissolved in CH<sub>2</sub>Cl<sub>2</sub> with retention times measured by GC using a Cyclosil B column (Agilent Technologies J&W Scientific, 30 m x 0.25 mm) under the following conditions: the initial oven temperature was 35 °C, and held for 15 min, followed by a ramp from 35 to 60 °C at a rate of 1 °C/min, and another ramp to 170 °C at a rate of 10 °C/min, and held at 170 °C for 5 min. The retention time of the H3mpa residue in acid hydrolysate of **1** matched with 2*S*,3*S*-H3mpa (45.63 min; 2*S*,3*R*-H3mpa, 44.86 min; 2*R*,3*S*-H3mpa, 45.06; 2*R*,3*R*-H3mpa, 45.26).

#### 2.4.6 Preparation and GCMS Analysis of Methyl 3-Hydroxy-2-Methyloctanoate (Hmoaa)

2*S*,3*S*-Hmoaa and 2*S*,3*R*-Hmoaa were synthesized following literature conditions.<sup>18</sup> A sample of 5 mg of each product was dissolved in 2 mL of dry CH<sub>2</sub>Cl<sub>2</sub> and treated with 0.122 mmol of triethylamine, 16.4 mmol of DMAP and each was separately treated with 0.126 mmol of both *R*-MTPA-Cl and *S*-MTPA-Cl for 17 h at room temperature. Each reaction was quenched with 2.5 mL of 1N HCl and extracted with Et<sub>2</sub>O to produce the four diastereomeric standards. An aliquot of the hydrolysate of veraguamide A (**1**, 0.3 mg) was dissolved in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and treated with 7.32 μmol of triethylamine, 0.964 mol of DMAP and 7.56 μmol of *S*-MTPA for 18 h at room temperature.

The four stereoisomeric standards of Hmoaa as well as the derivatized hydrolysate product of compound **1** were dissolved in CH<sub>2</sub>Cl<sub>2</sub> and analyzed by GCMS as described below. A DB-5MS GC column (Agilent Technologies J&W Scientific, 30 m x 0.25 mm) was used with the following conditions: initial oven temperature was 35 °C, held for 2 min, followed by a ramp

from 35 to 140 °C at a rate of 25 °C/min, followed by another ramp to 165 °C at a rate of 1 °C/min and held for 15 min before it was finally ramped up to a temperature of 190 °C at 1 °C/min. The retention time of the Hmoaa residue from the derivatized hydrolysate mixture of **1** matched that of 2*S*,3*R*-Hmoaa which was reacted with *S*-MTPA-Cl (47.13 min; 2*S*,2*S*-Hmoaa reacted with *S*-MTPA-Cl, 48.17 min; 2*S*,3*R*-Hmoaa reacted with *R*-MTPA-Cl, 48.13 min; 2*S*,3*S*-Hmoaa reacted with *R*-MTPA-Cl, 47.63 min).

#### 2.4.7 Tandem Mass Spectrometry Data Acquisition and Preprocessing

For the ion-trap data acquisition, each compound was prepared to a 1 µM solution using 50:50 MeOH:H<sub>2</sub>O with 1% AcOH as solvent, and underwent nanoelectrospray ionization on a Biversa Nanomate (pressure: 0.3 p.s.i., spray voltage: 1.4–1.8 kV). Ion trap spectra were acquired on a Finnigan LTQ-MS (Thermo-Electron Corporation) running Tune Plus software version 1.0. Ion tree datasets were collected using automatic mode, in which, the [M+H]<sup>+</sup> of each compound was set as the parent ion. MS<sup>n</sup> data were collected with the following parameters: maximum breadth, 50; maximum MS<sup>n</sup> depth, 3. At n = 2, isolation width, 4; normalized energy, 50. At n = 3, isolation width, 4; normalized energy 30. The Thermo-Finnigan files (in RAW format) were then converted to an mzXML file format using the ReAdW (<http://tools.proteomecenter.org/>) and subject to analysis using algorithms as well as manual interpretation.<sup>23-25</sup>

#### 2.4.8 Cytotoxicity Assay

H-460 cells were added to 96-well plates at 3.33 x 10<sup>4</sup> cells/mL of Roswell Park Memorial Institute (RPMI) 1640 medium with fetal bovine serum (FBS) and 1% penicillin/streptomycin. The cells, in a volume of 180 µL per well, were incubated overnight (37 °C, 5% CO<sub>2</sub>) to allow recovery before treatment with test compounds. Compounds were dissolved in DMSO to a stock concentration of 10 mg/mL. Working solutions of the compounds

were made in RPMI 1640 medium without FBS, with a volume of 20  $\mu\text{L}$  added to each well to give a final compound concentration of either 30  $\mu\text{g}/\text{mL}$  or 3  $\mu\text{g}/\text{mL}$ . An equal volume of RPMI 1640 medium with FBS was added to wells designated as negative controls for each plate. Plates were incubated for approximately 48 h before staining with MTT. Using a ThermoElectron Multiskan Ascent plate reader, plates were read at 570 and 630 nm.

#### 2.4.9 DNA Extraction, Amplification, and Sequencing

Algal biomass (~50 mg) was partly cleaned under an Olympus VMZ dissecting microscope. The biomass was pretreated using TE (10 mM Tris; 0.1M EDTA; 0.5 % SDS; 20  $\mu\text{g}/\text{mL}$  RNase)/lysozyme (1 mg/mL) at 37 °C for 30 min followed by incubation with proteinase K (0.5 mg/mL) at 50 °C for 1 h. Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega) following the manufacturer's specifications. DNA concentration and purity was measured on a DU® 800 spectrophotometer (Beckman Coulter). The 16S rRNA genes were PCR-amplified from isolated DNA using the modified lineage-specific primers, OT106F 5'-GGACGGGTGAGTAACGCGTGA-3' and OT1445R 5'-AGTAATGACTTCGGGCGTG-3'. The PCR reaction volumes were 25  $\mu\text{L}$  containing 0.5  $\mu\text{L}$  (~50 ng) of DNA, 2.5  $\mu\text{L}$  of 10  $\times$  PfuUltra IV reaction buffer, 0.5  $\mu\text{L}$  (25 mM) of dNTP mix, 0.5  $\mu\text{L}$  of each primer (10  $\mu\text{M}$ ), 0.5  $\mu\text{L}$  of PfuUltra IV fusion HS DNA polymerase and 20.5  $\mu\text{L}$  dH<sub>2</sub>O. The PCR reactions were performed in an Eppendorf® Mastercycler® gradient as follows: initial denaturation for 2 min at 95 °C, 25 cycles of amplification, followed by 20 sec at 95 °C, 20 sec at 55 °C and 1.5 min at 72 °C, and final elongation for 3 min at 72 °C. PCR products were purified using a MinElute® PCR Purification Kit (Qiagen) before subcloning using the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen) following the manufacturer's specifications. Plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (Qiagen) and sequenced with M13 primers. The 16S rRNA gene sequences are available in the DDBJ/EMBL/GenBank databases under acc. No. HQ900689.

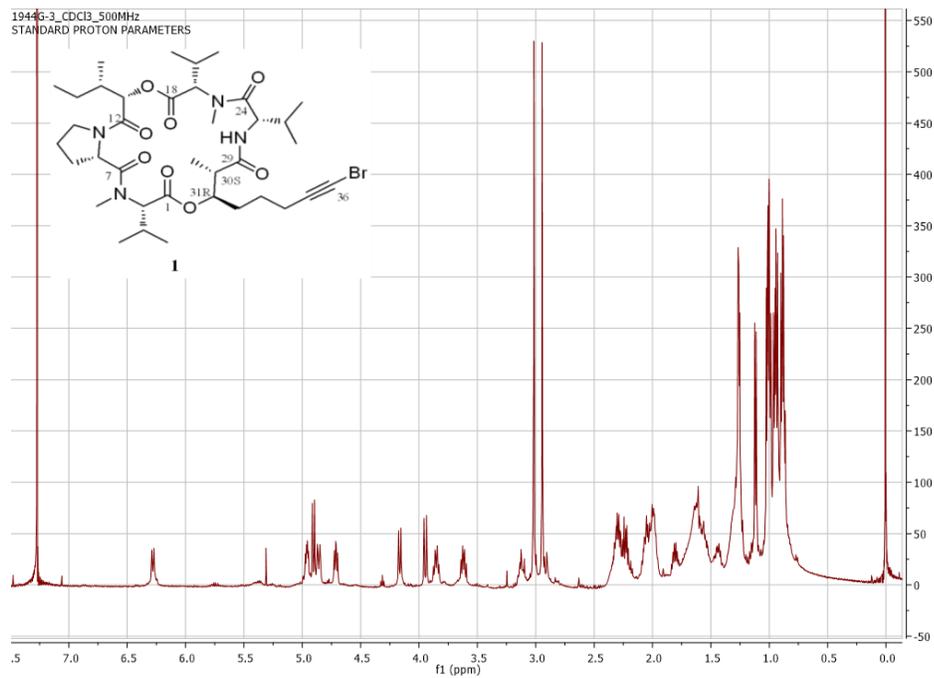
#### 2.4.10 Phylogenetic Inference

The 16S rRNA gene sequence of PAC-17-FEB-10-2 was aligned with evolutionary informative cyanobacteria using the L-INS-I algorithm in MAFFT 6.717<sup>36</sup> and refined using the SSU secondary structures model for *Escherichia coli* J01695<sup>37</sup> without data exclusion. The best-fitting nucleotide substitution model optimized by maximum likelihood was selected using corrected Akaike/Bayesian Information Criterion (AIC<sub>c</sub>/BIC) in jModeltest 0.1.1.<sup>38</sup> The evolutionary histories of the cyanobacterial genes were inferred using Maximum likelihood (ML) and Bayesian inference algorithms. The ML inference was performed using GARLI 1.0<sup>39</sup> for the GTR+I+G model assuming a heterogeneous substitution rates and gamma substitution of variable sites (proportion of invariable sites (pINV) = 0.494, shape parameter ( $\alpha$ ) = 0.485, number of rate categories = 4) with 1,000 bootstrap-replicates. Bayesian inference was conducted using MrBayes 3.1<sup>40</sup> with four Metropolis-coupled MCMC chains (one cold and three heated) ran for 3,000,000 generations. The first 25% were discarded as burn-in and the following data set were sampled with a frequency of every 100 generations. The MCMC convergence was detected by AWTY.<sup>41</sup>

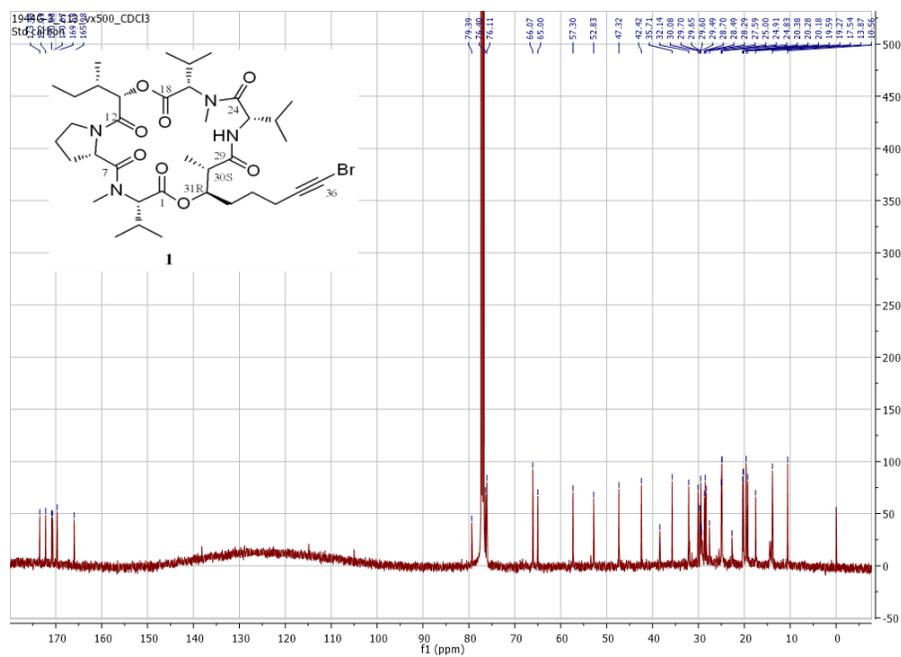
## 2.5 Acknowledgements

We thank the government of Panama for permission to make the cyanobacterial collections, W. T. Liu, N. Engene, H. Mohimani, T. Byrum, P. A. Pevzner, P. C. Dorrestein, C. Spadafora, and W. H. Gerwick for their contributions to this work. K. Tidgewell, H. Choi, P. Boudreau, M. Balunas, and J. Nunnery for helpful discussions, and the UCSD mass spectrometry facilities for their analytical services. The 500 MHz NMR  $^{13}\text{C}$  cold probe was supported by NSF CHE-0741968. We are grateful for data analysis using the NSF-funded CIPRES cluster at the San Diego Supercomputer Center. Financial support was provided by the Fogarty International Center's International Cooperative Biodiversity Groups (ICBG) program in Panama (ICBG TW006634). Funding of the Dorrestein Laboratory was provided by the Beckman Foundation and NIGMS grant NIHGM086283 (to P.C.D. and P.A.P.). W.T.L. was supported, in part, by study abroad grant SAS-98116-2-US-108 from Taiwan.

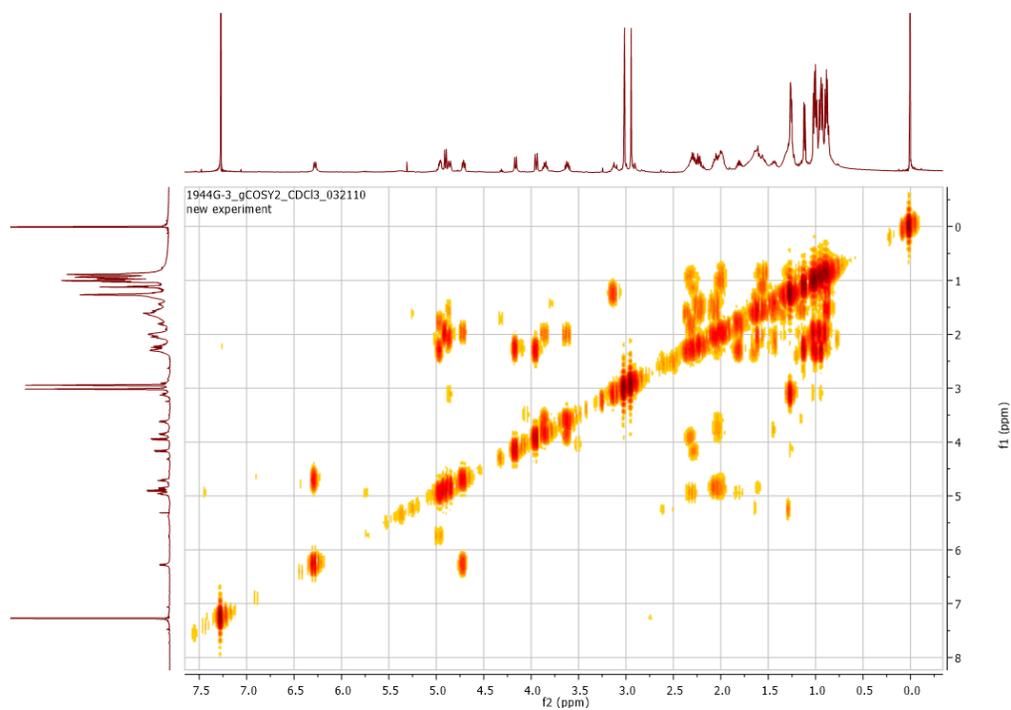
## 2.6 Appendix



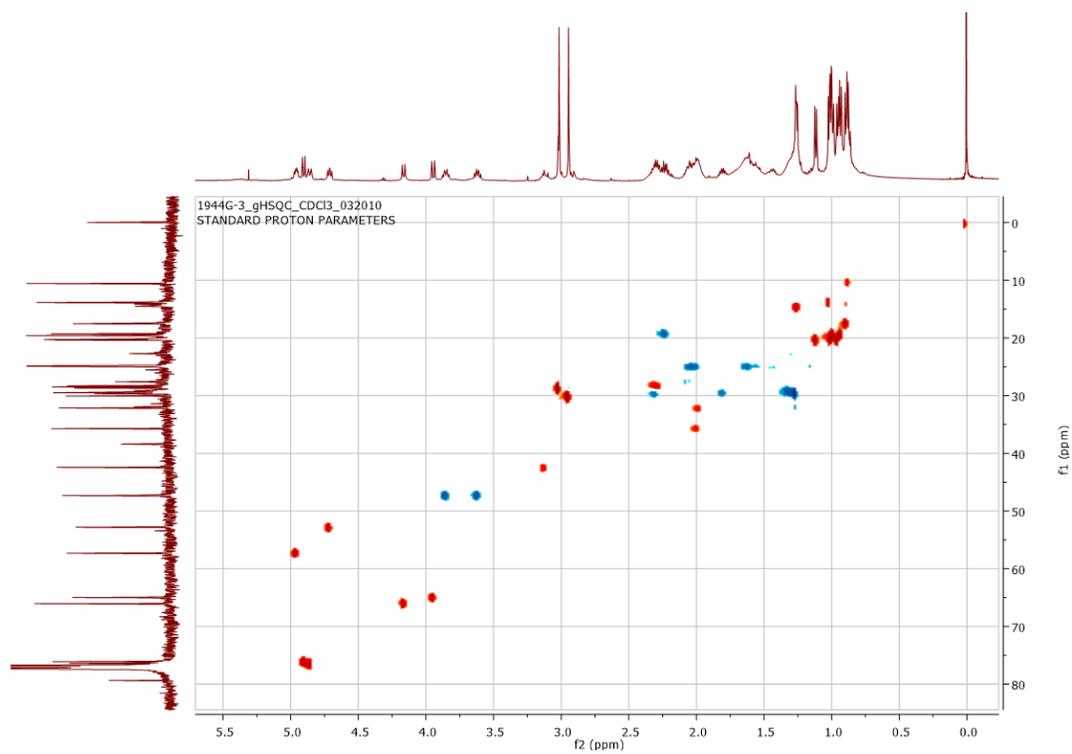
**Figure 2.6.1:**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide A



**Figure 2.6.2:**  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide A



**Figure 2.6.3:** COSY (500 MHz, CDCl<sub>3</sub>) spectrum of veraguamide A



**Figure 2.6.4:** HSQC (<sup>1</sup>H 500 MHz, CDCl<sub>3</sub>) spectrum of veraguamide A

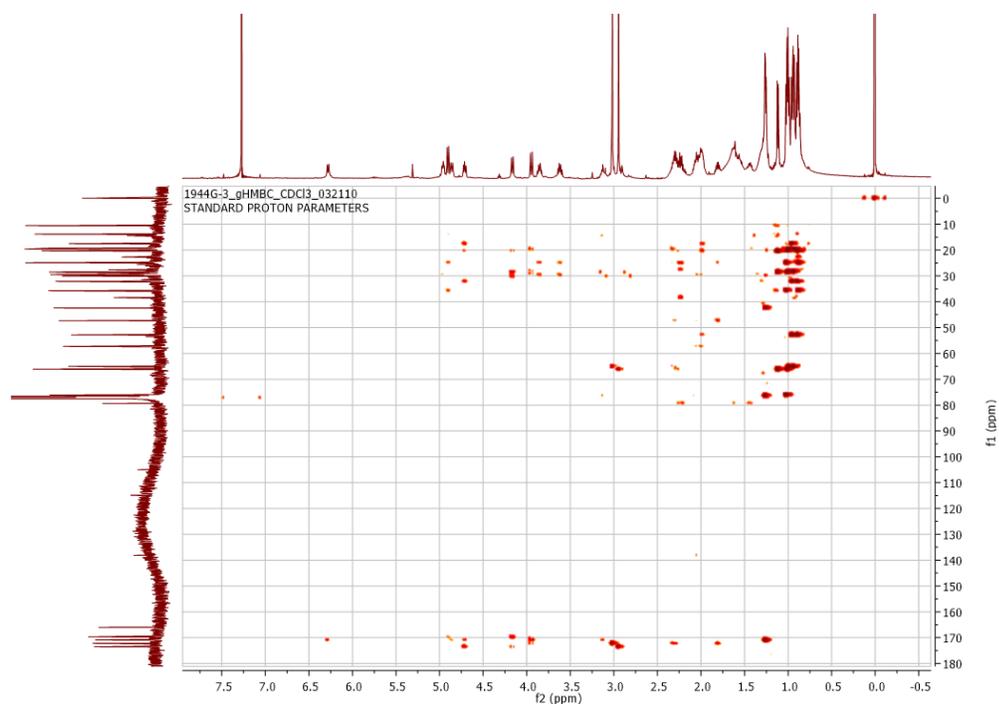


Figure 2.6.5: HMBC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide A

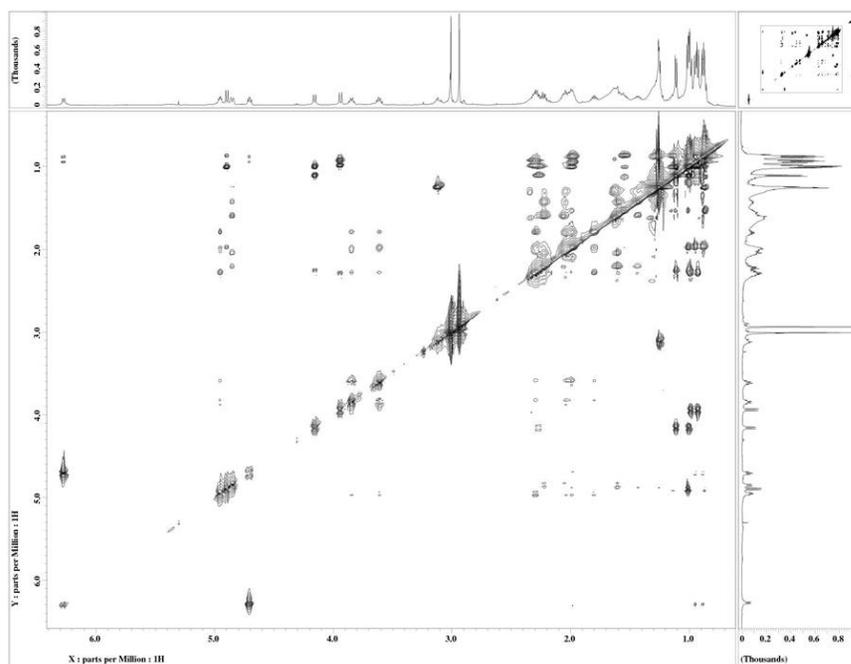
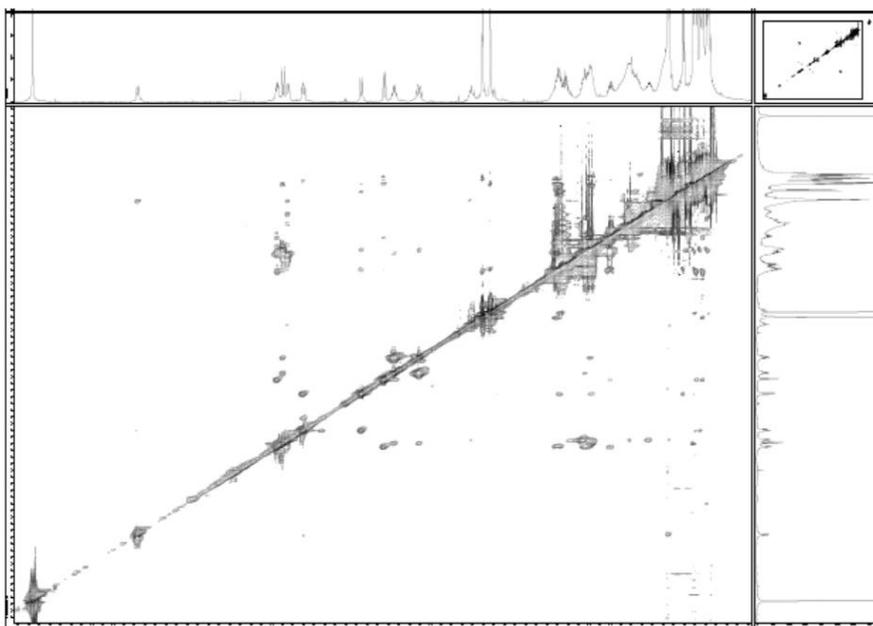


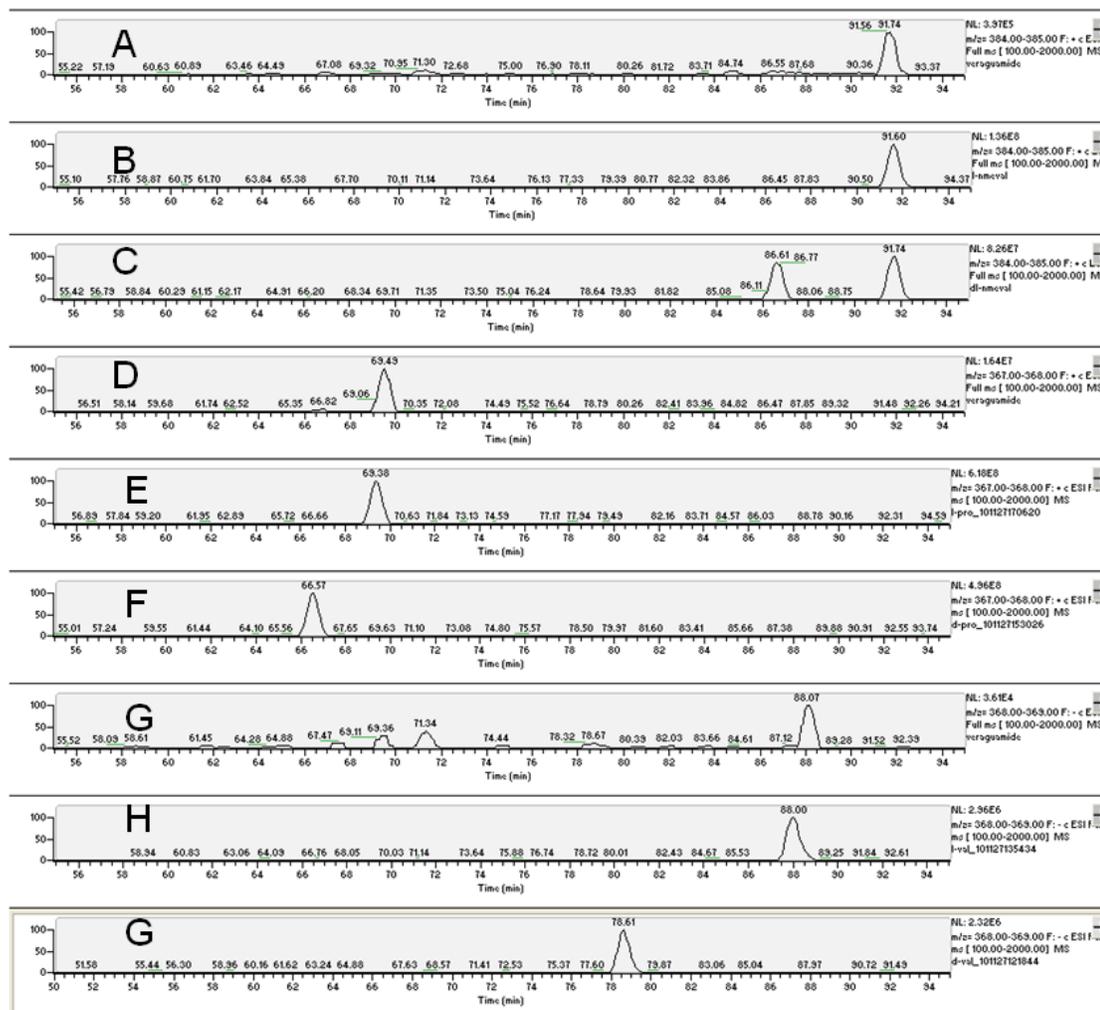
Figure 2.6.6: TOCSY (500 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide A



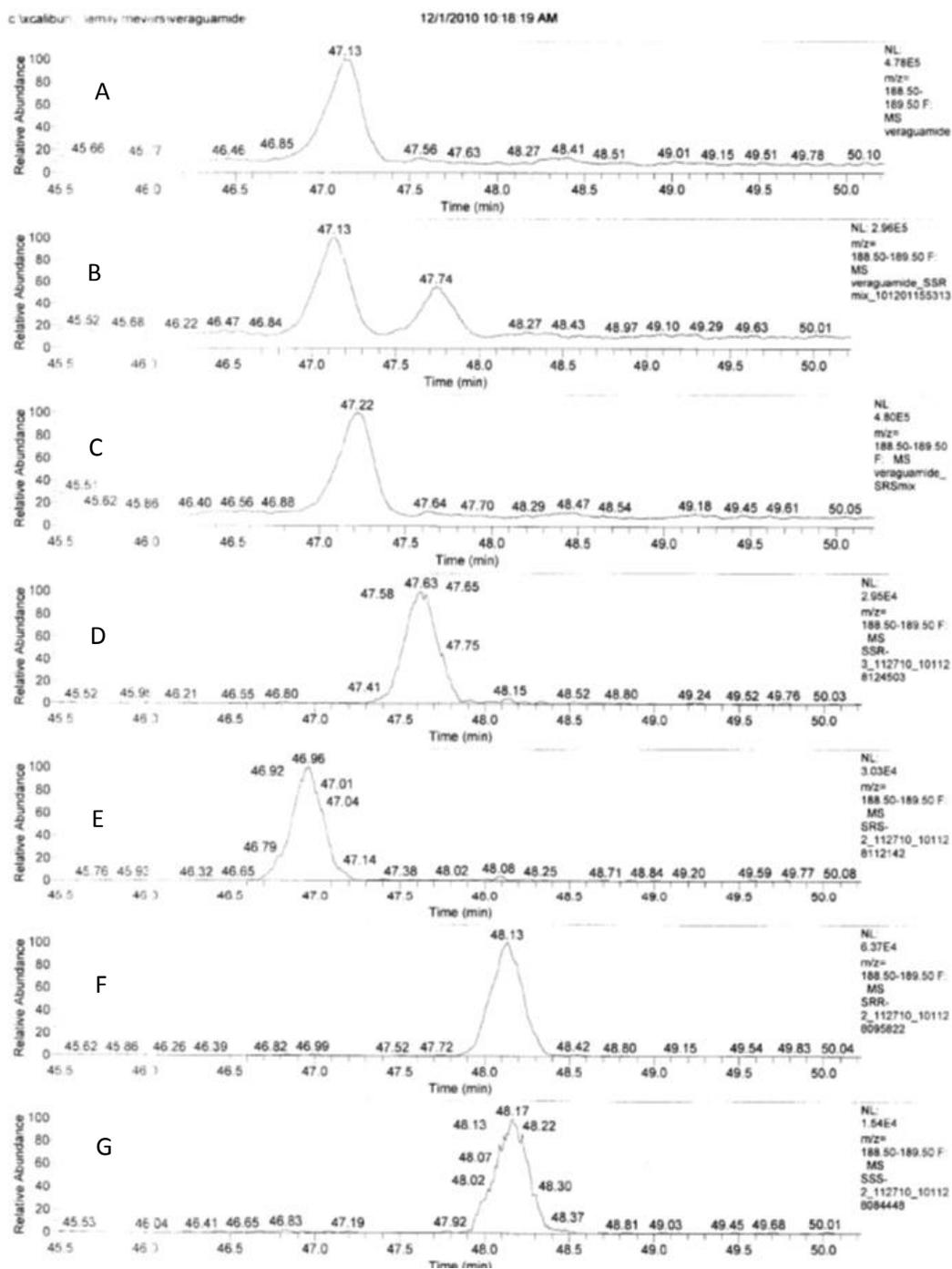
**Figure 2.6.7:** ROESY (500 MHz, CDCl<sub>3</sub>) spectrum of veraguamide A

d:\data\...ld-val\_101127121844

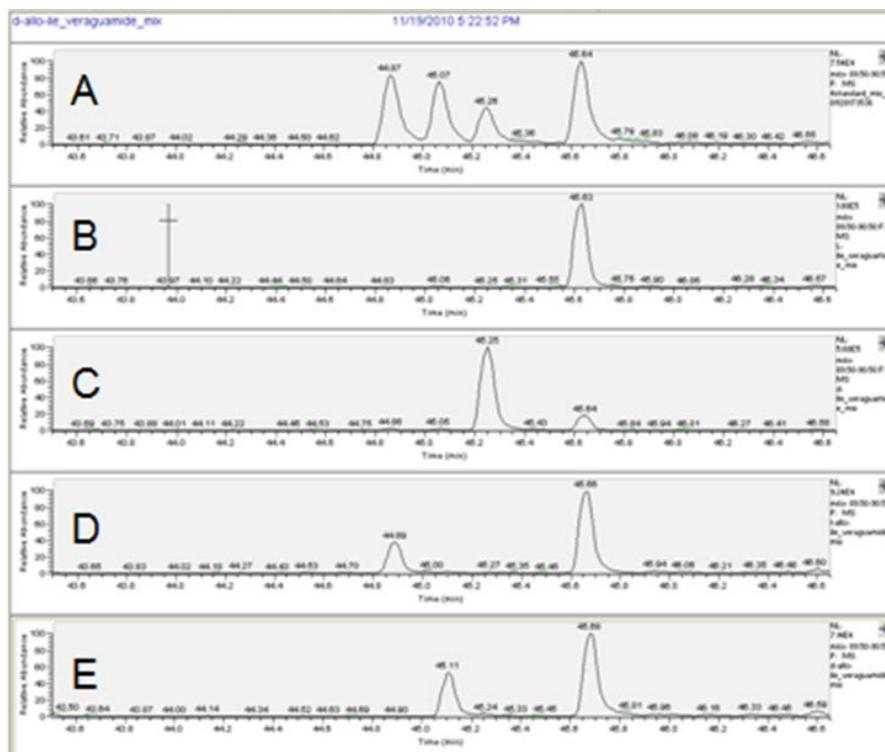
11/27/2010 12:18:44 PM



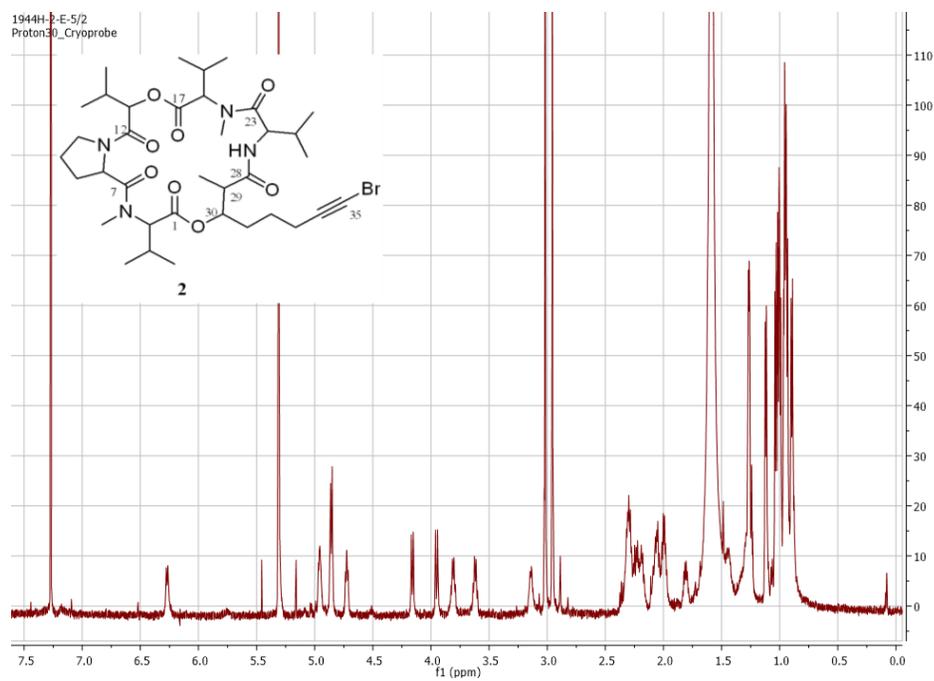
**Figure 2.6.8:** Marfey's analysis of the amino acids in veraguamide A on LCMS: (A) Marfey's derivatized hydrolysis product (ion chromatogram of 384-385 m/z); (B) L-N-MeVal (ion chromatogram of 384-385 m/z); (C) DL-N-MeVal (ion chromatogram of 384-385 m/z); (D) Marfey's derivatized hydrolysis product (ion chromatogram of 367-368 m/z); (E) L-Pro (ion chromatogram of 367-368 m/z); (F) D-Pro (ion chromatogram of 367-368 m/z); (G) Marfey's derivatized hydrolysis product (ion chromatogram of 368-369 m/z); (H) L-Val (ion chromatogram of 368-369 m/z); (I) D-Val (ion chromatogram of 368-369 m/z)



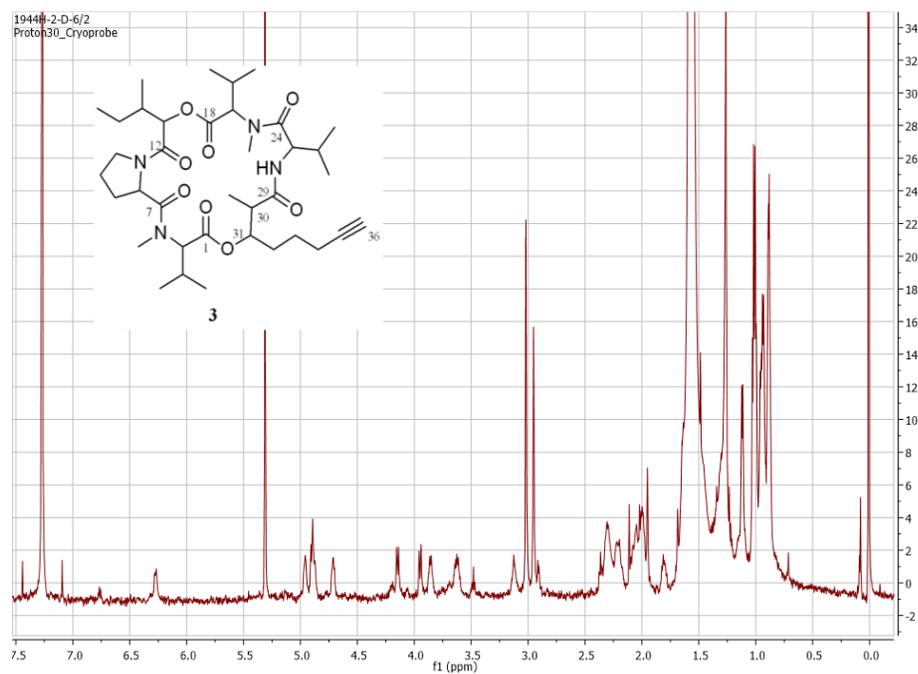
**Figure 2.6.9:** Analysis of the Mosher acid derivatized Hmoaa via GCMS in veraguamide A: (A) Hydrolysis product derivatized with *S*-MTPA-Cl; (B) Mixture of *2S,3S*-Hmoaa-3-(*S*-MTPA) and derivatized hydrolysis product; (C) Mixture of *2S,2R*-Hmoaa-3-(*R*-MTPA) and derivatized hydrolysis product; (D) *2S,3S*-Hmoaa-3-(*S*-MTPA); (E) *2S,3R*-Hmoaa-3-(*R*-MTPA); (F) *2S,2R*-Hmoaa-3-(*S*-MTPA); (G) *2S,2S*-Hmoaa-3-(*R*-MTPA)



**Figure 2.6.10:** Chiral GCMS analysis of the H3mpa residue in veraguamide A: (A) Mixture of *2S,3S*-H3mpa, *2S,3R*-H3mpa, *2R,3S*-H3mpa and *2R,3R*-H3mpa; (B) *2S,3S*-H3mpa co-injected with methylated hydrolysis product; (C) *2R,3R*-H3mpa co-injected with methylated hydrolysis product; (D) *2S,3R*-H3mpa co-injected with methylated hydrolysis product; (E) *2R,3S*-H3mpa co-injected with methylated hydrolysis product



**Figure 2.6.11:**  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide B



**Figure 2.6.12:**  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide C

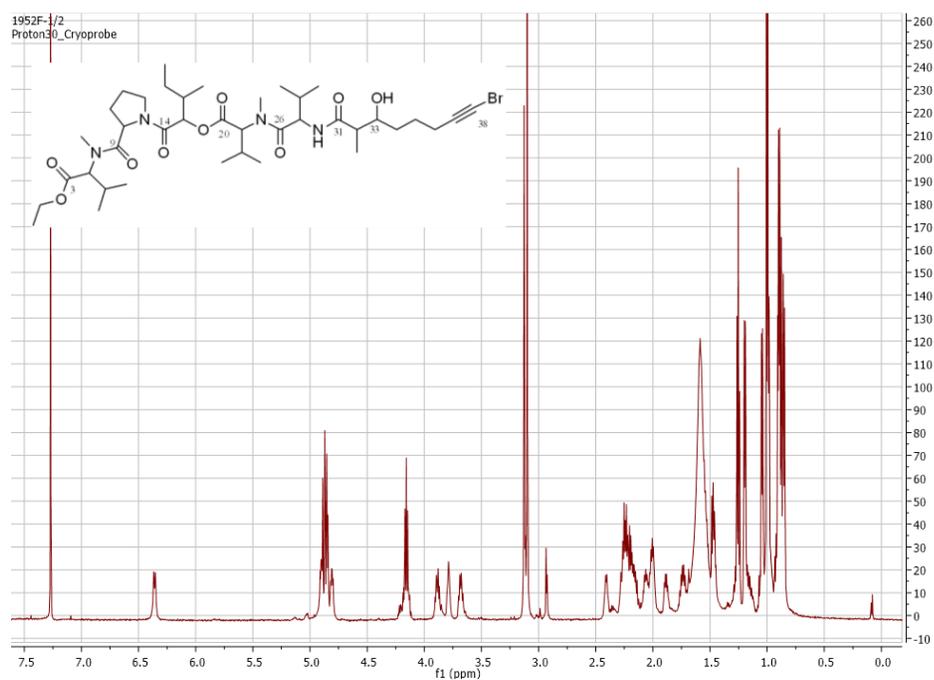


Figure 2.6.13:  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide K

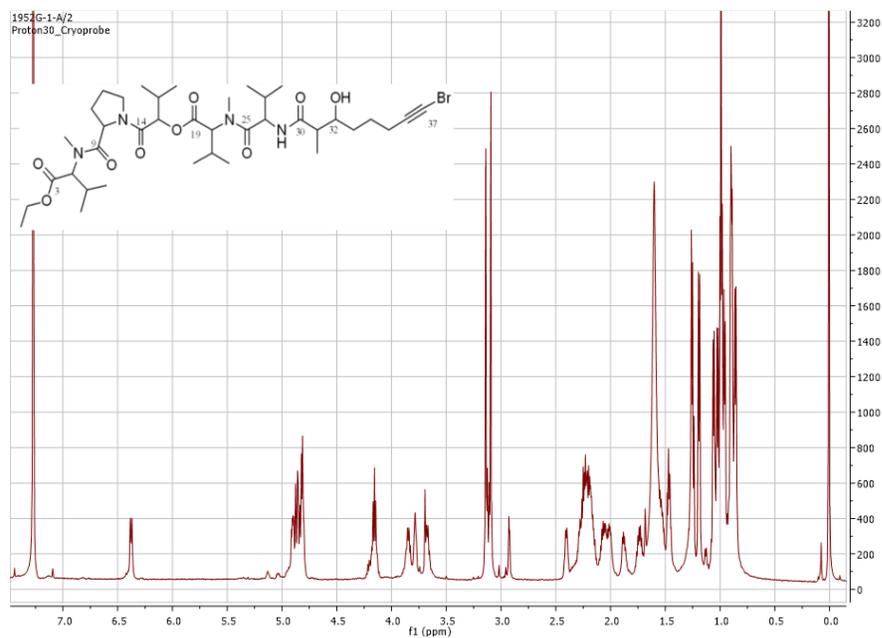
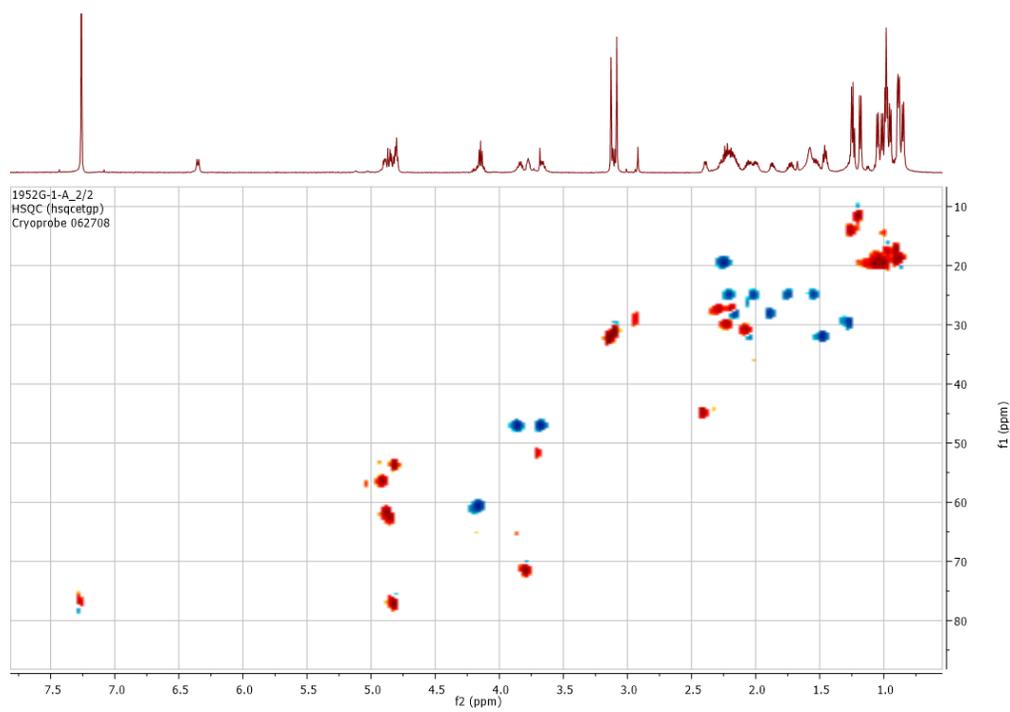
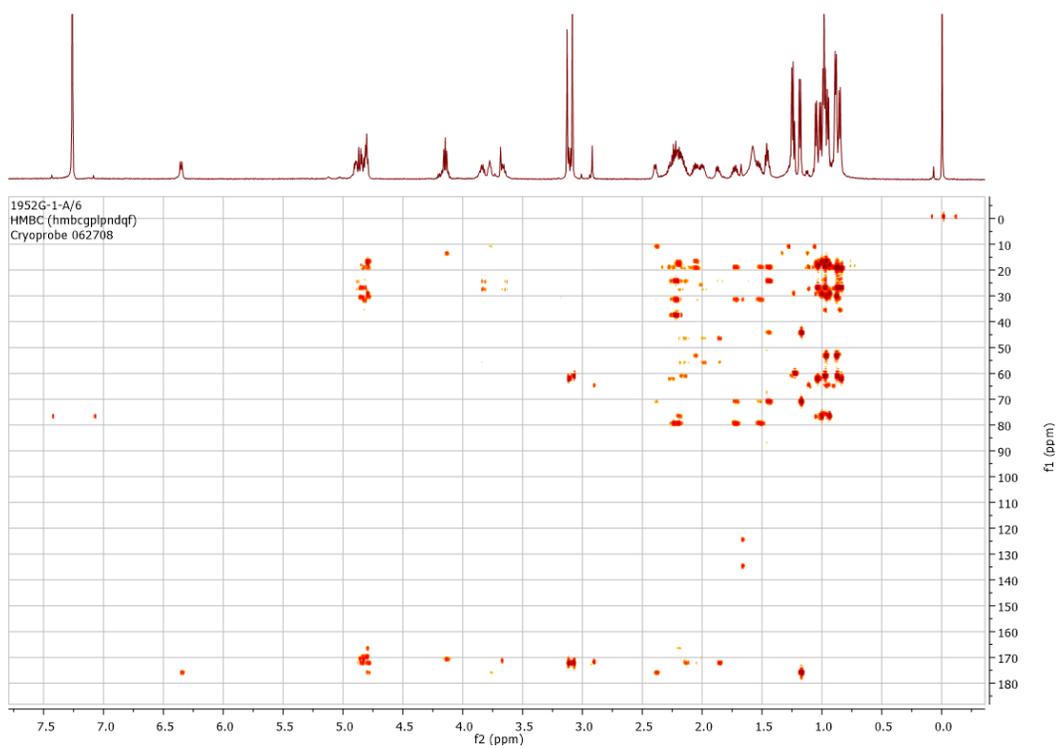


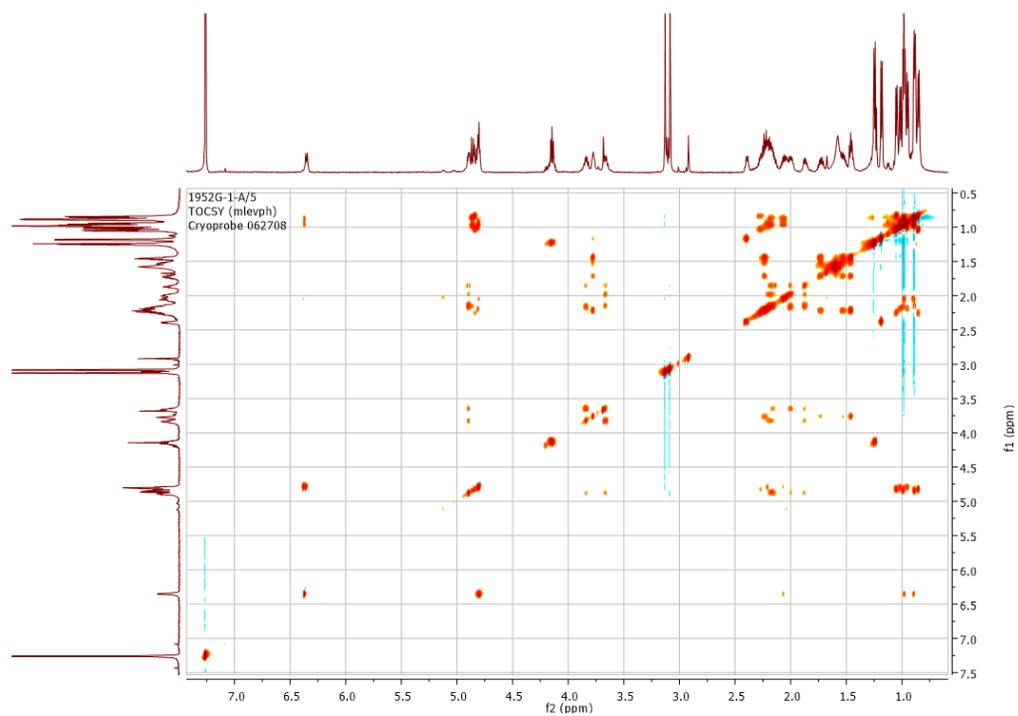
Figure 2.6.14:  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide L



**Figure 2.6.15:** HSQC ( $^1\text{H}$  600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide L



**Figure 2.6.16:** HMBC ( $^1\text{H}$  600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide L

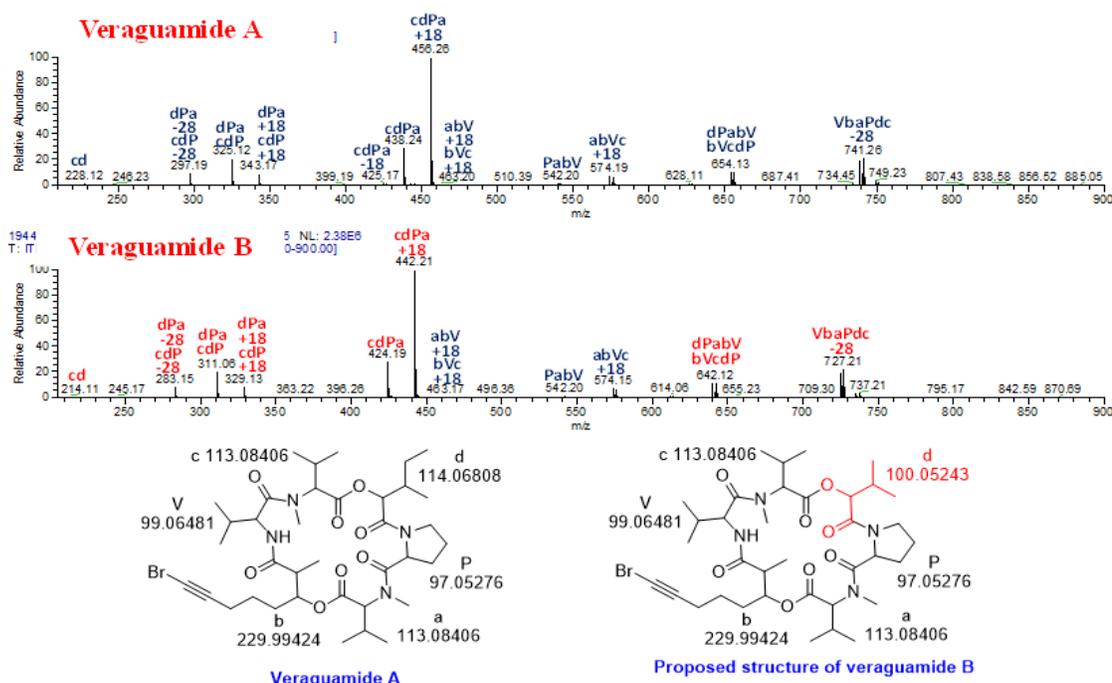


**Figure 2.6.17:** TOCSY (600 MHz,  $\text{CDCl}_3$ ) spectrum of veraguamide L

**Table 2.6.1:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for veraguamide L in  $\text{CDCl}_3$ 

Residue	position	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}}$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	
<i>N</i> -MeVal-1	1	14.3	1.23, t (7.12)	2	
	2a	60.6	4.17, m	1, 3	
	2b		4.14, m	1, 3	
	3	171.0			
	4	62.4	4.85, d (10.5)	3, 5, 6, 7, 9	
	5	27.6	2.26, m	4, 7	
	6	20.5	1.04, d (6.5)	4, 5	
	7	19.0	0.85, d (6.7)	4, 6	
Pro	8	32.2	3.13, s	4, 9	
	9	172.5			
	10	56.4	4.90, dd (8.5, 6.3)	11, 12	
	11a	25.4	2.19, m	12	
	11b		2.00, m	10, 11, 13	
	12a	28.4	1.87, m	10, 13	
	12b		2.15, m	10, 11, 13	
	13a	47.1	3.85, dt (7.7, 7.5)	10, 11, 12	
Hmba	13b		3.67, dt (7.7, 7.5)	11, 12	
	14	166.7			
	15	77.3	4.82, d (8.6)	14, 16, 17	
	16	30.1	2.20, m	14, 15, 17, 18	
	17	20.0	1.00, d (6.9)	15, 16, 18	
	18	17.3	0.95, d (6.6)	15, 16, 17	
	<i>N</i> -MeVal-2	19	170.0		
		20	61.7	4.87, d (10.3)	19, 21, 23, 25
21		27.6	2.22, m	20, 23	
22		17.8	0.97, d (6.5)	20, 21, 23	
23		18.9	0.88, d (6.7)	20, 21, 22	
24		30.7	3.09, s	20, 25	
Val		25	172.5		
		26	53.6	4.81, dt (6.2, 7.8)	25, 27, 28, 29, 30
	27	30.9	2.06, m	26, 28, 29	
	28	19.8	0.98, d (6.2)	26, 27, 29	
	29	17.5	0.89, d (6.6)	26, 27, 28	
	NH-1		6.37, d (8.8)	30	
	Brominated HMOYA	30	176.0		
		31	44.5	2.39, q (7.3)	30, 32, 38
32		71.7	3.78, t (6.6)	30, 38	
33		32.0	1.46, dt (7.1, 6.9)	31, 32, 34, 35	
34a		25.3	1.72, m	32, 33, 35, 36	
34b			1.51, m	32, 33, 35, 36	
35		19.8	2.24, m	33, 34, 36, 37, 38	
36		79.7			
37		38.0			
38		11.8	1.17, d (7.1)	30, 31, 32	

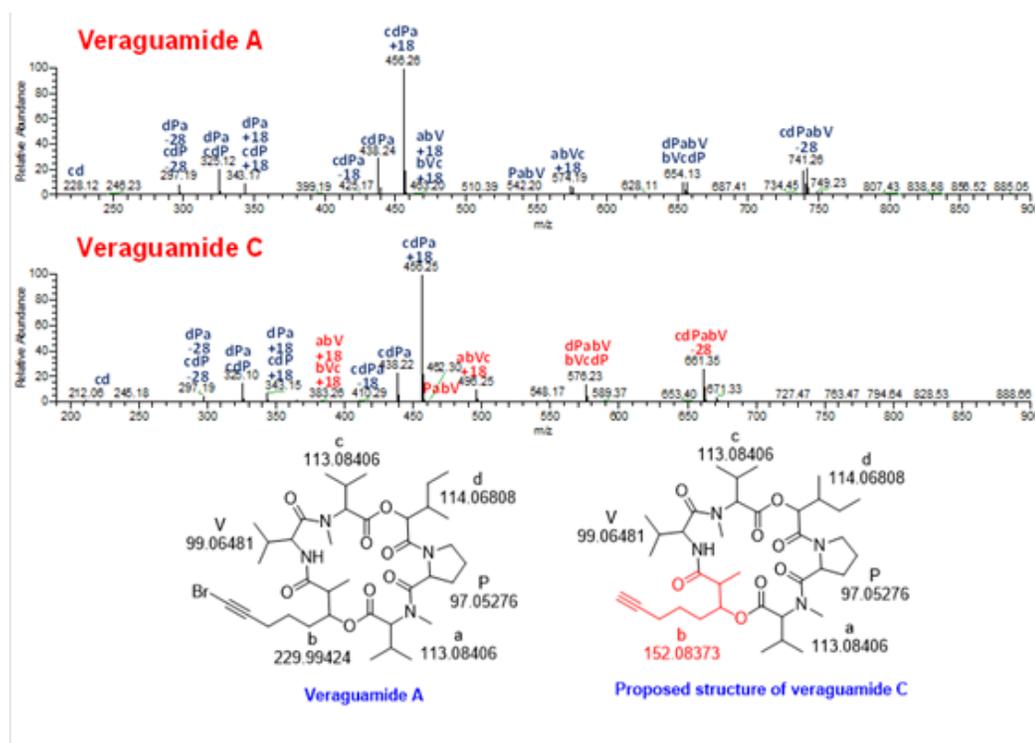
<sup>a</sup> 600 MHz for  $^1\text{H}$  NMR and HMBC; <sup>b</sup> Interpreted from HMBC and HSQC correlations



**Figure 2.6.18:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide A to B - Localize demethylation on residue d. Veraguamide B showed 14 Da loss compared to the standard compound, veraguamide A. To localize the residue that bears the 14 Da loss, fragments that bear the 14 Da shift are labeled in red, with the non-shifting fragments labeled in blue. By comparing the shifted and non-shifted ions it suggests the offset mass is on residue d. All of the fragments agreed well with this new mass annotation.

**Table 2.6.2:** Analysis of the algorithm results for veraguamide B

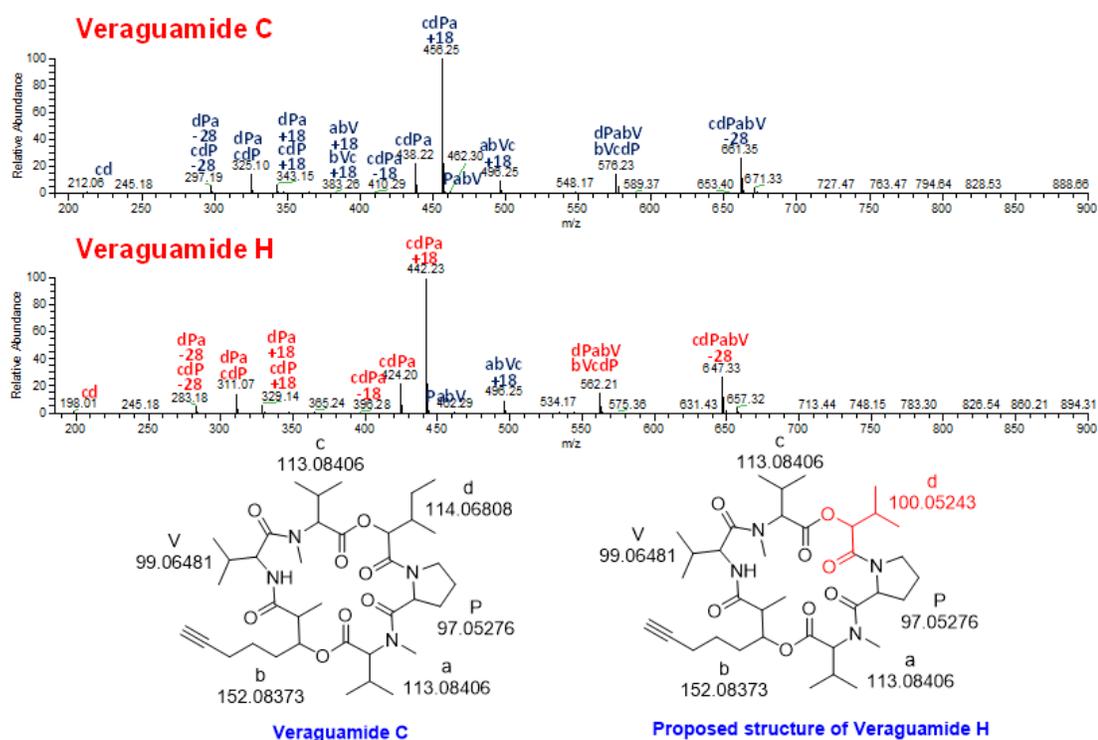
Peptide	Modified?	Mod mass	Mod position	Score
<b>Veraguamide A</b>	Yes	-14.1	5	25
<b>Demethyl Veraguamide A</b>	No	0		25
<b>Enniatin B</b>	Yes	112.8	2	24
<b>Enniatin B1</b>	Yes	98.8	6	24
<b>Enniatin B4</b>	Yes	98.8	2	24
<b>Destruxin C</b>	Yes	156.0	5	22
<b>Hydroxydestruxin B</b>	Yes	156.0	5	22
<b>Destruxin B</b>	Yes	172.9	5	21
<b>Destruxin E</b>	Yes	172.9	5	21
<b>Enniatin H</b>	Yes	98.8	1	19



**Figure 2.6.19:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide A to C - Localize debromination on residue b. Veraguamide C showed 78 Da loss compared to the standard compound, veraguamide A. In a similar manner, to localize the residue bearing the 78 Da loss, fragments that bear the 78 Da shift are labeled in red, with the non-shifting fragments labeled in blue. Results suggested loss of 78 Da on residue b or V. Since this 78 Da loss correlated well with the loss of Br atom, plus isotope pattern also suggest a non-brominated species. The best guess is the loss of 78 Da comes from residue b as a result of loss of bromine. Proposed structure is showed below.

**Table 2.6.3:** Analysis of the algorithm results for veraguamide C

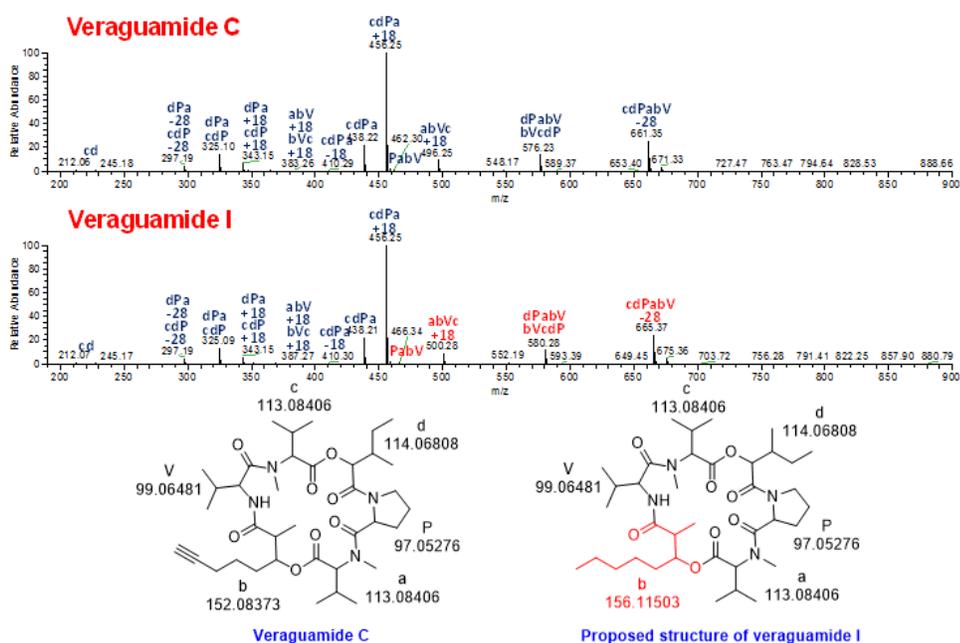
Peptide	Modified?	Mod mass	Mod position	Score
<b>Veraguamide A</b>	Yes	-78.0	2	29
<b>Kulomo opunalide 2</b>	Yes	-14.1	5	20
<b>Destruxin B1</b>	Yes	95.0	4	17
<b>Destruxin E1</b>	Yes	95.0	4	17
<b>Enniatin H</b>	Yes	35.0	4	16
<b>Enniatin B1</b>	Yes	35.0	3	14
<b>Enniatin B4</b>	Yes	35.0	5	14
<b>Enniatin A</b>	Yes	6.9	2	13
<b>Enniatin A1</b>	Yes	20.9	2	13
<b>Enniatin C</b>	Yes	6.9	2	13



**Figure 2.6.20:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide C to H - Localize demethylation on residue d. Veraguamide H showed a 14 Da loss compared to veraguamide C. To localize the residue bearing the 14 Da loss, fragments that bearing the 14 Da shift are labeled in red, with the non-shifting fragments labeled in blue. Comparing shifted and non-shifted ions suggests the offset mass is on residue d. All of the fragments agreed well with this new mass annotation. Possible structure of analog 675 is proposed as below.

**Table 2.6.4:** Analysis of the algorithm results for veraguamide H

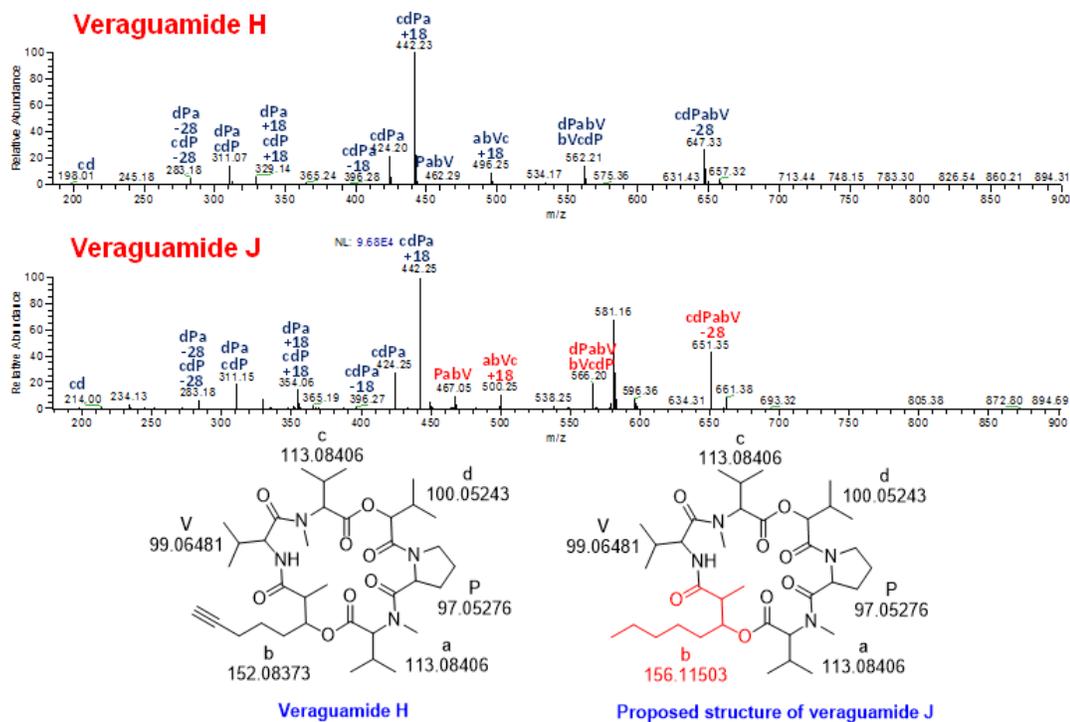
Peptide	Modified?	Mod mass	Mod position	Score
<b>Demethyl veraguamide A</b>	Yes	-77.9	2	32
<b>Enniatin B</b>	Yes	35.1	2	29
<b>Enniatin B1</b>	Yes	21.0	6	29
<b>Enniatin B4</b>	Yes	21.0	2	29
<b>Enniatin H</b>	Yes	21.0	1	27
<b>Enniatin L</b>	Yes	5.0	1	27
<b>Axinastatin 2</b>	Yes	-92.0	2	19
<b>Axinastatin 3</b>	Yes	-106.0	2	19
<b>Beauvericin E</b>	Yes	-60.9	3	19
<b>Destruxin B</b>	Yes	95.1	4	18



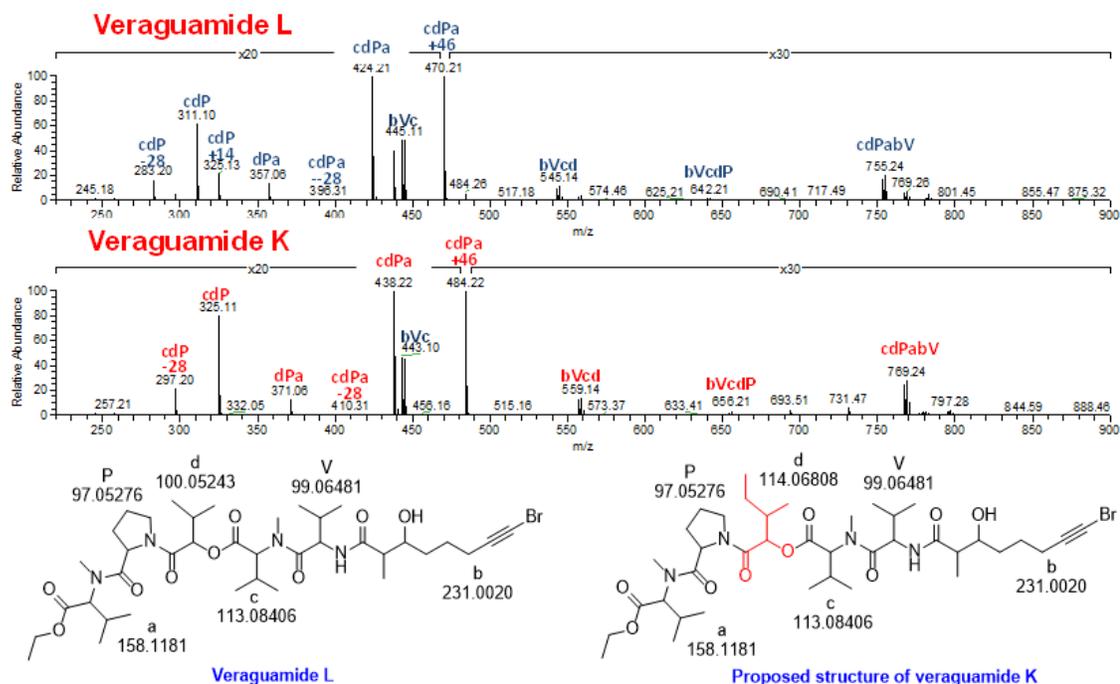
**Figure 2.6.21:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide C to I—veraguamide I showed a 4 Da adduct compared to Veraguamide C. In a similar manner, to localize the residue that bears the 4 Da adduct, fragments bearing the 4 Da shift are labeled in red, with the non-shifting fragments labeled in blue. Results suggested the gain of 4 Da is on residue b or V. The best guess is the reduction of the triple bond on residue b.

**Table 2.6.5:** Analysis of the algorithm results for veraguamide I

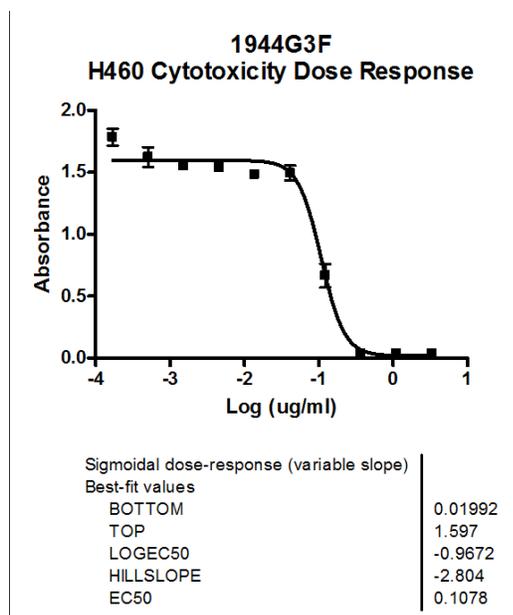
Peptide	Modified?	Mod mass	Mod position	Score
<b>Veraguamide A</b>	Yes	-74.0	2	29
<b>Destruxin B1</b>	Yes	99.0	5	19
<b>Destruxin E1</b>	Yes	99.1	5	19
<b>Enniatin H</b>	Yes	39.0	4	17
<b>Enniatin I</b>	Yes	25.0	5	17
<b>Enniatin N</b>	Yes	-5.0	1	17
<b>MK1688</b>	Yes	11.0	1	17
<b>Isariin</b>	Yes	55.0	1	17
<b>Kahalalide E</b>	Yes	-143.1	1	16
<b>Phakelistatin 13</b>	Yes	-106.0	7	15



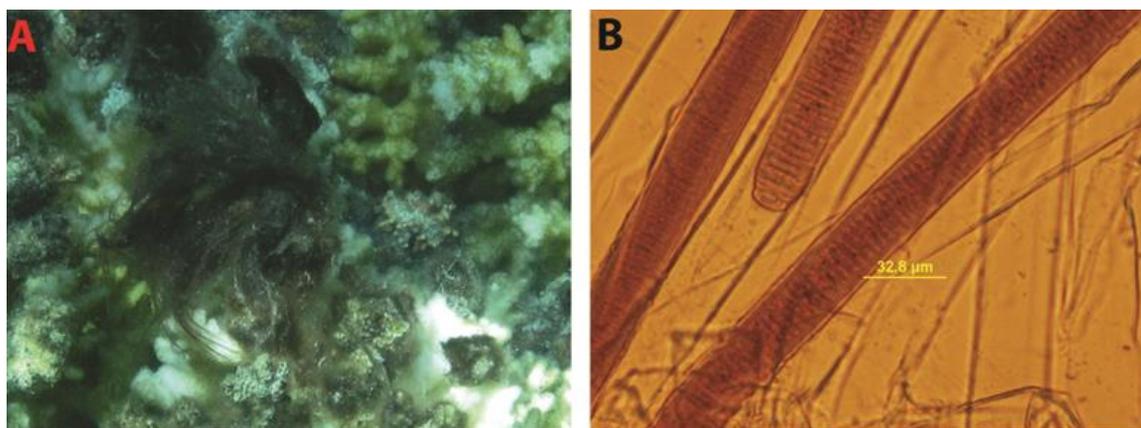
**Figure 2.6.22:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide H to J - Veraguamide J showed a 4 Da adduct compared to veraguamide H. In a similar manner, to localize the residue that bears the 4 Da adduct, fragments bearing the 4 Da shift are labeled in red, with the non-shifting fragments labeled in blue. Results suggested the gain of 4 Da is on residue b or V. The best guess is the reduction of the triple bond on residue b.



**Figure 2.6.23:** Comparison of the MS<sup>2</sup> fragmentation of veraguamide L to K - Veraguamide L showed a 14 Da adduct compared to veraguamide K. In a similar manner, to localize the residue that bears the 14 Da adduct, fragments bearing the 14 Da shift are labeled in red, with the non-shifting fragments labeled in blue. Results suggested the gain of 14 Da is on residue d.



**Figure 2.6.24:** H460 bioassay results for veraguamide A



**Figure 2.6.25:** Morphological description of veraguamide-producer (PAC-17/FEB/10-2) (a) Underwater picture (b) Photomicrographs (40X) of the cyanobacterial filaments.

## 2.7 References

- (1) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. *Curr. Opin. Biotechnol.* **2010**, *21*, 787-793.
- (2) Tan, L. T. *J. Appl. Phycol.* **2010**, *22*, 659-676.
- (3) Gerwick, W. H.; Coates, R. C.; Engene, N.; Gerwick, L.; Grindberg, R. V.; Jones, A. C.; Sorrels, C. M. *Microbe* **2008**, *3*, 277-284.
- (4) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. In *Comprehensive Natural Products Chemistry II*, Moore, B., Crews, P., Eds.; Elsevier: Oxford, UK, 2010; Vol. 8, pp 141– 188.
- (5) Gerwick, W. H.; Tan, L. T.; Sitachitta, N. *Alkaloids Chem. Biol.* **2001**, *57*, 75-184.
- (6) Tan, L. T. *Phytochemistry* **2007**, *68*, 954-979.
- (7) Hooper, G. J.; Orjala, J.; Schatzman, R. C.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 529-533.
- (8) Wan, F.; Erickson, K. L. *J. Nat. Prod.* **2001**, *64*, 143-146.
- (9) Luesch, H.; Pangilinan, R.; Yoshida, W. Y.; Moore, R.; Paul, V. J. *J. Nat. Prod.* **2001**, *64*, 304-307.
- (10) Sitachitta, N.; Williamson, R. T.; Gerwick, W. H. *J. Nat. Prod.* **2000**, *63*, 197-200.
- (11) Nogle, L. M.; Gerwick, W. H. *J. Nat. Prod.* **2002**, *65*, 21-24.
- (12) Bunyajetpong, S.; Yoshida, W. Y.; Sitachitta, N.; Kaya, K. *J. Nat. Prod.* **2006**, *69*, 1539-1542.
- (13) Tripathi, A.; Puddick, J.; Prinsep, M. R.; Lee, P. P. L.; Tan, L. K. *J. Nat. Prod.* **2009**, *72*, 29-32.
- (14) Han, B.; Goeger, D.; Maier, C. S.; Gerwick, W. H. *J. Nat. Prod.* **2005**, *70*, 3133-3139.
- (15) Balunas, M. J.; Linington, R. G.; Tidgewell, K.; Fenner, A. M.; Ureña, L. D.; Togna, G. D.; Kyle, D. E.; Gerwick, W. H. *J. Nat. Prod.* **2010**, *73*, 60-66.
- (16) Simmons, T. L.; Engene, N.; Ureña, D. L.; Romero, L. I.; Ortega-Barría, E.; Gerwick, L.; Gerwick, W. H. *J. Nat. Prod.* **2008**, *71*, 1544-1550.
- (17) Rodriguez, J.; Fernández, R.; Quiñoá, E.; Riguera, R.; Debitus, C.; Bouchet, P. *Tetrahedron Lett.* **1994**, *35*, 9239-9242.
- (18) Fernández, R.; Rodríguez, J.; Quiñoá, E.; Riguera, R.; Muñoz, Fernández-Suárez, M.; Debitus, C. *J. Am. Chem. Soc.* **1996**, *118*, 11635-11642.

- (19) Nakao, Y.; Yoshida, W. Y.; Szabo, C. M.; Baker, B. J.; Scheuer, P. J. *J. Org. Chem.* **1998**, *63*, 3272-3280.
- (20) <http://www.icbg.org/>. (Accessed March 30, 2011).
- (21) <http://whc.unesco.org/en/list/1138>. (Accessed March 30, 2011).
- (22) Mevers, E.; Tidgewell, K.; Byrum, T.; Spadafora, C.; Gerwick, W.H. In *Proceedings of the Joint Annual Meeting of the American Society of Pharmacognosy and the Phytochemical Society of North America*, St. Petersburg Beach, FL, July 10-14, 2010; Abstract P-52.
- (23) Liu, W. T.; Ng, J.; Meluzzi, D.; Bandeira, N.; Gutierrez, M.; Simmons, T. L.; Schultz, A. W.; Linnington, R. G.; Moore, B. S.; Gerwick, W. H.; Pevzner, P. A.; Dorrestein, P. C. *Anal. Chem.* **2009**, *81*, 4200-4209.
- (24) Mohimani, H.; Yang, Y. L.; Liu, W. T.; Hsieh, P. W.; Doerrestein, P. C.; Pevzner, P. A. *Proteomics* **2011**, *11*, 3642-3650.
- (25) Ng, J.; Bandeira, N.; Liu, W. T.; Ghassemian, M.; Simmons, T. L.; Gerwick, W. H.; Linnington, R. G.; Dorrestein, P. C.; Pevzner, P. A. *Nat. Methods* **2009**, *6*, 596-599.
- (26) Salvador, L. A.; Biggs, J. S.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2011**, *74*, 917-927.
- (27) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Geoger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (28) Mamer, O. A. *Methods Enzymol.* **2000**, *324*, 3-10.
- (29) Caboche, S.; Pupin, M.; Leclre, V.; Fontaine, A.; Jacques, P.; Kucherov, G. *Nucleic Acids Res.* **2008**, *36*, 326-331.
- (30) Gomont, M. M. *Ann. Sci. Nat. Bot. Ser.* **1892**, *16*, 91-264.
- (31) Johansen, J. R.; Casamatta, D. A. *Algol. Stud.* **2005**, *117*, 71-93.
- (32) Castenholz, R. W.; Rippka, R.; Herdman, M. In *Bergey's Manual of Systematic Bacteriology*; Boone, D. R., Castenholz, R. W., Eds.; Springer: New York, 2001; Vol. 1, pp 473-599.
- (33) Engene, N.; Coates, R. C.; Gerwick, W. H. *J. Phycol.* **2010**, *46*, 591-601.
- (34) Wang, D.; Jia, X.; Zhang, A. *Org. Biomol. Chem.* **2012**, *10*, 7027-7030.
- (35) Komárek, J.; Anagnostidis, K. In *Süßwasserflora von Mitteleuropa*; Büdel, B.; Gärtner, G.; Krienitz, L.; Schagerl, M., Eds.; Gustav Fischer: Jena, 2005, Vol. 19/2, pp 576-606.
- (36) Katoh, K.; Toh, H. *Brief Bioinform.* **2008**, *9*, 286-298.

- (37) Cannone, J. J., Subramanin, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pnde, N., Schang, Z., Yu, N.; Gutell R. R. *BMC Bioinformatics*. **2002**, *3*, 1471–2105.
- (38) Posada, D. *Mol. Biol. Evol.* **2008**, *25*, 1253-1256.
- (39) Zwickl, D. J. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion. Ph.D. Thesis, The University of Texas at Austin, 2006.
- (40) Ronquist, F.; Huelsenbeck, J. P. *Bioinformatics (Oxf)*. **2003**, *12*, 1572-1574.
- (41) Nylander, J. A. A.; Wilgenbusch, J. C.; Warren, D. L.; Swofford, D. L. *Bioinformatics (Oxf)* **2008**, *15*, 581-583.

## Chapter 3:

### Lyngbyabellin N, A Cytotoxic Secondary Metabolite from a Palmyra Atoll Collection of the Marine Cyanobacterium *Moorea bouillonii*

#### 3.0.1 Abstract

A new lipopeptide, lyngbyabellin N, was isolated from an extract of a filamentous marine cyanobacteria collected from Palmyra Atoll in the Central Pacific Ocean. Its planar structure and absolute configuration was elucidated by the combination of spectroscopic and chromatographic analyses as well as chemical synthesis of fragments. In addition to structural features typical of the lyngbyabellins, such as two thiazole rings and a chlorinated 2-methyloctanoate residue, this new compound possesses several interesting aspects, including an unusual *N,N*-dimethylvaline terminus and a leucine statine. Lyngbyabellin N exhibits strong cytotoxic activity against the HCT-116 colon cancer cell line ( $IC_{50} = 40.9 \pm 3.3$  nM).

#### 3.1 Introduction

Over the last 20 years, marine cyanobacteria have emerged as exceptionally prolific producers of biologically active secondary metabolites rivaling the metabolic richness of the actinobacteria.<sup>1-4</sup> Because they lack other more visible defense mechanisms, such as a hardened exterior or a cryptic habitat, and have an overall macroscopic structure, it is thought that cyanobacteria derive value from the biosynthesis of these structurally intriguing secondary metabolites for their chemical defense.<sup>5</sup> The genus *Moorea* (formally *Lyngbya* spp) is one of the most chemically prolific and has yielded such important metabolites as the apratoxins,<sup>6-10</sup> antillatoxin A,<sup>11</sup> lyngbyatoxin A,<sup>12</sup> curacin A,<sup>13</sup> barbamide,<sup>14</sup> the jamaicamides,<sup>15</sup> and the malyngamides<sup>16</sup>. In general, these structurally diverse metabolites exhibit a range of interesting biological activities, such as anti-cancer,<sup>17</sup> anti-feedant,<sup>18</sup> molluscicidal,<sup>19</sup> anti-inflammatory,<sup>20</sup>

and neuromodulatory<sup>21</sup>. The lyngbyabellins are another family of metabolites produced by *Moorea* sp. which are NRPS/PKS derived peptides and have a recognizable architecture composed of thiazole rings, hydroxy acid residues, and an acyl group with distinctive chlorination at the penultimate carbon atom.<sup>22-24</sup> Several of the lyngbyabellins are reported to exhibit moderate to potent cytotoxicity to various cancer cell lines and to exert this activity through interference with the actin system.<sup>22-24</sup>

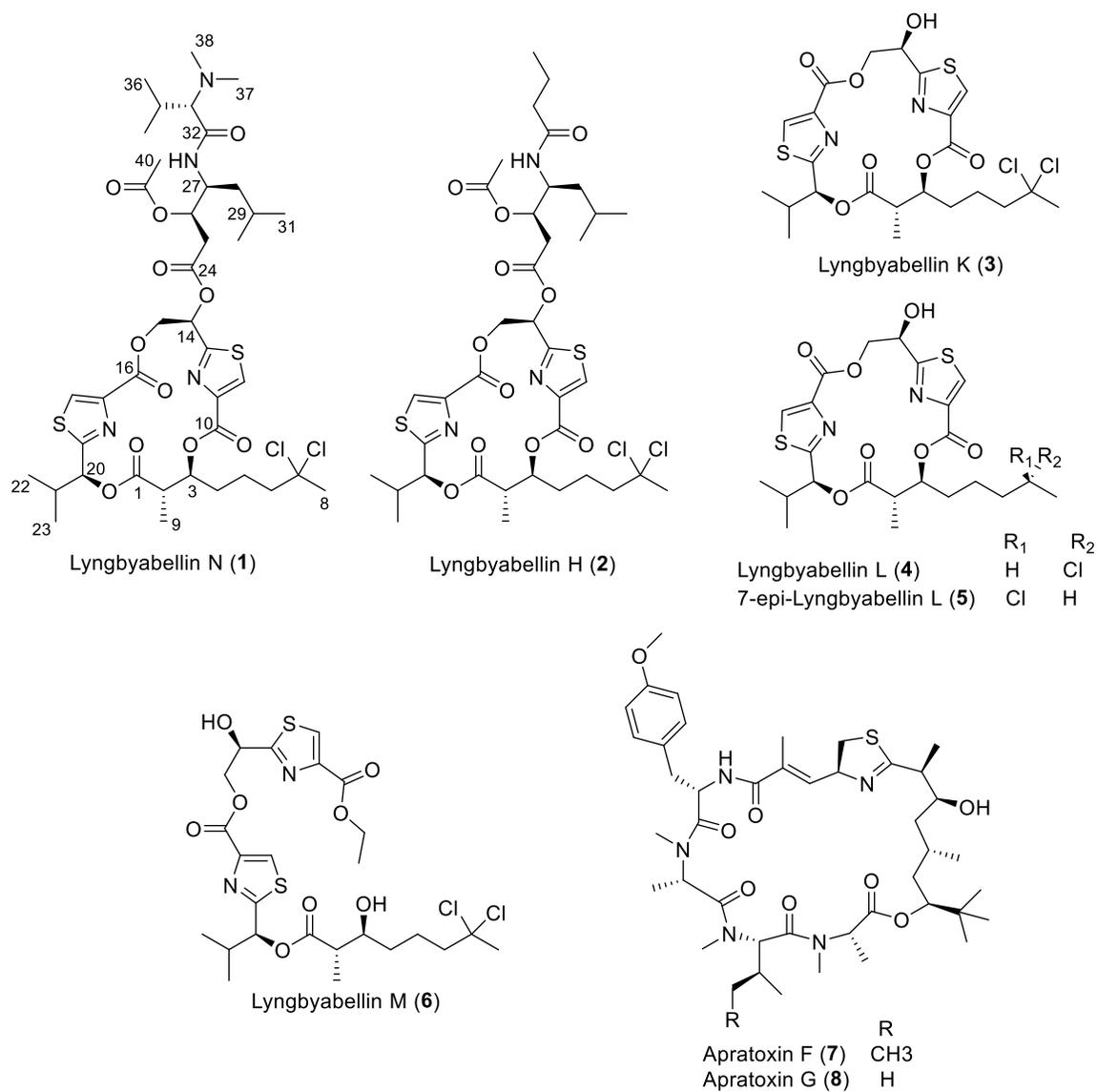
In the present work, a number of filamentous marine cyanobacteria were collected from Palmyra Atoll (approximately 1000 miles SSW of Hawaii) in 2008, and their extracts were evaluated in several biological assays. A reduced complexity fraction from an extract of *Moorea bouillonii*, was found to be highly cytotoxic to H-460 human lung cancer cells in vitro [20% survival at 3 µg/mL (fraction H)], and this was chosen for further investigation. Bioassay-guided fractionation of this extract yielded a new peptide, lyngbyabellin N (**1**), and was fully structurally defined by spectrochemical methods. Additionally, the known highly cytotoxic metabolites apratoxins F and G were also isolated from these fractions.

## 3.2 Results and Discussion

### 3.2.1 Lyngbyabellin N (**1**) Collection and Isolation

A sample of *M. bouillonii* (PAL 8/16/08-3) was collected by SCUBA in 2008 from reefs 9-15 m deep surrounding Palmyra Atoll. The ethanol-preserved material was repetitively extracted (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 2:1) and fractionated using normal-phase VLC to yield nine fractions. Three polar fractions (100% EtOAc, 25% EtOAc/MeOH, and 75% EtOAc/MeOH) were strongly cytotoxic to H-460 cancer cells (20% survival at 3 µg/mL) and were thus fractionated with reverse phase solid phase extraction (RP-SPE) followed by preparative thin layer chromatography (prepTLC) to yield 4.3 mg of highly purified lyngbyabellin N (**1**) as a pale yellow oil. These

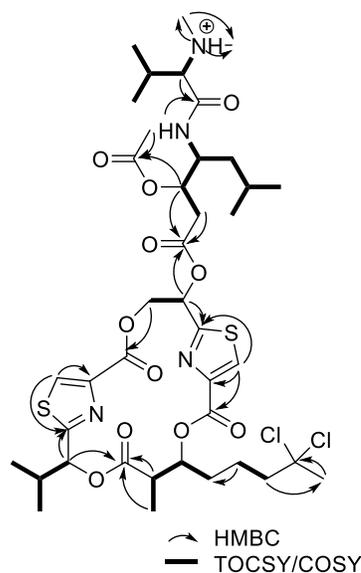
same fractions also contained apratoxins F (7) and G (8), which are known to have potent cytotoxicity.<sup>6-10</sup>



**Figure 3.1:** Lyngbyabellin N (1) along with other related metabolites

### 3.2.2 Lyngbyabellin N (**1**) Planar Structure Determination

The HR-ESITOFMS of lyngbyabellin N (**1**) showed an ion cluster at  $m/z$  905.2997/907.2977/909.2950 (calcd for  $C_{40}H_{59}Cl_2N_4O_{11}S_2$  905.2993) in a ratio of 100:80:20, indicating the presence of two chlorine atoms and 13 degrees of unsaturation. The IR spectrum of **1** suggested the presence of NH and ester/amide functionalities with absorption bands at 3436 and 1742  $cm^{-1}$ , respectively. The  $^{13}C$  NMR spectra revealed the presence of eight downfield-shifted signals of quaternary carbon atoms ( $\delta_C$  173.2, 169.4, 168.9, 167.2, 165.8, 165.5, 160.2, and 159.4), along with four carbons indicative of conjugated olefins ( $\delta_C$  145.4, 145.0, 130.2, and 129.7). In the  $^1H$  NMR spectrum, there were four downfield methyl groups ( $\delta_H$  2.77, 2.72, 2.09, and 1.91), two amide protons ( $\delta_H$  9.38, and 8.81), and two sharp downfield-shifted singlet protons at 8.46 and 8.44 ppm, which could be attributed to the presence of two 2,4,-disubstituted thiazole rings.



**Figure 3.2:** Select 2D NMR data for lyngbyabellin N

The two thiazole ring structures were confirmed by HMBC correlations from H-12 ( $\delta_{\text{H}}$  8.44) to C-11 ( $\delta_{\text{C}}$  145.4) and C-13 ( $\delta_{\text{C}}$  165.5), and H-18 ( $\delta_{\text{H}}$  8.46) to C-17 ( $\delta_{\text{C}}$  145.0) and C-19 ( $\delta_{\text{C}}$  167.2). The HMBC correlations from H-12 and H-18 of the two thiazole rings to carbonyl carbon atoms C-10 ( $\delta_{\text{C}}$  159.4) and C-16 ( $\delta_{\text{C}}$  160.2), respectively, indicated that carboxylic acid derivatives were directly attached to the 4-position of each of the thiazole rings. In addition, COSY correlations between H-14 ( $\delta_{\text{H}}$  6.27) and H-15a/b ( $\delta_{\text{H}}$  4.81/4.57), as well as HMBC correlations from H-15a/b to C-13 and C-16, established that a 1,2-dihydroxyethyl moiety formed a linkage between the two thiazole-4-carboxylate groups.

Further inspection of the  $^1\text{H}$  NMR spectrum of **1** revealed a series of upfield and highly coupled resonances reflective of an aliphatic chain. Additionally, a downfield methyl singlet at 2.09 ppm (H-8) showed HMBC correlations to a quaternary carbon atom at 92.1 ppm (C-7) as well as to a signal of a methylene carbon atom at 48.2 ppm (C-6). The chemical shift of C-7 was indicative of a *gem*-dichloro substituent, as observed in dolabellin,<sup>25</sup> hectochlorin,<sup>26</sup> and the lyngbyabellins,<sup>22-24</sup> and thus, accounted for the two chlorine atoms in the molecular formula. This moiety was extended to include an additional six carbon atoms (C-1 to C-5 and C-9) by integrated reasoning of COSY, TOCSY, HSQC, and HMBC data and identified this moiety as 7,7-dichloro-3-acyloxy-2-methyloctanoate (DCAMO). Additional HMBC correlations from H-3 of the DCAMO residue to a carbonyl carbon, C-10, of the first thiazole-4-carboxylate unit, allowed connection between these atoms through an ester bond.

Sequential COSY correlations between adjacent methine H-20 ( $\delta_{\text{H}}$  5.55) and H-21 ( $\delta_{\text{H}}$  2.24) to both doublet methyl groups H-22 and H-23, along with HMBC correlations from these two methyls groups back to C-20 and C-2, defined the side chain of a 2-hydroxy-3-methylbutyric acid (Hmba) residue. The HMBC correlations from H-20 to C-1 ( $\delta_{\text{C}}$  173.2) and C-19 ( $\delta_{\text{C}}$  167.2) supported the position of this Hmba-derived residue between C-19 and C-1, as shown in figure 3.1. From the above discussion, the sequence of residues in the macrocyclic ring was defined as

cyclo - [DCAMO—Hmba—thiazole-1-carboxylate—glyceric acid—thiazole-2-carboxylate], accounting for 10 of the 13 degrees of unsaturation.

Further analysis of the 2D NMR spectra led to the identification of two additional modified amino acids, an *N,N*-dimethyl-valine (*N,N*-DiMeVal) and an acetylated leucine statine. Key HMBC correlations between NH-27 and C-32 connected the *N,N*-DiMeVal and statine residues, though no correlations were observed between the carbonyl in the statine (C-24) and any of the residues in the macrocyclic ring. However, based on reasonable deduction and structure similarity to known analogs, it was rational to assume the presence of an ester linkage between C-24 and C-14 of the macrocycle to complete the planar structure.

