Title
Synthetic Approaches to Indole Alkaloids and Biosynthetic Investigations of the Echinocandins and Chaetoviridins

Author
Chiou, Grace

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Synthetic Approaches to Indole Alkaloids and Biosynthetic Investigations of the Echinocandins and Chaetoviridins

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemistry

by

Grace Chiou

2014

ABSTRACT OF THE DISSERTATION

Synthetic Approaches to Indole Alkaloids and

Biosynthetic Investigations of the Echinocandins and Chaetoviridins

by

Grace Chiou

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2014

Professor Neil K. Garg, Chair

Chapter one discusses the previous use of synthetic chemistry in biosynthetic studies of natural products. Emphasis is given to more recent work concerning the brevianamide, notoamide, and communesin fungal indole alkaloids. Chapters two and three highlight our explorations of the interrupted Fischer indolization reaction to access indoline containing natural products. Specifically, chapter two focuses on the use of the interrupted Fischer indolization reaction to prepare the Alzheimer's therapeutic candidate, (+)-phenserine (a.k.a. Posiphen®), in enantioenriched form. Chapter three is a discussion of our studies to use the interrupted Fischer indolization reaction towards syntheses of the communesin alkaloids and perophoramidine.

Chapters four, five, and six present our use of chemical synthesis in the discussion of elucidating the biosynthetic pathways of natural products. In particular, chapter four focuses on the echinocandin lipopeptides and our efforts to identify and characterize the biochemical

pathway utilized by nature. Chapters five and six highlight our work to determine the biosynthesis of the chaetoviridin polyketides. In addition, a strategy for introducing structural diversity into polyketides is reported.

The dissertation of Grace Chiou is approved.

Yi Tang

Tatiana Segura

Neil K. Garg, Committee Chair

University of California, Los Angeles

2014

For my mother and sister, Shiang Chiou and Tiffany Nuckols

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

APPENDIX ONE

CHAPTER THREE

APPENDIX TWO

LIST OF SCHEMES

CHAPTER ONE

CHAPTER SIX

LIST OF TABLES

CHAPTER FOUR

LIST OF ABBREVIATIONS

$[\alpha]_D$	specific rotation at wavelength of sodium D line
Ac	acetyl, acetate
app.	apparent
aq.	aqueous
atm	atmosphere
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
br	broad
Bu	butyl
n-Bu	butyl
t-Bu	*tert*-Butyl
c	concentration for specific rotation measurements
°C	degrees Celsius
calc'd	calculated
CI	chemical ionization
d	doublet
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
dec	decomposition
DIEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
EC_{50}	50% effective concentration
ee	enantiomeric excess
equiv	equivalent
ESI	electrospray ionization
Et	ethyl
g	gram(s)
h	hour(s)
HMDS	hexamethyldisilylazide
HRMS	high resolution mass spectroscopy
HPLC	high performance liquid chromatography
hν	light

Hz	hertz
IC_{50}	half maximal inhibitory concentration
IMDA	intramolecular Diels–Alder
IR	infrared (spectroscopy)
J	coupling constant
λ	wavelength
L	liter
LAH	Lithium aluminum hydride
m	multiplet or milli
m	meta
m/*z*	mass to charge ratio
μ	micro
μM	micromolar
Me	methyl
MHz	megahertz
MIC	minimal inhibitory concentration
min	minute(s)
mol	mole(s)
mp	melting point
MS	molecular sieves
NaHMDS	Sodium hexamethyldisilazide
NBS	*N*-bromosuccinimide
NMO	*N*-methylmorpholine *N*-oxide
NMR	nuclear magnetic resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Enhancement Spectroscopy
nNOS	neural nitric oxide synthase
Ns	Nosyl
[O]	oxidation
o-tol	ortho tolyl
p	para
PCC	pyridinium chlorochromate
Ph	phenyl
pH	hydrogen ion concentration in aqueous solution
PhH	benzene

ppm	parts per million
PPTS	pyridinium *p*-toluenesulfonate
Pr	propyl
i-Pr	isopropyl
pyr	pyridine
q	quartet
RED-Al	sodium bis(2-methoxyethoxy)aluminumhydride
rt	room temperature
R_f	retention factor
s	singlet or strong
t	triplet
TBAF	tetrabutylammonium fluoride
TBS	*tert*-butyldimethylsilyl
Tf	trifluoromethanesulfonyl (trifyl)
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMS	trimethylsilyl
Ts	*p*-toluenesulfonyl (tosyl)
UV	ultraviolet
w	weak

ACKNOWLEDGEMENTS

Although the following dissertation is an individual work, it would have never reached this point without the help, guidance, and support from a number of people. I want to first thank my research advisor, Professor Neil K. Garg. I am so fortunate to have followed my instinct during my visitation at UCLA to join his research group. There are no words to express my gratitude for his unwavering support and guidance. I can truly say I would not be the person I am today without his mentorship.

Along with my research advisor, I would like to thank the rest of the chemistry department at UCLA. Professor Yi Tang, in particular, has made a tremendous impact on my life by allowing a synthetic chemist, like myself, to collaborate with his group. Our research collaboration is a constant reminder of why I became a chemist. In addition, I would like to thank Professor Miguel Garcia-Garibay for including me in the Organization for Cultural Diversity in Chemistry (OCDC) at UCLA. OCDC's supportive environment contributed to what I consider a unique graduate education. Also, I have been very fortunate to have Professor Tatiana Segura, and Professor Richard Weiss on my committee.

In company with these individuals, I have also been lucky to have the continued support of colleagues from my undergraduates studies. I am forever grateful to Professor Michael Doyle, "Dr. D." He took a chance by taking me into his lab and giving me my early and wonderfully fantastic experiences in research. His unlimited zeal encouraged me to follow my passion. Without his mentorship over the years, this thesis would not be in existence. I want to thank Dr. Emily McLaughlin and Dr. Hojae Choi, for being great mentors during my time in the Doyle Research Group as well.

The members of the Garg group have contributed immensely to my personal and professional time at UCLA. The group has been a source of friendships, and I am especially grateful to Tejas Shah, Tehetena Mesganaw, and Joel Smith. My relationships with these individuals have given me more than words can describe, and I hope that they know they mean the world to me. Even if we drift apart, I will always hold a special place for them.

I am also grateful to every member of the Garg lab that I have had the honor to work with, past and present. Alex Schammel was the first person I got to work with and his friendship, camaraderie, resilience, and pretty much everything about him, will stay with me. I am grateful for the fantastic Garg group members who traveled on this journey with me: Adam Goetz, Amanda Silberstein, and Stephen Ramgren. I want to say thank you to Ben Boal for teaching me what a real hipster is, Kyle Quasdorf for letting me pour Coke in his hair, and Sarah Bronner for her love of cats. This experience taught me so much, and I feel fortunate to have met every one of you. Other members include Noah Fine Nathel, Jesus Moreno, Mike Corsello, Josh Demeo, Laura Shim, Dr. Travis McMahon, Emma Baker, and the numerous visiting students who have come through the lab.

Last but not least, I want to acknowledge my family and friends. I am so blessed to have an amazing mother and sister, whose unwavering support during the ups and downs of graduate school has been incredible. To my nephews, Billy Nuckols and Jimmy Nuckols, I'm sorry I was not around for those first few years. Hopefully if you are reading this, I've made up for it by spoiling you rotten. I have also been blessed with a special group of childhood friends: Gillian Horn, Sloane Hickey, Julia Dennis, and Madeleine Rubenstein – you are all my forever friends. Thank you for putting up with me over the past five years. Without these girls and my family, none of this work would have been possible. Thank you. Seriously.

Chapter 2 is a version of Schammel, A. W.; Chiou, G.; Garg, N. K. *J. Org. Chem.* **2012**, *77*, 725–728. Schammel and Chiou were responsible for experimental work.

Chapter 3 is a version of Schammel, A. W.; Chiou, G.; Garg, N. K. *Org. Lett.* **2012**, *14*, 4556–4559. Schammel and Chiou were responsible for experimental work.

Chapter 4 is a version of Jiang, W.; Cacho, R.; Chiou, G.; Garg, N. K.; Tang, Y.; Walsh, C. T. *J. Am. Chem. Soc.* **2013**, *135*, 4457–4466. Jiang, Cacho, Chiou were responsible for experimental work.

Chapter 5 is a version of Winter, J. M.; Sato, M.; Sugimoto, S.; Chiou, G.; Garg, N. K.; Tang, Y.; Watanabe, K. *J. Am. Chem. Soc.* **2012**, *134*, 17900–17903. Winter, Sato, Sugimoto, Chiou were responsible for experimental work.

Chapter 6 is a version of Winter, J. M.; Chiou, G.; Bothwell, I. R.; Xu, W.; Garg, N. K.; Luo, M.; Tang, Y. *Org. Lett.* **2013**, *15*, 3774–3777. Winter, Chiou, Bothwell, Xu were responsible for experimental work.

BIOGRAPHICAL SKETCH

Education:

University of Maryland, College Park, MD

- Bachelor of Science, Biochemistry with Honors, December 2008
- Cumulative GPA: 3.6/ 4.0

Professional and Academic Experience:

Graduate Research Assistant: University of California, Los Angeles, CA

- July 2009 to present, studying under the direction of Professor Neil K. Garg
- Developed an efficient synthesis of the Alzheimer's therapeutic, Posiphen®
- Constructed the complex tetracyclic indoline scaffold of perophoramidine and the communesin alkaloids
- Investigated the biosynthetic pathway of communesin alkaloid natural products in collaboration with Professor Yi Tang (UCLA)
- Synthesized natural and unnatural amino acids to probe the biosynthesis of echinocandin B, a potent antifungal agent, in collaboration with Professor Yi Tang (UCLA) and Professor Christopher Walsh (Harvard)
- Prepared α-substituted pyrones to elucidate the gene cluster of the natural products chaetoviridin A (antifungal) and chaetomugilin A (cytotoxic) in collaboration with Professor Yi Tang (UCLA) and Professor Kenji Watanabe (University of Shizuoka)

Teaching Assistant: University of California, Los Angeles, CA

- Undergraduate organic chemistry
  - September 2009 to June 2010
  - Supervised and taught discussion classes of >30 students

Undergraduate Research Assistant: University of Maryland, College Park, MD

- January 2007 to December 2008
- Developed new *tert*-butyl hydroperoxide transformations using dirhodium catalysis under the guidance of Professor Michael P. Doyle

Undergraduate Research Assistant: University of Maryland, College Park, MD

- September 2006 to December 2006
- Investigated *Mycobacterium marinum* pathogenesis with Professor Lian-Yong Gao

Teaching Assistant: University of Maryland, College Park, MD

- Undergraduate organic chemistry laboratory
  - September 2006 to June 2007
  - Supervised and taught laboratory classes of >30 students

Honors and Awards:

- National Science Foundation Graduate Research Fellowship – 2010–2013
- Frank and Sarah McKnight Prize in Chemistry, Semifinalist, University of Texas, Southwestern – 2008
- University of Maryland, College Park – Alpha Chi Sigma Award – 2008
- Beckman Scholars Award Recipient – May 2007

Publications:

7. **Expanding the Structural Diversity of Polyketides by Exploring the Cofactor Tolerance of an Inline Methyltransferase Domain.** Winter, J. M.; Chiou, G.; Bothwell, I. R.; Xu, W.; Garg, N. K.; Luo, M.; Tang, Y. *Org. Lett.* **2013**, *15*, 3774–3777.

6. **EcdGHK are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold.** Jiang, W.; Cacho, R.; Chiou, G.; Garg, N. K.; Tang, Y.; Walsh, C. T. *J. Am. Chem. Soc.* **2013**, *135*, 4457–4466.

5. **Identification and Characterization of the Chaetoviridin and Chaetomugilin Gene Cluster in Chatomium globosum Reveal Dual Functions of an Iterative Highly-Reducing Polyketide Synthase.** Winter, J. M.; Sato, M.; Sugimoto, S.; Chiou, G.; Garg, N. K.; Tang, Y.; Watanabe, K. *J. Am. Chem. Soc.* **2012**, *134*, 17900–17903.

4. **Interrupted Fischer Indolization Approach Toward the Communesin Alkaloids and Perophoramidine.** Schammel, A. W.; Chiou, G.; Garg, N. K. *Org. Lett.* **2012**, *14*, 4556–4559.

3. **Synthesis of (+)-Phenserine Using an Interrupted Fischer Indolization Strategy.** Schammel, A. W.; Chiou, G.; Garg, N. K. *J. Org. Chem.* **2012**, *77*, 725–728.

2. **Dirhodium-Catalyzed Phenol and Aniline Oxidations with T-HYDRO. Substrate Scope and Mechanism of Oxidation.** Ratnikov, M. O.; Farkas, L. E.; McLaughlin, E. C.; Chiou, G.; Choi, H.; El-Khalafy, S. H., Doyle, M. P. *J. Org. Chem.* **2011**, *76*, 2585–2593.

1. **Allylic Oxidations Catalyzed by Dirhodium Caprolactamate via Aqueous *tert*-Butyl Hydroperoxide.** McLaughlin, E. C.; Choi, H.; Wang, K.; Chiou, G.; Doyle, M. P. *J. Org. Chem*. **2009**, *74*, 730–738.

Presentations:

8. **A Synergistic Approach to the Communesin Alkaloids.** *Seaborg Symposium,* Los Angeles, CA, October 26, 2013.

7. **Synthetic Studies Pertaining to the Biosynthesis of Natural Products.** *ACS National Meeting*, Indianapolis, IN, September 2013 (oral).

6. **Progress Towards the Total Synthesis of Communesin F and Perophoramidine.** *ACS National Meeting*, San Diego, CA, March 2012 (poster).

5. **Progress Towards the Total Synthesis of Communesin F.** *ACS National Meeting*, Anaheim, CA, March 2011 (poster).

4. **Total Syntheses of Nakadomarin A.** University of California, Second Year Organic Student Seminar Series, Los Angeles, CA. January 2011.

3. **Dirhodium Caprolactamate in the Peroxidation of Phenols.** *ACS National Meeting*, Philadelphia, PA, August 2008 (poster).

2. **Oxidations Catalyzed by Aqueous *tert*-Butyl Hydroperoxide By Dirhodium Caprolactamate.** *Beckman Scholars Symposium*, Irvine, CA, July 2008 (invited speaker).

1. **Allylic Oxidations by Aqueous *tert*-Butyl Hydroperoxide Catalyzed by Dirhodium Caprolactamate.** *ACS National Meeting*, New Orleans, LA, March 2008 (poster).

Outreach Activities:

Organization for Cultural Diversity in Chemistry - University of California, Los Angeles

- Member from 2010 to present
- Aided in the organization of events, including Bridges Day and LATTC Exchange
- Served as a graduate mentor to promote undergraduate research for current UCLA undergraduates afflicted with socioemotional problems in collaboration with the UCLA Program for Education and Enrichment of Relational Skills (PEERS®)

Intel Science and Engineering Fair - LA Convention Center

- UCLA/Garg Group coordinator and liaison – May 12, 2011
- Organized chemistry demonstrations that were continuously shown throughout the day to >2000 children in L.A. County

L.A. County Science Fair - Pasadena Convention Center

- Garg Group liaison and judge – April 15, 2011
- Recruited UCLA graduate students to serve as judges for the L.A. County Science Fair

Alpha Chi Sigma - University of Maryland, College Park, MD

- Member from December 2005 to present
- Vice President from September 2006 – June 2007; organized recruitment events and weekly meetings during Vice Presidency term to enlist new members
- Annual volunteer at Maryland Day (2006–2008)

- Participated in Boy Scouts of America program: helped boy scouts earn their chemistry merit badge each year by performing demonstrations and providing assessments

CHAPTER ONE

Synthetic Studies Pertaining to the Biosynthesis of Fungal Indole Alkaloids

1.1 Abstract

This review describes recent biosynthetic studies that rely on the use of synthetic organic chemistry. Specifically, we highlight work pertaining to fungal indole alkaloids within the brevianamide, notoamide, and communesin natural products. Synthetic chemistry has played an integral role in probing biosynthetic pathways, and we hope to demonstrate its importance in understanding how nature creates complex natural products.

1.2 Introduction

Within the field of natural product research, fungal indole alkaloids continue to inspire developments in drug discovery.[1] Fungal-derived natural products often demonstrate architectural complexity and a broad range of biological activities (e.g., **1.1**–**1.2**, Figure 1.1). Furthermore, the presence of the indole motif itself is considered a "privileged structure" in drug discovery.[2] With the growing need to discover new secondary metabolites for drug discovery, there has been a great interest to investigate fungal indole-containing natural products.

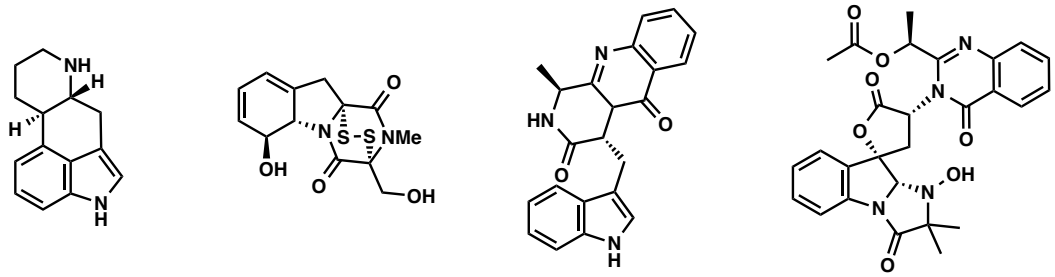


Figure 1.1. Representative indole alkaloids from fungi.

In recent years, there have been remarkable advances in understanding the biosynthesis of natural products as the distinction between classic total synthesis and biosynthesis has become less clear.[3] Interdisciplinary approaches are beginning to bridge the gap between total synthesis and molecular biology. Synthetic chemistry is an essential tool in biosynthesis in many ways including, but not limited to; a) providing isotopically labeled intermediates to be used in precursor feeding studies, b) providing authentic synthetic samples of potential intermediates, and c) verifying natural product structure through total synthesis for informed biosynthetic pathways to be proposed. In a complimentary sense, biosynthesis can inspire biomimetic total syntheses of natural products. This review provides a summarized account of the pivotal role of synthetic chemistry in understanding the biosynthetic pathways of fungal indole alkaloids; in particular, the brevianamides, notoamides, and communesin alkaloids are discussed.

1.3 Brevianamides

The brevianamides, A–F, (**1.5**–**1.10**, Figure 1.2) are a class of biologically active indole alkaloids. The more complex family members feature a bicyclo[2.2.2]diazo-octane ring system, which is also common to the stephacidins,[4,5,6] paraherquamides,[7,8] and asperparalines.[9,10] The

brevianamide family of fungal metabolites was initially isolated from *Penicillium brevicompactum* over thirty years ago.[39] Since then, their structural complexity and notable bioactivity have intrigued many in the field of synthetic chemistry and natural products biosynthesis.[11]

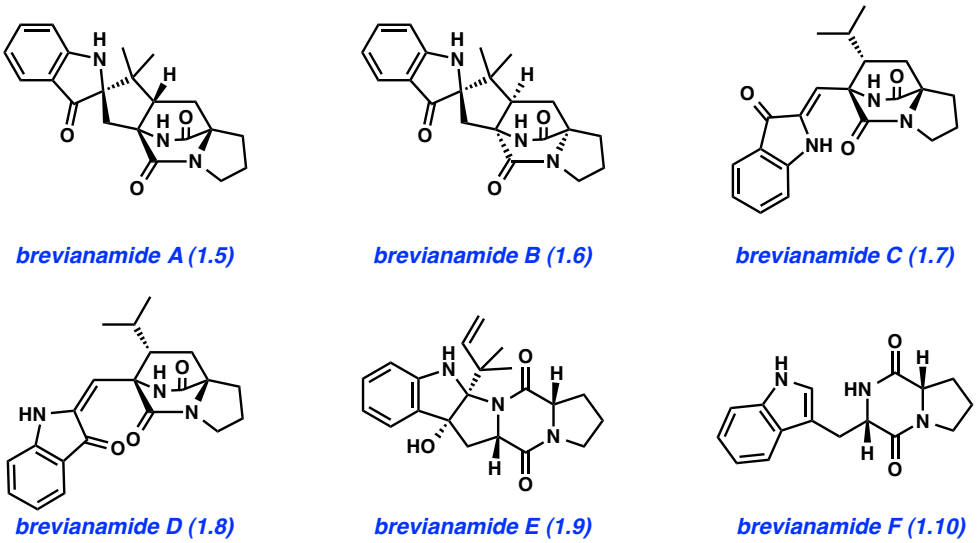


Figure 1.2. The brevianamide family of natural products.

In terms of their therapeutic potential, both natural and post-synthetic derivatives of brevianamides exhibit a wide range of biological activity. Brevianamide A (**1.5**) and D (**1.8**) have been shown to possess modest insecticidal activity.[12] Meanwhile, brevianamide F (**1.10**) possesses antibacterial and antifungal activity, and has shown potential for the treatment of cardiovascular dysfunction[13] and in cognitive enhancement.[14] With their intriguing architectural complexity and biological activity, the brevianamides have piqued the interest of multiple groups over the years.[15,16,17,18] Highlights of brevianamide biosynthetic studies that rely on synthetic chemistry have been directed towards determination of starting materials and intermediates in the

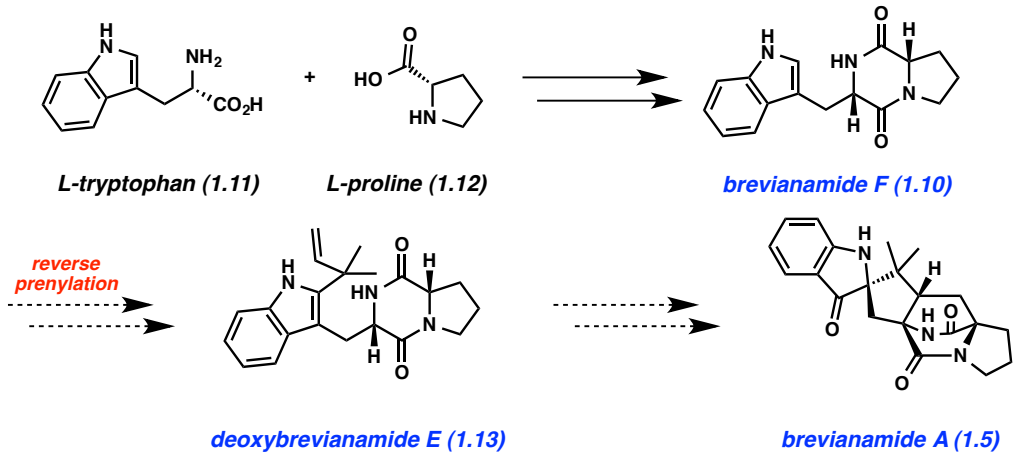
pathway, stereochemical assignments, and studies involving formation of the bicyclo[2.2.2]diazo-octane ring system, all of which will be summarized below.

1.3.1 Biosynthetic Studies of the Brevianamides

1.3.2 Birch's Proposed Biosynthetic Pathway

The first biosynthetic proposal for the brevianamides, by Birch in 1970, is summarized in Scheme 1.1.[19] This early work utilized isotopically labeled small molecules in feeding experiments. Whereas isotopically labeled tryptophan (**1.11**) and proline (**1.12**) were incorporated into brevianamide A (**1.5**), incorporation of isotopically labeled mevalonate and acetate was not observed. Therefore, this provided evidence that the building blocks of the brevianamides are tryptophan (**1.11**) and proline (**1.12**). Later work by Birch in 1974, revealed that brevianamide F (**1.10**) is a biosynthetic precursor to brevianamide A (**1.5**).[20,21] It was hypothesized in this work that brevianamide F (**1.10**) could undergo a reverse prenylation to afford deoxybrevianamide E (**1.15**), which could subsequently be elaborated into brevianamide A (**1.5**); however, deoxybrevianamide E (**1.13**) was not validated as an intermediate until later studies by the Williams group.[22,23] Although determination of the initial building blocks was a breakthrough, additional studies were necessary to understand the chemical transformations involved in brevianamide biosynthesis.

Scheme 1.1. Early proposal of the biosynthetic pathway by Birch.



1.3.3. Williams' Biosynthetic Studies of the Brevianamides

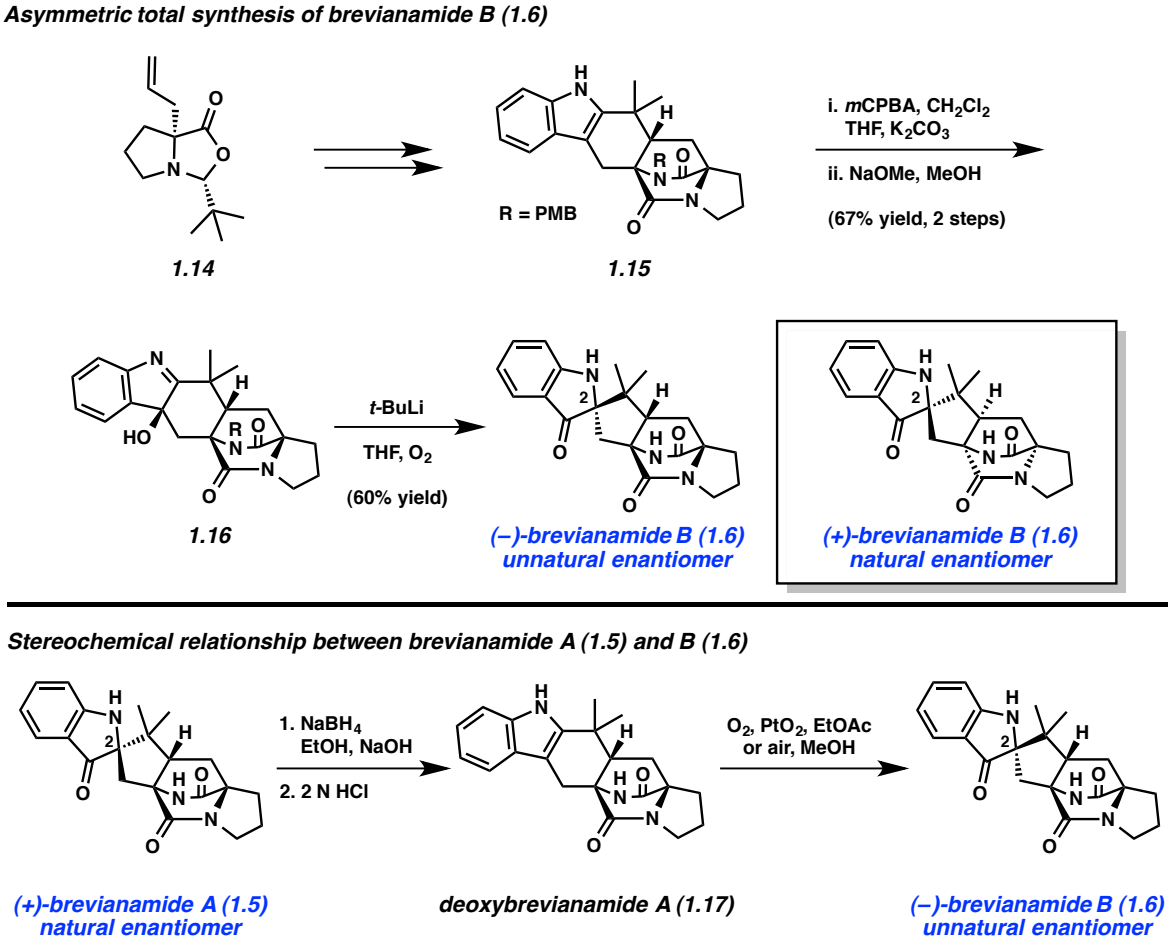
Nearly a decade would pass after Birch's preliminary efforts before the brevianamides were revisited.[23,24] It was not until 1986 that the Williams group began extensive studies to gain a clearer picture of how nature transforms tryptophan (**1.11**) and proline (**1.12**) to **1.5**–**1.10**.[24] Although these simple amino acids were validated using Birch's isotope labeling studies, there were many unanswered questions in the biosynthetic pathway. This included the stereochemical relationship between **1.5**–**1.10**, what other intermediates are formed after incorporation of **1.11** and **1.12**, and how the bicyclo[2.2.2]diazo-octane ring system is formed.[24] To answer some of these questions, the Williams group has judiciously utilized synthetic chemistry.[11]

One highlight of their efforts to use synthetic chemistry in biosynthetic investigations is their asymmetric total synthesis of brevianamide B (**1.6**), which led to the determination of the absolute configuration of **1.6**.[11,26] Shown in Scheme 1.2, known heterocycle **1.14**[25] was elaborated to hexacyclic indole **1.15**. Oxidation and base-induced rearrangement of **1.15** afforded the stable hydroxylindolenine **1.16**. Removal of the *p*-methoxybenzyl protecting group afforded (–)-brevianamide B (**1.6**). Interestingly, the specific rotation of the synthesized sample of

brevianamide B (**1.6**) was of equal magnitude, but of opposite sign than of their authentic sample isolated from *P. brevicompactum* fungus. This suggested that the absolute configuration of naturally occurring brevianamide B (**1.6**) is as shown.[26]

Williams and coworkers further validated this observation using semisynthetic studies. Synthetically useful quantities of natural (+)-brevianamide A (**1.5**), whose structure was confirmed earlier by single-crystal X-ray analysis,[27] was isolated from the fungal culture. The C2 stereocenter of natural (+)-brevianamide A (**1.5**) was ablated via a reduction and dehydration process to afford deoxybrevianamide A (**1.17**) (Scheme 1.2). This was subsequently oxidized to semisynthetic (–)-brevianamide B (**1.6**) following Birch's protocol.[19] The specific rotation of the semisynthetic (–)-brevianamide B (**1.6**) shared the same sign as the synthetic material, but again, was opposite in sign compared to the naturally occurring alkaloid. This is indicative that naturally occurring brevianamides A (**1.5**) and B (**1.6**) share an (*R*)-configuration at the C2-spirocyclic stereogenic center. However, in regards to the bicyclo[2.2.2]diazo-octane ring system, **1.5** and **1.6** are enantiomorphic.[11,23] The determination of the stereochemical relationship between naturally occurring brevianamides A (**1.5**) and B (**1.6**), will play into Williams' biomimetic approach to these natural products discussed later in this chapter.

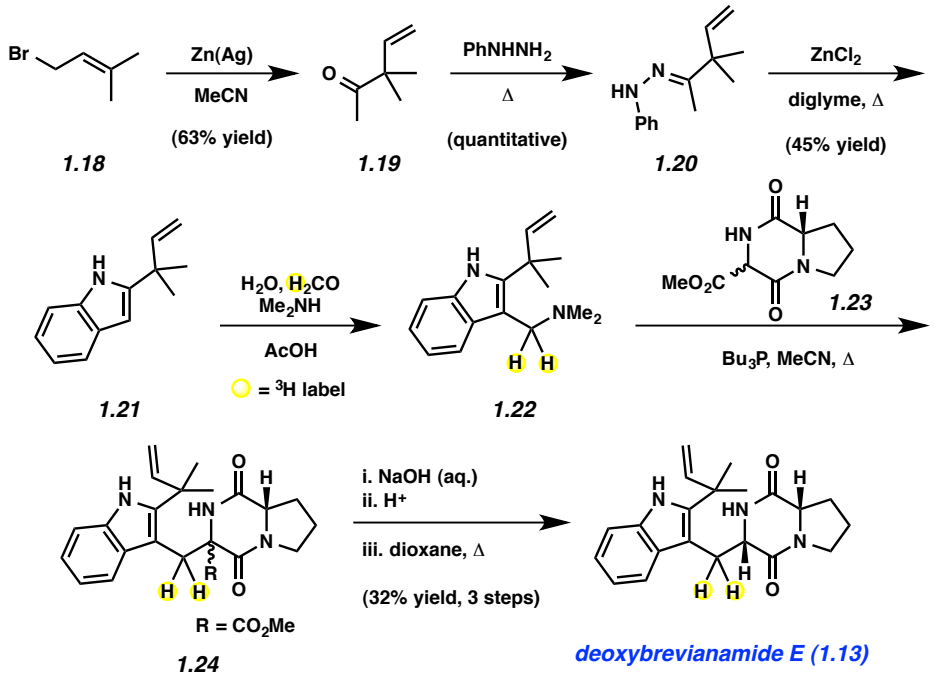
Scheme 1.2. Williams' synthesis of brevianamide B (**1.6**) and stereochemical relationship between brevianamide A (**1.5**) and B (**1.6**).



The Williams group continued to elucidate the biosynthetic pathway of the brevianamides by determining what other intermediates are formed after incorporation of tryptophan (**1.11**) and proline (**1.12**). Deoxybrevianamide E (**1.13**), an intermediate initially proposed by Birch (see Scheme 1.1),[19] was synthesized by Williams for feeding studies. To facilitate analysis, the Williams group synthesized tritium-labeled deoxybrevianamide E (**1.13**), in a similar approach to Birch's use of tritium-labeled amino acids, as shown in Scheme 1.3.[24] Commercially available alkyl bromide **1.18** was transformed to ketone **1.19** using Zn(Ag) in

CH_3CN. **1.19** then underwent a classic Fischer indolization reaction by a two-step sequence with initial formation of the hydrazone **1.20**, followed by treatment with zinc chloride at elevated temperatures to afford indole **1.21**. Reaction of **1.21** with tritium-labeled formaldehyde and dimethylamine then furnished gramine derivative **1.22**. This gramine derivative then underwent a Somei coupling[28] with piperazine **1.23** to produce compound **1.24**. Subsequent hydrolysis and decarboxylation afforded tritium-labeled **1.13**.

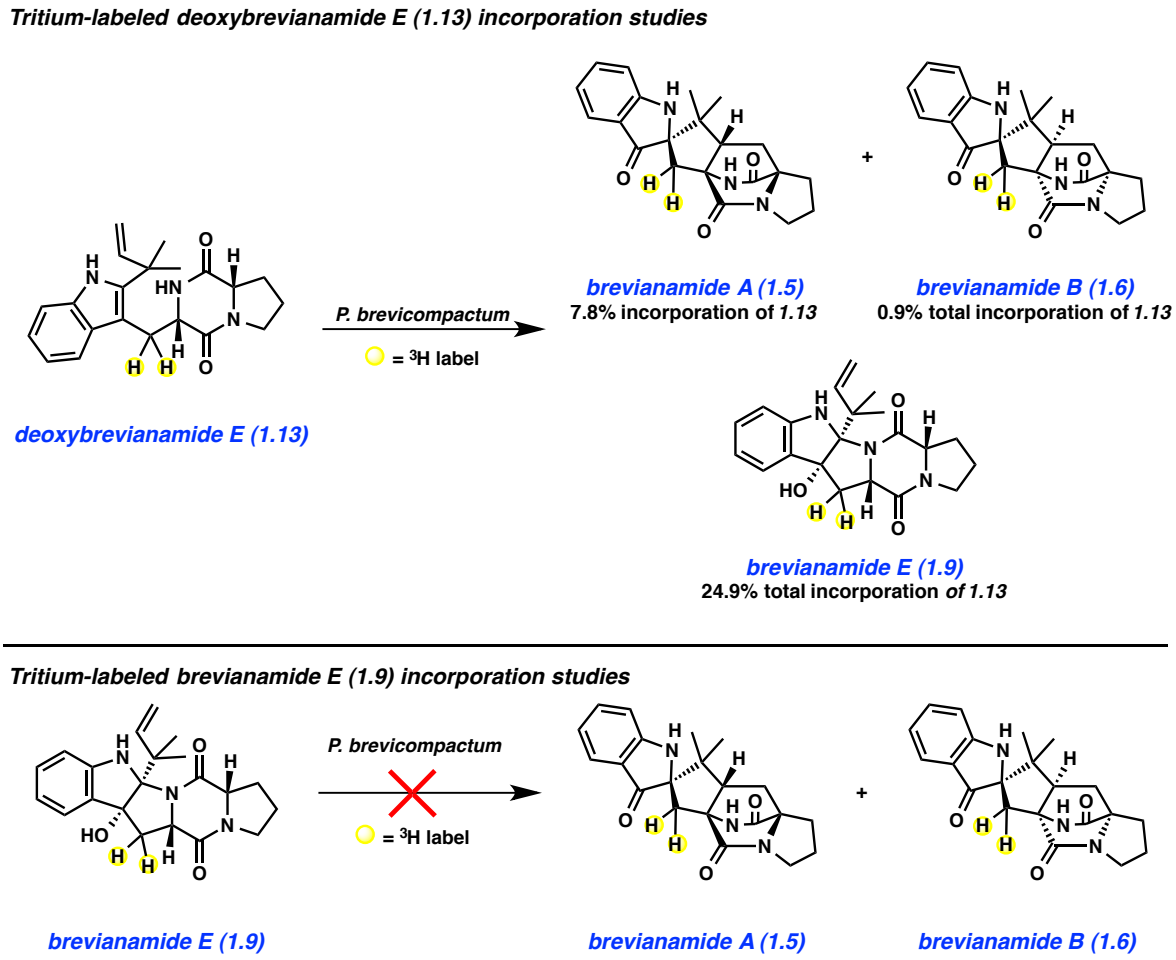
Scheme 1.3. Williams' synthesis of tritium-labeled **1.13**.



With tritium-labeled deoxybrevianamide E (**1.13**) available, it was administered to *P. brevicompatum* cultures. Incorporation of **1.13** into brevianamides A (**1.5**), B (**1.6**), and E (**1.9**) was observed, thus validating its intermediacy in the biosynthesis of **1.5**, **1.6**, and **1.7**. Tritium-labeled brevianamide E (**1.9**) was then reintroduced into *P. brevicompactum* cultures; however, there was no incorporation into brevianamide A (**1.5**) or B (**1.6**) was observed (Scheme 1.4).

This suggested that brevianamide E (**1.9**) is a dead-end metabolite and that the bicyclo[2.2.2]diazo-octane ring system common to **1.5** and **1.6** is constructed from deoxybrevianamide E (**1.13**).[29]

Scheme 1.4. Tritium-labeled deoxybrevianamide E (**1.13**) incorporation studies.

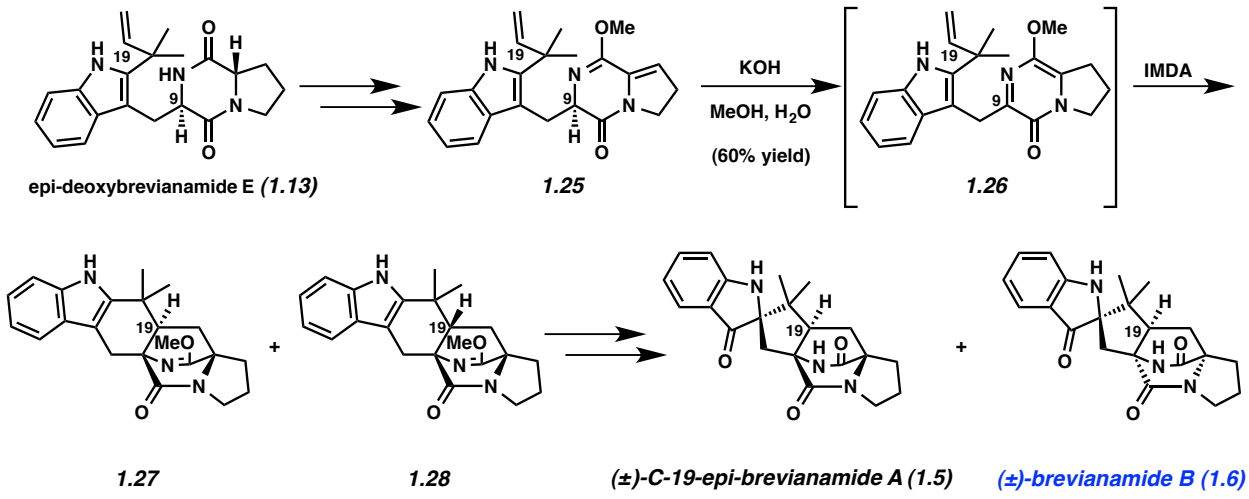


1.3.4 Williams' Biomimetic Approach to the Brevianamides

Biosynthetic studies described above prompted the Williams group to pursue a biomimetic approach toward the bicyclo[2.2.2]diazo-octane ring system in brevianamides A (**1.5**) and B (**1.6**). Based on literature precedent[30] and preliminary model system studies in with

corresponding author Sanz-Cervera,[31,32] it was suggested that an intramolecular Diels–Alder (IMDA) reaction could install the bicyclo[2.2.2]diazo ring system found in the brevianamides and related structures. Williams' biomimetic approach utilizing this idea is summarized in Scheme 1.5.[31] 9-epi-deoxybrevianamide E (**1.13**) was prepared and transformed to ether **1.25** using a two-step sequence. Base-mediated tautomerization of ether **1.25** to azadiene **1.26** led to spontaneous cyclization to afford epimeric cycloadducts **1.27** and **1.28**. The cycloadducts were subsequently elaborated to (±)-C-19-epi-brevianamide A (**1.5**) and (±)-brevianamide B (**1.6**), respectively. The study demonstrated that an intramolecular Diels–Alder reaction could be used to assemble the core structure of these natural products.[32]

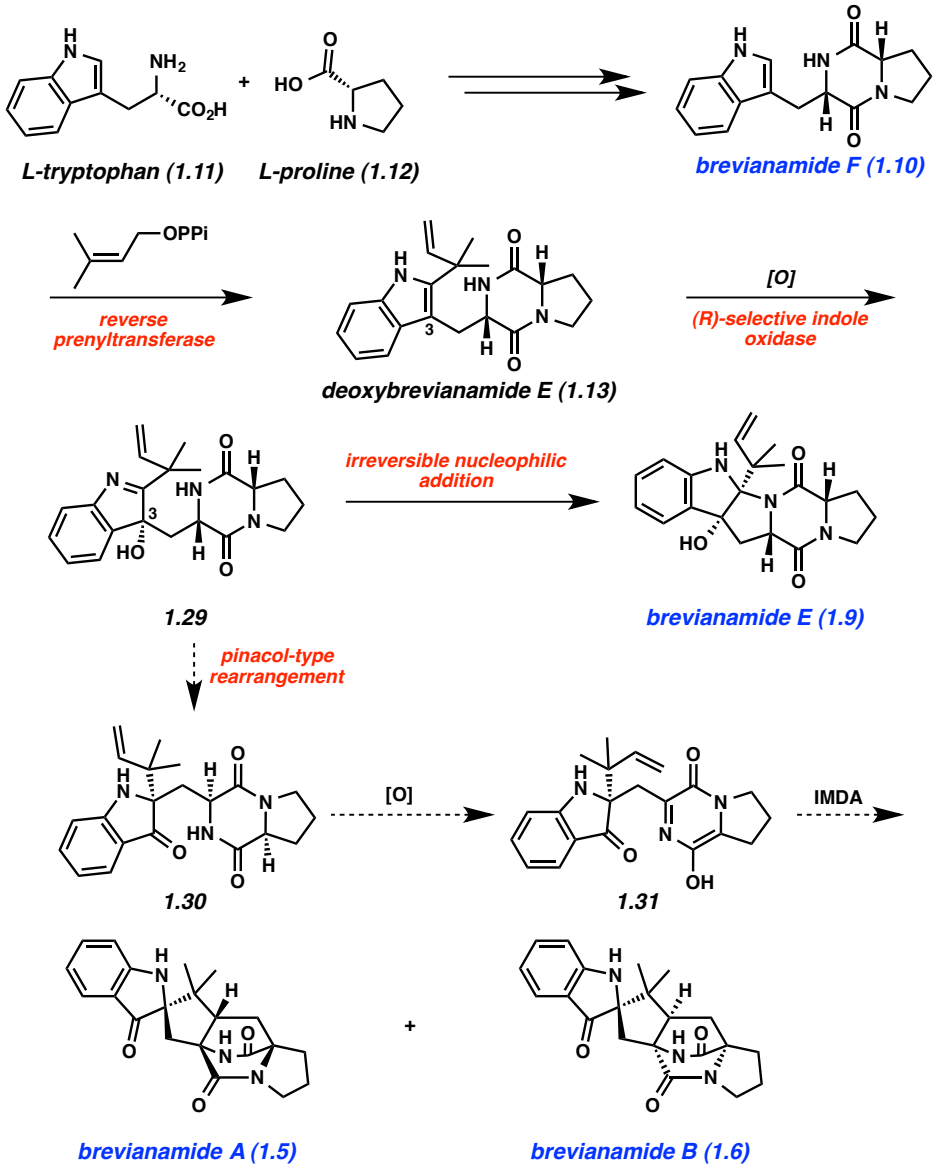
Scheme 1.5. Williams' biomimetic synthesis of (±)-brevianamide B (**1.6**).



With Birch's early work[19,20,21] and their own compiled studies, the Williams group developed the biosynthetic proposal shown in Scheme 1.6.[11] Following the union of **1.11** and **1.12**, brevianamide F (**1.10**) undergoes a reverse prenylation to give deoxybrevianamide E (**1.13**). This metabolite goes through an (*R*)-selective oxidation at the C3 position of the indole to

afford (*R*)-hydroxyindolenine (**1.29**), which can undergo irreversible nucleophilic addition by the amide nitrogen to form brevianamide E (**1.9**). Alternatively, a pinacol-type rearrangement of **1.29** forms metabolite **1.30**. Subsequent oxidation of the diketopiperazine motif in **1.30** forms the azadiene **1.31**, which then undergoes an intramolecular Diels–Alder reaction to form brevianamides A (**1.5**) and B (**1.6**).

Scheme 1.6. Current understanding of brevianamide biosynthesis.



Overall, substantial progress in understanding the biosynthesis of the brevianamide natural products has been made. Access to these compounds via total synthesis has provided insight into the absolute configuration of brevianamide B (**1.6**), which in turn, prompted investigations regarding the assembly of the bicycle present in brevianamide A (**1.5)** and B (**1.6**). Also, the application of synthetic chemistry to produce complex, isotopically-labeled compounds permitted feeding studies that validated the intermediacy of deoxybrevianamide E (**1.13**) in the pathway. Alongside this work, Williams' successful biomimetic strategy to the brevianamides lends credibility to their proposed biosynthesis of the bicyclo[2.2.2]diazo-octane framework via an intramolecular Diels–Alder reaction.

There have been numerous advances made in elucidating the biosynthesis of the brevianamides; however, questions still remain. This includes validation of a number of intermediates, such as **1.30** and **1.31**, and also complete isolation and characterization of the brevianamide gene cluster. Nonetheless, the aforementioned contributions from Birch, Williams and other groups demonstrate the integral role of synthetic chemistry to investigate biosynthetic pathways.

1.4 Notoamides

Originally isolated by Tsukamoto in 2007 from *Aspergillus* sp., notoamides A–D (**1.32**–**1.35**, Figure 1.3) have emerged as interesting metabolites related to the brevianamides, stephacidins, paraherquamide, and other bicyclo[2.2.2]diazo-octane containing alkaloids.[33] Since the original isolation paper, over a dozen family members have been isolated, with representative members depicted below.[34,35,46] Of note, notoamides C (**1.34**), D (**1.35**), and S (**1.37**) are considered to be precursors to mature notoamides A (**1.32**), B (**1.33**), and T (**1.38**), which contain

the complex bicyclo[2.2.2]diazo-octane ring system. In general, notoamides are very structurally similar to the brevianamides discussed earlier, except that notoamides A–E (**1.32**–**1.36**) are distinguished by a fused pyran moiety at C6 and C7 of the indole. The notoamides are also bioactive isolates with notoamides A–C (**1.32**–**1.34**), in particular, displaying cytotoxicity against HeLa and L1210 cells with IC_{50} values in the range of 22–52 μg/mL.[33]

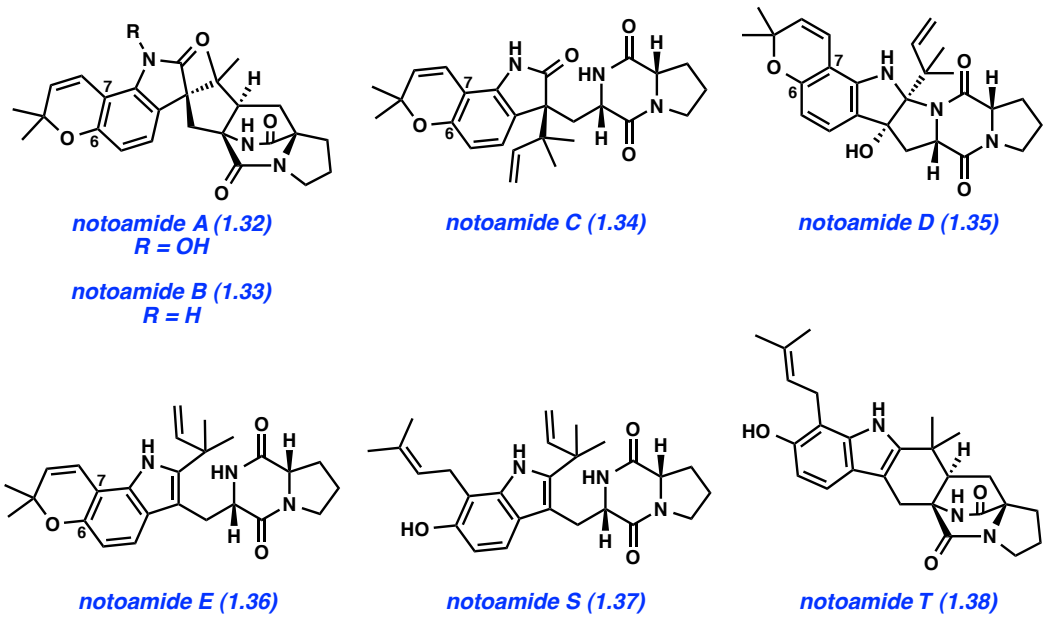


Figure 1.3. Representative notoamide natural products (**1.32**–**1.38**).

In their continued studies of bicyclo[2.2.2]diazo-octane containing alkaloids, the Williams laboratory has performed a thorough analysis of the notoamide biosynthetic pathway,[36,37,47] while Sherman[38,47] and Tsukamoto[35] have also provided detailed insight. This section will focus on the overall impact of synthetic chemistry in the determination of the enzymes involved in biosynthesis, the biochemical relationships between different natural product families, and the discovery of intermediates in the pathway to mature notoamides A (**1.32**), B (**1.33**), and T (**1.38**).

1.4.1 In Vitro Reconstitution Studies

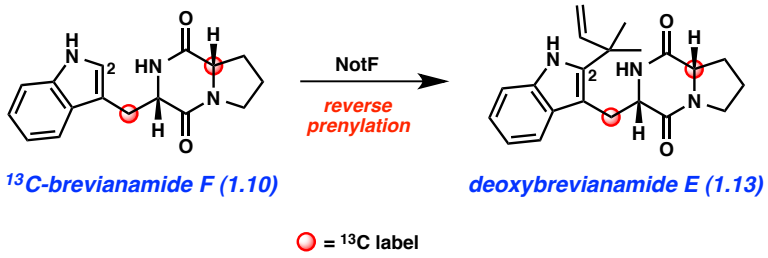
In recent years, technological advances in genetics and molecular biology have tremendously impacted biosynthetic studies. Along with conventional feeding studies, newer techniques such as genome mining have emerged to elucidate how nature makes these compounds enzymatically.[3] In the case of genome mining, known gene sequences are used to predict and discover homologous genes involved in natural product biosynthesis. Evidence for the utility of these methods can be found in the recently reported isolation and characterization of the notoamide gene cluster. In combination with synthetic chemistry, a number of metabolites formed in the biosynthetic pathway were validated as intermediate.[38]

Sherman, in collaboration with Williams and Tsukamoto, identified and characterized the notoamide gene cluster through genome mining of *Aspergillus* sp. It was reasoned that *Aspergillus* sp., which also produces deoxybrevianamide E (**1.13**), uses **1.13** as a common intermediate in the early steps of both brevianamide and notoamide biosynthesis.[37] Subsequent bioinformatic analysis for the notoamide gene cluster uncovered two predicted prenyltransferases, NotF and NotC, which were used in reconstitution studies to corroborate the direct relationship between the brevianamides and notoamides.

NotF was identified as a deoxybrevianamide E synthase and reconstituted for in vitro studies. ^{13}C-labeled brevianamide F (**1.10**) was prepared following Birch's synthesis,[39] and used as a substrate in NotF precursor incorporation studies (Scheme 1.7). Incorporation of ^{13}C-labeled brevianamide F (**1.10**) into deoxybrevianamide E (**1.13**) was observed, and NotF was shown to selectively catalyze the reverse prenylation at C2 of brevianamide F (**1.10**) to afford deoxybrevianamide E (**1.13**).[37] This prenylation is thought to ultimately play a central role in the assembly of the bicyclo[2.2.2]diazo-octane ring system found in notoamides A (**1.32**) and B

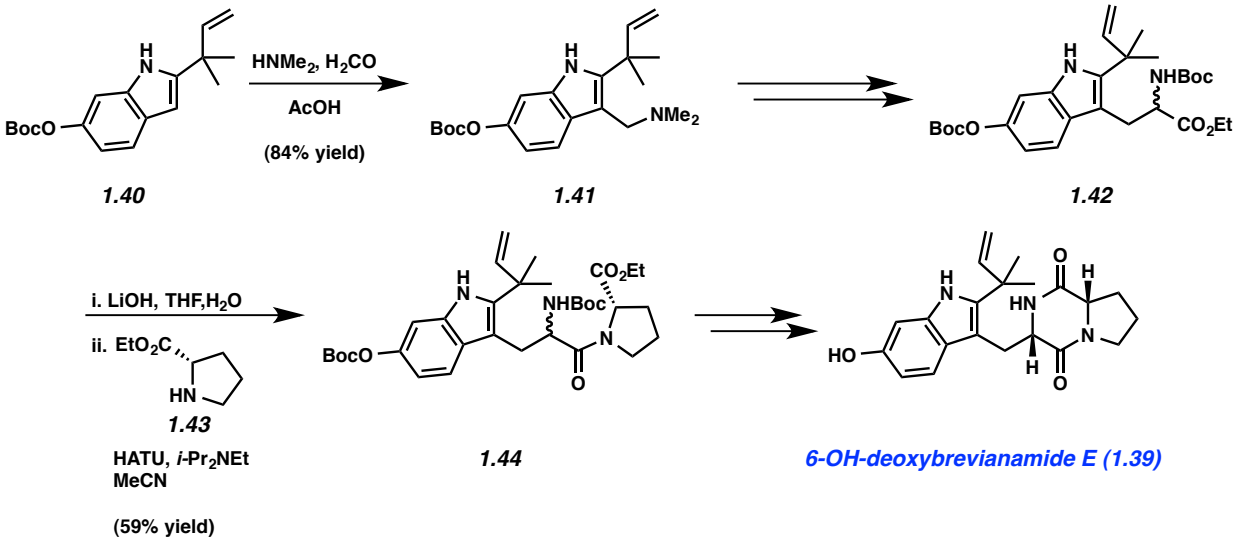
(**1.33**), analogous to the brevianamide natural products.[11] In addition to the structural resemblance of **1.10** and **1.13** with the notoamides, these studies revealed a direct biosynthetic relationship between the brevianamide and notoamide natural products.[40]

Scheme 1.7. Proposed biosynthetic relationship of brevianamide F (**1.10**) and deoxybrevianamide E (**1.13**).



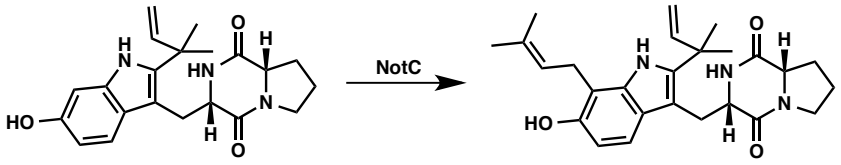
The prenyltransferase activity of NotC was also examined. In order to validate the proposed prenyltransferase activity of NotC, a number of substrates, including deoxybrevianamide E (**1.13**) and 6-OH-deoxybrevianamide E (**1.39**) were synthesized. **1.13** was synthesized as described previously (see Scheme 1.3). 6-OH-deoxybrevianamide E (**1.39**) was chemically synthesized in an analogous fashion as shown in Scheme 1.8.[39,41] Mannich reaction of prenylated indole **1.40** afforded gramine derivative **1.41**. Elaboration of **1.41** to tryptophan derivative **1.42**, followed by saponification and peptide coupling with **1.43**, delivered intermediate **1.44**. Deprotection and cyclization forged the diketopiperazine scaffold to give 6-OH-deoxybrevianamide E (**1.39**).

Scheme 1.8. Williams's synthesis of 6-OH-deoxybrevianamide E (**1.39**).



With deoxybrevianamide E (**1.13**) and 6-OH-deoxybrevianamide E (**1.39**) available, they were independently administered to NotC in vitro reconstitution assays. No incorporation of deoxybrevianamide E (**1.13**) was observed; however, 6-OH-deoxybrevianamide E (**1.39**) was identified as the exclusive substrate to NotC (Scheme 1.9). It is likely that **1.13** is converted to 6-OH-deoxybrevianamide E (**1.39**) by an as-of-yet unidentified oxidase, and then used by NotC as a substrate.[38] This subsequently produced what is now referred to as notoamide S (**1.37**), which remarkably has never been isolated from traditional natural products isolation studies. This is likely due to its role as a short-lived biosynthetic intermediate.[38] Nevertheless, its formation in NotC reconstitution studies provided significant evidence that **1.37** is a precursor to other notoamides in the biosynthetic pathway. Furthermore, it was also deduced that the pyran motif present in notoamide A–E (**1.32–1.36**) is installed after construction of notoamide S (**1.37**), following both prenylation events.[38,47]

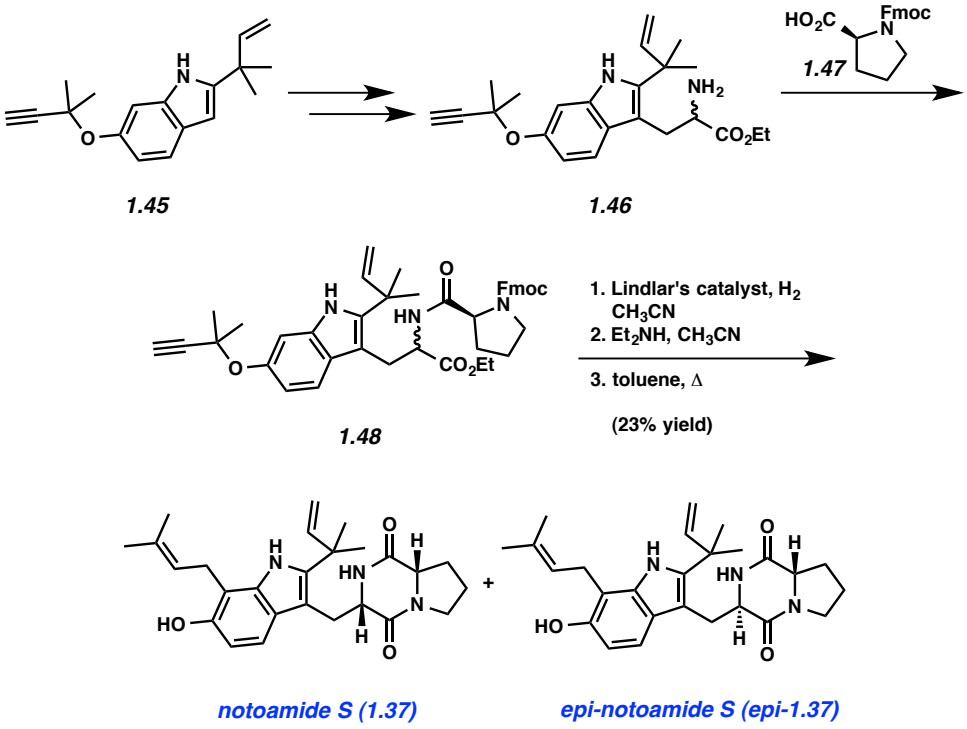
Scheme 1.9. Biosynthetic role of 6-OH-deoxybrevianamide E (**1.39**).



6-OH-deoxybrevianamide E (1.39) *notoamide S (1.37)*

To confirm the structure of notoamide S (**1.37**) obtained from NotC feeding studies, an authentic synthetic sample was prepared. Williams' total synthesis of notoamide S (**1.37**), summarized in Scheme 1.10,[42] commenced with transformation of known indole **1.45** into tryptophan derivative **1.46**. Somei coupling[28] of **1.46** with proline derivative **1.47** afforded amide **1.48**. Reduction of the propargyl alkyne in **1.48** using Lindlar's catalyst and H_2, followed by Fmoc-removal, Claisen rearrangement, and cyclization, afforded notoamide S (epi-**1.37**) and epi-notoamide S (**1.37**).[41] The structure of synthetic notoamide S (**1.37**) matched the sample prepared from the in vitro reconstitution assay.[38]

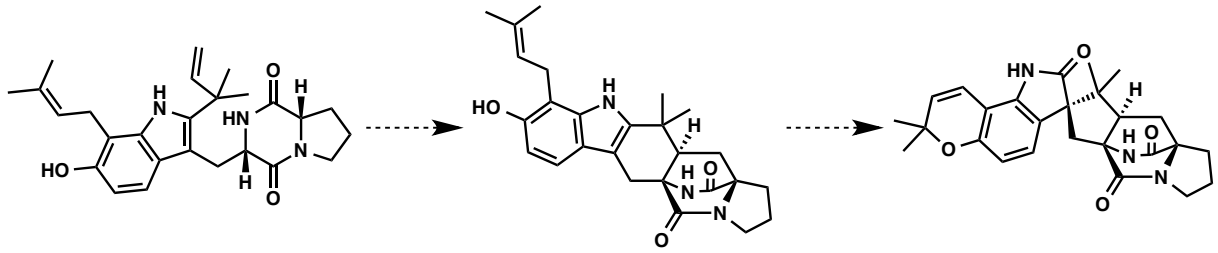
Scheme 1.10. Williams's synthesis of notoamide S (**1.37**).



1.4.2 Williams' Synthesis of Notoamide T (1.38) and Biosynthetic Studies

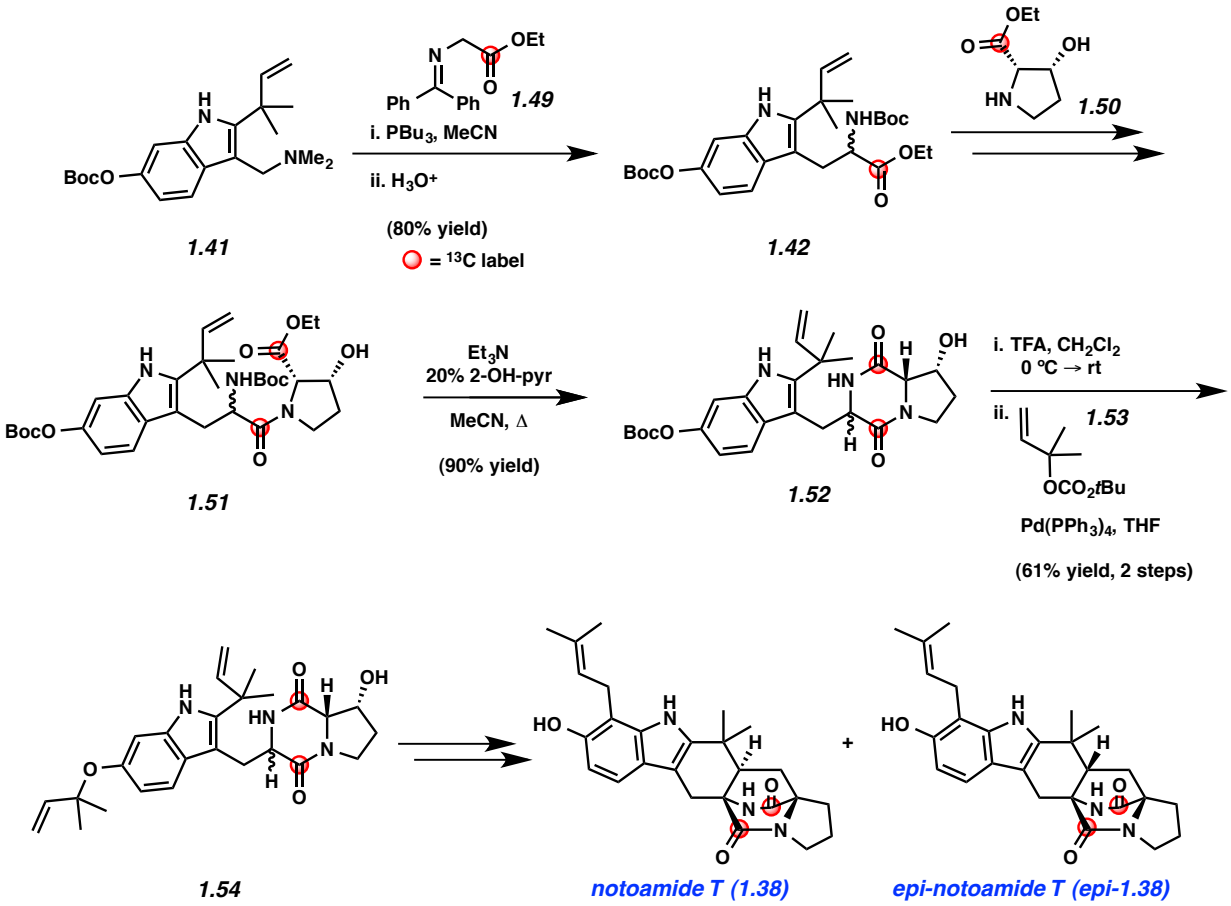
It was proposed that notoamide S (**1.37**) could serve as the biosynthetic precursor to other notoamides as shown in Scheme 1.11.[40] In an analogous fashion to the brevianamides, notoamide S (**1.37**) could undergo an oxidation and subsequent intramolecular Diels–Alder reaction to afford notoamide T (**1.38**), which contains a bicyclo[2.2.2]diazo-octane ring system, but lacks the pyran ring found in notoamide A–E (**1.32–1.36**). In turn, notoamide T (**1.38**) would serve as a biosynthetic precursor to notoamide B (**1.33**). Similar to notoamide S (**1.37**), notoamide T (**1.38**) had not been isolated from fungal extracts.[43]

Scheme 1.11. Proposed biosynthetic relationship between **1.37**, **1.38**, and **1.33**.



To investigate its potential intermediacy in notoamide biosynthesis, the Williams group synthesized doubly ^{13}C-labeled notoamide T (**1.38**) for precursor incorporation experiments.[43] As shown in Scheme 1.12, Somei coupling[28] of gramine **1.41** with ^{13}C-labeled glycine derivative **1.49**, followed by imine hydrolysis, afforded tryptophan **1.42**.[44] Peptide coupling of tryptophan **1.42** with ^{13}C-labeled **1.50** afforded dipeptide **1.51**. Subsequent base-catalyzed cyclization to install the diketopiperazine led to intermediate **1.52**. Of note, the *N*-Boc group was also removed in this process. Following *O*-Boc cleavage, Pd-catalyzed allylic alkylation[45] with carbonate **1.53** installed the requisite reverse prenyl group and furnished ether **1.54**. Finally, mesylation of the alcohol in **1.54**, followed by a sequence involving elimination, tautomerization, IMDA, and Claisen rearrangement then afforded notoamide T (**1.38**) and the inconsequential isomer epi-**1.38**.

Scheme 1.12. Williams' synthesis of ^{13}C- labeled notoamide T (**1.38**).



With access to ^{13}C-labeled notoamide T (**1.38**), it was used for feeding studies to test the hypothesis that notoamide T (**1.38**) is a precursor for notoamide B (**1.33**). Incorporation of notoamide T (**1.38**) into notoamide B (**1.33**) was observed. This result provided direct evidence that notoamide T (**1.38**) is formed during the notoamide biosynthetic pathway, specifically, in the biosynthesis of notoamide B (**1.33**).[43]

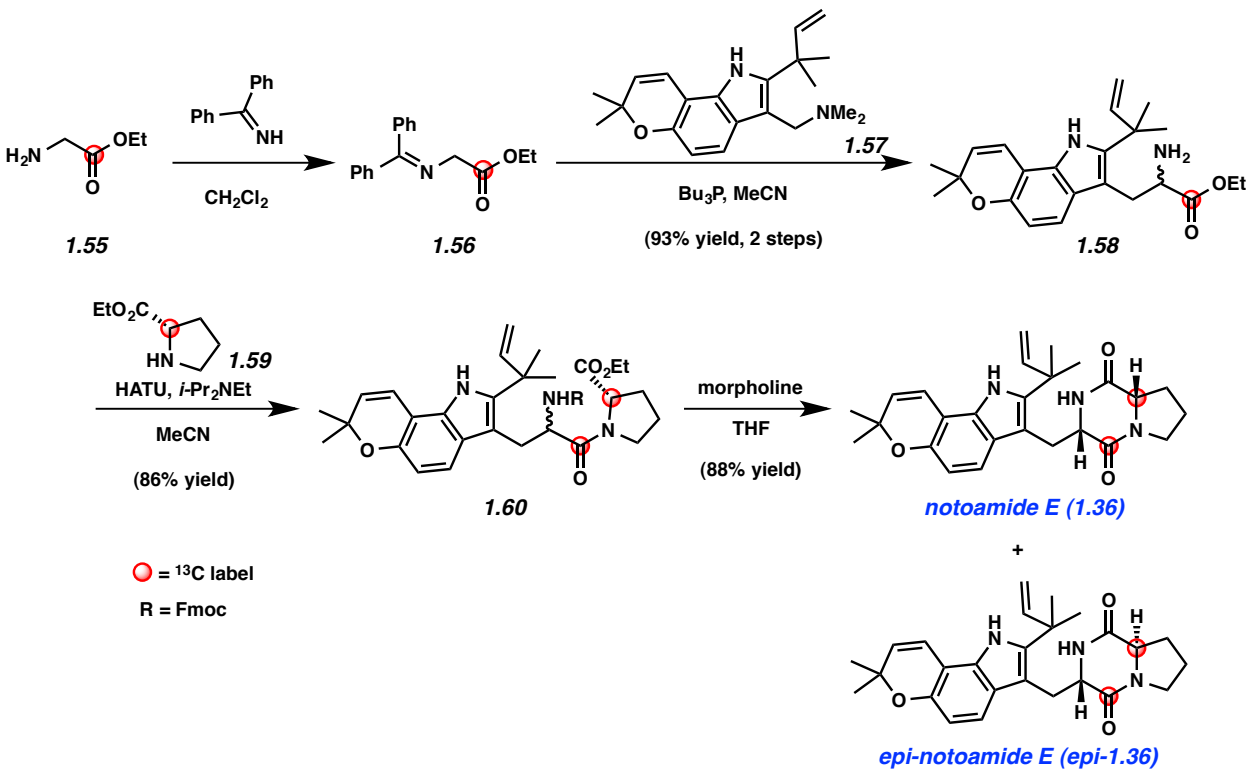
1.4.3 Total Synthesis of Notoamide E (1.36) and Biosynthetic Studies

In contrast to notoamide S (**1.37**) and T (**1.38**), it was possible to isolate notoamide E (**1.36**) from fungal extracts; however, it was found to exist for only one day during a 20-day

fermentation time period, suggesting it is an intermediate in the synthesis of other notoamides.[46] Notoamide E (**1.36**) is distinct from its related family members because of its pyran motif; however, it lacks a bicyclo[2.2.2]diazo-octane ring system. Williams and Tsukamoto proposed that notoamide E (**1.36**) was a precursor to notoamide A (**1.32**) and B (**1.33**).[46]

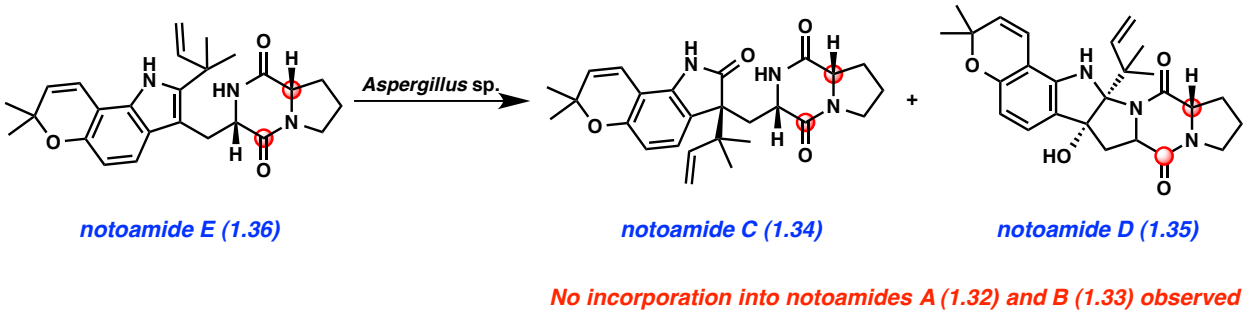
In order to verify this hypothesis, doubly ^{13}C-labeled notoamide E (**1.36**) was synthesized for feeding studies (Scheme 1.13). Condensation of ^{13}C-glycine **1.55** furnished benzophenone imine **1.56**. Somei coupling[28] with **1.57** afforded tryptophan derivative **1.58**, which underwent peptide coupling with **1.59** to provide **1.60**. Cyclization with concomitant deprotection afforded doubly ^{13}C-labeled notoamide E (**1.36**) and epi-**1.36**.[46]

Scheme 1.13. Williams' synthesis of doubly ^{13}C-labeled notoamide E (**1.36**).



With ^{13}C-labeled **1.36** available, Williams and coworkers performed feeding studies (Scheme 1.14). Contrary to their original proposal, no incorporation into notoamide A (**1.32**) and B (**1.33**) was observed.[40] Instead, notoamide E (**1.36**) was converted into notoamides C (**1.34**) and D (**1.35**), most likely through a pinacol-like rearrangement, analogous to that proposed for the brevianamides (see Scheme 1.6).[47,48]

Scheme 1.14. Notoamide E (**1.36**) incorporation studies.



