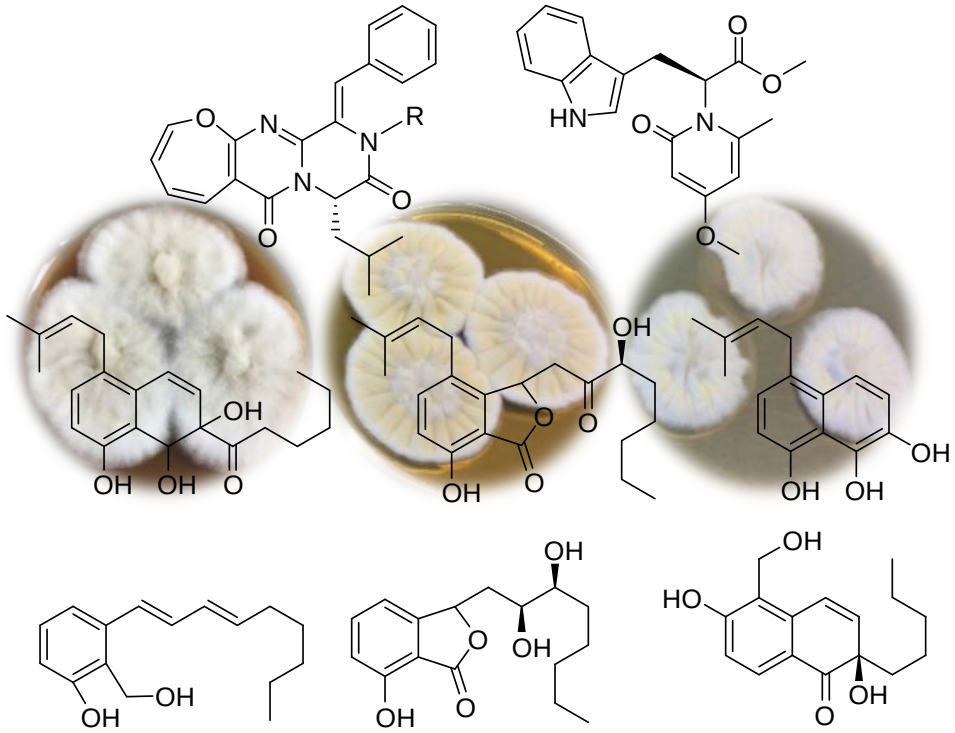
Mining of Cryptic Secondary Metabolism in Aspergillus

Guo, Yaojie

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Mining of Cryptic Secondary Metabolism in *Aspergillus*



Yaojie Guo

PhD Thesis

Department of Biotechnology and Biomedicine

Technical University of Denmark

September 2020

Mining of Cryptic Secondary Metabolism in *Aspergillus*

Yaojie Guo

PhD Thesis

Department of Biotechnology and Biomedicine

Technical University of Denmark

September 2020

Supervisors:

Professor Thomas Ostenfeld Larsen

Professor Uffe Hasbro Mortensen

Preface

This thesis is submitted to the Technical University of Denmark (Danmarks Tekniske Universitet, DTU) in partial fulfillment of the requirements for the Degree of Doctor of Philosophy. This work was carried out from September 2017 to September 2020 under the supervision of the main supervisor Professor Thomas Ostenfeld Larsen and co-supervisor Professor Uffe Hasbro Mortensen at the Department of Biotechnology and Biomedicine (DTU Bioengineering).

Throughout the three years´ research, I have received tons of help without which I wouldn´t be able to complete this thesis. First of all, I would like to give my sincere and deep gratitude to my supervisor Thomas, who is leading the Natural Product Discovery Group at Section for Microbial and Chemical Ecology, DTU Bioengineering. I was always feeling lucky to have such an experienced, open-minded, and easy-going supervisor. During the whole PhD study, Thomas not only gave me numerous scientific guidance, valuable discussions, and cheerful encouragements which kept me focused and motivated in the research but also supported me to learn new things beyond traditional natural product chemistry to expand my expertise.

Secondly, I would like to thank my co-supervisor Uffe Hasbro Mortensen and Dr. Fabiano Jares Contesini. Uffe offered me an opportunity to perform molecular engineering research in B223 which was an amazing environment to study fungi. The professional suggestions by him always helped me out when I met trouble in the *Aspergillus* engineering world. Fabiano taught me a lot of methodologies and practical operations during the genetic manipulation process. I also want to express my thanks to my collaborator Malgorzata Ewa Futyma for the amazing cooperation on several projects during which l got to learn molecular biology. A huge thanks go to Assoc. Prof. Charlotte Held Gotfredsen and Dr. Kasper Enemark-Rasmussen for all the NMR related stuff.

Special thanks go to Professor Jens Christian Frisvad for his experienced advice on all the fungal stuff. I am also very thankful to people in natural product discovery lab Ling Ding, Thomas Isbrandt Petersen, Mario Wibowo, Karolina Subko, Xinhui Wang, Aaron Andersen, Johan Vormsborg Christiansen, Lisette Knoth-Nielsen, Mette Amfelt, Elisabeth Varga, Andreas Heidemann, Christopher Phippen, as well as people in *Aspergillus* lab Jakob Blæsbjerg Hoof,

Andreas Møllerhøj Vestergaard and Jakob Kræmmer Haar Rendsvig. Besides, I feel like to extend my gratitude to all the people in B221 and B223 who offered a fantastic and collaborative atmosphere for research.

Financial support for this study from the China Scholarship Council and DTU Bioengineering is highly acknowledged. Finally, I want to thank my family for their continuous and boundless support thousand miles away all over the past three years.

Yaojie Guo

Kgs. Lyngby, Denmark

September 2020

Table of contents

Contents

Summary

Aspergillus species are ubiquitous filamentous fungi that can produce industrially important enzymes and a broad spectrum of structurally diversified secondary metabolites including pharmaceutical products like statins and echinocandin-type antibiotics. With the rapid evolvement of sequencing technologies, whole-genome sequences of *Aspergillus* species are becoming available. The access to the genomic data has made genome mining guided discovery of secondary metabolites (SMs) and linking the known and new products to the related biosynthetic gene clusters (BGCs) possible. However, only a limited number of *Aspergillus* species have been well studied for chemical discovery purposes. A plethora of new compounds still awaits discovery from the less known species or rare species which represent an untouched natural product arsenal. This PhD project focuses on two poorly studied species *A. californicus* and *A. homomorphus*.

The first aim of this PhD study has been to explore new secondary metabolites from the rare species *A. californicus* for antibacterial and cytotoxic screening. *A. californicus* was first isolated from Chamise chaparral (*Adenostoma fasciculatum*) soil in California in 1978. This asexual reproducing and poor sporulating species belongs to section *Cavernicolarum* of subgenus *Nidulantes* but resembles typical section *Usti* species because of its long light brown conidiophores. With only one isolate, this rare species has not been chemically investigated before. Through multiple chromatographic purification steps of the organic extracts, we were able to obtain fifteen compounds including ten previously unknown ones (Chapter 2.1). The structures of the new compounds were elucidated by a combination of techniques such as mass spectrometry, nuclear magnetic resonance, optical rotation, infrared spectroscopy, and electronic circular dichroism. Based on the biosynthetic origins, these compounds were categorized as eleven polyketides, two nonribosomal peptides, one sesterterpene, and one polyketide-nonribosomal peptide hybrid. Among the polyketides, nine octaketides were closely related and thus were proposed to be biosynthesized from the same pathway. The octaketide calidiol A showed moderate activities against methicillin-resistant *Staphylococcus aureus* MB5393 with a minimum inhibitory concentration of 48 μg/ml (Chapter 2.1.1 and Appendix 1).

Two other polyketides were elucidated as naphthyl derivatives (Chapter 2.1.2 and Appendix 2). The two nonribosomal peptides bear a unique oxepine-pyrimidinone-ketopiperazine (OPK) scaffold that has only been reported from fungal sources so far, particularly 70% from *Aspergillus* species (Chapter 2.1.3 and Appendixes 3 - 4).

The second aim of this study was to investigate the biosynthetic pathways of newly reported metabolites from this project as well as known structures with uncommon biosynthetic origins (Chapter 2.2). Genome mining tools like antiSMASH were used to predict the BGCs for the selected products. The hybrid compound calipyridone A is comprised of a 2-pyridone moiety and an indole substructure, the latter originating from a tryptophan residue. Its biosynthetic pathway was studied by engineering the host strain *A. californicus* with CRISPR-Cas9, which was the first genetic manipulating work on this non-model species. The results indicated that the 2-pyridone moiety was formed through aldol condensation without enzymatic ring expansion catalyzed by P450s, distinguishing calipyridone A from the biosynthesis of other 2-pyridone analogs. Besides, we successfully discovered two new precursors of calipyridone A using a molecular networking strategy and MSMS fragmentation analysis (Chapter 2.2.1 and Appendix 5). *A. homomorphus* is a protoheterothallic reproducing fungus belonging to section *Nigri* of subgenus *Circumdati*. Biosynthetic investigation of homopyrones from this species showed their synthesis to be of mixed biosynthetic origin starting with cinnamoyl-CoA as the starter unit, which was likely a product of uncommon phenylalanine ammonia-lyase in fungi, followed by three extension reactions facilitated by the polyketide synthase to give the aromatic α-pyrones (Chapter 2.2.2 and Appendix 6). Similar types of products are very commonly found in plants but rarely in filamentous fungi. Besides, we also present our effort on attempting to elucidate the biosynthetic pathway of the two OPK compounds, as well as linking of 5 genes, which were initially mis-proposed as the BGC for the two OPK products, to their encoding metabolites (Chapter 2.2.3 and Appendix 7).

Altogether, this PhD thesis has cast new insights into the secondary metabolism in *Aspergillus*, in particular in two less-studied species *A. californicus* and *A. homomorphus*. Regarding the secondary metabolite discovery, we have characterized fifteen compounds from *A. californicus*

including ten new ones. These compounds include polyketides (PKs), nonribosomal peptides (NRPs), sesterterpene, as well as two unusual hybrid types of natural products. Furthermore, in the biosynthesis investigation, we have succeeded in characterizing the biosynthetic pathway of calipyridone A in *A. californicus* representing a rare 2-pyridone formation without a P450-catalyzed ring expansion from a tetramic acid intermediate. Besides, the biosynthesis of two yellow homopyrones in *A. homomorphus* has been demonstrated to start with an unusual cinnamoyl-CoA unit which is rare in fungi. Importantly, our work represents the first genetic manipulations in both *A. californicus* and *A. homomorphus,* and thus as a huge benefit sets the scene for further biochemical investigations in these species.

Sammenfatning

Aspergillus er en allestedsnærværende gruppe af filamentøse svampe, som kan producer industrielt vigtige enzymer, samt et bredt spektrum af strukturelt diverse sekundære metabolitter, så som statiner og echinocandin-lignende antibiotika. Som følge af den hurtige udvikling af gensekventeringsteknologier, er fuldgenomskevenser af *Aspergillus* arter blevet lettere tilgængelige. Denne nye adgang til genomdata, har gjort genom baseret opdagelse af sekundære metabolitter, samt opdagelse af sammenhænge mellem kendte og nye naturstoffer og deres ansvarlige biosyntetiske gener mulig. Det er dog kun et begrænset antal *Aspergillus*-arter, hvis kemiske potential er blevet undersøgt ved hjælp af denne teknik. Et utal af nye naturstoffer derfor venter stadig på at blive opdaget fra mindre beskrevne og sjældne arter, der udgør et stadig uberørt lager. Dette PhD projekt fokuserer på to ringe studerede arter, *A. californicus* og *A. homomorphus*.

Det første mål med dette PhD studie var at udforske nye sekundære metabolitter fra den sjældne art *A. californicus* med henblik på at finde bioaktive stoffer med antibakteriel eller cytotoksisk virkning. *A. californicus* blev først isoleret fra jordprøver taget ved planten *Adenostoma fasciculatum* i Californien i 1978. Denne aseksuelt reproducerende og dårligt sporulerende art, tilhører sektion *Cavernicolarum* i underslægten *Nidulantes*, med har ligheder med arter fra sektion *Usti* på grund af dens lange lysebrune konidioforer. Der findes kun et enkelt isolat af denne art, og derfor er den ikke tidligere blevet kemisk karakteriseret. Ved hjælp af flere kromatografiske oprensningstrin af det organiske ekstrakt, kunne vi oprense femten forskellige forbindelser, heriblandt ti hidtil ukendte (Kapitel 2). Strukturerne af de nye forbindelser blev opklaret ved hjælp af en kombination af teknikker som massespektrometri, kernemagnetisk resonansspektroskopi, optisk drejning, infrarød spektroskopi, og elektronisk cirkulær dikromisme. Forbindelserne blev baseret på deres biosyntetiske oprindelse, kategoriseret som elleve polyketider, to non-ribosomale peptider, en sesterterpen, og en polyketid-non-ribosomal peptid hybrid. Blandt polyketiderne, var ni oktaketider nært beslægtede og disse var derfor foreslået at tilhøre samme biosyntesevej. Oktaketidet calidiol A udviste moderat aktivitet mod methicillin-resistent *Staphylococcus aureus* MB5393, med en minimum hæmmende koncentration på 48

µg/mL. To andre polyketider blev karakteriseret som værende naphtyl-afledte. De to non-ribosomale peptider viste sig at indeholde unikke oxepine-pyrimidinone-ketopiperaziner (OPK) strukturer, hidtil kun rapporteret fra svampe, særligt *Aspergillus* arter.

Det andet mål med dette studie var at undersøge biosynteseveje for de nye forbindelser karakteriseret i dette projekt, såvel som for kendte strukturer med ualmindelige biosynteseveje (Kapitel 2). Ved hjælp af bioinformatiske værktøjer som eksemplevis antiSMASH forudsagde vi de biosyntetiske gener for de udvalgte naturstoffer. Hybridforbindelsen calipyridone A består af en 2-pyridone del og en tryptofan-afledt indolstruktur og biosyntesevejen blev undersøgt ved, for første gang at benytte CRISPR-Cas9-baseret genmodifikation i værtsorganismen *A. californicus*. Resultaterne indikerede at 2-pyridone-delen dannes via en aldolkondensation uden enzymatisk ringudvidelse katalyseret af P450'ere, og dermed gør calipyridone A biosyntesen særskilt i forhold til andre 2-pyridone afledte forbindelser. Derudover, blev to calipyridone A forløbere opdaget ved hjælp af molekylære netværk og analyse af tandem MS fragmentationsdata. *A. homomorphus* er en protoheterothallisk reproducerende svamp tilhørende sektion *Nigri* i underslægt *Circumdati*. Biosyntetiske undersøgelser af homopyroner fra denne art afslørede en blandet biosyntese med cinnamoyl-CoA som startenhed, sandsynligvis et produkt af en, i svampe, usædvanlig phenylalanine ammonium lyase, efterfulgt af tre forlængelser faciliteret af polyketidsyntasen, resulterende i de aromatiske α-pyroner. Lignende forbindelser findes ofte i planter, men er sjældne i filamentøse svampe. Derforuden, præsenterer vi vores forsøg på at opklare biosyntesevejen for de to OPK forbindelser, såvel som at sammenkæde produkter til fem gener, der først blev foreslået som værende ansvarlige for biosyntesen af de to OPK forbindelser, men dog viste sig ikke at være involveret i biosyntesen af disse forbindelser.

Samlet har denne PhD afhandling givet ny indsigt i de sekundære metabolitter fra *Aspergillus*, navnlig de to ringe beskrevne arter *A. californicus* og *A. homomorphus*. Det lykkedes at karakterisere femten kemiske forbindelser *A. californicus*, heraf ti ikke før beskrevne. Disse forbindelser inkluderer polyketider, non-ribosomale peptider, en sesterterpen, såvel som to usædvanelige hybrid forbindelser. Derudover, har vi ved hjælp af biosyntetiske studier,

karakteriseret biosyntesevejen for calipyrodone A i *A. californicus*, som værende en sjælden 2-pyridone, dannet uden P450-katalyseret ringudvidelse af et tetraminsyreintermediat. Yderligere, har vi vist at biosyntesen af to gule homopyroner i *A. homomorphus* starter med en cinnamoyl-CoA enhed, usædvanlig for svampe. Slutteligt, viser vi den første genetiske manipulation i både *A. californicus* og *A. homomorphus*, og har dermed åbner muligheden for fremtidige biokemiske studier af disse arter.

List of publications and conference contributions

Publications

Paper 1 **Yaojie Guo**, Ling Ding, Simone Ghidinelli, Charlotte H. Gotfredsen, Mercedes de la Cruz, Thomas A. Mackenzie, Maria C. Ramos, Pilar Sánchez, Francisca Vicente, Olga Genilloud, Sonia Coriani, René W. Larsen, Jens C. Frisvad, and Thomas O. Larsen. Taxonomy and Rareness Driven Discovery of Polyketides from *Aspergillus californicus* (*Journal of Natural Products*, submitted)

Paper 2 **Yaojie Guo**, Simone Ghidinelli, Mercedes de la Cruz, Thomas A. Mackenzie, Maria C. Ramos, Pilar Sánchez, Francisca Vicente, Olga Genilloud, and Thomas O. Larsen. New naphthyl derivatives from *Aspergillus californicus* (*Journal of Antibiotics*, accepted)

Paper 3 **Yaojie Guo**, Simone Ghidinelli, Mercedes de la Cruz, Thomas A. Mackenzie, Maria C. Ramos, Pilar Sánchez, Francisca Vicente, Olga Genilloud, and Thomas O. Larsen. Oxepinamides L and M, two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides from *Aspergillus californicus* (*Natural Product Research*, submitted)

Paper 4 **Guo, Y.;** Frisvad, J. C.; Larsen, T. O. Review of Oxepine-Pyrimidinone-Ketopiperazine Type Nonribosomal Peptides. *Metabolites* **2020**, *10* (6), 246.

Paper 5 **Yaojie Guo**[1], Fabiano J. Contesini[1], Simone Ghidinelli, Uffe H. Mortensen, and Thomas O. Larsen. Biosynthesis of Calipyridone A Indicates a Fungal 2-pyridone Product Formation Without Ring Expansion (1 contributed equally, *Organic Letters*, submitted)

Paper 6 Malgorzata E. Futyma[1], **Yaojie Guo**[1], Casper Hoeck, Jakob B. Hoof, Charlotte H. Gotfredsen, Uffe H. Mortensen, Thomas O. Larsen. Genetic origin of homopyrones, a rare type of phenylpropanoid and polyketide derived yellow pigments from *Aspergillus homomorphus*. (1 contributed equally, *Applied Microbiology and Biotechnology*, submitted)

Conference contributions

Yaojie Guo, Ling Ding, Charlotte H. Gotfredsen, Thomas O. Larsen. Chemical compositions from *Aspergillus californicus* and proposed biosynthetic pathway of phthalide products. Nordic Natural Product Conference 2019, Flúðir, Iceland.

Yaojie Guo, Simone Ghidinelli, Sonia Coriani, Charlotte H. Gotfredsen, Thomas O. Larsen. NMR in natural products: solving the absolute structures of novel fungal enantiomers and diastereomers. 41st Danish NMR meeting 2020, Korsør, Denmark.

List of abbreviations

1D NMR	one-dimensional NMR spectroscopy
2D NMR	two-dimensional NMR spectroscopy
A domain	Adenylation domain
AC	Absolute configuration
ACP	Acyl carrier protein
AT	Acyltransferase domain
B3LYP	Becke three parameters Lee–Yang–Parr
BGC	Biosynthetic Gene Cluster
C domain	Condensation domain
C18	Octadecylsilane
C6-Ph	Phenyl-hexyl
C8	Octylsilane
Cas9	CRISPR-associated protein 9
CD	Circular Dichroism
CoA	Co-enzyme A
COSY	Correlation Spectroscopy
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAD	Diode Array Detector
DFT	Density Functional Theory
DH	Dehydratase domain
DMAPP	Dimethylallyl Pyrophosphate
dsDNA	Double-stranded DNA
E domain	Epimerization domain
ECD	Electronic Circular Dichroism
edHSQC	Multiplicity Edited Heteronuclear Single Quantum Correlation
EI	Electron ionization
ER	Enoyl Reductase domain

ESI	Electrospray Ionization
EtOAc	Ethylacetate
Phe	Phenylalanine
H2BC	Heteronuclear 2-Bond Correlation
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High-Performance Liquid Chromatography
IPP	Isopentenyl Pyrophosphate
iPKS	Iterative Polyketide Synthase
KR	Ketoreductase domain
KS	Ketosynthase domain
m/z	Mass-to-charge ratio
MeOH	Methanol
MIC	Minimum Inhibitory Concentration
MMFF	Merck Molecular force field
mPKS	Modular Polyketide Synthase
MS	Mass spectrometry
MT	Methyltransferase
NGS	Next-Generation Sequencing
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Enhancement
NOESY	Nuclear Overhauser Effect Spectroscopy
NP	Natural Product
NR-PKS	Non-reducing Polyketide Synthase
NRP	Nonribosomal Peptide
NRPS	Nonribosomal Peptide Synthetase
OPK	Oxepine-Pyrimidinone-Ketopiperazine
OR	Optical Rotation
PCP	Peptidyl Carrier Protein domain
PK	Polyketide

PKS	Polyketide Synthase
Q	Quadrupole
SM	Secondary Metabolite
T domain	Thiolation domain
TD	Terminal reductase domain
TDDFT	time-dependent Density Functional Theory
TE	Thioesterase domain
ToF	Time of Flight
Uri	Uridine
Ura	Uracil
USER	Uracil-Specific Excision Reagent
UV-vis	Ultraviolet-visible
WT	Wild Type

1. Introduction

1.1 Fungi

Fungi is one of the largest eukaryotic kingdoms in the tree of life consisting of a huge amount of diverse species including yeasts, mushrooms, molds, etc. These cosmopolitan organisms can be found almost everywhere on Earth and have a profound impact on the ecosystem by decomposing biopolymers and other biomolecules from dead or live hosts, and synthesizing a broad range of biological molecules for various purposes.[1] Currently, the number of accepted fungal species stays around 120 000 which is only 3% to 8% of the estimated figure from 2.2 to 3.8 million. A large number of unrecognized fungal species await discovery from biodiversity hot spots, little-explored habitats, etc.[2]

1.2 *Aspergillus*

Aspergillus is a diverse genus of filamentous fungi with 446 known species that have been accepted.[3] Belonging to phylum Ascomycota, class Eurotiomycetes, order Eurotiales, family Trichocomaceae, the speciation of genus *Aspergillus* species has been over 200 million years.[4] The characteristic morphological structure of *Aspergillus* species is the conidiophore (Figure 1) consisting of an unbranched stipe (usually aseptate) with a swollen apex (vesicle). Phialide borne directly on the vesicle (uniseriate) or the metulae (biseriate). Vesicle, phialide, metulae (if present), and conidia (asexual spores) form the conidial head.[5] The genus acquired its name because its conidiophore is similar to an aspergillum. However, it´s notable that the presence of an aspergillum-like conidial head does not guarantee a given species belonging to this genus as some other types of conidium-bearing structures exist beyond *Aspergillus* species.[6] *Aspergillus* species are mostly aerobic, and many are capable of growing in vital nutrients-depleted environments. In nature, they exist as endophytes, saprophytes, parasites, human and animal pathogens.[7]

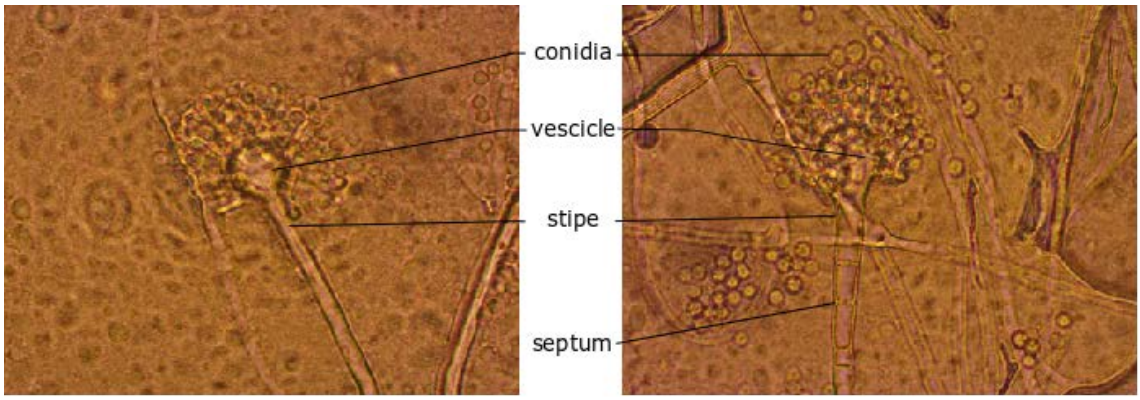


Figure 1. Morphological structures of the conidiophores and conidia of *Aspergillus californicus* IBT 16748. (Picture was taken using a Motic BA310E Microscope with 1000× magnification. Samples were collected after 9 days growing on minimal medium.)

Aspergilli are widely used organisms for the industrial production of organic acids and enzymes, and they are also famous for the ability to produce various secondary metabolites (Figure 2).[8] Generally, they show good tolerance to acidic conditions and a broad spectrum of substrates because many are saprophytic organisms, and they possess a natural ability for high enzyme secretion of hydrolytic enzymes like glucoamylase and α-amylase produced by *A. niger* and *A. oryzae*. Specifically, *A. niger* is applied for the commercial synthesis of citric acid because of its better production yield and it produced 1.6 million tons of citric acid in 2007.[9] In the pharmaceutical industry, *A. niger* and *A. ochraceus* are used in cortisone production.[10] *A. terreus* produces lovastatin which helps to reduce blood cholesterol.[11] Echinocandin B, first isolated from *A. spinulosporus* (= *A. delacroxii*, formerly *Aspergillus nidulans var. echinulatus*), is a lead compound for the antifungal agent anidulafungin (Figure 3).[12,13] In the food industry, *A. oryzae* can be used for producing sake, soy sauce, and miso, and *A. niger* is used in the bread and beer-making industries.[9,10] *A. oryzae and A. egypticus* are used in the production of Douchi, a fermented soybean product.[14,15] However, Aspergilli also show negative aspects. Some species from section *Flavi* can produce aflatoxins and orchratoxins which can be carcinogenic.[16] *A. fumigatus* is the most common prevalent airborne fungal pathogen causing severe invasive infections to humans.[17] Members of *Aspergillus* section *Nigri* are known as destructive degraders

of foods and feeds.[18] Due to the significance of *Aspergillus*, an *Aspergillus* Secondary Metabolites Database (A2MDB) has been recently established.[19]

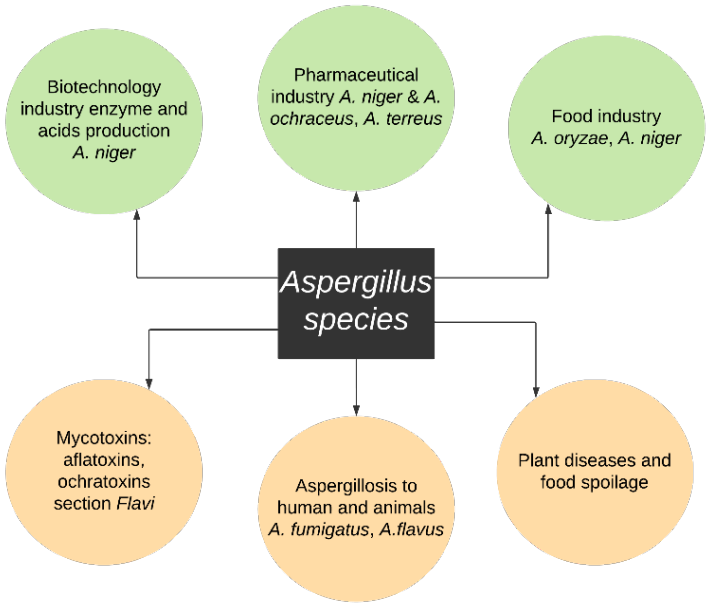


Figure 2. Overview of positive applications and negative aspects of *Aspergillus* species

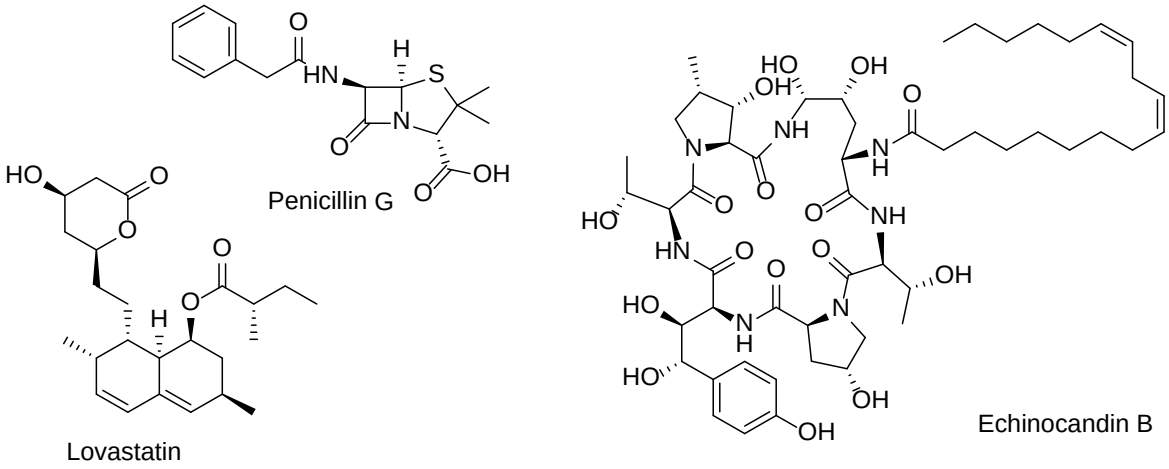


Figure 3. Structures of selected key natural products from fungal sources

1.3 Natural products

The biochemical pathways of fungi follow the central dogma of molecular biology where, typically, DNA is transcripted into RNA and further translated to proteins some of which are important enzymes for the production of low-molecular-weight metabolites in primary and secondary metabolisms (Figure 4). Unlike primary metabolites which are almost universally distributed in the intermediary metabolism, secondary metabolites (SMs), also known as natural products (NPs), are only found in specific organisms and under specific conditions. Generally, SMs are not essential for basic life processes as a single living organism like growth and reproduction, but they benefit the well-being of their producer in many other different ways.[20–22] For example, triterpenoid saponins help plants in resistance against fungi because of their ability to complex with sterols in fungal membranes leading to the loss of fungal membrane integrity.[23] Toxic products provide defense against pathogens and predators while volatiles and colorants attract or warn other species.[22]. A recent study showed that soil-smelling terpenoids geosmin and 2-methylisoborneol emitted by *Streptomyces* colonies facilitate the dispersal of *Streptomyces* spores by attracting soil arthropods.[24] Despite various ecological functions in the environment, NP research has also been driven by the findings that they are important and valuable agents that can potentially be developed as pharmaceuticals, herbicides, insecticides, etc. Globally among all the newly approved drugs between 1st January 1981 and 30th September 2019, 49.2% were NP related including unaltered NPs, botanical drugs, NP derivatives, NP mimics, and those synthetic drugs with an NP scaffold. When it comes to small-molecule approved drugs, the figure rises to 66.7%.[25]

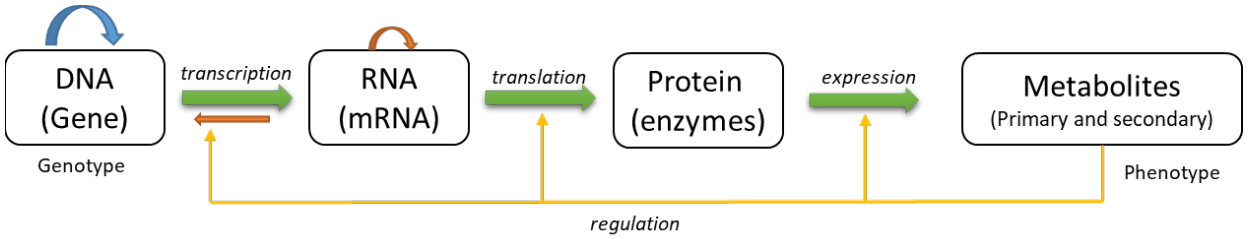


Figure 4. General biochemical pathways in fungi, inspired by ref[21,26].

Based on the strategies and methodologies used for NP discovery, the history of NP research can be segmented into three overlapping stages: 1. the 1940s - 1970s, phenotypic screening. In this period, a phenotypic screen employs a cell line (bacteria, yeast, eukaryotic cells or tissues in culture, etc.) and a method to determine a response to an applied test compound without prior knowledge of its mechanism of action (MOA). More than 1000 NPs with antibacterial or antifungal activities have been discovered including dozens that were eventually approved as drugs; 2. the 1970s - 2000s, the knowledge-based approaches. Target-based screening including biochemical and whole-cell assays as well as selective phenotypic screening methodologies was vastly used during this time; 3. the 2000s and beyond, genomics-based strategies. With the inexpensive sequencing techniques, advances in understanding SM biosynthesis, and improvement in analytical instruments and bioinformatics tools, it is possible to predict the partial structures of some new SMs which are related to known SMs. It is now estimated that only less than ten percent of secondary metabolites gene clusters (SMGCs) are expressed in sufficient quantities to be observed under routine fermentation analyses, and many SMGCs require special culture conditions and/or genetic manipulations to reveal their products.[11] Thus, there´s a huge potential to find a lot more new SMs by realizing the whole potential of the underlying silent or cryptic SMGCs.

Since the discovery of penicillin in 1929, the number of published NPs was estimated close to half a million (between 300 000 to 600 000) including 60 000 - 80 000 microbial metabolites.[20] In terms of originals of NPs, only a small fraction of 250 000 - 300 000 living species have been currently investigated, identified, and deposited compared to at least 1.5 million and nearly half a million fungal and higher plant species existing on the earth.[27] Of all known microbial metabolites nearly half (47%) exhibit some kind of biological activity, which is much higher than those plant- (7%) and animal-derived (3%) compounds. The largest group of reported microbial metabolites was produced by various fungi (45%), and filamentous fungi including *Aspergillus*, *Penicillium,* and *Trichoderma* and hundreds of other species represent almost 99% of all fungal metabolites.[27] A recent review showed that from January 2015 to December 2019, 187 products out of 362 secondary metabolites isolated from 17 endophytic *Aspergillus* species exhibited

various biological activities.[7] Thus, filamentous fungi from the genus *Aspergillus* are an important source for bioactive natural product discovery.

1.4 Natural products from a biosynthetic perspective

1.4.1 Polyketides

Polyketides (PKs) are a class of metabolites with various structures and bioactivities. One well-known polyketide is lovastatin, one of the most successful cholesterol-lowering agents on the market. Bacterial polyketides such as the tetracycline antibiotics or the daunomycin-inspired doxorubicin anticancer agents have been prescribed for decades. Notorious polyketides like aflatoxin B1, patulin, fumonisin B1, and zearalenone mycotoxins (Figure 5) produced by *Aspergillus* and *Fusarium* fungi have negative impacts on agriculture and human health.[28,29] The biosynthesis of PKs is similar to the fatty acid biosynthesis as they both use the same precursors acetyl-coenzyme A (CoA) and malonyl-CoA and extend chain length through Claisen Condensation. But the polyketide syntheses (PKSs) which catalyze the biosynthesis of PKs differ from fatty acid synthases (FASs) in three main ways, 1) alternative starter and/or extender units, 2) various chain lengths, and 3) flexible reductive steps. Based on the architecture and enzymatic assembly lines, PKSs can be distinguished as three types: Type I, II, and III (Table 1).

Type I PKSs can be further divided into modular and iterative type I PKSs. Modular type I PKS (mPKS), typically found in prokaryotes, contains multiple repetitions (modules) of the same active sites and each module is for single-use only. Classic *cis*-AT type I PKSs are linearly arranged and covalently fused catalytic domains within huge multifunctional enzymes. For each module, there are at least three functional domains: ketosynthase (KS), acyltransferase (AT), and acyl carrier protein domain (ACP). The AT domain acts as a ´gatekeeper´ for which substrates the enzyme uses and dominates the number of the keto units in the final product. In addition to the three minimal domain architectures, PKS modules can harbor other optional functional domains including the ketoreductase (KR) domain for the reduction of the β-keto function to a hydroxyl group, dehydratase (DH) domain that catalyzes the dehydration of the hydroxyl group to yield an α,β-double bond, and enoyl reductase (ER) domain that further reduces the double bond to a

fully saturated acyl structure. Actinomycetes can provide substituted malonyl-CoA as an extender unit while many other bacteria, including mycobacteria and cyanobacteria, use methyltransferase (MT) domains to generate α-methylated branches. In addition to the *cis*-AT PKS, the second type of modular system named *trans*-AT PKS has also been discovered. This type of PKS exhibits unique characteristics including a free-standing AT architecture, unique domains, and unusual domain orders and actions. A study revealed that 38% PKS genes of all sequenced bacterial genomes belong to trans-AT PKS.[30–32]

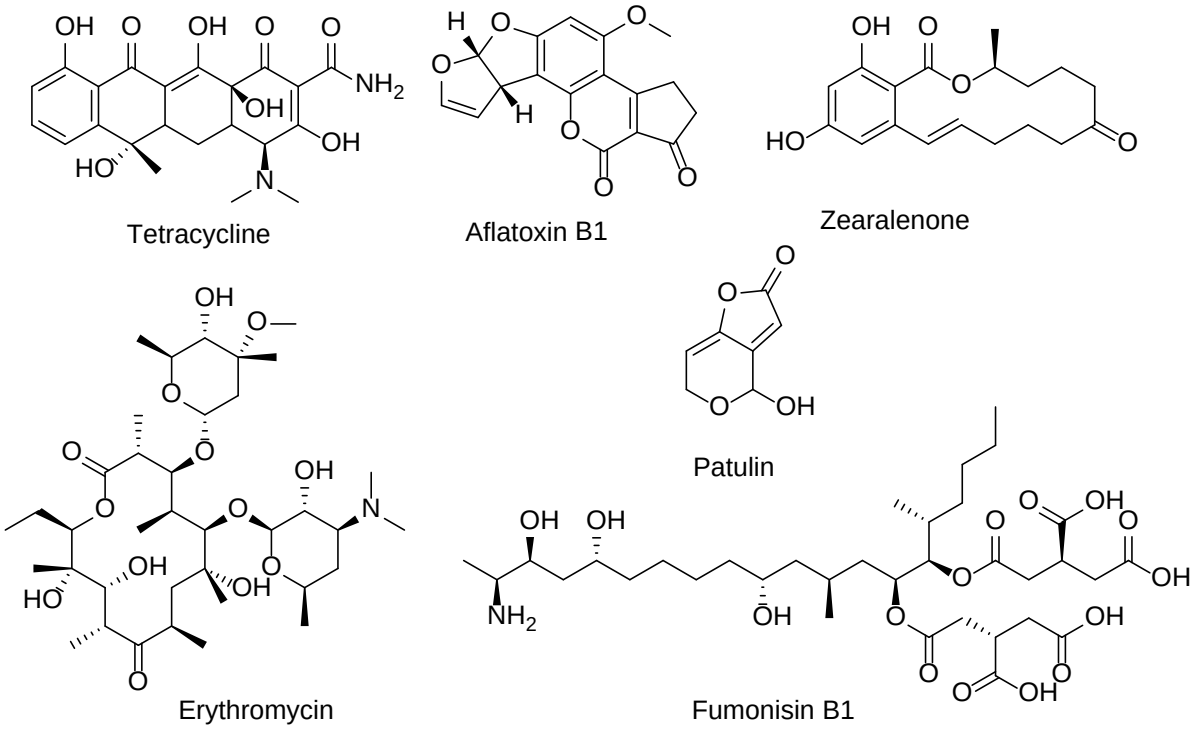


Figure 5. Structures of selected polyketides.

Iterative type I PKS (iPKS), typically found in fungi, consists of a single copy of each active site that can be used multiple times during the polyketide chain growing. While initially thought to exist only in fungal systems, iterative type I PKSs have also been found in many bacteria.[33] During the elongation process, domains like KR, DH, ER, and MT (if any) are optionally used in every extension round resulting in the high diversity of final products. Based on different reductive behaviors of β-keto groups, type I iPKS can be further distinguished as three subtypes: non-

reducing (NR), partially reducing (PR), and highly reducing (HR).[30,34] Figure 6 shows the schematic biosynthetic steps of fungal iPKS.

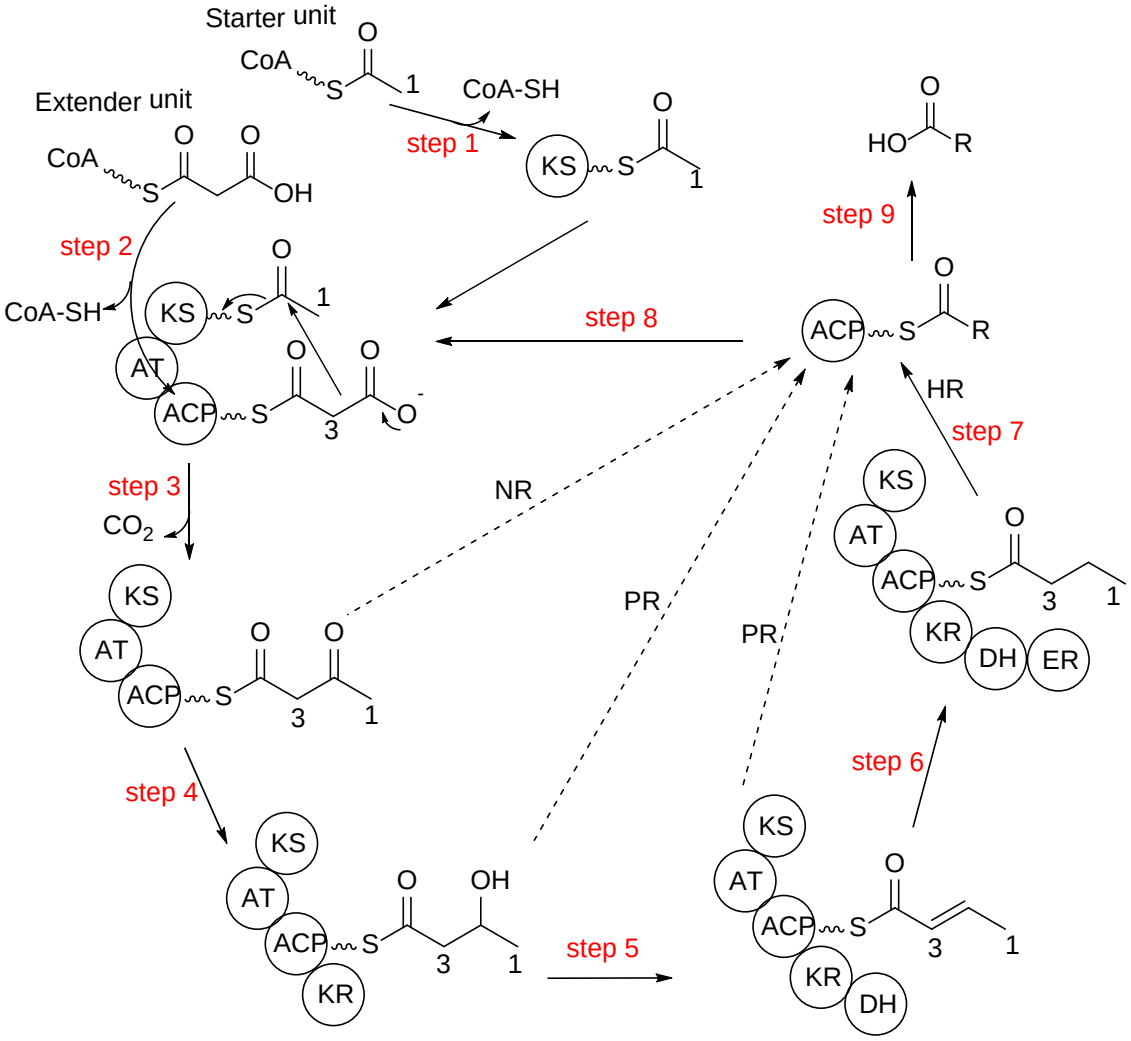


Figure 6. Schematic flow of biosynthesis of polyketides by iterative type I PKS.

Step 1: loading of a starter unit acetyl-CoA; step 2: loading of an extender unit malonyl-CoA; step 3: decarboxylative, Claisen-like condensation; step 4: reduction of previous keto group to form an alcohol group by KR; step 5: dehydration of alcohol group by DH; step 6: enoyl reduction of the double bond by ER; step 7: fully reduced intermediate/compound ready for another round or the final release; step 8: transfer to another elongation cycle; step 9: final product release, modified from ref[35].

Type II PKSs are quite common in bacteria like actinomycetes, and they consist of a dissociable complex of discrete and usually mono-functional enzymes which can be usually iteratively. A

minimal Type II PKS includes KSα, KSβ, and ACP. Type III PKSs are a single enzyme consisting of a homodimer of two KS domains without ACP or AT since the starter and extender units are transferred directly back and forth between CoA and KS. Type III PKSs can be found in plants, bacteria, and several fungi.[30,31,33]

Table 1. Classifications of polyketide syntheses.

PKS types	Selected compounds	Organisms
Modular type I PKS, cis-AT	Erythromycin, Avermectin	Typically in bacteria
Modular type I PKS, trans-AT	Mupirocin, Streptogramins	Bacteria
Iterative type I PKS (NR, PR, HR)	Blennolide (NR), Norsolorinic acid (NR), Zearalenone (HR/PR+NR), Lovastatin (PR)	Mainly fungi, also bacteria
Iterative Type II PKS	Actinorhodin	Bacteria
Interactive type III PKS	Alkylpyrone[36], Csypyrone B1[37]	Mostly plants, bacteria, and fungi
Hybrids	Cytochalasans[38], PTMs*[39]	Fungi and bacteria

* PTM: Polycyclic tetramate macrolactam.

1.4.2 Nonribosomal peptides

Nonribosomal peptides (NRPs) are derived from proteinogenic amino acids (AAs) and/or non-proteinogenic AAs synthesized by nonribosomal peptide synthetases (NRPSs). They represent an important class of pharmaceutically relevant drugs. NRPSs are multi-modular enzymes or enzyme complexes from bacteria and fungi and are capable of producing a large variety of structures, several of which are used clinically including penicillin, cyclosporin, vancomycin, and daptomycin. They harbor a modular architecture where each module consists of different catalytic domains. A minimal NRPS elongation module consists of three domains including an adenylation (A) domain, a condensation (C) domain, and a peptidyl carrier protein (PCP) domain, the latter was also referred as 4′-phosphopantetheinylated thiolation (T) domain.[40]

A putative biosynthetic pathway of Echinocandin B is drawn below (Figure 7) to show the detailed biosynthetic steps of NRPs.[41,42] In the peptidyl chain initiation, an AA is activated as the aminoacyl-AMP mixed anhydride on an ATP-dependent manner (where synthetases are from). A domain controls the AA selection, and the aminoacyl group was bounded covalently to the P-

pantetheine thiolate arm on the PCP domain. During the elongation process, a new peptide bond is formed catalyzed by the C domain. Two common optional reactions can happen in this step. One is the epimerization of L-aminoacyl- and L-peptidyl thioesters to the corresponding D-residues before transferring to the next module leading to subclassification of C domains into $^{L}C_{L}$ and $^{D}C_{L}$ types. The other usual modification is N-methylation catalyzed by nitrogen methyltransferase (*N*-MT) domains during NRPS assembly line elongation steps. The appropriately processed intermediate peptide is then passed to the next module(s) in line for further elongation and modification before being released from the assembly line.

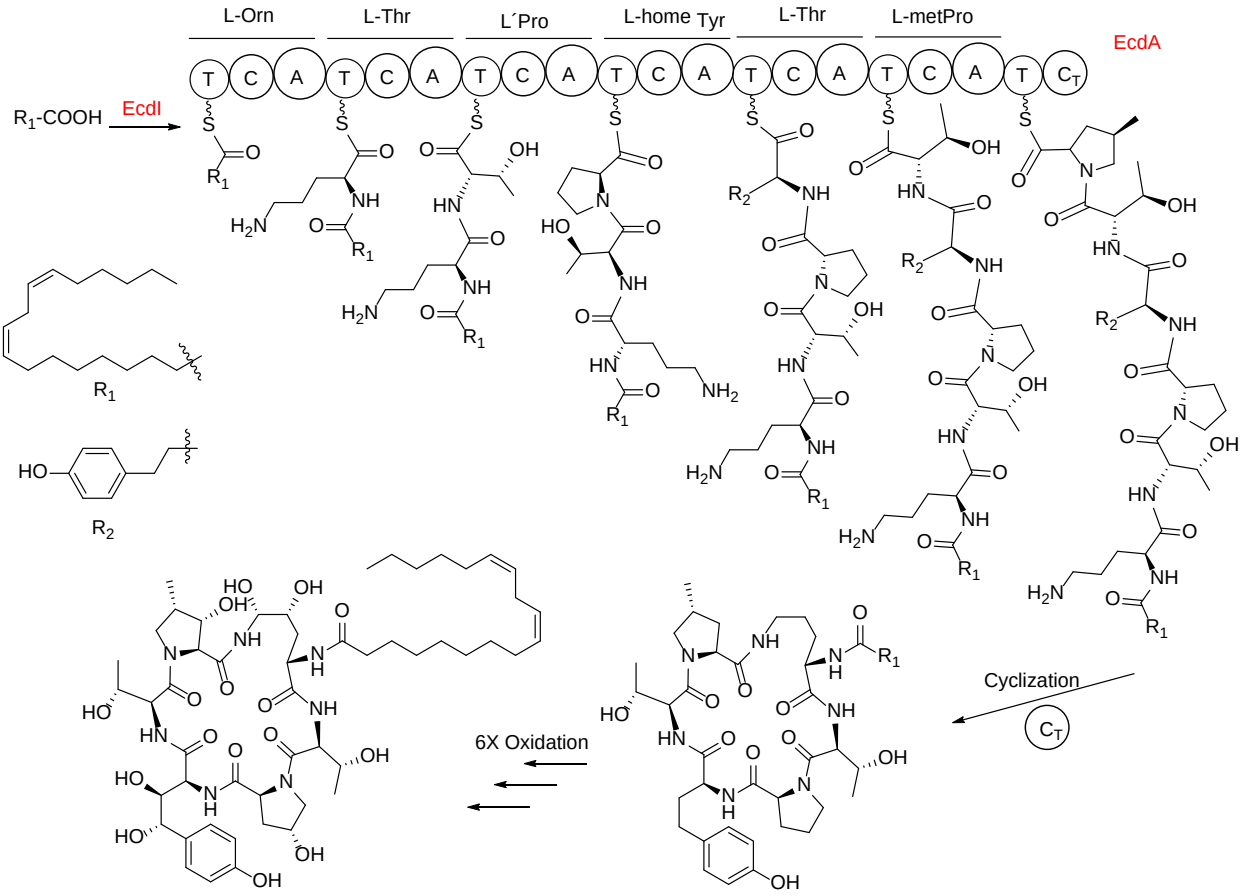


Figure 7. Putative biosynthetic pathway of Echinocandin B.[41,43] (T: thiolation domain; A: adenylation domain, C: condensation domain, C_T: terminal condensation domain)

Several other additional structural adjustments may occur after chain extensions such as the formation of the heterocyclic rings by heterocyclization domains and oxidations by oxidation

domains. Many bacterial NRPS have a thioesterase (TE) domain in the last module to act as a release catalyst releasing a free acid. On the contrary, in fungi, many NRPS harbor a terminal condensation (C-term) domain that mediates chain release yielding conformationally constrained macrocycles. In addition to these two releasing mechanisms, one NADPH-mediated terminal reduction route is found in the yeast lysine biosynthetic pathway yielding an aldehyde product. Some tailoring enzymes are often encoded in the biosynthetic gene clusters (BGCs) to provide unusual AAs as building blocks, tailor elongating chain on the assembly lines, or function as post-assembly line tailoring catalysts.[40,44,45]

1.4.3 Terpenoids

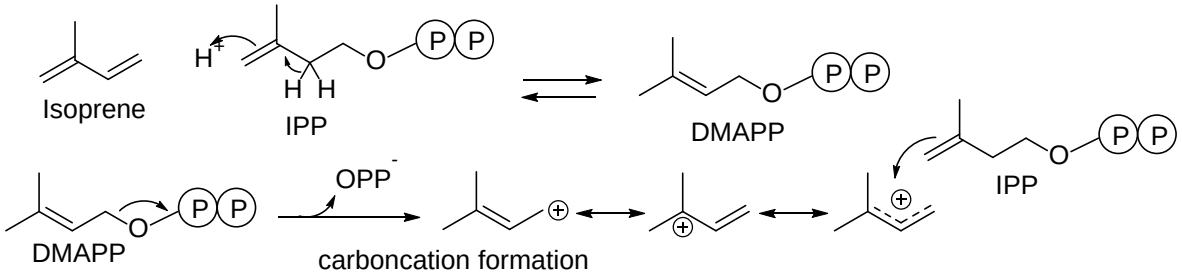


Figure 8. Building blocks of terpenoids, modified from ref[46].

Terpenoids, also called isoprenoids, form a large group of most diverse NPs, and they are composed of several isoprene units that are connected in a head-to-tail way and further modified by cyclizations. Based on the number of isoprene units, they can be classified as hemiterpenes (C_5), monoterpenes (C_{10}), sesquiterpene (C_{15}), diterpenes (C_{20}), sesterterpenes (C_{25}), triterpenes (C_{30}), and tetraterpenes (C_{40}). The best-known terpenes are odoriferous plant metabolites such as menthol, turpentine, and camphor, but bacteria and fungi also synthesize many terpenoids such as geosmin, aristolochenes, carotenoids, gibberellins, and trichothecenes. So far around 50 000 terpenoid metabolites have been isolated from plants, liverworts, fungi, and bacteria.[20,47] The building blocks for terpenoids are two biochemically active isoprenoid units: isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) which may be derived from mevalonic acid (MVA) or methylerythritol phosphate (MEP) pathways. DMAPP harbors a good leaving group, the diphosphate, and can ionize easily to produce an allylic carbocation which is

stabilized by charge delocalization. IPP with its terminal double bond is more likely to act as a nucleophile (Figure 8).[46]

1.4.4 Others

Meroterpenoids. Many NPs contain partial terpenoid elements in their molecules. A particularly common terpenoid fragment is a single C_5 unit, usually a dimethylallyl substituent. Compounds containing these isolated isoprene units are referred to as ´meroterpenoids´. Fungal meroterpenoids exhibit diverse chemical structures and a broad spectrum of bioactivities. For example, polyketide-terpenoid hybrid fumagillin, isolated from *Aspergillus fumigatus*, is used for the treatment of microsporidiosis and amebiasis. Terpendole E and lolitrem B both contain an indole ring and the latter is a potent tremorgenic neurotoxin.[46,48]

PK-NRP hybrids. These hybrid molecules contain both PK and NRP chemical elements which are synthesized by PKSs-NRPSs. The enzyme modules are usually connected in two ways. The first connection is typically found in bacteria that the PKS modules and NRPS units are located together (in either order) in the processive assembly line. One example is found in the biosynthesis of the antibiotic bacillaene. BaeJ comprises of an unusual starter module followed by an NRPS module. Two PKS modules are located after the NRPS domain with different architectures.[49] The second way is found mostly in fungi that an iPKS module is followed by a single NRPS module and a releasing domain. Many examples were well summarized including the *fusA* in the biosynthesis of fusarin C, *apdA* in the production of aspyridone A, and cheA in the biosynthesis of cytochalasans, etc.[39,50–52] In the course of the hybrid biosynthesis (Figure 9A), the PKS module synthesizes the polyketide substructure and the NRPS unit selects and activates a specific AA by A domain, which passes the aminoacyl residue onto the PCP domain. After the completion of the polyketide substructure, the C domain in the NRPS module catalyzes the fusion of the polyketide part to the activated AA residue to yield a new amide.[39,50] Recently, a rare fungal NRPS-PKS hybrid enzyme TAS1, encoding the mycotoxin tenuazonic acid, has been identified that begins with an NRPS module and only contains KS domain in the PKS module (Figure 9B).[53]

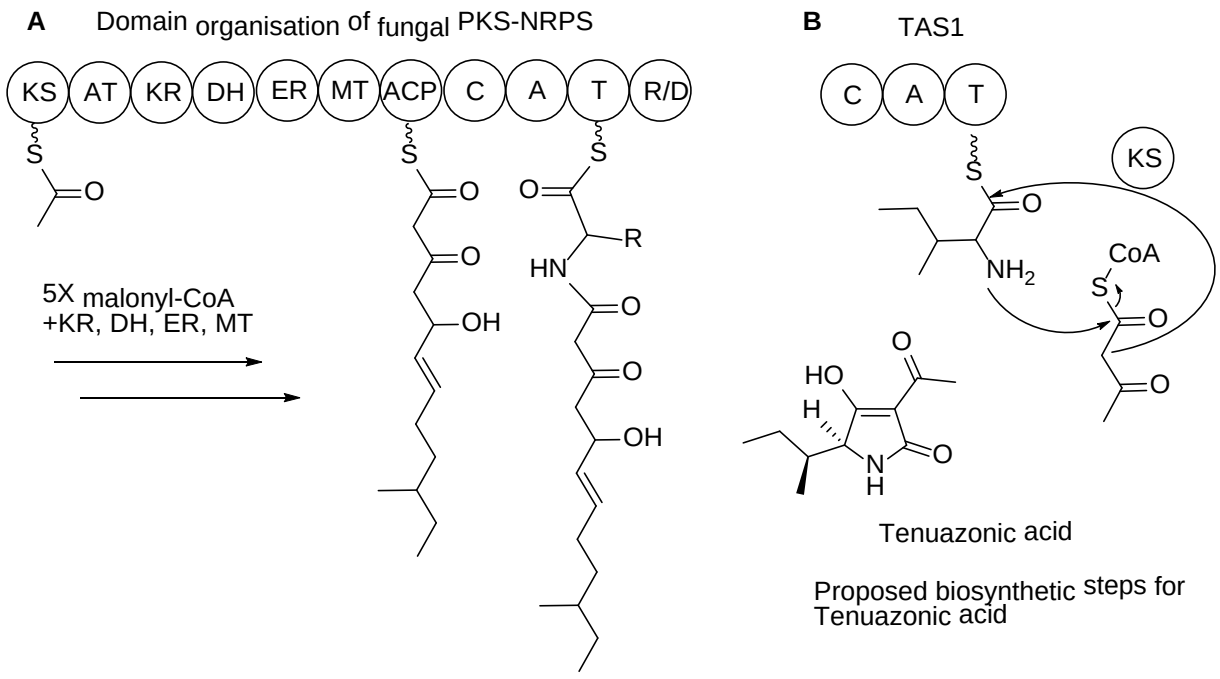


Figure 9. Domain organizations of fungal PKS-NRPS (**A**) and NRPS-PKS (**B**), modified from ref[50,53].

1.3 Natural product discovery techniques

1.3.1 High-performance liquid chromatography (HPLC)

The components in the extract from organisms/substrates are always complex and require multiple separations to obtain a pure natural product. The separation techniques are based on the physical or chemical properties of the target products including absorption property, partition coefficient, molecular size, ionic strength, etc. Chromatography, especially column chromatography, is the major separation method used in natural product discovery.[54] HPLC is a very powerful chromatographic technique for the separation and purification of NPs from matrices such as crude extracts.[55] Most LC separations are performed with liquid-solid isolation techniques such as various forms of planar chromatography and column chromatography. The use of preparative HPLC has become a mainstay in the isolation of most NPs. A typical HPLC system consists of five major parts including a pump, an injector, a column, a detector, and a data acquisition system (Figure 10). The pump delivers the mobile phase, and the sample injector/sampler injects the sample to the MP stream before the column. The column is the core

part of the whole system where different components in the sample mixture get separated based on different behaviors with the column matrices/stationary phase. The component with a shorter retention time gets detected earlier by the UV-vis detector and/or MS detector.

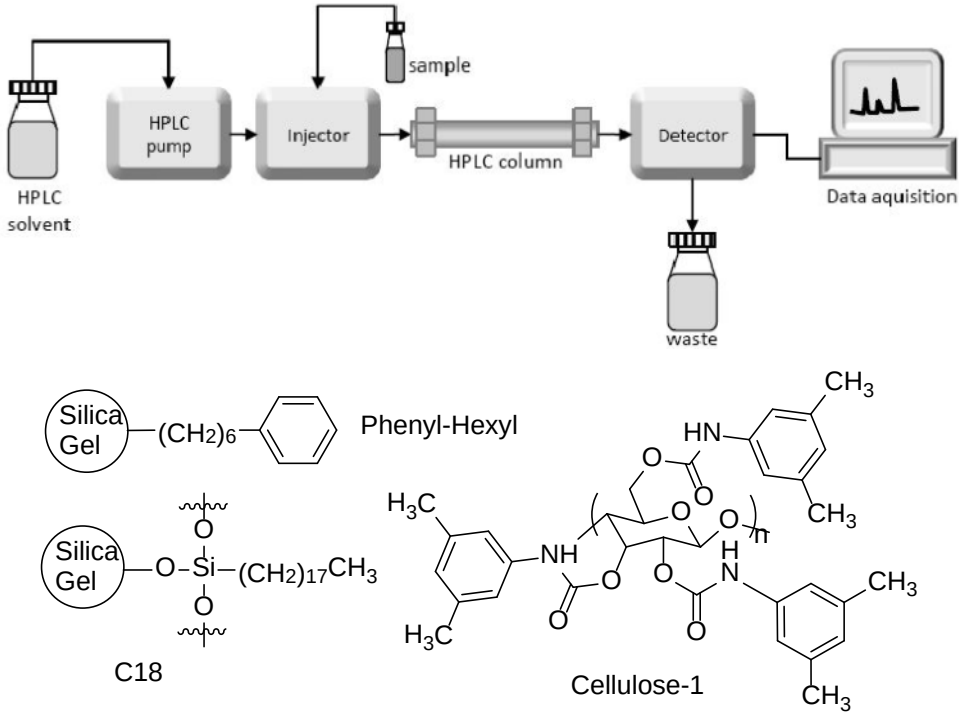


Figure 10. Illustration of components of HPLC (top)[56], and common column matrices (bottom).

Based on the separation mechanism there are different modes of LC such as reversed-phase, normal-phase, ion-exchange, size-exclusion chromatography, etc. Currently, octadecyl silane (RP-C18) columns (Figure 10) are the most widely used in the NP isolation and purification. Enantioseparation is also an application of HPLC. The most widely used principle for direct separation is based on enantioselective complexation in cavities of a chiral selector.[57] One useful chiral separation column is the cellulose-1 phase. The cellulose phenyl carbamate derivative is essential to any chiral separation. The linear structure of the cellulose adds grooves and cavities for steric interactions, and the selector constituents offer hydrogen bonding and π-π interactions that affect the chiral stationary phases ability to discern and separate the enantiomers.[58]

1.3.2 Mass spectrometer (MS)

Since the first mass spectrometer (MS) was constructed in 1912, this technique has been a powerful analytical tool for both quantitative and qualitative analysis of small organic molecules and biological macromolecules. MS can also be interfaced with other separation techniques, such as HPLC (HPLC-MS) and gas chromatography (GC-MS). A typical mass spectrometer consists of three major components: an ion source, mass analyzer, and detector.[59]

In the ion source, the analyte is ionized and transferred to the mass analyzer by magnetic and electronic fields. The traditional ionization method electron ionization (EI) utilizes an energetic electron beam, and the accelerated electrons collide with the analytes in the gas phase causing electron expulsion of the analytes and subsequent formation of positively charged radical cations. EI yields a large degree of fragmentation of the analytes and it is called hard ionization. Besides, there are also soft ionization methods such as chemical ionization (CI), fast atom bombardment (FAB), atmospheric pressure chemical ionization (APCI), electrospray ionization (ESI), and matrix-assisted laser desorption ionization (MALDI). Since the introduction of ESI and APCI, HPLC was able to couple to MS.[60] ESI is now the most commonly used ionization mode in the HPLC-MS system.[61] In the ESI source, HPLC eluent enters the spray chamber at atmospheric pressure through a nebulizer which is concentrically surrounded by a tube and the nebulizing gas enters the spray chamber from the tube. The sample solution breaks into droplets by the nebulizing gas and strong voltage (2-6 kV). As the droplets disperse, charges preferentially migrate to the surface due to electrostatic forces. As a result, the sample is simultaneously charged and dispersed into a fine spray of charged droplets. A countercurrent of neutral, heated drying gas, typically N_2, evaporates the solvent and decreases the droplet size. The droplet explodes when the force of the Coulomb repulsion equals or exceeds that of the surface tension thus producing smaller charged droplets. This process repeats itself till droplets with a high density of surface charges are formed. Gas-phase ions can be ejected from the droplet surface based on ion evaporation theory, or the droplet explosion/fission can continue until gas-phase ions are formed based on the charged residue theory.[59,62,63]

A mass analyzer is a part where ions are separated based on their mass-to-charge ratio (*m/z)* values. The isolation is usually driven by electrical and/or magnetic field. There are four main analyzer systems widely used including quadrupole (Q), quadrupole ion trap (QIT), time of flight (ToF), and Fourier transform ion cyclotrone resonance (FT-ICR). The Q analyzer is composed of four parallel electrical rods. A direct current (DC) potential (U) is applied to two of these rods, and the other two are linked to an alternating radio-frequency (RF) potential (V, frequency ω). Ions formed in the ionization chamber are pulsed towards a quadrupole by an electrical field and undergo a complex oscillation trajectory in the circular cross-section formed by four rods. With the appropriate values of V, U, and ω, only ions within a narrow range of *m/z* will survive the path towards the detector. The remaining ions will eventually collide with one of the rods. A ToF analyzer relies on the free flight of the ions in a tube of 1-2 m in length. An electric field was applied to accelerate the ions through the same potential, and the time they take to reach the detector is measured. According to the equation $E=1/2mv^2$ where the kinetic energy (E) is identical (as all ions are singly charged in the same electric field), m and v represent the mass and velocity of the ions respectively, ions with lower mass have a higher velocity thus reaching the detector earlier than the heavier ions.[59,64]

For the identification of unknown NPs, tandem MS (MS/MS or MS^n) is often used. The most successful tandem MS systems include the triple quadrupole (QQQ), quadrupole-time of flight (Q-TOF), quadrupole ion trap (QIT).[60] Figure 11 shows an Agilent Q-TOF LC/MS system that performs MS/MS using a quadrupole, a hexapole collision cell, and a time-of-flight analyzer to produce spectra. The quadrupole selects precursor ions are next fragmented in the collision cell into product ions, and the product ions are then impelled to the flight tube, at an angle perpendicular to the original path. The ions accelerated by the ion pulser travel through the flight tube. At the top end of the flight tube is an ion “mirror”, which reflects the ions to the bottom end with forward momentum to the ion detector.[62,63,65]

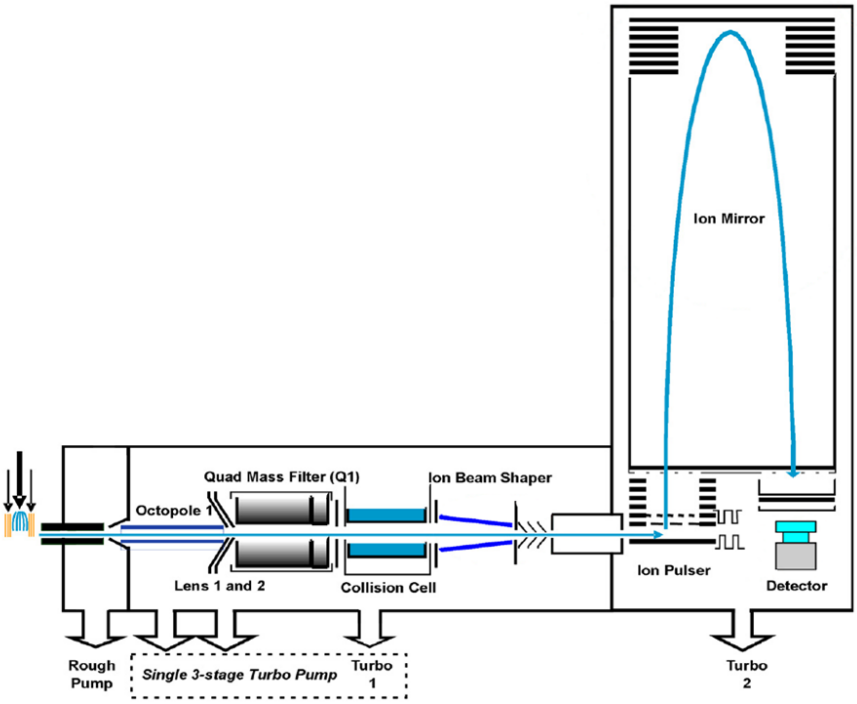


Figure 11. Schematic of Q-TOF MS/MS system, picture from ref[62].

1.3.3 Nuclear magnetic resonance (NMR) spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy has become an extremely powerful technique and has been very widely used for the structure elucidation of NPs since 1984 as there have been a lot of advances in NMR methodology and hardware.[66] Many atomic nuclei with either odd mass, odd atomic number, or odd both have a quantized spin angular moment and a magnetic moment. The number of the allowed spin state is $2I+1$, in which I, the spin quantum number, is a physical constant of an atom. The spin quantum number of a proton is 1/2 indicating it has two spin states. Proton also has a magnetic moment μ generated by its charge and spin. In the presence of an applied magnetic field B_0, the proton will either have an aligned (+1/2, low state) or opposed (-1/2, high state) spin. There are always slightly more nuclei in the low spin state than those in the high spin state. For protons in a magnetic field of 1.41 Tesla, there are 9 nuclei more in the lower spin state (excess nuclei) for every million protons occupying a higher spin state. In the applied magnetic field, the nuclei precess about its axis of spin with angular frequency ω (Larmor frequency), generating an oscillating electric field of the same frequency. If the precession protons are supplied with a radiofrequency of this frequency, energy can be transferred from the

incoming radiation to the nucleus, causing a spin change. This condition is called resonance. The excess nuclei allow people to observe the resonance. Not all protons in a molecule have resonance at the same frequency as different protons exist in a slightly different electronic environment. In the presence of B_0, the valence electrons are caused to circulate generating diamagnetic shielding. Shielded protons precess at a lower frequency and resonate at a lower radiofrequency. Thus different nuclei give different values of chemical shifts δ in the NMR spectrum which are diagnostic. The resonance of a proton can be split by the spin states of its neighbor proton(s) and this phenomenon is called spin-spin coupling which can be measured by coupling constant J. The splitting multiplicity pattern follows the $2nI+1$ rule. The principle of ^{13}C NMR is similar to ^{1}H NMR but more difficult to observe due to its low natural abundance and small magnetogyric ratio. Modern NMR spectrometer uses a decoupler to saturate the neighboring protons of ^{13}C obliterating all the interactions between ^{1}H and ^{13}C. In the ^{13}C NMR spectrum, the intensities of carbons with hydrogens directly attached increase due to nuclear overhauser enhancement (NOE).[67]

In two-dimensional NMR spectroscopy (2D NMR) data are plotted in a space defined by two frequency axes. The commonly used 2D NMR correlations are drawn in Figure 12. The first homonuclear sequence is ^{1}H-^{1}H Correlation Spectroscopy (COSY) the chemical shift range of the proton spectrum is plotted on both axes. The equivalent of a 1D spectrum appears along the diagonal of the 2D plot from the lower left to the upper right. The important signals in the spectrum are the off-diagonal/cross-peaks which represent two groups of protons on two adjacent carbons.[68] Another homonuclear sequence is Nuclear Overhauser Effect Spectroscopy (NOESY) which is used to determine the signals of two groups of protons that are close to each other in space even if they might not be bonded to adjacent carbons. The spectrum appearance is similar to a COSY spectrum both with diagonal peaks and cross-peaks. However, the cross-peaks in NOESY indicate resonances from nuclei that are spatially close (less than 5 Å) rather than those that are through bonds coupled to each other. NOESY experiment is an important tool to identify the stereochemistry of NPs.[66]

Heteronuclear Single Quantum Correlation (HSQC) is a heteronuclear 2D NMR with one axis for proton and the other for a heteronucleus usually ^{13}C (or ^{15}N). The spectrum contains peaks for the bounded nuclei. HSQC can be obtained in edited mode (ed HSQC) with CH_2 peaks upright and CH and CH_3 peaks inverted.[68] Heteronuclear Multiple Bond Correlation (HMBC) experiment correlates a proton with a heteronucleus usually ^{13}C (or ^{15}N) that are two or three bonds apart from each other. Sometimes four-bond or even five-bond HMBC correlations may be observed.[69] HMBC spectra are important for the detection of quaternary carbons and linking the isolated fragments in structure elucidation. Heteronuclear Two-Bond Correlation (H2BC) almost exclusively correlates protons and proton-bearing carbon spins separated by two covalent bonds as it relies on $^{3}J_{HH}$ coupling constants instead of $^{n}J_{CH}$. Complementary to HMBC, H2BC solves the problem of missing two-bond correlations in HMBC spectra. H2BC experiment is also a very good supplement to the COSY especially when there are overlapping protons in the COSY cross-peaks.[70]

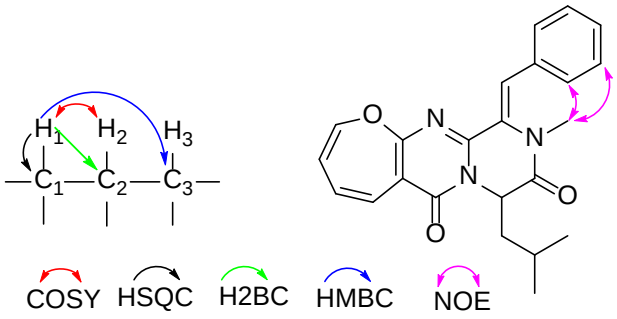


Figure 12. Common 2D NMR correlations used in natural product elucidation.

1.3.4 Chiroptical spectroscopy and electronic circular dichroism (ECD)

Chirality or handedness is an important character in nature and only one type of handedness is preferred by the biological molecules like *L*-amino acids and *D*-sugars. Chirality determination is of extreme importance in drug development to avoid disasters like the Thalidomide Tragedy. Absolute configuration (AC) determination is also a challenging aspect in the structure elucidation of NPs. Recently, chiroptical spectroscopic methods including optical rotation (OR), electronic circular dichroism (ECD), vibrational CD (VCD), and vibrational Raman optical activity

(VROA) are often used in the determination of chiral molecules, and specifically, ECD has been widely used in the NP elucidation.[71–73]

ECD is the difference in the absorption of left-handed circularly polarised light (*L*-CPL) and right-handed circularly polarised light (*R*-CPL) and occurs when a molecule contains one or more chiral chromophores. If a chiral molecule absorbs more *L*-CPL than *R*-CPL, the CD signal is positive, and vice versa. If the absorption of *L*-CPL equals that of *R*-CPL, the CD signal is zero. Recently, ECD calculation by time-dependent density functional theory (TDDFT) has been used commonly in the AC determination of NPs. The principle is computational-calculated spectra to compare with experimental ECD spectra. If the two data sets match each other, the chiral center of an NP is assigned. ECD measurements were recorded in the UV-vis region and ECD calculations generally involve six main steps (Figure 13) including conformational search, geometry optimization, and re-optimization, Boltzmann distribution, TDDFT ECD/UV calculation, Boltzmann-averaged ECD spectrum generation, and correction.[72,74]

The first step conformational search is to obtain the possible conformers of a starting diastereomeric structure at a low computational level, typically molecular mechanics Merck Molecular force field (MMFF). Several programs are available like Spartan and Macromodel.[75,76] The generated conformers are subjected to geometric optimization at the density functional theory (DFT) level using Becke three parameters Lee–Yang–Parr (B3LYP) and the 6-31G(d) basis set using programs such as Gaussian 16.[77] The energies of optimized geometries at this level are compared and the lowest energy conformers within a certain energy range are retained for further geometry optimization at a higher level of theory. Boltzmann distribution is then applied to the conformers obtained after re-optimization. Populated conformers above a certain threshold are selected for the TDDFT ECD/UV calculation, which is the most computationally demanding step in the whole process. The accuracy of TDDFT calculations depends mainly on the functional and basis set selected. In ECD calculations, the B3LYP functional and the basis sets such as 6-31G* or aug-cc-pVDZ are commonly used and give satisfactory results.[72] The results of TDDFT calculations are excitation energies, and their corresponding oscillator strength and rotational strength. The oscillator strengths are used to simulate the UV curve, and rotational strengths are

to simulate the ECD curve. The Boltzmann-averaged ECD spectrum is generated by the ECD curves of the populated conformers and their Boltzmann weight. UV correction/wavelength shift may be needed to take the systematic error/underestimation of transition energies into account. The corrected averaged spectrum is then ready for comparison with the experimental curve.[74]

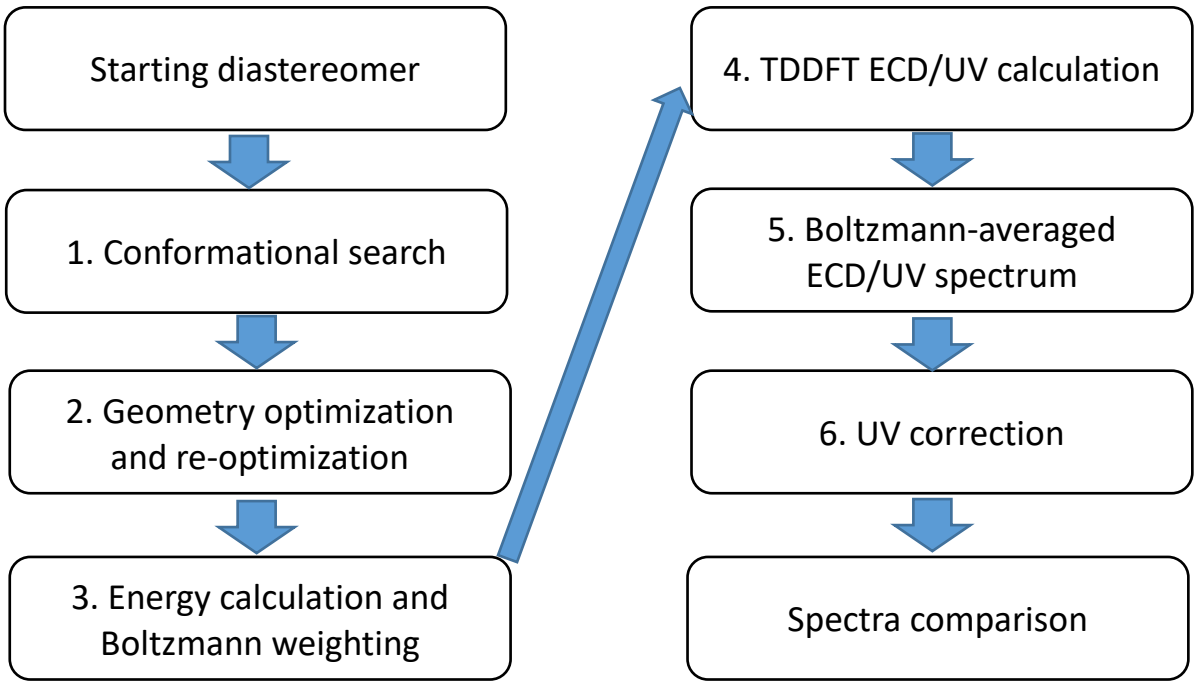


Figure 13. Main steps using ECD calculation to solve the absolute structure of natural products.

