

UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Exploration of Non-enzymatic Natural Products and the Identification of a Novel Type III Secretion System Inhibitor

Permalink

<https://escholarship.org/uc/item/4bm944b2>

Author

Bouthillette, Leah Marie

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

**EXPLORATION OF NON-ENZYMATIC NATURAL PRODUCTS AND THE
IDENTIFICATION OF A NOVEL TYPE III SECRETION SYSTEM INHIBITOR**

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

DOCTOR OF PHILOSOPHY
in
CHEMISTRY

By

Leah M. Bouthillette

September 2024

The Dissertation of Leah Bouthillette is
approved:

Professor John B. MacMillan, Chair

Professor Laura Sanchez

Professor Timothy Johnstone

Peter Biehl
Vice Provost and Dean of Graduate Studies

Copyright © by
Leah M. Bouthillette
2024

Table of Contents

	Page
Title Page	i
Copyright Notice	ii
Table of Contents	iii
List of Figures	vi
List of Schemes	xiii
List of Tables	Xiv
Abstract	xv
Dedication	xvii
Acknowledgments	xviii
Chapter 1. Overview of Non-enzymatic Transformations in Natural Product	1
Biosynthesis	
1.1 Introduction	2
1.1.1 Influence of biosynthetic studies	2
1.1.2 Foundations of natural product biosynthesis	3
1.1.3 Non-enzymatic reactions in the cellular milieu	4
1.1.4 Incomplete categorization of non-enzymatic processes in natural products	5
1.2 Classification of non-enzymatic mechanisms	7
1.3 Intramolecular reactions	8
1.4 Multi-component reactions	13
1.4.1 Homologous multi-component dimerization reactions	13
1.4.2 Heterologous cascading reactions	30
1.5 Tailoring reactions	46

1.6 Light-induced reactions	53
1.7 Non-enzymatic reactions for medicinal chemistry	61
1.8 Discussion of common precursors	63
1.9 Conclusion	67
1.10 References	70
 Chapter 2. Discovery Platform to Harness Non-enzymatic Scaffolds in Natural Products	 80
2.1 Introduction	81
2.2 Discovery platform and selection of precursors	85
2.3 Initial evaluation of selected precursors	87
2.3.1 Incorporation of ^{15}N anthranilic acid in discoipyrrole A	87
2.3.2 Incorporation of methyl-3-oxobutanoate-2,4- $^{13}\text{C}_2$ into novel ammosamide analog	89
2.3.4 Incorporation of 4-fluoroaniline into novel iminoquinone	90
2.4 Incorporation of 5-fluoroanthranilic acid into discoipyrrole A	92
2.5 Survey of microbial library with fluorinated precursors	102
2.6 Conclusion	107
2.7 Experimental procedures	108
2.8 References	115
 Chapter 3. Biomimetic Synthesis and Investigation of the Non-enzymatic Formation of the Tricrilactones	 119
3.1 Introduction	120
3.2 Putative biosynthetic pathways and preliminary mechanistic studies of the tricrilactones	122

3.3 Synthesis of amine fragment 3.22	128
3.4 Synthesis of protilactone A	128
3.4.1 Synthesis of acid coupling partner 3.28	129
3.4.2 Synthesis of fragment 3.29	131
3.4.3 Final coupling for protilactone A	132
3.5 Dimerization efforts for the synthesis of tricilactone G	134
3.6 Conclusion	142
3.7 Experimental procedures	143
3.8 References	158
 Chapter 4. Bioactivity Guided Isolation of the Tricyclic Lasso Peptide Aborycin with Novel Type III Secretion System Inhibitory Activity Against <i>Pseudomonas aeruginosa</i>	 161
4.1 Introduction	161
4.2 Structural analysis	164
4.3 Biosynthetic gene cluster comparative analysis	175
4.4 Bioactivity assays against Type III Secretion Systems	177
4.4.1 Primary screens for T3SS inhibition	178
4.4.2 Orthogonal screen for secretion inhibition	179
4.4.3 <i>Galleria mellonella</i> infection model	181
4.5 Conclusion	182
4.6 Experimental procedures	184
4.7 References	190
 Appendix A: Supplementary Experimental Data for Chapter 2	 195
A.1 Spectral data for open intermediate of fluorinated discoipyrrole A	196

A.2 Spectral data for fluorinated discoipyrrole A	201
A.3 Spectral data for the synthesis of fluorinated discoipyrrole A	205
Appendix B: Supplementary Experimental Data for Chapter 3	211
B.1 Spectral data for amine fragment 3.22	212
B.2 Spectral data for fragment 3.28	215
B.3 Spectral data for fragment 3.29	217
B.4 Spectral data for protilactone A	221
Appendix C: Supplementary Experimental Data for Chapter 4	225
C.1 Spectral data for aborycin	226
C.2 Aborycin-fluorophore spectral data	234

List of Figures

Chapter 1

Figure 1.1. Simplified models depicting the types of biosynthetic pathways observed in natural product biosynthesis	7
Figure 1.2. The biosynthetic mechanism for the endiandric acids	9
Figure 1.3. The original family of the elansolids	10
Figure 1.4. The proposed mechanism for elansolid A1/A2	12
Figure 1.5. Exploration of the carpanone biosynthetic mechanism	14
Figure 1.6. The proposed mechanism for the biosynthesis of oxazinin A	15
Figure 1.7. Progression of the investigation of paracaseolide A	17
Figure 1.8. Representative family members of the securinega alkaloids which include both monomeric and dimeric structures	20

Figure 1.9. Original proposal for the non-enzymatic biosynthetic mechanism of flueggeanine A	21
Figure 1.10. Synthetic dimerization attempts of securinega monomers	22
Figure 1.11. Biosynthetic mechanism of ($\pm$)-homodimericin A	25
Figure 1.12. A sample of natural products from diverse sources that feature dimerization through a methylene bridge	28
Figure 1.13. Biosynthetic proposal for the dimerization of bohemamine A to form dibohemamine A	29
Figure 1.14. Mechanism for the non-enzymatic biosynthesis of the jadomycins	32
Figure 1.15. Interrogation of the discoipyrrole family and their origin	34
Figure 1.16. Proposed mechanism for the non-enzymatic biosynthesis of discoipyrrole A	35
Figure 1.17. Investigation of the rubrolone biosynthesis	37
Figure 1.18. Conceptual overview of the non-enzymatic construction for the maximiscins and related compounds	39
Figure 1.19. Mechanistic details for the non-enzymatic formation of maximiscin from various precursors as suggested by in vitro studies and DFT calculations	40
Figure 1.20. The non-enzymatic formation of the pyonitrins	42
Figure 1.21. Non-enzymatic Pictet-Spengler mechanism of pyonitrin A formation	43
Figure 1.22. Rhabdoplanin family and its formation	45
Figure 1.23. Non-enzymatic derivatization of the ammosamide family via the highly reactive iminium functional group of ammosamide C	47
Figure 1.24. Proposed mechanism for the non-enzymatic conversion of DHNA to ACNQ	49
Figure 1.25. Proposed mechanism for the biosynthesis of juanlimycin A and B	50

Figure 1.26. The progression of stealthin C discovery	52
Figure 1.27. UV-induced formation of γ -pyrone natural products	55
Figure 1.28. UV-induced non-enzymatic formation of phototridachiahydroprone	56
Figure 1.29. Validation of [1,3]-sigmatropic rearrangement of the cembranes	58
Figure 1.30. Investigation of intricarene biosynthesis	59
Figure 1.31. Experimental validation of non-enzymatic [2 + 2] cycloaddition of dactylone to form aplydactylone by biological UV light	60
Figure 1.32. Recapitulated figures of the elansolids and rubrolones	62
Figure 1.33. Natural product scaffolds that non-enzymatically incorporate anthranilic acid	64
Figure 1.34. Representation of different biological applications of non-enzymatic mechanisms	66

Chapter 2

Figure 2.1. Examples of natural products that incorporate non-enzymatic transformations in their biosynthesis	81
Figure 2.2. Non-enzymatic formation of amidine containing ammosamide analogs	83
Figure 2.3. General methodology of the precursor-directed discovery platform	87
Figure 2.4. ^1H – ^{15}N HMBC spectral comparison of ^{15}N -anthranilic acid and supplemented <i>B. hunanensis</i> extract	88
Figure 2.5. Proposed formation of novel ammosamide analog	89
Figure 2.6. ^{19}F spectral comparison and structure of novel iminoquinone	91
Figure 2.7. Structure determination of novel iminoquinone	92
Figure 2.8. ^{19}F NMR spectral comparison of various extracts, highlighting the differences in δ_{F}	94

Figure 2.9. Analytical data for 2.20	97
Figure 2.10. Analytical data and structure of 2.21	98
Figure 2.11. Structural assignments of 2.21	99
Figure 2.12. Key HSQC and ^1H – ^{13}C HMBC correlations for structural assignment of 2.21	100
Figure 2.13. Comparison of the aromatic region in the ^1H NMR spectra of 2.20 and 2.21	101
Figure 2.14. Putative structures obtained from DFT calculations	102
Figure 2.15. ^{19}F NMR spectral data for the microbial survey	106
 Chapter 3	
Figure 3.1. Representation of structural diversity and broad bioactivity of macrolides	120
Figure 3.2. The tricrilactone family of natural products	122
Figure 3.3. Structural similarities between protricrilactone A and decarestrictine analogs that undergo non-enzymatic interconversion	123
Figure 3.4. Proposed biosynthetic pathway of tricrilactone A	124
Figure 3.5. Proposed biosynthetic pathway for tricrilactone B–D	125
Figure 3.6. Proposed biosynthetic pathway for tricrilactone G	127
Figure 3.7. Retrosynthesis of protricrilactone A	129
Figure 3.8. Summary of preliminary dimerization conditions	135
Figure 3.9. UV chromatogram of the amine-catalyzed dimerization at 280 nm	137
Figure 3.10. UV chromatogram of simplified amine reaction at 280 nm and corresponding mass spectra	139
Figure 3.11. Mass spectrometry results from HPLC purification	141

Chapter 4

Figure 4.1. Structure and peptide sequence of siamycin I and aborycin	164
Figure 4.2. HRESIMS analysis of SNE-013 bioactive metabolite	166
Figure 4.3. Magnified regions of the ^1H NMR spectrum of aborycin in various solvents highlighting improved resolution	167
Figure 4.4. Identification of Phe ¹⁰ –Ala ¹¹ –Gly ¹² sequence through HMBC correlations	169
Figure 4.5. Magnified HMBC spectrum highlighting critical correlations between Val ¹⁷ and Val ¹⁸ residues for aborycin	171
Figure 4.6. Predicted aborycin gene cluster from <i>Streptomyces</i> sp. SNE-013	176
Figure 4.7. Aborycin-containing SNE-013 extract inhibits T3SS gene expression in <i>Pseudomonas aeruginosa</i> without effecting cell viability or luciferase production	179
Figure 4.8. Aborycin and siamycin I inhibit secretion of T3SS effector proteins	181
Figure 4.9. Aborycin protects <i>Galleria mellonella</i> larvae from <i>P. aeruginosa</i> -induced death	182

Appendix A

Figure A.1. ^1H NMR spectrum of 2.21	197
Figure A.2. ^{13}C NMR spectrum of 2.21	197
Figure A.3. COSY spectrum of 2.21	198
Figure A.4. ^1H – ^{13}C HSQC spectrum of 2.21	198
Figure A.5. ^1H – ^{13}C HMBC spectrum of 2.21	199
Figure A.6. Full ^{19}F spectrum of 2.21	199
Figure A.7. UV, TIC, and mass spectrum of 2.21	200
Figure A.8. ^1H NMR spectrum of 2.20	202

Figure A.9. ^{13}C NMR spectrum of 2.20	202
Figure A.10. COSY spectrum of 2.20	203
Figure A.11. ^1H – ^{13}C HSQC spectrum of 2.20	203
Figure A.12. Full ^{19}F NMR spectrum of 2.20	204
Figure A.13. TIC chromatogram and mass spectrum of 2.20	204
Figure A.14. ^1H NMR spectrum of 2.16	205
Figure A.15. ^{13}C NMR spectrum of 2.16	206
Figure A.16. UV chromatogram and mass spectrum of 2.16	206
Figure A.17. ^1H NMR spectrum of 2.17	207
Figure A.18. ^{13}C NMR spectrum of 2.17	207
Figure A.19. UV chromatogram and mass spectrum of 2.17	208
Figure A.20. ^1H NMR spectrum of 2.18	208
Figure A.21. ^{13}C NMR spectrum of 2.18	209
Figure A.22. UV chromatogram and mass spectrum of 2.18	209
Figure A.23. ^1H NMR spectrum of 2.19	210

Appendix B

Figure B.1. ^1H NMR spectrum of 3.26	212
Figure B.2. ^{13}C NMR spectrum of 3.26	212
Figure B.3. ^1H NMR spectrum of 3.27	213
Figure B.4. ^{13}C NMR spectrum of 3.27	213
Figure B.5. ^1H NMR spectrum of 3.22	214
Figure B.6. ^1H NMR spectrum of 3.22	214
Figure B.7. ^1H NMR spectrum of 3.35	215
Figure B.8. ^{13}C NMR spectrum of 3.35	215
Figure B.9. ^1H NMR spectrum of 3.28	216

Figure B.10. ^{13}C NMR spectrum of 3.28	216
Figure B.11. ^1H NMR spectrum of 3.37	217
Figure B.12. ^1H NMR spectrum of 3.38	217
Figure B.13. ^1H NMR spectrum of 3.39	218
Figure B.14. ^1H NMR spectrum of 3.40	218
Figure B.15. ^1H NMR spectrum of 3.41	219
Figure B.16. ^1H NMR spectrum of 3.42	219
Figure B.17. ^1H NMR spectrum of 3.43	220
Figure B.18. ^1H NMR spectrum of 3.29	220
Figure B.19. ^1H NMR spectrum of 3.44	221
Figure B.20. ^{13}C NMR spectrum of 3.44	221
Figure B.21. ^1H NMR spectrum of 3.45	222
Figure B.22. ^{13}C NMR spectrum of 3.45	222
Figure B.23. ^1H NMR spectrum of botryolide B (3.9)	223
Figure B.24. ^{13}C NMR spectrum of botryolide B (3.9)	223
Figure B.25. ^1H NMR spectrum of protricrilactone A	224

Appendix C

Figure C.1. Annotated ^1H NMR spectrum of aborycin	226
Figure C.2. Expanded ^1H NMR spectrum of aborycin without annotations	227
Figure C.3. Annotated ^{13}C NMR spectrum of aborycin	227
Figure C.4. Expanded ^{13}C NMR spectrum of aborycin and magnified image of the carbonyl region	228
Figure C.5. ^1H – ^{13}C HSQC NMR spectrum of aborycin	228
Figure C.6. COSY NMR spectrum of aborycin	229
Figure C.7. Short-range (4 Hz) ^1H – ^{13}C HMBC spectrum of aborycin	229

Figure C.8. Long-range (8 Hz) ^1H – ^{13}C HMBC spectrum of aborycin	230
Figure C.9. TOCSY spectrum of aborycin	230
Figure C.10. NOSEY spectrum of aborycin	231
Figure C.11. HSQC–TOCSY spectrum of aborycin	231
Figure C.12. ^1H – ^{15}N HSQC spectrum of aborycin	232
Figure C.13. ^1H – ^{15}N HMBC spectrum of aborycin	232
Figure C.14. HPLC chromatogram of aborycin at 280 nm	233
Figure C.15. TIC chromatogram of aborycin	233
Figure C.16. Structure of aborycin with rhodamine cadaverine fluorescent tag	234
Figure C.17. Mass spectrum of aborycin-fluorophore showing $[\text{M}+3\text{H}]^{3+}$ and $[\text{M}+2\text{H}]^{2+}$ ions	234
Figure C.18. HPLC chromatogram of aborycin-fluorophore at 280 nm	235

List of Schemes

Chapter 2

Scheme 2.1. Synthesis of fluorinated discoipyrrole A	96
--	----

Chapter 3

Scheme 3.1. Synthetic scheme for the amine coupling partner 3.22	128
Scheme 3.2. Original synthetic scheme for the synthesis of fragment 3.28	130
Scheme 3.3. Optimal synthetic pathway for fragment 3.28	130
Scheme 3.4. Synthetic pathway for fragment 3.29	132
Scheme 3.5. Final coupling pathway to protricrilactone A	133
Scheme 3.6. Amine catalyzed dimerization pathway to tricrilactone G	136
Scheme 3.7. Final dimerization attempt with simplified amine fragment	138

List of Tables

Chapter 3

Table 3.1. Summary of dimerization attempts	156
---	-----

Chapter 4

Table 4.1. NMR assignments for elucidation of aborycin	172
--	-----

Table 4.2. Putative functions of genes encoded within aborycin biosynthetic gene cluster of SNE-013	177
--	-----

Appendix A

Table A.1. NMR assignments for elucidation of 2.21	196
--	-----

Table A.2. NMR assignments for elucidation of 2.20	201
--	-----

Exploration of Non-enzymatic Natural Products and the Identification of a Novel Type

III Secretion System Inhibitor

By

Leah M. Bouthillette

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

Doctor of Philosophy in Chemistry

Abstract

Natural products are secondary metabolites produced from a variety of organisms found in nature that have been and continue to be a rich source for a wide range of therapeutics playing a key role in drug discovery. They are also an abundant source of structurally complex and diverse molecules that can be difficult to imagine and achieve through combinatorial chemistry alone. Exploring the depth of natural product biosynthesis studies has opened up new avenues in the field of natural products from our understanding of chemistry, protein function and dynamics, and genetic regulation, which aid in the development of drugs and chemical tools. It is well known that the biosynthesis of natural products is primarily carried out by a series of protein catalysts encoded in the genome, however, a small but growing number of molecules involve one or more protein independent or non-enzymatic steps. As we continue to build a strong foundation of how microbes and higher organisms produce natural products, therein lies a need for understanding how non-enzymatic biosynthetic steps occur. Presented here is the development of a nuclear magnetic resonance (NMR)-based discovery platform to identify natural products scaffolds that can undergo non-enzymatic reactions in microbial cultures

and the investigation of the non-enzymatic formation of the tricrilactone family of natural products through biomimetic synthesis. Separately, this thesis also discusses the re-isolation of aborycin, a tricyclic lasso peptide, from a *Streptomyces* sp. and its biological evaluation against Type III Secretion System (T3SS) in *Pseudomonas aeruginosa*. This is the first report to date of aborycin as an effective T3SS inhibitor.

To My Dearest Parents
Denise and Louie Bouthillette

Acknowledgements

First and foremost, I would like to express a sincere thank you to my research advisor, Professor John B. MacMillan. You have played a pivotal role in my development as a scientist and none of this would have been possible without your encouragement. Over the years, you have provided me with countless opportunities to expand my knowledge and grow both as a chemist and as an individual, for which I am deeply grateful. Although I never quite got a handle on the “sports ball”, the lessons you have taught me in both science and life have been invaluable. Your support has inspired me to strive for excellence in my research. Under your mentorship, I have become a more confident and capable scientist. It has truly been a pleasure having the opportunity to work for you, and I will forever be grateful for all you have done for me. Thank you for trusting and believing in me.

I would also like to acknowledge my committee members, Professor Laura Sanchez and Professor Timothy Johnstone. Laura, you have been a pillar of support since you moved to UCSC. Thank you for always providing detailed answers to my questions, including me in the outreach programs you organize, and of course letting me hang out with Teddy (especially on the particularly rough days). You have helped create this wonderful little natural products community at UCSC, one that I have been so happy to be a part of. Thank you for your endless support and for being on my committee. Tim, you have had made a tremendous impact on me over the years and have been a true inspiration. Whether it is talking about radial distribution of electron density on a Saturday or delving into the intricacies of different font styles, I have learned so much from you. I appreciate you for always making time for me and for including me in your lab’s activities. Thank you for all your help along the way.

Great science is not accomplished by an individual alone; it takes a village. For that I owe many thanks to all the members of the MacMillan lab, both past and present. Every one of you has made a lasting impression on me, and I am deeply grateful for you all. Jocelyn, thank you for teaching me all there needs to know about HPLC, especially taking them apart

and putting them back together again because that has happened way more often than I'd like to admit. In addition to being a great mentor, you are a great friend and I have missed our Chinese food movie nights so much. Victor, one of the most fascinating people I have met. I cherish our bizarre cat-like friendship and our nonsensical Southern banter. Thank you for taking me under your synthetic chemistry wing and letting me take over your lucky fume hood, it worked. Duy, I would not have made it this far if it weren't for you. You are the go-to guy, always solving problems and helping others out when you are swamped yourself. You have helped me through the entirety of my time here and I cannot express how grateful I am for you. You are a dedicated scientist and an even more dedicated friend. Rahul, I would not have been able to do any of my synthetic work if it weren't for your guidance. I don't know how you did it, but you remained patient with me every step of the way. You taught me what it means to have "good hands" and I'm grateful you have passed those skills down to be. In addition to you being a synthetic genius, you are a kind-spirited person and a true joy to have worked with.

Riley, it has been so great watching you grow as a scientist and as a rock climber. Sharing 387 with you has been absolutely hilarious. I know I can count on you if I need a good laugh session or someone to commiserate with. Thank you for keeping me up to date with the youthful lingo and style of this generation. Aswad, you have been a source of levity in the lab with your great sense of humor and killer dance moves. You and Ananya opened your family to me and invited me to celebrate birthdays and holidays with you all means more than you know. Also, watching Aroh become a 7-year-old adult has been a real treat. Thank you for your generosity. Sahar, you are one of the most caring people I have ever met and have such a kind soul. Since joining the lab, you have brought a sense of compassion that I think everyone benefitted from. Your hardworking nature is admirable. You are going to do great things.

Anam, Namsies, you are a superstar. I have looked up to you since I first joined the lab and I'm pretty sure I've said that to you an embarrassing number of times by now. You are brilliant with big dreams that I know you will accomplish. Life has certainly thrown us some

curveballs, but we have always been there for each other. I was so happy to have found someone who shares the same eclectic, if you will, music interests and it's been so fun going to shows together. As I always say to you: Thank you for being you, don't ever change.

Rebecca, not only are you the best lab mate (self-proclaimed), but the best friend and sister I always wished I had. You have been my partner in crime since day 1 and I could never have asked for a better person to go through the good, the bad, and the ugly with other than you. You have made my graduate school experience one to be treasured. We have had so much fun together going on vacation, camping, escaping fire season, and binge-watching reality TV together, and I'm excited for that to continue (maybe not the fire season part). You have made all the great times greater and the bad times better. Thank you for making grad school "lit" and being my friend for life.

I must also acknowledge my loving partner, Brent Lindquist-Kleissler. Your constant support, patience, and understanding have been my rock throughout this journey. Thank you for being there through the late nights, the tough moments, and the celebrations. Your belief in me has meant the world, and I don't know how I would have done it if you weren't by my side. We've grown so much together during our time in grad school and have shared countless memorable moments—from breaking four bones around my qualifying exam to adopting three 10-lb rabbits. I wouldn't have had it any other way. You make me feel cherished and understood, and you've helped me become a better person. Brent, you inspire me every single day. Thank you for everything you do.

Finally, I would like to express my deepest gratitude to my family, to whom I owe everything. From the very beginning, you all have believed in me and provided the foundation I needed to succeed. To my brothers Christopher and Anthony, our relationships have been a source of strength and comfort. I have looked up to you both as role models and relied on you as support systems throughout my life. Thank you both for always being there for me, for the countless laughs, and for shaping me into the person I am today. To my mother and father,

you have instilled in me the values of hard work, perseverance, and the importance of appreciating the little things in life. It is with your unwavering support that has allowed me to achieve all that I have. Thank you for always believing in me, even when I didn't believe in myself, and for the sacrifices you made that have allowed me to pursue my dreams. I could not have accomplished any of this without you both in my corner. I dedicate this thesis to you. I am who I am because of your unconditional love and support. Thank you, from the bottom of my heart.

Chapter 1

Overview of Non-enzymatic Transformations in Natural Product Biosynthesis

Published in: Bouthillette, L. M.; Aniebok, V.; Colosimo, D. A.; Brumley, D.; MacMillan, J. B.
Nonenzymatic Reactions in Natural Product Formation. *Chemical Reviews* **2022**, 122 (18),
14815-14841.

1.1 Introduction

Biosynthetic mechanisms of natural products primarily depend on systems of protein catalysts. However, within the field of biosynthesis, there are cases in which the inherent chemical reactivity of metabolic intermediates and substrates evades the involvement of enzymes. These reactions are difficult to characterize based on their reactivity and occlusion within the milieu of the cellular environment. We have classified such reactions into four categories: intramolecular, multi-component, tailoring, and light-induced reactions. Intramolecular reactions are one of the most well studied in the context of biomimetic synthesis, consisting of cyclizations and cycloadditions due to the innate reactivity of intermediates. There are two subclasses that make up multi-component reactions, one being homologous multi-component reactions that result in dimeric and pseudodimeric natural products, and the other being heterologous cascading reactions where one or more precursors from independent biosynthetic pathways undergo a variety of transformations to produce the mature natural product. The third type of reactions discussed are tailoring reactions, where post-modifications occur on the natural product after the work of the biosynthetic machinery is completed. The last category consists of light-induced reactions involving ecologically relevant UV light rather than high intensity UV irradiation that is traditionally used in synthetic chemistry. This chapter will review recent non-enzymatic biosynthetic mechanisms and include sources for those reviewed previously.

1.1.1 Influence of biosynthetic studies

The depth of studies to understand how natural products are produced has opened new avenues in the field of natural products from our understanding of chemistry, protein dynamics and function, genetic regulation amongst other scientific disciplines.¹⁻³ These studies can aid

in the development of drugs and chemical tools. While most of the transformations in natural product biosynthesis can be found encoded in the genome and rely on a series of protein catalysts, or enzymes, there is a small but growing number of molecules in which the inherent chemical reactivity of metabolic intermediates and substrates that does not require the involvement of enzymes. These reactions are difficult to characterize based on their reactivity and occlusion within the milieu of the cellular environment. As the community continues to build a strong foundation for how microbes and higher organisms produce natural products, there is a need to understand how protein independent, or non-enzymatic, biosynthetic steps can occur as well.

1.1.2 Foundations of natural product biosynthesis

The anabolic pathways used to derive natural products can be best explained in the context of the central dogma of molecular biology first described by Francis Crick.⁴ Genes are transcribed to produce RNA, which is translated into proteins that perform catalytic functions to covalently link basic metabolic substrates made *de novo* or obtained from the environment.

The importance of biosynthetic gene organization cannot be overstated. Early work in the 1980s by D. A. Hopwood on the actinomycete *Streptomyces coelicolor* A3(2) showed that secondary metabolite production and sporulation, which are both hallmarks of defense, were dependent on each other.⁵ This was one of the earliest pieces of evidence to suggest that microbial defense tactics were a part of a dedicated system intimately tied to each other and the genes involved in any such mechanisms were likely to be spatially clustered on the genome.⁵ We now know that most genes responsible for the production of natural products lie in contiguous sequences known as biosynthetic gene clusters (BGCs) and are frequently accompanied by associated genes involved in active transport, resistance, and regulation.⁶

Advancements in next generation sequencing and increases in transcriptomic capabilities are leading to major gains in the understanding of more complex plant biosynthesis.⁷

The functional units of biosynthetic machinery are the proteins responsible for carrying out the catalytic steps to create new bonds in chemical substrates. Most commonly, after translation, these proteins form large multimeric complexes that serve as localized natural product factories. These systems have been extensively studied *in vitro* and *in vivo* from a diverse range of natural product families. This has built a strong knowledge base for secondary metabolic enzymes. However, unveiling a truly complete picture of this enzymatic machinery is a long and complicated process that will grow in complexity as researchers deviate from canonical protein scaffolds.

These types of studies will cover most of the global secondary metabolome, but there are exceptions that deserve attention. Specifically, the use of abiotic, or non-enzymatic, mechanisms for covalent bond construction is a recently exemplified process in natural product biosynthesis but has been well established in other areas of biology for years.

1.1.3 Non-enzymatic reactions in the cellular milieu

In the packed environment of any cell, there are chemical reactions that may occur in the absence of enzymes due to inherent chemical reactivity and physical proximity. These reactions are known as non-enzymatic reactions and have been studied particularly in the context of primary metabolism. These non-enzymatic reactions in microbes and humans have been thoroughly reviewed in recent years.⁸⁻¹⁰ A biophysical explanation of how their utility is amplified by cellular compartmentalization has also recently been discussed.¹¹ Finally, the Balskus laboratory wrote an intriguing opinion on how to manipulate non-enzymatic reactions within primary metabolism to study the cellular biology of microbes.¹² This work established a

role of non-enzymatic chemistry as an important intermediate in the crosstalk between primary metabolic flux and cellular function.

Gravel and Poupon wrote a micro-review over 10 years ago detailing many non-enzymatic mechanisms in natural product biosynthesis, focused primarily on the justification of these reactions using biomimetic total synthesis.¹³ Since then, a growing body of literature has accumulated, highlighting non-enzymatic steps in secondary metabolism. Not only do these studies allow for trends in substrates and mechanisms to be identified but also beg the question: are these reactions biologically relevant?

1.1.4 Incomplete categorization of non-enzymatic processes in natural products

The history of non-enzymatic chemistry in natural product biosynthesis is convoluted due to the inconsistent terminology and experimental details reported over the greater part of the last century. Commonly, when a new natural product is discovered, a biosynthetic hypothesis is made to support a logical path to the proposed structure. A lack of molecular biology efforts in chemo-centric natural products laboratories have historically prevented a thorough experimental process for publication. This works against the accurate definition of non-enzymatic chemistry in a two-fold manner.

First, authors could describe biosynthetic processes with no precedent or obvious catalyst as being spontaneous without thorough exploration for a possible enzyme. Identification of a racemic mixture of isomers and sensible arrow pushing is commonly used to justify such claims, as these are a necessary first step for exploring non-enzymatic reactions. Reactive intermediates are often unstable and/or produced in low titer, therefore the difficulty of these experiments may deter interrogation.^{14, 15} These technical issues can result in a lack of biosynthetic validation and thus mislabeled non-enzymatic mechanisms.

In direct contrast, mislabeling can occur if authors describe a process as being enzymatically driven without vetting of the proposed BGC. When gene clusters are identified, it is acceptable to predict the role of each gene based on sequence alignment to previously studied biosynthesis. This is a powerful practice that is the conceptual basis for the natural product community to identify BGCs in microbial genomes.¹⁶ However, sequence similarity does not always accurately predict the output of a biosynthetic pathway, as evident by the continued discoveries of unique BGCs.¹⁷ Rigorous studies are needed to confirm BGC predictions, but fortunately this process has been aided by the increased sophistication of scientific tools.

In addition to the lack of experimental validation, misrepresentation has also worked against the accurate description of non-enzymatic processes. The definition of non-enzymatic biosynthesis lies on a razor's edge between a purposeful function and an artifact of scientific intervention. We argue that there is a distinction between non-enzymatic chemistry that is functional in the context of an organism and non-enzymatic chemistry that occurs serendipitously as isolation artifacts created by various extraction methods. While this distinction is difficult to parse, it remains an important characterization of any non-enzymatic reaction regardless of its characterization as artifactual or functional, the identification can further scientific endeavors particularly in the case of medicinal chemistry. There lies a responsibility to explore the biological context without ignoring the chemical potential.

Finally, a fundamental component inhibiting the growth of studying non-enzymatic mechanisms is the disjointed language used to describe said reactions. Inconsistent terms used to label these include abiotic, domino, spontaneous, cascade, concerted, chemoreactive, non-proteogenic, and more. Accurately and wholly identifying potential non-enzymatic mechanisms has thus been difficult. As more examples of non-enzymatically derived molecules are being discovered, we hope to unify the language established in this chapter and used to assure future identification and consolidation.

1.2 Classification of non-enzymatic mechanisms

As the literature of this field continues to mount, it is necessary to define the types of non-enzymatic mechanisms observed in natural products. In Gravel and Poupon's 2008 review of non-enzymatic biosynthesis, they highlighted four main categories of such mechanisms as: intramolecular rearrangement, dimerization, multi-component reactions, and light induced changes.¹³ We have suggested a slightly expanded version in which there are five major categories: 1) intramolecular cycloaddition reactions, 2) homologous multi-component reactions, 3) heterologous cascading reactions, 4) light induced reactions, and 5) tailoring reactions. These categories will be described below using examples of well-studied non-enzymatic reactions, with a focus on recent publications. Categorization is not necessarily mutually exclusive, and exceptions will be noted where a molecule fits multiple classifications.

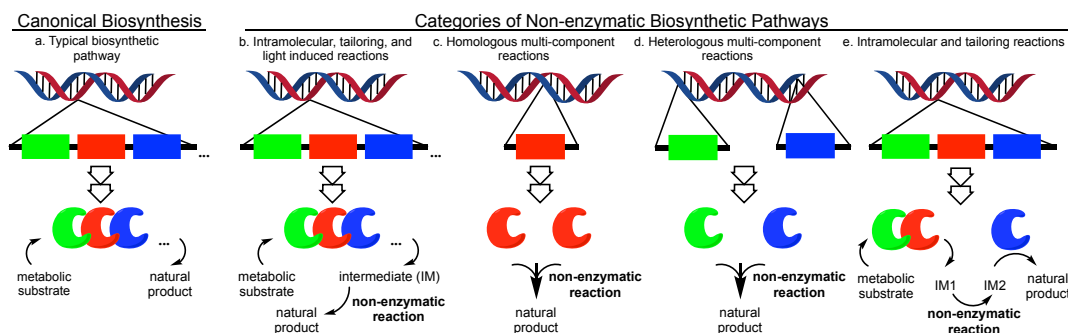


Figure 1.1. Simplified models depicting the types of biosynthetic pathways observed in natural product biosynthesis. a) The canonical biosynthetic pathway where natural products are made entirely from the biosynthetic gene cluster that is encoded in the organism's genome. b) The biosynthetic gene cluster produces an intermediate which undergoes a non-enzymatic transformation to the final natural product. c) The non-enzymatic dimerization of two monomeric natural products. d) Two or more metabolites produced from individual pathways undergo a non-enzymatic reaction to yield the final product. e) The biosynthetic enzymes produce a reactive intermediate that undergoes a non-enzymatic transformation that then reengages in the pathway. This transformation is seen in intramolecular and tailoring reactions.

Figure 1.1 depicts a simplified schematic of the general mechanisms non-enzymatic processes typically follow in natural product biosynthesis. Figure 1.1a shows the canonical mechanism of natural product biosynthesis where the genome encodes for BGCs, giving rise to enzymes that efficiently convert metabolites into the final product. Figure 1.1b is the most straightforward classification of non-enzymatic transformation, whereby a traditional BGC terminates to provide a reactive intermediate, which is further modified in a single transformation to give the final product. The next two mechanisms, Figure 1.1c and 1.1d, arise from a non-enzymatic transformation between two or more genetically encoded natural products leading to a molecule of higher complexity. In Figure 1.1c, two equivalents of the same metabolite participate in the non-enzymatic reaction to give a dimeric or pseudo-dimeric species. In contrast, natural products following the pathway of Figure 1.1d arise from two or more different metabolites, usually from independent pathways, combine to make a complex hybrid structure. These homologous multi-component and heterologous cascading pathways represent the most ubiquitous non-enzymatic reactions to date. The final mechanistic scenario, Figure 1.1e, occurs when an enzymatically derived intermediate undergoes a non-enzymatic reaction to form a new product that subsequently re-engages with the biosynthetic machinery for further transformations. An example would be glycosylation or other post-translational modifications.

1.3 Intramolecular reactions

Intramolecular reactions are often hard to characterize due to the difficulties associated with isolating the putative precursors. However, these are one of the most important types of non-enzymatic mechanisms in natural product biosynthesis in that they have been used extensively in biomimetic total synthesis.¹⁸⁻²⁰

The most well studied intramolecular reactions are cycloadditions between electron-rich π systems commonly found in conjugated polyenes. Even as early as 1980, David Black postulated that due to the racemic mixture of the isolated polyenes, the endiandric acids (**1.1**) (Figure 1.2) potentially arose from so-called “non-enzymatic” reactions.²¹ These hypotheses were validated in K.C. Nicolaou’s synthesis of the endiandric acids and their precursors.²²

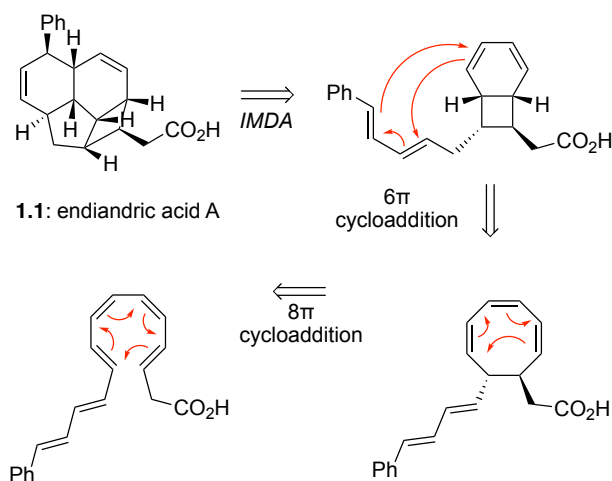


Figure 1.2. The biosynthetic mechanism for the endiandric acids. The mechanism features a series of non-enzymatic of cycloadditions, validated using biomimetic synthesis by K. C. Nicolaou soon after the original hypothesis.²²

The endiandric acids (**1.1**) feature one of the most utilized cyclization reactions in total synthesis of natural products, the [4 + 2] Diels-Alder reaction (Figure 1.2).²³ Several of the works highlighted in the Stocking and Williams review of biomimetic Diels-Alder reactions exist as clear examples of how to distinguish enzymatic processes from their non-enzymatic counterpart.²⁴ Additionally, understanding the distribution of isomers and its reflection on enzymatic involvement are commonly discussed. Knowledge of the role of enzymes in catalyzing pericyclic reactions, including Diels-Alder reactions, has been evolving over the last 10 years, and it is clear that many of these transformations in complex natural products do not

follow a non-enzymatic pathway.²⁵⁻³¹ Through the use of microbial genomics, computational, and experimental (including synthesis and structural biology) approaches, the number of natural product families with characterized Diels-Alderase enzymes has grown with the understanding that Diels-Alderase enzymes can act as functional or templating enzymes.^{32, 33} Importantly, whether the pericyclic reaction under investigation is non-enzymatic or enzymatic, the tools to study the transformations are similar. These and other classical cycloadditions have been reviewed in a synthetic chemistry context by many others, including the Trauner and Nicolaou groups.^{18, 19} Since these reviews, elegant syntheses of natural products have continued to be completed with cyclizations attributed to biomimetic mechanisms, however, not all of these mechanisms have been validated in a biological setting.^{34, 35}

A detailed modern study of non-enzymatic cyclizations features the elansolids, which represent the dichotomy of artifacts and natural products in non-enzymatic biosynthesis. The elansolids are a family of polyketide antibiotics from the gliding bacteria, *Chitinophaga sancti*.³⁶ Initially, alongside the original atropoisomers elansolids A1/A2 (**1.2**), the derivatives elansolid B1 (**1.3**), B2 (**1.4**), and B3 (not shown) were identified in the bacterial fermentation extract (Figure 1.3).³⁷⁻³⁹

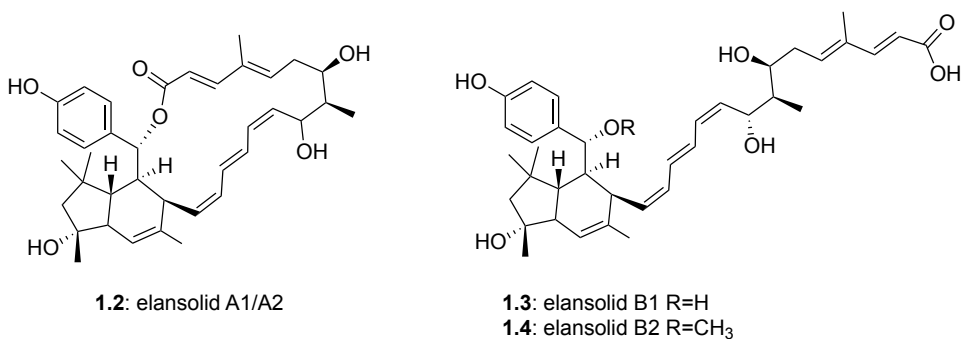


Figure 1.3. The original family of the elansolids.³⁶⁻⁴⁰

In the hunt for the reactive precursor of these artifacts, Jansen and co-workers carried out an extremely careful isolation to avoid contact of the unknown, sensitive intermediate with common nucleophilic agents such as water and methanol.³⁷ The resulting compound was elansolid A3 (**1.7**) and was found in high yield (6.9 mg/L), containing a highly electrophilic para-quinone methide functionality thus giving credence to its transient nature during extraction (Figure 1.4). It was then shown that not only does this intermediate undergo non-enzymatic nucleophilic attack by water and methanol to form elansolid B1 (**1.3**) and B2 (**1.4**), but many nucleophilic amines could be used in a medicinal chemistry effort to construct a library of synthetic analogs with varying antibacterial activity.^{38, 39}

To fully elucidate the biosynthesis of the transient **1.7**, Dehn and co-workers constructed a cosmid gene library of the *Chitinophaga* strain and utilized known polyketide synthase (PKS) sequences to identify putative BGCs.⁴⁰ Consequently, a trans-AT PKS cluster was found whose predicted enzymes matched the known chemical functionalities of the elansolids, particularly in the context of methylation and β -branching. Once the genes within the BGC were defined, there was no putative enzyme identified that would catalyze the final step of the biosynthesis from **1.5** to **1.7** (Figure 1.4). The comparison of the two structures suggested that **1.5** is the precursor that undergoes a series of enzymatic dehydrations to yield **1.6**. Which subsequently participates in a non-enzymatic intramolecular Diels-Alder (IMDA) cycloaddition to form **1.7**.⁴⁰

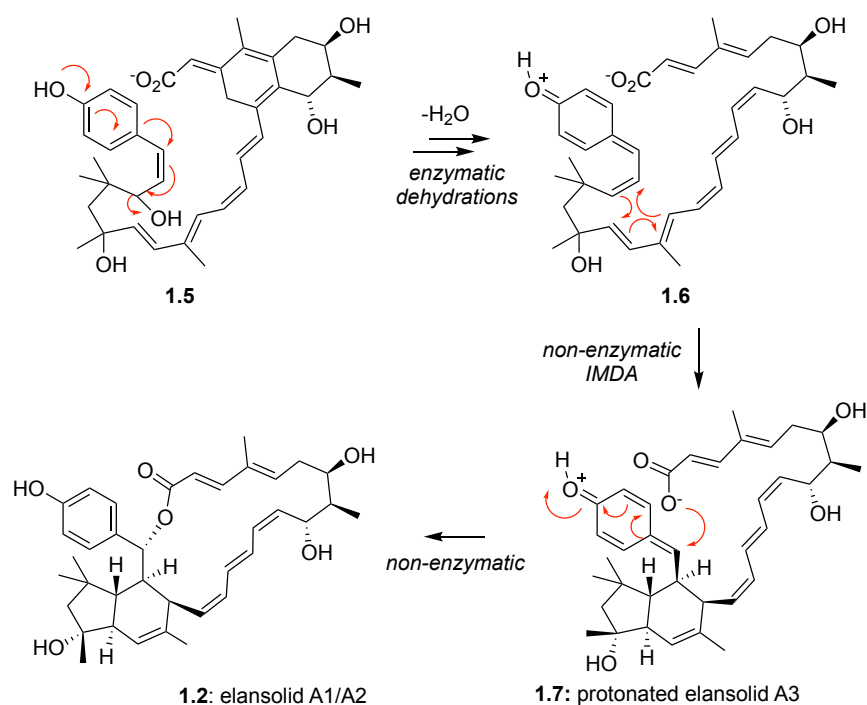


Figure 1.4. The proposed mechanism for elansolid A1/A2. The enzymatic precursor **1.6** undergoes an intramolecular Diels-Alder (IMDA) reaction to provide elansolid A3 that then non-enzymatically cyclizes, giving rise to the final natural product.⁴⁰

Intramolecular reactions are historically the most well studied biomimetic non-enzymatic mechanisms. The cascades involved in these mechanisms are remarkable in their ability to construct complex structures without the need for enzymatic catalysts or lengthy synthetic steps, thus drawing intense interest in the field.²⁰ However, many of these reactions cannot be studied extensively in biological systems due to their inherent reactivity upon formation of the intermediate biosynthetic precursor. The studies rely on synthetic formation of such intermediates, followed by measuring their ability to progress to the following biosynthetic step in moderate conditions. Studies like the elansolids performed by Dehn and co-workers represent a holistic way to approach these types of non-enzymatic reactions that involve a heavy synthetic chemistry effort alongside BGC identification to access transient molecules.⁴⁰

1.4 Multi-component reactions

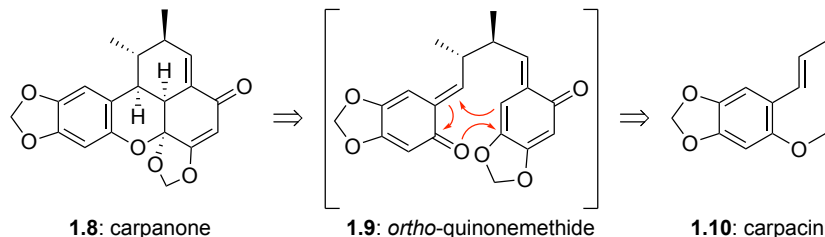
Recently, there has been a notable increase in the number of reported natural product biosynthetic pathways that feature multi-component non-enzymatic reactions. These are characterized by two or more components from enzymatically derived pathways that come together and/or react with simple nucleophiles to form the final natural product. This category is broadly split into two subclasses. Those that involve chemically similar components (dimerizations and pseudo-dimerizations) and those that involve functionally disparate molecules.

1.4.1 Homologous multi-component dimerization reactions

Non-enzymatic homologous multi-component reactions have been prevalent in natural products for many years. One of the earliest works in this category was in 1969 when Brophy and co-workers suggested that the novel lignan carpanone (**1.8**), discovered from a Carpano tree bark extract, might be derived from the dimerization of carpacin (**1.10**) (Figure 1.5a).⁴¹ The dimerization was suggested to occur through a phenolic coupling of **1.10** to generate *ortho*-quinone methide (**1.9**) followed by an IMDA to give **1.8**. In 1971, the Chapman laboratory completed a biomimetic synthesis to investigate this hypothesis.⁴² Researchers used the precursor **1.11** and isomerized it to **1.12**, which is structurally similar to **1.10** (Figure 1.5b). They observed the immediate IMDA product, thus validating the proposed biosynthesis put forth by the Brophy laboratory. Examples of these types of non-enzymatic dimerizations have been previously observed in the case of the bisorbicillinoids, torreyanic acids, stephacidin B, stylissadines, and the complex rugulin bisanthroquinolones, which have been extensively

reviewed but will not be discussed here.⁴³⁻⁵² Rather, this section will focus on modern examples of dimerizations and pseudodimerization.

a. Carpanone biosynthetic hypothesis by Brophy *et al.* after original extraction



b. Dimerization step of carpanone biomimetic synthesis by Chapman *et al.*

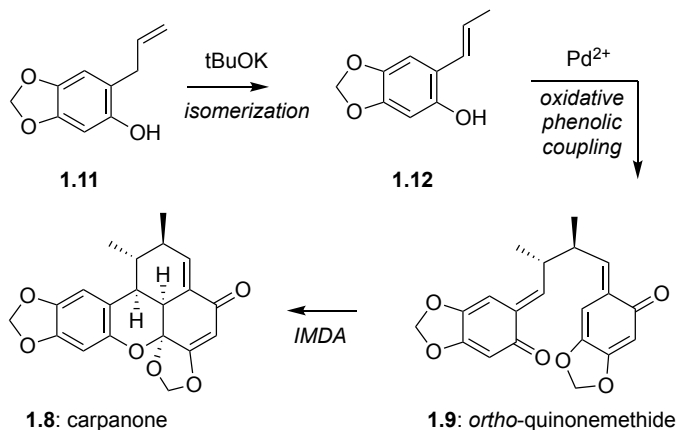


Figure 1.5. Exploration of the carpanone biosynthetic mechanism. a) The original hypothesis for carpanone biosynthesis from carpacin or desmethyl-derivative.⁴¹ b) The biomimetic synthesis used as evidence for the original hypothesized mechanism.⁴²

In 2014, Lin and co-workers isolated a racemic polyketide, oxazinin A (**1.13**) from a novel fungal *Eurotiomycetes* strain from the ascidian *Lissoclinum patella* in Papua New Guinea (Figure 1.6).⁵³ These marine invertebrates have been found to have symbiotic relationships with bacteria and fungi, of which produce structurally unique and biologically relevant secondary metabolites, much like oxazinin A as it possess modest antibacterial activity.⁵³ The biosynthesis of **1.13** was suggested to be a non-enzymatic dimerization between two modified

monomers (**1.16**) (Figure 1.6). The polyketide skeleton is prenylated to form **1.14**, containing a reactive aldehyde that undergoes a condensation reaction with anthranilic acid (**1.15**) to form an electrophilic imine (**1.16**). **1.16** then undergoes a complex cascading dimerization to the final product **1.13**.⁵³

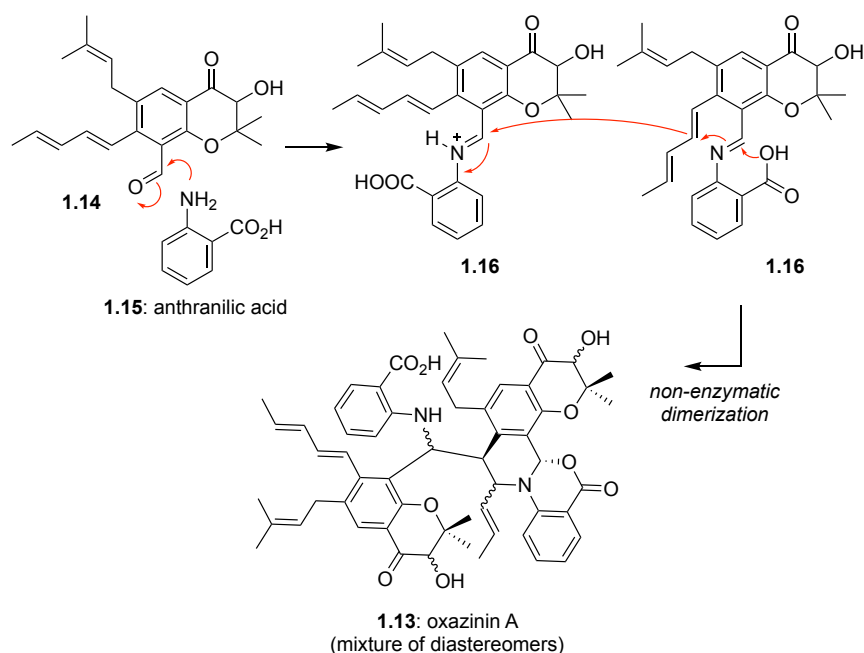


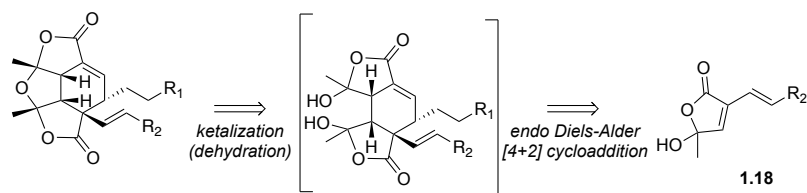
Figure 1.6. The proposed mechanism for the biosynthesis of oxazin A. The prenylated polyketide monomer **1.14** undergoes a condensation reaction with anthranilic acid followed by a final dimerization.⁵³

The pushing of electrons is thought to result in both cyclization of a benzoxazine ring and dimerization of another prenylated polyketide monomer, **1.16**, via an iminium intermediate (Figure 1.6). Suggestion of the biosynthesis occurring through a non-enzymatic dimerization is due to the diastereomeric mixture of **1.13** found in the fungal extract. As prenylated polyketides of similar structure have been isolated before, the reactivity and formation needed to be observed in a system devoid of protein catalysts.^{54, 55} In 2022, Aniebok and Shingare from our

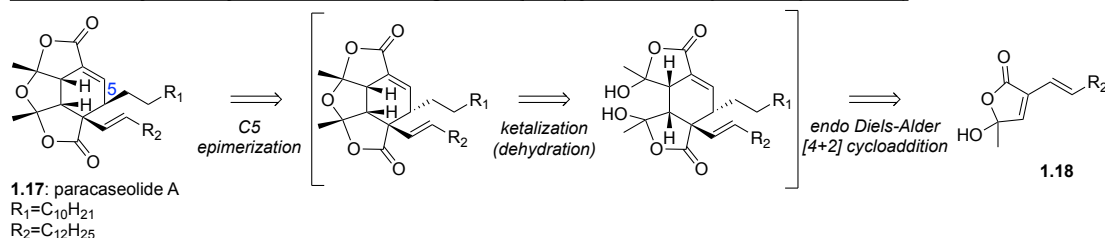
lab completed the first total synthesis of **1.13** and discovered the same 5:1:1:1 diastereomeric ratio in their synthetic approach as the isolates from the fungal fermentation originally reported by Lin and co-workers.^{53, 56} Researchers carried out real-time ^1H - ^{15}N HMBC NMR experiments to observe the stepwise mechanism of formation and confirmed the non-enzymatic dimerization of **1.16** to **1.13**.⁵⁶ The use of ^1H - ^{15}N HMBC NMR studies have shown to be a powerful tool in understanding and validating non-enzymatic pathways in multiple examples, which are discussed further.

One of the fastest moving discoveries of a natural product whose biosynthesis features a non-enzymatic dimerization involves the study of the paracaseolides, which originate from the bark of a *Sonneratia paracaseolaris* Chinese mangrove tree.⁵⁷ Their structure features a unique tetracyclic core comprised of a tetraquinane oxa-cage bislactone with two long chain alkyl appendages (Figure 1.7).⁵⁷ The original biosynthetic hypothesis by Chen and co-workers in 2011 suggested that paracaseolide A (**1.17**) was derived from a Diels-Alder [4 + 2] cycloaddition-based dimerization of the monomeric α -alkylbutenolide, **1.18** (Figure 1.7a).⁵⁷ In 2012, Noutsias and Vassilikogiannakis completed the first total synthesis of **1.17** using a similar biomimetic Diels-Alder as the last step.⁵⁸ In their synthesis, monomeric **1.18** was heated in the absence of solvent to give **1.17** in 59% yield (Figure 1.7b). A year later, Yin and co-workers took a similar approach involving high temperature dimerizations to yield a variety of paracaseolide analogs whose bioactivity as phosphate inhibitors were evaluated.⁵⁹ These along with other explorations of the paracaseolides indicated that the Diels-Alder [4 + 2] cycloaddition butenolides were chemically tractable.^{60, 61}

a. Original discovery and biosynthetic proposal Chen *et al.*



b. Biomimetic synthesis by Noutsias and Vassilakogiannakis (*endo* cycloaddition requires C5 epimerization)



c. Mechanistic studies by Wang and Hove (*exo* cycloaddition requires C7 epimerization)

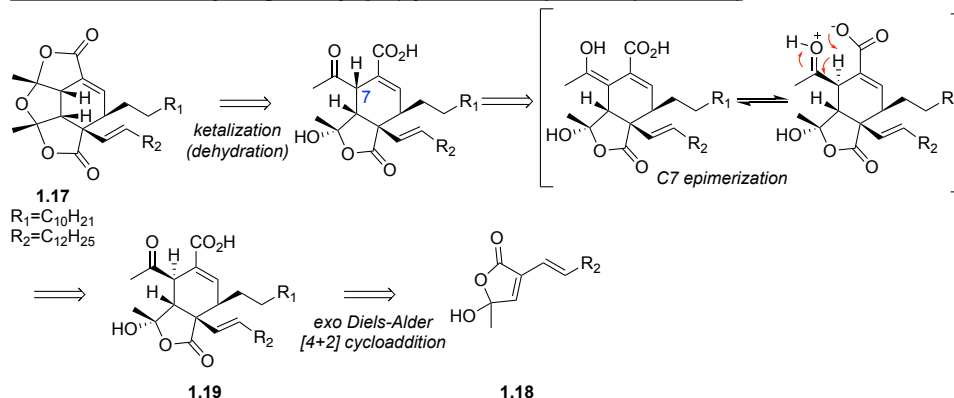


Figure 1.7. Progression of the investigation of paracaseolide A. a) Original proposed biosynthesis of paracaseolide A, which left out the epimerization needed for the correct stereochemistry at C5, based on an *endo*-Diels-Alder cycloaddition.⁵⁷ b) An updated proposal of the biosynthetic pathway of paracaseolide A based on a biomimetic synthesis starting with an *endo* cycloaddition and C5 epimerization.⁵⁸ c) A second proposed biosynthetic pathway based on computational calculations starting with an *exo* cycloaddition and C7 epimerization.⁶³

The bowl-like structure of **1.17** and the *cis* orientation of the alkyl chains has prompted detailed mechanistic studies for this dimerization.^{62, 63} It was first hypothesized that the cycloaddition was a concerted *endo*-Diels-Alder [4 + 2] cycloaddition. However, this reaction would result in the *trans* configuration of the alkyl chains, an isomer not reported by the initial authors.^{57, 61} While an epimerization of these stereocenters could explain their dynamic nature,

several alternate groups participating in total synthetic routes reported that epimerization of the final paracaseolide product was not observed. Vasamsetty and co-workers demonstrated in their 2014 report that the dimerization was non-enantioselective and these diastereomers could not epimerize, an experimental finding supported by DFT calculations.⁶² Further computational evaluation of the energetics of this system suggested a stepwise cycloaddition in contrast to a concerted mechanism. This prediction could explain the weighted isolation of the *cis*-paracaseolide A as a stepwise biradical mechanism would favor this diastereomer by an estimated 2.3 kcal/mol.

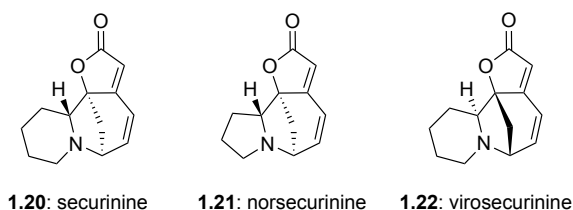
In 2015, Wang and Hoye further explored the non-enzymatic mechanism of **1.17** through various experimental and computational methods.⁶³ The authors initially showed that the non-enzymatic dimerization occurred at lower temperatures than previously reported, thus supporting its feasibility in a biological setting. One of the most important findings was the isolation of the intermediate **1.19** as a result of the Diels-Alder cycloaddition before dehydration to yield **1.17** (Figure 1.7c). This intermediate had been previously predicted but never isolated. Interestingly, the stereochemistry of **1.19** revealed that the cycloaddition occurred with an *exo* orientation of the carbonyl, which is a generally unfavorable mechanism contradicting the initial proposals that were made without experimental evidence. The *exo* configuration suggested by the isolation of intermediate **1.19** (Figure 1.7c) was supported by computational calculation of the energies for each transition state (TS_{endo} vs TS_{exo}) and an argument for a stabilized bis-pericyclic transition state that was made for increased *exo* favorability.⁶³

Previous reports have studied the epimerization of **1.17** at the C5 position, as the proposed *endo* configuration would result in the opposite orientation of the final product.⁶³ Their finding, as previously stated, declared that epimerization was not possible. However, the C5 stereochemistry of **1.19**, assuming an *exo* [4 + 2] addition is set correctly, while the orientation of C7 is opposite to the final product (Figure 1.7c). The authors provided evidence for this novel epimerization using deuterium labeling experiments of **1.19** and a model system in which the

alkyl substitution was simplified. In these experiments, the C7 proton was shown to exchange with deuterium at ambient temperatures quite rapidly (50% conversion in 4 hours).⁶³ The authors suggested this mechanism is due to keto-enol tautomerization that leads to the formation of a zwitterionic carboxylate ion capable of intramolecularly abstracting the C7 proton. The result of this abstraction would be an enol olefin capable of protonation before the final dehydration step. This finding was key as it demonstrated there is a feasible non-enzymatic mechanism for generating a planar environment at the C7 position, amenable to an obligate step for the preferred stereochemistry of **1.17**. Finally, the catalytic function of the carboxylate ion was supported by the decreased yields (35% conversion in 10 days) when esterified substrates were used.⁶³

The mechanistic studies performed by Wang and Hoyer exemplify the detail needed to study multistep non-enzymatic biosynthetic reactions. In addition to these findings, the authors raised more questions that require attention. Most importantly, they showed that a monomeric butanolide was irreversibly opened, containing a favorable dienophile with two electron withdrawing groups, could not only undergo dimerization, but could do so much more efficiently than its closed form.⁶³ Even though the exact nature of the biological dienophile remains to be determined, the work done by the Hoyer group and other synthetic chemists over the past 10 years is an impressive feat that has brought a necessary level of experimentation to a preliminary hypothesis of non-enzymatic biosynthesis. Where these studies lack is the biological exploration of the nature of these molecules in the context of the Chinese mangrove, which will be even more informative.

Examples of monomeric securinega alkaloids



Examples of dimeric securinega alkaloids

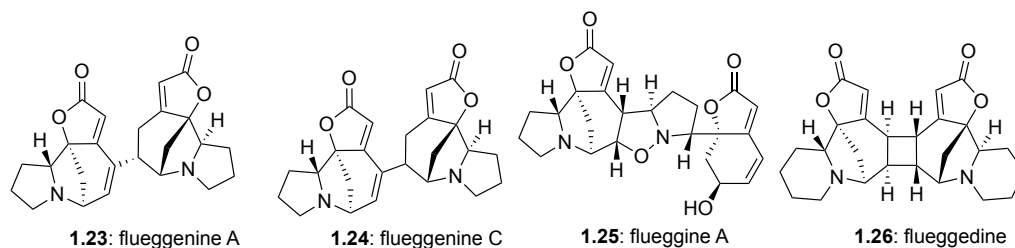


Figure 1.8. Representative family members of the securinega alkaloids which include both monomeric and dimeric structures.⁶⁴⁻⁷⁰

The alkaloid natural product, securinine (**1.20**), was isolated from the leaves of a *Securinega suffructicosa* plant in 1956 (Figure 1.8).⁶⁴ Since its discovery, over 60 related natural products have been discovered from 5 different genera of related *Euphorbiaceae* sp.⁶⁵ These compounds are characterized by their bridged tetracyclic core consisting of an unsaturated 5-membered lactone ring and either pyrrolidine or piperidine rings, as seen in norsecurinine (**1.21**) or virosecurinine (**1.22**), respectively (Figure 1.8). The chemical diversity of the securinega alkaloids expanded exponentially after 2006, as dimeric, trimeric, and higher order oligomeric derivatizations have been discovered.⁶⁶ The complicated structures of the securinega monomers consisting of multiple substitutions, compact ring structures, and short chain bridges were confounded even further in other dimeric analogs.⁶⁶⁻⁶⁹ These dimers included unsymmetrical structures as seen in flueggenine A-I and flueggine A-B, as well as bilaterally symmetrical structures seen in flueggedine in Figure 1.8 (not all listed are shown).

The hypothesis that these oligomeric molecules could be constructed via non-enzymatic mechanisms began when Gan and co-workers first isolated flueggenines A (**1.23**)

and B (**1.24**) from the roots of a *Flueggea virosa* plant (Figure 1.9).⁶⁷ The extract's most abundant metabolite, norsecurinine (**1.21**), was proposed to be a biosynthetic precursor to the flueggenines. The authors suggested an intramolecular vinylogous Morita-Baylis-Hillman (MBH) reactions, also known as the Rauhut-Currier (RC) reaction, as a mechanistic rationale for this dimerization.⁶⁷

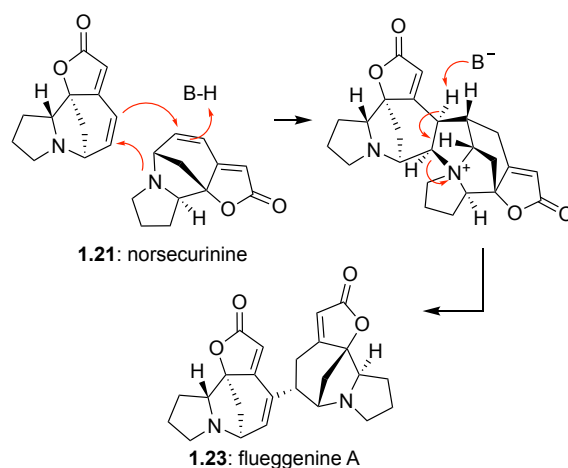


Figure 1.9. Original proposal for the non-enzymatic biosynthetic mechanism of flueggenine A from a co-isolated molecule, norsecurinine.⁶⁷

A conventional MBH reaction involves an aldehyde as the ultimate electrophilic acceptor of the anionic carbon, whereas in the vinylogous MBH reaction, or RC, a second Michael acceptor is the carbon source. While RC reactions typically involve phosphine catalysts, the classification of the securinega alkaloids as having an RC-like mechanism is more appropriate due to the identity of the electrophile.^{66, 67} Unlike the conventional RC reaction, where a nucleophile catalyst is used, three necessary components (activated alkene, electrophilic carbon, and nucleophilic amine) are all contained in the securinega monomers.⁶⁶⁻
⁶⁸ This provides a self-catalyzed system, as seen in several other non-enzymatic mechanisms discussed later (Figure 1.9).

The same laboratory later proposed in 2012 that there should be a possibility for continued oligomerization due to the regeneration of the nitrogen catalyst and an abundance of unhindered enone systems after putative dimerizations.⁶⁹ Three new trimeric securinega compounds, including flueviroisine A (**1.27**), were discovered whose oligomer state supported the author's hypothesis regarding repetitive RC reactions (Figure 1.10a). Stirring of the monomers in organic solvent did not lead to oligomerization to the flueviroisines or flueggenines, suggesting the necessity of an enzyme or a particular chemical environment, although these studies were limited.^{67, 69}

Attempts of *in vitro* securinega dimerization

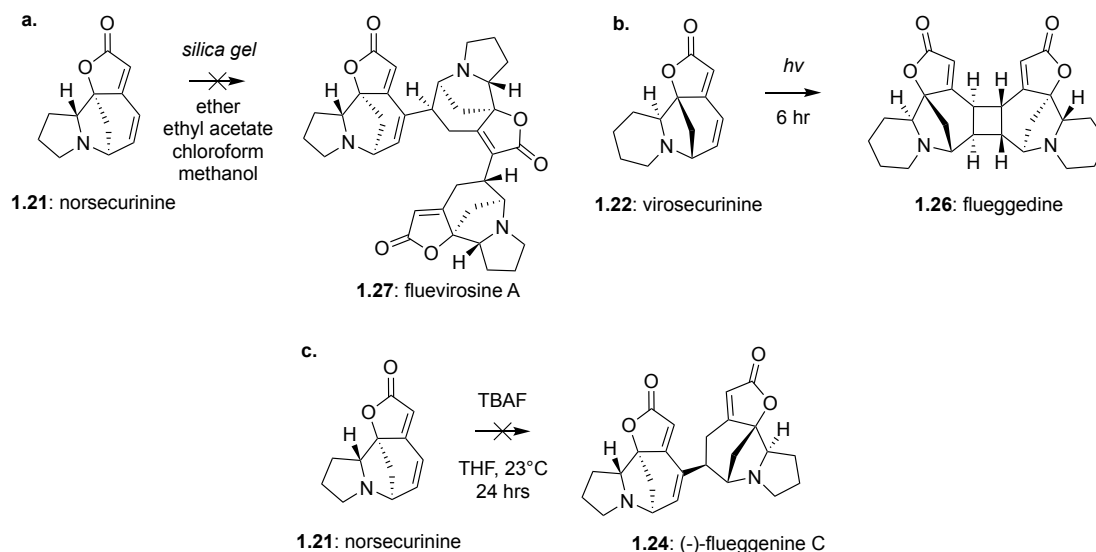


Figure 1.10. Synthetic dimerization attempts of securinega monomers. a) Stirring the norsecurinine monomers in organic solvents did not provide the oligomerized product.⁶⁹ b) Virosecurinine dimerized to flueggedine in low yields after 6 hours under UV light 350 nm.⁷⁰ c) Additional unsuccessful attempt at dimerization of norsecurinine in organic solvents.⁶⁶

Zhao and co-workers reported a new securinega dimer, flueggedine (**1.26**), which contained a novel cyclobutane ring connecting the two identical virosecurinine (**1.22**) subunits (Figure 1.10b).⁷⁰ The authors proposed that the cyclobutane ring was constructed via a [2 + 2]

cycloaddition, a hypothesis tested by irradiating **1.22** for several hours at room temperature under UV light (350 nm). The authors saw formation of **1.26**, albeit in low yield (5% over 6 hours), suggesting that this UV-based cycloaddition could occur. More so, the stereochemistry of **1.26** suggested that the cycloaddition occurred predominantly in the *exo* mode, yielding a caged structure.⁷⁰ This was the first experimental evidence for the dimerization of securinega monomers under biological conditions. Interestingly, mass spectrometry of both leaves and roots of the *F. virosa* plant demonstrated that monomer **1.22** was isolated from both sites, but dimeric **1.26** was only isolated in the leaves suggesting that the UV-induced dimerization seen *in vitro* may be relevant during conventional plant growth.⁷⁰ While a total synthetic route to flueggine A (**1.25**) (Figure 1.8) had been completed, no biomimetic studies were performed with the proposed RC reaction before the recent publication by Jeon and Han in 2017.⁶⁶ In their total synthetic effort to (-)-flueggene C (**1.24**), Jeon and Han incorporated a biomimetic RC step early in their route, whereas biomimetic coupling is used as terminal steps in the majority of these syntheses.⁶⁶ The authors constructed a model system of the proposed biosynthesis that differed quite drastically from endogenous substrates in the hopes of making the reaction more efficient. After optimizing the reaction conditions for their model substrate, the authors tried to dimerize norsecurinine (**1.21**) using TBAF as a base in THF at ambient temperatures with no success after 24 hours (Figure 1.10c). This agrees with Gan and co-workers who found no dimerization of **1.21** in various organic solvents.^{66, 67}

The studies of the flueggenines and related compounds introduce many interesting questions about their biosynthesis and the role of non-enzymatic reactions within. First, Zhao showed that flueggedine (**1.26**) dimerization could occur non-enzymatically under UV light in low yields *in vitro* (Figure 1.10b).⁶⁸ However, the proposed [2 + 2] cycloaddition is dramatically different from the suggested mechanisms in every other example of the 16 known securinega dimers, so conclusions from this report must then be tempered. The value of these studies is the demonstration of the RC reaction of the model dienone system can occur in the absence

of enzymes. Secondly, two groups have shown that mixing the monomers in organic solvents does not lead to dimerization.^{66, 67} This could be due to the fact that the RC mechanism requires enzymes, or it could be that it requires a non-proteogenic promoter that would be endogenous to the plant environment such as UV or an oxidation treatment, neither of which were tested for the norsecurinine monomer **1.21**.⁶⁶⁻⁷¹ Relatedly, a proposed mechanism for the formation of the dimers flueggine A (**1.25**) and B suggests that exogenous oxidation is required for a cyclization cascade.^{70, 71} Regardless of the conditions, it is likely that the non-enzymatic RC reaction will be inefficient as it is known to be of limited reactivity in a flask due to the transient nature of the nucleophilic trigger. The story of securinega oligomerization is still developing, however, the chemical complexity and diversity of these molecules raises much interest in their biosynthetic origins.

Recently, Mevers and co-workers isolated the complex fungal-derived natural product, (±)-homodimericin A (**1.29**), from a *Trichoderma* fungal strain (Figure 1.11).⁷² The elucidation of **1.29** pushed the limits of modern NMR spectroscopy by being extremely carbon-rich and isolated as a racemate. Due to structural similarity, **1.29** appears to be a complex dimeric derivative of the hydroquinone monomer **1.28**, and because of its racemic nature, authors suggest the biosynthetic dimerization to be non-enzymatic under oxidative conditions. However, no appreciable amount of **1.28** was isolated to validate that hypothesis.⁷²

(±)-Homodimericin A (**1.29**) is similar in structure to the bisorbicillinoids (**1.32-1.34**), which are a well-studied family of fungal natural products that are derived from various dimerizations of the sorbicillinoids (**1.30** and **1.31**) (Figure 1.11a).⁷³ Although many sorbicillinoid analogs have been discovered, they are proposed to be derived from a single common intermediate, sorbicillinol (**1.31**). The exact identity of **1.31** was unknown for 20 years after the discovery of the original dimeric species due to its extreme, inherent reactivity,^{73, 74} which may explain the inability to isolate the putative monomeric starting unit, **1.28**, by Mevers and co-workers.

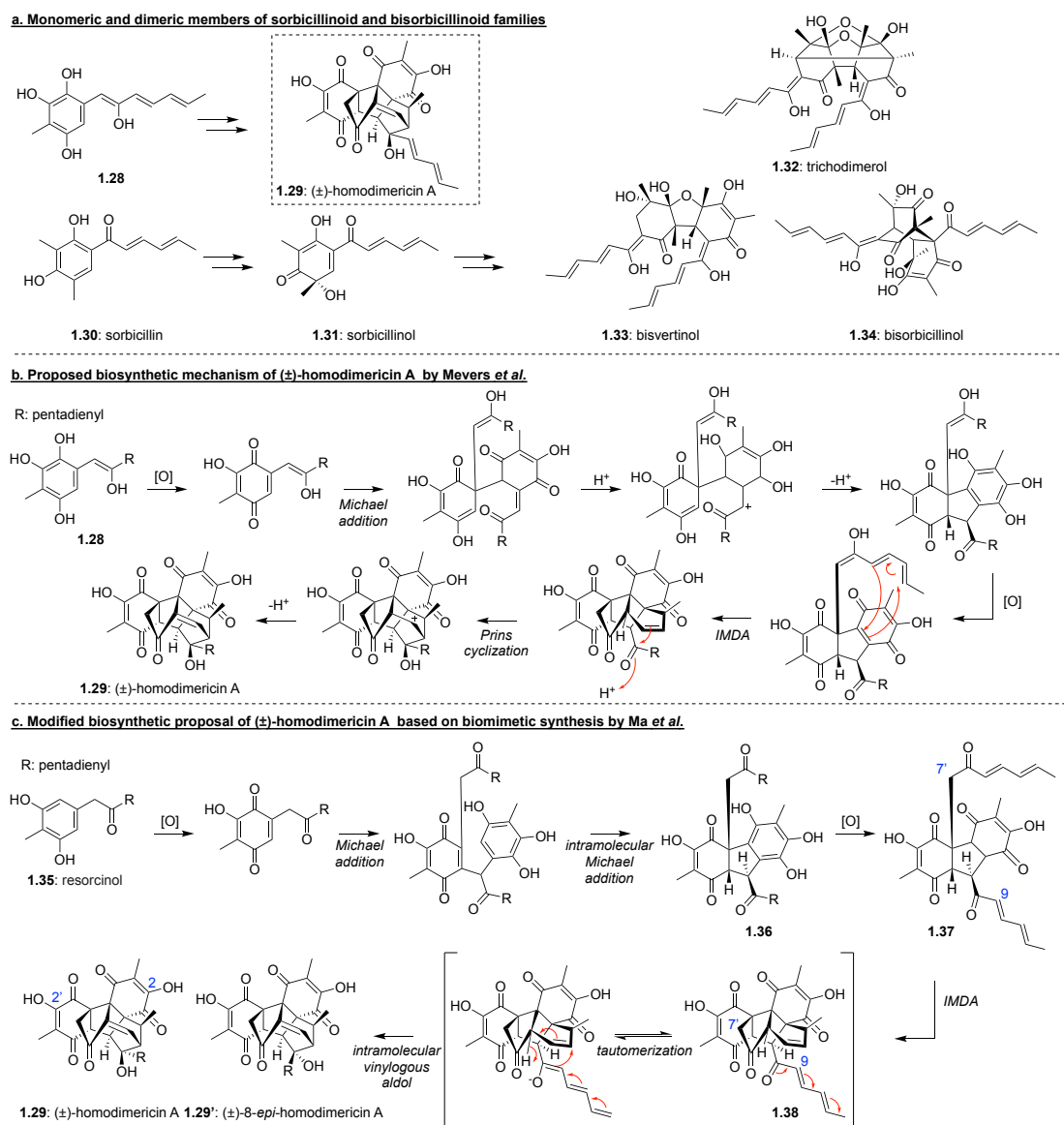


Figure 1.11. Biosynthetic mechanism of (±)-homodimericin A. a) Representation of the monomeric and dimeric members of the sorbicillinoid and bisorbicillinoid families. b) Proposed mechanism for the non-enzymatic biosynthesis of (±)-homodimericin A based on similar pseudo-dimerizations found in the bisorbicillinoids.⁷²⁻⁷⁴ c) Modified biosynthetic mechanism.⁷⁵

There have been multiple non-enzymatic mechanisms suggesting **1.31**, derived from **1.30**, as the precursor to three different bisorbicillinoid skeletons, represented by trichodimerol (**1.32**), bisvertinol (**1.33**), and bisorbicillinol (**1.34**) which are thoroughly discussed in Gravel and

Poupon's 2008 review (Figure 1.11a).¹³ The mechanism put forth by Mevers and co-workers for **1.29** is a chimera of the mechanism suggested for bisvertinol/trichodimerol and bisorbicillinol, with a few key differences.^{13, 72-74} The initial steps in the proposed biosynthesis of **1.29** features a Michael addition to form the initial C-C bond, similar to the mechanism of **1.32** and **1.33** (Figure 1.11b). An intramolecular Diels-Alder cycloaddition proposed later in the biosynthesis resembles the IMDA suggested for **1.34**, however, this cycloaddition is distinct with respect to the location of the diene. The conjugated polyene tail of **1.29** is involved in the [4 + 2] cycloaddition, whereas in contrast, the two quinone-based monomers are responsible for the proposed cycloaddition in **1.34**. Lastly, authors hypothesized a Prins cyclization as the final ring-forming step, which was not proposed for the bisorbicillinoids (Figure 1.11b).^{13, 72-74}

The structural complexity of (±)-homodimericin A makes the biosynthetic hypothesis more intricate. In 2017, Ma and co-workers completed the first biomimetic synthesis for **1.29** and slightly modified the proposed mechanism to incorporate an intramolecular vinylogous aldol reaction of **1.38** to **1.29**, rather than the Prins cyclization.⁷⁵ In their 7-step synthesis using resorcinol (**1.35**) as the starting monomer, they achieved a 26% yield of **1.29** and a 31% yield of the stereoisomer (±)-8-*epi*-homodimericin A (**1.29'**) (Figure 1.11c). During the structural assignment of **1.29'**, the side chain C9-C13 from **1.37** to **1.29'** maintained its diene moiety whereas side chain C7'-C13' was found to undergo a series of cyclization reactions, indicative of an intramolecular vinylogous aldol reaction with an IMDA.^{72, 75} Additional biomimetic syntheses from the Gong and Tang laboratories were published in 2017 where they used DFT calculations to support the energetically favorable double Michael addition and subsequent IMDA cascading reaction proposed by Ma and co-workers.^{72, 76, 77} The Tang laboratory used methoxy protection groups at the C2 and C2' for their synthetic strategy and were able to obtain crystal structures of the methoxy derivatives of **1.36**, **1.37**, and **1.29**, which confirmed the original structure of **1.29** from Mevers' work.⁷⁷

Several strains of fungi produce (±)-homodimericin A (**1.29**), however, one particular strain only produced it when co-cultured with a *Streptomyces* bacterium known to secrete the bafilomycin family of anti-fungals and pure bafilomycin C1 induced the production of **1.29** in this strain. Most intriguing is the fact that **1.29** does not possess anti-fungal or antibacterial activity but is strictly induced in the *Trichoderma* fungus by the presence of bafilomycin C1.⁷² Mevers and co-workers suggest this might be due to the mechanism of action of the bafilomycin family, which are known eukaryotic inhibitors of vacuolar ATPase, an effect linked to oxidative stress.^{78, 79} A highly oxidative environment has shown to be crucial for bisorbicillinoid dimerization. Additionally, a concerted effort by Abe and co-workers in three separate publications have suggested the bisorbicillinoids are effective radical scavengers.^{43, 80, 81} However, why other members of the bafilomycin family that have the same effect on eukaryotic machinery do not induce the production of (±)-homodimericin A remains a mystery. Perhaps this difference is explained by varying cell wall permeability or solubility. Whether these homodimericin products are the result of oxidative stress caused by the anti-fungal effect of the bafilomycins remains to be determined, but the experimental induction is convincing that these compounds are produced as a biological response to bacterial challenge.

The bohemamines are a family of pyrrolizidine alkaloids, a large and structurally diverse array of over 650 molecules with similar heterocyclic structures.⁸² In 2016, our laboratory published the first known example of bohemamine dimers, dibohemamine B (**1.40**) and C (**1.41**), isolated from a *Streptomyces spinoverrucosus* species (Figure 1.12).⁸³ Interestingly, these were not true dimers, as an extra methylene carbon was found to be bridging the two monomeric structures. These types of dimers, where two symmetric units connected by sp²-sp² bonds via a single methylene bridge, have been reported from various sources but primarily from plants.^{60, 84-90} Figure 1.12 depicts only a few of the many pseudo-dimers found with similar connectivity through a methylene bridge. Several of their monomers have been explored biosynthetically, but the dimerization process remains speculative.