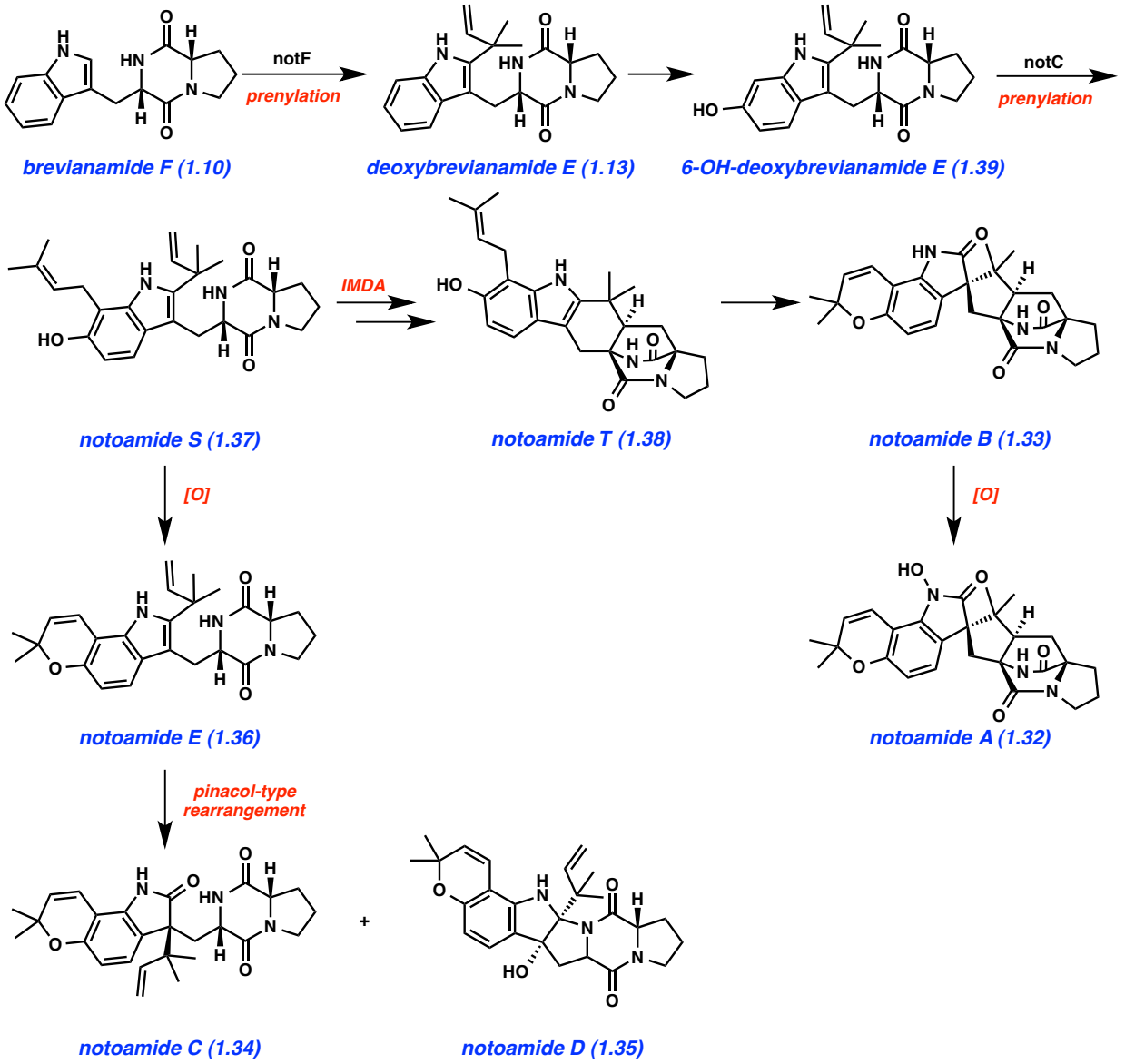
1.4.4 Proposed Biosynthetic Pathway of the Notoamides

As more information was uncovered about the notoamide biosynthesis, a greater number of intermediates were proposed.[49,50,51,52,53] The Williams laboratory was able to synthesize numerous structural analogues and many of these were later confirmed to be intermediates, despite not being found in fungal extracts using conventional natural product isolation methods. Also, access to these intermediates permitted Williams and collaborators to demonstrate the direct relationship between the notoamide and brevianamide family of natural products.

A summary of the currently proposed biosynthetic pathway for the notoamides is depicted in Scheme 1.15. Sharing a common intermediate with the brevianamides, the notoamides arise from brevianamide F (**1.10**). Enzymatic conversions of **1.10**, via **1.13** and **1.39**, lead to the pivotal metabolite notoamide S (**1.37**). Notoamide S (**1.37**) serves as a divergent point

in notoamide biosynthesis. One pathway involves oxidative cyclization of notoamide S (**1.37**) into notoamide E (**1.36**). A pinacol-type rearrangement, as proposed for the brevianamides,[11] could then afford notoamide C (**1.34**) or notoamide D (**1.35**). Another pathway involves conversion of notoamide S (**1.37**) to notoamide T (**1.38**) via an intramolecular Diels–Alder reaction.[43] In turn, notoamide T (**1.38**) is elaborated to notoamide A (**1.32**) and notoamide B (**1.33**), with notoamide B (**1.33**) believed to be the precursor to notoamide A (**1.32**).[47]

Scheme 1.15. Proposed biosynthetic pathway of the notoamides.



While the enzymatic function of the entire notoamide gene cluster has not been completed, significant insight into the biosynthetic steps has been gained. The technological advances in the biological sciences during the past two decades have undeniably fueled this progress. However, as evident by the studies discussed herein, synthetic chemistry has also played a pivotal role. In the case of the notoamides, synthetic chemistry was integral for the

confirmation of several proposed intermediates and provided key substrates needed for feeding studies. There are remaining questions though, including how notoamide B (**1.33**) is converted to notoamide A (**1.32**), and the characterization of a Diels–Alderase enzyme involved in catalyzing an IMDA reaction. The breakthroughs in understanding notoamide biosynthesis demonstrate that close collaborations between the fields of genetics, biochemistry, and synthetic chemistry can greatly enhance our understanding of biosynthetic pathways.

1.5 Communesins

The communesins (**1.61**–**1.68**, Figure 1.4) are another class of intriguing fungal indole alkaloids. Since their isolation from a strain of *Penicillium* sp. in 1993, communesins A (**1.61**) and B (**1.62**) have been of great interest to the chemical community.[54] Over the years, a number of related communesin natural products C–H (**1.63**–**1.68**) have been isolated,[55] in addition to the structurally related compound perophoramidine (**1.69**).[56] These metabolites are distinguished by their complex polycyclic skeletons, with communesins A–H (**1.61**–**1.68**) heptacyclic ring systems. With respect to communesin B (**1.62**), there are six stereocenters, four of which are contiguous and two are vicinal quaternary stereocenters. Also, the presence of bis(aminal) linkages adds an extra degree of complexity to communesin B (**1.62**) and its relatives.

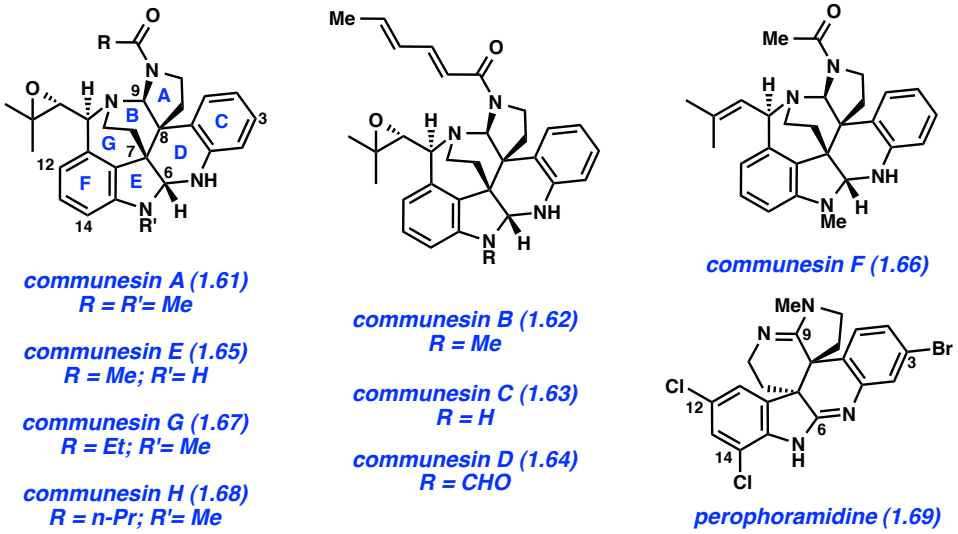


Figure 1.4. The communesin alkaloids (**1.61**–**1.68**) and perophoramidine (**1.69**).

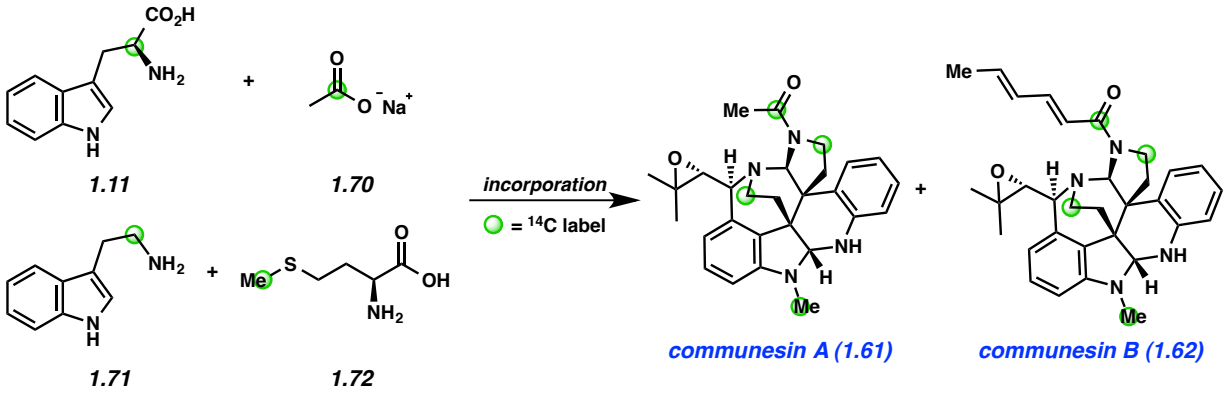
In addition to their structural complexity, the communesins possess interesting biological activity. Initial studies showed that some communesins exhibit moderate insecticidal activity. Communesins A (**1.61**) and B (**1.62**) also display modest activity against the P-388 leukemia cell line with ED_{50} values of 3.5 μg/mL and 0.45 μg/mL, respectively.[54] Communesin B (**1.62**) displays modest activity against the LoVo and KB cell lines with minimal inhibitory concentration (MIC) values of 3.9 and 8.8 μM, respectively.[57] The attractive structural and biological profiles of the communesins have made them formidable targets for synthetic chemists over the past two decades. While there have been elegant syntheses reported,[58] this work will be discussed in Chapter 3.

In contrast to the brevianamides and notoamides, there is no comprehensive study about the communesin biosynthetic pathway reported. There has only been limited work to confirm the initial building blocks and to clarify an interesting structural misassignment, which was done in the context of biomimetic synthetic approaches. These efforts are discussed herein.

1.5.1 Wigley's Biosynthetic Studies of the Communesins

The first biosynthetic study, by Wigley in 2006,[59] utilized isotopically-labeled small molecules in feeding experiments with *Penicillium* sp. cultures to determine the early steps in communesin biosynthesis (Scheme 1.16). Feeding studies with commercially available ^{14}C-tryptophan (**1.11**), ^{14}C-acetate (**1.70**), (methyl-^{14}C)-methionine (**1.72**), and ^{14}C tryptamine (**1.71**) showed incorporation into communesins A (**1.61**) and B (**1.62**), indicating that all of these compounds are the basic building blocks of communesin biosynthesis (Scheme 1.16). Whereas the *N*-acyl group and *N*-Me groups can be traced to building blocks **1.70** and **1.72**, respectively, it was not clear how **1.11** and **1.71** mapped onto the natural products.

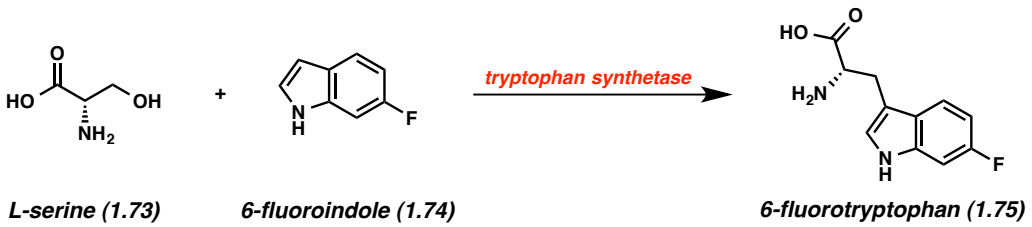
Scheme 1.16. ^{14}C-tryptophan (**1.11**), ^{14}C-acetate (**1.70**), (methyl-^{14}C)-methionine (**1.72**), and ^{14}C tryptamine (**1.71**) incorporation studies.



Substituted tryptophans were also used in feeding studies to help clarify where tryptophan gets incorporated into the communesins and to give unnatural communesin analogues. Radiolabeled halogenated tryptophans were also used, but none were incorporated into the communesins. Due to the substrate selectivity, non-radiolabeled 6-fluorotryptophan (**1.75**) was used in the feeding studies. 6-Fluorotryptophan (**1.75**) was accessed biosynthetically

from serine (**1.73**), 6-fluoroindole (**1.74**) and using a known tryptophan synthetase (Scheme 1.17).[60]

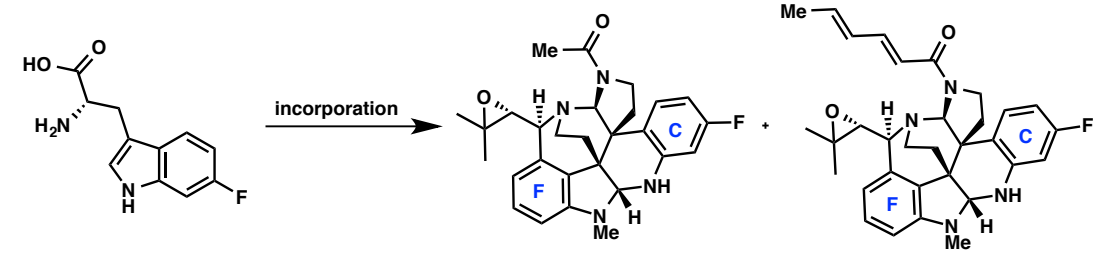
Scheme 1.17. Biosynthesis of 6-fluorotryptophan (**1.75**).



L-serine (1.73) *6-fluoroindole (1.74)* *6-fluorotryptophan (1.75)*

In the key feeding studies with 6-fluorotryptophan (**1.75**), monofluorinated analogues of communesins A (**1.61**) and B (**1.62**) accumulated (Scheme 1.18). These unnatural communesins were analyzed using mass spectroscopy, in light of the absence of the radiolabel (Scheme 1.18). Interestingly, incorporation was seen only in the C ring of the communesins. No incorporation of 6-fluorotryptophan (**1.75**) into the F ring was observed. These feeding studies indicated that the enzymes involved in communesin biosynthesis are able to differentiate tryptophan derivatives at an early stage. As will be discussed later, it is postulated that the F ring arises from the structurally related ergot alkaloid, aurantioclavine (**1.83**) (see Scheme 1.21).[61,62]

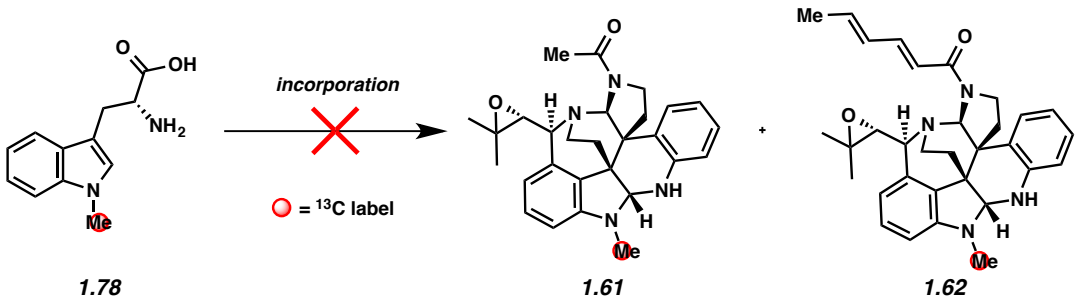
Scheme 1.18. 6-Fluorotryptophan (**1.75**) incorporation studies.



6-fluorotryptophan (1.75) *1.76* *1.77*

It was also proposed that the indole nitrogen of tryptophan (**1.11**) could be methylated at an early step in the biosynthesis. The resulting tryptophan derivative could then serve as a precursor to other communesins. To enable feeding studies, the authors prepared ^{13}C-labeled *N*-Me tryptophan (**1.78**) following literature precedent.[63] Using indole-*N*-(methyl-^{13}C)-tryptophan (**1.78**), no incorporation into either **1.61** or **1.62** was observed (Scheme 1.19). Therefore, it was deduced that methylation of tryptophan is not an early biosynthetic step and must occur later in the biosynthetic pathway.

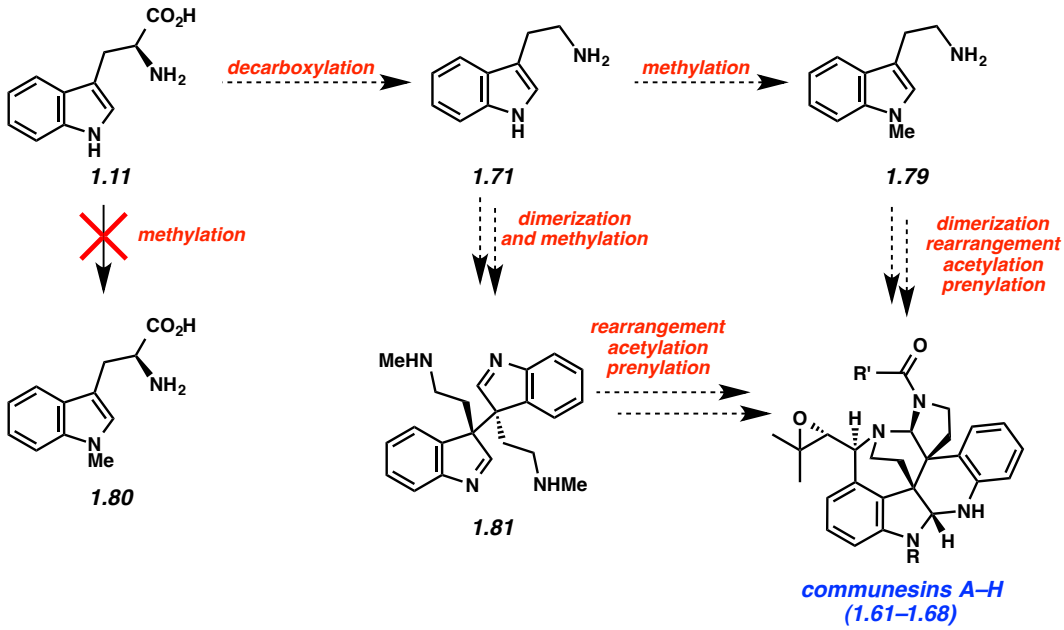
Scheme 1.19. ^{13}C-labeled *N*-Me tryptophan (**1.78**) incorporation studies.



Based on these basic observations, Wigley developed a general biosynthetic proposal for the communesins, which is summarized in Scheme 1.20. Wigley proposed that the first biosynthetic step is performed by a tryptophan decarboxylase, which converts tryptophan (**1.11**) to tryptamine (**1.71**).[64] One of two pathways could then be operative. **1.71** could be methylated to give **1.79**, which is then dimerized and further elaborated to mature communesin A–H (**1.61**–**1.68**) by a rearrangement, acetylation, and prenylation sequence. Another possibility is that the dimerization event could happen prior to the methylation to give **1.81**. Subsequent rearrangement, acetylation, and prenylation would give rise to **1.61**–**1.68**. From these preliminary studies, Wigley generalized that communesins A–H (**1.61**–**1.68**) arise from two molecules of

tryptophan-based intermediates, which later undergo a dimerization event.[58,60] Many details remain unknown.

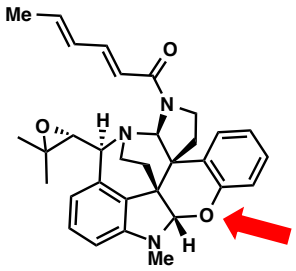
Scheme 1.20. Wigley's biosynthetic proposal.



1.5.2 Structural Misassignment and Biomimetic Approaches Toward the Communesins

In 2001, a compound named nomofungin (**1.82**) was isolated from a fungus growing on the bark of *Ficus microcarpa* by Hemscheidt and co-workers (Figure 1.5).[56] The proposed structure of **1.82** was nearly identical to that of communesin B (**1.62**), but with an *N*,*O*-acetal moiety instead of an aminal linkage. The questionable structural similarity between nomofungin (**1.82**) and communesin B (**1.62**), along with their identical ^{1}H and ^{13}C spectroscopic data, sparked the interest of multiple research groups. Synthetic studies by Stoltz and Funk independently showed that the structural assignment for nomofungin (**1.82**) was incorrect. In fact, nomofungin (**1.82**) was determined to be communesin B (**1.62**), and the nomofungin

isolation report was eventually retracted.[65] Even in the context of general structural assignment, chemical methods are an invaluable tool to validate proposed structures.



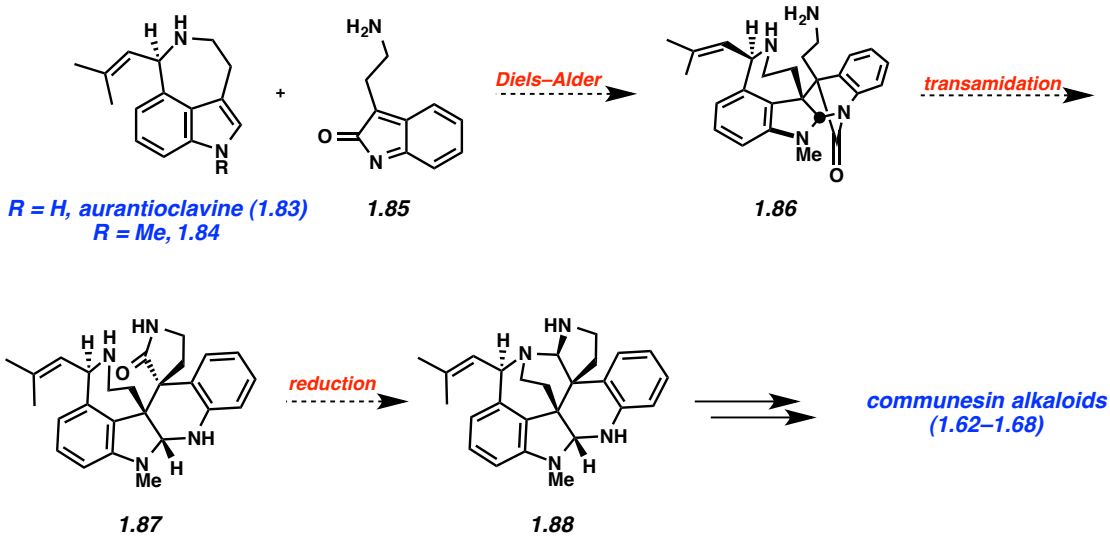
nomofungin (1.82)

Figure 1.5. Proposed structure of nomofungin (**1.82**).

1.5.3 Stoltz's Biomimetic Approach to the Communesins

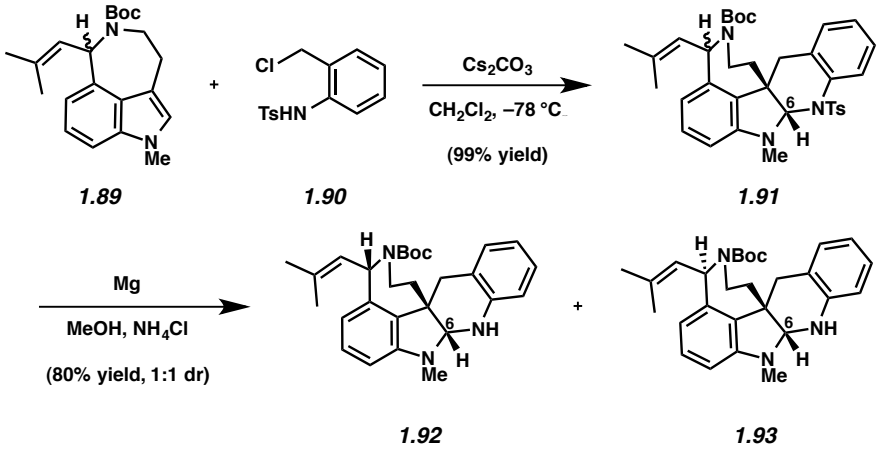
In their efforts to dismiss the proposed structure of nomofungin (**1.82**) and also provide access to the communesins, the Stoltz group disclosed a biomimetic strategy to prepare the framework of these compounds in 2003.[62,66] It was thought that **1.85**, an oxidized tryptamine derivative, could undergo a hetero-Diels–Alder reaction with aurantioclavine (**1.83**) or its methylated counterpart **1.84,** to generate intermediate **1.86** (Scheme 1.21). Transamidation of amine **1.86** could afford spiro-lactam **1.87**. Reduction and cyclization of aminolactam **1.87** could then furnish heptacyclic product **1.88** en route to communesin alkaloids **1.62**–**1.68**.

Scheme 1.21. Stoltz's biomimetic approach to the communesins.



Stoltz's effort to test the biomimetic hetero-Diels–Alder approach to the tetracyclic skeleton of the communesins is summarized in Scheme 1.22. Chloroaniline **1.90** and aurantioclavine derivative **1.89**, which was prepared synthetically,[67] were treated with cesium carbonate to afford pentacycle **1.91** using Corey's general hetero-Diels–Alder proposal.[68] Desulfonylation of **1.91** with magnesium and methanol afforded diastereomers **1.92** and **1.93** (1:1 dr), which could be separated by preparative thin layer chromatography. The ^{13}C chemical shifts at C6 for each diastereomer (84.8, 83.9, respectively in $CDCl_3$) were in agreement with the reported carbon resonance of communesin B (**1.62**) and nomofungin (**1.82**).[62] This corroborated the hypothesis that the structure of nomofungin (**1.82**) was in fact incorrect.

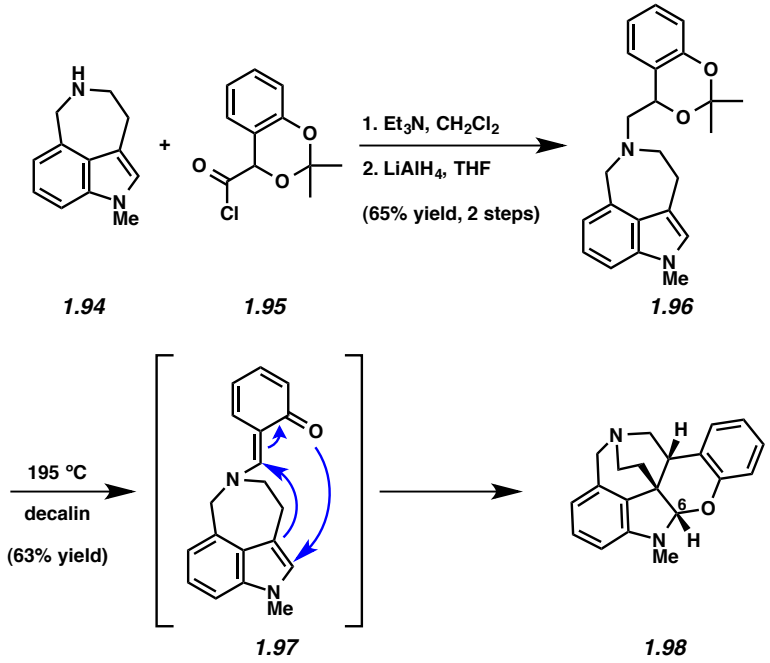
Scheme 1.22. Stoltz's biomimetic synthesis of **1.92** and **1.93**.



1.5.4 Funk's Biomimetic Approach to Nomofungin (1.82) and Communesin B (1.62)

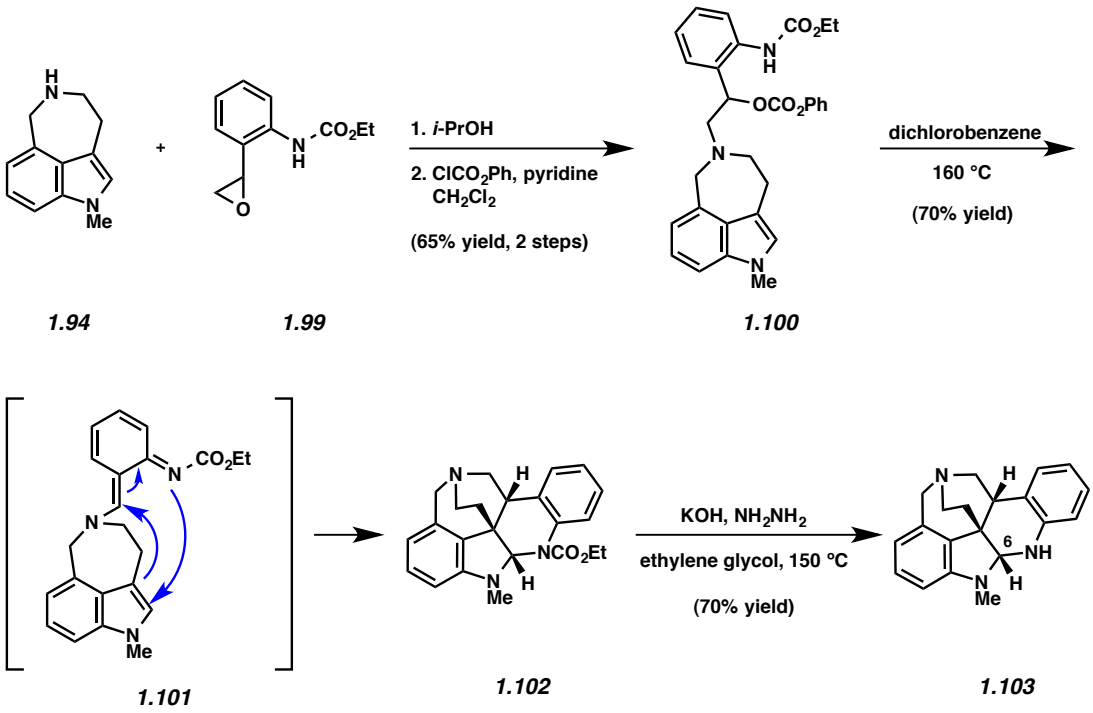
Funk performed complementary studies to investigate the structural assignments of communesin B (**1.62**) and nomofungin (**1.82**). Funk synthesized both *N,O*-acetal and aminal substrates, which were then compared to spectroscopic data for communesin B (**1.62**) and nomofungin (**1.82**).[69] *N,O*-acetal **1.98** was synthesized as shown in Scheme 1.23. Amine **1.94** was treated with Et_3N and **1.95** to give an intermediate amide. Upon reduction of the amide with $LiAlH_4$, benzodioxin **1.96** was formed as previously described.[70] Thermolysis of **1.96** facilitated an intramolecular hetero-Diels–Alder reaction (see transition structure **1.97**), analogous to Stoltz's method, to furnish *N,O*-acetal **1.98**.

Scheme 1.23. Funk's synthesis of the proposed nomofungin core.

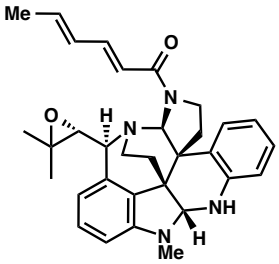


As shown in Scheme 1.24, the corresponding aminal was synthesized in an analogous fashion.[69] Nucleophilic attack of amine **1.94** onto epoxide **1.99**, gave an alcohol intermediate. Subsequent treatment with phenyl chloroformate yielded **1.100**. Thermolysis of **1.100** generated diene **1.101**, which underwent the hetero-Diels–Alder reaction to deliver aminal **1.102**. Cleavage of the ethyl carbamate with potassium hydroxide and hydrazine furnished aminal **1.103**.[69]

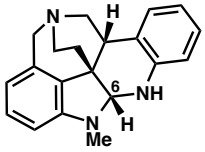
Scheme 1.24. Funk's synthesis of the communesin core.



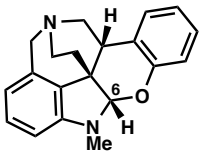
A careful inspection of the ^{1}H and ^{13}C NMR spectra for *N,O*-acetal **1.98** and aminal **1.103** showed that aminal **1.103** was more closely aligned with the reported data for nomofungin (**1.82**) and communesin B (**1.62**), as summarized in Figure 1.6. Careful examination of both ^{1}H NMR spectrums showed that the diagnostic H6 resonance for *N,O*-acetal **1.98** had a major discrepancy compared to the isolation data. *N,O*-acetal **1.98** had a C6 ^{1}H NMR shift of δ 5.4 (in $CDCl_3$), which was significantly downfield compared to the H6 resonance reported (δ 4.7, $CDCl_3$). Meanwhile, aminal **1.103** showed a ^{1}H NMR resonance at H6 of δ 4.6. The ^{13}C resonance at C6 for *N,O*-acetal **1.98** also was in major disagreement (δ 101.1 vs 82.4, respectively in $CDCl_3$). Meanwhile, the ^{13}C NMR data more strongly matched at C6 for aminal **1.103** (δ 84.4 vs. 82.4, respectively in $CDCl_3$).[69]



communesin B (1.62)



1.103



1.98

	Resonance of *1.62*	**Resonance of *1.103***	**Resonance of *1.98***
(^{1}H NMR)	4.7 ppm	4.6 ppm	5.4 ppm
(^{13}C NMR)	82.4 ppm	84.4 ppm	101.1 ppm

Figure 1.6. Spectroscopic comparison of **1.98** and **1.103**.

1.6. Future Directions in Communesin Biosynthesis

To date, there has only been a preliminary examination of the biosynthetic pathway of the communesin alkaloids. All of this work has heavily depended on chemical methods, either in the context of providing isotopically labeled compounds or providing synthetic compounds to verify natural product structure. Also, the success of the biomimetic approaches of Stoltz and Funk lend some credibility to the notion that the communesins may arise from Diels–Alder type processes. Although there is a general consensus about the role of tryptophan (**1.11**) and aurantioclavine (**1.83**) as biosynthetic intermediates, there has not been biosynthetic evidence beyond the work mentioned herein. With the remarkable technological advances utilized earlier to discern notoamide biosynthesis, we expect a detailed biosynthetic pathway for the communesins will be reported in due course.

1.7. Conclusion

The biosynthetic origin of fungal indole alkaloids, with their complex structures and alluring biological profiles, is a broad and ever-developing field. With respect to the brevianamides, notoamides, and communesin alkaloids, the foregoing sections highlight the various efforts to understand the biosynthesis of these intriguing molecules. Throughout this discussion, emphasis has been placed on synthetic chemistry and its impact on our understanding of biosynthetic pathways. There are still many unanswered questions that will likely be actively investigated in the years to come with the help of new technological advances and the enabling power of synthetic chemistry.

1.8 Notes and References

(1) Li, S.-M. *Nat. Prod. Rep.* **2010**, *27*, 57–78.

(2) Bandini, M.; Eichholzer, A. *Angew. Chem. Int. Ed.* **2009**, *48*, 9608–9644.

(3) Winter, J. M.; Tang, Y. *Curr. Opin. Biotech.* **2012**, *5*, 736–743.

(4) Qian-Cutrone, J.; Huang, S.; Shu, Y.-Z.; Vyas, D.; Fairchild, C.; Menende, A.; Krampitz, K.; Dalterio, R.; Khlor, S. E.; Gao, Q. *J. Am. Chem. Soc.* **2002**, *124*, 14556–14557.

(5) Kito, K.; Ookura, R.; Kusumi, T.; Namikoshi, M.; Ooi, T. *Heterocycles* **2009**, *78*, 2101–2106.

(6) Cai, S.; Luan, Y.; Kong, X.; Zhu, T.; Gu, Q.; Li, D. *Org. Lett.* **2013**, *15*, 2168–2171.

(7) Yamazaki, M.; Fujimoto, H.; Okuyama, E.; Ohta, Y. *Mycotoxins* **1980**, *10*, 27–28.

(8) Yamazaki, M.; Okuyama, E.; Kobayashi, M.; Inoue, H. *Tetrahedron Lett.* **1981**, *22*, 135–136.

(9) Hayashi, H.; Nishimoto, Y.; Nozaki, H. *Tetrahedron Lett.* **1997**, *38*, 5655–5658.

(10) Hayashi, H.; Nishimoto, Y.; Akiyama, K.; Nozaki, H. *Biosci., Biotechnol., Biochem.* **2000**, *64*, 111–115.

(11) Williams, R. M.; Cox, R. J. *Acc. Chem. Res.* **2003**, *36*, 127–139 and references therein.

(12) Paterson, R. R. M.; Simmonds, M. J. S.; Kemmelmeier, C.; Blaney, W. M. *Mycol. Res.* **1990**, *94*, 538–542.

(13) Jamie, H.; Kilian, G.; Dyason, K.; Milne, P. J. *J. Pharm. Pharmacol.* **2002**, *54*, 1659–1665.

(14) Preciado, S.; Mendive-Tapia, L.; Torres-García, C.; Zamudio-Vásquez, R.; Soto-Cerrato, V.; Pérez-Tomás, R.; Albericio, R.; Nicolás, E.; Lavilla, R. *Med. Chem. Commun.* **2013**, *4*, 1171–1174.

(15) Schkeryantz, J. M.; Woo, J. C. G.; Siliphaivanh, P.; Depew, K. M.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 11964–11975.

(16) Zhao, L.; May, J. P.; Huang, J., Perrin, D. M. *Org. Lett.* **2012**, *14*, 90–93.

(17) Yin, S.; Yu, X.; Wang, Q.; Liu, X.-Q.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1649–1660.

(18) Gunawan, S.; Hulme, C. *Tetrahedron Lett.* **2013**, *54*, 4467–4470.

(19) Birch, A. J.; Wright, J. J. *Tetrahedron* **1970**, *26*, 2329–2344.

(20) Birch, A. J.; Wright, J. J. *Tetrahedron* **1972**, *28*, 2999–3008.

(21) Baldas, J.; Birch, A. J.; Russell, R. A. *J. Chem. Soc., Perkin Trans. 1* **1974**, 50–52.

(22) a) Adams, L. A.; Valente, M. W. N.; Williams, R. M. *Tetrahedron* **2006**, *62*, 5195–5200. b) Adams L. A.; Gray, C. R.; Williams, R. M. *Tetrahedron* **2004**, *45*, 4489–4493.

(23) Williams, R. M.; Glinka, T.; Kwast, E.; Coffman, H.; Stille, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 808–821.

(24) Williams, R. M.; Glinka, T. *Tetrahedron Lett.* **1986**, *27*, 3581–3584.

(25) Seebach, D.; Boes, M.; Naef, R.; Schweizer, W. B. *J. Am. Chem. Soc.* **1983**, *105*b, 5390–5398.

(26) Williams, R. M.; Glinka, T.; Kwast, E. *J. Am. Chem. Soc.* **1988**, *110*, 5927–5929.

(27) Coetzer, J. *Acta Crystallogr.* **1974**, *30*, 2254–2256.

(28) Somei, M.; Kawasawa, Y.; Kaneko, C. *Heterocycles* **1981**, *16*, 941–949.

(29) Sanz-Cervera, J. F.; Glinka, T.; Williams, R. M. *Tetrahedron* **1993**, *49*, 8471–8482.

(30) Porter, A. E. A.; Sammes, P. G. *J. Chem. Soc., Chem. Commun.* **1970**, 1103.

(31) Domingo, L. R.; Sanz-Cervera, J. F.; Williams, R. M.; Picher, M. T.; Marco, J. A. *J. Org. Chem.* **1997**, *62*, 1662–1667.

(32) Williams R. M.; Sanz-Cervera, J. F.; Sancenon, F.; Marco, J. A.; Halligan, K. M. *Bioorg. Med. Chem.* **1998**, *6*, 1233–1241.

(33) Kato, H.; Yoshida, T.; Tokue, T.; Nojiri, Y.; Hirota, H.; Ohta, T.; Williams, R. M.; Tsukamoto, S. *Angew. Chem. Int. Ed.* **2007**, *46*, 2254–2256.

(34) Tsukamoto, S.; Umaoka, H.; Yoshikawa, K.; Ikeda, T.; Hirota, H. *J. Nat. Prod.* **2010**, *73*, 1438–1440.

(35) Tsukamoto, S.; Kato, H.; Samizo, M.; Nojiri, Y.; Onuki, H.; Hirota, H.; Ohta, T. *J. Nat. Prod.* **2008**, *71*, 2064–2067.

(36) Grubbs, A. W.; Artman, G. D.; Tsukamoto, S.; Williams, R. M. *Angew. Chem. Int. Ed.* **2007**, *46*, 2257–2261.

(37) Greshock, T. J.; Grubbs, A. W.; Tsukamoto, S.; Williams, R. M. *Angew. Chem. Int. Ed.* **2007**, *46*, 2262–2265.

(38) Ding, Y.; R. de Wet, J.; Cavalcoli, J.; Li, S.; Greshock, T. J.; Miller, K. A.; Finefield, J. M.; Sunderhaus, J. D.; McAfoos, T. J.; Tsukamoto, S.; Williams, R. M.; Sherman, D. H. *J. Am. Chem. Soc.* **2010**, *132*, 12733–12740.

(39) Birch, A. J.; Wright, J. J. *J. Chem. Soc. D,* **1969**, *12*, 644–645.

(40) Sunderhaus, J. D.; Sherman, D. H.; Williams, R. M. *Isr. J. Chem.* **2011**, *51*, 442–452.

(41) Finefield, J. M.; Williams, R. M. *J. Org. Chem.* **2010**, *75*, 2785–2789.

(42) McAfoos, T. J.; Li, S.; Tsukamoto, S.; Williams, R. M. *Heterocycles* **2010**, *82*, 461–463.

(43) Sunderhaus, J. D.; McAfoos, T. J.; Finefield, J. M.; Kato, H.; Li, S.; Tsukamoto, S.; Sherman, D. H.; Williams, R. M. *Org. Lett.* **2013**, *15*, 22–25.

(44) Somei, M.; Karasawa, Y.; Kaneko, C. *Heterocycles* **1981**, *16*, 941–942.

(45) Chantarasriwong, O.; Cho, W. C.; Batova, A.; Chavasiri, W.; Moore, C.; Rheingold, A. L.; Theodorakis, E. A. *Org. Biomol. Chem.* **2009**, *7*, 4886–4894.

(46) Tsukamoto, S.; Kato, H.; Greshock, T. J.; Hirota, H.; Ohta, T.; Williams, R. M. *J. Am. Chem. Soc.* **2009**, *131*, 3834–3835.

(47) Li, S.; Srinivasan, K.; Tran, H.; Yu, F.; Finefield, J. M.; Sunderhaus, J. D.; McAfoos, T. J.; Tsukamoto, S.; Williams, R. M.; Sherman, D. H. *MedChemComm* **2012**, *3*, 987–996.

(48) Finefield, J. M. Greshock, T. J.; Sherman, D. H.; Tsukamoto, S.; Williams, R. M. *Tetrahedron Lett.* **2011**, *52*, 1987–1989.

(49) Finefield, J. M.; Sherman, D. H.; Tsukamoto, S.; Williams, R. M. *J. Org. Chem.* **2011**, 5954–5958.

(50) Greshock, T. J.; Grubbs, A. W.; Jiao, P.; Wicklow, D. T.; Gloer, J. B.; Williams, R. M. *Angew. Chem. Int. Ed.* **2008**, *47*, 3573–3577.

(51) Tsukamoto, S.; Umaoka, H.; Yoshikawa, K.; Ikeda, T.; Hirota, H. *J. Nat. Prod.* **2010**, *73*, 1438–1440.

(52) Williams, J. M.; Williams, R. M. *J. Org. Chem.* **2010**, *75*, 2785–2789.

(53) Finefield, J. M.; Kato, H.; Greshock, T. J.; Sherman, D. H.; Tsukamoto, S.; Williams, R. M. *Org. Lett.* **2011**, *13*, 3802–3805.

(54) Numata, A.; Takahashi, C.; Ito, Y.; Takada, T.; Kawai, K.; Usami, Y.; Matsumura, E.; Imachi, M.; Ito, T.; Hasegawa, T. *Tetrahedron Lett.* **1993**, *34*, 2355–2358.

(55) a) Hayashi, H.; Matsumoto, H.; Akiyama, K. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 753–756. b) Dalsgaard, P. W.; Blunt, J. W.; Munro, M. H. G.; Frisvad, J. C.; Christophersen, C. *J. Nat. Prod.* **2005**, *68*, 258–261. c) Kerzaon, I.; Pouchus, Y. F.; Monteau, F.; Le Bizec, B.; Nourrisson, M.-R.; Biard, J.-F.; Grovel, O. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 3928–3938. d) Jadulco, R.; Edrada, R. A.; Ebel, R.; Berg, K.; Schaumann, K.; Wray, V.; Steube, K.; Proksch, P. *J. Nat. Prod.* **2004**, *67*, 78–81.

(56) Verbitski, S. M.; Mayne, C. L.; Davis, R. A.; Concepcion, G. P.; Ireland, C. M. *J. Org. Chem.* **2002**, *67*, 7124–7126.

(57) Ratnayake, A. S.; Yoshida, W. Y.; Mooberry, S. L.; Hemscheidt, T. K. *J. Org. Chem.* **2001**, *66*, 8717–8721.

(58) For select approaches to the communesins, see: a) Schammel, A. W.; Chiou, G.; Garg, N. K. *Org. Lett.* **2012**, *14*, 4556–4559. b) Zuo, Z.; Ma, D. *Angew. Chem. Int. Ed.* **2011**, *50*, 12008–12011. c) Zuo, Z.; Xie, W.; Ma, D. *J. Am. Chem. Soc.* **2010**, *132*, 13226–13228. d) Yang, J.; Wu, H.; Shen, L.; Qin, Y. *J. Am. Chem. Soc.* **2007**, *129*, 13794–13795. e) Wu, H.; Xue, F.; Xiao, X.; Qin, Y. *J. Am. Chem. Soc.* **2010**, *132*, 14052–14054. f) Zuo, Z.; Ma, D. *Angew. Chem. Int. Ed.* **2011**, *50*, 12008–12011. g) Zuo, Z.; Xie, W.; Ma, D. *J. Am. Chem. Soc.* **2010**, *132*, 13226–13228.

(59) Wigley, L. J.; Mantle, P. G.; Perry, D. A. *Phytochemistry* **2006**, *67*, 561–569.

(60) Matthew, S. J.; Jandu, S. K.; Leatherbarrow, R. J. *Biochemistry* **1993**, *32*, 657–662.

(61) Wigley, L. J.; Perry, D. A.; Mantle, P. G. *Mycol Res* **2008**, *112*, 131–137.

(62) May, J. A.; Zeidan, R. K.; Stoltz, B. M. *Tetrahedron Lett.* **2003**, *44*, 1203–1205.

(63) Yamada, S.; Shiori, T.; Itaya, T.; Hara, T.; Matsueda, P. *Chem. Pharm. Bull.* **1965**, *13*, 88–93.

(64) This partial biosynthetic pathway is also proposed on the basis of the known biosynthesis of Catharanthaceous alkaloids. For the use of a tryptophan decarboxylase see: Noe, W.; Mollenschott, C.; Berlin, J. *Plant Mol. Bio.* **1984**, *3*, 281–288.

(65) Ratanayake, A. S.; Yoshida, W. Y.; Mooberry, S. L.; Hemscheidt, T. K. *J. Org. Chem.* **2003**, *68*, 1640.

(66) May, J. A.; Stoltz, B. M. *Tetrahedron* **2006**, *62*, 5262–5271.

(67) For syntheses of aurantioclavine, see: a) Suetsugu, S.; Nishiguchi, H.; Tsukano, C.; Takemoto, Y. *Org. Lett*. **2014**, *ASAP* doi: 10.1021/ol4037314. b) Xu, Z.; Hu, W.; Liu, Q.; Zhang, L.; Jia, Y. *J. Org. Chem.* **2010**, *75*, 7626–7635. c) Brak, K.; Ellman, J. *Org. Lett*. **2010**, *12*, 2004–2007. d) Yamada, K.; Namerikawa, Y.; Haruyama, T.; Miwa, Y.; Yanada, R.; Ishikura, M. *Eur. J. Org. Chem.* **2009**, *33*, 5752–5759. e) Krishnan, S.; Bagdanoff, J. T.; Ebner, D. C.; Ramtohul, Y. K.; Tambar, U. K.; Stoltz, B. M. *J. Am. Chem. Soc.* **2008**, *130*, 13745–13754. e) Somei, M.; Yamada, F. *Heterocycles* **2007**, *74*, 943–950. f) Hegedus, L. S.; Toro, J. L.; Miles, W. H.; Harrington, P. J. *J. Org. Chem.* **1987**, *52*, 3319–3322.

(68) Steinhagen, H.; Corey, E. J. *Angew. Chem. Int. Ed*. **1999**, *38*, 1928–1931.

(69) Crawley, S. L.; Funk, R. L. *Org. Lett*. **2003**, *5*, 3169–3171.

(70) Clark, R. D.; Weinhardt, K. K.; Berger, J.; Fisher, L. E.; Brown, C.; MacKinnon, A. C.; Kilpatrick, A. T.; Spedding, M. *J. Med. Chem*. **1990**, *33*, 633–641.

CHAPTER TWO

Synthesis of (+)-Phenserine using an Interrupted Fischer Indolization Reaction

Alex W. Schammel, Grace Chiou, and Neil K. Garg.

J. Org. Chem. **2012**, 77, 725–728.

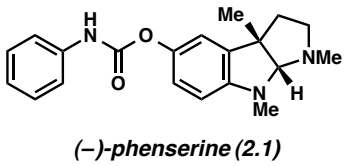
2.1 Abstract

A concise synthesis of the Alzheimer's therapeutic (+)-phenserine is described. The approach features an interrupted Fischer indolization to construct the pyrrolidinoindoline core, in addition to a classical resolution to arrive at phenserine in enantioenriched form.

2.2 Introduction

Small molecules that possess a pyrrolidinoindoline scaffold have drawn substantial interest from the scientific community.[1] One particular compound that has garnered significant attention is phenserine (**2.1**, Figure 2.1),[2,3] first accessed by semisynthesis from the naturally occurring alkaloid physostigmine by Polonovski in 1916.[4] The readily accessible (–)-enantiomer of **2.1**[5] has been explored as a potential therapeutic for Alzheimer's Disease and other neurodegenerative disorders.[1b,6] Despite it's promising biological profile, featuring acetylcholinesterase (AChE) inhibition and the inhibition of amyloid precursor protein (APP) expression, (–)-**2.1** ultimately failed in phase III clinical trials.[6b] The (+)-enantiomer (a.k.a., Posiphen®), however, possesses similar activities and is now undergoing clinical trials for the treatment of early phase Alzheimer's disease.[7,8] In Phase I/II trials oral Posiphen® was found to significantly lower two subtypes of APP by 44% and 45% in patients with mild cognitive

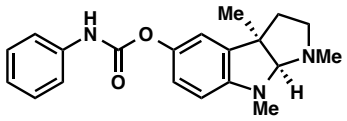
impairment.[8] Prompted by this promising data, and our own interest in indole-containing compounds,[9] we have developed a concise route to (+)-phenserine (**2.1**).



(–)-phenserine (2.1)

- potent AChE inhibitor
- APP synthase inhibitor

Failed in phase III clinical trials



(+)-phenserine (2.1)
"Posiphen®"

- modest AChE inhibitor
- APP synthase inhibitor

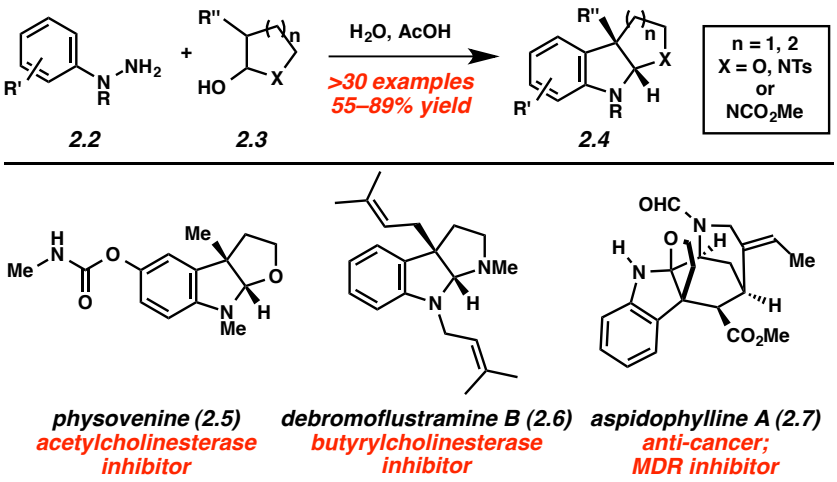
Currently in phase II clinical trials

Figure 2.1. (–)- and (+)-phenserine (**2.1**) and potential therapeutic value for the treatment of Alzheimer's disease.

2.3 The Interrupted Fischer Indolization Approach

Our laboratory has previously reported an efficient approach to pyrrolidinoindolines and related structures (Scheme 2.1).[9a,b] The method, termed 'interrupted Fischer indolization' [10] is broad in scope and has proven useful in the synthesis of physovenine (**2.5**), debromoflustramine B (**2.6**), and the complex pentacyclic alkaloid aspidophylline A (**2.7**), albeit in racemic form.[9b,d] We envisioned that (+)-phenserine (**2.1**) could also be prepared effectively using this methodology.

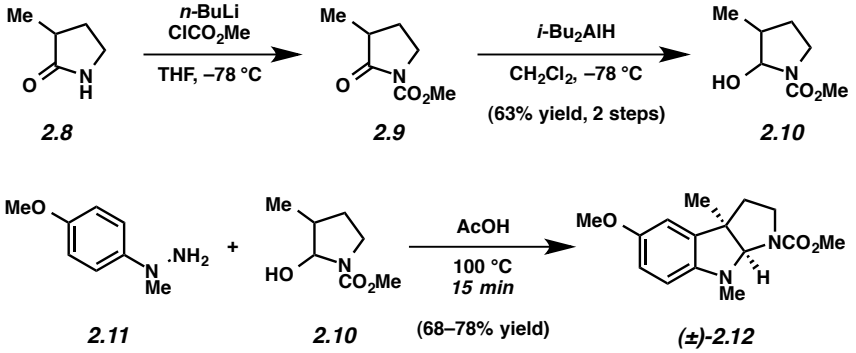
Scheme 2.1. Interrupted Fischer indolization methodology and fused indoline-containing targets.



2.4 Construction of the Pyrrolidinoindoline Core

In our initial approach, we sought to carry out a diastereoselective variant of the interrupted Fischer indolization using an arylhydrazine bearing a chiral auxiliary. Although the approach provided some diastereoselectivity in furoindoline formation,[9b] the corresponding transformation was less successful in the generation of pyrrolidinoindolines.[11] Thus, an alternative strategy for accessing (+)-phenserine (**2.1**) was pursued, whereby racemic material would be optically resolved at a late-stage. Pyrrolidinone **2.8**[12] was protected to furnish carbamate **2.9** (Scheme 2.2). Subsequent reduction with *i*-Bu_2AlH provided hemiaminal **2.10**. In the key interrupted Fischer indolization, treatment of hydrazine **2.11** with hemiaminal **2.10** in AcOH delivered pyrrolidinoindoline racemic **2.12** in 68–78% yield. The transformation is complete in just 15 min at 100 °C, and can be performed smoothly on gram scale.

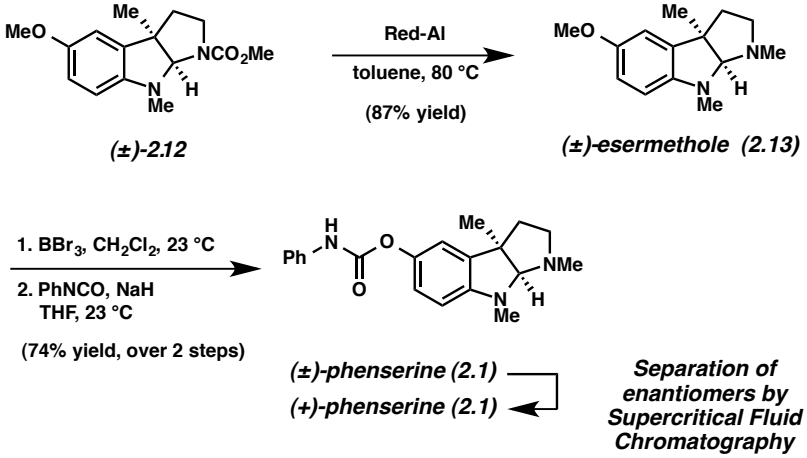
Scheme 2.2. Synthesis of hemiaminal **2.10** and the key interrupted Fischer indolization reaction to form **2.12**.



2.5 Synthesis of (±)-Phenserine and Separation of Enantiomers

With pyrrolidinoindoline (±)-**2.12** available, completion of the synthesis was achieved as shown in Scheme 2.3. Reduction of carbamate **2.12** with Red-Al gave esermethole (**2.13**), which is known to serve as a precursor to a variety of unnatural pyrrolidinoindolines.[2g] Using a modification of Overman's procedure,[2f] (±)-phenserine (**2.1**) was obtained by a deprotection/carbamylation sequence. Resolution of the racemate by preparative supercritical fluid chromatography (SFC) provided the desired (+)-enantiomer of phenserine (**2.1**), along with its antipode.[13]

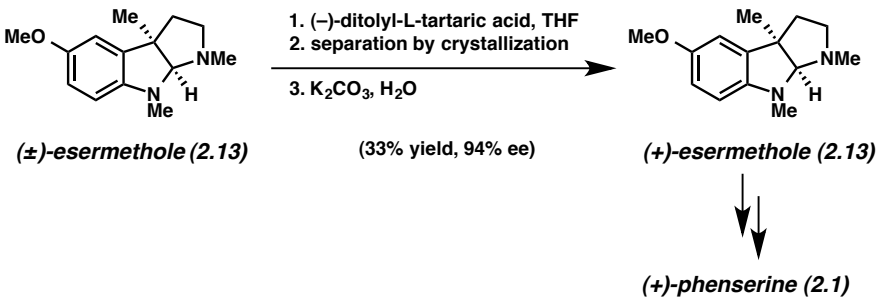
Scheme 2.3. Synthesis of (+)-phenserine (**2.1**) using preparative chiral SFC for separation of phenserine enantiomers.



2.6 Classical Resolution of (±)-Esermethole

As an alternative to SFC separation, we also explored classical resolution of a phenserine precursor. Surprisingly, no reports describing the resolution of esermethole are available. Nonetheless, we found that treatment of (±)-esermethole (**2.13**) with (–)-ditolyl-L-tartaric acid in THF led to crystallization (Scheme 2.4).[14] After collecting the filtrate, and free-basing the salt with aqueous K_2CO_3, (+)-esermethole (**2.13**) was obtained in 33% yield (94% ee). Elaboration of (+)-esermethole (**2.13**) to (+)-phenserine (**2.1**) proceeded smoothly using the two step procedure described above.

Scheme 2.4. Synthesis of (+)-phenserine (**2.1**) by classical resolution of racemic esermethole.



2.7 Conclusion

In summary, we have developed a concise and practical synthesis of (+)-phenserine (**2.1**), a compound currently in clinical trials for the treatment of Alzheimer's disease. The route features an interrupted Fischer indolization to construct the pyrrolidinoindoline core, in addition to a late-stage classical resolution. The strategy requires seven steps from pyrrolidinone **2.8** and provides efficient access to the desired target. We expect that our route will enable the synthesis of ample quantities of (+)-**2.1**, along with its derivatives.

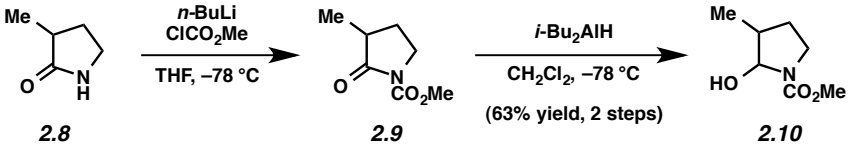
2.8 Experimental Section

2.8.1 Materials and Methods

Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of nitrogen using anhydrous solvents (either freshly distilled or passed through activated alumina columns). All commercially available reagents were used as received unless otherwise specified. (–)-ditolyl-L-tartaric acid was obtained from VWR (Alfa-Aesar). Diisobutylaluminium hydride, *n*-butyl lithium, methyl chloroformate, Red-Al, sodium hydride, phenylisocyanate, and boron tribromide were obtained from Sigma–Aldrich. Reaction temperatures were controlled using an IKAmag temperature modulator, and unless stated otherwise, reactions were performed at room temperature (rt, approximately 23 °C). Thin-layer chromatography (TLC) was conducted with EMD gel 60 F254 pre-coated plates (0.25 mm) and visualized using a combination of UV, anisaldehyde, iodine, and potassium permanganate staining. EMD silica gel 60 (particle size 0.040–0.063 mm) was used for flash column chromatography. ^{1}H NMR spectra were recorded on Bruker spectrometers (at 300 MHz, 400 MHz, or 500 MHz) and are reported relative to deuterated solvent signals. Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz) and integration. ^{13}C NMR spectra are reported in terms of chemical shift. For mixtures of diastereomers and rotamers, the major diastereomer is reported with the minor diastereomer in parentheses for both ^{1}H NMR and ^{13}C NMR spectra. IR spectra were recorded on a Perkin-Elmer 100 spectrometer and are reported in terms of frequency absorption (cm^{-1}). High resolution mass spectra were obtained from the UC Irvine Mass Spectrometry Facility. Determination of enantiopurity was carried out on a Mettler Toledo SFC (supercritical CO_2 liquid

chromatography) using a chiralpack OD-H column. Resolution of phenserine was conducted by Lotus Separations (Princeton, NJ) using a OJ-H column.

2.8.2 Experimental Procedures

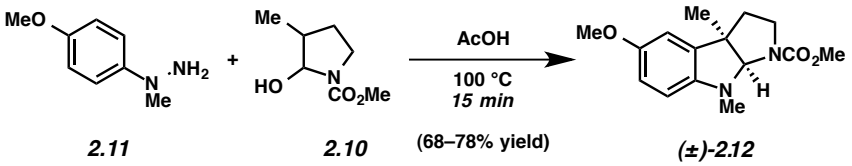


2.8 *2.9* *2.10*

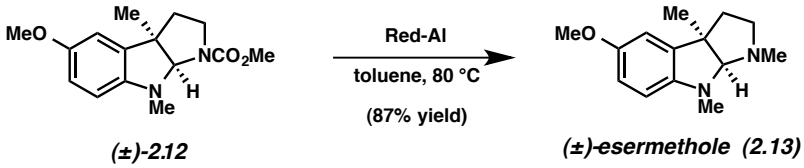
Hemiaminal 2.10: *n*-BuLi (6.82 mL, 7.43 mmol) was added to a solution of methyl pyrrolidinone **2.8** (670 mg, 6.68 mmol) in THF (30 mL) at –78 °C. The resulting mixture was stirred for 0.5 h and methyl chloroformate (679 μL, 8.79 mmol) was added. After stirring for 30 min, the mixture was warmed to 23 °C. The reaction was quenched with sat. aq. NH_4Cl (30 mL) and extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine (50 mL), dried over $MgSO_4$, and evaporated under reduced pressure. The crude product was used in the subsequent step without further purification.

To a solution of the crude product in CH_2Cl_2 (24 mL) at –78 °C was added *i*-Bu_2AlH (1.0 M in hexanes, 20 mL) over 1 h via syringe pump. The reaction was stirred for 30 min, quenched with sat. aq. NH_4Cl (40 mL), and then diluted with EtOAc (50 mL). Sat. aq. Na–K tartrate (150 mL) was added and the resulting mixture was stirred for 0.5 h. The layers were separated and the aqueous layer was extracted with EtOAc (3 x 100 mL). The combined organic layers were washed with brine (100 mL), dried over $MgSO_4$, and evaporated under reduced pressure. Purification by flash chromatography (1:1 hexanes:EtOAc) furnished hemiaminal **2.10** as a yellow oil (745 mg, 63% yield, over 2 steps, 58:42 mixture of diastereomers). R_f 0.2 (1:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, C_6D_6 @ 335 K) δ 5.23 (5.01) (s, 1H), 4.40 (4.40) (s, 1H), 3.48 (3.48) (s, 3H), 3.28–3.39 (3.28–3.39) (m, 1H), 3.28–3.33 (2.95–3.07) (m, 1H), 1.65–1.72 (1.99–2.03) (m, 1H), 1.44–1.49 (1.84–1.90) (m, 1H), 1.56–1.65 (1.01–1.09) (m, 1H), 0.96 (0.74) (d, J = 7.0, 3H); ^{13}C NMR (125 MHz, C_6D_6 @ 335 K) δ 155.6 (155.4), 82.4 (87.7), 51.4 (51.4),

44.1 (44.8), 37.9 (40.0), 29.4 (29.9), 12.4 (15.9); IR (film): 3421, 2961, 2885, 1683, 1450, 1379, 1121 cm^{-1}; HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_7H_{13}NO_3Na$, 182.0793; found 182.0789.

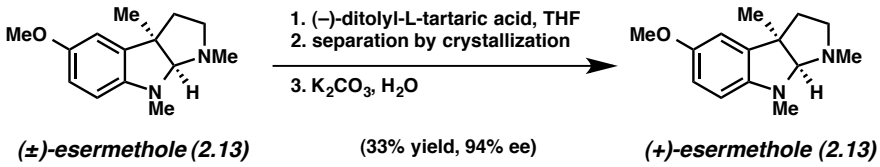


Pyrrolidinoindoline 2.12: Hydrazine **2.11** was freshly prepared before use.[15] A solution of hemiaminal **2.10** (1.20 g, 6.28 mmol) in acetic acid (12.5 mL) at 23 ºC, was stirred for 15 min. Phenylhydrazine **2.11** (1.33 g, 8.70 mmol) was then added and the resulting reaction mixture was heated to 100 ºC for 15 min. The reaction was then cooled to 23 ºC and was diluted with EtOAc (20 mL). The reaction mixture was quenched with sat. aq. $NaHCO_3$ (100 mL), and the layers were separated. The aqueous layer was extracted with EtOAc (3 x 25 mL). The combined organic layers were washed with brine (20 mL), dried over $MgSO_4$, and concentrated under reduced pressure to afford crude pyrrolidinoindoline **2.12**. Purification by flash chromatography (31:1:1 → 23:1:1 benzene:Et_2O:CH_2Cl_2) afforded pyrrolidinoindoline **2.12** as a red oil (1.36 g, 68% yield); R_f 0.2 (2:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, C_6D_6 @ 335 K): δ 6.61–6.64 (m, 2H), 6.18 (d, *J* = 8.0, 1H), 5.10 (bs, 1H), 3.50–3.60 (m, 1H), 3.50 (s, 3H), 3.46 (s, 3H), 2.95–3.01 (1H), 2.88 (s, 3H), 1.74–1.79 (m, 1H), 1.44–1.50 (m, 1H), 1.33 (s, 3H); ^{13}C NMR (125 MHz, C_6D_6 @ 335 K): δ 153.4, 144.9, 135.5, 127.8, 112.7, 109.85, 106.5, 89.6, 55.3, 51.4, 45.9, 38.4, 33.4, 33.4 23.9. Spectral data matches those previous reported.[16]



(±)-2.12 *(±)-esermethole (2.13)*

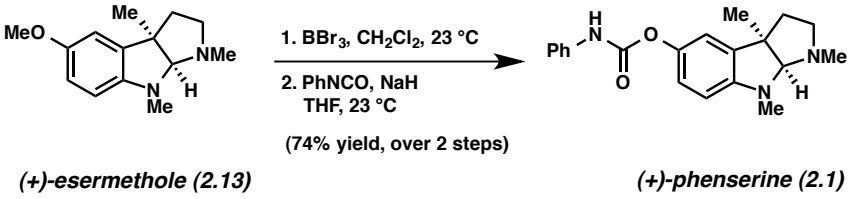
(±)-Esermethole (2.13): To a solution of pyrrolidinoindoline **2.12** (1.66 g, 6.01 mmol) in toluene (30 mL) was added Red-Al (65% in toluene, 5.90 mL) dropwise over 2 min. The resulting reaction mixture was heated to 80 °C. After 1 h, the reaction was cooled to room temperature, diluted with EtOAc (30 mL), and quenched with sat. aq. Na–K tartrate (40 mL). The resulting mixture was vigorously stirred for 1 h at 23 °C. The layers were separated and the aqueous layer was extracted with EtOAc (3 x 40 mL). The combined organic layers were washed with brine (75 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resulting residue was purified by flash column chromatography (1:1 EtOAc:hexanes with 1% Et_3N) to provide (±)-esermethole (**2.13**) (1.22 g, 87% yield) as brown oil. R_f 0.2 (1:9 MeOH:CH_2Cl_2); ^{1}H NMR (500 MHz, C_6D_6): δ 6.73 (d, *J* = 2.5, 1H), 6.70 (t, *J* = 2.5, 1H), 6.29 (d, *J* = 2.5, 1H), 3.96 (s, 1H), 3.44 (s, 3H), 2.61 (s, 3H), 2.54–2.56 (m, 2H), 2.36 (s, 3H) 1.77–1.85 (m, 2H), 1.32 (s, 3H); ^{13}C NMR (125 MHz, C_6D_6): 153.4, 146.9, 138.1, 112.2, 109.7, 107.4, 98.0, 55.1, 52.7, 52.5, 41.1, 37.5, 37.2, 27.2. Spectral data match those previously reported.[16]



(±)-esermethole (2.13) *(+)-esermethole (2.13)*

(+)-esermethole (2.13): To a solution of (±)-esermethole (**2.13**) (312 mg, 1.35 mmol) in THF (1.3 mL) was added (–)-ditolyl-L-tartaric acid (623 mg, 1.62 mmol) in 1.3 mL of THF. After 1 h, the resulting salt was heated to 80 °C with vigorous stirring for 40 min until the mixture became homogeneous. The solution was cooled to room temperature, while maintaining

vigorous stirring for 15 h. A solid precipitate formed, which was collected by filtration and washed with THF (3 x 2 mL) to afford the tartaric salt of (+)-esermethole (**2.13**). The resulting tartaric salt was suspended in EtOAc (30 mL) and washed with sat. aq. K_2CO_3 (30 mL) to afford (+)-esermethole (**2.13**) (131 mg, 33% yield, 94% ee). SFC (CHIRALPAK OD-H, CO_2/MeOH = 19/20 with 0.1% Et_2NH, flow 1.5 mL/min, 23 °C, detection at 254 nm) t_R 5.63 min (minor) and t_R 6.14 min (major); R_f 0.2 (1:9 MeOH:CH_2Cl_2); ^{1}H NMR (500 MHz, C_6D_6): δ 6.73 (d, J = 2.5, 1H), 6.70 (t, J = 2.5, 1H), 6.29 (d, J = 2.5, 1H), 3.96 (s, 1H), 3.44 (s, 3H), 2.61 (s, 3H), 2.54–2.56 (m, 2H), 2.36 (s, 3H) 1.77–1.85 (m, 2H), 1.32 (s, 3H); ^{13}C NMR (125 MHz, C_6D_6): δ 153.4, 146.9, 138.1, 112.2, 109.7, 107.4, 98.0, 55.1, 52.7, 52.5, 41.1, 37.5, 37.2, 27.2; $[\alpha]_D^{23.8}$ +130.0 (*c* 0.35, C_6H_6). Spectral data match those previously reported.[17]



(+)-esermethole (2.13) ***(+)-phenserine (2.1)***

(+)-phenserine (2.1): To a solution of (+)-esermethole (**2.13**) (74.2 mg, 0.32 mmol) in CH_2Cl_2 (4 mL) was added BBr_3 (0.15 mL, 1.60 mmol) as a solution in CH_2Cl_2 (1.60 mL) over 3 min. The reaction was stirred for 1.5 h, and concentrated by purging the reaction with N_2. The residue was dissolved in 5 mL of MeOH, and the resulting solution was stirred for 5 min and was then concentrated. The resulting residue was diluted with H_2O (5 mL) and sat. aq. $NaHCO_3$ (15 mL). The aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (20 mL), dried over Na_2SO_4, and concentrated under reduced pressure. The crude product was used in the subsequent step without further purification.

To a solution of the crude residue in THF (1.4 mL) was added NaH (1.1 mg, 0.03 mmol). The reaction was purged with N_2 for 1 min and then phenylisocyanate (38 mL, 0.35 mmol) was added dropwise over 5 sec. After stirring for 14 h, the reaction was quenched with sat. aq. $NaHCO_3$ (15 mL) and the aqueous layer was extracted with EtOAc (3 x 15 mL). The combined organic layers were washed with brine (50 mL), dried over Na_2SO_4, and concentrated under reduced pressure. The resulting residue was purified by flash chromatography (1:1 EtOAc:hexanes with 1% Et_3N) to afford (+)-phenserine (**2.1**) (79 mg, 81% yield) as a pink foam. R_f 0.3 (1:9 MeOH: CH_2Cl_2); ^{1}H NMR (500 MHz, C_6D_6): δ 7.37 (d, J = 8.0, 2H), 7.07–7.10 (m, 2H), 6.93–7.01 (m, 1H), 6.89 (s, 1H), 6.81–6.84 (m, 2H), 6.18 (d, J = 3.0, 1H), 3.99 (s, 1H), 2.53 (s, 3H), 2.44–2.50 (m, 2H), 2.30 (s, 3H), 1.71–1.80 (m, 2H), 1.24 (s, 3H); ^{13}C NMR (125 MHz, $C_6D_{6)}$): δ 152.2, 150.0, 143.0, 138.2, 137.6, 128.8, 128.8, 123.0, 120.6, 118.5, 116.4, 106.3, 97.7, 52.7, 52.3, 40.9, 40.9, 37.5, 36.4, 26.9; $[\alpha]_D^{23.8}$ +70.0 (*c* 0.005, $CHCl_3$); m.p. 138–140 °C. Spectral data match those previously reported.[2g]

(±)-Phenserine (**2.1**) was also prepared from (±)-esermethole (**2.13**) using the above procedure (74% yield). The racemic mixture was resolved by Supercritical Fluid Chromotography (SFC). SFC conditions: CHIRALPAK OJ-H (2 x 25 cm), CO_2/MeOH (85:15) with 1% Et_2NH, flow 50 mL/min, detection at 220 nm. Analatical detection was conducted under the same conditions, except at a flow rate of 3 mL/min on a 15 x 0.46 cm column: t_R 2.26 min ((–)-phenserine (**2.1**)) and t_R 3.00 min ((+)-phenserine (**2.1**)).

2.9 Notes and References

(1) a) Takano, S.; Ogasawara, K. *Alkaloids* **1989**, *36*, 225–251. b) Greig, N. H.; Pei, X.-F.; Soncrant, T. T. *Med. Res. Rev*. **1995**, *15*, 3–31.