1.4 Genome editing technologies

Along with the completion of the human genome project and advances of high-throughput sequencing technologies, programmable genome editing systems with engineered nucleases like meganucleases, zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have been developed. ZFNs and TALENs fuse a locator (zinc finger (ZF) motifs and transcription activator-like effector (TALE) repeats) to a heterogeneous *Fok*I endonuclease as the effector for genome editing. However, both systems have their limitations like the crosstalk challenge between adjacent ZF motifs and the complicated construction of TALEN vectors.[78]

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein (Cas) systems are RNA-mediated immune system in bacteria and archaea that protects

them against incoming bacteriophage infection and horizontal plasmid transmission. Without fusion to a heterogeneous endonuclease, CRISPR-Cas systems have both binding (locator) and cleavage (effector) activities. Class 2 type II CRISPR-Cas9 technology has been developed and expanded for editing the genome of a wide range of organisms including bacteria, yeast, and animals.[79,80] CRISPR-Cas9 targets the double-stranded (ds) DNA using the nuclease Cas9, a CRISPR RNA (crRNA), and trans-activating CRISPR RNA (tracrRNA), with the latter two fused as the chimeric sgRNA. Cas9 binds to the DNA sequences complementary to the sgRNA spacer (20-nucleotide) which was 3-5 bp upstream of the 3-nucleotide Protospacer Adjacent Motif (PAM). With the correct base pairing, the Cas9 activates its two individual endonuclease domains HNH and RuvC to cleave the target (complementary to the spacer region of sgRNA) and nontarget (cognate to the spacer region of sgRNA) DNA strands yielding the specific DNA double-strand breaks (DSBs) (Figure 14).[81,82] Advances on CRISPR based genome editing tools are evolving fast these days, and new genome editors like CRISPR-CasΦ are being developed.[83–85]

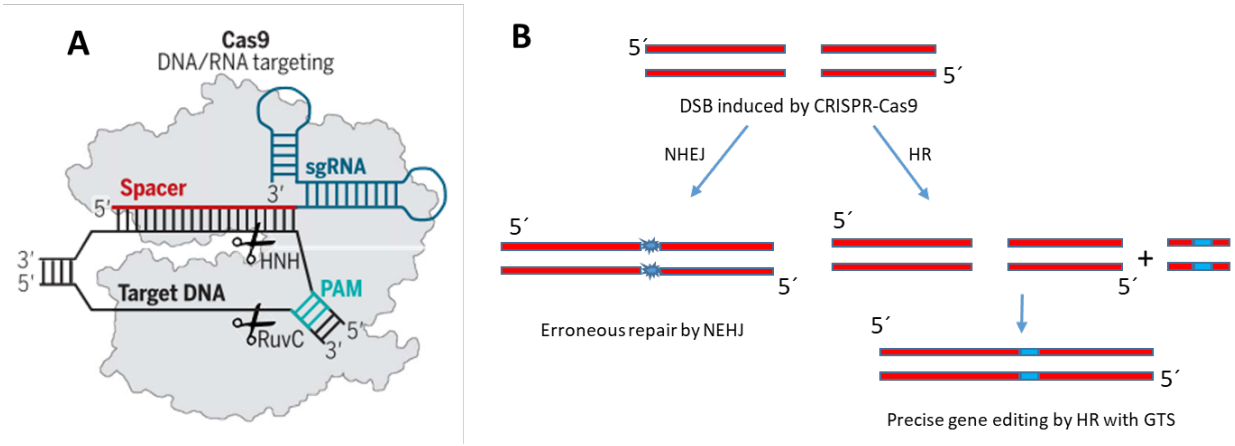


Figure 14 (**A**) Schematic of CRISPR-Cas9 from ref[82], and (**B**) two DNA repair strategies inspired by ref[86]. (DSB: double-strand breaks, NHEJ: nonhomologous end-joining, HR: homologous recombination, GTS: gene-targeting substrates like PCR fragment or oligonucleotide)

Recently, CRISPR-Cas9 has been adopted as a genetic tool in engineering the filamentous fungi *Aspergillus*, and with four AMA1-based CRISPR-Cas9 vectors Pfc330-333 successes have been achieved on several species including *A. nidulans* and *A. aculeatus*.[87,88] A following research indicated a single-strand oligonucleotide could be used as a high efficient gene-target substrate

for CRISPR-Cas9 mediated gene editing, and an updated vector system could deliver multiple sgRNAs thus targeting more than one gene at once.[89] Despite that the CRISPR-Cas9 system was originally developed for genome editing for *Aspergillus* species, it could be used to engineer a phylogenetically distinct species *Talaromyces atroroseus* without any modifications.[90]

1.5 Establishment of the link between secondary metabolites and biosynthetic gene clusters

With the rapid evolvement of next-generation sequencing (NGS) technologies, the average cost and time-consuming for whole-genome sequencing (WGS) have been dropped significantly in recent years. Meanwhile, since *Aspergillus* genomics started in January 2003 large fungal sequencing projects have been initiated to obtain more genome sequences thus facilitating a better understanding and full utilization of fungal resources such as the 1K Fungal Genomes Project, *Aspergillus* whole genus sequencing project, and Genomic Encyclopedia of Fungi.[91–93] Right now, 228 whole-genome sequences of *Aspergillus* species/strains have been released on Joint Genome Institute.[94] With more genomic data available, it is possible to find the corresponding BGCs of SMs. Several computational tools have been developed to predict the BGCs or the structures of SMs such as CLUSEAN[95], PRISM[96], antiSMASH[97], and others. Elucidation of the biosynthetic pathways and linking of the BGCs to the biosynthetic products are of high importance as they will help to predict the structures of SMs, facilitate the discovery of similar products in related species by comparative genomic analysis, and improve the production of products of interest.[98,99]

However, to the best of our knowledge, none of the abovementioned in silico platforms guarantees an automated prediction of precise structures of SMs so far. The retro-biosynthetic approach starting with a fully characterized structure is still the most common way to establish connections between BGCs and SMs. The overall workflow of establishing the link between BGCs and SMs is illustrated (Figure 15): structure of SMs – retro-biosynthesis – bioinformatic analysis of the genome sequence – BGC(s) prediction – experimental confirmation – chemical analysis – the original structure. In this workflow, the starting point is the chemically characterized SM

structure with knowledge of NP biosynthesis and an annotated whole-genome sequence. A BGC containing all the essential genes for synthesizing a specific compound is proposed by deducing what enzyme activities are needed in the hypothetic biosynthetic routes and then searching the genome sequence for those activities. A series of lab work is performed including genetic manipulation in the host strain and reconstitution of the biosynthesis including in vivo heterologous expression and in vitro enzymatic reactions, which were followed by chemical profiles analysis.[100]

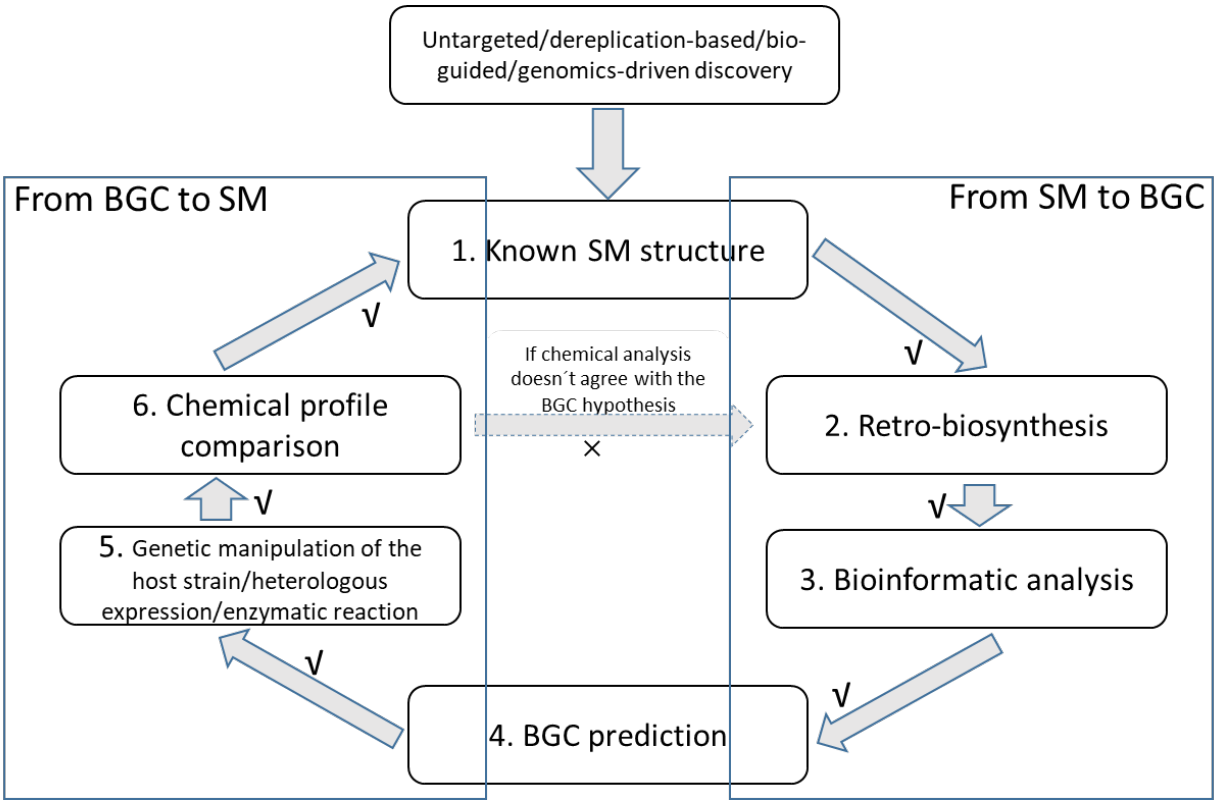


Figure 15. The general workflow to link known SMs to the corresponding BGC.

1.5.1 From secondary metabolites to biosynthetic gene clusters

Retro-biosynthesis. This approach starts with a characterized structure *A* and attempts to identify which enzymes might be involved in the biosynthetic process. Using the knowledge of enzymatic reactions and chemical transformation, an immediate *B* which is a one-step reaction before the final product *A* is predicted. The predicted intermediate *B* is then used for another round of retro-

biosynthetic enumeration to predict an immediate *C* which is the precursor of molecule *B*. This cycle of reaction enumeration repeats itself until a known putative starting compound *Z* is obtained.[101] With the starting molecule and all intermediates, the enzymatic activities needed are listed to search the corresponding responsible genes in the annotated genome sequence using a bioinformatic analysis based approach. A most promising BGC containing all the necessary elements can be predicted. However, it´s notable that as the biosynthetic pathway is not always a linear and single-direction process, bypass pathway(s) and shunt products may cause difficulties in the retro-biosynthetic analysis.

In addition to retro-biosynthesis, other methods can also be used to propose the BGC of a given structure. If the BGC of similar compounds has been characterized, a homology search will be a good alternative to search the responsible BGC for the metabolite of interest. The characterized genes of similar compounds can be used as a query/bait to search for similar genes in the genome sequence of the organism that produces the target compound. The third way to predict the BGC of a specific compound is comparative genomics which relies on the availability of whole-genome sequences of a list of producers of the given compound. The merit of this strategy is that no prior knowledge about the BGC of the compound of interest is needed.[100,102]

1.5.2 From biosynthetic gene clusters to secondary metabolites

A typical method to link the proposed BGC to the given compound is the genetic manipulation of the host strain. This is performed by deleting or disrupting the genes encoding backbone synthases or synthetases as well as the tailoring enzymes followed by subsequent biochemical analysis by LC-MS. If one product is missing (or reduced compared to the unmodified strain) in the chemical spectrum of one deletion strain, this gene is highly likely to be responsible for encoding the specific enzyme synthesizing this product. Intermediate accumulation may also be observed when comparing the chemical profiles of the gene deletion or disruption strains with the reference strain. A commonly used genetic editing technology is CRISPR-Cas9 and this technology has been adopted as a genetic tool in engineering filamentous fungi as indicated earlier. To overcome false results from gene deletion experiments, a complementation

experiment where the mutated genes are re-inserted to the deletion strain should be accompanied.[87,89,100]

Heterologous expression is another strategy to connect the BGC to the corresponding compound when the engineering of native strain is difficult or not established, or when a clean SM background is preferred. Before heterologous expression, the first thing that needs to be considered is the selection of a new host. Though many heterologous bacterial and fungal hosts for NP production have been used, *Escherichia coli* is a main choice for the prokaryotic expression while *Saccharomyces cerevisiae* and *A. nidulans* have often been used as hosts expressing genes from fungi. To facilitate the activation of pathways in a heterologous host, promoters are usually needed to be exchanged to drive gene expressions.[100,102–104] In vitro enzymatic reaction is the third method to elucidate the function of a specific enzyme (and thereby related gene) in the production of a specific compound, often used as a complementary method of the first two approaches. The reaction is usually performed with the purified enzyme with a precursor/intermediate compound under certain conditions, and the generation of another intermediate or the final product proves the function of the enzyme.[105]

2. Overall results and discussions

2.1. Natural products discovery from *Aspergillus californicus*

2.1.1 Characterization of octaketides from *Aspergillus californicus*

During the initial search for possible echinocandin-type of compounds from *A. californicus* IBT 16748, we detected many unknown compounds from its organic extract, which became the starting point of investigating this species. An MS guided natural product discovery lead to the isolation of nine polyketides including two new diastereomers califuranone A_1 and A_2 and a new racemate calitetralintriol (Figure 16). Compound (-)-calitetralintriol A showed weak cytotoxicity against HL-60 (leukemia) cell line with IC_{50} value 20 µg/ml (58.1 µM). Compound (+)-calitetralintriol A displayed activity against MCF-7 (breast) and HL-60 with IC_{50} 20 and 10 µg/ml (58.1 and 29.0 µM). Emefuran A showed activities against HepG2 (liver)and HL-60 with IC_{50} 9.9 and 20 µg/ml (28.4 and 57.4 µM), respectively. Calidiol A showed moderate activities against methicillin-resistant *Staphylococcus aureus* MB5393 with a MIC value of 48 µg/ml. (**Appendix 1**)

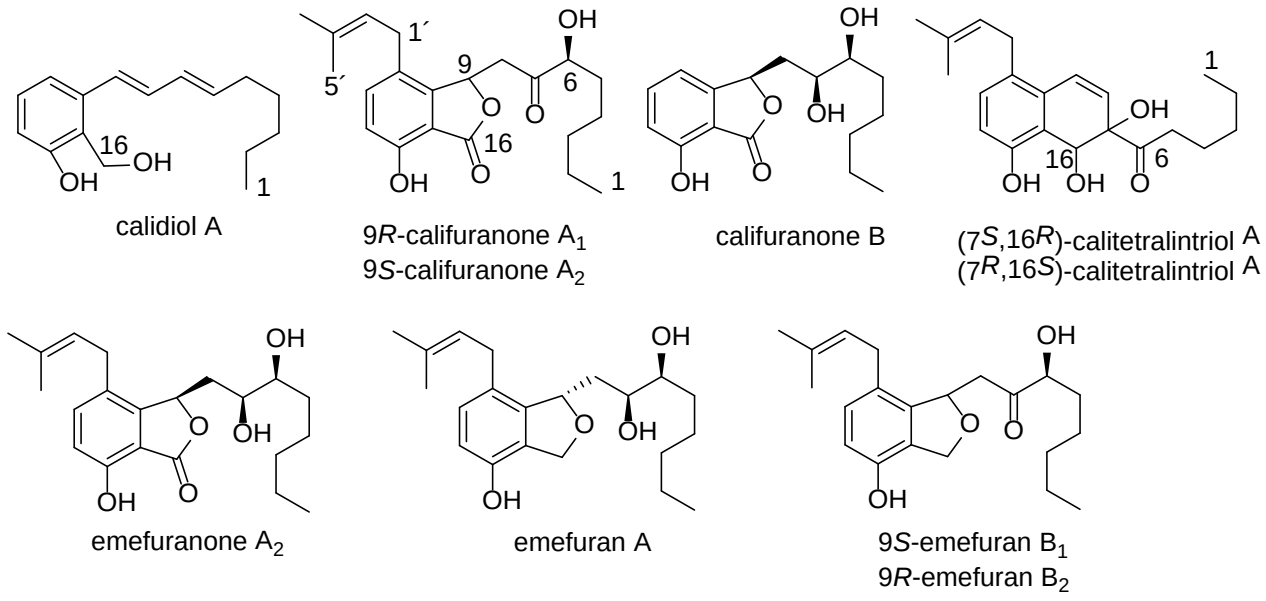


Figure 16. Structures of nine octaketides isolated from *Aspergillus californicus* in this project.

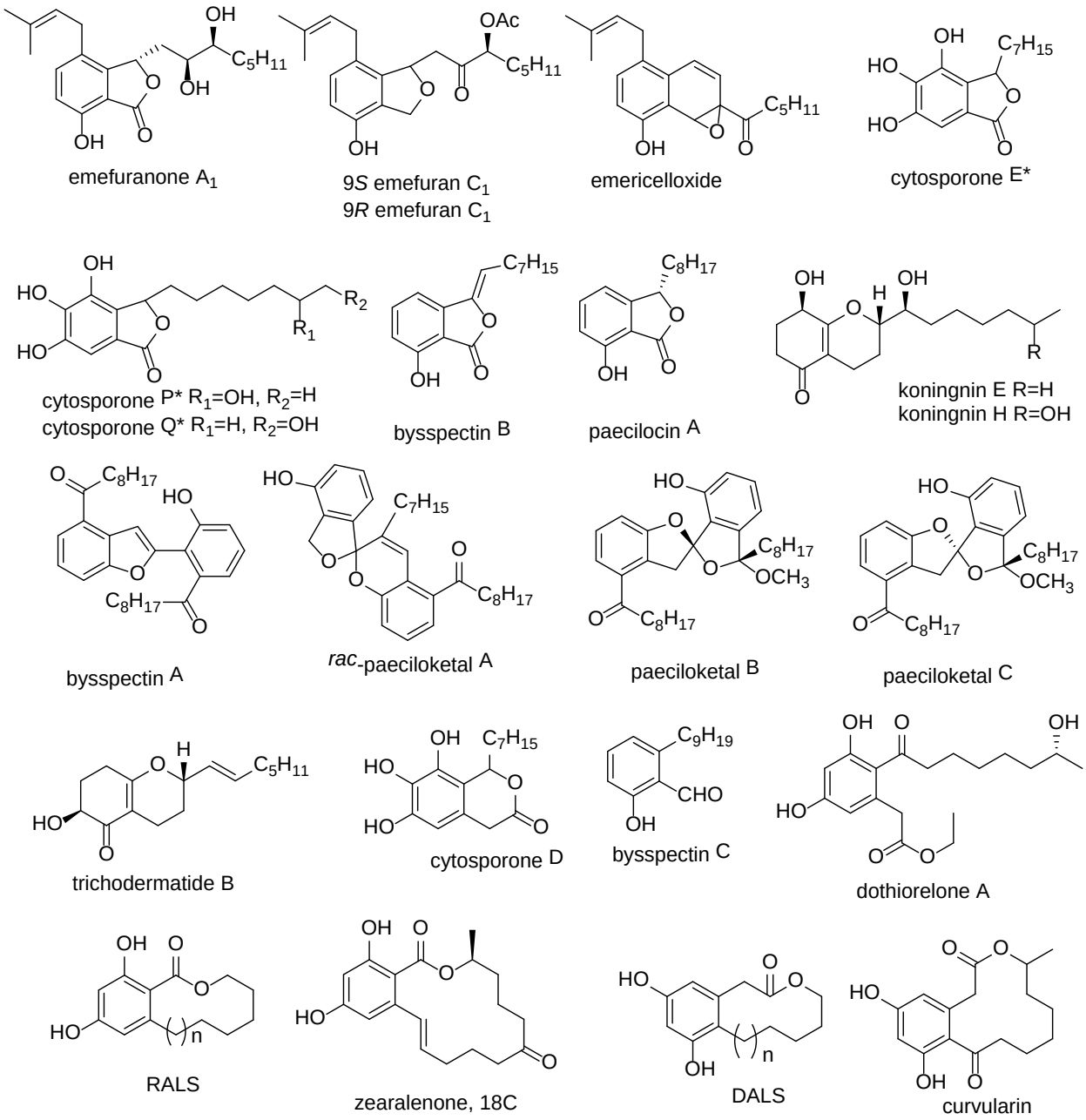


Figure 17. Selected known octaketides and related compounds.

All these compounds share a backbone of sixteen carbons, thus octaketides, and they show high similarities with the phthalide and phthalane derivatives isolated as anti-*Bacillus subtilis* agents from *Emericella* sp. IFM57991.[106] Under the "one fungus: one name" nomenclatural system, strain IFM57991 was proposed belonging to subgenus *Nidulantes* section *Usti* or *Cavernicolarum* due to the production of hülle cells. The unavailability of its β-tubulin gene sequence hampered

a further investigation of its relationship with *A. californicus* IBT 16748. Based on retro-biosynthetic considerations, these octaketides were supposed to be from the same biosynthetic pathway, where the racemate calitetralintriol A is likely constructed through a folding pattern where an intermediate emericelloxide might be involved, which distinguishes the other octaketides.

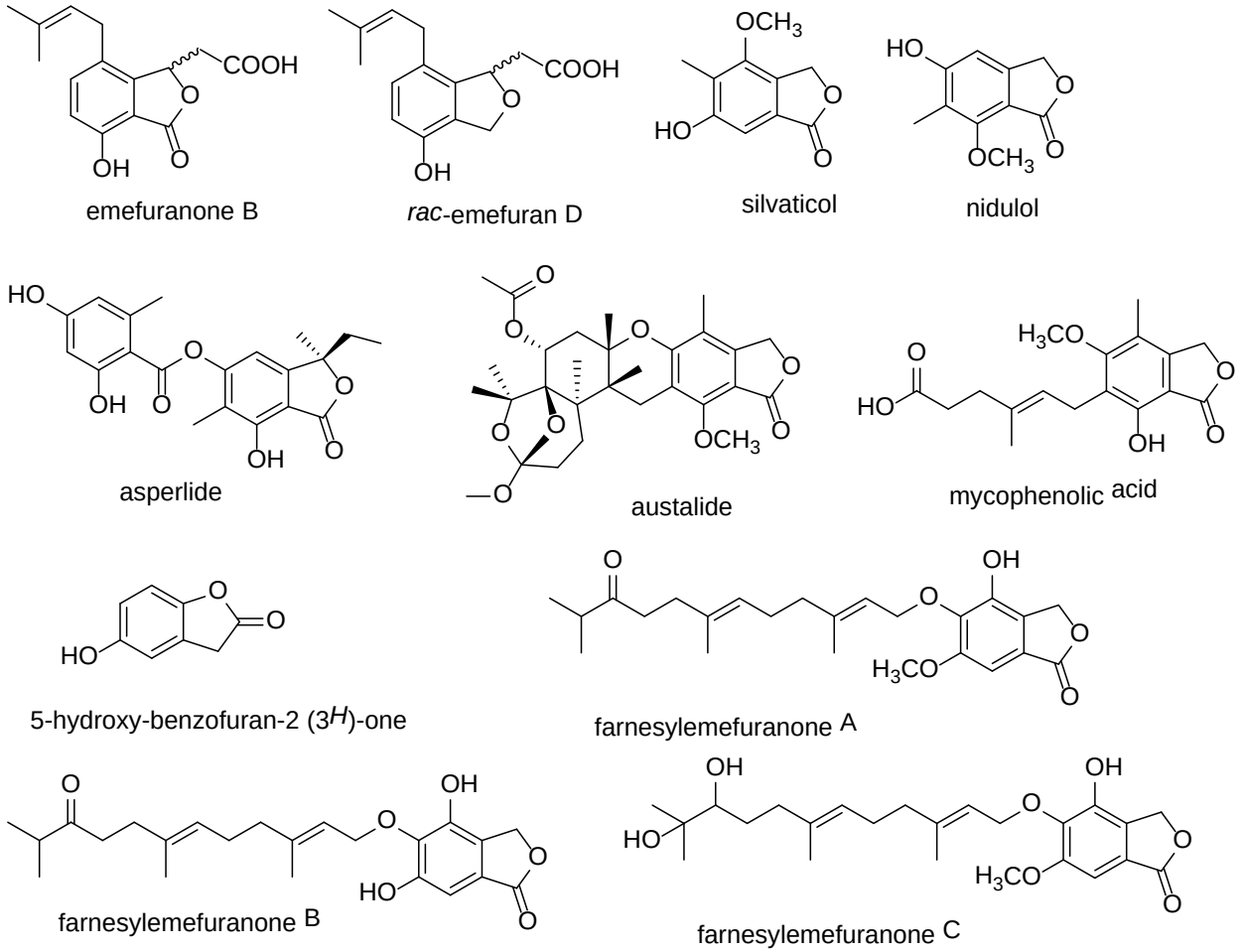


Figure 18. Selected structures of phthalide and phthalane derivatives from fungal sources.

Phthalide and phthalane octaketides have been discovered from other fungal sources (Figure 17) including monomers bysspectin B and paecilocin A from the endophytic fungus *Byssochlamys spectabilis*,[107] cytosporones E, P and Q with 15 carbons from endophytic fungus *Cytospora* sp.,[108,109] as well as dimers bysspectin A,[107] and paeciloketals A-C from the jellyfish-derived fungus *Paecilomyces variotii*.[110] Besides, many other fungal octaketides were biosynthesized by

different folding patterns without forming a phthalide or phthalane nucleus but showing various biological potential such as fungicidal, allelopathic, bactericidal and, cytotoxic activities. These compounds include cytosporones A-D,[108] cytosporones J-O,[109,111] dothiorelones A-G,[111,112], phomopsin C,[112] dihydroxyphenylacetic acid lactones (DALs) and resorcylic acid lactones (RALs),[112] koninginins A-H,[113–115] trichodermatides B-D,[116], etc. However, none of these reported octaketides have a prenyl substitution except for the emefuranones and emefurans from the *Aspergillus* ("*Emericella*") species strain IFM57991. Other metabolites from *Aspergillus* also contain a phthalide or phthalane moiety (Figure 18) including asperlide, porriolides, austalides, mycophenolic acids, emefuranones, emefurans, farnesylemefuranones, etc.[7,114,117–120] The prenylated octaketides with a phthalide or phthalane core discovered in this study enrich the chemical diversity of *Aspergillus* subgenus *Nidulantes*.[114]

2.1.2 Two new naphthyl derivatives from *Aspergillus californicus*

In this section, two new naphthyl products calinaphthyltriol A and calinaphthalenone A were isolated from *A. californicus* together with one known compound ophiobolin X (Figure 19). Calinaphthyltriol A and ophiobolin X showed moderate cytotoxiciteisagainst HL-60 cell line with IC_{50} values 18 and 24 µg/ml, respectively. Calinaphthyltriol A is a planar structure bearing a naphthyl scaffold with three hydroxyl groups and one prenyl group. Interestingly, a similar product xylarinol C was isolated from a close strain IFM57991 (chapter 2.1.1). Both compounds have 10-carbon backbones which were likely from a pentaketide pathway but with different folding patterns as proposed in Figure 19. Calinaphthyltriol A was possibly formed through C-3/C-8 and C-1/C-10 connections while xylarinol C was folded by C-2/C-7 linking. Both compounds have a prenyl substitution. More evidence from genome sequence analysis and experimental work is needed to fully understand the biosynthetic mechanisms leading to both compounds. (**Appendix 2**)

Calinaphthalenone A contains a 1(2*H*)-naphthalenone backbone with a hydroxymethyl group and a pentyl substitution. Its absolute configuration was determined by comparison of the calculation of the optical rotation (OR) with the measured specific OR. Its backbone comprises of 15 carbons, one less compared to the octaketides described in chapter 1.1, indicating that one carbon might

be lost during the polyketide maturation. Ophiobolin X was first reported from mangrove fungus *Aspergillus ustus*.[121] It belongs to a group of sesquiterpenoids with a tricyclic 5-8-5 ring system. Currently, there are over 61 ophiobolins reported from different fungal genera including *Bipolaris*, *Aspergillus*, *Sarocladium,* and *Drechslera*.[121–123]

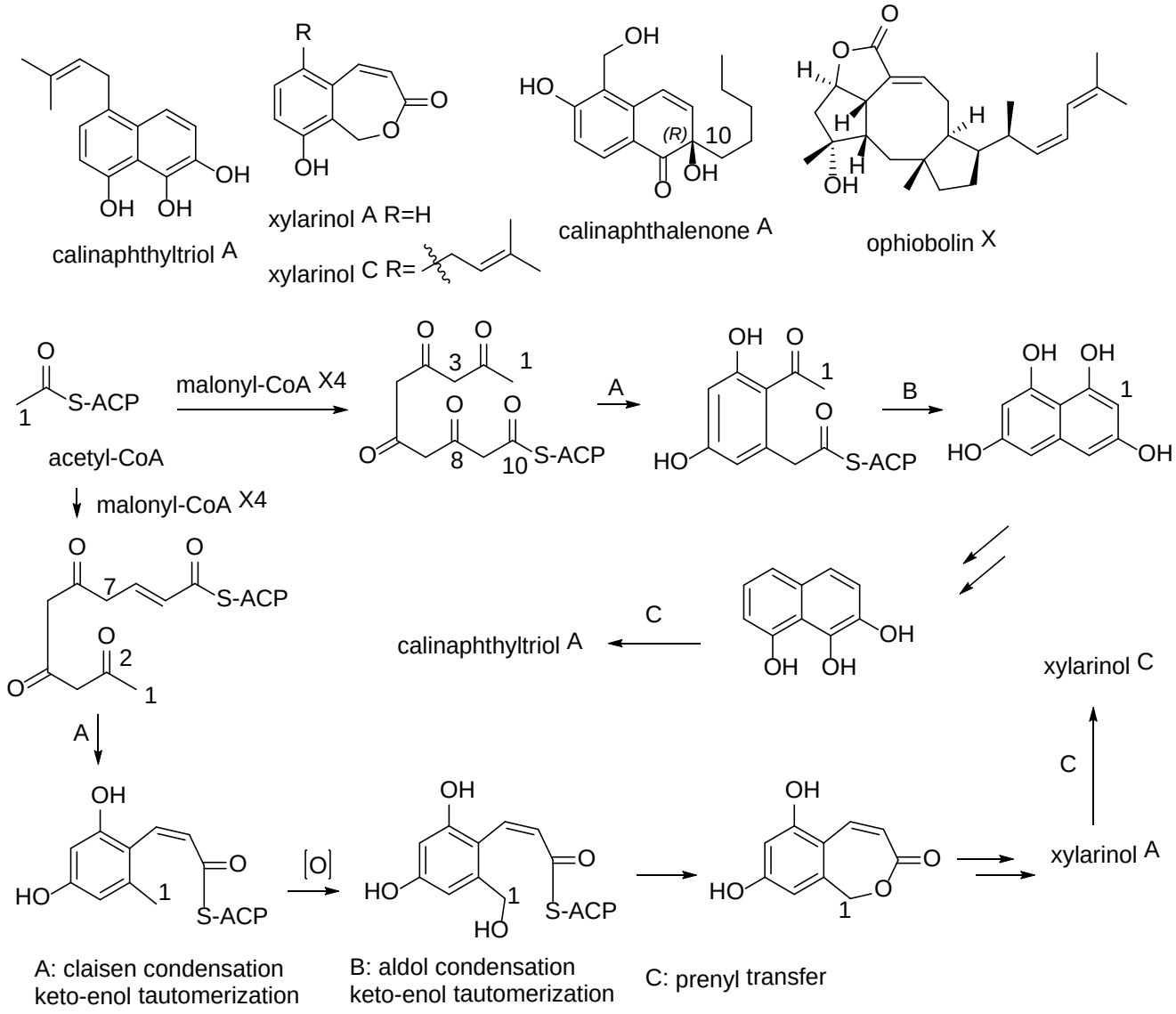


Figure 19. Structures of calinaphthyltriol A, calinaphthalenone A, and ophiobolin X from *Aspergillus californicus* as well as proposed biosynthesis of calinaphthyltriol A and xylarinols C.

2.1.3 Characterization of two new OPK NRPs from *Aspergillus californicus*

Oxepine-pyrimidinone-ketopiperazine (OPK) type nonribosomal peptides (NRPs) are a small class of natural products that have only been characterized from fungal sources before, in particular, 70% from *Aspergillus* species. These OPK compounds are likely active against plant-pathogenic fungi such as *Fusarium graminearum* and *Colletotrichum acutatum*. Oxepinamides D-K all exhibited transactivation effects on liver X receptor α (LXRα), which suggested their potential use as novel LXR agonists in the treatment of atherosclerosis, diabetes, and Alzheimer's disease (**Appendix 4**).[124] During our ongoing exploration of new chemicals from *A. californicus*, two OPK metabolites, oxepinamides L and M, were characterized (Figure 20). Oxepinamide L was the first OPK compound that has been methylated on N-2, and oxepinamide M was a demethylated form of oxepinamide L (**Appendix 3**).

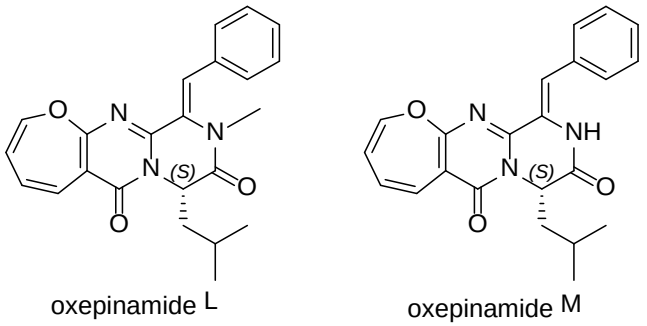


Figure 20. Structures of oxepinamides L and M.

2.2. Investigation of the biosynthetic pathways of secondary metabolites in Aspergilli

2.2.1 Investigation of the biosynthesis of calipyridone A in *Aspergillus californicus*

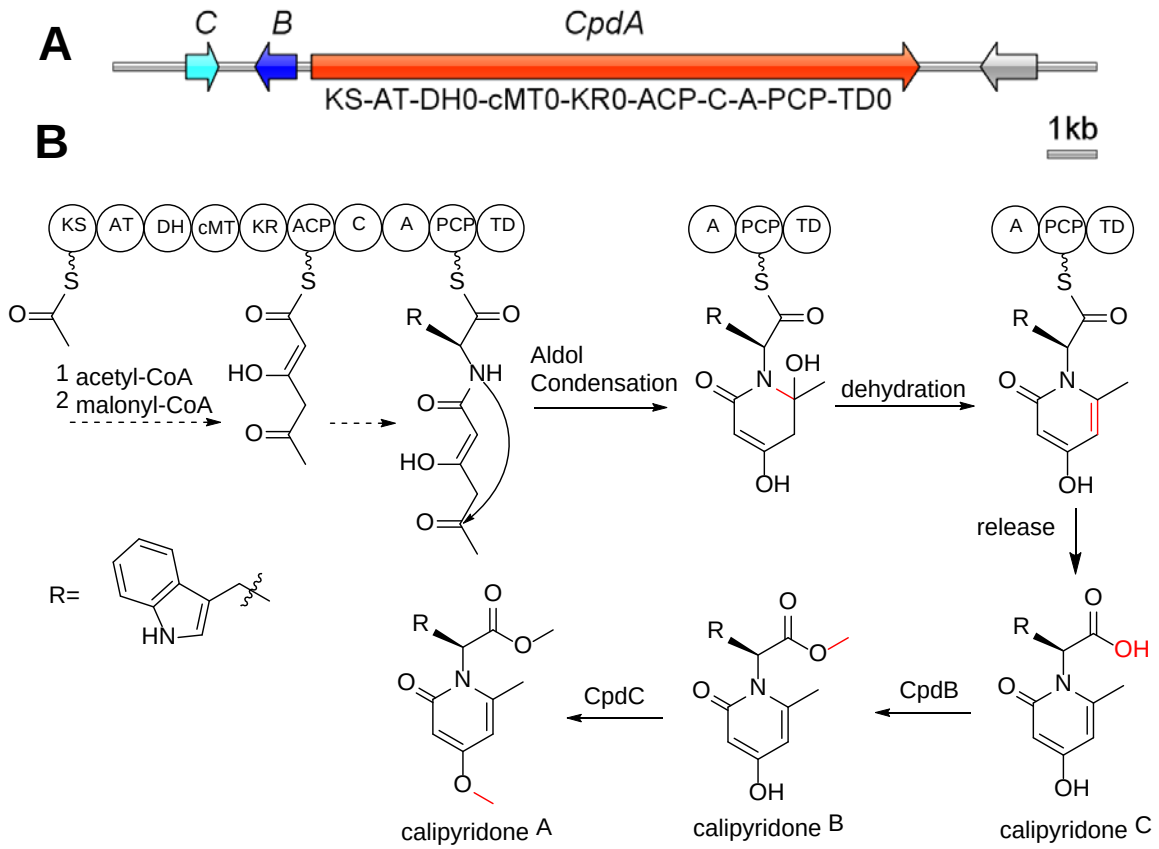


Figure 21. Proposed biosynthetic pathway of calipyridone A in *Aspergillus californicus*.

Polyketide-nonribosomal peptide (PK-NRP) hybrids represent a range of diverse natural products that are biosynthesized by polyketide synthase-nonribosomal peptide synthetases (PKS-NRPSs). Many of these compounds exhibit various biological activities. During the chemical investigation of the non-model filamentous fungus *A. californicus*, a 2-pyridone hybrid calipyridone A was elucidated. Based on genome mining using AntiSMASH[97], a promising biosynthetic gene cluster (BGC) was discovered, and the biosynthesis of calipyridone A was studied by gene deletion experiments in the native producer (Figure 21). Two precursors named calipyridones B (*m/z* 327) and C (*m/z* 313) were discovered by molecular networking (Figure 22),[125], and their structures

were proposed based on the MSMS fragmentation comparisons with that of calipyridone A. The biosynthetic investigation indicated that the 2-pyridone moiety of calipyridone A was formed from aldol condensation without ring expansion of tetramate intermediates which has shown to be catalyzed by P450 enzymes in other fungal 2-pyrones biosyntheses such as TenA, ApdE, LepH, AsolA, IliC, and HarG (Figure 23).[126–131] A second formation mechanism of 2-pyridone has been found in actinobacteria through Dieckmann cyclization such as kirromycin and factumycin (Figure 24).[132] Furthermore, the two *O*-methyltransferases in the biosynthesis showed substrate specificity as no compensation effects were observed between mutants ΔcpdB and ΔcpdC. The KR, DH, and *C*-MT domains of CpdA were inactive in the biosynthesis of calipyridone and these inactivities were supported by antiSMASH analysis and domain sequence comparison. The precursor calipyridone C was released from CpdA without reduction, which hinted that the TD domain within CpdA was not functional. (**Appendix 5**)

In conclusion, calipyridone A is the first microbial 2-pyridone metabolite that was formed without a ring expansion reaction catalyzed by P450. One could speculate that the only reason that a 2-pyridone is the final product of this BGC, is because the KR domain of the PKS is nonfunctioning, altogether meaning that the originally intended product of this BGC has a completely different structure. Unfortunately, we didn´t observe promising bioactivities calipyridone A in the cytotoxic and antibacterial screening. More efforts are needed to understand its ecological significance as it can be produced and secreted to surroundings on many agar plates.

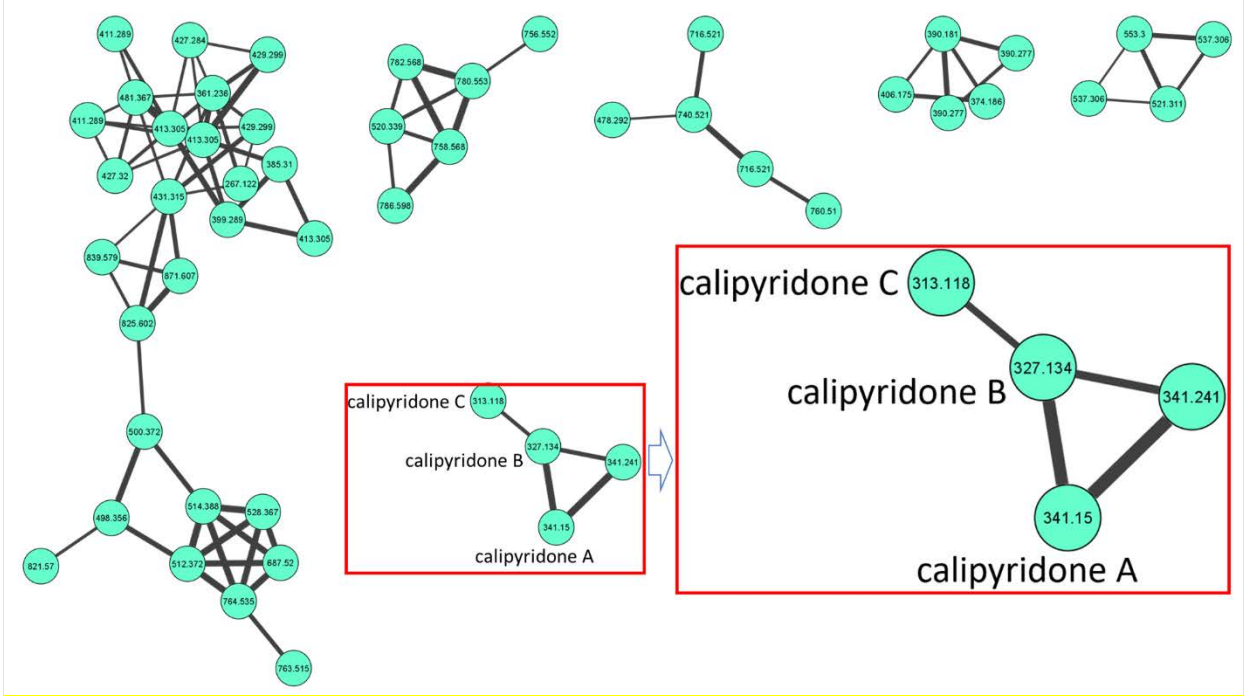


Figure 22. Molecular networking based discovery of two new compounds calipyridones B and C.

(Ion *m/z* 341.241 was proposed to be as calipyridone A as they both had the identical MSMS fragments, and the mass inaccuracy was possibly caused by ion saturation)

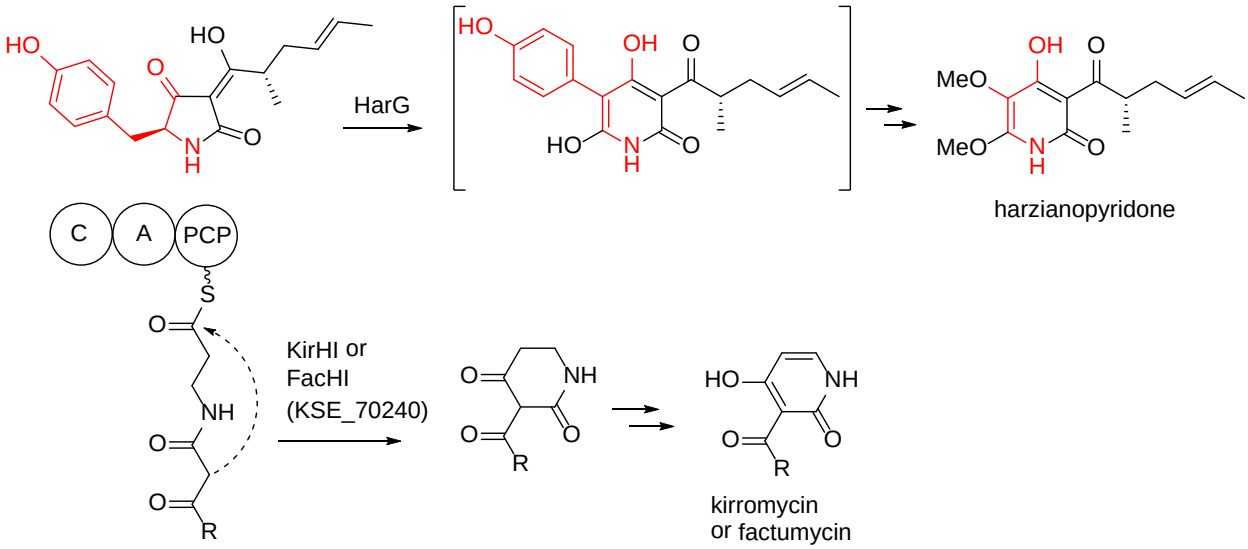


Figure 24. Continued 2-pyridone formation in harzianopyridone, kirromycin and factumycin.

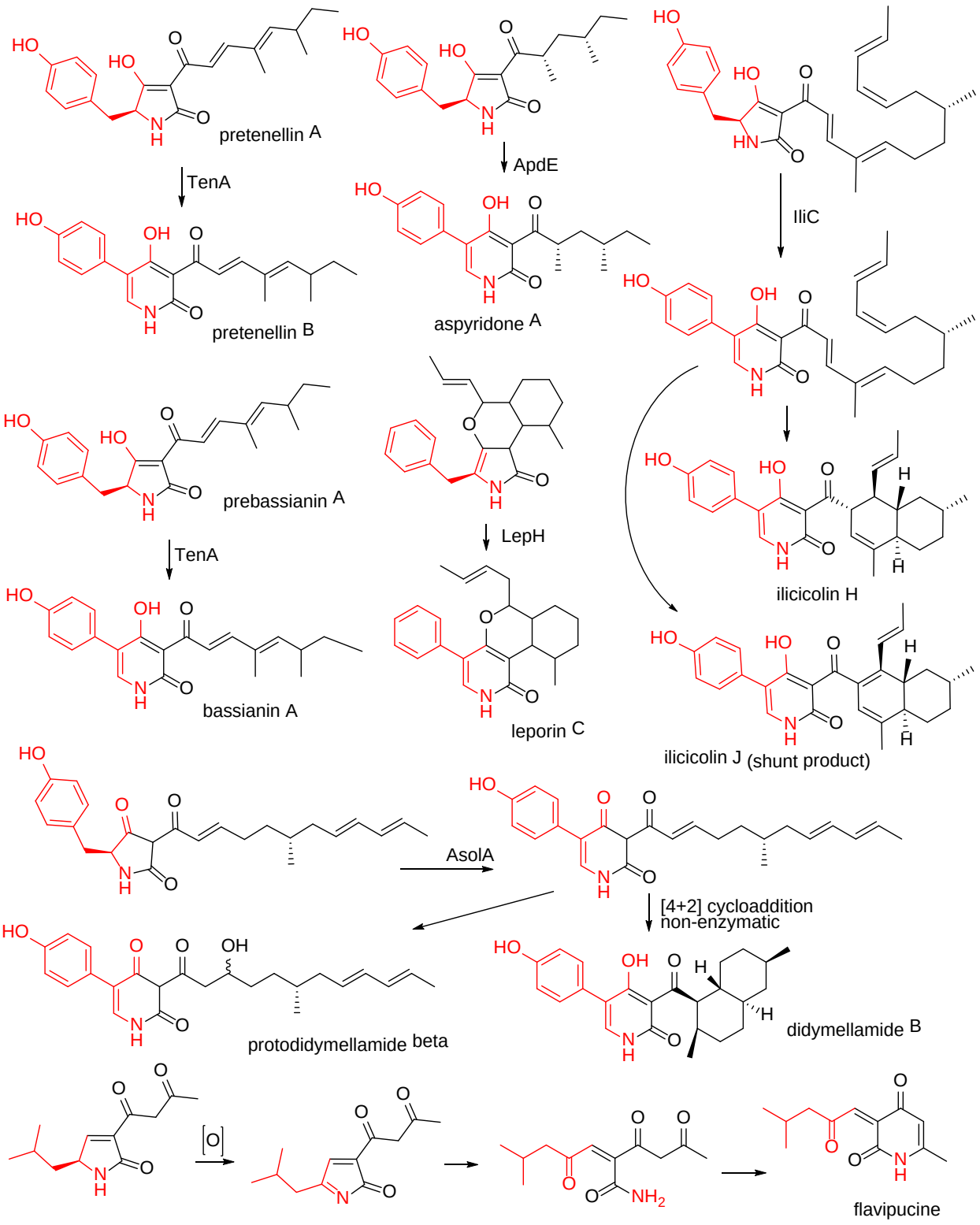


Figure 23. Formation of 2-pyridone analogs through ring expansion of tetramic acid intermediate.

2.2.2 Genetic origin of homopyrones, a rare type of phenylpropanoid and polyketide derived yellow pigments from *Aspergillus homomorphus*

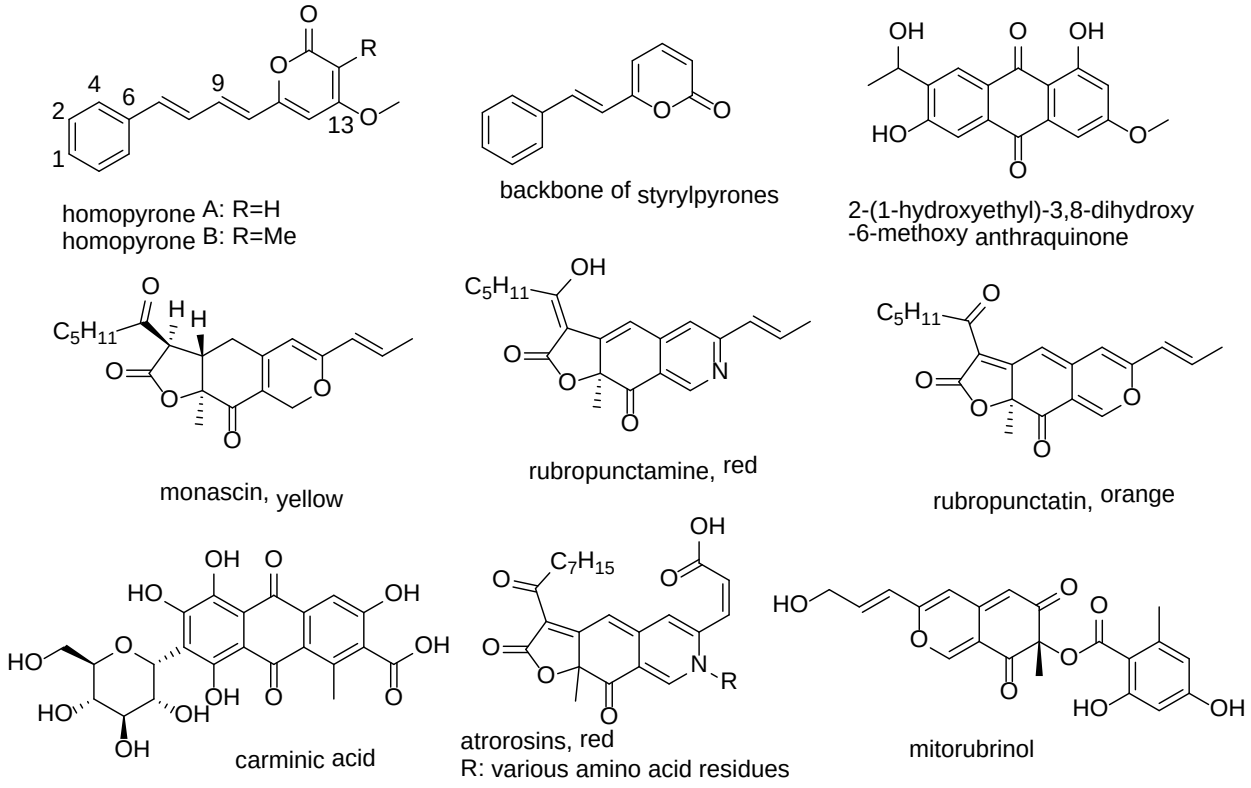


Figure 25. Structures of various pigments and backbone of styrylpyrones from fungi.

In recent years there has been an increasing demand for the replacement of synthetic food colorants with alternatives derived from natural sources. Filamentous fungi are good natural producers of various colorants (Figure 25) like genera *Monascus* (monascin, rubropunctatin, rubropunctamine), *Penicillium* (mitorubrinol), *Talaromyces* (atrorosins), and *Fusarium* (anthraquinones) and some species have been used as cell factories for the production of pigments such as carminic acid and azophilones.[133–140] *Aspergillus homomorphus* is a filamentous fungus belonging to the series *Homomorphi* section *Nigri* of the *Aspergillus* genus (´black aspergilli').[3] Two yellow compounds, homopyrones A and B, have been characterized from this species recently. Their structures turned out to be the same as synthetic ones, but they haven´t been isolated from natural sources before.[141,142] The structures of homopyrones A and B resemble some styrylpyrones from fungi *Phellinus* and *Inonotus* spp as well as *Penicillium*

glabrum but have longer a bridging conjugation between the aromatic ring and pyrone moiety.[143,144]

The intriguing unusual origin of the homopyrone formation prompted us to investigate their biosynthesis in *A. homomorphus*. The cinnamoyl-CoA as the starter unit in the biosynthesis of polyketide is quite uncommon in filamentous fungi, whereas both cinnamoyl-CoA and its derivative *p*-coumaroyl-CoA are widely used in the biosynthesis of flavonoids in plants.[145] Based on retro-synthetic considerations and genome mining using antiSMASH, BGC 68.2 on scaffold 68 containing a PKS gene *AhpA* and a phenylalanine ammonia-lyase (PAL) encoding gene *AhpB* was selected as the most promising cluster for the biosynthesis of both compounds. The elimination of the production of both compounds in the PKS deletion mutant HOM32 indicated that we have predicted the correct gene cluster. A significant decrease in the generation of two compounds in the mutant HOM45 with *AhpB* truncated indicated the involvement of PAL AhpB in their biosynthesis, and the remaining production was like due to the compensation of additional PAL (Asphom1_476219, identity 53.7%, coverage 90.9%) by BLASTP against the database Asphom1_GeneCatalog_proteins_20140716.aa on the JGI genome portal using the default settings.[146,147]

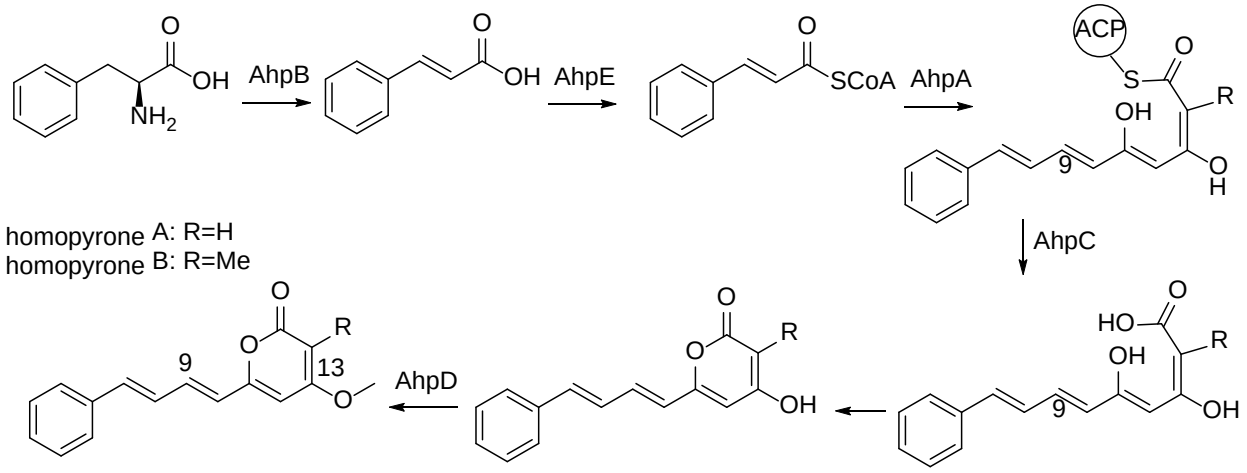


Figure 26. Proposed biosynthetic pathway of homopyrones in *Aspergillus homomorphus*.

Together with the analysis of the functions of surrounding genes, the biosynthetic steps of homopyrones A and B can thus be proposed as (Figure 26): phenylalanine is de-aminated by AhpB

to produce cinnamic acid which is likely ligated by AhpE, showing high similarity (56.99% identity, 99% coverage) with 4-coumarate-CoA ligase-like (4CL) enzyme sequence (accession number: GCB23709.1) from *A. awamori*, yielding the cinnamoyl-CoA as the starter unit.[148,149] Three malonyl-CoA units are then incorporated into the starter unit during which the former keto group of the cinnamic acid was reduced twice by KR and DH domains of AphA, respectively. Finally, *C*-methylation happens on C-14 of the third extender unit. The efficiency of *C*-MT is relaxed due to the conserved SAM binding motif GXGXXG in AhpA replaced by a serine GAGTGS which is likely why some intermediate skips this process resulting in one final homopyrone A with methyl on C-14 and homopyrone B not.[150] A standalone thioesterase (TE) AhpC with hydrolase signature GWSLG homologous to that of canonical type I TEs (GXSXG) catalyzes the hydrolytic release of the intermediate following by the spontaneous generation of the pyrone possibly driven by the keto-enol tautomerism of C-11.[151,152] The maturation of the product is completed after the *O*-methylation of on C-13 catalyzed by AhpD showed high similarity (59.47% identity, 98% coverage) with a putative S-adenosyl-L-methionine dependent methyltransferase (accession number: XP_025385220.1) from *A. eucalypticola* CBS 122712 by BLASTP.[147]

2.2.3 Linking of oxepinamide L to its biosynthetic gene cluster and linking of five genes to the cryptic metabolites in *Aspergillus californicus*

As mentioned above, two oxepine-pyrimidinone-ketopiperazine (OPK) metabolites oxepinamides L and M (Figure 27) were characterized from *Aspergillus californicus* in Appendix 3. The OPK class of nonribosomal peptides has only been discovered from fungal sources so far, and their biosynthesis remains unknown.[124] The various bioactivities of similar OPK NRPS and the intriguing biosynthesis of the oxepine moiety prompted the investigation of the biosynthesis of oxepinamide L. To investigate the biosynthetic pathway of oxepinamides, a collection of strategies were conducted to characterize their biosynthetic pathway including retro-biosynthesis, genome mining, and deletions of five proposed biosynthetic genes based on the *pyrG*- and *ckuAΔ* mutant obtained in Appendix 5 with CRISPR-Cas9 technologies.

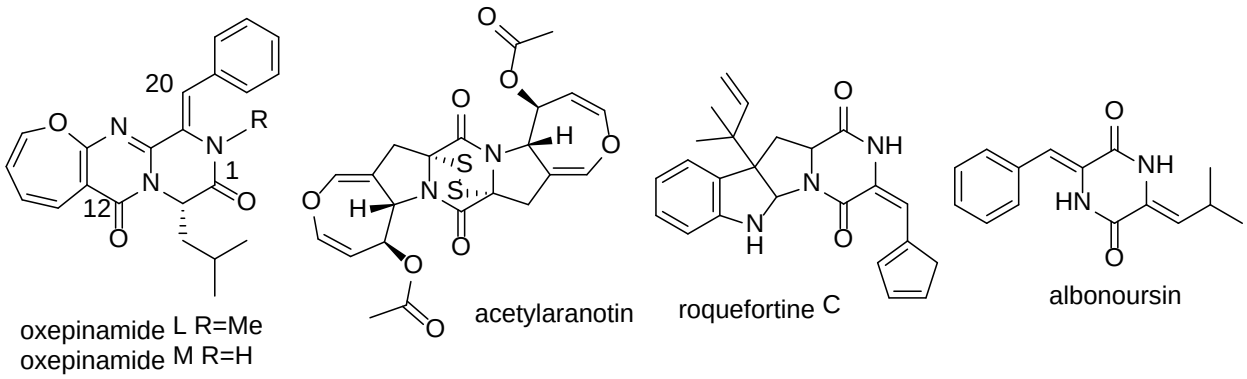


Figure 27. Structures of oxepinamides L and M, and other molecules with similar biosynthetic steps.

Retrobiosynthesis

From the perspective of biosynthesis, the tricyclic core of oxepinamide L is proposed from the condensation of three amino acids including anthranilic acid, phenylalanine, and leucine similar to the well-described biosynthesis of fumiquinazolines.[153–155] Next we hypothesize that epoxidation on the benzene ring of the anthranilic acid residue followed by a ring expansion generates the oxepine moiety (Figure 28). Moreover, the double bond between the C-3 and C-20 carbons of the phenylalanine residue can be proposed to be derived from the hydroxylation on the α carbon and followed by dehydration between the α and β carbons C-3 and C-20. This would be in line with the characterization of the OPK analogs with a hydroxylated α carbon. A second hypothesis of how the double bond formation might take place could be that it is catalyzed by a monooxygenase such as Pc21g15470 (albonoursin biosynthesis) or aminoacyl α, β-dehydrogenase with a FAD co-factor (acetylaranotin).[156,157]

The generation of the oxepine moiety in natural products is still unclear. A study indicated that AtaF, a cytochrome P450 monooxygenase by gene ATEG_03471.1, and AtaY, a putative *p*-hydroxylase encoded by ATEG_03468.1, are responsible for the dihydrooxepine formation from a hydrolyzed benzyl ring during the biosynthesis of acetylaranotin in *A. terreus* through gene deletion experiments.[158] Benzoyl-CoA epoxidase (BoxB), a dinuclear iron enzyme catalyzing the epoxidation reaction of the aromatic ring of benzoyl-CoA, generates the oxepine structure, which is a catalytic strategy by many bacteria to cleave the aromatic pollutants.[159,160] Similar enzymes

phenyl acetyl-CoA epoxidase can also generate oxepine intermediates.[161] A bacterial P450-catalyzed oxidation study in vitro showed that the equilibrium between the oxpine product and arene oxide was related to the solvent and temperature, non-polar solvent, and high temperature favoring the oxepine state (Figure 28).[162] Though no biosynthetic P450s have been associated with the oxepine production in natural products, an oxidase was assumed essential for its generation from the aromatic ring.

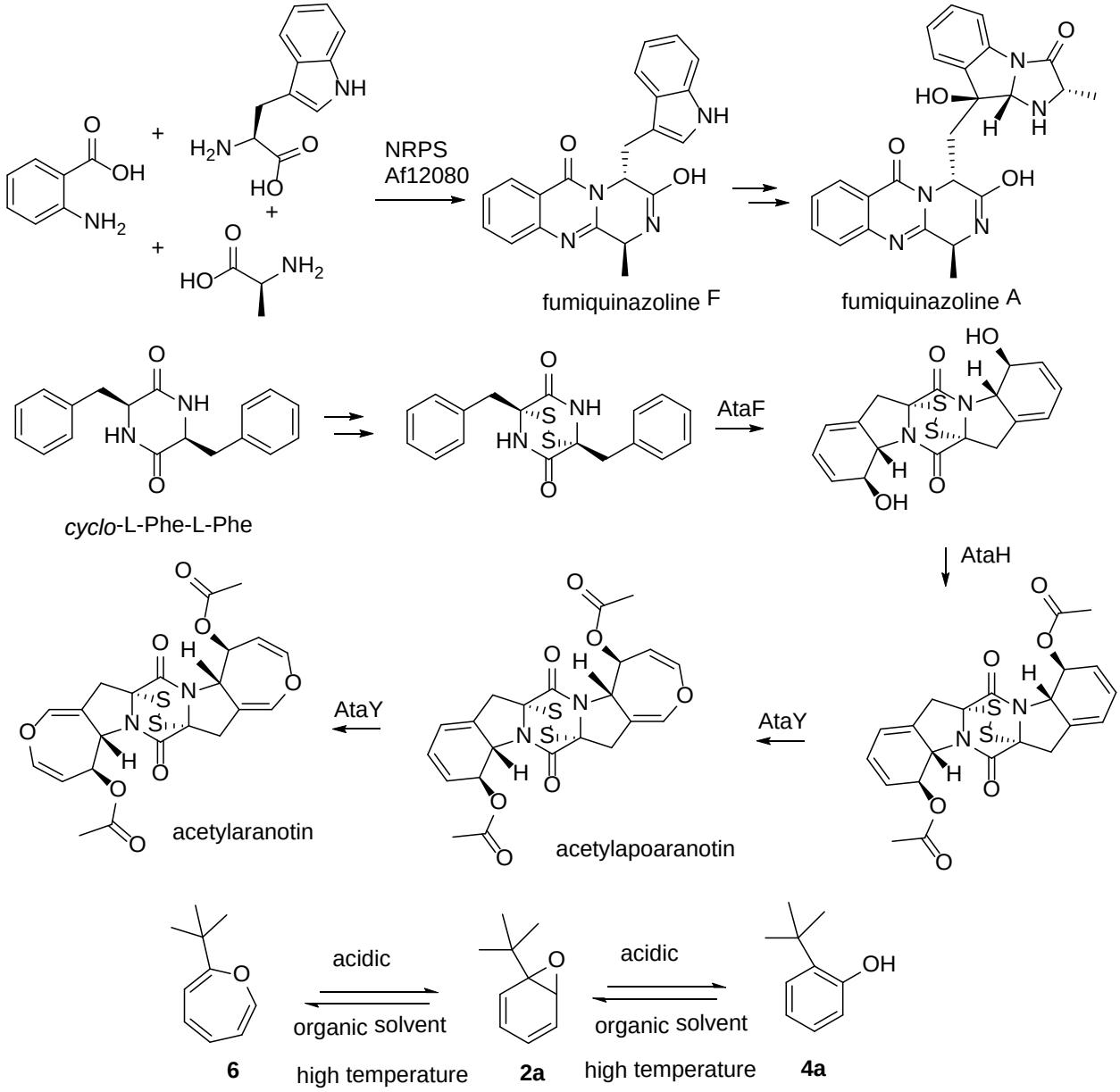


Figure 28. Biosynthesis of fumiquinazolines and oxepine-related products in microorganisms.

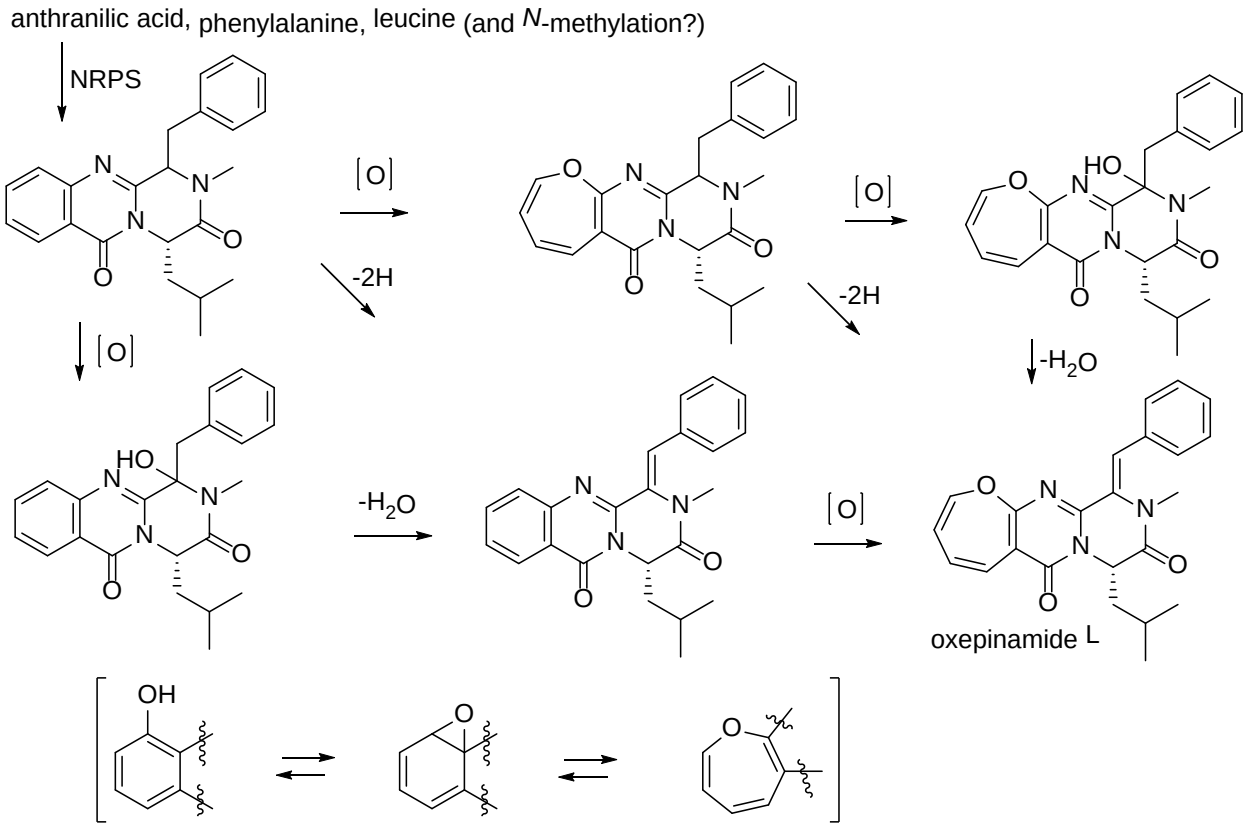


Figure 29. Proposed late-stage biosynthetic steps of oxepinamides L.

BGC prediction and results of gene deletion experiments

Bioinformatic analysis of the *A. californicus* whole-genome sequence by antiSMASH revealed that BGC 2.2 contains one NRPS gene with domain organization A-PCP-C-A-PCP-C-A--(*N*-MT)-PCP-C, two P450 encoding genes, one gene encoding short-chain dehydrogenase/reductase (SDR), one *O*-MT gene (might not be used), and several other surrounding genes. This cluster was selected as the most promising gene cluster out of eight candidates with three A domain based on the consideration that one P450 might contribute to the oxepine formation, an SDR could catalyze the double bond formation.[163–166] Five mutants (CAL005-CAL009) were obtained by deleting five selected genes including jgi.p_Aspcalif1_252836 (P450), jgi.p_Aspcalif1_252838 (NRPS), jgi.p_Aspcalif1_252843 (P450), jgi.p_Aspcalif1_82477 (short-chain dehydrogenase/reductase, SDR), and jgi.p_Aspcalif1_82488 (TF) using CRISPR-Cas9 tools developed for filamentous fungi.[87,89] However, chemical analysis by LC-MS analysis of the five mutants didn´t show any

decrease in the production of oxepinamide L, which unfortunately indicated that we have selected the wrong gene cluster. (**Appendix 7**)

Discussions

The generation of the oxepine sub-structure in natural products remains unknown though some bacteria could produce benzoyl-CoA epoxidase and phenyl acetyl-CoA epoxidase catalyzing the formation of oxepine during the cleave of aromatic pollutants.[159,160] Since BGC 2.2 proved not to be the responsible gene cluster for the biosynthesis of oxepinamide L, a parallel gene deletion experiments targeting the NRPS genes in three BGCs 84.1, 33.1, and 156.1 that contain both NRPS gene and MT module or gene might pinpoint to responsible NRPS thus facilitating the investigation of oxepine generation in natural products. Besides, considering whole genome sequences of 228 *Aspergillus* isolates are available on Joint Genome Institute,[94] comparative genomics using the genome sequences of the known OPK NRPS producers would also facilitate finding the correct BGC.[100,124]

To explore what metabolites those five genes were encoding, six solid media were used to cultivate the mutants (CAL005-CAL009) including CYA, MEAox, OAT, PDA, YES, and YPD, all supplemented with 10 mM Uri+Ura.[88,167] However, the chemical profiles of mutants CAL001 and CAL005-009 were practically identical, like no phenotype differences were observed between the different mutants. The obvious LC-MS difference was observed between CAL000 (wild-type) and CAL001 possibly due to the stress of the excess uridine and uracil to mutants with *pyrG* deleted, which could be seen from the phenotype (**Appendix 7** Figure S3). The inhibition effect of excess uracil and its derivatives has also been reported on *A. nidulans* pyrG89.[168] To find the cryptic metabolites encoded by the five genes that have been deleted, more work is needed like 1) One Strain-Many Compounds (OSMAC)[169] method using liquid media or different temperatures; 2) co-culture with other microorganisms; 3) overexpression of the two transcriptional factors (Gene IDs: jgi.p_Aspcalif1_82484 and jgi.p_Aspcalif1_217404).

3. Conclusions and perspectives

Overall, this PhD project has focused on two aspects related to the secondary metabolism of *Aspergillus*: natural product discovery and characterization of two intriguing metabolite biosynthetic pathways. During the chemistry discovery from *Aspergillus* species, we characterized 15 compounds including polyketides, nonribosomal peptides, terpenoids, and PK-NRP hybrids. Two new polyketides califuranones A_1 and A_2 were diastereomers and calitetralintriol was a racemate. A new compound calidiol A showed moderate activity against MRSA with MIC at 48 μg/ml. As for the cytotoxicity test, all the compounds that have been tested in the cytotoxic assays were regarded as inactive by following the strict criteria of cytotoxicity with the IC_{50} values less than 10 μM. Nine octaketides including five new ones calidiol A, califuranones A_1 and A_2, califuranone B, and calitetralintriol A as well as four known ones emefuranone A_2, emefuran A, emefurans B_1, and B_2 were proposed to be from the same biosynthetic pathway. The fifth known compound was ophiobolin X belonging to a group of sesquiterpenoids with an unusual tricyclic 5-8-5 ring system. Interestingly, all the five known compounds isolated in this study were reported within the last four years - 2016 and 2018. Considering that we obtained 10 new compounds out of 15 in this project, the rare species *A. californicus* was a good source for the discovery of new natural products.

The second major part of this project was the investigation of the biosynthetic pathways of secondary metabolites in *Aspergillus* species. In *A. californicus*, we extended our research from chemistry to molecular biology due to the intriguing structures of oxepinamides L and M, and the PK-NRP hybrid calipyridone A. Based on retro-biosynthesis, genome mining, molecular networking, and multiple gene deletion experiments using CRISPR-Cas9 tools, we successfully characterized the biosynthetic pathway of the hybrid calipyridone A in the native producer which has never been engineered before. To the best of our knowledge, calipyridone A was the first microbial 2-pyridone metabolite that was formed without ring expansion catalyzed by P450s. Besides, this PhD project also worked on the biosynthetic pathway elucidation of two homopyrones in *A. homomorphus*. The deletion of the PKS gene abolished the production of the two compounds indicating that we linked the correct gene cluster to the two yellow products. However, our efforts on deleting five predicted biosynthetic genes for oxepinamides L and M didn´t produce the expected results due to the prediction of the wrong BGC in the first place. A

new round of BGC selection and gene deletion work is needed to elaborate on the biosynthetic pathway of these two OPK compounds. Also, more work is needed to uncover the cryptic SMs encoded by the five genes.

In summary, this PhD project has elucidated the structures of several new natural products from *A. californicus* with one exhibiting anti-MRSA activities, characterized the biosynthetic pathway of a new PK-NRP hybrid in *A. californicus*, and linked the biosynthetic gene cluster to homopyrone polyketides in *A. homomorphus*. Given that both species are now ready for further genetic manipulations, this PhD thesis work sets the scene for additional investigations into the complex and fascinating world of secondary metabolism in the genus *Aspergillus*.

References

(1) Choi, J. J.; Kim, S. H. A Genome Tree of Life for the Fungi Kingdom. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (35), 9391–9396. https://doi.org/10.1073/pnas.1711939114.

(2) Hawksworth, D. L.; Lücking, R. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol. Spectr.* **2017**, *5* (4), 1–17. https://doi.org/10.1128/microbiolspec.FUNK-0052-2016.Correspondence.

(3) Houbraken, J.; Kocsubé, S.; Visagie, C. M.; Yilmaz, N.; Wang, X. C.; Meijer, M.; Kraak, B.; Hubka, V.; Bensch, K.; Samson, R. A.; et al. Classification of *Aspergillus*, *Penicillium*, *Talaromyces* and Related Genera (*Eurotiales*): An Overview of Families, Genera, Subgenera, Sections, Series and Species. *Stud. Mycol.* **2020**, *95*, 5–169. https://doi.org/10.1016/j.simyco.2020.05.002.

(4) Jane, N.; C, V. T.; Sebastian, T.; C, F. J.; O, L. T.; V, G. I.; E, B. S.; R, A. M. Speciation over 200 Million Years – What Makes an Aspergillus Species. In *Book of abstracts from the 13th European Conference on Fungal Genetics*; 2016; p 433.

(5) Robert A Samson, Jos Houbraken, Ulf Thrane, Jens C Frisvad, B. A. *Food and Indoor Fungi*; CBS-KNAW Fungal Biodiversity Centre, 2010.

(6) Samson, R. A.; Visagie, C. M.; Houbraken, J.; Hong, S. B.; Hubka, V.; Klaassen, C. H. W.; Perrone, G.; Seifert, K. A.; Susca, A.; Tanney, J. B.; et al. Phylogeny, Identification and Nomenclature of the Genus *Aspergillus*. *Stud. Mycol.* **2014**, *78* (1), 141–173. https://doi.org/10.1016/j.simyco.2014.07.004.

(7) El-Hawary, S. S.; Moawad, A. S.; Bahr, H. S.; Abdelmohsen, U. R.; Mohammed, R. Natural Product Diversity from the Endophytic Fungi of the GenusAspergillus. *RSC Adv.* **2020**, *10* (37), 22058–22079. https://doi.org/10.1039/d0ra04290k.

(8) Schueffler, A.; Anke, T. Fungal Natural Products in Research and Development. *Nat. Prod. Rep.* **2014**, *31* (10), 1425–1448. https://doi.org/10.1039/c4np00060a.

(9) Knuf, C.; Nielsen, J. Aspergilli: Systems Biology and Industrial Applications. *Biotechnol. J.* **2012**, *7* (9), 1147–1155. https://doi.org/10.1002/biot.201200169.

(10) Aspergillus & Aspergillosis Website https://www.aspergillus.org.uk/content/aspergillus-and-their-applications-food-industry.

(11) Katz, L.; Baltz, R. H. Natural Product Discovery: Past, Present, and Future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43* (2–3), 155–176. https://doi.org/10.1007/s10295-015-1723-5.

(12) Benz, V. F.; Knüsel, F.; Nüesch, J.; Treichler, H.; Voser, W.; Nyfeler, R.; Keller-Schierlein, W. Echinocandin B, Ein Neuartiges Polypeptid-Antibioticum Aus *Aspergillus Nidulans* Var. *Echinulatus*: Isolierung Und Bausteune. *Helv. Chim. ACTA* **1974**, *57* (8), 2459–2477.

(13) Denning, D. W. Echinocandin Antifungal Drugs. *Lancet* **2003**, *362* (9390), 1142–1151. https://doi.org/10.1016/S0140-6736(03)14472-8.

(14) Chen, J.; Cheng, Y. Q.; Yamaki, K.; Li, L. Te. Anti-α-Glucosidase Activity of Chinese Traditionally Fermented

Soybean (Douchi). *Food Chem.* **2007**, *103* (4), 1091–1096. https://doi.org/10.1016/j.foodchem.2006.10.003.

(15) Zhang, J. H.; Tatsumi, E.; Fan, J. F.; Li, L. Te. Chemical Components of *Aspergillus*-Type Douchi, a Chinese Traditional Fermented Soybean Product, Change during the Fermentation Process. *Int. J. Food Sci. Technol.* **2007**, *42* (3), 263–268. https://doi.org/10.1111/j.1365-2621.2005.01150.x.

(16) Frisvad, J. C.; Hubka, V.; Ezekiel, C. N.; Hong, S. B.; Nováková, A.; Chen, A. J.; Arzanlou, M.; Larsen, T. O.; Sklenář, F.; Mahakarnchanakul, W.; et al. Taxonomy of *Aspergillus* Section *Flavi* and Their Production of Aflatoxins, Ochratoxins and Other Mycotoxins. *Stud. Mycol.* **2019**, *93*, 1–63. https://doi.org/10.1016/j.simyco.2018.06.001.

(17) Latgé, J. P. *Aspergillus fumigatus* and Aspergillosis. *Clin. Microbiol. Rev.* **1999**, *12* (2), 310–350. https://doi.org/10.1128/cmr.12.2.310.

(18) Vesth, T. C.; Nybo, J. L.; Theobald, S.; Frisvad, J. C.; Larsen, T. O.; Nielsen, K. F.; Hoof, J. B.; Brandl, J.; Salamov, A.; Riley, R.; et al. Investigation of Inter- and Intraspecies Variation through Genome Sequencing of *Aspergillus* Section *Nigri*. *Nat. Genet.* **2018**, *50* (12), 1688–1695. https://doi.org/10.1038/s41588-018-0246-1.

(19) Vadlapudi, V.; Borah, N.; Yellusani, K. R.; Gade, S.; Reddy, P.; Rajamanikyam, M.; Vempati, L. N. S.; Gubbala, S. P.; Chopra, P.; Upadhyayula, S. M.; et al. *Aspergillus* Secondary Metabolite Database, a Resource to Understand the Secondary Metabolome of Aspergillus Genus. *Sci. Rep.* **2017**, *7* (1), 1–10. https://doi.org/10.1038/s41598-017-07436-w.

(20) Keller, N. P.; Turner, G.; Bennett, J. W. Fungal Secondary Metabolism - From Biochemistry to Genomics. *Nat. Rev. Microbiol.* **2005**, *3* (12), 937–947. https://doi.org/10.1038/nrmicro1286.

(21) Smedsgaard, J.; Nielsen, J. Metabolite Profiling of Fungi and Yeast: From Phenotype to Metabolome by MS and Informatics. *J. Exp. Bot.* **2005**, *56* (410), 273–286. https://doi.org/10.1093/jxb/eri068.

(22) Dewick, P. M. *Medicinal Natural Products A Biosynthetic Approach 3rd Edition*; John Wiley & Sons, Ltd.: Chichester, 2009.

(23) Papadopoulou, K.; Melton, R. E.; Leggett, M.; Daniels, M. J.; Osbourn, A. E. Compromised Disease Resistance in Saponin-Deficient Plants. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (22), 12923–12928. https://doi.org/10.1073/pnas.96.22.12923.

(24) Becher, P. G.; Verschut, V.; Bibb, M. J.; Bush, M. J.; Molnár, B. P.; Barane, E.; Al-bassam, M. M.; Chandra, G.; Song, L.; Challis, G. L.; et al. Developmentally Regulated Volatiles Geosmin and 2-Methylisoborneol Attract a Soil Arthropod to *Streptomyces* Bacteria Promoting Spore Dispersal. https://doi.org/10.1038/s41564-020-0697-x.

(25) 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. https://doi.org/10.1021/acs.jnatprod.9b01285.

(26) Wolff, P. B. Genomics-Driven Discovery, Characterization, and Engineering of Fungal Secondary Metabolites, Technical University of Denmark, 2020.

(27) Bérdy, J. Thoughts and Facts about Antibiotics: Where We Are Now and Where We Are Heading. *J. Antibiot. (Tokyo).* **2012**, *65* (8), 385–395. https://doi.org/10.1038/ja.2012.27.

(28) Crawford, J. M.; Townsend, C. A. New Insights into the Formation of Fungal Aromatic Polyketides. *Nat. Rev. Microbiol.* **2010**, *8* (12), 879–889. https://doi.org/10.1038/nrmicro2465.

(29) Zhang, G.-L.; Feng, Y.-L.; Song, J.-L.; Zhou, X.-S. Zearalenone: A Mycotoxin With Different Toxic Effect in Domestic and Laboratory Animals' Granulosa Cells. *Front. Genet.* **2018**, *9* (December), 1–8. https://doi.org/10.3389/fgene.2018.00667.

(30) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chemie - Int. Ed.* **2009**, *48* (26), 4688–4716. https://doi.org/10.1002/anie.200806121.

(31) Helfrich, E. J. N.; Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33* (2), 231–316. https://doi.org/10.1039/c5np00125k.

(32) Nivina, A.; Yuet, K. P.; Hsu, J.; Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem. Rev.* **2019**, *119* (24), 12524–12547. https://doi.org/10.1021/acs.chemrev.9b00525.

(33) Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G. Biosynthesis of Polyketide Synthase Extender Units. *Nat. Prod. Rep.* **2009**, *26* (1), 90–114. https://doi.org/10.1039/b801658p.

(34) Herbst, D. A.; Townsend, C. A.; Maier, T. The Architectures of Iterative Type I PKS and FAS. *Nat. Prod. Rep.* **2018**, *35* (10), 1046–1069. https://doi.org/10.1039/c8np00039e.

(35) Rasmus John Normand Frandsen. https://www.rasmusfrandsen.dk/polyketide_synthases.htm.

(36) Nakano, C.; Ozawa, H.; Akanuma, G.; Funa, N.; Horinouchi, S. Biosynthesis of Aliphatic Polyketides by Type III Polyketide Synthase and Methyltransferase in *Bacillus subtilis*. *J. Bacteriol.* **2009**, *191* (15), 4916–4923. https://doi.org/10.1128/JB.00407-09.

(37) Seshime, Y.; Juvvadi, P. R.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Identification of Csypyrone B1 as the Novel Product of Aspergillus Oryzae Type III Polyketide Synthase CsyB. *Bioorganic Med. Chem.* **2010**, *18* (12), 4542–4546. https://doi.org/10.1016/j.bmc.2010.04.058.

(38) Schümann, J.; Hertweck, C. Molecular Basis of Cytochalasan Biosynthesis in Fungi: Gene Cluster Analysis and Evidence for the Involvement of a PKS-NRPS Hybrid Synthase by RNA Silencing. *J. Am. Chem. Soc.* **2007**, *129* (31), 9564–9565. https://doi.org/10.1021/ja072884t.

(39) Fisch, K. M. Biosynthesis of Natural Products by Microbial Iterative Hybrid PKS-NRPS. *RSC Adv.* **2013**, *3* (40), 18228–18247. https://doi.org/10.1039/c3ra42661k.

(40) Walsh, C. T. Insights into the Chemical Logic and Enzymatic Machinery of NRPS Assembly Lines. *Nat. Prod. Rep.* **2016**, *33* (2), 127–135. https://doi.org/10.1039/c5np00035a.

