Title

Structural Elucidation of Metabolites via Electron Diffraction at the Interface of Synthetic Chemistry and Biology

Permalink

https://escholarship.org/uc/item/1h5577w9

Author

Kim, Lee Joon

Publication Date

2022

UNIVERSITY OF CALIFORNIA

Los Angeles

Structural Elucidation of Metabolites *via* Electron Diffraction at the Interface of Synthetic Chemistry and Biology

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

by

Lee Joon Kim

2022

ABSTRACT OF THE DISSERTATION

Structural Elucidation of Metabolites *via* Electron Diffraction at the Interface of Synthetic Chemistry and Biology

by

Lee Joon Kim
Doctor of Philosophy in Chemistry
University of California, Los Angeles, 2022
Professor Hosea Martin Nelson, Co-Chair
Professor Yi Tang, Co-Chair

This dissertation describes the application and development of electron crystallography in the realm of small molecules and, in particular, natural products (NPs). Natural products serve as a major inspiration for drug discovery, but the rate of discovering novel NPs is hindered by limitations in existing characterization techniques. This dissertation highlights microcrystal electron diffraction (MicroED) as a powerful approach to overcome these limitations and structurally characterize NPs and other small molecules. Overall, this work demonstrates the general applicability of MicroED in natural product and small molecule characterization. Moreover, using MicroED as a platform, an efficient workflow for NP identification is developed, with the goal of accelerating the rate of novel natural product discovery as well as drug discovery.

Chapter One is a brief overview of the current state of research on using electron diffraction to elucidate structures of natural products. This chapter serves as both a prelude to and a summary of the remaining chapters, and will be referenced throughout this dissertation. Brief history of electron diffraction and background of MicroED is provided, focusing on the scope of MicroED in small molecule regime. Furthermore, the application of MicroED in natural products and currently remaining challenges are highlighted, drawing on our experiences as well as others'.

Chapters Two through Five describe the applications of MicroED in NP characterization. Chapter Two highlights the structural elucidation of a glucosyluric metabolite in *C. elegans*, whose MicroED structure was used to complement the NMR data that was ambiguous due to poor line shapes. Chapter Three showcases the structure determination of a biosynthetic intermediate to a fungal metabolite (–)-sambutoxin, whose relative stereochemistry was validated using MicroED. Chapter Four discusses the synergy between MicroED and genome mining by rediscovering the biosynthetic pathway to fischerin and revising its relative stereochemistry that was incorrectly assigned using computational studies. In this chapter, a novel compound is isolated and characterized using MicroED, demonstrating the first example of a novel, non-peptidyl NP structurally elucidated *via* MicroED. Additionally, this study highlights MicroED's impressive sensitivity by detecting structural information from only 3 ng of material. Chapter Five reports the structural revision of the lomaiviticins enabled by MicroED analysis. The lomaiviticins are highly cytotoxic bacterial metabolites that have been extensively studied for decades, with multiple total synthesis efforts that failed. This structural revision, supported by high-field NMR as well as DFT calculations, enables scientists to embark on future synthetic efforts and studies on mechanisms of action with more confidence.

Chapter Six highlights the development of a MicroED platform to enable high-throughput screening for crystalline NPs. Initial studies demonstrate MicroED’s ability to obtain crystal structures in complex mixtures of natural products. These preliminary experiments are used to envision a workflow coupling liquid chromatography and MicroED, which is further developed into a high-throughput strategy. Using this platform, NPs produced in insufficient amounts for other analytical techniques can be detected and characterized, accelerating the rate of NP discovery.

Chapter Seven describes the expansion of the scope of MicroED further in the small molecule regime by solving structures of reactive organometallic species and synthetic reaction products. These examples showcase the ability of MicroED in obtaining crystal structures that were previously unattainable, and demonstrate the structural integrity of MicroED structures when compared with structures obtained by X-ray crystallography. Notably, we discuss the utility of room temperature MicroED experiments, which provide a facile and quick method of screening for diffraction and become an integral part of the workflow in structural elucidation of small molecules.

The dissertation of Lee Joon Kim is approved.

Neil Kamal Garg

José Alfonso Rodriguez

Yi Tang, Committee Co-Chair

Hosea Martin Nelson, Committee Co-Chair

University of California, Los Angeles

2022

This dissertation is dedicated to my parents and younger sister Rachel

for their unconditional love, trust and patience.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

LIST OF SCHEMES

CHAPTER TWO

CHAPTER THREE

CHAPTER FIVE

LIST OF TABLES

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

LIST OF ABBREVIATIONS

Å = ångstrom(s)

Ac = acetyl, acetate

AcOH = acetic acid

α = alpha

app. = apparent

aq. = aqueous

Ar = aryl

β = beta

br = broad

Bu = butyl

t-Bu = *tert*-butyl

°C = degrees Celsius

calc’d = calculated

cat. = catalytic

CC = correlation coefficient

CCDC = Cambridge Crystallographic Data Centre

CIF = crystallographic information file

Cl_{11} = $[CHB_{11}Cl_{11}]^{-}$

CMSO = complementary metal oxide semiconductor

COSY = correlated spectroscopy

cRED = continuous rotation electron diffraction

CSD = Cambridge Structural Database

cryoEM = cryoelectron microscopy

d = doublet

dd = doublet of double

dr = diastereomeric ratio

δ = chemical shift

DAD = diode array detection

DCB = dichlorobenzene

DFB = difluorobenzene

DMF = dimethylformamide

DMSO = dimethylsulfoxide

ED = electron diffraction

equiv = equivalent

EM = electron microscopy

ESI = electrospray ionization

Et = ethyl

F_{15} = tris(pentafluorophenyl)borane

F_{20} = tetrakis(pentafluorophenyl)borate

g = gram(s)

γ = gamma

GCMS = gas chromatography-mass spectrometry

GooF = goodness of fit

h = hour(s)

HMDS = hexamethyldisilazane

HPLC = high performance liquid chromatography

HRMS = high resolution mass spectroscopy

Hz = Hertz

IR = infrared spectroscopy

i-Pr = isopropyl

J = coupling constant

K = kelvin

L = liter

LCMS = liquid chromatography-mass spectrometry

Li = lithium

LDA = lithium diisopropylamide

m = multiplet

m = meta

M = molar

m/z = mass to charge ratio

μ = micro

Me = methyl

MeCN = acetonitrile

MeOH = methanol

MHz = megahertz

MicroED = microcrystal electron diffraction

min = minutes

mol = mole(s)

MOM = methoxymethyl ether

mp = melting point

MR = molecular replacement

MS = mass spectrometry

n = nano

NMR = nuclear magnetic resonance

NOESY = Nuclear Overhauser Effect spectroscopy

NP = natural product

o = ortho

ORTEP = Oak Ridge Thermal-Ellipsoid Plot Program

OTf = trifluoromethanesulfonate

OTs = *p*-toluenesulfonate

p = pico

p = para

PDC = piezo dispense capillary

Ph = phenyl

pin = pinacolato

ppm = parts per million

Pr = propyl

q = quartet

RMS = root mean square

rt = room temperature

s = singlet

SI = supplementary information

t = triplet

TEM = transmission electron microscope/microscopy

temp = temperature

TES = triethylsilyl

Tf = trifluoromethanesulfonyl

TFA = trifluoroacetic acid

THF = tetrahydrofuran

TIPS = triisopropylsilyl

TLC = thin layer chromatography

TMS = trimethylsilyl

Ts = *p*-toluenesulfonyl (tosyl)

TS = transition state

UPLC = ultra performance liquid chromatography

UV = ultraviolet

XRD = X-ray diffraction

YM = yeast malt

ACKNOWLEDGEMENTS

As a 1.5-generation Korean-American immigrant and the first person in my family to experience the education system in the United States, I have many people to acknowledge for guiding and supporting me throughout the years.

I would like to begin by thanking my advisor Professor Hosea Nelson. I still remember the first group meeting that I attended as a rotation student: while I could not understand most of the conversations about the C–H functionalization project, I was amazed by the originality of the DNA catalysis project and decided to join the lab to perform research on fun ideas that not many others in the field are working on, or are even capable of imagining. Working in the Nelson lab has exposed me to vastly different fields within chemistry, and it has been an incredible journey witnessing the development of diverse, creative research projects in the group. I'm grateful for the unique experiences I have had as part of the MicroED research subgroup, such as traveling to Oregon to work with folks from Thermo Fisher Scientific, Inc. and working with the American Chemical Society in helping to write a course on the technique. Thank you, Hosea, for letting me work on exciting collaborations both interally and externally, from which I was able to establish meaningful connections with scientists from various institutions. I also appreciated that you checked on my health from time to time after I had come back from my leave of absence in the summer of 2018, and I'm glad I was able to keep up my pace since then and finish my graduate studies.

I would also like to thank the rest of my doctoral committee members: Professors Yi Tang, José Rodriguez, and Neil Garg. They have all provided me with excellent advice and mentorship, and demonstrated exemplary professionalism and empathy during the course of my graduate studies. I'm extremely grateful for the tremendous support from Prof. Tang, who is the co-chair of my committee as well as my research collaborator of many years. With his permission, I was able to work with students and postdocs in his lab to learn molecular biology techniques, which not only expanded my skill sets but also enabled me to make a more informed decision about switching fields from chemistry to biology. I would like to thank him for always being open to meet with me, for the kind words, and for asking casual questions about my gaming hobbies that made it so comfortable for me to chat with him. Prof. Rodriguez has been an amazing instructor, mentor and a collaborator. Without him, my early days as an electron microscopist would have been extremely difficult, and I am thankful for his patience and thoroughness in his teachings. From nearly every conversation or interaction I have with him, I have been able to walk away with new knowledge, and his insight has been absolutely crucial on many occasions throughout the years. Prof. Garg has been a patient, supportive committee member as well as the department chair. From volunteering with him and his lab at a local elementary school to talking with him in his office to seek advice, Prof. Garg has always left a kind and caring impression on me. I consider myself incredibly fortunate to have these professors in my doctoral committee.

Next, I would like to thank the members of the UCLA–DOE institute, particularly Dr. Duilio Cascio and Dr. Michael "Mikey" Sawaya. Duilio has been a parental figure to me with his wisdom and big heart. Some of my fondest memories while in graduate school involve working late at night in Duilio's office, collecting and processing synchrotron data with Duilio and Mikey with all sorts of music playing in the background. I have learned so much from them, from crystallography to music and more.

I'm also thankful for my many academic collaborators who I've had amazing experiences working with, including, but not limited to: Dr. Brian Curtis and Prof. Frank Schroeder from Cornell University; Dr. Mengzhao "Lucy" Xue, Dr. Brandon Mercado, and Prof. Seth Herzon from Yale University; Dr. Gina Porras and Prof. Cassandra Quave from Emory University; Dr. Eun Bin Go, Dr. Masao "Masa" Ohashi, Dr. Zhuan Zhang, Dr. Bruno Perlatti, Jason Q. L. Williams, and Prof. Yi Tang from UCLA. There have been many more collaborators who have been extremely patient while their samples were backlogged due to the sheer number of samples we had to get through. I am grateful for everyone's understanding and scientific contributions, as well as sharing of exciting scientific ideas. In particular, I would like to thank to Masa for his tremendous support in supervising me on molecular biology and his scientific contribution for our published work over the years.

I would also like to express my gratitute for the education and support I've received from my undergraduate institution, Harvey Mudd College. The Mudd curriculum made me a well-rounded scholar that wasn't afraid to tackle problems in unknown territories. I learned the value of communicating with people from diverse fields, which was crucial in my highly interdisciplinary graduate research. I am extremely grateful for my undergraduate research advisor, Professor David Vosburg, who took me in as an undergraduate researcher, taught me the joy of performing research, and set my standards for what it means to be an amazing mentor. I am also thankful for all the other professors in the chemistry and biology departments including, but not limited to: Prof. Karl

Haushalter, Prof. Bill Daub, Prof. Lelia Hawkins, Prof. Kerry Karukstis, and Prof. Katherine Van Heuvelen. Thank you all for supporting me and cheering me on – it meant a lot to me to see so many of you attend Hosea's talk when he gave a seminar at the Claremont Colleges.

I am also thankful for my labmates in the Nelson lab, who have all been very helpful and supportive throughout my graduate studies. Dr. Alex Bagdasarian and Dr. Brian Shao were the most senior students. At first, I found it difficult to talk to them at first because I was both introverted and intimidated by their seniority, but I learned very quickly that they were both very friendly and easy to talk to. Alex and Brian set the bar for the students who came after them, and I learned so much about synthesis as well as presentation and writing skills from them. In the next class of students, Dr. Stasik Popov and Dr. Sydnee Green deserve the highest praise. Even though we were never working together directly, Stasik went above and beyond by giving me an orientation to the lab when I started rotating and lending me a hand whenever I needed help. I miss our evening chats in the scope room and our walks home, when we would talk about all sorts of topics such as our projects and the latest drama. It's always a joy to hear about how well he's doing after graduation, and I know that wherever he goes, he'll always be doing great science. Sydnee was the senior student I had the pleasure of working directly with when I first started in the lab. Her outgoing personality made it very easy to talk to her, and she was thorough in teaching and demonstrating labwork to me. I enjoyed our late-night labwork and hearing her distinct and frequent laughter in the lab. When I switched projects, I wasn't able to work with her as much, but it was inspiring and exciting to watch her spearhead the DNA catalysis project. I would also like to thank other students in my class: Benjamin Wigman, Sepand Nistanaki, Jessica Burch, and Christopher Jones. Most of my memories with Ben involve getting help, from getting a ride to the campsite to setting up the new lab space at Caltech, or listening to each other vent about something frustrating that happened. It was also nice to graduate at the same time as him and commiserate about writing our thesis, and I'm excited for him to start his new job after graduation. Sepand is an excellent scientist whose opinion I highly value; I always sent my drafts to him for feedback, and they always came back with thoughtful comments and suggestions for improvement. His passion for science and an eye for detail will make him a great asset in academia or in industry. Jess has been an irreplaceable friend and a project mate, and I wouldn't have asked for a better person to sit next to for years during graduate school. Her kind and humorous attitude, as well as our coffee and "hydration" breaks, have played a major part in helping me move past the gloomier days of graduate school and made each day a brighter day to come into work. Next, Chris has spearheaded the MicroED project to success, and without his achievement, I would not have started learning this technique in the first place. His talent for seeing the forest has helped me with storytelling in presentations and writing, and I wish him the best of luck in his own field transition. Next, I would like to thank the classes below me: Chloe Williams, Woojin Lee, and Zhenqi "Steven" Zhao. While Chloe and I have not worked together directly, she has made our interactions delightful with her cheerful personality. As she finishes up her fourth year of graduate school, her ability to brighten up the conversation and excellent work ethics will make her a desirable candidate for any workforce. Woojin and I first met during our summer internship at Genentech in 2017, and had an unexpected reunion in the Nelson lab the following year. Since our very first meeting, he has remained one of the most polite and positive people I know of, and I'm extremely thankful for his friendship and empathy over these years. Steven and I became closer friends during my last year when he gave me rides to and from Caltech, so that I wouldn't have to spend 5–6 hours a day on public transit commuting back and forth from Westwood late at night or very early in the morning. I enjoyed our conversations on various topics such as traveling, the lab, and the pandemic, and I was very surprised and grateful when he managed to find the time and energy to give me feedback on my thesis chapter drafts while preparing for his candidacy exam. I would also like to acknowledge the newest members of the MicroED project: Isabel Hernandez, Joshua Signore, Dr. David Delgadillo, and Dr. Lygia Silva de Moraes, who are all Caltech-affiliated. I haven't worked with them for as long as I did with the other labmates, but they have all proven to be supportive friends and coworkers, showing empathy for my long commutes and understanding for my lack of presence in the lab while writing my thesis. I'm excited to see all the amazing science they'll publish about in the future. Finally, I would like to thank Melody Tang, who is a high school intern I've had the pleasure of working with. Working with Melody has been very refreshing and enjoyable, and I wish her the best in starting her undergraduate studies in the upcoming school year.

I would also like to thank Nick Baerg and Annie Carpenter from the Graduate Office of UCLA Chemistry and Biochemistry. Nick has met with me on multiple occasions to give me advice, and each meeting provided me with clear directions on what to do next. It was a relief to have someone to go to whenever I felt clueless and confused. Annie is one of the major reasons why I am still here and was able to finish my PhD. At the end of my first year, when I took a sudden leave of absence due to a mix of mental health conditions including anxiety and depression, Annie was the person who offered to walk with me to the on-campus counseling center to make sure that I took my first step to receive help. I have been grateful for this simple yet hard-to-take action every since, and I'm glad that I was able to return to lab and reach my finish line. On a related note, I would also acknowledge my therapist Dr. Tracy Ballardo, without whom I would not have made it through my final year and a half. Her insightful comments and questions have been a tremendous help in recognizing the need for self-care and managing my eating disorder, anxiety and depression.

Additionally, I would like to thank my friends outside of the lab. One of my major hobbies is gaming, and playing games with my online or offline friends every night have been an incredible source of happiness that allowed me to move through the good and bad of each day. Also, even though I'm extremely bad at replying to texts and messages, my friends have shown true understanding of me and my situation with their patience and, at times, appropriate concern. Talking to my friends who are in different fields or outside of academia has also helped me get my mind off the feeling of being "stuck" as a PhD student, and recognize all of the opportunities that exist outside of and after graduate school. I have many friends to thank, but the ones that have supported me most consistently include Annisa Dea, Brandon Veles, Joshua Ryan "Rosh" Lam, Solhee "Suri" Kim, and Eun Bin Go. Out of them, I would like to especially thank Eun Bin Go for our long-lasting friendship and her support in various forms ranging from omakase dinner plans to dreaming up random science-related ideas. I have been fortunate to meet her in the Vosburg lab during my undergraduate studies, and every year she has demonstrated what it means to be an excellent friend. I would also like to acknowledge my dear friend from Harvey Mudd, Raunak Pednekar. I had heard about the news while preparing for my candidacy exam, and could not properly grieve for the world's loss. He was an excellent friend and a scholar: always curious about other people's ideas and willing to engage in a long debate about a topic he was passionate in. His love for life has always lifted my spirits, and will continue to inspire me to find my own joy in life.

Finally, I would like to acknowledge my family, both my immediate family in the United States, and my extended family back in South Korea. My family has been unbelievably patient and understanding. They have demonstated their support through various means, such as visiting me once a month with home-cooked meals and occasionally texting me to make sure that I was doing okay. My parents have often said that they feel bad for not being able to help with me navigate the U.S. education system; however, they have given me the best guidance I could have possibly asked for by letting me take the reins and believing in my choices. I'm especially grateful for their saying (in Korean), "It's okay to stop if you're having too hard of a time – your health is your number one priority." Ironically, this statement always made me more motivated, and has steered me away from quitting many times, but at the same time it let me put down some of the emotional burden I was feeling. I'm also thankful for my father's demonstration of following his passion in his own career and life, as it gives me courage to search for and follow my own passions in life even when I think it might be too late. As for my family back in South Korea, I've always been appreciative of them cheering me on in the form of video calls, and I'm excited to visit them for the first time in almost ten years once I graduate.

BIOGRAPHICAL SKETCH

Education:

University of California, Los Angeles, CA

- Ph.D. in Organic Chemistry, anticipated Spring 2022
- Advanced to Candidacy, Aug. 2019

Harvey Mudd College, Claremont, CA

- Joint Degree in Bachelor of Science in Chemistry and Biology, May 2017
- Undergraduate Researcher Jan. 2015 – May 2017

Professional and Academic Experience:

Graduate Student Researcher: University of California, Los Angeles, CA

- Sept. 2017 – present; Advisor: Prof. Hosea M. Nelson
- Spearheaded development of microcrystal electron diffraction for structural elucidation of natural products.
- Experienced in project lead position and demonstrated success on team projects.

Summer Research Intern: Genentech, Inc., South San Francisco, CA

- Jun. 2017 – Sept. 2017; Supervisors: Dr. Huiyong Hu and Dr. Michael Koehler
- Synthesized natural product analogs to be tested for biological activity.

Amgen Scholar and Summer Undergraduate Student Researcher: University of California, San Francisco, CA

- Jun. 2016 – Aug. 2016; Advisor: Prof. Jack Taunton
- Synthesized and tested covalent kinase inhibitors using *in vitro* and cell-based assays.

Undergraduate Student Researcher: Harvey Mudd College, Claremont, CA

- Jan. 2015 – May 2017; Advisor: Prof. David A. Vosburg
- Worked on synthesizing endiandric acid derivatives using iterative cross-coupling.

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

- Undergraduate organic chemistry labs and biochemistry discussion section (Fall 2017 – Spring 2020): Taught students laboratory techniques, general and organic chemistry concepts, and biochemistry.

Honors and Awards:

- UCLA Dissertation Fellowship (Jun. 2021)
- NSF BioPACIFIC MIP Fellowship (Jan. 2021)
- ACS Division of Organic Chemistry Summer Undergraduate Research Fellowship (Mar. 2016, Declined for Amgen Scholar Program)
- ACS Division of Organic Chemistry Travel Award (Jan. 2016)

Publications:

1. Go, E. B.; **Kim, L. J.**; Nelson, H. M.; Ohashi, M.; Tang, Y. Biosynthesis of the Fusarium Mycotoxin (–)-Sambutoxin. *Org. Lett.,* **2021**, *23*, 7819–7823.
2. **Kim, L. J.**; Ohashi, M.; Zhang, Z.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J. A.; Tang, Y.; Nelson, H. M. Prospecting for Natural Products by Genome Mining and Microcrystal Electron Diffraction. *Nat. Chem. Biol.*, **2021**, *17*, 872–877.
3. **Kim, L. J.**; Xue, M.; Li, X.; Paulson, E.; Mercado, B.; Nelson, H. M.; Herzon, S. B. Structure Revision of the Lomaiviticins. *J. Am. Chem. Soc.*, **2021**, *143*, 6578–6585.
4. Curtis, B. J.; **Kim, L. J.**; Wrobel, C. J. J.; Eagan, J. M.; Smith, R. A.; Burch, J. E.; Le, H. H.; Artyukhin, A. B.; Nelson, H. M.; Schroeder, F. C. Identification of Uric Acid Gluconucleoside–Ascaroside Conjugates in *Caenorhabditis elegans* by Combining Synthesis and MicroED. *Org. Lett.* **2020**, *22*, 6724–6728.
5. Jones, C. G.; Asay, M.; **Kim, L. J.**; Kleinsasser, J. F.; Saha, A.; Fulton, T. J.; Berkley, K. R.; Cascio, D.; Malyutin, A. G.; Conley, M. P.; Stoltz, B. M.; Lavallo, V.; Rodriguez, J. A.; Nelson, H. M. Characterization of Reactive Organometallic Species via MicroED. *ACS Cent. Sci.* **2019**, *5*, 1507–1513.
6. Go, E. B.; Wetzler, S. P.; **Kim, L. J.**; Chang, A. Y.; Vosburg, D. A. Concise, Diastereoconvergent Synthesis of Endiandric-type Tetracycles by Iterative Cross Coupling. *Tetrahedron.* **2016,** *72,* 3790–3794.

CHAPTER ONE

Microcrystal Electron Diffraction in Small Molecule Regime and Its Prospect in Natural Product Chemistry

1.1 Abstract

The bioactivity and high structural diversity of natural products (NPs) have inspired breakthroughs in drug development. While identification of novel NPs will undoubtedly accelerate the field of drug discovery, this process has been lagging due to challenges in NP structural determination. To this end, microcrystal electron diffraction (MicroED) has been highlighted as a way to address the challenges that arise from inherent limitations in traditional analytical techniques. However, despite the potential of MicroED in NP characterization, MicroED analysis of previously unknown non-peptidyl NPs has until recently remained rare in the literature. This chapter will outline the recent examples of utilizing MicroED to solve crystal structures of non-peptidyl NPs isolated from various biological sources where other analytical techniques have failed. The broad applicability and reliability of MicroED in unambiguous structure determination, even for novel NPs, will be highlighted. Finally, the future outlook of MicroED in NP chemistry will be addressed, detailing current challenges in MicroED workflow and key unanswered questions.

1.2 Introduction

Natural products (NPs) play a crucial role in the development of small molecule pharmaceuticals due to their structural complexity, diverse molecular frameworks, and bioactivity. For the last four decades, NPs, NP derivatives, and NP-inspired compounds have comprised the majority of all approved small molecule drugs.[1] Therefore, accelerating the rate of NP discovery, particularly compounds with novel scaffolds, will greatly aid the development of drug candidates. However, the field of NP chemistry is hindered by inherent challenges in structure elucidation, creating a major bottleneck in novel compound discovery.

NP characterization relies heavily on a combination of analytical techniques such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy.[2,3] Structural details such as atom connectivities and molecular formula can be accessed from these methods, allowing chemists to infer structures of NPs. However, while these techniques have advanced our knowledge of the NP chemical space for decades, they also exhibit limitations of their own. While MS boasts superb sensitivity for detecting thousands of NPs in a typical MS-based metabolomics experiment,[4] identification of these metabolites lags due to incomplete spectral databases and difficulty in identifying bond connectivity and stereoisomers.[5,6] To elucidate structures of these unknown NPs, NMR spectroscopy can be used. However, NMR experiments often require pure samples of milligram quantities, which is time- and labor-intensive to achieve for many unknown NPs that are produced in miniscule amounts. As a result, only a subset of NPs that are detected by MS can be structurally elucidated by NMR. Additional challenges, such as low number of protons and/or coupling, can hamper structure determination *via* NMR.[7–9] With the trend of drug candidates and novel natural products moving towards ever-increasing structural complexity,[10] NMR data may become insufficient to probe these intricate 3D structures.

Instead of inferring structures from spectroscopic techniques, single crystal X-ray crystallography can be used for unambiguous structural determination of NPs without any prior chemical knowledge. However, growing single crystals of sufficient size for X-ray diffraction studies, typically about 50 μm in all dimensions, can be an arduous process, and even more challenging when the NPs are produced in insufficient quantities. In some cases, poor physico-chemical properties of compounds can prevent the growth of large, high-quality crystals suitable for single crystal X-ray diffraction experiments. Overcoming these challenges in structural elucidation of NPs would unlock new opportunities for NP discovery.

1.3 Microcrystal Electron Diffraction

Microcrystal electron diffraction, or MicroED, is a cryoelectron microscopy technique that can help overcome these limitations. Unlike X-ray crystallography, MicroED experiments utilize electron beams inside transmission electron microscopes (TEMs) operating at a high voltage, ranging from 120 to 300 keV. Due to the strong interactions between electrons and matter, diffraction patterns can be obtained from crystals that are orders of magnitude smaller than those needed for single crystal X-ray diffraction.[11] These microcrystals may be present even in seemingly amorphous samples, such as powder obtained from purification and subsequent solvent evaporation, or in failed crystallization screens that yield only tiny precipitates. Moreover, microcrystals require much less material to generate than crystals suitable for X-ray crystallography, making this technique ideal for probing structures of molecules of precious amounts. These key advantages open up new avenues for discovery, as it provides chemists a method to unambiguously determine structures of their target molecules when the samples are unsuitable for other techniques.

In MicroED experiments, the sample – often pre-screened for diffraction using powder diffraction – is inserted into a transmission electron microscope (TEM), and a crystal is located and isolated by a selected area aperture. As the stage and the crystal are continuously rotated, a coherent electron beam projected at the crystal produces diffraction patterns, which are recorded in the form of movies containing numerous frames of diffraction pattern at different rotational angles (Figure 1.1).[12] This diffraction data can be processed using standard X-ray crystallographic software such as XDS,[13] DIALS,[14] and SHELX[15–18] to solve *ab initio* structures.

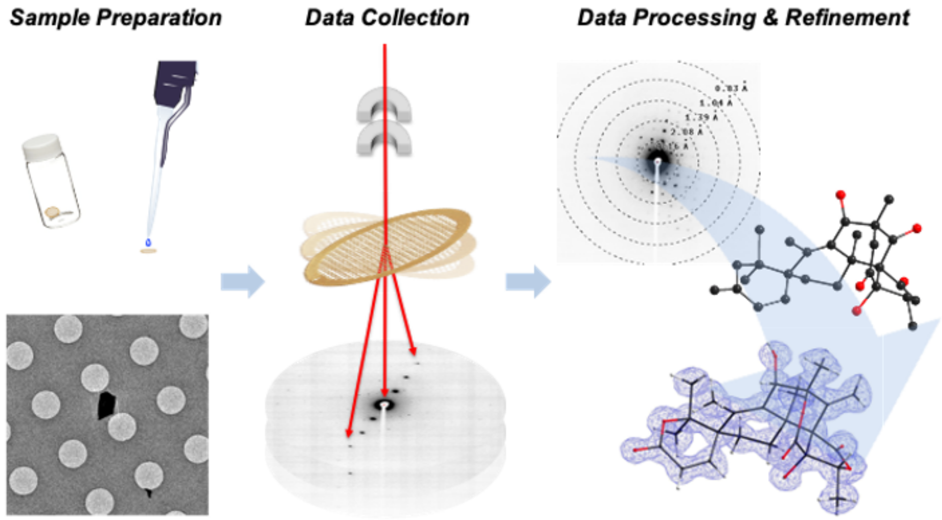


Figure 1.1 MicroED workflow from sample preparation to data processing and refinement.

1.3.1. Brief History of Electron Diffraction

While MicroED as a technique may be young (having been demonstrated in 2013),[19] the field of electron diffraction spands decades in history (Figure 1.2). Electron diffraction was discovered by Davisson and Germer in 1927 by scattering the electron beam from a nickel single crystal.[20] This famous experiment validated the wave nature of electrons, as hypothesized by de Broglie. Later in the same year, electron diffraction with higher kinetic energy was observed by Thomson and Reid,[21] and the first electron microscope was created by Ruska and Knoll in 1931, marking the beginnings of electron microscopy.[22] However, dynamical effects were a major concern that hindered structural analysis by electron crystallography, as electrons can scatter multiple times through a crystal and cause significant changes in the observed intensities.[23] Despite this phenomenon, Dorset and Hauptman showed in 1976 that crystallographic phases of electron diffraction data can be successfully recovered, enabling *ab initio* structure determination.[24] Structures of various organic and inorganic compounds solved fron electron diffraction data using direct methods were published in the literature,[25] as well as a protein structure from a 2D crystal solved using molecular replacement,[26] a common method that uses homology models to phase diffraction data and solve macromolecular structures. Many of these early accomplishments have been highlighted in the scientific community, as seen by the Nobel Prize awards received by Davisson, Germer, and Ruska, demonstrating the profound impact of their work.[27] From these earlier results, various electron diffraction methods, such as continuous rotation electron diffraction (cRED) and MicroED, were developed, and are generally described as "3D electron diffraction (3D ED)."

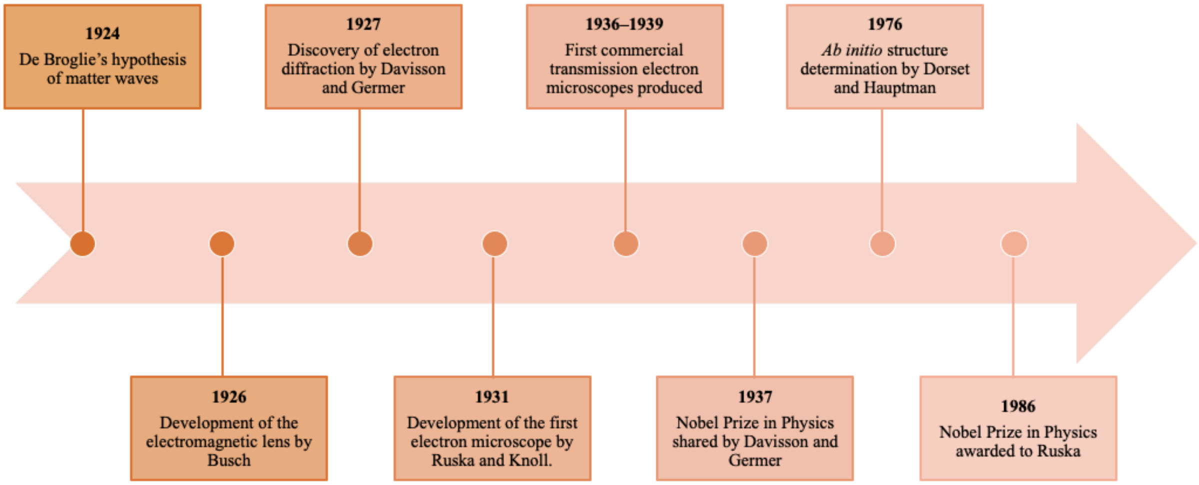


Figure 1.2 Timeline of selected historical events relevant to electron diffraction.

1.3.2 The Scope of MicroED in Small Molecule Regime

Since the initial report in 2013, MicroED has successfully been used to elucidate structures of various macromolecules, such as lysozyme and catalase.[19,28] Several early reports were limited to proteins whose structures have previously been solved using X-ray crystallography. However, more recent reports have used MicroED to solve structures of a novel protein R2lox and peptide fragments.[29] Many of these previous studies have relied on molecular replacement to phase the electron diffraction data; however, the few examples utilizing *ab initio* methods have demonstrated a convenient workflow of processing electron diffraction data using XDS and the SHELX suite,[30] which are common X-ray crystallographic software.

In 2018, MicroED became popularized among the field of structural chemistry when two papers, published back-to-back, demonstrated facile structure determination of various organic molecules in their commercial, seemingly amorphous powder form.[31,32] Both reports featured a simple workflow of preparing grids for analysis, which involved loading the dry powder samples onto a grid before plunge-freezing into liquid nitrogen, rather than vitrification in liquid ethane as seen in standard cryoEM experiments. The refined structures obtained from electron diffraction experiments were overlaid with previously solved X-ray structures to demonstrate good agreement between the two crystallographic methods. In many cases, hydrogen atoms could be resolved, and even a mixture of microcrystalline powders was able to be analyzed to afford crystal structures of individual components.

Shortly after, the technique was also applied to reactive organometallic species, demonstrating MicroED's ability to probe structures of chemical compounds whose structures are not as easily accessible due to their high reactivity and paramagnetism, preventing NMR analysis.[33] Notably, this study reported the use of fast and sensitive detectors that enabled MicroED experiments at ambient temperature, eliminating the need for the sample to be cooled to cryogenic temperature during data collection. This simplified workflow has since been commonly employed in our laboratory to shorten screening time, as preparing and inserting samples using the single-tilt sample holder at ambient temperature is easier and faster than mounting the sample in cryogenic conditions using the Gatan 626 cryo-holder.

Since these early applications to small molecules, MicroED has been used to obtain structures from various applications, ranging from metal-organic frameworks, pharmaceuticals, natural products, and more.[34] Many of these examples relied on MicroED due to limitations of other techniques, using the structure obtained from MicroED analysis to complement other experimental data for higher confidence in structure assignment. The technique's relevance in the small molecule regime is reflected by the increase in the number of small molecule structure deposition to the Cambridge Structural Database from 2013 to 2020 (Figure 1.3).[34g]

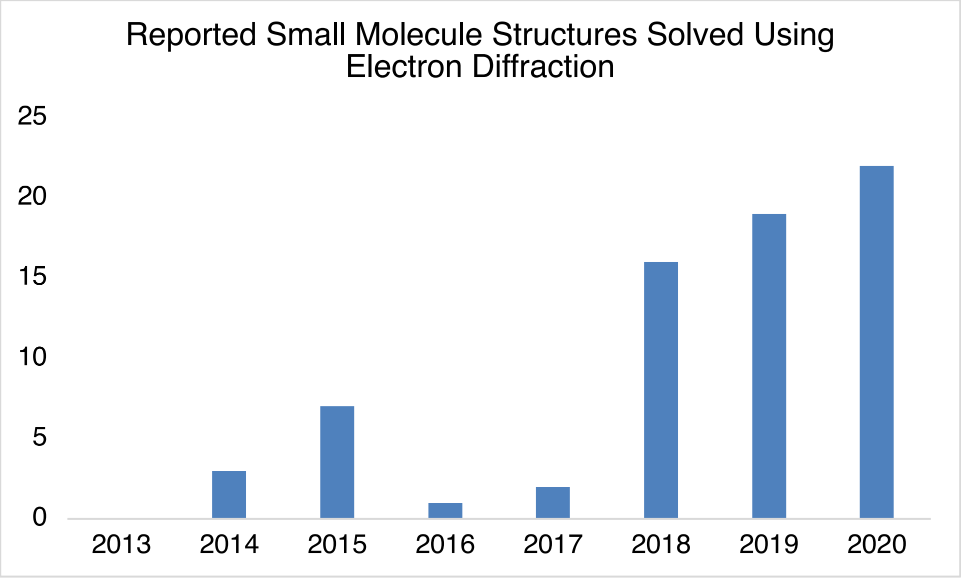


Figure 1.3 Number of small molecule structures obtained *via* electron diffraction and deposited to the Cambridge Structural Database (CSD) each year.

1.3.3 MicroED in NP Structure Elucidation

MicroED offers many advantages in structural characterization of NPs. Analogous to X-ray diffraction studies, MicroED experiments require crystalline material; however, as described earlier in this chapter, crystals that are orders of magnitude smaller can be used.The ability to probe crystals with dimensions on the order of nano- and micrometers provides a major advantage for studying NPs that are produced in infinitesimal quantities or are difficult to crystallize. Additionally, MicroED has been shown to be capable of screening mixtures of different compounds,[31,32,34o,34q] opening up the possibilities for detecting impurities, trace components, and/or new crystalline phases of NP mixtures that could be overlooked by other analytical techniques.

Despite its potential in NP chemistry, MicroED has not been used widely to elucidate structures of complex and/or novel NPs. Reports of NP characterization using MicroED has mainly been limited to peptide-based NPs.[31,34h] However, the chemical space of NPs is tremendously large; these highly diverse molecular frameworks demonstrate a need for structure determination that has not been well-represented in MicroED experiments. While the 2018 reports showed MicroED's potential for NP characterization, all chemical structures were previously known, with 3 of them representing novel forms of known compounds.[31,32] Solving crystal structures of novel non-peptidyl NPs to discover their chemical structures has not been reported in the scope of MicroED studies until very recently.

1.4 Recent Applications of MicroED in NP Structural Elucidation

In the past few years, MicroED has been highlighted in the chemical literature as a tool for overcoming limitations of existing structural characterization methods, particularly in the scope of NPs. In these cases, MicroED experiments contributed key structural information that enabled confident structure assignment, even of novel NPs, and revised previously proposed structures based on NMR analyses. These recent studies also showcase structures of various non-peptidyl NPs that represent a more diverse range of molecular weight and structural complexity, and are isolated from a wide range of biological organisms spanning from fungi to *Caenorhabditis elegans*. Taken together, these NPs demonstrate the broad applicability of MicroED in NP discovery and characterization.

1.4.1 Novel NP Identification

As pathogenic organisms continue to evolve and grow resistant to existing treatment methods, discovery of novel natural products has become paramount to fuel the development of new classes of pharmaceuticals with various mechanisms of action or molecular scaffolds.[35] Towards this effort, advancements in synthetic biology have enabled the production of novel, complex metabolites such as Py-469 (**1.2**). Py-469 is a novel fungal secondary metabolite that was biosynthesized using the *icc* biosynthetic gene cluster (BGC) from *Penicillium variabile*.[34o] This BGC is capable of producing ilicicolin H (**1.1**), an antifungal NP exhibiting a potent inhibition of the eukaryotic respiratory chain;[36,37] however, three additional enzymes were found in the BGC that were not involved in the biosynthesis of **1.2**. To investigate the possibility of further derivatizing ilicicolin H, these extra genes were heterologously expressed with the five other genes involved in the biosynthesis of **1.2**. The presence of the three additional enzymes led to the formation of **1.2**, which was purified from the host *via* high-performance liquid chromatography (HPLC) and subjected to MicroED analysis after lyophilization. MicroED experiments afforded a preliminary structure that unambiguously revealed the skeletal framework of Py-469, albeit with randomly assigned atoms (Figure 1.4). Notably, this structure was refined almost completely just based on bond angles and relative bond lengths, with no prior knowledge about the compound other than its molecular formula. NMR spectroscopic data confirmed the identity and the structure of this novel natural product, with one discrepancy in the location of a single nitrogen atom, which is a difficult element to distinguish from carbon with the current state of MicroED data quality. The corrected and fully refined MicroED structure revealed that Py-469 (**1.2**) is a novel epoxydiol-containing derivative of ilicicolin H, obtained as a result of oxidative dearomatization, epoxidation, and two reduction transformations catalyzed by the three additional *icc* enzymes. This example represents the first report of a novel, non-peptidyl natural product whose structure was elucidated using MicroED.

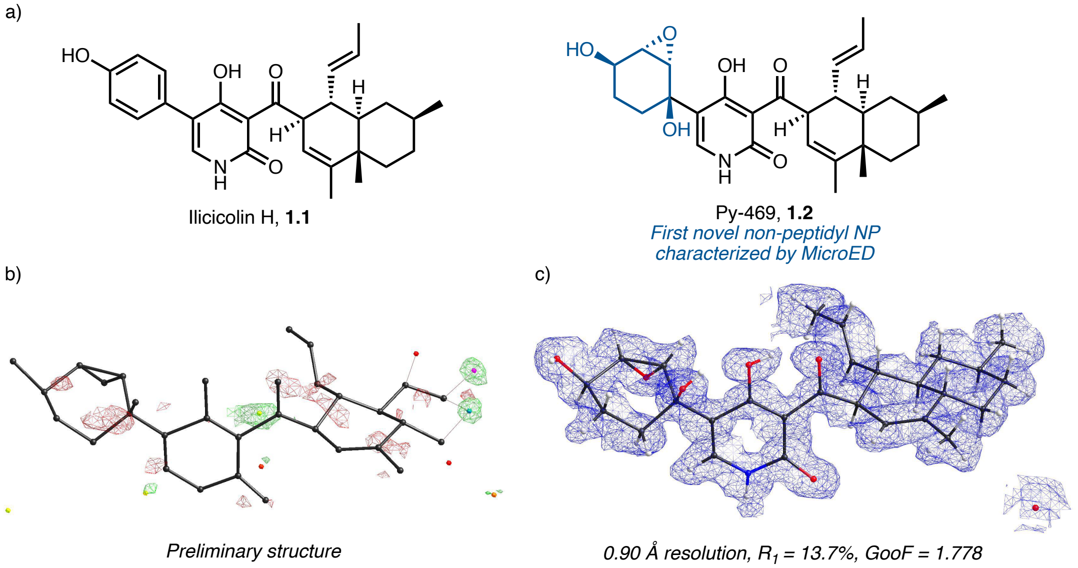
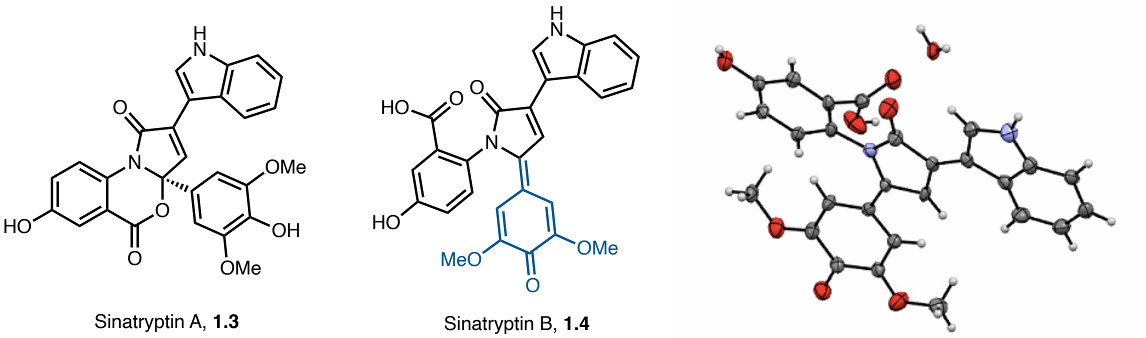


Figure 1.4. MicroED analysis of novel NP Py-469. (a) Chemical structures of ilicicolin H and its derivative Py-469. (b) Preliminary structure solution revealing the molecular framework of Py-469. (c) Fully refined structure of Py-469 (CCDC deposition number: 2038723).

Another method of obtaining novel natural products is the modulation of growth conditions for microorganisms.[34q] In their investigation of the algal-bacterial symbiosis between the haptophyte *Emiliania huxleyi* and α-proteobacterium *Phaeobacter inhibens*, Seyedsayamdost and coworkers detected a mutualist-to-parasite switch. During the mutualistically beneficial phase, the algae secretes molecules for promoting bacterial growth, and the bacteria produces compounds to promote and protect algal growth in return.[38,39] However, when *E. huxleyi* produces phenylpropanoids such as sinapic acid, *P. inhibens* responds by releasing algaecidal secondary metabolites.[38,40] In the laboratory, a mutualist-to-parasite switch in the system was mimicked by exposing *P. inhibens* to sinapic acid in the presence of tryptophan, which was hypothesized to be involved in the parasitic phase based on its significance in creating a feedback loop in the mutualistic phase. This lab-induced alteration in the growth conditions led to production of several metabolites incorporating tryptophan. Of these tryptophan-incorporating NPs, sinatryptin B (**1.4**) was isolated in amounts insufficient for spectroscopic analysis or growth of crystals amenable for X-ray crystallography. To overcome these challenges, MicroED experiments were performed to elucidate the structure of **1.4**, revealing a quinone derivative of its cometabolite sinatryptin A (**1.3**, Figure 1.5), and showcasing an example where MicroED was used as a key structural characterization method when other methods failed. Notably, work marks the second novel non-peptidyl NP whose structure was determined *ab initio* using MicroED.



Sinatryptin A, **1.3**

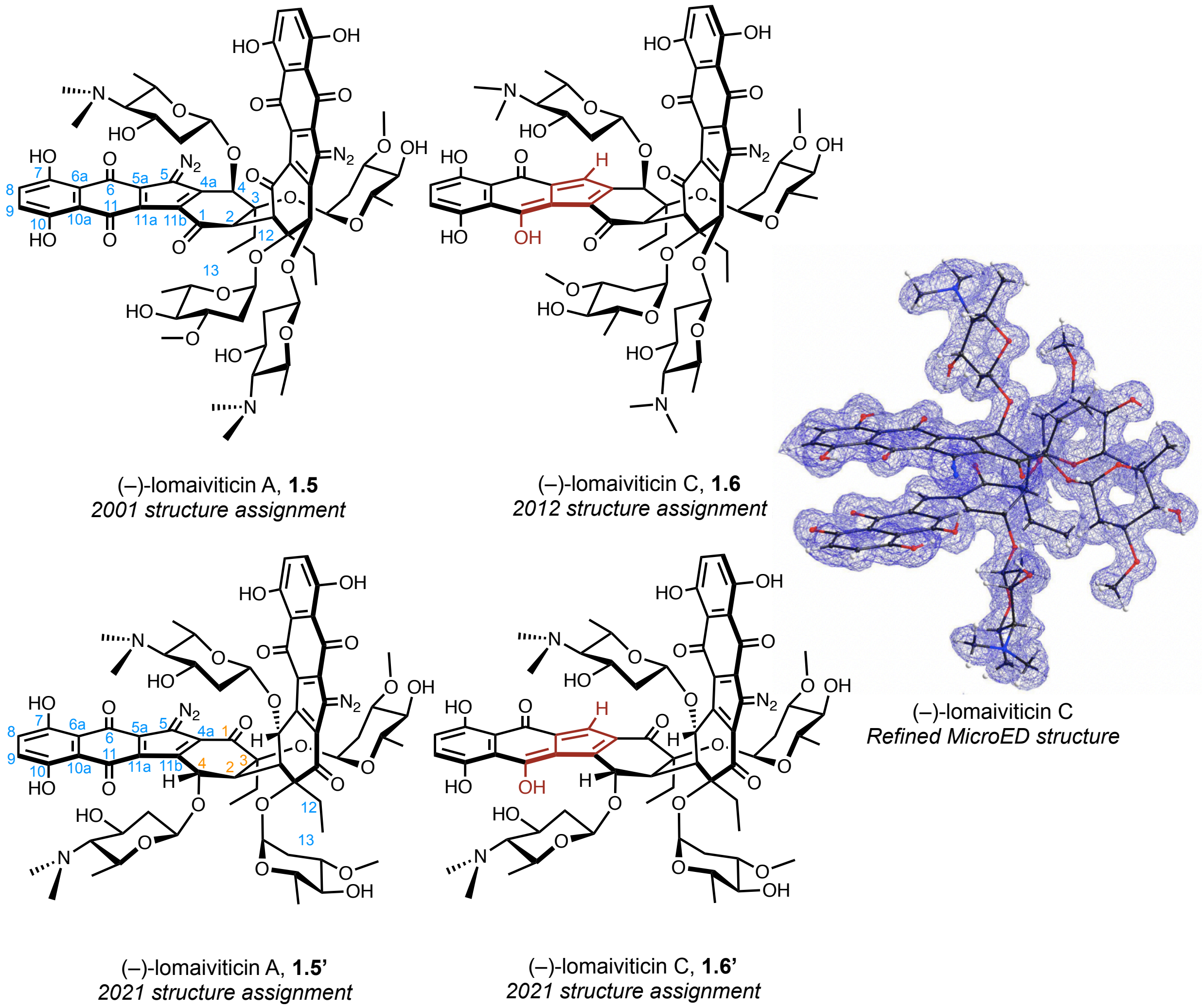
Sinatryptin B, **1.4**

Figure 1.5. Sinatryptin B, a quinone derivative of sinatryptin A, and its MicroED structure (CCDC deposition number: 2118442). Thermal ellipsoids are drawn at 30% probability.

1.4.2 Structural Revisions

Structure assignment of highly complex NPs can be a challenging process. The most routine characterization method is NMR spectroscopy, which enables structure determination by inference based on the data. However, this indirect view of the molecular structure can be misleading, as seen by various reports of NP structure revision,[41] and some of these incorrectly assigned structures may go unnoticed without lengthy total synthesis campaigns or obtaining a crystal structure of the natural molecule.[41c–f] Revising these incorrectly assigned structures is critical to ensure that synthetic efforts are biologically relevant and that its mode of action can be studied with higher confidence.

A notable example of structure revision include the collaborative work between our laboratory and the Herzon group at Yale University, wherein we revised the structure of the lomaiviticins, a family of bioactive NPs isolated from marine bacteria.[34n] Due to (–)-lomaiviticin A's potent cytotoxicity, this family of NPs has been extensively studied, featured in numerous articles discussing their bioactivity and biosynthesis.[42] However, due to the low number of proton-attached carbons to guide NMR analysis, confirmation of the originally proposed structure[43] (**1.5**, Figure 1.6) has proven exceedingly difficult. Moreover, despite two decades of synthetic efforts, total synthesis of the lomaiviticins have been unsuccessful. The reason for this became clear when (–)-lomaiviticin C (**1.6**), a precursor to lomaiviticin A that shares a similar core ring structure,[44] was subjected to MicroED analysis; we discovered that the previously proposed structures of the lomaiviticins, inferred from NMR data, was incorrect. In light of these findings, computational modeling, high-field (800 MHz) NMR spectroscopy, and density functional theory (DFT) calculations were used to show that the experimental NMR data is in better agreement with the revised structures (**1.5'** and **1.6'**) as opposed to the previously proposed structures. This example highlights a major limitation in structure determination by NMR, showing the difficulty of studying complex NPs with only a few proton-attached carbons and low coupling signals observed between the protons.



(–)-lomaiviticin A, **1.5**
2001 structure assignment

(–)-lomaiviticin C, **1.6**
2012 structure assignment

(–)-lomaiviticin C
Refined MicroED structure

(–)-lomaiviticin A, **1.5’**
2021 structure assignment

(–)-lomaiviticin C, **1.6’**
2021 structure assignment

Figure 1.6. Initially proposed structures of (–)-lomaiviticin A and C and their revised structures based on MicroED analysis (CCDC deposition number: 2062671). Core carbons are numbered in blue, and atoms exchanged during structure revision are colored in orange.

Another example of structure revision facilitated by MicroED is fischerin (**1.7**),[34o] which is an α-pyridone-containing fungal NP isolated from *Neosartorya fischeri* over 25 years ago that exhibits high bioactivity.[45] The chemical structure of fischerin was elucidated by NMR experiments, but the relative stereochemistry assignment proved challenging due to the freely rotatable bonds connecting the distal, stereochemically complex ring systems. A stereochemical assignment has been made using computational chemistry by Amini,[46] but no crystal structure of fischerin has been reported since its initial discovery. To probe the structure of fischerin using

MicroED without having access to the reported fungal strain, genome mining was used to identify a BGC that successfully produced fischerin. Consistent with the initial report by the isolation chemists, attempts to characterize fischerin *via* X-ray crystallography or NMR spectroscopy were unsuccessful; however, MicroED analysis of fischerin yielded a crystal structure revealing the relative stereochemistry of the epoxydiol ring substituents. Remarkably, this relative stereochemistry assignment was different than that of Amini's structure based on NMR shift calculations (Figure 1.7). This result demonstrates the utility of MicroED in complementing analytical techniques such as NMR, and highlights the importance of experimentally validating computed structures when possible.

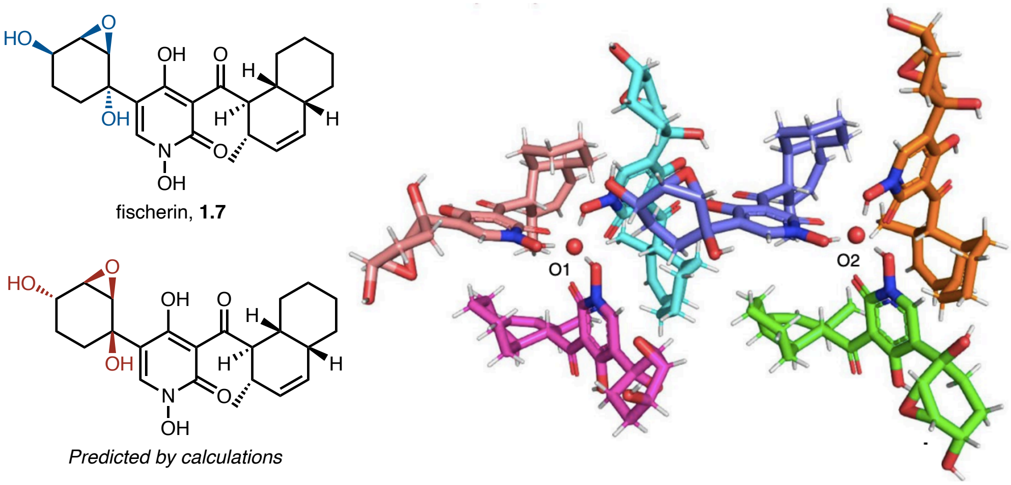


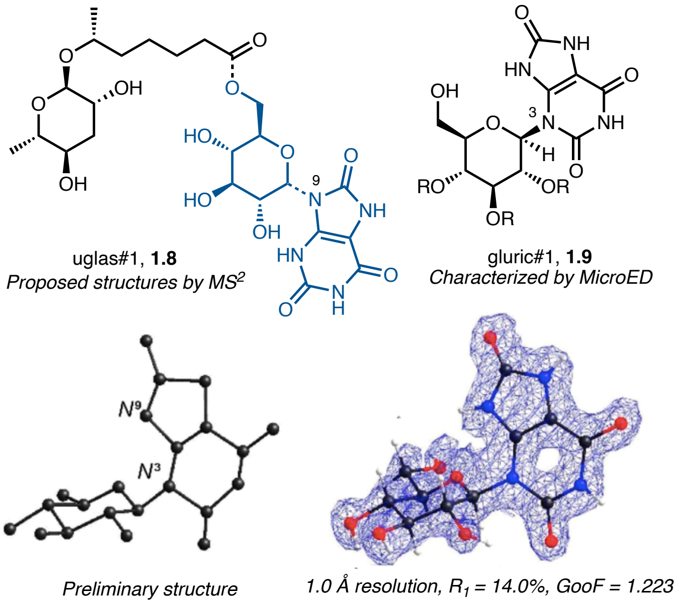
Figure 1.7. Fischerin's initially proposed relative stereochemistry based on NMR shift calculations, and the corrected relative stereochemistry based on MicroED structure (CCDC deposition number: 2020516).

1.4.3 Determining Connectivity

Determining the connectivity of structural motifs and functional groups within a natural product is crucial in understanding its molecular framework. Knowledge about a compound's molecular scaffold can guide drug discovery campaigns through docking experiments and studies on mechanisms of action based on the molecule's interaction within the active site of a protein.[47]

Molecular structures can often be elucidated from NMR spectroscopic analysis; however, this section focuses on cases where NMR data alone could not elucidate the full structure due to reasons such as poor line shapes and the difficulty of separating compounds present in a mixture. As seen in the earlier example of (–)-lomaiviticin C,[34n] where the connectivity of the atoms in the core structure was unambiguously determined and revised using MicroED, characterization of molecular scaffolds can advance synthetic efforts, biosynthetic understanding, and investigations on bioactivities.

Another example of elucidating the atom connectivity using MicroED is reported in the collaborative work between our group and the Schroeder laboratory, where MicroED was combined with a synthetic approach to characterize the uric acid-incorporating gluconucleoside NP, uglas#1 (**1.8**).[34m] Uglas#1 was isolated from *C. elegans*, a popular modeling organism for biological studies, and has been implicated to be involved in insulin signaling.[48] The proposed structure based on mass spectrometry data included uric acid, glucose, and an ascaroside moiety, but due to the broad line shapes in the 2D NMR spectra, the connectivity of the glucose and uric acid could not be unambiguously assigned. This challenge was overcome with a MicroED analysis of gluric#1 (**1.9**), a glucosyl uric acid NP used as a synthetic intermediate to uglas#1, whose MicroED structure revealed that glucosylation of uric acid is favored at the N^3 substitution (Figure 1.8). This study showcases another example of MicroED's utility in NP characterization, ability to address limitations in techniques such as NMR spectroscopy, and synergy with other analytical methods like mass spectrometry. Notably, the structure of gluric#1 was obtained with minimal human bias, where the only user input consisted of the compound's molecular formula prior to *ab initio* structure determination using direct methods.



uglas#1, **1.8**
Proposed structures by MS^2

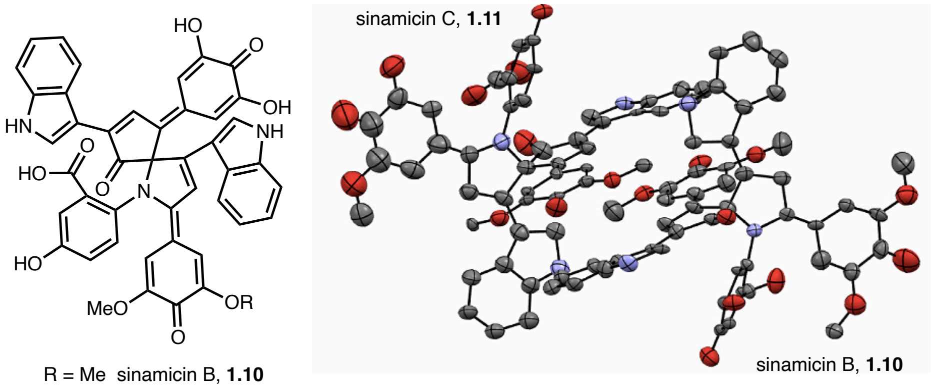
gluric#1, **1.9**
Characterized by MicroED

Preliminary structure

1.0 Å resolution, R_1 = 14.0%, GooF = 1.223

Figure 1.8. Proposed structure of uglas#1 and the MicroED structure (CCDC deposition number: 2020283) of gluric#1, a glucosyl uric acid NP used as a synthetic intermediate to **1.8**. Preliminary structure clearly reveals the N^3 connectivity.

Determining atom connectivity using NMR spectroscopy can also be challenging when the analyte solution contains a mixture of two or more compounds, especially when these compounds have few proton-attached carbon atoms to guide structure analysis.[7–9,34n] In the same report where the structure of sinatryptin A was elucidated, Seyedsayamdost and coworkers faced difficulties in separating sinamicin B (**1.10**) and sinamicin C (**1.11**), which are cometabolites of sinatryptin A and other novel alkaloids produced in the presence of sinapic acid and tryptophan.[34q] In addition to the difficult purification, sinamicin B and C were also produced in miniscule quantities and contained a low proton-to-carbon ratio that exacerbated the challenges in structure determination *via* NMR spectroscopy. Nevertheless, this NP mixture was successfully characterized using MicroED (Figure 1.9), where the two secondary metabolites crystallized together to generate microcrystals, allowing for two structure determinations in one experiment.



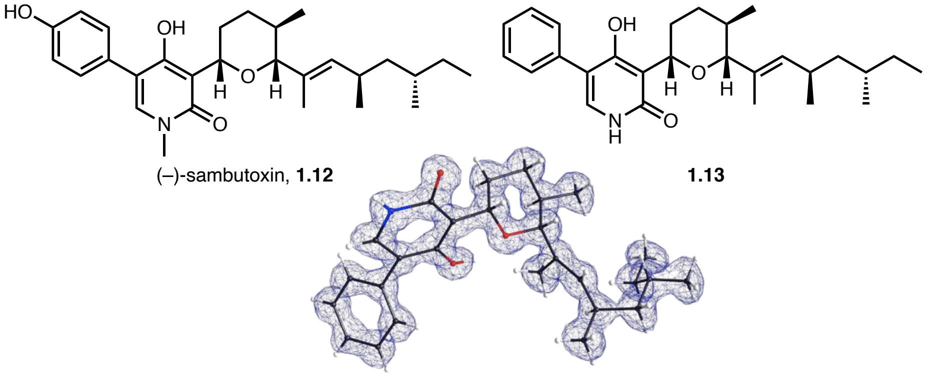
R = Me sinamicin B, **1.10**
R = H sinamicin C, **1.11**

Figure 1.9. MicroED structure showcasing co-crystallized sinamicin B and sinamicin C. In the asymmetric unit, sinamicin B is on the right and sinamicin C is on the left. Thermal ellipsoids drawn at 30% probability. Hydrogen atoms were not found in the CIF (CCDC deposition number: 2118443).

1.4.4. Assignment of Relative Stereochemistry

Assignment of relative stereochemistry is an important step in characterizing small molecules, particularly bioactive molecules such as secondary metabolites. One stereoisomer can exhibit entirely different bioactivity;[49] classic examples include propoxyphene and thalidomide. D-Propoxyphene exhibits analgesic activity,[50,51] while L-propoxyphene has antihistaminic activity,[51] and the (R)-enantiomer of thalidomide exhibits sedative effects and is effective against morning sickness, while the other enantiomer is teratogenic.[52] NMR spectroscopy often provides sufficient information to assign relative stereochemistry; however, some molecules lack suitably positioned hydrogen atoms, such as those containing distal stereocenters interrupted by rigid submoiety with rotatable bonds. Two examples disclosed earlier in this chapter, fischerin and Py-469,[34o] both featured these structural characteristics, preventing their relative stereochemistries from being confidently assigned. However, MicroED experiments clearly revealed the stereochemical information, even revising a previously proposed structure in the process.

In another collaborative work between our laboratory and the Tang group, MicroED was used to confirm the relative stereochemistry assignments in a biosynthetic intermediate of a fungal secondary metabolite (Figure 1.10).[34p] (–)-Sambutoxin (**1.12**) is a 4-hydroxy-2-pyridone toxin isolated from the potato parasite *Fusarium sambucinum*.[53] Of the 4-hydroxy-2-pyridone alkaloids, the 3-ether subfamily, which contains a characteristic tetrahydropyran motif, has remained understudied unlike its other subfamilies.[54] Using genome mining and stepwise reconstitution to study the biosynthetic pathway to **1.12**, key biosynthetic intermediates were isolated and characterized. Of these intermediates, **1.13** is the product directly after the tetrahydropyran formation, which subsequently undergoes a late-stage, P450-catalyzed oxidation and *N*-methylation to forge the final product (–)-sambutoxin.[34p] To characterize this key intermediate and confirm the relative stereochemistry between the tetrahydropyran motif and the methyl groups in the polyketide chain, MicroED and NMR spectroscopy were combined. These analyses confirmed that the relative stereochemistry is identical to that reported for (–)-sambutoxin, supporting that **1.13** is likely to be on path for the biosynthesis of **1.12**.



(–)-sambutoxin, **1.12**

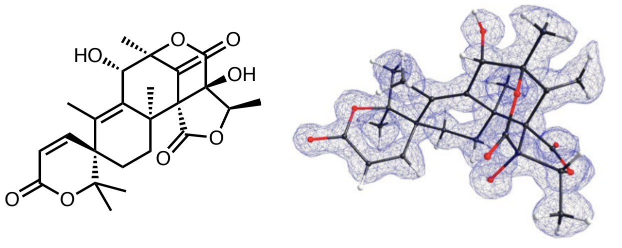
1.13

Figure 1.10. Structure of (–)-sambutoxin biosynthetic intermediate **1.13**, featuring the characteristic tetrahydropyran motif (CCDC deposition number: 2097629).

1.4.5 Detection of Trace Impurities

Natural product extract contains numerous metabolites, many of which are not produced in abundant amounts and may go unnoticed by characterization methods such as NMR. Even if a molecule is present in the solution, poor solubility in deuterated NMR solvents may prevent its detection. One of the key advantages of MicroED is its ability obtain structures from infinitesimal amount of crystalline sample, which makes MicroED a complementary technique to NMR spectroscopy by detecting and characterizing trace amounts of material.

A clear example can be seen in the collaborative work between the Tang group and our laboratory. During our analysis of fischerin, small amounts of a highly crystalline impurity were detected. Determination of its crystal structure revealed its identity to be austinol (**1.14**), a complex, polycyclic meroterpenoid NP that was co-purified with fischerin in trace amounts.[34o] The presence of austinol went unnoticed in the NMR analysis of fischerin, due to its low abundance dissolved in the NMR solvent. This finding highlighted that MicroED can be used to detect impurities in seemingly pure samples in NMR experiments, where minimal signals can be obtained from analytes with low solubility in the solvent used. Serial dilution experiments determined that merely 3 ng of sample could be detected, demonstrating MicroED's impressive sensitivity and potential for detecting trace impurities or minor components in NP mixtures (Figure 1.11).



austinol, **1.14** *1.0 Å resolution, R_1 = 14.7%, GooF = 1.401*

Figure 1.11. Structure of trace impurity austinol (CCDC deposition number: 2020510).

1.5 Challenges and Outlook in NP Characterization Using MicroED

In 2020 alone, over 20 small molecule crystal structures solved from electron diffraction methods have been released in the Cambridge Structural Database, with a steady increase in numbers since 2014 (Figure 1.3).[34g] However, none of these structures represented novel NPs whose structures were previously unknown. The lack of novel NP structures using electron diffraction data can be attributed to many reasons: 1) the expertise and instrumentation requirements of the technique; 2) the lack of optimized, automated MicroED workflow; 3) and the difficulty of obtaining high quality diffraction patterns from an unknown compound in a mixture of highly complex NP extract. Currently, MicroED is not routinely used for analysis due to the high level of expertise required in crystallography and electron microscopy, preventing a wide adoption of the method.[34g,55] These barriers in knowledge and instrumentation hinder more chemists from adopting MicroED in their routine analysis of NPs. Even if a researcher is equipped with the high level of expertise necessary for routine MicroED experiments, the rate of NP characterization using MicroED is heavily limited by the time- and personnel-intensive nature of the workflow. In our group's experience, MicroED analysis of complex NP samples have been unpredictable and largely "hit-or-miss". Some samples can be extremely challenging to crystallize, due to the molecule's high flexibility or disadvantageous properties such as poor solubility. Screening recrystallization conditions can add up to weeks or even months for a difficult sample; for example, the crystal structure of fischerin took six months of focused research effort. A number of methods have been reported to overcome these challenges, such as "crystalline sponge" and encapsulated nanodroplet crystallization (EnaCt) techniques.[56–58] These challenges are exacerbated when the purity is low. Low purity samples typically do not crystallize well, unless the compound exhibits high rigidity and is prone to easy crystallization, such as austinol that was present in the fischerin sample. The majority of our attempts on samples with low purity levels (<85%) have not been successful in microcrystal generation. Even after a successful growth of microcrystals, some crystals diffract at sub-ångstrom resolution even at room temperature and take less than an hour of data collection before structure solution, while others require cryogenic conditions to attenuate radiation damage, resulting in hours of manual screening per sample. Additionally, due to the high sensitivity of MicroED, trace impurities such as silicates (in the case of samples purified using liquid chromatography) can often hamper structure determination of target molecules. These challenges that prevent routine MicroED workflow can be addressed with increased automation and higher throughput of experiments that lower the knowledge barrier and manual labor required. The reduced cost of resources will enable a greater number of samples and more complex systems to be analyzed using MicroED, and as more NPs and their mixtures are studied, the rate of NP structure determination and novel NP discoveries can be accelerated with MicroED.

Since the breakthroughs that led to the birth of MicroED, this emerging methodology is still undergoing development, aided by technological developments in transmission electron microscopes and detectors that improve the diffraction data quality and pushing the resolution past its former limits. These advancements are heavily relied on by MicroED today, where high-resolution diffraction data are collected on transmission electron microscopes equipped with fast and sensitive detectors, then phased using software programs such as SHELX package.[15–18] However, key questions remain largely unanswered, such as determination of absolute stereochemistry in complex NPs. A few studies have shown examples of absolute stereochemistry determination, but relied on internal standards of known configuration.[34h,i] One report suggested dynamical refinement, which takes multiple scattering events of electrons into account, as a method to determine absolute configuration.[59] While this approach has not yet been widely used and tested, it suggests a possible future for absolute structure determination using MicroED. Another development in the field includes nanobeam diffraction studies that enable sampling of crystalline domains even smaller than those used in MicroED analysis.[60] This finer sampling may provide a closer look into inhomogeneities across a crystal and help gain a better understanding of the effects of dynamical scattering observed in MicroED experiments.

1.6 Conclusion

Unambiguous structure determination using MicroED is a powerful technique in characterizing NPs when other methods fail. The recent reports demonstrated the possible pitfalls when using NMR spectroscopy, including structural misassignments, inability to detect compounds with poor physico-chemical properties, and preclusion of relative stereochemistry information in structures with distal stereocenters. These studies also illustrated the broad applicability of MicroED in NP chemistry by solving crystal structures of NPs isolated from a variety of biological sources, ranging from fungi to *C. elegans*. With these findings and further technological and methodological developments, MicroED can begin to establish itself as a routine structural technique for characterizing NPs, and serve to solve crystal structures of numerous novel NPs in the future, expanding our knowledge base of the NP chemical space.

1.7 Notes and References

(1) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803.

(2) Breton, R. C.; Reynolds, W. F. Using NMR to identify and characterize natural products. *Nat. Prod. Rep.* **2013**, *30*, 501–524.

(3) Bouslimani, A.; Sanchez, L. M.; Garg, N.; Dorrestein, P. C. Mass spectrometry of natural products: Current, emerging and future technologies. *Nat. Prod. Rep.* **2014**, *31*, 718–729.

(4) Emwas, A.-H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Nagana Gowda, G. A.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; Wishart, D. S. NMR spectroscopy for metabolomics research. *Metabolites* **2019**, *9*, 123.

(5) Boiteau, R. M.; Hoyt, D. W.; Nicora, C. D.; Kinmonth-Schultz, H. A.; Ward, J. K.; Bingol, K. Structure elucidation of unknown metabolites in metabolomics by combined NMR and MS/MS prediction. *Metabolites* **2018**, *8*, 8.