Numerous groups have utilized synthetic strategies under mild conditions to demonstrate that the reaction can occur, giving credence to this non-enzymatic mechanism in a biological setting.⁹¹⁻⁹⁴

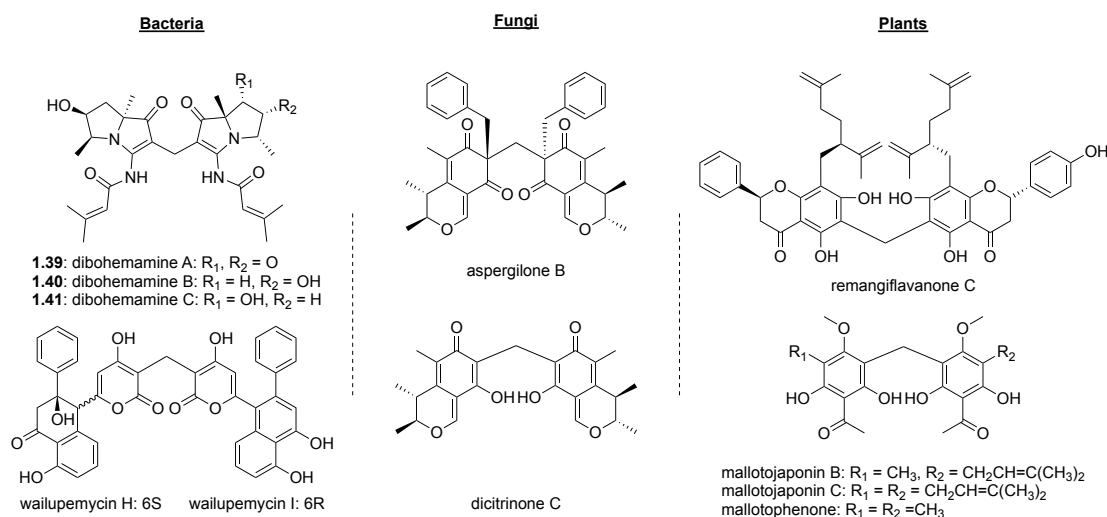


Figure 1.12. A sample of natural products from diverse sources that feature dimerization through a methylene bridge.⁸³⁻⁹⁰

The utilization of a C1 unit in secondary metabolism is rather limited, as these building blocks are commonly incorporated as terminal modifications via SAM-based methylation. Formaldehyde is a common metabolic by-product, and it was hypothesized that excess formaldehyde could be responsible for this dimerization.⁹⁵ The reaction of bohemamine A (**1.42**) with formaldehyde (**1.43**) in THF led to the formation of two products, the hydroxyl derivative **1.44** and diboheamine A (**1.39**) (Figure 1.13).⁸³ Figure 1.13 depicts the non-enzymatic dimerization where the enol resonance structure of **1.42** could participate in a nucleophilic addition with **1.43** to generate **1.44**. Under acidic conditions, **1.44** could then be protonated to generate a reasonable electrophile for a second intramolecular nucleophilic attack from **1.42**. Isolation of **1.44** and exposure to acid produced the dimeric **1.39** within an

hour, suggesting that acidification is essential, and that the original formaldehyde condensation is reversible. When exploring different aryl and alkyl aldehydes, it was shown that the dimerization was robust to aryl substitution under normal atmosphere, but the alkyl aldehydes led to facile oxidation unless performed under nitrogen.⁸³ These experiments suggest that the mechanism of dimerization is tuned to C1 metabolism.

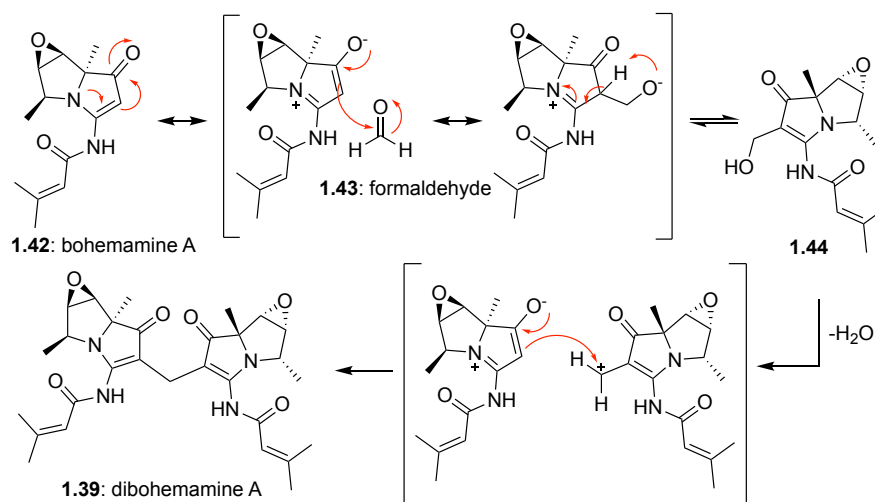


Figure 1.13. Biosynthetic proposal for the dimerization of bohemiaamine A to form dibohemiaamine A.⁸³

By using *O*-(2,3,4,5-pentafluorobenzyl)hydroxylamine (PFBHA) to derivatize aldehydes and ketones existing in trace amounts, it allowed for the quantification of extracellular formaldehyde in dibohemiaamine producing bacteria, which was found to be four times higher than in similar non-dibohemiaamine producing *Streptomyces* strains.⁸³ As formaldehyde is highly diffusible across the biological membrane, this result suggests that internal levels of formaldehyde were also elevated. Formaldehyde is toxic due to its highly electrophilic nature resulting in cross-linking of macromolecular structures within the cell, particularly after combining with proteogenic thiols and amines to form thiazolidine and

electrophilic imines, respectfully.⁹⁵ Furthermore, formaldehyde levels are known to be dependent on metabolic alterations, particularly as recent studies have shown that flux through the pentose phosphate pathway results in excess accumulation of the toxic C1 metabolite.⁹⁵ This has challenged canonical views that mainly methanotrophs produce formaldehyde as a product of methanol fixation. However, the metabolic data is sensible as genetic evidence of formaldehyde in bacteria is known to occur via four major mechanisms in order of abundance 1) thiol-based systems, 2) the ribose monophosphate or RuMP system, 3) the pterin-dependent system, and 4) direct oxidation system.⁹⁴ In the thiol and pterin systems, non-enzymatic reactions between sulfur and nitrogen nucleophiles and formaldehyde have been shown to trigger downstream reactions. It is enticing to ascribe the use of these highly reactive monomeric subunits found in bacteria, fungi, and plants as detoxification mechanisms of endogenously high levels of formaldehyde.⁹⁵

In conclusion, the examples provided here demonstrate that homologous multi-component reactions are an important set of non-enzymatic reactions in biosynthesis. In addition to the classic examples of cyclizations, these new studies represent new discoveries in carbon and nitrogen-based nucleophiles catalyzing cascades of varying intramolecular non-enzymatic reactions. Most importantly, the last three dimerization stories of flueggeidine (**1.25**) (Figure 1.8), (±)-homodimericin A (**1.29**) (Figure 1.11), and the dibohemamines (**1.39- 1.41**) (Figure 1.12) have generated many questions about the potential biological and ecological function of these reactive molecules.

1.4.2 Heterologous cascading reactions

An important addition to the field of non-enzymatic biosynthesis over the last 5 years has been the discovery of more than five families of compounds that are postulated or proven to be the

result of heterologous cascading reactions. These are characterized by the formation of multiple covalent bonds between distinct substrates and these studies have greatly expanded the complexity and chemical diversity on non-enzymatic mechanisms.

The jadomycins are a family of polyketide antibiotics produced by a *Streptomyces venezuela* strain under stressed conditions such as organic solvent exposure and heat.⁹⁶⁻⁹⁸ Through both genetic sequencing of the BGC of the jadomycins and biomimetic synthesis, it has been previously suggested that the biosynthesis of jadomycin A (**1.45**) involves a canonical polyketide extension of malonate subunits to form the intermediate, UWM6 (**1.46**) (Figure 1.14).⁹⁸⁻¹⁰¹ **1.46** can be oxidatively cleaved to form the aldehyde in **1.47**, which can non-enzymatically combine with L-isoleucine (**1.48**) via nucleophilic addition to afford the imine intermediate **1.49**. Intermediate **1.49** then reacts with the nucleophilic carboxylate of the amino acid substrate to trigger a Michael addition into the attached quinone to cyclize oxidative decarboxylation. Biosynthetic studies of this molecule have used homologous recombination and mutagenesis to show that no enzyme involvement is needed for the incorporation of L-isoleucine.⁹⁸

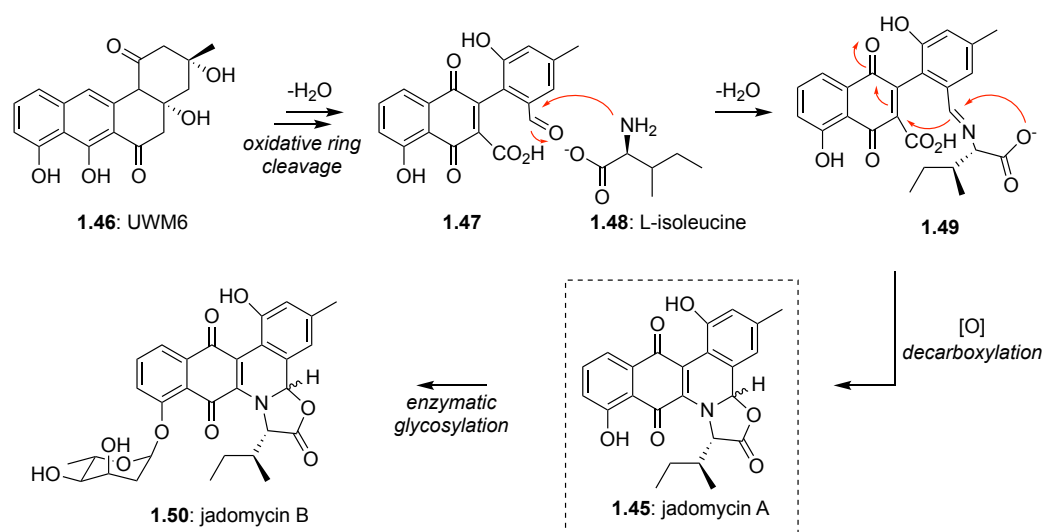


Figure 1.14. Mechanism for the non-enzymatic biosynthesis of the jadomycins. A unique amino acid addition of L-isoleucine (**1.48**) followed by a rearrangement with protein machinery.¹⁰⁰

The groups of Jakeman and Rohr have harnessed the reactivity of intermediate **1.49** to incorporate a slew of amino acids, whose common primary amine serves as a potent nucleophile.⁹⁵⁻¹⁰⁵ For these reactions, amines that lack the spatial proximity of a carboxylic acid found in amino acids would hypothetically not result in the cycloaddition of the full jadomycins. While experiments such as these would further support the non-enzymatic mechanism, they have not been performed. Additionally, intercepting the reactive iminium group of intermediate **1.49** with alternative nucleophiles would help determine how the proximity of the carboxylic acid effects the reactivity of the imine. That is, do proximally located aryl/alkyl primary amines and carboxylic acids have privileged reactivity compared to other nucleophiles.^{101, 103} This question is intriguing as the examples in which these substrates (anthranilic acid and amino acids) constitute the majority of the known non-enzymatic biosynthetic mechanisms.

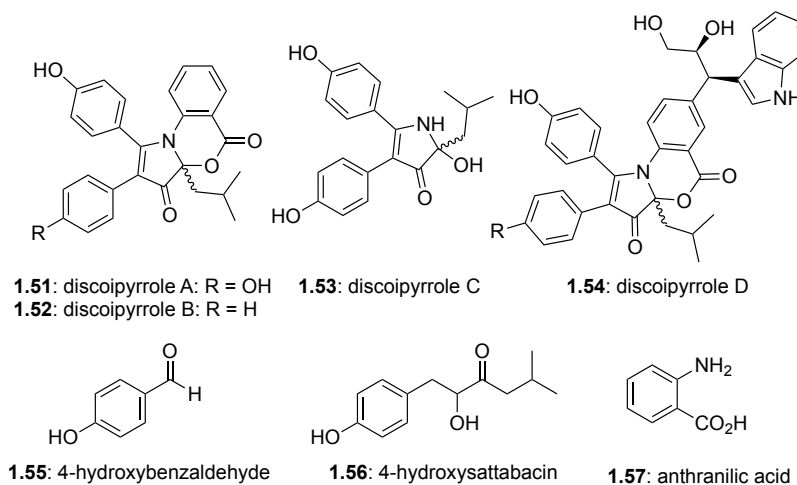
Interestingly, jadomycin A (**1.45**) is the only known analog that is not glycosylated and it is suggested that this glycosylation occurs after the non-enzymatic incorporation of L-isoleucine to produce jadomycin B (**1.50**) (Figure 1.14).¹⁰³ This is a rarity in non-enzymatic

biosynthesis, where it is more common for these non-enzymatic transformations to occur on fully matured natural products, or those that no longer required enzymatic involvement. Furthermore, aglycone jadomycins produced naturally or unnaturally do not retain the same level of bactericidal activity as other jadomycin analogs.^{98, 99, 102} This suggests that the non-enzymatic incorporation of amino acids, which is necessary for said glycosylation, is required for environmental fitness.

The discoipyrrole (**1.51-1.54**) family of natural products were discovered by researchers in our laboratory from the fermentation extract of a marine *Bacillus hunanensis* bacterial strain (Figure 1.15a).^{14, 106} During the isolation of these novel compounds, the simple substrates 4-hydroxybenzaldehyde (**1.55**), 4-hydroxysattabacin (**1.56**), and anthranilic acid (**1.57**) were abundant in the extract as well.¹⁰⁶ The structures of these compounds are similar to the discoipyrroles, particularly discoipyrrole A (**1.51**), and it was postulated that they were involved in the biosynthesis of the final natural products. This hypothesis was confirmed by separate isotopic labeling studies carried out by feeding in commercially available ¹³C-4-hydroxybenzaldehyde. ¹⁵N-anthranilate, as well as synthetically derived ¹³C-4-hydroxysattabacin into the fermentation. It was originally thought that secreted proteins were responsible for the production of the discoipyrroles.¹⁴

To test the involvement of proteins, pure substrates **1.55-1.57** were incubated in “spent” media from a 7-day culture of an unrelated *Bacillus* bacterium, SNE-038, that had been boiled and filtered through a 0.2 μm filter. Discoipyrrole A (**1.51**) was indeed formed from this incubation, as detected by mass spectrometry shown in Figure 1.15b, suggesting a non-enzymatic biosynthetic origin.¹⁴ Great detail was taken to find mild biomimetic conditions in organic solvent that would allow for high yield (up to 80%) of a synthetic discoipyrrole and permit mechanistic studies of this reaction.

a. Compounds isolated from *Bacillus hunanensis* strain SNA-048



b. *In vitro* experiments demonstrate conversion in aqueous media

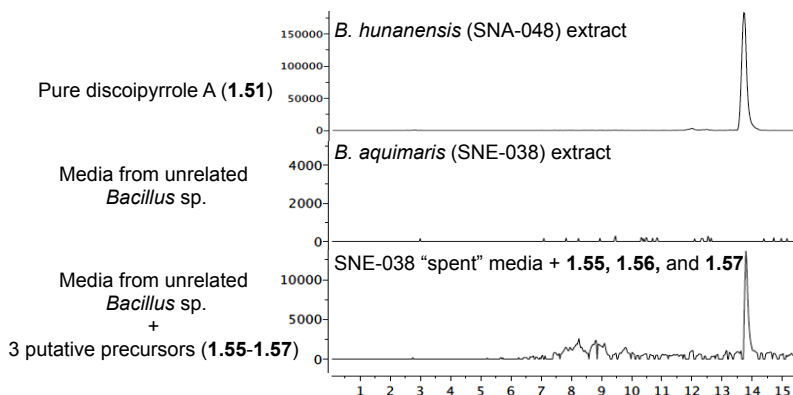


Figure 1.15. Interrogation of the discoipyrrole family and their origin. a) The family of discoipyrroles and their biosynthetic substrates. b) Mass spectrometry identification of the presence of discoipyrrole A in *Bacillus* strain SNA-048, absence in *Bacillus* strain SNE-038, and the appearance of discoipyrrole A in the "spent" media of SNA-038 after the addition of the precursors 4-hydroxybenzaldehyde, 4-hydroxysattabacin, and anthranilic acid.^{14, 102}

The summary of our findings is depicted in Figure 1.16, where anthranilic acid (**1.57**) undergoes a nucleophilic addition into the oxidized sattabacin diketone, **1.58**. The oxidation of 4-hydroxysattabacin (**1.56**) to **1.58** is thought to increase the acidity of the α -proton, allowing for its abstraction. The formed enol, **1.59**, then participates in a Claisen condensation into 4-hydroxybenzaldehyde (**1.55**) followed by a dehydration and nucleophilic attack by the

secondary amine derivative of anthranilate (Figure 1.16).^{14, 106} Details of the mechanism of formation were determined via NMR studies, allowing for the detection of intermediates **1.61** and **1.63**. A biomimetic synthesis was carried out under a strict atmosphere of $^{18}\text{O}_2$ to confirm this hypothesis in combination with mass spectrometry.¹⁴ After oxidation, the final form of the discoipyrroles is thought to exist in equilibrium between a lactone open (**1.63**) and closed form (**1.51**). This is a result of the free electrons of the tri-substituted amine, **1.63**, pushing into the heterocycle to form an iminium ion (**1.64**) and induce the loss of water. The iminium of **1.64** is open to nucleophilic addition by the carboxylate to form **1.51**. Isolation of each enantiomer of **1.51** via chiral high performance liquid chromatography and stirring in water results in the racemic form after 18 hours.¹⁴ The rate of this equilibrium is dictated by the electronics of the anthranilic ring structure in that more electron withdrawing aryl substitutions decrease the nucleophilicity of the carboxylic acid and stabilize the open ring form, **1.64**, a revelation from an unpublished medicinal chemistry effort that has generated over 120 analogs of the discoipyrroles.

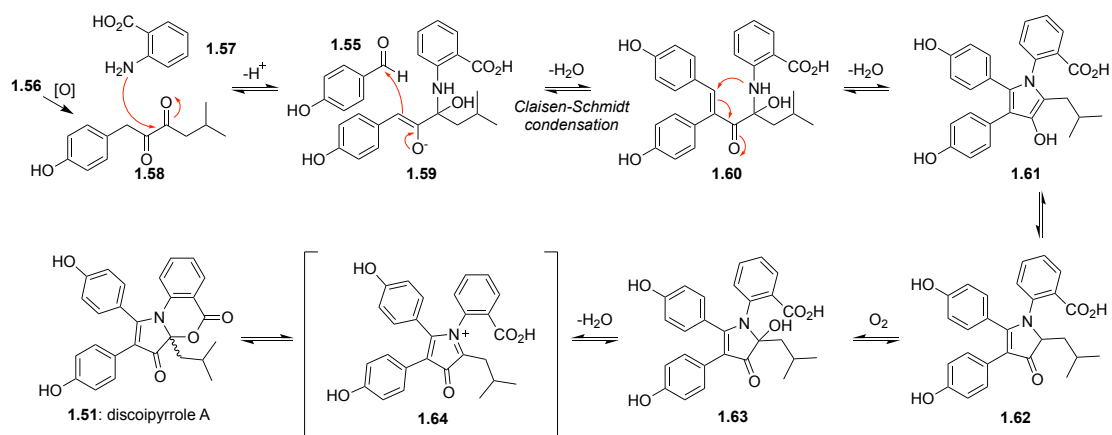


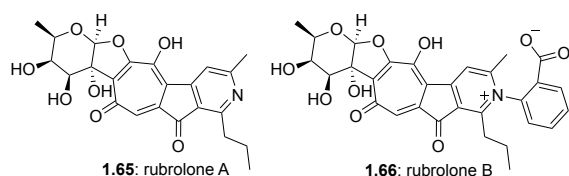
Figure 1.16. Proposed mechanism for the non-enzymatic biosynthesis of discoipyrrole A.¹⁴

The discoipyrroles and jadomycins have distinct similarities in their substrates, but interesting differences in their mechanism. Primarily, there is a similar use of a benzylic aldehyde electrophile and a proteogenic amine nucleophile, either L-isoleucine or the amino acid derived anthranilate in the jadomycins and discoipyrroles, respectively.^{10, 98, 102, 106} In the jadomycin biosynthesis, the attack of the aldehyde in **1.47** happens directly with the amino group of L-isoleucine (Figure 1.14). However, the mechanistic studies of the discoipyrroles show that 4-hydroxybenzaldehyde (**1.55**) and anthranilate (**1.57**) do not undergo aldimine formation in the absence of satabacin (**1.58**) in biomimetic conditions, even under the scrutiny of ¹⁵N enhanced two-dimensional NMR (Figure 1.16).¹⁴ Additionally, using non-carboxylated aniline derivatives in the biomimetic synthesis results in the accumulation of the condensation product, **1.61**, that does not progress to form the discoipyrrole-like compounds. This difference could be explained by the tempered nucleophilicity of the substituted anthranilic acid or the role of the acid in substrate selection between **1.55** and **1.56**.¹⁴ These comparisons suggest that while certain moieties are more susceptible for non-enzymatic reactions, the surrounding substrates and environment dictate the exact reaction mechanism.

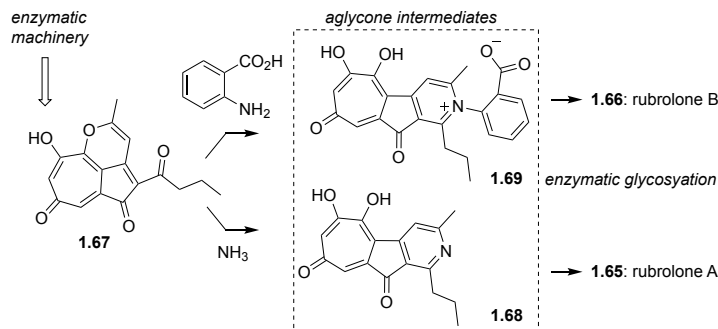
Continuing the theme of nucleophilic amines, rubrolone A (**1.65**) and B (**1.66**) are pyridine containing type II polyketides isolated from different *Streptomyces* strains in 1978 and 2016, respectively (Figure 1.17a).¹⁰⁷⁻¹⁰⁹ While these compounds represent interesting chemical complexity, their lack of bioactivity has potentially stymied biosynthetic studies until recently. In 2016, Yan and co-workers used a unique dehydratase gene, previously determined to be effective for mining of BGCs, to search for the putative rubrolone BGC within the sequenced genome of *Streptomyces* sp. KIB-H033.¹⁰⁹ After defining the biosynthetic pathway needed for the synthesis of the late stage intermediate **1.67**, it was thought that this molecule reacted non-enzymatically with different endogenous nucleophilic nitrogen species, ammonia for **1.68** and anthranilic acid for **1.69** (Figure 1.17b). In their heterologous expression of the BGC in the model bacterium, *Streptomyces albus*, which produces anthranilate in low titer, the yields of

1.68 were diminished unless supplemented with exogenous ^{15}N -anthranilate. ^{15}N was readily incorporated into rubrolone B (**1.66**) but not rubrolone A (**1.65**). This data and the lack of any appropriate aminotransferase in the BGC provided biological evidence of this non-enzymatic mechanism and was tested empirically by reacting isolated substrates in an aqueous phosphate buffer at the appropriate fermentation conditions *in vitro*.¹⁰⁹

a. Rubrolone structures reported



b. Biosynthetic mechanism of rubrolones



c. Proposed mechanism for pyridine formation

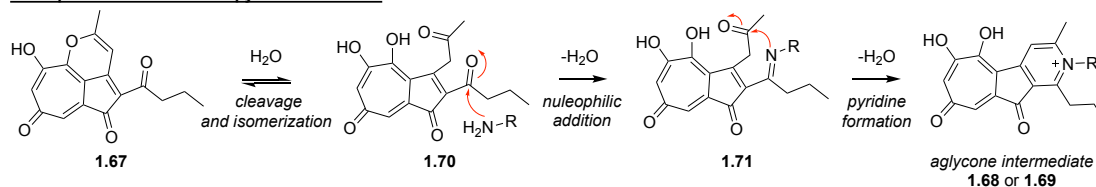


Figure 1.17. Investigation of the rubrolone biosynthesis. a) Structures of rubrolone A and B. b) The proposed non-enzymatic biosynthesis featuring the nucleophilic addition of amines to form pyridine rings in an unprecedented fashion. c) Suggested mechanism of the pyridine formation of aglycone intermediates of rubrolone A and B.¹⁰⁹

Interestingly, the non-enzymatic nucleophilic addition is not the final step in the biosynthesis. Rather, the initial attack causes a secondary reaction to generate the pyridine

ring of the rubrolones. Specifically, **1.67** is proposed to exist in equilibrium with **1.70** due to a hydrolytic ring opening demonstrated by deuterium exchange experiments (Figure 1.17c).¹⁰⁹ Nucleophilic attack of **1.70** with an amine substrate occurs to form the intermediate **1.71**. The reactive 1,5-dione of **1.71** can trap the non-enzymatically formed imine nitrogen resulting in the subsequent formation of the pyridine ring. This pyridine forming mechanism had never been reported before in natural product biosynthesis.¹⁰⁹ Reactions between amines and 1,5-diones have been reported for the formation of pyridine rings in chemistry efforts under high temperatures. This suggests the inherent properties of **1.71** could provide a uniquely selective advantage for cyclization reactions.

A unique component of the multi-step non-enzymatic pyridine formation is that it occurs prior to enzyme engagement downstream in the biosynthetic pathway.¹⁰⁹ This was demonstrated when glycosylated rubrolone levels directly correlated with the pool of pyridine formed intermediates, but not acyclic intermediates. This is similar to the case of the jadomycins in which non-enzymatic reactions occur before the natural product is reengaged with the catalytic proteins.⁹⁹⁻¹⁰⁵ Speaking to the substrate specificity of said enzymes in the rubrolone biosynthesis, a variety of functionalized anthranilic acid analogs were successfully incorporated into mature rubrolones *in vitro*. This surprising finding could be explained by a function of the rubrolones that would allow promiscuity in the substrate binding pocket if driven by an environmental pressure, albeit there is no known biological function attributed to these natural products.¹⁰⁹

The fungal based family of compounds, the maximiscins, starkly differ from the other multi-component biosynthetic reactions in that these reactions have demonstrated a unique biological function (Figure 1.18). The maximiscins have been shown to have potent activity against various types of cancer cell lines *in vitro* and *in vivo*.^{110, 111} These molecules have been studied extensively by the Cichewicz group since their discovery in 2014 after a series of

OSMAC-like experiments successfully elicited altered metabolite profiles of natural product producing *Tolypocladium* fungus.¹¹⁰⁻¹¹²

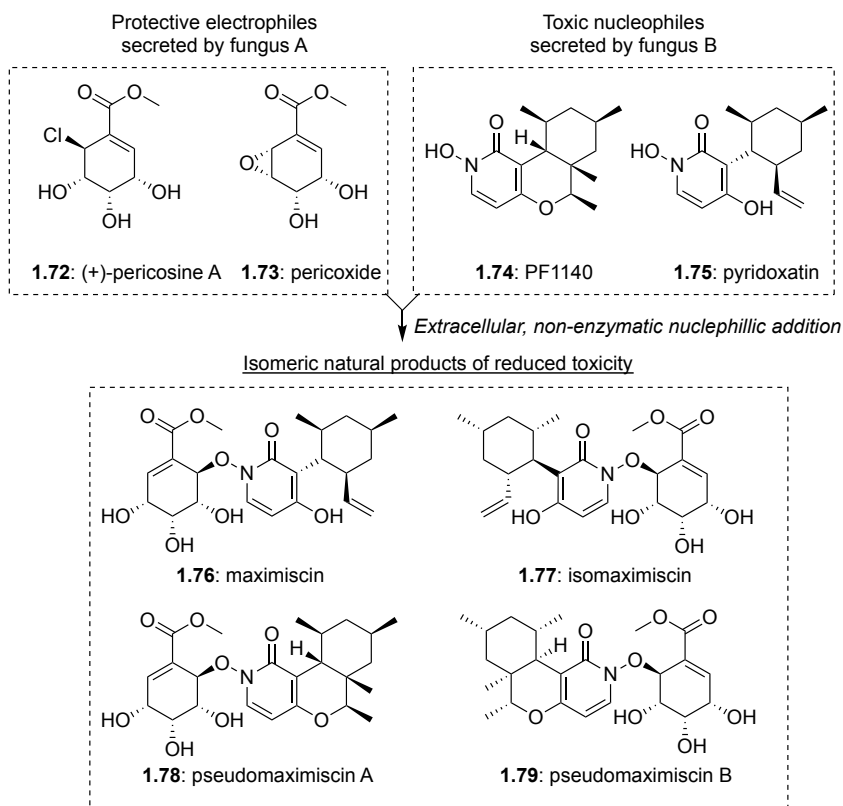


Figure 1.18. Conceptual overview of the non-enzymatic construction for the maximiscins and related compounds.¹¹⁰⁻¹¹²

In their 2014 report, Du et. al took advantage of isotopic labeling experiments to derive the biosynthetic origins of every carbon atom in the molecule, which features a unique amalgam of precursors (hence the name maximiscin meaning “from many sources”).¹¹⁰ The final step of the biosynthetic synthesis was proposed to be a bond formation between a shikimate-derived sugar and a polyketide-NRPS derived subunit. This reaction was reversible as demonstrated by the slow isomerization of maximiscin (**1.76**) to the isomer, isomaximiscin (**1.77**) (Figure

1.18). This data alongside careful mass spectrometry observations in co-culture experiments of the same fungal strain led the others to study the ability of the electrophilic precursors, (+)-pericosine A (**1.72**) and pericoxide (**1.73**), to participate in non-enzymatic nucleophilic additions to form the maximiscins via pyridoxatin (**1.75**) and various related molecules through exogenously added nucleophiles such as pseudomaximiscin A (**1.78**) and B (**1.79**) (Figure 1.18).^{111, 112}

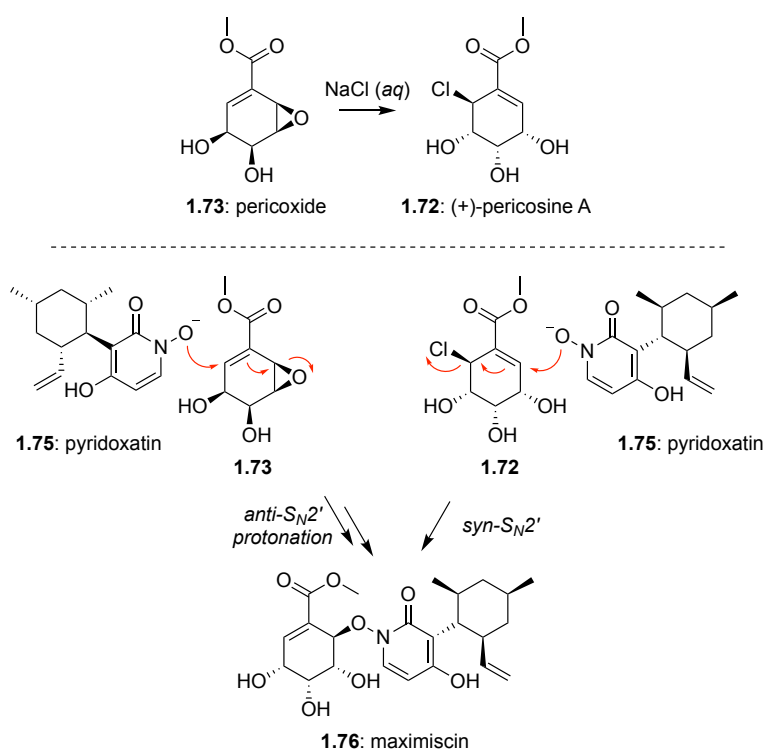


Figure 1.19. Mechanistic details for the non-enzymatic formation of maximiscin from various precursors as suggested by *in vitro* studies and DFT calculations.¹¹⁰⁻¹¹³

The formation of maximiscin and analogs from hydroxamic acids were found to be stereoselective. Du and co-workers used their substrate diversification studies alongside computational DFT calculations to examine the stereoselectivity of different nucleophiles.¹¹⁰⁻¹¹²

They were able to characterize the mechanism by which hydroxamic acids such as **1.75**, undergo nucleophilic addition to **1.72** and **1.73** which proceeds via anti-S_N2' and syn-S_N2' to form the non-racemic **1.76** (Figure 1.19).¹¹²

The story of the electrophilic natural products (+)-pericosine A (**1.72**) and pericoxide (**1.73**) are prime examples of how biological function can be attributed to a non-enzymatic reaction. While studying the *Tolypocladium* fungal strain T1, it was found that many of the nucleophiles that combined non-enzymatically with **1.72** and **1.73** possessed anti-fungal activity and that these new molecules were non-toxic.¹¹² Supplementation of “spent” media from T1 containing the electrophilic nucleophiles suppressed the effect of exogenously added nucleophiles. This suggests that the T1 strain has developed a non-enzymatic detoxification system to protect against competing fungi.¹¹² While the defense system remains to be studied on a genetic level, it suggests an exciting non-enzymatic relationship between secondary metabolism and basic microbial function.

In 2019, the Clardy laboratory isolated the family of antifungals, the pyonitrins A-D (**1.80-1.83**), from an insect-associated bacterial strain of *Pseudomonas protegens* (Figure 1.20).¹¹³ The pyonitrins were screened as a crude mixture *in vivo* using a mouse model which showed promising anti-fungal activity, leading the authors to further investigate the extent of pyonitrin production. Detailed analysis of the *P. protegens* genome showed the BGC for pyochelin (**1.84**) was significantly far from the BGC for pyrrolnitrin (**1.85**), suggesting that the biosynthetic pathways were not merged to form the pyonitrins (Figure 1.20).¹¹³ Therefore, it was hypothesized that the pyonitrins are formed through a non-enzymatic Pictet-Spengler condensation between aeruginaldehyde (**1.86**), an intermediate in the production of pyochelin, and an aminopyrrolnitrin derivative (**1.87**) of the pyrrolnitrin pathway. The condensation first forms the imine intermediate **1.89** that undergoes an intramolecular cyclization to yield a 1,2-dihydroquinoline structure (**1.90**), which re-aromatizes into the pyonitrins (Figure 1.21).^{113, 114}

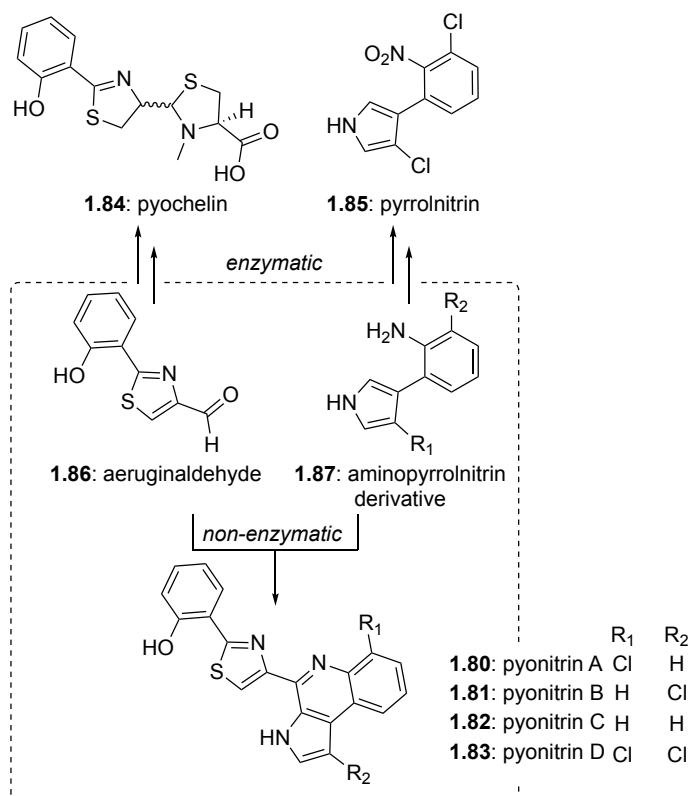


Figure 1.20. The non-enzymatic formation of the pyonitrins. The pyonitrins are derived from the shunt product, aeruginaldehyde, from the pyochelin biosynthetic pathway with aminopyrrolnitrin, an intermediate from the pyrrolnitrin biosynthetic pathway.^{113, 114}

The hypothesis was confirmed by conducting feeding studies using exogenous aminopyrrolnitrin derivatives (**1.87**) with **1.86**, isolated from the host species, in abiotic culture conditions and the formation of pyonitrin A (**1.80**) and D (**1.83**) was observed after 24 hours.¹¹³ This approach led to the first biomimetic synthesis of **1.80-1.83** in only three steps.¹¹⁴ Additionally, by isotopically labeling the aniline moiety in dechloroaminopyrrolnitrin (**1.88**), NMR-based studies were used to observe the formation of **1.80** and identify the transient intermediates in real time. The results showed immediate formation of the 1,2-dihydroquinoline moiety (**1.90**) and complete conversion to **1.80** after 96 hours. No evidence of the imine intermediate **1.89** was observed, which is attributed to the inherent reactivity and spatial

proximity of the starting precursors allowing for immediate cyclization of the imine (Figure 1.21).¹¹⁴ The identification of this rapid cyclization allows for the development of novel pyonitrin analogs and the ability to probe for structural activity relationships (SAR) in a synthetically efficient manner, which is highly valuable from a medicinal chemistry perspective. However, most notably, this is the first characterized example in natural product biosynthesis where a Pictet-Spengler condensation is not enzymatically catalyzed.^{113, 114}

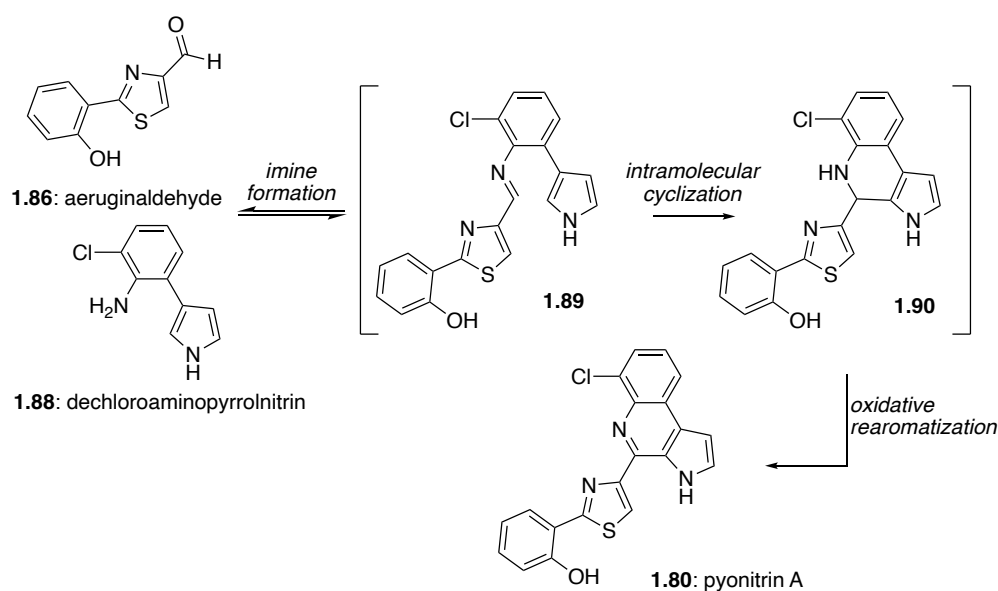
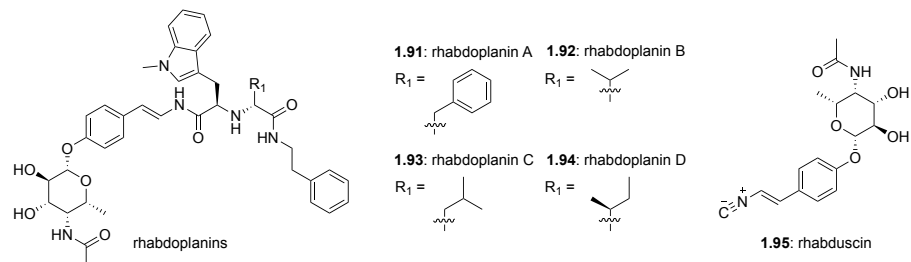


Figure 1.21. Non-enzymatic Pictet-Spengler mechanism of pyonitrin A formation.¹¹⁴

The last multi-component non-enzymatic biosynthetic mechanism that will be discussed are the rhabdoplanins A-D (**1.91-1.94**) (Figure 1.22). The rhabdoplanins were isolated from a *Xenorhabdus nematophila* bacterial insect pathogen in 2019 by the Crawford laboratory and contains unprecedented opine-glyco-peptide scaffold.¹¹⁵ It was proposed that the rhabdoplanins arise via a non-enzymatic, diastereoselective, “Ugi-like” reaction between the isocyanide-functionalized metabolite, rhabduscin (**1.95**) and a reactive imine species

(**1.101**) (Figure 1.22). To confirm this hypothesis, Oh and co-workers performed an *in vitro* biomimetic synthesis of rhabdoplanin A (**1.91**) by isolating the putative precursor **1.95** along with other precursors including phenylpyruvate (**1.96**), phenylethylamine (**1.97**), and *N*-methyltryptophan (**1.99**) from the host species. It was shown that **1.96** and **1.97** initially react to form the amide intermediate, **1.98**, which undergoes a condensation reaction with **1.99** to generate the imine intermediate **1.100**. Decarboxylation of **1.100** to give **1.101** allows for nucleophilic attack from the isocyanide group of **1.95** to form the opine core of the rhabdoplanins. Finally, the intermediate **1.102** undergoes hydrolytic oxidation to form **1.91**.¹¹⁵ As the rhabdoplanins are generated in high diastereoselectivity in the bacterial culture, it was proposed that the stereoselective condensation is similar to that observed in Ugi reaction with chiral amines. This hypothesis was validated through combinatorial biomimetic synthesis of **1.91**.¹¹⁵

a) Rhabdoplanin family of natural products



b) Biosynthetic mechanism of rhabdoplanin A

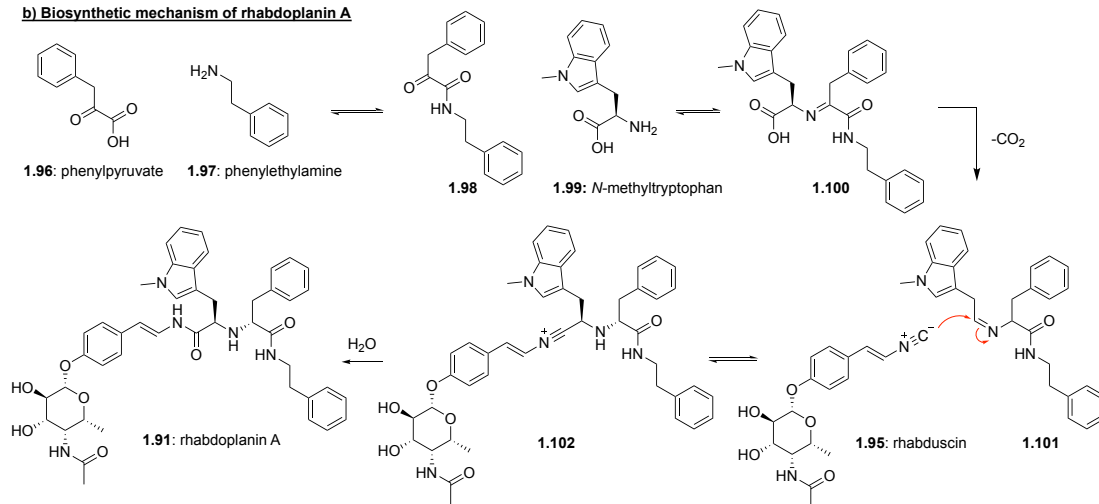


Figure 1.22. Rhabdoplanin family and its formation. a) The structure of the rhabdoplanin family of natural products and the putative precursor rhabduscin. b) Proposed non-enzymatic “Ugi-like” biosynthetic pathway for the rhabdoplanins.¹¹⁵

The isocyanide group on **1.95** is essential to this novel non-enzymatic “Ugi-like” formation of the rhabdoplanins. Numerous isocyanide-containing natural products have been isolated since 1950 and play important roles in synthetic chemistry.^{116, 117} These molecules also have biological relevance as rhabduscin inhibits the phenoloxidase pathway, which contributes to bacterial pathogenesis. The rhabdoplanins themselves have also shown to be a new type of G protein-coupled receptor (GPCR) agonists.¹¹⁵ Not only do they present a new lead in GPCR agonists but due to the abundance of isocyanide-functionalized natural products, there is potential for these types of multi-component reactions to be more ubiquitous than we think.

1.5 Tailoring reactions

The multi-component reactions discussed above involve the formation of chemical bonds that often trigger subsequent intramolecular reactions. However, non-enzymatic reactions can also occur on mature natural products, where there are no subsequent reactions and the transformation results in a modest change to the structure of the molecule. Based on the complexity of the substrate reacted with the mature natural product, we can divide this category into two main classes: reactions involving simple molecular species (O_2 , SH_2 , NH_3 , etc.) and those involving higher molecular weight species.

The simplest non-enzymatic tailoring reaction is between a secondary metabolite and a reactive oxygen, nitrogen, or sulfur species (ROS, RNS, RSS). These molecules are common by-products of primary metabolism and their ability to participate in non-enzymatic reactions is demonstrated by their documented reactivity with the proteome, which has been previously reviewed.¹⁰ Canonical enzymatic natural product biosynthesis uses simple building blocks, such as O_2 for oxidation, therefore the delineation between non-enzymatic and enzymatic processes in complex pathways is difficult to make. What stands out are the reactions that occur between these reactive species and mature natural products. In this context, a mature natural product is one that has undergone the majority of its biosynthetic pathway. In many cases, the identification of these transformations is assisted by the co-isolation of the mature natural product and their non-enzymatic derivative.

The biosynthesis of the ammosamides (**1.103-1.106**) is one of the most well studied cases of a pathway featuring a non-enzymatic tailoring reaction (Figure 1.23).¹¹⁸ After the original description of the ammosamide family in the extracts of a *Streptomyces* strain from the Bahamian islands, it was found that the BGC in *Salinispora* and *Streptomyces* did not contain a gene for an oxidoreductase that would catalyze the final oxidation step of ammosamide C (**1.104**) to ammosamide B (**1.105**) (Figure 1.23).^{118, 119} Relatedly, our group found that **1.104**

could easily react with several nucleophilic amines to form non-biogenetic derivatives.^{118, 120} These two independent reports suggest that **1.104**, containing an iminium functional group, was a precursor for a non-enzymatic derivatization to form the remaining ammosamides.

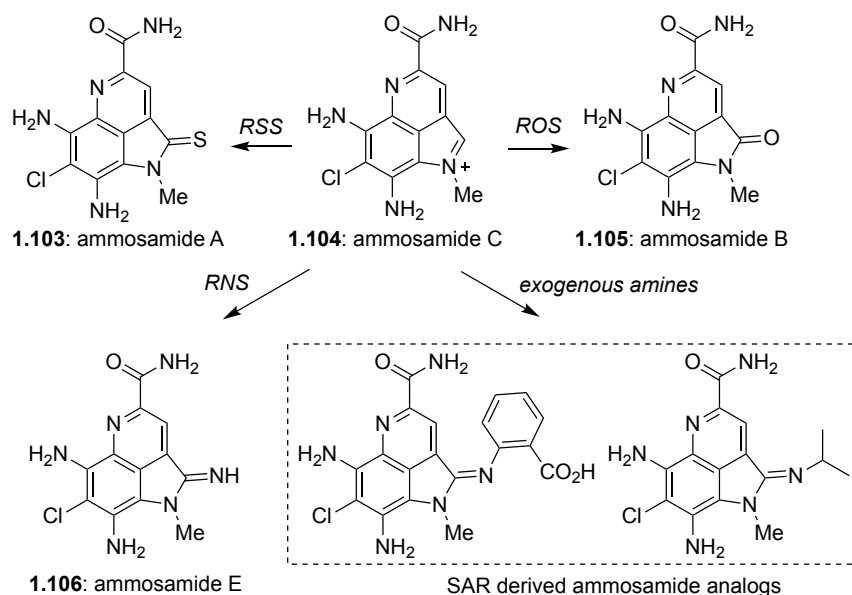


Figure 1.23. Non-enzymatic derivatization of the ammosamide family via the highly reactive iminium functional group of ammosamide C.^{118, 120}

Ammosamide A (**1.103**), B, (**1.105**), and E (**1.106**) are single atom derivatives of **1.104**, suggesting that these originated from within a *RSS*, *ROS*, or *RNS* produced by the bacteria.¹¹⁹ While the idea of having an iminium-based electrophile to either neutralize toxins or respond to metabolic by-products is alluring, further biological studies are needed to support that the ammosamides are a functional tool. Function aside, the ammosamides are an ideal example of how identifying non-enzymatic mechanisms can be used to perform medicinal chemistry quickly and efficiently on biologically active compounds as over 20 analogs were constructed using benign conditions and a variety of nucleophilic amine substrates, of which some had improved bioactivity over the natural products themselves.^{119, 120}

Redox reactions, which are essential to all life, also fall under the category of tailoring reactions as these occur after the biosynthesis of a mature natural product. The most recent study of this type of non-enzymatic transformation is seen in the case of the quinone-like molecule, 2-amino-3-carboxy-1,4-naphthoquinone (ACNQ) (**1.107**) (Figure 1.24).^{121, 122} ACNQ was isolated from a *Shewanella* species and identified as the essential electron shuttling molecule to the terminal electron acceptor (TEA) for anaerobic respiration.¹²³ The TEA in most organisms is oxygen, however, anaerobic microbes must employ an alternative mechanism to transport electrons for respiration. The ability for microbes to extracellularly transfer electrons to insoluble TEAs are of particular interest in the context of bioenergy development such as bioremediation, biofuel production, and biomining, thus making the identification of this elusive electron shuttle highly desired.^{121, 122}

In 2000, Newman and Kolter discovered that *Shewanella* mutants with genetic disruptions in the biosynthetic pathway of menaquinone (MK) (**1.108**), a naphthoquinone that mediates redox reactions in anaerobic respiration, were not able to grow under anaerobic conditions independently. However, respiration was restored when cultured with the wildtype strain.¹²² It was determined that a small molecule was being excreted into the fermentation media and oxidized for the purpose of shuttling electrons to the TEA, but the molecule itself hadn't been identified until 2019 by Mevers and co-workers.^{121, 122}

ACNQ (**1.107**) is derived from 1,4-dihydroxy-2-naphthoic acid (DHNA) (**1.109**), a biosynthetic precursor to MK (**1.108**), as studies report mutations inhibiting the production of **1.109** also prevented **1.107** production, but the genes responsible for this conversion were unable to be identified after various genetic screens (Figure 1.24).¹²¹ This led to the hypothesis that conversion happens via a non-enzymatic oxidation of **1.109** to **1.110**, which then undergoes a nucleophilic addition of ammonia to intermediate **1.111**. Intermediate **1.111** tautomerizes to its enol form and the hydroquinone product (**1.112**) is oxidized to **1.107**. To explore this proposed pathway, researchers incubated **1.109** in several different aqueous

solutions supplemented with various nitrogen sources. They observed consistent conversion to **1.107** regardless of the aqueous conditions, further supporting the proposed non-enzymatic pathway.¹²¹

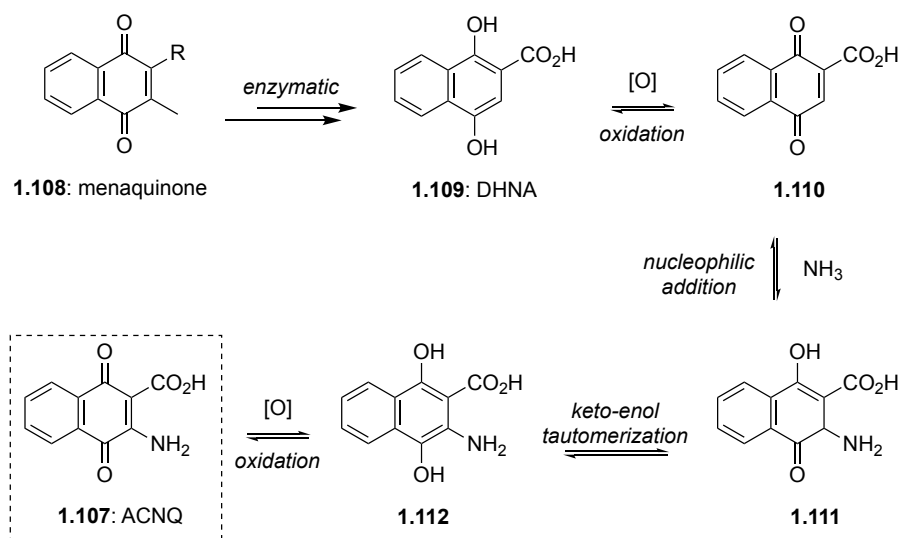


Figure 1.24. Proposed mechanism for the non-enzymatic conversion of DHNA to ACNQ.¹²¹

The second class of tailoring reactions involves the modification of mature natural products with common metabolic by-products of higher molecular weight substrates. Instead of reactions involving ROS, RNS, RSS, the following reactions involve amino acids or their derivatives. Though similar to multi-component reactions, these tailoring reactions occur on the periphery of the molecule and are not commonly followed by further reactions.

The anamycins are a well characterized and biomedically important family of macrolactams, which include common antibiotics such as rifamycin and geldanamycin (not shown), that exhibit high molecular weight tailoring reactions. Recently, Zhang and co-workers used the unique gene structure of the BGC for this family of molecules to search for its presence

in a collection of various *Streptomyces* strains.¹²³ The pentaketides, known as the juanlimycins, were discovered in a bacterium associated with a Chinese plant of the *Kandelia candel* family. The final structures of the juanlimycin A (**1.113**) and B (**1.114**) were postulated to be derived from a reaction between the thioesterification intermediate **1.115** and aryl amines, anthranilic acid and 3-hydroxy-5-aminobenzoic acid (Figure 1.25).¹²³ While the authors did not experimentally interrogate this hypothesis, the known availability and reactivity of such aryl amines in previous non-enzymatic mechanisms is intriguing.

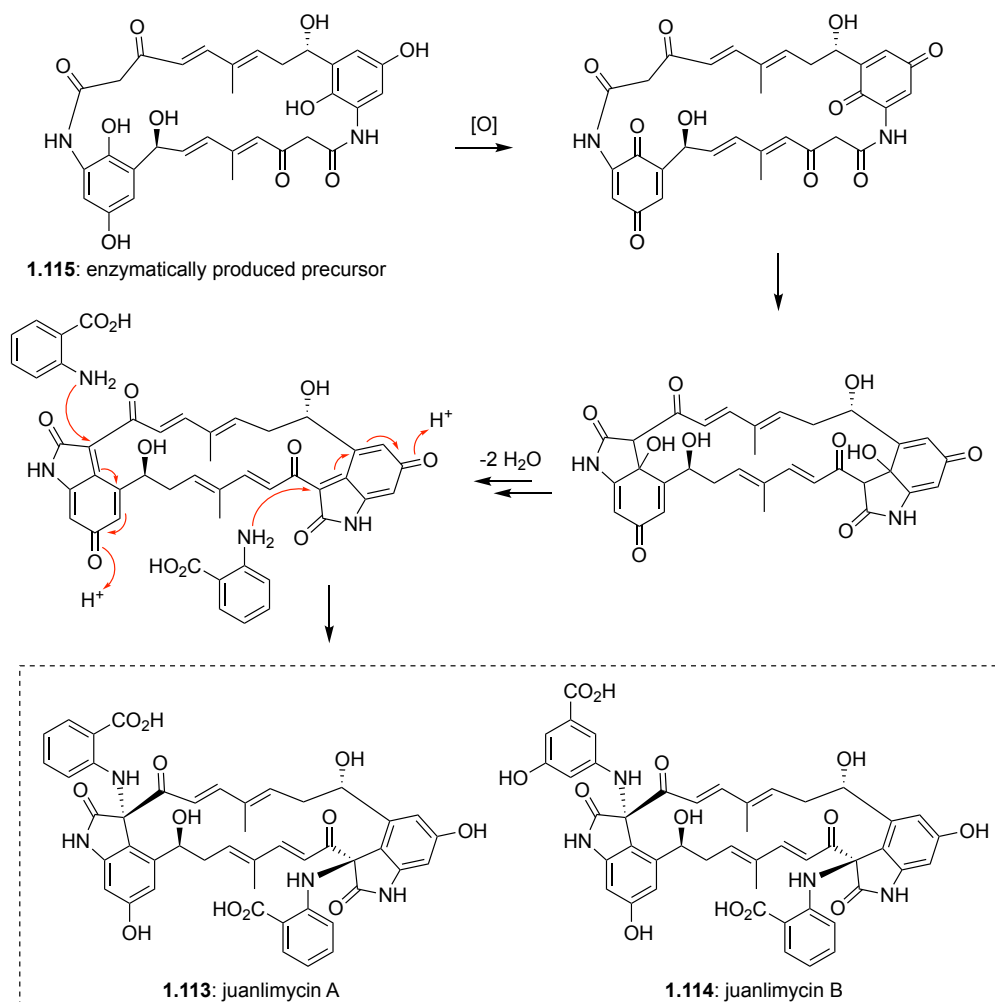


Figure 1.25. Proposed mechanism for the biosynthesis of juanlimycin A and B.¹²³

Another example of these tailoring reactions is by Wang and co-workers from the Balskus laboratory, who published a biosynthetic interrogation of kinamycin-like molecules (**1.116**) in *Streptomyces murayamaensis*, which feature several non-enzymatic steps in their biosynthesis.¹²⁴ Their studies focused on the construction of seongomycin (**1.117**) and stealthin C (**1.118**) from the enzymatically derived precursor for many angucycline polyketides, dehydrorabelomycin (**1.119**) (Figure 1.26a). Researchers discovered through *in vitro* studies that when **1.119** is in the presence of pure oxidase enzymes AlpJ and Fre along with L-cysteine, **1.118** and the final product **1.120** are both generated (Figure 1.26b).¹²⁴ It was proposed that the flavin enzyme AlpJ and the flavin reductase Fre catalyzed a Baeyer-Villiger based C-C bond cleavage/ring contraction and subsequent oxidation to produce an intermediate that is susceptible to non-enzymatic nucleophilic addition of thiols such as L-cysteine. The result of this non-enzymatic addition was thought to be *N*-deacetylseongomycin (**1.121**), which could undergo an intramolecular S-N-type Smiles arrangement to form **1.120** that is subject to non-enzymatic oxidation and hydrolysis to form **1.118** (Figure 1.26c). The non-enzymatic thiol addition was validated by the incorporation of various thiols into new dehydrorabelomycin derivatives, exemplifying a lack of enzyme specificity. The non-enzymatic oxidation and hydrolysis to **1.118** was validated by boiling solutions of **1.120** to degrade present enzymes before observing conversion to **1.118**.¹²⁴

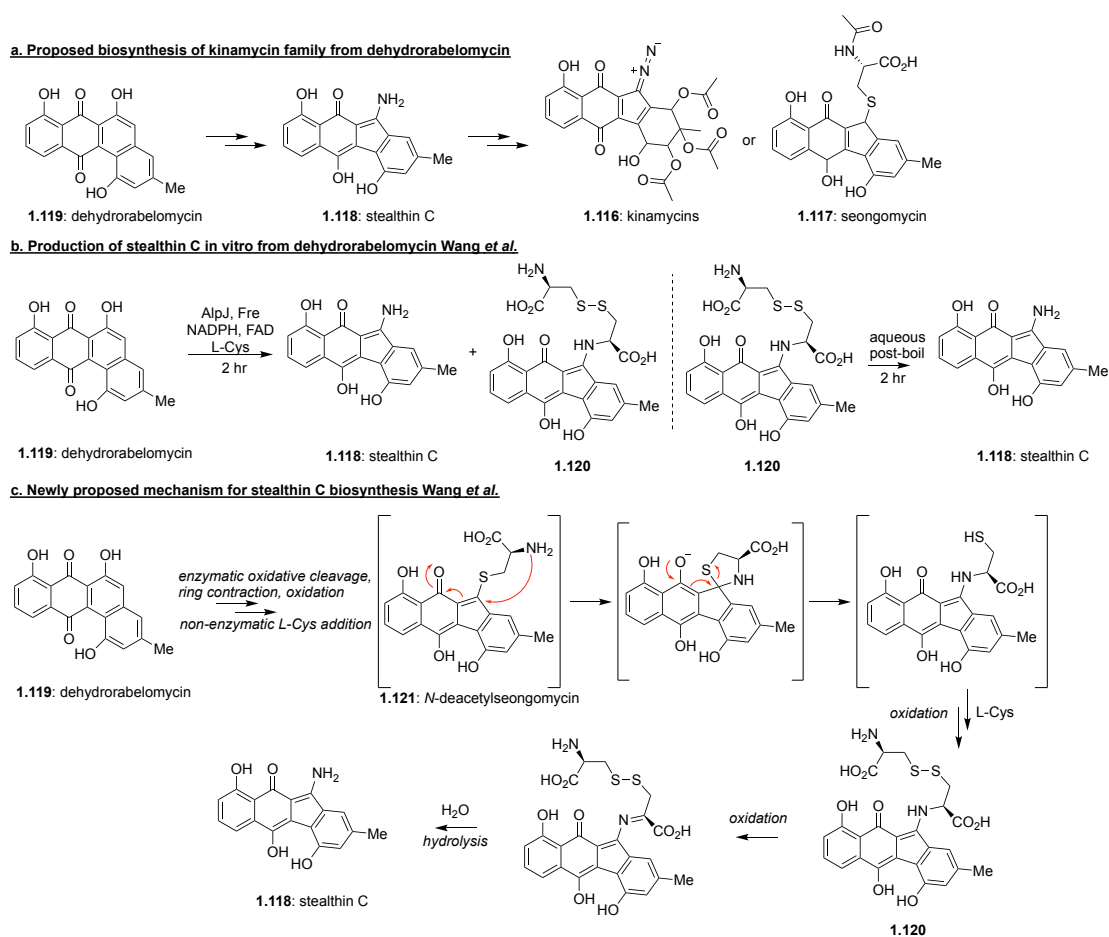


Figure 1.26. The progression of stealthin C discovery. a) Original description of stealthin C as an intermediate in kinamycin biosynthesis.¹²⁶ b) Representation of the experimental data gathered to suggest stealthin C was formed non-enzymatically.¹²⁴ c) Non-enzymatic mechanism proposed for stealthin C.¹²⁴

The discovery of stealthin C being derived from non-enzymatic mechanisms called into question an important component of angucycline polyketide biosynthesis.¹²⁵⁻¹²⁷ It has been suggested that **1.118** is a biosynthetic intermediate of the kinamycins (**1.116**) based on ¹³C isotope feeding studies and the accumulation of **1.118** in heterologous expression of the partial kinamycin BGC.¹²⁷ However, these experiments did not completely conclude the direct conversion of **1.118** to **1.116**. Furthermore, Wang and co-workers hypothesized that the reactivity of the non-enzymatic mechanism would make **1.118** unstable as an intermediate in

the kinamycin pathway.¹²⁴ They validated this alternative hypothesis using ¹⁵N isotopic labeling, suggesting that **1.118** is not an intermediate to the kinamycins, but rather a diverted pathway from the dehydroabelomycin (**1.119**) that features non-enzymatic diversification.

The investigation of the stealthins by the Balskus laboratory highlights how studying non-enzymatic transformations can lead to broader scientific discovery. Identification and experimental validation of these non-enzymatic mechanisms to provide insight into a biosynthetic pathway central to a large family of natural products. Also, the stealthins join the jadomycins as angucycline polyketides that participate in non-enzymatic reactions. While it remains to be seen if other members of this family are also constructed with non-enzymatic steps, the angucycline polyketides may represent a privileged scaffold for chemical reactivity.

1.6 Light-induced reactions

The utility of UV-induced transformations in synthetic chemistry is widely known.¹²⁸ However, its use in natural product biosynthesis is questionable, as the ecology of the natural product source is often not explored in the biomimetic synthesis. A truly biomimetic source of light is rarely used, but rather a high intensity of UV irradiation source is traditionally used to conduct these reactions.¹²⁸ That being said, there are cases of ecologically relevant UV-induced biosynthesis, including the brevianamides which feature a Norrish-type cleavage, the biyouyanagins undergo [2 + 2] cycloaddition, and most importantly, vitamin D3 which is formed after a rearrangement of 7-deoxycholesterol, all facilitated by UV (none shown).^{13, 128} These are thoroughly discussed by Gravel and Poupon's review of non-enzymatic biosynthesis, therefore here will discuss modern examples of UV-induced non-enzymatic biosynthesis.¹³

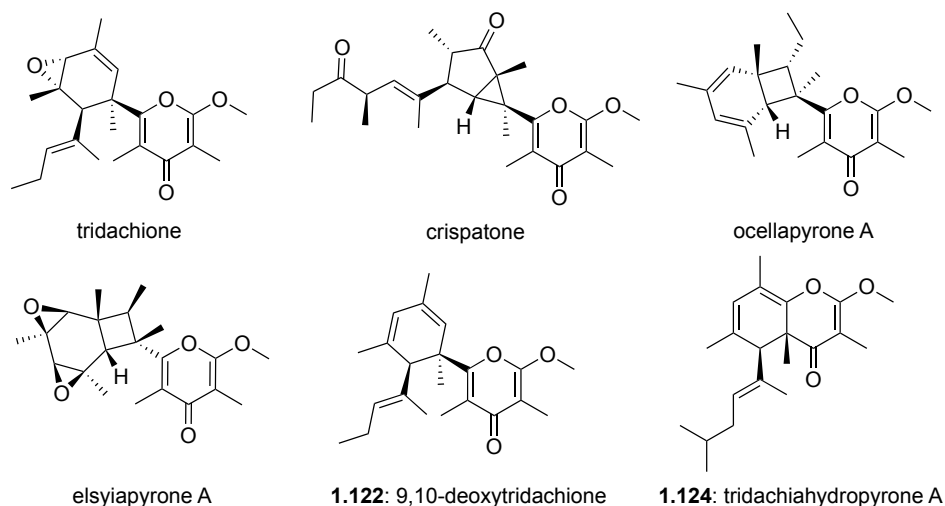
The production of flueggidine (**1.26**), previously mentioned as an example of a homologous multi-component dimerization, has also been shown to be synthesized from its

monomeric precursor by a UV light-induced [2 + 2] cycloaddition, although in low yields (5%) (Figure 1.10b).⁷⁰ Additionally, this dimer was only found in the leaves of the plant, not in the roots, suggesting this could be an ecologically relevant non-enzymatic reaction. Although flueggidine differs from the rest of the other securinene dimers in its constructed mechanism, the whole family is mainly categorized as non-enzymatic dimerizations, with this single example classified secondarily as a light-induced reaction.⁷⁰

When considering UV-induced non-enzymatic biosynthesis, it is obvious that not all ecological niches are created equal. In the case of flueggidine (**1.26**), the sequestering of the reactive monomer, virosecurinene (**1.22**), within the leaves is intuitive given the role of the foliage in carbon fixation (Figure 1.10b).⁶⁹ Likewise, UV-induced reactions have been implicated in marine invertebrates that are subject to continuous sunlight.

In the late 1970s, Ireland and Scheuer were investigating the Hawaiian based sacoglossan mollusk, *Placobranchius ocellatus*, and their ability to assimilate chloroplasts from marine algae when they discovered a family of molecules featuring a γ -pyrone ring that could undergo photochemical isomerization.¹²⁹ These molecules have been previously described from *Tridachia* and *Tridachiella* mollusks (Figure 1.27a).¹³⁰ Specifically, 9,10-deoxytridachione (**1.122**) could isomerize to photodeoxytridachione (**1.123**) (Figure 1.27b).^{129, 130} The authors found that this could happen *in vitro* and *in vivo*, as detailed by radioisotopic labeling experiments in which light exposed mollusks showed increased flux of ¹⁴C from carbon dioxide to **1.122** and subsequently to **1.123** compared to mollusks kept in the dark. Their rationale was that sacoglossan mollusks, which lack a protective shell, could use by-products of algae photosynthesis to generate γ -pyrones that could then serve as sunscreen against harmful radiation.¹²⁹

a. Examples of γ -pyrone natural products isolated from marine invertebrates



b. Original work by Ireland and Scheuer

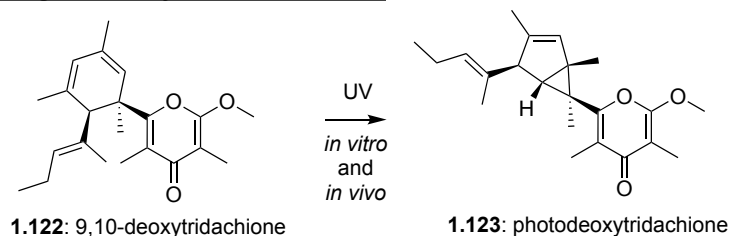


Figure 1.27. UV-induced formation of γ -pyrone natural products. a) Breadth of γ -pyrone natural products isolated from various marine organisms.¹³⁰ b) Seminal studies demonstrating photochemical isomerization of 9,10-deoxytridachione to photodeoxytridachione.^{129, 130}

The chemical diversity of these γ -pyrones has been expanding over the last 40 years as more are being discovered. However, in 2008, Eade and co-workers suggested that despite their diversity, most γ -pyrones could be derivatives of a single precursor consisting of a conserved γ -pyrone heterocycle linked to serial (*E*)-polyenes.¹³¹ They hypothesized that an (*E*)-polyene could undergo a series of *Z/E* isomerizations, thermal or photochemical electrocyclizations, [4 + 2] cycloadditions, and [2 + 2] concerted rearrangements to yield a variety of γ -pyrone family members. The authors demonstrated this with several γ -pyrones, albeit not consistently under truly biologically relevant conditions with no exogenous catalysts or organic solvents. They undertook a biomimetic synthesis of the complete set of

tridachiahypopyrones the following year and, interestingly enough, demonstrated that each of the tridachiahypopyrones could be derived photochemically from the precursor **1.125** (Figure 1.28).¹³² The authors showed that exposing **1.125** to “UK sunlight in the month of March for 3 consecutive and particularly sunny days” led to a 75% conversion to three products, intermediate **1.126**, tridachiahypopyrone A (**1.124**) and phototridachiahypopyrone (**1.127**) in a 20:7:1 ratio. Furthermore, photodeoxytridachione (**1.123**) irradiated for 24 hours led to quantitative conversion to **1.127**, presumably through a light-induced [1,3]-sigmatropic alkyl migration (Figure 1.28).¹³²

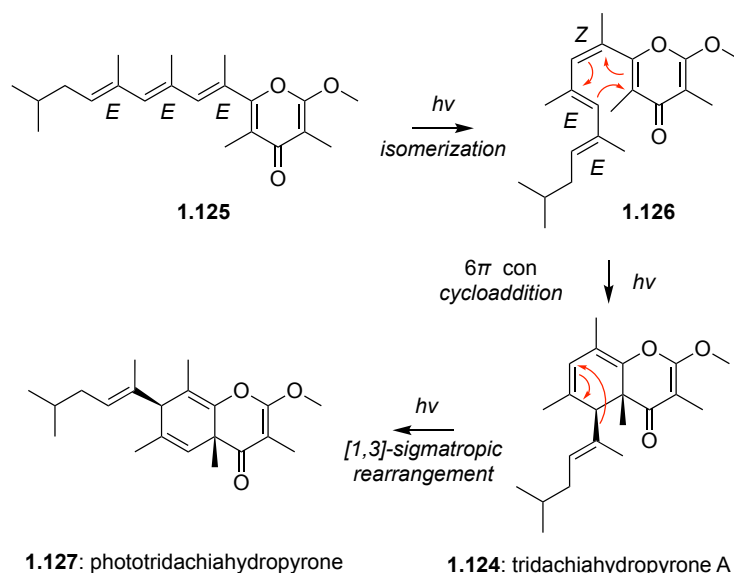


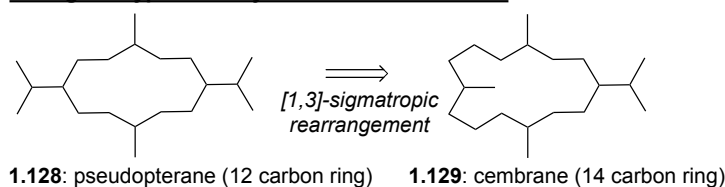
Figure 1.28. UV-induced non-enzymatic formation of phototridachiahypopyrone. Exposure to UV catalyzes isomerization and [1,3]-sigmatropic rearrangement of precursor **1.125** to phototridachiahypopyrone.^{132, 133}

While the extent to which these isomerization, cycloadditions, and rearrangements occur non-enzymatically in the context of marine invertebrates has yet to be identified, they remain an open inquiry on the matter. It is reasonable to predict that many of the initial steps in these biosynthetic pathways occur enzymatically with the involvement of isomerases and

related proteins. However, the final steps in which UV light can be absorbed by γ -pyrones, particularly in the exposed dorsal body of sea slugs as originally demonstrated by Ireland and Scheuer in their isotopic labeling experiments, can occur non-enzymatically.^{129, 130}

Early work by the Fenical laboratory in the 1980s uncovered the family of macrocyclic diterpenes known as the pseudopteranes (**1.128**) (Figure 1.29a).¹³³ In their original publication, it was suggested that the structure of these molecules could be a derivative of a related family of natural products, the cembranes (**1.129**), via a ring contraction. Rodriguez and Shi's isolation of the novel furanocembranolide, bipinnatin J (**1.130**), and the known pseudopterane, kallolide A (**1.131**), from the Eastern Caribbean Sea plume, *Pseudopterogorgia bipinnata*, gave credence to this hypothesis (Figure 1.29b).¹³⁴ The presence of both **1.128** and **1.129** in the same plume led authors to determine if the 14 membered carbon ring of **1.130** could undergo a [1,3]-sigmatropic rearrangement to yield the 12 membered carbon ring of **1.131** (Figure 1.29b). Additionally, the authors contemplated that UV light could provide the energy for the rearrangement. They found that the incubation of **1.130** in UV light for several hours led to a respectable 40% stereoselective conversion to **1.131**. While the number of ecological studies on the cembranes does not allow for comparison to the γ -pyrones but it could serve a similar sunscreen-like function for these coral-like invertebrates.^{133, 134}

a. Original hypothesis by Fenical and co-workers



b. Experimental non-enzymatic [1,3]-sigmatropic rearrangement

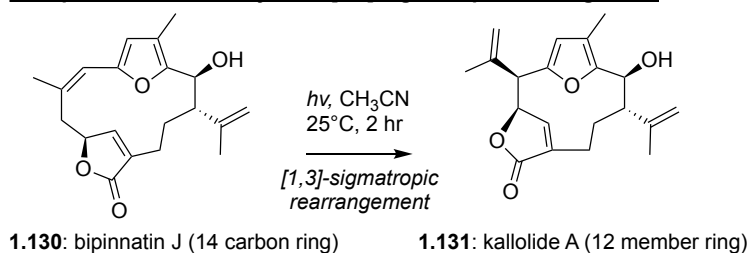
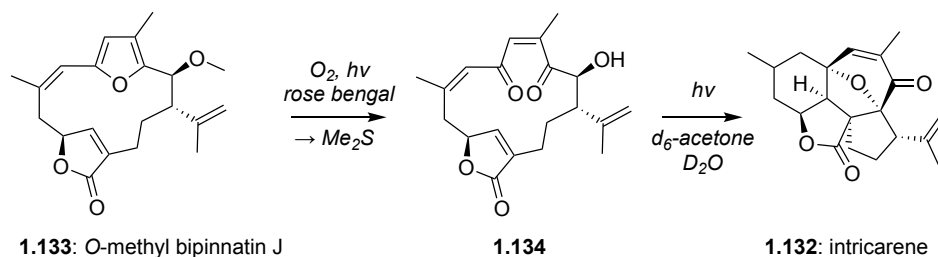


Figure 1.29. Validation of [1,3]-sigmatropic rearrangement of the cembranes. a) Original hypothesized mechanism for the rearrangement of cembrane scaffolds to create pseudopterane skeletons.¹³³ b) Experimental validation of this hypothesis with the bipinnatin J scaffold.¹³⁴

The Trauner laboratory and others have used UV-induced photochemistry to convert cembrane molecules to a family member of the pseudopterane molecules known as intricarene (**1.132**) by a [1,3]-dipolar cycloaddition (Figure 1.30a).¹³⁵⁻¹³⁷ Prior to 2014, the conditions to carry out this photochemistry were non-biological for they required extremely high temperatures and organic solvents. However, in 2014, Stichnoth and co-workers undertook an elegant and serendipitous biomimetic synthetic effort to **1.132** by *O*-methyl bipinnatin J (**1.133**) (Figure 1.30a), an oxidized derivative of the precursor used previously for the sigmatropic rearrangement photoconversion seen in kallolide A (**1.131**) (Figure 1.29b).¹³⁷ The researchers decided to use a UV lamp commonly used in reptile terrariums as the light source and an aqueous acetone solvent system, an altogether milder setting than had been used previously for bipinnatin-based synthetic efforts which resulted in a 25% yield of **1.132**. These experiments suggest a biosynthetic mechanism where **1.130** is enzymatically oxidized to give rise to the intermediate **1.134** that could then undergo a non-enzymatic, [1,3]-dipolar cycloaddition, however, this has yet to be justified *in vivo* (Figure 1.30b).

a. Trauner lab biomimetic synthesis of intricarene



b. Proposed biosynthesis of intricarene

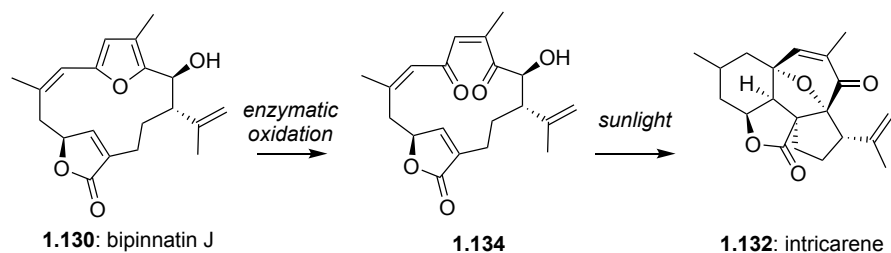


Figure 1.30. Investigation of intricarene biosynthesis. a) Biomimetic synthesis of intricarene from the building block O-methyl bipinnatin J. b) Proposed non-enzymatic UV-induced biosynthesis of intricarene from bipinnatin J based on synthetic efforts.¹³⁷

The chamigrenes are a unique family of brominated sesquiterpenes, of which there have been more than 50 molecules isolated from terrestrial and marine plants as well as marine invertebrates. Dactylone (**1.135**) and aplydactylone (**1.136**) are two members of this family discovered in the sea hare *Aplysia dactylomela* in the early 2000s by Fedorov and co-workers (Figure 1.31).^{138, 139} The researchers hypothesized that **1.136** could be derived from **1.135** via an enzymatic [2 + 2] cycloaddition, however, their attempts to cyclize these molecules under UV light were unsuccessful.

In 2016, the Burns laboratory took on the challenge of proposing a general approach for the enantioselective synthesis of the family of brominated chamigrenes and took this opportunity to examine the relationship between **1.135** and **1.136** (Figure 1.31a).¹⁴⁰ Conversion of **1.135** to **1.136** occurred under a variety of UV irradiation conditions, including an impressive

100% conversion after 15 minutes in THF that was irradiated at 254 nm. However, the more biologically relevant conditions, letting the compound sit in chloroform for 8 days in “Californian sunlight”, led to a 15% yield (Figure 1.31a). These experiments indicated that the cyclobutane present in **1.136** could in fact be formed non-enzymatically during exposure to UV light.¹⁴⁰

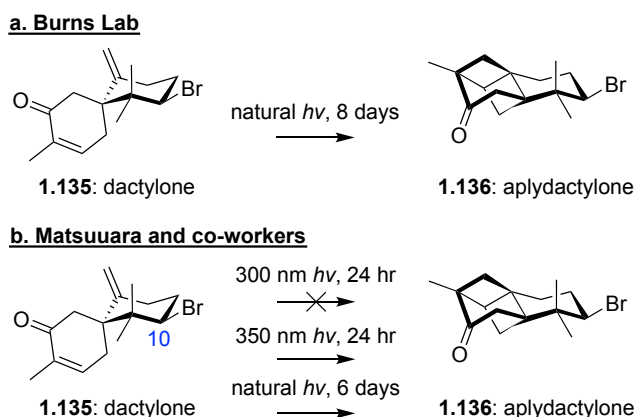


Figure 1.31. Experimental validation of non-enzymatic [2 + 2] cycloaddition of dactylone to form aplydactylone. a) Dactylone exposed to natural “Californian sunlight” showed 15% conversion to aplyactylone after 8 days.¹⁴⁰ b) Dactylone exposed to multiple artificial UV light showed no conversion but when exposed to “Munich sunlight”, conversion was observed after 6 days.¹⁴¹

Later that year, Matsuura and co-workers defined the mechanistic characteristics of the non-enzymatic UV-driven [2 + 2] cycloaddition reaction of **1.135** to **1.136**, as well as the potential cyclization of the diastereomer 10-*epi*-dactylone (Figure 1.31b).¹⁴¹ After synthesizing the respective precursors, the researchers took great care in assessing the type of UV light used in their photochemical reactions, selecting wavelengths that would be available in a natural light setting. Exposing **1.135** and 10-*epi*-dactylone to 350 nm UV light resulted in complete conversion over 24 hours, while 300 nm UV light led to the degradation of reagents (Figure 1.31b). Additionally, 6 days of “Munich sunlight” led to the conversion of **1.135** to **1.136**, further demonstrating that the [2 + 2] cycloaddition could occur non-enzymatically in response

to biological solar exposure.¹⁴¹ An important component to these studies was the justification for cyclobutane ring formation over a more relaxed cyclopentane structure and the authors carried out detailed quantum calculations for these determinations that will not be discussed here. Nevertheless, the work discussed in this section add to the growing examples of mechanistic studies being used to justify exactly how these non-enzymatic biosynthetic reactions are occurring.

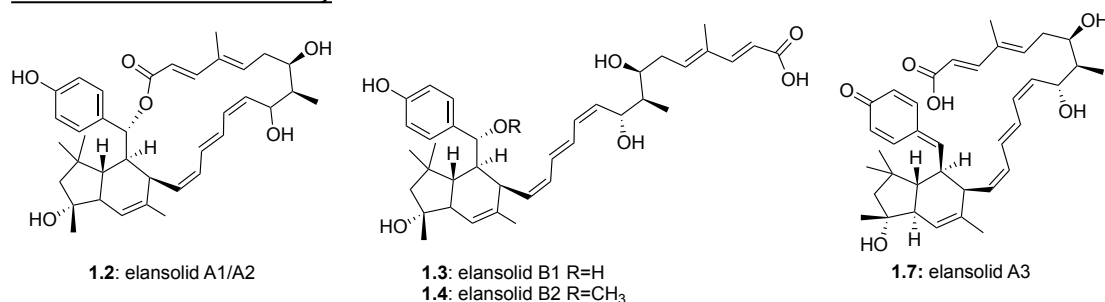
1.7 Non-enzymatic reactions for medicinal chemistry

The generation of new metabolites through the exploitation of non-enzymatic reactions is a potential approach to increase the chemical diversity of natural products. However, there lies an important distinction between processes such as these and natural biosynthetic approaches to produce biologically functional molecules non-enzymatically. There is a dichotomy of utility here: a medicinal chemistry perspective that suggests the circumstances of nature do not matter in the push to expand a chemical repertoire and a biosynthetic perspective that suggests the biological reconciliation of a non-enzymatic reaction is relevant. Either way, both perspectives are pipelines to impactful science.