(41) Cacho, R. A.; Jiang, W.; Chooi, Y. H.; Walsh, C. T.; Tang, Y. Identification and Characterization of the Echinocandin b Biosynthetic Gene Cluster from *Emericella rugulosa* NRRL 11440. *J. Am. Chem. Soc.* **2012**, *134* (40), 16781–16790. https://doi.org/10.1021/ja307220z.

(42) Huttel, W.; Youssar, L.; Grüning, B. A.; Günther, S.; Hugentobler, K. G. Echinocandin B Biosynthesis: A

Biosynthetic Cluster from Aspergillus Nidulans NRRL 8112 and Reassembly of the Subclusters Ecd and Hty from *Aspergillus pachycristatus* NRRL 11440 Reveals a Single Coherent Gene Cluster. *BMC Genomics* **2016**, *17* (1), 1–8. https://doi.org/10.1186/s12864-016-2885-x.

(43) Jiang, W.; Cacho, R. A.; Chiou, G.; Garg, N. K.; Tang, Y.; Walsh, C. T. EcdGHK Are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold. *J. Am. Chem. Soc.* **2013**, *135* (11), 4457–4466. https://doi.org/10.1021/ja312572v.

(44) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C. P.; Bode, H. B. De Novo Design and Engineering of Non-Ribosomal Peptide Synthetases. *Nat. Chem.* **2018**, *10* (3), 275–281. https://doi.org/10.1038/NCHEM.2890.

(45) Weissman, K. J. The Structural Biology of Biosynthetic Megaenzymes. *Nat. Chem. Biol.* **2015**, *11* (9), 660–670. https://doi.org/10.1038/nchembio.1883.

(46) Dewick, P. M. *Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition*; 2009. https://doi.org/10.1002/9780470742761.

(47) Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-ya, K.; Omura, S.; Cane, D. E.; Ikeda, H. Terpene Synthases Are Widely Distributed in Bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (3), 857–862. https://doi.org/10.1073/pnas.1422108112.

(48) Matsuda, Y.; Abe, I. Biosynthesis of Fungal Meroterpenoids. *Natural Product Reports*. Royal Society of Chemistry 2016, pp 26–53. https://doi.org/10.1039/c5np00090d.

(49) Chen, X. H.; Vater, J.; Piel, J.; Franke, P.; Scholz, R.; Schneider, K.; Koumoutsi, A.; Hitzeroth, G.; Grammel, N.; Strittmatter, A. W.; et al. Structural and Functional Characterization of Three Polyketide Synthase Gene Clusters in *Bacillus amyloliquefaciens* FZB 42. *J. Bacteriol.* **2006**, *188* (11), 4024–4036. https://doi.org/10.1128/JB.00052-06.

(50) Boettger, D.; Hertweck, C. Molecular Diversity Sculpted by Fungal PKS-NRPS Hybrids. *ChemBioChem* **2013**, *14* (1), 28–42. https://doi.org/10.1002/cbic.201200624.

(51) Skellam, E. The Biosynthesis of Cytochalasans. *Nat. Prod. Rep.* **2017**, *34* (11), 1252–1263. https://doi.org/10.1039/c7np00036g.

(52) HE, J.; NIU, X.; YANG, X. Research Progress on Fungal PKS-NRPS Hybrid Metabolites. *Sci. Sin. Vitae* **2019**, *49* (7), 848–864. https://doi.org/10.1360/ssv-2019-0105.

(53) Yun, C. S.; Motoyama, T.; Osada, H. Biosynthesis of the Mycotoxin Tenuazonic Acid by a Fungal NRPS-PKS Hybrid Enzyme. *Nat. Commun.* **2015**, *6*. https://doi.org/10.1038/ncomms9758.

(54) Zhang, Q. W.; Lin, L. G.; Ye, W. C. Techniques for Extraction and Isolation of Natural Products: A Comprehensive Review. *Chinese Med. (United Kingdom)* **2018**, *13* (1), 1–26. https://doi.org/10.1186/s13020-018-0177-x.

(55) Wolfender, J. L. HPLC in Natural Product Analysis: The Detection Issue. *Planta Med.* **2009**, *75* (7), 719–734. https://doi.org/10.1055/s-0028-1088393.

(56) High-Performance Liquid Chromatography (HPLC) https://microbenotes.com/high-performance-liquid-chromatography-hplc/.

(57) Bucar, F.; Wube, A.; Schmid, M. Natural Product Isolation-How to Get from Biological Material to Pure Compounds. *Nat. Prod. Rep.* **2013**, *30* (4), 525–545. https://doi.org/10.1039/c3np20106f.

(58) Phenomenex. Cellulose-1 Phase Information. *https://www.phenomenex.com/Products/HPLCDetail/Lux/Cellulose-1 Cellulose-*.

(59) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **2009**, *44* (3), 210–230. https://doi.org/10.1080/05704920902717872.

(60) Xing, J.; Xie, C.; Lou, H. Recent Applications of Liquid Chromatography-Mass Spectrometry in Natural Products Bioanalysis. *J. Pharm. Biomed. Anal.* **2007**, *44* (2), 368–378. https://doi.org/10.1016/j.jpba.2007.01.010.

(61) Kruve, A.; Kaupmees, K.; Liigand, J.; Oss, M.; Leito, I. Sodium Adduct Formation Efficiency in ESI Source. *J. Mass Spectrom.* **2013**, *48* (6), 695–702. https://doi.org/10.1002/jms.3218.

(62) Agilent Technologies. Agilent 6200 Series TOF and 6500 Series Q-TOF LC/MS System Concepts Guide The Big Picture. **2015**, 1–142.

(63) Demarque, D. P.; Crotti, A. E. M.; Vessecchi, R.; Lopes, J. L. C.; Lopes, N. P. Fragmentation Reactions Using Electrospray Ionization Mass Spectrometry: An Important Tool for the Structural Elucidation and Characterization of Synthetic and Natural Products. *Nat. Prod. Rep.* **2016**, *33* (3), 432–455. https://doi.org/10.1039/c5np00073d.

(64) Fjeldsted, J. Time-of-Flight Mass Spectrometry Technical Overview Gilent Technologies. **2003**.

(65) Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. An Introduction to Quadrupole-Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **2001**, *36* (8), 849–865. https://doi.org/10.1002/jms.207.

(66) Breton, R. C.; Reynolds, W. F. Using NMR to Identify and Characterize Natural Products. *Nat. Prod. Rep.* **2013**, *30* (4), 501–524. https://doi.org/10.1039/c2np20104f.

(67) Zhang, X.; Huang, Q. Z. *Introduction to Spectroscopy*, Fourth Edi.; Cengage Learning, 2009. https://doi.org/10.4028/www.scientific.net/AMM.448-453.4285.

(68) Reynolds, W. F. *Natural Product Structure Elucidation by NMR Spectroscopy*; Elsevier Inc., 2017. https://doi.org/10.1016/B978-0-12-802104-0.00029-9.

(69) Maturana, R. A.; Pessoa, H.; López, B. W. Very Long-Range Correlations (NJC,H n > 3) in HMBC Spectra. *Nat. Prod. Commun.* **2008**, 445–450.

(70) Nyberg, N. T.; Duus, J.; Sørensen, O. W. Heteronuclear Two-Bond Correlation: Suppressing Heteronuclear Three-Bond or Higher NMR Correlations While Enhancing Two-Bond Correlations Even for Vanishing 2JCH. *J. Am. Chem. Soc.* **2005**, *127* (17), 6154–6155. https://doi.org/10.1021/ja050878w.

(71) POLAVARAPU, P. L. Why Is It Important to Simultaneously Use More Than One Chiroptical Spectroscopic

Method for Determining the Structures of Chiral Molecules? *Chirality* **2008**, *43*, 664–672. https://doi.org/10.1002/chir.

(72) Nugroho, A. E.; Morita, H. Circular Dichroism Calculation for Natural Products. *J. Nat. Med.* **2014**, *68* (1), 1–10. https://doi.org/10.1007/s11418-013-0768-x.

(73) Polavarapu, P. L. Determination of the Absolute Configurations of Chiral Drugs Using Chiroptical Spectroscopy. *Molecules* **2016**, *21* (8). https://doi.org/10.3390/molecules21081056.

(74) GENNARO PESCITELLI, T. B. Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra. *Chirality* **2016**, *28* (6), 466–474. https://doi.org/10.1002/chir.

(75) Spartan. Wavefunction Inc.: Irvine, CA. https://www.wavefun.com/.

(76) Macromodel https://www.schrodinger.com/macromodel.

(77) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, D. J. Gaussian 16. Gaussian, Inc.: Wallingford CT 2016.

(78) Yang, L.; Chen, J. A Tale of Two Moieties: Rapidly Evolving CRISPR/Cas-Based Genome Editing. *Trends Biochem. Sci.* **2020**, *xx* (xx), 1–15. https://doi.org/10.1016/j.tibs.2020.06.003.

(79) Horvath, P.; Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science (80-.).* **2010**, *327* (5962), 167–170. https://doi.org/10.1126/science.1179555.

(80) Singh, V. *An Introduction to Genome Editing CRISPR-Cas Systems*; Elsevier Inc., 2020. https://doi.org/10.1016/b978-0-12-818140-9.00001-5.

(81) Koonin, E. V.; Makarova, K. S.; Zhang, F. Diversity, Classification and Evolution of CRISPR-Cas Systems. *Curr. Opin. Microbiol.* **2017**, *37*, 67–78. https://doi.org/10.1016/j.mib.2017.05.008.

(82) Knott, G. J.; Doudna, J. A. CRISPR-Cas Guides the Future of Genetic Engineering. *Science (80-.).* **2018**, *361* (6405), 866–869. https://doi.org/10.1126/science.aat5011.

(83) Makarova, K. S.; Wolf, Y. I.; Iranzo, J.; Shmakov, S. A.; Alkhnbashi, O. S.; Brouns, S. J. J.; Charpentier, E.; Cheng, D.; Haft, D. H.; Horvath, P.; et al. Evolutionary Classification of CRISPR–Cas Systems: A Burst of Class 2 and Derived Variants. *Nat. Rev. Microbiol.* **2020**, *18* (2), 67–83. https://doi.org/10.1038/s41579-019-0299-x.

(84) Hanna, R. E.; Doench, J. G. Design and Analysis of CRISPR–Cas Experiments. *Nat. Biotechnol.* **2020**, *38* (7), 813–823. https://doi.org/10.1038/s41587-020-0490-7.

(85) Patrick Pausch1, 2 †, Basem Al-Shayeb1, 3 †, Ezra Bisom-Rapp4, Connor A. Tsuchida1, 5; Brady F. Cress1, 2, Gavin J. Knott1, 2, Jillian F. Banfield1, 6, Jennifer A. Doudna1, 2, 7, 8. CRISPR-CasΦ from Huge Phages Is a Hypercompact Genome Editor. *Science (80-.).* **2020**, *337* (July), 333–337.

(86) Bhushan, K. *Evolution and Molecular Mechanism of CRISPR/Cas9 Systems*; Elsevier Inc., 2020. https://doi.org/10.1016/b978-0-12-818140-9.00002-7.

(87) Nødvig, C. S.; Nielsen, J. B.; Kogle, M. E.; Mortensen, U. H. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLoS One* **2015**, *10* (7), 1–18. https://doi.org/10.1371/journal.pone.0133085.

(88) Jakob B. Hoof, Christina S. Nødvig, and U. H. M. Chapter 11 Genome Editing: CRISPR-Cas9. In *Fungal Genomics*; Humana Press: New York, NY., 2018; Vol. 1775, pp 119–132. https://doi.org/10.1007/978-1-4939-7804-5.

(89) Nødvig, C. S.; Hoof, J. B.; Kogle, M. E.; Jarczynska, Z. D.; Lehmbeck, J.; Klitgaard, D. K.; Mortensen, U. H. Efficient Oligo Nucleotide Mediated CRISPR-Cas9 Gene Editing in Aspergilli. *Fungal Genetics and Biology*. 2018, pp 78–89. https://doi.org/10.1016/j.fgb.2018.01.004.

(90) Nielsen, M. L.; Isbrandt, T.; Rasmussen, K. B.; Thrane, U.; Hoof, J. B.; Larsen, T. O.; Mortensen, U. H. Genes Linked to Production of Secondary Metabolites in *Talaromyces atroroseus* Revealed Using CRISPR-Cas9. *PLoS One* **2017**, *12* (1), 1–9. https://doi.org/10.1371/journal.pone.0169712.

(91) Archer, D. B.; Dyer, P. S. From Genomics to Post-Genomics in *Aspergillus*. *Curr. Opin. Microbiol.* **2004**, *7* (5), 499–504. https://doi.org/10.1016/j.mib.2004.08.003.

(92) Mikael R. Andersen, Jane L.N. Rasmussen, Martin E. Kogle, Ellen K. Lyhne, Jens C. Frisvad, U. H. M. Whole-genus sequencing of the *Aspergillus* genus https://www.aspergillus.org.uk/content/whole-genus-sequencing-aspergillus-genus.

(93) Kjærbølling, I.; Vesth, T.; Frisvad, J. C.; Nybo, J. L.; Theobald, S.; Kildgaard, S.; Petersen, T. I.; Kuo, A.; Sato, A.; Lyhne, E. K.; et al. A Comparative Genomics Study of 23 *Aspergillus* Species from Section *Flavi*. *Nat. Commun.* **2020**, *11* (1), 1–12. https://doi.org/10.1038/s41467-019-14051-y.

(94) Nordberg, H.; 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* (D1), 26–31. https://doi.org/10.1093/nar/gkt1069.

(95) Weber, T.; Rausch, C.; Lopez, P.; Hoof, I.; Gaykova, V.; Huson, D. H.; Wohlleben, W. CLUSEAN: A Computer-Based Framework for the Automated Analysis of Bacterial Secondary Metabolite Biosynthetic Gene Clusters. *J. Biotechnol.* **2009**, *140* (1–2), 13–17. https://doi.org/10.1016/j.jbiotec.2009.01.007.

(96) Skinnider, M. A.; Merwin, N. J.; Johnston, C. W.; Magarvey, N. A. PRISM 3: Expanded Prediction of Natural Product Chemical Structures from Microbial Genomes. *Nucleic Acids Res.* **2017**, *45* (W1), W49–W54. https://doi.org/10.1093/nar/gkx320.

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

(98) Brown, P. D.; Lawrence, A. L. The Importance of Asking "How and Why?" In Natural Product Structure Elucidation. *Nat. Prod. Rep.* **2017**, *34* (10), 1193–1202. https://doi.org/10.1039/c7np00025a.

(99) Weber, T.; Kim, H. U. The Secondary Metabolite Bioinformatics Portal: Computational Tools to Facilitate Synthetic Biology of Secondary Metabolite Production. *Synth. Syst. Biotechnol.* **2016**, *1* (2), 69–79. https://doi.org/10.1016/j.synbio.2015.12.002.

(100) Kjærbølling, I.; Mortensen, U. H.; Vesth, T.; Andersen, M. R. Strategies to Establish the Link between

Biosynthetic Gene Clusters and Secondary Metabolites. *Fungal Genet. Biol.* **2019**, *130* (June), 107–121. https://doi.org/10.1016/j.fgb.2019.06.001.

(101) Khater, S.; Anand, S.; Mohanty, D. In Silico Methods for Linking Genes and Secondary Metabolites: The Way Forward. *Synth. Syst. Biotechnol.* **2016**, *1* (2), 80–88. https://doi.org/10.1016/j.synbio.2016.03.001.

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

(103) Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R. Heterologous Expression of Bacterial Natural Product Biosynthetic Pathways. *Nat. Prod. Rep.* **2019**, *36* (10), 1412–1436. https://doi.org/10.1039/c8np00091c.

(104) Alberti, F.; Foster, G. D.; Bailey, A. M. Natural Products from Filamentous Fungi and Production by Heterologous Expression. *Appl. Microbiol. Biotechnol.* **2017**, *101* (2), 493–500. https://doi.org/10.1007/s00253-016-8034-2.

(105) Matsuda, Y.; Bai, T.; Phippen, C. B. W.; Nødvig, C. S.; Kjærbølling, I.; Vesth, T. C.; Andersen, M. R.; Mortensen, U. H.; Gotfredsen, C. H.; Abe, I.; et al. Novofumigatonin Biosynthesis Involves a Non-Heme Iron-Dependent Endoperoxide Isomerase for Orthoester Formation. *Nat. Commun.* **2018**, *9* (1), 1–10. https://doi.org/10.1038/s41467-018-04983-2.

(106) Saito, T.; Itabashi, T.; Wakana, D.; Takeda, H.; Yaguchi, T.; Kawai, K. I.; Hosoe, T. Isolation and Structure Elucidation of New Phthalide and Phthalane Derivatives, Isolated as Antimicrobial Agents from *Emericella* sp. IFM57991. *J. Antibiot. (Tokyo).* **2016**, *69* (2), 89–96. https://doi.org/10.1038/ja.2015.85.

(107) Wu, Y. Z.; Zhang, H. W.; Sun, Z. H.; Dai, J. G.; Hu, Y. C.; Li, R.; Lin, P. C.; Xia, G. Y.; Wang, L. Y.; Qiu, B. L.; et al. Bysspectin A, an Unusual Octaketide Dimer and the Precursor Derivatives from the Endophytic Fungus *Byssochlamys spectabilis* IMM0002 and Their Biological Activities. *Eur. J. Med. Chem.* **2018**, *145*, 717–725. https://doi.org/10.1016/j.ejmech.2018.01.030.

(108) Brady, S. F.; Wagenaar, M. M.; Singh, M. P.; Janso, J. E.; Clardy, J. The Cytosporones, New Octaketide Antibiotics Isolated from an Endophytic Fungus. *Org. Lett.* **2000**, *2* (25), 4043–4046. https://doi.org/10.1021/ol006680s.

(109) Abreu, L. M.; Phipps, R. K.; Pfenning, L. H.; Gotfredsen, C. H.; Takahashi, J. A.; Larsen, T. O. Cytosporones O, P and Q from an Endophytic *Cytospora* sp. *Tetrahedron Lett.* **2010**, *51* (13), 1803–1805. https://doi.org/10.1016/j.tetlet.2010.01.110.

(110) Wang, H.; Hong, J.; Yin, J.; Moon, H. R.; Liu, Y.; Wei, X.; Oh, D. C.; Jung, J. H. Dimeric Octaketide Spiroketals from the Jellyfish-Derived Fungus *Paecilomyces variotii* J08NF-1. *J. Nat. Prod.* **2015**, *78* (11), 2832–2836. https://doi.org/10.1021/acs.jnatprod.5b00594.

(111) Xu, J.; Kjer, J.; Sendker, J.; Wray, V.; Guan, H.; Edrada, R. A.; Müller, W. E. G.; Bayer, M.; Lin, W.; Wu, J.; et al. Cytosporones, Coumarins, and an Alkaloid from the Endophytic Fungus *Pestalotiopsis* sp. Isolated from the Chinese Mangrove Plant *Rhizophora mucronata*. *Bioorganic Med. Chem.* **2009**, *17* (20), 7362–7367. https://doi.org/10.1016/j.bmc.2009.08.031.

(112) Meza, A.; Anjos dos Santos, E.; Silva Gomes, R.; de Lima, D.; Beatriz, A. Cytosporones and Related Compounds, A Review: Isolation, Biosynthesis, Synthesis and Biological Activity of Promising Fungal Resorcinolic Lipids. *Curr. Org. Synth.* **2015**, *12* (5), 618–638. https://doi.org/10.2174/1570179412051508211528 55.

(113) Tarawneha, A. H.; León, F.; Radwan, M. M.; Rosa, L. H.; Cutler, S. J. Secondary Metabolites from the Fungus *Emericella nidulans*. *Nat. Prod. Commun.* **2013**, *8* (9), 1285–1288. https://doi.org/10.1177/1934578x1300800925.

(114) Alburae, N. A.; Mohammed, A. E.; Alorfi, H. S.; Jamanturki, A.; Asfour, H. Z.; Alarif, W. M.; Abdel-Lateff, A. Nidulantes of Aspergillus (Formerly Emericella): A Treasure Trove of Chemical Diversity and Biological Activities. *Metabolites* **2020**, *10* (2), 1–28. https://doi.org/10.3390/metabo10020073.

(115) Khan, R. A. A.; Najeeb, S.; Hussain, S.; Xie, B.; Li, Y. Bioactive Secondary Metabolites from *Trichoderma* spp. Against Phytopathogenic Fungi. *Microorganisms* **2020**, *8* (6), 1–22. https://doi.org/10.3390/microorganisms8060817.

(116) Shigehisa, H.; Kikuchi, H.; Suzuki, T.; Hiroya, K. The Revised Structure of Trichodermatide A. *European J. Org. Chem.* **2015**, *2015* (35), 7670–7673. https://doi.org/10.1002/ejoc.201501281.

(117) Klaiklay, S.; Rukachaisirikul, V.; Aungphao, W.; Phongpaichit, S.; Sakayaroj, J. Depsidone and Phthalide Derivatives from the Soil-Derived Fungus *Aspergillus unguis* PSU-RSPG199. *Tetrahedron Lett.* **2016**, *57* (39), 4348–4351. https://doi.org/10.1016/j.tetlet.2016.08.040.

(118) Antipova, T. V.; Zaitsev, K. V.; Oprunenko, Y. F.; Ya. Zherebker, A.; Rystsov, G. K.; Zemskova, M. Y.; Zhelifonova, V. P.; Ivanushkina, N. E.; Kozlovsky, A. G. Austalides V and W, New Meroterpenoids from the Fungus *Aspergillus ustus* and Their Antitumor Activities. *Bioorganic Med. Chem. Lett.* **2019**, *29* (22), 126708. https://doi.org/10.1016/j.bmcl.2019.126708.

(119) Sklenář, F.; Jurjević; Zalar, P.; Frisvad, J. C.; Visagie, C. M.; Kolařík, M.; Houbraken, J.; Chen, A. J.; Yilmaz, N.; Seifert, K. A.; et al. Phylogeny of Xerophilic Aspergilli (Subgenus *Aspergillus*) and Taxonomic Revision of Section *Restricti*. *Stud. Mycol.* **2017**, *88*, 161–236. https://doi.org/10.1016/j.simyco.2017.09.002.

(120) Saito, T.; Itabashi, T.; Wakana, D.; Takeda, H.; Yaguchi, T.; Kawai, K. I.; Hosoe, T. Isolation and Structure Elucidation of New Phthalide and Phthalane Derivatives, Isolated as Antimicrobial Agents from *Emericella* sp. IFM57991. *Journal of Antibiotics*. 2016, pp 89–96. https://doi.org/10.1038/ja.2015.85.

(121) Zhu, T.; Lu, Z.; Fan, J.; Wang, L.; Zhu, G.; Wang, Y.; Li, X.; Hong, K.; Piyachaturawat, P.; Chairoungdua, A.; et al. Ophiobolins from the Mangrove Fungus *Aspergillus ustus*. *J. Nat. Prod.* **2018**, *81* (1), 2–9. https://doi.org/10.1021/acs.jnatprod.7b00335.

(122) Tian, W.; Deng, Z.; Hong, K. The Biological Activities of Sesterterpenoid-Type Ophiobolins. *Mar. Drugs* **2017**, *15* (7), 1–21. https://doi.org/10.3390/md15070229.

(123) Choi, B. K.; Trinh, P. T. H.; Lee, H. S.; Choi, B. W.; Kang, J. S.; Ngoc, N. T. D.; Van, T. T. T.; Shin, H. J. New Ophiobolin Derivatives from the Marine Fungus *Aspergillus flocculosus* and Their Cytotoxicities against Cancer Cells. *Mar. Drugs* **2019**, *17* (6). https://doi.org/10.3390/md17060346.

(124) Guo, Y.; Frisvad, J. C.; Larsen, T. O. Review of Oxepine-Pyrimidinone-Ketopiperazine Type Nonribosomal Peptides. *Metabolites* **2020**, *10* (6), 11–14. https://doi.org/10.3390/metabo10060246.

(125) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837. https://doi.org/10.1038/nbt.3597.

(126) 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* (4), 213–217. https://doi.org/10.1038/nchembio869.

(127) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. Late Stage Oxidations during the Biosynthesis of the 2-Pyridone Tenellin in the Entomopathogenic Fungus *Beauveria bassiana*. *J. Am. Chem. Soc.* **2008**, *130* (52), 17988–17996. https://doi.org/10.1021/ja807052c.

(128) Cary, J. W.; Uka, V.; Han, Z.; Buyst, D.; Harris-Coward, P. Y.; Ehrlich, K. C.; Wei, Q.; Bhatnagar, D.; Dowd, P. F.; Martens, S. L.; et al. An *Aspergillus flavus* Secondary Metabolic Gene Cluster Containing a Hybrid PKS-NRPS Is Necessary for Synthesis of the 2-Pyridones, Leporins. *Fungal Genet. Biol.* **2015**, *81*, 88–97. https://doi.org/10.1016/j.fgb.2015.05.010.

(129) Ugai, T.; Minami, A.; Gomi, K.; Oikawa, H. Genome Mining Approach for Harnessing the Cryptic Gene Cluster in *Alternaria solani*: Production of PKS-NRPS Hybrid Metabolite, Didymellamide B. *Tetrahedron Lett.* **2016**, *57* (25), 2793–2796. https://doi.org/10.1016/j.tetlet.2016.05.043.

(130) Lin, X.; Yuan, S.; Chen, S.; Chen, B.; Xu, H.; Liu, L.; Li, H.; Gao, Z. Heterologous Expression of Ilicicolin H Biosynthetic Gene Cluster and Production of a New Potent Antifungal Reagent, Ilicicolin J. *Molecules* **2019**, *24* (12), 1–10. https://doi.org/10.3390/molecules24122267.

(131) 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* (19), 8550–8554. https://doi.org/10.1021/jacs.0c03438.

(132) Gui, C.; Li, Q.; Mo, X.; Qin, X.; Ma, J.; Ju, J. Discovery of a New Family of Dieckmann Cyclases Essential to Tetramic Acid and Pyridone-Based Natural Products Biosynthesis. *Org. Lett.* **2015**, *17* (3), 628–631. https://doi.org/10.1021/ol5036497.

(133) Christiana, N. O. Production of Food Colourants by Filamentous Fungi. *African J. Microbiol. Res.* **2016**, *10* (26), 960–971. https://doi.org/10.5897/ajmr2016.7904.

(134) Frisvad, J. C.; Yilmaz, N.; Thrane, U.; Rasmussen, K. B.; Houbraken, J.; Samson, R. a. *Talaromyces Atroroseus*, a New Species Efficiently Producing Industrially Relevant Red Pigments. *PLoS One* **2013**, *8* (12), e84102. https://doi.org/10.1371/journal.pone.0084102.

(135) Frandsen, R. J. N.; Khorsand-Jamal, P.; Kongstad, K. T.; Nafisi, M.; Kannangara, R. M.; Staerk, D.; Okkels, F. T.;

Binderup, K.; Madsen, B.; Møller, B. L.; et al. Heterologous Production of the Widely Used Natural Food Colorant Carminic Acid in *Aspergillus nidulans*. *Sci. Rep.* **2018**, *8* (1), 1–10. https://doi.org/10.1038/s41598-018-30816-9.

(136) Tolborg, G.; Ødum, A. S. R.; Isbrandt, T.; Larsen, T. O.; Workman, M. Unique Processes Yielding Pure Azaphilones in *Talaromyces atroroseus*. *Appl. Microbiol. Biotechnol.* **2020**, *104* (2), 603–613. https://doi.org/10.1007/s00253-019-10112-w.

(137) Isbrandt, T.; Tolborg, G.; Ødum, A.; Workman, M.; Larsen, T. O. Atrorosins: A New Subgroup of Monascus Pigments from *Talaromyces atroroseus*. *Appl. Microbiol. Biotechnol.* **2020**, *104* (2), 615–622. https://doi.org/10.1007/s00253-019-10216-3.

(138) Isbrandt, T.; Frisvad, J. C.; Madsen, A.; Larsen, T. O. New Azaphilones from *Aspergillus neoglaber*. *AMB Express* **2020**, *10* (1), 145. https://doi.org/10.1186/s13568-020-01078-4.

(139) Chen, W.; Feng, Y.; Molnár, I.; Chen, F. Nature and Nurture: Confluence of Pathway Determinism with Metabolic and Chemical Serendipity Diversifies: Monascus Azaphilone Pigments. *Nat. Prod. Rep.* **2019**, *36* (4), 561–572. https://doi.org/10.1039/c8np00060c.

(140) Akilandeswari, P.; Pradeep, B. V. Exploration of Industrially Important Pigments from Soil Fungi. *Appl. Microbiol. Biotechnol.* **2016**, *100* (4), 1631–1643. https://doi.org/10.1007/s00253-015-7231-8.

(141) Li, C.; Cheng, B.; Fang, S.; Zhou, H.; Gu, Q.; Xu, J. Design, Syntheses and Lipid Accumulation Inhibitory Activities of Novel Resveratrol Mimics. *Eur. J. Med. Chem.* **2018**, *143*, 114–122. https://doi.org/10.1016/j.ejmech.2017.11.017.

(142) Oliver Kayser, W. Ray Waters, Keith M. Woods, Steve J. Upton, Janet S. Keithly, A. F. K. Evaluation of in Vitro Activity of Aurones and Related Compounds against *Cryptosporidium parvum*. *Planta Med.* **2001**, *3* (9), 722–725.

(143) Lee, I. K.; Yun, B. S. Styrylpyrone-Class Compounds from Medicinal Fungi *Phellinus* and *Inonotus* Spp., and Their Medicinal Importance. *J. Antibiot. (Tokyo).* **2011**, *64* (5), 349–359. https://doi.org/10.1038/ja.2011.2.

(144) Hammerschmidt, L.; Wray, V.; Lin, W.; Kamilova, E.; Proksch, P.; Aly, A. H. New Styrylpyrones from the Fungal Endophyte *Penicillium glabrum* Isolated from *Punica granatum*. *Phytochem. Lett.* **2012**, *5* (3), 600–603. https://doi.org/10.1016/j.phytol.2012.06.003.

(145) Abe, I.; Morita, H. Structure and Function of the Chalcone Synthase Superfamily of Plant Type III Polyketide Synthases. *Nat. Prod. Rep.* **2010**, *27* (6), 809–838. https://doi.org/10.1039/b909988n.

(146) Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otillar, R.; Riley, R.; Salamov, A.; Zhao, X.; Korzeniewski, F.; et al. MycoCosm Portal: Gearing up for 1000 Fungal Genomes. *Nucleic Acids Res.* **2014**, *42* (D1), D699–D704. https://doi.org/10.1093/nar/gkt1183.

(147) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; et al. BLAST: A More Efficient Report with Usability Improvements. *Nucleic Acids Res.* **2013**, *41* (Web Server issue), W29–W33. https://doi.org/10.1093/nar/gkt282.

(148) Dao, T. T. H.; Linthorst, H. J. M.; Verpoorte, R. Chalcone Synthase and Its Functions in Plant Resistance. *Phytochem. Rev.* **2011**, *10* (3), 397–412. https://doi.org/10.1007/s11101-011-9211-7.

(149) Bang, H. B.; Lee, Y. H.; Kim, S. C.; Sung, C. K.; Jeong, K. J. Metabolic Engineering of *Escherichia coli* for the Production of Cinnamaldehyde. *Microb. Cell Fact.* **2016**, *15* (1), 16. https://doi.org/10.1186/s12934-016-0415-9.

(150) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Functional Expression of the *Aspergillus flavus* PKS-NRPS Hybrid CpaA Involved in the Biosynthesis of Cyclopiazonic Acid. *Bioorganic Med. Chem. Lett.* **2009**, *19* (12), 3288–3292. https://doi.org/10.1016/j.bmcl.2009.04.073.

(151) Horsman, M. E.; Hari, T. P. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202. https://doi.org/10.1039/c4np00148f.

(152) Zhai, Y.; Bai, S.; Liu, J.; Yang, L.; Han, L.; Huang, X.; He, J. Identification of an Unusual Type II Thioesterase in the Dithiolopyrrolone Antibiotics Biosynthetic Pathway. *Biochem. Biophys. Res. Commun.* **2016**, *473* (1), 329–335. https://doi.org/10.1016/j.bbrc.2016.03.105.

(153) Ames, B. D.; Walsh, C. T. Anthranilate-Activating Modules from Fungal Nonribosomal Peptide Assembly Lines. *Biochemistry* **2010**, *49* (15), 3351–3365. https://doi.org/10.1021/bi100198y.

(154) Ames, B. D.; Liu, X.; Walsh, C. T. Enzymatic Processing of Fumiquinazoline F: A Tandem Oxidative-Acylation Strategy for the Generation of Multicyclic Scaffolds in Fungal Indole Alkaloid Biosynthesis. *Biochemistry* **2010**, *49* (39), 8564–8576. https://doi.org/10.1021/bi1012029.

(155) Gao, X.; Chooi, Y. H.; Ames, B. D.; Wang, P.; Walsh, C. T.; Tang, Y. Fungal Indole Alkaloid Biosynthesis: Genetic and Biochemical Investigation of the Tryptoquialanine Pathway in *Penicillium aethiopicum*. *J. Am. Chem. Soc.* **2011**, *133* (8), 2729–2741. https://doi.org/10.1021/ja1101085.

(156) García-Estrada, C.; Ullán, R. V.; Albillos, S. M.; Fernández-Bodega, M. Á.; Durek, P.; Von Döhren, H.; Martín, J. F. A Single Cluster of Coregulated Genes Encodes the Biosynthesis of the Mycotoxins Roquefortine C and Meleagrin in *Penicillium chrysogenum*. *Chem. Biol.* **2011**, *18* (11), 1499–1512. https://doi.org/10.1016/j.chembiol.2011.08.012.

(157) Gondry, M.; Lautru, S.; Fusai, G.; Meunier, G.; Ménez, A.; Genet, R. Cyclic Dipeptide Oxidase from S *Treptomyces noursei*. *Eur. J. Biochem.* **2001**, *268* (6), 1712–1721. https://doi.org/10.1046/j.1432-1327.2001.02038.x.

(158) Guo, C. J.; Yeh, H. H.; Chiang, Y. M.; Sanchez, J. F.; Chang, S. L.; Bruno, K. S.; Wang, C. C. C. Biosynthetic Pathway for the Epipolythiodioxopiperazine Acetylaranotin in *Aspergillus terreus* Revealed by Genome-Based Deletion Analysis. *J. Am. Chem. Soc.* **2013**, *135* (19), 7205–7213. https://doi.org/10.1021/ja3123653.

(159) Rather, L. J.; Weinert, T.; Demmer, U.; Bill, E.; Ismail, W.; Fuchs, G.; Ermler, U. Structure and Mechanism of the Diiron Benzoyl-Coenzyme A Epoxidase BoxB. *J. Biol. Chem.* **2011**, *286* (33), 29241–29248. https://doi.org/10.1074/jbc.M111.236893.

(160) Liao, R. Z.; Siegbahn, P. E. M. Mechanism and Selectivity of the Dinuclear Iron Benzoyl-Coenzyme A Epoxidase BoxB. *Chem. Sci.* **2015**, *6* (5), 2754–2764. https://doi.org/10.1039/c5sc00313j.

(161) Grishin, A. M.; Ajamian, E.; Zhang, L.; Rouiller, I.; Bostina, M.; Cygler, M. Protein-Protein Interactions in the β-Oxidation Part of the Phenylacetate Utilization Pathway: Crystal Structure of the PaaF-PaaG Hydratase-Isomerase Complex. *J. Biol. Chem.* **2012**, *287* (45), 37986–37996. https://doi.org/10.1074/jbc.M112.388231.

(162) Stok, J. E.; Chow, S.; Krenske, E. H.; Farfan Soto, C.; Matyas, C.; Poirier, R. A.; Williams, C. M.; De Voss, J. J. Direct Observation of an Oxepin from a Bacterial Cytochrome P450-Catalyzed Oxidation. *Chem. - A Eur. J.* **2016**, *22* (13), 4408–4412. https://doi.org/10.1002/chem.201600246.

(163) Loss, E. M. O.; Lee, M. K.; Wu, M. Y.; Martien, J.; Chen, W.; Amador-Noguez, D.; Jefcoate, C.; Remucal, C.; Jung, S.; Kim, S. C.; et al. Cytochrome P450 Monooxygenase-Mediated Metabolic Utilization of Benzo[a]Pyrene by *Aspergillus* Species. *MBio* **2019**, *10* (3), 1–15. https://doi.org/10.1128/mBio.00558-19.

(164) Shin, J.; Kim, J. E.; Lee, Y. W.; Son, H. Fungal Cytochrome P450s and the P450 Complement (Cypome) of *Fusarium graminearum*. *Toxins (Basel).* **2018**, *10* (3), 76–91. https://doi.org/10.3390/toxins10030112.

(165) Kavanagh, K. L.; Jörnvall, H.; Persson, B.; Oppermann, U. Medium- and Short-Chain Dehydrogenase/Reductase Gene and Protein Families: The SDR Superfamily: Functional and Structural Diversity within a Family of Metabolic and Regulatory Enzymes. *Cell. Mol. Life Sci.* **2008**, *65* (24), 3895–3906. https://doi.org/10.1007/s00018-008-8588-y.

(166) Hussain, R.; Ahmed, M.; Khan, T. A.; Akhter, Y. Fungal P450 Monooxygenases - the Diversity in Catalysis and Their Promising Roles in Biocontrol Activity. *Appl. Microbiol. Biotechnol.* **2020**, *104* (3), 989–999. https://doi.org/10.1007/s00253-019-10305-3.

(167) Samson, Robert A. ; Houbraken, Jos ; Thrane, Ulf ; Frisvad, Jens Christian ; Andersen, B. *Food and Indoor Fungi: Second Edition.*; CBS-KNAW Fungal Biodiversity Centre: Utrecht, The Netherlands., 2010.

(168) Sun, X. Y.; Zhu, J. F.; Bao, L.; Hu, C. C.; Jin, C.; Harris, S. D.; Liu, H. W.; Li, S. J. PyrG Is Required for Maintaining Stable Cellular Uracil Level and Normal Sporulation Pattern under Excess Uracil Stress in *Aspergillus nidulans*. *Sci. China Life Sci.* **2013**, *56* (5), 467–475. https://doi.org/10.1007/s11427-013-4480-6.

(169) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem* **2002**, *3* (7), 619–627. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9.

Appendixes

Appendix 1 Taxonomy and Rareness Driven Discovery of Polyketides from *Aspergillus californicus*

Appendix 2 New naphthyl derivatives from *Aspergillus californicus*

Appendix 3 Oxepinamides L and M, two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides from *Aspergillus californicus*

Appendix 4 Review of Oxepine-Pyrimidinone-Ketopiperazine Type Nonribosomal Peptides

Appendix 5 Biosynthesis of Calipyridone A Represents a Rare Fungal 2-pyridone Product Formation without Ring Expansion in *Aspergillus californicus*

Appendix 6 Genetic origin of homopyrones, a rare type of phenylpropanoid and polyketide derived yellow pigments from *Aspergillus homomorphus*

Appendix 7 Efforts on linking oxepinamide L to its biosynthetic gene cluster and linking five genes to the cryptic metabolites in *Aspergillus californicus*

Taxonomy and Rareness Driven Discovery of Polyketides from *Aspergillus californicus*

Yaojie Guo,[†] Ling Ding,[†] Simone Ghidinelli,[‡] Charlotte H. Gotfredsen,[§] Mercedes de la Cruz,[^] Thomas A. Mackenzie,[^] Carmen R. Martín,[^] Pilar Sánchez,[^] Francisca Vicente,[^] Olga Genilloud,[^] Sonia Coriani,[§] René Wugt Larsen,[§] Jens C. Frisvad,[†] and Thomas O. Larsen[*†]

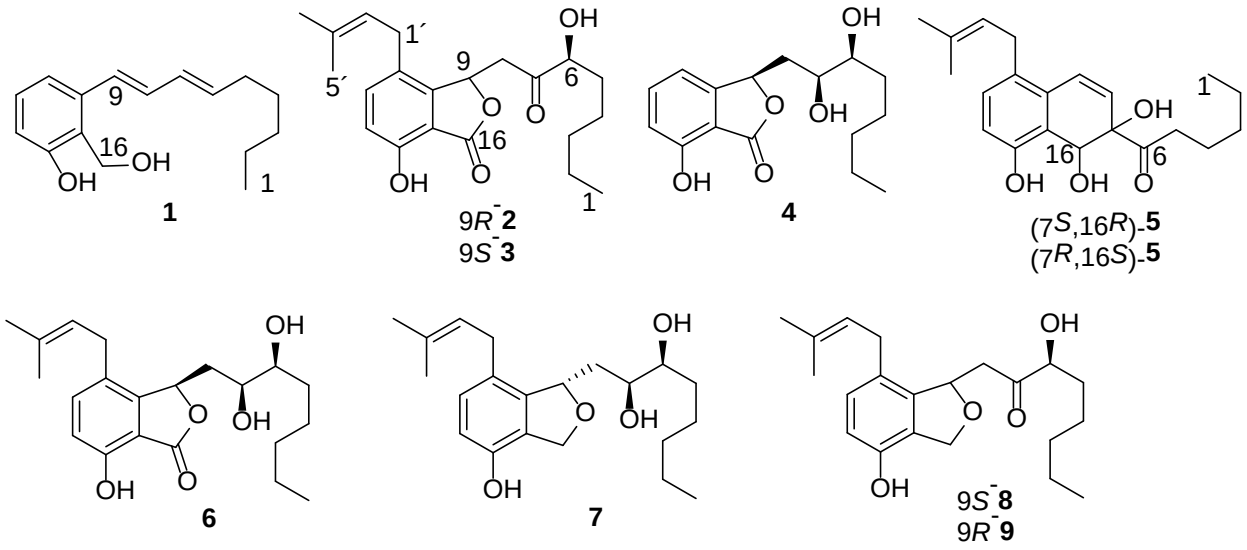
[†]Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800, Kgs. Lyngby, Denmark

[‡]Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, I-25123, Brescia, Italy

[§]Department of Chemistry, Technical University of Denmark, Kemitorvet, DK-2800, Kgs. Lyngby, Denmark

[^]Fundación MEDINA, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016, Armilla, Granada, Spain

ABSTRACT: Five new polyketides were isolated from the rare species *Aspergillus californicus* IBT 16748 including calidiol A (**1**), three phthalide derivatives califuranone A_1, A_2 and B (**2**-**4**), and a pair of enantiomers (-)-calitetralintriol A (-**5**) and (+)-calitetralintriol A (+**5**) together with four known ones (**6**-**9**). The structures of the new products were established by extensive spectroscopic analyses including HRMS, 1D and 2D NMR. The absolute configurations of two diastereomers (**2**, **3**) and the enantiomers (**5**) were solved by comparing their experimental and calculated ECD data, and the absolute configuration of **4** was assigned by NOESY and comparison with other analogues. The antibacterial and cytotoxic properties of **1**-**3** and **5**-**7** were evaluated. Compound **1** showed moderate activity against methicillin-resistant *Staphylococcus aureus*.



Aspergillus is a diverse genus of filamentous fungi with approximately 446 known species that have been accepted. Species of this genus are known to be a rich source of secondary metabolites (SMs) and some species have been extensively used in industry to produce organic acids, medicines, and enzymes.[1–3] For example, *Aspergillus niger* is used to produce citric acid and the annual yield production was up to 1.6 million tons in 2007.[2,4] In the pharmaceutical industry, *A. terreus* is known to produce lovastatin, which is a very successful cholesterol-lowering drug on the market.[5] Echinocandin B was first isolated from *A. spinulosporus* (= *A. delacroxii*, formerly *Aspergillus nidulans var. echinulatus*) and it led to the discovery of the fourth class of antifungal drugs against the human pathogenic fungi *Candida spp*.[6,7] One of the *Aspergillus* subgenera containing most species is subgenus *Nidulantes*, which compromises the sections *Aenei, Bispori, Cavernicolarum, Nidulantes, Ochraceorosei, Raperorum, Silvatici, Sparsi and Usti*.[3,8] The *Nidulantes* subgenus encompasses 128 species of which 86 species are very rare (67 %), with only 1 or 2 isolates known for each species.[3,8–23]. Such rare species often have a rich diversity of new natural products.

Here we explore the rare species *Aspergillus californicus*, of which only one isolate is known despite that it was isolated in 1978 Within subgenus *Nidulantes*, this species was previously placed in section *Usti* but later re-assigned to section *Cavernicolarum* based on a multigene phylogeny using concatenated 4 loci ITS, BenA, CaM and RPB2.[16,18] Recently, the entire genome sequence data of this species became available (https://genome.jgi.doe.gov/portal/). AntiSMASH[24] results of its whole genome sequence predicted that this species contains 91 secondary metabolite regions, and in-house UHPLC-DAD-HRMS dereplication based analysis of the extracts using OSMAC (One Strain - Many Compounds) approach[25] indicated that this strain can produce many possible unknown metabolites. This paper deals with the characterization of nine polyketides from *A.*

californicus including five previously unknown compounds calidiol A (**1**), califuranone A_1, A_2 and B (**2-4**), calitetralintriol A (**5**) as well as known compounds emefuranone A_2 (**6**), emefuran A (**7**) and a mixture of diastereomers emefuan B_1 and B_2 (**8**, **9**). Compounds **2-4** and **6** bear a phthalide moiety while **7-9** contain a phthalane scaffold. All compounds reported here are likely to originate from the same biosynthetic pathway. The antibacterial and cytotoxic properties of **1-3** and **5-7** were evaluated.

RESULTS AND DISCUSSION

The molecular formula of **1** was assigned as $C_{16}H_{22}O_2$ based on HRMS analysis. The ^{1}H NMR exhibited 20 proton signals (**Table 1**) including one methyl group δ_H 0.90 (t, J = 7.2 Hz, H_3-1), eight alkyl hydrogens δ_H 1.31-2.14, seven olefinic hydrogens δ_H 5.84-7.14, and two oxygenated benzylic methylene protons δ_H 4.99 (s, H_2-16). The COSY correlations (**Figure 1**) between δ_H 7.14 (t, J = 7.8 Hz, H-12) and 6.76 (d, J = 7.8 Hz, H-13), and δ_H 7.00 (d, J = 7.8 Hz, H-11) suggested three adjacent hydrogens on a benzene ring. The HMBC correlations from H-13 and H-11 to δ_C 121.8 (C-15), and from H-12 to δ_C 156.7 (C-14) and 137.3 (C-10) indicated that the remaining three positions on the aromatic ring were all substituted. More HMBC resonances from H_2-16 to C-15, C-14 and C-10, δ_H 6.58 (H-9) to C-11, and 6.57 (H-8) to C-10 suggested that C-15 was substituted with an oxygenated methylene group and that C-10 was attached with an olefinic unit. The COSY correlations δ_H 6.20 (ddq, J = 15.2, 7.4, 1.1 Hz, H-7) through H_3-1 determined the long chain. In $CDCl_3$ the resonances of H-8 and H-9 are overlapped. However, a ^{1}H NMR spectrum acquired in CD_3OD allowed for the differentiation of the two by showing a clear correlation between δ_H 6.83 (d, J = 15.6 Hz, H-9) and 6.68 (dd, J = 15.6, 10.4 Hz, H-8), and similar between H-8 and δ_H 6.28 (ddq, J = 15.0, 10.4, ~1 Hz, H-7). The double bond between C-6 and C-7 was assigned to be *E* configuration based on their coupling constant value 15.2 Hz and the same

resonance applies to the double bond between C-8 and C-9. C-14 was determined to be substituted by a phenol group and C-16 was attached by a primary alcohol group due to the two exchangeable hydrogen signals in the ^{1}H NMR and the downfield of C-14 and C-16 signals as well as two C-O stretching vibration peaks at 1206 and 1144 cm^{-1} in the IR spectrum (**Figure S1**). Thus, the structure of **1** named calidiol A was elucidated, which was close to a known salicylic alcohol derivative 1-hydroxy-2-hydroxymethyl-3-pent-1,3-dienylbenzene.[26,27] It´s notable that several similar compounds bearing a phenol and an adjacent primary alcohol group were likely misassigned as oxetanes.[28]

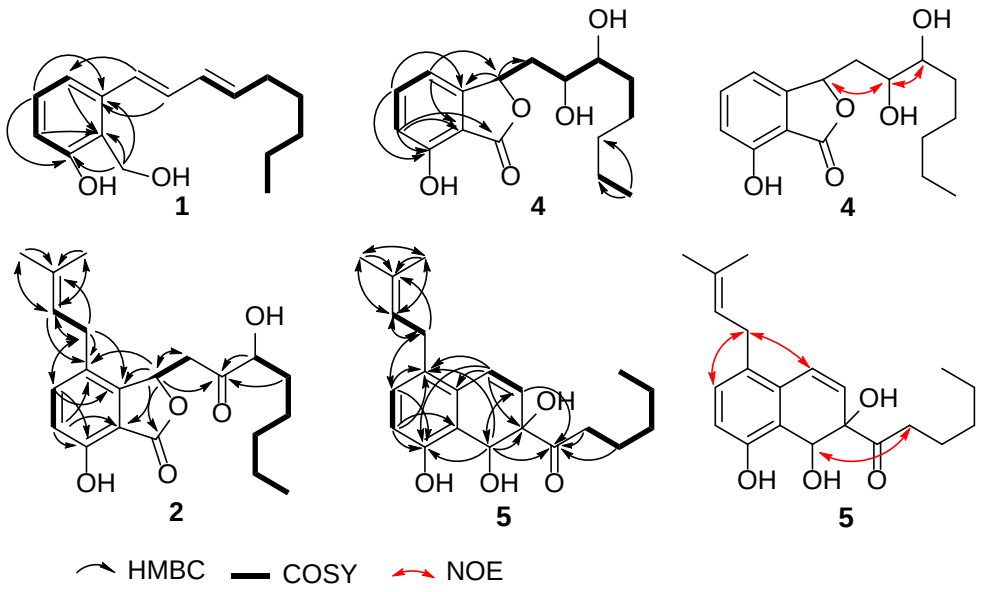


Figure 1. Selected HMBC, COSY and NOE correlations of **1**, **2**, **4** and **5**.

The formula of **2** was established as $C_{21}H_{28}O_5$ according to HRMS analysis. The ^{1}H NMR spectrum of **2** showed twenty-six hydrogens (**Table 1**) and the ^{13}C NMR displayed a keto carbon δ_C 208.0 (C-7) and an ester carbon δ_C 171.7 (C-16). The phthalide core was determined based on HMBC correlations (**Figure 1**) from H-9 to C-16, δ_C 146.2 (C-10) and 110.9 (C-15), H-12 to C-10 and δ_C 155.3 (C-14), and from H-13 to δ_C 127.3 (C-11) and C-15. The position of the prenyl

group was determined to be on C-11 by HMBC correlations from H_2-1´ to C-10, C-11 and C-12, and H-2´ to C-11. The long chain was assigned through HMBC correlations from H-9 to C-7, from H_2-5 to C-7 together with COSY correlations of the two spin systems H-9 through H_2-8 and H-6 through H_3-1. The main structure of **2** turned out to be an oxygenated form of emefuranone A_2 at C-7.[29] The absolute configuration of **2** was assigned by comparison of its experimental electronic circular dichroism (ECD) spectrum with the calculated data of the structure (see **Computational Details**).[30] As shown in **Figure 2A**, the calculated ECD curve of (6*S*, 9*R*)-**2** matched the experimental spectrum recorded in acetonitrile. The signals including a positive band at around 200 nm, two negative bands near 220 and 240 nm, and a broad positive band around 280 nm were well reproduced by the time-dependent density-functional-theory (TDDFT) calculation both in sign and relative intensities. Thus, the structure of **2** califuranone A_1 was determined.

Compound **3** was determined as a diastereomer of **2** based on HRMS, MS/MS, IR and 1D NMR comparisons (**Table 1, Figure S8**-**S11** and **S16**-**S19**). The chiral centers C-6 and C-9 were both established to be *S* configuration as the calculated ECD spectrum of (6*S*, 9*S*)-**3** (**Figure 2B**) agreed well with the experimental data.

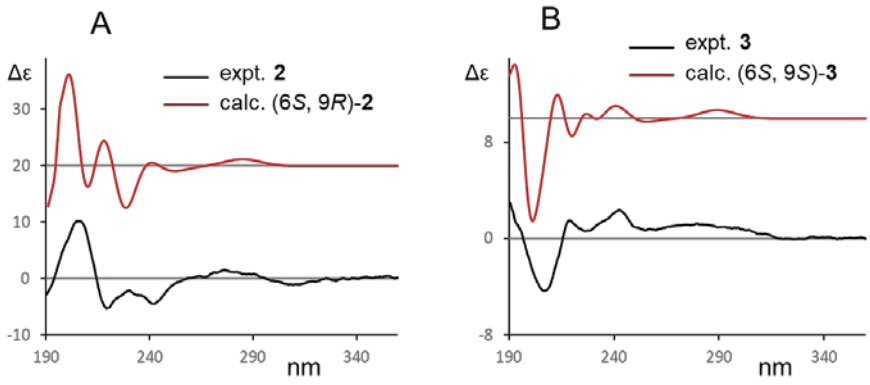


Figure 2. Comparison of the experimental and calculated ECD spectra of **2** (A) and **3** (B).

The chiral center at C-6 of **2** and **3** are both *S* configuration, which is consistent with other analogues from *Emericella* sp. IFM57991 including emefuranone A_1, emefuranone A_2, emefuran A, emefuran B_1, emefuran B_2, emefuran C_2, and emefuran C_2.[29] Under the "one fungus: one name" nomenclatural system, *Emericella* species have all been transferred to *Aspergillus* subgenus *Nidulantes*.[8,16] The strain IFM57991 was reported to have curved hülle cells, but no conidia and ascomata. So this strain belongs to subgenus *Nidulantes* section *Usti* or *Cavernicolarum* as several species in *Usti* and *Cavernicolarum* have curved hülle cells.[16,18,19] Considering *A. californicus* belonging to section *Cavernicolarum*, we assume all similar phthalide and phthalan products from a sixteen-carbon polyketide biosynthetic pathway (**Figure S40**) in sections *Usti* and *Cavericolarum* share *S* configuration at C-6.

The chemical formula of compound **4** was established as $C_{16}H_{22}O_5$ based on a $[M+H]^+$ ion *m/z* 295.1549. Its structure was also similar to emefuranone A_2[29] but without no prenyl group substitution on C-11. It is plausible to assign C-6 to have *S* configuration using a similar approach as described above. Besides, the NOE correlations (**Figure 1**) between H-6 and H-7, H-7 and H-9 suggested all these three hydrogens to be α-oriented resulting in configurations of C-7 and C-9 to be *S* and *R*, respectively. Thus, the structure of **4** named califuranone B was solved.

Compound **5** was first isolated as a racemate, and to explore the potential bioactivities of each enantiomer it was separated by a chiral column to yield (-)-**5** and (+)-**5**. The NMR data of **5** described here were obtained with the mixture. The molecular formula of **5** was established as $C_{21}H_{28}O_4$ with eight degrees of unsaturation based on a $[M + Na]^+$ ion at m/z 367.1883 and $[M - H]^-$ at m/z 343.1915. The ^{1}H NMR data (**Table 1**) showed twenty-five hydrogens including three methyl groups δ_H 0.86 (t, J = 7.3 Hz, H_3-1), 1.70 (d, J = 1.4 Hz, H_3-4´), 1.73 (s, H_3-5´), ten methylene hydrogens δ_H 1.19 (m, H_2-3), 1.24 (m, H_2-2), 1.54 (quintet, J = 7.4, H_2-4), 2.59 (m, H_2-

5), 3.36 (H_a-1´), 3.30 (H_b-1´), one methine hydrogen δ_H 5.35 (s, H-16), five olefinic hydrogens δ_H 5.16 (tq, *J* = 7.0, 1.4 Hz, H-2´), 5.83 (d, *J* = 10.1 Hz, H-8), 6.83 (d, *J* = 10.1, H-9), 6.97 (d, *J* = 8.5 Hz, H-12) and 6.66 (d, *J* = 8.5 Hz, H-13). The tetralin nucleus was assigned by the HMBC correlations from H-16 to δ_C 127.3 (C-8) and 154.3 (C-14), from H-8 to δ_C 73.3 (C-16), from H-9 to δ_C 80.0 (C-7), 129.9 (C-10), C-11 and C-14, from H-12 to C-14, and from H-13 to δ_C 129.6 (C-15) as well as the COSY relationships between H-8 and H-9, H-12 and H-13. The position of the prenyl group was determined based on the HMBC correlation from H_2-1´ to C-12, which was supported by the NOE correlations between H_2-1´ and H-9, H_2-1´ and H-12. The 1-keto-hexyl chain on C-7 was determined by HMBC correlations from H-16, H-8 and H_2-4 to δ_C 209.9 (C-6), together with the COSY resonances between H_2-5 and H_2-4, H_2-4 and H_2-3, H_2-3 and H_2-2, and H_2-2 and H_3-1. NOE correlations between H-16 and H_2-5 indicated H-16 and H_2-5 should be on the same side while 16-OH and 7-OH on the other side. Thus, the two enantiomers should be (7*S*, 16*R*) and (7*R*, 16*S*). ECD measurement and theoretical predictions were applied to solve the configuration of the two chiral centers at C-7 and C-16. The TDDFT calculated spectrum of (7*S*, 16*R*)-configuration matched the experimental curve of (-)-**5** and the calculated curve of (7*R*, 16*S*)-configuration agreed with the experimental data of (+)-**5** (**Figure S39**). Thus, the chiral centers of (-)-**5** were determined as (7*S*, 16*R*)-configuration, and (+)-**5** were assigned as (7*R*, 16*S*)-configuration.

Besides, the structure of **6** was confirmed to be identical to that of emefuranone A_2 based on identical ^{1}H NMR, MS and optical rotation data to the once reported in literature[29]. Similarly compound **7** was verified by comparing its ^{1}H NMR, MS to be emefuran A[29]. Finally, the mixture of **8** and **9** was identified to be a mixture of emefuran B_1 and B_2[29] based on ^{1}H NMR and HRMS data.

Table 1. ^{1}H (800 MHz) and ^{13}C NMR (200 MHz) Data for 1-6

position	**1**[a]		**2**[a]		**3**[a]		**4**[b]		**5**[b]	
	δ_C[C], type	δ_H, (J in Hz)	δ_C, type	δ_H, (J in Hz)	δ_C, type	δ_H, (J in Hz)	δ_C, type	δ_H (J in Hz)	δ_C[C], type	δ_H, (J in Hz)
1	14.2, CH_3	0.90, t (7.2)	14.1, CH_3	0.88, t (7.1)	14.1, CH_3	0.89, t (7.0)	14.4, CH_3	0.92, t (7.1)	12.9, CH_3	0.86, t (7.3)
2	22.7, CH_2	1.32, m	22.6, CH_2	1.31, m	22.6, CH_2	1.32, m	23.7, CH_2	1.35, m	22.0, CH_2	1.24, m
3	31.6, CH_2	1.31, m	31.7, CH_2	1.31, m	31.7, CH_2	1.32, m	33.1, CH_2	1.35, m	31.1, CH_2	1.19, m
4	29.1, CH_2	1.43, quintet (7.4)	24.7, CH_2	1.47, m 1.37, m	24.6, CH_2	1.44, m, overlapped	26.6, CH_2	1.56, m 1.35, m	22.9, CH_2	1.54, quintet (7.4)
5	33.0, CH_2	2.14, m	33.2, CH_2	1.79, m 1.53, m	33.6, CH_2	1.81, m 1.56, m	33.9, CH_2	1.60, m 1.35, m	36.7, CH_2	2.59, m
6	137.0, CH	5.84, dt (15.2, 7.4)	77.5, CH	4.23, dd (7.9, 3.7)	77.0, CH	4.21, dd (8.0, 3.8)	76.1, CH	3.44, m	209.9, C	
7	130.4, CH	6.20, ddq (15.2, 7.4, 1.1)	208.0, C		207.8, C		72.3, CH	3.75, (ddd, 10.8, 5.6, 2.0)	80.0, C	
8	133.1, CH	6.58, overlapped	41.7, CH_2	3.11, dd (17.0, 2.2) 2.81, dd (17.0, 9.7)	41.7, CH_2	3.08, dd (17.5, 2.5) 2.95, dd (17.5, 9.2)	39.7, CH_2	2.04, ddd (14.7, 10.8, 2.0) 1.81, ddd (14.7, 10.8, 2.0)	127.3, CH	5.83, d (10.1)
9	126.3, CH	6.58, overlapped	77.8, CH	6.03, dd (9.7, 2.2)	77.3, CH	6.06, dd (9.2, 2.5)	79.8, CH	5.69, dd (10.8, 2.0)	127.0, CH	6.83, d (10.1)
10	137.3, C		146.2, C		146.2, C		154.2[C], C		129.9, C	
11	118.4, CH	7.00, d (7.8)	127.3, C		127.3, C		113.7, CH	6.97, d (7.5)	118.7, C	
12	129.1, CH	7.14, t (7.8)	138.0, CH	7.35, d (8.3)	137.9, CH	7.35, d (8.3)	137.7, CH	7.55, dd (8.2, 7.5)	130.1, CH	6.97, d (8.5)
13	115.6, CH	6.76, d (7.8)	116.5, CH	6.90, d (8.3)	116.4, CH	6.90, d (8.3)	116.6, CH	6.88, d (8.2)	116.4, CH	6.66, d (8.5)
14	156.7, C		155.3, C		155.3, C		158.3[C], C		154.3, C	
15	121.8, C		110.9, C		110.9, C		112.5[C], C		129.6, C	
16	60.3, CH_2	4.99, s	171.7, C		171.7, C		171.8[C], C		73.3, CH	5.35, s
1´			30.3, CH_2	3.30, dd (15.9, 6.9) 3.23, dd (15.9, 6.9)	30.3, CH_2	3.29, dd (16.2, 7.2) 3.07, dd (16.2, 7.2)			30.6, CH_2	3.36[d] 3.30[d]
2´			121.2, CH	5.16, m	121.1, CH	5.16, m			123.7, CH	5.16, tq (7.0, 1.4)
3´			134.7, C		134.7, C				131.1, C	
4´			25.8, CH_3	1.75, d (1.3)	25.8, CH_3	1.75, d (1.3)			24.5, CH_3	1.70, d (1.4)
5´			18.2, CH_3	1.68, s	18.1, CH_3	1.67, s			16.6, CH_3	1.73, s

[a] Recorded in $CDCl_3$; [b] Recorded in CD_3OD; [c] Chemical shifts assigned by HSQC and HMBC correlations. [d] Signals partially obscured by the solvent peak, see **Figure S34 – S36**.

The compounds elucidated here add to the large chemical diversity of *Aspergillus* subgenus *Nidulantes*.[1,16,31] So far in genus *Aspergillus,* polyketides with a phthalide moiety have been reported from subgenus *Fumigati* section *Fumigati* (*Aspergillus duricaulis*),[32] several phthalides from subgenus *Nidulantes* including asperlide in *A. unguis*,[33] porriolides (silvaticols, and the nitrogen-containing cichorine) in *A. nidulans* and *A. silvaticus*,[34,35] emefuranones in an *Aspergillus* ("*Emericella*") species from section *Usti* or *Cavernicolarum*,[29] and austalides in *A. ustus*,[36] *A. aureolatus*[37] and *A. alabamensis*[38], and mycophenolic acids from subgenus *Aspergillus* section *Restricti* (*A. glabripes, A. gracilis, A. pachycaulis, A. penicillioides* and *A. tardicrescens*),[39] section *Aspergillus* (*A. pseudoglaucus*),[40] and also from subgenus *Fumigati* section *Fumigati* (*A. unilaterialis*).[41] However, none of those phthalides resembles **2-4** in *A. californicus* except the emefuranones. Considering that compounds **1-9** all contain a backbone of sixteen carbons we hypothesize that they originate from the same polyketide biosynthetic pathway (**Figure S40**).

Compounds **1-3** and **5-7** were subjected to antibacterial and cytotoxic assays. To assess their antibacterial properties they were tested against two Gram-positive bacteria (methicillin-resistant *Staphylococcus aureus* MB5393 and methicillin-sensitive *Staphylococcus aureus* ATCC 29213), and two Gram-negative bacteria (*Escherichia coli* ATCC 25922 and *Klebsiella pneumoniae* ATCC 700603). Compound **1** showed moderate activities against both Gram-positive bacteria with minimum inhibitory concentration (MIC) values of 48 μg/ml, whereas no activities were observed against Gram-negative bacteria. The remaining compounds were inactive with MIC values higher than 64 μg/ml against all microorganisms tested (**Table S1**). Six tumor cell lines A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast), Mia PaCa-2 (pancreas), HL-60 (leukemia) were used in the cytotoxicity test. None of them exhibited promising cytotoxic activities as their IC_{50} values exceeded 10 μM (**Table S2**).

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were measured on a PerkinElmer 341 polarimeter. ECD spectra were recorded in acetonitrile with a 2 mm path length cuvette in a JASCO J-1500 CD spectrophotometer. The infrared attenuated-total-reflectance (ATR) spectra were collected by a Bruker VERTEX 80v Fourier Transform vacuum spectrometer employing a single-reflection diamond ATR accessory. The apparatus was configured with a Ge on KBr beam splitter, a liquid nitrogen cooled HgCdTe detector and a thermal globar radiation source. Extended ATR corrections were applied to account for the wavelength-dependent penetration depth of the infrared probe beam. The NMR spectra were recorded on Bruker AVANCE III 800 MHz spectrometer equipped with 5 mm TCI Cryoprobes using standard pulse sequences. Chemical shifts were reported in ppm with reference to the solvent signals ($CDCl_3$ δ_H 7.26 and δ_C 77.16, CD_3OD δ_H 3.31 and δ_C 49.00). UHPLC-DAD-HRESIMS was recorded on an Agilent Infinity 1290 UHPLC - 6545 QTOF MS system. The separation was performed on an Agilent Poroshell 120 phenyl-hexyl column (250 × 2.1 mm, 2.7 μm) at 60 °C with a linear gradient program: mobile phase (MP) A H_2O and B MeCN, both buffered with 20 mM formic acid, at a flow rate of 0.35 mL/min: 0-15 min 10%-100% MP B, 15-17min 100% MP B, 17.1-20 min 10% MP B. MS was detected in the positive ion mode with an Agilent Dual Jet Stream electrospray ion (ESI) source with drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature of 300 °C and flow of 12 L/min. The capillary voltage was set to 4000 V, and nozzle voltage to 500 V. Mass spectra were recorded in the range *m*/*z* 100-1700, with a scan rate of 10 spectra/s. Automated data-dependent acquisition MS/HRMS analysis was performed for ion detected in the full scan above 5 000, applying fixed collision energies of 10, 20, and 40 eV with a maximum of three selected precursor ions per cycle. Lock mass solution in 70% MeOH was infused in the second sprayer

using an extra LC pump at a flow of 15 μL/min; the solution contained 1 μM tributylamine (Sigma-Aldrich) and 10 μM hexakis (2,2,3,3-tetrafluoropropoxy) phosphazene (Apollo Scientific Ltd.) with *m/z* 186.2216 and 922.0098, respectively. Flash chromatography was carried out on a Biotage Isolera One system. Semi-preparative HPLC was performed on either a Dionex Ultimate 3000 HPLC system with a DAD or a Waters 600 HPLC with a Photodiode Array (PDA) detector. The columns used in the isolation were Phenomenex Kinetex C_{18} column (250 × 10 mm, 5 μm), Phenomenex Luna phenyl-hexyl column (250 × 10 mm, 5 μm) and Phenomenex Lux cellulose-1 (100 × 4.6 mm, 3 μm) column. For extraction and separation, the solvents were HPLC grade and for HRMS the chemicals were LCMS grade.

Fungal Material. *Aspergillus californicus* IBT 16748 was from IBT culture collection, Department of Biotechnology and Biomedicine, Technical University of Denmark. The strains IBT 16748 = CBS 123895, are both ex-type cultures of *Aspergillus californicus*.

Incubation, Extraction and Isolation. The medium used was medium 2^{42} with 2% (w/v) agar. The strain was incubated (three-point stabs) on 270 plates at 25°C for 15 days. The plate contents were extracted twice with EtOAc containing 1% formic acid (FA). The crude extract (2.6 g) was dissolved in 90% MeOH and extracted with heptane. The 90% MeOH part was diluted to 50% MeOH before extracted with dichloromethane (DCM) twice. The DCM extract (1.1 g) was subjected to Isolera One with a Diol column (25 g, 33 mL) gradient elution DCM-EtOAc-MeOH at 25 mL/min to afford nine fractions. Fr.3 (0.15 g) was subjected to the C_{18} column ($MeCN/H_2O$, 50-80% in 9 min, 4 mL/min) to yield **7** (8.1 mg, t_R = 7.3 min) and a sub-fraction (12.1 mg) which was further applied to the cellulose-1 column ($MeCN/H_2O$, 40-43.5% in 10 min, 2 mL/min) to yield **2** (2.2 mg, t_R = 8.3 min) and **3** (2.2 mg, t_R = 9.1 min). Fr.4 (0.1 g) was applied to the C_{18} column to offer a sub-fraction (1.3 mg) which was further purified on the phenyl-hexyl column

(MeCN/H_2O, 63% isocratic in 20 min, 4 mL/min) to afford **1** (0.87 mg, t_R = 11.8 min). Fr.5 (0.2 g) was subjected to flash chromatography with a self-packed C_{18}-T column (Phenomenex Sepra, 25 g, 33 mL) to yield 14 sub-fractions, and sub-fr.4 (11.5 mg) was further applied to the Waters HPLC first with the phenyl-pexyl column and then C_{18} column (MeCN/H_2O, 30% isocratic in 20 min, 4 mL/min) to yield **4** (0.5 mg, t_R = 13.0 min). Compounds **5** and **7**-**9** were isolated from the extract of 120 CYA and 120 YES plates for 11 days at 25℃. The DCM part (1.3 g) was applied to flash chromatography with a self-packed NH_2 column (Sepra Phenomenex, 50g, 66 mL) to yield eight fractions. Fr6 (53 mg) was subjected to flash chromatography with a self-packed C_{18}-T column (Sepra Phenomenex, 10 g, CV 15 mL) to give nine sub-fractions. Sub-fr7 (3.63 mg) was applied to a the C_{18} column (MeCN/H_2O, 70-100% in 20 min, 4 mL/min) to yield **5** (t_R = 16.3 min) which was further purified using the cellulose-1 column (MeCN/H_2O, 50-60% in 10 min, 2 mL/min) to give (-)-**5** (0.58 mg, t_R = 4.9 min) and (+)-**5** (0.51 mg, t_R = 5.5 min). Fr4 (80.4 mg) was subjected to flash chromatography with a self-packed C_{18}-T column (Sepra Phenomenex, 25 g, 33 mL) to give eight sub-fractions. Sub-fr5 (32.04 mg) was first applied to the C_{18} column and then to the phenyl-hexyl column (MeCN/H_2O, 60% isocratic in 15min, 4 mL/min) to yield **7** (4.8 mg, t_R = 10.5 min) and a mixture of **8** and **9** (0.76 mg, t_R = 7.0 min).

*Calidiol A (**1**):* white amorphous solid; HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 236, 286 nm; IR (dry film) ν_{max} 3345, 3019, 2956, 2926, 2855, 1677, 1606, 1579, 1467, 1378, 1272, 1206, 1144, 988, 847, 804, 726, 518 cm^{-1}; ^1H and ^{13}C NMR data see **Table 1**. HRESIMS *m*/*z* $[M+H]^+$ calcd for $C_{16}H_{23}O_2^+$, 247.1693, found 247.1693; $[M-H]^-$ calcd for $C_{16}H_{21}O_2^-$, 245.1552, found 245.1552;

*Califuranone A_1 (**2**):* white amorphous solid, $[\alpha]_D^{20}$ = +100 (*c* = 0.04, MeOH); HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 222, 308 nm; ECD (0.514 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 309 (-1.2), 278 (+1.1), 240 (-4.2), 221 (-5.4), 207 (+10.1) nm; IR (dry film) ν_{max} 3440, 2957, 2931, 2860, 1742,

1678, 1626, 1503, 1445, 1366, 1316, 1201, 1144, 1091, 1054, 1029, 829, 802, 725 cm^{-1}; ^{1}H and ^{13}C NMR data see **Table 1**. HRESIMS *m/z* $[M+H]^+$ calcd for $C_{21}H_{29}O_5^+$, 361.2009, found 361.2006; $[M+Na]^+$ calcd for $C_{21}H_{28}O_5Na^+$, 383.1829, found 383.1825;

*Califuranone A₂ (**3**):* white amorphous solid, $[\alpha]_D^{20}$ = -100 (*c* = 0.04, MeOH); HPLC-UV (MeCN/ H_2O + 20 mM FA) λ_{max} 222, 308 nm; ECD (0.375 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 280 (+1.2), 243 (+2.4), 219 (+1.5), 208 (-4.3) nm; IR (dry film) ν_{max} 3438, 2956, 2930, 2860, 1742, 1673, 1627, 1503, 1446, 1366, 1316, 1200, 1159, 1088, 1053, 1028, 846, 803, 726 cm^{-1}; ^{1}H and ^{13}C NMR data see **Table 1**. HRESIMS *m/z* $[M+H]^+$ calcd for $C_{21}H_{29}O_5^+$, 361.2009, found 361.2013; $[M+Na]^+$ calcd for $C_{21}H_{28}O_5Na^+$, 383.1829, found 383.1827;

*Califuranone B (**4**):* white amorphous solid, $[\alpha]_D^{20}$ = +8.7 (*c* = 0.046, MeOH); HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 236, 301 nm; IR (dry film) ν_{max} 3438, 2959, 1740, 1441, 1384, 1269, 1239, 1196, 1086, 935, 782 cm^{-1}; ^{1}H and ^{13}C NMR data see **Table 1**. HRESIMS *m/z* $[M+H]^+$ calcd for $C_{16}H_{23}O_5^+$, 295.1540, found 295.1549; $[M+Na]^+$ calcd for $C_{16}H_{22}O_5Na^+$, 317.1359, found 317.1359;

(-)-*calitetralintriol A (**5**):* yellow amorphous solid, $[\alpha]_D^{20}$ = -32.8 (c = 0.058, MeCN); HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 236, 270, 318 nm; ECD (2.11 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 310 (-9.4), 267 (+7.0), 220 (-5.3) nm; ^{1}H and ^{13}C NMR data see **Table 1**. HRESIMS *m/z* $[M+Na]^+$ calcd for $C_{21}H_{28}O_4Na^+$, 367.1880, found 367.1883; $[M-H]^-$ calcd for $C_{21}H_{27}O_4^-$, 343.1915, found 343.1915.

(+)-*calitetralintriol A (**5**):* $[\alpha]_D^{20}$ = +37.3 (*c* = 0.051, MeCN); ECD (2.22 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 310 (+9.4), 267 (-7.0), 220 (+5.3) nm.

Computational Details. To reduce to computational cost, the conformational search and the DFT calculations of **2**, **3**, (-)-**5** and (+)-**5** were carried out by cutting three adjacent methylenes (-$CH_2CH_2CH_2$-) of the chain as they are distant from the chiral centers and the electronic chromophore. Molecular Mechanics search of conformers was performed using Spartan program[43] with MMFF94s as force field. Conformers within 5 kcal/mol from the most stable one have been optimized at DFT level using the B3LYP functional and the TZVP basis set and including acetonitrile solvent effects at the polarizable continuum model in the integral equation formalism (IEF-PCM).[44–46] The Gaussian16 program package was used for all DFT and TDDFT calculations.[47] Oscillator strengths (for the UV spectra) and rotatory strengths (for the ECD spectra) of the first 60 excited states of each conformer have been calculated at TDDFT CAM-B3LYP/TZVP level of theory in IEF-PCM approximation. The final theoretical UV and ECD spectra were obtained as Boltzmann averages of the spectra of the individual conformers. A Gaussian band shape with a bandwidth of 0.2 eV was chosen to plot UV and ECD spectra using SpecDis package.[48] All calculated ECD and UV absorption spectra were redshifted by 20 nm.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

HRESIMS, IR, 1D and 2D NMR data of **1**-**6**; the predominant conformers of **2**, **3** and (-)-**5**; experimental procedures and results for antibacterial and cytotoxic assays; proposed biosynthetic pathways of the polyketides **1**-**9**.

AUTHOR INFORMATION

Corresponding Author

*Tel: +45 45252632. Fax: +45 45884922. E-mail: tol@bio.dtu.dk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Kasper Enemark-Rasmussen for acquiring NMR data, and the NMR Center • DTU and the Villum Foundation are acknowledged for access to the 800 MHz spectrometers. Yaojie Guo is financially supported by the China Scholarship Council (No. 201709110107) and the Department of Biotechnology and Biomedicine at the Technical University of Denmark. Sonia Coriani acknowledges support from the Independent Research Fund Denmark – Natural Sciences, DFF-RP2 (grant No. 7014-00258B). Simone Ghidinelli thanks the University of Brescia and MIUR for a visiting grant to DTU under the auspices of the Doctorate program. The work in the present study was supported by the Novo Nordisk Foundation (grant No. NNF15OC0016610, TOL).

REFERENCES

(1) Frisvad, J. C.; Larsen, T. O. Chemodiversity in the genus *Aspergillus*. *Appl. Microbiol. Biotechnol.* **2015**, *99* (19), 7859–7877.

(2) Frisvad, J. C.; Møller, L. L. H.; Larsen, T. O.; Kumar, R.; Arnau, J. Safety of the fungal workhorses of industrial biotechnology: update on the mycotoxin and secondary metabolite potential of *Aspergillus niger*, *Aspergillus oryzae*, and *Trichoderma reesei*. *Appl. Microbiol.*

Biotechnol. **2018**, *102* (22), 9481–9515.

(3) Houbraken, J., Kocsubé, S., Visagie, C.M., Yilmaz, N., Wang, X.-C., Meijer, M., Kraak, B., Hubka, V., Samson, R.A., Frisvad, J. C. Classification of *Aspergillus*, *Penicillium*, *Talaromyces* and related Genera (Eurotiales), an overview of families, genera, subgenera, sections, series and species. *Stud. Mycol.* **2020**, *in press*, 1-165.

(4) Knuf, C.; Nielsen, J. Aspergilli: systems biology and industrial applications. *Biotechnol. J.* **2012**, *7* (9), 1147–1155.

(5) Butler, M. S. The role of natural product chemistry in drug discovery. *J. Nat. Prod.* **2004**, *67* (12), 2141–2153.

(6) Benz, V. F.; Knüsel, F.; Nüesch, J.; Treichler, H.; Voser, W.; Nyfeler, R.; Keller-Schierlein, W. Metabolites of Microorganisms, 143: Echinocandin B, a novel polypeptide-antibiotic from *Aspergillus ndulans* var echinulatus-isolation and structural components. *Helv. Chim. ACTA* **1974**, *57* (8), 2459–2477.

(7) Denning, D. W. Echinocandin antifungal drugs. *Lancet* **2003**, *362* (9390), 1142–1151.

(8) Samson, R. A.; Visagie, C. M.; Houbraken, J.; Hong, S. B.; Hubka, V.; Klaassen, C. H. W.; Perrone, G.; Seifert, K. A.; Susca, A.; Tanney, J. B.; et al. Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Stud. Mycol.* **2014**, *78* (1), 141–173.

(9) Romero, S. M.; Comerio, R. M.; Barrera, V. A.; Romero, A. I. *Aspergillus fuscicans* (Aspergillaceae, Eurotiales), a new species in section *Usti* from Argentinean semi-arid soil. *Phytotaxa* **2018**, *343* (1), 67–74.

(10) Siqueira, J. P. Z.; Wiederhold, N.; Gené, J.; García, D.; Almeida, M. T. G.; Guarro, J. Cryptic *Aspergillus* from clinical camples in the USA and description of a new species in section *Flavipedes*. *Mycoses* **2018**, *61* (11), 814–825.

(11) Jurjevic, Z.; Peterson, S. W. *Aspergillus Asper* sp. nov. and *Aspergillus collinsii* sp. nov., from *Aspergillus* section *Usti*. *Int. J. Syst. Evol. Microbiol.* **2016**, *66* (7), 2566–2572.

(12) Kozlovskii, A. G.; Antipova, T. V.; Zhelifonova, V. P.; Baskunov, B. P.; Ivanushkina, N. E.; Kochkina, G. A.; Ozerskaya, S. M. Secondary metabolites of fungi of the *Usti* section, genus *Aspergillus* and their application in chemosystematics. *Microbiol. (Russian Fed.* **2017**, *86* (2), 176–182.

(13) Nováková, A.; Hubka, V.; Saiz-Jimenez, C.; Kolarik, M. *Aspergillus baeticus* sp. nov. and *Aspergillus thesauricus* sp. nov., two species in section *Usti* from Spanish caves. *Int. J. Syst. Evol. Microbiol.* **2012**, *62* (11), 2778–2785.

(14) Long, W. Four new records of *Aspergillus* section *Usti* from Shandong Province. *Mycotaxon* **2012**, *129* (3), 373–384.

(15) Sutton, D. A.; Wickes, B. L.; Romanelli, A. M.; Rinaldi, M. G.; Thompson, E. H.; Fothergill, A. W.; Dishop, M. K.; Elidemir, O.; Mallory, G. B.; Moonnamakal, S. P.; et al. Cerebral Aspergillosis caused by *Aspergillus granulosus*. *J. Clin. Microbiol.* **2009**, *47* (10), 3386–3390.

(16) Chen, A. J.; Frisvad, J. C.; Sun, B. D.; Varga, J.; Kocsubé, S.; Dijksterhuis, J.; Kim, D. H.; Hong, S. B.; Houbraken, J.; Samson, R. A. *Aspergillus* section *Nidulantes* (formerly *Emericella*): polyphasic taxonomy, chemistry and biology. *Stud. Mycol.* **2016**, *84*, 1–118.

(17) Bills, G. F.; Yue, Q.; Chen, L.; Li, Y.; An, Z.; Frisvad, J. C. *Aspergillus mulundensis* sp. nov., a new species for the fungus producing the antifungal echinocandin lipopeptides, mulundocandins. *J. Antibiot. (Tokyo).* **2016**, *69* (3), 141–148.

(18) Samson, R. A.; Varga, J.; Meijer, M.; Frisvad, J. C. New taxa in *Aspergillus* section *Usti*. *Stud. Mycol.* **2011**, *69*, 81–97.

(19) Houbraken, J.; Due, M.; Varga, J.; Meijer, M.; Frisvad, J. C.; Samson, R. A. Polyphasic taxonomy of *Aspergillus* section *Usti*. *Stud. Mycol.* **2007**, *59*, 107–128.