(6) Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A. Untargeted metabolomics strategies-challenges and emerging directions. *J. Am. Soc. Mas. Spectrom.* **2016**, *27*, 1897–1905.

(7) Buevich, A. V.; Williamson, R. T.; Martin, G. E. NMR structure elucidation of small organic molecules and natural products: Choosing ADEQUATE vs HMBC. *J. Nat. Prod.* **2014**, *77*, 1942–1947.

(8) Kutateladze, A. G.; Kuznetsov, D. M.; Beloglazkina, A. A.; Holt, T. Addressing the challenges of structure elucidation in natural products possessing the oxirane moiety. *J. Org. Chem.* **2018**, *83*, 8341–8352.

(9) Riccio, R.; Bifulco, G.; Cimino, P.; Bassarello, C.; Gomez-Paloma, L. Stereochemical analysis of natural products. Approaches relying on the combination of NMR spectroscopy and computational methods. *Pure Appl. Chem.* **2003**, *75*, 295–308.

(10) Eastgate, M. D.; Schmidt, M. A.; Fandrick, K. R. On the design of complex drug candidate syntheses in the pharmaceutical industry. *Nat. Rev. Chem.* **2017**, *1*, 0016.

(11) Henderson R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **1995**, *28*, 171–193.

(12) Nannenga, B. L.; Shi, D.; Leslie, A. G. W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930.

(13) Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.

(14) Clabbers, M. T. B.; Gruene, T.; Parkhurst, J. M.; Abrahams, J. P.; Waterman, D. G. Electron diffraction data processing with DIALS. *Acta Cryst.* **2018**, *D74*, 506–518.

(15) Schneider, T. R.; Sheldrick, G. M. Substructure solution with SHELXD. *Acta Cryst.* **2002**, *D58*, 1772–1779.

(16) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(17) Sheldrick, G. M. SHELXT-Integrated space-group and crystal-structure determination. *Acta Cryst.* **2015**, *A71*, 3–8.

(18) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(19) Shi, D.; Nannenga, B. L.; Iadanza, M. G.; Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife* **2013**, *2*, e01345.

(20) Davisson, C.; Germer, L. H. Diffraction of electrons by a crystal of nickel. *Phys. Rev.* **1927**, *30*, 705–740.

(21) Thomson, G. P.; Reid, A. Diffraction of cathode rays by a thin film. *Nature* **1927**, *119*, 890.

(22) (a) Freundlich, M. M. Origin of the electron microscope. *Science* **1963,** *142*, 185–188. (b) Ruska, E. The development of the electron microscope and of electron microscopy. *Biosci. Rep.* **1987**, *7*, 607–629.

(23) Cowley, J. M.; Moodie, A. F. The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Cryst.* **1957**, *10,* 609–619.

(24) Dorset, D. L.; Hauptman, H. A. Direct phase determination for quasi-kinematical electron diffraction intensity data from organic microcrystals. *Ultramicroscopy* **1976**, *1*, 195–201.

(25) (a) Dorset, D. L. Electron crystallography. *Acta Cryst.* **1996**, *B52,* 753–769. (b) Weirich, T. E.; Ramlau, R.; Simon, A.; Hovmöller, S.; Zou, X. A crystal structure determined with 0.02 Å accuracy by electron microscopy. *Nature* **1996**, *382*, 144–146.

(26) Gonen, T.; Cheng, Y.; Sliz, P.; Hiroaki, Y.; Fujiyoshi, Y.; Harrison, S. C.; Walz, T. Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **2005**, *438,* 633–638.

(27) (a) The Nobel Prize in Physics, 1937. Nobel Prize Outreach AB 2022. https://www.nobelprize.org/prizes/physics/1937/summary/ (accessed May 21, 2022). (b) The Nobel Prize in Physics, 1986. Nobel Prize Outreach AB 2022. https://www.nobelprize.org/prizes/physics/1986/summary/ (accessed May 21, 2022).

(28) Nannenga, B. L.; Shi, D.; Hattne, J.; Reyes, F. E.; Gonen, T. Structure of catalase determined by MicroED. *eLife* **2014**, *3*, e03600.

(29) (a) Xu, H.; Lebrette, H.; Clabbers, M. T. B.; Zhao, J.; Griese, J. J.; Zou, X.; Högbom, M. Solving a new R2lox protein structure by microcrystal electron diffraction. *Sci. Adv.* **2019**, *5*, (b) Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; Jiang, L.; Arbin, M. A.; Nannenga, B. L.; Hattne, J.; Whitelegge, J.; Brewster, A. S.; Messerschmidt, M.; Boutet, S.; Sauter, N. K.; Gonen, T.; Eisenberg, D. S. Structure of the toxic core of α-synuclein from invisible crystals. *Nature* **2015**, *525*, 486–490. (c) de la Cruz, M. J.; Hattne, J.; Shi, D.; Seidler, P.; Rodriguez, J.; Reyes, F. E.; Sawaya, M. R.; Cascio, D.; Weiss, S. C.; Kim, S. K.; Hinck, C. S.; Hinck, A. P.; Calero, G.;

Eisenberg, D.; Gonen, T. Atomic-resolution structures from fragmented protein crystals with the cryoEM method MicroED. *Nat. Methods* **2017**, *14*, 399–402.

(30) Sawaya, M. R.; Rodriguez, J.; Cascio, D.; Collazo, M. J.; Shi, D.; Reyes, F. E.; Hattne, J.; Gonen, T.; Eisenberg, D. S. Ab initio structure determination from prion nanocrystals at atomic resolution by MicroED. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 11232–11236.

(31) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(32) Gruene, T.; Wennmacher, J. T. C.; Zaubitzer, C.; Holstein, J. J.; Heidler, J.; Fecteau-Lefebvre, A.; De Carlo, S.; Müller, E.; Goldie, K. N.; Regeni, I.; Li, T.; Santiso-Quinones, G.; Steinfeld, G.; Handschin, S.; van Genderen, E.; van Bokhoven, J. A.; Clever, G. H.; Pantelic, R. Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed.* **2018**, *57*, 16313–16317.

(33) Jones, C. G.; Asay, M.; Kim, L. J.; Kleinsasser, J. F.; Saha, A.; Fulton, T. J.; Berkley, K. R.; Cascio, D.; Malyutin, A. G.; Conley, M. P.; Stoltz, B. M.; Lavallo, V.; Rodríguez, J. A.; Nelson, H. M. Characterization of reactive organometallic species via MicroED. *ACS Cent. Sci.* **2019**, *5*, 1507–1513.

(34) (a) Banihashemi, F.; Bu, G.; Thaker, A.; Williams, D.; Lin, J. Y. S.; Nannenga, B. L. Beam-sensitive metal-organic framework structure determination by microcrystal electron diffraction. *Ultramicroscopy* **2020**, *216*, 113048. (b) Aykanat, A.; Jones, C. G.; Cline, E.; Stolz, R. M.; Meng, Z.; Nelson, H. M.; Mirica, K. A. Conductive stimuli-responsive coordination network linked with bismuth for chemiresistive gas sensing. *ACS Appl. Mater. Interfaces* **2021**, *13*, 60306–60318. (c) Meng, Z.; Jones, C. G.; Farid, S.; Khan, I. U.; Nelson, H. M.; Mirica, K. A.

Unraveling the electrical and magnetic properties of layered conductive metal-organic framework with atomic precision. *Angew. Chem. Int. Ed.* **2021**, *60*, 2–11. (d) Jellen, M. J.; Liepuoniute, I.; Jin, M.; Jones, C. G.; Yang, S.; Jiang, X.; Nelson, H. M.; Houk, K. N.; Garcia-Garibay, M. A. Enhanced gearing fidelity achieved through macrocyclization of a solvated molecular spur gear. *J. Am. Chem. Soc.* **2021**, *143*, 7740–7747. (e) Samkian, A.; Kiel, G. R.; Jones, C. G.; Bergman, H. Oktawiec, J.; Nelson, H. M.; Tilley, T. D. Elucidation of diverse solid-state packing in a family of electron-deficient expanded helicenes via microcrystal electron diffraction (MicroED). *Angew. Chem. Int. Ed.* **2020**, *5*, 2493–2499. (f) Burch, J. E.; Smith, A. G.; Caille, S.; Walker, S. D.; Wurz, R.; Cee, V. J.; Rodriguez, J.; Gostovic, D.; Quasdorf, K.; Nelson, H. M. Putting MicroED to the test: An account of the evaluation of 30 diverse pharmaceutical compounds. *ChemRxiv* **2021**. This content is a preprint and has not been peer-reviewed. DOI:10.26343/chemrxiv-2021-h3tqz. (g) Bruhn, J. F.; Scapin, G.; Cheng, A.; Mercado, B. Q.; Waterman, D. G.; Ganesh, T.; Dallakyan, S.; Read, B. N.; Nieusma, T.; Lucier, K. W.; Mayer, M. L.; Chiang, N. J.; Poweleit, N.; McGilvray, P. T.; Wilson, T. S.; Mashore, M.; Hennessy, C.; Thomson, S.; Wang, B.; Potter, C. S.; Carragher, B. Small molecule microcrystal electron diffraction for the pharmaceutical industry-lessons learned from examining over fifty samples. *Front. Mol. Biosci.* **2021**, *8*, 648603. (h) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; van der Donk, W. A. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **2019**, *365*, 280–284. (i) Dick, M.; Sarai, N. S.; Martynowycz, M. W.; Gonen, T.; Arnold, F. H. Tailoring tryptophan synthase TrpB for selective quaternary carbon bond formation. *J. Am. Chem. Soc.* **2019**, *141*, 19817–19822. (j) Gleason, P. R.; Nannenga, B. L.; Mills, J. H. Rapid structural analysis of a synthetic non-canonical amino acid by microcrystal electron diffraction. *Front. Mol. Biosci.* **2021**, *7*, 609999. (k) Nelson, H. M.; Siu,

J. C.; Saha, A.; Cascio, D.; MacMillan, S. N.; Wu, S.-B.; Lu, C.; Rodriguez, J. A.; Houk, K. N.; Lin, S. Isolation and X-ray crystal structure of an electrogenerated TEMPO-N_3 charge-transfer complex. *Org. Lett.* **2021**, *23*, 454–458. (l) Zee, C.-T.; Glynn, C.; Gallagher-Jones, M.; Miao, J.; Santiago, C. G.; Cascio, D.; Gonen, T.; Sawaya, M. R.; Rodriguez, J. A. Homochiral and racemic MicroED structures of a peptide from the ice nucleation protein InaZ. *IUCrJ* **2019**, *6*, 197–205. (m) Curtis, B. J.; Kim, L. J.; Wrobel, C. J. J.; Eagen, J. M.; Smith, R. A.; Burch, J. E.; Le, H. H.; Artyukhin, A. B.; Nelson, H. M.; Schroeder, F. C. Identification of uric acid gluconucleoside-ascaroside conjugates in *Caenorhabditis elegans* by combining synthesis and MicroED. *Org. Lett.* **2020**, *22*, 6724–6728. (n) Kim, L. J.; Xue, M.; Li, X.; Xu, Z.l Paulson, E.; Mercado, B. Q.; Nelson, H. M.; Herzon, S. Structural revision of the lomaiviticins. *J. Am. Chem. Soc.* **2021**, *143*, 6578–6585. (o) Kim, L. J.; Ohashi, M.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J.; Tang, Y.; Nelson, H. M. Prospecting for natural products by genome mining and microcrystal electron diffraction. *Nat. Chem. Biol.* **2021**, *17*, 872–877. (p) Go, E. B.; Kim, L. J.; Nelson, H. M.; Ohashi, M.; Tang, Y. Biosynthesis of the Fusarium mycotoxin (–)-sambutoxin. *Org. Lett.* **2021**, *23*, 7819–7823. (q) Park, J.-D.; Li, Y.; Moon, K.; Han, E. J.; Lee, S. R.; Seyedsayamdost, M. R. Structural elucidation of cryptic algaecides in marine algal-bacterial symbioses by NMR spectroscopy and MicroED. *Angew. Chem. Int. Ed.* **2022**, *61*, e202114022.

(35) Perfect, J. R. The antifungal pipeline: A reality check. *Nat. Rev. Drug Discov.* **2017**, *16*, 603–616.

(36) Singh, S. B.; Liu, W.; Li, X.; Chen, T.; Shafiee, A.; Card, D.; Abruzzo, G.; Flattery, A.; Gill, C.; Thompson, J. R.; Rosenbach, M.; Dreikorn, S.; Hornak, V.; Meinz, M.; Kurtz, M.; Kelly, R.; Onishi, J. C. Antifungal spectrum, in vivo efficacy, and structure-activity relationship of ilicicolin H. *ACS Med. Chem. Lett.* **2012**, *3*, 814–817.

(37) Zhang, Z.; Jamieson, C. S.; Zhao, Y.-L.; Li, D.; Ohashi, M.; Houk, K. N.; Tang, Y. Enzyme-catalyzed inverse-electron demand Diels-Alder reaction in the biosynthesis of antifungal ilicicolin H. *J. Am. Chem. Soc.* **2019**, *141*, 5659–5663.

(38) Seyedsayamdost, M. R.; Case, R. J.; Kolter, R.; Clardy, J. The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis. Nat. Chem.* **2011**, *3*, 331–335.

(39) (a) Wilson, M. Z.; Wang, R.; Gitai, Z.; Seyedsayamdost, M. R. Mode of action and resistance studies unveil new roles for tropodithietic acid as an anticancer agent and the γ–glutamyl cycle as a proton sink. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1630–1635. (b) Seymour, J. Simó, R.; Ahmed, T.; Stocker, R. Chemoattraction to dimethylsulfoniopropionate throughout the marine microbial food web. *Science* **2010**, *329*, 342–345. (c) Brinkhoff, T.; Bach, G.; Heidorn, T.; Liang, L.; Schlingloff, A.; Simon, M. Antibiotic production by a *Roseobacter* clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl. Environ. Microbiol.* **2004**, *70*, 2560–2565. (d) Bruhn, J. B.; Nielsen, K. F.; Hjelm, M.; Hansen, M.; Bresciani, J.; Schulz, S.; Gram, L. Ecology, inhibitory activity, and morphogenesis of a marine antagonistic bacterium belonging to the *Roseobacter* clade. *Appl. Environ. Microbiol.* **2005**, *71*, 7263–7270. (e) Thiel, V.; Brinkhoff, T.; Dickschat, J. S.; Wickel, S.; Grunenberg, J.; Wagner-Döbler, I.; Simon, M.; Schulz, S. Identification and biosynthesis of tropone derivatives and sulfur volatiles produced by bacteria of the marine *Roseobacter* clade. *Org. Biomol. Chem.* **2010**, *8*, 234–246. (f) Segev, E.; Wyche, T. P.; Kim, K. H.; Petersen, J.; Ellebrandt, C.; Vlamakis, H.; Barteneva, N.; Paulson, J. N.; Chai, L.; Clardy, J.; Kolter, R. Dynamic metabolic exchange governs a marine algal-bacterial interaction. *eLife* **2016**, *5*, e17473. (g) Bramucci, A. R.; Case, R. J. *Phaeobacter inhibens* induces apoptosis-like programmed cell death in calcifying *Emiliania huxleyi. Sci. Rep.* **2019**, *9*, 5215. (h) Wagner-Döbler, I.;

Ballhausen, B.; Berger, M.; Brinkhoff, T.; Buchholz, I.; Bunk, B.; Cypionka, H.; Daniel, R.; Drepper, T.; Gerdts, G.; Hahnke, S.; Han, C.; Jahn, D.; Kalhoefer, D.; Kiss, H.; Klenk, H.-P.; Kyrpides, N.; Liebl, W.; Liesegang, H.; Meincke, L.; Pati, A.; Petersen, J.; Piekarski, T.; Pommerenke, C.; Pradella, R.; Pukall, R.; Rabus, R.; Stackebrandt, E.; Thole, S.; Thompson, L.; Tielen, P.; Tomasch, J.; von Jan, M.; Wanphrut, N.; Wichels, A.; Zech, H.; Simon, M. The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: A hitchhiker's guide to life in the sea. *ISME J.* **2010**, *4*, 61–77.

(40) (a) Seyedsayamdost, M. R.; Carr, G.; Kolter, R.; Clardy, J. Roseobacticides: Small molecule modulators of an algal-bacterial symbiosis. *J. Am. Chem. Soc.* **2011**, *133*, 18343–18349. (b) Seyedsayamdost, M. R.; Wang, R.; Kolter, R.; Clardy, J. Hybrid biosynthesis of roseobacticides from algal and bacterial precursor molecules. *J. Am. Chem. Soc.* **2014**, *136*, 15150–15153. (c) Wang, R.; Seyedsayamdost, M. R. Roseochelin B, an algaecidal natural product synthesized by the *Roseobacter Phaeobacter inhibens* in response to algal sinapic acid. *Org. Lett.* **2017**, *19*, 5138–5141. (d) Wang, J.; Gallant, É.; Wilson, M. Z.; Wu, Y.; Li, A.; Gitai, Z.; Seyedsayamdost, M. R. Algal *p*-coumaric acid induces oxidative stress and siderophore biosynthesis in the bacterial symbiont *Phaeobacter inhibens. Cell Chem. Biol.* **2021**, *29*, 1–10.

(41) (a) Chhetri, B. K.; Lavoie, S.; Sweeney-Jones, A. M.; Kubanek, J. Recent trends in the structural revision of natural products. *Nat. Prod. Rep.* **2018**, *35*, 514–531. (b) Yoo, H.-D.; Nam, S.-J.; Chin, Y.-W.; Kim, M.-S. Misassigned natural products and their revised structures. *Arch. Pharm. Res.* **2015**, *39*, 143–153. (c) Nicolaou K. C.; Snyder, S. A. Chasing molecules that were never there: Misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem. Int. Ed.* **2005**, *44*, 1012–1044. (d) Chen, Z.; Rizzacasa, M. A. 2019 highlights of the structural revision of natural product via total synthesis. *Front. Chem. Sci. Eng.*

2021, *15*, 595–601. (e) Lapinskaite, R.; Malatinec, Š.; Mateus, M.; Rycek, L. Cross-coupling as a key step in the synthesis and structure revision of the natural products selagibenzophenones A and B. *Catalysts* **2021**, *11*, 708. (f) Tian, X.; Jaber, J. J.; Rychnovsky, S. D. Synthesis and structure revision of calyxin natural products. *J. Org. Chem.* **2006**, *71*, 3176–3183. (g) Ren, F.-C.; Wang, L.-X.; Lv, Y.-F.; Hu, J.-M.; Zhou, J. Structure revision of four classes of prenylated aromatic natural products based on a rule for diagnostic ^{13}C NMR chemical shifts. *J. Org. Chem.* **2021**, *86*, 10982–10990. (h) Greco, C.; Mattos-Shipley, K.; Bailey, A. M.; Mulholland, N. P.; Vincent, J. L.; Willis, C. L.; Cox, R. J.; Simpson, T. J. Structure revision of cryptosporioptides and determination of the genetic basis for dimeric xanthone biosynthesis in fungi. *Chem. Sci.* **2019**, *10*, 2930–2939.

(42) (a) Colis, L. C.; Woo, C. M.; Hegan, D. C.; Li, Z.; Glazer, P. M.; Herzon, S. B. The cytotoxicity of (–)-lomaiviticin A arises from inductioni of double-strand breaks in DNA. *Nat. Chem.* **2014**, *6*, 504–510. (b) Janso, J. E.; Haltli, B. A.; Eustáquio, A. S.; Kulowski, K.; Waldman, A. J.; Zha, L.; Nakamura, H.; Bernan, V. S.; He, H.; Carter, G. T.; Koehn F. E.; Balskus, E. P. Discovery of the lomaiviticin biosynthetic gene cluster in *Salinispora pacifica. Tetrahedron* **2014**, *70*, 4156–4164. (c) Kersten, R. D.; Lane, A. L.; Nett, M.; Richter, T. K. S.; Duggan, B. M.; Dorrestein, P. C.; Moore, B. S. Bioactivity-guided genome mining reveals the lomaiviticin biosynthetic gene cluster in *Salinispora tropica. ChemBioChem* **2013**, *14*, 955–962.

(43) He, H.; Ding, W.- D.; Berman, V. S.; Richardson, A. D.; Ireland, C. M.; Greenstein, M.; Ellestad, G. A.; Carter, G. T. Lomaiviticins A and B, potent antitumor antibiotics from *Micromonospora lomaivitiensis. J. Am. Chem. Soc.* **2001**, *123*, 5362–5363.

(44) Woo, C. M.; Beizer, N. E.; Janso, J. E.; Herzon, S. B. Isolation of lomaiviticins C–E, transformation of lomaiviticin C to lomaiviticin A, complete structure elucidation of lomaiviticin A, and structure-activity analyses. *J. Am. Chem. Soc.* **2012**, *134*, 15285–15288.

(45) Fujimoto, H.; Ikeda, M.; Yamamoto, K.; Yamazaki, M. Structure of fischerin, a new toxic metabolite from an ascomycete, *Neosartorya fischeri* var. *fischeri*. *J. Nat. Prod.* **1993**, *56*, 1268–1275.

(46) Amini, S. K. Assignment of the absolute configuration of fischerin by computed NMR chemical shifts. *J. Struct. Chem.* **2015**, *56*, 1334–1341.

(47) (a) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 146–147. (b) Pinzi, L.; Rastelli, G. Molecular docking: Shifting paradigms in drug discovery. *Int. J. Mol. Sci.* **2019**, *20*, 4331. (c) Mengist, H. M.; Dilnessa, T.; Jin, T. Structural basis of potential inhibitors targeting SARS-CoV-2 main protease. *Front. Chem.* **2021**, *9*, 622898. (d) Leissing, T. M.; Luh, L. M.; Cromm, P. M. Structure driven compound optimization in targeted protein degradation. *Drug Discov. Today Technol.* **2020**, *37*, 73–82.

(48) (a) Wan, Q. L.; Fu, X.; Dai, W.; Yang, J.; Luo, Z.; Meng, X.; Liu, X.; Zhong, R.; Yang, H.; Zhou, Q. Uric acid induces stress resistance and extends the life span through activating the stress response factor DAF-16/FOXO and SKN-1/NRF2. *Aging* **2020**, *12*, 2840–2856. (b) Kenyon, C. J. The genetics of ageing. *Nature* **2010**, *464*, 504–512. (c) Patel, D. S.; Garza-Garcia, A.; Nanji, M.; McElwee, J. J.; Ackerman, D.; Driscoll, P. C.; Gems, D. Clustering of genetically defined allele classes in the *Caenorhabditis elegans* DAF-2 insulin/IGF-1 receptor. *Genetics* **2008**, *178*, 931–946.

(49) Chhabra, N.; Aseri, M. L.; Padmanabhan, D. A review of drug isomerism and its significance. *Int. J. Appl. Basic Med. Res.* **2013**, *3*, 16–18.

(50) Molina, N.; Vettore, O.; Lorenzetti, B. B.; Ferreira, S. H. The peripheral analgesic effect of morphine, codeine, pentazocine and D-propoxyphene. *Braz. J. Med. Biol. Res.* **1983**, *16*, 345–352.

(51) Murphy, P. J.; Nickander, R. C.; Bellamy, G. M.; Kurtz, W. L. Effect of L-propoxyphene on plasma levels and analgesic activity of D-propoxyphene in the rat. *J. Pharmacol. Exp. Ther.* **1976**, *199*, 415–422.

(52) Agranat, L.; Caner, H.; Cadwell, J. Putting chirality to work: The strategy of chiral switches. *Nat. Rev. Drug Discov.* **2002**, *1*, 753–768.

(53) Kim, J.-C.; Lee, Y.-W.; Tamura, H.; Yoshizawa, T. Sambutoxin: A new mycotoxin isolated from *Fusarium sambucinum*. *Tetrahedron Lett.* **1995**, *36*, 1047–1050.

(54) Jessen, H. J.; Gademann, K. 4-Hydroxy-2-pyridone alkaloids: Structures and synthetic approaches. *Nat. Prod. Rep.* **2010**, *27*, 1168–1185.

(55) Danelius, E.; Halaby, S.; van der Donk, W. A.; Gonen, T. MicroED in natural product and small molecule research. *Nat. Prod. Rep.* **2021**, *38*, 423–431.

(56) Inokuma, Y.; Yoshioka, S.; Ariyoshi, J.; Arai, T.; Hitora, Y.; Takada, K.; Matsunaga, S.; Rissanen, K.; Fujita, M. X-ray analysis on the nanogram to microgram scale using porous complexes. *Nature* **2013**, *495*, 461–466.

(57) Hoshino, M.; Khutia, A.; Xing, H.; Inokuma, Y.; Fujita, M. The crystalline sponge method updated. *IUCrJ* **2016**, *3*, 139–151.

(58) Tyler, A. R.; Ragbirsingh, R.; McMonagle, C. J.; Waddell, P. G.; Heaps, S. E.; Steed, J. W.; Thaw, P.; Hall, M. J.; Probert, M. R. Encapsulated nanodroplet crystallization of organic-soluble small molecules. *Chem* **2020**, *6*, 1755–1765.

(59) Brázda, P.; Palatinus, L.; Babor, M. Electron diffraction determines molecular absolute configuration in a pharmaceutical nanocrystal. *Science* **2019**, *364*, 667–669.

(60) (a) Gallagher-Jones, M.; Ophus, C.; Bustillo, K. C.; Boyer, D. R.; Panova, O.; Glynn, C.; Zee, C.- T.; Ciston, J.; Mancia, K. C.; Minor, A. M.; Rodriguez, J. A. Nanoscale mosaicity revealed in peptide microcrystals by scanning electron nanodiffraction. *Commun. Biol.* **2019**, *2*, 26. (b) Gallagher-Jones, M.; Bustillo, K. C.; Ophus, C.; Richards, L. S.; Ciston, J.; Lee, S.; Minor, A. M.; Rodriguez, J. A. Atomic structures determined from digitally defined nanocrystalline regions. *IUCrJ* **2020**, *7*, 490–499.

CHAPTER TWO

Identification of Uric Acid Gluconucleoside–Ascaroside Conjugates in *Caenorhabditis elegans* by Combining Synthesis and MicroED

Adapted from: Brian J. Curtis, Lee Joon Kim, Chester J. J. Wrobel, James M. Eagan, Rubin A. Smith, Jessica E. Burch, Henry H. Le, Alexander B. Artyukhin, Hosea M. Nelson, Frank C. Schroeder *Org. Lett.* **2020**, *22*, 6724–6728.

2.1 Abstract

Few nucleoside- or uric acid-derived natural products have been isolated from animals, despite the ubiquity of these small molecules in living organisms. In this work, a synthetic approach is used in combination with microcrystal electron diffraction (MicroED) to determine the structures of various N^3-(β-glucopyranosyl)uric acid derivatives in *Caenorhabditis elegans.* Additionally, a shortend synthetic route is reported for these noncanonical gluconucleosides incorporating an ascaroside moiety. Phosphorylated gluconucleosides is produced due to evolutionarily conserved insulin signaling pathway.

2.2 Introduction

The nematodes *Caenorhabditis elegans* and *Pristionchus pacificus* are model systems extensively studied for physiological and developmental biology due to their genetic manipulability, well-characterized genome and developmental program, and shoft and prolific life cycles.[1–4] Metabolomic studies and mutant screening of these organisms have led to the discovery of a large, intricate signaling network of small molecules involved in the regulation of aging, behavior, and development.[5,6] These molecules are biochemically assembled using building blocks from all major primary metabolic pathways. These pathways include nucleoside metabolism, demonstrating that this metabolic process is important not only for the biosynthesis of canonical nucleosides and different types of RNA, but also for broader biological relevance.

An example of a signaling molecule derived from nucleosides is npar#1 (**2.1**), which contains a xyloadenoside and a glycoside of paratose (Figure 2.1A). npar#1 (**2.1**) and other ascaroside derivatives are key signaling molecules in regulating development in *P. pacificus* and *C. elegans*.[5,6] Comparitive metabolomics of peroxisomal *β*-oxidation (p*β*o) mutants and wild-type *C.* elegans[7] uncovered additional nucleoside derivatives such as nuclas#3 (**2.2**) and pugl#1 (**2.3**). Of these gluconucleoside derivatives from the p*β*o mutant screen, those that appeared to incorporate uric acid was chosen for study: uglas#1 (**2.4**) and its phosphorylated analogue uglas#11 (**2.5**), which were proposed to feature an N^9-linked uric acid moiety in analogy to pugl#1 (**2.3**). Even though uric acid is a product of purine degradation and is a common small molecule in biology, only a small number of nucleoside derivatives has been known to incorporate uric acid, with the only molecule being N^3-(ribofuranosyl)uric acid **2.6** isolated in the 1960s from bovine blood.[8–11] Despite the lack of uric acid incorporating natural products, uric acid is a biologically significant molecule due to its ability to serve as an antioxidant to capture reactive oxygen species involved in various signaling pathways.[12] Interestingly, uric acid supplementation was found to increase the life span of *C. elegans*, and long-lived mutants of the insulin/IGF-1 receptor *daf-2* exhibited an upregulation of the uglas#11 production while the levels of other ascarosides and putative building blocks remained unchanged (Figure 2.1B).[13–15]

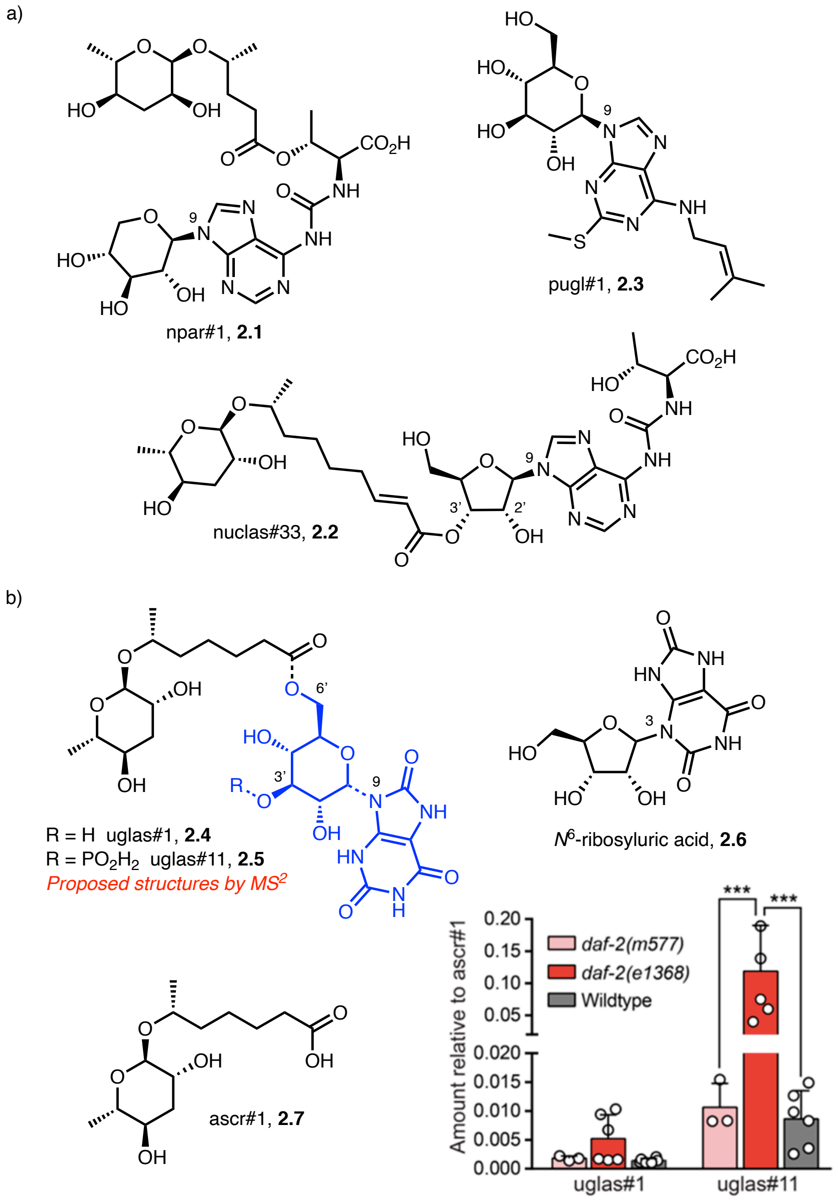


Figure 2.1 Nucleoside derivatives from nematodes. (a) Unusual purine nucleoside and gluconucleosides from *C. elegans* and other nematodes. (b) Proposed structures of uglas#1 and uglas#11 based on MS^2 analysis and analogy to other glucosides and gluconucleosides in *C. elegans,*[7] with dashed bonds referring to tentative assignments. *N*^{3}-(Ribofuranosyl)uric acid (**2.6**) has been reported from bovine blood.[8–11] Amounts of uglas#1 and uglas#11 relative to ascr#1 (**2.7**) were quantified in the wild type and mutants of the insulin/IGF-1 receptor *daf-2*. ****p* < 0.001.

2.3 Structural Characterization *via* Synthesis and MicroED

While the structural motifs of uglas#1 (**2.4**) are proposed to include uric acid, glucose, and an ascaroside moiety based on mass spectrometry, the connectivity of these moieties could not be elucidated as this metabolite was unable to be isolated in sufficient quantity for NMR spectroscopic characterization.[7] To overcome this challenge and determine the structure of **2.4**, a synthetic approach was used. Based on previous literature on guanine glucosylation, silyl Hilbert–Johnson glucosylation of **2.8** with protected glucose **2.9** was expected to forge the thermodynamically favored N^9 attachment.[16] However, the product mixture consisted of three regioisomers (Figure 2.2A), with the major product representing the N^3 regioisomer **2.10** and two minor products that were N^1-substituted (**2.11**) and N^7-substituted (**2.12**). Aminolysis of **2.10–12**, which produced samples of the three glucosyl uric acid isomers, and subsequent HPLC-HRMS comparison with *C. elegans exo*-metabolome samples indicated that the N^3 regioisomer **2.13** (named gluric#1) is a major glucosyluric acid metabolite (Figure 2.2B).

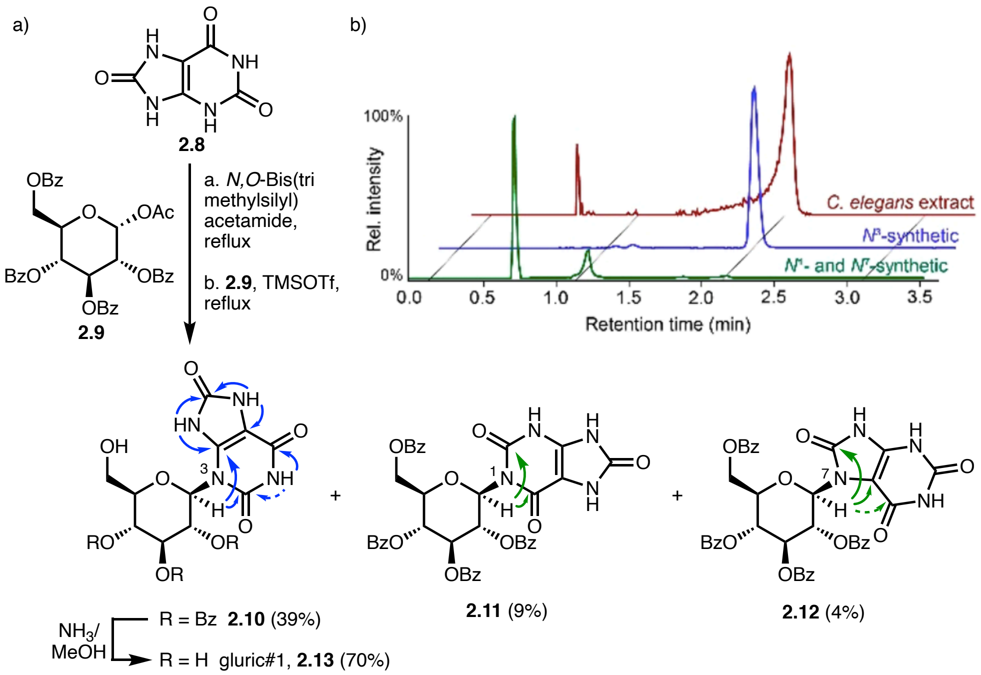
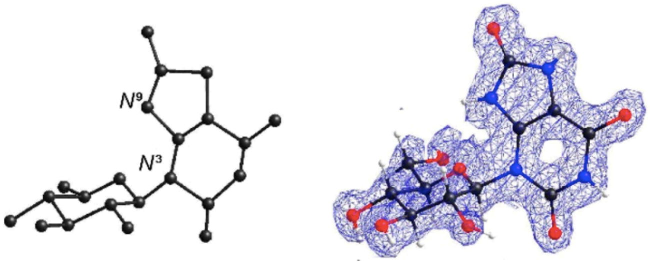


Figure 2.2 Synthesis of gluric#1. (a) Silyl Hilbert–Johnson glucosylation of uric acid (**2.8**) favors N^3 substitution (**2.10**) with the N^1 (**2.11**) and N^7 (**2.12**) isomers as minor products. Arrows indicate relevant HMBC correlations, with the weak correlations shown as dashed lines. (b) Comparison with *C. elegans exo*-metabolome by HPLC–HRMS revealed **2.13** as the major glucosyluric metabolite in *C. elegans*.

Due to the broad line shapes of the N–H protons in the HMBC spectra, N-linkage was unable to be assigned unambiguously using 2D-NMR. To confirm the identity of metabolite **2.13**, microcrystal electron diffraction (MicroED) was used. MicroED has been recently demonstrated as a powerful cryoelectron microscopy technique for elucidating structures of small molecules, but has remained largely underutilized in natural product chemistry.[17,18] Synthetic gluric#1 (**2.13**) was isoalted by reversed-phase flash chromatography to afford seemingly amorphous white powder. This powder sample was deposited on a Quantifoil holey-carbon TEM grid, and screening under the electron beam revealed numerous microcrystalline domains, exhibiting prisms 1–3 μm long. Continuous rotation selected area diffraction patterns were recorded from 104 crystals, and owing to the low symmetry of P 1, four data sets were merged to provide a high-resolution (1.0 Å) and high-completeness (87%) solution using direct methods (R_1 = 14.0%). Importantly, the preliminary structure solution, obtained directly from SHELXD[19] without any human input other than the molecular formula, clearly confirmed N^3 glucosylation, despite lacking some atoms on the sugar fragment (Figure 2.3). Refinement of this structure allowed for placement of all atoms of gluric#1 (**2.13**), providing an unambiguous confirmation of its structure.



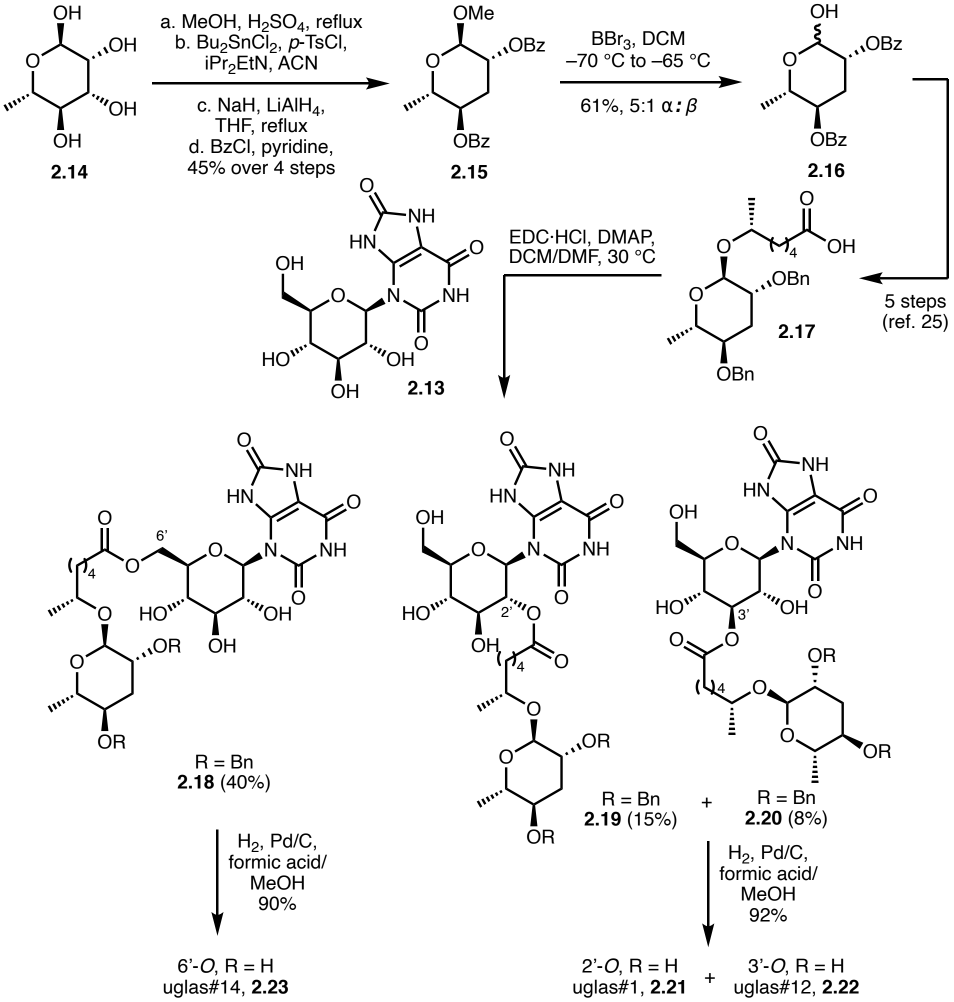
Preliminary structure

1.0 Å resolution, R1 = 14.0%, GooF = 1.223

Figure 2.3 Initial MicroED structural solution of **2.13** directly from SHELXD[19] (left) and final refined structure with electron potential map overlaid (right).

The observed preference for N^3 glucosylation is consistent with calculated π-electron densities at purine nitrogen atoms, as well as previously reported N^3 glucosylation and ribosylation of uric acid using a silver-based silyl Hilbert–Johnson method.[20,21] These results also suggest that uglas#1 and uglas#11 are also N^3-linked, as opposed to the original proposal[7] and canonical N^9-linked purine nucleosides in DNA and RNA; instead, these nucleosides are in agreement with the connectivity observed in the previously reported N^3-linked **2.6**.[8–11]

The synthetic route to the proposed 2,4-dideoxyglycoside (ascaroside) moiety in uglas#1 and uglas#11 (Scheme 2.1) began by methylation of α-L-rhamnose **2.14**, followed by tosylation to afford a 3-O-tosylated intermediate.[22,23] This intermediate was then reduced and subsequently benzoylated to yield **2.15**. Formation of a branched-chain deoxyfuranose side product derived from ring contraction, previously reported for similar deoxygenation reactions,[24] was prevented by optimizing reaction conditions. Demethylation of **2.15** led to the target precursor **2.16**, which was converted to dibenzylated ascaroside **2.17** as previously described.[25]



Scheme 2.1 Nonselective synthesis of uglas#1 isomers.

The position of the ascaroside side chain attachment to *N*^{3}-(*β*-glucopyranosyl)uric acid (**2.13**) was probed using a nonselective approach. Steglich esterification of dibenzylated ascr#1 (**2.17**)[25] and gluric#1 (**2.13**) led to a mixture of three dibenzylated uglas#1 isomers. The product mixture was found to contain the 6'-O-substituted isomer (**2.18**) as the major product and two minor isomers derived from acylation of secondary alcohols. The ^{1}H NMR signals of one of the two minor isomers exhibited extreme line broadening of the glucose-attached protons, preventing full NMR spectroscopic characterization. However, double quantum-filtered COSY (dqfCOSY) spectra was processed to identify 2'-O-acylation product **2.19**. The other minor isomer was identified as the 3'-O isomer **2.20**. Hydrogenation of the three O-substituted isomers **2.18–20** provided three uglas#1 isomers, and comparison of their retention times and MS^2 spectra with those of natural uglas#1 revealed that the synthetic 2'-O isomer **2.21** matched the natural metabolite. Additionally, small amounts of the other two isomers **2.22** and **2.23**, named uglas#12 and uglas#14, were detected in *C. elegans endo*-metabolome samples in minor amounts (Figure 2.4).

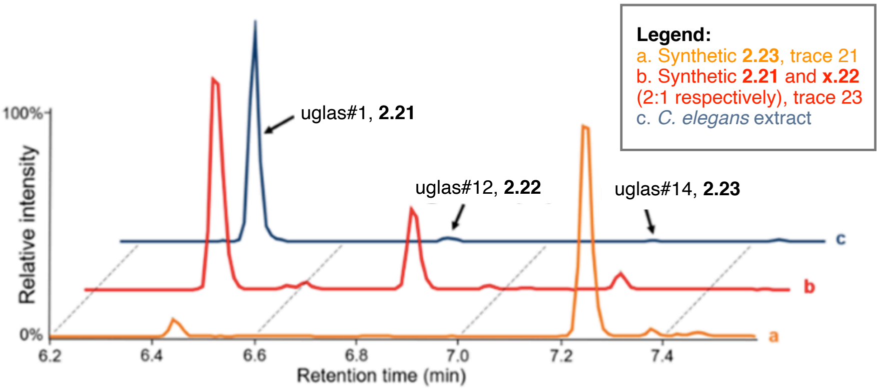
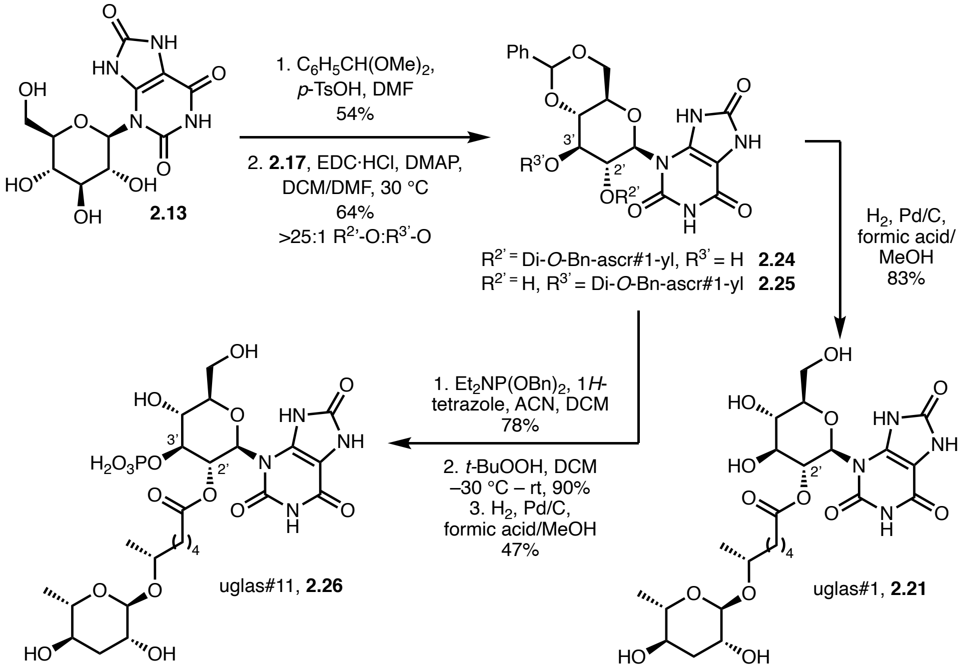


Figure 2.4 HPLC-HRMS ion chromatograms showing the elution of synthetic and natural uglas#1 isomers.

To selectively synthesize the major natural isomer, 2'-O-acylated **2.21**, gluric#1 (**2.13**) was protected and coupled with dibenzylated ascr#1 (**2.17**) *via* Steglich esterification to afford 2'-O-acylated isomer **2.24** and a trace amount of 3’-O isomer **2.25** (Scheme 2.2). Subsequent hydrogenation yielded uglas#1 (**2.21**) in high yield. Using the benzylidene derivative **2.24**, which exhibited better line shapes in the ^{1}H NMR spectra, the earlier assignment of uglas#1 as the 2’-O-acylated isomer was further confirmed. Phosphitylation of **2.24** followed by oxidation with *t*-butyl hydroperoxide selectively yielded the protected 3’-O-phosphate over two steps. Deprotection of benzyl and benzylidene moieties resulted in uglas#11 (**2.26**), whose HPLC retention times and MS^2 spectra were in good agreement with those of natural uglas#11 in the *C. elegans* extract. Using the synthetic uglas#11 as a standard, uglas#11 concentration in the wild type and *daf-2(e1368)* mutant worms were found to range from ~25 nM to ~250 nM, respectively.



Scheme 2.2 Selective synthesis of uglas#1 and uglas#11.

2.4 Conclusion

Taken together, these results elucidate the connectivity of the building blocks in the uglas family of metabolites. This family of metabolites exhibits connectivity that deviates from that found in canonical nucleosides and previously identified *C. elegans* metabolites. Furthermore, a simplified route for the synthesis of protected ascarylose (**2.16**) is reported, greatly facilitating the synthesis of diverse ascaroside-derivatives involved in signaling pathways.

While NMR spectroscopy provides sufficient characterization for a majority of newly discovered natural products, the cases presented in this work demonstrate the challenges that can arise from broad line shapes or the absence of non-exchangeable protons near heteroatomatic cores in NMR spectroscopy. Moreover, growing crystals large enough for single-crystal X-ray crystallography is frequently not feasible. These challenges are addressed using MicroED, which provided unambiguous confirmation of N^3 substitution of **2.13** that was tentatively assigned from inference using solution-state NMR spectra of the precursor intermediate **2.10**. The ability to obtain crystal structures from microcrystals of natural products without formal crystallization efforts will greatly facilitate identification of various novel metabolites.

The work presented in this study is in agreement with earlier reports that natural glycosylated uric acid derivatives represent the N^3-linked isomers[8–11] in contrast to many other nucleoside derivatives obtained from animals and the *C. elegans* metabolite pugl#1 (**2.3**). Notably, this report represents the first example of a N^3-glucosyluric acid moiety obtained from natural sources, despite the ubiquity of the building blocks uric acid and glucose. The most abundant members of the uglas family, uglas#1 (**2.21**) and uglas#11 (**2.26**), exhibit 2’-O attachment of the ascaroside side chain, unlike all previously identified ascarosyl glucosides such as glas#10 (**2.28**) or iglas#1 (**2.29**),[7,26] demonstrating the diversity of substitution specificity and suggesting that uglas#1 (**2.21**) and its related metabolites are produced by highly specific biosynthetic pathways. These results facilitate ongoing studies investigating the roles of uglas#1, uglas#11, and their related metabolites in signaling pathways such as insulin signaling.[14]

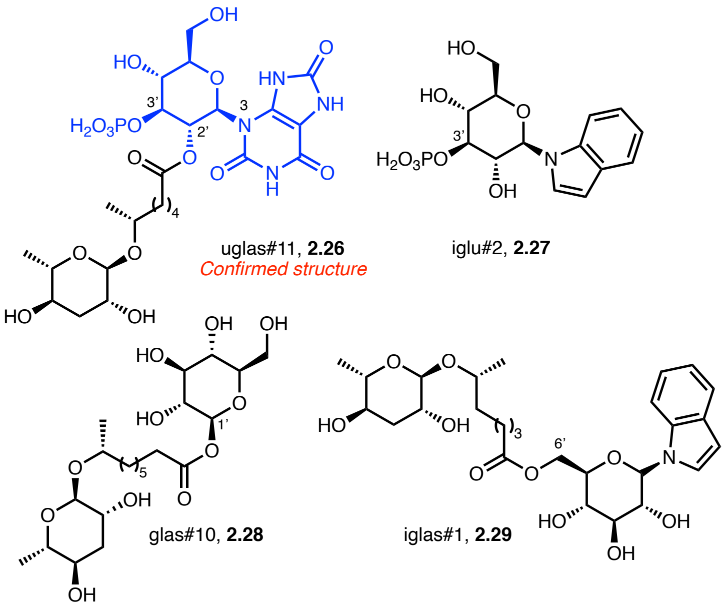


Figure 2.5 Examples of modular glucosyl metabolites/ascaroside *O*-glycoside esters produced by *C. elegans* with the indicated variation in substitution specificity.

2.5 Notes and References

(1) Bose, N.; Ogawa, A.; von Reuss, S. H.; Yim, J. J.; Ragsdale, E. J.; Sommer, R. J.; Schroeder, F. C. Complex small-molecule architectures regulate phenotypic plasticity in a nematode. *Angew. Chem.* **2012**, *124*, 12606−12611.

(2) Kaletta, T.; Hengartner, M. O. Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discovery* **2006**, *5*, 387−399.

(3) Bumbarger, D. J.; Riebesell, M.; Rödelsperger, C.; Sommer, R. J. System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes. *Cell* **2013**, *152*, 109−119.

(4) Bento, G.; Ogawa, A.; Sommer, R. J. Co-option of the hormone-signalling module dafachronic acid–DAF-12 in nematode evolution. *Nature* **2010**, *466*, 494−497.

(5) Von Reuss, S. H.; Schroeder, F. C. Combinatorial chemistry in nematodes: modular assembly of primary metabolism-derived building blocks. *Nat. Prod. Rep.* **2015**, *32*, 994−1006.

(6) Butcher, R. A. Natural products as chemical tools to dissect complex biology in *C. elegans. Curr. Opin. Chem. Biol.* **2019**, *50*, 138−144.

(7) Artyukhin, A. B.; Zhang, Y. K.; Akagi, A. E.; Panda, O.; Sternberg, P. W.; Schroeder, F. C. Metabolomic "dark matter" dependent on peroxisomal β-oxidation in *Caenorhabditis elegans. J. Am. Chem. Soc.* **2018**, *140*, 2841−2852.

(8) Hatfield, D.; Rinehart, R. R.; Forrest, H. S. 3-Ribosyluric acid. Part II. Isolation of the corresponding nucleotide from beef blood. *J. Chem. Soc.* **1963**, 899−902.

(9) Forrest, H. S.; Hatfield, D.; Lagowski, J. M. Uric acid riboside. Part I. Isolation and reinvestigation of the structure. *J. Chem. Soc.* **1961**, 963−968.

(10) Hatfield, D.; Forrest, H. S. Biosynthesis of 3-ribosyluric acid (uric acid riboside). *Biochim. Biophys. Acta* **1962**, *62*, 185−187.

(11) Lohrmann, R.; Lagowski, J. M.; Forrest, H. S. 3-Ribosyluric acid. Part III. Unambiguous syntheses of 3-ribosyluric acid and related compounds. *J. Chem. Soc.* **1964**, 451−459.

(12) Balaban, R. S.; Nemoto, S.; Finkel, T. Mitochondria, oxidants, and aging. *Cell* **2005**, *120*, 483−495.

(13) Wan, Q. L.; Fu, X.; Dai, W.; Yang, J.; Luo, Z.; Meng, X.; Liu, X.; Zhong, R.; Yang, H.; Zhou, Q. Uric acid induces stress resistance and extends the life spam through activating the stress response factor DAF-16/FOXO and SKN-1/NRF2. *Aging* **2020**, *12*, 2840−2856.

(14) Kenyon, C. J. The genetics of ageing. *Nature* **2010**, *464*, 504−512.

(15) Patel, D. S.; Garza-Garcia, A.; Nanji, M.; McElwee, J. J.; Ackerman, D.; Driscoll, P. C.; Gems, D. Clustering of genetically defined allele classes in the *Caenorhabditis elegans* DAF-2 insulin/IGF-1 receptor. *Genetics* **2008**, *178*, 931−946.

(16) Garner, P.; Ramakanth, S. A regiocontrolled synthesis of N7-and N9-guanine nucleosides. *J. Org. Chem*. **1988**, *53*, 1294−1298.

(17) Kahn, K.; Serfozo, P.; Tipton, P. A. Identification of the true product of the urate oxidase reaction. *J. Am. Chem. Soc*. **1997**, *119*, 5435−5442.

(18) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci*. **2018**, *4*, 1587−1592.

(19) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; Van Der Donk, W. A. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **2019**, *365*, 280−284.

(20) Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, *A64*, 112−122.

(21) Miyaki, M.; Shimizu, B. N→N alkyl and glycosyl migration of purines and pyrimidines. III. N →N alkyl and glycosyl migration of purine derivatives. *Chem. Pharm. Bull.* **1970**, *18*, 1446−1456.

(22) Birkofer, L.; Ritter, A. Neuere Methoden der präparativen organischen Chemie IV. Die Silylierung als Hilfsmittel in der organischen Synthese. *Angew. Chem.* **1965**, *77*, 414−426.

(23) Jeong, P. Y.; Jung, M.; Yim, Y. H.; Kim, H.; Park, M.; Hong, E.; Lee, W.; Kim, Y. H.; Kim, K.; Paik, Y. K. Chemical structure and biology activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **2005**, *433*, 541−545.

(24) Muramatsu, W.; Tanigawa, S.; Takemoto, Y.; Yoshimatsu, H.; Onomura, O. Organotin-catalyzed highly regioselective thiocarbonylation of nonprotected carbohydrates and synthesis of deoxy carbohydrates in a minimum number of steps. *Chem. - Eur. J.* **2012**, *18*, 4850−4853.

(25) Martinelli, M. J.; Vaidyanathan, R.; Van Khau, V. Selective monosulfonylation of internal 1,2-diols catalyzed by di-n-butyltin oxide. *Tetrahedron Lett.* **2000**, *41*, 3773−3776.

(26) Baer, H. H.; Astles, D. J.; Chin, H.-C.; Siemsen, L. The formation of branched-chain deoxypentofuranosides by ring contraction in the reductive desulfonyloxylation of hexopyranoside *p*-toluenesulfonates. *Can. J. Chem.* **1985**, *63*, 432−439.

(27) Zhang, Y. K.; Sanchez-Ayala, M. A.; Sternberg, P. W.; Srinivasan, J.; Schroeder, F. C. Improved synthesis for modular ascarosides uncovers biological activity. *Org. Lett.* **2017**, *19*, 2837−2840.

(28) Stupp, G. S.; Von Reuss, S. H.; Izrayelit, Y.; Ajredini, R.; Schroeder, F. C.; Edison, A. S. Chemical detoxification of small molecules by *Caenorhabditis elegans. ACS Chem. Biol.* **2013**, *8*, 309−313.

(29) Von Reuss, S. H.; Bose, N.; Srinivasan, J.; Yim, J. J.; Judkins, J. C.; Sternberg, P. W.; Schroeder, F. C. Comparative metabolomics reveals biogenesis of ascarosides, a modular library of small-molecule signals in *C. elegans. J. Am. Chem. Soc.* **2012**, *134*, 1817−1824.

2.6 Experimental Section

This section outlines the experimental procedure and statistics for electron diffraction studies. Al supplementary information, including procedures, figures, tables, and spectra, for isolation and synthesis of metabolites are reported in the adapted article.

2.6.1 MicroED Data and Statistics of 2.13

Quantifoil holey-carbon EM grids were placed in a dram vial with purified **2.13** and shaken lightly. Residual compound was removed by tapping lightly against the surface of a filter paper.

All diffraction data was collected on FEI Tecnai F200C electron microscope with an operating voltage of 200 keV, corresponding to a wavelength of 0.025 Å, using Gatan 626 cryo-holder under cryogenic temperature (100 K). During data acquisition, the crystal of interest was isolated using a selected area aperture and continuously rotated at a rate of -0.3°/s over a tilt range of 50–100°. Continuous rotation diffraction data was recorded using rolling shutter mode with a Ceta-D CMOS 4k x 4k camera, integrating at a rate of 3 s per frame and binning by 2 to produce final images of 2k x 2k.[1]

Diffraction movies saved as SER files were converted to SMV format using ser2smv software as described previously.[2] Frames were indexed and integrated in XDS.[3] Data from four crystals were scaled and merged together using XSCALE[4] to produce the final data set. Finally, intensities were converted to SHELX format using XDSCONV.[4]

The structure of **2.13** was solved *ab initio* using direct methods in SHELXD[5] and refined with SHELXL[6] in ShelXle. All non-hydrogen atoms were refined anisotropically, and hydrogen atoms were placed using the riding model.

Crystallographic information file (CIF) for compound **2.13** has been deposited at the Cambridge Crystallographic Data Center (Deposition Number: 2020283).

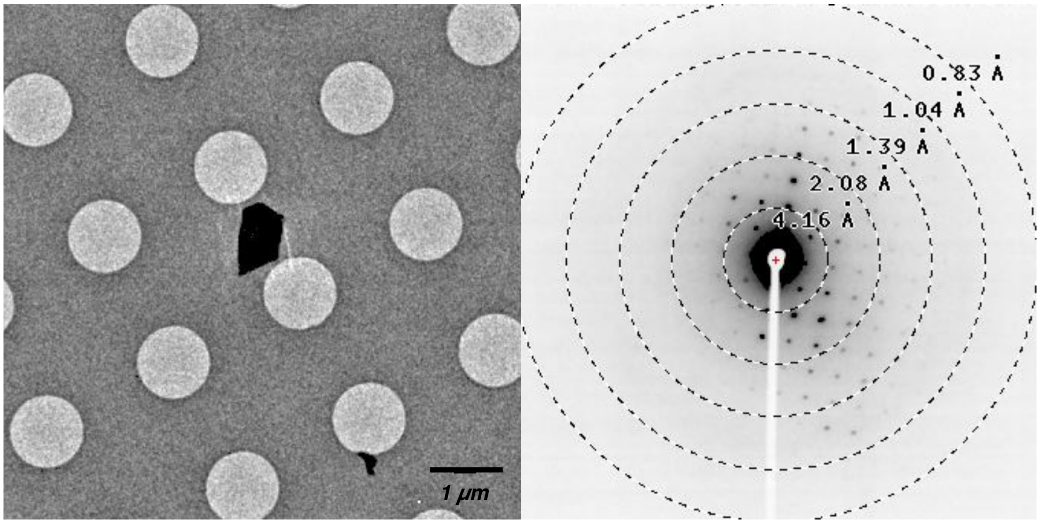


Figure 2.6 Gluric#1 microcrystal and diffraction pattern at SA 2600x magnification. Grid holes are 1 μm in diameter.

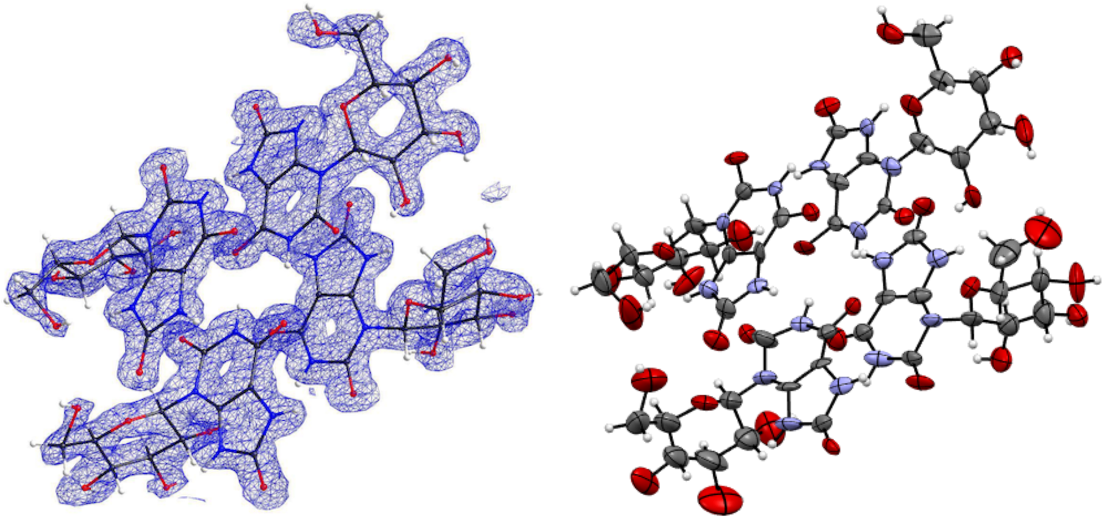


Figure 2.7 Asymmetric unit of gluric#1 (**2.13**) crystal structure (CCDC 2020283). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.36 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{11}H_{14}N_4O_8$
Formula weight	330.25
Temperature (K)	100(2)
Space group	P 1
Cell dimensions	

a, b, c (Å)	8.990(9), 9.730(2), 14.16(3)
α, β, γ (°)	93.05(9), 94.62(5), 108.54(5)
Resolution (Å)	1.0
Observed reflections	8144
Unique reflections	2122
Completeness (%)	87.1
R_{obs} (%)	15.5
R_{meas} (%)	17.7
I/σI	5.79
$CC_{1/2}$ (%)	95.4
R_1 (%)	14.01
wR_2 (%)	31.99
GooF	1.223

Table 2.1 Data collection and refinement statistics for **2.13**.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4999	Data was truncated to a resolution of 1.0 angstrom due to poor diffraction data quality above this threshold. The calculated value of sin(theta_max)/wavelength is 0.4999, which is still acceptable.
PLAT029_ALERT_3_A _diffrn_measured_fraction_theta_full value Low . 0.872 Why?	Full sampling of the reciprocal space was hindered by the limited stage tilt of the electron microscope stage and low symmetry (crystal system is triclinic).
PLAT089_ALERT_3_A Poor Data / Parameter Ratio (Zmax < 18) .... 3.05 Note	The structure has low symmetry and large number of non-hydrogen atoms to refine (92 atoms). The completeness of the data set is low due to limitations in stage rotation in MicroED.

PLAT415_ALERT_2_A Short Inter D-H..H-X H16A_1 ..H12A_3 . 1.88 Ang. x,-1+y,z = 1_545	These hydrogen atoms, placed using the riding model, are on alcohol groups that can rotate around.
PLAT415_ALERT_2_A Short Inter D-H..H-X H14_2 ..H16_4 . 1.48 Ang. x,-1+y,z = 1_545 Check	These hydrogen atoms, placed using the riding model, are on alcohol groups that can rotate around.
PLAT417_ALERT_2_A Short Inter D-H..H-D H12A_1 ..H17_4 . 1.41 Ang. x,y,-1+z = 1_554 Check	These hydrogen atoms, placed using the riding model, are on alcohol groups that can rotate around.
PLAT417_ALERT_2_A Short Inter D-H..H-D H13A_2 ..H13_3 . 1.48 Ang. -1+x,-1+y,1+z = 1_446 Check	These hydrogen atoms, placed using the riding model, are on alcohol groups that can rotate around.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.0475 Ang.	Diffraction data from four different crystals were merged to achieve a desirable completeness, resulting in some structural disordering.
PLAT360_ALERT_2_B Short C(sp3)-C(sp3) Bond C14_4 - C15_4 . 1.30 Ang.	Diffraction data from four different crystals were merged to achieve a desirable completeness, resulting in some structural disordering.
PLAT415_ALERT_2_B Short Inter D-H..H-X H14_1 ..H13A_4 . 1.96 Ang. x,-1+y,-1+z = 1_544 Check	H13A_4 is a hydrogen atom placed using riding model on an alcohol, which is free to rotate.
PLAT416_ALERT_2_B Short Intra D-H..H-D H9_1 ..H17_1 . 1.41 Ang. x,y,z = 1_555 Check	H17_1 is a hydrogen atom placed using riding model on an alcohol, which is free to rotate.
PLAT417_ALERT_2_B Short Inter D-H..H-D H11A_2 ..H9_4 . 1.91 Ang. x,y,z = 1_555 Check	Some structural disordering caused by merging of multiple data sets.
PLAT417_ALERT_2_B Short Inter D-H..H-D H17_2 ..H7_3 . 2.03 Ang. -1+x,-1+y,z = 1_445 Check	H17_2 is a hydrogen atom placed using riding model on an alcohol, which is free to rotate.
PLAT430_ALERT_2_B Short Inter D...A Contact O8_1 ..O15_3 . 2.83 Ang. x,y,z = 1_555 Check	Some structural disordering caused by merging of multiple data sets.

Table 2.2 Justification for alerts in gluric#1 (**2.13**) crystal structure (CCDC 2020283).

2.6.2 Supplementary Notes and References

(1) Nannenga, B. L.; Shi, D.; Leslie, A. G. W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930.

(2) Hattne, J; Reyes, F. E.; Nannenga, B. L.; Shi, D.; de la Cruz, M. J.; Leslie, A. G. W.; Gonen, T. MicroED data collection and processing. *Acta Cryst.* **2015**, *A71*, 353–360.

(3) Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.

(4) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010,** *D66*, 133–144.

(5) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(6) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(7) Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. ShelXle: a Qt graphical user interface for SHELXL. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.

CHAPTER THREE

Biosynthesis of the Fusarium Mycotoxin (–)-Sambutoxin

Adapted from: Eun Bin Go, Lee Joon Kim, Hosea M. Nelson, Masao Ohashi, Yi Tang *Org. Lett.* **2021**, *23*, 7819–7823.

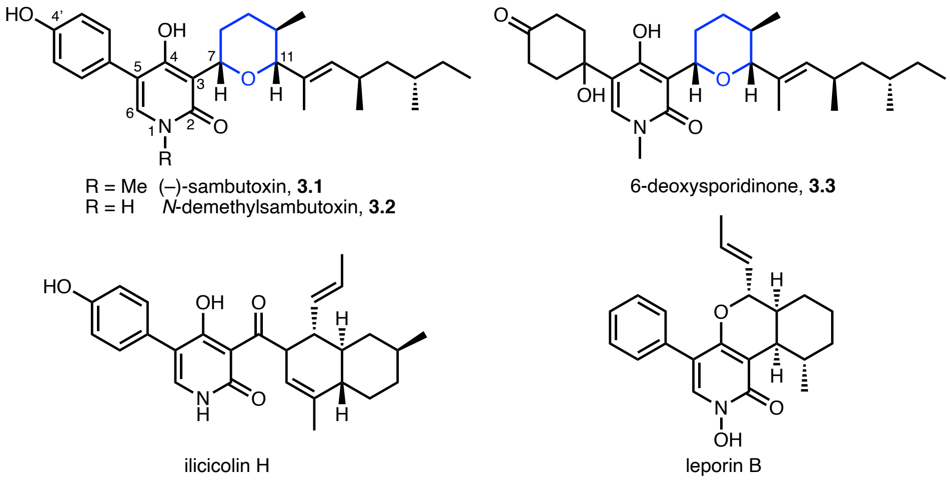
3.1 Abstract

4-Hydroxy-2-pyridone alkaloids have garnered synthetic and biosynthetic attention due to their impressive range of biological activities and structural diversity. In this work, the biosynthetic pathway of (–)-sambutoxin (**3.1**) is elucidated using genome mining approach, characterizing metabolic intermediates using a combination of NMR and MicroED. Notably, the enzymatic origin of the tetrahydropyran moiety characteristic of the 3-ether subfamily of 4-hydroxy-2-pyridone alkaloids is uncovered. Moreover, this study marks the first example of the *p*-hydroxyphenyl group installation *via* a late-stage, P450-catalyzed oxidation of the phenylalanine-derived side chain rather than the incorporation of tyrosine.

3.2 Introduction

4-Hydroxy-2-pyridone alkaloids are a family of fungal natural products that boasts a wide range of biological activities and structural diversity of their molecular framework.[1] This family of secondary metabolites can be classified into 3-acyl, 3-alkyl, and 3-ether subfamilies depending on their linkage at the 3-position of their 2-pyridone cores.[1] In contrast to the 3-acyl (e.g. ilicicolin H) and 3-alkyl subfamilies (e.g. leporins), the 3-ether subfamily remains understudied. This subfamily is characterized by the tetrahydropyran motif, as seen in the case of (–)-sambutoxin

(**3.1**) and its analogues such as *N*-demethylsambutoxin (**3.2**) and 6-deoxysporidinone (**3.3**, Figure 3.1).