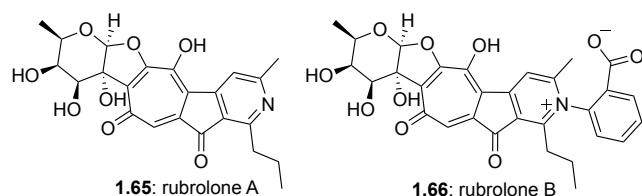
Validation is imperative to claims of non-enzymatic biosynthesis, as highlighted by the case of the elansolids. Elansolids A1/A2 (**1.2**) do not require exogenous sources for formation and were isolated in high yields from fermentation (Figure 1.32a). Discovery of the isolation artifacts elansolid B1 (**1.3**) and B2 (**1.4**) provided evidence that a non-enzymatic reaction could occur with a variety of different substrates.³⁹ This metabolome suggested an intermediate existed that predominantly reacts with endogenous substrates but could also form artificial derivatives during isolation. **1.2** provides a survival function to their host bacteria as antibiotics, therefore, the non-enzymatic formation is beneficial yet can still result in artifacts under lab

conditions. The authors experimentally validated this by characterizing the intermediate, elansolid A3 (**1.7**) (Figure 1.32a).³⁷ Knowledge of **1.7** was then exploited in a medicinal chemistry effort and a biosynthetic investigation in the producing organism. This exemplifies how determining the presence of these reactions, endogenous or not, is crucial for further exploration.

a. Structures of the elansolid family



b. Structure of the rubrolones



c. Abbreviated mechanism for the pyridine formation of the rubrolones

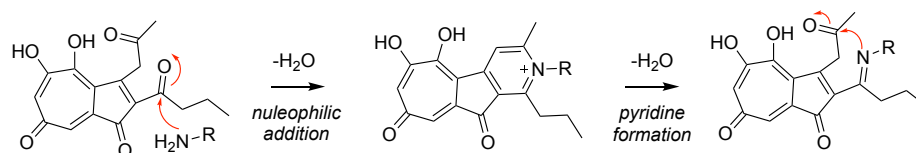


Figure 1.32. Recapitulated figures of the elansolids and rubrolones. a) Structures of the elansolid family and the key reactive intermediate elansolid A3. b) Structures of the rubrolones. c) Partial mechanism of the pyridine core formation of the rubrolones.^{36-39, 109}

The core of this story is the proper validation of non-enzymatic reactions, specifically the two major constituents of any claim are mechanism and function. It is paramount that one demonstrates analytically that a reaction can occur in solvent free of protein catalysts and an

effort must be made to determine that the reaction serves a biological benefit to the organism. The latter pursuit is much more difficult and is commonly attempted with indirect qualitative data. An example of these indirect validation is represented in the story of the rubrolones (**1.65** and **1.66**) (Figure 1.32b).¹⁰⁹ A non-enzymatic reaction is required to generate the substrate for downstream glycosylation. As glycosylation is needed for antibiotic activity, the non-enzymatic pyridine formation is essential for the biological function of the rubrolones (Figure 1.32c). This suggests that this non-enzymatic reaction can truly be considered a component of the biosynthesis rather than a side product.¹⁰⁹ Regardless of their complexity, mechanistic and functional studies should be pursued for all proposed non-enzymatic biosynthetic reactions.

1.8 Discussion of common precursors

Just as enzymes have evolved to improve the efficiency of biologically important reactions, it is possible that chemical reactivity is inherent in particular substrates. An attractive hypothesis for such a design would be the repurposing of primary metabolites for secondary metabolic function that would eliminate the need for a new set of anabolic genes. Evidence for this claim is shown by the reoccurrence of several substrates in diverse biosynthetic pathways.

Anthranilic acid appears in oxazinin A (**1.13**), the discoipyrroles (**1.51**), rubrolone B (**1.66**) and the juanlimycins (**1.113**) (Figure 1.33).^{14, 53, 56, 106, 109, 123} Additionally, this compound has shown to react and incorporate into the elansolids (Figure 1.2) and the ammosamides (Figure 1.23) for the generation of non-endogenous metabolites.^{40, 119} The presence of anthranilic acid is not rare, as it is a common by-product of primary metabolism via the kynurinate and the tryptophanase pathways. However, it might be possible that upregulation of anthranilate, which results in an environmental fitness, could allow for said upregulation to be conserved. The power of anthranilate as well as other nucleophilic amines lies in their dual

function as 1. initiators of new bonds by nucleophilic addition into carbonyls (juanlimycins, discoipyrroles, rubrolone B, oxazinin A, and jadomycins) and 2. generators of secondary reactive sites including electrophilic moieties (imine and iminium carbons: discoipyrroles, oxazinin A, ammosamides) and secondary nucleophiles (secondary amines: discoipyrroles, oxazinin A, and rubrolone B).

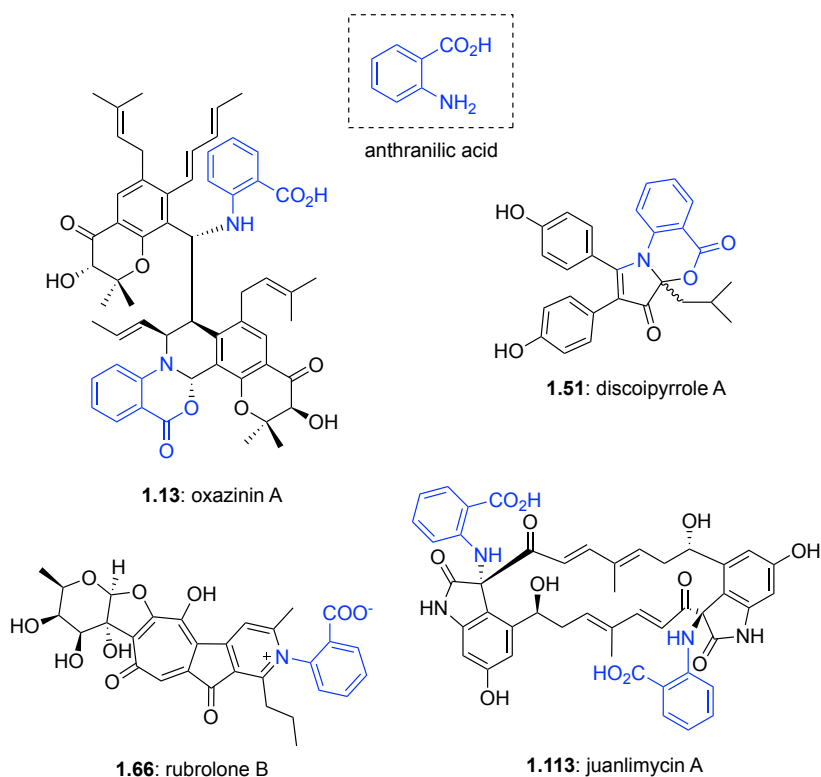


Figure 1.33. Natural product scaffolds that non-enzymatically incorporate anthranilic acid. Highlighted in blue is the incorporation of anthranilic acid into the mature natural products of oxazinin A, discoipyrrole A, rubrolone B, and juanlimycin A. ^{14, 53, 56, 106, 109, 123}

Another common reactive functionality found in non-enzymatic natural product biosynthesis are polyenes, particularly in the context of cycloaddition reactions as discussed in the intramolecular, dimerizations, and light-induced categories. The power of olefins in

synthetic chemistry as a whole is well-known and the number of enzymatic processes involving conjugated systems of double bonds is overwhelming, therefore, it is inaccurate to say that this functional group is a signature of non-enzymatic processes. However, when mature natural products with multiple olefins are found alongside heavily cyclized natural products in extracts of ecologically sensible organisms, it is important to interrogate the relationships between those reactions and their biological functions. It could be said that any light-induced non-enzymatic reaction shown to occur in a biological setting reveals a potential functional role for these natural products as molecular sunscreens or radical scavengers, a relationship that would bolster an argument for conserved non-enzymatic biosynthesis.

A third fundamental group found in both non-enzymatic biosynthesis and organic synthesis are the α,β -unsaturated ketones. These functional groups are encountered as both nucleophilic and electrophilic components, commonly as Michael acceptors. As with the polyenes, α,β -unsaturated ketones are commonly derived from polyketide biosynthesis, suggesting that most reactions should be enzymatically driven. However, conjugated cyclic systems especially are indicative of mature or late-stage natural products that can engage in non-enzymatic reactions.

The last two examples of common moieties subject to non-enzymatic chemistry demonstrate the intimate relationship between synthetic chemistry and natural product biochemistry. It is no surprise to any chemist that reactions in biological systems in the absence of protein catalysts are the same as those that are integral to performing basic organic chemistry. This is the most important reasoning for the number of concerted non-enzymatic reactions that are discovered through the eyes of observant synthetic chemists. Beyond the simple classification of reactions as either enzymatic or non-enzymatic, it is important to incorporate these synthetic tools to identify reactions that can be applied in total synthesis and/or medicinal chemistry, two fields that directly benefit biomedicine.

In addition to the fundamental aspect of identifying non-enzymatic reactions, biological insights can be gained from studying these mechanisms. This leads to the discussion of the common functional group that is harder to define, that of the biological toxin. This was exemplified by four major stories: the tridachiahypopyrones (**1.124**), the dibohemamines (**1.39**), ($\pm$)-homodimericin A (**1.29**), and the maximiscins (**1.76**), and the tridachiahypopyrones (Figure 1.34). Each of these natural products are derived from a presumed biological detoxification of a common toxin whether it be formaldehyde, reactive oxygen species, UV-light, or other general toxins (Figure 1.34). The nuances of their biological functions require further mechanistic studies, although the preliminary data for each of these natural products is certainly intriguing and these are informative for future studies. When examining natural product biosynthesis, it is imperative to consider the incorporation of these common toxins among others in the absence of genetic information regarding their source.

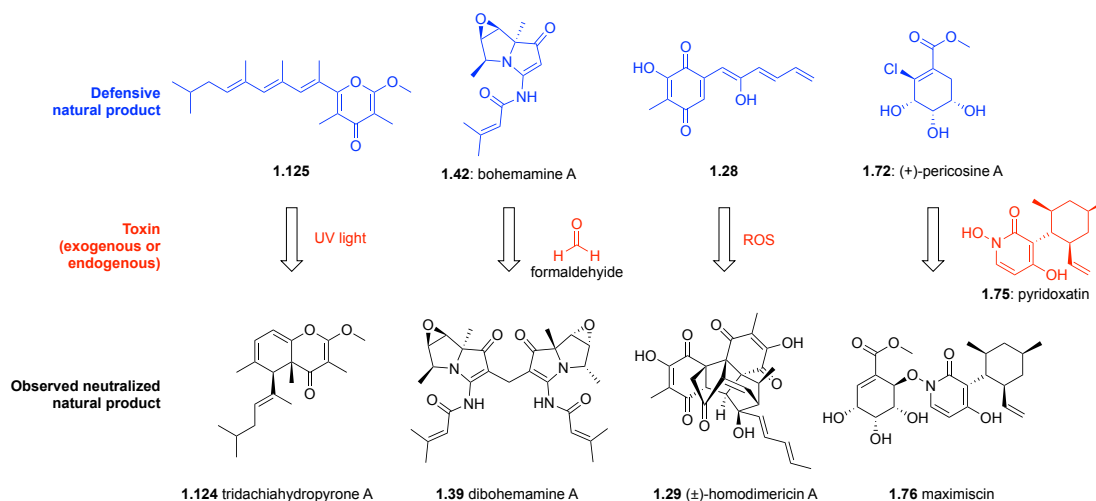


Figure 1.34. Representation of different biological applications of non-enzymatic mechanisms. Examples of precursor molecules that undergo a non-enzymatic transformation in response to a toxin or other biological factor to create new molecular scaffolds.^{72, 83, 95, 110, 111, 132}

While there are functional moieties that are prevalent in the chemistry described above, it is also worth noting that a large number of scaffolds that take place in this chemistry are polyketides. This is likely due to the presence of polyenes, enones, and other functional groups that promote non-enzymatic chemistry.

1.9 Conclusion

Although non-enzymatic mechanisms have been historically relevant, the growing number of occurrences reported in recent years demonstrates the need for a conscious effort for the accurate and mechanistic assessment of these reactions in natural product biosynthesis. Reactions ranging from synthetically important reactions, such as intramolecular cyclizations, to novel multi-component reactions that participate in biological functions can all be arranged within the concept of non-enzymatic chemistry.

In conclusion, it is most important to understand the scope and breadth of these types of reactions. Non-enzymatic biosynthetic steps are the vast minority of construction mechanisms for natural products. The superior efficiency, regulations, and complexity of protein machinery cannot be overstated. Enzymatic processes have evolved over time to be responsible for the chemistry we isolate daily. However, there exists the possibility that the inherent reactivity of the biological milieu is used to convey important signals for cell survival as seen in many eukaryotic systems such as the non-enzymatic acetylation of proteins, detoxification of glutathione, and oxidative regulation of the HIF system.¹⁴²⁻¹⁴⁶ With that being said, every detail, as minute or rare as it may be, must be considered in order to have a holistic understanding of how organisms use simple building blocks to construct complex molecules.

1.10 References

1. Li, D.; Azoulay, P.; Sampat, B. N. The applied value of public investments in biomedical research. *Science* **2017**, 356 (6333), 78-81.
2. Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **2002**, 417 (6885), 141-147.
3. Rutledge, P. J.; Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology* **2015**, 13 (8), 509-523.
4. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, 227 (5258), 561-563.
5. Hopwood, D. A. The Leeuwenhoek Lecture, 1987 - Towards an understanding of gene switching in *Streptomyces*, the basis of sporulation and antibiotic production. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **1988**, 235 (1279), 121-138.
6. Jensen, P. R. Natural Products and the Gene Cluster Revolution. *Trends in Microbiology* **2016**, 24 (12), 968-977.
7. Courdavault, V.; O'Connor, S. E.; Jensen, M. K.; Papon, N. Metabolic engineering for plant natural products biosynthesis: new procedures, concrete achievements and remaining limits. *Natural Product Reports* **2021**, 38 (12), 2145-2153.
8. de Lorenzo, V.; Sekowska, A.; Danchin, A. Chemical reactivity drives spatiotemporal organisation of bacterial metabolism. *FEMS Microbiology Reviews* **2015**, (39), 96-119.
9. Shi, L.; Tu, B. P. Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. *Current Opinion in Cell Biology* **2015**, 33, 125-131.
10. Keller, M. A.; Piedrafitá, G.; Ralser, M. The widespread role of non-enzymatic reactions in cellular metabolism. *Current Opinion in Biotechnology* **2015**, 34, 153-161.
11. Weiss, M. Crowding, Diffusion, and Biochemical Reactions. In *New Models of the Cell Nucleus: Crowding, Entropic Forces, Phase Separation, and Fractals*, International Review of Cell and Molecular Biology, 2014; pp 383-417.
12. Wallace, S.; Schultz, E. E.; Balskus, E. P. Using non-enzymatic chemistry to influence microbial metabolism. *Current Opinion in Chemical Biology* **2015**, 25, 71-79.
13. Gravel, E.; Poupon, E. Biogenesis and Biomimetic Chemistry: Can Complex Natural Products Be Assembled Spontaneously? *European Journal of Organic Chemistry* **2008**, 2008 (1), 27-42.

14. Colosimo, D. A.; MacMillan, J. B. Detailed Mechanistic Study of the Non-enzymatic Formation of the Discoipyrrole Family of Natural Products. *Journal of the American Chemical Society* **2016**, *138* (7), 2383-2388.
15. Czechowski, T.; Larson, T. R.; Catania, T. M.; Harvey, D.; Brown, G. D.; Graham, I. A. Artemisia annua mutant impaired in artemisinin synthesis demonstrates importance of nonenzymatic conversion in terpenoid metabolism. *Proceedings of the National Academy of Sciences* **2016**, *113* (52), 15150-15155.
16. Leao, T.; Castelh o, G.; Korobeynikov, A.; Monroe, E. A.; Podell, S.; Glukhov, E.; Allen, E. E.; Gerwick, W. H.; Gerwick, L. Comparative genomics uncovers the prolific and distinctive metabolic potential of the cyanobacterial genus Moorea. *Proceedings of the National Academy of Sciences* **2017**, *114* (12), 3198-3203.
17. Jordan, P. A.; Moore, B. S. Biosynthetic Pathway Connects Cryptic Ribosomally Synthesized Posttranslationally Modified Peptide Genes with Pyrroloquinoline Alkaloids. *Cell Chemical Biology* **2016**, *23* (12), 1504-1514.
18. Beaudry, C. M.; Malerich, J. P.; Trauner, D. Biosynthetic and Biomimetic Electrocyclizations. *Chemical Reviews* **2005**, *105* (12), 4757-4778.
19. Nicolaou, K. C.; Montagnon, T.; Snyder, S. A. Tandem reactions, cascade sequences, and biomimetic strategies in total synthesis. *Chemical Communications* **2003**, (5), 551-564.
20. Razzak, M.; De Brabander, J. K. Lessons and revelations from biomimetic syntheses. *Nature Chemical Biology* **2011**, *7* (12), 865-875.
21. Bandaranayake, W. M.; Banfield, J. E.; Black, D. S. C. Postulated electrocyclic reactions leading to endiandric acid and related natural products. *Journal of the Chemical Society, Chemical Communications* **1980**, (19), 902-903.
22. Nicolaou, K. C.; Petasis, N. A.; Zipkin, R. E. The endiandric acid cascade. Electrocyclizations in organic synthesis. 4. Biomimetic approach to endiandric acids A-G. Total synthesis and thermal studies. *Journal of the American Chemical Society* **1982**, *104* (20), 5560-5562.
23. Juhl, M.; Tanner, D. Recent applications of intramolecular Diels–Alder reactions to natural product synthesis. *Chemical Society Reviews* **2009**, *38* (11), 2983-2992.
24. Stocking, E. M.; Williams, R. M. Chemistry and Biology of Biosynthetic Diels–Alder Reactions. *Angewandte Chemie International Edition* **2003**, *42* (27), 3078-3115.
25. Chook, Y. M.; Ke, H.; Lipscomb, W. N. Crystal structures of the monofunctional chorismate mutase from Bacillus subtilis and its complex with a transition state analog. *Proceedings of the National Academy of Sciences* **1993**, *90* (18), 8600-8603.
26. Kasahara, K.; Miyamoto, T.; Fujimoto, T.; Oguri, H.; Tokiwano, T.; Oikawa, H.; Ebizuka, Y.; Fujii, I. Solanapyrone Synthase, a Possible Diels–Alderase and Iterative Type I Polyketide Synthase Encoded in a Biosynthetic Gene Cluster from Alternaria solani. *ChemBioChem* **2010**, *11* (9), 1245-1252.

27. Kim, R.-R.; Illarionov, B.; Joshi, M.; Cushman, M.; Lee, C. Y.; Eisenreich, W.; Fischer, M.; Bacher, A. Mechanistic Insights on Riboflavin Synthase Inspired by Selective Binding of the 6,7-Dimethyl-8-ribityllumazine Exomethylene Anion. *Journal of the American Chemical Society* **2010**, *132* (9), 2983-2990.
28. Medvedev, M. G.; Zeifman, A. A.; Novikov, F. N.; Bushmarinov, I. S.; Stroganov, O. V.; Titov, I. Y.; Chilov, G. G.; Svitanko, I. V. Quantifying Possible Routes for SpnF-Catalyzed Formal Diels–Alder Cycloaddition. *Journal of the American Chemical Society* **2017**, *139* (11), 3942-3945.
29. Jeon, B.-s.; Ruzsyczky, M. W.; Russell, W. K.; Lin, G.-M.; Kim, N.; Choi, S.-h.; Wang, S.-A.; Liu, Y.-n.; Patrick, J. W.; Russell, D. H.; Liu, H.-w. Investigation of the mechanism of the SpnF-catalyzed [4+2]-cycloaddition reaction in the biosynthesis of spinosyn A. *Proceedings of the National Academy of Sciences* **2017**, *114* (39), 10408-10413.
30. Cai, Y.; Hai, Y.; Ohashi, M.; Jamieson, C. S.; Garcia-Borras, M.; Houk, K. N.; Zhou, J.; Tang, Y. Structural basis for stereoselective dehydration and hydrogen-bonding catalysis by the SAM-dependent pericyclase Lepl. *Nature Chemistry* **2019**, *11* (9), 812-820.
31. Serafimov, J. M.; Gillingham, D.; Kuster, S.; Hilvert, D. The Putative Diels–Alderase Macrophomate Synthase is an Efficient Aldolase. *Journal of the American Chemical Society* **2008**, *130* (25), 7798-7799.
32. Watanabe, K. Discovery and investigation of natural Diels–Alderase. *Journal of Natural Medicines* **2021**, *75* (3), 434-447.
33. Klas, K.; Tsukamoto, S.; Sherman, D. H.; Williams, R. M. Natural Diels–Alderase: Elusive and Irresistible. *The Journal of Organic Chemistry* **2015**, *80* (23), 11672-11685.
34. Raju, R.; Piggott, A. M.; Conte, M. M.; Capon, R. J. Heronamides A–C, new polyketide macrolactams from an Australian marine-derived *Streptomyces* sp. A biosynthetic case for synchronized tandem electrocyclization. *Organic & Biomolecular Chemistry* **2010**, *8* (20), 4682-4689.
35. Webster, R.; Gaspar, B.; Mayer, P.; Trauner, D. Studies toward the Biomimetic Total Synthesis of (–)-PF-1018. *Organic Letters* **2013**, *15* (8), 1866-1869.
36. Steinmetz, H.; Gerth, K.; Jansen, R.; Schläger, N.; Dehn, R.; Reinecke, S.; Kirschning, A.; Müller, R. Elansolid A, a Unique Macrolide Antibiotic from *Chitinophaga sancti* Isolated as Two Stable Atropisomers. *Angewandte Chemie International Edition* **2011**, *50* (2), 532-536.
37. Jansen, R.; Gerth, K.; Steinmetz, H.; Reinecke, S.; Kessler, W.; Kirschning, A.; Müller, R. Elansolid A3, a Unique p-Quinone Methide Antibiotic from *Chitinophaga sancti*. *Chemistry – A European Journal* **2011**, *17* (28), 7739-7744.
38. Steinmetz, H.; Zander, W.; Shushni, M. A. M.; Jansen, R.; Gerth, K.; Dehn, R.; Dräger, G.; Kirschning, A.; Müller, R. Precursor-Directed Syntheses and Biological Evaluation of New Elansolid Derivatives. *ChemBioChem* **2012**, *13* (12), 1813-1817.

39. Wang, L.-L.; Kirschning, A. Total synthesis of elansolids B1 and B2. *Beilstein Journal of Organic Chemistry* **2017**, *13*, 1280-1287.
40. Dehn, R.; Katsuyama, Y.; Weber, A.; Gerth, K.; Jansen, R.; Steinmetz, H.; Höfle, G.; Müller, R.; Kirschning, A. Molecular Basis of Elansolid Biosynthesis: Evidence for an Unprecedented Quinone Methide Initiated Intramolecular Diels–Alder Cycloaddition/Macrolactonization. *Angewandte Chemie International Edition* **2011**, *50* (17), 3882-3887.
41. Brophy, G. C.; Mohandas, J.; Slaytor, M.; Sternhell, S.; Watson, T. R.; Wilson, L. A. Novel lignans from a sp. from bougainville. *Tetrahedron Letters* **1969**, *10* (59), 5159-5162.
42. Chapman, O. L.; Engel, M. R.; Springer, J. P.; Clardy, J. C. Total synthesis of carpanone. *Journal of the American Chemical Society* **2002**, *93* (24), 6696-6698.
43. Abe, N.; Murata, T.; Hirota, A. Novel DPPH Radical Scavengers, Bisorbicillinol and Demethyltrichodimerol, from a Fungus. *Bioscience, Biotechnology, and Biochemistry* **1998**, *62* (4), 661-666.
44. Lee, J. C.; Strobel, G. A.; Lobkovsky, E.; Clardy, J. Torreyanic Acid: A Selectively Cytotoxic Quinone Dimer from the Endophytic Fungus *Pestalotiopsis microspora*. *The Journal of Organic Chemistry* **1996**, *61* (10), 3232-3233.
45. Lee, J. C.; Yang, X.; Schwartz, M.; Strobel, G.; Clardy, J. The relationship between an endangered North American tree and an endophytic fungus. *Chemistry & Biology* **1995**, *2* (11), 721-727.
46. Artman, G. D.; Grubbs, A. W.; Williams, R. M. Concise, Asymmetric, Stereocontrolled Total Synthesis of Stephacidins A, B and Notoamide B. *Journal of the American Chemical Society* **2007**, *129* (19), 6336-6342.
47. Baran, P. S.; Hafensteiner, B. D.; Ambhaikar, N. B.; Guerrero, C. A.; Gallagher, J. D. Enantioselective Total Synthesis of Avrainvillamide and the Stephacidins. *Journal of the American Chemical Society* **2006**, *128* (26), 8678-8693.
48. Baran, P. S.; Guerrero, C. A.; Hafensteiner, B. D.; Ambhaikar, N. B. Total Synthesis of Avrainvillamide (CJ-17,665) and Stephacidin B. *Angewandte Chemie International Edition* **2005**, *44* (25), 3892-3895.
49. Herzon, S. B.; Myers, A. G. Enantioselective Synthesis of Stephacidin B. *Journal of the American Chemical Society* **2005**, *127* (15), 5342-5344.
50. Grube, A.; Immel, S.; Baran, P. S.; Köck, M. Massadine Chloride: A Biosynthetic Precursor of Massadine and Stylissadine. *Angewandte Chemie International Edition* **2007**, *46* (35), 6721-6724.
51. Nicolaou, K. C.; Papageorgiou, C. D.; Piper, J. L.; Chadha, R. K. The Cytoskyrin Cascade: A Facile Entry into Cytoskyrin A, Deoxyrubroskyrin, Rugulin, Skyrin, and Flavoskyrin Model Systems. *Angewandte Chemie International Edition* **2005**, *44* (36), 5846-5851.

52. Nicolaou, K. C.; Lim, Y. H.; Piper, J. L.; Papageorgiou, C. D. Total Syntheses of 2,2'-epi-Cytoskyrin A, Rugulosin, and the Alleged Structure of Rugulin. *Journal of the American Chemical Society* **2007**, 129 (13), 4001-4013.
53. Lin, Z.; Koch, M.; Abdel Aziz, M. H.; Galindo-Murillo, R.; Tianero, M. D.; Cheatham, T. E.; Barrows, L. R.; Reilly, C. A.; Schmidt, E. W. Oxazinin A, a Pseudodimeric Natural Product of Mixed Biosynthetic Origin from a Filamentous Fungus. *Organic Letters* **2014**, 16 (18), 4774-4777.
54. Tanaka, K.; Sasaki, A.; Cao, H. Q.; Yamada, T.; Igarashi, M.; Komine, I.; Nakahigashi, H.; Minami, N.; Kuwahara, S.; Nukina, M.; Kiyota, H. Synthesis and Biotransformation of Plausible Biosynthetic Intermediates of Salicylaldehyde-Type Phytotoxins of Rice Blast Fungus, *Magnaporthe grisea*. *European Journal of Organic Chemistry* **2011**, 2011 (31), 6276-6280.
55. Teles, H. L.; Sordi, R.; Silva, G. H.; Castro-Gamboa, I.; Bolzani, V. d. S.; Pfenning, L. H.; Magalhães de Abreu, L.; Costa-Neto, C. M.; Young, M. C. M.; Araújo, A. R. Aromatic compounds produced by *Periconia atropurpurea*, an endophytic fungus associated with *Xylopiia aromatica*. *Phytochemistry* **2006**, 67 (24), 2686-2690.
56. Aniebok, V.; Shingare, R. D.; Wei-Lee, H.; Johnstone, T. C.; MacMillan, J. B. Biomimetic Total Synthesis and Investigation of the Non-Enzymatic Chemistry of Oxazinin A. *Angewandte Chemie International Edition* **2022**, 61 (38), e202208029.
57. Chen, X.-L.; Liu, H.-L.; Li, J.; Xin, G.-R.; Guo, Y.-W. Paracaseolide A, First α -Alkylbutenolide Dimer with an Unusual Tetraquinane Oxa-Cage Bislactone Skeleton from Chinese Mangrove *Sonneratia paracaseolaris*. *Organic Letters* **2011**, 13 (19), 5032-5035.
58. Noutsias, D.; Vassilikogiannakis, G. First Total Synthesis of Paracaseolide A. *Organic Letters* **2012**, 14 (14), 3565-3567.
59. Yin, J.-P.; Tang, C.-L.; Gao, L.-X.; Ma, W.-P.; Li, J.-Y.; Li, Y.; Li, J.; Nan, F.-J. Design and synthesis of paracaseolide A analogues as selective protein tyrosine phosphatase 1B inhibitors. *Organic & Biomolecular Chemistry* **2014**, 12 (21), 3441-3445.
60. Giera, D. S.; Stark, C. B. W. Total synthesis of ($\pm$)-paracaseolide A and initial attempts at a Lewis acid mediated dimerization of its putative biosynthetic precursor. *RSC Adv.* **2013**, 3 (44), 21280-21284.
61. Vasamsetty, L.; Khan, F. A.; Mehta, G. Total synthesis of a novel oxa-bowl natural product paracaseolide A via a 'putative' biomimetic pathway. *Tetrahedron Letters* **2013**, 54 (27), 3522-3525.
62. Vasamsetty, L.; Sahu, D.; Ganguly, B.; Khan, F. A.; Mehta, G. Total synthesis of novel bioactive natural product paracaseolide A and analogues: computational evaluation of a 'proposed' biomimetic Diels–Alder reaction. *Tetrahedron* **2014**, 70 (45), 8488-8497.
63. Wang, T.; Hoye, T. R. Diels–Alderase-free, bis-pericyclic, [4+2] dimerization in the biosynthesis of ($\pm$)-paracaseolide A. *Nature Chemistry* **2015**, 7 (8), 641-645.

64. Gupta, K.; Chakrabarti, A.; Rana, S.; Ramdeo, R.; Roth, B. L.; Agarwal, M. L.; Tse, W.; Agarwal, M. K.; Wald, D. N. Securinine, a Myeloid Differentiation Agent with Therapeutic Potential for AML. *PLoS ONE* **2011**, 6 (6), e21203.
65. Weinreb, S. M. Total synthesis of the Securinega alkaloids. *Natural Product Reports* **2009**, 26 (6), 758-775.
66. Jeon, S.; Han, S. An Accelerated Intermolecular Rauht–Currier Reaction Enables the Total Synthesis of (–)-Flueggeine C. *Journal of the American Chemical Society* **2017**, 139 (18), 6302-6305.
67. Gan, L.-S.; Fan, C.-Q.; Yang, S.-P.; Wu, Y.; Lin, L.-P.; Ding, J.; Yue, J.-M. Flueggeines A and B, Two Novel C,C-Linked Dimeric Indolizidine Alkaloids from *Flueggea virosa*. *Organic Letters* **2006**, 8 (11), 2285-2288.
68. Zhao, B.-X.; Wang, Y.; Zhang, D.-M.; Jiang, R.-W.; Wang, G.-C.; Shi, J.-M.; Huang, X.-J.; Chen, W.-M.; Che, C.-T.; Ye, W.-C. Flueggines A and B, Two New Dimeric Indolizidine Alkaloids from *Flueggea virosa*. *Organic Letters* **2011**, 13 (15), 3888-3891.
69. Zhang, H.; Zhang, C.-R.; Zhu, K.-K.; Gao, A.-H.; Luo, C.; Li, J.; Yue, J.-M. Fluevirosines A–C: A Biogenesis Inspired Example in the Discovery of New Bioactive Scaffolds from *Flueggea virosa*. *Organic Letters* **2012**, 15 (1), 120-123.
70. Zhao, B.-X.; Wang, Y.; Li, C.; Wang, G.-C.; Huang, X.-J.; Fan, C.-L.; Li, Q.-M.; Zhu, H.-J.; Chen, W.-M.; Ye, W.-C. Flueggedine, a novel axisymmetric indolizidine alkaloid dimer from *Flueggea virosa*. *Tetrahedron Letters* **2013**, 54 (35), 4708-4711.
71. Wei, H.; Qiao, C.; Liu, G.; Yang, Z.; Li, C. c. Stereoselective Total Syntheses of (–)-Flueggine A and (+)-Virosaine B. *Angewandte Chemie International Edition* **2013**, 52 (2), 620-624.
72. Mevers, E.; Saurí, J.; Liu, Y.; Moser, A.; Ramadhar, T. R.; Varlan, M.; Williamson, R. T.; Martin, G. E.; Clardy, J. Homodimericin A: A Complex Hexacyclic Fungal Metabolite. *Journal of the American Chemical Society* **2016**, 138 (38), 12324-12327.
73. Harned, A. M.; Volp, K. A. The sorbicillinoid family of natural products: Isolation, biosynthesis, and synthetic studies. *Natural Product Reports* **2011**, 28 (11), 1790-1810.
74. Trifonov, L. S.; Dreiding, A. S.; Hoesch, L.; Rast, D. M. Isolation of Four Hexaketides from *Verticillium intertextum*. Preliminary communication. *Helvetica Chimica Acta* **1981**, 64 (6), 1843-1846.
75. Ma, D.; Liu, Y.; Wang, Z. Biomimetic Total Synthesis of (±)-Homodimericin A. *Angewandte Chemie International Edition* **2017**, 56 (27), 7886-7889.
76. Huang, J.; Gu, Y.; Guo, K.; Zhu, L.; Lan, Y.; Gong, J.; Yang, Z. Bioinspired Total Synthesis of Homodimericin A. *Angewandte Chemie International Edition* **2017**, 56 (27), 7890-7894.
77. Feng, J.; Lei, X.; Guo, Z.; Tang, Y. Total Synthesis of Homodimericin A. *Angewandte Chemie International Edition* **2017**, 56 (27), 7895-7899.

78. Nishikawa, H.; Miyazaki, T.; Nakayama, H.; Minematsu, A.; Yamauchi, S.; Yamashita, K.; Takazono, T.; Shimamura, S.; Nakamura, S.; Izumikawa, K.; et al. Roles of vacuolar H⁺-ATPase in the oxidative stress response of *Candida glabrata*. *FEMS Yeast Research* **2016**, *16* (5), fow054.
79. Yokomakura, A.; Hong, J.; Ohuchi, K.; Oh, S.-E.; Lee, J.-Y.; Mano, N.; Takahashi, T.; Hwang, G.-W.; Naganuma, A. Increased production of reactive oxygen species by the vacuolar-type (H⁺)-ATPase inhibitors bafilomycin A1 and concanamycin A in RAW 264 cells. *The Journal of Toxicological Sciences* **2012**, *37* (5), 1045-1048.
80. Abe, N.; Murata, T.; Yamamoto, K.; Hirota, A. Bisorbibetanone, a novel oxidized sorbicillin dimer, with 1,1-diphenyl-2-picrylhydrazyl radical scavenging activity from a fungus. *Tetrahedron Letters* **1999**, *40* (28), 5203-5206.
81. Abe, N.; Hirota, A. Chemical studies of the radical scavenging mechanism of bisorbicillinol using the 1,1-diphenyl-2-picrylhydrazyl radical. *Chemical Communications* **2002**, (6), 662-663.
82. Robertson, J.; Stevens, K. Pyrrolizidine alkaloids: occurrence, biology, and chemical synthesis. *Natural Product Reports* **2017**, *34* (1), 62-89.
83. Fu, P.; Legako, A.; La, S.; MacMillan, J. B. Discovery, Characterization, and Analogue Synthesis of Bohemamine Dimers Generated by Non-enzymatic Biosynthesis. *Chemistry – A European Journal* **2016**, *22* (10), 3491-3495.
84. Piel, J.; Hertweck, C.; Shipley, P. R.; Hunt, D. M.; Newman, M. S.; Moore, B. S. Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate 'Streptomyces maritimus': evidence for the derailment of an aromatic polyketide synthase. *Chemistry & Biology* **2000**, *7* (12), 943-955.
85. Kalaitzis, J. A. Discovery, Biosynthesis, and Rational Engineering of Novel Enterocin and Wailupemycin Polyketide Analogues. In *Metabolomics Tools for Natural Product Discovery*, Methods in Molecular Biology, 2013; pp 171-189.
86. Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. Enzymatic total synthesis of enterocin polyketides. *Nature Chemical Biology* **2007**, *3* (9), 557-558.
87. Kalaitzis, J. A.; Cheng, Q.; Thomas, P. M.; Kelleher, N. L.; Moore, B. S. In Vitro Biosynthesis of Unnatural Enterocin and Wailupemycin Polyketides. *Journal of Natural Products* **2009**, *72* (3), 469-472.
88. Kalaitzis, J. A.; Izumikawa, M.; Xiang, L.; Hertweck, C.; Moore, B. S. Mutasynthesis of Enterocin and Wailupemycin Analogues. *Journal of the American Chemical Society* **2003**, *125* (31), 9290-9291.
89. Shao, C.-L.; Wang, C.-Y.; Wei, M.-Y.; Gu, Y.-C.; She, Z.-G.; Qian, P.-Y.; Lin, Y.-C. Aspergilones A and B, two benzylazaphilones with an unprecedented carbon skeleton from the gorgonian-derived fungus *Aspergillus* sp. *Bioorganic & Medicinal Chemistry Letters* **2011**, *21* (2), 690-693.
90. Arisawa, M.; Fujita, A.; Hayashi, T.; Hayashi, K.; Ochiai, H.; Morita, N. Cytotoxic and antiherpetic activity of phloroglucinol derivatives from *Mallotus japonicus* (Euphorbiaceae). *Chemical and Pharmaceutical Bulletin* **1990**, *38* (6), 1624-1626.

91. Eaton, A. L.; Dalal, S.; Cassera, M. B.; Zhao, S.; Kingston, D. G. I. Synthesis and Antimalarial Activity of Mallatojaponin C and Related Compounds. *Journal of Natural Products* **2016**, 79 (6), 1679-1683.
92. Chauthe, S. K.; Bharate, S. B.; Sabde, S.; Mitra, D.; Bhutani, K. K.; Singh, I. P. Biomimetic synthesis and anti-HIV activity of dimeric phloroglucinols. *Bioorganic & Medicinal Chemistry* **2010**, 18 (5), 2029-2036.
93. Minassi, A.; Cicione, L.; Koeberle, A.; Bauer, J.; Laufer, S.; Werz, O.; Appendino, G. A Multicomponent Carba-Betti Strategy to Alkylidene Heterodimers – Total Synthesis and Structure–Activity Relationships of Arzanol. *European Journal of Organic Chemistry* **2012**, 2012 (4), 772-779.
94. Penttilä, A.; Kapadia, G. J.; Fales, H. M. The Biosynthesis in Vivo of Methylenebisphloroglucinol Derivatives. *Journal of the American Chemical Society* **1965**, 87 (19), 4402-4403.
95. Chen, N. H.; Djoko, K. Y.; Veyrier, F. J.; McEwan, A. G. Formaldehyde Stress Responses in Bacterial Pathogens. *Frontiers in Microbiology* **2016**, 7.
96. Doull, J. L.; Ayer, S. W.; Singh, A. K.; Thibault, P. Production of a novel polyketide antibiotic, jadomycin B, by *Streptomyces venezuelae* following heat shock. *The Journal of Antibiotics* **1993**, 46 (5), 869-871.
97. Doull, J. L.; Singh, A. K.; Hoare, M.; Ayer, S. W. Conditions for the production of jadomycin B by *Streptomyces venezuelae* ISP5230: Effects of heat shock, ethanol treatment and phage infection. *Journal of Industrial Microbiology* **1994**, 13 (2), 120-125.
98. Sharif, E. U.; O'Doherty, G. A. Biosynthesis and Total Synthesis Studies on the Jadomycin Family of Natural Products. *European Journal of Organic Chemistry* **2012**, 2012 (11), 2095-2108.
99. Jakeman, D. L.; Bandi, S.; Graham, C. L.; Reid, T. R.; Wentzell, J. R.; Douglas, S. E. Antimicrobial Activities of Jadomycin B and Structurally Related Analogues. *Antimicrobial Agents and Chemotherapy* **2009**, 53 (3), 1245-1247.
100. Rix, U.; Zheng, J.; Remsing Rix, L. L.; Greenwell, L.; Yang, K.; Rohr, J. The Dynamic Structure of Jadomycin B and the Amino Acid Incorporation Step of Its Biosynthesis. *Journal of the American Chemical Society* **2004**, 126 (14), 4496-4497.
101. Robertson, A. W.; Martinez-Farina, C. F.; Smithen, D. A.; Yin, H.; Monroe, S.; Thompson, A.; McFarland, S. A.; Syvitski, R. T.; Jakeman, D. L. Eight-Membered Ring-Containing Jadomycins: Implications for Non-enzymatic Natural Products Biosynthesis. *Journal of the American Chemical Society* **2015**, 137 (9), 3271-3275.
102. Forget, S. M.; Robertson, A. W.; Hall, S. R.; MacLeod, J. M.; Overy, D. P.; Kerr, R. G.; Goralski, K. B.; Jakeman, D. L. Isolation of a jadomycin incorporating L-ornithine, analysis of antimicrobial activity and jadomycin reactive oxygen species (ROS) generation in MDA-MB-231 breast cancer cells. *The Journal of Antibiotics* **2018**, 71 (8), 722-730.

103. Dupuis, S. N.; Veinot, T.; Monro, S. M. A.; Douglas, S. E.; Syvitski, R. T.; Goralski, K. B.; McFarland, S. A.; Jakeman, D. L. Jadomycins Derived from the Assimilation and Incorporation of Norvaline and Norleucine. *Journal of Natural Products* **2011**, *74* (11), 2420-2424.
104. Martinez-Farina, C. F.; Jakeman, D. L. Jadomycins, put a bigger ring in it: isolation of seven- to ten-membered ring analogues. *Chemical Communications* **2015**, *51* (78), 14617-14619.
105. Jakeman, D. L.; Graham, C. L.; Reid, T. R. Novel and expanded jadomycins incorporating non-proteogenic amino acids. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15* (23), 5280-5283.
106. Hu, Y.; Potts, M. B.; Colosimo, D.; Herrera-Herrera, M. L.; Legako, A. G.; Yousufuddin, M.; White, M. A.; MacMillan, J. B. Discoipyrroles A–D: Isolation, Structure Determination, and Synthesis of Potent Migration Inhibitors from *Bacillus hunanensis*. *Journal of the American Chemical Society* **2013**, *135* (36), 13387-13392.
107. Palleroni, N. J.; Reichelt, K. E.; Mueller, D.; Epps, R.; Tabenkin, B.; Bull, D. N.; Schüep, W.; Berger, J. Production of a novel red pigment, rubrolone, by *Streptomyces echinoruber* sp. nov. I. Taxonomy, fermentation and partial purification. *The Journal of Antibiotics* **1978**, *31* (12), 1218-1225.
108. Schüep, W.; Blount, J. F.; Williams, T. H.; Stempel, A. Production of a novel red pigment, rubrolone, by *Streptomyces echinoruber* sp. nov. II. Chemistry and structure elucidation. *The Journal of Antibiotics* **1978**, *31* (12), 1226-1232.
109. Yan, Y.; Yang, J.; Yu, Z.; Yu, M.; Ma, Y.-T.; Wang, L.; Su, C.; Luo, J.; Horsman, G. P.; Huang, S.-X. Non-enzymatic pyridine ring formation in the biosynthesis of the rubrolone tropolone alkaloids. *Nature Communications* **2016**, *7* (1).
110. Du, L.; Robles, A. J.; King, J. B.; Powell, D. R.; Miller, A. N.; Mooberry, S. L.; Cichewicz, R. H. Crowdsourcing Natural Products Discovery to Access Uncharted Dimensions of Fungal Metabolite Diversity. *Angewandte Chemie International Edition* **2013**, *53* (3), 804-809.
111. Robles, A. J.; Du, L.; Cichewicz, R. H.; Mooberry, S. L. Maximiscin Induces DNA Damage, Activates DNA Damage Response Pathways, and Has Selective Cytotoxic Activity against a Subtype of Triple-Negative Breast Cancer. *Journal of Natural Products* **2016**, *79* (7), 1822-1827.
112. Du, L.; You, J.; Nicholas, K. M.; Cichewicz, R. H. Chemoreactive Natural Products that Afford Resistance Against Disparate Antibiotics and Toxins. *Angewandte Chemie International Edition* **2016**, *55* (13), 4220-4225.
113. Mevers, E.; Saurí, J.; Helfrich, E. J. N.; Henke, M.; Barns, K. J.; Bugni, T. S.; Andes, D.; Currie, C. R.; Clardy, J. Pyonitrins A–D: Chimeric Natural Products Produced by *Pseudomonas protegens*. *Journal of the American Chemical Society* **2019**, *141* (43), 17098-17101.
114. Shingare, R. D.; Aniebok, V.; Lee, H.-W.; MacMillan, J. B. Synthesis and Investigation of the Abiotic Formation of Pyonitrins A–D. *Organic Letters* **2020**, *22* (4), 1516-1519.

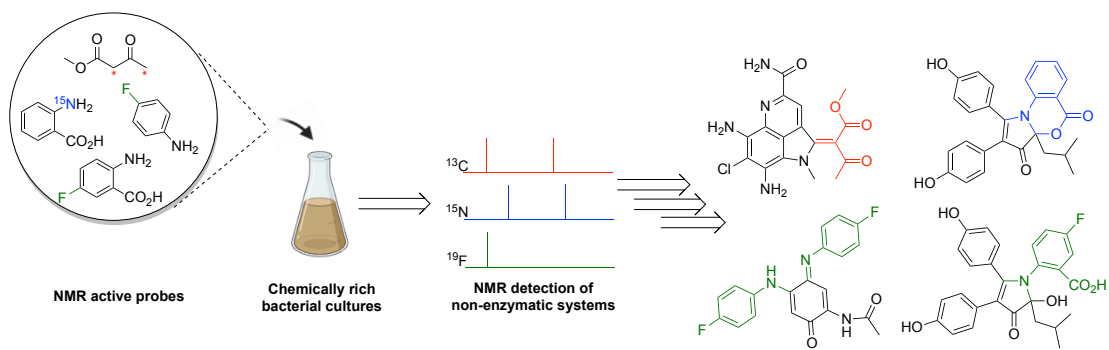
115. Oh, J.; Kim, N. Y.; Chen, H.; Palm, N. W.; Crawford, J. M. An Ugi-like Biosynthetic Pathway Encodes Bombesin Receptor Subtype-3 Agonists. *Journal of the American Chemical Society* **2019**, *141* (41), 16271-16278.
116. Massarotti, A.; Brunelli, F.; Aprile, S.; Giustiniano, M.; Tron, G. C. Medicinal Chemistry of Isocyanides. *Chemical Reviews* **2021**, *121* (17), 10742-10788.
117. Scheuer, P. J. Isocyanides and cyanides as natural products. *Accounts of Chemical Research* **1992**, *25* (10), 433-439.
118. Hughes, C. C.; MacMillan, J. B.; Gaudêncio, S. P.; Fenical, W.; La Clair, J. J. Ammosamides A and B Target Myosin. *Angewandte Chemie International Edition* **2009**, *48* (4), 728-732.
119. Pan, E.; Oswald, N. W.; Legako, A. G.; Life, J. M.; Posner, B. A.; MacMillan, J. B. Precursor-directed generation of amidine containing ammosamide analogs: ammosamides E–P. *Chem. Sci.* **2013**, *4* (1), 482-488.
120. Hughes, C. C.; MacMillan, J. B.; Gaudêncio, S. P.; Jensen, P. R.; Fenical, W. The Ammosamides: Structures of Cell Cycle Modulators from a Marine-Derived Streptomyces Species. *Angewandte Chemie International Edition* **2009**, *48* (4), 725-727.
121. Mevers, E.; Su, L.; Pishchany, G.; Baruch, M.; Cornejo, J.; Hobert, E.; Dimise, E.; Ajo-Franklin, C. M.; Clardy, J. An elusive electron shuttle from a facultative anaerobe. *eLife* **2019**, *8*, e48054.
122. Newman, D. K.; Kolter, R. A role for excreted quinones in extracellular electron transfer. *Nature* **2000**, *405* (6782), 94-97.
123. Zhang, J.; Qian, Z.; Wu, X.; Ding, Y.; Li, J.; Lu, C.; Shen, Y. Juanlimycins A and B, Ansamycin Macrodilactams from Streptomyces sp. *Organic Letters* **2014**, *16* (10), 2752-2755.
124. Wang, P.; Hong, G. J.; Wilson, M. R.; Balskus, E. P. Production of Stealthin C Involves an S–N-Type Smiles Rearrangement. *Journal of the American Chemical Society* **2017**, *139* (8), 2864-2867.
125. Gould, S. J.; Hong, S.-T.; Carney, J. R. Cloning and Heterologous Expression of Genes from the Kinamycin Biosynthetic Pathway of Streptomyces murayamaensis. *The Journal of Antibiotics* **1998**, *51* (1), 50-57.
126. Gould, S. J.; Chen, J.; Cone, M. C.; Gore, M. P.; Melville, C. R.; Tamayo, N. Identification of Prekinamycin in Extracts of Streptomyces murayamaensis. *The Journal of Organic Chemistry* **1996**, *61* (17), 5720-5721.
127. Gould, S. J.; Melville, C. R.; Cone, M. C.; Chen, J.; Carney, J. R. Kinamycin Biosynthesis. Synthesis, Isolation, and Incorporation of Stealthin C, an Aminobenzo[b]fluorene. *The Journal of Organic Chemistry* **1997**, *62* (2), 320-324.
128. Bach, T.; Hehn, J. P. Photochemical Reactions as Key Steps in Natural Product Synthesis. *Angewandte Chemie International Edition* **2011**, *50* (5), 1000-1045.

129. Ireland, C.; Sheuer, P. J. Photosynthetic marine mollusks: in vivo ^{14}C incorporation into metabolites of the sacoglossan *Placobranchius ocellatus*. *Science* **1979**, 205 (4409), 922-923.
130. Ireland, C.; Faulkner, D. J.; Solheim, B. A.; Clardy, J. Tridachione, a propionate-derived metabolite of the opisthobranch mollusc *Tridachia diomedea*. *Journal of the American Chemical Society* **1978**, 100 (3), 1002-1003.
131. Eade, S. J.; Walter, M. W.; Byrne, C.; Odell, B.; Rodriguez, R.; Baldwin, J. E.; Adlington, R. M.; Moses, J. E. Biomimetic Synthesis of Pyrone-Derived Natural Products: Exploring Chemical Pathways from a Unique Polyketide Precursor. *The Journal of Organic Chemistry* **2008**, 73 (13), 4830-4839.
132. Sharma, P.; Lygo, B.; Lewis, W.; Moses, J. E. Biomimetic Synthesis and Structural Reassignment of the Tridachiahydropyrones. *Journal of the American Chemical Society* **2009**, 131 (16), 5966-5972.
133. Bandurraga, M. M.; Fenical, W.; Donovan, S. F.; Clardy, J. Pseudopterolide, an irregular diterpenoid with unusual cytotoxic properties from the Caribbean sea whip *Pseudopterogorgia acerosa* (Pallas) (Gorgonacea). *Journal of the American Chemical Society* **1982**, 104 (23), 6463-6465.
134. Rodríguez, A. D.; Shi, J.-G. The First Cembrane–Pseudopterane Cycloisomerization. *The Journal of Organic Chemistry* **1998**, 63 (3), 420-421.
135. Tang, B.; Bray, C. D.; Pattenden, G. A biomimetic total synthesis of (+)-intricarene. *Tetrahedron Letters* **2006**, 47 (36), 6401-6404.
136. Roethle, P. A.; Hernandez, P. T.; Trauner, D. Exploring Biosynthetic Relationships among Furanocembranoids: Synthesis of (–)-Bipinnatin J, (+)-Intricarene, (+)-Rubifolide, and (+)-Isoepilophodione B. *Organic Letters* **2006**, 8 (25), 5901-5904.
137. Stichnoth, D.; Kölle, P.; Kimbrough, T. J.; Riedle, E.; de Vivie-Riedle, R.; Trauner, D. Photochemical formation of intricarene. *Nature Communications* **2014**, 5 (1), 5597.
138. Fedorov, S. N.; Radchenko, O. S.; Shubina, L. K.; Kalinovsky, A. I.; Gerasimenko, A. V.; Popov, D. Y.; Stonik, V. A. Aplydactone, a New Sesquiterpenoid with an Unprecedented Carbon Skeleton from the Sea Hare *Aplysia dactylomela*, and Its Cargill-Like Rearrangement. *Journal of the American Chemical Society* **2000**, 123 (3), 504-505.
139. Fedorov, S. N.; Shubina, L. K.; Kalinovsky, A. I.; Lyakhova, E. G.; Stonik, V. A. Structure and absolute configuration of a new rearranged chamigrane-type sesquiterpenoid from the sea hare *Aplysia* sp. *Tetrahedron Letters* **2000**, 41 (12), 1979-1982.
140. Burckle, A. J.; Vasilev, V. H.; Burns, N. Z. A Unified Approach for the Enantioselective Synthesis of the Brominated Chamigrene Sesquiterpenes. *Angewandte Chemie International Edition* **2016**, 55 (38), 11476-11479.
141. Matsuura, B. S.; Kölle, P.; Trauner, D.; de Vivie-Riedle, R.; Meier, R. Unravelling Photochemical Relationships Among Natural Products from *Aplysia dactylomela*. *ACS Central Science* **2017**, 3 (1), 39-46.

142. Baez, S.; Segura-Aguilar, J.; Widersten, M.; Johansson, A.-S.; Mannervik, B. Glutathione transferases catalyse the detoxication of oxidized metabolites (o-quinones) of catecholamines and may serve as an antioxidant system preventing degenerative cellular processes. *Biochemical Journal* **1997**, 324 (1), 25-28.
143. Kolm, R. H.; Danielson, U. H.; Zhang, Y.; Talalay, P.; Mannervik, B. Isothiocyanates as substrates for human glutathione transferases: structure-activity studies. *Biochemical Journal* **1995**, 311 (2), 453-459.
144. Satoh, K. The high non-enzymatic conjugation rates of some glutathione S-transferase (GST) substrates at high glutathione concentrations. *Carcinogenesis* **1995**, 16 (4), 869-874.
145. Tarhonskaya, H.; Rydzik, A. M.; Leung, I. K. H.; Loik, N. D.; Chan, M. C.; Kawamura, A.; McCullagh, J. S. O.; Claridge, T. D. W.; Flashman, E.; Schofield, C. J. Non-enzymatic chemistry enables 2-hydroxyglutarate-mediated activation of 2-oxoglutarate oxygenases. *Nature Communications* **2014**, 5 (1), 3423.
146. Kim, S. C.; Sprung, R.; Chen, Y.; Xu, Y.; Ball, H.; Pei, J.; Cheng, T.; Kho, Y.; Xiao, H.; Xiao, L.; et al. Substrate and Functional Diversity of Lysine Acetylation Revealed by a Proteomics Survey. *Molecular Cell* **2006**, 23 (4), 607-618.

Chapter 2

Discovery Platform to Harness Non-enzymatic Scaffolds in Natural Products



2.1 Introduction

There are a growing number of natural products that have been experimentally validated to involve one or more non-enzymatic steps within their biosynthetic pathway. As discussed in Chapter 1, the complexity of the non-enzymatic step(s) is highly variable, ranging from a simple modification to complex multi-component cyclization cascades. Examples such as (±)-homodimericin A (**2.1**) and oxazinin A (**2.2**) demonstrate how reactive functional groups can lead to complex molecules in the cellular environment of fungal fermentation (Figure 2.1).^{1, 2}

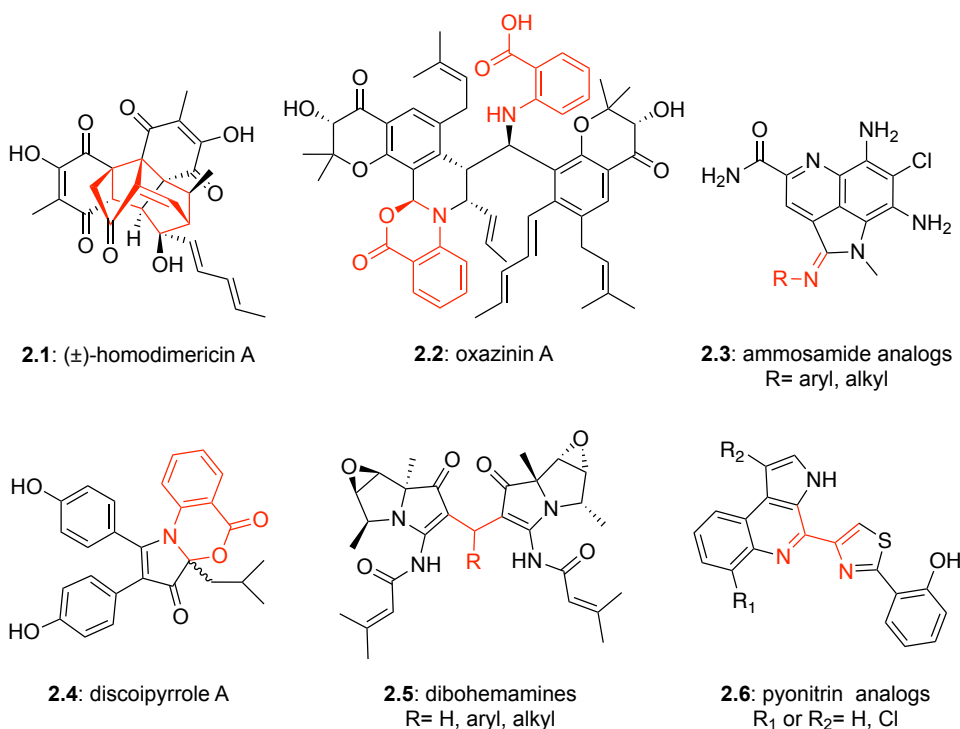


Figure 2.1. Examples of natural products that incorporate non-enzymatic transformations in their biosynthesis.^{1, 3-7} Highlighted in red are the reactive centers that participate in the non-enzymatic reaction(s).

In pursuit of discovering novel natural products from a library of marine-derived bacteria using comprehensive phenotypic screens, researchers in our lab identified three

families of natural products, the ammosamides (**2.3**), discoipyrroles (represented by discoipyrrole A (**2.4**), and dibohemamines (**2.5**) to possess potent cytotoxicity against non-small cell lung cancer (NSCLC) cell lines (Figure 2.1).^{4-6, 8} Separately, the Clardy laboratory screened a collection of insect-associated bacterial strains against fungal pathogens and through bioactivity guided isolation discovered the family of pyonitrins (**2.6**).⁷ Despite being unrelated and independent discoveries, thorough investigation into the chemistry behind these molecules revealed a commonality: all involve at least one non-enzymatic transformation in the pathway to the final product. Moreover, non-enzymatic transformations observed in these families are also unique from one another. Described in greater detail in Chapter 1, the ammosamide scaffold is subject to a non-enzymatic nucleophilic addition of a nitrogen, sulfur, or oxygen-based nucleophile to an *N*-methyl iminium ion to afford a corresponding amidine, thiolactone, or lactone (Figure 1.23).⁹ The dibohemamines are generated via a Baylis-Hilman type dimerization with formaldehyde for detoxification purposes (Figure 1.13).⁶ The discoipyrroles undergo a complex three-component cyclization cascade initiated by an imine formation between anthranilic acid and a diketone substrate (Figure 1.16).¹⁰ Lastly, biomimetic synthesis of the pyonitrins led to the revelation that these compounds undergo a non-enzymatic Pictet-Spengler condensation between two intermediates, aeruginaldehyde and an aminopyrrolnitrin, from independent biosynthetic pathways forming an imine intermediate that intramolecularly cyclizes into a 1,2-dihydroquinone that then rearomatizes to give the final product (Figure 1.21).^{7, 11}

The therapeutic properties and scaffold diversity of natural products have influenced synthetic methodology and medicinal chemistry alike. However, the same structural complexity that has inspired chemists is also responsible for some of the many challenges in the field of natural products.^{12, 13} Their structures are commonly enriched with multiple stereocenters, heteroatoms, and contain highly substituted frameworks all of which are difficult to achieve through combinatorial chemistry methods alone.^{12, 14} The total synthesis of natural products has been notoriously difficult as these synthetic routes typically require numerous steps and

often result in low yields. Consequently, the generation of natural product libraries has historically been a major challenge, especially in the context of SAR studies and probing for new biological activity.^{14, 15} However, the identification of natural product scaffolds that have the ability to undergo non-enzymatic transformations presents an opportunity to capitalize on the inherent reactivity of these molecules as a way to circumvent some of the difficulties associated with generating novel libraries. This was beautifully demonstrated by the work done with the ammosamide producing *Streptomyces variabilis* strain and truly shows the power of how identifying non-enzymatic transformations within natural product biosynthesis can lead to the rapid generation of analogs and expedite SAR studies.⁹

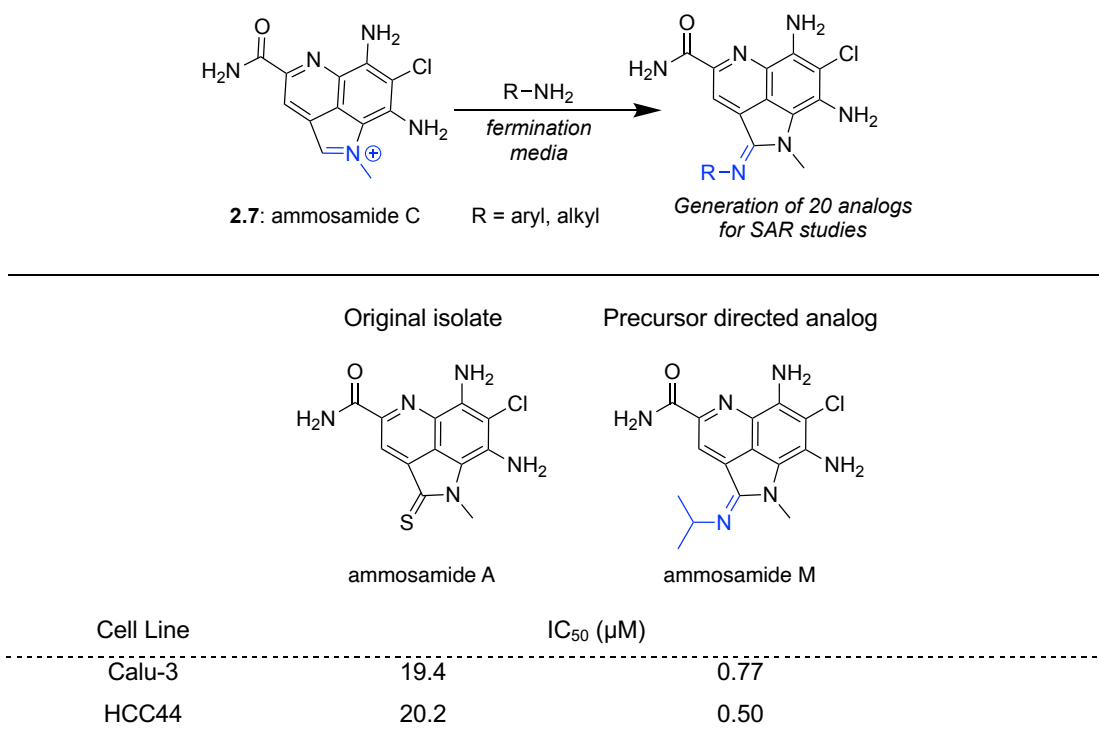


Figure 2.2. Non-enzymatic formation of amidine containing ammosamide analogs. Feeding in various aryl and alkyl amines into the fermentation media of *S. variabilis* led to the production of 20 analogs that were used for SAR analysis against NSCLC cell lines. The improved potency of the precursor directed analogs is represented by ammosamide M compared to the natural product ammosamide A.^{4, 9}

The ammosamides have a wide range of biological activity including but not limited to activity against colon carcinoma cell line HCT-116, NSCLC cell lines (HCC44, Calu-3), and have been incorporated as chemopreventive agents through inhibition of quinone reductase-2 (QR2, NQO2).^{3, 4, 16} With this expansive activity, it was of extreme interest to start building a library of analogs to increase potency as well as investigate the SAR. After synthetic strategies and isolation of the highly reactive iminium ion, ammosamide C (**2.7**), previous members in the lab were able to exploit the reactivity by feeding various aryl and alkyl amines into the fermentation media.⁹ As a result, they isolated 20 new analogs that were subject to bioactivity and found incorporation of lipophilic alkylamines significantly improved the potency against NSCLC cell lines when compared to the originally isolated ammosamides with sub μM IC_{50} values (Figure 2.2). This study highlights the significance of non-enzymatic chemistry and the desire to find more examples within natural products.

In addition to the ammosamides, precursor directed feeding studies to generate analogs were employed in the work done on the rubrolones (**1.65**, **1.66**), jadomycins (**1.45**, **1.50**), elansolids (**1.2–1.7**), and more which were discussed in length in Chapter 1.¹⁷⁻²⁰ This concept of feeding studies to generate novel analogs is not new, it is also a commonly used strategy for diversifying enzymatically derived natural products.^{13, 21} However, the breadth of analogs that can be made is limited by substrate specificity / promiscuity of the enzyme. Either a deep understanding of biosynthetic enzymes or of the native precursor building block is essential for successful production, which can be a time-consuming process. Additionally, promiscuous enzymes are generally less efficient than selective enzymes and often result in lower yields, meaning modifications to native substrates could hinder the already low production of the natural product. With that being said, the identification of non-enzymatic transformations and reactive scaffolds that can participate would help circumvent those issues in the context of novel library generation. This further substantiates the benefits of non-enzymatic chemistry and is the motivation behind the following work discussed in this chapter.

2.2 Discovery platform and selection of precursors

The discovery of natural products scaffolds that have the potential to undergo non-enzymatic reactions would provide a better understanding of how widespread these transformations are and be a powerful tool for rapidly carrying out SAR studies as mentioned above. In attempts to do this, we have taken an NMR-based approach to develop a discovery platform and general methodology for identifying natural product frameworks prone to non-enzymatic chemistry in microbial culture. For a successful and unbiased platform, the following components must be taken into consideration: (I) sensitivity of detection, (II) level of throughput, and (III) distinguishing between molecules resulting from non-enzymatic reactions versus those from endogenous production. With those features in mind, we have designed an NMR-based discovery platform that employs ^{13}C , ^{15}N , and ^{19}F containing substrates that have the added ability to be multiplexed. Due to the low abundance of ^{13}C and ^{15}N isotopes and the scarcity of fluorinated natural products, enriching the fermentation culture with these modified substrates will streamline the tracking of incorporation by NMR, facilitating the identification of reactive scaffolds, novel analogs, and potentially identify non-enzymatic transformations (Figure 2.3).

We have selected nitrogen and carbon substrates that are nucleophilic at the pH of microbial fermentation. ^{15}N -anthranilic acid (**2.8**) was selected as our nitrogen nucleophile for three reasons (Figure 2.3). The first being that it is an intermediate of tryptophan catabolism in the kynurenine pathway through tryptophan catabolism and is a building block for many important metabolites such as quinolinic acid, the precursor to nicotinamide adenine dinucleotide (NAD^+).²²⁻²⁵ As tryptophan is one of the essential amino acids, the kynurenine pathway is highly conserved across many species including microorganisms, plants, and mammals, making the selection of anthranilic acid biologically relevant.^{26, 27} Secondly, numerous examples of non-enzymatically derived natural products have incorporated anthranilic acid such as oxazinin A (**2.2**), discoipyrrole A (**2.4**) (Figure 2.1), and more that were mentioned in Chapter 1 such as rubrolone B (**1.68**) (Figure 1.17), and juanlimycin (**1.113**)

(Figure 1.25).^{1, 10, 17, 28} Lastly, ¹⁵N-anthranilic acid is commercially available and easily accessible.

For the carbon-based nucleophile, methyl 3-oxobutanoate-2,4-¹³C₂ (**2.9**) was chosen due in part to the formation of an enolate under aqueous fermentation conditions, which was confirmed by former lab members through ¹H NMR. Furthermore, **2.9** strongly resembles fundamental acyl building blocks used in the biosynthesis of various classes of natural products, (i.e., polyketides (PKs), non-ribosomal peptides (NRPs), and ribosomally synthesized and post-translationally modified peptides (RiPPs)), potentially serving as a substitute for a variety of intermediates.²⁹⁻³¹

There was particular interest in exploring the use of fluorinated probes as well given that fluorinated natural products are extremely rare, as previously mentioned, despite it being the 13th most abundant element and the most abundant halogen on earth.^{32, 33} Fluorine-containing compounds are also relevant in the context of drug development for there are more than 300 FDA approved drugs as of 2022.^{33, 34} Therefore, our rationale is that if a fluorine signal is observed in the NMR spectrum of our microbial extracts, there is a high likelihood that a new reaction occurred, which could lead to the identification of a novel compound. Furthermore, fluorine has only one stable isotope, which is NMR active with a spin of ½, which contributes to ¹⁹F NMR experiments being nearly as sensitive as ¹H NMR (0.83 relative sensitivity). Maintaining a wide spectral window (–220 to 500 ppm), the use of ¹⁹F NMR makes screening the extracts feasible in a short experiment allowing for higher throughput.³⁵ For those reasons, we have selected both 4-fluoroaniline (**2.10**) and 5-fluoroanthranilic acid (**2.11**) as our fluorinated substrates (Figure 2.3). 4-fluoroaniline (**2.10**) and other fluoroaromatics are commonly used in manufacturing and agriculture industries and are prevalent environmental toxins making this substrate environmentally and biologically relevant.^{36, 37} 5-Fluoroanthranilic acid (**2.11**) was chosen for the same reasons as **2.8** mentioned above, but the added fluorine handle combats the relatively low sensitivity of ¹⁵N NMR and allows for easily tracking incorporation and subsequent isolation.

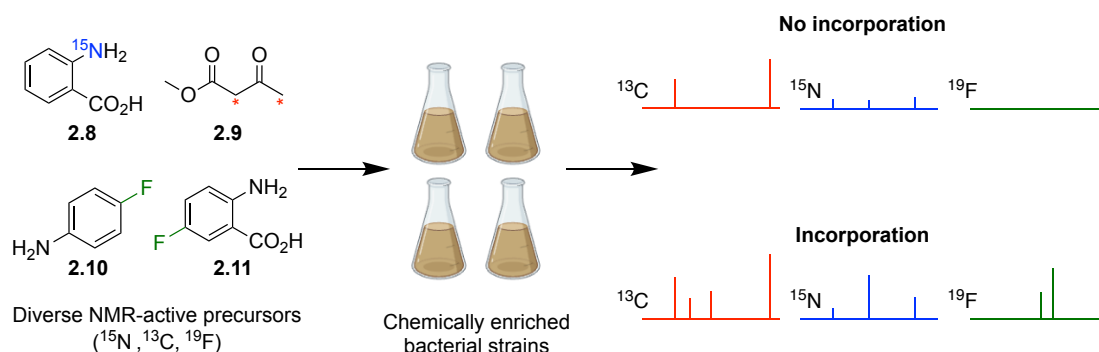


Figure 2.3. General methodology of the precursor-directed discovery platform. Isotopically labeled and fluorinated substrates are supplemented into microbial fermentation to identify scaffold subject to non-enzymatic chemistry via NMR after extraction. *Indicates ^{13}C label.

2.3 Initial evaluation of selected precursors

This project has been a joint effort between former members of the MacMillan lab, David Brumley and Dominic Colosimo, and myself over the years.³⁸ The preliminary feeding studies initially focused on bacterial strains already known for their production of non-enzymatically derived natural products, serving as a proof of principle. Subsequently, a random subset of microbial strains from our microbial library was surveyed to assess the platform's broader utility.

2.3.1 Incorporation of ^{15}N anthranilic acid in discoipyrrole A: Proof of concept

The non-enzymatic incorporation of anthranilic acid into the discoipyrrole family of natural products have been extensively studied, therefore the discoipyrrole producing strain *Bacillus hunanensis* strain SNA-048 served as an ideal model system to establish the basis for this platform. Therefore, previous lab members supplemented cultures of SNA-048 with ^{15}N -anthranilic acid (**2.8**) (0.1–1.0 g/L) after stationary phase was reached and allowed the cultures to grow for an additional 6 days before extracting with ethyl acetate.³⁸ The extracts were then analyzed using ^1H – ^{15}N HMBC experiments, as the indirect detection of nitrogen through proton-

nitrogen correlations offers greater sensitivity than traditional ^{15}N NMR.³⁹ The ^1H – ^{15}N HMBC spectrum of pure **2.8** showed correlations from the aromatic protons at δ_{H} 6.71 and 7.67 ppm to a nitrogen at δ_{N} –311.0 ppm, which served as a reference point for tracking incorporation (Figure 2.4a). The researchers observed multiple new correlations in the ^1H – ^{15}N HMBC spectrum of the ethyl acetate extract with significant shifts from δ_{N} –311.0 ppm to δ_{N} –292.0, –295.0, and –300.0 ppm (Figure 2.4b).^{10, 38} They were able to identify the presence of ^{15}N –discoipyrrole A (**2.4**) by the correlations from the signal at δ_{N} –292.0 ppm to signals at δ_{H} 1.72 and 1.56 ppm, corresponding to the methylene protons on the isobutyl side chain of **2.4**. This work demonstrates the feasibility of **2.8** and its effectiveness as a probe within this platform.

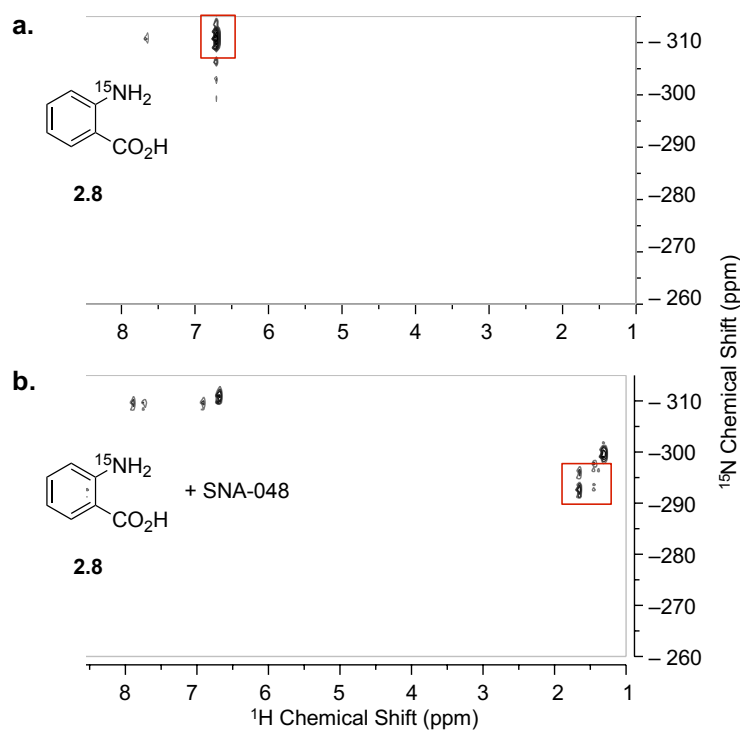


Figure 2.4. ^1H – ^{15}N HMBC spectral comparison of ^{15}N -anthranilic acid and supplemented *B. hunanensis* extract. a) Spectrum of pure ^{15}N -anthranilic acid with its chemical shift of δ_{N} –311.0 ppm highlighted in red. b) Spectrum of *B. hunanensis* (SNA-048) extract fermented with ^{15}N -anthranilic acid, with the identifiable correlations of ^{15}N -discoipyrrole A highlighted in red. Performed and analyzed by David Brumley and Dominic Colosimo.^{10, 38}

2.3.2 Incorporation of methyl 3-oxobutanoate-2,4-¹³C₂ into novel ammosamide analog

Former members of the lab explored the practicality of carbon-based nucleophiles by synthesizing a series of 1,3-diones, however only methyl 3-oxobutanoate-2,4-¹³C₂ (**2.9**) displayed incorporation.³⁸ Synthetic **2.9** was fed into the ammosamide producing *S. variabilis* strain, SNA-020, and extracted after 6 days of fermentation. Unfortunately, the synthesis of **2.9** resulted in low yields, which posed challenges for detection via NMR. However, the incorporation of **2.9** was successfully detected by LC-UV-MS, revealing a new ammosamide analog with a [M+2]⁺ ion at *m/z* 392, consistent with their synthetic prediction. This experiment was repeated using commercially available unlabeled **2.9** to circumvent the low abundance issue, resulting in the isolation and full characterization of the novel analog **2.12** (Figure 2.5).³⁸

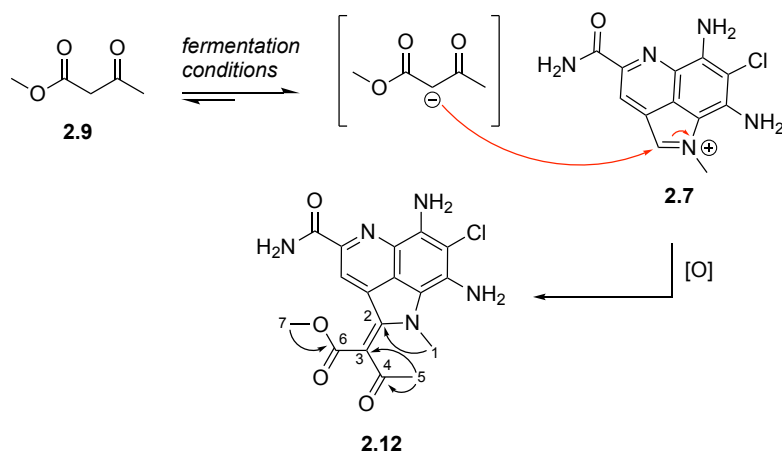


Figure 2.5. Proposed formation of novel ammosamide analog. A nucleophilic attack of the enolate form of **2.9** and subsequent oxidation to form novel **2.12**. Critical HMBC correlations for the elucidation of **2.12** indicated by black arrows. Performed and analyzed by David Brumley.³⁸

As previously noted, **2.9** behaves as an enolate under our fermentation conditions, suggesting that the generation of **2.12** occurs via nucleophilic attack of **2.9** on the iminium carbon of ammosamide C (**2.7**) followed by an oxidation event to provide the olefin bridge, a

similar mechanism proposed for the other ammosamide analogs with amine nucleophiles (Figure 2.5).⁹ Resonances at δ_{H} 8.76, 7.43, and 7.23 ppm in the ^1H NMR spectrum of **2.12** coincide with the aryl amines and amide in the ammosamide scaffold, which are conserved in other analogs, indicating that **2.9** had in fact incorporated into this motif.^{9, 38} Additional confirmation of the incorporation of **2.9** was the presence of two singlets at δ_{H} 3.48 and δ_{H} 2.34 ppm, corresponding to the methyl ester (H7) and methyl ketone protons (H5) respectively, further substantiated by the HMBC correlations to both carbonyl carbons (C4, C6) of **2.9**. Lastly, they observed HMBC correlations from the N-methyl (H1) at δ_{H} 3.73 ppm to C2 and methyl ketone protons (H5) at δ_{H} 2.34 ppm and C3 of the olefin bridge.³⁸ Despite the initial challenges with detection, this study highlights the advantage of integrating a carbon nucleophile in our platform as a way to generate new carbon-carbon bonds and further diversify scaffolds.

2.3.4 Incorporation of 4-fluoroaniline into novel iminoquinone

The next substrate that they investigated was 4-fluoroaniline (**2.10**), aimed at assessing the effectiveness of a fluorine-based probe.³⁸ They were also curious whether the incorporation of **2.10** would give rise to notable chemical shift differences in the ^{19}F NMR spectrum as was observed in the ^1H – ^{15}N HMBC spectrum in the ^{15}N -anthranilic acid study. A significant shift would indicate the presence of a new molecule, facilitating its isolation and preventing the reisolation of the substrate.

A similar approach was taken by supplementing a variety of microbial cultures with **2.10** but at a higher concentration ranging from 0.5 to 1.0 g/L. The cultures were fermented for five days, extracted with ethyl acetate, and the extracts were analyzed by ^{19}F NMR experiments. The most promising result from their screen was also from the ammosamide producing *S. variabilis* strain (SNA-020), for there was complete disappearance of the signal δ_{F} –126.9 ppm, corresponding to **2.10** (Figure 2.6a), and the emergence of two new signals at δ_{F} –116.6 and –118.3 ppm (Figure 2.6b).³⁸ Although the change in chemical shift wasn't as

dramatic as in the ^1H – ^{15}N HMBC experiments mentioned above, the observed shift and depletion of the **2.10** signal strongly suggest the presence of a new molecule. They proceeded with ^{19}F NMR guided isolation of SNA-020 in conjunction with normal and reverse phase chromatography and were able to characterize the novel iminoquinone, **2.13** (Figure 2.6c), the identity of which was uncovered as described below.³⁸

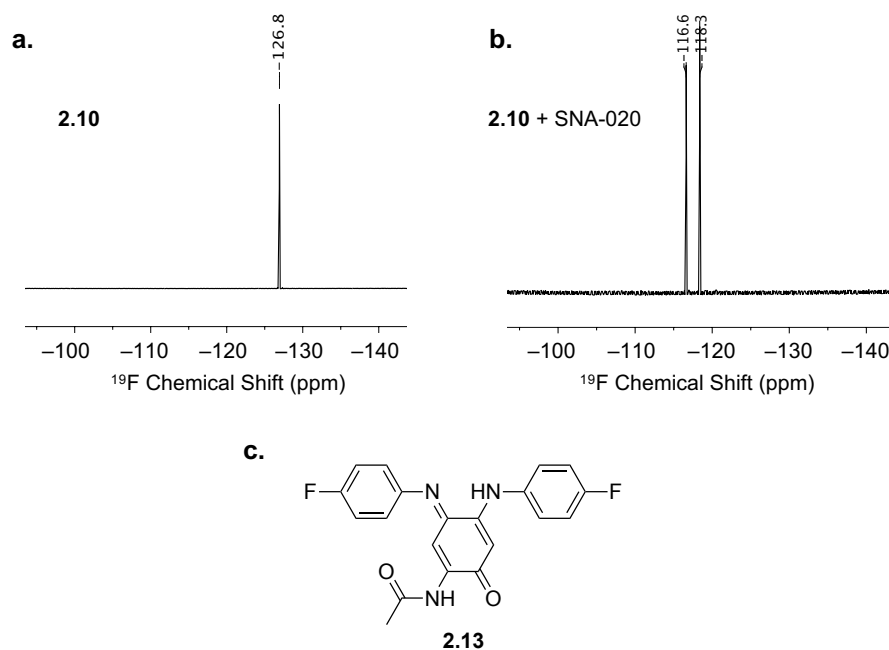


Figure 2.6. ^{19}F spectral comparison and structure of novel iminoquinone. a) ^{19}F NMR spectrum of pure 4-fluoroaniline with a chemical shift of $\delta_{\text{F}} -126.9$ ppm. b) ^{19}F NMR spectrum of *S. variabilis* (SNA-020) extract fermented with 4-fluoroaniline giving rise to two new signals at $\delta_{\text{F}} -116.9$ and $\delta_{\text{F}} -118.3$ ppm. c) Final structure of novel **2.13**. Performed and analyzed by David Brumley.³⁸ (Expanded view)

After determining the molecular formula to be $\text{C}_{20}\text{H}_{15}\text{F}_2\text{N}_3\text{O}_2$ and a m/z 368 $[\text{M}+\text{H}]^+$, the incorporation of two **2.10** moieties was confirmed by resonances at δ_{H} 7.42, 7.30, 7.28, and 7.03 ppm observed in the ^1H NMR spectrum.³⁸ Those signals are indicative of the four aromatic protons from each **2.10** subunit, which is further corroborated by the signals observed at $\delta_{\text{F}} -116.62$ and -118.38 ppm that fall within the aromatic region of ^{19}F NMR spectrum. The assignments of the aniline systems were validated primarily by ^{13}C – ^{19}F one and two bond

coupling constants and 2D correlations. The assignment of the quaternary amino carbon to be C4 was the key feature in finalizing the first aniline ring system (Figure 2.7a). This was based on the ^{13}C – ^{19}F coupling constant of 2.7 Hz, suggesting the position to be furthest from the fluorine substituent, and by observed HMBC correlations between H3/H3' and H2/H2'. Lastly, correlations in the HMBC spectrum from the amine proton at δ_{H} 9.26 ppm to both C4 and C3/C3' of the aniline ring solidified this assignment. The same methods were used to assign the second ring system as well as the iminoquinone core, all of which was substantiated by X-ray diffraction data (2.7b).³⁸ This study shows the success of how useful a fluorine handle can be in both tracking incorporation through ^{19}F NMR as well as simplifying the isolation process.