**Table 3.1:** NMR spectral data for lyngbyabellin N (1) in  $d_6$ -DMSO

position	$\delta_C^b$	$\delta_H$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
1	173.2			
2	42.9	2.84, dd (9.4, 6.8)	1, 3, 9	3, 9
3	74.5	5.13, m		2, 4a, 4b
4a	30.0	1.80, m	3, 6	3
4b		1.69, m	3, 6	3
5	29.0	1.24, m	4	6
6a	48.3	2.27, m	7, 8	4a, 4b, 5
6b		2.20, m	7, 8	4a, 4b, 5
7	92.1			
8	37.0	2.09, s	6, 7	
9	14.6	1.16, d (7.1)	1, 2, 4	2
10	159.4			
11	145.4			
12	129.7	8.44, s	17, 19	
13	165.5			
14	70.2	6.27, t (6.1)	15, 24	15
15a	63.6	4.81, dd (11.2, 5.2)	14	14
15b		4.57, dd (11.7, 6.8)	14	14
16	160.2			
17	145.0			
18	130.2	8.46, s	17, 19	
19	167.2			
20	76.2	5.55, d (8.1)	1, 19, 21, 22, 23	21
21	31.9	2.24, m	20, 23	20, 22, 23
22	18.4	0.80, d (6.5)	20, 21, 23	21
23	17.9	1.00, d (6.5)	20, 21, 23	21
24	168.9			
25a	34.0	2.91, m	24	26

**Table 3.1: continued**

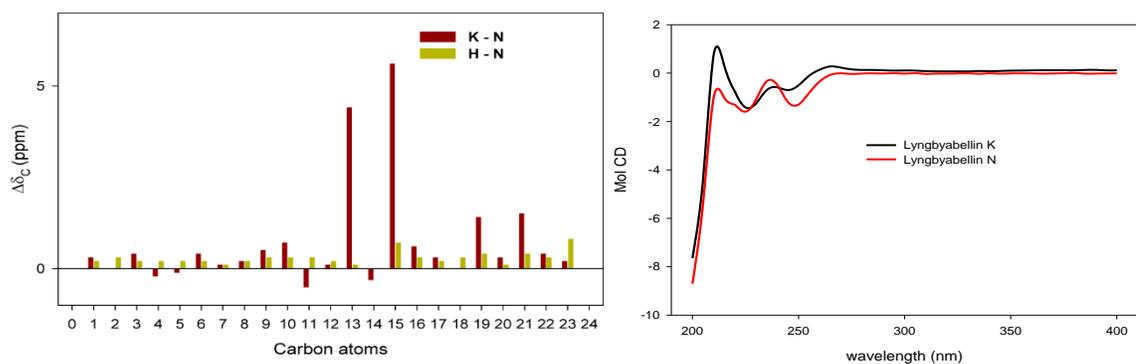
position	$\delta_C^b$	$\delta_H$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
25b		2.70, m	24, 26	21
26	71.6	5.12, m		25a, 25b, 27
27	47.9	4.35, t (4.4)		26, 28a, 28b, NH-1
28a	37.5	1.41, m	29, 31	29
28b		1.23, m	29, 31	29
29	24.4	1.55, m		28, 30, 31
30	24.0	0.90, d (6.4)	28, 29, 31	29
31	20.8	0.81, d (6.2)	28, 29, 31	29
NH-1		8.81, d (8.7)		27
32	165.8			
33	71.8	3.63, t (6.7)	32, 34, 35, 36	34
34	26.3	2.31, m	32, 33, 35, 36	33, 35, 36
35	19.5	1.07, d (6.8)	33, 34, 36	34
36	16.5	0.94, d (6.6)	33, 34, 36	34
37	41.5	2.72, d (4.0)	33, 38	
38	41.0	2.77, d (4.2)	33, 37	
NH-2		9.38, br s	39	
39	169.4			
40	20.8	1.91, s	39	

<sup>a</sup>500 MHz for <sup>1</sup>H NMR, HMBC, and COSY. <sup>b</sup>75 MHz for <sup>13</sup>C NMR.

### 3.2.3 Stereochemical Analysis of Lyngbyabellin N (**1**)

The planar structure of lyngbyabellin N (**1**) is closely related to that of lyngbyabellin H (**2**) except for the replacement of the polyketide portion with a *N,N*-DiMeVal residue. The comparison of optical rotations and carbon chemical shifts (macrocyclic lactone portion) strongly supported that the absolute configuration of the macrocyclic lactone in **1** and **2** were identical. Further support of this configuration was obtained by comparison of the CD absorption curve of **1**

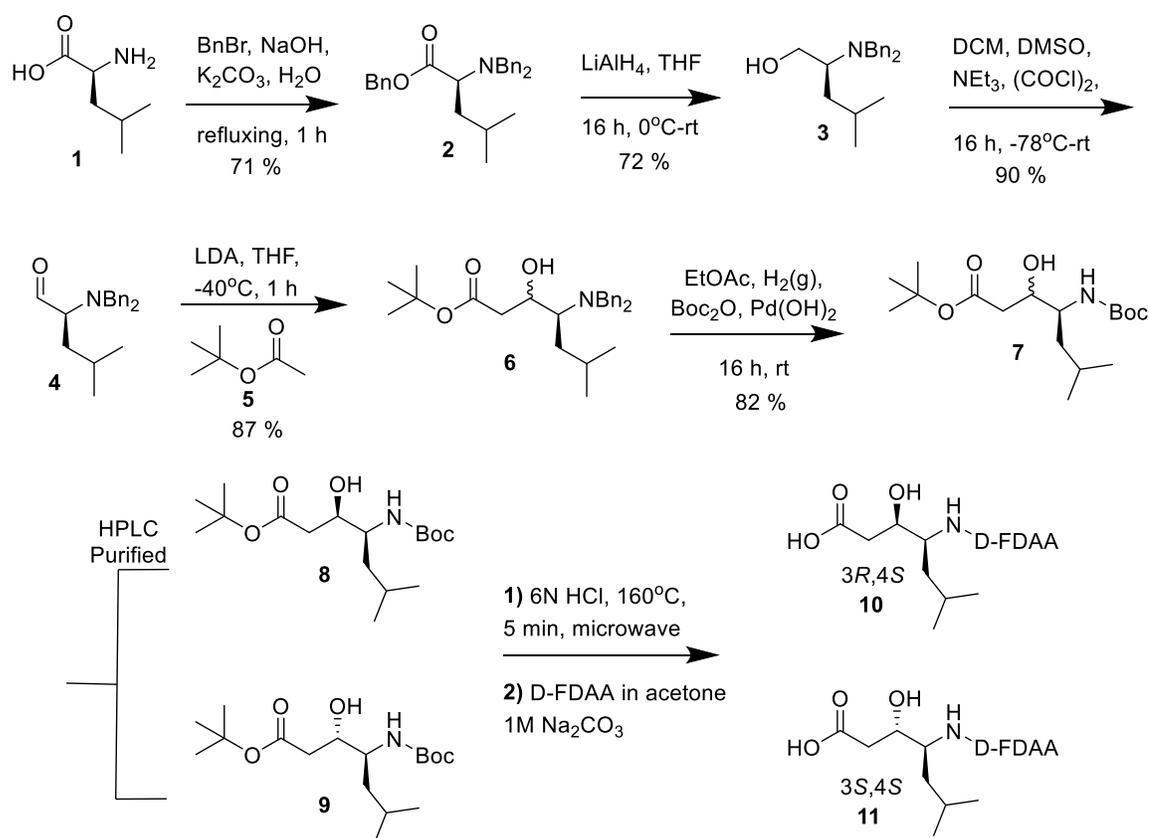
with that of lyngbyabellin K (**3**), which was isolated around the same time as **1** and was successfully crystallized.<sup>27</sup> Both of the CD curves were nearly identical, confirming that **1** has the 2*S*, 3*S*, 14*R*, 20*S* configuration (figure 3.3).



**Figure 3.3:** (left)  $^{13}\text{C}$  NMR comparison between lyngbyabellin N and both H and K; **Figure 3.4.** (right) Comparison of CD spectrums of lyngbyabellin K and N

The absolute configuration of the leucine statine in **1** was determined by LC-MS analysis of the acid hydrolysate appropriately derivatized with Marfey's reagent (D-FDAA). The four standards, 3*S*,4*S*-statine (Sta), 3*R*,4*S*-Sta, 3*S*,4*R*-Sta, and 3*R*,4*R*-Sta, were all synthesized beginning with their corresponding amino acids (L- and D-leucine, respectively), and following literature procedures, converted to their respective *N*-benzyl protected aldehydes (figure 3.3).<sup>28</sup> Each standard was then reacted with tert-butyl 2-bromoacetate in the presence of *N*-BuLi to yield a mixture of diastereomeric protected leucine statines. However, these diastereomers were inseparable by HPLC and were thus converted to Boc protected analogs, which were readily purified by RP HPLC.<sup>20</sup> Hydrolysis, followed by the derivatization with Marfey's reagent, yielded the four standards which each possessed a distinct retention time by LC-MS [(3*R*,4*R*)-Sta-D-FDAA (78.2 min), (3*S*,4*R*)-Sta-D-FDAA (80.9 min), (3*S*,4*S*)-Sta-D-FDAA (92.2 min), and (3*R*,4*S*)-Sta-D-FDAA (93.1 min)]. From the retention time of the natural product statine derivative (93.29 min), it was clear that this residue was of the 3*R*, 4*S* configuration.

The absolute configuration of the *N,N*-dimethylvaline (DiMeVal) residue in compound **1** was determined by comparing the chiral GC-MS retention time of the methylated residue liberated by acid hydrolysis with authentic standards. The two standards, *L-N,N*-DiMeVal and *D-N,N*-DiMeVal, were synthesized from *L*- and *D*-Val, respectively, following literature procedures.<sup>30</sup> The two standards each possessed distinctly different retention times by GC-MS [*L*-DiMeVal (63.7 min) and (*D*-DiMeVal (64.2 min)]. The methylated residue from the acid hydrolysate gave a single peak at 63.8 min, thus indicating its configuration as *L*, and was confirmed by co-injection with the standards. In summary, the above experiments established that lyngbyabellin N (**1**) consisted of *2S*, *3S*, *14R*, *20S*, *26R*, *27S*, and *33S*.

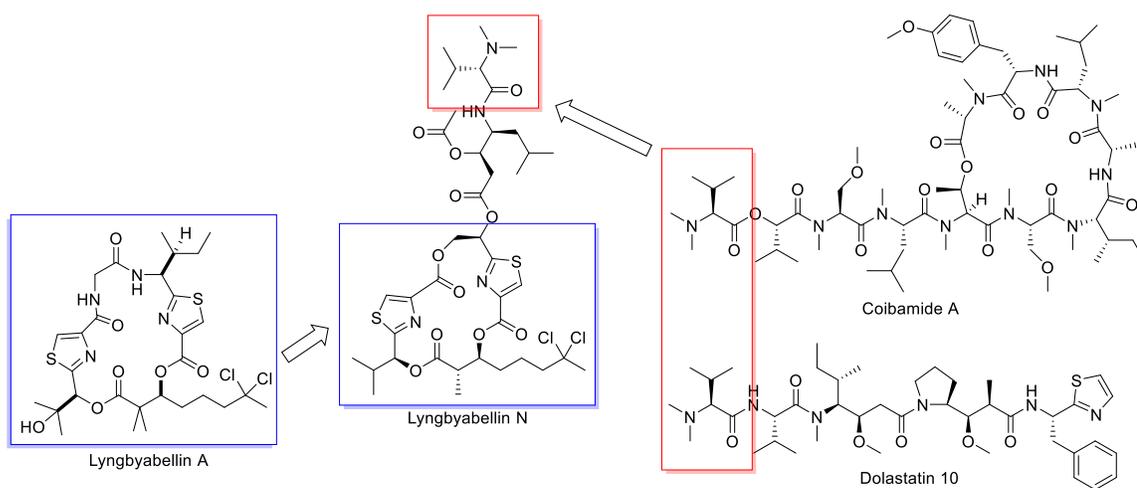


**Figure 3.5:** Synthetic scheme for the synthesis of the four statine residues

### 3.2.4 Bioassay Results

The lyngbyabellin family of compounds are known to exhibit moderate to potent cytotoxicity against a number of different cancer cell types through the promotion of actin polymerization.<sup>15</sup> Thus, after the completion of the structural analysis, compound **1** along with lyngbyabellin K-L (**3-5**) were evaluated in an H-460 human lung carcinoma cell cytotoxicity assay. Compound **1** showed strong yet variable cytotoxicity ( $IC_{50}$  0.0048-1.8  $\mu$ M, perhaps due to solubility problems), while compounds **3-6** were inactive. However, in the HCT-116 colon cancer cell line, reproducible  $IC_{50}$  values were obtained for lyngbyabellin N of  $40.9 \pm 3.3$  nM, confirming the potent cytotoxic effect, and suggesting that the side chain of lyngbyabellin N is an essential structural feature for this potent activity. However, this trend is not entirely consistent within this structure class as other lyngbyabellin analogs lacking the side chain exhibit sub-micromolar activity against HT29 and HeLa cells.<sup>22-24</sup>

It is interesting to note the increasing structural complexity in the lyngbyabellin family of metabolites, with that of lyngbyabellin N (**1**) being the most complex to date. While it has the recognizable core of the lyngbyabellins, the side chain and *N,N*-dimethylvaline terminus resembles that of the dolastatin 10 and coibamide A, two biologically potent families of metabolites, and in this regard, it has a hybrid structure between these cyanobacterial natural product classes (figure 3.5). Additionally, from a biosynthetic logic perspective, these more complex lyngbyabellins are perplexing as they possess two logical points for the initiation of molecule construction: the polyketide chain represents one such point and the *N,N*-dimethylvaline a second. Thus, these complex lyngbyabellins may indeed represent the hybridization and co-joining of two natural product structure classes.



**Figure 3.6:** Structure similarity of lyngbyabellin N to coibamide A, dolastatin 10, and lyngbyabellin A

### 3.3 Conclusion

A chemical investigation of anti-cancer active extracts of *Moorea bouillonii* collected in Palmyra Atoll led to the isolation of lyngbyabellin N (**1**). The planar structure and absolute configurations was determined by the combination of various techniques in spectroscopy, chromatography and synthetic chemistry. Compound **1** exhibited strong yet variable cytotoxicity ( $IC_{50}$  0.0048-1.8  $\mu$ M) against H-460 cells, while related compounds **3-6** were completely inactive. We do not understand the basis for this variable level of activity, however, it most likely relates to solubility issues in the assay buffer. However, in the HCT-116 colon cancer cell line, **1** was shown to be potent and exhibited an  $IC_{50}$  of  $40.9 \pm 3.3$  nM. This new lyngbyabellin metabolite possesses a unique structural feature where it has two conceptual points of biosynthetic chain initiation, which reflect unique metabolic reactions not yet characterized nor understood.

### 3.4 Experimental Methods

#### 3.4.1 General Experimental Procedures

Optical rotations were measured on a JASCO P-2000 polarimeter, CD spectra were taken in EtOH using a JASCO J-810 spectropolarimeter. UV spectra and IR spectra were recorded on a Beckman Coulter DU800 spectrophotometer and a Nicolet ThermoElectron Nicolet IR100 FT-IR spectrometer using KBr plates, respectively. NMR spectra were recorded with DMSO as internal standards ( $\delta_C$  39.5,  $\delta_H$  2.50) on a Varian 500 MHz spectrometer (500 and 125 MHz for  $^1H$  and  $^{13}C$  NMR, respectively) and Varian 300 MHz spectrometer (300 and 75 MHz for  $^1H$  and  $^{13}C$  NMR, respectively). HR ESIMS spectra were obtained on an Agilent 6230 ESI-TOF mass spectrometer. HPLC was carried out using Waters 515 pumps system with a Waters 996 PDA detector. Acid hydrolysis was performed using a Biotage (Initiator) microwave reactor equipped with high-pressure vessels.

#### 3.4.2 Cyanobacterial Collection and Taxonomic Identification

Lyngbyabellin N producing cyanobacterium (collection code: PAL 8/16/08-3) was collected by SCUBA on reefs 9-15 m deep around Palmyra Atoll, USA. The environmental samples were stored in EtOH/H<sub>2</sub>O (1:1) at -20 °C, while the genetic materials were preserved in RNA stabilization solution at -20 °C (RNAlater, Ambion Inc.). Morphological characterization was performed using an Olympus IX51 epifluorescent microscope (100×) equipped with an Olympus U-CMAD3 camera. Taxonomic identification of cyanobacterial specimens was performed in accordance with current phycological systems.<sup>31-32</sup>

#### 3.4.3 Isolation of Lyngbyabellin N (**1**)

One liter of cyanobacterial tissue (previously identified as *Moorea bouillonii*)<sup>3</sup> was repetitively extracted with 2:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH to afford 4.2 g of crude extract. The extract was

fractionated by silica gel VLC with a stepwise gradient solvent system of increasing polarity starting from 10% EtOAc in hexanes to 100% MeOH, to produce nine fractions (A–I). The fraction eluting with 75% EtOAc in MeOH (fraction H) was subsequently separated using a 5 g RP SPE with a stepwise gradient solvent system decreasing in polarity starting from 55% CH<sub>3</sub>CN in H<sub>2</sub>O to 100% CH<sub>2</sub>Cl<sub>2</sub>, to produce five fractions (1-5). The fraction eluting with 70% CH<sub>3</sub>CN in H<sub>2</sub>O (fraction 3) was further separated using prepTLC, with an isocratic solvent system of 100% EtOAc, to yield pure lyngbyabellin N (**1**, 5.4 mg, 0.12 %).

**Lyngbyabellin N (1):** pale yellow oil;  $[\alpha]_D^{27}$ -24.0 (*c* 1.05, MeOH); UV (MeOH)  $\lambda_{\max}$  202 nm (log  $\epsilon$  4.36), 235 (log  $\epsilon$  3.99); CD (MeOH)  $\lambda_{\max}$  ( $\Delta\epsilon$ ), 212 nm (-0.64), 224 (-1.60), 237 (-0.27), 248 (-1.36), 268 (+0.03); IR (KBr)  $\gamma_{\max}$  3436, 2961, 2933, 1742, 1677, 1467, 1372, 1321, 1233, 1166, 1097, 1037cm<sup>-1</sup>; <sup>1</sup>H, <sup>13</sup>C and 2D NMR data, see Table 2; HRESIMS *m/z* [M+H]<sup>+</sup> 905.2997 (calcd for C<sub>40</sub>H<sub>59</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>11</sub>S<sub>2</sub> 905.2993,  $\Delta$  +0.4 mmu).

#### 3.4.4 Ozonolysis, and Acid Hydrolysis of Lyngbyabellin N (**1**)

A portion (1 mg) of **1** was dissolved in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> at – 78 °C and O<sub>3</sub> was bubbled through the sample for 10 min. The pale blue solution was dried under N<sub>2</sub> (g). The products were re-suspended in 500  $\mu$ L of 6 N HCl and reacted at 160 °C for 5 min in a microwave reactor.

#### 3.4.5 Modified Marfey's Analysis of the Statine Unit

An aliquot (~500  $\mu$ g) of the acid hydrolysate was dried under N<sub>2</sub> (g), and dissolve in 1 mL of 1 M sodium bicarbonate, and 12  $\mu$ L of 1% D-FDAA (1-fluoro-2,4-dinitrophenyl-5-D-alanine amide) was added in acetone. The solution was maintained at 40 °C for 90 min, at which time the reaction was quenched by the addition of CH<sub>3</sub>CN, and 10  $\mu$ L of the solution was analyzed by LC-ESIMS. The Marfey's derivatives of the hydrolysate and standards were analyzed by RP HPLC using Phenomenex Luna 5  $\mu$  C<sub>18</sub> column (4.6 x 250 mm). The HPLC

conditions began with 10% CH<sub>3</sub>CN/H<sub>2</sub>O acidified with 0.1% formic acid (FA) followed by a gradient profile to 50% CH<sub>3</sub>CN/H<sub>2</sub>O acidified with 0.1% FA over 85 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. The retention times of the authentic acid D-FDAA derivatives were (3*R*,4*R*)-Sta-D-FDAA (78.2 min), (3*S*,4*R*)-Sta-D-FDAA (80.9 min), (3*S*,4*S*)-Sta-D-FDAA (92.2 min), and (3*R*,4*S*)-Sta-D-FDAA (93.1 min); the hydrolysate product gave a peak with retention time of 93.3 min, indicating an absolute configuration of 3*R*,4*S*.

Correspondingly, (3*S*,4*S*)-statine, (3*S*,4*R*)-statine, (3*R*,4*R*)-statine and (3*R*,4*S*)-statine were synthesized from L-leucine (1 g) and D-leucine (1 g), respectively, following Reetz *et al.* and to yield the benzyl protected aldehyde.<sup>28</sup> An aliquot (3.38 mmol) of the benzyl protected aldehydes dissolved in 15 mL were then added to a solution of with 3.72 mmol *tert*-butyl acetate and freshly prepared lithium diisopropylamine (5.06 mmol). The reaction was warmed to – 40 °C for 1 h and then quenched with 85 mL NaHCO<sub>3</sub> (*aq*), filtered and the aqueous layer was separated and washed 2x with 20 mL of Et<sub>2</sub>O. The organic layer was then washed with brine and dried with NaSO<sub>4</sub> and filtered to yield diastereomeric *tert*-butyl 4-(dibenzylamino)-3-hydroxy-6-methylheptanoate. The mixture of diastereomers were inseparable, thus the method outlined by Andrés *et al.* was used, where the benzyl protected amine was converted to the Boc protection allowing for purification of a small amount of each of the diastereomers.<sup>29</sup> Each of the diastereomers were identified by the comparison of <sup>1</sup>H NMR each of the pure standards to the known compounds in the above paper. An aliquot (2 mg) of each protected statins were treated with 1 mL of 6 N HCl and heated to 160 °C for 5 min in a microwave reactor. Each of the hydrolysate products were dried down by N<sub>2</sub> (g) and then reacted with the Marfey's reagent as mentioned above, to yield the four Marfey's derived statine standards.

#### 3.4.6 Preparation and GCMS Analysis of 2-Hydroxy-3-methylbutyric Acid (Hmba)

An aliquot (~500  $\mu\text{g}$ ) of the hydrolysate was treated with excessive  $\text{CH}_2\text{N}_2$  for 30 min at room temperature, and dried under  $\text{N}_2$  (g). Correspondingly, L-Hmba and D-Hmba were synthesized as described above. The products were dissolved in DCM and injected over chiral-phase GC-MS using a Chirasil-Val column (Agilent Technologies J&W Scientific, 30 m x 0.25 mm) under the following conditions: the initial oven temperature was 32  $^\circ\text{C}$ , held for 15 min, followed by a ramp from 32  $^\circ\text{C}$  to 60  $^\circ\text{C}$  at a rate of 10  $^\circ\text{C}/\text{min}$ , followed by another ramp to 200  $^\circ\text{C}$ , at a rate of 15  $^\circ\text{C}/\text{min}$  and held at 200  $^\circ\text{C}$  for 5 min. The retention time of products resulting from the acid hydrolysate of **1** matched the synthetic 2*S*-Hmba standard (9.7 min; 2*R*-Hmba, 10.2 min).

#### 3.4.7 Preparation and GCMS Analysis of *N,N*-Dimethylvaline

The methylated hydrolysate product of **1** was analyzed by Chiral GC-MS using a Cyclosil B column (Agilent Technologies J&W Scientific, 30 m x 0.25 mm) under the following conditions: the initial oven temperature was 34  $^\circ\text{C}$  and was held for 68 min, followed by a ramp from 34  $^\circ\text{C}$  to 100  $^\circ\text{C}$  at a rate of 30  $^\circ\text{C}/\text{min}$  and held at 100  $^\circ\text{C}$  for 5 min. Synthetic standards of 2*S-N,N*-dimethylvaline and 2*R-N,N*-dimethylvaline were first methylated by dissolving 10 mg of each starting material into 433  $\mu\text{L}$  of  $\text{H}_2\text{O}$ , followed by the addition of 27  $\mu\text{L}$  of formaldehyde and 10.4 mg of 10% Pd/C. The system was then treated with  $\text{H}_2$  (g) for 16 h. After 16 h the reactions were brought to a boil and then concentrated via rotorvap. Each of the synthetic standards were then treated with  $\text{CH}_2\text{N}_2$  for 5 min and then dried down under  $\text{N}_2$  (g), re-suspended in  $\text{CH}_2\text{Cl}_2$ , then analyzed with chiral GC-MS. The retention time of products resulting from the acid hydrolysate of **1** matched to that of the authentic 2*S-N,N*-dimethylvaline standard (63.7 min; 2*R-N,N*-dimethylvaline, 64.2 min).

#### 3.4.8 Cytotoxicity Assay

H-460 cells were added to 96-well plates at  $3.33 \times 10^4$  cells/mL of Roswell Park Memorial Institute (RPMI) 1640 medium with 10% fetal bovine serum (FBS) and 1% Penicillin/Streptomycin. The cells, in a volume of 180  $\mu$ L per well, were incubated overnight (37°C, 5% CO<sub>2</sub>) to allow recovery before treatment with test compounds. Compounds were dissolved in DMSO to a stock concentration of 10 mg/mL. Working solutions of the compounds were made in RPMI 1640 medium without FBS, with a volume of 20  $\mu$ L added to each well to give a final compound concentration of either 30 or 3  $\mu$ g/mL. An equal volume of RPMI 1640 medium without FBS was added to wells designated as negative controls for each plate. Plates were incubated for approximately 48 h before staining with MTT. Using a ThermoElectron Multiskan Ascent plate reader, plates were read at 570 and 630 nm. Concentration response graphs were generated using GraphPad Prism (GraphPad Software Inc., San Diego, CA).

### 3.5 Acknowledgements

I thank A.C. Jones, A.R. Pereira, and R.C. Coates for assistance with the collection of the lyngbyabellin-producing cyanobacterial samples from the Palmyra Atoll, as well as the UCSD mass spectrometry facilities for their analytical services, and H. Choi, T. Byrum, F. A. Valeriote, and W. H. Gerwick for their scientific contribution to this work. We further acknowledge support of this research from the UC San Diego and the NIH (CA092143). We also acknowledge The Growth Regulation & Oncogenesis Training Grant NIH/NCI (T32A009523-24).

## 3.6 Appendix

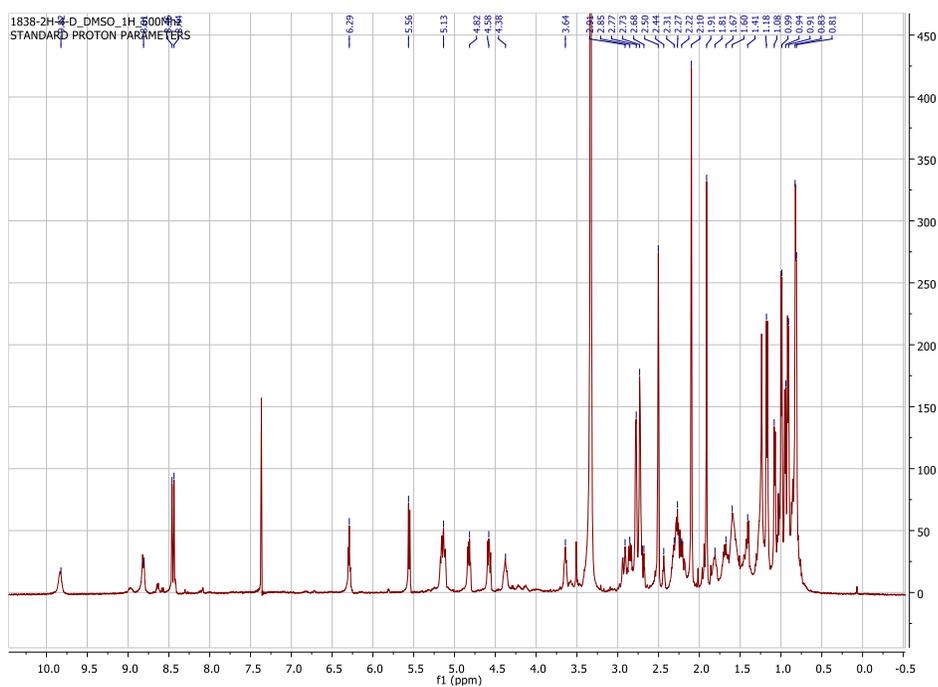


Figure 3.6.1:  $^1\text{H}$  NMR (500 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N

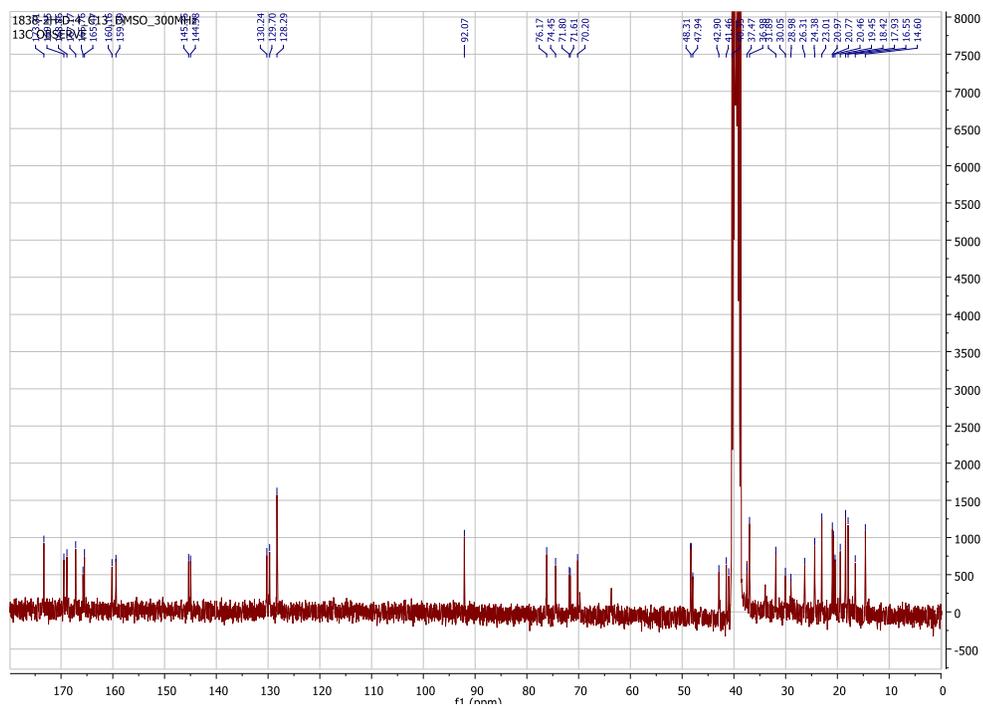
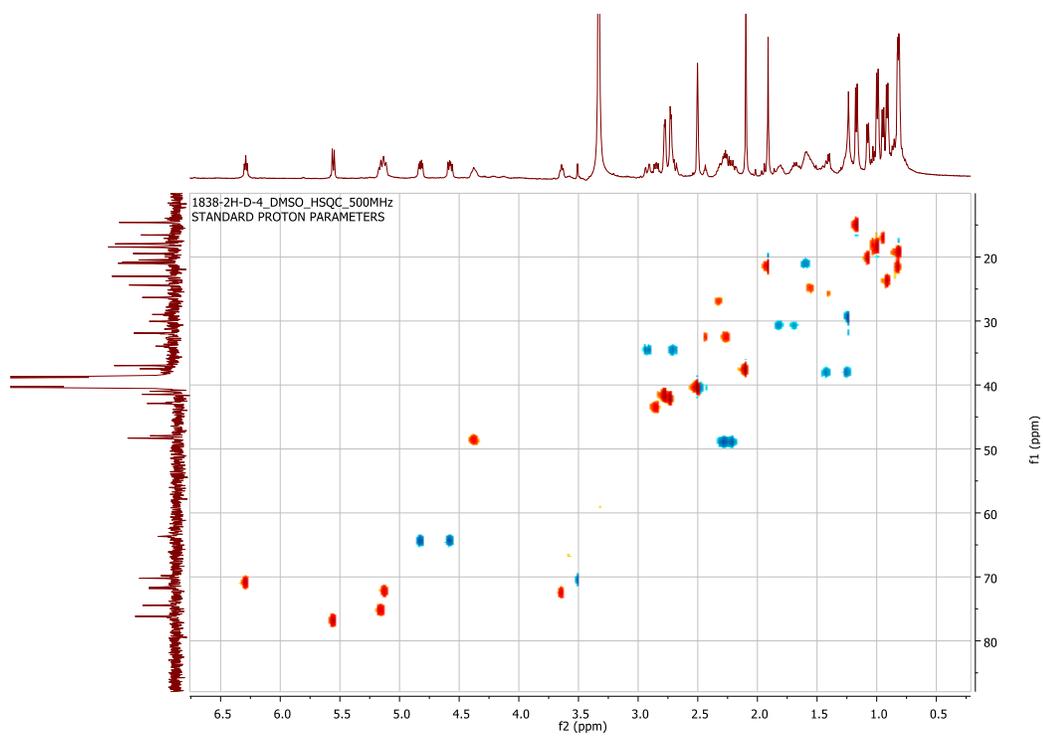
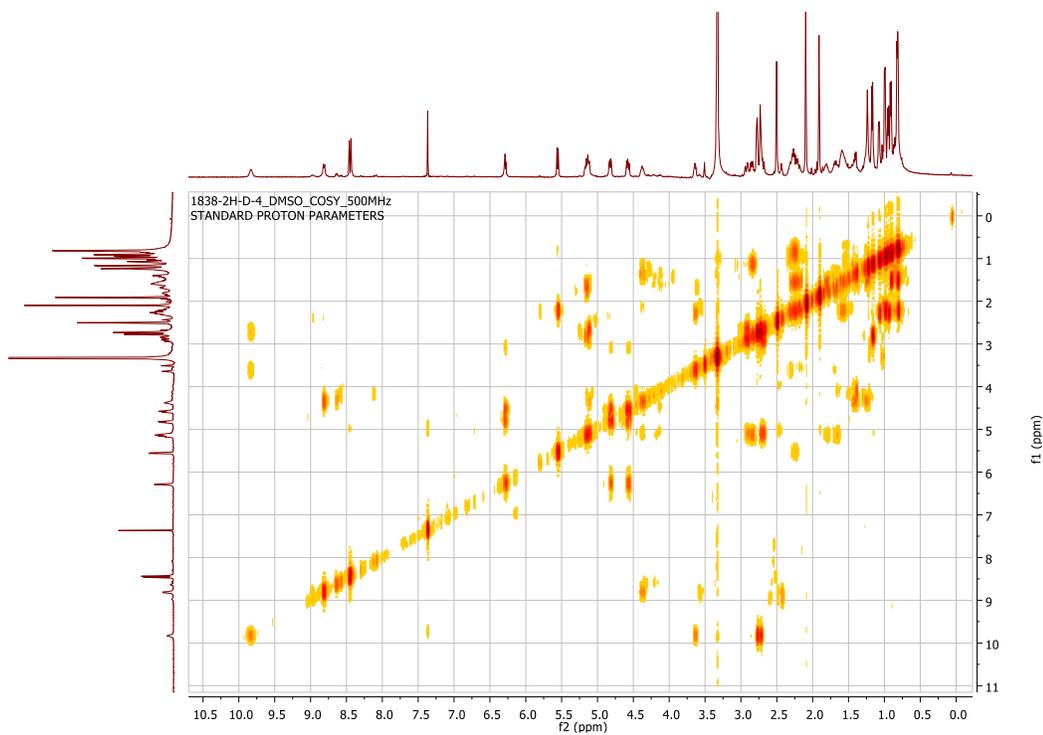


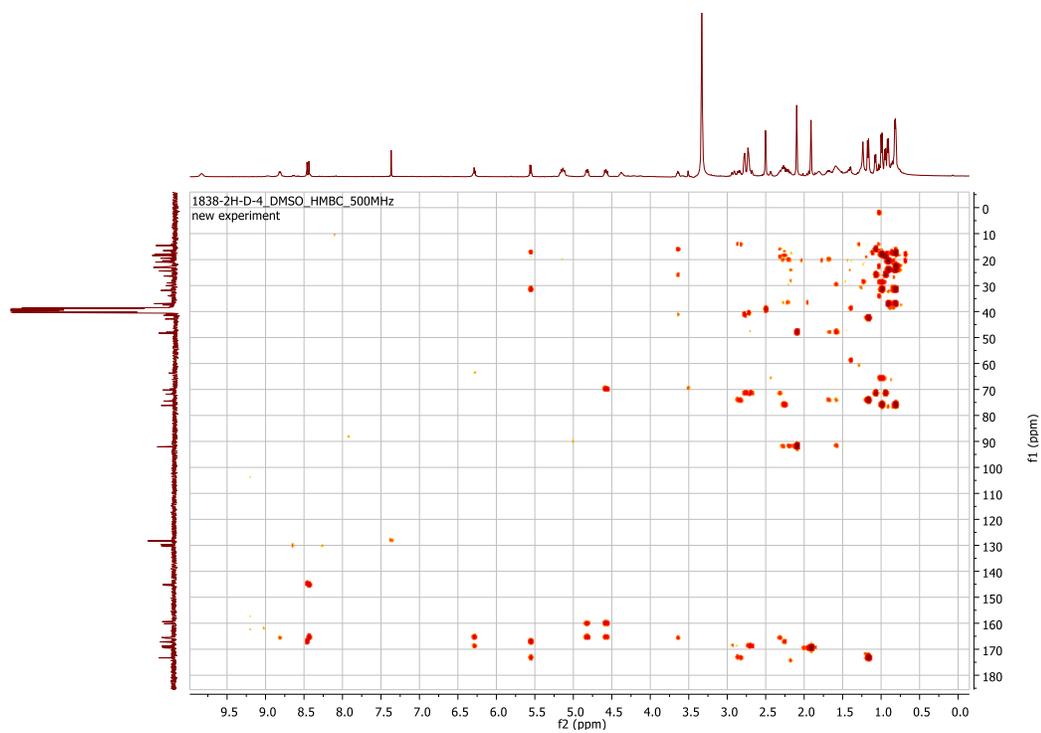
Figure 3.6.2:  $^{13}\text{C}$  NMR (75 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N



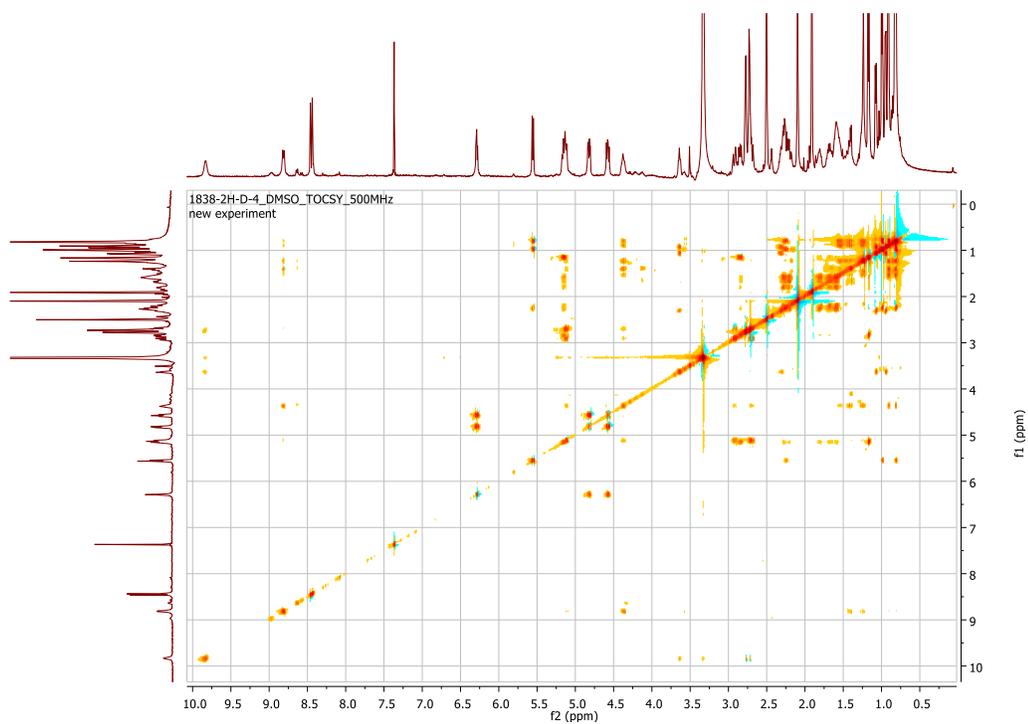
**Figure 3.6.3:** gHSQC ( $^1\text{H}$  500 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N



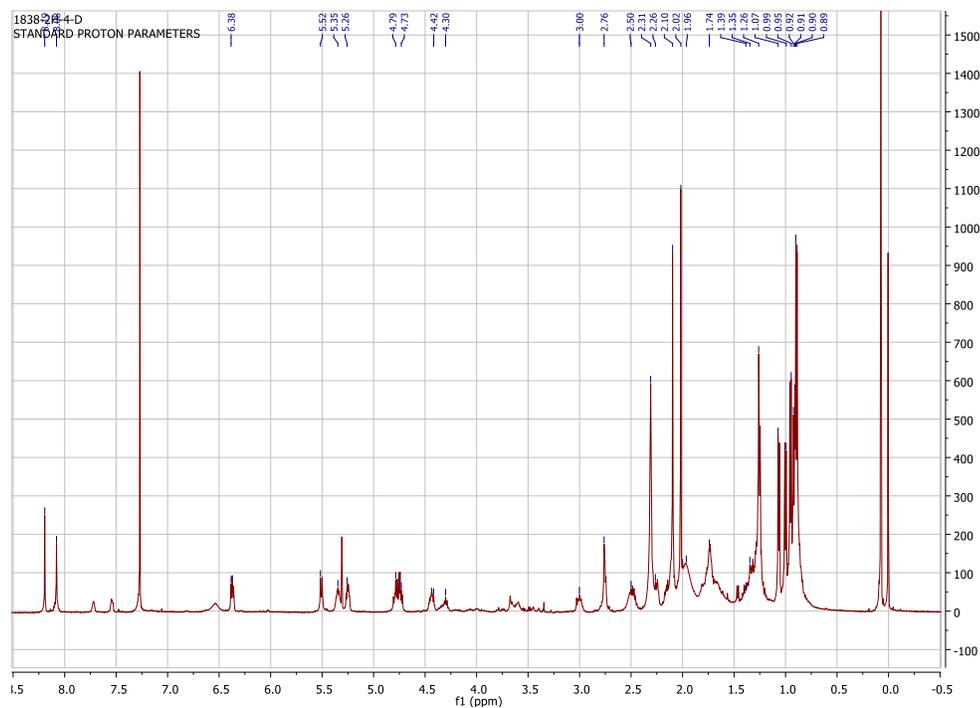
**Figure 3.6.4:** COSY ( $^1\text{H}$  500 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N



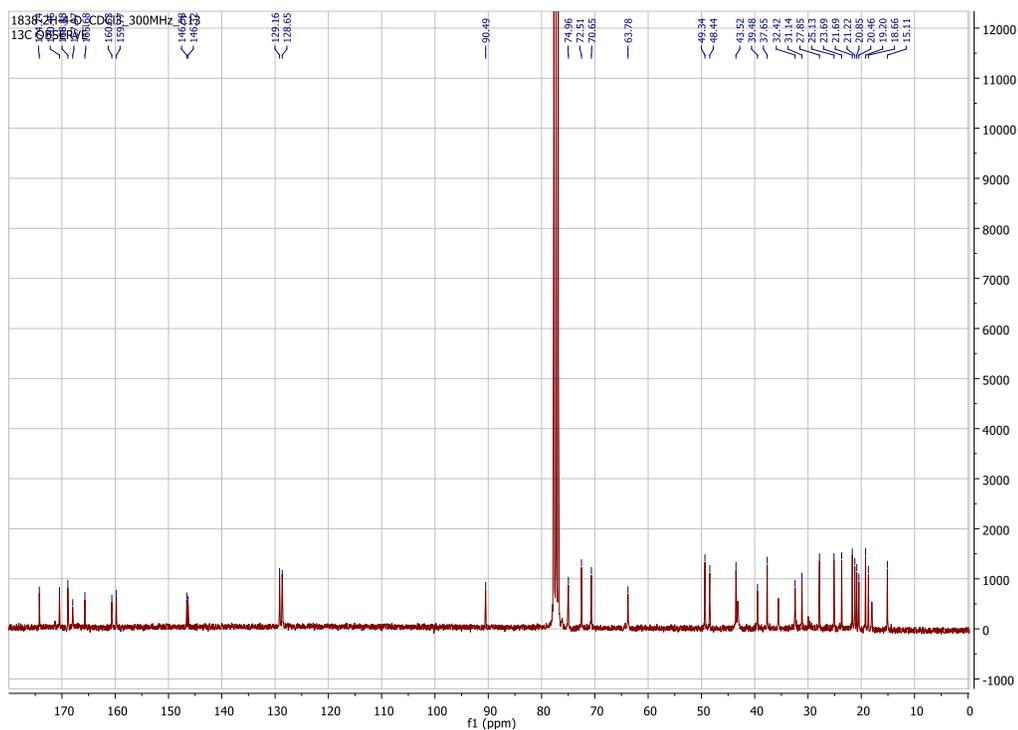
**Figure 3.6.5:** HMBC ( $^1\text{H}$  500 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N



**Figure 3.6.6:** TOCSY ( $^1\text{H}$  500 MHz,  $d_6$ -DMSO) spectrum of lyngbyabellin N



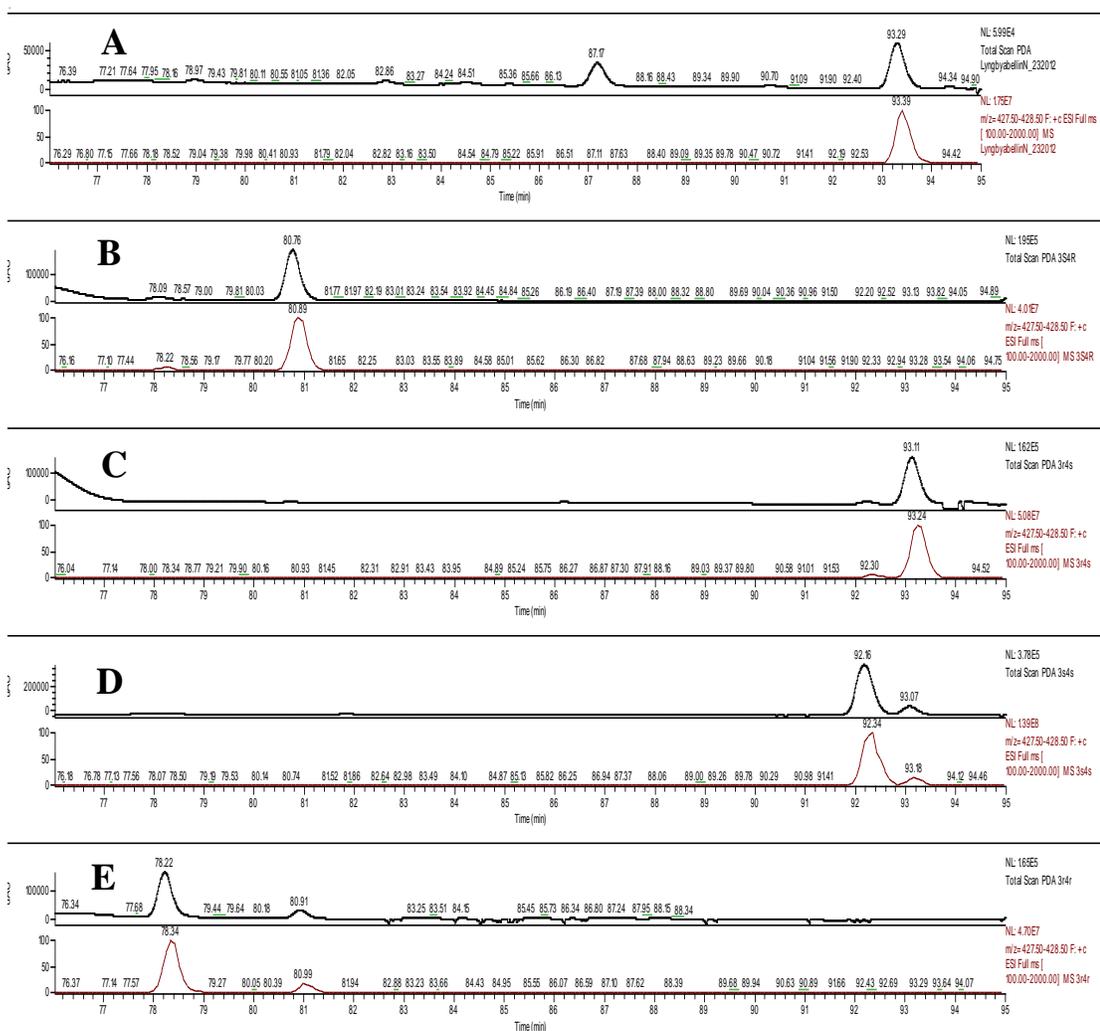
**Figure 3.6.7:**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of lyngbyabellin N



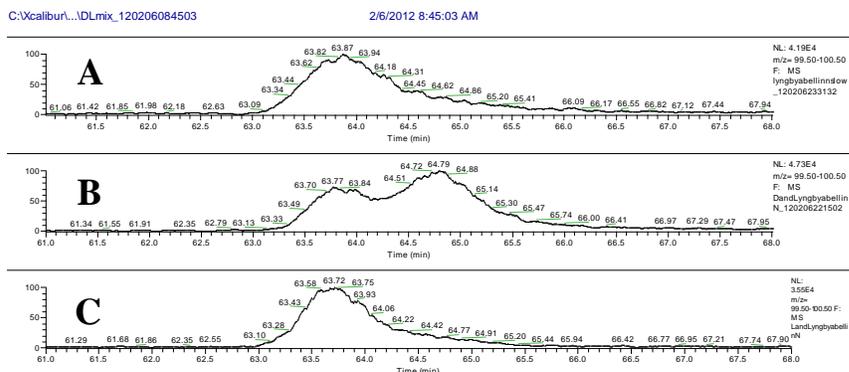
**Figure 3.6.8:**  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ) spectrum of lyngbyabellin N

d:\data\mevrs3r4r

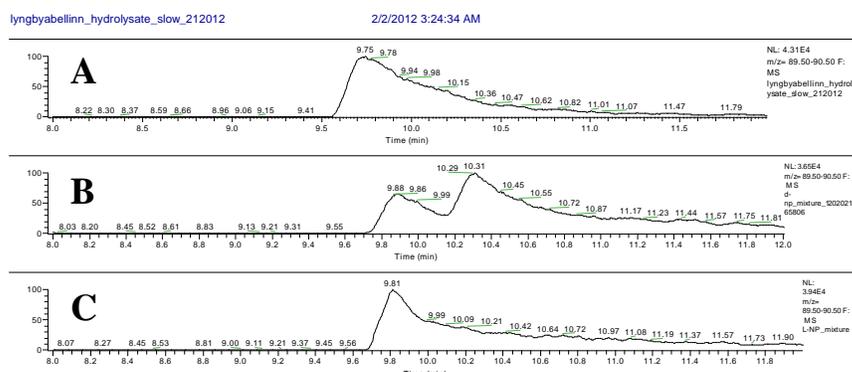
2/3/2012 9:46:01 PM



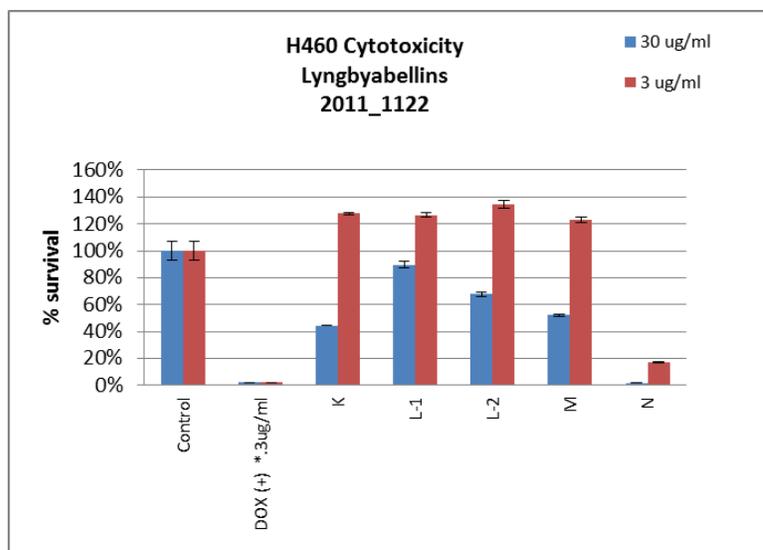
**Figure 3.6.9:** Analysis of the leucine statine residue via LCMS: (A) Natural product hydrolysis product derivatized with D-FDAA (Marfey's Reagent); (B) 3*S*,4*R*-leucine statine standard; (C) 3*R*,4*S*-leucine statine standard; (D) 3*S*,4*S*-leucine statine standard; (E) 3*R*,4*R*-leucine statine standard



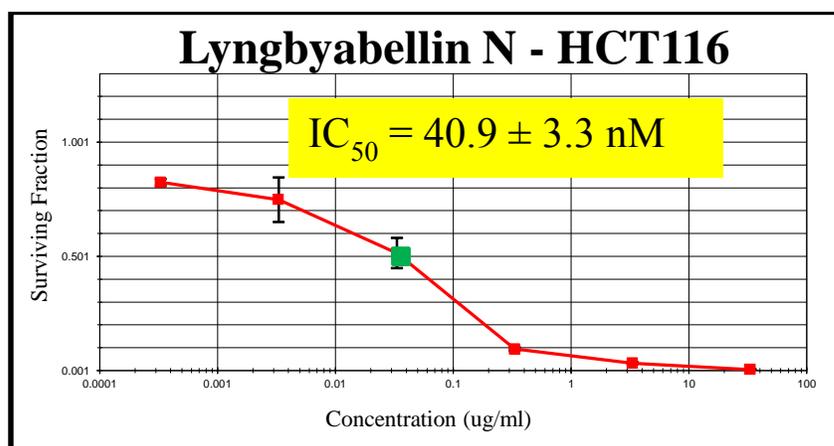
**Figure 3.6.10:** Analysis of *N,N*-dimethylvaline via chiral GCMS: (A) Natural product hydrolysis product derivatized with diazomethane; (B) Mixture of *R-N,N*-dimethylvaline standard and derivatized hydrolysis product; (C) Mixture of *S-N,N*-dimethylvaline standard and derivatized hydrolysis product



**Figure 3.6.11:** Analysis of the Hmba residue via chiral-phase GCMS: (A) Natural product hydrolysis product derivatized with diazomethane; (B) Mixture of *R*-Hmba standard and derivatized hydrolysis product; (C) Mixture of *S*-Hmba standard and derivatized hydrolysis product



**Figure 3.6.12:** H-460 cytotoxicity assay results



**Figure 3.6.13:** HCT-116 dose response

## 3.7 References

- (1) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. *Curr. Opin. Biotechnol.* **2010**, *21*, 787–793.
- (2) Tan, L. T. *J. Appl. Phycol.* **2010**, *22*, 659–676.
- (3) Gerwick, W. H.; Coates, R. C.; Engene, N.; Gerwick, L.; Grindberg, R. V.; Jones, A. C.; Sorrels, C. M. *Microbe* **2008**, *3*, 277–284
- (4) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. in *Comprehensive Natural Products Chemistry II* (Eds.: B. Moore, P. Crews), Elsevier, Oxford, **2010**, vol. 8, pp. 141–188.
- (5) Nagle, D. G.; Paul, V. J. *J. Phycol.* **1999**, *35*, 1412–1421.
- (5) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418–5423
- (6) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *Bioorg. Med. Chem.* **2002**, *10*, 1973–1978
- (7) Gutiérrez, M.; Suyama, T. L.; Engene, N.; Wingerd, J. S.; Matainaho, T.; Gerwick, W. H. *J. Nat. Prod.* **2008**, *71*, 1099–1103
- (8) Matthew, S.; Schupp, P. J.; Luesch, H. *J. Nat. Prod.* **2008**, *71*, 1113–1116
- (9) Tidgewell, K.; Engene, N.; Byrum, T.; Media, J.; Takayuki, D.; Valeriote, F. A.; Gerwick, W. H. *ChemBioChem* **2010**, *11*, 1458–1466.
- (10) Orjala, J.; Nagle, D. G.; Hsu, V. L.; Gerwick, W. H. *J. Am. Chem. Soc.* **1995**, *117*, 8281–8282.
- (11) Cardellina, J. H.; Marner, F. J.; Moore, R. E. *Science* **1979**, *204*, 193–195.
- (12) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. *J. Org. Chem.* **1994**, *59*, 1243–1245.
- (13) Orjala, J.; Gerwick, W. H. *J. Nat. Prod.* **1996**, *59*, 427–430.
- (14) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817–833.
- (15) Gerwick, W. H.; Tan, L. T.; Sitachitta, N. in *The Alkaloids* (Ed.: G. A. Cordell), Academic Press, San Diego, **2001**, vol. 57, pp. 75–184.
- (16) Simmons, T. L.; Andrianasolo, E.; McPhail, K.; Flatt, P.; Gerwick, W. H. *Mol. Cancer Ther.* **2005**, *4*, 333–342.
- (17) Matthew, S.; Ratnayake, R.; Becerro, M. A.; Ritson-Williams, R.; Paul, V. J.; Luesch, H. *Mar. Drugs* **2010**, *8*, 1803–1816.

- (18) Pereira, A.; Etzbach, L.; Engene, N.; Müller, R.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 1175–1181.
- (19) Choi, H.; Mascuch, S. J.; Villa, F. A.; Byrum, T.; Teasdale, M. E.; Smith, J. E.; Preskitt, L. B.; Rowley, D. C.; Gerwick, L.; Gerwick, W. H. *Chem. Biol.* **2012**, *19*, 589–598.
- (20) Grindberg, R. V.; Shuman, C. F.; Sorrels, C. M.; Wingerd, J.; Gerwick, W. H. in *Modern Alkaloids: Structure, Isolation, Synthesis and Biology* (Eds.: E. Fattorusso, O. Taglialetela-Scafati), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, **2008**, vol. 1, pp. 139–170.
- (21) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Mooberry, S. L. *J. Nat. Prod.* **2000**, *63*, 611–615
- (22) Matthew, S.; Salvador, L. A.; Schupp, P. J.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2010**, *73*, 1544–1552
- (23) Han, B.; McPhail, K. L.; Gross, H.; Goeger, D. E.; Mooberry, S. L.; Gerwick, W. H. *Tetrahedron* **2005**, *61*, 11723–11729.
- (24) Sone, H.; Kondo, T.; Kiryu, M.; Ishiwata, H.; Ojika, M.; Yamada, K. *J. Org. Chem.* **1995**, *60*, 4774–4781.
- (25) Marquez, B. L.; Watts, K. S.; Yokochi, A.; Roberts, M. A.; Verdier-Pinard, P.; Jimenex, J. I.; Hamel, E.; Scheuer, P. J.; Gerwick, W. H. *J. Nat. Prod.* **2002**, *65*, 866–871.
- (26) Choi, H.; Mevers, E.; Byrum, T.; Valeriote, F. A.; Gerwick, W. H. *Eur. J. Org. Chem.* **2012**, 5141–5150.
- (27) Reetz, M. T.; Drewes, M. W.; Schwickardi, R. *Org. Synth.* **1999**, *76*, 110.
- (28) Andrés, J. M.; Pedrosa, R.; Pérez, A.; Pérez-Encabo, A. *Tetrahedron* **2001**, *57*, 8521–8530.
- (29) Bowman, R. E.; Stroud, H. H. *J. Chem. Soc.* **1950**, 1322–1345.
- (30) Castenholz, R. W.; Rippka, R.; Herdman, M. in *Manual of Systematic Bacteriology* (Eds.: Boone, D. R.; Castenholz, R. W.), Springer, New York, **2001**, vol. 1, pp. 473–599.
- (31) Komárek, J.; Anagnostidis K. in *Süßwasserflora von Mitteleuropa* (Eds.: Büdel, B.; Gärtner, G.; Krienitz, L.; Schagerl, M.), Gustav Fischer, Jena, **2005**, vol. 19/2, pp. 576–606.

## Chapter 4:

### Lipopeptides from the Tropical Marine Cyanobacterium *Symploca* sp.

#### 4.0.1 Abstract

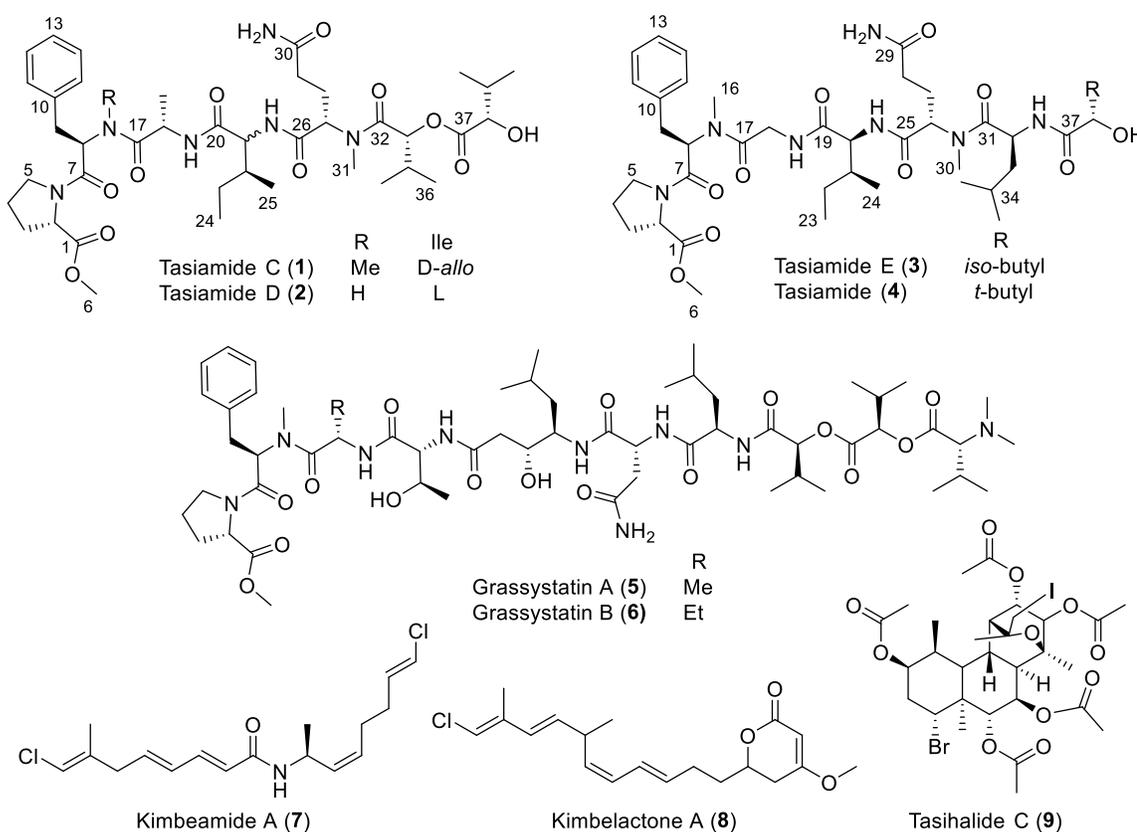
A collection of the tropical marine cyanobacterium *Symploca* sp., collected near Kimbe Bay, Papua New Guinea, previously yielded several new metabolites including kimbeamides A-C, kimbelactone A and tasihalide C. Investigations into a more polar cytotoxic fraction yielded three new lipopeptides, tasiamides C-E (**1-3**). The planar structures were deduced by 2D NMR spectroscopy and tandem mass spectrometry, and their absolute configurations were determined by a combination of Marfey's and chiral-phase GC-MS analysis. These new metabolites are similar to several previously isolated compounds, including tasiamide (**4**), grassystatins (**5-6**), and symplocin A, all of which were isolated from similar marine filamentous cyanobacteria.

#### 4.1 Introduction

Recent sequencing efforts of marine cyanobacterial genomes have revealed their exceptional capacity to produce a large diversity of intriguing secondary metabolites, representing a range of distinct and unrelated biosynthetic pathways.<sup>1,2</sup> This has also been observed from chemical investigations of such cyanobacterial species as *Moorea bouillonii* and *Moorea producens*, formerly known as *Lyngbya bouillonii* and *Lyngbya majuscula*, respectively.<sup>3</sup> Both of these latter organisms are known to produce a plethora of metabolites, including the lyngbyabellins,<sup>4-7</sup> apratoxins,<sup>8-11</sup> laingolide,<sup>12</sup> lyngbyaloxide,<sup>13-14</sup> apramide A<sup>15</sup> and palau'imide<sup>16</sup> from *M. bouillonii* and lyngbyatoxin A,<sup>17</sup> the jamaicamides,<sup>18</sup> carnabin,<sup>19</sup> various malyngamides,<sup>20</sup> and barbamide<sup>21</sup> from *M. producens*. These examples of highly productive strains and improved genomic insights into cyanobacterial biosynthetic capacity demonstrates the utility of rigorously examining

collections of cyanobacteria for novel natural products, even when a given strain has already been extensively investigated.

In the present investigation into a filamentous tuft-forming cyanobacterium from Kimbe Bay, Papua New Guinea, which already yielded kimbeamides A-C, kimbealactone A, and tasihalide C,<sup>22,23</sup> another chromatography fraction exhibited strong cytotoxicity against several cancer cell lines and was thus chosen for further evaluation. A subsequent NMR-guided fractionation process yielded three new lipopeptides, tasiamides C-E (**1-3**), two (C, D) of which were evaluated for cytotoxicity and found to be inactive. Both the planar and absolute configurations of these metabolites were determined, and have led to some intriguing insights into the biosynthetic capability of this particular collection.



**Figure 4.1:** Tasiamides C-E and grassystatin A-B

## 4.2 Results and Discussion

### 4.2.1 Tasiamides C-E Collection and Isolation

Orange tufts of a tropical marine *Symploca* sp. were collected in approximately 20 m of water near Kimbe Bay, Papua New Guinea, in July 2007. The preserved collection was repetitively extracted (2:1 CH<sub>2</sub>Cl<sub>2</sub>—MeOH) and fractionated using normal-phase vacuum liquid chromatography (VLC). Previously, a middle polarity fraction (60% hexanes/40% EtOAc) of this extract yielded several biologically active metabolites, including kimbeamides A-C, kimbelactone A and tasihalide C.<sup>22-23</sup> Additionally, a relatively polar fraction eluting with 25% MeOH/EtOAc exhibited cytotoxic activity against H-460 human lung cancer cells (81% toxicity at 3 µg/mL). Further chromatography of this fraction using normal-phase solid phase extraction (SPE) and reversed-phase HPLC yielded 1.9 mg of tasiamide C (**1**), 2.5 mg of tasiamide D (**2**), and 0.7 mg of tasiamide E (**3**).

### 4.2.2 Tasiamide C Absolute Structure Determination

HR-ESITOFMS of **1** yielded an [M+Na]<sup>+</sup> at *m/z* 839.4541, giving a molecular formula C<sub>41</sub>H<sub>64</sub>N<sub>6</sub>O<sub>11</sub> (calcd for C<sub>41</sub>H<sub>64</sub>N<sub>6</sub>NaO<sub>11</sub>, 839.4525, 1.9 ppm), with 13 degrees of unsaturation. The IR spectrum featured absorptions indicative of the presence of NH or OH protons and the presence of amide or ester carbonyls (3371 and 1737 cm<sup>-1</sup>, respectively). Further evidence of ester or amide carbonyls were present as eight downfield signals in the <sup>13</sup>C NMR spectrum (δ<sub>c</sub> 167.8, 169.2, 169.9, 170.8, 172.1, 172.5, 173.7, and 175.7). Additionally, the presence of a mono-substituted phenyl group was evident from four downfield carbon signals, two of which were composed of two carbons each as indicated by their relative peak height (δ<sub>c</sub> 126.4, 128.0 x 2, 129.2 x 2, and 136.4). Further analysis of the <sup>1</sup>H NMR spectrum revealed the presence of three singlet methyls at shifts indicative of an *O*-methyl (δ<sub>H</sub> 3.67) and two *N*-methyl groups (δ<sub>H</sub> 2.93 and 3.03), along with two broad downfield signals suggestive of NH protons (δ<sub>H</sub> 6.69 and 6.78).

Analysis of the 2D NMR data (COSY, TOCSY, ROESY, HSQC, and HMBC) enabled the assignment of seven COSY spin systems consisting of five amino and two hydroxy acid residues [proline methyl ester (Pro-Me ester), *N*-MePhe, Ala, Ile, *N*-MeGln, and two 2-hydroxy-3-methylbutyric acids (Hmba)], accounting for all 13 degrees of unsaturation, and thereby signifying an overall linear arrangement (Figure 4.1). HMBC correlations from the NH and *N*-Me groups to the carbonyl of the neighboring residues allowed for the assignment of connections between *N*-MePhe and Ala (C-16 to C-17), *N*-MeGln and Hmba-1 (C-31 to C-32), Ala and Ile (NH-1 to C-20), and Ile and *N*-MeGln (*N*-H-2 to C-26), leading to three partial structures, *O*-MePro, *N*-MePhe – Ala – Ile – *N*-MeGln – Hmba-1, and Hmba-2. Key ROESY correlations revealed further connections between these fragments, one between H-5a/b ( $\delta$  3.34/3.15) and H-8 ( $\delta$  5.52) connecting Pro-Me ester to *N*-MePhe and the other between H-34 ( $\delta$  2.15) and H-38 ( $\delta$  4.09), making the final connection between the two Hmba residues. Thus, tasiamide C was deduced to have an overall linear structure consisting of Pro-Me ester– *N*-MePhe – Ala – Ile – *N*-MeGln – Hmba-1 – Hmba-2. This planar constitution was supported by tandem mass spectroscopy analysis (Figure 4.3).

The absolute configuration of several of the amino acids (Pro, *N*-MePhe, Ala, and *N*-MeGln) were determined by Marfey's analysis. Authentic D and L standards for Pro, *N*-MePhe, and Ala were each derivatized with D-(1-fluoro-2,4-dinitrophenyl-5-D-alanine amide) (D-FDAA). Unfortunately, authentic D-*N*-MeGlu was unavailable; therefore, chromatographic standards were prepared by derivatizing L-*N*-MeGlu with both D and L-FDAA (upon acidic hydrolysis, amides *N*-MeGln becomes *N*-MeGlu). Compound **1** was hydrolyzed and derivatized with D-FDAA and analyzed by LC-MS in comparison with the retention times of authentic standards. From this analysis it was clear that three of the four amino acids (Pro, Ala, and *N*-MeGln) were of the L configuration, while the *N*-MePhe residue was of the D configuration.

The absolute configuration of the Ile residue was determined by chiral-phase GC-MS comparison of *N*-Boc, *O*-Me derivatized authentic standards against the similarly derivatized Ile residue released by acid hydrolysis of **1**. The four protected standards of L-Ile, L-*allo*-Ile, D-Ile, and D-*allo*-Ile, were prepared by first synthesizing the *N*-Boc protected amino acids, followed by methyl esterification of the carboxylic acid using diazomethane. From retention time comparison and co-injection experiments it was clear that the Ile residue was of the D-*allo* configuration.

The absolute configuration of the final two Hmba stereocenters in **1** presented a challenge. Comparison of authentic standards of *S*- and *R*-Hmba with the methylated hydrolysate revealed that the natural product contained both *S*- and *R*-Hmba residue. A similar situation has been previously reported in closely related metabolites of this compound family.<sup>24</sup> Following mild base treatment (1:1, 0.5 N NaOH (*aq*)—MeOH), only the terminal Hmba residue was released; this residue was subsequently methyl esterified using diazomethane,<sup>24</sup> and by retention time comparison and co-injections with authentic standards, was identified as of the *S* configuration. The penultimate Hmba must therefore be of the *R* configuration. In summary, the above experiments established that tasiamide C (**1**) possessed a 2*S*, 8*R*, 18*S*, 21*R*, 22*S*, 27*S*, 33*R*, and 38*S* absolute configuration.

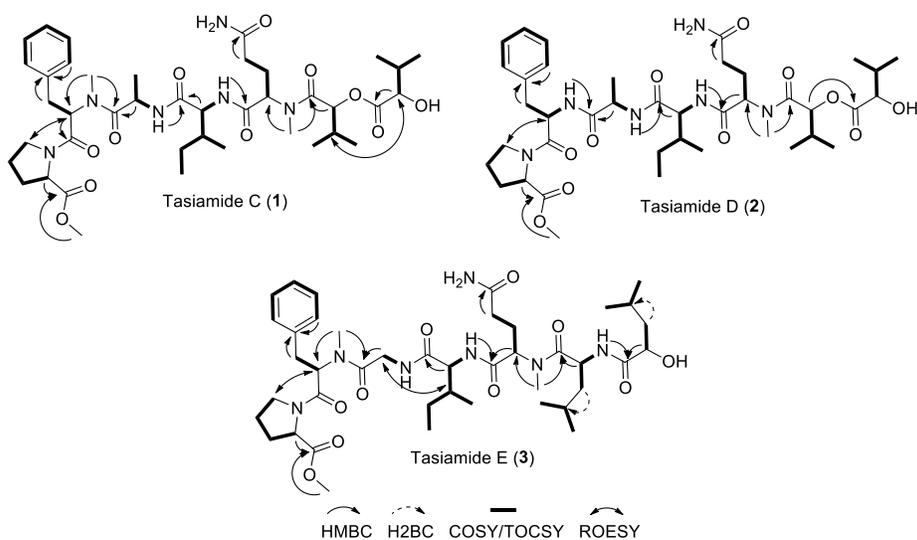
**Table 4.1:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for tasiamide C (**1**) in  $\text{CDCl}_3$ 

residue	position	$\delta_{\text{C}}^{\text{c}}$	$\delta_{\text{H}}$ ( $J$ in Hz) <sup>a</sup>	HMBC <sup>b</sup>	COSY <sup>b</sup>
Pro-Me ester	1	172.1			
	2	59.5	4.37, dd (7.5, 7.5)	1, 3	3a, 3b
	3a	28.7	2.16, m	1, 2, 4	2, 3b, 4a
	3b		1.86, m	1, 2, 4, 5	2, 3a, 4b
	4a	25.2	1.92, m	5	3a, 4b, 5a, 5b
	4b		1.75, m	3, 5	3b, 4a, 5a, 5b
	5a	47.1	3.34	3	4a, 4b, 5b
	5b		3.15, m	3, 4	4a, 4b, 5a
	6	52.3	3.67, s	1	
<i>N</i> -MePhe	7	167.8			
	8	55.9	5.52, dd (9.8, 6.1)	7, 9, 16	9a, 9b
	9a	34.7	3.14, m	7, 8, 19, 11, 15	8, 9b
	9b		3.87, dd (14.3, 9.7)	8, 10, 11, 15	8, 9a
	10	136.4			
	11/15	129.2	7.10, m	9, 10, 12, 14	12, 14
	12/14	128.0	7.15, m	10, 11, 13, 15	11, 13, 15
	13	126.4	7.11, m	12, 14	12, 14
	16	30.5	2.93, s	8, 17	
Ala	17	172.5			
	18	45.1	4.67, dq (7.4, 7.3)	17, 19, 20	19
	19	17.4	0.76, d (7.1)	17, 18	18
NH-1		6.69, d (8.6)	20	18	
Ile	20	169.9			
	21	58.0	4.14, dd (8.2, 7.3)	20, 22, 23, 25	22
	22	36.1	1.91, m	21, 23, 24, 25	21, 25
	23a	24.7	1.30, m	25	23b, 24
	23b		1.02, m	21, 22, 24	23a, 24

**Table 4.1: continued**

<b>residue</b>	<b>position</b>	<b><math>\delta_C^c</math></b>	<b><math>\delta_H</math> (J in Hz)<sup>a</sup></b>	<b>HMBC<sup>b</sup></b>	<b>COSY<sup>b</sup></b>
	24	11.1	0.78, t (7.4)	22, 23	23a, 23b
	25	15.4	0.74, d (6.8)	21, 22, 23	22
	NH-2		6.78, d (7.5)	26	21
<i>N</i> -MeGlu	26	169.2			
	27	56.7	4.93, dd (8.0, 7.0)	26, 28, 29, 32	28a, 28b
	28a	24.5	2.35, m	26, 27, 29, 30	27, 28b
	28b		1.86, m	26, 27, 29, 30	27, 28a, 29a
	29a	32.7	2.26, m	27, 28, 30	28b, 29b
	29b		2.10, m	30	29a
	30	173.7			
	NH <sub>2</sub> a		6.62, bs		
	NH <sub>2</sub> b		6.46, bs		
	31	31.2	3.03, s	27, 32	
Hmba-1	32	170.8			
	33	76.5	4.85, d (7.9)	32, 34, 35, 36, 37	34
	34	30.0	2.15, m	33, 35, 36	33, 35, 36
	35	18.3	0.99, d (6.7)	33, 34, 36	34
	36	18.2	0.93, d (7.0)	33, 34, 35	34
Hmba-2	37	175.7			
	38	74.6	4.09, d (3.7)	37, 39	39
	39	31.7	2.07, m	38, 40, 41	38, 40, 41
	40	18.9	0.97, d (6.9)	39, 41	39
	41	16.2	0.83, d (6.9)	39, 40	39

<sup>a</sup>600 MHz for <sup>1</sup>H NMR; <sup>b</sup>500 MHz for HMBC and COSY; <sup>c</sup>125 MHz for <sup>13</sup>C NMR



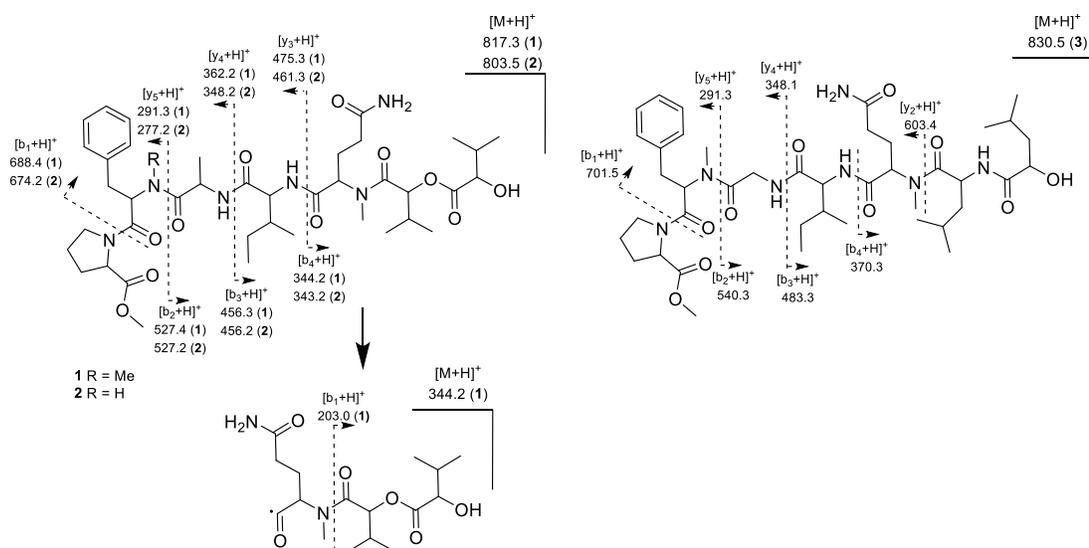
**Figure 4.2:** Select 2D NMR data for tasiamides C-E

#### 4.2.3 Tasiamide D Absolute Structure Determination

HR-ESITOFMS of **2** yielded an  $[M+Na]^+$  at  $m/z$  825.4371 for a molecular formula of  $C_{40}H_{62}N_6O_{11}$  (calcd for  $C_{40}H_{62}N_6NaO_{11}$ , 825.4369, 0.2 ppm). The IR,  $^1H$  and  $^{13}C$  NMR spectra were similar to those of **1**; however, the molecular formula indicated a reduction of 14 amu (e.g. a  $CH_2$  unit) relative to tasiamide C. Inspection of the  $^1H$  NMR spectrum revealed the presence of only two singlet methyl groups (e.g. one *N*-methyl at  $\delta_H$  3.02 and one *O*-methyl at  $\delta_H$  3.58) and three NH protons ( $\delta_H$  6.77, 6.89, and 6.98). From the 2D NMR data, it was clear that the only modification between **1** and **2** was the loss of the *N*-methyl on the Phe residue, thus yielding a planar constitution of Pro-Me ester–Phe–Ala–Ile–*N*-MeGln–Hmba-1–Hmba-2 for tasiamide D (**2**). This assembly was corroborated by tandem mass spectrometry analysis (Figure 4.3).

The absolute configurations of the residues in compound **2** were determined in an identical fashion as described above for compound **1**. Analysis of the retention times of the Marfey's derivatized hydrolysate and authentic standards by LC-MS revealed that three of the four amino acids (Pro, Ala, and *N*-MeGln) were of the L configuration and that the Phe residue was of the D

configuration. Chiral-phase GC-MS analysis of the Ile residue, *N*-Boc and *O*-methylated as with **1**, showed by retention time comparison and co-injections that this residue was of the *L* configuration. Chiral-phase GC-MS analysis of the methylated hydrolysate of **2** exhibited peaks matching both *S*- and *R*-Hmba. Analysis of the methylated mild base hydrolysate confirmed that, as with **1**, the terminal Hmba was of *S* configuration, and so the penultimate Hmba was of *R* configuration; establishing the absolute configuration of tasiamide D (**2**) as 2*S*, 8*R*, 17*S*, 20*S*, 21*S*, 26*S*, 32*R*, and 37*S*.



**Figure 4.3:** Select low resolution MS fragmentation cleavages for tasiamides C-E

**Table 4.2:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for tasiamide D (**2**) in  $\text{CDCl}_3$ 

residue	position	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}}$ ( $J$ in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
Pro-Me ester	1	172.2			
	2	59.1	4.13, m	1, 3	3
	3	29.0	1.78, m	2, 4	3, 4a, 4b, 5a, 5b
	4a	24.6	1.35, m	5	3, 4b, 5b
	4b		1.07, m	3, 5	3, 4b, 5b
	5a	46.9	3.47, dd (8.7, 5.1)	4, 5	3, 4b, 5b
	5b		2.68, m	5	4a, 4b, 5a
Phe	6	52.2	3.58, s	1	
	7	170.4			
	8	52.5	4.76, q (7.6)	7, 9	9, NH-1
	9	39.0	2.94, m	7, 8, 10, 11, 15	8
	10	136.2			
	11/15	129.1	7.09, m	9, 10, 12, 13, 14	12, 14
	12/14	128.5	7.17, m	10, 11, 13, 15	11, 12, 15
13	127.0	7.13, m	11, 12, 14, 15	12, 14	
NH-1		6.98, d (6.8)	17	8	
Ala	17 <sup>c</sup>	171.8			
	18	48.8	4.43, dq (8.0, 6.8)	17, 19	19, NH-2
	19	18.1	1.26, d (7.2)	17, 18	18
	NH-2		6.89, d (7.7)	20	18
Ile	20	170.7			
	21	59.4	4.12, m	20, 22, 25	22, NH-3
	22	35.9	1.97, m	21, 23, 24, 25	21, 23a, 23b, 25
	23a	24.7	1.35, m	22, 24, 25	22, 23b, 24
	23b		1.07, m	22, 24, 25	22, 23a, 24
	24	11.5	0.82, t (6.8)	22, 23	23a, 23b
	25	15.7	0.82, d (6.3)	21, 22, 23	22

**Table 4.2: continued**

residue	position	$\delta_C^b$	$\delta_H$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
	NH-3		6.77, d (6.7)	26	21
N-MeGln	26	170.4			
	27	57.1	5.08, dd (7.6, 6.7)	26, 28, 29	28a, 28b
	28a	24.0	2.48, m	26, 27, 29, 30	28b, 29a
	28b		1.81, m	26, 27, 29, 30	28a, 29b
	29a	33.0	2.32, m	27, 28, 30	28a, 29b
	29b		2.25, m	27, 28, 30	28b, 29a
	30	174.6			
	NH <sub>2</sub> a		5.98, bs		
	NH <sub>2</sub> b		5.62, bs		
		31	31.1	3.02, s	27, 32
Hmba-1	32	171.2			
	33	76.7	4.85, d (7.5)	32, 34, 35, 36, 37	34
	34	29.9	2.14, m	33, 35, 36	33, 35, 36
	35	18.7	0.93, d (7.0)	33, 34, 36	34
	36	18.1	0.99, d (6.1)	33, 34, 35	34
Hmba-2	37	175.8			
	38	75.1	4.11, m	37, 39, 40, 41	39
	39	32.1	2.07, m	38, 40, 41	38, 40, 41
	40	18.4	0.97, d (6.2)	38, 39, 41	39
	41	16.1	0.82, d (6.3)	38, 39, 40	39

<sup>a</sup>500 MHz for <sup>1</sup>H NMR, HMBC, and COSY; <sup>b</sup>75 MHz for <sup>13</sup>C NMR; <sup>c</sup>Carbon 16 was intentionally skipped in order to be consistent with the numbering displayed in figure 4.1

#### 4.2.4 Tasiamide E Absolute Structure Determination

HR-ESITOFMS of **3** yielded an [M+Na]<sup>+</sup> at  $m/z$  852.4842 in agreement with a molecular formula of C<sub>42</sub>H<sub>67</sub>N<sub>7</sub>O<sub>10</sub> (calcd for C<sub>42</sub>H<sub>67</sub>N<sub>7</sub>NaO<sub>10</sub>, 852.4842, 0 ppm), requiring 13 degrees of

unsaturation. The IR,  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra again featured similarities to **1** and **2**; however, the molecular weight was 13 and 27 amu greater than **1** and **2**, respectively, and thus could not be readily attributed to a single modification. Closer inspection of the  $^1\text{H}$  NMR spectrum revealed the presence of three singlet methyls ( $\delta_{\text{H}}$  2.98, 2.99, and 3.72), similar to **1** and three amide protons ( $\delta_{\text{H}}$  6.78, 6.82, and 7.01), as seen in **2**; suggesting that one of the hydroxy acids was replaced by an amino acid. Further evidence of this was seen by comparison of the  $^{13}\text{C}$  NMR spectra of **1** and **3**; in tasiamide C there were two  $\alpha$ -carbons with shifts indicative of hydroxy acids at  $\delta_{\text{C}}$  74.6 and 76.4 whereas there was only one such peak in tasiamide E at  $\delta_{\text{C}}$  71.1.

Further analysis of 1D and 2D NMR data ( $^1\text{H}$ ,  $^{13}\text{C}$ , COSY, TOCSY, ROESY, HSQC, HMBC, and H2BC) confirmed the presence of seven residues, six amino acids, and one hydroxy acid [Pro-Me ester, *N*-MePhe, Gly, Ile, *N*-MeGln, Leu, and 2-hydroxy-4-methylpentaic acid (H4mpa)]. In a similar fashion to **1**, each of the connections between residues in **3** were revealed by key HMBC and ROESY correlations. HMBC correlations from both *N*-methyl groups to the carbonyls on the neighboring residues allowed for the assignment of connections between *N*-MePhe to Gly (C-16 to C-17) and *N*-MeGln to Leu (C-30 to C-31). Correspondingly, HMBC correlations from amide protons to adjacent carbonyls allowed for the assignment of connections between Ile and *N*-MeGln (NH-2 to C-25), and Leu and H4mpa (*N*-H-3 to C-37), leading to three partial structures, Pro-Me ester, *N*-MePhe – Gly, and Ile – *N*-MeGln – Leu – H4mpa. Two ROESY correlations provided the final connections between these fragments, one between H-5a/b ( $\delta$  3.36/3.31) and H-8 ( $\delta$  5.55), connecting Pro-Me ester and *N*-MePhe, and the other between H-18a/b ( $\delta$  4.05/3.81) and H-21 ( $\delta$  1.82), making the final connection between Gly and Ile. Thus, tasiamide E was deduced to have a planar linear structure consisting of Pro-Me ester – *N*-MePhe – Gly – Ile – *N*-MeGln – Leu – H4mpa and this was supported by tandem mass spectrometry analysis (Figure 2).

The absolute configurations of the residues in **3** were determined as described above for compounds **1** and **2** using appropriate authentic standards. Analysis of the Marfey's derivatized hydrolysate with the retention times of authentic standards on LC-MS revealed that three of the four amino acids (Pro, Leu, and *N*-MeGln) were of the L configuration while the *N*-MePhe residue was of the D configuration. As for the chiral GC-MS analysis of the Ile residue, retention time comparison and co-injections confirmed that it was also of the L configuration. The absolute configuration of the H4mpa residue was determined using chiral-phase GC-MS, both comparing retention time and co-injections with authentic standards, thus confirming its *S* configuration, and establishing the absolute configuration of tasiamide E (**3**) as *2S*, *8R*, *20S*, *21S*, *26S*, *32S*, and *38S*.

**Table 4.3:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for tasiamide E (**3**) in  $\text{CDCl}_3$ 

residue	position	$\delta_{\text{C}}^{\text{c}}$	$\delta_{\text{H}}$ ( $J$ in Hz) <sup>a</sup>	HMBC <sup>b</sup>	COSY <sup>a</sup>
Pro-Me ester	1	172.5			
	2	58.9	4.39, dd (8.1, 5.2)	1, 3, 4	3a, 3b
	3a	28.8	2.12, m	2	2, 3b, 4b
	3b		1.86, m	2, 4	2, 3a, 4a
	4a	25.0	1.93, m	2, 3, 5	3b, 4b, 5b
	4b		1.80, m	2, 3	3a, 4a, 5a
	5a	46.8	3.36, m	4	4b, 5b
	5b		3.31, m	4	4a, 5a
	6	52.3	3.72, s	1	
<i>N</i> -MePhe	7	167.8			
	8	56.2	5.55, t (7.4)	7, 9, 16	9a, 9b
	9a	35.1	3.27, dd (13.6, 8.2)	7, 8, 10, 11	8, 9b
	9b		2.82, dd (13.6, 6.9)	7, 8, 10, 11, 16	8, 9a
	10	136.8			
	11/15	129.4	7.22, m	9, 10, 12	12, 14
	12/14	128.4	7.26, m	11, 13, 15	11, 13, 15
16	29.7	2.98, s	8, 9, 17		
Gly	17	167.5			
	18a	41.2	4.05, dd (17.7, 4.6)	17	18b, NH-1
	18b		3.81, dd (17.7, 3.3)	17	18a, NH-1
	NH-1		6.78, d (4.36)		18a, 18b
Ile	19	171.5			
	20	57.6	4.29, dd (8.8, 6.3)	19, 21, 22	21, NH-2
	21	37.2	1.82, m	20, 23, 24	20, 22b, 24
	22a	24.7	1.43, m	23, 24	22b
	22b		1.11, m	23, 24	21, 22a, 23

**Table 4.3: continued**

residue	position	$\delta_C^c$	$\delta_H$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
	23	11.3	0.87, t (7.7)	21, 22	22b
	24	15.6	0.88, d (7.2)	21, 22	21
	NH-2		6.82, d (9.4)	25	20
<i>N</i> -MeGln	25	169.7			
	26	56.2	5.06, t (7.4)	25, 27, 30	27a, 27b
	27a	23.4	2.31, m	28	26, 27b, 28a
	27b		1.99, m	27	26, 27a, 28a
	28a	31.9	2.24, m	27, 29	27b, 28b
	28b		2.18, m	29	27a, 28a
	29	174.0			
	NH <sub>2</sub> a		5.61, bs		
	NH <sub>2</sub> b		5.25, bs		
	30	31.0	2.99, s	26, 31	
Leu	31	174.4			
	32	56.9	4.95, q (7.9)	33	33, NH-3
	33	41.2	1.60, m	32, 34	32
	34	24.8	1.60, m	33, 35, 36	35, 36
	35	21.4	0.92, d (6.4)	34	34
	36	22.7	0.95, d (6.4)	34	34
	NH-3		7.01, d (9.3)	37	
H4mpa	37	175.2			
	38	71.1	4.09, dd (9.1, 4.5)	37, 39	39
	39	43.1	1.54, m	38, 40	38, 40
	40	24.6	1.87, m	38, 41, 42	39, 41, 42
	41	22.3	0.96, d (6.6)	40	40
	42	23.0	0.96, d (6.6)	40	40

<sup>a</sup> 500 MHz for <sup>1</sup>H NMR and COSY; <sup>b</sup> 600 MHz for HMBC; <sup>c</sup> 125 MHz for <sup>13</sup>C NMR

#### 4.2.5 Structural Similarities of Tasiamides C-E to Known Metabolites

Tasiamides C-E (**1-3**) are of close structural relation to several families of known metabolites, the grassystatins, tasiamides, and symplocin A. In this regard, **1-3** are structurally most similar to tasiamide (**4**), with **3** varying only in the terminal residue, while compounds **1** and **2** possess other relatively simple modifications. In 2008, Li and co-workers revised the original configuration of the *N*-MeGln residue in **4** from L to D based on the comparison of analytical data (<sup>13</sup>C NMR and specific rotation) of the natural product with four stereoisomers obtained from a total synthesis (containing both L and D stereoisomers of *N*-MeGln and Leu; Table 2).<sup>25</sup> However, the original assignments for all of the residues in **4** are consistent with that of **3**, and a comparison of <sup>13</sup>C shifts revealed no significant differences between **3** and **4**, along with the four additional synthetic stereoisomers (**A-D**, Table 2).<sup>26</sup> Furthermore, the opposite specific rotation signs for tasiamide C and D suggests a structural variance between these two metabolites (**3**,  $[\alpha]_{\text{D}}^{25} -22.2$ ; **4**,  $[\alpha]_{\text{D}}^{21} +15.0$ ). Therefore, a more detailed analysis of the <sup>13</sup>C NMR data for tasiamide and the four synthetic analogs was conducted using DP4 probability calculations, and revealed that the carbon data alone is insufficient to deduce the correct configuration in tasiamide (**4**).<sup>27</sup> The probability of the original experimental data matching each analog was calculated as: 1.3% for analog **A** (**4**), 56.0% for analog **B**, 3.0% for analog **C**, and 39.7% for analog **D**. Because analogs **B** and **D** have the same sign and magnitude of specific optical rotation, and are indistinguishable by DP4 calculations, it would be impossible to conclusively assign the configuration of **4** solely based on this data. This also suggests that the misassignment in tasiamide (**4**) may involve residues other than just the *N*-MeGln, and thus, a broader investigation into the configuration of **4** is necessary to clarify its correct absolute configuration.

It is interesting to note that there appears to be considerable biosynthetic flexibility in this family of metabolites. Insights are thus provided by comparing the amino- or hydroxy-acid residues in each member of this family (Table 2). The first three residues (A, B, and C) starting

from the carboxy terminus of these molecules are rather well conserved, with each analog containing an Pro-Me ester followed by the incorporation of D-Phe, which is *N*-methylated in all but one analog, and then the incorporation of an Ala, Aba, or Gly as the third residue. Residue E, H, and J are also conserved residues (statine, Hmba, and an *N,N*-dimethylated amino acid, respectively) but they occur only in a subset of the natural products. Residue F and I are slightly less conserved, with three different polar amino acids (*N*-MeGln, Asn, and Ser) incorporating into residue F and four different hydroxy acids (Hmpa, H4mpa, Hpa, and H3mpa) in residue I. Also observed are variations in the absolute configurations of these residues. For example, residue D and J in eight of nine metabolites incorporate an L residue, while a single metabolite has a *D-allo* residue at this position. At residue F all of the amino acids were originally deduced as L; however, in tasiamide and tasiamide B these residues were reassigned as D based on total synthesis. Although non-ribosomal peptide synthetase (NRPS)-derived peptides can show variations in the incorporated amino acid, it is difficult to imagine a single biosynthetic pathway that would be capable of producing the all of these metabolites, especially those of varying absolute configurations, which usually requires an epimerase.<sup>28</sup>

#### 4.2.6 Bioassay Results

Both the grassystatins and symplocin A were reported to possess exceptional inhibitory activity toward cathepsin E (grassystatin A:  $IC_{50} = 886$  pM, symplocin A:  $IC_{50} = 300$  pM),<sup>24,29</sup> however, this activity is likely due to the presence of a statine residue, which is absent in the tasiamides. Additionally, tasiamide and tasiamide B both showed moderate cytotoxicity against KB cells, with  $IC_{50}$  values of 0.48 and 0.8  $\mu$ M, respectively.<sup>26,30</sup> Due to limited isolated quantities of tasiamide E (**3**), only tasiamide C (**1**) and D (**2**) were evaluated for their cytotoxicity against the HCT-116 colon cancer cell line and were found to be inactive (tasiamide C,  $IC_{50} > 25$   $\mu$ M; tasiamide D,  $IC_{50} \approx 25$   $\mu$ M).