R = Me (–)-sambutoxin, **3.1**
R = H *N*-demethylsambutoxin, **3.2**

6-deoxysporidinone, **3.3**

ilicicolin H

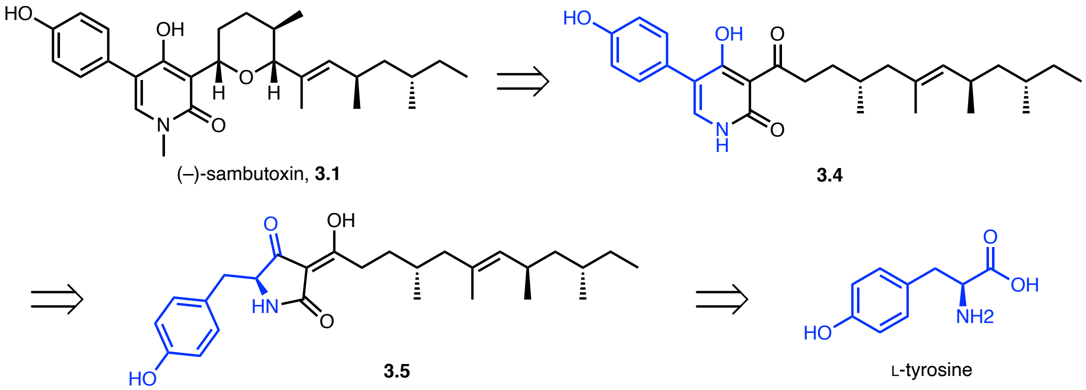
leporin B

Figure 3.1 4-Hydroxy-2-pyridone natural product (–)-sambutoxin (**3.1**) and related metabolites.

3.3 Elucidating the Biosynthesis of (–)-Sambutoxin

(–)-Sambutoxin (**3.1**) was first isolated from the potato parasite *Fusarium sambucinum* in 1995.[2] This natural product exhibits toxicity in chicken embryos and human tumor cells, and has been shown to cause hemorrhage in rats.[2–3] Due to its potent biological activity and the unique tetrahydropyranyl motif at the C3 position, **3.1** has been the target of total synthesis campaigns that led to the assignment of its absolute stereochemistry.[4]

Prior to experimental studies, a retrobiosynthesis of **3.1** was proposed based on its biosynthetic relationship to other 4-hydroxy-2-pyridone natural products (Scheme 3.1).[5–9] *N*-methylation is a modification seen only in **3.1** and its cometabolite 4-hydroxy-2-pyridones, and this reaction is proposed to be catalyzed by an *N*-methyltransferase. The tetrahydropyran motif is predicted to be formed by oxidative cyclization of the linear pyridone precursor **3.4**, whose structure is similar to known natural products such as aspyridone A and pretenellin B.[10–12] To forge **3.4**, P450-catalyzed ring expansion reaction of the tetramic acid **3.5** can occur. Tetramic acid **3.5** can be synthesized by polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) and *trans*-enoyl reductase (ER). Since **3.1** contains a hydroxyphenyl motif, the NRPS module of the PKS-NRPS can activate and amidate the polyketide acyl chain with L-tyrosine rather than L-phenylalanine, avoiding th need for an extra step in the biosynthesis to install the hydroxyl group.



Scheme 3.1 Proposed retrobiosynthesis of **3.1** featuring tyrosine incorporation.

Because the originally published strains capable of producing **3.1** were inaccessible, genome mining approach was used to identify other biosynthetic gene clusters (BGCs) of **3.1** in other *Fusarium* species from the National Center for Biotechnology Information (NCBI) fungal genome database. Protein sequences of the PKS-NRPS (LepA), *trans*-ER (LepG), and P450$_{\text{RE}}$ (LepH) from the biosynthetic pathway of a related natural product leporin B[5] (Figure 3.1) was used to search the NCBI genome database. Since (–)-sambutoxin (**3.1**) has an *N*-methyl group that leporin B lacks, the *N*-methyltransferase was added as a requirement to arrive at a well-conserved BGC in several species of *Fusarium* (Figure 3.2). These clusters also included genes that code for a short-chain dehydrogenase/reductase (SDR) as well as two additional P450 monooxygenases. Notably, one of the species these clusters were found in was *Fusarium oxysporum*, which has been previously reported as a source of **3.1** and its cometabolites **3.2** and **3.3**.[3]

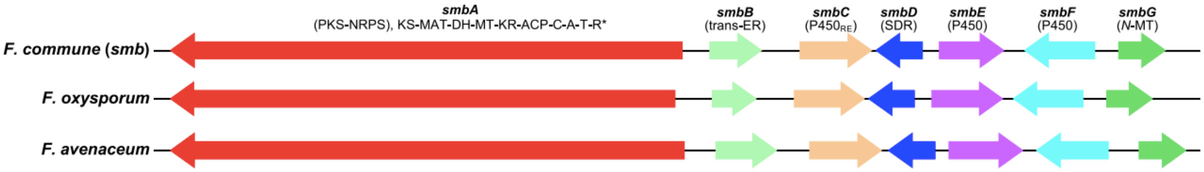
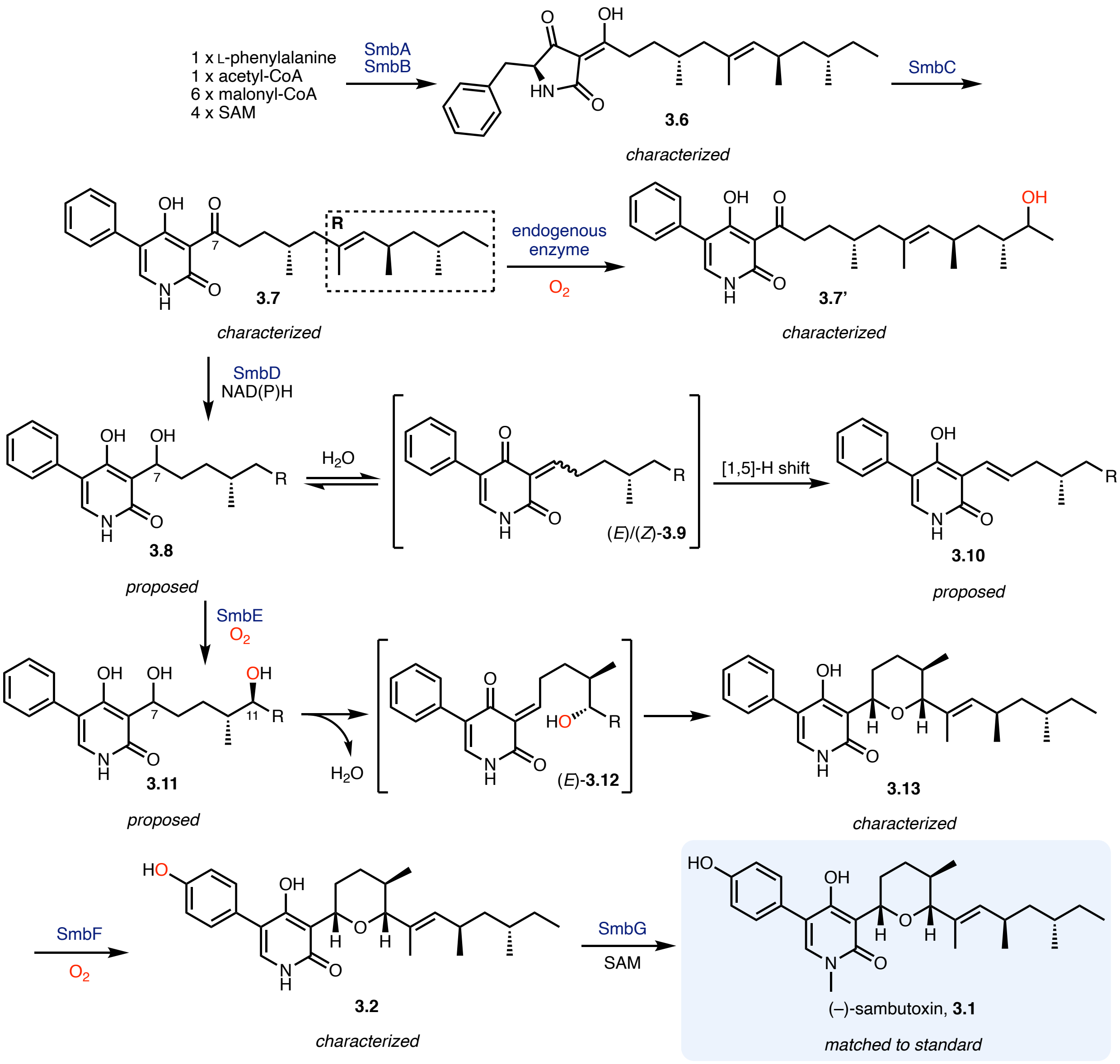


Figure 3.2. A conserved biosynthetic gene cluster from various species of *Fusarium* proposed to be involved in the biosynthesis of **3.1**.

To examine the products of this *Fusarium* BGC (named *smb*), the cluster was reconstituted from *F. commune*, whose genomic DNA was obtained. Then, the *smb* genes were heterologously expressed in the engineered strain of *Aspergillus nidulans* Δ*S*TΔ*E*M,[13] which provides experimental advantages due to its reduced endogenous metabolite background. Co-expression of the PKS-NRPS (SmbA) and *trans*-ER (SmbB) gave rise to a new compound at a titer of 2 mg/L. Isolation and structural characterization of this new compound after scaling up cultures led to the identification of this metabolite as the tetramic acid **3.6** (Scheme 3.2). While the first step of the biosynthesis does indeed involve the tetramic acid as initially proposed, metabolite **3.6** contained a phenyl ring rather than a *p*-hydroxyphenyl group. This discrepancy suggested that the tetramic acid is derived from phenylalanine instead of the expected tyrosine. No MS signal was observed for *m/z* of 442 $[M+H]^+$ corresponding to the tetramic acid **3.5** derived from tyrosine, indicating that SmbA exclusively accepts phenylalanine to derive the early stage intermediate towards **3.1**. This intermediate formation is consistent with the 2019 report from Õmura and Shiomi, who isolated fusaramin, which is a cometabolite of compounds **3.1**–**3** from *Fusarium* sp. In this preceding literature, fusaramin was found to be derived *via* hydroxylation of the C_β of phenylalanine in tetramic acid **3.6**.[14] Taken together, these results indicate that the biosynthesis of **3.1** starts with phenylalanine, and the *p*-hydroxy group is introduced at a later stage during the biosynthetic pathway.



Scheme 3.2. Proposed biosynthetic pathway for **3.1** based on stepwise reconstitution.

SmbC has sequence similarity to other characterized ring expansion P450s ($P450_{RE}$), and co-expression of this gene along with SmbA and SmbB resulted in the production of the next intermediate, which was structurally characterized to be the 2-pyridone ketone **3.7** (Scheme 3.2). This transformation is consistent with the functions of $P450_{RE}$ enzymes in other 2-pyridone alkaloid pathways.[5–7,15] A cometabolite **3.7'** with +16 mu was also present in the product mixture, and **3.7'** was isolated and characterized to contain a secondary alcohol in the polyketide chain of

3.7. The addition of a hydroxyl group may be due to the endogenous alkyl hydroxylases present in *A. nidulans*, and this activity was observed in other reconstitution studies.[16,17]

As proposed initially, 2-pyridone ketone **3.7** can serve as a starting material for the series of reactions to forge the tetrahydropyran motif. In the biosynthesis of other cyclic 2-pyridone natural products such as leporins, the C7 ketone is reduced by an SDR to an alcohol, followed by dehydration to yield a reactive *o*-quinone methide (*o*-QM), which can serve as a diene or a (di)enophile in pericyclic reactions.[5,6,8,18] Co-expressing SmbD with SmbA–C led to the formation of two new metabolites **3.8** and **3.10** with *m/z* 426 $[M+H]^+$ and 408 $[M+H]^+$, respectively (Scheme 3.2). While these metabolites were not isolated due to their low amounts and relative instability, metabolite **3.8** is proposed to be the C7 alcohol formed from the ketoreduction of **3.7**, and compound **3.10** is proposed to be the [1,5]-hydride shifted shunt product based on mass spectrometry data. This shunt product may be formed by dehydration of **3.8** to either the (*E*)- or (*Z*)-*o*-QM **3.9**, followed by a [1,5]-hydride shift observed in the synthetic study of 2-pyridones.[19] To support these tentative assignments, **3.7** was reduced with $NaBH_4$, following prior literature that used this method to verify SDR functions.[5,8] Chemical reduction of **3.7** led to the formation of the same products **3.8** and **3.10**, with consistent MS and UV profiles. This result supports that SmbD is a ketoreductase similar to its homologues that were previously reported.[5,8]

Based on the ability of *o*-QMs to serve as Michael acceptors in nucleophilic additions,[20,21] (*E*)-*o*-QM is proposed to act as a Michael acceptor for the tetrahydropyran formation. Once a (*R*)-hydroxyl group at C11 is installed by one of the remaining P450 enzymes to forge (*E*)-**3.12**, the hydroxyl group could serve as an internal nucleophile, and a stereospecific Michael addition could take place due to the the bulky groups on C2 and C6 preferentially assuming the equatorial positions.[4,22] Co-expression of SmbE with SmbA–D led to the formation of a new metabolite **x.13** with *m/z* of 424 $[M+H]^+$ as **3.8** was consumed (titer of 3.7 mg/L, Scheme 3.2). NMR analysis of the compound **3.13** revealed the desired tetrahydropyran motif found in **3.1**. To further support the NMR analysis of this natural product intermediate and confirm the relative stereochemistries of the tetrahydropyran and methyl groups in the polyketide chain, microcrystal electron diffraction (MicroED) was used to probe its structure. MicroED enables crystallographic characterization of small molecules from microcrystals, eliminating the need for arduous crystallization to produce single crystals amenable for single crystal X-ray crystallography.[23] Microcrystals of metabolite **3.13** were generated by slow evaporation of a pure HPLC fraction containing **3.13** in a mixture of acetonitrile and water. Diffraction data were merged to afford a high completeness data set, and *ab initio* structure was solved using direct methods (Figure 3.3). MicroED structure revealed the relative stereochemistry of **3.13** identical to that reported for (–)-**3.1**, confirming that **3.13** is likely to be on path for the biosynthesis of **3.1** and that SmbE forges the tetrahydropyran motif characteristic of the 3-ether subfamily of 4-hydroxy-2-pyridone alkaloids *via* an intramolecular Michael addition.

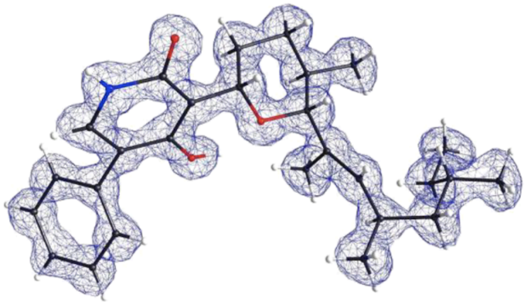


Figure 3.3. MicroED structure of **3.13**.

Up through intermediate **3.13**, the aromatic ring at C4’ has remained an unoxidized phenyl group as opposed to the initially proposed biosynthetic pathway. The remaining enzymes in the cluster were another P450 (SmbF) and an *N*-MT (SmbG), suggesting that SmbF would oxidize the C4’ position of the phenyl ring to give rise to intermediate **3.2**. Co-expression of SmbF with

SmbA–E led to the production of a new compound (titer of 2 mg/L) that was confirmed to be *N*-demethylsambutoxin (**3.2**, Scheme 3.2). Based on this result, SmbF can be assigned as the C4'-hydroxylase in the biosynthetic pathway of **3.1**. The tetrahydropyran formation facilitated by SmbE must occur prior to the phenyl hydroxylation by SmbF, as co-expressing SmbF with SmbA–D in the absence of SmbE led to no activity.

The hydroxylation reaction of the phenyl ring catalyzed by SmbF can proceed through two possible mechanisms. In one proposed mechanism, abstraction of hydrogen from the weak amide N1–H bond can generate a free radical. This radical can be delocalized to C4' on the phenyl ring before reacting with the iron-bound hydroxyl radical in SmbF. In the second mechanism, epoxidation at C3'–C4' followed by 1,2-hydride (NIH) shift can form the hydroxyl group at C4' similar to the mechanism of cinnamate hydroxylase.[24] This late-stage C–H oxidation by SmbF is not seen in the biosynthesis of most other polyketide-nonribosomal peptide hybrid molecules: for tenellin, aspyridone, and illicicolin H, the *p*-hydroxyphenyl groups are installed in the beginning as a result of tyrosine incorporation by the PKS-NRPS.[6,9,10] A possible explanation for the incorporation of phenylalanine rather than tyrosine by SmbA may be due to the ability of some P450$_{\mathrm{RE}}$ enzymes to dephenylate tyrosine-containing tetramic acids.[8,25] Therefore, by delaying the phenyl group oxidation to a later stage in the pathway, intermediate **3.5** could be protected from dephenylation reaction to arrive at a successful biosynthesis of **3.1**.

Finally, the remaining *N*-methyltransferase (SmbG) can perform the *N*-methylation reaction to arrive at **3.1** from **3.2** (Scheme 3.2). Co-expression of SmbG with SmbA–F led to a small peak with *m/z* 454 $[\mathrm{M+H}]^+$ with the same retention time as a commercial standard of **3.1**. However, purifying sufficient amount of **3.1** for NMR analysis proved difficult, possibly due to an unknown degradation or detoxification pathways in the host organism. MicroED attempts also proved unsuccessful, presumably due to the high number of impurities that prevented the formation of ordered microcrystals. Nonetheless, the recombinantly expressed SmbG was demonstrated to be capable of *N*-methylating **3.2** *in vitro* ($K_M = 54.7 \pm 11.4$ μM, $k_{cat} = 1.70 \pm 0.13$ min^{-1}, and $k_{cat}/K_M = 3.1 \times 10^4$ $min^{-1}M^{-1}$) to give **3.1** in the presence of *S*-adenosylmethionine (SAM). SmbG was found to specifically methylate **3.2**, as it was inactive on intermediate **3.13**. These results confirmed that the *N*-methylation of **3.2** to **3.1** is the final step to forge the natural product **3.1** biosynthetically.

3.4 Conclusion

In conclusion, the biosynthetic pathway towards **3.1** has been uncovered, elucidating nature's strategy for forging the tetrahydropyran motif characteristic of the 4-hydroxy-2-pyridone alkaloid 3-ether subfamily. Notably, this study marks the first example of 4-hydroxy-2-pyridone alkaloid biosynthesis where the *p*-hydroxyphenyl group at C4' is derived from phenylalanine rather than tyrosine, and a late-stage oxidation is employed to install the hydroxyl group. This work also utilizes MicroED as a complementary technique to characterize natural product intermediates to confirm the relative stereochemistries of **3.13**. Taken together, full reconstitution of the biosynthesis of **3.1** enables genome-based mapping of fungi capable of making this toxic metabolite, and further investigations of related fungal natural products.

3.5 Notes and References

(1) Jessen, H. J.; Gademann, K. 4-Hydroxy-2-pyridone alkaloids: Structures and synthetic approaches. *Nat. Prod. Rep.* **2010**, *27*, 1168−1185.

(2) Kim, J.-C.; Lee, Y.-W.; Tamura, H.; Yoshizawa, T. Sambutoxin: A new mycotoxin isolated from *Fusarium sambucinum*. *Tetrahedron Lett.* **1995**, *36*, 1047−1050.

(3) Jayasinghe, L.; Abbas, H. K.; Jacob, M. R.; Herath, W. H. M. W.; Nanayakkara, N. P. D. *N*-Methyl-4-hydroxy-2-pyridinone analogues from *Fusarium oxysporum. J. Nat. Prod*. **2006**, *69*, 439−442.

(4) Williams, D. R.; Turske, R. A. Construction of 4-hydroxy-2-pyridinones. Total synthesis of (+)-sambutoxin. *Org. Lett*. **2000**, *2*, 3217−3220.

(5) Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y. SAM-dependent enzyme catalysed pericyclic reactions in natural product biosynthesis. *Nature* **2017**, *549*, 502−506.

(6) Zhang, Z.; Jamieson, C. S.; Zhao, Y.-L.; Li, D.; Ohashi, M.; Houk, K. N.; Tang, Y. Enzyme-catalyzed inverse-electron demand Diels−Alder reaction in the biosynthesis of antifungal ilicicolin H. *J. Am. Chem. Soc*. **2019**, *141*, 5659−5663.

(7) Zhang, Z.; Qiao, T.; Watanabe, K.; Tang, Y. Concise biosynthesis of phenylfuropyridones in fungi. *Angew. Chem., Int. Ed*. **2020**, *59*, 19889−19893.

(8) Ohashi, M.; Jamieson, C. S.; Cai, Y.; Tan, D.; Kanayama, D.; Tang, M.-C.; Anthony, S. M.; Chari, J. V.; Barber, J. S.; Picazo, E.; Kakule, T. B.; Cao, S.; Garg, N. K.; Zhou, J.; Houk, K. N.; Tang, Y. An enzymatic Alder-ene reaction. *Nature* **2020**, *586*, 64−69.

(9) Fisch, K. M. Biosynthesis of natural products by microbial iterative hybrid PKS−NRPS. *RSC Adv*. **2013**, *3*, 18228−18247.

(10) Wasil, Z.; Pahirulzaman, K. A. K.; Butts, C.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. One pathway, many compounds: Heterologous expression of a fungal biosynthetic pathway reveals its intrinsic potential for diversity. *Chem. Sci*. **2013**, *4*, 3845− 3856.

(11) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus *Beauveria bassiana. ChemBioChem* **2007**, *8*, 289−297.

(12) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol*. **2007**, *3*, 213−217.

(13) Liu, N.; Hung, Y.-S.; Gao, S.-S.; Hang, L.; Zou, Y.; Chooi, Y.- H.; Tang, Y. Identification and heterologous production of a benzoyl-primed tricarboxylic acid polyketide intermediate from the zaragozic acid A biosynthetic pathway. *Org. Lett*. **2017**, *19*, 3560−3563.

(14) Sakai, K.; Unten, Y.; Iwatsuki, M.; Matsuo, H.; Fukasawa, W.; Hirose, T.; Chinen, T.; Nonaka, K.; Nakashima, T.; Sunazuka, T.; Usui, T.; Murai, M.; Miyoshi, H.; Asami, Y.; Ōmura, S.; Shiomi, K. Fusaramin, an antimitochondrial compound produced by *Fusarium* sp., discovered using multidrug-sensitive *Saccharomyces cerevisiae. J. Antibiot*. **2019**, *72*, 645−652.

(15) Kim, L. J.; Ohashi, M.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J.; Tang, Y.; Nelson, H. Structural determination of an orphan natural product using microcrystal electron diffraction and genome mining. *ChemRxiv* **2020**.

(16) Liu, N.; Abramyan, E. D.; Cheng, W.; Perlatti, B.; Harvey, C. J. B.; Bills, G. F.; Tang, Y. Targeted genome mining reveals the biosynthetic gene clusters of natural product CYP51 inhibitors. *J. Am. Chem. Soc*. **2021**, *143*, 6043−6047.

(17) Zhu, Y.; Wang, J.; Mou, P.; Yan, Y.; Chen, M.; Tang, Y. Genome mining of cryptic tetronate natural products from a PKS-NRPS encoding gene cluster in *Trichoderma harzianum* t-22. *Org. Biomol. Chem*. **2021**, *19*, 1985−1990.

(18) Chen, Q.; Gao, J.; Jamieson, C.; Liu, J.; Ohashi, M.; Bai, J.; Yan, D.; Liu, B.; Che, Y.; Wang, Y.; Houk, K. N.; Hu, Y. Enzymatic intermolecular hetero-Diels−Alder reaction in the biosynthesis of tropolonic sesquiterpenes. *J. Am. Chem. Soc*. **2019**, *141*, 14052− 14056.

(19) Fotiadou, A. D.; Zografos, A. L. Accessing the structural diversity of pyridone alkaloids: Concise total synthesis of *rac*-citridone A. *Org. Lett*. **2011**, *13*, 4592−4595.

(20) Purdy, T. N.; Kim, M. C.; Cullum, R.; Fenical, W.; Moore, B. S. Discovery and biosynthesis of tetrachlorizine reveals enzymatic benzylic dehydrogenation via an ortho-quinone methide. *J. Am. Chem. Soc*. **2021**, *143*, 3682−3686.

(21) Doyon, T. J.; Perkins, J. C.; Baker Dockrey, S. A.; Romero, E. O.; Skinner, K. C.; Zimmerman, P. M.; Narayan, A. R. H. Chemoenzymatic *o*-quinone methide formation. *J. Am. Chem. Soc*. **2019**, *141*, 20269−20277.

(22) Nakamura, T.; Harachi, M.; Kano, T.; Mukaeda, Y.; Hosokawa, S. Concise synthesis of reduced propionates by stereoselective reductions combined with the Kobayashi reaction. *Org. Lett*. **2013**, *15*, 3170−3173.

(23) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587−1592.

(24) Ortiz de Montellano, P. R.; Nelson, S. D. Rearrangement reactions catalyzed by cytochrome P450s. *Arch. Biochem. Biophys*. **2011**, *507*, 95−110.

(25) Bat-Erdene, U.; Kanayama, D.; Tan, D.; Turner, W. C.; Houk, K. N.; Ohashi, M.; Tang, Y. Iterative catalysis in the biosynthesis of mitochondrial complex II inhibitors harzianopyridone and atpenin B. *J. Am. Chem. Soc*. **2020**, *142*, 8550−8554.

3.6 Experimental Section

This section outlines the experimental procedure and statistics for electron diffraction studies. Al supplementary information, including procedures, figures, tables, and spectra, for isolation and synthesis of metabolites are reported in the adapted article.

3.6.1 Electron Diffraction Data Collection, Processing, and Refinement

Compound **3.13** was recrystallized by slow evaporation of the HPLC fraction in MeCN/H_2O. Microcrystals were deposited onto a pure carbon EM grid and transferred to the FEI Tecnai F200C transmission electron microscope (operating voltage of 200 keV, wavelength of 0.025 Å) on a Gatan 626 cryo-holder at ambient temperature (273 K). The sample was cooled to cryogenic temperature (100 K) after insertion. Diffraction data acquisition, conversion and reduction were performed as described previously.[1] Data from two crystals were scaled and merged together to produce the final data set for **3.13**, before converting the intensities to SHELX format using XDSCONV.[2] Structure of **3.13** was solved *ab initio* using direct methods in SHELXD[3] and refined with SHELXL[4] in ShelXle[5], refining all non-hydrogen atoms anisotropically and placing hydrogen atoms using the riding model. Residual density belonging to solvent molecules was observed, but solvent molecules could not be placed due to disorder.

Crystallographic information file (CIF) for compound **3.13** has been deposited at the Cambridge Crystallographic Data Center (Deposition Number: 2097629).

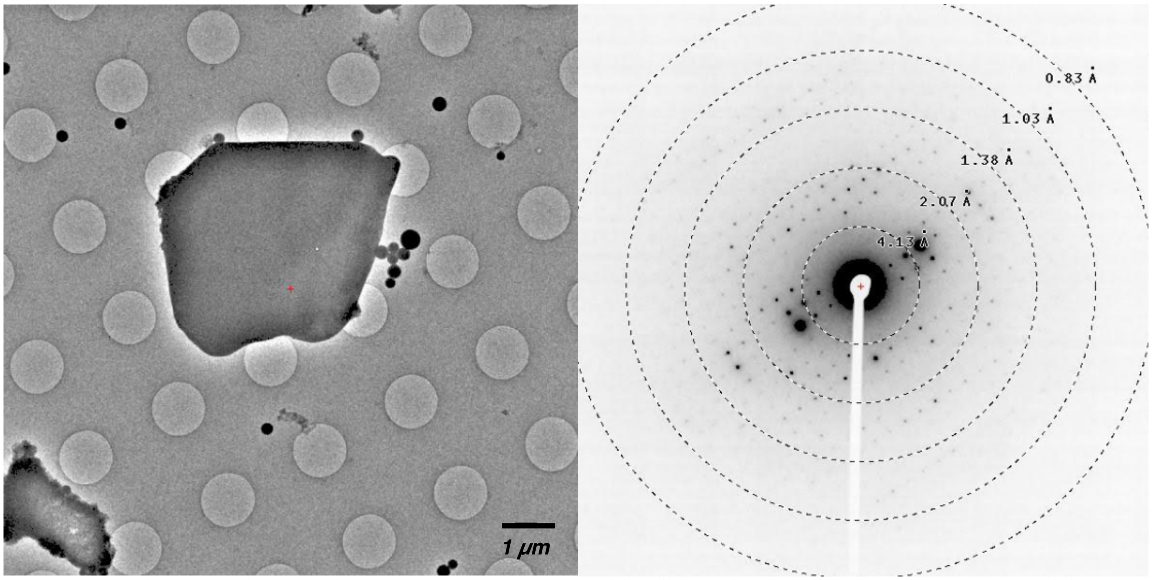


Figure 3.**4** Electron micrograph of **3.13** microcrystal and its diffraction pattern at SA 2600x magnification. Grid holes are 1 μm in diameter.

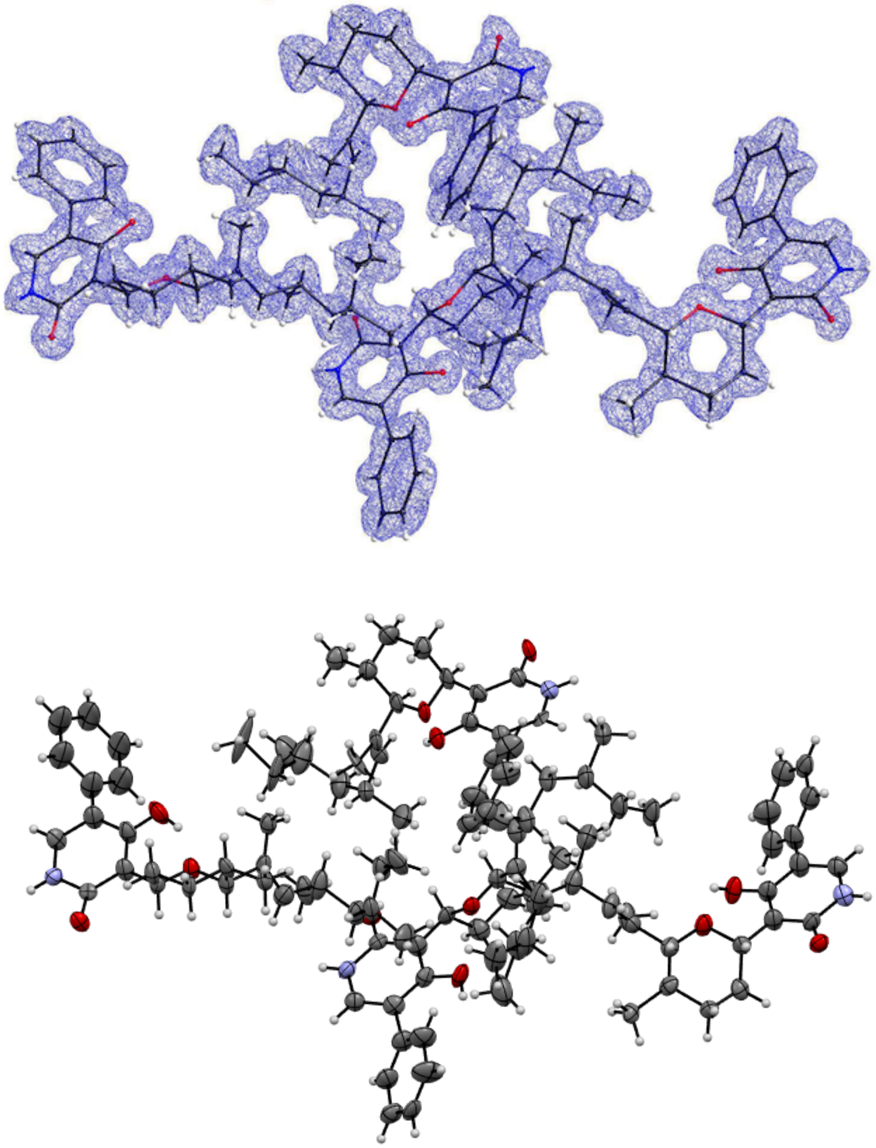


Figure 3.5 Asymmetric unit of **3.13** crystal structure (CCDC 2097629). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.30 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{27}H_{37}NO_3$
Formula weight	423.60
Temperature (K)	100(2)
Space group	C 2
Cell dimensions	

a, b, c (Å)	37.50(4), 10.91(5), 23.36(4)
α, β, γ (°)	90, 97.5(1), 90
Resolution (Å)	1.00(1.05–1.00)
Observed reflections	19453(1374)
Unique reflections	4788(441)
Completeness (%)	90.3(63.8)
R_{obs} (%)	17.1(63.0)
R_{meas} (%)	19.7(76.0)
I/σI	5.07(1.81)
$CC_{1/2}$ (%)	98.8(76.8)
R_1 (%)	16.17
wR_2 (%)	39.19
GooF	1.519

Table 3.1 Data collection and refinement statistics for **3.13**. *Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4999	Data was truncated to a resolution of 1.0 angstrom due to poor diffraction data quality above this threshold. Of the two diffraction data sets merged, one diffracted with reasonable R-meas values to 1.0 angstrom, and the other diffracted up to 1.1 angstrom.
PLAT312_ALERT_2_A Strange C-O-H Geometry (C-O < 1.25 Ang) ..... O1# Check	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT312_ALERT_2_A Strange C-O-H Geometry (C-O < 1.25 Ang) ..... O10# Check	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

PLAT602_ALERT_2_A Solvent Accessible VOID(S) in Structure .... ! Check	Diffraction data from two crystals were merged to obtain a desirable completeness, and we observed high solvent disorder occurring in the structural void.

Table 3.2 Justification for alerts in **3.13** crystal structure (CCDC 2097629).

3.6.2 Supplementary Notes and References

(1) Kim, L. J.; Xue, M.; Li, X.; Xu, Z.; Paulson, E.; Mercado, B.; Nelson, H. M.; Herzon, S. B. Structure revision of the lomaiviticins. *J. Am. Chem. Soc*. **2021**, *143*, 6578–6585.

(2) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010,** *D66*, 133–144.

(3) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(4) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(5) Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. ShelXle: a Qt graphical user interface for SHELXL. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.

CHAPTER FOUR

Prospecting for Natural Product Structural Complexity Using Genome Mining and Microcrystal Electron Diffraction

Adapted from: Lee Joon Kim[†], Masao Ohashi[†], Zhuan Zhang[†], Dan Tan, Matthew Asay, Duilio Cascio, José A. Rodriguez, Yi Tang, Hosea M. Nelson *Nat. Chem. Biol.* **2021**, *17*, 872–877.

4.1 Abstract

Natural products (NPs) continue to fuel the development of pharmaceutical compounds due to their high structural diversity and potent bioactivity. However, the rate of NP discovery has been hampered by limitations in traditional analytical methods for structural analysis. The work described in this chapter combines microcrystal electron diffraction (MicroED) with genome mining to overcome these limitations and accelerate the rate of NP discovery and characterization. This study demonstrates that MicroED can be used to unambiguously elucidate structures of novel NPs, such as the α-pyridone containing NP Py-469, as well as trace impurities that may go undetected by other techniques. Notably, MicroED is used to revise the structure of fischerin, a cytotoxic NP isolated more than 25 years ago. Taken together, the combination of MicroED and genome mining showcases a powerful approach to accelerate the rate of novel NP discovery and, ultimately, drug discovery.

4.2 Introduction

For decades, natural products (NPs) have served as a key inspiration for medicinal chemists to develop new bioactive drugs.[1] As pathogenic organisms continue to evolve and become resistant to existing treatment methods, it has become increasingly important to combat their resistance by discovering therapeutic molecules with novel scaffolds or mechanisms of action.[2] Towards this effort, methods in synthetic biology have enabled the production of novel, highly complex metabolites and rediscovery of previously isolated NPs that are no longer accessible.[3] However, despite these advancements, structural determination of NPs remains a major bottleneck in NP discovery.

The challenges associated with structural characterization of NPs can be attributed to limitations in traditional analytical methods such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. While solution-state NMR spectroscopy is an integral part of structural determination of organic molecules, analytes of insufficient quantities or those with poor physical properties that are insoluble or unstable in deuterated NMR solvents cannot be characterized.[4] Additionally, relative stereochemistry can be difficult to assign for molecules with distal stereocenters interrupted by rigid submoiety with rotatable bonds.[5] X-ray crystallography enables unambiguous structure determination and stereochemistry assignment, but its major limitation is the need for large, highly ordered crystals, as described in Chapter 1.[6] Growing crystals that meet the stringent requirements for X-ray crystallography can prove to be an arduous task, and this process is further hindered in the case of NPs that are produced in infinitesimal amounts. In light of these challenges, microcrystal electron diffraction (MicroED) was sought out to be a complementary technique to traditional analytical methods and overcome their limitations. By enabling unambiguous structure determination from microcrystals that may be present in even seemingly amorphous powders,[7] MicroED could characterize NPs whose structures were previously inaccessible, and accelerate the rate of novel NP discovery.

4.3 Discovery and Structural Characterization of Py-469

To evaluate the combination of MicroED and genome mining[8] in NP discovery, α-pyridone containing fungal metabolites, such as ilicicolin H (**4.1**),[9] fischerin (**4.2**),[10] maximiscin (**4.3**),[11] and atpenin A5 (**4.4**)[12] were targeted due to their high structural complexity and bioactivity (Figure 4.1).

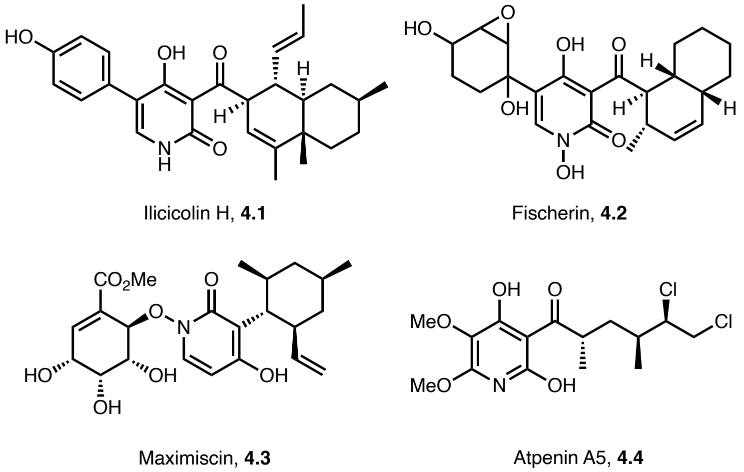

Ilicicolin H, **4.1**
Fischerin, **4.2**
Maximiscin, **4.3**
Atpenin A5, **4.4**

Figure 4.1 α-Pyridone containing fungal metabolites.

Ilicicolin H (**4.1**) is an antifungal natural product originally isolated from *Cylindrocladium ilicicola* that exhibits a potent inhibition of the eukaryotic respiratory chain.[9,13] Comparison of the biosynthetic gene clusters (BGCs) capable of producing Ilicicolin H (**4.1**) revealed that the *icc* BGC from *Penicillium variabile* encodes for three enzymes in addition to the five that produce **4.1**: cytochrome P450 (*iccF*), a short-chain dehydrogenase/reductase (SDR) (*iccH*), and a flavin-dependent oxidoreductase with homology to old-yellow ene reductase (OYE) (*iccG*) (Figure 4.2). The genetic encoding of these three biosynthetically relevant proteins imparted the possibility of discovering an even more complex derivative of **4.1**.

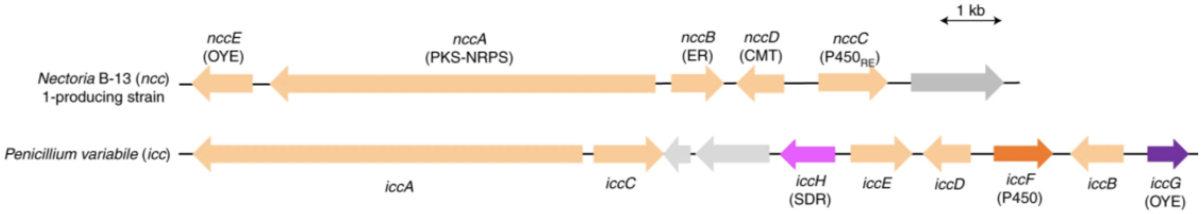


Figure 4.2 Biosynthetic gene clusters that can produce ilicicolin H (**4.1**). The *P. variabile* cluster contains three enzymes in addition to the five genes for producing ilicicolin H (**4.1**).

Heterologous expression of the three genes with the five genes responsible for synthesizing **4.1** led to the identification of a new compound Py-469 (**4.5**) with an increased molecular weight compared to that of ilicicolin H (**4.1**). After purification from the host *via* high-performance liquid chromatography (HPLC), the metabolite **4.5** sample was lyophilized and immediately analyzed by MicroED. The seemingly amorphous white powder sample contained crystalline domains as observed under the electron beam. Two movies collected from a single data collection session were merged to obtain a high-completeness data set, from which a preliminary structure was obtained. The preliminary structure unambiguously revealed the skeletal framework of Py-469 (**4.5**), albeit with randomly assigned elements (Figure 4.3A). Remarkably, with no prior knowledge about the compound other than its molecular formula, the preliminary structure could be refined almost completely just based on bond angles and relative bond lengths. NMR data of Py-469 (**4.5**) corroborated the MicroED data except for the location of one heteroatom, nitrogen, which is a challenging element to distinguish from carbon with the current state of MicroED data quality. The corrected and fully refined MicroED structure elucidated the structure of Py-469 (**4.5**) as a novel epoxydiol-containing derivative of ilicicolin H (**4.1**) (Figure 4.3B). The transformations catalyzed by the three additional *icc* enzymes were confirmed by further experiments: P450 IccF catalyzes oxidative dearomatization and subsequent epoxidation, IccG catalyes ene reduction, and IccH catalyzes the final reduction to yield **4.5** (Figure 4.3C).

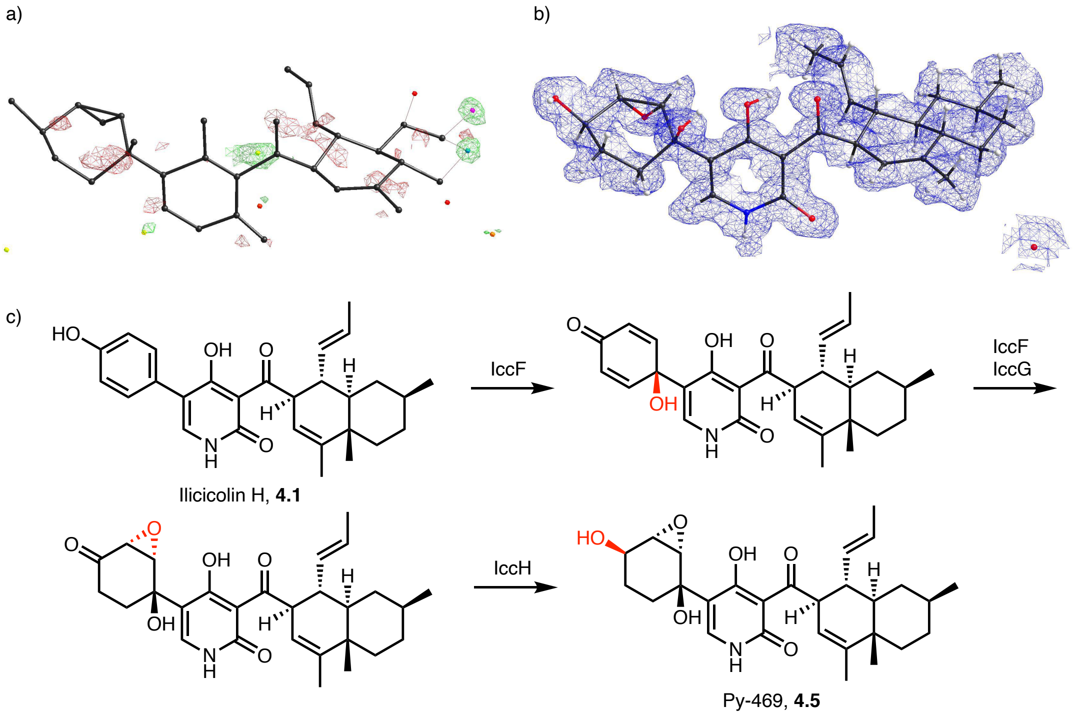
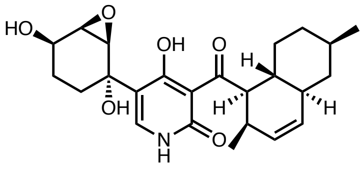


Figure 4.3 Novel derivative of iliciolin H elucidated *via* genome mining and MicroED. (a) Preliminary structure solution of **4.5**. (b) Fully refined MicroED structure of **4.5**. (c) The roles of three additional *icc* enzymes in the biosynthesis of **4.5** from ilicicolin H (**4.1**).

Notably, while the NMR data proved unsuccessful, MicroED structure unambiguously assigned the relative configuration of the epoxydiol with respect to the decalin moiety. This limitation in stereochemical assignment by NMR is attributed to the distal stereocenters that are interrupted by a rigid α-pyridone moiety containing freely rotatable bonds. Similar challenges have been reported for related NPs such as apiosporamide (**4.6**), whose structure was assigned only through total synthesis almost 30 years after its initial isolation (Figure 4.4).[14]



Apiosporamide, **4.6**

Figure 4.4 The revised structure of apiosporamide (**4.6**).

4.4 Rediscovery and Structural Revision of Fischerin

Having successfully demonstrated the ability of MicroED to complement and overcome limitations in NMR spectroscopy, another α-pyridone containing fungal metabolite, fischerin (**4.2**), was targeted. Fischerin (**4.2**) was first isolated from *Neosartorya fischeri* more than 25 years ago, and has been shown to exhibit potent cytotoxicity, causing acute peritonitis in mice.[10] While the two-dimensional structure of fischerin has been determined by the isolation chemists *via* NMR, the relative stereochemistries between the epoxydiol ring and the decalin moiety could not be assigned. In 2015, a structure was proposed based on NMR shift calculations (Figure 4.5),[15] but it has yet to be supported by unambiguous experimental data. Another notable feature of fischerin (**4.2**) is its *cis*-decalin moiety proposed by the isolation chemists, which is rare for this family of natural products. Despite these proposed structures, fischerin (**4.2**) has remained an elusive target not only in structural determination but also in isolation, as no other isolation of this compound has been reported.

Since the original strain reported to produce fischerin (**4.2**) was inaccessible, genome mining approach was leveraged to reconstruct the biosynthesis of fischerin. Upon searching the National Center for Biotechnology Information (NCBI), a BGC responsible for producing a related compound, *N*-hydroxyapiosporamide (**4.7**), was found to be promising (Figure 4.5). Because *N*-hydroxyapiosporamide contains a *trans*-decalin moiety and an extra methyl group at C12, a polyketide synthase–nonribosomal peptide synthetase (PKS–NRPS) containing an inactive methyl transferase domain was identified and carried forward for fischerin biosynthesis. Expressing the gene cluster produced a compound with a molecular weight consistent with that of fischerin. At 98% purity based on liquid chromatography-mass spectrometry (LC-MS) and NMR, this compound exhibited a strong match with fischerin in its proton and carbon NMR spectra, as well as its optical rotation ($[a]_D^{20}$ –28° (c=0.10, $CHCl_3$)). These results strongly support the successful rediscovery of fischerin biosynthesis.

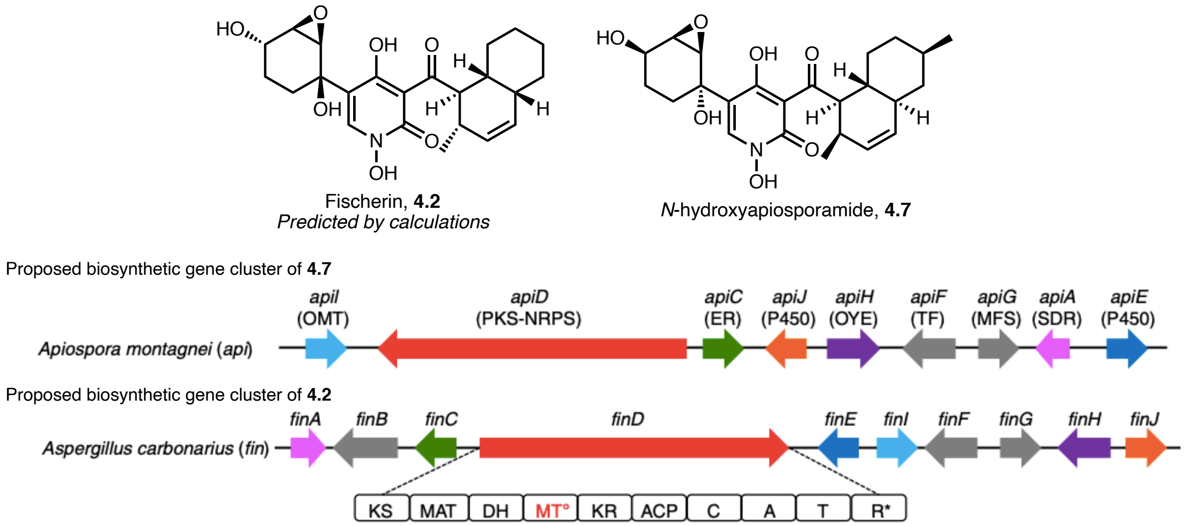


Figure 4.5 Structures (top) and biosynthetic gene clusters (bottom) of fischerin (**4.2**) and *N*-hydroxyapiosporamide (**4.7**).

Similar to the initial isolation study, the relative stereochemistry of fischerin could not be determined using NMR or X-ray crystallography.[10] While crystallization attempts to grow large, single crystals suitable for X-ray diffraction studies were unsuccessful, they could have generated microcrystals that would enable structure determination *via* electron diffraction. Screening various crystallization conditions revealed that pale-yellow particles can be precipitated from a mixture of acetonitrile and water. Under the electron beam, these particles could be characterized as triangular plates (~0.5–3µm longest dimension) that were so thin that they almost blended in the background; however, they were often found stacked on top of each other, leading to a difficulty in collecting clean diffraction pattern from an isolated crystal (Figure 4.6). This challenge was further exacerbated by subpar statistics: diffraction from these triangular plates were limited to a resolution of 1.3–1.5 Å, which is too poor for structure determination using direct methods, and the overall completeness of the data set could not be increased beyond 60% due to the preferred orientation of these flat microcrystals. Preferred orientation of crystals on the grid can hinder structure determination in MicroED experiments due to the physical limits of the stage tilt (–65° to +65°).[16]

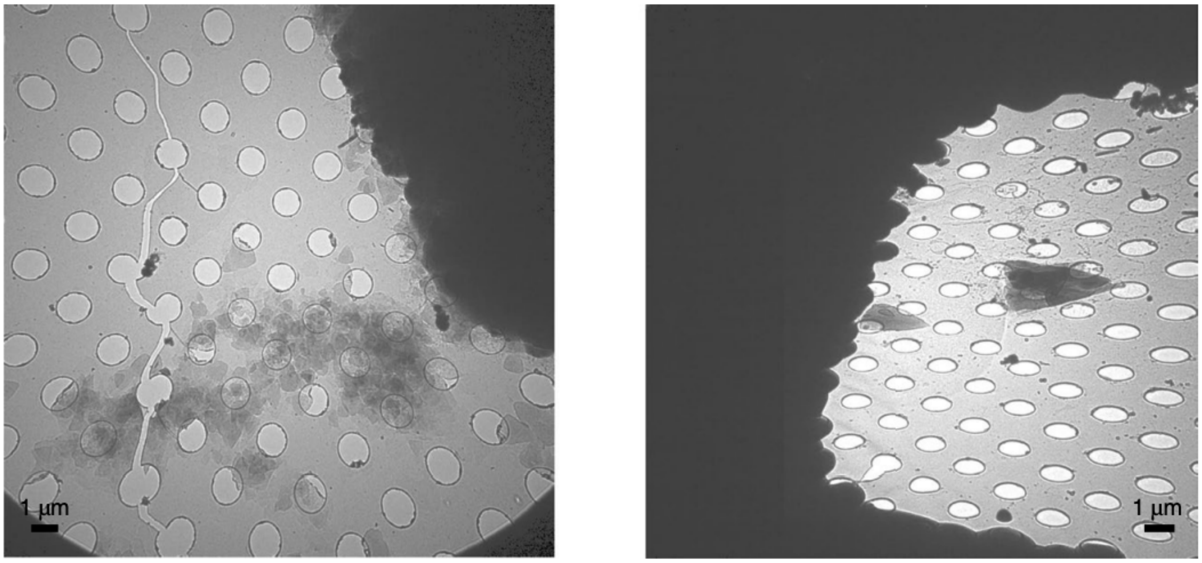


Figure 4.6 Electron micrographs of fischerin (**4.2**) microcrystalline aggregates at 0° tilt (left) and at maximum tilt (approximately 60°, right). Holes are 1 µm wide in diameter.

To overcome these challenges, hundreds of crystallization conditions were screened, followed by optimization of experimental conditions, including: crushing the crystals into smaller crystalline domains using vigorous pipetting or sonication; utilizing different grids such as holey carbon grids so that crystals would partially fall through the bigger holes; and dual loading, where amorphous crystals were pre-loaded onto a grid before depositing a droplet containing the fischerin (**4.2**) crystals (Figure 4.7). Ultimately, various vitrification conditions[17] were screened in an attempt to trap microcrystals in vitreous ice before they fully land on the grid surface, thereby giving access to various orientations.[18] These efforts, in addition to an extensive data collection that resulted in hundreds of movies, led to an increase in both completeness and resolution.

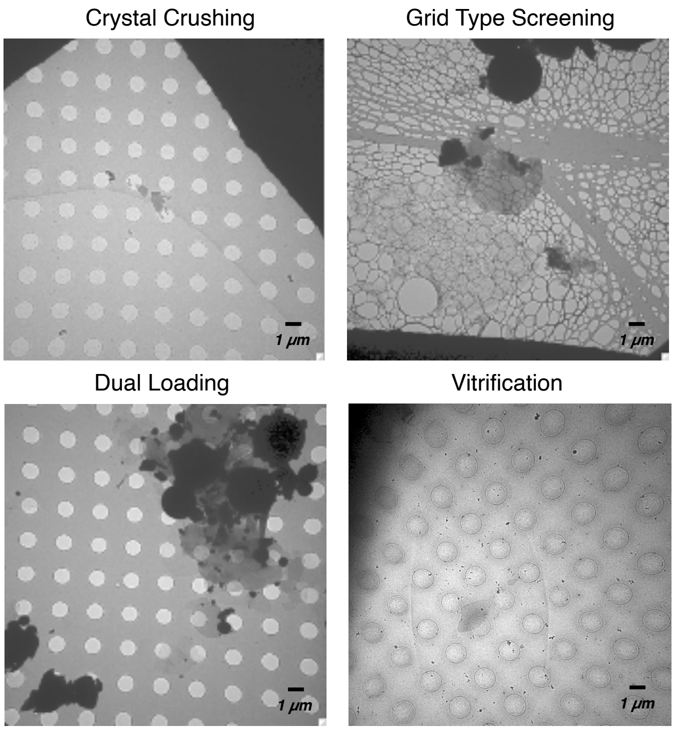

Crystal Crushing
Grid Type Screening
Dual Loading
Vitrification

Figure 4.7 Experimental conditions screened to overcome preferred orientation of microcrystals.

In the end, data from four crystals were merged together to form a high completeness data set that produced an *ab initio* solution at 1.05 Å resolution, which was refined anisotropically to an R_1 value of 13.8%. The asymmetric unit of the crystal structure contains six fischerin molecules with different degrees of rotation about the freely rotatable C–C bond between the epoxydiol unit and the pyridone moiety (C1'–C5, Figure 4.8). This free rotation about the C–C bond prevents any additional symmetry elements and adds disorder, hindering crystallization and structure determination of fischerin *via* crystallographic or spectroscopic methods.

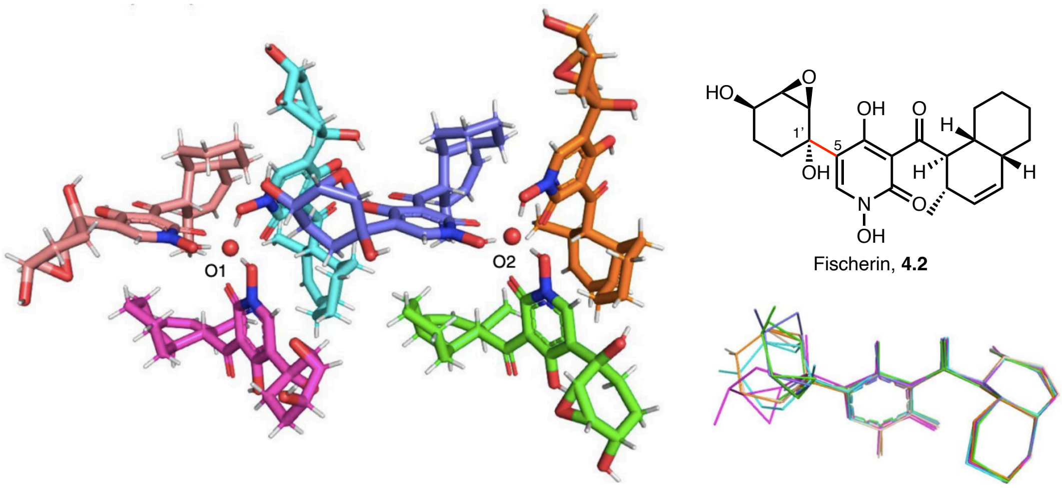


Figure 4.8 Crystal structure of fischerin (**4.2**). Asymmetric unit consists of two trimers (left). Proposed relative stereochemistry (top right) and overlay of six fischerin (**4.2**) molecules showing various degrees of epoxydiol ring rotation (hydrogens omitted for clarity, bottom right).

Interestingly, the crystal structure of fischerin consists of two trimers coordinating to a common central atom (O1 and O2, Figure 4.8). Unlike X-ray crystallography, MicroED cannot unambiguously distinguish one element from another at the current state of MicroED data quality. Refinement procedure is heavily dependent on the molecular formula, although number of covalent bonds and relative bond angles or lengths can provide some guidance, as seen in the case of Py-469. In the case of these central atoms, the data was ambiguous and the identities of these atoms could not be determined reliably. To probe this mystery, an atom substitution test was devised, based on a method described by Kato and coworkers.[19] From the initial SHELXD output, a single refinement cycle was run with X-ray or electron scattering factors of various possible elements. The resulting residual values, R_1 and wR_2, converged to a minimum when oxygen was substituted as the central atoms, suggesting that the fischerin trimers are arranged around disordered water molecules (Figure 4.9 and Table 4.7).

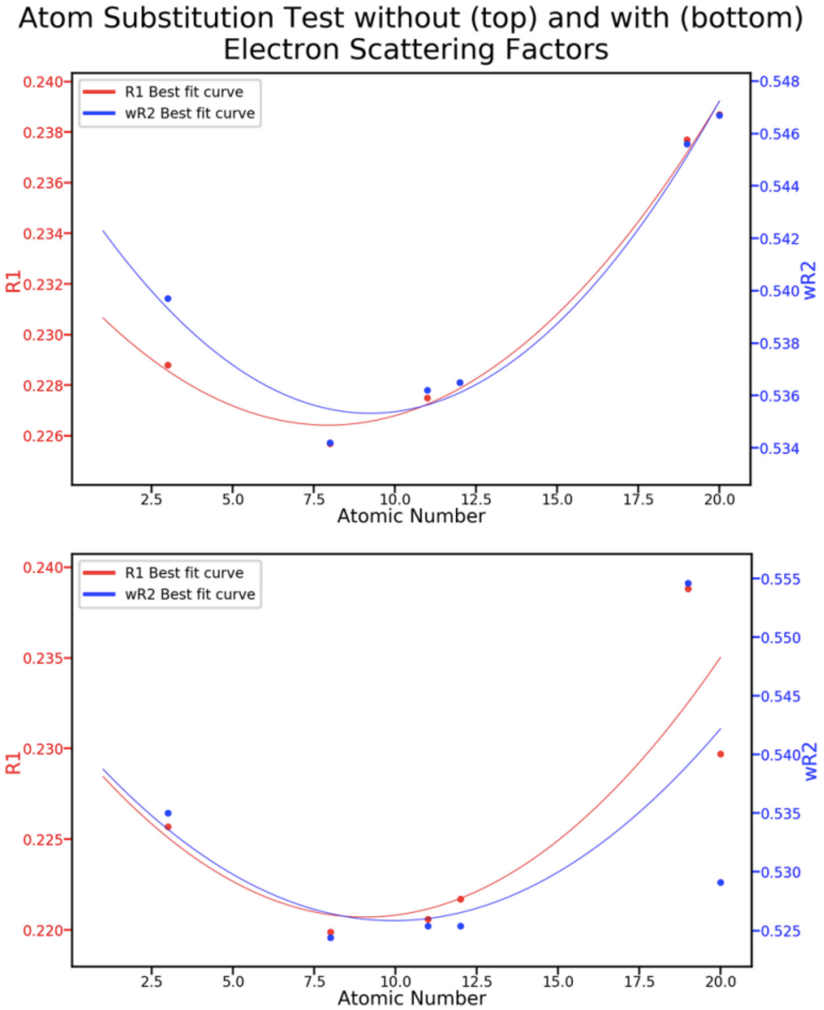
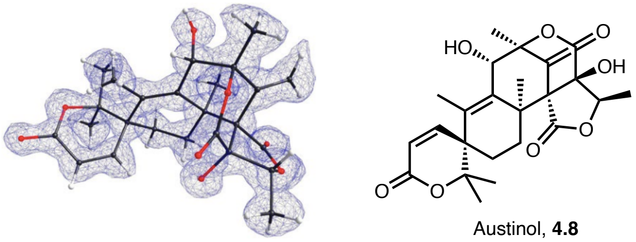


Figure 4.9 Atom substitution test for fischerin (**4.2**) with (top) and without (bottom) electron scattering factors.

From the fully refined crystal structure, the relative stereochemistry of the epoxydiol ring was unambiguously assigned, and the presence of *cis*-decalin moiety, as well as its relative stereochemistry to the epoxydiol unit, was confirmed (Figure 4.8). Importantly, however, the relative stereochemistry assigned using the *ab initio* MicroED structure differs from the predictions made using NMR shift calculations.[15] This disparity demonstrates the limitations in using NMR shift calculations to predict stereochemistries of complex and dynamic functional groups, as well as the importance of experimentally validating computed structures.

4.5 Detection of Trace Impurity Austinol

During the initial screening of crystallization conditions for fischerin (**4.2**), precipitation of acetonitrile:water solution under N_2 gas led to formation of needles that were visible under an optical microscope. Electron diffraction data from two microcrystals were combined to form a high completeness data set, with resolution extending beyond 1.0 Å, and the resulting preliminary structure solution revealed that the identity of these crystals were not of fischerin (**4.2**), but of austinol (**4.8**, Figure 4.10).



Austinol, **4.8**

Figure 4.10 Structural determination of minor impurity austinol (**4.8**).

Austinol (**4.8**) is a polycyclic meroterpenoid NP produced by *A. nidulans*,[20] but a trace amount had co-eluted with fischerin (**4.2**) during purification. Despite the fact that the presence of this trace metabolite was readily detected by electron diffraction, ^1H NMR experiments failed to elucidate this impurity due to its low concentration. To determine MicroED's limit of detection, the solution containing austinol (**4.8**) was serially diluted, deposited onto a grid, and let dry before screening under the electron beam. Even at a concentration of 1.5 ng μL^{-1}, microcrystals with diffraction pattern consistent with that of austinol (**4.8**) could be observed (Figure 4.11), albeit with longer screen time due to sparsity of crystals found on the grid. While microcrystals from lower concentrations could not be detected, this is not necessarily due to the technique's limit of detection, as crystals could have been found after an even lengthier screening. Nonetheless, this serial dilution study demonstrated an impressive sensitivity of MicroED, showcasing its applicability in not only detecting trace impurities, but also obtaining structures of metabolites that may be produced in infinitesimal amounts and undetected by other analytical techniques.

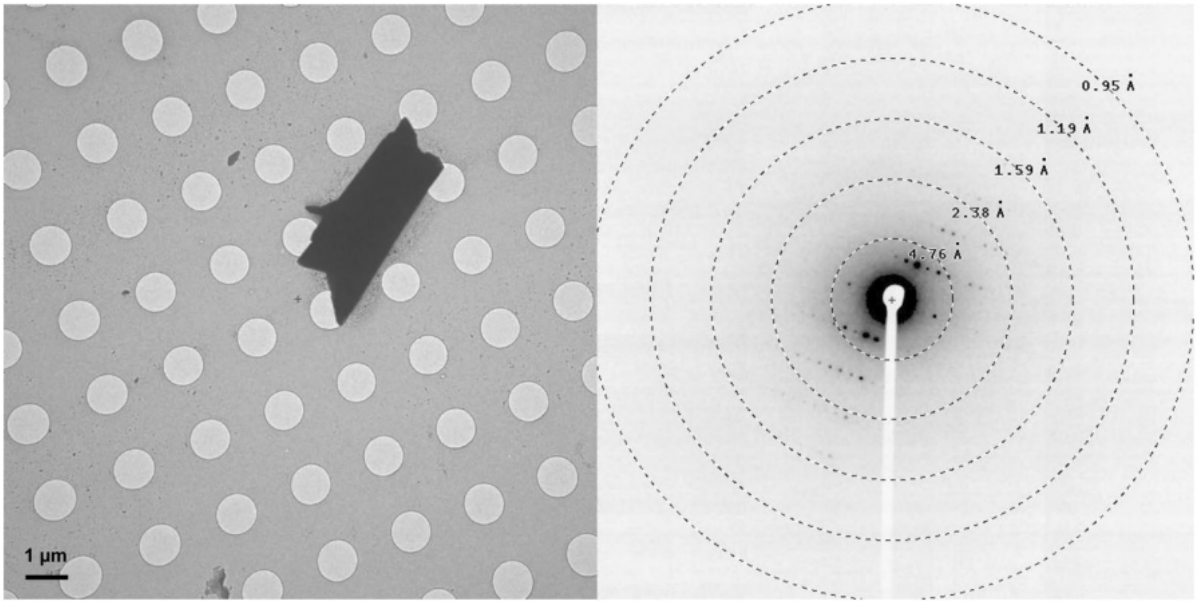


Figure 4.11 Electron micrograph of austinol (**4.8**) crystal and its diffraction pattern from 3 ng of sample. Holes are 1 µm wide in diameter.

4.6 Conclusion

In conclusion, the results described in this chapter demonstrate how the synergy between MicroED and genome mining can be leveraged to elucidate structures of various NPs. The combination of these two techniques enabled the discovery and full characterization of a novel NP Py-469, as well as the rediscovery of fischerin, whose complete structure had remained elusive for decades. In both of these examples, MicroED played an integral role in unambiguously assigning both the connectivity and relative stereochemistry. Additionally, the structure of a co-metabolite austinol was obtained from a sample previously deemed pure by initial ^{1}H NMR experiments. This sensitivity was demonstrated by a serial dilution study, where diffraction pattern was obtained from merely 3 ng of material directly recrystallized on-grid. Taken together, these results highlight the applicability of MicroED in NP discovery and characterization, and its ability to complement existing analytical techniques to overcome their inherent limitations. Further developments, such as enabling automation or increasing the throughput of experiments, will establish MicroED as a powerful method in quickly and unambiguously determining structures of novel and complex NPs that may have been inaccessible by other techniques.

4.7 Notes and References

(1) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803.

(2) Perfect, J. R. The antifungal pipeline: A reality check. *Nat. Rev. Drug Discov.* **2017**, *16*, 603–616.

(3) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. Rational domain swaps decipher programming in fungal highly reducing polyketide synthases and resurrect an extinct metabolite. *J. Am. Chem. Soc.* **2011**, *133*, 16635–16641.

(4) (a) Nicolaou, K. C.; Synder, S. A. Chasing molecules that were never there: Misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem. Int. Ed.* **2005**, *44*, 1012–1044. (b) Maier, M. E. Structural revisions of natural products by total synthesis. *Nat. Prod. Rep.* **2009**, *26*, 1105–1124.

(5) Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. Determination of relative configuration in organic compounds by NMR spectroscopy and computational methods. *Chem. Rev.* **2007**, *107*, 3744–3779.

(6) See Chapter 1.2 for a detailed discussion.

(7) (a) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592. (b) Gruene, T., Wennmacher, J. T. C.; Zaubitzer, C.; Holstein, J. J.; Heidler, J.; Fecteau-Lefebvre, A.; De Carlo, S.; Müller, E.; Goldie, K. N.; Regeni, I.; Li, T.; Santiso-Quinones, G.; Steinfeld, G.; Handschin, S.; van Genderen, E.; van Bokhoven, J. A.; Clever, G. H.; Pantelic, R. Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed.* **2018**, *57*, 16313–16317.

(8) Ziemert, N.; Alanjary, M.; Weber, T. The evolution of genome mining in microbes – a review. *Nat. Prod. Rep.* **2016**, *33*, 988–1005.

(9) Hayakawa, S.; Minato, H.; Katagiri, K. The ilicicolins, antibiotics from *Cylindrocladium ilicicola*. *J. Antibiot.* **1971**, *24*, 653–654.

(10) Fujimoto, H.; Ikeda, M.; Yamamoto, K.; Yamazaki, M. Structure of fischerin, a new toxic metabolite from an ascomycete, *Neosartorya fischeri* var. *fischeri*. *J. Nat. Prod.* **1993**, *56*, 1268–1275.

(11) Du, L.; Robles, A. J.; King, J. B.; Powell, D. R.; Miller, A. N.; Mooberry, S. L.; Cichewicz, R. H. Crowdsourcing natural products discovery to access uncharted dimensions of fungal metabolite diversity. *Angew. Chem. Int. Ed.* **2014**, *53*, 804–809.

(12) Miyadera, H.; Shiomi, K.; Ui, H.; Yamaguchi, Y.; Masuma, R.; Tomoda, H.; Miyoshi, H.; Osanai, A.; Kita, K.; Ōmura, S. Atpenins, potent and specific inhibitors of mitochondrial complex II (succinate-ubiquinone oxidoreductase). *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 473–477.

(13) Singh, S. B.; Liu, W.; Li, X.; Chen, T.; Shafiee, A.; Card, D.; Abruzzo, G.; Flattery, A.; Gill, C.; Thompson, J. R.; Rosenbach, M.; Dreikorn, S.; Hornak, V.; Meinz, M.; Kurtz, M.; Kelly, R.;

Onishi, J. C. Antifungal spectrum, in vivo efficacy, and structure-activity relationship of ilicicolin H. *ACS Med. Chem. Lett.* **2012**, *3*, 814–817.

(14) Williams, D. R.; Kammler, D. C.; Donnell, A. F.; Goundry, W. R. F. Total synthesis of (+)-apiosporamide: Assignment of relative and absolute configuration. *Angew. Chem. Int. Ed.* **2005**, *44*, 6715–6718.