(20) Hubka, V.; Nováková, A.; Peterson, S. W.; Frisvad, J. C.; Sklenář, F.; Matsuzawa, T.; Kubátová, A.; Kolařík, M. A reappraisal of *Aspergillus* section *Nidulantes* with descriptions of two new Sterigmatocystin-producing species. *Plant Syst. Evol.* **2016**, *302* (9), 1267–1299.

(21) Varga, J., Frisvad, J.C. and Samson, R. A. *Aspergillus* sect. *Aenei* sect. nov., a new Section of the genus for *A. karnatakaensis* sp. nov. and some allied fungi. *IMA Fungus* **2010**, *1* (2), 197–205.

(22) Varga, J., Frisvad, J.C. and Samson, R. A. Polyphasic taxonomy of *Aspergillus* section sparsi. *IMA Fungus* **2010**, *1* (2), 187–195.

(23) Sklenář, F.; Jurjević; Peterson, S. W.; Kolařík, M.; Nováková, A.; Flieger, M.; Stodůlková, E.; Kubátová, A.; Hubka, V. Increasing the species diversity in the *Aspergillus* section *Nidulantes*: six novel species mainly from the indoor environment. *Mycologia* **2020**, *112* (2), 342–370.

(24) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87.

(25) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem* **2002**, *3* (7), 619–627.

(26) Gehrt, A.; Erkel, G.; Anke, H.; Anke, T.; Sterner, O. New hexaketide inhibitors of eukaryotic signal transduction. *Nat. Prod. Lett.* **1997**, *9* (4), 259–264.

(27) Zhao, Z.; Ying, Y.; Hung, Y. S.; Tang, Y. Genome mining reveals *Neurospora crassa* can

produce the Salicylaldehyde Sordarial. *J. Nat. Prod.* **2019**, *82* (4), 1029–1033.

(28) Kutateladze, A. G.; Holt, T.; Reddy, D. S. Natural products containing the oxetane and related moieties present additional challenges for structure elucidation: A DU8+ Computational Case Study. *J. Org. Chem.* **2019**, *84* (12), 7575–7586.

(29) Saito, T.; Itabashi, T.; Wakana, D.; Takeda, H.; Yaguchi, T.; Kawai, K. I.; Hosoe, T. Isolation and structure elucidation of new phthalide and phthalane derivatives, isolated as antimicrobial agents from *Emericella sp.* IFM57991. *J. Antibiot. (Tokyo).*2016, 89–96.

(30) Nugroho, A. E.; Morita, H. Circular dichroism calculation for natural products. *J. Nat. Med.* **2014**, *68* (1), 1–10.

(31) Alburae, N. A.; Mohammed, A. E.; Alorfi, H. S.; Jamanturki, A.; Asfour, H. Z.; Alarif, W. M.; Abdel-Lateff, A. Nidulantes of *Aspergillus* (formerly *Emericella*): a treasure trove of chemical diversity and biological activities. *Metabolites* **2020**, *10* (2), 1–28.

(32) Achenbach, H.; Muhlenfeld, A.; Uv, G. Phthalide Und Chromanole Aus *Aspergillus duricaulis*. *Liebgs Ann. der Chemie* **1985**, *1985*, 1596–1628.

(33) Klaiklay, S.; Rukachaisirikul, V.; Aungphao, W.; Phongpaichit, S.; Sakayaroj, J. Depsidone and phthalide derivatives from the soil-derived fungus *Aspergillus unguis* PSU-RSPG199. *Tetrahedron Lett.* **2016**, *57* (39), 4348–4351.

(34) Fujita, M., Yamada, M., Nakajima, S., Kawai, K. O-Methylation Effect on the C-13 Nuclear magnetic resonance signal of ortho-disubstítuted phenols and its application to structure determination of new phthalides from *Aspergillus silvaticus*. *Chem. Pharm. Bull.* **1984**, *32* (7), 2622–2627.

(35) Sanchez, J. F.; Entwistle, R.; Corcoran, D.; Oakley, B. R.; Wang, C. C. C. Identification and molecular genetic analysis of the cichorine gene cluster in *Aspergillus nidulans*.

Medchemcomm **2012**, *3* (8), 997–1002.

(36) Antipova, T. V.; Zaitsev, K. V.; Oprunenko, Y. F.; Ya. Zherebker, A.; Rystsov, G. K.; Zemskova, M. Y.; Zhelifonova, V. P.; Ivanushkina, N. E.; Kozlovsky, A. G. Austalides V and W, new meroterpenoids from the fungus *Aspergillus ustus* and their antitumor activities. *Bioorganic Med. Chem. Lett.* **2019**, *29* (22), 126708.

(37) Peng, J.; Zhang, X.; Wang, W.; Zhu, T.; Gu, Q.; Li, D. Austalides S-U, new meroterpenoids from the sponge-derived fungus *Aspergillus aureolatus* HDN14-107. *Mar. Drugs* **2016**, *14* (7), 1–9.

(38) Zhou, Y.; Mándi, A.; Debbab, A.; Wray, V.; Schulz, B.; Müller, W. E. G.; Lin, W.; Proksch, P.; Kurtán, T.; Aly, A. H. New austalides from the aponge-associated fungus *Aspergillus* sp. *European J. Org. Chem.* **2011**, No. 30, 6009–6019.

(39) Sklenář, F.; Jurjević; Zalar, P.; Frisvad, J. C.; Visagie, C. M.; Kolařík, M.; Houbraken, J.; Chen, A. J.; Yilmaz, N.; Seifert, K. A.; et al. Phylogeny of xerophilic Aspergilli (subgenus *Aspergillus*) and taxonomic revision of section *Restricti*. *Stud. Mycol.* **2017**, *88*, 161–236.

(40) Chen, A. J.; Hubka, V.; Frisvad, J. C.; Visagie, C. M.; Houbraken, J.; Meijer, M.; Varga, J.; Demirel, R.; Jurjević; Kubátová, A.; et al. Polyphasic taxonomy of *Aspergillus* section *Aspergillus* (formerly *Eurotium*), and its occurrence in indoor environments and food. *Stud. Mycol.* **2017**, *88*, 37–135.

(41) Frisvad, J.C. and Samson, R. A. Chemotaxonomy and morphology of *Aspergillus* fumigatus and related taxa. in modern concepts in *Penicillium* and *Aspergillus* classification; Plenum Press: New York, 1990.

(42) Cacho, R. A.; Jiang, W.; Chooi, Y. H.; Walsh, C. T.; Tang, Y. Identification and characterization of the echinocandin B biosynthetic gene cluster from *Emericella rugulosa*

NRRL 11440. *J. Am. Chem. Soc.* **2012**, *134* (40), 16781–16790.

(43) Spartan. Wavefunction Inc.: Irvine, CA. https://www.wavefun.com/.

(44) Gennaro, P.; Torsten, B. Good computational practice in the assignment of absolute configurations by TDDFT calculations of ECD spectra. *Chirality* **2016**, *28* (6), 466–474.

(45) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105* (8), 2999–3093.

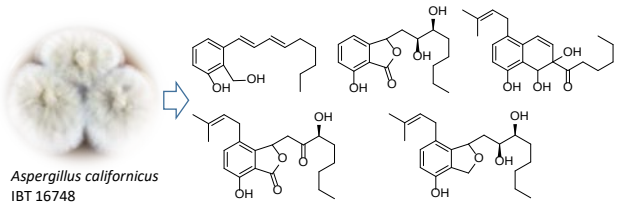
(46) Warnke, I.; Furche, F. Circular dichroism: electronic. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 150–166.

(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision C.01, Gaussian, Inc.: Wallingford CT 2016.

(48) Torsten, B.; Anu, S.; Yasmin, H.; Gerhard, B. SpecDis: Quantifying the comparison of calculated and experimental electronic circular dichroism spectra. *Chirality* **2013**, *25*, 243–

249.

Table of Contents (TOC)/Abstract Graphic



Aspergillus californicus
IBT 16748

Supporting information for

Taxonomy and Rareness Driven Discovery of Polyketides from *Aspergillus californicus*

Yaojie Guo,[†] Ling Ding,[†] Simone Ghidinelli,[‡] Charlotte H. Gotfredsen,[§] Mercedes de la Cruz,[^] Thomas A. Mackenzie,[^] Carmen R. Martín,[^] Pilar Sánchez,[^] Francisca Vicente,[^] Olga Genilloud,[^] Sonia Coriani,[§] René Wugt Larsen,[§] Jens C. Frisvad,[†] and Thomas O. Larsen[*†]

[†]Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800, Kgs. Lyngby, Denmark

[‡]Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, I-25123, Brescia, Italy

[§]Department of Chemistry, Technical University of Denmark, Kemitorvet, DK-2800, Kgs. Lyngby, Denmark

[^]Fundación MEDINA, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016, Armilla, Granada, Spain

Table of Contents

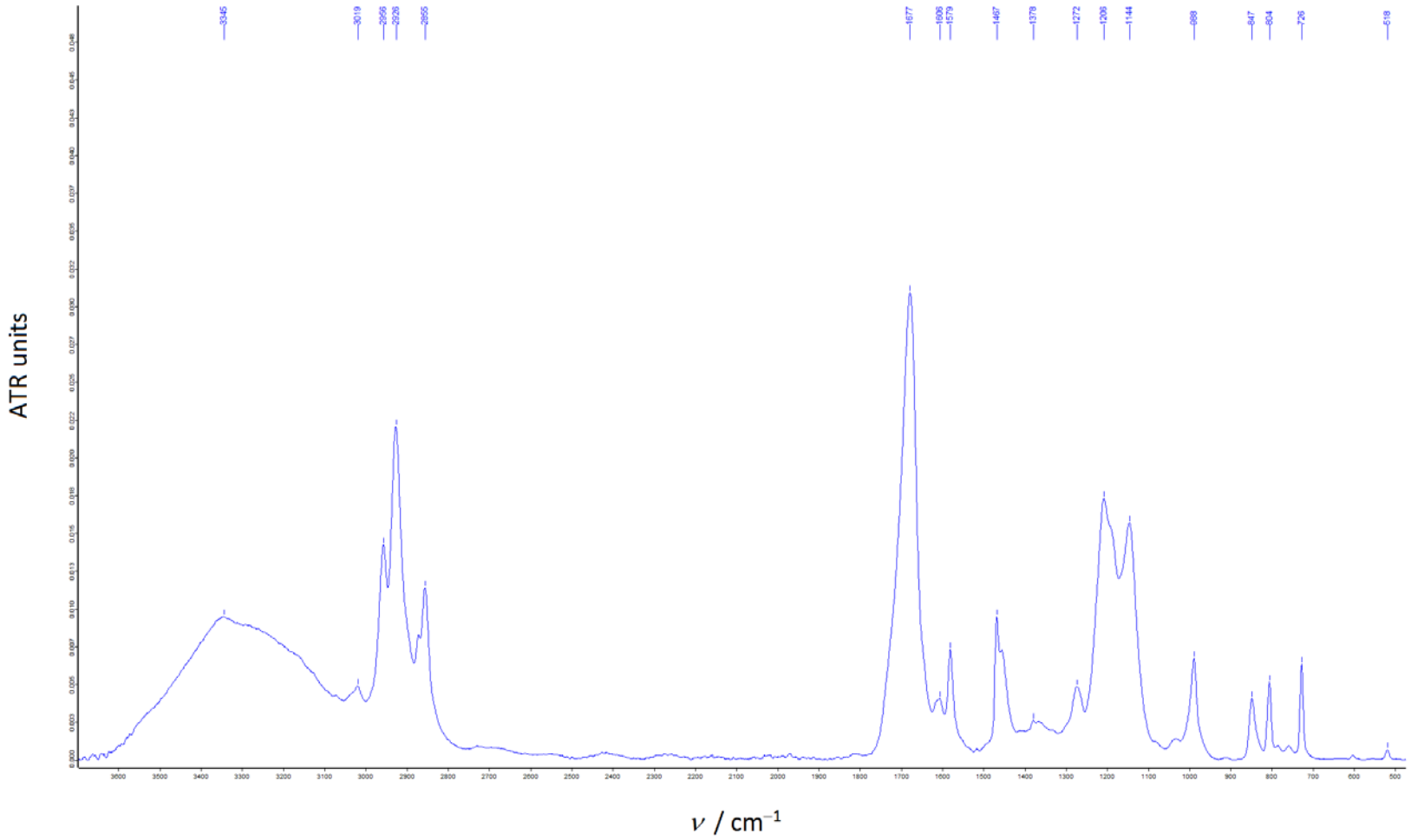


Figure S1. IR spectrum of **1**.

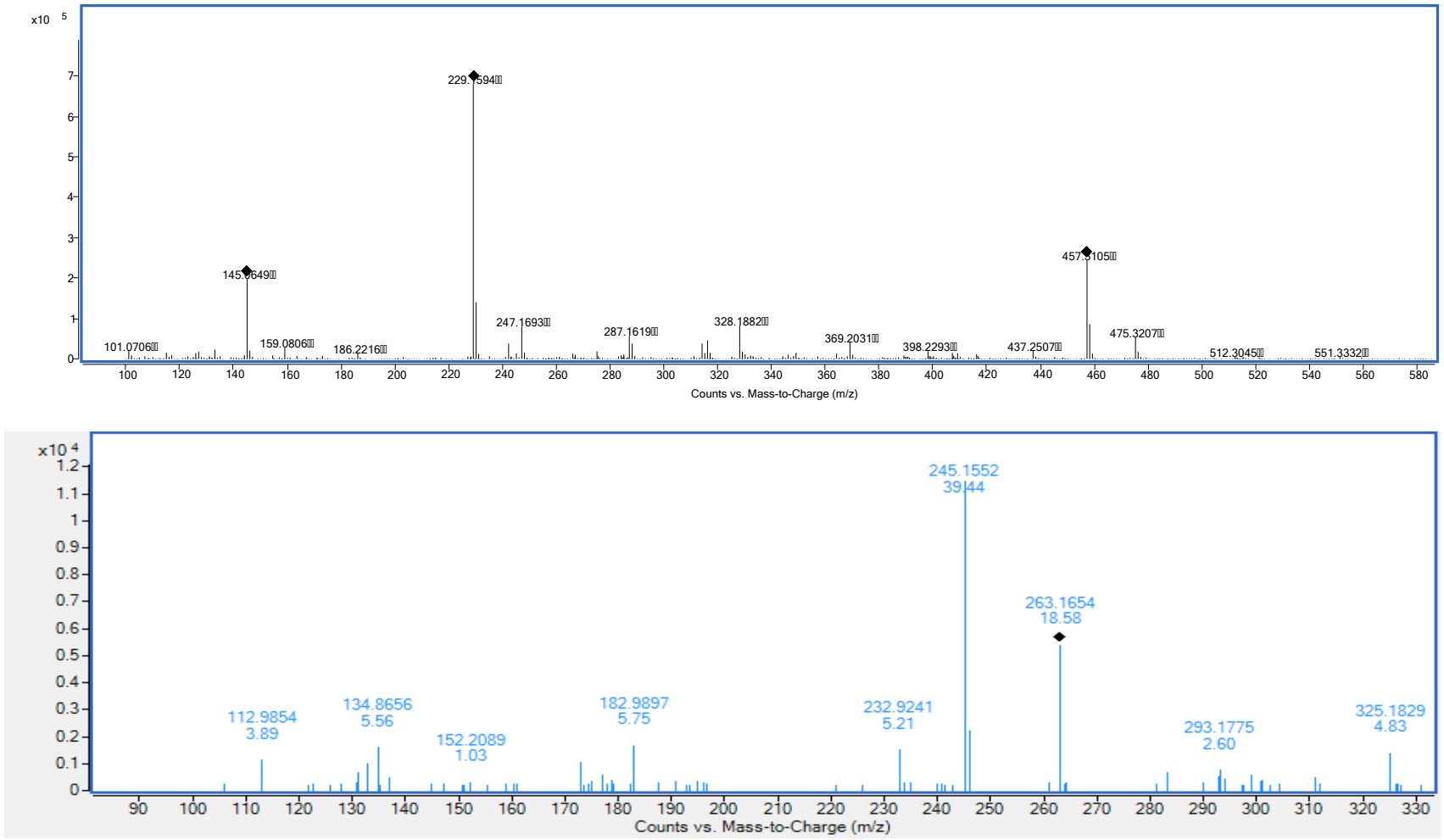


Figure S2. HRESIMS (ESI^+ top, ESI^- bottom) spectrum of **1**.

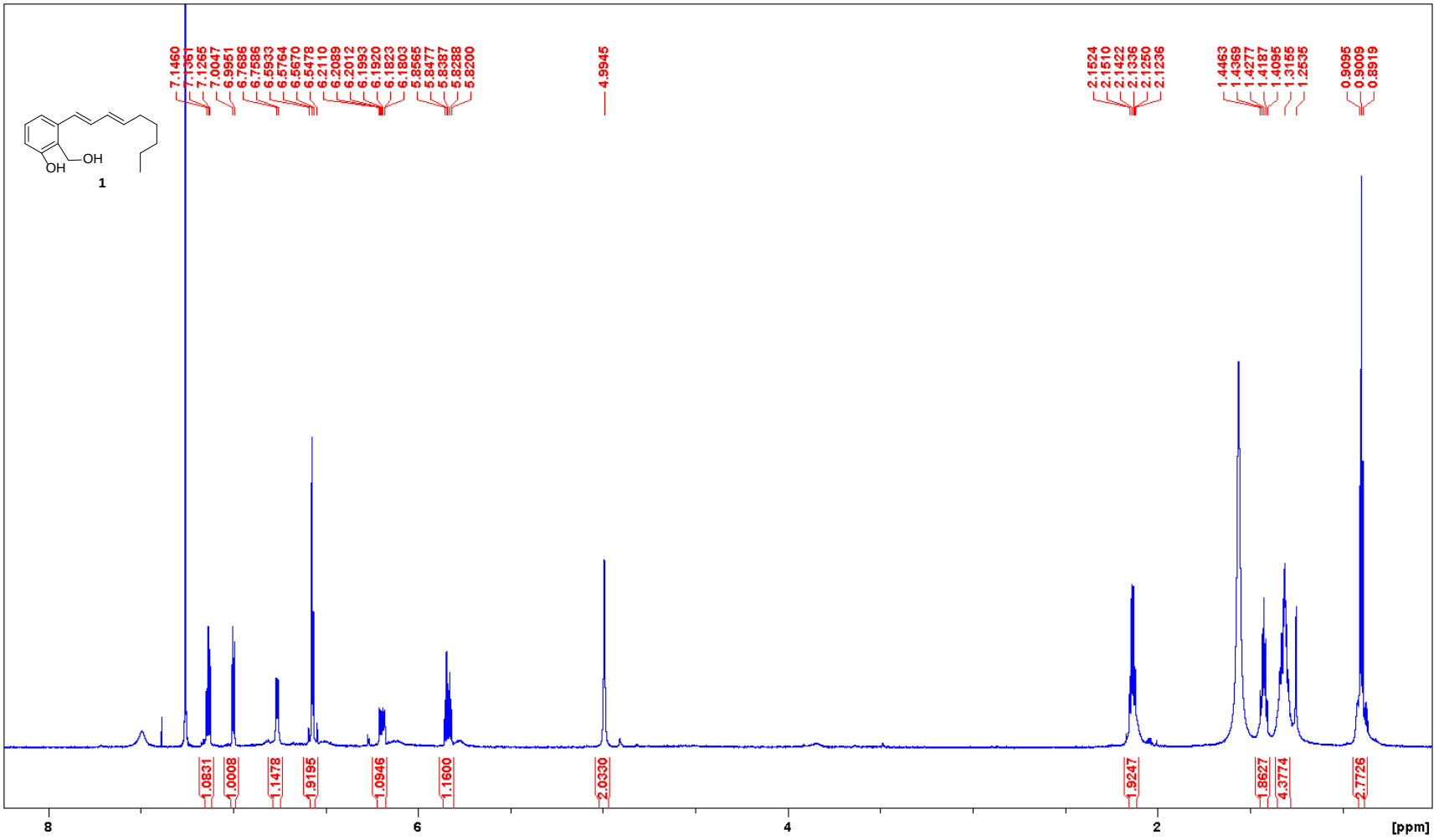


Figure S3. ^{1}H NMR spectrum of **1** (800 MHz, $CDCl_3$).

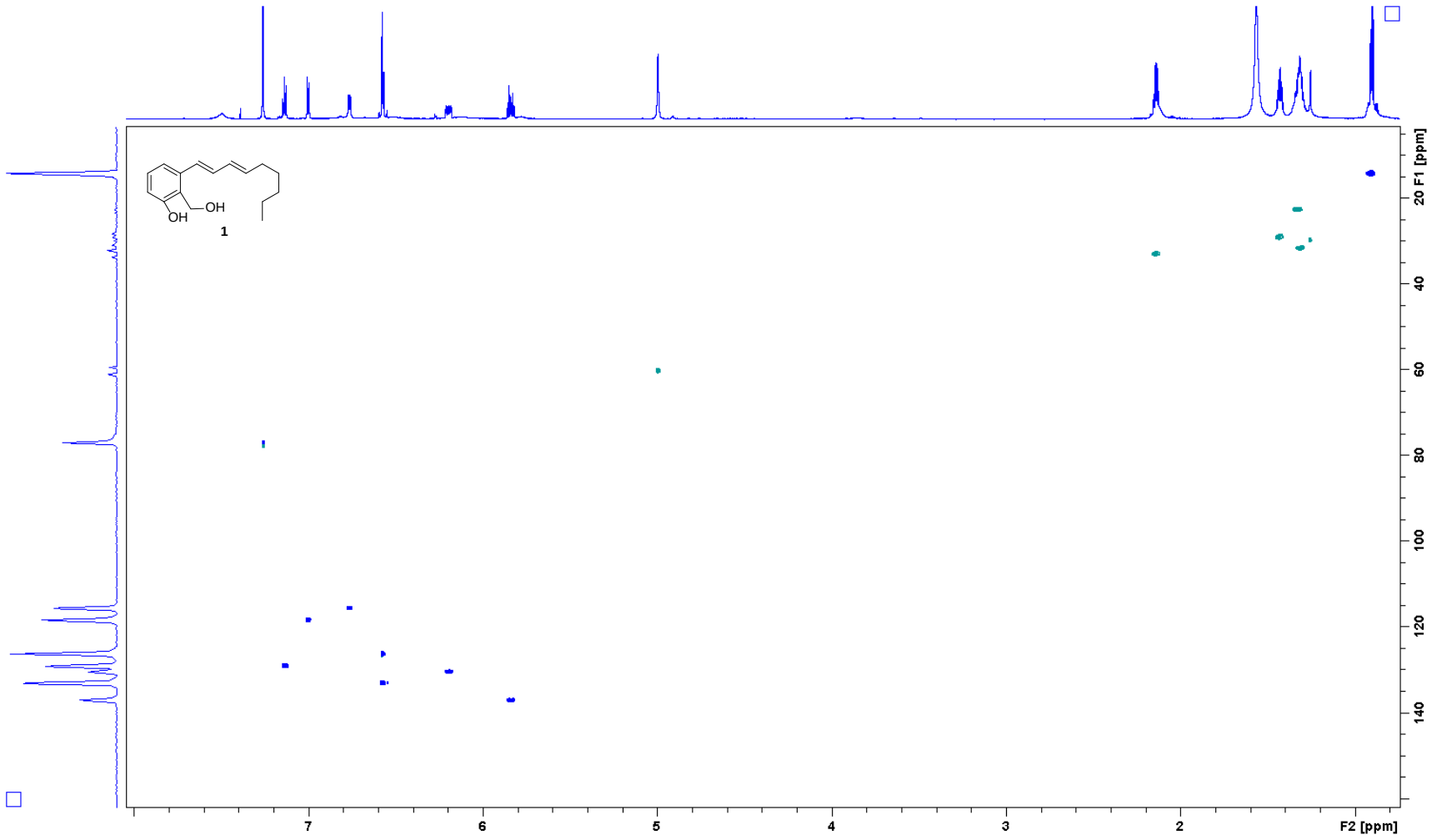


Figure S4. Edited HSQC spectrum of **1**.

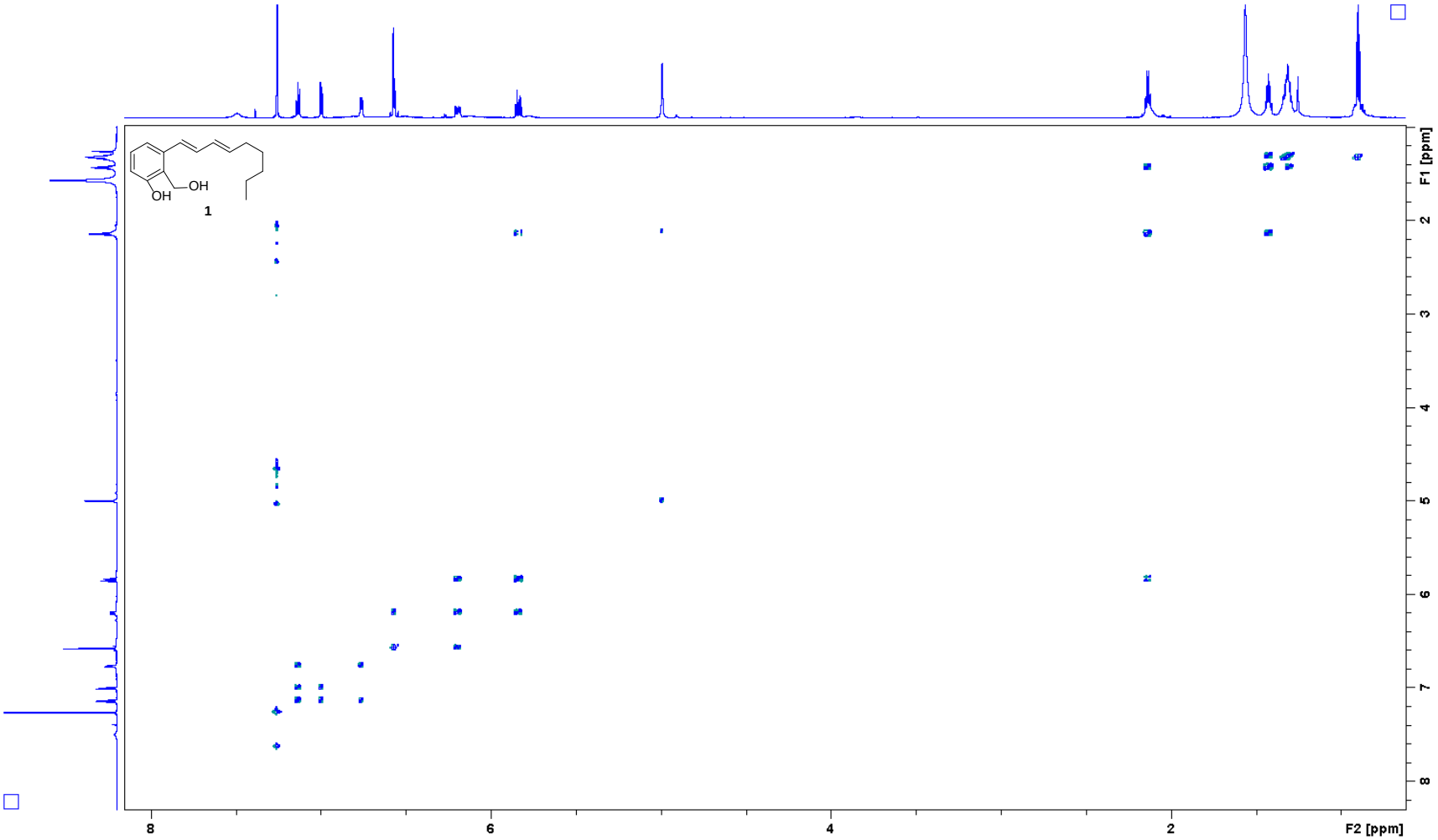


Figure S5. H2BC spectrum of **1**.

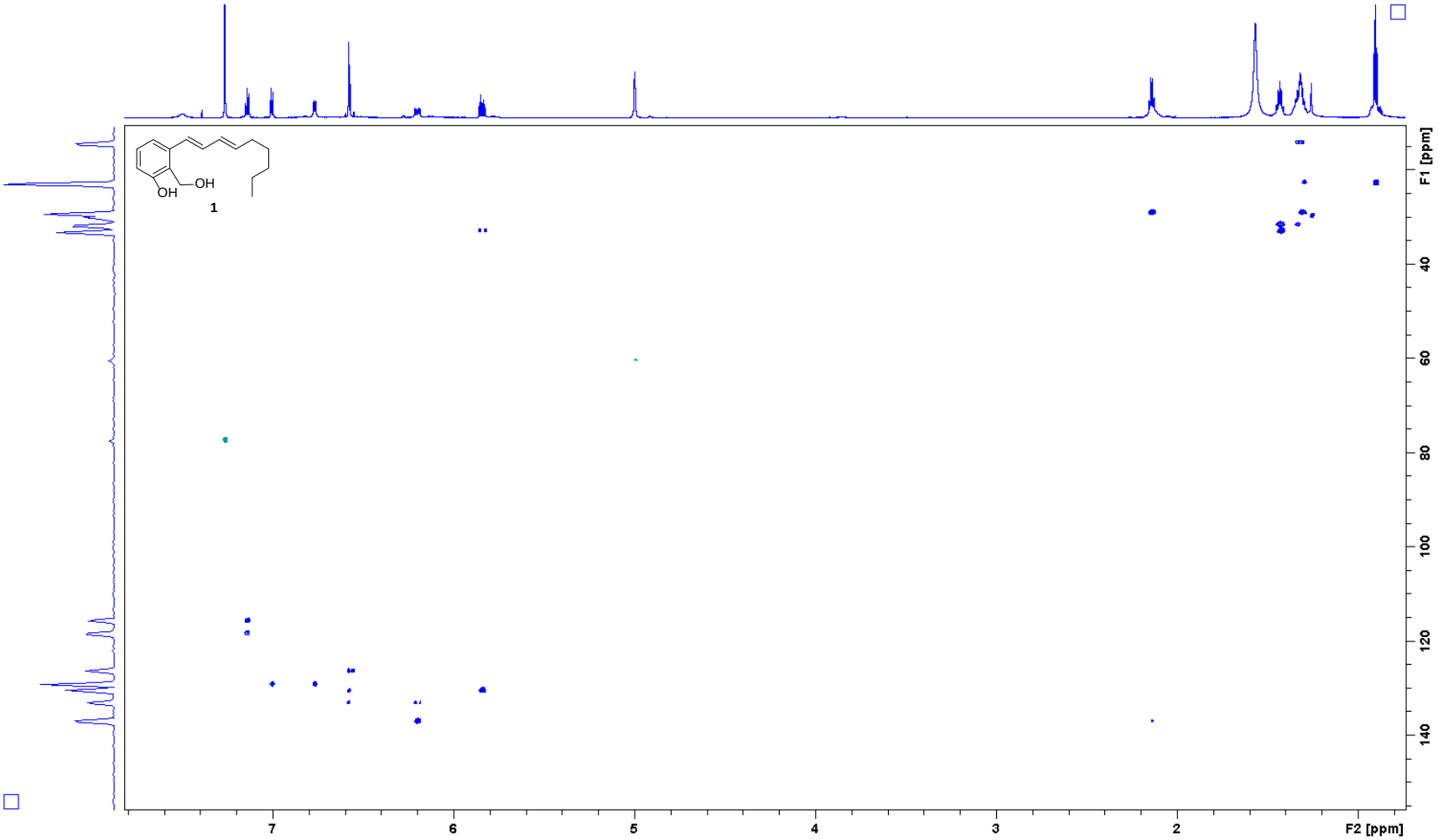


Figure S6. H2BC spectrum of **1**.

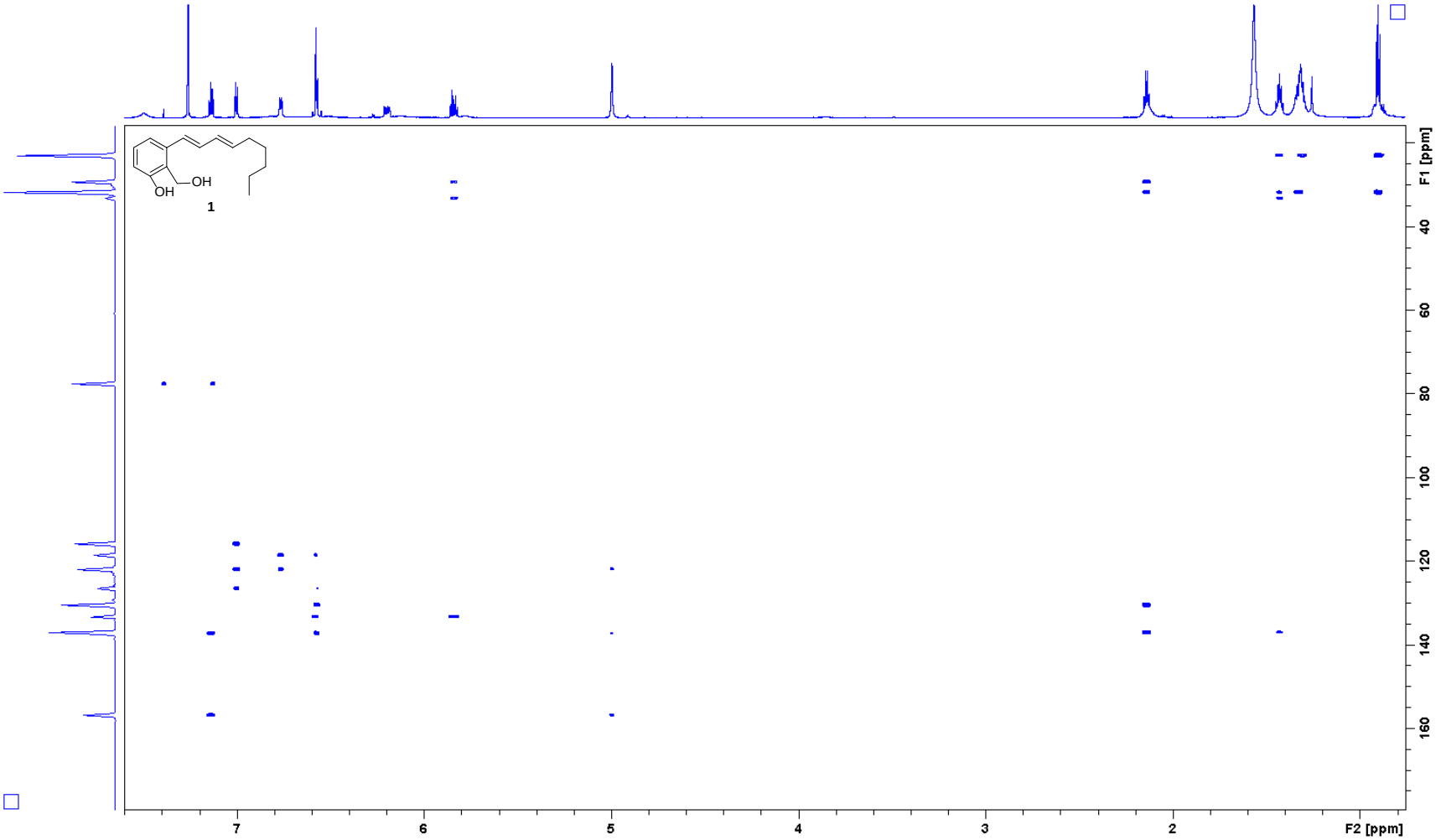


Figure S7. HMBC spectrum of **1**.

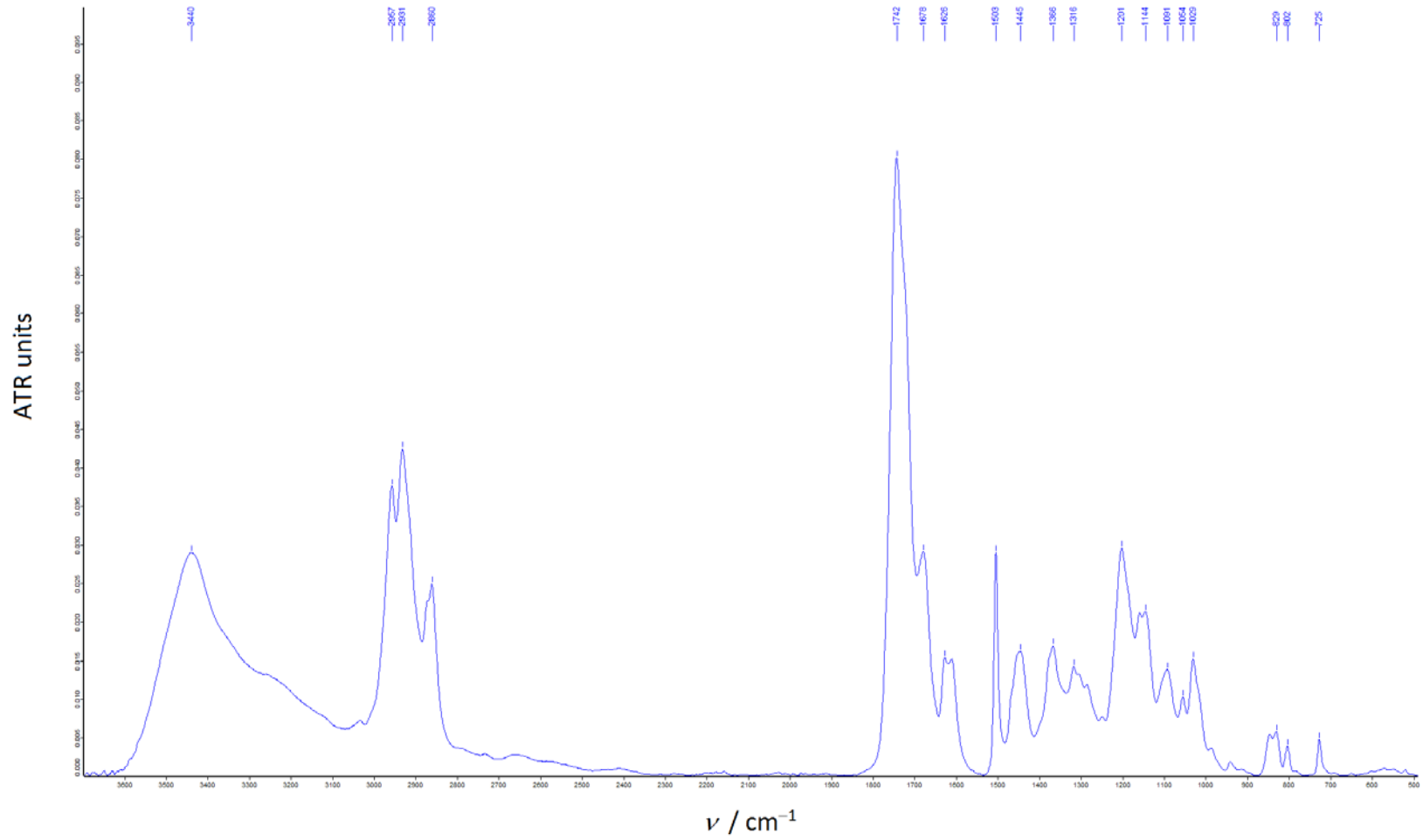


Figure S8. IR spectrum of **2**.

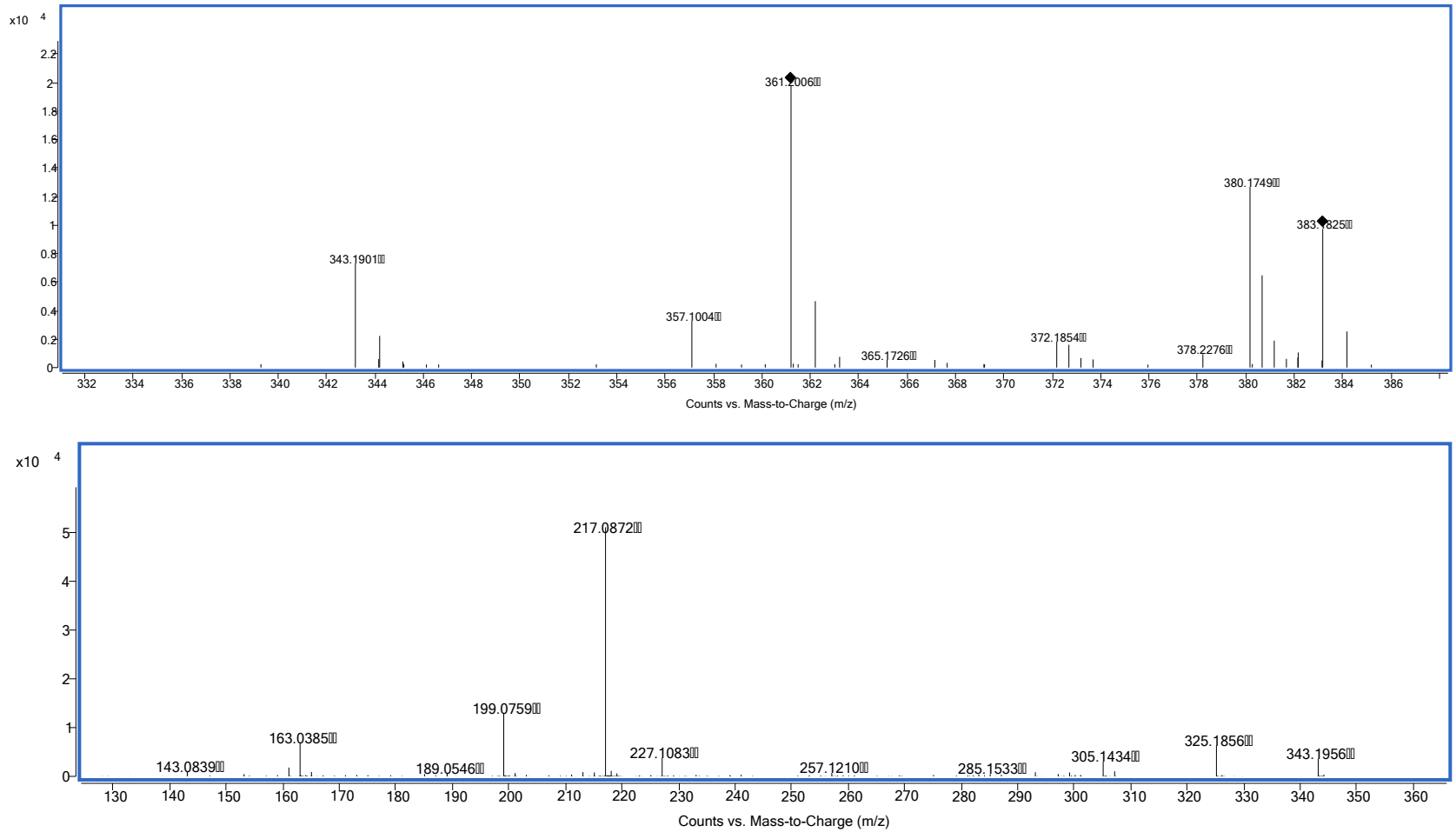


Figure S9. HRMS (ESI^+ top) and MSMS (CID 10 eV bottom) spectra of **2**.

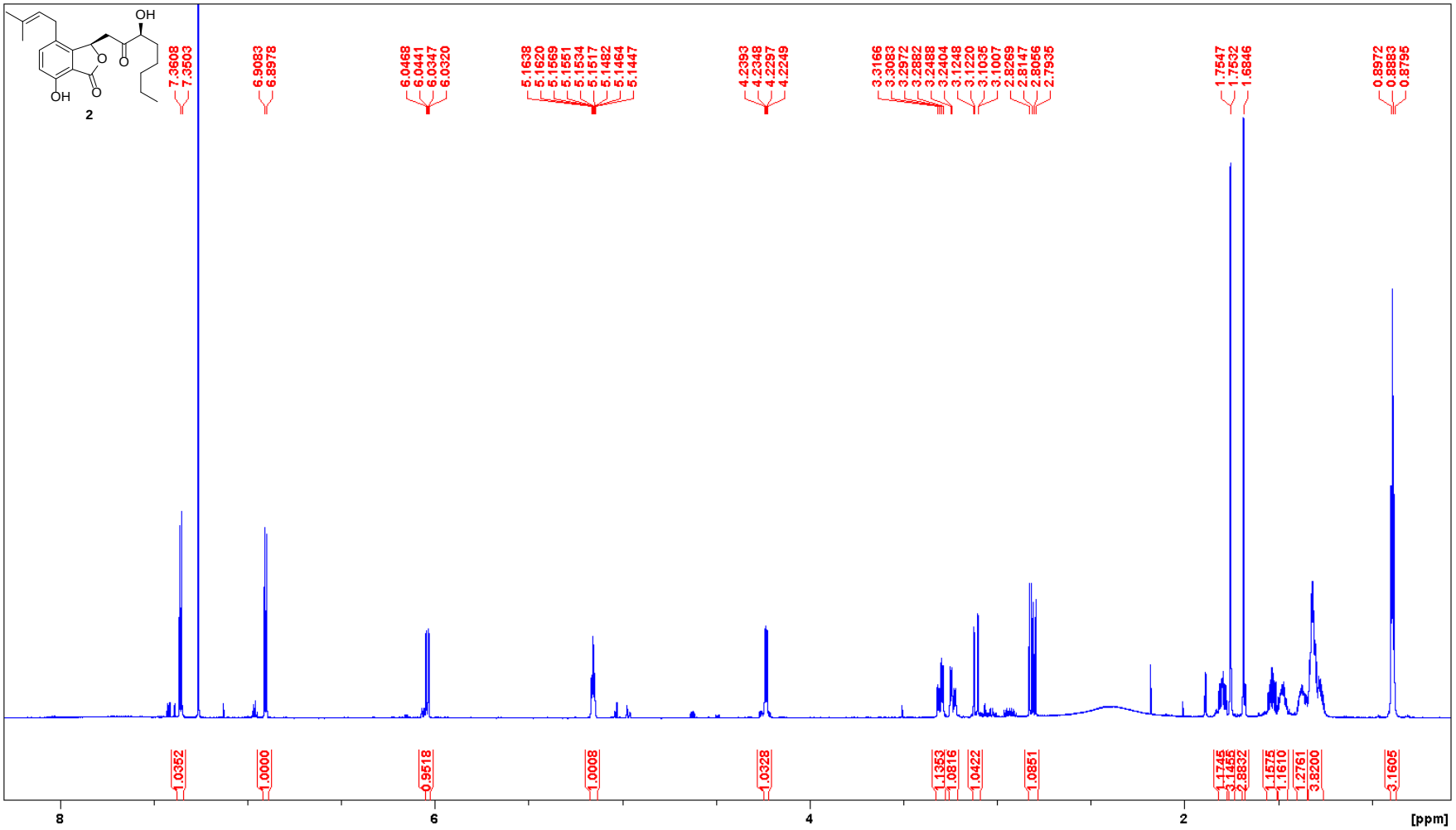


Figure S10. ^{1}H NMR spectrum of **2** (800 MHz, $CDCl_3$).

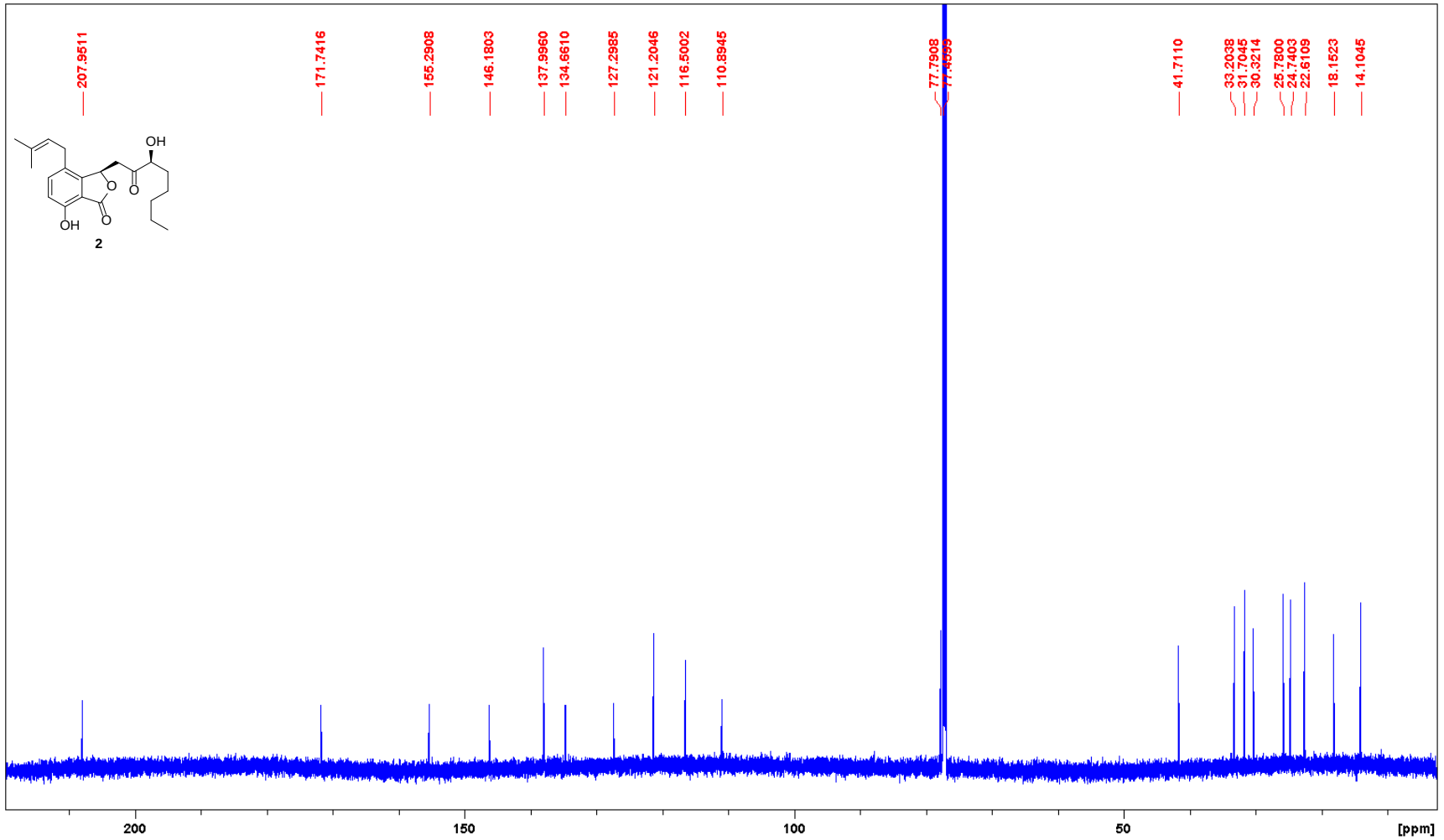


Figure S11. ^{13}C NMR spectrum of **2** (200 MHz, $CDCl_3$).

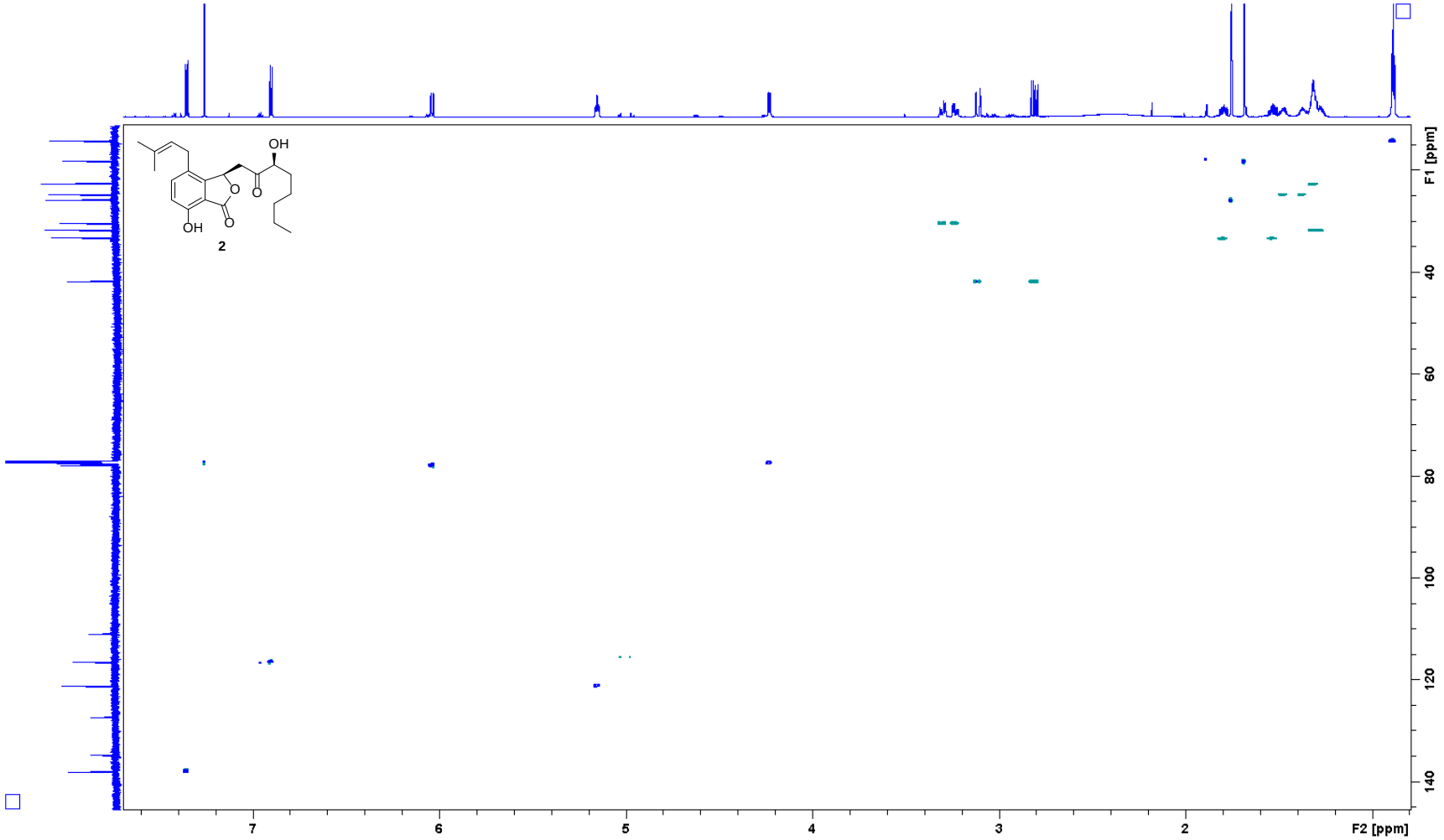


Figure S12. HSQC spectrum of **2**.

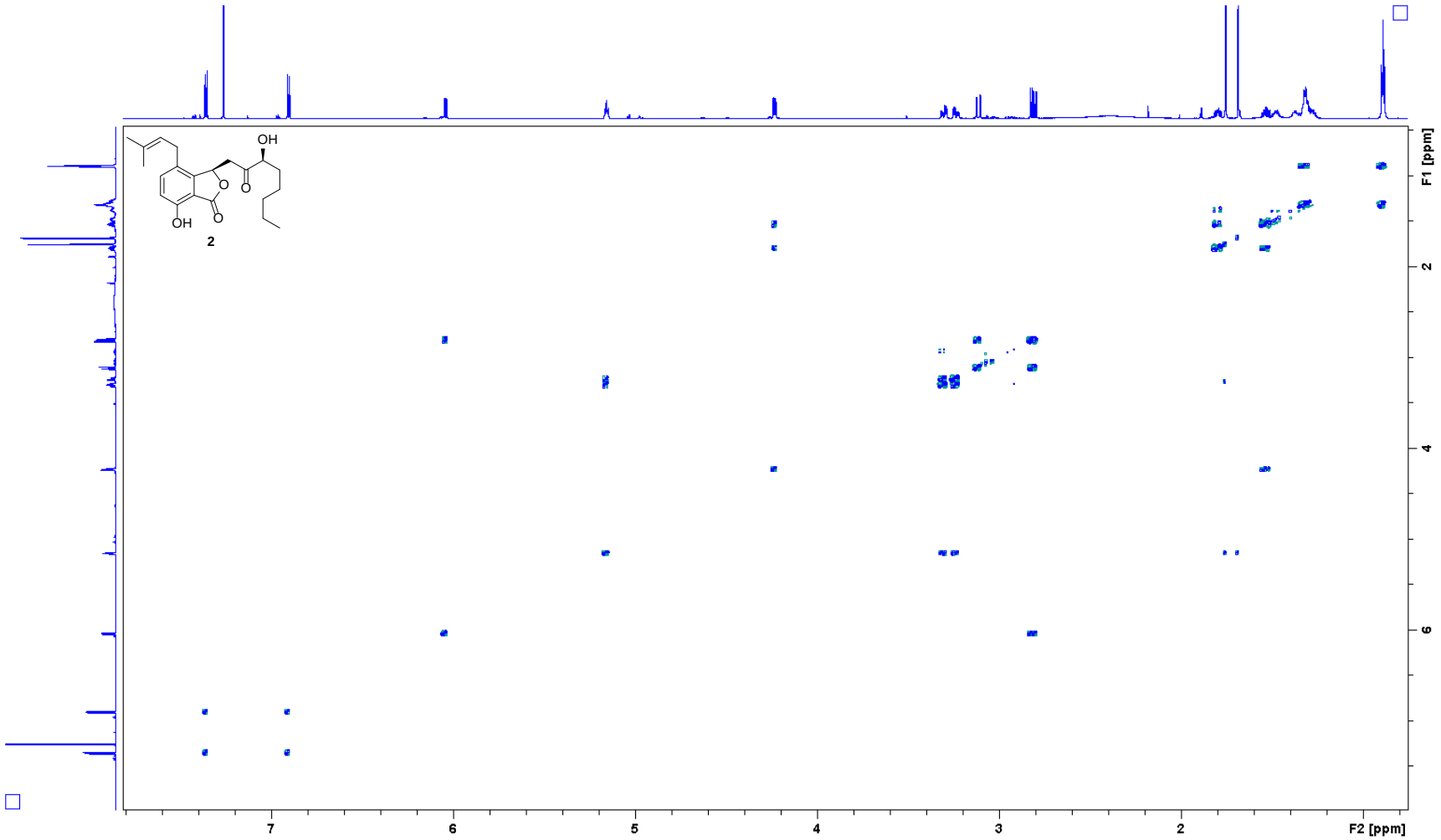


Figure S13. COSY spectrum of **2**.

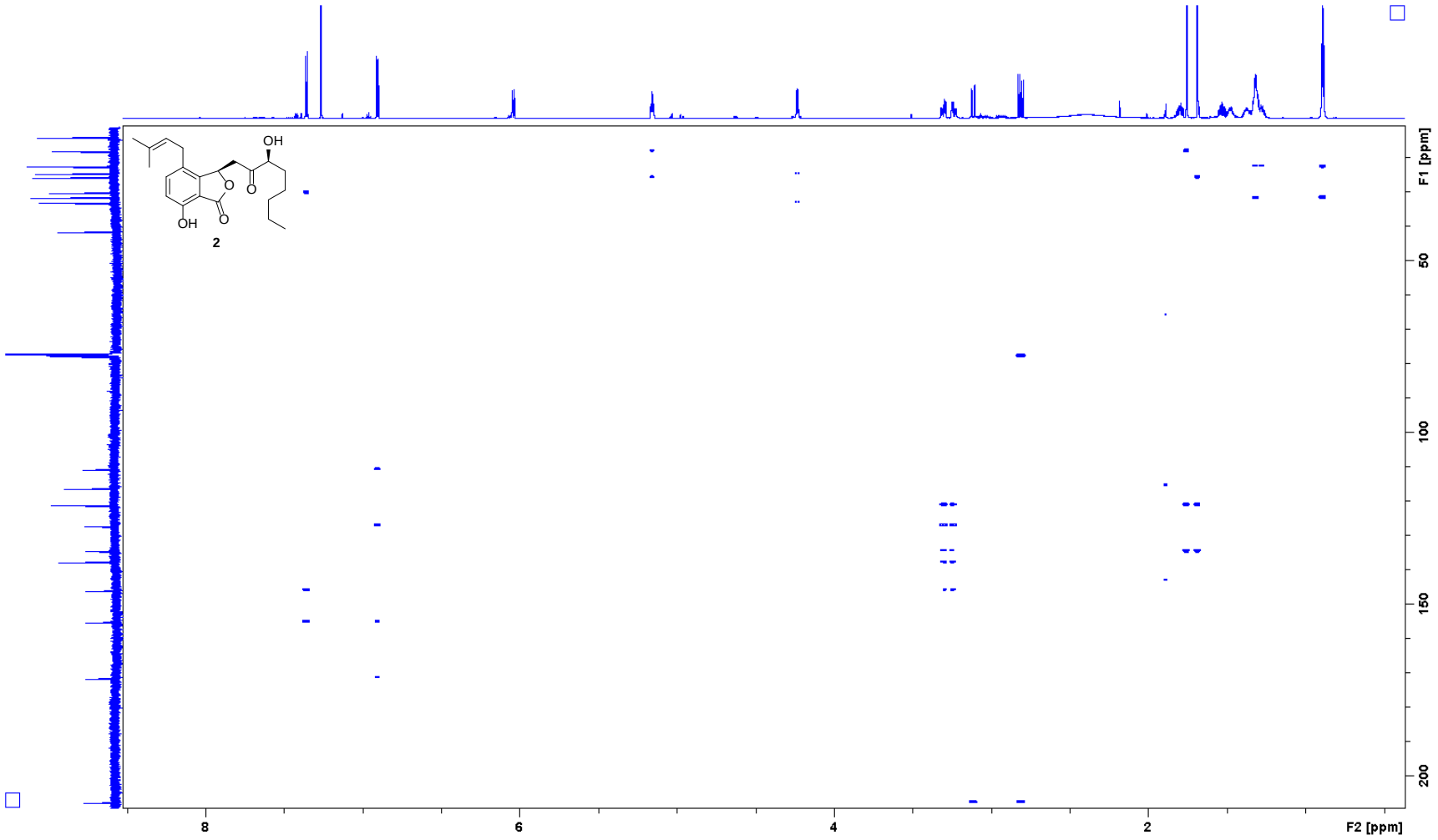


Figure S14. HMBC spectrum of **2**.

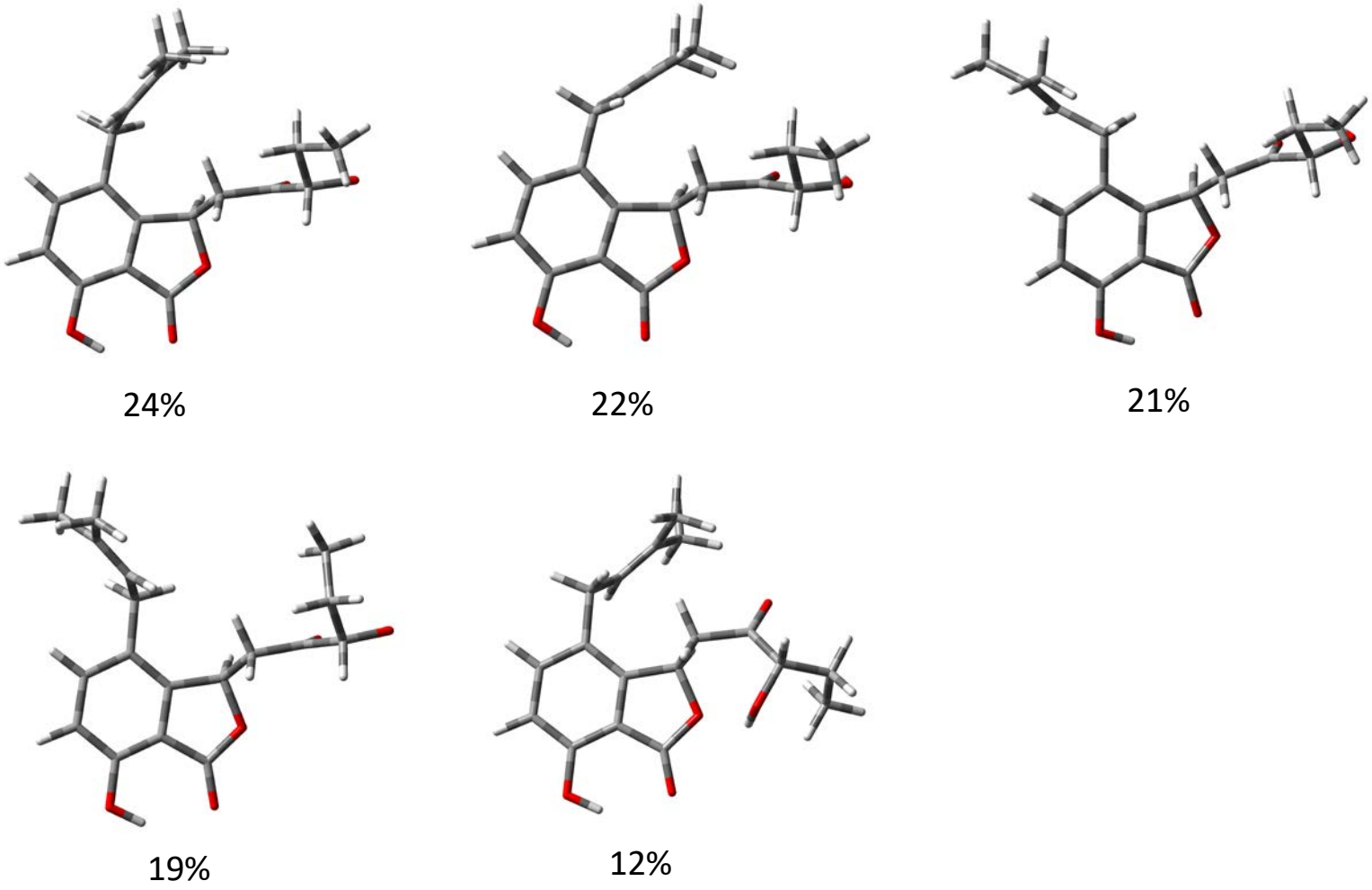


Figure S15. Structures and relative populations of the predominant conformers of (6*S*, 9*R*)-**2**.

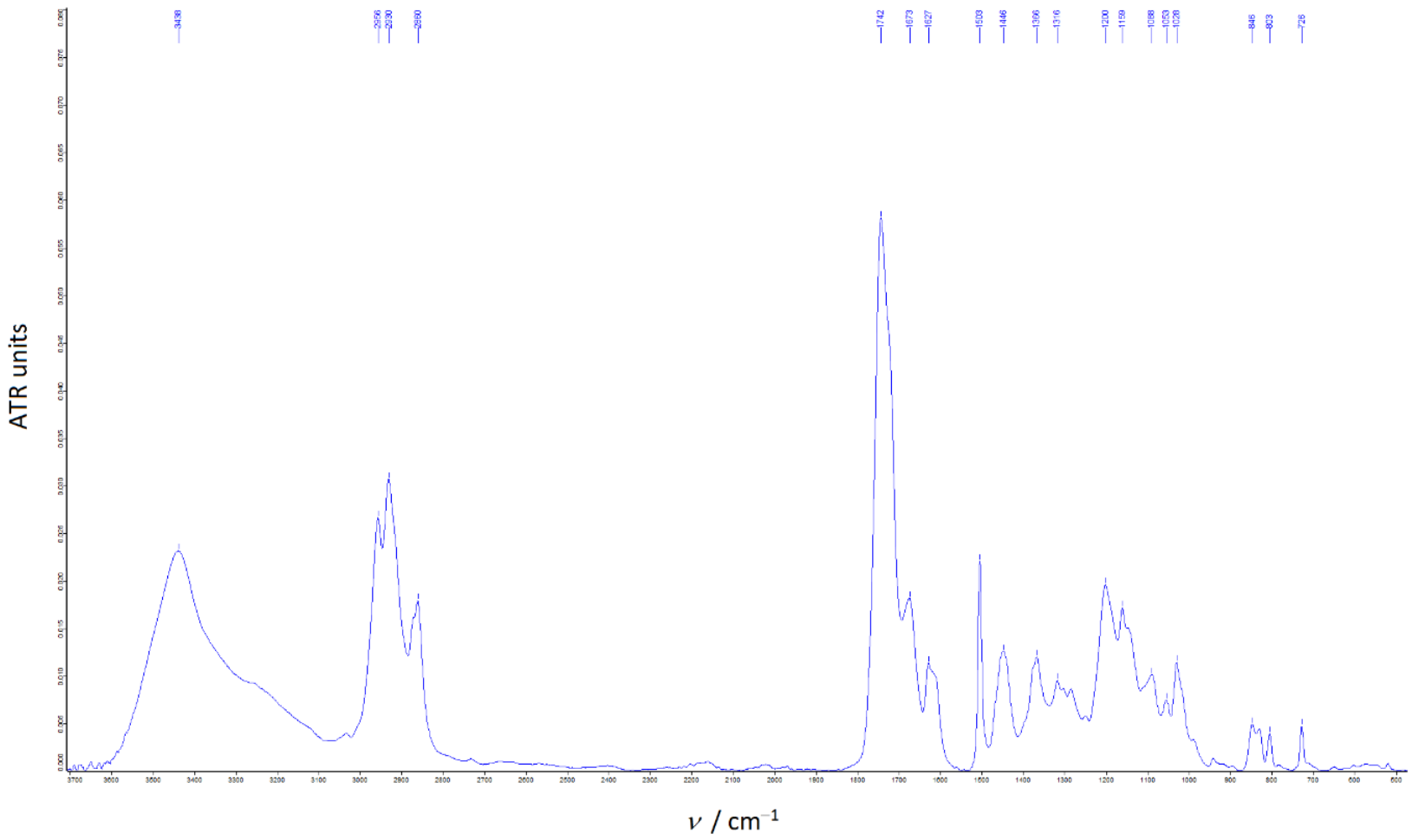


Figure S16. IR spectrum of **3**.

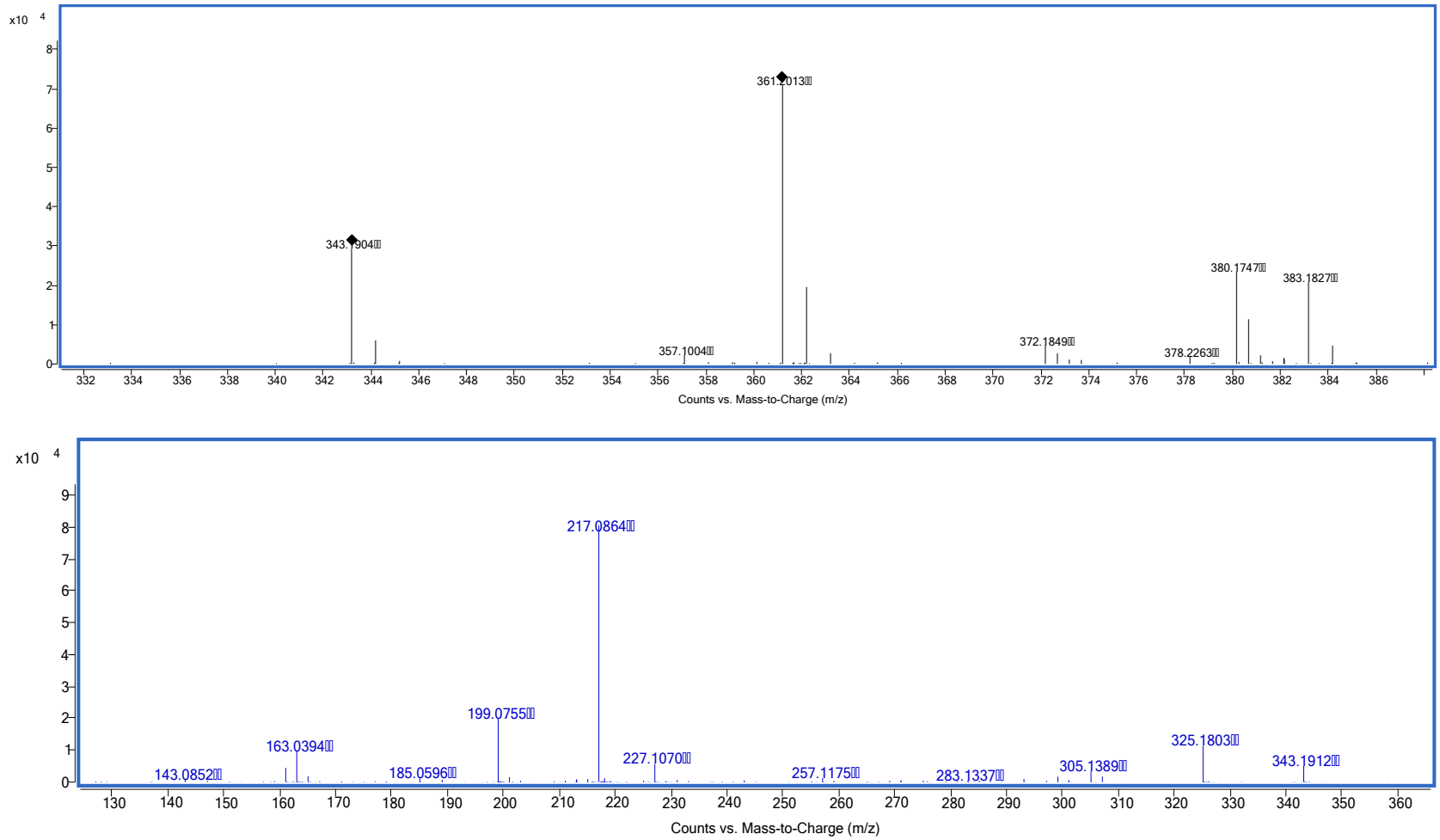


Figure S17. HRMS (ESI$^+$ top) and MSMS (CID 10 eV bottom) spectra of **3**.

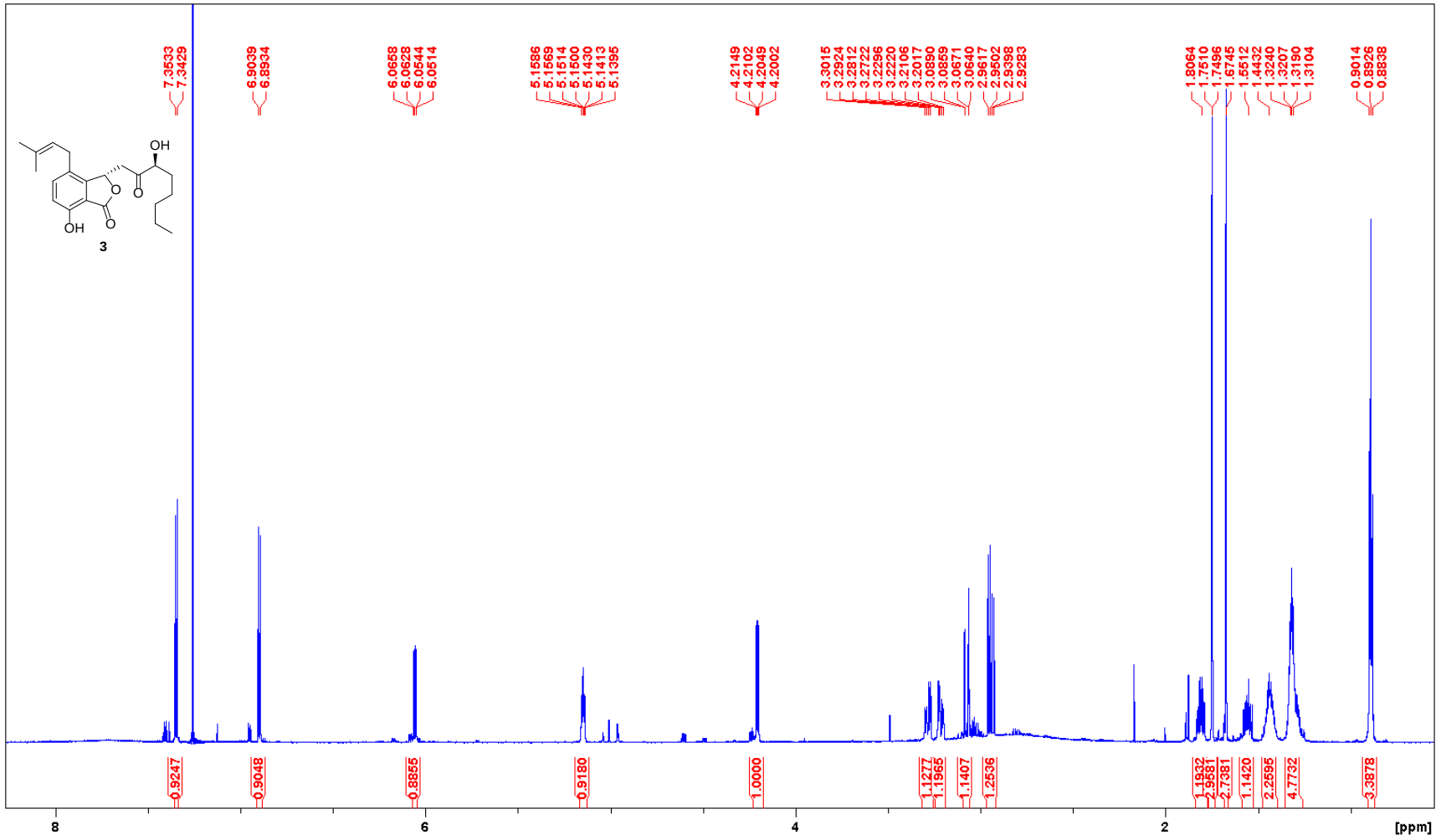


Figure S18. ^{1}H NMR spectrum of **3** (800 MHz, $CDCl_3$).

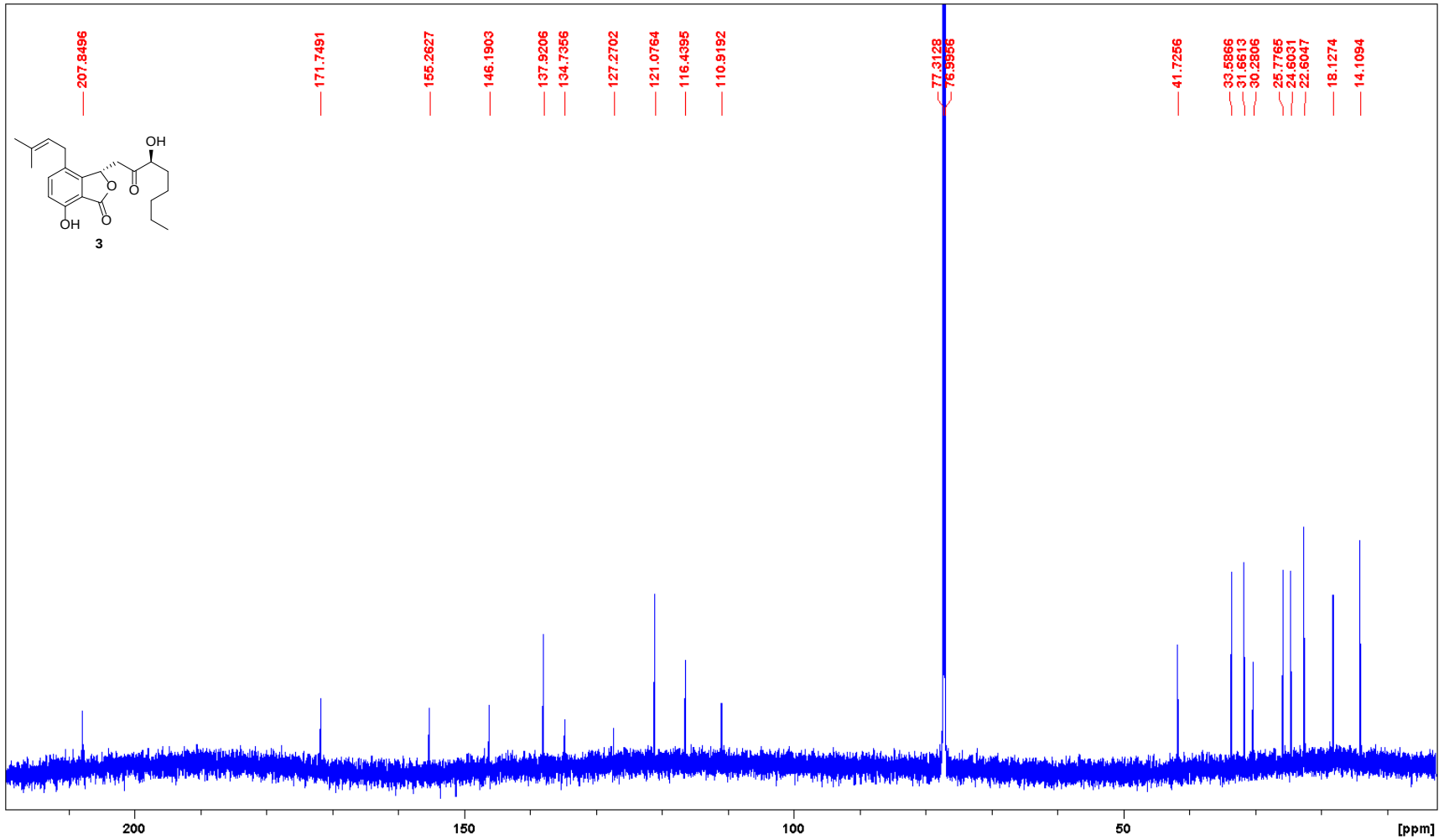


Figure S19. ^{13}C NMR spectrum of **3** (200 MHz, $CDCl_3$).