(2) For examples of enantiospecific routes to phenserine and closely related pyrrolidinoindolines, see: a) Schönenberger, B.; Brossi, A.; George, C.; Flippen–Anderson, J. L. *Helv. Chim. Acta* **1986**, *69*, 283–287. b) Takano, S.; Moriya, M.; Ogasawara, K. *J. Org. Chem.* **1991**, *56*, 5982–5984. c) Pei, X.-F.; Greig, N. H.; Flippen–Anderson, J. L.; Bi, S.; Brossi, A. *Helv. Chim. Acta* **1994**, *77*, 1412–1422. d) Node, M.; Hao, X.-J.; Nishide, K.; Fuji, K. *Chem. Pharm. Bull*. **1996**, *44*, 715–719. e) Yu, Q.-S.; Pei, X.-F.; Holloway, H. W.; Greig, N. H. *J. Med. Chem.* **1997**, *40*, 2895–2901. f) Matsuura, T.; Overman, L. E.; Poon, D. J. *J. Am. Chem. Soc.* **1998**, *120*, 6500–6503. g) Yu, Q.-S.; Brossi, A. *Heterocycles* **1988**, *27*, 745–750. h) Huang, A.; Kodanko, J. J.; Overman, L. E. *J. Am. Chem. Soc.* **2004**, *126*, 14043–14053. i) Nishida, A.; Ushigome, S.; Sugimoto, A.; Arai, S. *Heterocycles* **2005**, *66*, 181–185. j) Trost, B. M.; Quancard, J. *J. Am. Chem. Soc.* **2006**, *128*, 6314–6315. k) Nakao, Y.; Ebata, S.; Yada, A.; Hiyama, T.; Ikawa, M.; Ogoshi, S. *J. Am. Chem. Soc.* **2008**, *130*, 12874–12875. l) Bui, T.; Syed, S.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2009**, *131*, 8758–8759. m) Trost, B. M.; Zhang, Y. *Chem. Eur. J.* **2011**, *17*, 2916–1922.

(3) For a recent synthesis of (±)-phenserine, see: Rigby, J. H.; Sidique, S. *Org. Lett.* **2007**, *9*, 1219–1221.

(4) Polonovski, M. *Bull. Soc. Chim. Fr.* **1916**, *19*, 47.

(5) Yu, Q.-S.; Schönenberger, B.; Brossi, A. *Heterocycles* **1987**, *26*, 1271–1275.

(6) a) Greig, N. H.; Sambamurti, K.; Yu, Q.-S.; Brossi, A.; Bruinsma, G. B.; Lahiri, D. K. *Curr. Alzeheimer Res.* **2005**, *2*, 281–290. b) Becker, R. E.; Greig, N. H. *Curr. Alzeheimer Res.* **2010**, *7*, 642–651.

(7) Lahiri, D. K.; Chen, D.; Maloney, B.; Holloway, H. W.; Yu, Q.-S.; Utsuki, T.; Giordano, T.; Sambamurti, K.; Greig, N. H. *J. Pharmacol. Exp. Ther*. **2007**, *320*, 396–396.

(8) QR Pharma, Inc. Home Page. http://www.qrpharma.com/ (accessed September 15, 2011).

(9) a) Boal, B. W.; Schammel, A. W.; Garg, N. K. *Org. Lett.* **2009**, *11*, 3458–3461. b) Schammel, A. W.; Boal, B. W.; Zu, L.; Mesganaw, T.; Garg, N. K. *Tetrahedron* **2010**, *66*, 4687–4695. c) Çelebi-Ölçüm, N.; Boal, B. W.; Huters, A. D.; Garg, N. K.; Houk, K. N. *J. Am. Chem. Soc*. **2011**, *133*, 5752–5755. d) Zu, L.; Boal, B. W.; Garg, N. K. *J. Am. Chem. Soc*. **2011**, *133*, 8877–8879.

(10) For the classic Fischer indole synthesis and select examples of related methodologies to construct fused indolines, see: a) Fischer, E.; Jourdan, F. *Ber.* **1883**, *16*, 2241–2245. b) Fischer, E.; Hess, O. *Ber.* **1884**, *17*, 559–568. c) Grandberg, I. I.; Zuyanova, T. I.; Afonina, N. I.; Ivanova, T. A. *Dokl. Akad. Nauk SSSR* **1967**, *176,* 583–585. d) Grandberg, I. I.; Tokmakov, G. P. *Khim. Geterotsikl. Soedin*. **1975**, 207–210. e) Takano, S.; Moriya, M.; Iwabuchi, Y.; Ogasawara, K. *Chem. Lett*. **1990**, 109–112. f) Takano, S.; Ogasawara, K.; Iwabuchi, R.; Moriya, M. JP 03112989 A 19910514, 1991. g) Rosenmund, P.; Gektidis, S.; Brill, H.; Kalbe, R. *Tetrahedron Lett.* **1989**, *30*, 61–62.

(11) Mesganaw, T.; Zu, L.; Garg, N. K. University of California, Los Angeles, CA. Unpublished work, 2010. For a chiral auxilary approach toward (–)-esermethole and other related pyrrolidinoindolines, see reference 2i.

(12) α-Methylpyrrolidinone **2.8** can be obtained from commercial sources or laboratory preparation; for preparation, see: a) Menezes, R.; Smith, M. B. *Synth. Commun.* **1988**, *18*, 1625–1636. b) Mader, M.; Helquist, P. *Tetrahedron Lett.* **1988**, *29,* 3049–3052.

(13) Separation was conducted using an OJ-H column and 15% MeOH (0.1% Et_2NH)/CO_2 solvent.

(14) Dibenzoyltartaric acid has previously been employed to resolve a similar pyrrolidinoindoline; see reference 2e.

(15) Srivastava, S.; Ruane, P. H.; Toscano, J. P.; Sullivan, M. B.; Cramer, C. J.;Chiapperino, D.; Reed, E. C.; Falvey, D. E. *J. Am. Chem. Soc*. **2000**, *122*, 8271–8278.

(16) Tsuji, R.; Nakagawa, M.; Nishida, A. *Heterocycles* **2002**, *58*, 587–594.

(17) Pallavicini, M.; Valoti, E.; Villa, L.; Lianza, F. *Tetrahedron: Asymmetry* **1994**, *5*, 111–116.

APPENDIX ONE

Spectra Relevant to Chapter Two:

Synthesis of (+)-Phenserine using an Interrupted Fischer Indolization Reaction

Alex W. Schammel, Grace Chiou, and Neil K. Garg.

J. Org. Chem. **2012**, *77*, 725–728.

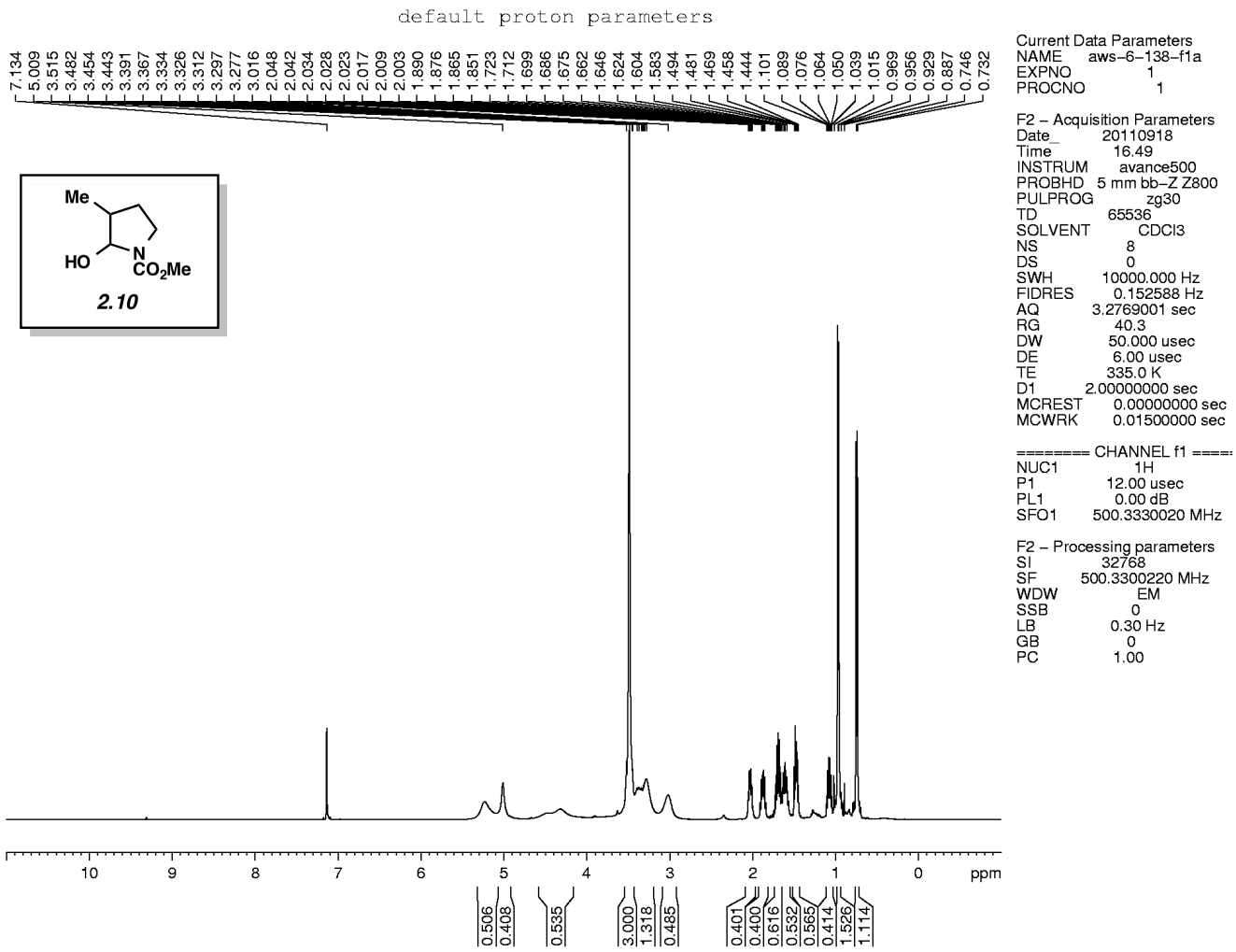


Figure A1.1 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **2.10**.

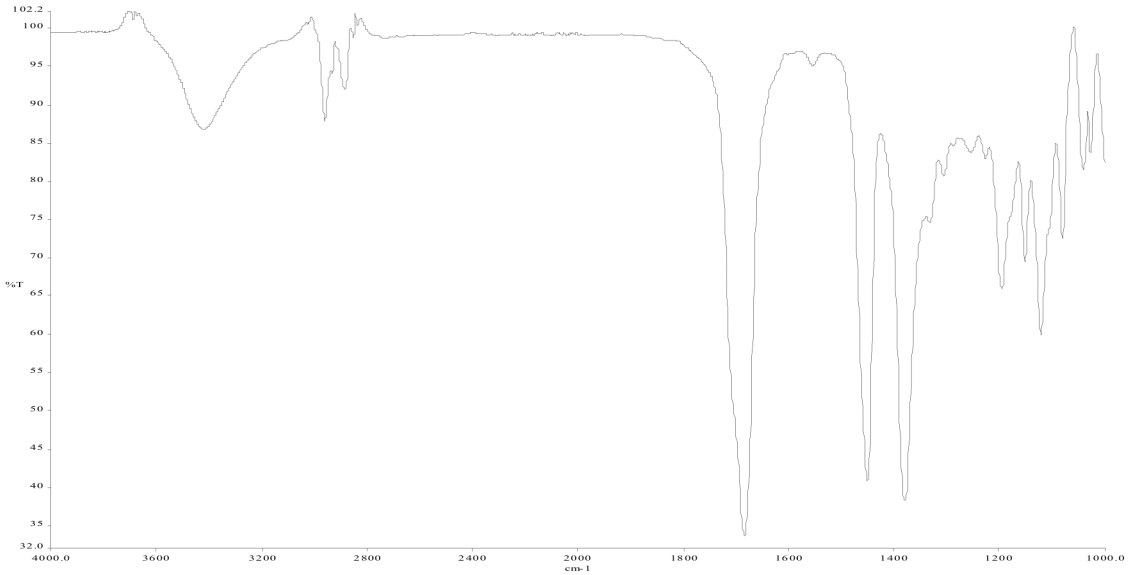


Figure A1.2 Infrared spectrum of compound **2.10**.

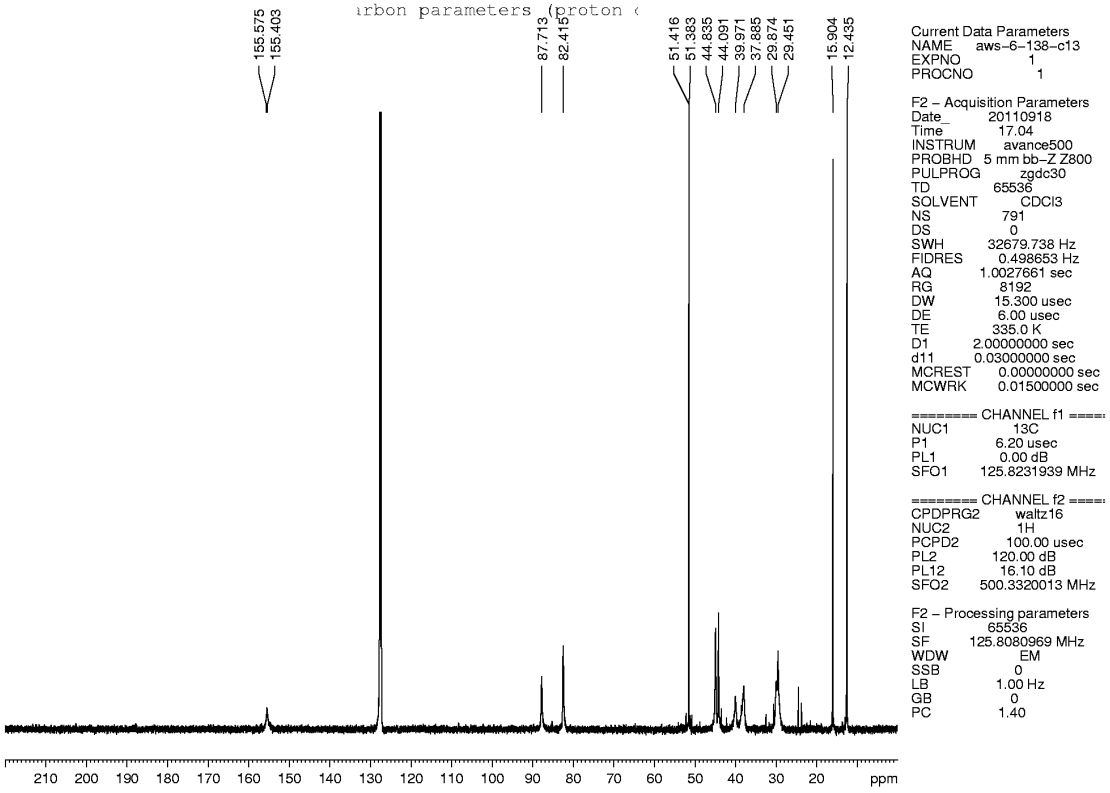


Figure A1.3 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **2.10**.

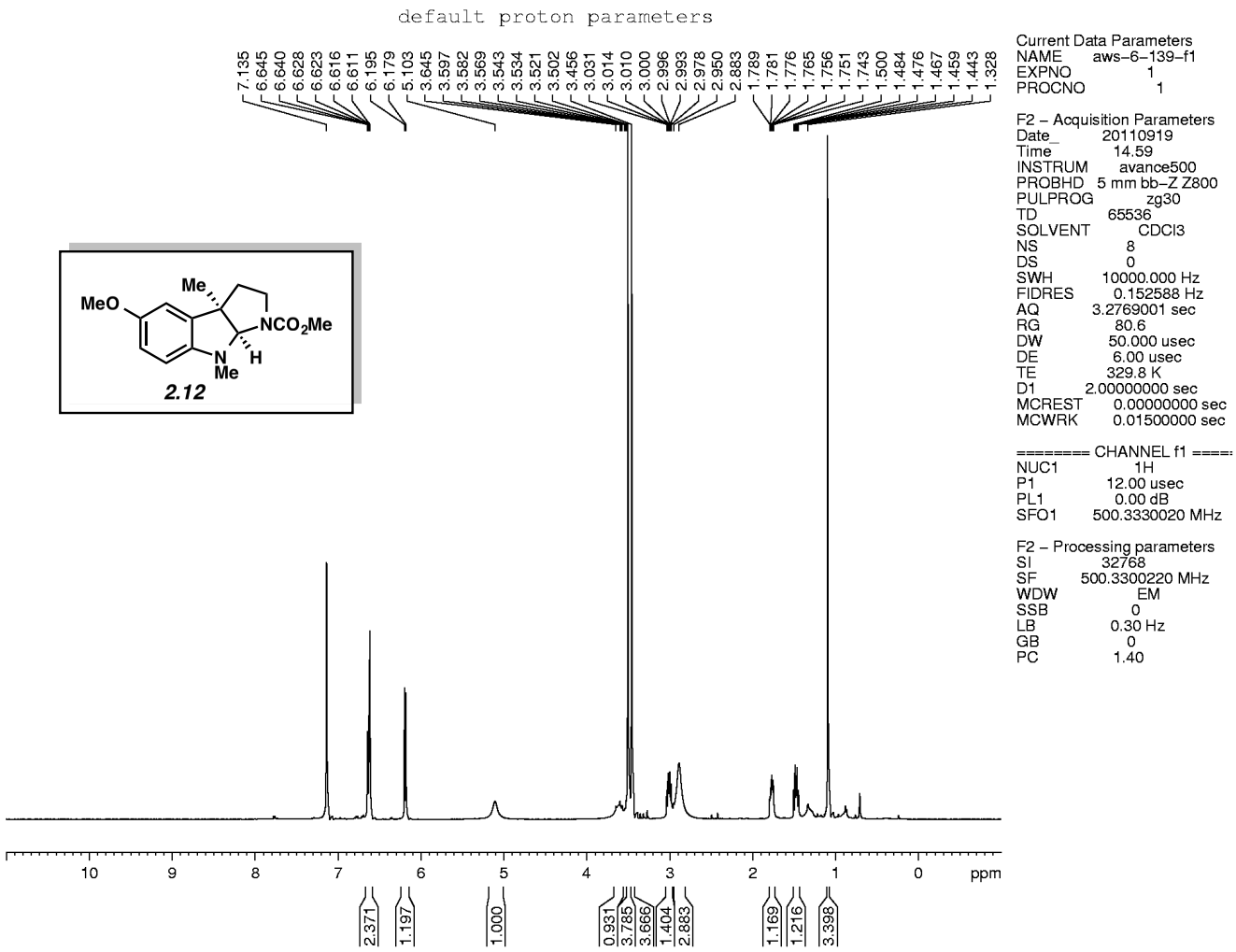


Figure A1.4 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **2.12**.

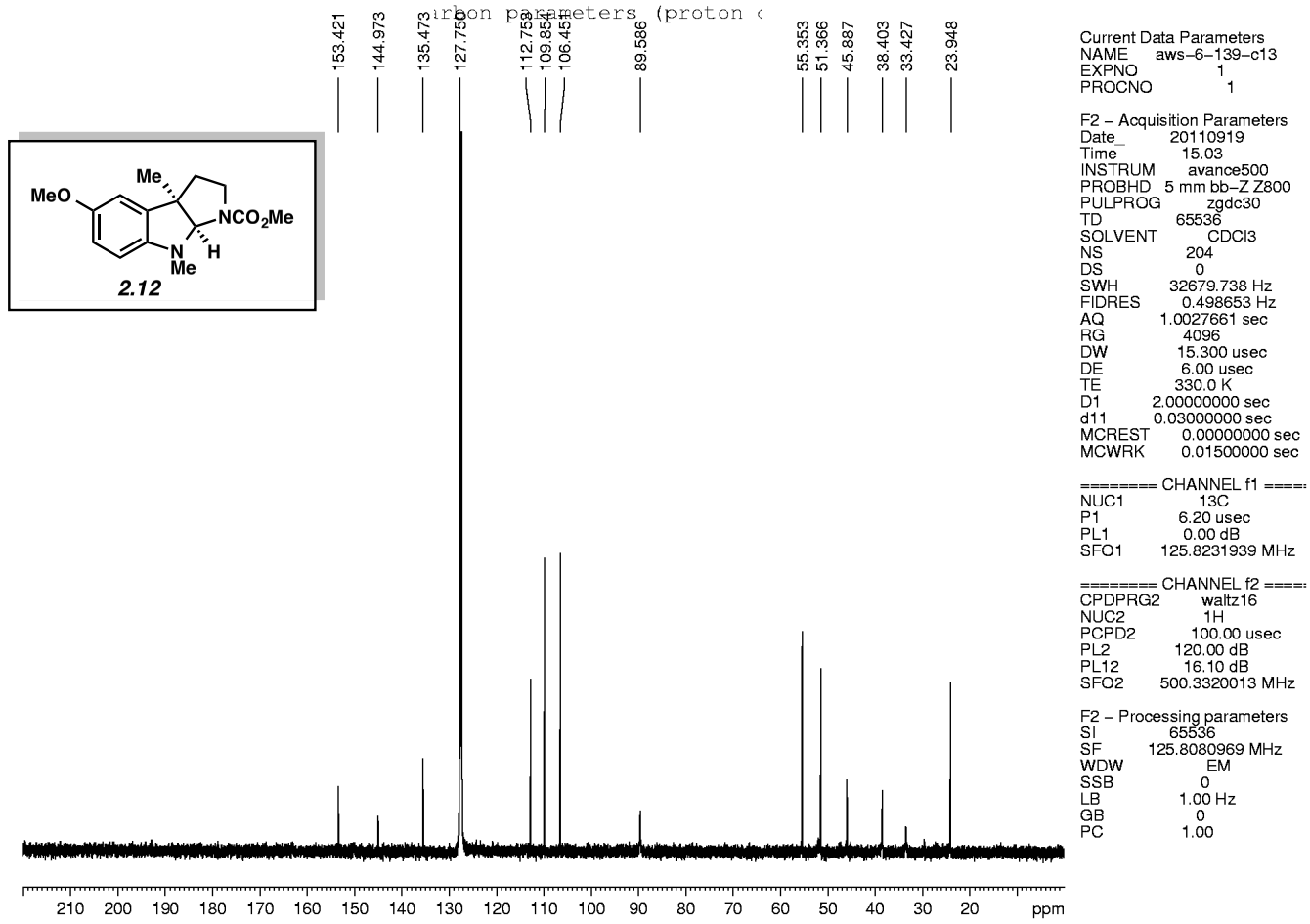


Figure A1.5 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **2.12**.

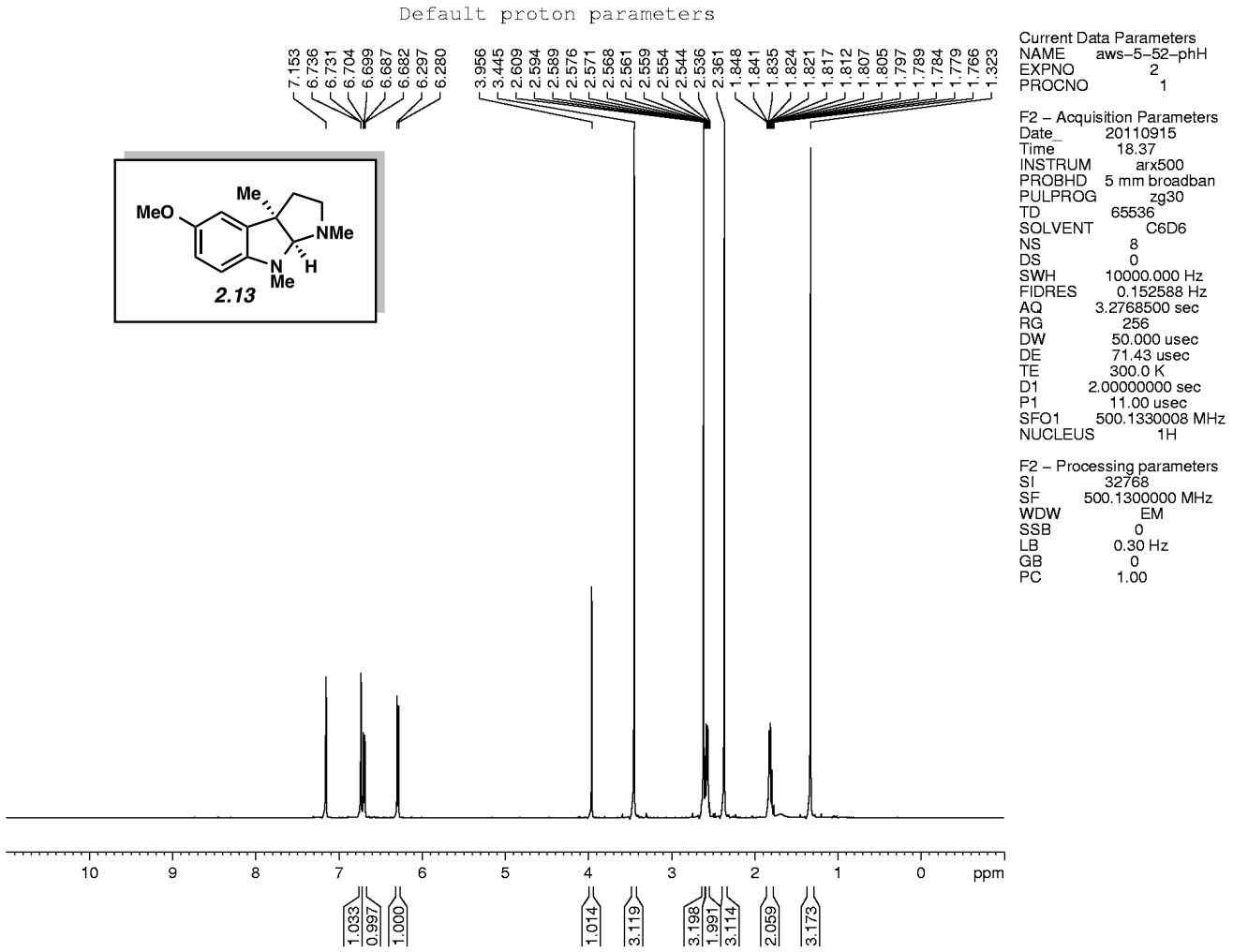


Figure A1.6 ^{1}H NMR (500 MHz,C_6D_6) of compound **2.13**.

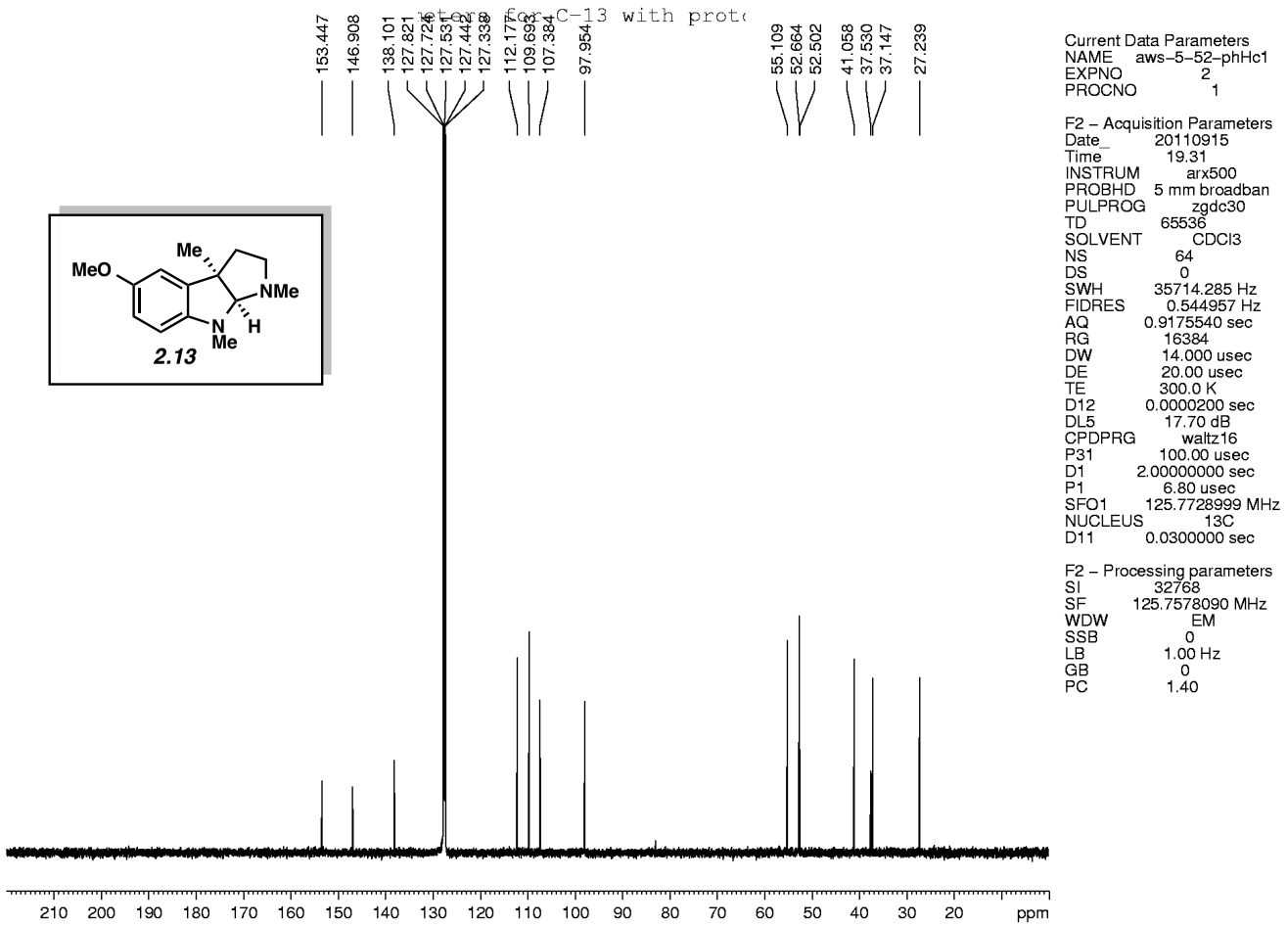


Figure A1.7 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **2.13**.

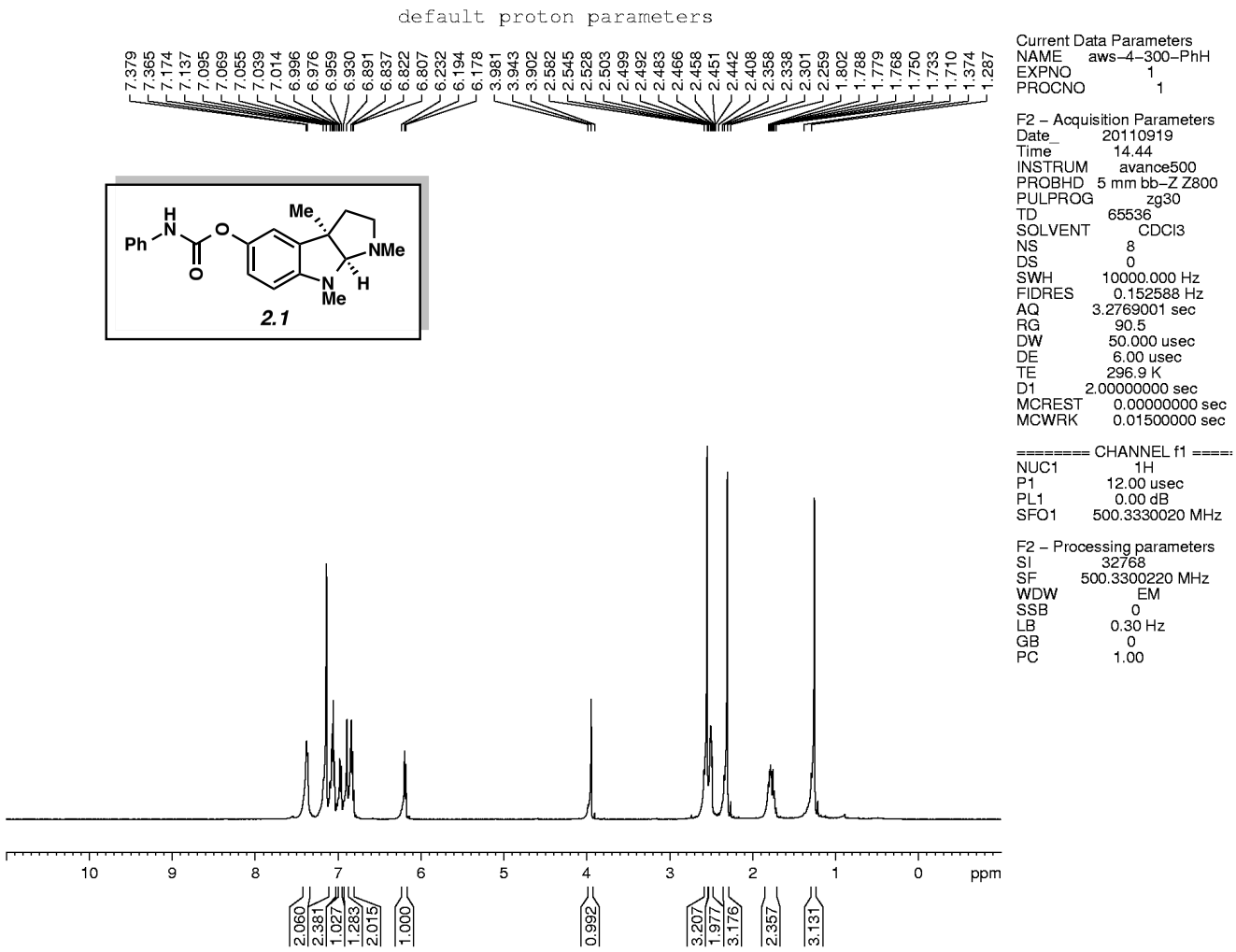


Figure A1.8 ^{1}H NMR (500 MHz, C_6D_6) of compound **2.1**.

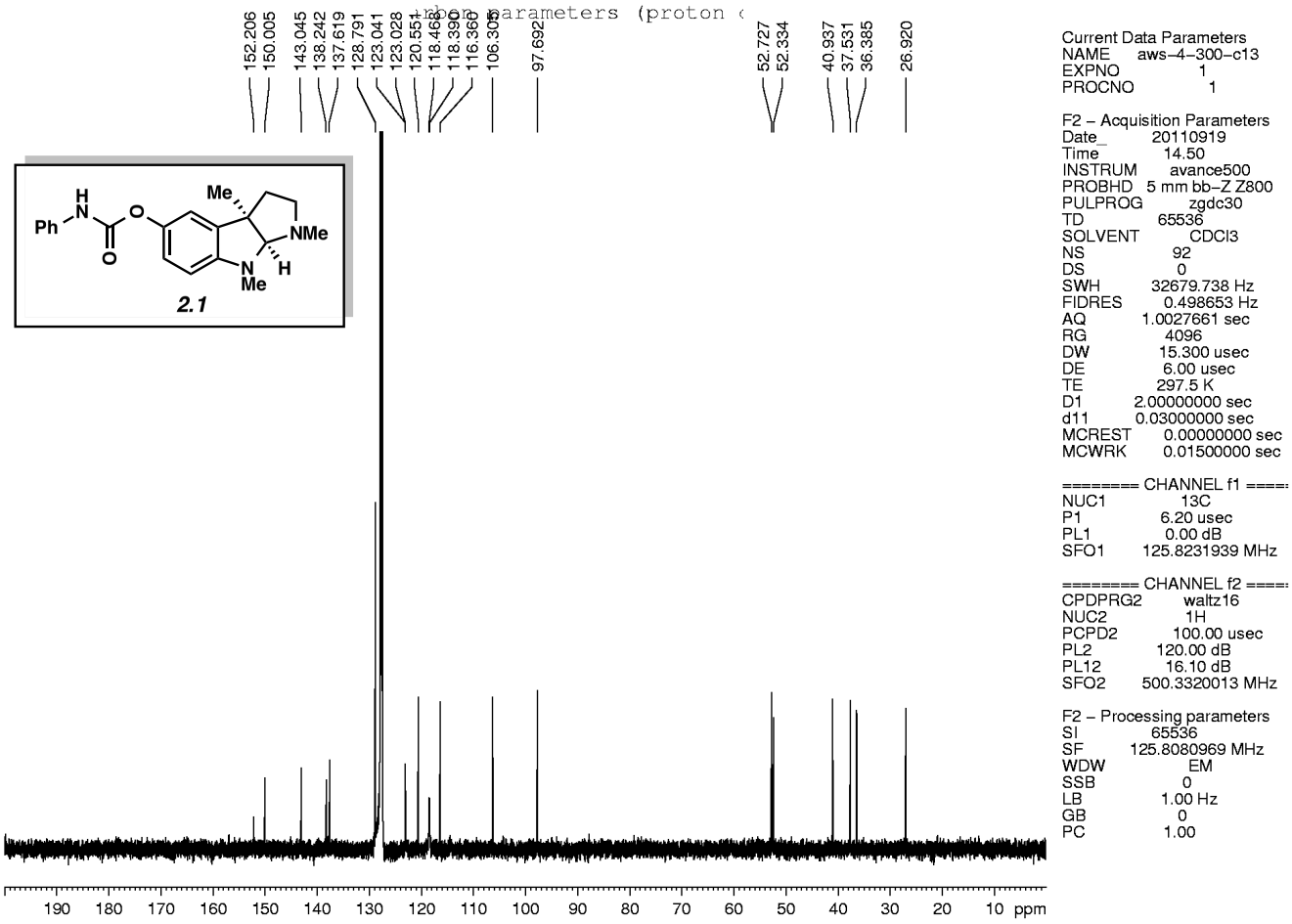
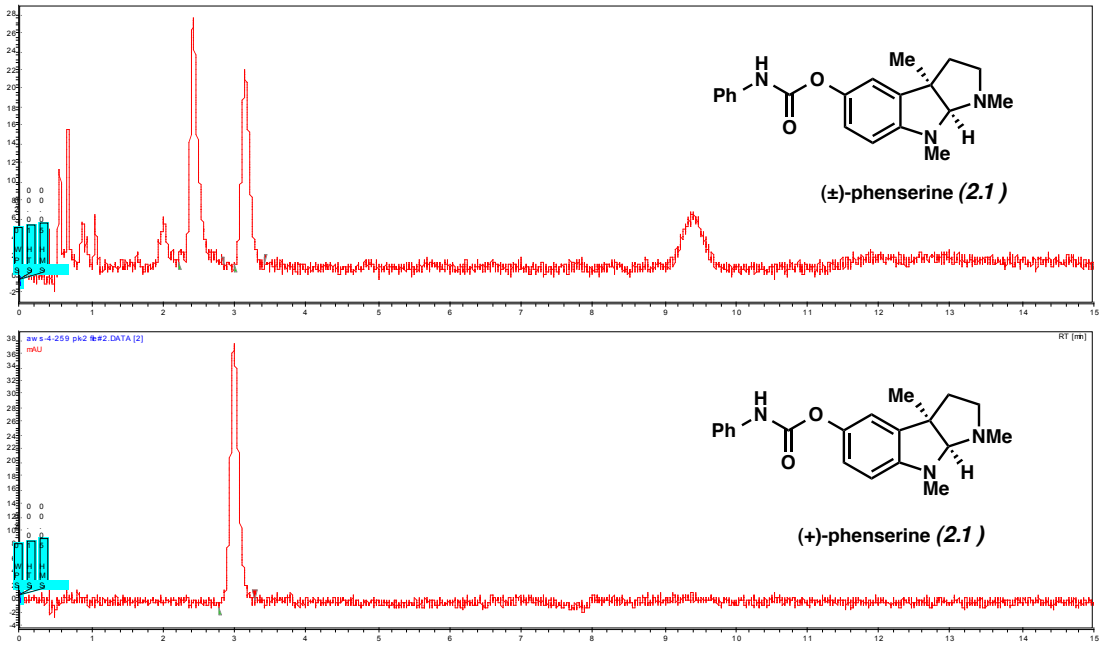


Figure A1.9 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **2.1**.

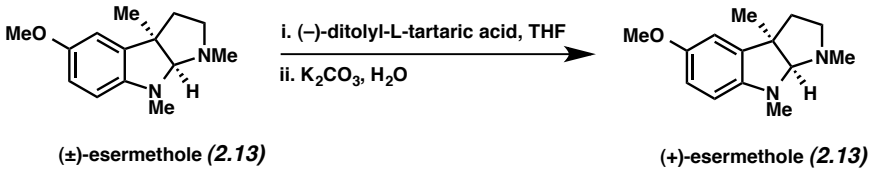
(±)-phenserine *(2.1)* (+)-phenserine *(2.1)*

SFC Preparative resolution of racemic phenserine (Lotus separations):
OJ-H (2 x 25 cm)
15% Methanol (0.1% diethylamine)/CO_2 100 bar
flow 50 mL/min, 23 °C, detection at 220 nm
Analytical method (Lotus separations):
OJ-H (15 x 0.46 cm)
flow 3 mL/min, 23 °C, detection at 220 nm
t_R 2.26 min ((–)-phenserine (**2.1**)) and t_R 3.00 min ((+)-phenserine(**2.1**))



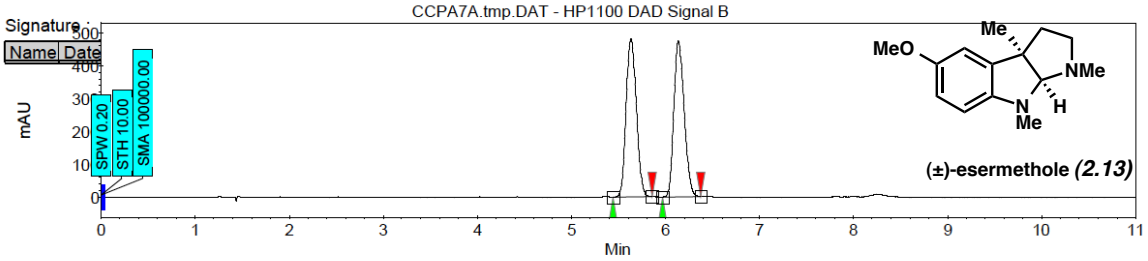
Index	Time (min)	Area (%)
Peak 1	2.26	0.00
Peak 2	3.00	100.00
Total		100.00

Figure A1.10. SFC traces of **2.1**.

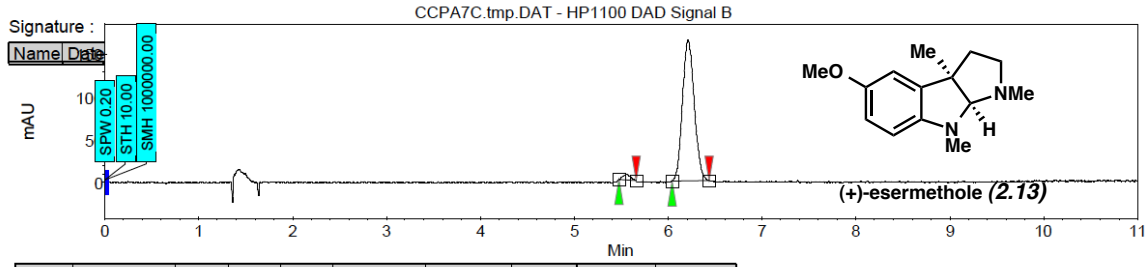


(±)-esermethole *(2.13)* (+)-esermethole *(2.13)*

SFC analytical method for determining the ee of (+)-esermethole:
CHIRALPAK OD-H column
CO_2/MeOH = 19/20 with 0.1% Et_2NH
flow 1.5 mL/min, 23 °C, detection at 254 nm
t_R 5.63 min ((–)-esermethole (**2.13**)) and t_R 6.14 min ((+)-esemethole (**2.13**))



Index	Name	Start	Time	End	RT Offset	Quantity	Height	Area	Area
		[Min]	[Min]	[Min]	[Min]	[% Area]	[µV]	[µV.Min]	[%]
1	UNKNOWN	5.44	5.63	5.86	0.00	50.17	481.3	61.0	50.167
2	UNKNOWN	5.97	6.14	6.38	0.00	49.83	474.5	60.6	49.833
Total						100.00	955.9	121.5	100.000



Index	Name	Start	Time	End	RT Offset	Quantity	Height	Area	Area
		[Min]	[Min]	[Min]	[Min]	[% Area]	[µV]	[µV.Min]	[%]
1	UNKNOWN	5.48	5.57	5.66	0.00	2.80	7.4	0.7	2.805
2	UNKNOWN	6.04	6.21	6.44	0.00	97.20	165.6	24.2	97.195
Total						100.00	173.0	24.9	100.000

Figure A1.11. SFC traces of **2.13**.

CHAPTER THREE

Interrupted Fischer Indolization Approach Toward the Communesin Alkaloids and Perophoramidine

Alex W. Schammel, Grace Chiou, Neil K. Garg

Org. Lett. **2012**, *14*, 4556–4559.

3.1 Abstract

We report a concise approach toward the total synthesis of the communesin alkaloids and perophoramidine. The strategy relies on the use of the interrupted Fischer indolization to build the tetracyclic indoline core of the natural products. Studies to probe the scope and limitations of this plan are presented. Although the methodology does not tolerate a C8-allyl substituent en route to the challenging vicinal quaternary stereocenters, variation at C7 and on the C ring is permitted.

3.2 Introduction

The communesin family of natural products and perophoramidine have been popular targets for chemical synthesis (Figure 3.1).[1] Communesins A (**3.1**) and B (**3.2**) were isolated in 1993 by Numata and co-workers from a *Penicillium* mold found growing on the marine algae *Enteromorpha intestinalis*.[2] Over the past decade, several additional communesin alkaloids (e.g., **3.3**–**3.8**) have been discovered,[3] in addition to perophoramidine (**3.9**).[4] Many of the communesins exhibit moderate insecticidal activity, whereas communesin B (**3.2**) also shows activity against the P-388 leukemia cell line (ED_{50} = 0.88 μM).[5] Furthermore, **3.9** is cytotoxic toward the HCT116 colon cancer cell line (IC_{50} = 60 μM).[iv]

The alkaloids **3.1**–**3.9** possess a number of structural features that render them attractive and challenging targets for total synthesis. Communesins **3.1**–**3.8** bear a heptacyclic skeleton, which contains a heterocycle-fused indoline core. Other characteristics of these natural products include the vicinal quaternary centers at C7 and C8 and two aminal linkages at C6 and C9.[6] Perophoramidine (**3.9**) is similar to the communesins, although it lacks the seven-membered G ring and has an opposite sense of relative stereochemistry at C7 and C8. Additionally, perophoramidine (**3.9**) is more highly oxidized in comparison to the communesins as noted by the aromatic halogenation pattern and presence of amidine functional groups instead of aminals. Synthetic endeavors toward these natural products have led to several promising routes,[7] in addition to a few recently completed total syntheses by the groups of Weinreb,[8] Funk,[9] Ma,[10] and Qin.[11]

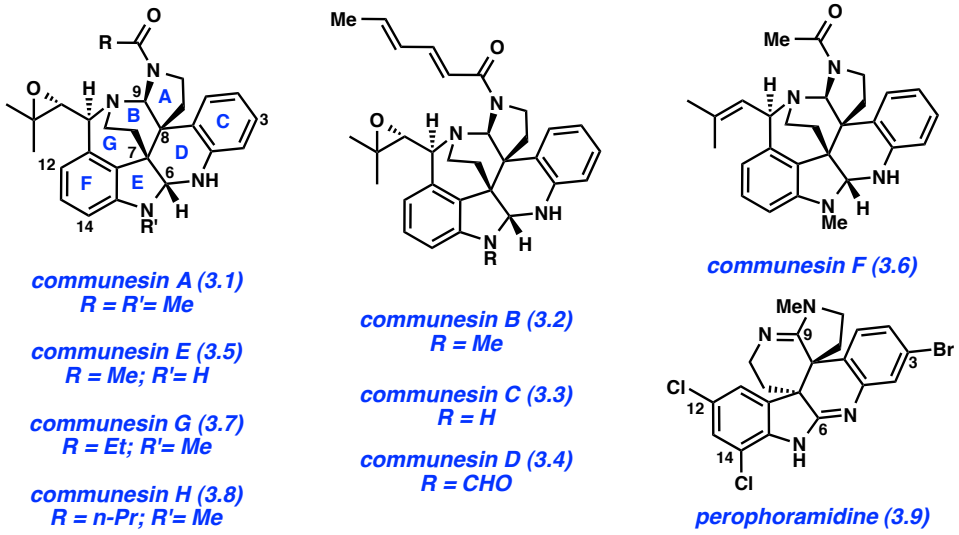


Figure 3.1. Communesins (**3.1**–**3.8**) and perophoramidine (**3.9**).

3.3 The Interrupted Fischer Indolization Approach

En route to the total synthesis of compounds **3.1**–**3.9** and a variety of other alkaloids, we have developed a cascade reaction that provides access to fused indoline scaffolds.[12] The transformation, termed the "interrupted Fischer[13,14,15] indolization", allows for aryl hydrazines **3.10** and latent aldehydes **3.11** to be converted to indoline products **3.12** (Figure 3.2). The reaction is operationally simple and broad in scope, requires mild reaction conditions, and has shown utility in synthesis.[12] In our initial attempts, we found that the interrupted Fischer indolization of *N*-methylphenylhydrazine (**3.13**) with *N,O*-acetal **3.14** efficiently delivered tetracycle **3.15**.[12a,b] Of note, this reaction provides rapid access to the 6,5,6,6 F–E–D–C ring system of the communesins and perophoramidine and provides a unique approach to these highly sought after alkaloids.

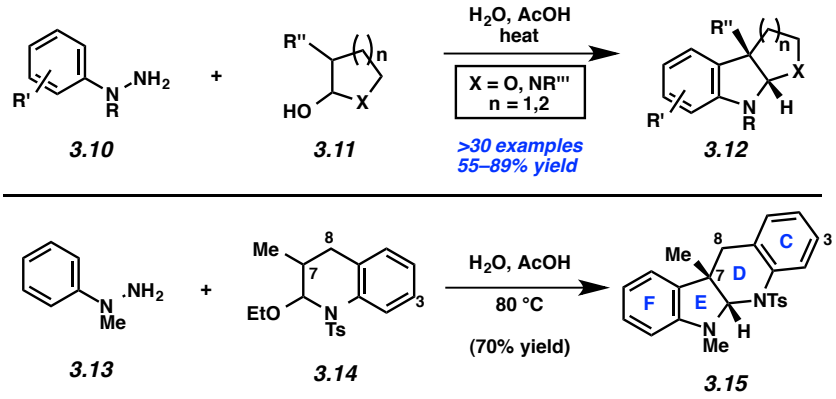


Figure 3.2. Interrupted Fischer indolization methodology and approach toward alkaloids **3.1**–**3.9**.

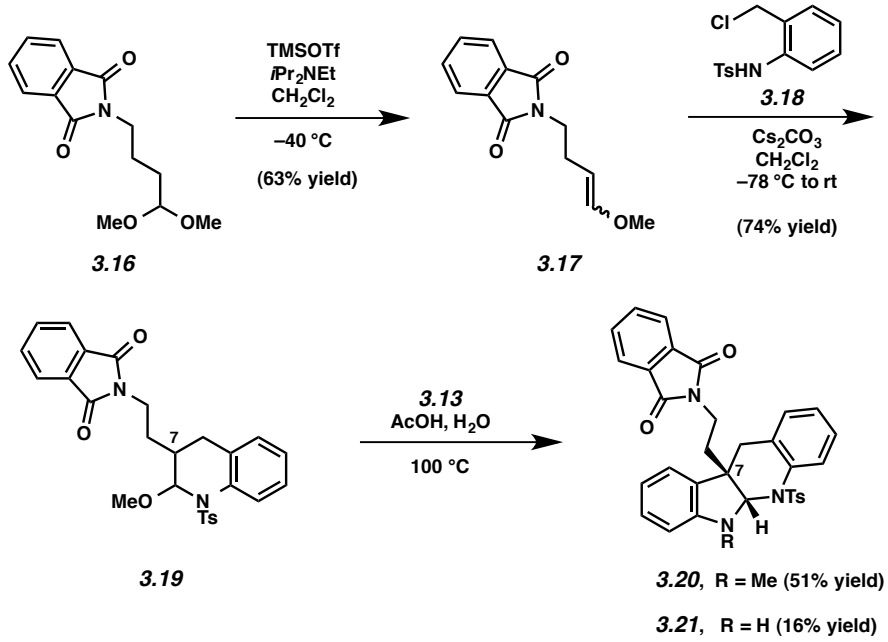
3.4 Construction of the CDEF Ring System of the Communesins and Perophoramidine

By further varying the reaction partners in the key interrupted Fischer indolization reaction, we envisioned that more highly functionalized derivatives of **3.15** could be readily accessible. This would not only provide advanced intermediates for our total synthesis

objectives, but would also provide an opportunity to study the scope and limitations of the interrupted Fischer indolization methodology. Herein, we report the outcome of the interrupted Fischer indolization reaction as variations in the *N*,*O*-acetal coupling fragment (at C7, C8, or C3) are considered.

The first point of variation we explored was the substituent at C7 of the *N*,*O*-acetal component, as a C7-aminoethyl substituent would be necessary for the synthesis of the communesin family of natural products and also for perophoramidine (Scheme 3.1). Acetal **3.16**, a readily available known compound,[16] was treated with TMSOTf in the presence of Hünig's base to afford enol ether **3.17** as an inconsequential mixture of E/Z isomers. Subsequently, enol ether **3.17** underwent a hetero-Diels–Alder reaction with sulfonamide **3.18**[17] to deliver the desired *N*,*O*-acetal substrate **3.19** in 70% yield. In the interrupted Fischer indolization reaction, treatment of **3.19** with *N*-methylphenylhydrazine (**3.13**) furnished tetracycles **3.20** and **3.21** in a combined 67% yield (approximately 3:1 ratio of **3.20** to **3.21**).[18] Thus, we were able to construct the key indoline scaffold, which bears a nitrogen functional handle useful for the synthesis of the communesin alkaloids.

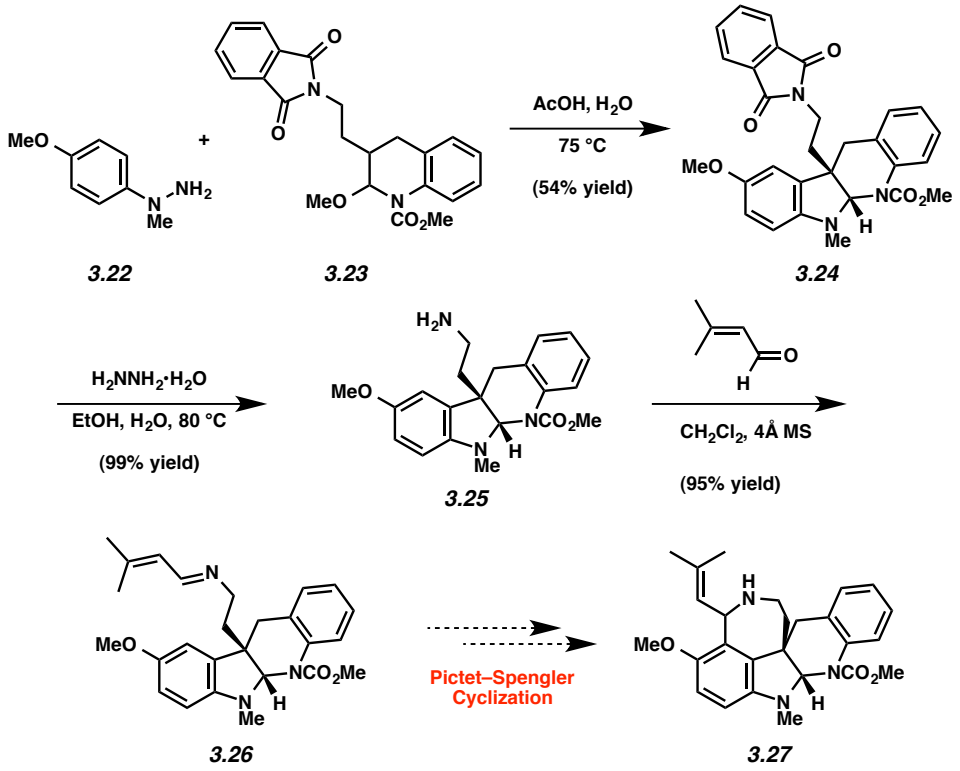
Scheme 3.1. Interrupted Fischer indolization of **3.19**.



3.5 Synthesis of Pictet–Spengler Precursor

Encouraged by these results, we also explored the interrupted Fischer indolization of two reaction partners that could plausibly allow for assembly of the 7-membered G ring of the communesins. Thus, hydrazine **3.22**[19] underwent reaction with *N*,*O*-acetal **3.23**[20] to provide tetracyclic indoline **3.24** in 54% yield (Scheme 3.2). Subsequent deprotection of the phthalyl group was achieved upon treatment of **3.24** with hydrazine monohydrate to give primary amine **3.25**. With the aim of forging the G ring via a Pictet–Spengler cyclization,[21] amine **3.25** was condensed with 3-methylbut-2-enal to furnish imine **3.26**. Unfortunately our efforts to arrive at pentacycle **3.27** through a Pictet–Spengler cyclization have been unsuccessful to date.[22]

Scheme 3.2. Interrupted Fischer indolization of **3.22**+**3.23** and further elaboration to **3.26**.



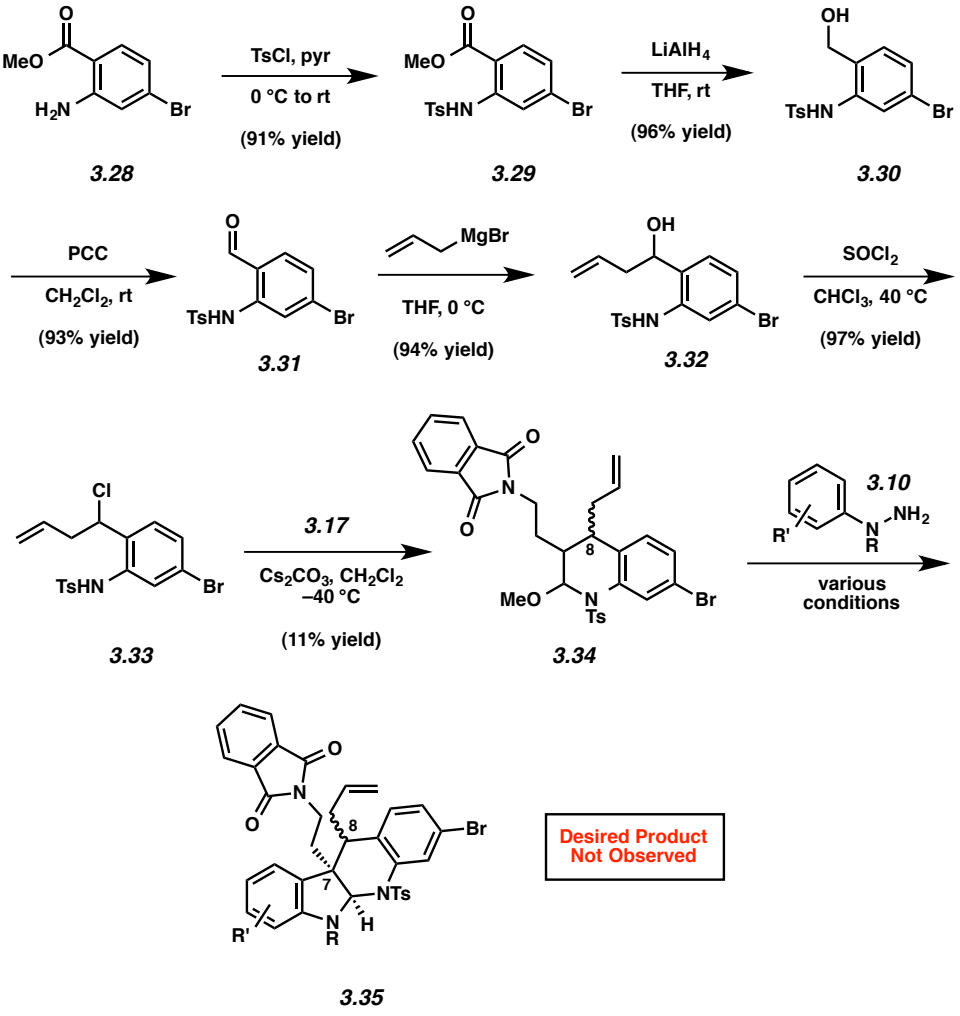
3.22 ***3.23*** ***3.24*** ***3.25*** ***3.26*** ***3.27***

3.6 Attempted Interrupted Fischer Indolization on C8 Substituted *N,O*-acetal

Other aspects of the interrupted Fischer indolization we aimed to test involved substitution at C8 for the eventual installation of the C7/C8 vicinal quaternary stereocenters, in addition to the aryl bromide substituent present in perophoramidine. Our route to a suitable substrate is shown in Scheme 3.3. Commercially available aniline **3.28** was converted to sulfonamide **3.29**, which in turn, underwent reduction upon treatment with lithium aluminum hydride to furnish alcohol **3.30**. Oxidation of alcohol **3.30** with PCC yielded aldehyde **3.31**, which was elaborated to alcohol **3.32** by the addition of allyl magnesium bromide. Treatment of alcohol **3.32** with thionyl chloride provided benzylic chloride **3.33**, which was the substrate for the hetero-Diels–Alder reaction. In the event, the desired C8 substituted acetal **3.34** was

obtained, albeit in modest yield.[23] Attempts to affect the interrupted Fischer indolization on this substrate to arrive at **3.35** proved fruitless, despite exhaustive efforts using a variety of acid-mediated conditions, in addition to several arylhydrazine coupling partners.[24] Although this route proved ineffective for our planned total synthesis, these efforts reveal a steric limitation of the interrupted Fischer indolization methodology.

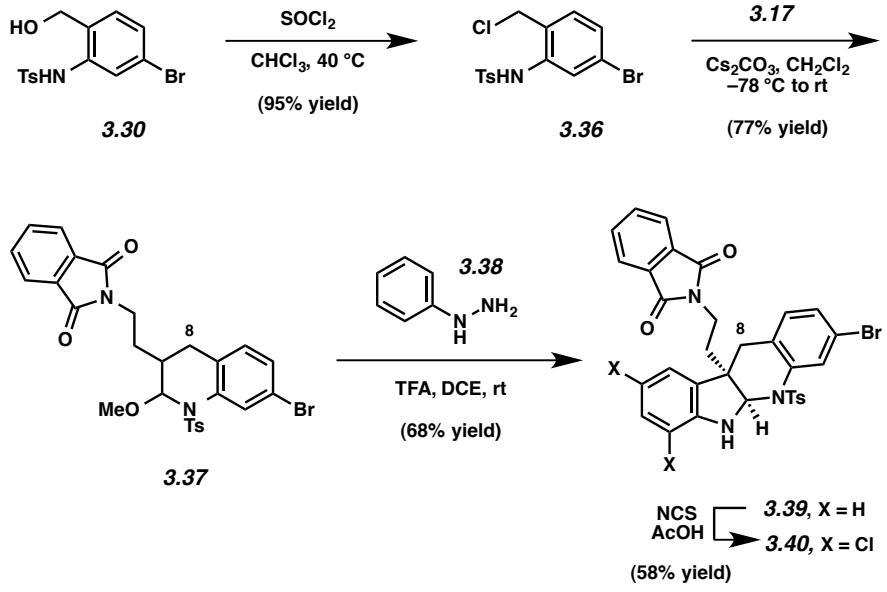
Scheme 3.3. Attempted interrupted Fischer indolization of C8-allylated compound **3.34**.



3.7 Synthesis of the Trihalotetracycle of Perophoramidine

We also prepared a derivative of *N,O*-acetal **3.34** lacking the C8 allyl substituent and tested its viability in the interrupted Fischer indolization reaction (Scheme 3.4). Alcohol **3.30** was treated with thionyl chloride to provide benzylic chloride **3.36** in 95% yield. Subsequent reaction of alkyl chloride **3.36** with enol ether **3.17** under basic conditions provided *N,O*-acetal **3.37** in significantly higher yield compared to the reaction of the C8-allylated material (see Scheme 3.3, **3.33**+**3.17**→**3.34**).[23] Fortunately, treatment of *N,O*-acetal **3.37** with phenylhydrazine (**3.38**) in the presence of TFA in dichloroethane generated tetracycle **3.39** in a 68% yield.[25] The success of this transformation supports the notion that the interrupted Fischer indolization reaction of **3.34** fails (see Scheme 3.3) due to steric considerations. Finally, tetracycle **3.39** was exposed to NCS to furnish trihalotetracycle **3.40**.[9] Intermediate **3.40**, which is available in only 6 steps from commercially available materials, is expected to serve as a precursor to perophoramidine (**3.9**).[26]

Scheme 3.4. Synthesis of trihalotetracycle **3.40**.



3.8 Conclusion

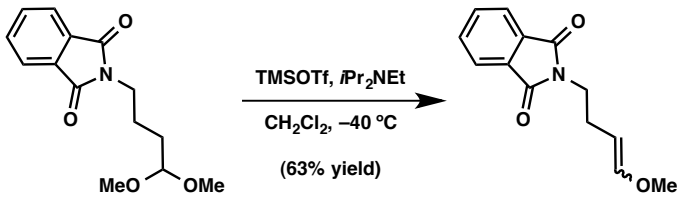
In summary, we have established a promising synthetic strategy toward the communesin alkaloids and perophoramidine. The approach involves the use of the interrupted Fischer indolization methodology, which in turn enables the rapid assembly of tetracyclic scaffolds that are reminiscent of the cores of the desired natural products. Although the methodology does not tolerate a C8-allyl substituent en route to the challenging vicinal quaternary stereocenters, variation at C7 and on the C ring is permitted. Current efforts are dedicated to further elucidating the subtleties of the interrupted Fischer indolization methodology in the context of natural product total synthesis.

3.9 Experimental Section

3.9.1 Materials and Methods

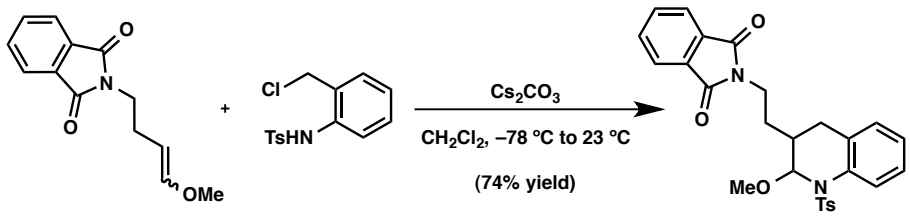
Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of nitrogen using anhydrous solvents (either freshly distilled or passed through activated alumina columns). All commercially available reagents were used as received unless otherwise specified. 2-amino benzylalcohol (TCI), and $CsCO_3$ (Alfa-Aeser), were obtained from VWR. Acetic Acid was obtained from Fisher Scientific. Hydrazine monohydrate, tosyl chloride, lithium aluminum hydride, and were obtained from Sigma–Aldrich. Reaction temperatures were controlled using an IKAmag temperature modulator, and unless stated otherwise, reactions were performed at room temperature (rt, approximately 23 °C). Thin-layer chromatography (TLC) was conducted with EMD gel 60 F254 pre-coated plates (0.25 mm) and visualized using a combination of UV, anisaldehyde, iodine, and potassium permanganate staining. EMD silica gel 60 (particle size 0.040–0.063 mm) was used for flash column chromatography. ^{1}H NMR spectra were recorded on Bruker spectrometers (at 300 MHz, 400 MHz, or 500 MHz) and are reported relative to deuterated solvent signals. Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz) and integration. ^{13}C NMR spectra are reported in terms of chemical shift. For mixtures of diastereomers and rotamers, the major diastereomer is reported with the minor diastereomer in parentheses for both ^{1}H NMR and ^{13}C NMR spectra. IR spectra were recorded on a Perkin-Elmer 100 spectrometer and are reported in terms of frequency absorption (cm^{-1}). High resolution mass spectra were obtained from the UC Irvine and UCLA Mass Spectrometry Facilities.

3.9.2 Experimental Procedures



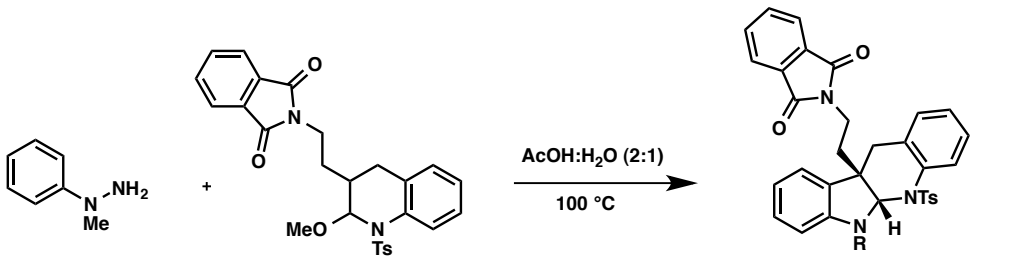
3.16 *3.17*

Vinyl Ether 3.17: To a solution of acetal **3.16**[27] (5.07 g, 19.25 mmol, 1 equiv) and *i*-Pr_2NEt (4.7 mL, 26.95 mmol, 1.4 equiv) in CH_2Cl_2 (130 mL), TMSOTf (4.2 mL, 23.10 mmol, 1.2 equiv) was added dropwise over 3 min at –40 °C. After stirring for 50 min, the reaction was quenched with sat. aq. $NaHCO_3$ (130 mL). The layers were separated and the aqueous layer was extracted with EtOAc (3 x 100 mL). The combined organic layers were washed with brine (200 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resultant residue was purified by flash chromatography (6:1 hexanes:EtOAc) to afford vinyl ether **3.17** (2.79 g, 63% yield) as a white solid and an 82:18 mixture of trans:cis isomers. R_f 0.75 (1:1 hexanes:EtOAc); ^1H NMR (500 MHz, $CDCl_3$): Trans isomer: δ 7.76–7.79 (m, 2H), 7.64–7.70 (m, 2H), 5.84 (d, *J* = 6.1, 1H), 4.29 (ddd, *J* = 14.0, 7.2, 6.8, 1H), 3.40–3.69 (m, 2H), 3.35 (s, 3H), 2.40–2.52 (m, 1H); Cis isomer: 7.76–7.79 (m, 2H), 7.64–7.70 (m, 2H), 6.25 (d, *J* = 12.5, 1H), 4.63 (ddd, 7.5, 5.2, 2.2), 3.41 (s, 3H), 2.22–2.38 (m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): Trans isomer: δ 168.4, 148.4, 133.7, 132.0, 123.0, 101.9, 59.4, 37.7, 23.2; Cis isomer: 168.4, 148.9, 133.7, 132.0, 123.1, 98.4, 55.7, 38.6, 27.1; IR (neat): 2933, 1706, 1665, 1436, 1385, 1328, 1109 cm^{-1}; m.p. 51–52 °C; HRMS-ESI (*m/z*) $[M + H]^+$ calcd for $C_{13}H_{13}NO_3$, 232.0974; found, 232.0976.



3.17 *3.18* *3.19*

***N,O* Acetal 3.19:** To a solution of vinyl ether **3.17** (517 mg, 2.23 mmol, 1 equiv) in CH_2Cl_2 (12 mL) at –78 ºC was added Cs_2CO_3 (2.55 g, 7.82 mmol, 3.5 equiv). Chloroaniline **3.18** in CH_2Cl_2 (16 mL) was added dropwise over 3.5 h via a syringe pump. After the addition of chloroaniline **3.18**, the reaction mixture was allowed to warm naturally to room temperature over 12 h. The reaction was then quenched with H_2O (30 mL) and the layers separated. The aqueous layer was extracted with EtOAc (3 x 40 mL). The combined organic layers were washed with brine (50 mL), dried over $MgSO_4$, and concentrated under reduced pressure. Purification by column chromatography (6:1 → 3:1 hexanes:EtOAc) afforded *N,O*-acetal **3.19** (806 mg, 74% yield) as a white solid. R_f 0.35 (2:1 hexanes/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.88 (7.88) (d, *J* = 8.0, 2H), 7.87 (7.87) (dd, *J* = 5.5, 3.0, 2H), 7.75 (7.75) (dd, *J* = 5.5, 3.0, 2H), 7.44 (7.44) (d, *J* = 8.0, 2H), 7.15 (7.15) (t, *J* = 8.0, 1H), 7.02–7.07 (7.02–7.07) (m, 4H), 5.40 (4.75) (d, *J* = 2.5, 1H), 3.67 (3.67) (t, *J* = 8.0, 2H), 3.41 (3.49) (s, 3H), 2.69 (2.69) (dd, *J* = 17.0, 7.4, 1H), 2.60 (2.60) (dd, *J* = 16.9 12.4, 1H), 2.26 (2.34) (s, 3H), 1.92 (1.92) (m, 1H), 1.71 (1.71) (m, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 168.2 (168.2), 143.7 (143.8), 136.2 (136.9), 134.1 (134.1), 133.3 (134.0), 132.0 (132.0), 129.5 (129.5), 129.3 (127.5), 128.2 (127.4), 127.0 (126.9), 126.5 (126.8), 124.8 (127.0), 124.1 (124.1), 123.3 (123.2), 87.5 (92.1), 55.8 (55.3), 35.5 (35.6), 33.3 (32.9), 31.0 (29.9), 28.6 (8.6), 21.6 (21.5); IR (film): 2900, 1769, 1707, 1491, 1394, 1348, 1167, 1090, 1011 cm^{-1}; m.p. 143–146 °C; HRMS-ESI (*m/z*) $[M + Na]^+$ calc $C_{27}H_{26}N_2O_5S$, 513.1460; found, 513.1459.