(15) Amini, S. K. Assignment of the absolute configuration of fischerin by computed nmr chemical shifts. *J. Struct. Chem.* **2015**, *56*, 1334–1341.

(16) Nannenga, B. L. MicroED methodology and development. *Struct. Dyn.* **2020**, *7*, 014304, https://doi.org/10.1063/1.5128226.

(17) (a) de la Cruz, M. J.; Hattne, J.; Shi, D.; Seidler, P.; Rodriguez, J.; Reyes, F. E.; Sawaya, M. R.; Cascio, D.; Weiss, S. C.; Kim, S. K.; Hinck, C. S.; Hinck, A. P.; Calero, G.; Eisenberg, D.; Gonen, T. Atomic-resolution structures from fragmented protein crystals with the cryoEM method MicroED. *Nat. Methods* **2017**, *14*, 399–402. (b) Dubochet, J.; Adrian, M.; Chang, J. J.; Homo, J. C.; Lepault, J.; McDowall, A. W.; Schultz, P. Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **1988**, *21*, 129–228.

(18) Natesh, R. *Structural Bioinformatics: Applications in Preclinical Drug Discovery Process*, Edition 1, Vol. 27 (ed. Mohan, C. G.) 375–400 (Springer Nature, 2019).

(19) Kato, K.; Tanaka, H.; Sumizawa, T.; Yoshimura, M.; Yamashita, E.; Iwasaki, K.; Tsukihara, T. A vault ribonucleoprotein particle exhibiting 39-fold dihedral symmetry. *Acta Cryst.* **2008**, *D64*, 525–531.

(20) Matsuda, Y.; Abe, I. Biosynthesis of fungal meroterpenoids. *Nat. Prod. Rep.* **2016**, *33*, 26–53.

4.8 Experimental Section

*This section outlines the experimental procedure and statistics for electron diffraction studies. All supplementary information, including procedures, figures, tables, and spectra, for genome mining and isolation of **4.2**, **4.5**, and **4.8** are reported in the adapted article.*

4.8.1 Recrystallization

Dried, purified fischerin (**4.2**) sample (1.0 mg, 2.3 μmol) was suspended in 0.1 mL of tetrahydrofuran (THF), and H_2O was added until the solution became clear and homogeneous, changing color to light yellow. Crystals were grown at room temperature by slow evaporation in a dram vial.

0.8 mg (1.7 μmol) of sample containing austinol (**4.8**) was dissolved in 0.8 mL of 1:1 MeCN–H_2O solution, then precipitated at room temperature under a stream of air.

4.8.2 Serial Dilution

0.4 mg (0.9 μmol) of vacuum-dried sample containing austinol (**4.8**) was redissolved in a 2:1 MeCN–H_2O solution at 333 μM concentration. The sample was then diluted 100-fold in H_2O, vortexed, and centrifuged before the subsequent dilution.

4.8.3 Electron Diffraction Data Acquisition and Processing

Recrystallized fischerin (**4.2**) microcrystals were diluted to optimize crystal density and deposited onto a Quantifoil holey-carbon EM grid as a 2 μL drop before blotting and plunge freezing in liquid ethane using Vitrobot Mark IV (FEI). The vitrified grid was then transferred to the TEM on a Gatan 626 cryo-transfer holder, maintaining cryogenic temperature (100 K) throughout data acquisition.

Py-469 (**4.5**) powder was deposited onto a continuous carbon EM grid as previously described.[1] The grid was then transferred to the TEM on a single-tilt sample holder for room temperature data acquisition.

2 µL of recrystallized austinol (**4.8**) sample was aliquotted from the top of the solution to account for crystal density before deposition onto a continuous carbon EM grid. 2 µL of serial dilution sample was obtained from the bottom of the solution and deposited onto a Quantifoil holey-carbon EM grid. In both cases, solvent was wicked away manually, and the dried grid was transferred to the TEM for room temperature data acquisition on a single-tilt sample holder.

All diffraction data were collected using the TEM User Interface on an FEI Tecnai F200C electron microscope with an operating voltage of 200 keV, corresponding to a wavelength of 0.025 Å. During data acquisition, the crystal of interest was isolated using a selected area aperture and continuously rotated at a rate of $-0.3° \text{ s}^{-1}$ over a tilt range of 50–100°. Continuous rotation diffraction data was recorded on TEM Imaging & Analysis (TIA) using rolling shutter mode with a Ceta-D CMOS 4k × 4k camera, integrating at a rate of 3 s per frame and binning by 2 to produce final images of 2k × 2k.[2]

Diffraction movies saved as SER files were converted to SMV format using ser2smv[3] software, as described previously[4]. Frames were indexed and integrated in XDS[5]. Data from two to four crystals were scaled and merged together using XSCALE[6] to produce the final data sets for **4.2**, **4.5** and **4.8** (Tables 4.1, 4.3, and 4.5). Finally, intensities were converted to SHELX format using XDSCONV[4].

4.8.4 Structure Determination and Refinement

Structures of **4.2** and **4.5** were solved ab initio using direct methods in SHELXD[7]. The structure of **4.8** was also solved ab initio using SHELXT[8]. All structures were refined with SHELXL[9] in ShelXle[10]. Non-hydrogen atoms were refined anisotropically, and hydrogen atoms were placed using the riding model. Residual density corresponding to solvent molecules was observed during refinement; however, solvent disorder hindered unambiguous identification, and oxygen atoms were placed in lieu of water molecules. Overall data quality and refinement statistics are reported in Tables 4.1, 4.3, and 4.5. Structure overlays were illustrated using Pymol[11]. Data from atom substitution tests were visualized using Python[12].

4.8.5 Crystal structures of Fungal Metabolites

4.8.5.1 Py-469 (4.5)

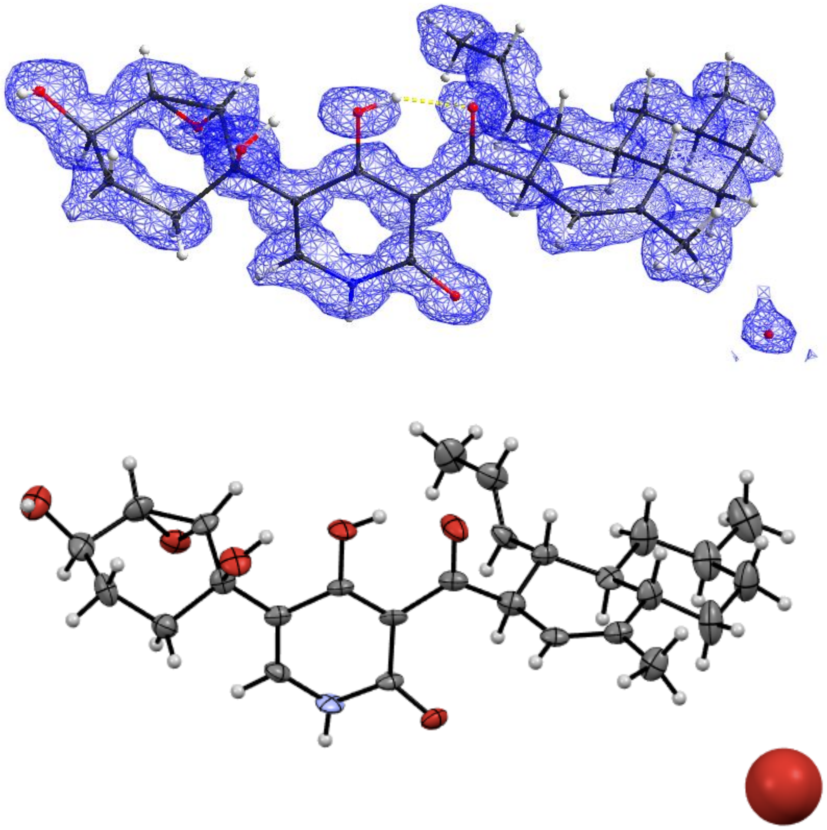


Figure 4.12 Asymmetric unit of Py-469 (**4.5**) crystal structure (CCDC 2038723). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.33 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{27}H_{35}NO_6$
Data Collection	
Temperature (K)	293(2)
Space group	C 2 2 2_1
Cell dimensions	
a, b, c (Å)	5.26(1), 26.64(4), 34.53(9)
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	0.90 (0.95–0.90)
R_{obs} (%)	15.84 (40.1)

R_{meas} (%)	16.6 (43.0)
I/σI	7.83 (3.74)
$CC_{1/2}$ (%)	99.1 (86.0)
Completeness (%)	83.6 (83.3)
Data Refinement	
Resolution (Å)	0.90
Observed reflections (#)	12884 (1796)
Unique reflections (#)	1707 (230)
R_1 (%)	13.68
wR_2 (%)	35.05
GooF	1.778

Table 4.1 Data collection and refinement statistics for **4.5**. Highest resolution shell is shown in parenthesis.

A-level Alerts	**Justifications**
PLAT029_ALERT_3_A _diffrn_measured_fraction_theta_full value Low	Full sampling of reciprocal space was hindered by the limited stage tilt of the electron microscope.
PLAT417_ALERT_2_A Short Inter D-H..H-D H1 ..H1 . 1.62 Ang. 1-x,y,3/2-z = 3_656 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.

B-level Alerts	**Justifications**
THETM01_ALERT_3_B The value of sine(theta_max)/wavelength is less than 0.575 Calculated sin(theta_max)/wavelength = 0.5556	Data was truncated to a resolution of 0.9 angstrom, due to higher resolution data falling outside of the edge of the detector.

PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing ?) .... O1A Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and place. O1A is an oxygen we put in lieu of a solvent molecule.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.02324 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 4.2 Justification for alerts in Py-469 (**4.5**) crystal structure (CCDC 2038723).

4.8.5.2 Fischerin (4.2)

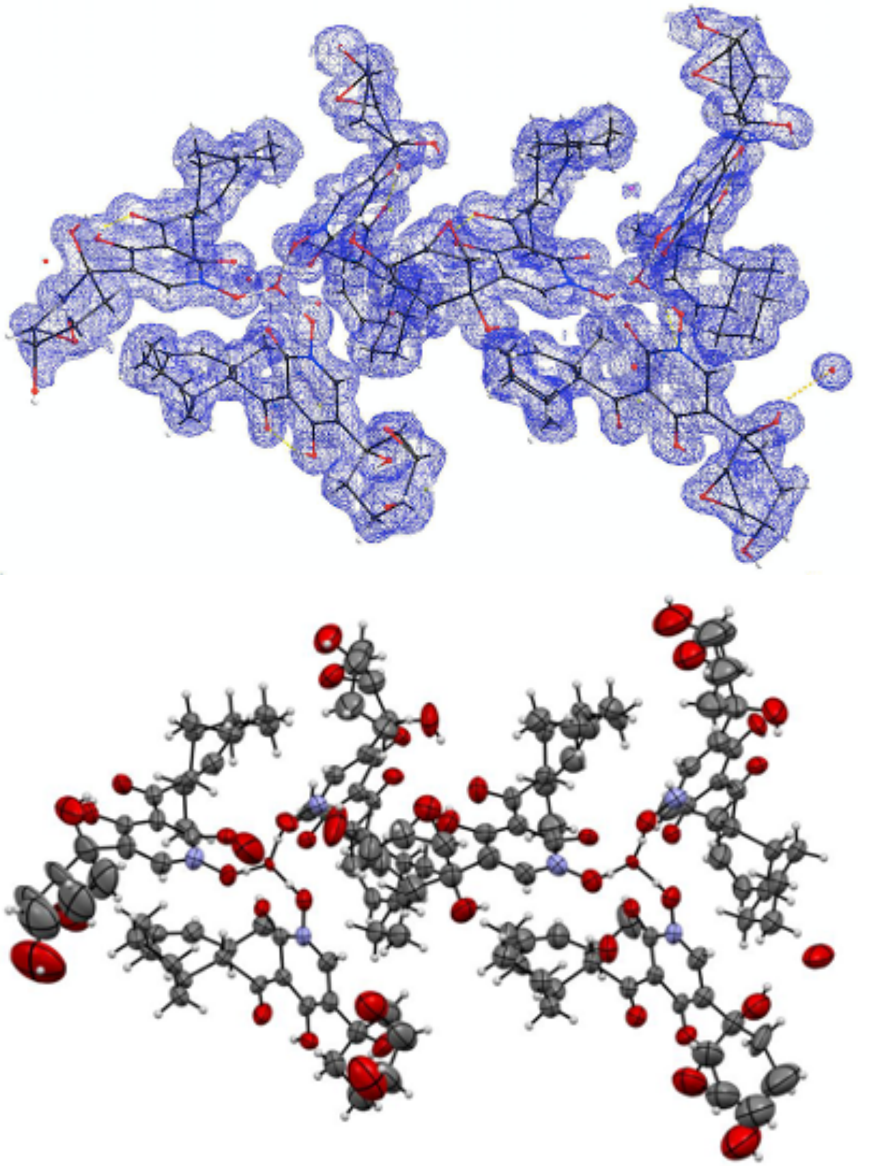


Figure 4.13 Asymmetric unit of fischerin (**4.2**) crystal structure (CCDC 2020516). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.39 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{23}H_{29}NO_7$
Data Collection	
Temperature (K)	100(2)
Space group	C 2
Cell dimensions	

a, b, c (Å)	36.59(2), 20.800(9), 18.9(2)
α, β, γ (°)	90.00, 96.3(2), 90.00
Resolution (Å)	1.05 (1.10–1.05)
R_{obs} (%)	19.0 (94.6)
R_{meas} (%)	20.1 (111.2)
I/σI	5.87 (1.19)
$CC_{1/2}$ (%)	98.9 (64.9)
Completeness (%)	86.5 (72.5)
Data Refinement	
Resolution (Å)	1.05
Observed reflections (#)	52478 (2220)
Unique reflections (#)	5789 (621)
R_1 (%)	13.78
wR_2 (%)	32.23
GooF	1.432

Table 4.3 Data collection and refinement statistics for **4.2**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4763	Data was truncated to a resolution of 1.05 angstrom due to poor diffraction data quality above this threshold. Of the four diffraction data sets merged, only one diffracted with reasonable R-meas values to 1.05 angstrom, with the other three diffracting up to ranges 1.1-1.15.

PLAT029_ALERT_3_A _diffrn_measured_fraction_theta_full value Low . 0.865 Why?	Full sampling of reciprocal space was hindered by the limited stage tilt of the electron microscope. The crystals also suffered from preferred orientation due to flat, plate-like morphology, as well as low symmetry (C2).
PLAT089_ALERT_3_A Poor Data / Parameter Ratio (Zmax < 18) .... 3.95 Note	The structure has low symmetry and a very large number of non-hydrogen atoms to refine (191 atoms). The completeness of the data set is low due to limitations in stage rotation in MicroED and preferred orientation of the crystals due to their morphology.
PLAT375_ALERT_2_A Strange C-O-H Geometry (C-O > 1.45 Ang) ..... O19# Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT410_ALERT_2_A Short Intra H...H Contact H17_2 ..H28A_2 . 1.79 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.
PLAT414_ALERT_2_A Short Intra D-H..H-X H17_4 ..H29_4 1.76 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.
PLAT414_ALERT_2_A Short Intra D-H..H-X H20_6 ..H28B_6 1.53 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.
PLAT415_ALERT_2_A Short Inter D-H..H-X H31_3 ..H27B_6 . 1.69 Ang. 2-x,y,1-z = 2_756 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.
PLAT601_ALERT_2_A Unit Cell Contains Solvent Accessible VOIDS of . 214 Ang**3	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and place.

B-level Alerts	**Justifications**

PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) .... 0.36 Report	Electron diffraction data has higher R-factors compared to X-ray data. Four diffraction data sets were merged.
PLAT149_ALERT_3_B s.u. on the beta Angle is Too Large .... 0.20 Degree	Diffraction data from four crystals were merged to obtain a desirable completeness, and high levels of disorder was expected due to multiple rotatable bonds within a big unit cell consisting of multiple molecules.
PLAT220_ALERT_2_B NonSolvent Resd 1 O Ueq(max)/Ueq(min) Range 7.0 Ratio	Diffraction data from four crystals were merged to obtain a desirable completeness, and high levels of disorder was expected due to multiple rotatable bonds within a big unit cell consisting of multiple molecules.
PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors of C27_6 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT260_ALERT_2_B Large Average Ueq of Residue Including O05 0.559 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and place. O05 is an oxygen we put in lieu of a solvent molecule.
PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing ?) .... O05 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and place. O05 is an oxygen we put in lieu of a solvent molecule.
PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing ?) .... O04 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and place. O05 is an oxygen we put in lieu of a solvent molecule.
PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing ?) .... O03 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder, whose atoms we could not unambiguously identify and

	place. O05 is an oxygen we put in lieu of a solvent molecule.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.04507 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT363_ALERT_2_B Long C(sp3)-C(sp2) Bond C18_2 - C23_2 . 1.75 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT369_ALERT_2_B Long C(sp2)-C(sp2) Bond C12_2 - C14_2 . 1.60 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT417_ALERT_2_B Short Inter D-H..H-D H31_3 ..H31_4 . 1.90 Ang. 5/2-x,-1/2+y,1-z = 4_746 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.
PLAT420_ALERT_2_B D-H Without Acceptor O29_2 --H29_2 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder. Presence of nearby solvent molecules was not observed.
PLAT420_ALERT_2_B D-H Without Acceptor O31_3 --H31_3 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder. Presence of nearby solvent molecules was not observed.
PLAT420_ALERT_2_B D-H Without Acceptor O31_5 --H31_5 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder. Presence of nearby solvent molecules was not observed.
PLAT420_ALERT_2_B D-H Without Acceptor O20_6 --H20_6 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we observed a high level of solvent disorder. Presence of nearby solvent molecules was not observed.
PLAT420_ALERT_2_B D-H Without Acceptor O31_6 --H31_6 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and we

	observed a high level of solvent disorder. Presence of nearby solvent molecules was not observed.

Table 4.4 Justification for alerts in fischerin (**4.2**) crystal structure (CCDC 2020516).

4.8.5.3 Austinol (4.8)

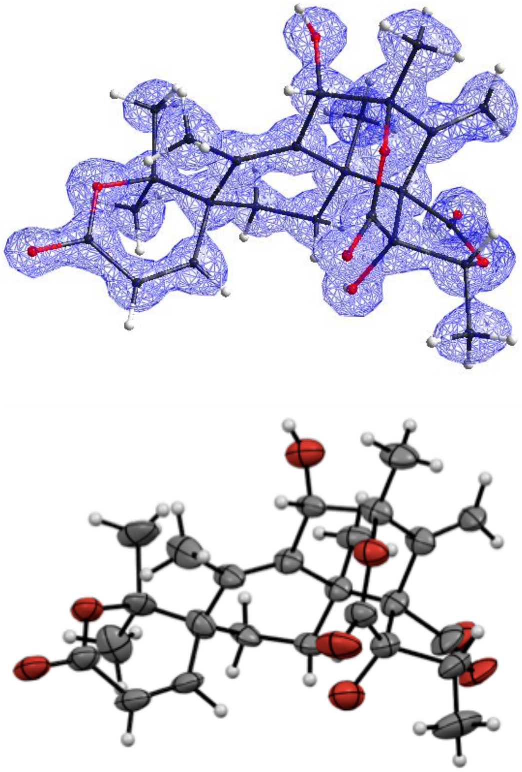


Figure 4.14 Asymmetric unit of austinol (**4.8**) crystal structure (CCDC 2020510). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.51 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{25}H_{30}O_8$
Data Collection	
Temperature (K)	293(2)
Space group	P 2_1 2_1 2_1
Cell dimensions	

a, b, c (Å)	7.910(6), 11.110(8), 23.89(1)
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	1.00 (1.05–1.00)
R_{obs} (%)	16.0 (71.8)
R_{meas} (%)	18.3 (80.4)
I/σI	5.20 (1.74)
$CC_{1/2}$ (%)	98.8 (67.4)
Completeness (%)	99.3 (100.0)
Data Refinement	
Resolution (Å)	1.00
Observed reflections (#)	6179 (856)
Unique reflections (#)	1315 (173)
R_1 (%)	14.69
wR_2 (%)	33.56
GooF	1.401

Table 4.5 Data collection and refinement statistics for **4.8**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4999	Data was truncated to a resolution of 1.0 angstrom due to poor diffraction data quality above this threshold. The calculated value of sin(theta_max)/wavelength is 0.4999, which is still acceptable.
PLAT414_ALERT_2_A Short Intra D-H..H-X H11B ..H28 1.40 Ang. x,y,z = 1_555 Check	H28 is a hydrogen atom placed using the riding model on an alcohol, which is free to rotate.

B-level Alerts	Justifications
PLAT089_ALERT_3_B Poor Data / Parameter Ratio (Zmax < 18) .... 4.24 Note	The data was truncated due to poor diffraction above a resolution of 1.0 angstrom. Despite limitations in stage tilt in MicroED, we were still able to obtain a solution with completeness of 99.3%.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.02615 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT420_ALERT_2_B D-H Without Acceptor O28 --H28 . Please Check	Diffraction data from four crystals were merged to obtain a desirable completeness, and solvent disorder is expected. Presence of nearby solvent molecules was not observed.

Table 4.6 Justification for alerts in austinol (**4.8**) crystal structure (CCDC 2020510).

4.8.6 Atom Substitution Test

Without Electron Scattering Factors			With Electron Scattering Factors		
Atomic Number	R_1	wR_2	Atomic Number	R_1	wR_2
3	0.2288	0.5397	3	0.2257	0.5350
8	0.2257	0.5342	8	0.2199	0.5244
11	0.2275	0.5362	11	0.2206	0.5254
12	0.2281	0.5365	12	0.2217	0.5254
19	0.2377	0.5456	19	0.2388	0.5546
20	0.2387	0.5467	20	0.2297	0.5291

Table 4.7 Data coordinates for atom substitution test.

4.8.6.1 Automated Atom Substitution Test Procedure

Prior to running the script, the computer should have a properly installed SHELX suite[7–9]. For substituting a single atom with user-defined list of elements, a single directory must contain .ins,

.hkl, an Excel spreadsheet containing the list of electron scattering factors[13], and atom_substitution.py prior to calling “python3 atom_substitution.py” to run Python3.8[12]. If running a batch atom substitution, where every atom in the .ins file is substituted with user-defined list of elements, the directory must also contain batch_atom_substitution.py to call “python3 batch_atom_substitution.py.”

4.8.6.2 atom_substitution.py

```
"""
Created on Thu Mar 26 16:54:56 2020
This code will take an atom and substitute it with another atom then run one refinement cycle with
ShelxL.
Finally, it will grab the R1 and wR2 outputs and plot them using Matplotlib.

@author: Lee Joon Kim
"""

import os
import subprocess
import shutil
import xlrd
from pathlib import Path
import matplotlib.pyplot as plt

def main(original_atom = "", original_atom_number = "", new_atom = "", scattering = ""):

    # input for atom type you'll be substituting
    periodic_table = ['H', 'HE', 'LI', 'BE', 'B', 'C', 'N', 'O', 'F', 'NE', 'NA', 'MG', 'AL', 'SI', 'P', 'S', 'CL',
'AR',
                  'K', 'CA', 'SC', 'TI', 'V', 'CR', 'MN', 'FE', 'CO', 'NI', 'CU', 'ZN', 'GA', 'GE', 'AS', 'SE',
'BR', 'KR',
                  'RB', 'SR', 'Y', 'ZR', 'NB', 'MO', 'TC', 'RU', 'RH', 'PD', 'AG', 'CD', 'IN', 'SN', 'SB', 'TE',
'I', 'XE',
                  'CS', 'BA', 'LA', 'CE', 'PR', 'ND', 'PM', 'SM', 'EU', 'GD', 'TB', 'DY', 'HO', 'ER', 'TM',
'YB', 'LU',
                  'HF', 'TA', 'W', 'RE', 'OS', 'IR', 'PT', 'AU', 'HG', 'TL', 'PB', 'BI', 'PO', 'AT', 'RN',
                  'FR', 'RA', 'AC', 'TH', 'PA', 'U', 'NP', 'PU']
    original_atom = input("Atom type to change (ex: C or Br): ")
    while original_atom.upper() not in periodic_table:
        print("Check if that's the right element")
```

```
            original_atom = input("Atom type to change (ex: C or Br): ")

    # input for atom number you'll be substituting
    original_atom_number = input("Corresponding atom number as written in .ins or .res file (ex: 001 or 02): ")

    # input for the new atom you'll test
    new_atom = input("Atom(s) to replace, separated by a comma (ex: C or C,BR,CL,H or all): ")
    if new_atom.lower() in ['all']:
        list_new_atom = periodic_table
        print("Substituting " + str(original_atom.upper()) + str(original_atom_number) + " with all 94 elements")
    else:
        list_new_atom = (new_atom.split(','))
        for new_atom_item in range(len(list_new_atom)):
            while list_new_atom[new_atom_item].upper() not in periodic_table:
                print("Check if that's the right element")
                new_atom = input("Atom(s) to replace, separated by a comma (ex: C or C,BR,CL,H): ")
                list_new_atom = (new_atom.split(','))
        print("Substituting " + str(original_atom.upper()) + str(original_atom_number) + " with " + str(new_atom.upper()))

    # input for whether or not you'll use electron scattering factors
    # currently only using ____ version
    scattering = input("Use electron scattering? (Y/N) ")

    # makes separate directories for each atom substitution to make
    for new_atom_item in range(len(list_new_atom)):
        path = os.getcwd()
        files = os.listdir(".")
        new_atom_path = os.path.join(path, str(list_new_atom[new_atom_item].upper()))
        os.makedirs(new_atom_path, exist_ok = True)

    # copies .ins and .hkl files into the subdirectories
    for atom_dir in os.scandir(path):
        full_path = os.path.join(path,atom_dir.name)
        if os.path.isdir(full_path) == True:
            for file in files:
                if file.endswith(".ins") or file.endswith(".hkl"):
                    shutil.copyfile(file, full_path + "/" +  file)
    print("Making subdirectories and copying .ins and .hkl files... ")

    # replaces the desired atom in each subdirectory
    for root, dirs, files in os.walk("."):
        results = open("Results.txt", "w+")
        results.write("Element \t Atomic Number \t R1 \t wR2 \n")
```

```
        for folder in dirs:
            os.chdir(str(path)+"/"+str(folder))
            subdir_files = os.listdir(".")
            for file in subdir_files:
                if file.endswith(".ins"):
                    shutil.move(file, file+".original")
                    destination = open(file, "w")
                    source = open(file+".original", "r")
                    for line in source:
                        if line.startswith('SFAC '):
                            if scattering.lower() in ['n', 'no']:
                                SFAC_list = line.split()
                                SFAC_list.append(str(folder.upper())+ "\n")
                                new_line = " ".join(SFAC_list)
                                destination.write(line.replace(str(line),str(new_line)))
    # if using electron scattering factors, grabs them from the excel spreadsheet
                            elif scattering.lower() in ['y', 'yes']:
                                excel_file_location = path
                                excel_file                                              =
xlrd.open_workbook(os.path.join(str(excel_file_location),'electron scattering table.xlsx'))
                                sheet = excel_file.sheet_by_index(1)

                                SFAC_list = line.split()
                                SFAC_list.append(str(folder.upper()))
                                atom_list = SFAC_list[1:]
                                new_SFAC_list = []
                                for i in range(len(atom_list)):
                                    for atom_row in range(sheet.nrows):
                                        if atom_list[i] == sheet.cell_value(atom_row,0).upper():
                                            scattering_factors_1 = sheet.row_values(atom_row,2,10)
                                            scattering_factors_2 = sheet.row_values(atom_row,10,16)
                                            scattering_line_1 = "".join(str(scattering_factors_1))
                                            scattering_line_1 = scattering_line_1.replace(",", " ")
                                            scattering_line_1 = scattering_line_1.replace("[", " ")
                                            scattering_line_1 = scattering_line_1.replace("]", " ")
                                            scattering_line_2 = "".join(str(scattering_factors_2))
                                            scattering_line_2 = scattering_line_2.replace(",", " ")
                                            scattering_line_2 = scattering_line_2.replace("[", " ")
                                            scattering_line_2 = scattering_line_2.replace("]", " ")
                                            new_SFAC_list.append("SFAC "+atom_list[i]+scattering_line_1+"="
+ "\n"+" "+scattering_line_2+"\n")

                                new_line = "".join(new_SFAC_list)
                                destination.write(line.replace(str(line),str(new_line)))

                        elif line.startswith('UNIT '):
```

```
                        UNIT_list = line.split()
                        UNIT_list.append("1" + "\n")
                        new_line = " ".join(UNIT_list)
                        destination.write(line.replace(str(line),str(new_line)))
                    elif line.startswith(str(original_atom.upper()+original_atom_number+"  ")):
                        original_line = line
                        atom_line = line.split()
                        old_SFAC_number = len(UNIT_list)-2
                      #  old_SFAC_number = atom_line[1]
                        atom_line[1] = str(int(old_SFAC_number)+1)
                        atom_line[0] = str(folder.upper()+'99')
                        atom_line.append("\n")
                        new_atom_line = " ".join(atom_line)
                        destination.write(line.replace(str(original_line),str(new_atom_line)))
                    else:
                        destination.write(line)

                source.close()
                destination.close()

                os.remove(str(file)+".original")

# runs shelxl cycle once in each subdirectory
        shelxl_call = Path(file).stem
        subprocess.run(["shelxl", shelxl_call])

# fills out Results.txt file with outputs
    for folder in dirs:
        os.chdir(str(path)+"/"+str(folder))

        atomic_number = periodic_table.index(str(folder.upper()))
        atomic_number = atomic_number + 1

        new_subdir_files = os.listdir(".")
        for file in new_subdir_files:
            if file.endswith(".res"):
                res_file = open(file, "r")
                data_list = [str(folder), str(atomic_number)]
                for line in res_file:
                    if line.startswith('REM R1 = '):
                        R1_list = line.split()
                        R1 = R1_list[3]
                        data_list.insert(2,R1)
                    elif line.startswith('REM wR2 = '):
                        wR2_list = line.split()
                        wR2_nocomma = wR2_list[3].replace(",","")
```

```
                    data_list.insert(3,wR2_nocomma)
                    data_list.append("\n")
                res_file.close()
                data_line = "\t".join(data_list)
                results.write(data_line)
    results.close()

os.chdir('..')
files = os.listdir('.')

if "Results.txt" in files:
    data=open("Results.txt","r")
    lines = data.readlines()
    x = []
    y_R1 = []
    y_wR2 = []

    for line in lines:
        if line.startswith("Element") == False:
            data_list = line.split()
            x.append(int(data_list[1]))
            y_R1.append(float(data_list[2]))
            y_wR2.append(float(data_list[3]))

    R1_sort = sorted(zip(x,y_R1))
    wR2_sort = sorted(zip(x,y_wR2))

    new_x,new_R1 = zip(*R1_sort)
    new_x,new_wR2 = zip(*wR2_sort)

    graph, r1 = plt.subplots()
    r1.set_xlabel('Atomic Number')
    r1.set_ylabel('R1')
    r1.yaxis.label.set_color('red')
    r1.scatter(new_x,new_R1,c='red')
    r1.tick_params(axis='y', colors='red')

    wr2 = r1.twinx()
    wr2.set_ylabel('wR2')
    wr2.yaxis.label.set_color('blue')
    wr2.scatter(new_x,new_wR2, c='blue')
    wr2.tick_params(axis='y', colors='blue')

    graph.tight_layout()
    if scattering.lower() in ['n', 'no']:
```

```
            plt.title('Atom Substitution Test for ' + str(original_atom.upper()) +
str(original_atom_number) +
                      ' without Electron Scattering Factors', y=1.01)
        elif scattering.lower() in ['y', 'yes']:
            plt.title('Atom Substitution Test for ' + str(original_atom.upper()) +
str(original_atom_number) +
                      ' with Electron Scattering Factors', y=1.01)
        plt.savefig("Results-as-graph.png", format='png', dpi=300, bbox_inches='tight')
        data.close()
        print("Compiling results into Results.txt and plotting it as a scatterplot... ")

if __name__== "__main__":
    main(original_atom = "", original_atom_number = "", new_atom = "", scattering = "")


    # need to edit .ins file (SFAC and the atom name itself) and
    # run shelxl, and read output from .res file
```

4.8.6.3 batch_atom_substitution.py

```
"""
Autosetup script for running atom_substitution.py for all atoms in a given .ins file.
Will run both scattering and no scattering substitution.
Must have .ins, .hkl, electron scattering factor Excel spreadsheet, and atom_substitution.py
in a single directory.

@author: Lee Joon Kim
"""

import os
import shutil
import re
import pexpect
import sys

def main():
    path = os.getcwd()
    files = os.listdir(".")

    # reads .ins file and makes separate subdirectories for all atoms in the .ins file
    # looks for lines with 7 or more words after 15 lines from the top
    atom_files = []

    for file in files:
        if file.endswith(".ins"):
```

```
            ins_file = open(file, "r")
            ins_line_after_15 = ins_file.readlines()[15:]
            for line in ins_line_after_15:
                ins_line_split = line.split()
                if len(ins_line_split) >= 7 and line.startswith("REM") == False:
                    atoms = ins_line_split[0]
                    atom_files.append(atoms)
            ins_file.close()

    # makes scattering and noscattering directories
    print("Setting up subdirectories and copying files... ")
    for entries in range(len(atom_files)):
        new_path_scattering = os.path.join(path, str(atom_files[entries].upper() + "/scattering"))
        new_path_noscattering = os.path.join(path, str(atom_files[entries].upper() +
"/noscattering"))
        os.makedirs(new_path_scattering, exist_ok = True)
        os.makedirs(new_path_noscattering, exist_ok = True)
        print("Still working... ")

    # copies necessary files into the subdirectories
        for atom_dir in os.scandir(path):
            full_path = os.path.join(path,atom_dir.name)
            if os.path.isdir(full_path) == True:
                for file in files:
                    if file.endswith(".ins") or file.endswith(".hkl"):
                        shutil.copyfile(file, new_path_scattering + "/" +  file)
                        shutil.copyfile(file, new_path_noscattering + "/" + file)
                    if os.path.exists("electron scattering table.xlsx"):
                        shutil.copyfile("electron scattering table.xlsx", new_path_scattering + "/electron
scattering table.xlsx")
                        shutil.copyfile("electron scattering table.xlsx", new_path_noscattering + "/electron
scattering table.xlsx")
                    if os.path.exists("atom_substitution.py"):
                        shutil.copyfile("atom_substitution.py", new_path_scattering
+"/atom_substitution.py")
                        shutil.copyfile("atom_substitution.py", new_path_noscattering
+"/atom_substitution.py")

    # goes into subdirectories and runs atom_substitution.py
    for root, dirs, files in os.walk("."):
        for folder in dirs:
            os.chdir(str(path)+"/"+str(folder))
            current_path = os.getcwd()
            current_path_split = current_path.split("/")  # Note: on Windows this has to be split by \\
            last_term = str(current_path_split[len(current_path_split)-1])
            atom_sub_input = re.findall(r"[^\W\d_]+|\d+", last_term)
```

```
    atom_sub_name = atom_sub_input[0]
    atom_sub_number = "".join(atom_sub_input[1:])

    os.chdir(current_path+"/scattering")

    atom_sub = pexpect.spawnu('python3 atom_substitution.py', timeout = None)
    print("Running atom_substitution.py... ")
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Atom type to change", timeout = None)
    atom_sub.sendline(atom_sub_name)
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Corresponding atom number", timeout = None)
    atom_sub.sendline(atom_sub_number)
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Atom(s) to replace", timeout = None)
    atom_sub.sendline("all")
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Use electron scattering?", timeout = None)
    atom_sub.sendline("y")
    atom_sub.logfile_read= sys.stdout

    atom_sub.expect("Compiling results into Results.txt and plotting it as a scatterplot", timeout = None)

    os.chdir("../noscattering")

    atom_sub = pexpect.spawnu('python3 atom_substitution.py', timeout = None)
    print("Running atom_substitution.py... ")
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Atom type to change", timeout = None)
    atom_sub.sendline(atom_sub_name)
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Corresponding atom number", timeout = None)
    atom_sub.sendline(atom_sub_number)
    atom_sub.logfile_read = sys.stdout

    atom_sub.expect_exact("Atom(s) to replace", timeout = None)
    atom_sub.sendline("all")
    atom_sub.logfile_read = sys.stdout
```

```
        atom_sub.expect_exact("Use electron scattering?", timeout = None)
        atom_sub.sendline("n")
        atom_sub.logfile_read= sys.stdout

        atom_sub.expect("Compiling results into Results.txt and plotting it as a scatterplot",
timeout = None)

   os.chdir(path)
   print("Atom substitution is complete. Results.txt and Results-as-graph.png files are now
available.")

if __name__== "__main__":
   main()
```

4.8.6.4 Electron Scattering Factors Spreadsheet

		a1	b1	a2	b2	a3	b3	a4	b4	c	f'	f"	mu	r	wt
H	1	0.0367	0.5608	0.1269	3.7913	0.236	13.5557	0.129	37.7229	0	0	0	0	0.23	1.008
He	2	0.0406	0.314	0.1276	2.0952	0.1738	7.1369	0.0758	19.4462	0	0	0	0	1.5	4.003
Li	3	0.1198	0.5908	0.3952	6.1114	1.3794	36.7672	1.3889	117.0314	0	0	0	0	1.28	6.941
Be	4	0.118	0.4599	0.4394	4.4816	1.4273	21.5831	1.0661	66.115	0	0	0	0	0.96	9.012
B	5	0.1298	0.418	0.52	3.9214	1.3767	16.6634	0.7666	51.7511	0	0	0	0	0.83	10.811
C	6	0.1361	0.3731	0.5482	3.2814	1.2266	13.0456	0.5971	41.0202	0	0	0	0	0.68	12.011
N	7	0.1372	0.3287	0.5344	2.6733	1.0862	10.3165	0.4547	32.7631	0	0	0	0	0.68	14.007
O	8	0.1433	0.3055	0.5103	2.2683	0.937	8.2625	0.3923	25.6645	0	0	0	0	0.68	15.999
F	9	0.1516	0.2888	0.5193	2.0619	0.822	7.2628	0.3081	22.0262	0	0	0	0	0.64	18.998
Ne	10	0.1575	0.2714	0.5041	1.8403	0.733	6.3059	0.2572	19.164	0	0	0	0	1.5	20.18
Na	11	0.3319	0.495	0.9857	4.0855	1.4885	31.5107	1.9657	118.494	0	0	0	0	1.66	22.99
Mg	12	0.3248	0.4455	0.9243	3.5744	2.0039	24.2702	1.9507	80.7304	0	0	0	0	1.41	24.305
Al	13	0.3582	0.4529	0.9754	3.7745	2.6393	23.3862	1.9103	80.5019	0	0	0	0	1.21	26.982
Si	14	0.3626	0.4281	0.9737	3.557	2.7209	19.3905	1.766	64.3334	0	0	0	0	1.2	28.086
P	15	0.354	0.3941	0.9397	3.181	2.6203	15.6579	1.5707	49.5239	0	0	0	0	1.05	30.974
S	16	0.3478	0.3652	0.9158	2.8915	2.5066	13.0522	1.3884	40.1848	0	0	0	0	1.02	32.065
Cl	17	0.3398	0.3379	0.8908	2.619	2.3878	11.0684	1.2376	33.5378	0	0	0	0	0.99	35.453
Ar	18	0.3409	0.3229	0.8966	2.4778	2.2636	9.7408	1.0786	28.9354	0	0	0	0	1.51	39.098
K	19	0.5658	0.5061	2.4151	5.7656	2.4655	31.9169	3.524	151.259	0	0	0	0	2.03	39.948
Ca	20	0.5474	0.47	2.2793	5.0494	3.1934	29.6928	3.8824	109.5608	0	0	0	0	1.76	40.078
Sc	21	0.5389	0.4446	2.2102	4.5701	3.1187	26.4597	3.4302	98.1283	0	0	0	0	1.7	44.956
Ti	22	0.5398	0.4281	2.1568	4.2236	2.9961	24.1928	3.0751	90.6685	0	0	0	0	1.6	47.867

V	23	0.5412	0.4132	2.1063	3.9256	2.8525	22.3625	2.7967	84.4689	0	0	0	0	1.53	50.942
Cr	24	0.5478	0.4032	2.0737	3.7014	2.3527	20.0893	1.9866	87.923	0	0	0	0	1.39	51.996
Mn	25	0.5552	0.3947	2.0073	3.48	2.5678	19.5862	2.3695	74.9201	0	0	0	0	1.61	54.938
Fe	26	0.5627	0.3859	1.9685	3.3103	2.4527	18.7003	2.175	71.6638	0	0	0	0	1.52	55.845
Co	27	0.573	0.3799	1.9219	3.1572	2.3358	17.8168	2.0177	68.4867	0	0	0	0	1.26	58.693
Ni	28	0.5785	0.3704	1.8679	2.9964	2.2229	16.8507	1.894	65.2843	0	0	0	0	1.24	58.933
Cu	29	0.5932	0.368	1.8344	2.8959	1.808	15.6333	1.3586	70.68	0	0	0	0	1.32	63.546
Zn	30	0.5996	0.3602	1.7763	2.7665	2.024	15.5278	1.6598	60.283	0	0	0	0	1.22	65.39
Ga	31	0.6737	0.3901	1.8457	3.0289	2.6395	19.0448	1.9413	74.1674	0	0	0	0	1.22	69.723
Ge	32	0.666	0.3741	1.7662	2.839	2.8876	17.2911	2.051	62.5149	0	0	0	0	1.17	72.64
As	33	0.6504	0.355	1.6706	2.6169	2.993	15.11	2.0018	50.8113	0	0	0	0	1.21	74.922
Se	34	0.6343	0.3374	1.5698	2.4	3.0362	13.1385	1.9607	42.4344	0	0	0	0	1.22	78.96
Br	35	0.6216	0.3216	1.4943	2.2338	3.0763	11.7075	1.8642	36.7639	0	0	0	0	1.21	79.904
Kr	36	0.624	0.3144	1.4352	2.1394	3.0847	10.606	1.7511	32.4403	0	0	0	0	1.5	83.8
Rb	37	1.018	0.485	2.8882	5.1504	3.5125	26.1762	4.3375	155.1291	0	0	0	0	2.2	85.468
Sr	38	1.0127	0.4721	2.9403	4.9802	3.992	26.8565	5.1441	116.0307	0	0	0	0	1.95	87.62
Y	39	39.9722	0.4427	2.8705	4.5552	4.1571	24.0646	4.6466	99.1688	0	0	0	0	1.9	88.906
Zr	40	0.9592	0.4279	2.8531	4.3237	4.1761	22.151	4.1491	89.7694	0	0	0	0	1.75	91.224
Nb	41	0.9337	0.4074	2.8218	4.0572	3.9092	19.4778	2.988	85.5042	0	0	0	0	1.64	92.906
Mo	42	0.9334	0.3996	2.8581	3.9476	3.8278	18.6153	2.6302	82.5062	0	0	0	0	1.54	95.94
Tc	43	0.9282	0.3894	2.8832	3.8131	3.8985	18.2715	3.1237	74.4844	0	0	0	0	1.47	98
Ru	44	0.9175	0.3776	2.8691	3.6434	3.548	16.4538	2.203	74.459	0	0	0	0	1.46	101.07
Rh	45	0.9046	0.3646	2.8679	3.4888	3.3997	15.5106	2.051	71.3737	0	0	0	0	1.42	102.906
Pd	46	0.7832	0.3094	2.1753	2.5857	3.2003	9.8481	1.4127	37.4035	0	0	0	0	1.39	106.42
Ag	47	0.9085	0.3541	2.9433	3.3319	3.1021	14.7723	1.7092	70.5441	0	0	0	0	1.45	107.868
Cd	48	0.9056	0.3466	2.943	3.2215	3.1814	14.7749	2.1946	62.3544	0	0	0	0	1.54	112.411
In	49	49.966	0.3648	3.1971	3.4534	3.6243	18.3358	2.6351	75.8859	0	0	0	0	1.42	114.818
Sn	50	0.9481	0.3518	3.1011	3.2474	3.8863	17.4348	2.9142	66.0384	0	0	0	0	1.39	118.71
Sb	51	0.9323	0.3405	2.9938	3.0545	4.0829	16.2843	2.9578	56.5387	0	0	0	0	1.39	121.76
Te	52	0.906	0.3256	2.8392	2.817	4.1639	14.6041	3.0782	48.088	0	0	0	0	1.47	126.905
I	53	0.8912	0.315	2.7442	2.6616	4.3316	13.6958	2.9328	43.2296	0	0	0	0	1.4	127.6
Xe	54	0.8749	0.3046	2.6276	2.4959	4.4112	12.5908	2.8758	38.5666	0	0	0	0	1.5	131.293
Cs	55	1.2968	0.4462	3.8609	4.3194	5.6613	24.1136	5.6489	168.3481	0	0	0	0	2.44	132.906
Ba	56	1.3111	0.4432	3.9542	4.3226	5.8693	24.3857	7.0998	128.5976	0	0	0	0	2.15	137.327
La	57	1.2872	0.4283	3.8478	4.0782	6.0581	22.4363	6.5606	109.8403	0	0	0	0	2.07	138.906
Ce	58	1.3019	0.4256	3.8491	4.0288	5.8568	22.002	6.3222	107.6524	0	0	0	0	2.04	140.116
Pr	59	1.3354	0.4305	3.8652	4.0368	5.3	22.1273	6.4296	117.0334	0	0	0	0	2.03	140.908
Nd	60	1.3473	0.4266	3.8589	3.9844	5.1226	21.8395	6.2264	114.8891	0	0	0	0	2.01	144.24

Pm	61	1.3631	0.4253	3.8462	3.9386	4.9578	21.633	6.0228	112.8589	0	0	0	0	1.99	145
Sm	62	1.3702	0.4204	3.827	3.8747	4.8068	21.384	5.8402	110.9527	0	0	0	0	1.98	150.36
Eu	63	1.3963	0.423	3.8201	3.8621	4.6869	21.4604	5.6321	110.3872	0	0	0	0	1.98	151.964
Gd	64	1.3784	0.4105	3.7349	3.7014	4.8549	20.2877	5.2564	96.0364	0	0	0	0	1.96	157.25
Tb	65	1.4257	0.4182	3.7865	3.7931	4.4336	21.3185	5.2698	107.633	0	0	0	0	1.94	158.925
Dy	66	1.4104	0.4076	3.6934	3.6339	4.3012	20.3594	5.1904	103.6978	0	0	0	0	1.92	162.5
Ho	67	1.3805	0.3938	3.5395	3.4131	4.4117	18.4989	5.0148	86.9074	0	0	0	0	1.92	164.93
Er	68	1.4391	0.4038	3.6351	3.5707	4.0572	20.0727	4.9059	100.4842	0	0	0	0	1.89	167.259
Tm	69	1.4561	0.4028	3.6041	3.5468	3.936	19.9521	4.7711	98.6682	0	0	0	0	1.9	168.934
Yb	70	1.4617	0.3986	3.5735	3.4988	3.8563	19.905	4.6254	97.9782	0	0	0	0	1.87	173.04
Lu	71	1.4381	0.3874	3.4689	3.3304	4.0824	18.8077	4.4599	84.5078	0	0	0	0	1.87	174.967
Hf	72	1.4209	0.3772	3.3984	3.2109	4.2315	17.772	4.1	76.0916	0	0	0	0	1.75	178.49
Ta	73	1.4163	0.3714	3.3456	3.1289	4.3281	16.8719	3.7406	70.3632	0	0	0	0	1.7	180.948
W	74	1.4109	0.3653	3.2997	3.0517	4.3698	15.963	3.4408	65.7545	0	0	0	0	1.62	183.84
Re	75	1.3898	0.3548	3.2373	2.9382	4.3683	14.85	3.2472	60.6048	0	0	0	0	1.51	186.207
Os	76	1.3877	0.3502	3.2082	2.8796	4.3731	14.1535	2.9958	57.3692	0	0	0	0	1.44	190.23
Ir	77	1.3783	0.344	3.1637	2.7975	4.3549	13.364	2.805	54.3356	0	0	0	0	1.41	192.217
Pt	78	1.3489	0.3322	3.0674	2.6601	4.2095	12.0128	2.1738	50.3882	0	0	0	0	1.36	195.078
Au	79	1.3413	0.3262	3.0559	2.6076	4.1659	11.5043	2.0022	48.4954	0	0	0	0	1.36	196.967
Hg	80	1.3592	0.3265	3.1327	2.6458	4.1897	11.6919	2.2748	47.9347	0	0	0	0	1.32	200.59
Tl	81	1.4735	0.35	3.6368	3.0849	4.395	14.9338	2.5803	65.0072	0	0	0	0	1.45	204.383
Pb	82	1.47	0.3452	3.6387	3.034	4.5674	14.8562	2.9103	60.2174	0	0	0	0	1.46	207.2
Bi	83	1.4616	0.3396	3.6426	2.9739	4.7876	14.8388	3.1935	56.7362	0	0	0	0	1.48	208.98
Po	84	1.4255	0.3274	3.5214	2.8031	4.9058	13.8305	3.4925	50.1382	0	0	0	0	1.4	209
At	85	1.417	0.3224	3.4829	2.7289	5.0951	13.4904	3.4639	46.2094	0	0	0	0	1.21	210
Rn	86	1.4001	0.3153	3.4082	2.6261	5.2516	12.9003	3.425	42.3929	0	0	0	0	1.5	222
Fr	87	1.7697	0.3969	4.682	3.995	6.6701	23.0274	5.5268	154.8886	0	0	0	0	2.6	223
Ra	88	1.7704	0.3938	4.6854	3.9313	6.8034	22.4168	7.2352	120.9717	0	0	0	0	2.21	226
Ac	89	1.7346	0.3816	4.5959	3.7454	7.0436	21.133	7.0521	102.1159	0	0	0	0	2.15	227
Th	90	1.7012	0.3704	4.5007	3.5712	7.2504	19.8632	6.6077	89.3562	0	0	0	0	2.06	231.036
Pa	91	1.7266	0.3729	4.6123	3.6159	6.7855	19.6433	6.3876	95.272	0	0	0	0	2	232.038
U	92	1.7417	0.3729	4.7063	3.6337	6.6509	19.6579	5.9929	95.0805	0	0	0	0	1.96	237
Np	93	1.7327	0.3683	4.6782	3.5375	6.4465	18.8556	5.8528	91.5521	0	0	0	0	1.9	238.029
Pu	94	1.7491	0.3683	4.7933	3.5651	5.9841	18.8727	5.6139	100.8062	0	0	0	0	1.87	243
Am	95	1.7519	0.3654	4.8096	3.5215	5.8155	18.512	5.4119	98.8221	0	0	0	0	1.8	244
Cm	96	1.7284	0.3577	4.7162	3.3776	5.9338	17.7168	5.2896	86.4936	0	0	0	0	1.69	247
Bk	97	1.7298	0.3554	4.7198	3.3269	5.7527	17.3659	5.1576	84.3294	0	0	0	0	1.54	247
Cf	98	1.6401	0.3332	4.3905	2.983	5.4504	14.6892	5.4542	73.0723	0	0	0	0	1.83	251

Table 4.8 Electron scattering factors formatted in an Excel spreadsheet for atom substitution test.

4.8.7 Supplementary Notes and References

(1) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The CryoEM Method MicroED as a Powerful Tool for Small Molecule Structure Determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(2) Nannenga, B. L.; Shi, D.; Leslie, A. G. W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930.

(3) Program ser2smv obtained from https://cryoem.ucla.edu/downloads/snapshots.

(4) Hattne, J; Reyes, F. E.; Nannenga, B. L.; Shi, D.; de la Cruz, M. J.; Leslie, A. G. W.; Gonen, T. MicroED data collection and processing. *Acta Cryst.* **2015**, *A71*, 353–360.

(5) Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.

(6) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010,** *D66*, 133–144.

(7) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(8) Sheldrick, G. M. SHELXT – Integrated space-group and crystal-structure determination. *Acta Cryst.* **2015**, *A71*, 3–8.

(9) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(10) Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. ShelXle: a Qt graphical user interface for SHELXL. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.

(11) Delano, W. *The PyMOL Molecular Graphics System* version 2.3.3 (Schrödinger LLC, 2019); http://www.pymol.org.

(12) Van Rossum, G.; Drake, F. L. *Python 3 Reference Manual* (CreateSpace, 2009).

(13) Peng, L.-M. Electron atomic scattering factors and scattering potentials of crystals. *Micron* **1999**, *30*, 625–648.

CHAPTER FIVE

Structural Revision of the Lomaivitolicins

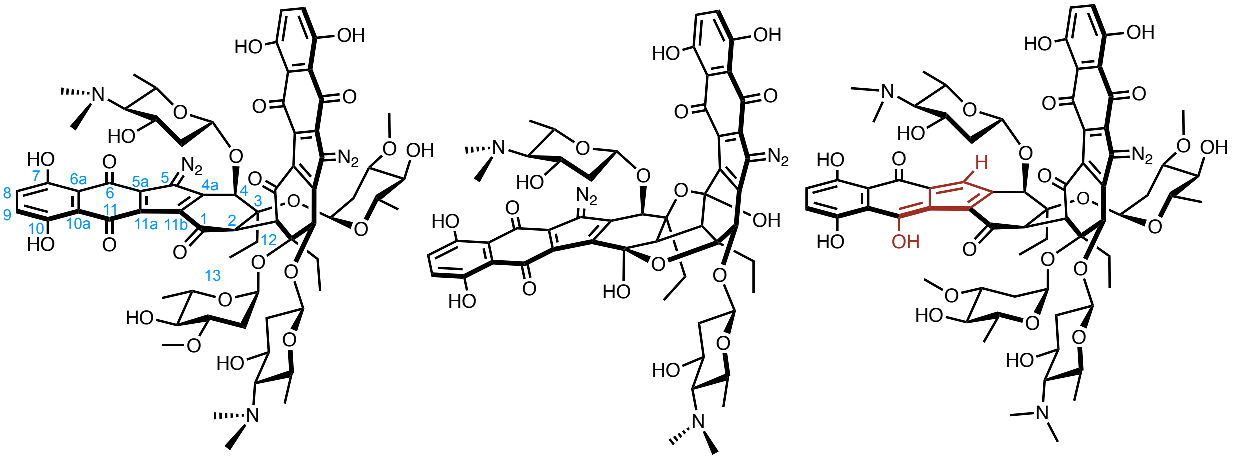
Adapted from: Lee Joon Kim, Mengzhao Xue, Xin Li, Zhi Xu, Eric Paulson, Brandon Mercado, Hosea M. Nelson *J. Am. Chem. Soc.* **2021**, *143*, 6578–6585.

5.1 Abstract

The lomaiviticins are dimeric bacterial metabolites that exhibit potent cytotoxicity and contain unusual diazocyclopentadiene functional groups decorated with deoxyglycoside residues. Due to the few proton-attached carbon atoms in the core structure, structure determination using NMR spectroscopy has been challenging. Prior structure elucidation efforts proposed that the two halves of the lomaiviticins are connected *via* a single carbon–carbon bond at the C6 position of a 4,5-dihydroxycyclohex-2-ene-1-one, but this proposal has not been tested as a result of unsuccessful attempts at total synthesis or X-ray crystallography. This study reports the first solid-state structure of lomaiviticins obtained through MicroED studies, unambiguously revealing that (–)-lomaiviticin C contains a 4,6-dihydroxy-cyclohex-2-ene-1-one residue with a bridging carbon–carbon bond at C5, and that the orientation of the cyclohexenone ring and configuration of the secondary glycoside are reversed compared to the original structure assignment. The structure revision is supported by high-field (800 MHz) NMR analysis and DFT calculations of the expected ^{13}C chemical shifts and C–H coupling constants. As the interconversion of lomaiviticins A, B, and C has been previously demonstrated, these findings apply to each isolate, elucidating the structures of this natural product family for the first time and demonstrating the power of MicroED analysis in natural product structure determination.

5.2 Introduction

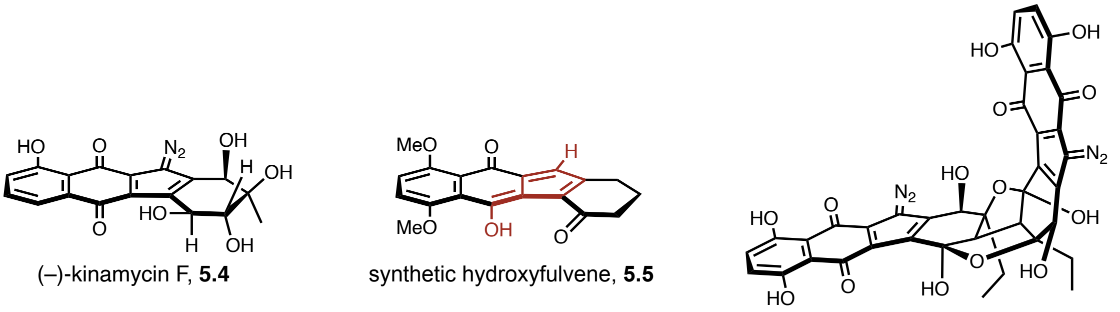
The lomaiviticins are cytotoxic metabolites isolated from various strains of marine bacteria *Salinispora*.[1] This family of natural products is characterized by a dimeric core structure that contains diazocyclopentadiene functional groups and deoxyglycoside residues. (–)-Lomitivicins A and B (**5.1** and **5.2**, respectively) were first reported by He and coworkers, who probed their structures using HRMS and extensive NMR experiments.[2] Since then, all aspects of the lomaiviticin structure have been studied, such as the relative stereochemistry and anomeric configuration of the carbohydrate residues.[3] The presence of the diazofluorene unit was inferred from kinamycins, such as (–)-kinamycin F (**5.4**), which are monomeric diazofluorene natural products whose structures have already been elucidated.[4] The structure determination of the 2,2'-bicyclohexyl core was particularly challenging due to the low number of proton-attached carbons for NMR analysis: only 6 out of 19 carbon atoms are proton-attached, and two of these carbons (C8, C9) are separated due to the diazofluorene unit, leaving the proton signals from only C2, C4, C12, and C13 to guide structure elucidation. 2-D NMR experiments such as NOESY, HMQC, and HMBC led to the proposed structures of (–)-lomaiviticins A and B as shown in Figure 5.1, where the core structure is connected by a bridging C–C bond between C2 and C2'.



(–)-lomaiviticin A, **5.1**
2001 structure assignment

(–)-lomaiviticin B, **5.2**
2001 structure assignment

(–)-lomaiviticin C, **5.3**
2012 structure assignment



(–)-kinamycin F, **5.4**

synthetic hydroxyfulvene, **5.5**

aglycon structure, **5.6**
synthesized in 2011

Figure 5.1 Originally proposed structures for (–)-lomaiviticins A, B, and C (**5.1**, **5.2**, and **5.3**, respectively), and the structures of related molecules.

In 2012, (–)-lomaiviticin C (**5.3**) was isolated by the Herzon laboratory, who showed that (–)-lomaiviticin C is the product of formal hydrodediazotization of (–)-lomaiviticin A, and that semisynthesis of (–)-lomaiviticin A can be achieved by an azide transfer reaction of natural (–)-lomaiviticin C (Scheme 5.1).[5] Based on this successful conversion from (–)-lomaiviticin C to A, and experimental data from NMR and mass spectrometry, the structure of (–)-lomaiviticin C was proposed to have a hydroxyfulvene. The presence of a hydroxyfulvene unit was supported by the NMR data of fully synthetic hydroxyfulvenes, such as **5.5**. Furthermore, the carbohydrate residues were isolated by acid digestion of (–)-lomaiviticin C, and their absolute stereochemistry was determined using optical rotation data and synthetic standards.

natural (–)-lomaiviticin C $\xrightarrow[\text{TFA}]{\text{TfN}_3}$ semisynthetic (–)-lomaiviticin A $\equiv$ natural (–)-lomaiviticin A

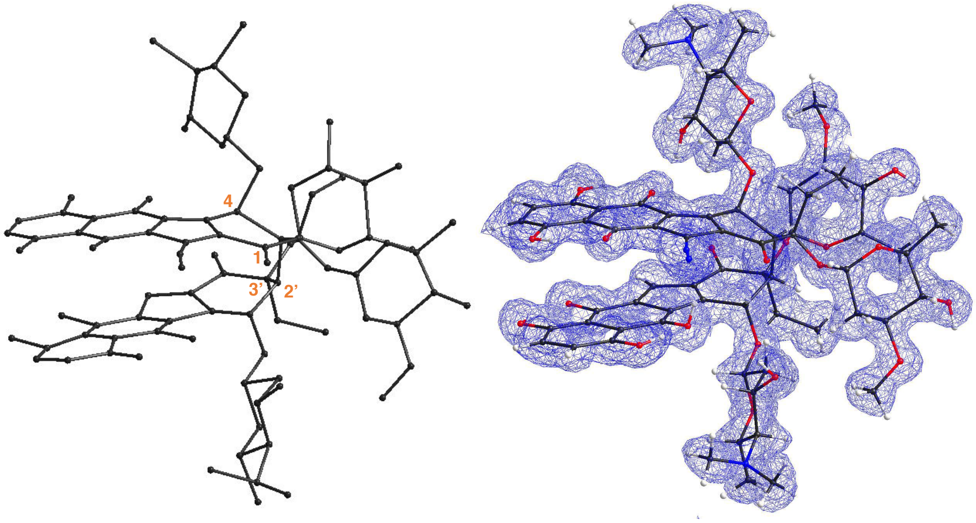
natural (–)-lomaiviticin A $\xrightarrow[\text{CH}_3\text{OH}]{\text{TFA}}$ semisynthetic (–)-lomaiviticin B $\equiv$ natural (–)-lomaiviticin B

Scheme 5.1 The structural homology between (–)-lomaiviticin A–C established by semisynthesis.[5,6]

Despite decades of extensive structural studies of these natural products, single crystal X-ray structures of these compounds have remained elusive. Moreover, synthetic efforts to forge structures **5.1**, **5.2**, and **5.3** have been unsuccessful,[6,7] and the synthetic lomaiviticins could not be linked to their natural forms with certainty. Unambiguous structural determination of these metabolites would establish the biological and medicinal relevance of these synthetic reports, as well as facilitate mechanistic studies to elucidate the lomaiviticins' mode of action. To this end, this report utilizes microcrystal electron diffraction, or MicroED, to probe the structure of these metabolites. MicroED is an emerging cryoelectron microscopy technique capable of determining structures from microcrystals present even in seemingly amorphous samples.[8] This sensitivity may provide a unique advantage to obtain structures of these extensively studied natural products.

5.3 Structural Determination and Revision of (–)-Lomaiviticin C

From the family of the lomaiviticins, (–)-lomaiviticin C (**5.3**) was chosen for structural analysis due to its higher abundance,[5] better stability,[9] and less cytotoxicity.[5] During the initial studies, 2.0 µL droplet containing a 0.1 mg/mL solution of **5.3** (~200 ng) in a mixture of tetrahydrofuran and acetonitrile was deposited onto a carbon TEM grid. The droplet was then allowed to evaporate at ambient temperature to yield microcrystalline domains that diffracted to >2 Å resolution. While this was a promising result, the resolution of these preliminary diffraction patterns was not sufficiently high for *ab initio* structural determination. Screening various conditions revealed that slow evaporation over 30 days of the **5.3** solution at –20 °C slowed the crystal growth sufficiently to generate microcrystals that diffracted to 1.05 Å. Due to the loss of high resolution reflections from radiation damage, data collection was performed under cryogenic temperature. Data from four crystals were merged to solve the crystal structure of **5.3** using direct methods (Figure 5.2).

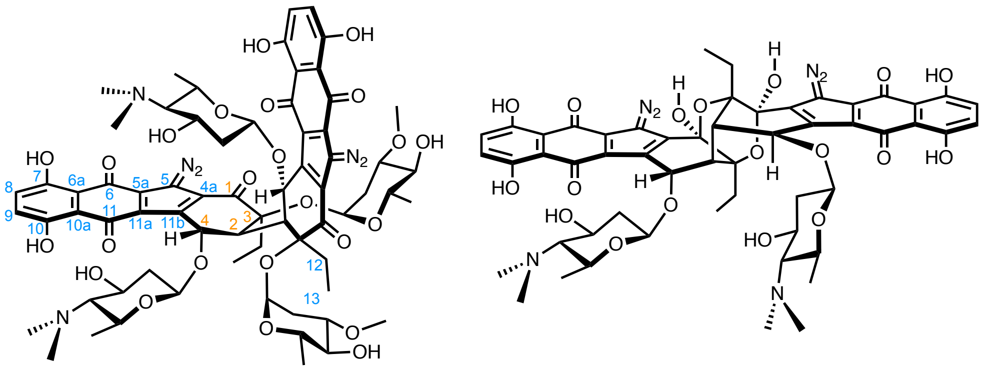


Preliminary solution directly from SHELXD

(–)-lomaiviticin C, **5.3'** *Refined MicroED structure*

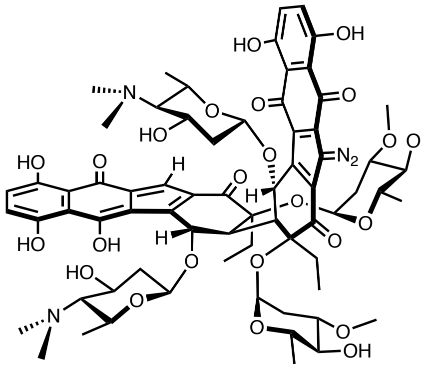
Figure 5.2 Preliminary (left) and fully refined (right) MicroED structures of (–)-lomaiviticin C.

To our surprise, the preliminary solution from SHELXD, obtained without any human input other than the molecular formula, revealed major structural discrepancies from the structure of the lomaiviticin core proposed based on NMR analysis. C1/C4 (C1'/C4') and C2/C3 (C2'/C3') atom positions were exchanged, and the configurations of C4 and C4' were inverted. While the connectivity of the atoms were easily determined using the MicroED data, the heteroatom positions of the aminosugar residues were more difficult to assign using the MicroED data alone, and results from earlier chemical degradation studies were used to complement the crystal structure and assign the heteroatoms.[5] Using the MicroED data, structures of lomaiviticins A–C can be revised as **5.1'**, **5.2'**, and **5.3'**, as seen in Figure 5.3.



(–)-lomaiviticin A, **5.1’**
2021 structure assignment

(–)-lomaiviticin B, **5.2’**
2021 structure assignment



(–)-lomaiviticin C, **5.3’**
2021 structure assignment

Figure 5.3 Revised structures of (–)-lomaiviticins A, B, and C (**5.1’**, **5.2’**, and **5.3’**, respectively) based on data reported in this study. The assignments that are exchanged in the revised structures are shown in orange in **5.1’**.

The structure revision was also supported by Monte Carlo molecular dynamics simulation[10] carried out on the two structures of (–)-lomaiviticin C based on the original NMR analysis and MicroED (**5.3** and **5.3’** respectively, Figure 5.4). Notably, the geometry of the lowest energy conformer of **5.3’** was in qualitative agreement with the geometry of the MicroED structure. High-field (800 MHz) NMR data was acquired for (–)-lomaiviticin C in acetone-d_6 and revealed key correlations that further support the structural revision (Figure 5.5). These correlations are discussed in depth in the adapted article.[11]

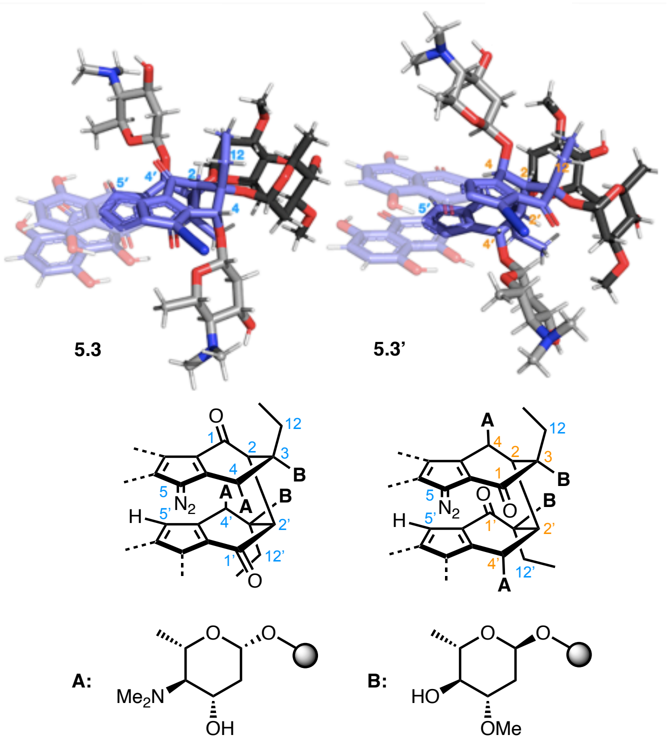


Figure 5.4 Lowest-energy conformers (top) and line structures (bottom) of structures **5.3** and **5.3'**. For clarity, the naphthoquinone residues are shaded in the lowest-energy conformers and omitted from the line structures.