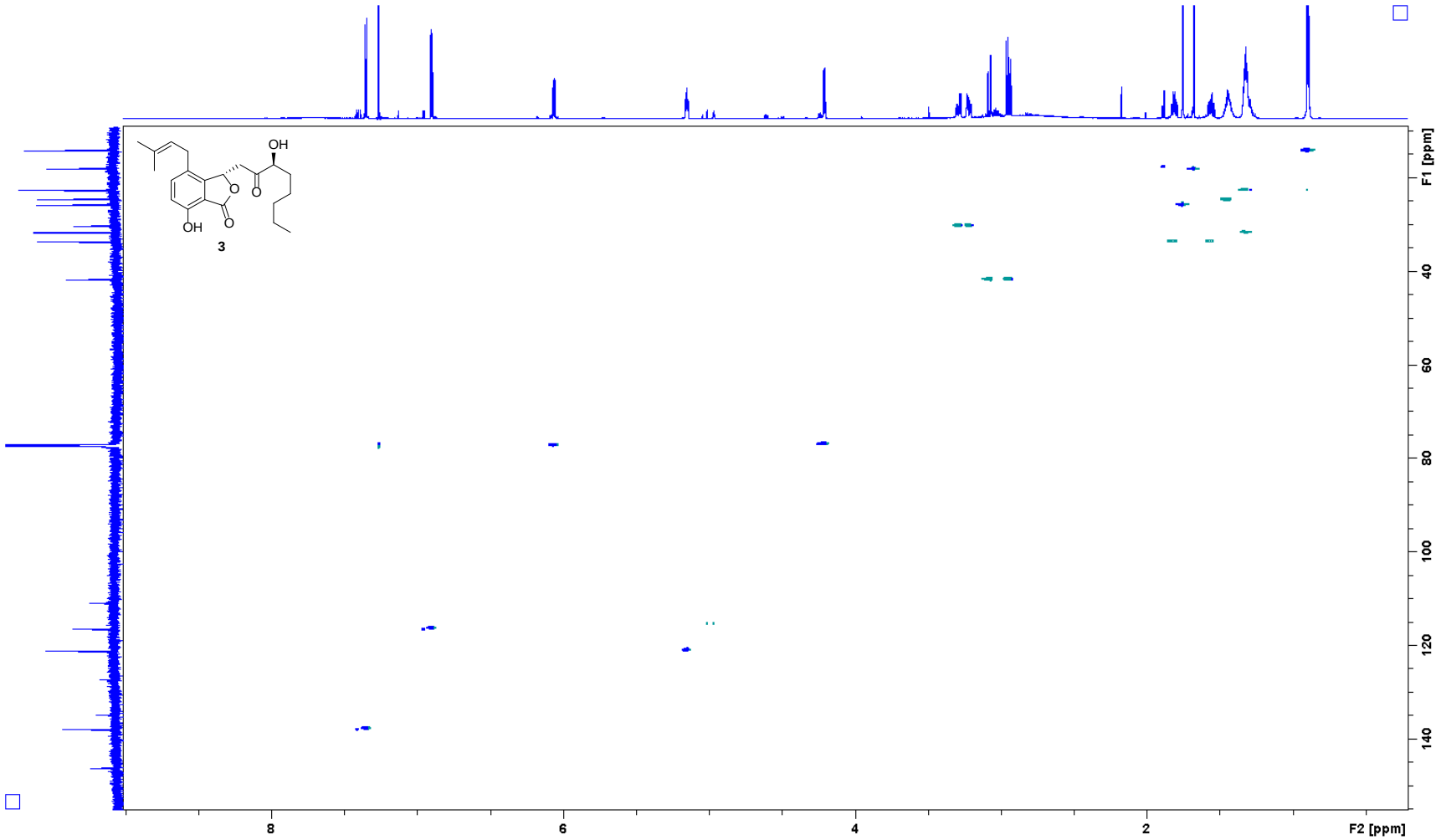


Figure S20. HSQC spectrum of **3**.

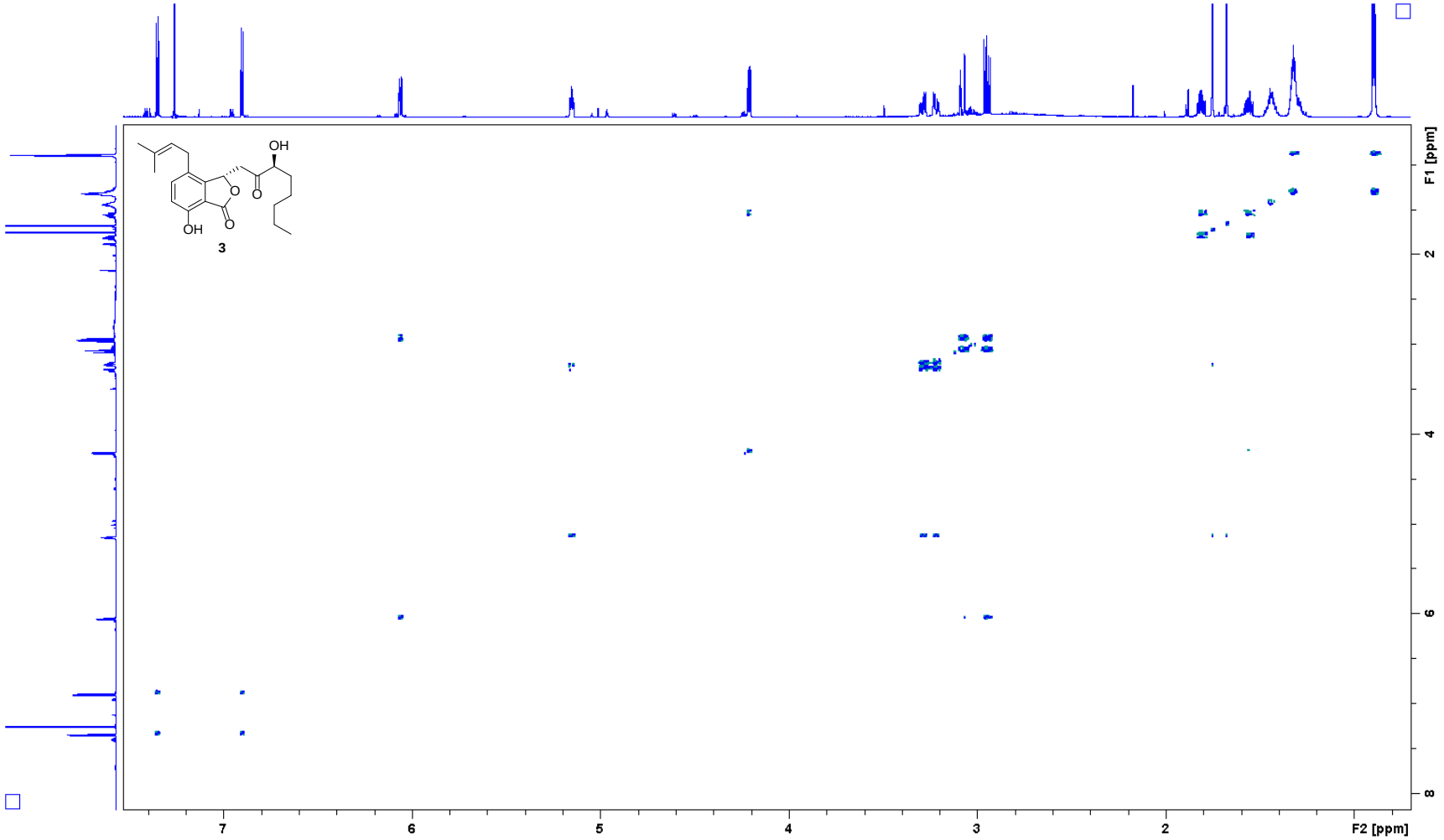


Figure S21. COSY spectrum of **3**.

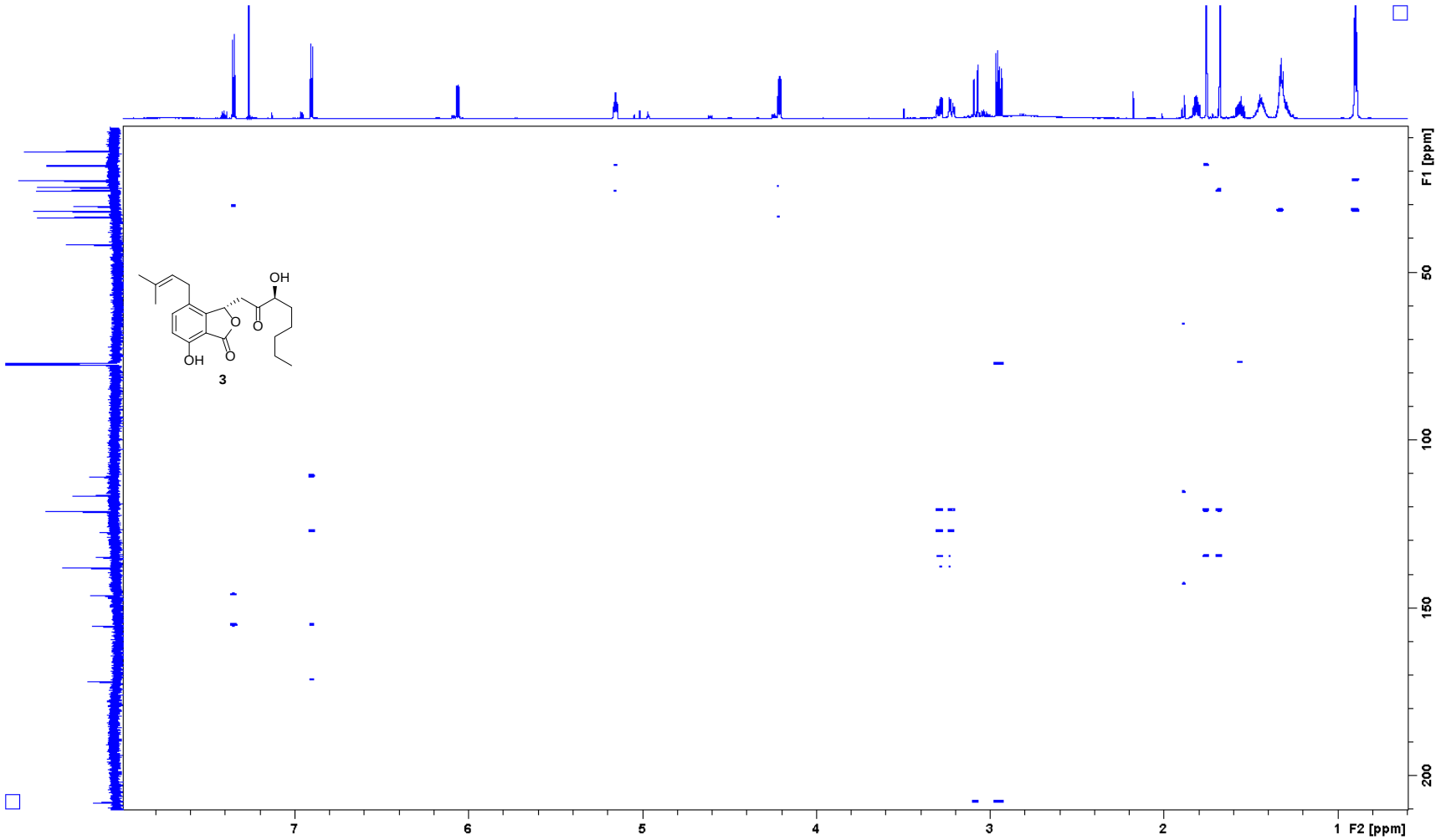


Figure S22. HMBC spectrum of **3**.

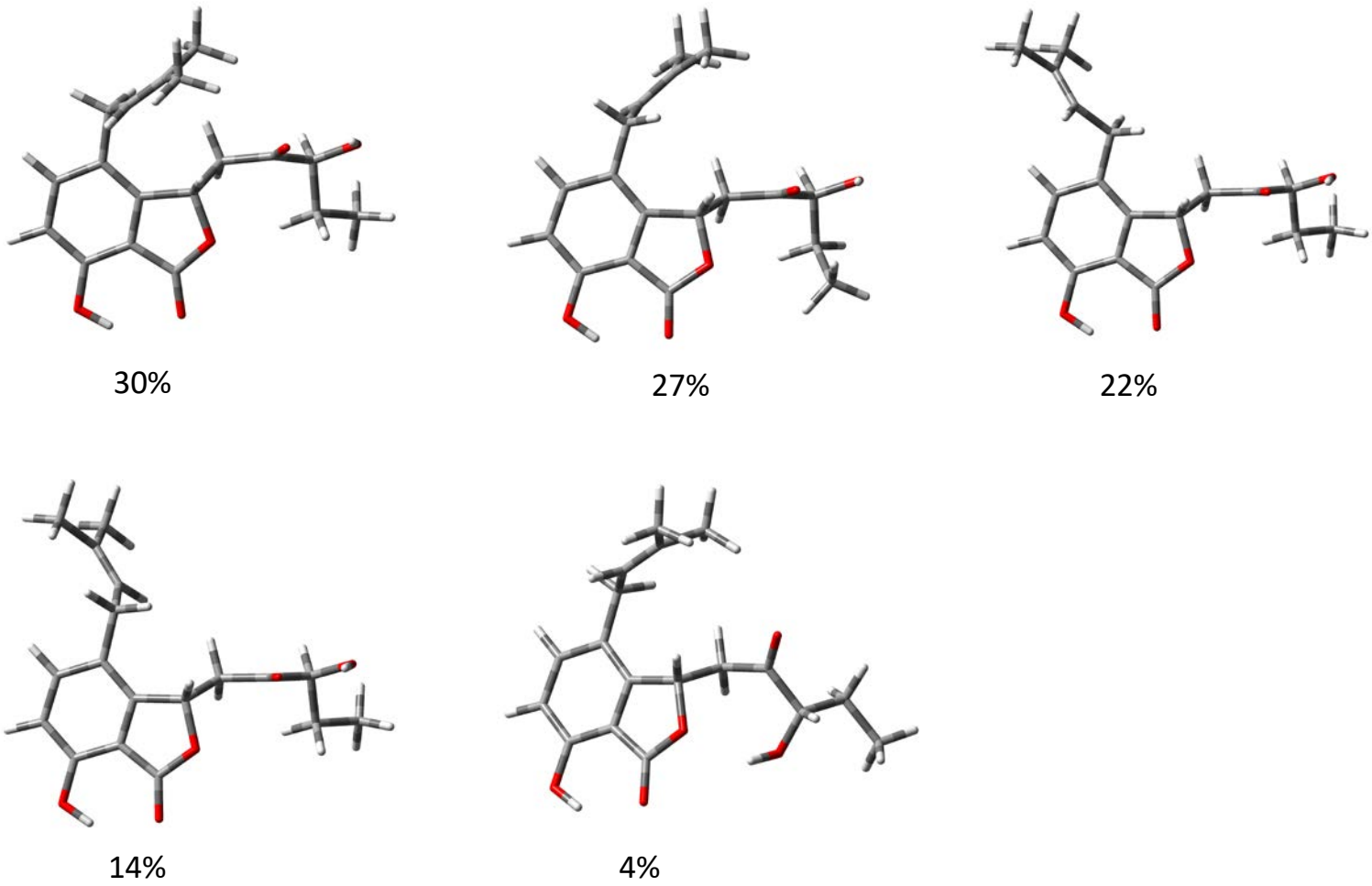


Figure S23. Structures and relative populations of the predominant conformers of (6*S*, 9*S*)-**3**.

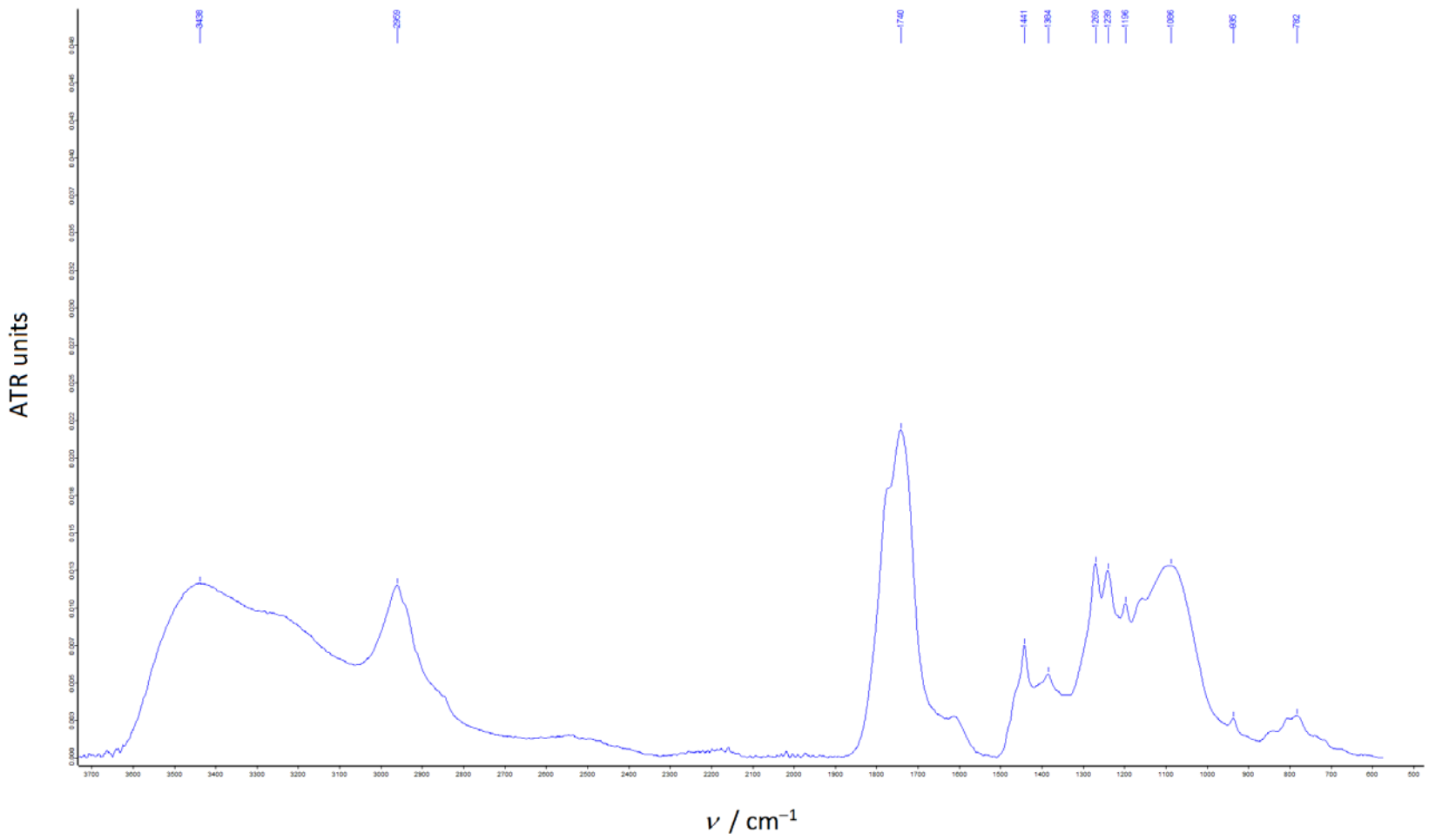


Figure S24. IR spectrum of **4**.

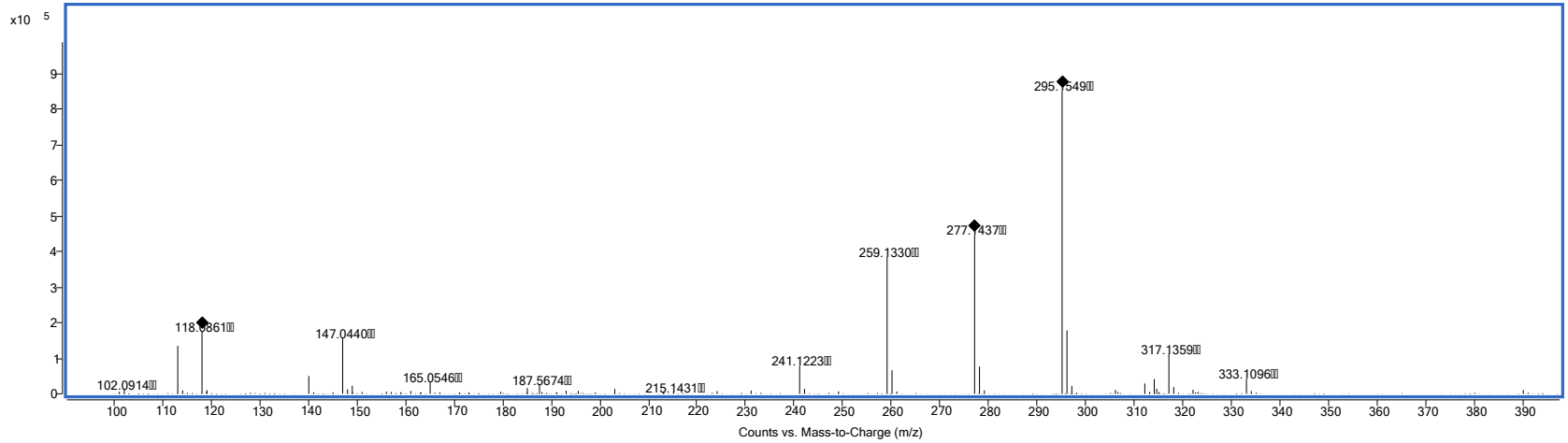


Figure S25. HRESIMS spectrum of **4**.

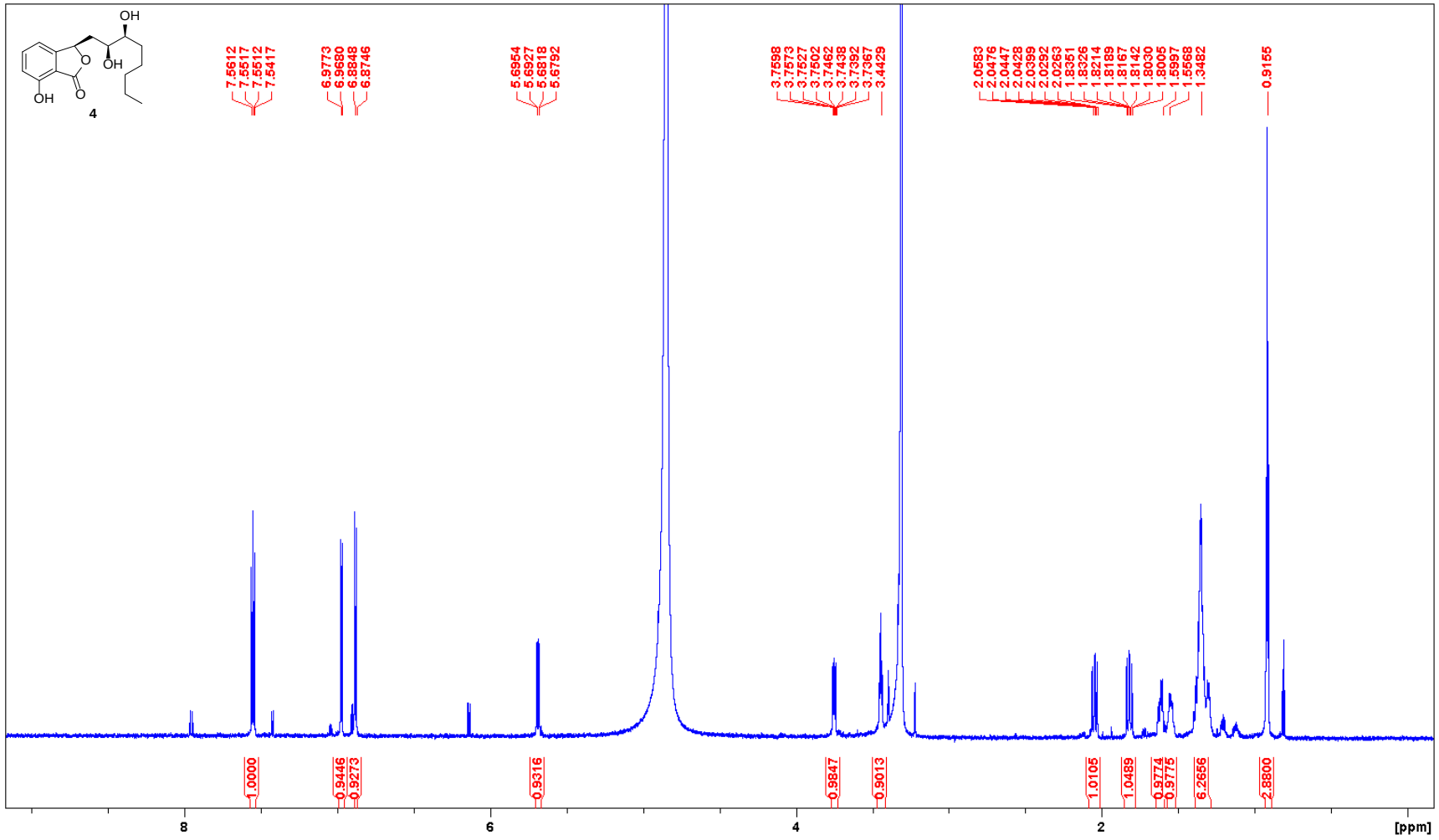


Figure S26. ^{1}H NMR spectrum of **4** (800 MHz, CD_3OD).

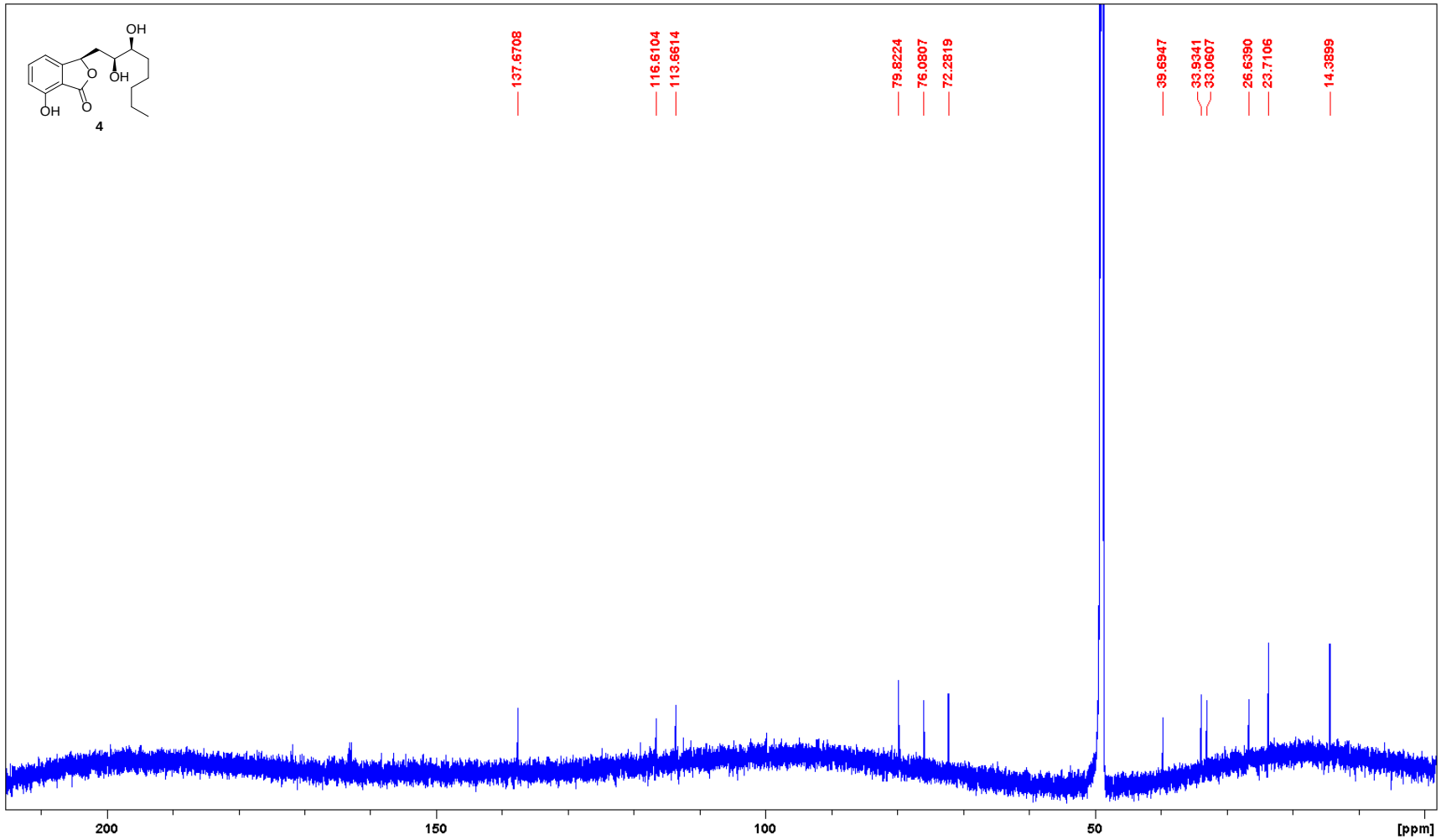


Figure S27. ^{13}C NMR spectrum of **4** (200 MHz, CD_3OD)

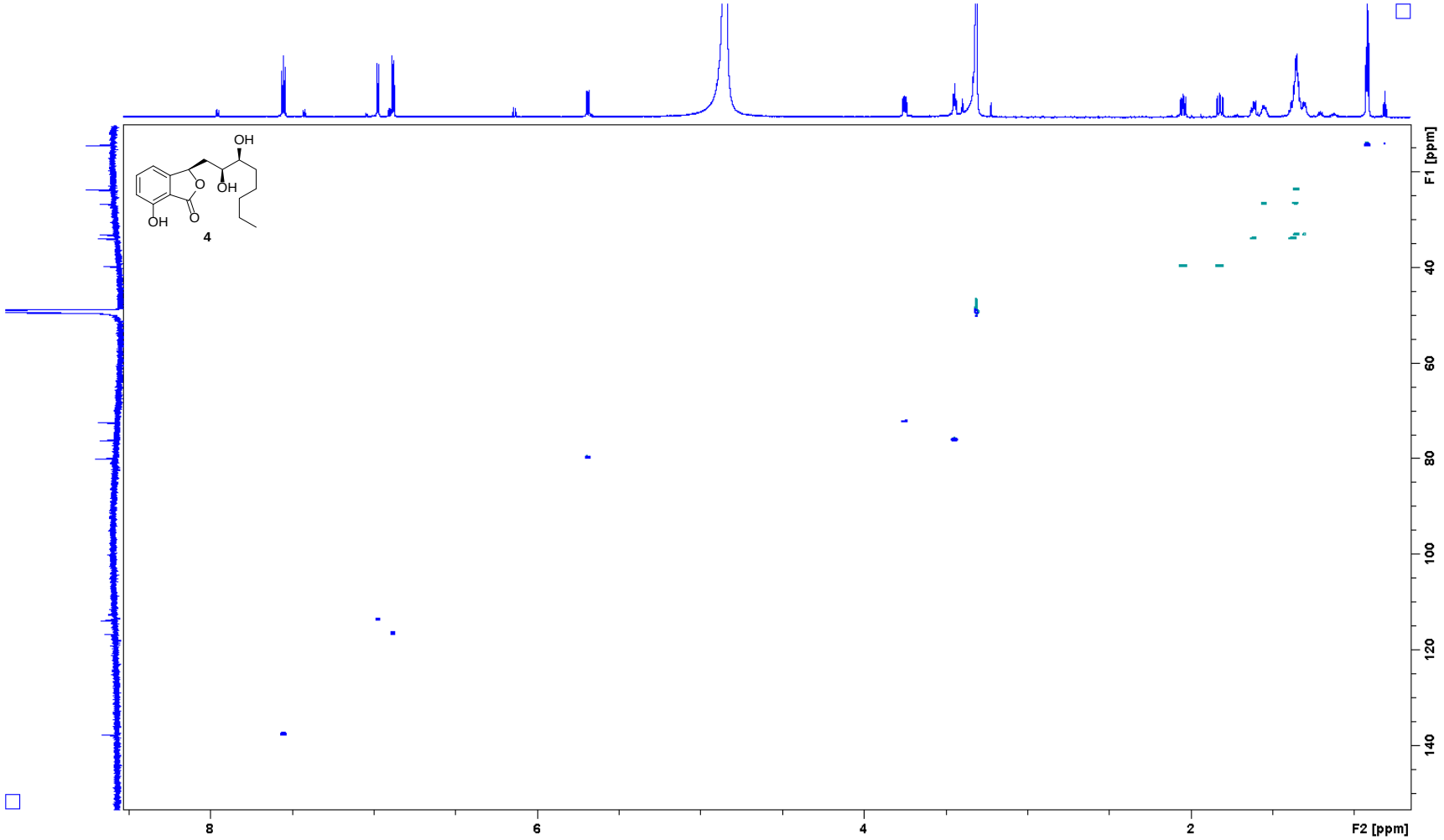


Figure S28. HSQC spectrum of **4**.

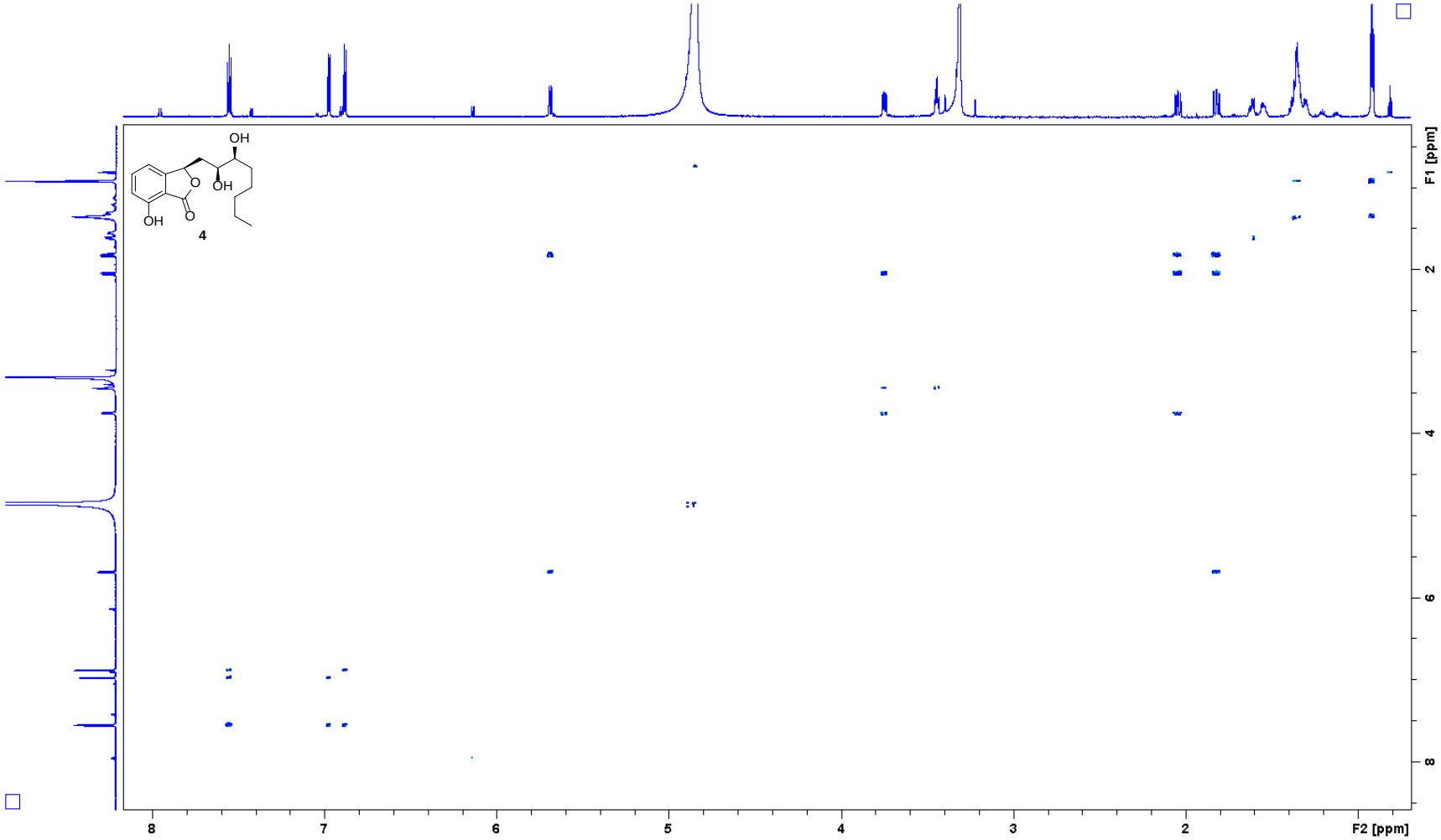


Figure S29. COSY spectrum of **4**.

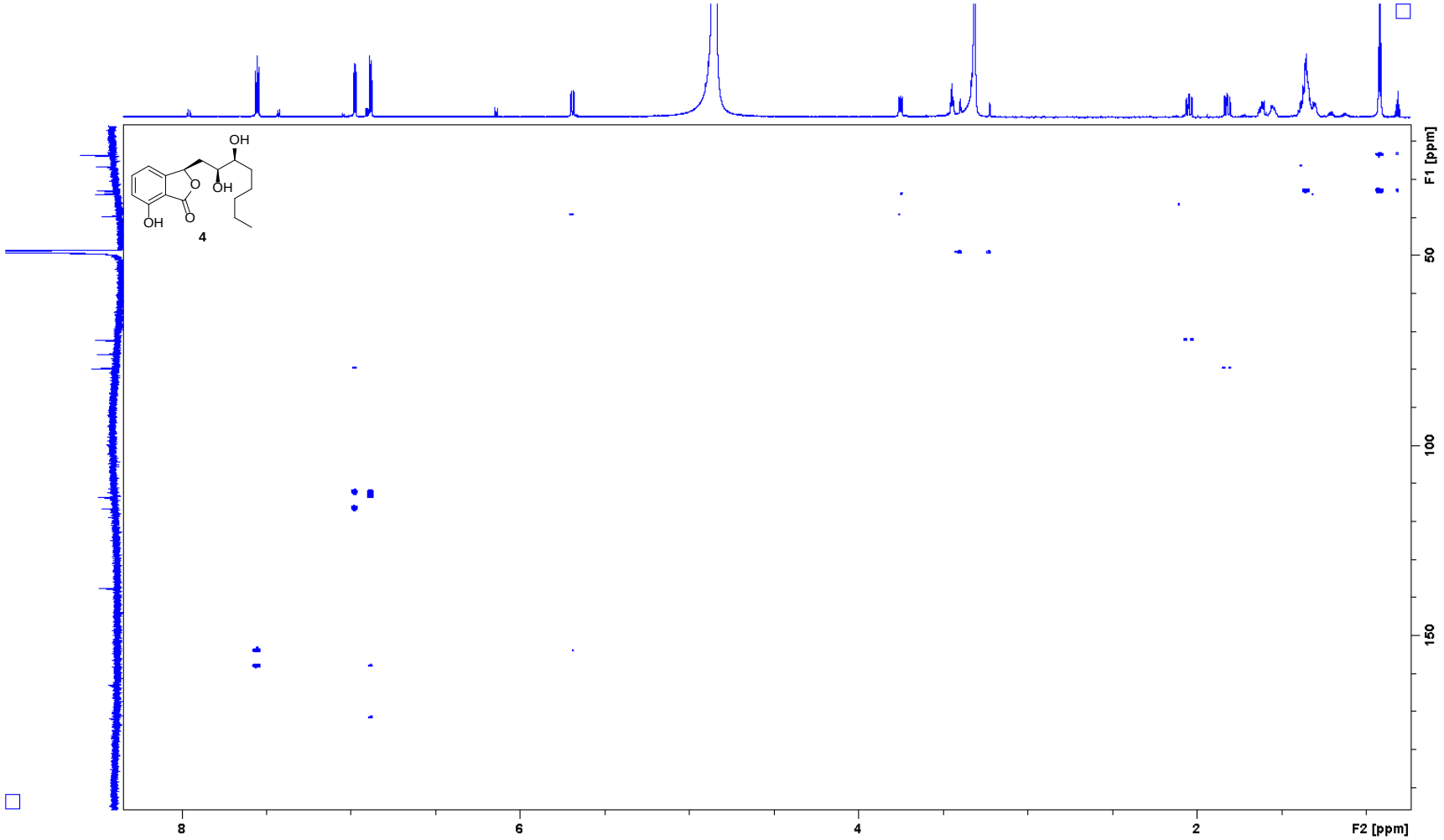


Figure S30. HMBC spectrum of **4**.

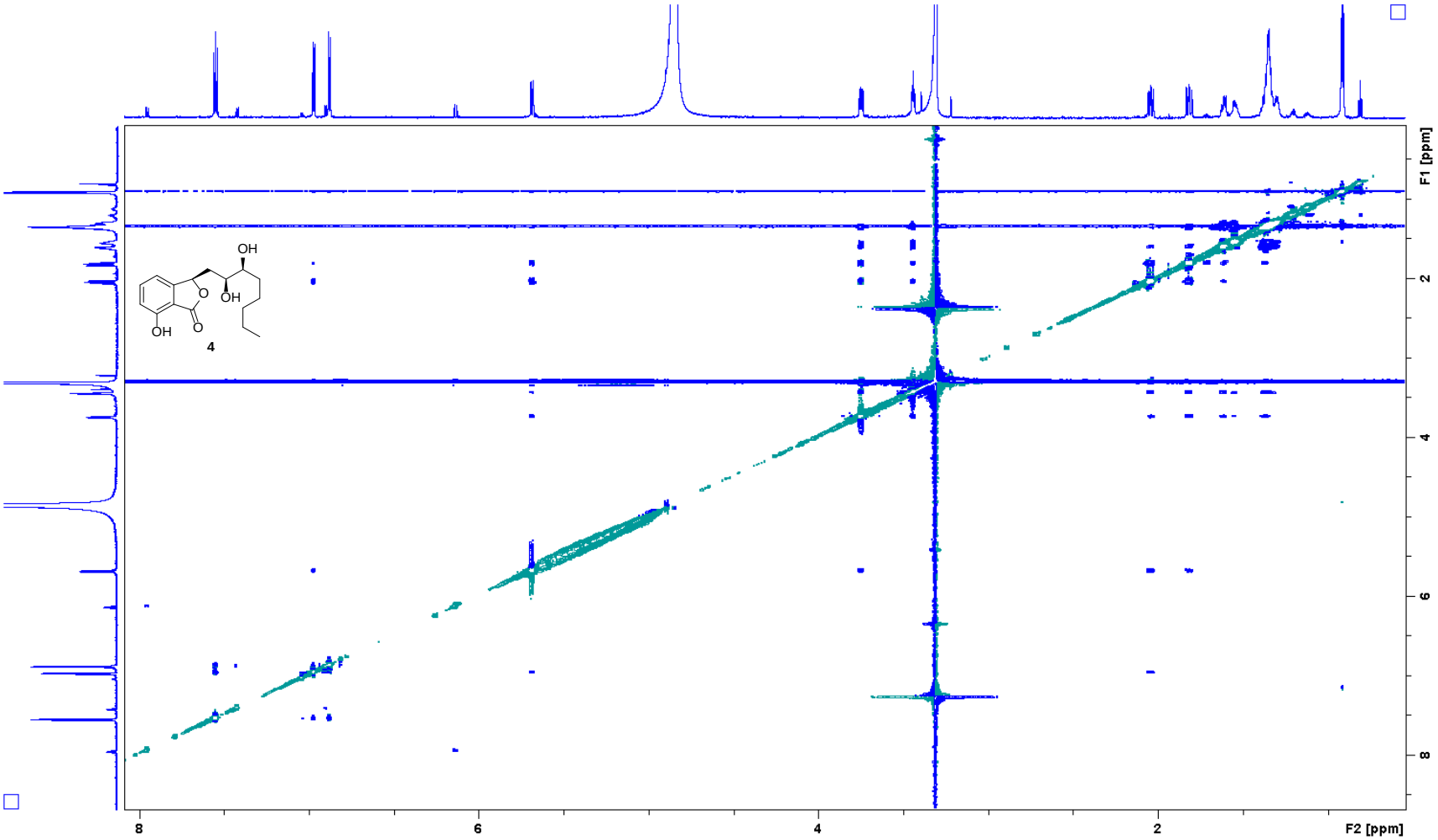


Figure S31. NOESY spectrum of **4**.

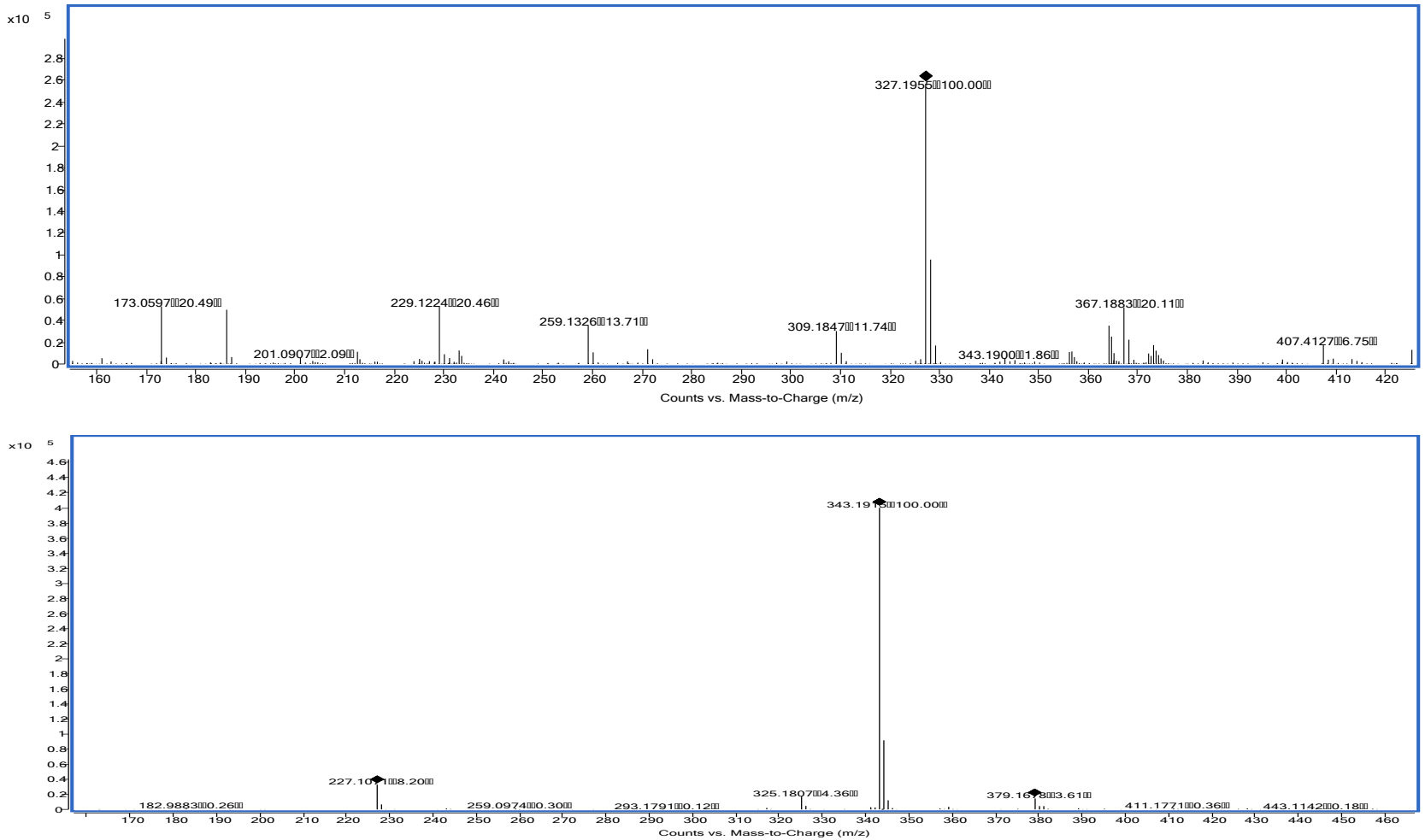


Figure S32. HRESIMS (ESI^+ top, ESI^- bottom) spectra of **5**.

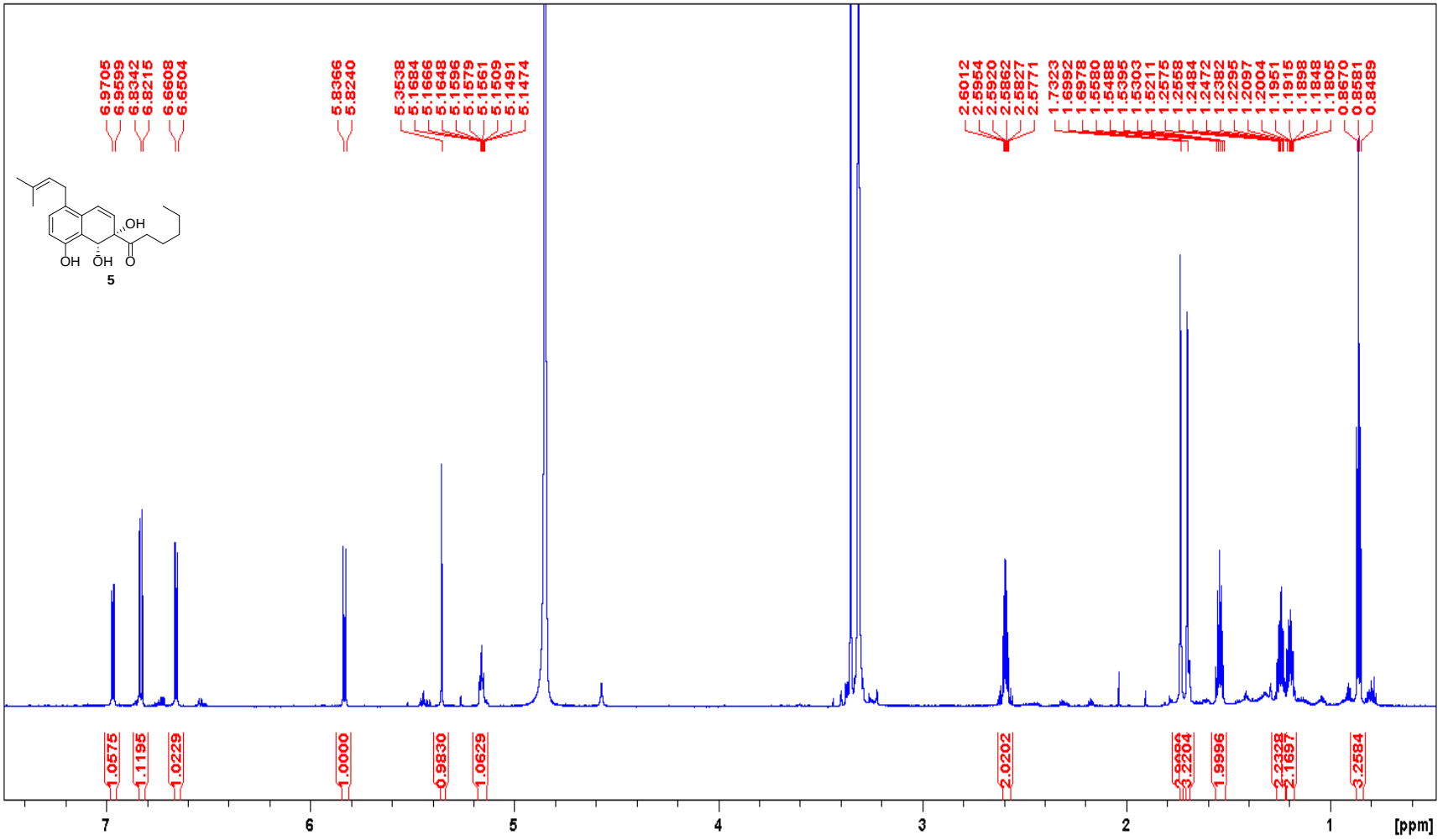


Figure S33. ^{1}H NMR spectrum of **5** (800 MHz, CD3OD).

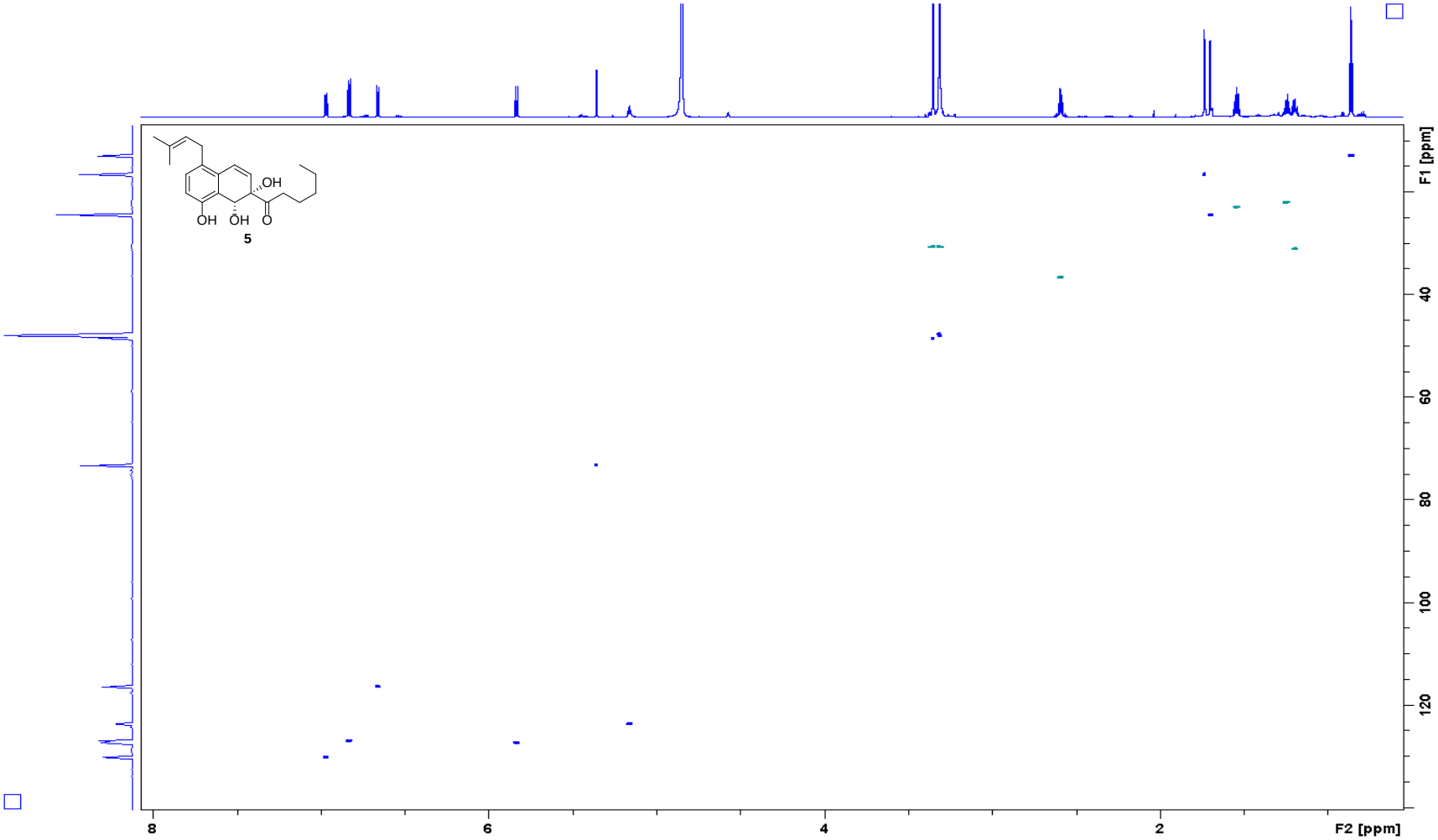


Figure S34. HSQC spectrum of **5**.

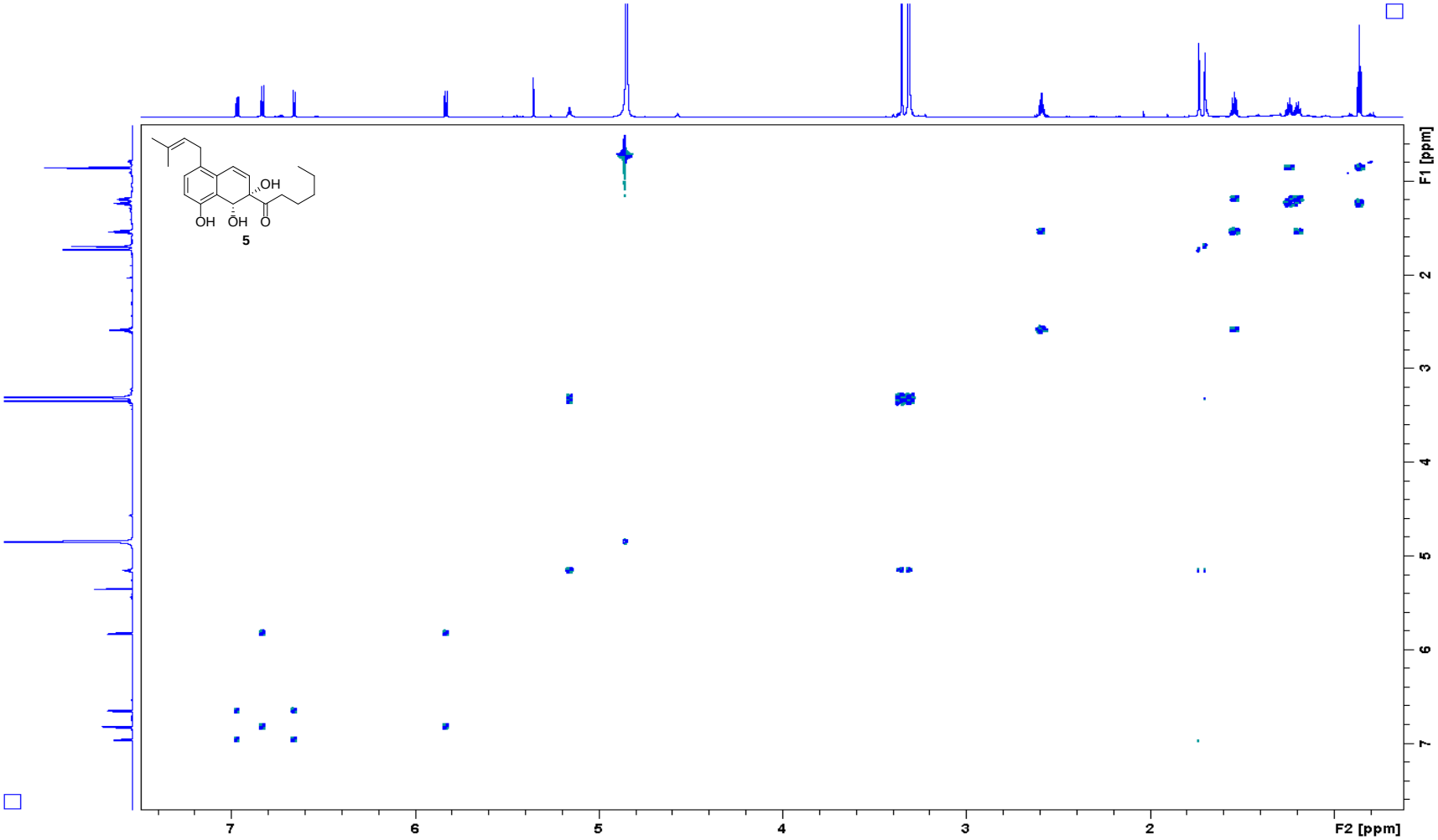


Figure S35. COSY spectrum of **5**.

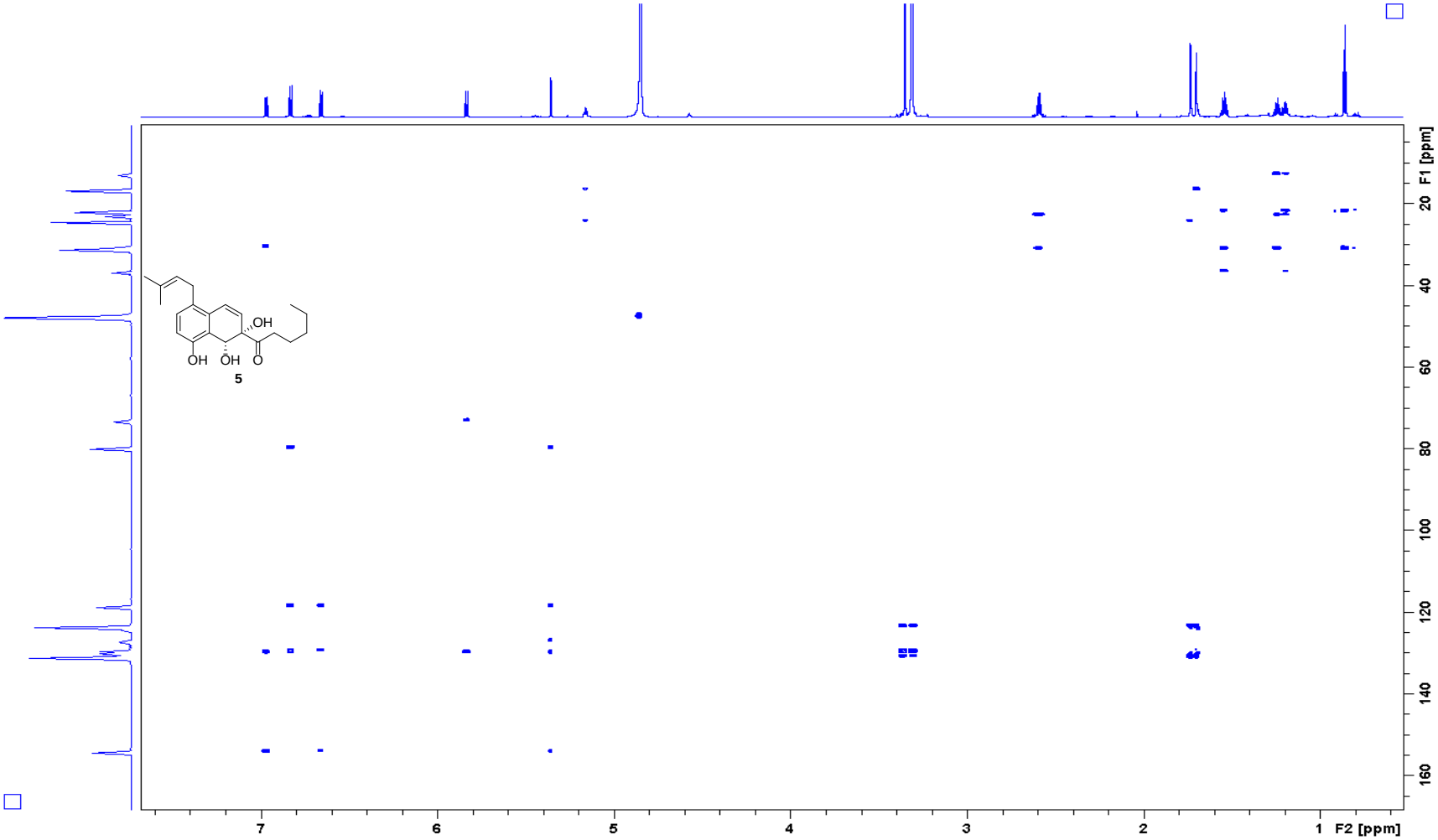


Figure S36. HMBC spectrum of **5**.

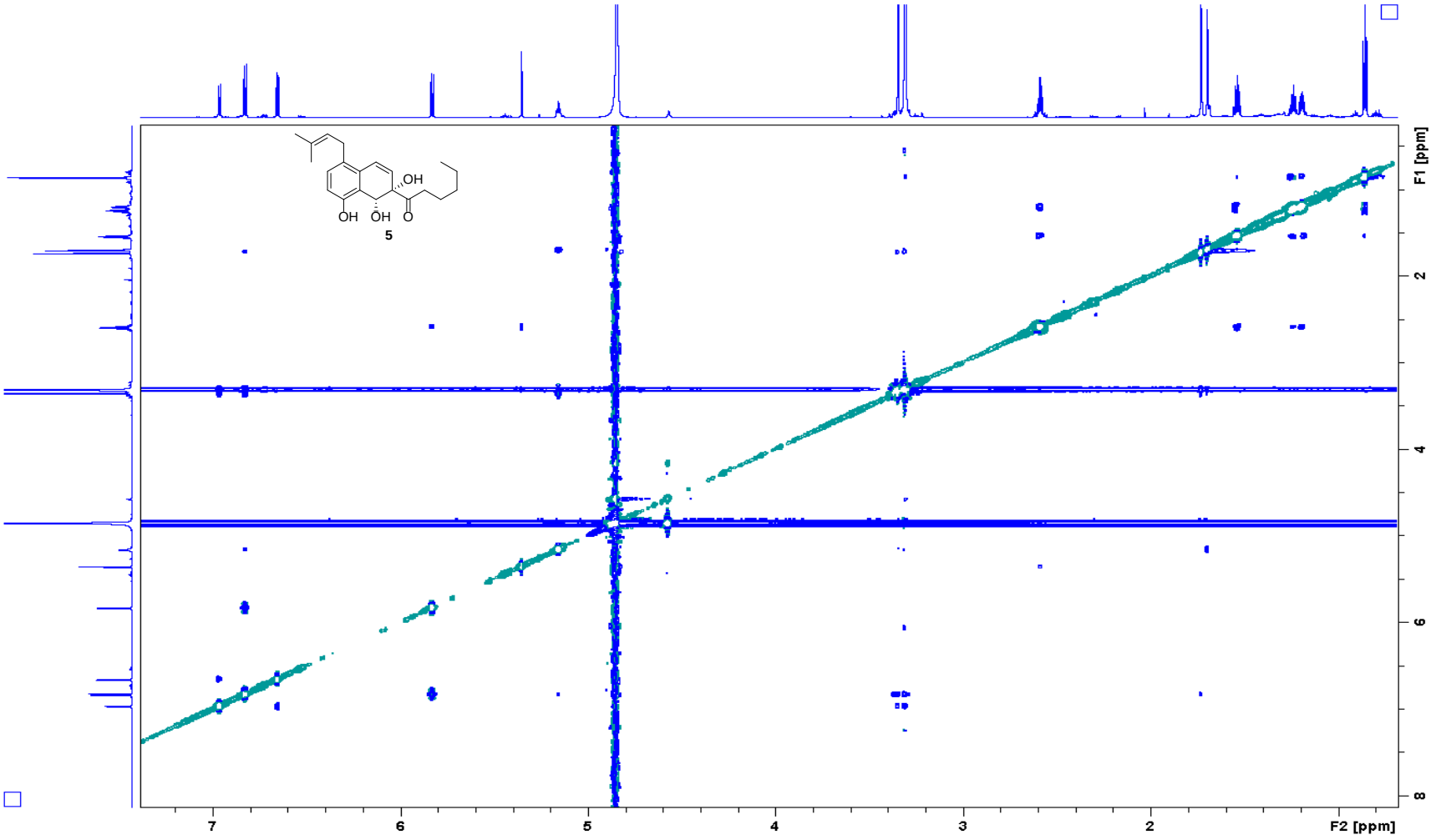


Figure S37. NOESY spectrum of **5**.

Figure S38. Structures of the predominant conformers of (-)-**5**.

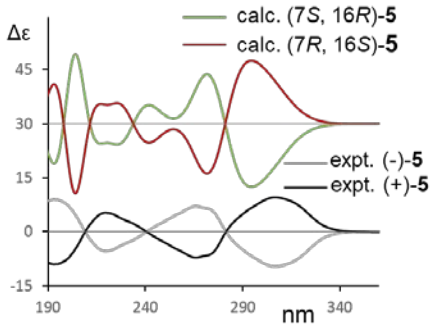
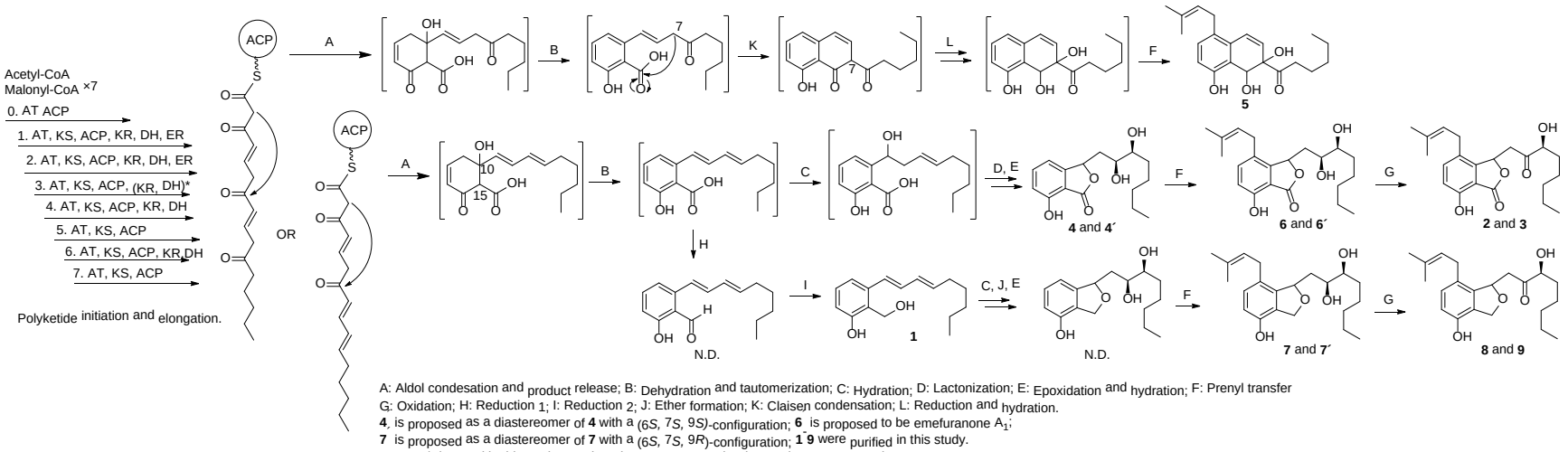


Figure S39. Comparison of the experimental and calculated ECD spectra of (-)-**5** and (+)-**5**.



A: Aldol condesation and product release; B: Dehydration and tautomerization; C: Hydration; D: Lactonization; E: Epoxidation and hydration; F: Prenyl transfer
G: Oxidation; H: Reduction 1; I: Reduction 2; J: Ether formation; K: Claisen condensation; L: Reduction and hydration.
4, is proposed as a diastereomer of **4** with a (6*S*, 7*S*, 9*S*)-configuration; **6** is proposed to be emefuranone A_1;
7 is proposed as a diastereomer of **7** with a (6*S*, 7*S*, 9*R*)-configuration; **1**-**9** were purified in this study.
N.D.: not detected in this study. Bracketed compounds can be detected. *optional reaction.

Figure S40. Proposed biosynthetic pathways of the polyketides.

Experimental procedures for antibacterial and cytotoxic assays

Antibacterial assay. The strains used were: Methicillin-resistant *Staphylococcus aureus* (MRSA MB5393) and methicillin sensitive *Staphylococcus aureus* (MSSA ATCC 29213) as examples of gram-positive bacteria model and *Escherichia coli* ATCC 25922*)* and *Klebsiella pneumoniae* ATCC 700603 as examples of Gram-negative bacteria model. All compounds were dissolved in DMSO to a stock solution of 6 mg/mL and were serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 96 µg/mL for all the antibacterial assays. For MRSA, MSSA, *K. pneumoniae* and *E. coli* liquid assays 90 µL of the appropriately diluted inocula in BHI (Brain Heart Infusion), MHII (Muller Hinton II) or LB (Luria Broth) medium are mixed with 1.6 µL of compound stock solution at 6 mg/mL plus 8.4 µL of the appropriate medium depending on the microorganism to be tested. Positive control compounds included in the assays: Aztreonan (2-0.25 µg/mL) was used as positive control *E. coli* ATCC 25922; Gentamycin (32-0.25 µg/mL) was used for *K. pneumonia*; Vancomycin (32-0.25 µg/mL) was used as a positive control for MRSA, MSSA.

For colorimetric assays absorbance at OD612nm was measured at zero time (T0) and at the final time (Tf). Absorbance was measured in an Envision reader (Perkin Elmer). Percentage of growth inhibition was calculated using the following normalization: % Inhibition = 100 × {1 − [(Tf Sample − T0 Sample) − (Tf Blank − T0 Blank)]/[(Tf Growth − T0 Growth) − (Tf Blank − T0 Blank)]}. The *Genedata Screener* software (Genedata, Inc., Basel, Switzerland) was used to process and analyze the data. The activity of the compounds was expressed as a percentage of inhibition growth, where (100%) represented the inhibition growth of the target microorganisms and 0% represented the total growth of target microorganisms on the assay. The Z' factor predicts the robustness of an assay by considering the mean and standard deviation of both positive and negative controls. The

robust Z′ factor (RZ′ factor) is based on the Z′ factor, but standard deviations and means were replaced by the robust standard deviations and medians, respectively. In all experiments performed in this work, the RZ′ factor obtained was higher than 0.90 in all cases.[1-3]

Cytotoxicity assay. The tumor cell lines used were A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast) and MiaPaca-2 (pancreas) and HL-60 (leukemia). All the compounds were assayed for cytotoxicity from 30 mg/ml, or 20 mg/ml (Calinaphthalenetriol A2) in 10 points of dilutions 1:2, per triplicated. Data obtained were analyzed using Genedata Screener Software and IC50 was calculated. Cytotoxicities against A549, A2058, HepG2, MCF-7 and MiaPaca-2 were assayed using MTT test. Cells were seeded at 10 000 cells/well in a 96-well plate. After 24 h, cells were treated with compounds for 72 h. MMS (methyl methanesulfonate) 4mM was used as a positive control of cell death. Then, MTT was added for 2 h and then absorbance was measured at 570 nm; Cytotoxicities against HL-60 was assayed using CCK-8 test. Cells were seeded at 5 000 cells/well in a 384-well plate. After 24 h, cells were treated with compounds for 72 h. MMS 4mM was used as a positive control of cell death. Then, WST-8 was added for 2 h and then absorbance was measured at 450 nm.[1-3]

Table S1 Antibacterial assay results of **1-3** and **5-7**.

Compounds	MICs (μg/mL)			
	MRSA MB5393	*MSSA* ATCC 29213	*E. coli* ATCC 25922	*K. pneumoniae* ATCC 700603
1	48	48	>96	>96
2	96	>96	>96	>96
3	96	>96	>96	>96
(-)-**5**	>96	>96	>96	>96
(+)-**5**	>64	>64	>64	>64
6	>96	>96	>96	>96
7	96	>96	>96	>96

Table S2 Cytotoxic assay results of **1-3** and **5-7**.

Compounds	IC_{50} (μM)					
	A549	A2058	HepG2	MCF-7	MiaPaca-2	HL-60
1	>121	>121	>121	>121	>121	121
2	>83.2	>83.2	>83.2	>83.2	>83.2	80.4
3	>83.2	>83.2	>83.2	>83.2	>83.2	74.9
(-)-**5**	>87.1	>87.1	>87.1	>87.1	>87.1	58.1
(+)-**5**	>58.1	>58.1	>58.1	58.1	>58.1	29.0
6	>82.8	>82.8	>82.8	>82.8	>82.8	74.5
7	>86.1	80.3	28.4	80.3	>86.1	57.4

Supplementary references

(1) Zhang, J. H.; Chung, T. D. Y.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4*, 67–73.

(2) Kümmel, A.; Gubler, H.; Gehin, P.; Beibel, M.; Gabriel, D.; Christian, N. P. Integration of Multiple Readouts Into the Z' Factor for Assay Quality Assessment. *J. Biomol. Screen.* **2010**, *15*, 95–101.

(3) Audoin, C.; Bonhomme, D.; Ivanisevic, J.; De La Cruz, M.; Cautain, B.; Monteiro, M. C.; Reyes, F.; Rios, L.; Perez, T.; Thomas, O. P. Balibalosides, an Original Family of Glucosylated Sesterterpenes Produced by the Mediterranean Sponge *Oscarella balibaloi*. *Mar. Drugs* **2013**, *11* (5), 1477–1489.

New naphthyl derivatives from *Aspergillus californicus*

Yaojie Guo[1] · Simone Ghidinelli[2] · Mercedes de la Cruz[3] · Thomas A. Mackenzie[3] · Maria C. Ramos[3] · Pilar Sánchez [3] · Francisca Vicente[3] · Olga Genilloud[3] · Thomas O. Larsen[1]

[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800, Kgs. Lyngby, Denmark

[2] Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy

[3] Fundación MEDINA, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016, Armilla, Granada, Spain

Abstract

Two new naphthyl-products calinaphthyltriol A (**1**) and calinaphthalenone A (**2**) were isolated from *Aspergillus californicus* IBT 16748 together with one known compound ophiobolin X (**3**). Their structures were elucidated by extensive spectroscopic analyses. The absolute configuration of **2** was solved by comparing its optical rotation with data for the known compounds **4**, **5** and **6** as well as theoretical calculations. The antibacterial and cytotoxic activities of **1** and **3** were evaluated. Both compounds did not show antibacterial activity (MIC > 96 ug/mL) against a few selected clinically relevant Gram positive and Gram negative bacterial strains. However, they showed moderate cytotoxicity against HL-60 cell line with IC_{50} values of 18 and 24 µg/ml, respectively.

Introduction

Filamentous fungi in genus *Aspergillus* are well known for their abilities to produce various extrolites including secondary metabolites (SMs) and enzymes [1–3]. In our continuing search for novel bioactive compounds we set out to investigate the chemical potential of the rare species *Aspergillus californicus* IBT 16748 which was isolated from Chamise chaparral (*Adenostoma fasciculatum*) soil in California [4, 5]. In this paper, we report the isolation and characterization of two new compounds naphthyl compounds calinaphthyltriol A (**1**), calinaphthalenone A (**2**) and a known product ophiobolin X (**3**) [6] (Fig. 1), as well as their antibacterial and cytotoxic activities.

Results and discussion

The chemical formula of **1** was assigned as $C_{15}H_{16}O_3$ based on detection of a $[M+H]^+$ ion at 245.1179 Da and a $[M+Na]^+$ ion at 267.0996 Da. The ^1H NMR spectrum (Table 1) exhibited 13 hydrogens including two methyl groups δ_H 1.72 (d, J = 1.1 Hz, H_3-12) and 1.74 (s, H_3-13), two methylene hydrogens δ_H 3.39 (d, J = 6.8 Hz, H_2-9), one olefinic hydrogen δ_H 5.17 (tq, J = 7.1, 1.4 Hz, H-10), four aryl protons δ_H 6.35 (d, J = 12.3 Hz, H-3), 7.47 (d, J = 12.3 Hz, H-4), 7.13 (d, J = 8.3 Hz, H-6) and 6.88 (d, J = 8.3 Hz, H-7). The naphthyl scaffold was assigned through a group of HMBC correlations (Fig. 2) from H-3 to δ_C 171.2 (C-1) and 136.2 (C-4a), H-4 to δ_C 123.6 (C-8a), H-6 to C-4a and δ_C 154.3 (C-8), H-7 to C-8a, C-8 and C-2. The prenyl unit was determined to be attached on C-5 due to the HMBC resonances from H-9 to C-4a, C-5 and C-6, which was supported by the NOE correlations observed between H_2-9 and H-4, H-9 and H-6. Compound **1** bears the same core skeleton as the synthetic compound 1-(3-Methylbut-2-enyl)naphthalene [7], but with three hydroxyl groups at C-1, C-2 and C-8. Thus, the structure of **1** calinaphthyltriol A was assigned.

Compound **2** was obtained as white amorphous solid and its molecular formula was determined as $C_{16}H_{20}O_4$ due to detection of a $[M+H]^+$ ion at 277.1439 Da and a $[M+Na]^+$ ion at 299.1253 Da. The ^1H NMR (Table 1) exhibited 17 hydrogens including one methyl group δ_H 0.83 (t, J = 7.2 Hz, H_3-14), eight methylene hydrogens δ_H 1.19 (m, H-11b), δ_H 1.19 (m, H_2-12), 1.23 (m, H_2-13), 1.34 (m, H-11a), 1.60 (m, H-10b) and 1.79 (m, H-10a), two oxygenated methylene hydrogens δ_H 4.80 (d, J = 11.8 Hz, H-9b) and 4.83 (d, J = 11.8 Hz, H-9a), four olefinic protons δ_H 6.26 (d, J = 10.2 Hz, H-3), 7.00 (d, J = 10.2 Hz, H-4), 6.81 (d, J = 8.5 Hz, H-7) and 7.79 (d, J = 8.5 Hz, H-8). The 1(2H)-naphthalenone nucleus was assigned through the multiple HMBC correlations from H-3 to δ_C 203.6 (C-1) and 140.1 (C-4a), H-4 to δ_C 76.9 (C-2), 123.23 (C-8a) and C-4a, H-7 to δ_C 124.2 (C-5), 163.8 (C-6) and C-8a, H-8 to C-4a, C-6 and C-1. The location of hydroxymethyl group was determined based on the HMBC correlations from H_2-9 to C-4a, C-5 and C-6 while the location of the pentyl group was determined based on the HMBC correlations from H_2-10 to C-1 and C-2. The absolute configuration of **2** was determined by comparing its optical rotation (OR) with its analogue biatriosporin E (**4**) [8] since both compounds share a single chiral center at the same position of the 1(2H)-naphthalenone skeleton. Biatriosporin E (**4**) showed negative OR in $CHCl_3$ and the chiral center was assigned as *S* configuration, which implied that **2** on the contrary displaying a positive OR in $CHCl_3$ should be determined as *R* configuration. This assignment was supported by TD-DFT calculation of OR of *R*-**2** (Fig. 3) which indicated a positive OR signal at 589 nm. Thus, **2** with a 2*R*-configuration was assigned.

However, our conclusion was inconsistent with the data given in the literature of the two synthetic compounds (*S*)-(+)-*ortho*-quinol (**5**) and (*R*)-(-)-*ortho*-quinol (**6**) [9], whose absolute configurations were assigned by vibrational circular dichroism (VCD). This drove us to look into these compounds by VCD calculations. Surprisingly, our calculated VCD spectrum for (*S*)-*ortho*-

quinol matches the experimental data (-)-*ortho*-quinol better than the calculated one for (*R*)-*ortho*-quinol reported in reference [9]. The comparison of our calculations with the original data is reported in Figure S17. Therefore, we propose that the absolute configurations of **5** and **6** were likely mis-reported in the original work and they should be revised to (*S*)-(-)-*ortho*-quinol (**5**) and (*R*)-(+)-*ortho*-quinol (**6**).

Table 1 ^{1}H NMR (800 MHz) and ^{13}C NMR Data of **1** and **2** in CD_3OD

Position	**1**		**2**	
	δ_{C}[a], type	δ_{H}, multi (*J* in Hz)	δ_{C}, type	δ_{H}, multi (*J* in Hz)
1	171.2, C		203.6, C	
2	133.2, C		76.9, C	
3	122.2, CH	6.35, d (12.3)	138.9, CH	6.26, d (10.2)
4	140.0, CH	7.47, d (12.3)	123.21, CH	7.00, d (10.2)
4a	136.2, C		140.1, C	
5	133.4, C		124.2, C	
6	132.2, CH	7.13, d (8.3)	163.8, C	
7	118.1, CH	6.88, d (8.3)	115.9, CH	6.81, d (8.5)
8	154.3, C		129.8, CH	7.79, d (8.5)
8a	123.6, C		123.23, C	
9	32.5, CH_2	3.39, d (6.8)	55.1, CH_2	4.83[b], d (11.8); 4.80[b], d (11.8)
10	124.1, CH	5.17, tq (7.1, 1.4)	42.5, CH_2	1.79, m; 1.60, m
11	133.5, C		24.1, CH_2	1.34, m; 1.19, m
12	25.8, CH_3	1.72, d (1.1)	33.2, CH_2	1.19, m
13	17.9, CH_3	1.74, s	23.4, CH_2	1.23, m
14			14.2, CH_3	0.83, t (7.2)

[a] Chemical shifts assigned by HSQC and HMBC correlations

[b] Signals partially obscured by water peak, **Figure S13**-**S14** and **S16**.