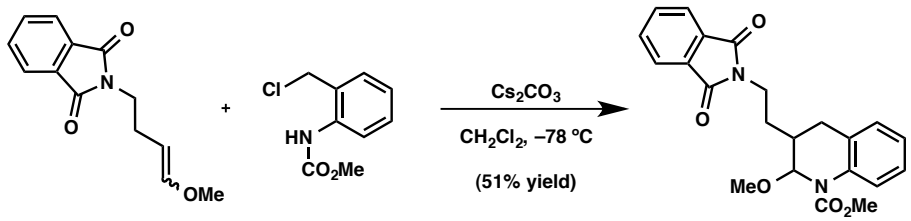


3.13 *3.19*

3.20, **R = Me (51% yield)**
3.21**, R = H (16% yield)**

Indolines 3.20 and 3.21: To a solution of *N*,*O*-acetal **3.19** (75.8 mg 0.155 mmol) in 580 μL of AcOH, was added *N*-methylphenylhydrazine (**3.13**) (22 μL, 0.187 mmol) followed by H_2O (190 μL). The reaction was heated to 100 °C for 2 h. The reaction mixture was then cooled to room temperature, diluted with EtOAc (15 mL), and quenched with sat. aq. $NaHCO_3$ (15 mL). The layers were separated and the aqueous layer was extracted with EtOAc (3 x 15 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$, and concentrated under reduced pressure to give crude indolines **3.20** and **3.21**. Purification by flash chromatography (1:1:16 → 1:1:8 Et_2O:CH_2Cl_2:hexanes) provided indoline **3.20** (44.8 mg, 51% yield) as a yellow foam and indoline **3.21** (13.7 mg, 16% yield) as a white foam. Indoline **3.20**: R_f = 0.50 (1:1:4 Et_2O:CH_2Cl_2:hexanes); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.80 (dd, *J* = 5.5, 2.5, 2H), 7.70 (dd, *J* = 5.5, 2.5, 2H), 7.51–7.69 (m, 3H), 7.20 (d, *J* = 8.0, 2H), 7.10 (t, *J* = 7.5, 1H), 6.96 (t, *J* = 7.5, 1H), 6.74–6.80 (m, 3H), 6.39 (t, *J* = 7.5, 1H), 6.06 (d, *J* = 7.5, 1H), 5.88 (s, 1H), 3.50–3.57 (m, 1H), 3.46–3.48 (m, 1H), 2.94 (s, 3H), 2.50 (d, *J* = 14.0, 1H), 2.36 (s, 3H), 2.04–2.36 (m, 1H), 1.87–1.89 (m, 1H), 1.61 (d, *J* = 14.0, 1H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.0, 150.5, 143.8, 137.8, 135.3, 134.1, 133.9, 132.1, 129.8, 128.9, 128.3, 128.2, 128.1, 127.3, 127.2, 126.8, 123.1, 122.0, 117.1, 104.6, 83.9, 53.6, 39.2, 37.4, 33.7, 29.6, 21.6; IR (neat): 2810, 1771, 1707, 1609, 1491, 1397, 1349, 1164; HRMS-ESI (*m/z*) $[M + H]^+$ calcd for $C_{33}H_{30}N_3O_4S$, 564.1957; found, 564.1954. Indoline **3.21**: R_f = 0.33 (1:1:4 Et_2O:CH_2Cl_2:hexanes); ^{1}H NMR (500

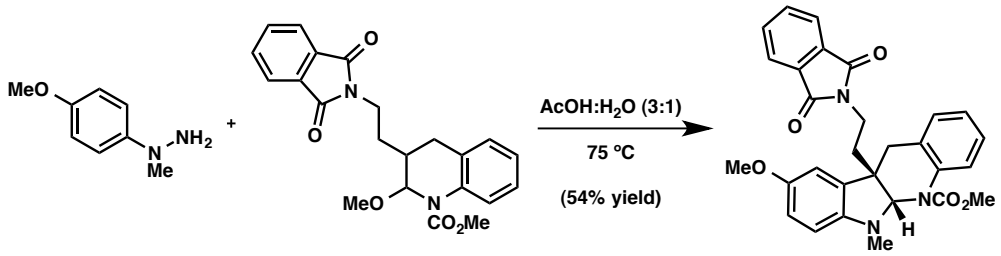
MHz, $CDCl_3$): δ 7.81 (dd, *J* = 5.5, 2.5, 2H), 7.70 (dd, *J* = 5.5, 2.5, 2H), 7.57 (d, *J* = 8.5, 2H), 7.53 (d, *J* = 8.0, 1H), 7.22 (d, *J* = 8.0, 2H), 7.12 (t, *J* = 7.5, 1H), 6.98 (d, *J* = 7.5, 1H), 6.76–6.81 (m, 3H), 6.49 (t, *J* = 7.5, 1H), 6.33 (d, *J* = 1H), 5.99 (s, 1H), 4.35 (s, 1H), 3.64 (ddd, *J* = 10.5, 4.5, 1.0, 1H), 3.44 (ddd, *J* = 10.5, 4.5, 1.0, 1H), 2.38 (s, 3H), 2.05 (ddd, *J* = 11.0, 6.0, 3.5, 1H), 1.88 (ddd, *J* = 11.0, 6.0, 3.5, 1H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.0, 148.9, 143.8, 137.6, 135.2, 133.9, 133.3, 132.1, 129.9, 128.6, 128.2, 128.1, 128.0, 127.1, 126.9, 126.8, 123.2, 122.5, 118.8, 108.0, 78.9, 54.5, 39.6, 36.9, 33.6, 21.6; IR (neat): 3362, 1766, 1702, 1610, 1487, 1396, 1346; HRMS-ESI (*m/z*) $[M + Na]^+$ calcd for $C_{32}H_{27}N_3O_4SNa$, 574.1643; found, 574.1617.



3.17 ***3.41*** ***3.23***

***N,O*-Acetal 3.23:** Vinyl ether **3.17** (1.24 g, 5.36 mmol 1 equiv) and Cs_2CO_3 was dissolved in CH_2Cl_2 (15mL). The solution was cooled to –78 ºC. Chloroaniline **3.41**[28] (1.86 g, 4.96 mmol, 1.7 equiv) in CH_2Cl_2 (32 mL) was added dropwise to the vigorously stirred reaction mixture over 4.5 h via a syringe pump. The reaction mixture was then allowed to warm to room temperature over 15 h. The reaction was then quenched with H_2O (50 mL) and the layers separated. The aqueous layer was extracted with EtOAc (3 x 100 mL). The combined organic layers were washed with brine (100 mL), dried over $MgSO_4$, and concentrated under reduced pressure. Purification by column chromatography (5:1 hexanes:EtOAc) afforded *N,O*-acetal **3.23** (1.08 g, 51% yield) as a white foam as a 63:37 mixture of diasteromers. R_f 0.60 (1:1 hexanes/EtOAc); ^1H NMR (500 MHz, $CDCl_3$): δ 7.82 (7.82) (dd, *J* = 5.5, 3.0, 2H), 7.70 (7.70) (dd, *J* = 5.5, 3.0, 2H), 7.61 (7.28)

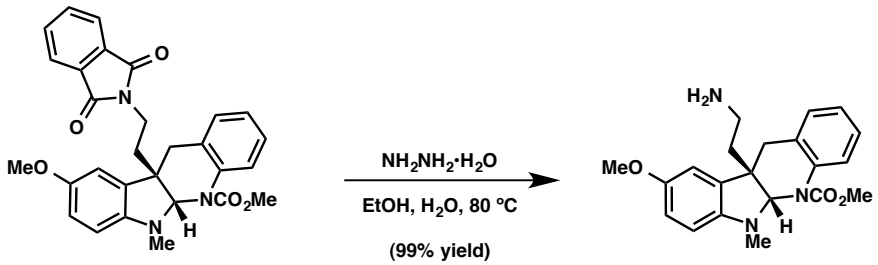
(d, *J* = 8.0, 1H), 7.08–7.15 (7.08–7.15) (m, 2H), 6.98–7.04 (6.98–7.04) (m, 1H), 5.61 (4.67) (d, *J* = 2.0, 1H), 3.80 (3.75) (s, 3H), 3.77–3.82 (3.77–3.82) (m, 1H) 3.29 (3.33) (s, 3H), 3.03 (2.52) (dd, *J* = 17.0, 6.5, 1H), 2.71 (2.71) (dd, *J* = 17.0, 6.5), 1.96–2.03 (1.96–2.03) (m, 2H), 1.77–1.88 (1.66–1.69), (m, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 168.4 (168.3), 154.8 (154.5), 134.3 (134.9), 134.0 (134.0), 132.1 (132.0), 128.9 (129.1), 128.8 (128.7), 127.8 (127.8), 125.9 (126.2), 123.9 (124.7), 124.0 (124.6), 123.3 (123.2), 85.1 (64.3), 55.5 (55.3), 53.1 (53.2), 35.4 (52.3), 31.6 (31.0), 28.9 (28.8); IR (film): 3465, 2951, 1771, 1707, 1593, 1530 cm^{-1}; HRMS-ESI (*m/z*) [M – OMe]$^+$ calcd for $C_{21}H_{19}N_2O_4$, 363.1345; found, 363.1345.



3.22 ***3.23*** ***3.24***

Indoline 3.24: *N,O*-acetal **3.23** (133 mg, 0.337 mmol, 1 equiv) was dissolved in AcOH (1.3 mL). Aryl hydrazine **3.22**[29] (102 mg, 0.674 mmol, 2 equiv) and H_2O (400 μL) was added and the mixture was stirred for 5 min at rt and then for 3 h at 75 °C. After cooling to room temperature, the reaction mixture was diluted with EtOAc (30 mL) and quenched with saturated aqueous $NaHCO_3$ (50 mL). The layers were separated and the resulting aqueous layer was extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine (50 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (3:1 → 2:1 hexanes/EtOAc) to afford indoline **3.24** (90.5 mg, 54% yield) as an yellow foam. R_f 0.32 (2:1 Et_2O:hexanes); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.76 (dd, *J* = 6.5, 3.5, 2H), 7.66 (dd, *J* = 5.5, 3.5, 2H), 7.22 (br s, 1H), 7.07 (td, *J* = 8.5, 2.0, 1H), 6.89–6.93 (m, 2H),

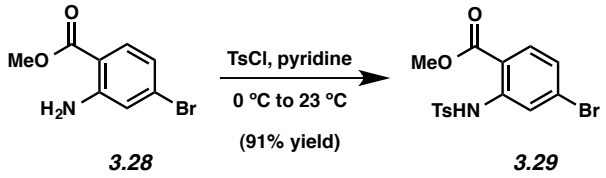
6.52 (d, *J* = 2.5, 1H), 6.26 (dd, *J* = 7.5, 2.5, 1H), 6.15 (br s, 1H), 5.90 (d, *J* = 8.0, 1H), 3.82 (br s, 3H), 3.62–3.78 (m, 2H), 3.58 (s, 3H), 2.84 (d, *J* = 3.0, 2H), 2.79 (s, 3H), 2.42 (ddd, *J* = 14.0, 3.5, 2.5, 1H), 2.18 (ddd, *J* = 14.0, 7.0, 7.0, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 167.8, 151.9, 145.2, 137.1, 133.6, 133.4, 132.4, 132.0, 131.4, 127.9, 126.6, 125.5, 125.3, 122.9, 112.2, 109.8, 104.5, 81.3, 55.7, 53.5, 52.9, 38.3, 38.2, 33.9, 30.6; IR (film): 2950, 1771, 1710, 1496, 1439, 1399, 1328, 1251, 1126, 1029 cm^{-1}; HRMS-ESI (*m/z*) [M + Na]$^{+}$ calc $C_{29}H_{27}N_3O_5S$, 520.1848; found, 520.1849.



3.24 *3.25*

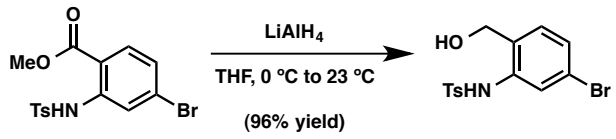
Indoline 3.25: Indoline **3.24** (29.6 mg, 0.059 mmol, 1 equiv) was dissolved in EtOH (400 μL) and H_2O (5 mL). Hydrazine monohydrate (28 μL) was added and reaction was heated to 80 ºC for 1 h. The reaction mixture was then cooled to room temperature, and the resulting solid was filtered through a plug of celite ($CHCl_3$ eluent, 15 mL) to provide crude indoline **3.25**. Purified by column chromatography (1:3 MeOH:CH_2Cl_2) afforded indoline **3.25** (23 mg, 99% yield) as a white foam. R_f 0.00 (1:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, CD_3COD); δ 7.21 (br s, 1H), 7.04 (t, *J* = 8.0, 1H), 6.91–6.97 (m, 2H), 6.61 (d, *J* = 2.0, 1H), 6.46 (dd, *J* = 8.5, 2.5, 1H), 6.05 (d, *J* = 8.5, 1H), 5.98 (br s, 1H), 3.79 (br s, 3H), 3.62 (s, 3H), 3.16 (ddd, *J* = 4.5, 3.5, 2.5, 1H), 2.95 (d, *J* = 14.5, 1H), 2.88 (d, *J* = 15.0, 1H), 2.77 (s, 3H), 2.26–2.32 (m, 1H); ^{13}C NMR (125 MHz, CD_3COD): δ 180.0, 152.5, 145.4, 136.9, 132.8, 131.2, 127.9, 126.6, 125.7, 125.5, 112.5, 110.3,

105.1, 81.8, 55.0, 53.1, 52.9, 37.8, 36.6, 36.2, 29.8; IR (film): 2820, 1710, 1553, 1420, 1273 cm^{-1}; HRMS-ESI (*m/z*) $[M + H]^+$ calc $C_{21}H_{25}N_3O_3$, 368.1974; found, 368.1972.



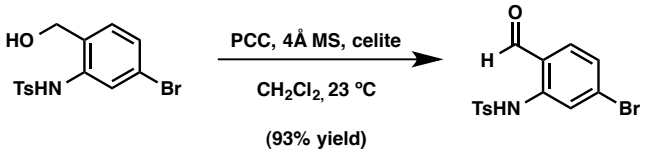
3.28 *3.29*

Ts-Aniline 3.29: To a solution of tosyl chloride (9.05 g, 43.50 mmol, 1.1 equiv) in pyridine (5 mL) at 0 °C, was added a solution of aniline **3.28** (9.93 g, 43.10 mmol, 1 equiv) in pyridine (23 mL) over 7 min. Once the addition was complete, the reaction was warmed to room temperature and stirred for another 2 h. The reaction mixture was quenched with 1M HCl (30 mL) and the resulting solids were removed by filtration. The resulting filter cake was washed with 1M HCl (30 mL) and H_2O (10 mL). The filtrate was evaporated under reduced pressure to provide Ts-aniline **3.29** (15.32 g, 91% yield) as a tan solid, which was used without further purification. R_f 0.64 (2:1 hexanes/EtOAc); ^1H NMR (500 MHz, $CDCl_3$): δ 10.71 (br s, 1H), 7.89 (d, *J* = 2.0, 1H), 7.76 (dd, *J* = 8.3, 2.0, 2H), 7.26 (d, *J* = 8.3, 3H), 7.14 (dd, *J* = 8.5, 2.0, 1H), 3.88 (s, 3H), 2.38 (s, 3H); ^{13}C NMR (125 MHz, CDCl3): δ 168.0, 144.4, 141.7, 136.2, 132.4, 129.9, 129.6, 127.5, 126.1, 121.7, 114.3, 52.8, 21.7; IR (neat): 3109, 2954, 1685, 1590, 1485, 1087 cm^{-1}; m.p. 149–150 °C; HRMS-ESI (*m/z*) $[M + H]^+$ calcd for $C_{14}H_{12}BrNO_4S$, 371.9729; found, 371.9727.



3.29 *3.30*

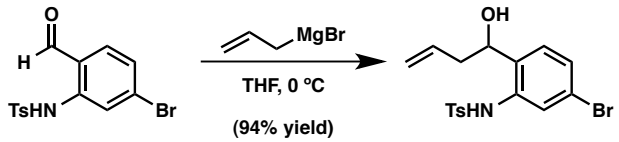
Benzyl alcohol 3.30: To a solution of Ts-aniline **3.29** (15.32 g, 39.36 mmol, 1.0 equiv) in THF (240 mL) at 0 °C was added $LiAlH_4$ (2.94 g, 78.70 mmol, 2 equiv) portionwise over 5 min. After 20 min, the reaction was carefully quenched with 1M HCl (10 mL) followed by addition of a saturated aqueous solution of Na–K tartrate (30 mL). The reaction mixture stirred vigorously for 1 h at 23 ºC. The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 40 mL). The combined organic layers were washed with brine (75 mL), and dried over $MgSO_4$. Evaporation of the solvent under reduced pressure afforded benzyl alcohol **3.30** (13.50 g, 96% yield) as a tan solid, which was used in the subsequent reaction without further purification. R_f 0.43 (4:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.62 (d, *J* = 8.5, 2H), 7.55 (d, *J* = 3.2, 2H), 7.22 (d, *J* = 8.0, 2H), 7.15 (dd, *J* = 8.5, 2.0, 1H), 6.91 (d, *J* = 8.0, 1H), 4.33 (s, 2H), 2.38 (s, 3H); ^{13}C NMR (125 MHz, CDCl3): δ 144.3, 138.0, 136.7, 130.3, 129.9, 129.8, 128.2, 127.2, 125.9, 122.8, 63.7, 21.7; IR (neat): 3600, 1478, 1326, 1153, 1088, 1042 cm^{-1}; m.p. 135–137 °C; HRMS-ESI (*m/z*) $[M + H]^+$ calcd for $C_{14}H_{14}BrNO_3S$, 355.9957; found, 355.9956.



3.30 *3.31*

Aldehyde 3.31: Activated 4Å molecular sieves, celite, and benzyl alcohol **3.30** (13.5 g, 37.01 mmol, 1 equiv) were suspended in CH_2Cl_2 (190 mL). PCC (9.8 g, 45.48 mmol, 1 equiv) was then added in one portion. The reaction mixture stirred vigorously over 2 h, and then poured over an

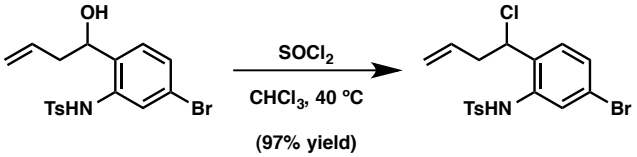
SiO_2/celite plug (4 cm x 5 cm, Et_2O eluent, 1L). The filtrate was concentrated under reduced pressure to afford aldehyde **3.31** as a tan solid (12.49 g, 93% yield), which was used in the subsequent step without further purification. R_f 0.45 (4:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.83 (s, 1H), 9.77 (s, 1H), 7.89 (d, *J* = 1.0, 2H), 7.78 (d, *J* = 8.5, 2H), 7.42 (d, *J* = 8.0, 1H), 7.27–7.30 (m, 3H), 2.37 (s, 3H); ^{13}C NMR (125 MHz, CDCl3): δ 194.2, 144.8, 141.0, 137.1, 136.1, 131.6, 130.1, 127.5, 126.3, 120.8, 120.5, 21.7; IR (neat): 3142, 1672, 1594, 1568, 1486, 1161, 1089 cm^{-1}; m.p. 145–150 °C; HRMS-ESI (*m/z*) [M + H]$^+$ calcd $C_{14}H_{12}BrNO_3S$, 355.9704; found, 355.9780.



3.31 ***3.32***

Alcohol 3.32: To a solution of aldehyde **3.31** (940 mg, 2.65 mmol, 1 equiv) in THF (13 mL) was cooled to 0 ºC and allyl magnesium bromide (1.0 M in THF, 5.84 mL, 5.84 mmol, 2.2 equiv) was added dropwise over 4 min. An additional portion of allyl magnesium bromide (800 µL, 0.80 mmol, 0.30 equiv) was added after 19 min. After stirring for 10 min, the reaction was quenched with a solution of sat. aq. NH_4Cl (20 mL) and transferred to a separatory funnel with EtOAc (20 mL). The layers separated and the aqueous layer was extracted with EtOAc (3 x 15 mL). The combined organic layers were washed with brine (20 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resultant residue was purified by flash chromatography (9:1 → 6:1 hexanes:EtOAc) to afford alcohol **3.32** (951 mg, 94% yield) as a colorless solid. R_f 0.48 (2:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.62 (br s, 1H), 7.71 (d, J = 4.7, 2H), 7.68 (d, *J* = 1.9, 2H), 7.27 (s, 1H), 7.14 (dd, *J* = 8.2, 1.9, 1H), 6.87 (d, *J* = 2.8, 1H), 5.62 (dddd, *J*

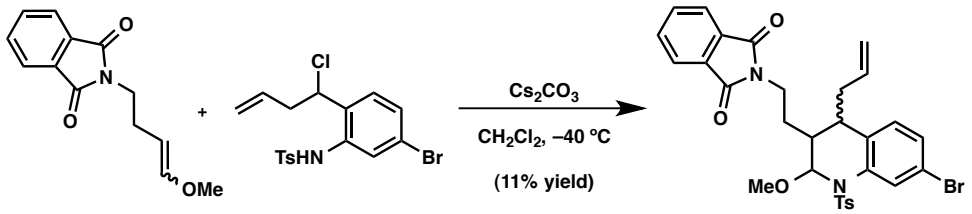
= 7.8, 6.8, 4.2, 1.0, 1H), 5.14 (dd, *J* = 10.2, 0.8, 1H), 5.04 (dd, *J* = 17.2, 1.6, 1H), 4.60 (dd, *J* = 8.5, 5.6, 1H), 2.39 (s, 3H), 2.21-2.31 (m, 2H); ^{13}C NMR (125 MHz, CDCl3): δ 144.3, 137.4, 136.7, 133.3, 130.3, 129.9, 129.3, 127.3, 127.2, 124.5, 122.3, 119.9, 73.5, 41.4, 21.7; IR (neat): 3485, 3233, 1641, 1595, 1490, 1329, 1159, 1090 cm^{-1}; HRMS-ESI (*m/z*) [M + Na]$^+$ calcd for $C_{17}H_{18}NO_3SBr$, 418.0089; found, 418.0115.



3.32 *3.33*

Chloride 3.33:1 To a solution of alcohol **3.32** (90.8 mg, 0.2303 mmol, 1 equiv) in anhydrous $CHCl_3$ (1.2 mL) was added Thionyl chloride (28 μL, 0.381 mmol, 1.6 equiv) . The reaction mixture was then heated to 40 ºC. After 16 h, the reaction was cooled to room temperature and quenched with H_2O (15 mL). The resulting mixture was transferred to a separatory funnel and the layers were separated. The aqueous layer was extracted with $CHCl_3$ (3 x 15 mL). The combined organic layers were washed with brine (20 mL) and dried over $MgSO_4$. Evaporation of the solvent under reduced pressure afforded crude chloride **3.33** as a tan solid (60 mg, 97% yield), which was subsequently used in the next reaction without purification. R_f 0.75 (2:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.65 (d, *J* = 6.7, 2H), 7.46 (d, *J* = 2.0, 1H), 7.34 (dd, *J* = 8.4, 2.0, 1H), 7.28 (d, *J* = 8.0, 2H), 7.21 (d, *J* = 8.4, 1H), 6.69 (s, 1H), 5.55 (dddd, *J* = 10.3, 6.8, 6.8, 3.3, 1H), 5.07 (dd, *J* = 10.3, 1.5, 2H), 4.72 (dd, *J* = 8.4, 6.3, 1H), 2.73-2.79 (m, 1H), 2.60 (m, 1H), 2.41 (s, 3H), 1.33 (t, *J* = 7.1, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 144.6, 136.2, 135.2, 133.8, 133.2, 130.1, 130.0, 129.0, 128.8, 127.3, 122.9, 119.1, 57.2, 41.4, 21.7; IR

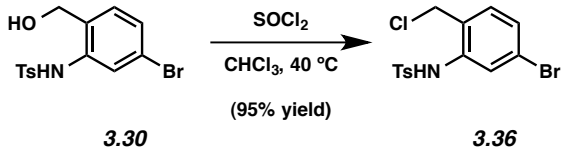
(neat): 3256, 1643, 1593, 1395, 1330, 1161, 1091 cm^{-1}; HRMS-ESI (*m/z*) $[M - Cl]^+$ calcd for $C_{17}H_{15}BrNO_2S$, 377.9987; found, 376.0007.



3.17 *3.33* *3.34*

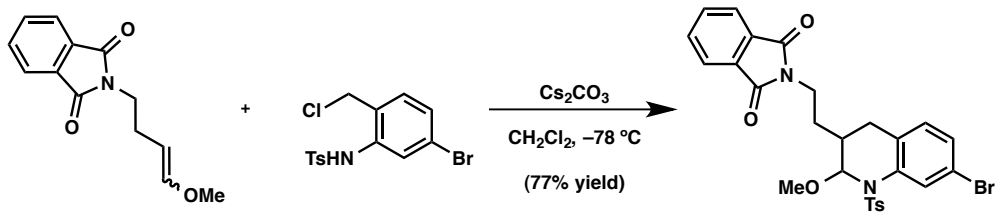
***N,O* acetal 3.34:** Vinyl ether **3.17** (218 mg, 0.943 mmol, 1.7 equiv) and Cs_2CO_3 (217 mg, 0.665 mmol, 1.2 equiv) was dissolved in CH_2Cl_2 (1.1 mL). The solution was cooled to –40 ºC to which the allyl chloride **3.33** (230 mg, 0.554 mmol, 1 equiv) in CH_2Cl_2 (2.6 mL) was added dropwise over 6.5 h via a syringe pump. The reaction mixture was then quenched with H_2O (15 mL) and transferred to a separatory funnel. The layers were separated and the aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resulting crude material was purified by flash chromatography (98:1:1 benzene:CH_2Cl_2:Et_2O) to afford *N,O* acetal **3.34** (36.2 mg, 11% yield) as a white foam. R_f 0.5 (1:1:1 Et_2O:CH_2Cl_2:benzene); ^1H NMR (500 MHz, $CDCl_3$): δ 7.95 (d, *J* = 1.7, 1H), 7.83 (dd, *J* = 5.1, 3.0, 2H), 7.72 (dd, *J* = 5.1, 3.0, 2H), 7.40 (d, *J* = 8.1, 2H), 7.20 (dd, *J* = 8.1, 1.7, 2H), 7.11 (d, *J* = 8.1, 1H), 6.88 (d, *J* = 8.1, 1H) 5.54 (dddd, *J* = 10.2, 6.9, 6.9, 3.5, 1H), 5.50 (d, *J* = 4.9, 1H), 4.85 (dd, *J* = 10.3, 1.0, 1H), 4.81 (dd, *J* = 16.4, 1.0, 1H) 3.63 (ddd, *J* = 14.3, 7.1, 2.7), 3.51 (s, 3H), 2.29–2.37 (m, 4H), 2.04–2.10 (m, 1H), 2.09 (dd, *J* = 11.8, 5.4, 1H), 2.04–2.10 (m, 1H), 1.51–1.62 (m, 2H); ^{13}C NMR (125 MHz, CDCl3): δ 168.5, 144.2, 137.2, 135.8, 135.3, 134.2, 133.6, 132.2, 129.7, 129.0, 128.4, 127.3, 123.3, 120.2, 116.8, 89.1,

56.5, 39.5, 37.5, 37.5, 34.6, 25.6, 21.6; IR (film): 3055, 1712, 1592, 1397, 1353, 1264, 1168, 1090 cm^{-1}; HRMS-ESI (*m/z*) $[M + H]^+$ calcd for $C_{29}H_{26}BrN_2O_4S$, 579.0780; found, 579.0778.



3.30 *3.36*

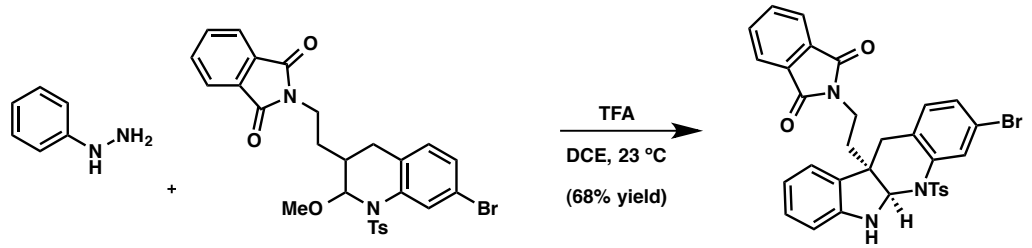
Chloroaniline 3.36: To a solution of benzyl alcohol **3.30** (5.31 g, 14.91 mmol, 1 equiv) in chloroform (75 mL) was added thionyl chloride (1.73 mL, 23.85 mmol, 1.6 equiv) was added. The reaction mixture was heated to 40 ºC for 12 h and then cooled to room temperature. The reaction mixture was diluted with EtOAc (30 mL) and quenched with $NaHCO_3$ (30 mL). The layers were separated and the resulting aqueous layer was extracted with EtOAc (3 x 25 mL). The combined organic layers were washed with brine (40 mL) and dried over $MgSO_4$, and concentrated under reduced pressure to afford benzyl choride **3.36** as a tan solid (5.30 g, 95% yield), which was used without further purification. R_f 0.66 (2:1 hexanes:EtOAc); ^1H NMR (500 MHz, $CDCl_3$): δ 7.67 (d, *J* = 8.3, 2H), 7.56 (d, *J* = 1.9, 1H), 7.27 (d, *J* = 8.3, 2H), 7.09 (d, *J* = 8.3, 1H), 6.80 (br s, 1H), 4.24 (s, 2H), 2.41 (s, 3H); ^{13}C NMR (125 MHz, CDCl3): δ 144.7, 136.6, 136.2, 131.7, 130.1, 129.4, 129.2, 127.7, 127.3, 123.7, 42.6, 21.7; IR (neat): 3291, 1595, 1574, 1490, 1393, 1339, 1161, 1089; m.p. 104–107 °C; HRMS-ESI (*m/z*) $[M - H]^{-1}$ calcd $C_{14}H_{12}BrClNO_2$, 373.9438; found 373.9438.



3.17 *3.36* *3.37*

***N,O*-Acetal 3.37:** To a solution of vinyl ether **3.17** (675 mg, 2.92 mmol, 1 equiv) and Cs_2CO_3 (3.33 g, 10.2 mmol, 3.5 equiv) in CH_2Cl_2 (4 mL), was added chloroaniline **3.36** (1.86 g, 4.96 mmol, 1.7 equiv) in CH_2Cl_2 (32 mL) was added dropwise to the vigorously stirred reaction mixture over 4.5 h via a syringe pump at –78 °C. The reaction was allowed to warm to room temperature over 14 h and was quenched with H_2O (50 mL). The layers were separated and the aqueous layer was extracted with EtOAc (3 x 70 mL). The combined organic layers were washed with brine (100 mL), dried over $MgSO_4$, and concentrated under reduced pressure. Purification by column chromatography (10:1 → 10:2.5 hexanes/EtOAc) afforded *N,O*-acetal **3.37** (1.29 g, 77% yield, 8.7:1 mixture of diastereomers) as a white solid. R_f 0.35 (2:1 hexanes/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.08 (8.08) (d, *J* = 1.9, 1H), 7.86 (7.82) (dd, *J* = 9.2, 3.8, 2H), 7.76 (7.71) (dd, 5.6, 3.1, 2H), 7.47 (7.47) (d, *J* = 8.3, 2H), 7.12 (7.17) (dd, *J* = 8.3, 1.8, 1H), 7.06 (7.06) (d, *J* = 8.1, 2H), 6.88 (6.90) (d, *J* = 8.2, 2H), 5.40 (5.40) (d, *J* = 2.4, 1H), 3.67 (3.67) (apt t, *J* = 7.2, 2H), 3.40 (3.47) (s, 3H), 2.64 (2.64) (dd, *J* = 21.3, 6.7, 1H), 2.53 (2.53) (dd, *J* = 16.7, 11.8, 1H), 2.27 (2.35) (s, 3H), 1.91 (1.91), (m, 1H), 1.71 (1.71) (m, 1H), 1.35-1.44 (1.35-1.44) (m, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 168.4 (168.3), 154.8 (154.5), 134.3 (134.9), 134.0 (134.0), 132.1 (132.0), 128.9 (129.1), 128.8 (128.7), 127.8 (127.8), 125.9 (126.2), 124.0 (124.6), 123.9 (124.7), 123.3 (123.3), 85.1 (87.4), 64.3 (64.4), 55.5 (55.3), 53.1 (53.2), 52.3 (52.3), 35.6 (35.4), 35.0 (35.0), 31.6 (31.1), 28.9 (28.8); IR (film): 2935, 1771, 1708, 1592, 1396, 1349,

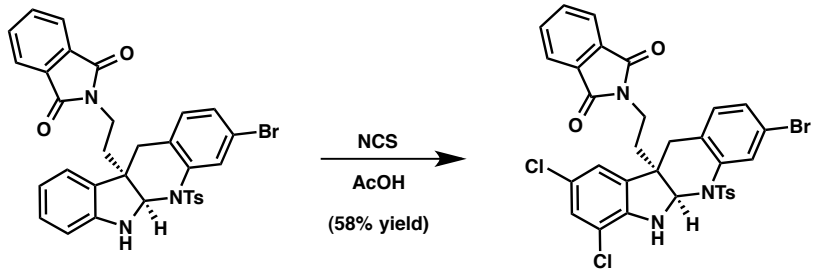
1243, 1165, 1089 cm^{-1}; m.p. 125–127 °C; HRMS-ESI (*m/z*) $[M + Na]^+$ calcd for $C_{27}H_{25}BrN_2O_5S$, 591.0565; found, 593.0548.



3.38 ***3.37*** ***3.39***

Indoline 3.39: To a solution of *N,O*-acetal **3.37** (1.56 g, 2.73 mmol, 1 equiv) in 1,2-dichloroethane (14 mL) was added phenylhydrazine (**3.38**) (1.35 mL, 13.70 mmol, 5 equiv). The mixture was stirred vigorously and trifluoroacetic acid (2.11 mL) was added dropwise over 2 min. After stirring over 22 h, the reaction was diluted with EtOAc (20 mL) and quenched with saturated aqueous $NaHCO_3$ (30 mL). The layers were separated and the resulting aqueous layer was extracted with EtOAc (3 x 40 mL). The combined organic layers were washed with brine (40 mL), dried over $MgSO_4$, and concentrated under reduced pressure to afford the crude product. Purification by flash chromatography (19:1:1 benzene:Et_2O:CH_2Cl_2) afforded indoline **3.39** (1.17 g, 68% yield) as a yellow foam. R_f 0.55 (5:1:1 benzene:Et_2O:CH_2Cl_2); ^1H NMR (500 MHz, $CDCl_3$): δ 7.81 (dd, *J* = 5.5, 3.0, 2H), 7.71 (s, 1H) 7.70 (dd, *J* = 5.5, 3.0, 2H), 7.60 (d, *J* = 8.5, 2H), 7.24 (d, *J* = 8.5, 2H), 7.10 (dd, *J* = 8.0, 2.0, 1H), 6.79–6.82 (m, 2H), 6.68 (d, *J* = 8.0, 1H), 6.36 (d, *J* = 7.5, 1H), 5.98 (s, 1H), 4.37 (s, 1H), 3.62 (ddd, *J* = 10, 4.5, 1.0, 1H), 3.43 (ddd, *J* = 10.5, 6.5, 3.5, 1H), 2.73 (d, *J* = 14.0, 1H), 2.39 (s, 3H), 2.04 (ddd, *J* = 10.5, 6.0, 4.5, 1H), 1.85 (ddd, *J* = 10.5, 4.5, 3.0, 1H), 1.70 (d, *J* = 14.5, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 167.9, 148.7, 144.2, 137.4, 136.6, 133.9, 132.2, 132.1, 130.6, 130.0, 129.9, 129.3, 128.3, 128.2, 126.8, 123.2, 122.4, 120.1, 118.9, 108.3, 78.8, 54.5, 39.5, 36.5, 33.5, 21.6; IR (film): 3362, 1702, 1609,

1487, 1468, 1396, 1346, 1161, 1062, 1026 cm^{-1}; HRMS-ESI (*m/z*) $[M + Na]^+$ calc $C_{32}H_{27}N_3O_4S$, 572.1620; found, 572.1619.



3.39 *3.40*

Indoline 3.40: To a solution of indoline **3.39** (105.8 mg, 0.168 mmol, 1 equiv) in AcOH (1.7 mL) was added *N*-chlorosuccinimide (67.4 mg, 0.505 mmol, 3 equiv). After stirring for 16 h at rt, the reaction was then diluted with EtOAc (20 mL) and quenched with saturated aqueous $Na_2S_2O_3$ (20 mL). The layers were separated and the resulting aqueous layer was extracted with EtOAc (3 x 30 mL). The combined organic layers were washed with brine (40 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (6:1 → 4:1 hexanes/EtOAc) to afford indoline **3.40** (68.4 mg, 58% yield) as an white foam. R_f 0.48 (2:1 hexanes:EtOAc); ^1H NMR (500 MHz, $CDCl_3$): δ 7.80 (dd, *J* = 6.5, 4.0, 2H), 7.70–7.73 (m, 3H), 7.60 (d, *J* = 8.5, 2H), 7.26 (d, *J* = 8.0, 2H), 7.17 (dd, *J* = 8.0, 2.0, 1H), 6.69–6.71 (m, 2H), 6.62 (d, *J* = 1.5, 1H), 6.08 (d, *J* = 1.0, 1H), 4.63 (br s, 1H), 3.55–3.61 (m, 2H), 2.49 (d, *J* = 14.5, 1H), 2.41 (s, 3H), 2.36–2.42 (m, 1H), 2.13 (ddd, *J* = 9.0, 4.0, 1.5, 1H), 1.85 (ddd, *J* = 9.0, 4.0, 1.5), 1.71 (d, *J* = 14.5, 1H); ^{13}C NMR (125 MHz, CDCl3): δ 167.9, 144.6, 144.4, 137.1, 136.2, 134.1, 131.8, 131.4, 131.3, 130.9, 130.4, 130.2, 130.1, 129.3, 127.8, 126.8, 123.4, 123.2, 121.1, 120.7, 113.8, 55.9 38.9, 36.7, 33.4, 21.7; IR (film): 3405, 2926, 1772, 1710, 1595, 1474 cm^{-1}; HRMS-ESI (*m/z*) $[M + H]^+$ calc $C_{32}H_{24}N_3BrO_4Cl_2S$, 698.0106; found, 698.0104.

3.10 Notes and References

(1) a) Siengalewicz, P.; Gaich, T.; Mulzer, J. *Angew. Chem. Int. Ed.* **2008**, *47*, 8170–8176. b) Zhang, D.; Song, H.; Qin, Y. *Acc. Chem. Res.* **2011**, *44*, 447–457. c) Zuo, Z.; Ma, D. *Isr. J. Chem.* **2011**, *51*, 434–441.

(2) Numata, A.; Takahashi, C.; Ito, Y. Takada, T.; Kenzo, K.; Usami, Y.; Matsumura, E.; Imachi, M.; Ito, T.; Hasegawa, T. *Tetrahedron Lett.* **1993**, *34*, 2355–2358.

(3) a) Jadulco, R.; Edrada, R.A.; Ebel, R.; Berg, A.; Schaumann, K.; Wray, V.; Steube, K.; Proksch. *J. Nat. Prod.* **2004**, *67*, 78–81. b) Hayashi, H.; Matsumoto, H.; Akiyama, K. *Biosci., Biotechnol., Biochem.* **2004**, *68*, 753–756. c) Dalsgaard, P. W.; Blunt, J. W.; Munro, M. H. G.; Frisvad, J. C.; Christophersen, C. *J. Nat. Prod.* **2005**, *68*, 258–261.

(4) Verbitski, S. M.; Mayne, C. L.; Davis, R. A.; Concepcion, G. P.; Ireland, C. M. *J. Org. Chem.* **2002**, *67*, 7124–7126.

(5) a) Ratnayake, A. S.; Yoshida, W.; Mooberry, S. L.; Hemscheidt, T. K. *J. Org. Chem.* **2001**, *66*, 8717. b) Erratum: Ratnayake, A. S.; Yoshida, W.Y.; Mooberry, S. L.; Hemscheidt, T. K. *J. Org. Chem.* **2003**, *68*, 1640.

(6) Compounds **3.1**–**3.9** are numbered in a manner that is consistent with that of Numata et al. (see reference 2).

(7) a) May, J. A.; Zeidan, R. K.; Stoltz, B. M. *Tetrahedron Lett.* **2003**, *44*, 1203–1205. b) Crawley, S. L.; Funk, R. L. *Org. Lett.* **2003**, *5*, 3169–3171. c) May, J. A.; Stoltz, B. M. *Tetrahedron* **2006**, *62*, 5262–5271. d) Yang, J.; Song, H.; Xiao, X.; Wang, J.; Qin, Y. *Org. Lett.* **2006**, *8*, 2187–2190. e) Crawley, S. L.; Funk, R. L. *Org. Lett.* **2006**, *8*, 3995–3998. f)

Sabahi, A.; Novikov, A.; Rainier, J. D. *Angew. Chem. Int. Ed.* **2006**, *45*, 4317–4320. g) George, J. H.; Adlington, R. M. *Synlett* **2008**, 2093–2096.

(8) Liu, P.; Seo, J. H.; Weinreb, S. M. *Angew. Chem. Int. Ed.* **2010**, *49*, 2000–2003.

(9) Fuchs, J. R.; Funk, R. L. *J. Am. Chem. Soc.* **2004**, *126*, 5068–5069.

(10) a) Zuo, Z.; Ma, D. *Angew. Chem. Int. Ed.* **2011**, *50*, 12008–12011. b) Zuo, Z.; Xie, W.; Ma, D. *J. Am. Chem. Soc.* **2010**, *132*, 13226–13228.

(11) a) Yang, J.; Wu, H.; Shen, L.; Qin, Y. *J. Am. Chem. Soc.* **2007**, *129*, 13794–13795. b) Wu, H.; Xue, F.; Xiao, X.; Qin, Y. *J. Am. Chem. Soc.* **2010**, *132*, 14052–14054.

(12) a) Boal, B. W.; Schammel, A. W.; Garg, N. K. *Org. Lett.* **2009**, *11*, 3458–3461. b) Schammel, A. W.; Boal, B. W.; Zu, L.; Mesganaw, T.; Garg, N. K. *Tetrahedron* **2010**, *66*, 4687–4695. c) Çelebi–Ölçüm, N.; Boal, B. W.; Huters, A. D.; Garg, N. K; Houk, K. N. *J. Am. Chem. Soc.* **2011**, *133*, 5752–5755. d) Zu, L.; Boal B. W.; Garg, N. K. *J. Am. Chem. Soc.* **2011**, *133*, 8877–8879. e) Schammel, A. W.; Chiou, G.; Garg, N. K. *J. Org. Chem.* **2012**, *77*, 725–728.

(13) For reviews of the classic Fischer indole synthesis, see: a) Robinson, B. *Chem. Rev.* **1963**, *63*, 373–401. b) Robinson, B. *Chem. Rev.* **1969**, *69*, 227–250.

(14) For seminal studies of the Fischer indole synthesis, see: a) Fischer, E.; Jourdan, F. *Ber.* **1883**, *16*, 2241–2245. b) Fischer, E.; Hess, O. *Ber.* **1884**, *17*, 559–568.

(15) For the use of the interrupted Fischer indolization to access flustramine derivatives, see: Bunders, C.; Cavanagh, J.; Melander, C. *Org. Biolmol. Chem.* **2011**, *9*, 5476–5481.

(16) a) Paz Muñoz, M.; de la Torre, M. C.; Sierra, M. A. *Adv. Synth. Catal.* **2010**, *352*, 2189–2194. b) Kowalski, P.; Jaśkowska, J. *Arch. Pharm. (Weinheim, Ger.)* **2012**, *345*, 81–85.

(17) Steinhagen, H.; Corey, E. J. *Angew. Chem. Int. Ed*. **1999**, *38*, 1928–1931.

(18) The formation of NMe and NH products is consistent with previous observations in the interrupted Fischer indolization; see reference 12a,b.

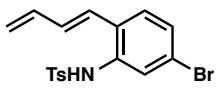
(19) Srivastava, S.; Ruane, P. H.; Toscano, J. P.; Sullivan, M. B.; Cramer, C. J.; Chiapperino, D.; Reed, E. C.; Falvey, D. E. *J. Am. Chem*. *Soc.* **2000**, *122*, 8271−8278.

(20) Carbamylated *N,O*-acetal **3.23** was employed in this study because it was found to be more readily removable compared to the *N*-Ts group.

(21) For reviews of the Pictet–Spengler reaction, see: a) Cox, E. D.; Cook, J. M. *Chem. Rev.* **1995**, *95*, 1797–1842. b) Stökigt, J.; Antonchick, A. P.; Wu, F.; Waldmann, H. *Angew. Chem. Int. Ed.* **2011**, *50*, 8538–8564. For Pictet–Spengler reactions on similar systems, see: c) Mahboobi, S.; Eluwa, S.; Koller, M.; Popp, A.; Schollmeyer, D. *J. Heterocycl. Chem.* **2000**, *37*, 1177–1185. d) Somei, M.; Teranishi, S.; Yamada, K.; Yamada, F. *Chem. Pharm. Bull.* **2001**, *49*, 1159–1165. e) Yamada, K.; Namerikawa, Y.; Haruyama, T.; Miwa, Y.; Yanada, R.; Ishikura, M. *Eur. J. Org. Chem.* **2009**, 5752–5759.

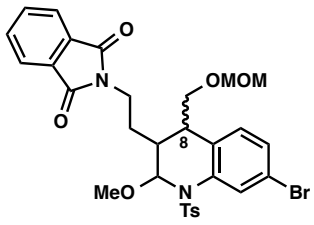
(22) Attempts to form the G ring through intermediate **3.26** led to hydrolysis of the imine, thus returning **3.25**. Base induced Pictet–Spengler reaction attempts led to the recovery of starting material (see ref. 21e). Further efforts to introduce the G ring of the communesins are underway.

(23) This hetero-Diels–Alder reaction was hampered by the competitive formation of diene **i**, which presumably forms from elimination of the benzylic chloride. Efforts to suppress the formation of **3.41** were not undertaken because of the difficulty of the subsequent interrupted Fischer indolization.



3.41

(24) MOM ether **3.42** was also prepared, but unfortunately proved ineffective in the interrupted Fischer indolization reaction under a variety of conditions.



3.42

(22) The use of AcOH/H_2O conditions for this transformation failed to deliver the desired product **3.39**.

(26) Intermediate **3.40** may plausibly be elaborated by initial benzylic oxidation to functionalize C8, followed by further manipulations.

(27) Paz Muñoz, M.; de la Torre, M. C.; Sierra, M. A. *Adv. Synth. Catal.* **2010**, *352*, 2189–2194. b) Kowalski, P.; Jaśkowska, J. *Arch. Pharm. (Weinheim, Ger.)* **2012**, *345*, 81–85.

(28) Chong, P.Y.; Janicki, S. Z.; Petillo, P. A. *J. Org. Chem.* **1998**, *63*, 8515–8521.

(29) Srivastava, S.; Ruane, P. H.; Toscano, J. P.; Sullivan, M. B.; Cramer, C. J.; Chiapperino, D.; Reed, E. C.; Falvey, D. E. *J. Am. Chem.Soc*. **2000**, *122*, 8271−8278.

APPENDIX TWO

Spectra Relevant to Chapter Three:

Interrupted Fischer Indolization Approach Toward the Communesin Alkaloids and Perophoramidine

Alex W. Schammel, Grace Chiou, and Neil K. Garg.

Org. Lett. **2012**, *14*, 4556–4559.

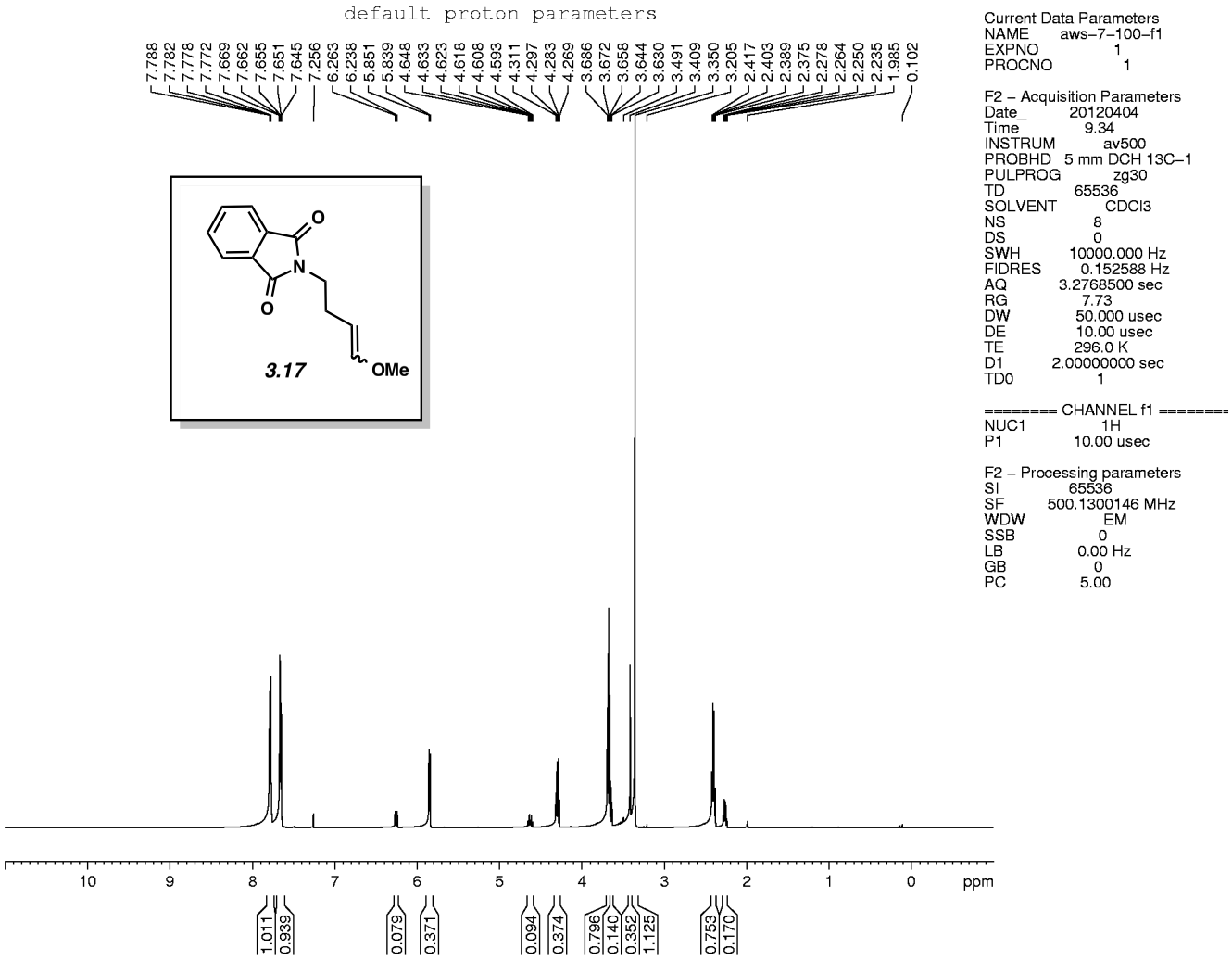


Figure A2.1 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.17**.

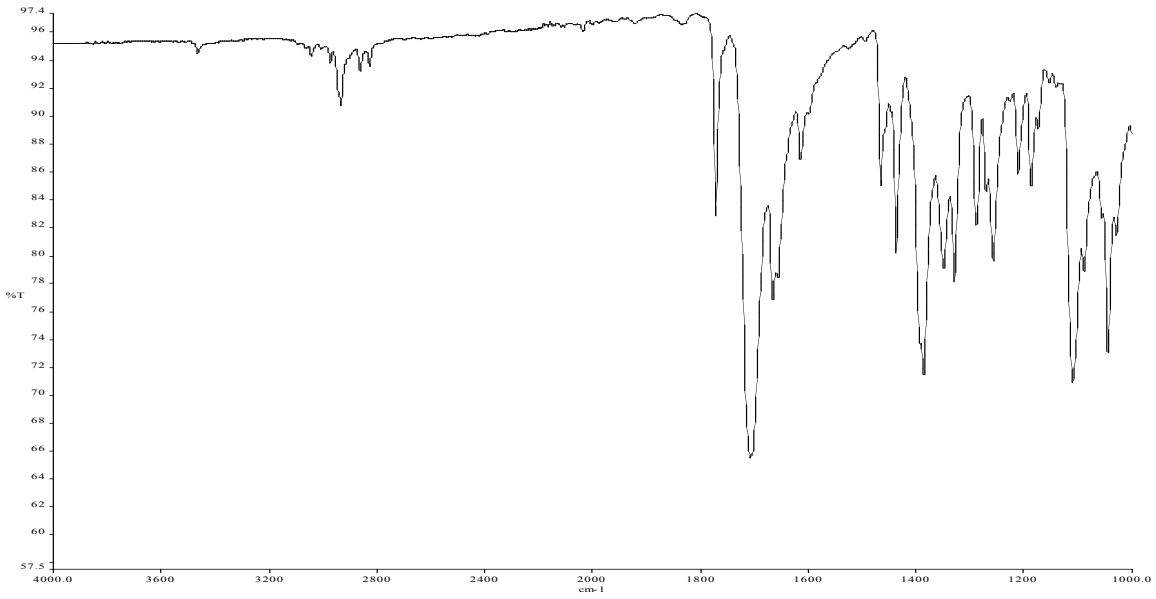


Figure A2.2 Infrared spectrum of compound **3.17**.

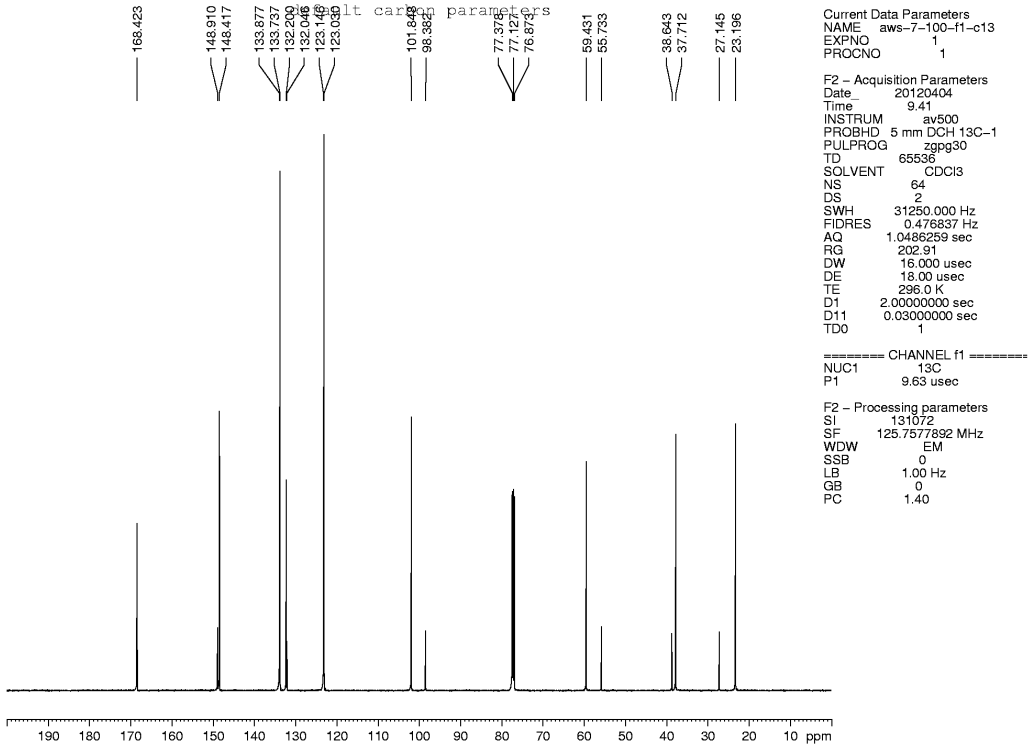


Figure A2.3 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.17**.

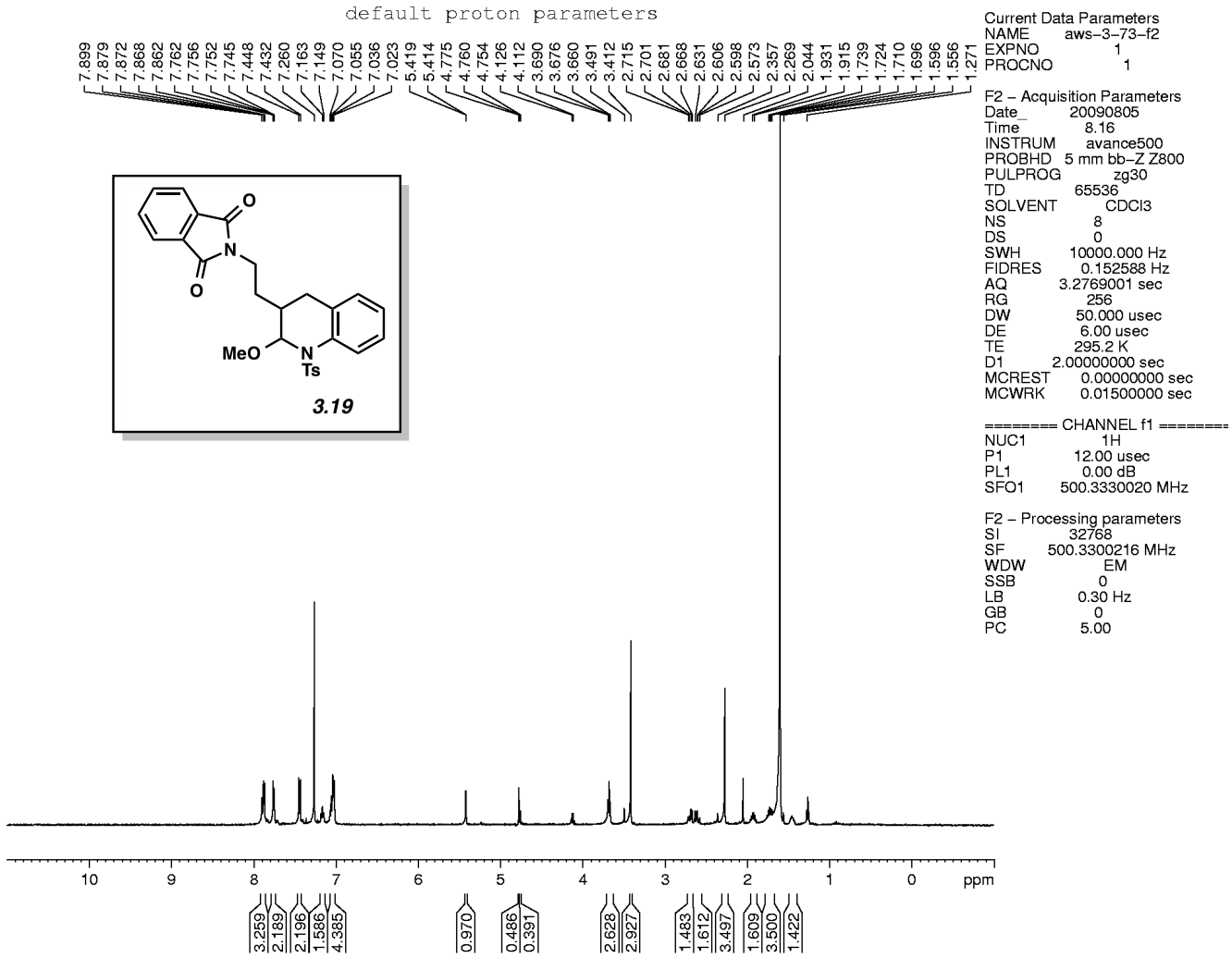


Figure A2.4 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.19**.

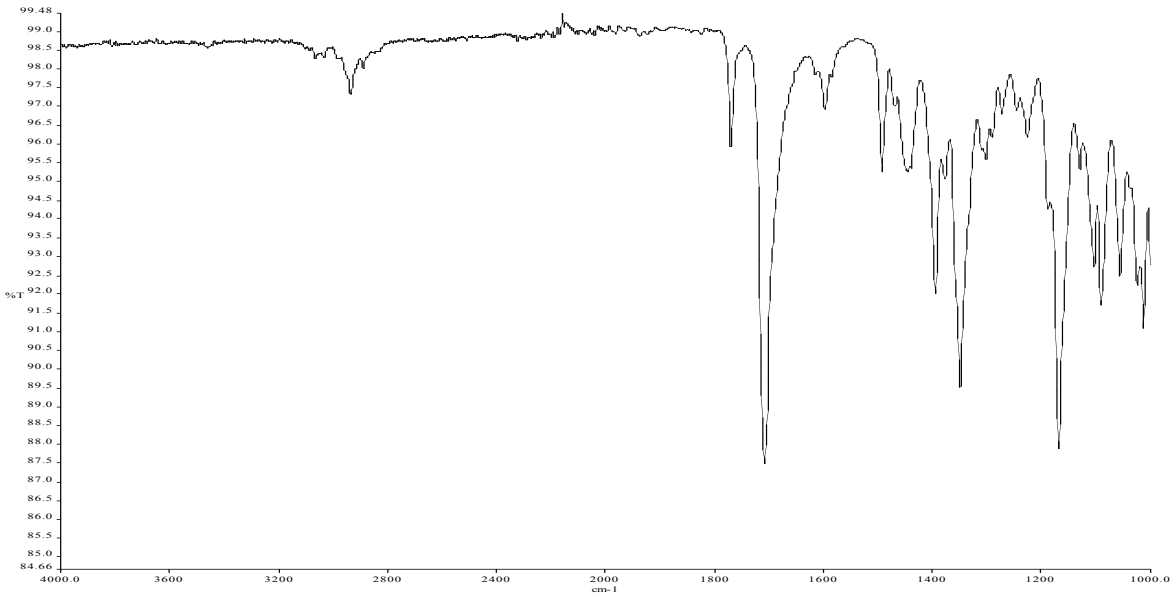


Figure A2.5 Infrared spectrum of compound **3.19**.

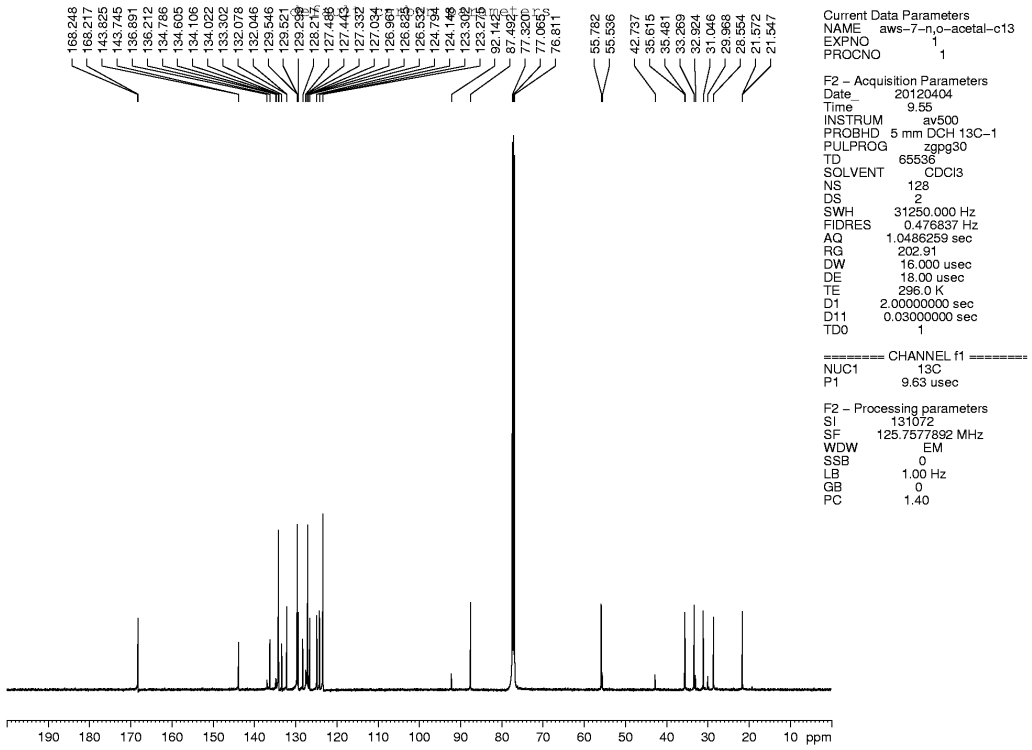


Figure A2.6 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.19**.

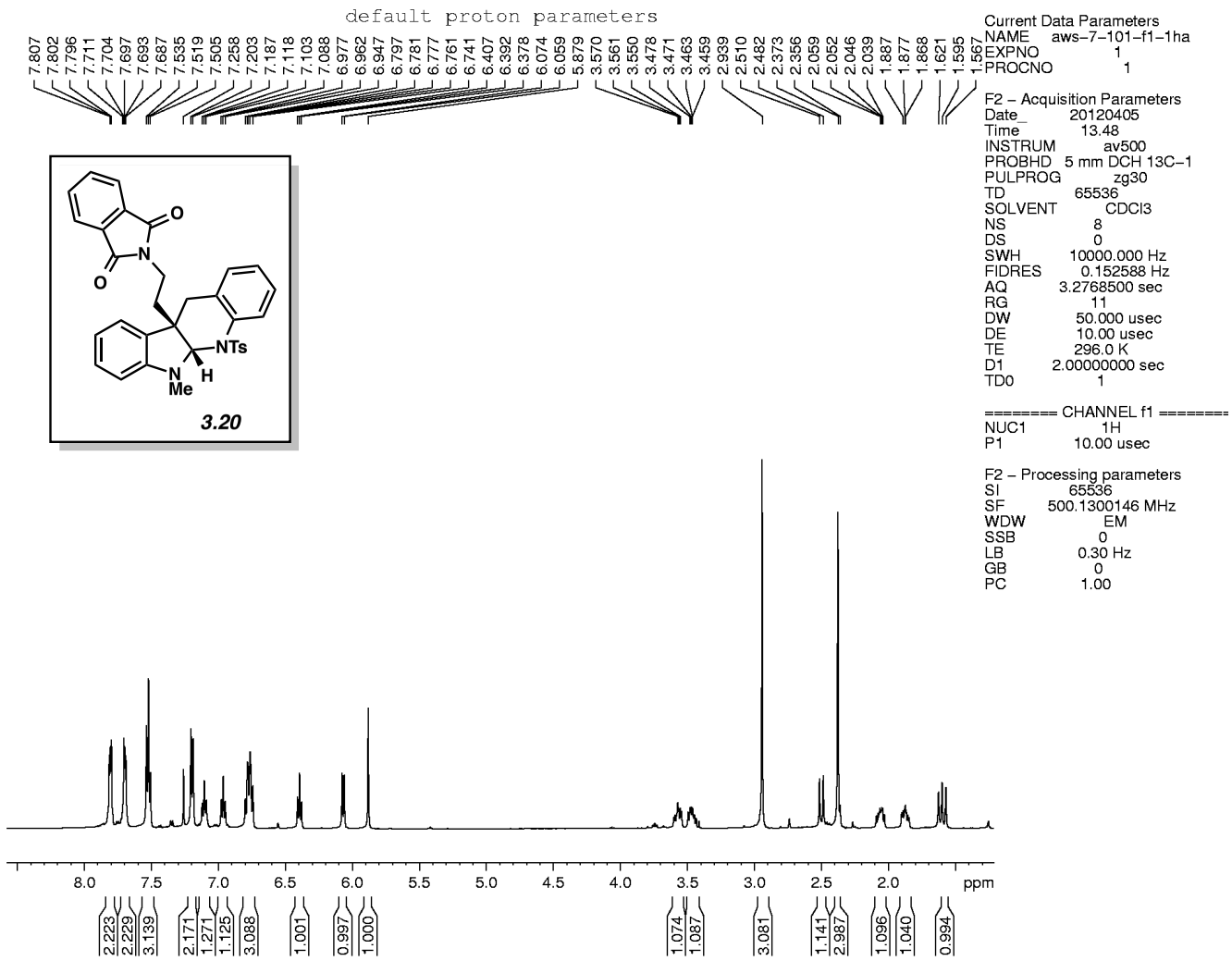


Figure A2.7 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.20**.

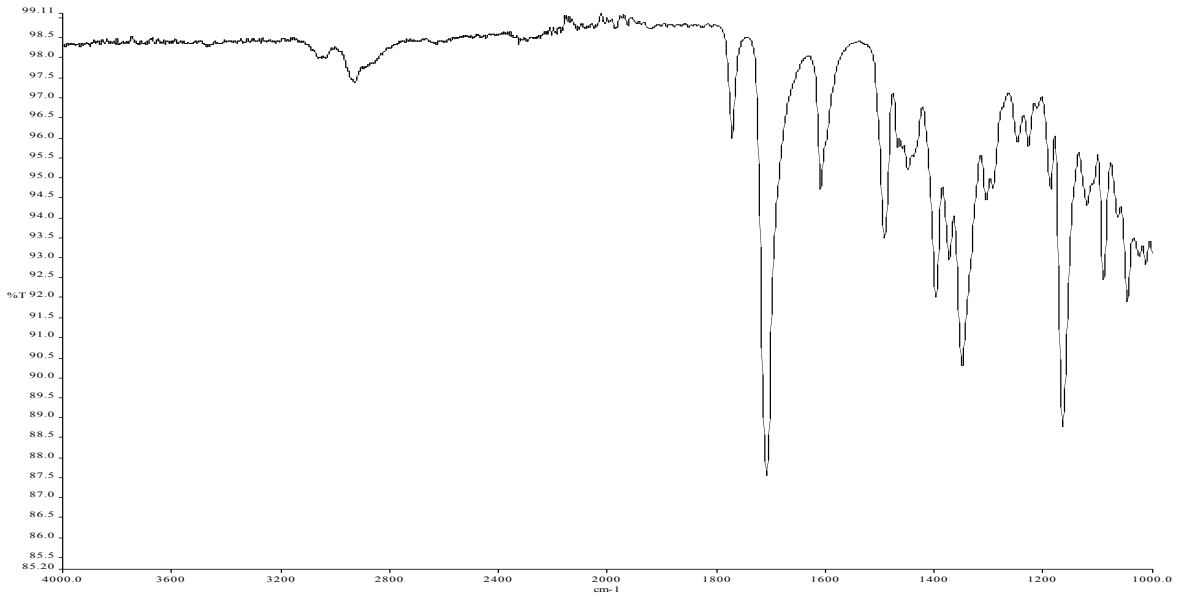


Figure A2.8 Infrared spectrum of compound **3.20**.

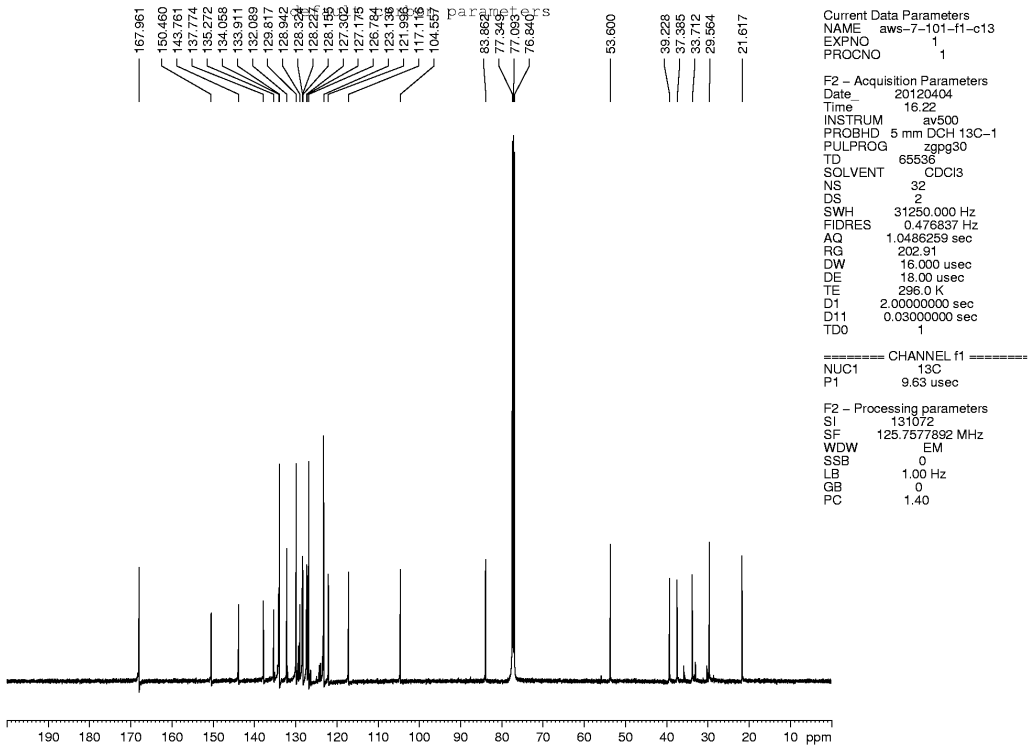


Figure A2.9 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.20**.

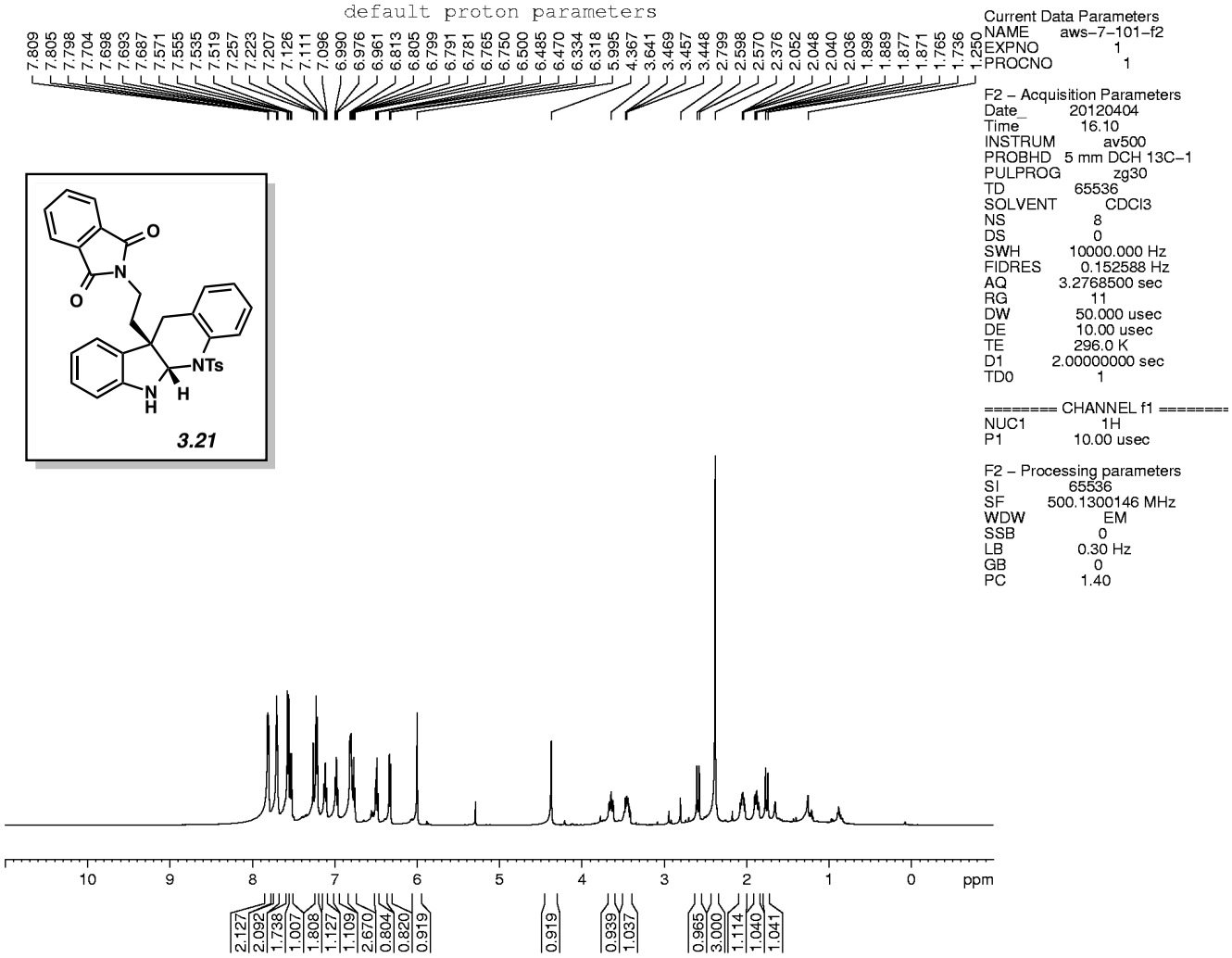


Figure A2.10 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.21**.