**Table 4.4.** Residue sequence comparisons of the tasiamide/grassystatin/symplocin superfamily of metabolites.

Compound	A	B	C	D	E	F	G	H	I	J
Grassystatin A (5) <sup>24</sup>	L-OMe-Pro	D-NMe-Phe	L-Ala	L-Thr	(S,S)-Sta	L-Asn	L-Leu	D-Hmba	L-Hmba	L-N,N-diMe-Val
Grassystatin B (6) <sup>24</sup>	L-OMe-Pro	D-NMe-Phe	L-Aba	L-Thr	(S,S)-Sta	L-Asn	L-Leu	D-Hmba	L-Hmba	L-N,N-diMe-Val
Grassystatin C <sup>24</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile	(S,S)-Sta	L-NMe-Gln	L-Leu	D- <i>allo</i> -H3mpa	D- <i>allo</i> -H3mpa	
Symplocin A <sup>29</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Val	(R,S)-Sta	L-Ser	L-Tyr	D-Hmba		D-N,N-diMe-Ile
Tasiamide (4) <sup>26,a</sup>	OMe-Pro	NMe-Phe	Gly	Ile		NMe-Gln	Leu		Hmba	
Tasiamide B <sup>30,a</sup>	OMe-Pro	NMe-Phe	Ala	Leu	Ahppa	NMe-Gln	Val		Hpa	
Tasiamide C (1)	L-OMe-Pro	D-NMe-Phe	L-Ala	D- <i>allo</i> -Ile		L-NMe-Gln		D-Hmba	L-Hmba	
Tasiamide D (2)	L-OMe-Pro	D-Phe	L-Ala	L-Ile		L-NMe-Gln		D-Hmba	L-Hmba	
Tasiamide E (3)	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile		L-NMe-Gln	L-Leu		L-H4mpa	
Tasiamide Synthetic Analog A (4) <sup>25,b</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile		L-NMe-Gln	L-Leu		L-Hmba	
Tasiamide Synthetic Analog B <sup>17,c</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile		D-NMe-Gln	L-Leu		L-Hmba	
Tasiamide Synthetic Analog C <sup>17</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile		L-NMe-Gln	D-Leu		L-Hmba	
Tasiamide Synthetic Analog D <sup>17</sup>	L-OMe-Pro	D-NMe-Phe	Gly	L-Ile		D-NMe-Gln	D-Leu		L-Hmba	

<sup>a</sup> Stereochemical assignments for tasiamide and tasiamide B intentionally omitted due to unresolved stereochemical questions, <sup>b</sup> Matches original configuration proposed by Williams et al.<sup>26</sup>, <sup>c</sup> Proposed revised structure of tasiamide according to Ma et al.<sup>25</sup>

### 4.3 Conclusion

The new metabolites, tasiamides C-E (**1-3**), were isolated from a Kimbe Bay, Papua New Guinea collection of the tropical marine cyanobacterium *Symploca* sp. These metabolites add to the growing family of structurally homologous natural products [tasiamide (**4**), grassystatins (**5-6**) and symplocin A] that have all been isolated from similar tuft-forming cyanobacteria. This particular *Symploca* sp. is a prolific producer of secondary metabolites as several structurally diverse natural products were isolated from this collection including kimbeamide A-C, kimbelactone, tasihalide C, and the three tasiamides reported herein.<sup>22-23</sup>

### 4.4 Experimental Methods

#### 4.4.1 General Experimental Procedures

Optical rotations were measured on a JASCO P-2000 polarimeter, UV spectra on a Beckman Coulter DU-800 spectrophotometer, and IR spectra on a Nicolet IR-100 FT-IR spectrophotometer using KBr plates. NMR spectra were recorded with chloroform as an internal standard ( $\delta_C$  77.0,  $\delta_H$  7.26 for CHCl<sub>3</sub>) on a Varian Unity 500 MHz spectrometer (500 and 125 MHz for <sup>1</sup>H and <sup>13</sup>C NMR, respectively), a Varian Unity 300 MHz spectrometer (300 and 75 MHz for <sup>1</sup>H and <sup>13</sup>C NMR, respectively), a Varian VNMRS (Varian NMR System) 500 MHz spectrometer equipped with a cold probe (500 and 125 MHz for <sup>1</sup>H and <sup>13</sup>C NMR), and Bruker 600 MHz spectrometer equipped with a 1.7 mm MicroCryoProbe (600 and 150 MHz for <sup>1</sup>H and <sup>13</sup>C NMR). LR- and HR-ESIMS data were obtained on ThermoFinnigan LCQ Advantage Max and Thermo Scientific LTQ Orbitrap-XL mass spectrometers, respectively. Tandem mass spectroscopy experiments were run with a Biversa Nanomate (Advion Biosystems) electrospray source for a Finnigan LTQ-FTICR-MS instrument (Thermo-Electron Corporation) running Tune Plus software version 1.0. HPLC was carried out using a Waters 515 pump system with a Waters 996 PDA detector. GC-MS was conducted with a Thermo Electron Corp. DSQ/TRACE-GC-Ultra GCMS system. All solvents were

either distilled or of HPLC quality. Acid hydrolysis was performed using a Biotage (Initiator) microwave reactor equipped with high pressure vessels.

#### 4.4.2 Extraction and Isolation

The cyanobacterial biomass (101.7 g dry wt) was extracted with 2:1 CH<sub>2</sub>Cl<sub>2</sub>—MeOH to afford 1.8 g of dried extract. A portion of the extract was fractionated by silica gel VLC using a stepwise gradient solvent system of increasing polarity starting from 100% hexanes to 100% MeOH (nine fractions, A-I). The fraction eluting with 25% MeOH/75% EtOAc (fraction H) was separated further using RP HPLC [4 $\mu$  Synergi Hydro, isocratic 65% MeCN/35% H<sub>2</sub>O] to yield pure tasiamide C (**1**, 1.9 mg), tasiamide D (**2**, 2.5 mg) and tasiamide E (**3**, 0.7 mg).

**Tasiamide C (1):** White amorphous solid;  $[\alpha]_D^{25}$  -36.6 (*c* 1.17, CHCl<sub>3</sub>); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 211.0 (3.84); IR (neat)  $\nu_{\max}$  3371, 2965, 2933, 2877, 1737, 1644, 1521, 1453, 1263, 1201, 1178, 1098, 1031 cm<sup>-1</sup>; <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>), see Table 4.1; HRESIMS [M+Na]<sup>+</sup> *m/z* 839.4541 (calcd for C<sub>41</sub>H<sub>64</sub>N<sub>6</sub>O<sub>11</sub>Na, 839.4525).

**Tasiamide D (2):** White amorphous solid;  $[\alpha]_D^{25}$  -84.7 (*c* 1.67, CHCl<sub>3</sub>); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 220.0 (3.86) nm; IR (neat)  $\nu_{\max}$  3318, 2965, 2931, 1739, 1648, 1526, 1453, 1203, 1033 cm<sup>-1</sup>; <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>), see Table 4.2; HRESIMS [M+Na]<sup>+</sup> *m/z* 825.4371 (calcd for C<sub>40</sub>H<sub>62</sub>N<sub>6</sub>O<sub>11</sub>Na, 825.4369).

**Tasiamide E (3):** White amorphous solid;  $[\alpha]_D^{25}$  -22.2 (*c* 0.533, CHCl<sub>3</sub>); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 207.0 (4.10) nm; IR (neat)  $\nu_{\max}$  3329, 2926, 2958, 1743, 1645, 1521, 1454, 1281, 1177 cm<sup>-1</sup>; <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>), see Table 4.3; HRESIMS [M+Na]<sup>+</sup> *m/z* 852.4842 (calcd for C<sub>42</sub>H<sub>67</sub>N<sub>7</sub>O<sub>10</sub>Na, 852.4842).

#### 4.4.3 Acid Hydrolysis and Marfey's Analysis of Tasiamides C-E

Tasiamides C-E (**1-3**, 0.2 mg) were separately treated with 400  $\mu\text{L}$  of 6 N HCl in a microwave reactor at 160  $^{\circ}\text{C}$  for 5 min. An aliquot of the reaction product was dissolved in 500  $\mu\text{L}$  of 1 mg/mL solution of D-FDAA (1-fluoro-2,4-dinitrophenyl-5-D-alanine amide) in acetone followed by the addition of 20  $\mu\text{L}$  1 N  $\text{NaHCO}_3$ . The solution was maintained at 40  $^{\circ}\text{C}$  for 1 h at which time the reaction was quenched by the addition of 40  $\mu\text{L}$  of 1 N HCl. The reaction mixture was then dried down under  $\text{N}_2$  (g), re-suspended in 200  $\mu\text{L}$  of 50%  $\text{H}_2\text{O}$ /50%  $\text{CH}_3\text{CN}$  and 10  $\mu\text{L}$  of the solution was analyzed by LC-ESIMS.

The Marfey's derivatives of the hydrolysate and standards reacted with D-FDAA (Ala, Phe, *N*-MePhe, Pro, and Leu) were analyzed by RP HPLC using a Phenomenex Luna 5  $\mu\text{m}$   $\text{C}_{18}$  column (4.6 x 250 mm). The HPLC conditions began with 10%  $\text{CH}_3\text{CN}$ /90%  $\text{H}_2\text{O}$  acidified with 0.1% formic acid (*aq*) (FA) followed by a gradient profile to 50%  $\text{CH}_3\text{CN}$ /50%  $\text{H}_2\text{O}$  acidified with 0.1% FA (*aq*) over 85 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. Because authentic D-*N*MeGlu was not available, L-*N*MeGlu standard was derivatized with both the L-FDAA and D-FDAA. The *N*MeGlu residue was analyzed by RP HPLC using a Kinetex 5 $\mu$   $\text{C}_{18}$  100A column (4.6 x 100 mm). The HPLC condition began with 5% MeOH/95%  $\text{H}_2\text{O}$  acidified with 0.1% FA (*aq*) followed by a gradient profile to 45% MeOH/55%  $\text{H}_2\text{O}$  acidified with 0.1% FA (*aq*) over 125 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. The retention times of the derivatives of the authentic amino acids when analyzing for **1** were D-Ala (64.8 min), L-Ala (71.1 min), D-Pro (66.6 min), L-Pro (69.4 min), D-*N*MePhe (78.4 min), L-*N*MePhe (79.7 min), L-*N*MeGlu reacted with L-FDAA (93.3 min), and L-*N*MeGlu reacted with D-FDAA (91.9 min); the derivatives of the hydrolysate product of **1** gave peaks with retention times of 71.5, 69.8, 78.3, and 91.8 min, according to L-Ala, L-Pro, D-*N*MePhe, and L-*N*MeGlu,. The retention times of the derivatives of the authentic amino acids when analyzing for **2** were D-Ala (59.3 min), L-Ala (63.6 min), D-Pro (60.4 min), L-Pro (62.6 min), D-Phe (78.7 min), L-Phe (83.8 min), L-*N*MeGlu reacted with L-FDAA (93.3 min) and L-*N*MeGlu reacted with D-FDAA (91.9 min); the derivatives of the hydrolysate

product of **2** gave peaks with retention times of 64.0, 62.6, 83.8 and 91.7 min, according to L-Ala, L-Pro, D-Phe, and L-Glu. The retention times of the derivatives of the authentic amino acids when analyzing for **3** were D-Leu (68.4 min), L-Leu (85.8 min), D-Pro (60.4 min), L-Pro (62.6 min), D-NMePhe (78.6 min), L-NMePhe (79.2 min), L-NMeGlu reacted with L-FDAA (107.6 min), and L-NMeGlu reacted with D-FDAA (103.9 min); the derivatives of the hydrolysate product of **3** gave peaks with retention times of 85.7, 62.2, 78.2, and 104.5 min, according to L-Leu, L-Pro, D-NMePhe, and L-NMeGlu.

#### 4.4.4 Preparation and GCMS Analysis of Isoleucine (Ile) in Tasiamides C-E

An aliquot (~0.1 mg) of the above hydrolysate product was dissolved in 100  $\mu\text{L}$   $\text{H}_2\text{O}$  and then treated with 0.263 mg (0.429  $\mu\text{mol}$ ) of  $\text{NaHCO}_3$  followed by 0.936 mg (0.429  $\mu\text{mol}$ ) of *tert*-butyl dicarbonate. After 16 h at room temperature, the reaction mixture was quenched by 500  $\mu\text{L}$  of 5%  $\text{KHSO}_4$ , back extracted 3 x 1 mL of EtOAc and the combined organic layers were dried over  $\text{MgSO}_4$ . The reaction product was then treated with freshly prepared diazomethane in diethyl ether. Each of the four Ile standards were prepared in a similar fashion.

The derivatized hydrolysate product and standards were analyzed by chiral-phase GC-MS using a Chirasil-Val (Agilent Technologies J&W Scientific, 30 m X 0.25 mm) under the following conditions: the initial oven temperature was 40  $^\circ\text{C}$ , kept for 2 min, followed by a ramp from 40 to 75  $^\circ\text{C}$  at a rate of 10  $^\circ\text{C}/\text{min}$ , kept for 5 min, followed by another ramp to 110  $^\circ\text{C}$ , at a rate of 0.5 $^\circ\text{C}/\text{min}$ , followed by a final ramp to 200  $^\circ\text{C}$ , at a rate of 25  $^\circ\text{C}/\text{min}$ , kept for 2 min. The retention times for the four authentic standards when analyzing **1** were D-*allo*-Ile (49.6 min), L-*allo*-Ile (49.9), D-Ile (51.5 min), and L-Ile (51.7 min); the derivatized hydrolysis product of **1** yielded a peak at 49.6 min, according to D-*allo*-Ile. The retention times for the four authentic standards when analyzing **2** and **3** were D-*allo*-Ile (47.7 min), L-*allo*-Ile (48.2), D-Ile (50.1 min), and L-Ile (50.5

min); the derivatized hydrolysis products of **2** and **3** each yielded a peak at 50.6 min, according to L-Ile.

#### 4.4.5 Base Hydrolysis of Hmba Units in Tasiamides C and D

Tasiamide C (**1**) and D (**2**) (0.150 mg) was treated with 150  $\mu$ L of 1:1 0.5 N NaOH (*aq*)—MeOH (1:1) solution at room temperature for 72 h.<sup>24</sup> The reaction mixture was neutralized by the addition of 40  $\mu$ L of 1 N HCl (*aq*), back extracted with 3 x 1 mL of EtOAc and the combined organic layers were dried over MgSO<sub>4</sub>. The product was then treated with freshly prepared diazomethane in diethyl ether for 5 min at rt. The derivatized acid hydrolysate was analyzed by chiral-phase GC-MS using a Cyclosil B column (Agilent Technologies J&W Scientific, 30 m X 0.25 mm) under the following conditions: the initial oven temp was 35 °C, kept for 15 min, followed by a ramp to 60 °C, at a rate of 1.5 °C, followed by ramp to 170 °C, at a rate of 5 °C/min, kept for 5 min. The retention times for the authentic standards were *R*-Hmba (34.9 min) and *S*-Hmba (35.8 min); the derivatized hydrolysis product for **1** and **2** each exhibited a single peak at 35.8 and 35.9 min, respectively, according to *S*-Hmba.

#### 4.4.6 GCMS Analysis of H4mpa in Tasiamide E

The derivatized hydrolysate product used in the Ile analysis from **3** was also used in the H4mpa analysis. Authentic standards were synthesized following a previously published method.<sup>31</sup> The derivatized hydrolysis product and authentic standards were analyzed by chiral-phase GC-MS using a Cyclosil B column under the following conditions: the initial oven temp was 40 °C, kept for 15 min, followed by a ramp to 90 °C, at a rate of 1.5 °C/min, followed by a ramp to 200 °C, at a rate of 10 °C/min, kept for 5 min. The retention times for the authentic standards were *R*-H4mpa (43.2 min) and *S*-H4mpa (43.8 min); the derivatized hydrolysis product yielded a peak at 43.7 min, consistent with *S*-H4mpa.

#### 4.4.7 H-460 Cytotoxicity Assay

H-460 cells were added to 96-well plates at  $3.33 \times 10^4$  cells/mL of Roswell Park Memorial Institute (RPMI) 1640 medium with 10% fetal bovine serum (FBS) and 1% Penicillin/Streptomycin. The cells, in a volume of 180  $\mu$ L per well, were incubated overnight (37°C, 5% CO<sub>2</sub>) to allow recovery before treatment with test compounds. Compounds were dissolved in DMSO to a stock concentration of 10 mg/mL. Working solutions of the compounds were made in RPMI 1640 medium without FBS, with a volume of 20  $\mu$ L added to each well to give a final compound concentration of either 30 or 3  $\mu$ g/mL. An equal volume of RPMI 1640 medium without FBS was added to wells designated as negative controls for each plate. Plates were incubated for approximately 48 h before staining with MTT. Using a ThermoElectron Multiskan Ascent plate reader, plates were read at 570 and 630 nm.

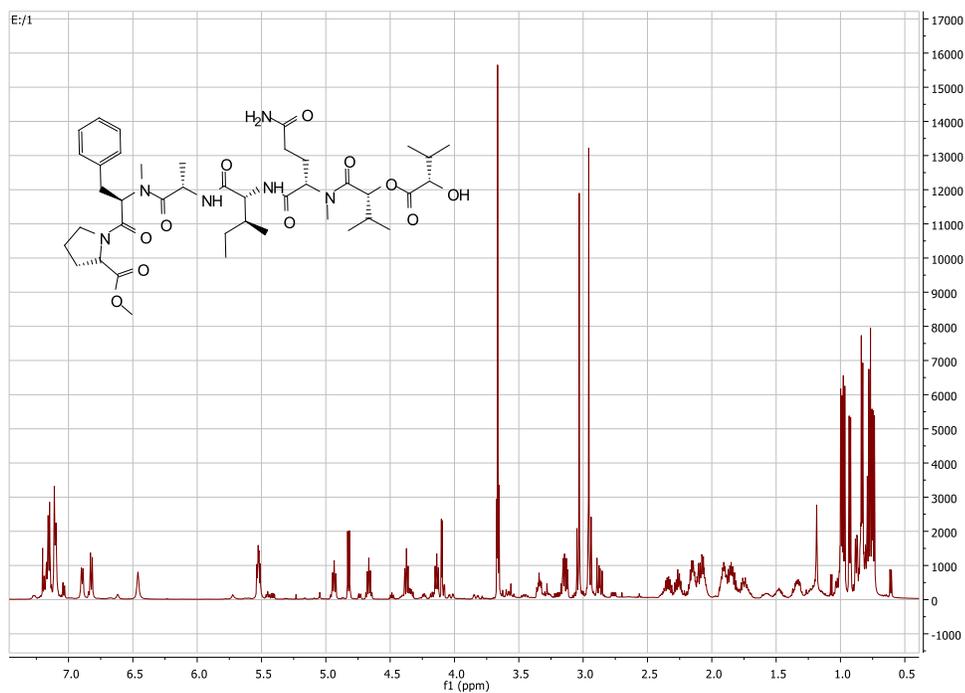
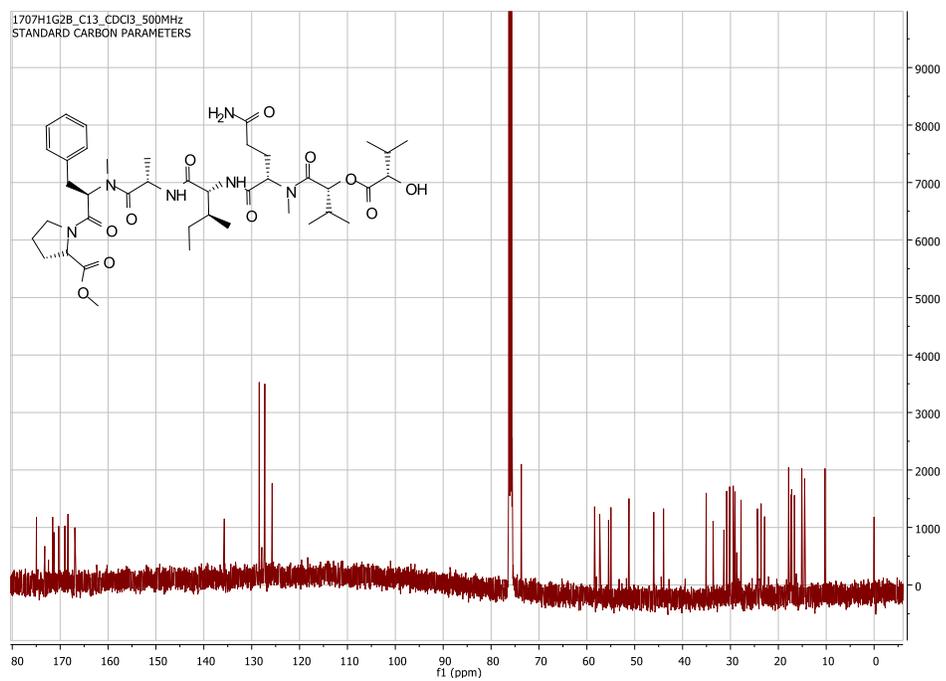
#### 4.4.8 HCT-116 Cytotoxicity Assay

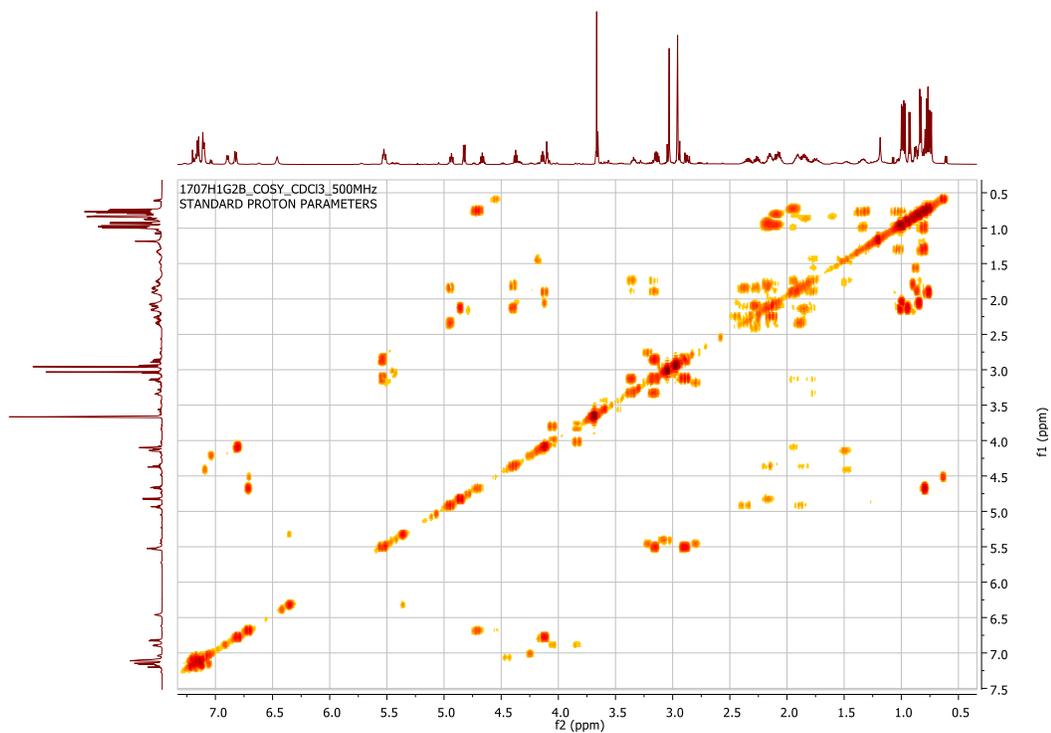
Cytotoxicity was measured in HCT-116 cells using a hemocytometer. These cells were grown in culture medium [5 mL; RPMI-1640 containing FBS (15 %), penicillin—streptomycin (1 %) and glutamine (1 %)] at 37 °C and CO<sub>2</sub> (5 %) at a starting cell density of  $5 \times 10^4$  cells per T25 flask. On day 3, cells were exposed to different concentrations of the metabolite. Flasks were incubated for 120 h (5 d) in a CO<sub>2</sub> (5 %) incubator at 37 °C, and the cells were harvested with trypsin, washed once with HBSS, and then re-suspended in HBSS and counted using a hemocytometer.

#### 4.5 Acknowledgements

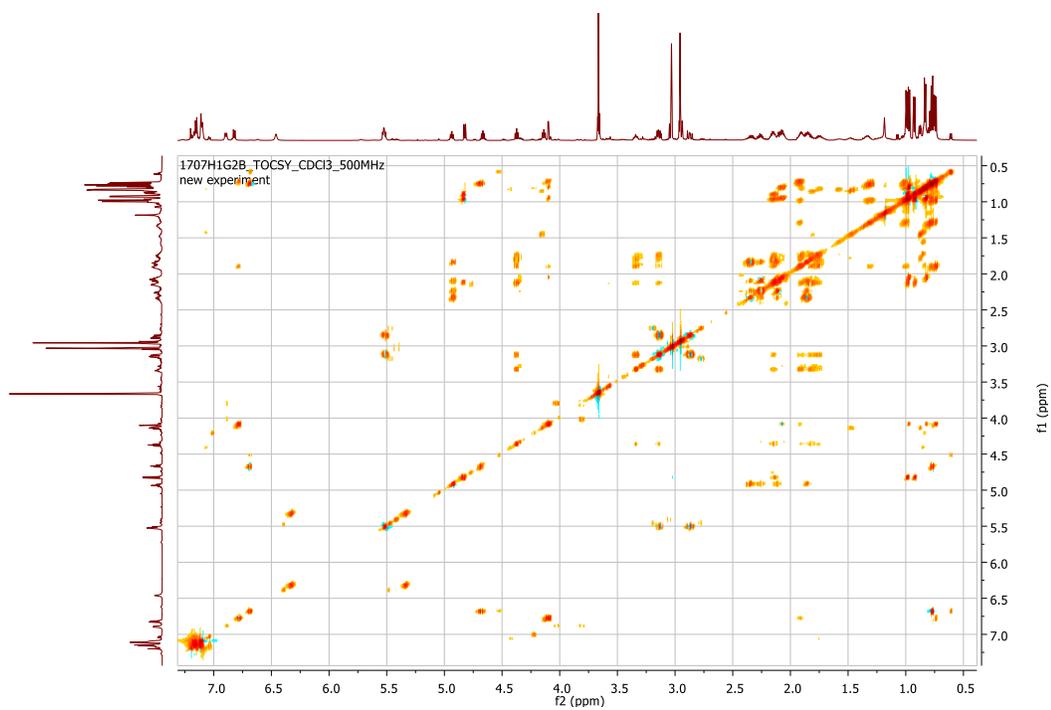
We thank the government of Papua New Guinea for permission to collect the cyanobacterial specimens, F. P. J. Haeckl, P. D. Boudreau, T. Byrum, P. C. Dorrestein, F. A. Valeriote, and W. H. Gerwick for their scientific contributions, J. Nunnery for initial purification work, E. Theodorakis for use of the microwave reactor, and the UCSD mass spectrometry facility for their analytical services. The 500 MHz NMR  $^{13}\text{C}$  XSens cold probe was supported by NSF CHE-0741968. We also acknowledge The Growth Regulation & Oncogenesis Training Grant NIH/NCI (T32A009523-24) for a fellowship to E. M., and NIH Grants (CA100851 and GM107550) for support of the research.

## 4.6 Appendix

**Figure 4.6.1:**  $^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide C**Figure 4.6.2:**  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide C



**Figure 4.6.3:** COSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide C



**Figure 4.6.4:** TOCSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide C

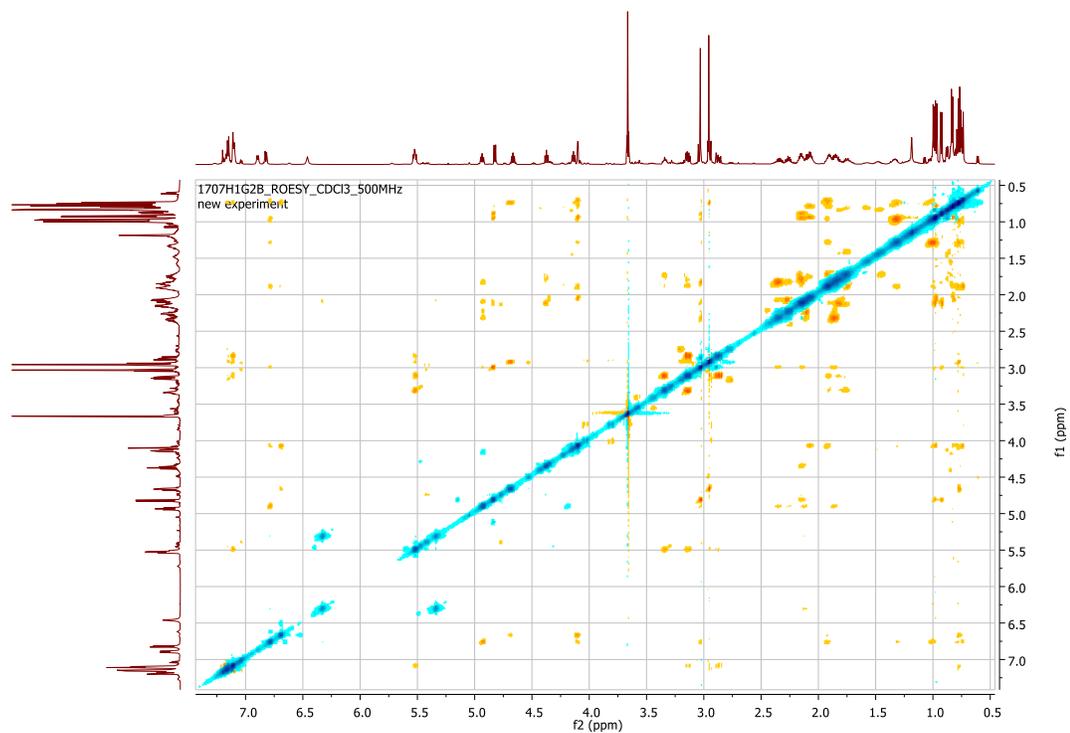


Figure 4.6.5: ROESY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide C

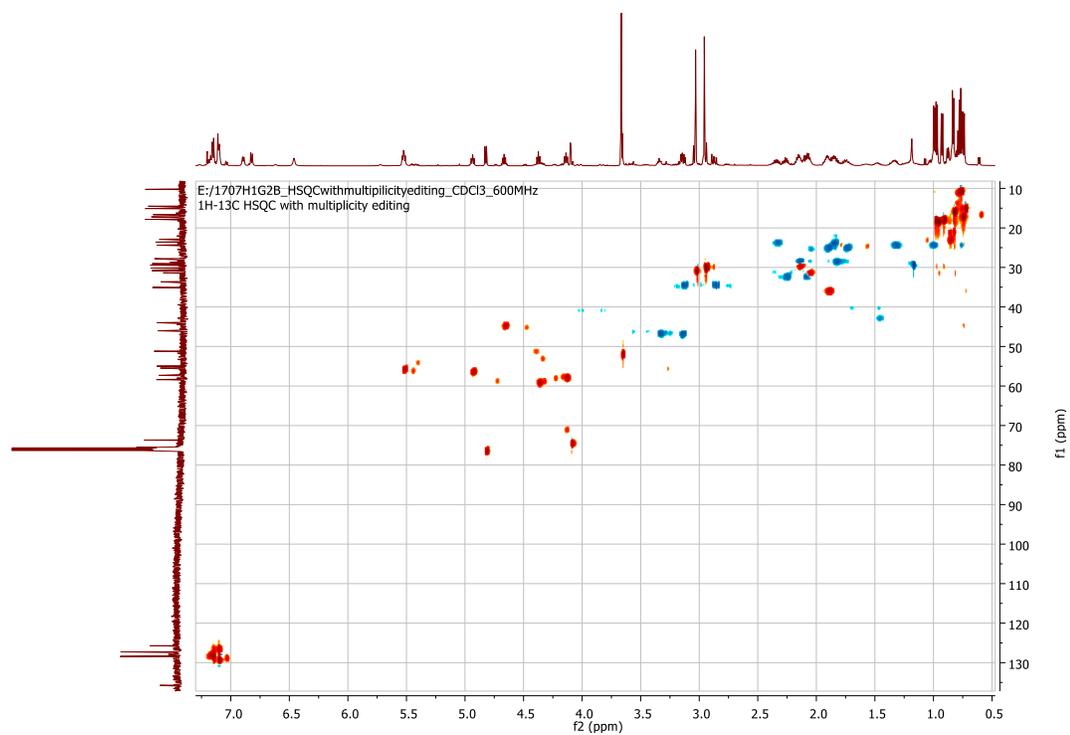
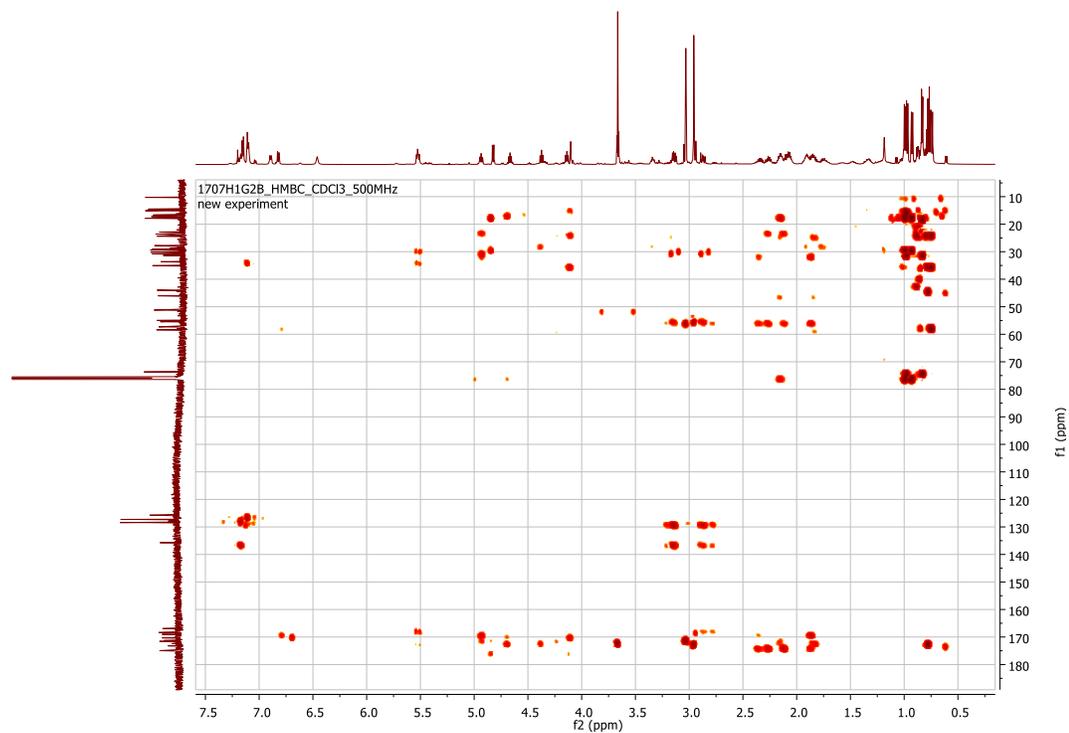
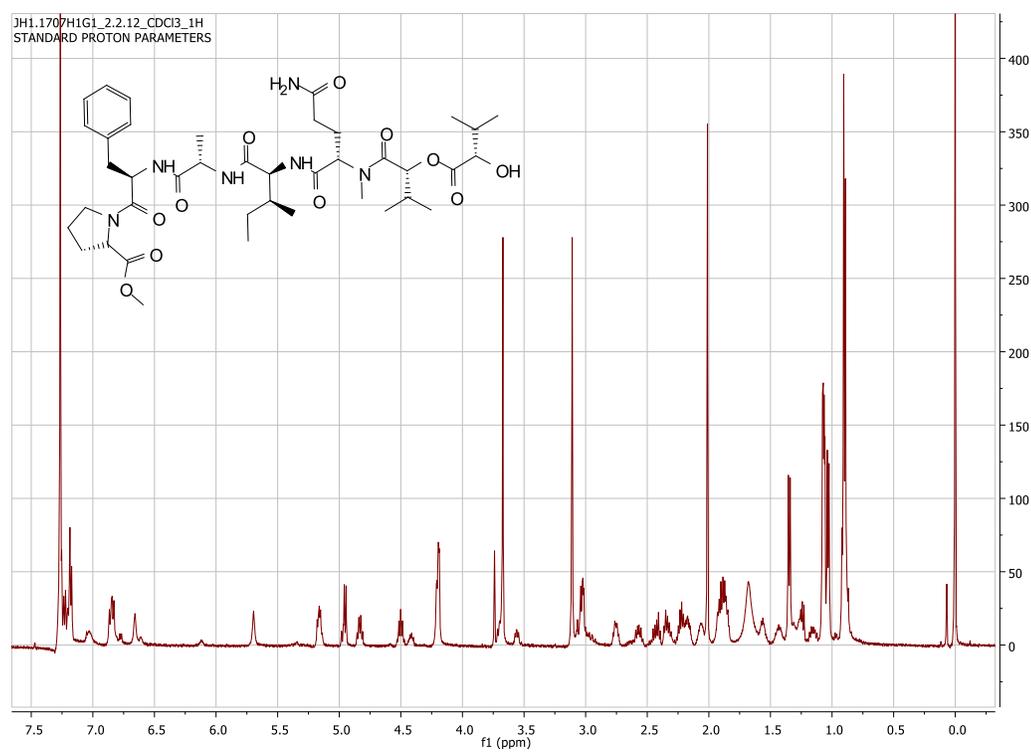


Figure 4.6.6: HSQC (<sup>1</sup>H 600 MHz, CDCl<sub>3</sub>) spectrum of tasiamide C



**Figure 4.6.7:** HMBC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide C



**Figure 4.6.8:**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide D

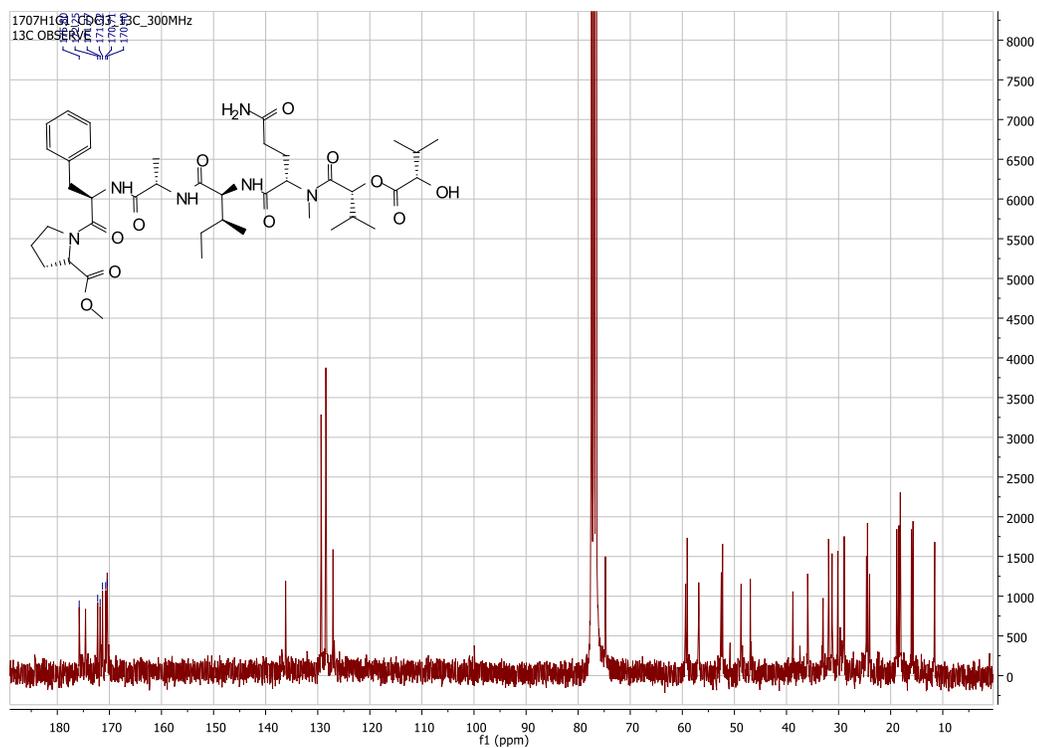


Figure 4.6.9: <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) spectrum of tasiamide D

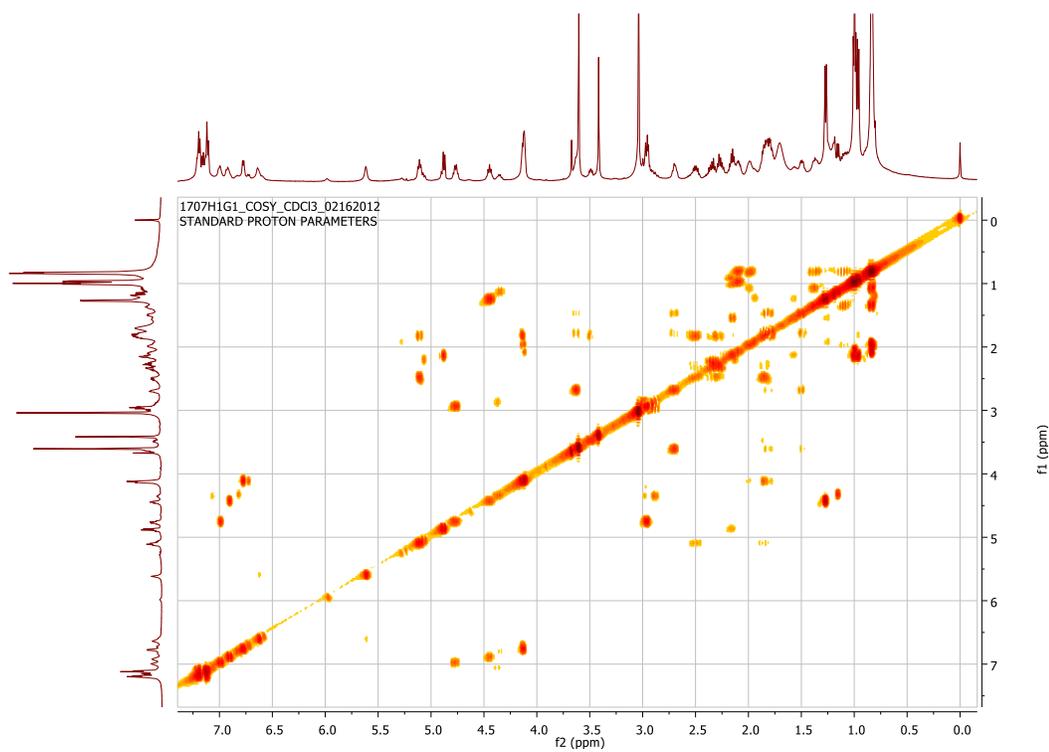
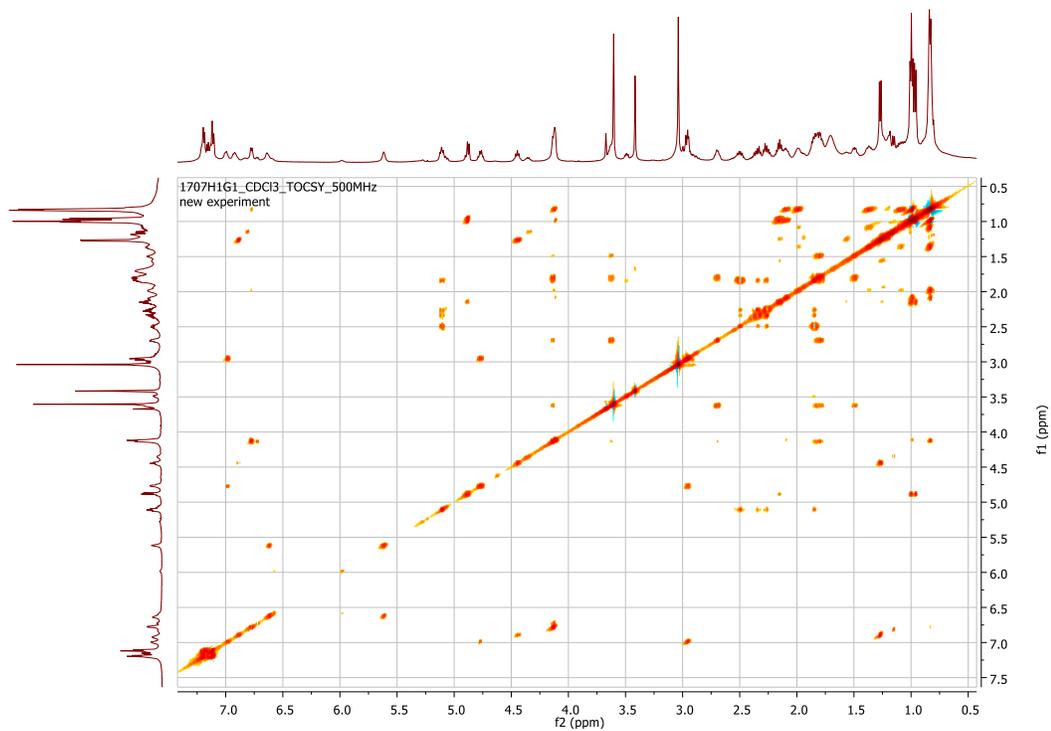
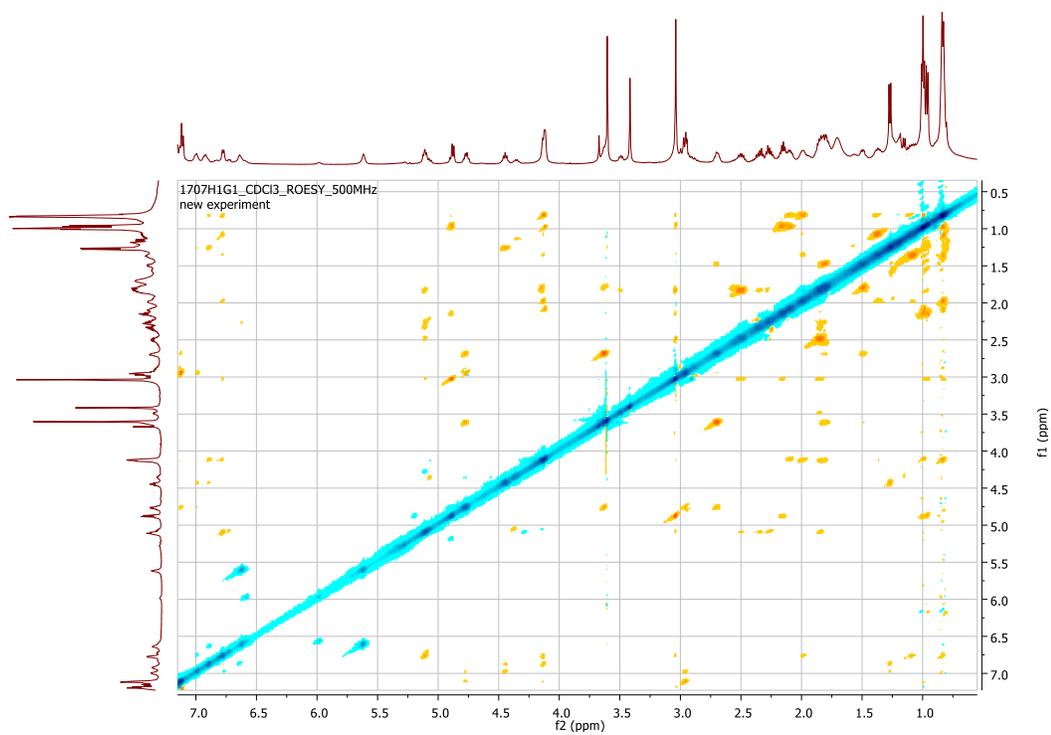


Figure 4.6.10: COSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide D



**Figure 4.6.11:** TOCSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide D



**Figure 4.6.12:** ROESY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide D

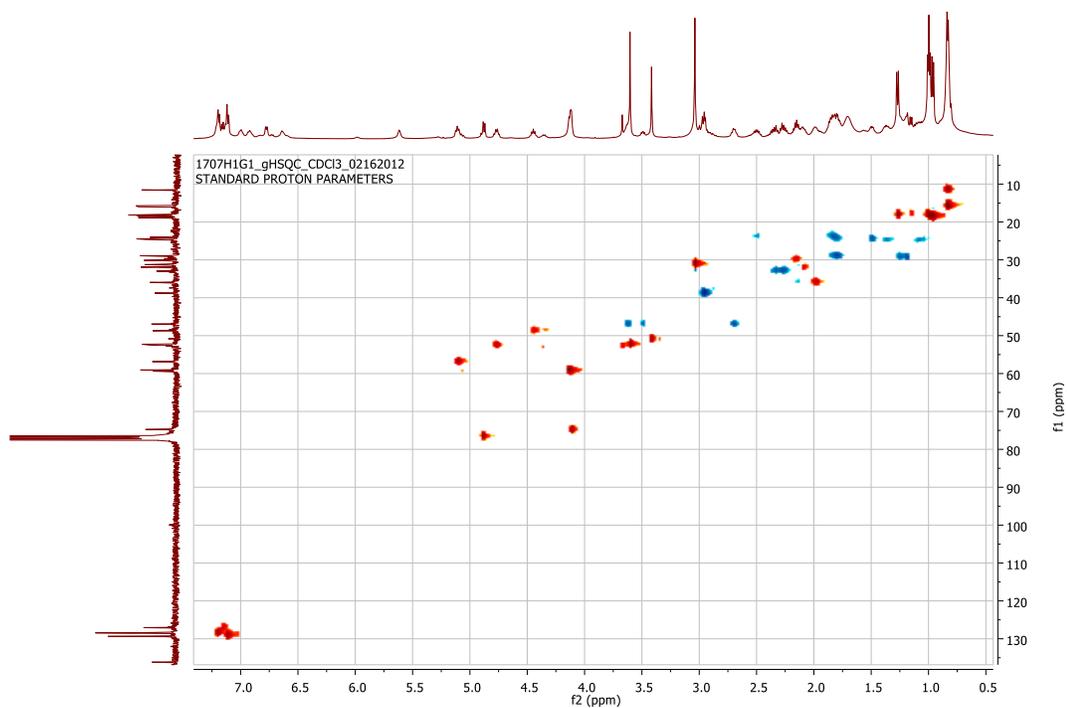


Figure 4.6.13: HSQC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide D

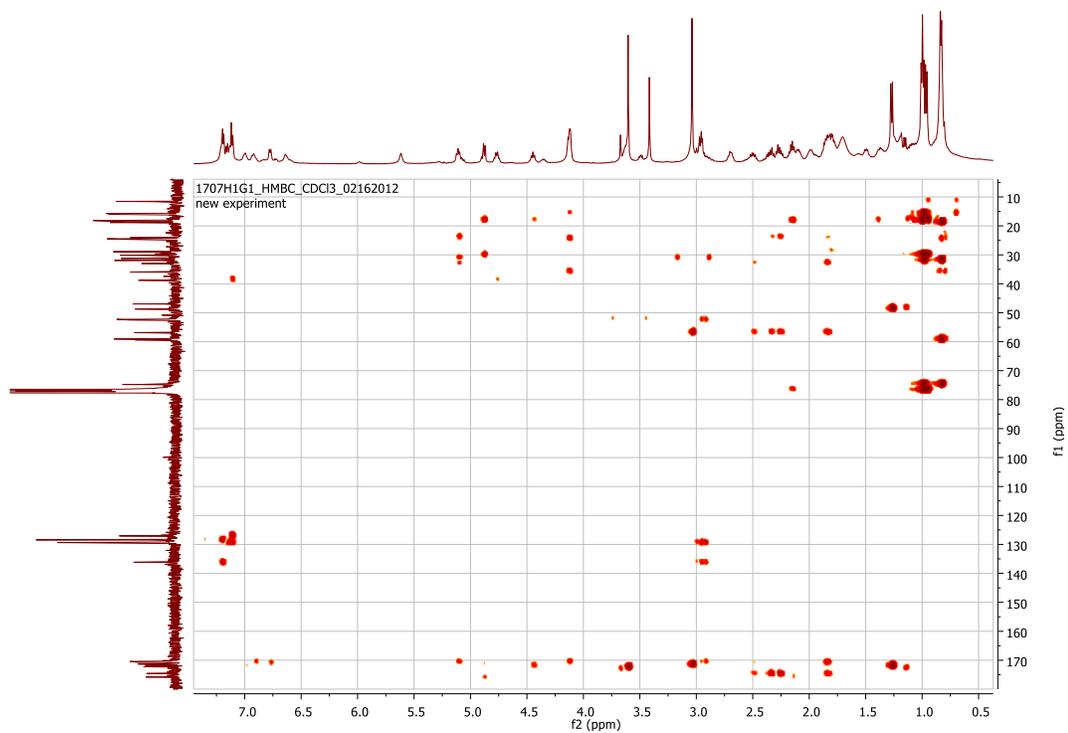


Figure 4.6.14: HMBC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide D

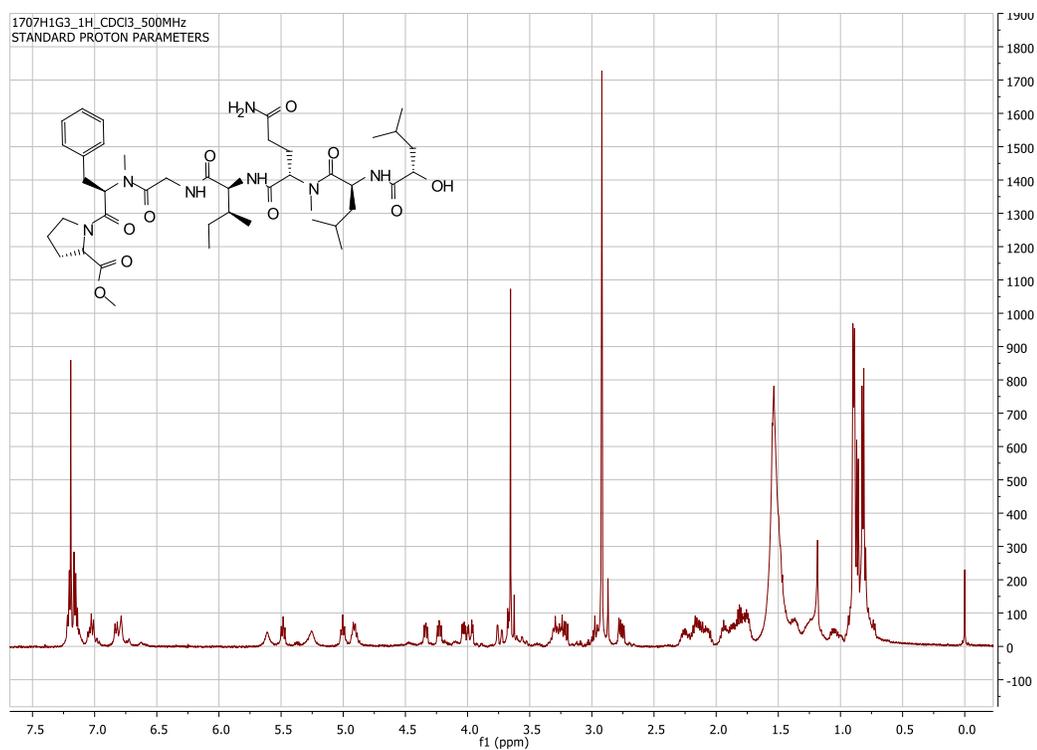


Figure 4.6.15: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E

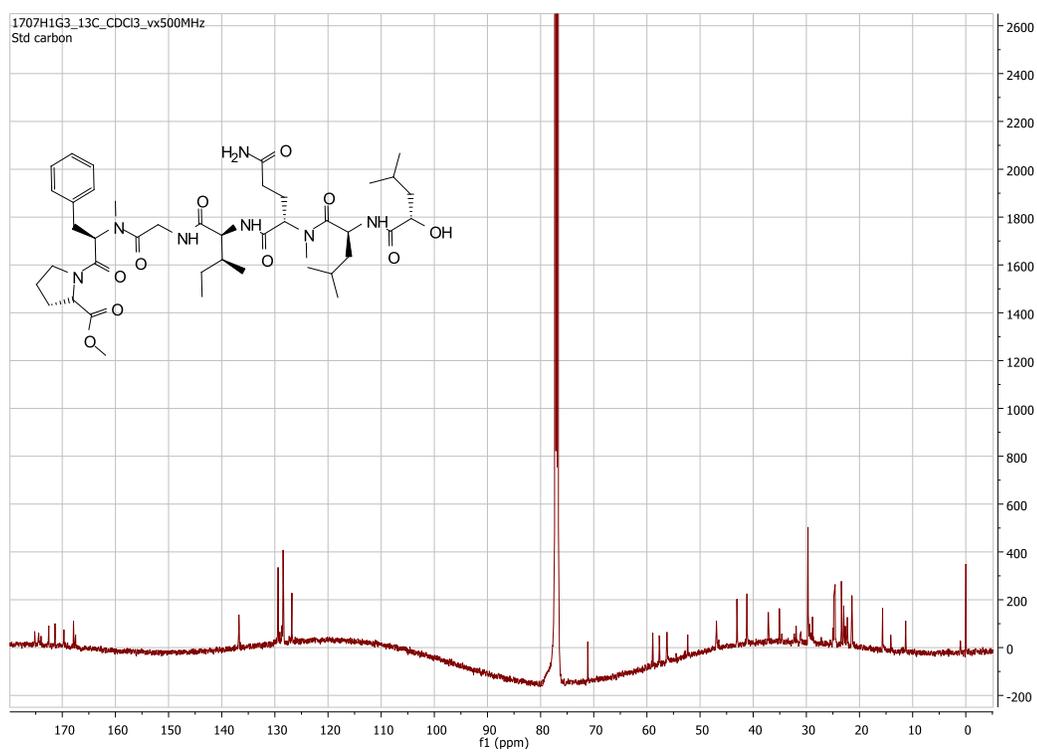
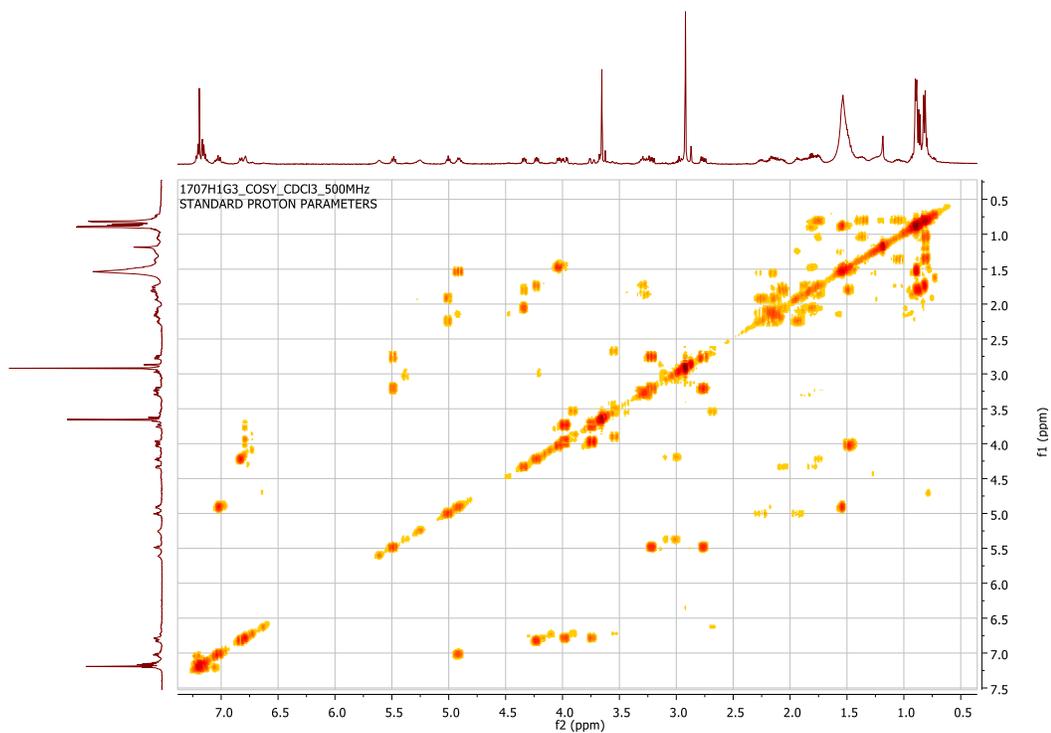
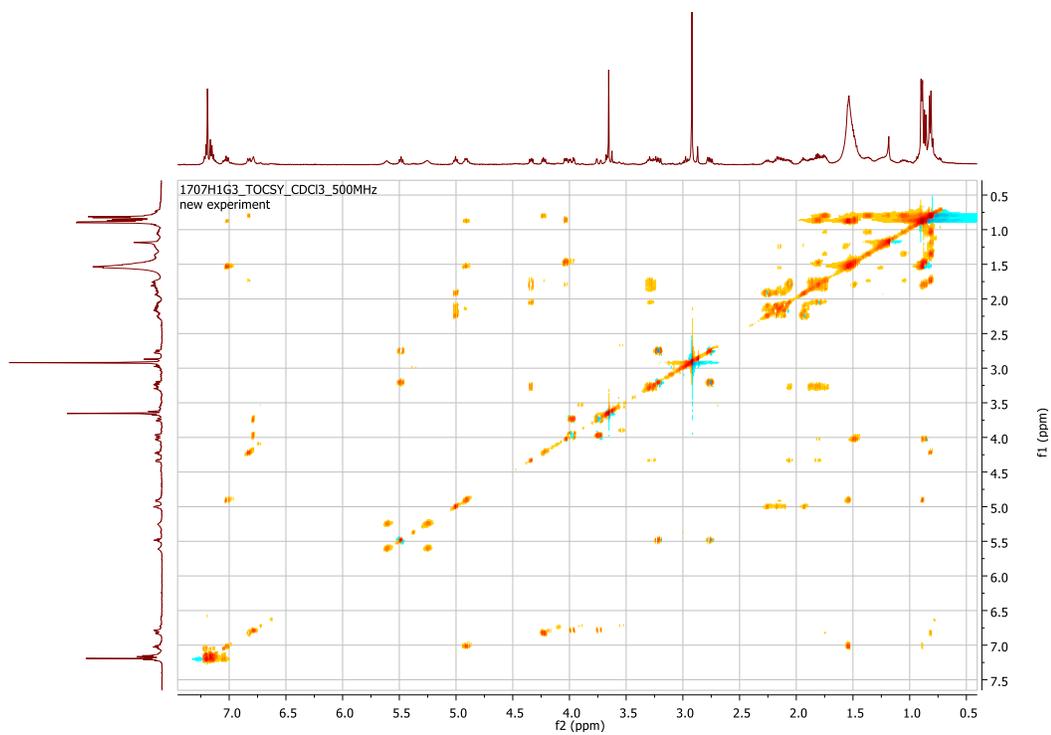


Figure 4.6.16: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E



**Figure 4.6.17:** COSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E



**Figure 4.6.18:** TOCSY (500 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E

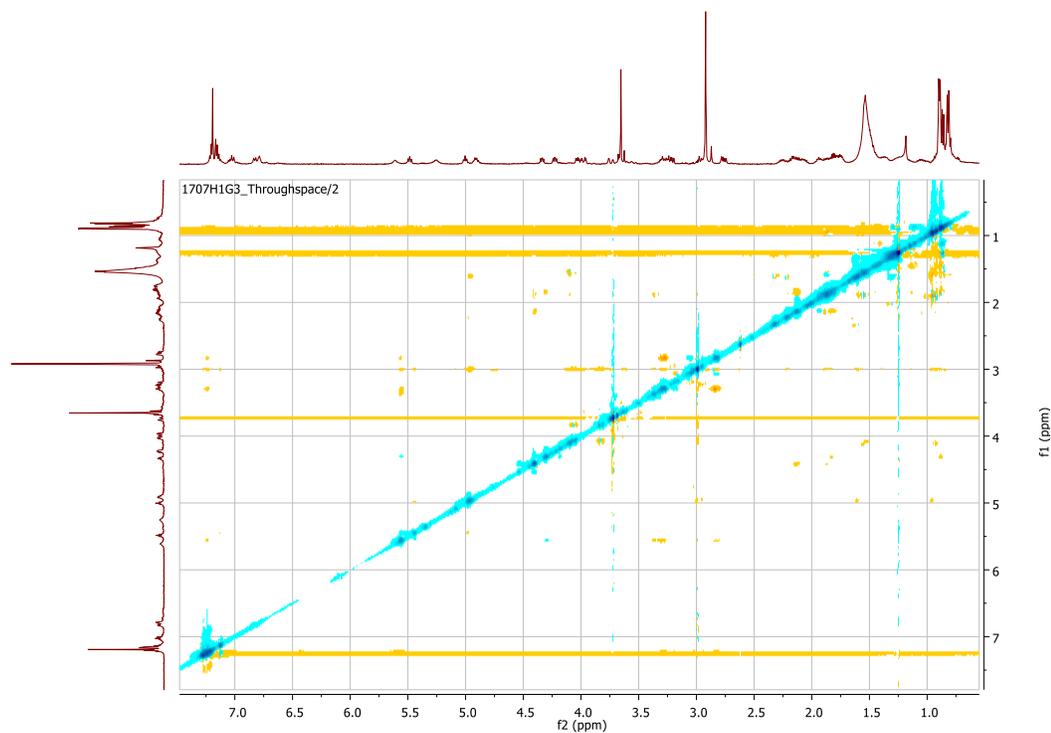


Figure 4.6.19: ROESY (600 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E

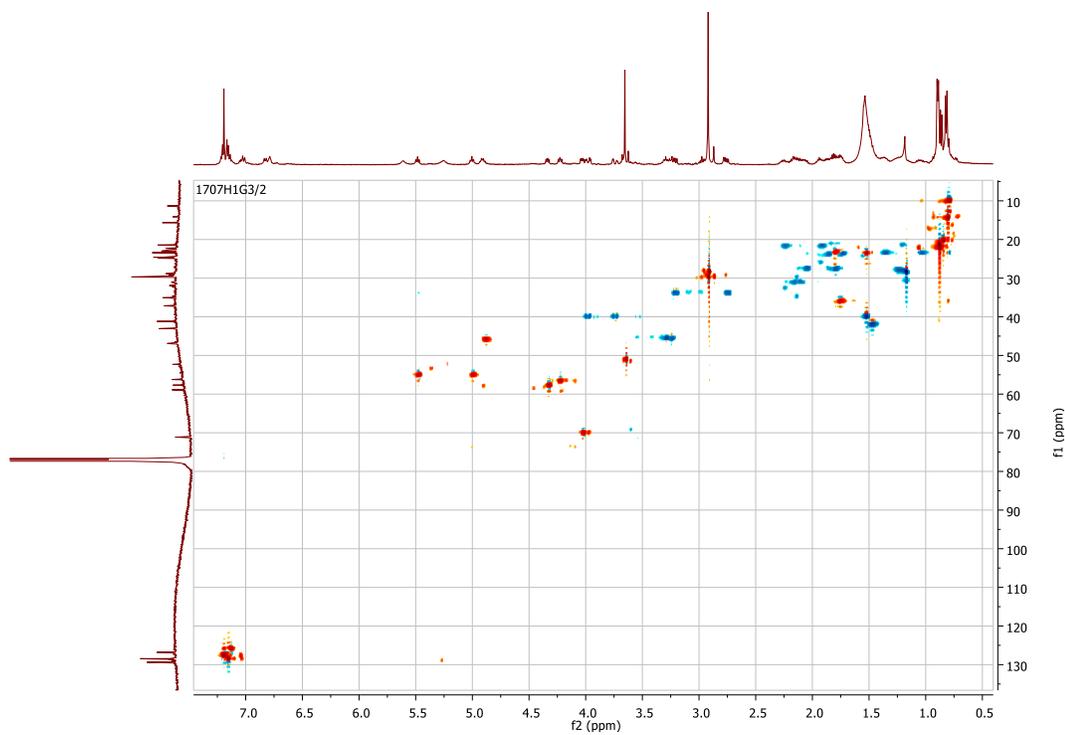


Figure 4.6.20: HSQC (<sup>1</sup>H 600 MHz, CDCl<sub>3</sub>) spectrum of tasiamide E

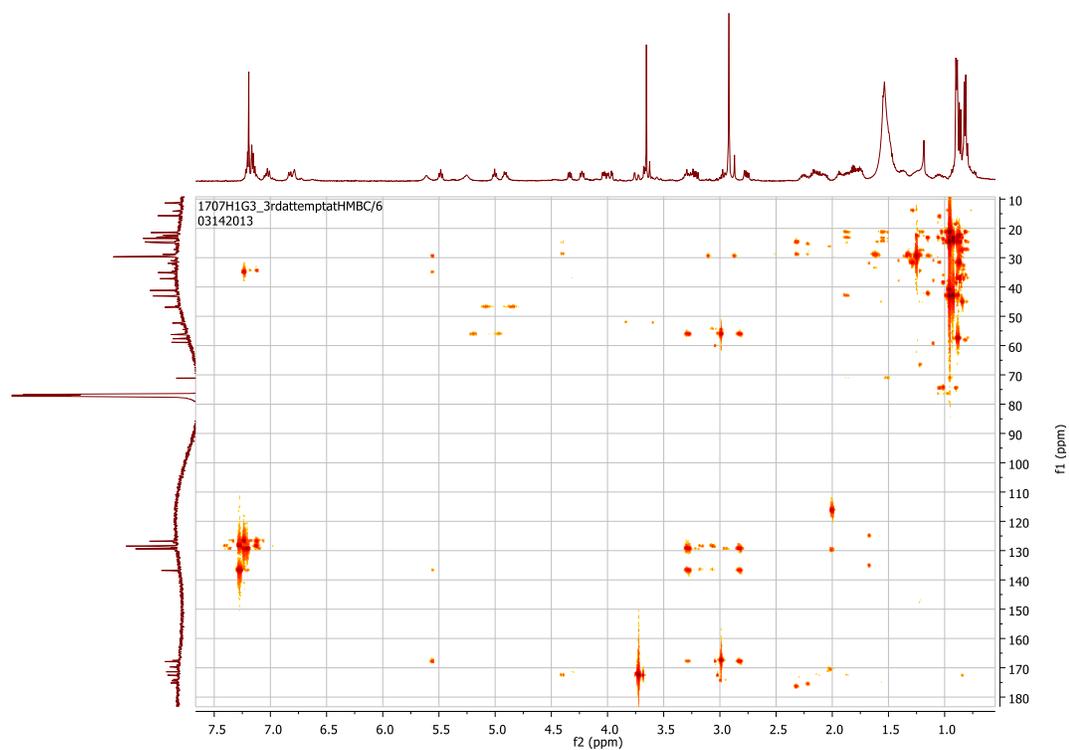


Figure 4.6.21: HMBC ( $^1\text{H}$  600 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide E

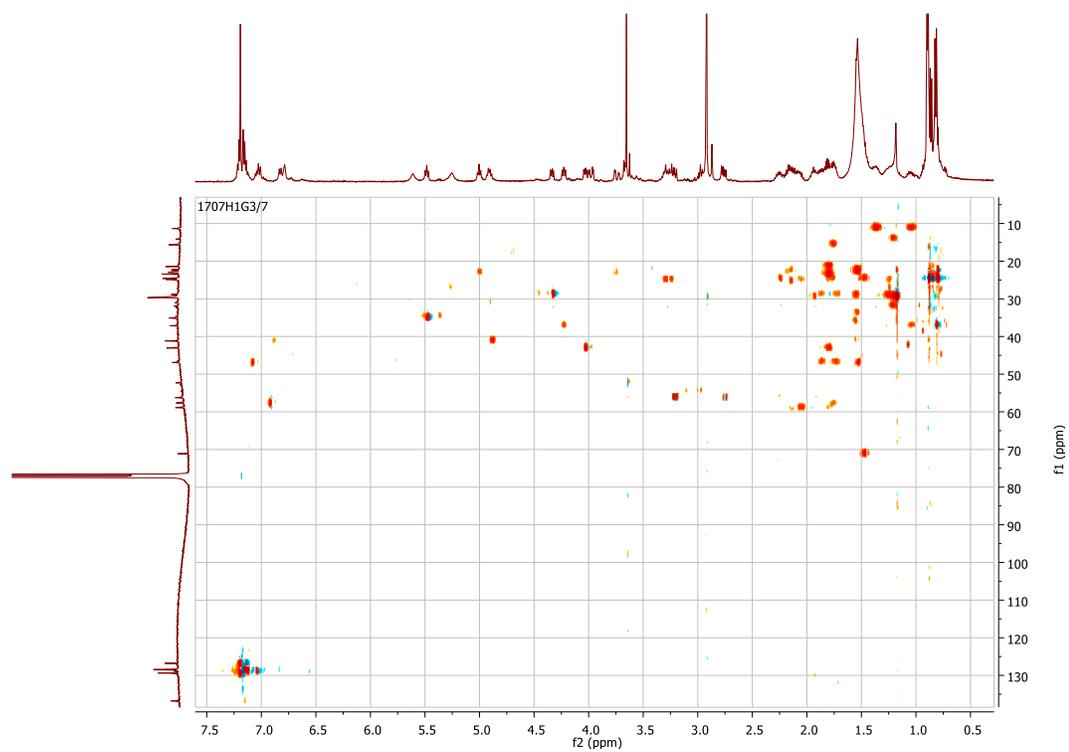
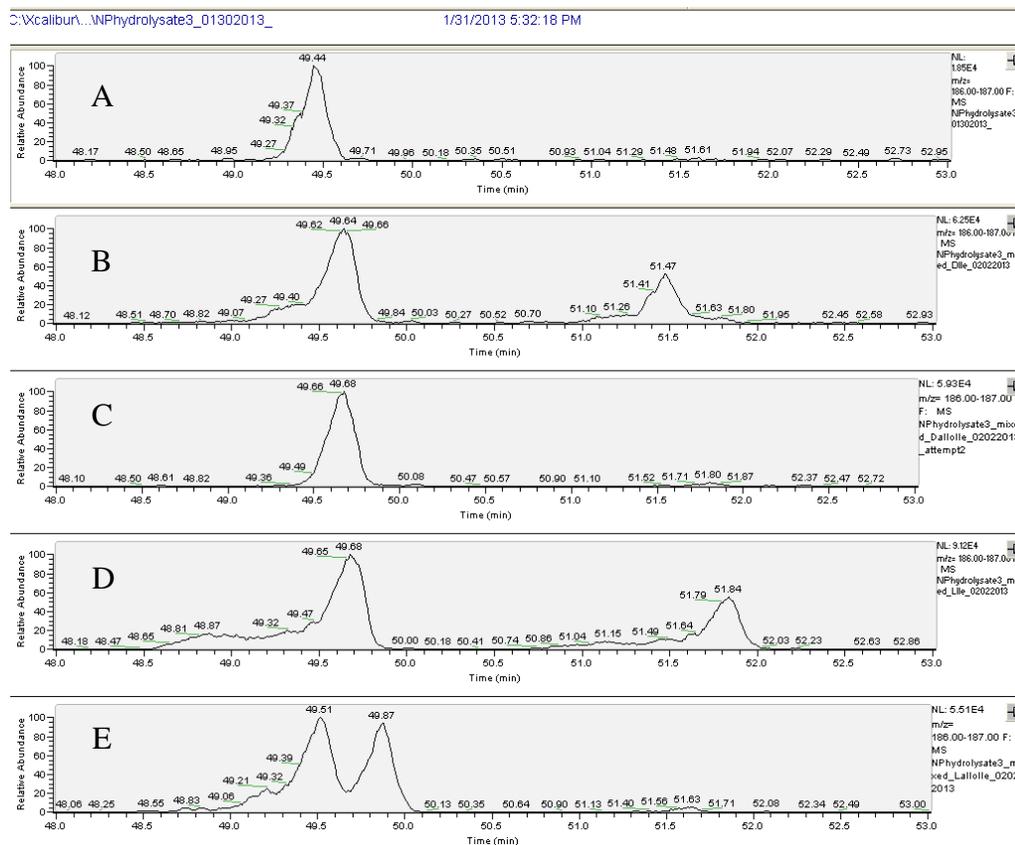
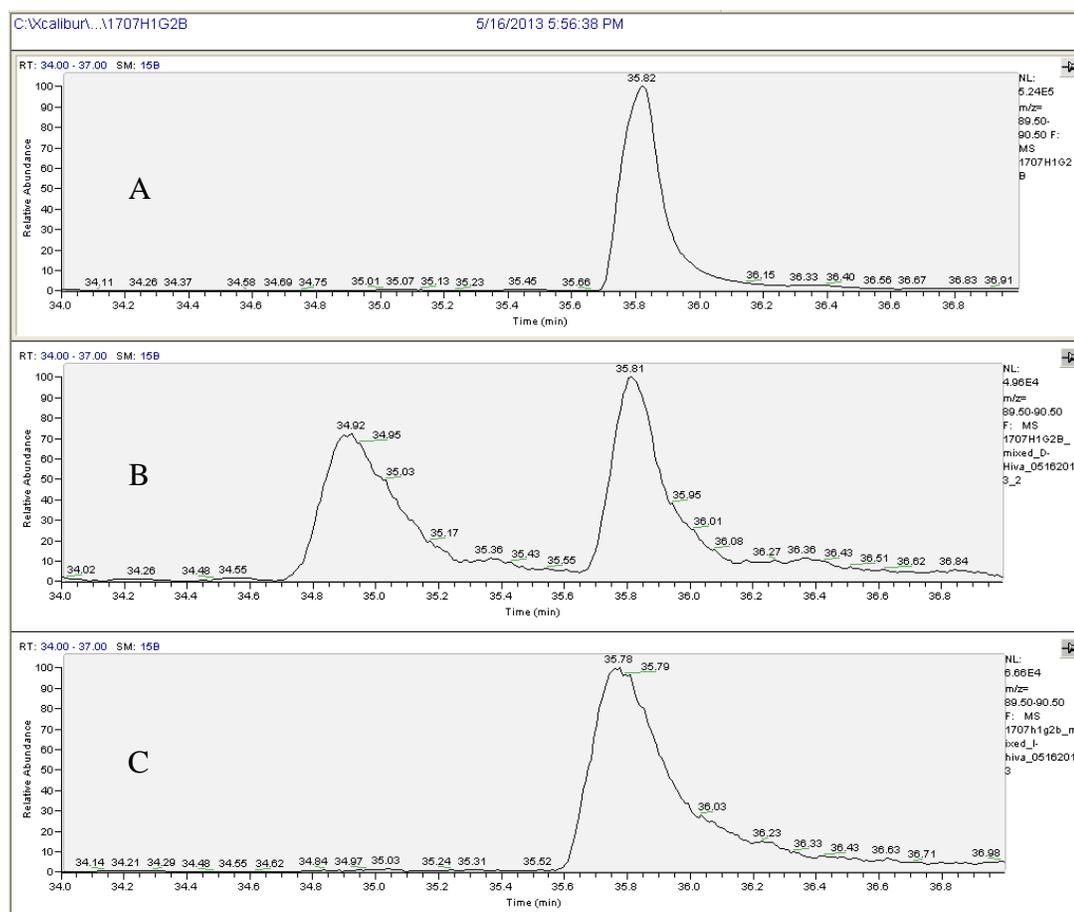


Figure 4.6.22: H2BC ( $^1\text{H}$  600 MHz,  $\text{CDCl}_3$ ) spectrum of tasiamide E

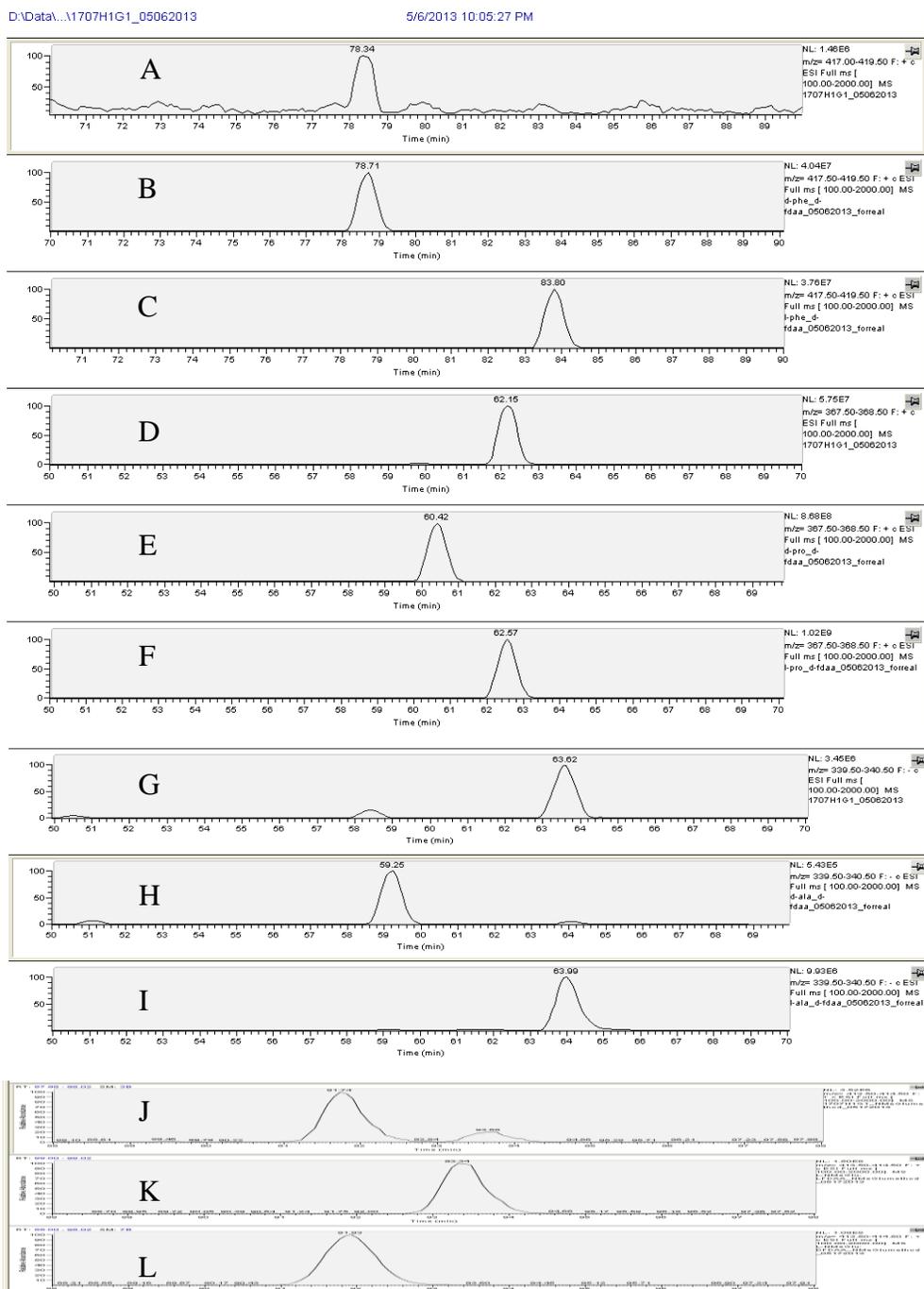




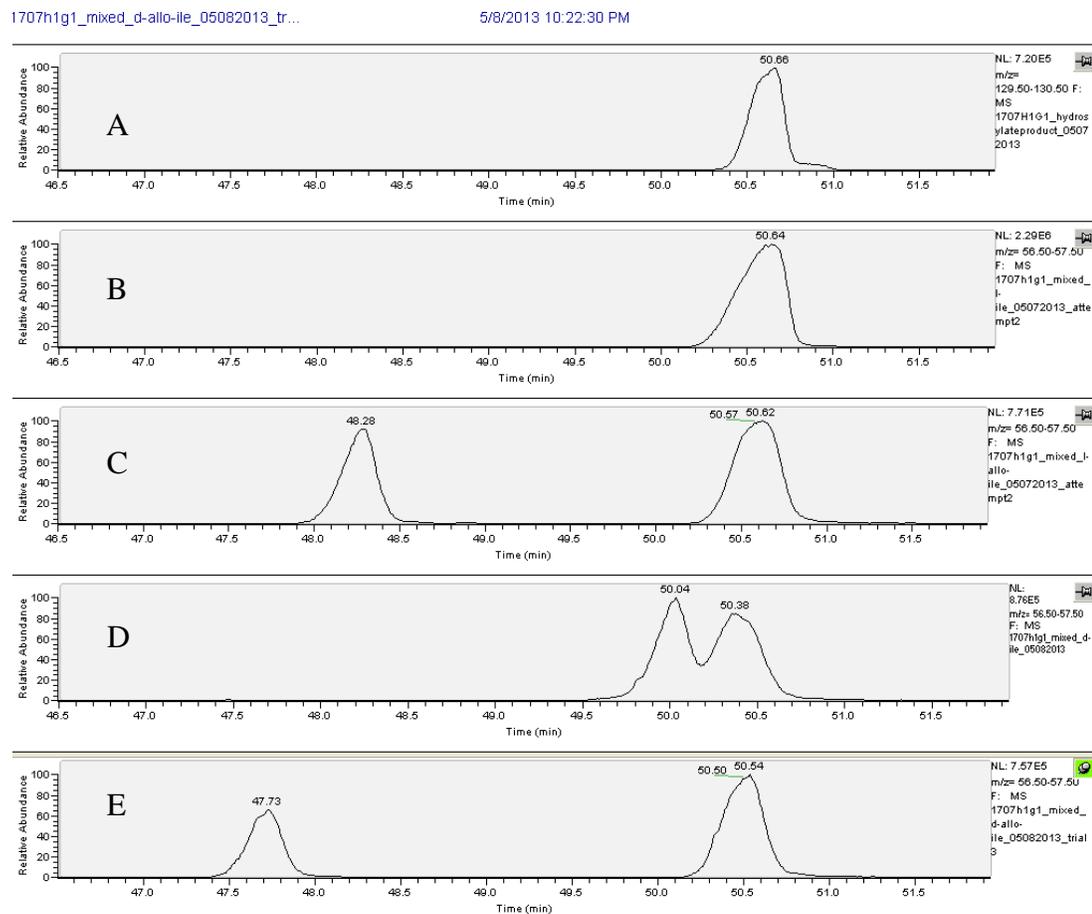
**Figure 4.6.24:** Chiral GCMS analysis of isoleucine in tasiamide C (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with D-Ile; (C) Derivatized hydrolysis product co-injected with D-allo-Ile; (D) Derivatized hydrolysis product co-injected with L-Ile; (E) Derivatized hydrolysis product co-injected with L-allo-Ile



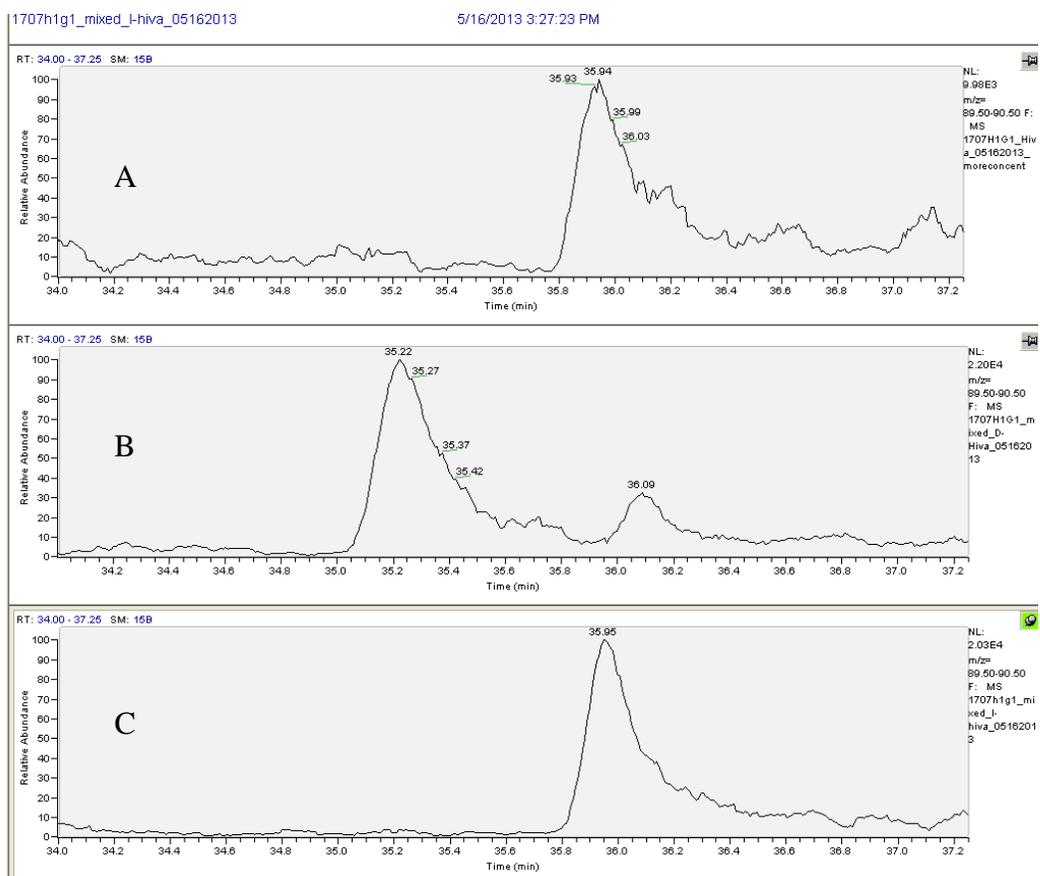
**Figure 4.6.25:** Chiral GCMS analysis of the terminal Hmba in tasiamide C (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with *R*-Hmba; (C) Derivatized hydrolysis product co-injected with *S*-Hmba



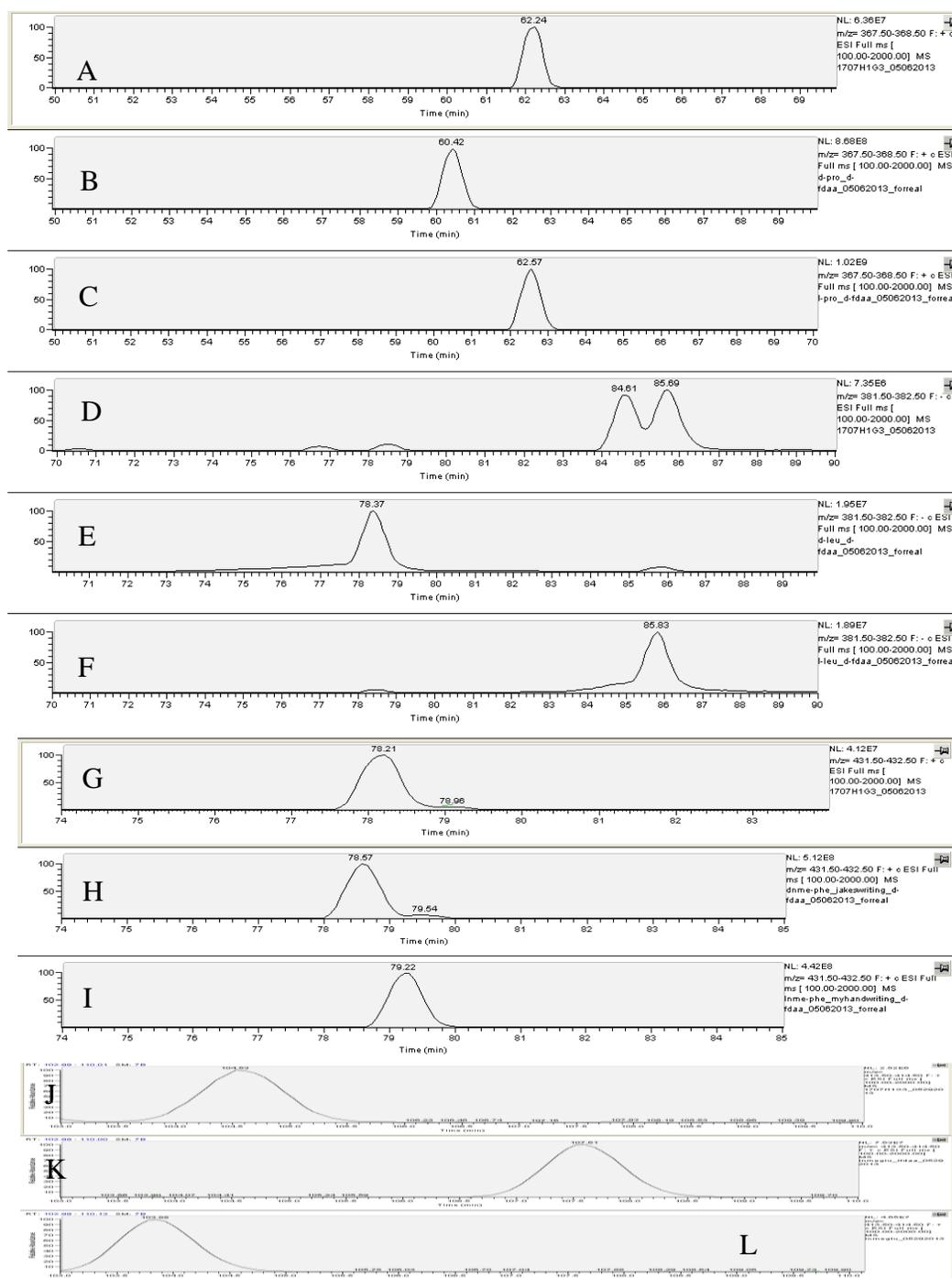
**Figure 4.6.26:** Marfey's analysis of the amino acids in tasiamide D on LCMS: (A) Marfey's derivatized hydrolysis product (ion chromatogram of 417-419 m/z); (B) D-Phe; (C) L-Phe; (D) Marfey's derivatized hydrolysis product (ion chromatogram of 367.5-368.5 m/z); (E) D-Pro; (F) L-Pro; (G) Marfey's derivatized hydrolysis product (ion chromatogram of 339.5-340.5); (H) D-Ala; (I) L-Ala; (J) Marfey's derivatized hydrolysis product (ion chromatogram of 413.5-414.5 m/z); (K) D-Glu; (L) L-Glu



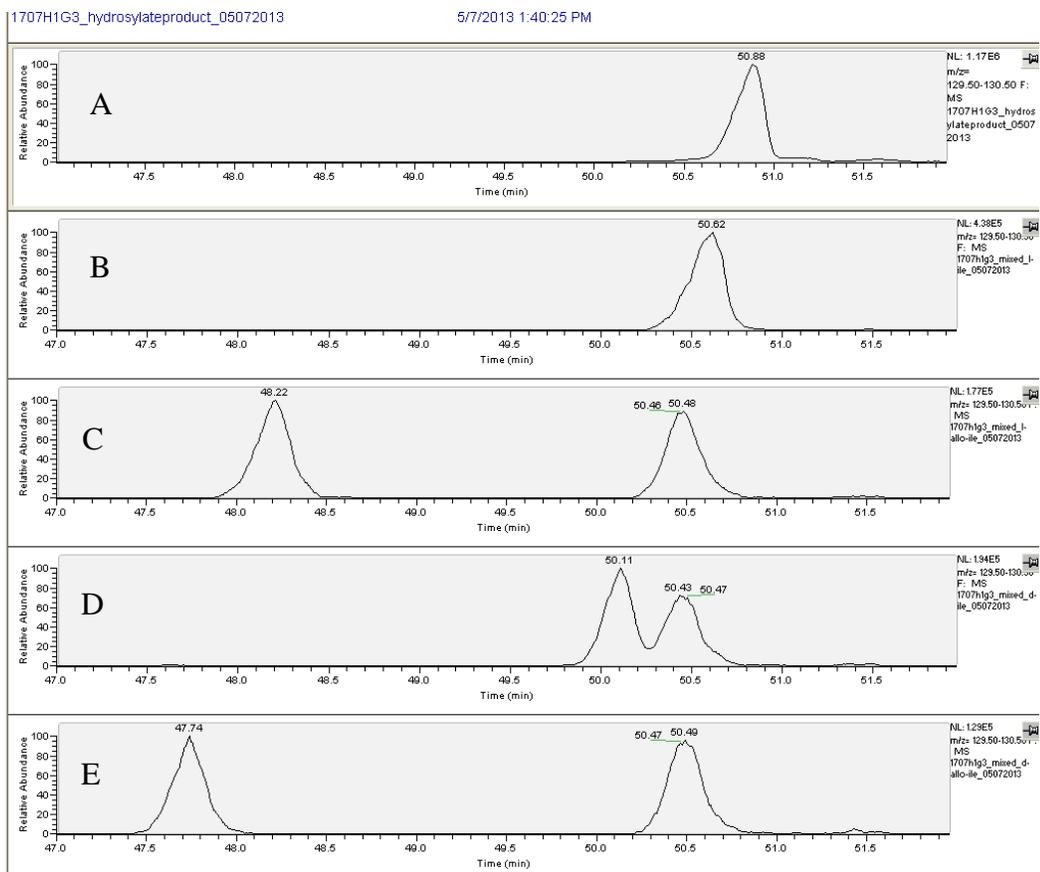
**Figure 4.6.27:** Chiral GCMS analysis of isoleucine in tasiamide D (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with L-Ile; (C) Derivatized hydrolysis product co-injected with L-*allo*-Ile; (D) Derivatized hydrolysis product co-injected with D-Ile; (E) Derivatized hydrolysis product co-injected with D-*allo*-Ile