Compounds **1** and **3** were subjected to antibacterial and cytotoxic assays. To assess their antibacterial properties they were tested against two Gram-positive bacteria (methicillin-resistant *Staphylococcus aureus* MB5393 and methicillin-sensitive *Staphylococcus aureus* ATCC 29213), and two Gram-negative bacteria (*Escherichia coli* ATCC 25922 and *Klebsiella pneumoniae* ATCC 700603). Both compounds were inactive with MICs above 96 µg/ml. Six tumor cell lines A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast), Mia PaCa-2 (pancreas), HL-60 (leukemia) were used in the cytotoxicity test. Both showed moderate cytotoxicity against HL-60 cell line with IC_{50} values 18 and 24 µg/ml, respectively.

Material and methods

General experimental procedure

Optical rotations were measured on a PerkinElmer 341 polarimeter. IR data were registered on a Bruker IFS 120 FTIR spectrometer. The NMR spectra were performed on a Bruker AVANCE 800 MHz system (DTU800), equipped with a 5 mm TCI Cryoprobe. Chemical shifts were reported in ppm with reference to the solvent signals (CD_3OD δ_H 3.31 and δ_C 49.00). UHPLC-DAD-HRESIMS was recorded on an Agilent Infinity 1290 UHPLC - 6545 QTOF MS system. Flash chromatography was carried out on a Biotage Isolera One system. Semipreparative HPLC was performed on a Waters 600 controller with a photodiode array detector. The columns used for purification were Phenomenex Kinetex C_{18} column (250 × 10 mm, 5 µm) and Phenomenex Luna phenyl-hexyl column (250 × 10 mm, 5 µm).

Fungal material

Aspergillus californicus IBT 16748 was obtained from the IBT culture collection, Department of Biotechnology and Biomedicine, Technical University of Denmark.

Fermentation and the isolation of compounds

The medium used was medium 2 [10] with 2% (w/v) agar. The strain was grown on 270 plates at 25 °C for 15 days and the plate contents were extracted twice with ethyl acetate containing 1% formic acid. The extract was subjected to liquid-liquid separation to yield dichloromethane (DCM) part (1.1 g) which was applied to the flash chromatography with a self-packed diol column (25 g, 33 mL) using a stepwise gradient elution DCM-EtOAc-MeOH at 25 mL/min to afford nine fractions. Fr.5 (0.2 g) was separated with a self-packed C18-T column (Phenomenex Sepra, 25 g, 33 mL) to yield 14 sub-fractions of which sub-fr.5 (18.8 mg) was further applied to the C_{18} column (MeCN/H_2O, 40-47% in 20 min, 4 mL/min) to yield compound **1** (0.74 mg, t_R = 17.5 min), and sub-fr.4 (11.5 mg) was applied to the phenyl-hexyl column and then the C_{18} column (MeCN/H_2O, 30% isocratic in 20 min, 4 mL/min) to yield compound **2** (1.0 mg, t_R = 14.0 min). Compounds **3** was isolated from the extract of *A. californicus* IBT 16748 on 120 CYA plates and 120 YES plates for 11 days at 25°C in dark conditions. The combined DCM part (1.266 g) was applied to the flash chromatography with a self-packed NH_2 column (Sepra Phenomenex, 50g, CV 66 mL) to yield eight fractions. Fr3 (10.8 mg) was subjected to the C_{18} column (MeCN/H_2O, 60-90% in 20 min, 4 mL/min) to get a sub-fraction (1.24 mg, t_R = 15.0 min) which was purified on the phenyl-hexyl column (MeCN/H_2O, 75-79% in 16 min, 4 mL/min) to give **3** (0.45 mg, t_R = 11.0 min).

Calinaphthyltriol A (**1**): white amorphous solid; HPLC-UV (MeCN/H_2O + 20 mM formic acid) λ_{max} 228, 248, 280, 327 nm; IR (dry film) ν_{max} 3382, 2933, 2872, 1718, 1669, 1636, 1591, 1456, 1387, 1263, 1199, 1117, 1084, 1044, 864, 823 cm^{-1}; ^1H and ^{13}C NMR data see Table 1. HRESIMS m/z $[M+H]^+$ 245.1179 (calcd for $C_{15}H_{17}O_3^+$, 245.1172); $[M+Na]^+$ 267.0996 (calcd for $C_{15}H_{16}O_3Na^+$, 267.0992).

Calinaphthalenone A (**2**): white amorphous solid; $[\alpha]_D^{20}$= +50 (c = 0.014, $CHCl_3$); HPLC-UV (MeCN/H_2O + 20 mM formic acid) λ_{max} 256, 316 nm; IR (dry film) ν_{max} 3210, 2957, 2928, 2871,

2856, 1733, 1637, 1457, 1380, 1271, 1082, 810 cm^{-1}; ^{1}H and ^{13}C NMR data see Table 1. HRESIMS m/z $[M+H]^+$ 277.1439 (calcd for $C_{16}H_{21}O_4^+$, 277.1434); $[M+Na]^+$ 299.1253 (calcd for $C_{16}H_{20}O_4Na^+$, 299.1254).

Computational Details. Molecular Mechanics (MM) search of conformers within 5 kcal/mol was performed using Spartan program [11] with MMFF94s as the force field. Conformers have been optimized at DFT level with B3LYP/TZVP through Gaussian 16 platform [12] in IEF-PCM approximation [13]. VCD intensities were calculated at DFT level with the B3LYP functional and 6-311G+(d,p) basis set [9]. Calculations of ORD spectra were performed at the TD-DFT level of theory using CAM-B3LYP/6-311++(2d,2p) in IEF-PCM approximation [14,15].

Acknowledgements The authors thank Dr. Kasper Enemark-Rasmussen and Associate Professor René Wugt Larsen from Department of Chemistry at the Technical University of Denmark for acquiring NMR and IR data, respectively. Yaojie Guo is financially supported by the China Scholarship Council and Department of Biotechnology and Biomedicine at the Technical University of Denmark.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Frisvad JC, Larsen TO. Chemodiversity in the genus *Aspergillus*. Appl Microbiol Biotechnol. 2015;99:7859–7877.
2. Alburae NA, Mohammed AE, Alorfi HS, Jamanturki A, Asfour HZ, Alarif WM, et al. Nidulantes of *Aspergillus* (Formerly *Emericella*): A treasure trove of chemical diversity and

biological activities. Metabolites;2020:10:1–28.

3. Sanchez JF, Somoza AD, Keller NP, Wang CCC. Advances in *Aspergillus* secondary metabolite research in the post-genomic era. Nat Prod Rep. 2012;29:351–371.

4. Samson RA, Varga J, Meijer M, Frisvad JC. New taxa in *Aspergillus* section *Usti*. Stud Mycol. 2011;69:81–97.

5. Chen AJ, Frisvad JC, Sun BD, Varga J, Kocsubé S, Dijksterhuis J, et al. *Aspergillus* section *Nidulantes* (formerly *Emericella*): Polyphasic taxonomy, chemistry and biology. Stud Mycol. 2012;84:1–118.

6. Zhu T, Lu Z, Fan J, Wang L, Zhu G, Wang Y, et al. Ophiobolins from the mangrove fungus *Aspergillus ustus*. J Nat Prod. 2018;81:2–9.

7. Tsukamoto H, Sato M, Kondo Y. Palladium(0)-catalyzed direct cross-coupling reaction of allyl alcohols with aryl- and vinyl-boronic acids. Chem Commun. 2004;10:1200–1201.

8. Zhou YH, Zhang M, Zhu RX, Zhang JZ, Xie F, Li X Bin, et al. Heptaketides from an endolichenic fungus *Biatriospora* sp. and their antifungal activity. J Nat Prod. 2016;79:2149–2157.

9. Quideau S, Lyvinec G, Marguerit M, Bathany K, Ozanne-Beaudenon A, Buffeteau T, et al. Asymmetrie hydroxylative phenol dearomatization through in situ generation of iodanes from chiral iodoarenes and m-CPBA. Angew Chemie - Int Ed. 2009;48:4605–4609.

10. Cacho RA, Jiang W, Chooi YH, Walsh CT, Tang Y. Identification and characterization of the echinocandin b biosynthetic gene cluster from *Emericella rugulosa* NRRL 11440. J Am Chem Soc. 2012;134:16781–16790.

11. Wavefunction Inc. Spartan. Irvine, CA. https://www.wavefun.com/.

12. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman,

et al. Gaussian 16. Gaussian, Inc. Wallingford CT.

13. Mennucci B, Tomasi J, Cammi R, Cheeseman JR, Frisch MJ, Devlin FJ, et al. Polarizable Continuum Model (PCM) calculations of solvent effects on optical rotations of chiral molecules. J. Phys. Chem. A. 2002;106:6102−6113.

14. Rossi D, Nasti R, Collina S, Mazzeo G, Ghidinelli S, Longhi G, et al. The role of chirality in a set of key intermediates of pharmaceutical interest, 3-aryl-substituted-γ-butyrolactones, evidenced by chiral HPLC separation and by chiroptical spectroscopies. J Pharm Biomed Anal. 2017;10:144:41-51.

15. Mazzeo G, Cimmino A, Masi M, Longhi G, Maddau L, Memo M, et al. Importance and Difficulties in the use of chiroptical methods to assign the absolute configuration of natural products: the case of phytotoxic pyrones and furanones produced by diplodia corticola. J. Nat. Prod. 2017;80:2406−2415.

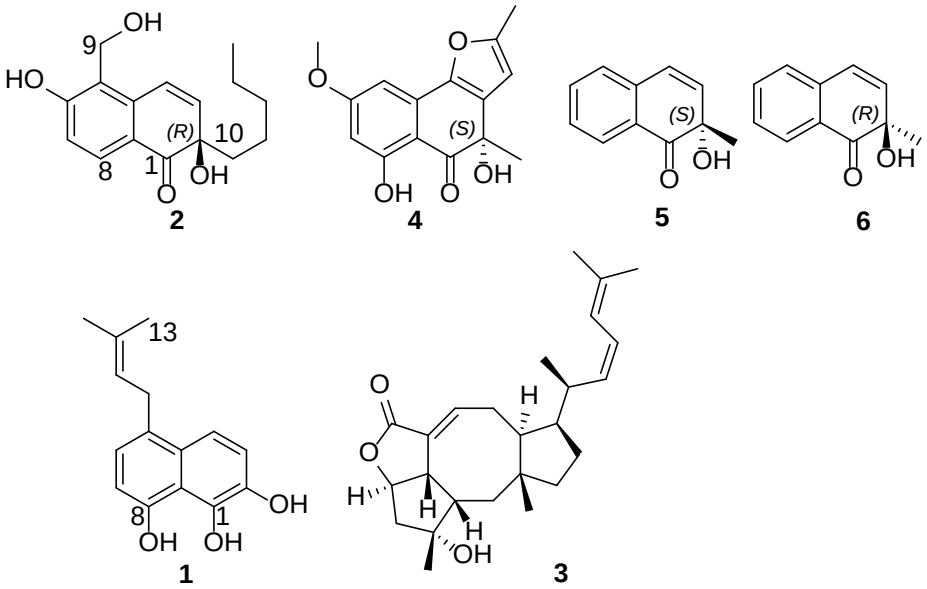


Fig. **1** Structures of compounds **1**-**6**

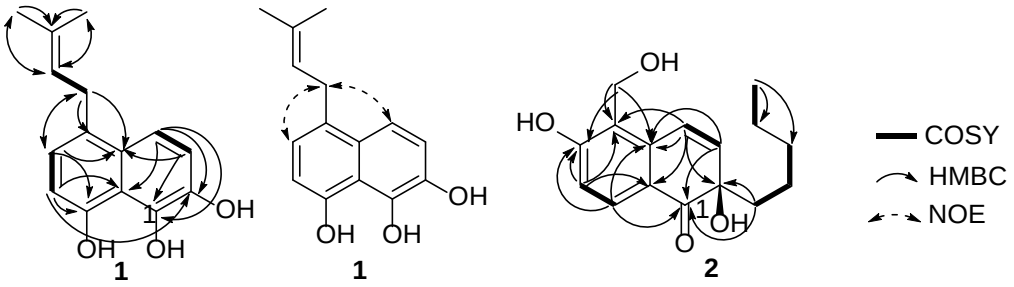


Fig. 2 Selected HMBC, COSY and NOE correlations of **1** and **2.**

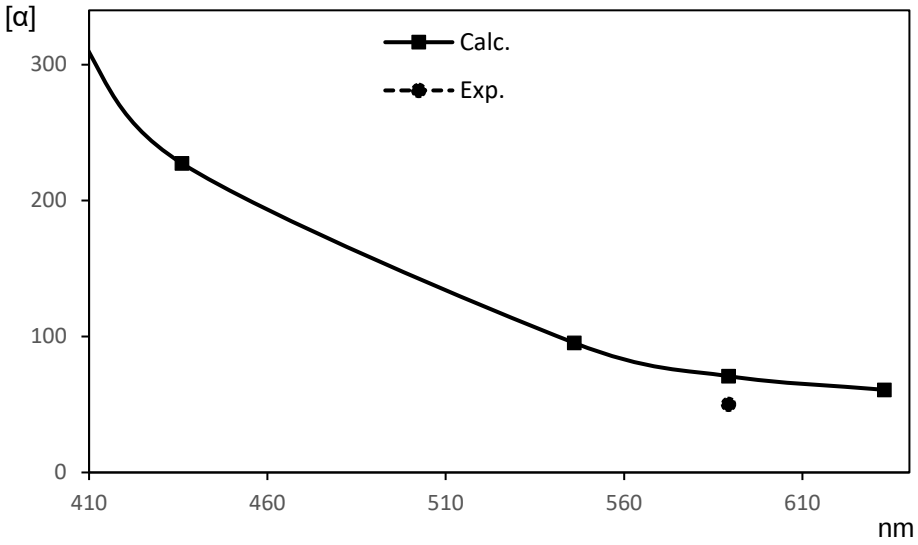


Fig. 3 Comparison of the experimental and calculated optical rotation for (*R*)-**2**

Supplementary information for

New naphthyl derivatives from *Aspergillus californicus*

Yaojie Guo[1] · Simone Ghidinelli[2] · Mercedes de la Cruz[3] · Thomas A. Mackenzie[3] · Maria C. Ramos[3] · Pilar Sánchez[3] · Francisca Vicente[3] · Olga Genilloud[3] · Thomas O. Larsen[1]

[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800, Kgs. Lyngby, Denmark

[2] Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy

[3] Fundación MEDINA, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016, Armilla, Granada, Spain

Table of Contents

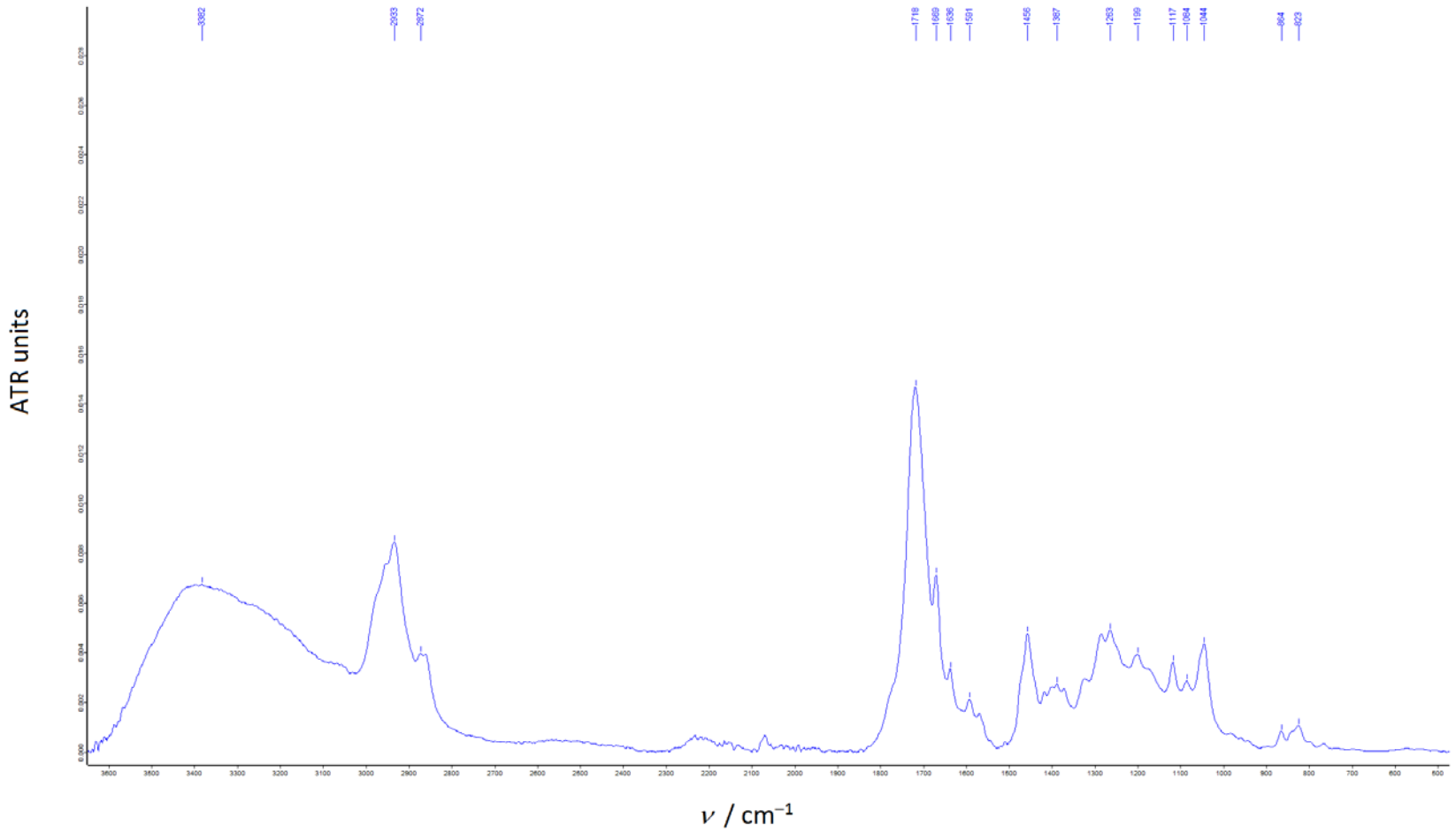


Figure S1. IR spectrum of **1**.

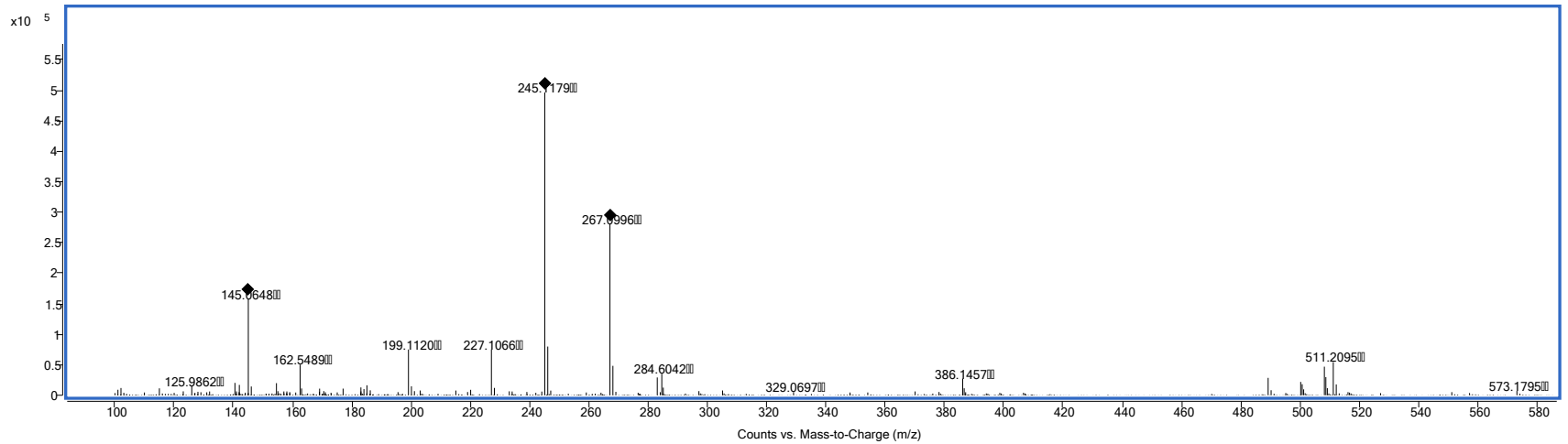


Figure S2. HRMS (ESI^+) spectrum of **1**.

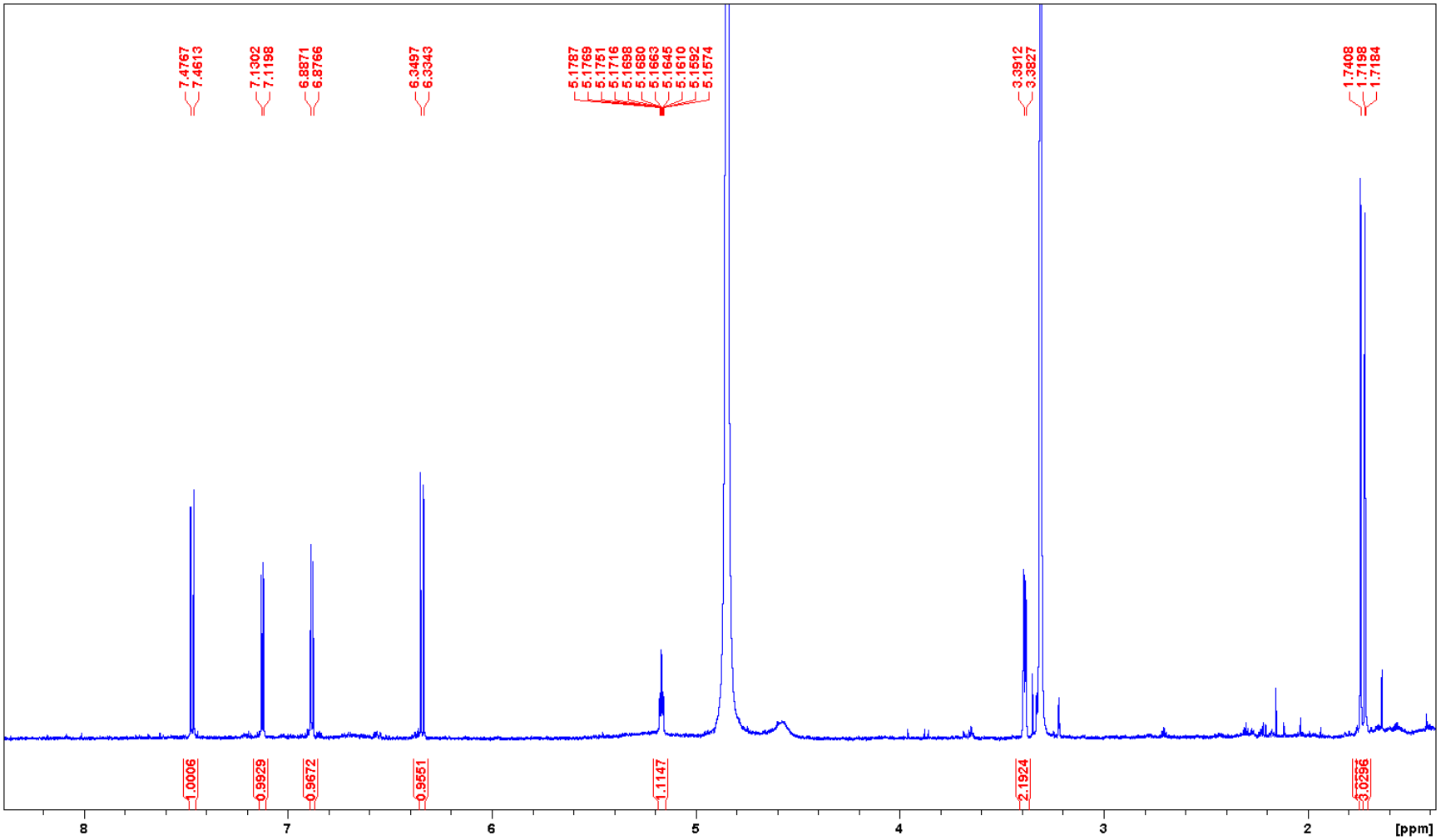


Figure S3. ^{1}H NMR spectrum of **1** (800 MHz, CD_3OD).

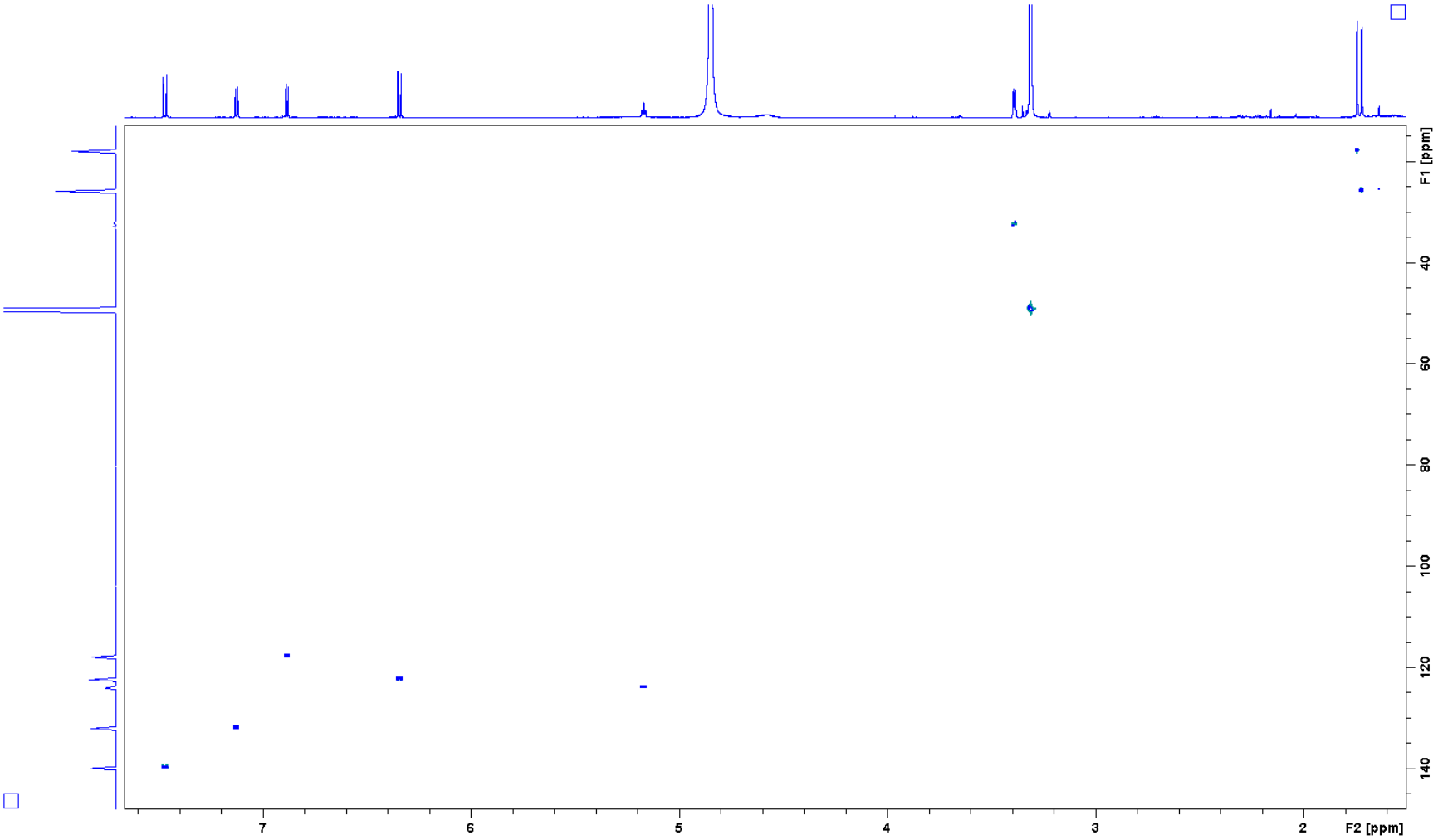


Figure S4. HSQC spectrum of **1**.

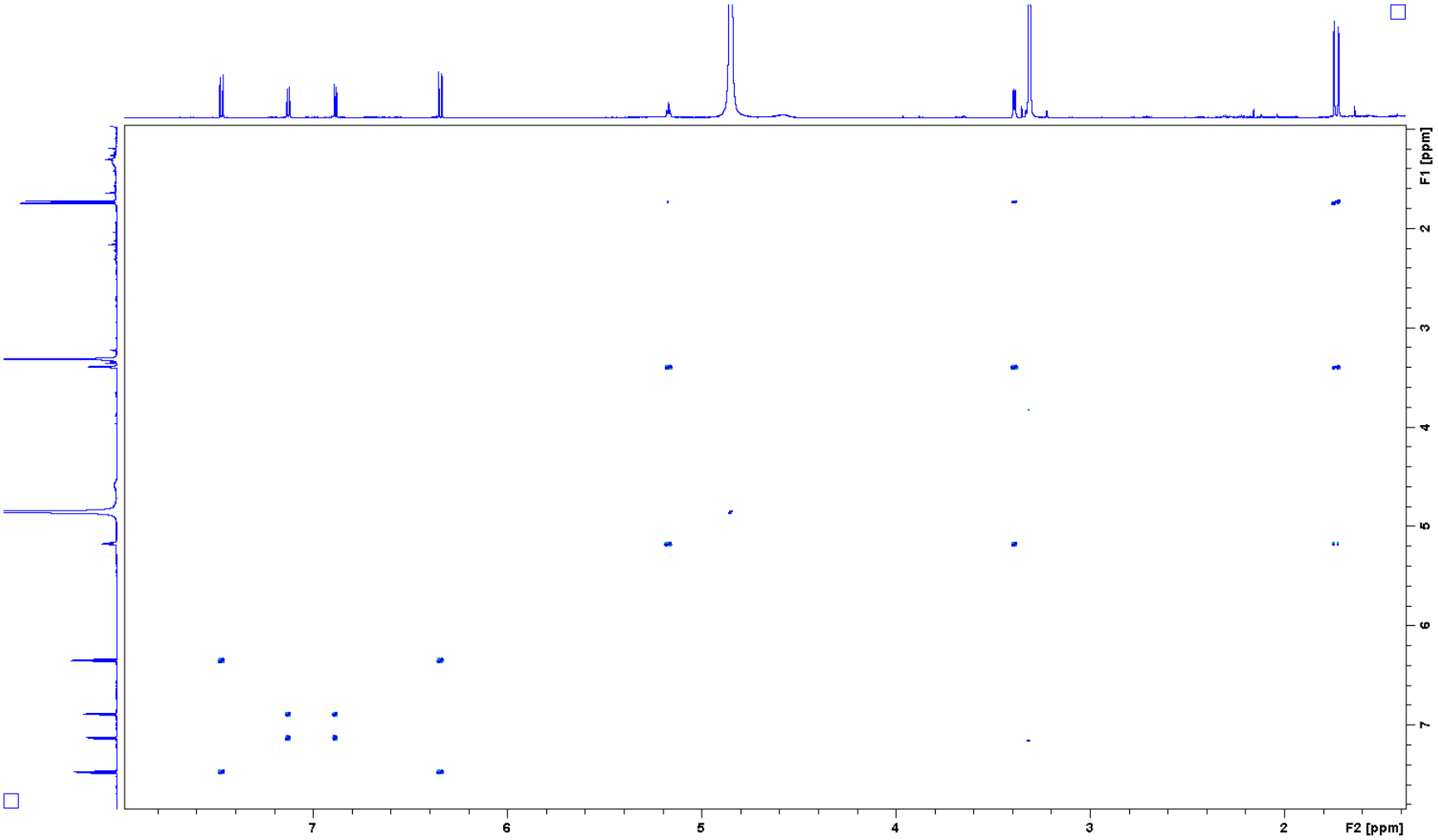


Figure S5. COSY spectrum of **1**.

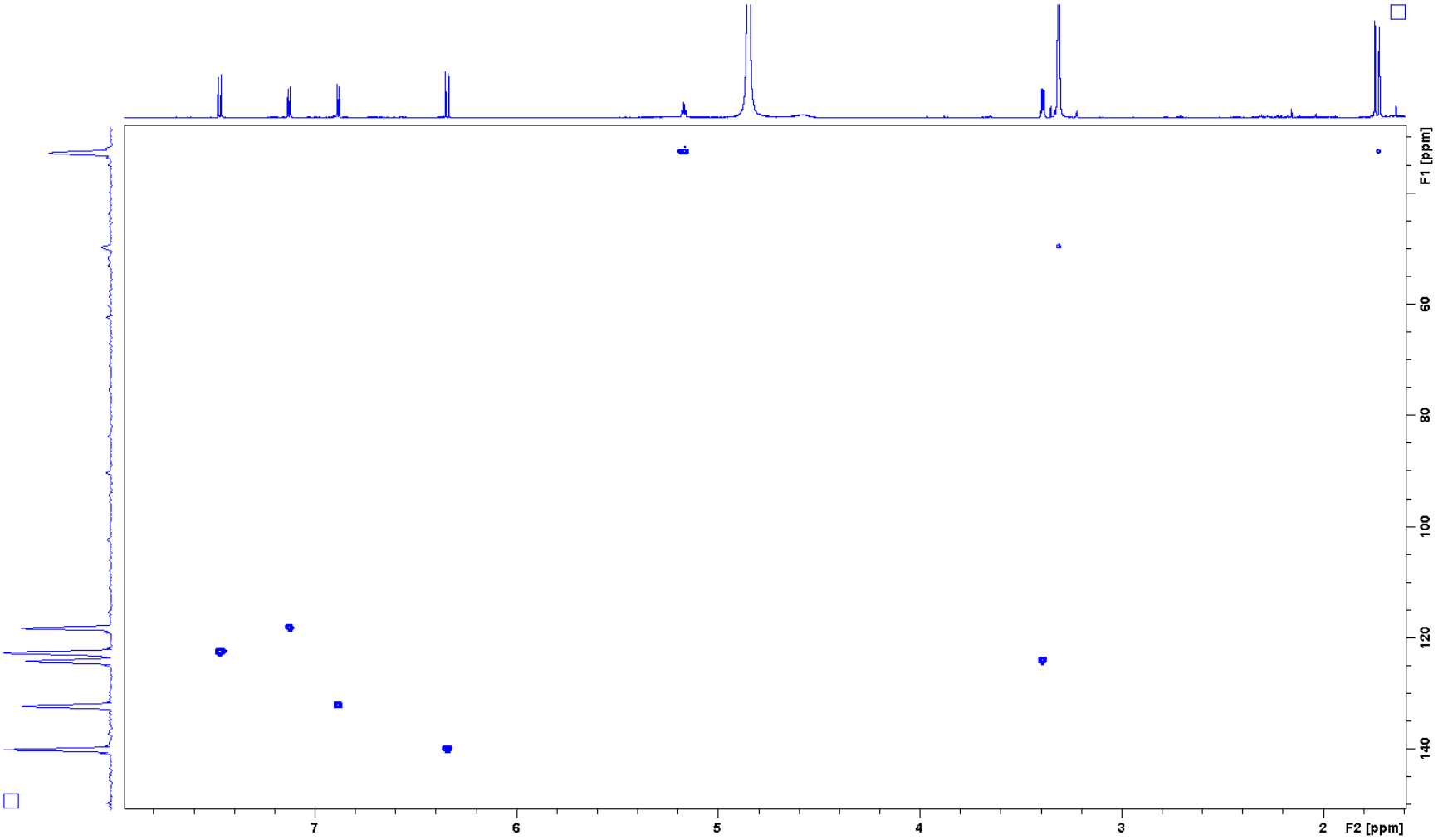


Figure S6. H2BC spectrum of **1**.

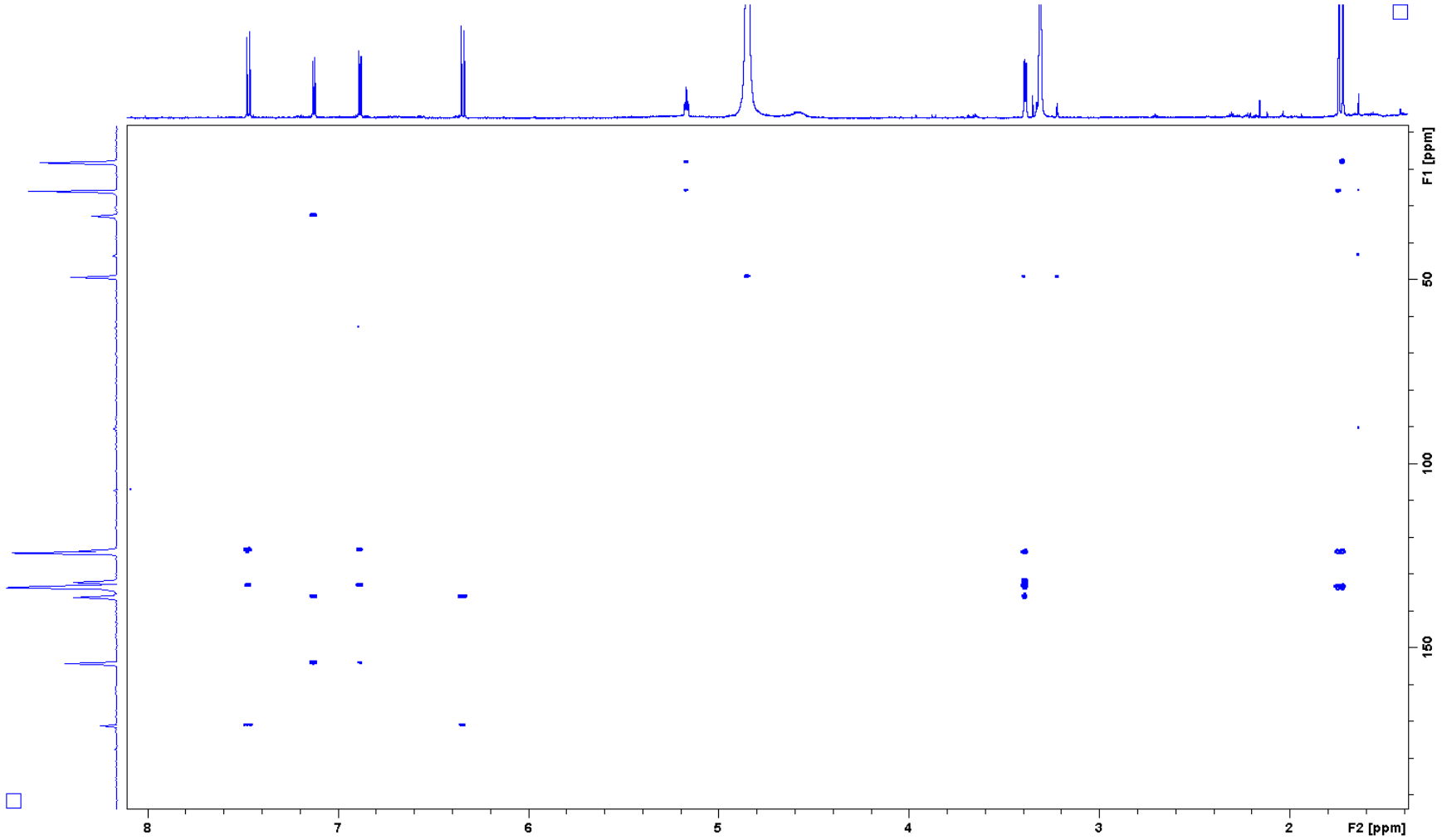


Figure S7. HMBC spectrum of **1**.

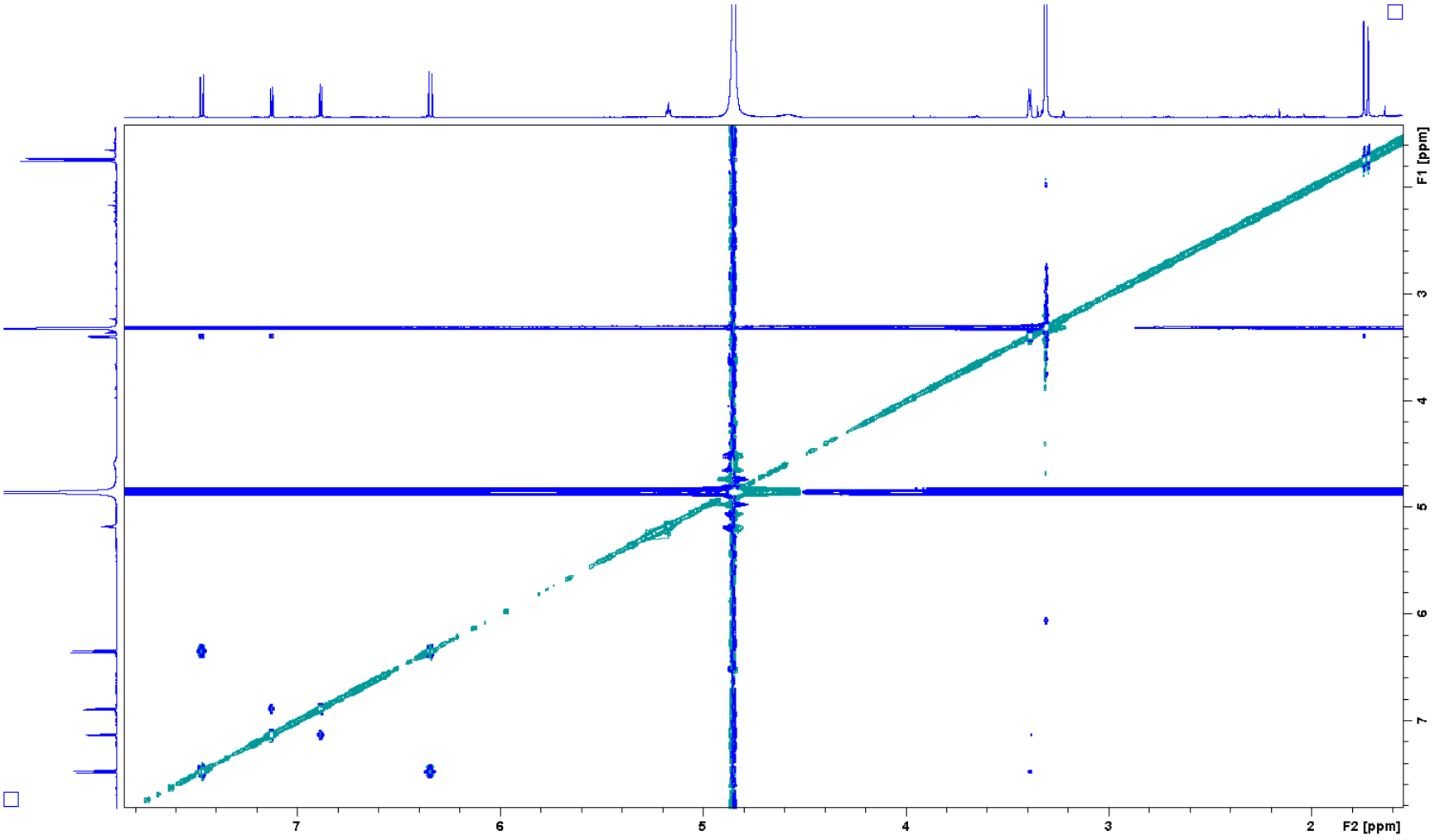


Figure S8. NOESY spectrum of **1**.

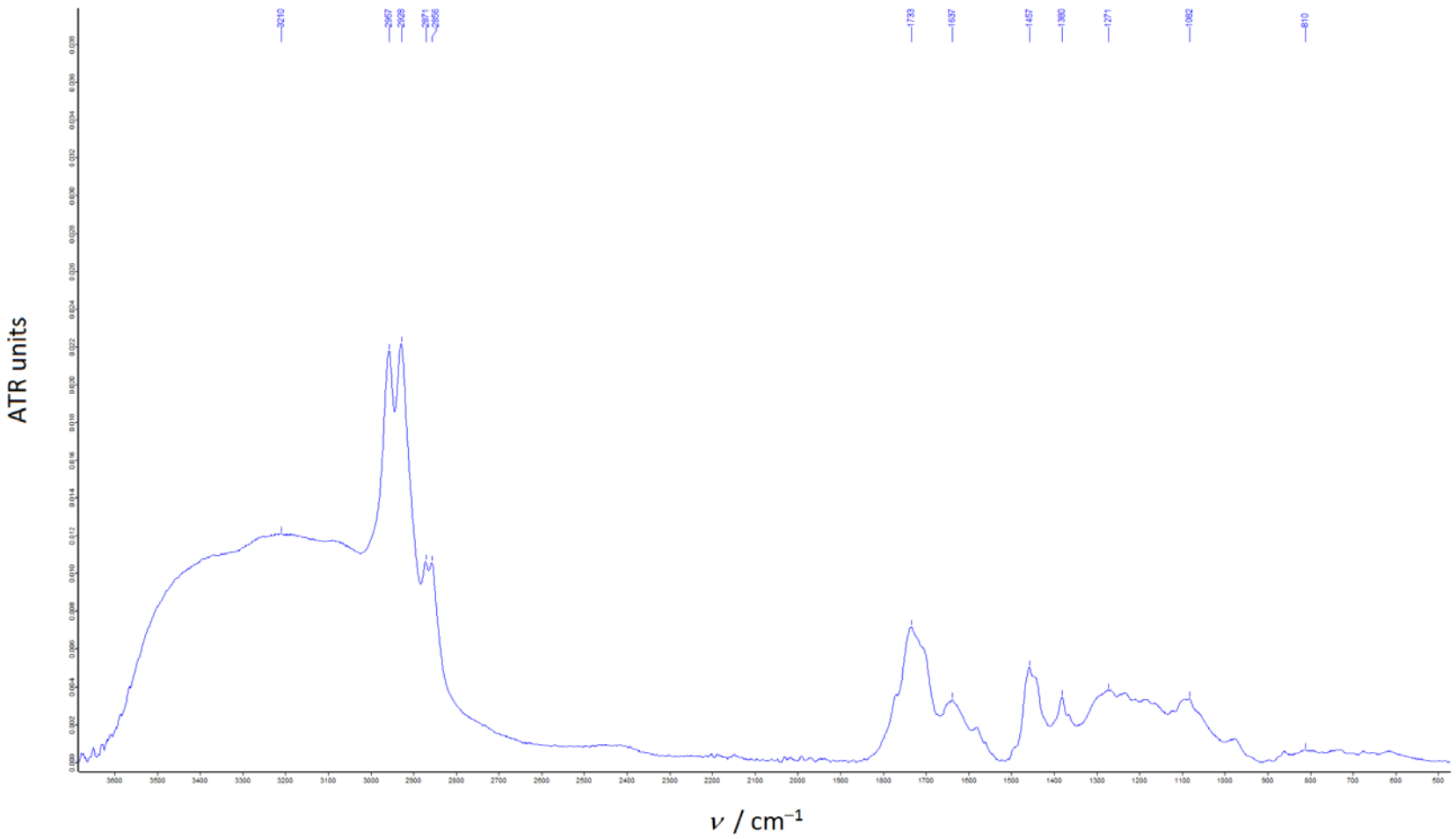


Figure S9. IR spectrum of **2**.

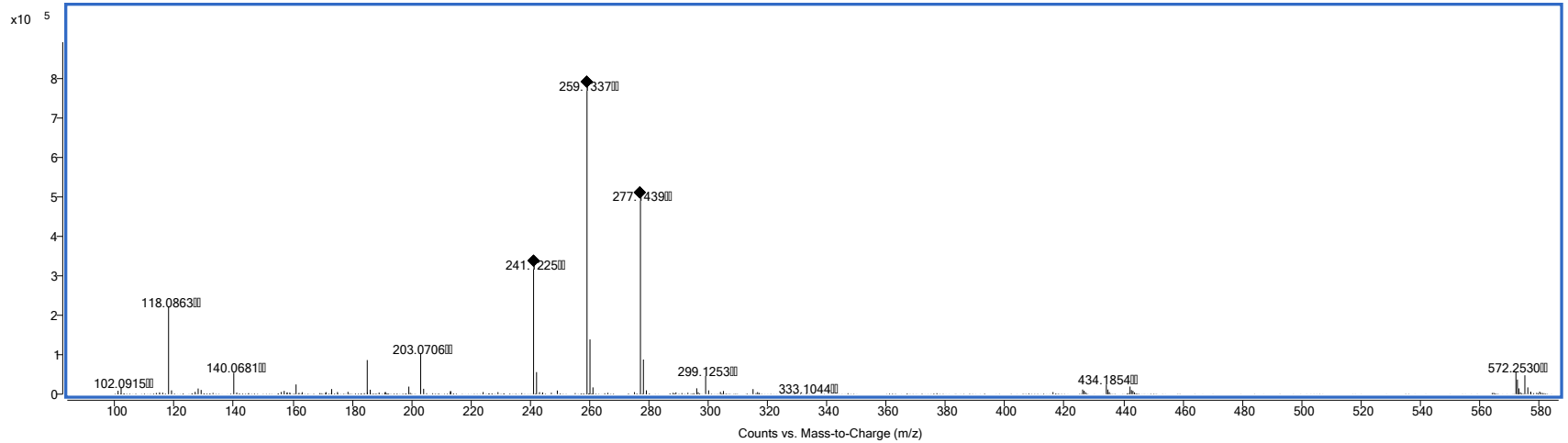


Figure S10. HRMS (ESI^+) spectrum of **2**.

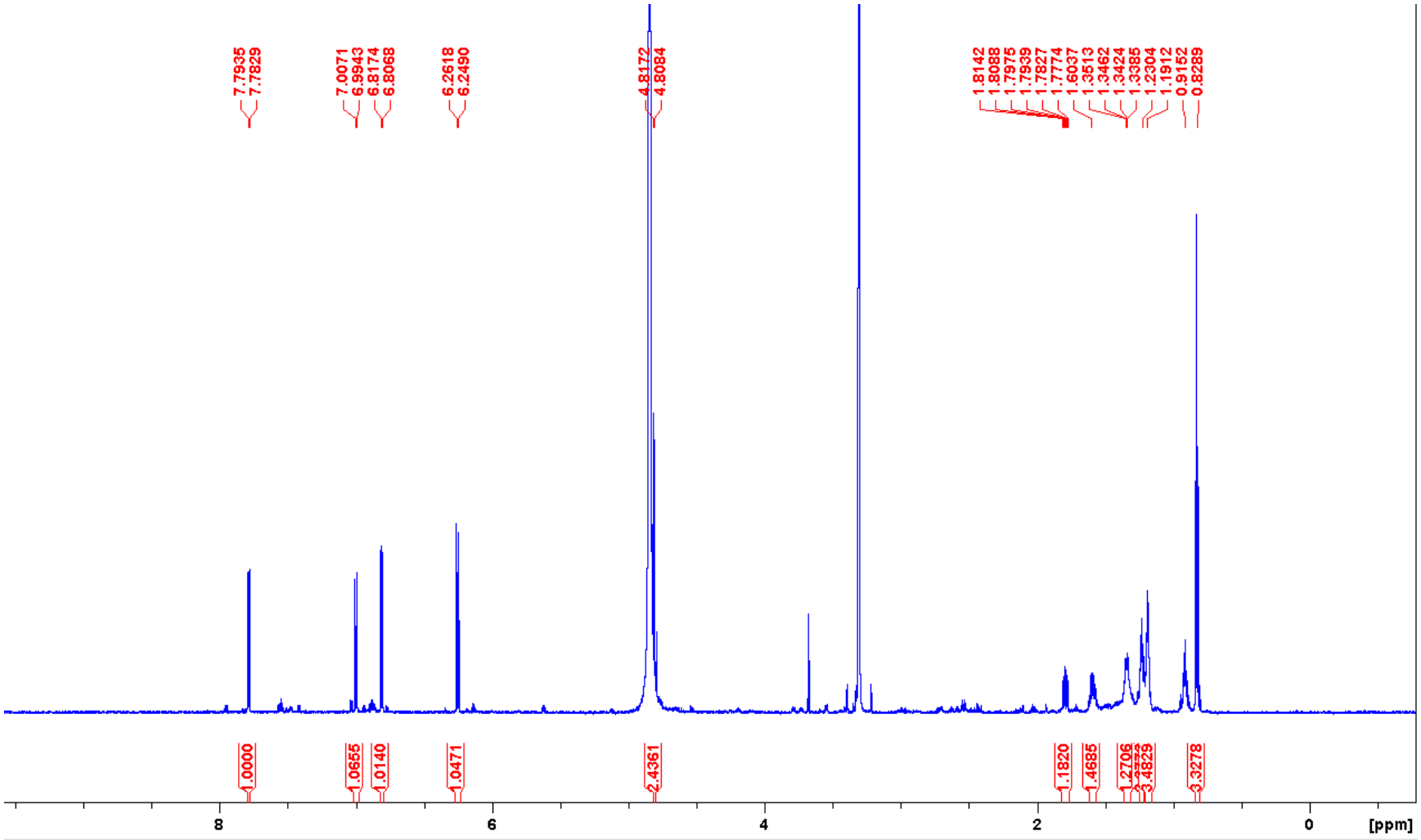


Figure S11. ^{1}H NMR spectrum of **2** (800 MHz, CD_3OD).

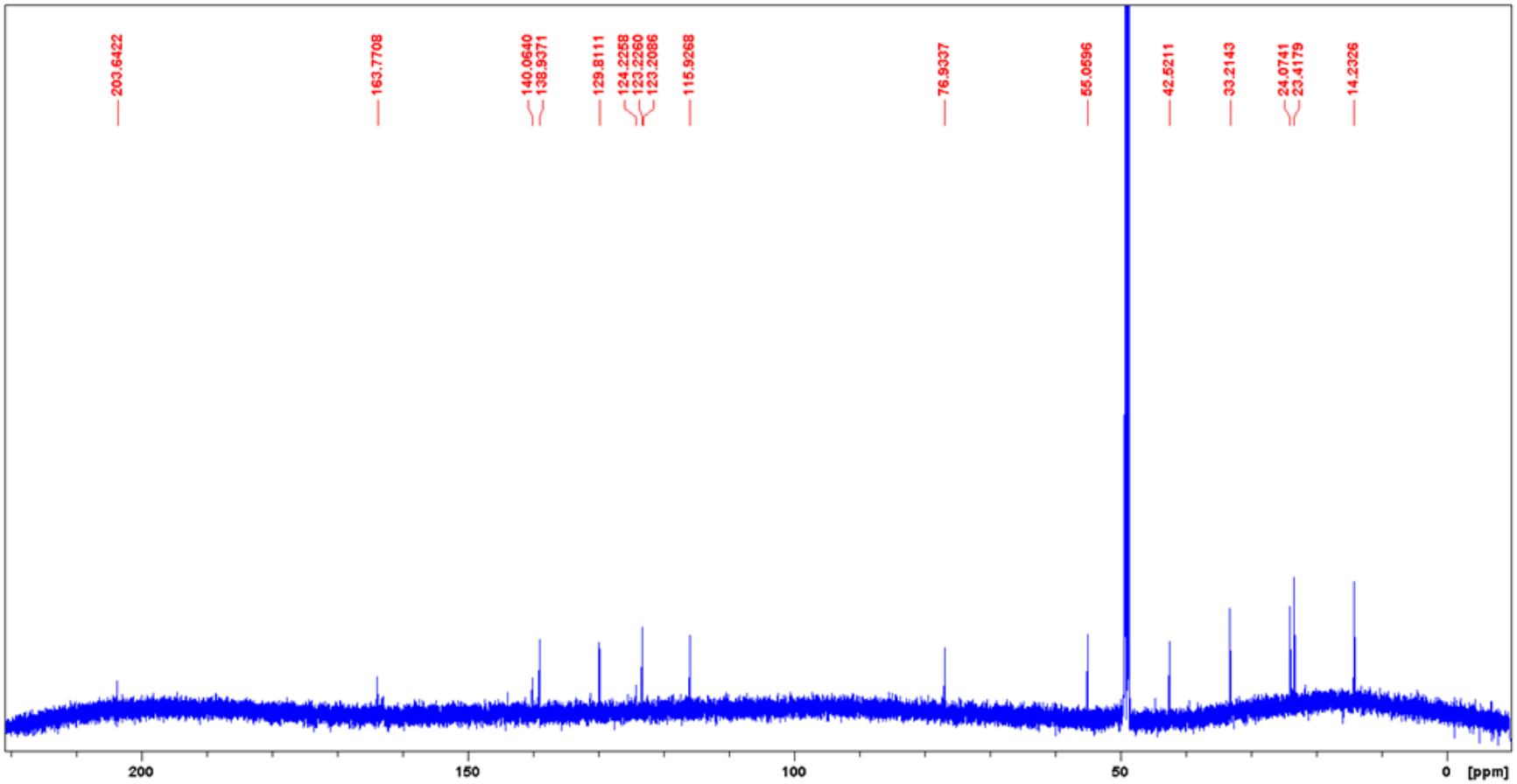


Figure S12. ^{13}C NMR spectrum of **2** (200 MHz, CD_3OD).

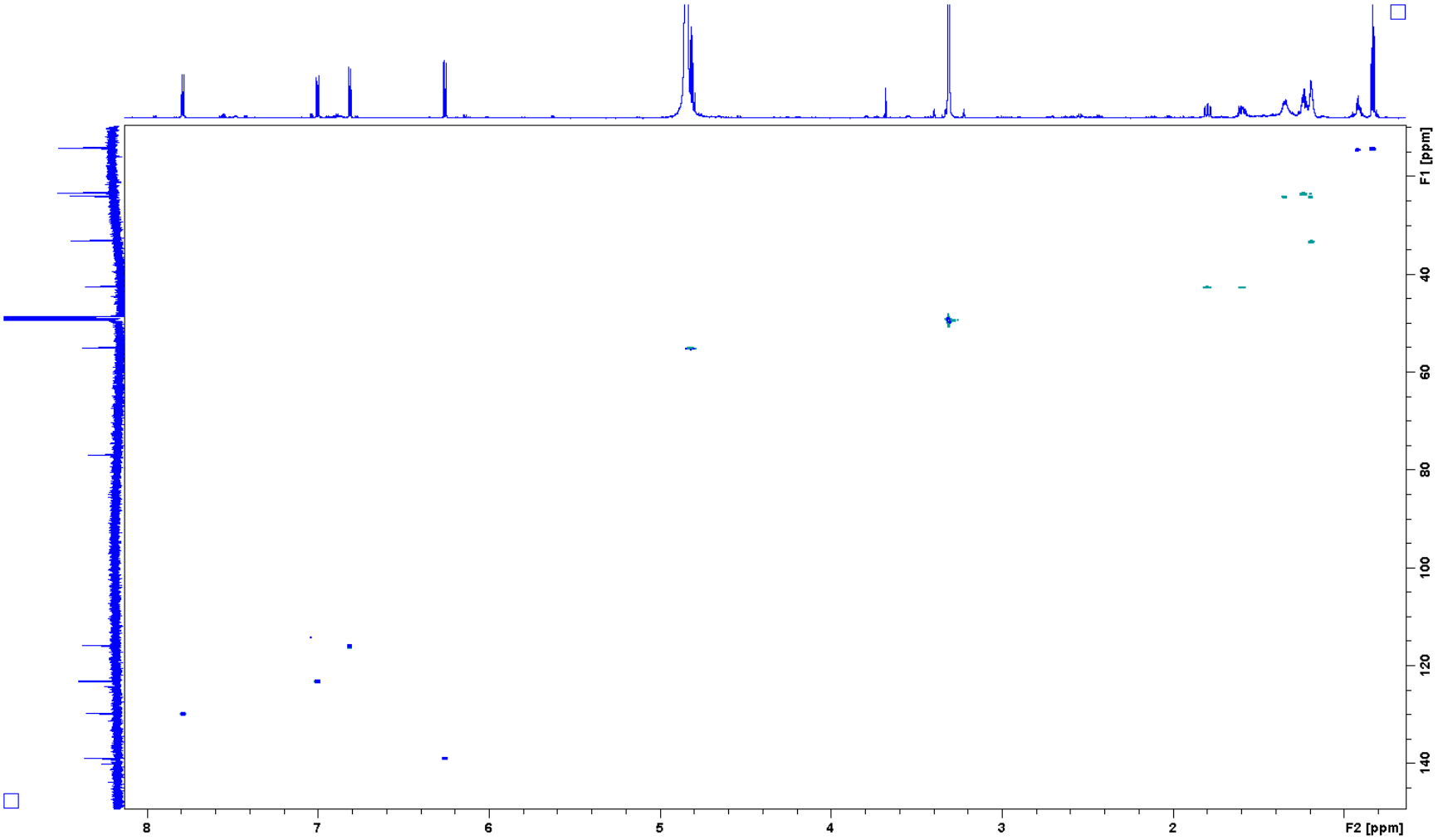


Figure S13. HSQC spectrum of **2**.

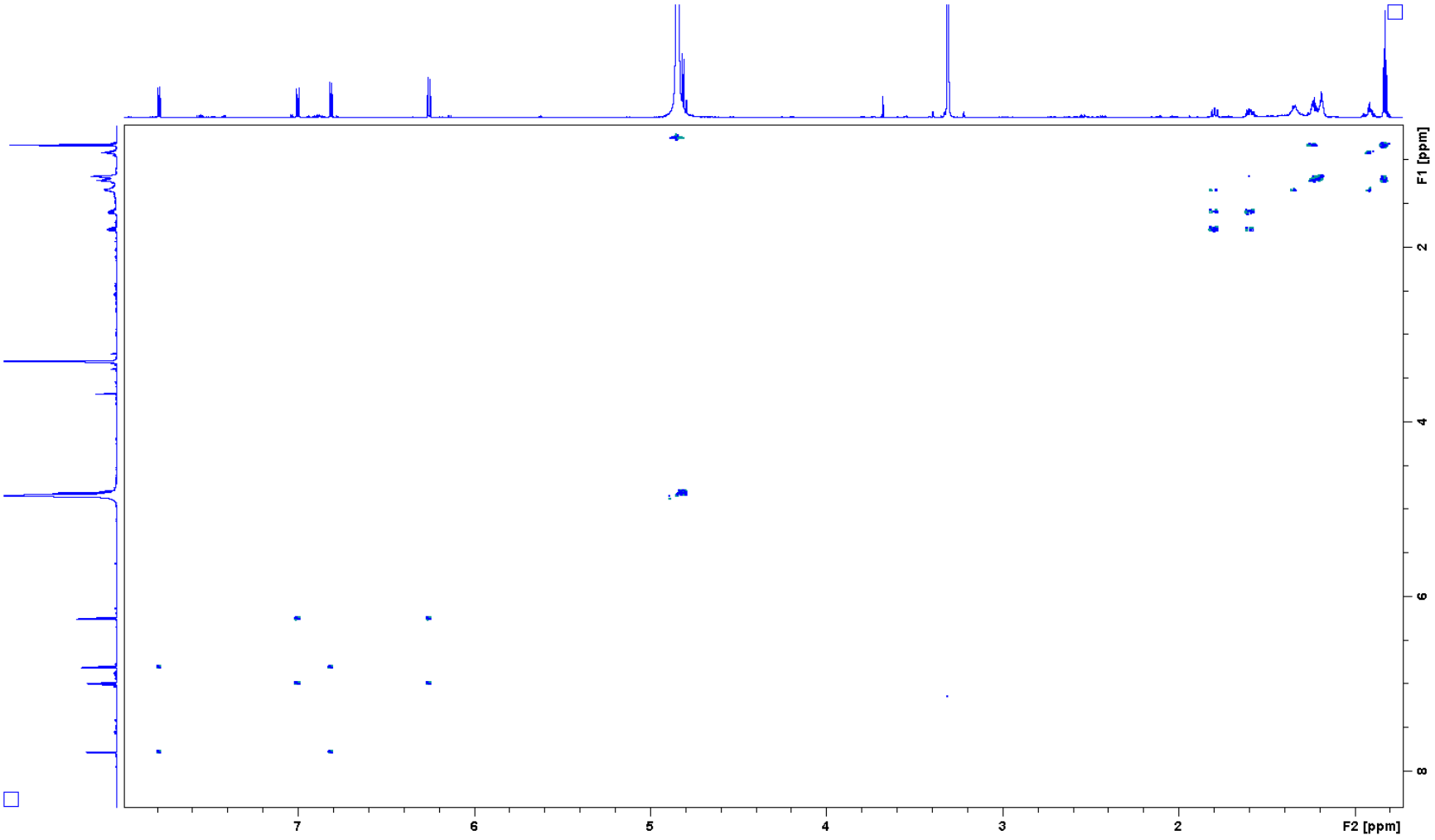


Figure S14. COSY spectrum of **2**.

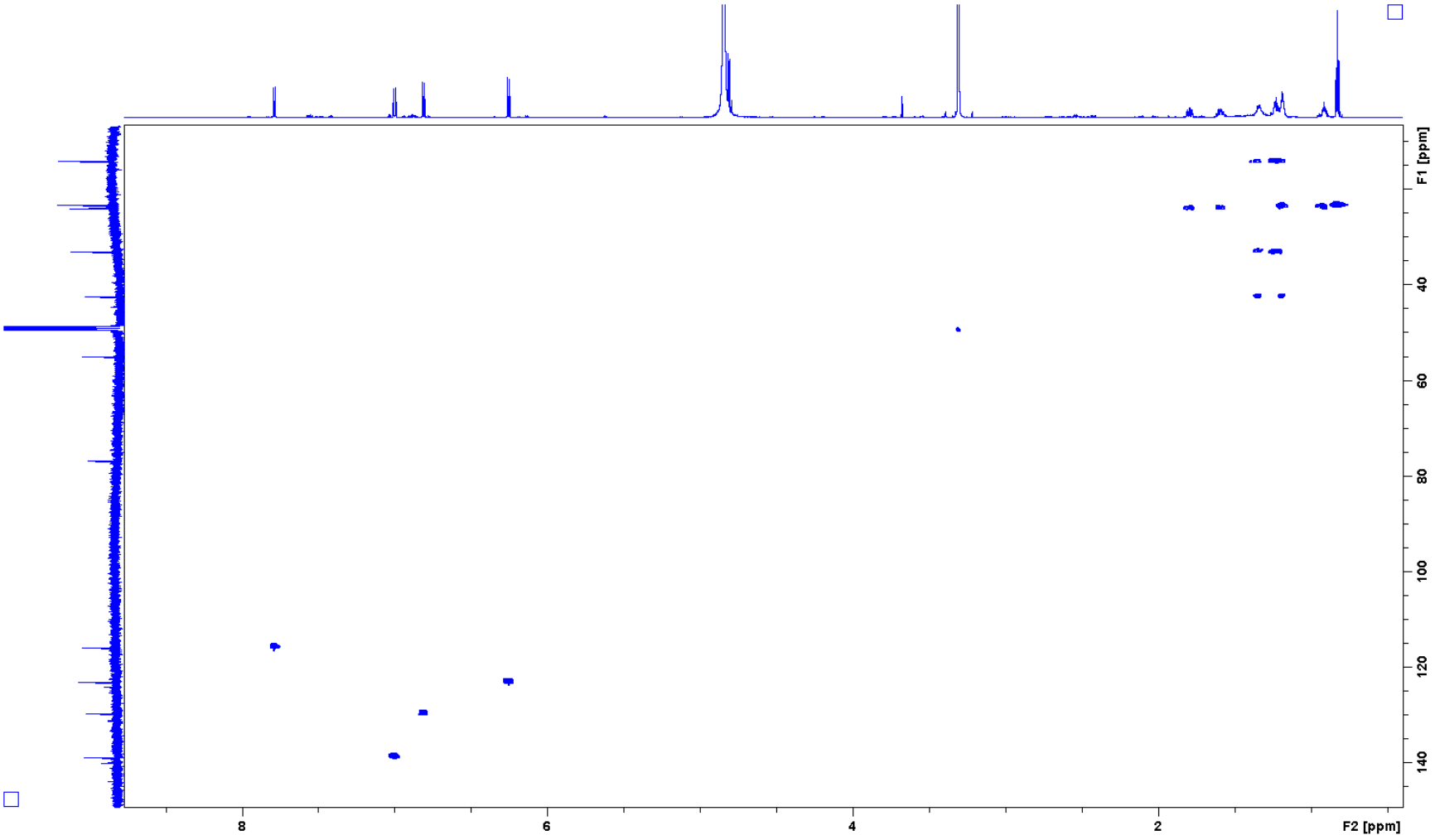


Figure S15. H2BC spectrum of **2**.

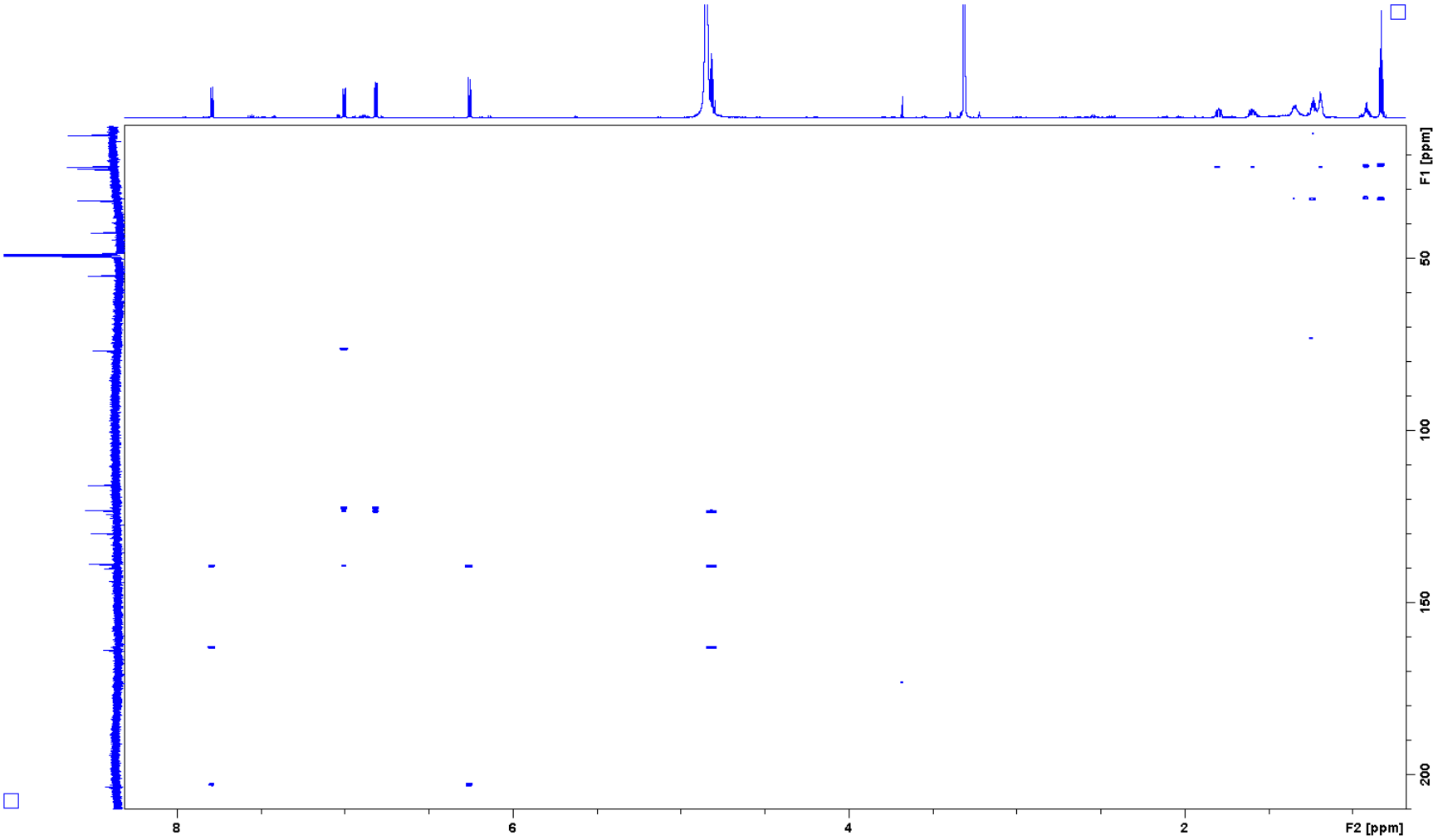


Figure S16. HMBC spectrum of **2**.

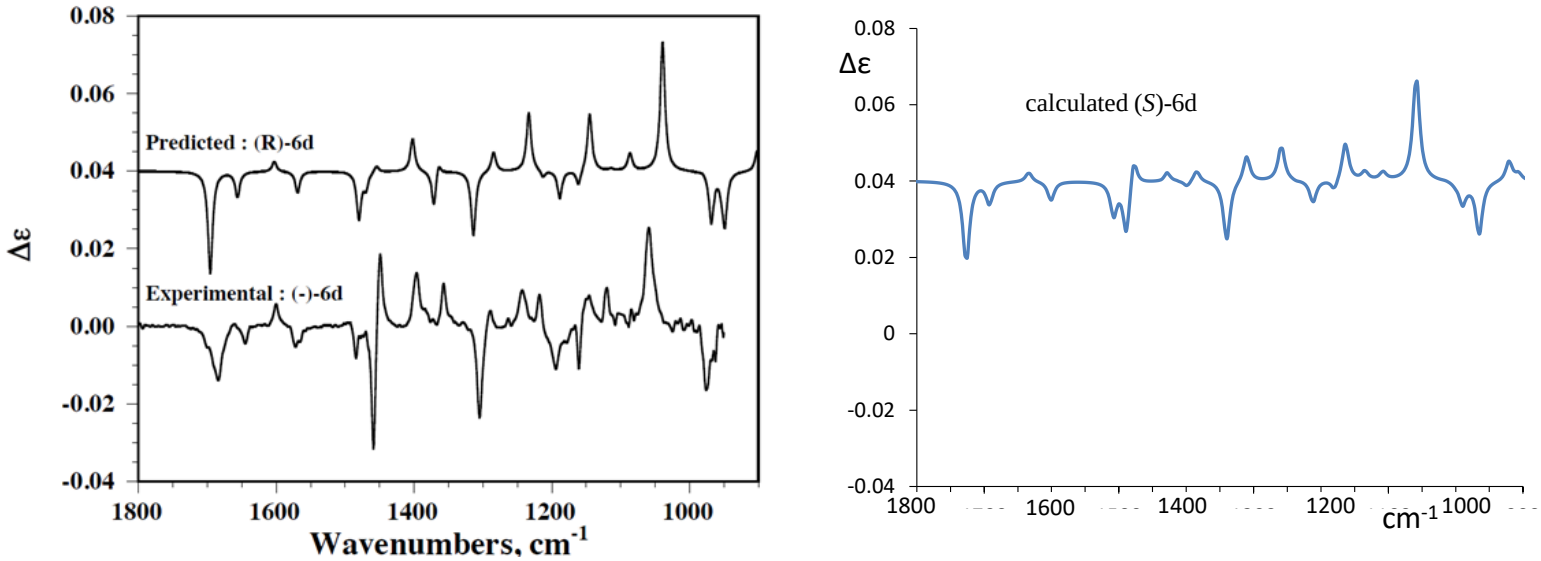


Figure S17. *Left:* Experimental (-)-6d (6d = *ortho*-quinol) and predicted (*R*)-6d VCD spectra adapted from reference [1]; *Right*: Calculated (DFT B3LYP/6-311G*) VCD spectrum of (*S*)-6d carried out in the present study.

Antibacterial and cytotoxic experimental procedures

Antibacterial assay. The strains used were: Methicillin-resistant *Staphylococcus aureus* MRSA MB5393 and methicillin sensitive *Staphylococcus aureus* MSSA ATCC 29213 as examples of Gram-positive bacteria model and *Escherichia coli* ATCC 25922 and *Klebsiella pneumoniae* ATCC 700603 as examples of Gram-negative bacteria model. Both compounds were dissolved in DMSO to a stock solution of 6 mg/mL and were serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 96 µg/mL for all the antibacterial assays. For MRSA, MSSA, *K. pneumoniae* and *E. coli* liquid assays 90 µL of the appropriate diluted inocula in BHI (Brain Heart Infusion), MHII (Muller Hinton II) or LB (Luria Broth) medium are mixed with 1.6 µL of compound stock solution at 6 mg/mL plus 8.4 µL of the appropriate medium depending on the microorganism to be tested. Positive control compounds included in the assays: Aztreonan (2-0.25 µg/mL) was used as positive control *E. coli* ATCC 25922; Gentamycin (32-0.25 µg/mL) was used for *K. pneumonia*; Vancomycin (32-0.25 µg/mL) was used as positive control for MRSA, MSSA.

For colorimetric assays absorbance at OD612nm was measured at zero time (T0) and at final time (Tf). Absorbance was measured in an Envision reader (Perkin Elmer). Percentage of growth inhibition was calculated using the following normalization: % Inhibition = 100 × {1 − [(Tf Sample − T0 Sample) − (Tf Blank − T0 Blank)]/[(Tf Growth − T0 Growth) − (Tf Blank − T0 Blank)]}

The *Genedata Screener* software (Genedata, Inc., Basel, Switzerland) was used to process and analyze the data. The activity of the compounds was expressed as a percentage of inhibition growth, where (100%) represented the inhibition growth of the target microorganisms and 0%

represented the total growth of target microorganisms on the assay. The Z′ factor predicts the robustness of an assay by considering the mean and standard deviation of both positive and negative controls.[2–4] The robust Z′ factor (RZ′ factor) is based on the Z′ factor, but standard deviations and means were replaced by the robust standard deviations and medians, respectively. In all experiments performed in this work, the RZ′ factor obtained was higher than 0.90 in all cases.

Cytotoxicity assay. The tumor cell lines used were A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast) and MiaPaca-2 (pancreas) and HL-60 (leukemia). Both compounds were assayed for cytotoxicity from 30 mg/ml in 10 points of dilutions 1:2, per triplicated. Data obtained were analyzed using Genedata Screener Software and IC50 was calculated. Cytotoxicities against A549, A2058, HepG2, MCF-7 and MiaPaca-2 were assayed using MTT test. Cells were seeded at 10 000 cells/well in a 96-well plate. After 24 h, cells were treated with compounds for 72 h. MMS (methyl methanesulfonate) 4mM was used as positive control of cell death. Then, MTT was added for 2 h and then absorbance was measured at 570 nm; Cytotoxicities against HL-60 was assayed using CCK-8 test. Cells were seeded at 5 000 cells/well in a 384-well plate. After 24 h, cells were treated with compounds for 72 h. MMS 4mM was used as positive control of cell death. Then, WST-8 was added for 2 h and then absorbance was measured at 450 nm.[2–4]

Table S1 Antibacterial assay results of **1** and **3**

Compounds	MICs (μg/mL)			
	MRSA MB5393	*MSSA* ATCC 29213	*E. coli* ATCC 25922	*K. pneumoniae* ATCC 700603
1	>96	>96	>96	>96
3	>96	>96	>96	>96

Table S2 Cytotoxic assay results of **1** and **3**

Compounds	IC_{50} (μg/mL)					
	A549	A2058	HepG2	MCF-7	MiaPaca-2	HL-60
1	>30	>30	>30	>30	>30	18
3	>30	>30	>30	>30	>30	24

Supplementary references

1. Quideau S, Lyvinec G, Marguerit M, Bathany K, Ozanne-Beaudenon A, Buffeteau T, et al. Asymmetrie hydroxylative phenol dearomatization through in situ generation of iodanes from chiral iodoarenes and m-CPBA. Angew Chemie - Int Ed. 2009;48:4605–4609.

2. Zhang JH, Chung TDY, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen. 1999;4:67–73.

3. Anne K, Hanspeter G, Patricia G, Martin B, Daniela G, Christian NP. Integration of Multiple Readouts Into the Z' Factor for Assay Quality Assessment. J Biomol Screen. 2010;15:95–101.

4. Audoin C, Bonhomme D, Ivanisevic J, De La Cruz M, Cautain B, Monteiro MC, et al. Balibalosides, an original family of glucosylated sesterterpenes produced by the Mediterranean sponge *Oscarella balibaloi*. Mar Drugs. 2013;11:1477–1489.

Oxepinamides L and M, two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides from *Aspergillus californicus*

Yaojie Guo[a], Simone Ghidinelli[b], Mercedes de la Cruz[c], Thomas A. Mackenzie[c], Maria C. Ramos[c], Pilar Sánchez[c], Francisca Vicente[c], Olga Genilloud[c] & Thomas O. Larsen[a]*

[a] *Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark;* [b] *Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy;* [c] *Fundación MEDINA, Granada, Spain.*

*Tel: +45 45252632. Fax: +45 45884922. E-mail: tol@bio.dtu.dk.

Oxepinamides L and M, two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides from *Aspergillus californicus*

ABSTRACT

A chemical investigation of *Aspergillus californicus* IBT 16748 led to the isolation of two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides oxepinamides L (**1**) and M (**2**). Their structures were characterised by spectroscopic analysis including HRESIMS, 1D and 2D NMR. The absolute structure of **1** was assigned by ECD calculation. The antibacterial and cytotoxic properties of **1** were evaluated.

KEYWORDS *Aspergillus californicus*; oxepine; nonribosomal peptides; oxepinamide

1. Introduction

Oxepine-pyrimidinone-ketopiperazine (OPK) type nonribosomal peptides (NRPs) are a small class of rare natural products that up to now have only been isolated from fungal sources and in particular from *Aspergillus* species (Guo et al. 2020). *Aspergillus californicus* was first isolated from Chamise chaparral (*Adenostoma fasciculatum*) soil in California over forty years ago. This poor sporulating species is assigned to section *Cavernicolarum* but it resembles typical section *Usti* species as it produces long light brown conidiophores (Samson et al. 2011; Chen et al. 2016). During the ongoing discovery of novel natural products from filamentous fungi in genus *Aspergillus*, *A. californicus* IBT 16748 showed a high potential to produce many previously unknown compounds using a One Strain-Many Compounds (OSMAC) strategy (Bode et al. 2002). In this paper, we report the isolation and characterisation of the two previously unknown OPK NRPs oxepinamides L (**1**) and M (**2**)

(**Figure 1)** as well as the antibacterial and cytotoxic evaluation of **1**.

Figure 1. Structures and selected HMBC, COSY and NOE correlations of **1** and **2**.