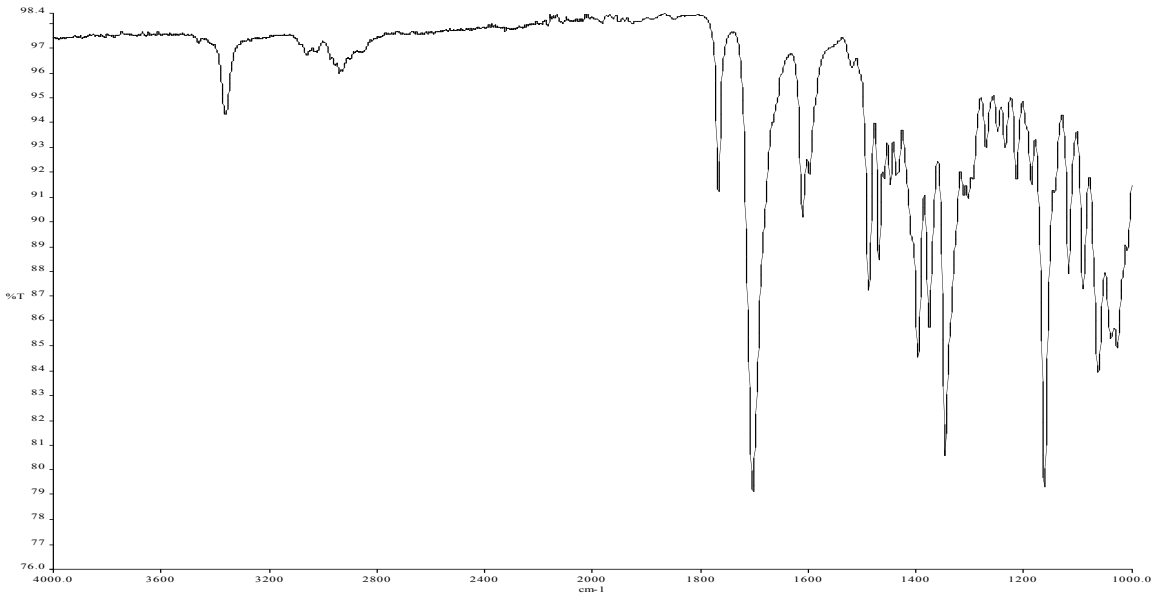


Figure A2.11 Infrared spectrum of compound **3.21**.

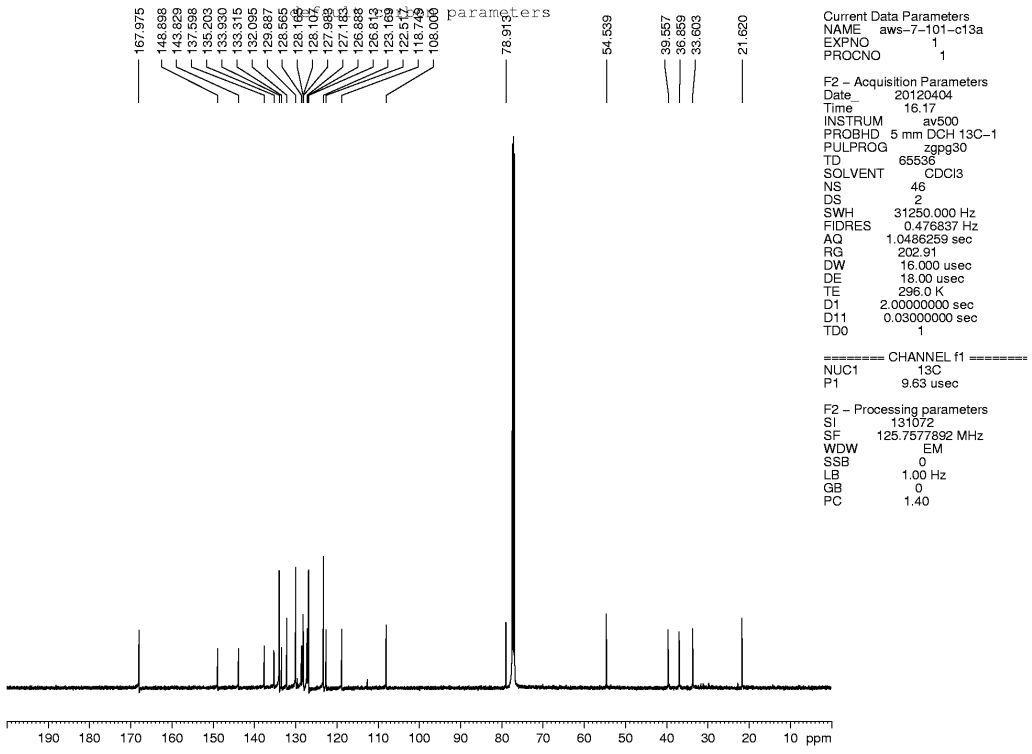


Figure A2.12 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.21**.

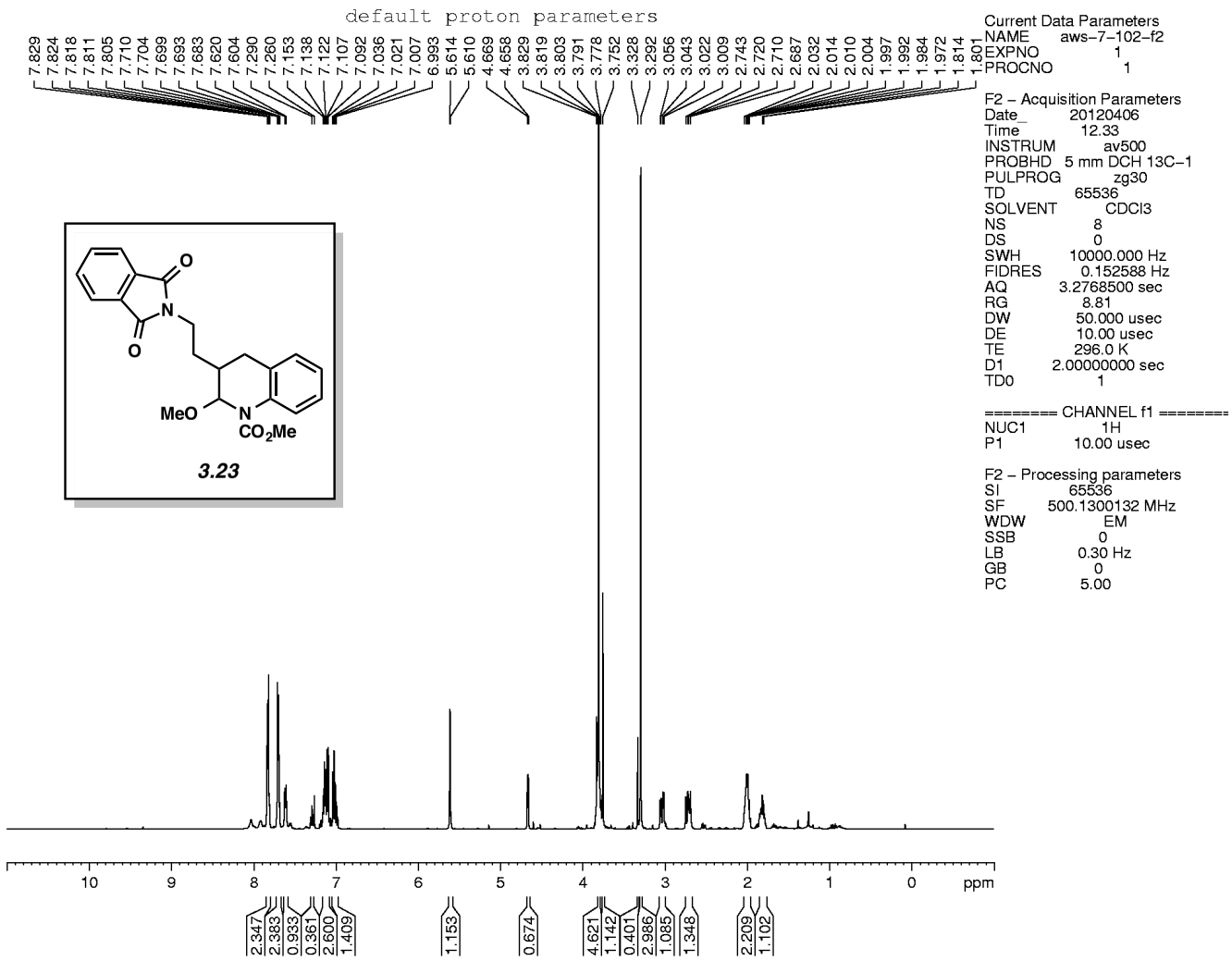


Figure A2.13 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.23**.

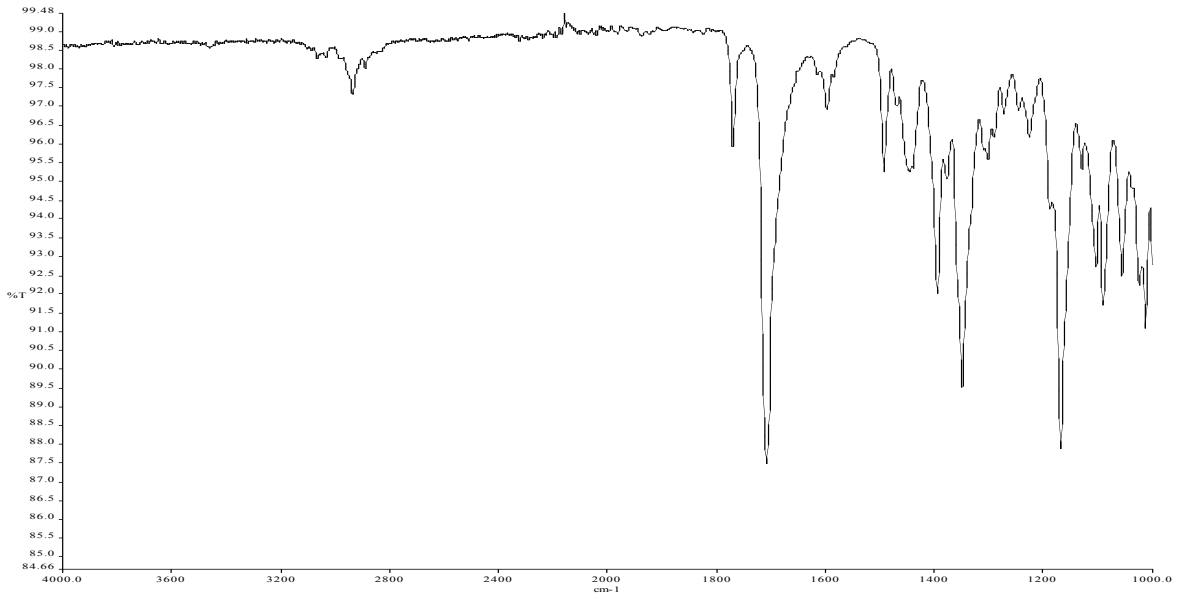


Figure A2.14 Infrared spectrum of compound **3.23**.

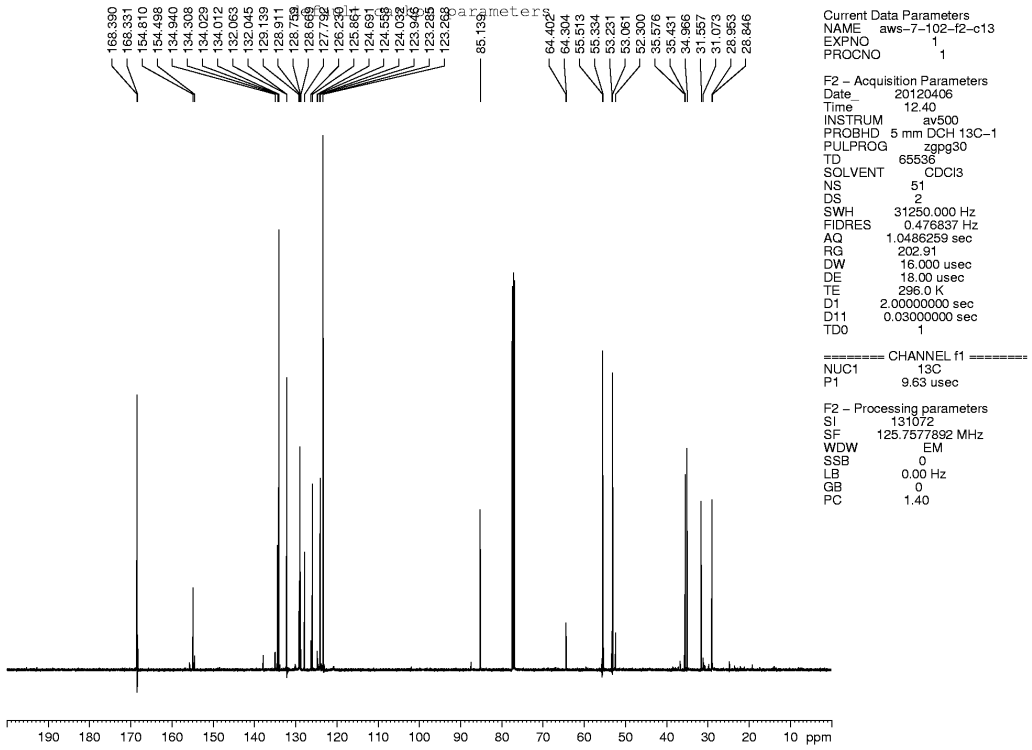


Figure A2.15 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.23**.

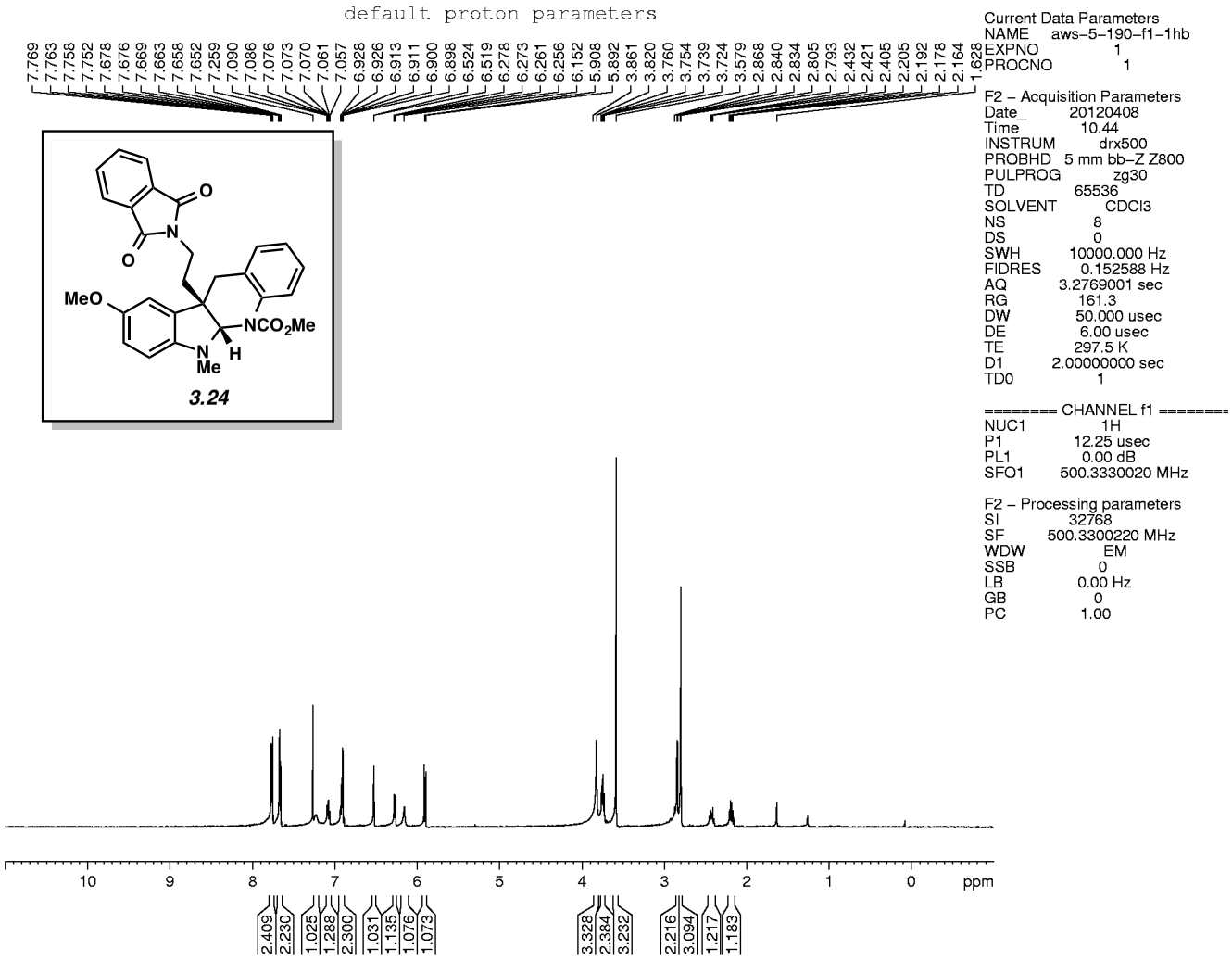


Figure A2.16 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **3.24**.

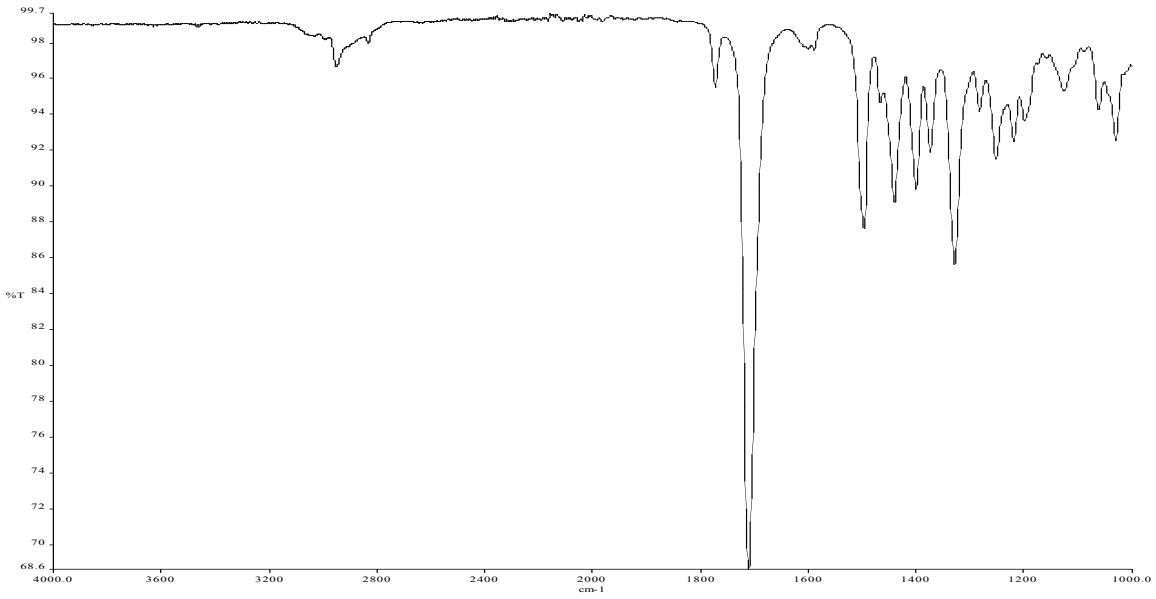


Figure A2.17 Infrared spectrum of compound **3.24**.

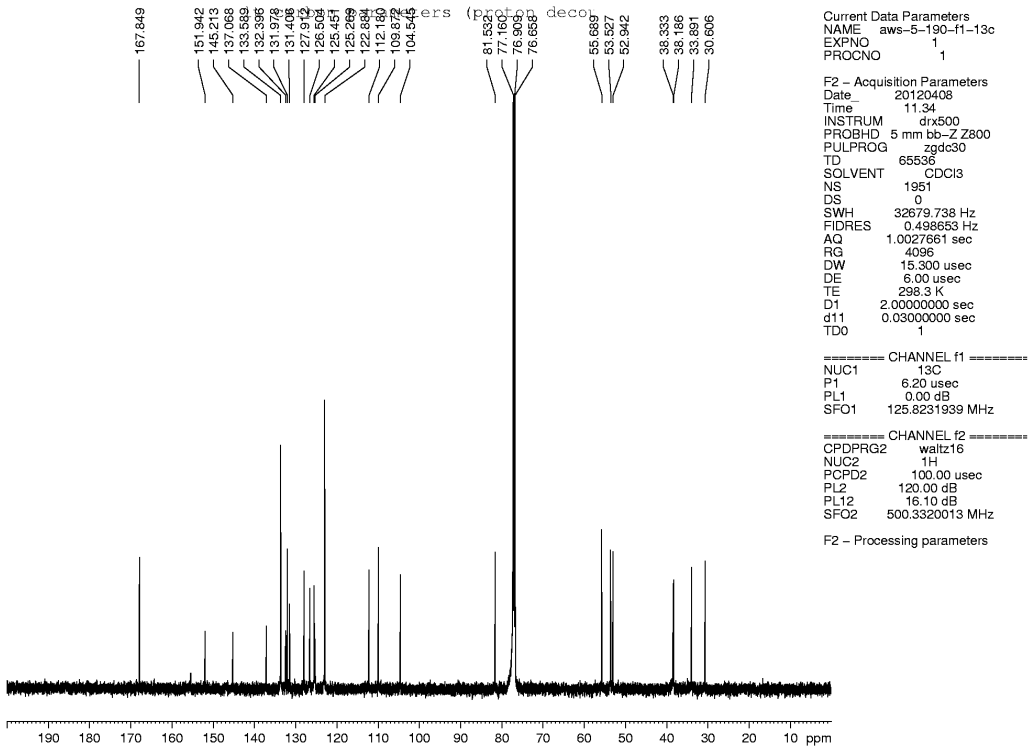


Figure A2.18 ^{13}C NMR (125 MHz, $CDCl_3$) of compound **3.24**.

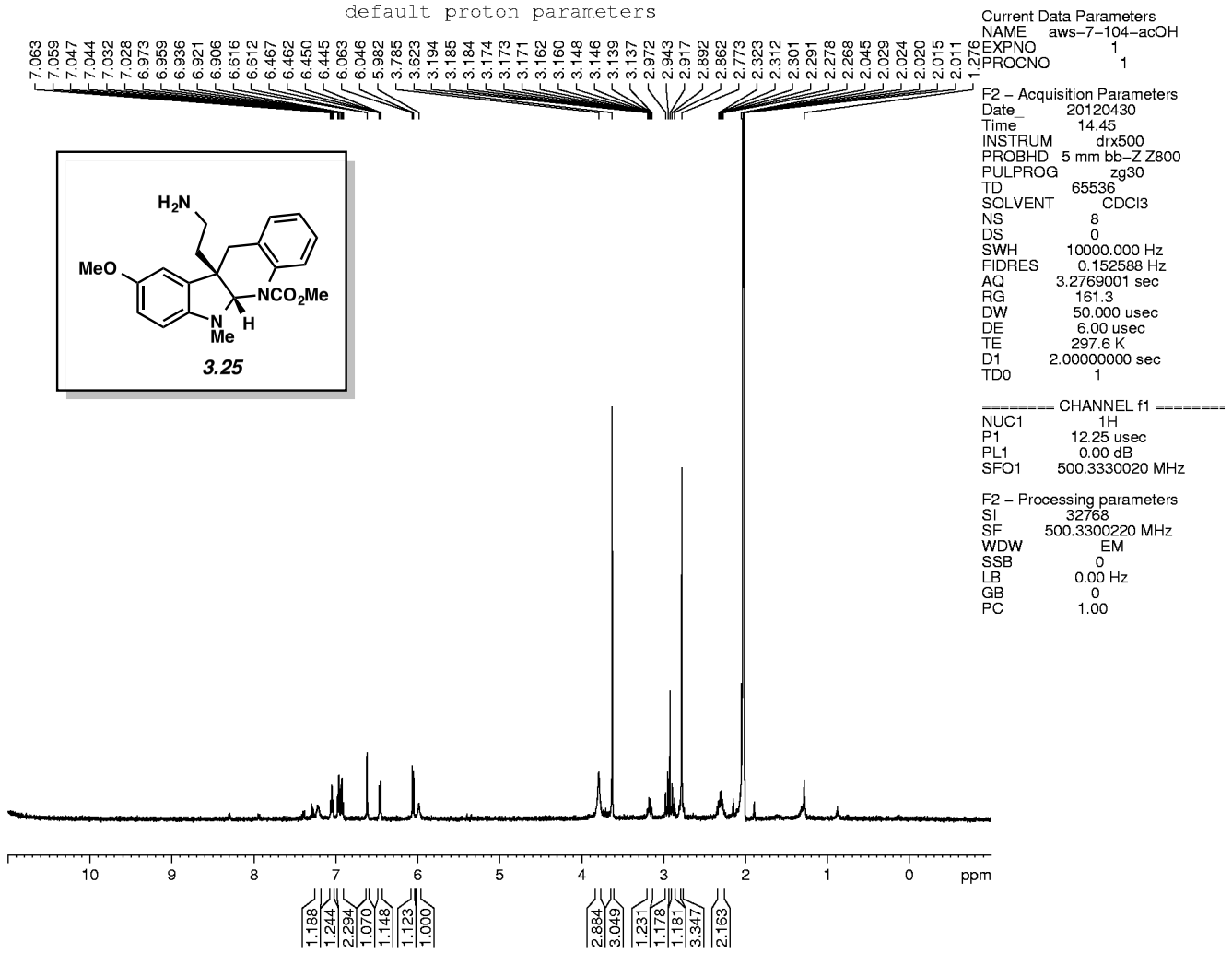


Figure A2.19 ^{1}H NMR (500 MHz, CD_3OD) of compound **3.25**.

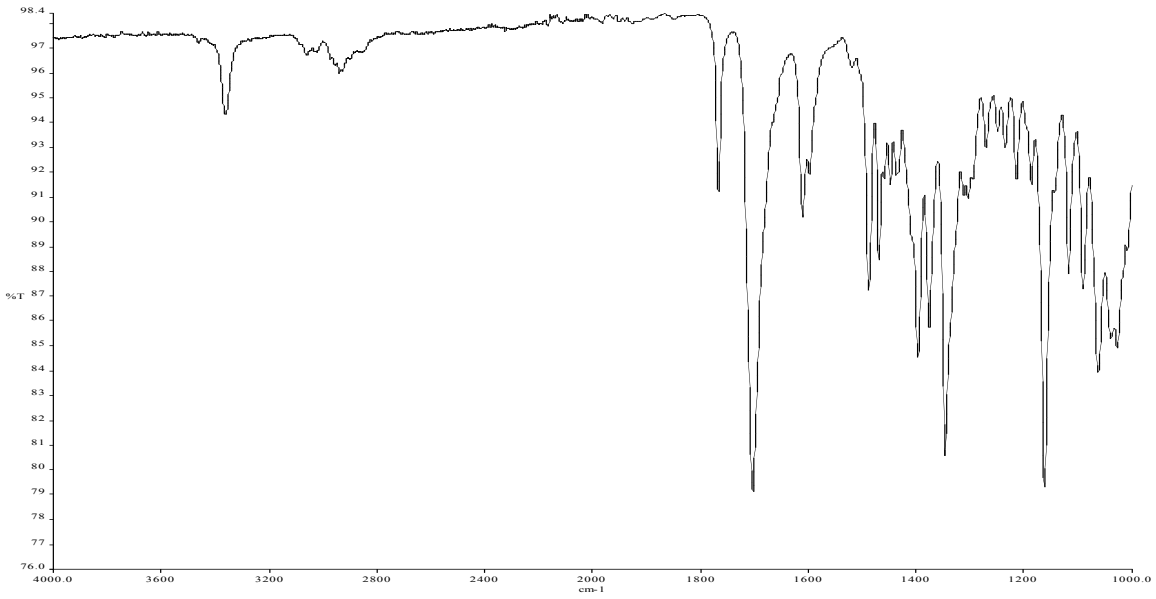


Figure A2.20 Infrared spectrum of compound **3.25**.

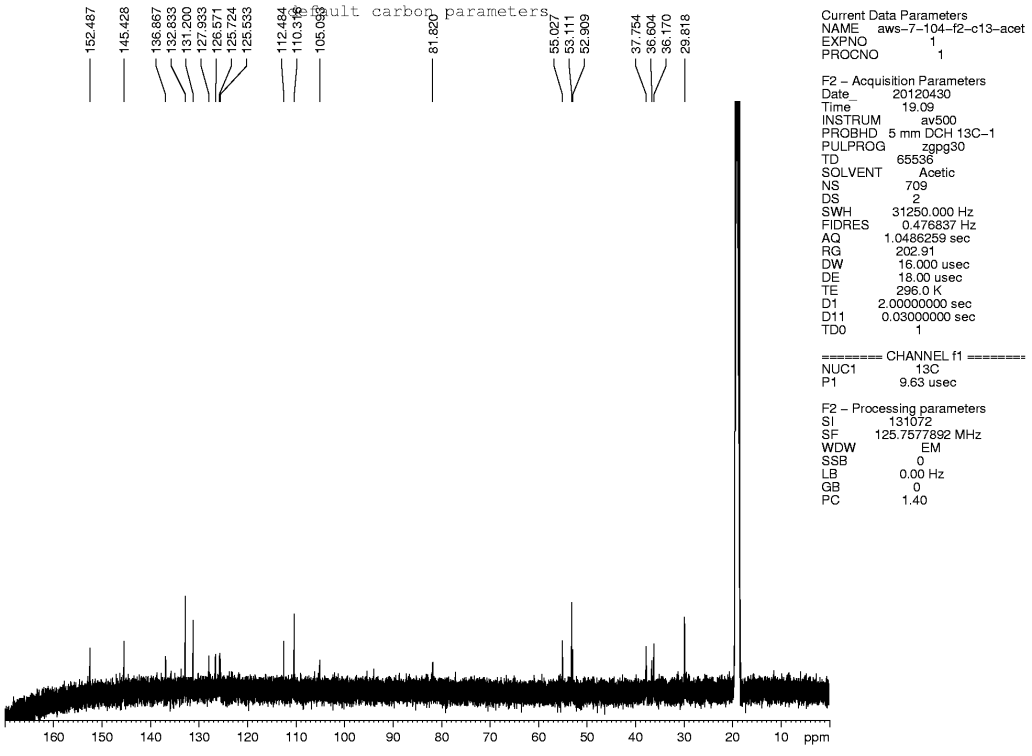


Figure A2.21 ^{13}C NMR (125 MHz, CD_3OD) of compound **3.25**.

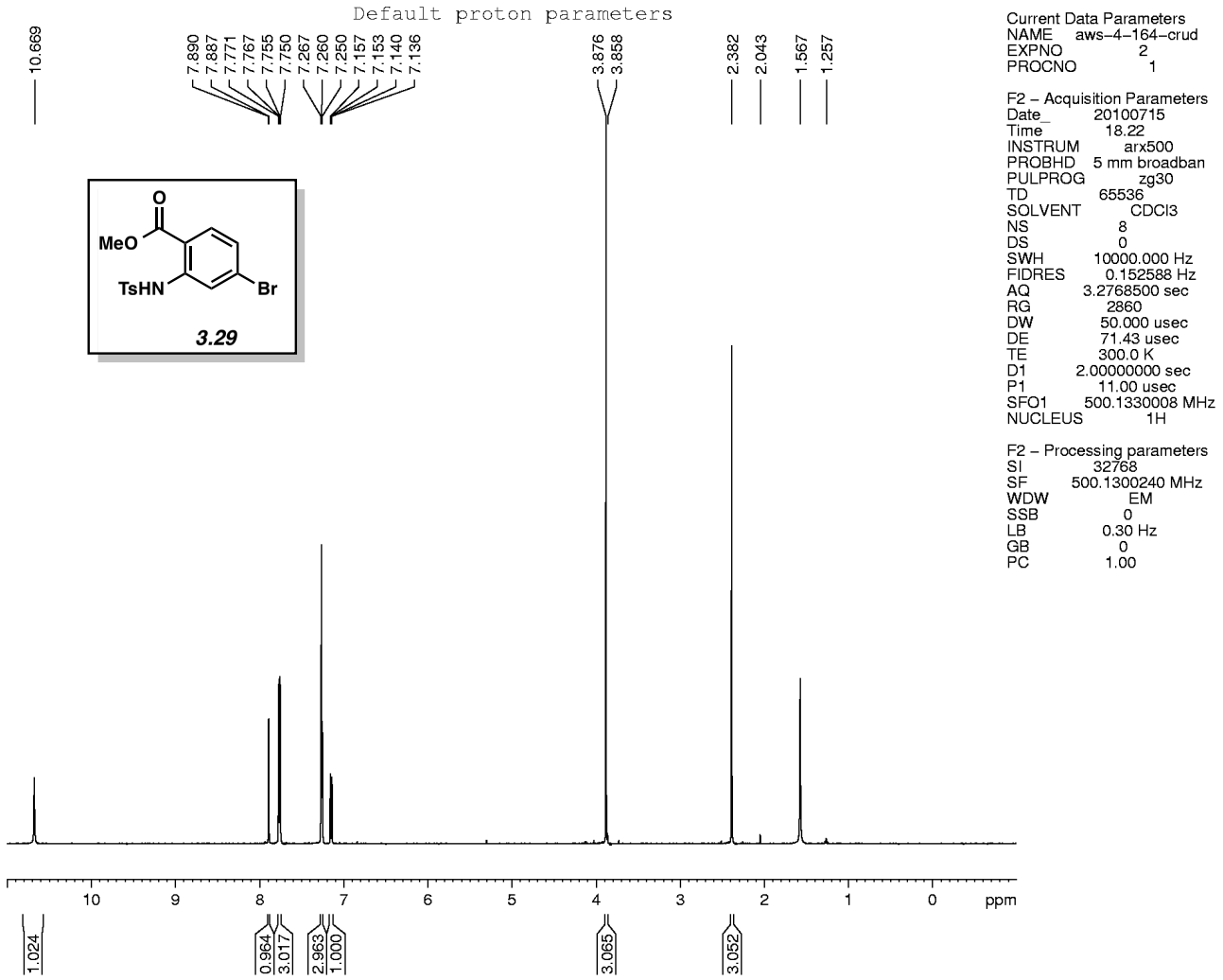


Figure A2.22 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.29**.

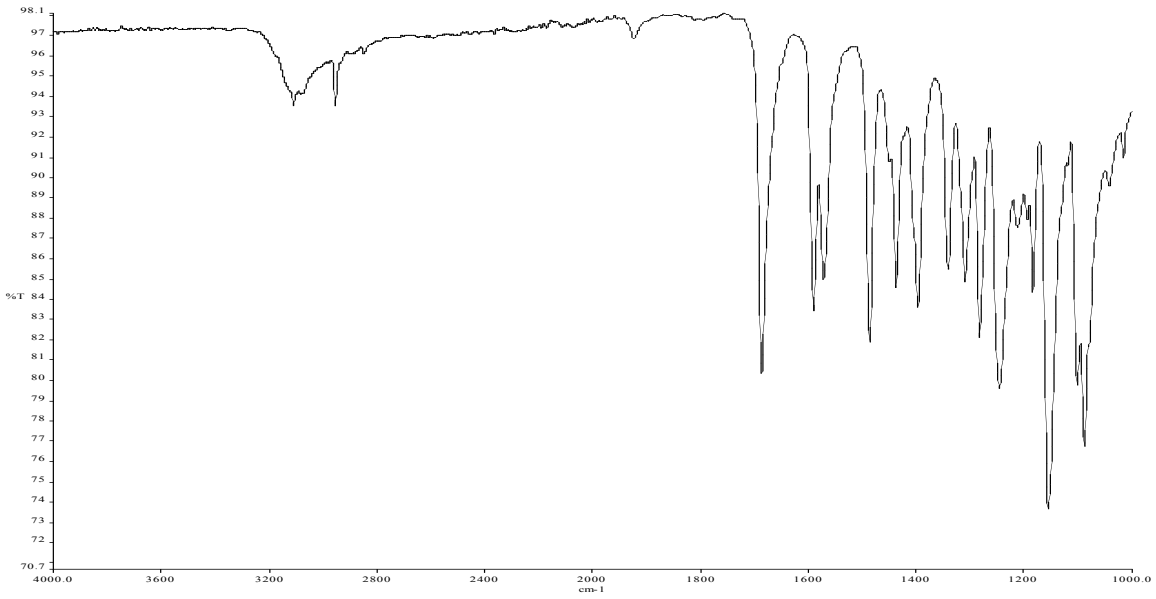


Figure A2.23 Infrared spectrum of compound **3.29**.

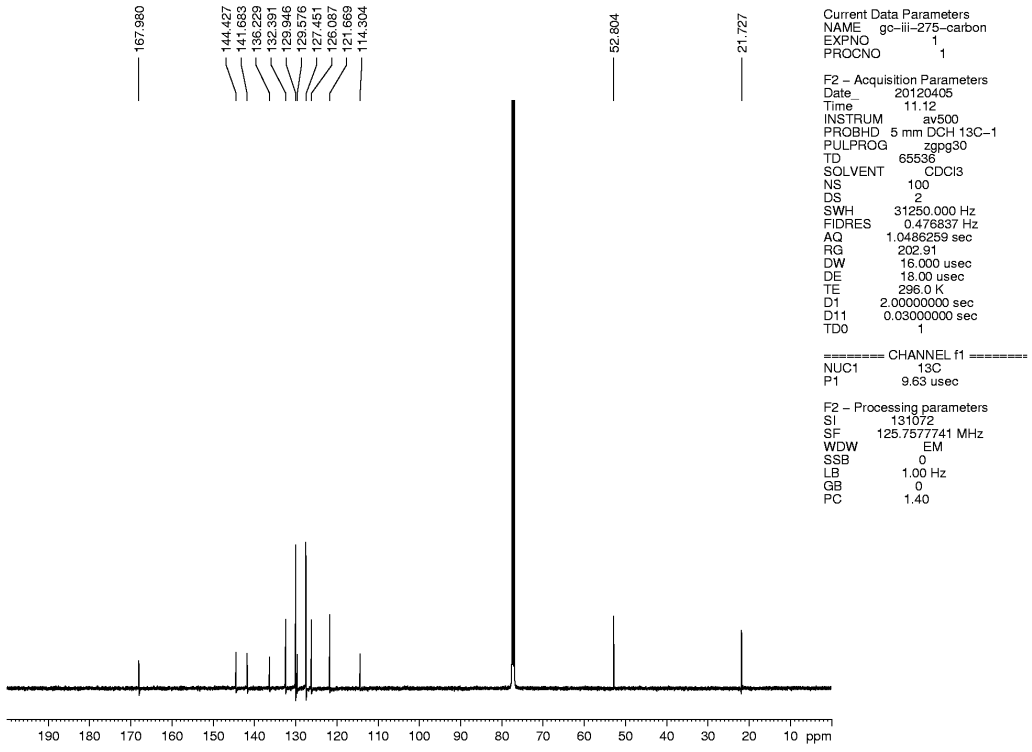


Figure A2.24 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.29**.

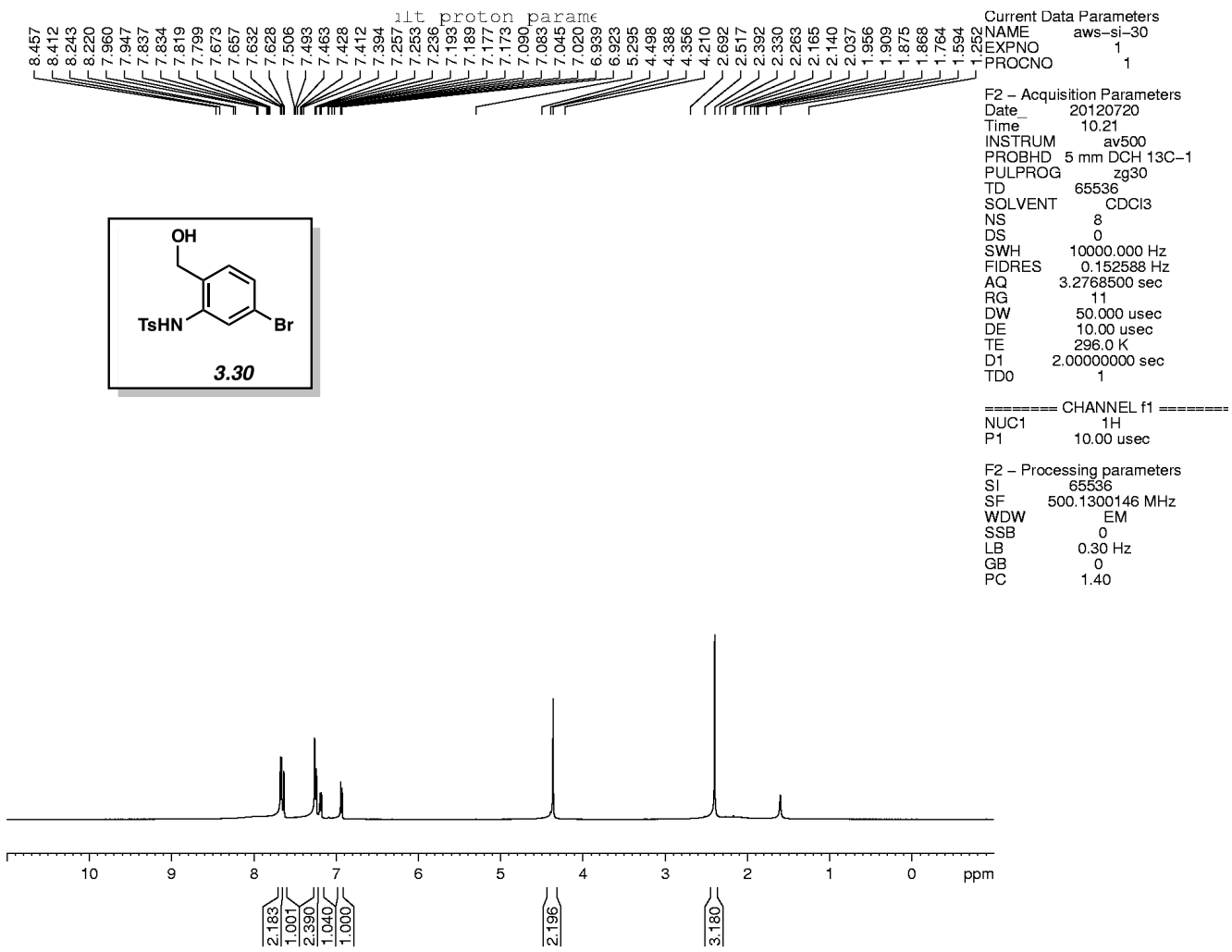


Figure A2.25 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.30**.

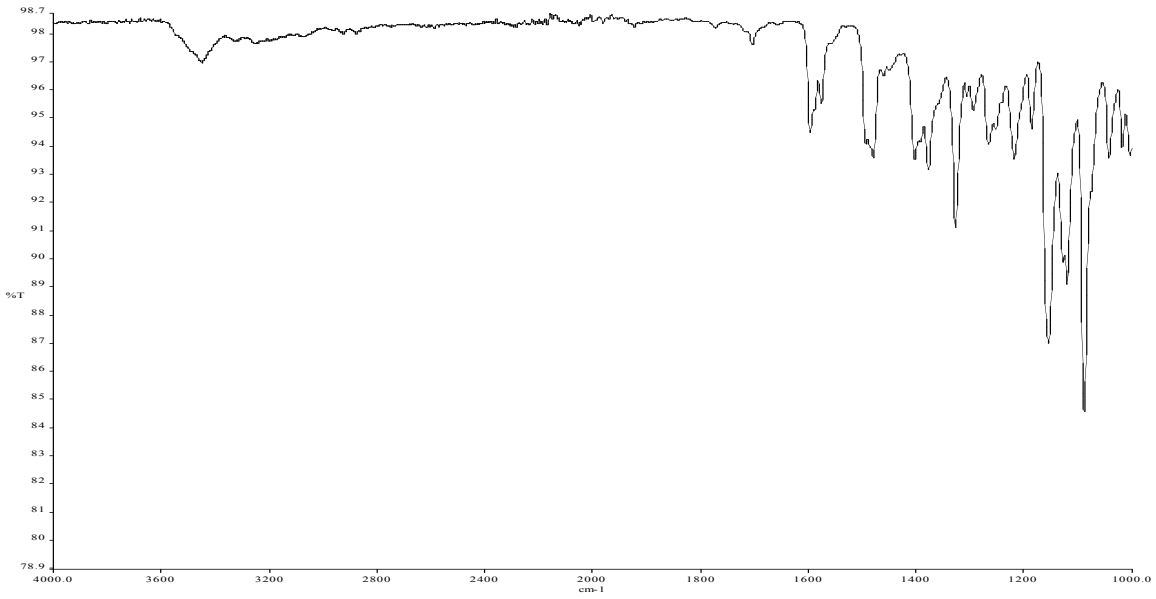


Figure A2.26 Infrared spectrum of compound **3.30**.

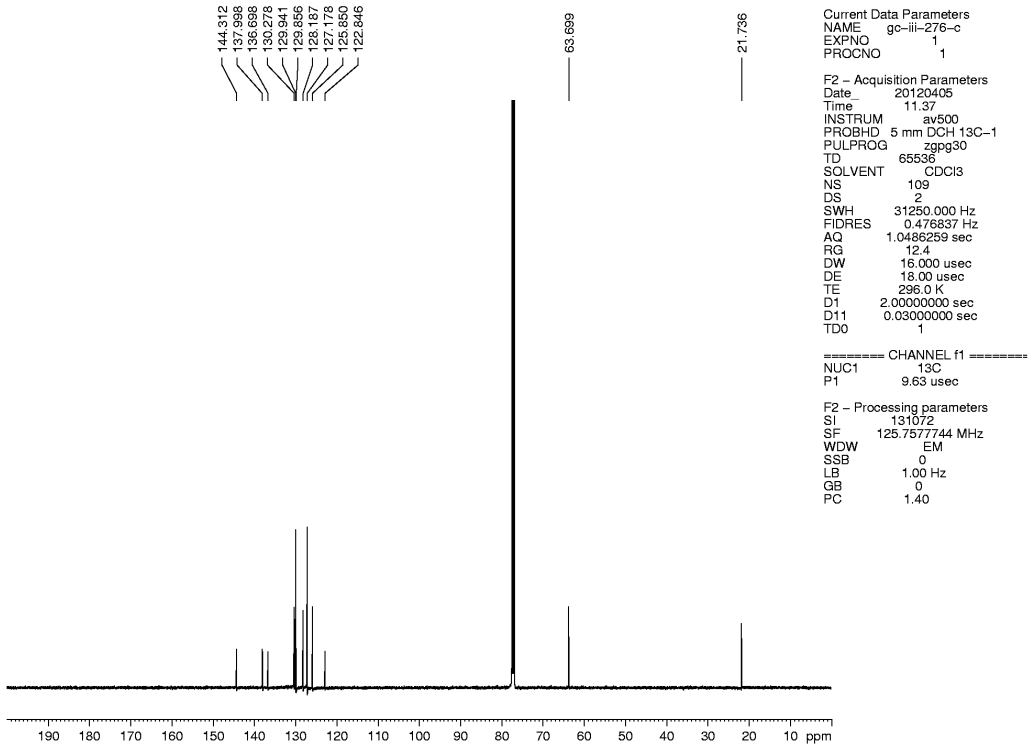


Figure A2.27 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.30**.

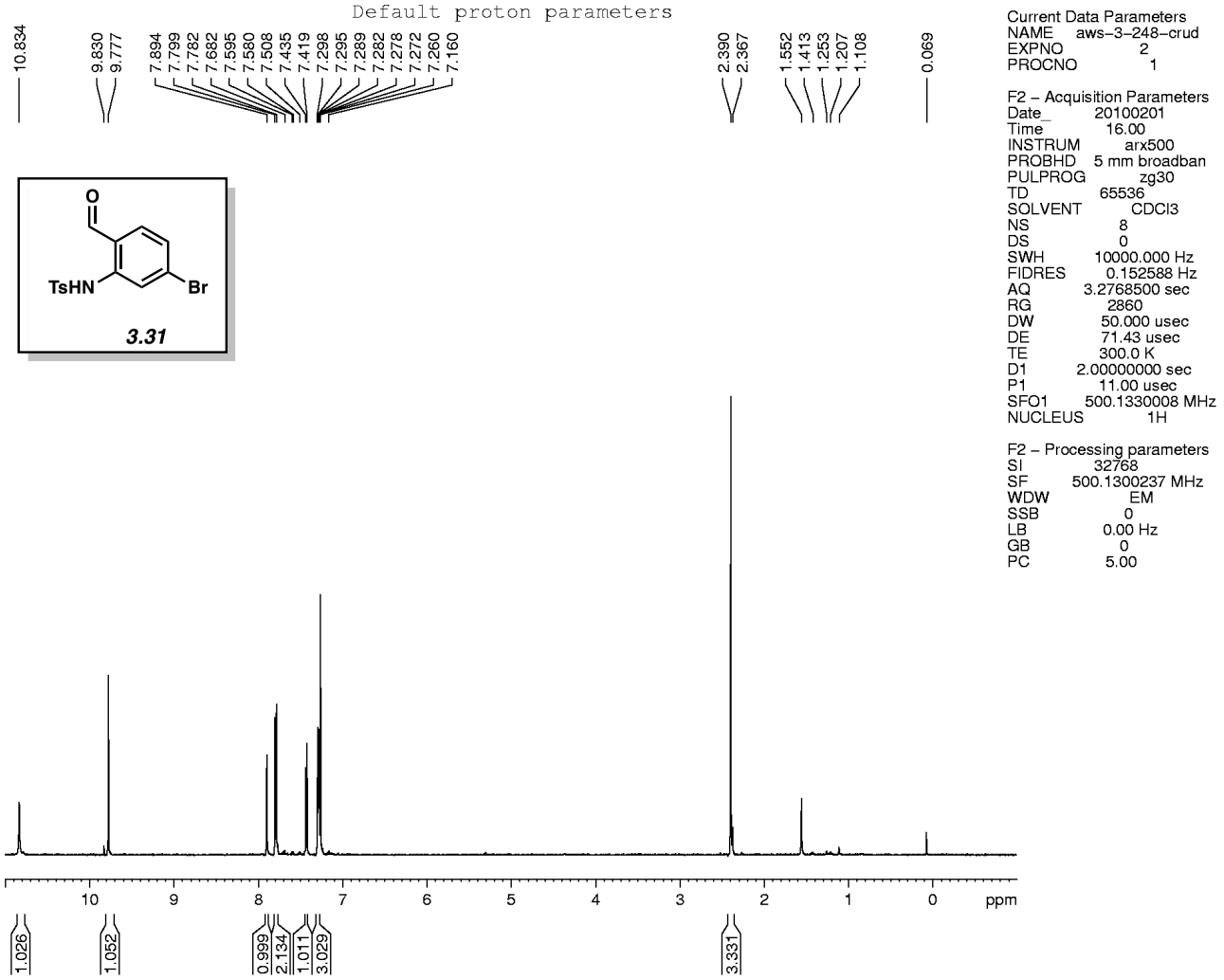


Figure A2.28 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.31**.

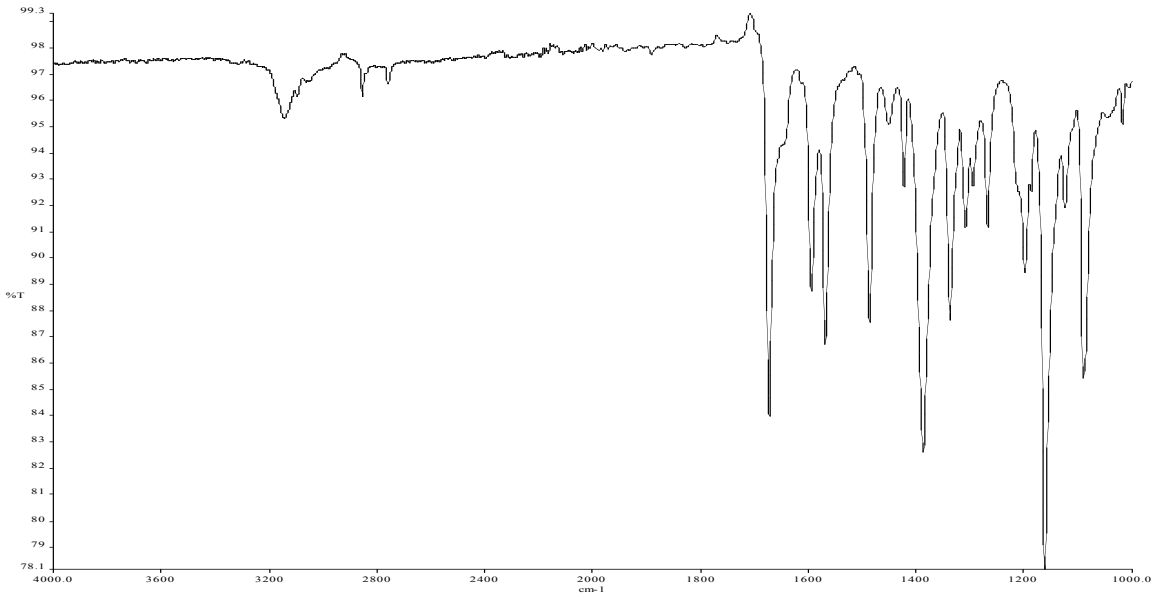


Figure A2.29 Infrared spectrum of compound **3.31**.

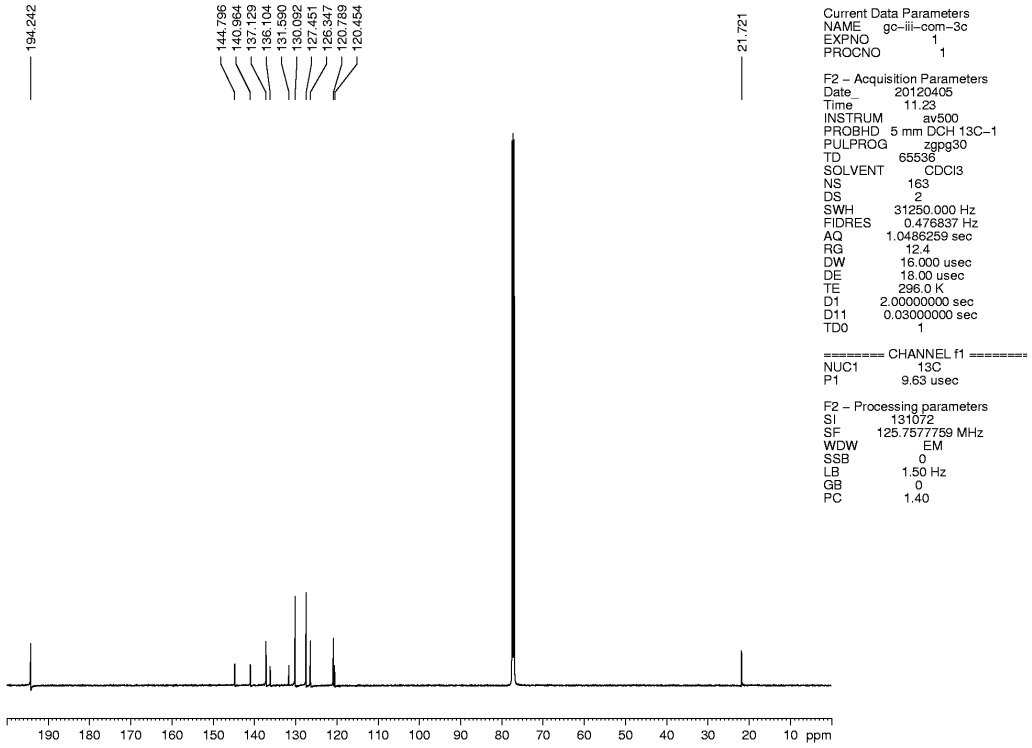


Figure A2.30 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.31**.

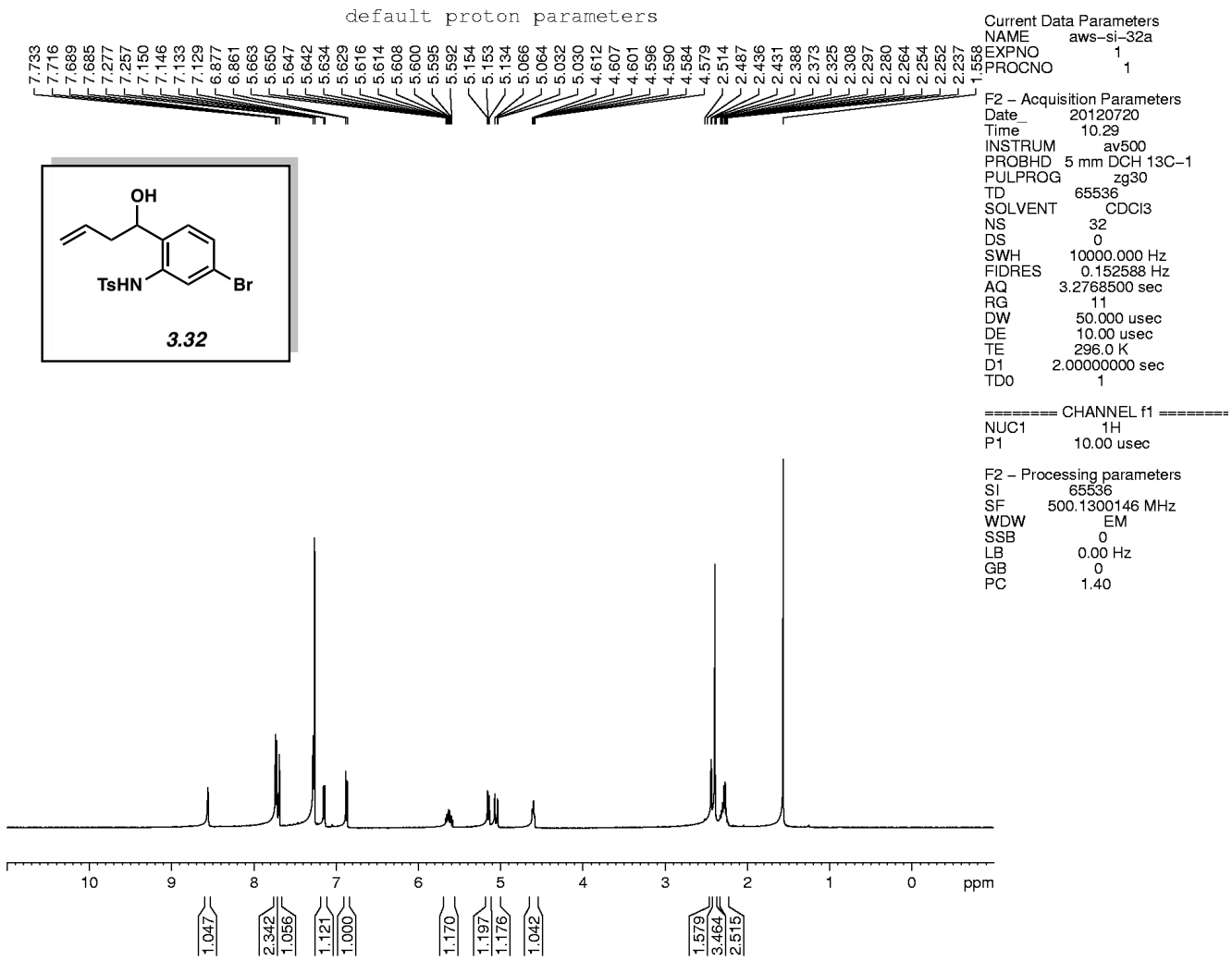


Figure A2.31 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.32**.

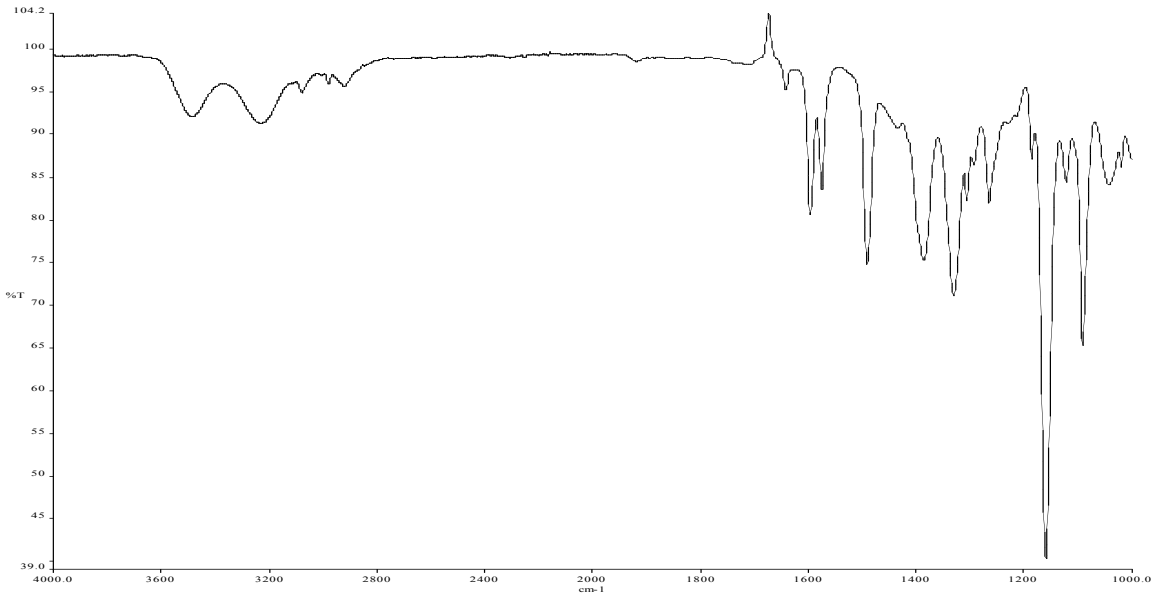


Figure A2.32 Infrared spectrum of compound **3.32**.

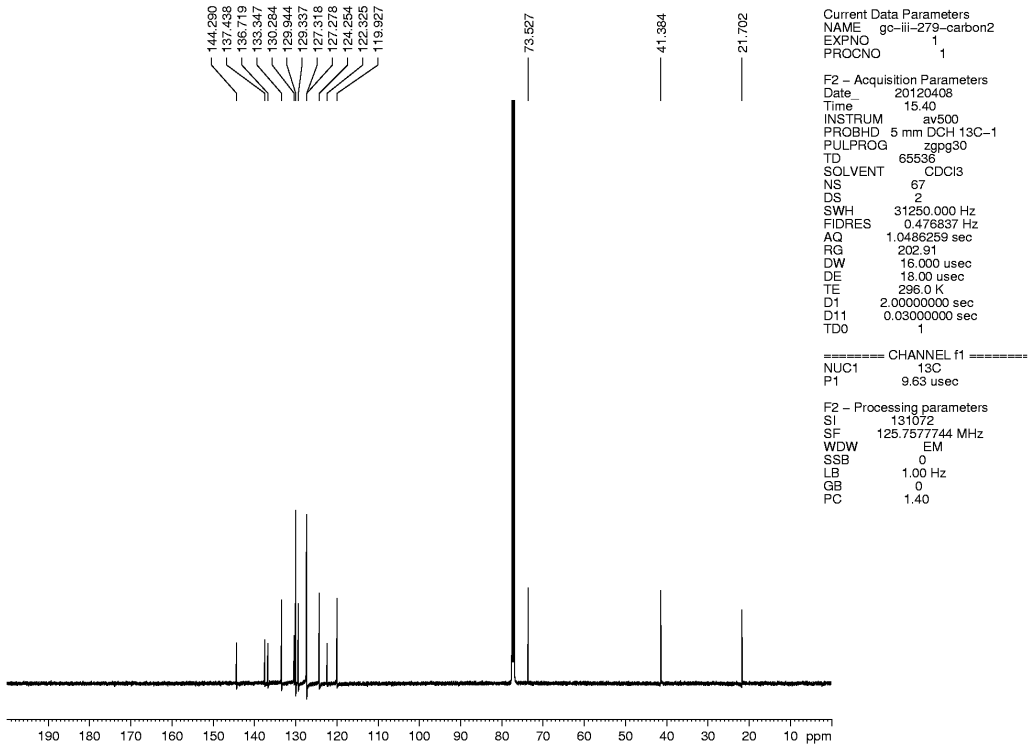


Figure A2.33 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.32**.

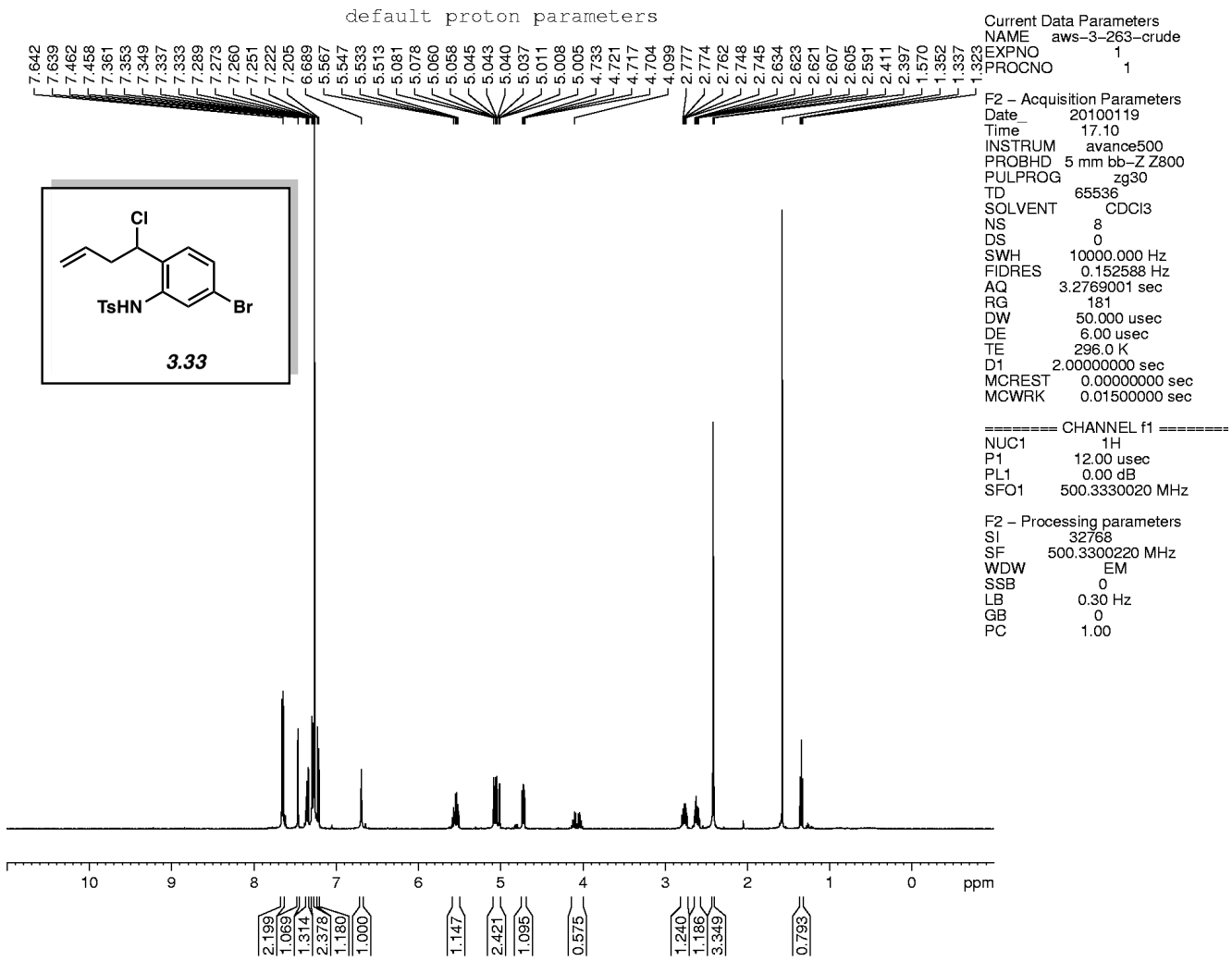


Figure A2.34 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.33**.

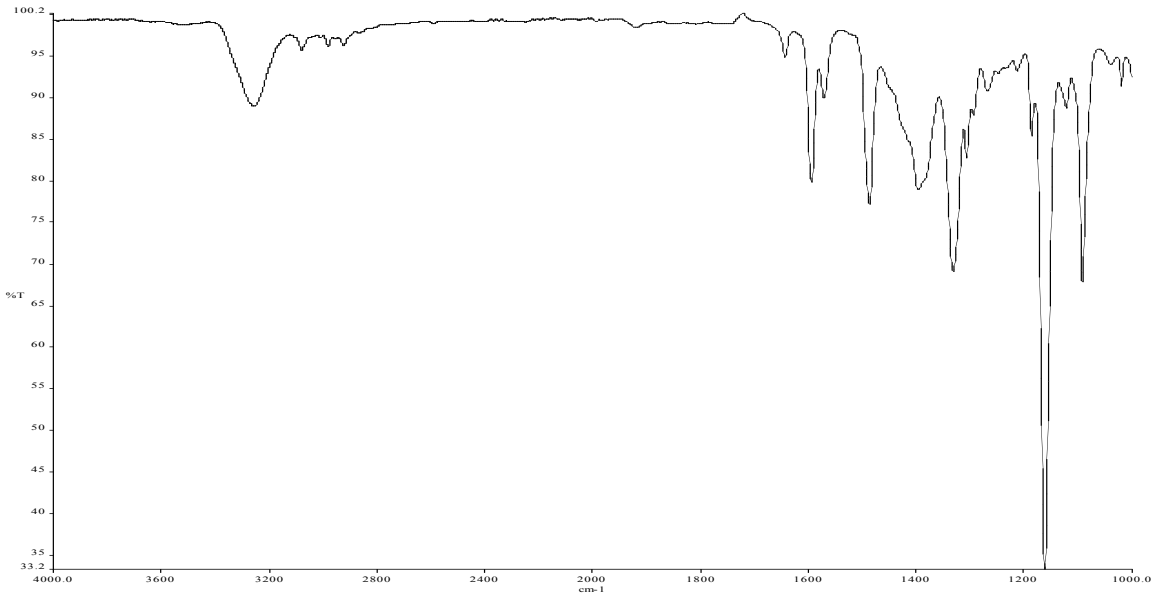


Figure A2.35 Infrared spectrum of compound **3.33**.

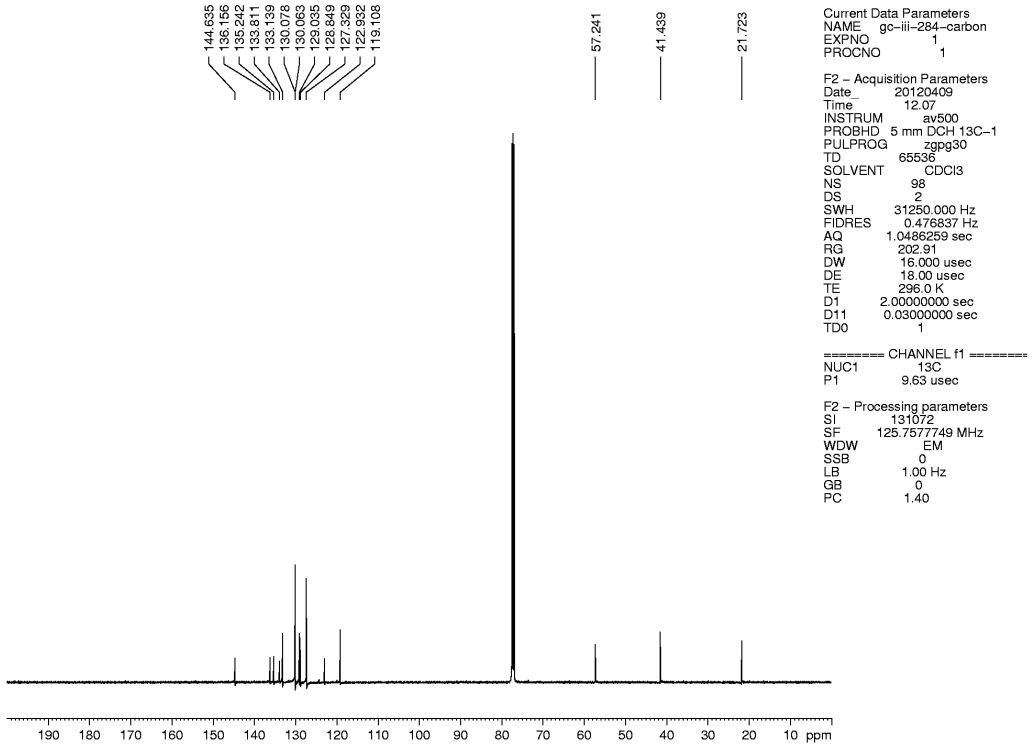


Figure A2.36 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.33**.