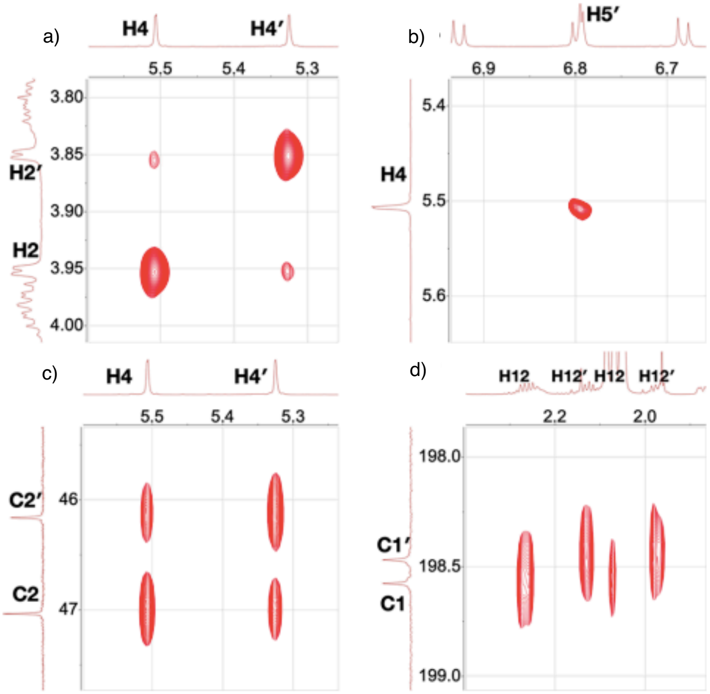
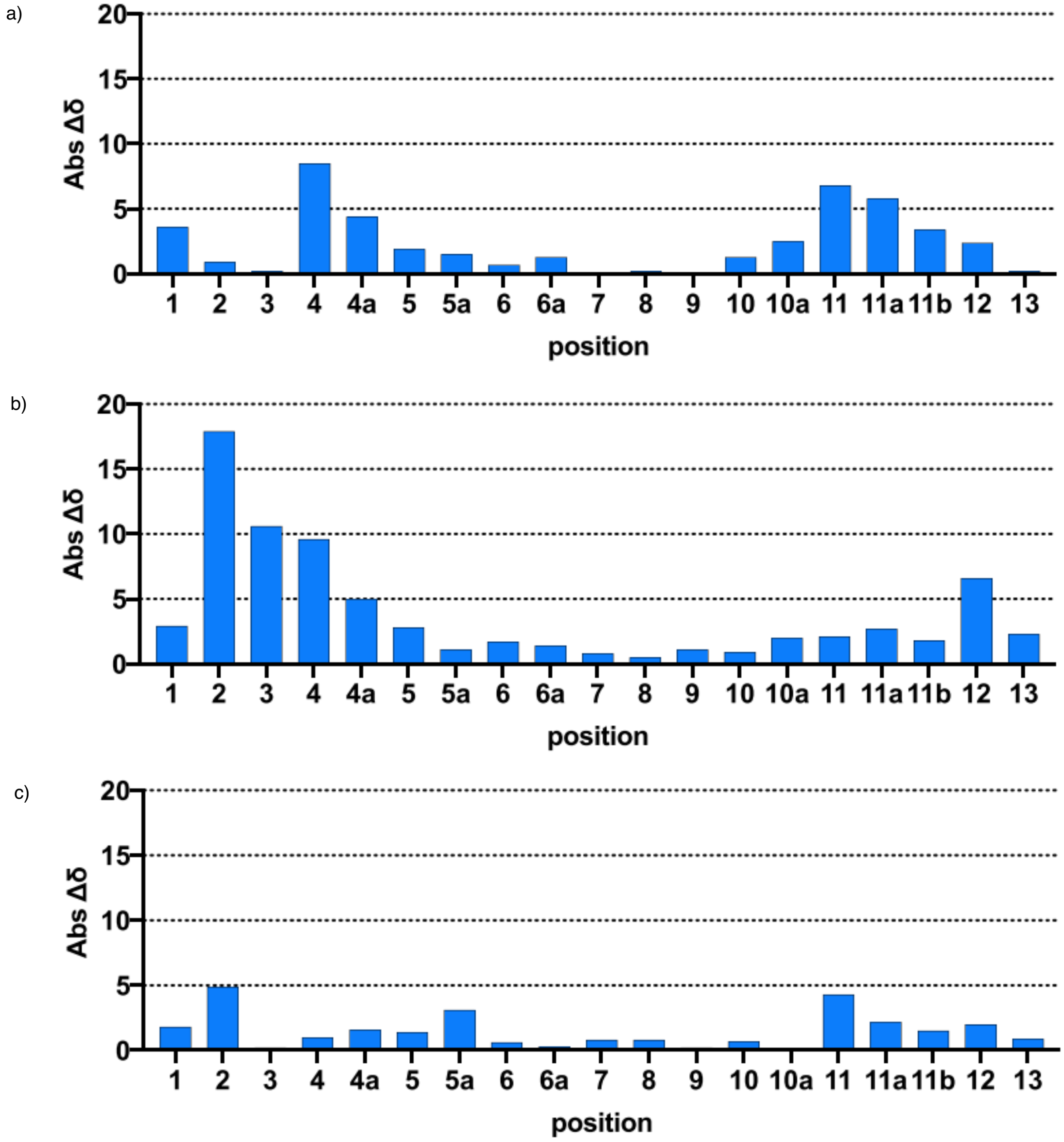
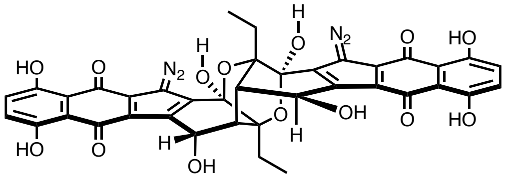


Figure 5.5 Diagnostic ROESY and HMBC correlations in the NMR spectra of (–)-lomaiviticin C supporting structure **5.3'**. (a) ROESY correlations between (H4, H2), (H4', H2'), (H4, H2'), and (H4', H2). (b) ROESY correlation between (H4, H5'). (c) HMBC correlations between (H4, C2'), (H4, C2), (H4', C2'), and (H4', C2). (d) HMBC correlations between (H12, C1) and (H12', C1').

Computational modeling was also carried out to provide further support for the structure revision.[12] Lomaiviticin aglycon (**5.6**, Figure 5.1 and **5.6'**, Figure 5.6) and (–)-lomaiviticin B (**5.2**, Figure 5.1 and **5.2'**, Figure 5.3), which are less conformationally flexible than (–)-lomaiviticin A and C, were used to reduce the computational cost. The expected ^{13}C NMR chemical shifts of **5.6** was calculated and compared with the experimental ^{13}C shifts[6,7i] to benchmark the computational approach. Good agreement observed between calculated and experimental values (overall root-mean-square = 3.39) supported this computational approach as a viable strategy to determine the expected ^{13}C spectra of the lomaiviticins.





aglycon structure **5.6’**

Figure 5.6 Comparison of calculated and experimental ^{13}C chemical shifts between aglycon structures and (–)-lomaiviticin B. (a) Absolute difference between calculated and experimental ^{13}C chemical shifts of the aglycon structure **5.6**. RMS = 3.39. (b) Absolute difference between experimental ^{13}C chemical shifts of the aglycon structure **5.6** and natural (–)-lomaiviticin B. RMS = 5.82. (c) Absolute difference between calculated and experimental ^{13}C chemical shifts of the aglycon structure **5.6’** and (–)-lomaiviticin B. RMS = 1.99.

Experimental ^{13}C chemical shifts of **5.6** were compared with those of natural (–)-lomaiviticin B (Figure 5.6) to obtain high root-mean-square (RMS) value (RMS = 5.82). In particular, large differences were observed between carbons C2, C3 and C4. Then, the expected ^{13}C chemical shifts of **5.6’** were determined, whose structure reflects the structural revision. When the expected chemical shifts of **5.6’** were compared to those of natural (–)-lomaiviticin B, they showed good agreement with an overall RMS value of 1.99 and all experimental shifts within 5 ppm of the calculated values.

Similarly, the expected ^{13}C NMR chemical shifts of **5.2** and **5.2’** were calculated and compared with the experimental chemical shifts of natural (–)-lomaiviticin B (Figure 5.7). While the structure **5.2** showed large discrepancies, particularly at key positions such as C1, C2, C3, C4, C11b, and C12, structure **5.2’** showed good agreement with the chemical shifts of the natural metabolite (RMS = 2.83).

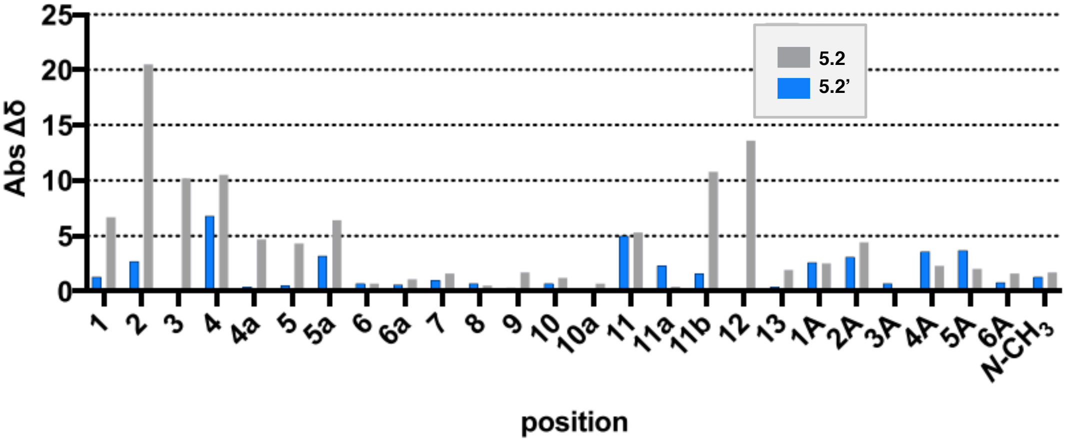


Figure 5.7 Absolute difference between calculated ^{13}C chemical shifts of structures **5.2** and **5.2’** and natural (–)-lomaiviticin B. Gray series: **5.2** (RMS = 6.63); blue series: **5.2’** (RMS = 2.83).

Finally, DFT calculations were carried out to determine the expected values of $^{n}J_{C,H}$ coupling constants for H2, H2’, H4, and H4’ of **5.3** and **5.3’** to compare with those of natural (–)-lomaiviticin C. The single lowest energy conformers of **5.3** and **5.3’** identified by molecular dynamics were used as input and optimized by DFT (ωB97X-D/6-31G*) before determining the couplings. The theoretical couplings derived from **5.3’** were in better agreement (RMS = 1.80 vs. RMS = 3.87 for **5.3**), with the couplings between (H2, C4’) and (H2’, C4) being the most diagnostic: the calculated ${}^{3}J_{\mathrm{C,H}}$ and ${}^{4}J_{\mathrm{H,H}}$ couplings for structure **5.3’** were in reasonable agreement with the experimental values, as opposed to those of structure **5.3**.

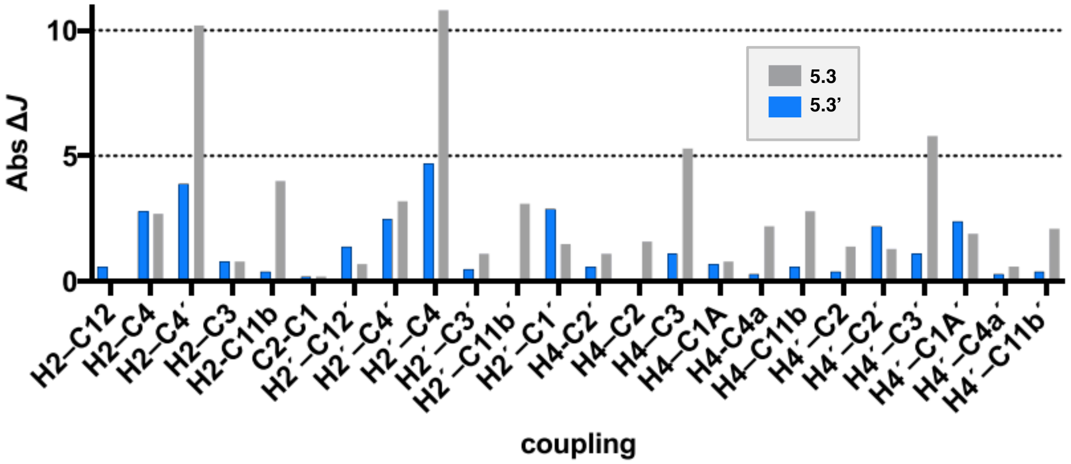


Figure 5.8 Absolute difference between experimental ${}^{3}J_{\mathrm{C,H}}$ coupling constants for H2, H2’, H4, and H4’ of natural (–)-lomaiviticin C and those expected for structures **5.3** and **5.3’** based on DFT calculations.

5.4 Conclusion

The lomaiviticins are highly cytotoxic bacterial metabolites containing an unusual diazotetrahydrobenzo[*b*]fluorene functional group as well as the deoxyglycosides L-oleandrose and *N,N*-dimethyl-L-pyrrolosamine. Due to the low number of proton signals in the core structure, determining structures of the lomaiviticins using NMR analysis alone was exceedingly difficult. For two decades since their initial disclosure in 2001, synthetic campaigns to forge these natural products have been unsuccessful, preventing spectroscopic data comparison between synthetic material and natural sample to determine the structures of these metabolites. Moreover, single crystal X-ray crystallographic data has not been reported for any of the lomaiviticins.

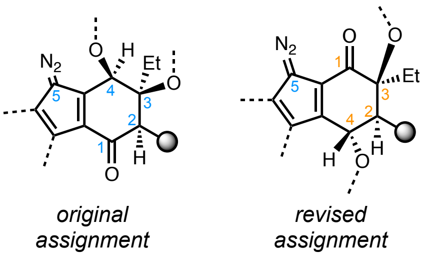
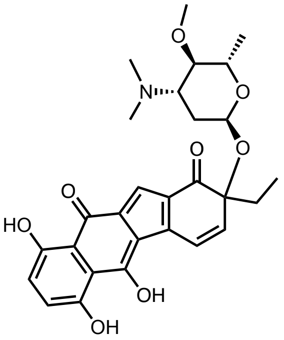


Figure 5.9 Original (left) and revised (right) core structures of the lomaiviticins.

In this work, MicroED was used to determine the structure of (–)-lomaiviticin C and revise the core structure of this family of metabolites. The structure revision (Figure 5.9) is supported by high-field NMR spectroscopy and computational modeling. Because natural (–)-lomaiviticin C has been shown to be converted into semisynthetic (–)-lomaiviticin A, the findings of this work apply to not only (–)-lomaiviticin C but also its related metabolites. The oxidation pattern found in the revised structures of **5.1'**, **5.2'**, and **5.3'** has been reported in other isolates, such as nenestatin A (**5.7**, Figure 5.10)[13] and in intermediates in kinamycin biosynthesis.[4c] Interestingly, biosynthetic precursors to nenestatin A have been found to undergo nucleophilic addition and dimerization,[13b] which may have implications for lomaiviticin biosynthesis. Taken together, these results demonstrate the importance of unambiguous structure determination of these complex metabolites. This structure revision will facilitate future synthetic efforts as well as further investigations into the lomaiviticins' mode of interaction with DNA to induce double-strand breaks.[14]



nenestatin A, **5.7**

Figure 5.10 Structure of nenestatin A.

5.5 Notes and References

(1) For a recent review, see: Herzon, S. B.; Woo, C. M. The diazofluorene antitumor antibiotics: Structural elucidation, biosynthetic, synthetic, and chemical biological studies. *Nat. Prod. Rep.* **2012**, *29*, 87–118.

(2) He, H.; Ding, W. D.; Bernan, V. S.; Richardson, A. D.; Ireland, C. M.; Greenstein, M.; Ellestad, G. A.; Carter, G. T. Lomaiviticins A and B, potent antitumor antibiotics from *Micromonospora lomaivitiensis*. *J. Am. Chem. Soc*. **2001**, *123*, 5362–5363.

(3) (a) Burg, R. W.; Miller, B. M.; Baker, E. E.; Birnbaum, J.; Currie, S. A.; Hartman, R.; Kong, Y. L.; Monaghan, R. L.; Olson, G.; Putter, I.; Tunac, J. B.; Wallick, H.; Stapley, E. O.; Oiwa, R.; Omura, S. Avermectins, new family of potent anthelmintic agents: Producing organism and fermentation. *Antimicrob. Agents Chemother.* **1979**, *15*, 361–367. (b) Lam, K. S.; Hesler, G. A.; Gustavson, D. R.; Berry, R. L.; Tomita, K.; MacBeth, J. L.; Ross, J.; Miller, D.; Forenza, S. Pyrrolosporin A, a new antitumor antibiotic from *Micromonospora* sp. C39217-r109−7. I. Taxonomy of producing organism, fermentation and biological activity. *J. Antibiot.* **1996**, *49*, 860–864. (c) Schroeder, D. R.; Colson, K. L.; Klohr, S. E.; Lee, M. S.; Matson, J. A.; Brinen, L. S.; Clardy, J. Pyrrolosporin A, a new antitumor antibiotic from *Micromonospora* sp. C39217-r109−7. II. Isolation, physico-chemical properties, spectroscopic study and x-ray analysis. *J. Antibiot.* **1996**, *49*, 865–872.

(4) (a) Cone, M. C.; Seaton, P. J.; Halley, K. A.; Gould, S. J. New products related to kinamycin from streptomyces murayamaensis. I. Taxonomy, production, isolation and biological properties. *J. Antibiot.* **1989**, *42*, 179–188. (b) Seaton, P. J.; Gould, S. J. New products related to kinamycin from streptomyces murayamaensis. II. Structures of pre-kinamycin, keto-anhydrokinamycin, and kinamycins e and f. *J. Antibiot*. **1989**, *42*, 189–197. (c) Gould, S. J. Biosynthesis of the

kinamycins. *Chem. Rev.* **1997**, *97*, 2499–2510. (d) Marco-Contelles, J.; Molina, M. T. Naturally occurring diazo compounds: The kinamycins. *Curr. Org. Chem.* **2003**, *7*, 1433–1442.

(5) Woo, C. M.; Beizer, N. E.; Janso, J. E.; Herzon, S. B. Isolation of lomaiviticins C–E. Transformation of lomaiviticin C to lomaiviticin A, complete structure elucidation of lomaiviticin A, and structure-activity analyses. *J. Am. Chem. Soc.* **2012**, *134*, 15285–15288.

(6) Woo, C. M.; Gholap, S. L.; Lu, L.; Kaneko, M.; Li, Z.; Ravikumar, P. C.; Herzon, S. B. Development of enantioselective synthetic routes to (−)-kinamycin f and (−)-lomaiviticin aglycon. J. Am. Chem. Soc. 2012, 134, 17262.

(7) (a) Nicolaou, K. C.; Denton, R. M.; Lenzen, A.; Edmonds, D. J.; Li, A.; Milburn, R. R.; Harrison, S. T. Stereocontrolled synthesis of model core systems of lomaiviticins A and B. *Angew. Chem., Int. Ed.* **2006**, *45*, 2076–2081. (b) Morris, W. J.; Shair, M. D. Stereoselective synthesis of 2-deoxy-β-glycosides using anomeric O-alkylation/arylation. *Org. Lett.* **2009**, *11*, 9–12. (c) Krygowski, E. S.; Murphy-Benenato, K.; Shair, M. D. Enantioselective synthesis of the central ring system of lomaiviticin A in the form of an unusually stable cyclic hydrate. *Angew. Chem., Int. Ed.* **2008**, *47*, 1680–1684. (d) Zhang, W.; Baranczak, A.; Sulikowski, G. A. Stereocontrolled assembly of the C3/C3′ dideoxy core of lomaiviticin A/B and congeners. *Org. Lett.* **2008**, *10*, 1939–1941. (e) Gholap, S. L.; Woo, C. M.; Ravikumar, P. C.; Herzon, S. B. Synthesis of the fully glycosylated cyclohexenone core of lomaiviticin A. *Org. Lett.* **2009**, *11*, 4322–4325. (f) Nicolaou, K. C.; Nold, A. L.; Li, H. Synthesis of the monomeric unit of the lomaiviticin aglycon. *Angew. Chem., Int. Ed.* **2009**, *48*, 5860–5863. (g) Morris, W. J.; Shair, M. D. Synthesis of the *N*-(*tert*-butyloxycarbonyl)-*O*-triisopropylsilyl-D-pyrrolosamine glycal of lomaiviticins A and B via epimerization of L-threonine. *Tetrahedron Lett.* **2010**, *51*, 4310–4312. (h) Lee, H. G.; Ahn, J. Y.; Lee, A. S.; Shair, M. D. Enantioselective synthesis of the lomaiviticin

aglycon full carbon skeleton reveals remarkable remote substituent effects during the dimerization event. *Chem. - Eur. J.* **2010**, *16*, 13058–13062. (i) Herzon, S. B.; Lu, L.; Woo, C. M.; Gholap, S. L. 11-step enantioselective synthesis of (−)-lomaiviticin aglycon. *J. Am. Chem. Soc.* **2011**, *133*, 7260–7263. (j) Scully, S. S.; Porco, J. A. Asymmetric total synthesis of the epoxykinamycin FL-120B′. *Angew. Chem., Int. Ed*. **2011**, *50*, 9722–9726. (k) Baranczak, A.; Sulikowski, G. A. Synthetic studies directed toward dideoxy lomaiviticinone lead to unexpected 1,2-oxazepine and isoxazole formation. *Org. Lett*. **2012**, *14*, 1027–1029. (l) Scully, S. S.; Porco, J. A. Studies toward the synthesis of the epoxykinamycin FL-120B′: Discovery of a decarbonylative photocyclization. *Org. Lett*. **2012**, *14*, 2646–2649. (m) Feldman, K. S.; Selfridge, B. R. Enantioselective synthesis of the *ent*-lomaiviticin A bicyclic core. *Org. Lett*. **2012**, *14*, 5484–5487. (n) Lee, A. S.; Shair, M. D. Synthesis of the C4-*epi*-lomaiviticin B core reveals subtle stereoelectronic effects. *Org. Lett*. **2013**, *15*, 2390–2393. (o) Feldman, K. S.; Selfridge, B. R. Synthesis studies on the lomaiviticin A aglycone core: Development of a divergent, two-directional strategy. *J. Org. Chem*. **2013**, *78*, 4499–4511. (p) Rose, J. A.; Mahapatra, S.; Li, X.; Wang, C.; Chen, L.; Swick, S. M.; Herzon, S. B. Synthesis of the bis(cyclohexenone) core of (−)-lomaiviticin A. *Chem. Sci*. **2020**, *11*, 7462–7467. (q) Nicolaou, K. C.; Chen, Q.; Li, R.; Anami, Y.; Tsuchikama, K. Total synthesis of the monomeric unit of lomaiviticin A. *J. Am. Chem. Soc*. **2020**, *142*, 20201–20207. (r) Kaneko, M.; Li, Z.; Burk, M.; Colis, L.; Herzon, S. B. Synthesis and biological evaluation of (2*S*,2′*S*)-lomaiviticin A. *J. Am. Chem. Soc.* **2021**, *143*, 1126–1132.

(8) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(9) Xue, M.; Herzon, S. B. Mechanism of nucleophilic activation of (−)-lomaiviticin A. *J. Am. Chem. Soc*. **2016**, *138*, 15559–15562.

(10) *Spartan 18*; Wavefunction, Inc., Irvine, CA.

(11) Kim, L. J.; Xue, M.; Li, X.; Xu, Z.l Paulson, E.; Mercado, B. Q.; Nelson, H. M.; Herzon, S. Structural revision of the lomaiviticins. *J. Am. Chem. Soc.* **2021**, *143*, 6578–6585.

(12) Hehre, W.; Klunzinger, P.; Deppmeier, B.; Driessen, A.; Uchida, N.; Hashimoto, M.; Fukushi, E.; Takata, Y. Efficient protocol for accurately calculating ^{13}C chemical shifts of conformationally flexible natural products: Scope, assessment, and limitations. *J. Nat. Prod.* **2019**, *82*, 2299–2306.

(13) (a) Jiang, X.; Zhang, Q.; Zhu, Y.; Nie, F.; Wu, Z.; Yang, C.; Zhang, L.; Tian, X.; Zhang, C. Isolation, structure elucidation and biosynthesis of benzo[*b*]fluorene nenestatin A from deep-sea derived *Micromonospora echinospora* SCSIO 04089. *Tetrahedron* **2017**, *73*, 3585–3590. (b) Huang, C.; Yang, C.; Zhang, W.; Zhang, L.; De, B. C.; Zhu, Y.; Jiang, X.; Fang, C.; Zhang, Q.; Yuan, C.-S.; Liu, H.-W.; Zhang, C. Molecular basis of dimer formation during the biosynthesis of benzofluorene-containing atypical angucyclines. *Nat. Commun.* **2018**, *9*, 2088.

(14) (a) Herzon, S. B. The mechanism of action of (−)-lomaiviticin A. *Acc. Chem. Res.* **2017**, *50*, 2577–2588. (b) Woo, C. M.; Li, Z.; Paulson, E. K.; Herzon, S. B. Structural basis for DNA cleavage by the potent antiproliferative agent (−)-lomaiviticin A. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2851–2856.

5.6 Experimental Section

This section outlines the experimental procedure and statistics for electron diffraction studies. All other supplementary information are reported in the adapted article.

5.6.1 Recrystallization

(–)-Lomaiviticin C was recrystallized by slow evaporation of a tetrahydrofuran–acetonitrile solution (initial concentration: ~0.1 mg/mL) at –20 °C.

5.6.2 Electron Diffraction Data Collection, Processing, and Refinement

Crystals of (–)-lomaiviticin C were loaded onto a Quantifoil holey-carbon EM grid. Using Gatan 626 cryo-holder, the grid was inserted into FEI Tecnai F200C electron microscope (operating voltage of 200 keV, wavelength of 0.025 Å), and cooled to cryogenic temperature (100 K). Crystals of interest were isolated using a selected area aperture and continuously rotated at a rate of –0.3° s-1 over a tilt range of 50–100°. Diffraction data was recorded on a Ceta-D CMOS 4k x 4k camera in rolling shutter mode (3 s integration per frame, binned by 2) and converted from SER to SMV format using ser2smv software, as previously described.[1] Data from 4 crystals were indexed and integrated in XDS, merged using XSCALE, and converted to SHELX format using XDSCONV.[2] A preliminary structure was obtained using direct methods in SHELXD[3] and refined with SHELXL[4] in ShelXle.[5] All non-hydrogen atoms were refined anisotropically, and hydrogen atoms were placed using the riding model. Crystallographic information file (CIF) for (–)-lomaiviticin C containing atomic coordinates and structure factors have been deposited at the Cambridge Crystallographic Data Center (CCDC 2062671).

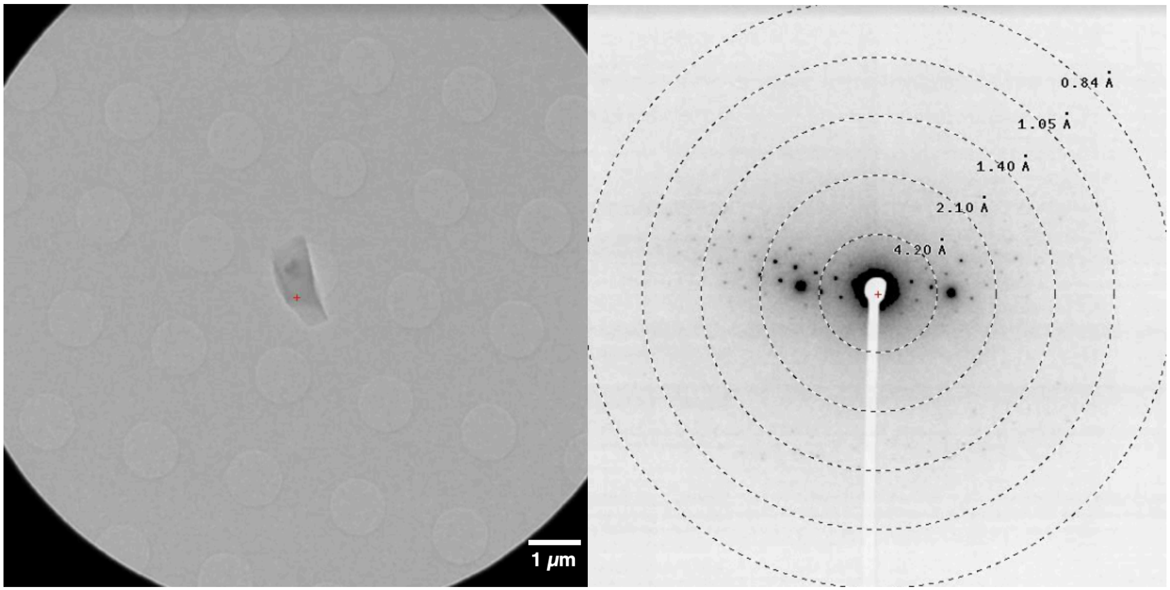


Figure 5.11 (–)-Lomaiviticin C microcrystal and diffraction pattern at SA 2600x magnification. Grid holes are 1 μm in diameter.

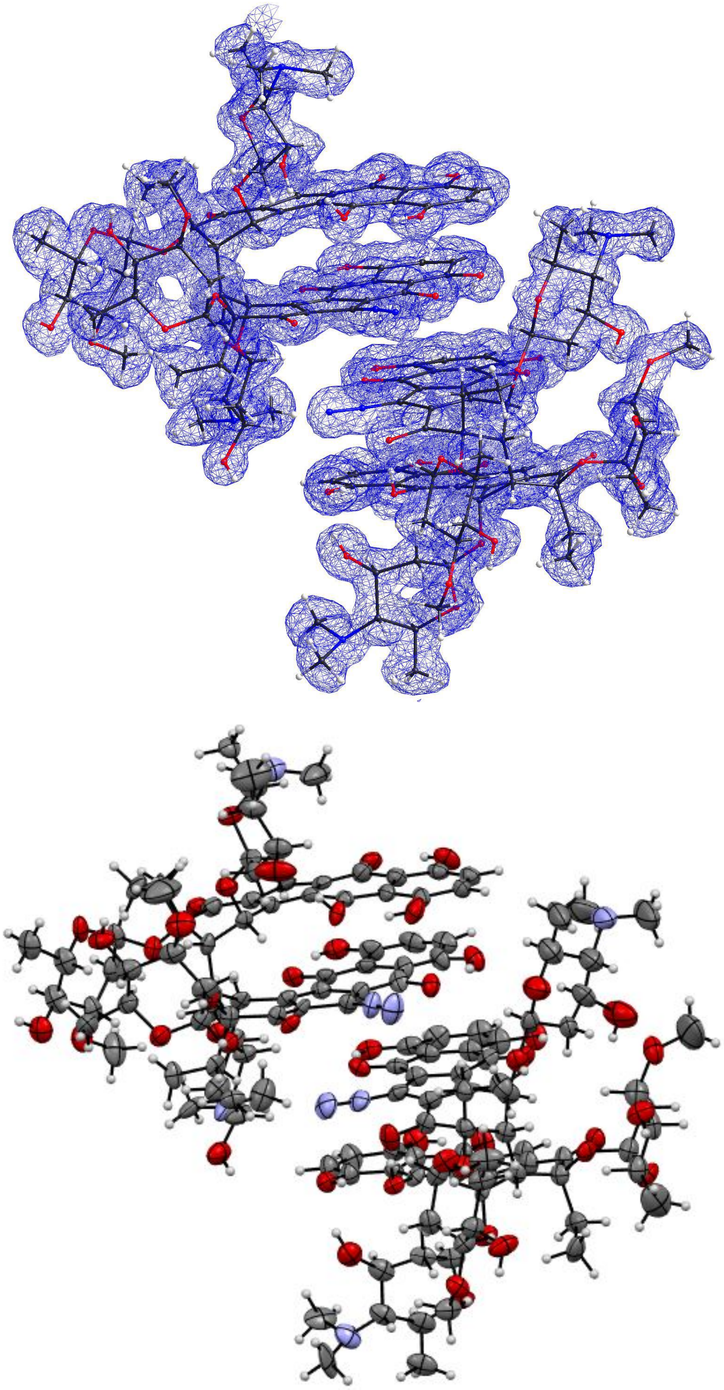


Figure 5.12 Asymmetric unit of (–)-lomaiviticin C (**5.3'**) crystal structure (CCDC 2062671). The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.35 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{64}H_{82}N_4O_{16}$
Formula weight	1339.41

Temperature (K)	100(2)
Space group	$P\ 2_1$
Cell dimensions	
a, b, c (Å)	19.46(2), 14.32(3), 22.94(2)
α, β, γ (°)	90, 112.47(6), 90
Resolution (Å)	1.05(1.10–1.05)
Observed reflections	38577(5242)
Unique reflections	5632(723)
Completeness (%)	99.5(99.9)
R_{obs} (%)	19.8(72.2)
R_{meas} (%)	21.5(77.7)
$I/\sigma I$	6.49(2.85)
$CC_{1/2}$ (%)	99.2(79.9)
R_1 (%)	11.86
wR_2 (%)	29.27
GooF	1.333

Table 5.1 Data collection and refinement statistics for **5.3'**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4763	Data was truncated to a resolution of 1.05 Ångstrom due to poor diffraction data quality above this threshold.

PLAT089_ALERT_3_A Poor Data / Parameter Ratio (Zmax < 18) .... 3.21 Note	The structure has low symmetry and a very large number of non-hydrogen atoms to refine (192 atoms).
PLAT312_ALERT_2_A Strange C-O-H Geometry (C-O < 1.25 Ang) ..... O30# Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller).
PLAT312_ALERT_2_A Strange C-O-H Geometry (C-O < 1.25 Ang) ..... O40# Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller).
PLAT312_ALERT_2_A Strange C-O-H Geometry (C-O < 1.25 Ang) ..... O42# Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller).
PLAT410_ALERT_2_A Short Intra H...H Contact H1C_1 ..H12B_1 . 1.78 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.
PLAT410_ALERT_2_A Short Intra H...H Contact H17_2 ..H9D2_2 . 1.78 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.
PLAT412_ALERT_2_A Short Intra XH3 .. XHn H8Z1_1 ..H6D3_1 . 1.54 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition,

	diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.
PLAT413_ALERT_2_A Short Inter XH3 .. XHn H22_1 ..H15A_2 . 1.70 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.
PLAT413_ALERT_2_A Short Inter XH3 .. XHn H22_1 ..H15A_2 . 1.70 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.
PLAT413_ALERT_2_A Short Inter XH3 .. XHn H22_1 ..H15A_2 . 1.70 Ang. x,y,z = 1_555 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, diffraction data from microED analysis results in unit cell dimensions and bond lengths that deviate from the real values (often smaller). These hydrogen atoms were also placed using the riding model.

Table 5.2 Justification for alerts in (–)-lomaiviticin C (**5.3’**) crystal structure (CCDC 2062671).

5.6.3 Supplementary Notes and References

(1) Hattne, J.; Reyes, F. E.; Nannenga, B. L.; Shi, D.; de la Cruz, M. J.; Leslie, A. G.; Gonen, T. MicroED data collection and processing. *Acta Cryst.* **2015**, *A71*, 353–360.

(2) Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.

(3) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(4) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst*. **2015**, *C71*, 3–8.

(5) Hübschle, C. B.; Sheldrick, G. M.; Dittrich, B. Shelxle: A qt graphical user interface for SHELXL. *J. Appl. Crystallogr*. **2011**, *44*, 1281–1284.

CHAPTER SIX

Mining for Natural Products in Complex Biological Extracts

6.1 Contributions

This chapter is an unpublished collaborative work. Plant extract fractions were provided by Dr. Gina Porras in the Quave lab at Emory University. Fungal culture growth and time-fractionation using liquid chromatography were performed and optimized by Dr. Bruno Perlatti in the Tang lab at University of California, Los Angeles. Electron diffraction and microarrayer experiments were conducted and optimized by Lee Joon Kim in the Nelson lab at University of California, Los Angeles and California Institute of Technology. Crystal structure determination and analyses were performed by Lee Joon Kim.

6.2 Abstract

Discoveries of novel natural products (NPs) with complex molecular scaffolds and potent bioactivity fuel various fields of chemistry, biology and medicine. However, purification and identification of NPs remain time- and resource-intensive, and are faced with difficulties due to limitations in existing characterization methods. To overcome these challenges, microcrystal electron diffraction (MicroED) is employed as a key method in simplifying NP purification and identification workflow. This study demonstrates that MicroED is capable of determining NP structures from partially purified NP extract sample consisting of ten or more different compounds. Stemming from these initial experiments, a new workflow combining liquid chromatography with MicroED is designed and applied to a fungal NP extract to solve structures after a single

fractionation into a 96-well plate. This platform is further developed into a high-throughput approach, where thousands of fractionated NP extract samples can be screened to quickly identify microcrystalline NPs to be characterized. This high-throughput MicroED platform opens up numerous possibilities in the realm of natural product discovery by enabling efficient structure determination of metabolites that were previously undetected or unable to be characterized, greatly accelerating the rate of novel NP and NP-inspired drug discovery.

6.3 Introduction

Various fields of research including untargeted metabolomics, natural product isolation chemistry, and synthetic biology focus their efforts on identifying or isolating primary and/or secondary metabolites, leading to discoveries of novel natural products (NPs). These areas contribute to the broad scientific knowledge base in many ways, such as connecting genetic information with phenotypes to uncover biomarkers for diseases,[1] inspiring biochemical mechanism of action studies,[2] and challenging synthetic organic chemists to synthesize novel scaffolds by developing new methodologies.[3] The high complexity and molecular diversity of NPs with potent bioactivity have long inspired chemists in drug discovery and design: For the last four decades, NPs, NP derivatives, and NP-inspired compounds have comprised the majority of all approved small molecule drugs.[4] Thus, advancing the rate at which novel NPs are discovered will also accelerate the rate of discovering drugs with novel scaffolds or mechanisms of action, both of which are important design elements for combatting drug resistance of various pathogens.

However, a major bottleneck exists in the characterization and discovery of natural products, particularly in the case of novel compounds, due to the time and resources involved in purifying NP extracts until the target analyte is sufficiently pure for structural analysis using NMR spectroscopy or X-ray crystallography. A natural extract sample can contain hundreds of thousands to millions of compounds,[5] making this endeavor akin to the phrase "searching for the needle in a haystack."[6] To discover natural products in these unbelievably complex environments, the extract sample must undergo multiple rounds of purification using liquid chromatography, producing fractions that are each purer than the last. Even after purification, however, major limitations in existing analytical methods hinder facile structure determination for NP identification, as discussed in Chapter 1.[7]

Microcrystal electron diffraction, or MicroED, is an emerging cryoelectron microscopy technique that complements existing analytical methods and help overcome these limitations. MicroED has been utilized to solve crystal structures of small molecules from various sources, including commercial chemical bottles and biological specimen.[8] This technique is capable of probing structural information from microcrystals generated from miniscule amounts of material,[8q,8r,8t] as well as analyzing heterogeneous mixtures of small molecules.[8a,8b,8r,8t] Taking advantage of this method can extend our structure determination capabilities to a vast realm to include natural products that 1) are isolated in amounts insufficient for NMR analysis or X-ray crystallography, 2) are produced in such infinitesimal amounts that they evade detection entirely during liquid chromatography, 3) are in mixtures of compounds difficult to separate, 4) are difficult to crystallize due to inherent molecular flexibility, 5) are challenging to characterize using NMR spectroscopy due to low proton-to-carbon ratio, and more.

In this study, this powerful technique was harnessed to envision a high-throughput approach to identify NPs from complex natural product extracts. In our initial studies, MicroED was demonstrated to be capable of simplifying NP identification workflow in the most simplistic sense, by elucidating structures from partially purified samples. This preliminary experiment was further developed by designing a NP identification workflow, coupling high-performance liquid chromatography (HPLC) with MicroED. This workflow was improved using a high-throughput approach, where thousands of samples from a natural product extract could be screened using MicroED for efficient screening and crystallographic structure determination.

6.4 NP Identification from Partially Purified Samples Using MicroED

While MicroED has been demonstrated to be capable of studying heterogeneous mixtures, reported examples in literature have been artificial,[8a,8b] where two or more diffracting microcrystals were manually mixed together during formulation, or composed of only two compounds.[8r,8t] Solving structures from more complex mixtures of samples from natural sources would increase MicroED's relevance to NP identification. To demonstrate that MicroED can be leveraged to identify NPs from partially purified natural samples, a highly complex extract derived from the fruits of *Schinus terebinthifolia* was studied. *S. terebinthifolia* is a native plant to South and Central America that was introduced to the United States over 100 years ago.[9] While this plant is regarded as an invasive species in the southern United States,[10] it is valued as medicine in Brazil for its antimicrobial, anti-inflammatory, antioxidant and anti-tumor activities.[11] Previously, Quave and coworkers reported the isolation and identification of three NPs from the sample collected in Florida (Figure 6.1) and demonstrated these compounds' anti-virulence effects against *Staphylococcus aureus*, whose antibiotic resistance is jeopardizing public health in community and healthcare settings.[12] Compounds **6.1**, **6.2** and **6.3** were identified only after multiple rounds of separation by liquid chromatography followed by a combination of techniques such as NMR spectroscopy, high-resolution mass spectometry, and X-ray crystallography. Successful characterization of these triterpenoid acids *via* MicroED without the need to arrive at a completely pure, isolated sample would demonstrate the possibility of alleviating the substantial time and labor required to discover NPs.

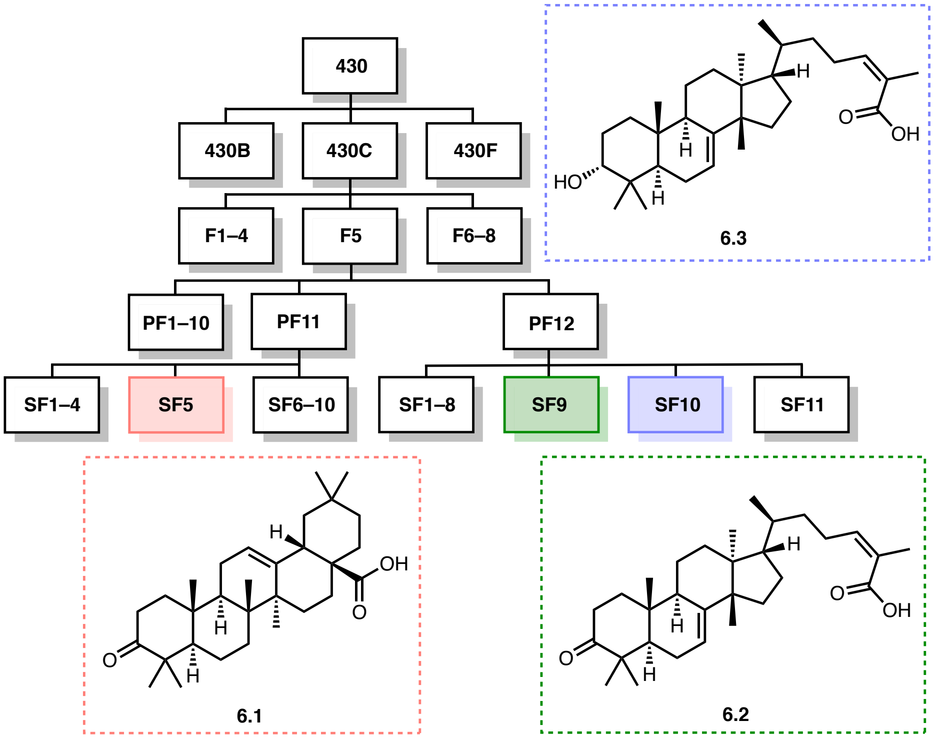


Figure 6.1 Liquid chromatography purification schematic for *S. terebinthifolia* extract. The three bioactive triterpenoid acids 3-oxo-olean-12-en-28-oic acid (**6.1**), 3-oxotirucalla-7,24*Z*-dien-26-oic acid (**6.2**), and 3α-hydroxytirucalla-7,24*Z*-dien-26-oic acid (**6.3**) were found in fractions 430F-F5-PF11 and 430F-F5-PF12.

The fractionated samples containing *S. terebinthifolia* extract at various purification stages were screened starting from the bottom row (430F-F5-PF11-SF1–10 and 430F-F5-PF12-SF1–11, Figure 6.1), which contains the most pure fractions. Analyses of 430F-F5-PF11-SF5, 430F-F5-PF12-SF9, and 430F-F5-PF12-SF10 yielded diffraction patterns that were characteristic of small molecules (See SI, Figures 6.12–6.14). Notably, the diffraction patterns from fraction 430F-F5-PF12-SF9 was processed to yield unit cell parameters that were consistent with those of the published X-ray crystal structure of **6.2** obtained from this fraction (Figure 6.2, top).[13] MicroED data sets collected during the screening process was sufficiently high in resolution and completeness to yield a crystal structure that also matched the X-ray structure (Figure 6.2, bottom).

	Space Group	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
MicroED	P 2_1	6.948(5)	19.988(7)	19.789(7)	90	94.70(9)	90
X-ray	P 2_1	6.81969(10)	19.7846(3)	19.4400(3)	90	94.4735(13)	90
Absolute Error	N/A	0.12831	0.2034	0.3490	N/A	0.2265	N/A

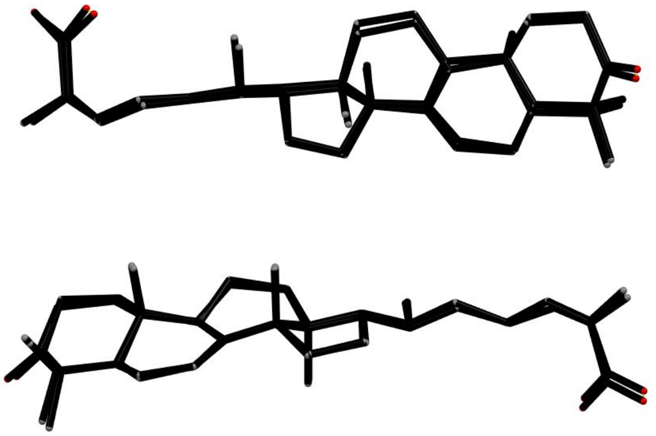


Figure 6.2 Comparison of MicroED and X-ray unit cell parameters (top) and structures (bottom, RMS = 0.279) for **6.2**. Estimated standard deviations of cell parameters from each data set is shown in parenthesis.

Encouraged by these results, the fractions 430F-F5-PF11 and 430F-F5-PF12 were screened for diffraction. These samples have undergone one fewer purification step and may contain tens of compounds in each. From fraction 430F-F5-PF12, which contains triterpenoid **6.2** along with a minimum of ten other major peaks (SF1–11), diffraction pattern consistent with that of **6.2** was obtained in sufficient quality to afford the crystal structure of the compound (See SI, Figure 6.16). The successful identification of triterpenoid acid **6.2** in this NP mixture highlights the power of MicroED in obtaining structural information from an impure natural product sample, consisting of the highest number of components reported to date.

In an attempt to push the limit of MicroED's ability to detect NPs in a complex mixture, fraction 430F-F5 was screened. Diffraction patterns were observed from this highly complex sample; however, they did not match the diffraction pattern of **6.2** (See SI, Figure 6.17). The lack of matching diffraction pattern even after multiple screening sessions is presumably due to the lack of suitable microcrystals generated under the screening conditions, indicating a need for recrystallization. Thus, recrystallization screens were attempted using a number of solvents, but diffraction patterns continued to be elusive, suggesting that a more extensive screening of conditions is necessary, or perhaps the mixture was exceedingly complex even for microcrystal generation. While this study demonstrated that MicroED can be successfully used to mine for crystalline NPs from an extract without the need to fully isolate the compounds, it also highlighted the need for a better screening workflow for NP identification.

6.5 NP Identification After Single Fractionation by Liquid Chromatography

Having successfully demonstrated that MicroED can substantially simplify purification procedures for NPs before elucidating their structures, a streamline workflow for structural elucidation of NPs was envisioned by combining liquid chromatography and MicroED (Figure 6.3). Instead of undergoing multiple rounds of purification, complex natural product extracts were fractionated into 96-well plates. These fractionated solutions were then loaded on transmission electron microscopy (TEM) grids to be screened for diffraction. As previously established by our study with *S. terebinthifolia* and prior literature,[8a,8r,8t] the solutions in the wells do not need to be completely pure for the generation of microcrystals amenable for MicroED. Fractionating into 96-well plates would divide up the number of NPs in each well to prevent too many compounds from being present. Owing to high sensitivity of MicroED, this workflow will enable the detection of crystalline NPs that are produced in amounts insufficient for NMR or X-ray crystallographic analyses, and reduce the number of purification steps to a single fractionation strategy.

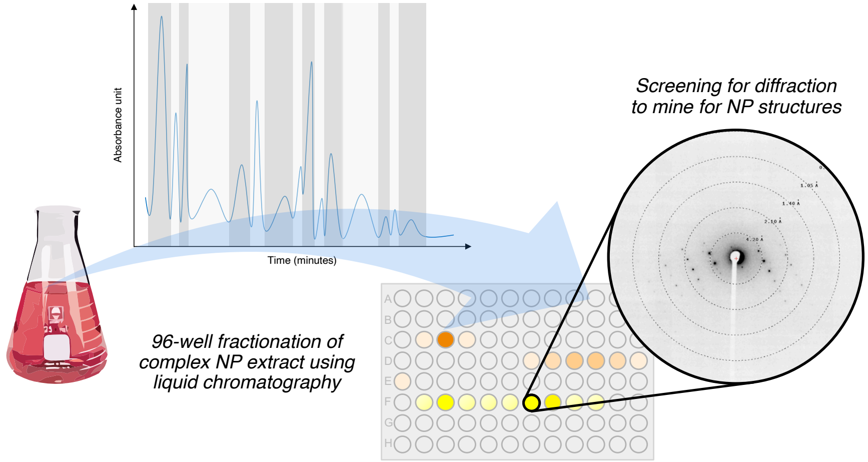


Figure 6.3 Representative workflow for structural elucidation of NPs by combining liquid chromatography and MicroED.

As a proof-of-principle, *Trichoderma afroharzianum* t-22 was chosen as a model system. *Trichoderma afroharzianum* t-22, previously classified as *Trichoderma harzianum* t-22, is a biocontrol fungus characterized by strong competitiveness and mycoparasitic activity against other fungi such as *Rhizoctonia*, *Fusarium*, *Phytophthora* species.[14] While a number of metabolites produced by *T. afroharzianum* t-22 have been previously identified,[15] the currently known compounds consist of a minor fraction in the possible pool of NPs.[16] By choosing this strain as a model system, a mixture of both known and unknown NPs could be elucidated through the liquid chromatography-MicroED workflow.

The complex NP extract was subjected to a time-based fractionation into a 96-well plate, using a collection increment of 15 seconds. This process yielded a 96-deep well plate filled with colored and colorless solutions, and the 2 μL droplets of the well solutions were deposited onto a Quantifoil TEM grid for on-grid microcrystallization before inserting into the 200 keV electron microscope Talos F200C. Under the electron beam, microcrystals with various morphologies were observed (See SI, Figure 6.18). Notably, high resolution diffraction patterns were obtained from wells H3–H5 to yield crystal structures of two anthraquinone-derivative NPs, chrysophanol (**6.4**) and pachybasin (**6.5**) (Figure 6.4). The crystal structure of chrysophanol, obtained from well H3, is a new P-orthorhombic polymorph of previously reported X-ray structure of chrysophanol,[17] although diffraction pattern consistent with the reported unit cell parameters of this X-ray structure was also observed during data collection (See SI, Figure 6.20). Interestingly, the second crystal structure solved from well H3 revealed that **6.4** and **6.5** were co-crystallized, whereas the diffraction patterns collected from wells H4 and H5 yielded crystal structures of only **6.5**, due to the earlier elution of **6.4**. Additionally, the MicroED structure of **6.5** represents the first solid-state structure of pachybasin. Chrysophanol and pachybasin are known NPs that have been isolated from *T. harzianum* before;[15d] however, unlike previous literature that utilized two silica column chromatography purification steps along with a combination of characterization methods such as various 1D- and 2D NMR experiments, the structures of these compounds were elucidated after a single round of fractionation and a single screening session on the TEM. These results demonstrate a powerful approach to mine for NPs that would reduce the amount of time and number of resources required for purification and isolation of compounds prior to structural identification.

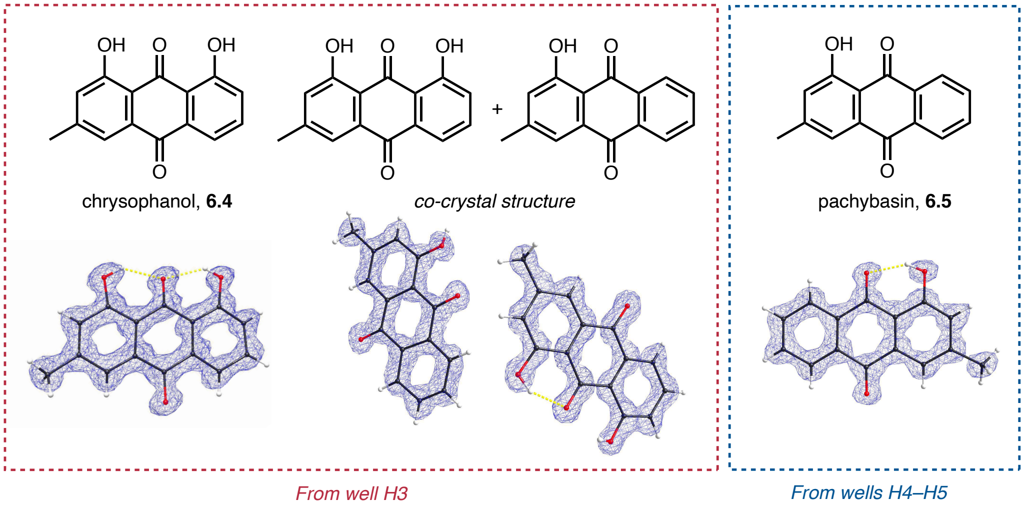


Figure 6.4 Crystal structures of chrysophanol (**6.4**) and pachybasin (**6.5**) obtained using the proposed workflow.

6.6 High-Throughput Screening Platform for NP Discovery

While fractionating complex extracts into 96-well plates reduced sample purification time significantly, it created a major bottleneck at the screening stage. Talos F200C is a side-entry electron microscope that allows for a single grid to be analyzed at a time using the standard single-tilt sample holder. This limitation, combined with the sheer number of wells to be analyzed, led to an arduous screening process for each extract. To increase the throughput of our experiments, two instruments were employed: Scienion sciFLEXARRAYER S3 and Talos Arctica. Scienion sciFLEXARRAYER S3 (SCIENION AG, Berlin, Germany) is a piezoelectric liquid handling robot that enables automated non-contact dispensing of solutions with volumes in the picoliter to microliter range.[18] Using the piezoelectric dispenser, all 96 solutions in a well plate could be microarrayed onto a single grid, enabling the screening of 96 well solutions at a time and greatly increasing the throughput of our MicroED experiments (Figure 6.5). Additionally, Talos Arctica equipped with an autoloader allows for a simultaneous loading of up to 12 grids and minimizes the time it takes to switch to a new grid. Combined together, the new workflow enables the screening of thousands of different solutions while minimizing wait time in between switching out the grids.

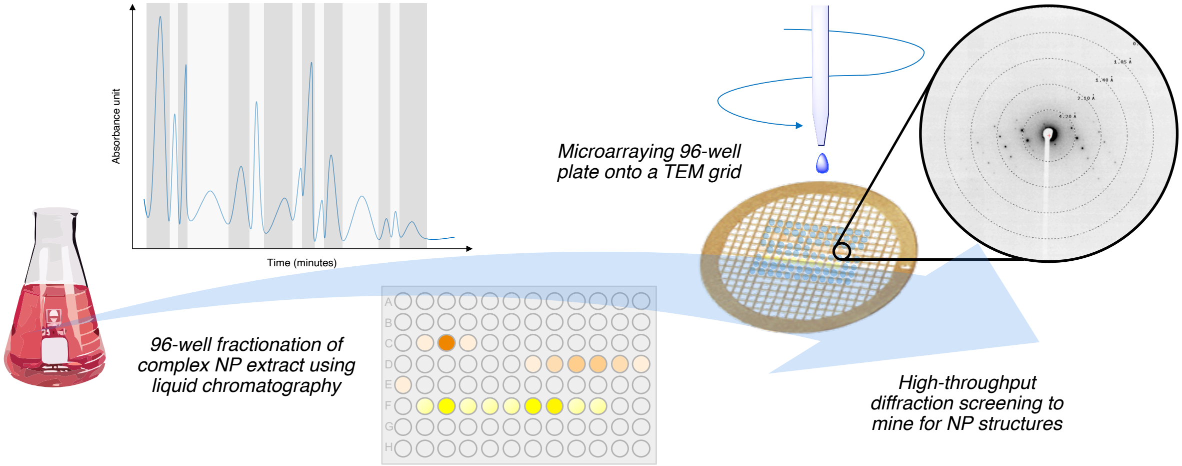


Figure 6.5 Representative workflow of the high-throughput screening platform for NP discovery.

To demonstrate the feasibility of this high-throughput workflow, the same 96-well plate containing time-fractionated *T. afroharzianum* t-22 extract was loaded onto the sciFLEXARRAYER S3 system. Using the S3 system, ~250 picoliter-sized droplets were deposited onto the grid from each well in a fully automated, 2 hour-long process. The microarrayed grid was clipped at room temperature and inserted into the Talos Arctica, followed by acquisition of a full grid montage using SerialEM[19] (version 3.8.18) for facile orientation of the samples on the grid. Using the grid's center feature and our labeling solution containing either carbamazepine or sodium chloride, the orientation of the grid could be easily distinguished, and microcrystals from individual grid squares could be traced to a well (Figure 6.6).

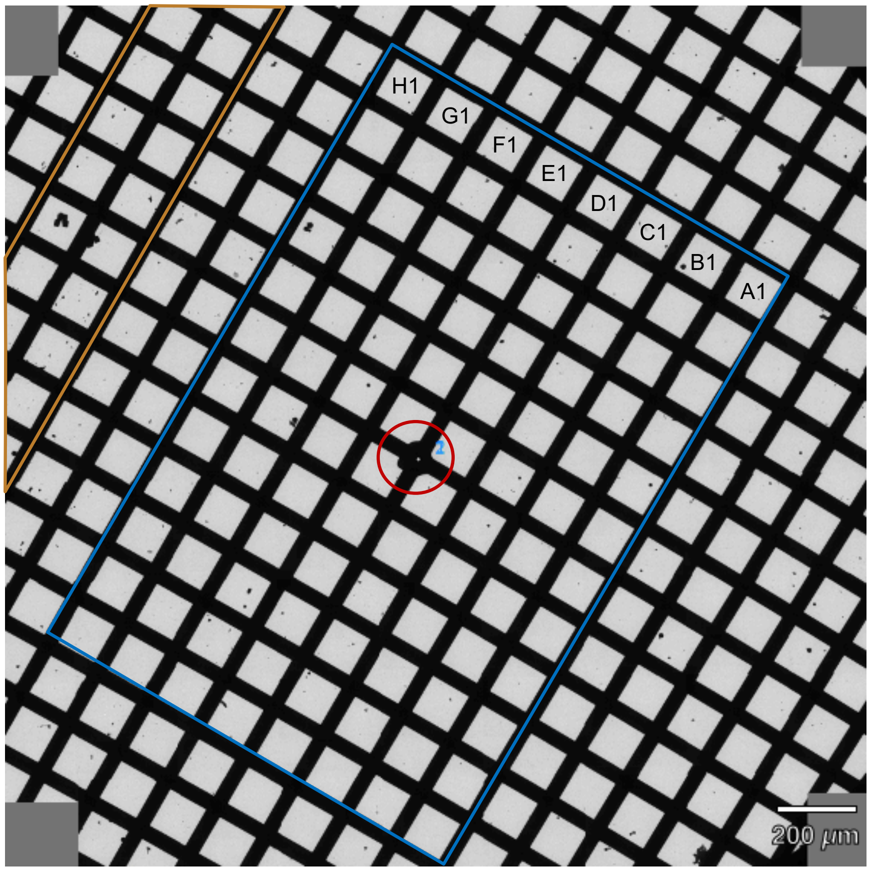


Figure 6.6 Microarrayed grid alignment using a labeling solution (orange, sodium chloride shown) or the grid's center feature (red). The blue rectangle indicates the expected location of the 96 well plate samples.

Once oriented, the microarrayed grids were manually screened, assisted by the full grid montage. To our delight, the grid squares where droplets from wells H3–H5 were deposited contained small crystalline needles (Figure 6.7). Compared to the results from the low-throughput method described in section 6.5, these needles were much smaller in size (~1–5 microns in length) and diffracted poorly, presumably due to the miniscule amount of material that was used to generate microcrystals from (~ 250 pL droplet per well as opposed to 2 μL in the case of the low-throughput method. See SI, Figures 6.19 and 6.20–6.25 for comparison). Nevertheless, the diffraction patterns were identifiable as those from small molecules rather than impurities such as ice or small inorganic salts based on approximate unit cell dimensions, demonstrating that diffraction screening of microarrayed grids can quickly identify grid squares, and therefore wells, containing microcrystals of potential interest.

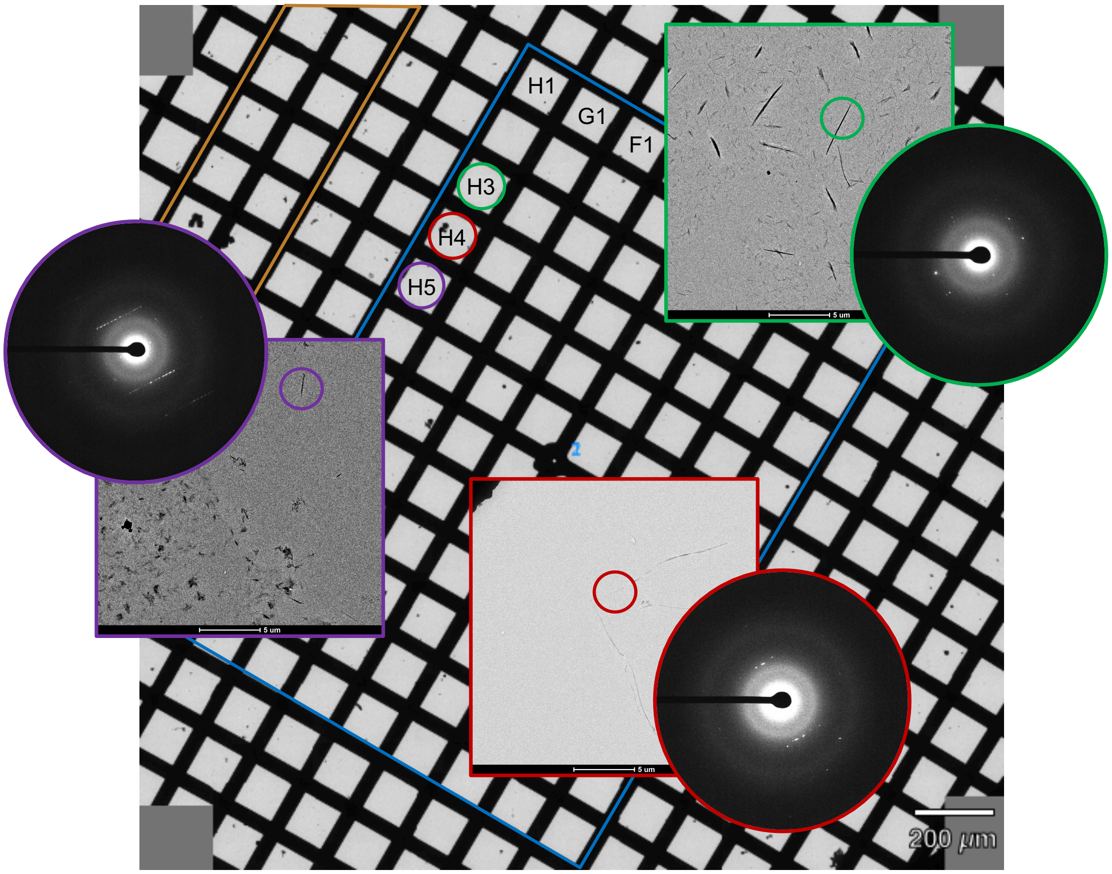


Figure 6.7 Diffraction patterns obtained from microarrayed grids. Microcrystals (square) and diffraction patterns (circle) obtained from wells H3 (green), H4 (red), and H5 (purple) are shown. Blue: expected location of the 96 well plate samples. Orange: sodium chloride labeling solutions. Diffraction patterns are zoomed in for visibility.

6.7 Conclusion

In conclusion, the studies described in this chapter highlight different approaches in simplifying workflow in NP purification and subsequent identification. In the most simplistic way, MicroED is used to reduce the number of purification steps by demonstrating that structure determination of a natural product can occur from partially purified samples consisting of ten or more different compounds. This initial study led to the design of a new workflow combining liquid chromatography with MicroED, enabling structure determination after a single fractionation into a 96-well plate. However, the resulting bottleneck in screening necessitated a high-throughput workflow. The improved design, employing an autoloader capable of storing 12 grids and a piezoelectric dispenser that microarrays an entire 96-well plate onto a single TEM grid, opens up opportunities to screen thousands of samples in a much more efficient manner. Future investigations involve utilizing this platform to screen various natural product extracts, mining for diffraction and crystalline microcrystals to solve structures of. Moreover, as efforts to automate electron diffraction experiments continue,[20] incorporating automation into this workflow will maximize the number of microcrystalline natural products to be characterized, increasing the chances of novel NP discovery and greatly accelerating NP and drug discovery campaigns.

6.8 Notes and References

(1) (a) Cox, D. G.; Oh, J.; Keasling, A.; Colson, K. L.; Hamann, M. T. The utility of metabolomics in natural product and biomarker characterization. *Biochim. Biophys. Acta Gen.* Identification of biomarkers to diagnose diseases and find adverse drug reactions by metabolomics. *Drug Metab. Pharmacokinet.* **2021**, *37*, 100373. (c) Geng, C.; Cui, C.; Guo, Y.; Wang, C.; Zhang, J.; Han, W.; Jin, F.; Chen, D.; Jiang, P. Metabolomic profiling reveals biomarkers in patients with moyamoya disease. *Front. Neurosci.* **2020**, *14*, 308.

(2) (a) Baptista, R.; Fazakerley, D. M.; Beckmann, M.; Baillie, L.; Mur, L. A. J. Untargeted metabolomics reveals a new mode of action of pretomanid (PA-824). *Sci. Rep.* **2018**, *8*, 5084. (b) Wang, H.; Liu, A.; Zhao, W.; Zhao, H.; Gong, L.; Chen, E.; Cui, N.; Ji, X.; Wang, S.; Jiang, H. Metabolomics research reveals the mechanism of action of astragalus polysaccharide in rats with digestive system disorders. *Molecules* **2018**, *23*, 3333.

(3) (a) Nicolaou, K. C.; Snyder, S. A. The essence of total synthesis. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11929–11936. (b) Mohr, J. T.; Krout, M. R.; Stoltz, B. M. Natural products as inspiration for the development of asymmetric catalysis. *Nature* **2008**, *455*, 323–332.

(4) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803.

(5) (a) Dixon, R. A.; Strack, D. Phytochemistry meets genome analysis, and beyond. *Phytochemistry*, **2003**, *62*, 815–816. (b) Rai, A.; Saito, K.; Yamazaki, M. Integrated omics analysis of specialized metabolism in medicinal plants. *Plant J. Cell Mol. Biol.* **2017**, *90*, 764–787. (c) Afendi, F. M.; Okada, T.; Yamazaki, M.; Hirai-Morita, A.; Nakamura, Y.; Nakamura, K.; Ikeda, S.; Takahashi, H.; Altaf-Ul-Amin, M.; Darusman, L. K,; Saito, K.; Kanaya, S. *Plant Cell Physiol.* **2012**, *53*, e1.

(6) Chevrette, M. G.; Handelsman, J. Needles in haystacks: reevaluating old paradigms for the discovery of bacterial secondary metabolites. *Nat. Prod. Rep.* **2021**, *38*, 2083.

(7) See Chapter 1.2 for a detailed discussion.

(8) (a) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592. (b) Gruene, T.; Wennmacher, J. T. C.; Zaubitzer, C.; Holstein, J. J.; Heidler, J.; Fecteau-Lefebvre, A.; De Carlo, S.; Müller, E.; Goldie, K. N.; Regeni, I.; Li, T.; Santiso-Quinones, G.; Steinfeld, G.; Handschin, S.; van Genderen, E.; van Bokhoven, J. A.; Clever, G. H.; Pantelic, R. Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed.* **2018**, *57*, 16313–16317. (c) Jones, C. G.; Asay, M.; Kim, L. J.; Kleinsasser, J. F.; Saha, A.; Fulton, T. J.; Berkley, K. R.; Cascio, D.; Malyutin, A. G.; Conley, M. P.; Stoltz, B. M.;

Lavallo, V.; Rodríguez, J. A.; Nelson, H. M. Characterization of reactive organometallic species via MicroED. *ACS Cent. Sci.* **2019**, *5*, 1507–1513. (d) Banihashemi, F.; Bu, G.; Thaker, A.; Williams, D.; Lin, J. Y. S.; Nannenga, B. L. Beam-sensitive metal-organic framework structure determination by microcrystal electron diffraction. *Ultramicroscopy* **2020**, *216*, 113048. (e) Aykanat, A.; Jones, C. G.; Cline, E.; Stolz, R. M.; Meng, Z.; Nelson, H. M.; Mirica, K. A. Conductive stimuli-responsive coordination network linked with bismuth for chemiresistive gas sensing. *ACS Appl. Mater. Interfaces* **2021**, *13*, 60306–60318. (f) Meng, Z.; Jones, C. G.; Farid, S.; Khan, I. U.; Nelson, H. M.; Mirica, K. A. Unraveling the electrical and magnetic properties of layered conductive metal-organic framework with atomic precision. *Angew. Chem. Int. Ed.* **2021**, *60*, 2–11. (g) Jellen, M. J.; Liepuoniute, I.; Jin, M.; Jones, C. G.; Yang, S.; Jiang, X.; Nelson, H. M.; Houk, K. N.; Garcia-Garibay, M. A. Enhanced gearing fidelity achieved through macrocyclization of a solvated molecular spur gear. *J. Am. Chem. Soc.* **2021**, *143*, 7740–7747. (h) Samkian, A.; Kiel, G. R.; Jones, C. G.; Bergman, H. Oktawiec, J.; Nelson, H. M.; Tilley, T. D. Elucidation of diverse solid-state packing in a family of electron-deficient expanded helicenes via microcrystal electron diffraction (MicroED). *Angew. Chem. Int. Ed.* **2020**, *5*, 2493–2499. (i) Burch, J. E.; Smith, A. G.; Caille, S.; Walker, S. D.; Wurz, R.; Cee, V. J.; Rodriguez, J.; Gostovic, D.; Quasdorf, K.; Nelson, H. M. Putting MicroED to the test: An account of the evaluation of 30 diverse pharmaceutical compounds. *ChemRxiv* **2021**. This content is a preprint and has not been peer-reviewed. DOI:10.26343/chemrxiv-2021-h3tqz. (j) Bruhn, J. F.; Scapin, G.; Cheng, A.; Mercado, B. Q.; Waterman, D. G.; Ganesh, T.; Dallakyan, S.; Read, B. N.; Nieusma, T.; Lucier, K. W.; Mayer, M. L.; Chiang, N. J.; Poweleit, N.; McGilvray, P. T.; Wilson, T. S.; Mashore, M.; Hennessy, C.; Thomson, S.; Wang, B.; Potter, C. S.; Carragher, B. Small molecule microcrystal electron diffraction for the pharmaceutical industry-lessons learned from examining over fifty samples. *Front. Mol. Biosci.* **2021**, *8*, 648603. (k) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; van der Donk, W. A. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **2019**, *365*, 280–284. (l) Dick, M.; Sarai, N. S.; Martynowycz, M. W.; Gonen, T.; Arnold, F. H. Tailoring tryptophan synthase TrpB for selective quaternary carbon bond formation. *J. Am. Chem. Soc.* **2019**, *141*, 19817–19822. (m) Gleason, P. R.; Nannenga, B. L.; Mills, J. H. Rapid structural analysis of a synthetic non-canonical amino acid by microcrystal electron diffraction. *Front. Mol. Biosci.* **2021**, *7*, 609999. (n) Nelson, H. M.; Siu, J. C.; Saha, A.; Cascio, D.; MacMillan, S. N.; Wu, S.-B.; Lu, C.; Rodriguez, J. A.; Houk, K. N.; Lin, S. Isolation and X-ray crystal structure of an electrogenerated TEMPO-N$_3$ charge-transfer complex. *Org. Lett.* **2021**, *23*, 454–458. (o) Zee, C.-T.; Glynn, C.; Gallagher-Jones, M.; Miao, J.; Santiago, C. G.; Cascio, D.; Gonen, T.; Sawaya, M. R.; Rodriguez, J. A. Homochiral and racemic MicroED structures of a peptide from the ice nucleation protein InaZ. *IUCrJ* **2019**, *6*, 197–205. (p) Curtis, B. J.; Kim, L. J.; Wrobel, C. J. J.; Eagen, J. M.; Smith, R. A.; Burch, J. E.; Le, H. H.; Artyukhin, A. B.; Nelson, H. M.; Schroeder, F. C. Identification of uric acid gluconucleoside-ascaroside conjugates in *Caenorhabditis elegans* by combining synthesis and MicroED. *Org. Lett.* **2020**, *22*, 6724–6728. (q) Kim, L. J.; Xue, M.; Li, X.; Xu, Z.l Paulson, E.; Mercado, B. Q.; Nelson, H. M.; Herzon, S. Structural revision of the lomaiviticins. *J. Am. Chem. Soc.* **2021**, *143*, 6578–6585. (r) Kim, L. J.; Ohashi, M.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J.; Tang, Y.; Nelson, H. M. Prospecting for natural products by genome mining and microcrystal electron diffraction. *Nat. Chem. Biol.* **2021**, *17*, 872–877. (s) Go, E. B.; Kim, L. J.; Nelson, H. M.; Ohashi, M.; Tang, Y. Biosynthesis of the Fusarium mycotoxin (–)-sambutoxin. *Org. Lett.* **2021**, *23*, 7819–7823. (t) Park, J.-D.; Li, Y.; Moon, K.; Han, E. J.; Lee, S. R.; Seyedsayamdost, M. R. Structural elucidation of cryptic algaecides in marine algal-bacterial symbioses by NMR spectroscopy and MicroED. *Angew. Chem. Int. Ed.* **2022**, *61*, e202114022.