2. Results and Discussion

The molecular formula of oxepinamide L (**1**) was established as $C_{23}H_{23}N_3O_3$ with 14 degrees of unsaturation based on a $[M + H]^+$ ion *m/z* 390.1816 and $[M + Na]^+$ *m/z* 412.1629. ^{1}H NMR data **(Table 1)** showed 23 hydrogens including two methyl groups δ_H 0.88 (d, J = 6.3 Hz, H_3-18), 1.01 (d, J = 6.3 Hz, H_3-19), one nitrogenated methyl group δ_H 4.05 (s, H_3-21), one nitrogenated methine δ_H 5.40 (dd, 8.2, 5.0, H-15), five aromatic hydrogens δ_H 8.09 (d, J = 7.4 Hz, H-2′/H-6′), 7.41 (t, J = 7.4 Hz, H-3′/H-5′), 7.35 (t, J = 7.4 Hz, H-4′), four adjacent olefinic protons δ_H 6.14 (d, J = 5.6 Hz, H-8), 5.72 (t, J = 5.6 Hz, H-9), 6.17 (dd, J = 11.1, 5.6 Hz, H-10) and 6.70 (d, J = 11.1 Hz, H-11). The ^{13}C NMR spectrum showed 20 signals among which δ_C 129.5 was proved to be two overlapping carbons C-3′/5′ and δ_C 132.7 was C-2′/6′ by HSQC. The missing carbon δ_C 130.8 (C-3) was assigned by the weak HMBC correlation from H-2′/H-6′ (**Figure S8)**. All the NMR data showed high similarity to a known compound oxepinamide K (Liang et al. 2019). One major difference was that C-15 in **1** was substituted by an isobutyl group instead of a methyl group in oxepinamide K. The second difference was the methyl substitution on N-2 in **1**. These two differences were supported by the COSY correlations of one spin system H-15 through H_3-18/19 and the HMBC correlation from H_3-21 to δ_C 165.3 (C-1). The double bond between C-3 and C-20 was determined to be *Z* configuration based on the NOE correlations between H_3-21 and H-6′. The absolute configuration of the chiral center C-15 was elucidated by comparing the electronic circular dichroism (ECD) spectrum calculated by time-dependent density functional theory (TDDFT)

formalism (see Computation Details) with the measured data (Mazzeo et al. 2013; Gennaro et al. 2016; Polavarapu 2016). The predicted ECD spectrum of (15*S*)-**1** matched the experimental curve (**Figure S10)**. Thus, C-15 was established as *S*-configuration indicating an *L*-leucine residue.

Table 1. NMR data (800 MHz ^{1}H; 200 MHz ^{13}C) of **1** and **2** in CD_3OD

It´s notable that the oxepinamide K was reported as a likely artificial product that was stable under acidic conditions. In the present study, **1** appeared as a major peak in the UHPLC-MS profiles of *A. californicus* inoculated on Czapeck Yeast extract Agar (CYA) and several other solid media plates (Samson et al. 2010) (**Figure S12**). Besides, several other OPK NRPs analogues also contain a double bond between the α and β carbons of an amino acid residue such as cinereain, janoxepin, protuboxepins F and G (Cutler et al. 1988; Sprogøe et al. 2005; Dmitrenok 2013; Luo et al. 2019). Therefore **1** was deduced to be a real natural product.

The molecular formula of oxepinamide M (**2**) was assigned as $C_{22}H_{21}N_3O_3$ by HRESIMS. Both the ^{1}H and ^{13}C NMR data showed a high similarity with those of **1** except one obvious difference that the methyl group δ_H 4.05 was missing in the ^{1}H NMR spectrum of **2**. The number of hydrogen signals in the ^{1}H NMR spectrum was 20 indicating the existence of an exchangeable proton which was determined to be on N-2. This compound decomposed during the trial to measure the NOE correlations between H-2 and the aromatic hydrogens in DMSO-d_6. Thus, the stereochemistry of **2** was proposed to be the same as **1** considering that both exhibited positive optical rotations in methanol and they likely originated from the same biosynthetic pathway. Thus, the structure of **2** was assigned.

Compound **1** was subjected to antibacterial and cytotoxic screening. In the antibacterial test, **1** didn´t show activities against two Gram-positive bacteria (methicillin-resistant *Staphylococcus aureus* MB5393 and methicillin-sensitive *Staphylococcus aureus* ATCC 29213), and two Gram-negative bacteria (*Escherichia coli* ATCC 25922 and *Klebsiella pneumoniae* ATCC 700603) with the minimum inhibitory concentration (MIC) values over 96 μg/ml. Six tumor cell lines A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast), Mia PaCa-2 (pancreas), HL-60 (leukemia) were used in the cytotoxicity test. **1** didn´t show promising activities against the six cell lines as the IC_{50} values were above 30 μg/ml.

3. Experimental

3.1 General Experimental Procedures

Optical rotations were measured on a PerkinElmer 341 polarimeter. ECD spectra were recorded in acetonitrile with a 2 mm path length cuvette on a JASCO J-1500 CD spectrophotometer. The infrared attenuated-total-reflectance (ATR) spectrum was collected on a Bruker VERTEX 80v Fourier Transform vacuum spectrometer employing a single-reflection diamond ATR accessory. The NMR spectra were recorded on Bruker AVANCE III 800 MHz spectrometer equipped with a 5 mm TCI cryoprobe using standard pulse sequences. Chemical shifts were reported in ppm with reference to the solvent signals (CD_3OD δ_H 3.31 and δ_C 49.00). HRESIMS data of two compounds were acquired on a Dionex Ultimate 3000 UHPLC system (Thermo Scientific) coupled to a maXis 3G QTOF orthogonal mass spectrometer (Bruker Daltonics). Chemical analysis of the plug extractions was performed on an Agilent 1290-6545 UHPLC-QTOF-MS system with an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 150 mm, 1.9 μm) at 60 °C with a linear gradient program: mobile phase (MP) A H_2O and B MeCN, both buffered with 20 mM formic acid

(FA), at a flow rate of 0.35 mL/min: 0-10.0 min 10%-100% MPB, 10.0-12.0 min 100% MPB, 12.1-14.0 min 10% MPB. Flash chromatography was carried out on a Biotage Isolera One system. Semi-preparative separation was performed on Waters 600 HPLC with a Photodiode Array detector. The HPLC columns used in the isolation were a Phenomenex Kinetex C_{18} column (250 × 10 mm, 5 μm) and a Phenomenex Luna phenyl-hexyl column (250 × 10 mm, 5 μm). For extraction and separation, the solvents were HPLC grade and for HRMS analysis the chemicals were LC-MS grade.

3.2 Fungal Material

A. californicus IBT 16748 was from IBT culture collection, Department of Biotechnology and Biomedicine, Technical University of Denmark.

3.3 Extraction and Isolation.

A. californicus IBT16748 was cultivated on 120 CYA plates and 120 YES plates for 11 days at 25°C. All the plates were extracted twice with ethyl acetate with 1% FA. The crude extract was dissolved in 90% MeOH, extracted with heptane, and further diluted to 50% MeOH before being extracted with dichloromethane (DCM). The DCM part (1.266 g) was applied to the flash chromatography with a self-packed NH_2 column (Sepra Phenomenex, 50g, CV 66 mL) to yield eight fractions. Fr1 (24.5 mg) was subjected to semi-preparative HPLC with a C_{18} column (MeCN/H_2O, 70-85% in 20 min, 4 mL/min) to yield **1** (0.71 mg, t_R = 12.6 min). Fr3 (10.8 mg) was also subjected to HPLC with a phenyl-hexyl column (MeCN/H_2O, 60-85% in 17 min, 4 mL/min) to yield four sub-fractions among which sub-fr3 (0.34 mg) was purified with a C_{18} column (MeCN/H_2O, 65% isocratic, 4 mL/min) to yield **2** (0.27 mg, t_R = 7.8 min).

Oxepinamide L (**1**): yellow amorphous solid; $[\alpha]_D^{20}$ = +35.2 (c = 0.071, MeOH); HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 386 nm; ECD (0.535 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 420 (-0.5), 318 (+0.76), 255 (+3.6), 212 (+5.1) nm; IR (dry film) ν_{max} 3355, 2953, 1681, 1508, 1445, 1363, 1212, 1141, 847, 804, 726 cm^{-1}; ^1H and ^{13}C data see **Table 1**. HRESIMS *m/z* 390.1816 $[M+H]^+$ (calcd for $C_{23}H_{24}N_3O_3^+$, 390.1812), 412.1629 $[M+Na]^+$ (calcd for $C_{23}H_{23}N_3O_3Na^+$, 412.1632).

Oxepinamide M (**2**): yellow amorphous solid; $[\alpha]_D^{20}$ = +25.9 (*c* = 0.027, MeOH); HPLC-UV (MeCN/H_2O + 20 mM FA) λ_{max} 384 nm; ^1H and ^{13}C data see **Table 1**. HRESIMS *m/z* 376.1657 $[M+H]^+$ (calcd for $C_{22}H_{22}N_3O_3^+$ 376.1656), 398.1475 $[M+Na]^+$ (calcd for $C_{22}H_{21}N_3O_3Na^+$, 398.1475).

3.4 Computational Details

Conformational analysis was performed at the Molecular Mechanics (MM) level, considering all conformers in the range of 5 kcal/mol from the most stable one using Maestro software. Geometry optimization was carried out at B3LYP/TZVP within the PCM approximation in the Gaussian 16 package (Tomasi et al. 2005; Frisch et al. 2016). ECD calculated spectra were obtained by Gaussian 16 using TDDFT CAM-B3LYP/TZVP level of theory. 0.2 eV wide Gaussian bands were used to simulate the spectra.

4. Conclusions

In conclusion, two OPK NRPS oxepinamides L (**1**) and M (**2**) were characterised from *A. californicus* IBT 16748. The structure of **1** was determined by HRMS, 1D and 2D NMR, and ECD calculation. To the best of our knowledge, **1** is the first OPK compound that has been methylated on N-2. Compound **2** was a demethylated form of **1**. Though no promising

antibacterial or cytotoxic activities were observed in the bioassays in this study, **1** and **2** might be active in activating the liver X receptor α (Lu et al. 2011; Liang et al. 2019).

Acknowledgments

The authors thank Dr. Kasper Enemark-Rasmussen and Associate Professor René Wugt Larsen from the Department of Chemistry at the Technical University of Denmark for acquiring NMR and IR data, respectively. Emil H. Kristiansen is acknowledged for his work on strain cultivation and chemical extraction.

Disclosure statement

The authors declare no potential conflicts of interest.

Funding

This work was supported by the China Scholarship Council under grant number 201709110107.

References

Bode HB, Bethe B, Höfs R, Zeeck A. 2002. Big effects from small changes: possible ways to explore nature's chemical diversity. ChemBioChem. 3(7):619–627.

Chen AJ, Frisvad JC, Sun BD, Varga J, Kocsubé S, Dijksterhuis J, Kim DH, Hong SB, Houbraken J, Samson RA. 2016. *Aspergillus* section *Nidulantes* (formerly *Emericella*): polyphasic taxonomy, chemistry and biology. Stud Mycol. 84:1–118.

Cutler HG, Springer JP, Arrendale RF, Arison BH, Cole PD, Roberts RG. 1988. Cinereain: a novel metabolite with plant growth regulating properties from *Botrytis cinerea*. Agric Biol Chem. 52(7):1725–1733.

Olesya IZ, Shamil SA, Ekaterina AY, Vladimir AD, Natalya NK, Pavel SD. 2013. New metabolites from the algal associated marine-derived fungus *Aspergillus carneus*. Nat Prod Commun. 8(8):1071–1074.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, et al. 2016. Gaussian 16.

Gennaro P, Torsten B. 2016. Good computational practice in the assignment of absolute configurations by TDDFT calculations of ECD Spectra. Chirality. 28(6):466–474.

Guo Y, Frisvad JC, Larsen TO. 2020. Review of oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides. Metabolites. 10(6):246.

Liang X, Zhang X, Lu X, Zheng Z, Ma X, Qi S. 2019. Diketopiperazine-type alkaloids from a deep-sea-derived *Aspergillus puniceu*s fungus and their effects on liver X receptor α. J Nat Prod. 82(6):1558–1564.

Lu XH, Shi QW, Zheng ZH, Ke AB, Zhang H, Huo CH, Ma Y, Ren X, Li YY, Lin J, et al. 2011. Oxepinamides: novel liver X receptor agonists from *Aspergillus puniceus*. European J Org Chem. 2011(4):802–807.

Luo X, Chen C, Tao H, Lin X, Yang B, Zhou X, Liu Y. 2019. Structurally diverse diketopiperazine alkaloids from the marine-derived fungus: *Aspergillus versicolor* SCSIO 41016. Org Chem Front. 6(6):736–740.

Mazzeo G, Santoro E, Andolfi A, Cimmino A, Troselj P, Petrovic AG, Superchi S, Evidente A, Berova N. 2013. Absolute configurations of fungal and plant metabolites by chiroptical methods. ORD, ECD, and VCD studies on phyllostin, scytolide, and oxysporone. J Nat Prod. 76(4):588–599.

Polavarapu PL. 2016. Determination of the absolute configurations of chiral drugs using chiroptical spectroscopy. Molecules. 21(8):1056.

Samson RA, Houbraken J, Thrane U, Frisvad JC, Andersen B. 2010. Food and indoor fungi: second edition. Utrecht, The Netherlands: CBS-KNAW Fungal Biodiversity Centre.

Samson RA, Varga J, Meijer M, Frisvad JC. 2011. New taxa in *Aspergillus* section *Usti*. Stud Mycol. 69:81–97.

Sprogøe K, Manniche S, Larsen TO, Christophersen C. 2005. Janoxepin and brevicompanine B: antiplasmodial metabolites from the fungus *Aspergillus janus*. Tetrahedron. 61(36):8718–8721.

Tomasi J, Mennucci B, Cammi R. 2005. Quantum mechanical continuum solvation models. Chem Rev. 105(8):2999–3093.

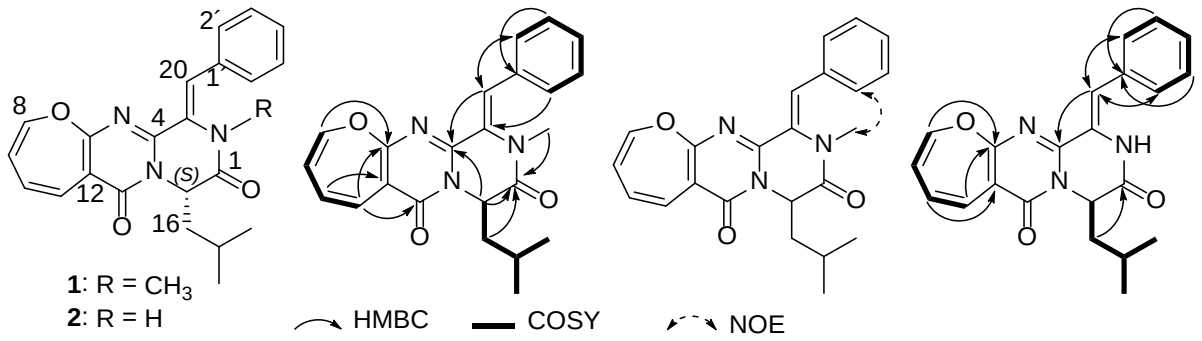


Figure 1. Structures and selected HMBC, COSY and NOE correlations of **1** and **2**.

Table 1. NMR data (800 MHz ^{1}H; 200 MHz ^{13}C) of **1** and **2** in CD_3OD

No.	**1**		**2**	
	$\delta_{C,\ type}$	$\delta_{H,}$ multi (*J* in Hz)	$\delta_{C}{}^{a,\ type}$	$\delta_{H,}$ multi (*J* in Hz)
1	165.3, C		167.6, C	
3	130.8[a], C		130.3, C	
4	153.3, C		151.9, C	
6	165.0, C		164.6, C	
8	144.4, CH	6.14, d (5.6)	144.4, CH	6.14, d (5.6)
9	118.4, CH	5.72, t (5.6)	118.5, CH	5.73, t (5.6)
10	128.7, CH	6.17, dd (11.1, 5.6)	129.1, CH	6.20, dd (11.1, 5.6)
11	126.2, CH	6.70, d (11.1)	126.3, CH	6.71, d (11.1)
12	110.8, C		111.5, C	
13	162.7, C		159.9, C	
15	52.1, CH	5.40, dd (8.2, 5.0)	56.0, CH	5.36, dd (8.5, 5.8)
16	43.4, CH_2	a, 1.72, m b, 1.65, m	43.5, CH_2	a, 1.82, m b, 1.69, m
17	26.1, CH	1.69, m	26.1, CH	1.81, m
18	23.6, CH_3	0.88, d (6.3)	23.5, CH_3	0.98, d (6.4)
19	22.3, CH_3	1.01, d (6.3)	22.2, CH_3	1.07, d (6.4)
20	127.8, CH	7.45, s	120.6, CH	7.33, s
21	55.3, CH_3	4.05, s		
1´	136.4 C		134.5, C	
2´, 6´	132.7, CH	8.09, d (7.4)	130.6, CH	7.59, d (7.6)
3´, 5´	129.5, CH	7.41, t (7.4)	130.1, CH	7.47, t (7.6)
4´	130.2, CH	7.35, t (7.4)	130.3, CH	7.38, t (7.6)

[a] carbon chemical shifts assigned by HMBC correlations.

SUPPLEMENTARY MATERIAL

Oxepinamides L and M, two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides from *Aspergillus californicus*

Yaojie Guo[a], Simone Ghidinelli[b], Mercedes de la Cruz[c], Thomas A. Mackenzie[c], Maria C. Ramos[c], Pilar Sánchez[c], Francisca Vicente[c], Olga Genilloud[c] & Thomas O. Larsen[a]*

[a] *Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark;* [b] *Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy;* [c] *Fundación MEDINA, Granada, Spain.*

*Tel: +45 45252632. Fax: +45 45884922. E-mail: tol@bio.dtu.dk.

ABSTRACT

A chemical investigation of *Aspergillus californicus* IBT 16748 led to the isolation of two new oxepine-pyrimidinone-ketopiperazine type nonribosomal peptides oxepinamides L (**1**) and M (**2**). Their structures were characterised by spectroscopic analysis including HRESIMS, 1D and 2D NMR. The absolute structure of **1** was assigned by ECD calculation. The antibacterial and cytotoxic properties of **1** were evaluated.

KEYWORDS *Aspergillus californicus*; oxepine; nonribosomal peptides; oxepinamide

Contents

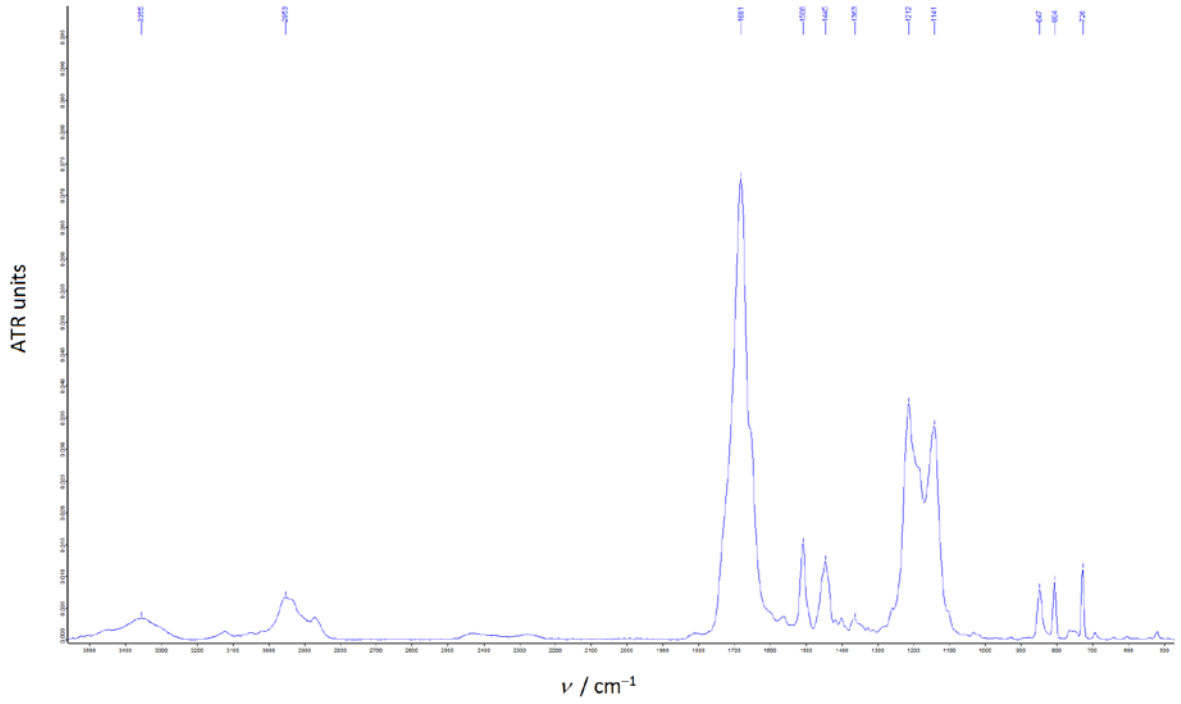


Figure S1. IR spectrum of **1**.

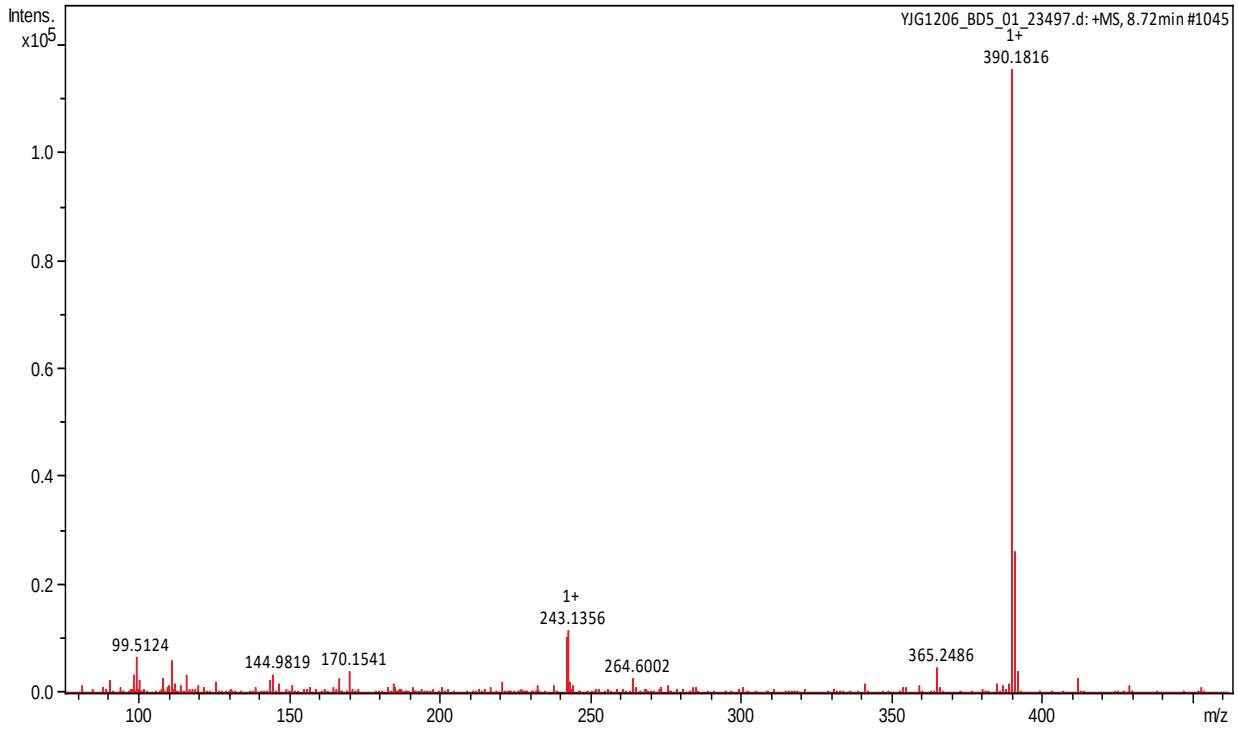


Figure S2. HRESIMS spectrum of **1**.

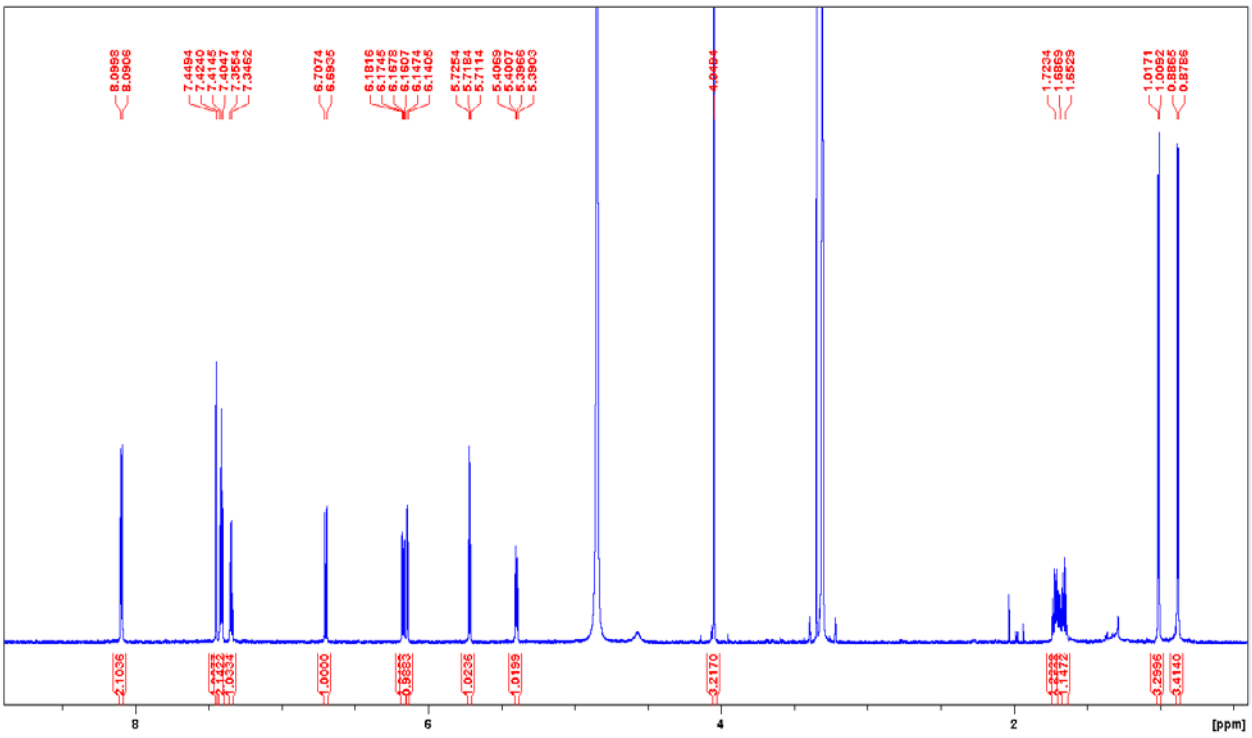


Figure S3. ^{1}H NMR spectrum of **1** (800 MHz, CD_3OD).

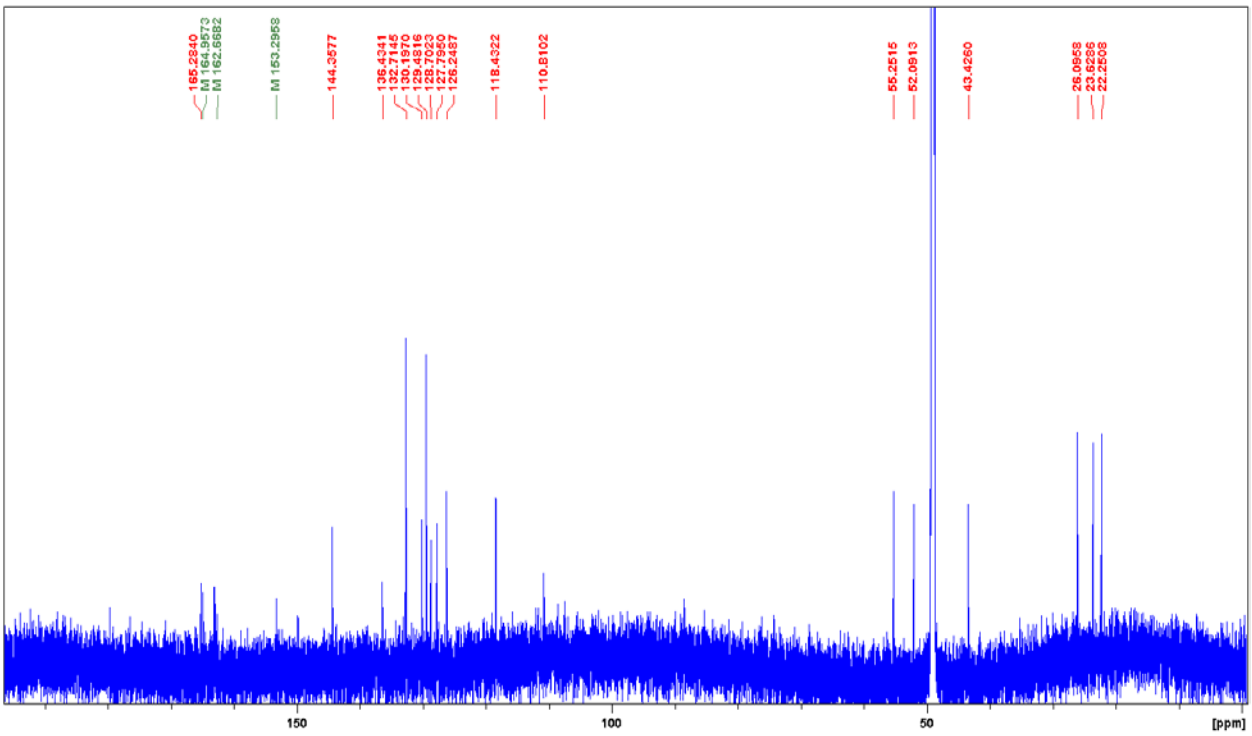


Figure S4. ^{13}C NMR spectrum of **1** (200 MHz, CD_3OD).

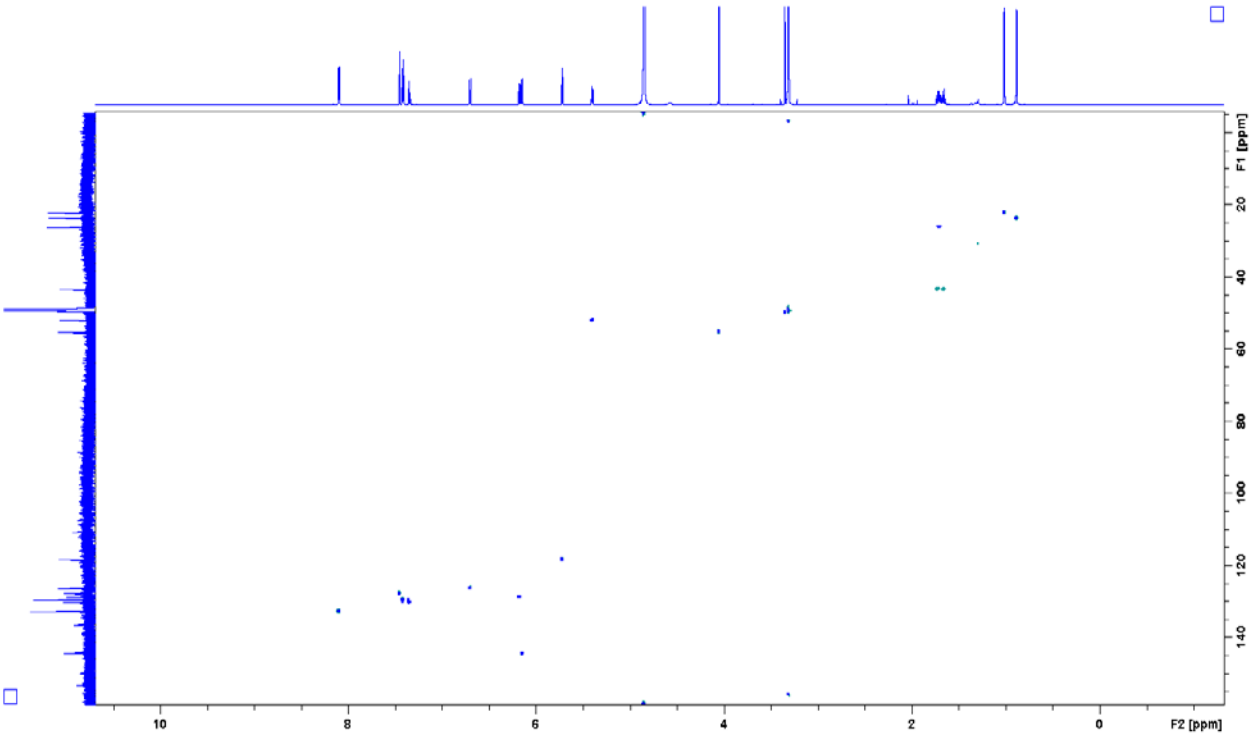


Figure S5. HSQC spectrum of **1**.

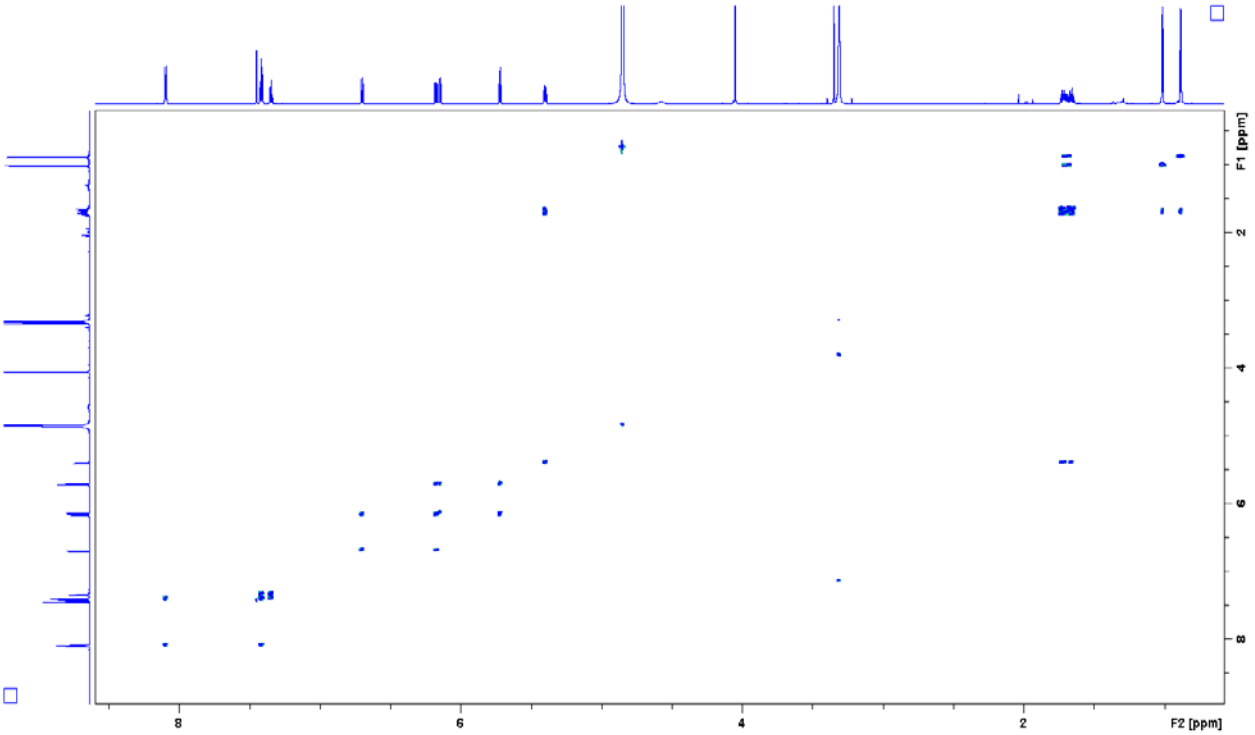


Figure S6. COSY spectrum of **1**.

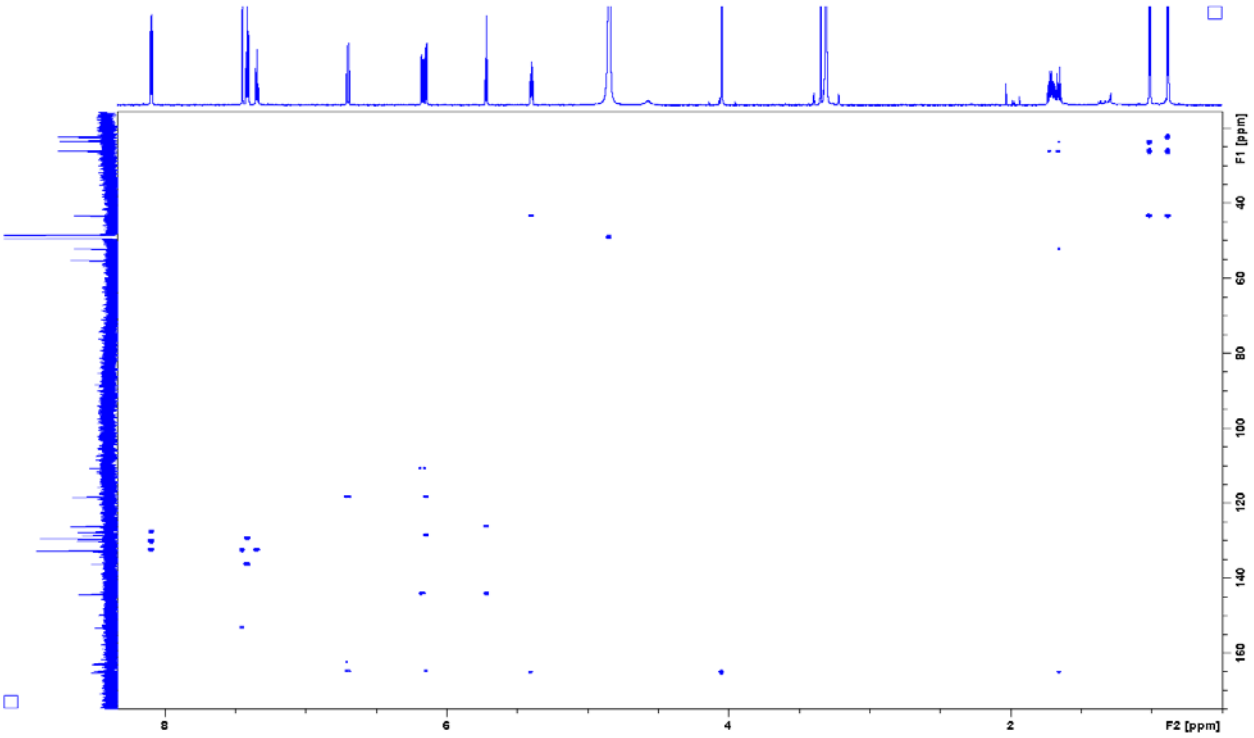


Figure S7. HMBC spectrum of **1**.

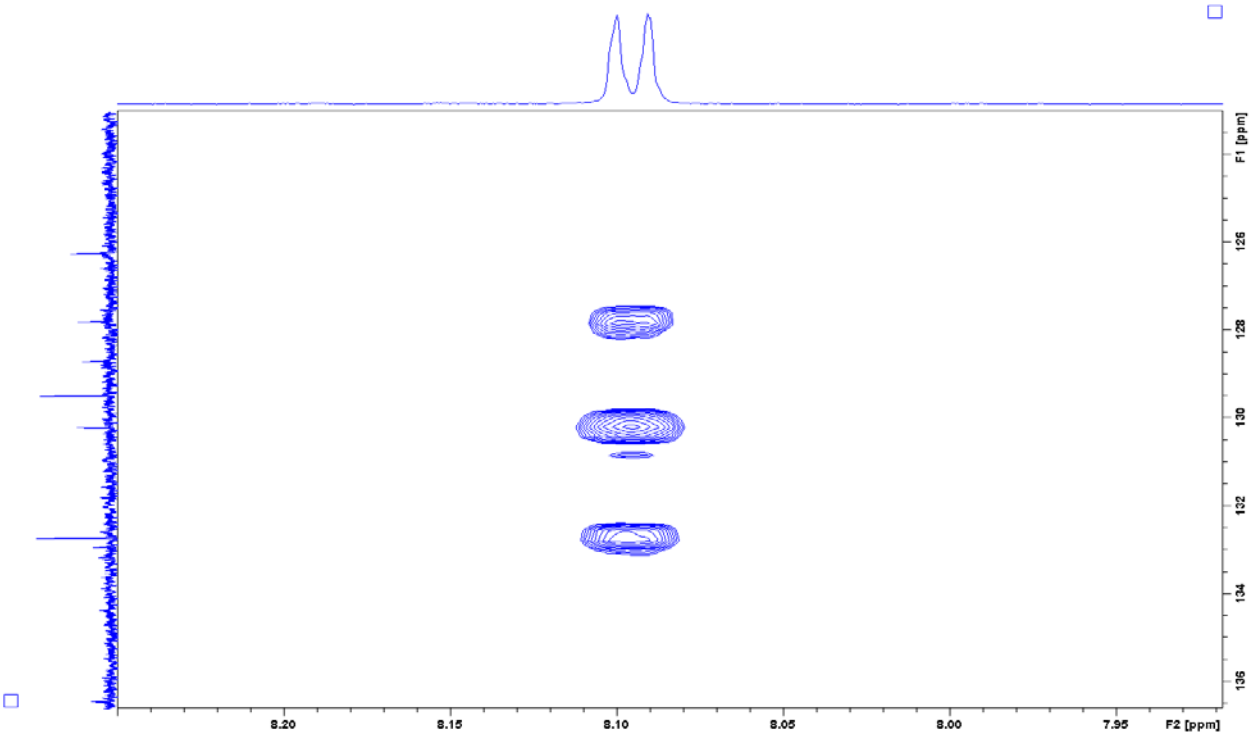


Figure S8. HMBC correlations of H-2´/6´ in **1**.

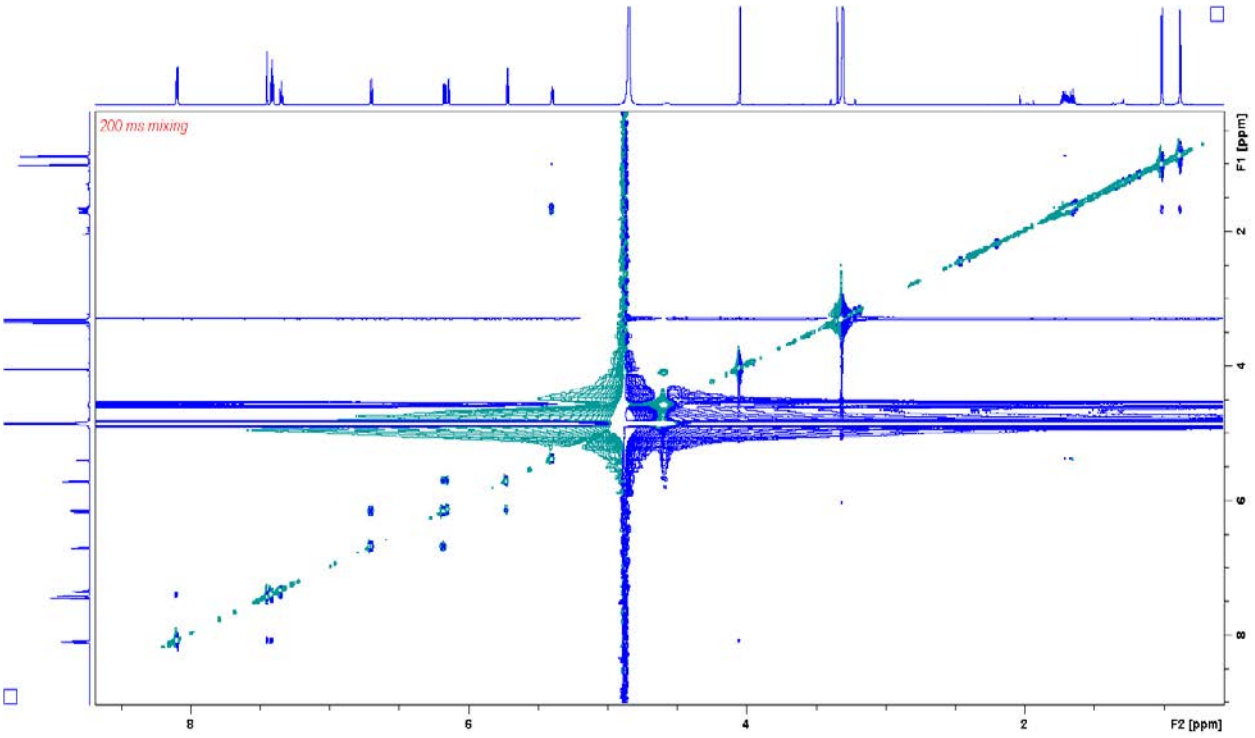


Figure S9. NOESY spectrum of **1**.

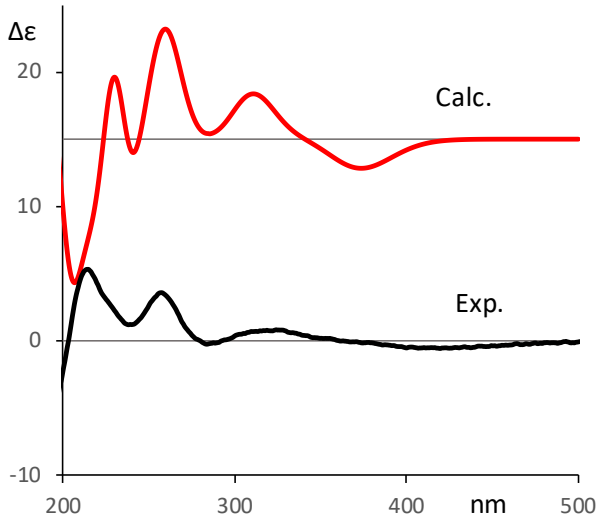


Figure S10. Comparison of the experimental and calculated ECD spectra of **1**.

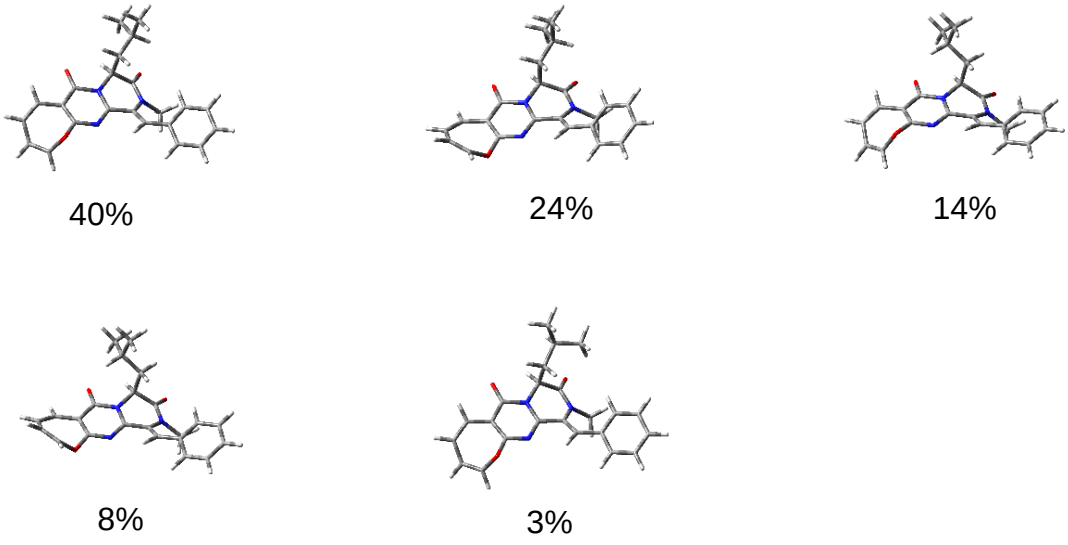


Figure S11. Structure and Boltzmann population of the predominant conformers of 15*S*-**1**.

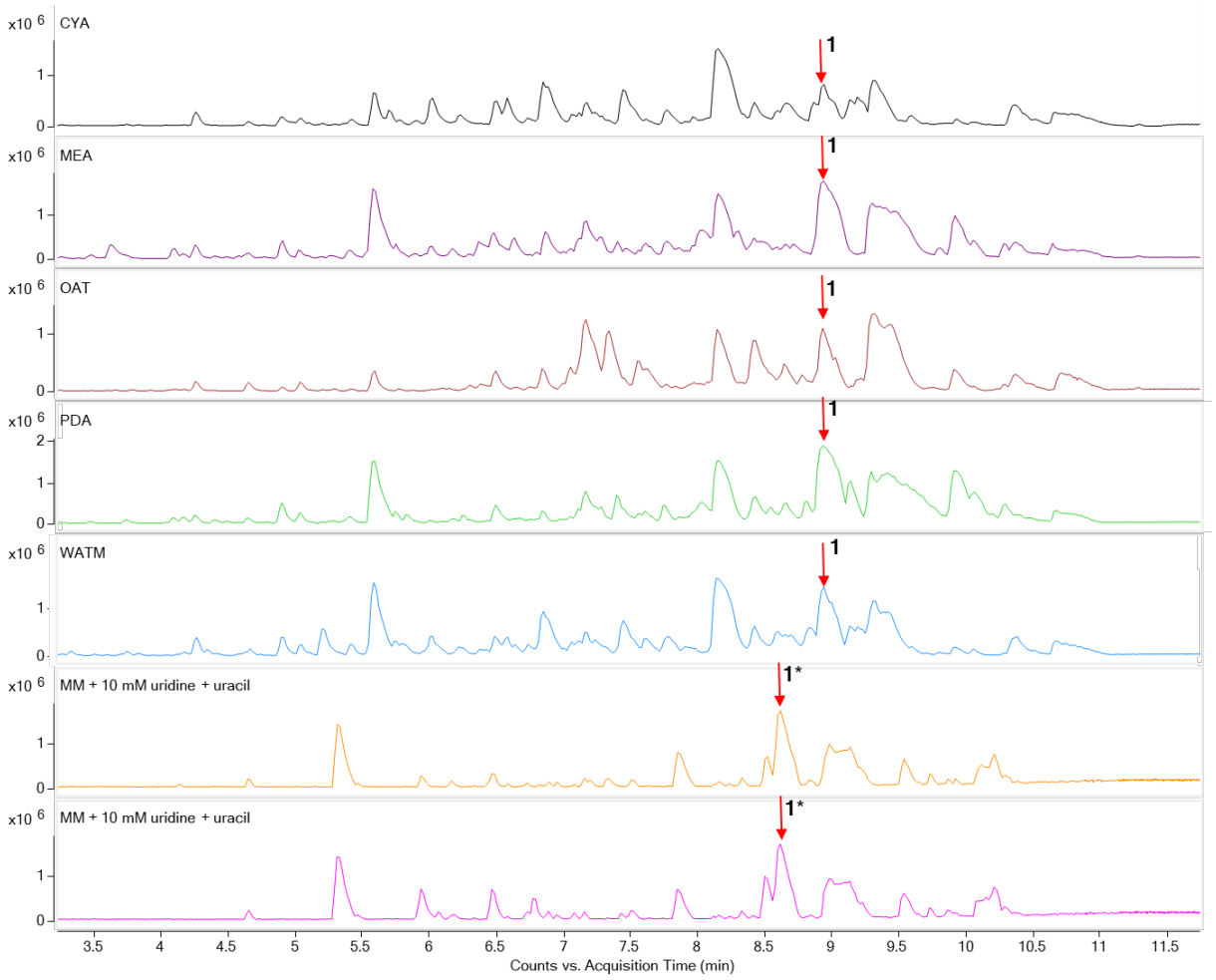


Figure S12 Production of **1** on six solid media in *Aspergillus californicus*.

A. californicus was inoculated as three-point stabs on CYA, MEA, OAT, PDA, WATM and MM with 10 mM uridine and uracil plates at 25 °C for 11 days (Samson et al. 2010; Hoof et al. 2018). Five plugs were harvested and transferred to a tube and extracted in ultra-sonication bath for 30 min with 1.5 mL ethyl acetate: isopropanol (3:1) solvent containing 1% formic acid, except that the bottom sample was extracted without formic acid. 1 mL liquid was transferred into a new tube and evaporated in a nitrogen stream. Dried samples were re-suspended with 200 μL of methanol, vortexed and centrifuged. 180 μL supernatant was transferred to HPLC vials for LC-MS analysis. *The retention time shift of the bottom two samples was due to the instrument pressure change as they were run on different days from the top five.

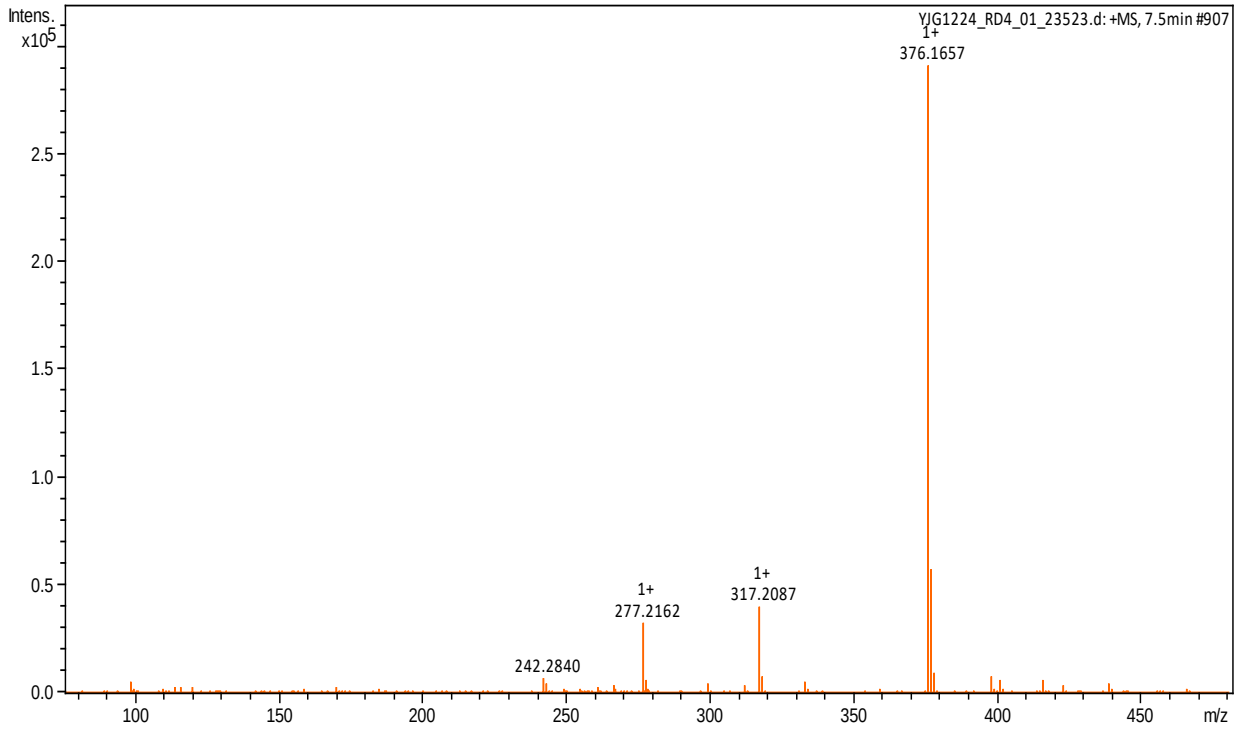


Figure S13. HRESIMS spectrum of **2**.

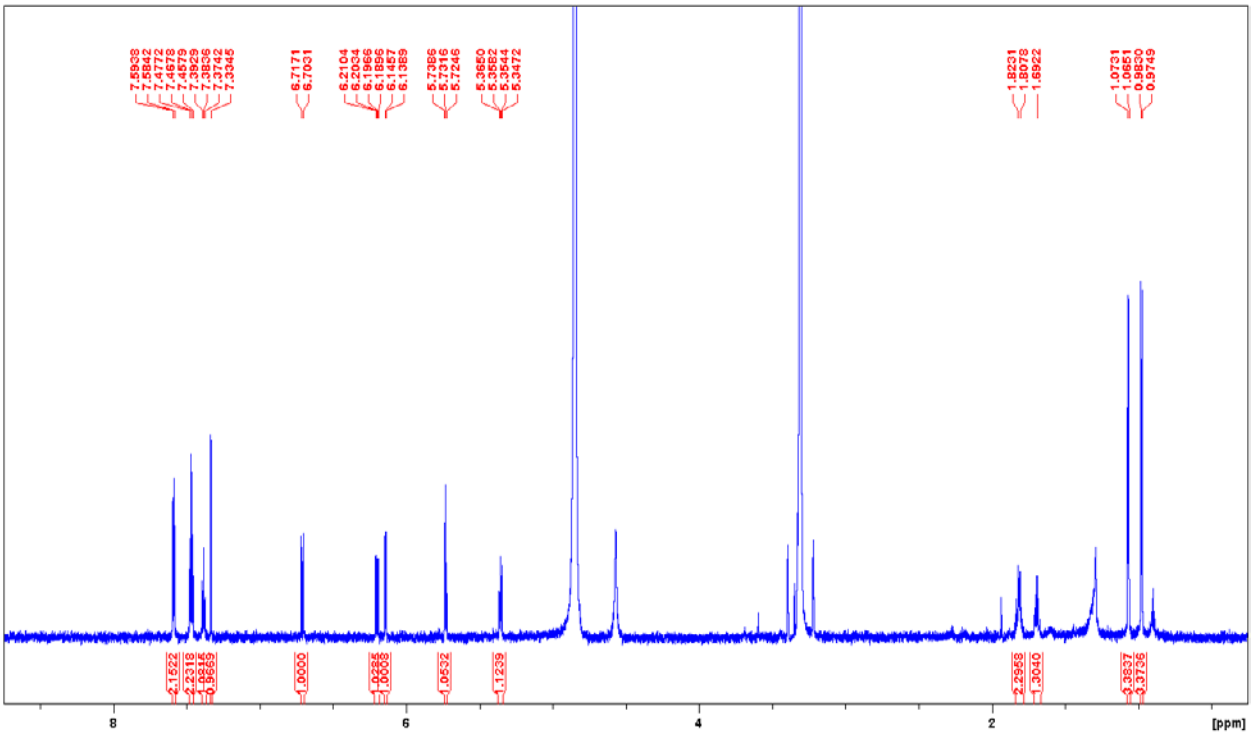


Figure S14. ^{1}H NMR spectrum of **2** (800 MHz, CD_3OD).

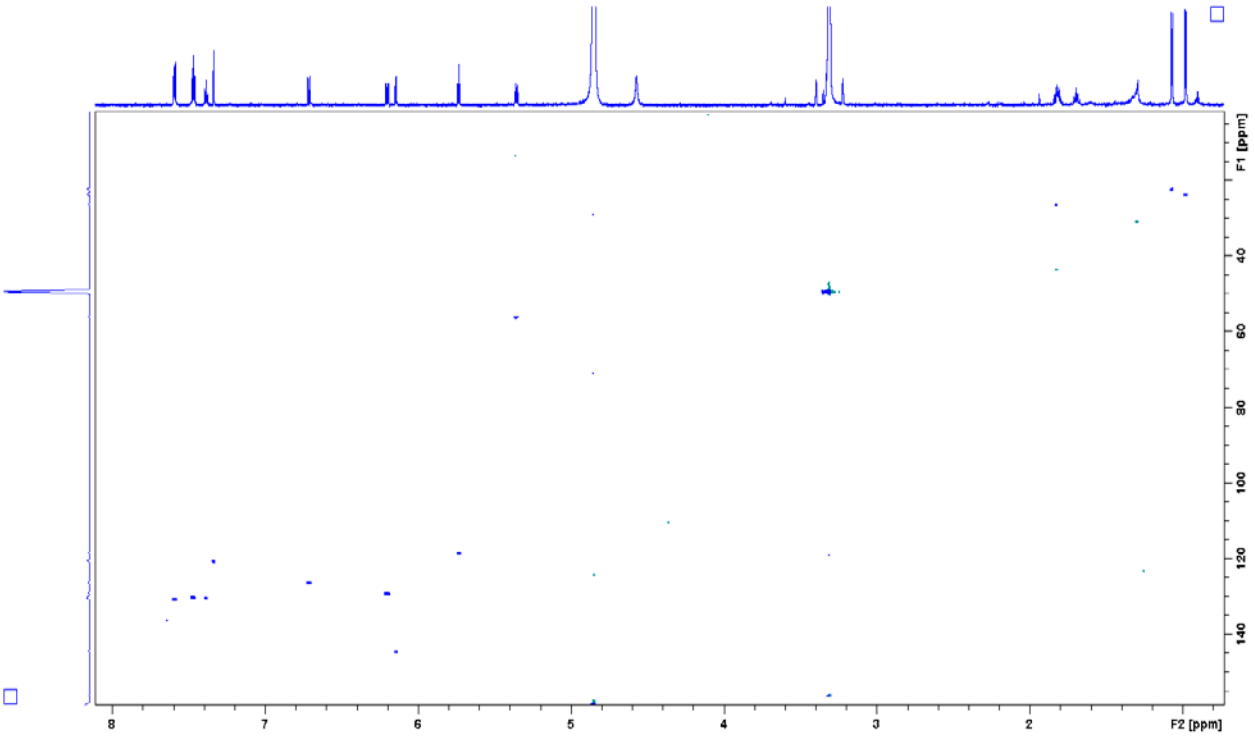


Figure S15. HSQC spectrum of **2**.

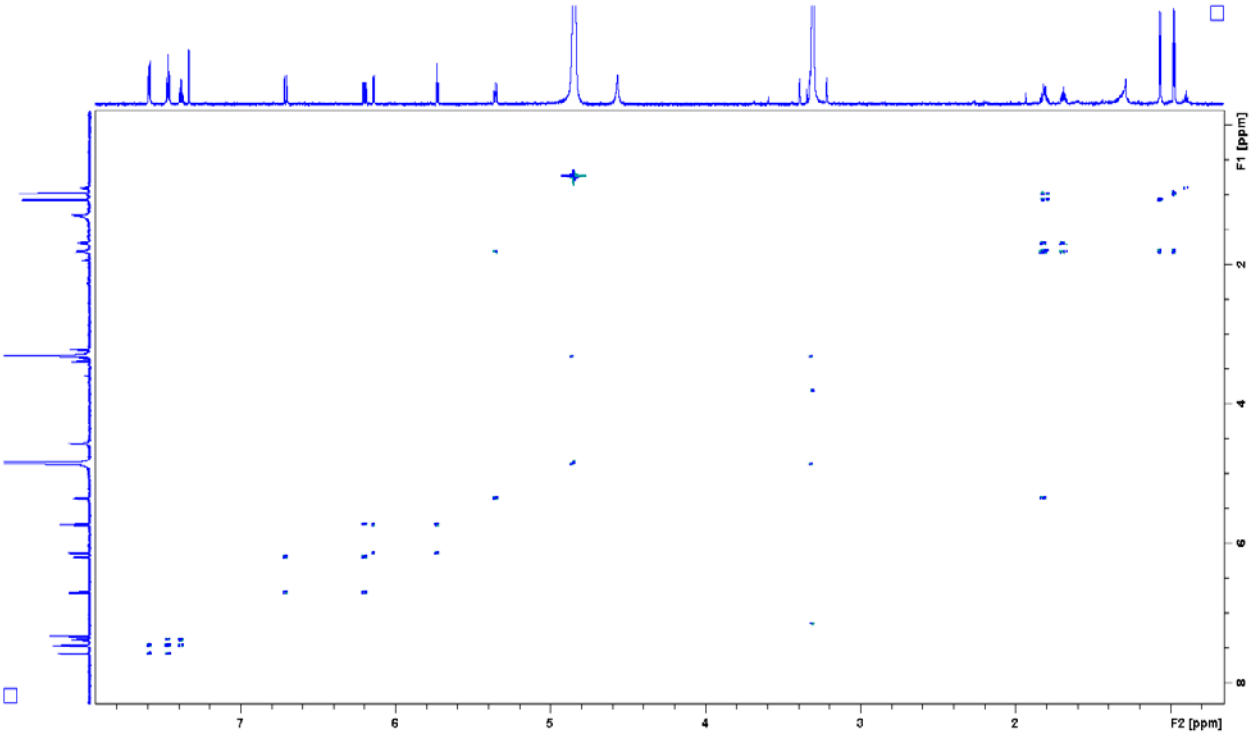


Figure S16. COSY spectrum of **2**.

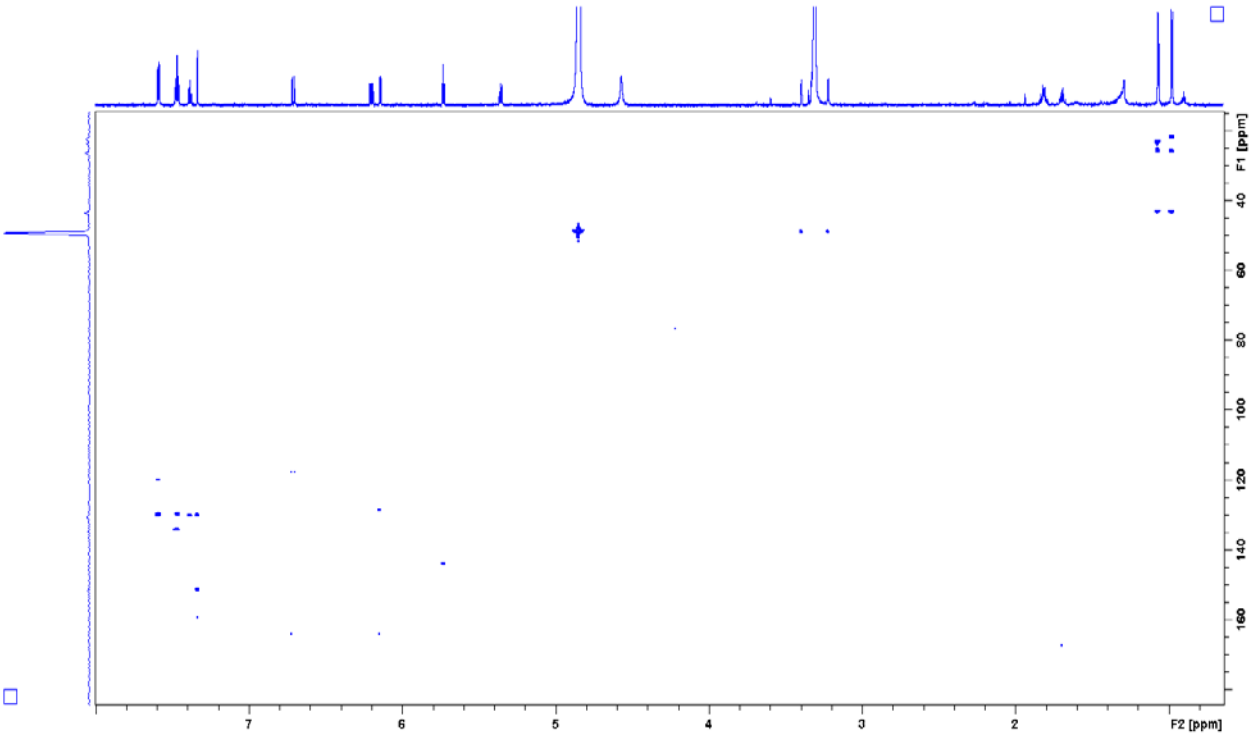


Figure S17. HMBC spectrum of **2**.

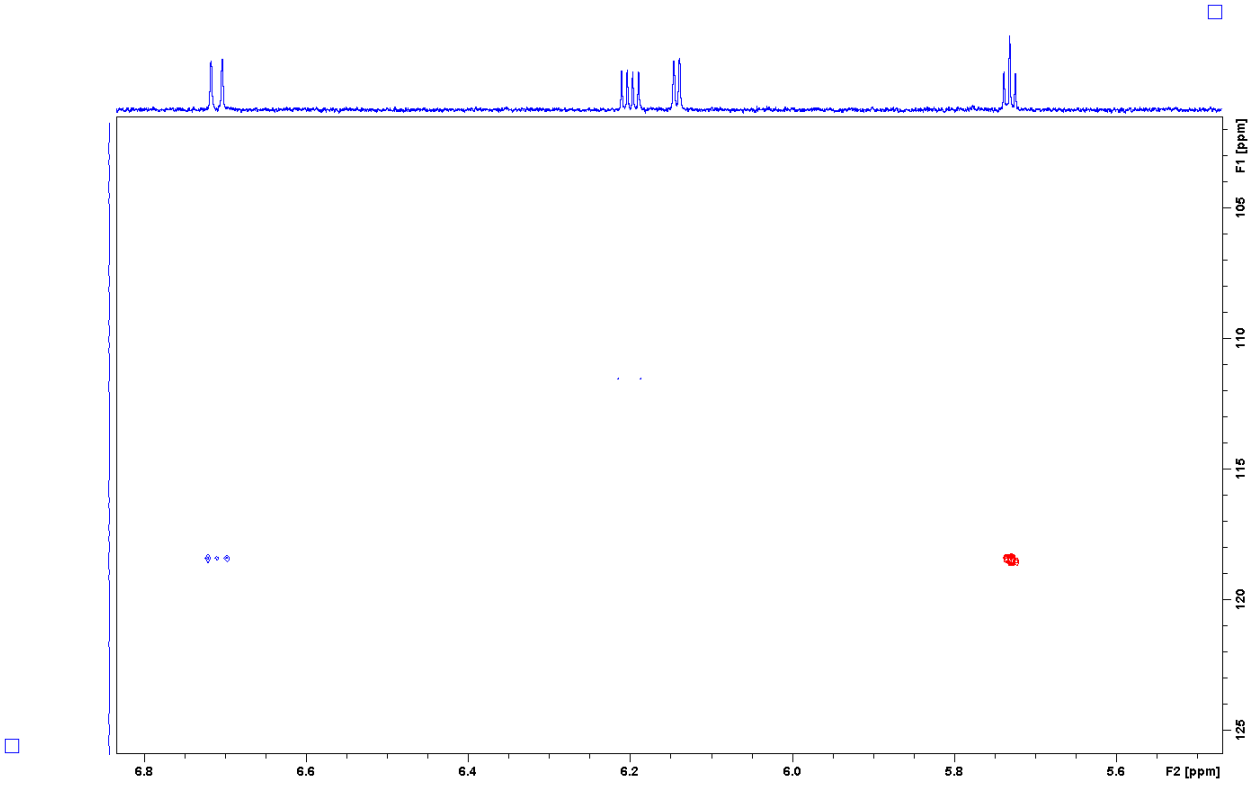


Figure S18. Overlapping display of the selected HSQC (red) and HMBC (blue) spectral region of **2**.

Antibacterial and cytotoxic assays

Antibacterial and cytotoxic tests were performed by Foundación Medina (Granada, Spain).

Supplementary references

Hoof JH, Nødvig CS, Mortensen UH. 2018. Fungal Genomics. vol. 1775. New York (NY): Humana Press. Chapter 11, Genome editing: CRISPR-Cas9; p. 119–132.

Samson RA, Houbraken J, Thrane U, Frisvad JC, Andersen B. 2010. Food and indoor fungi: second edition. Utrecht, The Netherlands: CBS-KNAW Fungal Biodiversity Centre.

metabolites

MDPI

Review

Review of Oxepine-Pyrimidinone-Ketopiperazine Type Nonribosomal Peptides

Yaojie Guo, Jens C. Frisvad and Thomas O. Larsen *

Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, Building 221, DK-2800 Kgs. Lyngby, Denmark; yaguo@dtu.dk (Y.G.); jcf@bio.dtu.dk (J.C.F.)
* Correspondence: tol@bio.dtu.dk; Tel.: +45-4525-2632

Received: 12 May 2020; Accepted: 8 June 2020; Published: 15 June 2020

Abstract: Recently, a rare class of nonribosomal peptides (NRPs) bearing a unique Oxepine-Pyrimidinone-Ketopiperazine (OPK) scaffold has been exclusively isolated from fungal sources. Based on the number of rings and conjugation systems on the backbone, it can be further categorized into three types A, B, and C. These compounds have been applied to various bioassays, and some have exhibited promising bioactivities like antifungal activity against phytopathogenic fungi and transcriptional activation on liver X receptor α. This review summarizes all the research related to natural OPK NRPs, including their biological sources, chemical structures, bioassays, as well as proposed biosynthetic mechanisms from 1988 to March 2020. The taxonomy of the fungal sources and chirality-related issues of these products are also discussed.

Keywords: oxepine; nonribosomal peptides; bioactivity; biosynthesis; fungi; *Aspergillus*

1. Introduction

Nonribosomal peptides (NRPs), mostly found in bacteria and fungi, are a class of peptidyl secondary metabolites biosynthesized by large modularly organized multienzyme complexes named nonribosomal peptide synthetases (NRPSs) [1]. These products are amongst the most structurally diverse secondary metabolites in nature; they exhibit a broad range of activities, which have been exploited in treatments such as the immunosuppressant cyclosporine A and the antibiotic daptomycin [2,3]. Due to their high importance, a lot of bioengineering studies have been carried out to elucidate their biosynthetic pathways, increase their yields, and generate novel homologs [4,5]. Within the recent decades, a rarely observed class of NRPs containing an Oxepine-Pyrimidinone-Ketopiperazine (OPK) scaffold comprising three amino acids, including one or two anthranilic acid(s), has emerged since the isolation of cinereain 32 years ago [6]. Interestingly, the structures of OPK NRPs are close to some quinazolinone alkaloids, specifically types Q12 to Q18 quinazolinones, such as fumiquinazolines and benzomalvins mostly produced by *Aspergillus* and *Penicillium* species as summarized in a recent review covering 157 compounds [7]. One major difference of the core skeleton between OPK NRPs and those specific quinazolinones is that OPK compounds bear a unique oxepine moiety instead of a phenyl group. Additionally, the OPK compounds were also described as diketopiperazine alkaloids [8–10]. However, they were not included in recent reviews on quinazolinones or diketopiperazines [7,11–14]. More attention should be paid to this class of compounds, considering their various bioactivities and intriguing structures, although some synthetic efforts have already been made [15,16]. To get a comprehensive perspective, here we review different aspects of these OPK NRPs, including their biological sources, structures, bioactivities, and proposed biosynthesis, for the first time.

2. Results

2.1. Biological Sources and Chemical Structures

Up to March 2020, thirty-five products bearing OPK backbone (Figure 1, Tables 1–3) have been isolated from natural sources, surprisingly all from fungi. The first compound reported was cinereain (**1**) from fungus *Botrytis cinerea* ATCC 64157 cultured on shredded wheat medium [6] followed by the isolation of asperloxin A (**2**) [17] and B (**3**) [18] from *Aspergillus ochraceus* DSM 7428, which was a part of One-Strain-Many-Compounds (OSMAC) approach to release the chemical diversity of this strain in A. Zeeck´s group [19]. Oxepinamides A–C (**4**–**6**) were reported to be isolated from the organic extract of the culture broth and mycelia of filamentous fungus *Acremonium* sp. grown in static liquid culture containing seawater-based medium [20]. Janoxepin (**7**) with a rare D-leucine residue was obtained from *Aspergillus janus* IBT 22274 cultivated on yeast extract sucrose (YES) medium [21]. Circumdatins A (**2**) and B (**8**), first reported to be benzodiazepines with two benzyl groups from *Aspergillus ochraceus* IBT 12704 as good chemotaxonomic markers [22], were later isolated from a marine-derived fungus *Aspergillus ostianus* strain 01F313, and their structures were revised to be oxepine-containing benzodiazepine alkaloids by X-ray crystallography [23]. The structure of circumdatin A was finally established to be the same as reported for asperloxin A (**2**) [17]. The first oxepine-containing alkaloid with a phenylalanine residue brevianamide L (**9**) containing a 12-hydroxyl dihydro-oxepine ring, together with brevianamides O and P (**10**–**11**), was isolated from the solid-state fermented rice culture of *Aspergillus versicolor* (AS 3.4186) [8,9]. Oxepinamide D (**12**) and oxepinamides E–G (**13**–**15**), containing a 12-oxygenated-oxepine ring, were isolated from *Aspergillus puniceus* F02Z-1744 grown on solid media containing rice and soybean [24]. Protuboxepins A (**16**) and B (**17**) were isolated from the marine-derived fungus *Aspergillus* sp. SF-5044, whose 28S rRNA gene (Genbank accession number FJ935999) showed a high-sequence identity of 99.64% with that of *Aspergillus protuberus* (FJ176897) [25]. Circumdatin L (**18**) was isolated from the solid rice culture of *Aspergillus westerdijkiae* DFFSCS013 [26]. Dihydrocinereain (**19**) with cinereain (**1**) was characterized from a marine strain of *Aspergillus carneus* KMM 4638 grown on modified rice medium with seawater [27]. Varioxepine A (**20**) bearing a unique oxa-cage was isolated from the marine algal-derived fungus *Paecilomyces variotii* EN-291 fermented in potato dextrose broth medium [28]. Varioloids A and B (**21**–**22**) with protuboxepin B (**17**) were also isolated from *Paecilomyces variotii* EN-291 fermented in the same condition [10]. Versicoloids A and B (**23**–**24**) were isolated from the deep-sea-derived fungus *Aspergillus versicolor* SCSIO 05879 grown in liquid medium containing starch and polypeptone [29]. Versicomide D (**25**) was isolated from *Aspergillus versicolor* XZ-4 fermented in liquid medium with seawater [30]. Protuboxepins C and D (**26**–**27**) were isolated from the sponge-derived fungus *Aspergillus* sp. SCSIO XWS02F40, which was found to belong to a clade related to *Aspergillus austroafricanus* NRRL 233 with an identity of 99.4% using ITS1-5.8S-ITS2 sequence region [31,32]. Chrysopiperazines A and B (**28**–**29**) with versicoloids A and B (**23**–**24**) were obtained from a gorgonian-derived *Penicillium chrysogenum* strain (CHNSCLM-0019), and their absolute configurations were completely solved by NOESY, Marfey's reaction, and electronic circular dichroism (ECD) and vibrational circular dichroism (VCD) methods [33]. Protuboxepins F (**30**) and G (**31**) were isolated from the marine sponge-derived fungus *Aspergillus versicolor* SCSIO 41016 grown on solid rice media with artificial sea salt [34]. Oxepinamides H−K (**32**–**35**) were isolated from a deep-sea-derived *Aspergillus puniceus* SCSIO z021 fermented in liquid medium with sea salt [35].

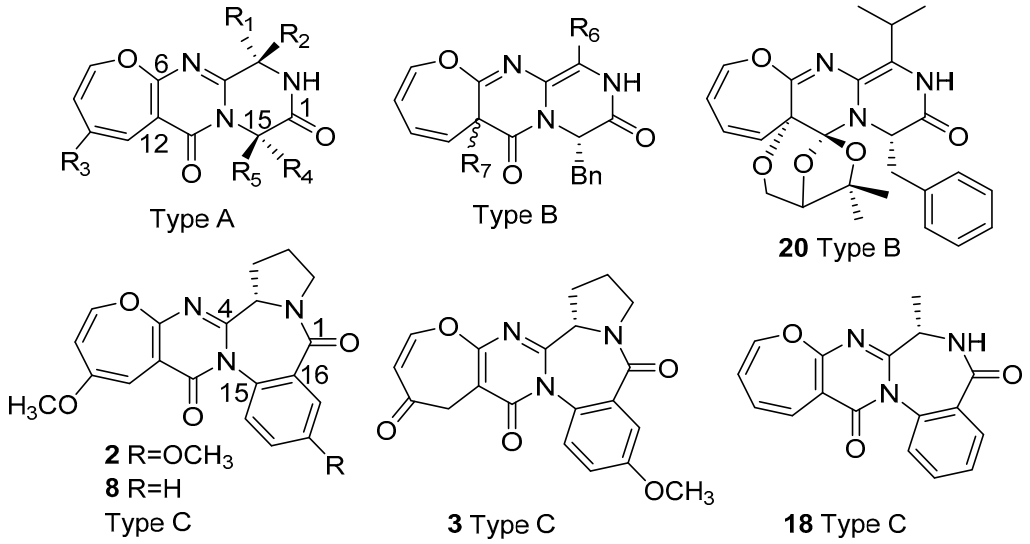


Figure 1. Structures of three types (A, B, and C) of Oxepine-Pyrimidinone-Ketopiperazine (OPK) nonribosomal peptides (NRPs).

Table 1. Structures and Biological sources of Type A Oxepine-Pyrimidinone-Ketopiperazine NRPs.