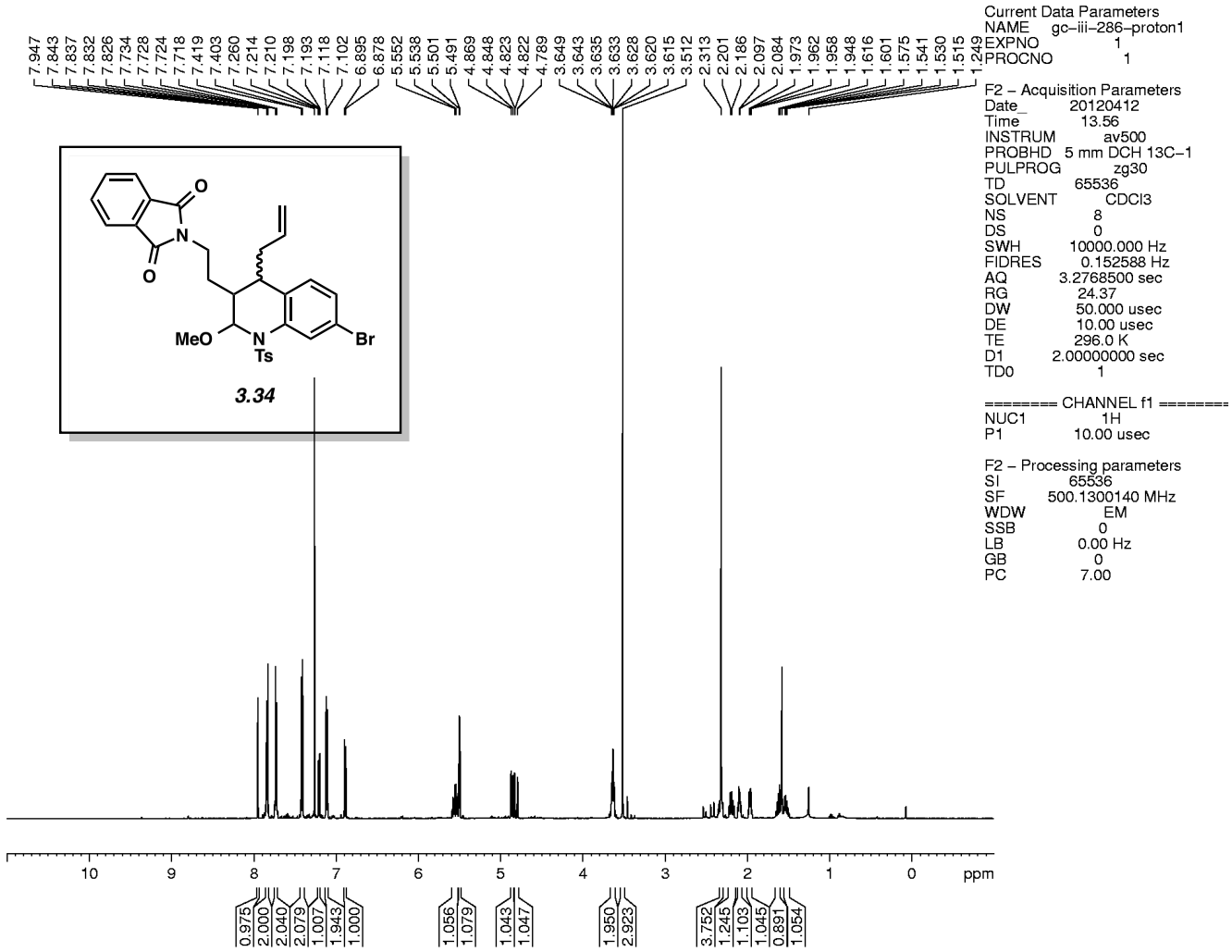


Figure A2.37 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.34**.

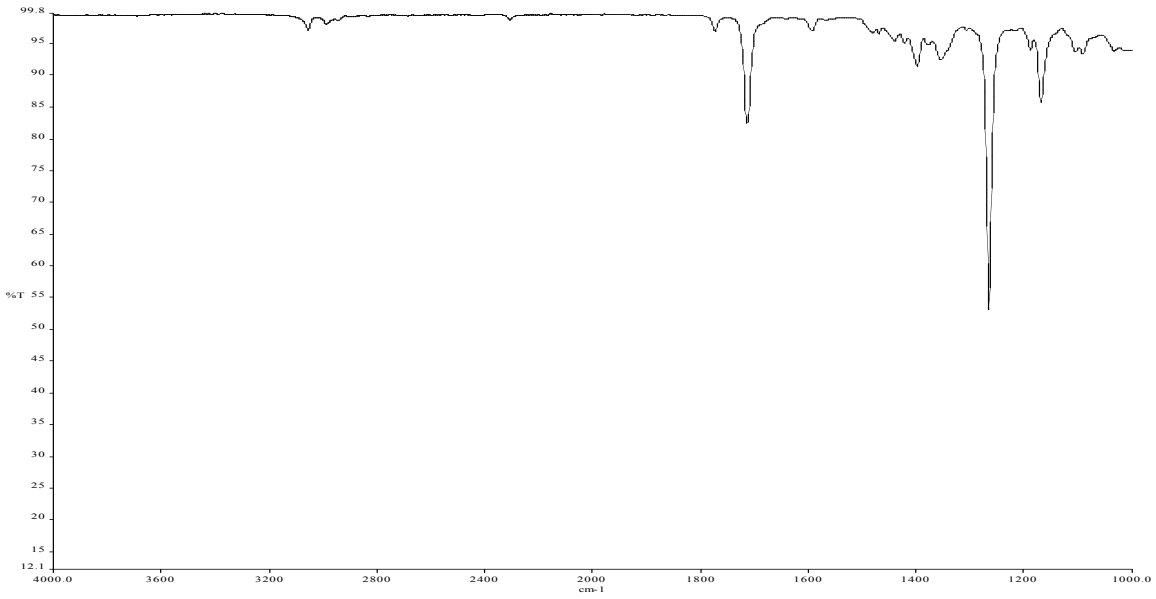


Figure A2.38 Infrared spectrum of compound **3.34**.

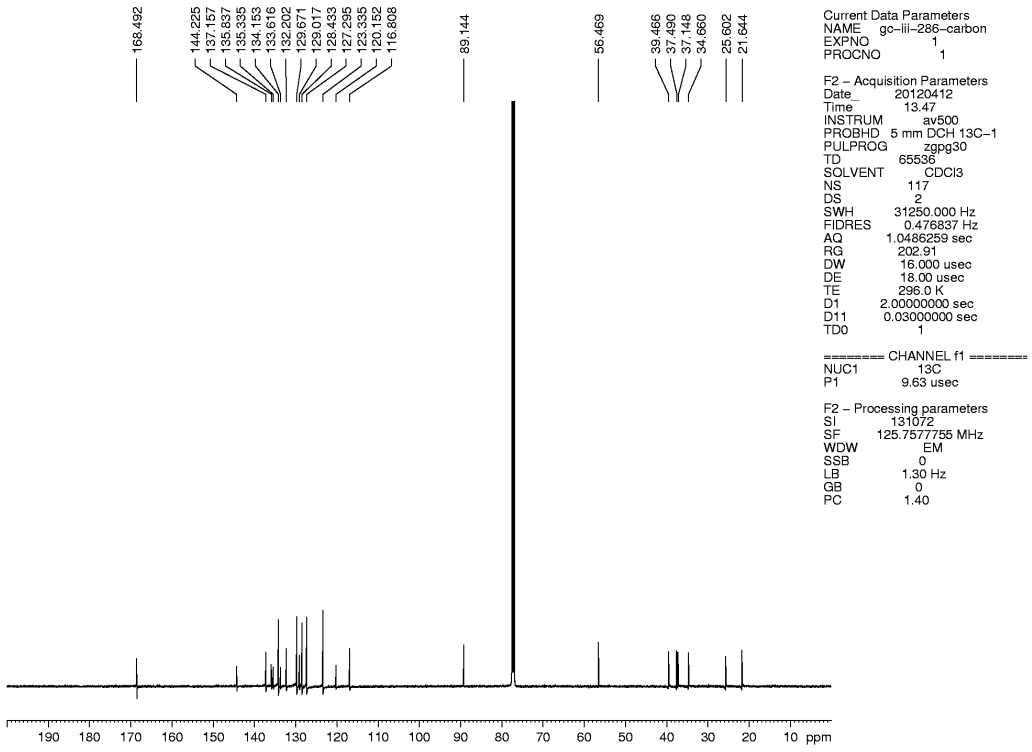


Figure A2.39 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.34**.

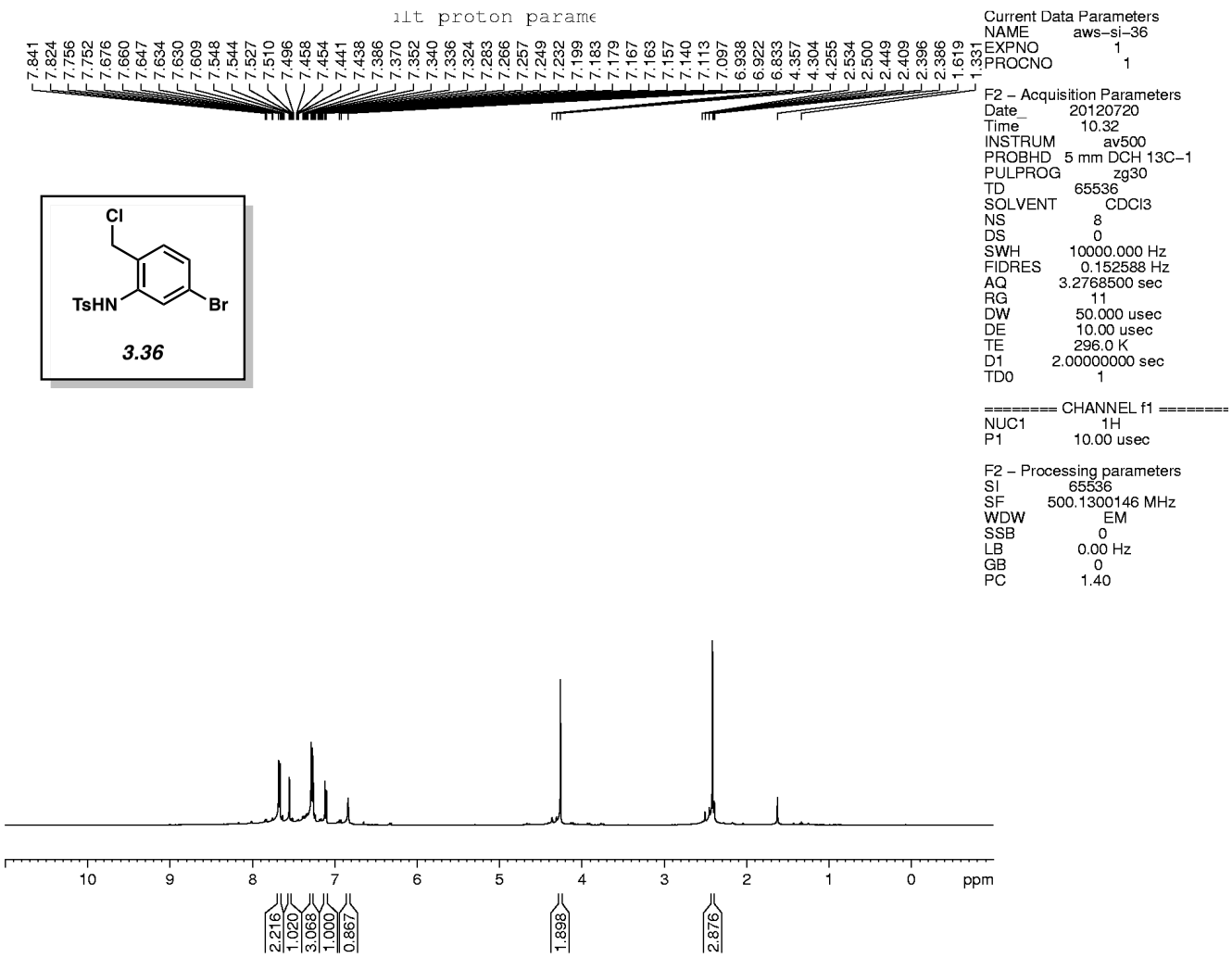


Figure A2.40 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.36**.

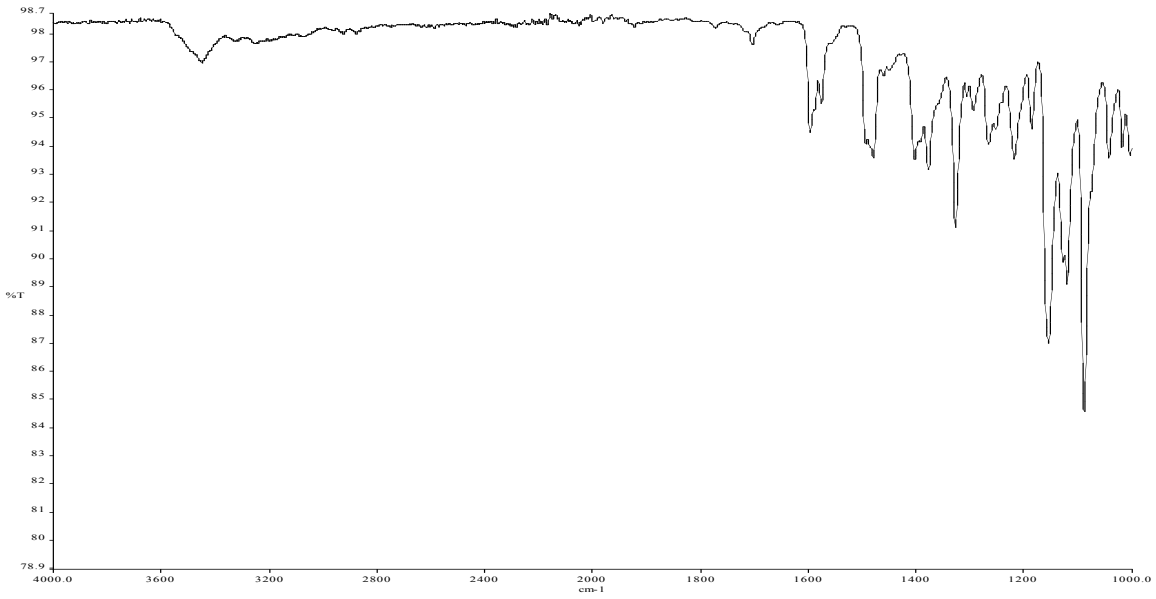


Figure A2.41 Infrared spectrum of compound **3.36**.

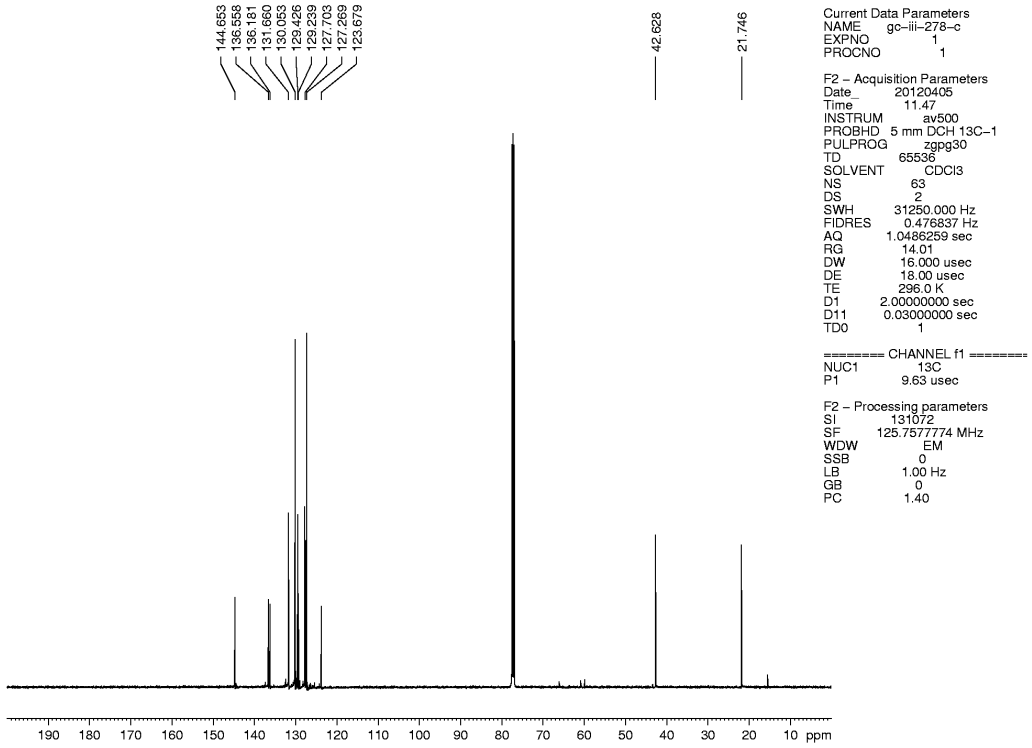


Figure A2.42 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.36**.

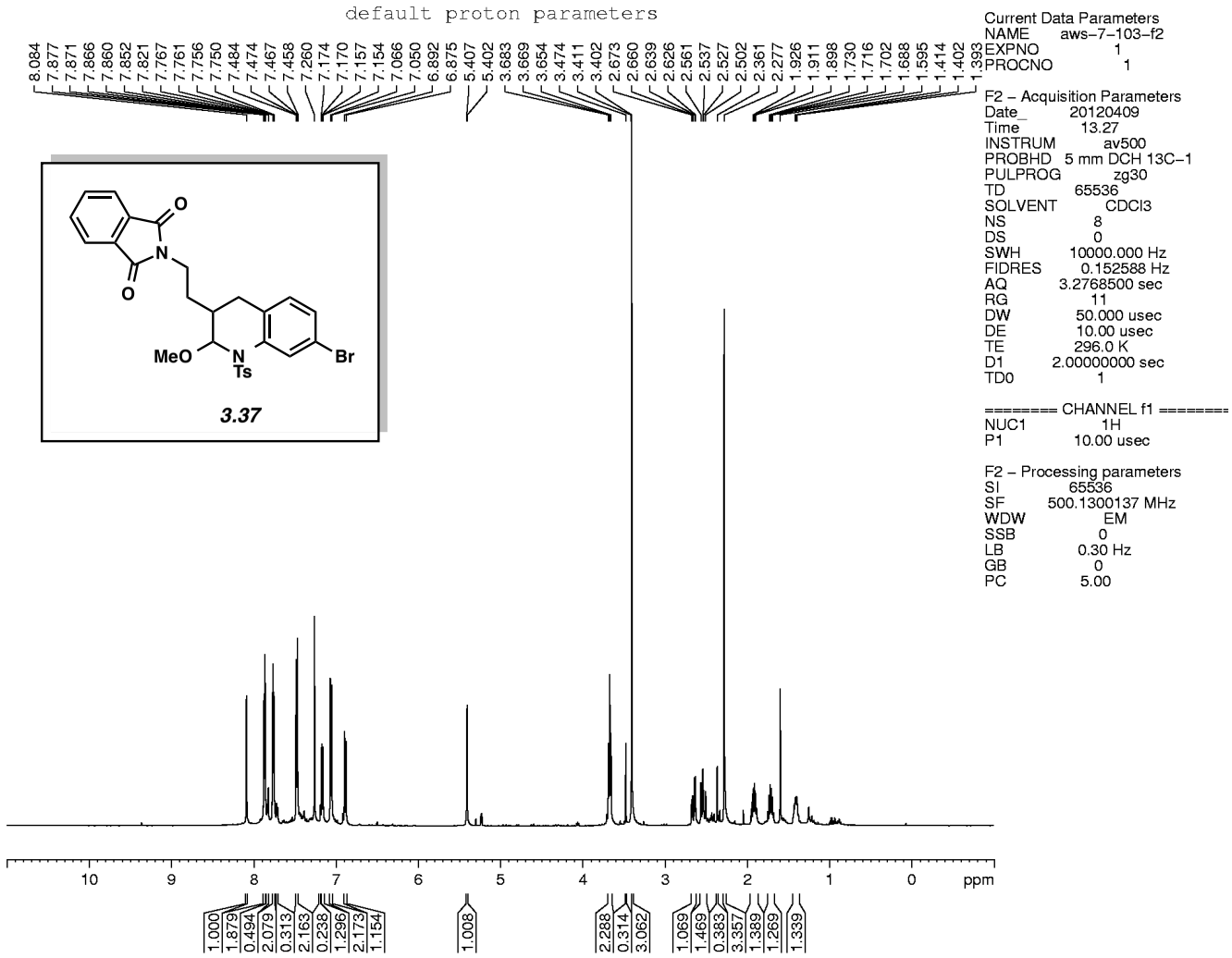


Figure A2.43 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.37**.

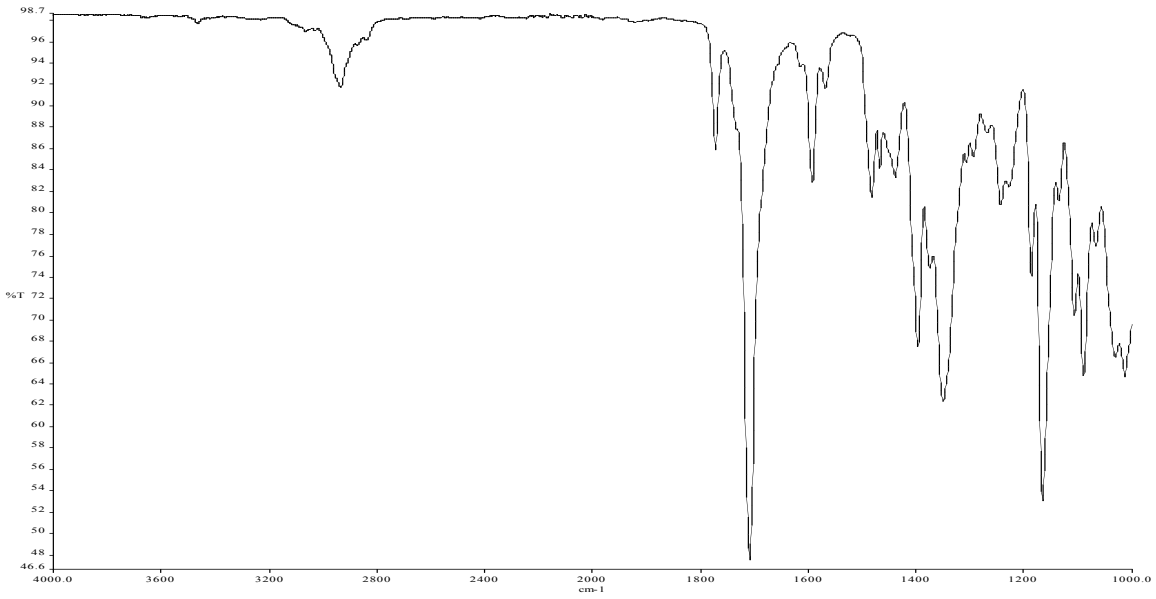


Figure A2.44 Infrared spectrum of compound **3.37**.

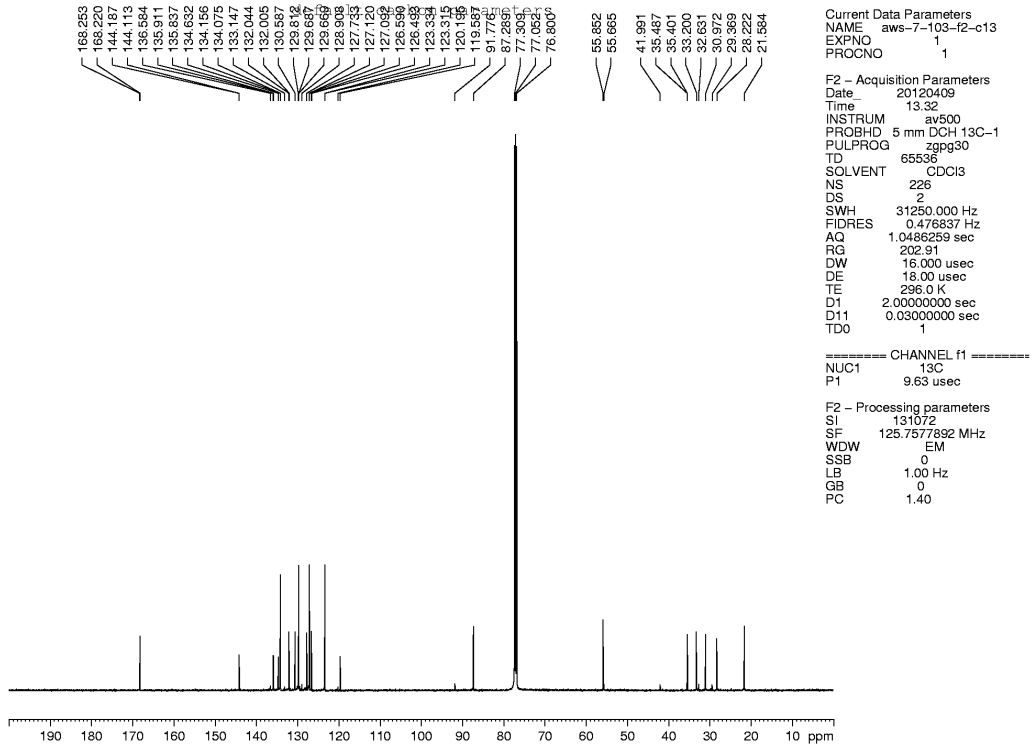


Figure A2.45 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.37**.

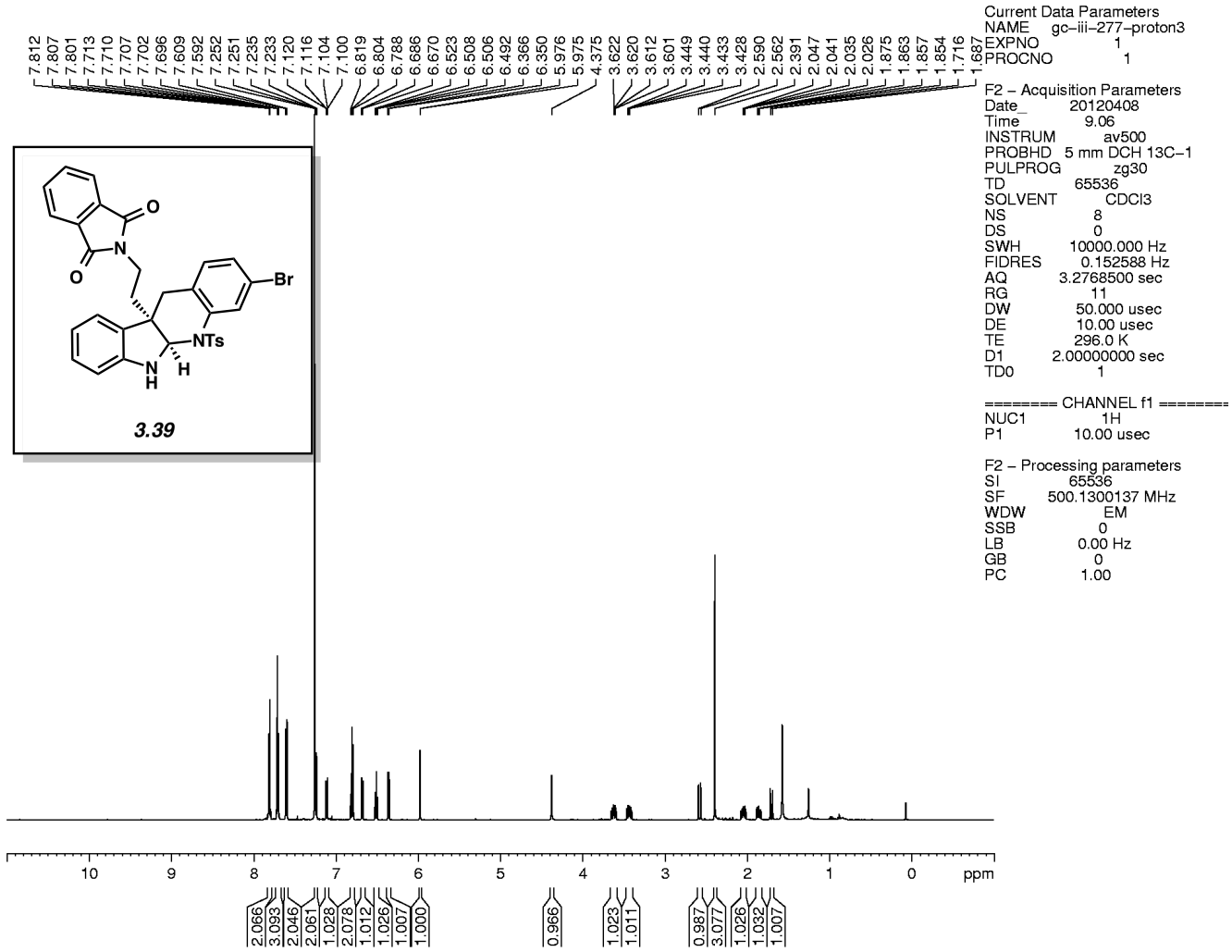


Figure A2.46 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.39**.

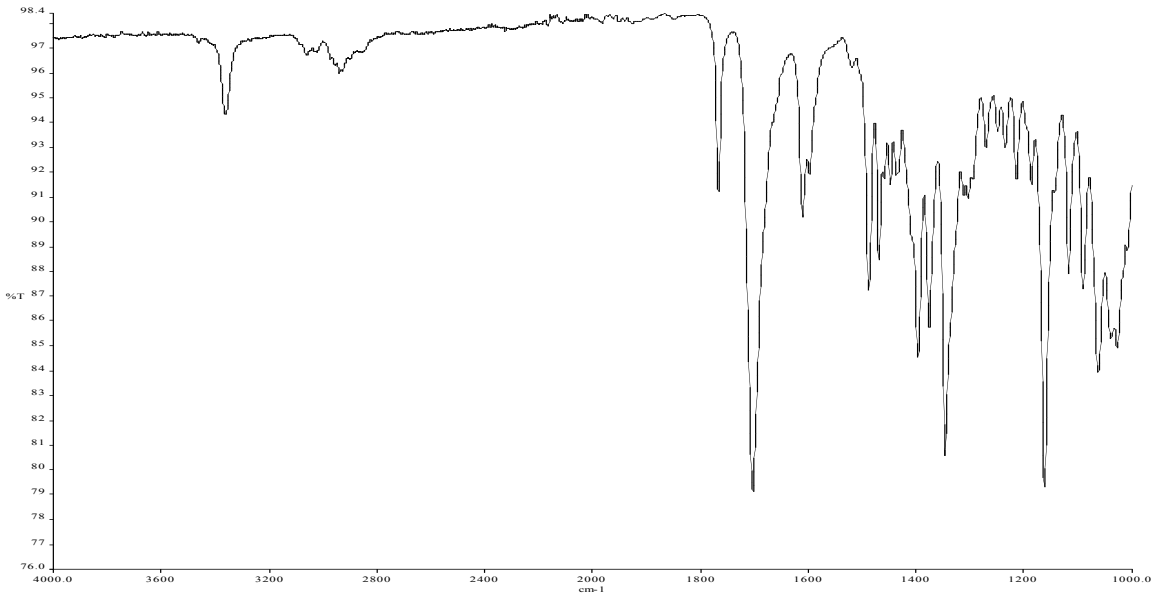


Figure A2.47 Infrared spectrum of compound **3.39**.

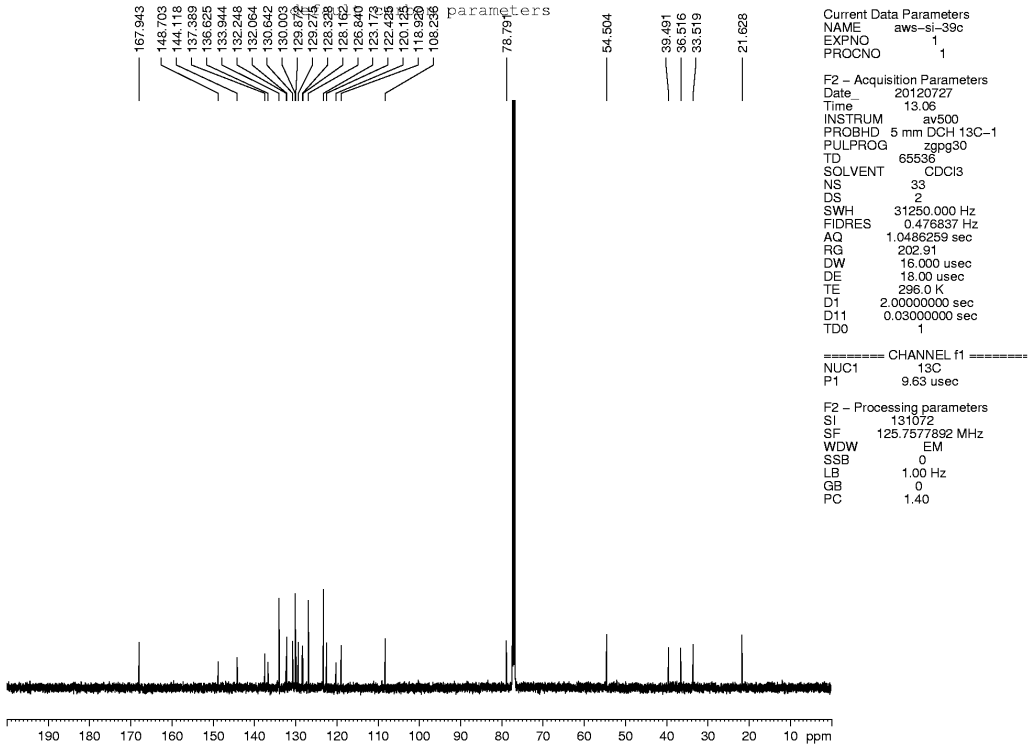


Figure A2.48 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.39**.

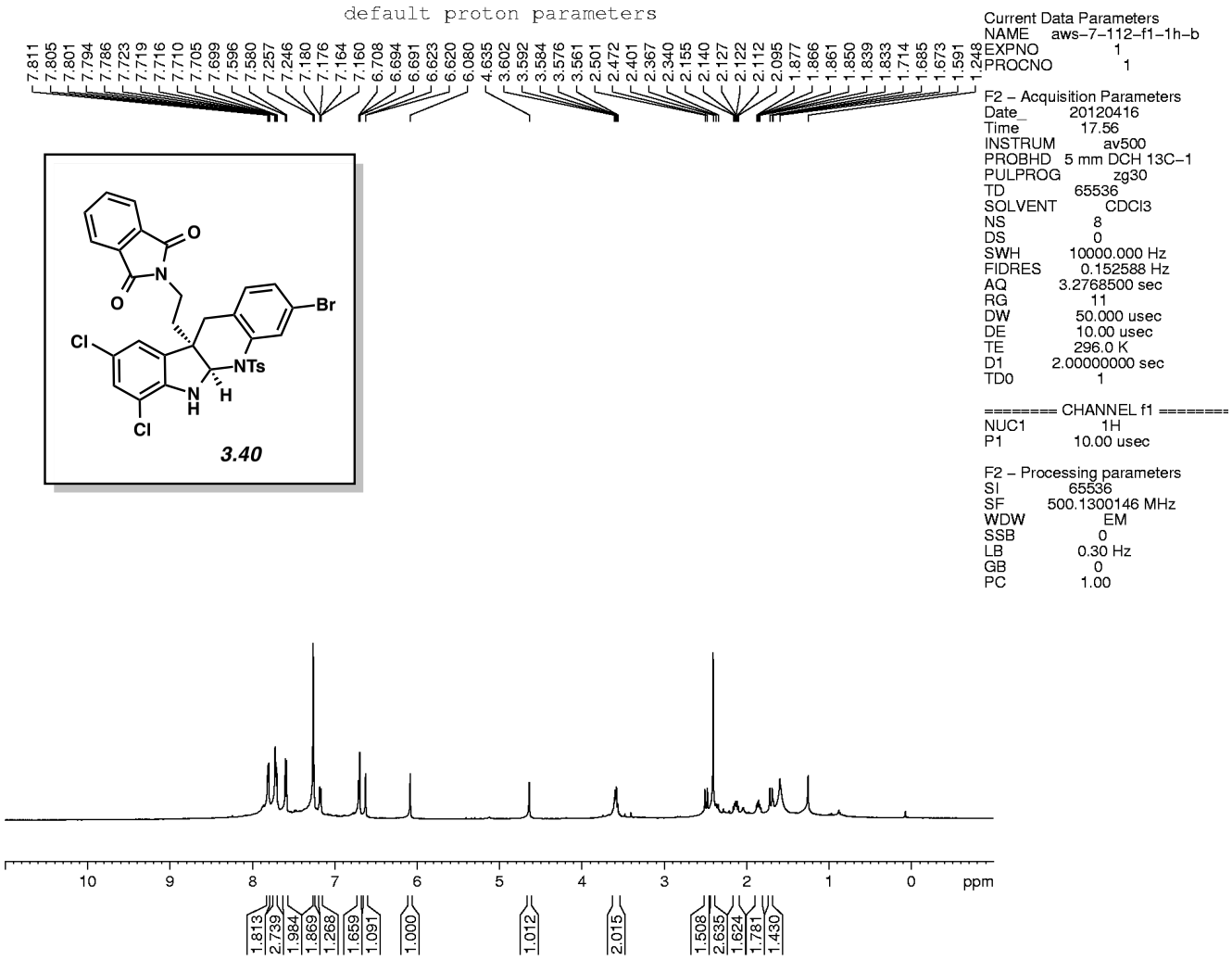


Figure A2.49 ^{1}H NMR (500 MHz,$CHCl_3$) of compound **3.40**.

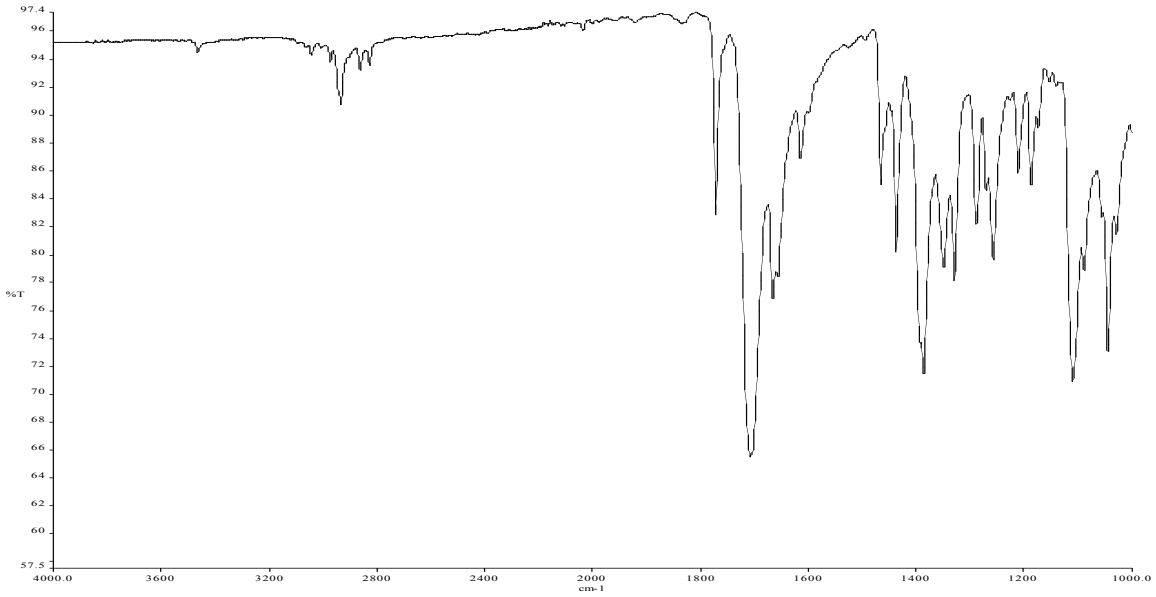


Figure A2.50 Infrared spectrum of compound **3.40**.

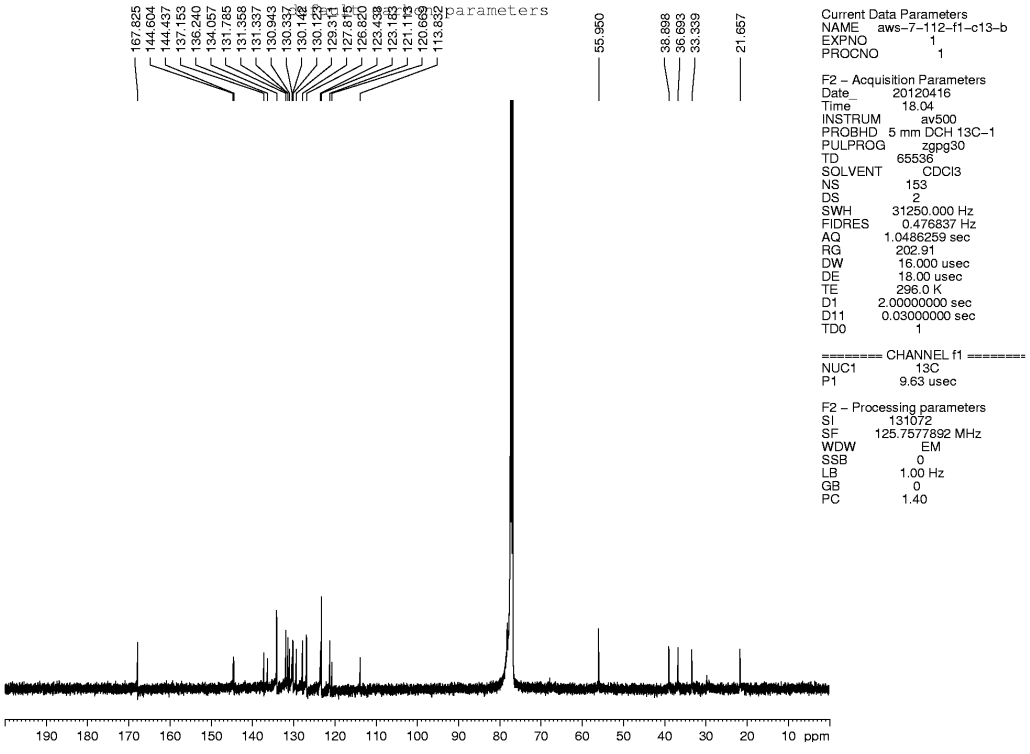


Figure A2.51 ^{13}C NMR (125 MHz, $CHCl_3$) of compound **3.40**.

CHAPTER FOUR

EcdGHK are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold

Wei Jiang, Ralph A. Cacho, Grace Chiou, Neil K. Garg, Yi Tang, Christopher T. Walsh.

J. Am. Chem. Soc. **2013**, *135*, 4457–4466.

4.1 Abstract

The echinocandins (**4.1**–**4.3**, Figure 4.1A) are a small group of fungal *N*-acylated cyclic hexapeptides that are fungicidal for candida strains and fungistatic for aspergilli by targeting cell wall 1,3-β-glucan synthases. The side chains of all six amino acid building blocks have hydroxyl groups, including the nonproteinogenic 4*R*,5*R*-dihydroxy-Orn_1, 4*R*-OH-Pro_3, 3*S*,4*S*-dihydroxy-$homoTyr_4$ and 3*S*-OH-4*S*-Me-Pro_6 (Figure 4.1A). The echinocandin (*ecd*) gene cluster contains two predicted nonheme mononuclear iron oxygenase genes (*ecdG,K*) and one encoding a P450 type heme protein (*ecdH*). Deletion of the *ecdH* gene in the producing *Emericella rugulosa* generates an echinocandin scaffold (echinocandin D (**4.3**)) lacking both hydroxyl groups on Orn_1. Correspondingly, the *ΔecdG* strain failed to hydroxylate C_3 of the homoTyr residue, and purified EcdG hydroxylated free L-homoTyr at C_3. The *ΔecdK* strain failed to generate mature echinocandin unless supplemented with either 4-Me-Pro or 3-OH-4-Me-Pro, indicating blockage of a step upstream of Me-Pro formation. Purified EcdK is a Leu 5-hydroxylase, acting iteratively at C_5 to yield γ-Me-Glu-γ-semialdehyde in equilibrium with the cyclic imine product. Evaluation of deshydroxyechinocandin scaffolds in *in vitro* anticandidal assays revealed up to 5-

fold loss of potency for the *ΔecdG* scaffolds, but a 3-fold gain of potency for the *ΔecdH* scaffold, in line with prior results on deoxyechinocandin homologs.

4.2 Introduction

The echinocandins (**4.1**–**4.3**) are a small group of fungal lipopeptides that act to kill other fungi by targeting the 1,3-glucan synthase that builds the 1,3-glucans that are key structural components of fungal cell walls.[1,2] Three semisynthetic versions (with the natural fatty acyl chains replaced by synthetic hydrophobic groups) of the naturally occurring pneumocandin or echinocandin scaffolds have been commercialized and approved for human clinical use as fungicidal and fungistatic agents for the treatment of candidiasis and aspergillosis, respectively.[3,4,5]

For naturally occurring echinocandin B (**4.1**, Figure 4.1A) the lipid is the di-olefinic linolenic acid. It has an amide linkage to a modified L-ornithine residue that is part of a cyclic hexapeptide framework (Figure 4.1). Four of the six residues in the hexapeptide framework are nonproteinogenic: only Thr_2 and Thr_5 are proteinogenic amino acids. The 4*R*,5*R*-dihydroxy-L-Orn_1, 4*S*-OH-Pro_3, 3*S*,4*S*-dihydroxy-homotyrosine$_4$, and 3*S*-OH-4*S*-methyl-Proline$_6$ (3-OH-MePro) residues are typically not found in proteins (although collagen has hydroxyproline residues introduced by posttranslational hydroxylation). Echinocandin C (**4.2**) has the C_3-OH, but lacks the C_4-OH on the homoTyr residue, while echinocandin D (**4.3**) is a trideoxy form that has the 3-OH-homotyrosine$_4$ (3-OH-homoTyr_4) but, a nonhydroxylated Orn_1.[6]

The unusual building block inventory suggests that echinocandins and the closely related pneumocandin are constructed by fungal nonribosomal peptide synthetase (NRPS) assembly lines. This expectation has been recently validated by the identification of a six-module NRPS

biosynthetic gene *ecdA* in *A. nidulans var roseus* (recently renamed *Emericella rugulosa*[7]), whose disruption abolishes production of **4.1**.[8] Among the coding sequences in the vicinity of the *ecdA* gene are *ecdG, ecdH, ecdK,* two of which are predicted to be α-ketoglutarate-dependent mononuclear, nonheme iron oxygenases genes (*ecdG*, *ecdK*), with the third predicted to be a P450 type of hemeprotein monooxygenase gene (*ecdH*) (Figure 4.1B). Given that all six of the echinocandin B amino acid residues have at least one side chain hydroxyl group (Orn_1 and $homoTyr_4$ each have two such –OH groups), it is likely that these oxygenases may fulfill tailoring functions at some stage to create the hydrophilic surface of the mature echinocandin B (**5.1**) antifungal agent. In addition, it is likely that the Me-Pro framework at residue 6 arises from the proteinogenic amino acid building block L-leucine by hydroxylation at C_5,[9] suggesting another, albeit cryptic, step for hydroxylase action in echinocandin producers.

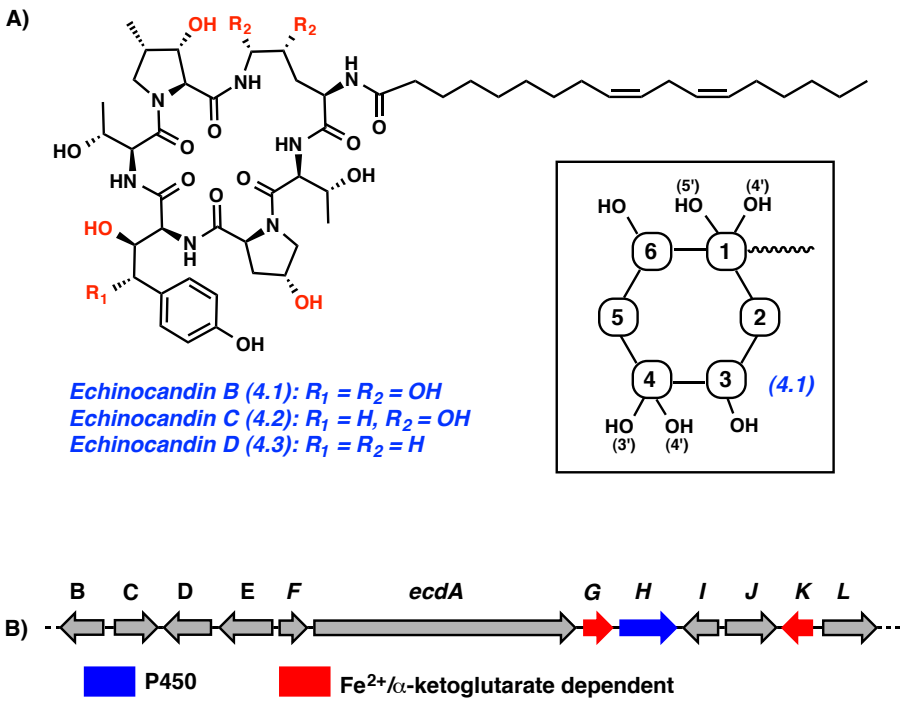


Figure 4.1: A) Structure of echinocandin B (**4.1**), C (**4.2**), and D (**4.3**). A schematic representation of **4.1** used in Figures 4.2 and 4.3 is shown in the inset. B) Proposed nonheme mononuclear iron oxygenase and P450 type heme genes in *ecd* and *hty* gene clusters.

To address the nature of the involvement of *ecdG*, *ecdH*, and *ecdK* in the maturation of the lipopeptide scaffold of echinocandin B (**4.1**), we report the gene disruptions in each of these three genes and analysis by mass spectrometry and NMR of the des-hydroxy forms of echinocandin that accumulate in the media in such fermentations. Studies with purified EcdG and EcdK confirm regiospecific oxygenase action on L-homoTyr and L-Leu, respectively. The results allow assessment of the specificity of each of the three oxygenases and the consequences for biological activity (antifungal action) as the nascent echinocandin scaffold is progressively matured.

4.3 Echinocandin Production from wild type *E. rugulosa*

In preparation for analysis of the echinocandin forms produced in gene deletion mutants, we first examined the forms of echinocandin generated by the native *E. rugulosa* producer under echinocandin production conditions.[10] In addition to the fully hydroxylated **4.1**, (*m/z* calcd for $C_{52}H_{81}N_7O_{16}$ $[M+H]^+[-H_2O]$ 1042.5707, found 1042.5700), there is a monodeoxy form **4.2** (=echinocandin C) noted previously by earlier workers[6] that corresponds to the lack of a hydroxyl group at the 4-position of the homoTyr residue (*m/z* calcd for $C_{52}H_{81}N_7O_{15}$ $[M+H]^+[-H_2O]$ 1026.5758, found 1026.5757) (Figure 4.2A). Presumably, conversion of homoTyr to 3,4-dihydroxyhomoTyr, either as the free monomer or on a hexapeptide intermediate, is incomplete and either occurs late in maturation or does not prevent hydroxylation at other sites during maturation. The presence of **4.2** along with **4.1** suggests hydroxylation at the 3-position of $homoTyr_4$ can occur before hydroxylation at C^4.

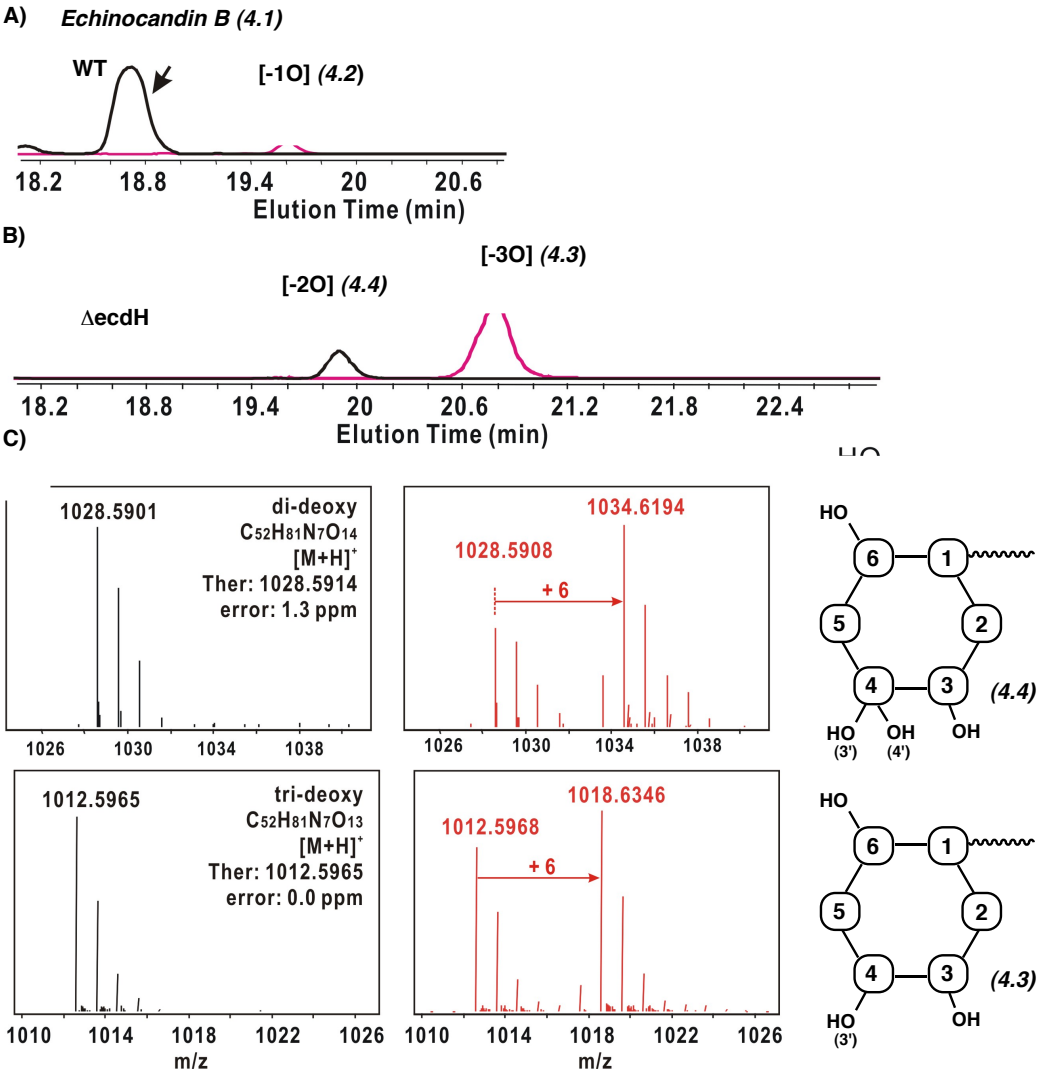


Figure 4.2*:* A) Extracted ion chromatogram of echinocandin B (**4.1**) and C (**4.2**) from *Emericella rugulosa* profile. B) Extracted ion chromatogram of echinocandin mutants from Δ*ecdH* profile. C) MS of parent ions from echinocandin mutants produced by Δ*ecdH* (**4.3** & **4.4**), either from regular media (black) or fed with D_6-L-ornithine (red). The proposed structure of each compound is on the right.

4.4 Echinocandin Product Profile from a *ΔecdH Emericella rugulosa* strain

Construction and characterization of the *ΔecdH* strain is shown and is representative of the approaches used also for the *ΔecdG* and *ΔecdK* genes. Deletion mutants in each of the *ecdGHK* genes were screened for insertion of *bar* knockout cassettes followed by PCR validation of *bar* gene insertion into the desired *ecd* gene.[11,12] Mutant cultures were grown under echinocandin production conditions,[10] compounds were extracted and evaluated by MS and NMR.

The *ΔecdH* strain loses production of **4.1** and **4.2** while yielding dideoxy (*m*/*z* calcd for $C_{52}H_{81}N_7O_{14}$ $[M+H]^+$ 1028.5914; found 1028.5901) and trideoxy forms (*m*/*z* calcd for $C_{52}H_{81}N_7O_{13}$ $[M+H]^+$ 1012.5965; found 1012.5965) of **4.1** (Figure 4.2B). The trideoxy compound is the major component with yields of ~10 mg from 8 L culture. NMR characterization reveals the loss of hydroxyl group at 4-homoTyr$_4$ and both of the 4- and 5-OH of Orn$_1$ (Table S1).[29] This is identical to echinocandin D, compound **4.3**, which is one of the naturally occurring minor echinocandins (Table S2).[6,13,29]

The yield of dideoxy compound **4.4** from *ΔecdH* was too low for NMR assignment. We thus carried out feeding studies with deuterated amino acid monomers to evaluate the hydroxylation state at certain residues. 10 mg of 3,3-,4,4-5,5-perdeutero (D_6)-L-ornithine was used in 10 mL culture to assess the hydroxylation of the Orn$_1$ residue. **4.1** and **4.2** extracted from the WT strain exhibits an M+4 parent ion (Figure S2),[29] consistent with loss of a deuterium at each of C_4 and C_5 as 4-OH, 5-OH-L-Orn is incorporated in the antibiotic framework. Variants **4.3** and **4.4** isolated from *ΔecdH* strain fed with (D_6)-L-Orn, on the other hand, both had prominent peaks at M+6, indicating that none of the Orn side chain deuterium atoms had been

lost (Figure 4.2C). Therefore, compound **4.4** must retain all of the hydroxylations except C_4 and C_5 at Orn_1.

We also tested feeding with perdeuteroproline (D_7) or perdeuteroTyr (D_7) as a parallel control. While no incorporation was observed in rich medium, in minimal medium, echinocandins were still produced and a low but detectable level of (D_6)-L- proline incorporation (loss of one deuterium at C_4 during conversion to 4-OH-Pro) was obtained in **4.1** from the native producer (Figure S3).[29] Compound **4.3** from the *ΔecdH* strain exhibited the same M+6 parent ion (Figure S3),[29] indicating no structural difference between them on the Pro_4 residue. On the other hand, there was still no detectable incorporation of perdeutero-Tyr. Proline is not expected to be a precursor to 3-OH-4-$MePro_6$ since precedent suggests a route from leucine,[14] and accordingly no incorporation of deuterium from (D_6)-L-proline is found at that residue.

Thus, one can unambiguously assign the hemeprotein P450 EcdH as the Orn oxygenase acting to hydroxylate two adjacent carbons. The *ecdH* gene knockout studies are silent on two aspects of the timing of C_4 vs C_5 hydroxylation of Orn_1: (a) whether the free amino acid undergoes hydroxylation or only the Orn residue after incorporation into the nascent peptide is oxygenated; and (b) whether C_4 or C_5 is oxygenated first. Efforts to isolate soluble EcdH from *E. coli* expression were unsuccessful (overproduction yielded insoluble inactive protein). Because EcdH may be membrane-associated, expression in yeast microsomes was conducted. Neither free L-Orn nor the trideoxy form **4.3** of echinocandin was detectably oxygenated on incubation with those microsomal preparations, so no conclusion can be drawn about point (a).

However, given that C_5 hydroxylation of Orn_1 in mature **4.1** is a key feature of the hemiaminal linkage, one expects this would not have happened before Orn was incorporated into the echinocandin peptide backbone. Otherwise a free C_5-OH-L-Orn tetrahedral addition

compound would unravel to the aldehyde and in that oxidation state will not be able to form the amide bond with 3-OH-MePro$_6$. The relative timing of C_5 vs C_4 hydroxylation can be intuited in the *ΔecdG* studies described below, from the trideoxy echinocandin **4.7** NMR where the C_5 of Orn has been hydroxylated but C_4 has not. In aggregate the results suggest EcdH acts on Orn only after incorporation into the nascent echinocandin chain and acts first at C_5, then at C_4. There are precedents for a single P450 enzyme acting iteratively on its substrate. Notably, the P450 side chain cleavage enzyme in the adrenal steroid hormone pathway converts cholesterol to 22*R*-OH-cholesterol, then to the 20*R*, 22*R*-diol, before a third P450$_{scc}$ oxygenative step cleaves between the diol to yield pregnenolone and isocaproic aldehyde.[15]

4.5 Echinocandin Product Profile from a *ΔecdG Emericella rugulosa* strain

The *ΔecdG* strain (Figure S4)[29] gave a more complex echinocandin metabolite profile, producing a mixture of mono- (**4.5**), di- (**4.6**), tri- (**4.7**) and tetra-deoxy (**4.8**) compounds, with **4.6** and **4.7** as the major components (Figure 4.3A). Both of them were successfully isolated (4.0 mg of **4.6** and 5.5 mg of **4.7** from 8 L culture) and characterized by NMR. The dideoxy form **4.6** (*m/z* calcd for $C_{52}H_{81}N_7O_{14}$ $[M+H]^+$ -$[H_2O]$ 1010.5809; found 1010.5819) is missing both hydroxyl groups on 3- and 4-homoTyr$_4$ (Table S1).[29] The trideoxy **4.7** (*m/z* calcd for $C_{52}H_{81}N_7O_{13}$ $[M+H]^+$ -$[H_2O]$ 994.5859; found 994.5826), in addition to missing two -OH groups on homoTyr$_4$, also is lacking an -OH at C_4 of Orn$_1$ (Table S1).[29] Consistent with NMR analysis, feeding with (D_6)-L-Orn to the *ΔecdG* strain led to M+4 and M+5 shifts in compound **4.6** and **4.7**, respectively (Figure 4.3B).

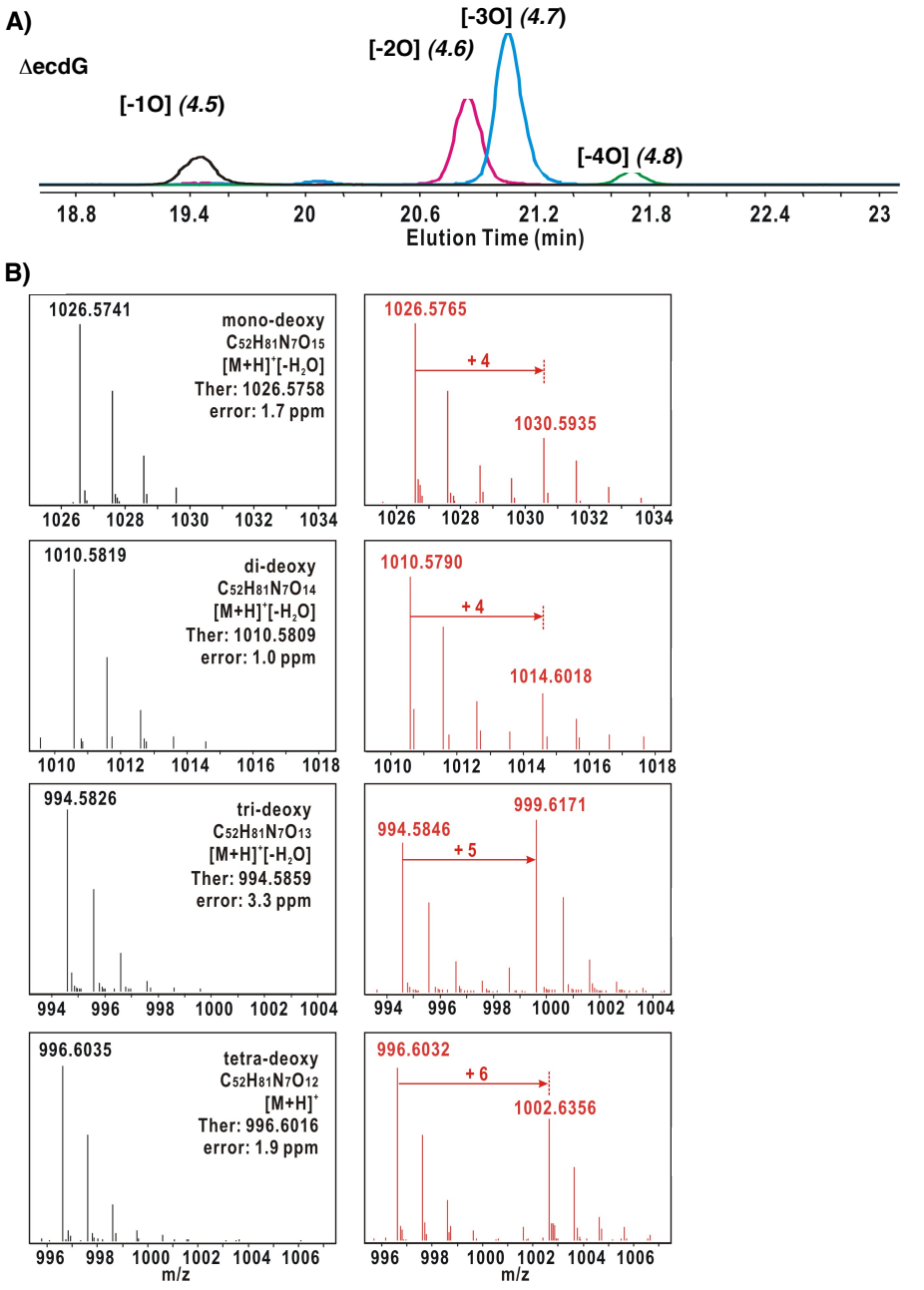


Figure 4.3: A) Extracted ion chromatogram of echinocandin mutants from Δ*ecdG* profile. B) MS of parent ions from echinocandin mutants produced by Δ*ecdG* (**4.5** - **4.8**), either from regular media (black) or fed with D_6-L-ornithine (red). The proposed structure of each compound is on the right.

The yield of **4.8** (*m/z* calcd for $C_{52}H_{81}N_7O_{12}$ $[M+H]^+$ 996.6016; found 996.6035) was too low for NMR analysis; however, its structure can be deduced from feeding results. It exhibits an M+6 peak upon (D_6)-L-Orn addition, indicating both hydroxyl groups on Orn_1 are absent. Together with the confirmed structure of **4.6** and **4.7**, it is reasonable to propose that Orn_1 and $homoTyr_4$ side chains in the tetradeoxy compound **4.8** are completely unmodified. Consistent with this scheme, **4.8** shows an M+6 pattern upon (D_7)-L-proline feeding (Figure S3),[29] indicating hydroxylation on the proline moiety. As expected, **4.7** present the same M+6 pattern on proline as well (the incorporation level in **4.5** and **4.6** is too low to be identified).

The presence of all three variations of the Orn_1 moiety (des-hydroxy in **4.8**, monohydroxy in **4.7**, dihydroxy in **4.5** and **4.6**) suggests that the hydroxylation of Orn does not *require* EcdG but is *affected* upon the absence of EcdG. This could reflect altered conformation of the hexapeptide macrocycle when particular-OH groups are missing. Therefore, the target of EcdG, as suggested by missing hydroxyl groups on homoTyr in **4.6** and **4.7**, should be homoTyr. Because only one hydroxyl group is missing in compound **4.5**, we expect it to be the specific hydroxyl group (either 3OH- or 4OH-$homoTyr_4$) installed by EcdG. Resolving the structure of **4.5** would be one key to this question, however, the isolated quantity was too low for detailed NMR analysis. Since 4-OH on $homoTyr_4$ is not ubiquitous in the echinocandins obtained from WT and gene deletion mutants, we favored EcdG as a homotyrosine 3-hydroxylase, and the *in vitro* assay described next confirmed it.

4.6 Purified EcdG is a Homotyrosine 3-Hydrolase

The mononuclear nonheme iron enzyme EcdG could be expressed and purified in soluble form from *E. coli* (Figure 4.4A) with a yield of ~30 mg/L. Assay with the amino acid building block L-homoTyr (Figure 4.4B) followed by Fmoc derivatization revealed that EcdG acts to

introduce one hydroxyl group on free L-homoTyr (Fmoc-homoTyr: *m/z* calcd for $C_{25}H_{23}NO_5$ $[M+Na]^+$ 440.1468, found 440.1460; Fmoc-hydroxy-homoTyr: *m/z* calcd for $C_{25}H_{23}NO_6$ $[M+Na]^+$ 456.1418, found 456.1411). As has been seen with other nonheme iron oxygenases that generate high valent oxoiron intermediate,[16,17] EcdG inactivates itself after a small number of turnovers. Variation of the amount of pure EcdG relative to excess substrate indicates a proportional amount of product formation (Figure 4.4C) and indicates that about 3-3.5 turnovers occur per inactivation event. The estimated initial rate of hydroxylation is higher than 5 catalytic events per minute.

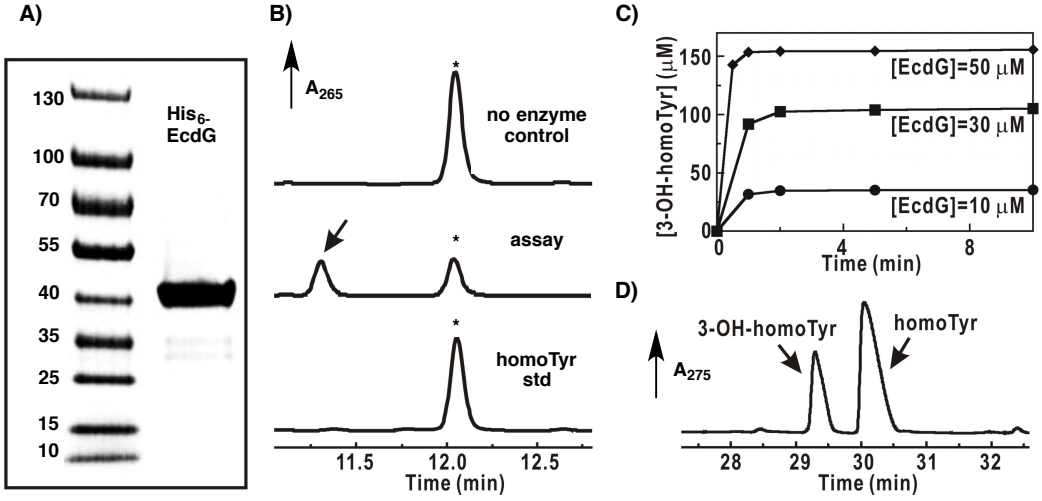


Figure 4.4: A) SDS-PAGE gel showing purified His_6-EcdG enzyme (expected molecular weight: 40.0 kD). B) HPLC-UV (259 nm) traces indicate formation of hydroxy-homoTyr (arrow, middle trace) relative to negative control without enzyme (top) and homoTyrosine standard (star, bottom trace) as Fmoc-derivatives. C) Time course of formation of hydroxy-homoTyr at different enzyme concentrations. D) HPLC-UV (275 nm) trace showing isolation of hydroxyl-homoTyr from EcdG assay solutions for further NMR analysis.

Although only 3-3.5 oxygenation events per EcdG enzyme molecule occur before enzyme suicide, it was possible to scale up the monohydroxy homoTyr product and separate it from unreacted L-homoTyr by Hypercarb column using 20 mM perfluoropentanoic acid / acetonitrile as mobile phase (Figure 4.4D, homoTyr: *m/z* calcd for $C_{10}H_{13}NO_3$ $[M+H]^+$ 196.0968, found 196.0975; hydroxy-homoTyr: *m/z* calcd for $C_{10}H_{13}NO_4$ $[M+H]^+$ 212.0917, found 212.0925). Proton NMR and gCOSY clearly indicated the enzymatic product is a 3-OH-homoTyr (Figure S5 and S6)[29] and correlated with the chemical shifts reported for synthetic (2*S*, 3*R*)-OH-homoTyr.[18] We have not determined the absolute stereochemical configuration at the C3-OH of the EcdG *in vitro* product but anticipate it will be 3*R*.

Pure EcdG showed no activity with the deshydroxy echinocandin scaffold isolated from the *ΔecdG* mutant strain (**4.6** & **4.7**), confirming that it does not act after the NRPS assembly line has produced the hexapeptide framework. Moreover, EcdG will not take the 3-OH-homoTyr product and convert it on to the 3,4-dihydroxy amino acid (data not shown), making it unlikely to be the hydroxylase for introducing the 4-OH group in addition to the hydroxylation at C3. Therefore, the prediction is that the adenylation domain in module 4 of the six module EcdA would activate 3-OH-homoTyr (or 3,4-dihydroxy-homoTyr but less likely, see discussion below) for incorporation into the growing echinocandin chain.

4.7 Deletion of *ecdK* abolishes production of echinocandins

Given the above results from *ecdG* and *ecdH* deletion strains, we anticipated that the remaining putative oxygenase, the nonheme mononuclear iron protein EcdK, could act either on Pro_3 or 4R-$MePro_6$, or both, since these are the remaining two hydroxylated nonproteinogenic residues in **4.1**. It was also possible that EcdK could be involved in the biosynthesis of 4-MePro,

which in cyanobacteria is known to involve initial hydroxylation at C_5 of L-leucine.[14] When metabolites of *ΔecdK* mutants were analyzed by LCMS, indeed, production of **4.1** was abolished. Feeding experiments with homoTyr or 4-OH-Pro did not restore production of **4.1**. On the other hand, either 3*S*-OH-4R-MePro or 4R-MePro did restore robust echinocandin production (Figure 4.5A). These results clearly implicate EcdK in a step upstream of 4R-MePro formation. The first step in MePro construction is proposed in other systems to be hydroxylation of Leu to 5-OH-Leu (Figure 4.5B).[14] Further oxidation (either by a second hydroxylation at C_5 or via action of an alcohol oxidase) to γ-methyl-glutamic acid-γ-semialdehyde (γ-Me-Glu-γ-semialdehyde) would set up nonenzymatic cyclization to 3-methyl-D^1-pyrroline-5-carboxylic acid (MeP5C). Subsequent reduction of the 1-pyrroline by a nicotinamide-utilizing reductase would give 4-MePro (Figure 4.5B).

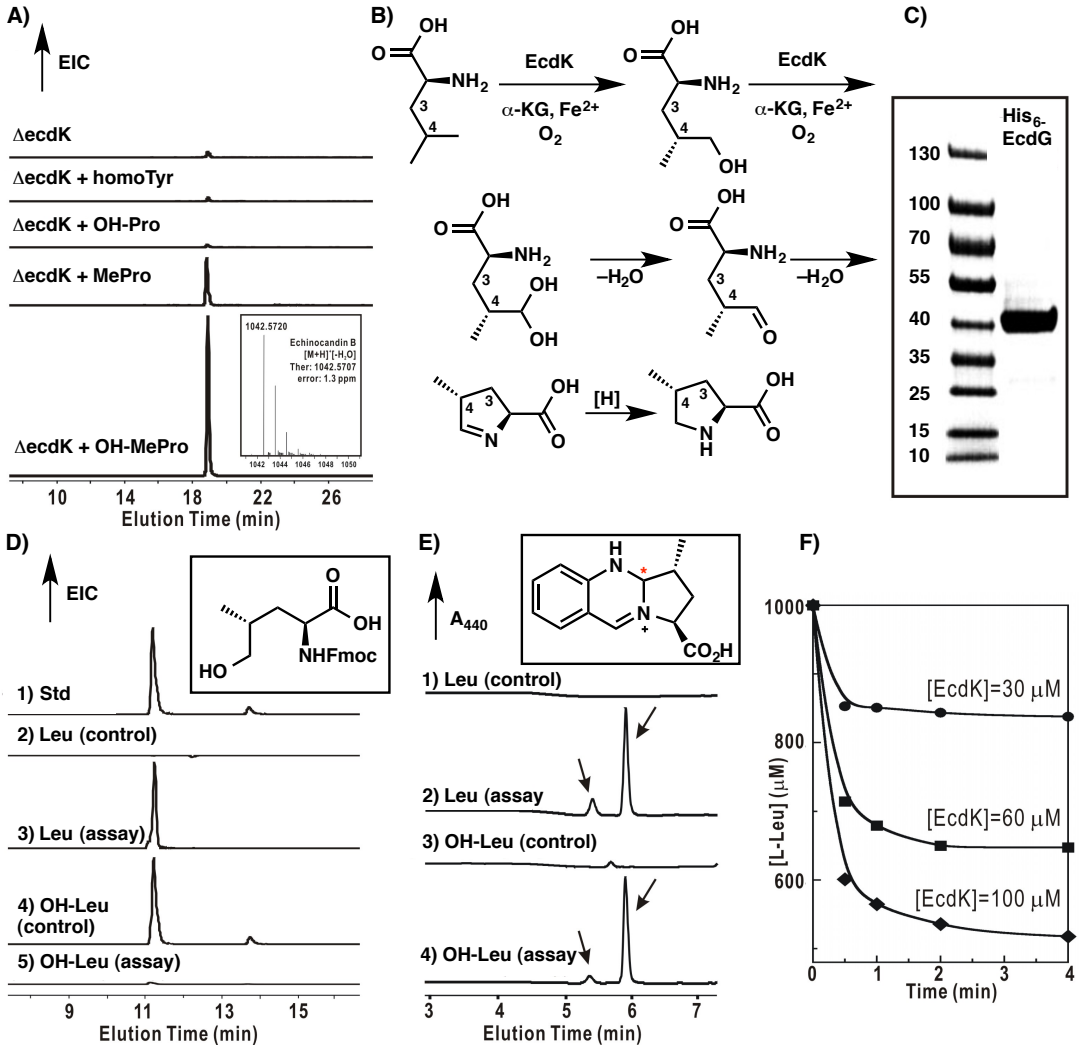


Figure 4.5: A) Extracted ion chromatogram of Echinocandin B (**4.1**) from Δ*ecdK* mutant. The culture media is supplemented with amino acid as labeled (OH-Pro = 4*R*-OH-L-Pro; MePro = 4*R*-Me-L-Pro; OH-MePro: 3*S*-OH-4*S*-Me-L-Pro). B) Proposed biosynthesis pathway for 4*R*-methyl-proline from L-Leu by tandem hydroxylation at C_5. C) SDS-PAGE gel showing purified His_6-EcdK enzyme (expected molecular weight: 39.7 kD). D) Extracted ion chromatogram of Fmoc-derivatized 5-OH-Leu (structure shown in inset). Top to bottom: 1) synthetic standard of Fmoc-5-OH-Leu; 2) assay with Leu, no enzyme; 3) assay with Leu, with enzyme; 4) assay with 5-OH-Leu, no enzyme; 5) assay with 5-OH-Leu, with enzyme, reflecting complete consumption

of 5-OH-Leu. E) HPLC-UV (440 nm) traces showing the formation of MeP5C after *o*-AB derivatization (derivatized product structure shown in inset). Arrows indicate a pair a diastereomers with a ratio of ~ 1:6. Top to bottom: 1) assay with Leu, no enzyme; 2) assay with Leu, with enzyme; 3) assay with 5-OH-Leu, no enzyme; 4) assay with 5-OH-Leu, with enzyme. (F) Time course of utilization of L-leucine at different EcdK concentrations. Residual Leu is quantified in its Fmoc-derivatized form by UV/Vis absorption.