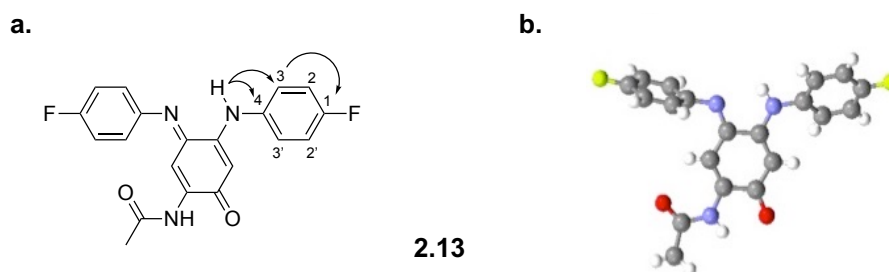


Figure 2.7. Structure determination of novel iminoquinone. a) Notable HMBC correlations for the structural assignment of **2.13** indicated by black arrows. b) Crystal structure of **2.13**. Performed and analyzed by David Brumley.³⁸

2.4 Incorporation of 5-fluoroanthranilic acid

The success of tracking ^{15}N -anthranilic acid (**2.8**) into the discoipyrroles and the use of 4-fluoroaniline (**2.10**) to reveal the power of ^{19}F NMR in identifying reactive centers inspired us to explore the utility of 5-fluoroanthranilic acid (**2.11**). This study served as an additional proof of concept using the discoipyrrole producing *B. hunanensis* strain SNA-048.

Following a modified protocol, liquid cultures of SNA-048 were supplemented with 0.5 g/L of **2.11** in DMSO after one day of fermentation and extracted with resin after an additional

five days of growth. The crude material was then partitioned with hexanes, dichloromethane (DCM), and methanol (MeOH) to provide three semi-crude fractions. To rule out the possibility of ^{19}F NMR signals resulting from **2.11** reacting with media components, **2.11** was added to the media used for microbial fermentation without bacteria and allowed to incubate for five days. Though the likelihood of SNA-048 natively producing fluorinated compounds is low, it was important to confirm that this was indeed the case. As an additional control, SNA-048 was cultured for the same amount of time without supplementing the cultures with **2.11** and extracted with resin.

The ^{19}F NMR data did not reveal any new signals and no changes in the chemical shift of pure **2.11** ($\delta_{\text{F}} -130.2$ ppm) (Figure 2.8a) to that of the extract from the media supplemented with **2.11** were observed (Figure 2.8b). Unsurprisingly, ^{19}F signals were not observed in either the extract of SNA-048 that was not supplemented with **2.11** (Figure 2.8c) nor in the hexane partition of the resin extract of SNA-048 that was supplemented with **2.11** (Figure 2.8d). However, there was a surprising number of new and unique ^{19}F NMR signals in the DCM (Figure 2.8e) and MeOH partitions of the resin extract of SNA-048 with **2.11** (Figure 2.8f). Most notably, the lack of ^{19}F signals observed in the spectrum of media supplemented with **2.11** validated that the appearance of the new signals observed in the spectrum of both the DCM and MeOH partitioned fractions were in fact a result of reactions occurring and new molecules forming during microbial fermentation and not due to reactions within the media itself.

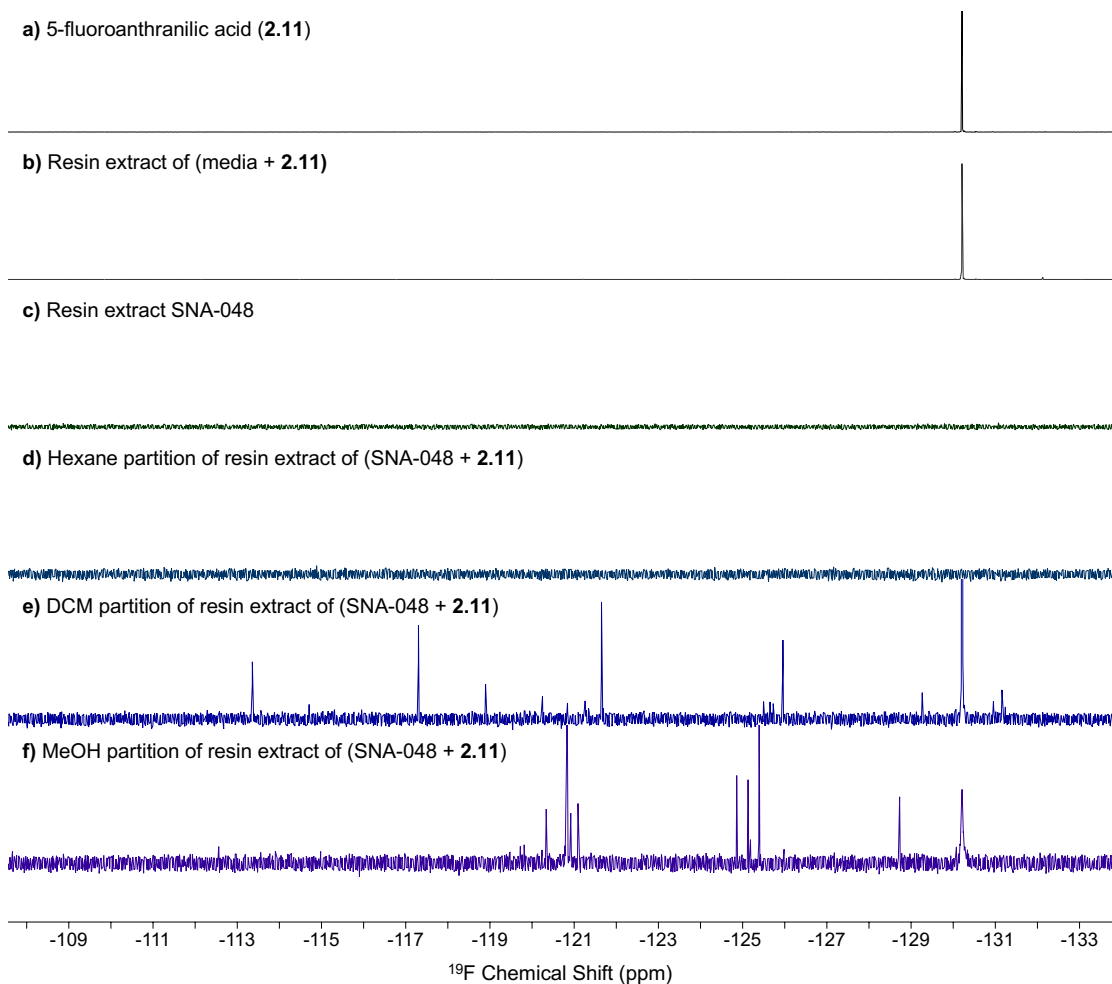
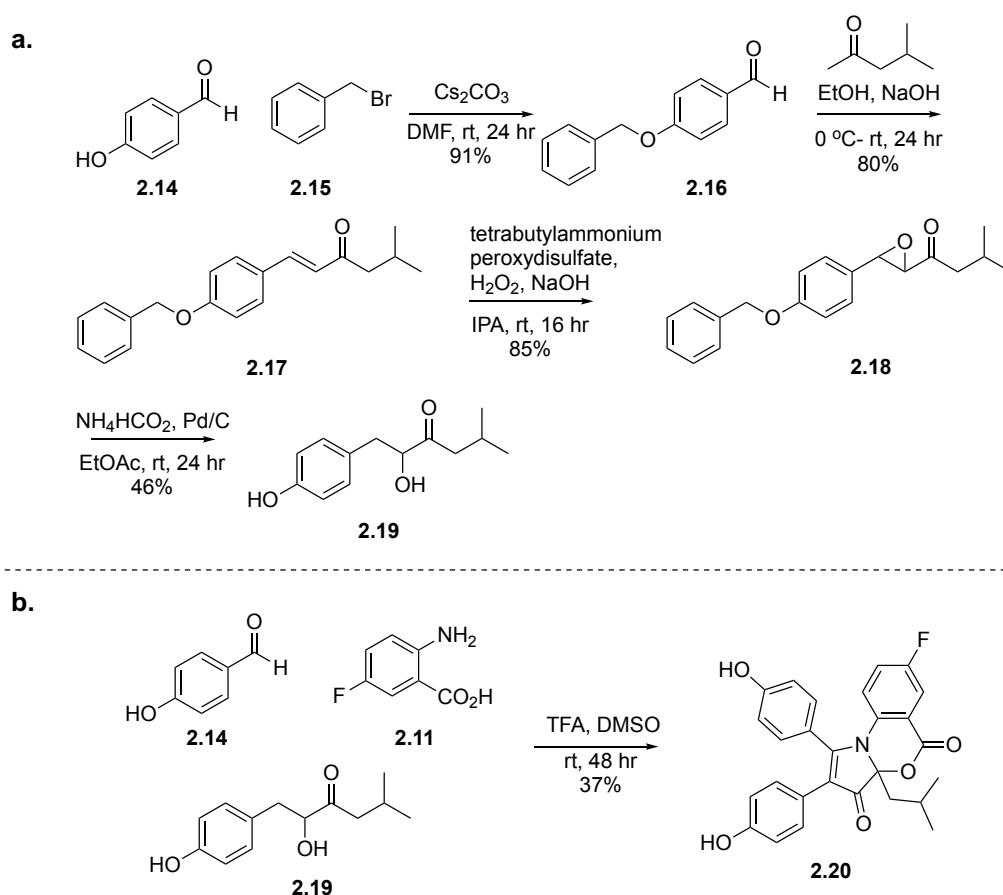


Figure 2.8. ^{19}F NMR spectral comparison of various extracts, highlighting the differences in δ_{F} . a) Pure 5-fluoroanthranilic acid (**2.11**), δ_{F} -130.2 ppm. b) Resin extract of media supplemented with **2.11** with a signal at δ_{F} -130.2 ppm. c) Resin extract of SNA-048 without **2.11**. d-e) Solvent partitions of resin extracts of SNA-048 supplemented with **2.11**. The appearance of new ^{19}F signals indicates **2.11** is successfully incorporating into scaffolds. (470 MHz, $\text{DMSO-}d_6$). (Expanded view, not to scale)

Analysis of the ^{19}F spectra of the MeOH and DCM partitioned fractions of SNA-048 revealed that the chemical shift of **2.11** (δ_{F} -130.2 ppm) remained present, unlike in the experiments with 4-fluoroaniline and SNA-020. The relatively high intensity of this signal in the NMR spectrum overshadowed the signals of novel metabolites produced in low titer. Additionally, the DCM and MeOH partitions of SNA-048 above showed a significantly greater

number of signals, resulting in more complex spectra and a more challenging isolation process compared to the 4-fluoroaniline studies. While these observations did not impede the subsequent experiments, they raised interesting questions about the methodology. To address these complexities, experiments with lower concentrations of **2.11** may reduce the intensity of its signal and the number of observed signals, thereby simplifying the ^{19}F NMR spectrum for a more streamlined isolation process. Alternatively, extending the fermentation period for SNA-048 supplemented with **2.11**, or allowing fermentation until complete consumption of **2.11**, could reveal whether the intensity of less abundant signals—likely corresponding to fluorinated metabolites produced in low titer—increases over time and allow for the observation of any changes in the number of ^{19}F signals.

Given that SNA-048 is known to produce discoipyrroles, we synthesized discoipyrrole A using **2.11** to identify the corresponding ^{19}F chemical shift and deconvolute the complex spectra of the partitioned fractions. This approach also served as a proof of concept to assess the incorporation of **2.11** under fermentation conditions. It is worth noting that the many unexplored ^{19}F signals should be pursued in future studies for the identification of novel fluorinated compounds would significantly add to this platform.



Scheme 2.1. Synthesis of fluorinated discoipyrrole A. a) Synthetic pathway of 4-hydroxysattabacin. b) Non-enzymatic cyclization with 5-fluoroanthranilic acid to form fluorinated discoipyrrole A.

Previous mechanistic studies determined that discoipyrrole A is the product of a non-enzymatic cascade between anthranilic acid, 4-hydroxybenzaldehyde (**2.14**), and 4-hydroxysattabacin (**2.19**).^{5, 10} While **2.11** and **2.14** are commercially available, **2.19** had to be synthesized. The first step was a straightforward benzyl protection of **2.14** with benzyl bromide (**2.15**) to give compound **2.16**, followed by a Claisen condensation with methyl isobutyl ketone to yield compound **2.17** (Scheme 2.1a). Epoxidation of the olefin in **2.17** provided **2.18** which was subject to a reduction with Pd/C and NH_4HCO_3 to provide the precursor 4-hydroxysattabacin (**2.19**) in moderate yield for the following non-enzymatic cyclization. All three substrates, **2.11**, **2.14**, and **2.19**, were dissolved in DMSO and TFA was added to account for

pH of the fermentation conditions (Scheme 2.1b). After 48 hours, the crude material was purified using reverse phase high performance liquid chromatography (HPLC) to provide fluorinated discoipyrrole A (**2.20**) confirmed by HRESIMS m/z 460.1563 $[M+H]^+$ (Figure 2.9b) and was fully characterized by NMR (Table A.1, Appendix A). The ^{19}F NMR spectrum of **2.20** showed a signal at δ_{F} -117.2 ppm when taken in methanol- d_4 (CD_3OD) (Figure 2.9a). The structure was validated through 1D and 2D NMR experiments and reported in Appendix A (Table A.2).

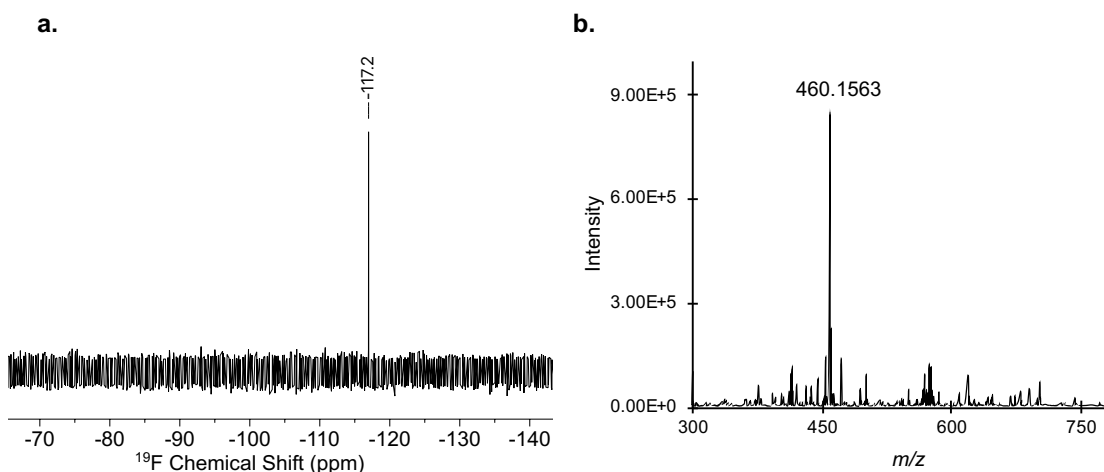


Figure 2.9. Analytical data for 2.20. a) Expanded view of ^{19}F NMR spectrum of **2.20** (470 MHz, CD_3OD). b) HRESIMS mass spectrum of **2.20**, with $[M+H]^+$ ion at m/z 460.1563 (calculated for $\text{C}_{27}\text{H}_{23}\text{FNO}_6$, 460.1516; error: +10.2 ppm).

When comparing the ^{19}F spectrum of **2.20** to both the MeOH and DCM partitioned extracts, the spectrum of the DCM fraction displayed the same resonance as synthetic **2.20** at δ_{F} -117.2 ppm, in addition to multiple others. Therefore, the DCM fraction was selected for NMR guided purification. After multiple rounds of purification, the final ^{19}F NMR spectrum revealed a minor signal at δ_{F} -117.2 ppm, consistent with **2.20**, and a major signal at δ_{F} -115.8 ppm that could not be separated (Figure 2.10a).

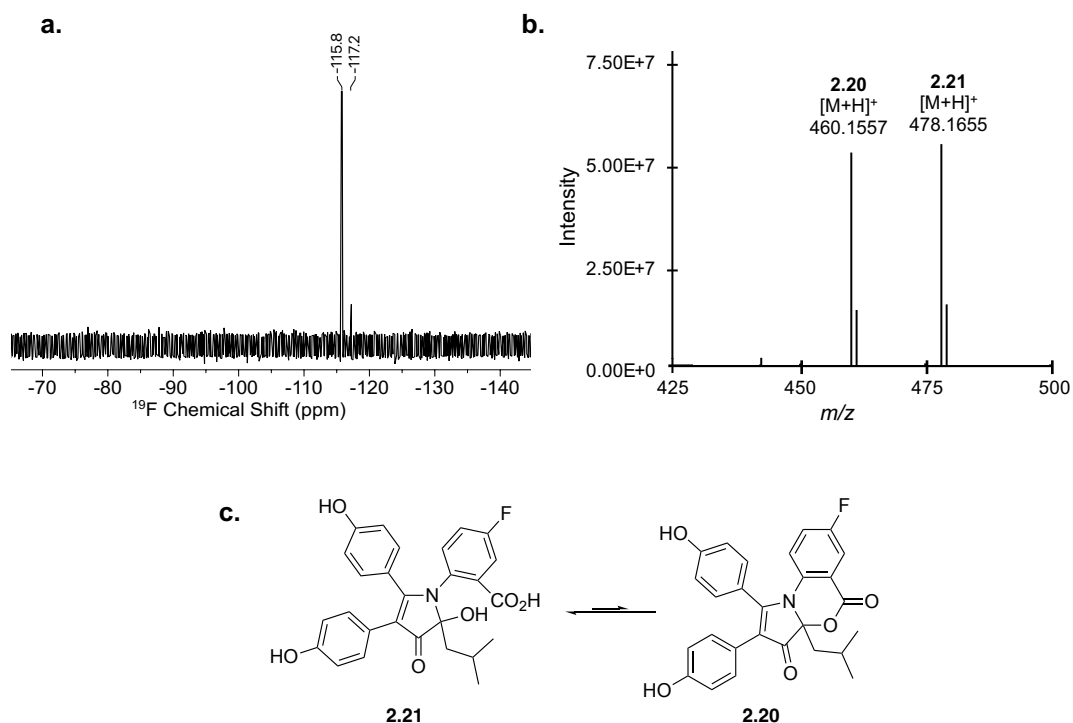


Figure 2.10. Analytical data and structure of 2.21. a) Expanded view of ^{19}F NMR spectrum of isolated **2.21** with δ_{F} -115.8 ppm and δ_{F} -117.3 corresponding to **2.20** (470 MHz, CD_3OD). b) HRESIMS of **2.21** with m/z of 478 and 460 corresponding to **2.20**. c) Structure of the open intermediate, **2.21**, in equilibrium with **2.20**.

HRESIMS analysis of the purified material provided an m/z 460.1577 $[\text{M}+\text{H}]^+$, also consistent with **2.20**, and an m/z 478.1655 $[\text{M}+\text{H}]^+$ corresponding to a free carboxylic acid, or open, intermediate, **2.21** (Figure 2.10b and 2.10c). This open intermediate, without the fluorine substituent, had been previously identified in mechanistic studies to be highly stable and in pH-dependent equilibrium with discoipyrrole A and the isolation of one over the other varied upon fermentation conditions (Figure 2.10c). In this case however, we consistently isolated **2.21** in higher abundance than **2.20** with each repeated fermentation. During the synthesis of **2.20**, the addition of TFA was intended to loosely recapitulate biological conditions but also to facilitate the dehydration-dependent lactone ring closure as previously determined. This shifted the equilibrium to favor **2.20** to where **2.21** was undetectable, suggesting that the pH of the

fermentation culture is considerably more basic at the time of extraction than the original isolation of discoipyrrole A.

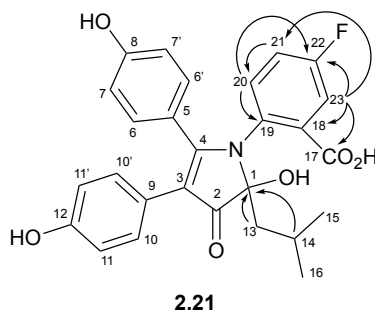


Figure 2.11. Structural assignments of 2.21. Critical HMBC correlations represented by black arrows.

Although the position where **2.11** would be incorporated can be assumed based on previous studies, it was confirmed through 1D and 2D NMR assignments. Analysis of the ^1H NMR of **2.21** revealed a carboxylic acid proton resonance at δ_{H} 8.54 ppm and aromatic proton resonances at δ_{H} 7.46, 7.21, and 7.20 ppm, which were determined to be H20, H21, and H23, respectively (Figure 2.11). H21 and H23 presented as overlapping signals in the ^1H NMR spectrum, however, HMBC correlations allowed us to differentiate the two. The HSQC spectrum and ^{13}C – ^{19}F one bond coupling constants allowed us to assign C20 (δ_{C} 134.1 ppm, J_{CF} = 8.2 Hz) to the meta position and both C21 (δ_{C} 117.3 ppm, J_{CF} = 21.5 Hz) and C23 (δ_{C} 118.2 ppm, J_{CF} = 23.2 Hz) to the ortho positions relative to the fluorine substituent (Figure 2.12a). The fluorine substituted quaternary carbon was identified to be C22 (δ_{C} 162.8 ppm) based on the J_{CF} of 243.2 Hz obtained from the HMBC spectrum (Figure 2.12b). These assignments were solidified based on key HMBC correlations observed from H23 to the carboxylic acid carbon, C17 (δ_{C} 173.0 ppm) and C18 (δ_{C} 144.1 ppm), while H21 correlated to C20. The observed correlations from H13 (δ_{H} 1.92 ppm) and H14 (δ_{H} 2.02 ppm) of the isopropyl tail to C1 (δ_{C} 94.0 ppm) were consistent with reported values of discoipyrrole A.^{5, 10} The remaining features were assigned in the same fashion (Table A.1, Appendix A).

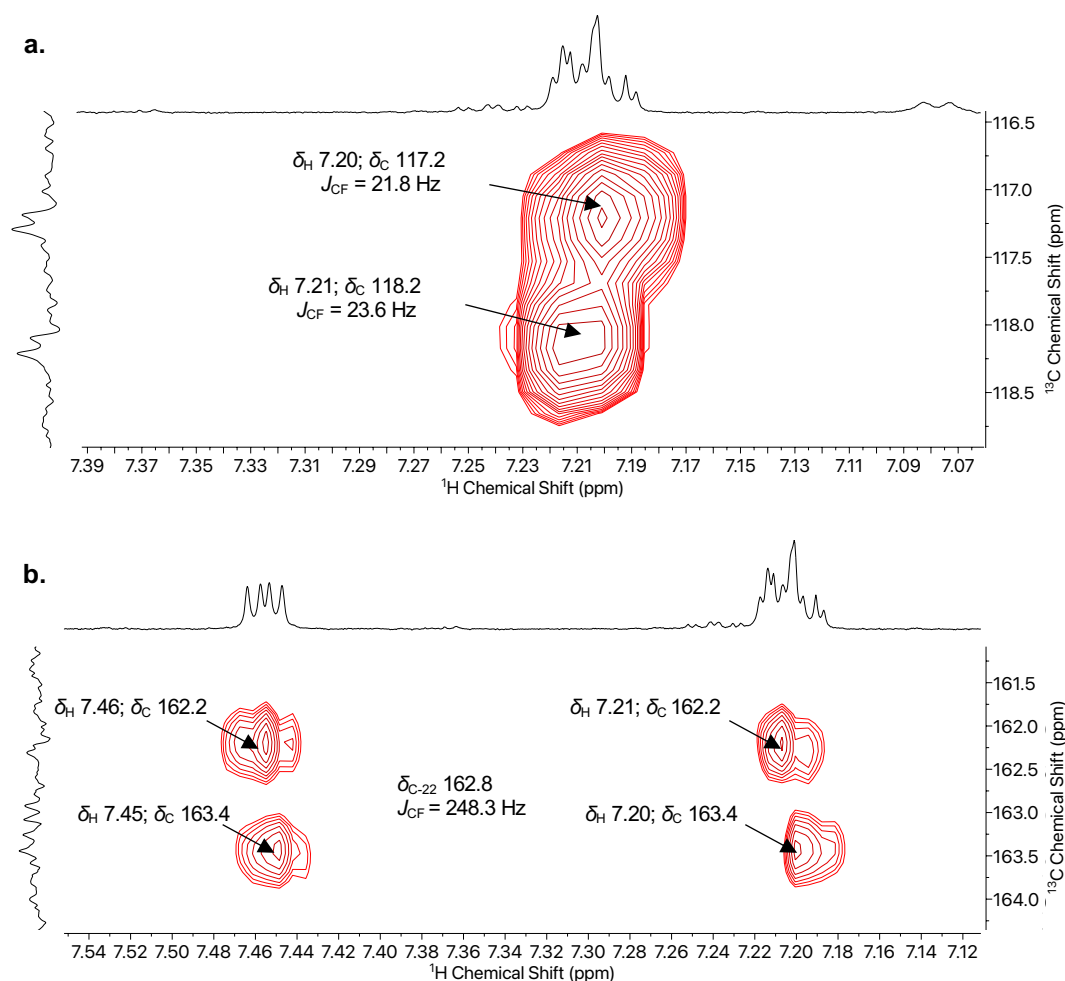


Figure 2.12. Key HSQC and ^1H – ^{13}C HMBC correlations for structural assignment of **2.21.** a) HSQC correlations of H21 to C21 and H23 to C23 along with the associated ^{13}C – ^{19}F coupling constants. b) HMBC correlations identifying the fluorine substituted carbon and corresponding ^{13}C – ^{19}F coupling constant. (800 MHz, CD_3OD). (Expanded view)

Interestingly, when comparing the ^1H NMR spectrum of **2.20** to **2.21**, there were substantial differences in the observed resonances of H20 and H23 of the fluorine substituted ring, whereas H21 maintained a similar chemical shift of $\delta_{\text{H}} 7.24 \text{ ppm}$ in **2.20** and $\delta_{\text{H}} 7.21 \text{ ppm}$ in **2.21**. H23 is notably less shielded in **2.20** with a chemical shift of $\delta_{\text{H}} 7.74 \text{ ppm}$ than in **2.21** ($\delta_{\text{H}} 7.20 \text{ ppm}$) and H20 is significantly more shielded in **2.20** with a resonance at $\delta_{\text{H}} 6.51 \text{ ppm}$ compared to **2.21** ($\delta_{\text{H}} 7.46 \text{ ppm}$) (Figure 2.13). Though the differences in chemical shifts clearly indicate that the lactone in **2.20** has a stronger delocalizing effect on the aromatic ring than the

carboxylic acid substituent in **2.21**, this observation inspired us to further investigate the structural differences between these molecules.

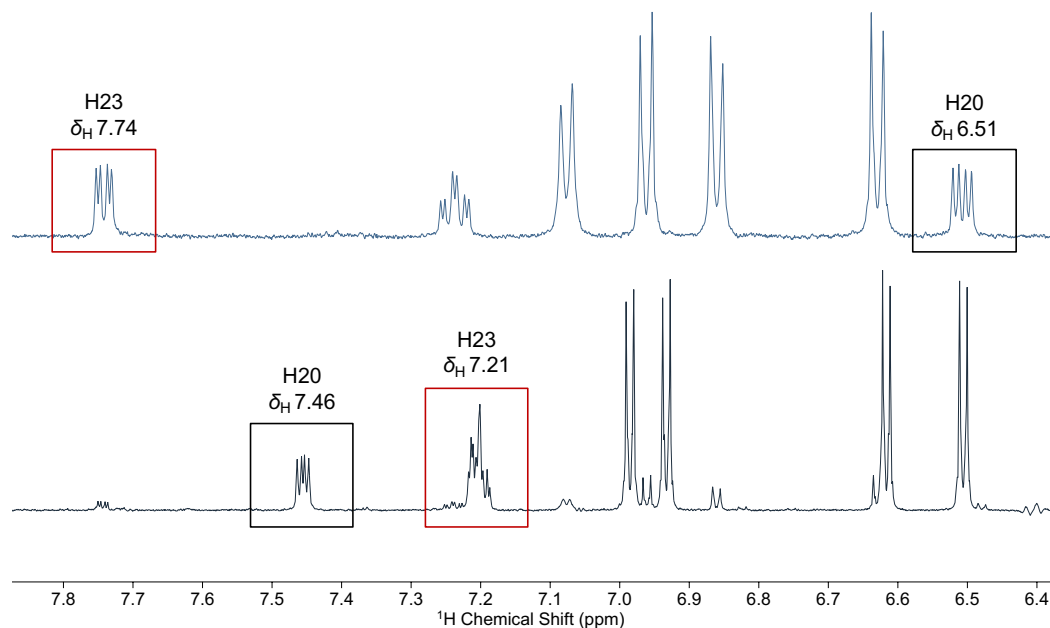


Figure 2.13. Comparison of the aromatic region in the ^1H NMR spectra of **2.20 and **2.21**.** *Top:* spectrum of **2.20**. *Bottom:* spectrum of **2.21**. H23 is noticeably more shielded in **2.21** (*top*) than **2.20** (*bottom*), highlighted by red boxes. On the contrary, H20 is significantly more shielded in **2.20** than **2.21**, indicated by black boxes. (800 MHz, CD_3OD).

With the help from the Johnstone lab at UC Santa Cruz, we were able to use DFT calculations to optimize the geometry of both **2.20** and **2.21**. The DFT results revealed that the pyrrole core remains pyramidalized in both molecules, however, there is a major difference in the planarity of the functional groups in the fluorine substituted aromatic ring. The lactone moiety in **2.20** is almost co-planar with respect to the fluorine substituted aromatic ring (Figure 2.14a), while the carboxylic acid functional group at the same position in **2.21** is rotated out of the plane of the ring (Figure 2.14b). A planar geometry allows the carbonyl unit to more effectively delocalize into the aromatic ring, and the greater planarity of **2.20** is therefore consistent with the observed shielding of its H20 and H23 as compared to those of **2.21**.

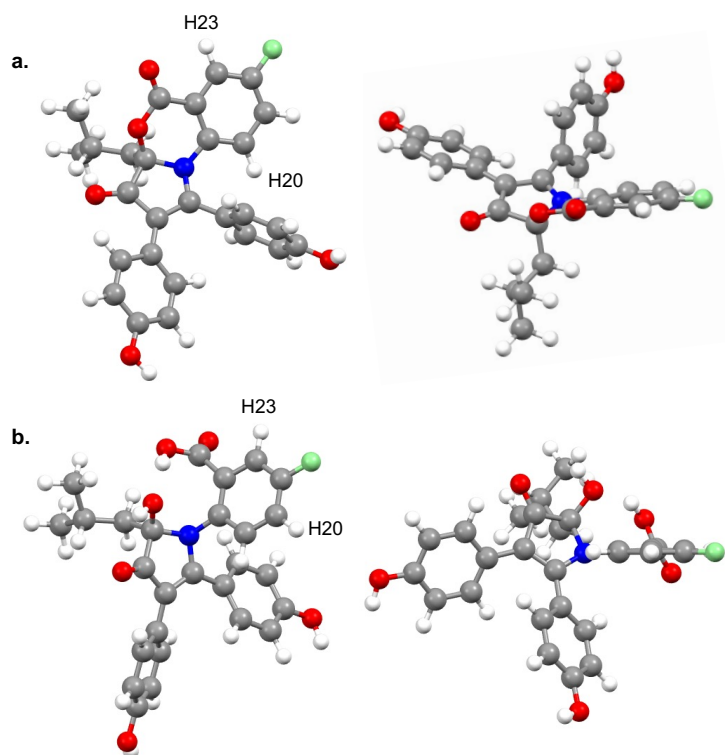


Figure 2.14. Putative structures obtained from DFT calculations. a) Structure of **2.20**. b) Structure of **2.21**. *Left*: Frontal view with pyrrole core in the center. *Right*: Side view with respect to the fluorine substituted aromatic ring. Fluorine atom represented in green, nitrogen in blue, oxygen atoms in red, carbons in gray, and hydrogens in white.

2.5 Survey of microbial library with fluorinated precursors

After successfully isolating novel compounds and demonstrating the utility of selected precursors, we established the foundation of the platform and went on to conduct a preliminary survey of various microbes from our library. We proceeded with the fluorinated probes for this screen since their incorporation was the most distinguishable by NMR compared to the others, allowing for a more streamlined and higher throughput survey. We broadened the scope of the fluorinated probes to include 4-bromo-2-fluoroaniline (**2.22**), as the bromine provides a mass spectrometry handle for identification, and 4-fluorobenzaldehyde (**2.23**), which introduces an electrophilic substrate to our platform (Figure 2.15b and c). Additionally, **2.23** is another

environmental pollutant known to be degraded by microbes to form fluorocatechols and fluoroacetates, similar to **2.10**.⁴⁰⁻⁴²

The microbial strains were arbitrarily to eliminate bias towards strains known to produce non-enzymatically derived natural products and to explore whether similar trends are observed. The panel of microbes consisted of six different strains: *Salinispora arenicola* sp. (SNA-091), a well-studied prolific producer of secondary metabolites,⁴³ *Streptomyces* sp. (RL10-292-AIS-A, abbreviated RL10), *Erythrobacter flavus* sp. (SNC-119), which are less known for producing bioactive compounds but more so for diverse sulfur containing carotenoids,⁴⁴ *Streptomyces camosus* sp. (SNB-014), *Arthrobacter aurescens* sp. (SNG-016), known for its ability to degrade environmental pollutants but not as renowned for production of diverse natural products,^{45, 46} and *Micromonospora sediminicola* sp. (SNC-018), another notable producer.⁴⁷ RL10, given to us by the Linington Lab at Simon Fraser University, was of particular interest as it is known to produce the abyssomicin class of molecules, which contain potential Michael acceptors for non-enzymatic addition to occur.^{48, 49} This non-enzymatic potential has yet to be fully explored, making RL10 a great model for this platform.

The precursors were arbitrarily assigned to the strains: SNA-091 and RL10 were treated with 4-fluoroaniline (**2.10**), while and SNC-119 and RL10 were treated with 5-fluoroanthranilic acid (**2.11**). For the exploratory probes, SNB-014 and RL10 were supplemented with **2.22** and SNC-018 and SNG-016 were supplemented with **2.23**. On a 5 L scale, all cultures were supplemented with 0.5 g/L of their respective probe after one day of growth and allowed to ferment for six additional days. The cultures were then extracted with EtOAc, and the crude material was analyzed by ¹⁹F NMR. To our excitement, we observed the same trend as before where the supplementation gave rise to an abundance of new and unique signals in all cases (Figure 2.15).

Analysis of the ¹⁹F NMR spectrum for both SNA-091 and RL10 supplemented with **2.10** showed the complete disappearance of **2.10** at δ_F –126.8 ppm (Figure 2.15a), as was observed in the isolation of the novel iminioquinone (**2.13**) in section 2.3.4. In the crude extract of SNA-

091, four prominent signals observed at δ_F -117.4, δ_F -118.2, δ_F -119.8, and δ_F -131.2 ppm (Figure 2.15a, middle). However, these signals were not abundant, likely due to the small scale of this survey. It is recommended that this study be repeated on a larger scale for isolation in future work. In the ^{19}F NMR spectrum of RL10 crude extract, there were two very pronounced signals were observed at δ_F -120.7 and δ_F -129.3 ppm (Figure 2.15a, bottom). The distinct signals observed from the relatively small fermentation scale of RL10 supplemented with **2.10** suggest great potential for the efficient isolation of a fluorinated compound.

The opposite was observed in the ^{19}F spectrum of the **2.11** study, where **2.11** at δ_F -131.4 was still present (Figure 2.15b). The ^{19}F NMR spectrum of SNC-119 revealed a plethora of robust signals, spanning a wide range in the spectral window about δ_F -104 to δ_F -140 ppm (Figure 2.15b, middle). The significant number of observed resonances has the potential to complicate the isolation process but would still be extremely interesting for future students to follow up on, especially given the unique carotenoids produced by *E. flavus*.⁴⁴ Additionally, RL10 supplemented with **2.11** resulted in another set of notable and relatively intense signals at δ_F -121.0, -126.2, and -129.7 ppm, again allowing for a more streamlined NMR guided isolation process (Figure 2.15b, bottom).

Similarly, **2.22**, with a chemical shift of δ_F -134.3 ppm, remained present in both crude extracts of SNB-014 and RL10 (Figure 2.15c). There were significantly fewer signals in the ^{19}F NMR spectrum of SNB-014 than observed in the other strains, with only two distinct peaks at δ_F -124.8 and δ_F -126.8 ppm, albeit with high intensity (Figure 2.15c, middle). The spectrum of RL10 revealed a fair number of new signals, however, there wasn't substantial change in chemical shift from that of **2.22** (Figure 2.15c, bottom). While it is unclear if a significant change in chemical shift is needed to isolate a unique compound, this would be ranked as a lower priority compared to the strains from the other feeding studies. It can be reasoned that the presence of two strong electron-withdrawing groups, bromine and fluorine, and their inductive effects on the aromatic ring system decreases the nucleophilicity and reactivity of the amine

substituent on **2.22**, which could explain the fewer and less diverse signals detected in this feeding study.⁵⁰ Therefore, it is advised to use the other fluorinated substrates in future studies.

Lastly, the ¹⁹F NMR spectrum of **2.23** revealed a major signal at δ_F –104.9 ppm and a minor one at δ_F –116.3 ppm (Figure 2.15d, top). Based on a prior report, the minor signal observed at δ_F –116.3 ppm is consistent with 4-fluorobenzoate (δ_F –116.4 ppm), a common oxidized degradation product of **2.23**.⁴⁰ This is not surprising for **2.23**, like most aldehydes, is reactive and sensitive to external conditions such as air, water, and light. The ¹⁹F NMR spectrum of both SNC-018 and SNG-016 supplemented with **2.23** showed numerous signals in the range of δ_F –114 to δ_F –119 ppm (Figure 2.15d middle and bottom). These chemical shifts fall within the window of 4-fluorobenzoate and the other molecules isolated in this chapter (δ_F –116.6 and –119 ppm (**2.13**); δ_F –117.2 ppm (**2.20**); δ_F –115.8 ppm (**2.23**)), making it potentially challenging to discern between a natural product and an inherent degradation product. This should be taken into consideration especially in the case of *Arthrobacter* sp. (SNG-016) for its ability to metabolize and breakdown environmental pollutants such as **2.23**, as previously mentioned.^{45, 46} Although this claim cannot be made definitively, it might be of more interest to pursue the signals up or downfield of that range or use an alternative electrophile for this platform to mitigate the signal overlap observed in this experiment.

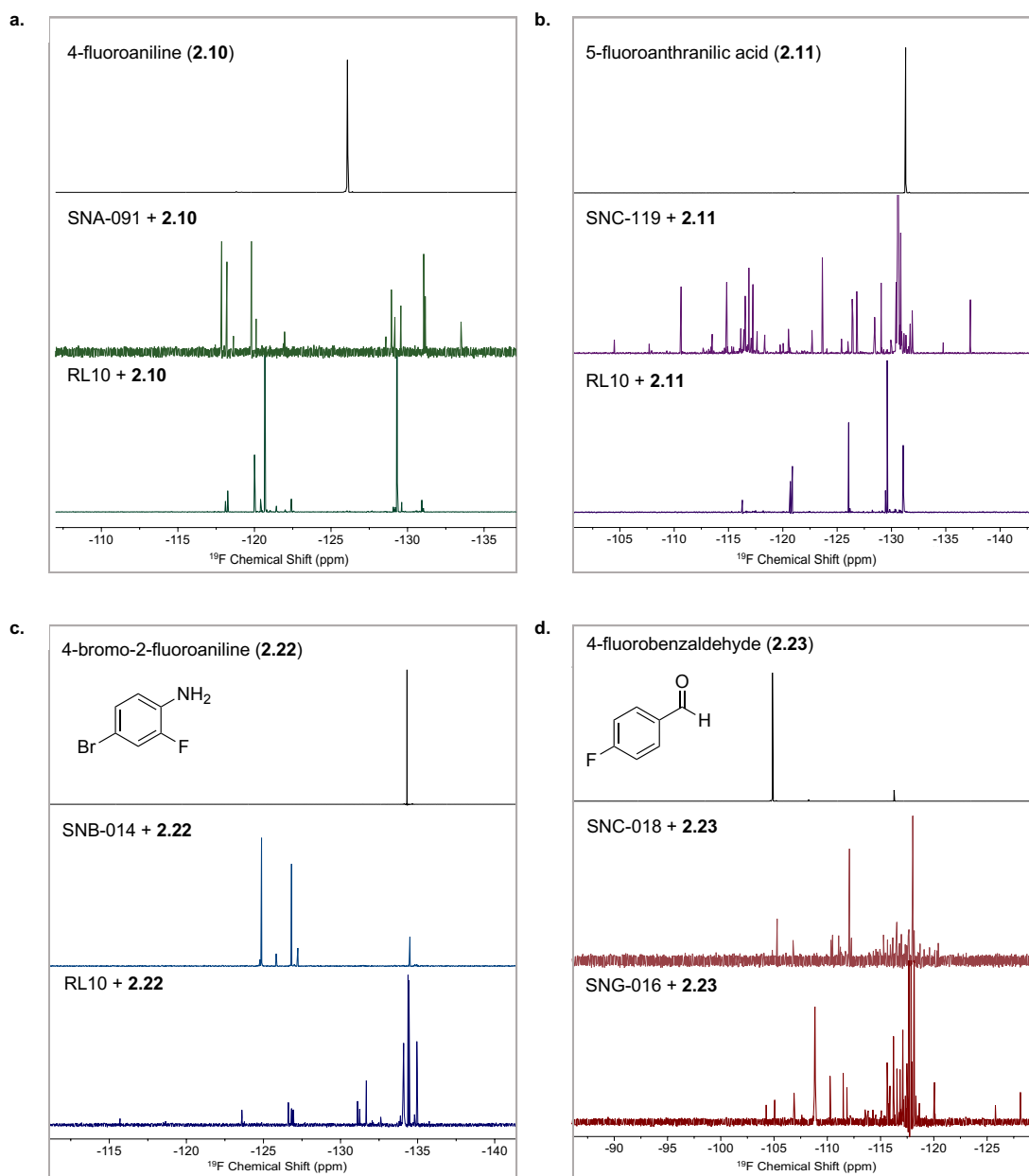


Figure 2.15. ^{19}F NMR spectral data for the microbial survey. (Expanded view, not to scale) a) Top: 4-fluoroaniline (**2.10**) and crude extracts of *Salimispora* sp. SNA-091 (middle) and *Streptomyces* sp. RL10 (bottom), both supplemented with **2.10**. b) Top: 5-fluoroanthranilic acid (**2.11**) and crude extracts of *Erythrobacter* sp. SNC-119 (middle) and RL10 (bottom), both supplemented with **2.11**. c) Top: 4-bromo-2-fluoroaniline (**2.22**) and crude extracts of *Streptomyces* sp. SNB-014 (middle) and RL10 (bottom), both supplemented with **2.22** d) Top: 4-fluorobenzaldehyde (**2.23**) and crude extracts of *Micromonospora* sp. SNC-018 (middle) and *Arthrobacter* sp. SNG-016 (bottom), both supplemented with **2.23**. All strains show unique ^{19}F chemical shifts than that of the respective probe. (470 MHz, CD_3OD).

While this survey involved a small sample size, it suggests that these fluorine feeding studies can be employed across a variety of microbial strains, resulting in promising leads. Despite minor complications, such as convoluted NMR spectra or insignificant changes in chemical shifts, the prominent signals observed in each strain present a valuable opportunity for future work to be done on this platform.

2.6 Conclusion

To summarize, we were able to build the foundation of an NMR-based discovery platform to identify natural products with reactive scaffolds that can be harnessed for structure diversification and, if applicable, SAR studies. Using isotopically labeled ^{15}N and ^{13}C and fluorinated substrates led us to a successful proof of concept in the re-isolation of discoipyrrole A (**2.4**) and the isolation of a novel ammosamide analog (**2.12**), novel iminoquinone (**2.13**) and a fluorinated analog of discoipyrrole A (**2.21**). Additionally, we conducted a basic survey of several microbes supplemented with different fluorinated probes to explore the broader application of both this platform and precursor variability which resulted in ^{19}F NMR signals observed in all strains with all probes, albeit some more distinguishable than others, further supporting the validity of this platform.

Several factors still need to be considered to increase the scalability of this discovery platform including the cost of probes and NMR detection limits for the probe of choice. For instance, ^{15}N -anthranilic acid (**2.8**) involved costly experiments (\$935 per gram), however, sensitivity was improved by employing 2D ^1H - ^{15}N HMBC experiments rather than 1D ^{15}N NMR experiments. Similarly, the synthesis of the carbon nucleophile (**2.9**) resulted in a 37% yield with the starting material costing \$147 per gram. Additionally, a considerable amount of **2.9** is needed as well as the reactive intermediate would need to be produced in relatively high titer in order to accurately detect incorporation due to the inherently low sensitivity of ^{13}C NMR. Despite those challenges, there are still benefits to including carbon nucleophiles for the

functionalization of natural products and it would be worthwhile to explore alternatives that are more easily scalable and result in higher yield. Evaluation of the fluorinated probes, 4-fluoroaniline (**2.10**) and 5-fluoroanthranilic acid (**2.11**), were the least expensive of the probes (\$0.66 and \$15 per gram, respectively) and ^{19}F NMR provided greater sensitivity. The sensitivity combined with the scarcity of naturally occurring fluorinated natural products made the incorporation of these probes easily distinguishable and significantly aided in the NMR guided isolation process.

Overall, the work presented in this chapter strongly supports the potential of this NMR based discovery platform to have substantial promise for future natural product discovery. Identifying reactive scaffolds and/or scaffolds that can participate in non-enzymatic chemistry offers an opportunity for structural diversification and the generation of natural product libraries. In addition, this platform allows for rapid and efficient SAR studies for biologically active molecules. We believe this is a highly promising approach and could be seamlessly integrated with high-throughput biological screening platforms.

2.7 Experimental procedures

General methods

All solvents and reagents were commercially available unless otherwise noted. Chemicals were purchased from Aldrich, Fisher, Alfa Aesar, TCI, or Oakwood chemicals and used without purification. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance III HD 800 MHz and Bruker Avance III HD 500 MHz spectrometers and deuterated solvents were purchased from Cambridge Isotope Laboratories. Chemical shifts (δ) are reported in ppm relative to residual solvent chloroform-*d* (CDCl_3 : δ_{H} = 7.26 ppm, δ_{C} = 77.16 ppm), dimethyl sulfoxide-*d*₆ ($(\text{CD}_3)_2\text{SO}$: δ_{H} = 2.50 ppm, δ_{C} = 39.52 ppm), and methanol-*d*₄ (CD_3OD : δ_{H} = 3.31 ppm, δ_{C} = 49.0 ppm) and coupling constants (*J*) in Hz. Multiplicity is denoted as s for singlet, d for doublet, t for triplet, q for quartet, m for multiplet, and br. s for when the signal in question is

broadened. For ^{19}F NMR experiments, 64 – 128 scans were collected depending on the amount of sample over a spectral width of 0 – 200 ppm. Flash chromatography was performed using E. Merck silica gel 60 (240–400 mesh). High resolution electrospray ionization mass spectrometry (HRESIMS) was conducted on LTQ-Orbitrap Velos Pro MS. Low resolution mass spectrometry was recorded on Agilent Technologies 6130 series Quadrupole LC-MS with electrospray ionization (ESI). HPLC purification was performed on Agilent Technologies 1200 series consisting of a UV–vis detector (G1315D), dual pumps (G1312A), and a dynamic mixer (G1312A). Reactions were performed under an atmosphere of nitrogen with magnetic stirring unless otherwise noted. Thin layer chromatography (TLC) was performed using precoated plates purchased from E. Merk (Silica gel 60 PF₂₅₄, 0.25 mm).

General mass spectrometry analysis

Methods for both HRESIMS and LC-UV-MS were identical. Samples were diluted with MeOH to a concentration of 1 mg/mL with an injection volume of 5–10 μL on a Phenomenex Luna Omega C18 (100 $\times$ 2.1 mm, 1.6 μm) LC column with a flow rate of 0.3 mL/min and a gradient solvent system of 10–100% ACN in H₂O (0.1% formic acid) over 12 minutes. MS scans were operated from either m/z 150–1500 or m/z 300. LC-UV-MS was monitored at UV = 254 nm.

Cultivation and extraction of *Bacillus hunanensis* SNA-048 cultures

From a glycerol stock, SNA-048 was streaked onto agar plates (15 g agar, 10 g starch, 4 g yeast extract, 2 g peptone, 1 g CaCO₃, 40 mg Fe₂(SO₄)₃ 4H₂O, 100 mg KBr, 1 L seawater) and incubated for 6 days at 30 °C. One colony of SNA-048 was inoculated in a seawater-based media (10 g starch, 4 g yeast extract, 2 g peptone, 1 g CaCO₃, 40 mg Fe₂(SO₄)₃ 4H₂O, 100 mg KBr, 1 L seawater) and shaken in a rotary shaker at 200 rpm at 30 °C for 24 hrs. To do a large-scale fermentation, 10 mL/L of the overnight culture was sterilely transferred into 21 x 2.8 L Fernbach flasks and shaken under the same conditions as above. After 1 day of the large-scale fermentation, 0.5 g in DMSO of 5-anthranilic acid was sterilely filtered into 20 of the 1 L

cultures and allowed to shake for an additional 5 days. Activated Amberlite XAD-7-HP resin (20 g/L) was added to the cultures and shaken at 200 rpm for 2 hours. The resin was then filtered through cheesecloth, washed with deionized water, and eluted with acetone. The acetone soluble material was concentrated under reduced pressure to yield 3.0 g crude material.

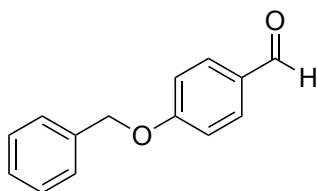
Isolation and purification of **2.21**

The crude extract (3.0 g) was resuspended in MeOH:H₂O and partitioned with hexanes and DCM (4 X 200 mL). The DCM soluble layer (0.22 g) was fractionated via flash chromatography (Silica, stepwise to 70% EtOAc in hexanes) with a final wash of 10% MeOH in EtOAc. Every 10 fractions were pooled together to provide Fr.1–Fr.8 and the EtOAc:MeOH flush (Fr.9). Fr.9 was further fractionated by flash chromatography on ODS (C18 gold, 30 g, UV = 254 nm) with a stepwise gradient of 10–90% MeOH in H₂O over 30 minutes to afford 25 subfractions that were combined based on UV activity to provide a total of 13 subfractions (Fr.S1–Fr.S13). Subfraction Fr.S8 (0.012 mg, RT = 13.5 min) was purified by semipreparative reverse phase HPLC (Phenomenex Synergi Hydro-RP, 250 × 10 mm, 4 μm, 1.8 mL/min, UV = 254) with a gradient solvent system of 35–70% ACN in H₂O (0.1% formic acid) over 30 minutes to afford **2.21** (0.0012 g, RT = 20.2 min). HRESIMS [M+H]⁺ *m/z* 478.1649 (calcd for C₂₇H₂₅FNO₆, 478.1621; error: +5.86 ppm). ¹H NMR (800 MHz, CD₃OD) δ 8.54 (br. s, 1H), 7.46 (dd, *J* = 8.4, 4.9 Hz), 7.21 (m, 1H), 6.99 (d, *J* = 8.7 Hz, 2H), 6.93 (d, *J* = 8.7 Hz, 2H), 6.62 (d, *J* = 8.7 Hz, 2H), 6.51 (d, *J* = 8.7 Hz, 2H), 2.02 (m, 1H), 1.92 (m, 2H), 1.16 (d, *J* = 6.7 Hz, 3H), 0.90 (d, *J* = 6.7 Hz, 3H) ¹³C NMR (200 MHz, CD₃OD) δ 202.2, 176.7, 173.0, 162.8, 160.5, 156.5, 144.1, 134.1, 133.7, 133.1, 131.6, 124.9, 122.2, 118.2, 117.3, 115.9, 115.7, 113.9, 94.0, 45.2, 25.0, 24.9, 24.3 ¹⁹F NMR (470 MHz, CD₃OD) δ –115.8

Cultivation and extraction of microbial strains for survey experiment

The same fermentation media and conditions described above were used to culture *Salinispora arenicola* sp. (SNA-091), *Streptomyces* sp. (RL10), *Erythrobacter flavus* sp. (SNC-119), *Streptomyces camosus* sp. (SNB-014), *Anthrobacter aurescens* sp. (SNG-016), *Micromonospora sediminicola* (SNC-018). From an overnight 50 mL culture, 10 mL/L of each culture was sterilely transferred into 21 x 2.8 L Fernbach flasks containing the seawater-based media and allowed to grow for 1 day shaking at 200 rpm. 500 mg of the respective probe was sterilely filtered into the cultures, and the cultures were allowed to grow for an additional 6 days. The cultures were extracted with EtOAc and concentrated under reduced pressure. The crude extracts were then subject to NMR analysis.

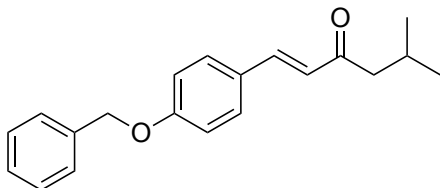
Synthesis of fluorinated discoipyrrole A (2.20)



4-(benzyloxy)benzaldehyde (2.16)

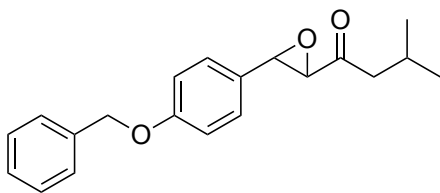
To a solution of 4-hydroxybenzaldehyde (4 g, 32.7 mmol) and cesium carbonate (10.7 g, 32.3 mmol) in DMF (75 mL), benzyl bromide (6.7 g, 39.3 mmol) was added, and the reaction stirred at room temperature for 24 hours under N₂. 100 mL of DI water was added, and the product was extracted with DCM (3 X 30 mL). The organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to provide a pale-yellow oil. The crude material was purified by column chromatography (Silica, stepwise to 20% EtOAc in hexanes) to give **2.16** as an off-white solid (6.3 g, 91% yield). LC-MS [M+H]⁺ *m/z* 213.2, [M+Na]⁺ *m/z* 235.2, [M+K]⁺ *m/z* 235.2. ¹H NMR (500 MHz, CDCl₃) δ 9.89 (s, 1H), 7.84 (d, *J* = 8.3 Hz, 2H), 7.46 – 7.39 (m, 4H),

7.38 – 7.33 (m, 1H), 7.08 (d, $J = 8.3$ Hz, 2H), 5.15 (s, 2H) ^{13}C NMR (200 MHz, CDCl_3) δ 190.9, 163.9, 136.1, 132.1, 130.2, 128.9, 128.5, 127.6, 122.1, 122.0, 115.3, 70.4



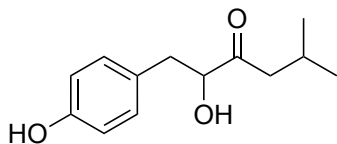
(E)-1-(4-(benzyloxy)phenyl)-5-methylhex-1-en-3-one (2.17)

4-Methylpentan-2-one (0.57 g, 5.66 mmol) and **2.16** (1.0 g, 4.72 mmol) were dissolved in a mixture of $\text{H}_2\text{O}:\text{EtOH}:\text{THF}$: (1:5:10 mL). The mixture was cooled to 0 °C and 2.5N NaOH (2.5 mL) was added and the reaction was brought to room temperature to stir for 24 hours. The reaction was again cooled to 0 °C and 2 N HCl was added dropwise until pH 7. The material was concentrated under reduced pressure, resulting in a yellow precipitate. The water was removed by filtration and the crude material was purified by column chromatography (Silica, stepwise to 85% EtOAc in hexanes) to give **2.17** as an off-white solid (1.10 g, 80%). LC-MS $[\text{M}+\text{H}]^+ m/z$ 295.3, $[\text{M}+\text{Na}]^+ m/z$ 317.2. ^1H NMR (500 MHz, CDCl_3) δ 7.53 – 7.47 (m, 3H), 7.45 – 7.37 (m, 4H), 7.34 (t, $J = 7.0$ Hz, 1H), 6.99 (d, $J = 8.4$ Hz, 2H), 6.63 (d, $J = 16.1$ Hz, 1H), 5.10 (s, 2H), 2.51 (d, $J = 7.0$ Hz, 2H), 2.27 – 2.19 (m, 1H), 0.98 (d, $J = 6.6$ Hz, 6H) ^{13}C NMR (200 MHz, CDCl_3) δ 200.5, 160.8, 142.3, 136.6, 130.1, 128.8, 128.3, 127.7, 127.6, 124.8, 115.4, 70.2, 50.0, 25.5, 22.9



1-(3-(4-(benzyloxy)phenyl)oxiran-2-yl)-3-methylbutan-1-one (2.18)

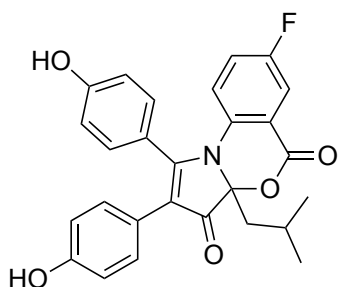
A solution of **2.17** (0.87 g, 2.96 mmol) in IPA (20 mL) was sonicated for 10 minutes and tetrabutylammonium peroxydisulfate (2.0 g, 2.96 mmol), H₂O₂ (30% in water, 0.5 mL) and NaOH (0.12 g, 2.96 mmol) were added. The reaction stirred for 16 hours and was quenched with saturated aq. NH₄Cl and extracted with ether. The organic layers were combined, dried over Na₂SO₄, and concentrated under reduced pressure to provide pale yellow crystals. The crude material was washed with cold EtOH to give pure **2.18** as colorless crystals (0.78 g, 85%). LC-MS [M+Na]⁺ *m/z* 333.3. ¹H NMR (800 MHz, CDCl₃) δ 7.42 (d, *J* = 7.5 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.33 (t, *J* = 7.3 Hz, 1H), 7.20 (d, *J* = 8.5 Hz, 2H), 6.97 (d, *J* = 8.5 Hz, 2H), 5.07 (s, 2H), 3.90 (s, 1H), 3.50 (s, 1H), 2.44 (dd, *J* = 16.3, 6.8 Hz, 1H), 2.29 (dd, *J* = 16.3, 7.2 Hz, 1H), 2.23 – 2.18 (m, 1H), 0.96 (t, *J* = 6.8 Hz, 6H) ¹³C NMR (200 MHz, CDCl₃) δ 206.0, 159.6, 136.8, 128.8, 128.2, 127.6, 127.3, 115.3, 70.2, 63.4, 58.0, 46.8, 24.4, 22.8, 22.7



2-hydroxy-1-(4-hydroxyphenyl)-5-methylhexan-3-one (4-hydroxysattabacin) (2.19)

A catalytic amount of Pd/C (0.02 g, 0.19 mmol) was added to a solution of **2.18** (0.10 g, 0.33 mmol) and ammonium formate (0.15 g, 2.37 mmol) in ethyl acetate (30 mL) and MeOH (10 mL) under N₂ and allowed to stir for 24 hours. The crude reaction mixture was then filtered through celite and washed with EtOAc (2 X 20 mL) then purified by column chromatography (Silica, stepwise to 30% EtOAc in hexanes). This provided pure **2.19** (0.03 g, 46%) as a yellow oil. Analytical data is consistent to what has been reported.¹⁰ ¹H NMR (500 MHz, CDCl₃) δ 7.05

(d, $J = 7.6$ Hz, 2H), 6.68 (d, $J = 7.6$ Hz, 2H), 5.97 (br. s, 1H), 4.35 (dd, $J = 7.6, 4.4$ Hz, 1H), 3.62 (br. s, 1H), 3.07 (dd, $J = 14.4, 4.4$ Hz, 1H), 2.74 (ddd, $J = 14.4, 7.6, 1.6$ Hz, 1H), 2.38 (d, $J = 6.8$ Hz, 2H), 2.23 – 2.14 (m, 1H), 0.92 (dd, $J = 6.8, 1.6$ Hz, 6H)



Fluorinated discoipyrrole A (2.20)

A solution of **2.19** (0.04 g, 0.17 mmol), 4-hydroxybenzaldehyde (0.04 g, 0.36 mmol), and 5-fluoroanthranilic acid (0.06 g, 0.36 mmol) in 5 mL DMSO was allowed to stir at room temperature for 24 hours, upon which 102 μ L of TFA was added. The reaction stirred for an additional 48 hours then quenched with saturated aq. NH_4Cl (10 mL) and extracted with DCM (3 X 20 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude material was then purified by reverse phase HPLC (Phenomenex Synergi Hydro-RP, 250 $\times$ 10 mm, 4 μ , 2 mL/min, UV = 254) using a gradient solvent system from 40% to 95% CH_3CN (0.1% formic acid) over 35 minutes to elute pure **2.20** (RT = 20.9 minutes, 0.003 g, 37% yield). HRESIMS $[\text{M}+\text{H}]^+$ m/z 460.1563 (calcd for $\text{C}_{27}\text{H}_{23}\text{FNO}_6$, 460.1516; error: 10.2 ppm) **^1H NMR (500 MHz, CD_3OD)** δ 7.74 (dd, $J = 8.0, 3.0$ Hz, 1H), 7.24 (td, $J = 8.6, 3.0$ Hz, 1H), 7.08 (d, $J = 8.3$ Hz, 2H), 6.96 (d, $J = 8.6$ Hz, 2H), 6.86 (d, $J = 8.3$ Hz, 2H), 6.63 (d, $J = 8.6$ Hz, 2H), 6.51 (dd, $J = 9.0, 4.3$ Hz, 1H), 2.21 (dd, $J = 14.1, 6.1$ Hz, 1H), 2.00 (dd, $J = 14.1, 6.5$ Hz, 1H), 1.76 (m, 1H), 0.95 (d, $J = 6.7$ Hz, 3H), 0.85 (d, $J = 6.7$ Hz, 3H). **^{13}C NMR (200 MHz, CD_3OD)** δ 196.2, 162.2, 161.2, 160.8, 157.7, 141.4, 138.1, 135.0, 131.7, 131.4, 126.0, 121.6, 120.9, 117.3, 117.0, 116.4, 115.9, 92.9, 42.8, 25.3, 24.2, 23.4, **^{19}F NMR (470 MHz, CD_3OD)** δ -117.2

2.8 References

1. Lin, Z.; Koch, M.; Abdel Aziz, M. H.; Galindo-Murillo, R.; Tianero, M. D.; Cheatham, T. E.; Barrows, L. R.; Reilly, C. A.; Schmidt, E. W. Oxazinin A, a Pseudodimeric Natural Product of Mixed Biosynthetic Origin from a Filamentous Fungus. *Organic Letters* **2014**, *16* (18), 4774-4777.
2. Mevers, E.; Saurí, J.; Liu, Y.; Moser, A.; Ramadhar, T. R.; Varlan, M.; Williamson, R. T.; Martin, G. E.; Clardy, J. Homodimericin A: A Complex Hexacyclic Fungal Metabolite. *Journal of the American Chemical Society* **2016**, *138* (38), 12324-12327.
3. Pan, E.; Jamison, M.; Yousufuddin, M.; MacMillan, J. B. Ammosamide D, an Oxidatively Ring Opened Ammosamide Analog from a Marine-Derived *Streptomyces variabilis*. *Organic Letters* **2012**, *14* (9), 2390-2393.
4. Hughes, C. C.; MacMillan, J. B.; Gaudêncio, S. P.; Jensen, P. R.; Fenical, W. The Ammosamides: Structures of Cell Cycle Modulators from a Marine-Derived *Streptomyces* Species. *Angewandte Chemie International Edition* **2009**, *48* (4), 725-727.
5. Hu, Y.; Potts, M. B.; Colosimo, D.; Herrera-Herrera, M. L.; Legako, A. G.; Yousufuddin, M.; White, M. A.; MacMillan, J. B. Discoipyrroles A–D: Isolation, Structure Determination, and Synthesis of Potent Migration Inhibitors from *Bacillus hunanensis*. *Journal of the American Chemical Society* **2013**, *135* (36), 13387-13392.
6. Fu, P.; Legako, A.; La, S.; MacMillan, J. B. Discovery, Characterization, and Analogue Synthesis of Bohemamine Dimers Generated by Non-enzymatic Biosynthesis. *Chemistry – A European Journal* **2016**, *22* (10), 3491-3495.
7. Mevers, E.; Saurí, J.; Helfrich, E. J. N.; Henke, M.; Barns, K. J.; Bugni, T. S.; Andes, D.; Currie, C. R.; Clardy, J. Pyonitrins A–D: Chimeric Natural Products Produced by *Pseudomonas protegens*. *Journal of the American Chemical Society* **2019**, *141* (43), 17098-17101.
8. Hughes, C. C.; MacMillan, J. B.; Gaudêncio, S. P.; Fenical, W.; La Clair, J. J. Ammosamides A and B Target Myosin. *Angewandte Chemie International Edition* **2009**, *48* (4), 728-732.
9. Pan, E.; Oswald, N. W.; Legako, A. G.; Life, J. M.; Posner, B. A.; MacMillan, J. B. Precursor-directed generation of amidine containing ammosamide analogs: ammosamides E–P. *Chem. Sci.* **2013**, *4* (1), 482-488.
10. Colosimo, D. A.; MacMillan, J. B. Detailed Mechanistic Study of the Non-enzymatic Formation of the Discoipyrrole Family of Natural Products. *Journal of the American Chemical Society* **2016**, *138* (7), 2383-2388.
11. Shingare, R. D.; Aniebok, V.; Lee, H.-W.; MacMillan, J. B. Synthesis and Investigation of the Abiotic Formation of Pyonitrins A–D. *Organic Letters* **2020**, *22* (4), 1516-1519.

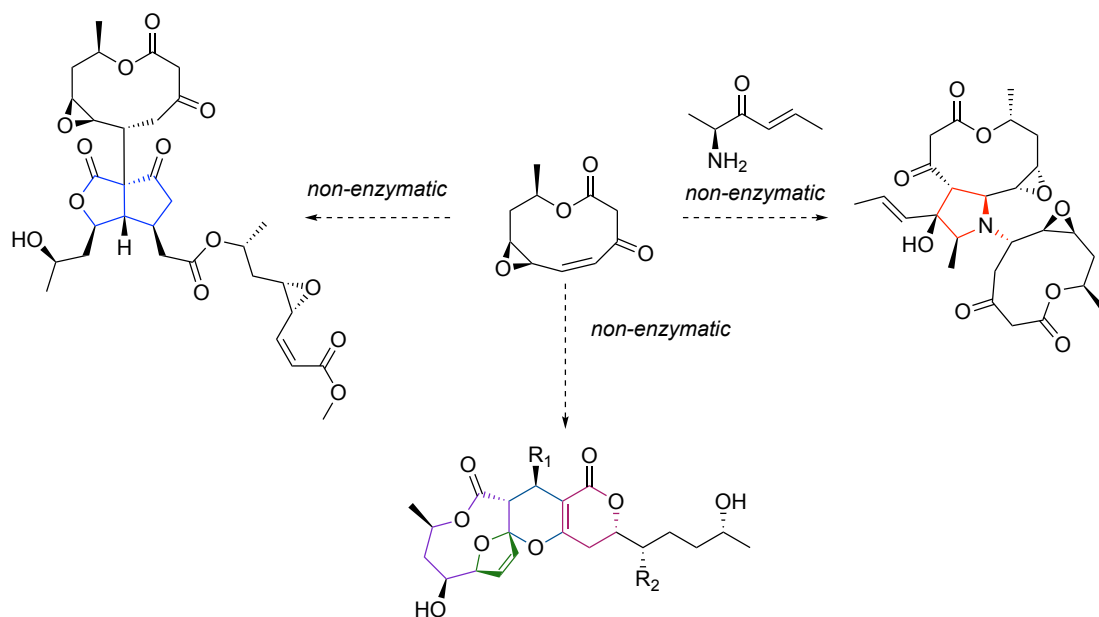
12. Grabowski, K.; Baringhaus, K.-H.; Schneider, G. Scaffold diversity of natural products: inspiration for combinatorial library design. *Natural Product Reports* **2008**, 25 (5), 892-904.
13. Ang, E. L.; Sun, H.; Liu, Z.; Zhao, H. Recent advances in combinatorial biosynthesis for drug discovery. *Drug Design, Development and Therapy* **2015**, 9, 823-833.
14. Feher, M.; Schmidt, J. M. Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. *Journal of Chemical Information and Computer Sciences* **2002**, 43 (1), 218-227.
15. Barnes, E. C.; Kumar, R.; Davis, R. A. The use of isolated natural products as scaffolds for the generation of chemically diverse screening libraries for drug discovery. *Natural Product Reports* **2016**, 33 (3), 372-381.
16. Buryanovskyy, L.; Fu, Y.; Boyd, M.; Ma, Y.; Hsieh, T.-c.; Wu, J. M.; Zhang, Z. Crystal Structure of Quinone Reductase 2 in Complex with Resveratrol. *Biochemistry* **2004**, 43 (36), 11417-11426.
17. Yan, Y.; Yang, J.; Yu, Z.; Yu, M.; Ma, Y.-T.; Wang, L.; Su, C.; Luo, J.; Horsman, G. P.; Huang, S.-X. Non-enzymatic pyridine ring formation in the biosynthesis of the rubrolone tropolone alkaloids. *Nature Communications* **2016**, 7 (1), 13083.
18. Dupuis, S. N.; Veinot, T.; Monroe, S. M. A.; Douglas, S. E.; Syvitski, R. T.; Goralski, K. B.; McFarland, S. A.; Jakeman, D. L. Jadomycins Derived from the Assimilation and Incorporation of Norvaline and Norleucine. *Journal of Natural Products* **2011**, 74 (11), 2420-2424.
19. Borissow, C. N.; Graham, C. L.; Syvitski, R. T.; Reid, T. R.; Blay, J.; Jakeman, D. L. Stereochemical Integrity of Oxazolone Ring-Containing Jadomycins. *ChemBioChem* **2007**, 8 (10), 1198-1203.
20. Robertson, A. W.; Martinez-Farina, C. F.; Smithen, D. A.; Yin, H.; Monroe, S.; Thompson, A.; McFarland, S. A.; Syvitski, R. T.; Jakeman, D. L. Eight-Membered Ring-Containing Jadomycins: Implications for Non-enzymatic Natural Products Biosynthesis. *Journal of the American Chemical Society* **2015**, 137 (9), 3271-3275.
21. Namdeo, A. G.; Jadhav, T. S.; Rai, P. K.; Gavali, S.; Mahadik, K. R. Precursor feeding for enhanced production of secondary metabolites. *Pharmacognosy Reviews* **2007**, 1 (2), 227-231.
22. Zhang, Y.; Colabroy, K. L.; Begley, T. P.; Ealick, S. E. Structural Studies on 3-Hydroxyanthranilate-3,4-dioxygenase: The Catalytic Mechanism of a Complex Oxidation Involved in NAD Biosynthesis. *Biochemistry* **2005**, 44 (21), 7632-7643.
23. Davis, I.; Liu, A. What is the tryptophan kynurenine pathway and why is it important to neurotherapeutics? *Expert Review of Neurotherapeutics* **2015**, 15 (7), 719-721.
24. Lovelace, M. D.; Varney, B.; Sundaram, G.; Franco, N. F.; Ng, M. L.; Pai, S.; Lim, C. K.; Guillemin, G. J.; Brew, B. J. Current Evidence for a Role of the Kynurenine Pathway of Tryptophan Metabolism in Multiple Sclerosis. *Frontiers in Immunology* **2016**, 7, 246.

25. Savitz, J. The kynurenine pathway: a finger in every pie. *Molecular Psychiatry* **2019**, 25 (1), 131-147.
26. Cristina, B.; Veronica, R.; Silvia, A.; Andrea, G.; Sara, C.; Luca, P.; Nicoletta, B.; M.C, B. J.; Silvio, B.; Fabio, T. Identification and characterization of the kynurenine pathway in the pond snail *Lymnaea stagnalis*. *Scientific Reports* **2022**, 12 (1), 15617.
27. Jamshed, L.; Debnath, A.; Jamshed, S.; Wish, J. V.; Raine, J. C.; Tomy, G. T.; Thomas, P. J.; Holloway, A. C. An Emerging Cross-Species Marker for Organismal Health: Tryptophan-Kynurenine Pathway. *International Journal of Molecular Sciences* **2022**, 23 (11), 6300.
28. Zhang, J.; Qian, Z.; Wu, X.; Ding, Y.; Li, J.; Lu, C.; Shen, Y. Juanlimycins A and B, Ansamycin Macrolactams from *Streptomyces* sp. *Organic Letters* **2014**, 16 (10), 2752-2755.
29. Miyanaga, A. Structure and function of polyketide biosynthetic enzymes: various strategies for production of structurally diverse polyketides. *Bioscience, Biotechnology, and Biochemistry* **2017**, 81 (12), 2227-2236.
30. Liou, G. Building-block selectivity of polyketide synthases. *Current Opinion in Chemical Biology* **2003**, 7 (2), 279-284.
31. Franck, B. Key Building Blocks of Natural Product Biosynthesis and Their Significance in Chemistry and Medicine. *Angewandte Chemie International Edition in English* **2003**, 18 (6), 429-439.
32. O'Hagan, D.; Harper, D. Fluorine-containing natural products. *Journal of Fluorine Chemistry* **1999**, 100 (1-2), 127-133.
33. Chandra, G.; Singh, D. V.; Mahato, G. K.; Patel, S. Fluorine-a small magic bullet atom in the drug development: perspective to FDA approved and COVID-19 recommended drugs. *Chemical Papers* **2023**, 77 (8), 4085-4106.
34. Rizzo, C.; Amata, S.; Pibiri, I.; Pace, A.; Buscemi, S.; Palumbo Piccionello, A. FDA-Approved Fluorinated Heterocyclic Drugs from 2016 to 2022. *International Journal of Molecular Sciences* **2023**, 24 (9), 7728.
35. Gimenez, D.; Phelan, A.; Murphy, C. D.; Cobb, S. L. ^{19}F NMR as a tool in chemical biology. *Beilstein Journal of Organic Chemistry* **2021**, 17, 293-318.
36. Kiel, M.; Engesser, K.-H. The biodegradation vs. biotransformation of fluorosubstituted aromatics. *Applied Microbiology and Biotechnology* **2015**, 99 (18), 7433-7464.
37. Wang, M.; Xu, J.; Wang, J.; Wang, S.; Feng, H.; Shentu, J.; Shen, D. Differences between 4-fluoroaniline degradation and autoinducer release by *Acinetobacter* sp. TW: implications for operating conditions in bacterial bioaugmentation. *Environmental Science and Pollution Research* **2013**, 20 (9), 6201-6209.
38. Brumley, D. A. Precursor-Directed Identification of Natural Product Scaffolds. *Master's Thesis*. UT Southwestern Medical Center, 2017.

39. Martin, G. E.; Hadden, C. E. Long-Range ^1H - ^{15}N Heteronuclear Shift Correlation at Natural Abundance. *Journal of Natural Products* **2000**, 63 (4), 543-585.
40. Hidde Boersma, F. G.; Colin McRoberts, W.; Cobb, S. L.; Murphy, C. D. A ^{19}F NMR study of fluorobenzoate biodegradation by *Sphingomonas* sp. HB-1. *FEMS Microbiology Letters* **2004**, 237 (2), 355-361.
41. Murphy, C. D. Biodegradation and biotransformation of organofluorine compounds. *Biotechnology Letters* **2009**, 32 (3), 351-359.
42. Murphy, C. D. The Application of ^{19}F Nuclear Magnetic Resonance to Investigate Microbial Biotransformations of Organofluorine Compounds. *OMICS: A Journal of Integrative Biology* **2007**, 11 (3), 314-324.
43. Williams, D. E.; Morgan, K. D.; Dalisay, D. S.; Matainaho, T.; Perrachon, E.; Viller, N.; Delcroix, M.; Gauchot, J.; Niikura, H.; Patrick, B. O.; et al. Natural Products Produced in Culture by Biosynthetically Talented *Salinispora arenicola* Strains Isolated from Northeastern and South Pacific Marine Sediments. *Molecules* **2022**, 27 (11), 3569.
44. Setiyono, E.; Heriyanto; Pringgenies, D.; Shioi, Y.; Kanasaki, Y.; Awai, K.; Brotosudarmo, T. H. P. Sulfur-Containing Carotenoids from A Marine Coral Symbiont *Erythrobacter flavus* Strain KJ5. *Marine Drugs* **2019**, 17 (6), 349.
45. Doroghazi, J. R.; Metcalf, W. W. Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* **2013**, 14 (1), 611.
46. Mongodin, E. F.; Shapir, N.; Daugherty, S. C.; DeBoy, R. T.; Emerson, J. B.; Shvartzbeyn, A.; Radune, D.; Vamathevan, J.; Riggs, F.; Grinberg, V.; et al. Secrets of Soil Survival Revealed by the Genome Sequence of *Arthrobacter aureus* TC1. *PLoS Genetics* **2006**, 2 (12), e214.
47. Hifnawy, M. S.; Fouda, M. M.; Sayed, A. M.; Mohammed, R.; Hassan, H. M.; AbouZid, S. F.; Rateb, M. E.; Keller, A.; Adamek, M.; Ziemert, N.; Abdelmohsen, U. R. The genus *Micromonospora* as a model microorganism for bioactive natural product discovery. *RSC Advances* **2020**, 10 (35), 20939-20959.
48. León, B.; Navarro, G.; Dickey, B. J.; Stepan, G.; Tsai, A.; Jones, G. S.; Morales, M. E.; Barnes, T.; Ahmadyar, S.; Tsiang, M.; et al. Abyssomicin 2 Reactivates Latent HIV-1 by a PKC- and HDAC-Independent Mechanism. *Organic Letters* **2015**, 17 (2), 262-265.
49. Savic, V. Abyssomicins. In *Studies in Natural Products Chemistry*, Studies in Natural Products Chemistry, Vol. 40; Elsevier, 2013; pp 133-172.
50. Ni, C.; Hu, J. The unique fluorine effects in organic reactions: recent facts and insights into fluoroalkylations. *Chemical Society Reviews* **2016**, 45 (20), 5441-5454.

Chapter 3

Biomimetic Synthesis and Investigation of the Non-enzymatic Formation of the Tricrilactones



3.1 Introduction

Macrolides are a ubiquitous class of natural products produced from a variety of plants and microorganisms, primarily derived from polyketide synthase (PKS) pathways.¹⁻³ These compounds are characterized by their intricate large macrocyclic lactone framework and exhibit expansive chemical and biological diversity (Figure 3.1).³⁻⁵ Most commonly, macrolides are comprised of 12–16 membered ring systems, however, larger macrolides harboring 28, 38, and 66-membered rings have also been reported.^{2, 4} Additionally, macrolides exhibit a wide range of pharmacological activities, including antibacterial, antifungal, antiparasitic, antiviral, herbicidal, and cytotoxic properties.^{2, 5-7} Given their structural complexity and vast applicability in drug discovery, macrolides remain a dynamic and evolving area of research in medicinal chemistry and natural product synthesis.

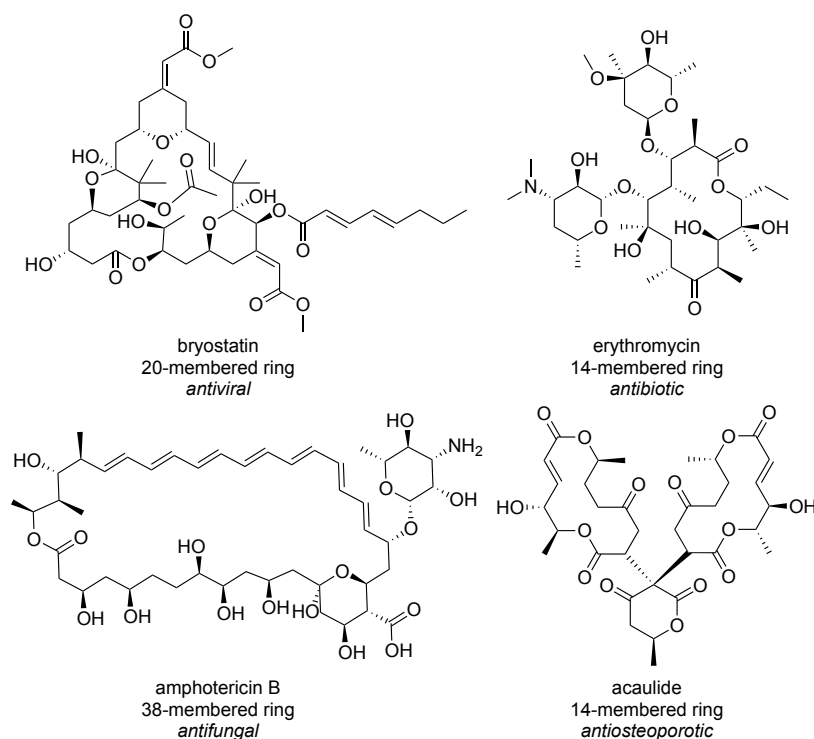


Figure 3.1. Representation of structural diversity and broad bioactivity of macrolides.^{2-4, 7}

In 2023, a new family of structurally intriguing macrolides, the tricrilactones, were discovered from the understudied Alpine-moss associated fungal species *Trichocladium crispatum* (Figure 3.2).⁸ Molecular networking guided fractionation led to the isolation of the 10-membered macrolides, tricrilactone A–H, which were then subject to bioactivity testing. Interestingly, all compounds exhibited potent antiosteoporotic activity, reducing bone loss in zebrafish models more effectively than the clinically approved antiosteoporosis drug, alendronate.^{8, 9} Most notable, however, is their unprecedented ring systems and uniquely complex core scaffolds.