(9) Williams, D. A.; Muchugu, E.; Overholt, W. A.; Cuda, J. P. Colonization patterns of the invasive Brazilian peppertree, *Schinus terebinthifolius*, in Florida. *Heredity* **2007**, *98*, 284–293.

(10) Pernas, T.; Giardina, D.; Duquesnel, J.; Franck, A.; Hammer, R. L.; Kunzer, J.; Lange, J.; Langeland, K.; Lieurance, D.; Lockhart, C.; McCollom, J.; Nelson, G.; Possley, J.; Sadle, J. L.; Sowell, D.; Spencer, J.; Stilles, A.; Wunderlin, R. P. Florida exotic pest plant council's 2019 list of invasive plant species, 2019. Florida Exotic Pest Plant Council. https://floridainvasivespecies.org/plantlist2019.cfm (accessed May 21, 2022).

(11) (a) da Silva Dannenberg, G.; Funck, G. D.; Mattei, F. J.; da Silva, W. P.; Fiorentini, Â. M. Antimicrobial and antioxidant activity of essential oil from pink pepper tree (*Schinus terebinthifolius* Raddi) *in vitro* and in cheese experimentally contaminated with *Listeria monocytogenes. Innov. Food Sci. Emerg. Technol.* **2016**, *36*, 120–127. (b) Silva, A. B.; Silva, T.; Franco, E. S.; Rabelo, S. A.; Lima, E. R.; Mota, R. A.; da Câmara, C. A. G.; Pontes-Filho, N. T.; Lima-Filho, J. V. Antibacterial activity, chemical composition, and cytotoxicity of leaf's essential oil from Brazilian pepper tree (*Schinus terebinthifolius*, Raddi). *Braz. J. Microbiol.* **2010**, *41*, 158–163. (c) da Silva, M. M.; Iriguchi, E. K. K.; Kassuya, C. A. L.; Vieira M. C.; Foglio, M. A.; Carvalho, J. E.; Ruiz, A. L. T. G.; Souza, K. P.; Formagio, A. S. N. *Schinus terebinthifolius*: phenolic constituents and *in vitro* antioxidant, antiproliferative and *in vivo* anti-inflammatory activities. *Rev. Bras. Farmacogn.* **2017**, *27*, 445–452. (d) Jeribi, C.; Karoui, I. J.; Hassine, D. B.; Abderrabba, M. Comparative study of bioactive compounds and antioxidant activity of *Schinus terebinthifolius* Raddi fruits and leaves essential oils. *Int. J. Sci. Res.* **2014**, *3*, 453–458. (e) Ramos, D. M. B.; Araújo, M. T. M. F.; Araújo, T. C. L.; Neto, O. G. S.; Silva, M.

G.; Silva, Y. A.; Torres, D. J. L.; Patiota, L. L. S.; Melo, C. M. L.; Lorena, V. M. B.; Paiva, P. M. G.; Mendes, R. L.; Napoleão, T. H. Evaluation of antitumor activity and toxicity of *Schinus terebinthifolia* leaf extract and lectin (SteLL) in sarcoma 180-bearing mice. *J. Ethnopharmacol.* **2019**, *233*, 148–157.

(12) (a) Wilcox, M.; Al-Obeid, S.; Gales, A.; Kozlov, R.; Martínez-Orozco, Rossi, F.; Sidorenko, S.; Blondeau, J. Reporting elevated vancomycin minimum inhibitory concentration in methicillin-resistant *Staphyloccocus aureus*: consensus by an International Working Group. *Future Microbiol.* **2019**, *14*, 4. (b) King, M. D.; Humphrey, B. J.; Wang, Y. F.; Kourbatova, E. V.; Ray, S. M.; Blumberg, H. M. Emergence of community-acquired methicillin-resistant *Staphylococcus aureus* USA 300 clone as the predominant cause of skin and soft-tissue infections. *Ann. Intern. Med.* **2006**, *144*, 309–317. (c) Stryjewski, M. E.; Chambers, H. F. Skin and soft-tissue infections caused by community-acquired methicillin-resistant *Staphylococcus aureus. Clin. Infect. Dis.* **2008**, *46*, S368–S377.

(13) Tang, H.; Porras, G.; Brown, M. M.; Chassagne, F.; Lyles, J. T.; Bacsa, J.; Horswill, A. R.; Quave, C. L. Triterpenoid acids isolated from *Schinus terebinthifolia* fruits reduce *Staphylococcus aureus* virulence and abate dermonecrosis. *Sci. Rep.* **2020**, *10*, 8046.

(14) (a) Popiel, D. I.; Koczyk, G.; Dawidziuk, A.; Gromadzka, K.; Blasczyk, L.; Chelkowski, J. Zearalenone lactonohydrolase activity in Hypocreales and its evolutionary relationships within the epoxide hydrolase subset of a/b-hydrolases. *BMC Microbiol.* **2014**, *14*, 82. (b) Ferrigo, D.; Raiola, A.; Piccolo, E.; Scopel, C.; Causin, R. *Trichoderma harzianum* T22 induces in maize systemic resistance against *Fusarium verticillioides. J. Plant Pathol.* **2014**, *96*, 133–142. (c) Harman, G. E. Overview of mechanisms and uses of *Trichoderma* spp. *Phytopathology* **2015**, *96*, 190–194.

(15) (a) Xie, L.; Zang, X.; Cheng, W; Zhang, Z.; Zhou, J.; Chen, M.; Tang, Y. Harzianic acid from *Trichoderma afroharzianum* is a natural product inhibitor of acetohydroxyacid synthase. *J. Am. Chem. Soc.* **2021**, *143*, 9575–9584. (b) Zhang, J.-L.; Tang, W.-L.; Huang, Q.-R.; Li, Y.-Z.; Wei, M.-L.; Jiang, L.-L.; Liu, C.; Yu, X.; Zhu, H.-W.; Chen, G.-Z.; Zhang, X.-X. *Trichoderma*: a treasure house of structurally diverse secondary metabolites with medicinal importance. *Front. Microbiol.* **2021**, *12*, 723828. (c) Degenkolb, T.; Nielsen, K. F.; Dieckmann, R.; Branco-Rocha, F.; Chaverri, P.; Samuels, G. J.; Thrane, U.; Döhren, H.; Vilcinskas, A.; Brückner, H. Peptaibol, secondary-metabolite, and hydrophobin pattern of commercial biocontrol agents formulated with species of the *Trichoderma harzianum* complex. *Chem. Biodivers.* **2015**, *12*, 662–684. (d) Liu, S.-Y.; Lo, C.-T.; Chen, C.; Liu, M.-Y.; Chen, J.-H.; Peng, K.-C. Efficient isolation of anthraquinone-derivatives from *Trichoderma harzianum* ETS 323. *J. Biochem. Biophys. Methods* **2007**, *70*, 391–395.

(16) Henrik, N.; Cantor, M.; Dusheyko, S.; Hua, S.; Poliakov, A.; Shabalov, I.; Smirnova, T.; Grigoriev, I. V.; Dubchak, I. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* **2014**, *42*, D26–D31.

(17) Argay, G.; Kalman, A.; Kovacevic, N.; Grubisic, D.; Ribar, B. Crystal structure of 1,8-dihydroxy-3-methyl-9,10-anthracendion, $C_{15}H_{10}O_4$. *Zeitschrift fuer Kristallographie – Crystalline Materials*, **1996**, *211*, 723–724.

(18) (a) Patterson, J. P.; Parent, L. R.; Cantlon, J.; Eickhoff, H.; Bared, G.; Evans, J. E.; Gianneschi, N. C. Picoliter drop-on-demand dispensing for multiplex liquid cell transmission electron microscopy. *Microsc. Microanal.* **2016**, *22*, 507–514. (b) Gong, X.; Gnanasekaran, K.; Ma, K.; Forman, C. J.; Wang, X.; Su, S.; Farha, O. K.; Gianneschi, N. C. Rapid generation of

metal-organic framework phase diagrams by high-throughput transmission electron microscopy. *J. Am. Chem. Soc.* **2022**, *144*, 6674–6680.

(19) Mastronarde, D. N. A program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc. Microanal.* **2003**, *9*, 1182–1183.

(20) (a) de la Cruz, M. J. Automation of continuous-rotation data collection for MicroED. *Methods Mol. Biol.* **2021**, *2215*, 321–327. (b) Roslova, M.; Smeets, S.; Wang, B.; Thersleff, T.; Xu, H.; Zou, X. *InsteaDMatic*: towards cross-platform automated continuous rotation electron diffraction. *J. Appl. Cryst.* **2020**, *53*, 1217–1224. (c) de la Cruz, Martynowycz, M. W.; Hattne, J.; Gonen, T. MicroED data collection with SerialEM. *Ultramicroscopy* **2019**, *201*, 77–80.

6.9 Experimental Section

6.9.1 Sample Preparation

6.9.1.1 *Schinus terebinthifolia*

This section pertains to section 6.4. The purification, isolation and characterization procedure for ***6.2*** *are described in literature.* [1]

6.9.1.2 *Trichoderma afroharzanium* T-22

This section contains supplementary information for section 6.5 and 6.6.

Trichoderma afroharzianum t-22 (ATCC 20847) frozen stock cultures were kept in 20% glycerol solution at –80 °C. For the seed culture, frozen agar plugs were transferred to YM agar plates (dextrose 10.0 g, peptode 5.0 g, malt extract 2.5 g, yeast extract 2.5 g, agar 20.0 g, 1000 mL deionized H_2O) maintained at 25 °C. Two agar discs from 1-week old YM agar cultures were inoculated into 50 mL YM semisolid medium (dextrose 10.0 g, peptode 5.0 g, malt extract 2.5 g, yeast extract 2.5 g, agar 3.0 g, 1000 mL deionized H_2O) in a 250 mL flask. Seed cultures were grown at 25 °C, 220 rpm for 4 d. Then, 0.5 mL of seed culture was used to inoculate 15.0 mL of CYS80 medium (sucrose 80.0 g, yellow cornmeal 50.0 g, yeast extract 1.0 g in 1000 mL deionized H_2O) that were incubated statically with vials slanted at a 45 ° angle for 14 days at 23 °C. Cultures were extracted by adding the same volume of ethyl acetate followed by shaking for 2 hours. Organic phases were transferred to vial and dried. The resulting crude was dissolved in 0.5 mL DMSO. A 50 µL aliquot was diluted with 50 µL of MeOH and submitted to semipreparative HPLC microfractionation (100 µL injection; Cosmosil-5C_{18}-MS-II column 10 x 250 mm, 40 °C; gradient elution, 25% A for 2 minutes, 25–100% of A over 18 minutes, and hold at 100% A for 1.5 min, with 0.1% formic acid in acetonitrile (A) and 0.1% aqueous formic acid (B); 3.0 mLmin^{-1}). Aliquots of 750 µL were collected in 15 s intervals on a 96-well polypropylene DeepWell plate.

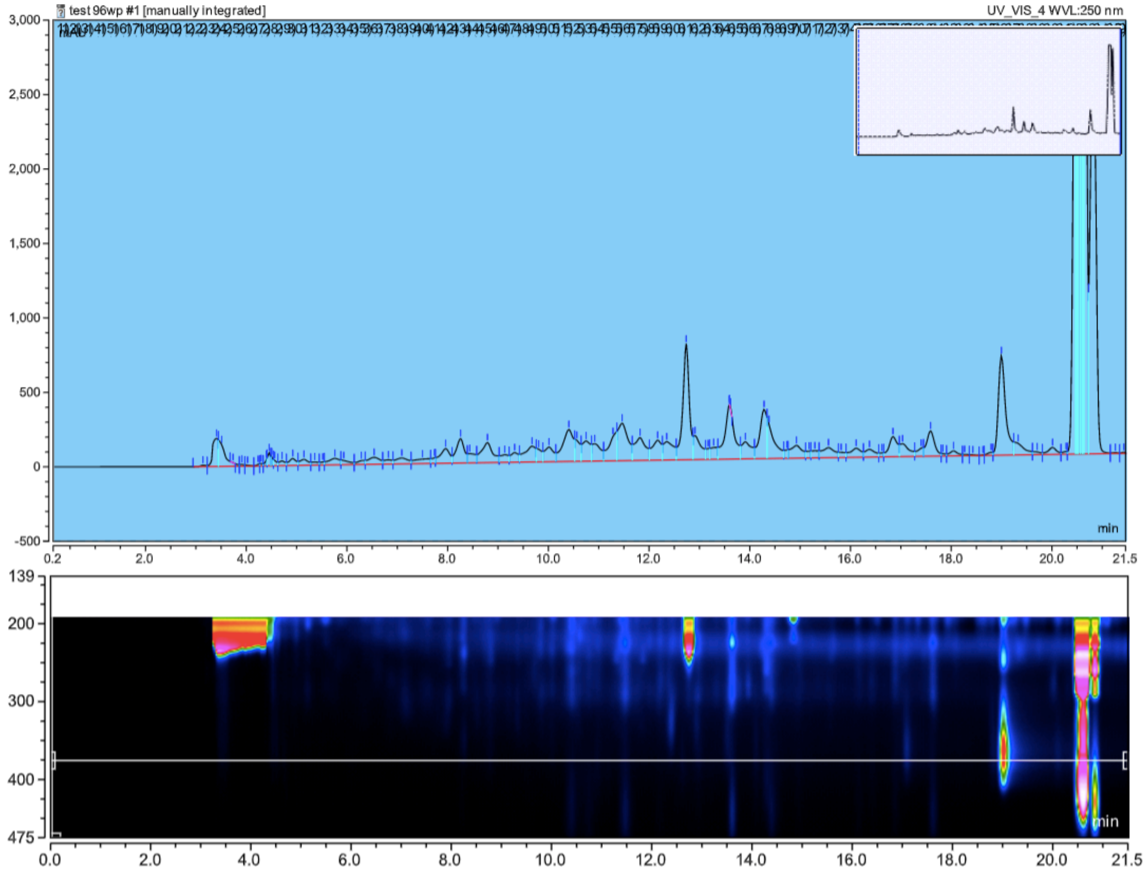


Figure 6.8 Semipreparative HPLC chromatogram of *Trichoderma afroharzianum* t-22 crude extract in CYS80.

Following MicroED analysis, the presence of metabolites **6.4** and **6.5** was further confirmed by HRMS. A small amount of solid from the bottom of the dried wells (H3 and H4) were transferred to a microcentrifuge tube, dissolved in MeOH and submitted to UPLC-DAD-HRMS. As expected, the major constituent of H3 showed a peak with m/z 255.0651 (expected for $[C_{15}H_{10}O_4+H]^+$: 255.0657; -2.35 ppm error) and DAD profile consistent with anthraquinones (Figure 6.9, top). Similarly, H4 showed a major peak with m/z 239.0709 (expected for $[C_{15}H_{10}O_3+H]^+$: 239.0708; 0.42 ppm error) and similar DAD profile (Figure 6.9, bottom).

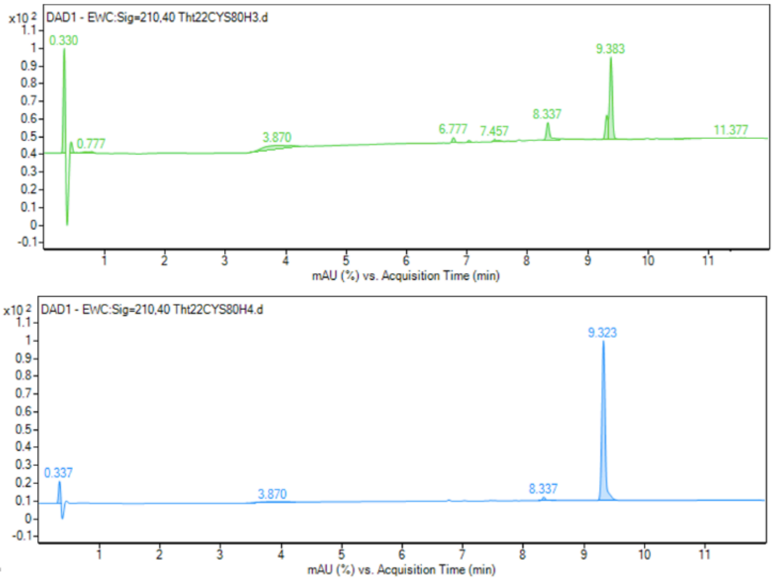


Figure 6.9 UPLC-DAD chromatograms for H3 (top, green) and H4 (bottom, cyan).

6.9.2 Microarraying Conditions

The 350 µL of each fractionated well samples in varying ratios of acetonitrile:water contained in a 96-well polypropylene DeepWell plate were transferred to a shallow 96-well polypropylene or glass-coated plate. The sciFLEXARRAYER S3 system (SCIENON AG, Berlin, Germany) was equipped with a PDC 60 Type 3 capable of depositing drop volume between 220–300 pL. The microarrayer was operated at 30% humidity, and the deck and probe plate temperatures were controlled to be at the dew point (~11 °C). Prior to the runs, nozzle and camera were aligned, and the grids were placed on a glass slide resting next to a spacer on the deck. The grids were manually aligned using the head camera to ensure that the grid squares were perpendicular to the X and Y axes of the microarrayer. After the alignment, two runs were utilized for each full 96-well plates: the first run deposits the 96 samples onto a grid, followed by a second run that deposits a row (or any desired shape) of saturated labeling solution containing either saturated aqueous carbamazepine solution or saturated aqueous sodium chloride solution.

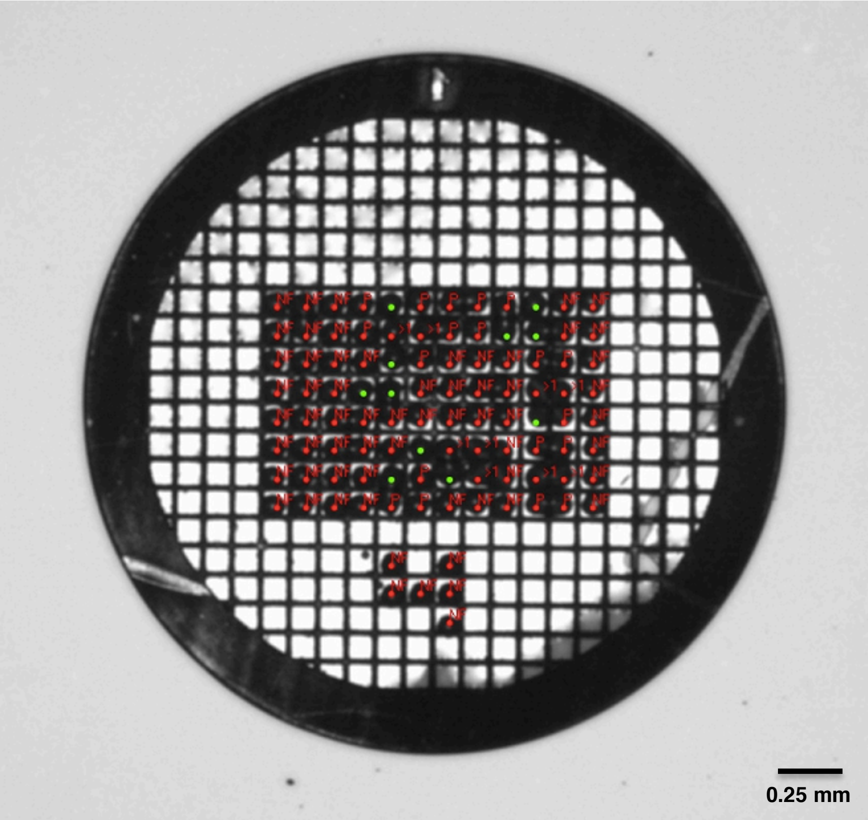


Figure 6.10 Representation of microarrayed droplet placement on a TEM grid. Red and green dots represent the expected drop deposition. Depicted labeling is in the shape of a “4”, which was used to determine the “up” or “down” orientation of the grid during early experiments.

6.9.2.1 Probe Settings

“Probe” refers to the 96-well plate containing the samples to be deposited.

Dimension X	86000
Dimension Y	128000
Dimension Z	9900
Well No. X	8
Well No. Y	12
Well Distance X	9000
Well Distance Y	9000
Well Depth	8000
X Position	5970
Y Position	281000
Z Position (above)	27000
Z Position (into)	35000
Z Vel. [µm/s] into	5000
Probe No. X	1
Probe Offset X	0

Probe No. Y	1
Probe Offset Y	0
Module No. X	1
Module Offset X	0

6.9.2.2 Target Settings

"Target" refers to the TEM grid, where the 96-well plate is deposited on.

Dimension X	3100
Dimension Y	3100
Dimension Z	1200
X Position	138926
Y Position	287303
Z Position	28000
Z Position (into)	0
Spot-Area X	3100
Spot-Area Y	3100
Spot-Area No. X	5
Spot-Area Offset X	5000
Spot-Area No. Y	1
Spot-Area Offset Y	0
Rotate (No = 0 < Yes)	0
Module No. X	1
Module Offset X	0
Module No. Y	1
Module Offset Y	0
Field Velocity X	5000
Field Velocity Y	5000
Spot-Area Teach Corner (0 = LeftFront, CW)	0
Module Height	0
Display (1 = 90°CW)	0
Z Offsets (Yes = 1)	0
Head Cam Focus (0 = into, 1 = Z)	0
Z Into Field (Yes = 1)	0

6.9.2.3 Task List for 96-Well Plate Microarraying

TargetPosition
BeginLoop
TakeProbe_7ul
SpotProbeRun
WashFlush_Medium
EndLoop

QC

6.9.3 Setting Up SerialEM Grid Montage

SerialEM (version 3.8.18)^2 was used to collect the full grid montage. Under the “Navigator” drop-down menu, “Setup Full Montage” feature within the “Montaging & Grids” tab was opened. The well-aligned microscope was set to magnification of LM 210x. 9 pieces in X and 9 pieces in Y direction were collected, binning by 2. The minimum overlap between the pieces was set to 25% and 0.5 micron. Using these settings, total area of 13984 x 13984 pixels, or 2071.7 x 2071.7 microns, were sampled.

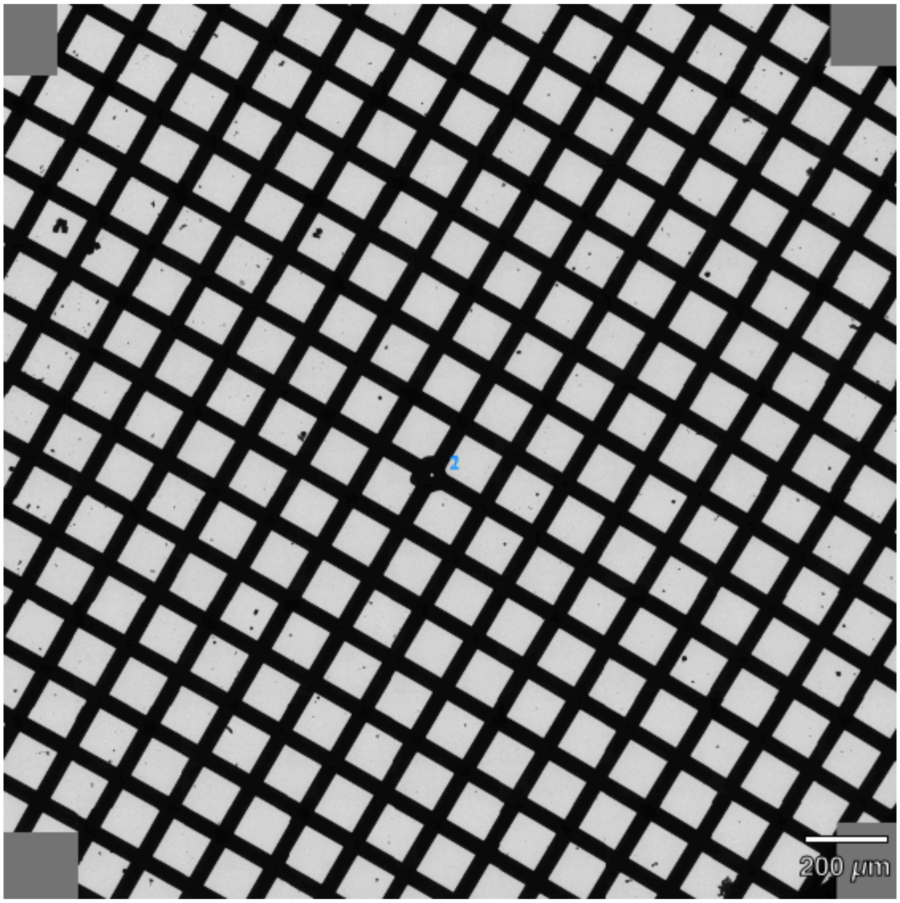


Figure 6.11 An example SerialEM grid montage.

6.9.4 Electron Diffraction Data Acquisition and Processing

All samples from *S. terebinthifolia* were initially screened using the following procedure. If the sample looked solid (430F-F5, 430F-F5-PF11, 430F-F5-PF12, 430F-F5-PF11-SF5, 430-F5-PF12-SF9, 430-F5-PF12-SF10), the fraction was deposited onto a continuous carbon EM grid as previously described.[3] If screening of the solid sample did not yield diffraction, the sample was reconstituted using various solvents. If the sample was oily (430, 430F), suitable solvent was added to the fraction. In both of these cases, 2 μL drop was loaded onto a continuous carbon EM grid and let dry before screening. The grid was then transferred to the TEM on a single-tilt sample holder to be screened at room temperature. The grids containing fractions 430F-F5-PF12 and 430-F5-PF12-SF9 were transferred to the TEM on a Gatan 626 cryo-transfer holder, and cooled to cryogenic temperatures inside the TEM before data acquisition.

All well solutions of fractionated *T. afroharzianum* t-22 extract were initially screened using the following procedure: 2 μL drops were deposited onto a Quantifoil holey-carbon EM grid and let dry before being transferred to the TEM on a single-tilt sample holder to be screened at room temperature. *T. afroharzanium* t-22 extract was microarrayed using the procedure previously described in section 6.9.2. Microarrayed grids were clipped and inserted into a cassette to be loaded onto the TEM. After loading, SerialEM grid montage was set up as described in section 6.9.3.

Diffraction data for *S. terebinthifolia* and *T. afroharzianum* t-22 were collected using the TEM User Interface on an FEI Tecnai F200C electron microscope, and diffraction data for microarrayed grids containing *T. afroharzianum* t-22 were collected using the TEM User Interface on an FEI Arctica electron microscope. Both electron microscopes have an operating voltage of 200 keV, corresponding to a wavelength of 0.025 Å. During data acquisition, the crystal of interest was isolated using a selected area aperture and continuously rotated at a rate of –0.3° s^{-1} over a tilt range of 50–100°. Continuous rotation diffraction data was recorded on TEM Imaging & Analysis (TIA) using rolling shutter mode with a Ceta-D CMOS 4k × 4k camera, integrating at a rate of 3 s per frame and binning by 2 to produce final images of 2k × 2k.[4]

Diffraction movies saved as SER files were converted to SMV format using ser2smv[5] software, as described previously[6]. Frames were indexed and integrated in XDS[7]. Data from multiple crystals were scaled and merged together using XSCALE[8] to produce the final data sets (4, 2, 3, and 3 crystals merged for **6.2**, **6.4**, **6.4**+**6.5** cocrystal, and **6.5**, respectively). Finally, intensities were converted to SHELX format using XDSCONV[8].

6.9.5 Structure Determination and Refinement

Structure of **6.2** was solved *ab initio* using direct methods in SHELXD[9]. The structures of **6.4**, **6.4**+**6.5** cocrystal, and **6.5** were solved *ab initio* using SHELXT[10]. All structures were refined with SHELXL[11] in ShelXle[12]. Non-hydrogen atoms were refined anisotropically, and hydrogen atoms were placed using the riding model. Overall data quality and refinement statistics are reported in Table 6.1, 6.3, 6.5, and 6.7.

6.9.6 Electron Micrographs and Diffraction Patterns of Microcrystals

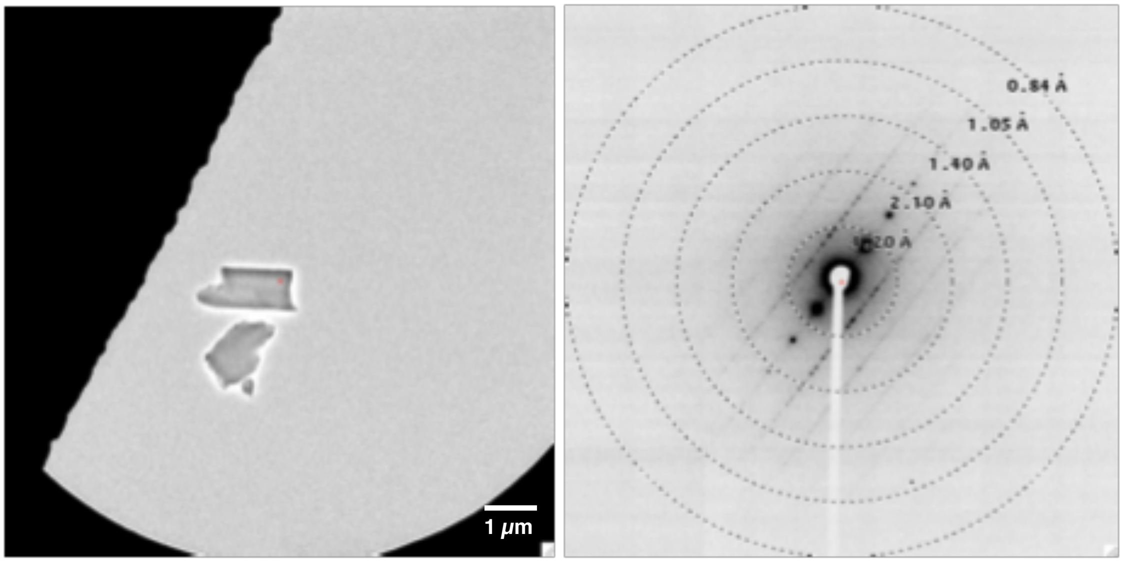


Figure 6.12 Electron micrograph and diffraction pattern of 430F-F5-PF11-SF5 (containing **6.1**) at SA 2600x magnification, 960 mm detector distance.

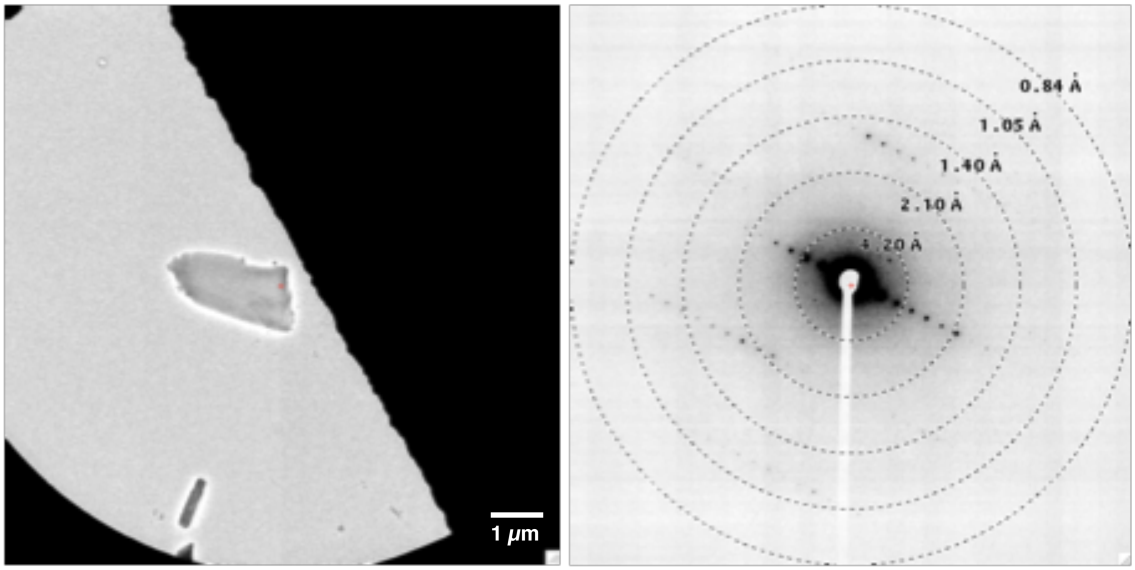


Figure 6.13 Electron micrograph and diffraction pattern of 430F-F5-PF12-SF9 (containing **6.2**) at SA 2600x magnification, 960 mm detector distance.

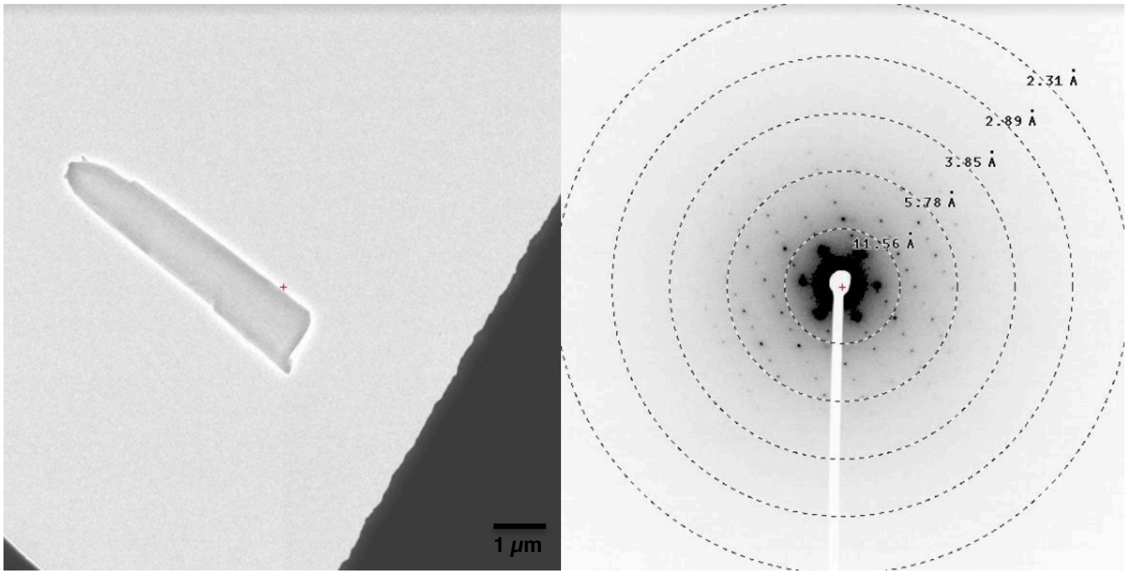


Figure 6.14 Electron micrograph and diffraction pattern of 430F-F5-PF12-SF10 (containing **6.3**) at SA 2600x magnification, 2640 mm detector distance.

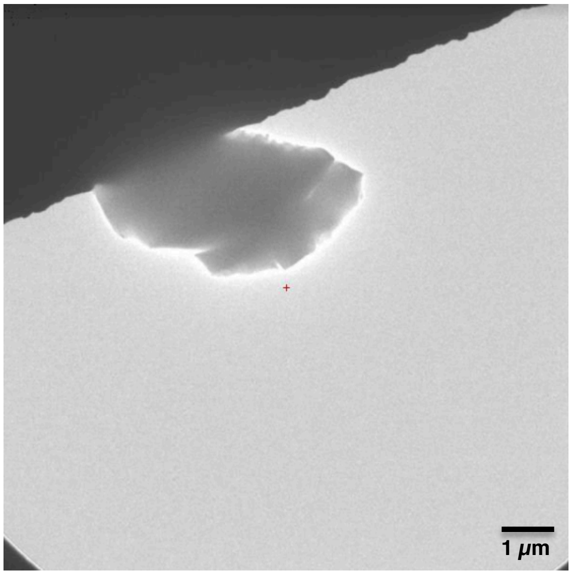
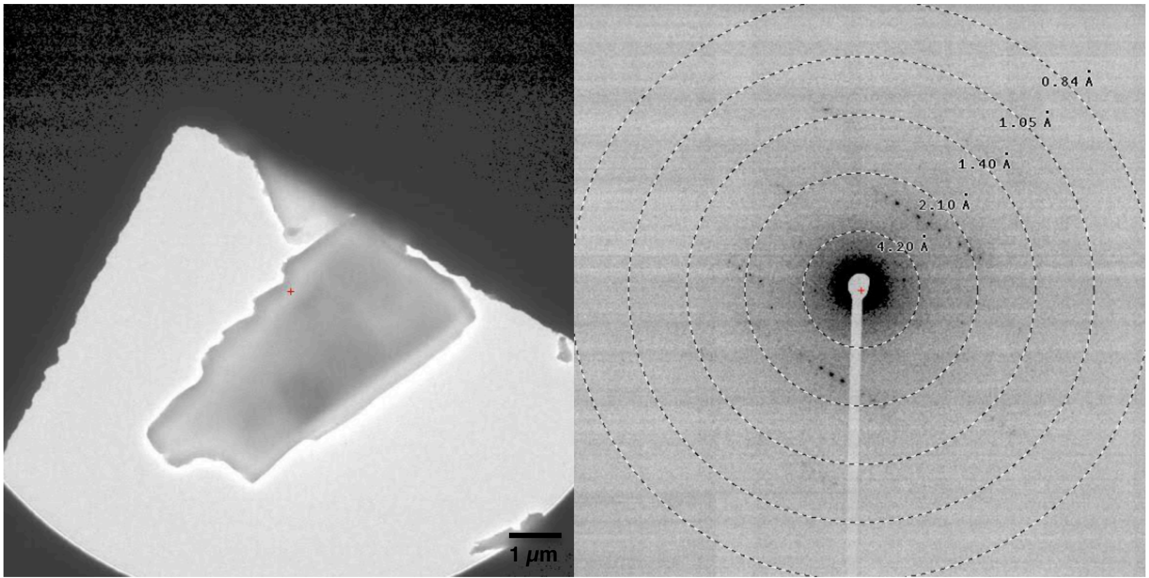


Figure 6.15 Electron micrograph of 430F-F5-PF11 (containing **6.1** among many others) at SA 2600x magnification. Diffraction was not observed.



	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
430F-F5-PF12	7.00(1)	20.11(3)	19.782(9)	90	94.70(9)	90
430F-F5-PF12-SF9	6.948(5)	19.988(7)	19.789(7)	90	94.70(9)	90
Absolute Error	0.052	0.122	0.007	N/A	0.00	N/A

Figure 6.16 Electron micrograph and diffraction pattern of 430F-F5-PF12 (containing **6.2** and **6.3** among many others) at SA 2600x magnification, 960 mm detector distance. Unit cell parameters are consistent with those obtained from 430F-F5-PF12-SF9.

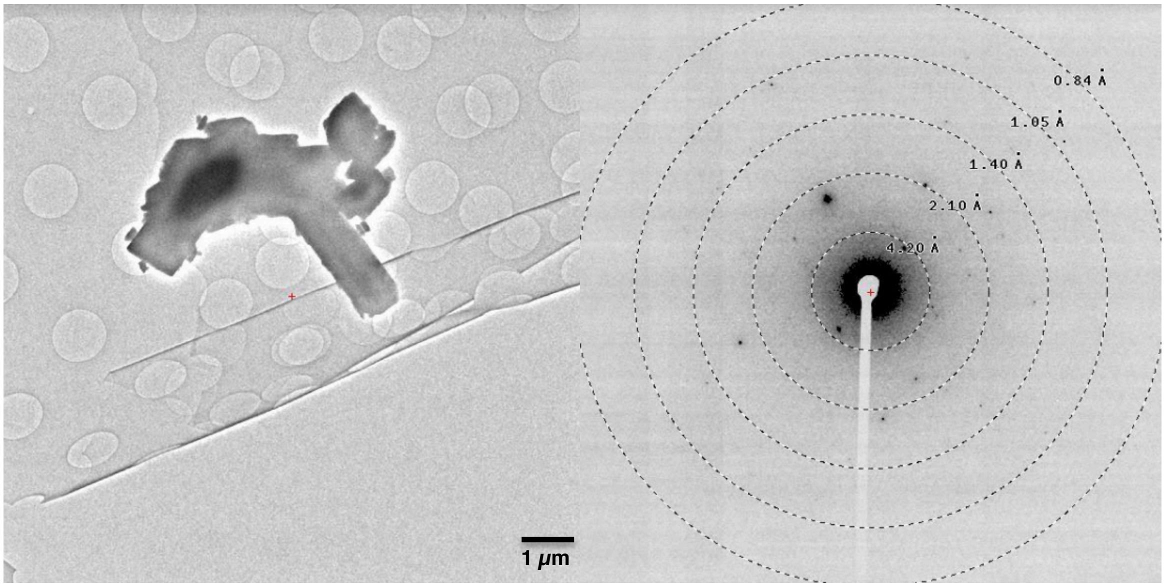


Figure 6.17 Electron micrograph and diffraction pattern of 430F-F5 (containing **6.1**, **6.2** and **6.3** among many others) at SA 2600x magnification, 960 mm detector distance.

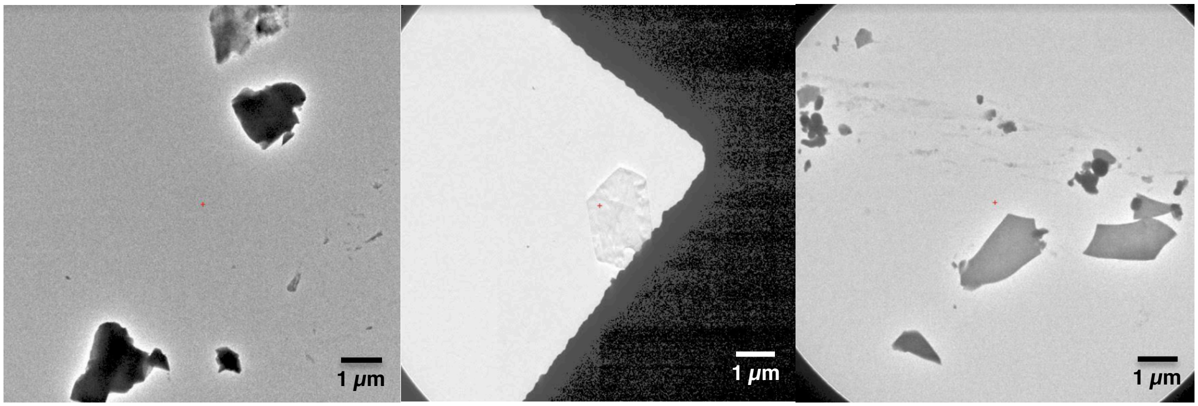


Figure 6.18 Electron micrographs from various wells (low-throughput method) at SA 2600x magnification.

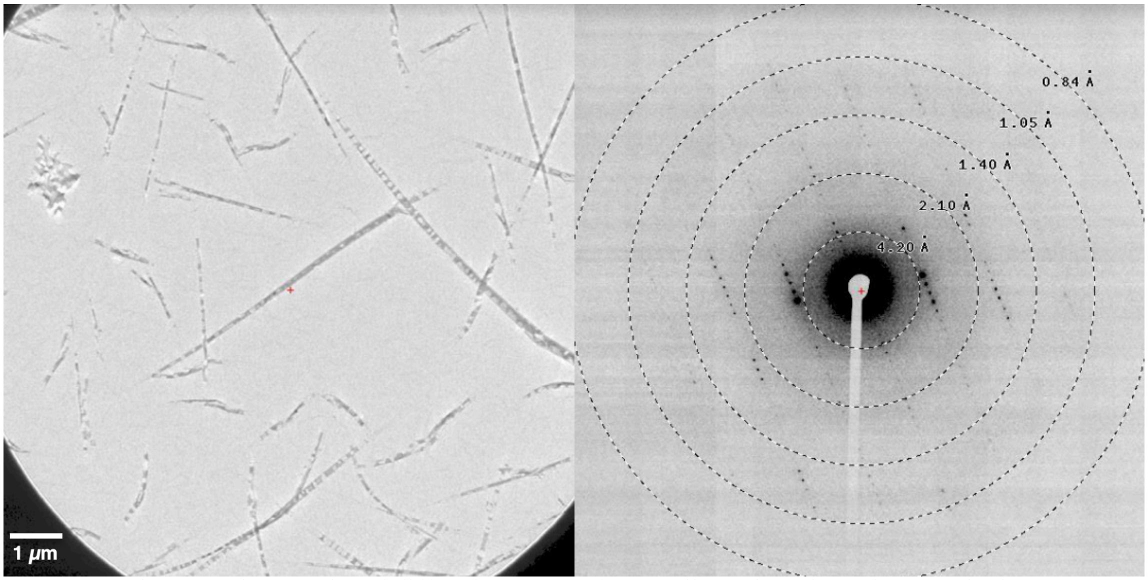
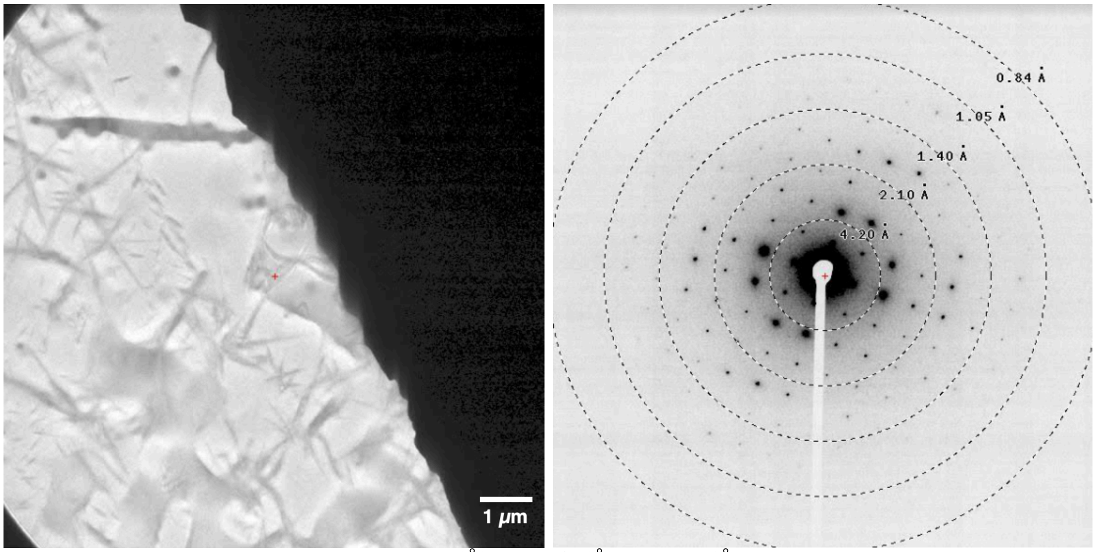


Figure 6.19 Electron micrograph and diffraction pattern from well H3 (low-throughput method) at SA 2600x magnification, 960 mm detector distance.



	Space Group	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
MicroED	C 2/c	10.133(9)	9.452(3)	22.1(2)	90	91.9(2)	90
X-ray	C 2/c	10.737(2)	9.818(1)	21.541(3)	90	92.03(1)	90
Absolute Error	N/A	0.604	0.366	0.559	N/A	0.13	N/A

Figure 6.20 Electron micrograph and diffraction pattern from well H3 at SA 2600x magnification, 960 mm detector distance, that matches the unit cell parameters of previously reported X-ray structure.[17]

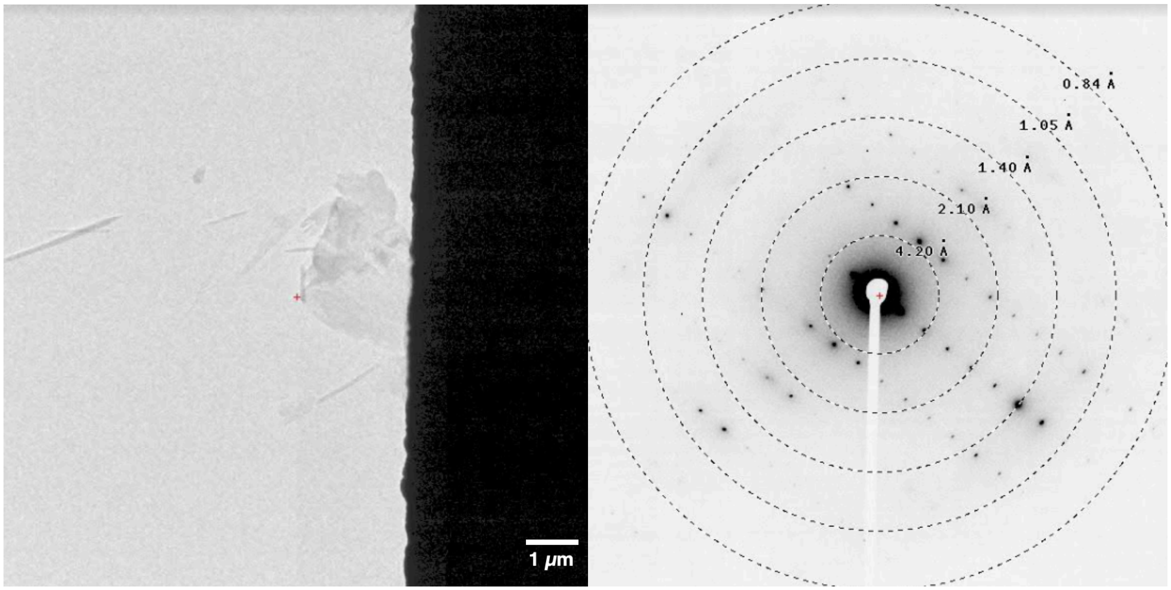


Figure 6.21 Electron micrograph and diffraction pattern from well H4 (low-throughput method) at SA 2600x magnification, 960 mm detector distance.

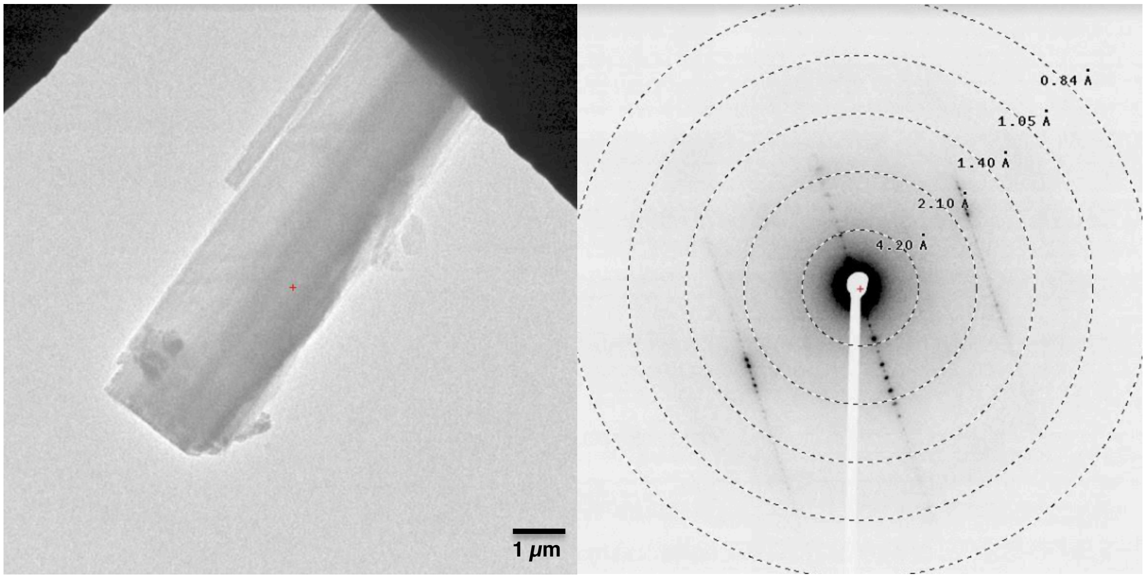


Figure 6.22 Electron micrograph and diffraction pattern from well H5 (low-throughput method) at SA 2600x magnification, 960 mm detector distance.

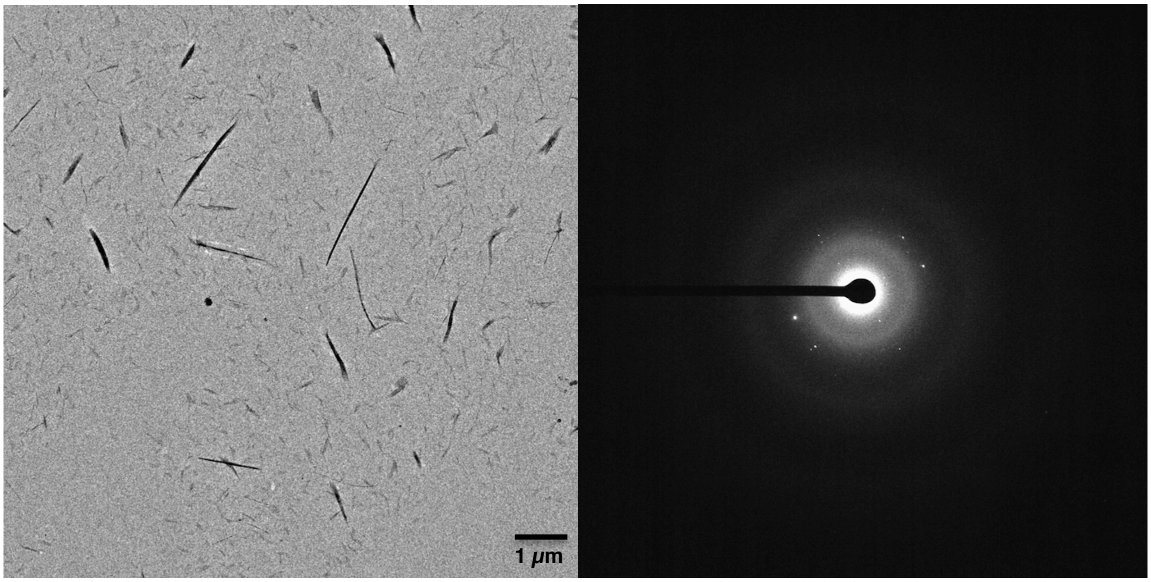


Figure 6.23 Electron micrograph and diffraction pattern from well H3 (high-throughput method) at SA 2600x magnification, 960 mm detector distance. Resolution rings not shown but are identical to Figure 6.18.

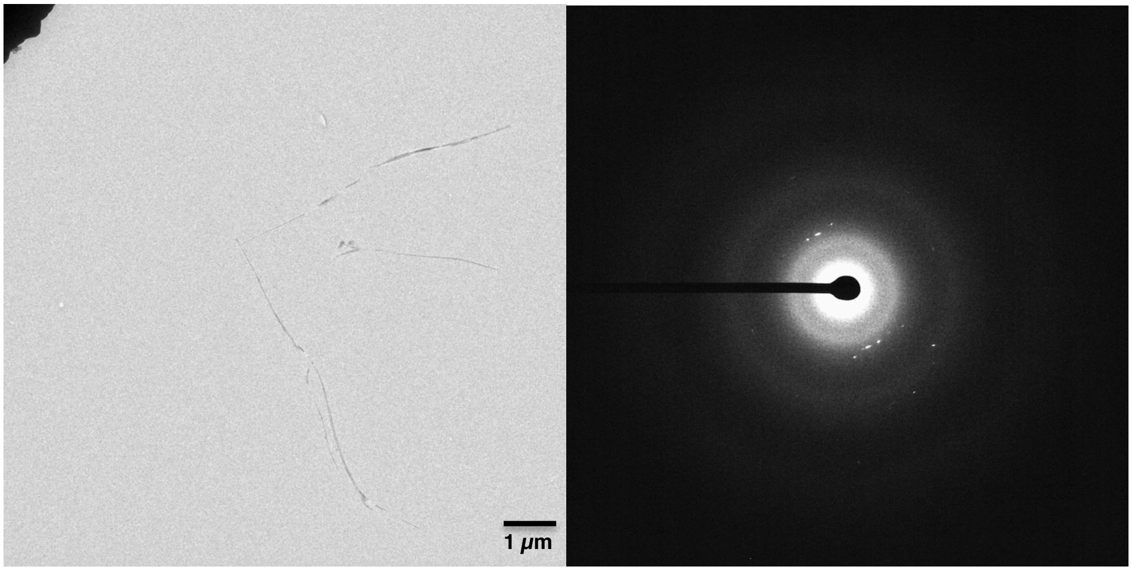


Figure 6.24 Electron micrograph and diffraction pattern from well H4 (high-throughput method) at SA 2600x magnification, 960 mm detector distance. Resolution rings not shown but are identical to Figure 6.19.

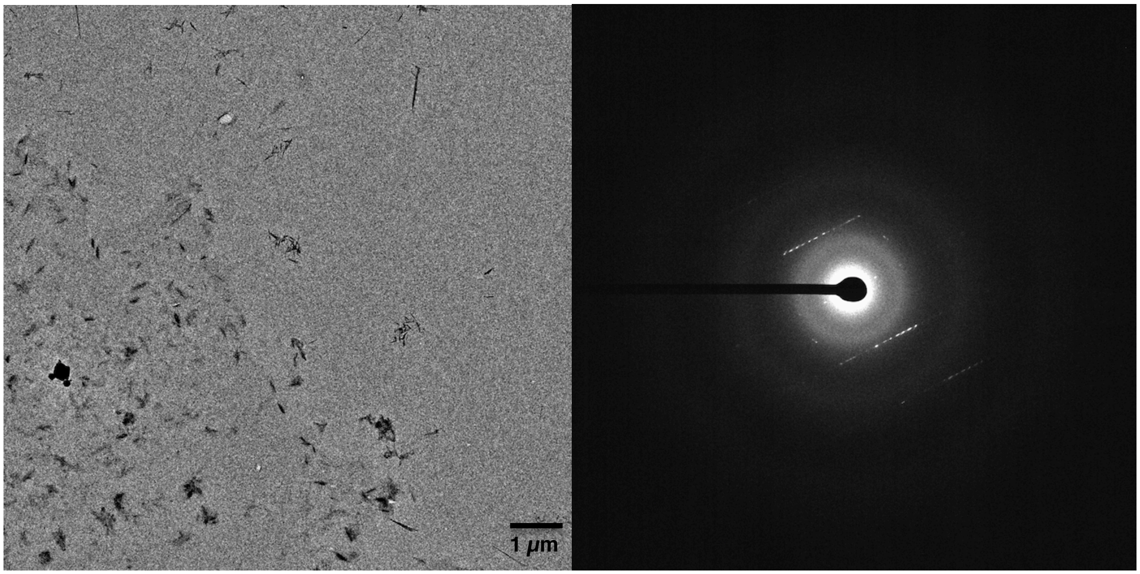


Figure 6.25 Electron micrograph and diffraction pattern from well H5 (high-throughput method) at SA 2600x magnification, 960 mm detector distance. Resolution rings not shown but are identical to Figure 6.20.

6.9.7 Crystal Structures of Natural Products

6.9.7.1 3-Oxotirucalla-7,24*Z*-dien-26-oic acid (6.2)

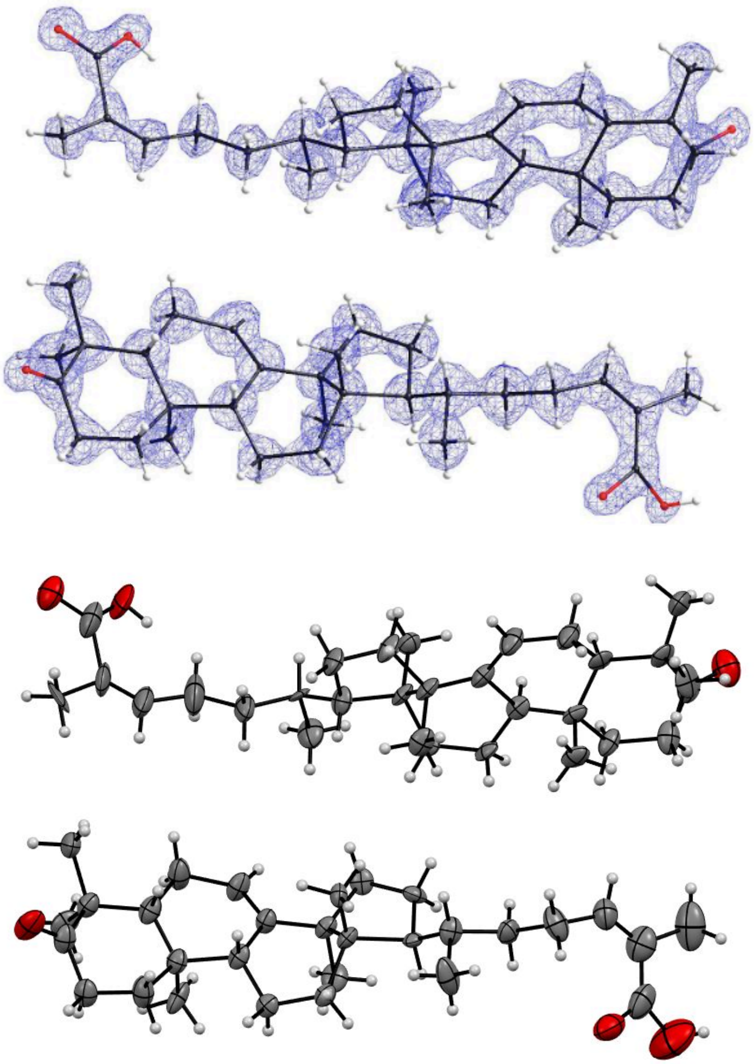


Figure 6.26 Asymmetric unit of 3-oxotirucalla-7,24*Z*-dien-26-oic acid (**6.2**) crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.65 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{30}H_{46}O_3$
Data Collection	
Temperature (K)	100(2)
Space group	P 2(1)
Cell dimensions	
a, b, c (Å)	6.96(7), 20.110(6), 19.89(1)
α, β, γ (°)	90.00, 95.4(2), 90.00

Resolution (Å)	1.05 (1.10–1.05)
R_{obs} (%)	17.9 (89.3)
R_{meas} (%)	19.8 (99.3)
I/σI	5.92 (1.84)
$CC_{1/2}$ (%)	98.3 (60.4)
Completeness (%)	91.1 (82.3)
Data Refinement	
Resolution (Å)	1.05
Observed reflections (#)	9402 (1238)
Unique reflections (#)	2173 (288)
R_1 (%)	11.44
wR_2 (%)	29.88
GooF	1.298

Table 6.1 Data collection and refinement statistics for **6.2**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4763	Data was truncated to a resolution of 1.0 angstrom, due to higher resolution data falling outside of the edge of the detector.
PLAT213_ALERT_2_A Atom C065 has ADP max/min Ratio ..... 6.0 prolat	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering and atom displacement.
PLAT414_ALERT_2_A Short Intra D-H..H-X H5 ..H04L . 1.57 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in

x,y,z = 1_555 Check	some structural disordering. In addition, these hydrogens were placed using the riding model.

B-level Alerts	Justifications
PLAT149_ALERT_3_B s.u. on the beta Angle is Too Large .... 0.20 Degree	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.0353 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT369_ALERT_2_B Long C(sp2)-C(sp2) Bond C044 - C047 . 1.63 Ang.	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT420_ALERT_2_B D-H Bond Without Acceptor O5 --H5 . Please Check	Presence of a solvent molecule to engage in hydrogen bonding was not observed.
PLAT420_ALERT_2_B D-H Bond Without Acceptor O6 --H6 . Please Check	Presence of a solvent molecule to engage in hydrogen bonding was not observed.
PLAT430_ALERT_2_B Short Inter D...A Contact O2 ..O4 . 2.60 Ang. x,1+y,z = 1_565 Check	Diffraction data from four crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.

Table 6.2 Justification for alerts in 3-oxotirucalla-7,24*Z*-dien-26-oic acid (**6.2**) crystal structure.

6.9.7.2 Chrysophanol (6.4) from Well H3

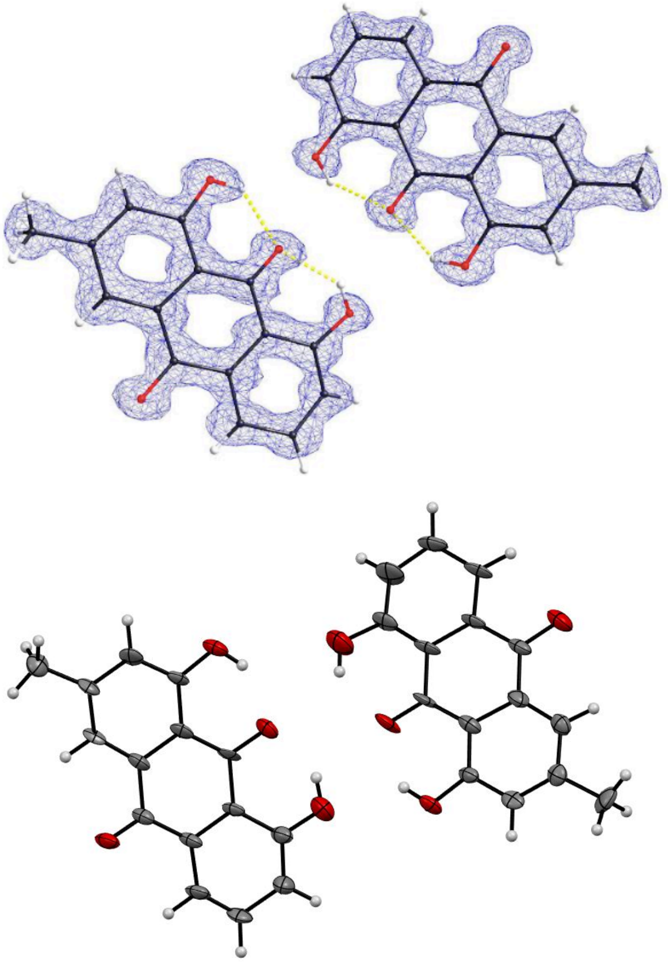


Figure 6.27 Asymmetric unit of chrysophanol (**6.4**) crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.46 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{15}H_{10}O_4$
Data Collection	
Temperature (K)	293(2)
Space group	$P\ 2_1\ 2_1\ 2_1$
Cell dimensions	
a, b, c (Å)	3.71(7), 22.250(6), 23.670(1)
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	0.90 (0.95–0.90)

R_{obs} (%)	22.6 (39.0)
R_{meas} (%)	25.5 (45.1)
I/σI	4.37 (2.24)
$CC_{1/2}$ (%)	98.4 (92.7)
Completeness (%)	85.0 (60.4)
Data Refinement	
Resolution (Å)	0.90
Observed reflections (#)	6454 (573)
Unique reflections (#)	1482 (151)
R_1 (%)	16.16
wR_2 (%)	38.56
GooF	1.447

Table 6.3 Data collection and refinement statistics for **6.4**. Highest resolution shell is shown in parenthesis.

B-level Alerts	Justifications
RINTA01_ALERT_3_B The value of Rint is greater than 0.18 Rint given 0.214	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
THETM01_ALERT_3_B The value of sine(theta_max)/wavelength is less than 0.575 Calculated sin(theta_max)/wavelength = 0.5556	Data was truncated to a resolution of 0.9 angstrom, due to higher resolution data falling outside of the edge of the detector.
PLAT020_ALERT_3_B The Value of Rint is Greater Than 0.12 .... 0.214 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.

PLAT082_ALERT_2_B High R1 Value .... 0.16 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) .... 0.39 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT213_ALERT_2_B Atom C006 has ADP max/min Ratio ..... 4.1 prolat	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering and atom displacement. Additionally, data collection occurred at ambient temperature, resulting in more thermal displacement.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C006 -C00Q . 1.34 Ang.	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.03824 Ang.	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 6.4 Justification for alerts in chrysophanol (**6.4**) crystal structure.

6.9.7.3 Chrysophanol (6.4) and Pachybasin (6.5) from Well H3

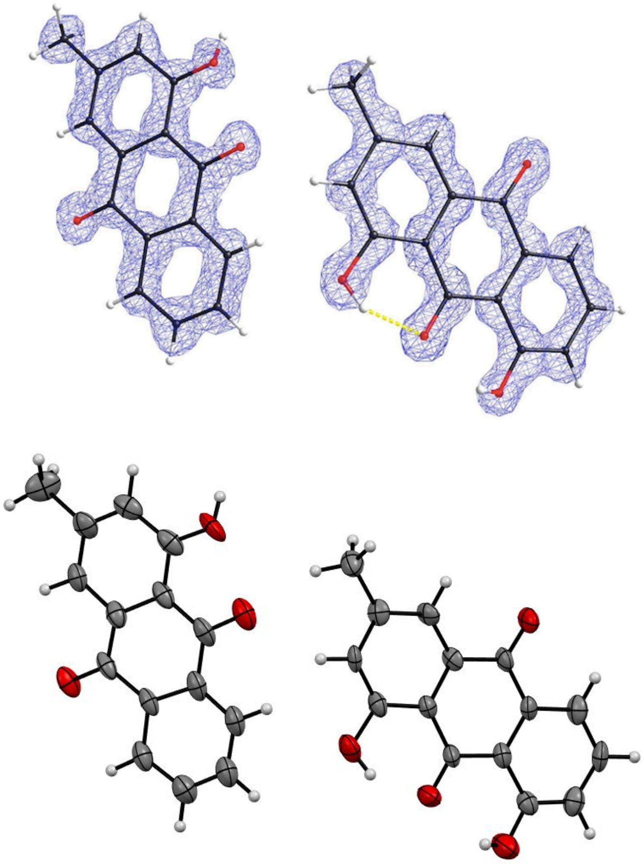


Figure 6.28 Asymmetric unit of co-crystallized chrysophanol (**6.4**) and pachybasin (**6.5**) structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.46 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{15}H_{10}O_4$, $C_{15}H_{10}O_3$
Data Collection	
Temperature (K)	100(2)
Space group	C 2/c
Cell dimensions	
a, b, c (Å)	36.72(7), 3.670(6), 32.3(1)
α, β, γ (°)	90.00, 116.2(2), 90.00
Resolution (Å)	1.00 (1.05–1.00)
R_{obs} (%)	21.2 (66.7)

R_{meas} (%)	24.1 (63.6)
I/σI	3.84 (1.99)
$CC_{1/2}$ (%)	98.5 (91.0)
Completeness (%)	87.4 (90.3)
Data Refinement	
Resolution (Å)	1.00
Observed reflections (#)	9402 (1238)
Unique reflections (#)	2173 (288)
R_1 (%)	15.73
wR_2 (%)	37.34
GooF	1.668

Table 6.5 Data collection and refinement statistics for **6.4**+**6.5** co-crystal structure. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4999	Data was truncated to a resolution of 1.0 angstrom, due to higher resolution data falling outside of the edge of the detector.
PLAT415_ALERT_2_A Short Inter D-H..H-X H0K ..H00H . 1.79 Ang. 1-x,1+y,3/2-z = 2_666 Check	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. In addition, these hydrogens were placed using the riding model.