4.8 EcdK oxygenates C_5 of L-Leucine iteratively.

In order to verify the results observed in the knockout of ecdK, we sought to reconstitute *in vitro* the oxidation of L-Leu to γ-Me-Glu-γ-semialdehyde by EcdK. EcdK could be overproduced in and purified from *E. coli* as an N-His-tagged enzyme (Figure 4.5C) as a soluble protein in a yield of 33 mg/L. Assay of pure EcdK with L-leucine followed by analysis of products revealed O_2-dependent conversion to two new products. One was identified as a hydroxy-Leu by LC/MS after Fmoc derivatization (Figure 4.5D, trace 2 & 3) (Fmoc-OH-Leu: *m/z* calcd for $C_{21}H_{23}NO_5$ $[M+Na]^+$ 392.1468, found 392.1465). A synthetic standard of Fmoc-5-OH-Leu, which has the same retention time, confirms the assignment of the enzymatic product (Figure 4.5D, trace 1). The second proposed product, MeP5C, could be captured by ortho-aminobenzaldehyde (*o*-AB)[19] to yield the characteristic A_{440} nm chromophore of the dihydroquinazolinium derivative in a pair of products (major to minor = 6/1, Figure 4.5D, trace 1 & 2) with identical mass (MeP5C: *m/z* calcd for $C_{13}H_{15}N_2O_2$ $[M]^+$ 231.1128, found 231.1121) and UV spectrum (Figure S8).[29] This is consistent with a pair of diastereomers created at C_1 on capture of the C_1 imine (Figure S9).[29] Scale up of the EcdK reaction was challenging, given that EcdK underwent suicidal inactivation during turnover (Figure 4.5F). This is characteristic of a subset of the nonheme iron oxygenases, as exemplified also by EcdG (Figure 4.4C), and most

probably reflects adventitious, uncontrolled reaction of an Fe^{IV}=O high valent intermediate in the oxygenation process. From the three traces in Figure 4.5F, there are about 5-6 turnovers/inactivation event for EcdK. Even with the low turnover/inactivation ratio it was possible to use enough pure EcdK to generate sufficient dihydroquinazolinium product adduct. It was purified by HPLC (Figure S8),[29] collected for NMR analysis of the major isomer, and the structure was validated (SI figure S10-S14).[29] In turn this supports the scheme in Figure 4B that EcdK acts in tandem to oxygenate C_5 of L-Leu. The *gem* diol product is the carbonyl hydration product of γ-Me-Glu-γ-semialdehyde, in turn in equilibrium with the cyclic imine that is trappable with *o*-aminobenzaldehyde.

To further corroborate the view that EcdK can act at C_5 of L-Leu iteratively, 5-OH-Leu was synthesized from an intermediate in MePro synthesis.[14] The 5-OH-Leu served as a substrate for EcdK yielding the same imine as seen from L-Leu (Figure 4.5D, trace 4 & 5; Figure 4.5E, trace 3 & 4). The cyclic imine, the penultimate compound in biosynthesis of the cyclic amino acid Me-Pro, presumably then undergoes reduction (e.g. by an NADH-dependent oxidoreductase) to afford Me-Pro. Recently, a comparable mononuclear iron-dependent L-Leu 5-hydroxylase has been characterized from *Nostoc punctiforme*.[20] Apparently, L-Leu 5-hydroxylase from *N. punctiforme* stops after the first hydroxylation, in contrast to EcdK, which performs both oxidation steps.

4.9 Anticandidal Activities of Echinocandin Variants

Earlier reports noted that **4.2** and **4.3** retained similar activity against *Candida* as **4.1**, albeit without detailed data.[6,21] Studies on the semisynthetic cilofungin, which has the same hexapeptide core as echinocandin but different fatty acid side chain, indicated a sixteen fold drop

in potency from the fully oxygenated hexapeptide scaffold to the tetra-deshydroxy type scaffold.[21] It is known that the nature of the hydrophobic acyl chain in the semisynthetic variants is an important determinant of their antifungal activity and toxicity.[5,22,23] Given that we had access to **4.1** from wild type and **4.3** from the *ΔecdH* strain, we examined their antifungal activity directly side by side along with other des-hydroxy analogs against *Candida albicans* strain ATCC 90234 with the results shown in Table 5.1. Under these assay conditions **4.3** is three fold more potent, while the compounds lacking the 3-hydroxy on homotyrosine (**4.5**–**4.7**) are 1.3-5.1 fold less active. Zambias et al noted from work on synthetic derivatives of cilofungin that the hydroxy groups on Pro_3 and Me-OH-Pro_6, along with the homoTyr residue, were the most important hydroxylation sites for anticandidal activity.[21] Our results are consistent with observations on cilofungin but we do not yet see production of deshydroxy Pro_3 or $MePro_6$ variants.

Table 4.1: MIC measurement of compound echinocandin B (**4.1**) and its deshydroxy analogs against *Candida albicans*.

Compound	MIC (μg/mL)
4.1 (wt)	0.6
4.3 (Δ*ecdH*)	0.2
4.5 (Δ*ecdG*)	3.1
4.6 (Δ*ecdG*)	0.8
4.7 (Δ*ecdG*)	1.6

4.10 Conclusion

Six hydroxyl groups are distributed over the four nonproteinogenic building blocks in **4.1**. Two additional hydroxylations, at C_5 of L-Leu have been presumed to initiate the pathway to 4-Me-Pro (as a precursor to the 3-OH-4-$MePro_6$ residue). Thus, at least eight hydroxylations

are required for generation of the mature echinocandin B scaffold, albeit the two hydroxyls introduced at C_5 of L-Leu are cryptic, *i.e.* not persisting in the mature echinocandin scaffold.

Among the questions we had at the outset was which of the putative three EcdGHK oxygenases, if any, act iteratively on Orn_1 and $homoTyr_4$ and which may act on Pro_3 and/or 4-Me-Pro_6 or the L-Leu precursor to MePro. The timing of side chain oxygenations in the antifungal scaffold assembly were also unknown: free amino acid, amino acyl- or peptidyl-S-NRPS stage, or on a des-hydroxy echinocandin framework. There was also ambiguity about whether hydroxylations must go in a prescribed order and whether independent routes of oxygenative maturation can go in parallel.

This study identifies five oxygenation events by these three iron enzymes. The nonheme iron EcdG works on free L-homotyrosine to generate 3-OH-L-homoTyr. This step occurs at an early stage during echinocandin biosynthesis before the building block is loaded onto the hexamodular NRPS EcdA. The hydroxylase for C_4 of homotyrosine remains unidentified. Knockout of *ecdG* or *ecdH* significantly decreases the hydroxylation efficiency at this position in the final echinocandin scaffold, although neither enzyme directly hydroxylates it. We thus propose that the homoTyr-4-hydroxylation reaction occurs after EcdH, probably as the last step during echinocandin B biosynthesis (Figure 4.6). That placement implies this particular hydroxylase takes the partially deshydroxy macrocycle as a substrate and can produce minor compounds **4.4** and **4.5**, with a preference for generating the fully hydroxylated echinocandin B scaffold.

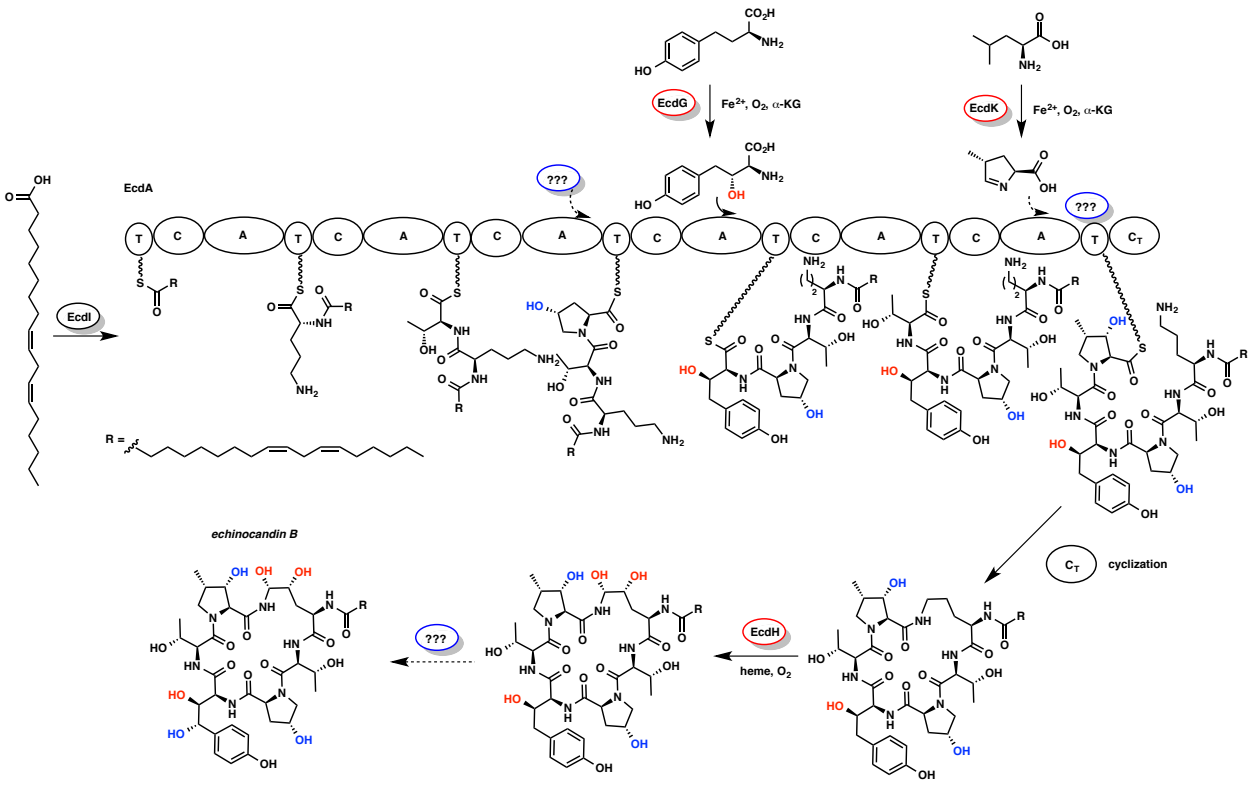


Figure 4.6. Proposed timing of post-translational hydroxylations for biosynthesis of echinocandin (**4.1**). Confirmed steps and modifications are highlighted in red, while proposed ones are in blue.

The hemeprotein EcdH, which iteratively works at C_5 and then C_4 of Orn_1, probably functions at a late stage after the release of the macrocyclic peptide from the NRPS EcdA, since a free C_5-OH-L-Orn would unravel to the aldehyde and not be competent to be joined to 3*S*-OH-$MePro_6$. EcdH is sensitive to the conformational change of the hexapeptide macrocycle when particular-OH groups are missing. When *ecdG* is deleted, EcdH becomes less efficient and produces three variants at Orn_1 (compound **4.5**–**4.8**). The EcdH-mediated hydroxylation at C_5 of Orn_1 creates the hydrolytically labile hemiaminal linkage: a subsequent elimination of water to the cyclic imine would be on pathway to hydrolytic linearization and deactivation of the

hexapeptide macrocycle. The post NRPS generation of that hemiaminal linkage is in contrast to known routes to imine linkages in other nonribosomal cyclic peptides such as nostocyclopeptide[24] and koranimine.[25] In those cases the NRPS assembly lines have C-terminal reductase domains that release the linear peptide aldehydes as nascent products. Imine formation with the N-terminal-NH_2 is the cyclization step. In contrast EcdA releases the cyclic hexapeptide apparently via a fungal macrocyclizing C_T domain and the hemiaminal (precursor to imine) is a post assembly line maturation step.

EcdK, in analogy to EcdG, works as a nonheme mononuclear iron oxygenase at the level of a free amino acid. In this case L-Leu is substrate and undergoes tandem hydroxylation at C_5 in the first and second step in the pathway to 4*R*-Me-proline assembly. Thus, both EcdH and EcdK do iterative hydroxylations, one on the same carbon (EcdK), the other on adjacent carbons (EcdH) of their substrates.

Three echinocandin oxygenation events are still to be characterized. To date we have no information on the identity of the hydroxylase(s) for Pro_3 (oxygenation at C_4) and 4*R*-Me-Pro_6 (oxygenation at C_3) and at what stage they occur. Given the absence of any deshydroxy forms of echinocandins at either of these two residues, hydroxylations on the macrocycle after release by the nonribosomal peptide synthetase EcdA is disfavored. Until the Pro/MePro hydroxylases are found, we cannot tell if the free amino acids or peptidyl-S-NRPS species are substrates. Precedent does suggest conversion of Pro to 4-OH-Pro can occur at the level of the free amino acid.[26] It may be that is also a reasonable expectation for the transformation of 4-Me-Pro to 3-OH-4-MePro. As noted earlier, the homoTyr-4-hydroxylase is also not yet identified. We have noted that two genes annotated as *htyE* and *htyF* are adjacent to the *htyA-D* genes required for

conversion of Tyr to homoTyr.[8] Whether either HtyE or HtyF are any of the missing three oxygenation catalysts for echinocandin B maturation is a subject for further study.

4.11 Experimental Section

4.11.1 Materials and Methods

Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of nitrogen using anhydrous solvents (either freshly distilled or passed through activated alumina columns). All commercially obtained reagents were used as received with the following exceptions. Reaction temperatures were controlled using an IKAmag temperature modulator, and unless stated otherwise, reactions were performed at room temperature (rt, approximately 23 °C). Thin-layer chromatography (TLC) was conducted with EMD gel 60 F254 pre-coated plates, (0.25 mm) and visualized using a combination of UV, anisaldehyde, ceric ammonium molybdate, and potassium permanganate staining. EMD silica gel 60 (particle size 0.040–0.063 mm) was used for flash column chromatography. ^{1}H NMR spectra were recorded on Bruker spectrometers (at 500 MHz). ^{13}C NMR spectra were recorded on Bruker Spectrometers (at 125 MHz).

4.11.2 Experimental Procedures

Gene knock-out in *Emericella rugulosa*: Procedure for the gene knockout for *E rugulosa* is based on a protocol established for *A. nidulans* A4 with modification.[12] The linear knockout cassette was constructed as done in previous studies[12] except that the glufosinate resistance gene driven by the *trpC* promoter was used as selection marker.

Transformation of *E. rugulosa* was done as follows: *E. rugulosa* conidiospores were inoculated to 250 mL of glucose minimal media with 10 mM ammonium tartrate as nitrogen source (GMMT)[8] at a concentration of 10^{7} conidiospores/mL. The culture was shaken at 28 °C for 16 hrs and 250 rpm. Thereafter, 2 g of germlings was digested at 28 °C while shaking at 100

rpm with 3 g of Vinotaste Enzyme Mixture (Novozyme) in 8 mL GMMT supplemented with 8 mL of 1.2 M KCl and 50 mM citric acid (pH 5.5). After 3 hrs, the digestion mixture was overlaid on top of an equal volume of 1.2 M sucrose and centrifuged at 1800 g and 4 °C. The protoplasts trapped on top of the sucrose cushion were isolated, washed with 0.6 M KCl three times and resuspended in 500 μL 0.6 M KCl and 50 mM $CaCl_2$ solution. Ten micrograms of the linear knockout cassette was added to 100 μL of resuspended protoplasts and the mixture was briefly vortexed for 3 seconds. 50 μL of PEG solution (0.6 M KCl, 50 mM $CaCl_2$, 75 mM polyethylene glycol with average molecular weight of 3250 g/mol, 10 mM Tris-HCl, pH 7.5) was added to the protoplast and the resulting mixture was incubated at 4 °C for 20 minutes. Thereafter, 1 mL of PEG solution was added to the protoplast mixture and incubated at room temperature for an additional 20 minutes and subsequently plated to GMMT plates supplemented with 1.2 M sorbitol and 10 mg/mL glufosinate.

Genomic DNA from the resulting transformants was isolated using ZR Fungal/Bacterial DNA Miniprep kit (Zymo) and was used as template for PCR screening. Primers for the PCR screening are given in Table S3.[29]

Culturing of *Emericella rugulosa* and mutants in large scale: The media and process used for ***Emericella rugulosa*** growth are adapted from US Patent 4,288,549.[10] Wild-type or verified mutant *E. rugulosa* cultures were grown in solid GMMT plates (or solid GMMT plates supplemented with 10 mg/mL glufosinate in the case of the mutants) at 28 °C for 4 days. The harvested conidiospores from the plates were inoculated to 8 L of medium 2 ([2.5% (w/v) glucose (Sigma), 1% peptone (BD Biosciences), 1% (w/v) starch (Sigma), 1%(w/v) molasses,

0.4% (w/v) N-Z Amine A (Sheffield Biosciences), 0.2% (w/v) calcium carbonate (Sigma)] and was shaken at 250 rpm and 28 °C for 7 days.

Preparative Scale Isolation of Echinocandins for NMR analysis: The extraction process for the echinocandins was adapted from US Patent 4,288,549.[10] The whole cell culture and broth was mixed with an equal volume of methanol and stirred thoroughly for one hour at 16 °C. It was then filtered through a layer of Miracloth and adjusted to pH 4.0 by the addition of concentrated HCl. The acidified filtrate was extracted twice with equal volumes of chloroform. The chloroform extracts were combined and concentrated *in vacuo* to 5-10 mL with a dark yellow color. The solution was centrifuged and the supernatant was loaded onto a Sephadex LH20 size-exclusion column (400x40 mm). The compounds were eluted by using 1:1 chloroform/methanol (v/v) with a rate of 1.5 mL/min. Fractions with desired compounds were combined and dried *in vacuo*. The mixture was re-suspended in methanol, centrifuged, and the supernatant was injected onto a Phenomenex Luna prep-HPLC C18 column (250 x 21.1 mm, 10 micron, part # 00G-4253-P0-AX) using a solvent gradient of 40 to100% B over 30 min (solvent A, 0.1% TFA/H2O; solvent B, 0.1% TFA/MeCN). Elution of echinocandins was monitored at 280 nm. The fractions with desired compounds were lyophilized for further analysis.

Culturing, feeding and extraction of *Emericella rugulosa* with deuterated amino acids for MS analysis: 10 mL *Emericella rugulosa* culture, either in medium 2 described above or GMMT, was inoculated and incubated at 28 °C without shaking. Deuterated amino acids (D_6-L-ornithine, D_7-L-proline, and D_7, ^{15}N-L-tyrosine) were purchased from Cambridge Isotope Laboratories (Andover, MA). Orn and Pro for feeding studies were prepared as neutral aqueous

solutions with a concentration of 60 mg /mL and sterile filtered. Tyrosine solution with a concentration of 36 mg / mL was prepared in a similar way but at pH 13 due to tyrosine's poor solubility at neutral pH. 250 μL of compound solutions were added to 10 mL *Emericella rugulosa* culture at Day 4 after inoculation, and harvested at Day 7. For compound extraction, an equal volume of methanol was added to the culture broth, mixed well, incubated at room temperature for 15 min and filtered. An equal volume of chloroform was then added to the mixture. The organic layer was separated, dried with anhydrous Na_2SO_4, and filtered. The crude extract was dried *in vacuo* and stored at –20 °C for further use.

LC-MS methods for natural compound profiles: All LCMS analysis was carried out on an Agilent 1200 Series HPLC coupled to an Agilent 6520 QTOF spectrometer. The crude material was re-dissolved in 200 μL methanol and filtered through a 0.45 μM membrane. 10 μL were injected for each analytical run. Separation was achieved on a 50 x 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 20 to 100% B over 30 min (solvent A: 0.1% formic acid in H_2O; solvent B: 0.1% formic acid in acetonitrile). Temperature was 350 °C.

Cloning of EcdG and EcdK: RNA from 7-day old wild-type *E. rugulosa* liquid static culture was isolated using Ribopure-Yeast kit (Ambion). First strand cDNA synthesis was performed with SuperScript III-First Strand Synthesis SuperMix using the RNA from wild-type *E. rugulosa* as template and oligo-dT as reverse-transcriptase primer. EcdG and EcdK cDNA were amplified using primers listed in Table S3 and cloned into p*ET*30-Xa/LIC vector. The resulting constructs

were transformed into *E.coli* Xl1 Blue cells (Stratagene) and sent for sequencing to verify correct splicing of the cDNA.

Reconstitution of EcdG/EcdK and hydroxylation assays: Anaerobic reconstitution of EcdG/EcdK with Fe(II), α-ketoglutarate (α-KG), and DTT was conducted as for other nonheme Fe(II)-dependent halogenases.[27] Then the hydroxylation assay was initiated by adding 100 μM or otherwise indicated concentration of reconstituted EcdG or EcdK to a 300 μL solution containing 50 mM HEPES (pH 7.0), 2 mM a-KG, and 1 mM homoTyr (EcdG) or Leu or 5-hydroxyl-Leu (EcdK). The mixture was incubated at ambient temperature and 50 μL aliquot was quenched by addition of 50 μL acetonitrile at various time points. The solution was centrifuged at 13k rpm for 5 min and the supernatant was subjected for further use.

Fmoc derivatization and LC/MS analysis: To a 100 μL of quenched hydroxylation assay mixture described above was added 50 μL of 200 mM sodium borate (pH 8.0) and 20 μL of 10 mM 9-fluorenylmethyl chloroformate (Fmoc-Cl). After incubation at room temperature for 5 min, 20 μl of 100 mM 1-adamantanamine was added to stop the reaction. 10 μL of the mixture were injected onto a 50 x 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 20 to 100% B over 20 min (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile).

Isolation of 3-OH-homoTyrosine: 1.2 mL of the quenched EcdG hydroxylation assay mixture described above was dried *in vacuo* and re-dissolved in 200 μL ddH_2O. 50 μL was injected for each run to a 100 *2.1 mm Thermo Hypercarb column (5 micron, part # 35005-102130). Solvent

A was composed of 20 mM perfluoropentanoic acid in water; solvent B was composed of acetonitrile without any additive. Separation was achieved by using the LC gradient 0% B for 5 min followed by 0-70% B over 30 min.

The process was monitored at 275 nm. Solutions from two major absorption peaks were collected and subjected to MS to confirm that hydroxyl-homoTyr was eluted prior to homoTyr. Hydroxyl-homoTyr from 10 runs was collected, combined, lyophilized and subjected to NMR.

Ortho-aminobenzaldehyde derivatization and LC/MS analysis: The hydroxylation assay containing 50 mM HEPES (pH 7.0), 2 mM α-KG, 1 mM Leu or 5-hydroxyl-Leu, and 0.2 mM reconstituted EcdK proceeded for 15 min at ambient temperature. The solution was filtered through a Microcon YM-3 (Amicon) microcentrifuge filter to remove the enzyme. The flow-through was collected, to which solid ortho-aminobenzaldehyde (*o*-AB) was directly added at 0.5 mg per 100 μL, mixed well and incubated at ambient temperature for 1 hour. Afterwards, the insoluble *o*-AB was removed by filtration and the solution was subjected to LC/MS or HPLC (see below).10 μL were injected onto a 50 x 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 1 to 100% B over 20 min (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile).

Isolation of *o*-AB-derivatized 3-methyl-D^1-pyrroline-5-carboxylic acid (MeP5C) by HPLC: 100 μL of *o*-AB derivatized solution described above was injected for each run to a 25 cm *4.6 mm Supelco Discovery C18 HPLC column (5 micron, part # 504971). Separation was achieved by using a very mild solvent gradient of 5-8% B over 20 min (solvent A: 0.1% TFA in water; solvent B: 0.1% TFA in acetonitrile). The process was monitored at 440 nm. Solutions from the

major absorption peak (the 2nd) were collected. Material from ~20 runs was combined, lyophilized and subjected to NMR studies.

Antifungal susceptibility test by broth microdilution method: The potency of echinocandin B (**5.1**) and its analogs against *Candida albicans* was measured following a published protocol using 96-well plate.[29] MIC_{90} was designated as the lowest concentration that produced an increase of less than 10% in OD over that of the adjacent sterile control.

4*R*-methyl proline (4.9): Following the protocol disclosed by Rapoport et al,[27] 4*R*-methyl proline was synthesized starting from commercially available materials.

3*S*-OH-4*R*-methyl proline (4.10): Following the protocol disclosed by Gurjar et al, 3*S*-OH-4*R*-methyl proline was synthesized from commercially available materials.[28]

4-hydroxyl-L-leucine: Following the protocol described by Moore et al,[14] 4-hydroxyl-L-leucine was synthesized starting from commercially available materials.

4.12 Notes and References

(1) Benz, F.; Knüsel, F.; Nüesch, J.; Treichler, H.; Voser, W.; Nyfeler, R.; Keller-Schierlein, W. *Helv. Chim. Acta* **1974**, *57*, 2459.

(2) Sawistowska-Schröder, E. T.; Kerridge, D.; Perry, H. *FEBS Lett.* **1984**, *173*, 134.

(3) Bouffard, F. A.; Zambias, R. A.; Dropinski, J. F.; Balkovec, J. M.; Hammond, M. L.; Abruzzo, G. K.; Bartizal, K. F.; Marrinan, J. A.; Kurtz, M. B. *J. Med. Chem.* **1994**, *37*, 222.

(4) Tomishima, M.; Oohki, H.; Yamada, A.; Takasugi, H.; Maki, K.; Tawara, S.; Tanaka, H. *J. Antibiot (Tokyo)* **1999**, *52*, 674.

(5) Debono, M.; Turner, W. W.; LaGrandeur, L.; Burkhardt, F. J.; Nissen, J. S.; Nichols, K. K.; Rodriguez, M. J.; Zweifel, M. J.; Zeckner, D. J. *J. Med. Chem.* **1995**, *38*, 3271.

(6) Traber, R.; Keller-Juslén, C.; Loosli, H.-R.; Kuhn, M.; Von Wartburg, A. *Helv. Chim. Acta* **1979**, *62*, 1252.

(7) Tóth, V.; Nagy, C.; Miskei, M.; Pócsi, I.; Emri, T. *Folia Microbiol.* **2011**, *56*, 381.

(8) Cacho, R. A.; Jiang, W.; Chooi, Y.-H.; Walsh, C. T.; Tang, Y. *J. Am. Chem. Soc.* **2012**, *134*, 16781.

(9) Adefarati, A. A.; Giacobbe, R. A.; Hensens, O. D.; Tkacz, J. S. *J. Am. Chem. Soc.* **1991**, *113*, 3542.

(10) Boeck, L. D.; Kastner, R. E. (Eli Lilly and Co.) US Patent 4,288,549 **1981**.

(11) Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R. *Genetics* **2006**, *172*, 1557.

(12) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. *Nat. Protocols* **2007**, *1*, 3111.

(13) Evans, D. A.; Weber, A. E. *J. Am. Chem. Soc.* **1987**, *109*, 7151.

(14) Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. *J. Org. Chem* **2003**, *68*, 83.

(15) Strushkevich, N.; MacKenzie, F.; Cherkesova, T.; Grabovec, I.; Usanov, S.; Park, H.-W. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10139.

(16) Galonić, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M.; Krebs, C. *Nat. Chem. Biol.* **2007**, *3*, 113.

(17) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10111.

(18) Kurokawa, N.; Ohfune, Y. *J. Am. Chem. Soc.* **1986**, *108*, 6041.

(19) Mezl, V. A.; Knox, W. E. *Anal. Biochem.* **1976**, *74*, 430.

(20) Hibi, M.; Kawashima, T.; Sokolov, P.; Smirnov, S.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. *Appl Microbiol Biotechnol* **2012**, 1.

(21) Zambias, R. A.; Hammond, M. L.; Heck, J. V.; Bartizal, K.; Trainor, C.; Abruzzo, G.; Schmatz, D. M.; Nollstadt, K. M. *J. Med. Chem.* **1992**, *35*, 2843.

(22) Boeck, L. D.; Fukuda, D. S.; Abbott, B. J. *J. Antibiot. (Tokyo)* **1989**, *42*, 382.

(23) Debono, M.; Abbott, B. J.; Fukuda, D. S.; Barnhart, M.; Willard, K. E.; Molloy, R. M.; Michel, K H. *J. Antibiot (Tokyo)* **1989**, *42*, 389.

(24) Becker, J. E.; Moore, R. E.; Moore, B. S. *Gene* **2004**, *325*, 35.

(25) Evans, B. S.; Ntai, I.; Chen, Y.; Robinson, S. J.; Kelleher, N. L. *J. Am. Chem. Soc.* **2011**, *133*, 7316.

(26) Lawrence, C. C.; Sobey, W. J.; Field, R. A.; Baldwin, J. E.; Schofield, C. J. *Biochem. J.* **19 96**, *313*, 185.

(27) Mohapatra, D. K.; Mondal, D.; Chorghade, M. S.; Gurjar, M. K. *Tetrahedron Lett*. **2006**, 47, 9215.

(28) Koskinen, A. M. P.; Rapoport, H. *J. Org. Chem* **1989**, *54*, 1859.

(29) For supplementary figures, please refer to the original manuscript's Supporting Information.

APPENDIX THREE

Spectra Relevant to Chapter Four:

EcdGHK are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold

Wei Jiang, Ralph A. Cacho, Grace Chiou, Neil K. Garg, Yi Tang, Christopher T. Walsh.

J. Am. Chem. Soc. **2013**, *135*, 4457–4466.

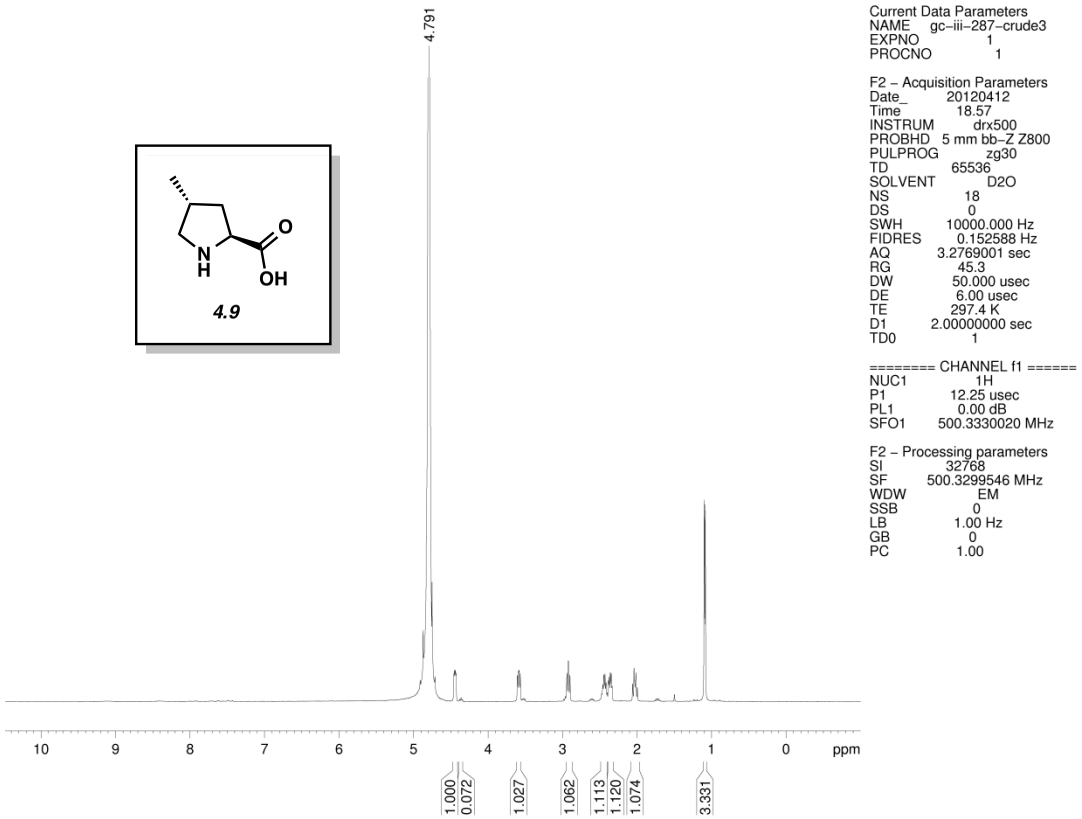


Figure A3.1 ^{1}H NMR (500 MHz, D_2O) of compound **4.9**.

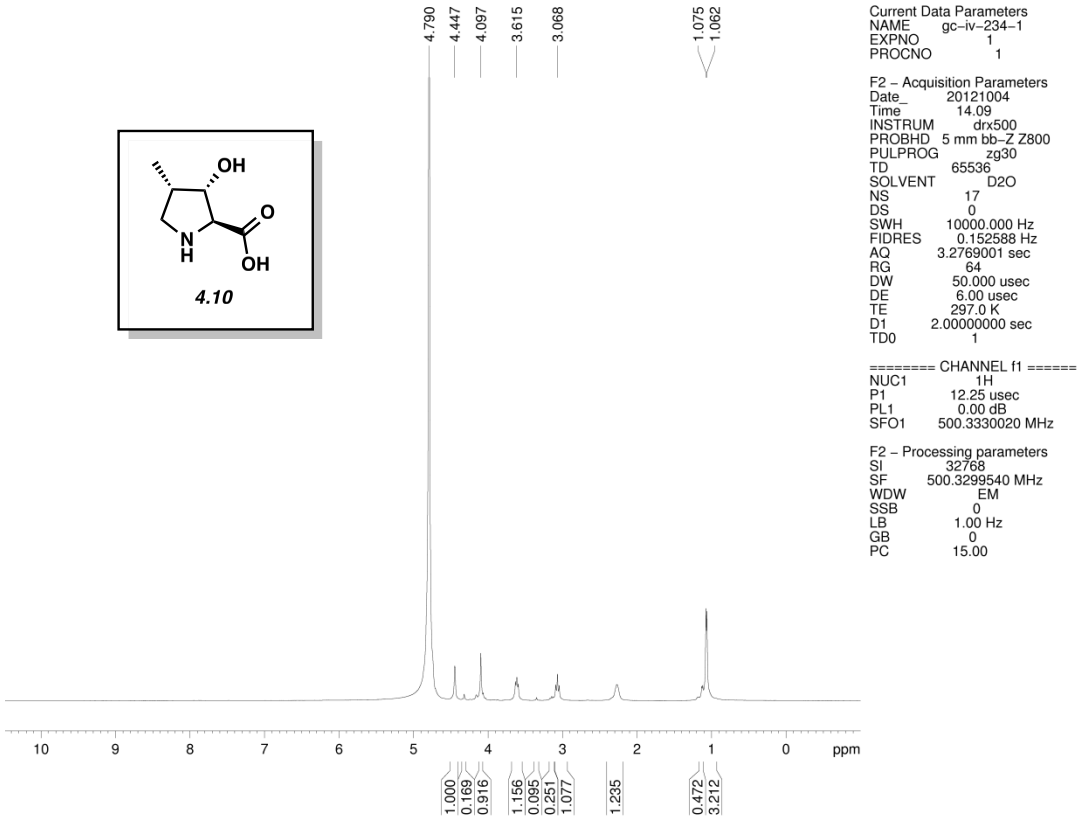


Figure A3.2 ^{1}H NMR (500 MHz, D_2O) of compound **4.10**.

CHAPTER FIVE

Identification and Characterization of the Chaetoviridin and Chaetomugilin Gene Cluster in *Chaetomium globosum* Reveal Dual Functions of an Iterative Highly-Reducing Polyketide Synthase

Jaclyn M. Winter, Michio Sato, Satoru Sugimoto, Grace Chiou, Neil K. Garg, Yi Tang, and Kenji Watanabe.

J. Am. Chem. Soc. **2012**, *134*, 17900–17903.

5.1 Abstract

We report the identification and characterization of the *caz* biosynthetic cluster from *C. globosum* and the characterization of a highly-reducing polyketide synthase (PKS) that acts in both a sequential and convergent manner with a nonreducing PKS to form the chaetomugilin and chaetoviridin azaphilones. Genetic inactivation studies verified the involvement of individual *caz* genes in the biosynthesis of the azaphilones. Through in vitro reconstitution, we demonstrated the in vitro synthesis of chaetoviridin A (**5.3**) from the pyrano-quinone intermediate cazisochromene **5.9** using the highly-reducing PKS and an acyltransferase.

5.2 Introduction

Polyketides produced by fungi possess a diverse array of structural features and biological activities.[1,2] These compounds are assembled by iterative polyketide synthases (IPKSs), which are megasynthases containing single copies of catalytic domains that are programmed to function repeatedly in different combinations.[3] Additional programming complexity and structural diversity are generated when multiple IPKSs function collaboratively as illustrated in the biosynthesis of the azaphilones asperfuranone (**5.1**) and azanigerone A (**5.2**). This partnership between tandem IPKS can occur in a sequential manner in which the product of

a highly-reducing PKS (HR-PKS) is transferred downstream to a nonreducing PKS (NR-PKS) for further chain elongation, as observed with **5.1**;[4] or in a convergent manner in which the HR-PKS and NR-PKS function independently of one another and their individual products are combined, as demonstrated with **5.2**.[5] Understanding the functions of individual IPKSs, as well as how collaborative IPKSs are recruited to function together, is therefore an important step towards decoding the relationship between fungal IPKSs and their products. In this work, we discovered a single HR-PKS from *Chaetomium globosum* that can partner in both sequential and convergent fashions with a NR-PKS to synthesize two structurally distinct fragments that are incorporated into the antifungal chaetoviridin A (**5.3**) and cytotoxic chaetomugilin A (**5.4**) (Figure 5.1). This finding uncovers a previously unknown mode of collaboration among IPKSs and further demonstrates how a limited number of IPKSs can be combined to generate highly complex scaffolds.

5.3 Azaphilones and Polyketide Synthases (PKSs)

The common bicyclic pyranoquinone scaffold observed in the azaphilones family of polyketides is biosynthesized by an NR-PKS. The reduced substituents, such as those in **5.1** and **5.2** are synthesized by a partnering HR-PKS found in each gene cluster.[4,5] *C. globosum* is a filamentous fungi that has been reported to produce a large variety of azaphilones, including the chaetoviridins A (**5.3**),[6a] B (**5.5**),[6a,b] and C (**5.6**),[6c] as well as chaetomugilins A (**5.4**),[6c] and M **5.7**)[6d] (Figure 6.1). Unlike **5.1** and **5.2**, the chlorinated **5.3**, **5.6**, and **5.7** contain an angular lactone ring attached to the isochromenone core, whereas **5.4** and **5.5** contain a tetracyclic isochromenone-lactol-lactone structure.[6b,7] On the basis of their structural similarities, we propose that **5.3–5.7** are assembled via a common biosynthetic pathway. Feeding experiments

with [1,2-$^{13}C_2$]acetate confirmed that the entire molecule of **5.4** is derived from a polyketide backbone (Figures 6.1 and S1[8]), which is consistent with previous isotopic experiments with structurally related azaphilones.[9] Structural inspection of **5.3** and **5.4** reveals that in contrast to **5.1** and **5.2**, additional IPKS machinery must be required for their biosynthesis. It can be speculated that assembly of the northern portions of **5.3** and **5.4** follows the biosynthetic logic of **5.1**, in which a sequential collaboration of HR- and NR-PKS is required; whereas the southern portions mimic that of **5.2**, in which the product of a different IPKS is combined with the northern portion in a convergent manner. Hence, we hypothesized that up to three IPKSs may be found in the biosynthetic gene cluster of **5.3–5.7**.

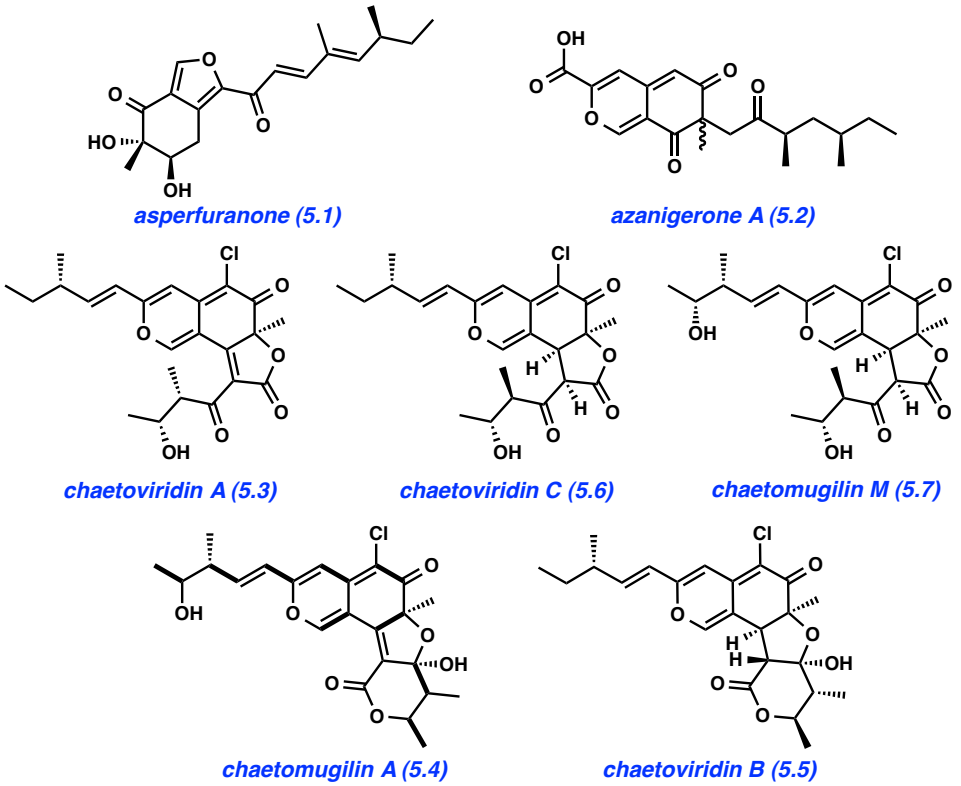


Figure 5.1. Structures of selected azaphilones and azaphilone-like molecules. The bold lines in chaetomugilin A (**5.4**) show enrichment from [1,2-$^{13}C_2$] acetate.

5.4 Identification of the Gene Cluster

Using the HR-PKS and NR-PKS sequences from the gene clusters of **5.1** and **5.2** identified in *Aspergillus nidulans*[4] and *A. niger*,[5] respectively, the genome of *C. globosum* was scanned for putative azaphilone biosynthetic clusters containing two or more IPKSs. Three clusters containing a HR-PKS and a NR-PKS in close proximity to each other were identified. Additionally, each cluster contained a predicted $FADH_2$-dependent halogenase that would introduce a chlorine atom onto the bicyclic core. Interestingly, none of these clusters contained three IPKSs as we hypothesized. Only one cluster (hereby named the *caz* cluster) encoded a putative acyltransferase (*cazE*) that would be required for transferring the convergent polyketide fragment (southern portion) to the pyranoquinone substrate (northern portion) to form **5.3–5.7**. This 65-kb *caz* cluster was annotated to encode 16 open reading frames, including genes encoding the HR-PKS (*cazF*) and NR-PKS (*cazM*) (Figures 5.2A and S2[8]).

To confirm the involvement of the *caz* cluster in the biosynthesis of **5.3**–**5.7**, gene inactivation of the NR-PKS *caz*M was accomplished via a fusion PCR gene targeting cassette (Figure S3[8]).[10] The Δ*caz*M inactivation completely abolished production of the azaphilones. Inactivation of the HR-PKS *caz*F also completely abolished production of **5.3**–**5.7** without the recovery of any pyranoquinone intermediates, suggesting an essential role of both CazM and CazF in the synthesis of the northern portion of **5.3**–**5.7**. This is consistent with the collaborative mode of interaction between the two IPKSs, in which the 4-methyl-hex-2-enoate **5.8** (Figure 5.3) is the most likely product synthesized by CazF and is transferred to CazM via the starter-unit:ACP transacylase (SAT) domain of CazM as a starter unit for further chain elongation (Figure 5.3). In this model, the NR-PKS is unable to produce any polyketide product in the absence of the starter unit.[11]

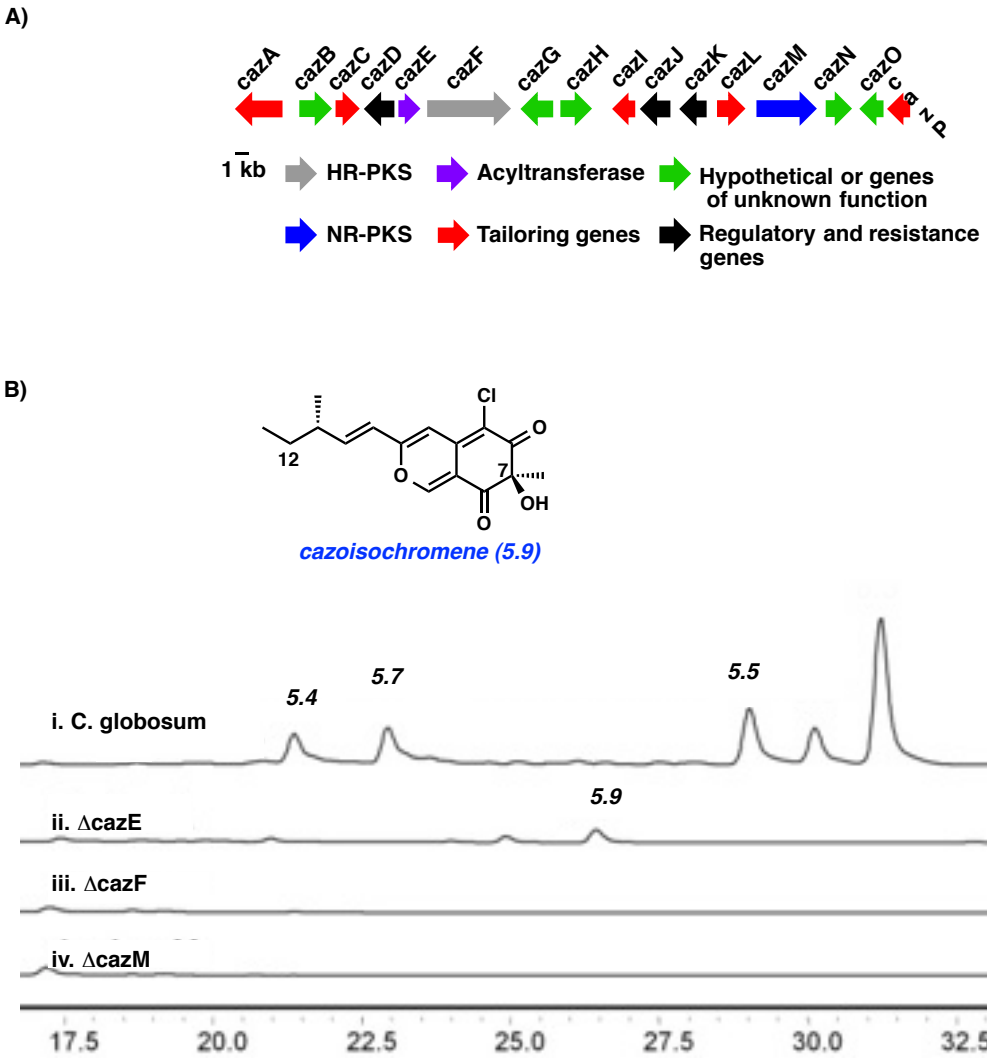


Figure 5.2. Genetic verification of the *caz* biosynthetic cluster in *C. globosum*. A) Organization of the 65 kb *caz* cluster. B) LC-MS analysis (observed at 360 nm) of (i) azaphilones produced by wild-type *C. globosum*, (ii) formation of **5.9** in the Δ*caz*E mutant; and (iii, iv) abolishment of the production of **5.3**–**5.7** in the (iii) Δ*caz*F and (iv) Δ*caz*M mutants.

To obtain further confirmation of this mode of interaction between the two *caz* IPKSs, we inactivated the acyltransferase *caz*E that may be involved in connecting the northern and southern portions of **5.3**–**5.7**. We reasoned that inactivation of *caz*E should therefore have no effect on the sequential collaboration between CazM and CazF. Indeed, the Δ*caz*E mutant was

unable to produce **5.3–5.7** and instead accumulated an intermediate **5.9** with UV absorbance consistent with that of a pyranoquinone (Figure S4–S7).[8] The structure of **5.9**, which we have named cazisochromene, was elucidated by 1D and 2D NMR spectroscopy (Figure 5.2B). The structure of the reduced substituent in **5.9** is consistent with that of **5.3**, **5.5**, and **5.6**, further confirming **5.9** is an authentic intermediate in the *caz* pathway. However, the 12-hydroxyl derivative of **5.9** could not be detected in the culture suggesting that the hydroxyl moiety present at this position in **5.9** and **5.9** may be installed by post-PKS oxygenases from the *caz* cluster rather than formed from incomplete reduction by the HR-PKS CazF. This hypothesis was supported by monitoring the time course of metabolite production, and observing that the amount of **5.3** dramatically decreased while the amount of **5.4** increased over time (Figure S8).[8]

5.5 Proposed Biosynthetic Pathway

With **5.9** in hand, we propose the biosynthetic pathway of **5.3–5.7** shown in Figure 5.3. Following transfer of **5.8** from CazF to CazM, the polyketide undergoes four additional rounds of elongation and is methylated once, forming the heptaketide intermediate **5.10**. While tethered to the acyl carrier protein (ACP), **5.10** undergoes a product template (PT)-domain-mediated regioselective cyclization via a C2–C7 aldol condensation followed by reductive release to form benzaldehyde **5.11**.[12,13] Chlorination of the aromatic ring catalyzed by the halogenase CazI and subsequent hydroxylation-catalyzed annulation at C7[5,14] by the predicted monooxygenase CazL then affords **5.9**. Given that the *caz* cluster only contains two IPKSs instead of the initially hypothesized three, either CazF or CazM must provide the additional southern triketide fragment. Since the *caz* cluster contains only two IPKSs instead of the initially hypothesized three, either CazF or CazM must provide the additional southern triketide fragment. We propose

that CazF could act as a dual functioning IPKS that supplies not only the reduced triketide **5.8** for CazM, but also a more oxidized intermediate such as **5.12** for CazE (Figure 5.3). Following CazE catalyzed acylation, intramolecular aldol condensation of the newly formed β-keto ester yields **5.3.** Further modification of **5.3** through reduction (**5.3→5.6**), lactone cleavage and rearrangement, and C12 oxidation furnishes products **5.4**–**5.7**.

5.6 Elucidation of the Biosynthetic Pathway

To test the involvement of CazF and CazE in the conversion of **5.9** into **5.3**, both enzymes were cloned and expressed from heterologous hosts for in vitro reconstitution. CazE (54 kDa) was solubly expressed as an *N*-terminal octahistidine-tagged protein from *Escherichia coli* BL21(DE3)/pJW07637.[15] CazF was solubly expressed as a *C*-terminal hexahistidine-tagged protein from *Saccharomyces cerevisiae* BJ5464-NpgA/pJW07638. Both proteins were purified to near homogeneity with yields of 52 mg/L and 9.5 mg/L, respectively (Figures S9–S11).[8,16] The minimal PKS activity of CazF was evaluated by incubating the enzyme with 2 mM malonyl-CoA. Treatment of the reaction mixture with 1 M NaOH (base hydrolysis) followed by LC–MS analysis of the organic extract showed production of the triacetic acid lactone **5.13** (Figure 5.4A), which is formed from the spontaneous enolization and cyclization of the unreduced triketide.[11,17] Intriguingly, in the absence of base hydrolysis, significantly lower levels of **5.13** were detected, indicating that the triketide may be fixed in an extended conformation by CazF to prevent lactonization and release. However, when CazE was added to the reaction, levels of **5.13** comparable to that of the base-hydrolyzed reaction were produced (Figure S12).[8] This suggested a likely interaction between CazE and CazF, in which a conformational change in CazF after it

interacts with CazE may expose the unreduced triketide to the aqueous environment, resulting in its subsequent lactonization and release.

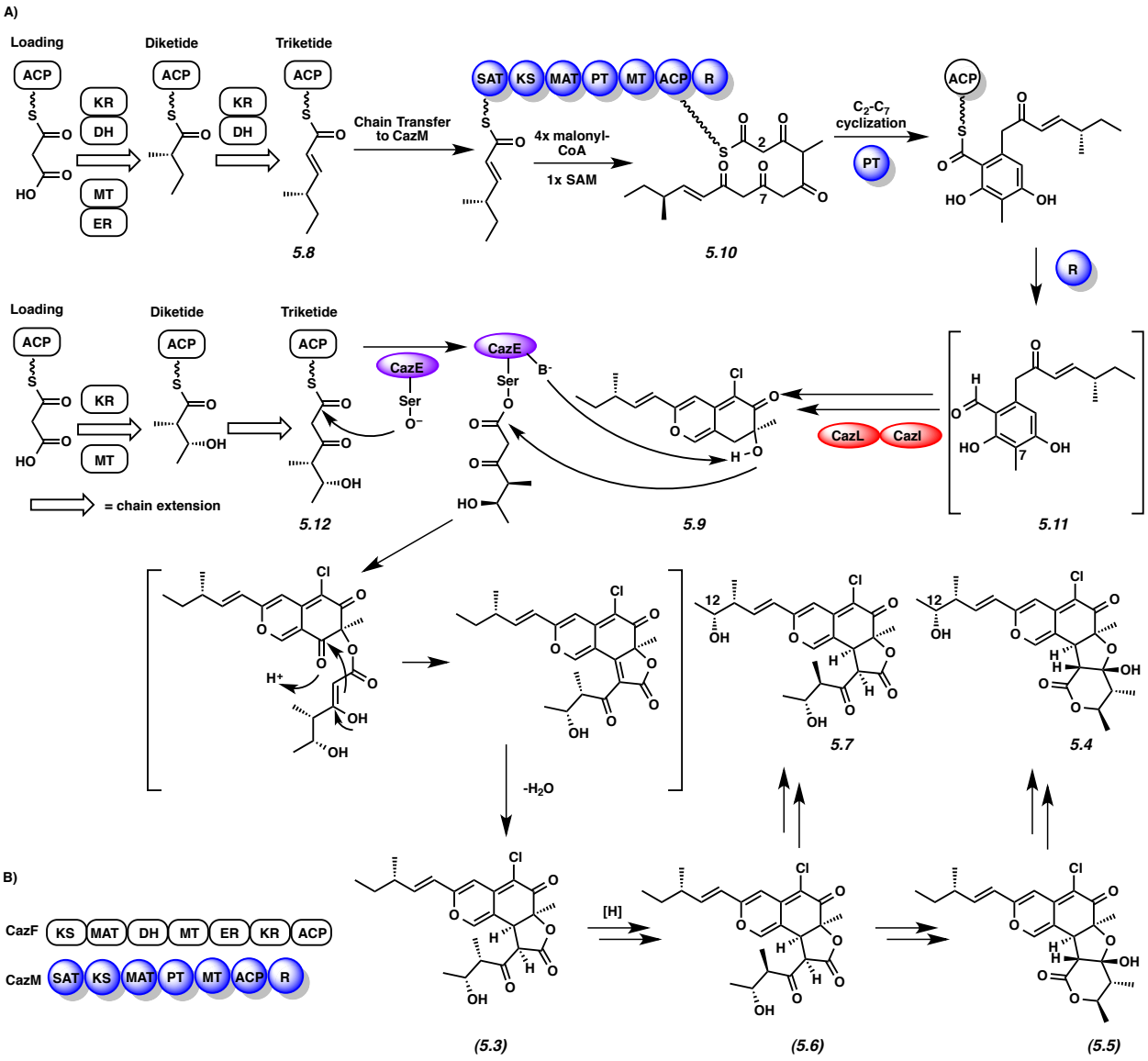


Figure 5.3. Proposed biosynthesis of the chaetoviridins and chaetomugilins from *C. glbosoum*. A) Biosynthesis of **5.3**–**5.7** by the *caz* pathway. B) Domain organization of the HR-PKS CazF and NR-PKS CazM. CazF consists of a ketosynthase (KS), malonyl-CoA:ACP acyltransferase (MAT), dehydratase (DH), methyltransferase (MT), enoyl reductase (ER), ketoreductase (KR)

and acyl carrier protein (ACP). CazM consists of a starter-unit:ACP transacylase (SAT), KS, MAT, product template (PT), MT, ACP and reductive domain (R).

We next examined the tailoring domains of CazF. The activity of the MT domain was assessed by adding 2 mM *S*-adenosyl-L-methionine (SAM) to CazF together with malonyl-CoA. Following incubation and base hydrolysis, dimethylated α-pyrone **5.14** (Figure 5.4A) was isolated and found to match a synthetically prepared standard (Figure S13–14).[8,18] The formation of **5.14** is indicative of the correct timing of the MT domain following formation of diketide intermediates on CazF, which is required for the formation of both **5.8** and **5.12**. To assay the functions of the KR, DH and ER domains, acetoacetyl-*S*-*N*-acetyl cysteamine (acetoacetyl-SNAC) was incubated with CazF and NADPH. LC–MS analysis of the organic extract showed accumulation of 3-hydroxybutanoyl-SNAC and butyryl-SNAC, confirming that all reductive domains were functional for the purified enzyme (Figure S15).[8]

After confirming that all of the domains of CazF are catalytically active, we attempted to synthesize **5.3** from **5.9**. Equimolar amounts of CazF and CazE were incubated with **5.9**, NADPH, SAM, and malonyl-CoA. LC–MS analysis of the organic extract revealed the formation of a new product whose retention time, UV profile, and isotopic mass ratio were identical to those of **5.3** (Figure 5.4B). When either CazF or CazE was omitted from the reaction, the formation of **5.3** could not be detected. Additionally, when only malonyl-CoA was added to the reaction or when NADPH was omitted, no new product formation was detected. However, when only SAM was omitted from the reaction, a new product with a shorter retention time than **5.3** and loss of a methyl group in mass was detected by the in vitro assay. On the basis of these observations, we assigned this product to be 4′-desmethylchaetoviridin A (**5.15**).

Therefore, the in vitro reactions unequivocally confirmed the involvement of the HR-PKS in providing the southern portion of **5.3**–**5.7** and provided another example of an acyltransferase-mediated convergent synthesis of fungal polyketides.[19] The requirement of NADPH to produce the chaetoviridins suggests that CazE may possess substrate specificity toward a triketide reduced at the δ position, while the formation of **5.15** in the absence of SAM indicates that methylation at the γ position is not essential.

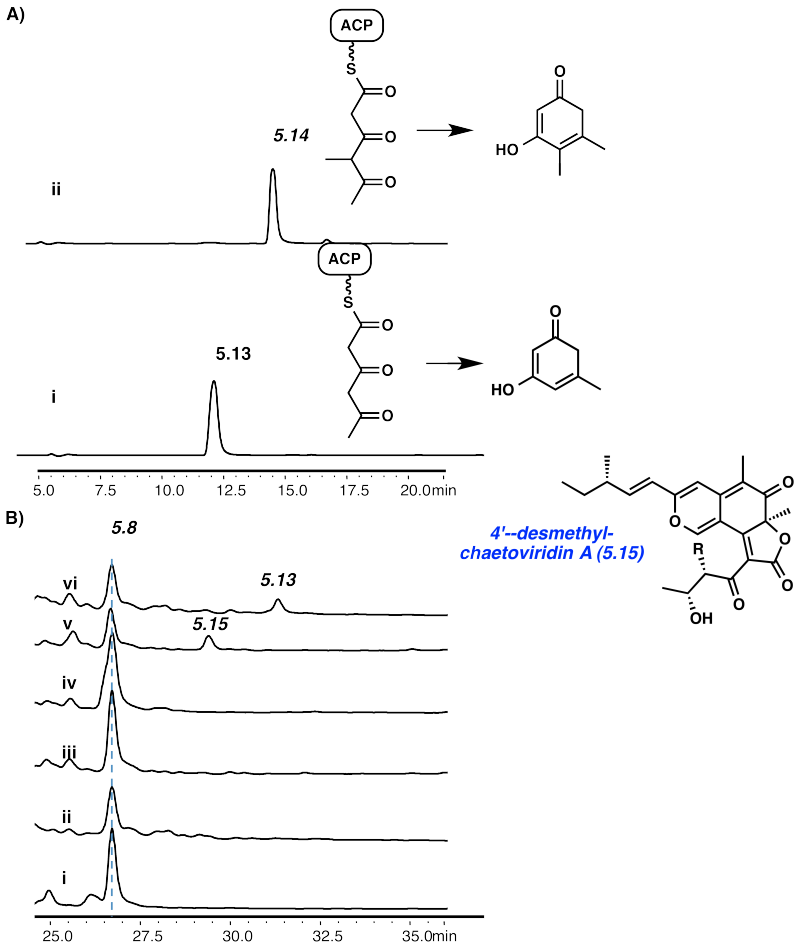


Figure 5.4. Reconstitution of CazE and CazF activity in vitro. A) HPLC analysis (280 nm) of α-pyrones biosynthesized by CazF in the presence of (i) malonyl-CoA and (ii) malonyl-CoA and SAM. B) HPLC analysis (360 nm) of chaetoviridins enzymatically synthesized in vitro when (i) CazF was incubated with **5.9**, malonyl-CoA, NADPH and SAM; (ii) CazE was incubated with **5.9**, malonyl-CoA, NADPH and SAM; (iii) CazE and CazF were incubated with malonyl-CoA; (iv) CazE and CazF were incubated with malonyl-CoA and SAM; (v) CazE and CazF were incubated with malonyl-CoA and NADPH; (vi) CazE and CazF were incubated with malonyl-CoA, NADPH and SAM.

5.7 Conclusion

In summary, we have identified and characterized the caz biosynthetic cluster that is responsible for the production of the chaetoviridin and chaetomugilin azaphilones from *C. globosum*. Through in vivo gene inactivation experiments, we have established that the HR-PKS CazF most likely supplies a highly reduced methylated triketide product such as **5.8** to the NR-PKS CazM. Since CazM cannot be solubly expressed at this point, we cannot exclude the possibility that CazF transfers a polyketide product of different length to CazM as a starter unit. Additionally, through in vitro reconstitution experiments with the acyltransferase CazE, we established that CazF can also produce a more oxidized triketide product, **5.12**, that is transacylated by CazE to a pyranoquinone intermediate, **5.9**, to afford **5.3**. Hence, CazF serves as an unprecedented dual function in completing the biosynthesis of a polyketide product. The timing of substrate (**5.8** or **5.12**) transfer from CazF must be precisely orchestrated by the two interacting acyltransferases in the form of the freestanding CazE and the SAT domain of CazM. Our work therefore unveils a new level of programming complexity among the fungal PKSs.

5.8 Experimental Section

5.8.1 Materials

Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of nitrogen using anhydrous solvents (either freshly distilled or passed through activated alumina columns). All commercially available reagents were used as received unless otherwise specified. Reaction temperatures were controlled using an IKAmag temperature modulator, and unless stated otherwise, reactions were performed at room temperature (rt, approximately 23 °C). Thin-layer chromatography (TLC) was conducted with EMD gel 60 F254 pre-coated plates (0.25 mm) and visualized using a combination of UV, anisaldehyde, iodine, and potassium permanganate staining. EMD silica gel 60 (particle size 0.040–0.063 mm) was used for flash column chromatography.

Strains: *Chaetomium globosum* CBS 148.51 was obtained from The American Type Culture Collection. *Chaetomium globosum ΔCgligD*[1] was used for all gene inactivation experiments. *Saccharomyces cerevisiae* BY4741 was used for constructing the *caz*E inactivation cassette. *Escherichia coli* BL21(DE3) was used for expressing the acyltransferase CazE. *Saccharomyces cerevisiae* BJ5464-NpgA[2] (*MATα ura3-52 trp1 leu2-Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*), which contains a chromosomal copy of the phosphopantetheinyl transferase NpgA,[3] was used for expressing the HR-PKS CazF.

Culture conditions: *Chaetomium globosum* CBS 148.51 was cultured in 1 L of PD medium (24 g Potato Dextrose from Fluka Analytical and Milli-Q filtered deionized water) in 150 x 15 mm petri dishes. Each plate contained 75 mL of media and was incubated at 28ºC for 3–7 days.

Inoculation of each plate was carried-out with 1% of a 1 x 10^8 spore suspension. For the [1,2-$^{13}C_2$]acetate feeding study, 100 mg of labeled precursor was added to 1 L of PD medium at the time of inoculation. To purify cazisochromene **5.9** for structure elucidation and in vitro assays, the Δ*lig*D *C. globosum* strain was cultured in 1 L of MYG medium (10 g/L malt extract, 4 g/L yeast extract, 4 g/L glucose). Cultivation was carried out at 30 ºC for four days at 250 rpm.

Chemicals and spectral analysis: Labeled [1,2-$^{13}C_2$]acetate was purchased from Cambridge Isotope Laboratories, Inc. The 4-hydroxy-5,6-dimethylpyran-2-one standard for in vitro assays was prepared as described previously.[4] All solvents and other chemicals used were of analytical grade. ^{1}H NMR spectra were obtained on Bruker 400-MHz and 500-MHz spectrometers a 500 MHz Bruker AV500 spectrometer and a JEOL JNM-ECA 500 MHz spectrometer (125 MHz for ^{13}C). ^{13}C NMR spectrum for [1,2-$^{13}C_2$]acetate-enriched chaetomugilin A was obtained on a 500 MHz Bruker AV500 spectrometer with a 5mm dual cryoprobe.

HPLC analysis and purification of chaetoviridins and chaetomugilins*:* All petri dishes containing PD media were pooled together and the mycelium and media were extracted with ethyl acetate, dried over anhydrous $MgSO_4$ and concentrated *in vacuo*. From the 1 L culture, the crude extract was fractionated using reverse phase C18 flash column chromatography (Fisher Scientific, PrepSep C18 1 g/6 ml) with 20% MeCN:H_2O, 40% MeCN:H_2O, 60% MeCN:H_2O, 80% MeCN:H_2O, and 100% MeCN as the mobile phase. The fractions were analyzed in positive and negative mode on a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer (gas flow set to 4.2 mL/min, drying temperature set to 200 °C, and capillary voltage set to 2000 V). A Luna C18(2) 100 x 2 mm column was used with a flow rate of 0.1 mL/min and a linear solvent

gradient of 5–95% MeCN:H_2O containing 0.1% formic acid over a period of 30 min followed by 15 minutes of 95% MeCH:H_2O.

Isolation of cazisochromene (5.9): The MYG culture was centrifuged and the supernatant was extracted with 2 x 1 L volumes of ethyl acetate. The ethyl acetate extracts were combined and concentrated *in vacuo,* which yielded an oily residue. The residue was fractionated by silica gel flash column chromatography using 99% $CHCl_3$:MeOH. Fractions containing the target product were collected and further purified by reverse phase HPLC (Nacalai Tesque Inc.) using a COSMOSIL 5 C18 MS-II 5μm 20 x 250 mm column. Compound **5.9** was isolated at 1 mg/L and its chemical structure was elucidated from the spectroscopic data given in SI Table 4.[8]

Isolation of chaetomugilin A (5.4) and chaetomugililn M (5.7): The PD cultures were extracted after seven days of incubation. Compound **5.4** and **5.7** eluted in the 40% MeCN:H_2O fraction and were purified by HPLC (Beckman Coulter) using a Luna C18 column (250 x 10 mm; 5μm particle size) and an isocratic condition of 45% MeCN:H_2O containing 0.1% TFA with a flow rate of 2.0 mL/min. The t_R of compound **5.4** was 13 min and the t_R of compound **5.7** was 23 min. The coupling constants for [1,2-$^{13}C_2$]-enriched **5.4** are shown in SI Table 2.[8]

Isolation of chaetoviridin A (5.3), chaetoviridin B (5.5) and chaetoviridin C (5.6): The PD cultures were extracted after three days of incubation. The azaphilone metabolites **5.3**, **5.5**, and **5.6** eluted in the 60% MeCN:H_2O fraction described above and were purified by HPLC (Beckman Coulter) using a Luna C18 column (250 x 10 mm; 5μm particle size) and an isocratic

condition of 50% MeCN:H_2O containing 0.1% TFA with a flow rate of 2.0 mL/min. The t_R of **5.3**, **5.5**, and **5.6** were 35 minutes, 28 minutes and 32 minutes, respectively.

General molecular biology experiments: PCR reactions were carried out using Platinum Pfx DNA polymerase (Invitrogen), Phusion DNA polymerase (New England Biolabs), and PrimeSTAR GXL DNA polymerase (Takara). Restriction enzymes were purchased from New England Biolabs and used according to the manufacturer's instructions. pCR®-blunt (Invitrogen) was used to construct recombinant DNA products and DNA sequencing was performed by Retrogen, California, USA and by Macrogen Japan Corporation. DNA manipulation using standard techniques[5] were performed in *E. coli* TOP10 (Invitrogen) and *E. coli* XL1-Blue (Stratagene).

Construction of gene inactivation cassettes: Disruption cassettes containing the hygromycin phosphotransferase *hph* gene, which confers resistance to hygromycin, and the *tub* promoter were designed by fusion PCR[6] for *caz*F, and *caz*M. The *hph* resistance marker was amplified from plasmid pKW3202[7] using primer pair cazF_hpfF and cazF_hpfR for inactivation of *caz*F and primer pair cazM_hpfF and cazM_hpfR for inactivation of *caz*M. The knockout cassettes were constructed as previously described[6] except 2 kb of homologous DNA located upstream and downstream of the targeted gene were attached to the resistance marker for the *caz*F inactivation, whereas 1 kb was attached for the *caz*M cassette. For construction of the *caz*F cassette, primer pairs CazF_ko_P1 and CazF_ko_P2 were used for the upstream region, while CazF_ko_P3 and CazF_ko_P4 were used for the downstream region. For inactivation of *caz*M,

primer pairs CazM_ko_P1 with CazM_ko_P2 and CazM_ko_P3 with CazM_ko_P4 were used for constructing the upstream and downstream regions, respectively.

The disruption cassette for inactivation of *caz*E was also designed to contain the *hph* resistance marker and amplified from the pKW3202 cassette using primer pairs CazE_hphF and CazE_hphR. The homologous upstream and downstream regions (1 kb) were amplified from genomic DNA (gDNA) using primer pairs CazE_ko_P1 with CazE_ko_P2 and CazE_ko_P3 with CazE_ko_P4, respectively. For in vivo homologous recombination in yeast, the P1–P2 region, P3–P4 region and *hph* resistance cassette (50 ng to 150 ng each) were mixed with 40 ng of the linearized delivery vector pRS426[7] in a total volume of 45 μL. pRS426 was linearized by digesting with *Kpn* I and *Sac* I at 37 ºC for eight hours. The mixture was transformed into *S. cerevisiae* BY4741 and the four DNA fragments were joined *in situ* by the endogenous homologous recombination activity of *S. cerevisiae* through the 15-bp homologous sequences present in the DNA fragments. The desired transformants were selected on a uracil-deficient plate using the presence of the selection marker URA_3. The resulting plasmid pKW19053 containing the *cazE-hph* gene inactivation cassette was recovered from the yeast transformant and transferred to *E. coli*. pKW19053 was amplified in *E. coli* for subsequent characterization by restriction enzyme digestion and DNA sequencing to confirm its identity. The disruption cassette was amplified from pKW19053 using primer pair cazE_ko_P1 and cazE_ko_P4.

Transformation of disruption cassettes into C. globosum ΔligD: Linearized inactivation cassettes were transformed into *C. globosum* Δ*lig*D as described previously.[7] Transformants were grown on MYG agar supplemented with 200 μg/mL hygromycin B. To confirm correct integration of the disruption cassettes into the genome, gDNA was extracted from the

transformants and used as template for PCR using primer pairs CazE_mt_ver_P0 with CazE_mt_ver_P5, CazE_mt_ver_P6 with CazE_mt_ver_P7, CazF_mt ver_P0 with CazF_mt_ver_P5, CazF_mt_ver_P6 with CazF_mt_ver_P7, CazM_mt_ver_P0 with CazM_mt_ver_P5, and CazM_mt_ver_P6 with CazM_mt_ver_P7 (Figure S3).[8]

Protein expression and purification: *caz*E did not contain any introns and was amplified from *C. globosum* gDNA using primer pair CazE_F and CazE_R, which contained EcoRI and NotI restriction sites, respectively. The PCR product was digested with EcoRI and NotI and ligated into the pHis8 expression vector,[8] which was previously digested with the same restriction enzymes. The expression plasmid pJW07637 was transformed into *E. coli* BL21 (DE3) electro competent cells. A single colony was used to inoculate 10 mL Luria-Bertani (LB) media supplemented with 35 μg/mL kanamycin and grown overnight with constant shaking at 37 ºC. A 5 mL aliquot of the seed culture was used to inoculate 1 L of Terrific Broth (TB) (24 g yeast extract, 12 g tryptone and 4 mL glycerol) supplemented with 35 mg/L kanamycin. The cells were grown at 37 ºC to an OD600 of 0.4. To induce protein expression, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM and the culture was shaken at 16 ºC for 16 hours. The cells were harvested by centrifugation (3,750 rpm at 4 ºC for 20 mins) and the cell pellet was resuspended in 20 mL lysis buffer (50 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 500 mM NaCl, 5 mM imidazole, pH 7.9). The cells were sonicated on ice in 30 second intervals until homogenous. To remove cellular debris, the homogenous mixture was centrifuged at 17,000 rpm for 20 min at 4 ºC. Ni-NTA agarose resin was added to the supernatant (1 mL) and the solution was incubated at 4 ºC for 4 hours. Soluble CazE was purified by gravity-flow column chromatography using increasing concentrations of imidazole in

buffer A (50 mM Tris-HCl, 500 mM NaCl, 20 mM–250 mM imidazole, pH 7.9). Purified CazE was concentrated and buffer exchanged into Buffer B (50 mM Tris-HCl, 2 mM EDTA, 100 mM NaCl, pH 8.0) using an Ultracel 30,000 MWCO centrifugal filter (Amicon Inc.) and stored in 10% glycerol. The protein concentration was calculated to be 52 mg/L by Bradford assay using BSA as a standard.