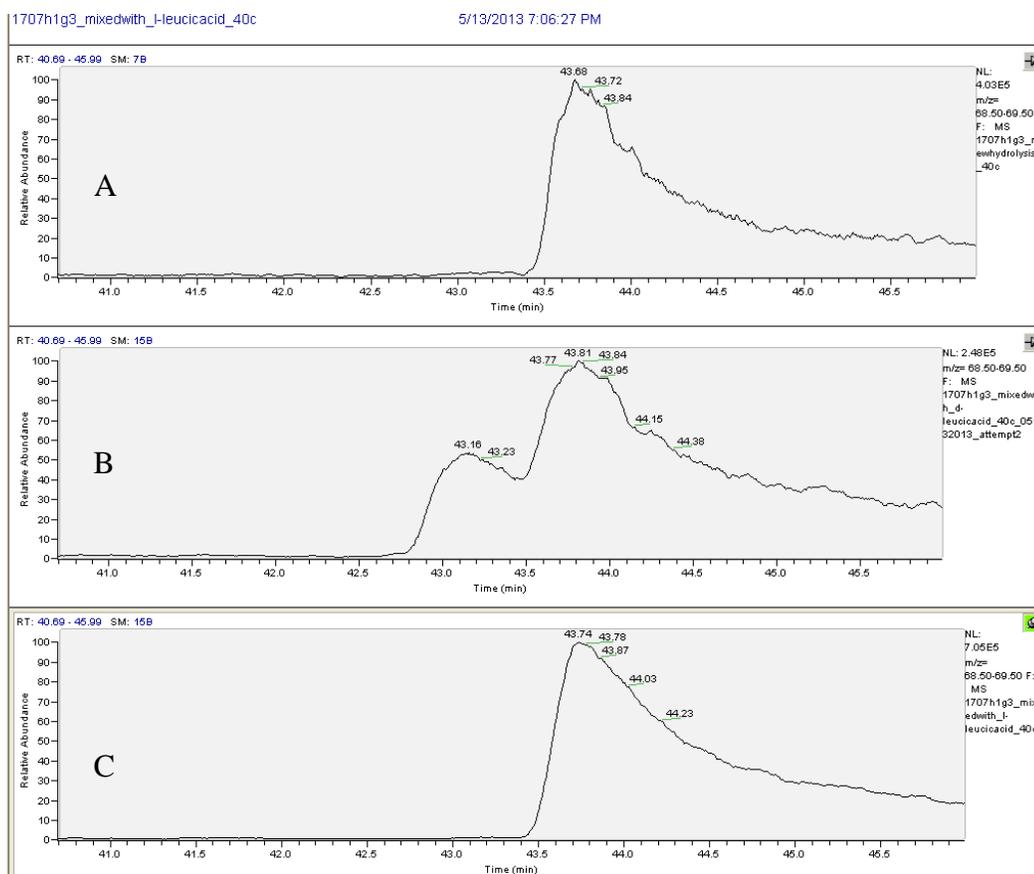
**Figure 4.6.28:** Chiral GCMS analysis of the terminal Hmba in tasiamide D (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with *R*-Hmba; (C) Derivatized hydrolysis product co-injected with *S*-Hmba



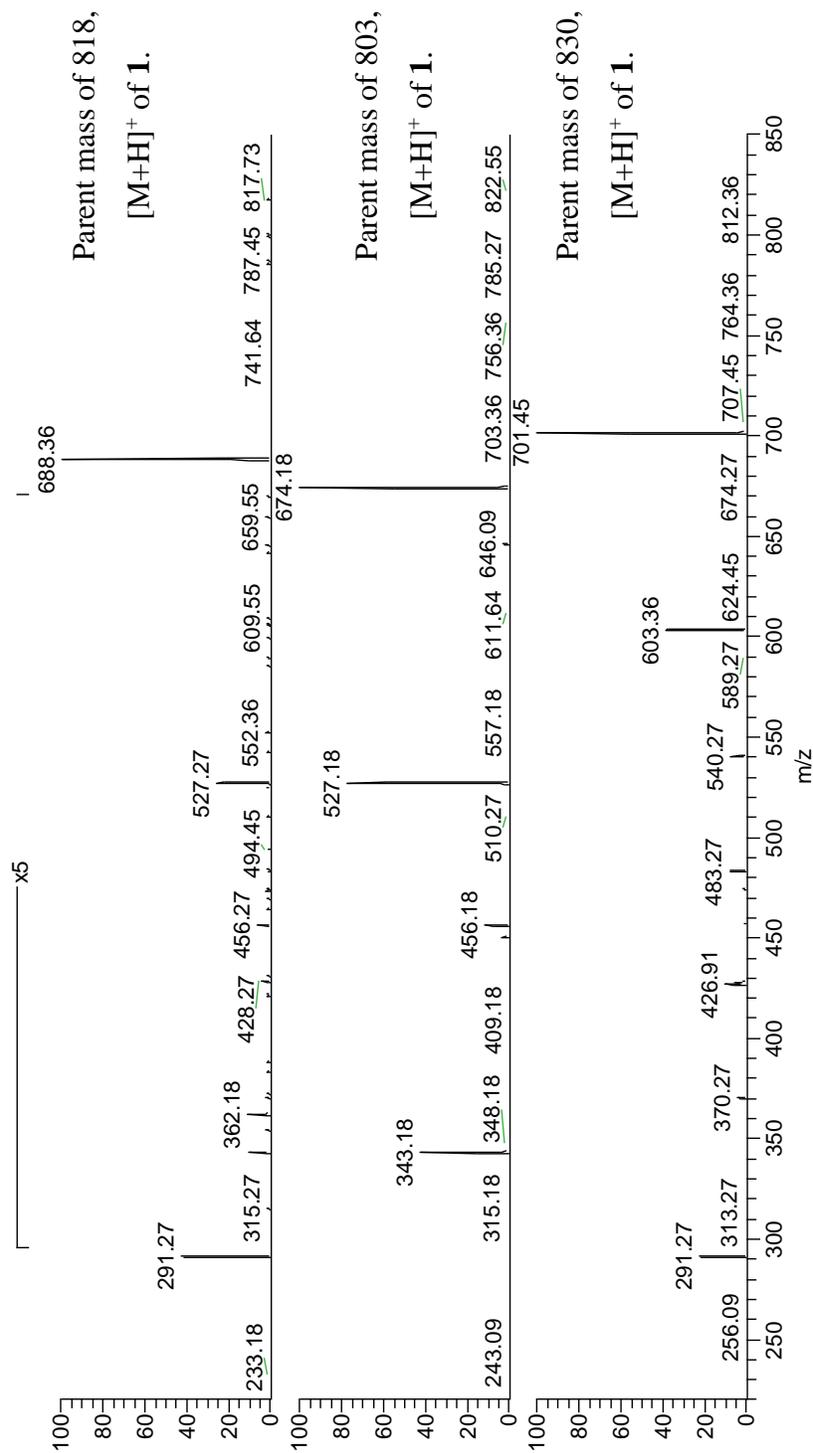
**Figure 4.6.29:** Marfey's analysis of the amino acids in tasiamide E on LCMS: (A) Marfey's derivatized hydrolysis product (ion chromatogram of 367.5-368.5 m/z); (B) D-Pro; (C) L-Pro; (D) Marfey's derivatized hydrolysis product (ion chromatogram of 381.5-382.5 m/z); (E) D-Leu; (F) L-Leu; (G) Marfey's derivatized hydrolysis product (ion chromatogram of 431.5-432.5); (H) D-NMePhe; (I) L-NMePhe; (J) Marfey's derivatized hydrolysis product (ion chromatogram of 413.5-414.5 m/z); (K) D-NMeGlu; (L) L-NMeGlu

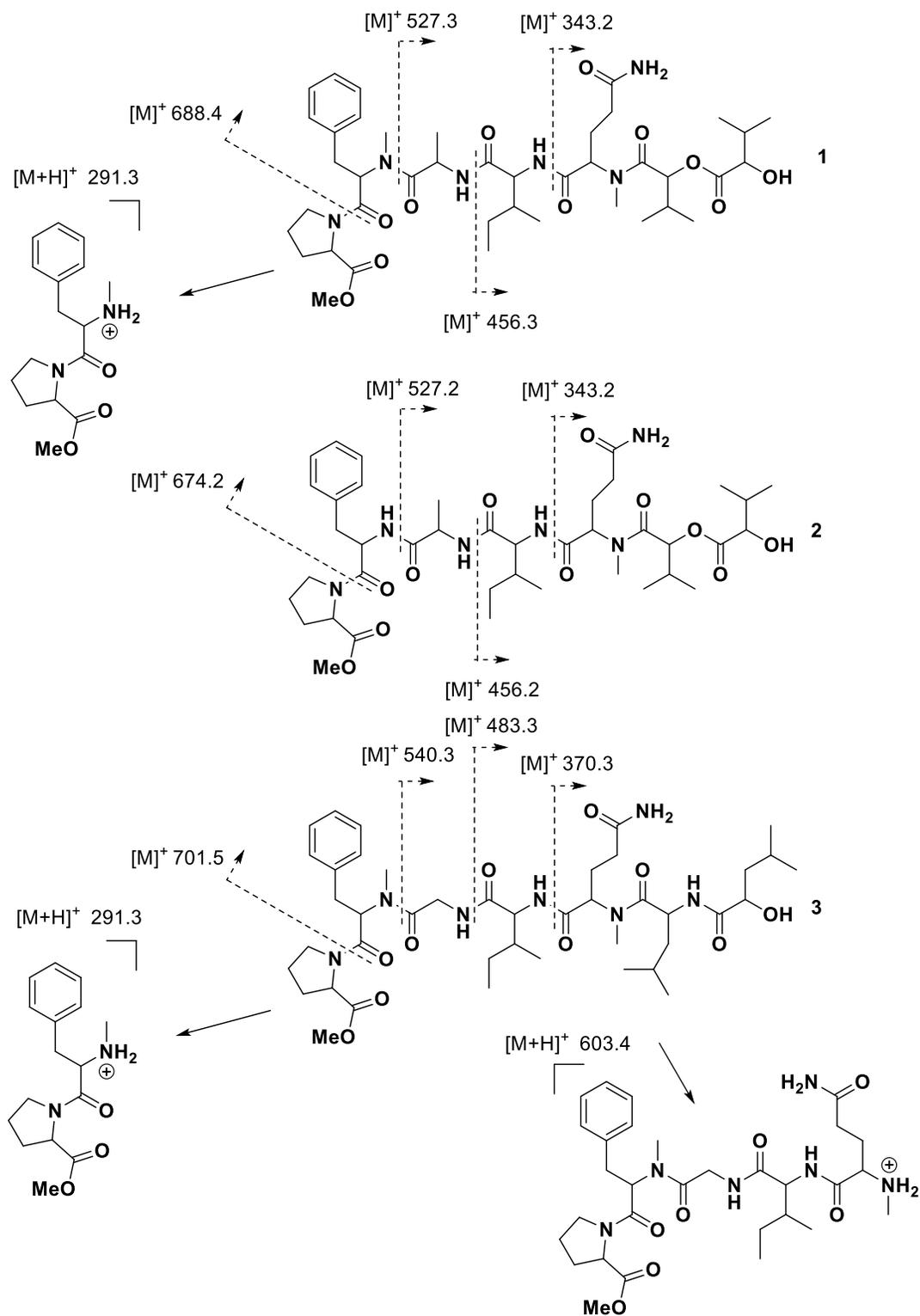


**Figure 4.6.30:** Chiral GCMS analysis of isoleucine in tasiamide E (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with L-Ile; (C) Derivatized hydrolysis product co-injected with L-*allo*-Ile; (D) Derivatized hydrolysis product co-injected with D-Ile; (E) Derivatized hydrolysis product co-injected with D-*allo*-Ile

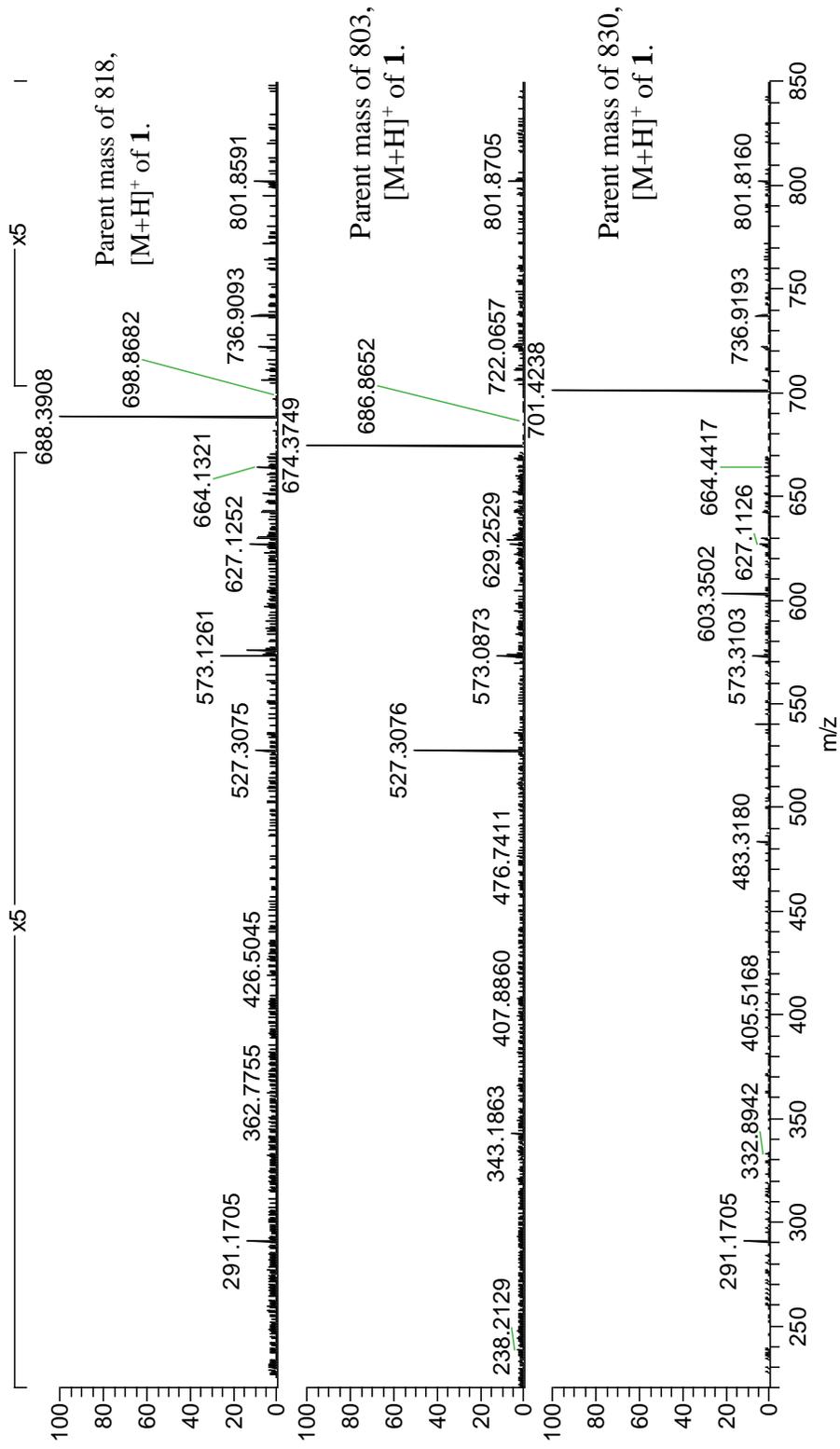


**Figure 4.6.31:** Chiral GCMS analysis of isoleucine in tasiamide E (A) Derivatized hydrolysis product; (B) Derivatized hydrolysis product co-injected with *R*-H4mpa; (C) Derivatized hydrolysis product co-injected with *S*-H4mpa

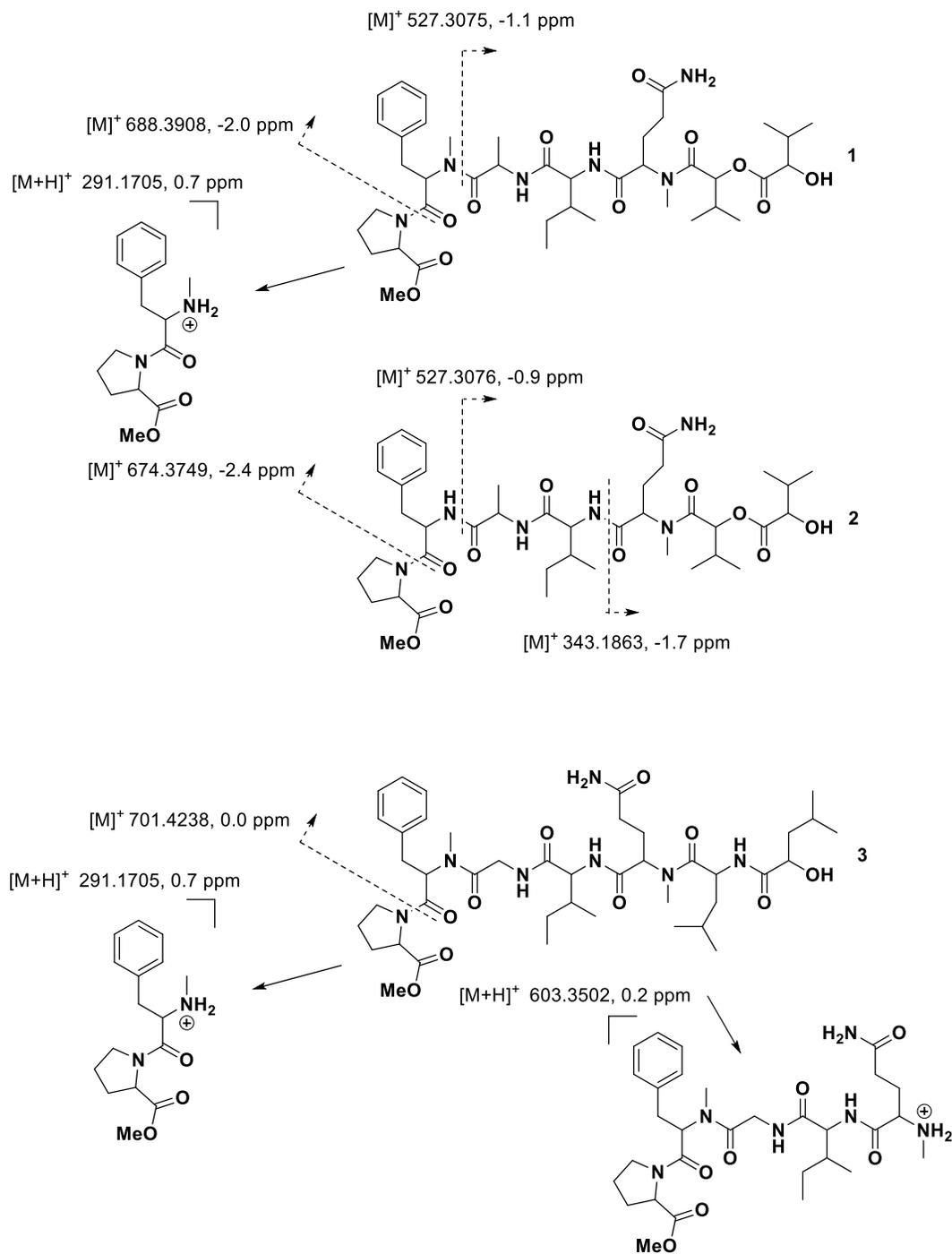
**Figure 4.6.32:** LTQ Ion-Trap MS<sup>2</sup> spectra of tasiamides C-E



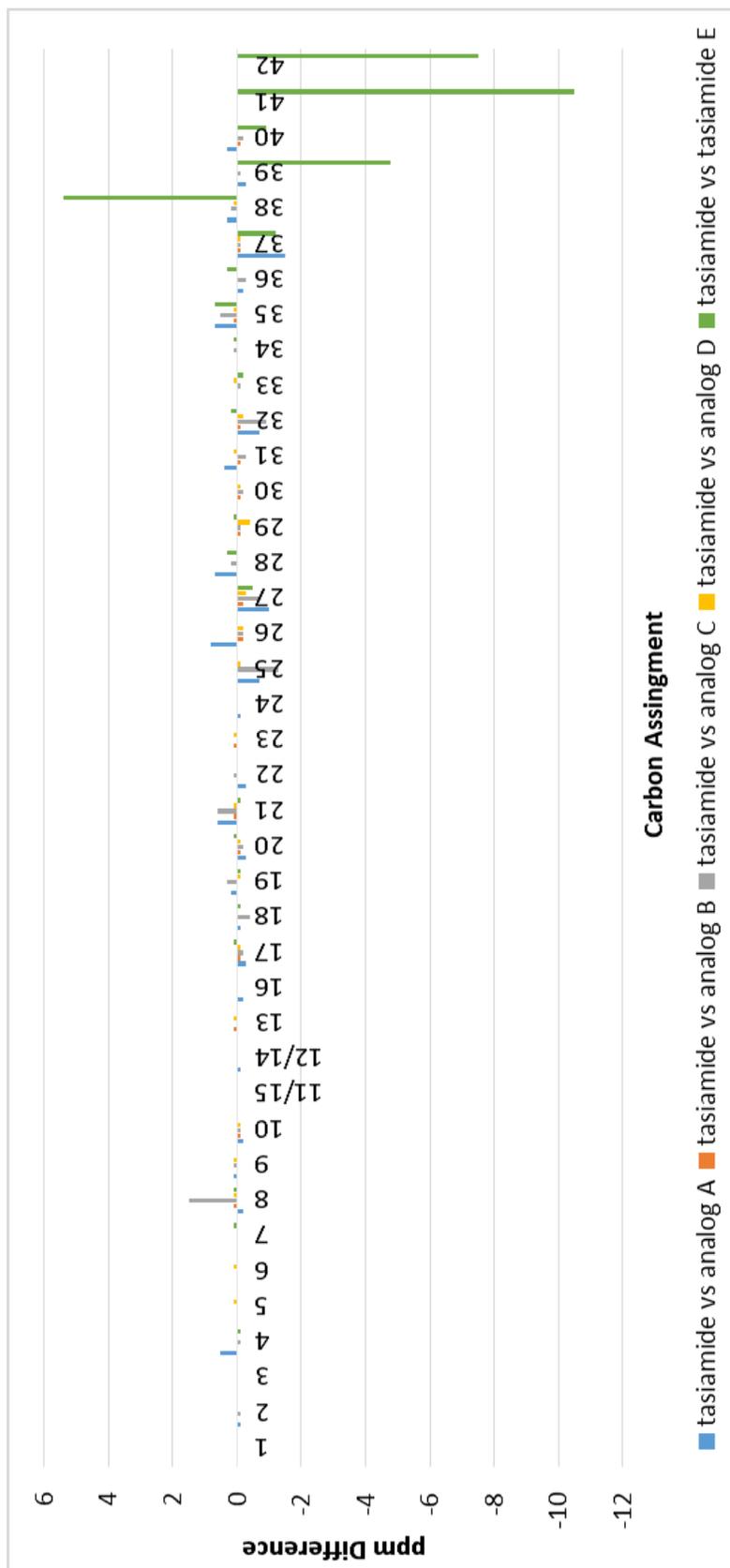
**Figure 4.6.33:** Comparison of Ion-Trap MS<sup>2</sup> fragments for tasiamides C-E



**Figure 4.6.34:** FT-MS MS<sup>2</sup> spectra of tasiamides C-E



**Figure 4.6.35:** Comparison of FT-MS MS<sup>2</sup> fragments of tasiamides C-E



**Figure 4.6.36:** Carbon comparison of tasiamide, tasiamide E and three synthetic stereoisomers

Please select version of database to use:

DP4-original  
**DP4-database2**

Select probability distribution:

t distribution (recommended)  
 normal distribution

**13C Calc:**

C1,C2,C3,C4,C5,C6,C7,C8,C9,C10,C11,C12,C13,  
172.5,59.0,28.8,24.4,46.8,52.3,167.9,56.5,35.0,137  
172.5,58.9,28.8,24.9,46.8,52.3,167.9,56.2,35.1,136  
172.5,59.0,28.8,25.0,46.8,52.3,167.9,54.8,35.0,136  
172.5,58.9,28.8,24.9,46.7,52.2,167.9,56.2,35.0,136

**1H Calc:**

**13C Expt:**

172.5(C1), 58.9(C2), 28.8(C3), 24.9(C4), 46.8(C5),

**1H Expt:**

Read Data Show Assignments Calculate Clear

This calculation will use the DP4-database2 version of the database and the t distribution.  
(To change these options select the desired database and distribution from the menus at the top of the applet and then click Calculate).

Results of DP4 using the carbon data only:

Isomer 1: 1.3%  
Isomer 2: 56.0%  
Isomer 3: 3.0%  
Isomer 4: 39.7%

(c) Jonathan M Goodman and Steven G Smitt

**Figure 4.6.37:** DP4 calculation results comparing the four synthetic analogs to tasiamide

## 4.7 References

- (1) Jones, C. A.; Monroe, E. A.; Podell, S.; Hess, W. R.; Klages, S.; Esquenazi, E.; Niessen, S.; Hoover, H.; Rothmann, M.; Lasken, R. S.; Yates III, J. R.; Reinhardt, R.; Kube, M.; Burkart, M. D.; Allen, E. E.; Dorrestein, P. C.; Gerwick, W. H.; Gerwick, L. *PNAS* **2011**, *108*, 8815-8820.
- (2) Kalaitzis, J. A.; Lauro, F. M.; Neilan, B. A. *Nat. Prod. Rep.* **2009**, *26*, 1447-1465.
- (3) Engene, N.; Rottacker, E. C.; Kaštovský, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komárek, J.; Gerwick, W. H. *Int. J. Syst. Evol. Micr.* **2012**, *62*, 1171-1178.
- (4) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Mooberry, S. L. *J. Nat. Prod.* **2000**, *63*, 611-615.
- (5) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2000**, *63*, 1437-1439.
- (6) Han, B.; McPail, K. L.; Gross, H.; Goeger, D. E.; Mooberry, S. L.; Gerwick, W. H. *Tetrahedron* **2005**, *61*, 11723-11729.
- (7) Choi, H.; Mevers, E.; Byrum, T.; Valeriote, F. A.; Gerwick, W. H. *Eur. J. Org. Chem.* **2012**, *27*, 5141-5150.
- (8) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418-5423.
- (9) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *Bioorg. Med. Chem.* **2002**, *10*, 1973-1978.
- (10) Matthew, S.; Schupp, P. J.; Luesch, H. *J. Nat. Prod.* **2008**, *71*, 1113-1116.
- (11) Tidgewell, K.; Engene, N.; Byrum, T.; Media, J.; Doi, T.; Valeriote, F. A.; Gerwick, W. H. *ChemBioChem* **2010**, *11*, 1458-1466.
- (12) Klein, D.; Braekman, J. C.; Daloz, D.; Hoffmann, L.; Castillo, G.; Demoulin, V. *J. Nat. Prod.* **1999**, *62*, 934-936.
- (13) Klein, D.; Braekman, J. C.; Daloz, D.; Hoffmann, L.; Demoulin, V. *J. Nat. Prod.* **1997**, *60*, 1057-1059.
- (14) Luesch, H.; Yoshida, W. Y.; Harrigan, G. G.; Doom, J. P.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2002**, *65*, 1945-1948.
- (15) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2000**, *63*, 1106-1112.
- (16) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *Tetrahedron* **2002**, *58*, 7959-7966.
- (17) Cardellina II, J. H.; Marner, F. J.; Moore, R. E. *Science* **1979**, *204*, 193-195.

- (18) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (19) Hooper, G. J.; Orjala, J.; Schatzman, R. C.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 529-533.
- (20) Gerwick, W. H.; Tan, L. T.; Satachitta, N. *Alkaloids Chem. Biol.* **2001**, *57*, 75-184.
- (21) Orjala, J.; Gerwick, W. H. *J. Nat. Prod.* **1996**, *59*, 427-430.
- (22) Nunnery, J. K.; Engene, N.; Byrum, T.; Cao, Z.; Jabba, S. V.; Pereira, A. R.; Matainaho, T.; Murray, T. F.; Gerwick, W. H. *J. Org. Chem* **2012**, *77*, 4190-4208.
- (23) Nunnery, J. K. PhD. Dissertation, University of California San Diego, 2012.
- (24) Kwan, J. C.; Eksioglu, E. A.; Liu, C.; Paul, V. J.; Luesch, H. *J. Med. Chem.* **2009**, *52*, 5732-5747.
- (25) Ma, Z.; Song, N.; Li, C.; Zhang, W.; Wang, P.; Li, Y. *J. Pept. Sci.* **2008**, *14*, 1139-1147.
- (26) Williams, P. G.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2002**, *65*, 1336-1339.
- (27) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *37*, 12946-12959.
- (28) von Döhren, H.; Dieckmann, R.; Pavela-Vrancic, M. *Chem. Biol.* **1999**, *10*, R273-279.
- (29) Molinski, T. F.; Reynolds, K. A.; Morinaka, B. I. *J. Nat. Prod.* **2012**, *75*, 425-431.
- (30) Williams, P. G.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2003**, *66*, 1006-1009.
- (31) Mevers, E.; Liu, W.; Engene, N.; Mohimani, H.; Byrum, T.; Pevzner, P. A.; Dorrestein, P. C.; Spadafora, C.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 928-936.

## Chapter 5:

### Parguerene, Precarriebowmide, and Mooreamide: New Lipopeptides from the Marine

#### Cyanobacterium *Moorea* sp.

##### 5.0.1 Abstract

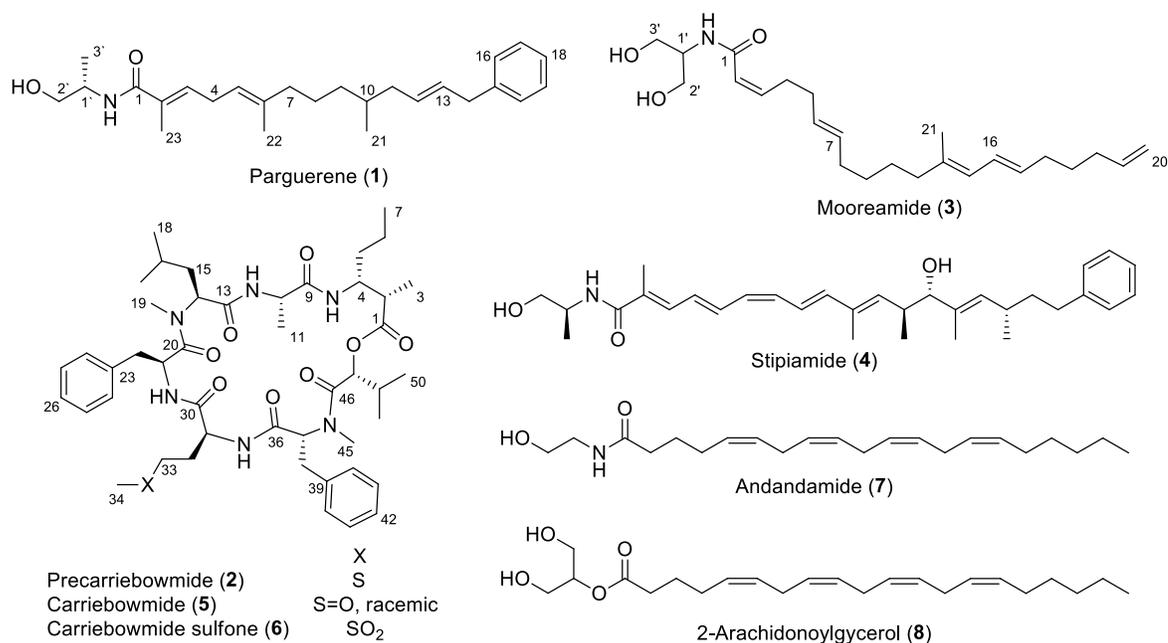
Three new marine cyanobacterial natural products, parguerene (**1**), precarriebowmide (**2**), and mooreamide (**3**), were isolated from two separate collections of *Moorea* sp., one from Puerto Rico and the other from Papua New Guinea. The planar structures of each were deduced by 2D NMR spectroscopy and mass spectrometry. Parguerene and mooreamide are modified alkyl amides, whereas precarriebowmide is a lipopeptide and represents a minor modification compared to two known metabolites, carriebowmide (**5**) and carriebowmide sulfone (**6**). The identification of precarriebowmide led to an investigation into whether carriebowmide and carriebowmide sulfone were true secondary metabolites or isolation artifacts. Both parguerene and mooreamide are structurally reminiscent of the endocannabinoids anadamide, and 2-arachidonoylglycerol and thus it was hypothesized that each would exhibit some cannabinoid receptor binding activity. Unfortunately, parguerene decomposed prior to being evaluated but mooreamide exhibit moderate selective binding affinity towards CB<sub>1</sub> over CB<sub>2</sub> ( $K_i = 0.47 \mu\text{M}$  and  $K_i > 25 \mu\text{M}$ , respectively).

##### 5.1 Introduction

Tropical marine cyanobacteria are exceptionally prolific producers of structurally diverse secondary metabolites, many of which have intriguing biological properties.<sup>1-4</sup> The cyanobacterial genus *Moorea*, formally known as *Lyngbya*,<sup>5</sup> is one of the most prolific producers of these secondary metabolites which include the jamaicamides,<sup>6</sup> malyngamides<sup>7</sup> and apratoxins.<sup>8</sup> These compounds represent a number of different structure classes (alkaloid, polyketide, peptide, or

mixed NRPS/PKS metabolites) and have a number of different biological properties (anticancer, antimicrobial, neurotoxic, and anti-inflammatory). Such diversity within a single genus makes this an ideal organism for continued study for structurally unique and biologically active secondary metabolites.

In the current effort, several filamentous tuft-forming species of marine cyanobacteria were collected from two different locations, one off the south coast of Puerto Rico in March 2011 and the other off the east coast of Papua New Guinea in May 2005. Their extracts and reduced complexity chromatography fractions were evaluated in a number of biological assays, with two such fractions, one from a collection of *Moorea producens* and another from a collection of *Moorea bouillonii*, were found to be either cytotoxic to H-460 human lung cancer cells in vitro (41% survival at 30  $\mu\text{g/mL}$  for the *M. producens* collection) or selective against brain and pancreatic cell lines, and were thus chosen for further investigation. As a result of a NMR-guided fractionation process, three new secondary metabolites were isolated and structurally defined; two are linear alkyl amides and the other a lipopeptide, along with the isolation of previously identified apratoxins A-C and E. The planar structure, double bond geometry, and one of two stereocenters was determined for parguerene (**1**); unfortunately, it decomposed before the second stereocenter could be defined. The lipopeptide, precarriebowmide (**2**), and mooreamide (**3**) were more stable metabolites and were thus fully characterized including the absolute configurations of all stereogenic centers and double bond geometry.



**Figure 5.1:** Structures of the new metabolites parguerene, mooreamide, and precarriewamide, along with structurally related metabolites

## 5.2 Results and Discussion

### 5.2.1 Collection and Isolation

The cyanobacterium *M. producens* was collected by hand from shallow water (1 m) where it was found growing on mangrove roots near La Parguera, Puerto Rico, in March 2011. The cyanobacterium was identified by 16S rRNA analysis, where it shares a 100% maximum identity with what is annotated as *Lyngbya majuscula* 3L (Accession # EU315909.1)(currently *M. producens*). The isopropanol-preserved collection was repetitively extracted (CH<sub>2</sub>Cl<sub>2</sub>—MeOH, 2:1) and fractionated using normal-phase vacuum liquid chromatography (VLC). Further fractionation using reversed-phase HPLC yielded 4.5 mg of parguerene (**1**), a pale yellow oil, and 2.1 mg of precarriewamide (**2**), an amorphous solid.

The cyanobacterium *M. bouillonii* was collected by SCUBA in 10-30 feet of water where it was growing tangled within the coral *Stylophora pistillata* off Pigeon Island, Papua New

Guinea, in May 2005. Previously, a middle polarity fraction from this extract exhibited potent molluscicidal activity and yielded the polyglycosylated macrolide, cyanolide A.<sup>10</sup> A relatively more polar fraction eluting with 25% MeOH/EtOAc exhibited potent toxicity against brain and pancreatic cancer cell lines and was thus chosen for further fractionation. Using a combination of reverse-phase solid phase extraction (SPE), along with both normal and reverse-phase HPLC, yielded 0.7 mg of mooreamide A (**3**), a pale yellow oil, along with the previously identified apratoxins A-C, and E.

### 5.2.2 Structure Elucidation of Parguerene

HR-ESIMS of **1** gave a  $[M+H]^+$  at  $m/z$  398.3056, indicating a molecular formula of  $C_{26}H_{39}NO_2$  and requiring eight degrees of unsaturation. IR spectroscopy suggested the presence of an amide or ester bond and an NH or OH functionality with strong absorption bands at 1646  $cm^{-1}$  and 3283  $cm^{-1}$ . The  $^{13}C$  NMR spectrum also suggested the presence of a mono-substituted phenyl ring ( $\delta_C$  141.1, 128.3 x 2, 128.4 x 2, and 125.8), three olefins ( $\delta_C$  120.0, 130.0, 130.3, 130.5, 135.3, and 137.4) and one amide or ester carbonyl ( $\delta_C$  170.3). The  $^1H$  NMR spectrum corroborated this with protons resonating at  $\delta_H$  7.28, 7.19, and 7.18 for the mono-substituted phenyl ring and four olefinic protons at  $\delta_H$  6.29, 5.54, 5.49, and 5.10 plus two deshielded olefinic methyl groups at  $\delta_H$  1.87 and 1.61.

Analysis of the 2D NMR spectra of **1** (COSY, HSQC, and HMBC) led to the identification of four partial structures (**a-d**)(Figure 5.2). The first (**a**) was comprised of one hydroxy group ( $\delta_H$  2.93) and one amide NH proton ( $\delta_H$  5.82) along with a methylene ( $\delta_H$  3.56/3.70), a methine ( $\delta_H$  4.12) and a methyl group ( $\delta_H$  1.21). Assembly of these atoms by 2D NMR yielded a fragment that resembled alanine but with the carboxy group reduced to a primary alcohol. The second fragment (**b**) consisted of an  $\alpha,\beta$ -unsaturated ketone in which the olefin was  $\alpha$ -substituted with a methyl group and possessed a methylene group at the  $\beta$ -position. This was

followed by a second carbon-carbon double bond that was substituted at the distal location with one more vinyl methyl group. The third partial structure (**c**) contained three consecutive methylene groups adjacent to a methine bearing a methyl and a fourth methylene group; the latter was downfield shifted due to an adjacent di-substituted olefin. An even further downfield shifted methylene was at the distal side of this olefinic bond. The final fragment (**d**) contained the final four degrees of unsaturation present as a mono-substituted phenyl ring.

Assembly of partial structures **a-d** was accomplished by HMBC (Figure 5.2). Partial structures **a** and **b** were linked by correlations from the NH proton of **a** as well as the  $\alpha$ -methyl protons and  $\beta$ -olefinic proton of **b** to the amide carbonyl. Partial structures **b** and **c** were connected through HMBC correlations from the *bis*-allylic methylene and distal olefinic proton of **b** as well as the allylic methylene of **c** to the quaternary olefinic carbon; this connection was reinforced by an HMBC correlation from the vinyl methyl protons to the nearby methylene carbon atom. Converging HMBC correlations from the C-14 allylic-benzylic methylene of **c** and the C-16 aromatic proton of **d** to the quaternary C-15 carbon served to connect these last two partial structures, thus completing the planar structure of parguerene (**1**).

The configurations of the three olefins were determined by a mixture of  $^{13}\text{C}$  NMR analysis and by  $^3J$  coupling. The olefin between C-12 and C-13 exhibited a coupling constant of 15.5 Hz, indicative of *E* configuration. The other two olefins (C-2/C-3 and C-5/C-6) were each tri-substituted; thus, the distinctive carbon shifts of the two vinyl methyl groups were used to infer their geometry.<sup>6</sup> Both olefinic methyl groups showed upfield-shifted carbon resonances (C-23,  $\delta_{\text{H}}$  12.7; C-22,  $\delta_{\text{H}}$  16.8), and thus their corresponding olefins were deduced to be of *E* configuration.

The absolute configuration of the reduced alanine residue in **1** was determined by LC-MS analysis of the oxidized acid hydrolysate appropriately derivatized with Marfey's reagent (D-FDAA). The two standards, L- and D-Ala, were also reacted with D-FDAA and compared to the

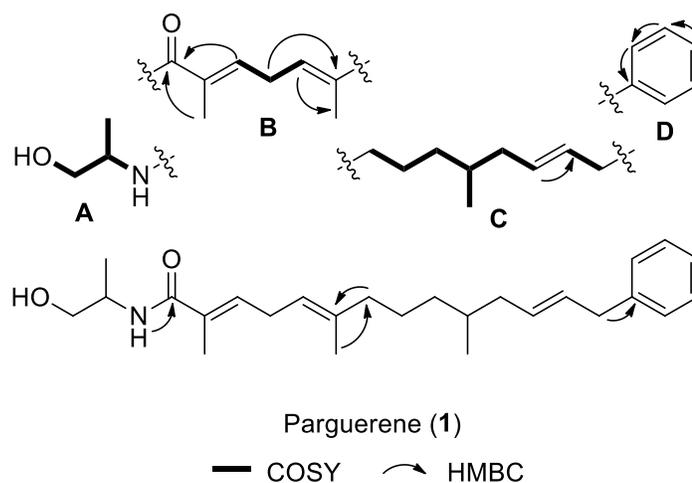
derivatized hydrolysate by LC-MS. From the retention times it was clear that the alanine residue produced from compound **1** was of the L-configuration. Unfortunately, once purified, parguerene (**1**) proved to be unstable and decomposed shortly after acquiring NMR spectra. As a result, the absolute configuration of C-10 was not determined, and thus either a total synthesis or re-collection of the producing organism are required to determine the full absolute configuration of **1**.

Parguerene is structurally reminiscent of stipiamide (**4**), a natural product isolated from the Gram-negative soil bacterium *Myxococcus stipitatus*.<sup>11</sup> Stipiamide and other related compounds exhibit a broad range of biological activities including reversal of multidrug resistance (MDR) in cancer cells,<sup>11</sup> anti-HIV,<sup>12</sup> antifungal and antibacterial<sup>13</sup>. From several structure-activity relationship studies, it was shown that reducing the number of conjugated double bonds in stipiamide significantly reduced its overall toxicity [ED<sub>50</sub> 0.01 nM to 14 μM against adriamycin resistant breast cancer cells (MDR-7adrR)] while the MDR reversing activity was maintained.<sup>14,15</sup> Although parguerene decomposed prior to being evaluated for biological properties, based on these previous studies and comparison of its structure with that of stipiamide, we hypothesize that it would have less cellular toxicity than stipiamide but perhaps retain its MDR reversing properties. Thus, for a variety of reasons, including confirmation of structure and clarification of the one unresolved stereocenter, re-isolation or production of parguerene via total synthesis is needed.

**Table 5.1:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for parguerene in  $\text{CDCl}_3$ 

Position	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}}$ (J in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY <sup>a</sup>
1	170.3			
2	130.3			
3	135.3	6.29, t (7.3)	1, 4, 5, 23	4, 23
4	27.3	2.85, t (7.2)	1, 3, 5, 6	3, 5, 23
5	120.0	5.10, t (7.0)	3, 4, 7, 22	4, 7, 22
6	137.4			
7	39.8	1.95, t (7.4)	8, 9, 22, 6	4, 5, 8
8	25.3	1.40, m	7, 9, 10	7, 9a, 9b
9a	36.2	1.29, m	7, 10, 11, 21	8, 9b, 10
9b		1.09, m	7, 10, 11, 21	8, 9a, 10
10	33.0	1.47, m	8, 9, 11, 21	9a, 9b, 11a, 11b, 21
11a	39.9	2.02, dt (14.0, 6.3)	9, 10, 12, 21	10, 11b, 12
11b		1.86, dt (14.0, 6.9)	9, 10, 12, 21	10, 11a, 12
12	139.5	5.49, dt (15.5, 6.3)	10, 11, 13, 14	11a, 11b
13	130.0	5.54, dt (15.5, 6.5)	14, 15, 16, 17	14
14	39.1	3.34, d (6.5)	13, 15, 16	11a, 11b, 13
15	141.1			
16/20	128.3	7.28, m	15, 17	
17/19	128.4	7.18, m	16, 18	
18	125.8	7.19, m	17	
21	19.5	0.88, d (6.8)	9, 10, 11	10
22	16.1	1.61, s	5, 6, 7	5, 7
23	12.8	1.87, s	3	3, 4
1'	48.1	4.12, m	2', 3'	2a', 2b', 3', NH
2a'	67.7	3.70, dd (10.9, 3.2)	1', 3'	1', 2b', OH
2b'		3.56, dd (10.9, 6.5)	1', 3'	1', 2a', OH
3'	17.1	1.21, m	1', 2'	1'
OH		2.93, s		2a', 2b'
NH		5.82, d (5.5)		1'

<sup>a</sup>500 MHz for  $^1\text{H}$  NMR, HMBC, and COSY. <sup>b</sup>75 MHz for  $^{13}\text{C}$  NMR.



**Figure 5.2:** Four partial structures and select 2D NMR data for parguerene

### 5.2.3 Structure Eluidation of Precarriebowmide

HR-ESIMS of compound **2** gave a  $[M+Na]^+$  at  $m/z$  887.4706, indicating a molecular formula of  $C_{46}H_{68}N_6O_8S$  and requiring 16 degrees of unsaturation (inactive in H-460 cancer cell assay with an  $IC_{50} > 10 \mu M$ ). IR spectroscopy suggested that **2** was peptidic in nature with a strong absorption band at  $1646 \text{ cm}^{-1}$ , and was supported by the observation of seven amide or ester type carbonyls by  $^{13}C$  NMR analysis ( $\delta_C$  175.5, 174.1, 174.0, 172.9, 172.0, 170.6, and 170.0). The  $^1H$  NMR spectrum also suggested a peptide with four amide (NH) protons resonating at  $\delta_H$  7.23, 7.38, 8.60, 8.92, and two *N*-methyl groups at  $\delta_H$  2.62 and 3.09. The  $^{13}C$  and  $^1H$  NMR spectrum indicated the presence of two mono-substituted benzene rings with six peaks, two of which were composed of four carbons each, as indicated by relative peak height ( $\delta_C$  138.1, 137.5,  $130.6 \times 4$ ,  $129.9 \times 4$ , 128.4, and 128.3) and numerous protons resonating between  $\delta_H$  7.20 and 7.40.

Analysis of 1D and 2D NMR spectra (COSY, TOCSY, ROESY, HSQC, and HMBC) led to the identification of five amino acids [alanine (Ala), phenylalanine (Phe), methonine (Met), *N*-methyl-phenylalanine (*N*-MePhe), and *N*-methyl-leucine (*N*-MeLeu)], one hydroxy acid [2-

hydroxy-3-methylbutanic acid (Hmba)], and one extended chain polyketide [3-amino-2-methylhexanoic acid (Amha)]. These residues accounted for 15 of the 16 degrees of unsaturation, indicating that the final degree of unsaturation must arise from **2** having an overall cyclic constitution. This conclusion was also apparent from the residue connectivities observed by HMBC and ROESY, along with the comparison to a known compound, carriebowmide (**4**), as described below.

ROESY correlations from NH protons to adjacent residue  $\alpha$ -protons were used to sequentially connect the Ala, Amha, and Hmba residues. HMBC correlations from the *N*-methyl of the *N*-MePhe residue (C-45) to the carbonyl of the Hmba residue (C-46) extended this sequence to Ala – Amha – Hmba – *N*-MePhe. Another fragment was constructed by an HMBC cross-peak between the *N*-methyl of the *N*-MeLeu residue (C-19) and the carbonyl of Phe (C-20). A third fragment was comprised of a Met residue that showed no long-range correlations to any of the other residues. These partial structures were deployed in a search for related compounds in MarinLit©, which revealed that compound **2** was very similar to the known cyanobacterial metabolite, carriebowmide (**5**).<sup>16</sup> A comparison of their <sup>13</sup>C NMR spectra revealed that the planar structures were identical except for the oxidation state of the sulfur atom in the Met residue. Thus, precarriebowmide was deduced to have a cyclo-[Amha – Ala – *N*-MeLeu – Phe – Met – *N*-MePhe – Hmba] structure (Figure 5.3).<sup>17</sup>

The absolute configuration of the L-Ala and L-Met residues in precarriebowmide (**2**) were determined by LC-MS analysis of the acid hydrolysate appropriately derivatized with Marfey's reagent (D-FDAA). As both **2** and **5** had the same L configured Ala and Met or Met(SO<sub>2</sub>) residues, and the same relative configurations as indicated by their nearly superimposable <sup>13</sup>C NMR spectra, it was deduced that the two compounds should have the same absolute configurations in the remaining residues [L-Phe, D-*N*-MePhe, L-*N*-MeLeu, (2*S*,3*R*)-Amha, and *R*-Hmba].

The similarities between compound **2**, **5** and carriebowmide sulfone (**6**), suggested that **2** might be the actual natural product with the others perhaps representing artifacts as a result of exposure to atmospheric oxygen. Precarriebowmide was extracted and purified from the collected cyanobacterial mass within a few days of collection and preservation, thus potentially preventing **2** from oxidizing to carriebowmide. Close inspection of the original LC-MS chromatogram of the semi-crude fraction containing **2** revealed a trace amount of carriebowmide (**5**); however, the major metabolites were compounds **1** and **2**. Upon purification of **2**, there was no indication of carriebowmide; however, after two weeks in CD<sub>3</sub>OD the sample was found to be comprised of a mixture of precarriebowmide, carriebowmide and carriebowmide sulfone, in an approximate 60:40:<1 ratio, respectively. These observations taken together with the facile oxidation of the sulfide in methionine, it is conceivable that **2** represents the true natural metabolite and that **5** and **6** are artifacts of the isolation process. Moreover, this conclusion is consistent with the finding that the sulfoxide in carriebowmide (**5**) is racemic, and in fact, represents a mixture of two diastereomeric compounds.

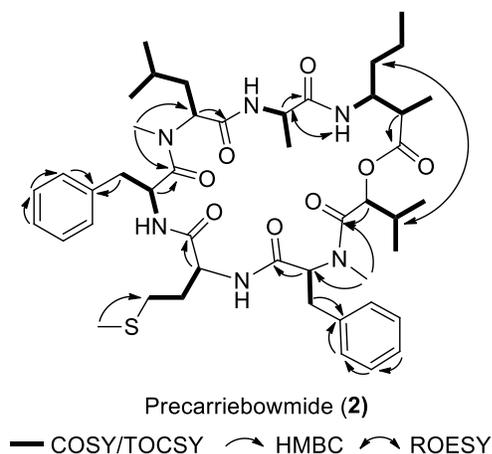
**Table 5.2:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for precarriebowmide (**2**) in  $\text{CD}_3\text{OD}$ 

residue	position	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}}$ ( $J$ in Hz) <sup>a</sup>	HMBC <sup>a</sup>	COSY
Amha	1	175.5			
	2	43.2	2.65, m	1, 3, 4	3, 4
	3	9.1	0.90, d (6.9)	1, 2, 4	2
	4	52.0	4.28, m	2, 3, 5	2, 5a, 5b
	5a	35.0	1.52, m	4, 6, 7	4
	5b		1.50, m	4, 6, 7	4
	6a	20.6	1.43, m	5, 7	7
	6b		1.37, m	5, 7	7
	7	13.8	0.97, d (7.4)	5, 6	6a, 6b
	8-NH		7.38		
Ala	9	174.1			
	10	48.8	4.50, m	9, 11, 13	11
	11	16.4	1.16, d (6.8)	9, 10	10
	12-NH		8.92, d (9.0)	13	
<i>N</i> -MeLeu	13	170.6			
	14	59.8	4.69, m	13, 15, 19	15a, 15b
	15a	37.8	1.74, m	14, 16, 17, 12-NH	14, 16
	15b		-0.23, td (10.9, 3.5)	14, 12-NH	14, 16
	16	25.7	1.44, m	15, 17	15a, 15b, 16, 17
	17	21.8	0.78, d (6.5)	15, 16, 18	16
	18	23.9	0.73, d (6.5)	17	16
	19	29.7	2.62, s	14, 20	
Phe	20	174.0			
	21	52.8	4.77, dd (9.9, 6.0)	22	22a, 22b
	22a	38.8	3.09, m	20, 21, 24	21
	22b		3.03, m	20, 21, 24	21
	23	137.5			
	24/28	230.6	7.21, d (7.5)	22, 23, 25, 26	

**Table 5.2: continued**

<b>residue</b>	<b>position</b>	$\delta_C^b$	$\delta_H$ (J in Hz) <sup>a</sup>	<b>HMBC<sup>a</sup></b>	<b>COSY<sup>a</sup></b>
	25/27	129.9	7.28, d (7.5)	24, 26	
	26	128.3	7.24, t (7.5)	24, 25	
	29-NH		7.23, m	30	
Met	30	172.9			
	31	52.0	4.55, m	30, 32	33a, 33b
	32a	30.1	2.19, m	34	33a, 33b
	32b		1.31, m	31, 33	33a, 33b
	33a	33.5	1.88, m	34	31, 32a, 32b
	33b		1.67, m	32	31, 32a, 32b, 34
	34	15.2	2.05, s	32, 33	33b
	35-NH		8.60, d (8.8)		
<i>N</i> -MePhe	36	170.0			
	37	62.2	4.62, m	36, 38, 45	38a, 38b, 45
	38a	37.7	3.41, dd (13.5, 9.9)	36, 37, 39	37
	38b		2.95, dd (13.5, 5.6)	36, 37, 39	37
	39	138.1			
	40/44	130.6	7.25, d (7.5)	39, 41, 42	
	41/43	129.9	7.37, t (7.5)	40, 42	
	42	128.4	7.27, t (7.5)	40, 41	
	45	30.1	3.09, s	37, 46	37
Hmba	46	172.0			
	47	76.1	5.16, d (2.81)	48, 49, 50	48
	48	30.6	1.70, m	47, 48, 50	47, 50
	49	19.7	1.18, d (6.8)	47, 48, 50	48, 50
	50	16.9	0.88, d (6.8)	47, 48, 49	48, 49

<sup>a</sup>500 MHz for <sup>1</sup>H NMR, HMBC, and COSY; <sup>b</sup>75 MHz for <sup>13</sup>C NMR



**Figure 5.3:** Select 2D NMR data for precarriebowmide (**2**)

#### 5.2.4 Structure Elucidation of Mooreamide (**3**)

HR-ESI/MS of **3** yielded an  $[M+H]^+$  peak at  $m/z$  390.3006, indicating a molecular formula of  $C_{24}H_{39}NO_3$  and requiring 6 degrees of unsaturation (calcd for  $C_{24}H_{40}NO_3$  390.3003). IR spectroscopy suggested the presence of a carbonyl along with *NH* and/or *OH* functionality with strong absorption bands at 1657 and 3399  $cm^{-1}$ . The  $^{13}C$  NMR spectrum suggested that the carbonyl was present as an amide or ester functionality with an observed shift at 167.2 ppm, and also indicated the presence of five olefins ( $\delta_C$  145.9, 139.1, 136.0, 132.5, 130.9, 129.2, 126.6, 124.8, 122.0, and 114.2), accounting for all six degrees of unsaturation. The  $^1H$  NMR spectrum supported the presence of both an *NH* and *OH* functionality with corresponding protons resonating at  $\delta_H$  6.25 and 2.50, respectively, and a number of olefins with ten protons resonating between  $\delta_H$  4.90 and 6.30 along with one deshielded vinyl methyl at  $\delta_H$  1.71.

Analysis of 2D NMR data for **3** (COSY, HSQC, HMBC, and NOESY) led to the identification of three partial structures (**a-c**), each consisting of separate spin systems (Figure 5.4). The first (**a**) was comprised of two hydroxy groups ( $\delta_H$  2.50 x 2), one amide proton ( $\delta_H$  6.25), two methylenes ( $\delta_H$  3.85 x 2) and a methine ( $\delta_H$  4.00). From  $^1H$ - $^1H$  COSY correlations it was clear that this fragment was an amino glycerol moiety and this was further supported by the symmetry

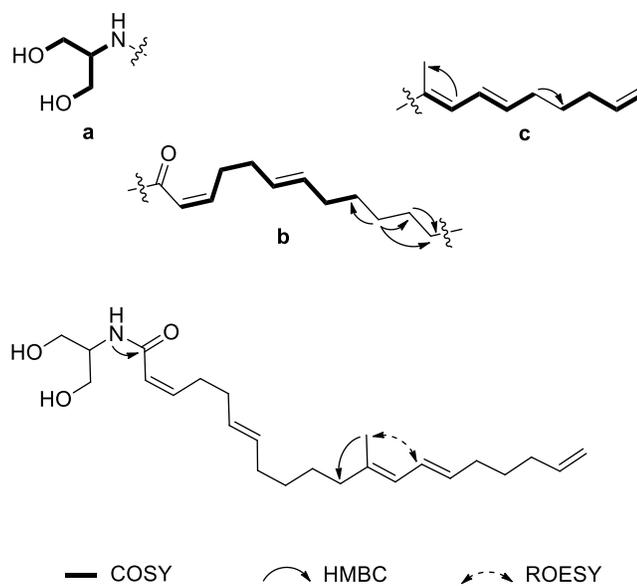
seen in both the  $^1\text{H}$  and  $^{13}\text{C}$  data. By peak integration, one signal represented both hydroxy groups and another for both methylenes. The second spin system (**b**) consisted of 11 substituted carbons in a linear fashion in which an  $\alpha,\beta$ -unsaturated ketone was followed by two methylenes in the  $\beta$  and  $\gamma$  positions. Adjacent to the distal methylene was a disubstituted carbon-carbon double bond followed by four consecutive methylenes. The third and final fragment (**c**) consisted of the remaining 10 carbons, including a conjugated diene which was substituted with a methyl group. This was followed by three consecutive methylene groups and a terminal carbon-carbon double bond.

The three partial structures were assembled by two key HMBC correlations. Partial structures **a** and **b** were linked by a correlation from the *NH* proton of **a** to the  $\alpha,\beta$ -unsaturated ketone in **b** (*NH* to C-1). A second key HMBC correlation was between the vinyl methyl in **c** and the distal methylene in **b** (H-21 to C-10), thus completing the planar structure of mooreamide A (**3**) as seen in figure 5.4.

The configurational assignment of the four internal olefins were determined by a combination of NOE correlations,  $^3J$  coupling, and  $^{13}\text{C}$  NMR analysis. The configuration of the tri-substituted olefin between C-14/C-15 was determined using a key NOE correlation between H-21 ( $\delta_{\text{H}}$  1.72) and the olefinic proton on C-14 ( $\delta_{\text{H}}$  6.22), establishing the configuration of the double bond as *E*. The configuration of the olefins between C-2/C-3 and C-16/C-17 were determined using  $^3J$  coupling analysis. The measured coupling constant between C-2/C-3 was 11.7 Hz which is indicative of *Z* configuration, whereas the coupling constant between C-16/C-17 of 15.1 Hz was indicative of *E* conformation.<sup>18</sup> The configuration of the final olefin between C-6/C-7 was more challenging to determine as both the vinyl protons and adjacent methylenes were overlapped, thus preventing the use of either  $^3J$  coupling constant analysis or NOE correlations. Thus, the distinctive carbon shifts of the two allylic methylenes (C-5 and C-8), adjacent to the

olefin, were used to infer its geometry.<sup>19, 20</sup> Both methylene groups showed downfield-shifted carbon resonances (C-5,  $\delta_C$  32.0; C-8,  $\delta_C$  31.1), thus indicative of *E* configuration.

Due to the isolation of highly cytotoxic apratoxins from the same fraction that yielded mooreamide A (**3**), and the limited quantity of **3** available, we were unable to evaluate it for cytotoxic activity.<sup>21,22</sup> However, based on structure homology between **3** and two endocannabinoid ligands, anandamide (**7**) and 2-arachidonoylglycerol (**8**), it was evaluated against both neuroreceptors, CB<sub>1</sub> and CB<sub>2</sub>. Compound **3** exhibited moderate affinity for CB<sub>1</sub> with a  $K_i$  value of 0.47  $\mu$ M, whereas it showed no affinity for CB<sub>2</sub> at concentrations up to 25  $\mu$ M, and thus appears to be strongly selective towards CB<sub>1</sub> (>10-fold). Mooreamide (**3**) exhibits the strongest affinity for CB<sub>1</sub> thus far from this class of marine metabolites.<sup>23,24</sup>



**Figure 5.4:** Three partial structures and select 2D NMR data for mooreamide (**3**)

**Table 5.3:**  $^1\text{H}$  and  $^{13}\text{C}$  NMR assignments for mooreamide (**3**) in  $\text{CDCl}_3$ 

Position	$\delta_{\text{C}}^{\text{c}}$	$\delta_{\text{H}}$ ( $J$ in Hz) <sup>a</sup>	HMBC <sup>b</sup>	COSY <sup>a</sup>
1	167.2			
2	122.0	5.74, d (11.7)	1, 4	3, 4
3	145.9	6.03, dt (11.5, 7.7)	1, 4, 5	2, 4
4	28.6	2.70, dt (7.5, 6.4)	1, 2, 3, 5	2, 3, 5
5	32.0	2.12, m	3, 4	4, 6, 7
6	129.2	5.43, m	4, 5, 7, 8	5, 8
7	130.9	5.43, m	5, 6, 8	5, 8
8	31.1	2.12	6, 7, 11	4, 6, 7
9	29.5	1.25, m	8, 11	
10	28.8	1.31, m	8, 11	
11	39.8	2.05, m	8, 12, 13, 21	
12	136.0			
13	124.8	5.78, d (11.0)	11, 14, 15, 21	14
14	126.6	6.22, dd (15.1, 11.0)	12, 13, 16	13, 15
15	132.5	5.57, dt (15.1, 7.0)	13, 16, 17	15, 16
16	32.9	2.07, m	14, 15, 17	15, 18
17	28.7	1.39, m	15, 16, 18	
18	33.7	2.04, m	17, 19, 20	17, 19
19	139.1	5.82, dt (10.3, 6.5)	17, 18, 20	18, 20
20a	114.2	4.99, d (17.0)	18	19
20b		4.93, d (10.3)	18	19
21	16.5	1.72, s	11, 13	
1'	52.1	4.00, dt (7.2, 4.1)	2'	NH
2'/3'	63.9	3.85, m	1'	OH
NH		6.25, m	1	1'
OH		2.50, bs		2'

<sup>a</sup>500 MHz for  $^1\text{H}$  NMR, and COSY; <sup>b</sup>600 MHz for HMBC; <sup>c</sup>125 MHz for  $^{13}\text{C}$  NMR.

### 5.2.5 Biosynthetic Origin

Although precarriebowmide, parguerene, and mooreamide are ostensibly of mixed biosynthetic origin, they clearly represent two very different structural classes, cyclic lipopeptide versus modified linear alkyl amide.<sup>25</sup> Furthermore, each metabolite has undergone modifications to the core structure, such as the incorporation of a hydroxy acid residue and *N*-methylation events (*N*-MePhe and *N*-MeLeu) in precarriebowmide versus integrated aromatic and aliphatic moieties as well as a presumed reductive offloading of alanine and serine in parguerene and mooreamide, respectively.<sup>26</sup> The location of the methyl groups on the alkyl chain of parguerene (**1**) suggests the possibility of their arising from one of two different biosynthetic pathways, both of which likely begin with phenyl acetic acid. In one scenario, the phenyl acetic acid is condensed with a sesquiterpene moiety whereas in the second phenyl acetic acid is the starter unit for six iterative polyketide synthase (PKS) additions of acetate. In the latter case, *C*-methylation must occur on the *C*-2 position of every other acetate unit. Differentiation between these two intriguing alternatives may be possible through genome sequencing of DNA preserved at the time of collection.

As for mooreamide (**3**), the C1-C20 chain is likely assembled by a polyketide synthetase (PKS), where four of the five olefins (C2/C3, C6/C7, C12/C13, and C14/C15) occur between the incorporation of predicted acetate units, and thus the modules responsible for each would be lacking the enoylreductase domain.<sup>27</sup> However, the terminal olefin (C19/C20) must be formed in a different manner as it involves both carbons of a single acetate unit. The mechanism for formation of this functionality may be similar to that responsible for forming the terminus of the jamaicamides which involve a fatty acid type desaturase.<sup>28,29</sup> Other known mechanisms to form terminal olefins, such as the olefin synthase (OLS) in the curacin A pathway, introduce terminal olefins at the carboxyl terminus, and thus are likely not involved in the biosynthesis of mooreamide (**3**).<sup>30,31</sup>

### 5.3 Conclusion

The new metabolites, parguerene (**1**), precarriebowmide (**2**), and mooreamide A (**3**), were each isolated from a collection of *Moorea* sp. that were found growing in the tropical marine environment. These two collections have been extraordinarily fruitful as numerous structurally intriguing metabolites have previously been isolated, such as apratoxins A-C and E, several lyngbyabellins<sup>23</sup> and cyanolide A.<sup>11</sup> Both parguerene and mooreamide A are structurally distinct from each of these other compound classes as they are alkyl amides, and consist of a modified fatty acid tail and an amide linkage. The planar structure of all three of these metabolites were determined by NMR and other spectroscopic techniques, while the absolute configurations of all but one stereocenter were determined by carbon chemical shift comparison, and Marfey's analysis. The structures of parguerene and mooreamide add to a growing family of alkyl amide metabolites from cyanobacteria, most of which possess interesting biological activities. In this case, mooreamide A (**3**) was shown to be a selective agonist to the CB<sub>1</sub> receptor.

### 5.4 Experimental Methods

#### 5.4.1 General Experimental Procedures

Optical rotations were measured on a JASCO P-2000 polarimeter, UV spectra on a Beckman Coulter DU-800 spectrophotometer, and IR spectra were obtained using a Nicolet IR-100 FT-IR spectrophotometer using KBr plates. NMR spectra were recorded with solvent peaks as internal standards ( $\delta_C$  77.0,  $\delta_H$  7.26 for CHCl<sub>3</sub>, and  $\delta_C$  49.0,  $\delta_H$  3.31 for CH<sub>3</sub>OH) on a Varian Unity 500 MHz spectrometer (500 and 125 MHz for <sup>1</sup>H and <sup>13</sup>C NMR, respectively) and Varian Unity 300 MHz spectrometer (300 and 75 MHz for <sup>1</sup>H and <sup>13</sup>C NMR, respectively). LR- and HR-ESIMS were obtained on ThermoFinnigan LCQ Advantage Max and Thermo Scientific LTQ Orbitrap-XL mass spectrometers, respectively. All solvents were either distilled or of HPLC

quality. Acid hydrolysis was performed using a Biotage (Initiator) microwave reactor equipped with high pressure vessels.

#### 5.4.2 Morphological Identification

Morphological characterization were performed using an Olympus IX51 epifluorescent microscope (1000X) equipped with an Olympus U-CMAD3 camera. Morphological comparison and putative taxonomic identification of the cyanobacterial specimen was performed in accordance with modern classification systems.<sup>32,33</sup>

#### 5.4.3 Collection, Extraction, and Isolation

The *Moorea producens* cyanobacterial biomass (10.4 g, dry wt) was extracted with 2:1 CH<sub>2</sub>Cl<sub>2</sub>—MeOH to afford 3.9 g of dried extract. A portion of the extract was fractionated by silica gel VLC using a stepwise gradient solvent system of increasing polarity starting from 100% hexanes to 100% EtOAc to 100% MeOH (nine fractions). The fraction eluting with 100% EtOAc was separated further using RP HPLC [4 $\mu$  Synergi Fusion, 65% CH<sub>3</sub>CN/H<sub>2</sub>O over 50 min to produce six fractions (1-6)] to yield pure parguerene (**1**, 4.5 mg) and precarriebowmide (**2**, 2.1 mg).

The second cyanobacterium, PNG 5/19/05-8, was collected by SCUBA in 10-30 feet of water off of Pigeon Island on the northeast coast of New Britain Island, Papua New Guinea and was previously identified as *Moorea bouillonii* by 16S RNA sequencing.<sup>10</sup> The biomass (37.9 g, dry wt) was extracted with 2:1 CH<sub>2</sub>Cl<sub>2</sub>—MeOH to afford 1.2 g of dried extract and was subsequently fractionated by silica gel VLC using a stepwise gradient solvent system of increasing polarity starting from 100% hexanes to 100% MeOH (nine fractions, A-I). The fraction eluting with 75% EtOAc/MeOH (fraction H) was separated further by a 1 g RP SPE using a stepwise gradient solvent system of decreasing polarity starting from 50% MeOH/H<sub>2</sub>O to

100% CH<sub>2</sub>Cl<sub>2</sub>. Fractions 3 and 4, which eluted with 70% MeOH/H<sub>2</sub>O and 80% MeOH/H<sub>2</sub>O, respectively, were combined and further purified using RP HPLC [4 $\mu$  Synergi Fusion, 60% CH<sub>3</sub>CN/H<sub>2</sub>O for 30 min, followed by 70% CH<sub>3</sub>CN/H<sub>2</sub>O for 15 min and then 100% CH<sub>3</sub>CN for 20 min at 3 mL/min to produce six fractions (1-6)]. The final step in the purification employed NP HPLC [5 $\mu$  Luna, holding 90% hexanes/EtOAc for 5 min and then changing to 100% EtOAc over 15 min at 3 mL/min to produce another eight fractions (A-H)] yielding 0.7 mg of pure mooreamide A (**1**).

**Parguerene (1):** pale yellow oil;  $[\alpha]_D^{23} +17.3$  (*c* 0.22, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 202 (4.50), 206 (2.66) nm; IR (neat)  $\nu_{\max}$  3283, 3050, 2958, 1646, 1541, 1449, 1239, 1185, 1134, 746; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>), see Table 1; HRESIMS [M+H]<sup>+</sup> *m/z* 398.3056 (calcd for C<sub>26</sub>H<sub>40</sub>NO<sub>2</sub>, 398.3054).

**Precarriebowmide (2):** amorphous solid;  $[\alpha]_D^{23} -52.6$  (*c* 0.15, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 201 (4.67), 256.0 (2.91) nm; IR (neat)  $\nu_{\max}$  3061, 3030, 2958, 1646, 1541, 1239, 1134, 746; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>), see Table 2; HRESIMS [M+Na]<sup>+</sup> *m/z* 887.4706 (calcd for C<sub>46</sub>H<sub>69</sub>N<sub>6</sub>O<sub>8</sub>S, 887.4712).

**Mooreamide (3):** pale yellow oil; UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 230 (4.12) nm; IR (neat)  $\nu_{\max}$  3399, 2924, 2853, 1657, 1543, 1441, 1383, 1265, 1197, 1076, 738 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) and <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>), see Table 3; HRESIMS [M+H]<sup>+</sup> *m/z* 390.3006 (calcd for C<sub>24</sub>H<sub>40</sub>NO<sub>3</sub> 390.3003).

#### 5.4.4 Oxidation, Acid Hydrolysis, and Marfey's Analysis of Parguerene

Parguerene (**1**, 0.5 mg) was dissolved in 200  $\mu$ L of acetone and cooled to 0 °C, then treated with 10  $\mu$ L of 0.25 M Jones reagent (CrO<sub>3</sub>, H<sub>2</sub>SO<sub>4</sub>). After 20 min the reaction was quenched with 200  $\mu$ L of isopropyl alcohol and the reaction mixture was extracted with EtOAc 5x to yield the desired product. The reaction product was then treated with 300  $\mu$ L of 6 N HCl in

a microwave reactor at 160 °C for 5 min. The reaction product was dissolved in 300 µL of 1 M sodium bicarbonate, and then 56 µL of 0.5% D-FDAA (1-fluoro-2,4-dinitrophenyl-5-D-alanine amide) was added in acetone. The solution was maintained at 40 °C for 80 min at which time the reaction was quenched by the addition of 300 µL of 1 N HCl. The reaction mixture was diluted with 300 µL of CH<sub>3</sub>CN and 10 µL of the solution was analyzed by LC-ESIMS.

The Marfey's derivatives of the hydrolysate and standards were analyzed by RP HPLC using a Phenomenex Luna 5 µm C<sub>18</sub> column (4.6 x 250 mm). The HPLC conditions began with 10% CH<sub>3</sub>CN/90% H<sub>2</sub>O + 0.1% formic acid (FA) followed by a gradient profile to 50% CH<sub>3</sub>CN/50% H<sub>2</sub>O +0.1% FA over 85 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. The retention times of the D-FDAA derivatives of the authentic amino acids were D-Ala (50.62 min), and L-Ala (56.44 min); the derivative of the hydrolysate product gave a peak with a retention time of 56.68 min, corresponding to L-Ala.

#### 5.4.5 Acid Hydrolysis and Marfey's Analysis of Precarriebowmide

Precarriebowmide (**2**, 0.3 mg) was treated with 300 µL of 6 N HCl in a microwave reactor at 160 °C for 5 min. The reaction product was dissolved in 200 µL of 1 M sodium bicarbonate, and then 32 µL of 0.5% D-FDAA was added in acetone. The solution was maintained at 40 °C for 70 min at which time the reaction was quenched by the addition of 100 µL of 2 N HCl. The reaction mixture was diluted with 200 µL of CH<sub>3</sub>CN and 10 µL of the solution was analyzed by LC-ESIMS.

The Marfey's derivatives of the hydrolysate and standards were analyzed by RP HPLC using a Phenomenex Luna 5 µm C<sub>18</sub> column (4.6 x 250 mm). The HPLC conditions were identical to the method described above. The retention times of the D-FDAA derivatives of the authentic amino acids were D-Ala (50.17), L-Ala (56.33), D-Met (62.35) and L-Met (70.49); the

hydrolysate product gave peaks with retention times of 56.51 and 70.56 min, according to L-Ala and L-Met, respectively.

#### 5.4.6 Polymerase Chain Reaction (PCR) and Cloning for PRM-25/Mar/11-2

Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega Inc.) following the manufacturer's specifications. DNA concentration and purity were measured on a DU 800 spectrophotometer (Beckman Coulter). The 16S rRNA genes were PCR-amplified from isolated DNA using the cyanobacterial specific primers 27F5'-AGAGTTTGATCCTGGCTCAG-3' and 809R 5'-GCTTCGGCACGGCTCGGGTCGATA-3'. The PCR reaction contained 1.0  $\mu$ L (~100 ng) of DNA, 2.5  $\mu$ L of 10  $\times$  PfuUltra IV reaction buffer, 1.0  $\mu$ L (10 mM) of dNTP mix, 1.0  $\mu$ L of each primer (10  $\mu$ M), 1.0  $\mu$ L of PfuUltra IV fusion HS DNA polymerase and 17.5  $\mu$ L H<sub>2</sub>O for a total volume of 25  $\mu$ L. The PCR reactions were performed in an Eppendorf Mastercycler gradient as follows: initial denaturation for 4 min at 95°C, amplification by 30 cycles of 30 sec at 95°C, 30 sec at 50°C and 1 min at 72°C, and final elongation for 7 min at 72°C. PCR products were purified using a MinElute PCR Purification Kit (Qiagen) before subcloning with the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) following the manufacturer's specifications. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen) and sequenced with M13 primers. The 16S rRNA gene sequence is available in the DDBJ/EMBL/GenBank databases under acc. No. KC790370.

#### 5.4.7 Phylogentic Inferences for PRM-25/Mar/11-2

All gene sequences were analyzed using Geneious Pro v.5.5.4.<sup>34</sup> The 16S rRNA gene sequences were aligned using the L-INS-I algorithm in MAFFT v6.814b.<sup>35</sup> Best-fitting nucleotide substitution models optimized by maximum likelihood were selected using corrected Akaike/Bayesian Information Criterion (AIC/BIC) in jModelTest v0.1.1.<sup>36</sup> The evolutionary

histories of the cyanobacterial genes were inferred using Maximum Likelihood (ML) and Bayesian inference algorithms. The Maximum Likelihood (ML) inference was performed using PhyML<sup>37</sup> in Geneious Pro v5.5.4. The analysis was run using the GTR+I+G model (selected by AIC and BIC criteria) assuming heterogeneous substitution rates and gamma substitution of variable sites (proportion of invariable sites (pINV) = 0.000, shape parameter ( $\alpha$ ) = 0.16, number of rate categories = 4). Bootstrap resampling was performed on 1,000 replicates. Bayesian analysis was conducted using MrBayes<sup>38</sup> in Geneious Pro v5.5.4 with four Metropolis-coupled MCMC chains (one cold and three heated) ran for 3,000,000 generations. The first 25% were discarded as burn-in and data set was sampled with a frequency of every 200 generations.

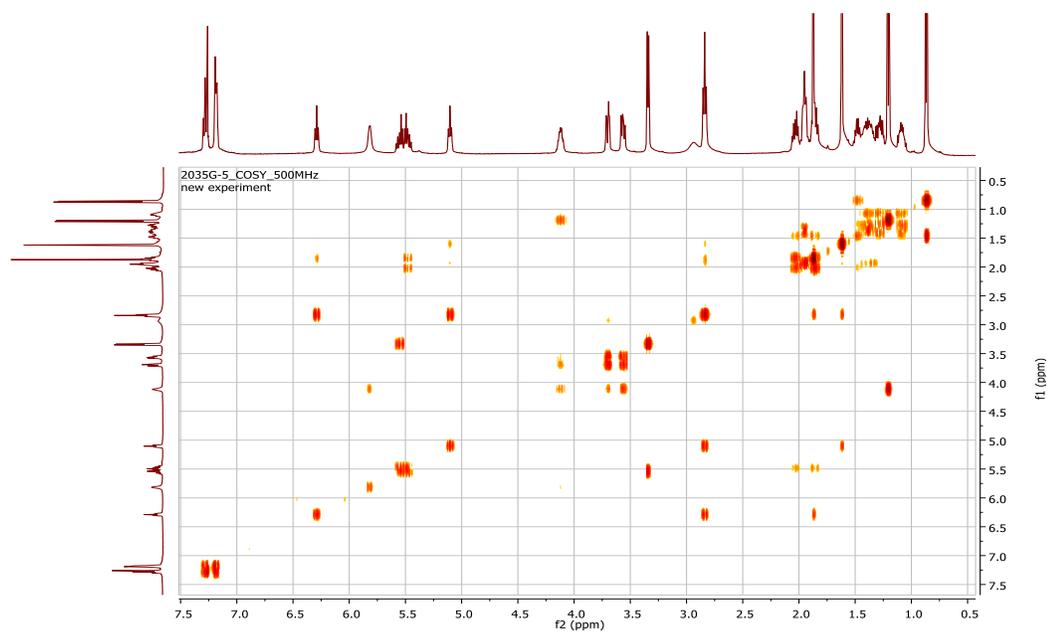
#### 5.4.8 H-460 Cytotoxicity Assay

H-460 cells were added to 96-well plates at  $3.33 \times 10^4$  cells/mL of Roswell Park Memorial Institute (RPMI) 1640 medium with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. The cells, in a volume of 180  $\mu$ L per well, were incubated overnight (37 °C, 5% CO<sub>2</sub>) to allow recovery before treatment with test compounds. Compounds were dissolved in DMSO to a stock concentration of 10 mg/mL. Working solutions of the compounds were made in RPMI 1640 medium without FBS, with a volume of 20  $\mu$ L added to each well to give a final compound concentration of either 30  $\mu$ g/mL or 3  $\mu$ g/mL. An equal volume of RPMI 1640 medium without FBS was added to wells designated as negative controls for each plate. Plates were incubated for approximately 48 h before staining with MTT. Using a ThermoElectron Multiskan Ascent plate reader, plates were read at 570 and 630 nm.

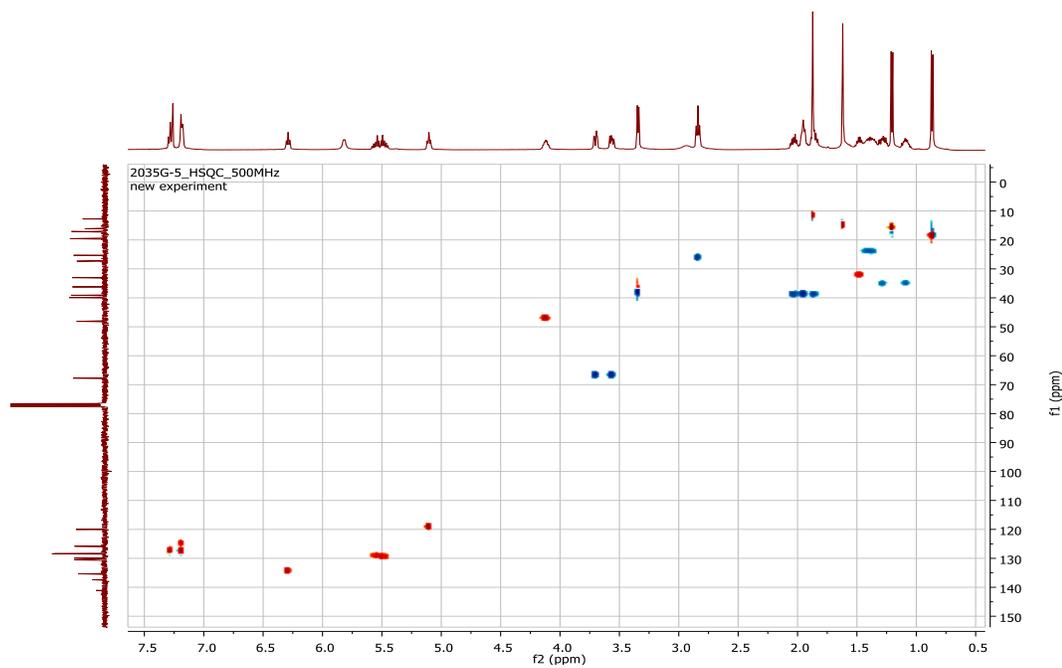
### 5.5 Acknowledgements

We thank the University of Puerto Rico, especially the Department of Marine Sciences Research Center on Magueyes and D. Ballantine for assistance in making the cyanobacterial collections from Puerto Rico, the government of Papua New Guinea for the permission to collect the cyanobacterial specimens, V. Di Marzo, T. Byrum,, W. H. Gerwick for their scientific contributions, P. Boudreau and A. Saitman for helpful discussions, and the UCSD mass spectrometry facilities for their analytical services. We also acknowledge the Growth Regulation & Oncogenesis Training Grant NIH/NCI (T32A009523-24) for a fellowship to E.M. and NIH Grant CA 100851 for support of the research.





**Figure 5.6.3:** COSY (500 MHz,  $\text{CDCl}_3$ ) spectrum of parguerene



**Figure 5.6.4:** gHSQC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of parguerene

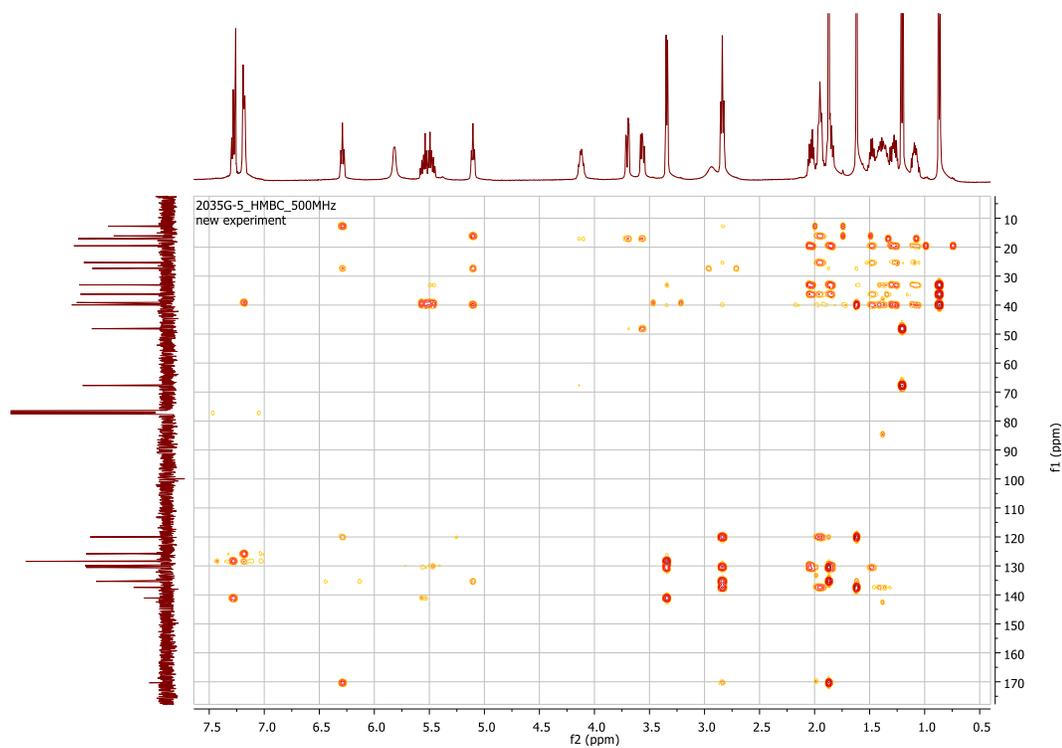


Figure 5.6.5: HMBC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of parguerene

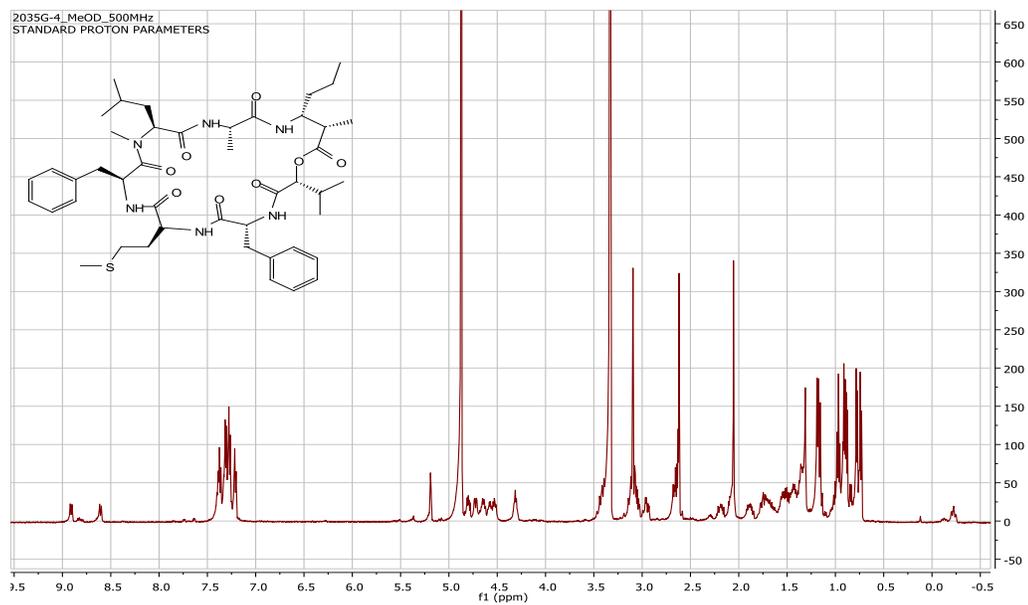


Figure 5.6.6:  $^1\text{H}$  NMR (500 MHz,  $\text{CD}_3\text{OD}$ ) spectrum of precarriebowmide

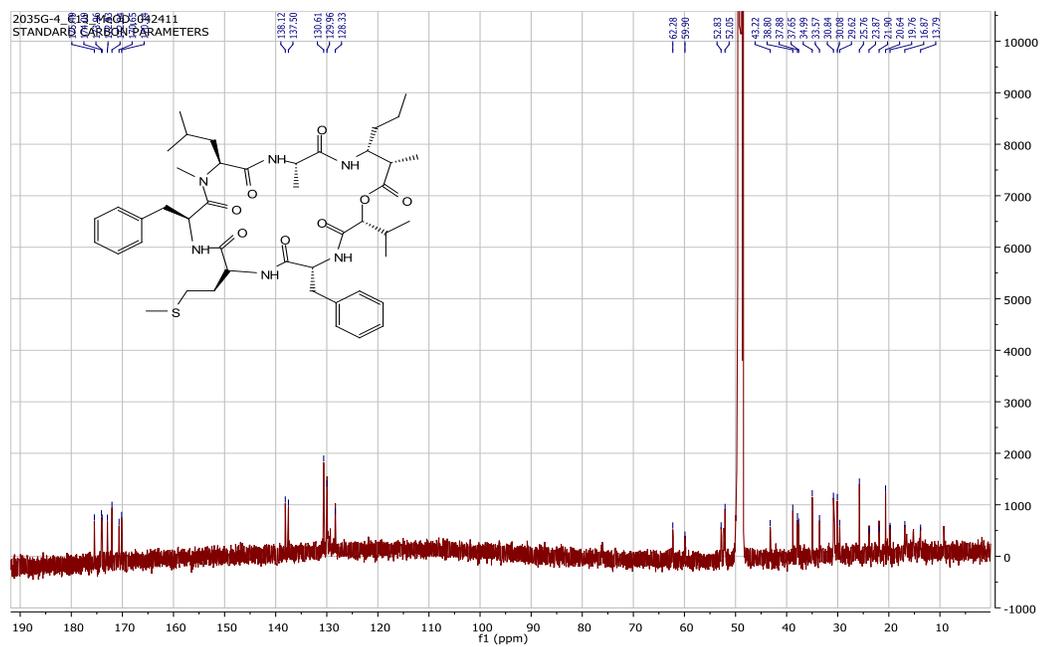


Figure 5.6.7:  $^{13}\text{C}$  NMR (75 MHz,  $\text{CD}_3\text{OD}$ ) spectrum of precarriebowmide

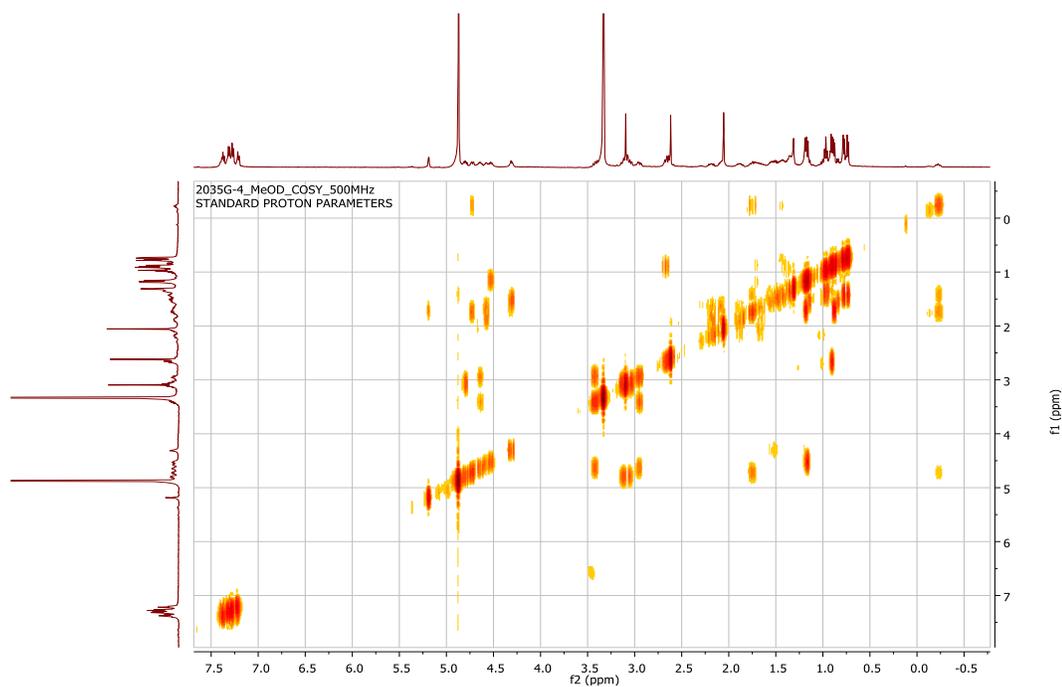


Figure 5.6.8: COSY (500 MHz,  $\text{CD}_3\text{OD}$ ) spectrum of precarriebowmide

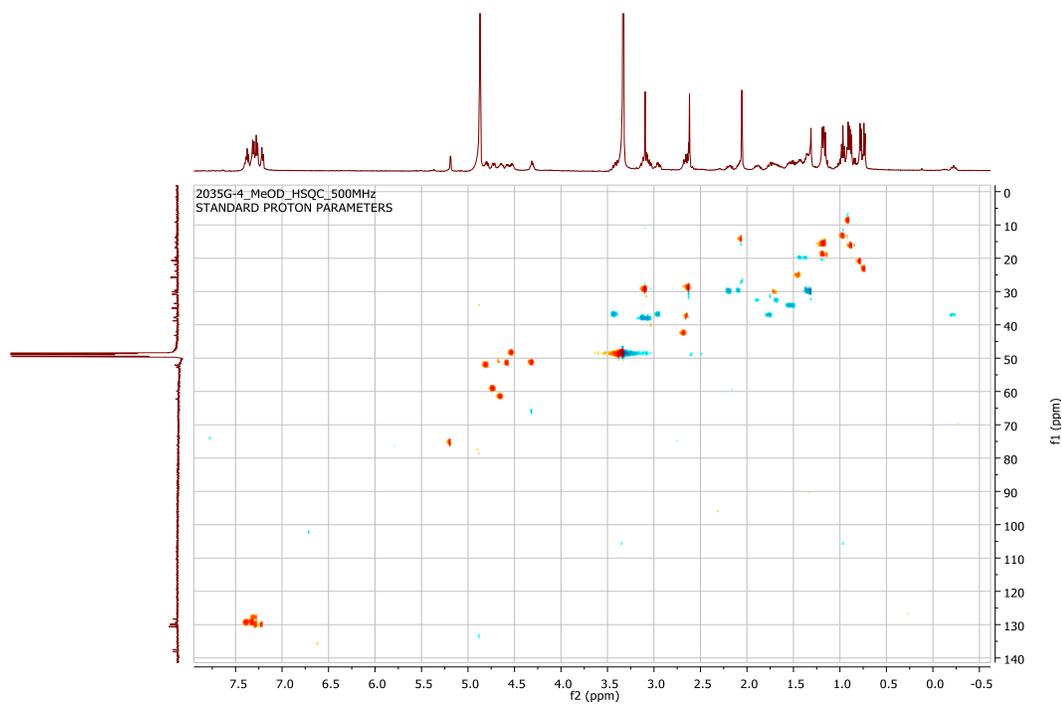


Figure 5.6.9: gHSQC ( $^1\text{H}$  500 MHz,  $\text{CD}_3\text{OD}$ ) spectrum of precarriebowmide