Tricrilactone A (**3.1**) presents the first of its kind as a pseudo trimeric macrolide with a C–C linkage between the monomeric species to form the unusual tricyclic core.⁸ Tricrilactone B–F (**3.2–3.6**) are pseudodimers connected by a methylene bridge, similar to that observed in the dibohemamines described in Chapter 1.¹⁰ Heterodimeric tricrilactone B (**3.2**), C (**3.3**), E (**3.5**), and F (**3.6**) contain a novel 9/5/6/6 scaffold possessing an uncommon 1,5-dioxanon-2-one moiety. Alternatively, tricrilactone D (**3.4**) is a homodimer with a fused 9/5/6/10/3 pentacyclic ring moiety, also an anomalous motif. Tricrilactone G (**3.7**) is the first of its class of hybrid polyketide–alkaloid structures with a pyrrolidine center that forms the bridged azatricyclic framework. Lastly, tricrilactone H is another pseudodimer, consisting of two 9-membered monomeric units bridged by a C–C bond, similar to tricrilactone A (Figure 3.2).⁸

What makes the tricrilactone family exceptionally appealing, and the inspiration of this project, is their postulated mechanism of formation, believed to arise from a series of non-enzymatic cascades. Herein, this work is an expansion of exploring non-enzymatic transformations and building tools for discovery and investigation through the biomimetic synthesis.

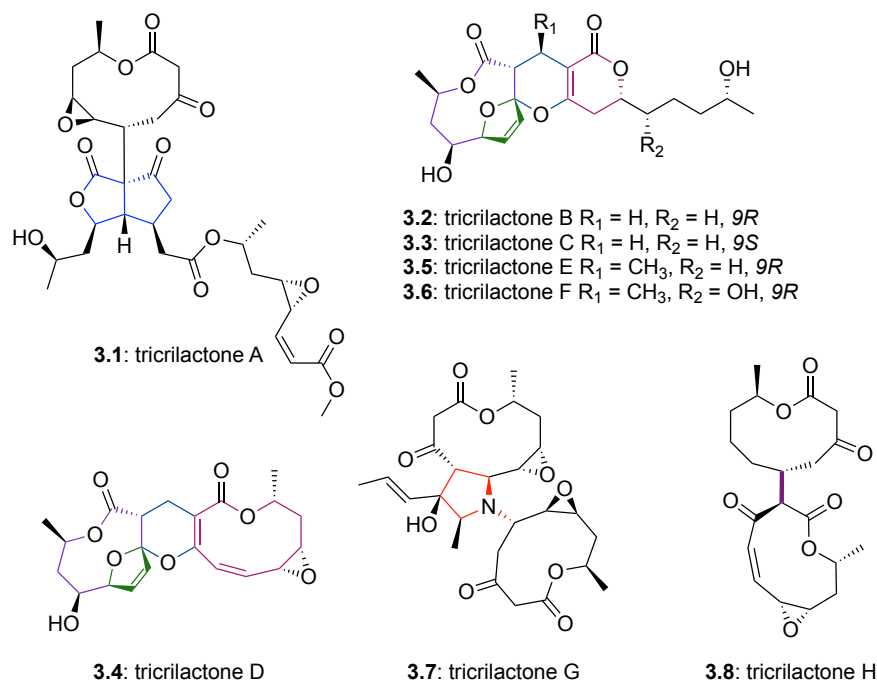


Figure 3.2. The tririlactone family of natural products. Outlined in color are the distinct structural features for each molecule.⁸

3.2 Putative biosynthetic pathways and preliminary mechanistic studies of the tririlactones

Investigation of the biosynthetic origin of the tririlactones led to the identification of a novel precursor protilactone A (**3.8**), which is structurally similar to the fungal metabolite, botryolide B, conceivably arising from an oxidation event (Figure 3.3).^{8, 11} Botryolide B is a unique member of the decarestrictine family of decanolide natural products, featuring a rare *cis* configured olefin whereas most other analogs contain *trans* olefins.^{11, 12} Biosynthetic studies have shown a single pentaketide intermediate serves as the precursor to all decarestrictine analogs that undergoes various modifications, much like the proposed pathway to the tririlactones.^{12, 13} Importantly, these analogs can non-enzymatically interconvert in a pH-dependent manner (Figure 3.3).^{12, 14} Given the structural similarities between botryolide B, the decarestrictine family, and the

proposed precursor to the tricrilactones, it is reasonable to assume that **3.8** can also engage in non-enzymatic chemistry, supporting this proposed pathway.

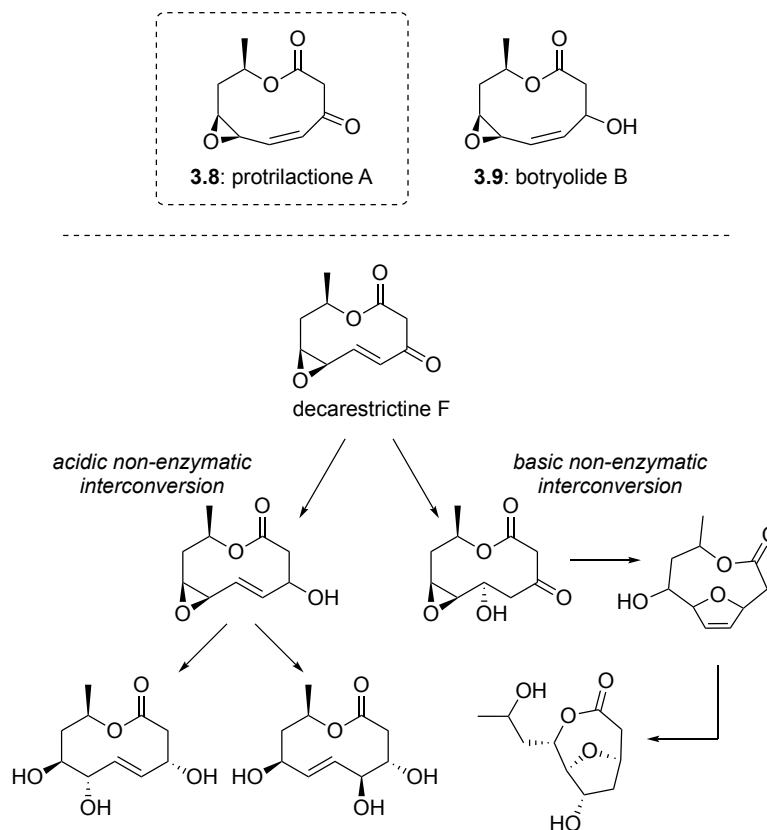


Figure 3.3. Structural similarities between protricrilactone A and decarestrictine analogs that undergo non-enzymatic interconversion.¹²⁻¹⁴

The identification of the precursor then led to the identification of two more intermediates, **3.11** and **3.14** by LC–MS and confirmed by single-crystal X-ray diffraction, gave rise to a proposed biosynthetic pathway (Figure 3.4).⁸ Derived from malonyl-CoA and acetyl-CoA, pentaketide-CoA (**3.10**) undergoes a cyclization and oxidation transformation to afford **3.8**, the novel precursor to all the tricrilactones. For the formation of tricrilactone A (**3.1**), it is proposed that the **3.8** intramolecularly cyclizes at the C5 to C9 position to provide **3.12** containing a 7/5-fused ring system. Intermediate **3.12** undergoes an intramolecular

transesterification at the C4 to C10 position to form the 5/5 fused bicyclic skeleton of **3.13** followed by two consecutive Michael additions of the monomeric precursor **3.8**, to first form protirilactone B (**3.14**) and subsequently **3.15**. A retro-Dieckmann reaction with methanol is proposed to be the final step in the formation of **3.1**, cleaving at the C29 position of the monomeric unit installed from the second Michael addition to provide the novel, pseudo trimeric macrolide scaffold.⁸

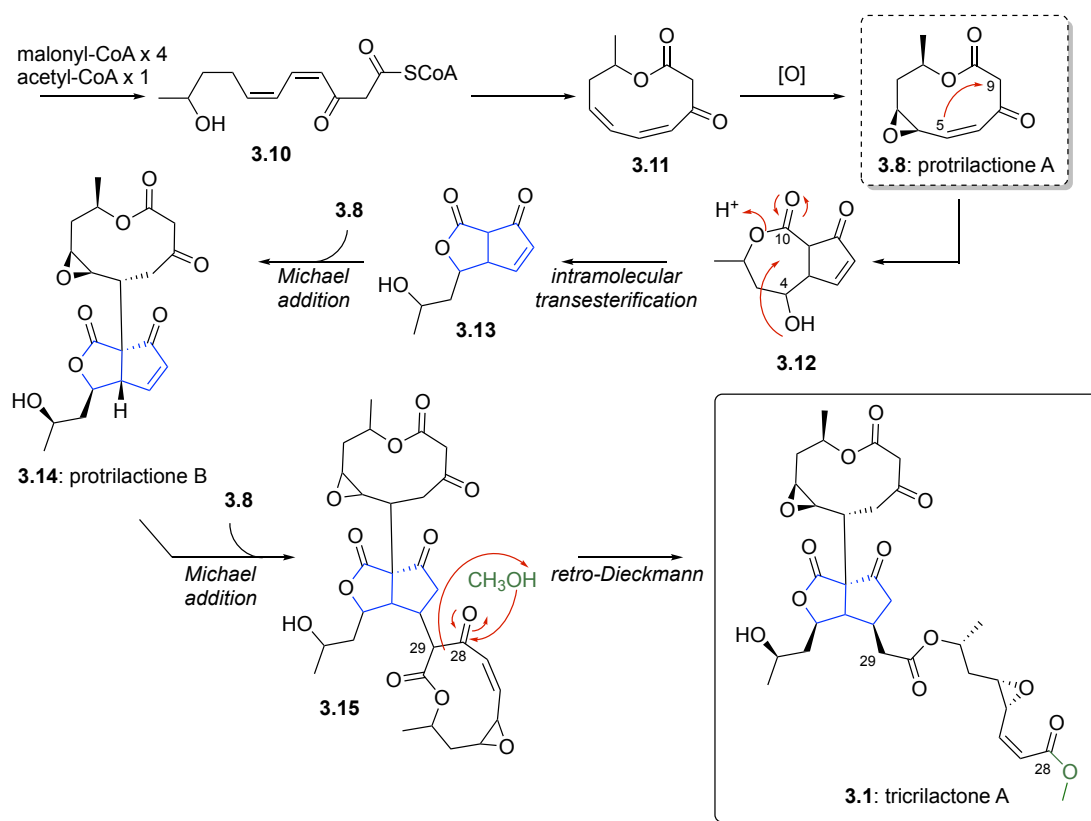


Figure 3.4. Proposed biosynthetic pathway of tririlactone A.⁸

The formation of tririlactone B–D follows a similar series of Michael additions and intramolecular cyclizations, however, the additional methylene bridge is not accounted for in the monomeric unit **3.8**.⁸ As fungi have the ability to produce formaldehyde by the oxidation of methanol, it is proposed that the C1 extension arises from the non-enzymatic incorporation of

formaldehyde in a similar fashion to other bridged dimers such as the dibohemamines and the fungal metabolites aspergilone B and dictrinone C depicted in Chapter 1 (Figure 1.12, section 1.4.1).^{10, 15-18} The authors suggest that **3.8** first reacts with formaldehyde to generate **3.16**, the precursor to **3.2–3.4**. For the formation of **3.2** and **3.3**, a Michael addition of **3.11** occurs at the terminal alkene of **3.16** to provide **3.17** that is subject to a sequential reduction and hydrolysis of the epoxide to afford **3.18**. Intermediate **3.18** then undergoes an intramolecular cyclization at the C5 to C8 and C8 to C18 positions followed by the addition of water to generate **3.19**. The final step in the formation of **3.2** and **3.3** is an intramolecular transesterification from the hydroxyl group of C16 to C20, yielding the two diastereomers. Alternatively, Michael addition of **3.8**, rather than **3.11**, with **3.16** occurs giving rise to intermediate **3.20**. Then the same hydration and intramolecular transesterification process is proposed to afford **3.4** (Figure 3.5).⁸

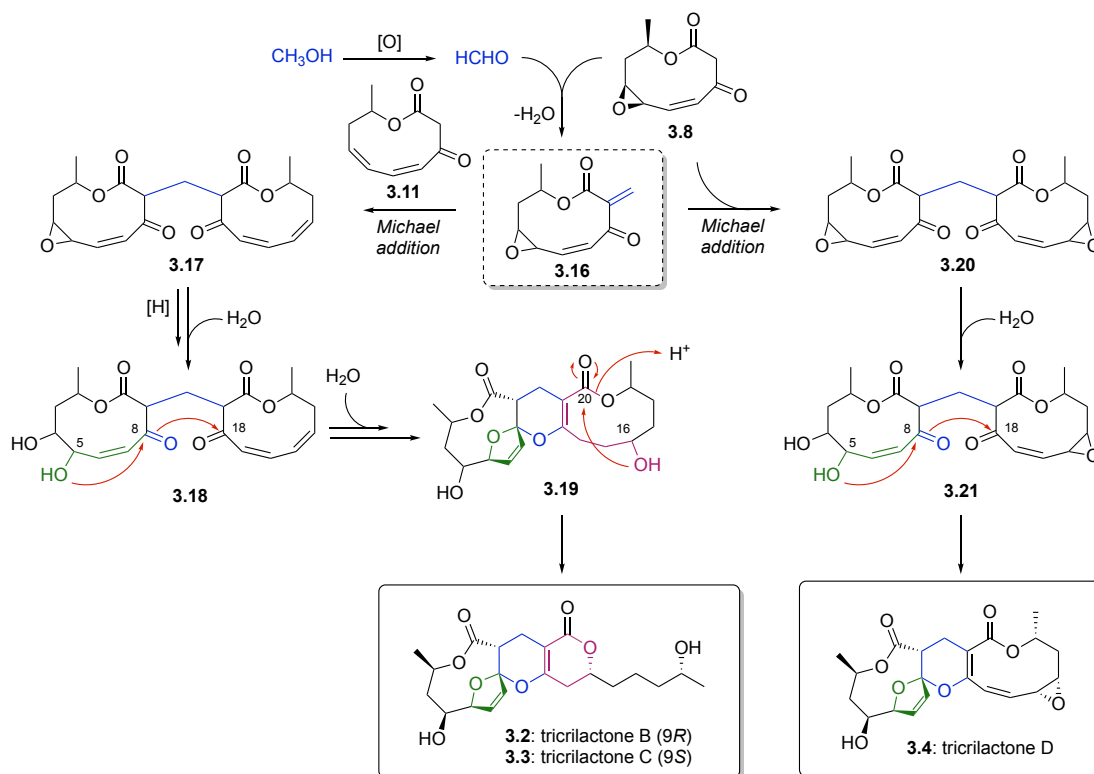


Figure 3.5. Proposed biosynthetic pathway for tricrilactone B–D.⁸

Preliminary studies using isotopic feeding experiments have been conducted to corroborate the hypothesis that the methylene bridge of compounds **3.2–3.4** results from the non-enzymatic incorporation of formaldehyde.⁸ Fungal cultures were supplemented with deuterated methanol at various stages of growth and analyzed by LC-ESI-MS, revealing *m/z* values with 2 Dalton (Da) increments compared to the original isolates. Furthermore, compound **3.8** was treated with an aqueous solution of formaldehyde under physiological conditions, forming another methylene bridged dimer. These initial studies suggest that formaldehyde-induced dimerization is initiated by the oxidation of methanol and can be achieved in the absence of enzymes, potentially emulating the biosynthetic pathway of **3.2–3.4**.⁸ Although these are encouraging results, more in-depth mechanistic studies are needed to further validate this hypothesis.

The formation of the novel azatricyclic skeleton of tricrilactone is proposed to arise from a unique pathway, which we were particularly interested in. A Claisen condensation between diketide-CoA and L-alanine forms an α,β -unsaturated ketoamine coupling partner, **3.22**, that participates in an aza-Michael addition to afford intermediate **3.23** (Figure 3.6).⁸ This unique Claisen condensation has been observed in the biosynthetic pathway of curvulamine and the fumonisins, however, is relatively uncommon.^{19, 20} Presumably, **3.23** can then undergo an aldol condensation between C7 and C4, generating **3.24**, followed by a second aza-Michael addition to yield and tricrilactone G (**3.7**).

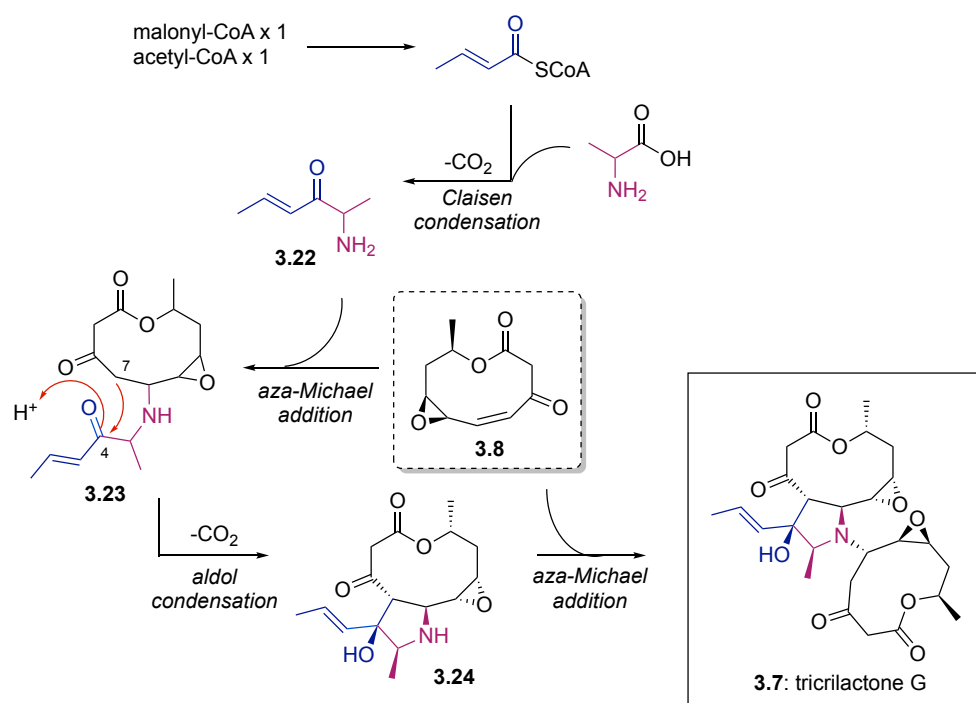
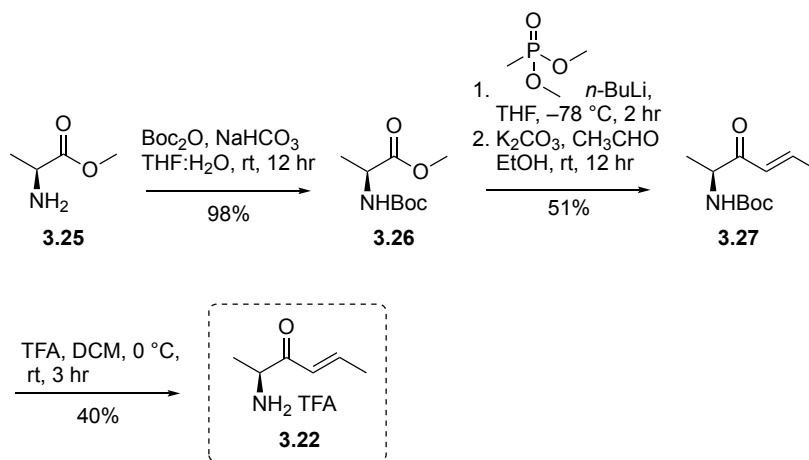


Figure 3.6. Proposed biosynthetic pathway for tricrilactone G.⁸

The studies described above provided preliminary evidence that the tricrilactones are formed through non-enzymatic chemistry, which we sought to explore further. We were particularly interested in tricrilactone G due to its unique pyrrolidine core originating from L-alanine and it being the first of its class of bridged azatricyclic scaffolds. Moreover, given the antiosteoporotic activity of tricrilactone G, understanding the non-enzymatic mechanism presents an opportunity for harnessing the inherent reactivity to generate novel compound libraries with diverse amino acid derivatives. Though we did not investigate the biological aspects, this research could pave the way for exploring and optimizing the medicinal potential of this natural product family. Therefore, we took a biomimetic approach by synthesizing the amine coupling partner **3.22** and the key intermediate protirlactone A, **3.8**.

3.3. Synthesis of amine fragment 3.22

The synthesis of **3.22** began with the *tert*-butoxycarbonyl (Boc)-protection of L-alanine methyl ester hydrochloride (**3.25**) with di-*tert*-butyl dicarbonate (Boc_2O) and sodium bicarbonate (NaHCO_3) to afford **3.26** (98%) (Scheme 3.1).²¹ The α,β -unsaturated ketoamine **3.22** was prepared by a Horner-Wadsworth-Emmons reaction between **3.26** and dimethyl methylphosphonate in the presence of *n*-butyllithium (*n*-BuLi) at -78°C , generating a phosphonium ylide, which was subjected to a Wittig olefination with acetyl aldehyde and potassium carbonate (K_2CO_3), resulting in a 51% yield of **3.27**.^{22, 23} A final Boc-deprotection with trifluoroacetic acid (TFA) provided the final amine fragment **3.22** as a TFA salt in 40% yield.



Scheme 3.1. Synthetic scheme for the amine coupling partner 3.22.²¹⁻²³

3.4 Synthesis of protilactone A

As previously mentioned, the monomer **3.8** is structurally similar to botryolide B (**3.9**), which has been well studied and many total syntheses for both decastrictine and **3.9** have been reported.^{11, 12, 24, 25} The most concise synthetic pathway to **3.9** includes a Yamaguchi

esterification between fragment **3.28** and fragment **3.29**, followed by a ring closing metathesis (RCM) (Figure 3.7).²⁴ Lastly, oxidation of **3.9** will provide the primary diketone precursor **3.8** for the investigation of the tricrilactones.

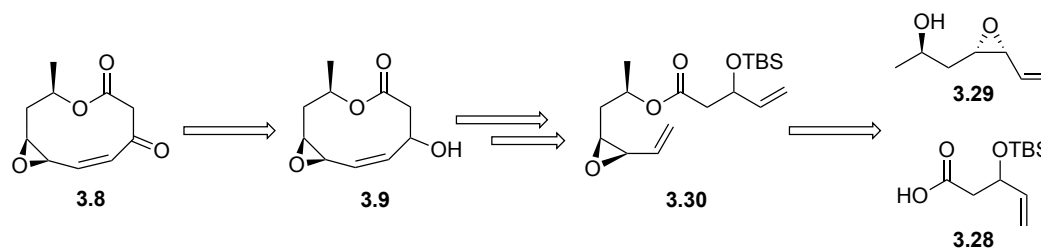
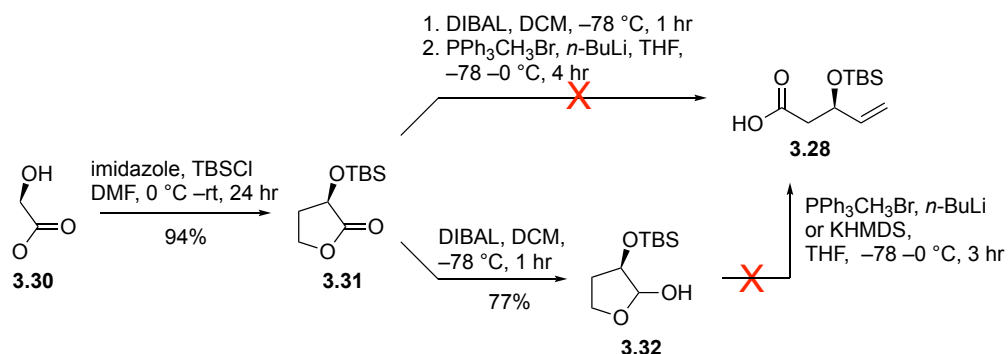


Figure 3.7. Retrosynthesis of protricrilactone A.^{24, 25}

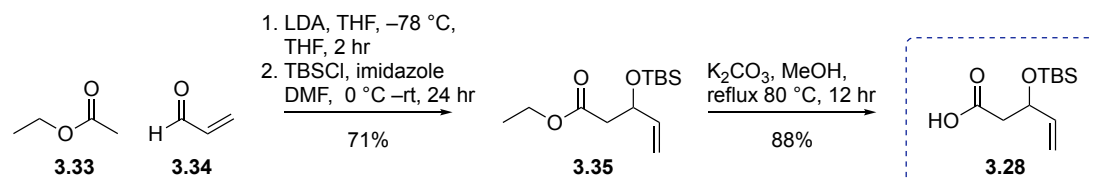
3.4.1 Synthesis of acid coupling partner **3.28**

We first attempted the synthesis of fragment **3.28** by following a previously established protocol by Mohapatra et al. in their total synthesis of botryolide B (Scheme 3.2).²⁴ Intermediate **3.31** was prepared by silyl protection of commercially available (*R*)-(+)- α -hydroxy- γ -butyrolactone (**3.30**), using *tert*-butyldimethylsilyl chloride (TBSCl) and imidazole at $-78\text{ }^{\circ}\text{C}$, resulting in a 94% yield. Reduction of lactone **3.31** to lactol **3.32** with diisobutylaluminum hydride (DIBAL) at $-78\text{ }^{\circ}\text{C}$ followed by one-carbon homologation via Wittig olefination with methyltriphenylphosphonium bromide ($\text{PPh}_3\text{CH}_3\text{Br}$) and *n*-BuLi, was reported to provide **3.28** in yields over 80% as a one-pot reaction.²⁴ However, we were unable to obtain the desired product in sufficient yields to proceed. Therefore, the protocol was then modified to become two independent steps. The DIBAL reduction afforded **3.31** in 77% yield after 1 hour at $-78\text{ }^{\circ}\text{C}$, but the subsequent olefination with $\text{PPh}_3\text{CH}_3\text{Br}$ and *n*-BuLi at $-78\text{ }^{\circ}\text{C}$ did not provide an appreciable amount of product. Repeating the Wittig olefination with potassium bis(trimethylsilyl)amide (KHMDs) as an alternative base also did not result in the desired product.



Scheme 3.2. Original synthetic scheme for the synthesis of fragment 3.28.²⁴

After the unsuccessful attempts with the lactone starting material, we adopted a different methodology, beginning with an aldol addition of ethyl acetate (**3.33**) and acrolein (**3.34**) in the presence of lithium diisopropylamide (LDA) (Scheme 3.3).^{26, 27} This provided a β -hydroxy ester that was subject to an immediate TBSCl protection of the hydroxyl group in the presence of imidazole, resulting in **3.35** as a racemate in 71% yield. Saponification of **3.35** with K_2CO_3 in methanol (MeOH) at $80\text{ }^{\circ}\text{C}$ afforded fragment **3.28** (88%).²⁷ It is worth noting that simultaneous hydrolysis of both the ethyl ester and silyl protecting group was occasionally observed. This could be attributed to the acidic work-up for this reaction with 2 N hydrochloric acid (HCl). Achieving a pH below 5 is necessary to acidify the carboxylate to provide the carboxylic acid, but excessive acid can consequently hydrolyze both functional groups. Therefore, careful acidification is advised.



Scheme 3.3. Optimal synthetic pathway for fragment 3.28.^{26, 27}

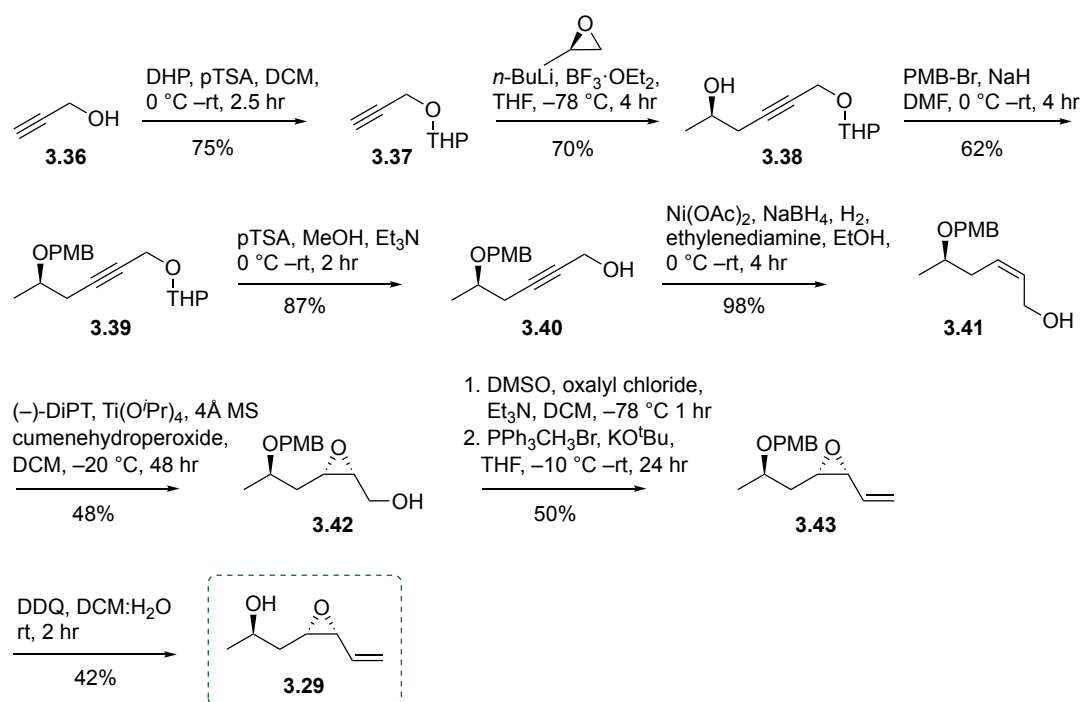
3.4.2 Synthesis of fragment 3.29

The synthesis of **3.29** began with the protection of the primary alcohol in propargyl alcohol (**3.36**) with 3,4-dihydropyran (DHP) in the presence of *p*-toluenesulfonic acid (pTSA), providing **3.37** in 75% yield (Scheme 3.4).²⁸ (S)-Propylene oxide underwent a stereoselective epoxide ring opening with **3.37** in the presence of *n*-BuLi, and boron trifluoride diethyl etherate (BF₃·OEt₂) to give **3.38** in 70% yield.²⁹ The secondary alcohol of **3.38** was protected with 4-methoxybenzyl bromide (PMB-Br) to give **3.39** in a 62% yield, followed by a selective deprotection of the THP group to afford the allylic alcohol **3.40** (87%).^{25, 29} A nickel-catalyzed partial reduction of **3.40** with nickel acetate (Ni(OAc)₂) and sodium borohydride (NaBH₄) under H₂ atmosphere provided the *cis*-allylic alcohol **3.41** in excellent yields (98%).

Sharpless asymmetric epoxidation of **3.41** with (–)-diisopropyltartrate (DiPT), titanium(IV) isopropoxide (Ti(O^{*i*}Pr)₄), cumenehydroperoxide, in the presence of 4Å molecular sieves (MS) at –20 °C under inert atmosphere for 48 hours afforded epoxy alcohol **3.42**, albeit in fairly low yields (14-48%) (Scheme 3.4).²⁵ Reported yields for this reaction average much higher (80-93%), with significantly shorter reaction times (5-12 hours).^{25, 29} Maintaining the temperature at –20 °C and avoiding moisture is imperative for this reaction, which was appropriately done during the setup. Despite our careful measures, no product was observed by TLC after the first 12 hours. The reaction mixture was placed in a –20 °C freezer overnight to ensure consistent temperature. However, no appreciable amount of product was observed the following day. It is possible that the reaction was accidentally exposed to moisture upon transfer to the freezer. It might be advantageous to explore alternative epoxidation strategies to procure better yields. For example, a Jacobsen epoxidation would provide the same enantioselectivity and can be achieved at 0 – 4 °C, allowing for better temperature control. Nevertheless, the amount of epoxide **3.42** obtained after 48 hours was adequate to proceed.

The allylic alcohol of **3.42** underwent a Swern oxidation with oxalyl chloride and dimethyl sulfoxide (DMSO) in the presence of triethylamine (Et₃N) at –78 °C for 1 hour to

generate a reactive aldehyde intermediate that was subjected to an immediate Wittig olefination with $\text{PPh}_3\text{CH}_3\text{Br}$ and potassium *tert*-butoxide (KO^tBu) for 24 hours to afford the epoxy alkene **3.43** (50%).²⁵ Finally, PMB deprotection with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) at room temperature and exposed to air, provided fragment **3.29** in 42% yield after 2 hours (Scheme 3.4).



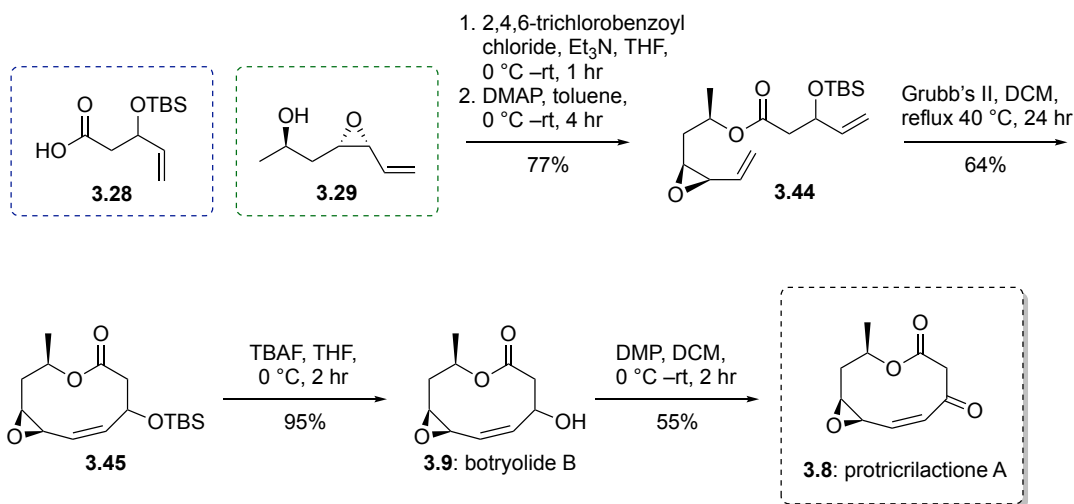
Scheme 3.4. Synthetic pathway for fragment 3.29.^{25, 29}

3.4.3 Final coupling for protilactone A

With both coupling partners in hand, a Yamaguchi esterification was performed under standard conditions. Fragment **3.28** was first treated with 2,4,6-trichlorobenzoyl chloride in the presence of Et_3N at 0 °C for 1 hour, followed by the addition fragment **3.29** and 4-dimethylaminopyridine (DMAP) in toluene (Scheme 3.5). This provided the dienolic Yamaguchi ester **3.44** in a comparable 77% yield. A RCM of **3.44** with Grubb's II catalyst at reflux in dichloromethane

(DCM) overnight generated the macrolide intermediate **3.45** in a 64% yield, consistent with the literature.²⁵ The silyl group was removed with tetra-*n*-butylammonium fluoride (TBAF) after 2 hours at 0 °C, providing botryolide B (**3.9**) in 95% yield.

Interestingly, literature reports on the total synthesis of decarestrictine I and botryolide B by Krishna et al. describe a concurrent deprotection and spontaneous ring closure via intramolecular epoxide opening, forming a dihydrofuran ring identical to the tricrilactone B–F scaffold (Figure 3.3).²⁵ It was determined that this cascade arises from spatial proximity, dictated by the stereochemistry of **3.28**, for this phenomenon was only observed in the (*S*)-enantiomer and not in the (*R*)-enantiomer. Although this is not something we explored in our studies, it gives credence to the proposed non-enzymatic biosynthetic pathway and is important to keep in mind for future investigation of tricrilactone B–F.



Scheme 3.5. Final coupling pathway to protricrilactone A.²⁵

The last step to the critical precursor protricrilactone A **3.8** was the oxidation of **3.9** with Dess-Martin periodinane (DMP) in DCM for 2 hours, resulting in a 55% yield. After repeating the synthesis on a slightly larger scale and with the help of a former postdoctoral researcher,

we ended up with a total of 60 mg of **3.8** to pursue the following dimerization conditions for tricrilactone G.

3.5 Dimerization efforts for the synthesis of tricrilactone G

All the following dimerization reactions were performed in a 2:1 molar equivalence (eq) ratio of **3.8** and the amine fragment **3.22**, respectively. The addition of either acid or base varied slightly between each reaction and will be specified below. Our initial dimerization attempt was performed in neutral conditions with solely **3.8** and **3.22** in MeOH at ambient temperature. The reaction was monitored by LC-UV-MS over the course of 48 hours; however, no reaction occurred. The amine fragment exhibited limited solubility in MeOH, presumably due to its TFA salt form. Moreover, we believe the TFA salt is precluding the necessary aza-Michael addition with **3.8**. To address this, we repeated the reaction in basic conditions using 0.5 eq Et₃N to hopefully free the amine in solution. While the addition of base improved the solubility, the mass corresponding to tricrilactone G (**3.7**) or any of the other tricrilactones were undetectable. We then substituted Et₃N with 0.5 eq 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and carried out the reaction in acetonitrile (ACN) with **3.8** and **3.22** and monitored the reaction for 5 days. None of the desired masses were detected and similar results were observed when conducted under acidic conditions using pTSA (0.1 eq).

A former postdoctoral researcher was able to synthesize fragment **3.22** as the free amine and attempted the reaction again in MeOH. After 48 hours, mass spectrometry results revealed that **3.22** underwent a Michael addition and intramolecular cyclization with itself rather than precursor **3.8**. This suggests that the TFA salt plays an important role in preventing intramolecular side reactions from occurring, but consequently obstructs the aza-Michael addition with **3.8**. Therefore, we decided to explore alternative reagents and catalysts.

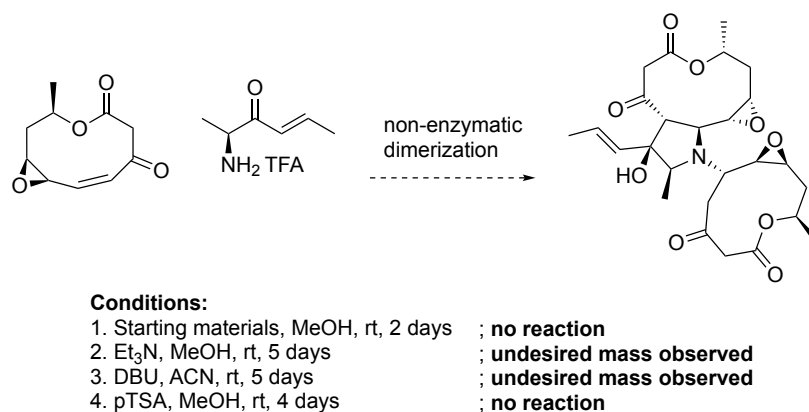
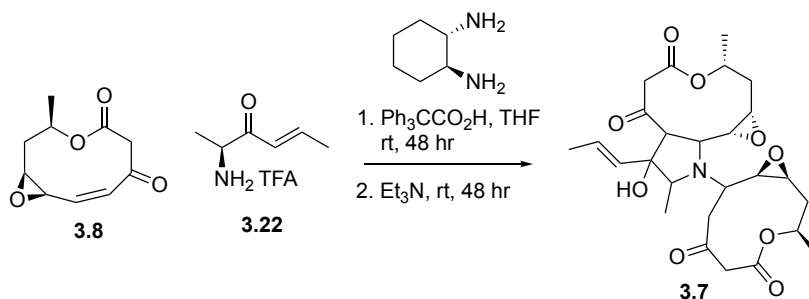


Figure 3.8. Summary of preliminary dimerization conditions.

We adopted an amine-catalyzed approach to perform an aza-Michael-aldol cascade reaction, which has been reported to provide highly functionalized chiral pyrrolizines in substantial quantities.³⁰ The success of this reaction is highly dependent on the pK_a of the starting primary amine, as it needs to be sufficiently low to be deprotonated by a moderate base in order to initiate the first aza-Michael addition. The α,β -unsaturated enone in **3.8** would then act as the electrophile for the aldol condensation cascade, resulting in the formation of the pyrrolidine core and simultaneous reduction to a secondary α -hydroxyl group depicted in Figure 3.6. The amine catalyst selected for this reaction was ($\pm$)-*trans*-1,2-diaminocyclohexane for it has shown to efficiently generate racemic pyrrolizines in respectable yields (Scheme 3.6).³⁰ At this stage, stereochemistry was not a concern as our primary objective was to optimize the reaction conditions. Triphenylacetic acid ($\text{Ph}_3\text{CCO}_2\text{H}$) was used as the acid additive as previously reported.



Scheme 3.6. Amine catalyzed dimerization pathway to tricrilactone G.³⁰

A reaction mixture comprising monomer **3.8**, **3.22**, the amine catalyst (0.25 eq), and $\text{Ph}_3\text{CCO}_2\text{H}$ (0.2 eq) in dry tetrahydrofuran (THF) was monitored for 48 hours. The UV chromatogram at 280 nm showed two prominent signals after 12 hours at retention times (RT) 2.8 and 3.2 minutes, for which the primary MS signals occurred at m/z values of 293 and 311, respectively (Figure 3.9). A signal at RT 1.5 minutes was observed after 48 hours that corresponded to m/z 205. However, we believe these to be unrelated to **3.7** or any of the predicted intermediates in the biosynthetic pathway. Under the assumption that the amine catalyst was not sufficient for deprotonating **3.8**, 0.5 eq of Et_3N was added and the reaction was monitored for an additional 48 hours. The signal observed at RT 1.5 minutes prior to the addition of Et_3N was undetectable after 2 hours and a new signal emerged at RT 1.8 minutes after 4 hours. These signals continued to interchange for the duration of the time course and unfortunately corresponded to the same undesired m/z 205.

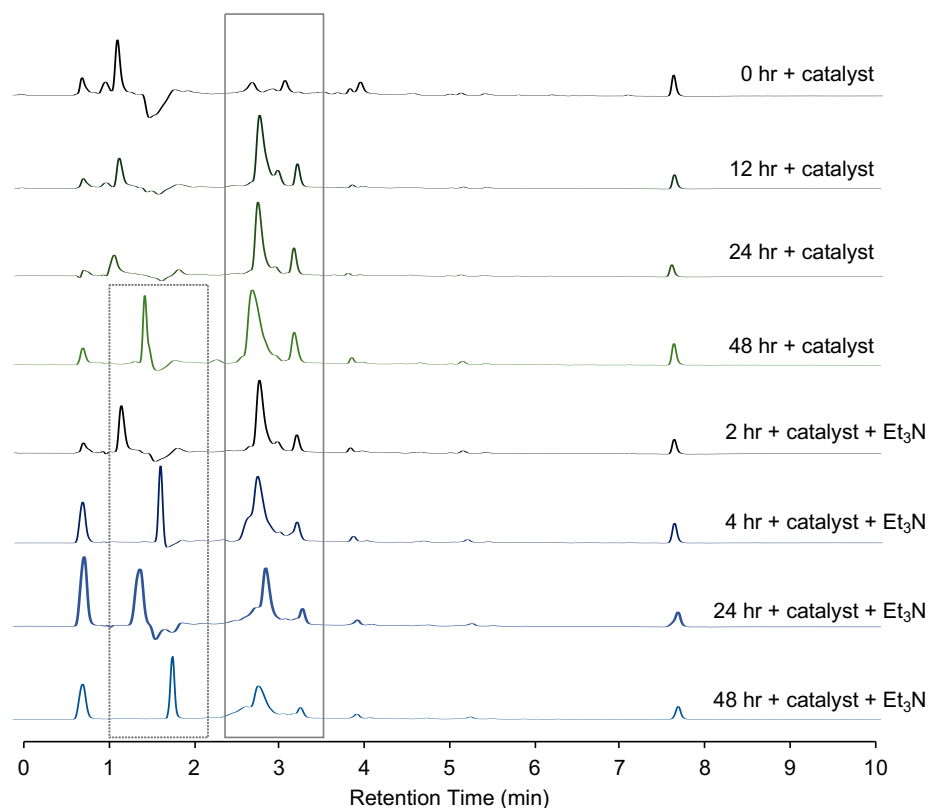
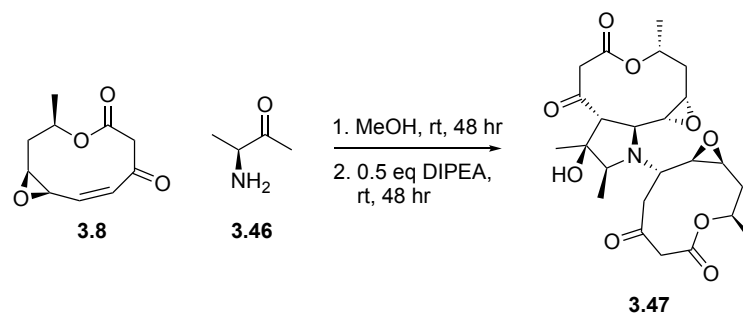


Figure 3.9. UV chromatogram of the amine-catalyzed dimerization at 280 nm. The reaction between **3.8** and **3.22** with ($\pm$)-*trans*-1,2-diaminocyclohexane as the catalyst was monitored over 2 days. Following the addition of Et₃N, the reaction was observed for an additional 2 days. Changes in signal intensities over time are outlined with boxes, although none correspond to the mass of interest.

We have considered several explanations for why **3.7** is not being formed. One possibility is that the enone moiety in **3.22** may be more reactive than that in **3.8**, leading to the unfavorable side reactions observed by LC-UV-MS. Steric interference between the α -methyl substituent of **3.22** and the epoxide of **3.8** could also hinder the dimerization cascade. Furthermore, it is plausible that conducting the reaction under basic conditions promotes the deprotonation of the substantially acidic β -keto ester protons in **3.8**, thus favoring **3.8** as the Michael acceptor over **3.22**, rather than the donor as anticipated. For these reasons, the commercially available and simplified amine fragment, (*S*)-3-aminobutan-2-one (**3.46**), was selected for subsequent attempts to synthesize compound **3.47** (Scheme 3.7).



Scheme 3.7. Final dimerization attempt with simplified amine fragment.

The new amine fragment **3.46** (0.5 eq) was added to a solution of **3.8** in MeOH at ambient temperature. Only the mass of **3.8** was detected at a RT of 1.99 minutes after 48 hours, therefore 0.5 eq of *N,N*-diisopropylethylamine (DIPEA) was added to the reaction to facilitate deprotonation and initiate the aza-Michael addition (Figure 3.10). To our excitement, *m/z* 480.2229 was observed at RT 3.46 minutes by HRESIMS 2 hours after the addition of DIPEA, which aligns with what is calculated for **3.47** [M+H]⁺ ion at *m/z* 480.2223. Another ion with a *m/z* 266.1388 was also observed, which could correspond to the product (**3.48**) of an aza-Michael-aldol condensation cascade between **3.46** with only one monomeric unit. However, a significant amount of **3.8** remained unreacted, so the reaction was left to continue. The starting material was no longer detected after 48 hours, yet the *m/z* 480 appeared to diminish and new masses emerged (Figure 3.10). The reaction was stopped by removing the solvent removed under reduced pressure and the crude material was subjected to reverse phase HPLC for purification.

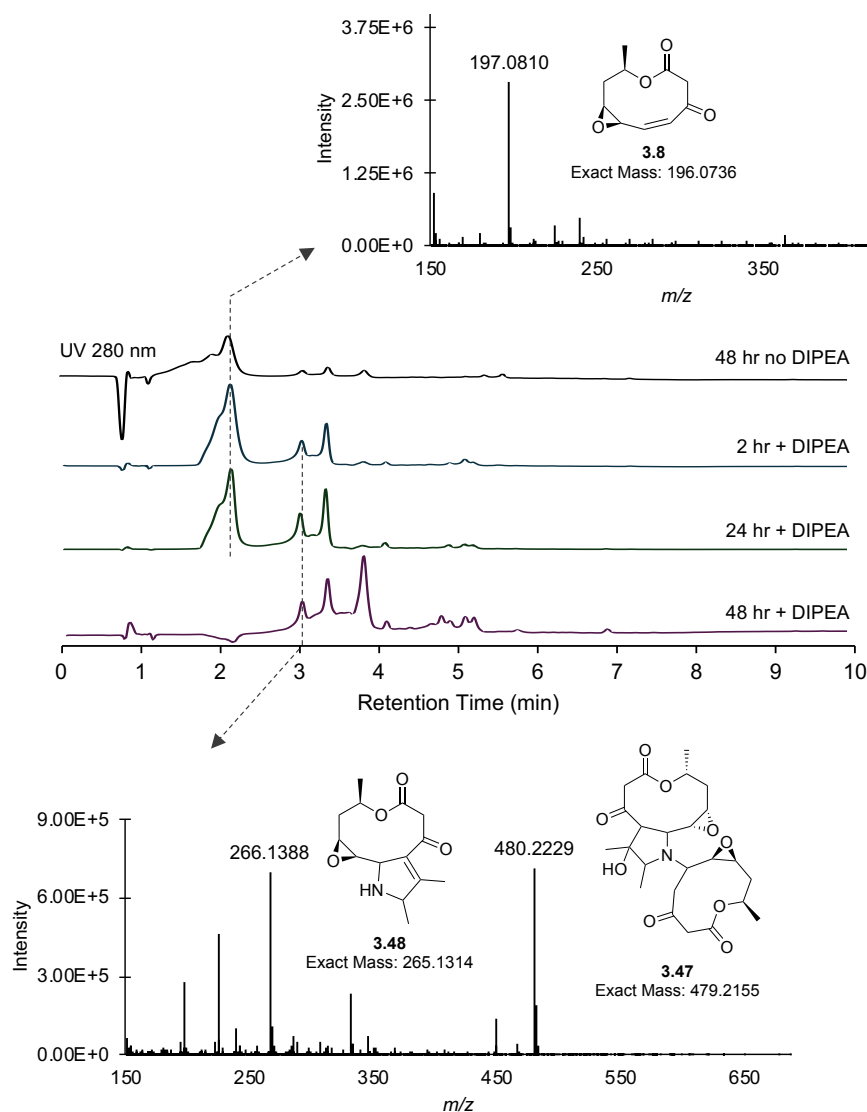


Figure 3.10. UV chromatogram of simplified amine reaction at 280 nm and corresponding mass spectra. The reaction between **3.8** and **3.46** was monitored 48 hours prior to the addition of DIPEA. The m/z indicative of **3.47** was observed after 2 hours in the presence of DIPEA. The reaction was monitored for an additional 48 hours.

HPLC purification of the crude reaction resulted in a total of 11 fractions (A–K) which were analyzed by HRESIMS, three of which were of particular interest (Figure 3.11, top). As previously mentioned, the m/z 480 was no longer the most abundant ion in any of the fractions as observed in the 2-hour time point, although it was observed in fraction D (RT 8.9 minutes)

along with m/z 512 and 448 (Figure 3.11a). While these other ions have not been accounted for, we have a strong reason to believe **3.47** was successfully made. Interestingly, fraction G, eluting at RT 10.1 minutes contained m/z 498.1737 and 476.1916, which aligns precisely to the $[M+H]^+$ and $[M+Na]^+$ ions calculated for the dimerization product maintaining the enone moiety (**3.49**) or a dimerization at the B keto ester position, although the latter does not seem likely (calcd for $C_{24}H_{30}NO_9$, 476.1915 and $C_{24}H_{29}NO_9Na$, 498.1734) (Figure 3.11b). In the same fraction m/z 405.1545 was also observed, which corresponds to the $[M+H]^+$ ion for tricrilactone D (**3.4**) (calcd for $C_{21}H_{25}O_8$, 405.1505). Lastly, the mass spectrum of fraction I (RT 11.5 minutes) revealed ions with a m/z 284.1494 and 266.1389 (**3.48**). The m/z 284.1494 is analogous to the $[M+H]^+$ ion for the aza-Michael addition intermediate (**3.50**) and/or the subsequent aldol condensation product (**3.51**) between **3.46** and a single monomeric unit of **3.9** (calcd for $C_{14}H_{22}NO_5$, 284.1498) (Figure 3.11c). Although all the observed masses discussed here have great promise, there was not enough material to obtain a complete data set for full characterization.

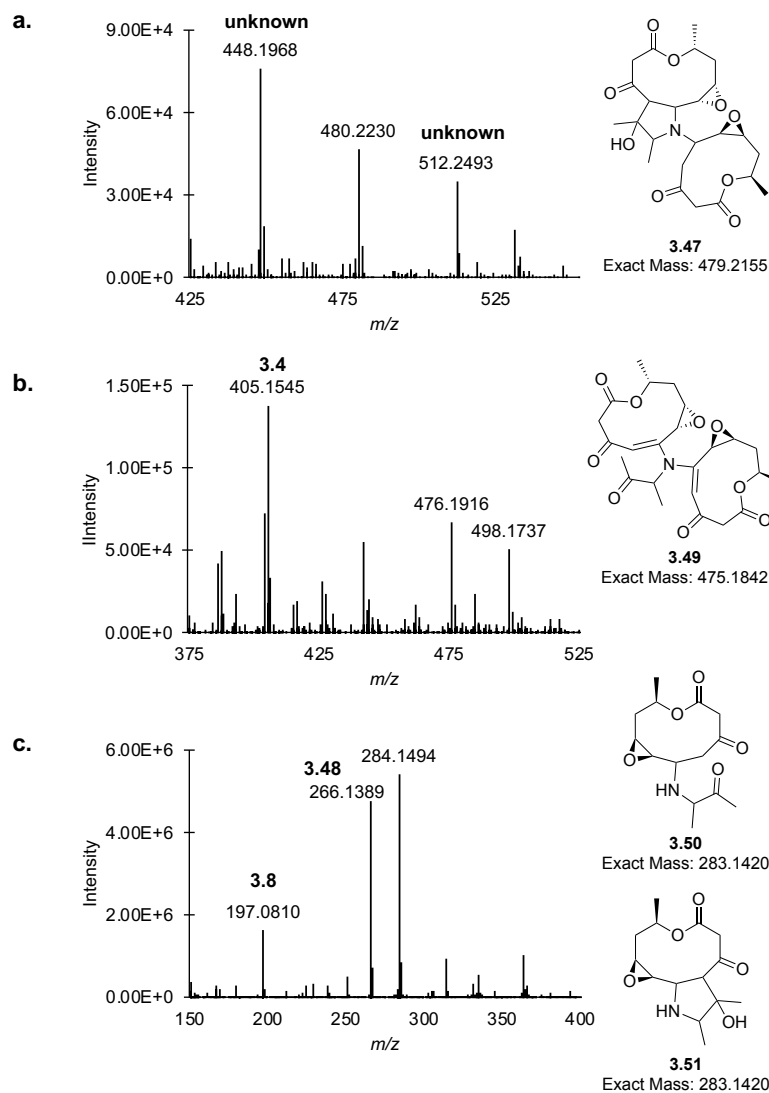
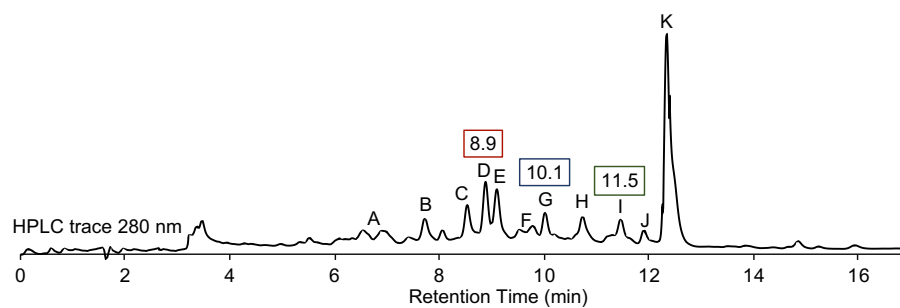


Figure 3.11 Mass spectrometry results from HPLC purification. *Top:* UV chromatogram of HPLC trace at 280 nm. *a)* Mass spectrum of fraction D (8.9 minutes) revealing the m/z $[M+H]^+$ ion of **3.47**. *b)* Mass spectrum of fraction G (10.1 minutes) corresponding to **3.4** and a potential analog of **3.47**. *c)* Mass spectrum of fraction I (11.5 minutes) containing m/z values equivalent to intermediates of **3.47**.

There are a few things to be taken into consideration for ongoing studies, one of which being the duration of the dimerization reaction in the presence of a base. Although the addition of DIPEA undoubtedly facilitated the reaction, it is possible that prolonged exposure resulted in the hydrolysis products observed in the mass spectrum. Therefore, it would be ideal to stop the reaction within the 2–12-hour window to minimize decomposition products or alternative side reactions from occurring. Moreover, improving the yield of the final dimerization reaction could presumably be achieved by changing the molar ratio of **3.8** and **3.46** from 2:1 to 1:1. This would allow for the first aza-Michael addition to occur and afford **3.50** as an intermediate. The intermediate could then be purified out, removing any side products, and added to another equivalent of **3.8**. Alternatively, another equivalent of **3.8** could be added directly to the reaction mixture upon detection of the intermediate.

3.6 Conclusion

We have refined existing methodologies for the synthesis of both fragment **3.28** and **3.29** en route to protricrilactone A (**3.8**), the critical monomeric unit for the formation of the tricrilactones.^{24-27, 29} A straightforward Yamaguchi esterification followed by RCM and oxidation afforded **3.8** allowing us to explore mechanistic insights into the non-enzymatic potential of the tricrilactone family.²⁵ Despite several attempts at aza-bridged dimerization under a variety of conditions and reagents, the mass of tricrilactone G was not observed. While this is unfortunate, it does not rule out the possibility of non-enzymatic chemistry to be occurring. It would be of interest to perform the dimerization under conditions that are more akin to the natural fermentation conditions of the producing organism.

By transitioning to a simplified ketoamine, HRESIMS data provided compelling evidence that the tricrilactone G analog **3.47** has been successfully synthesized. Additionally, we observed a mass that we believe to be the product of an aza-Michael-aldol reaction between **3.46** and a single monomeric unit **3.8**, supporting the proposed non-enzymatic cascade for the

formation of the pyrrolidine scaffold. Despite not having enough material for full structural elucidation, this synthetic pathway provides a concise route to **3.8** and NMR studies can be done to explore the mechanism of formation.

Although further work is required and dimerization conditions need to be optimized, the findings discussed in this chapter are encouraging. This work overall opens the door to many possibilities for future researchers to explore further and begin generating a small library of analogs using various keto amines or other amino acid derivatives to probe the biological activity.

3.7 Experimental procedures

General methods

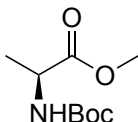
All solvents and reagents were commercially available unless otherwise noted. Chemicals were purchased from Aldrich, Fisher, Alfa Aesar, TCI, or Oakwood chemicals and used without purification. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance III HD 800 MHz and Bruker Avance III HD 500 MHz spectrometers and deuterated solvents were purchased from Cambridge Isotope Laboratories. Chemical shifts (δ) are reported in ppm relative to residual solvent chloroform-*d* (CDCl₃: δ_{H} = 7.26 ppm, δ_{C} = 77.16 ppm), dimethyl sulfoxide-*d*₆ ((CD₃)₂SO: δ_{H} = 2.50 ppm, δ_{C} = 39.52 ppm), and methanol-*d*₄ (CD₃OD: δ_{H} = 3.31 ppm, δ_{C} = 49.0 ppm) and coupling constants (*J*) in Hz. Multiplicity is denoted as s for singlet, d for doublet, t for triplet, q for quartet, p for pentent, m for multiplet, and br. s for when the signal in question is broadened. Flash chromatography was performed using E. Merck silica gel 60 (240 – 400 mesh). High resolution mass spectrometry (HRESIMS) was conducted on LTQ-Orbitrap Velos Pro MS and low resolution mass spectrometry was recorded on Agilent Technologies 6130 series Quadrupole LC-MS with electrospray ionization (ESI). HPLC purification was performed on Agilent Technologies 1200 series consisting of a UV-vis detector (G1315D), dual pumps (G1312A), and a dynamic mixer (G1312A). Reactions were performed

under an atmosphere of nitrogen with magnetic stirring unless otherwise noted. Thin layer chromatography (TLC) was performed using precoated plates purchased from E. Merk (Silica gel 60 PF₂₅₄, 0.25 mm).

Mass spectrometry analysis

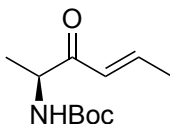
The methods for both HRESIMS and LC-UV-MS were identical. Samples were diluted with MeOH to a concentration of 1 mg/mL with an injection volume of 5–10 μ L on a Phenomenex Luna Omega C18 (100 $\times$ 2.1 mm, 1.6 μ m) LC column with a flow rate of 0.3 mL/min and a gradient solvent system of 10–100% ACN in H₂O (0.1% formic acid) over 12 minutes. MS scans were operated from either m/z 150–1500 or m/z 300–700. LC-UV-MS was monitored at UV = 280 nm.

Synthesis of amine fragment 3.22



methyl (*tert*-butoxycarbonyl)-*L*-alaninate (3.26)

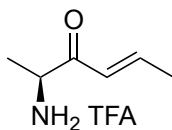
To a solution of L-alanine methyl ester (3.12 g, 22 mmol) in aq. THF (3:2, 30 mL), NaHCO₃ (5.70 g, 68 mmol) was added at room temperature. After 10 minutes, Boc₂O (5.37 g, 25 mmol) was added slowly and the reaction stirred for 24 hours at room temperature. The reaction was extracted with ether (3 X 50 mL), dried over Na₂SO₄ and concentrated under reduced pressure. The crude material was purified by column chromatography (Silica, stepwise to 10% EtOAc in hexanes) to give **3.26** (4.42 g, 98%) as a white solid. Analytical data is consistent with what has been reported.²¹ **¹H NMR (800 MHz, CDCl₃)** δ 5.04 (s, 1H), 4.31 (p, J = 7.3 Hz, 1H), 3.73 (s, 3H), 1.44 (s, 9H), 1.37 (d, J = 7.3 Hz) **¹³C NMR (200 MHz, CDCl₃)** δ 174.0, 155.2, 80.0, 52.4, 49.3, 28.4, 18.8



***tert*-butyl (*S,E*)-(3-oxohex-4-en-2-yl)carbamate (**3.27**)**

To a solution of dimethyl methylphosphonate (3.2 mL, 29.5 mmol) in dry THF (30 mL) at -78°C under N_2 , *n*-BuLi (2.5 M solution in hexanes, 12 mL, 29.5 mmol) was added dropwise and stirred. After 30 minutes, **3.26** (2 g, 9.8 mmol) in dry THF (10 mL) was added to the reaction mixture at -78°C and stirred for 2 hours. The reaction mixture was then brought to -30°C for 1 hour and was quenched with a saturated aq. solution of NH_4Cl (50 mL). The reaction mixture was extracted with EtOAc (2 X 100 mL), dried over Na_2SO_4 , and concentrated under reduced pressure to yield 2.6 g of the crude ylide reagent.

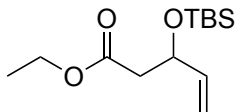
The ylide intermediate (2.6 g, 8.8 mmol) was resuspended in absolute EtOH (25 mL) and brought to 10°C . Acetyl aldehyde (0.5 mL, 8.8 mmol) was added and after 5 minutes, K_2CO_3 (1.22 g, 8.8 mmol) was added portion wise at 10°C and stirred until room temperature for 24 hours. The reaction was filtered and neutralized with acetic acid (10 mL) and the solvent was removed under reduced pressure. A saturated aq. solution of NaHCO_3 (40 mL) was added to the concentrated material which was then extracted with EtOAc (3 X 30 mL). The organic layers were combined, dried over Na_2SO_4 , concentrated under reduced pressure, and the crude material was purified by column chromatography (Silica, stepwise to 10% EtOAc in hexanes) to yield **3.27** (0.962 g, 51%) as a white solid. $^1\text{H NMR}$ (800 MHz, CDCl_3) δ 7.01 (dq, $J = 14.0, 6.9$ Hz, 1H), 6.20 (d, $J = 15.7$ Hz, 1H), 5.38 (d, $J = 7.2$, 1H), 4.59 – 4.52 (m, 1H), 1.93 (d, $J = 6.9$ Hz, 3H), 1.43 (s, 9H), 1.32 (d, $J = 7.2$ Hz, 3H) $^{13}\text{C NMR}$ (200 MHz, CDCl_3) δ 198.4, 155.3, 145.1, 128.2, 79.7, 53.3, 28.5, 18.8, 18.6



(S,E)-2-aminohept-4-en-3-one TFA salt (3.22)

To a solution of **3.27** (0.962 g, 4.5 mmol) in dry DCM (10 mL), 2,2,2-trifluoroacetic acid (2.57 g, 22.5 mmol) was added at 0 °C and the reaction stirred to room temperature for 3 hours until completion. The solvent was removed under reduced pressure to provide **3.22** (40%) as a white solid without purification. ¹H NMR (800 MHz, CDCl₃) δ 8.70 (br. s, 1H), 8.15 (s, 2H), 7.05 (dq, *J* = 14.2, 6.9 Hz, 1H), 6.18 (dd, *J* = 15.6, 2.3 Hz, 1H), 4.42 – 4.36 (m, 1H), 1.91 (d, *J* = 6.9 Hz, 3H), 1.49 (d, *J* = 7.3 Hz, 3H) ¹³C NMR (200 MHz, CDCl₃) δ 195.2, 161.7 (TFA), 148.4, 126.4, 116.2 (TFA), 53.4, 18.6, 16.1

Synthesis of fragment 3.28

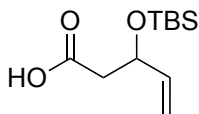


ethyl 3-((tert-butyldimethylsilyl)oxy)pent-4-enoate (3.35)

To a solution of diisopropylamine (5.5 g, 55 mmol) in dry THF (100 mL), *n*-BuLi (2.5M solution in hexanes, 22 mL, 55 mmol) was added dropwise at 0 °C under N₂ and stirred for 15 minutes. The reaction was brought to –78 °C and EtOAc (4.4 g, 50 mmol) was added dropwise and stirred for 45 minutes at 0 °C. The reaction mixture was brought to –78 °C and acrolein (2.85 g, 50 mmol) was added dropwise and stirred for 15 minutes. A saturated aq. solution of NH₄Cl (50 mL) was added slowly at –78 °C and stirred for an additional 15 minutes. The reaction mixture was extracted with ether (3 X 50 mL) and the organic layers were combined, washed with cold H₂O (2 X 100 mL) and brine (1 X 100 mL), dried over Na₂SO₄, and concentrated

under reduced pressure to provide ethyl 3-hydroxypent-4-enoate as a crude mixture (7.85 g). The crude material was moved forward to the subsequent reaction without purification.

Imidazole (9.38 g, 138 mmol) was added to a solution of the crude material from above (7.85 g, 54 mmol) in dry DMF (60 mL) at 0 °C under N₂. TBSCl (10.1 g, 67 mmol) was added portion wise and stirred at 0 °C for 10 minutes, then to room temperature for 24 hours. The reaction was quenched with cold H₂O (100 mL) and extracted with ether (1 X 200 mL). The organic layer was washed with a saturated aq. solution of NaHCO₃ (1 X 100 mL) and brine (1 X 100 mL), dried over Na₂SO₄, and concentrated under reduced pressure to provide **3.35** (10 g, 71%) as a yellow syrup. Analytical data is consistent with what has been reported.²⁷ **¹H NMR (500 MHz, CDCl₃)** δ 5.84 (ddd, *J* = 16.9, 10.4, 6.2 Hz, 1H), 5.22 (dt, *J* = 17.2, 1.4 Hz, 1H), 5.06 (dt, *J* = 10.4, 1.4 Hz, 1H), 4.58 (q, *J* = 6.3 Hz, 1H), 4.17 – 4.07 (m, 2H), 2.51 (dd, *J* = 14.5, 8.1 Hz, 1H), 2.43 (dd, *J* = 14.5, 5.2 Hz, 1H), 1.25 (t, *J* = 7.2 Hz, 3H), 0.87 (s, 9H), 0.05 (s, 3H), 0.04 (s, 3H) **¹³C NMR (200 MHz, CDCl₃)** δ 171.3, 140.5, 114.8, 71.0, 60.5, 43.9, 25.9, 18.2, 14.3, – 4.3, –5.0

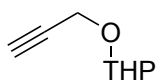


3-((*tert*-butyldimethylsilyl)oxy)pent-4-enoic acid (3.28**)**

To a solution of **3.35** (10 g, 39 mmol) in MeOH (100 mL), K₂CO₃ (10.7 g, 77 mmol) was added, and the reaction mixture was refluxed at 80 °C for 4 hours. The reaction was cooled to 0 °C and quenched with cold H₂O then extracted with ether (1 X 200 mL). The aqueous layer was acidified with 2 N HCl until pH was below 5 and extracted with EtOAc (2 X 200 mL). The organic layers were combined, washed with H₂O (1 X 100 mL), dried over Na₂SO₄, and concentrated under reduced pressure. The crude material was purified by column chromatography (Silica, stepwise to 30% EtOAc in hexanes) to provide compound **3.28** (7.83 g, 88%) as a colorless oil. Analytical data is consistent with what has been reported.²⁷ **¹H NMR (500 MHz, CDCl₃)** δ

5.85 (ddd, $J = 16.8, 10.4, 6.2$ Hz, 1H), 5.25 (d, $J = 17.1$ Hz, 1H), 5.10 (d, $J = 10.4$ Hz, 1H), 4.58 (q, $J = 6.3$ Hz, 1H), 2.55 (dd, $J = 14.8, 7.6$ Hz, 1H) 2.52 (dd, $J = 14.8, 5.3$ Hz, 1H), 0.88 (s, 9H), 0.07 (s, 3H), 0.05 (s, 3H) ^{13}C NMR (200 MHz, CDCl_3) δ 177.4, 140.0, 115.1, 70.8, 43.6, 25.9, 18.2, -4.3, -5.0

Synthesis of fragment 3.29



2-(prop-2-yn-1-yloxy)tetrahydro-2H-pyran (3.37)

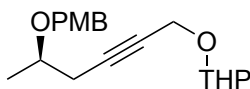
pTSA (3.01 g, 1.74 mmol) was added portion-wise to a solution of propargyl alcohol (10.1 g, 17.8 mmol) and 3,4-dihydropyran (18 g, 21.4 mmol) in dry DCM (100 mL) under N_2 at 0 °C. The reaction was brought to room temperature after 30 minutes and stirred for 2.5 hours. The reaction was quenched with H_2O and saturated aq. NaHCO_3 (1:1, 100 mL) and extracted with EtOAc (3 X 100 mL). The organic layers were combined, dried over Na_2SO_4 , concentrated under reduced pressure, and purified by column chromatography (Silica, stepwise to 5% EtOAc in hexanes) to give **3.37** (18.7 g, 75%) as a colorless syrup. Analytical data is consistent with what has been reported.²⁸ ^1H NMR (500 MHz, CD_3OD) δ 4.83 – 4.79 (m, 1H), 4.28 – 4.19 (m, 2H), 3.83 (t, $J = 9.5$ Hz, 1H), 3.55 – 3.49 (m, 1H), 2.79 (s, 1H), 1.87 – 1.78 (m, 1H), 1.76 – 1.68 (m, 1H), 1.64 – 1.49 (m, 4H)



(2R)-6-((tetrahydro-2H-pyran-2-yl)oxy)hex-4-yn-2-ol (3.38)

n-BuLi (2.5M solution in hexanes, 16.5 mL, 45.5 mmol) was added dropwise to a solution of **3.37** (5.31 g, 37.9 mmol) in dry THF (40 mL) under N_2 at -78 °C and stirred for 30 minutes then $\text{BF}_3 \cdot \text{OEt}_2$ (5 mL, 40.5 mmol) was added to the reaction mixture. After 15 minutes, a solution of

(*R*)-propylene oxide (2.31 g, 39.8 mmol) in dry THF (8 mL) was added to the reaction mixture and stirred for 4 hours. Saturated aq. solutions of NaHCO₃ (20 mL) and NH₄Cl (20 mL) were added to the reaction mixture at –78 °C and stirred to room temperature. The reaction mixture was extracted with EtOAc (2 X 100 mL), dried over Na₂SO₄, concentrated under reduced pressure, and the crude material was purified by column chromatography (Silica, stepwise to 30% EtOAc in hexanes) to give compound **3.38** (4.76 g, 70%) as a yellow syrup. Analytical data is consistent with what has been reported.²⁹ **¹H NMR (500 MHz, CDCl₃)** δ 4.72 (t, *J* = 3.5 Hz, 1H), 4.20 (dt, *J* = 15.4, 2.2 Hz, 1H) 4.12 (dt, *J* = 15.4, 2.2 Hz, 1H), 3.84 (sextet, *J* = 6.2 Hz, 1H), 3.77 – 3.71 (m, 1H), 3.46 – 3.41 (m, 1H), 2.31 – 2.27 (m, 2H), 1.77 – 1.68 (m, 1H), 1.63 (tt, *J* = 13.3, 3.5 Hz, 1H), 1.55 – 1.48 (m, 2H), 1.46 – 1.41 (m, 2H), 1.16 (d, *J* = 6.2 Hz, 3H)

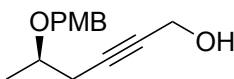


2-(((*R*)-5-((4-methoxybenzyl)oxy)hex-2-yn-1-yl)oxy)tetrahydro-2H-pyran (3.39)

PBr₃ (1.70 mL, 18.1 mmol) was added dropwise to a solution of 4-methoxybenzaldehyde (5 g, 36.2 mmol) in diethyl ether (50 mL) at 0 °C under N₂. After 15 minutes, the reaction was brought to room temperature and stirred for 45 minutes. The reaction was quenched with cold H₂O and saturated aq. NaHCO₃ (1:1, 100 mL) and extracted with diethyl ether (1 X 100 mL). The organic layer was washed with H₂O (1 X 50 mL), dried over Na₂SO₄ and concentrated under reduced pressure to provide 6.7 g of crude 4-methoxybenzyl bromide (PMB-Br) that was used in the subsequent reaction.

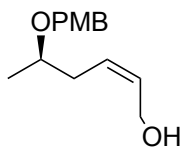
3.38 (2.77 g, 14 mmol) in dry DMF (10 mL) was added to a solution of NaH (1.4 g, 58 mmol, 60 % NaH) in dry DMF (10 mL) at 0 °C under N₂ and stirred for 25 minutes at room temperature. PMB-Br (3.41 g, 16.9 mmol) in dry DMF (10 mL) was added dropwise to the reaction mixture at 0 °C and the reaction stirred at room temperature for 4 hours. The reaction was quenched with cold H₂O (50 mL) and extracted with EtOAc (2 X 70 mL). The organic layers were combined and washed with a saturated aq. solution of NaHCO₃ (2 X 50 mL) and brine (1

X 20 mL), dried over Na₂SO₄ and concentrated under reduced pressure. The crude material was purified by column chromatography (Silica, stepwise to 10% EtOAc in hexanes) to provide **3.39** (4.44 g, 62%) as a yellow syrup. Analytical data is consistent with what has been reported.²⁹ **¹H NMR (500 MHz, CDCl₃)** δ 7.26 (d, *J* = 8.0 Hz, 2H), 6.86 (d, *J* = 8.0 Hz, 2H), 4.82 – 4.79 (m, 1H), 4.48 (s, 2H), 4.31 – 4.19 (m, 2H), 3.79 (s, 3H), 3.66 (sextet, *J* = 6.2 Hz, 1H), 3.55 – 3.48 (m, 1H), 2.58 – 2.48 (m, 1H), 2.36 (dd, *J* = 16.5, 6.8 Hz, 1H), 1.86 – 1.69 (m, 2H), 1.64 – 1.50 (m, 4H), 1.27 (d, *J* = 6.2 Hz, 3H)



(*R*)-5-((4-methoxybenzyl)oxy)hex-2-yn-1-ol (3.40)

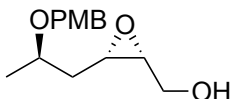
To a solution of **3.39** (2.76 g, 8.7 mmol, 1 eq) in MeOH (25 mL), pTSA (0.167 g, 0.08 mmol) was added at 0 °C under N₂ and the reaction stirred to room temperature for 2 hours. Et₃N (1 mL) was added to the reaction, stirred for 5 minutes, and concentrated under reduced pressure. The crude material was purified by column chromatography (Silica, stepwise to 40% EtOAc in hexanes) to provide **3.40** (2.03 g, 87%) as a yellow syrup. Analytical data is consistent with what has been reported.²⁹ **¹H NMR (500 MHz, CDCl₃)** δ 7.27 (d, *J* = 8.4 Hz, 2H), 6.87 (d, *J* = 8.4 Hz, 2H), 4.48 (s, 2H), 4.21 (s, 2H), 3.79 (s, 3H), 3.65 (sextet, *J* = 6.2 Hz, 1H), 2.73 (br. s, 1H), 2.55 – 2.45 (m, 1H), 2.39 (dd, *J* = 16.7, 6.2 Hz, 1H), 1.27 (d, *J* = 6.2 Hz, 3H)



(*R,Z*)-5-((4-methoxybenzyl)oxy)hex-2-en-1-ol (3.41)

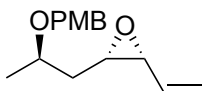
20 mL of a 1 M solution of NaBH₄ in EtOH (0.760 g, 11.3 mmol) was added nickel acetate tetrahydrate (1.87 g, 7.3 mmol) 0 °C under hydrogen atmosphere. The mixture was brought to room temperature and stirred for 30 minutes. Ethylenediamine (1.81 g, 30 mmol) was added

to the reaction mixture and stirred for 15 minutes at room temperature. **3.40** (1.76 g, 7.5 mmol) was dissolved in EtOH (2 mL) and added to the reaction mixture and stirred for 4 hours. The reaction was filtered through celite and washed with EtOAc (2 X 100 mL). The organic filtrate was concentrated under reduced pressure and filtered through celite and washed with EtOAc once more and the organic layer was concentrated again. The crude material was purified by column chromatography (Silica, stepwise to 40% EtOAc in hexanes) to provide **3.41** (1.74 g, 98%). Analytical data is consistent with what has been reported.²⁵ **¹H NMR (500 MHz, CDCl₃)** δ 7.25 (d, J = 8.5 Hz, 2H), 6.86 (d, J = 8.5 Hz, 2H), 5.58 – 5.50 (m, 2H), 4.47 (d, J = 11.3 Hz, 1H), 4.33 – 4.26 (m, 2H), 3.78 (s, 3H), 3.63 – 3.58 (m, 2H), 2.34 – 2.23 (m, 2H), 2.06 (br. s, 1H), 1.25 (d, J = 6.3 Hz, 3H)



((2R,3S)-3-((R)-2-((4-methoxybenzyl)oxy)propyl)oxiran-2-yl)methanol (3.42)

To a stirred solution of (–)-DiPT (0.30 g, 1.5 mmol) in dry DCM (15 mL) containing 4Å MS (9 g) at –20 °C under N₂ atmosphere, Ti(O^{*i*}Pr)₄ (0.21 g, 0.74 mmol) was added followed by 1.52 mL of an 80% solution of cumenehydroperoxide (1.55 g, 8.1 mmol). After 20 minutes, **3.41** (1.74 g, 7.3 mmol) in dry DCM (12 mL) was added to the reaction mixture and stirred for 2 days at –20 °C. The reaction was quenched with a 10% NaOH solution (24 mL) which stirred for 2 hours. The reaction mixture was filtered through celite and washed with EtOAc (2 X 50 mL), dried over Na₂SO₄, and purified by column chromatography (Silica, stepwise to 30% EtOAc in hexanes) to provide **3.42** (0.896 g, 48%). Analytical data is consistent with what has been reported.²⁵ **¹H NMR (500 MHz, CDCl₃)** δ 7.25 – 7.21 (m, 2H), 6.89 – 6.83 (m, 2H), 4.60 – 4.54 (m, 1H), 4.34 (dd, J = 11.3, 1.8 Hz, 1H), 3.80 – 3.74 (m, 4H), 3.73 – 3.66 (m, 1H), 3.47 (dd, J = 12.0, 7.5 Hz, 1H), 3.25 – 3.07 (m, 2H), 2.96 (quin., J = 4.4 Hz, 1H), 1.92 – 1.81 (m, 1H), 1.66 – 1.57 (m, 1H), 1.27 – 1.22 (m, 3H)

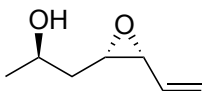


(2*S*,3*R*)-2-((*R*)-2-((4-methoxybenzyl)oxy)propyl)-3-vinyloxirane (3.43)

To a stirred solution of oxalyl chloride (0.496 g, 3.90 mmol) in dry DCM (11 mL), DMSO (0.610 g, 7.8 mmol) was added at $-78\text{ }^{\circ}\text{C}$ and stirred for 30 minutes. **3.42** (0.896, 3.5 mmol) dissolved in dry DCM (8 mL) was added to the reaction mixture at $-78\text{ }^{\circ}\text{C}$. After 1 hour, 3 mL of Et_3N (2.16 g, 21.3 mmol) was added at the same temperature and stirred for an additional 30 minutes. The reaction was quenched with 40 mL H_2O and extracted with DCM (4 X 40 mL), dried over Na_2SO_4 , and concentrated under reduced pressure to provide a yellow syrup. The crude aldehyde was used for the next step.