B-level Alerts	Justifications
RINTA01_ALERT_3_B The value of Rint is greater than 0.18 Rint given 0.201	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.

PLAT020_ALERT_3_B The Value of Rint is Greater Than 0.12 .... 0.201 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.
PLAT082_ALERT_2_B High R1 Value .... 0.16 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.
PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) .... 0.37 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.
PLAT149_ALERT_3_B s.u. on the beta Angle is Too Large .... 0.20 Degree	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C1 -C00C . 1.33 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Electron data also tends to have shorter bond lengths.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C2 -C00H . 1.34 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Electron data also tends to have shorter bond lengths.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C006 -C011 . 1.34 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Electron data also tends to have shorter bond lengths.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.02588 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT420_ALERT_2_B D-H Bond Without Acceptor O00K --H0K . Please Check	Presence of a solvent molecule to engage in hydrogen bonding was not observed.

Table 6.6 Justification for alerts in chrysophanol (**6.4**) and pachybasin (**6.5**) co-crystal structure.

6.9.7.4 Pachybasin (6.5) from Wells H4–H5

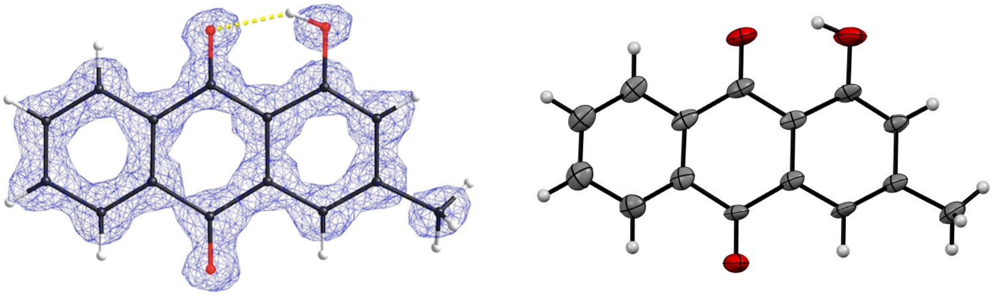


Figure 6.29 Asymmetric unit of pachybasin (**6.5**) crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.74 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{15}H_{10}O_3$
Data Collection	
Temperature (K)	293(2)
Space group	$P\ 2_1/c$
Cell dimensions	
a, b, c (Å)	7.840(6), 3.730(3), 33.69(8)
α, β, γ (°)	90.00, 93.57(8), 90.00
Resolution (Å)	0.90 (0.95–0.90)
R_{obs} (%)	18.0 (42.2)
R_{meas} (%)	19.4 (45.5)
I/σI	5.44 (3.14)
$CC_{1/2}$ (%)	99.0 (91.2)
Completeness (%)	84.5 (86.2)
Data Refinement	
Resolution (Å)	0.90
Observed reflections (#)	9315 (1555)

Unique reflections (#)	1405 (213)
R_1 (%)	17.32
wR_2 (%)	42.72
GooF	2.346

Table 6.7 Data collection and refinement statistics for **6.5**. Highest resolution shell is shown in parenthesis.

B-level Alerts	Justifications
THETM01_ALERT_3_B The value of sine(theta_max)/wavelength is less than 0.575 Calculated sin(theta_max)/wavelength = 0.5556	Data was truncated to a resolution of 0.9 angstrom, due to higher resolution data falling outside of the edge of the detector.
PLAT082_ALERT_2_B High R1 Value .... 0.17 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.
PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) .... 0.43 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged.
PLAT149_ALERT_3_B s.u. on the beta Angle is Too Large .... 0.08 Degree	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C003 -C008 . 1.33 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Electron data also tends to have shorter bond lengths.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C004 -C00D . 1.34 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Electron data also tends to have shorter bond lengths.

PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.01471 Ang.	Diffraction data from three crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 6.8 Justification for alerts in pachybasin (**6.5**) crystal structure.

6.9.8 Supplementary Notes and References

(1) Tang, H.; Porras, G.; Brown, M. M.; Chassagne, F.; Lyles, J. T.; Bacsa, J.; Horswill, A. R.; Quave, C. L. Triterpenoid acids isolated from *Schinus terebinthifolia* fruits reduce *Staphylococcus aureus* virulence and abate dermonecrosis. *Sci. Rep.* **2020**, *10*, 8046.

(2) Mastronarde, D. N. A program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc. Microanal.* **2003**, *9*, 1182–1183.

(3) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(4) Nannenga, B. L.; Shi, D.; Leslie, A. G. W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930.

(5) Program ser2smv obtained from https://cryoem.ucla.edu/downloads/snapshots.

(6) Hattne, J; Reyes, F. E.; Nannenga, B. L.; Shi, D.; de la Cruz, M. J.; Leslie, A. G. W.; Gonen, T. MicroED data collection and processing. *Acta Cryst.* **2015**, *A71*, 353–360.

(7) Kabsch, W. XDS. *Acta Cryst.* **2010**, *D66*, 125–132.

(8) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010,** *D66*, 133–144.

(9) Sheldrick, G. M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112–122.

(10) Sheldrick, G. M. SHELXT – Integrated space-group and crystal-structure determination. *Acta Cryst.* **2015**, *A71*, 3–8.

(11) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(12) Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. ShelXle: a Qt graphical user interface for SHELXL. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.

CHAPTER SEVEN

Characterization of Reactive Organometallic Species and Synthetic Reaction Products *via* Microcrystal Electron Diffraction

Adapted from: Christopher G. Jones†, Matthew Asay†, Lee Joon Kim, Jack F. Kleinsasser, Ambarneil Saha, Tyler J. Fulton, Kevin R. Berkley, Duilio Cascio, Andrey G. Malyutin, Matthew P. Conley, Brian M. Stoltz, Vincent Lavallo, José A. Rodríguez, Hosea M. Nelson *ACS Cent. Sci.* **2019**, *5*, 1507–1513.

Unpublished work: Zhenqi Zhao, Stanislav Popov, Lee Joon Kim, Matthew Asay, Hosea M. Nelson

7.1 Contributions

This chapter is a combination of published work and unpublished work. For the unpublished work, Zhenqi Zhao and Dr. Stanislav Popov performed the synthetic experiments. Lee Joon Kim performed the electron diffraction studies from sample preparation to refinement, and Dr. Matthew Asay refined structural data.

7.2 Abstract

Structure determination has fueled the fields of organic and organometallic chemistry. While common structural characterization methods such as solution-state NMR spectroscopy and X-ray crystallography exhibit limitations such as needing multiple experiments for NMR characterization of novel compounds and stringent requirements in crystal quality for X-ray diffraction experiments, microcrystal electron diffraction (MicroED) can circumvent these problems by taking advantage of microcrystals to provide a direct, unambiguous picture of a molecular structure. This work showcases the application of MicroED for characterization of

reactive organometallic complexes and synthetic organic reaction products. Even without the use of cryogenic cooling of the samples, this technique successfully solves the first crystal structure of privileged industrial catalyst, Schwartz's reagent. Additionally, five paramagnetic and diamagnetic transition metal complexes are studied, along with a novel Pd(II) complex whose structure could not be probed using X-ray crystallography or solution-state NMR. Moreover, four reaction products synthesized by organic reactions performed in research settings are characterized in a facile manner, demonstrating MicroED's potential to be used as routine structure determination method for these compounds.

7.3 Introduction

Structural arrangement of atoms forms the basis of chemistry, and the knowledge of these 2- and 3-dimensional structures is indispensable for the field of synthetic chemistry. In studying chemical processes, structures of starting materials and characterization of reaction products are essential, as detecting changes in chemical structures gives chemists an invaluable insight into reaction mechanisms and guides reaction condition optimization. Additionally, structural characterization of reactive catalysts can inform chemists in the design and development of new catalytic species.[1]

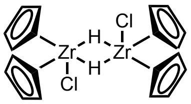
Many techniques exist for characterizing organic compounds, but the most routine method is undoubtedly NMR spectroscopy. Often, comparison of a published ^{1}H NMR spectrum with a pure in-house one can reveal the identity of the analyte. However, in the case of a novel compound, ^{1}H NMR experiments are complemented by ^{13}C NMR, as well as various 2D-NMR techniques such as COSY and NOESY if necessary. These additional NMR experiments can amount to longer periods of time until the structure is fully characterized with confidence. On the other hand, X-ray crystallography can avoid ambiguous NMR interpretations by providing a direct visualization of the molecular structure as well as assign the correct stereochemistry. However, X-ray crystallography is not typically considered a routine method for synthetic chemists due to the need to grow crystals of stringent quality.

In addition to providing unambiguous structure determination or stereochemical assignments for organic compounds, crystallography has also played a critical role in characterizing organometallic complexes and fueling the development of organometallic catalysis. By enabling solid-state structure determination, crystallography has been able to overcome challenges inherent in solution-state characterization of transition-metal complexes, such as paramagnetism and poor solubility or reactivity profiles.[2] While X-ray crystallography is widely used and regarded as the gold standard in solid-state structural analysis, it is limited by the necessity of growing large, single crystals, as described in Chapter 1.[3] As an alternative method, electron diffraction can be utilized to circumvent this problem by analyzing crystals that are orders of magnitude smaller than those required for X-ray crystallography. These microcrystals are much more facile to form and require less material for analysis.

An emerging technique in the field of electron diffraction is microcrystal electron diffraction (microED). Originally developed for macromolecular crystallography, this method has been recently utilized to successfully determine structures from microcrystals of pharmaceutical compounds that were present in seemingly amorphous powders.[4,5] Previously, MicroED studies were often performed under cryogenic temperatures to preserve the hydration state of biomolecules or to attenuate radiation damage from the electron beam. In contrast, this report utilizes a sensitive detector and low electron doses that enable the structure determination of organometallic complexes and synthetic compounds at ambient temperature. Moreover, by applying MicroED to not only commercial compounds but also synthetic molecules obtained as reaction products in real research settings, this study demonstrates the potential of MicroED to be a routine analytical method for solid-state analysis in synthetic organic chemistry and organometallic chemisty.

7.4 The First Solid-State Structure of Schwartz's Reagent

To demonstrate the applicability and facility of MicroED, this study focused on organometallic species whose structures have eluded traditional structural analysis. Chloridobis(η^5-cyclopentadienyl)hydridozirconium **7.1**, colloquially known as Schwartz's reagent, is a privileged reagent widely used for hydrozirconation reactions.[6] However, due to its poor solubility in hydrocarbon and ethereal solvents, as well as its reactivity with polar chlorinated solvents,[7] crystallization and NMR studies have been challenging. Instead, data from FTIR spectroscopy,[8] solid-state ^{35}Cl NMR spectroscopy, [6e,6f] and X-ray structures of derivatives[9] have been used to support an inferred structure of Schwartz's reagent, composed of a dimeric zirconocene bridged by two hydrides (Figure 7.1).

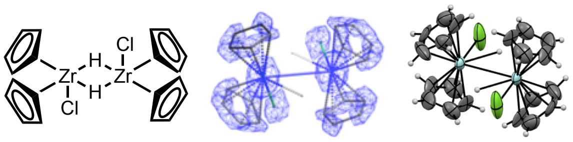


Schwartz's reagent, **7.1**

Structure proposed from FTIR, solid-state ^{35}Cl NMR, X-ray structures of derivatives

Figure 7.1 Proposed structure of Schwartz's reagent as a dimer.

To probe the solid-state structure of this well-known industrial catalyst, a commercial powder sample was acquired and loaded onto a 200 mesh Quantifoil grid for ambient temperature MicroED experiments using a Thermo Fischer Scientific Tecnai F30 transmission electron microscope. Continuous rotation data was collected with a TVIPS XF-416 detector operating at an accelerating voltage of 300 keV. Reducing and merging diffraction data from three microcrystals yielded a 1.15 Å resolution data set with high completeness. *Ab initio* structure solution, obtained using the SHELX suite, confirmed the solid-state structure of Schwartz's reagent as the expected centrosymmetric dimer (Figure 7.2).



Schwartz's reagent, **7.1**

Figure 7.2 Solid-state structure of Schwartz's reagent. Chemdraw structure (left), refined MicroED structure with potential map overlay (middle), and ORTEP structure (right). Thermal ellipsoids shown at 30% probability.

Notably, regions of electron potential corresponding to the bridging hydride ligands were clearly seen in the initial difference Fourier map (Figure 7.3A). To visualize the hydride atoms more closely, the unit cell was sliced into planes that run along the *a*-axis. Upon reaching the central mirror plane that bisects zirconium, chlorine and hydrogen atoms, the hydride atoms can be clearly seen, supporting the refined structure of **7.1** (Figure 7.3B).

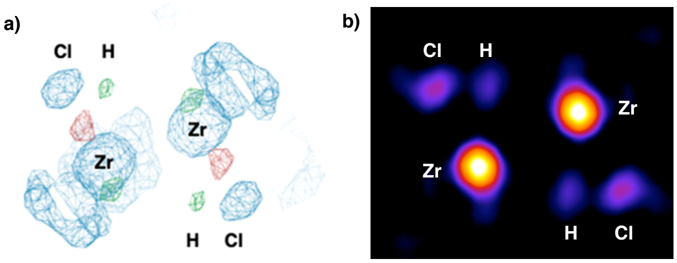


Figure 7.3 The initial difference Fourier map (left) and contour map (right) of Schwartz's reagent showing the bridging hydrides.

7.5 The Solid-State Structure of Ethylene Insertion Product

In demonstrating the applicability of MicroED in structure determination of organometallic compounds, an active research problem involving the reactivity of dimeric Pd(I) species was investigated. Dimeric palladium(I) complexes containig Pd–Pd bonds can react with small molecules to undergo homolytic cleavage or insertion.[10,11] Pd(I) dimer **7.2** was synthesized and characterized using X-ray crystallography and NMR studies (Figure 7.4). Reacting species **7.2** with ethylene gas led to a color change and precipitation of solid particles. However, these solid particles could not be characterized using X-ray or NMR, as reconstituting the precipitate with solvents led to rapid reversion of the product to starting material **7.2**.

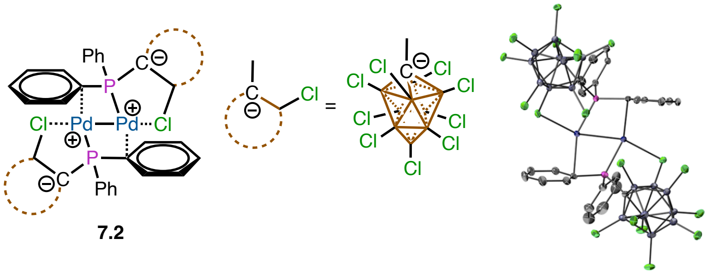


Figure 7.4 X-ray crystal structure of Pd(I) species **7.2**.

Taking advantage of MicroED's ability to determine structure from microcrystalline domains that may be present even in seemingly amorphous material, the precipitate was loaded on an EM grid for ambient temperature MicroED experiments. Under the electron beam, the microcrystals contained in the precipitate sample yielded diffraction patterns that were reduced to a data set with high completeness. The resulting data set produced an *ab initio* solution that revealed an unambiguous structure of the unexpected oxidative insertion product **7.3**, featuring an ethylene linker connecting the two Pd(II) atoms (Figure 2.5). The $^{13}\text{C}\{^{1}\text{H}\}$, $^{11}\text{B}\{^{1}\text{H}\}$, and $^{31}\text{P}\{^{1}\text{H}\}$ cross-polarization magic angle spinning (CPMAS) NMR spectroscopy experiments were consistent with the MicroED structure and shows the homogeneity of the sample.

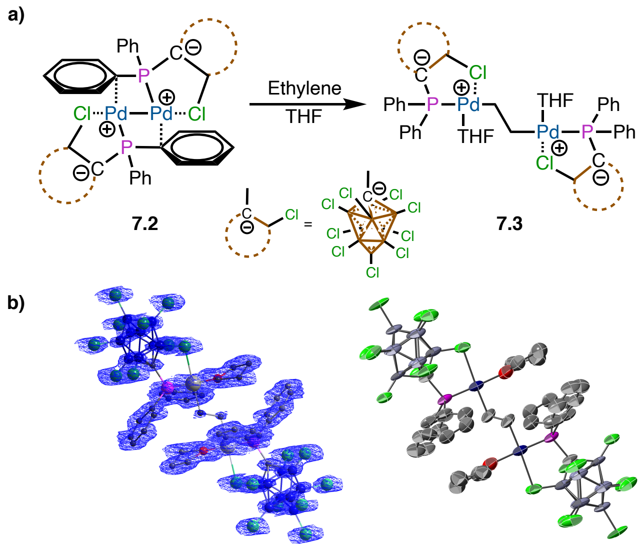
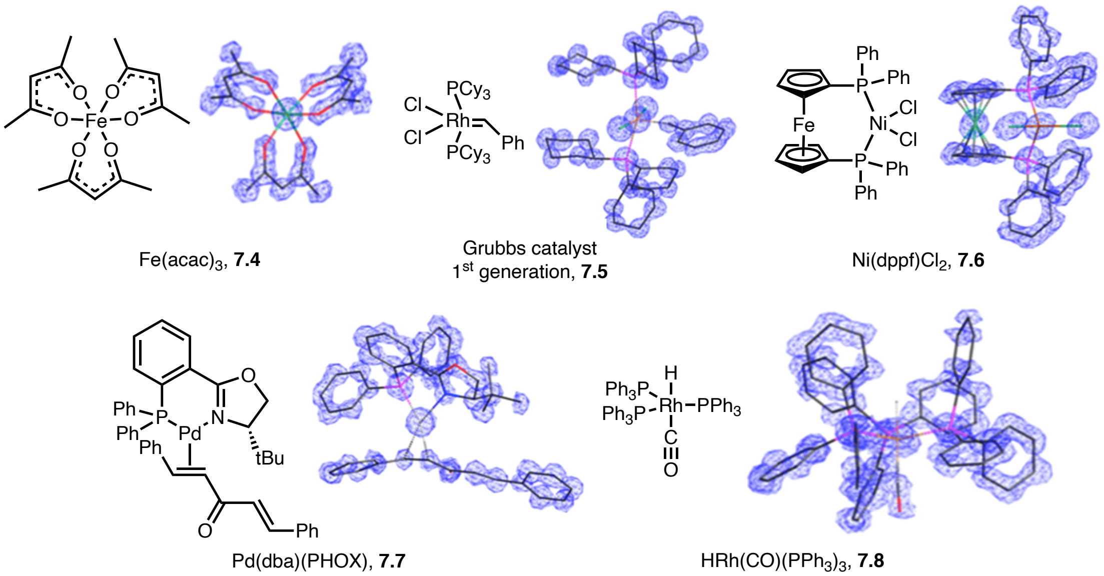


Figure 7.5 MicroED analysis of ethylene insertion product. (a) The reactivity of species **7.2** with ethylene gas forming insertion product **7.3**. (b) Refined MicroED structure with Coulomb potential map overlay and ORTEP structure of product **7.3** (right). Thermal ellipsoids are shown at 30% probability.

7.6 The Scope of Ambient Temperature MicroED for Organometallic Complexes

The two experiments in sections 7.4 and 7.5 demonstrate MicroED's applicability in studying reactive species that are not amenable to conventional structural analysis techniques such as X-ray crystallography and NMR spectroscopy. To establish that MicroED can be used as a general, routine approach for various organometallic complexes, five other commonly used complexes were investigated: tris(acetylacetonato)-iron(III) [Fe(acac)$_3$] **7.4**, benzylidene-bis-(tricyclohexylphosphino)-dichlororuthenium (Grubbs first generation catalyst) **7.5**, [1,1-bis(diphenylphosphino)-ferrocene]dichloronickel(II) **7.6**, Pd(dibenzylideneacetone)-((*S*)-4-*tert*-butyl-2-[2-(diphenylphosphino)phenyl]-2-oxazoline) [Pd(dba)(PHOX)] **7.7**, and carbonyl(hydrido)tris-triphenylphosphane)rhodium(I) [HRh(CO)(PPh$_3$)$_3$] **7.8** (Figure 7.6). The organometallic compounds and transition-metal complexes were loaded as bulk powder onto an EM grid and subjected to ambient temperature MicroED. Electron diffraction data was collected from one or more crystals and merged to obtain a high completeness data set. In all cases, direct methods yielded *ab initio* structures with high resolution (Table 7.1).



Fe(acac)$_3$, **7.4**

Grubbs catalyst 1st generation, **7.5**

Ni(dppf)Cl$_2$, **7.6**

Pd(dba)(PHOX), **7.7**

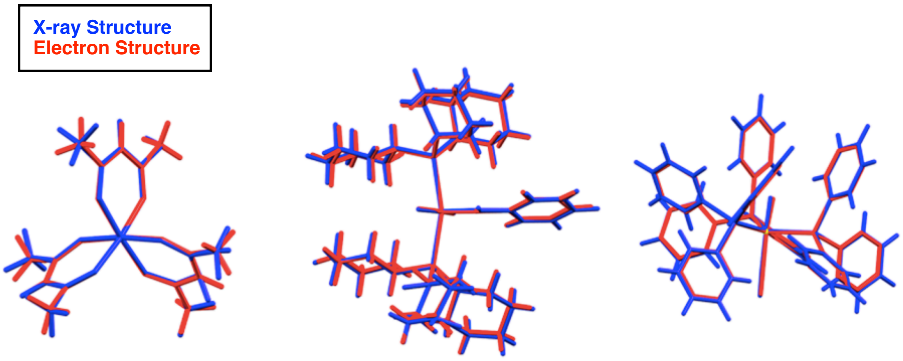
HRh(CO)(PPh$_3$)$_3$, **7.8**

Figure 7.6 Structures of five organometallic complexes obtained using ambient temperature MicroED.

Table 7.1. Data and statistics obtained by ambient temperature MicroED for organometallic complexes.

Entry	R_1 (%)	GooF	Completeness (%)	Resolution (Å)
7.4	16.1	1.826	92.1	0.90
7.5	16.0	1.802	99.3	0.85
7.6	11.3	1.468	86.7	0.90
7.7	14.3	1.249	99.1	1.00
7.8	13.2	1.665	76.7	1.00

To demonstrate MicroED's potential as a general approach for characterizing organometallic species, the electron diffraction structures were compared with published X-ray structures. In most cases, the space groups and unit cell parameters of both types of crystallographic data matched closely. The only exceptions were compounds **7.6** and **7.7**, where MicroED technique revealed crystal structures of new polymorphs.[12,13] To directly visualize the similarities between the X-ray and electron diffraction structures, the MicroED structures of compounds **7.4**, **7.5**, and **7.8** were overlaid with their X-ray counterparts.[14,15] Comparison between the two types of data was made by calculating root-mean-square (RMS) based on deviation of atomic position. For all three cases, the structures visually aligned well, and exhibited root-mean-square (RMS) values of 0.0514, 0.1311, 0.1733 Å and a maximum deviation of 0.1140, 0.2270, and 0.3289 Å for compounds **7.4**, **7.5**, and **7.8**, respectively (Figure 7.7). This consistency was also seen in the position of the hydride in the Rh–H complex **7.5**, where the hydride deviates from the published X-ray data only by 0.164 Å.[16] These small deviation values indicate that even though MicroED structures exhibit poorer statistics such as high R-values and GooF, they provide accurate atomic positions and unambiguous connectivity information comparable to that of X-ray structures. Importantly, these MicroED structures were obtained from bulk powders without additional recrystallizations or cryogenic cooling, demonstrating the ease at which structural solutions can be obtained. Taken together, these results showcase MicroED as a powerful approach for facile and routine structural determination of organometallic compounds that can be used as an alternative to X-ray crystallography.



Fe(acac)$_3$, **7.4**
RMSD: 0.0514
Max D: 0.1140

Grubbs catalyst
1st generation, **7.5**
RMSD: 0.1311
Max D: 0.2270

HRh(CO)(PPh$_3$)$_3$, **7.8**
RMSD: 0.1733
Max D: 0.3289

Figure 7.7 Overlay of ambient temperature ED (red) and previously reported X-ray diffraction (blue) structures for compounds **7.4**, **7.5**, and **7.8** with calculated root-mean-square deviation (Å) and maximum deviation (Å) of atomic coordinates.

7.7 The Scope of Ambient Temperature MicroED for Synthetic Reaction Products

Having demonstrated that MicroED can quickly characterize various reactive organometallic species at ambient temperature and without formal recrystallization efforts, we sought to test MicroED's potential as a routine characterization method for practicing organic chemists as well. In previous studies, various organic molecules were characterized using MicroED to demonstrate its relevance in organic chemistry;[4,5] however, these compounds mainly consisted of commercial materials that have already undergone batch recrystallization processes prior to packaging. To showcase the applicability of MicroED in organic synthesis in real research settings, this technique was used to characterize reaction products from in-house chemistry in the Nelson lab (Figure 7.8). The organic compounds **7.9–7.12** were loaded as seemingly amorphous powders onto an EM grid and subjected to ambient temperature MicroED. Similar to the organometallic complexes, electron diffraction data was collected from microcrystals and merged to obtain a high completeness data set. In all four cases, high resolution structures were obtained using direct methods (Table 7.2). With the exception of **7.9**, all synthetic products were characterized directly after purification by high-performance liquid chromatography (HPLC) without formal crystallization efforts, and all crystal structures were solved after just one or two screening sessions, demonstrating the facile nature of structure determination for these organic molecules using MicroED.

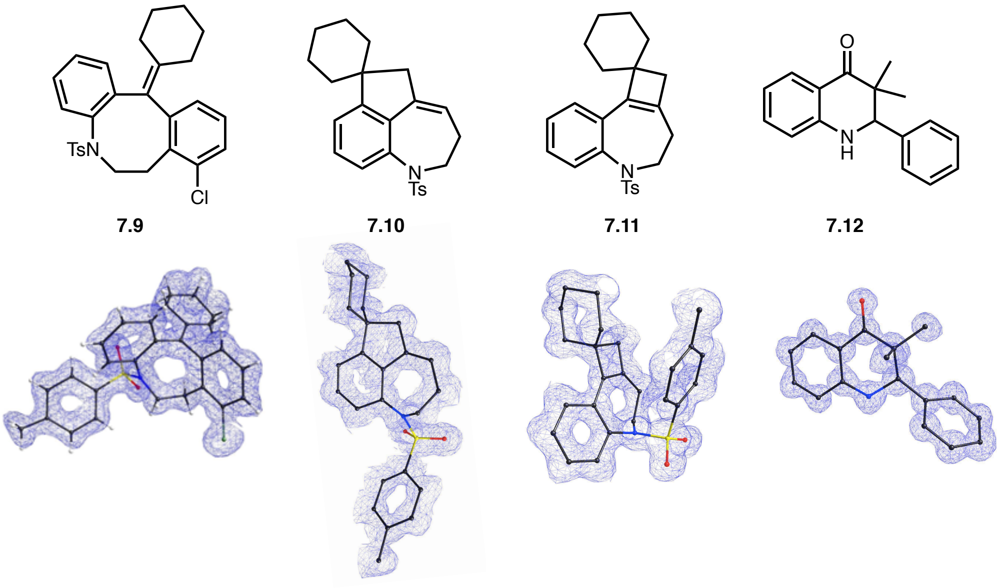
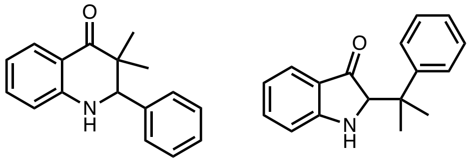


Figure 7.8 Structures of four synthetic reaction products obtained using ambient temperature MicroED. Hydrogens are omitted for clarity.

Table 7.2. Data and statistics obtained by ambient temperature MicroED for synthetic products.

Entry	R_1 (%)	GooF	Completeness (%)	Resolution (Å)
7.9	18.6	2.283	96.3	1.00
7.10	10.4	1.577	90.6	1.00
7.11	13.4	1.905	95.2	1.05
7.12	16.9	1.858	92.2	0.85

Notably, compound **7.12** was tentatively assigned to be indolone **7.13** after NMR analysis, but ambient temperature MicroED revealed its structure to be a quinolone. Compound **7.12** has been previously reported in literature by Long and coworkers.[17] However, due to the discrepancy between the reported and our experimental NMR data, the reported quinolone was synthesized in-house by reproducing the literature procedure. The crude NMR matched our NMR data of compound **7.12**, indicating that the previously reported NMR spectrum was erroneous. This result serves as a powerful example of the importance of unambiguous structure determination. By obtaining a crystal structure, the synthetic reaction product was identified to be quinolone **7.12** without being misled by preceding literature data that may be incorrect.



7.12
Elucidated by MicroED

7.13
Tentatively assigned after NMR analysis

Figure 7.9 Chemical structure of indolone **7.13** that was tentatively assigned after NMR analysis, then elucidated as quinolone **7.12** after MicroED analysis.

7.8 Conclusion

In conclusion, this chapter highlights MicroED as a powerful approach for practical and routine structural characterization of both organometallic complexes and synthetic organic molecules. Using ambient temperature MicroED, structures of several organometallic compounds were elucidated, including the well-known industrial catalyst Schwartz's reagent and a novel Pd(I) dimer whose structure could not be determined using X-ray crystallography or NMR. Additionally, five other organometallic and transition-metal complexes were investigated, and their MicroED structures were compared to published X-ray structures. The comparison between the electron and X-ray structures revealed that MicroED structures provide sufficiently accurate statistical parameters and unambiguous information on connectivity. To demonstrate MicroED's applicability in organic synthesis, four reaction products obtained from in-house chemistry were subjected to ambient temperature MicroED. In the process, the identity of a synthetic product was unambiguously determined despite misleading literature data. Taken together, these results demonstrate MicroED's generalizability across the fields of organic and organometallic chemistry. With the need for additional recrystallizations and cryogenic cooling of the samples eliminated, ambient temperature MicroED is shown to be a facile alternative to X-ray crystallography that can structurally determine various organometallic and organic compounds.

7.9 Notes and References

(1) (a) Deuss, P. J.; Heeten, R.; Laan, W.; Kamer, P. C. J. Bioinspired catalyst design and artificial metalloenzymes. *Chem. Eur. J.* **2011**, *17*, 4680–4698. (b) Gotico, P.; Boitrel, B.; Guillot, R.; Sircoglou, M.; Quaranta, A.; Halime, Z.; Leibl, W.; Aukauloo, A. Second-sphere biomimetic multipoint hydrogen-bonding patterns to boost CO_2 reduction of iron porphyrins. *Angew. Chem. Int. Ed.* **2019**, *58*, 4504–4509. (c) Nagashima, H. Catalyst design of iron complexes. *Bull. Chem. Soc. Jpn.* **2017**, *90*, 761–775.

(2) (a) Diebold, A..; Elbouadili, A.; Hagen, K. S. Crystal structures and solution behavior of paramagnetic divalent transition metal complexes (Fe, Co) of the sterically encumbered tridentate macrocycles 1,4,7-R_3-1,4,7-triazacyclononane: Coordination numbers 5 (R = *i*-Pr) and 6 (R = *i*-Bu). *Inorg. Chem.* **2000**, *39*, 3915–3923. (b) Jamieson, P. B.; Abrahams, S. C.; Bernstein, J. L. Crystal structure of the transition-metal molybdates and tungstates. V. Paramagnetic alpha-$Nd_2(MoO_4)_3$. *J. Chem. Phys.* **1968**, *50*, 86–94.

(3) See Chapter 1.2 for a detailed discussion.

(4) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(5) Gruene, T.; Wennmacher, J. T. C.; Zaubitzer, C.; Holstein, J. J.; Heidler, J.; Fecteau-Lefebvre, A.; De Carlo, S.; Müller, E.; Goldie, K. N.; Regeni, I.; Li, T.; Santiso-Quinones, G.; Steinfeld, G.; Handschin, S.; van Genderen, E.; van Bokhoven, J. A.; Clever, G. H.; Pantelic, R.

Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed.* **2018**, *57*, 16313–16317.

(6) (a) Hart, D. W.; Schwartz, J. Hydrozirconation. Organic synthesis via organozirconium intermediates. Synthesis and rearrangement of alkylzirconium(IV) complexes and their reaction with electrophiles. *J. Am. Chem. Soc.* **1974**, *96*, 8115–8116. (b) Schwartz, J.; Labinger, J. A. Hydrozirconation: A new transition metal reagent for organic synthesis. *Angew. Chem. Int. Ed. Engl.* **1976**, *15*, 333–340. (c) Wieclaw, M. M.; Stecko, S. Hydrozirconation of C = X functionalities with Schwartz's reagent. *Eur. J. Org. Chem.* **2018**, *2018*, 6601–6623. (d) Pinheiro, D. L. J.; de Castro, P. P.; Amarante, G. W. Recent developments and synthetic applications of nucleophilic zirconocene complexes from Schwartz's reagent. *Eur. J. Org. Chem.* **2018**, *2018*, 4828–4844. (e) Kautzner, B.; Wailes, P. C.; Weigold, H. Hydrides of bis(cyclopentadienyl)zirconium. *J. Chem. Soc. D* **1969**, 1105a. (f) Wailes, P. C.; Weigold, H. Hydrido complexes of zirconium I. Preparation. *J. Organomet. Chem.* **1970**, *24*, 405–411.

(7) Takahashi, T.; Suzuki, N.; Jayasuriya, N.; Wipf, P. Chlorobis-(cyclopentadienyl)hydridozirconium. *Encyclopedia of Reagents for Organic Synthesis* **2006.**

(8) Rossini, A. J.; Mills, R. W.; Briscoe, G. A.; Norton, E. L.; Geier, S. J.; Hung, I.; Zheng, S.; Autschbach, J.; Schurko, R. W. Solid-state chlorine NMR of group IV transition metal organometallic complexes. *J. Am. Chem. Soc.* **2009**, *131*, 3317–3330.

(9) Harlan, C. J.; Bott, S. G.; Barron, A. R. Methyl-hydride metathesis between [Zr(cp)$_2$Me$_2$] and [HAl(μ_3-NBu^t)]$_4$: Molecular structures of [Me$_{1-x}$HxAl(μ_3-NBu^t)]$_4$ (x = 0, 0.78 or 1) and [(cp)$_2$ZrMe(μ-H)]$_2$ (cp = η^5-C$_5$H$_5$). *J. Chem. Soc., Dalton Trans.* **1997**, 637–642.

(10) Fafard, C. M.; Adhikari, D.; Foxman, B. M.; Mindiola, D. J.; Ozerov, O. V. Addition of ammonia, water, and dihydrogen across a single Pd–Pd bond. *J. Am. Chem. Soc.* **2007**, *129*, 10318–10319.

(11) Huacuja, R.; Graham, D. J.; Fafard, C. M.; Chen, C.-H.; Foxman, B. M.; Herbert, D. E.; Alliger, G.; Thomas, C. M.; Ozerov, O. V. Reactivity of a Pd(I)–Pd(I) dimer with O_2: Monohapto Pd superoxide and dipalladiumperoxide in equilibrium. *J. Am. Chem. Soc.* **2011**, *133*, 3820–3823.

(12) Weng, S.-S.; Ke, C.-S.; Chen, F.-K.; Lyu, Y.-F.; Lin, G.-Y. Transesterificationi catalyzed by iron(III) *β*-diketonate species. *Tetrahedron* **2011**, *67*, 1640–1648.

(13) Sherden, N. H.; Behenna, D. C.; Virgil, S. C.; Stoltz, B. M. Unusual allylpalladium carboxylate complexes: Identification of the resting state of catalytic enantioselective decarboxylative allylic alkylation reactions of ketones. *Angew. Chem., Int. Ed.* **2009**, *48*, 6840–6843.

(14) Trnka, T. M. Catalyst for olefin metathesis: Ruthenium alkylidene complexes with phosphine and *N*-heterocyclic ligands. Ph.D. Dissertation, California Institute of Technology, Pasadena, CA, 2002.

(15) La Placa, S. J.; Ibers, J. A. Crystal and molecular structure of tritriphenylphosphine rhodium carbonyl hydride. *Acta Cryst.* **1965**, *18*, 511–519.

(16) Babra, I. S.; Morley, L. S.; Nyburg, S. C.; Parkins, A. W. The crystal and molecular structure of a new polymorph of carbonylhydridotris(triphenylphosphine)rhodium(I) having a Rh-H stretching absorption at 2013 cm^{-1}. *J. Crystallogr. Spectrosc. Res.* **1993**, *23*, 997–1000.

(17) Hu, W.; Lin, J.-P.; Song, L.-R.; Long, Y.-Q. Direct synthesis of 2-aryl-4-quinolones via transition-metal-free intramolecular oxidative C(sp^3)–H/C(sp^3)–H coupling. *Org. Lett.* **2015**, *17*, 1268–1271.

7.10 Experimental Section

7.10.1 Materials and Methods

7.10.1.1 General Considerations for Organometallic Complexes

This section pertains to electron diffraction studies of organometallic complexes (section 7.4–7.6). Grubbs Catalyst® 1st Generation and Schwartz's reagent were purchased from Sigma-Aldrich. Carbonyl(hydrido)tris(triphenylphosphane)rhodium(I) was purchased from Strem Chemicals. All commercial samples were used as received from the supplier. Pd(dba)(PHOX), Fe(acac)$_3$, and Ni(dppf)Cl$_2$ were prepared according to the reported literature procedures.[1–3] Synthetic samples were precipitated from solution and used as prepared with no formal recrystallization. Unless otherwise stated, all manipulations were carried out using standard Schlenk or glovebox techniques (O_2, H_2O < 1ppm) under a dinitrogen or argon atmosphere. Solvents were dried on K or CaH_2 and distilled under argon before use. (κ^2-P,Cl-$PPh_2CB_9Cl_9$)PdMe(THF) was prepared according to the literature procedure.[4] Solution state NMR spectra were recorded at room temperature on Bruker Avance 300MHz, Bruker Avance 400MHz, or Bruker Avance 600MHz spectrometers. NMR chemical shifts are reported in parts per million (ppm). ^{1}H NMR and ^{13}C NMR chemical shifts were referenced to residual protio solvent. ^{11}B NMR chemical shifts were externally referenced to BF_3OEt_2. ^{31}P NMR chemical shifts were externally referenced to 80% H_3PO_4 in H_2O. The $^{13}C\{^1H\}$ Cross Polarization Magic Angle Spinning (CPMAS) NMR spectrum of compound **7.3** was acquired on a Bruker NEO600 operating at an ^{1}H frequency of 600 MHz. The analyte was packed in a 4 mm zirconia rotor packed in an Argon filled glovebox. The spinning speed was 10 kHz at the magic angle, and contact time was 2 ms for both $^{13}C\{^1H\}$ and $^{31}P\{^1H\}$ CPMAS NMR spectra. $^{13}C\{^1H\}$ spectra were referenced to external adamantane, $^{31}P\{^1H\}$ spectra were referenced to external H_3PO_4, and $^{11}B\{^1H\}$ spectra were referenced to $NaBH_4$. FT-IR spectra were recorded as pressed pellets using a Bruker Alpha IR spectrometer in an argon-filled glovebox.

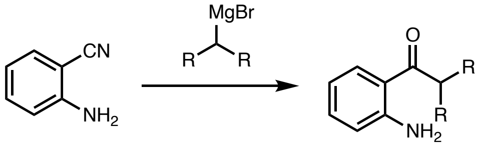
7.10.1.2 General Considerations for Synthetic Reaction Products

This section pertains to electron diffraction studies of synthetic reaction products (section 7.7).

Unless otherwise stated, all reactions were performed in a VAC glovebox under nitrogen atmosphere with ≤ 0.5 ppm O_2 levels. All glassware and stir-bars were dried in a 160 °C oven for at least 12 hours and dried *in vacuo* before use. All liquid substrates were rigorously dried before use. Ethyl ether, tetrahydrofuran, dichloromethane, dimethylformamide, toluene, 1,2-dichloroebenzene, and hexanes were degassed and dried in a JC Meyer solvent system. Acetonitrile, triethylamine, and pyridine were distilled over calcium hydride. Solid substrates were dried over P_2O_5. $[Li]^+[B(C_6F_5)_4]^-$ salts were synthesized according to literature procedure.[5] Preparatory thin layer chromatography (TLC) was performed using Millipore silica gel 60 F_{254} pre-coated plates (0.25 mm) and visualized by UV fluorescence quenching. SiliaFlash P60 silica gel (230-400 mesh) was used for flash chromatography. NMR spectra were recorded on a Bruker AV-300 (^1H, ^{13}C, ^{19}F), AV-400 (^1H, ^{13}C, ^{19}F), Bruker DRX-500 (^1H, ^{13}C), and Bruker AV-500 (^1H, ^{13}C). ^1H NMR spectra are reported relative to $CDCl_3$ (7.26 ppm) unless noted otherwise. Data for ^1H NMR spectra are as follows: chemical shift (ppm), multiplicity, coupling constant (Hz), integration. Multiplicities are as follows: s = singlet, d = doublet, t = triplet, dd = doublet of doublet, dt = doublet of triplet, ddd = doublet of doublet of doublet, td = triplet of doublet, tt = triplet of triplet, quint = quintet, sept = septet, m = multiplet. ^{13}C NMR spectra are reported relative to $CDCl_3$ (77.0 ppm) unless noted otherwise. GC spectra were recorded on an Agilent 6850 series GC using an Agilent HP-1 (50 m, 0.32 mm ID, 0.25 µm DF) column. GCMS spectra were recorded on a Shimadzu GCMS-QP2010 using a Restek XTI-5 (50 m, 0.25 mm ID, 0.25 µm DF) column interface at room temperature. IR Spectra were record on a Perkin Elmer 100 spectrometer and are reported in terms of frequency absorption (cm^{-1}). High resolution mass spectra (HR-MS) were recorded on a Waters (Micromass) GCT Premier spectrometer, a Waters (Micromass) LCT Premier, an Agilent GC EI-MS, and are reported as follows: m/z (% relative intensity). Purification by preparative HPLC was done on an Agilent 1200 series instrument with a reverse phase Alltima C_{18} (5m, 25 cm length, 1 cm internal diameter) column.

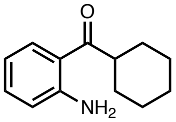
7.10.2 Synthetic Procedures

*Synthetic procedure for compounds **7.2** and **7.3** are reported in the adapted article.*



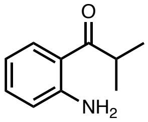
7.10.2.1 General Procedure for Grignard Reaction of Aryl Nitriles

Magnesium (3.0 equiv) was added to a flame-dried three-neck flask equipped with a condenser, followed by the addition of THF to generate a 1 M solution for the alkyl bromide. Alkyl bromide (1.0 equiv) was added slowly into the flask to form the Grignard reagent, keeping the solution under gentle reflux. The solution was then cooled down to 0 °C and 1 M solution of 2-aminobenzonitrile (3.0 equiv) in THF was added dropwise. The reaction was run overnight and quenched with water and concentrated hydrochloric acid until pH of 1. The reaction mixture was extracted with ethyl ether three times, and the combined organic phase was washed with saturated sodium bicarbonate solution and brine. Finally, the crude solution was dried with magnesium sulfate, filtered, and concentrated before purification *via* flash column chromatography.



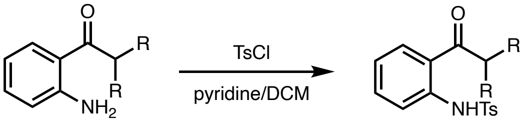
(2-aminophenyl)(cyclohexyl)methanone (7.14).

Synthesized according to general procedure 7.10.2.1 starting from 25.4 mmol of 2-aminobenzonitrile. Crude product was purified via flash column chromatography using 20% ethyl ether in hexanes to give the product as a yellow solid (2.51 g, 48.6% yield). Spectral data matched those reported in the literature.[6]



1-(2-aminophenyl)-2-methylpropan-1-one (7.15).

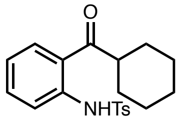
Synthesized according to general procedure 7.10.2.1 starting from 42.3 mmol of 2-aminobenzonitrile. Crude product was purified via flash column chromatography using 10% ethyl ether in hexanes to give the product as a yellow solid (5.45 g, 78.9% yield). Spectral data matched those reported in the literature.[7]



7.10.2.2 General Procedure for N-Tosylation of Anilines

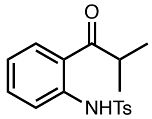
To a flame dried roundbottom flask was added aniline (1.0 equiv) followed by DCM (13.0 equiv) and pyridine (7.0 equiv). This was cooled to 0 °C and then tosyl chloride (1.42 equiv) was added.

The reaction was warmed up to room temperature and stirred for 16 hours. The reaction was diluted with additional DCM (~15 equiv) and water. The layers were separated and the aqueous later was extracted twice more with DCM. The combined organics were washed 1 M aqueous HCl, water and brine in that order and the dried over magnesium sulfate, filtered and concentrated. The crude product was purified by flash column chromatography to give pure material as a white solid.



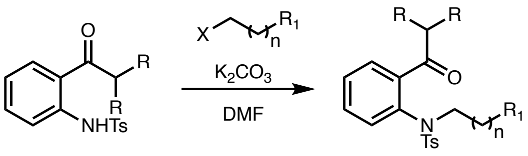
N-(2-(cyclohexanecarbonyl)phenyl)-4-methylbenzenesulfonamide (7.16).

Synthesized according to general procedure 7.10.2.2 starting from 73.3 mmol of the corresponding aniline **7.14**. Crude product was purified via flash column chromatography using 20% ethyl acetate in hexanes to give sulfonamide **7.16** as a white solid (26.2 g, 80% yield). Spectral data matched those reported in the literature.[7]



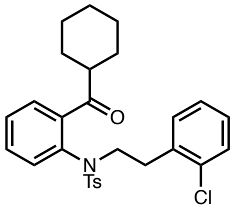
N-(2-isobutyrylphenyl)-4-methylbenzenesulfonamide (7.17).

Synthesized according to general procedure 7.10.2.2 starting from 24.5 mmol of the corresponding aniline **7.15**. Crude product was purified via flash column chromatography using 30% ethyl acetate in hexanes to give sulfonamide **7.17** as a white solid (5.15 g, 66% yield). Spectral data matched those reported in the literature.[7]



7.10.2.3 General Procedure for N-Alkylation of Sulfonamides

To an oven dried 20 mL scintillation vial was added sulfonamide (1.0 equiv) followed by DMF (to yield a 1 M solution). To the solution was added and potassium carbonate (2.0 equiv) and alkyl iodide (2.0 equiv) under a stream of N_2. The vial was sealed and heated to 100 °C for 24h. The reaction mixture was cooled to rt, diluted with water and ether. The layers were separated and the aqueous layer was extracted with ether (3x). The combined organics were washed with water (3x) and brine (1x) then dried over $MgSO_4$, filtered and concentrated to give crude product. The crude product was purified by flash column chromatography.



N-(2-chlorophenethyl)-N-(2-(cyclohexanecarbonyl)phenyl)-4-methylbenzenesulfonamide (7.18).

Synthesized according to general procedure 7.10.2.3 starting from 1.50 mmol of the corresponding sulfonamide **7.16** and 6.00 mmol of 1-(2-bromoethyl)-2-chlorobenzene. Crude product was purified via flash column chromatography using 25% ethyl ether in hexanes to give sulfonamide **7.18** as a white powder (0.50 g, 67% yield).

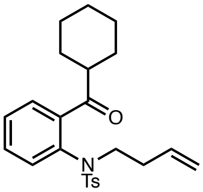
*NMR had poor resolution at room temperature, so ^{13}C NMR is reported below at 50 °C.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.59 (dd, J = 7.6, 1.8 Hz, 1H), 7.48 (d, J = 8.3 Hz, 2H), 7.39 (ddd, J = 7.5, 7.5, 1.3 Hz, 1H), 7.33 (ddd, J = 7.8, 7.8, 1.8 Hz, 1H), 7.29 (m, 1H), 7.23 (d, J = 8.0 Hz, 2H), 7.17 (m, 3H), 6.78 (dd, J = 7.9, 1.2 Hz, 1H), 3.95 (m, 1H), 3.49 (m, 1H), 3.42 (tt, J = 10.8, 3.4 Hz, 1H), 3.10 (m, 1H), 2.94 (m, 1H), 2.41 (s, 3H), 2.11 (m, 1H), 1.94-1.56 (m, 5H), 1.46-1.22 (m, 4H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 206.8, 143.6, 141.8, 136.8, 136.0, 135.5, 134.2, 131.0, 130.6, 129.5, 129.4, 129.3, 128.2, 128.1, 128.0, 127.0, 51.3, 49.3, 32.8, 29.2, 26.0, 25.9, 21.4.

FTIR (Neat film NaCl): 3065, 2928, 2853, 1690, 1596, 1444, 1351, 1159, 1092, 580.

HR-MS (ESI-MS) m/z: [M+Na]+ Calc'd for $C_{28}H_{30}ClNO_3SNa$ 518.1533; Found 518.1528.

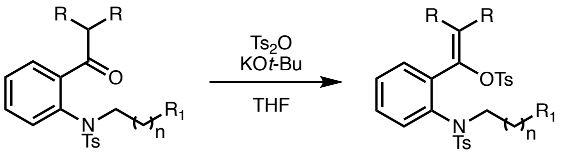


N-(but-3-en-1-yl)-N-(2-(cyclohexanecarbonyl)phenyl)-4-methylbenzenesulfonamide (7.19).

Synthesized according to general procedure 7.10.2.3 starting from 3.00 mmol of the corresponding sulfonamide **7.16** and 6.00 mmol of 4-bromobut-1-ene. Crude product was purified via flash column chromatography using 25% ethyl ether in hexanes to give sulfonamide **7.19** (1.12 g, 91% yield).

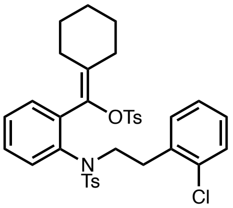
^{1}H NMR (400 MHz, CDCl$_3$) δ 7.55 (dd, J = 7.5, 1.9 Hz, 1H), 7.50 (dt, J = 8.4, 1.9 Hz, 2H), 7.34 (dtd, J = 22.3, 7.4, 1.3 Hz, 2H), 7.26 (dt, J = 8.0, 0.5 Hz, 2H), 6.74 (dd, J = 7.8, 1.4 Hz, 1H), 5.66 (ddt, J = 17.1, 10.4, 6.7 Hz, 1H), 5.02 (dq, J = 8.2, 1.5 Hz, 1H), 4.98 (t, J = 1.5 Hz, 1H), 3.81 (s, 1H), 3.36 (tt, J = 11.3, 3.3 Hz, 2H), 2.43 (s, 3H), 2.21 (s, 1H), 2.05 (s, 1H), 1.81 (s, 2H), 1.69 (d, J = 11.6 Hz, 1H), 1.58 – 1.21 (m, 5H).

HR-MS (ESI-MS) m/z: [M+H]+ Calc'd for $C_{24}H_{30}NO_3S$ 412.1946; Found 412.1960.



7.10.2.4 General Procedure for Tosylation of Ketones

The corresponding ketone (1 equiv) was dissolved in THF (0.33 M solution) and cooled to 0 °C. To this was added a solution of potassium *tert*-butoxide (1.5 equiv) in THF (1.0 M solution). This was stirred 1.5 hours and then tosic anhydride (1.5 equiv) was added and the reaction was warmed up to rt. After 4 hours, the reaction mixture (generally a thick slurry) was diluted with ethyl acetate. This was washed with water (x1) and brine (x1) then dried over $MgSO_4$, filtered and concentrated to give crude vinyl tosylate. This was purified by flash column chromatography to give pure vinyl tosylate.



(2-((N-(2-chlorophenethyl)-4-methylphenyl)sulfonamido)phenyl) (cyclohexylidene)methyl-4-methylbenzenesulfonate (7.20).

Synthesized according to general procedure 7.10.2.4 starting from 2.96 mmol of the corresponding ketone **7.18**. Crude product was purified via flash column chromatography using benzene to give vinyl tosylate **7.20** as a white solid (0.66 g, 34% yield).

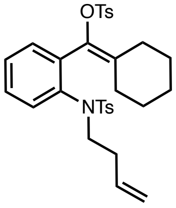
*NMR had poor resolution at room temperature, so ^{1}H NMR is reported below at 70 °C
and ^{13}C NMR is reported below at 50 °C.

^{1}H NMR (500 MHz, CDCl$_3$, 70 °C) δ 7.71 (d, J = 7.8 Hz, 2H), 7.57 (d, J = 7.8 Hz, 2H), 7.44 (br s, 1H), 7.31 (m, 2H), 7.25 (m, 3H), 7.09 (m, 2H), 7.04 (d, J = 7.8 Hz, 3H), 6.97 (m, 1H), 3.60 (m, 1H), 3.44 (m, 1H), 2.79 (m, 1H), 2.59 (m, 1H), 2.49 (m, 1H), 2.43 (s, 3H), 2.32 (m, 1H), 2.26 (s, 3H), 2.10 (m, 2H), 1.68 (br s, 2H), 1.59 (br s, 2H), 1.51 (br s, 2H).

^{13}C NMR (126 MHz, CDCl$_3$) δ 144.4, 143.4, 139.0, 136.8, 136.39, 136.36, 135.6, 134.5, 134.1, 133.6, 130.9, 129.4, 129.3, 128.4, 128.0, 127.79, 127.76, 126.8, 50.8, 32.3, 30.6, 28.7, 27.0, 26.7, 26.2, 21.4, 21.3.

FTIR (Neat film NaCl): 3066, 2973, 2928, 2855, 1597, 1475, 1444, 1356, 1176, 1160, 656, 571, 552.

HR-MS (ESI-MS) m/z: [M+Na]+ Calc'd for $C_{35}H_{36}ClNO_5S_2Na$ 672.1621; Found 672.1607.



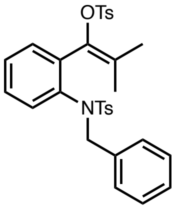
(2-((N-(but-3-en-1-yl)-4-methylphenyl)sulfonamido)phenyl)(cyclohexylidene)methyl 4-methylbenzenesulfonate (7.21).

Synthesized according to general procedure 7.10.2.4 starting from 2.72 mmol of the corresponding ketone **7.19**. Crude product was purified via flash column chromatography using benzene to give vinyl tosylate **7.21** (0.73 g, 47% yield).

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65 (d, J = 7.9 Hz, 2H), 7.56 (d, J = 8.1 Hz, 2H), 7.40 (s, 1H), 7.27 (d, J = 7.6 Hz, 4H), 7.15 (d, J = 8.1 Hz, 2H), 6.91 (s, 1H), 5.44 (ddt, J = 17.1, 10.4, 6.7 Hz, 1H), 4.93 – 4.78 (m, 2H), 3.30 (dddd, J = 55.0, 13.8, 11.4, 5.3 Hz, 2H), 2.43 (s, 3H), 2.40 (s, 1H), 2.35 (s, 3H), 2.27 (ddd, J = 13.0, 8.4, 4.3 Hz, 1H), 2.14 – 1.85 (m, 4H), 1.68 – 1.41 (m, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 144.5, 143.5, 138.9, 136.7, 136.5, 135.6, 134.7, 134.2, 133.6, 129.4, 129.4, 129.4, 128.2, 128.1, 127.8, 116.5, 77.4, 77.0, 76.7, 50.7, 32.2, 30.5, 28.6, 27.0, 26.7, 26.2, 21.6.

HR-MS (ESI-MS) m/z: $[M+NH_4]+$ Calc'd for $C_{31}H_{39}N_2O_5S_2$ 583.2300; Found 583.2321.



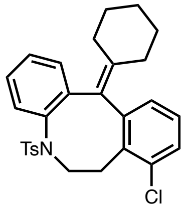
1-(2-((N-benzyl-4-methylphenyl)sulfonamido)phenyl)-2-methylprop-1-en-1-yl 4-methylbenzenesulfonate (7.22).

N-benzyl-N-(2-isobutyrylphenyl)-4-methylbenzenesulfonamide (2.7 mmol) was dissolved in THF (0.27 M solution) and cooled to –20 °C. To this was added a solution of potassium *tert*-butoxide (1.5 equiv) in THF (0.9 M solution) dropwise, turning the colorless reaction solution to yellow. The reaction was warmed to 0 °C for 2 hours and cooled back down to –20 °C before tosic anhydride (1.5 equiv) was added. This was warmed to rt to form a milky white precipitate. The reaction mixture was diluted with ethyl acetate, washed with water (x1) and brine (x1), then dried over $MgSO_4$ before filtering and concentrating to give crude vinyl tosylate **7.22**. This was initially purified by flash column chromatography using 2% ether in DCM then by a second flash column chromatography using a gradient from 20% ether in hexanes to 30% ether in hexanes to afford vinyl tosylate **7.22** as a white solid (1.11 g, 73% yield).

^{1}H NMR 500 MHz, CDCl3) δ 7.60 (t, J = 9.6 Hz, 3H), 7.34 (dt, J = 16.2, 8.4 Hz, 2H), 7.24 – 7.14 (m, 2H), 7.11 (dd, J = 7.7, 1.7 Hz, 2H), 6.92 – 6.87 (m, 2H), 4.75 (d, J = 14.8 Hz, 2H), 4.36 (d, J = 14.8 Hz, 2H), 2.39 (d, J = 13.0 Hz, 4H), 1.74 (s, 3H), 1.55 (s, 6H), 1.46 (s, 3H).

FTIR (Neat film NaCl): 3065, 3034, 2924, 1597, 1446, 1364, 1177, 1169, 1092, 811, 727, 702, 658, 561.

HR-MS (ESI-MS) m/z: [M+Na]+ Calc'd for $C_{31}H_{31}NO_5S_2Na$ 584.1541; Found 584.1532.



8-chloro-12-cyclohexylidene-5-tosyl-5,6,7,12-tetrahydrodibenzo[b,e]azocine (7.9).

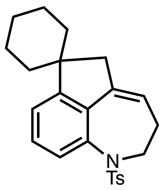
In the glovebox, lithium tetrakis(pentafluorophenyl)borate (0.1 equiv), lithium hydride (5.0 equiv), and vinyl tosylate **7.20** (0.231 mmol, 1.0 equiv) were dissolved into 1,2-dichlorobenzene to generate a 0.0143 M solution for the vinyl tosylate. The reaction was heated under 140 °C overnight. The reaction solution was directly purified via flash column chromatography using hexanes to get rid of 1,2-dichlorobenzene and then ethyl acetate to flush the remaining products off the column. Then the crude product was purified via flash column chromatography using 10% ethyl ether in hexanes to give the product **7.9** as a white powder (0.080 g, 73% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85 (d, J = 7.8 Hz, 2H), 7.36 (d, J = 7.9 Hz, 2H), 7.25 (m, 2H), 7.18 (m, 1H), 7.08 (m, 3H), 6.90 (d, J = 8.0 Hz, 1H), 4.31 (dd, J = 15.3, 6.9 Hz, 1H), 3.56 (dd, J = 14.9, 6.9 Hz, 1H), 3.36 (dd, J = 15.1, 9.4 Hz, 1H), 2.80 (dd, J = 15.3, 9.4 Hz, 1H), 2.46 (s, 3H), 2.21 (m, 2H), 2.06 (m, 2H), 1.87 (m, 1H), 1.57 (m, 5H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 146.1, 145.4, 143.2, 140.5, 139.4, 139.2, 136.7, 133.9, 130.1, 130.0, 129.8, 128.7, 128.1, 127.8, 127.7, 127.6, 127.4, 127.3, 50.3, 33.7, 31.8, 31.5, 28.1, 27.8, 26.5, 21.6.

FTIR (Neat film NaCl): 3062, 2925, 2852, 1560, 1482, 1446, 1349, 1158, 1092, 569.

HR-MS (ESI-MS) m/z: [M+Na]+ Calc'd for $C_{28}H_{28}ClNO_2SNa$ 500.1427; Found 500.1436.

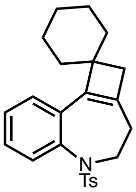


1'-tosyl-1',2',3',5'-tetrahydrospiro[cyclohexane-1,6'-indeno[7,1-bc]azepine] (7.10).

In the glovebox, lithium tetrakis(pentafluorophenyl)borate (0.1 equiv), lithium hydride (5.0 equiv), and vinyl tosylate **7.21** (1.0 equiv) were dissolved into 1,2-dichlorobenzene to generate a 0.0143 M solution for the vinyl tosylate. The reaction was heated under 120 °C overnight. The reaction solution was directly purified via flash column chromatography using 5% ethyl acetate in hexanes followed by preparative HPLC (90:10 MeCN:H_2O) to get multiple products (four peaks) that included **7.10** (21% yield).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.45 – 7.37 (m, 3H), 7.19 (t, J = 7.7 Hz, 1H), 7.10 (ddd, J = 11.5, 8.1, 0.9 Hz, 3H), 5.54 (tt, J = 4.4, 2.0 Hz, 1H), 3.80 (t, J = 5.5 Hz, 2H), 2.54 (dh, J = 7.3, 2.4 Hz, 2H), 2.47 (q, J = 2.3 Hz, 2H), 2.35 (s, 3H), 1.69 (ddt, J = 18.4, 13.1, 3.3 Hz, 3H), 1.59 – 1.23 (m, 7H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 157.64, 142.99, 138.33, 138.00, 137.00, 134.37, 129.11, 128.01, 126.98, 124.58, 121.35, 120.91, 77.29, 77.03, 76.78, 47.95, 46.02, 44.35, 37.80, 31.21, 25.78, 23.40, 21.49.

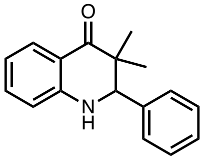


5-tosyl-2,3,4,5-tetrahydrospiro[benzo[b]cyclobuta[d]azepine-1,1'-cyclohexane] (7.11).

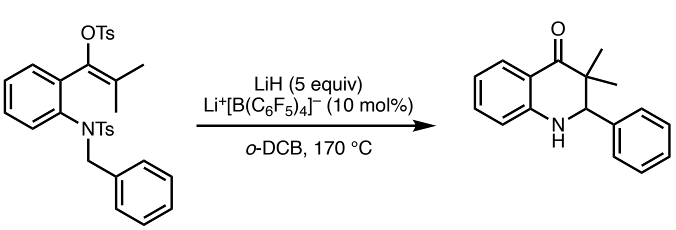
In the glovebox, lithium tetrakis(pentafluorophenyl)borate (0.1 equiv), lithium hydride (5.0 equiv), and vinyl tosylate **7.21** (1.0 equiv) were dissolved into 1,2-dichlorobenzene to generate a 0.0143 M solution for the vinyl tosylate. The reaction was heated under 120 °C overnight. The reaction solution was directly purified via flash column chromatography using 5% ethyl acetate in hexanes followed by preparative HPLC (90:10 MeCN:H_2O) to get multiple products (four peaks) that included **7.11** (21% yield).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.70 – 7.65 (m, 1H), 7.31 – 7.27 (m, 2H), 7.19 (tq, J = 4.8, 2.2 Hz, 3H), 7.09 – 7.02 (m, 2H), 4.55 (dd, J = 14.5, 5.6 Hz, 1H), 3.18 (td, J = 13.7, 3.3 Hz, 1H), 2.83 (td, J = 13.1, 7.0 Hz, 1H), 2.30 (s, 3H), 2.11 – 1.75 (m, 3H), 1.69 – 1.15 (m, 9H), 0.59 (d, J = 13.1 Hz, 1H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 144.41, 143.74, 142.85, 138.89, 137.88, 131.88, 129.93, 128.96, 127.09, 126.80, 126.42, 126.30, 77.28, 77.03, 76.77, 48.76, 46.94, 42.20, 36.01, 32.93, 31.97, 25.72, 24.23, 23.76, 21.29.



3,3-dimethyl-2-phenyl-2,3-dihydroquinolin-4(1H)-one (7.12).

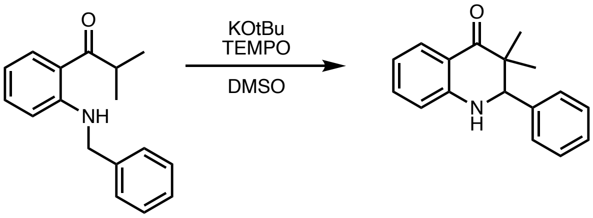


Procedure A: In the glovebox, lithium tetrakis(pentafluorophenyl)borate (0.1 equiv), lithium hydride (5.0 equiv), and vinyl tosylate **7.22** (0.15 mmol, 1.0 equiv) were dissolved into 1,2-dichlorobenzene to generate a 0.0143 M solution for the vinyl tosylate. The reaction was heated under 170 °C for 36 h. The reaction solution was then purified *via* flash column chromatography, first with 1% ether/hexanes then 1:1 ether/hexanes to remove 1,2-dichlorobenzene. 20% ether/hexanes column chromatography was used to isolate one of the products, and the rest of the material was purified by preparative HPLC (90:10 MeCN:H_2O). Isolated compound **7.12** was a minor product obtained as a white solid (19.6:11.4:35.9:21.2 ~2:1:3.6:2 ratio of products).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (dd, J = 8.0, 1.6 Hz, 1H), 7.50 – 7.41 (m, 2H), 7.42 – 7.35 (m, 3H), 7.33 (ddd, J = 8.4, 7.1, 1.6 Hz, 1H), 6.80 (ddd, J = 8.1, 7.0, 1.0 Hz, 1H), 6.70 (dd, J = 8.2, 1.0 Hz, 1H), 4.51 (s, 1H), 4.45 (s, 1H), 1.08 (s, 3H), 1.02 (s, 3H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 199.17, 151.17, 138.49, 135.18, 128.97, 128.82, 128.69, 118.79, 117.89, 115.94, 77.67, 77.41, 77.16, 67.45, 46.25, 21.20, 18.33.

GC-MS m/z: Calc'd for $C_{17}H_{17}NO$ 251.1; Found 251.1.



Procedure B: The literature procedure[8] was reproduced using (0.50 mmol), potassium *tert*-butoxide (2 equiv), and TEMPO (2 equiv) in DMSO (50 mM solution). **7.12** was obtained after purification *via* 10% ether/hexanes flash column chromatography to yield a white solid (67 mg, 53% yield). NMR matched the obtained experimental NMR data from procedure A.

7.10.3 Sample Preparation for Electron Diffraction Studies

Samples were prepped using Quantifoil R1/2 Cu200 mesh grids. For preparation of Fe(acac)$_3$ **7.4**, Ni(dppf)Cl$_2$ **7.6**, Pd(dba)(PHOX) **7.7**, and synthetic reaction product samples **7.10–12**, grids were placed in a dram vial with ~1 mg of compound and shaken lightly. The samples were tapped against the surface of a filter paper to remove residual compound and the TEM grid was subsequently transferred to a clean vial. Schwartz's reagent **7.1**, compound **7.3**, Grubbs Catalyst **7.5** and carbonyl(hydrido)tris (triphenylphosphane)rhodium(I) **7.8** samples were prepared in a similar fashion within a glovebox under N$_2$ as received from the supplier. All samples were transferred from sealed vials to the sample holder under ambient conditions with no additional experimental setup.

7.9 was recrystallized by slow evaporation of **7.9** solution in acetonitrile at ambient temperature. Following recrystallization, the resulting solids were loaded on a TEM grid as described above.

7.10.4 Instrument Parameters

All data for compounds **7.1–7.8** was collected on a FEI Tecnai TF-30 electron microscope operating at ambient temperature with an operating voltage of 300 keV, corresponding to a wavelength of ~0.0196 Å using a single-tilt sample holder. TEM grids were screened by operating the microscope in over focused diffraction mode.

All data for compounds **7.9–7.12** was collected on a FEI Talos F200C electron microscope at ambient temperature with an operating voltage of 200 keV, corresponding to a wavelength of ~0.0251 Å using a single-tilt sample holder.

7.10.5 Data Collection Procedure

Diffraction data for **7.1–7.8** was collected using rolling shutter mode with a TVIPS TemCam-XF416 CMOS 4k x 4k camera. Diffraction data for **7.9–7.12** was collected using rolling shutter mode with a Ceta-D CMOS 4k x 4k camera. Images were collected as movies by continuous rotation of crystals under a parallel electron beam using a constant tilt rate of ~0.3 deg s^{-1} over an angular wedge of ~50° between the minimum and maximum tilt ranges of –72° to +72° respectively.[9] During data collection, the camera integrated continuously at a rate of 3 s per frame. Crystals were isolated using a selected area aperture to reduce background noise and calibrated to eucentric height to remain within the aperture during continuous rotation over the tilt range.

Diffraction movies for **7.1–7.8** were saved as TVIPS files and were converted to SMV format using open source software freely available online (https://cryoem.ucla.edu/pages/MicroED). Diffraction movies for **7.9–7.12** were saved as SER files and were converted to SMV format as previously reported.[10] Frames were indexed and integrated in XDS and multiple datasets were scaled and merged using XSCALE.[11–12] Intensities were converted to SHELX format using XDSCONV.[12]

7.10.6 Diffraction Data and Statistics

*Diffraction data procedure and statistics for compound 7.**2** are reported in the adapted article.*

Electron diffraction data of compounds **7.1**, **7.3–7.12** were collected and treated in the above manner. All structures were solved *ab initio* using direct methods in SHELXT and refined with SHELXL using ShelXle.[13–15] All non-hydrogen atoms were refined anisotropically and all hydrogen atoms, aside from the metal hydrides in compound **7.1** and **7.8**, were refined using the riding model.

Crystallographic information files (CIF) for compounds **7.1–7.8** have been deposited at the Cambridge Crystallographic Data Center and are available free of charge under reference numbers 1908168–1908175 at:

http://www.ccdc.cam.ac.uk/Community/Requestastructure/Pages/DataRequest.aspx.

7.10.6.1 Crystal Structures of Organometallic Complexes

Schwartz's Reagent	
Stoichiometric Formula	$C_{24}H_{34}Cl_2Zr_2$
Temperature (K)	~293
Space Group	Pnnm
Unit cell lengths a, b, c (Å)	a = 7.21 Å, b = 9.20 Å, c = 13.84 Å
angles α, β, γ	α = 90.000°, β = 90.000°, γ = 90.000°
Reflections (#)	2139
Unique reflections (#)	307
Robs	17.8%
Rmeas	19.5%
CC1/2	97.8
Resolution (Å)	1.15
Completeness (%)	90.6%
Total exposure (e Å-2)	~3
R	0.1495
wR2	0.3785
GooF	2.254

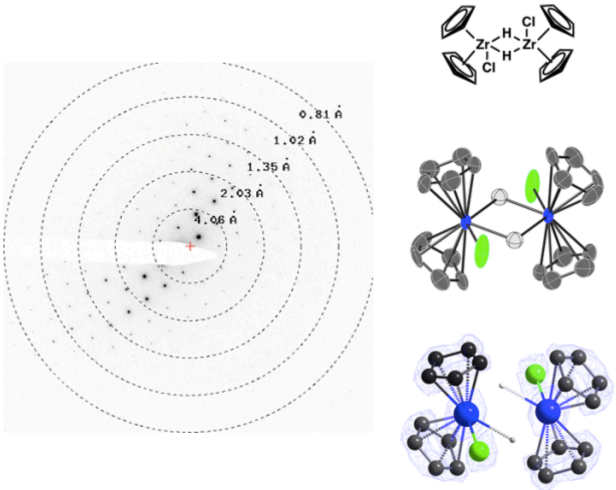


Figure 7.10. Data, statistics, sample diffraction pattern and structure for compound **7.1**.