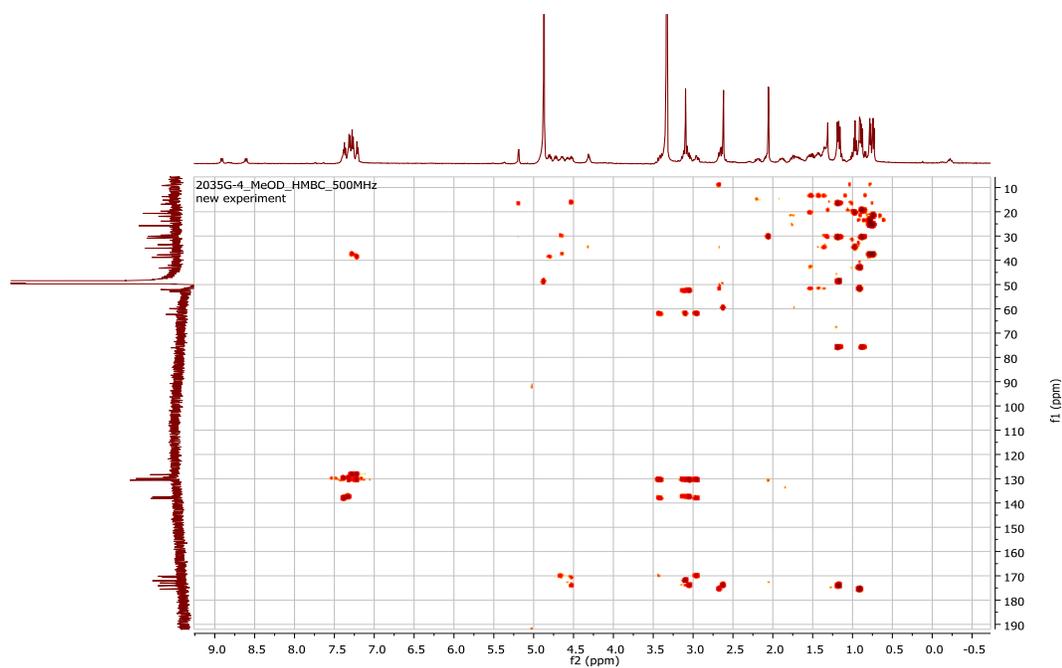
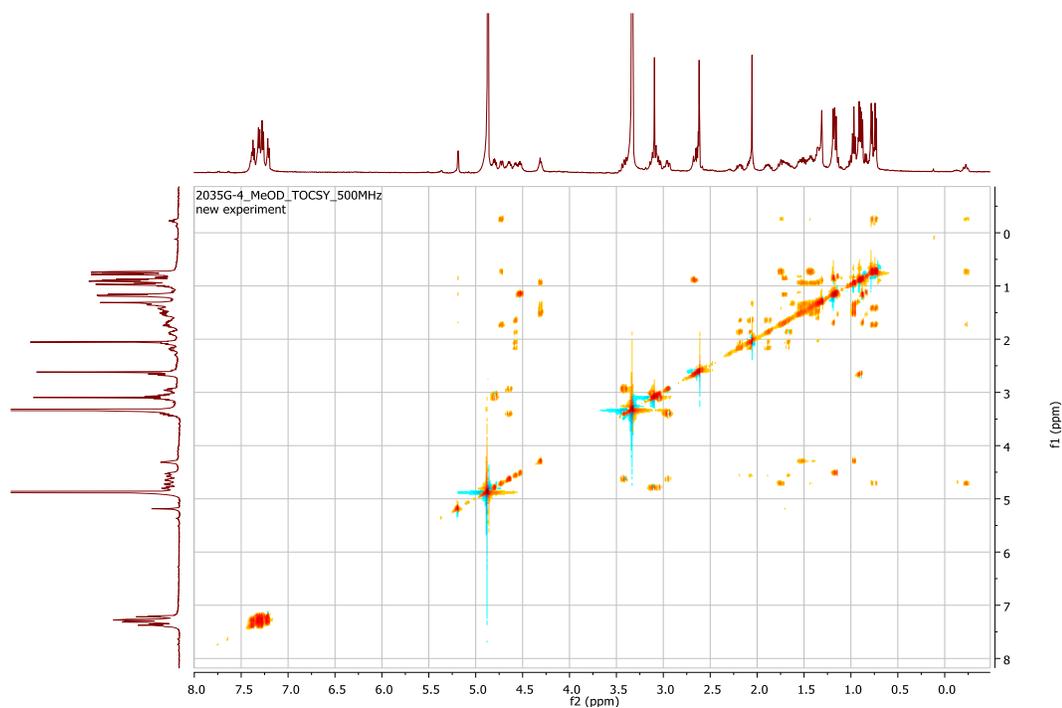
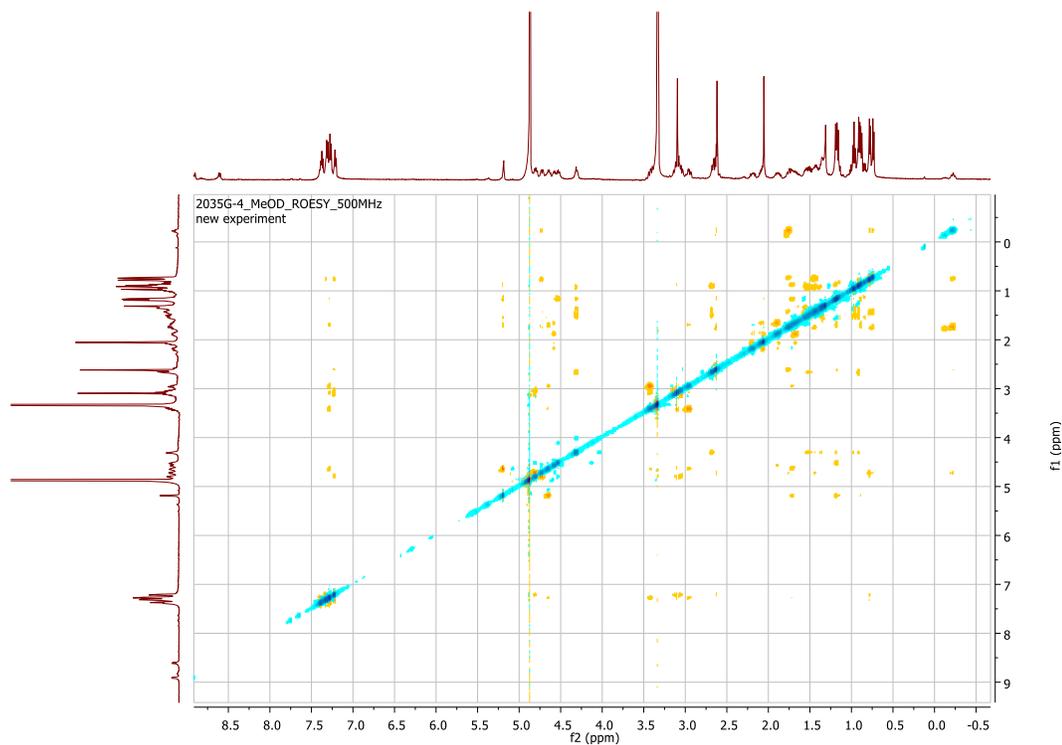


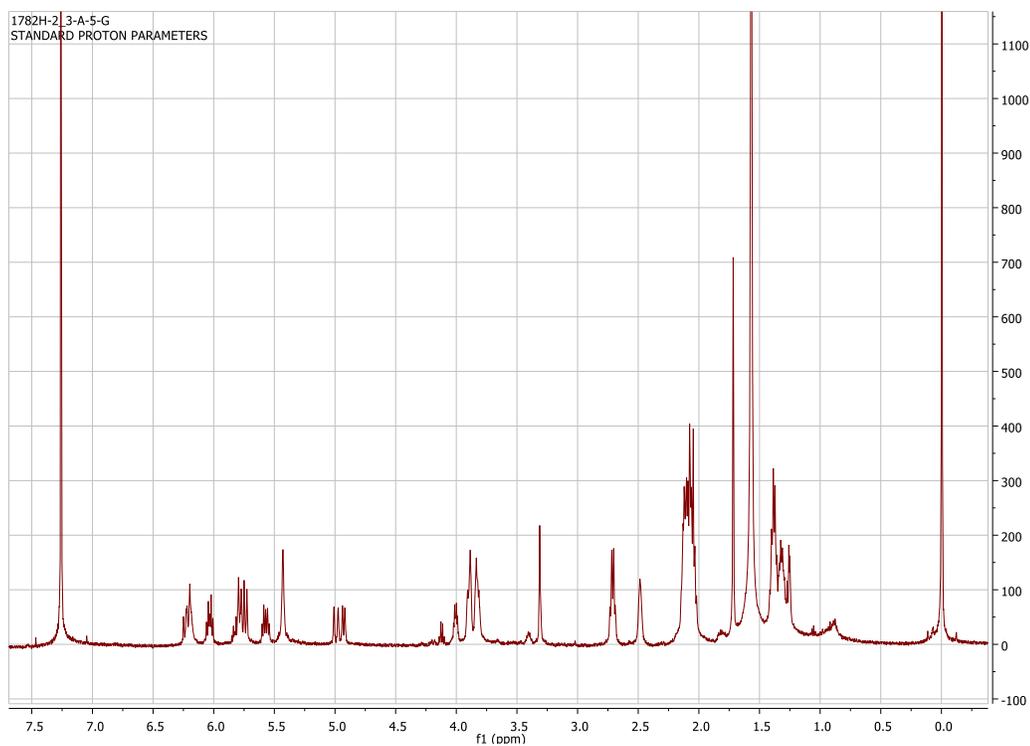
Figure 5.6.10: HMBC ( $^1\text{H}$  500 MHz,  $\text{CD}_3\text{OD}$ ) spectrum of precarriebowmide



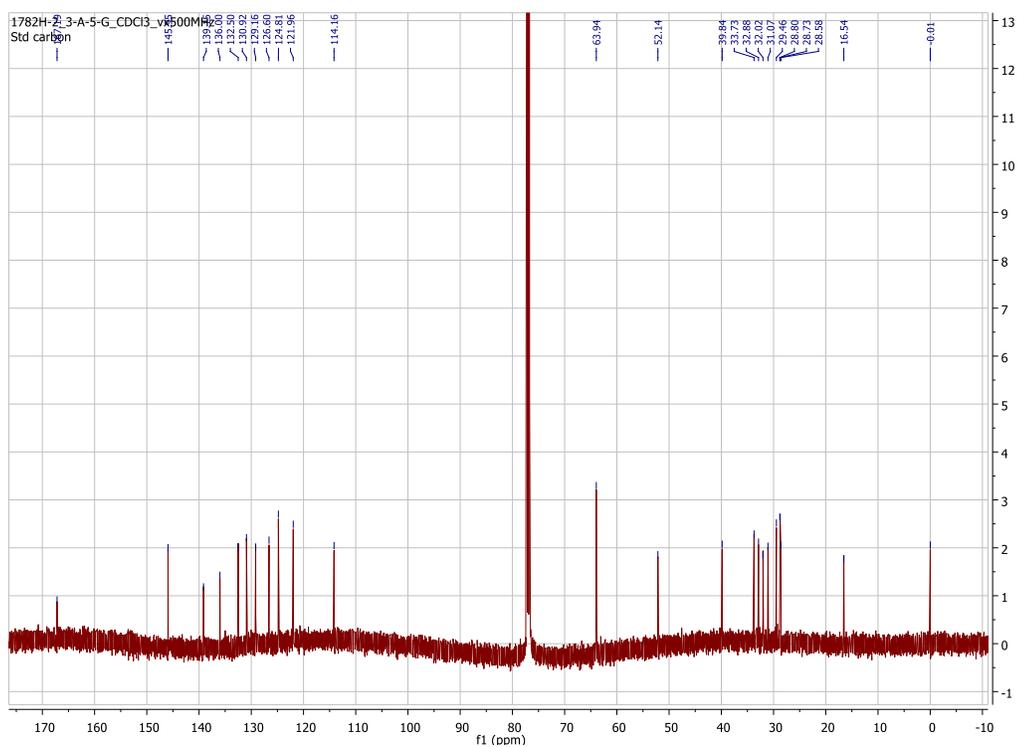
**Figure 5.6.11:** TOCSY (500 MHz, CD<sub>3</sub>OD) spectrum of precarriebowmide



**Figure 5.6.12:** ROESY (500 MHz, CD<sub>3</sub>OD) spectrum of precarriebowmide



**Figure 5.6.13:**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of mooreamide



**Figure 5.6.14:**  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of mooreamide

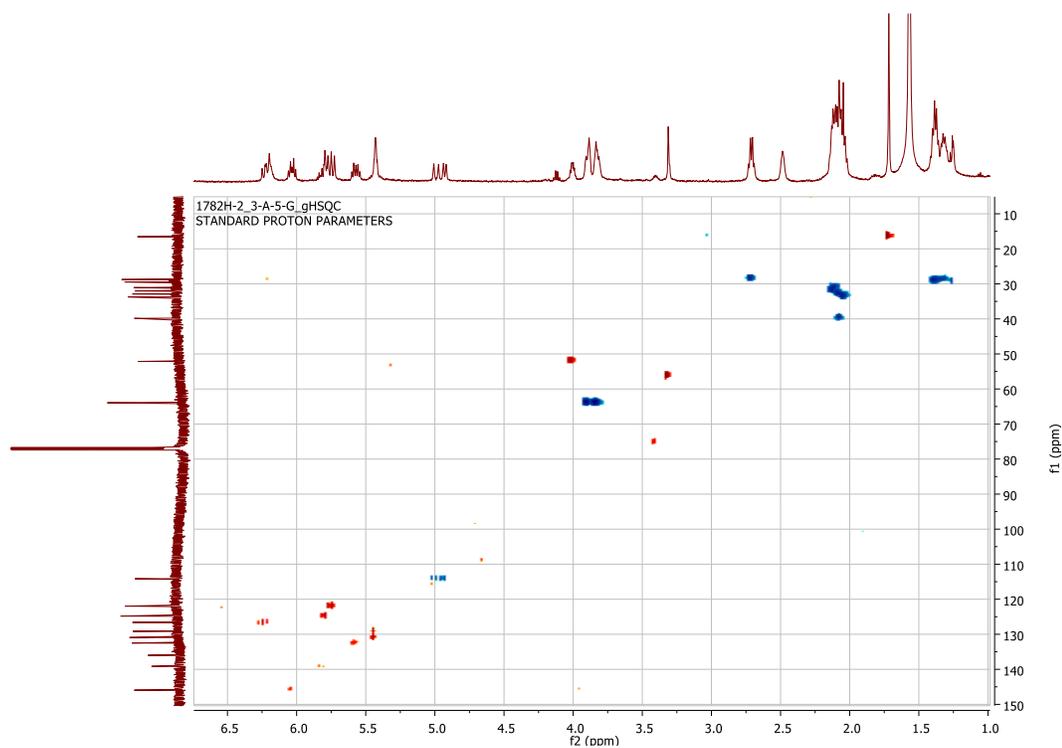


Figure 5.6.15: gHSQC ( $^1\text{H}$  500 MHz,  $\text{CDCl}_3$ ) spectrum of mooreamide

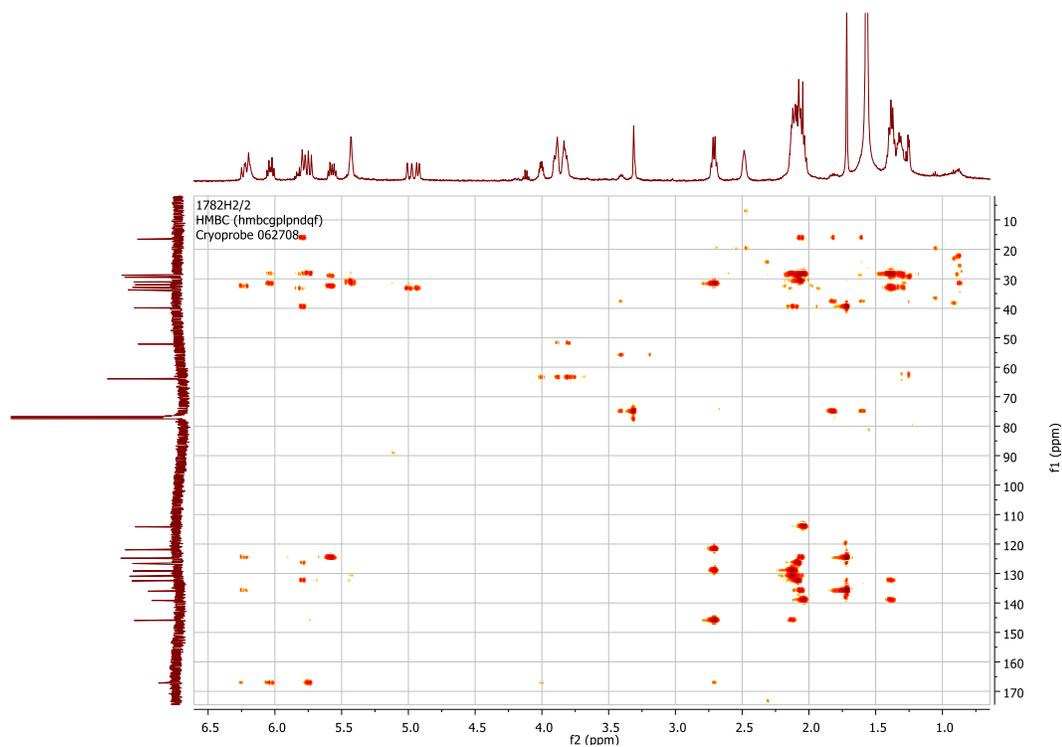
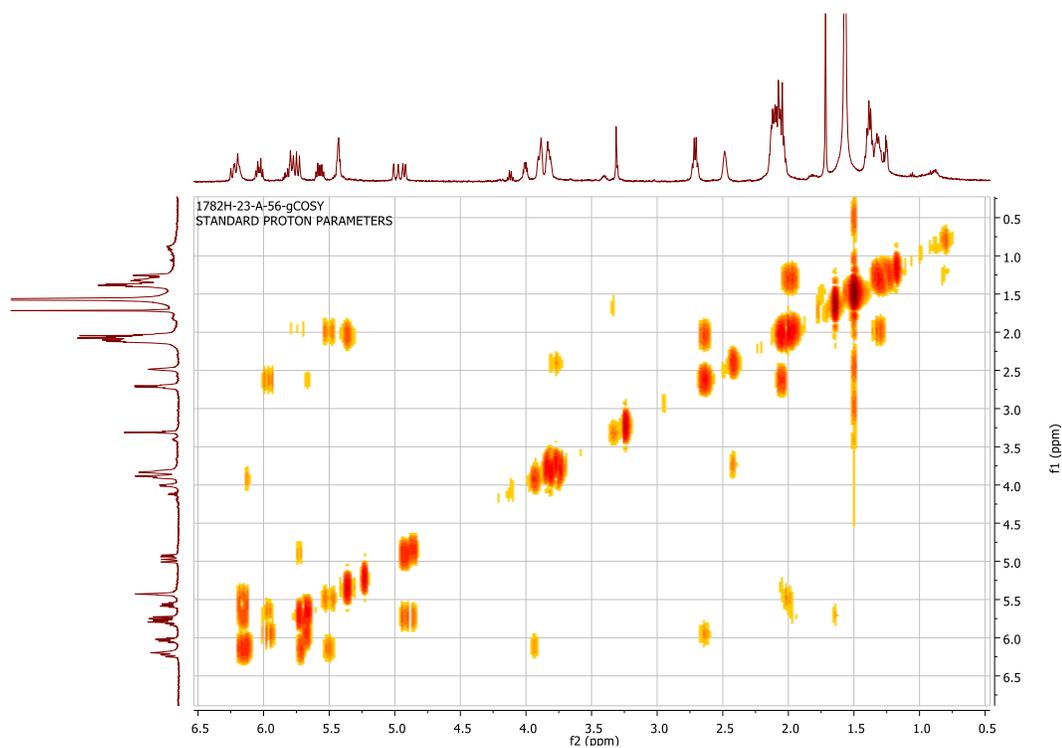
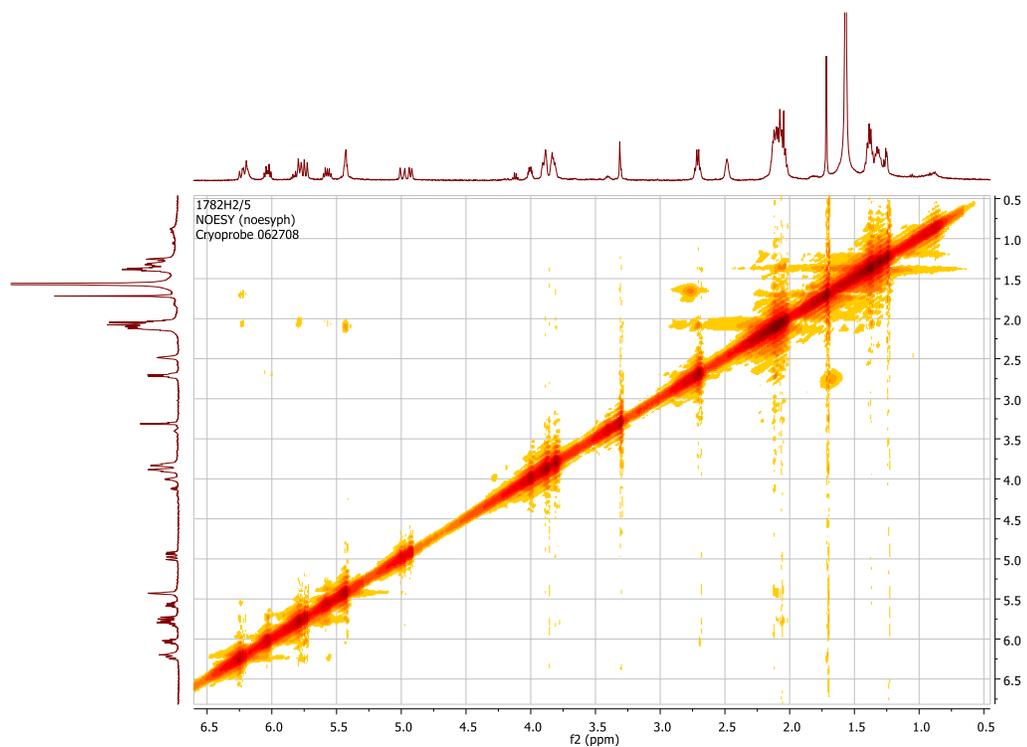


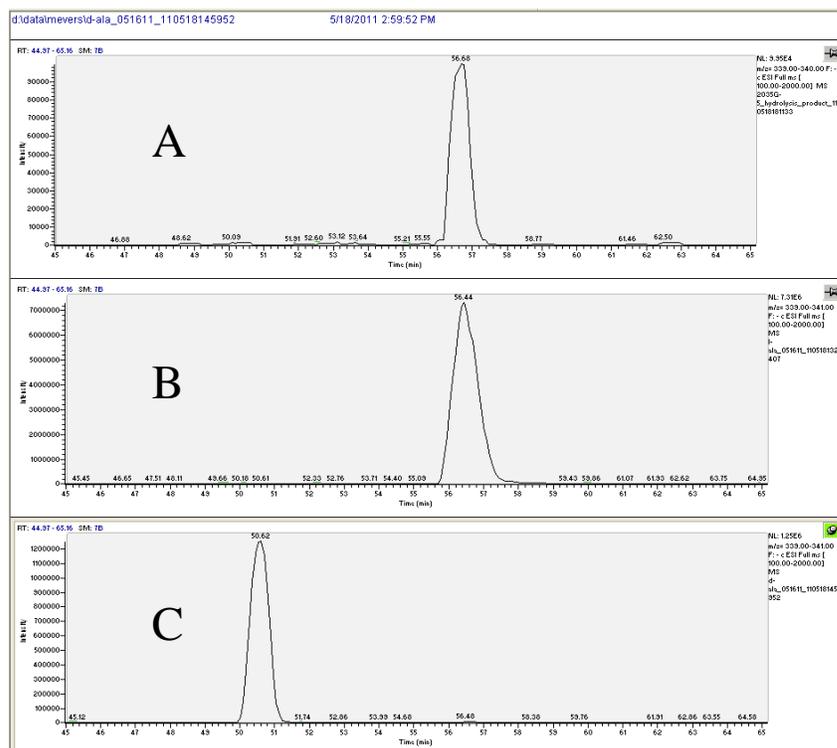
Figure 5.6.16: HMBC ( $^1\text{H}$  600 MHz,  $\text{CDCl}_3$ ) spectrum of mooreamide



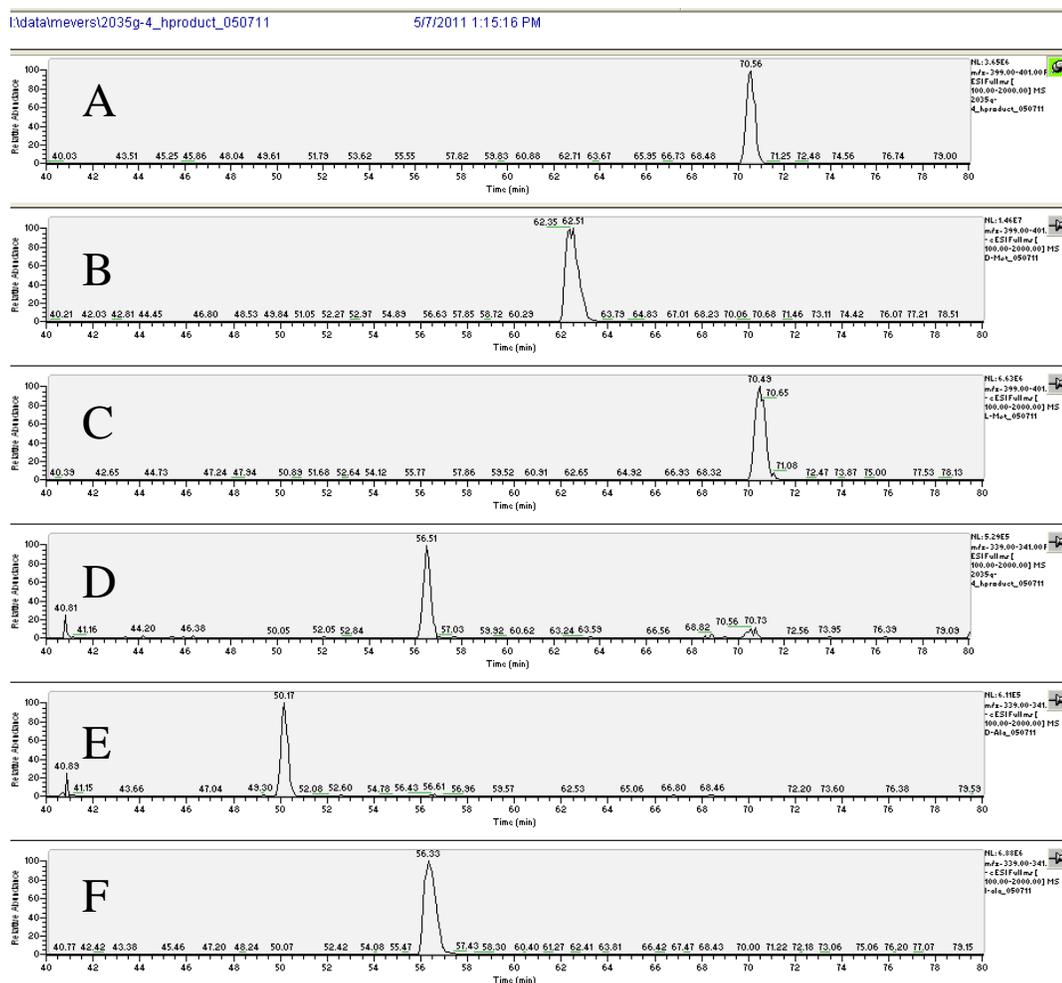
**Figure 5.6.17:** COSY (500 MHz, CDCl<sub>3</sub>) spectrum of mooramide



**Figure 5.6.18:** NOESY (600 MHz, CDCl<sub>3</sub>) spectrum of mooramide



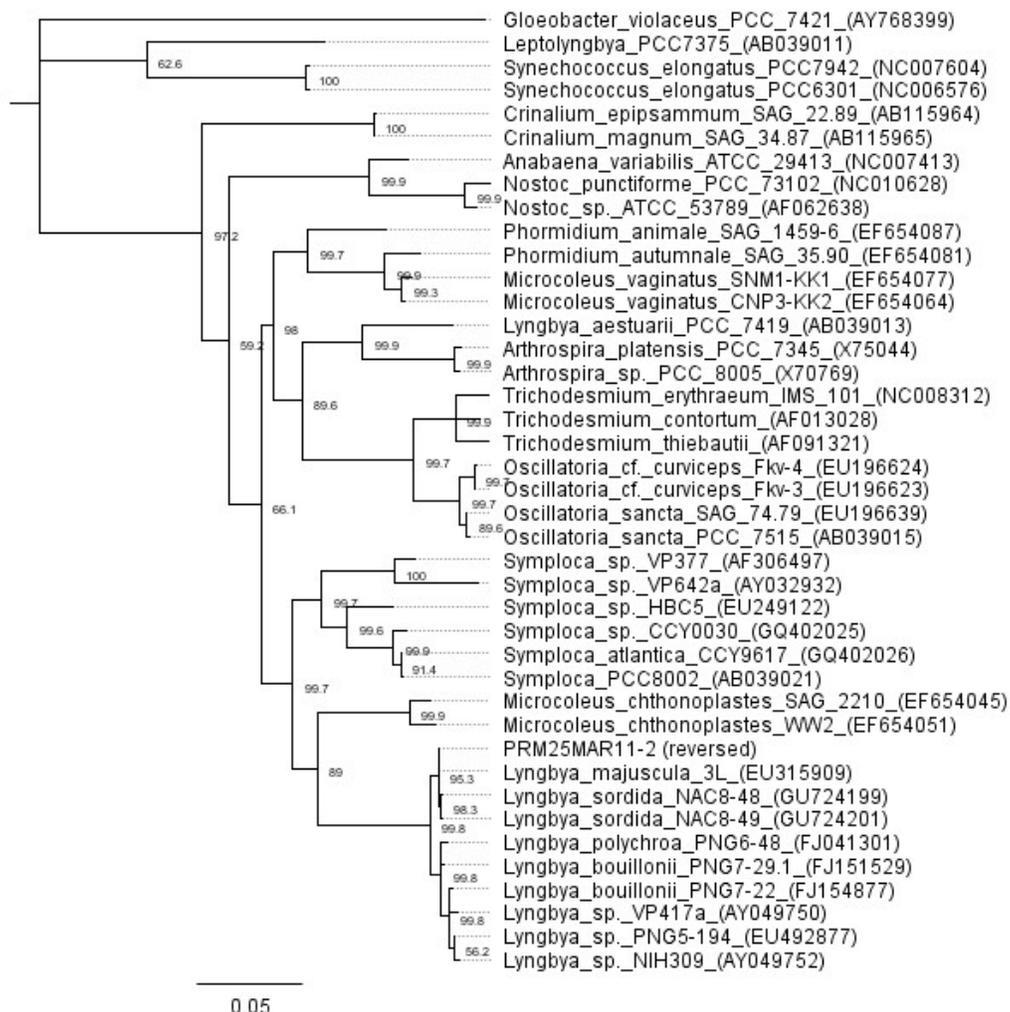
**Figure 5.6.19:** Marfey's analysis of parguerene on LCMS: (A) Marfey's derivatized hydrolysis product (ion chromatogram of 339-340 m/z); (B) L-Ala (ion chromatogram of 339-340 m/z); (C) D-Ala (ion chromatogram of 339-340 m/z)



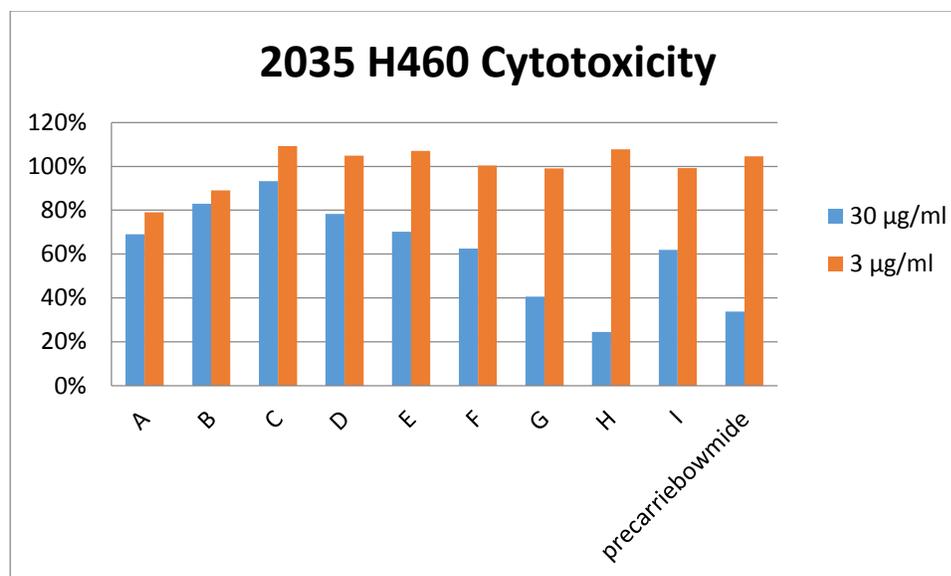
**Figure 5.6.20:** Marfey's analysis of the amino acids in precarriebowmide on LCMS: (A) Marfey's derivatized hydrolysis product (ion chromatogram of 399-341 m/z); (B) D-Met (ion chromatogram of 399-401 m/z); (C) L-Met (ion chromatogram of 399-401 m/z); (D) Marfey's derivatized hydrolysis product (ion chromatogram of 339-341 m/z); (E) D-Ala (ion chromatogram of 339-341 m/z); (F) L-Ala (ion chromatogram of 339-341 m/z)

**Figure 5.6.1:**  $^{13}\text{C}$   $\Delta$ -table for carriebowmide to precarriebowmide ( $\text{CD}_3\text{OD}$ )

Position	Carriebowmide	Precarriebowmide	Delta-Delta
1	175.6	175.5	-0.1
2	43.2	43.2	0.0
3	9.3	9.1	-0.2
4	51.9	52.0	0.1
5	34.9	35.0	0.1
6	20.6	20.6	0.0
7	13.8	13.8	0.0
8-NH			
9	174.1	174.1	0.0
10	49.0		
11	16.5	16.4	-0.1
12-NH			
13	170.7	170.6	-0.1
14	59.9	59.8	-0.1
15	37.9	37.8	-0.1
16	25.8	25.7	-0.1
17	21.9	21.8	-0.1
18	23.8	23.9	0.1
19-N-Me	29.7	29.7	0.0
20	173.9	174.0	0.1
21	52.7	52.8	0.1
22	38.7	38.8	0.1
23	137.3	137.5	0.2
24/28	130.5	130.6	0.1
25/27	129.9	129.9	0.0
26	128.3	128.3	0.0
29-NH			
		Skipped Methionine	
36	170.1	170.0	-0.1
37	62.2	62.2	0.0
38	37.3	37.7	0.4
39	138.3	138.1	0.2
40/44	130.6	130.6	0.0
41/43	130.0	129.9	-0.1
42	128.3	128.4	0.1
45-N-Me	30.0	30.1	0.1
46	172.1	172.0	-0.1
47	76.0	76.1	0.1
48	30.7	30.6	-0.1
49	19.8	19.7	-0.1
50	16.9	16.9	0.0



**Figure 5.6.21:** Phylogenetic tree for PRM-25/Mar/11-2



**Figure 5.6.22:** Cytotoxicity results for semi-crude fractions and precarrierbowmide

## 5.7 References

- (1) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. *Curr. Opin. Biotechnol.* **2010**, *21*, 787-793.
- (2) Tan, L. T. *J. Appl. Phycol.* **2010**, *22*, 659-676.
- (3) Gerwick, W. H.; Coates, R. C.; Engene, N.; Gerwick, L.; Grindberg, R. V.; Jones, A. C.; Sorrels, C. M. *Microbe* **2008**, *3*, 277-284.
- (4) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. In *Comprehensive Natural Products Chemistry II*, Moore, B., Crews, P., Eds.; Elsevier: Oxford, UK, 2010; Vol. 8, pp 141– 188.
- (5) Engene, N.; Rottacker, E. C.; Kaštovský, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komárek, J.; Gerwick, W. H. *Int. J. Syst. Evol. Micr.* **2012**, *62*, 1171-1178.
- (6) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Geoger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (7) Cardellina, J. H.; Marner, F. J.; Moore, R. E. *J. Am. Chem. Soc.* **1979**, *101*, 240-242.
- (8) Liu, Y.; Law, B. K.; Luesch, H. *Mol. Pharmacol.* **2009**, *76*, 91-104.
- (9) Maxwell, A.; Rampersad, D. *J. Nat. Prod.* **1989**, *52*, 614-618.
- (10) Pereira, A. R.; McCue, C. F.; Gerwick, W. H. *J. Nat. Prod.* **2010**, *73*, 217-220.
- (11) Kim, J. K.; Furihata, K.; Yamanaka, S.; Fudo, R.; Seto, H. *J. Antibiot.* **1991**, *44*, 553-556.
- (12) Trowitz-Kiensat, W.; Forche, E.; Wray, V.; Reichenbach, H.; Jurkiewicz, E.; Hunsmann, G.; Höfle, G. *Liebigs Ann. Chem.* **1992**, 659-664.
- (13) Jansen, R.; Reifenstahl, G.; Gerth, K.; Reichenbach, H.; Höfle, G. *Liebigs Ann. Chem.* **1983**, 1081-1095.
- (14) Andrus, M. B.; Lepore, S. D.; Turner, T. M. *J. Am. Chem. Soc.* **1997**, *119*, 12159-12169.
- (15) Andrus, M. B.; Turner, T. M.; Sauna, Z. E.; Ambudkar, S. V. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2275-2278.
- (16) Gunasekera, S.; Ritson-Williams, R.; Paul, V. J. *J. Nat. Prod.* **2008**, *71*, 2060-2063.
- (17) Jiménez, J. I.; Vansach, T.; Yoshida, W. Y.; Sakamoto, B.; Pörzgen, P.; Horgen, F. D. *J. Nat. Prod.* **2009**, *72*, 1573-1578.
- (18) Pretsch, E., Bühlmann, P., Badertscher, M. (2009) *Structure Determination of Organic Compounds: Tables of Spectral Data*, Springer Berlin
- (19) Crews, P.; Kho-Wiseman, E. *Tetrahedron Lett* **1978**, *19*, 2483-2486.

- (20) Timmers, M. A.; Dias, D. A.; Urban, S. *Mar Drugs* **2012**, *10*, 2089-2102.
- (21) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418-5423.
- (22) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J Bioorg Med Chem* **2002**, *10*, 1973-1978.
- (23) Guitierrez, M.; Pereira, A. R.; Deboni, H. M.; Ligresti, A.; Di Marzo, V.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 2313-2317.
- (24) Sitachitta, N.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 681-684.
- (25) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. In *Comprehensive Natural Products II*; Mander, L.; Liu, H. W., Eds.; Elsevier: New York, 2010; Vol. 2.06, pp 141-181.
- (26) Chhabra, A.; Haque, A. S.; Pal, R. K.; Goyal, A.; Rai, R.; Joshi, S.; Panjikar, S.; Pasha, S.; Sankaranarayanan, R.; Gokhale, R. S. *PNAS* **2012**, *109*, 5681-5686.
- (27) Dewick, P. M. In *Medicinal Natural Products: A Biosynthetic Approach*; John Wiley & Sons Ltd: West Sussex, 2009; Vol. 3, pp 39-131.
- (28) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (29) Dorrestein, P. C.; Blackhall, J.; Straight, P. D.; Fischbach, M. A.; Garneau-Tsodikova, S.; Edwards, D. J.; McLaughlin, S.; Lin, M.; Gerwick, W. H.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. *Biochemistry* **2006**, *45*, 1537-1546.
- (30) McCarthy, J. G.; Eisman, E. B.; Kulkarni, S.; Gerwick, L.; Gerwick, W. H.; Wipf, P.; Sherman, D. H.; Smith, J. L. *ACS Chem Biol* **2012**, *7*, 1994-2003.
- (31) Gu, L.; Wang, B.; Kulkarni, A.; Gehret, J. J.; Lloyd, K. R.; Gerwick, L.; Gerwick, W. H.; Wipf, P.; Håkansson, K.; Smith, J. L.; Sherman, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 16033-16035.
- (32) Castenholz, R. W.; Rippka, R.; Herdman, M. In *Bergey's Manual of Systematic Bacteriology*; Boone, D. R.; Castenholz, R. W., Eds.; Springer: New York, 2001; Vol. 1, pp 473-599.
- (33) Komárek, J.; Anagnostidis, K. In *Süßwasserflora von Mitteleuropa*; Büdel, B.; Gärtner, G.; Krienz, L.; Schagerl, M., Eds.; Gustav Fischer: Jena, 2005; Vol. 19/2, pp. 576-606.
- (34) Drummond, A. J.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Duran, C.; Field, M.; Heled, J.; Kearse, M.; Markowitz, S.; Moir, R.; Stones-Haras, S.; Thierer, T.; Wilson, A. *Geneious*, Version 5.5; Biomatters: Auckland, New Zealand, 2011. Available from <http://www.geneious.com>

- (35) Katoh, K.; Misawa, K.; Kuma, K. I.; Miyata, T. *Nucleic Acids Res.* **2002**, *30*, 3059-3066.
- (36) Posada, D. *Mol. Biol. Evol.* **2008**, *25*, 1253-1256.
- (37) Guindon, S.; Gascuel, O. *Syst. Biol.* **2003**, *52*, 696-704.
- (38) Huelsenbeck, J. P.; Ronquist, F. *Bioinformatics* **2001**, *17*, 754-755.

## Chapter 6:

### Structure-Activity Relationships of the Lyngbyamide Class of Marine Natural Product

#### 6.0.1 Abstract

Alkyl amides are a growing class of marine natural products that have both physiological and pharmacological importance. In recent years, several alkyl amides have been isolated from marine cyanobacteria and exhibit a broad range of biological activities, including ion channel modulation, brine shrimp toxicity and cannabinoid receptor binding ability. These interesting biological activities, coupled with the relative synthetic tractability of these alkyl amides, led to the design of a structure-activity relationship study into the lyngbyamides. In total, 50 analogs were synthesized, designed to probe the importance of several structural characteristics of the lyngbyamides. These compounds were tested in a wide array of biological assays, and a subset was found to possess strong activity in the stabilization of cathepsin L-mediated proteolysis, brine shrimp toxicity, and reduction in surface tension.

#### 6.1 Introduction

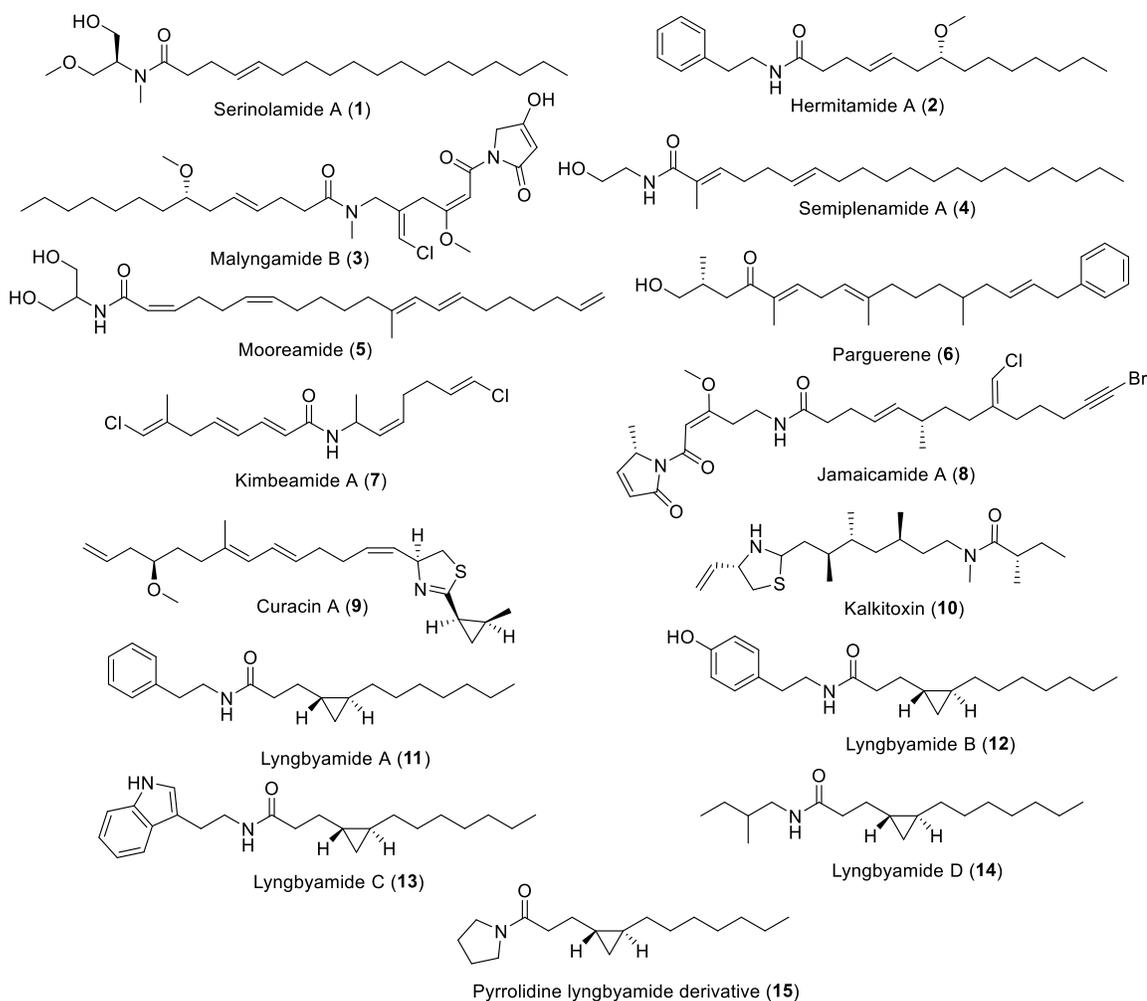
Marine filamentous cyanobacteria are prolific producers of interesting secondary metabolites, many of which possess intriguing biological activities.<sup>1-4</sup> One growing subset of cyanobacterial natural products is the alkyl amides, which includes the serinolamides (**1**),<sup>5</sup> hermitamides (**2**),<sup>6</sup> malyngamides (**3**),<sup>7</sup> semiplenamides (**4**),<sup>8</sup> mooreamide (**5**), parguerene (**6**),<sup>9</sup> the kimbeamides (**7**),<sup>10</sup> curacin A (**8**),<sup>11</sup> kalkitoxin (**9**),<sup>12</sup> the jamaicamides (**10**),<sup>13</sup> and the lyngbyamides (**11-14**).<sup>14,15</sup> The majority of these metabolites are produced by only one genus, *Moorea* (formerly *Lyngbya*),<sup>16</sup> and are reported to exhibit a broad range of biological activities, such as brine fish toxicity, gold fish toxicity, cannabinoid receptor binding, ion channel modulation, and moderate cancer cell cytotoxicity.<sup>17</sup> However, this subset of natural products is rather understudied, as most

of the above mentioned metabolites have not been exhaustively evaluated for their biological potential due to limited sample quantities, and thus requiring either recollection or the completion of total syntheses to provide the material needed for continued biological analysis.

One particularly interesting family of alkyl amides is the lyngbyamides, which consist of a twelve carbon fatty acid tail group, functionalized with a *trans*-cyclopropyl ring, and coupled to non-polar biogenic amines through an amide bond linkage. Lyngbyamide A (**11**) (also known as grenadamide) was originally isolated in 1998 by Sitachitta and Gerwick from a Grenada collection of *Moorea*, and was found to exhibit both moderate brine shrimp toxicity ( $LD_{50} = 5 \mu\text{g/mL}$ ) and cannabinoid receptor binding activity ( $K_i = 4.7 \mu\text{M}$ ).<sup>14</sup> Three additional natural analogs, along with lyngbic acid, were isolated by Nannini and Gerwick in 2002 from a Madagascar collection of *Moorea*, which also exhibits similar brine shrimp toxicity; however these analogs were not evaluated in the cannabinoid receptor binding assay.<sup>15</sup> Further semi-synthetic investigations into this family using the isolated lyngbic acid to make a pyrrolidine derivative of lyngbyamide A (**15**), resulted in a semi-synthetic derivative with 10-fold increased brine shrimp toxicity ( $LD_{50} = 0.3 \mu\text{g/mL}$ ).<sup>15</sup> This increase in potency was hypothesized to be attributed to the size of the pyrrolidine ring in relation to the tryptamine and tyramine functionalities, however further studies were needed to probe the validity of this hypothesis.<sup>16</sup>

In the present investigation into structure-activity relationships of the lyngbyamide A class of metabolite, a number of synthetic analogs were designed and synthesized in order to obtain a better understanding of the key structural features responsible for the observed brine shrimp toxicity and cannabinoid receptor binding activity. Additionally, the first round of analogs was also evaluated in a variety of biological assays, such as for cytotoxicity to H-460 human lung cancer cells, as ion channel modulators, as anti-inflammatory agents, and for the activation/inhibition of the cathepsin L protease enzyme, in order to obtain a broader understanding of their pharmaceutical potential. Through this project, we obtained a better understanding of both key structural features

responsible for the brine shrimp toxicity as well as shed light on the potential natural functions of tertiary amide secondary metabolites.



**Figure 6.1:** Alkyl amides from marine cyanobacteria

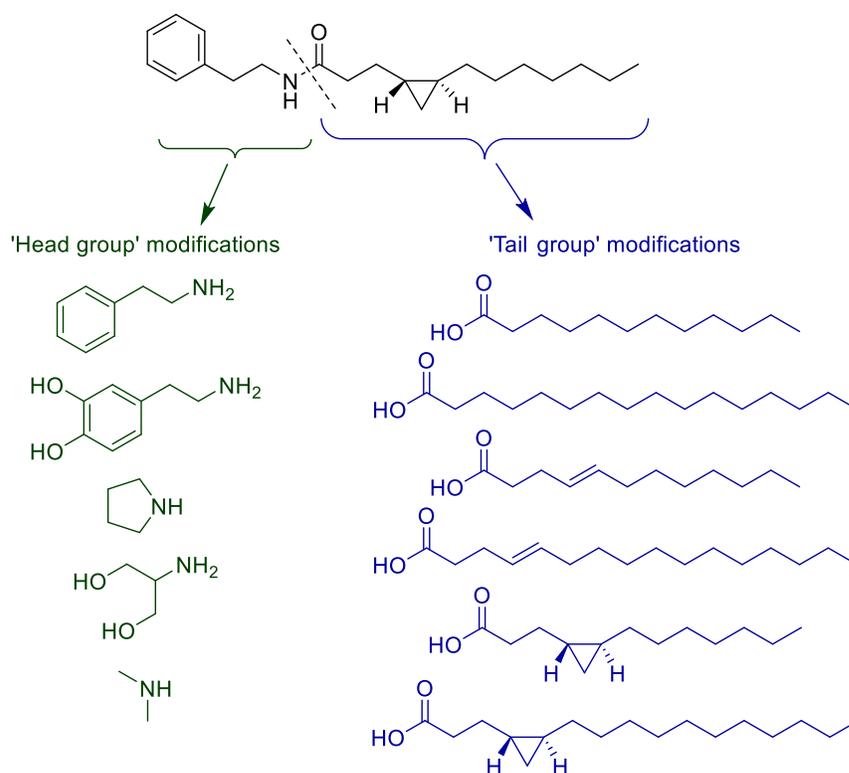
## 6.2 Results and Discussion

### 6.2.1 Synthetic Approach to Synthesizing Lyngbyamide Analogs

Inspired by the combination of interesting biological activity and the synthetic tractability of the lyngbyamide family of compounds, we designed a synthesis that could easily be modified to allow us to obtain a large number of analogs. Synthetically, lyngbyamide A has one obvious point of connection, the amide bond, which splits the molecule into two portions, the head and tail groups.

In many of the cyanobacterial alkyl amide natural products, the tail groups vary in carbon chain length from 12 to 20, and are often times functionalized with cyclopropane rings, *O*- and *C*-methyl groups, and various degrees of unsaturation. As for the head groups, there are several different types naturally observed, primarily deriving from either biogenic amines or amino glycerol. Thus, the first round of synthetic analogs were designed to probe the effect that the fatty acid chain length (12 and 16), fatty acid functionality (saturated, C4 *trans*-unsaturation, and C4 *trans*-cyclopropanation), and size/polarity of the head group amines (phenethylamine, dopamine, pyrrolidine, amino glycerol, and *N,N*-dimethyl amine) had on a number of biological assays, including brine shrimp toxicity, H-460 human lung cell carcinoma cytotoxicity, nitric oxide production in rat macrophages, ion channel modulation, cathepsin L activation/inhibition, bone regeneration, and in a number of neuro-receptor binding assays (serotonin, cannabinoid, and dopamine)(figure 6.2).

The synthetic approach to these analogs was split into three different target groups based on the functionality of their tail group (figure 6.3). The synthesis of the saturated fatty amide involved only one step, which was the coupling between the fatty acids [lauric acid (**16**) and palmitic acid (**17**)] and the five different amines [amino glycerol (**a**), phenethylamine (**b**), pyrrolidine (**c**), *N,N*-dimethylamine (**d**), and dopamine (**e**)] by coupling reagents *N,N*-diisopropylcarbodiimide (DIPC) and hydroxybenzotriazole (HOBt).<sup>18</sup> This produced the desired compounds **18a-e** and **19a-e** in yields ranging from 50-85%.

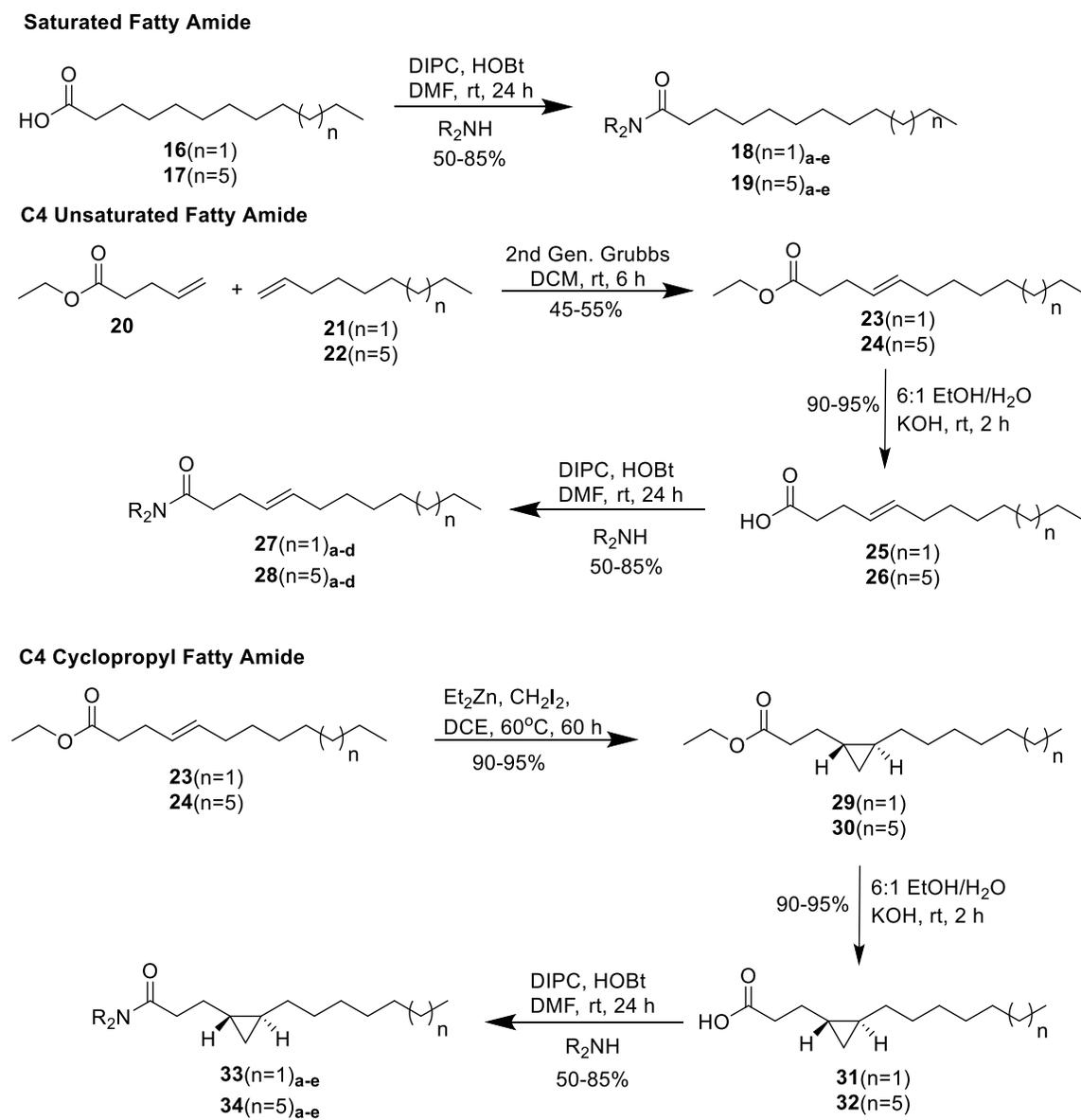


**Figure 6.2:** Round 1 analogs of lyngbyamide A

The second target group involved analogs with C4-unsaturation, and began with an olefin coupling reaction between ethyl pentenoate (**20**) and either nonene (**21**) or tridecene (**22**), using the Grubbs 2<sup>nd</sup> generation catalyst.<sup>19</sup> The carboxyl groups in the resulting ethyl esters, **23** and **24**, were then deprotected using mild base conditions, and subsequently linked to each of the amines using the same coupling reagents mentioned above, producing analogs **27a-d** and **28a-d**.<sup>20</sup>

The final target group included the analogs with the C4-*trans*-cyclopropyl ring, and began with the unsaturated ethyl esters **23** and **24**. The *trans*-cyclopropyl rings were installed by using a modified Simmons-Smith reaction, involving diethylzinc and diiodomethane in 1,2-dichloroethane (DCE).<sup>21</sup> In order to get the reaction to go to completion, both of the reagents were added at three different time points during the reaction (0, 24, and 48 h), resulting in a yield greater than 95% for compounds **29** and **30**. Although the Simmons-Smith reaction of *trans*-alkenes strongly favors the

formation of *trans*-cyclopropyl rings, there is no enantiomeric control; therefore, the resulting product is a mixture of enantiomers, which were not separated from one another. The carboxyl groups were then deprotected and coupled to the amines as described above to afford analogs **33a-d** and **33a-d**.



**Figure 6.3:** General synthetic scheme

### 6.2.2 Biological Evaluation of the First Round of Analogs

From the above synthesis, twenty-six analogs and ten intermediates were prepared and evaluated in range of bioassays, including for brine shrimp toxicity, H-460 human lung cell carcinoma cytotoxicity, nitric oxide production, ion channel modulation, cathepsin L protease activation or inhibition, cancer cell selectivity, and for neuro-receptor binding activity (cannabinoid, serotonin, and dopamine receptors). As expected, many of the analogs exhibited moderate to potent toxicity against brine shrimp (*Artemia salina*), and some showed cannabinoid receptor binding activity. Excitingly though, many of the analogs also exhibited strong activation of the cathepsin L protease enzyme. None of the analogs showed cytotoxicity against the H-460 cell line, cancer cell selectivity against numerous cell lines on a disc diffusion assay, nitric oxide production, ion channel modulation, or neuro-receptor binding activity against the serotonin or dopamine receptors.

Delving a little deeper into the observed activation of cathepsin L, and the brine shrimp toxicity revealed a potential key structural functionality that is important for these activities, the presence of tertiary amides. All of the tertiary amides and only a couple of the secondary amides possessed activity in both of these assays (figure 6.4), with most 100% lethal to brine shrimp at 3  $\mu\text{g/mL}$ . Furthermore, there appears to be no major differences between each of the tail group functionalities (saturated, C4-unsaturated, and C4-cyclopropyl ring); however, longer carbon chained tail groups may be slightly more active, but a larger sample size is needed in order to confirm this observation.

### Round 1: Cathepsin L and Brine Shrimp Activity

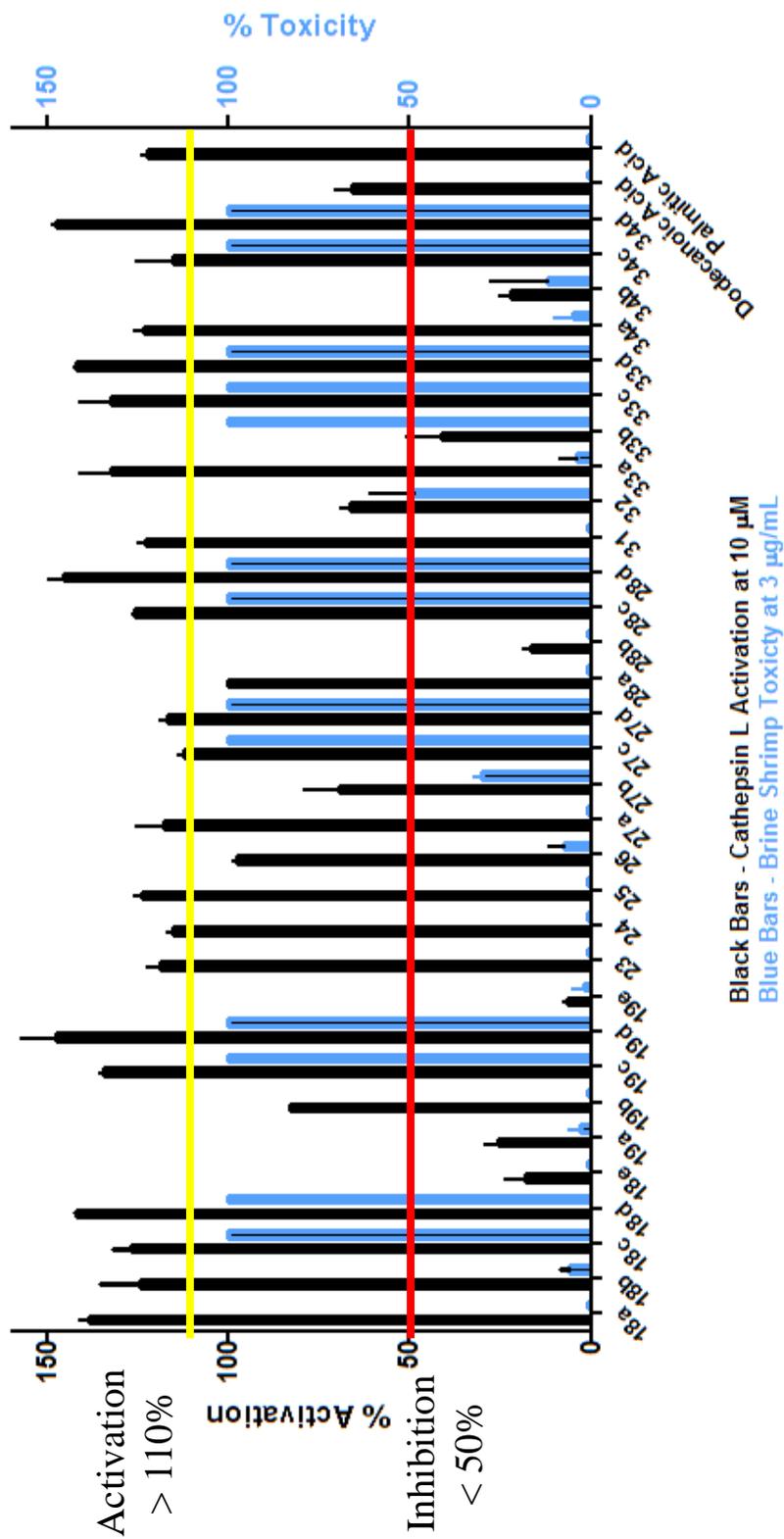
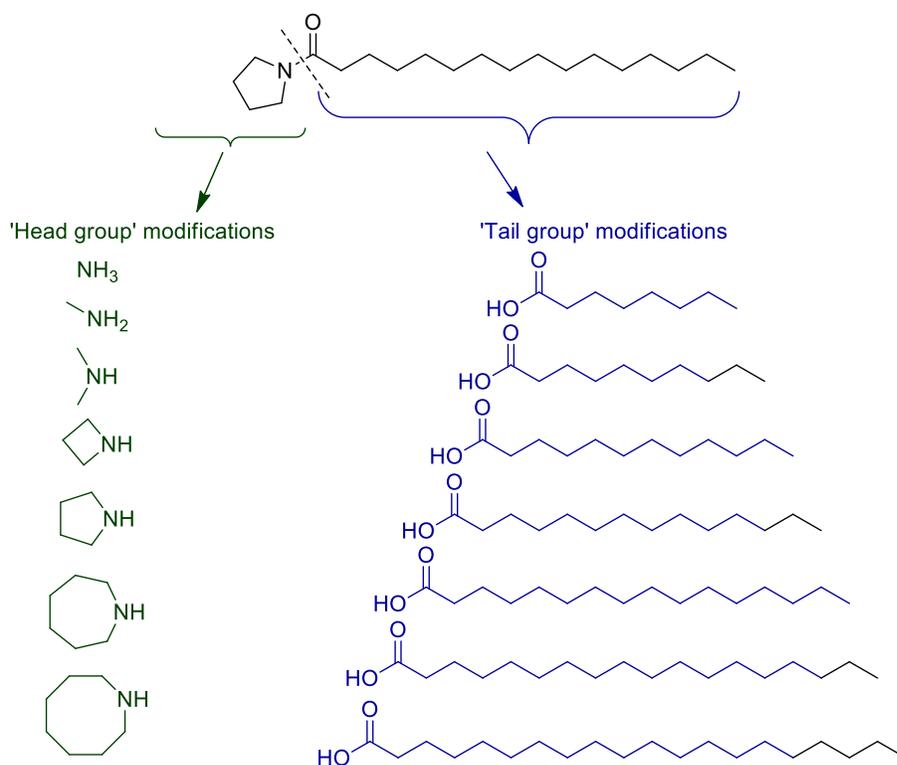


Figure 6.4: Cathepsin L modulation and brine shrimp toxicity for round 1 analogs

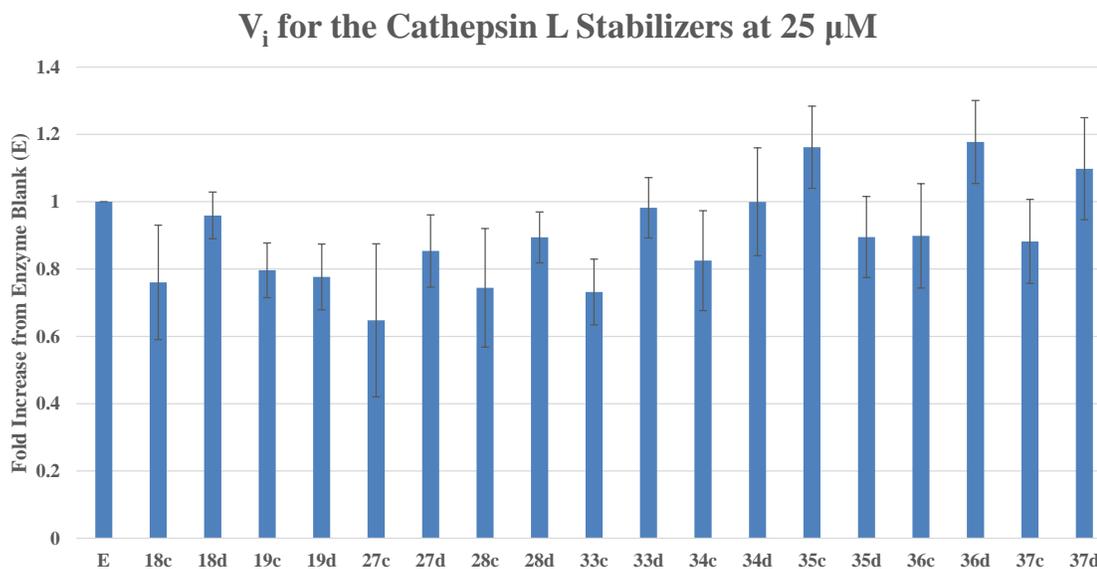


**Figure 6.5:** Round 2a (right) and 2b (left) of lyngbyamide analogs

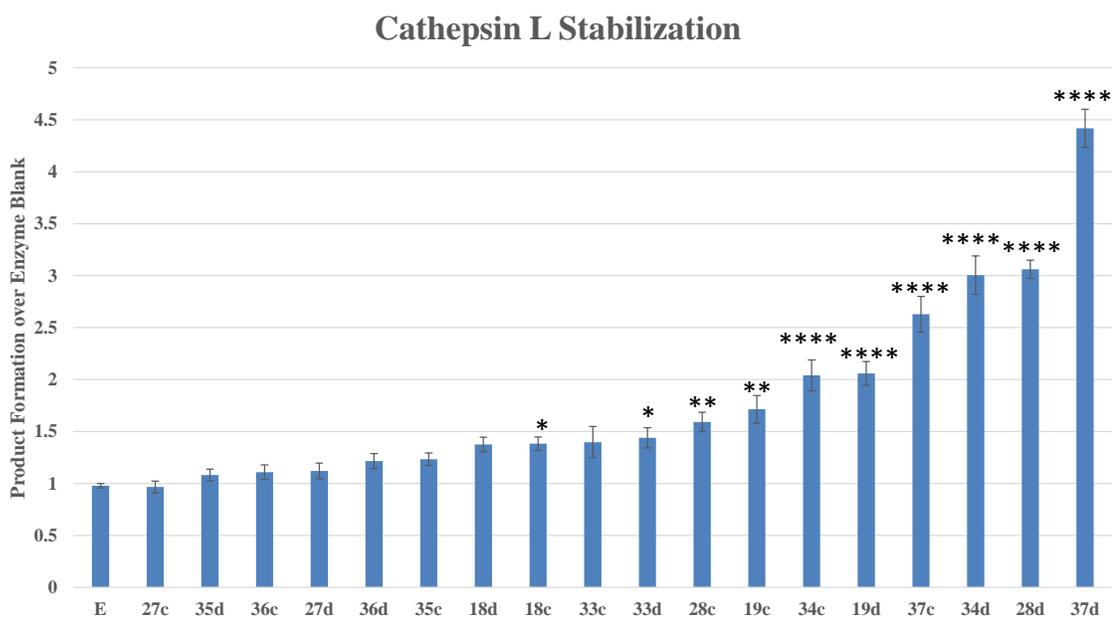
### 6.2.3 The Synthesis and Activity of Rounds 2A and 2B

Two additional rounds (2A and 2B) of analogs were designed based on the observed activity in the first round of compounds (figure 6.5). Where round 2A included saturated tail groups ranging in chain length from eight to twenty carbons, linked to either the *N,N*-dimethylamine or pyrrolidine head groups, round 2B involved coupling palmitic acid (C16) with amines varying in bulkiness, including ammonia, *N*-methylamine, *N,N*-dimethylamine, azetidine, pyrrolidine, hexamethyleneimine, and heptamethyleneimine. Each of these new analogs was synthesized using the coupling reagents HOBt and DIPC, as described above.<sup>18</sup> Furthermore, each new round of analogs was designed to probe separate questions pertaining to key functionalities, with hopes of improving and better understanding the observed activity.

Preliminary screening of each of these new analogs alongside the twelve tertiary amides from the first round showed that many of these compounds exhibit both brine shrimp toxicity and cathepsin L activation. Quantification of this observed activity by dose responses was easily obtained for the brine shrimp toxicity (table 6.1); however, when trying to obtain  $EC_{50}$  values for the cathepsin L assay, a number of complications arose with the biggest being the day-to-day irreproducibility of the dose response curves. Either solubility issues or variations in the timing of the assay likely caused this problem. Trying to eliminate the latter issue, the assay protocol was modified to ensure that the mixture of the substrate, enzyme, and analog occurred at the same time consistently using a kinetic assay setup. Upon starting the assay, the production of product [7-amino-4-methylcoumarin (AMC)] was monitored by fluorescence readings every minute for 150 min. Analyzing the initial velocity from the resulting progress curves revealed that there were no statistical differences between the analogs and the enzyme blank, and thus it is unlikely that these compounds are actually activating cathepsin L over its native initial velocity rate (figure 6.6). However, the progress curve of several of the analogs was similar to those reported for common biological surfactants acting on other enzymes,<sup>21</sup> thus suggesting that these analogs might be having a stabilizing effect on the enzyme by somehow keeping the enzyme active and in solution for longer periods of time. Quantification of this activity was accomplished by comparing the formation of AMC by the analogs to that of the enzyme blank at the 120 min time point (before the assay becomes substrate limited). This verified that several of the synthetic analogs do have significant capabilities to stabilize the cathepsin L enzyme, some even generating a fourfold increase in product (table 6.1) over the enzyme system in buffer alone (figure 6.7).



**Figure 6.6:** Initial velocity of the analogs compared to the enzyme blank (E)



**Figure 6.7:** Stabilization of cathepsin L  
 $*p \leq 0.05$ ,  $**p \leq 0.01$ ,  $****p \leq 0.001$  – ANOVA followed by TUKEY'S Multiple Comparison Method

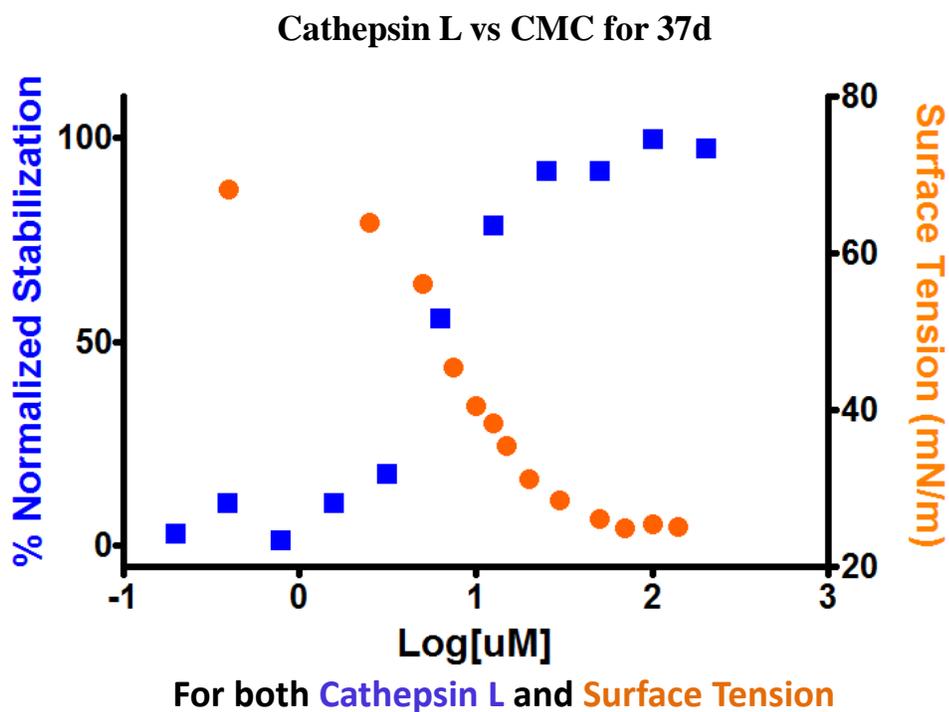


#### 6.2.4 Analysis of the Cathepsin L Stabilization, Brine Shrimp Toxicity, and CMC Data

Analysis of the brine shrimp LD<sub>50</sub>, cathepsin L stabilization, and CMC data for each of the tertiary amides showed similar trends in activity with correlations to specific structural features. Analogs with 14 carbon fatty acid chains had the highest activity and activity decreased corresponding to either increasing or decreasing chain length of the tail groups (C8 < C10 < C12 < C14 > 16). However, analogs with either 18 or 20 carbon chain length tail groups had significant solubility issues in all of the assay solvents (DMSO, buffered pH 5.5 water, and artificial seawater); therefore, the data for these analogs were variable and suspect. Another important structural feature for potent activity was the presence of tertiary amide head groups between the size of the *N,N*-dimethylamine and the pyrrolidine group. Tertiary head groups of larger size (hexamethyleneimine, and heptmethyleneimine) were completely inactive in both assays.

Statistical analysis comparing each of these different biological and biophysical data sets to one another showed that there were significant correlations between them. A rank linear regression analysis showed there to be a correlation between the brine shrimp toxicity and CMC data sets ( $p = 0.021$ ) as well as between brine shrimp toxicity and cathepsin L stabilization ( $p = 0.009$ ). Furthermore, a multiple linear regression model showed an adjusted R<sup>2</sup> of 0.773 demonstrating that 77.3% of the variability in cathepsin L stabilization can be explained by the regression analysis. The ANOVA table  $p$ -value <0.0001 shows that at least one independent variable is significant in the model. Examining the model parameters reveals that structures with a carbon chain length between 14 and 16 have the greatest effect on cathepsin L stabilization activity. That is to say that structures with carbon chain lengths of 14-16 have greater cathepsin L stabilization compared to those structures with carbon chain lengths of 8-12. Critical micelle concentration was also a significant variable ( $p = 0.049$ ); however, it appears clear that structural class is a more important variable in determining cathepsin L stabilization. While an adjusted R<sup>2</sup> of the model of 0.773 explains a great deal of the variability, 22.7% of the variability in cathepsin L

stabilization is not explained by this model and other variables that we have not tested may be involved. While our goal was to identify which of the independent variables analyzed are significantly correlated with cathepsin L stabilization, repeating the model with only the significant variables would be necessary to obtain the most appropriate regression equation.



**Figure 6.9:** Comparison of cathepsin L dose response (blue) and surface tension readings (orange) over the same logarithmic concentration scale for compound **37d**

**Table 6.1:** Biological data for the lyngbyamide A (**11**) analogs

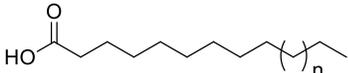
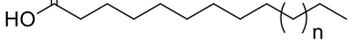
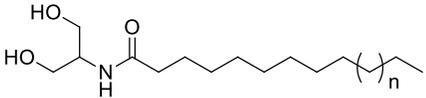
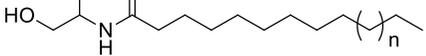
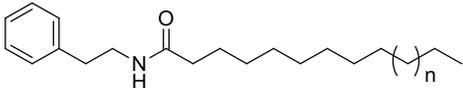
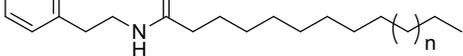
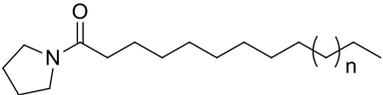
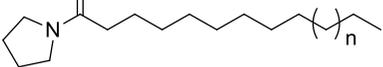
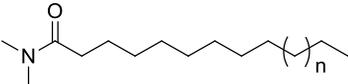
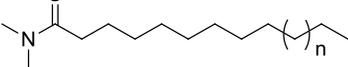
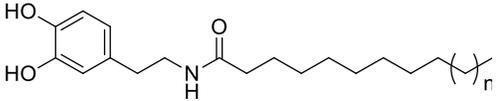
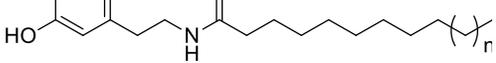
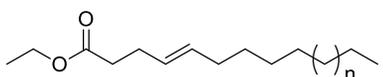
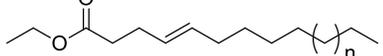
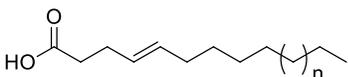
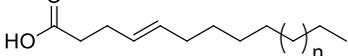
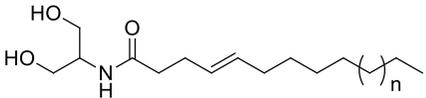
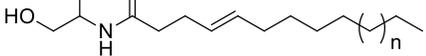
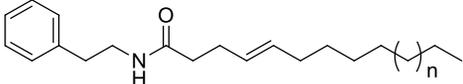
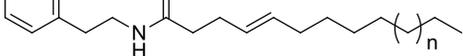
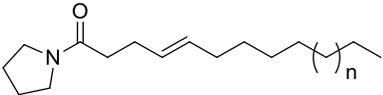
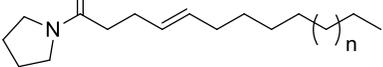
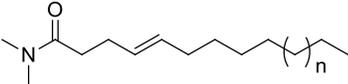
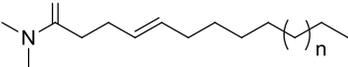
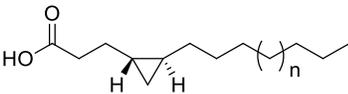
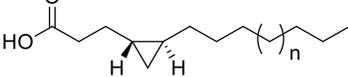
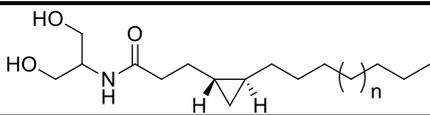
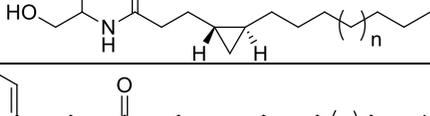
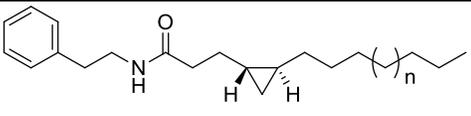
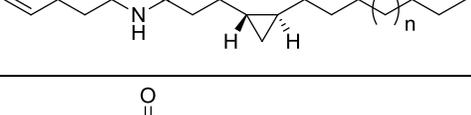
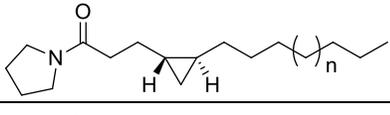
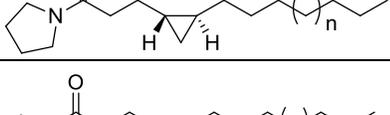
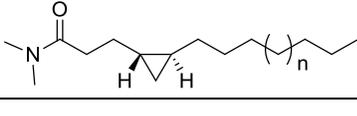
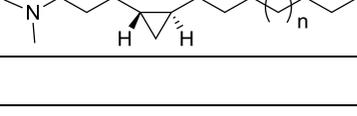
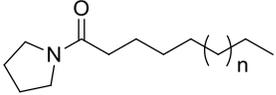
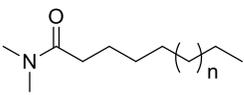
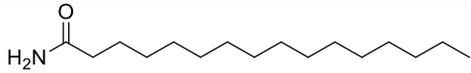
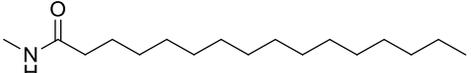
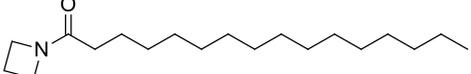
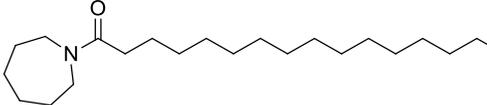
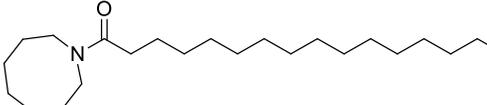
Compound	Chain length (n=)	Cathepsin L stabilization <sup>a</sup>	Brine Shrimp Toxicity (μM)	CMC (μM)
Round 1				
	1 ( <b>16</b> )	NA	NA	NT
	5 ( <b>17</b> )	NA	NA	NT
	1 ( <b>18a</b> )	NA	NA	NT
	5 ( <b>19a</b> )	NA	NA	NT
	1 ( <b>18b</b> )	NA	NA	NT
	5 ( <b>19b</b> )	NA	NA	NT
	1 ( <b>18c</b> )	1.38 ± 0.07	0.931 ± 0.092	27.6
	5 ( <b>19c</b> )	1.71 ± 0.13	0.766 ± 0.003	50-80
	1 ( <b>18d</b> )	1.38 ± 0.07	1.64 ± 0.031	34.9
	5 ( <b>19d</b> )	2.06 ± 0.11	0.923 ± 0.009	50-80
	1 ( <b>18e</b> )	NA	NA	NT
	5 ( <b>19e</b> )	NA	NA	NT
	1 ( <b>23</b> )	NA	NA	NT
	5 ( <b>24</b> )	NA	NA	NT
	1 ( <b>25</b> )	NA	NA	NT
	5 ( <b>26</b> )	NA	NA	NT
	1 ( <b>27a</b> )	NA	NA	NT
	5 ( <b>28a</b> )	NA	NA	NT
	1 ( <b>27b</b> )	NA	> 10	NT
	5 ( <b>28b</b> )	NA	NA	NT
	1 ( <b>27c</b> )	0.97 ± 0.06	8.32 ± 0.49	61.3
	5 ( <b>28c</b> )	1.59 ± 0.09	0.682 ± 0.009	41.3
	1 ( <b>27d</b> )	1.12 ± 0.08	10.42 ± 0.23	50-80
	5 ( <b>28d</b> )	3.06 ± 0.09	1.11 ± 0.099	34.8
	1 ( <b>31</b> )	NA	NA	NT
	5 ( <b>32</b> )	NA	> 11	NT

Table 6.1: continued

Compound	Chain length (n=)	Cathepsin L stabilization <sup>a</sup>	Brine Shrimp Toxicity ( $\mu\text{M}$ )	CMC ( $\mu\text{M}$ )
	1 ( <b>33a</b> )	NA	NA	NT
	5 ( <b>34a</b> )	NA	NA	NT
	1 ( <b>33b</b> )	NA	> 1	NF
	5 ( <b>34b</b> )	NA	NA	NT
	1 ( <b>33c</b> )	$1.40 \pm 0.15$	$1.55 \pm 0.086$	40.3
	5 ( <b>34c</b> )	$2.04 \pm 0.14$	$1.00 \pm 0.55$	32.5
	1 ( <b>33d</b> )	$1.44 \pm 0.10$	$2.88 \pm 0.23$	50-80
	5 ( <b>34d</b> )	$3.00 \pm 0.18$	$1.03 \pm 0.039$	34.9
Round 2A				
	1 ( <b>35c</b> )	$1.23 \pm 0.06$	NA	NF
	3 ( <b>36c</b> )	$1.11 \pm 0.07$	$5.21 \pm 0.35$	>140
	7 ( <b>37c</b> )	$2.63 \pm 0.17$	$0.451 \pm 0.022$	13.9
	13 ( <b>38c</b> ) <sup>b</sup>	NT	NA	NT
	15 ( <b>39c</b> ) <sup>b</sup>	NT	NA	NT
	1 ( <b>35d</b> )	$1.08 \pm 0.06$	NA	NF
	3 ( <b>36d</b> )	$1.22 \pm 0.07$	$17.9 \pm 0.83$	>140
	7 ( <b>37d</b> )	$4.42 \pm 0.18$	$0.710 \pm 0.025$	21.7
	13 ( <b>38d</b> ) <sup>b</sup>	NT	NA	NT
	15 ( <b>39d</b> ) <sup>b</sup>	NT	NA	NT
Round 2B				
	( <b>40</b> )	NA	NA	NT
	( <b>41</b> )	NA	NA	NT
	( <b>42</b> )	3.12	$1.58 \pm 0.58$	>70

**Table 6.1: continued**

Compound	Chain length (n=)	Cathepsin L stabilization <sup>a</sup>	Brine Shrimp Toxicity ( $\mu\text{M}$ )	CMC ( $\mu\text{M}$ )
	(43)	NA	NA	>140
	(44)	NA	NA	>140

<sup>a</sup> Fold change over enzyme blank, <sup>b</sup> Solubility issues in the assay solution

NA: Not active (Cathepsin L: 25  $\mu\text{M}$ ; brine shrimp: 30  $\mu\text{M}$ ); NT: not tested; NF: test but no reduction in surface tension at 140  $\mu\text{M}$

### 6.2.5 Significance of this Study

The surfactant industry is a multi-billion dollar global industry,<sup>25</sup> as they are incorporated into a number of daily household items, such as toothpaste, wax, washing detergents, and ink, and have important commercial utility with such uses as in herbicides, for the prevention of corrosion, detecting leaks, alkali polymers, oil dispersants, enzyme stabilization, and many others.<sup>26,27</sup> There are three different types of surfactants (ionic, non-ionic, and zwitterionic); however, each consists of a hydrophilic (head group) and hydrophobic (tail group) portion. All surfactants form micelles at a particular concentration (CMC), and thus reduce the surface tension at air-liquid, liquid-liquid, or liquid-solid interfaces. Once they have reached their CMC, there is very little to reduction in surface tension even with significant increases in the surfactant concentration. The majority of the commercially available non-ionic surfactants are long hydrocarbons with either polyethers or polyols.<sup>23</sup> Although there have been two reports of simple tertiary amides, like the ones discovered and developed in this study, that are able to reduce the surface tension of aqueous mixtures, they do not mention their potential as surfactants.<sup>28,29</sup>

One of the best known recent uses of a surfactant was during the cleanup efforts of the British Petroleum (BP) oil spill in the Gulf of Mexico between April 2<sup>nd</sup>, 2010 and July 15<sup>th</sup>, 2010.

During this time period over 210 million gallons of oil spewed into the Gulf, and in an effort to clean up the oil 1.84 million gallons of surfactants were used to disperse the oil.<sup>30</sup> The primary surfactant used was Corexit 9500A®, which is a mixture of surfactants, including Tween85, Tween80, Span80, di-(propylene glycol) butyl ether, kerosene, and 2-butoxyethanol, all of which are non-ionic polyether and polyols. Although, Corexit 9500A® did disperse the oil,<sup>31</sup> it is estimated that it actually made the oil 52 times more toxic.<sup>32</sup> Thus, there is a real need for the development of new surfactants with more environment friendly toxicity profiles.

This structure-activity relationship investigation into lnyngbyamide A (**11**) revealed that, although these analogs do exhibit some interesting biological activities, they are not just non-specifically active in a broad range of assays. They exhibit toxicity against brine shrimp and stabilization of cathepsin L; it is conceivable that both of these activities could be due to the surfactant properties of these tertiary amides. This can be clearly seen in figure 6.9 where the reduction in surface tension inversely correlates with the increase in cathepsin L stabilization; at the CMC it reaches the maximum cathepsin L stabilization effect. Several reports have shown the ability of common biologically important surfactants to stabilize cathepsin L or other enzymes by either preventing degradation or preventing the enzyme from sticking to walls of the assay plate.<sup>33-</sup>  
<sup>35</sup> In comparing two of the best synthetic analogs from this study with some of the most common biochemically important surfactants, these simple tertiary amides have the lowest CMC values and have significantly smaller molecular weights, thus requiring less material to elicit comparable responses (table 6.2).<sup>36</sup>

Furthermore, there are numerous reports of the toxic effects of surfactants on aquatic organisms, as surfactants are thought to have both acute and chronic toxicity on both fish and shrimp.<sup>37-39</sup> The acute toxicity is due to the surfactants interfering with the permeability of the gills and therefore suffocating the organism to death; however, the chronic toxicity is more likely caused

by the ability of surfactants to degrade their protective mucus layer, thus making it much more susceptible to contracting infections.<sup>40</sup>

**Table 6.2:** Comparing the CMC and MW of commercially available biologically important surfactants and two analogs

Compound	CMC ( $\mu\text{M}$ )	MW (g/mol)	Type
<b>37d</b>	21.7	255.5	Non-ionic
<b>37c</b>	13.9	281.5	Non-ionic
Tween20	60	~1228	Non-ionic
Triton X-100	450	~650	Non-ionic
SDS	8,500	288.5	Ionic
CHAPS	7500	615	Zwitterionic
BRIJ 35	90	1200	Non-ionic
NP-40	150	~650	Non-ionic

### 6.3 Conclusion

The structure-activity relationship investigation into lynchamide A (**11**) led to the production of 50 structural analogs including 10 synthetic intermediates, which were evaluated in a variety of different assays including H-460 cytotoxicity, brine shrimp toxicity, cathepsin L activation/inhibition, nitric oxide production in RAW cells, cancer cell selectivity, ion channel modulation and in several neuro-receptor binding assays (serotonin, dopamine, and cannabinoid). Overall, these alkyl amides were not that biologically active in these assays; however, a subset was active in both the brine shrimp toxicity and cathepsin L activation assays. The key structural functionalities that give optimal activity are the tertiary amide head group that is no larger than a pyrrolidine ring, and a fatty acid chain that is 14 carbons in length. The activity profiles of several analogs in the cathepsin L assay were remarkably similar to that of the biological important surfactant, tween20, thus indicating that these new synthetic analogs may be working in a similar fashion. Evaluation of the surface tension of aqueous mixtures of each of the tertiary amides

revealed that these compounds are indeed surfactants, thus explaining the mechanism behind both the brine shrimp toxicity and cathepsin L stabilization. These analogs have already shown utility in the ability to stabilize enzymes, however further investigations are needed in order to examine other potential uses as the surfactant industry is growing rapidly every year. Additionally, they should be evaluated for their potentially useful arthropod and insect killing activities, as these could be commercially and societally useful applications as well.

## 6.4 Experimental Methods

### 6.4.1 General Experimental Procedures

All reagents were commercially obtained (Aldrich, Alfa Aesar, or TCI America) at highest commercial quality and used without further purification. Air- and moisture-sensitive liquids and solutions were transferred via syringe. All non-aqueous reactions were carried out under anhydrous conditions using flame-dried glassware with an argon atmosphere. Yields refer to chromatographically and spectroscopically ( $^1\text{H}$  and  $^{13}\text{C}$  NMR) homogeneous materials, unless otherwise stated. Sigma silica gel (60, particle size 0.015-0.040 mm) was used for flash chromatography. IR spectra were recorded on a Nicolet ThermoElectron IR100 FT-IR spectrometer using KBr plates. NMR spectra were recorded with  $\text{CHCl}_3$ , DMSO, and MeOH as internal standards ( $\delta_{\text{C}}$  77.0/ $\delta_{\text{H}}$  7.26,  $\delta_{\text{C}}$  39.5/ $\delta_{\text{H}}$  2.50, and  $\delta_{\text{C}}$  49.0/ $\delta_{\text{H}}$  3.31, respectively) on a Varian 500 MHz spectrometer (500 and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively) and Varian 400 MHz spectrometer (400 and 100 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  NMR, respectively). HR ESIMS spectra were obtained on an Agilent 6230 ESI-TOF mass spectrometer. HPLC was carried out using Waters 515 pumps system with a Waters 996 PDA detector.

### 6.4.2 Amide bond formation

To a solution of the fatty acid (0.75 mmol) in *N,N*-dimethylformamide [DMF (2.5 mL)] at room temperature (rt) was added *N,N*-diisopropylcarbodiimide [DIPC (129.0  $\mu$ L, 0.82 mmol, 1.1 eq)] and hydroxybenzotriazole [HOBt (38% wet, 178.0 mg, 0.82 mmol, 1.1 eq)].<sup>18</sup> After stirring for 5 min, the amines were added (0.90 mmol, 1.2 eq) and the reaction was left stirring for 24 hours (h). The reaction was quenched with MeOH (1 mL) and H<sub>2</sub>O (10 mL). The aqueous was then back extracted 3x with DCM. The combined organic layers were washed once with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. Flash chromatography (up to 30% EtOAc in hexanes) of the crude mixtures gave amides with yields between 55%-85%.

#### 6.4.3 Olefin Metathesis

A flamed-dry pear shaped flask was charged with 10 mL of freshly distilled DCM, both ethyl pentenoate (1.11 mL, 7.14 mmol) and terminal alkene (either nonene or tridecene, 10.72 mmol, 1.5 eq) were added, followed by the Grubbs second generation catalyst (113 mg, 0.18 mmol, 2.5%).<sup>19</sup> The reaction was left stirring at rt for 5 h and then dried down under vacuum. Flash chromatography (up to 20% DCM in hexanes) of the crude mixture gave the pure unsaturated fatty ester with yields between 45%-55%.

#### 6.4.4 Base Hydrolysis

To a solution of KOH (1.06 g, 18.9 mmol) in 6:1 EtOH-H<sub>2</sub>O (5 mL) at rt was added the fatty esters (1.88 mmol).<sup>20</sup> The mixture was left vigorously stirring for 2 h, then acidified with H<sub>2</sub>SO<sub>4</sub> (2 M) and extracted 3x with diethyl ether. The combined organic layer was washed once with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. By <sup>1</sup>H NMR the resulting mixture was considered pure enough for the next reaction.

#### 6.4.4 Cyclopropanation

To a solution of the unsaturated fatty ester (1.17 mmol) in 1,2-dichloroethane (5 mL) at rt was added diiodomethane (233.6  $\mu$ L, 2.90 mmol, 2.5 eq). The mixture was then cooled to 0°C for the drop wise addition of diethylzinc (1 M in hexanes, 2.01 mL, 2.90 mmol, 2.5 eq).<sup>21</sup> The solution was stirred at rt for 60 h, however at 24 and 48 h equivalent amounts of diiodomethane and diethylzinc were added to the reaction mixture. The reaction was quenched with NH<sub>4</sub>Cl (10 mL) and subsequently back-extracted 3x with DCM. The combined organic layers were washed once with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The crude mixture was cleaned up using a 5 g normal-phase solid phase extraction, eluting with 30% DCM in hexanes with yields between 90%-95%.

#### 6.4.6 Analytical Data for Synthetic Analogs

**Compound 18a:** white solid; IR (KBr)  $\gamma_{\max}$  3298, 2921, 2851, 1640, 1549, 1060, 974  $\text{cm}^{-1}$ ; <sup>1</sup>H NMR (500 MHz, MeOD)  $\delta$ : 3.90 (quin, 1H,  $J = 5.6$ ), 3.58 (d, 4H,  $J = 5.6$ ), 2.20 (t, 2H,  $J = 7.6$ ), 1.59 (m, 2H), 1.34-1.24 (m, 16H), 0.88 (t, 3H,  $J = 7.1$ ); <sup>13</sup>C NMR (125 MHz, MeOD)  $\delta$ : 176.5, 62.0, 54.3, 37.2, 33.1, 30.7, 30.7, 30.6, 30.5, 30.5, 30.3, 27.0, 23.7, 14.4; HRESIMS  $m/z$  [M+H]<sup>+</sup> 274.2382 (calcd. for C<sub>15</sub>H<sub>32</sub>NO<sub>3</sub> 274.2382,  $\Delta$  0.13 mmu).

**Compound 18b:** white solid; IR (KBr)  $\gamma_{\max}$  3312, 2921, 2851, 1639, 1545, 1055, 1033, 1009, 669  $\text{cm}^{-1}$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$ : 7.30 (t, 2H,  $J = 7.4$ ), 7.23 (d, 1H,  $J = 7.2$ ), 7.18 (d, 2H,  $J = 7.5$ ), 5.69 (s, 1H), 3.50 (q, 2H,  $J = 6.6$ ), 2.81 (t, 2H,  $J = 7.0$ ), 2.11 (t, 2H,  $J = 7.8$ ), 1.58 (m, 2H), 1.36-1.26 (m, 16H), 0.88 (t, 3H, 6.6); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$ : 173.1, 138.9, 128.7, 128.5, 126.4, 40.4, 36.7, 35.6, 31.8, 29.5, 29.5, 29.4, 29.3, 29.3, 29.2, 25.7, 22.6, 14.0; HRESIMS  $m/z$  [M+H]<sup>+</sup> 304.2642 (calcd. for C<sub>20</sub>H<sub>34</sub>NO 304.2640,  $\Delta$  0.42 mmu).

**Compound 18c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2923, 2864, 1641, 1425, 1342, 1055, 1033, 1011  $\text{cm}^{-1}$ ; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$ : 3.46 (t, 2H,  $J = 6.8$ ), 3.41 (t, 2H,  $J = 6.8$ ), 2.25 (t, 2H,  $J = 7.9$ ), 1.95, (quin, 2H,  $J = 6.9$ ), 1.85 (quin, 2H,  $J = 6.9$ ), 1.64 (m, 2H), 1.36-1.26 (m, 16H), 0.88 (t, 3H,

6.4);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.7, 46.5, 45.4, 34.7, 31.8, 29.5, 29.5, 29.4, 29.4, 29.3, 29.2, 26.0, 24.8, 24.3, 22.5, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  254.2482 (calcd. for  $\text{C}_{16}\text{H}_{32}\text{NO}$  253.2484  $\Delta$  0.58 mmu).