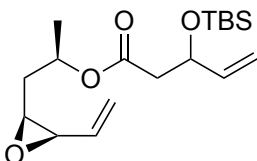
KO^tBu (1 M solution in THF, 8.88 mL, 8.87 mmol) was added dropwise to a solution of methyltriphenylphosphonium bromide (3.49 g, 12.6 mmol) in dry THF (10 mL) at $-10\text{ }^{\circ}\text{C}$ and stirred for 3 hours. The crude aldehyde (3.49 g, 13.9 mmol) in dry THF (10 mL) was added at $-10\text{ }^{\circ}\text{C}$ and the reaction stirred to room temperature for 24 hours. The reaction was quenched with a saturated aq. solution of NH_4Cl (40 mL) and extracted with EtOAc (3 X 40 mL). The organic layers were combined, dried over Na_2SO_4 , concentrated under reduced pressure, and purified by column chromatography (Silical, stepwise to 30% EtOAc in hexanes) to provide **3.43** (0.440 g, 50%) as a yellow oil. Analytical data is consistent with what has been reported.²⁵

^1H NMR (500 MHz, CDCl_3) δ 7.27 (d, $J = 7.8\text{ Hz}$, 2H), 6.87 (d, $J = 7.8\text{ Hz}$, 2H), 5.75 – 5.66 (m, 1H), 5.47 (d, $J = 17.2\text{ Hz}$, 1H), 5.35 (d, $J = 10.5\text{ Hz}$, 1H), 4.53 (d, $J = 11.2\text{ Hz}$, 1H), 4.43 (d, $J = 11.2\text{ Hz}$, 1H), 3.80 (s, 3H), 3.78 – 3.71 (m, 1H), 3.46 – 3.42 (m, 1H), 3.30 – 3.25 (m, 1H), 1.85 – 1.78 (m, 1H), 1.66 – 1.60 (m, 1H), 1.26 – 1.23 (m, 3H)



(R)-1-((2S,3R)-3-vinyloxiran-2-yl)propan-2-ol (3.29)

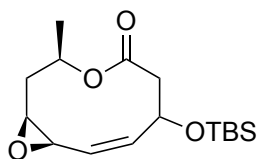
DDQ (0.445, 2 mmol) was added to a stirred solution of **3.43** (0.440 g, 1.7 mmol) in 5 mL of aq. DCM (19:1) and stirred at room temperature for 1 hour. The reaction was quenched with saturated aq. NaHCO₃ (10 mL) and extracted with EtOAc (3 X 30 mL). The organic layers were dried over Na₂SO₄, concentrated under reduced pressure, and purified by column chromatography (Silica, stepwise to 40% EtOAc in hexanes) to give **3.29** (0.095 g, 42%) of as a yellow oil. Analytical data is consistent with what has been reported.²⁵ **¹H NMR (500 MHz, CDCl₃)** δ 5.72 (ddd, J = 17.2, 10.5, 7.1 Hz, 1H), 5.48 (d, J = 17.2 Hz, 1H), 5.37 (d, J = 10.5, 1H), 4.09 – 4.02 (m, 1H), 3.46 (dd, J = 7.1, 4.4 Hz), 3.28 (dt, J = 7.4, 4.6 Hz, 1H), 1.75 – 1.70 (m, 1H), 1.68 (br. s, 1H), 1.67 – 1.61 (m, 1H), 1.27 (d, J = 6.3 Hz, 3H)



(R)-1-((2S,3R)-3-vinyloxiran-2-yl)propan-2-yl-3-((tert-butyldimethylsilyl)oxy)pent-4-enoate (3.44)

To a solution of **3.28** (0.131 g, 0.57 mmol) in dry THF (4 mL), Et₃N (0.115 g, 1.1 mmol) was added under N₂ and brought to 0 °C. 2,4,6-Trichlorobenzoyl chloride (0.193, 0.79 mmol) was then added and the reaction stirred for 1 hour at room temperature. The crude mixture was concentrated under reduced pressure and redissolved in dry toluene (2 mL). A solution of **3.29** (0.056 g, 0.44 mmol) and DMAP (0.173 g, 1.4 mmol) in dry toluene (4 mL) was added to the reaction at 0 °C and stirred to room temperature for 5 hours. The reaction was quenched with saturated aq. NH₄Cl (10 mL) and extracted with EtOAc (3 X 20 mL). The organic layers were combined, dried over Na₂SO₄ and concentrated under reduced pressure. The crude material

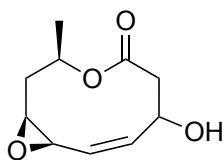
was purified by column chromatography (Silica, stepwise to 4% EtOAc in hexanes) to provide racemic **3.44** in a 1:1 ratio (0.117 g, 77%) as a colorless oil.²⁴ **¹H NMR (500 MHz, CDCl₃)** δ 5.83 (dddd, J = 16.7, 10.5, 6.2, 2.5 Hz, 1H), 5.70 (dddd, J = 17.2, 10.1, 7.0, 2.8 Hz, 1H), 5.46 (d, J = 17.2 Hz, 1H), 5.36 (d, J = 10.5 Hz, 1H), 5.21 (d, J = 17.2, 1H), 5.12 – 5.03 (m, 2H), 4.60 – 4.53 (m, 1H), 3.42 – 3.38 (m, 1H), 3.16 – 3.11 (m, 1H), 2.50 (ddd, J = 14.8, 7.6, 3.0 Hz, 1H), 2.41 (ddd, J = 14.8, 6.7, 5.7 Hz, 1H), 1.87 – 1.79 (m, 1H), 1.78 – 1.72 (m, 1H), 1.28 (dd, J = 6.3, 2.6 Hz, 3H), 0.86 (s, 9H), 0.05 (s, 3H), 0.04 (s, 3H) **¹³C NMR (200 MHz, CDCl₃)** δ 170.6, 170.5, 140.4, 132.4, 120.8, 114.9, 114.8, 70.9, 69.1, 69.0, 56.9, 55.6, 43.9, 34.4, 25.9, 20.4, 18.2, –4.3, –4.9



(1S,3R,10S,Z)-7-((*tert*-butyldimethylsilyl)oxy)-3-methyl-4,11-dioxabicyclo[8.1.0]undec-8-en-5-one (3.45)

To a solution of **3.44** (0.117 g, 0.34 mmol) in dry DCM (100 mL), Grubb's II catalyst (0.030 g, 0.04 mmol) was added under N₂ and refluxed for 24 hours at 40 °C. The reaction was stopped by bubbling in air to decompose the catalyst and the crude reaction mixture was concentrated under reduced pressure and purified by column chromatography (Silica, stepwise to 4% EtOAc in hexanes) to yield racemic **3.45** (0.069 g, 64%) as a pale brown syrup.²⁴ **¹H NMR (500 MHz, CDCl₃) Major isomer:** δ 5.68 (ddd, J = 11.6, 8.1, 1.9 Hz, 1H), 5.37 (ddd, J = 11.6, 5.5, 1.4 Hz, 1H), 5.28 (dq, J = 13.0, 6.5, 1.5 Hz, 1H), 4.78 (dddd, J = 11.6, 8.1, 3.7, 1.4 Hz, 1H), 3.45 (ddd, J = 5.7, 4.2, 1.9 Hz, 1H), 3.00 (dt, J = 10.5, 3.7 Hz, 1H), 2.55 (dd, J = 10.5, 3.7 Hz, 1H), 2.43 (t, J = 10.8 Hz, 1H), 2.20 – 2.14 (m, 1H), 1.30 – 1.28 (m, 4H), 0.89 (s, 9H), 0.13 (s, 3H), 0.11 (s, 3H), 0.10 (s, 3H), 0.10 (s, 3H) **Minor isomer:** δ 5.68 (ddd, J = 11.6, 8.1, 1.9 Hz, 1H, major + minor), 5.47 (dd, J = 12.3, 4.8, 1H), 5.28 (dq, J = 13.0, 6.5, 1.5 Hz, 1H, major + minor), 4.68

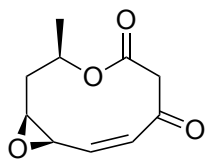
(td, $J = 6.1, 5.6, 2.1$), 4.16 – 4.13 (m, 1H),), 3.00 (dt, $J = 10.5, 3.7$ Hz, 1H, major + minor), 2.48 (dd, $J = 12.3, 5.9$, 1H), 2.41 – 2.39 (m, 1H), 2.20 – 2.14 (m, 1H, major + minor), 1.30 – 1.28 (m, 4H, major + minor) 0.13 (s, 3H), 0.11 (s, 3H), 0.92 (s, 9H) **^{13}C NMR (200 MHz, CDCl_3)** Major isomer: δ 170.6, 137.6, 126.3, 68.7, 67.7, 57.3, 53.6, 44.5, 38.2, 25.9, 20.6, 18.3, –4.5, –4.8. Minor isomer: δ 168.8, 133.1, 129.6, 69.3, 67.8, 57.2, 54.4, 42.1, 37.9, 26.0, 21.1, 18.5, –4.9, –5.0



Botryolide B ((1*S*,3*R*,10*S*,*Z*)-7-hydroxy-3-methyl-4,11-dioxabicyclo[8.1.0]undec-8-en-5-one) (3.9)

To a solution of **3.45** (0.030 g, 0.1 mmol) in dry THF (2 mL), TBAF (1 M solution in THF, 0.192 mL, 0.7 mmol) was added at 0 °C under N_2 and stirred at room temperature for 2 hours. The reaction was quenched with H_2O (1 mL) and extracted with EtOAc (2 X 2 mL). The organic layers were combined, dried over Na_2SO_4 , concentrated under reduced pressure, and purified by column chromatography (Silica, stepwise to 40% EtOAc in hexanes) to yield botryolide B, **3.9**, (0.018 g, 95%) as a colorless oil. Analytical data is consistent to what has been reported.²⁴

^1H NMR (500 MHz, CDCl_3) δ 5.74– 5.70 (m, 1H), 5.46 (dd, $J = 11.7, 5.5$ Hz, 1H), 5.29 (dq, $J = 12.7, 6.5$ Hz, 1H), 4.88 (ddd, $J = 11.7, 8.0, 3.4$ Hz, 1H), 3.51 – 3.45 (m, 1H), 3.00 (dt, $J = 10.8, 3.8$ Hz, 1H), 2.67 (dd, $J = 10.4, 3.6$ Hz, 1H), 2.44 (t, $J = 10.8$ Hz, 1H), 2.24 – 2.16 (m, 2H), 1.67 (br. s, 1H), 1.30 (d, $J = 6.5$ Hz, 3H), 1.28 – 1.23 (m, 1H) **^{13}C NMR (200 MHz, CDCl_3)** δ 170.4, 136.0, 127.7, 68.0, 67.7, 57.3, 53.6, 43.7, 38.1, 20.5



Protricrilactone A (1S,3R,10S,Z)-3-methyl-4,11-dioxabicyclo[8.1.0]undec-8-ene-5,7-dione (3.8)

To a solution of **3.9** (0.020 g, 0.1 mmol) in dry DCM (2 mL), Dess-Martin periodinane (0.090 g, 0.2 mmol) was added at 0 °C under N₂ and the reaction stirred for 2 hours at room temperature. The reaction mixture was quenched with H₂O (0.5 mL) and extracted with EtOAc (2 X 1 mL). The organic layers were combined, dried over Na₂SO₄, concentrated under reduced pressure, and purified by column chromatography (Silica, stepwise to 40% EtOAc in hexanes) to give **3.8** (0.011 g, 55%) as a colorless oil. ¹H NMR (500 MHz, CDCl₃) δ 6.22 (dt, *J* = 13.0, 1.5 Hz, 1H), 6.08 (dd, *J* = 13.0, 3.7 Hz, 1H), 5.46 (dq, *J* = 12.8, 6.5, 2.2 Hz, 1H), 3.69 – 3.63 (m, 2H), 3.21 (dt, *J* = 10.9, 3.5 Hz, 1H), 2.32 (dt, *J* = 13.9, 2.5 Hz, 1H), 1.34 (d, *J* = 6.5 Hz, 3H), 1.31 – 1.24 (m, 2H)

Table 3.1 Summary of dimerization attempts. Reaction conditions for all dimerization efforts.

Reaction number	Starting materials	Condition	Result
1	3.8 (0.010 g, 0.05 mmol) 3.22 (0.006 g, 0.025 mmol)	MeOH (0.3 mL) rt, 48 hr	No reaction
2	3.8 (0.010 g, 0.05 mmol) 3.22 (0.006 g, 0.025 mmol)	Et ₃ N (0.004 mL, 0.025 mmol) MeOH (0.5 mL) rt, 5 days	Undesired masses
3	3.8 (0.005 g, 0.025 mmol) 3.22 (0.002 g, 0.012 mmol)	DBU (0.002 g, 0.013 mmol) ACN (0.25 mL) rt, 5 days	Undesired masses
4	3.8 (0.005 g, 0.025 mmol) 3.22 (0.002 g, 0.012 mmol)	pTSA (0.001 g, 0.005 mmol) MeOH, rt, 4 days	No reaction
5	3.8 (0.010 g, 0.05 mmol) Free amine of 3.22 (0.003 g, 0.025 mmol)	MeOH (0.5 mL) rt, 48 hr	Intramolecular reactions with 3.22

6	3.8 (0.010 g, 0.05 mmol) 3.22 (0.006 g, 0.025 mmol)	1. (±)- <i>trans</i> -1,2-diaminocyclohexane (0.012 g, 0.013 mmol) triphenylacetic acid (0.003 g, 0.01 mmol) dry THF (0.5 mL) rt, 48 hr 2. Et ₃ N (0.004 mL, 0.025 mmol) rt, 48 hr	Undesired masses
7	3.8 (0.005 g, 0.025 mmol) 3.46 (0.002 g, 0.012 mmol)	1. MeOH (0.25 mL), 48 hr 2. DIPEA (0.0002 g, 0.025 mmol), 48 hr	Observed: <i>m/z</i> 480 <i>m/z</i> 498 <i>m/z</i> 476 <i>m/z</i> 404 <i>m/z</i> 284 <i>m/z</i> 266

HPLC purification of final dimerization reaction number 7

Reaction 7 was stopped by removing all solvent under reduced pressure and the crude material was resuspended in MeOH at a 10 mg/mL concentration and subject to reverse phase HPLC purification. The purification was conducted on a Phenomenex Luna C18 (150 × 4.6 mm, 5 μm) analytical column with a flow rate of 1.3 mL/min and a gradient solvent system of 10–100% 60–100% ACN in H₂O over 18 minutes (no formic acid), monitoring at UV = 280. This provided a total of 11 fractions (A–K) that were subject to HRESIMS analysis. The results are above and TIC chromatograms can be found in Appendix B.

3.8 References

1. Katz, L.; Mankin, A. S. Macrolides. In *Encyclopedia of Microbiology*, 2009; pp 529-558.
2. Lenz, K. D.; Klosterman, K. E.; Mukundan, H.; Kubicek-Sutherland, J. Z. Macrolides: From Toxins to Therapeutics. *Toxins* **2021**, *13* (5), 347.
3. Zhang, H.; Zou, J.; Yan, X.; Chen, J.; Cao, X.; Wu, J.; Liu, Y.; Wang, T. Marine-Derived Macrolides 1990–2020: An Overview of Chemical and Biological Diversity. *Marine Drugs* **2021**, *19* (4), 180.
4. Evidente, A. Fungal bioactive macrolides. *Natural Product Reports* **2022**, *39* (8), 1591-1621.
5. Du, L.; Robles, A. J.; King, J. B.; Powell, D. R.; Miller, A. N.; Mooberry, S. L.; Cichewicz, R. H. Crowdsourcing Natural Products Discovery to Access Uncharted Dimensions of Fungal Metabolite Diversity. *Angewandte Chemie International Edition* **2013**, *53* (3), 804-809.
6. Seiple, I. B.; Zhang, Z.; Jakubec, P.; Langlois-Mercier, A.; Wright, P. M.; Hog, D. T.; Yabu, K.; Allu, S. R.; Fukuzaki, T.; Carlsen, P. N.; et al. A platform for the discovery of new macrolide antibiotics. *Nature* **2016**, *533* (7603), 338-345.
7. Wang, T. T.; Wei, Y. J.; Ge, H. M.; Jiao, R. H.; Tan, R. X. Acaulide, an Osteogenic Macrodiolide from *Acaulium* sp. H-JQSF, an Isopod-Associated Fungus. *Organic Letters* **2018**, *20* (4), 1007-1010.
8. Han, W. B.; Zhai, Y. J.; Zhang, R.; Gong, X. S.; Li, J. Q.; Xu, G.; Lei, X.; Du, L.; Gao, J. M. Tricrilactones A–H, Potent Antiosteoporosis Macrolides with Distinctive Ring Skeletons from *Trichocladium crispatum*, an Alpine Moss-Associated Fungus. *Angewandte Chemie International Edition* **2023**, *62* (15), e202300773.
9. Villanueva-Martínez, A.; Hernández-Rizo, L.; Ganem-Rondero, A. Evaluating two nanocarrier systems for the transdermal delivery of sodium alendronate. *International Journal of Pharmaceutics* **2020**, *582*, 119312.
10. Fu, P.; Legako, A.; La, S.; MacMillan, J. B. Discovery, Characterization, and Analogue Synthesis of Bohemamine Dimers Generated by Non-enzymatic Biosynthesis. *Chemistry – A European Journal* **2016**, *22* (10), 3491-3495.
11. Sy, A. A.; Swenson, D. C.; Gloer, J. B.; Wicklow, D. T. Botryolides A–E, Decarestrictine Analogues from a Fungicolous *Botryotrichum* sp. (NRRL 38180). *Journal of Natural Products* **2007**, *71* (3), 415-419.
12. Riatto, V. B.; Pilli, R. A.; Victor, M. M. Fifteen years of biological and synthetic studies of decarestrictine family. *Tetrahedron* **2007**, *64* (10), 2279-2300.
13. Mayer, M.; Thiericke, R. Biosynthetic studies on the decarestrictine family. *Journal of the Chemical Society, Perkin Transactions 1* **1993**, (4), 495-500.

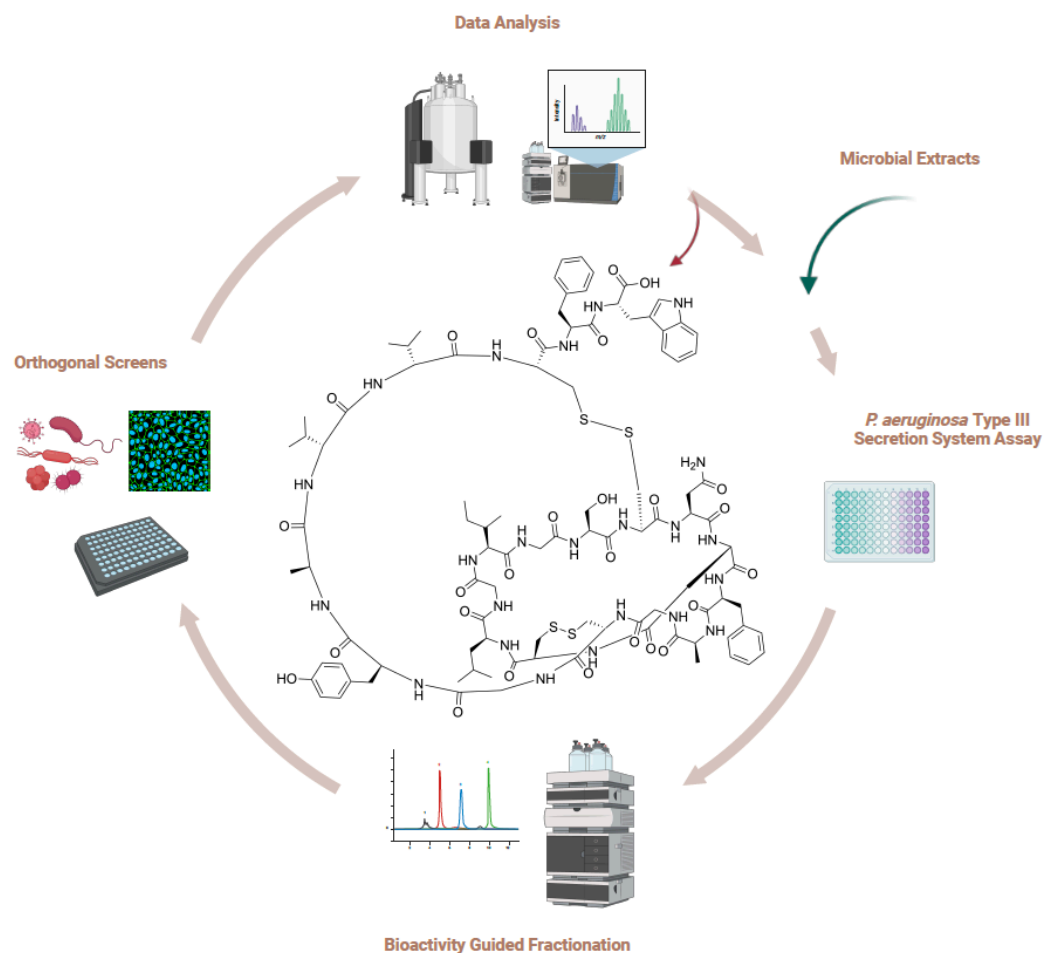
14. Mayer, M.; Thiericke, R. A non-enzymatic reaction in the late biosynthesis of the decarestrictine family. *The Journal of Antibiotics* **1993**, 46 (9), 1372-1380.
15. Chen, N. H.; Djoko, K. Y.; Veyrier, F. J.; McEwan, A. G. Formaldehyde Stress Responses in Bacterial Pathogens. *Frontiers in Microbiology* **2016**, 7, 257.
16. Colombo, L.; Gennari, C.; Ricca, G. S.; Scolastico, C.; Aragozzini, F. Detection of one symmetrical precursor during the biosynthesis of the fungal metabolite austdiol using [1,2-¹³C]acetate and [Me-¹³C]methionine. *Journal of the Chemical Society, Chemical Communications* **1981**, (11), 575-576.
17. Vartiainen, E.; Blomberg, P.; Ilmén, M.; Andberg, M.; Toivari, M.; Penttilä, M. Evaluation of synthetic formaldehyde and methanol assimilation pathways in *Yarrowia lipolytica*. *Fungal Biology and Biotechnology* **2019**, 6 (1), 27.
18. Shao, C.-L.; Wang, C.-Y.; Wei, M.-Y.; Gu, Y.-C.; She, Z.-G.; Qian, P.-Y.; Lin, Y.-C. Aspergilones A and B, two benzylazaphilones with an unprecedented carbon skeleton from the gorgonian-derived fungus *Aspergillus* sp. *Bioorganic & Medicinal Chemistry Letters* **2011**, 21 (2), 690-693.
19. Zhu, X.; Vogeler, C.; Du, L. Functional Complementation of Fumonisin Biosynthesis in FUM1-Disrupted *Fusarium verticillioides* by the AAL-Toxin Polyketide Synthase Gene ALT1 from *Alternaria alternata* f. sp. *Lycopersici*. *Journal of Natural Products* **2008**, 71 (6), 957-960.
20. Dai, G. Z.; Han, W. B.; Mei, Y. N.; Xu, K.; Jiao, R. H.; Ge, H. M.; Tan, R. X. Pyridoxal-5'-phosphate-dependent bifunctional enzyme catalyzed biosynthesis of indolizidine alkaloids in fungi. *Proceedings of the National Academy of Sciences* **2019**, 117 (2), 1174-1180.
21. Kobayashi, K.; Narumi, T.; Oishi, S.; Ohno, H.; Fujii, N. Amino Acid-Based Synthesis of Trifluoromethylalkene Dipeptide Isosteres by Alcohol-Assisted Nucleophilic Trifluoromethylation and Organozinc-Copper-Mediated SN2' Alkylation. *The Journal of Organic Chemistry* **2009**, 74 (12), 4626-4629.
22. Benedetti, F.; Miertus, S.; Norbedo, S.; Tossi, A.; Zlatoidzky, P. Versatile and Stereoselective Synthesis of Diamino Diol Dipeptide Isosteres, Core Units of Pseudopeptide HIV Protease Inhibitors. *The Journal of Organic Chemistry* **1997**, 62 (26), 9348-9353.
23. Wilson, D. J.; Shi, C.; Duckworth, B. P.; Muretta, J. M.; Manjunatha, U.; Sham, Y. Y.; Thomas, D. D.; Aldrich, C. C. A continuous fluorescence displacement assay for BioA: An enzyme involved in biotin biosynthesis. *Analytical Biochemistry* **2011**, 416 (1), 27-38.
24. Mohapatra, D. K.; Umamaheshwar, G.; Rao, M. M.; Umadevi, D.; Yadav, J. S. Concise total synthesis of botryolide B. *RSC Advances* **2014**, 4 (16), 8335-8340.
25. Radha Krishna, P.; Rao, T. J. Total synthesis of decarestrictine I and botryolide B via RCM protocol. *Organic & Biomolecular Chemistry* **2010**, 8 (14), 3130-3132.
26. Crimmins, M. T.; King, B. W.; Watson, P. S.; Guise, L. E. Synthesis and intramolecular photocycloadditions of 2-acyloxy-3-hexenoyl cyclohexenones: Diastereoselectivity in

the intramolecular [2+2] photocycloadditions of alkenes and cyclohexenones tethered by four atoms. *Tetrahedron* **1997**, 53 (26), 8963-8974.

27. Schmidt, B.; Kunz, O.; Petersen, M. H. Total Syntheses of Naturally Occurring Seimatopolide A and Its Enantiomer from Chiral Pool Starting Materials Using a Bidirectional Strategy. *The Journal of Organic Chemistry* **2012**, 77 (23), 10897-10906.
28. Shibakami, M.; Suzuki, H.; Mori, M. Synthesis of Ceramide Mimics with a Pseudo Cyclic Framework. *Synlett* **2003**, (14), 2163-2166.
29. Sharma, G. V. M.; Veera Babu, K. RCM mediated synthesis of macrosphelides I and G. *Tetrahedron: Asymmetry* **2007**, 18 (18), 2175-2184.
30. Lee, H.-J.; Cho, C.-W. Cinchona-Based Primary Amine-Catalyzed Asymmetric Cascade Aza-Michael–Aldol Reactions of Enones with 2-(1H-Pyrrol-2-yl)-2-oxoacetates: Synthesis of Chiral Pyrrolizines with Multistereocenters. *The Journal of Organic Chemistry* **2013**, 78 (7), 3306-3312.

Chapter 4

Bioactivity Guided Isolation of the Tricyclic Lasso Peptide Aborycin with Novel Type III Secretion System Inhibitory Activity Against *Pseudomonas aeruginosa*



4.1 Introduction

The antibiotic crisis, marked by the rapid emergence of antibiotic resistant bacteria, poses a significant threat to global health.^{1, 2} The overuse and misuse of antibiotic treatments have accelerated the development of resistance, rendering many traditional therapies ineffective and leading to increased mortality rates.¹⁻³ The U.S. Center for Disease Control and Prevention (CDC) and the World Health Organization (WHO) have recognized this as an urgent crisis, for

antibiotic resistance was responsible for the deaths of 1.27 million people worldwide in 2019.⁴⁻

⁶ Six multi-drug resistant bacteria, known collectively as the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) have been classified as exceptionally life-threatening.^{7, 8}

Various strategies have been employed to combat these opportunistic pathogens, including nanodelivery systems, bacteriophage therapy, adjuvant treatments, and drug repurposing.⁹⁻¹³ While these methods remain important, a promising approach that has emerged over the last decade is targeting virulence factors.^{14, 15} Virulence factors are molecules produced by pathogens that facilitate colonization of the host, evasion of the immune response or drug treatment, and damage to host tissues.^{16, 17} These factors include motility mechanisms, adhesion proteins, toxin secretion, and protective capsules against phagocytosis.^{15, 18, 19} Targeting virulence factors offers a method to disarm pathogens without compromising viability, potentially reducing the selective pressure for developing antibiotic resistance.^{11, 20, 21}

The Type III Secretion System (T3SS) is a critical virulence factor for Gram-negative pathogens, playing a vital role in the pathology of these opportunistic bacteria.²² It is largely absent in commensal populations and portions are highly conserved across Gram-negative species, making it an ideal target for virulence blockers. The T3SS is an extensively complex protein assembly that perpetuates colonization, survival, and reproduction.²³ It is comprised of a membrane-bound basal body, an external hollow needle, a translocation apparatus, effector toxins, regulatory, and chaperone proteins. The hollow needle functions as a molecular syringe, protruding from the outer membrane to transport effector proteins or toxins from the cytosol to the cellular environment, ultimately leading to infection and cell death.^{22, 24, 25}

Among the Gram-negative bacteria, the ESKAPE pathogen, *P. aeruginosa* is classified as one of the top three critical pathogens prioritized by global health officials for urgent treatment development and relies heavily on the T3SS for its virulence.^{25, 26} It is responsible for numerous hospital-related infections, including ventilator-associated pneumonia (VAP),

catheter-associated urinary tract infections, and post-operative surgical infections.²⁷ Immunocompromised individuals are especially at risk, as *P. aeruginosa* can cause acute and chronic infections in those with chronic obstructive pulmonary disease (COPD) and cystic fibrosis (CF). In 2017, the CDC reported 32,600 *P. aeruginosa* infections among hospitalized patients and 2,700 deaths in the United States.⁴ These numbers are predicted to drastically increase if new antimicrobial agents are not developed.

In collaboration with the Auerbuch Stone lab, we implemented a high-throughput screening pipeline to identify natural products with validated T3SS inhibitory activity using 6,080 extracts from our microbial library derived from marine and terrestrial bacteria. While our library has been used to discover bioactive metabolites against lung cancer, insects, and metabolic disorders, it had not been mined for T3SS inhibition. This screening process involves iterative rounds of bioactivity guided fractionation via high performance liquid chromatography (HPLC) and sequential antivirulent assays until a pure metabolite is identified using a combination of analytical tools such as mass spectrometry (MS), nuclear magnetic resonance (NMR), genome mining, and natural product databases.

Through this pipeline, we have identified aborycin, a class I lasso peptide from *Streptomyces albovinaceus* sp. (SNE-013), as an effective T3SS inhibitor with protective effects in an invertebrate model organism against *P. aeruginosa* infection (Figure 4.1). Additionally, we evaluated siamycin I, a structurally similar lasso peptide, for its inhibitory potential. Originally isolated in the 1990s from different *Streptomyces* sp., siamycin I (MS-271) and aborycin (RP71855) have been well studied for their efficacy against HIV and Gram-positive bacteria.²⁸⁻³⁰ However, no significant activity against Gram-negative bacteria has been published until now.³¹ This study presents the first evidence of aborycin and siamycin I's efficacy as T3SS inhibitors.

Herein, we present NMR spectral data with enhanced resolution and complete NMR assignments of aborycin. This chapter will discuss the detailed NMR analysis, genomic

insights, and the biological evaluation of aborycin and siamycin I successful inhibition of *P. aeruginosa* T3SS, highlighting their potential as novel antivirulence therapeutics.

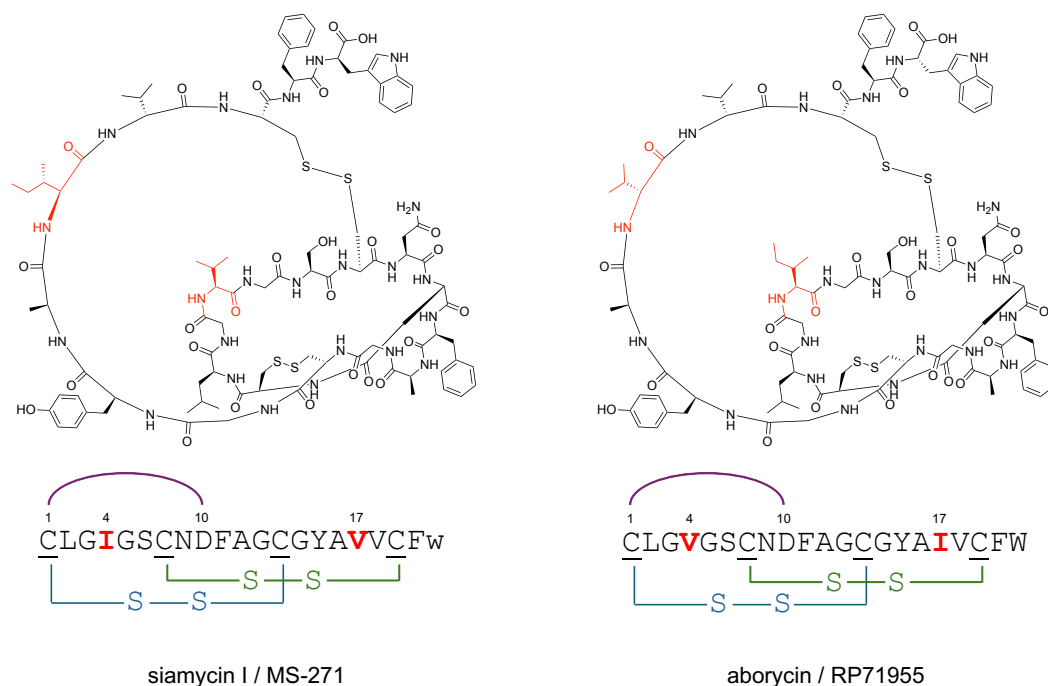


Figure 4.1. Structure and peptide sequence of siamycin I and aborycin.

4.2 Structural analysis

High-resolution electrospray ionization mass spectrometry (HRESIMS) analysis of the bioactive metabolite from SNE-013 revealed an $[M+2H]^{2+}$ ion at m/z 1082.4326, indicated by the mass difference of 0.5 Daltons (Da) observed within the isotope cluster, and the exact mass was calculated to be approximately 2161 Da (Figure 4.2b). Therefore, The Natural Product Atlas (npatlas) database and CHEMnetBASE Dictionary of Natural Products were used to search for known metabolites within the mass range of 2160–2164 Da.^{32, 33} The search provided matches to siamycin I (MS-271) and aborycin (RP 71995), both of which have a chemical formula of $C_{97}H_{131}N_{23}O_{26}S_4$ and an exact mass of 2161.8518 Da. Siamycin I and

aborycin are tricyclic ribosomally synthesized post-translationally modified peptides (RiPPs) classified as class I lasso peptides, characterized by two disulfide bonds within their scaffold.³⁴⁻³⁷ They have an identical 21-amino acid composition but differ slightly in their sequence, where Val⁴ and Ile¹⁷ in siamycin I are substituted by Ile⁴ and Val¹⁷ in aborycin (Figure 4.1). The remainder of the sequence, including the disulfide bridges between Cys¹–Cys¹³ and Cys⁷–Cys¹⁹, are identical.

The C-terminal Phe²⁰–Tryp²¹–OH fragment was observed in the MS/MS spectrum of *m/z* 1082, however, the remaining sequence was unclear due to the tricyclic nature of these compounds (Figure 4.2b and c). We performed a partial hydrolysis on the SNE-013 fraction with 2 N hydrochloric acid (HCl) at 110 °C for 12 hours to obtain individual fragments of the sequence and compare the data to what has been published for both peptides.^{31, 35, 38} Unfortunately, the precursor ions observed from this reaction were not consistent with the literature and the partial fragments obtained remained challenging to annotate. To conserve material, we did not repeat the hydrolysis but rather employed additional tools for the characterization of the bioactive metabolite.

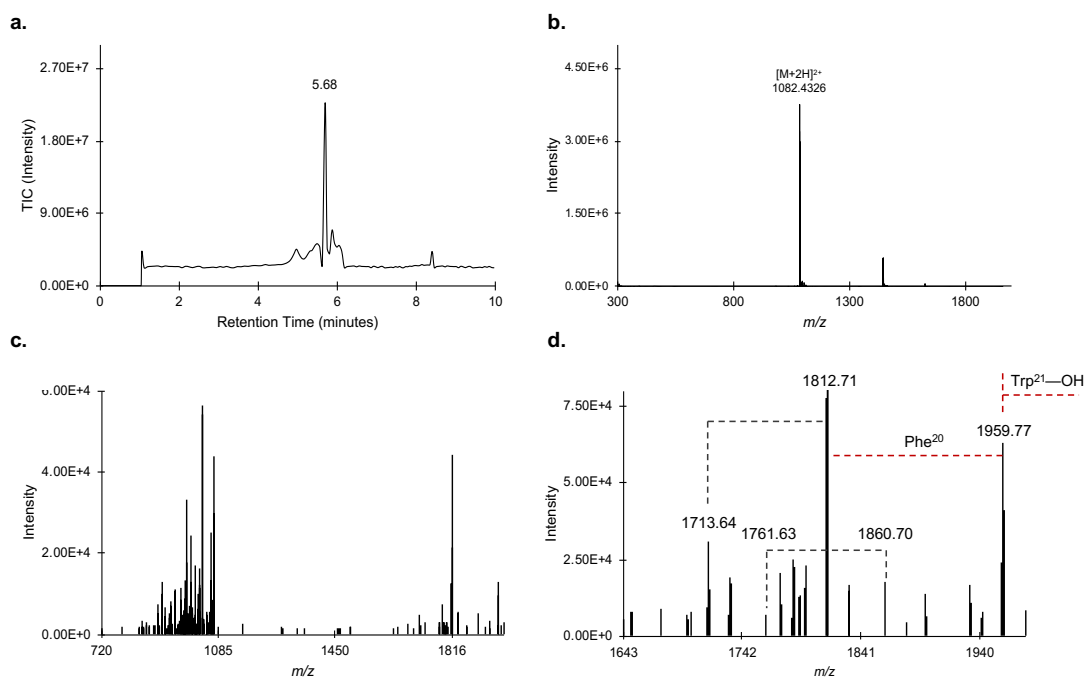


Figure 4.2. HRESIMS analysis of SNE-013 bioactive metabolite. a) TIC of purified aborycin with a retention time of 5.68 minutes. b) Mass spectrum of aborycin with a $[M+2H]^{2+}$ m/z 1082. c) Full MS/MS spectrum of m/z 1082. d) MS/MS spectrum of m/z 1082. Red lines indicate the identification of C-terminal Phe²⁰-Trp²¹-OH sequence based on the mass differences between each fragment. The spectral width did not exceed 2,000 Da, so the calculated monoisotopic m/z 2162 was used to identify Trp²¹-OH fragment. Black lines have a mass difference of 99.07 Da, suggesting valine residues but is not definitive.

We collected a comprehensive NMR dataset on the SNE-013 fraction, which included ^1H , ^{13}C , ^{15}N , ^1H - ^{13}C HSQC, long-range (8 Hz) and short-range (4 Hz) ^1H - ^{13}C HMBC, TOCSY, HSQC-TOCSY, NOSEY, and ^1H - ^{15}N HMBC experiments for structural elucidation (Table 4.1). The published spectral data for both peptides were collected in either dimethyl sulfoxide- d_6 (DMSO- d_6) or 50% methanol- d_3 (CD_3OH) in H_2O for aborycin, and either DMSO- d_6 or 50% DMSO- d_6 in H_2O for siamycin I, but these spectra are either not accessible, exhibited low resolution, or only select chemical shifts were reported.^{30, 31, 35, 36, 39} Moreover, our material exhibited limited solubility and resulted in poor resolution in these solvents (Figure 4.3), preventing direct comparison to the literature. The NMR experiments for full characterization

of the SNE-013 fraction were conducted in methanol- d_4 (CD_3OD), which significantly improved the resolution.

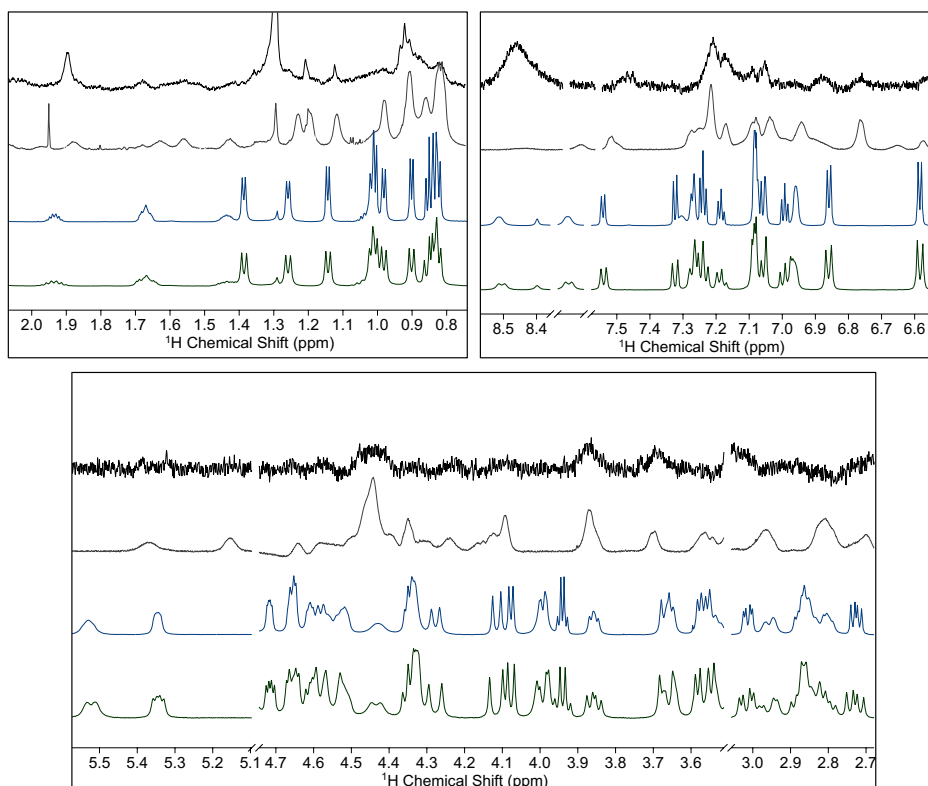


Figure 4.3. Magnified regions of the 1H NMR spectrum of aborycin in various solvents highlighting improved resolution. The spectrum in black was collected in 50% $DMSO-d_6$ in H_2O . The gray spectrum was collected in $DMSO-d_6$. The blue and green spectra were collected in CD_3OD at 20 °C and 30 °C, respectively. The residual solvent peaks were removed for visual clarity. (800 MHz)

Individual amino acid residues were assigned primarily based on TOCSY and HSQC-TOCSY correlations, as these spectra reveal all protons and carbons within a given spin system. The annotations for each amino acid are depicted in Table 4.1. The HMBC spectrum provided information on the amino acid sequence through the carbonyl region, although signal overlap prevented complete and confident annotation of the entire sequence. Nevertheless,

we were able to identify key features that differentiate aborycin from siamycin I and assemble partial fragments of the sequence.

The C-terminal Phe²⁰–Trp²¹ segment observed in the MS/MS spectrum was validated by HMBC correlations. The spectrum showed correlations from Trp-H α at δ 4.65 ppm to both δ_c 176.7 and δ_c 172.2 ppm, while the diastereotopic Trp-H β protons at δ 3.20;3.33 ppm correlated exclusively with δ_c 176.7 ppm, indicating that δ_c 176.7 ppm corresponds to the carboxylic acid of Trp²¹. Additionally, Phe-H α at δ 4.61 ppm displayed correlations to δ_c 172.2 ppm and δ_c 171.5 ppm, whereas a single cross-peak between Phe-H β at δ 3.01;2.73 ppm and δ_c 172.2 ppm was observed. Consequently, δ_c 172.2 ppm was assigned to the Phe²⁰ residue, thus confirming the Phe²⁰–Trp²¹ fragment. The signal observed at δ_c 171.5 ppm is presumed to be of Cys¹⁹, although resonances within that region were convoluted and could not be assigned with a high degree of confidence.

Another clearly identifiable partial sequence was Phe¹⁰–Ala¹¹–Gly¹². The long-range and short-range HMBC spectra revealed a series of strong correlations between δ_c 175.7 ppm and the H α (δ 3.94 ppm) and H β (δ 1.29 ppm) protons of an alanine residue, as well as the diastereotopic H α (δ 3.58;4.12 ppm) protons of a glycine residue (Figure 4.4a). Gly-H α also exhibited a significant correlation to δ_c 170.9 ppm in both spectra, while a secondary cross-peak observed between Ala-H α and a signal at δ_c 173.9 ppm was detectable only in the long-range HMBC spectrum (Figure 4.4b). The signal at δ_c 173.9 ppm was also strongly coupled to diastereotopic H β protons at δ 2.88;2.80 ppm of a second phenylalanine residue (Figure 4.4b). Based on these correlations, the resonances observed at δ_c 173.9, δ_c 175.7, and δ_c 170.9 ppm were attributed to phenylalanine, alanine, and glycine, respectively. Assuming this bioactive metabolite is either siamycin I or aborycin, this assignment is consistent with the Phe¹⁰–Ala¹¹–Gly¹² sequence of both peptides and were annotated accordingly.

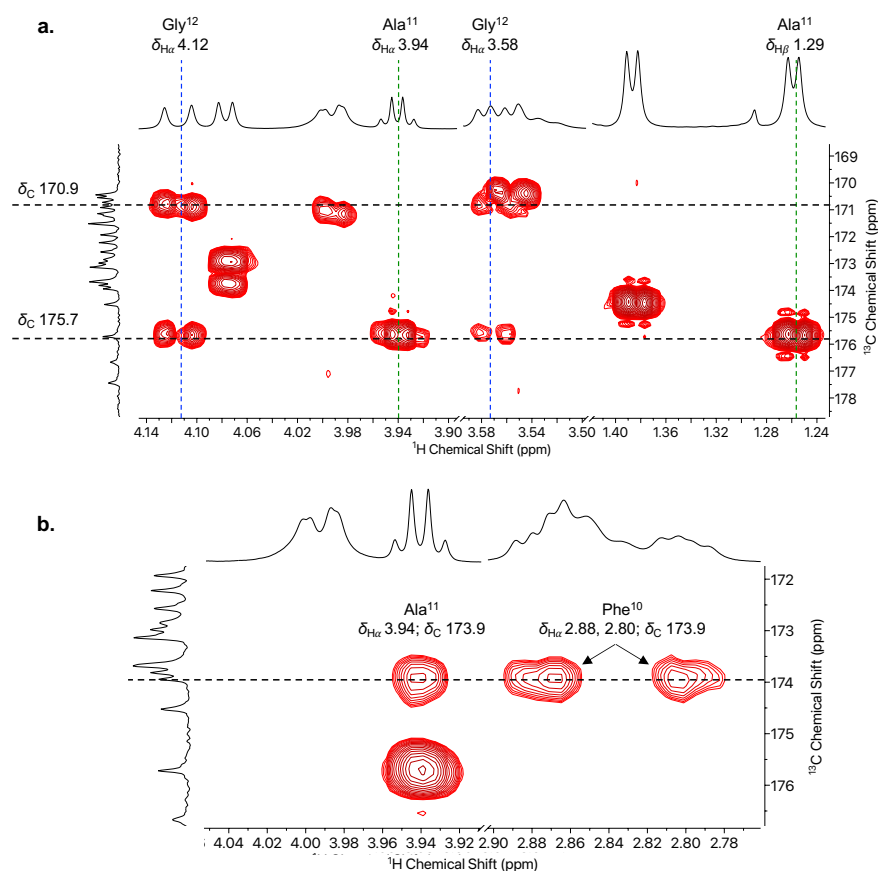


Figure 4.4. Identification of Phe¹⁰–Ala¹¹–Gly¹² sequence through HMBC correlations.

a) Short-range (4 Hz) HMBC spectrum depicting correlations from δ_C 175.7 ppm to Ala-Hα, Ala-Hβ, and Gly-Hα and correlations from δ_C 170.9 ppm to Gly-Hα, thus determining Ala-C to be δ_C 175.7 ppm and Gly-C to be δ_C 170.9 ppm. b) Long-range (8 Hz) HMBC spectrum revealing a new correlation from Ala-Hα to δ_C 173.9 ppm. Diastereotopic Phe-Hβ also correlate to δ_C 173.9, indicating δ_C 173.9 ppm corresponds to Phe-C. The observed correlations from both spectra validate the Phe¹⁰–Ala¹¹–Gly¹² partial sequence. (800 MHz, CD₃OD). *Spectra were expanded for clarity.

To distinguish whether this compound was aborycin or siamycin I, it was imperative to identify correlations between the isoleucine and valine residues and their adjacent amino acids to determine their positions as either 4 or 17 in the peptide sequence. The individual isoleucine and valine residues were readily characterized by the ¹H, COSY, and HSQC NMR data. However, substantial overlap in HMBC resonances for the isoleucine residue initially made accurate assignment challenging. Therefore, we focused on the valine correlations, which were slightly more distinct.

Analysis of one valine residue revealed cross-peaks in the long-range HMBC spectrum between Val-H α at δ 4.34 ppm and both δ_c 174.5 ppm and δ_c 173.8 ppm. The signal at δ_c 174.5 ppm also strongly coupled to resonances at δ_H 4.78 ppm and δ_H 1.39 ppm, corresponding to the H α and H β protons of an alanine residue, respectively. It is important to note that Val-H α (δ 4.34 ppm) and Ile-H α (δ 4.33 ppm) were closely overlapping in the 1H NMR spectrum, initially complicating the discernment of strong HMBC correlations. However, the short-range HMBC spectrum provided clarity by revealing a distinct correlation between Val-H α at δ 4.34 ppm and δ_c 173.8, and not to Ile-H α at δ 4.33 ppm (Figure 4.5a). Furthermore, a single cross-peak between Val-H β of the same residue at δ 2.16 ppm and δ_c 173.8 ppm was observed, indicating that the residue adjacent the alanine was valine rather than isoleucine. With Val-H β only coupling to δ_c 173.8 ppm, it was determined that δ_c 173.8 corresponds to the valine residue and δ_c 174.5 ppm to the alanine residue. This sequence aligns with the Ala¹⁶–Val¹⁷ sequence of aborycin rather than Ala¹⁶–Ile¹⁷ of siamycin I, and the residues were annotated as such.

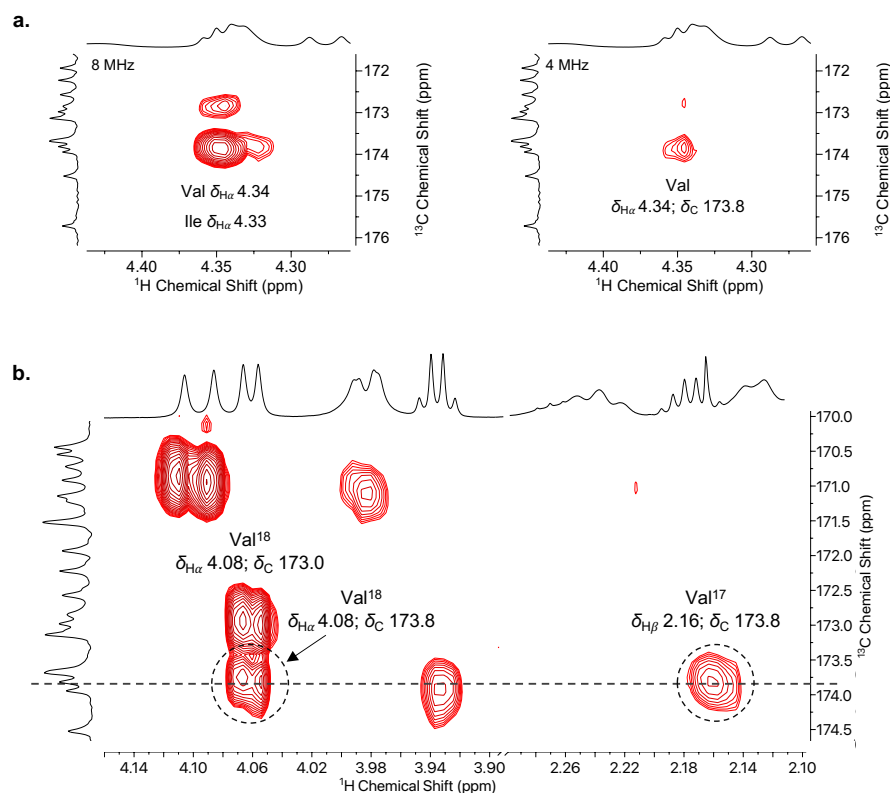


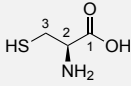
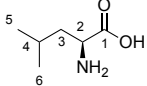
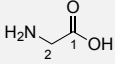
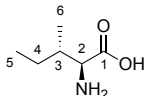
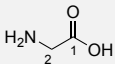
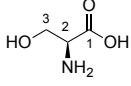
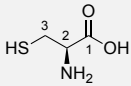
Figure 4.5. Magnified HMBC spectrum highlighting critical correlations between Val¹⁷ and Val¹⁸ residues for aborycin. a) Left: long-range (8 Hz) HMBC spectrum depicting indistinguishable correlations of Ile-H α and Val-H α to δ_C 173.8 ppm. Right: short-range (4 Hz) HMBC spectrum clarifying δ_C 173.8 ppm correlates to Val-H α and not Ile-H α . b) The pivotal correlations from δ_C 173.8 ppm to Val¹⁸-H α and Val¹⁷-H β that strongly indicate the isolation of aborycin. (800 MHz, CD₃OD).

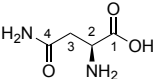
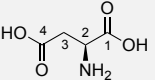
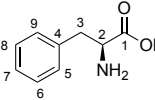
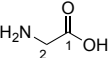
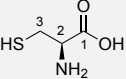
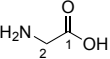
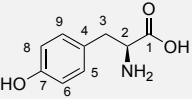
While the above correlations alone are convincing, the most notable observation was the predominant cross-peaks between the two valine residues (Figure 4.5b). The Val-H α (δ 4.08 ppm) of the second valine residue had strong HMBC correlations to δ_C 173.0 and δ_C 173.8 ppm. Since 173.8 ppm was assigned to Val¹⁷ based on the observed correlation to Val-H β , δ_C 173.0 ppm was attributed to the other valine residue and subsequently denoted Val¹⁸. The combined correlations substantiate the identification Ala¹⁶-Val¹⁷-Val¹⁸ sequence and provide a strong argument for SNE-013 to be producing aborycin.

Despite the compelling spectroscopic evidence from the NMR data, it remained possible that this compound could be another lasso peptide with the same composition but a

slightly different sequence. To further validate our findings and structural assignments, we took a genomic approach to identify the biosynthetic gene cluster (BGC) responsible for producing this lasso peptide.

Table 4.1. NMR assignments for elucidation of aborycin (800 MHz, CD₃OD).

Residue		Position	δ_C	δ_H , mult. (<i>J</i> in Hz)
Cys ¹		1	171.9	
		2 α	54.1	4.89, m
		3 β	43.9	3.44; 3.20, m
		NH		8.94, br. s
Leu ²		1	177.5	
		2 α	53.0	4.66, m
		3 β	41.9	2.23; 1.67, m
		4	25.8	2.12, m
		5	24.7	1.14, d (6.7)
		6	20.8	1.01, m
Gly ³		1	170.6	
		2 α	44.5	4.28, d (17.2); 3.67, m
		NH		8.93, br. s
Ile ⁴		1	170.7	
		2 α	57.5	4.33, m
		3 β	40.8	1.67, m
		4	25.8	1.43; 1.04, m
		5	12.0	0.84, m
		6	15.5	0.83, m
Gly ⁵		*NH		
		1	171.5	
		2 α	44.3	4.00; 3.25, m
Ser ⁶		NH		7.46, br. s
		1	171.1	
		2 α	57.7	4.72, m
		3 β	63.9	3.99; 3.86, m
Cys ⁷		NH		
		1	174.5	
		2 α	54.3	4.43
		3 β	39.5	3.53; 3.14, m
				8.35, br. s

Asn ⁸		1	173.3	
		2 α	53.0	5.35, d (8.6)
		3 β	38.5	2.46, m
		4	173.6	
		NH		8.92, br. s
Asp ⁹		1	173.6	
		2 α	53.7	4.52, m
		3 β	35.6	2.96; 2.84, m
		4	172.8	
		NH		8.06, br. s
Phe ¹⁰		1	173.9	
		2 α	56.1	4.35, t (7.0)
		3 β	40.7	2.80; 2.88, m
		4	137.9	
		5/9	131.0	7.28, d (7.5)
		6/8	129.3	7.24, t (7.5)
		7	127.8	7.19, t (7.5)
Ala ¹¹		1	175.7	
		2 α	52.7	3.94, q (7.0)
		3 β	16.2	1.26, d (7.0)
		*NH		
Gly ¹²		1	170.9	
		2 α	44.0	4.11, d (17.1); 3.57, d (17.1)
		NH		8.68, br. s
Cys ¹³		1	173.6	
		2 α	53.9	4.84, m
		3 β	43.3	3.67; 2.53, m
		NH		8.67, br. s
Gly ¹⁴		1	174.0	
		2 α	43.5	4.55; 3.56, m
		NH		7.30, br. s
Tyr ¹⁵		1	173.1	
		2 α	56.4	5.53, d (10.8)
		3 β	40.5	2.11; 2.85, m
		4	129.6	
		5/9	131.9	6.86, d (8.0)
		6/8	115.9	6.58, d (8.0)
		7	156.9	
		NH		9.46, br. s
Ala ¹⁶		1	174.5	
		2 α	49.2	4.78, t (7.5)

		3 β	21.3	1.39, d (7.0)
		NH		8.51, br. s
Val ¹⁷		1	173.8	
		2 α	61.5	4.34, m
		3 β	31.5	2.16, m
		4	20.2	1.02, m
		5	19.2	0.98, d (6.8)
		*NH		
Val ¹⁸		1	173.0	
		2 α	60.3	4.08, d (8.8)
		3 β	31.6	1.93, m
		4	19.8	0.82, d (6.7)
		5	19.0	0.90, d (6.7)
		*NH		
Cys ¹⁹		1	171.5	
		2 α	52.9	4.58, m
		3 β	43.5	2.48; 2.27, m
		*NH		
Phe ²⁰		1	172.2	
		2 α	55.7	4.61, m
		3 β	38.8	3.01, dd (14.0, 8.8); 2.73, dd (14.0, 8.8)
		4	138.2	
		5/9	130.3	6.96, d (6.5)
		6/8	129.3	7.08, m
		7	127.6	7.08, m
		NH		8.16, br. s
Trp ²¹		1	176.7	
		2 α	55.6	4.65, m
		3 β	28.7	3.20, d (7.3); 3.33, m
		4	111.2	
		5	124.6	7.05, br. s
		5-NH		10.22, br. s
		6	137.9	
		7	112.4	7.32, d (8.0)
		8	122.3	7.06, m
		9	119.9	6.99, t (8.0)
		10	119.5	7.54, d (8.0)
		11	129.3	
		*NH		

* Indicates unobserved exchangeable proton. Diastereotopic protons separated by ;

4.3 Biosynthetic gene cluster comparative analysis

To our advantage, the biosynthetic gene cluster (BGC) of aborycin was recently characterized in 2019 and is accessible in various BGC databases.³¹ Our collaborators in the Auerbuch Stone lab successfully sequenced the whole genome of *Streptomyces* sp. SNE-013 for analysis. For the purpose of this project, however, we focused solely on the lasso peptide gene cluster, while the remainder of the genome will be characterized in ongoing studies.

The DNA sequence of the lasso peptide gene cluster was uploaded to genome mining databases, antiSMASH v. 7.0 and MIBiG, for comparative analysis.^{40, 41} The sequence alignment shared 94% identity to the aborycin gene cluster from *Streptomyces* sp. ZS0098 and 82% identity with the siamycin I cluster from *Streptomyces* sp. MS-271. Each gene within this region was then evaluated using the Basic Local Alignment Search Tool (BLAST) database, yielding shared identity scores of 99% or higher for every gene (Table 4.2).⁴² Notably, an aborycin family lasso peptide gene (*aboA*) was identified, encoding a leader peptide (MTAIYEPPALQEIGDFDELTK) at its N-terminus and most importantly, the core 21-sequence (CLGIGSCNDFAGCGYAVVCFW) of aborycin at the C-terminus (Figure 4.6).^{31, 39, 42}

In addition to *aboA*, the core biosynthetic machinery was comprised of a lasso peptide isopeptide cyclase (*aboC*), precursor peptide recognition chaperone (*aboB1*), cysteine protease to cleave the precursor peptide (*aboB2*), and two thiol oxidoreductase associated genes (*aboE* and *aboF*).^{31, 40, 41} The regulatory genes consisted of transcription regulation factors (*aboR1* and *aboR2*), histidine kinase (*aboG*), ATP-binding domains (*aboD1* and *aboD3*), and ABC transport permeases (*aboD2* and *aboD4*). The other genes identified included DNA polymerase (*orf(-3)* and *orf(+3)*), nuclease (*orf(-2)*), unspecified hypothetical protein (*orf(+1)*), and an antibiotic biosynthesis regulator (*orf(+2)*) (Figure 4.6).

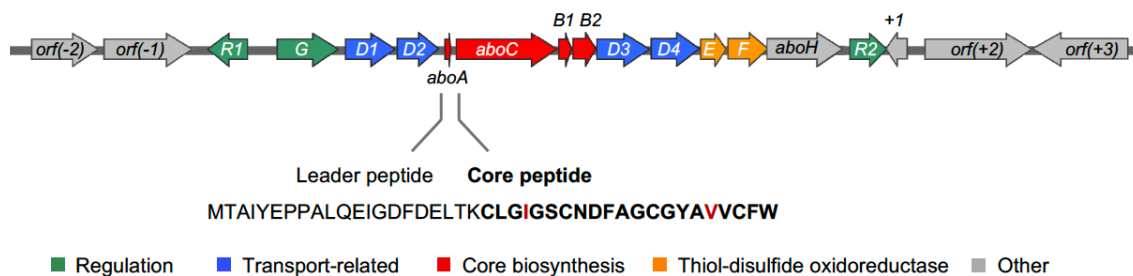


Figure 4.6. Predicted aborycin gene cluster from *Streptomyces* sp. SNE-013. Highlighted in red is where the amino acid sequence differs between aborycin and siamycin I.^{31, 39, 40}

Lastly, the BLAST results presented another interesting finding: a gene (*aboH*) in SNE-013 lasso peptide cluster shared 100% identity with a C-terminal tryptophan epimerase in the siamycin I BGC.⁴² This epimerase (*mslH*) was found to be an unprecedented metal-dependent enzyme responsible for the unusual D-Trp²¹ residue in siamycin I.^{43, 44} To the best of our knowledge, a D-configured Trp²¹ residue has not been reported for aborycin and the function of this gene as an epimerase in the aborycin BGC has not been proposed.^{28, 30, 31, 35, 36, 39, 45} While this discovery is beyond the scope of the current study, it opens the door for future experiments to determine the configuration of Trp²¹.

By integrating the genomic and spectroscopic data, we can now confidently confirm that the bioactive metabolite produced by SNE-013 is aborycin, thereby substantiating our NMR assignments. The biosynthetic genes were annotated according to the nomenclature used in the original characterization of the aborycin BGC (*aboA-F*), and a summary of their putative functions along with accession numbers can be found in Table 4.2.³¹

Table 4.2. Putative functions of genes encoded within aborycin biosynthetic gene cluster of SNE-013.^{31, 40, 42}

Size	Gene annotation	Function	% ID	Accession Number
1206 bp	<i>orf(-3)</i>	DNA polymerase III subunit delta'	99.8%	WP_156699219.1
1587 bp	<i>orf(-2)</i>	Alpha/beta hydrolase (nuclease)	100%	WP_163040000.1
702 bp	<i>aboR1</i>	Response regulation transcription factor / DNA-binding response regulator	100%	WP_048459576.1
1110 bp	<i>aboG</i>	Sensor histidine kinase	100%	WP_121720803.1
903 bp	<i>aboD1</i>	ATP-binding cassette domain-containing protein	100%	MUT89845.1
753 bp	<i>aboD2</i>	ABC transporter permease	99.6%	WP_048459497.1
129 bp	<i>aboA</i>	Aborycin family lasso peptide	100%	WP_106963308.1
1812 bp	<i>aboC</i>	Lasso peptide isopeptide bond-forming cyclase	99.8%	WP_121720801.1
264 bp	<i>aboB1</i>	Lasso peptide biosynthesis PqqD family chaperone	100%	WP_037640857.1
426 bp	<i>aboB2</i>	Lasso peptide biosynthesis B2 protein	99.3%	WP_048459500.1
966 bp	<i>aboD3</i>	ATP-binding cassette domain-containing protein	100%	WP_114873957.1
870 bp	<i>aboD4</i>	ABC transporter permease	100%	WP_121720800.1
474 bp	<i>aboE</i>	DoxX family redox-associated membrane protein	99.4%	MUT89853.1
726 bp	<i>aboF</i>	Thioredoxin domain-containing protein (disulfide bond formation)	100%	WP_163041243.1
1323 bp	<i>aboH</i>	Lasso peptide C-terminal Trp epimerase	100%	WP_048459505.1
678 bp	<i>aboR2</i>	Response regulator transcription factor	99.6%	WP_114873953.1
372 bp	<i>orf(+1)</i>	Hypothetical protein	99.2%	WP_048459507.1
1899 bp	<i>orf(+2)</i>	AfsR/SARP family transcriptional regulator	99.5%	WP_121720795.1
1692 bp	<i>orf(+3)</i>	Bifunctional 3'-5' exonuclease/DNA polymerase	99.8%	WP_156699228.1

4.4 Bioactivity assays against Type III Secretion Systems

As aforementioned, this project arose from a larger endeavor to identify bioactive molecules that inhibit *P. aeruginosa* T3SS. This is the first report to our knowledge of aborycin and

siamycin I to show activity as T3SS inhibitors. All of the following microbiology work was conducted and analyzed by our collaborators in the Auerbuch Stone lab and I will provide an overview of the assays that drove this project and their findings.

4.4.1 Primary screens for T3SS inhibition

The primary screen used to assess our 6,080 microbial extracts employs luciferase, an oxidative enzyme that induces bioluminescence, as a transcriptional reporter.^{21, 46, 47} By inserting a LUX Operon downstream of a T3SS promotor, the bacterial strain of interest will begin to luminesce when T3SS genes are being actively transcribed. Therefore, a significant decrease in luminescence when treated with microbial extracts indicates a T3SS inhibitor. Our collaborators used a *P. aeruginosa* strain that expressed an *exo-T-luxCDABE* transcriptional reporter for the screen, resulting in 1006 extracts exhibiting activity or “hits”. Of those hits, 320 were prioritized based on fold-inhibition and comparison to established inhibitors. These hits were then tested for general antibiotic activity and luciferase inhibition, for these would give rise to false positives in the primary screen.

Rather than using a bacterial strain that expresses luciferase in response to T3SS gene transcription, *P. aeruginosa* PAO1 Xen41 strain that encodes a constitutively expressed luciferase operon and little to no luminescence implies the fraction to be a luciferase inhibitor.⁴⁶ Of those 320 extracts, 207 showed luciferase inhibition and were excluded from the subsequent screens. The general antibiotic activity assay is a cell viability screen where wild-type (WT) *P. aeruginosa* is grown with alamarBlue fluorescent dye and activity is measured by the decrease in fluorescence.⁴⁸ A total of 27 extracts showed a decrease in fluorescence and were discarded, as our interest was specifically in selective T3SS inhibition. The remaining 86 hits were then categorized based on fold-inhibition from the transcriptional reporter screen and little to no activity in the alamarBlue or Xen41 assays.

The results from the primary screen, cell viability, and luciferase inhibition assays for the SNE-013 extract (SNE013-15) containing aborycin, prior to purification, are depicted in Figure 4.7. The extract showed a statistically significant decrease ($p = 0.001$) in luminescence in the primary *exo-T-lux* screen (Figure 4.7a) compared to the negative control (DMSO). The crude fraction did not display any antibiotic activity from the alamarBlue compared to the known antibiotic carbenicillin as the positive control (Figure 4.7b) nor did it inhibit luciferase (Figure 4.7c).

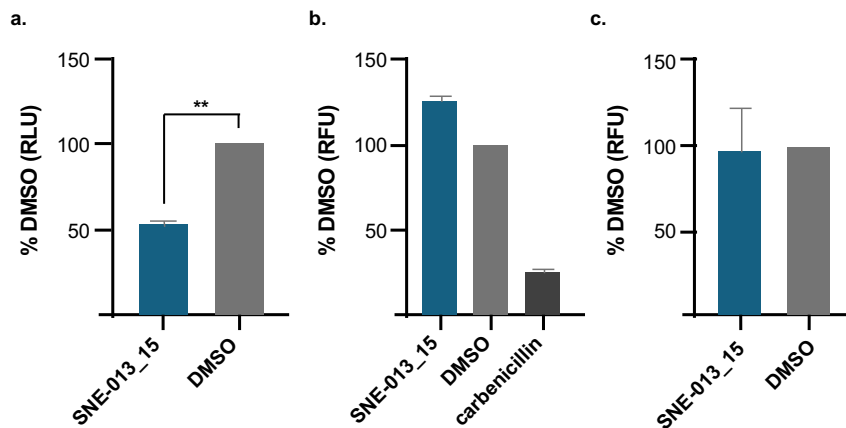


Figure 4.7. Aborycin-containing SNE-013 extract inhibits T3SS gene expression in *Pseudomonas aeruginosa* without effecting cell viability or luciferase production. a) Primary *exo-T-lux* screen of SNE-013 extract with *P. aeruginosa* shows significant decrease in T3SS expression (** $p = 0.001$; paired t test) compared to the negative DMSO control. The readout is measured in relative light units (RLU). b) *P. aeruginosa* was grown in the presence of either DMSO, SNE-013 extract, the antibiotic carbenicillin (positive control). Cell viability was quantified in relative fluorescence units (RFU), which showed SNE-013 does not have antibiotic activity. c) Luciferase inhibition assay with *P. aeruginosa* Xen41 revealed SNE-013 extract had no effect on luciferase production. Performed and analyzed by the Auerbuch Stone lab.

4.4.2 Orthogonal screen for secretion inhibition

An orthogonal inhibition assay, in which the secretion of T3SS effector proteins is measured rather than the general production of T3SS, was performed on the 69 of the 86 hits that did not show antibiotic activity.^{46, 49} Effector proteins are one of the numerous virulence factors

pathogenic bacteria, particularly one of the most important in respect to *P. aeruginosa*. The T3SS complex injects these effector proteins into the host cells, inducing serious infection while maintaining survival of the pathogen.²⁵ The effector protein, ExoU, used in this screen is fused to the small subunit of luciferase (HiBiT) to form the ExoU-HiBiT protein reporter which is then secreted by the T3SS needle. Upon secretion, the large luciferase subunit (LgBiT) and luciferin substrate are added, inducing luminescence production proportional to the amount of ExoU-HiBiT secreted.⁵⁰ Therefore, extracts that show a decrease in luminescence indicates inhibition. A total of 56 extracts inhibited T3SS activity and were prioritized into five tiers based on the combined data from all four screens.

The SNE-013 extract was of the top five prioritized samples and therefore was subject to iterative rounds of HPLC until pure aborycin was obtained. Siamycin I is commercially available and was purchased to incorporate into these bioactivity screens as well. The purified aborycin from SNE-013 and commercially available siamycin I were subjected to the orthogonal ExoU-HiBiT assay along with a known T3SS inhibitor, MBX1614. Aborycin and siamycin I exhibited a nearly identical inhibition of ExoU-HiBiT secretion (Figure 4.8a). To conserve the little amount of aborycin isolated for additional assays, this screen was repeated in triplicate with siamycin I to provide an IC₅₀ value of 7.3 μ M, consistent with other established classes of T3SS inhibitors (Figure 4.8b).^{21, 46, 50}

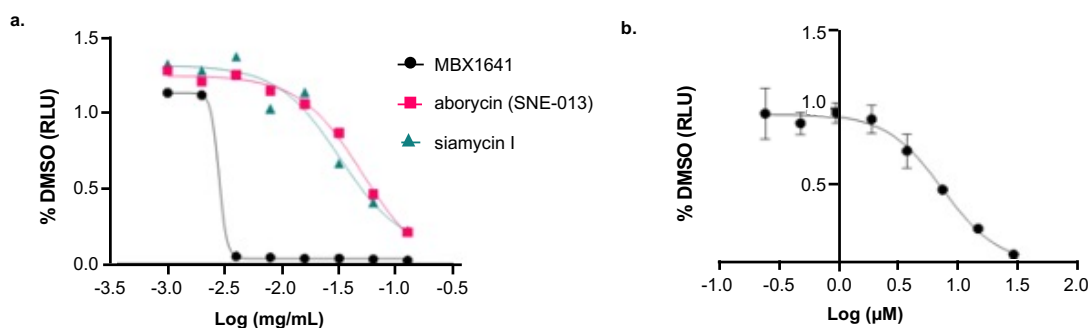


Figure 4.8. Aborycin and siamycin I inhibit secretion of T3SS effector proteins. a) *P. aeruginosa* *exoU-HiBiT* was treated with various concentrations of aborycin, siamycin I, and the known T3SS inhibitor, MBX1641, as a positive control. Both lasso peptides show inhibitory effect on ExoU-HiBiT secretion. b) The average of three independent experiments in the presence of different concentrations of siamycin I, yielding an IC_{50} value of 7.3 μ M. Performed and analyzed by the Auerbuch Stone lab.

4.4.3 *Galleria mellonella* infection model

The last bioassay that will be discussed is the live infection model with *Galleria mellonella* Greater Wax Moth larvae. Invertebrates have been shown to be excellent models for *P. aeruginosa* pathogenesis.⁵⁰⁻⁵² *G. mellonella* larvae in particular will die within the first 24 hours of infection, which is largely dependent on the *Pseudomonas* T3SS.⁵² Our collaborators injected either aborycin or the positive control T3SS inhibitor, MBX1641, into the larvae simultaneously with the pathogenic *P. aeruginosa* strain PAO1 and larval death was monitored by melanization over time (Figure 4.9). Additionally, a T3SS-deficient *P. aeruginosa* mutant, $\Delta pscF$, was used as a negative control.

To assess whether aborycin alone had any effect on survival, it was injected into healthy larvae along with phosphate buffered saline (PBS), which has no influence on viability. Larval death was not observed with aborycin and PBS alone, indicating that aborycin is not toxic to these organisms. The results showed that aborycin significantly protected the infected larvae from *P. aeruginosa*-induced death (Figure 4.9). Combining the results from all of the

bioactivity assays, the data strongly indicate that aborycin and siamycin I are effective T3SS inhibitors that can block acute *P. aeruginosa* infection.

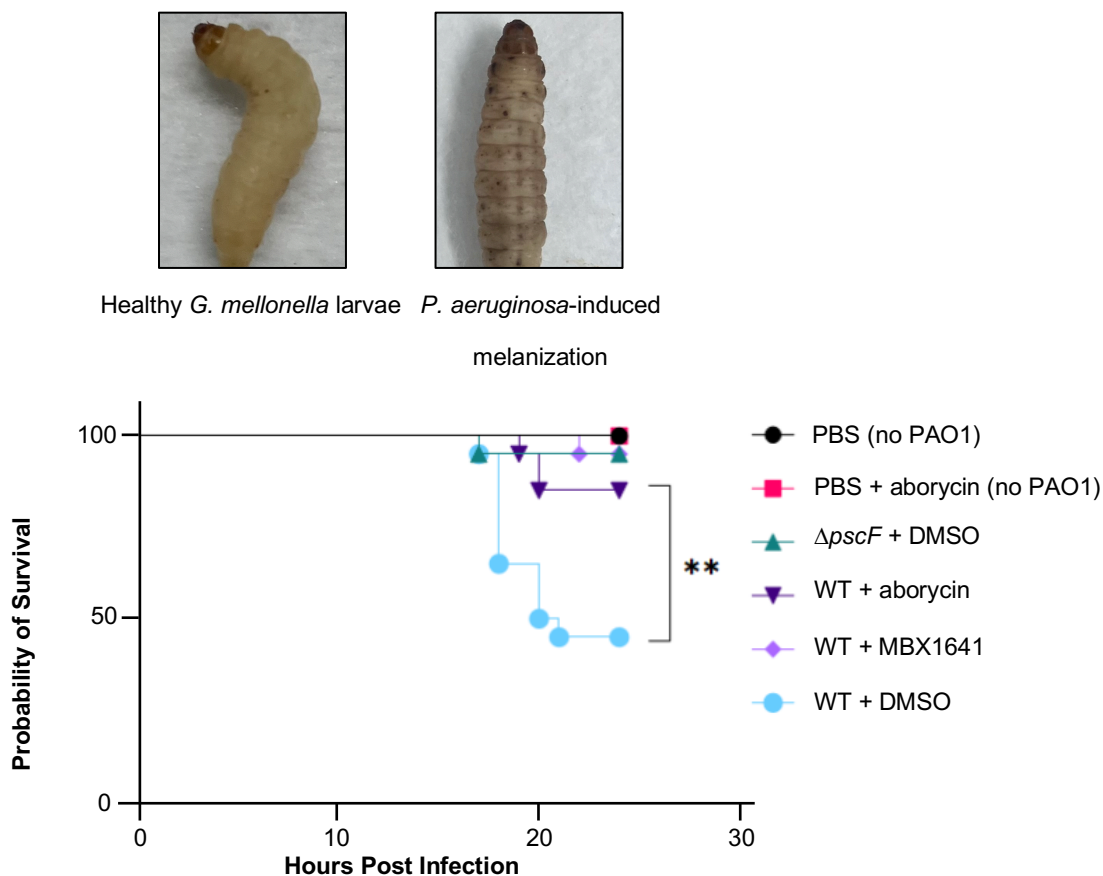


Figure 4.9. Aborycin protects *Galleria mellonella* larvae from *P. aeruginosa*-induced death. *G. mellonella* larvae were injected with DMSO, MBX1641, or aborycin and phosphate buffered saline (PBS) or WT or $\Delta pscF$ *P. aeruginosa* PAO1. The $\Delta pscF$ strain is deficient for T3SS activity and is used as a negative control. Larval melanization (death) was monitored over 24 hours. N = 20 larvae per group. **, p = 0.0064 Log-rank (Mantel-Cox test). Performed and analyzed by the Auerbuch Stone lab.

4.5 Conclusion

In summary, we successfully isolated aborycin from a *Streptomyces albovinaceus* sp. SNE-013 through bioactivity guided fractionation and evaluated its potential as a T3SS inhibitor. Using mass spectrometry and NMR data, we identified partial sequences of the bioactive lasso

peptide, suggesting aborycin over siamycin I. This was ultimately confirmed by the identification of the aborycin gene cluster predicted from whole genome sequencing. Additionally, we provided a comprehensive NMR annotation of the structure of aborycin with significant improvements in spectral resolution compared to the current literature.

As part of a larger screening project, we identified 56 bioactive extracts and created a prioritized five-tier list for future students to continue isolating T3SS inhibitors from our natural product library. To the best of our knowledge, this is the first report of aborycin and siamycin I demonstrating substantial inhibitory activity against *P. aeruginosa* T3SS, as well as effectiveness against Gram-negative pathogens overall. Notably, aborycin protected *G. mellonella* larvae from *P. aeruginosa*-induced death and enhanced survival rates. Preliminary data also suggest that both lasso peptides may be broad spectrum inhibitors, showing activity against *Yersinia* T3SS as well as *Salmonella* T3SS, one of the more divergent T3SS to *P. aeruginosa*. However, further validation of these findings is ongoing.

We have also begun exploring the mechanism of action of aborycin and identifying its target within the T3SS. We derivatized aborycin with a rhodamine cadaverine fluorescent tag, confirmed by mass spectrometry (Figure C.17, Appendix C), for use in confocal imaging microscopy. Future studies will also involve modifying the amino acid sequence through mutagenesis and heterologous expression to perform structural activity relationship (SAR) studies, as well as considering other lasso peptides. This will provide insights into the critical residues for bioactivity, informing the design of more potent derivatives.

Overall, the broader impact of this work contributes to the ongoing battle against antibiotic resistance by developing strategies that target pathogenesis mechanisms rather than viability. Our results not only validate the efficacy of these lasso peptides but also pave the way for the development of new anti-virulence therapeutics. By targeting a virulence factor rather than bacterial survival, aborycin and siamycin I emerge as promising candidates in combating the global antibiotic crisis, particularly against the notoriously difficult to treat Gram-negative

pathogens. While this project is still ongoing, there is still a rich reservoir of extracts remaining to be pursued, with some demonstrating even greater potency than SNE-013.

4.6 Experimental procedures

General methods

All solvents and reagents were commercially available unless otherwise noted. Chemicals were purchased from Aldrich, Fisher, Alfa Aesar, TCI, or Oakwood chemicals and used without purification. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance III HD 800 MHz and Bruker Avance III HD 500 MHz spectrometers and deuterated solvents were purchased from Cambridge Isotope Laboratories. Chemical shifts (δ) are reported in ppm relative to residual solvent chloroform-*d* (CDCl_3 : δ_{H} = 7.26 ppm, δ_{C} = 77.16 ppm), dimethyl sulfoxide-*d*₆ ($(\text{CD}_3)_2\text{SO}$: δ_{H} = 2.50 ppm, δ_{C} = 39.52 ppm), and methanol-*d*₄ (CD_3OD : δ_{H} = 3.31 ppm, δ_{C} = 49.0 ppm) and coupling constants (*J*) in Hz. Multiplicity is denoted as s for singlet, d for doublet, t for triplet, q for quartet, p for pentet, m for multiplet, and br. s for when the signal in question is broadened. High resolution mass spectrometry (HRESIMS) was conducted on LTQ-Orbitrap Velos Pro MS and low resolution mass spectrometry was recorded on Agilent Technologies 6130 series Quadrupole LC-MS with electrospray ionization (ESI). HPLC purification was performed on Agilent Technologies 1200 series consisting of a UV-vis detector (G1315D), dual pumps (G1312A), and a dynamic mixer (G1312A). Reactions were performed under an atmosphere of nitrogen with magnetic stirring unless otherwise noted.

General mass spectrometry analysis

Methods for both HRESIMS and LC-UV-MS were identical. Samples were diluted with MeOH to a concentration of 1 mg/mL with an injection volume of 5–10 μL on a Phenomenex Luna Omega C18 (100 $\times$ 2.1 mm, 1.6 μm) LC column with a flow rate of 0.3 mL/min and a gradient

solvent system of 10–100% ACN in H₂O (0.1% formic acid) over 12 minutes. MS scans were operated from either *m/z* 300–1500. LC-UV-MS was monitored at UV = 280 nm.

Collection and phylogenetic analysis of strain SNE-013

Strain SNE-013 was isolated from a sediment sample collected from an estuary on Kiawah Island, South Carolina (N 32° 35'10", W 80° 07'24") and isolated on a seawater agar based media (10 g starch, 4 g yeast extract, 2 g peptone, 8 g agar, 1 L of 75% sH₂O). Bacterial spores were collected via stepwise centrifugation as follows: 2 g of sediment was dried over 24 hours in an incubator at 35 °C and the resulting sediment added to 10 mL sea water (sH₂O) containing 0.05% Tween 20. After a vigorous vortex for 10 minutes, the sediment was centrifuged at 2500 rpm for 5 min (4 °C). The supernatant was removed and transferred into a new tube and centrifuged at 18,000 rpm for 25 minutes (4 °C) and the resulting spore pellet collected. The resuspended spore pellet (4 mL sH₂O) was plated on a humic acid media, giving rise to individual colonies of SNE-013 after two weeks. Genomic DNA of strain SNE-013 was isolated using standard methods and was amplified using PCR with the Universal 16S rRNA primers FC27 and RC 1492 using the method of Mincer. The partial 16S rRNA sequence (1432 of 1492 bp) was compared to sequences in available databases using the Basic Local Alignment Search Tool and revealed 99.9% identity to *Streptomyces albovinaceus*.

Cultivation and extraction of SNE-013 cultures

Bacterium SNE-013 was cultured in 10 × 2.8 L Fernbach flasks each containing 1 L of a seawater based medium (10 g starch, 4 g yeast extract, 2 g peptone, 1 g CaCO₃, 40 mg Fe₂(SO₄)₃·4H₂O, 100 mg KBr) and shaken at 200 rpm at 27 °C. After seven days of cultivation, sterilized Activated Amberlite XAD-7-HP resin (20 g/L) was added to adsorb the organic products, and the culture and resin were shaken at 200 rpm for 2 hours. The resin was filtered through cheesecloth, washed with deionized water, and eluted with acetone. The acetone soluble fraction was dried in vacuo to yield 7.0 g of extract from 10 L cultures.