Compound 3	
Stoichiometric Formula	$C_{36}H_{40}B_{18}O_2P_2Cl_{18}Pd_2$
Temperature (K)	~293
Space Group	P-1
Unit cell lengths a, b, c (Å)	13.050 12.040 13.060
angles α, β, γ	107.566 91.152 123.194
Reflections (#)	23883
Unique reflections (#)	3713
Robs	22.5%
Rmeas	23.8%
CC1/2	98.0
Resolution (Å)	0.90
Completeness (%)	80.5%
Total exposure (e Å-2)	~3
R	0.1808
wR2	0.4000
GooF	1.803

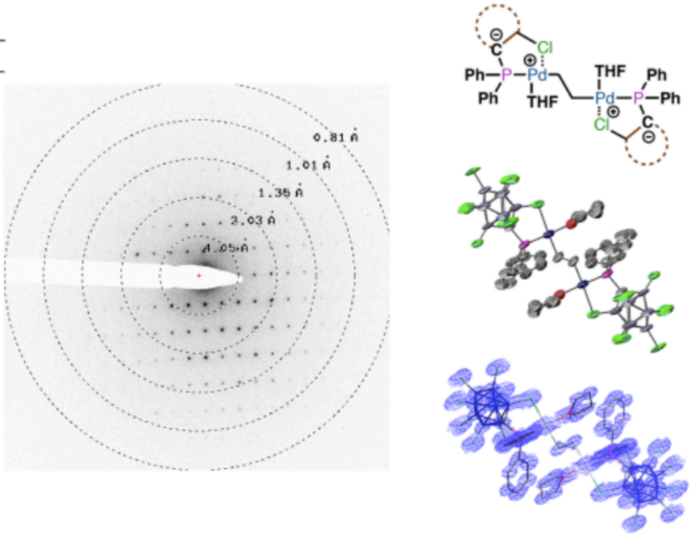


Figure 7.11. Data, statistics, sample diffraction pattern and structure for compound **7.3**.

$Fe(acac)_3$	
Stoichiometric Formula	$C_{15}H_{24}FeO_6$
Temperature (K)	~293
Space Group	Pbca
Unit cell lengths a, b, c (Å)	15.5200 13.4100 16.4900
angles α, β, γ	90.000 90.000 90.000
Reflections (#)	14159
Unique reflections (#)	2269
Robs	16.40%
Rmeas	17.90%
CC1/2	99
Resolution (Å)	0.9
Completeness (%)	92.10%
Total exposure (e Å-2)	~3
R	0.1607
wR2	0.3941
GooF	1.826

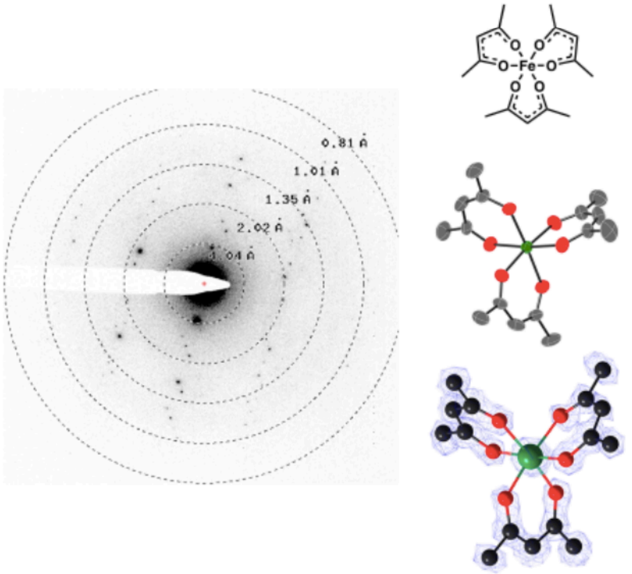


Figure 7.12. Data, statistics, sample diffraction pattern and structure for compound **7.4**.

Grubbs Catalyst 1st Generation	
Stoichiometric Formula	$C_{43}H_{72}Cl_2P_2Ru$
Temperature (K)	~293
Space Group	P2(1)/n
Unit cell lengths a, b, c (Å)	9.9700 26.0500 16.3400
angles α, β, γ	90.000 94.185 90.000
Reflections (#)	31201
Unique reflections (#)	6768
Robs	17.4%
Rmeas	19.6%
CC1/2	98.5
Resolution (Å)	0.85
Completeness (%)	99.3%
Total exposure (e Å-2)	~3
R	0.1595
wR2	0.3997
GooF	1.802

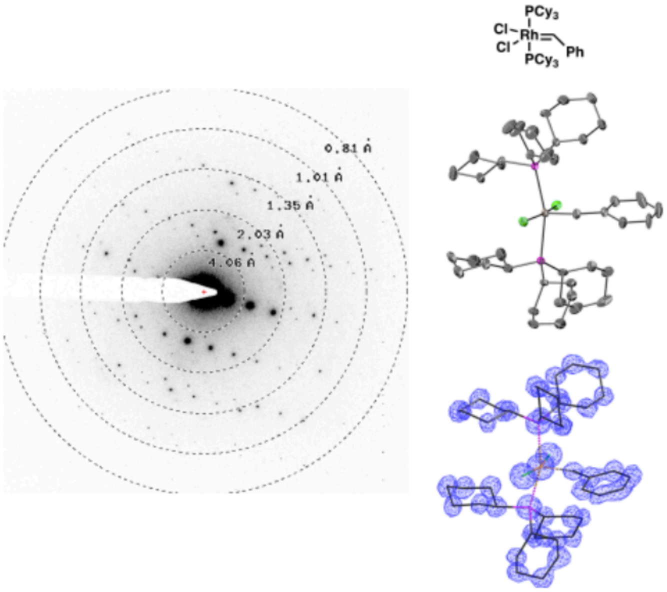


Figure 7.13. Data, statistics, sample diffraction pattern and structure for compound **7.5**.

Ni(dppf)Cl₂	
Stoichiometric Formula	$C_{34}H_{28}Cl_2FeNiP_2$
Temperature (K)	~293
Space Group	Pna2(1)
Unit cell lengths a, b, c (Å)	19.7600 10.1000 14.5800
angles α, β, γ	90.000 90.000 90.000
Reflections (#)	18929
Unique reflections (#)	1379
Robs	21.2%
Rmeas	22.1%
CC1/2	98.5
Resolution (Å)	1.00
Completeness (%)	86.3%
Total exposure (e Å-2)	~3
R	0.1126
wR2	0.2739
GooF	1.468

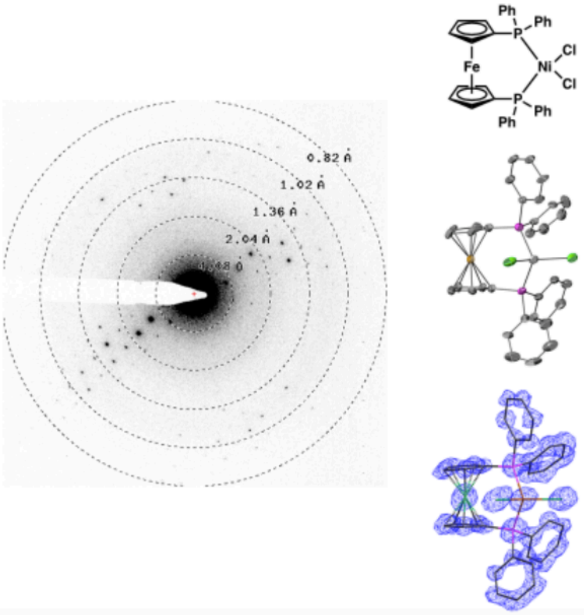


Figure 7.14. Data, statistics, sample diffraction pattern and structure for compound **7.6**.

Pd(dba)(PHOX)	
Stoichiometric Formula	$C_{43}H_{43}NO_2PPd$
Temperature (K)	~293
Space Group	P2(1)2(1)2(1)
Unit cell lengths a, b, c (Å)	10.2800 14.2700 26.5500
angles α, β, γ	90.000 90.000 90.000
Reflections (#)	12385
Unique reflections (#)	2102
Robs	26.0%
Rmeas	28.8%
CC1/2	93.5
Resolution (Å)	1.00
Completeness (%)	99.1%
Total exposure (e Å-2)	~3
R	0.1432
wR2	0.3504
GooF	1.249

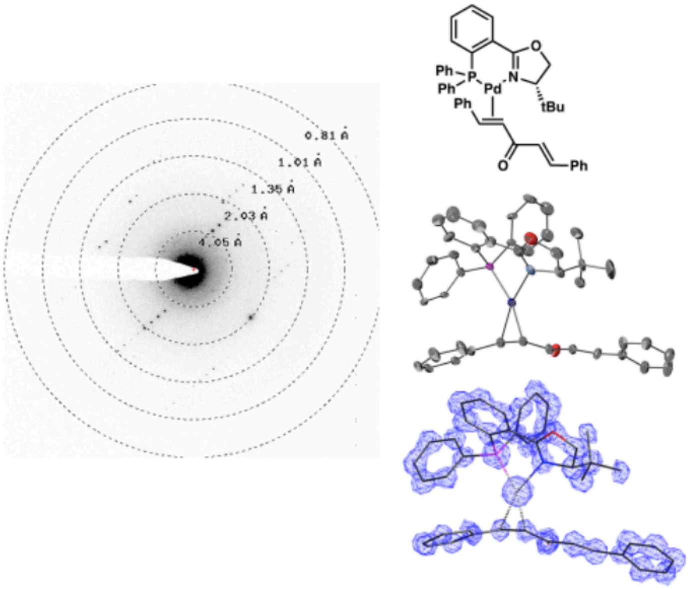


Figure 7.15. Data, statistics, sample diffraction pattern and structure for compound **7.7**.

$HRh(CO)(PPh_3)_3$	
Stoichiometric Formula	$C_{55}H_{46}OP_3Rh$
Temperature (K)	~293
Space Group	P2(1)/n
Unit cell lengths a, b, c (Å)	10.0100 31.7800 13.3400
angles α, β, γ	90.000 90.155 90.000
Reflections (#)	19748
Unique reflections (#)	3400
Robs	19.40%
Rmeas	21.30%
CC1/2	98.5
Resolution (Å)	1
Completeness (%)	76.70%
Total exposure (e Å-2)	~3
R	0.1324
wR2	0.3401
GooF	1.665

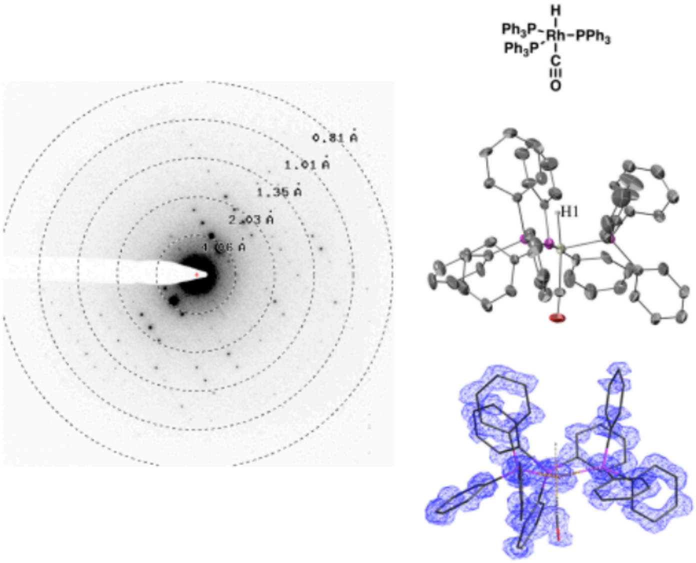


Figure 7.16. Data, statistics, sample diffraction pattern and structure for compound **7.8**.

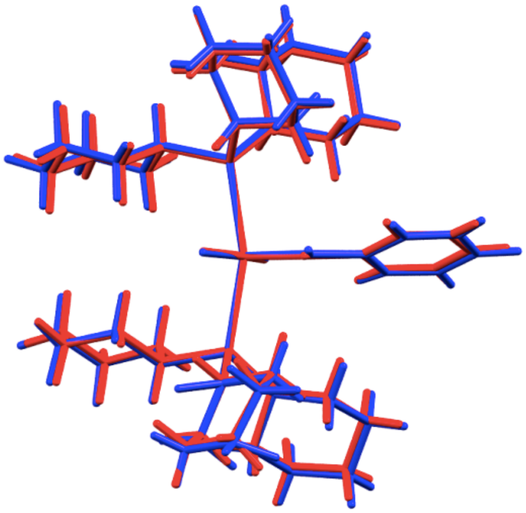


Figure 7.17. Overlay of electron diffraction structure obtained from Grubbs Catalyst (red) with previously reported X-ray structure (blue).[14] Root mean square (RMS) = 0.1311 Å, Max. Deviation = 0.2270 Å.

7.10.6.2 Crystal Structures of Synthetic Reaction Products

7.10.6.2.1 8-Chloro-12-cyclohexylidene-5-tosyl-5,6,7,12-tetrahydrodibenzo[b,e]azocine (7.9)

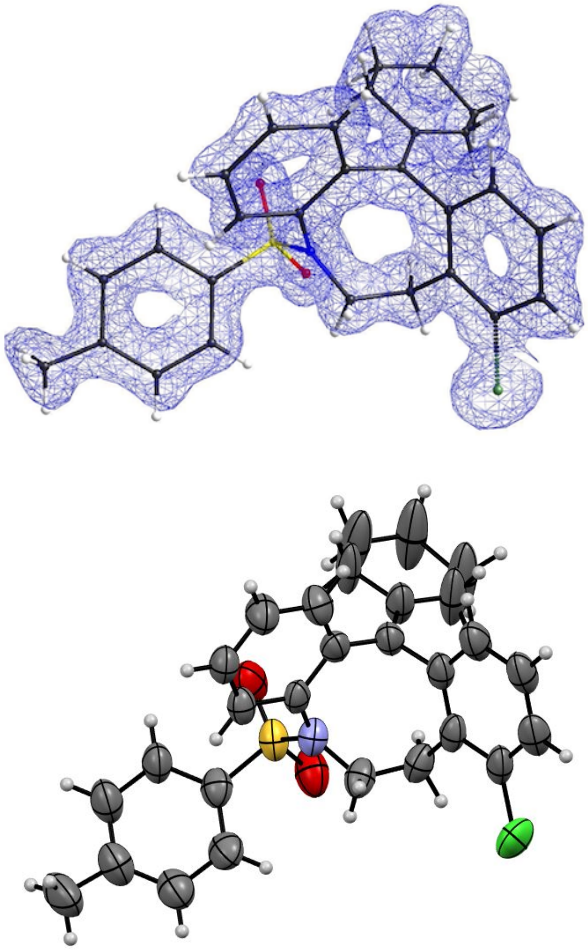


Figure 7.18. Asymmetric unit of **7.9** crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.30 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{28}H_{28}ClNO_2S$
Formula weight	478.02
Temperature (K)	100(2)
Space group	C 2/c
Cell dimensions	
a, b, c (Å)	28.02(4), 8.53(2), 19.40(2)
α, β, γ (°)	90, 107.98(5), 90
Resolution (Å)	1.00(1.05–1.00)

Observed reflections	10047(1570)
Unique reflections	2393(340)
Completeness (%)	96.3(97.7)
R_{obs} (%)	16.5(87.7)
R_{meas} (%)	18.7(99.2)
$I/\sigma I$	4.38(1.67)
$CC_{1/2}$ (%)	99.1(68.6)
R_1 (%)	18.82
wR_2 (%)	46.05
GooF	2.310

Table 7.3 Data collection and refinement statistics for **7.9**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4763	Data was truncated to a resolution of 1.05 Ångstrom due to poor diffraction data quality above this threshold.
PLAT084_ALERT_3_A High wR2 Value (i.e. > 0.25) .... 0.46 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT360_ALERT_2_A Short C(sp3)-C(sp3) Bond C1 - C2 . 1.09 Ang.	Diffraction data from two crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT360_ALERT_2_A Short C(sp3)-C(sp3) Bond C4 - C5 . 1.11 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

PLAT410_ALERT_2_A Short Intra H...H Contact H1A ..H2B . 1.66 Ang. x,y,z = 1_555 Check	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Additionally, the hydrogen atoms were placed using the riding model.
PLAT410_ALERT_2_A Short Intra H...H Contact H1B ..H2A . 1.67 Ang. x,y,z = 1_555 Check	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Additionally, the hydrogen atoms were placed using the riding model.
PLAT410_ALERT_2_A Short Intra H...H Contact H4A ..H5B . 1.67 Ang. x,y,z = 1_555 Check	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Additionally, the hydrogen atoms were placed using the riding model.
PLAT410_ALERT_2_A Short Intra H...H Contact H4B ..H5A . 1.66 Ang. x,y,z = 1_555 Check	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Additionally, the hydrogen atoms were placed using the riding model.

B-level Alerts	**Justifications**
PLAT082_ALERT_2_B High R1 Value .... 0.19 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.02053 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 7.4 Justification for alerts in **7.9** crystal structure.

7.10.6.2.2 1'-Tosyl-1',2',3',5'-tetrahydrospiro[cyclohexane-1,6'-indeno[7,1-bc]azepine] (7.10)

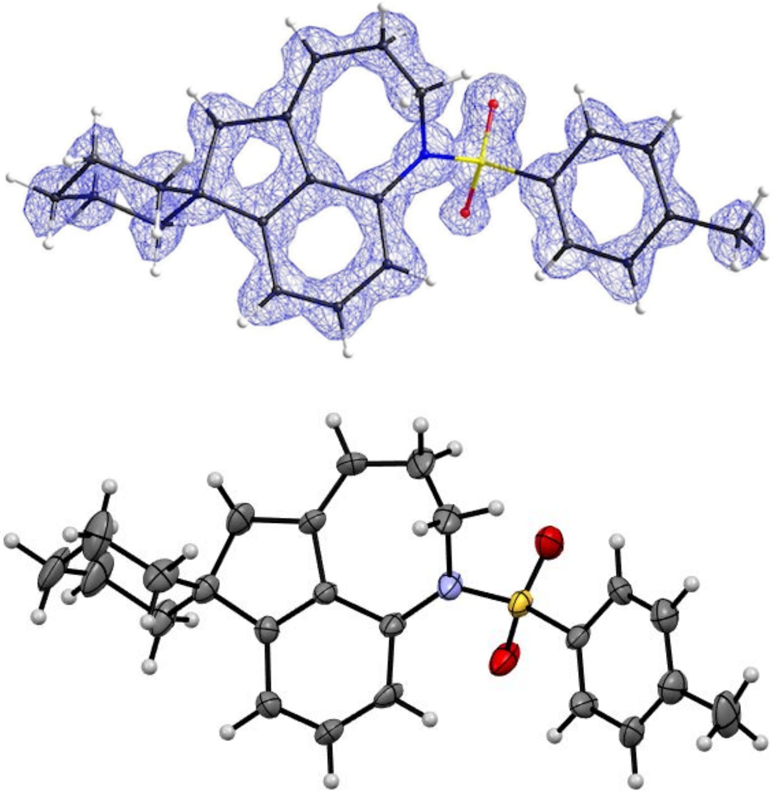


Figure 7.19. Asymmetric unit of **7.10** crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.60 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{24}H_{27}NO_2S$
Formula weight	393.53
Temperature (K)	293(2)
Space group	$P\ 2_1$
Cell dimensions	
a, b, c (Å)	11.26(1), 5.97(1), 14.30(1)
α, β, γ (°)	90, 109.18(4), 90
Resolution (Å)	1.00(1.05–1.00)
Observed reflections	4473(607)
Unique reflections	974(122)

Completeness (%)	90.6(89.7)
R_{obs} (%)	13.2(47.9)
R_{meas} (%)	14.9(53.1)
I/σI	8.95(4.19)
$CC_{1/2}$ (%)	98.7(88.4)
R_1 (%)	10.39
wR_2 (%)	27.77
GooF	1.577

Table 7.5 Data collection and refinement statistics for **7.10**. Highest resolution shell is shown in parenthesis.

A-level Alerts	Justifications
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4992	Data was truncated to a resolution of 1.00 Ångstrom due to poor diffraction data quality above this threshold.

B-level Alerts	Justification
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C2 -C5 . 1.34 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C16 -C21 . 1.33 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.02535 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT360_ALERT_2_B Short C(sp3)-C(sp3) Bond C12 - C13 . 1.30 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 7.6 Justification for alerts in **7.10** crystal structure.

7.10.6.2.3 5-Tosyl-2,3,4,5-tetrahydrospiro[benzo[b]cyclobuta[d]azepine-1,1'-cyclohexane] (7.11)

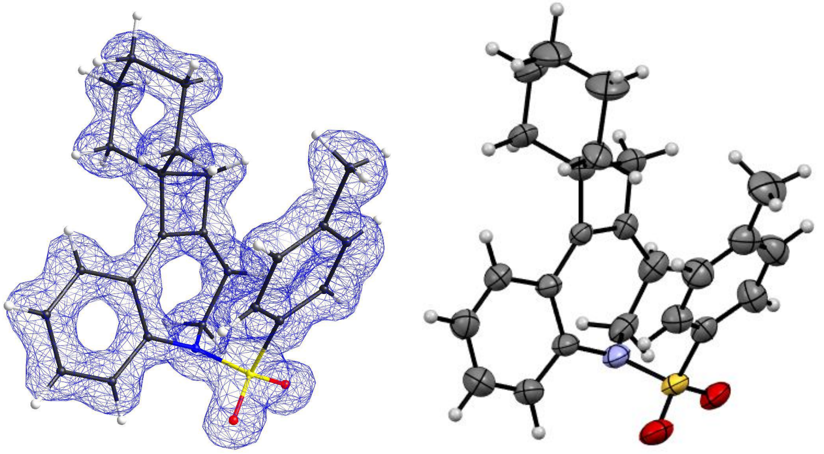


Figure 7.20. Asymmetric unit **7.11** crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.45 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{24}H_{27}NO_2S$
Formula weight	393.53
Temperature (K)	293(2)
Space group	$P\ 2_1 / c$
Cell dimensions	
a, b, c (Å)	14.010(4), 9.85(1), 15.29(2)
α, β, γ (°)	90, 116.16(7), 90
Resolution (Å)	1.05(1.10–1.05)
Observed reflections	7716(1075)
Unique reflections	1762(227)
Completeness (%)	95.2(96.6)

R_{obs} (%)	14.2(47.6)
R_{meas} (%)	16.2(53.2)
I/σI	6.52(3.66)
$CC_{1/2}$ (%)	99.4(96.5)
R_1 (%)	13.44
wR_2 (%)	35.12
GooF	1.905

Table 7.7 Data collection and refinement statistics for **7.11**. Highest resolution shell is shown in parenthesis.

A-level Alerts	**Justifications**
THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550 Calculated sin(theta_max)/wavelength = 0.4763	Data was truncated to a resolution of 1.05 Ångstrom due to poor diffraction data quality above this threshold.
PLAT334_ALERT_2_A Small <C-C> Benzene Dist. C2 -C5 . 1.31 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

B-level Alerts	**Justification**
PLAT149_ALERT_3_B s.u. on the beta Angle is Too Large .... 0.07 Degree	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering. Additionally, data collection occurred at ambient temperature.
PLAT334_ALERT_2_B Small <C-C> Benzene Dist. C16 -C21 . 1.33 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.01815 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 7.8 Justification for alerts in **7.11** crystal structure.

7.10.6.2.4 3,3-Dimethyl-2-phenyl-2,3-dihydroquinolin-4(1H)-one (7.12)

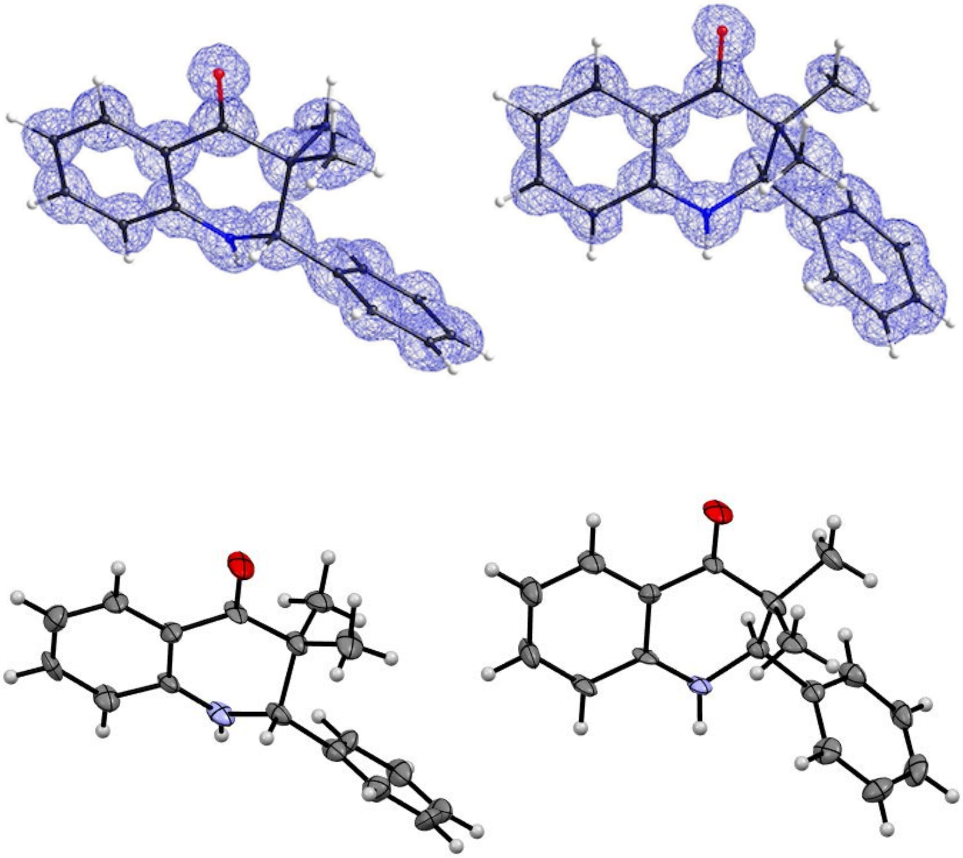


Figure 7.21. Asymmetric unit of **7.12** crystal structure. The structure is shown as ball-and-stick model overlaid with electron potential map (F_{obs}) contoured at 0.39 e Å^{-3} (top) and ORTEP diagram (bottom). Thermal ellipsoids are drawn at 30% probability.

Stoichiometric formula	$C_{17}H_{17}NO$
Formula weight	251.33
Temperature (K)	293(2)
Space group	P –1
Cell dimensions	
a, b, c (Å)	6.86(1), 9.64(1), 19.74(2)

α, β, γ (°)	101.54(6), 97.10(8), 88.9(1)
Resolution (Å)	0.85(0.90–0.85)
Observed reflections	14320(766)
Unique reflections	3988(427)
Completeness (%)	92.2(63.0)
R_{obs} (%)	15.0(60.1)
R_{meas} (%)	17.4(85.0)
I/σI	4.62(0.91)
$CC_{1/2}$ (%)	99.3(59.2)
R_1 (%)	16.86
wR_2 (%)	41.23
GooF	1.858

Table 7.9 Data collection and refinement statistics for **7.12**. Highest resolution shell is shown in parenthesis.

B-level Alerts	Justifications
PLAT082_ALERT_2_B High R1 Value .... 0.17 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) .... 0.41 Report	Electron diffraction data has higher R-factors compared to X-ray data. Multiple diffraction data sets were merged, and data collection occurred at ambient temperature.
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds .... 0.01189 Ang.	Diffraction data from multiple crystals were merged to obtain a desirable completeness, resulting in some structural disordering.

Table 7.10 Justification for alerts in **7.12** crystal structure.

7.10.7 Screened Coulombic Potential Computational Analysis and Isolation of Hydrides

The procedure and code for computational analysis is reported in the adapted article.

7.10.8 Movies

The movies of diffraction and Coulomb potential contour map is reported in the adapted article.

7.10.9 Relevant Spectra

Relevant spectra for compounds ***7.2*** *and* ***7.3*** *are reported in the adapted article.*

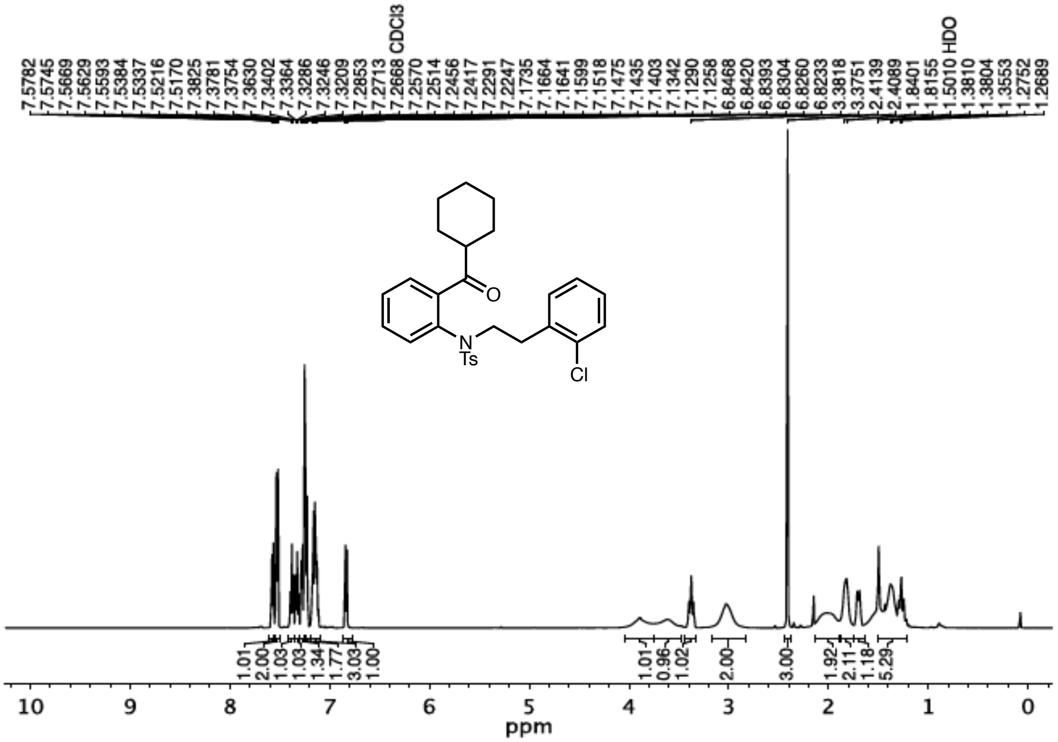


Figure 7.22 ^{1}H NMR (400 MHz, CDCl$_3$) of compound **7.18**.

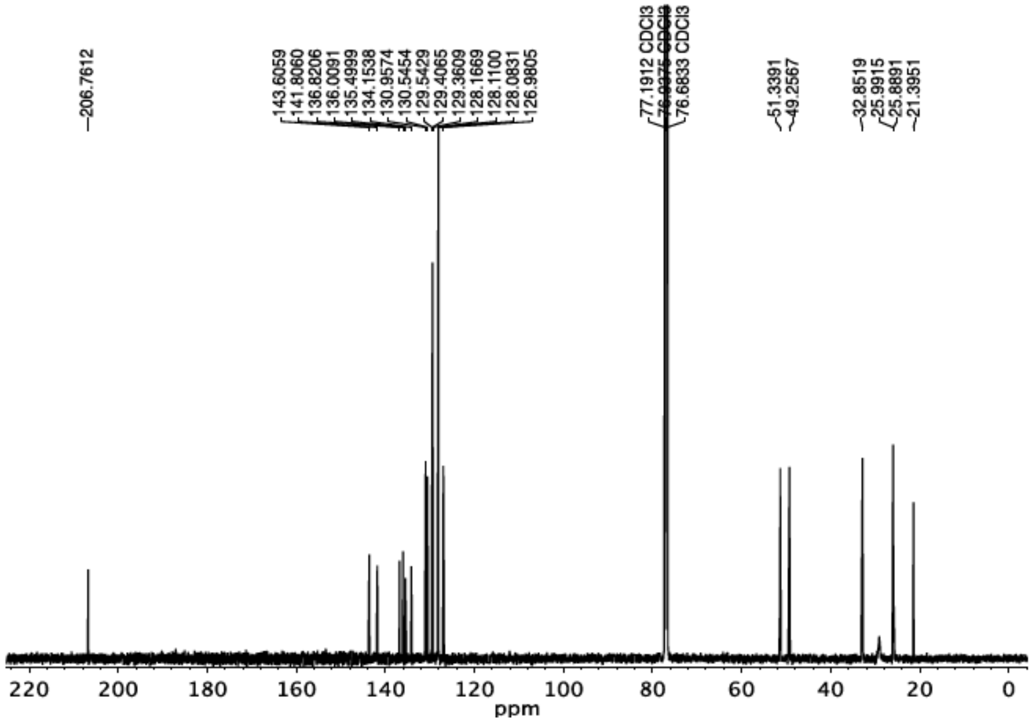


Figure 7.23 ^{13}C NMR (101 MHz, CDCl$_3$) of compound **7.18**.

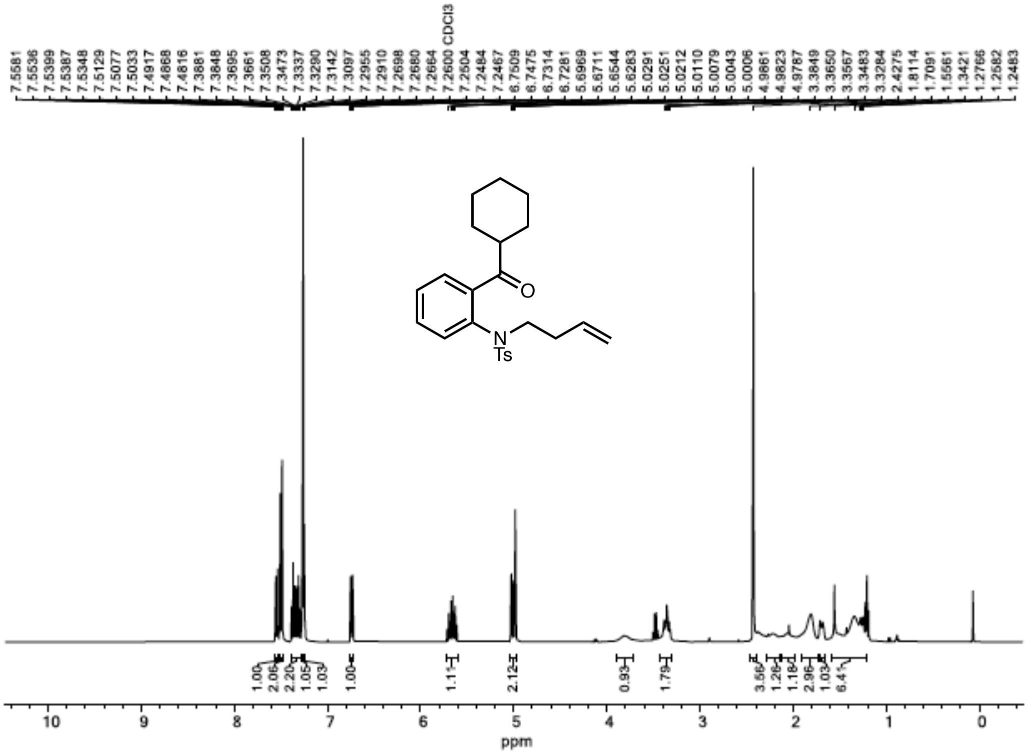


Figure 7.24 ^{1}H NMR (400 MHz, CDCl$_3$) of compound **7.19**.

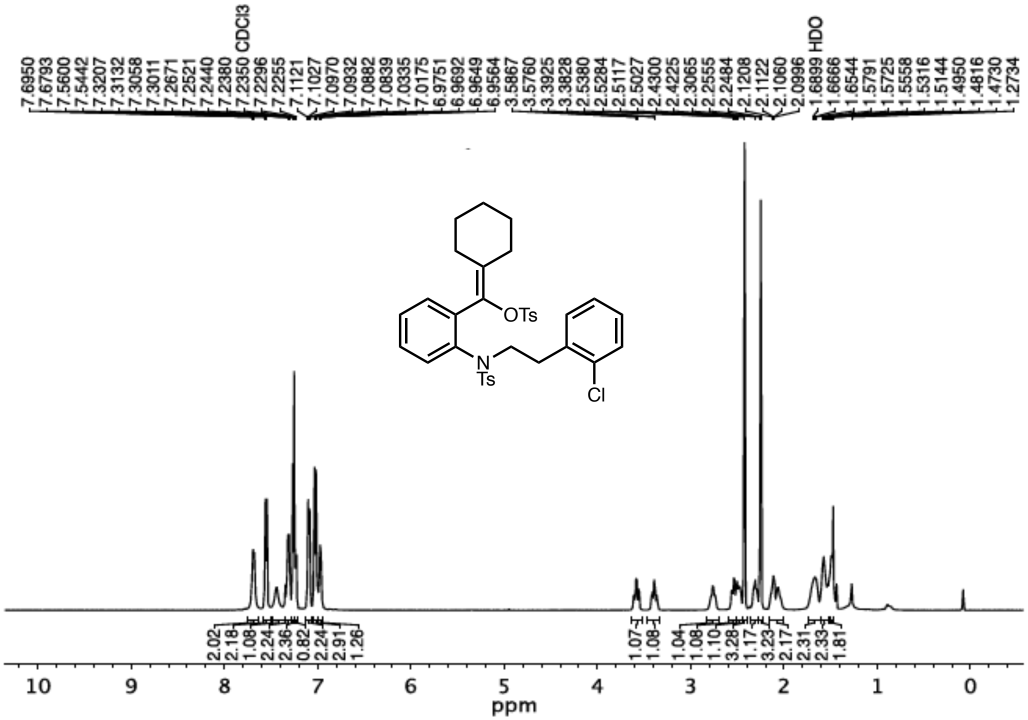


Figure 7.25 ^{1}H NMR (500 MHz, CDCl$_3$) of compound **7.20**.

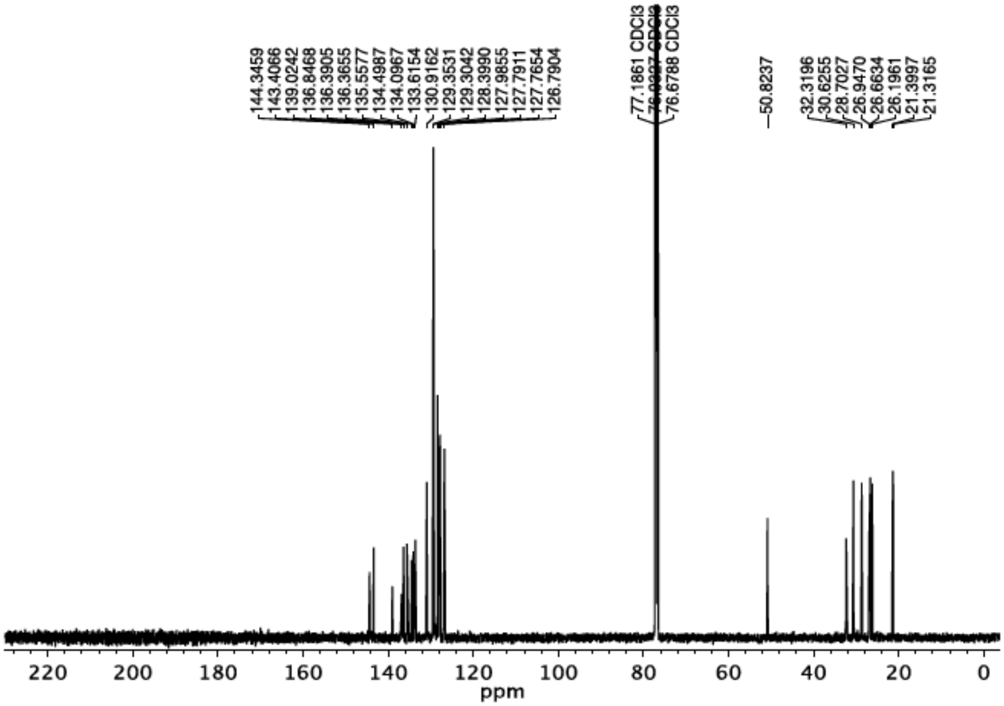


Figure 7.26 ^{13}C NMR (126 MHz, CDCl$_3$) of compound **7.20**.

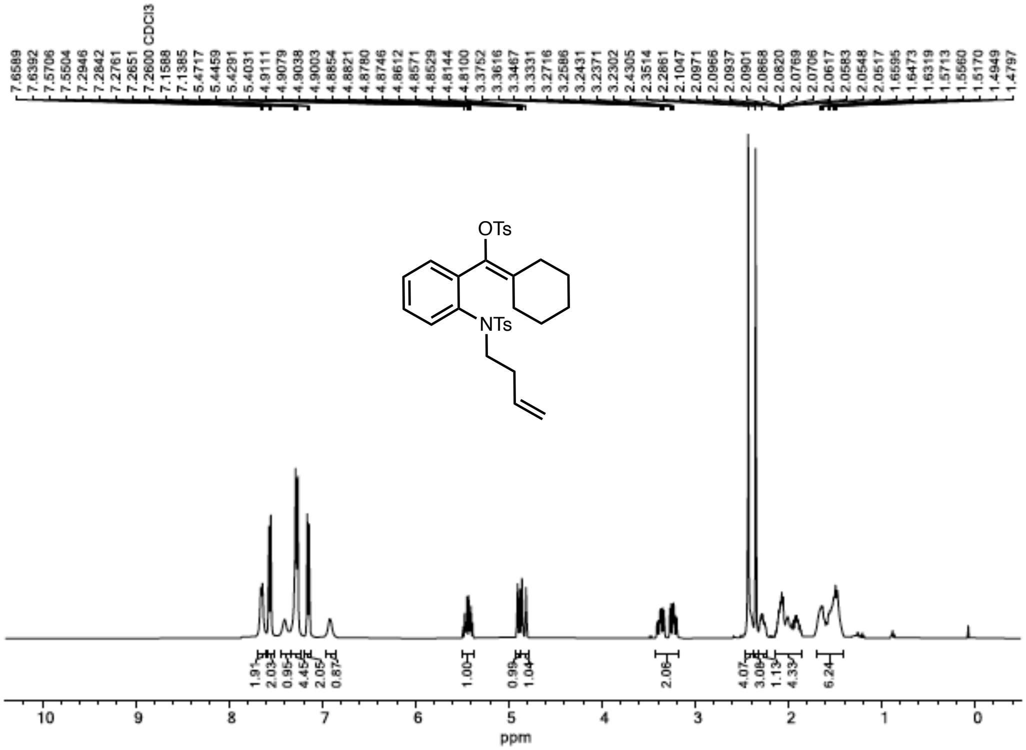


Figure 7.27 ^{1}H NMR (400 MHz, $CDCl_3$) of compound **7.21**.

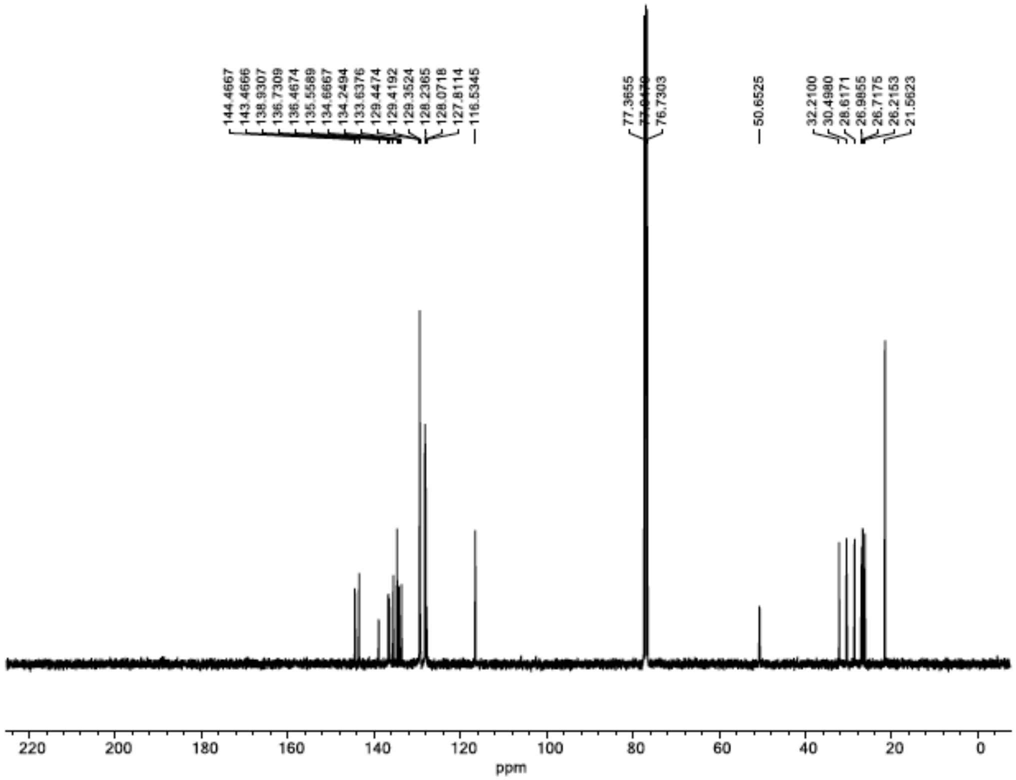


Figure 7.28 ^{13}C NMR (101 MHz, $CDCl_3$) of compound **7.21**.

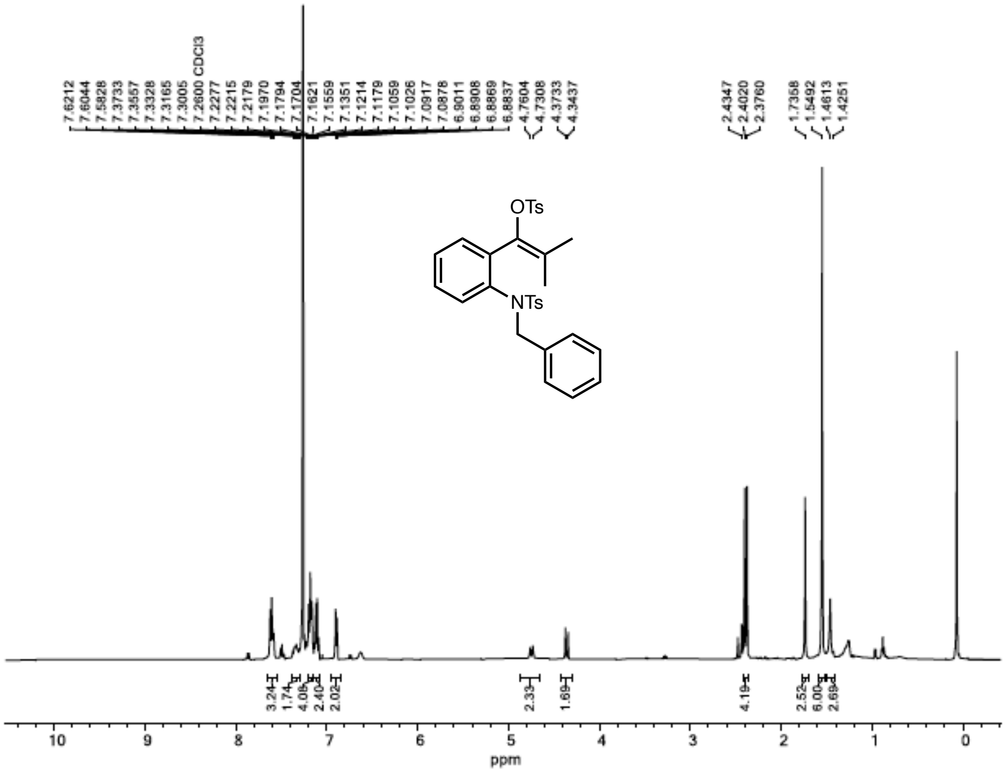


Figure 7.29 ^{1}H NMR (500 MHz, CDCl$_3$) of compound **7.22**.

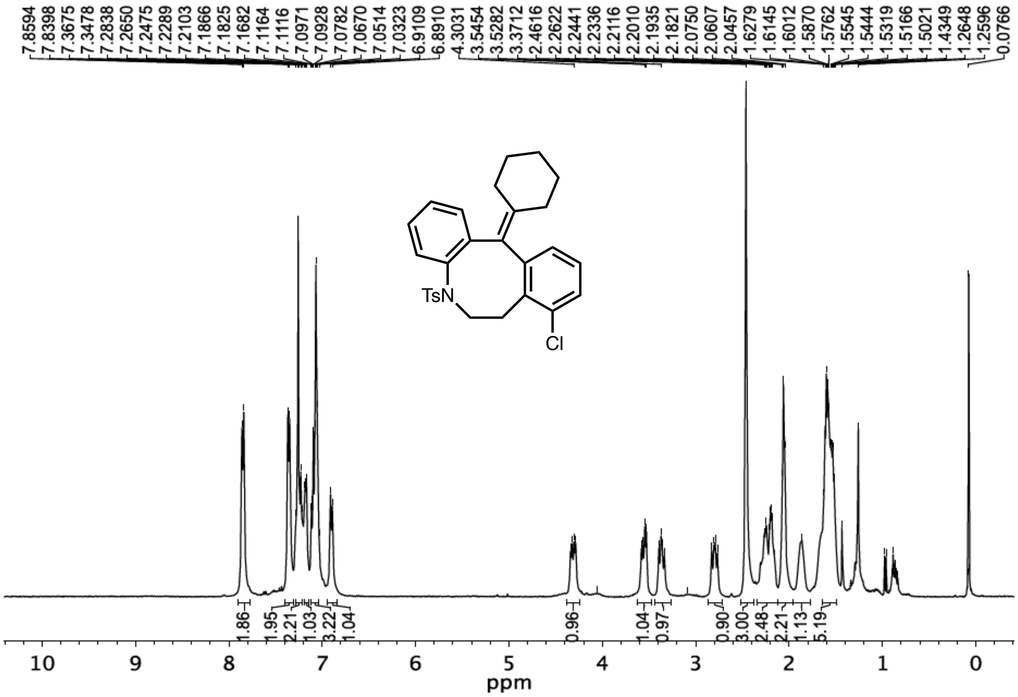


Figure 7.30 ^{1}H NMR (400 MHz, CDCl$_3$) of compound **7.9**.

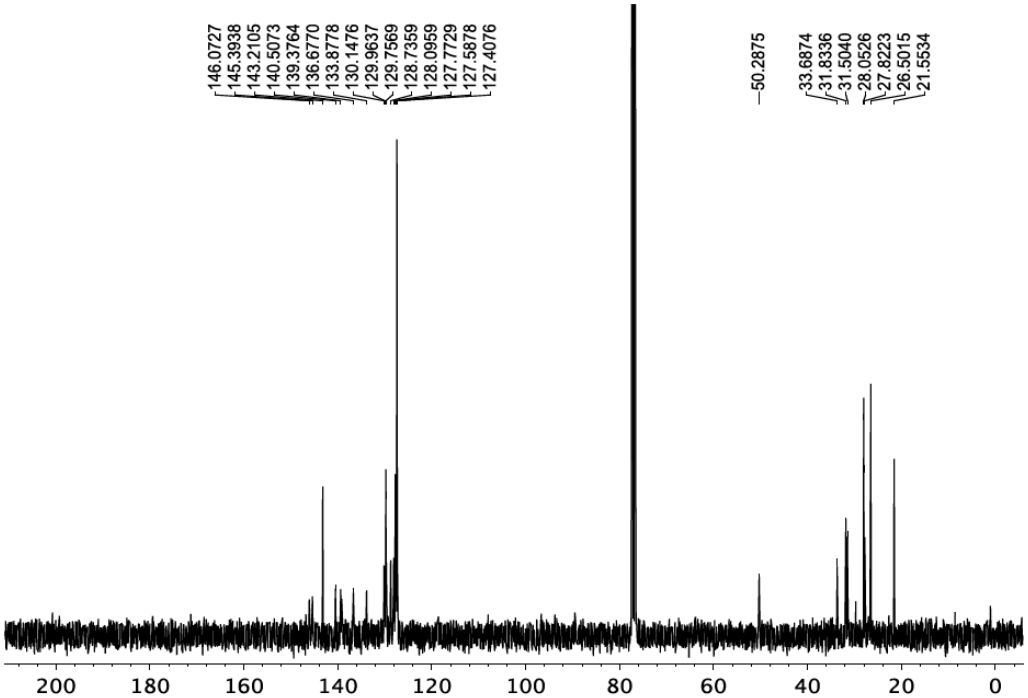


Figure 7.31 ^{13}C NMR (101 MHz, CDCl$_3$) of compound **7.9**.

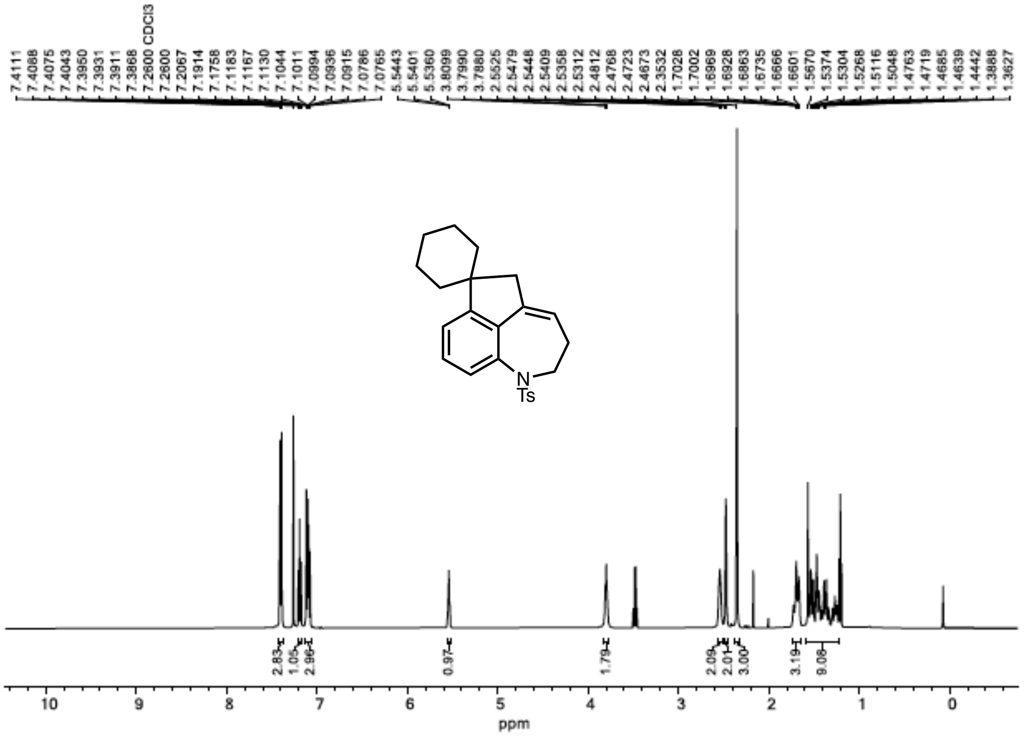


Figure 7.32 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **7.10**.

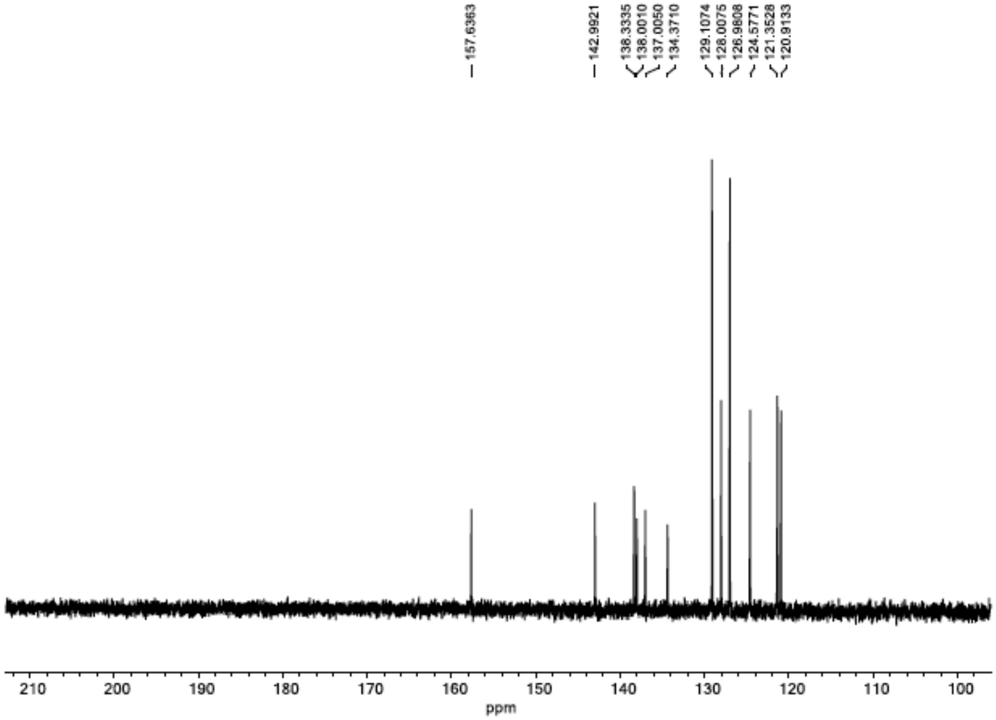


Figure 7.33 ^{13}C NMR (126 MHz, $CDCl_3$) of compound **7.10**.

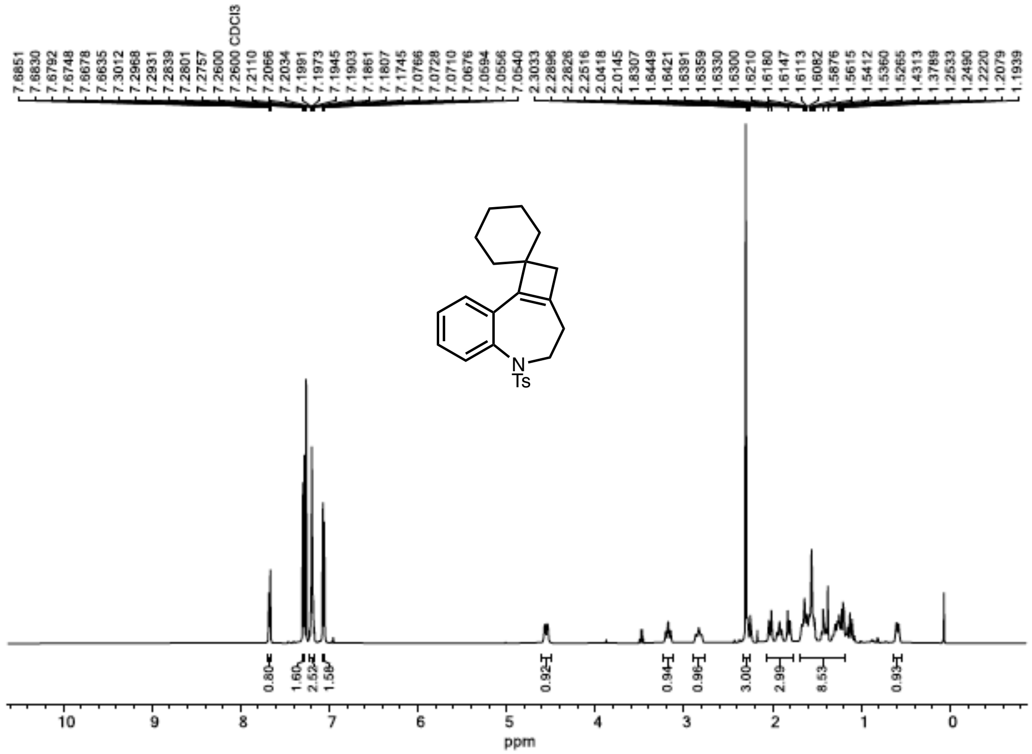


Figure 7.34 ^{1}H NMR (500 MHz, CDCl$_3$) of compound **7.11**.

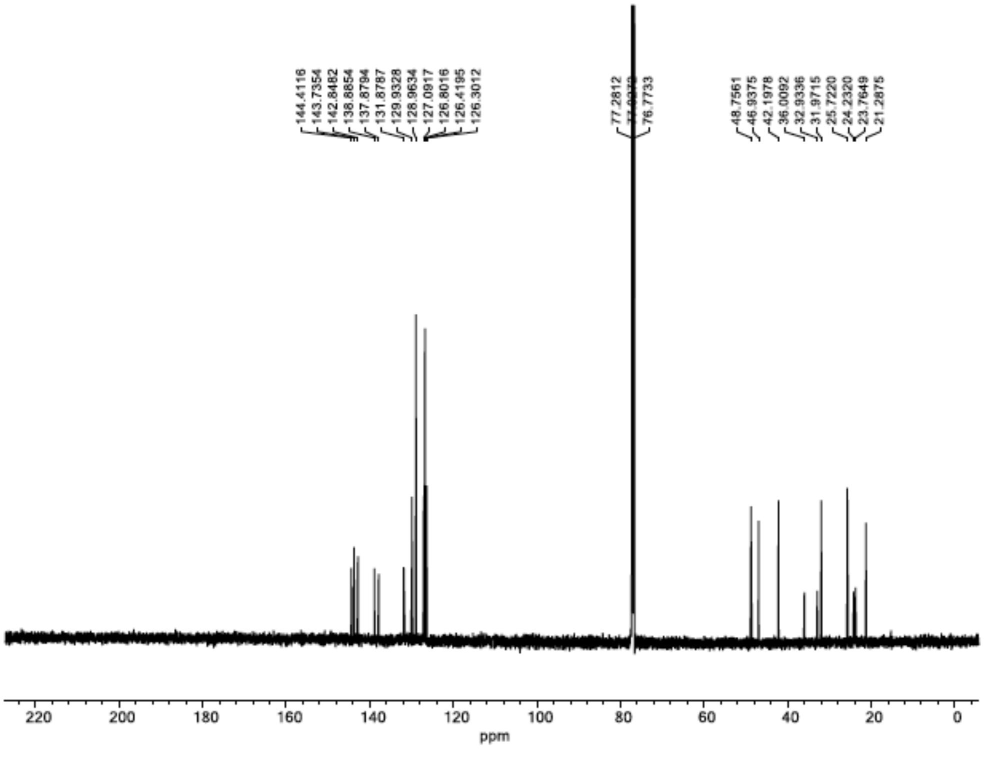


Figure 7.35 ^{13}C NMR (126 MHz, CDCl$_3$) of compound **7.11**.

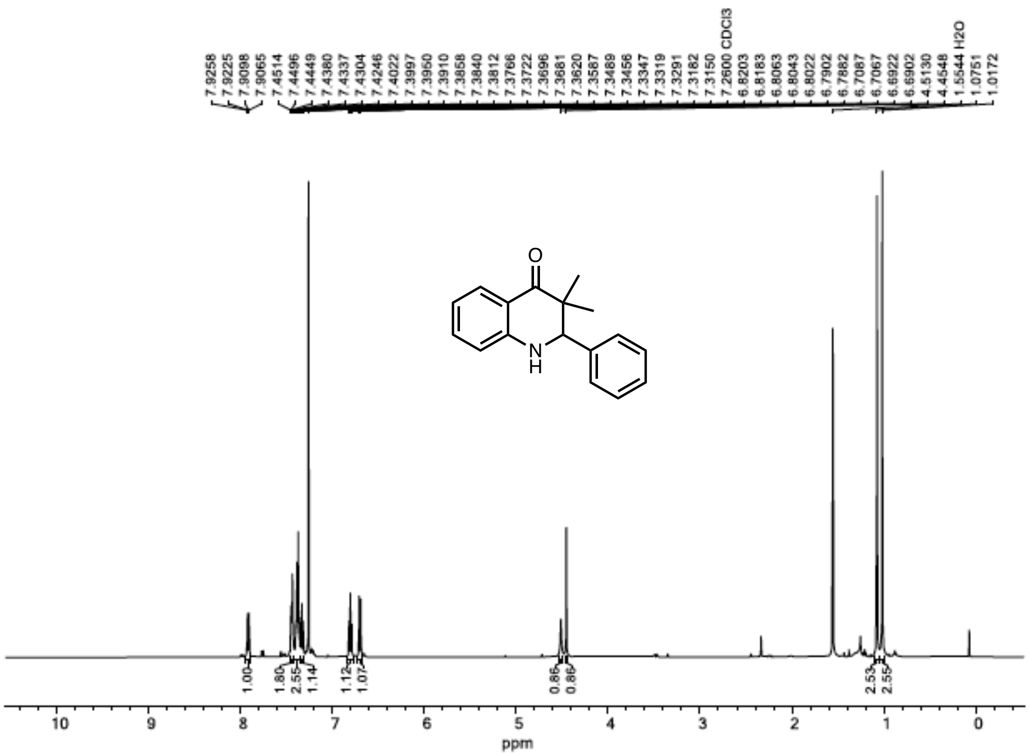


Figure 7.36 ^{1}H NMR (500 MHz, $CDCl_3$) of compound **7.12**.

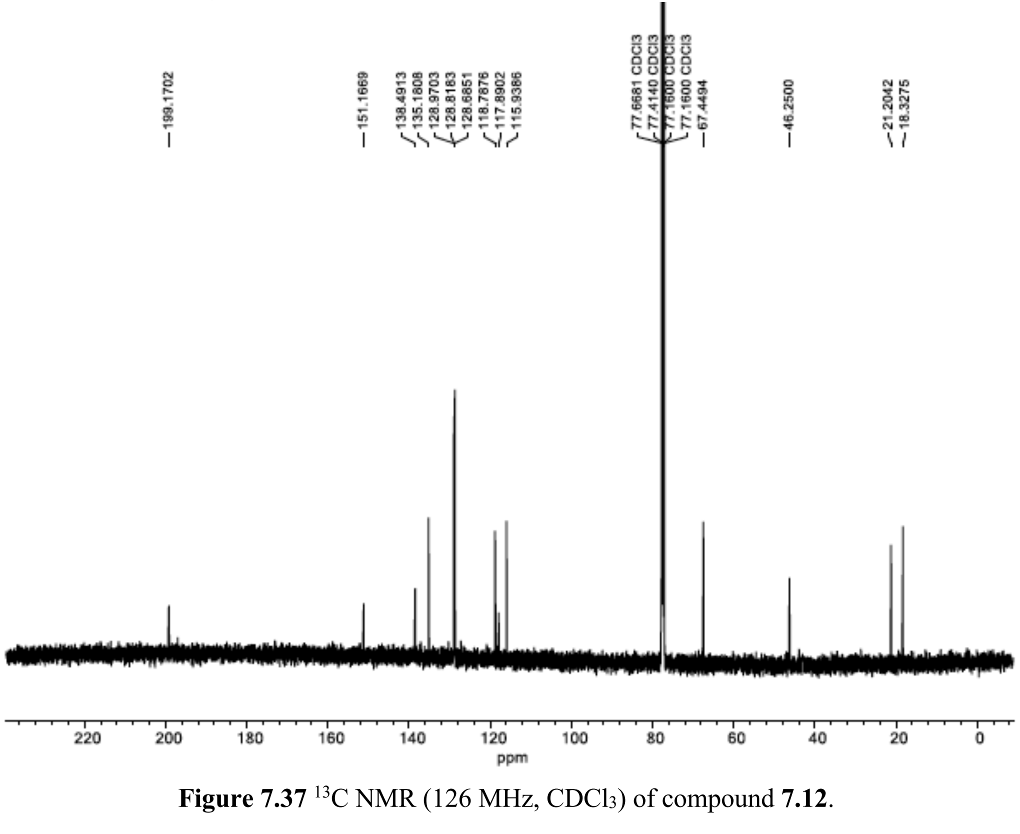


Figure 7.37 ^{13}C NMR (126 MHz, $CDCl_3$) of compound **7.12**.

7.10.10 Supplementary Notes and References

(1) Behenna, D. C.; Mohr, J. T.; Sherden, N. H.; Marinescu, S. C.; Harned, A. M.; Tani, K.; Seto, M.; Ma, S.; Novák, Z.; Krout, M. R.; McFadden, R. M.; Roizen, J. L.; Enquist Jr., J. A.; White, D. E.; Levine, S. R.; Petrova, K. V.; Iwashita, A.; Virgil, S. C.; Stoltz, B. M. Enantioselective decarboxylative alkylation reactions: Catalyst development, substrate scope, and mechanistic studies. *Science* **2011**, *17*, 14199–14223.

(2) Chaudhuri, M. K.; Ghosh, S. K. Novel synthesis of tris(acetylacetonato)iron(III). *Dalton Trans.* **1983**, 839–840.

(3) Pilloni, G.; Toffoletti, A.; Bandoli, G.; Longato, B. Homoleptic complexes of cobalt(0) and nickel(0,I) with 1,1‘-Bis(diphenylphosphino)ferrocene (dppf): Synthesis and characterization. *Inorg. Chem.* **2006**, *45*, 10321–10328.

(4) Kleinsasser, J. F.; Reinhart, E. D.; Estrada, J.; Jordan, R. F.; Lavallo, V. Ethylene oligomerization and polymerization by palladium(II) methyl complexes supported by phosphines bearing a perchlorinated 10-vertex closo-carborane anion substituent. *Organometallics* **2018**, *37*, 4773–4783.

(5) Kuprat, M.; Lehmann, M.; Shulz, A; Villinger, A. Synthesis of pentafluorophenyl silver by means of Lewis acid catalysis: Structure of silver solvent complexes. *Organometallics*, **2010**, *29*, 1421–1427.

(6) Colas, K.; Dos Santos, A. C. V. D.; Mendoza, A. *i*-Pr$_2$NMgCl·LiCl enables the synthesis of ketones by direct addition of Grignard reagents to carboxylate anions. *Org. Lett.* **2019**, *21*, 7908–7913.

(7) Bruneau-Voisine, A.; Pallova, L.; Bastin, S.; César, V.; Sortais, J.-B. Manganese catalyzed α-methylation of ketones with methanol as a C1 source. *Chem. Commun.* **2019**, *55*, 314–317.

(8) Hu, W.; Lin, J.-P.; Song, L.-R.; Long, Y.-Q. Direct synthesis of 2-aryl-4-quinolones via transition-metal-free intramolecular oxidative C(sp^3)–H/C(sp^3)–H coupling. *Org. Lett.* **2015**, *17*, 1268–1271.

(9) Nannenga, B. L.; Shi, D.; Leslie, A. G. W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930.

(10) Jones, C. G.; Martynowycz, M. W.; Hattne, J.; Fulton, T. J.; Stoltz, B. M.; Rodriguez, J. A.; Nelson, H. M.; Gonen, T. The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* **2018**, *4*, 1587–1592.

(11) Kabsch, W. Xds. *Acta Cryst.* **2010,** *D66*, 125–132.

(12) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **2010,** *D66*, 133–144.

(13) Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Cryst.* **2015**, *A71*, 3–8.

(14) Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3–8.

(15) Hübschle, C. B.; Sheldrick, G. M. & Dittrich, B. *ShelXle*: A Qt graphical user interface for *SHELXL*. *J. Appl. Cryst.* **2011,** *44*, 1281–1284.