**Compound 18d:** colorless oil; IR (KBr)  $\gamma_{\text{max}}$  2920, 2848, 1644, 1509, 1396, 1139, 1008  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.01 (s, 3H), 2.94 (s, 3H), 2.30 (t, 2H,  $J = 7.8$ ), 1.62 (quin, 2H,  $J = 7.5$ ), 1.36-1.26 (m, 16H), 0.88 (t, 3H, 6.7);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.1, 37.1, 35.1, 33.3, 31.8, 29.5, 29.5, 29.4, 29.4, 29.3, 29.2, 25.0, 22.5, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  228.2326 (calcd. for  $\text{C}_{14}\text{H}_{30}\text{NO}$  228.2327,  $\Delta$  0.52 mmu).

**Compound 18e:** pale yellow solid; IR (KBr)  $\gamma_{\text{max}}$  3306, 2921, 2851, 1640, 1555, 1056  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz, MeOD)  $\delta$ : 6.65 (d, 1H,  $J = 7.6$ ), 6.62 (d, 1H,  $J = 1.5$ ), 6.49 (dd, 1H,  $J = 7.6, 1.5$ ), 3.30 (m, 2H), 2.60 (t, 2H,  $J = 7.2$ ), 2.11 (t, 2H,  $J = 7.6$ ), 1.54 (m, 2H), 1.36-1.26 (m, 16H), 0.87 (t, 3H, 6.8);  $^{13}\text{C}$  NMR (125 MHz, MeOD)  $\delta$ : 176.2, 146.2, 144.7, 132.0, 121.0, 116.8, 116.3, 42.2, 37.2, 36.0, 33.1, 30.8, 30.7, 30.6, 30.5, 30.4, 30.3, 27.1, 23.7, 14.5; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  336.2531 (calcd for  $\text{C}_{20}\text{H}_{34}\text{NO}_3$  336.2539,  $\Delta$  2.4 mmu).

**Compound 19a:** white solid; IR (KBr)  $\gamma_{\text{max}}$  3298, 2919, 2850, 1640, 1546, 1057  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz, MeOD)  $\delta$ : 3.90 (quin, 1H,  $J = 5.5$ ), 3.58 (d, 4H,  $J = 5.5$ ), 2.20 (t, 2H,  $J = 7.4$ ), 1.59 (m, 2H), 1.34-1.24 (m, 16H), 0.88 (t, 3H,  $J = 6.2$ );  $^{13}\text{C}$  NMR (125 MHz, MeOD)  $\delta$ : 176.5, 62.0 x 2, 54.3, 37.2, 33.1, 30.8 (3 overlapping species), 30.8, 30.8, 30.7, 30.6, 30.5, 30.5, 30.3, 27.0, 23.7, 14.4; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  330.3006 (calcd. for  $\text{C}_{19}\text{H}_{40}\text{NO}_3$  330.3008,  $\Delta$  0.51 mmu).

**Compound 19b:** white solid; IR (KBr)  $\gamma_{\text{max}}$  3316, 2921, 2850, 1639, 1055, 1033, 1012  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.30 (t, 2H,  $J = 7.4$ ), 7.23 (d, 1H,  $J = 7.6$ ), 7.19 (d, 2H,  $J = 7.6$ ), 5.58, (s, 1H), 3.51 (q, 2H,  $J = 6.4$ ), 2.81 (t, 2H,  $J = 7.1$ ), 2.11 (t, 2H,  $J = 7.8$ ), 1.58 (m, 2H), 1.36-1.26 (m, 24H), 0.88 (t, 3H, 6.7);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.2, 138.9, 128.7, 128.5, 126.4, 40.5, 36.8, 35.7, 31.9, 29.6 (8 overlapping species), 29.6, 29.3, 29.3, 29.2, 25.7, 22.6, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  360.3268 (calcd. for  $\text{C}_{24}\text{H}_{42}\text{NO}$  360.3266,  $\Delta$  0.53 mmu).

**Compound 19c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2921, 2851, 1639, 1425, 1342, 1055, 1033, 1009, 670  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.46 (t, 2H,  $J = 6.9$ ), 3.41 (t, 2H,  $J = 6.9$ ), 2.25 (t, 2H,  $J = 7.9$ ), 1.95, (quin, 2H,  $J = 6.6$ ), 1.85 (quin, 2H,  $J = 6.6$ ), 1.64 (m, 2H), 1.36-1.26 (m, 24H), 0.88 (t, 3H, 7.0);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.8, 46.5, 45.5, 34.8, 31.9, 29.6 (5 overlapping species), 29.6, 29.5, 29.5, 29.4, 29.3, 26.1, 24.9, 24.4, 22.6, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  310.3114 (calcd. for  $\text{C}_{20}\text{H}_{40}\text{NO}$  310.3110,  $\Delta$  1.33 mmu).

**Compound 19d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2918, 2849, 1633, 1394, 1055, 1008  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.00 (s, 3H), 2.94 (s, 3H), 2.30 (t, 2H,  $J = 7.6$ ), 1.62 (m, 2H), 1.36-1.26 (m, 24H), 0.88 (t, 3H, 6.8);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.2, 37.2, 35.3, 33.4, 31.9, 29.6 (6 overlapping species), 29.5, 29.5, 29.4, 29.3, 25.1, 22.6, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  284.2959 (calcd. for  $\text{C}_{18}\text{H}_{38}\text{NO}$  284.2953,  $\Delta$  2.10 mmu).

**Compound 19e:** pale yellow solid; IR (KBr)  $\gamma_{\max}$  3307, 2920, 1639, 1554, 1055, 1033, 1011  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz, MeOD)  $\delta$ : 6.67 (d, 1H,  $J = 8.1$ ), 6.64 (d, 1H,  $J = 1.8$ ), 6.51 (dd, 1H,  $J = 8.1$ , 1.8), 3.32 (m, 2H), 2.62 (t, 2H,  $J = 7.6$ ), 2.14 (t, 2H,  $J = 7.2$ ), 1.57 (m, 2H), 1.36-1.26 (m, 24H), 0.90 (t, 3H, 6.5);  $^{13}\text{C}$  NMR (125 MHz, MeOD)  $\delta$ : 176.2, 146.2, 144.7, 132.0, 121.0, 116.8, 116.3, 42.2, 37.2, 36.0, 33.1, 30.8 (5 overlapping species), 30.7, 30.6, 30.5, 30.4, 30.3, 27.1, 23.8, 14.5; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  392.3168 (calcd for  $\text{C}_{24}\text{H}_{42}\text{NO}_3$  392.3165,  $\Delta$  0.84 mmu).

**Compound 23:** colorless oil; IR (KBr)  $\gamma_{\max}$  2926, 2856, 1739, 1462, 1373, 1345, 1248, 1163, 1121, 1039  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 5.40 (m, 1H), 5.36 (m, 1H), 4.08 (q, 2H,  $J = 6.9$ ), 2.28 (m, 4H), 1.92 (q, 2H,  $J = 7.0$ ), 1.35-1.20 (m, 13H), 0.84 (t, 3H, 7.1);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.0, 131.7, 127.8, 60.0, 34.3, 32.4, 31.8, 29.3, 29.1, 29.0, 27.9, 22.6, 14.1, 14.0; HRESIMS  $m/z$   $[\text{M}-\text{H}]^-$  227.2011 (calcd. for  $\text{C}_{14}\text{H}_{27}\text{O}_2$  227.1006,  $\Delta$  2.2 mmu).

**Compound 24:** colorless oil; IR (KBr)  $\gamma_{\max}$  2925, 2854, 1740, 1638, 1423, 1372, 1248, 1170, 1040, 968  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 5.44 (m, 1H), 5.40 (m, 1H), 4.12 (q, 2H,  $J = 7.4$ ), 2.32 (m, 4H), 1.96 (q, 2H,  $J = 7.0$ ), 1.35-1.20 (m, 21H), 0.88 (t, 3H, 7.0);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ :

173.1, 131.7, 127.9, 60.1, 34.4, 32.5, 31.9, 29.6, 29.6, 29.6, 29.5, 29.4, 29.3, 29.1, 27.9, 22.6, 14.2, 14.0; HRESIMS  $m/z$  [M-H]<sup>-</sup> 283.2635 (calcd. for C<sub>18</sub>H<sub>35</sub>O<sub>2</sub> 283.2631,  $\Delta$  1.1 mmu).

**Compound 25:** colorless oil; IR (KBr)  $\gamma_{\max}$  2926, 2856, 1742, 1712, 1439, 1286, 1210, 1166 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$ : 5.46 (m, 1H), 5.41 (m, 1H), 2.40 (t, 2H, J = 7.5), 2.31 (m, 2H), 1.96 (q, 2H, J = 7.0), 1.35-1.20 (m, 10H), 0.88 (t, 3H, J = 6.0); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  179.6, 132.2, 127.4, 34.2, 32.5, 31.8, 29.4, 29.2, 29.1, 27.6, 22.7, 14.1; HRESIMS  $m/z$  [M-H]<sup>-</sup> 197.1548 (calcd. for C<sub>12</sub>H<sub>23</sub>O<sub>2</sub> 197.1547,  $\Delta$  0.5 mmu).

**Compound 26:** colorless oil; IR (KBr)  $\gamma_{\max}$  2917, 2850, 1707, 1468, 1265, 1214, 964 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, MeOD-CDCl<sub>3</sub>)  $\delta$ : 5.44 (m, 2H), 2.26 (m, 2H), 1.96 (q, 2H, J = 7.0), 1.35-1.20 (m, 18H), 0.89 (t, 3H, J = 7.0); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$ : 178.0, 132.0, 127.6, 34.1, 32.5, 31.9, 29.6, 29.6, 29.5, 29.4, 29.3, 29.1, 27.6, 22.7, 18.1, 14.1; HRESIMS  $m/z$  [M-H]<sup>-</sup> 253.2176 (calcd. for C<sub>16</sub>H<sub>31</sub>O<sub>2</sub> 253.2173,  $\Delta$  1.2 mmu).

**Compound 27a:** white solid; IR (KBr)  $\gamma_{\max}$  3263, 2922, 2851, 1629, 1447, 1072, 972 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$ : 5.47 (m, 1H), 5.41 (m, 1H), 3.91 (quin, 1H, J = 5.3), 3.62 (m, 4H), 3.32 (bs, 2H), 2.28 (m, 4H), 1.98 (q, 2H, J = 6.7), 1.37-1.23 (m, 10H), 0.88 (t, 3H, 6.4); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$ : 173.7, 130.8, 127.3, 60.0, 52.0, 35.3, 31.6, 31.0, 28.6, 28.3, 28.2, 27.9, 21.7, 12.6; HRESIMS  $m/z$  [M+H]<sup>+</sup> 272.2219 (calcd. for C<sub>15</sub>H<sub>30</sub>NO<sub>3</sub> 272.2226,  $\Delta$  2.44 mmu).

**Compound 27b:** white solid; IR (KBr)  $\gamma_{\max}$  3301, 2923, 2853, 1639, 1549, 1454, 966 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$ : 7.31 (t, 2H, J = 7.7), 7.24 (d, 1H, J = 7.2), 7.19 (d, 2H, J = 7.5), 5.58 (s, 1H), 5.41 (m, 1H), 5.37 (m, 1H), 3.51 (q, 2H, J = 6.6), 2.81 (t, 2H, J = 7.0), 2.28 (q, 2H, J = 6.4), 2.18 (t, 2H, J = 7.1), 1.94 (q, 2H, J = 6.7), 1.34-1.22 (m, 10H), 0.88 (t, 3H, 6.8); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$ : 172.5, 138.9, 131.9, 128.7, 128.6, 128.1, 126.4, 40.5, 36.6, 35.7, 32.5, 31.8, 29.4, 29.1, 29.1, 28.6, 22.6, 14.1; HRESIMS  $m/z$  [M+H]<sup>+</sup> 302.2487 (calcd. for C<sub>20</sub>H<sub>32</sub>NO 302.2484,  $\Delta$  0.92 mmu).

**Compound 27c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2955, 2924, 2855, 1640, 1436  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 5.45 (m, 2H), 3.46 (t, 2H,  $J = 6.7$ ), 3.40 (t, 2H,  $J = 6.7$ ), 2.34 (m, 4H), 1.95 (m, 2H), 1.85 (m, 2H), 1.37-1.20 (m, 10H), 0.88 (t, 3H,  $J = 6.1$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.2, 131.4, 128.8, 46.6, 45.6, 34.9, 32.5, 31.6, 29.5, 29.2, 29.2, 28.0, 26.1, 24.4, 22.7, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  252.2333 (calcd. for  $\text{C}_{16}\text{H}_{30}\text{NO}$  252.2327,  $\Delta$  2.32 mmu).

**Compound 27d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2926, 2855, 1649, 1464, 1397, 1266, 1144, 970  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $d_6$ -DMSO)  $\delta$ : 5.39 (m, 2H), 2.92 (s, 3H), 2.78 (s, 3H), 2.48 (m, 2H), 2.28 (m, 2H), 2.14 (m, 2H), 1.91 (m, 2H), 1.32-1.17 (m, 10H), 0.83 (t, 3H, 6.7);  $^{13}\text{C}$  NMR (125 MHz,  $d_6$ -DMSO)  $\delta$ : 171.3, 130.4, 129.3, 36.7, 34.8, 32.5, 31.9, 31.3, 29.0, 28.6, 28.5, 27.7, 22.1, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  226.2173 (calcd. for  $\text{C}_{14}\text{H}_{28}\text{NO}$  226.2171,  $\Delta$  0.85 mmu).

**Compound 28a:** white solid; IR (KBr)  $\gamma_{\max}$  3291, 2921, 2850, 1640, 1465, 1074, 970  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 5.47 (m, 1H), 5.41 (m, 1H), 3.91 (m, 1H), 3.63 (m, 4H), 3.33 (bs, 2H), 2.29 (m, 4H), 1.98 (q, 2H,  $J = 7.1$ ), 1.37-1.23 (m, 18H), 0.88 (t, 3H, 7.5);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.8, 131.0, 127.3, 60.2, 52.1, 35.5, 31.7, 31.2, 28.9, 28.9, 28.9, 28.7, 28.7, 28.6, 28.4, 28.0, 21.9, 12.9; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  328.2856 (calcd. for  $\text{C}_{19}\text{H}_{38}\text{NO}_3$  328.2852,  $\Delta$  1.17 mmu).

**Compound 28b:** white solid; IR (KBr)  $\gamma_{\max}$  3301, 2920, 2851, 1638, 1549, 1454  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.30 (t, 2H,  $J = 7.4$ ), 7.23 (d, 1H,  $J = 7.4$ ), 7.18 (d, 2H,  $J = 7.4$ ), 5.65 (s, 1H), 5.41 (m, 1H), 5.36 (m, 1H), 3.50 (q, 2H,  $J = 6.6$ ), 2.80 (t, 2H,  $J = 6.90$ ), 2.28 (q, 2H,  $J = 6.7$ ), 2.18 (t, 2H,  $J = 7.2$ ), 1.94 (q, 2H,  $J = 7.1$ ), 1.34-1.22 (m, 18H), 0.88 (t, 3H, 7.0);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 172.5, 138.9, 131.9, 128.7, 128.5, 128.1, 126.4, 40.4, 36.6, 35.7, 32.5, 31.9, 29.6, 29.6, 29.6, 29.5, 29.4, 29.3, 29.1, 28.5, 22.6, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  358.3118 (calcd. for  $\text{C}_{24}\text{H}_{40}\text{NO}$  358.3110,  $\Delta$  2.34 mmu).

**Compound 28c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2956, 2923, 2853, 1641, 1434  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 5.45 (m, 2H), 3.46 (t, 2H,  $J = 6.6$ ), 3.41 (t, 2H,  $J = 6.6$ ), 2.32 (m, 4H), 1.94 (m, 2H), 1.86 (m, 2H), 1.38-1.20 (m, 18H), 0.88 (t, 3H, 6.5);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.2,

131.4, 128.7, 46.6, 45.6, 34.9, 32.5, 31.9, 29.7, 29.6, 29.6, 29.5, 29.5, 29.3, 29.1, 28.0, 26.1, 24.4, 22.7, 14.1; HRESIMS  $m/z$   $[M+H]^+$  308.2958 (calcd. for  $C_{20}H_{38}NO$  308.2953,  $\Delta$  1.56 mmu).

**Compound 28d:** colorless oil; IR (KBr)  $\gamma_{max}$  2924, 2853, 1650, 1463, 1397, 1142, 969  $cm^{-1}$ ;  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$ : 5.45 (m, 2H), 3.00 (s, 3H), 2.94 (s, 3H), 2.34 (m, 4H), 1.97 (m, 2H), 1.37-1.22 (m, 18H), 0.88 (t, 3H, 6.7);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$ : 172.6, 131.4, 128.6, 37.2, 35.3, 33.4, 32.4, 31.8, 29.6, 29.5, 29.5, 29.4, 29.4, 29.2, 29.1, 28.1, 22.6, 14.0; HRESIMS  $m/z$   $[M+Na]^+$  304.2628 (calcd. for  $C_{18}H_{35}NONa$  304.2616,  $\Delta$  0.44 mmu).

**Compound 29:** colorless oil; IR (KBr)  $\gamma_{max}$  3062, 2925, 2855, 1739, 1461, 1373, 1250, 1172  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 4.1 (q, 2H,  $J = 7.2$ ), 2.34 (t, 2H,  $J = 7.5$ ), 1.53 (m, 2H), 1.47 (m, 2H), 1.38-1.03 (m, 13H), 0.86 (t, 3H,  $J = 7.1$ ), 0.40 (m, 1H), 0.17 (m, 1H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$  173.7, 60.1, 34.5, 34.1, 31.9, 29.7, 29.6, 29.4, 29.3, 22.6, 18.8, 18.1, 14.2, 14.1, 11.7; HRESIMS  $m/z$   $[M-H]^-$  241.2166 (calcd. for  $C_{15}H_{37}O_2$  241.2162,  $\Delta$  1.7 mmu).

**Compound 30:** colorless oil; IR (KBr)  $\gamma_{max}$  2925, 2854, 1739, 1462, 1373, 1249, 1177  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 4.1 (q, 2H,  $J = 7.1$ ), 2.36 (t, 2H,  $J = 7.5$ ), 1.55 (m, 2H), 1.48 (m, 2H), 1.38-1.03 (m, 21H), 0.87 (t, 3H,  $J = 7.0$ ), 0.42 (m, 1H), 0.19 (m, 1H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$  173.7, 60.1, 34.5, 34.1, 31.9, 29.7-29.6 (6 overlapping species), 29.5, 29.4, 22.7, 18.8, 18.1, 14.2, 14.1, 11.8; HRESIMS  $m/z$   $[M-H]^-$  297.2789 (calcd. for  $C_{15}H_{37}O_2$  297.2788,  $\Delta$  0.3 mmu).

**Compound 31:** white solid; IR (KBr)  $\gamma_{max}$  2925, 2855, 1742, 1712, 1439, 1286, 1210, 1166  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 9.80 (bs, 1H), 2.43 (t, 2H,  $J = 7.6$ ), 1.54 (m, 2H), 1.40-1.20 (m, 19H), 1.14 (m, 1H), 0.88 (t, 3H,  $J = 6.7$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 180.2, 34.3, 34.1, 31.9, 29.6, 29.5, 29.4, 29.3, 22.7, 18.9, 18.0, 14.0, 11.7; HRESIMS  $m/z$   $[M-H]^-$  211.1706 (calcd. for  $C_{12}H_{25}O_2$  211.1704,  $\Delta$  0.9 mmu).

**Compound 32:** white solid; IR (KBr)  $\gamma_{max}$  2925, 2854, 1709, 1639, 1459, 1285  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 2.40 (t, 2H,  $J = 7.3$ ), 1.51 (m, 2H), 1.40-1.20 (m, 19H), 1.11 (m, 1H), 0.86 (t, 3H,  $J = 6.9$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 180.5, 34.3, 34.1, 32.0, 29.7 (3 overlapping species), 29.7,

29.6, 29.5, 29.4, 29.4, 22.7, 18.9, 18.1, 14.1, 11.8; HRESIMS  $m/z$   $[M-H]^-$  267.2334 (calcd. for  $C_{17}H_{33}O_2$  267.2330,  $\Delta$  1.5 mmu).

**Compound 33a:** white solid; IR (KBr)  $\gamma_{max}$  3299, 2920, 2852, 1641, 1546, 1462, 1056, 974, 692  $cm^{-1}$ ;  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$ : 3.91 (quin, 1H,  $J = 5.5$ ), 3.65 (m, 4H), 3.34 (m, 2H), 2.30 (t, 2H,  $J = 7.7$ ), 1.53 (m, 2H), 1.40-1.23 (m, 10H), 1.20 (m, 2H), 0.89 (t, 3H,  $J = 7.0$ ), 0.45 (m, 1H), 0.22 (m, 1H);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$ : 174.4, 60.3, 52.1, 35.7, 33.5, 31.2, 29.9, 28.9, 28.8, 28.6, 21.9, 18.1, 17.5, 13.0, 10.9; HRESIMS  $m/z$   $[M+H]^+$  286.2384 (calcd. for  $C_{21}H_{34}NO$  286.2382,  $\Delta$  0.69 mmu).

**Compound 33b:** white solid; IR (KBr)  $\gamma_{max}$  3309, 2921, 2853, 1639, 1546, 1459, 698  $cm^{-1}$ ;  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$ : 7.31 (t, 2H,  $J = 7.2$ ), 7.24 (d, 1H,  $J = 7.4$ ), 7.19 (d, 2H,  $J = 7.4$ ), 3.52 (q, 2H,  $J = 6.6$ ), 2.82 (t, 2H,  $J = 6.6$ ), 2.19 (t, 2H,  $J = 7.5$ ), 1.50 (dq, 2H,  $J = 7.5, 3.1$ ), 1.38-1.21 (m, 10H), 1.15 (m, 2H), 0.88 (t, 3H,  $J = 7.1$ ), 0.39 (m, 1H), 0.17 (m, 1H);  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$ : 173.0, 138.9, 128.7, 128.6, 126.5, 40.5, 36.9, 35.7, 34.1, 31.9, 30.3, 29.6, 29.5, 29.3, 22.7, 18.9, 18.2, 14.1, 11.7; HRESIMS  $m/z$   $[M+H]^+$  316.2647 (calcd. for  $C_{21}H_{34}NO$  316.2640,  $\Delta$  2.22 mmu).

**Compound 33c:** colorless oil; IR (KBr)  $\gamma_{max}$  2956, 2923, 2853, 1641, 1434  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 3.45 (t, 2H,  $J = 7.2$ ), 3.42 (t, 2H,  $J = 6.6$ ), 2.33 (dd, 4H,  $J = 8.1, 6.9$ ), 1.94 (m, 2H), 1.84 (m, 2H), 1.54 (q, 2H,  $J = 6.9$ ), 1.40-1.15 (m, 11H), 1.12 (m, 1H), 0.88 (t, 3H,  $J = 6.9$ ), 0.44 (m, 1H), 0.18 (m, 1H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 171.7, 46.6, 45.6, 34.8, 34.1, 31.9, 29.7, 29.6, 29.4, 29.4, 22.7, 18.9, 18.4, 14.1, 11.8; HRESIMS  $m/z$   $[M+H]^+$  266.2488 (calcd. for  $C_{17}H_{32}NO$  266.2484,  $\Delta$  1.51 mmu).

**Compound 33d:** colorless oil; IR (KBr)  $\gamma_{max}$  2922, 2852, 1646, 1458, 1396, 1265, 1148  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 3.00 (bs, 3H), 2.94 (bs, 3H), 2.38 (t, 2H,  $J = 7.6$ ), 1.52 (m, 2H), 1.39-1.17 (m, 11H), 1.12 (m, 1H), 0.88 (t, 3H,  $J = 6.6$ ), 0.43 (m, 1H), 0.18 (m, 1H);  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 173.1, 37.3, 35.4, 34.1, 33.4, 31.9, 29.9, 29.6, 29.4, 29.4, 22.7, 18.9, 18.4, 14.1, 11.8; HRESIMS  $m/z$   $[M+H]^+$  240.2325 (calcd. for  $C_{15}H_{30}NO$  240.2327,  $\Delta$  0.94 mmu).

**Compound 34a:** white solid; IR (KBr)  $\gamma_{\max}$  3297, 2919, 2850, 1640, 1546, 1074, 974  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.91 (quin, 1H,  $J = 5.3$ ), 3.63 (m, 4H), 3.33 (m, 2H), 2.30 (t, 2H,  $J = 7.6$ ), 1.53 (m, 2H), 1.43-1.24 (m, 18H), 1.20 (m, 2H), 0.88 (t, 3H,  $J = 7.1$ ), 0.44 (m, 1H), 0.21 (m, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 174.3, 60.1, 52.1, 35.6, 33.4, 31.1, 29.8, 28.9-28.9 (5 overlapping species), 28.7, 28.5, 21.8, 17.9, 17.4, 12.8, 10.8; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  342.3012 (calcd. for  $\text{C}_{20}\text{H}_{40}\text{NO}_3$  342.3008,  $\Delta$  1.09 mmu).

**Compound 34b:** white solid; IR (KBr)  $\gamma_{\max}$  3309, 2919, 2850, 1637, 1546, 1458, 697  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.31 (t, 2H,  $J = 7.3$ ), 7.24 (d, 1H,  $J = 7.3$ ), 7.19 (d, 2H,  $J = 7.4$ ), 3.52 (q, 2H,  $J = 6.8$ ), 2.82 (t, 2H,  $J = 7.0$ ), 2.19 (t, 2H,  $J = 7.5$ ), 1.50 (dq, 2H,  $J = 6.7, 2.7$ ), 1.38-1.21 (m, 18H), 1.14 (m, 2H), 0.88 (t, 3H,  $J = 7.2$ ), 0.39 (m, 1H), 0.17 (m, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.0, 138.9, 128.7, 128.6, 126.5, 40.5, 36.9, 35.7, 34.1, 31.9, 30.3, 29.7-29.6 (5 overlapping species), 29.5, 29.3, 22.7, 18.9, 18.2, 14.1, 11.7; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  372.3276 (calcd. for  $\text{C}_{25}\text{H}_{42}\text{NO}$  372.3266,  $\Delta$  2.71 mmu).

**Compound 34c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2923, 2853, 1641, 1434  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.45 (t, 2H,  $J = 6.9$ ), 3.43 (t, 2H,  $J = 6.9$ ), 2.33 (dd, 4H,  $J = 8.9, 7.4$ ), 1.94 (m, 2H), 1.84 (m, 2H), 1.54 (q, 2H,  $J = 7.9$ ), 1.40-1.16 (m, 19H), 1.12 (m, 1H), 0.88 (t, 3H,  $J = 6.8$ ), 0.44 (m, 1H), 0.19 (m, 1H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.8, 46.7, 45.5, 34.8, 34.1, 31.9, 29.7, 29.7, 29.7, 29.7, 29.6, 29.5, 29.3, 22.7, 18.9, 18.4, 14.1, 11.8; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  322.3116 (calcd. for  $\text{C}_{21}\text{H}_{40}\text{NO}$  322.3110,  $\Delta$  1.82 mmu).

**Compound 34d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2923, 2853, 1651, 1461, 1396, 1267, 1146  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.01 (s, 3H), 2.93 (s, 3H), 2.38 (t, 2H,  $J = 7.7$ ), 1.52 (m, 2H), 1.39-1.17 (m, 19H), 1.13 (m, 1H), 0.86 (t, 3H,  $J = 7.2$ ), 0.43 (m, 1H), 0.18 (m, 1H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.1, 37.3, 35.3, 34.1, 33.3, 31.9, 29.9, 29.7, 29.7, 29.7, 29.6, 29.6, 29.5, 29.3, 22.7, 18.9, 18.4, 14.1, 11.8; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  296.2958 (calcd. for  $\text{C}_{19}\text{H}_{38}\text{NO}$  296.2953,  $\Delta$  1.43 mmu).

**Compound 35c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2927, 2852, 1642, 1434  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.42 (t, 2H,  $J = 6.7$ ), 3.38 (t, 2H,  $J = 6.7$ ), 2.22 (dd, 2H,  $J = 8.6, 7.6$ ), 1.91 (m, 2H), 1.81 (m, 2H), 1.60 (m, 2H), 1.33-1.19 (m, 20H), 0.84 (t, 3H,  $J = 7.1$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.7, 46.5, 45.3, 34.7, 31.6, 29.3, 28.9, 26.0, 24.8, 24.2, 22.4, 13.9; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  198.1859 (calcd. for  $\text{C}_{12}\text{H}_{24}\text{NO}$  198.1856,  $\Delta$  0.31 mmu).

**Compound 35d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2927, 2853, 1645  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.01 (s, 3H), 2.94 (s, 3H), 2.31 (dd, 2H,  $J = 9.1, 7.8$ ), 1.63 (m, 2H), 1.35-1.21 (m, 8H), 0.88 (t, 3H,  $J = 7.1$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.2, 37.2, 35.2, 33.3, 31.6, 29.3, 29.0, 25.1, 22.5, 13.9; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  172.1703 (calcd. for  $\text{C}_{10}\text{H}_{22}\text{NO}$  172.1701,  $\Delta$  0.76 mmu).

**Compound 36c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2926, 2856, 1645, 1434  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.46 (t, 2H,  $J = 6.7$ ), 3.42 (t, 2H,  $J = 6.7$ ), 2.25 (dd, 2H,  $J = 8.6, 7.5$ ), 1.95 (m, 2H), 1.85 (m, 2H), 1.64 (m, 2H), 1.39-1.21 (m, 20H), 0.88 (t, 3H,  $J = 6.1$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.9, 46.6, 45.5, 34.8, 31.8, 29.4, 29.4, 29.4, 29.2, 26.0, 24.9, 24.3, 22.6, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  226.2173 (calcd. for  $\text{C}_{14}\text{H}_{28}\text{NO}$  226.2171,  $\Delta$  0.71 mmu).

**Compound 36d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2926, 2855, 1651, 1465, 1400, 1150  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.01 (s, 3H), 2.94 (s, 3H), 2.31 (dd, 2H,  $J = 8.9, 7.4$ ), 1.62 (m, 2H), 1.35-1.21 (m, 12H), 0.88 (t, 3H,  $J = 6.7$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.3, 37.2, 35.3, 33.4, 31.8, 29.4, 29.4, 29.4, 29.2, 25.1, 22.6, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  200.2016 (calcd. for  $\text{C}_{12}\text{H}_{26}\text{NO}$  200.2014,  $\Delta$  0.75 mmu).

**Compound 37c:** colorless oil; IR (KBr)  $\gamma_{\max}$  2924, 2854, 1645, 1431  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.46 (t, 2H,  $J = 6.7$ ), 3.41 (t, 2H,  $J = 6.7$ ), 2.25 (t, 2H,  $J = 7.7$ ), 1.95 (m, 2H), 1.85 (m, 2H), 1.64 (m, 2H), 1.35-1.21 (m, 20H), 0.88 (t, 3H,  $J = 6.0$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.9, 46.6, 45.5, 34.8, 31.8, 29.6-29.5 (4 overlapping species), 29.4, 29.4, 29.4, 29.3, 26.0, 24.9, 24.3, 22.6, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  282.2800 (calcd. for  $\text{C}_{18}\text{H}_{36}\text{NO}$  282.2797,  $\Delta$  0.93 mmu).

**Compound 37d:** colorless oil; IR (KBr)  $\gamma_{\max}$  2923, 2853, 1649  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.01 (s, 3H), 2.94 (s, 3H), 2.30 (dd, 2H,  $J = 8.9, 7.4$ ), 1.62 (m, 2H), 1.35-1.21 (m, 20H), 0.88 (t, 3H,  $J = 6.2$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.3, 37.3, 35.3, 33.4, 31.9, 29.6-29.5 (4 overlapping species), 29.5, 29.5, 29.4, 29.3, 25.2, 22.6, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  256.2645 (calcd. for  $\text{C}_{16}\text{H}_{34}\text{NO}$  256.2640,  $\Delta$  1.92 mmu).

**Compound 38c:** white solid; IR (KBr)  $\gamma_{\max}$  2918, 2850, 1637, 1467  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.43 (t, 2H,  $J = 6.7$ ), 3.38 (t, 2H,  $J = 6.7$ ), 2.22 (dd, 2H,  $J = 8.9, 7.4$ ), 1.91 (m, 2H), 1.82 (m, 2H), 1.61 (m, 2H), 1.39-1.21 (m, 32H), 0.85 (t, 3H,  $J = 7.2$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.8, 46.5, 45.5, 34.8, 31.9, 29.7-29.6 (7 overlapping species), 29.6, 29.6, 29.6, 29.5, 29.5, 29.4, 29.3, 24.9, 22.6, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  338.3430 (calcd. for  $\text{C}_{22}\text{H}_{44}\text{NO}$  338.3422,  $\Delta$  1.95 mmu).

**Compound 38d:** white solid; IR (KBr)  $\gamma_{\max}$  2918, 2849, 1643  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 2.97 (s, 3H), 2.91 (s, 3H), 2.27 (dd, 2H,  $J = 8.9, 7.8$ ), 1.60 (m, 2H), 1.35-1.21 (m, 8H), 0.85 (t, 3H,  $J = 7.3$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.2, 37.2, 35.3, 33.4, 31.2, 29.7-29.6 (7 overlapping species), 29.6, 29.5, 29.5, 29.4, 29.3, 25.1, 22.6, 14.0; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  312.3272 (calcd. for  $\text{C}_{20}\text{H}_{42}\text{NO}$  312.3272,  $\Delta$  0.00 mmu).

**Compound 39c:** white solid; IR (KBr)  $\gamma_{\max}$  2918, 2850, 1643, 1467  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.46 (t, 2H,  $J = 7.0$ ), 3.41 (t, 2H,  $J = 7.1$ ), 2.25 (dd, 2H,  $J = 8.8, 8.0$ ), 1.94 (m, 2H), 1.85 (m, 2H), 1.64 (m, 2H), 1.39-1.21 (m, 32H), 0.88 (t, 3H,  $J = 7.5$ );  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.9, 46.6, 45.5, 34.8, 31.9, 29.7-29.6 (10 overlapping species), 29.5, 29.5, 29.4, 29.3, 26.1, 24.9, 24.4, 22.7, 14.1; HRESIMS  $m/z$   $[\text{M}+\text{H}]^+$  366.3740 (calcd. for  $\text{C}_{24}\text{H}_{48}\text{NO}$  366.3736,  $\Delta$  1.12 mmu).

**Compound 39d:** white solid; IR (KBr)  $\gamma_{\max}$  2919, 2850, 1643  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$ : 3.00 (s, 3H), 2.94 (s, 3H), 2.30 (dd, 2H,  $J = 8.7, 7.5$ ), 1.62 (m, 2H), 1.35-1.21 (m, 32H), 0.85 (t, 3H,  $J = 6.4$ );  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$ : 173.3, 37.3, 35.3, 33.4, 31.9, 29.7-29.6 (10

overlapping species), 29.5, 29.5, 29.5, 29.4, 25.2, 22.7, 14.1; HRESIMS  $m/z$   $[M+H]^+$  340.3584 (calcd. for  $C_{22}H_{46}NO$  340.3579,  $\Delta$  1.27 mmu).

**Compound 41:** white solid; IR (KBr)  $\gamma_{\max}$  3300, 2918, 2849, 1635, 1562, 1463, 1266, 1161, 1019  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 2.74 (s, 3H), 2.43 (bs, 1H), 2.12 (t, 2H,  $J = 7.4$ ), 1.57 (m, 2H), 1.39-1.21 (m, 32H), 0.86 (t, 3H,  $J = 6.5$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 174.3, 36.5, 31.8, 29.6-29.5 (6 overlapping species), 29.5, 29.4, 29.3, 29.3, 29.2, 25.7, 22.6, 14.0; HRESIMS  $m/z$   $[M+H]^+$  270.2780 (calcd. for  $C_{17}H_{36}NO$  270.2797,  $\Delta$  6.2 mmu).

**Compound 42:** colorless oil; IR (KBr)  $\gamma_{\max}$  2918, 2850, 1645, 1466, 1303, 1240  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $DMSO-CDCl_3$ )  $\delta$ : 4.01 (t, 2H,  $J = 7.3$ ), 3.80 (t, 2H,  $J = 7.3$ ), 2.13 (quin, 2H,  $J = 7.9$ ), 1.90 (t, 2H,  $J = 7.9$ ), 1.41 (m, 2H), 1.20-1.05 (m, 24H), 0.76 (t, 3H,  $J = 7.0$ );  $^{13}C$  NMR (125 MHz,  $DMSO-CDCl_3$ )  $\delta$ : 171.1, 48.3, 45.9, 30.1, 29.2, 27.8-27.7 (x overlapping species), 27.7, 27.6, 27.5, 27.5, 23.0, 20.8, 13.2, 12.5; HRESIMS  $m/z$   $[M+H]^+$  296.2964 (calcd. for  $C_{19}H_{38}NO$  296.2953,  $\Delta$  3.4 mmu).

**Compound 43:** white solid; IR (KBr)  $\gamma_{\max}$  2925, 2854, 1644, 1460, 1374, 1266, 1196, 1167  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 3.47 (t, 2H,  $J = 6.2$ ), 3.38 (t, 2H,  $J = 6.2$ ), 2.26 (dd, 2H,  $J = 8.5$ , 8.0), 1.66 (m, 4H), 1.59 (m, 2H), 1.52 (m, 4H), 1.35-1.21 (m, 24H), 0.83 (t, 3H,  $J = 7.0$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 172.8, 47.8, 45.8, 33.2, 31.8, 29.6-29.5 (8 overlapping species), 29.5, 29.4, 29.4, 29.3, 29.1, 27.5, 27.0, 26.7, 25.3, 22.6, 14.0; HRESIMS  $m/z$   $[M+H]^+$  338.3437 (calcd. for  $C_{22}H_{44}NO$  338.3423,  $\Delta$  4.2 mmu).

**Compound 44:** white solid; IR (KBr)  $\gamma_{\max}$  2925, 2854, 1643, 1463, 1422, 1360, 1203  $cm^{-1}$ ;  $^1H$  NMR (500 MHz,  $CDCl_3$ )  $\delta$ : 3.43 (t, 2H,  $J = 6.4$ ), 3.37 (t, 2H,  $J = 5.9$ ), 2.28 (dd, 2H,  $J = 8.5$ , 7.6), 1.69 (m, 4H), 1.62 (m, 2H), 1.57 (m, 2H), 1.49 (m, 4H), 1.35-1.21 (m, 24H), 0.85 (t, 3H,  $J = 7.3$ );  $^{13}C$  NMR (125 MHz,  $CDCl_3$ )  $\delta$ : 172.9, 49.1, 47.2, 33.5, 31.9, 29.7-29.6 (9 overlapping species), 29.6, 29.4, 29.3, 27.5, 27.0, 26.1, 25.6, 25.4, 25.1, 22.6, 14.0; HRESIMS  $m/z$   $[M+H]^+$  352.3591 (calcd. for  $C_{23}H_{46}NO$  352.3579,  $\Delta$  3.30 mmu).

#### 6.4.7 Brine Shrimp Toxicity Assay

Brine shrimp (*Artemia salina*) were added to 120-well 5 mL plate with about 15-40 shrimp in 2.5 mL of artificial seawater. Compounds were dissolved in DMSO to a stock concentration of 10  $\mu$ M, with the working serial dilutions also in DMSO, with a volume of 7.5  $\mu$ L added to each well. Plates were kept at room temperature for 24 h before the dead shrimp were counted by a dissecting microscope. Then all shrimp were killed with acetone, and re-counted to obtain an accurate total shrimp count.

#### 6.4.8 Cathepsin L Assay

Z-Phe-Arg-AMC substrate and E-64-c were purchased from Bachem Americas. Human recombinant cathepsin L was purchased from R&D Systems. Assays were carried out using 30  $\mu$ M Z-Phe-Arg-AMC and 3.0 ng/mL human recombinant cathepsin L. Assay buffer consisted of 50 mM sodium acetate, 100 mM NaCl, 1.0 mM EDTA and 4 mM dithiothreitol, pH 5.5. The enzymatic reaction (25°C) was monitored on a SpectraMax Gemini or SpectraMax microplate reader (PerkinElmer Life Sciences) and the fluorescent signal was measured at the excitation and emission wavelengths of 365 and 450 nm, respectively.

Kinetic characterization of the interaction between the stabilizers and cathepsin L was performed by continuous monitoring of substrate hydrolysis in the presence of the stabilizer. Each analog was tested at 25  $\mu$ M with simultaneous mixing of enzyme, substrate and stabilizer. The total product formation at the 2 h time point in the presence of the stabilizer was compared to that of just the enzyme and substrate in order to determine fold increases in enzyme activity. This analysis was performed in technical triplicate with ANOVA significance analysis done on GraphPad (vs. 5.0, Prism).

#### 6.4.9 Critical Micelle Concentration Analysis

Stock solution of each compound was prepared in DMSO at several different concentrations (140, 100, 70, 50, 30, 20, 15, 12.5, 10, 7.5, 5, and 2.5 mM), with 20  $\mu$ L of each stock solution added to 20 mL of MilliQ water. This solution was transferred to disposable aluminum sample vessels (Kruss, part # SV13), where the surface tension was calculated 3 times on a Kruss K11 tensiometer equipped with a platinum plate. In between dilution series the samples vessels and the platinum plate were rinsed with MilliQ water. After the analysis of all of the dilutions of one analog the platinum plate was rinsed with acetone, ethanol, and MilliQ water. Determination of the critical micelle concentration (CMC) for each compound was performed using the best fit line function in Excel (vs. 2013).

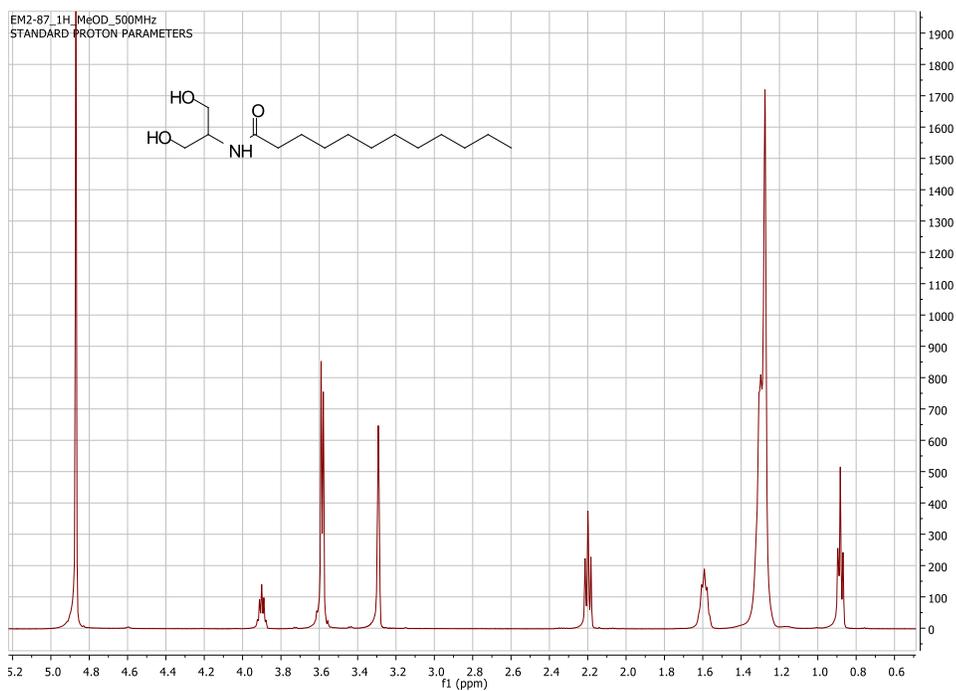
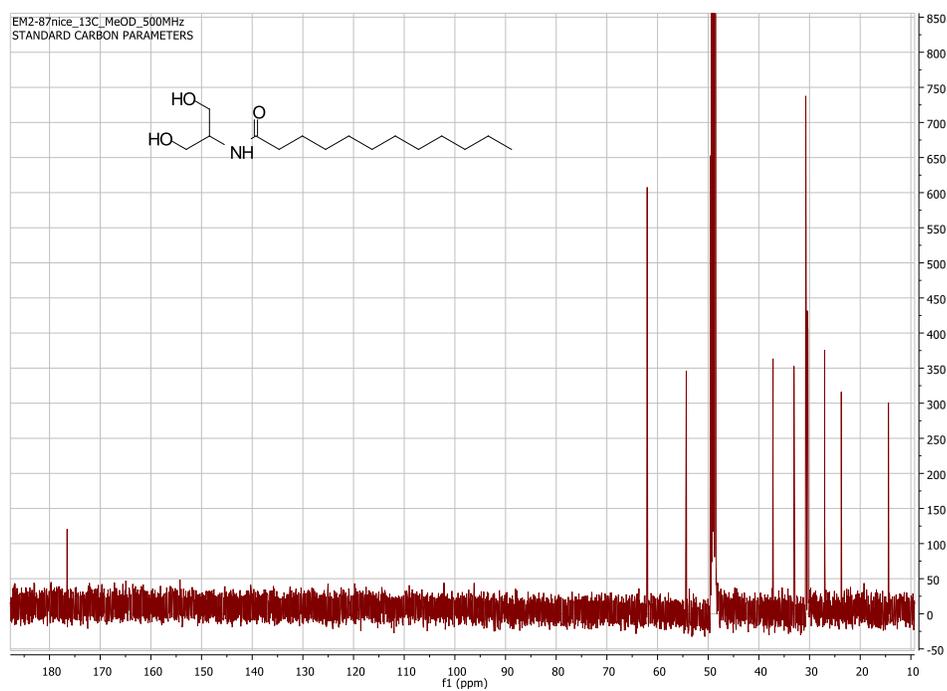
#### 6.4.10 Statistical Analysis

To examine possible correlation between CMC, surface tension, brine shrimp toxicity, and structural class on cathepsin activation, I employed multiple linear regression using ordinal ranked data in xLSTAT. Cathepsin activity, brine shrimp toxicity, critical micelle concentration, and surface tension values were ranked from highest to lowest for a total of 17 rankings. Compounds were categorized into two groups: those with carbon chain lengths of 8-12 carbons and those with carbon chain lengths of 14-16 carbons. Cathepsin activation was selected as the dependent variable and brine shrimp toxicity, CMC, and surface tension were selected as quantitative independent variables. Carbon chain lengths were set as independent qualitative variables (C-8-12 and C-14-16). The confidence interval was set at 95% and best model was selected with the criterion of adjusted  $R^2$ .

## 6.5 Acknowledgements

I thank the psychoactive drug screening program at UNC Chapel Hill, Fred Valeriote at Henry Ford Cancer Center, Thomas Murray at Creighton State University, and Samantha Mascuch and Tara Byrum both at Scripps Institution of Oceanography for preliminary screening of the first round of compounds. I further thank Dale Stokes (SIO) for use of the tensiometer, B. Miller, M. Bertin, V. Hook and W. H. Gerwick for their scientific contribution and A. Saitman, and P. Boudreau for helpful discussions. I acknowledge the use of the UCSD Chemistry and Biochemistry mass spectrometry and NMR facilities for analytical services. I also acknowledge The Growth Regulation & Oncogenesis Training Grant NIH/NCI (T32A009523-24) for a fellowship, and NIH Grants CA 100851 and GM 107550 for support of this research.

## 6.6 Appendix

**Figure 6.6.1:**  $^1\text{H}$  NMR (500 MHz, MeOD) spectrum of compound **18a****Figure 6.6.2:**  $^{13}\text{C}$  NMR (125 MHz, MeOD) spectrum of compound **18a**

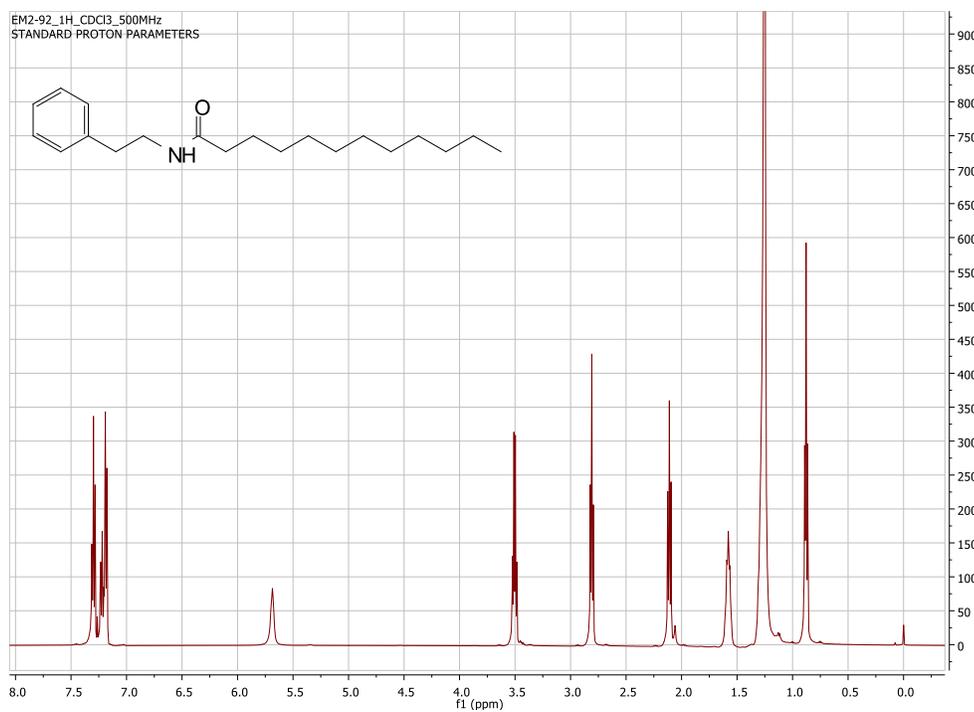


Figure 6.6.3: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 18b

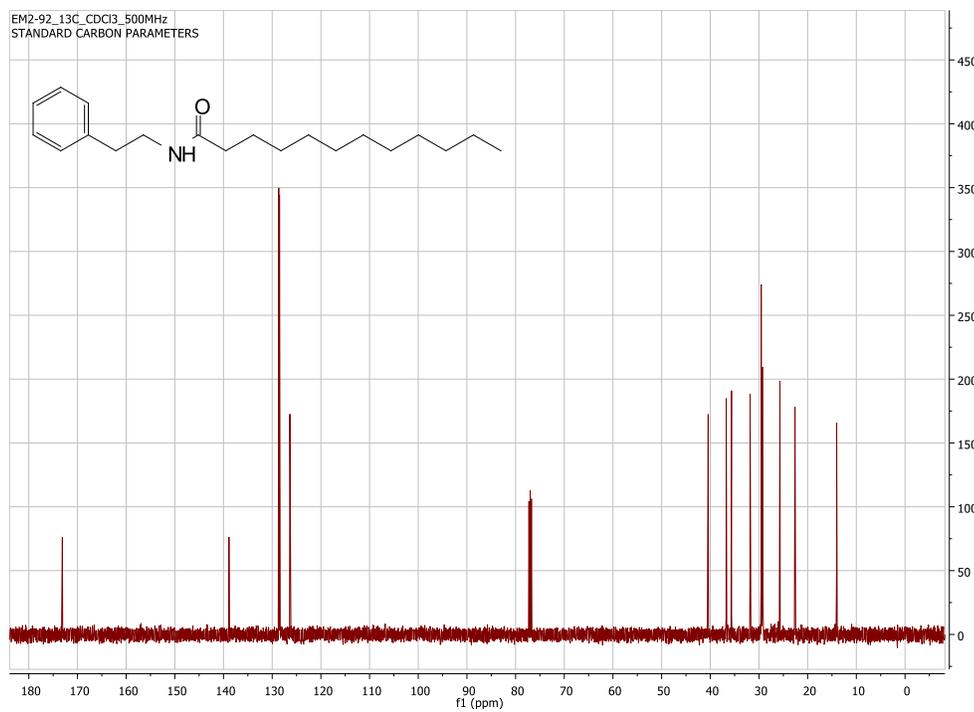
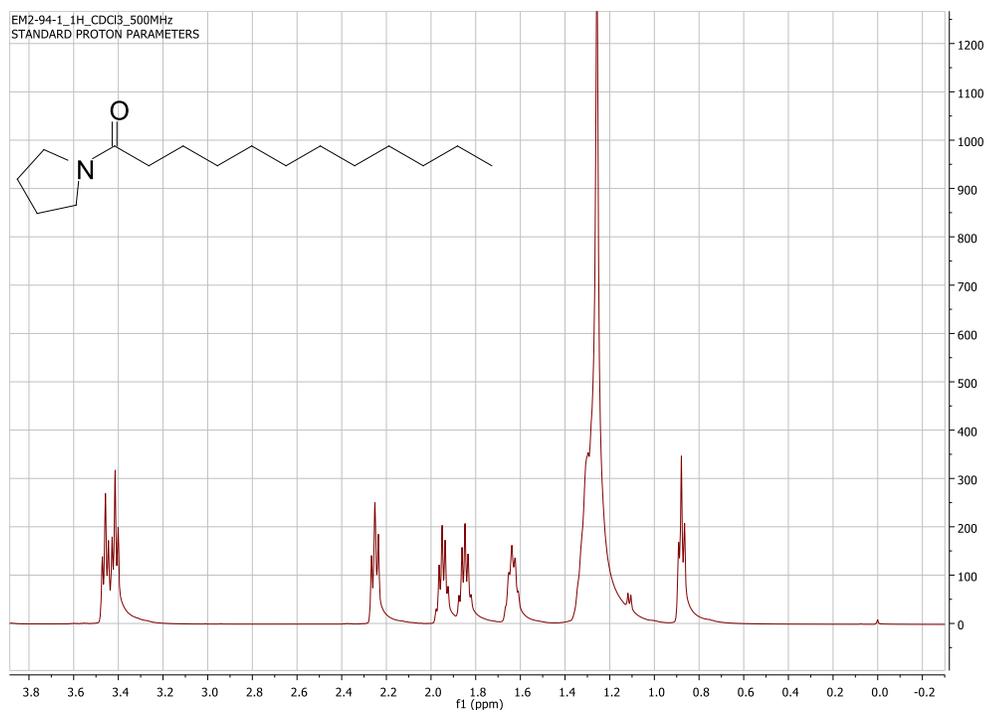
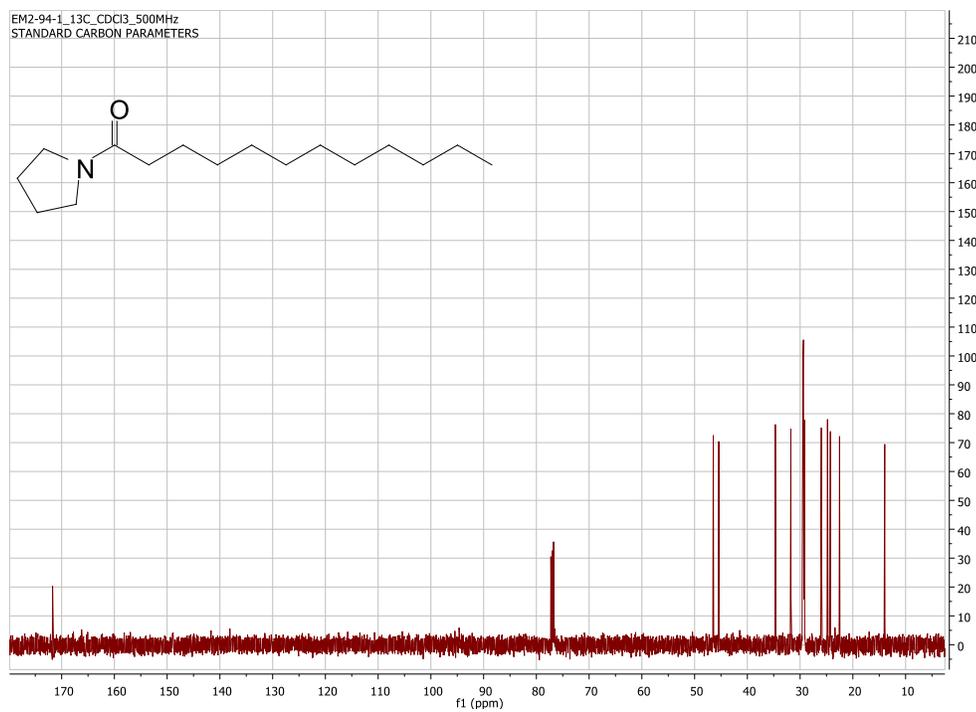


Figure 6.6.4: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 18b



**Figure 6.6.5:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **18c**



**Figure 6.6.6:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **18c**

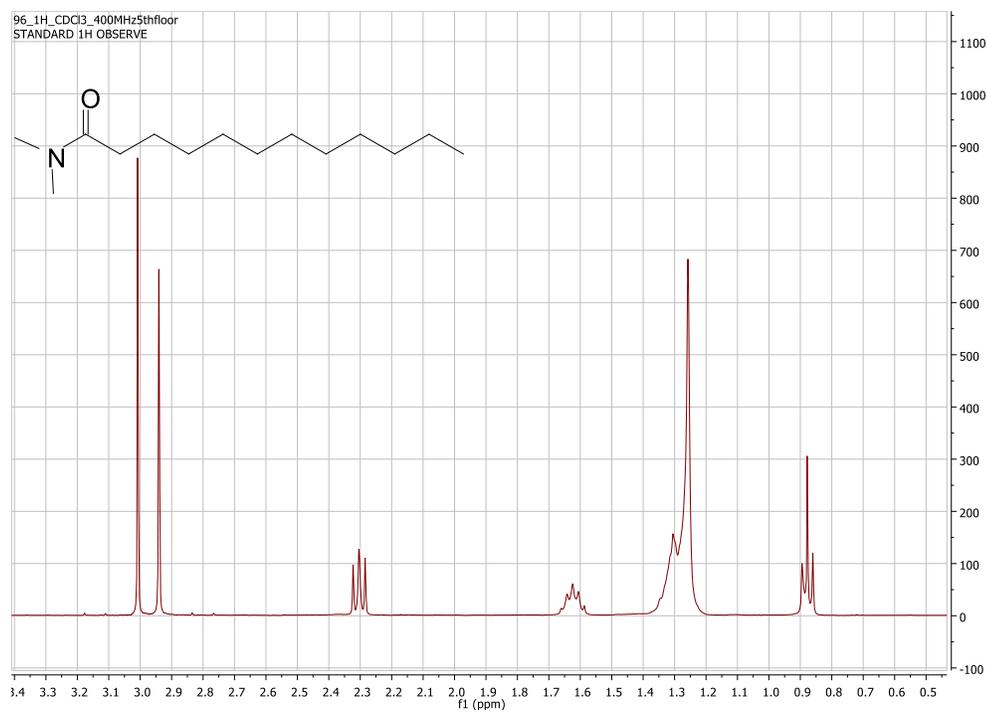


Figure 6.6.7: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 18d

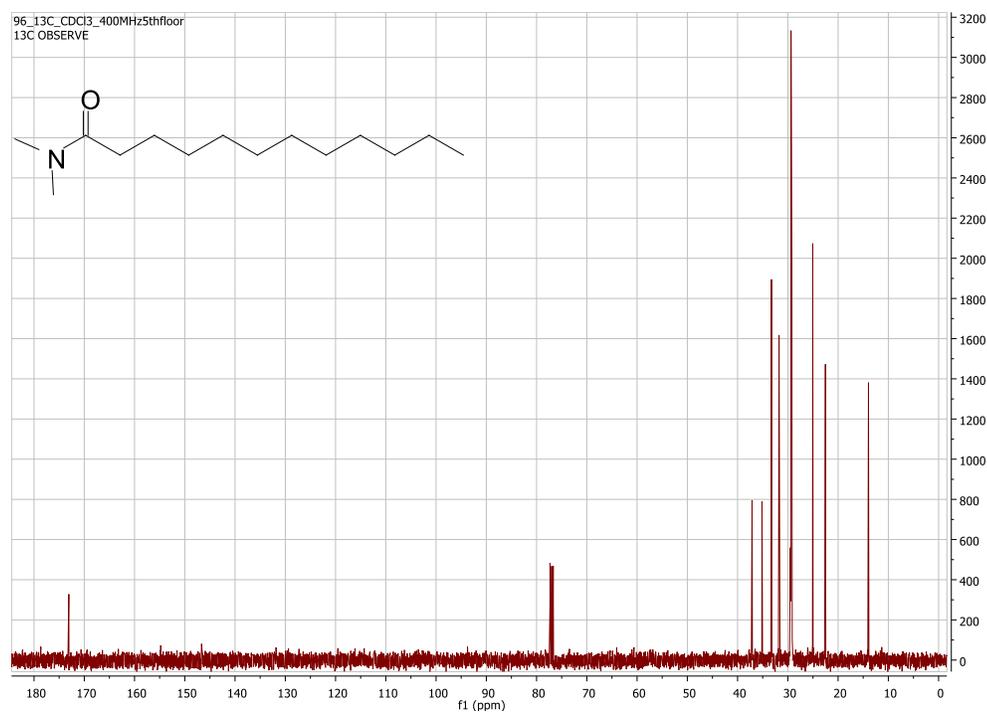
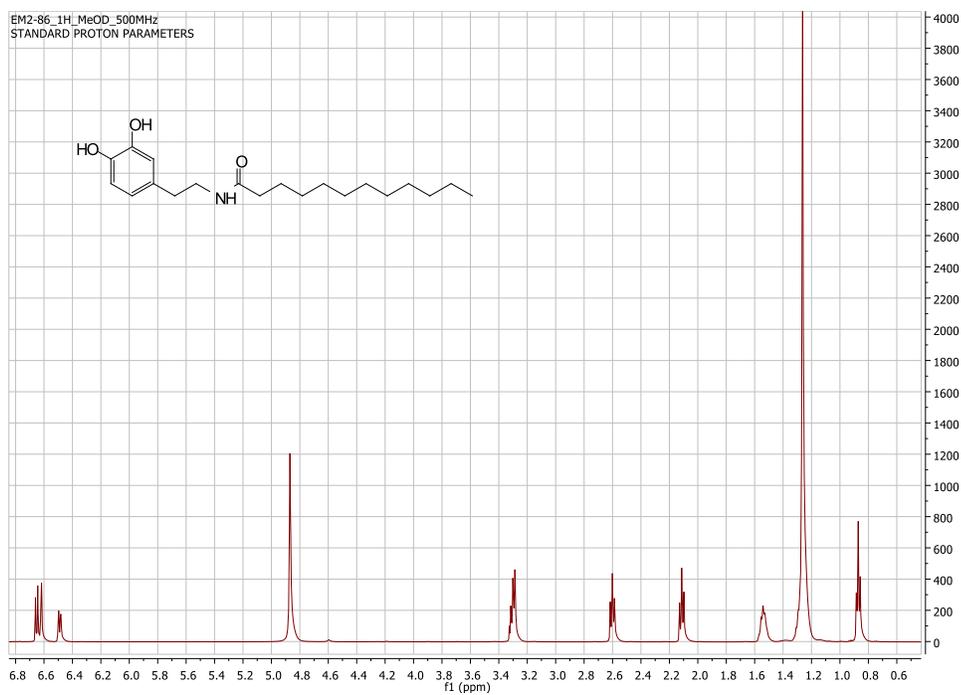
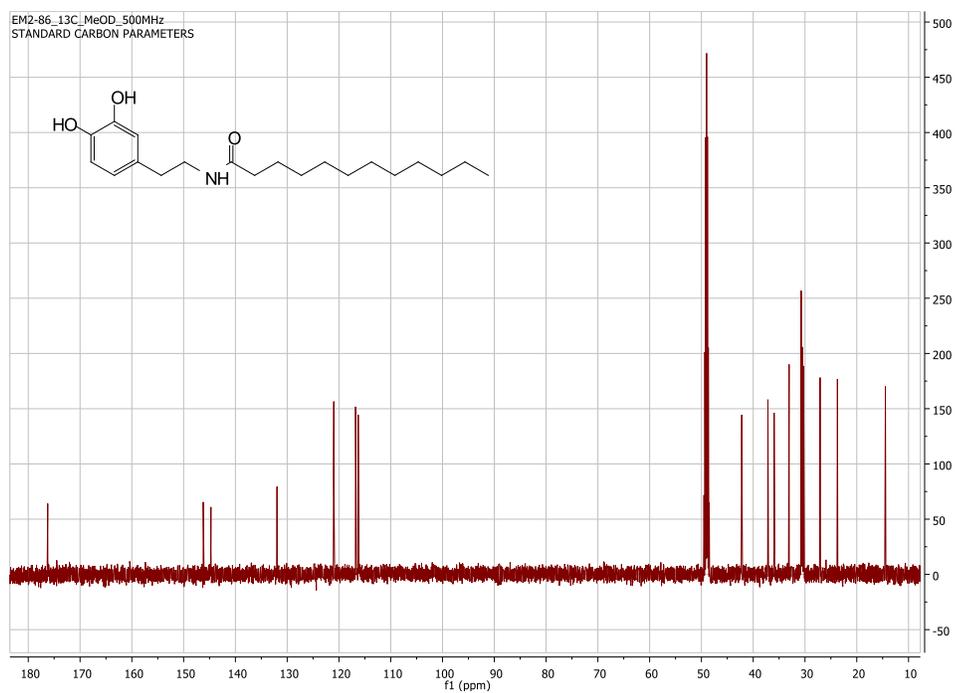


Figure 6.6.8: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 18d



**Figure 6.6.9:** <sup>1</sup>H NMR (500 MHz, MeOD) spectrum of compound **18e**



**Figure 6.6.10:** <sup>13</sup>C NMR (125 MHz, MeOD) spectrum of compound **18e**

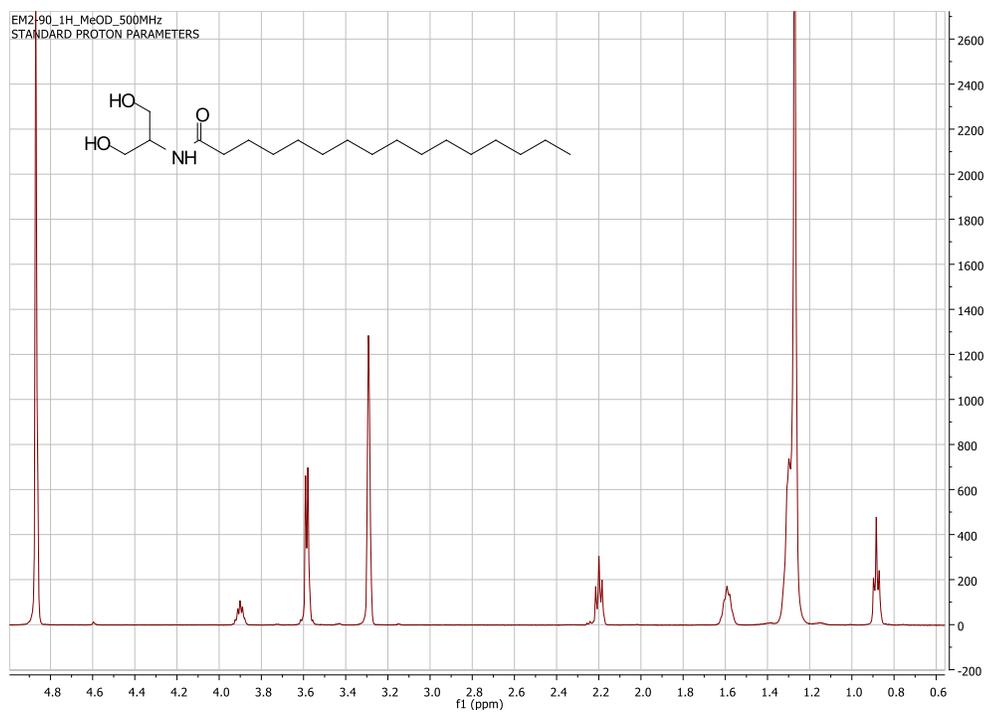


Figure 6.6.11:  $^1\text{H}$  NMR (500 MHz, MeOD) spectrum of compound **19a**

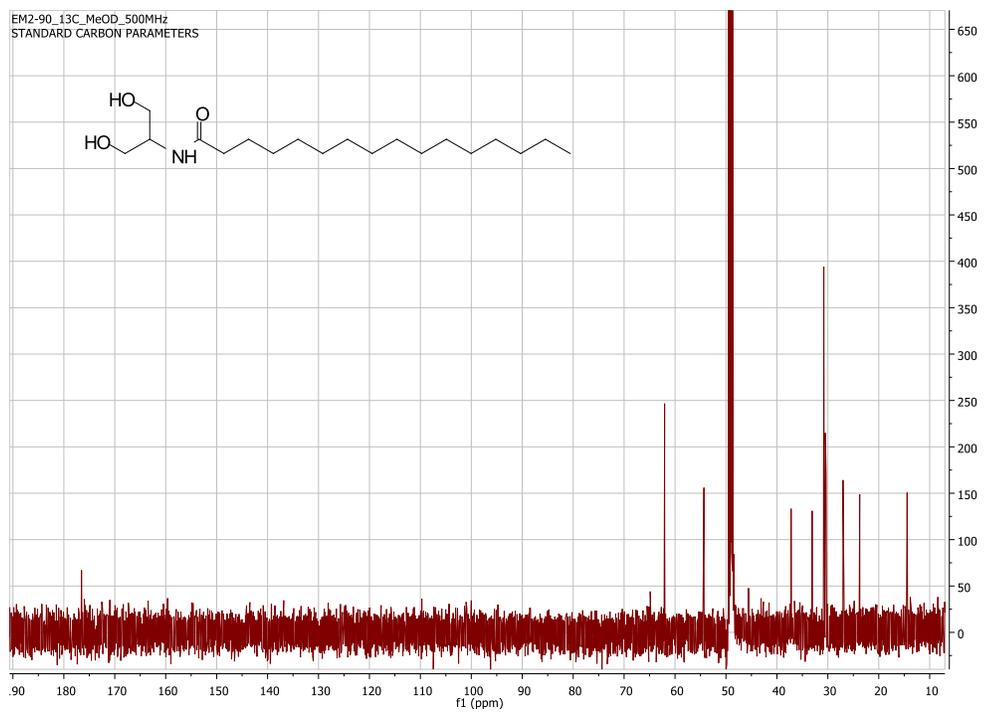


Figure 6.6.12:  $^{13}\text{C}$  NMR (125 MHz, MeOD) spectrum of compound **19a**

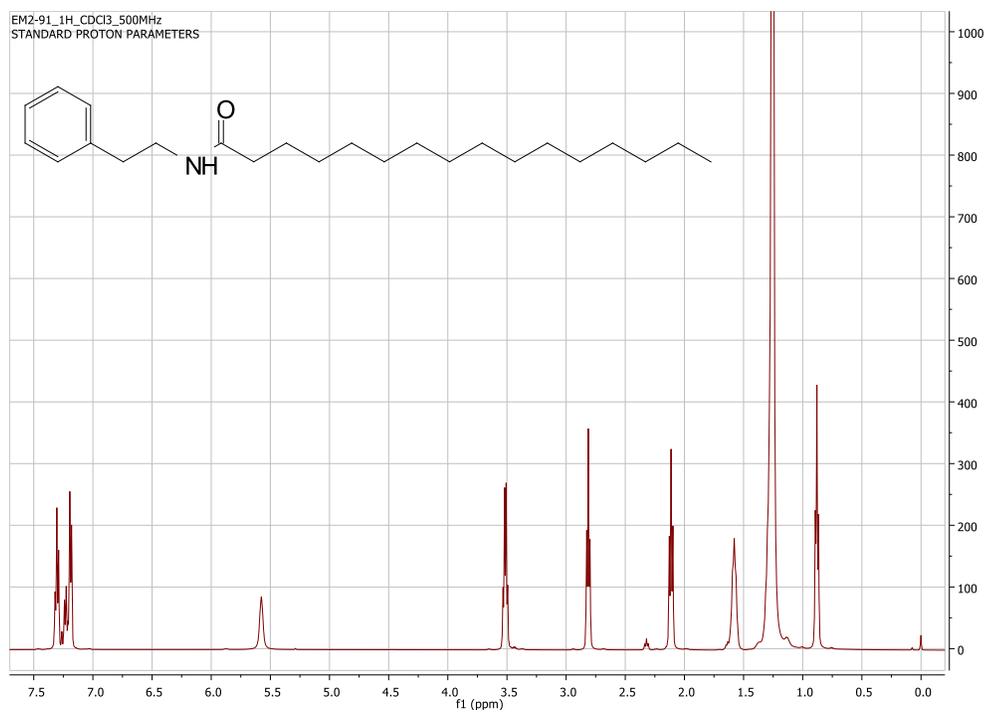


Figure 6.6.13: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 19b

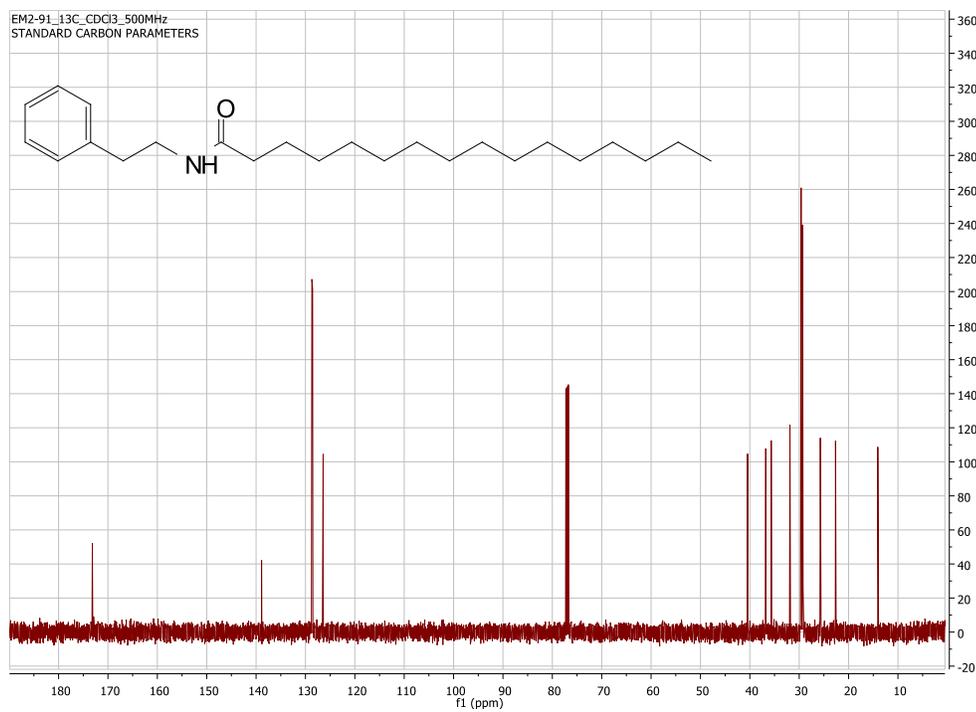


Figure 6.6.14: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 19b

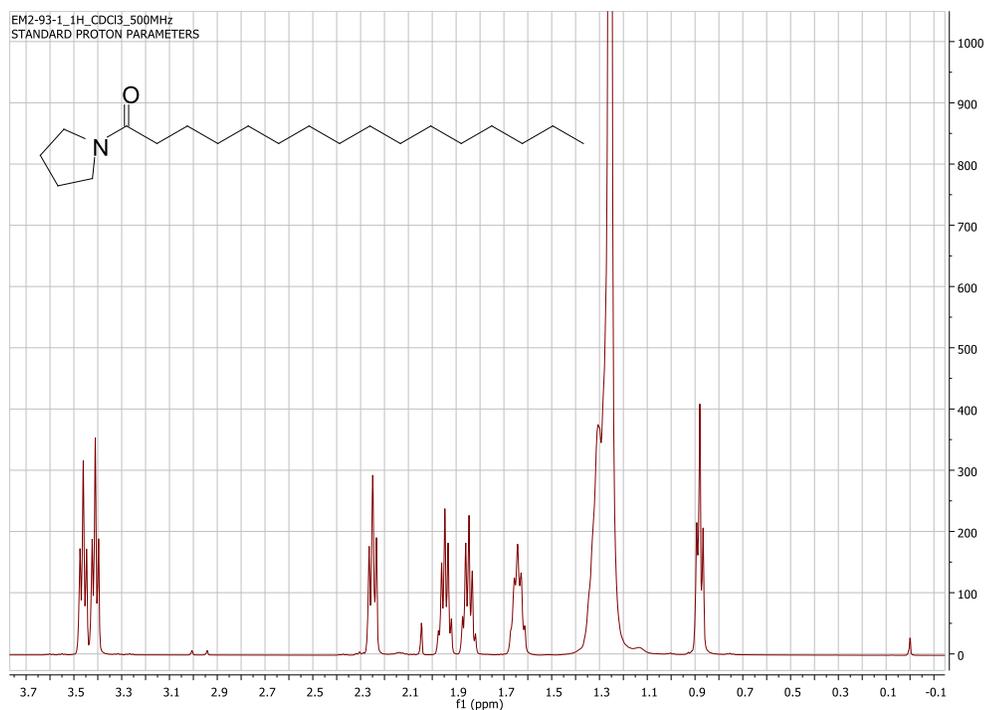


Figure 6.6.15: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 19c

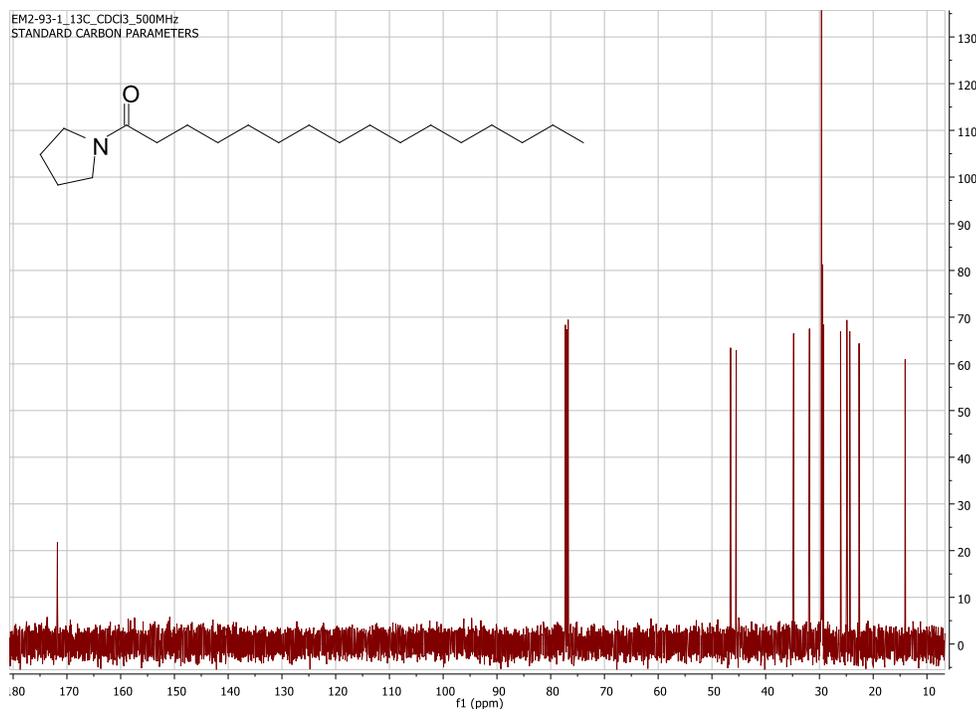
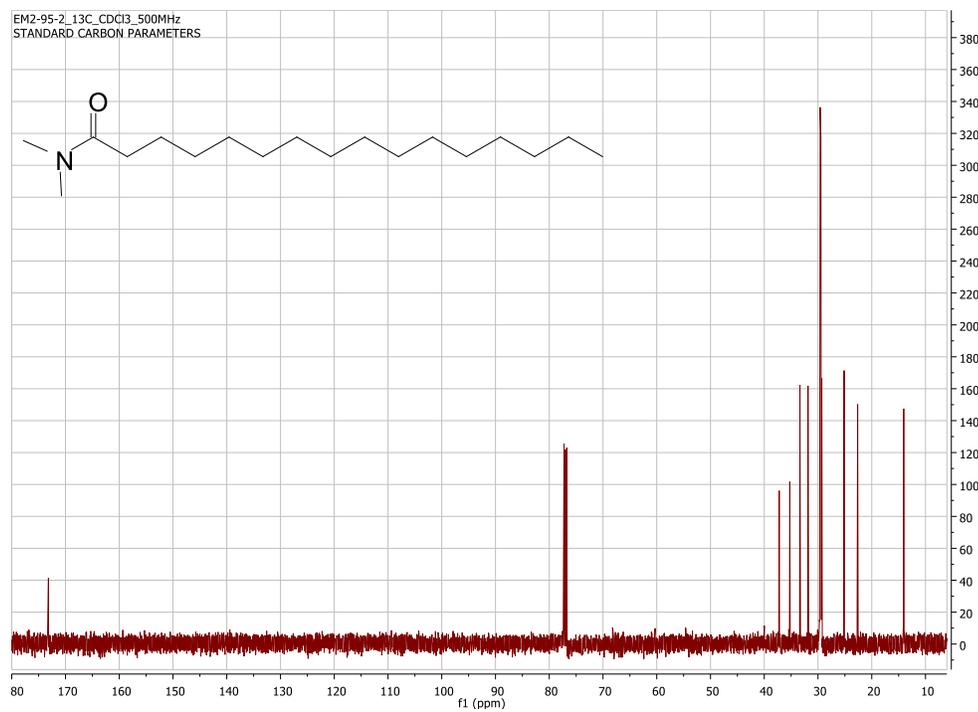
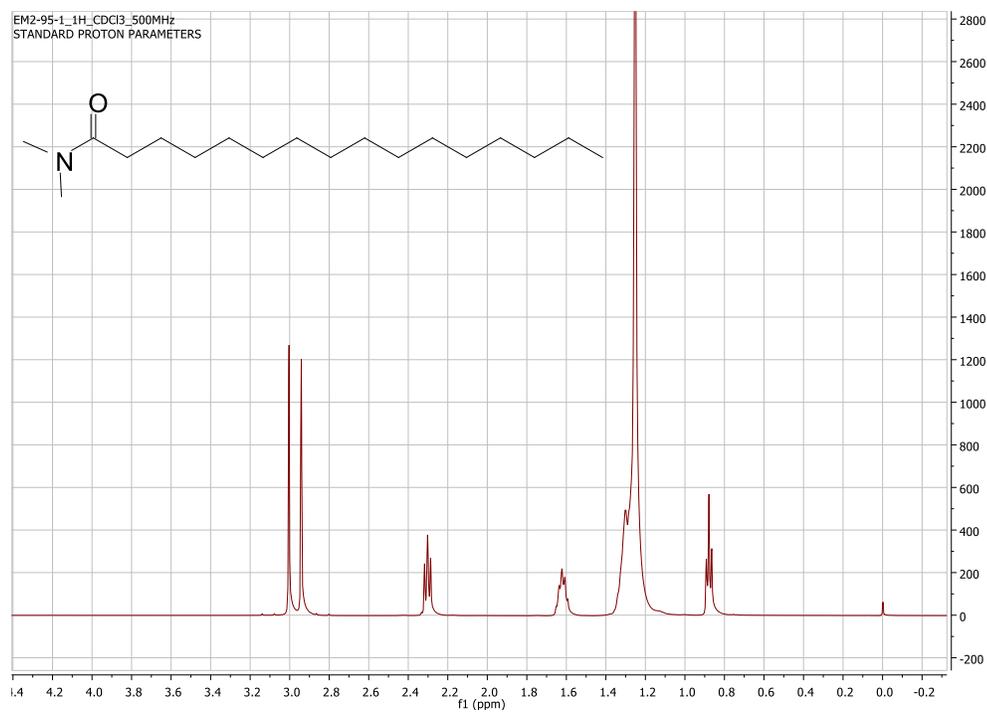
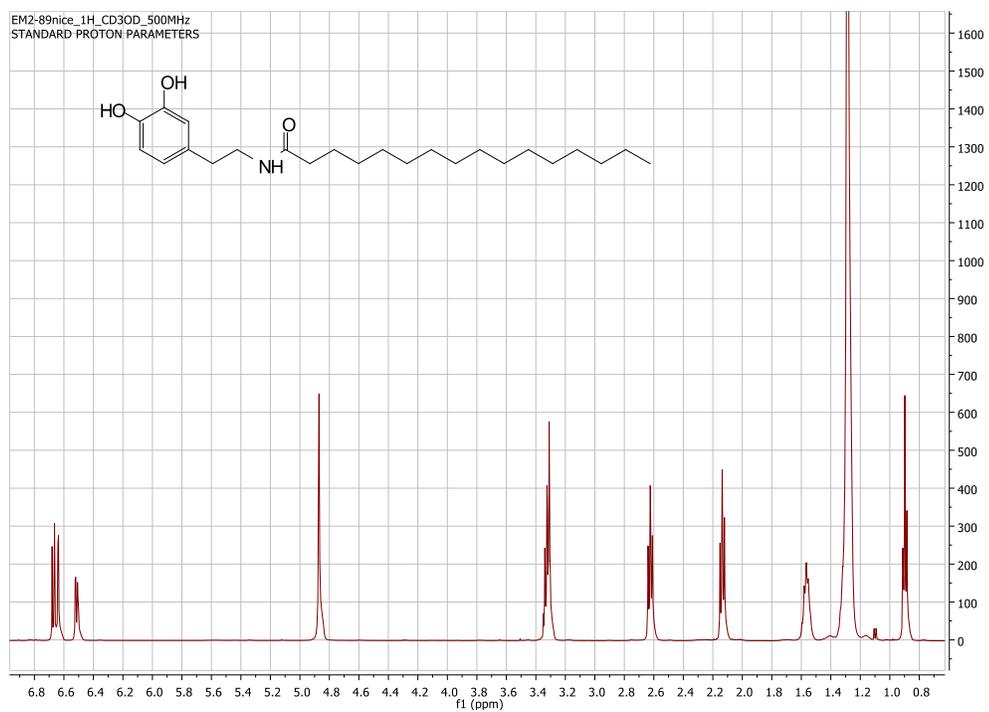
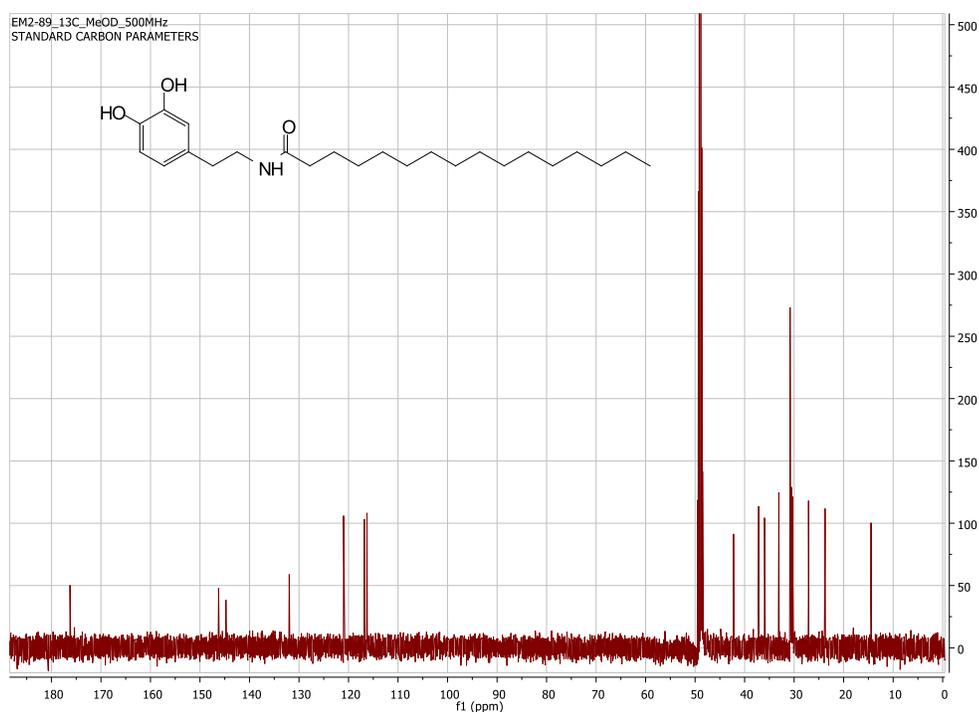


Figure 6.6.16: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 19c

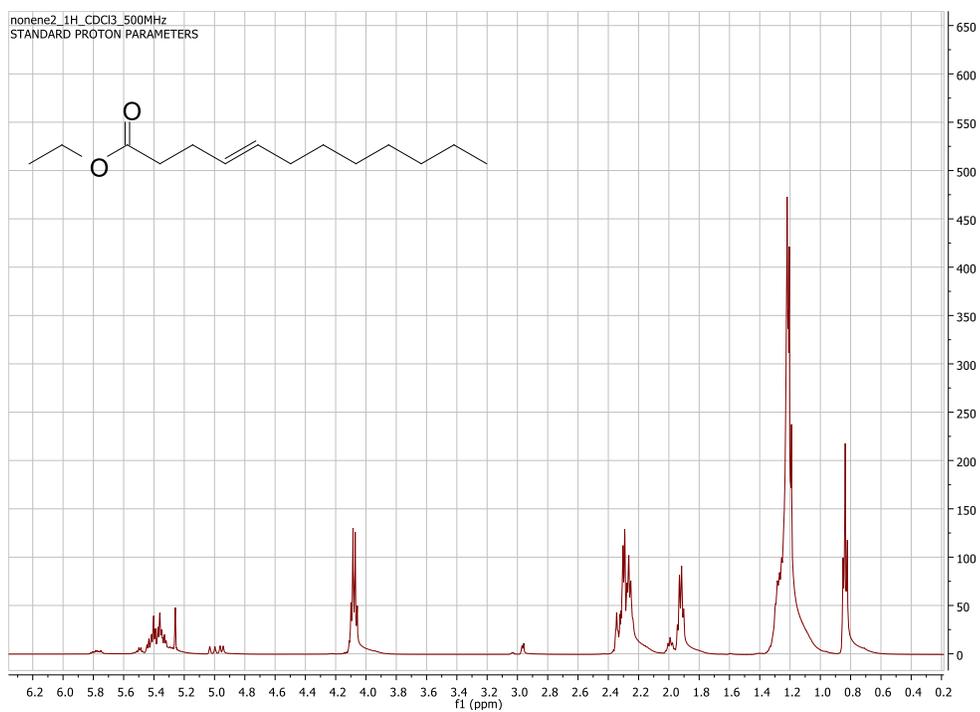




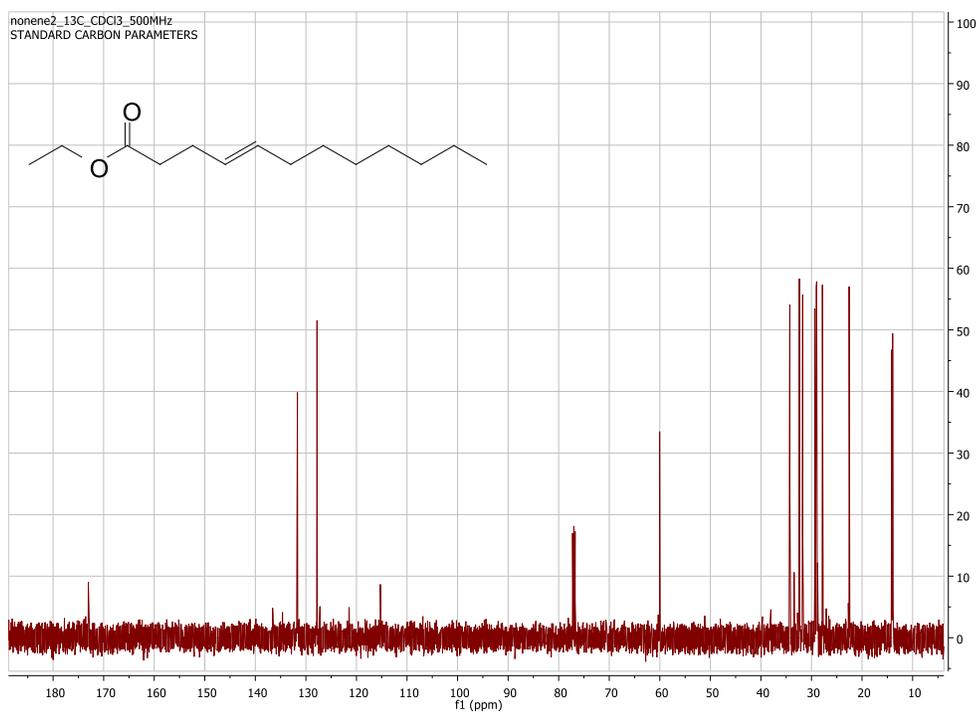
**Figure 6.6.19:** <sup>1</sup>H NMR (500 MHz, MeOD) spectrum of compound 19e