Isolation and purification

The extract (7.0 g) was resuspended in MeOH and was fractionated by flash column chromatography on ODS (C18 gold, 100 g, UV = 254 nm), eluting with a step gradient of H₂O and MeOH (100:0) over 30 minutes, and 25 fractions (Fr.1–Fr.25) were collected. All fractions were subject to bioactivity testing and Fr.20 (0.7 g) was most active. Fr.20 was fractionated by flash column chromatography on ODS (C18 gold, 30 g, UV = 254 nm), eluting with a step gradient of H₂O and MeOH (100:0) over 30 minutes to provide 20 fractions (Fr.S1–Fr.S20), subject to bioactivity testing. Fr.S-10–Fr.S12 (0.045 g) were combined purified by reversed phase HPLC (Phenomenex Synergi Hydro-RP, 250 × 10 mm, 4 μm, 2 mL/min, UV = 280) using a gradient solvent system of 0–100% ACN in H₂O (0.1% formic acid) over 35 minutes to afford aborycin (0.010 g, RT = 18.2 min) as a white solid. HRESIMS [M+2H]²⁺ *m/z* 1082.4326 (calcd for C₉₇H₁₃₂N₂₃O₂₆S₄, 1082.4323)

Aborycin-fluorophore synthesis and purification

As a one pot reaction, tetramethylrhodamine-5-(and-6) cadaverine (mixed isomers, 0.0007 g, 1.39 μmol), benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) (0.0016 g, 4.17 μmol), and isolated aborycin (0.003 g, 1.39 μmol) were dissolved in 300 μL of anhydrous dimethylformamide (DMF) and the reaction was allowed to stir for 10 minutes at ambient temperature, exposed to air, and protected from light. Diisopropylethylamine (DIPEA) (0.0011 g, 8.34 μmol) was added and the reaction was allowed to stir under the same conditions as above. The reaction mixture was cooled to –20 °C then diluted with 10 mL of ether and allowed to settle for 1 hour at –20 °C. The material was subject to centrifugation at 2000 g for 15 minutes and the supernatant was decanted. The precipitated aborycin-fluorophore was purified by reverse phase HPLC (Phenomenex Luna Phenyl-Hexyl, 150 × 4.6 mm, 5 μm, 1 mL/min, UV = 280) using a gradient solvent system of 35–65% ACN in H₂O (0.1% formic acid) over 14 minutes to afford both aborycin-FL isomer 1 (0.0025 g) and isomer

2 (0.001 g) with an overall 95% yield. HRESIMS $[M+2H]^{2+}$ m/z 1331.0528 (calcd for $C_{127}H_{167}N_{27}O_{29}S_4$, 1331.0647) and $[M+3H]^{3+}$ m/z 887.7052 (calcd for $C_{127}H_{168}N_{27}O_{29}S_4$, 887.7122).

Bacterial strains and culture conditions for T3SS induction

Cultures were grown with shaking at 250 rpm unless otherwise noted. *P. aeruginosa* was grown in Luria-Bertani (LB) or 2×YT (2× yeast extract and tryptone) at 37 °C overnight, supplemented with antibiotics where noted. To induce the *P. aeruginosa* T3SS, the cultures were subcultured to an optical density at 600 nm (OD600) of 0.2 into low-calcium medium (LB with 5 mM EGTA and 20 mM $MgCl_2$). SPI-1 T3SS secretion was assessed after subculturing into fresh LB at 37 °C.

Luciferase transcriptional reporter screen

Overnight cultures of WT *P. aeruginosa* *exoT-lux* (PAO1 *exoT-lux*) and T3SS deficient mutant *P. aeruginosa* *exoT-lux* (PAO1 Δ pscC *exoT-lux*) were back diluted 1:40 in LB with 10 µg/mL gentamicin and incubated for 1.5 hours at 37 °C. White 384-well plates (Greiner Bio-One 781207) were then prepared with one-half final volume of 2X low calcium LB (2X $MgCl_2$ and EGTA) supplemented with 10 µg/mL gentamicin and natural product extracts or equivalent volume DMSO. Extracts/DMSO were added using a Janus MDT robot (PerkinElmer). After incubation, the cultures were centrifuged for 5 min at >6,000 rpm. Supernatant was removed and the cell pellets were resuspended in LBgent10. Cultures were then normalized to an OD600 of 0.1 and added to the prepared 384-well plates. Plates were sealed with Breathe-Easy® gas permeable membranes (Diversified Biotech; BEM-1) and incubated at ambient room temperature for 4 hours without shaking. Luminescence was then measured using a plate reader (PerkinElmer, EnVision 2105).

Luciferase inhibitor counterscreen

Following overnight growth, cultures of PAO1 Xen41 were centrifuged for 5 minutes at >6,000 rpm. Supernatant was removed and the cell pellets were resuspended in LB. The cultures were then normalized to an OD600 of 0.2. Normalized cultures were added to white with clear bottom 384-well plates (Corning 3763). Extracts/DMSO were then added by hand or with a Janus MDT robot (PerkinElmer) and plates were placed in a static incubator at 37 °C for 4 hours. Luminescence and absorbance at 600 nm for each plate were then measured (PerkinElmer EnVision 2105).

Bacterial viability counterscreen

Cultures of WT PAO1 were grown in LB at 37 °C overnight. The cultures were centrifuged for 5 min at 14,800 rpm. Supernatant was removed, and the pellets were resuspended in medium. Cultures were then normalized to an OD600 of 0.0005 and added to extract treated 384-well plates (Extracts/DMSO were added using either a Janus MDT robot (PerkinElmer) or an Echo acoustic liquid handler (Beckman Coulter). Plates were incubated in a plate reader (PerkinElmer Envision 2105) at room temperature, and fluorescence was measured every hour for 12 hours.

ExoU-HiBiT assay

A plasmid carrying tetracycline resistance and an inactive ExoU-HiBiT reporter under control of the *exoU* promoter (pME6032 *exoU* D344A Rep) was provided by Erwin Bohn at the Tübingen University Hospital in Germany. The plasmid was cloned using *E. coli* DH5α competent cells. Isolated plasmid was then electroporated into WT *P. aeruginosa* PA103 and T3SS deficient mutant PA103 ΔpscJ to create the ExoU-HiBiT secretion reporter strains. Both reporter strains were grown overnight in LB with 50 µg/mL tetracycline. Cultures were then diluted to an OD600 of 0.2 with 50 µg/mL tetracycline and plated in black untreated 384-well plates with extract/DMSO (Corning 3573). Extracts/DMSO were added using either a Janus

MDT robot (PerkinElmer) or an Echo Acoustic liquid handler (Beckman Coulter). Plates were sealed with Breathe-Easy® gas permeable membranes (Diversified Biotech; BEM-1) and incubated at ambient room temperature for 4 hours. A 1:4 dilution of NanoGlo® HiBiT Extracellular Detection System reagent (Promega; N2421) in 1X PBS was then added to plates at 50% v/v. Plates were then resealed and shaken at 26°C for 5 minutes. Luminescence was then measured (PerkinElmer Envision 2105).

***Galleria mellonella* infection assay**

G. mellonella larvae (Carolina Biological Supply 143928) at the last instar developmental stage were selected and starved 24 hours prior to infection. 20 larvae were selected per treatment group. Cultures of either WT or Δ pscF *P. aeruginosa* (PAO1) were grown in LB at 37 °C overnight. Overnight cultures were then subcultured to an OD600 of 0.2 in LB, then serially diluted to appropriately 100 CFU/mL in 1xPBS. MBX1641, SNE0-13_Fr.S10-Fr.S12., or an equivalent volume of DMSO were added to the diluted PAO1 preparations. The *G. mellonella* larvae were injected with 10 μ l in the left pro-leg using an autopump (KDSscientific 780100) and a 26G needle. Larval melanization (death) was monitored over the following 24 hours.

DNA sequencing and assembly

DNA samples from SNE0-13 were prepared using a Wizard Genomic DNA Purification kit (Promega A1120) according to the provided protocol for Gram-positive bacteria. Bacterial Genome Sequencing was performed by Plasmidsaurus using Oxford Nanopore Technology and Illumina sequencing with custom analysis and annotation.

Statistical analysis

Prism 10 (GraphPad Software, La Jolla, CA, USA) was used to calculate the mean, standard deviation, standard error of the mean, median, standard error of median, T-test values, and Log-rank (Mantel-Cox test) values shown.

4.7 References

1. Lessa, F. C.; Sievert, D. M. Antibiotic Resistance: A Global Problem and the Need to Do More. *Clinical Infectious Diseases* **2023**, 77 (Supplement_1), S1-S3.
2. Church, N. A.; McKillip, J. L. Antibiotic resistance crisis: challenges and imperatives. *Biologia* **2021**, 76 (5), 1535-1550.
3. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T* **2015**, 40 (4), 277-283.
4. CDC. *Antibiotic Resistance Threats in The United States, 2019*; U.S. Department of Health and Human Services, CDC, Atlanta, GA, 2019. DOI: 10.15620/cdc:82532.
5. Murray, C. J. L.; Ikuta, K. S.; Sharara, F.; Swetschinski, L.; Robles Aguilar, G.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **2022**, 399 (10325), 629-655.
6. Organization, G. W. H. Global antimicrobial resistance and use surveillance system (GLASS) report. Organization, G. W. H., Ed.; 2022; Vol. CC BY-NC-SA 3.0 IGO.
7. Boucher, Helen W.; Talbot, George H.; Bradley, John S.; Edwards, John E.; Gilbert, D.; Rice, Louis B.; Scheld, M.; Spellberg, B.; Bartlett, J. Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **2009**, 48 (1), 1-12.
8. Miller, W. R.; Arias, C. A. ESKAPE pathogens: antimicrobial resistance, epidemiology, clinical impact and therapeutics. *Nature Reviews Microbiology* **2024**, Advance online publication.
9. Kamurai, B.; Mombeshora, M.; Mukanganyama, S.; Khamesipour, F. Repurposing of Drugs for Antibacterial Activities on Selected ESKAPE Bacteria *Staphylococcus aureus* and *Pseudomonas aeruginosa*. *International Journal of Microbiology* **2020**, 2020, 1-9.
10. Ma, Y. X.; Wang, C. Y.; Li, Y. Y.; Li, J.; Wan, Q. Q.; Chen, J. H.; Tay, F. R.; Niu, L. N. Considerations and Caveats in Combating ESKAPE Pathogens against Nosocomial Infections. *Advanced Science* **2019**, 7 (1), 1-43.
11. Pacios, O.; Blasco, L.; Bleriot, I.; Fernandez-Garcia, L.; González Bardanca, M.; Ambroa, A.; López, M.; Bou, G.; Tomás, M. Strategies to Combat Multidrug-Resistant and Persistent Infectious Diseases. *Antibiotics* **2020**, 9 (2), 65.
12. Sarker, S. A.; Sultana, S.; Reuteler, G.; Moine, D.; Descombes, P.; Charton, F.; Bourdin, G.; McCallin, S.; Ngom-Bru, C.; Neville, T.; et al. Oral Phage Therapy of Acute Bacterial Diarrhea With Two Coliphage Preparations: A Randomized Trial in Children From Bangladesh. *EBioMedicine* **2016**, 4, 124-137.

13. Steindl, T. M.; Schuster, D.; Laggner, C.; Chuang, K.; Hoffmann, R. D.; Langer, T. Parallel Screening and Activity Profiling with HIV Protease Inhibitor Pharmacophore Models. *Journal of Chemical Information and Modeling* **2007**, 47 (2), 563-571.
14. Dehbanipour, R.; Ghalavand, Z. Anti-virulence therapeutic strategies against bacterial infections: recent advances. *Germs* **2022**, 12 (2), 262-275.
15. Rasko, D. A.; Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nature Reviews Drug Discovery* **2010**, 9 (2), 117-128.
16. Wu, S.-C.; Liu, F.; Zhu, K.; Shen, J.-Z. Natural Products That Target Virulence Factors in Antibiotic-Resistant *Staphylococcus aureus*. *Journal of Agricultural and Food Chemistry* **2019**, 67 (48), 13195-13211.
17. Sharma, A. K.; Dhasmana, N.; Dubey, N.; Kumar, N.; Gangwal, A.; Gupta, M.; Singh, Y. Bacterial Virulence Factors: Secreted for Survival. *Indian Journal of Microbiology* **2016**, 57 (1), 1-10.
18. Miller, V.; Defoirdt, T. Antivirulence Therapy for Animal Production: Filling an Arsenal with Novel Weapons for Sustainable Disease Control. *PLoS Pathogens* **2013**, 9 (10), e1003603.
19. Morales, G.; Abelson, B.; Reasoner, S.; Miller, J.; Earl, A. M.; Hadjifrangiskou, M.; Schmitz, J.; Tyne, D. V. The Role of Mobile Genetic Elements in Virulence Factor Carriage from Symptomatic and Asymptomatic Cases of *Escherichia coli* Bacteriuria. *Microbiology Spectrum* **2023**, 11 (3), e0471022.
20. Calvert, M. B.; Jumde, V. R.; Titz, A. Pathoblockers or antivirulence drugs as a new option for the treatment of bacterial infections. *Beilstein Journal of Organic Chemistry* **2018**, 14, 2607-2617.
21. Lam, H. N.; Lau, T.; Lentz, A.; Sherry, J.; Cabrera-Cortez, A.; Hug, K.; Lalljie, A.; Engel, J.; Lokey, R. S.; Auerbuch, V. Developing Cyclic Peptomers as Broad-Spectrum Type III Secretion System Inhibitors in Gram-Negative Bacteria. *Antimicrobial Agents and Chemotherapy* **2021**, 65 (7), e00060-00017.
22. Deng, W.; Marshall, N. C.; Rowland, J. L.; McCoy, J. M.; Worrall, L. J.; Santos, A. S.; Strynadka, N. C. J.; Finlay, B. B. Assembly, structure, function and regulation of type III secretion systems. *Nature Reviews Microbiology* **2017**, 15 (6), 323-337.
23. Notti, R. Q.; Stebbins, C. E.; Kudva, I. T. The Structure and Function of Type III Secretion Systems. *Microbiology Spectrum* **2016**, 4 (1), 10.1128/microbiolspec.VMBF-0004-2015.
24. Ashida, H.; Kim, M.; Sasakawa, C. Manipulation of the host cell death pathway by *Shigella*. *Cellular Microbiology* **2014**, 16 (12), 1757-1766.
25. Hauser, A. R.; Ruiz, J. The type III secretion system of *Pseudomonas aeruginosa*: infection by injection. *Nature Reviews Microbiology* **2009**, 7 (9), 654-665.
26. Organization, W. H. *Global Priority List of Antibiotic-Resistance Bacteria to Guide Research, Discovery, and Development of New Antibiotics*; World Health Organization, 2017.

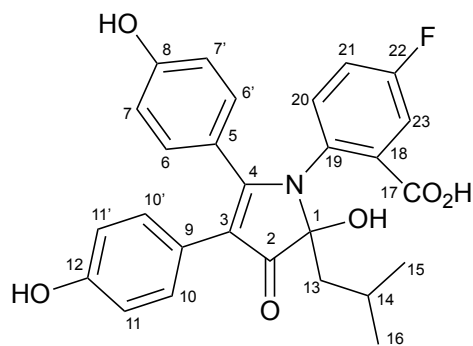
27. Qin, S.; Xiao, W.; Zhou, C.; Pu, Q.; Deng, X.; Lan, L.; Liang, H.; Song, X.; Wu, M. *Pseudomonas aeruginosa*: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. *Signal Transduction and Targeted Therapy* **2022**, 7 (1), 199.
28. Helynck, G.; Dubertret, C.; Mayaux, J.-F.; Leboul, J. Isolation of RP 71955, a new anti-HIV-1 peptide secondary metabolite. *The Journal of Antibiotics* **1993**, 46 (11), 1756-1757.
29. Tan, S.; Ludwig, K. C.; Müller, A.; Schneider, T.; Nodwell, J. R. The Lasso Peptide Siamycin-I Targets Lipid II at the Gram-Positive Cell Surface. *ACS Chemical Biology* **2019**, 14 (5), 966-974.
30. Constantine, K.; Friedrichs, M.; Detlefsen, D.; Nishio, M.; Tsunakawa, M.; Furumai, T.; Ohkuma, H.; Oki, T.; Hill, S.; Brucoleri, R.; et al. High-resolution solution structure of siamycin II: Novel amphipathic character of a 21-residue peptide that inhibits HIV fusion. *Journal of Biomolecular NMR* **1995**, 5 (3), 271-286.
31. Shao, M.; Ma, J.; Li, Q.; Ju, J. Identification of the Anti-Infective Aborycin Biosynthetic Gene Cluster from Deep-Sea-Derived *Streptomyces* sp. SCSIO ZS0098 Enables Production in a Heterologous Host. *Marine Drugs* **2019**, 17 (2), 127.
32. van Santen, J. A.; Jacob, G.; Singh, A. L.; Aniebok, V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N.; et al. The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery. *ACS Central Science* **2019**, 5 (11), 1824-1833.
33. Dictionary of Natural Products. CRC press, Taylor & Francis group. <https://dnp.chemnetbase.com/chemical/ChemicalSearch.xhtml?dswid=3149> (accessed April 2023).
34. Katahira, R.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. MS-271, A novel inhibitor of calmodulin-activated myosin light chain kinase from *Streptomyces* sp.—II. Solution structure of MS-271: Characteristic features of the 'lasso' structure. *Bioorganic & Medicinal Chemistry* **1996**, 4 (1), 121-129.
35. Potterat, O.; Stephan, H.; Metzger, J. W.; Gnau, V.; Zähler, H.; Jung, G. Aborycin – A Tricyclic 21-Peptide Antibiotic Isolated from *Streptomyces griseoflavus*. *Liebigs Annalen der Chemie* **2006**, 1994 (7), 741-743.
36. Frechet, D.; Guitton, J. D.; Herman, F.; Faucher, D.; Helynck, G.; Monegier du Sorbier, B.; Ridoux, J. P.; James-Surcouf, E.; Vuilhorgne, M. Solution structure of RP 71955, a new 21 amino acid tricyclic peptide active against HIV-1 virus. *Biochemistry* **1994**, 33 (1), 42-50.
37. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, 30 (1), 108-160.

38. Zubarev, R. A.; Chivanov, V. D.; Håkansson, P.; Sundqvist, B. U. R.; Ens, W. Peptide sequencing by partial acid hydrolysis and high resolution plasma desorption mass spectrometry. *Rapid Communications in Mass Spectrometry* **2005**, 8 (11), 906-912.
39. Li, J.-Y.; Liang, J.-Y.; Liu, Z.-Y.; Yi, Y.-Z.; Zhao, J.; Huang, Z.-Y.; Chen, J. Multicopy Chromosome Integration and Deletion of Negative Global Regulators Significantly Increased the Heterologous Production of Aborycin in *Streptomyces coelicolor*. *Marine Drugs* **2023**, 21 (10), 534.
40. Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J. N.; et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research* **2023**, 51 (W1), W46-W50.
41. Terlouw, B. R.; Blin, K.; Navarro-Muñoz, J. C.; Avalon, N. E.; Chevrette, M. G.; Egbert, S.; Lee, S.; Meijer, D.; Recchia, Michael J. J.; Reitz, Zachary L.; et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Research* **2023**, 51 (D1), D603-D610.
42. Altschul, S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **1997**, 25 (17), 3389-3402.
43. Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T. Biosynthetic Gene Cluster of a D-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem* **2018**, 19 (19), 2045-2048.
44. Nakashima, Y.; Kawakami, A.; Ogasawara, Y.; Maeki, M.; Tokeshi, M.; Dairi, T.; Morita, H. Structure of lasso peptide epimerase MslH reveals metal-dependent acid/base catalytic mechanism. *Nature Communications* **2023**, 14 (1), 4752.
45. Wang, Q.; Zhao, J.; Liu, Z.; Ding, S.; Huang, Z.; Chen, J. Genomic Insights and Synthetic Biology Applications of Marine Actinomycete *Streptomyces griseoincarnatus* HNS054. *International Journal of Molecular Sciences* **2024**, 25 (6), 3127.
46. Aiello, D.; Williams, J. D.; Majgier-Baranowska, H.; Patel, I.; Peet, N. P.; Huang, J.; Lory, S.; Bowlin, T. L.; Moir, D. T. Discovery and Characterization of Inhibitors of *Pseudomonas aeruginosa* Type III Secretion. *Antimicrobial Agents and Chemotherapy* **2010**, 54 (5), 1988-1999.
47. Morgan, J. M.; Lam, H. N.; Delgado, J.; Luu, J.; Mohammadi, S.; Isberg, R. R.; Wang, H.; Auerbuch, V. An Experimental Pipeline for Initial Characterization of Bacterial Type III Secretion System Inhibitor Mode of Action Using Enteropathogenic *Yersinia*. *Frontiers in Cellular and Infection Microbiology* **2018**, 8, 404.
48. Voytik-Harbin, S. L.; Brightman, A. O.; Waisner, B.; Lamar, C. H.; Badylak, S. F. Application and evaluation of the alamarblue assay for cell growth and survival of fibroblasts. *In Vitro Cellular & Developmental Biology - Animal* **1998**, 34 (3), 239-246.
49. Guo, H.; Geddes, E. J.; Opperman, T. J.; Heuck, A. P. Cell-Based Assay to Determine Type 3 Secretion System Translocon Assembly in *Pseudomonas aeruginosa* Using Split Luciferase. *ACS Infectious Diseases* **2023**, 9 (12), 2652-2664.

50. Ngo, T.-D.; Plé, S.; Thomas, A.; Barette, C.; Fortuné, A.; Bouzidi, Y.; Fauvarque, M.-O.; Pereira de Freitas, R.; Francisco Hilário, F.; Attrée, I.; et al. Chimeric Protein–Protein Interface Inhibitors Allow Efficient Inhibition of Type III Secretion Machinery and *Pseudomonas aeruginosa* Virulence. *ACS Infectious Diseases* **2019**, 5 (11), 1843-1854.
51. Jander, G.; Rahme, L. G.; Ausubel, F. M. Positive Correlation between Virulence of *Pseudomonas aeruginosa* Mutants in Mice and Insects. *Journal of Bacteriology* **2000**, 182 (13), 3843-3845.
52. Lereclus, D.; Nielsen-Leroux, C.; Ramarao, N. The Insect *Galleria mellonella* as a Powerful Infection Model to Investigate Bacterial Pathogenesis. *Journal of Visualized Experiments* **2012**, (70), e4392.

Appendix A: Supplementary Experimental Data for Chapter 2

A.1 Spectral data for open intermediate of fluorinated discoipyrrole A (2.21)



2.21

Table A.1. NMR assignments for elucidation of **2.21** (^1H 800 MHz, ^{13}C 200 MHz, CD_3OD).

No.	δ_{H} , mult. (J in Hz)	δ_{C} (J_{CF} in Hz)	COSY	HMBC
1		94.0		
2		202.2		
3		113.9		
4		176.7		
5		122.2		
6, 6'	6.99, d (8.7)	133.1	7, 7'	4, 7, 7', 8
7, 7'	6.51, d (8.7)	115.9	6, 6'	5, 8
8		160.5		
9		124.9		
10, 10'	6.93, d (8.7)	131.6	11, 11'	3, 11, 11', 12
11, 11'	6.62, d (8.7)	115.7	10, 10'	9, 12
12		156.5		
13	1.92, m	45.2	14	1, 2, 14, 15, 16
14	2.02, m	25.0	13, 15, 16	1, 13, 15, 16
15	1.16, d (6.7)	24.9	14	13, 14, 16
16	0.90, d (6.6)	24.3	14	13, 14, 15
17		173.0		
18		144.1		
19		133.7		
20	7.46, dd (8.5, 5.0)	134.1	21	18, 19, 22
21	7.20, m	117.3 (21.8)	20	18, 22, 23
22		162.8 (248.3)		
23	7.21, m	118.2 (23.6)		17, 18, 21, 22
CO ₂ H	8.54, br. s			

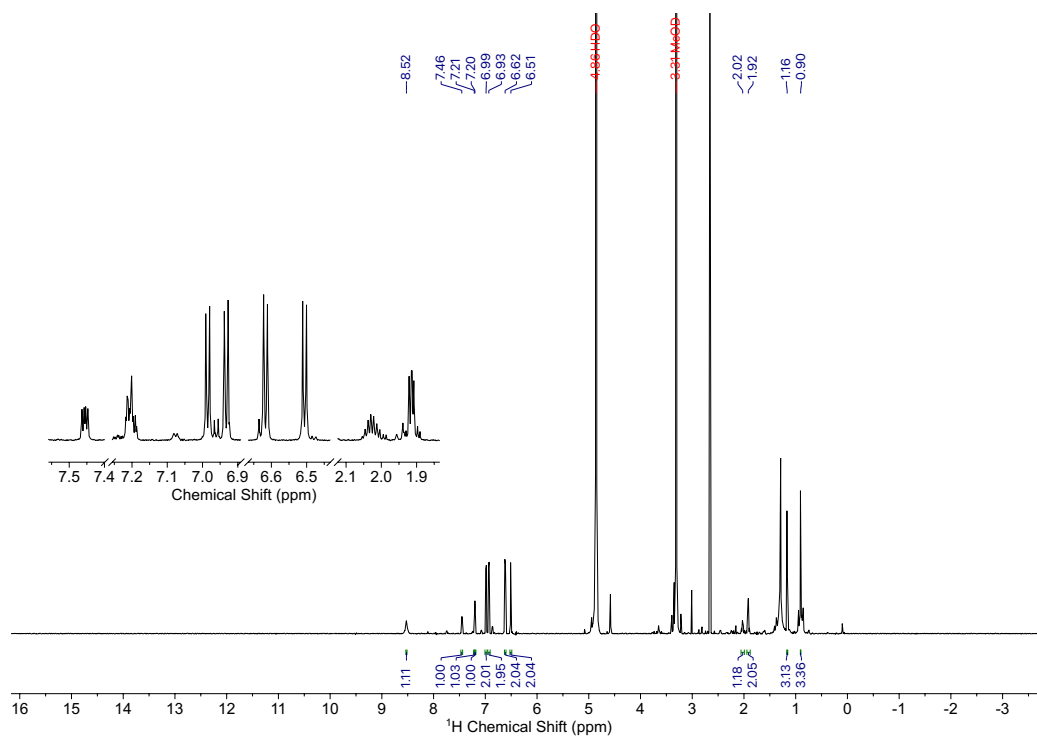


Figure A.1. ¹H NMR spectrum of **2.21** (800 MHz, CD₃OD).

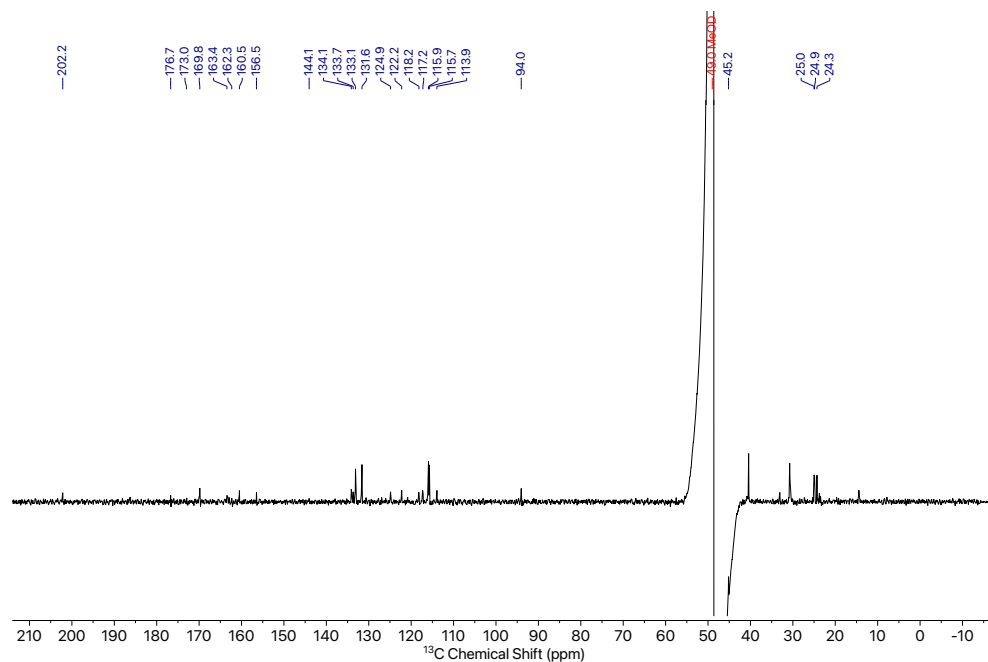


Figure A.2. ¹³C NMR spectrum of **2.21**. (200 MHz, CD₃OD).

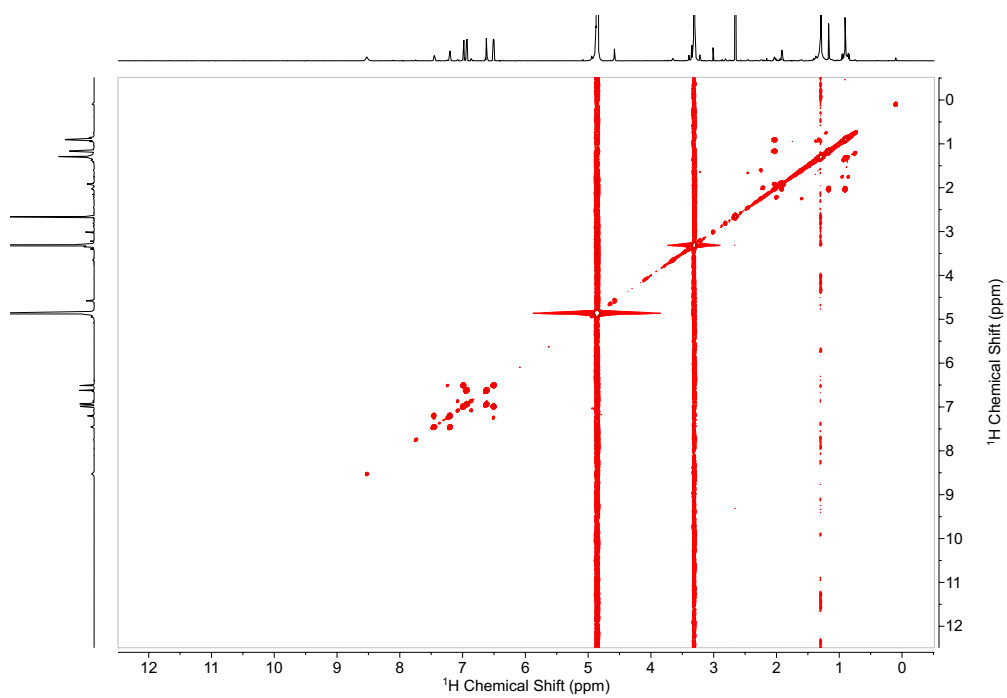


Figure A.3. COSY spectrum of **2.21** (800 MHz, CD_3OD).

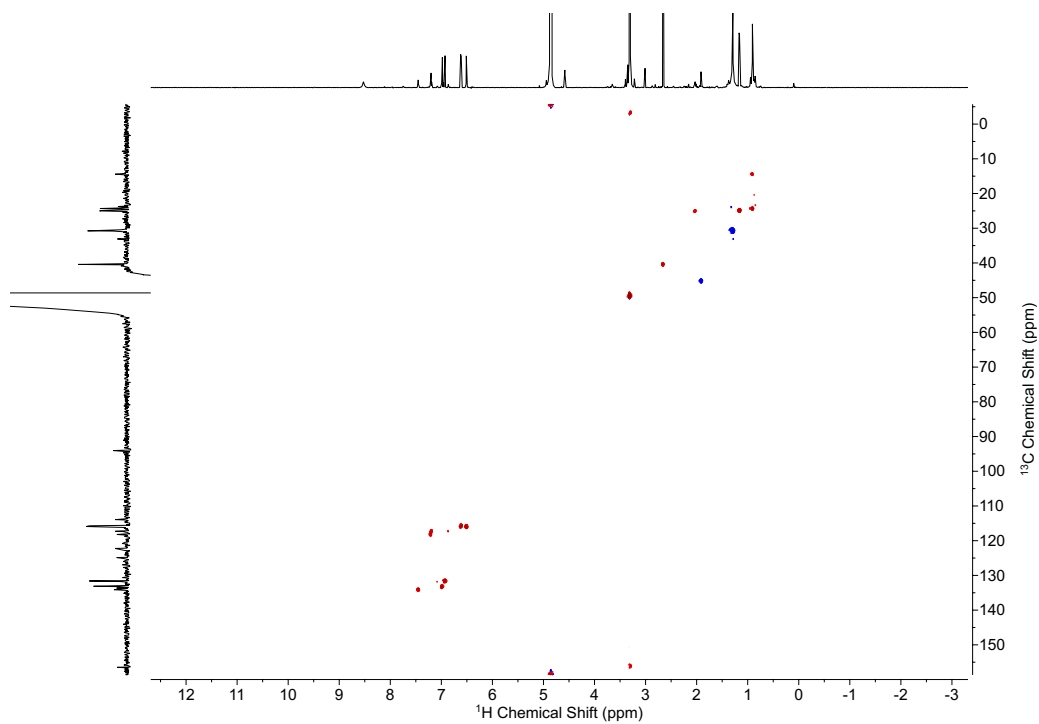


Figure A.4. ^1H – ^{13}C HSQC spectrum of **2.21** (800 MHz, CD_3OD).

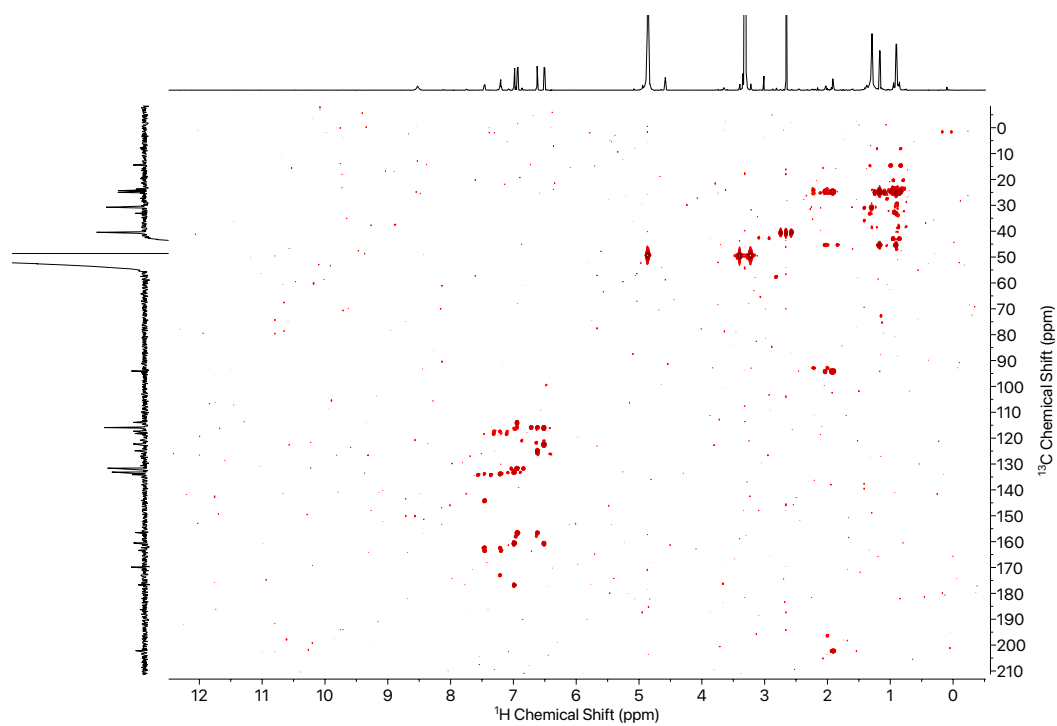


Figure A.5. $^1\text{H} - ^{13}\text{C}$ HMBC spectrum of **2.21** (800 MHz, CD_3OD).

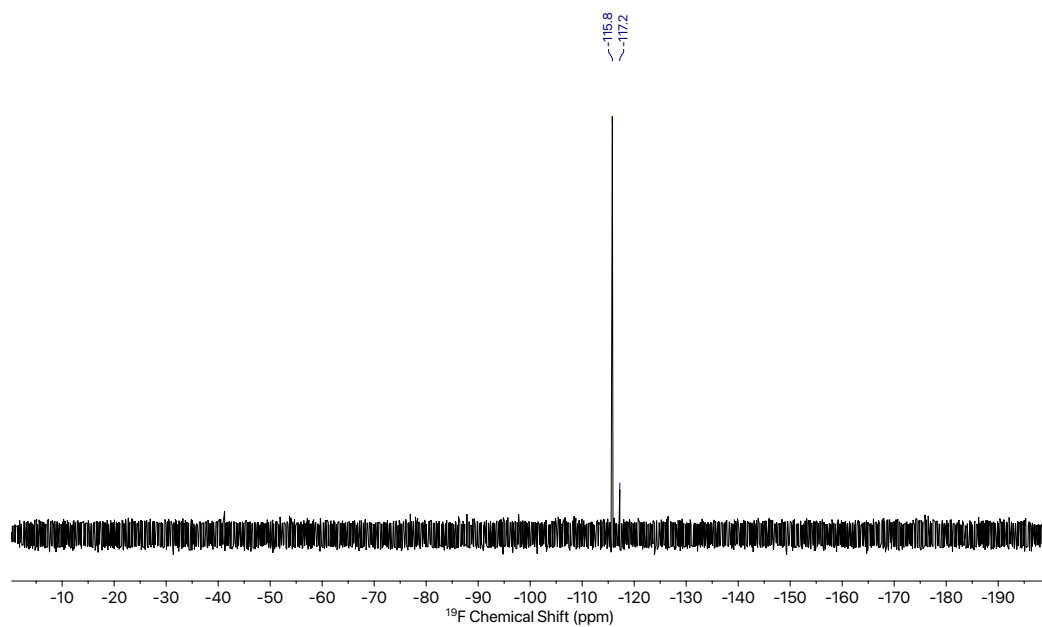


Figure A.6. Full ^{19}F spectrum of **2.21** (470 MHz, CD_3OD).

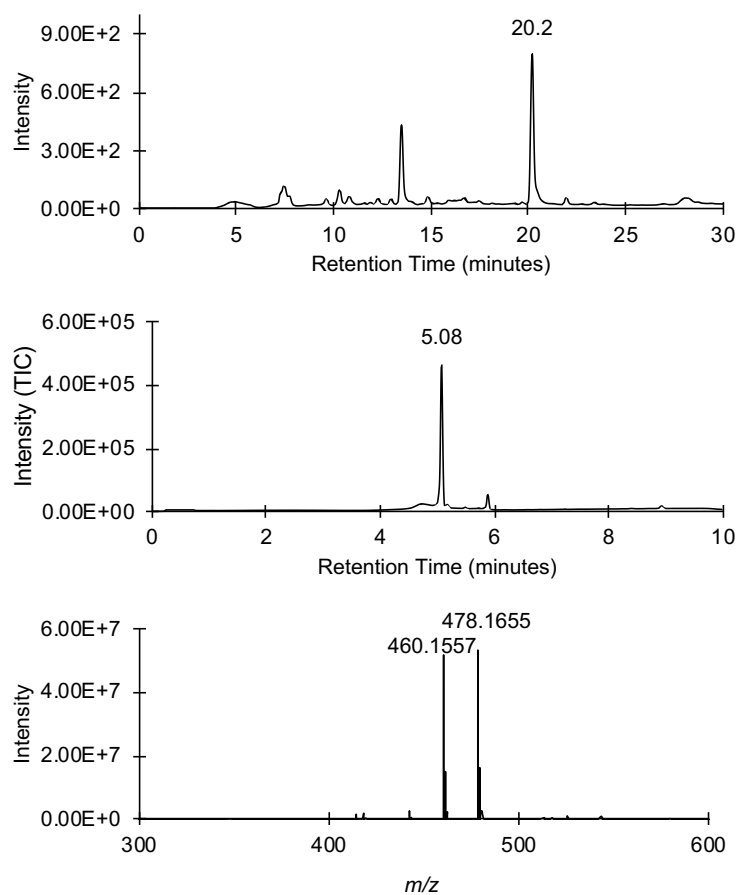
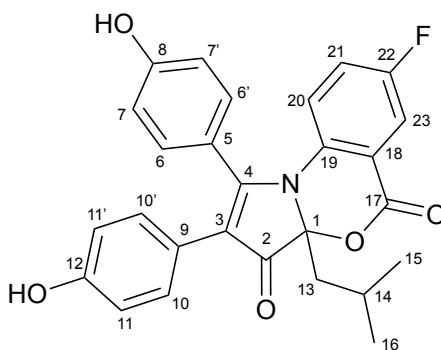


Figure A.7. *Top:* HPLC chromatogram of **2.21** at 254 nm. *Middle:* TIC chromatogram of **2.21** (absolute abundance). *Bottom:* Mass spectrum of **2.21**.

A.2 Spectral data for fluorinated discoipyrrole A (2.20)



2.20

Table A.2. NMR assignments for elucidation of **2.20** (^1H 800 MHz, ^{13}C 200 MHz, CD_3OD).

No.	δ_{H} , mult. (J in Hz)	δ_{C} (J_{CF} in Hz)	COSY	HMBC
1		92.9		
2		196.2		
3		116.4		
4		141.4		
5		120.9		
6, 6'	7.08, d (8.3)	131.7	7, 7'	4, 7, 7', 8
7, 7'	6.86, d (8.3)	117.0	6, 6'	5, 8
8		161.2		
9		121.6		
10, 10'	6.96, d (8.6)	131.4	11, 11'	3, 12
11, 11'	6.63, d (8.6)	115.9	10, 10'	3, 9, 12
12		157.7		
13	2.21, dd (14.1, 6.1) 2.00, dd (14.1, 6.5)	42.8	14, 13 14, 13	1, 14, 15, 16 1, 2, 14, 15, 16
14	1.76, m	25.3	13, 15, 16	15, 16
15	0.95, d (6.7)	24.2	14	13, 14, 16
16	0.85, d (6.7)	23.4	14	13, 14, 15
17		162.2		
18		138.1		
19		135.0		
20	6.51, dd (9.0, 4.3)	126.0	21	19, 22
21	7.24, td (8.6, 3.0)	123.4	20	18, 19
22		160.8 (251.0)		
23	7.74, dd (8.0, 3.0)	117.3 (25.2)		17

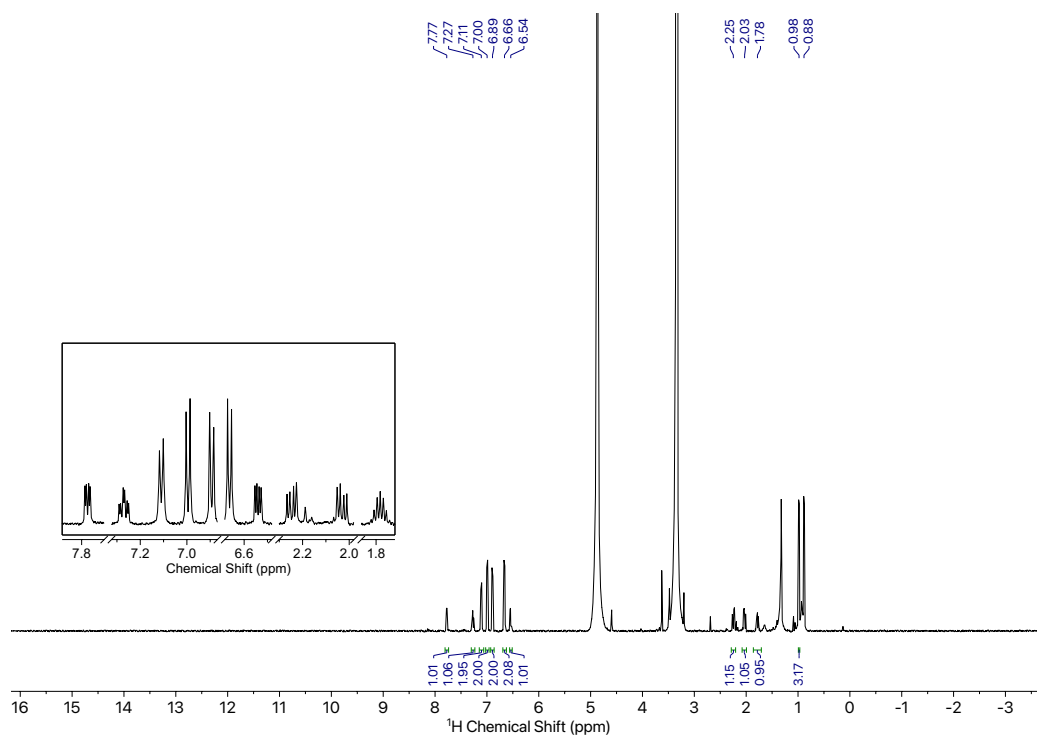


Figure A.8. ¹H NMR spectrum of **2.20** (800 MHz, CD₃OD).

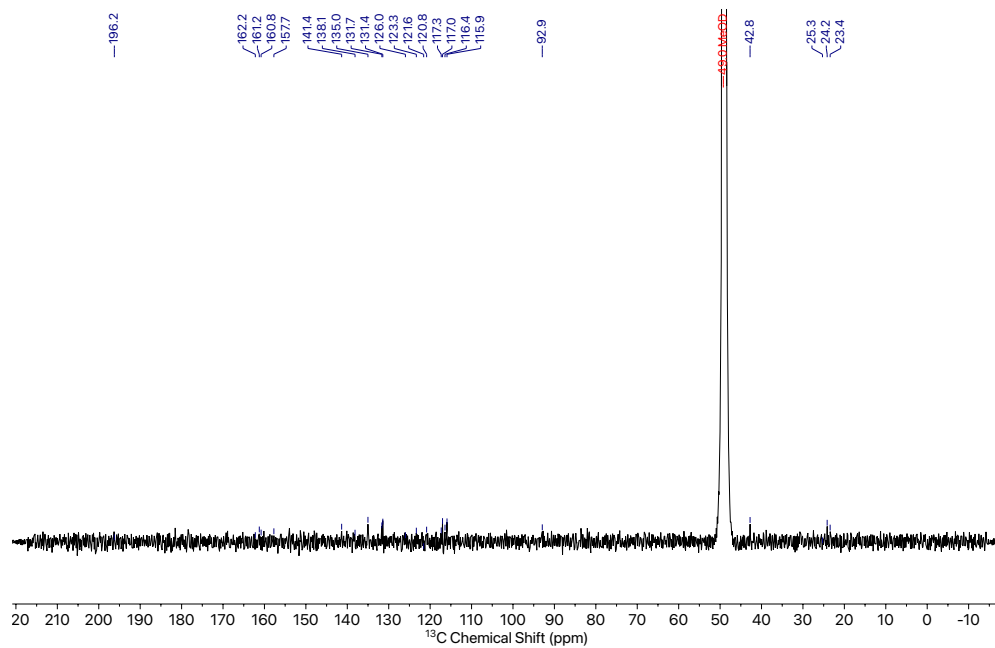


Figure A.9. ¹³C NMR spectrum of **2.20**. (200 MHz, CD₃OD).

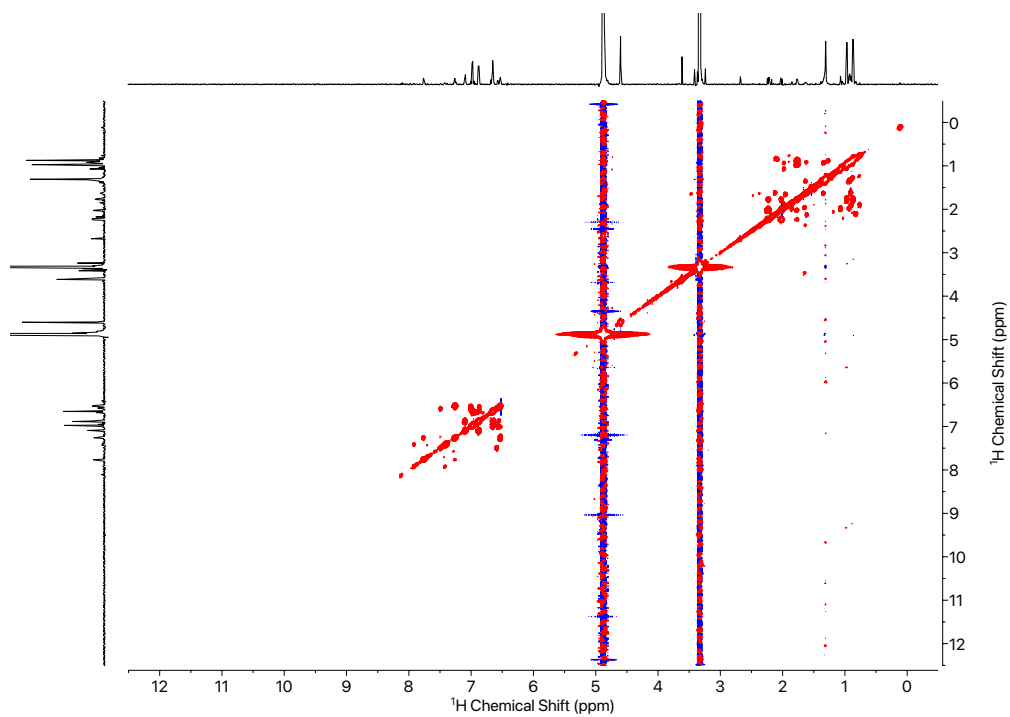


Figure A.10. COSY spectrum of **2.20** (800 MHz, CD₃OD).

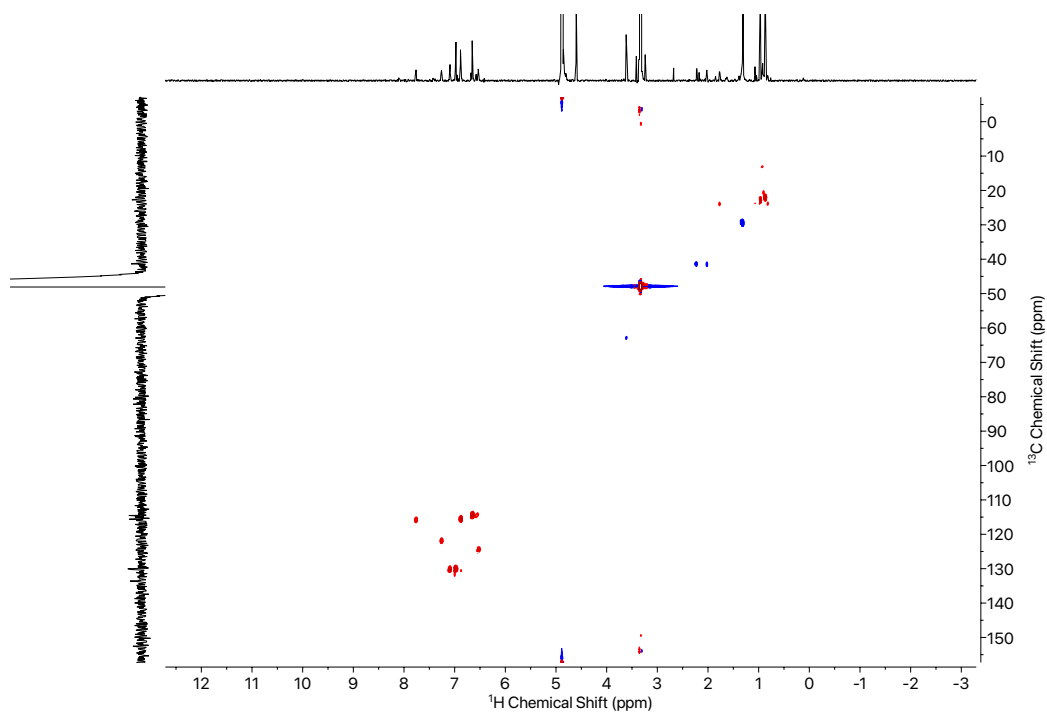


Figure A.11. ^1H – ^{13}C HSQC spectrum of **2.20** (800 MHz, CD₃OD).

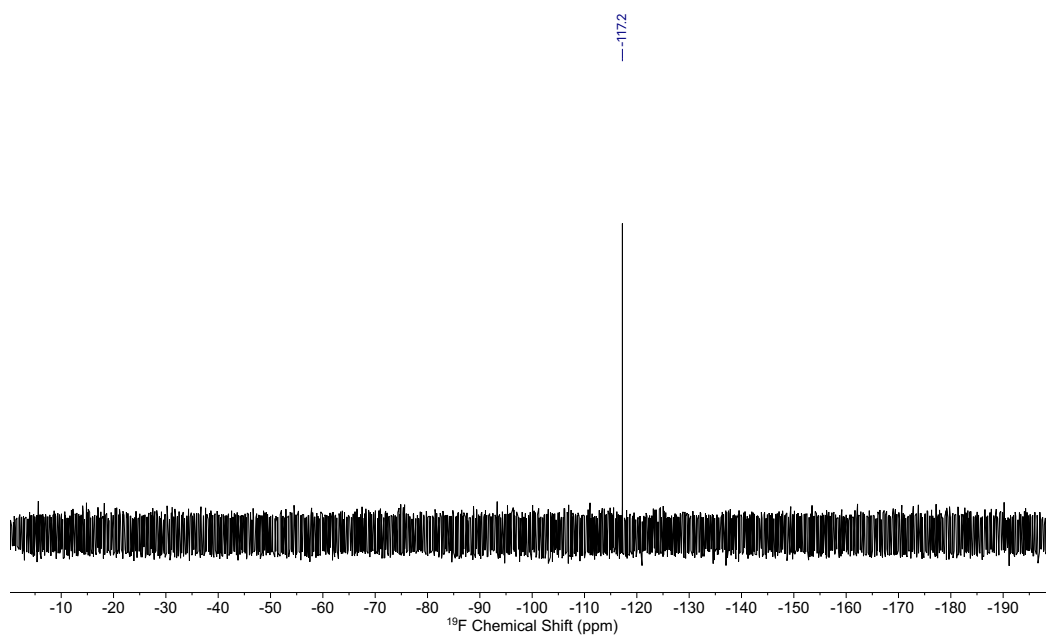


Figure A.12. Full ^{19}F NMR spectrum of **2.20** (470 MHz, CD_3OD).

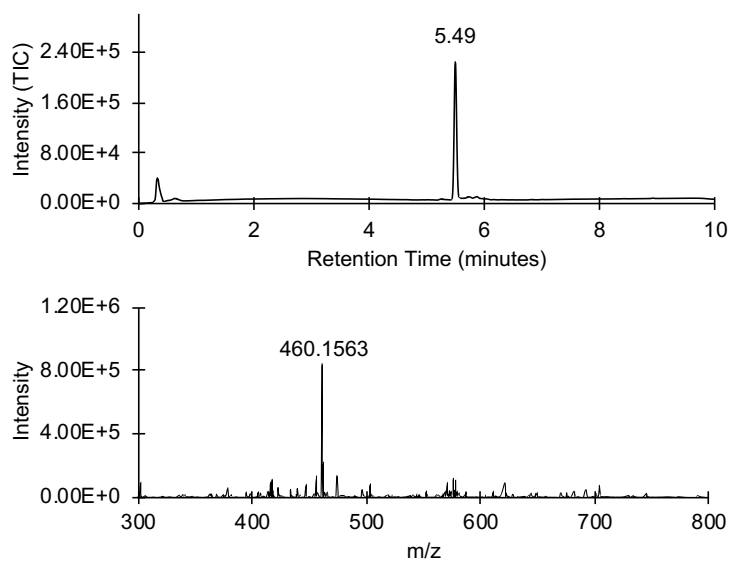


Figure A.13. *Top:* TIC chromatogram of **2.20** (absolute abundance). *Bottom:* Mass spectrum of **2.20**.

A.3 Spectral data for the synthesis of 2.20

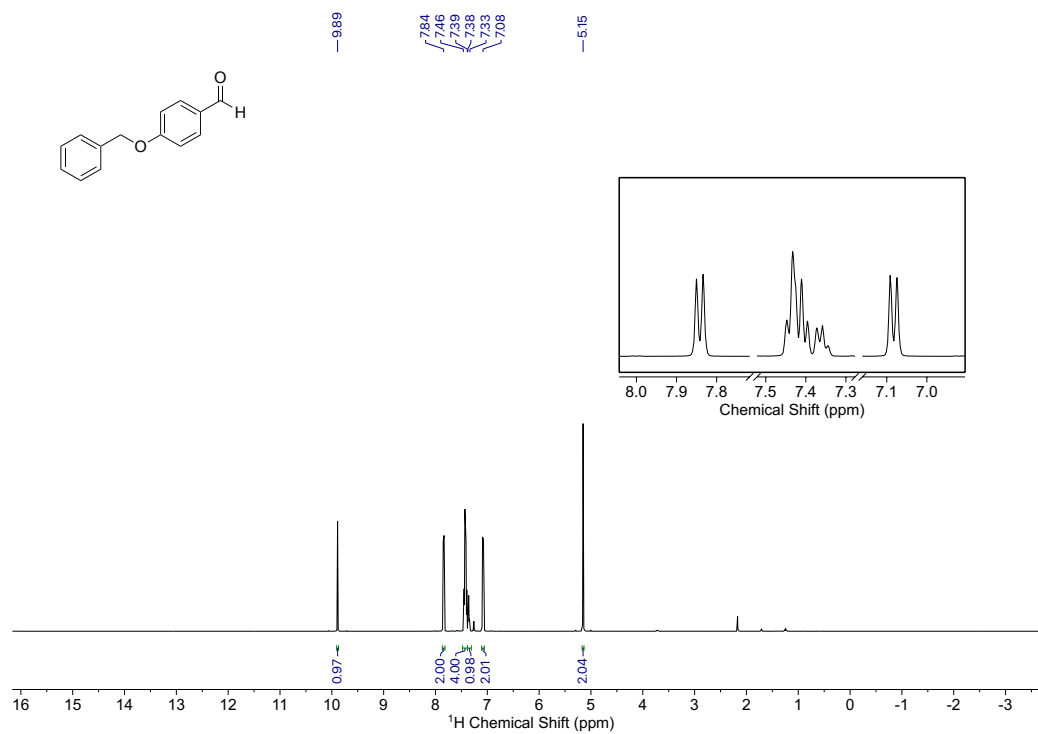


Figure A.14. ¹H NMR spectrum of **2.16** (500 MHz, CDCl₃).

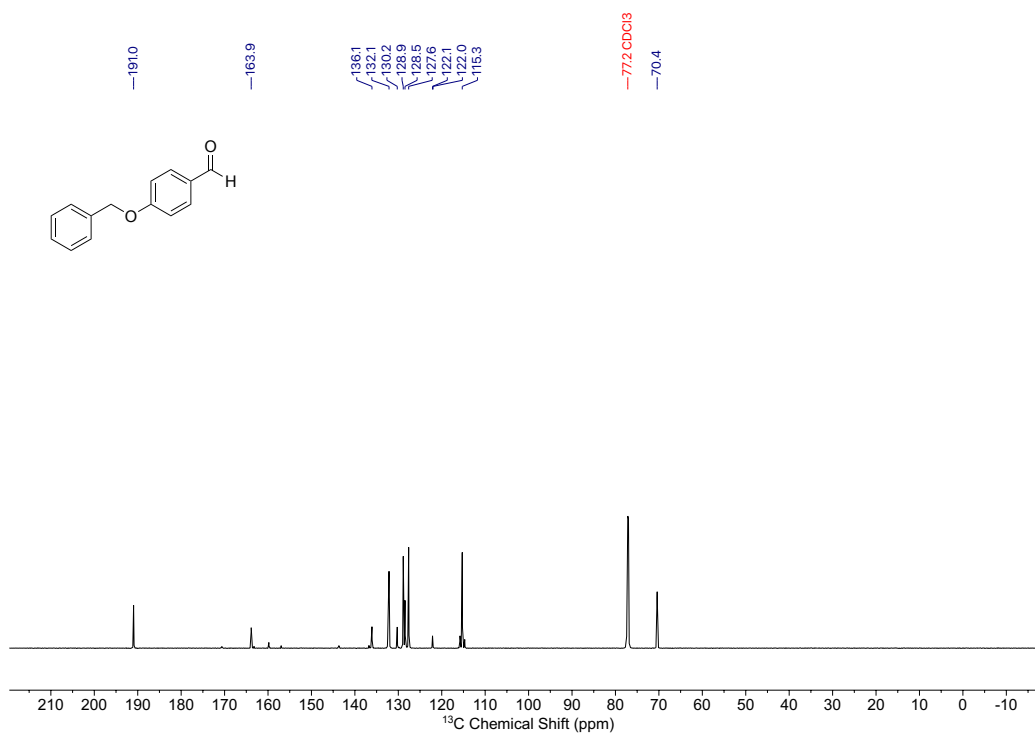


Figure A.15. ¹³C NMR spectrum of **2.16** (125 MHz, CDCl₃).

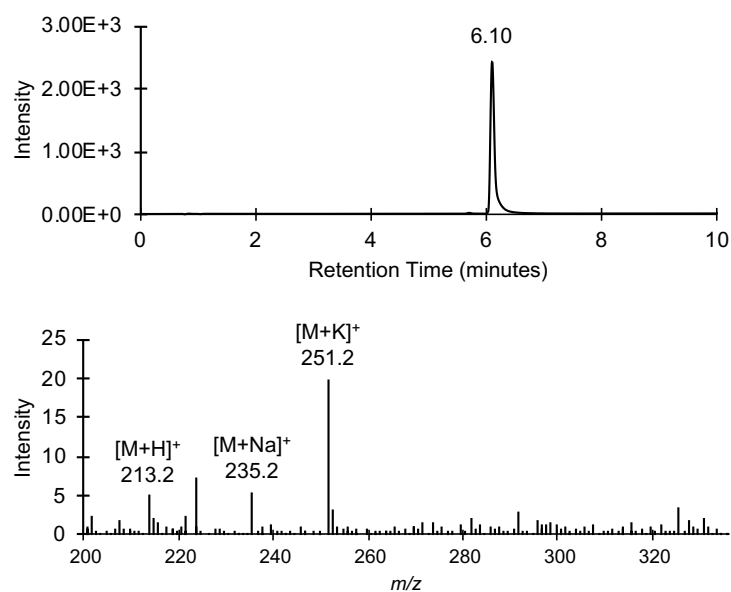


Figure A.16. *Top:* UV chromatogram of **2.16** at 254 nm. *Bottom:* Mass spectrum of **2.16**.

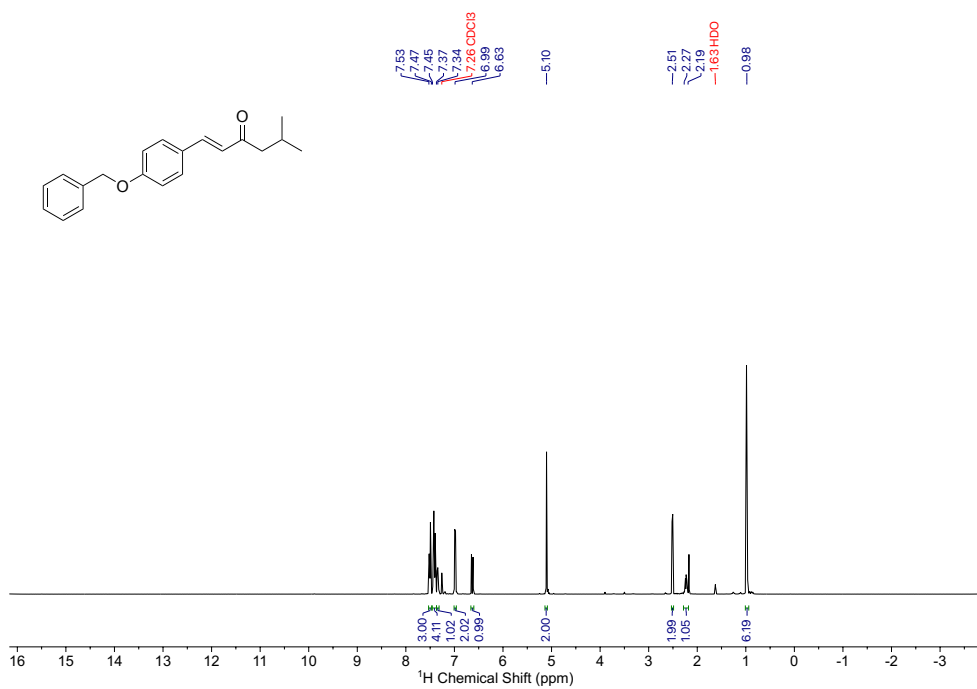


Figure A.17. ¹H NMR spectrum of **2.17** (500 MHz, CDCl₃).

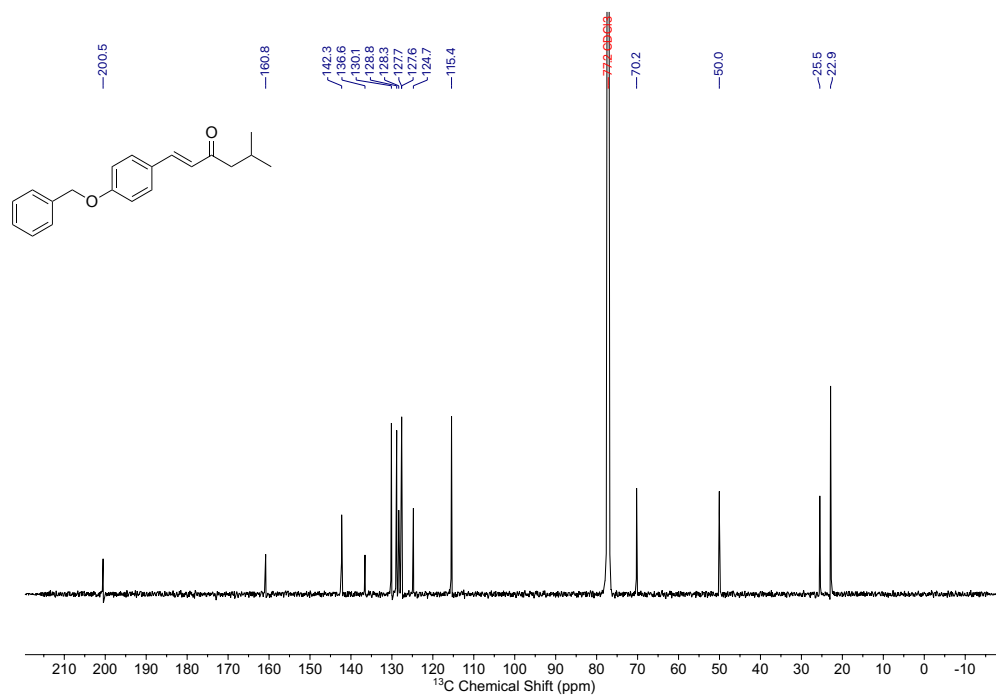


Figure A.18. ¹³C NMR spectrum of **2.17** (200 MHz, CDCl₃).

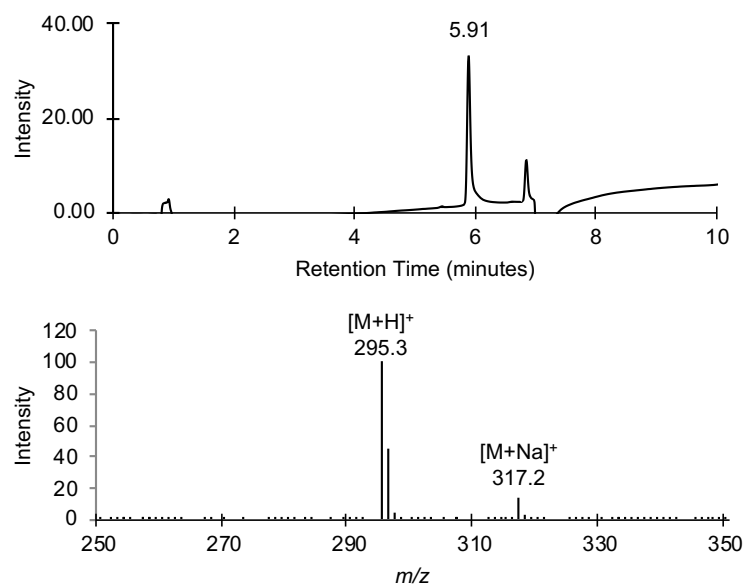


Figure A.19. *Top:* UV chromatogram of **2.17** at 330 nm. *Bottom:* Mass spectrum of **2.17**.

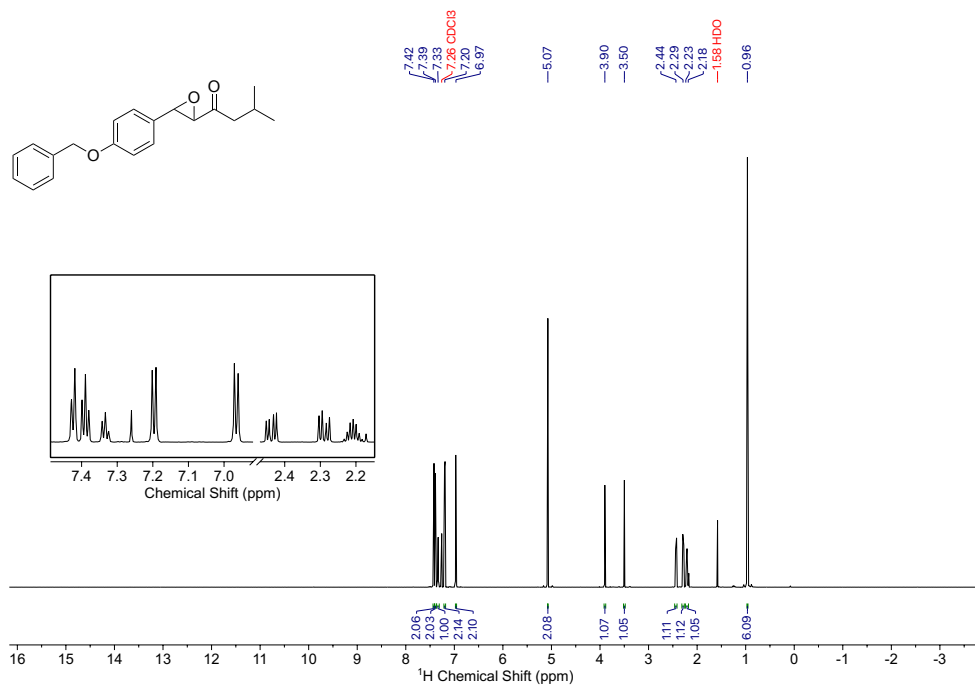


Figure A.20. ^1H NMR spectrum of **2.18** (500 MHz, CDCl_3).

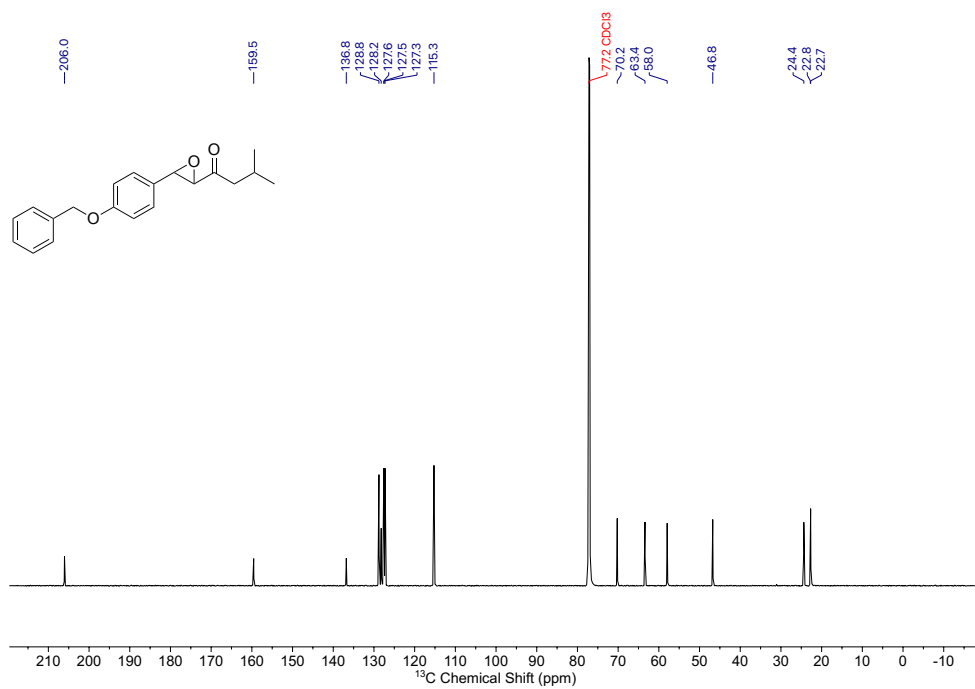


Figure A.21. ¹³C NMR spectrum of **2.18** (200 MHz, CDCl₃).

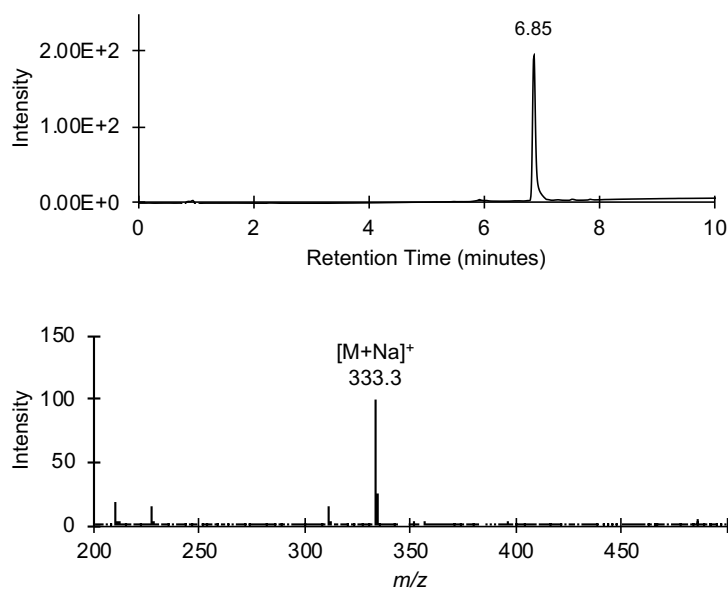


Figure A.22. *Top:* UV chromatogram of **2.18** at 254 nm. *Bottom:* Mass spectrum of **2.18**.

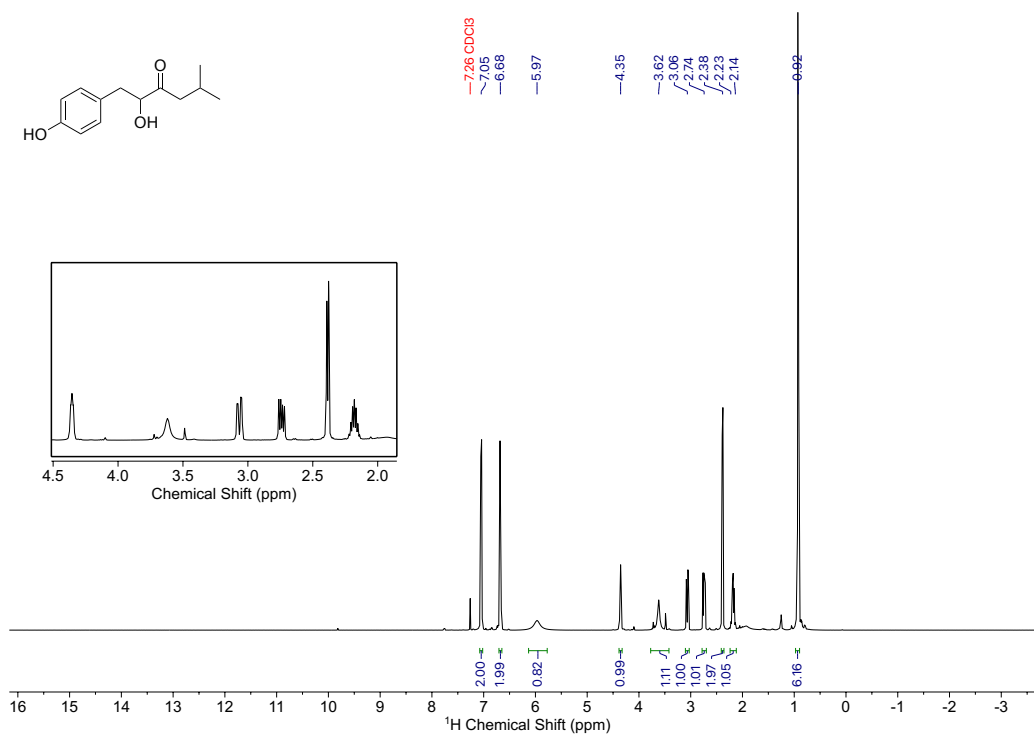


Figure A.23. ¹H NMR spectrum of **2.19** (800 MHz, CDCl₃).

Appendix B: Supplementary Experimental Data for Chapter 3

B.1 Spectral data for amine fragment 3.22

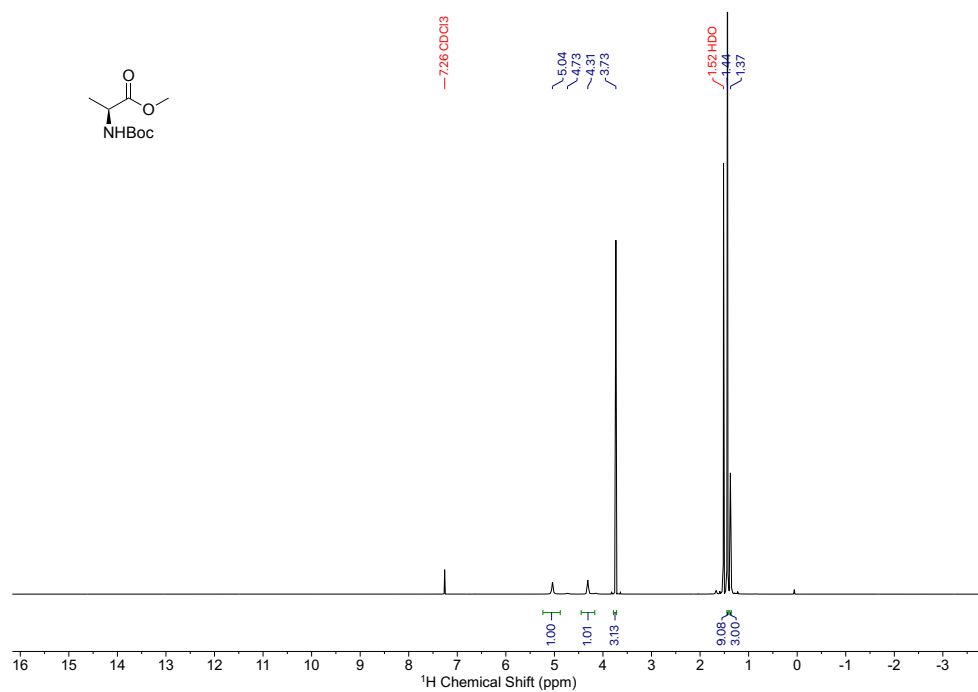


Figure B.1. ¹H NMR spectrum of **3.26** (800 MHz, CDCl₃).

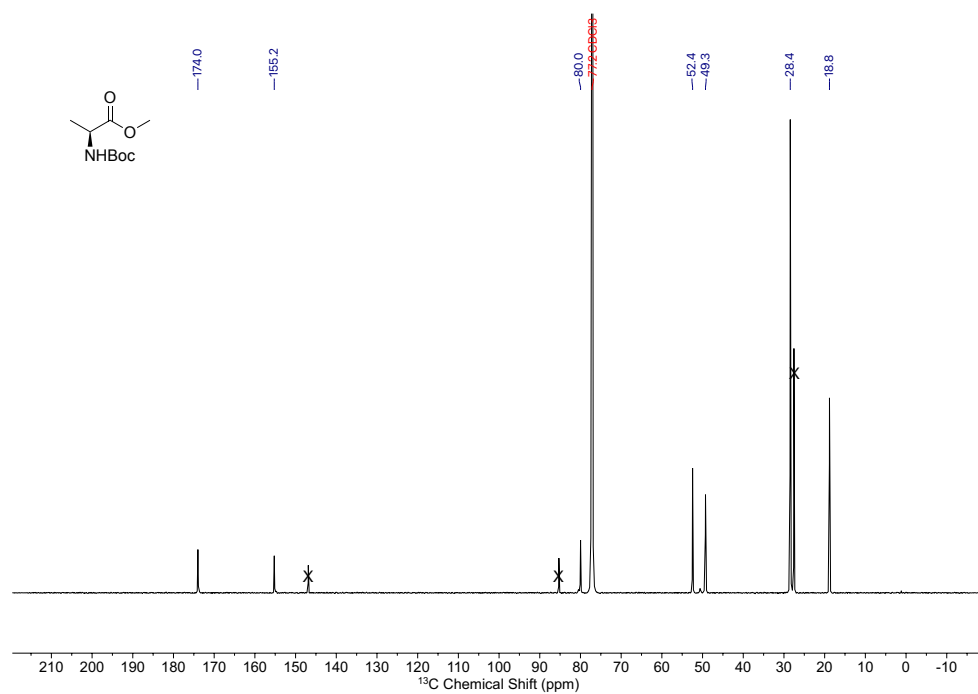


Figure B.2. ¹³C NMR spectrum of **3.26** (200 MHz, CDCl₃). X indicates unreacted Boc₂O

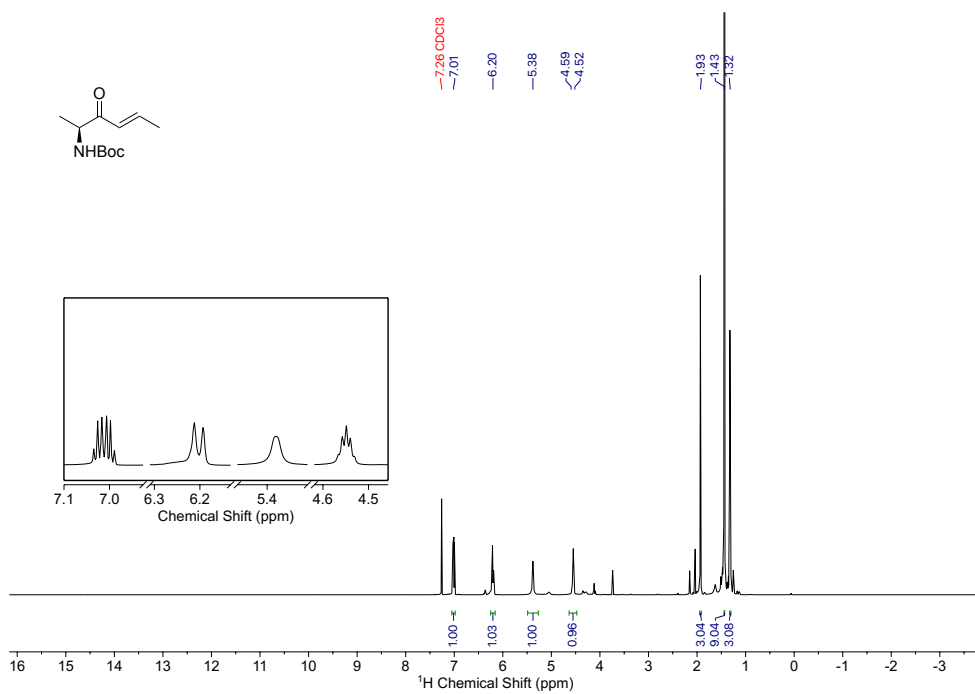


Figure B.3. ¹H NMR spectrum of **3.27** (800 MHz, CDCl₃).