The original DNA sequence of *caz*F contained a large unsequenced gap. Primer pair Nseq_07638fwd and Nseq_07638rev was used to amplify this region, which was subcloned into pCR®-blunt and sequenced (Figure S9).[19] To reconstitute an intron-free *caz*F, four overlapping pieces were constructed (Figure S10).[19] Piece two (2,585 bp) and piece four (3,194) did not contain any introns and were amplified from gDNA using primer pairs CazF_P2_F with CazF_P2_R and CazF_P4_F with CazF_P4_R, respectively. Piece one and three were initially annotated to contain introns and were thus amplified from cDNA. RNA was extracted from a five-day old PD culture using the RiboPure Yeast kit (Ambion) following the manufacturer's instructions and gDNA was digested with DNase (2 U/μL) (Invitrogen) at 37 ºC for four hours. Complementary DNA (cDNA) was synthesized from total RNA for piece one and three using SuperScript III First Strand Synthesis System with the reverse primers CazF_P1_R and CazF_P3_R, respectively. The cDNA was used as template for PCR and piece one and three were amplified using primer pairs CazF_P1_F with CazF_P1_R and CazF_P3_F with CazF_P3_R. The four exons were co-transformed with the 2 μm expression plasmid, which was previously linearized by digesting with Pml I and Nde I at 37 ºC overnight, in *S. cerevisiae* BJ5464-NpgA using an *S. c* EasyComp™ Transformation kit (Invitrogen) to yield the expression plasmid pJW07638. A single transformant was used to inoculate 3 mL SD_{ct} media (0.5 g Bacto Technical grade cassamino acids, 2 g Dextrose, and 88 mL Milli-Q water) supplemented with 1

mL adenine (40 mg/20 mL Milli-Q water), 1 mL tryptophan (40 mg/20 mL Milli-Q water) and 10 mL nitrogen base (1.7 g Nitrogen Base without amino acids, 5 g ammonium sulfate and 100 mL Milli-Q water) and grown for 72 hours with constant shaking at 28 ºC. A 1 mL aliquot of the seed culture was used to inoculate 1 L of YPD (10 g yeast extract, 20 g peptone and 950 mL Milli-Q water) supplemented with 1% dextrose and the culture was shaken at 28 ºC for 72 hours. The cells were harvested by centrifugation (3,750 rpm at 4 ºC for 10 mins) and the cell pellet was resuspended in 30 mL lysis buffer (50 mM NaH_2PO_4, 150 mM NaCl, 10 mM imidazole, pH 8.0). The cells were sonicated on ice in one-minute intervals until homogenous. To remove cellular debris, the homogenous mixture was centrifuged at 17,000 rpm for 1 hour at 4 ºC. Ni-NTA agarose resin was added to the supernatant (2 mL) and the solution was incubated at 4 ºC for 12–16 hours. Soluble CazF was purified by gravity-flow column chromatography using increasing concentrations of imidazole in buffer A (50 mM Tris-HCl, 500 mM NaCl, 20 mM–250 mM imidazole, pH 7.9). Purified protein was concentrated and buffer exchanged into Buffer B (50 mM Tris-HCl, 2 mM EDTA, 100 mM NaCl, pH 8.0) using an Ultracel 100,000 MWCO centrifugal filter (Amicon Inc.) and stored in 10% glycerol. The sequence of an intron-free *caz*F was confirmed using primers CazF_seq1–CazF_seq6 and its protein concentration was calculated to be 9.5 mg/L using the Bradford assay with BSA as a standard.

CazF in vitro activity assays: To assess the minimal activity of CazF, 25 μM of protein was incubated at room temperature with 2 mM malonyl-CoA in 100 mM phosphate buffer pH 7.4 in a total volume of 100 μL. After 16 hours, the reaction was base hydrolyzed by adding 20 μL 1M NaOH and incubated at 65 ºC for 10 minutes. The reaction was then extracted two times with 200 μL 99% ethyl acetate:1% acetic acid and the organic layer was dried by Speedvac. The

extract was then dissolved in 20 μL MeOH and analyzed on a Shimadzu 2010 EV LC-MS with a Phenomenex Luna 5μm 2.0 x 100 mM C18 column using positive and negative mode electrospray ionization with a linear gradient of 5–95% MeCN:H_2O over 30 minutes followed by 95% MeCN for 15 minutes and a flow rate of 0.1 mL/min. When the reaction was only extracted with ethyl acetate and not base hydrolyzed, lower levels of the triketide alpha-pyrone **5.13** was detected by LC-MS. However, when 25 μM of CazE was added to the reaction and extracted by ethyl acetate, comparable amounts of **5.13** were detected by LC-MS (Figure S12).[19]

The activity of the methyltransferase domain in CazF was analyzed by adding *S*-adenosyl-L-methionine chloride (SAM) to the minimal PKS activity assay. In a 100 μL reaction, 25 μM CazF was incubated with 2 mM SAM and 2 mM malonyl-CoA in 100 mM phosphate buffer pH 7.4. The reaction was incubated at room temperature for 16 hours, base hydrolyzed, extracted with 99% ethyl acetate:1% acetic acid, resuspended in 20 μL MeOH and analyzed by LC-MS as mentioned above.

The reduction activities of the ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains were analyzed by incubating 25 μM CazF with 2 mM NADPH, 2 mM acetoacetyl-SNAC and 100 mM phosphate buffer (pH 7.4) in a total volume of 100 μL. The reaction was incubated at room temperature for 12 hours, extracted two times with 200 μL ethyl acetate and the organic layer was dried by speedvac. The extract was resuspended in 20 μL MeOH and analyzed by LC-MS using the same conditions mentioned above (Figure S15).[8]

In vitro biosynthesis of chaetoviridin A (5.3) and 4'-desmethyl-chaetoviridin A (5.15) using CazF, CazE and cazisochromene (5.9): In a 100 μL reaction, 25 μM CazF was incubated with

2 mM NADPH, 2 mM SAM, 2 mM malonyl-CoA, 25 μM CazE and 2 mM **5.9** in 100 mM phosphate buffer pH 7.4. The reaction was incubated at room temperature for three hours and extracted two times with 200 μL ethyl acetate. The organic layer was dried by speedvac, resuspended in 20 μL MeOH and analyzed by LC-MS using the same conditions in the CazF in vitro activity assays. The in vitro biosynthesis of **5.3** was confirmed by comparing its retention time, UV profile and mass spectrometer analysis with an authentic standard.

For the production of **5.15**, the above reaction was repeated, except SAM was omitted from the reaction. The structure of **5.15** was proposed based on its retention time, UV profile and observed *m/z* to that of known chaetoviridins and chaetomugilins (Figure S16).[8]

5.9 Notes and References

(1) Hoffmeister, D.; Keller, N. P. *Nat. Prod. Rep.* **2007**, *24*, 393.

(2) Brase, S.; Encinas, A.; Keck, J.; Nising, C. F. *Chem. Rev.* **2009**, *109*, 3903.

(3) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.

(4) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965.

(5) Zabala, A.; Xu, W.; Chooi, Y. H.; Tang, Y. *Chem. Biol.* **2012**, *19*, 1049.

(6) a) Takahashi, M.; Koyama, K.; Natori, S. *Chem. Pharm. Bull.* **1990**, *38*, 625. (b) Kingsland,S. R.; Barrow, R. A. *Aust. J. Chem.* **2009**, *62*, 269. (c) Muroga, Y.; Takeshi, Y.; Atsushi, N.; Reiko, T. J. *Antibiot.* **2008**, *61*, 615. (d) Muroga, Y.; Yamada, T.; Numata, A.; Tanaka, R. *Tetrahedron* **2009**, *65*, 7580.

(7) a) Achard, M.; Beeler, A. B.; Porco, J. A., Jr. *ACS Comb. Sci.* **2012**, *4*, 236. b) Germain, A. R.; Bruggemeyer, D. M.; Zhu, J.; Genet, C.; O'Brien, P.; Porco, J. A., Jr. *J. Org. Chem.* **2011**, *76*, 2577. c) Musso, L.; Dallavalle, S.; Merlini, L.; Bava, A.; Nasini, G.; Penco, S.; Giannini, G.; Giommarelli, C.; De Cesare, A.; Zuco, V.; Vesci, L.; Pisano, C.; Castorina, M.; Milazzo, F.; Cervoni, M. L.; Dal Piaz, F.; De Tommasi, N.; Zunino, F. *Bioorg. Med. Chem.* **2010**, *18*, 6031.

(8) For supplementary figures, please refer to the original manuscript's Supplementary Information.

(9) a) Ogihara, J.; Kato, J.; Oishi, K.; Fujimoto, Y. J. *Biosci. Bioeng.* **2000**, *90*, 678. b) Hajjaj, H.; Klaebe, A.; Loret, M. O.; Goma, G.; Blanc, P. J.; Francois, J. *Appl. Environ. Microbiol.* **1999**, *65*, 311. c) Hajjaj, H.; Klaebe, A.; Goma, G.; Blanc, P. J.; Barbier, E.; Francois, *J.*

Appl. Environ. Microbiol. **2000**, *66*, 1120.

(10) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. *Nat. Protoc.* **2006**, *1*, 3111.

(11) Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 4530.

(12) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.

(13) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun.* **2007**, *39*, 4053.

(14) Davison, J.; al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. *J. Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 7642.

(15) Jez, J. M.; Ferrer, J. L.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Biochemistry* **2000**, *39*, 890.

(16) Lee, K. K.; Da Silva, N. A.; Kealey, J. T. *Anal. Biochem.* **2009**, *394*, 75.

(17) Ma, S. M.; Li, J. W.; Choi, J. W.; Zhou, H.; Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589.

(18) Yokoe, H.; Mitsuhashi, C.; Matsuoka, Y.; Yoshimura, T.; Yoshida, M.; Shishido, K. *J. Am. Chem. Soc.* **2011**, *133*, 8854.

(19) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.

CHAPTER SIX

Expanding the Structural Diversity of Polyketides by Exploring the Cofactor Tolerance of an In-line Methyltransferase Domain

Jaclyn M. Winter, Grace Chiou, Ian R. Bothwell, Wei Xu, Neil K. Garg, Minkui Luo, and Yi Tang.

Org. Lett. **2013**, *15*, 3774–3777.

6.1 Abstract

We report a strategy for introducing structural diversity into polyketides by exploiting the promiscuity of an in-line methyltransferase domain in a multidomain polyketide synthase. In vitro investigations using the highly-reducing fungal polyketide synthase CazF revealed that its methyltransferase domain accepts the non-natural cofactor propargylic *Se*-adenosyl-L-methionine and can transfer the propargyl moiety onto its growing polyketide chain. This propargylated polyketide product can then be further chain-extended and cyclized to form propargyl-α pyrone, or be processed fully into the alkyne-containing 4'-propargyl-chaetoviridin A.

6.2 Introduction

S-Adenosyl-L-methionine (SAM, **6.1**, Scheme 6.1)-dependent methyltransferases are a highly versatile class of enzymes that catalyze the methylation of DNA, RNA, protein and small molecules using the sulfonium methyl group of **6.1**.[1] Remarkably, these enzymes have been shown to accept synthetic SAM analogues and can catalyze the transalkylation of their corresponding proteins,[2] nucleic acids[3] and small molecules[4] with reactive bioorthogonal functional groups. In addition to freestanding methyltransferases, methyltranserase (MT)

domains can also be found within polyketide synthases (PKS) where they catalyze the α-methylation of a growing β-keto polyketide chain using **6.1** as the methyl-donating cofactor.[5]

Polyketide natural products are a structurally diverse class of molecules isolated from bacteria, fungi and plants and represent a rich source of medicinally important compounds.[6] The complex structures associated with polyketides arise from the utilization of simple building blocks, such as malonyl-CoA, and can undergo varying degrees of β-keto reduction during each catalytic cycle on the PKS.[7] As a result, enzyme-directed engineering approaches have been aimed at diversifying building block and starter unit selection as a way to enhance the structural variation of polyketides and to create analogues containing nonnative functional groups.[8] Recently, enzyme engineering efforts with bacterial acyltransferase (AT) domains[9] and acyl-CoA synthetases[10] have been successful at assimilating non-natural moieties into polyketides using synthetic malonic acid building blocks. While these synthetic biology strategies are promising for enhancing the structural diversity of polyketides, an alternative method for introducing structural variation can involve the use of non-natural cofactors. Thus, given the reported promiscuous nature of standalone methyltransferases, we wanted to investigate whether MT domains found within PKSs could behave in the same manner and facilitate the incorporation of different alkyl groups at the α-positions of polyketides.

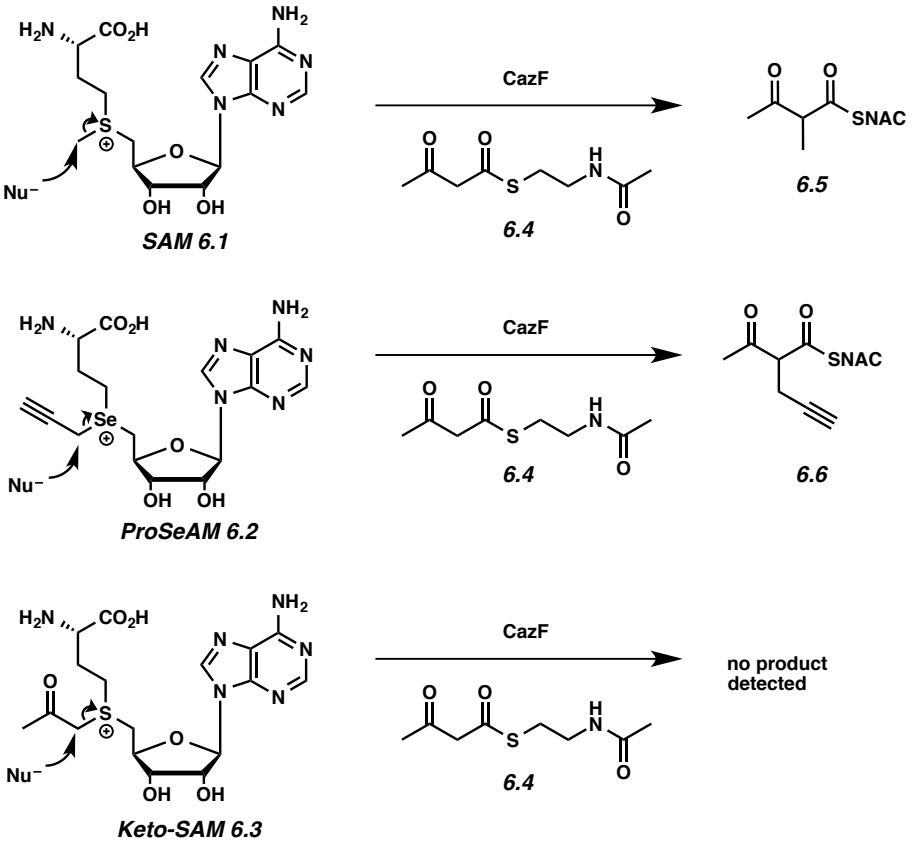
6.3 Determination of Substrate Specificity and Kinetic Assays of CazF

To examine the substrate specificity of a PKS MT domain, we turned our attention to the fungal highly-reducing PKS (HR-PKS) CazF, which is involved in the biosynthesis of the chaetoviridin and chaetomugilin azaphilones.[11] In vitro reconstitution experiments with recombinant CazF (284 kDa) expressed from *Saccharomyces cerevisiae* BJ5464-NpgA[12] verified

that the MT domain functions after the first chain-extending condensation step (Scheme 7.1), and the remaining domains are all catalytically active to generate a small polyketide product.[11] CazF therefore represented a good model system for exploring whether a MT domain can be used to introduce chemical diversity into polyketides. We selected the two unnatural SAM analogues propargylic *Se*-adenosyl-L-methionine (ProSeAM, **6.2**)[2d, 2h] and keto-SAM (**6.3**)[2d,13] for this study because both analogs can provide a reactive handle that can be further modified with mild, orthogonal chemistry. The selenium containing **6.2** was selected over the sulfur-based propargylic SAM derivative because it is more stable at physiological pH with a half-life of approximately one hour.[2d]

We first assayed the kinetics of the CazF MT domain towards the unnatural cofactors. Recombinant CazF was incubated with acetoacetyl-*S*-*N*-acetyl cysteamine (acetoacetyl-SNAC, **6.4**) in the presence of **6.1**, **6.2**, or **6.3** (Scheme 6.1). The production of alkylated acetoacetyl-SNAC products were analyzed by liquid chromatography and mass spectroscopy (LC-MS); and kinetic constants were calculated by fitting initial velocity data at various concentrations of **6.1** or **6.2** to Michaelis-Menten parameters using nonlinear least squares curve fitting. In the assays containing **6.3**, no product was detected suggesting that the MT domain would not accept the ketone derivative. The alkylated **6.5** and **6.6** were formed in the presence of **6.1** and **6.2**, respectively. The steady-state kinetic parameters were k_{cat} = 0.025 min^{-1} and K_M = 15.5 μM for **6.1**; and k_{cat} = 0.036 min^{-1} and K_M = 43.6 μM for **7.2**. Although the turnover rates were slow for both **6.1** and **6.2**, the comparable kinetic parameters indicate that the CazF MT displays similar preference towards **6.1** and **6.2**. The slow turnover rate may be attributed to the use of a small molecule SNAC-bound substrate instead of an acyl-carrier protein (ACP)-tethered molecule.

Scheme 6.1. CazF-Mediated Transalkylation of **6.4**.



6.4 Exploring the Cofactor Tolerance of CazF

After confirming that the MT domain would accept **6.2** and transfer the propargyl moiety onto **6.4**, we assessed whether the β-ketoacyl synthase (KS) domain of CazF would accept the propargylated-diketide intermediate and perform another decarboxylative Claisen condensation reaction with malonyl-CoA to form the propargylated triketide product. Previous in vitro characterization assays with CazF demonstrated that the enzyme can produce the triacetic acid lactone **6.7** when incubated with only malonyl-CoA, and the dimethylated α-pyrone **6.8** when SAM was included in the reaction (Figure 6.1).[11] To determine whether CazF could produce a propargyl-α-lactone, the enzyme was incubated with malonyl-CoA and 0.2 mM of **6.2**.

Following incubation, the reaction mixture was treated with 1 M NaOH (base hydrolysis) and the organic extract was analyzed by LC-MS. Production of the propargylated-α-pyrone **6.9** was observed verifying that the KS domain can indeed tolerate the bulkier propargyl substituent in the diketide and extend it to the corresponding triketide product (Figure 6.1).[14] To confirm the structure of **6.9** synthesized by CazF, a synthetic standard was prepared using a 4-step sequence (Scheme 2).[15] Alkylation of commercially available 2,4-pentanedione **6.19** with 3-bromoprop-1-yn-1-yl-trimethylsilane **6.20** provided the dione **6.10**, which in turn, was elaborated to the β,δ-diketoester **6.11** using sodium bis(trimethylsilyl)amide and dimethyl carbonate. A base-mediated cyclization of **6.11** with DBU yielded pyrone **7.12**, which underwent alkyne deprotection of **6.12** to furnish **6.9**.

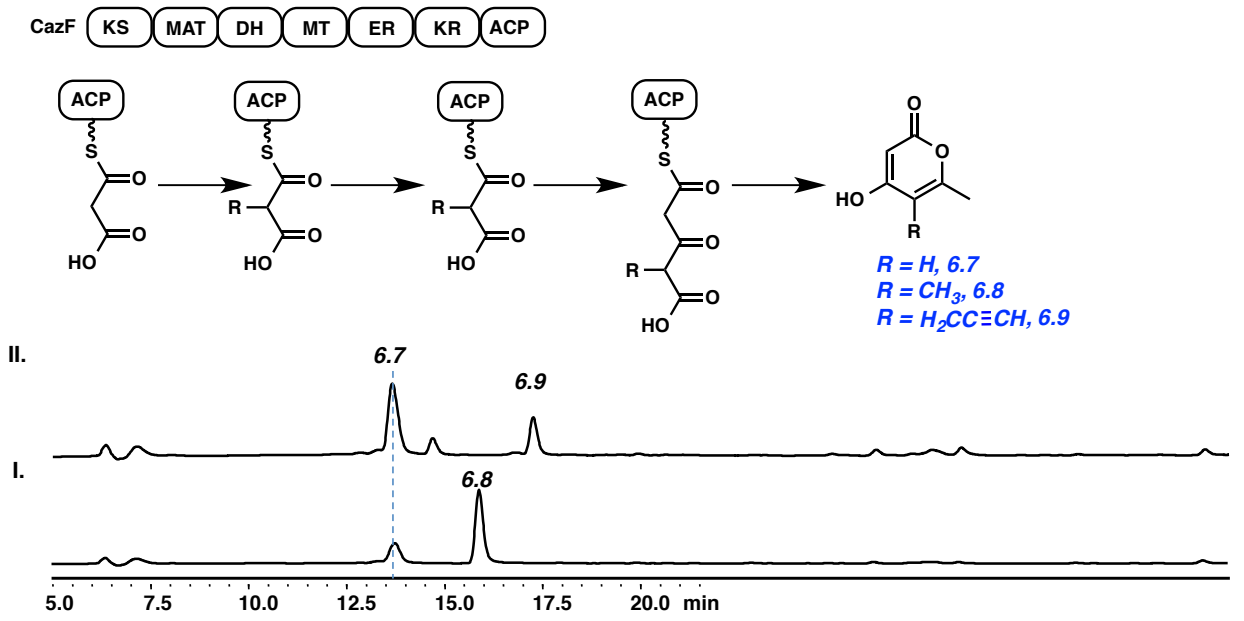
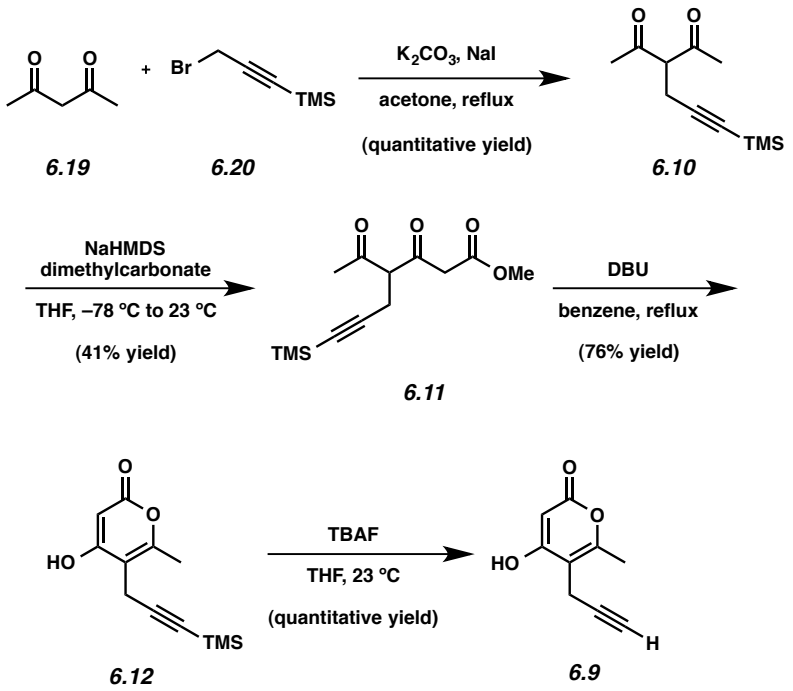


Figure 6.1. α-pyrones **6.7**–**6.9** biosynthesized by CazF in the presence of (i) malonyl-CoA and SAM **6.1** or (ii) malonyl-CoA and ProSeAM **6.2**. HPLC traces are shown in the same scale at λ= 280 nm. Domain organization of CazF consists of a ketosynthase (KS), malonyl-CoA:ACP

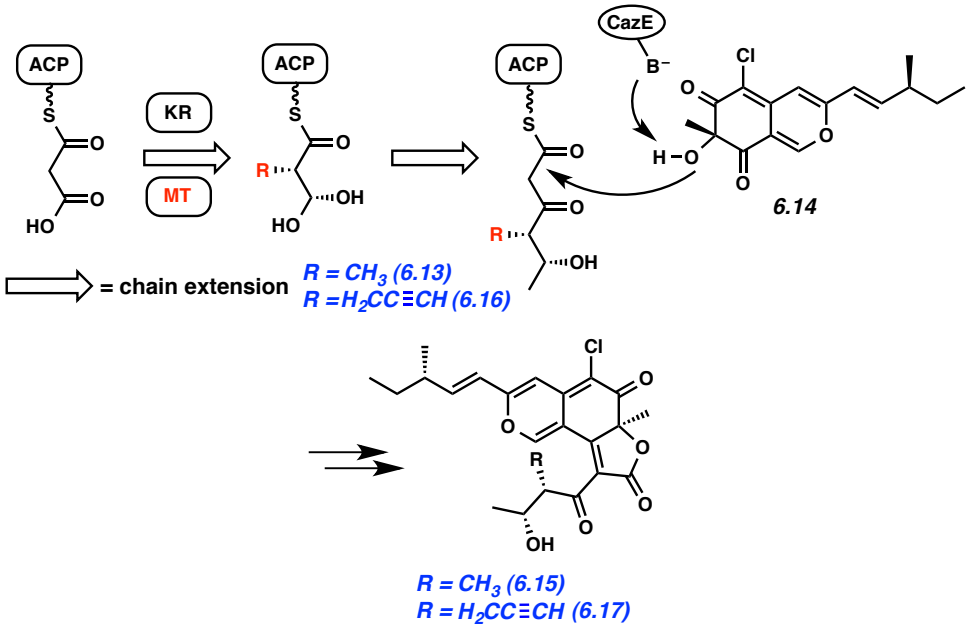
acyltransferase (MAT), dehydratase (DH), methyltransferase (MT), enoyl reductase (ER) ketoreductase (KR), and acyl carrier protein (ACP).

Scheme 6.2. Synthesis of **6.9**.



The higher amount of des-methylated compound **6.7** observed in the assay using **6.2** instead of **6.1** does suggest that under the reaction conditions, the rate of MT-alkylation is slower with the unnatural cofactor, and that the KS was able to outcompete the MT domain in extending the chain without alkylation. With the natural cofactor **6.1**, MT-catalyzed methylation is faster, which results in **6.8** as the dominant product. Increasing the concentration of **6.2** to 1 mM did not alter the ratio of **6.9** and **6.7**, which was expected since the K_M towards **6.2** was measured to be <50 mM.

Scheme 6.3. Biosynthesis of chaetoviridin A (**6.15**) and 4'-propargyl-chaetoviridin A (**6.17**) using the HR-PKS CazF, the acyltransferase CazE, and the *caz* pyrano-quinone intermediate cazisochromene (**6.14**) in the presence of the alkylating cofactor **6.1** or **6.2**.



Previous in vitro reconstitution experiments established that when **6.1**, along with NADPH used by the KR domain are present, the triketide product **6.13** is synthesized by CazF and is transacylated by the acyltransferase CazE to the pyrano-quinone intermediate cazisochromene (**6.14**) to form chaetoviridin A (**6.15**) (Scheme 6.3).[11] To further explore the tolerance of PKS domains, such as that of the KR towards the propargylated diketide and post-PKS accessory enzyme CazE towards the propargyl-containing polyketide intermediate **6.16**, we tested whether an alkyne-containing analog of **6.15**, such as **6.17**, can be synthesized using the same set of enzymes. Equimolar amounts of CazF and CazE were incubated with **6.14**, NADPH, malonyl-CoA and **6.2**. LC-MS analysis of the organic extract revealed the formation of a new

product with a shorter retention time than **6.15** and a $[M+H]^+$ m/z = 457 (Figure 6.2A) consistent with that of **6.17**. This new compound has an identical UV profile to **6.15** and an isotopic mass ratio consistent with the presence of one chlorine atom. Based on these observations, we assigned the product to be 4'propargyl-chaetoviridin A (**6.17**). To provide additional evidence toward the identify of the new product, a Cu(I)-catalyzed azide-alkyne Huisgen cycloaddition reaction[16] was performed to verify if **6.17** did indeed contain a terminal alkyne moiety. After reacting the organic extract containing **6.17** with azide-PEG_3-5(6)-carboxytetramethylrhodamine, the formation of a new peak, **6.18**, together with the disappearance of **6.17** was detected by LC-MS (Figure 6.2B). This new product exhibited UV absorption l_{max} at 550 nm, which is characteristic of the rhodamine dye; and a $[M+H]^+$ m/z = 1087, which is identical to the expected mass of the triazole-containing molecule **6.18** that forms as a result of alkyne mediated cycloaddition. Taken together, our experiments show that the propargyl functionality installed by the MT domain can indeed be propagated throughout the downstream reaction steps, and be found in the final natural product analog.

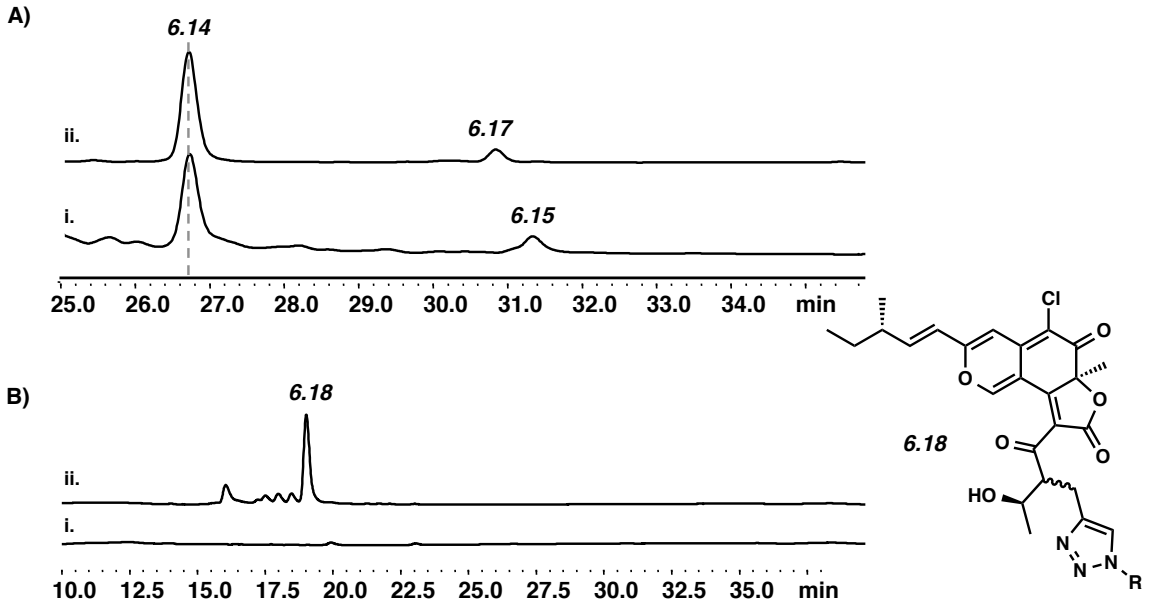


Figure 6.2. Analysis of the enzymatically synthesized chaetoviridins. A) HPLC analysis (360 nm) of chaetoviridins biosynthesized in vitro when (i) CazF and CazE were incubated with **6.14**, malonyl-CoA, NADPH and **7.1** and (ii) CazF and CazE were incubated with **6.14**, malonyl-CoA, NADPH and **6.2**. B) HPLC analysis (550 nM) of the triazole-containing **6.18** after labeling **6.17** with the rhodamine-azide.

6.5 Conclusion

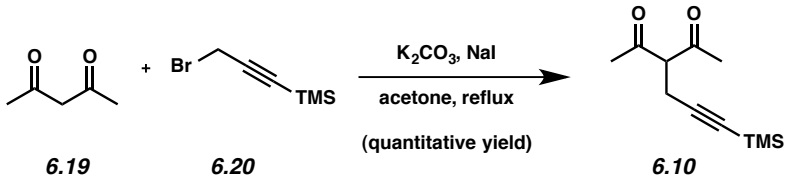
In summary, we demonstrated that the auxiliary MT domain within the highly-reducing fungal PKS CazF can act as a gateway for introducing orthogonally reactive functional groups into polyketides. Using the promiscuous nature of the MT domain, our study unveils an alternative approach for enhancing the structural diversity of natural products and opens the door for accessing and exploiting MT domains in other natural product systems.

6.6 Experimental Section

6.6.1 Materials and Methods

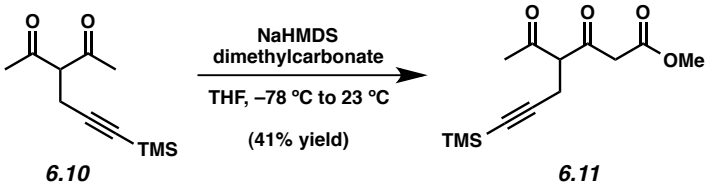
Unless stated otherwise, reactions were conducted in flame-dried glassware under an atmosphere of nitrogen using anhydrous solvents (either freshly distilled or passed through activated alumina columns). All commercially available reagents were used as received unless otherwise specified. Reaction temperatures were controlled using an IKAmag temperature modulator, and unless stated otherwise, reactions were performed at room temperature (rt, approximately 23 °C). Thin-layer chromatography (TLC) was conducted with EMD gel 60 F254 pre-coated plates (0.25 mm) and visualized using a combination of UV, anisaldehyde, iodine, and potassium permanganate staining. EMD silica gel 60 (particle size 0.040–0.063 mm) was used for flash column chromatography. All solvents and other chemicals used were of analytical grade. Azide-PEG_3-5(6)-carboxytetramethylrhodamine and the click chemistry protein reaction buffer kit was purchased from Click Chemistry Tools (Scottsdale, AZ). Acetoacetyl-*S*-*N*-acetyl cysteamine (acetoacetyl-SNAC) was a gift from Dr. Kangjian Qiao (UCLA). ProSeAM and Keto-SAM were prepared as previously described.[1] All ^{1}H NMR and ^{13}C NMR spectra were obtained on a 500 MHz Bruker AV500 spectrometer with a 5 mm dual cryoprobe. IR spectra were recorded on a Perkin-Elmer 100 spectrometer and are reported in terms of frequency absorption (cm^{-1}). High resolution mass spectra were obtained from the UC Irvine Mass Spectrometry Facility.

6.6.2 Experimental Procedures



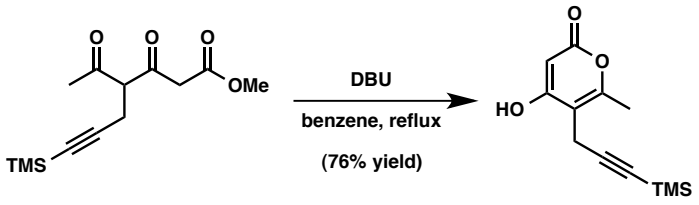
6.19 *6.20* *6.10*

Dione 6.10: To a flask containing 2,4-pentanedione **6.19** (18.5 mL, 80.0 mmol, 5.0 equiv) in acetone (32 mL) was added 3-bromoprop-1-yn-1-yl-trimethylsilane **6.20** (3.0 g, 16.0 mmol, 1.0 equiv). Potassium carbonate (11.0 g, 80.0 mmol, 5.0 equiv) was then added. The flask was topped with a reflux condenser and the system placed under N_2. The reaction was heated to reflux for 2 h. After cooling, the reaction was filtered and the filter cake was washed with acetone (3 x 20 mL). The filtrate was then concentrated under reduced pressure. The resultant oil was purified by flash chromatography (10:1 hexanes:EtOAc) to afford dione **6.10** (3.3 g, quantitative yield) as a pale yellow oil in a 1:1 mixture of keto:enol tautomers. R_f 0.8 (6:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) (Figure S1)[17]: Enol tautomer: δ 16.52 (s, 1H), 3.14 (s, 2H), 2.21 (s, 6H), 0.14 (s, 9H); Keto tautomer: δ 3.84 (t, J = 7.65 Hz, 1H), 2.71 (d, J = 7.65 Hz, 2H), 2.40 (s, 6H), 0.13 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 202.4, 190.9, 106.7, 103.8, 102.6, 87.6, 85.1, 66.7, 29.5, 23.2, 18.9, 18.8, 0.1, 0.01; IR (film): 2960, 2176, 1703, 1357, 1249 cm^{-1}; HRMS-ESI (m/z) $[M - H]^-$ calcd for $C_{11}H_{17}O_2Si$, 209.1076; found, 209.0998.



6.10 *6.11*

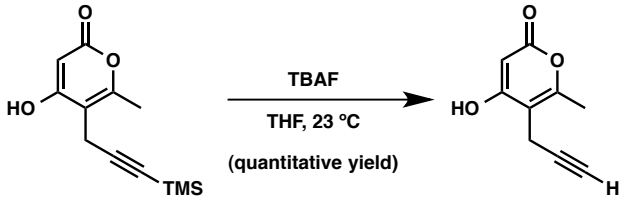
Methyl ester 6.11: To a flask containing a solution of **6.10** (198.0 mg, 0.942 mmol, 1.0 equiv) in THF (4.71 mL) at 78 ºC was added a freshly made 1M solution of NaHMDS in THF (2.82 mL, 2.82 mmol, 3.0 equiv). After stirring at –78 ºC for 1 h, dimethylcarbonate (95 mL, 1.03 mmol, 1.1 equiv) was added. The resulting mixture was removed from the cooling bath and warmed until the internal temperature of the reaction reached –10 ºC. The reaction was then quenched by the addition of 1 M HCl (10 mL). The reaction was transferred to a separatory funnel containing EtOAc (20 mL). The layers were separated, and the aqueous layer was extracted with EtOAc (3 x 10 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resultant yellow oil was purified by flash chromatography (10:1 hexanes:EtOAc) to afford methyl ester **6.11** (120.2 mg, 45% yield) as a clear oil. R_f 0.28 (6:1 hexanes:EtOAc); ^{1}H NMR (500 MHz, DMSO-*d6*) (Figure S2)[17]: δ 4.21 (t, J = 7.25, 1H), 3.77 (d, J = 16.9, 1H), 3.70 (d, J = 16.9, 1H), 3.63 (s, 3H), 2.64 (dd, J = 7.25, 1.95, 2H), 2.21 (s, 3H), 0.10 (s, 9H); ^{13}C NMR (125 MHz, DMSO-*d6*): δ 202.5, 198.6, 166.9, 104.0, 86.6, 63.45, 52.0, 48.9, 30.0, 18.1, -0.11); IR (film): 2958, 2177, 1727, 1707, 1626, 1249 cm^{-1}; HRMS-ESI (m/z) $[M + H]^+$ calcd for $C_{13}H_{21}O_4Si$, 269.1209; found 269.1201.



6.11 *6.12*

Pyrone 6.12: To a vial containing a solution of **6.11** (126.2 mg, 0.447 mmol, 1.0 equiv) in benzene (844 mL) was added DBU (134 mL, 0.894 mmol, 2.0 equiv) in one portion. The

resulting solution was stirred at 60 ºC for 3 h. After cooling, the reaction was quenched with saturated aqueous NH_4Cl (5 mL). The layers were separated and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine (50 mL), dried over $MgSO_4$, and concentrated under reduced pressure. The resultant solid was purified by flash chromatography (20:1 CH_2Cl_2:MeOH) to afford pyrone **6.12** (79.8 mg, 76% yield) as a white solid. Mp: 184–186 ºC; R_f 0.5 (10:1 CH_2Cl_2:MeOH); ^{1}H NMR (500 MHz, $CDCl_3$) (Figure S3)[17]: δ 5.64 (s, 1H), 3.36 (s, 2H), 2.34 (s, 3H), 0.13 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 171.2, 167.2, 161.6, 109.4, 102.2, 90.1, 85.5, 17.8, 14.9, 0.12; IR (film): 2958, 2176, 1727, 1645, 1608, 1560, 1248 cm^{-1}; HRMS-ESI (m/z) $[M + H]^+$ calcd for $C_{12}H_{17}O_3Si$, 237.0947; found, 237.0955.



6.12 *6.9*

Pyrone 6.9: To a vial containing a solution of **6.12** (16.6 mg, 0.058 mmol, 1.0 equiv) in THF (588 mL) was added a solution of TBAF (1.0M in THF, 88 mL, 0.088 mmol, 1.5 equiv). The reaction was stirred at 23 ºC for 14 h and then transferred to a separatory funnel containing EtOAc (10 mL) and 1 M HCl (25 mL). The layers were separated and the resulting aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine (40 mL), dried over $MgSO_4$, and concentrated under reduced pressure to afford the crude product. The resultant solid was purified by flash chromatography (2:1 EtOAc:hexanes) to afford **6.9** (9.2 mg, 97% yield) as a white solid. Mp: 209–210 ºC; R_f 0.25 (100% EtOAc); ^{1}H NMR (500 MHz, acetone-*d6*) (Figure S4)[17]: δ 10.70 (br s, 1H), 5.37 (s, 1H), 3.33 (d, J = 2.60, 2H), 2.42 (t, J

= 2.60, 1H), 2.28 (s, 3H); ^{13}C NMR (125 MHz, acetone-*d6*): δ 169.5, 163.6, 161.6, 108.1, 90.0, 81.6, 69.8, 17.5, 13.9; IR (film): 2959, 2924, 1706, 1674, 1561, 1259 cm^{-1}; HRMS-ESI (m/z) [M – H]$^-$ calcd for $C_9H_7O_3$, 163.0395; found, 163.0395.

CazF in vitro activity assays

CazF was expressed from *S. cerevisiae* BJ5464-NpgA[2] and purified by gravity-flow column chromatography as previously described.[3] The specificity of CazF's MT domain toward its natural SAM cofactor and unnatural analogs was analyzed by incubating 5 μM CazF with 2 mM acetoacetyl-SNAC (**6.4**) and 5 μM, 10 μM, 25 μM, 50 μM and 1 mM of SAM (**7.1**), ProSeAM (**6.2**), or Keto-SAM (**6.3**) in 100 mM phosphate buffer pH 7.4 in a total volume of 10 μL for one hour at room temperature. The reaction was quenched by adding 90 μL of MeOH and centrifuged for ten minutes at room temperature. A 10 μL aliquot was analyzed on a Shimadzu 2010 EV LC-MS with a Phenomenex Luna 5μ 2.0 x 100 mM C18 column using a linear gradient of 5–95% MeCN/H_2O containing 0.1% formic acid over 30 minutes with a flow rate of 0.1 mL/min. The corresponding *m*/z peaks for **6.5** and **6.6** were integrated and the amount of alkylated-SNAC product was quantified using a standard curve. Kinetic constants were calculated by fitting initial velocity data at various concentrations of **6.1** or **6.2** to Michaelis-Menten parameters using nonlinear least squares curve fitting in GraphPad Prism (Figure S5).[17]

To assess whether the KS domain of CazF would perform another round of elongation after the attachment of the alkyl group, 25 μM of protein was incubated at room temperature with 2 mM malonyl-CoA and 0.2 mM of **6.1** or **6.2** in 100 mM phosphate buffer pH 7.4 in a total volume of 100 μL. After 1 hour, the reaction was base hydrolyzed by adding 20 μL 1M NaOH,

incubated at 65 °C for 10 minutes, followed by addition of 40 μL 1N HCl. The reaction was then extracted two times with 200 μL of 99% ethyl acetate:1% acetic acid and the organic layer was dried using a Speedvac. The extract was resuspended in 20 μL MeOH and analyzed by LC-MS as mentioned above. The production of the propargyl-α-pyrone in the in vitro assay was confirmed using the synthetic standard **6.9** (Figure S6)[17] and the amounts of **6.8** or **6.9** produced in the in vitro assay were quantified from their respective standard curves. The in vitro production of **6.8** was previously confirmed.[3]

In vitro biosynthesis of 4'propargyl-chaetoviridin A (6.17) using CazF and CazE

In a 100 μL reaction, 25 μM CazF; 25 μM CazE, which was expressed and purified as previously described;[3] 2 mM NADPH; 2 mM malonyl-CoA; and 2 mM cazisochromene (**6.14**) were incubated with **6.1** or **6.2** in 100 mM phosphate buffer pH 7.4 for one hour at room temperature. Due to the instability of **6.1** and **6.2** at neutral conditions and at room temperature, 1 mM of **6.1** and **6.2** were added at the start of the reaction and after 30 minutes. The final concentration of the cofactors was 2 mM. The reaction was extracted twice with 300 μL ethyl acetate containing 1% acetic acid. The organic layer was dried using a Speedvac and resuspended in 20 μL MeOH. The extract was analyzed by LC-MS using the same conditions as in the CazF in vitro activity assays. The in vitro biosynthesis of chaetoviridin A (**6.15**) using **6.1** as the alkyl donor was confirmed previously.[3] The structure of 4'-propargyl-chaetoviridin A (**6.17**), which was formed using **6.2** as the alkyl donor, was proposed based on its retention time, UV profile and observed *m/z* isotopic ratio.

Copper-catalyzed azide-alkyne cycloaddition reaction

The in vitro reaction for the enzymatic synthesis of **6.17** was repeated using the same conditions mentioned above. After incubating at room temperature for one hour, the reaction was extracted twice with 300 μL ethyl acetate containing 1% acetic acid and the organic layer was dried using a Speedvac. The crude extract was then resuspended in 50 μL MeOH. Azide-PEG3-5(6)-carboxytetramethylrhodamine was resuspended in DMSO to a final concentration of 2.5 mM and the click labeling reaction was carried out using 2 μL of the 2.5 mM rhodamine-azide according to the manufacturer's instructions. After a four hour incubation at room temperature, the reaction was extracted twice with 250 μL ethyl-acetate containing 1% acetic acid and the organic layer was dried using a Speedvac. The extract was resuspended in 20 μL MeOH and analyzed by LC-MS using the same conditions as mentioned previously in the CazF in vitro activity assay.

6.7 Notes and References

(1) Struck, A. W.; Thompson, M. L.; Wong, L.; Micklefield, J. *ChemBioChem*. **2012**, *13*, 2642.

(2) a) Peters, W.; Willnow, S.; Duisken, M.; Kleine, H.; Macherey, T.; Duncan, K. E.; Litchfield, D. W.; Lüscher, B.; Weinhold, E. *Angew. Chem. Int. Ed. Engl.* **2010**, *49,* 5170. b) Islam, K.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *ACS Chem. Biol.* **2011**, *6*, 679. c) Wang, R.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2011**, *133*, 7648. d) Bothwell, I. R.; Islam, K.; Chen, Y.; Zheng, W.; Blum, G.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2012**, *134*, 14905. e) Binda, O.; Boyce, M.; Rush, J. S.; Palaniappan, K. K.; Bertozzi, C. R.; Gozani, O. *ChemBioChem*. **2011**, *12*, 330. f) Wang, R.; Islam, K.; Liu, Y.; Zheng, W.; Tang, H.; Lailler, N.; Blum, G.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2013**, *135*, 1048. g) Blum, G.; Bothwell, I. R.; Islam, K.; Luo, M. *Curr. Prot. Chem. Biol.* **2011**, *7*, 2970. h) Willnow, S.; Martin, M.; Luscher, B.; Weinhold, E. *ChemBioChem*. **2012**, *13*, 1167.

(3) a) Dalhoff, C.; Lukinavičius, G.; Klimašauskas, S.; Weinhold, E. *Nat. Chem. Biol.* **2006**, *2*, 31. b) Dalhoff, C.; Lukinavičius, G.; Klimašauskas, S.; Weinhold, E. *Nat. Protoc.* **2006**, *1*, 1879.

(4) Stecher, H.; Tengg, M.; Ueberbacher, B. J.; Remler, P.; Schwab, H.; Griengl, H.; Gruber-Khadjawi, M. *Angew. Chem. Int. Ed.* **2009**, *48*, 9546.

(5) Ansari, M. Z.; Sharma, J.; Gokhale, R. S.; Mohanty, D. *BMC Bioinformatics.* **2008**, *9*, 454.

(6) O'Hagan, D. *The Polyketide Metabolites*; Ellis Howard: Chichester, UK, 1991.

(7) a) Fischbach, M. A.; Walsh, C. T. *Chem. Rev.* **2006**, *106*, 3468. (b) Hertweck, C. *Angew. Chem. Int. Ed.* **2009**, *48,* 4688.

(8) a) Pickens, L. B.; Tang, Y.; Chooi, Y. H. *Annu. Rev. Chem. Biomol.* **2011**, *2*, 211. b) Winter,

J. M.; Tang, Y. *Curr. Opin. Biotech.* **2012**, *5*, 736. c) Wilson, M. C.; Moore, B. S. *Nat. Prod. Rep.* **2012**, *29*, 72. d) Kennedy, J. *Nat. Prod. Rep.* **2008**, *25*, 25.

(9) a) Sundermann, U.; Bravo-Rodriguez, K.; Klopries, S.; Kushnir, S.; Gomez, H.; Sanchez-Garcia, E.; Schulz, F. *ACS Chem. Biol.* **2013**, *8*, 443. b) Hughes, A. J.; Detelich, J. F.; Keatinge-Clay, A. T. *Med. Chem. Commun.* **2012**, *3*, 956.

(10) Koryakina, I.; McArthur, J.; Randall, S.; Draelos, M. M.; Musiol, E. M.; Muddiman, D. C.; Weber, T.; Williams, G. J. *ACS Chem. Biol.* **2013**, *8*. 200.

(11) Winter, J. M.; Sato, M.; Sugimoto, S.; Chiou, G.; Garg, N. K.; Tang, Y.; Watanabe, K. *J. Am. Chem. Soc.* **2012**, *134*, 17900.

(12) Lee, K. K.; Da Silva, N. A.; Kealey, J. T. *Anal. Biochem*. **2009**, *394*, 75.

(13) Lee, B. W. K.; Sun, H. G.; Zang, T.; Kim, B. J.; Alfaro, J. F.; Zhou, Z. S. *J. Am. Chem. Soc.* **2010**, *132*, 3642.

(14) In Figure 7.1 trace ii, there is a additional peak at 14.5 min. Based on its *m/z,* this compound is most likely propargylic *Se*-adenosyl.

(15) For specific details on the synthetic reactions, please see the supporting information.

(16) Huisgen, R. *1,3-Dipolar Cycloaddition Chemistry;* Padwa, A., Ed.; Wiley: New York, 1984.

(17) For supplementary figures, please refer to the original manuscript's Supplementary Information.

APPENDIX FOUR

Spectra Relevant to Chapter Six

Expanding the Structural Diversity of Polyketides by Exploring the Cofactor Tolerance of an In-line Methyltransferase Domain

Jaclyn M. Winter, Grace Chiou, Ian R. Bothwell, Wei Xu, Neil K. Garg, Minkui Luo, and Yi Tang.

Org. Lett. **2013**, *15*, 3774–3777.

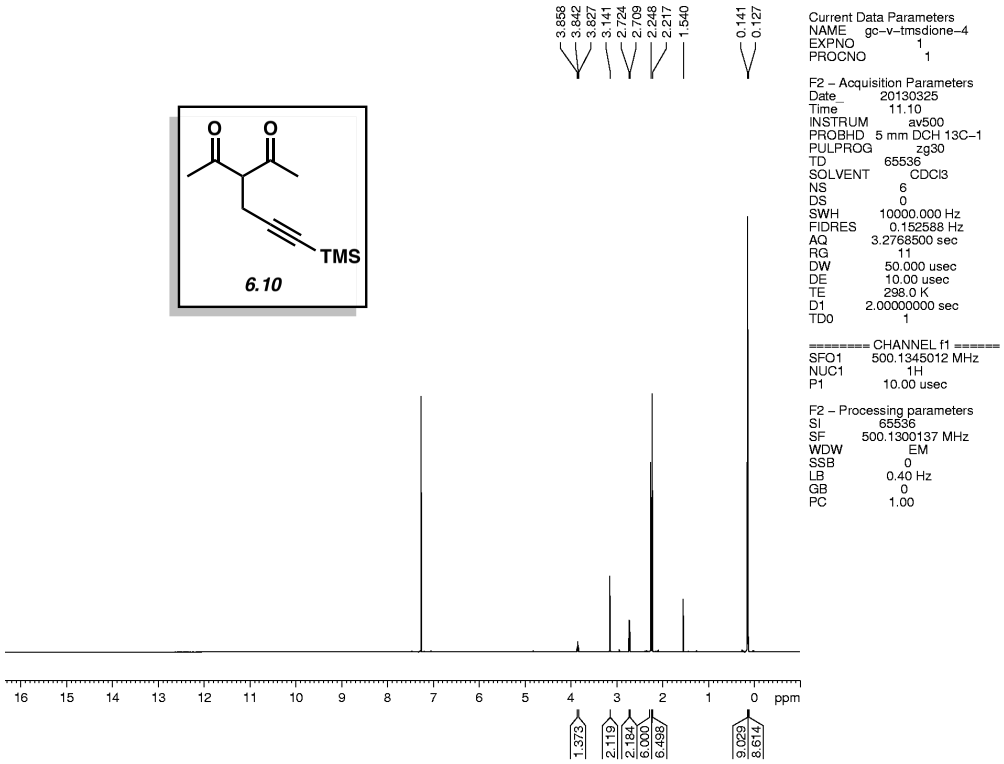


Figure A4.1 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **6.10**.

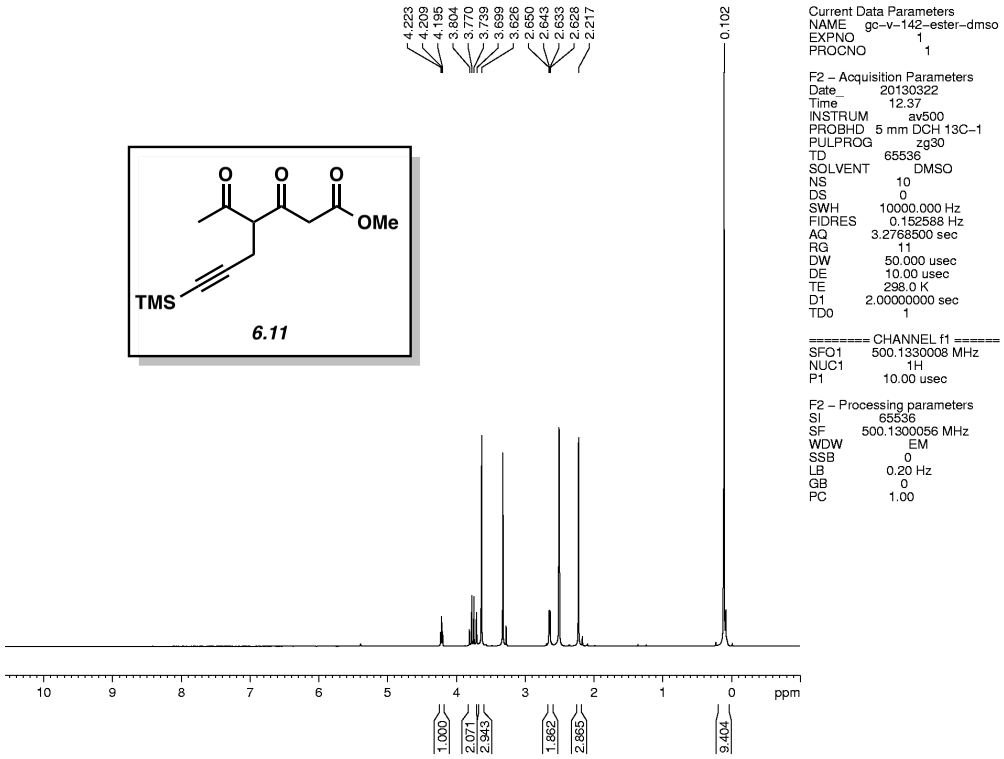


Figure A4.2 ^{1}H NMR (500 MHz, DMSO) of compound **6.11**.

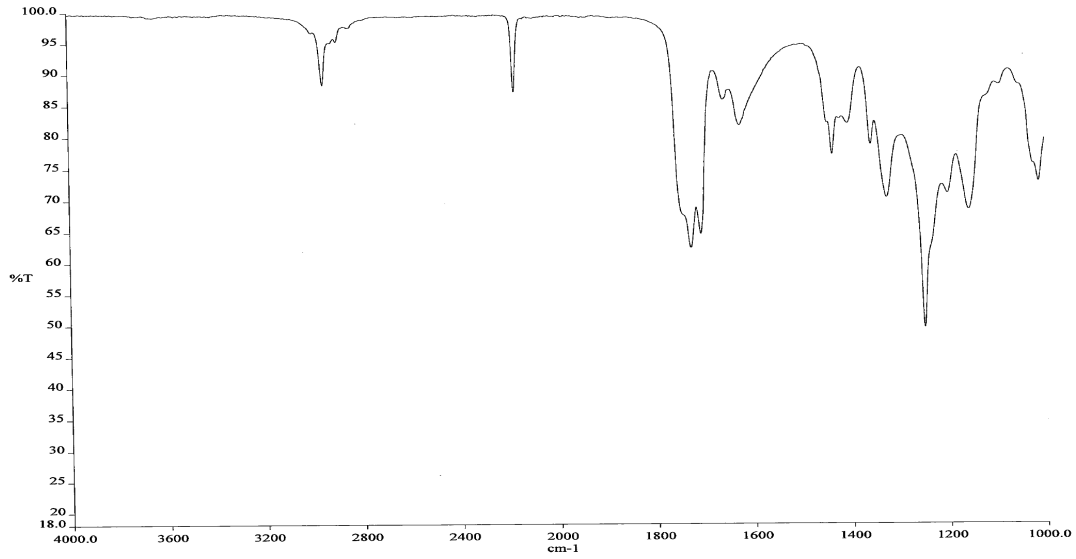


Figure A4.3 Infrared spectrum of compound **6.11**.

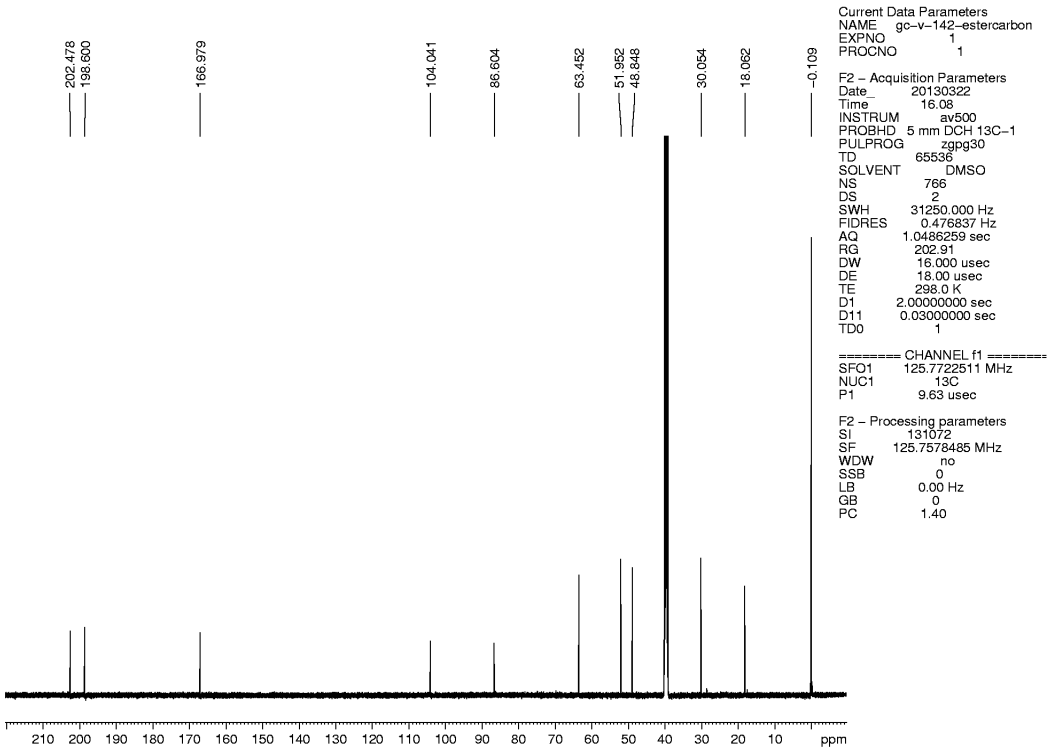


Figure A4.4 ^{13}C NMR (500 MHz, DMSO) of compound **6.11**.

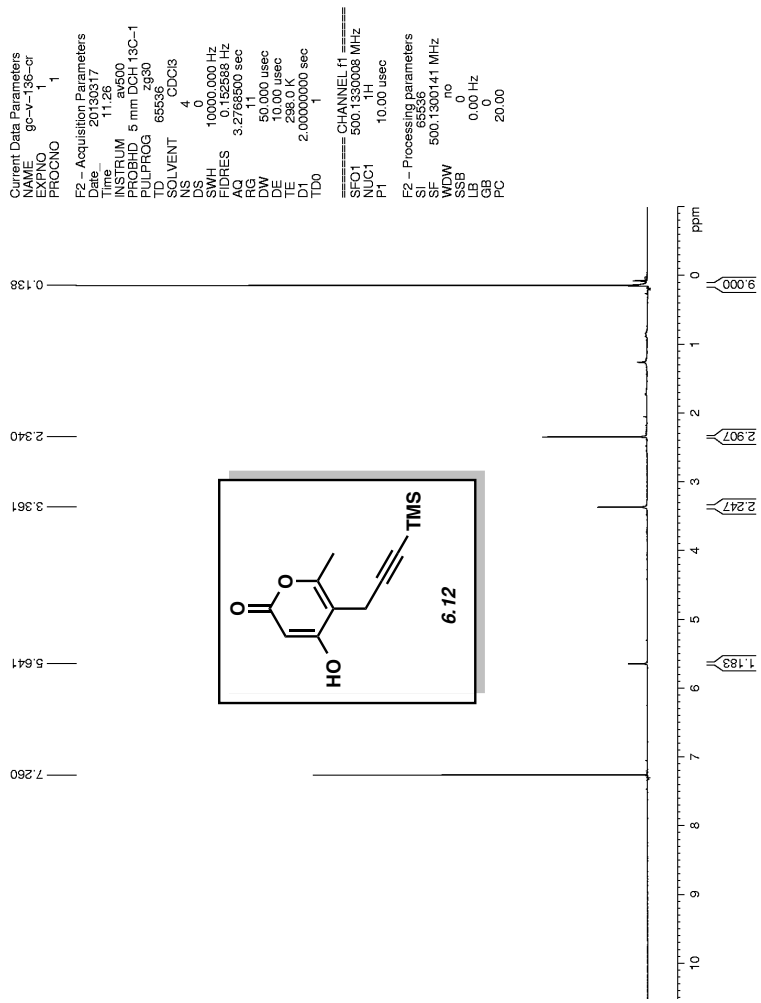


Figure A4.5 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **6.12**.

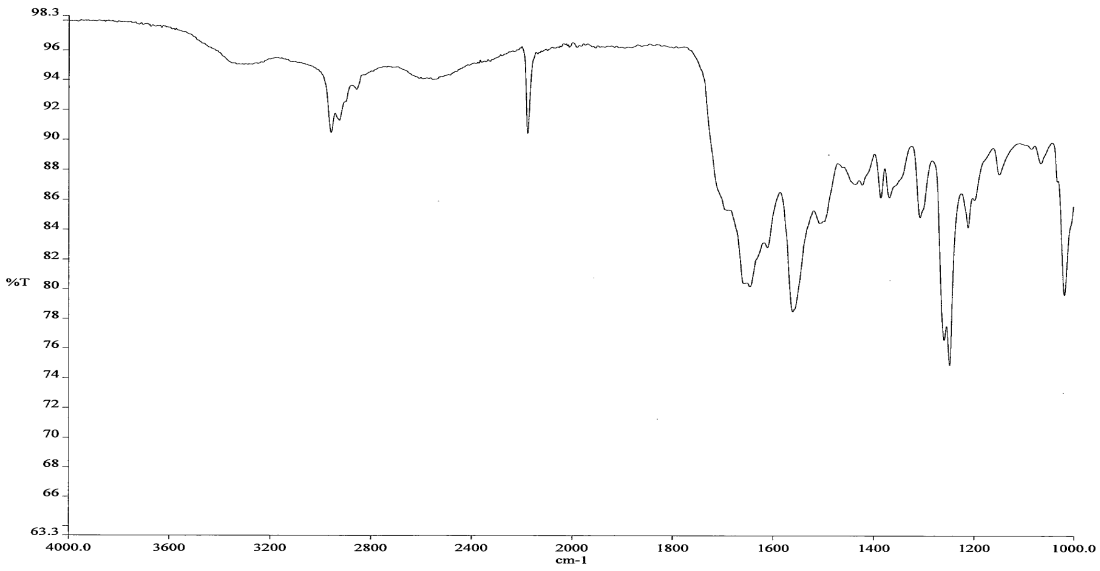


Figure A4.6 Infrared spectrum of compound **6.12**.

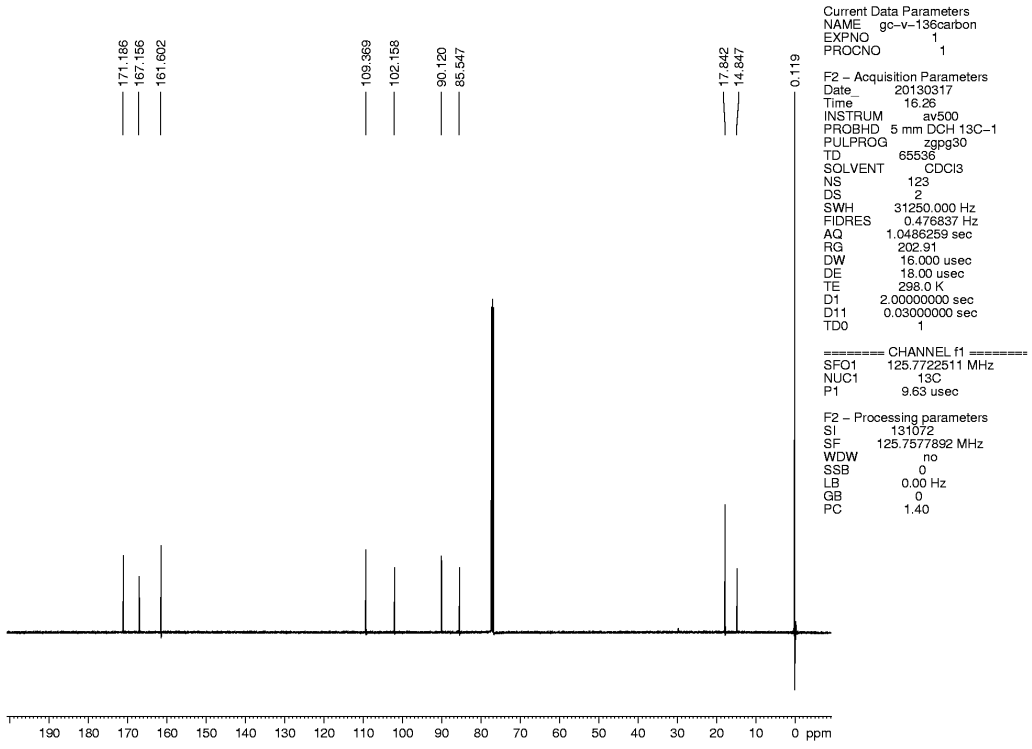


Figure A4.7 ^{13}C NMR (500 MHz, $CDCl_3$) of compound **6.12**.

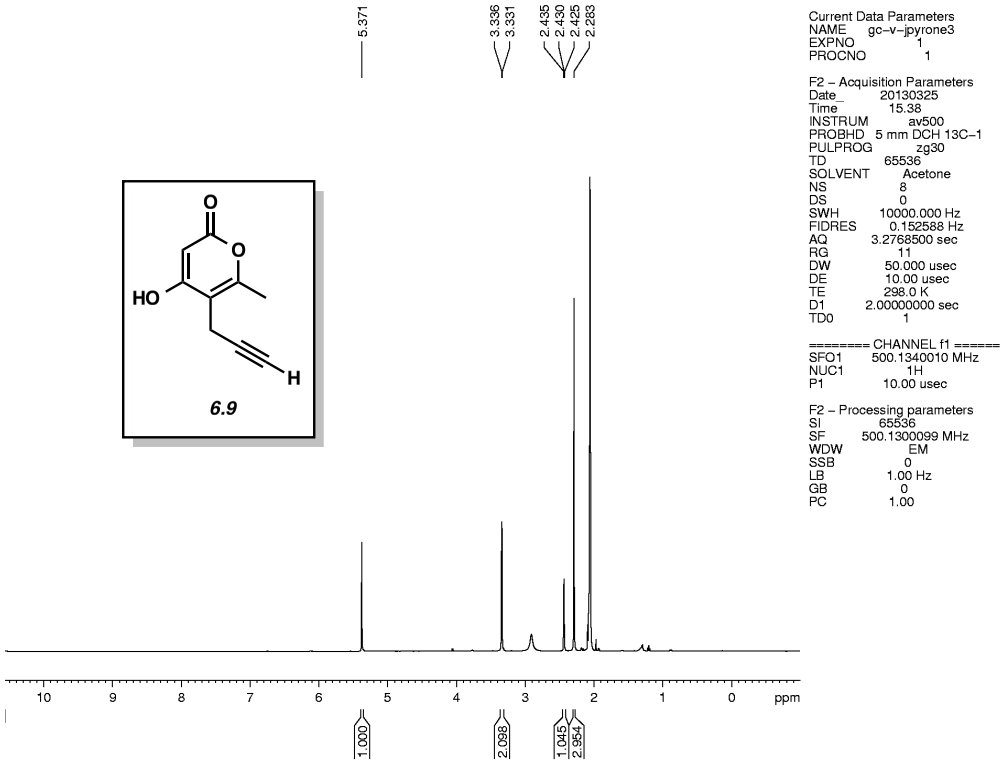


Figure A4.8 ^{1}H NMR (500 MHz, acetone) of compound **6.9**.

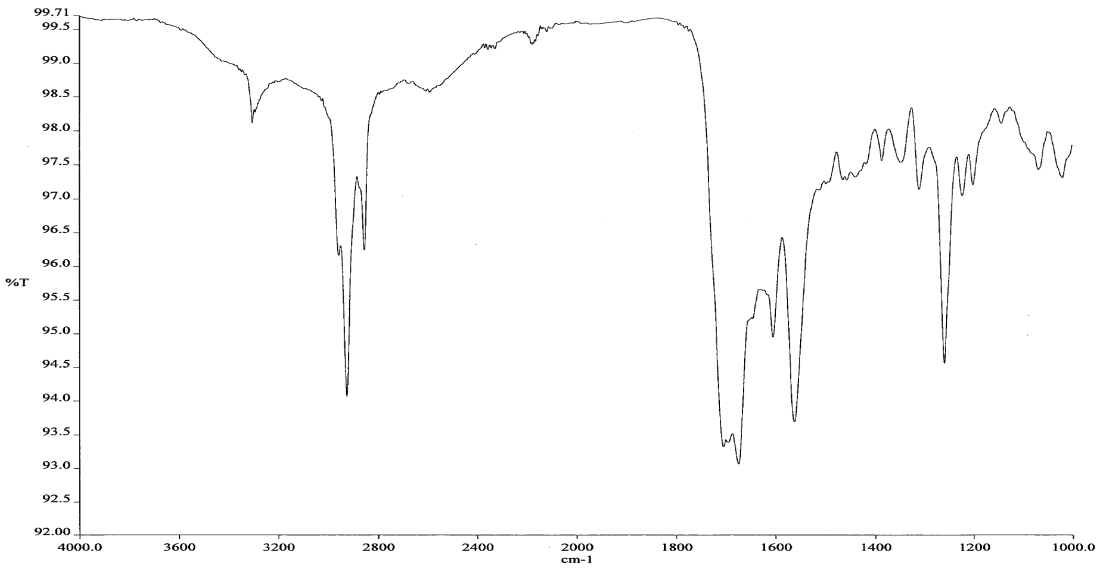


Figure A4.9 Infrared spectrum of compound **6.9**.

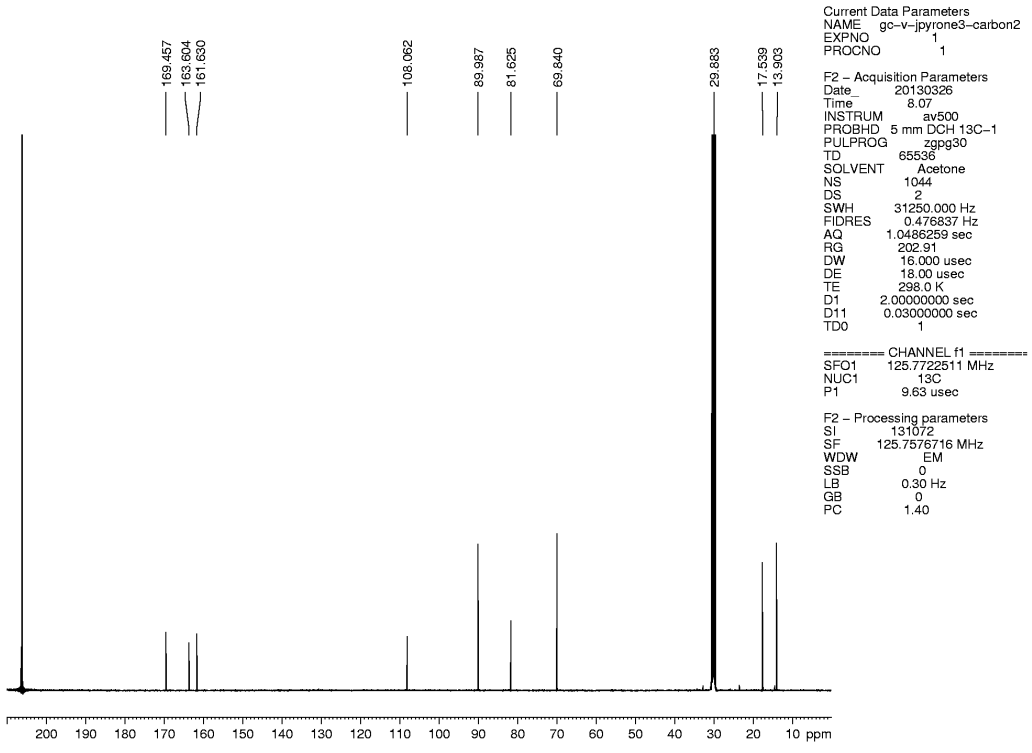


Figure A4.10 ^{13}C NMR (500 MHz, acetone) of compound **6.9**.