**Figure 6.6.20:** <sup>13</sup>C NMR (125 MHz, MeOD) spectrum of compound 19e



**Figure 6.6.21:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **23**



**Figure 6.6.22:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **23**

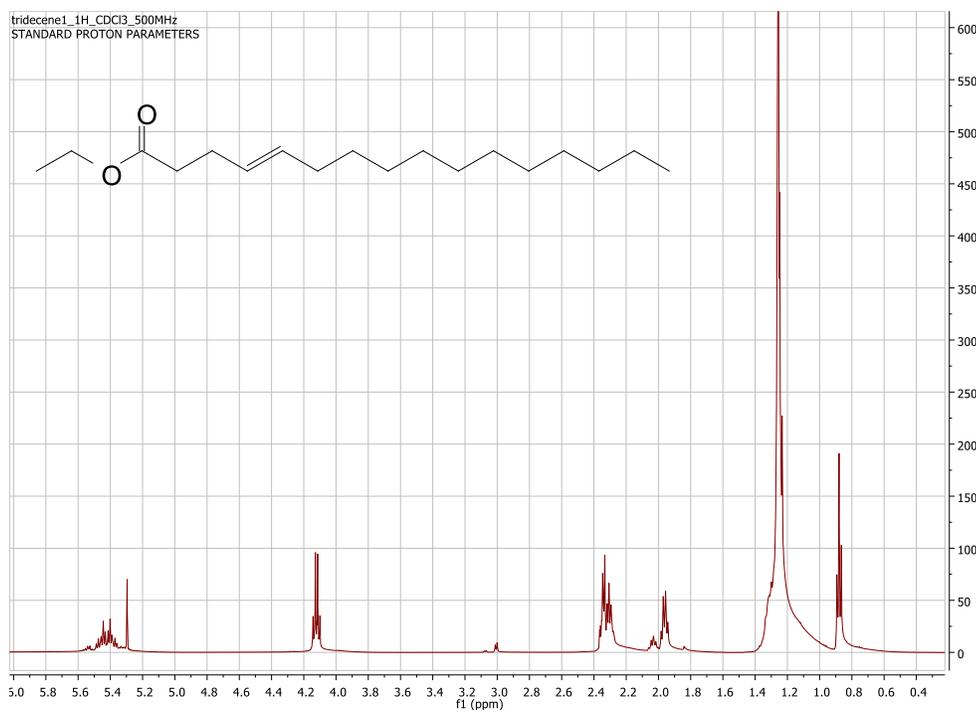


Figure 6.6.23: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 24

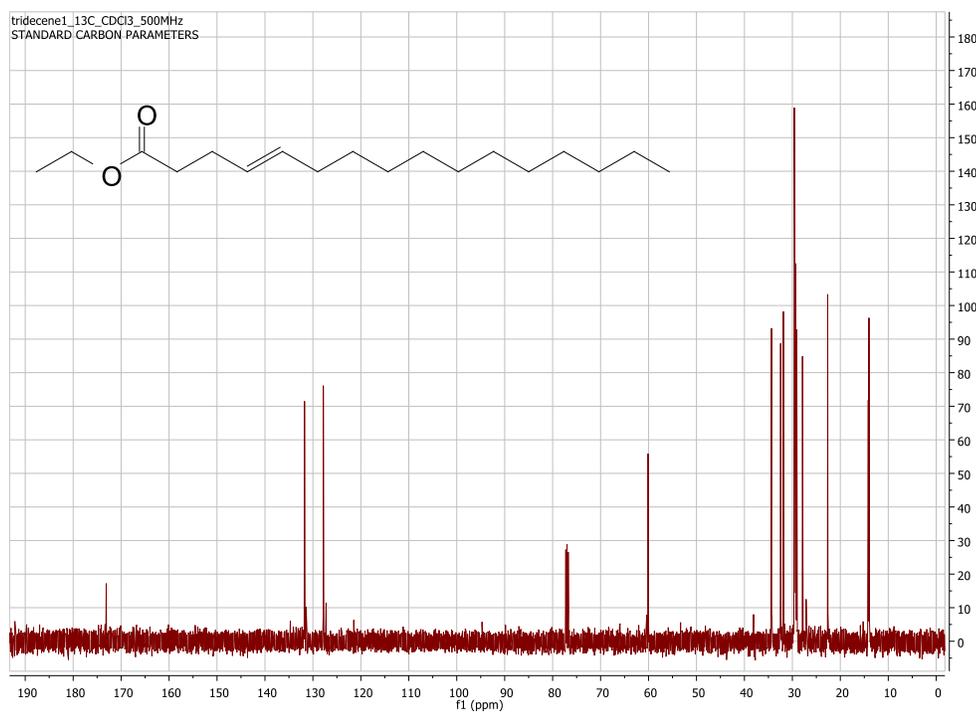


Figure 6.6.24: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 24

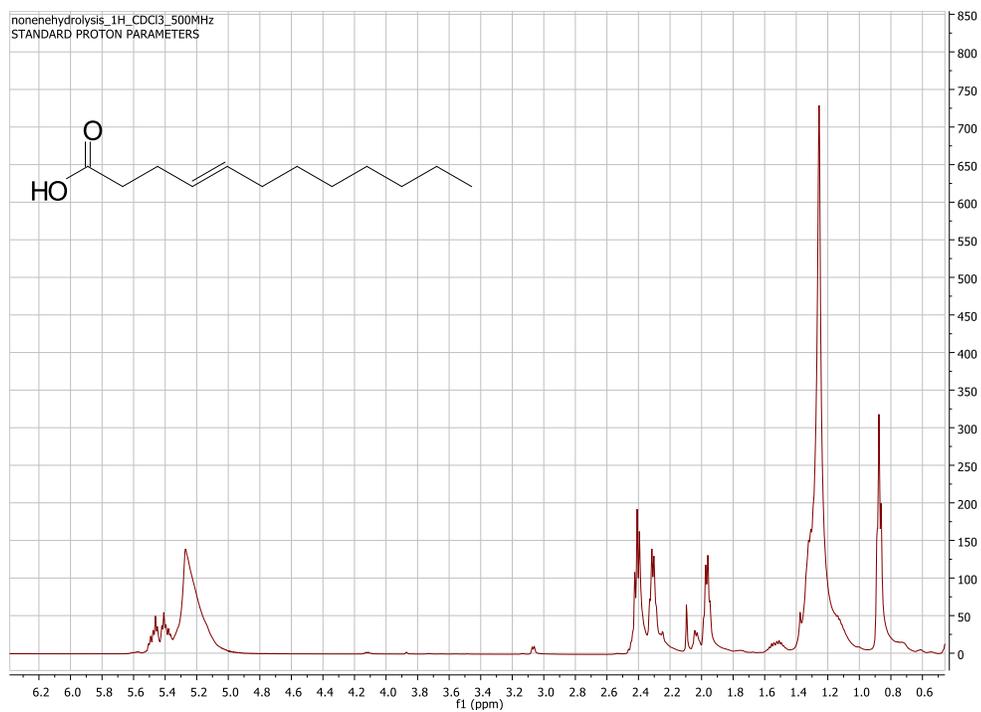


Figure 6.6.25: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 25

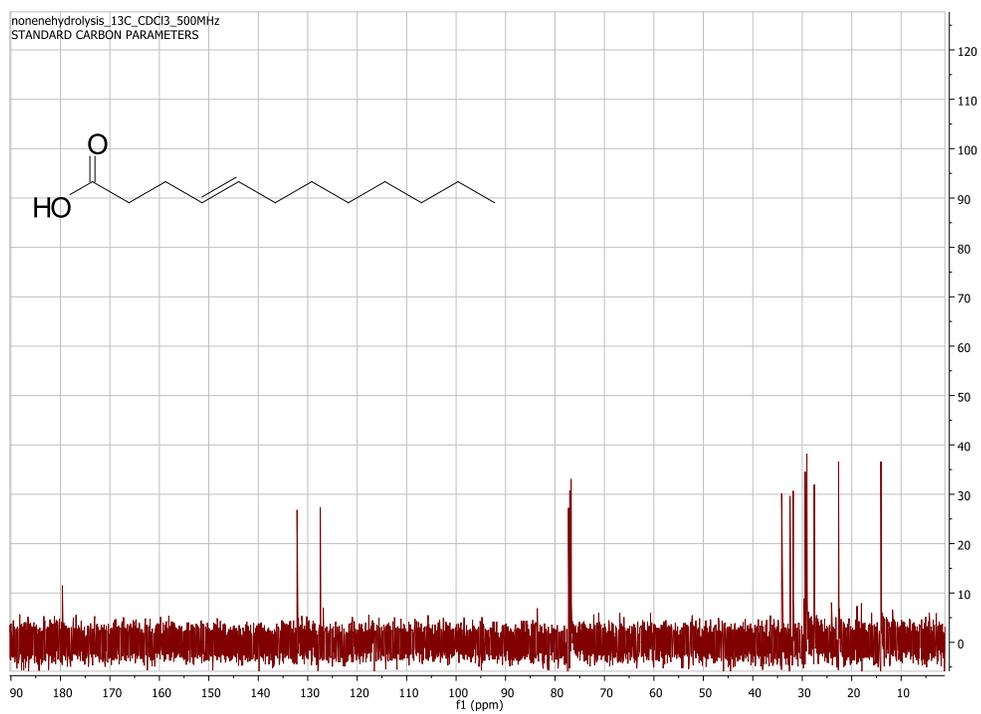


Figure 6.6.26: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 25

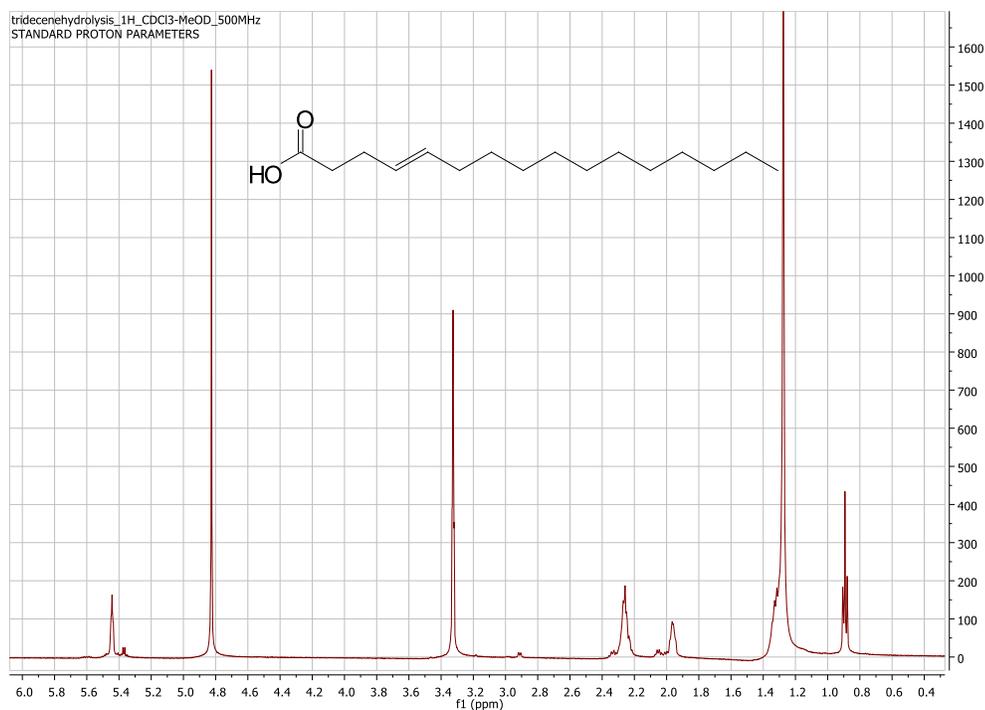


Figure 6.6.27: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 26

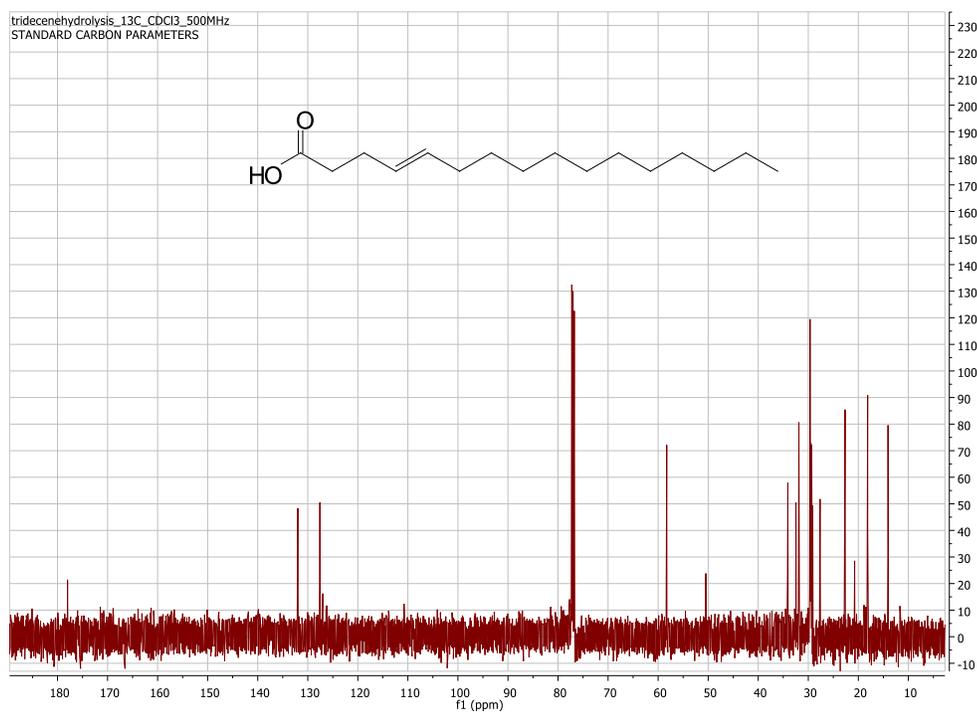


Figure 6.6.28: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 26

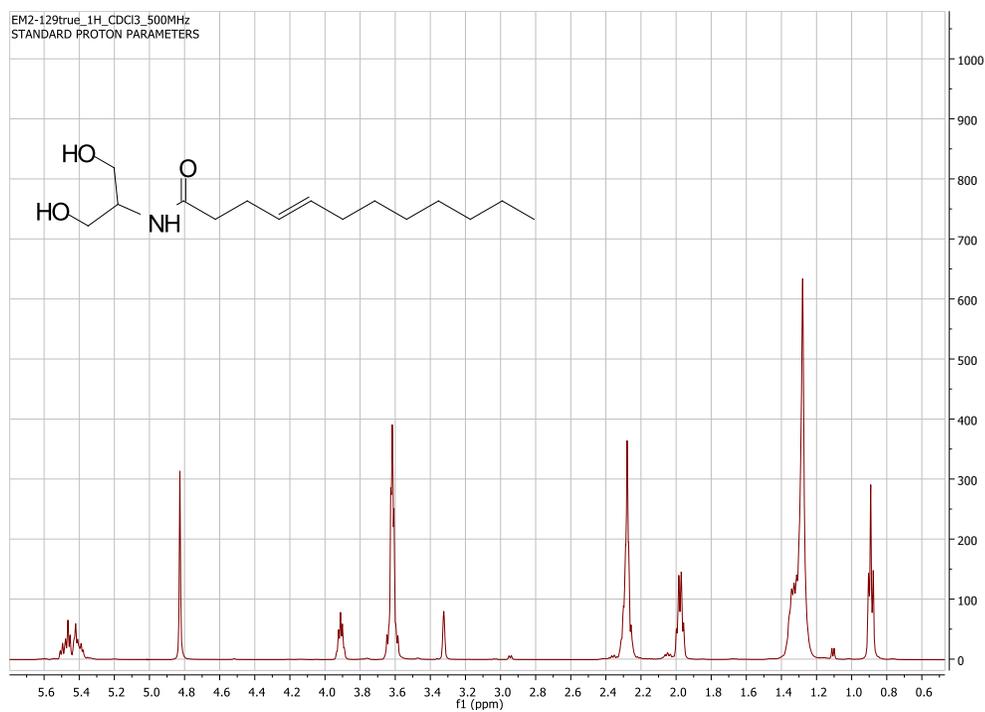


Figure 6.6.29: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 27a

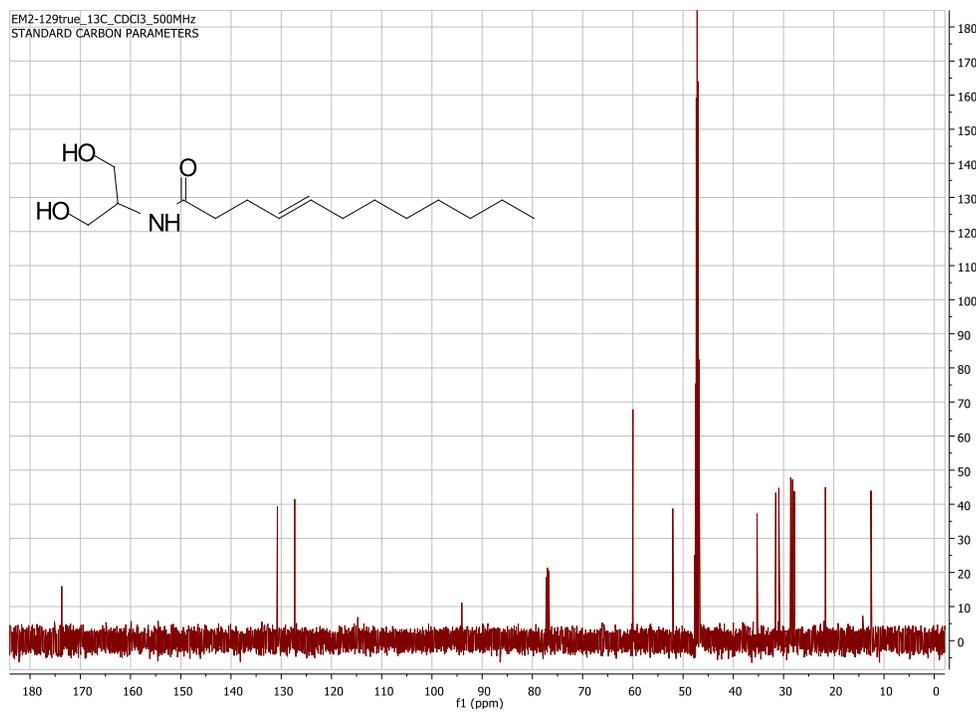


Figure 6.6.30: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 27a

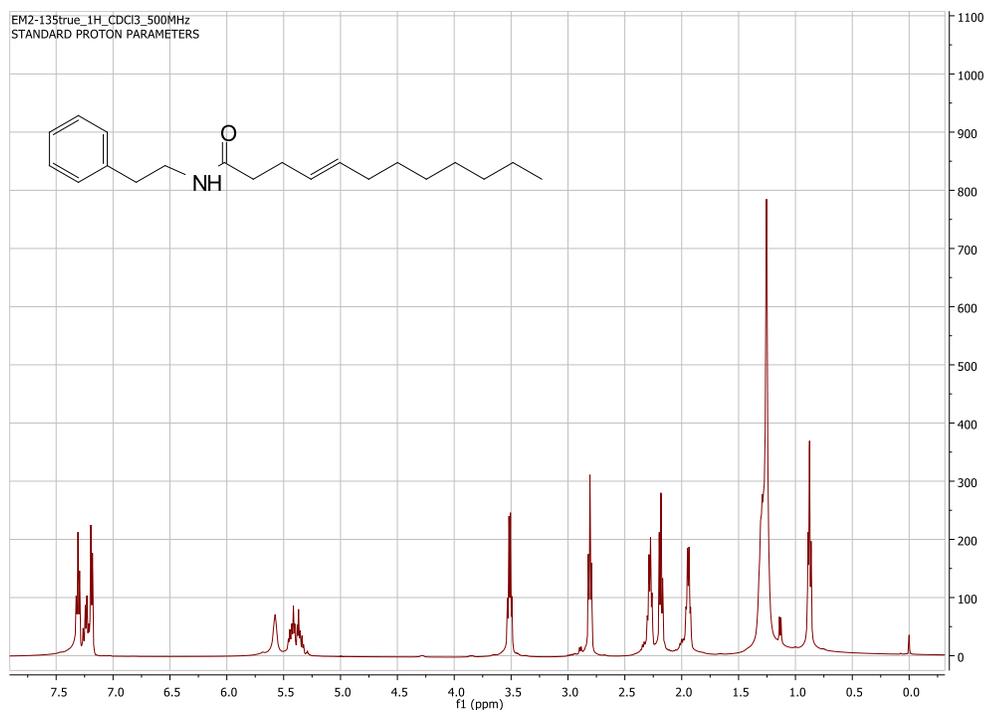


Figure 6.6.31: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 27b

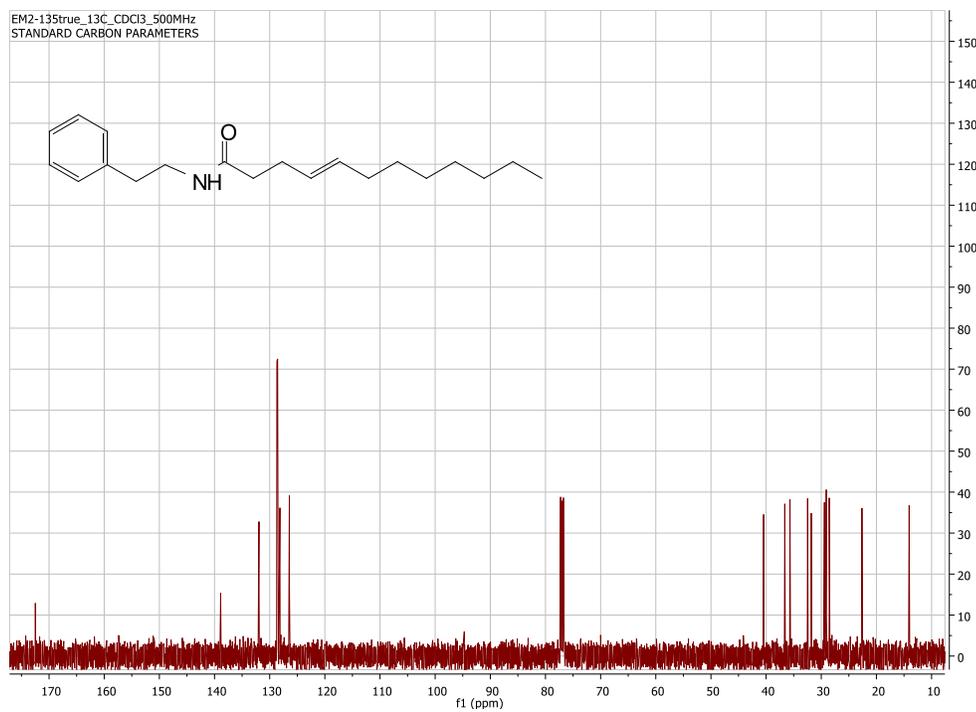


Figure 6.6.32: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 27b

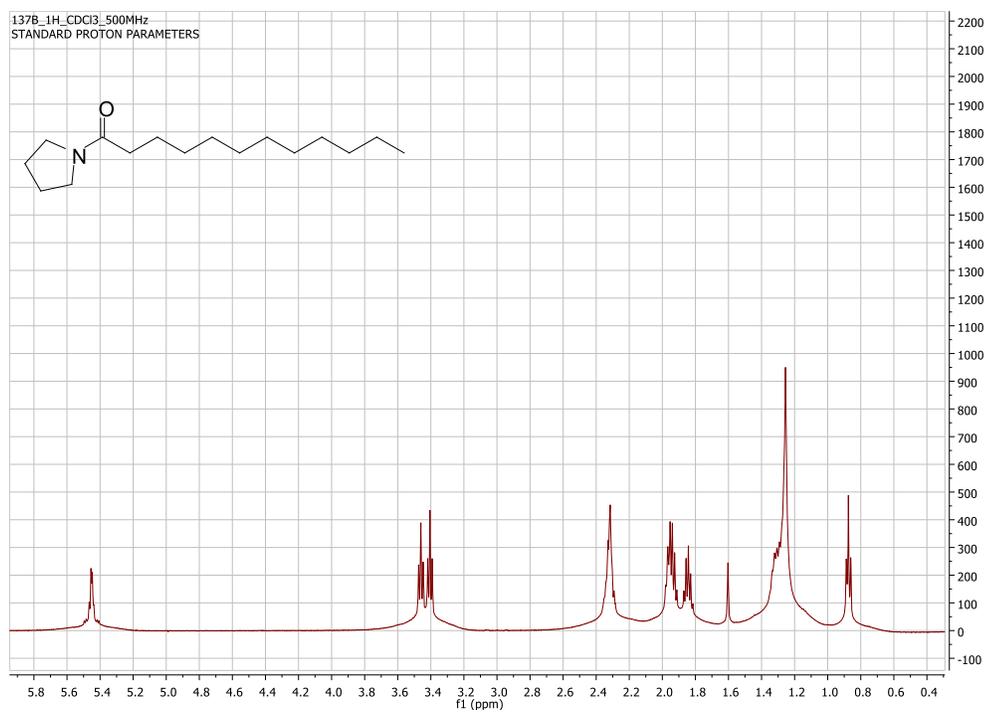


Figure 6.6.33:  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of compound 27c

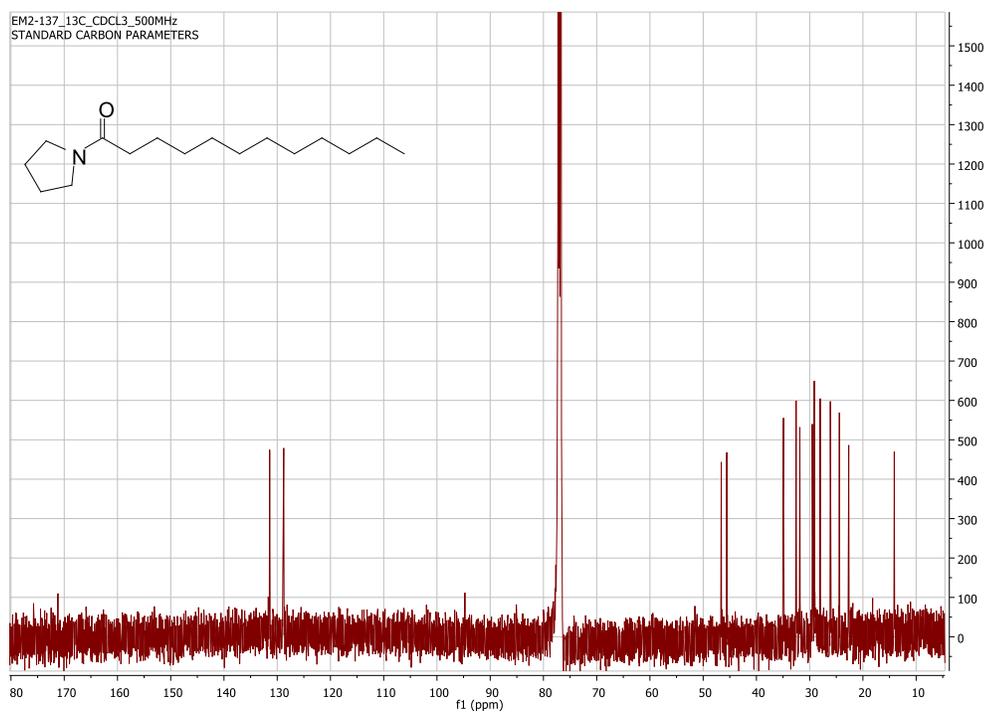
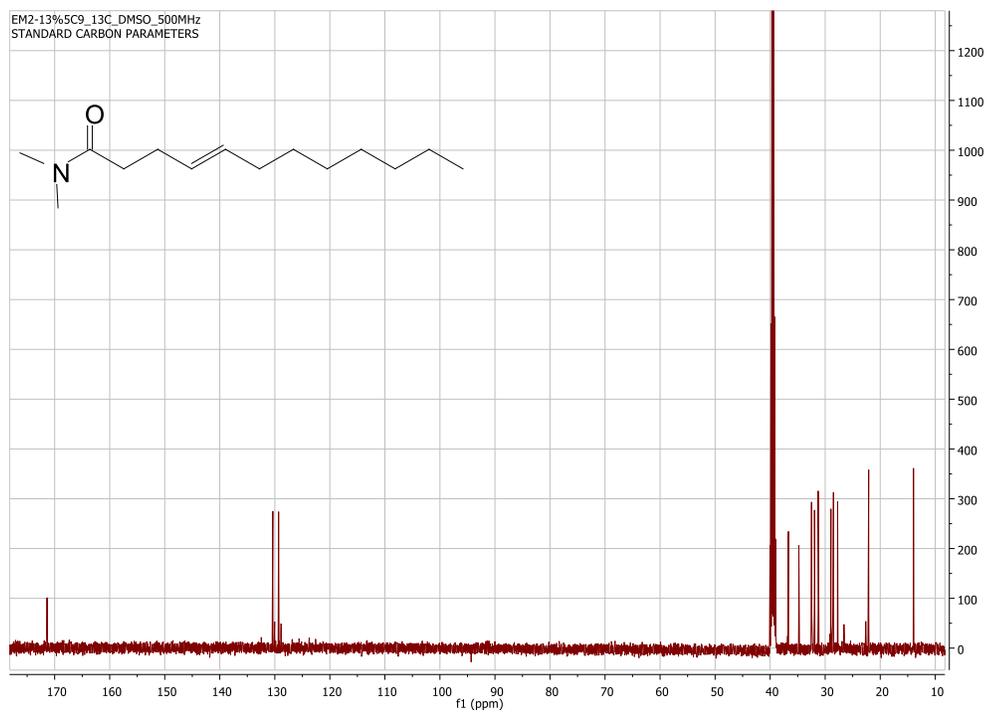
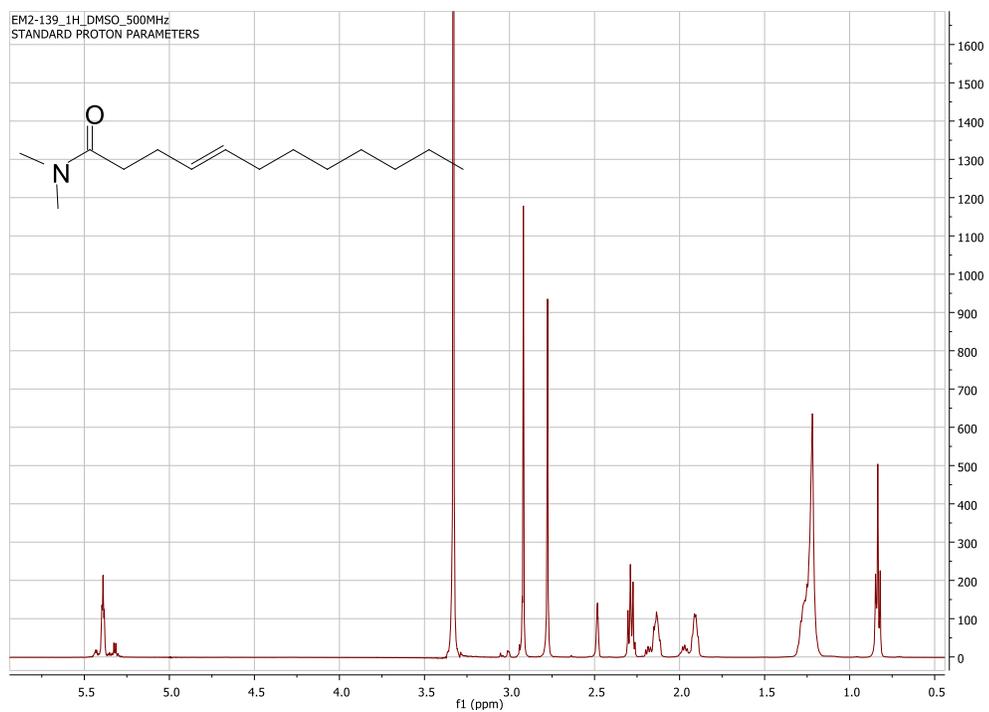


Figure 6.6.34:  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of compound 27c



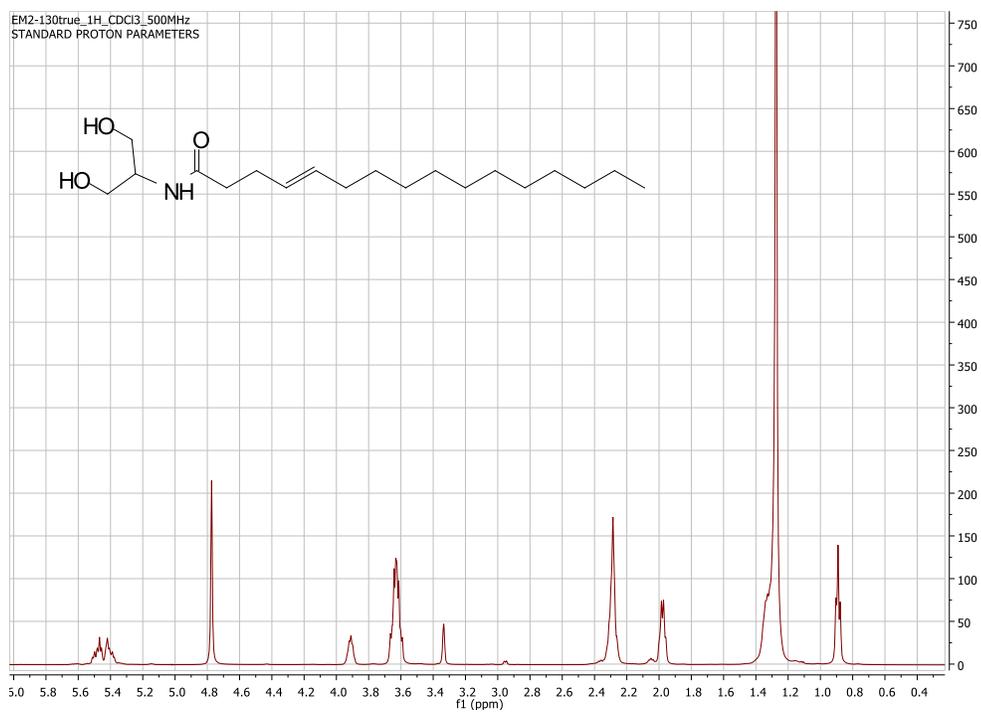


Figure 6.6.37:  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of compound **28a**

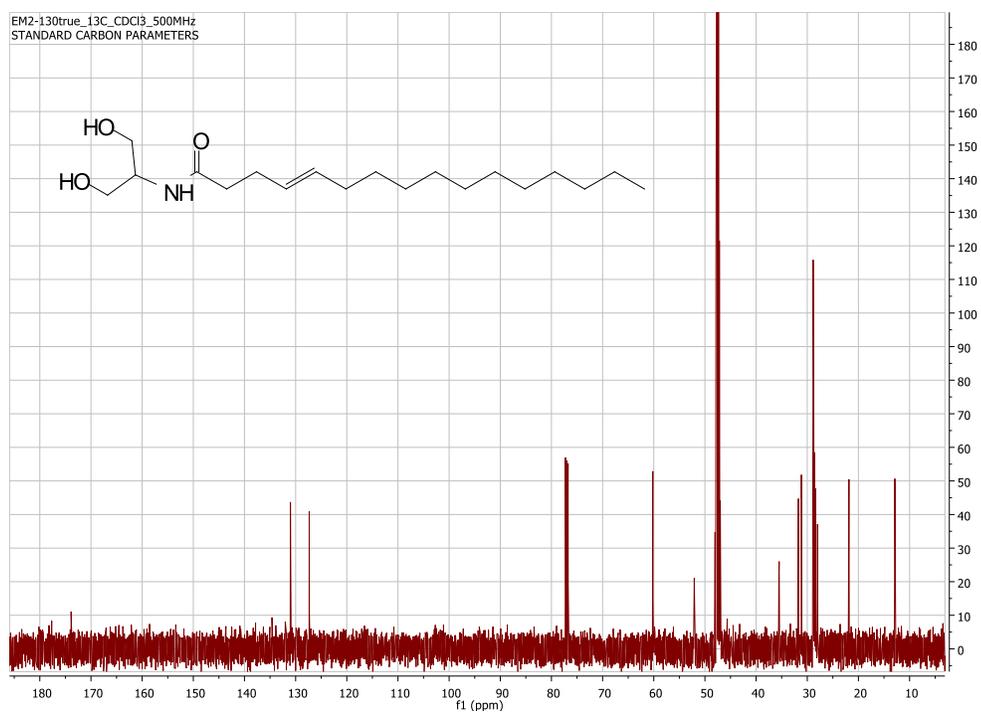


Figure 6.6.38:  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of compound **28a**

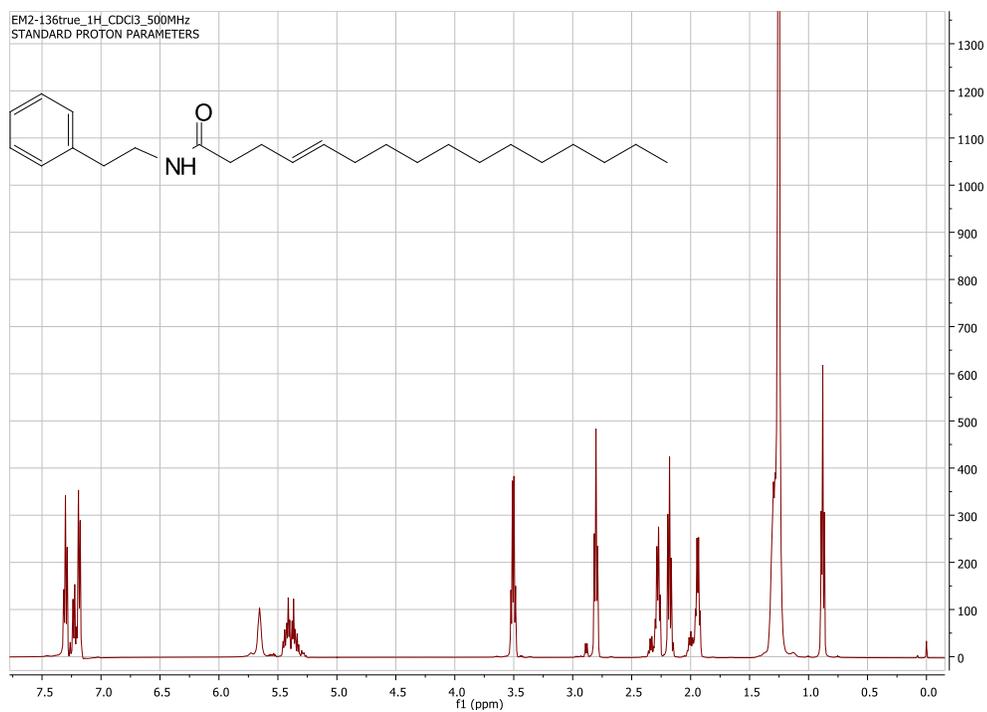


Figure 6.6.39: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 28b

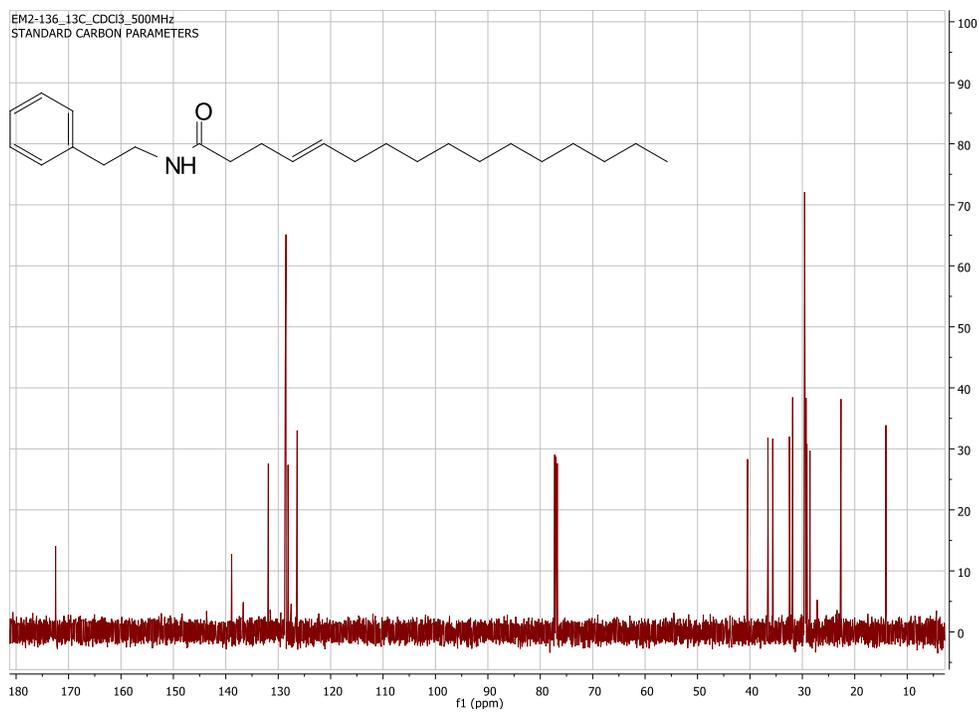


Figure 6.6.40: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 28b

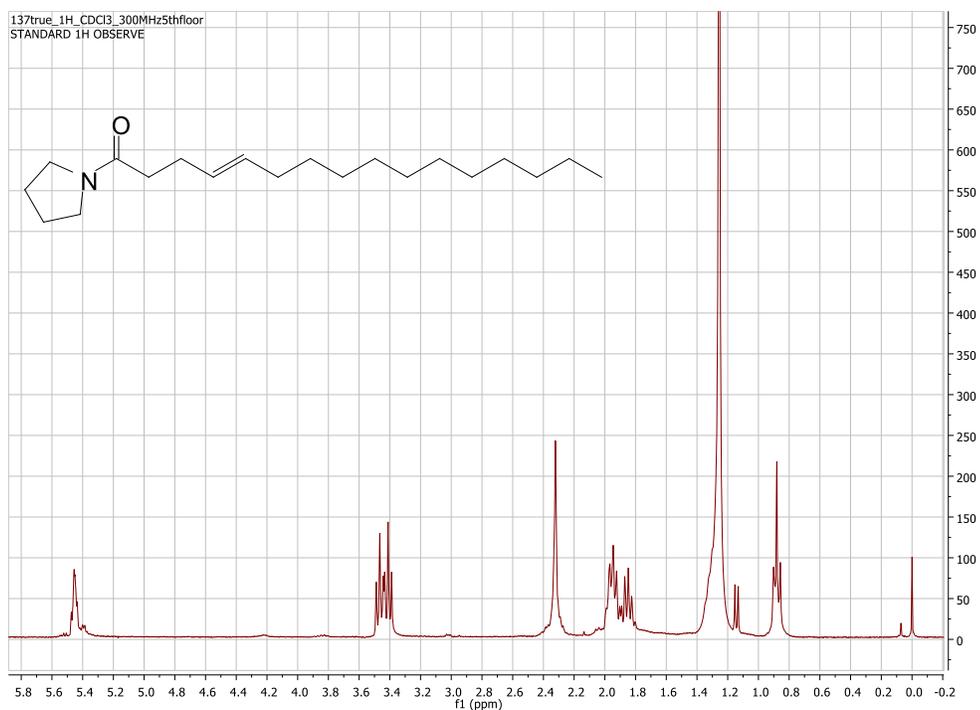


Figure 6.6.41: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 28c

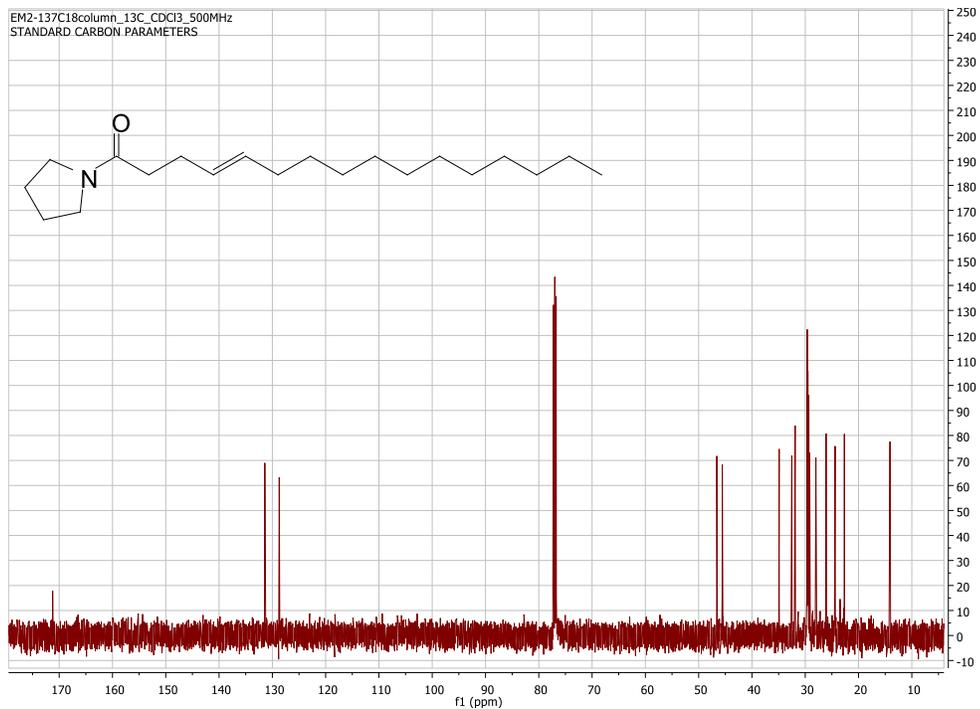
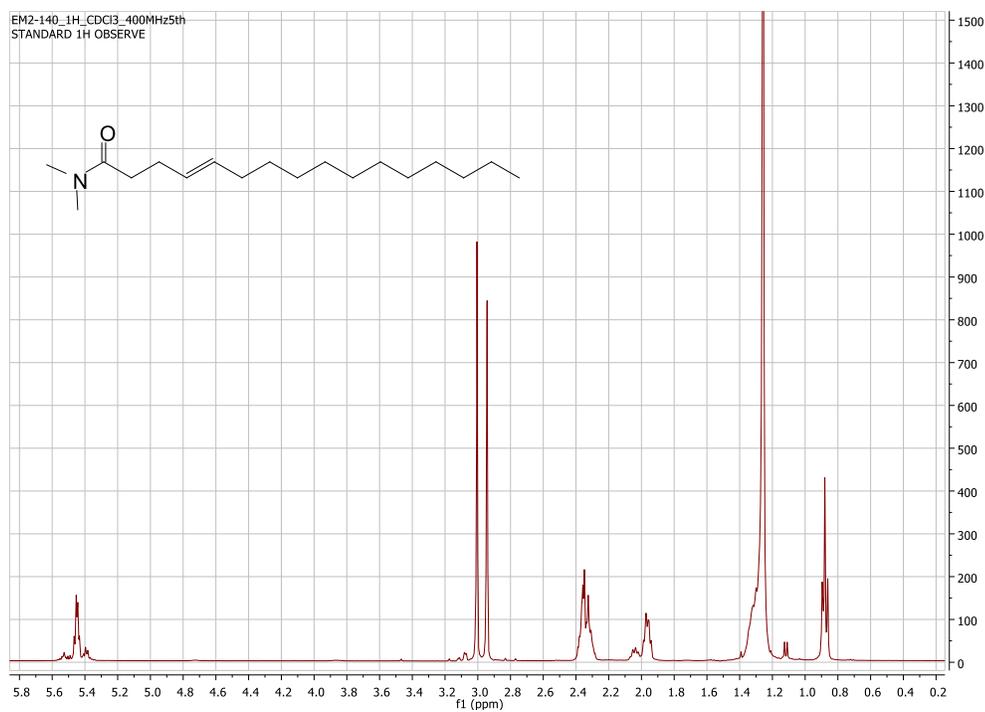
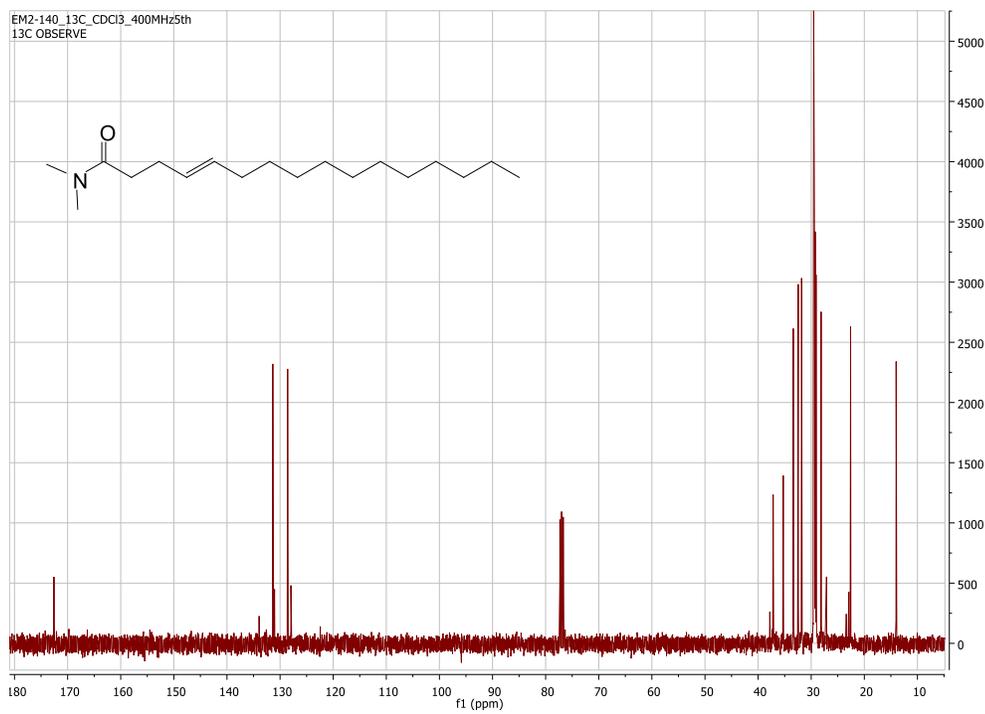


Figure 6.6.42: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 28c



**Figure 6.6.43:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound **28d**



**Figure 6.6.44:** <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound **28d**

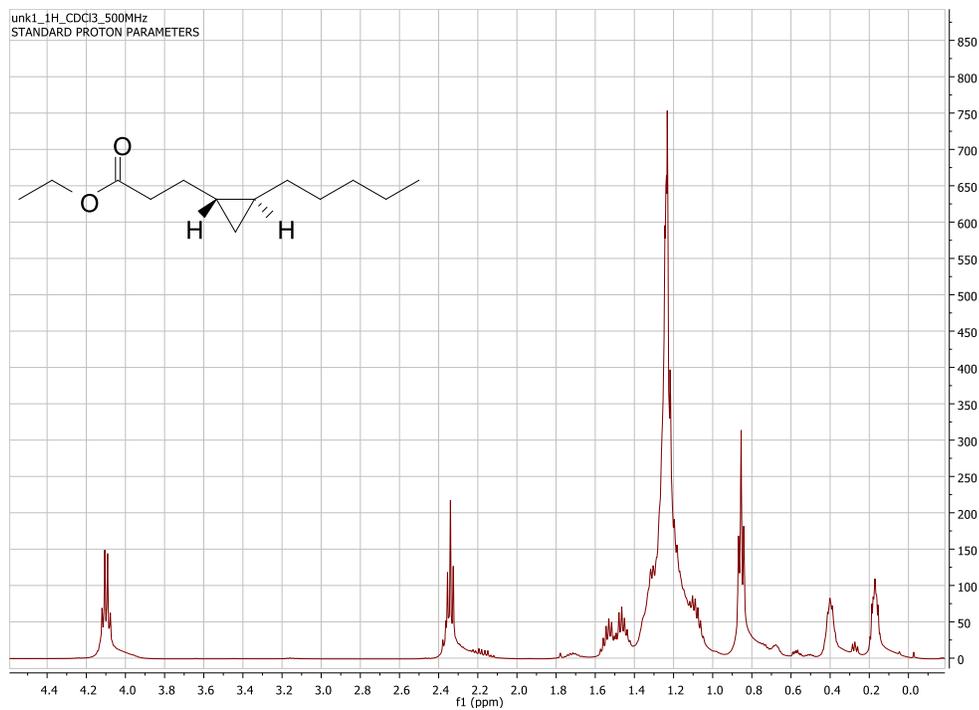


Figure 6.6.45: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 29

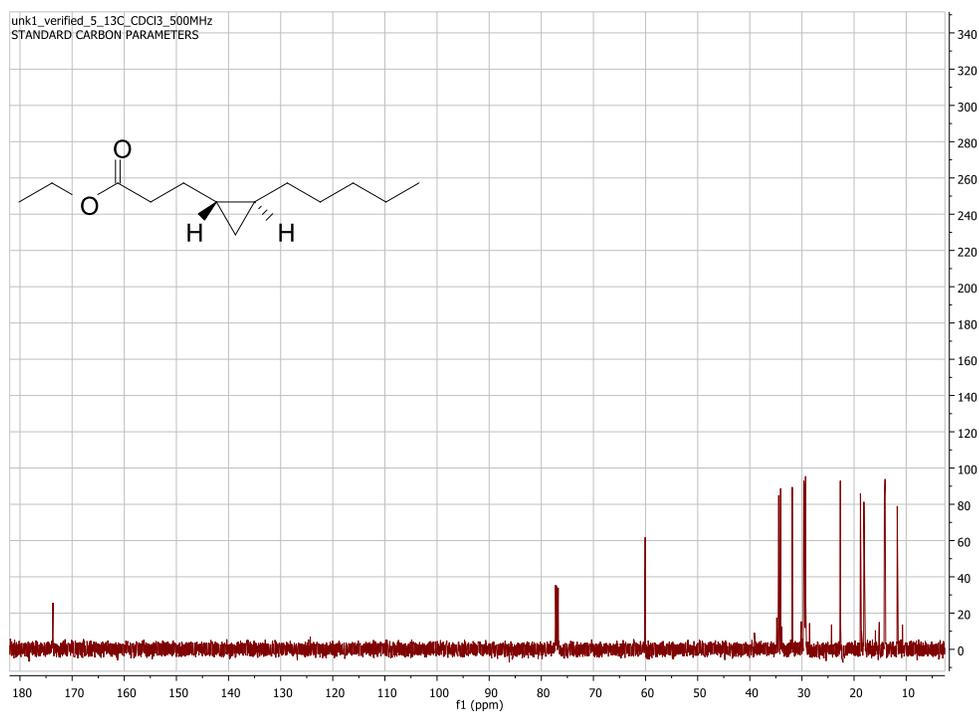
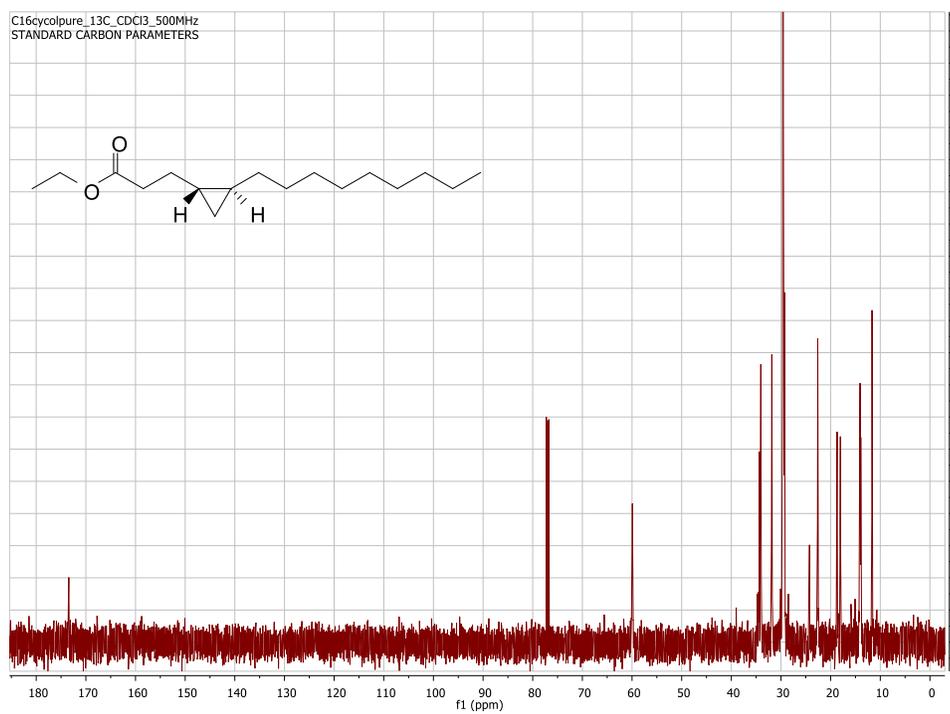
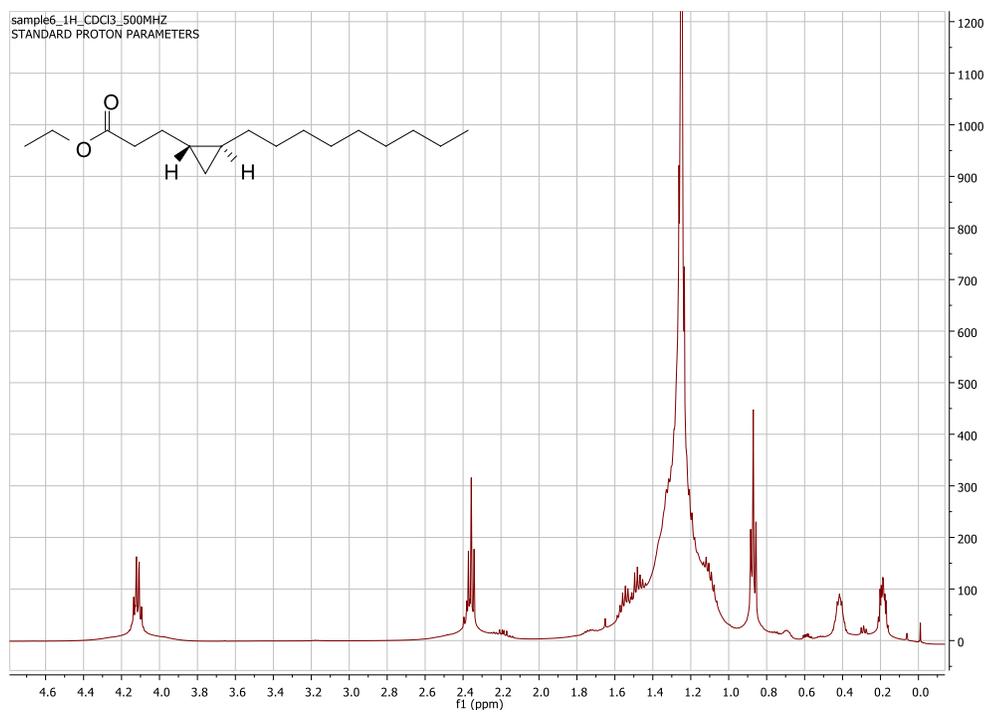


Figure 6.6.46: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 29



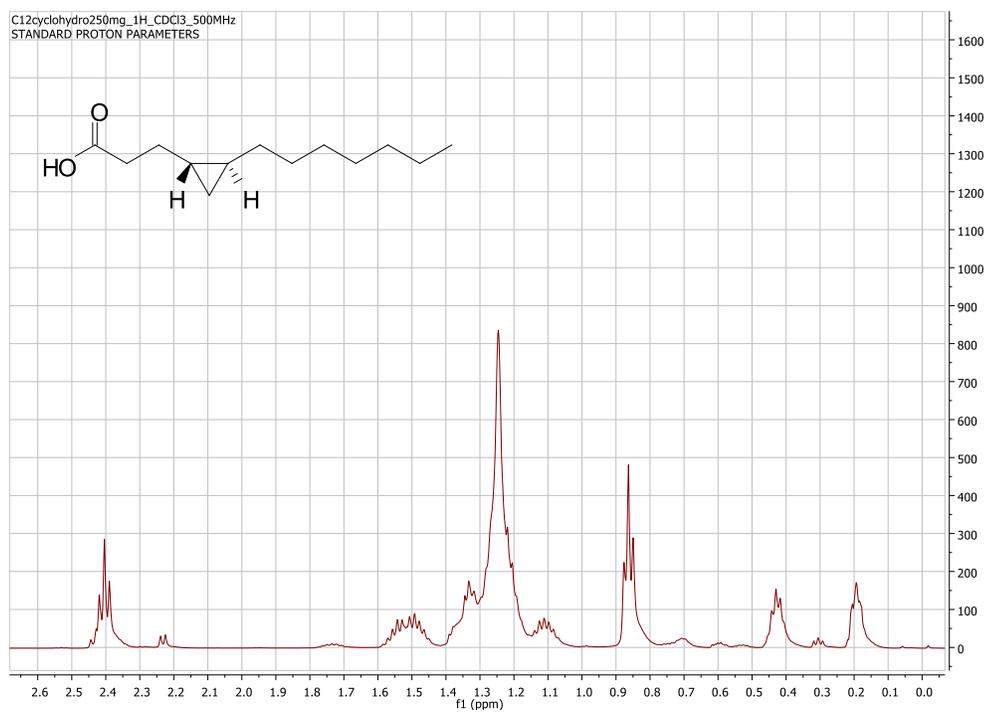


Figure 6.6.49: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 31

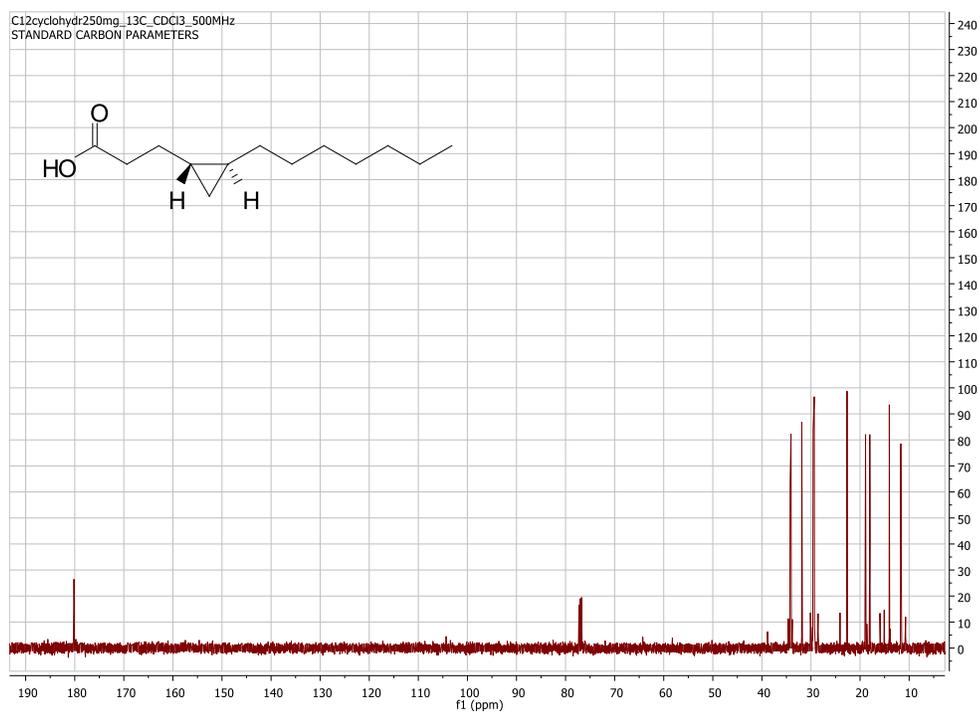
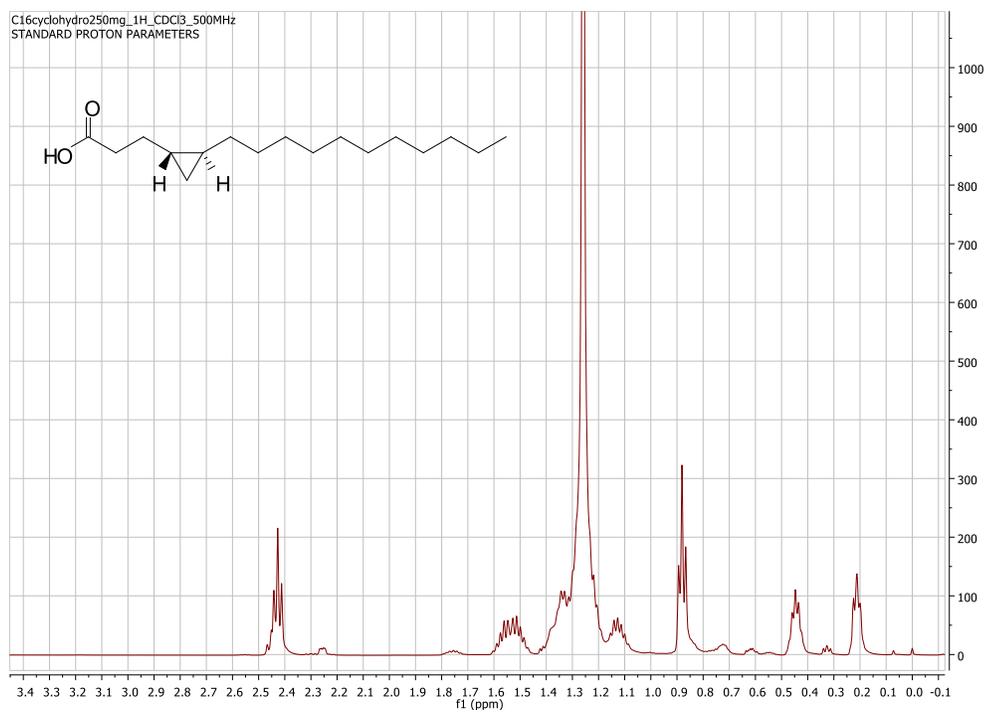
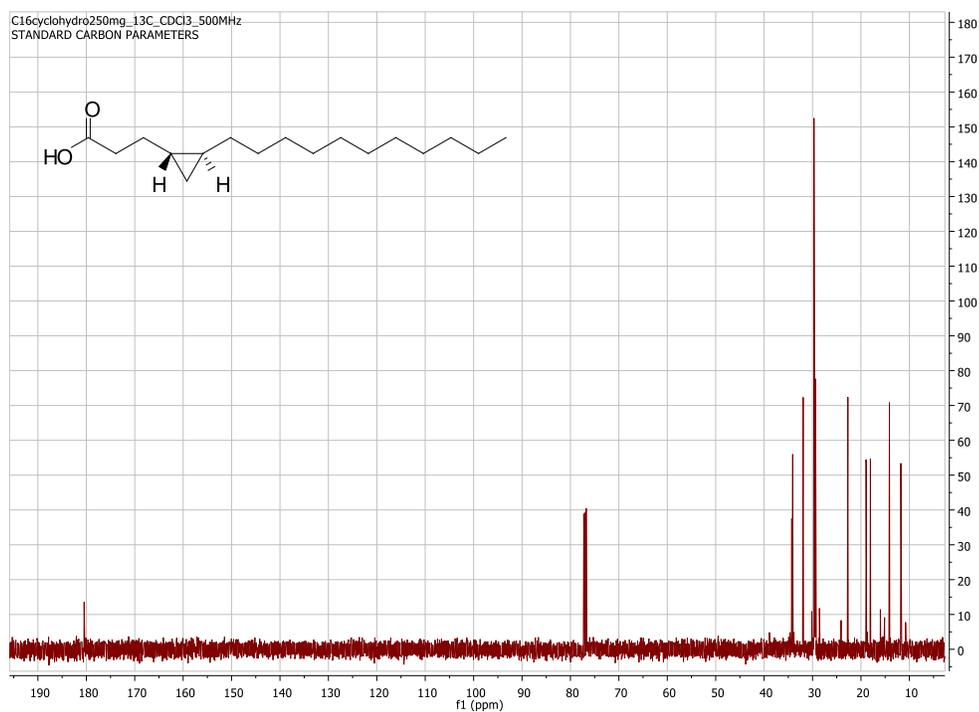


Figure 6.6.50: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 31



**Figure 6.6.51:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **32**



**Figure 6.6.52:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **32**



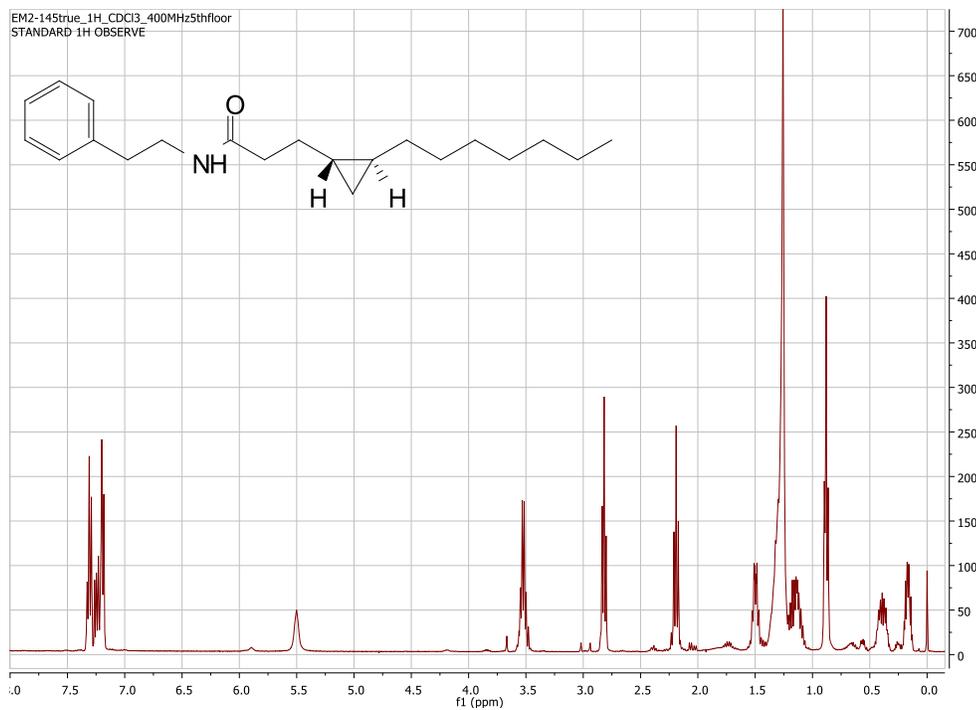


Figure 6.6.55: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 33b

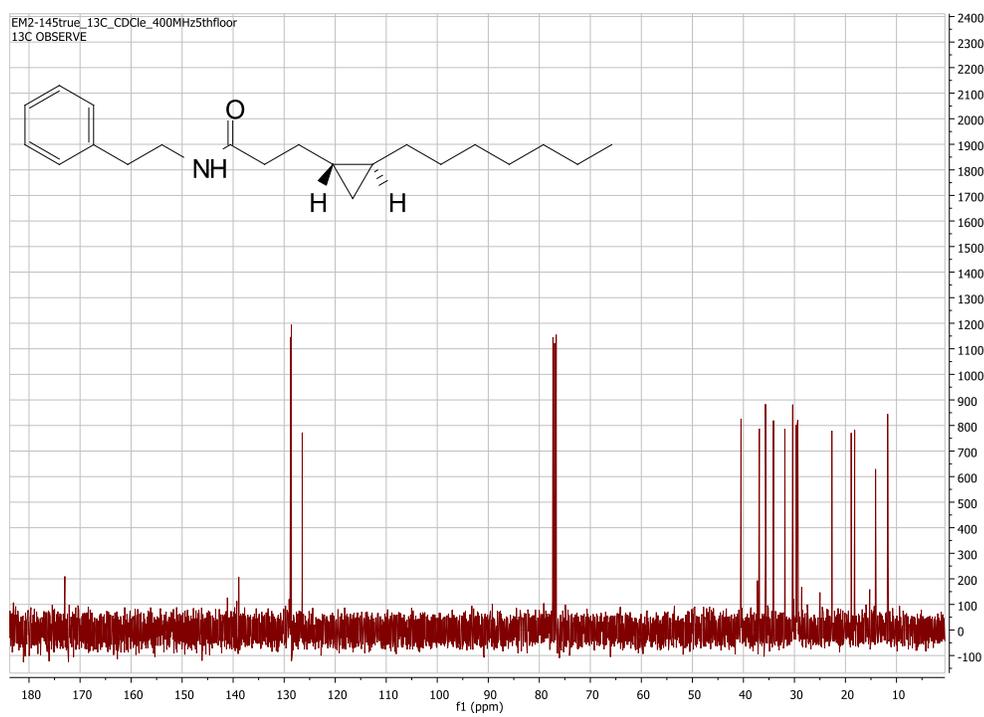


Figure 6.6.56: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 33b

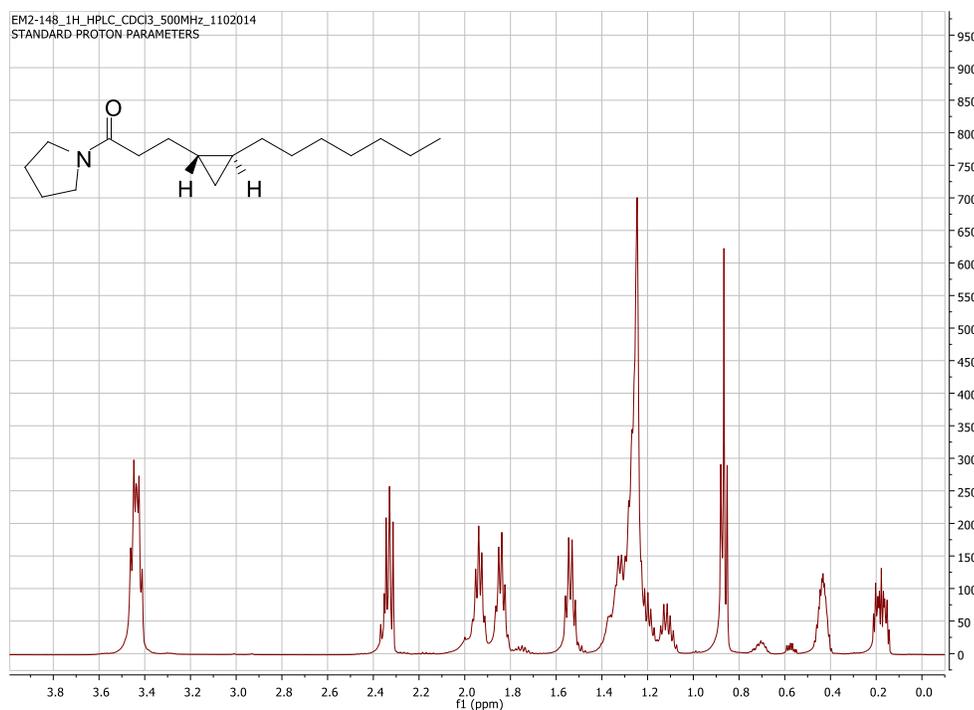


Figure 6.6.57: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 33c

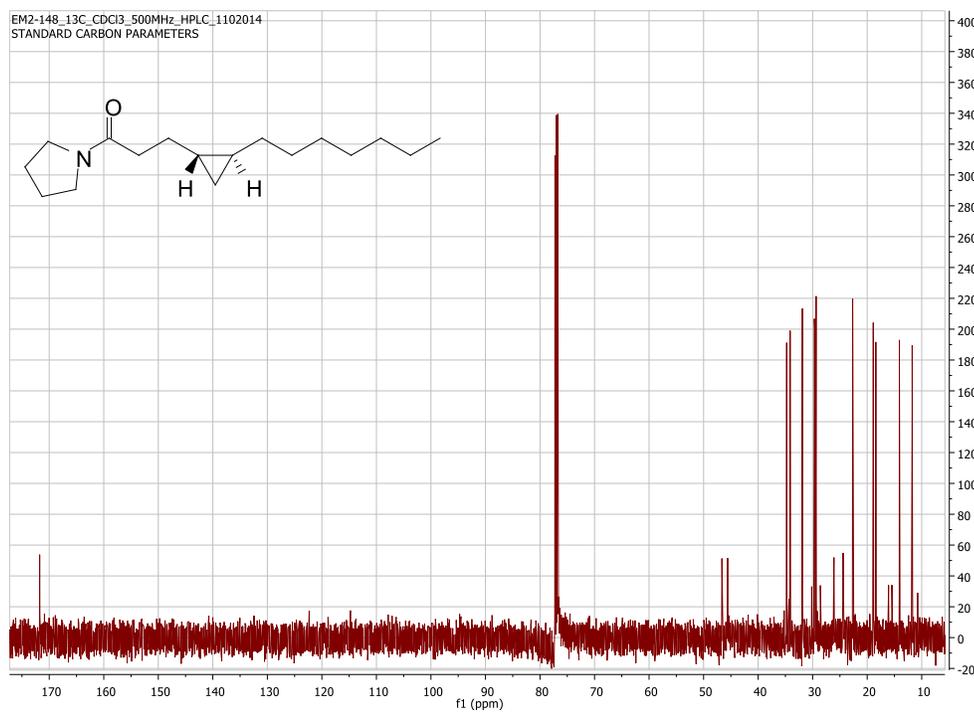
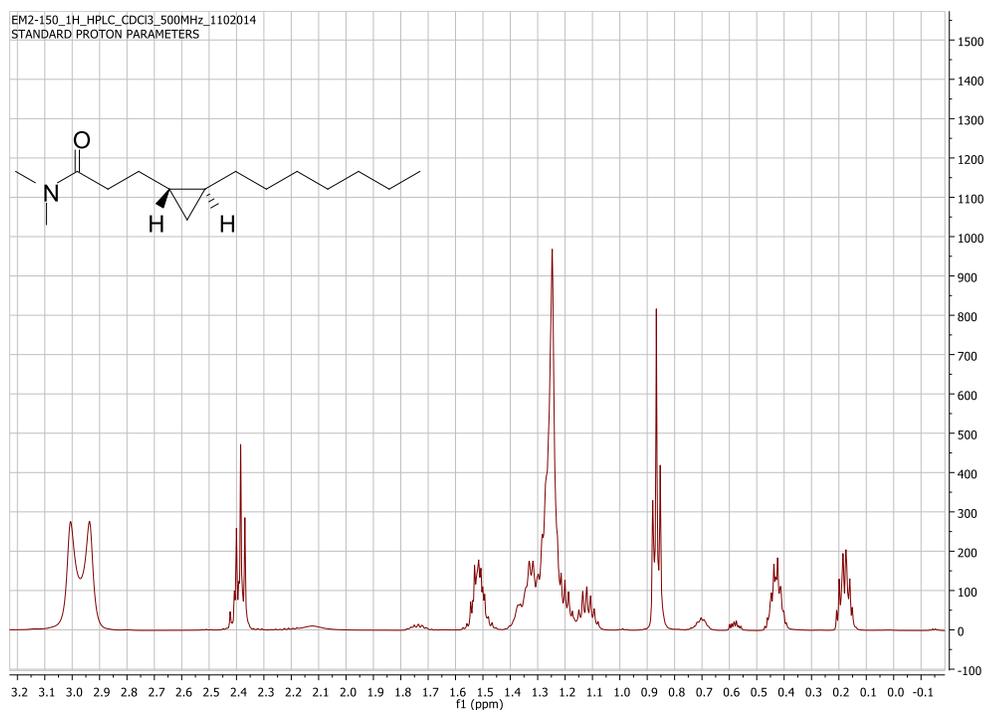
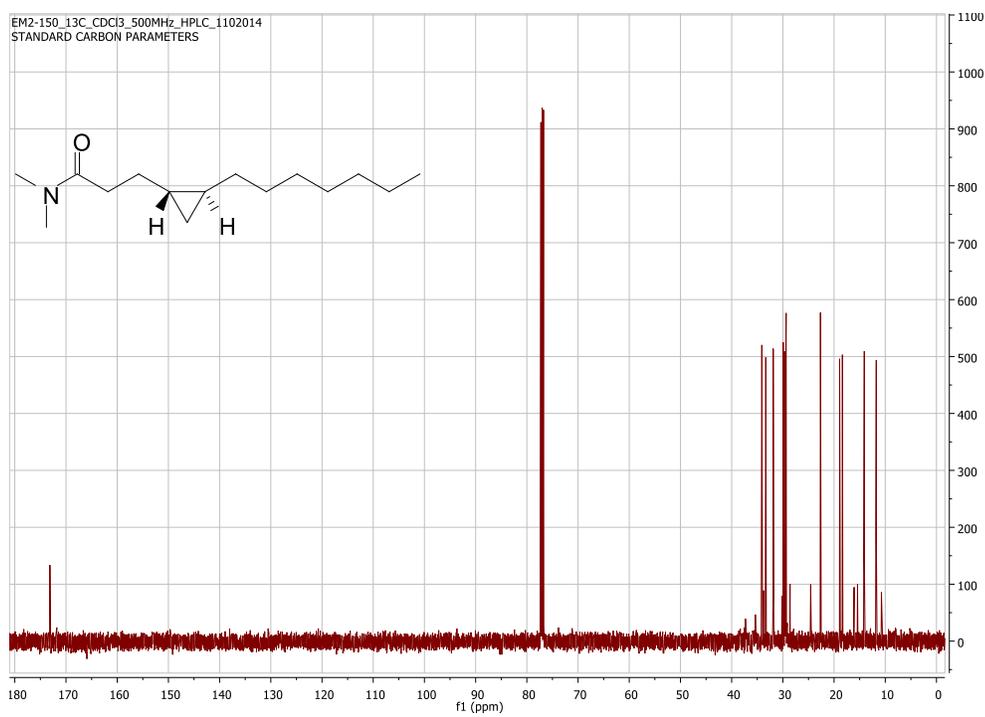


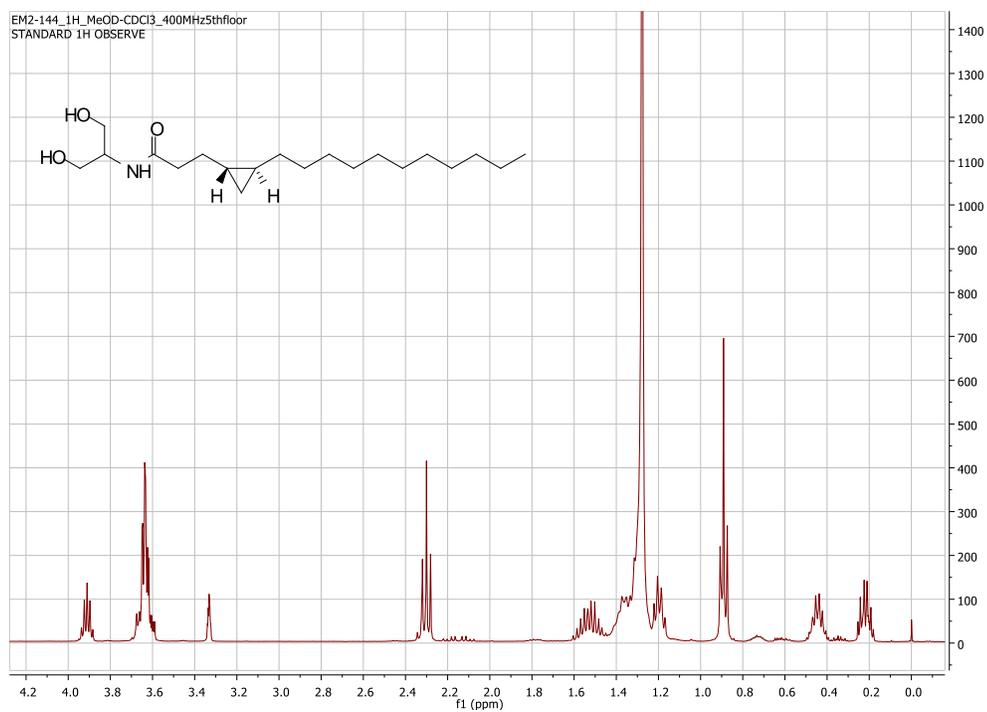
Figure 6.6.58: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 33c



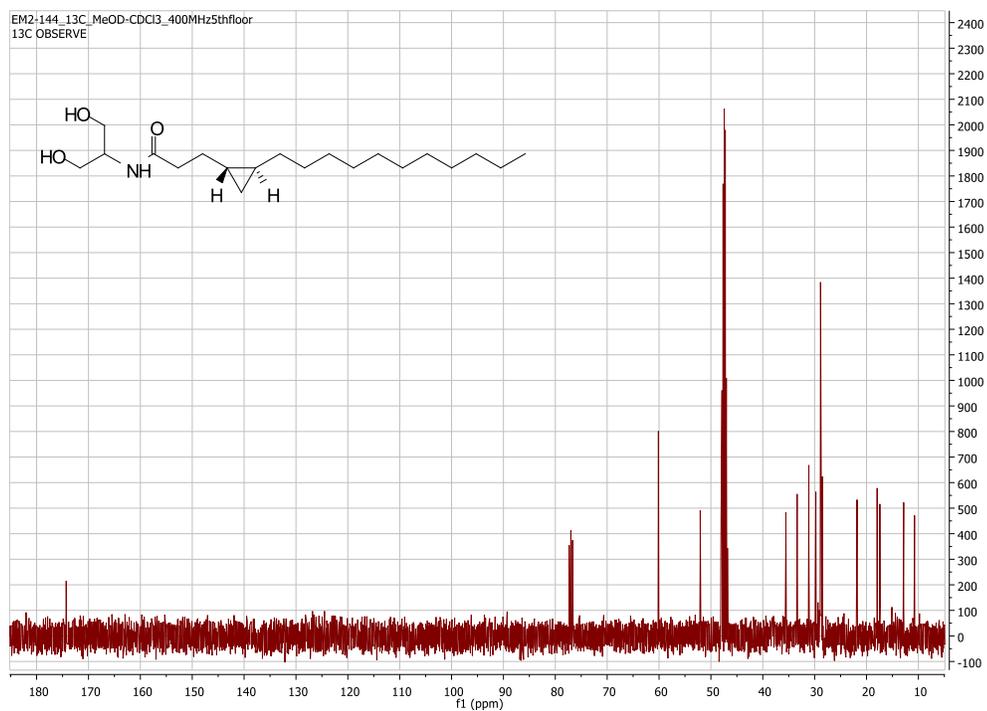
**Figure 6.6.59:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **33d**



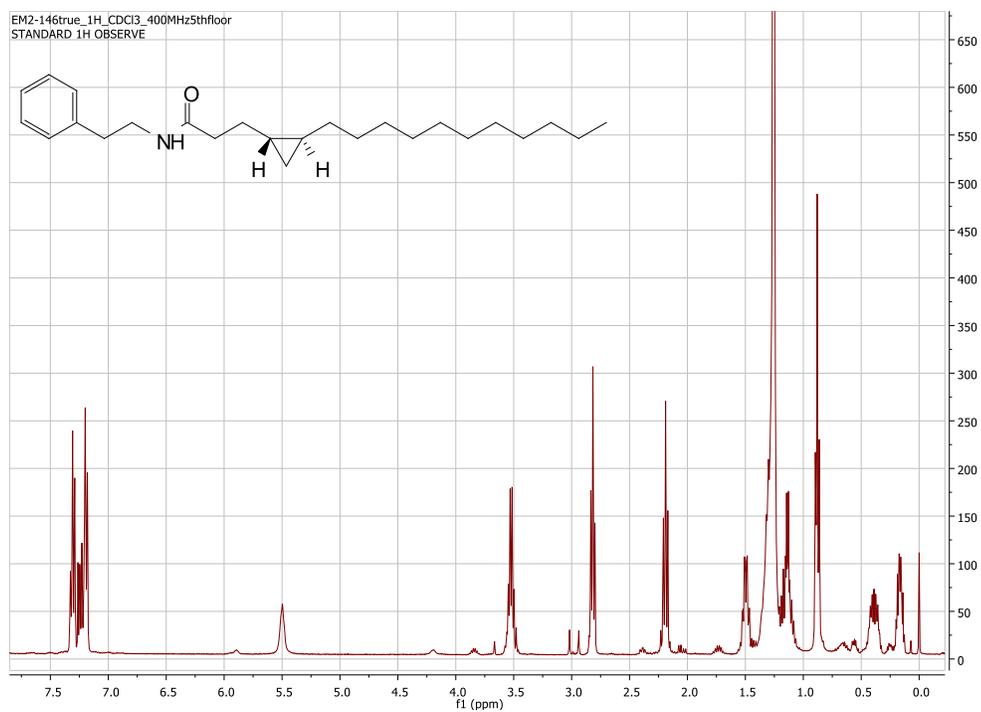
**Figure 6.6.60:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **33d**



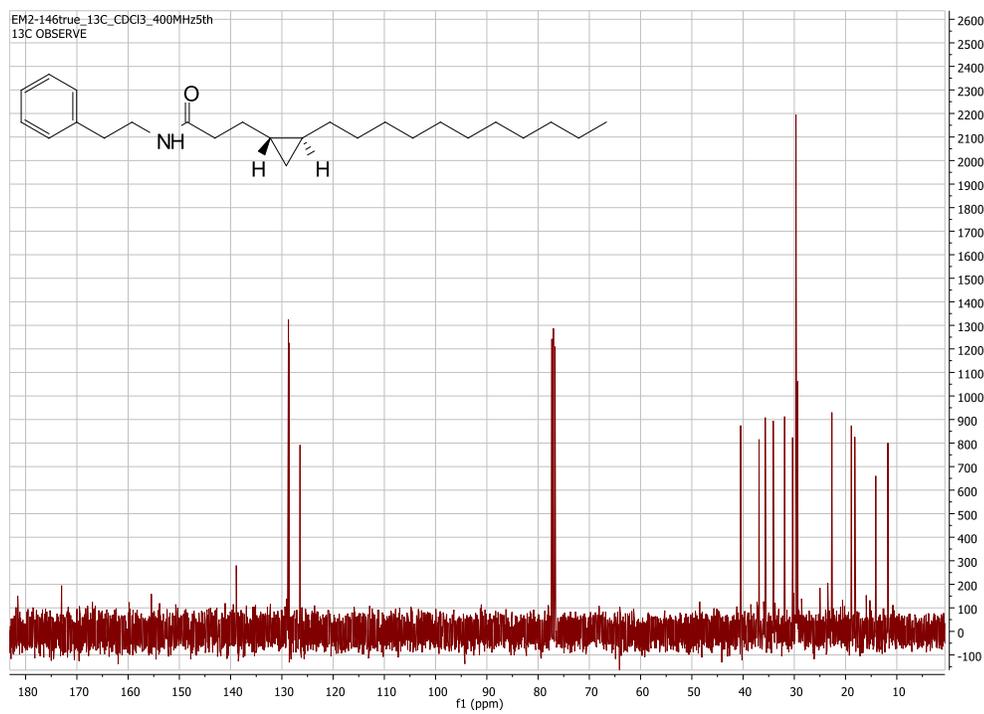
**Figure 6.6.61:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound **34a**



**Figure 6.6.62:** <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound **34a**



**Figure 6.6.63:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound **34b**



**Figure 6.6.64:** <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound **34b**

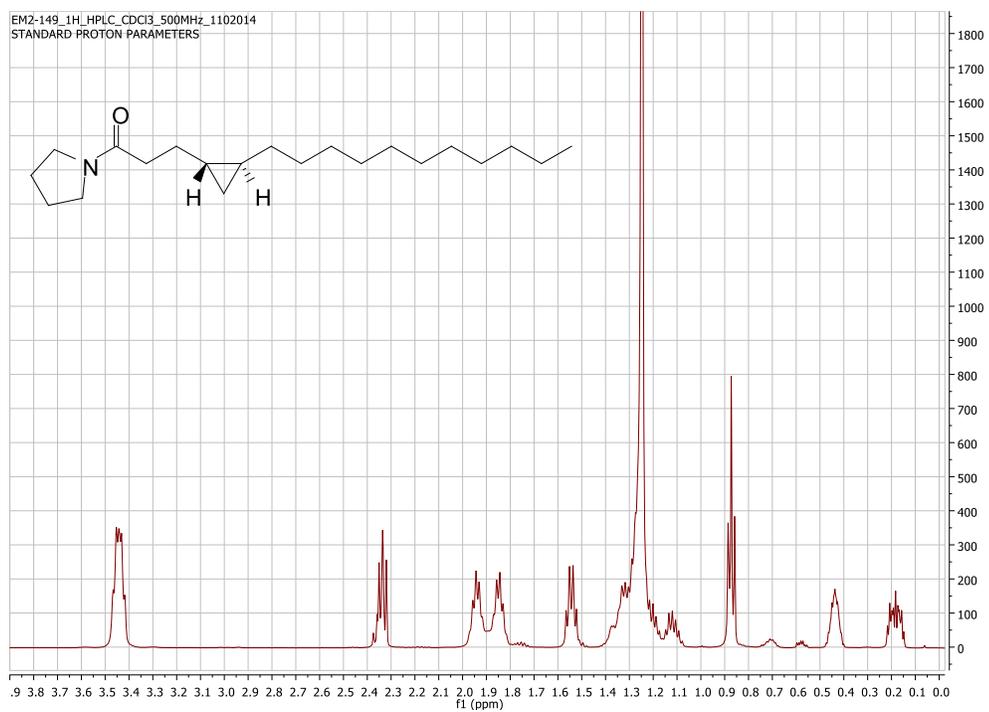


Figure 6.6.65: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 34c

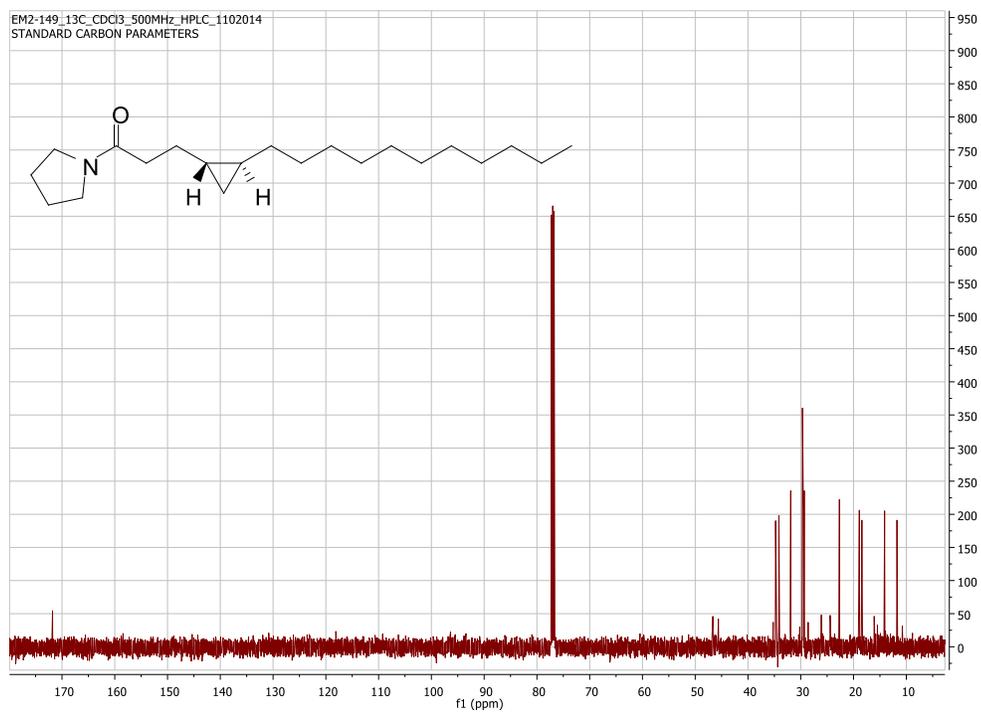
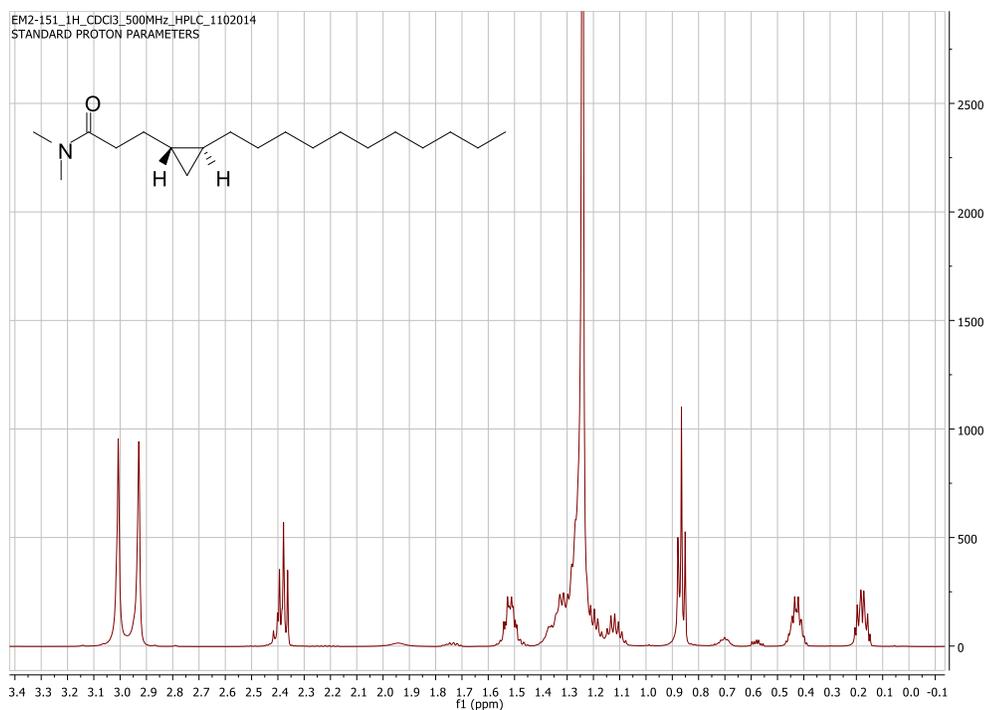
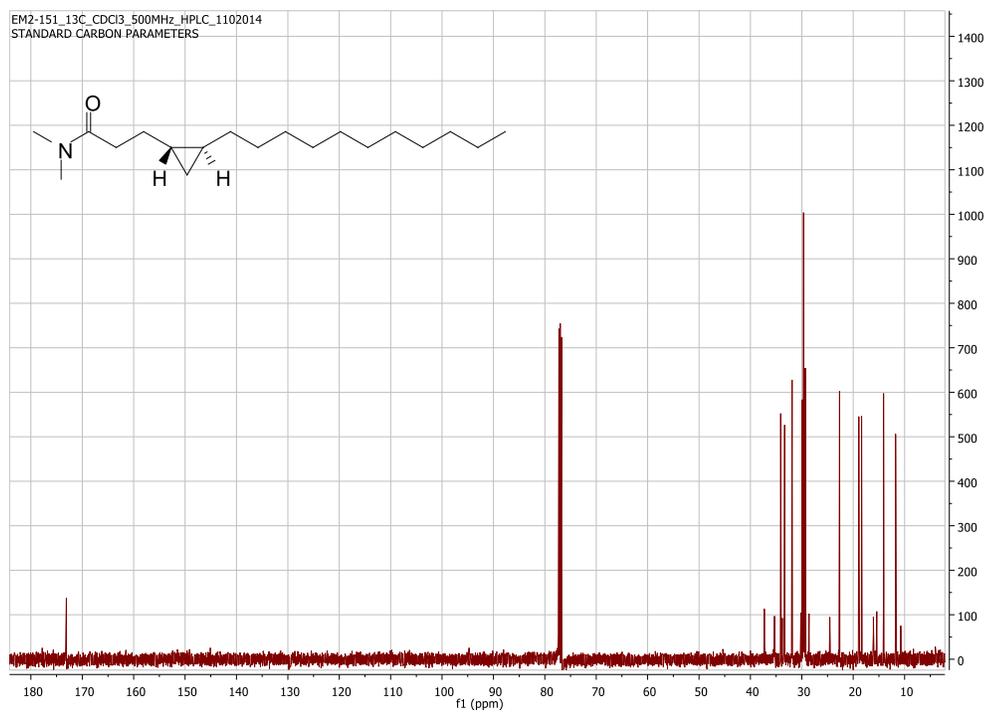