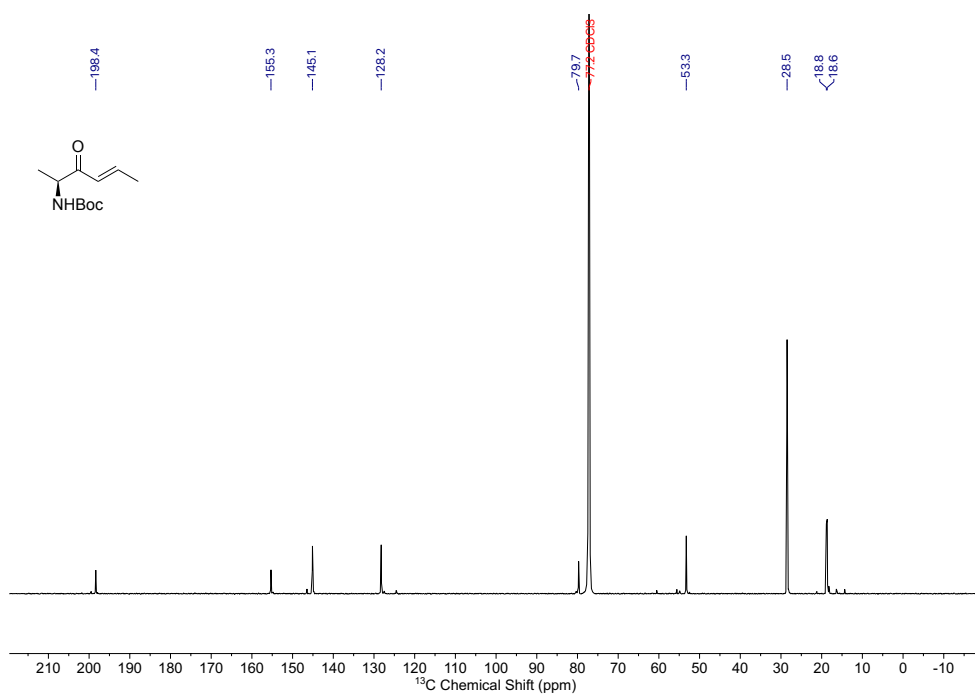


Figure B.4. ¹³C NMR spectrum of **3.27** (200 MHz, CDCl₃).

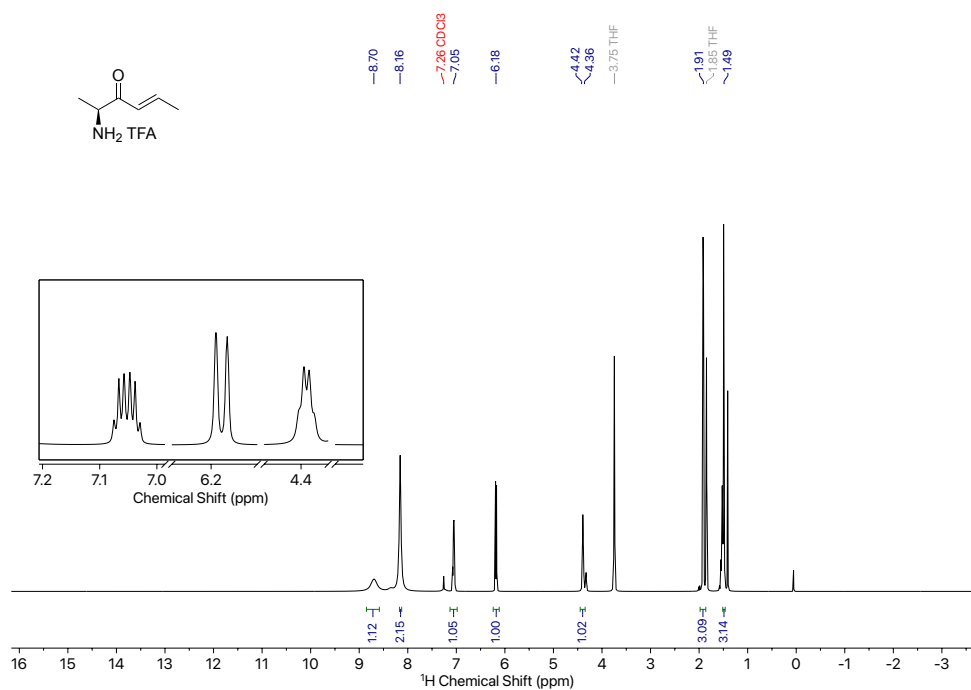


Figure B.5. ¹H NMR spectrum of **3.22** (800 MHz, CDCl₃).

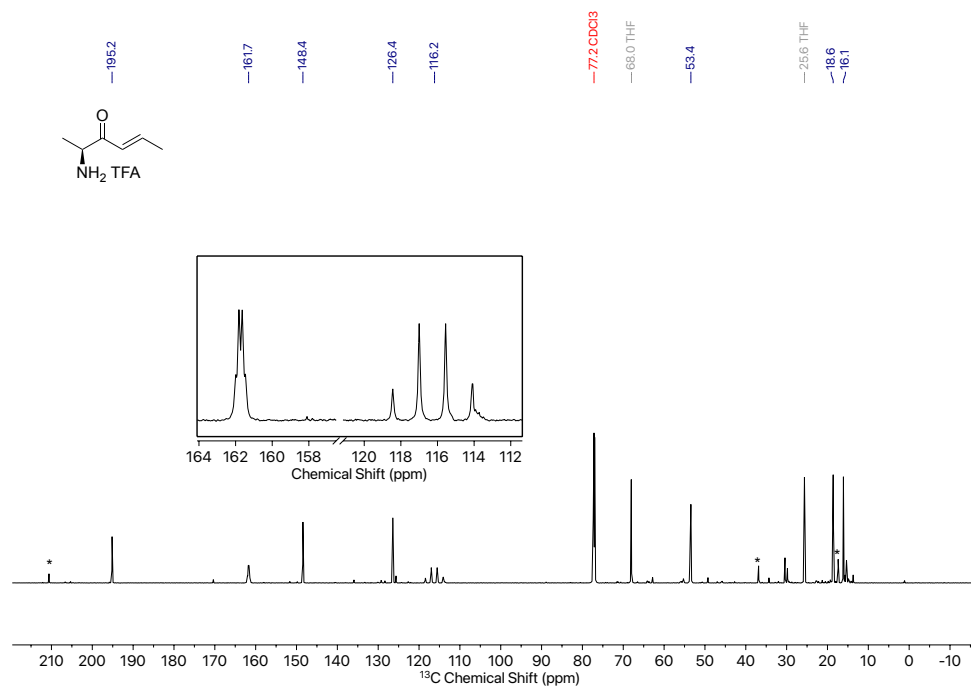


Figure B.6. ¹³C NMR spectrum of **3.22** (200 MHz, CDCl₃). * Indicates unknown impurities

B.2 Spectral data for fragment 3.28

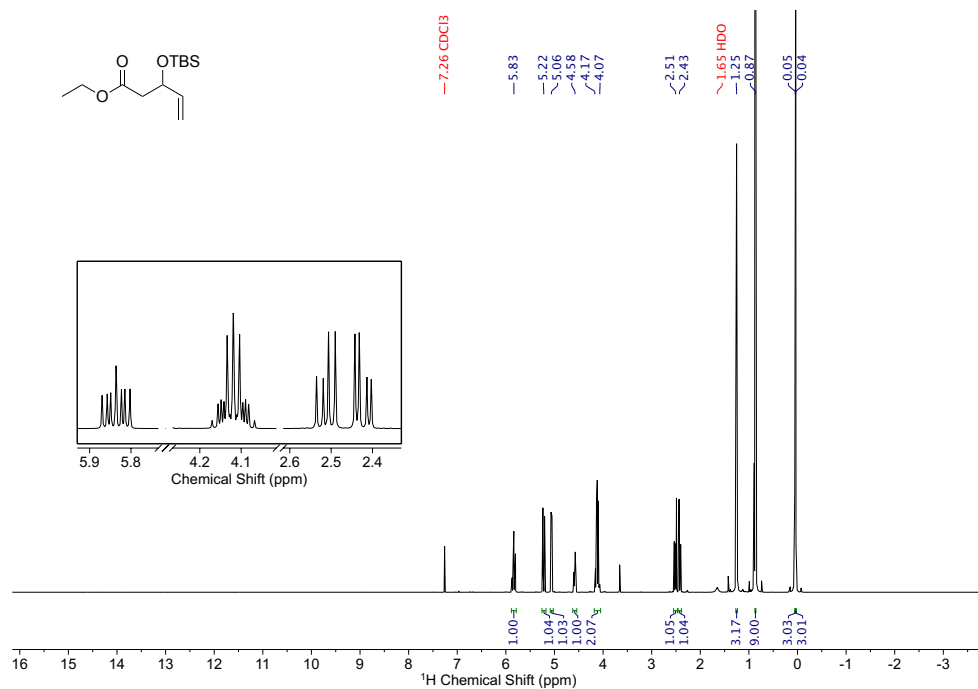


Figure B.7. ¹H NMR spectrum of **3.35** (500 MHz, CDCl₃).

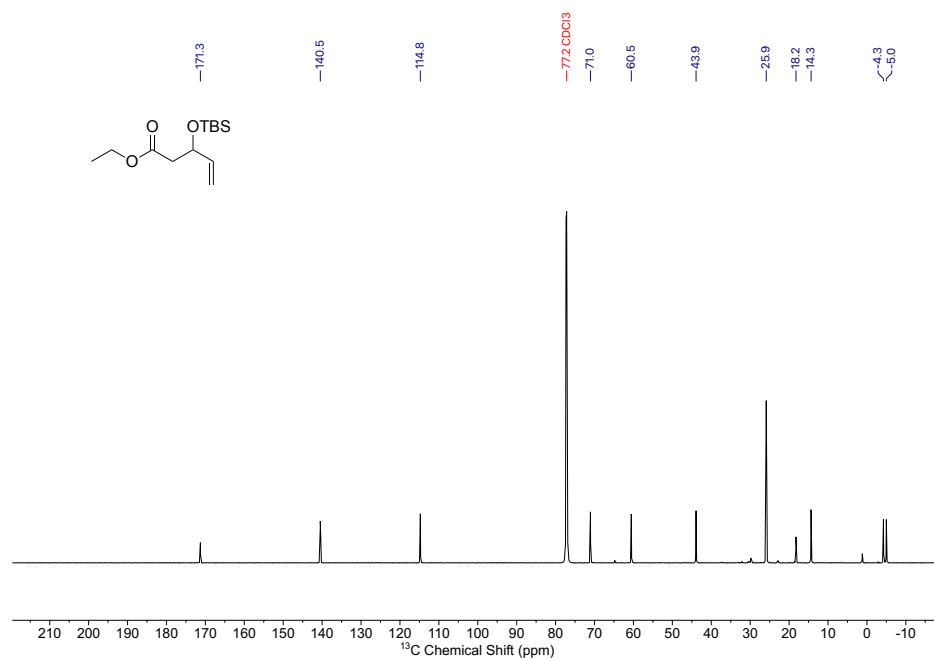


Figure B.8. ¹³C NMR spectrum of **3.35** (200 MHz, CDCl₃).

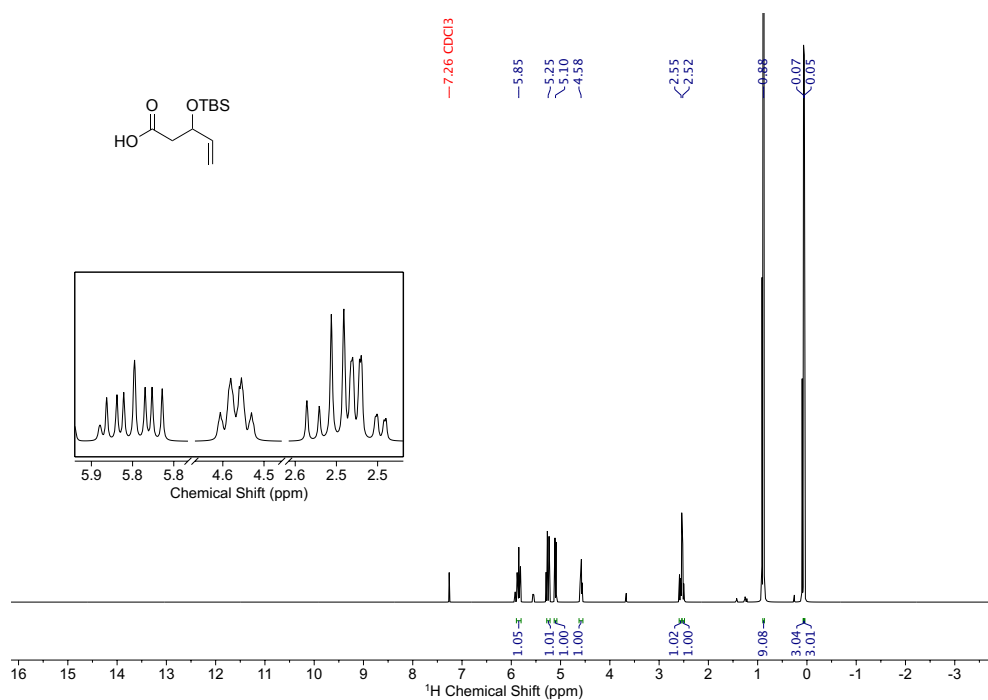


Figure B.9 ¹H NMR spectrum of **3.28** (500 MHz, CDCl₃).

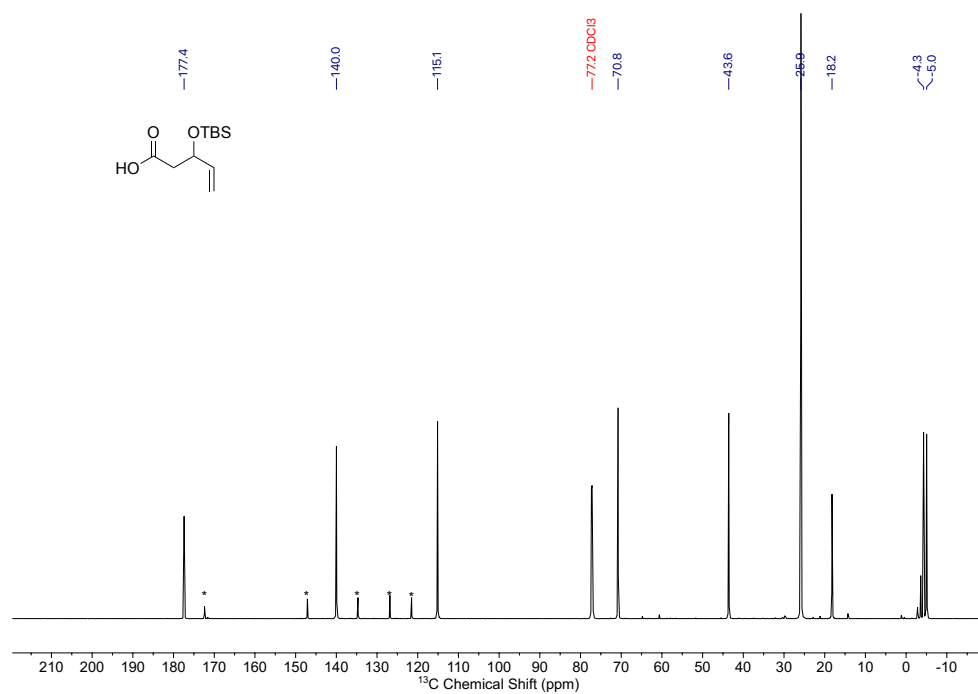


Figure B.10. ¹³C NMR spectrum of **3.28** (200 MHz, CDCl₃). * Indicates unknown impurities

B.3 Spectral data for fragment 3.29

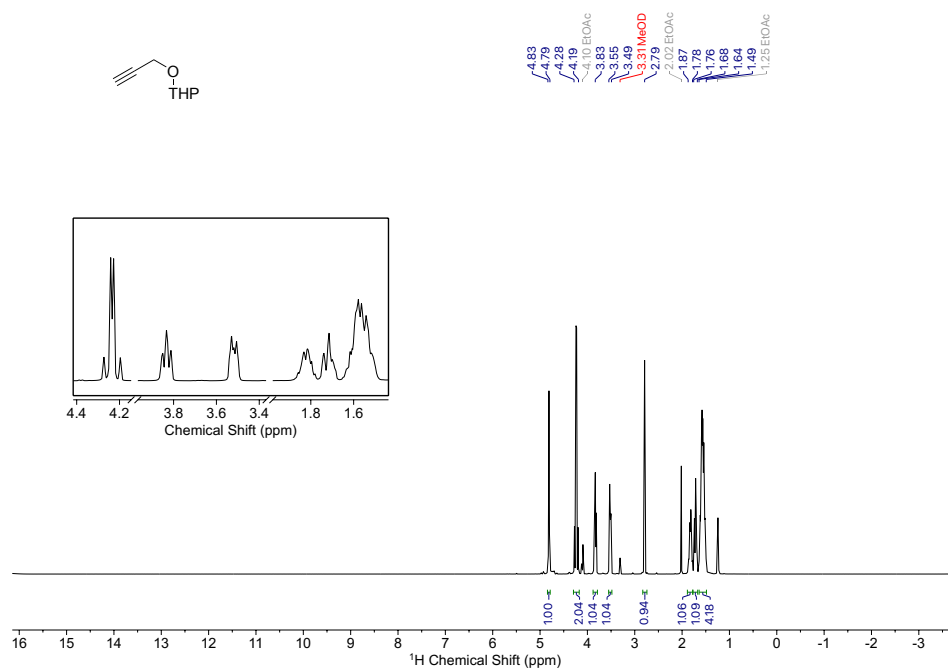


Figure B.11. ¹H NMR spectrum of **3.37** (500 MHz, CD₃OD).

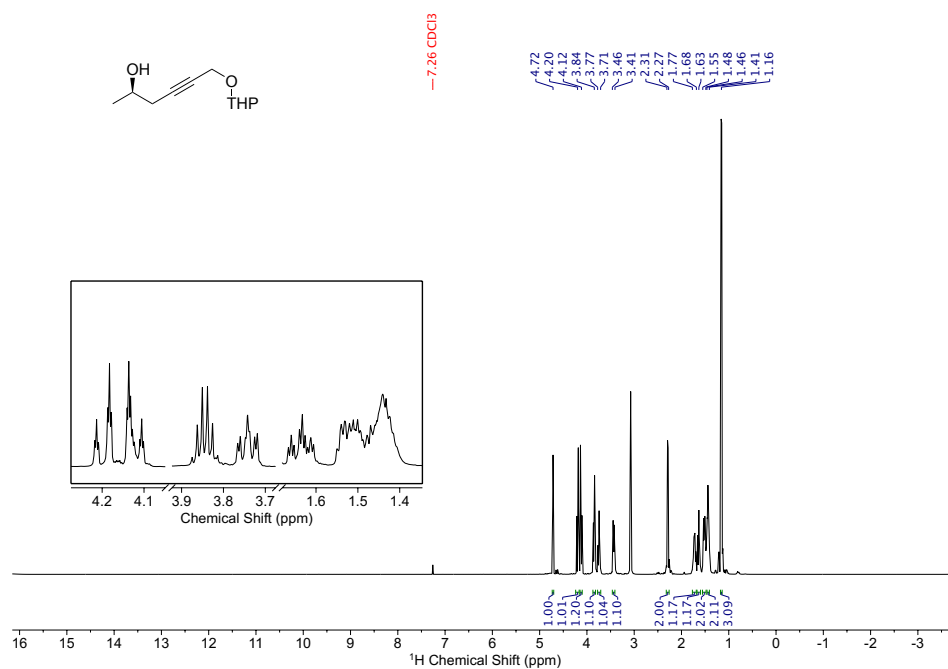


Figure B.12. ¹H NMR spectrum of **3.38** (500 MHz, CDCl₃).

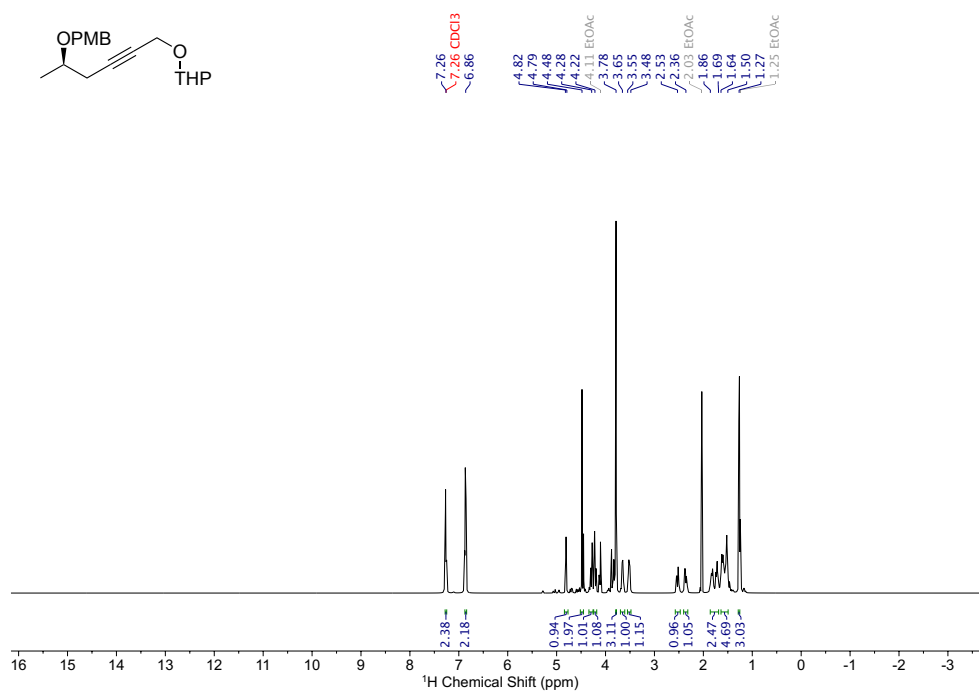


Figure B.13. ¹H NMR spectrum of **3.39** (500 MHz, CDCl₃).

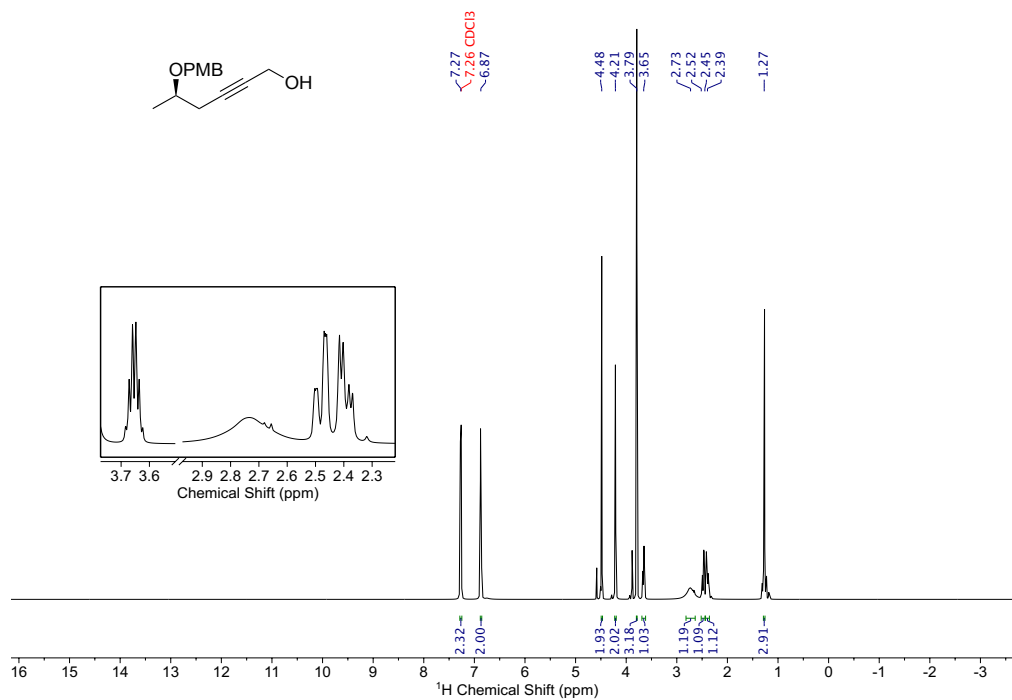


Figure B.14. ¹H NMR spectrum of **3.40** (500 MHz, CDCl₃).

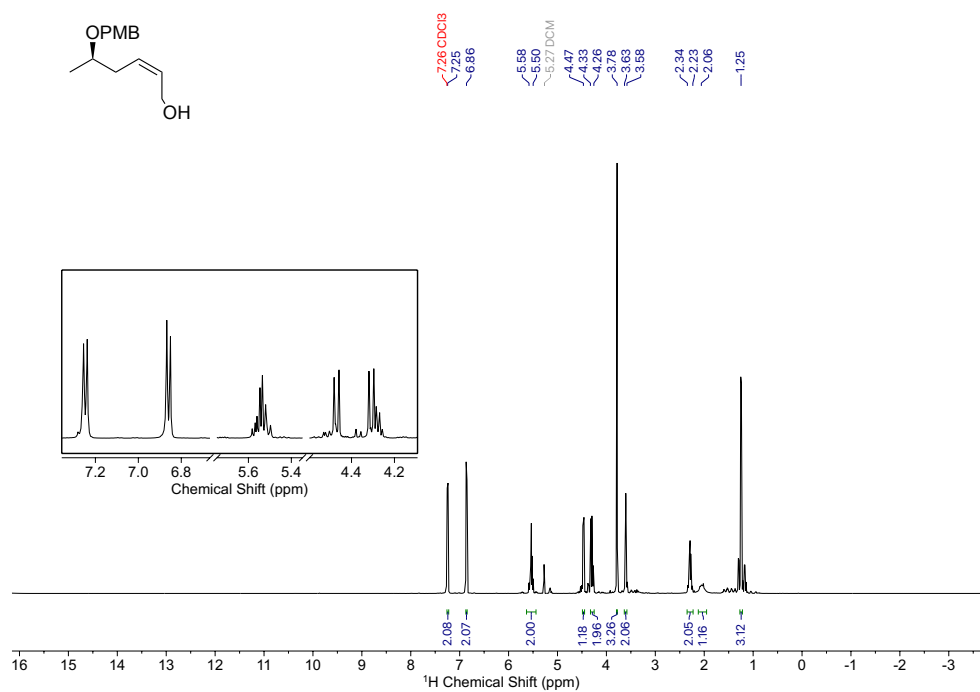


Figure B.15. ^1H NMR spectrum of **3.41** (500 MHz, CDCl_3).

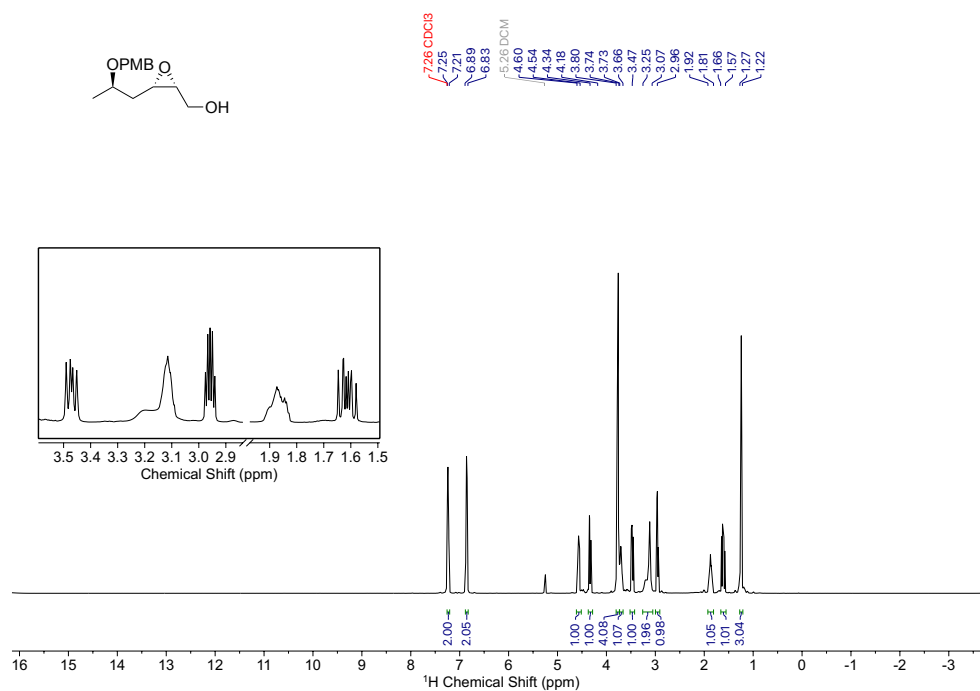


Figure B.16. ^1H NMR spectrum of **3.42** (500 MHz, CDCl_3).

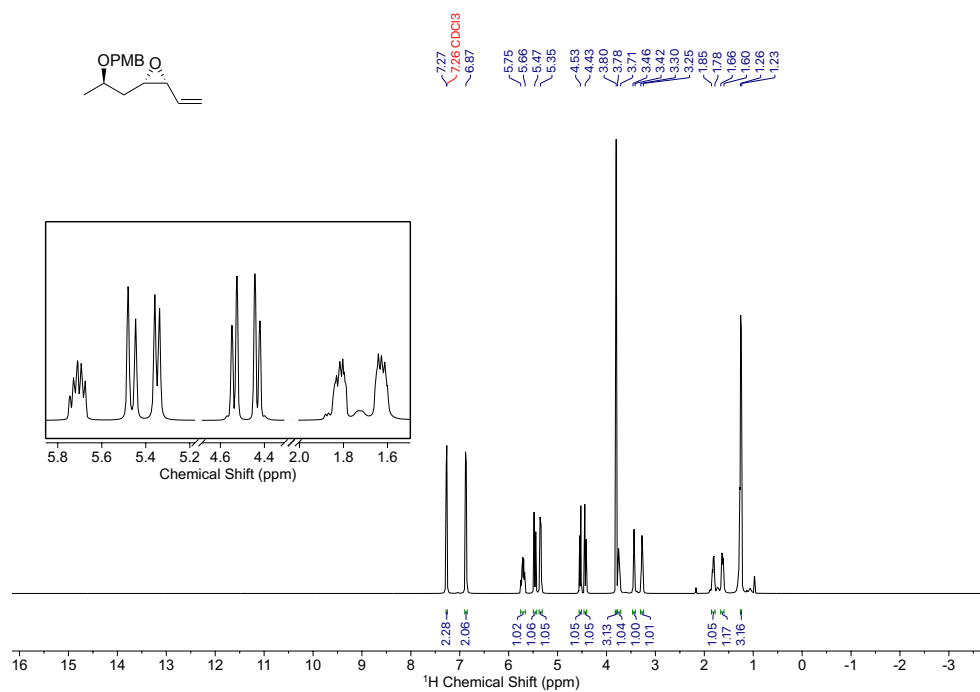


Figure B.17. ¹H NMR spectrum of **3.43** (500 MHz, CDCl₃).

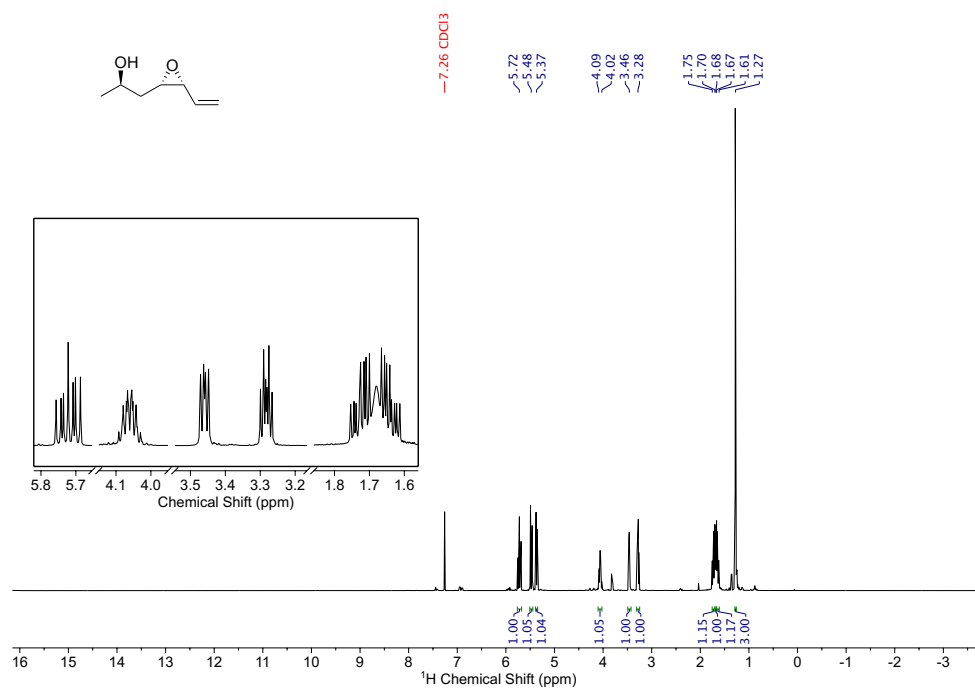


Figure B.18. ¹H NMR spectrum of **3.29** (500 MHz, CDCl₃).

Chemical structure of compound 10: CC(C)(C)C(=O)O[C@H]1C=C[C@H]1C=C

^1H NMR spectrum (400 MHz, CDCl_3) of compound 10. The spectrum shows peaks in the range 0 to 6 ppm. The inset shows the region from 1.7 to 5.9 ppm.

Peak labels (ppm):

- 5.83, 5.70, 5.46, 5.36, 5.21, 5.12, 5.03, 4.60, 4.53 (multiplet, 5H)
- 3.40, 3.14, 2.50, 2.31, 1.87, 1.79, 1.78, 1.72, 1.68, 1.66, 1.64, 1.62, 1.60, 1.58, 1.56, 1.54, 1.52, 1.50, 1.48, 1.46, 1.44, 1.42, 1.40, 1.38, 1.36, 1.34, 1.32, 1.30, 1.28, 1.26, 1.24, 1.22, 1.20, 1.18, 1.16, 1.14, 1.12, 1.10, 1.08, 1.06, 1.04, 1.02, 1.00, 0.98, 0.96, 0.94, 0.92, 0.90, 0.88, 0.86, 0.84, 0.82, 0.80, 0.78, 0.76, 0.74, 0.72, 0.70, 0.68, 0.66, 0.64, 0.62, 0.60, 0.58, 0.56, 0.54, 0.52, 0.50, 0.48, 0.46, 0.44, 0.42, 0.40, 0.38, 0.36, 0.34, 0.32, 0.30, 0.28, 0.26, 0.24, 0.22, 0.20, 0.18, 0.16, 0.14, 0.12, 0.10, 0.08, 0.06, 0.04, 0.02, 0.00 (multiplet, 1H)

Chemical structure of the compound is shown above the spectrum. The spectrum displays peaks corresponding to the chemical shifts (ppm) listed below:

Chemical Shifts (ppm): 170.6, 170.5, 140.4, 132.4, 120.8, 120.8, 114.9, 114.8, 77.2 (CDCl₃), 70.9, 70.9, 69.7, 69.7, 69.0, 56.9, 55.6, 55.6, 43.9, 34.4, 25.9, 20.4, 20.4, 18.2, 4.3, 4.9, 4.9.

13C Chemical Shift (ppm)

221

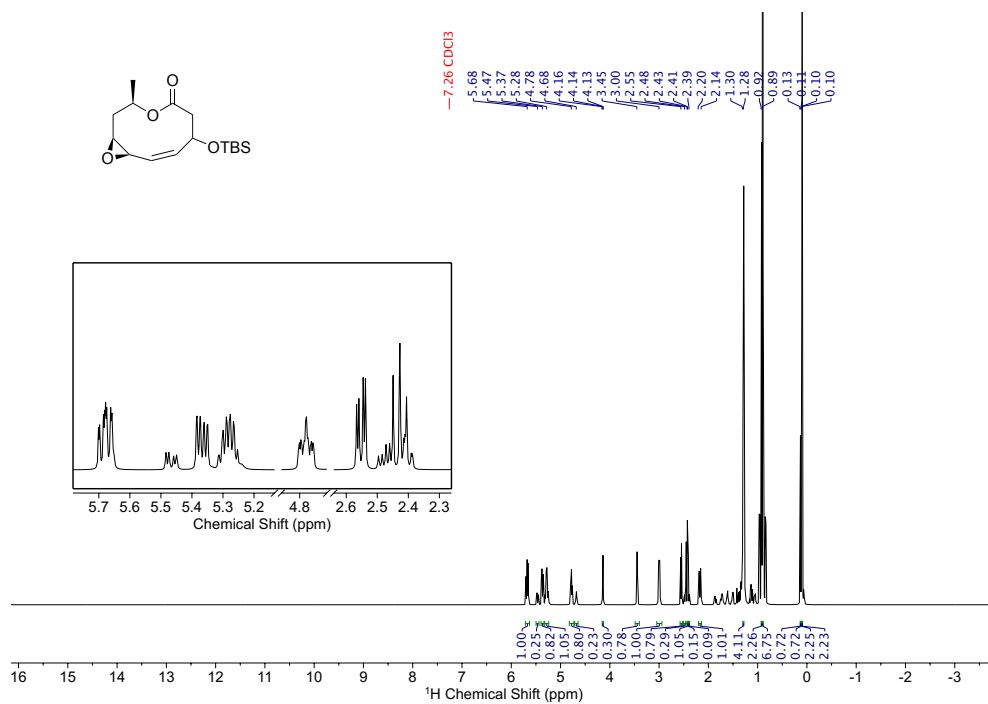


Figure B.21. ¹H NMR spectrum of **3.45** (500 MHz, CDCl₃).

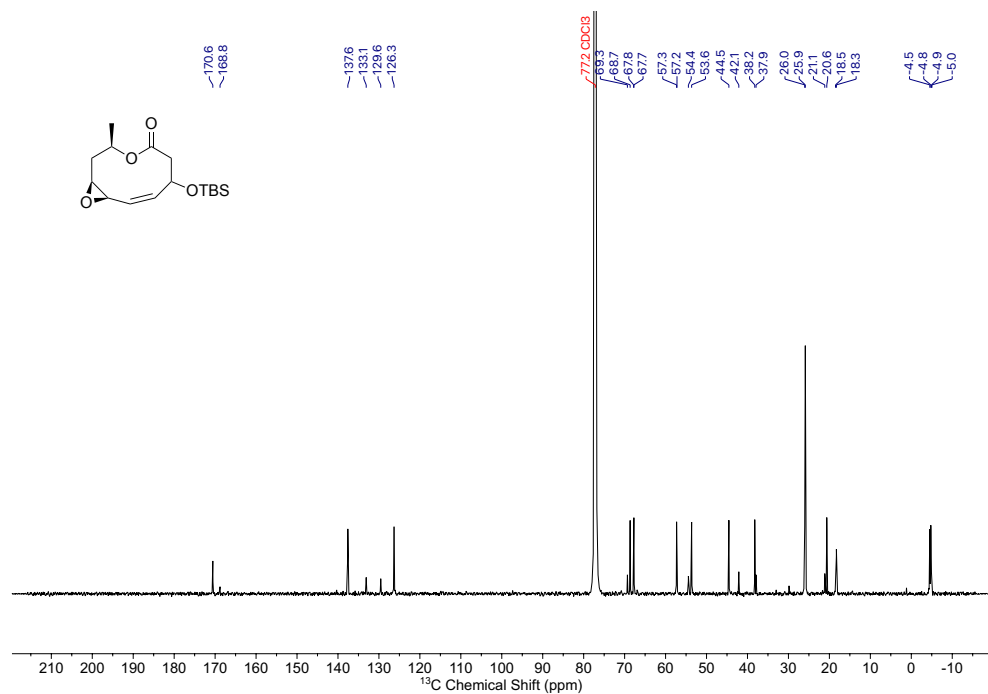


Figure B.22. ¹³C NMR spectrum of **3.45** (200 MHz, CDCl₃).

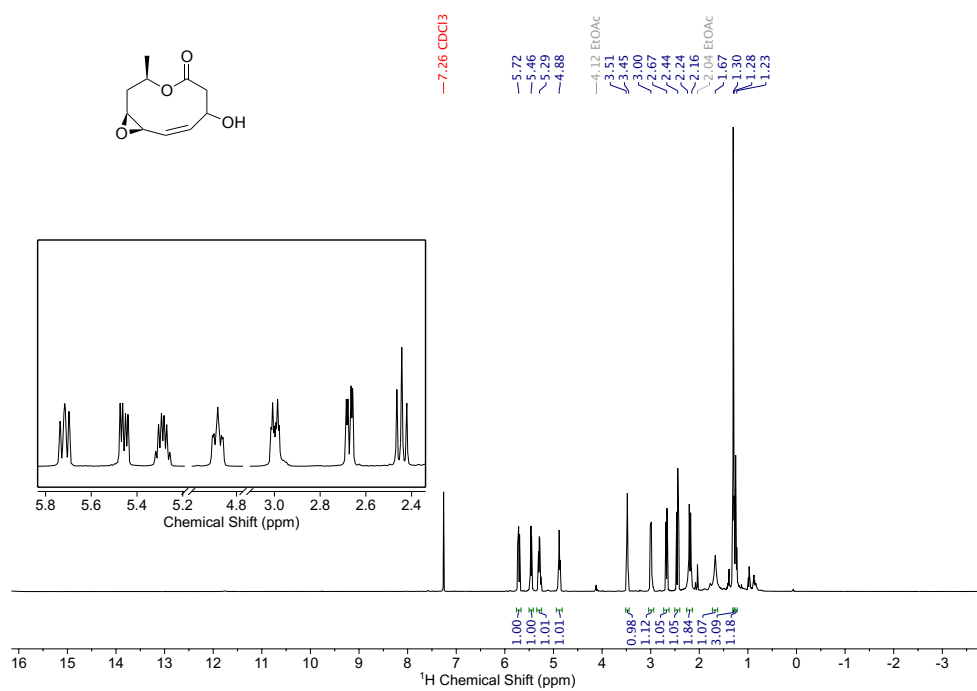


Figure B.23. ¹H NMR spectrum of botryolide B (3.9) (800 MHz, CDCl₃).

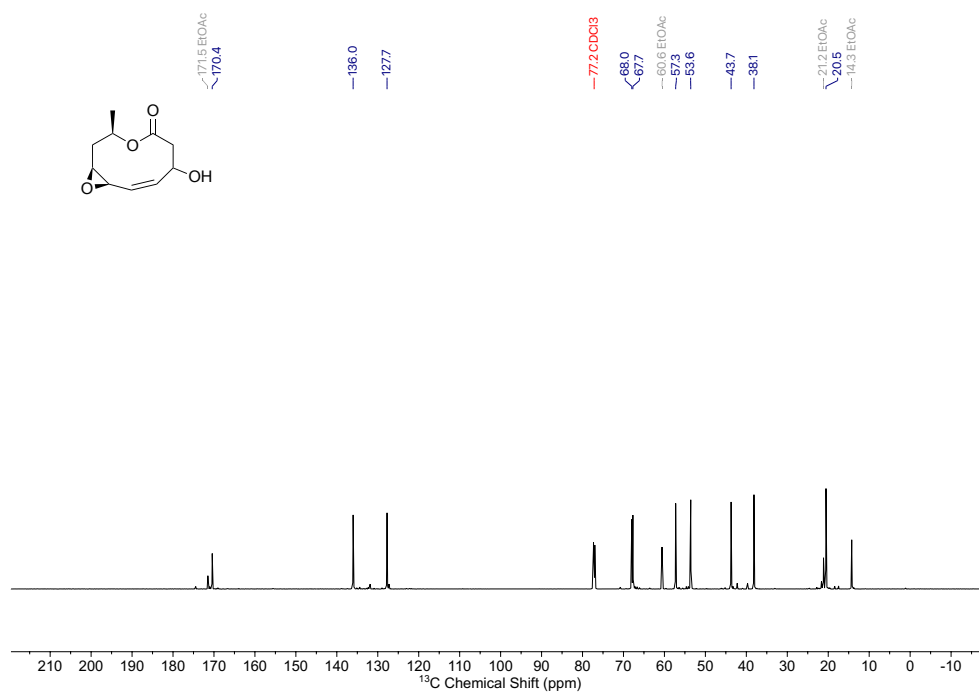


Figure B.24. ¹³C NMR spectrum of botryolide B (3.9) (200 MHz, CDCl₃).

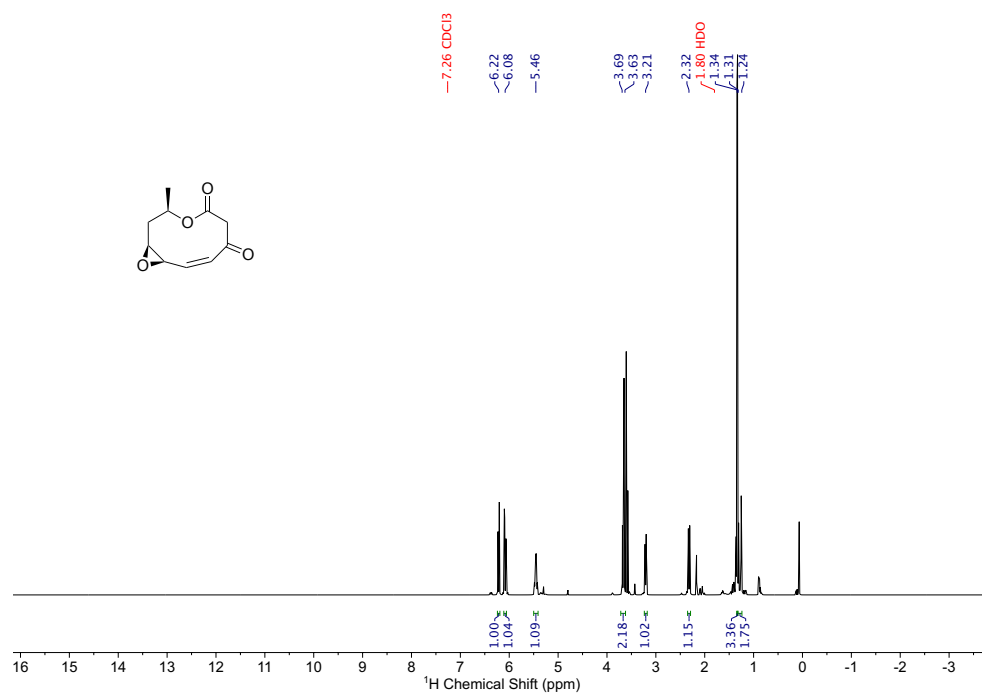


Figure B.25. ¹H NMR spectrum of protricilactone A (**3.8**) (800 MHz, CDCl₃).

Appendix C: Supplementary Experimental Data for Chapter 4

C.1 Spectral data for aborycin

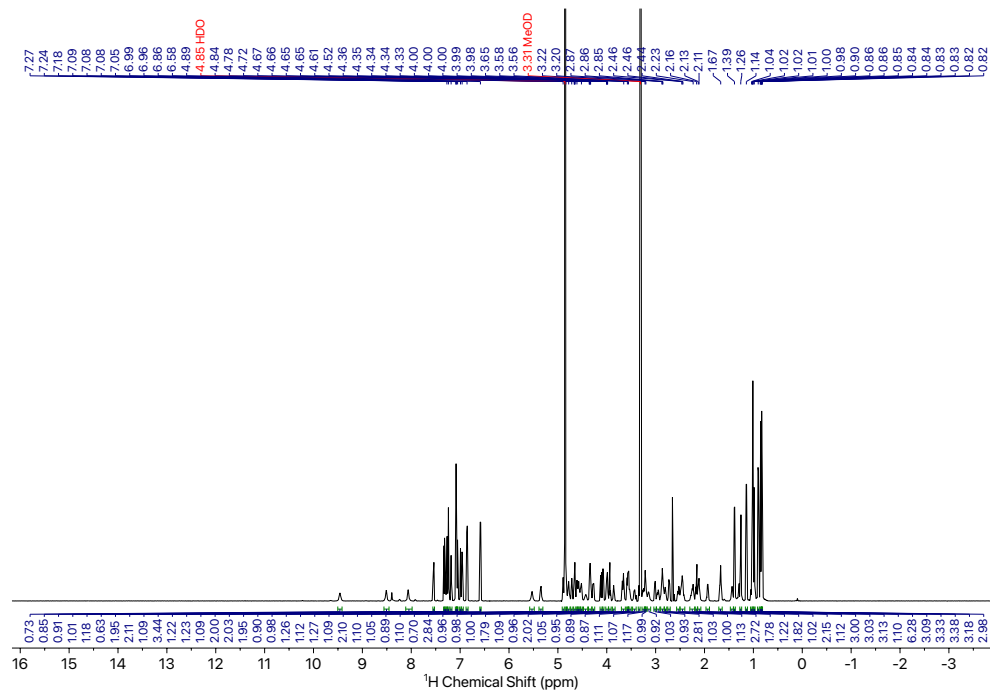


Figure C.1. Annotated ^1H NMR spectrum of aborycin (800 MHz, CD_3OD).

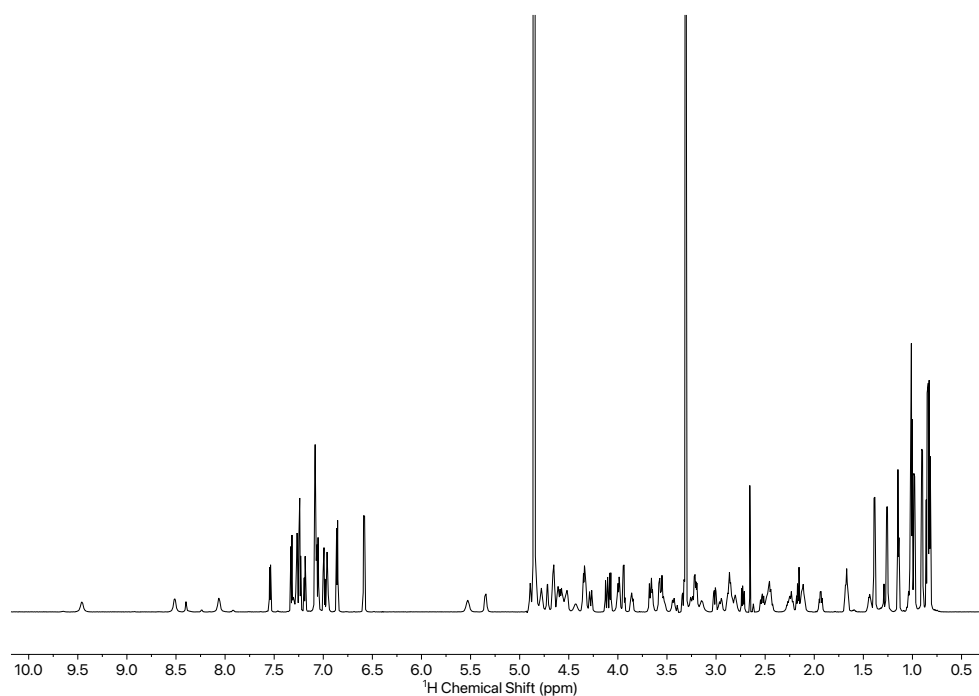


Figure C.2. Expanded ¹H NMR spectrum of aborycin without annotations (800 MHz, CD₃OD).

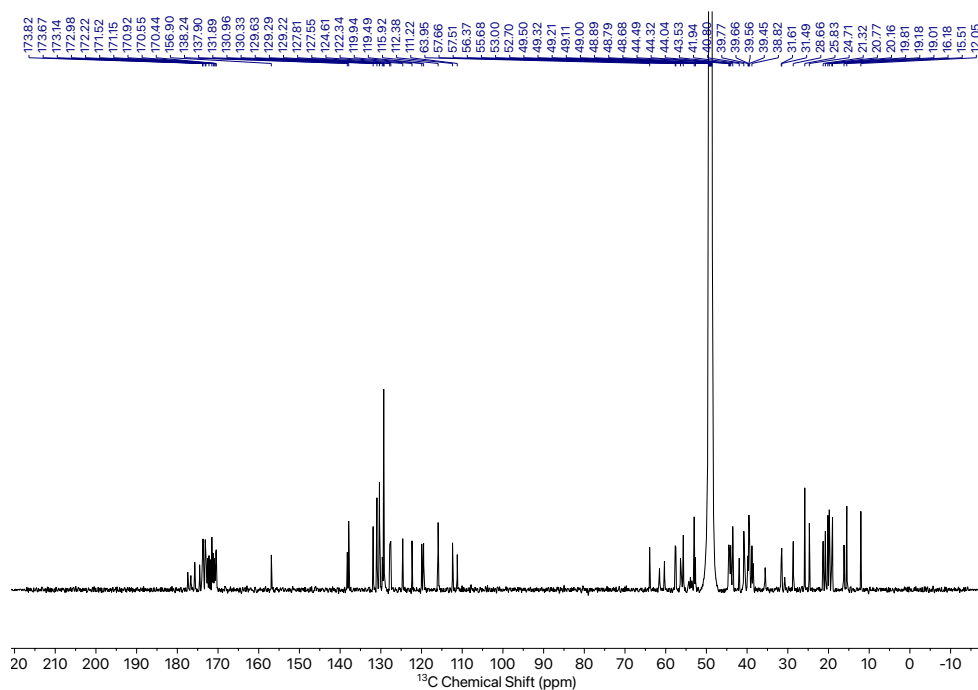


Figure C.3. Annotated ¹³C NMR spectrum of aborycin (200 MHz, CD₃OD).

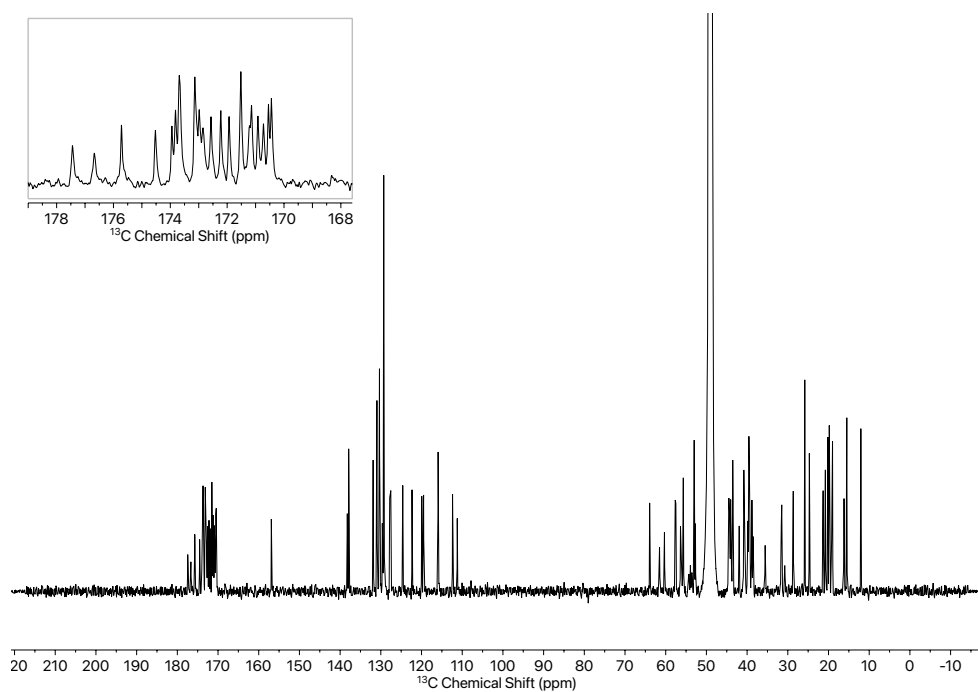


Figure C.4. Expanded ^{13}C NMR spectrum of aborycin and magnified image of the carbonyl region (200 MHz, CD_3OD).

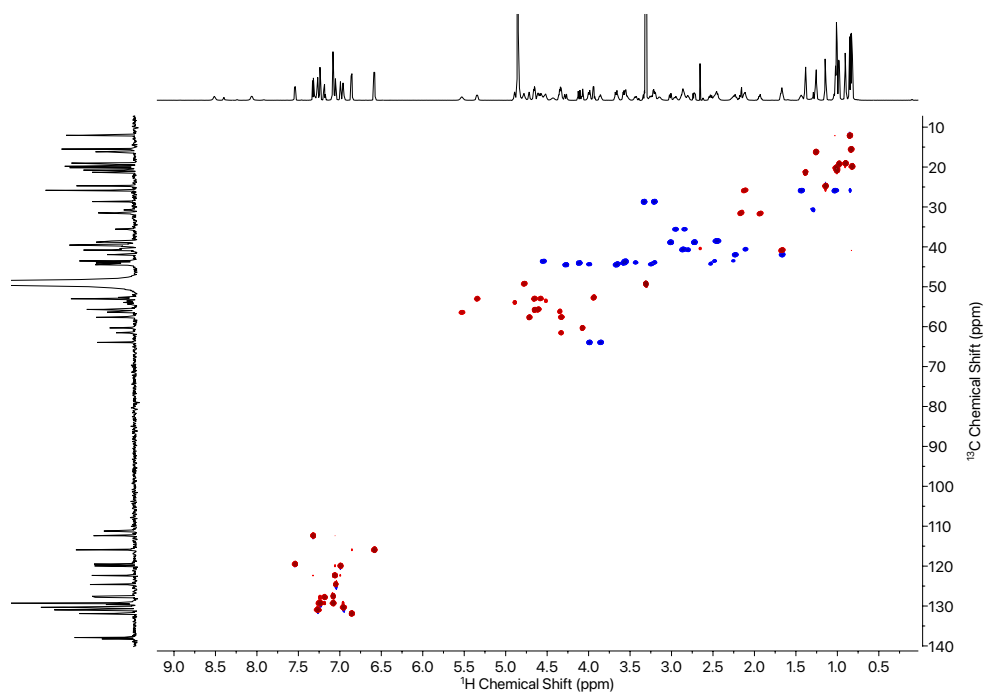


Figure C.5. ^1H - ^{13}C HSQC NMR spectrum of aborycin (800 MHz, CD_3OD).

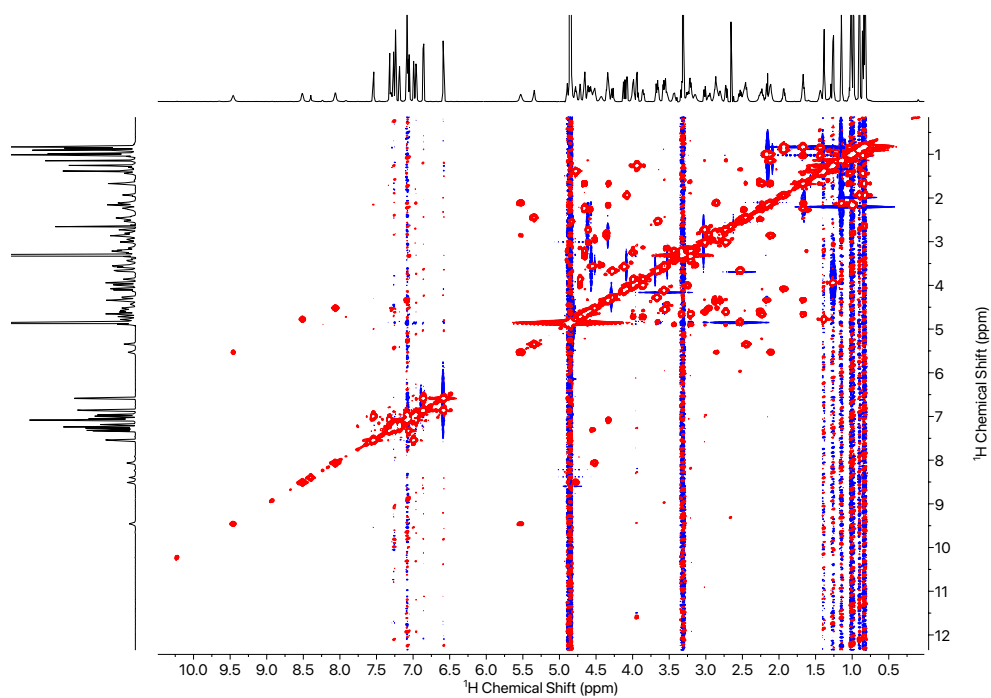


Figure C.6. COSY NMR spectrum of aborycin (800 MHz, CD_3OD).

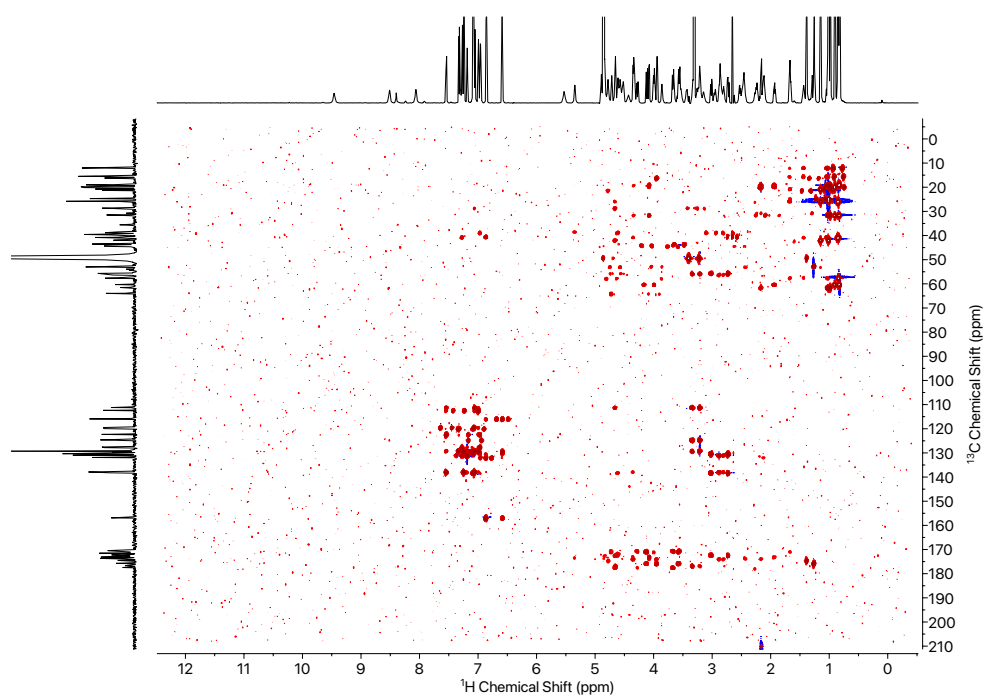


Figure C.7. Short-range (4 Hz) ^1H - ^{13}C HMBC spectrum of aborycin (800 MHz, CD_3OD).

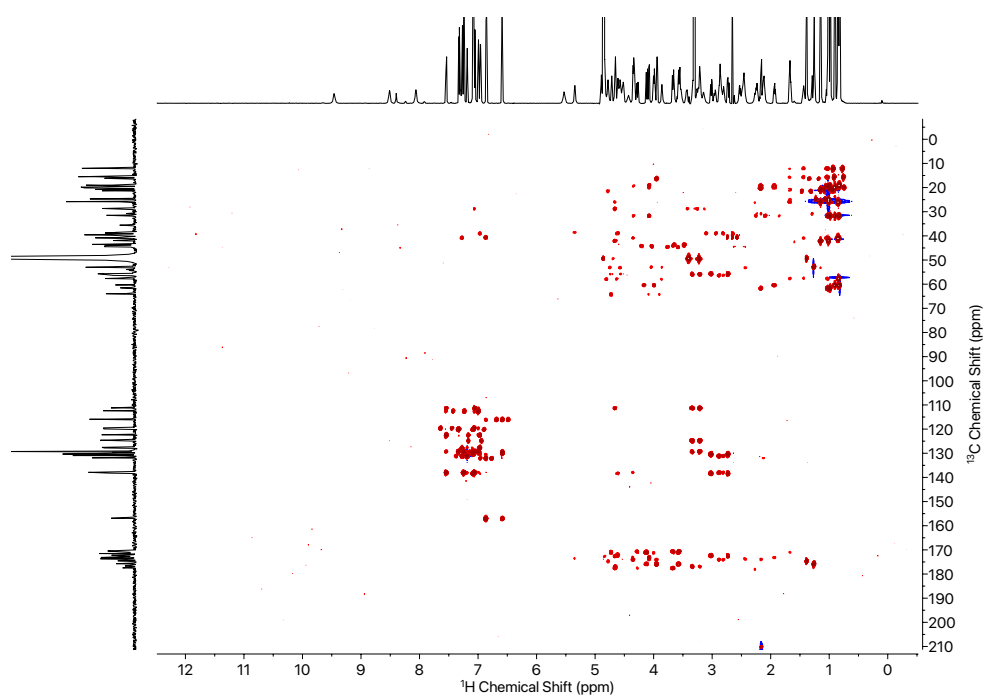


Figure C.8. Long-range (8 Hz) ^1H - ^{13}C HMBC spectrum of aborycin (800 MHz, CD_3OD).

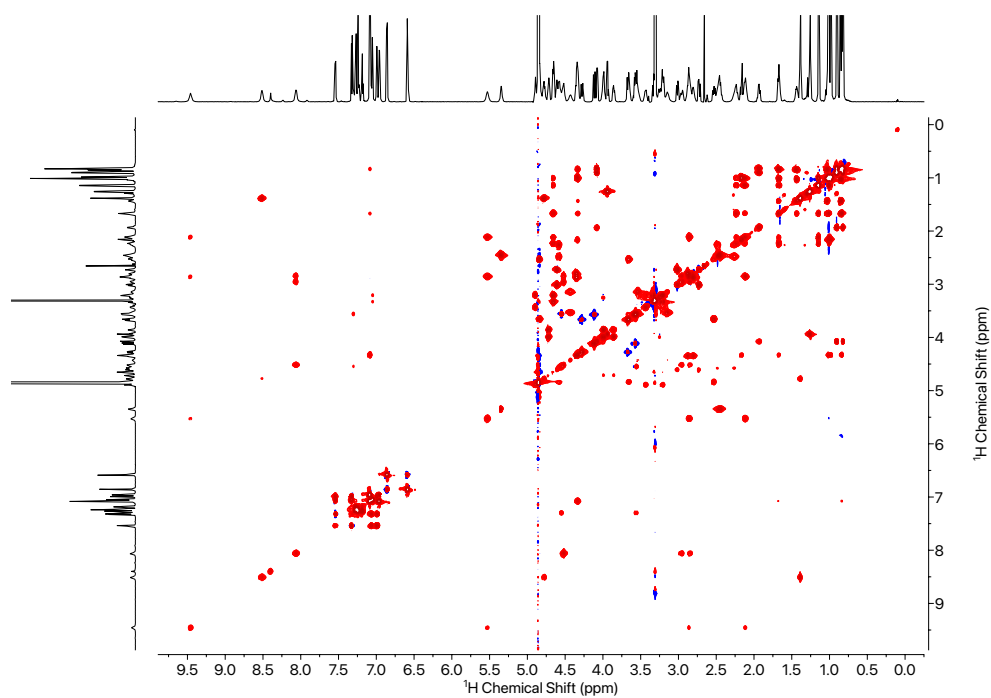


Figure C.9. TOCSY spectrum of aborycin (800 MHz, CD_3OD).

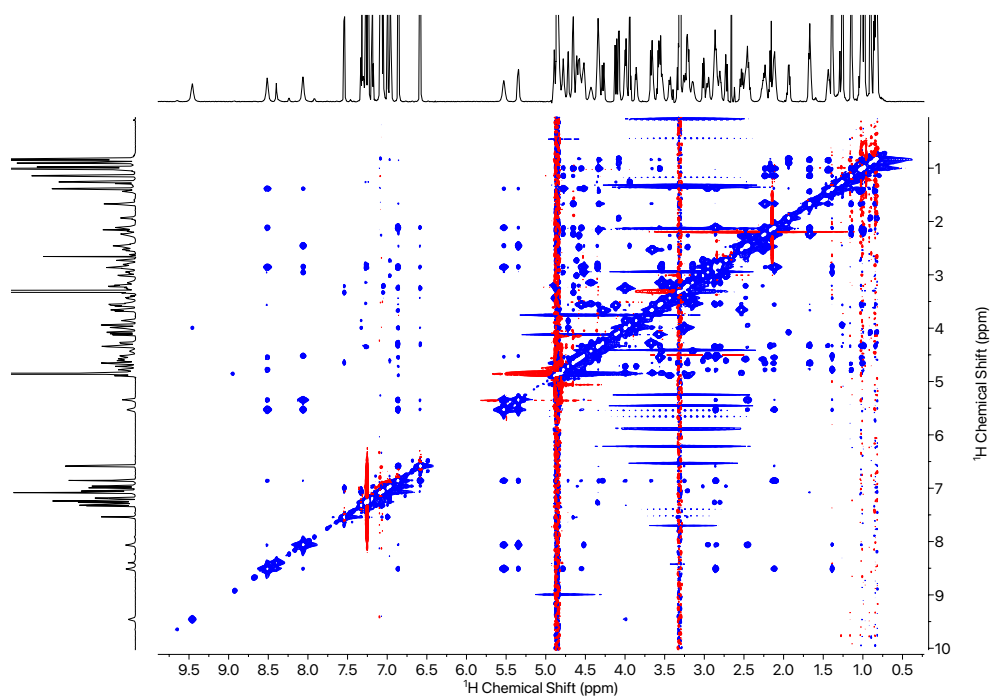


Figure C.10. NOSEY spectrum of aborycin (800 MHz, CD_3OD).

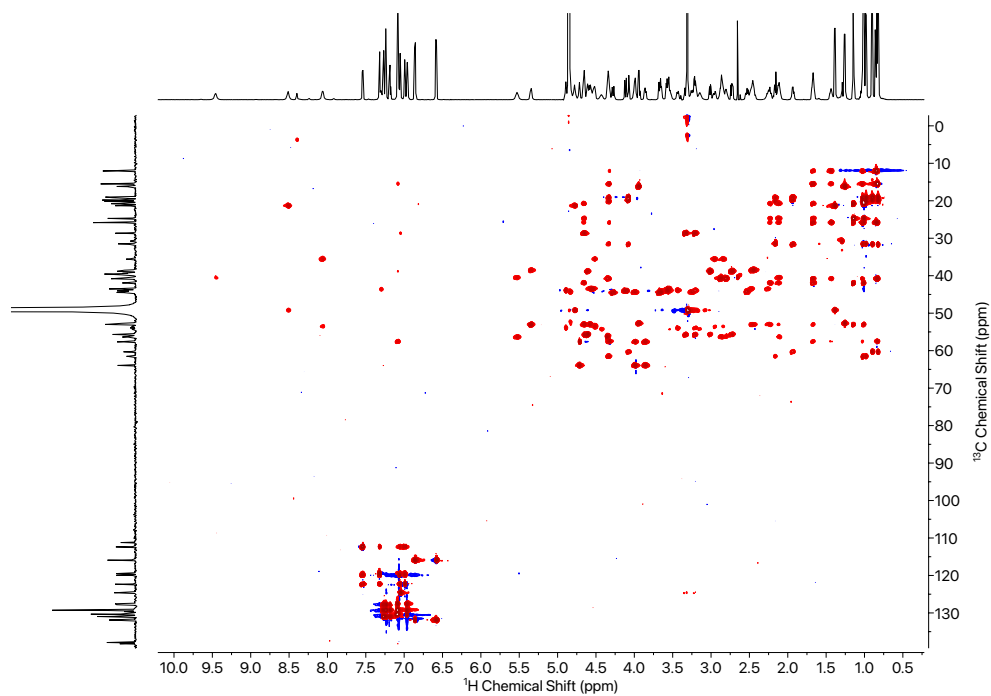


Figure C.11. HSQC-TOCSY spectrum of aborycin (800 MHz, CD_3OD).

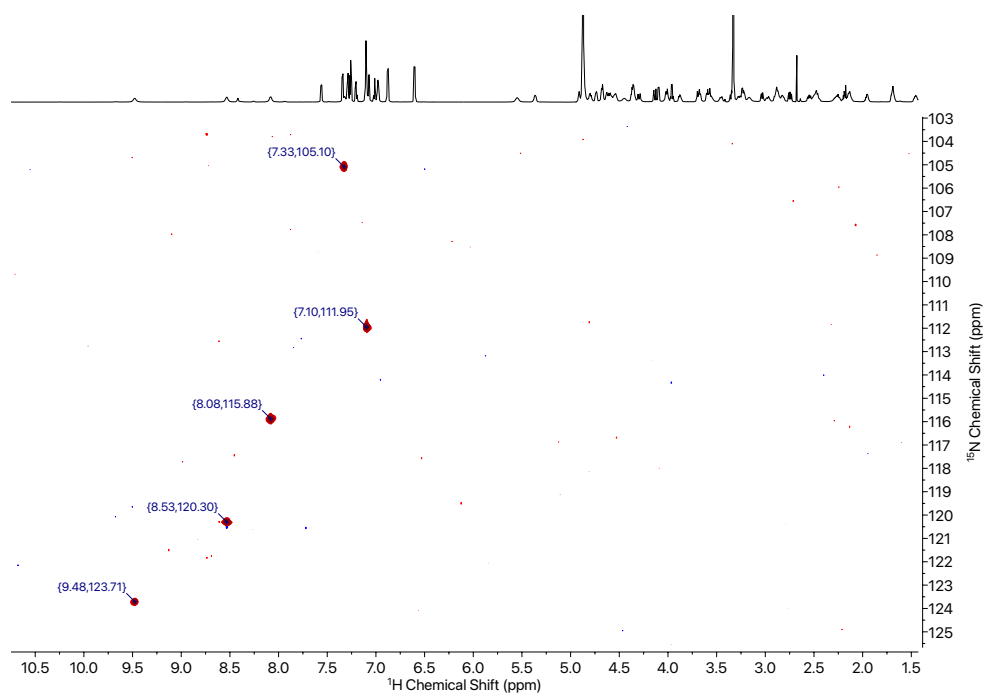


Figure C.12. ^1H - ^{15}N HSQC spectrum of aborycin (800 MHz, CD_3OD).

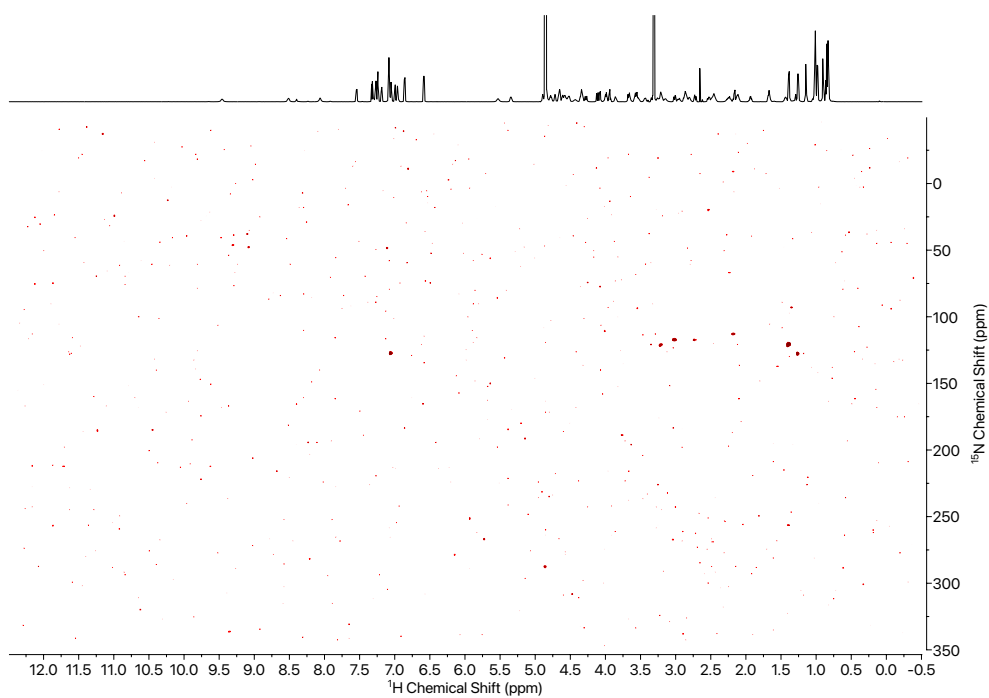


Figure C.13. ^1H - ^{15}N HMBC spectrum of aborycin (800 MHz, CD_3OD).

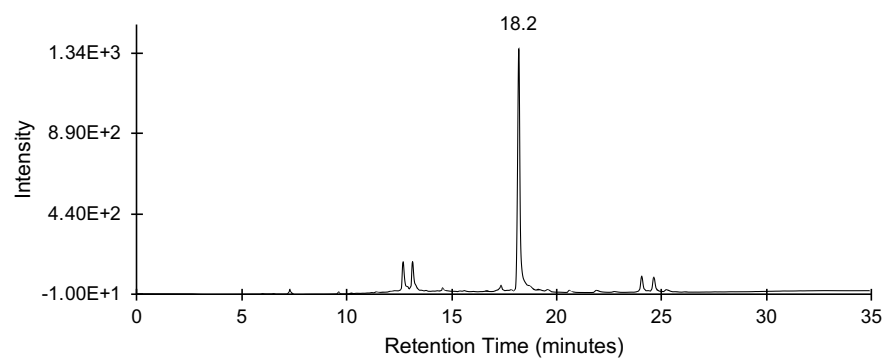


Figure C.14. HPLC chromatogram of aborycin at 280 nm.

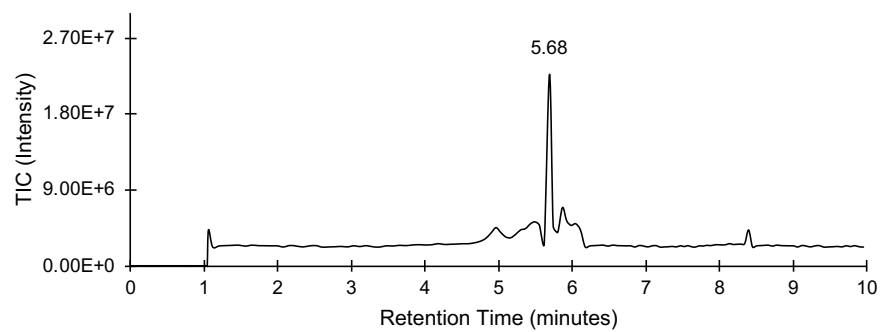


Figure C.15. TIC chromatogram of aborycin.

C.2. Aborycin-fluorophore spectral data

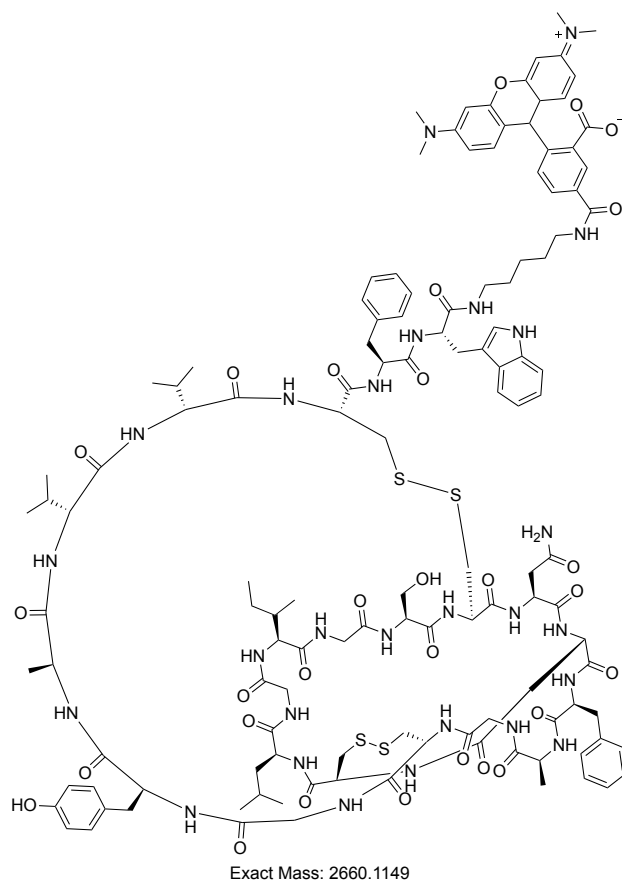


Figure C.16. Structure of aborycin with rhodamine cadaverine fluorescent tag.

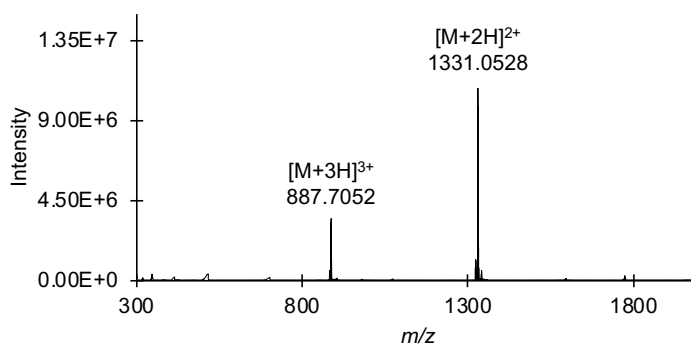


Figure C.17. Mass spectrum of aborycin-fluorophore showing $[M+3H]^{3+}$ and $[M+2H]^{2+}$ ions.

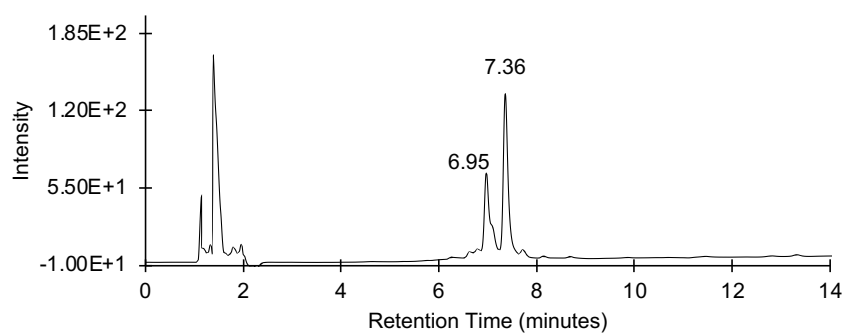


Figure C.18. HPLC chromatogram of aborycin-fluorophore at 280 nm. Rhodamine cadaverine was purchased as mixed isomers and each labeled signal represents one aborycin-fluorophore isomer.