Figure 6.6.66: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 34c



**Figure 6.6.67:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **34d**



**Figure 6.6.68:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **34d**

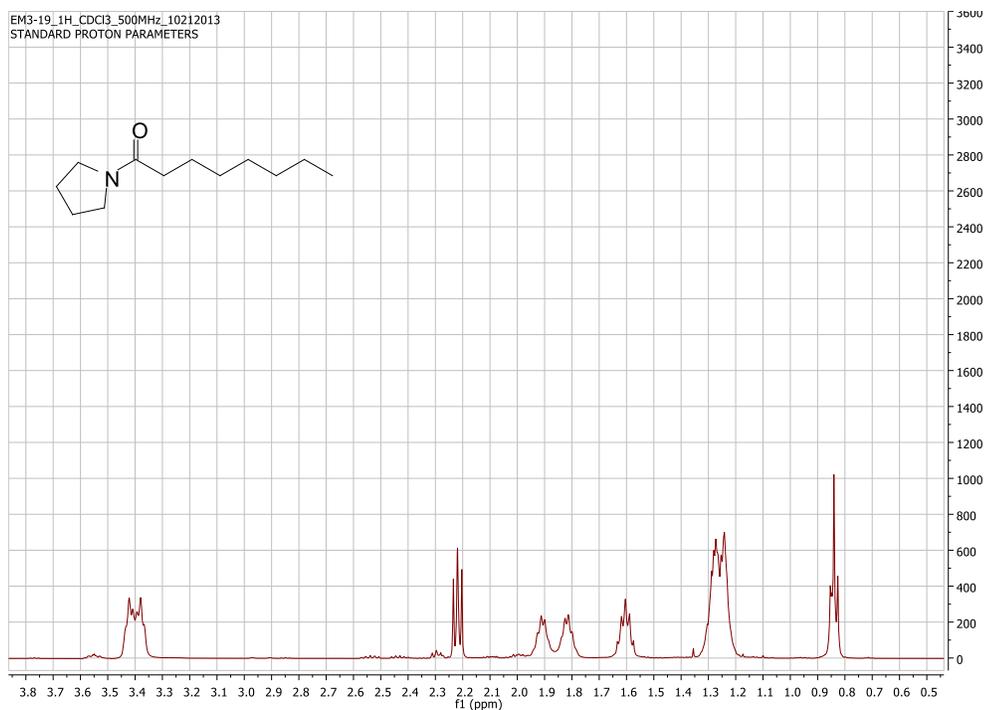


Figure 6.6.69: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 35c

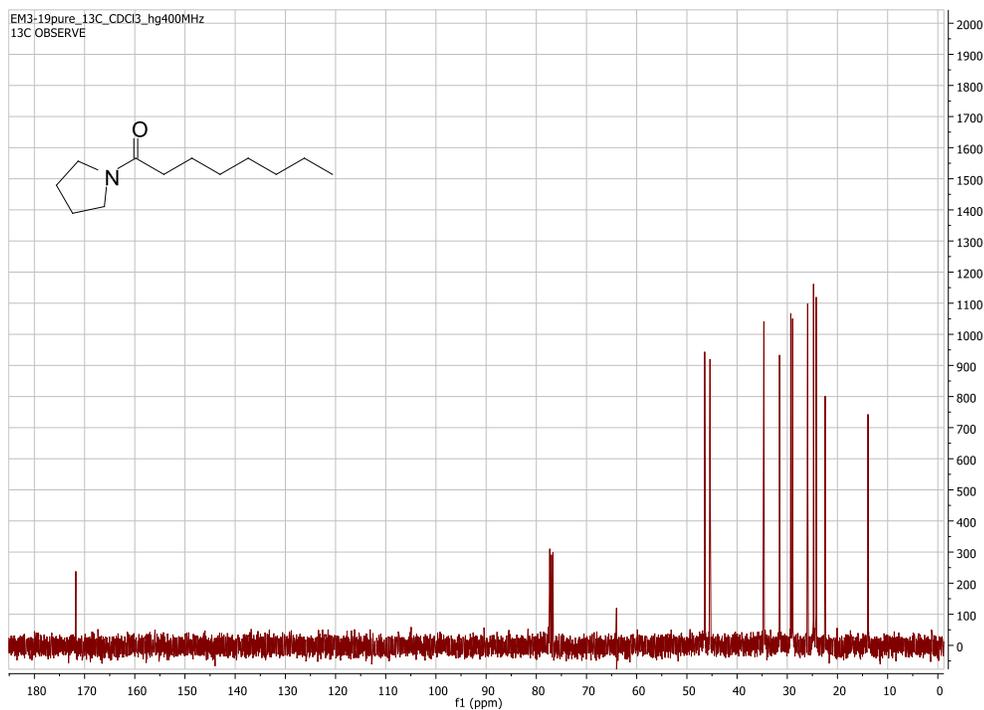


Figure 6.6.70: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 35c

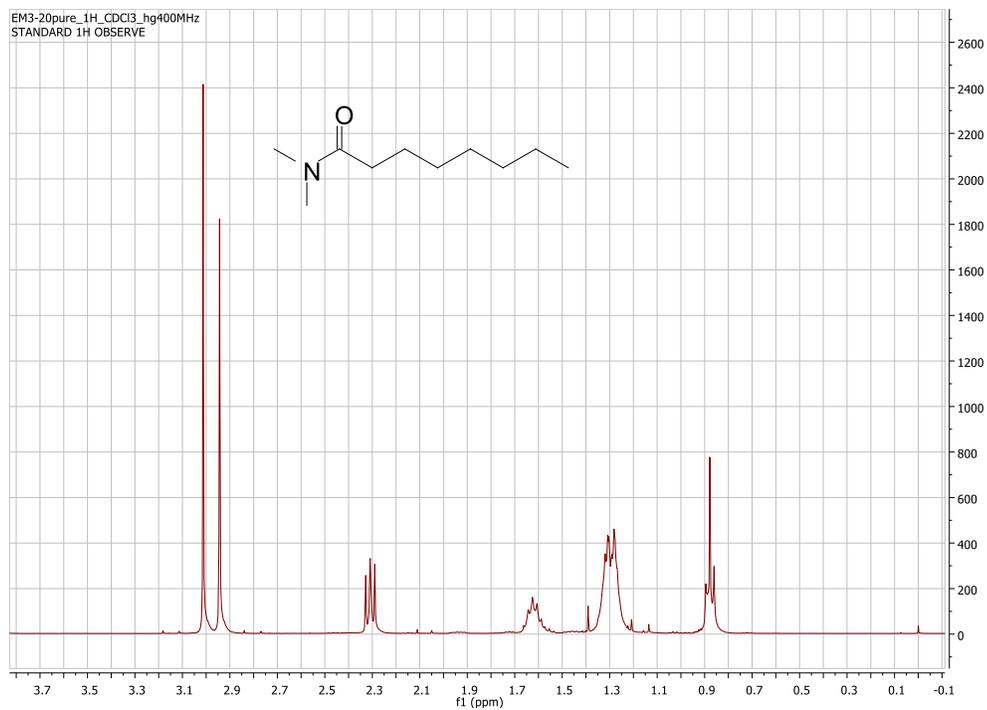


Figure 6.6.71: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 35d

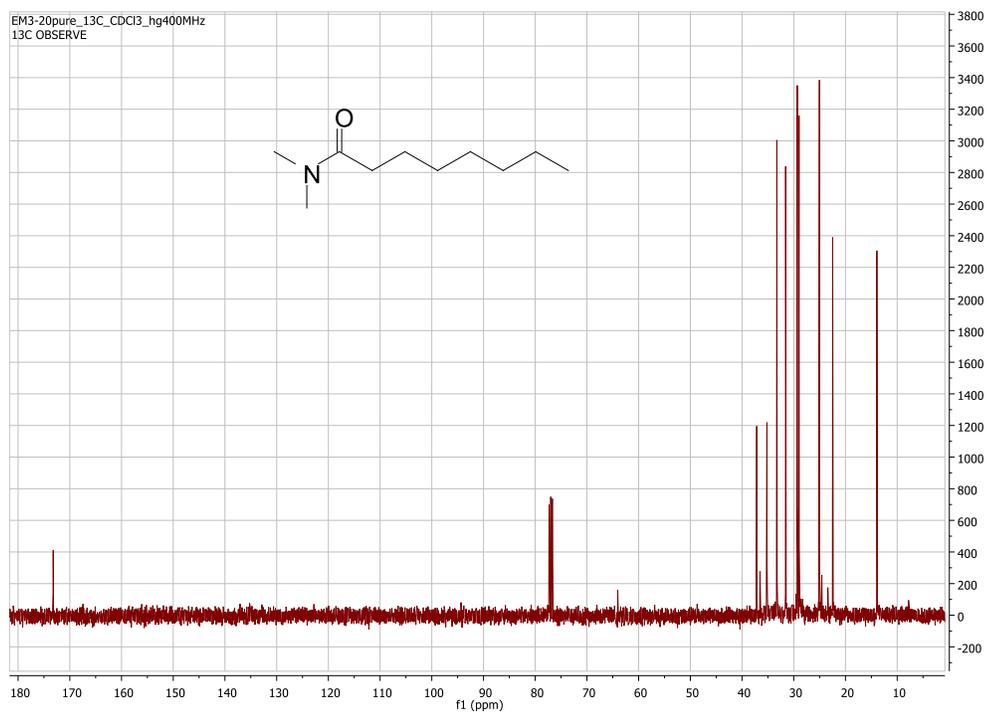


Figure 6.6.72: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 35d

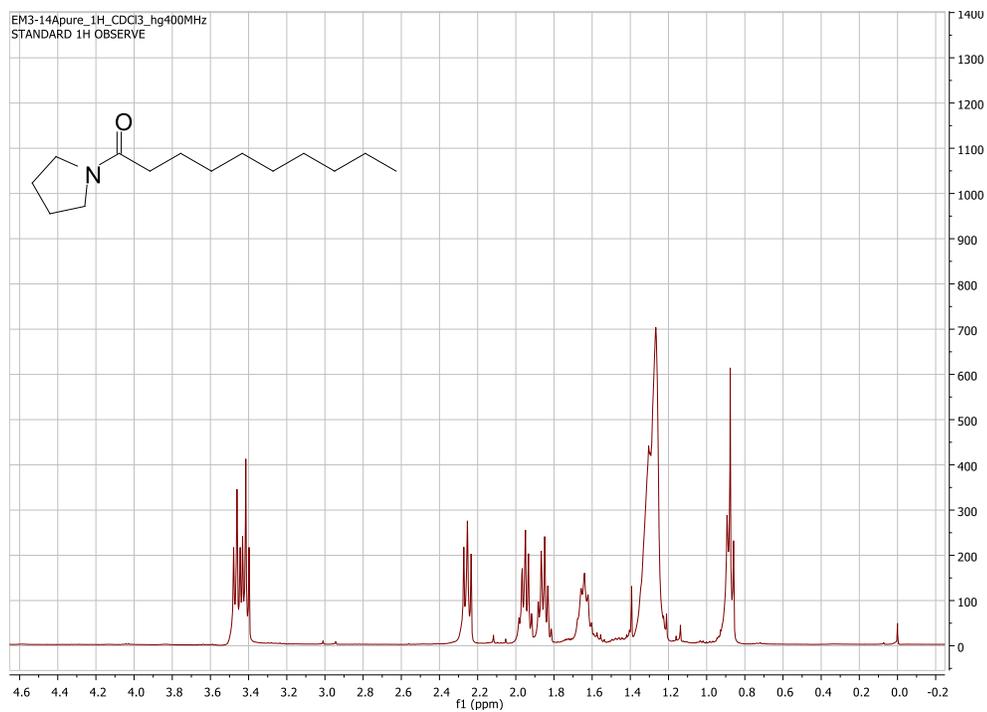


Figure 6.6.73:  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ) spectrum of compound **36c**

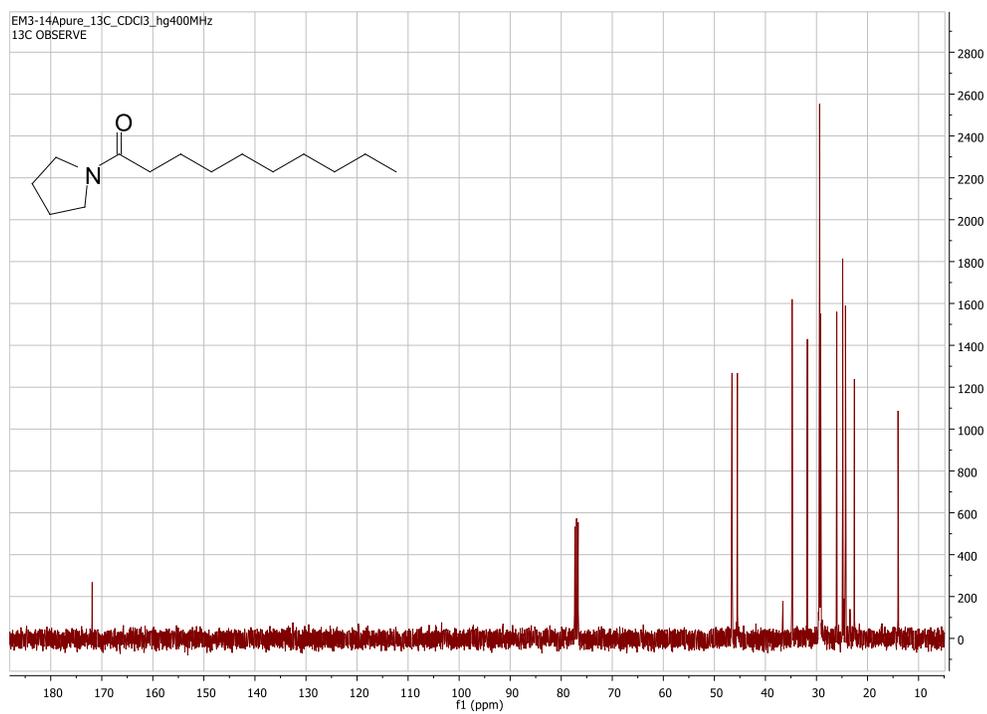


Figure 6.6.74:  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ) spectrum of compound **36c**

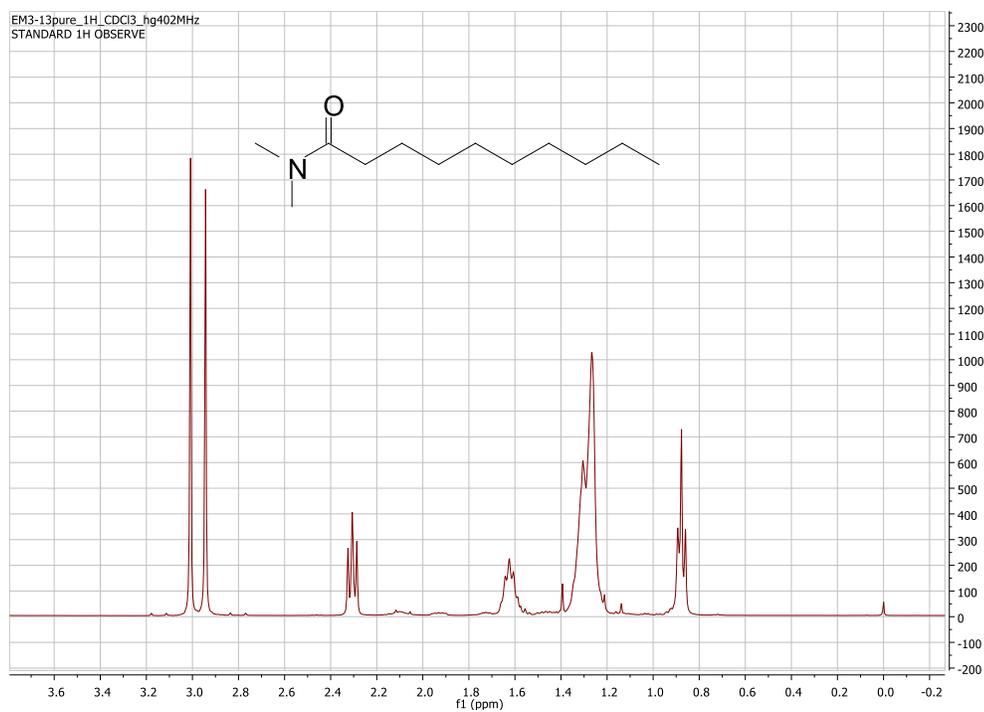


Figure 6.6.75: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 36d

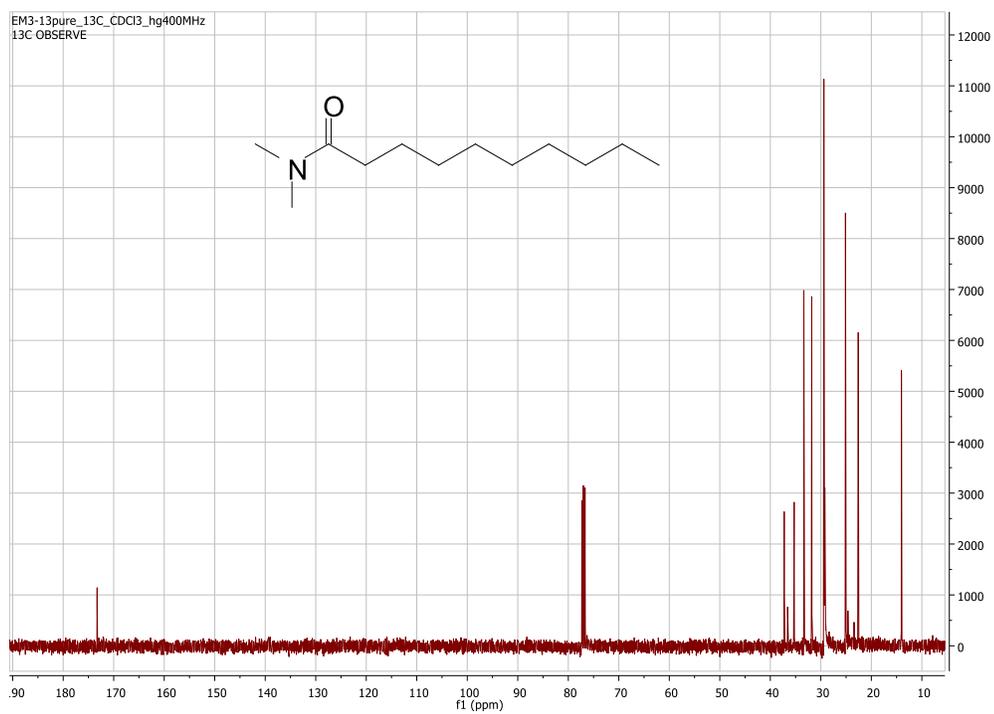


Figure 6.6.76: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 36d

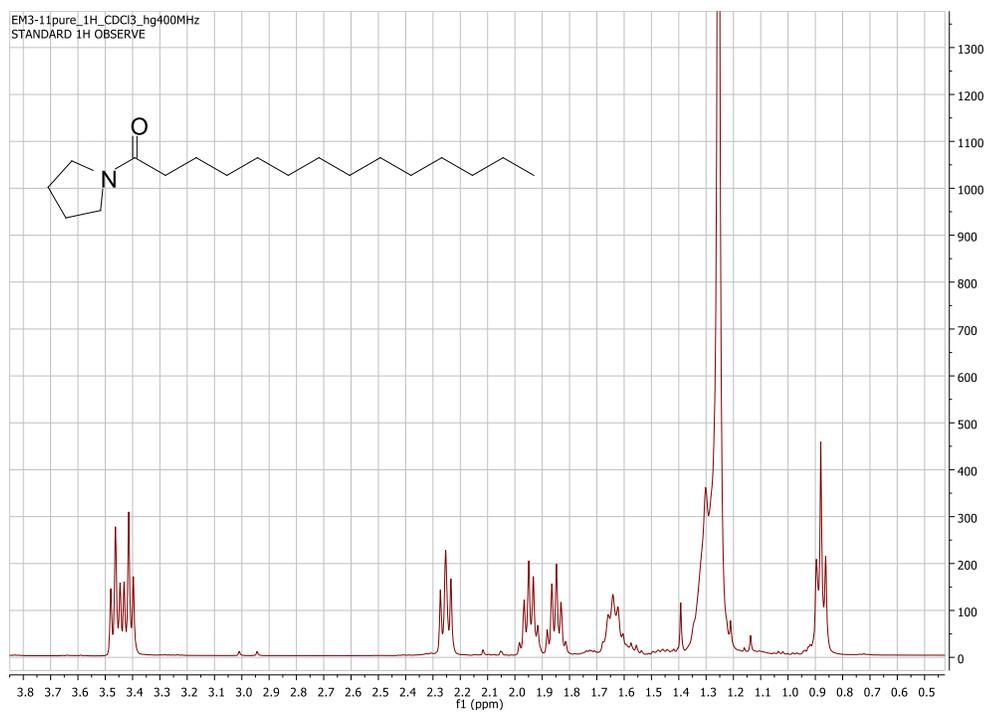


Figure 6.6.77: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 37c

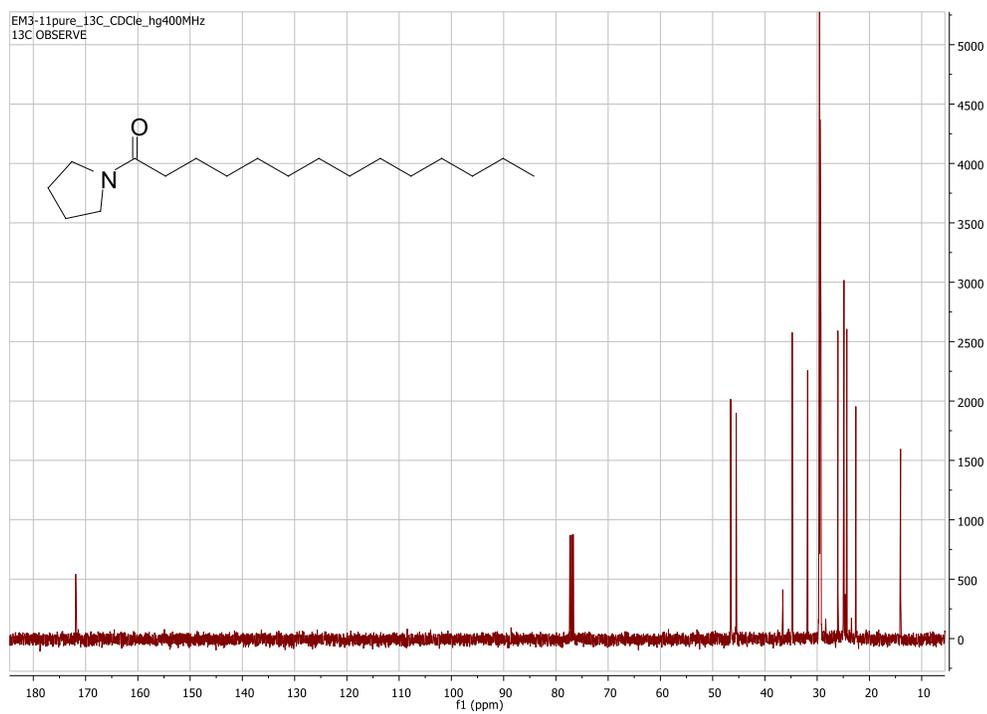
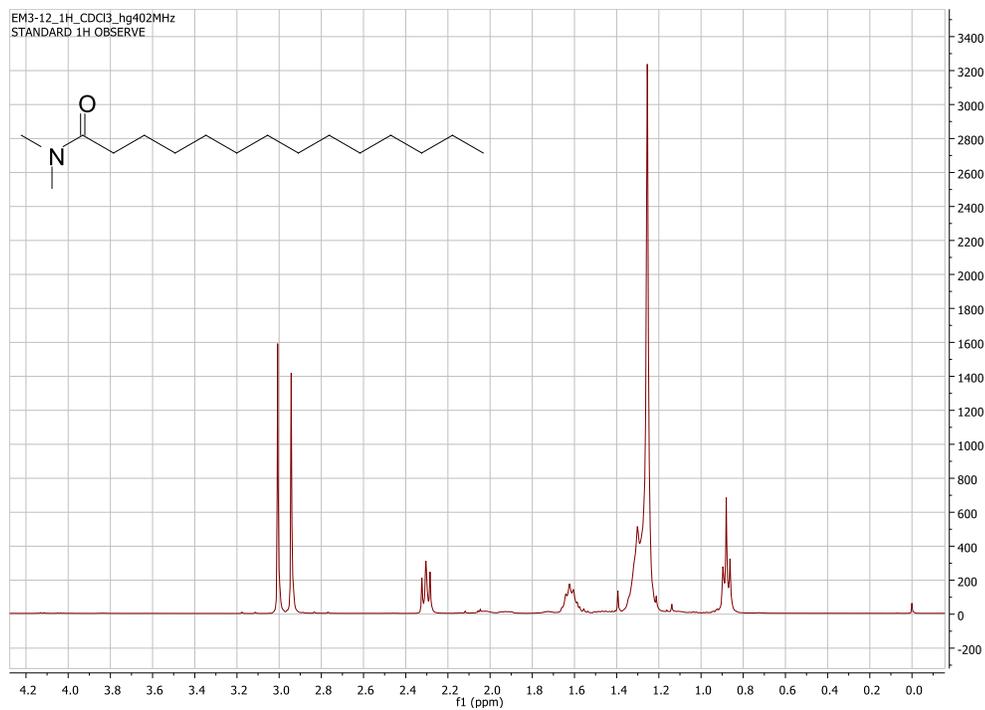
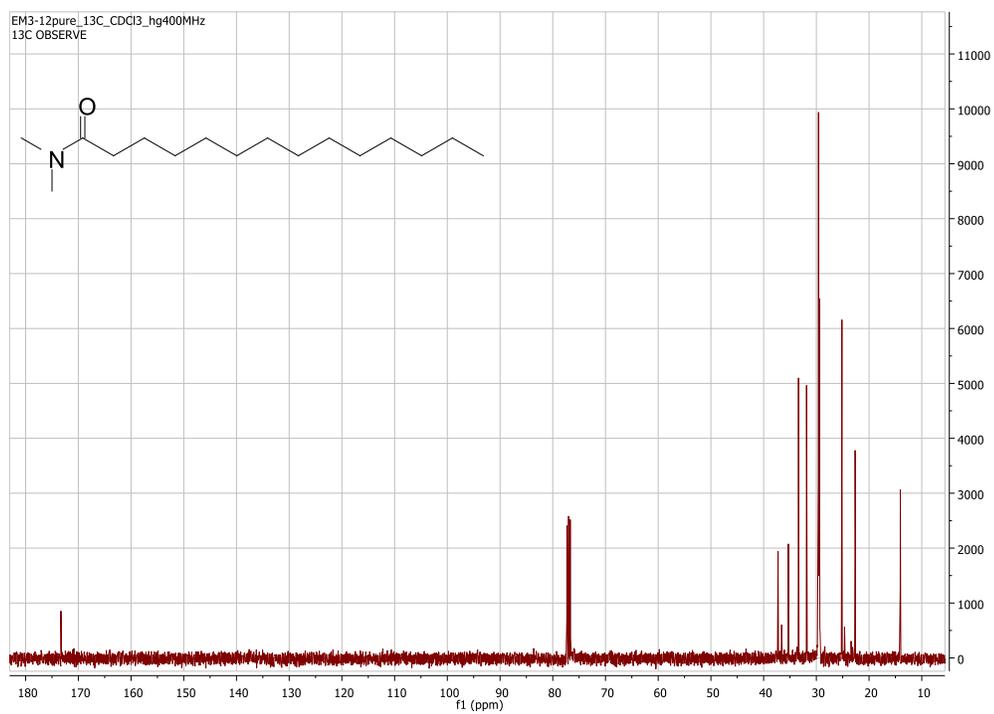


Figure 6.6.78: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 37c



**Figure 6.6.79:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound **37d**



**Figure 6.6.80:** <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound **37d**

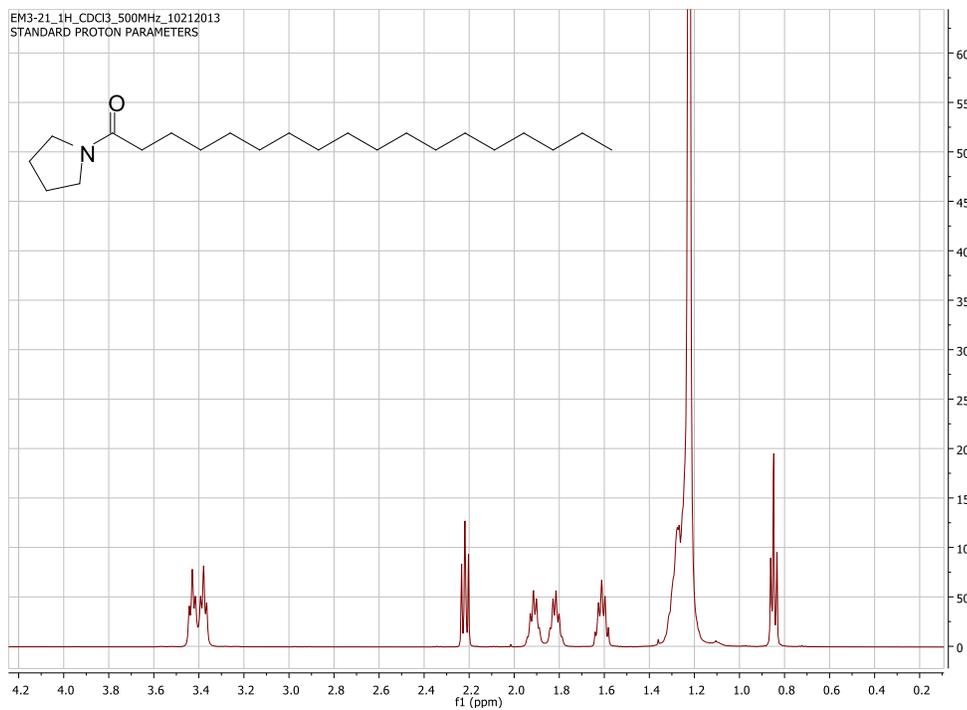


Figure 6.6.81: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 38c

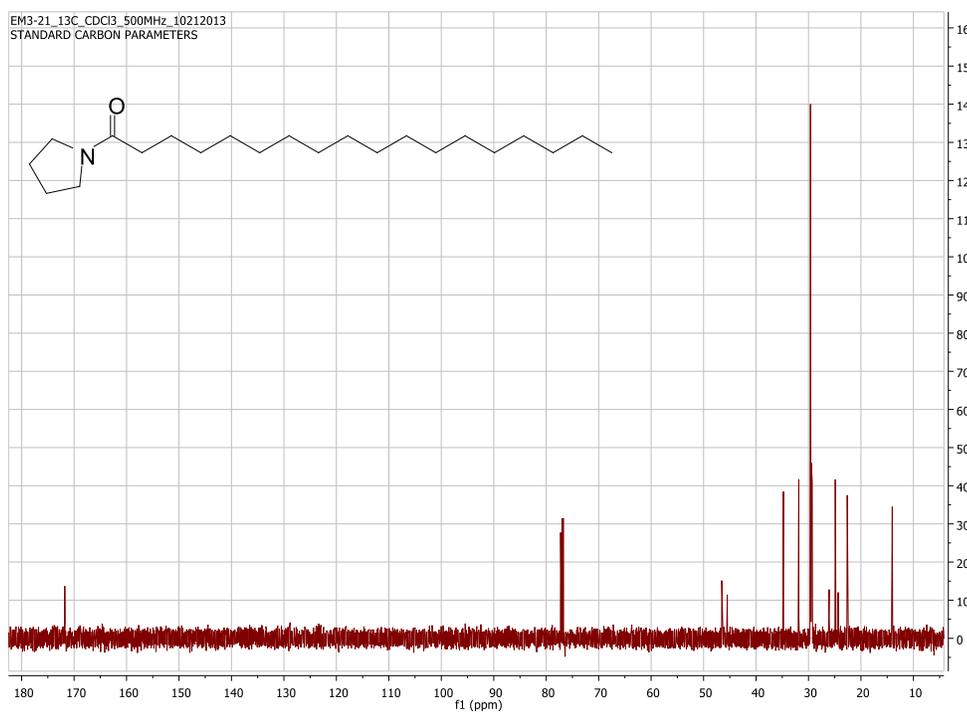
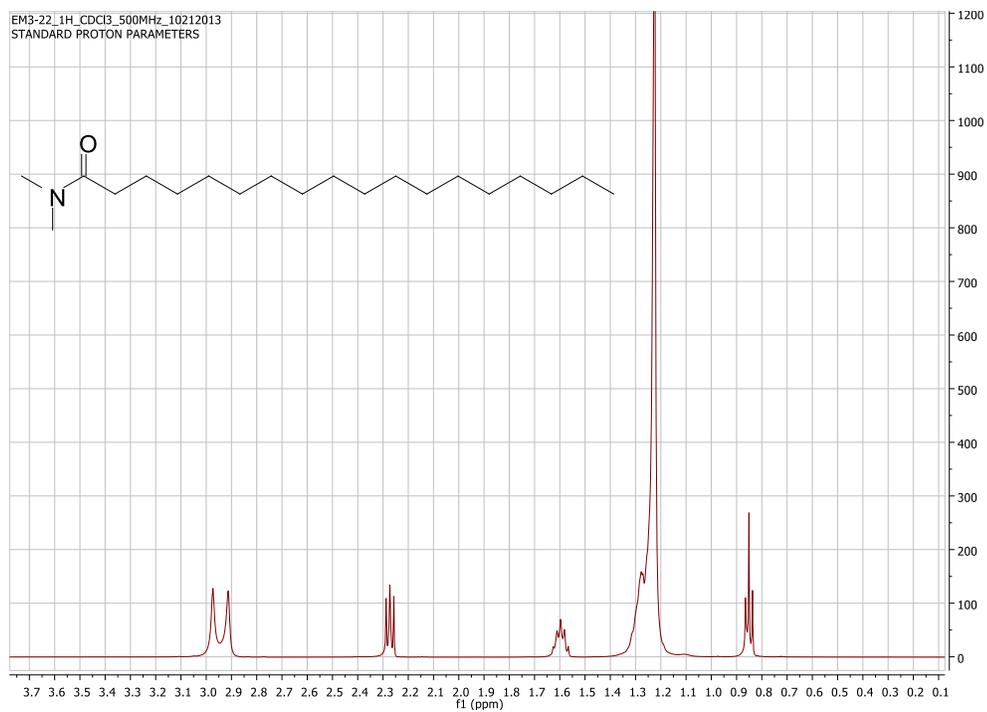
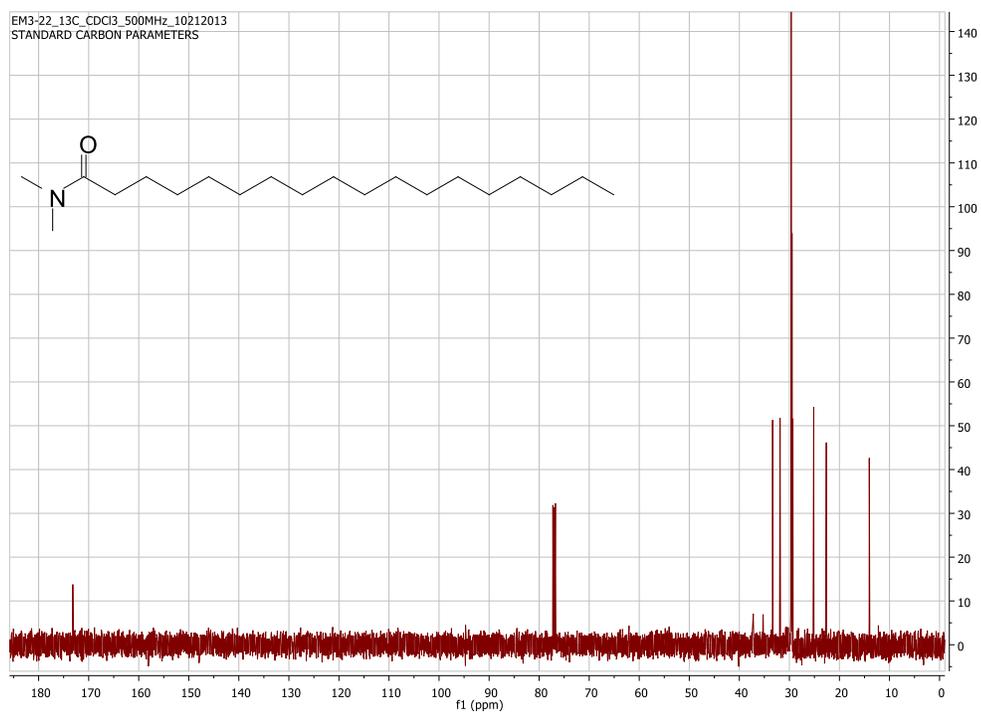


Figure 6.6.82: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 38c



**Figure 6.6.83:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **38d**



**Figure 6.6.84:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **38d**

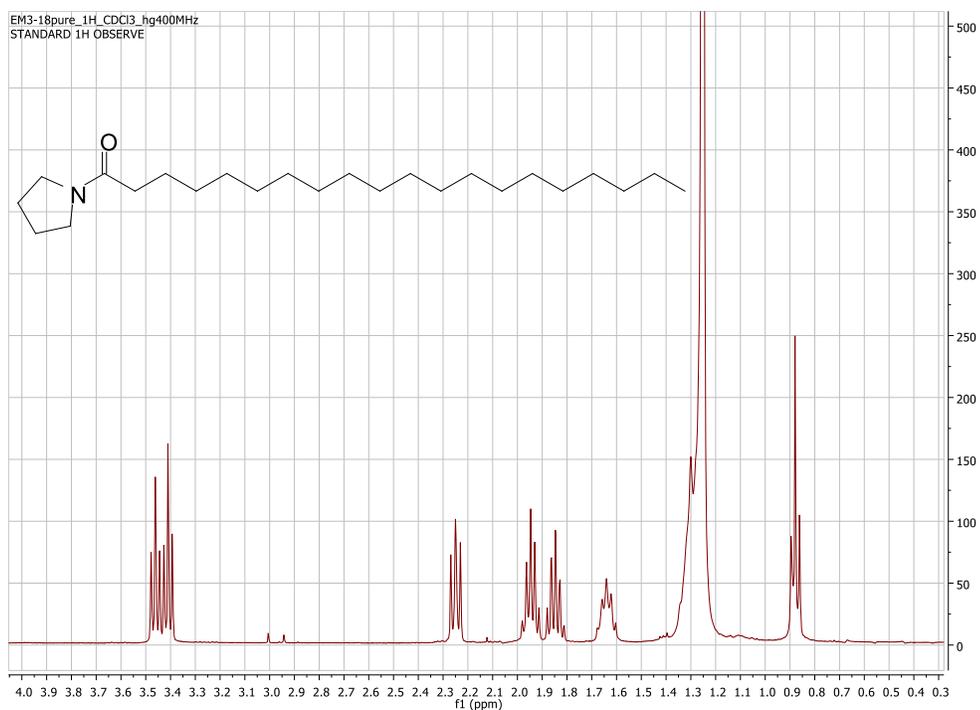


Figure 6.6.85: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) spectrum of compound 39c

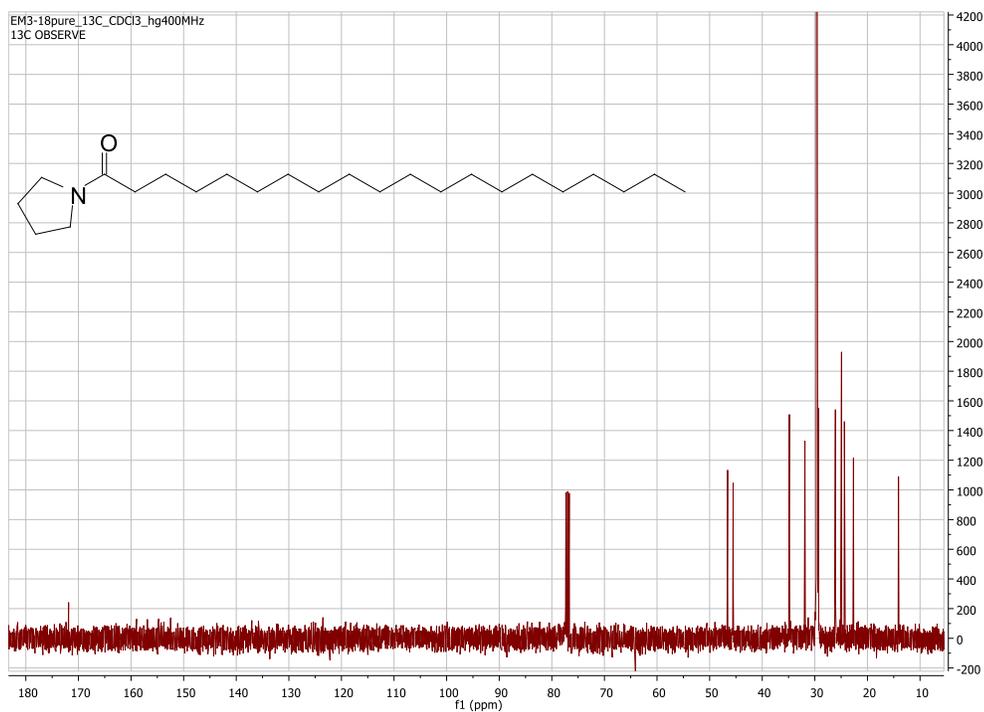
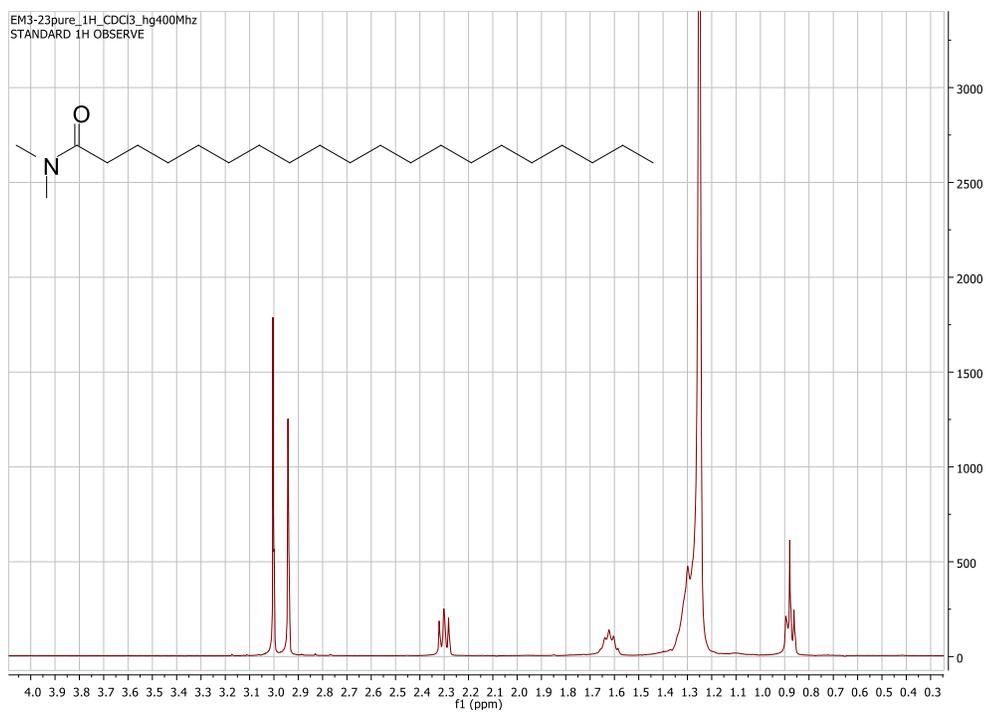
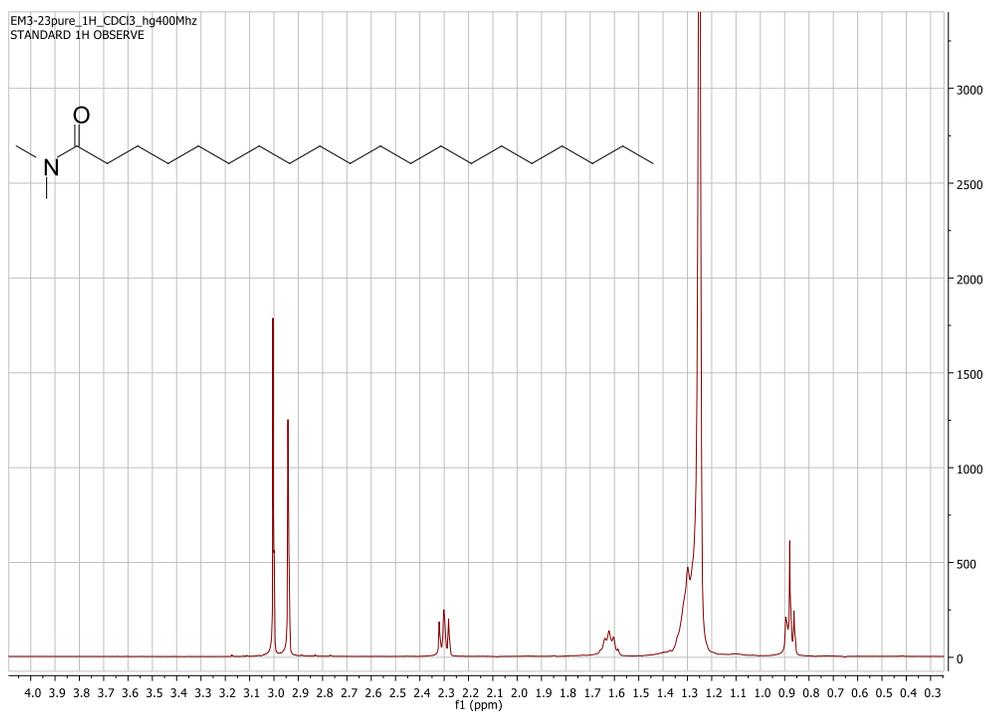


Figure 6.6.86: <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) spectrum of compound 39c



**Figure 6.6.87:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **39d**



**Figure 6.6.88:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **39d**

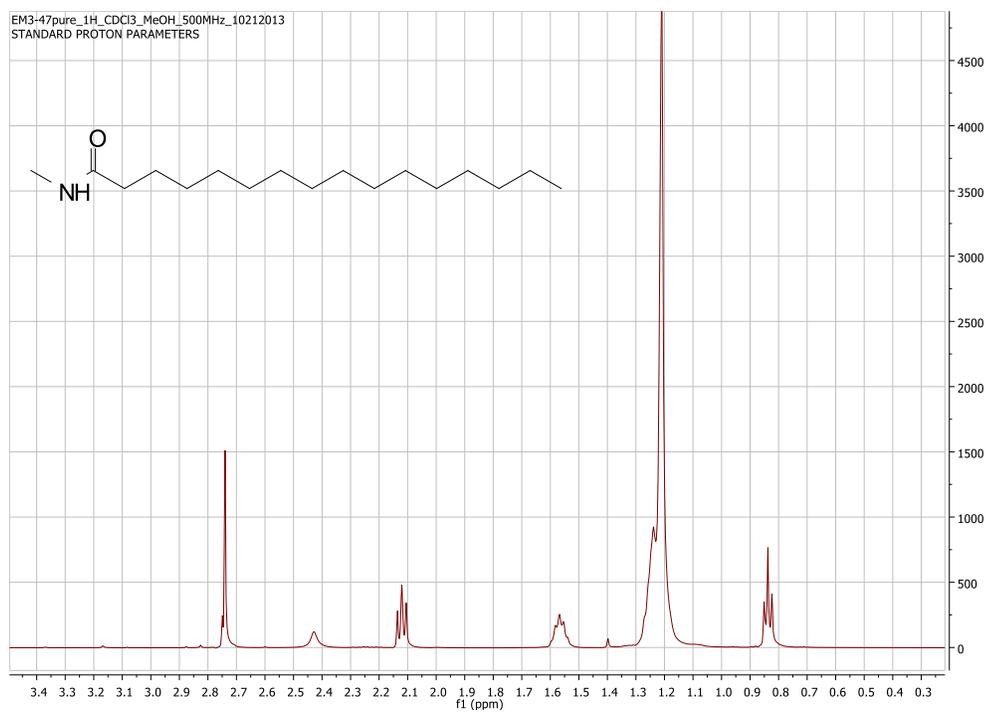


Figure 6.6.89: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 41

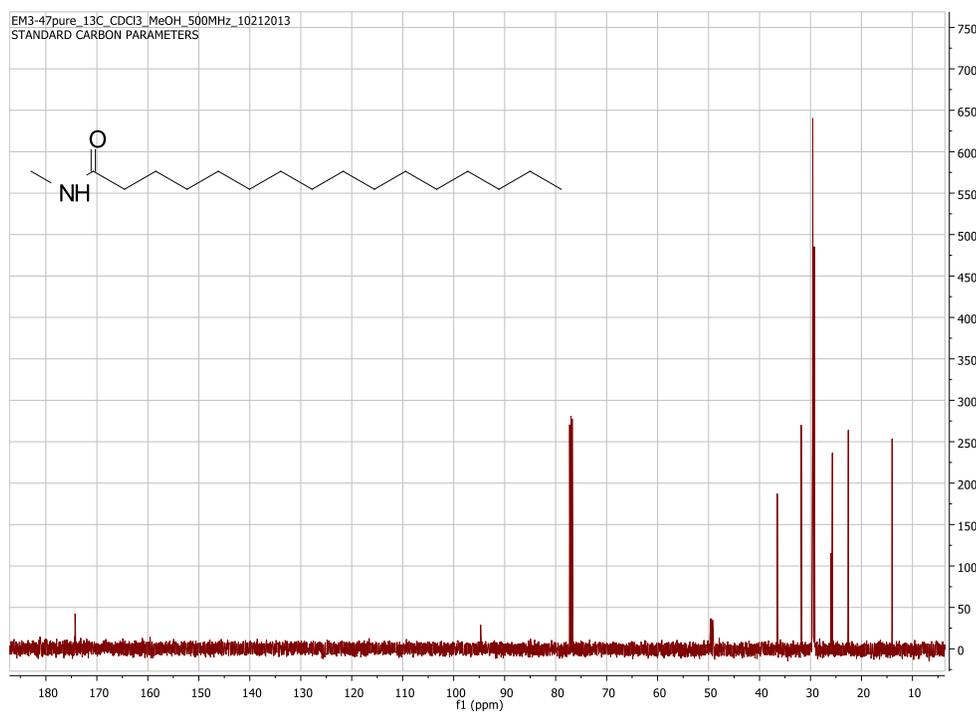


Figure 6.6.90: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 41

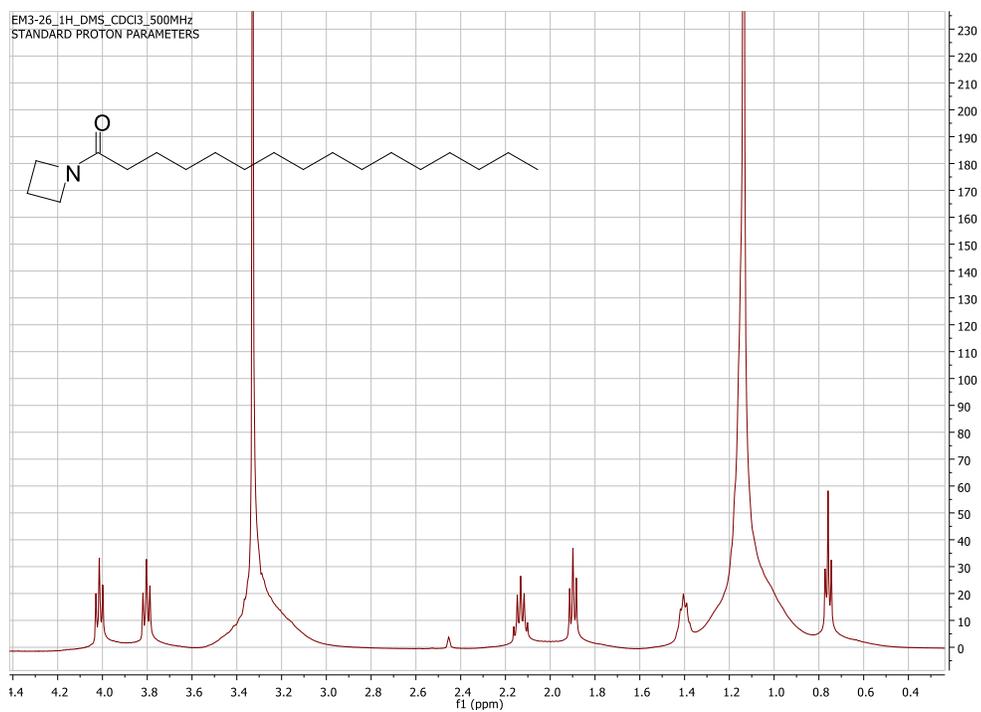


Figure 6.6.91: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound 42

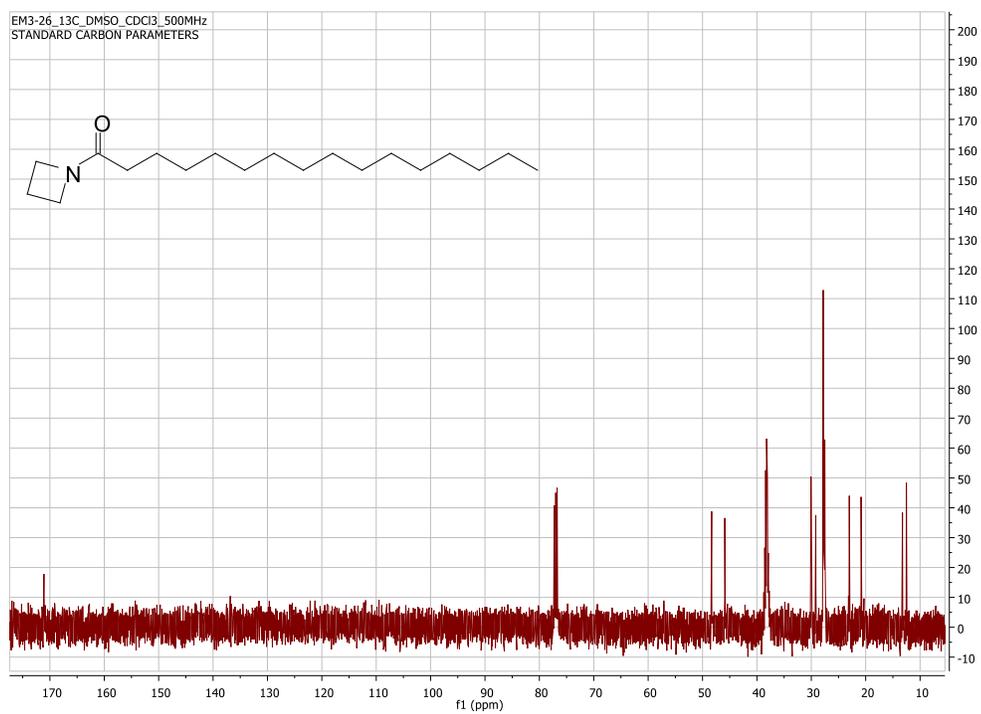
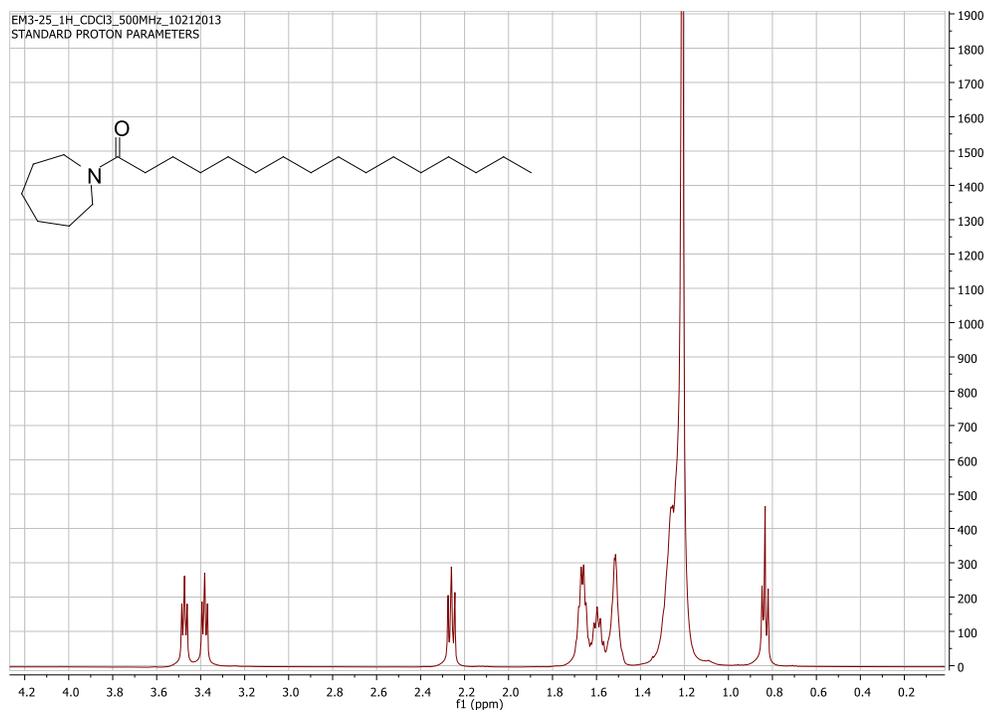
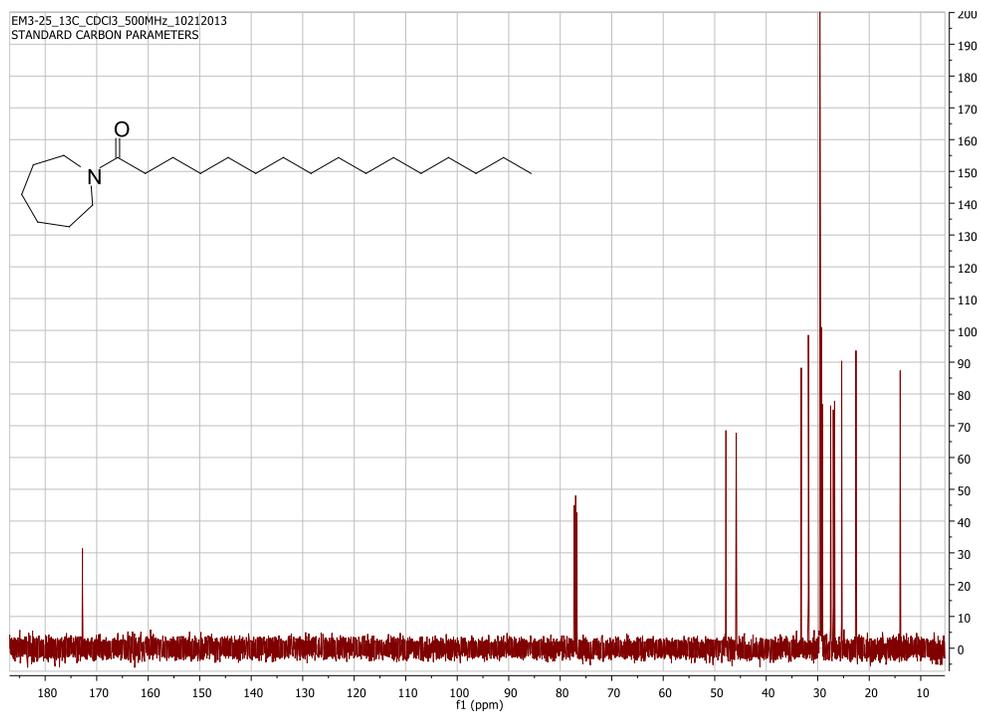


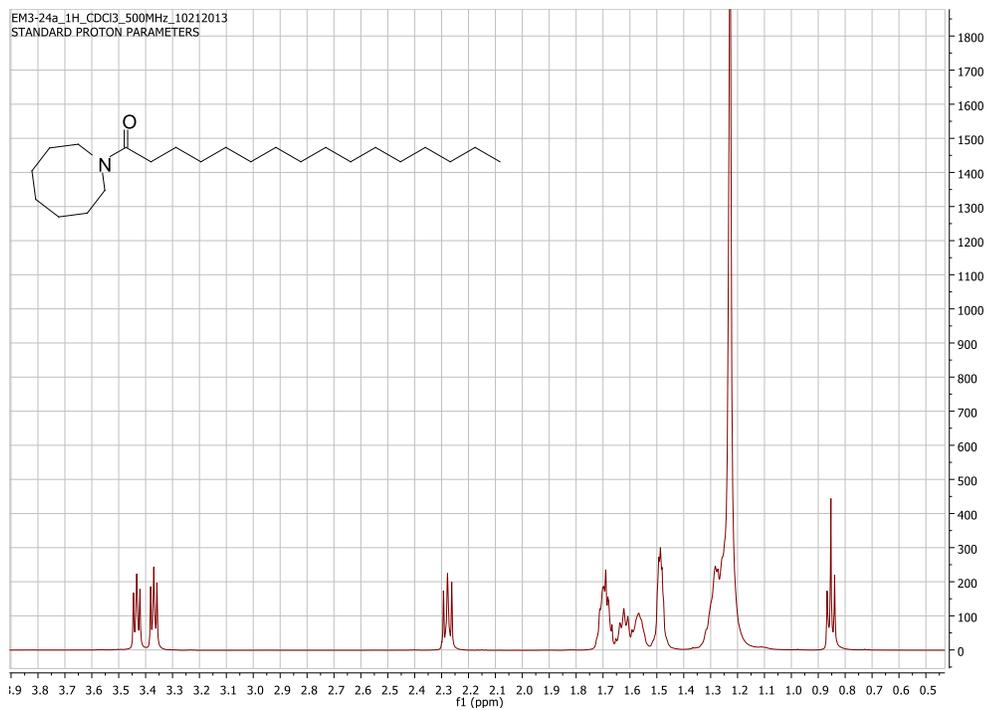
Figure 6.6.92: <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound 42



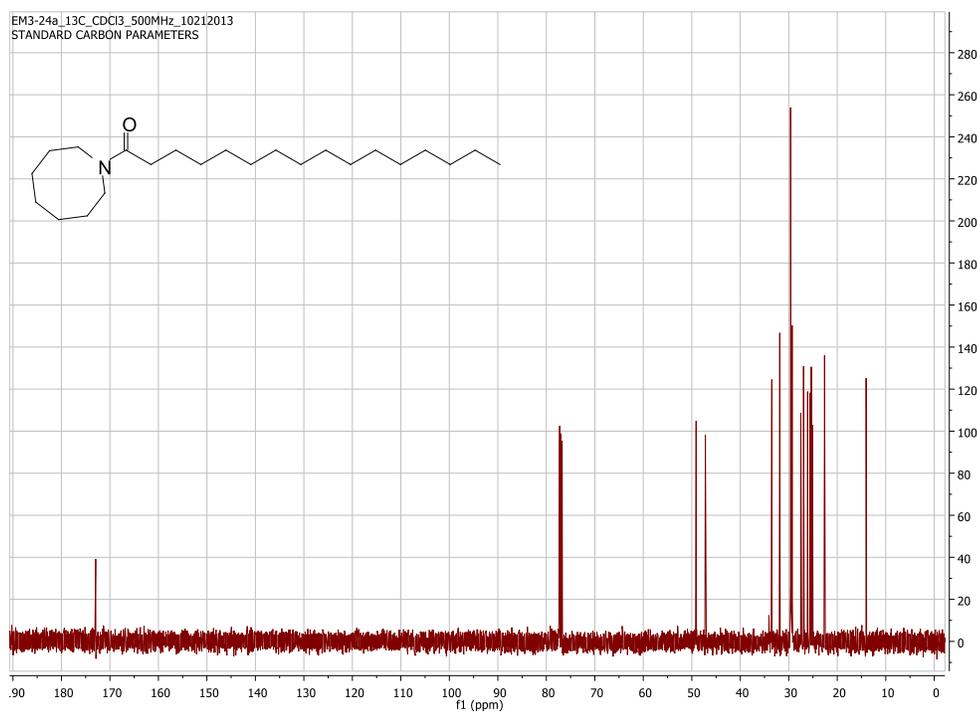
**Figure 6.6.93:**  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ) spectrum of compound **43**



**Figure 6.6.94:**  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ) spectrum of compound **43**



**Figure 6.6.95:** <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) spectrum of compound **44**



**Figure 6.6.96:** <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) spectrum of compound **44**

## 6.7 References

- (1) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. *Curr. Opin. Biotechnol.* **2010**, *21*, 787-793.
- (2) Tan, L. T. *J. Appl. Phycol.* **2010**, *22*, 659-676.
- (3) Gerwick, W. H.; Coates, R. C.; Engene, N.; Gerwick, L.; Grindberg, R. V.; Jones, A. C.; Sorrels, C. M. *Microbe* **2008**, *3*, 277-284.
- (4) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. In *Comprehensive Natural Products Chemistry II*, Moore, B., Crews, P., Eds.; Elsevier: Oxford, UK, 2010; Vol. 8, pp 141– 188.
- (5) Gutiérrez, M.; Pereira, A. R.; Debonsi, H. M.; Ligresti, A.; Di Marzo, V.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 2313-2317.
- (6) Tan, L. T.; Okino, T.; Gerwick, W. H. *J. Nat. Prod.* **2000**, *63*, 952-955.
- (7) Gerwick, W. H.; Tan, L. T.; Sitachitta, N. *Alkaloids Chem. Biol.* **2001**, *57*, 75-184.
- (8) Han, B.; McPhail, K. L.; Ligresti, A.; Di Marzo, V.; Gerwick, W. H. *J. Nat. Prod.* **2003**, *66*, 1364-1368.
- (9) Mevers, E.; Byrum, T.; Gerwick, W. H. *J. Nat. Prod.* **2013**, *76*, 1810-1814.
- (10) Nunnery, J. K.; Engene, N.; Byrum, T.; Cao, Z.; Jabba, S. V.; Pereira, A. R.; Matainaho, T.; Murray, T. F.; Gerwick, W. H. *J. Org. Chem.* **2012**, *77*, 4198-4208.
- (11) Gerwick, W. H.; Proteau, P. J.; Nagle, D. G.; Hamel, E.; Blokhin, A.; Slate, D. L. *J. Org. Chem.* **1994**, *59*, 1243-1245.
- (12) Wu, M.; Okino, T.; Nogle, L. M.; Marquez, B. L.; Williamson, R. T.; Sitachitta, N.; Berman, F. W.; Murray, T. F.; McGough, K.; Jacobs, R.; Colson, K.; Asano, T.; Yokokawa, F.; Shioiri, T.; Gerwick, W. H. *J. Am. Chem. Soc.* **2000**, *122*, 12041-12042.
- (13) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (14) Sitachitta, N.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 681-684.
- (15) Nannini, C. J. MS Thesis, Oregon State University, 2002.
- (16) Engene, N.; Rottacker, E. C.; Kaštovský, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komáerik, J.; Gerwick, W. H. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1171-1178.
- (17) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. **2010**, *21*, 1-7.
- (18) Somoza, A.; Terrazas, M.; Eritija, R. *Chem. Commun.* **2010**, *46*, 4270-4272.

- (19) Fu, R.; Ye, J. L.; Dai, X. J.; Ruan, Y. P.; Huang, P. Q. *J. Org. Chem.* **2010**, *75*, 4230-4243.
- (20) Theodorou, V.; Skobridis, K.; Tzakos, A. G.; Ragoussis, V. *Tetrahedron Lett.* **2007**, *48*, 8230-8233.
- (21) Senter, P.; Burgoyne, W. F.; Coates, R. M. *Organic Syntheses* **1988**, *6*, 327-333.
- (22) Novinec, M.; Lenarčič, B.; Baici, A. *FEBS J.* **2012**, *586*, 1062-1066.
- (23) Texter, J. In *Surfactants: A Practical Handbook*; Lange, K. R. (Ed.); Hanser Gardner Publications, Inc: Cincinnati, OH, 1999, Vol. 1: 1-66.
- (24) <http://www.kruss.de/services/education-theory/glossary/cmc/> (accessed Feb. 16<sup>th</sup>, 2014)
- (25) Reznik, G. O.; Vishwanath, P.; Pynn, M. A.; Sitnik, J. M.; Todd, J.; Wu, J.; Jiang, Y.; Keenan, B. G.; Castle, A. B.; Haskell, R. F.; Smith, T. F.; Somasundaran, P.; Jarrell, K. A. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1387-1397.
- (26) Lange, K. R. In *Surfactants: A Practical Handbook*; Lange, K. R. (Ed.); Hanser Gardner Publications, Inc: Cincinnati, OH, 1999, Vol. 1: 144-169.
- (27) Raney, O. In *Surfactants: A Practical Handbook*; Lange, K. R. (Ed.); Hanser Gardner Publications, Inc: Cincinnati, OH, 1999, Vol. 1: 171-202.
- (28) Texter, J. *J. Colloid. Interf. Sci.* **1999**, *217*, 208-210.
- (29) Vidyalakshmi, V.; Subramanian, M. S.; Rajeswari, S.; Srinivassan, T. G.; Vasudeva Rao, P. R. *Solvent Extr. Ion Exc.* **2003**, *21*, 399-412.
- (30) [http://www.uscg.mil/foia/docs/dwh/fosc\\_dwh\\_report.pdf](http://www.uscg.mil/foia/docs/dwh/fosc_dwh_report.pdf) (accessed Feb. 16<sup>th</sup>, 2014)
- (31) Place, B.; Anderson, B.; Meker, A.; Furlong, E. T.; Gray, J. L.; Tjeerdema, R.; Field, J. *Environ. Sci. Technol.* **2010**, *44*, 6016-6018.
- (32) Rico-Martínez, R.; Snell, T. W.; Shearer, T. L. *Environ. Pollut.* **2013**, *173*, 5-10.
- (33) Lippert, K.; Galinski, E. A. *Appl. Microbiol. Biotechnol.* **1992**, *37*, 61-65.
- (34) Krupa, J. C.; Mort, J. S. *Anal. Biochem.* **2000**, *283*, 99-103.
- (35) Olbrich, C.; Muller, R. H. *Int. J. Pharm.* **1999**, *180*, 31-39.
- (36) Bhairi, S. M. In *Detergents: A guide to the properties and uses of detergents in biological systems*; Calbiochem-Novabiochem Corporation: San Diego, CA, 2001, pp. 35-41.
- (37) Lindgren, Å.; Sjöström, M.; Wold, S. *Quant. Struct. Act. Relat.* **1996**, *15*, 208-218.
- (38) Lewis, M. A. *Wat. Res.* **1991**, *25*, 101-113.

- (39) Sandbacka, M.; Christianson, I.; Isomaa, B. **2000**, *14*, 61-68.
- (40) Talmage, S. S. In *Environmental and Human Safety of Major Surfactants: Alcohol Ethoxylates and Alkylphenol Ethoxylates*; CRC Press: New York, NY, 1994, Vol. 1: pp. 59-88.

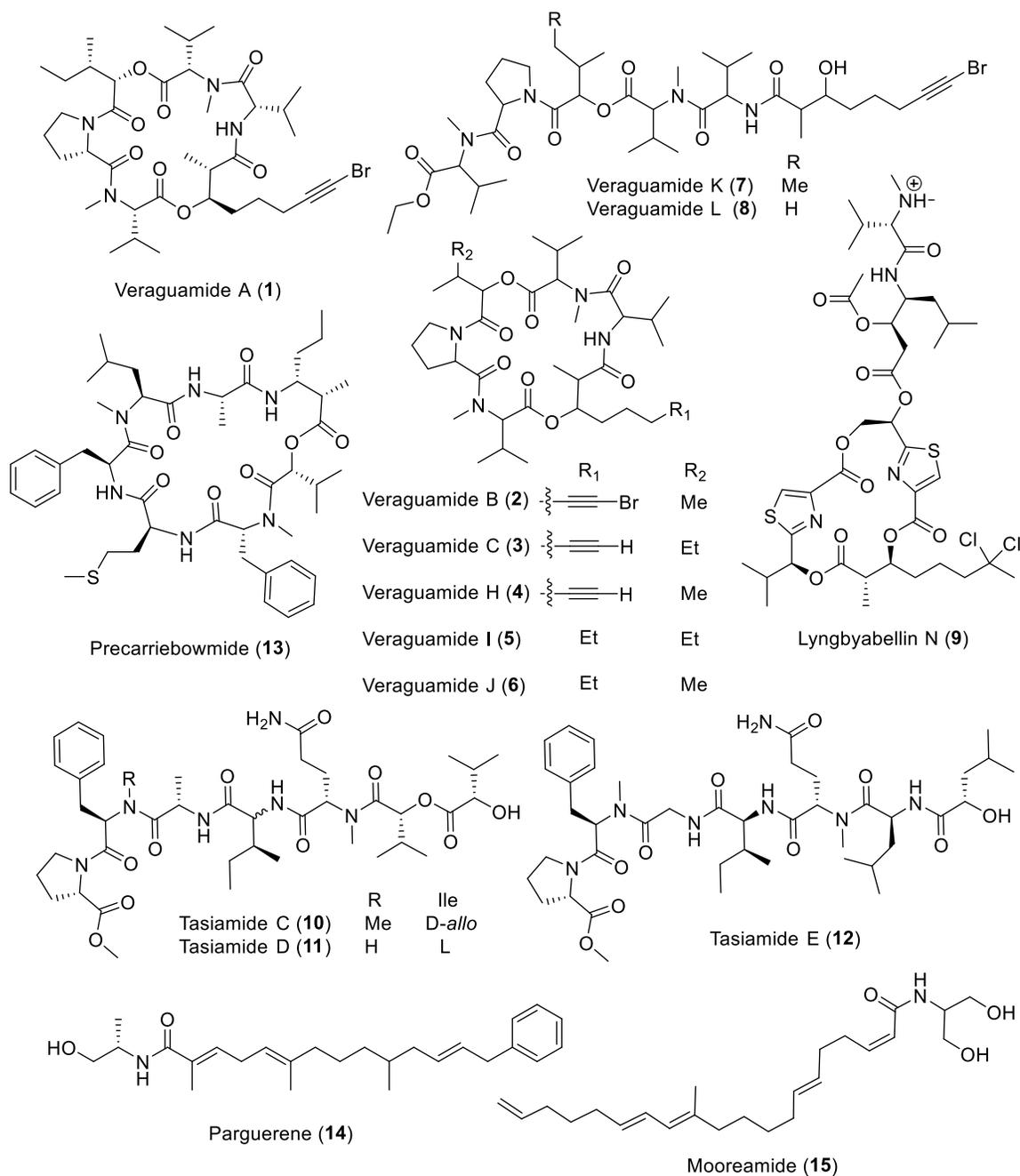
## Chapter 7: Conclusion & Future Work

### 7.1 Summary of the Research Presented in the Dissertation and Future Work

The major objective of the research present herein was to isolate novel, biologically active, secondary metabolites from marine cyanobacteria, with a focus on metabolites exhibiting cancer cell cytotoxicity, in hopes of discovering a new lead drug for the treatment of cancer. Chapter 2 through 5 addressed this major objective with the isolation and structure elucidation of 15 new cyanobacterial metabolites (figure 6.1), of which 8 were biologically evaluated with 2 compounds possessing potent cytotoxicity against HCT-116 (human colon cancer) and H-460 (human lung cancer), and one compound with cannabinoid receptor binding activity. A secondary objective involved a structure-activity relationship (SAR) investigation into lymbbyamide A, an alkyl amide, with the objective of both improving the efficacy of the previously reported brine shrimp toxicity and cannabinoid binding activity. At the same time, I desired to obtain a better understanding of the key functionalities responsible for this activity. A brief summary of each research chapter and future work for each project is described below.

Chapter 2 addressed this primary objective by describing the isolation and characterization of eight new metabolites, veraguamides A-C and H-L (**1-8**), which were isolated from a collection of *Oscillatoria margaritifera* from Coiba National Park (CNP) off of Panama's west coast, as part of the Panama International Cooperative Biodiversity Group program.<sup>1</sup> The planar structure of veraguamides A and L were fully deduced by 2D NMR spectroscopy and mass spectrometry, whereas the structures of veraguamides B, C, and H-K were mainly determined by a combination of <sup>1</sup>H NMR and MS<sup>2</sup>/MS<sup>3</sup> techniques. These new compounds are analogous to the mollusk-derived kulomo'opunalide natural products, with two of the veraguamides (C and H) containing the same terminal alkyne moiety.<sup>2</sup> However, four veraguamides, A, B, K, and L, also featured an alkynyl

bromide, a functionality that has been previously observed in only one other marine natural product, jamaicamide A.<sup>3</sup> Veraguamide A showed potent cytotoxicity to the H-460 human lung cancer cell line ( $LD_{50} = 141 \text{ nM}$ ).<sup>4</sup>



**Figure 7.1:** Cyanobacterial natural products that were discussed in the previous research chapters

Future directions for this research should include a total synthesis in order to obtain more material for biological evaluation as all of the natural product was used in the course of these studies. A synthetic route would also allow for the design of a structure-activity relationship study in order to probe the importance of key functional groups. From the natural analogs, it appears that the alkynyl bromide, and the overall cyclic constitution, are both vitally important for the potent activity; however, not much else can be learned from these co-metabolites. A total synthesis of veraguamide A would also help confirm the absolute stereochemical assignments, as the only published total synthesis raised questions because the  $^1\text{H}$  and  $^{13}\text{C}$  NMR chemical shifts between the natural product and synthetic material had significant differences, almost too drastic to be solely a stereochemical issue.<sup>5</sup> However, as another laboratory concurrently isolated several of the veraguamides from a Guam collection of cyanobacteria, and they independently confirmed the structural and stereochemical assignments, we feel confident of our original assignments.<sup>6</sup> Nevertheless, further synthetic investigations are needed to confirm this, and provide material for further biological evaluations.

Chapter 3 discusses the isolation and structure elucidation of a new lipopeptide, lyngbyabellin N (**9**), from an extract of the marine cyanobacterium *Moorea bouillonii* collected from Palmyra Atoll in the Central Pacific Ocean. The semi-crude fraction containing lyngbyabellin N showed strong cytotoxic activity in the H-460 assay; however, purification proved difficult as its peak shape on several chromatographic supports was extremely broad and unpredictable, even with the addition of a small percentage of acid. Ultimately, purification of lyngbyabellin N was accomplished using preparatory TLC as there was acceptable baseline resolution between neighboring bands. The planar structure of lyngbyabellin N was determined using 1D and 2D NMR techniques, which revealed an intriguing structural feature, an *N,N*-dimethylvaline residue, likely the cause of odd chromatographic characteristics. Two other interesting functionalities that lyngbyabellin N has include the presence of two thiazole rings and dichlorination on the polyketide

portion. Lyngbyabellin N exhibits strong cytotoxic activity against the HCT-116 colon cancer cell line ( $IC_{50} = 40.9 \pm 3.3$  nM).<sup>7</sup>

Future directions for this research would also include either a total synthesis or semi-synthesis in order to obtain more material for biological evaluation as there is no remaining natural product. Lyngbyabellin N exhibits potent cytotoxicity and is one of the most active compounds in this rather large family of metabolites.<sup>8-10</sup> This could potentially be caused by its hybrid nature, where it has the traditional lyngbyabellin highly modified cyclic NRPS/PKS portion which typically has anti-actin activity, but also has a tail group that consists of a protected leucine statine and the unique *N,N*-dimethylvaline terminus; these latter functionalities are structural features typical of antitubulin agents.<sup>8,11</sup> This might mean that lyngbyabellin N has two biological mechanisms of action, namely interaction with both microtubules and actin, which may have a synergistic effect; however, further investigations are needed to confirm this hypothesis. Since no more natural product remains, either a total synthesis or a much simpler semi-synthesis is needed to provide the material. Several other lyngbyabellin analogs are produced in larger quantities, thus it could be envisioned that one of these could be transformed into lyngbyabellin N by linking a synthetically-derived tail group which matches that of lyngbyabellin N.<sup>8</sup> This would reduce the number of overall synthetic steps and could be a renewable source of material for future biological evaluation.

Chapter 4 described the isolation and characterization of three new lipopeptides, tasiamides C-E (**10-12**), from a collection of the tropical marine cyanobacterium *Symploca* sp., collected near Kimbe Bay, Papua New Guinea. This collection has been particularly rich in secondary metabolites, such as kimbeamides A-C, kimbelactone A, and tasihalide C, which were previously characterized.<sup>12,13</sup> However, renewed investigations into a relatively polar and cytotoxic fraction of this extract yielded the three new lipopeptides. Their planar structures were deduced by traditional 2D NMR spectroscopy and tandem mass spectrometry, and their absolute configurations

were determined by a combination of Marfey's and chiral-phase GC-MS analysis. These new metabolites are similar to several previously isolated families of metabolites, including tasiamide, the grassystatins, and symplocin A, all of which were isolated from similar filamentous marine cyanobacteria.<sup>14-16</sup> Although both the grassystatins and symplocin A exhibit potent protease inhibition, tasiamide C and D were found to be inactive ( $IC_{50} > 20 \mu M$ ) against several cancer cell lines, suggesting that the statine residue present in the grassystatins and symplocin, but absent in the tasiamides, is important for the observed activity.<sup>17</sup>

Although the research described in chapter 4 did not directly involve the isolation of tasiamide, future directions for this project should involve deducing its correct absolute stereochemistry, as the isolation of tasiamide C-E put it in further doubt. Tasiamide was originally isolated by Williams et al. in 2002, with the same described absolute configuration as the new tasiamide E, however, they had opposite specific rotation signs.<sup>14</sup> A subsequent total synthesis suggested that the misassignment in tasiamide involved just the configuration of the *N*-MeGln residue, but their supporting evidence was solely NMR data (<sup>1</sup>H and <sup>13</sup>C) and specific rotation.<sup>18</sup> I performed DP4 calculations on the <sup>13</sup>C NMR data comparing the four synthetic analogs to that of the natural product (tasiamide), and this analysis revealed that the carbon data alone is insufficient to deduce the correct configuration.<sup>19</sup> The DP4 calculations also suggested that the misassignment in tasiamide may involve other residues, and thus, a broader investigation into the configuration of tasiamide is necessary to clarify its correct absolute configuration.

Chapter 5 discussed the isolation and structure elucidation of three new marine cyanobacterial natural products, precarriebowmide (**13**), parguerene (**14**), and mooreamide (**15**), from two separate collections of *Moorea* sp., one obtained from Puerto Rico and the other from Papua New Guinea. The planar structures of each were deduced by 2D NMR spectroscopy and mass spectrometry. Parguerene and mooreamide are modified alkyl amides, whereas precarriebowmide is a lipopeptide and represents only a minor modification compared to two other

known metabolites, carriebowmide and carriebowmide sulfone.<sup>20</sup> The identification of precarriebowmide led to an investigation into whether carriebowmide and carriebowmide sulfone were true secondary metabolites or isolation artifacts. Both parguerene and mooreamide are structurally reminiscent of the endocannabinoids, anadamide and 2-arachidonoylglycerol, and thus it was hypothesized that each would exhibit some cannabinoid receptor binding activity. Unfortunately, parguerene decomposed prior to being evaluated but mooreamide did exhibit moderate selective binding affinity towards CB<sub>1</sub> over CB<sub>2</sub> ( $K_i = 0.47 \mu\text{M}$  and  $K_i > 25 \mu\text{M}$ , respectively).<sup>21</sup>

Future directions for this research would be a total synthesis of parguerene in order to determine the absolute configuration of the isolated secondary methyl group (C-10) and to evaluate it for its biological potential, most importantly as a cannabinoid binding receptor ligand and for MDR reversing activity against Adriamycin resistant breast cancer cells (MDR-7adrR).<sup>22,23</sup> Parguerene is structurally similar to the natural product stipiamide, which was isolated from the Gram-negative soil bacterium, *Myxococcus stipitatus*.<sup>22</sup> From several structure-activity relationship studies, it was shown that reducing the number of conjugated double bonds in stipiamide significantly reduced its overall toxicity while the MDR reversing activity was maintained; because parguerene has fewer conjugated double bonds, it would likely have similar activity.<sup>23</sup> Thus, a total synthesis is needed to assign the C-10 methyl group and for biological evaluation studies.

Chapter 6 addressed the secondary objective of this thesis with an examination into the structure-activity relationship (SAR) of lyngbyamide A in a broad range of bioassays. Lyngbyamide A was isolated from a collection of *M. bouillonii* obtained from Grenada in 1995, and was shown to exhibit both brine shrimp toxicity and cannabinoid receptor binding activity.<sup>24</sup> Subsequently, three additional analogs were isolated from a Madagascar collection of *M. bouillonii*, and also shown to be toxic against brine shrimp.<sup>25</sup> During this latter study, Nannini and Gerwick

semi-synthesized a pyrrolidine derivative, which exhibited a 10-fold increase in brine shrimp potency, and thus suggested the potential to further improve the activity with other simple modifications.<sup>25</sup> The lynchbyamides are rather small alkyl amides consisting of a twelve carbon fatty acid tail group, which is functionalized with a *trans*-cyclopropyl ring at the C4 position, and a head group portion, which typically consists of a biogenic amine deriving from isoleucine, tyrosine, or phenylalanine. The synthetic tractability, coupled with their biological potential, made these targets for an interesting SAR study. In total, 50 analogs were designed and synthesized to probe the importance of key functional groups, such as the cyclopropyl ring, the chain length (both shorter and longer), head group polarity, and number of amine substituents. These analogs were constructed in three different rounds of synthesis and were evaluated for cathepsin L activation/inhibition, brine shrimp toxicity, cannabinoid receptor binding, cancer cell cytotoxicity, nitric oxide production in RAW cells, and ion channel modulation. Interestingly, a subset of these analogs showed strong activity toward the stabilization of cathepsin L and brine shrimp toxicity. The profile of some of the most active cathepsin L stabilizers resembled the reported activity of several commercially available surfactants, suggesting that these synthetic analogs may be surfactants. Using a tensiometer, it was shown that many of the tertiary amides have the ability to reduce the surface tension of an aqueous mixture and form micelles at concentrations as low as 13  $\mu\text{M}$ , which is lower than most commercially available surfactants.<sup>26</sup>

There are numerous directions that this project could foreseeably go in the future, such as continued biological evaluations targeting assays that are susceptible to surfactants, investigations into the exact type of surfactant these analogs are, and even looking into the ecological role similar natural products may have in their producing organism. Since the surfactant industry is a multi-billion dollar industry, there are many different types of surfactants and each has utility in various applications.<sup>27</sup> In order to potentially commercialize on their surfactant properties, it will be important to determine how these analogs fit into the current surfactant classifications, first by

determining what types of micelles they form. This will also likely aid in determining the overall utility that these tertiary amides might possess. However, from the work in chapter 6 it appears clear that these analogs have the ability to stabilize the cathepsin L enzyme and potentially other important enzymes, at lower concentrations than any of the common biochemically important surfactants (i.e. tween 20, SDS, triton X-100, BRIJ 35, and NP-40).<sup>26</sup>

## 7.2 Expected Future Direction of Marine Natural Products Chemistry

Natural products have had a huge impact on the pharmaceutical industry since its inception, with many of the hallmark drugs coming from a natural source.<sup>28</sup> However, in the 1980's the pharmaceutical industry began to shift away from natural products to synthetic combinatorial libraries in the search of new drugs. The idea behind this shift was that combinatorial libraries could be significantly larger (> 100,000 compounds), and thus a greater chance for a 'hit'; because these compounds are synthetically derived and there is no supply issue, they had a more obvious development path.<sup>29</sup> However, this has not translated into more drugs, and has overall been perceived as a failure, as it has produced only one drug (sorafenib) after many years of study.<sup>28</sup>

Fundamentally, there are issues in the druggability for many of the combinatorial library metabolites as they are derived from simple, high yielding synthetic reactions. Compared to natural products, combinatorial compounds have significantly less chiral centers (0 vs. 4), have higher molecular weights (389 vs. 362), increased rotatable bonds (6 vs. 3), fewer fused rings systems (1 vs. 2.5), increased degree of unsaturation (12 vs. 8), fewer Lipinski-type donors (1 vs. 2), higher in-ring Lipinski acceptors (2 vs. 1), and are less lipophilic with increased SlogP values (4.2 vs. 2.7)(all median values).<sup>30</sup> Overall, combinatorial libraries are simpler molecules, and incorporate a significant number of heteroaromatic rings to avoid chiral centers and are much more flexible than natural products, and thus incur higher detrimental entropic consequences.<sup>30</sup> In medicinal chemistry it is common knowledge that "the removal of chiral centers, introducing additional

flexibility into the molecule and degreasing its size generally leads to a less specific and weaker activity”;<sup>31</sup> however, that is exactly what happened when the pharmaceutical industry transitioned its focus away from natural products to synthetic combinatorial libraries.

Although the focus of the pharmaceutical industry has shifted away from natural products, natural product derived or inspired metabolites still have had a huge impact on the number of new chemical entities (NCE) approved for use worldwide. Between 1981 and 2010 only 29% of all NCEs were of truly synthetic origin, whereas 71% were either natural products, derived from a natural product, or synthetic but mimicking a natural product.<sup>28</sup> Natural products cannot, and should not be replaced by synthetic combinatorial libraries, as natural products provide structural diversity and complexity that is currently unattainable by combinatorial synthesis. The producing organisms of these natural products have evolved over thousands of years to enantiomerically biosynthesize complex metabolites that have co-evolved with receptors to develop high receptor—ligand specificity.<sup>30</sup> Furthermore, only a small percentage of all organisms have even been chemically investigated, and an even smaller percentage of marine invertebrates and especially microbes, thus highlighting the exceptional potential natural products still have in producing new drug leads.<sup>32</sup>

Several obstacles impede the development of natural products into drugs, with the biggest being the supply issue. Currently, it is possible to isolate and elucidate a structure on sub-milligram quantities of material, and thus it is feasible to work on minor components of an extract.<sup>33,34</sup> However, this significantly reduces the breath of biological evaluation and begs the question, what good is a compound with no reported biological activity? The goal of natural product chemists whose focus is on human health impacts should not only be to isolate structurally diverse metabolites but also ones that possess interesting biological activity. We not only need to find structurally diverse metabolites, but also enough material to biologically evaluate these compounds in a broad range of assays to reveal novel biological targets. Over the past decade multiple different

technological platforms have been developed to improve identification of both structurally diverse and biologically active compounds, including mass spectrometry techniques,<sup>35,36</sup> genomic techniques,<sup>37,38</sup> and biological screening methodologies;<sup>39</sup> however, neither of the isolation techniques (MS and genomic) incorporate biological screening protocols until the very end of the process.

The improvement in MS and genomic techniques have led to a significant increase in knowledge about the exceptional biosynthetic capabilities of natural product-producing organisms,<sup>40</sup> has provided access to ‘silent’ pathways,<sup>38</sup> and has led to the targeted isolation of structurally novel, secondary metabolites.<sup>41</sup> However, the lack of incorporation of biological assays until the end of a project makes it impractical for drug development by pharmaceutical companies. An ideal protocol would incorporate high-throughput, high-content bioassay screening along with an isolation technique that enriches for structurally novel metabolites. For example, using chromatography to reduce the complexity of the sample and analyzing these fractions on a LCMS that has a splitter prior to the MS, where one portion goes to the MS and another to a 96-well plate for bioassay, would yield both MS data along with bioactivity data on almost pure compounds in a high-throughput fashion.<sup>42</sup> Another emphasis could be focused on building a library of secondary metabolites, which focuses on the structural diversity of the metabolites, rather than on the number of compounds.<sup>28</sup> This will likely lead to more ‘hits’ than any combinatorial library, which can be subsequently followed up by a structure-activity relationship to refine the activity.

Although investigations into the marine realm is still a relatively young field, it has already led to 12 approved drugs, with the potential for many more. It has even been shown that marine natural products are likely our best source of new drugs, with approximately 1.7- to 3.3-fold more approved drugs per tested compounds over the industry average, which will only likely increase further with improvements in technology.<sup>43</sup> Overall, I strongly believe that the future of marine

natural products is extremely bright, with huge pharmaceutical potential, and thus worthy of our continued investment.

## 7.3 References

- (1) <http://www.icbg.org/> (Accessed March 30, 2011).
- (2) Nakao, Y.; Yoshida, W. Y.; Szabo, C. M.; Baker, B. J.; Scheuer, P. J. *J. Org. Chem.* **1998**, *63*, 3272-3280.
- (3) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Geoger, D. E.; Roberts, M. A.; Gerwick, W. H. *Chem. Biol.* **2004**, *11*, 817-833.
- (4) Mevers, E.; Liu, W. T.; Engene, N.; Mohimani, H.; Byrum, T.; Pevzner, P. A.; Dorrestein, P. C.; Spadafora, C.; Gerwick, W. H. *J. Nat. Prod.* **2011**, *74*, 928-936.
- (5) Wang, D.; Jia, X.; Zhang, A. *Org. Biomol. Chem.* **2012**, *10*, 7027-7030.
- (6) Salvador, L. A.; Biggs, J. S.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2011**, *74*, 917-927.
- (7) Choi, H.; Mevers, E.; Byrum, T.; Valeriote, F. A.; Gerwick, W. H. *Eur. J. Org. Chem.* **2012**, 5141-5150.
- (8) Han, B. N.; McPhail, K. L.; Gross, H.; Goeger, D. E.; Mooberry, S. L.; Gerwick, W. H. *Tetrahedron* **2005**, *61*, 11723-11729.
- (9) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Pau, V. J.; Mooberry, S. L. *J. Nat. Prod.* **2000**, *63*, 611-615.
- (10) Marquez, B. L.; Watts, K. S.; Yokochi, A.; Roberts, M. A.; Verdier-Pinard, P.; Jimenez, J. I.; Hamel, E.; Scheuer, P. J.; Gerwick, W. H. *J. Nat. Prod.* **2002**, *65*, 866-871.
- (11) 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. *J. Am. Chem. Soc.* **2008**, *130*, 6324-6325.
- (12) Nunnery, J. K.; Engene, N.; Byrum, T.; Cao, Z.; Jabba, S. V.; Pereira, A. R.; Matainaho, T.; Murray, T. F.; Gerwick, W. H. *J. Org. Chem.* **2012**, *77*, 4198-4208.
- (13) Nunnery, J. K. PhD. Dissertation, University of California San Diego, 2012.
- (14) Williams, P. G.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **2002**, *65*, 1336-1339.
- (15) Kwan, J. C.; Eksioglu, E. A.; Liu, C.; Paul, V. J.; Luesch, H. *J. Med. Chem.* **2009**, *52*, 5732-5747.
- (16) Molinski, T. F.; Reynolds, K. A.; Morinaka, B. I. *J. Nat. Prod.* **2012**, *75*, 425-431.
- (17) Mevers, E.; Haeckl, F. P. J.; Boudreau, P. D.; Byrum, T.; Dorrestein, P. C.; Valeriote, F. A.; Gerwick, W. H. *J. Nat. Prod.* **2014**, accepted.
- (18) Ma, Z.; Song, N.; Li, C.; Zhang, W.; Wang, P.; Li, Y. *J. Pept. Sci.* **2008**, *14*, 1139-1147.

- (19) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *37*, 12946-12959.
- (20) Gunasekera, S. P.; Ritson-Williams, R.; Paul, V. J. *J. Nat. Prod.* **2008**, *71*, 2060-2063.
- (21) Mevers, E.; Byrum, T.; Gerwick, W. H. *J. Nat. Prod.* **2013**, *76*, 1810-1814.
- (22) Kim, J. K.; Furihata, K.; Yamanaka, S.; Fudo, R.; Seto, H. *J. Antibiot.* **1991**, *44*, 553-556.
- (23) Andrus, M. B.; Lepore, S. D.; Turner, T. M. *J. Am. Chem. Soc.* **1997**, *119*, 12159-12169.
- (24) Sitachitta, N.; Gerwick, W. H. *J. Nat. Prod.* **1998**, *61*, 681-684.
- (25) Nannini, C. J. MS Thesis, Oregon State University, 2002.
- (26) Bhairi, S. M. In *Detergents: A guide to the properties and uses of detergents in biological systems*; Calbiochem-Novabiochem Corporation: San Diego, CA, 2001, pp. 35-41.
- (27) Reznik, G. O.; Vishwanath, P.; Pynn, M. A.; Sitnik, J. M.; Todd, J.; Wu, J.; Jiang, Y.; Keenan, B. G.; Castle, A. B.; Haskell, R. F.; Smith, T. F.; Somasundaran, P.; Jarrell, K. A. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1387-1397.
- (28) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311-335.
- (29) Carter G. T. *Nat. Prod. Rep.* **2011**, *28*, 1783-1789.
- (30) Feher, M.; Schmidt, J. M. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218-227.
- (31) Bohm, H. J.; Klebe, G.; Kubinyi, H. *Wirkstoffdesign*; Spektrum Akademischer Verlag: Heidelberg, 1996, p 153.
- (32) Cragg, G. M.; Newman, D. J. *Biochim. Biophys. Acta* **2013**, *1830*, 3670-3695.
- (33) Molinski, T. F. *Nat Prod. Rep.* **2010**, *27*, 321-329.
- (34) Molinski, T. F. *Curr Opin. Biotech.* **2010**, *21*, 819-826.
- (35) Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22-31.
- (36) Mohimani, H.; Liu, W. T.; Mylne, J. S.; Poth, A. G.; Colgrave, M. L.; Tran, D.; Selsted, M. E.; Dorrestein, P. C.; Pevzner, P. A. *J. Proteome Res.* **2011**, *10*, 4505-4512.
- (37) Eustáquio, A. S.; McGlinchey, R. P.; Liu, y.; Hazzard, C.; Beer, L. L.; Florova, G.; Alhamadsheh, N. M.; Lechner, A.; Kale, A. J.; Kobayashi, Y.; Reynolds, K. A.; Moore, B. S. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12295-12300.
- (38) Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *P. Natl. Acad. Sci. USA* **2014**, ASAP.

- (39) Schulze, C. J.; Bray, W. M.; Woerhman, M. H.; Stuart, J.; Lokey, R. S.; Linington, R. G. *Chem. Biol.* **2013**, *20*, 285-295.
- (40) Jones, A. C.; Monroe, E. A.; Podell, S.; Hess, W. R.; Kages, S.; Esquenazi, E.; Niessen, S.; Hoover, H.; Rothmann, M.; Lasken, R. S.; Yates III, J. R.; Reinhard, R.; Kube, M.; Burkart, M. D.; Allen, E. E.; Dorrestein, P. C.; Gerwick, W. H.; Gerwick, L. *P. Natl. Acad. Sci. USA* **2011**, *108*, 8815-8820.
- (41) Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22-31.
- (42) Wen, L.; Wei, Q.; Chen, G.; Liu, F.; Zhang, S.; You, T. *Pharmacogn. Mag.* **2013**, *9*, S25-31.
- (43) Gerwick, W. H.; Moore, B. S. *Chem. Biol.* **2012**, *19*, 85-98.
- (44) [http://www.phrma.org/sites/default/files/159/rd\\_brochure\\_020307](http://www.phrma.org/sites/default/files/159/rd_brochure_020307) (accessed Feb. 3rd, 2007)