No.	Name	Substitution groups					Sources
		R_1	R_2	R_3	R_4	R_5	
1	Cinereain	=CHCH(CH_3)$_2$, Z	n/a	H	CH(CH_3)$_2$	H	*Botrytis cinerea* ATCC 64157 [6] *Aspergillus carneus* KMM 4638 [27]
4	Oxepinamide A	CH(CH_3)CH_2CH_3	OH	OCH_3	H	CH_3	*Acremonium* sp. [20]
5	Oxepinamide B	OH	CH(CH_3)CH_2CH_3	OCH_3	H	CH_3	*Acremonium* sp. [20]
6	Oxepinamide C	CH_2CH(CH_3)$_2$	OCH_3	OCH_3	H	CH_3	*Acremonium* sp. [20]
7	Janoxepin	=CHCH(CH_3)$_2$, Z	n/a	H	H	CH_2CH(CH_3)$_2$	*Aspergillus janus* IBT 22274 [21]
10	Brevianamide O	OH	CH(CH_3)CH_2CH_3	H	Benzyl	H	*Aspergillus versicolor* (AS 3.4186) [9]
11	Brevianamide P	H	CH(CH_3)CH_2CH_3	H	Benzyl	H	*Aspergillus versicolor* (AS 3.4186) [9]
12	Oxepinamide D	OH	Benzyl	H	H	CH_3	*Aspergillus puniceus* F02Z-1744 [24]
16	Protuboxepin A	CH(CH_3)CH_2CH_3	H	H	H	Benzyl	*Aspergillus* sp. SF-5044 [25] *Penicillium expansum* Y32 [36]
17	Protuboxepin B	CH(CH_3)$_2$	H	H	H	Benzyl	*Aspergillus* sp. SF-5044 [25] *Paecilomyces variotii* EN-291 [10] *Penicillium expansum* Y32 [36]
19	Dihydrocinereain	H	CH_2CH(CH_3)$_2$	H	CH(CH_3)$_2$	H	*Aspergillus carneus* KMM 4638 [27]
22	Varioloid B	OCH_3	CH(CH_3)$_2$	H	Benzyl	H	*Paecilomyces variotii* EN-291 [10]
23	Versicoloid A	H	CH(CH_3)CH_2CH_3	OCH_3	CH(CH_3)$_2$	H	*Aspergillus versicolor* SCSIO 05879 [29] *Penicillium chrysogenum* CHNSCLM-0019 [33]
24	Versicoloid B	OH	CH(CH_3)CH_2CH_3	OCH_3	CH(CH_3)$_2$	H	*Aspergillus versicolor* SCSIO 05879 [29] *Penicillium chrysogenum* CHNSCLM-0019 [33]
25	Versicomide D	CH(CH_3)CH_2CH_3, 18*S*	H	OCH_3	CH(CH_3)$_2$	H	*Aspergillus versicolor* XZ-4 [30]
26	Protuboxepin C	CH(CH_3)CH_2CH_3, 16*S*	OCH_3	H	H	Benzyl	*Aspergillus* sp. SCSIO XWS02F40 [31]
27	Protuboxepin D	CH(CH_3)CH_2CH_3, 16*S*	OH	H	H	Benzyl	*Aspergillus* sp. SCSIO XWS02F40 [31]
28	Chryzopiperazine A	CH(CH_3)CH_2CH_3, 19*S*	OCH_3	OCH_3	H	CH(CH_3)$_2$	*Penicillium chrysogenum* CHNSCLM-0019 [33]
29	Chrysopiperazine B	OCH_3	CH(CH_3)CH_2CH_3, 19*S*	OCH_3	H	CH(CH_3)$_2$	*Penicillium chrysogenum* CHNSCLM-0019 [33]

Table 1. *Cont.*

No.	Name	Substitution groups					Sources
		R_1	R_2	R_3	R_4	R_5	
30	Protuboxepin F	=CHCH(CH_3)$_2$, *Z*	n/a	H	H	Benzyl	*Aspergillus versicolor* SCSIO 41016 [34]
31	Protuboxepin G	=CHCH(CH_3)$_2$, *E*	n/a	H	H	Benzyl	*Aspergillus versicolor* SCSIO 41016 [34]
32	Oxepinamide H	OCH_3	Benzyl	H	H	CH_3	*Aspergillus puniceus* SCSIO z021 [35]
33	Oxepinamide I	Benzyl	OCH_3	H	H	CH_3	*Aspergillus puniceus* SCSIO z021 [35]
34	Oxepinamide J	Benzyl	OH	H	H	CH_3	*Aspergillus puniceus* SCSIO z021 [35]
35	Oxepinamide K	=CH-Phenyl, *Z*	H	H	H	CH_3	*Aspergillus puniceus* SCSIO z021 [35]

Note: backbone numberings follow Figure 1, and the other numberings are based on the original publications. n/a: not applicable due to double bond substitution.

Table 2. Structures and Biological sources of Type B Oxepine-Pyrimidinone-Ketopiperazine NRPs.

No.	Name	Substitution Groups		Sources
		R_6	R_7	
9	Brevianamide L	$CH(CH_3)CH_2CH_3$	OH, 12*S*	*Aspergillus versicolor* (AS 3.4186) [8]
13	Oxepinamide E	$CH(CH_3)CH_2CH_3$, 17*S*	OH, 12*R*	*Aspergillus puniceus* F02Z-1744 [24]
14	Oxepinamide F	$CH(CH_3)CH_2CH_3$, 17*S*	OCH_3, 12*R*	*Aspergillus puniceus* F02Z-1744 [24]
15	Oxepinamide G	$CH(CH_3)_2$	OCH_3, 12*R*	*Aspergillus puniceus* F02Z-1744 [24]
20	Varioxepine A	$CH(CH_3)_2$	See Figure 1	*Paecilomyces variotii* EN-291 [28]
21	Varioloid A	$CH(CH_3)_2$	$O(CH_2)COCH(CH_3)_2$, 12*R*	*Paecilomyces variotii* EN-291 [10]

Table 3. Structures and Biological sources of Type C Oxepine-Pyrimidinone-Ketopiperazine NRPs.

No.	Name	Scaffold	Sources
2	Asperloxin A (Circumdatin A)	7/6/7/6/5, Figure 1	*Aspergillus ochraceus* DSM 7428 [17] *Aspergillus ochraceus* IBT 12704 [22] *Aspergillus ostianus* 01F313 [23]
3	Asperloxin B	7/6/7/6/5, Figure 1	*Aspergillus ochraceus* DSM 7428 [18]
8	Circumdatin B	7/6/7/6/5, Figure 1	*Aspergillus ochraceus* IBT 12704 [22] *Aspergillus ostianus* 01F313 [23]
18	Circumdatin L	7/6/7/6, Figure 1	*Aspergillus westerdijkiae* DFFSC S013 [26]

2.2. Bioactivities

2.2.1. Plant Growth Regulation

Cinereain (**1**), the first OPK peptide, could significantly inhibit the growth of etiolated wheat coleoptiles ($p < 0.01$) at 10^{-4} and 10^{-3} M and cause mild necrosis and chlorosis in corn, but it did not have any effect on intact greenhouse-grown bean and tobacco plants [6].

2.2.2. Anti-Inflammatory Activity

In a topical resiniferatoxin (RTX)-induced mouse ear edema assay, oxepinamide A (**4**) showed good topical anti-inflammatory activity with 82% inhibition of edema at the standard testing dose of 50 µg per ear [20].

2.2.3. Antifungal Activity

Oxepinamides A–C (**4**–**6**) showed no antifungal activity toward *Candida albicans* in a broth micro-dilution assay [20]. Janoxepin (**7**) showed no antifungal activity in an in-house disc diffusion assay [21]. Brevianamide L (**9**) showed no inhibitory activity against *Candida albicans* at a concentration of 100 µg/mL [8]. However, varioxepine A (**20**) and varioloids A and B (**21**–**22**) exhibited activity against the plant-pathogenic fungus *Fusarium graminearum* with MIC values of 4, 8 µg/ml, respectively [10,28]. Versicoloids A and B (**23**–**24**) exhibited antifungal activities against the three phytopathogenic fungi *Colletotrichum acutatum*, *Magnaporthe oryzae*, and *Fusarium oxysporum*, both with MICs of 1.6, 128, and 64 µg/mL. Their activity against *Colletotrichum acutatum* was even stronger than the positive control cycloheximide (MIC of 6.4 µg/mL), and they could be regarded as candidate agrochemical antifungal agents [29]. Chrysopiperazine A (**28**) did not show activity against *Candida albicans* at the concentration of 50 µM [33]. Oxepinamides H–K (**32**–**35**) showed low percent inhibition (< 50%) against the four phytopathogenic fungi—*Curvularia australiensis*, *Colletotrichum gloeosporioides*, *Fusarium oxysporum*, and *Pyricularia oryzae*—at a concentration around 0.6 mM [35].

2.2.4. Cytotoxicity

Oxepinamides A–C (**4**–**6**) showed no significant activity against any cell line in the National Cancer Institute´s 60 cell-line panel [20]. Circumdatin B (**8**) was also tested in the NCI´s 60 cancer cell line panel and did not show activity either [22]. Neither Circumdatin A (**2**) nor Circumdatin B (**8**) showed cytotoxicity against A548 lung cancer cells [23]. Brevianamides L, O, and P (**9**–**11**) showed no

cytotoxicity against human breast cancer (Bre04), human lung (Lu04), or human neuroma (N04) cell lines (GI_{50} > 10 μg/mL) [8,9]. Protuboxepin A (**16**) showed weak inhibitory activity against human acute promyelocytic leukemia cells (HL-60), human breast cancer adenocarcinoma cells (MDA-MB-231), hepatocellular carcinoma cells (Hep3B), rat fibroblast cells (3Y1), and chronic myelogenous leukemia cells (K562), with IC_{50} values of 75, 130, 150, 180, and 250 μM, respectively [25]. A further in vitro study revealed that this compound could bind to α- and β-tubulin and thereby stabilize tubulin polymerization, altogether disrupting microtubule dynamics. This disruption led to chromosome misalignment and metaphase arrest, inducing apoptosis in tumor cells [37]. The compound circumdatin L (**18**) did not show cytotoxicity toward the human carcinoma A549, HL-60, K562, and MCF-7 cell lines (IC_{50} > 10 μM) [26]. Dihydrocinereain (**19**) was tested against murine ascites Ehrlich carcinoma cells but did not show activity up to 100 μM [27]. Similarly, protuboxepins C and D (**26**-**27**) showed no inhibitory activity against A549 cells with IC_{50} values of 100 and 190 μM and weak activities against HeLa cells with IC_{50} values of 61 and 114 μM [31]. Protuboxepin G (**31**) displayed moderate cytotoxic activities against three renal carcinoma cell lines (ACHN, OS-RC-2, and 786-O cells) with the IC_{50} values 27.0, 47.1, and 34.9 μM, respectively [34].

2.2.5. Antibacterial Activity

In disk assays with cinereain (**1**) against *Bacillus subtilis*, *Bacillus cereus*, and *Mycobacterium thermosphactum* (Gram-positive), and *Escherichia coli* and *Citrobacter freundii* (Gram-negative), no effects were observed in concentrations up to 500 μg per disk [6]. Janoxepin (**7**) showed no antibacterial activity in an in-house agar overlay assay [21]. Circumdatins A (**2**) and B (**8**) were subjected to an inhibitory test against Methicillin-resistant *Staphylococcus aureus* (MRSA), but no activities were observed [23]. Brevianamide L (**9**) showed no inhibitory activity against *Escherichia coli*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa*, at a concentration of 100 μg/mL [8]. Varioxepine A (**20**) and Varioloids A and B (**21**–**22**) showed promising antibacterial activities against *Micrococcus luteus*, *Staphylococcus aureus*, *Escherichia coli*, and the aquacultural bacteria *Aeromonas hydrophila*, *Vibrio anguillarum*, *Vibro harveyi* and *Vibro parahaemolyticus*, with MIC values ranging from 16 to 64 μg/ml [10,28]. Versicomide D (**25**) was applied to three pathogenic bacteria (*E. coli*, *S. aureus* and *B. subtilis*), but no MIC values were reported. Chrysopiperazine A (**28**) did not show activity against *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Photobacterium halotolerans*, and *Enterobacter cloacae*, at the concentration of 50 μM [33].

2.2.6. Anti-Plasmodial Activity

Janoxepin (**7**) exhibited antiplasmodial activity against the malaria parasite *Plasmodium falciparum* 3D7 in the radioisotope assay with IC_{50} lower than 28 mg/mL [21].

2.2.7. Transcriptional Activation

Selective transactivation effects of oxepinamides D–G (**12**–**15**) were examined, and they selectively showed moderate transcriptional activation on Liver X Receptor α (LXRα) with EC_{50} values of 10.6, 12.8, 13.6, and 12.1 μM, but no agonistic effects were observed towards other seven nuclear receptors FXRα, PPARα, PPARβ, PPARγ, RARα, RXRα, or ERα [24]. Oxepinamides H−K (**32**–**35**) later also showed the same activation effects on LXRα with EC_{50} values of 15, 15, 16, and 50 μM, respectively, but did not show inhibition activity against other seven enzymes [35].

2.3. Biosynthesis

The biosynthesis of OPK NRPs remains unsolved despite the fact that some biosynthetic pathway studies have been performed on similar quinazolinone alkaloids [38–41]. Possible biosynthetic pathways of several OPK compounds have, however, been proposed. Janoxepin (**7**) was suggested to be derived from the condensation of anthranilic acid with a diketopiperazine ring formed between two leucine residues, followed by oxidation of the benzoyl derivative to give the oxepine derivative [21]. Similarly, oxepinamide D (**12**) was proposed to be biosynthesized by the condensation

of a diketopiperazine with an anthranilic acid and subsequent oxidation of the benzene ring to form an arene oxide, which was opened through a thermal 6π electrocyclic ring-opening process. Oxepinamides E–G (**13**–**15**) were formed by dehydration on the 2,5-diketo ring, followed by the addition of water between C-6 and C-12 [24]. Circumdatins A (**2**) and B (**8**) were proposed to be biosynthesized by oxidation of circumdatins H and J to form a benzene oxide, where a retro-pericyclic reaction (benzene oxide–oxepine tautomerism) took place to produce the final products [23,42]. Similar to janoxepin (**7**), the backbone of varioxepine A (**20**) has also been proposed to be from the condensation of anthranilic acid with a diketopiperazine, followed by epoxidation of the benzene ring to form the oxepine derivative. A series of reactions were proposed, including a second epoxidation, ring arrangement, epoxy opening, prenylation, dihydroxylation, and/or cyclization to yield the end product [28]. Protuboxepin D (**27**) was proposed to be formed by condensation of D-phenylalanine, L-isoleucine, and anthranilic acid, followed by oxidation of the benzene ring to form the oxepine derivative through an epoxy precursor and sequent oxidation at C-3 to form the hydroxyl group. Protuboxepin C (**26**) was a methylation product of protuboxepin D (**27**) [31]. A recent report proposed that additional opening and oxidation could happen on the oxepine ring, which then may undergo addition of water, cyclization, and methylation to yield unique (di/tetra)-hydropyran-pyrimidinone-ketopiperazine heterotricyclic products [34].

3. Discussion

In total, thirty-five OPK compounds have currently been characterized from natural sources. The speed of novel OPK product discovery has been increasing in recent years, as over half of the currently described products were isolated during the past eight years (Figure 2A). It is quite noteworthy that all these compounds were isolated from five fungal genera. Specifically, 70% of OPK NRPs, including the rediscovered cases, were obtained from the genus *Aspergillus*, followed by genus *Penicillium* accounting for 14%, *Acremonium* 7%, *Paecilomyces* 7%, and *Botrytis* 2% (Figure 2B). Interestingly, all type C producers are from *Aspergillus* section *Circumdati*, including *A. ochraceus*, *A. ostianus*, and *A. westerdijkiae*, and a large proportion of type A and B compounds were obtained from different isolates of in particular the two species *A. versicolor* and *A. protuberus*, both belonging to *A. versicolor* clade in section *Nidulantes* [43–45]. In general, OPK compounds have been reported from species in the closely related fungal families *Aspergillaceae* (*Aspergillus*, *Penicillium*) and *Trichocomaceae* (*Paecilomyces*). *Botrytis cinerea* and *Acremonium* species are distantly related to *Aspergillaceae* and *Trichocomaceae*, but they were also reported to produce OPK compounds. Unfortunately, several of the reported strains have not been deposited in any culture collections affiliated to the World Federation for Culture Collections (WFCC), which is possibly why their identity has not been validated. It is also notable that even though some species reported bear the initials of a collection center, their strain number cannot be traced in the corresponding collection system. For example, *Aspergillus ochraceus* DSM 7428 cannot be found in DSMZ collection, and *Aspergillus versicolor* (AS 3.4186) cannot be traced in CGMCC collection. While the identification of *Botrytis cinerea* (ATCC 64157) can be verified, the identification of *Acremonium* (strain unavailable) was based on fatty acid methyl ester (FAME) profiles, a method which has not been authenticated for identification purposes in filamentous fungi. Genome mining of *Botrytis* and *Acremonium* species will show whether OPK compounds are taxonomically widespread or restricted to *Aspergillaceae* and *Tricocomaceae*.

Based on the number of rings and conjugation systems on the backbone, OPK NRPs were categorized into three types: A, B, and C. Type A dominating the OPK NRPs with 25 compounds shares the same 7/6/6 backbone, whereas type B OPK's contains a larger conjugation system. Type C 7/6/7/6 backbone has one more ring than types A and B due to incorporation of a second anthranilic acid moiety, and some products even display a complex 7/6/7/6/5 ring system with an additional pyrrolidine-ring from proline. In nature, a lot of other OPK similar products have been isolated, such as the quinazolinones [7,13]. Due to their possible related biosynthetic pathways, mistakes might happen during structure elucidation [22,23]. One common issue with OPK compounds is the absolute configuration (AC) determination of α carbons and R groups of the amino acids. In many

reports, NOESY, Marfey´s reaction, X-ray crystallography, and ECD methods were applied. However, one might observe a mixture of D- and L- products after the acid hydrolysis and derivation process when using Marfey´s reaction method. Hydrolysis conditions thus may need to be optimized. In the case of a chiral center at a flexible position, it can be very challenging to solve the correct configuration. Success has recently been achieved by comparing the experimental VCD spectrum with calculated data [33]. The chiral centers within the R group of the isoleucine residue in eight OPK compounds (**4**, **5**, **9**, **10**, **11**, **16**, **23**, and **24**) remain uncharacterized. The chirality also makes it confusing when referring to a structure in a publication. For example, the drawings of oxepinamide E and F (**13**–**14**) showed a 17*R* configuration (wrong) but was described as 17*S* (correct configuration by X-ray Crystallography) in the same paper [24]. Additionally, the chiral center of janoxepin (**7**) was determined as *R* configuration by Marfey´s method, but the drawing mistakenly exhibited *S* configuration [21]. Such errors also happened when the structures were drawn in different publications, like the chirality of the two α carbons of both versicoloids A and B (**23**–**24**) drawn in a recent paper [33], which displayed opposite configurations from the original structures [29]. Care should be taken to avoid making such erroneous configurational drawings. Moreover, it is also notable that both D- and L- amino acids can participate in building the OPK products based on all the characterized structures. Therefore, proposing the chirality of α carbon from a biogenetic prospect can be challenging.

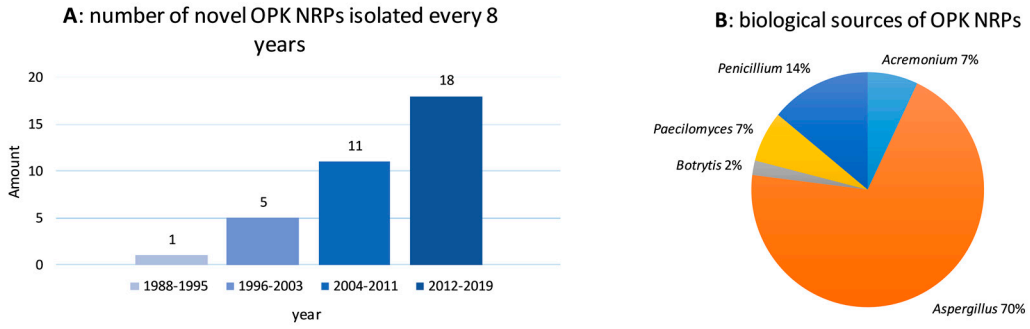


Figure 2. (**A**) number of novel Oxepine-Pyrimidinone-Ketopiperazine NRPs isolated every eight years, (**B**) biological sources of OPK NRPs at the genera level.

A wide range of bioassays have been applied to assess the potential bioactivity of the OPK type of compounds. Though they in general seem to be inactive against human pathogenic strains of *Candia albicans*, some showed potential in treating plant-pathogenic fungi such as *Fusarium graminearum* and *Colletotrichum acutatum*. Notably, protuboxepin A (**16**) has the potential to become a new and effective anticancer drug as it displayed antiproliferative activity by disrupting microtubule dynamics through the tubulin polymerizing in tumor cells [37] despite several other reports showing that OPK compounds did not seem to be active against cancer cells. Antibacterial tests have shown that varioloids A and B (**21**–**22**) exhibited promising activities against several species, while the rest of antibacterial tests did not display antibacterial activity effects. Interestingly, oxepinamides D-G and H–K (**12**–**15**, **32**–**35**) all selectively showed transactivation effects on LXRα, which implied their potential use as novel LXR agonists in the treatment of atherosclerosis, diabetes, and Alzheimer's disease.

Overall, this class of compounds seem to share similar biosynthetic steps to form the OPK backbone, which is likely biosynthesized by the condensation of three amino acids, including one or two anthranilic acids, to form the tricyclic core. Subsequent epoxidation on the benzene ring of the first anthranilic acid residue followed by a ring rearrangement then produces the oxepine moiety. Several successive tailoring reactions can happen before the full construction of the final product(s) (Figure 3). Based on the knowledge of the biosynthesis of fumiquinazolines, and their well documentated proposed biosynthetic pathways, we anticipate that the OPK NRPs biosynthetic gene cluster contains at least a tri-modular NRPS gene with three adenylation domains, including one or two anthranilate-activating domains, and a gene responsible for oxidizing the phenyl moiety of the

anthranilic residue to form the oxepine unit [39,40]. Additionally, an epimerization domain as part of NRPS is needed to convert L-amino acids to D-amino acids in the structures with a D-amino acid residue. Other tailoring genes are also required to encode for OPK related enzymes such as anthranilate synthase, oxidoreductases, and transporters.

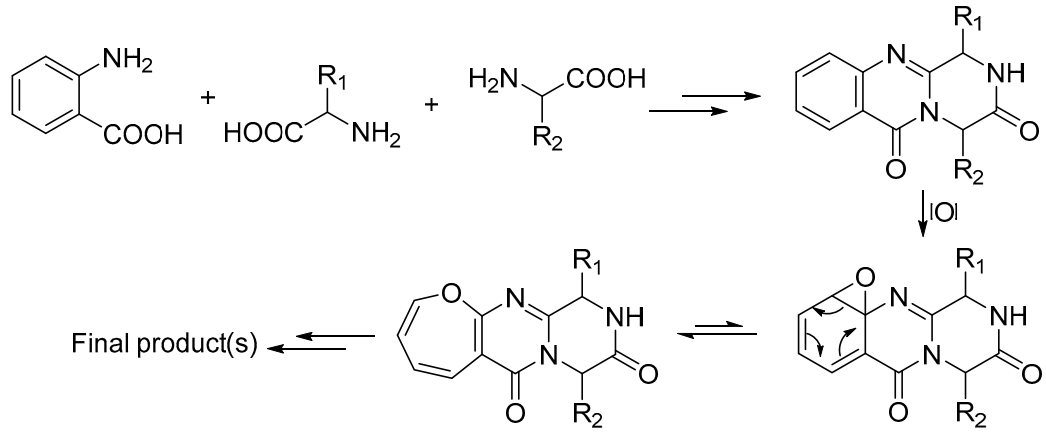


Figure 3. Proposed common biosynthetic steps of Oxepine-Pyrimidinone-Ketopiperazine NRPs.

4. Conclusions

All the OPK NRPs described here were isolated from fungal sources with most compounds reported from species within the families *Aspergillaceae* (*Aspergillus, Penicillium*) and *Trichocomaceae* (*Paecilomyces*). Type A and B compounds share the same 7/6/6 backbone, with the former dominating OPK NRPs with twenty-five reported compounds, while Type C OPKs have a larger 7/6/7/6 backbone with four products reported. In general, these compounds showed promising activities against various phytopathogenic fungi and exhibited transactivation effects on LXRα. In addition, the skeleton of OPK NRPs is likely derived from the condensation of three amino acids, including one or two anthranilic acid(s), and the oxepine moiety is formed by the epoxidation of the benzene ring followed by ring arrangement. However, experimental investigation is needed to support this hypothesis. With the advance of separation skills and spectroscopic techniques, more oxepine-containing compounds are likely to be discovered. Considering that many of these compounds were reported from Aspergilli, ongoing whole genome sequencing of all species in genus *Aspergillus* will possibly set the scene for genomic driven approaches towards new OPK NRPs [46,47].

Author Contributions: Wrote the manuscript, Y.G., J.C.F. and T.O.L.; proof reading, Y.G., J.C.F. and T.O.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Novo Nordisk Foundation (NNF15OC0016610).

Acknowledgments: The authors thank Axel Zeeck at Universität Göttingen for kindly providing the copies of two dissertations of compounds asperloxins A (2) and B (3). Yaojie Guo is financially supported by the China Scholarship Council and Department of Biotechnology and Biomedicine at the Technical University of Denmark.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schwarzer, D.; Finking, R.; Marahiel, M.A. Nonribosomal peptides: From genes to products. *Nat. Prod. Rep.* **2003**, *20*, 275–287. [CrossRef] [PubMed]
2. Cohen, D.J.; Loertscher, R.; Rubin, M.F.; Tilney, N.L.; Carpenter, C.B.; Strom, T.B. Cyclosporine: A new immunosuppressive agent for organ transplantation. *Ann. Intern. Med.* **1984**, *101*, 667–682. [CrossRef] [PubMed]
3. Raja, A.; LaBonte, J.; Lebbos, J.; Kirkpatrick, P. Fresh from the pipeline: Daptomycin. *Nat. Rev. Drug Discov.* **2003**, *2*, 943–944. [CrossRef] [PubMed]

4. Winn, M.; Fyans, J.K.; Zhuo, Y.; Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **2016**, *33*, 317–347. [CrossRef] [PubMed]
5. Niquille, D.L.; Hansen, D.A.; Mori, T.; Fercher, D.; Kries, H.; Hilvert, D. Nonribosomal biosynthesis of backbone-modified peptides. *Nat. Chem.* **2018**, *10*, 282–287. [CrossRef] [PubMed]
6. Cutler, H.G.; Springer, J.P.; Arrendale, R.F.; Arison, B.H.; Cole, P.D.; Roberts, R.G. Cinereain: A novel metabolite with plant growth regulating properties from *Botrytis cinerea*. *Agric. Biol. Chem.* **1988**, *52*, 1725–1733. [CrossRef]
7. He, D.; Wang, M.; Zhao, S.; Shu, Y.; Zeng, H.; Xiao, C.; Lu, C.; Liu, Y. Pharmaceutical prospects of naturally occurring quinazolinone and its derivatives. *Fitoterapia* **2017**, *119*, 136–149. [CrossRef]
8. Li, G.Y.; Yang, T.; Luo, Y.G.; Chen, X.Z.; Fang, D.M.; Zhang, G.L. Brevianamide J, a New indole alkaloid dimer from fungus *Aspergillus versicolor*. *Org. Lett.* **2009**, *11*, 3714–3717. [CrossRef]
9. Li, G.; Li, L.; Yang, T.; Chen, X.; Fang, D. Four New Alkaloids, Brevianamides O – R, from the fungus *Aspergillus versicolor*. *Helv. Chim. Acta* **2010**, *93*, 2075–2080. [CrossRef]
10. Zhang, P.; Li, X.M.; Wang, J.N.; Wang, B.G. Oxepine-containing diketopiperazine alkaloids from the algal-derived endophytic fungus *Paecilomyces variotii* EN-291. *Helv. Chim. Acta* **2015**, *98*, 800–804. [CrossRef]
11. Kshirsagar, U.A. Recent developments in the chemistry of quinazolinone alkaloids. *Org. Biomol. Chem.* **2015**, *13*, 9336–9352. [CrossRef]
12. Hameed, A.; Al-Rashida, M.; Uroos, M.; Ali, S.A.; Arshia; Ishtiaq, M.; Khan, K.M. Quinazoline and quinazolinone as important medicinal scaffolds: A comparative patent review (2011–2016). *Expert Opin. Ther. Pat.* **2018**, *28*, 281–297. [CrossRef]
13. Resende, D.I.S.P.; Boonpothong, P.; Sousa, E.; Kijjoa, A.; Pinto, M.M.M. Chemistry of the fumiquinazolines and structurally related alkaloids. *Nat. Prod. Rep.* **2019**, *36*, 7–34. [CrossRef] [PubMed]
14. Huang, R.M.; Yi, X.X.; Zhou, Y.; Su, X.; Peng, Y.; Gao, C.H. An update on 2,5-Diketopiperazines from marine organisms. *Mar. Drugs* **2014**, *12*, 6213–6235. [CrossRef] [PubMed]
15. Doveston, R.G.; Steendam, R.; Jones, S.; Taylor, R.J.K. Total synthesis of an oxepine natural product, (±)-janoxepin. *Org. Lett.* **2012**, *14*, 1122–1125. [CrossRef]
16. Doveston, R.G.; Taylor, R.J.K. An expedient synthesis of the proposed biosynthetic precursor of the oxepine natural product, janoxepin. *Tetrahedron Lett.* **2012**, *53*, 2533–2536. [CrossRef]
17. Fuchser, J. Beeinflussung der Sekundarstoffbildung bei *Aspergillus ochraceus* durch Variation der Kulturbedingungen sowie Isolierung, Strukturaufklarung und Biosynthese der neuen Naturstoffe. Ph. D. Thesis, Universität Göttingen, Göttingen, Germany, 1996.
18. Michael, B. Pyrralsäuren, eine neue Klasse pyrrolmaskierter Aminosäuren aus *Nocardia* sp. und neue Sekundärmetabolite aus *Aspergillus ochraceus*. Ph. D. Thesis, Universität Göttingen, Göttingen, Germany, 1999.
19. Bode, H.B.; Bethe, B.; Höfs, R.; Zeeck, A. Big effects from small changes: Possible ways to explore nature's chemical diversity. *ChemBioChem* **2002**, *3*, 619–627. [CrossRef]
20. Belofsky, G.N.; Anguera, M.; Jensen, P.R.; Fenical, W.; Köck, M. Oxepinamides A-C and fumiquinazolines H-I: Bioactive metabolites from a marine isolate of a fungus of the genus *Acremonium*. *Chem. Eur. J.* **2000**, *6*, 1355–1360. [CrossRef]
21. Sprogøe, K.; Manniche, S.; Larsen, T.O.; Christophersen, C. Janoxepin and brevicompanine B: Antiplasmodial metabolites from the fungus *Aspergillus janus*. *Tetrahedron* **2005**, *61*, 8718–8721. [CrossRef]
22. Rahbæk, L.; Breinholt, J.; Frisvad, J.C.; Christophersen, C. Circumdatin A, B, and C: Three new benzodiazepine alkaloids isolated from a culture of the fungus *Aspergillus ochraceus*. *J. Org. Chem.* **1999**, *64*, 1689–1692.
23. Ookura, R.; Kito, K.; Ooi, T.; Namikoshi, M.; Kusumi, T. Structure revision of circumdatins A and B, benzodiazepine alkaloids produced by marine fungus *Aspergillus ostianus*, by X-ray crystallography. *J. Org. Chem.* **2008**, *73*, 4245–4247. [CrossRef] [PubMed]
24. Lu, X.H.; Shi, Q.W.; Zheng, Z.H.; Ke, A.B.; Zhang, H.; Huo, C.H.; Ma, Y.; Ren, X.; Li, Y.Y.; Lin, J.; et al. Oxepinamides: Novel liver X receptor agonists from aspergillus puniceus. *Eur. J. Org. Chem.* **2011**, 802–807. [CrossRef]
25. Lee, S.U.; Asami, Y.; Lee, D.; Jang, J.H.; Ahn, J.S.; Oh, H. Protuboxepins A and B and protubonines A and B from the marine-derived fungus *Aspergillus sp.* SF-5044. *J. Nat. Prod.* **2011**, *74*, 1284–1287. [CrossRef] [PubMed]
26. Peng, J.; Zhang, X.Y.; Tu, Z.C.; Xu, X.Y.; Qi, S.H. Alkaloids from the deep-sea-derived fungus *Aspergillus westerdijkiae* DFFSCS013. *J. Nat. Prod.* **2013**, *76*, 983–987. [CrossRef]
27. Zhuravleva, O.I.; Afiyatullov, S.S.; Yurchenko, E.A.; Denisenko, V.A.; Kirichuk, N.N.; Dmitrenok, P.S. New metabolites from the algal associated marine-derived fungus *Aspergillus carneus*. *Nat. Prod. Commun.* **2013**, *8*, 1071–1074. [CrossRef]

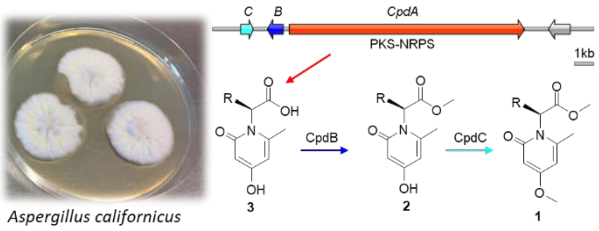
28. Zhang, P.; Mandi, A.; Li, X.M.; Du, F.Y.; Wang, J.N.; Li, X.; Kurtan, T.; Wang, B.G. Varioxepine a, a 3*H*-oxepine-containing alkaloid with a new oxa-cage from the marine algal-derived endophytic fungus *Paecilomyces variotii*. *Org. Lett.* **2014**, *16*, 4834–4837. [CrossRef] [PubMed]
29. Wang, J.; He, W.; Huang, X.; Tian, X.; Liao, S.; Yang, B.; Wang, F.; Zhou, X.; Liu, Y. Antifungal New oxepine-containing alkaloids and xanthones from the deep-sea-derived fungus *Aspergillus versicolor* SCSIO 05879. *J. Agric. Food Chem.* **2016**, *64*, 2910–2916. [CrossRef]
30. Pan, C.; Shi, Y.; Chen, X.; Chen, C.T.A.; Tao, X.; Wu, B. New compounds from a hydrothermal vent crab-associated fungus *Aspergillus versicolor* XZ-4. *Org. Biomol. Chem.* **2017**, *15*, 1155–1163. [CrossRef]
31. Tian, Y.Q.; Lin, S.N.; Zhou, H.; Lin, S.T.; Wang, S.Y.; Liu, Y.H. Protuboxepin C and protuboxepin D from the sponge-derived fungus *Aspergillus* sp. SCSIO XWS02F40. *Nat. Prod. Res.* **2018**, *32*, 2510–2515. [CrossRef]
32. Tian, Y.Q.; Lin, X.P.; Wang, Z.; Zhou, X.F.; Qin, X.C.; Kaliyaperumal, K.; Zhang, T.Y.; Tu, Z.C.; Liu, Y. Asteltoxins with antiviral activities from the marine sponge-derived fungus *Aspergillus* 308 sp. SCSIO XWS02F40. *Molecules* **2016**, *21*, 34. [CrossRef]
33. Xu, W.F.; Mao, N.; Xue, X.J.; Qi, Y.X.; Wei, M.Y.; Wang, C.Y.; Shao, C.L. Structures and absolute configurations of diketopiperazine alkaloids chrysopiperazines A-C from the gorgonian-derived *Penicillium chrysogenum* fungus. *Mar. Drugs* **2019**, *17*, 250. [CrossRef] [PubMed]
34. Luo, X.; Chen, C.; Tao, H.; Lin, X.; Yang, B.; Zhou, X.; Liu, Y. Structurally diverse diketopiperazine alkaloids from the marine-derived fungus: *Aspergillus versicolor* SCSIO 41016. *Org. Chem. Front.* **2019**, *6*, 736–740. [CrossRef]
35. Liang, X.; Zhang, X.; Lu, X.; Zheng, Z.; Ma, X.; Qi, S. Diketopiperazine-type alkaloids from a deep-sea-derived *Aspergillus puniceus* fungus and their effects on Liver X Receptor α. *J. Nat. Prod.* **2019**, *82*, 1558–1564. [CrossRef] [PubMed]
36. Fan, Y.Q.; Li, P.H.; Chao, Y.X.; Chen, H.; Du, N.; He, Q.X.; Liu, K.C. Alkaloids with cardiovascular effects from the marine-derived fungus *Penicillium expansum* Y32. *Mar. Drugs* **2015**, *13*, 6489–6504. [CrossRef]
37. Asami, Y.; Jang, J.H.; Soung, N.K.; He, L.; Moon, D.O.; Kim, J.W.; Oh, H.; Muroi, M.; Osada, H.; Kim, B.Y.; et al. Protuboxepin A, a marine fungal metabolite, inducing metaphase arrest and chromosomal misalignment in tumor cells. *Bioorganic Med. Chem.* **2012**, *20*, 3799–3806. [CrossRef] [PubMed]
38. Ames, B.D.; Walsh, C.T. Anthranilate-activating modules from fungal nonribosomal peptide assembly lines. *Biochemistry* **2010**, *49*, 3351–3365. [CrossRef] [PubMed]
39. Ames, B.D.; Liu, X.; Walsh, C.T. Enzymatic processing of fumiquinazoline F: A tandem oxidative-acylation strategy for the generation of multicyclic scaffolds in fungal indole alkaloid biosynthesis. *Biochemistry* **2010**, *49*, 8564–8576. [CrossRef]
40. Gao, X.; Chooi, Y.H.; Ames, B.D.; Wang, P.; Walsh, C.T.; Tang, Y. Fungal indole alkaloid biosynthesis: Genetic and biochemical investigation of the tryptoquialanine pathway in penicillium aethiopicum. *J. Am. Chem. Soc.* **2011**, *133*, 2729–2741. [CrossRef]
41. Yan, D.; Chen, Q.; Gao, J.; Bai, J.; Liu, B.; Zhang, Y.; Zhang, L.; Zhang, C.; Zou, Y.; Hu, Y. Complexity and diversity generation in the biosynthesis of fumiquinazoline-related peptidyl alkaloids. *Org. Lett.* **2019**, *21*, 1475–1479. [CrossRef]
42. Henderson, A.P.; Mutlu, E.; Leclercq, A.; Bleasdale, C.; Clegg, W.; Henderson, R.A.; Golding, B.T. Trapping of benzene oxide-oxepin and methyl-substituted derivatives with 4-phenyl- and 4-pentafluorophenyl-1,2,4-triazoline-3,5-dione. *Chem. Commun.* **2002**, *8*, 1956–1957. [CrossRef]
43. Samson, R.A.; Houbraken, J.A.M.P.; Kuijpers, A.F.A.; Frank, J.M.; Frisvad, J.C. New ochratoxin A or sclerotium producing species in *Aspergillus* section *Circumdati*. *Stud. Mycol.* **2004**, *50*, 45–61.
44. Chen, A.J.; Frisvad, J.C.; Sun, B.D.; Varga, J.; Kocsubé, S.; Dijksterhuis, J.; Kim, D.H.; Hong, S.B.; Houbraken, J.; Samson, R.A. *Aspergillus* section *Nidulantes* (formerly *Emericella*): Polyphasic taxonomy, chemistry and biology. *Stud. Mycol.* **2016**, *84*. [CrossRef]
45. Takahashi, C.; Matsushita, T.; Doi, M.; Minoura, K.; Shingu, T.; Kumeda, Y.; Numata, A. Fumiquinazolines A-G, novel metabolites of a fungus separated from a *Pseudolabrus* marine fish. *J. Chem. Soc. Perkin Trans. 1* **1995**, 2345–2353. [CrossRef]

46. Vesth, T.C.; Nybo, J.L.; Theobald, S.; Frisvad, J.C.; Larsen, T.O.; Nielsen, K.F.; Hoof, J.B.; Brandl, J.; Salamov, A.; Riley, R.; et al. Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nat. Genet.* **2018**, *50*, 1688–1695. [CrossRef] [PubMed]
47. Kjærbølling, I.; Vesth, T.; Frisvad, J.C.; Nybo, J.L.; Theobald, S.; Kildgaard, S.; Petersen, T.I.; Kuo, A.; Sato, A.; Lyhne, E.K.; et al. A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nat. Commun.* **2020**, *11*. [CrossRef] [PubMed]

Biosynthesis of Calipyridone A Represents a Rare Fungal 2-pyridone Product Formation without Ring Expansion in *Aspergillus californicus*

Yaojie Guo[†], Fabiano J. Contesini[†], Simone Ghidinelli, Uffe H. Mortensen, and Thomas O. Larsen[*]

Supporting Information Placeholder



ABSTRACT: A chemical investigation of the non-model filamentous fungus *Aspergillus californicus* led to the isolation of a polyketide-nonribosomal peptide hybrid calipyridone A (**1**). The related biosynthetic gene cluster was discovered by genome mining, and the biosynthesis of **1** was studied by gene deletion experiments in the host strain. The results indicate that the 2-pyridone moiety of **1** is formed directly from aldol condensation without ring expansion, which is different from the biosynthesis of other fungal 2-pyridone products reported so far. Furthermore, the two *O*-methyltransferases in the biosynthetic pathway show substrate specificity.

Polyketide-nonribosomal peptide (PK-NRP) hybrid molecules represents a diverse class of natural products that are biosynthesized by polyketide synthase-nonribosomal peptide synthetases (PKS-NRPSs).[1] Based on the domain organizations of PKS and NRPS modules, PKS-NRPS systems can be categorized into four types including tethered hybrid modular type, separate hybrid modular type, partially standalone type, and fully standalone type.[2] Besides, a rare NRPS-PKS hybrid enzyme TAS1 was discovered in fungus *Magnaporthe oryzae* whose PKS part contains only a ketosynthase domain.[3] Currently, over 250 tetramate and pyridone based natural products have been discovered from fungi, bacteria and sponges with various bioactivities.[4,5] In fungi, approximately 30 PKS-NRPS involved biosynthetic gene clusters (BGCs) have been studied and linked to their encoding metabolites.[1,6-11] Interestingly, all the 2-pyridone molecules were reported to be biosynthesized from tetramic acid intermediates[5] that have been through enzymatic ring expansion catalyzed by P450 enzymes such as ApdE,[12] TenA,[13] LepH,[14] AsolA,[15] and IliC.[16]

During our ongoing investigation of the natural product chemistry of a poorly studied filamentous fungus *Aspergillus californicus*, a previously unknown PK-NRP compound calipyridone A (**1**) was isolated. **1** is comprised by a 2-pyridone moiety and an indole substructure, the latter likely from the R group of tryptophan residue. Retro-biosynthesis and genome mining using AntiSMASH lead to the discovery of a promising BGC.[17] In this study, we report the structure characterization, gene annotation, and biosynthesis of **1** directly through aldol condensation without a ring rearrangement.

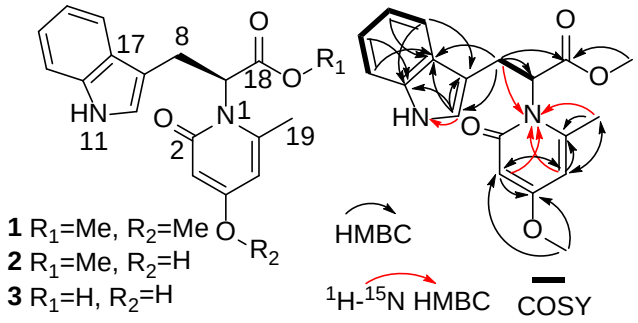


Figure 1. Structures of **1**-**3**, and selected HMBC, ^{1}H-^{15}N HMBC and COSY correlations of **1**.

The molecular formula of **1** was established as $C_{19}H_{20}N_2O_4$ based on HRESIMS data. The ^{1}H NMR spectrum showed 19 hydrogens that could be linked to their corresponding carbons based on ^{13}C NMR and HSQC spectra (**Table S1**). The tryptophan residue was determined by the HMBC correlations (**Figure 1**) from H-8 to C-18 (δ_C 171.9), C-7, C-10, C-17 (δ_C 128.6), H-10 to C-9 (δ_C 111.2), C-12 (δ_C 137.9) and C-17, H-13 to C-17, H-14 to C-12, H-15 to C-17, and from H-16 to C-12, together with two COSY spin systems H-13 through H-16, and H-7 to H-8. The 2-pyridone moiety was characterized by both HMBC and ^{1}H-^{15}N HMBC correlations from H-3 to δ_C 170.2 (C-4), C-5, N-1, and H-5 to C-3, C-6 (δ_C 149.1) and N-1. The keto group at C-2 (δ_C 167.0) was speculated to be vicinal to N-1 by its shielding on the keto group which was similar with

other fungal analogs such as aspyridones A and B, sambutoxin,[18] 4-hydroxy-3-(2, 4-dimethylhexanoyl) 2-pyridone,[19] as well as citridones K and L.[20] CH_3-19 was assigned by HMBC correlations from H-19 to C-5 and C-6, and the methoxyl substitution at C-4 was determined by HMBC correlations from H-20 to C-3 and C-4. The connection between the tryptophan residue and the 2-pyridone unit was solved by the ^{1}H-^{15}N HMBC relationship from H-8 to N-1 which was well supported by two MSMS fragments *m/z* 202 (radical ion) and 140 (adduct ion) from the heterolytic cleavage between N-1 and C7 through a charge retention fragmentation (**Figure S2)**.[21] The configuration of C-7 was determined by electronic circular dichroism (ECD).[22,23] The experimental ECD spectrum showed three main negative features at 203, 219 and 294 nm, which were well reproduced by the calculated ECD spectrum for (7*S*)-**1** (**Figure 2**). Thus, absolute structure of **1** with a 7*S*-configuration was assigned.

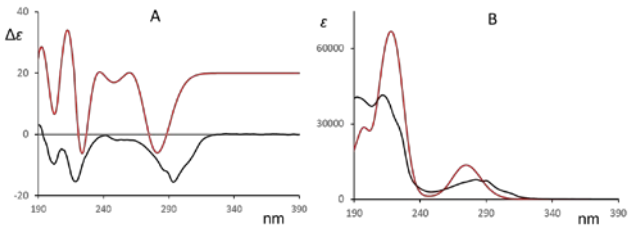


Figure 2. Experimental (black line) and calculated (red line) ECD (A) and UV spectra (B) for (7*S*)-**1**.

The discovery of **1** with a 2-pyridone moiety prompted us to investigate its biosynthesis considering that many fungal analogs showed various biological activities such as cytotoxic, mycotoxic and antifungal activities.[1,7,16] Since we only had **1** at hand, identification of other molecules involved in the pathway would benefit the understanding of each biosynthetic steps. A molecular networking based strategy[24] helped the identification of two related compounds *m/z* 327 (**2**) and *m/z* 313 (**3**) (**Figure S5**). By comparing their MS/MS fragmentation patterns with **1** (**Figure S6-S9)**, we were able to determine their structures: both **2** and **3** shared the same scaffold with **1,** but with different methylation levels. Thus, they were named as calipyridones B (**2**) and C (**3**), respectively.

Recently, the whole genome sequence of *A. californicus* CBS 123895 (= IBT 16748) became available at the Joint Genome Institute.[25] Based on retro-biosynthetic considerations, **1** was proposed to be bio-synthetized from the fusion of tryptophan and a short non-reducing triketide. Therefore, the corresponding BGC was hypothesized to contain one PKS-NRPS hybrid gene whose PKS portion was likely non-reducing and the NRPS module should contain one adenylation domain. Besides, two *O*-methyltransferase (*O*-MT) encoding genes were needed for the addition of two methyl groups.

We next identified BGCs that putatively encode these activities via an antiSMASH analysis.[17] Four BGCs (referred as ´regions´ by antiSMASH) that partially fulfilled the requirements were selected including BGCs annotated as 18.1, 34.3, 73.2 and 120.1. Both two BGCs 34.3 and 73.2 contain only one *C*-MT gene close to the PKS-NRPS gene, and BGC 120.1 has only one *O*-MT gene. Thus, these three BGCs were deprioritized. In contrast, BGC 18.1 on scaffold 18 contains a putative PKS-NRPS gene with a partially reducing PKS portion and a complete NRPS module as well as two *O*-MT genes positioned less than 3 kb upstream of the hybrid PKS-NRPS gene. The PKS module of the hybrid enzyme contains five domains including ketosynthase (KS), acyltransferase (AT), dehydratase (DH), *C*-MT, ketoreductase (KR) and acyl-carrier protein domain (ACP), and the NRPS module contains four domains including condensation domain (C), adenylation domain (A), peptidyl-carrier protein (PCP) and a terminal reductase domain (TD). As these genes encode all the essential elements to synthesize **1**, we envisioned BGC 18.1 as the most promising gene cluster for producing **1** despite that KR, DH, and *C*-MT domains were not expected to be functional (**Figure 3A**).

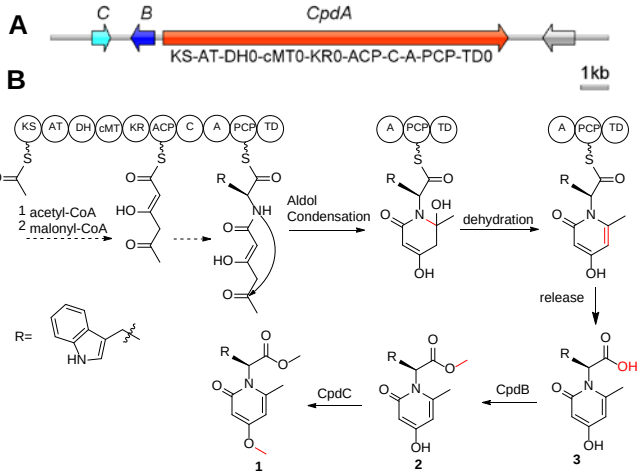


Figure 3. (A) Genetic organization of *cpd* gene cluster, ´0´ after the domain indicating non-functional, drawn with IBS.[26] (B) Proposed biosynthetic steps for **1**.

To investigate the function of each gene of the selected BGC in the biosynthesis of **1**, we individually deleted all the relevant genes. Since *A. californicus* has not been genetically engineered before, we first developed a mutant strain that is more amiable for gene targeting by eliminating the two genes *pyrG* and c*kuA*. In fungi, *pyrG*, encoding orotidine-5'-phosphate decarboxylase, is a classical marker gene that offers selectable and counter-selectable selection that can be used for iterative gene targeting.[27,28] Gene c*kuA* encodes a protein, which is essential for the repair of a DNA double strand break by the non-homologous end-joining (NHEJ) pathway. In the absence of NHEJ, integration of a gene-targeting substrate almost entirely proceeds by homologous recombination, hence, setting the stage for efficient gene targeting.[29–31]

Based on the *A. californicus* mutant (*pyrG*- and c*kuA*Δ, referred as CAL001), we targeted three genes in three individual transformation experiments including the PKS-NRPS gene *cpdA* (gene ID: jgi.p_Aspcalif1_255550) and the two predicted MT encoding genes *cpdB* and *cpdC* (gene IDs: jgi.p_Aspcalif1_222319 and jgi.p_Aspcalif1_222318). The mutant CAL002 (*cpdA*Δ) completely abolished the production of **1**, **2** and **3** (**Figure 4**). In strain CAL003 (*cpdB*Δ) both **1** and **2** disappeared, but the production of **3** could still be observed suggesting that **3** was the precursor of **1** and **2**. In the chemical profile of CAL004 (*cpdC*Δ), the production of **1** was eliminated but the yield of **2** accumulated significantly which indicated that **2** was the precursor of **1**. Therefore, a biosynthetic route was proposed (**Figure 3B**). Precursor **3** was synthesized and released from the assembly line of CpdA, and methylated by CpdB to yield an ester intermediate **2,** which was further methylated by CpdC to yield the final product **1**.

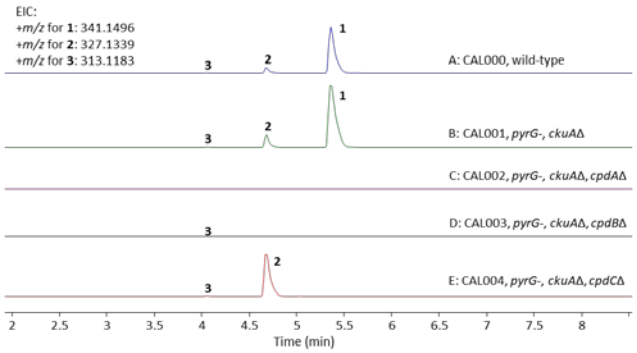


Figure 4. LC-MS analysis of *Aspergillus californicus* wild-type (trace A) and mutants (traces B-E), peaks with a compound number indicating detectable by LC-MS.

In the proposed biosynthetic process, the 2-pyridone moiety is constructed from aldol condensation via a nucleophilic attack of the amine N-1 to the carbonyl C-6. To the best of our knowledge, this is the first fungal metabolite whose 2-pyridone moiety was formed without any ring expansion catalyzed by a P450 enzyme. This is different from the formations of other fungal analogs including tenellin,[32] desmethylbassianin,[33] bassianin,[34] aspyridones A and B,[12] flavipucine and isoflavipucine,[35] leporins A-C,[14] didymellamides A-D,[15,36] protodidymellamide β,[15] and ilicicolins H and J.[16] In addition, the 2-pyridone formation described here also contrasted the formation of kirromycin and factumycin whose 2-pyridone synthesis was driven by Dieckmann cyclases during the product release in actinomycetes.[5,37,38] Before **3** was released from CpdA a double bond between C-5 and C-6 was formed possibly through a spontaneous dehydration during the 2-pyridone ring formation. This spontaneous dehydration was likely driven by the formation of a unsaturated conjugation system, which was similar to the biosynthesis of the flavipucine,[35] kirromycin,[38] and SEK34b.[39] The failure to identify a proposed intermediate with a 6-OH group at *m/z* 359.1601, may suggest that the dehydration reaction happens before release of the hybrid PK-NRPS product.

Both the KR and DH domains of NRPS were not active in the biosynthesis of **1** due to the inconclusive catalytic triads S, Y, and N, and H, G, and P, respectively as predicted by antiSMASH analysis. Similarly, the *C*-MT domain of the PKS module was not functional as no C-methylation was observed during the biosynthesis process. This was supported by the lack of conserved motif GXGXXG in the comparison of its SAM binding site with other *C*-methyltransferases.[40] Inactive *C*-MT domains were also found in the biosynthesis of other PK-NRP hybrids such as PKS3 (xyrrolin),[41] ATEG00325 (flavipucine),[35] CpaA (cyclopiazonic acid).[40] Two known release mechanisms for fungal PK-NRP hybrids were summarized before including reductive release by a catalytic triad of Ser-Tyr-Lys in the terminal reductive domain, and Dieckmann cyclization release.[1,6] In this study, the release of **1** is different from both two as neither reduction nor cyclization was observed.

The two tailoring *O*-MTs seem to be substrate-specific as no compensations were observed between mutants CAL003 and CAL004. Recent studies showed that the substrate specificities of several plant *O*-MTs were determined by key residues like F119, W123, I114, F140 and F162.[42] Engineering of the key residues of CpdB and CpdC might expand their substrate range. The gene (ID: jgi.p_Aspcalif1_255551) 1.2 kb downstream of *cpdA* did not seem to be involved in the biosynthesis of **1**.

In conclusion, calipyridone A (**1**), a new PK-NRP hybrid, was characterized from *A. californicus*. Furthermore, the two related precursor molecules calipyridones B (**2**) and C (**3**) could be discovered by molecular networking analysis, and their structures could be tentatively proposed based on MS/MS fragmentation pattern analysis. The detailed biosynthetic steps of **1** were elucidated by multiple gene deletion experiments in the native producer. Albeit having a relatively simple structure, **1** is the first fungal 2-pyridone metabolite that was formed without a ring expansion reaction of a tetramic acid intermediate catalyzed by a P450 enzyme.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.
Experimental procedures, spectroscopic data, supplementary figures and tables.

AUTHOR INFORMATION

Corresponding Author

***Thomas O. Larsen**: E-mail: tol@bio.dtu. dk.

Authors

Yaojie Guo - Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

Fabiano J. Contesini - Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

Simone Ghidinelli - Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy

Uffe H. Mortensen - Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

Thomas O. Larsen - Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

Author Contributions

[†]Y.G. and F.J.C. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank Dr. Kasper Enemark-Rasmussen from DTU Chemistry for acquiring NMR data and Emil H. Kristiansen for chemical extraction. Mercedes de la Cruz, Thomas A. Mackenzie, Carmen R. Martín and Pilar Sánchez from Fundación MEDINA, Spain are acknowledged for the antibacterial and cytotoxic tests. The authors thank the Novo Nordisk Foundation for their grant support (NNF15OC0016610). Y.G. is a scholarship recipient (201709110107) from the China Scholarship Council.

REFERENCES

(1) Boettger, D.; Hertweck, C. Molecular Diversity Sculpted by Fungal PKS-NRPS Hybrids. *ChemBioChem* **2013**, *14* (1), 28–42.

(2) Miyanaga, A.; Kudo, F.; Eguchi, T. Protein-Protein Interactions in Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrid Assembly Lines. *Nat. Prod. Rep.* **2018**, *35* (11), 1185–1209.

(3) Yun, C. S.; Motoyama, T.; Osada, H. Biosynthesis of the Mycotoxin Tenuazonic Acid by a Fungal NRPS-PKS Hybrid Enzyme. *Nat. Commun.* **2015**, *6*.

(4) Mo, X.; Li, Q.; Ju, J. Naturally Occurring Tetramic Acid Products: Isolation, Structure Elucidation and Biological Activity. *RSC Adv.* **2014**, *4* (92), 50566–50593.

(5) Gui, C.; Li, Q.; Mo, X.; Qin, X.; Ma, J.; Ju, J. Discovery of a New Family of Dieckmann Cyclases Essential to Tetramic Acid and Pyridone-Based Natural Products Biosynthesis. *Org. Lett.* **2015**, *17* (3), 628–631.

(6) Fisch, K. M. Biosynthesis of Natural Products by Microbial Iterative Hybrid PKS-NRPS. *RSC Adv.* **2013**, *3* (40), 18228–18247.

(7) HE, J.; NIU, X.; YANG, X. Research Progress on Fungal PKS-NRPS Hybrid Metabolites. *Sci. Sin. Vitae* **2019**, *49* (7), 848–864.

(8) Li, H.; Gilchrist, C. L. M.; Lacey, H. J.; Crombie, A.; Vuong, D.; Pitt, J. I.; Lacey, E.; Chooi, Y. H.; Piggott, A. M. Discovery and Heterologous Biosynthesis of the Burnettramic Acids: Rare PKS-NRPS-Derived Bolaamphiphilic Pyrrolizidinediones from an Australian Fungus, *Aspergillus burnettii*. *Org. Lett.* **2019**, *21* (5), 1287–1291.

(9) Wolff, P. B.; Nielsen, M. L.; Slot, J. C.; Andersen, L. N.; Petersen, L. M.; Isbrandt, T.; Holm, D. K.; Mortensen, U. H.; Nødvig, C. S.; Larsen, T. O.; et al. Acurin A, a Novel Hybrid Compound, Biosynthesized by Individually Translated PKS- and NRPS-Encoding Genes in *Aspergillus aculeatus*. *Fungal Genet. Biol.* **2020**, *139* (*2020*), 103378.

(10) Tang, S.; Zhang, W.; Li, Z.; Li, H.; Geng, C.; Huang, X.; Lu, X. Discovery and Characterization of a PKS-NRPS Hybrid in *Aspergillus terreus* by Genome Mining. *J. Nat. Prod.* **2020**, *83* (2), 473–480.

(11) Kato, S.; Motoyama, T.; Futamura, Y.; Uramoto, M.; Nogawa, T.; Hayashi, T.; Hirota, H.; Tanaka, A.; Takahashi-Ando, N.; Kamakura, T.; et al. Biosynthetic Gene Cluster Identification and Biological Activity of Lucilactaene from *Fusarium* sp. RK97-94. *Biosci. Biotechnol. Biochem.* **2020**, *84* (6), 1303–1307.

(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* (4), 213–217.

(13) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. Late Stage Oxidations during the Biosynthesis of the 2-Pyridone Tenellin in the Entomopathogenic Fungus *Beauveria bassiana*. *J. Am. Chem. Soc.* **2008**, *130* (52), 17988–17996.

(14) Cary, J. W.; Uka, V.; Han, Z.; Buyst, D.; Harris-Coward, P. Y.; Ehrlich, K. C.; Wei, Q.; Bhatnagar, D.; Dowd, P. F.; Martens, S. L.; et al. An *Aspergillus flavus* Secondary Metabolic Gene Cluster Containing a Hybrid PKS-NRPS is Necessary for Synthesis of the 2-Pyridones, Leporins. *Fungal Genet. Biol.* **2015**, *81*, 88–97.

(15) Ugai, T.; Minami, A.; Gomi, K.; Oikawa, H. Genome Mining Approach for Harnessing the Cryptic Gene Cluster in *Alternaria solani*: Production of PKS-NRPS Hybrid Metabolite, Didymellamide B. *Tetrahedron Lett.* **2016**, *57* (25), 2793–2796.

(16) Lin, X.; Yuan, S.; Chen, S.; Chen, B.; Xu, H.; Liu, L.; Li, H.; Gao, Z. Heterologous Expression of Ilicicolin H Biosynthetic Gene Cluster and Production of a New Potent Antifungal Reagent, Ilicicolin J. *Molecules* **2019**, *24* (12), 1–10.

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

(18) Kim, J. C.; Lee, Y. W.; Tamura, H.; Yoshizawa, T. Sambutoxin: A New Mycotoxin Isolated from *Fusarium sambucinum*. *Tetrahedron Lett.* **1995**, *36* (7), 1047–1050.

(19) 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* (10), 3845–3856.

(20) Yan, T.; Ding, W.; Liu, H.; Wang, P.-M.; Zheng, D.; Xu, J. New Pyridone Alkaloids from Marine-Derived Fungus *Penicillium* sp. *Tetrahedron Lett.* **2020**, 61 (19), 151843.

(21) Demarque, D. P.; Crotti, A. E. M.; Vessecchi, R.; Lopes, J. L. C.; Lopes, N. P. Fragmentation Reactions Using Electrospray Ionization Mass Spectrometry: An Important Tool for the Structural Elucidation and Characterization of Synthetic and Natural Products. *Nat. Prod. Rep.* **2016**, *33* (3), 432–455.

(22) Mazzeo, G.; Santoro, E.; Andolfi, A.; Cimmino, A.; Troselj, P.; Petrovic, A. G.; Superchi, S.; Evidente, A.; Berova, N. Absolute Configurations of Fungal and Plant Metabolites by Chiroptical Methods. ORD, ECD, and VCD Studies on Phyllostin, Scytolide, and Oxysporone. *J. Nat. Prod.* **2013**, *76* (4), 588–599.

(23) Salib, M. N.; Molinski, T. F. Six Trikentrin-like Cyclopentanoindoles from Trikentrion Flabelliforme. Absolute Structural Assignment by NMR and ECD. *J. Org. Chem.* **2018**, *83* (3), 1278–1286.

(24) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837.

(25) Nordberg, H.; 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* (D1), 26–31.

(26) Liu, W.; Xie, Y.; Ma, J.; Luo, X.; Nie, P.; Zuo, Z.; Lahrmann, U.; Zhao, Q.; Zheng, Y.; Zhao, Y.; et al. IBS: An Illustrator for the Presentation and Visualization of Biological Sequences. *Bioinformatics* **2015**, *31* (20), 3359–3361.

(27) Ballance D.J., Buxton F.P., T. G. Transformation of Aspergillus Nidulans by the Orotidine-5´-Phos- Phate Decarboxylase Gene of *Neurospora crassa*. *Biochem. Biophys. Res. Commun.* **1983**, *112* (1), 284–289.

(28) Boeke, J. D.; La Croute, F.; Fink, G. R. A Positive Selection for Mutants Lacking Orotidine-5′-Phosphate Decarboxylase Activity in Yeast: 5-Fluoro-Orotic Acid Resistance. *MGG Mol. Gen. Genet.* **1984**, *197* (2), 345–346.

(29) Meyer, V.; Arentshorst, M.; El-Ghezal, A.; Drews, A. C.; Kooistra, R.; van den Hondel, C. A. M. J. J.; Ram, A. F. J. Highly Efficient Gene Targeting in the *Aspergillus niger* KusA Mutant. *J. Biotechnol.* **2007**, *128* (4), 770–775.

(30) Zhang, J.; Mao, Z.; Xue, W.; Li, Y.; Tang, G.; Wang, A.; Zhang, Y.; Wang, H. Ku80 Gene Is Related to Non-Homologous End-Joining and Genome Stability in *Aspergillus niger*. *Curr. Microbiol.* **2011**, *62* (4), 1342–1346.

(31) Jakob B. Hoof, Christina S. Nødvig, and U. H. M. Chapter 11 Genome Editing: CRISPR-Cas9. In *Fungal Genomics*; 2018; Vol. 1775, pp 119–132.

(32) 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* (3), 289–297.

(33) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; et al. The Programming Role of Trans-Acting Enoyl Reductases during the Biosynthesis of Highly Reduced Fungal Polyketides. *Chem. Sci.* **2011**, *2* (5), 972–979.

(34) 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* (41), 16635–16641.

(35) Gressler, M.; Zaehle, C.; Scherlach, K.; Hertweck, C.; Brock, M. Multifactorial Induction of an Orphan PKS-NRPS Gene Cluster in *Aspergillus terreus*. *Chem. Biol.* **2011**, *18* (2), 198–209.

(36) Haga, A.; Tamoto, H.; Ishino, M.; Kimura, E.; Sugita, T.; Kinoshita, K.; Takahashi, K.; Shiro, M.; Koyama, K. Pyridone Alkaloids from a Marine-Derived Fungus, *Stagonosporopsis cucurbitacearum*, and Their Activities against Azole-Resistant

Candida albicans. *J. Nat. Prod.* **2013**, *76* (4), 750–754.

(37) Weber, T.; Laiple, K. J.; Pross, E. K.; Textor, A.; Grond, S.; Welzel, K.; Pelzer, S.; Vente, A.; Wohlleben, W. Molecular Analysis of the Kirromycin Biosynthetic Gene Cluster Revealed β-Alanine as Precursor of the Pyridone Moiety. *Chem. Biol.* **2008**, *15* (2), 175–188.

(38) Thaker, M. N.; García, M.; Koteva, K.; Waglechner, N.; Sorensen, D.; Medina, R.; Wright, G. D. Biosynthetic Gene Cluster and Antimicrobial Activity of the Elfamycin Antibiotic Factumycin. *Medchemcomm* **2012**, *3* (8), 1020–1026.

(39) Deng, M. R.; Li, Y.; He, H. H.; Zhou, X.; Zheng, X. L.; Wang, Y. H.; Zhu, H. An Aberrant Metabolic Flow toward Early Shunt Products in the Granaticin Biosynthetic Machinery of *Streptomyces vietnamensis* GIMV4.0001. *J. Antibiot. (Tokyo).* **2020**, *73* (4), 260–264.

(40) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Functional Expression of the *Aspergillus flavus* PKS-NRPS Hybrid CpaA Involved in the Biosynthesis of Cyclopiazonic Acid. *Bioorganic Med. Chem. Lett.* **2009**, *19* (12), 3288–3292.

(41) Phonghanpot, S.; Punya, J.; Tachaleat, A.; Laoteng, K.; Bhavakul, V.; Tanticharoen, M.; Cheevadhanarak, S. Biosynthesis of Xyrrolin, a New Cytotoxic Hybrid Polyketide/Non-Ribosomal Peptide Pyrroline with Anticancer Potential, in *Xylaria* sp. BCC 1067. *ChemBioChem* **2012**, *13* (6), 895–903.

(42) Bennett, M. R.; Shepherd, S. A.; Cronin, V. A.; Micklefield, J. Recent Advances in Methyltransferase Biocatalysis. *Curr. Opin. Chem. Biol.* **2017**, *37*, 97–106.

Supporting information for

Biosynthesis of Calipyridone A Represents a Rare Fungal 2-pyridone Product Formation without Ring Expansion in *Aspergillus californicus*

Yaojie Guo,[a,†] Fabiano J. Contesini,[a,†] Simone Ghidinelli,[b] Uffe H. Mortensen,[a] and Thomas O. Larsen [a,*]

[a] Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark;

[b] Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123, Brescia, Italy

[†]Y.G. and F.J.C. contributed equally.

Table of Contents

Experimental procedures.

General.

Optical rotations were measured on a PerkinElmer 341 polarimeter. ECD and UV spectra were recorded in acetonitrile on a JASCO J-1500 CD spectrophotometer with a 2 mm path length cuvette. The NMR experiments were performed on Bruker AVANCE III 800 MHz system equipped with a 5 mm TCI cryoprobe using standard pulse sequences. UHPLC-DAD-HRMS was acquired on an Agilent 1290-6545 UHPLC-QTOF-MS system equipped with a diode array detector. Separation was performed on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 150 mm, 1.9 μm) at 60 °C with a linear gradient program: mobile phase (MP) A H_2O and B MeCN, both buffered with 20 mM formic acid, at a flow rate of 0.35 mL/min: 0-10.0 min 10%-100% MPB, 10.0-12.0 min 100% MPB, 12.1-14.0 min 10% MPB. Flash chromatography was carried out on a Biotage Isolera One system. Semi-preparative HPLC was performed on a Waters 600 HPLC with a photodiode array detector. The purification column was Phenomenex Kinetex C_{18} column (250 × 10 mm, 5 μm). For extraction and separation the solvents were HPLC grade, and for HRMS the chemicals were LCMS grade.

Fungal and bacterial strains.

Aspergillus californicus IBT 16748 (= CBS 123895, both are ex- type cultures of *Aspergillus californicus*) was from IBT culture collection, Department of Biotechnology and Biomedicine, Technical University of Denmark. The fungal strains were cultivated on minimal media (MM), transformation medium (TM) or Czapek yeast extract agar (CYA) plates supplemented with both uridine and uracil at 10 mM, and 1.3 mg/ml 5-fluoroorotic acid when necessary. MM and TM were described by Nødvig et al,[1] and CYA media was prepared as described by Samson et al.[2] *Escherichia coli* DH5α was used as a host for propagating all plasmids using solid (2% agar) or liquid LB medium with 100 μg/ml ampicillin.

*Isolation of **1**.*

A. californicus IBT16748 was cultivated on 120 CYA plates and 120 YES plates for 11 days at 25℃. All the plates were extracted twice with ethyl acetate with 1% formic acid (FA). The crude extract was dissolved in 90% MeOH, extracted with heptane, and further diluted to 50% MeOH before being extracted with dichloromethane (DCM). The DCM fraction was applied to the flash

chromatography with a self-packed NH_2 column (Sepra Phenomenex, 50g, 66 mL) to yield eight fractions. Fr4 (80.4 mg) was subjected to the flash chromatography with a self-packed C_{18}-T column (Sepra Phenomenex, 25 g, 33 mL) to give eight sub-fractions. Sub-fr3 (6.87 mg) was applied to the C_{18} column (MeCN/H_2O, 35-40%, 20 min, 4 mL/min) to yield compound **1** (1.15 mg, t_R = 12.3 min).

Computational details.

Conformational analysis was performed at Molecular Mechanics (MM) level using Schrödinger suite. Conformers have been optimized at B3LYP/TZVP level in IEF-PCM approximation[3] using Gaussian 16 package.[4] Dipole and rotational strengths for the first 80 excited states have been calculated using time-dependent density functional theory (TDDFT) formalism with CAM-B3LYP as functional and TZVP as basis sets in IEF-PCM approximation. The calculated UV and ECD were generated using Gaussian functions with bandwidth of 0.20 eV.[5,6] A red shift of 20 nm has been applied to the calculated spectra.

Molecular networking details.

The HRMSMS data for networking were from the extract of CAL001. Data were converted to mzML format using MSConvert.[7] The processed data were applied to the Global Natural Products Social Molecular Networking (GNPS) platform using the default settings except for the precursor and fragment ion tolerances both as 0.005.[8] The output networking data were visualized using Cytocape.[9] The results for the molecular networking can be accessed at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=fc8728f4f2c2438e9b83be455526df36.

Genomic DNA, plasmid construction for fungal transformations.

Primers were designed based on the genome sequence of *A. californicus* CBS 123895 at JGI genome portal (https://genome.jgi.doe.gov/portal/). All primers (Integrated DNA Technology, Belgium) used for sequencing and PCR amplification are listed in **Table S2**. All plasmids were constructed using USER cloning and USER fusion.[10] Pfu X7 DNA polymerase[11] or Phusion U hot start DNA polymerase (Thermo Fisher Scientific) were used to generate each fragment for vector construction using pFC902 vector as template. The amplified fragments were purified on agarose gel (1% or 2%) and cloned into the pFC330 vector previously digested with PacI and Nt.BbvcI enzymes (New England Biolabs).[12] Single-stranded repair oligonucleotides (composed of 45

nucleotides upstream and 45 downstream of the target gene) were obtained from Integrated DNA Technology. The *pyrG* and c*kuA* deletion cassettes were constructed by amplifying the 3'UTR and 5'UTR (~2 kb) region of both genes and cloned into the pU2002 vector. Sequencing was performed by Macrogen.

Protoplastation and transformation.

Fungal protoplasts preparation and transformation were performed as described by Nielsen et al.[13] Plasmid pCas9-hyg-pyrgku5BB was used for the transformation of *A. californicus* wild-type (WT) for the deletion of *pyrG* (scaffold 37) and c*kuA* (scaffold 11) genes. The gene-targeting substrates pU2002-pyrgupdw and pU2002-kuupdw were digested by SwaI (New England Biolabs) before use. Twenty-eight candidate transformants were picked from the transformation plate, and three (col8, 12, and 28) showed correct band lengths by PCR. The spores of col12 were collected for monosporic purification and four colonies (col12 1-4) from single spores were diagnosed by PCR. The genomic DNA (gDNA) of col12-1 (= CAL001) was isolated via FastDNA SPIN Kit for Soil DNA extraction kit (MP Biomedicals, USA) and sequenced. For the deletion of the biosynthetic genes, single-stranded oligonucleotides were used as repair templates. Detailed transformation steps have been described by ref.[1,14]

Antibacterial and cytotoxicity tests.

Antibacterial and cytotoxic tests were performed by Fundación MEDINA, Spain. **1** didn´t show activities against two Gram-positive bacteria (methicillin-resistant *Staphylococcus aureus* MB5393 and methicillin-sensitive *Staphylococcus aureus* ATCC 29213) and two Gram-negative bacteria (*Escherichia coli* ATCC 25922 and *Klebsiella pneumoniae* ATCC 700603) with all MICs above 96 µg/ml. In the cytotoxic assays against six tumor cell lines A549 (lung), A2058 (skin), HepG2 (liver), MCF-7 (breast), Mia PaCa-2 (pancreas), HL-60 (leukemia), **1** didn´t show activities with IC_{50} values over 30 µg/ml.

*Physiochemical data of **1**.*

Calipyridone A (**1**): colorless amorphous $[\alpha]_D^{20}$= -59.7 (c = 0.067, Acetonitrile); UV (MeCN) λ_{max} (log ε) 292 (3.86), 210 (4.61) nm; ECD (1.97 mM, MeCN) λ_{max} ($\Delta\varepsilon$) 294 (-15.4), 219 (-15.2), 203 (-9.4) nm; ^{1}H and ^{13}C data see **Table 1**. HRESIMS m/z $[M+H]^+$ calcd for $C_{19}H_{21}N_2O_4^+$, 341.1496, found 341.1501; $[M+Na]^+$ calcd for $C_{19}H_{20}N_2O_4Na^+$, 363.1315, found 363.1311.

Table S1. NMR Data (800 MHz ^{1}H; 200 MHz ^{13}C) of **1** in CD_3OD.

Position	δ_C (or δ_N), type	δ_H, multi (*J* in Hz)
1	121.8, N	
2	167.0, C	
3	95.5, CH	5.84, d (2.7)
4	170.2, C	
5	102.6, CH	5.61, d (2.7)
6	149.1, C	
7	61.4, CH	4.97, m
8	24.9, CH_2	a, 3.72, dd (15.1, 10.8) b, 3.69, ddd (15.1, 3.8, 0.6)
9	111.2, C	
10	124.6, CH	6.81, s
11	77.2[a], NH	
12	137.9, C	
13	112.4, CH	7.29, ddd (8.1, 1.0, 0.8)
14	122.5, CH	7.06, ddd (8.1, 7.0, 1.1)
15	120.0, CH	6.94,ddd (8.0, 7.0, 1.0)
16	118.7, CH	7.41, ddd (8.0, 1.1, 0.8)
17	128.6, C	
18	171.9, C	
19	20.5, CH_3	1.56, s
20	56.1, CH_3	3.76, s
21	52.8, CH_3	3.74, s

[a]Nitrorgen chemical shifts were from ^{1}H-^{15}N HMBC.

Table S2. Primers used in this study.

Primers/90-mers	**Sequences (5' to 3')**
Acal.outPyrG.F	CGGAGCCAAAATTCCCCGAT
Acal.outPyrG.R	CGTCTATCCACACGGCAGTC
Acal.outku70F	TGCCCTCAAAATGGCTGACG
Acal.midku70F	AAGCATTAACTCCAAGTCCGTACC
Acal.outku70R	GCCACTGGAGGCAGCTATCT
Acal.up.pyrG.F	GGGTTTAAUGTCCGTATCCCAGCAACATCAC
Acal.up.pyrG.R	AAACCCTUGGGCTCGATATCGGCACAGAAA GCAAAGATCCCCACCATGACC
Acal.down.pyrG.F	AAGGGTTUCACCAGTTGATCAGTTGAGGGCGG ACTCGCAACTGGATCATGGGT
Acal.down.pyrG.R	GGTCTTAAUGGATGACTCCATCAGTTGCTGC
Acal.up.ku70.F	GGGTTTAAUATACGACACAAGGAGCGGTCTC
Acal.up.ku70.R	AAACCCTUGGGCTCGATATCGGCACAGAAA GGTACACTACGCACGGGTGA
Acal.down.ku70.F	AAGGGTTUCGCCCTTCCCATCTCAAGATGCGG ACCGCCTGATATACCAGCCG
Acal.down.ku70.R	GGTCTTAAUTCTCTCGTCCATTTCTCCCGC
Acal.pyrG.PS1.R	ATCGATGAATTUGCGGTC TGCATCATCCGTGAATCGAAC
Acal.pyrG.PS1.F	AAATTCATCGAUAT GTTTTAGAGCTAGAAATAGCAAG
Acal.pyrG.PS2.R	AGGACTCACCUCC TGCATCATCCGTGAATCGAAC
Acal.pyrG.PS2.F	AGGTGAGTCCUCAGCCGG GTTTTAGAGCTAGAAATAGCAAG
Acal.ku70.PS1.R	ATTGGATATGAUTCGT TGCATCATCCGTGAATCGAAC
Acal.ku70.PS1.F	ATCATATCCAAUCCCT GTTTTAGAGCTAGAAATAGCAAG
Acal.ku70.PS2.R	AATGCCATUGAAGCGCTA TGCATCATCCGTGAATCGAAC
Acal.ku70.PS2.F	AATGGCATUAT GTTTTAGAGCTAGAAATAGCAAG
Acal.pyr255550.F1	CGTCCTGTCAGAGAAACACAGC
Acal.pyr255550.R1	CAGGGCGAGTCTCAGTGGTT
Acal.pyr255550.F2	ATGAAGGACGATGAGGACGAGG
Acal.pyr255550.R2	GCGTTAGGTCCTCGGGAATGT
Acal.pyr255550.PS1.R	ATCGCAATAGGUTCAGT TGCATCATCCGTGAATCGAAC
Acal.pyr255550.PS1.F	ACCTATTGCGATAAT GTTTTAGAGCTAGAAATAGCAAG
Acal.pyr255550.PS2.R	ACCAGTCCAUGAAC TGCATCATCCGTGAATCGAAC
Acal.pyr255550.PS2.F	ATGGACTGGTCGATTA GTTTTAGAGCTAGAAATAGCAAG

Acal.pyr255550.90nt	TAATATATAGTCTGATGCTTGAGAAATTGCCCTGTTGCATAAACGATGAA GGTTCTGCTCCCGGTTTATGGGGTCTACTATATATCTATA
Acal.pyrGgup.out.F	TTCTGGACTCGCTGGGACAA
Acal.pryGdw.out.R	GACAGAAGAGGTCAGGTGGAGA
Acal.outku.F	CGGTCCACTCTGACCCAATCA
Acal.outku.R	CCGATAGGTCATGGTTGGTGGT
Acal.kuup.out.F	GCGATTGGCAAACGATGGGT
Acal.kudw.out.R	ACTTCTCCAGTCCCCTCTTGC
Acal.pyr222318.out.F	AGATTTGTCACTTGGCGTGAGC
Acal. pyr222318.out.R	AGTCAACCTTATCTGACAGGTGGT
Acal.pyr222318.PS1.R	ATTTGCCCUGGGGTGTG TGCATCATCCGTGAATCGAAC
Acal. pyr222318.PS1.F	AGGGCAAAUGCAGTTTTAGAGCTAGAAATAGCAAG
Acal. pyr222318.PS2.R	AAGGCTTCUATCA TGCATCATCCGTGAATCGAAC
Acal. pyr222318.PS2.F	AGAAGCCTUACATCATGTTTTAGAGCTAGAAATAGCAAG
Acal. pyr222318.90nt	TAATACTGATCGAGTATTAATCATCGCTAGGACGATTAATGCGATCATGG GAATGACTATTATAATGAGGTGCTTAACCAATGTAAGTGG
Acal.pyr222319.out.F	ACCACCTGTCAGATAAGGTTGACT
Acal. pyr222319.out.R	GCAACAGGGCAATTTCTCAAGCA
Acal.pyr222319.PS1.R	ATAACTATACUAATATTGCATCATCCGTGAATCGAAC
Acal. pyr222319.PS1.F	AGTATAGTTAUAGTAGTTTTAGAGCTAGAAATAGCAAG
Acal. pyr222319.PS2.R	ACCACCGCCUACCCTCATGCATCATCCGTGAATCGAAC
Acal. pyr222319.PS2.F	AGGCGGTGGUGGTGTTTTAGAGCTAGAAATAGCAAG
Acal. pyr222319.90nt	TAACTAGATTATGTTCCAAAGATAACAGGTACTAAATAAAACTAAAGTGA CTGGTCCTCAGATTGGGCGTCGCCGTCGTCGCGTTAGAAC
CSN438	GGGTTTAAUGATCACATAGATGCTCGGTTGACA
CSN790	GGTCTTAAUACCCTGAGAAGATAGATGTGAATGTG

Table S3. Plasmids constructed in this study.

Plasmids	**Backbones**	**Primers for PCR amplification**	**PCR templates**
pCas9-hyg-pyrgku	pFC332	CSN438/Acal.pyrG.PS1.R Acal.pyrG.PS1.F/Acal.pyrG.PS2.R Acal.pyrG.PS2.F/Acal.ku70.PS1.R Acal.ku70.PS1.F/Acal.ku70.PS2.R Acal.ku70.PS2.F/CSN790	pFC902
pU2002-pyrgupdw	pU2002c	Acal.up.pyrG.F/Acal.up.pyrG.R Acal.down.pyrG.F/Acal.down.pyrG.R	gDNA
pU2002- kuupdw	pU2002c	Acal.up.ku70.F/Acal.up.ku70.R Acal.down.ku70.F/Acal.down.ku70.R	gDNA
pCas9-pyrg-pyr50	pFC330	CSN438/Acal.pyr255550.PS1.R Acal.pyr255550.PS1.F/Acal.pyr255550.PS2.R Acal.pyr255550.PS2.F/CSN790	pFC902
pCas9-pyrg-pyr18	pFC330	CSN438/Acal.pyr222318.PS1.R Acal. pyr222318.PS1.F/Acal. pyr222318.PS2.R Acal. pyr222318.PS2.F/CSN790	pFC902
pCas9-pyrg-pyr19	pFC330	CSN438/Acal.pyr222319.PS1.R Acal. pyr222319.PS1.F/Acal. pyr222319.PS2.R Acal. pyr222319.PS2.F/CSN790	pFC902

Table S4. Mutant strains of *Aspergillus californicus* constructed in this study.

Strains	**Genotype**
CAL000	Wild type (WT)
CAL001	pyrG-, ku70Δ
CAL002	pyrG-, ku70Δ, CpdAΔ
CAL003	pyrG-, ku70Δ, CpdBΔ
CAL004	pyrG-, ku70Δ, CpdCΔ

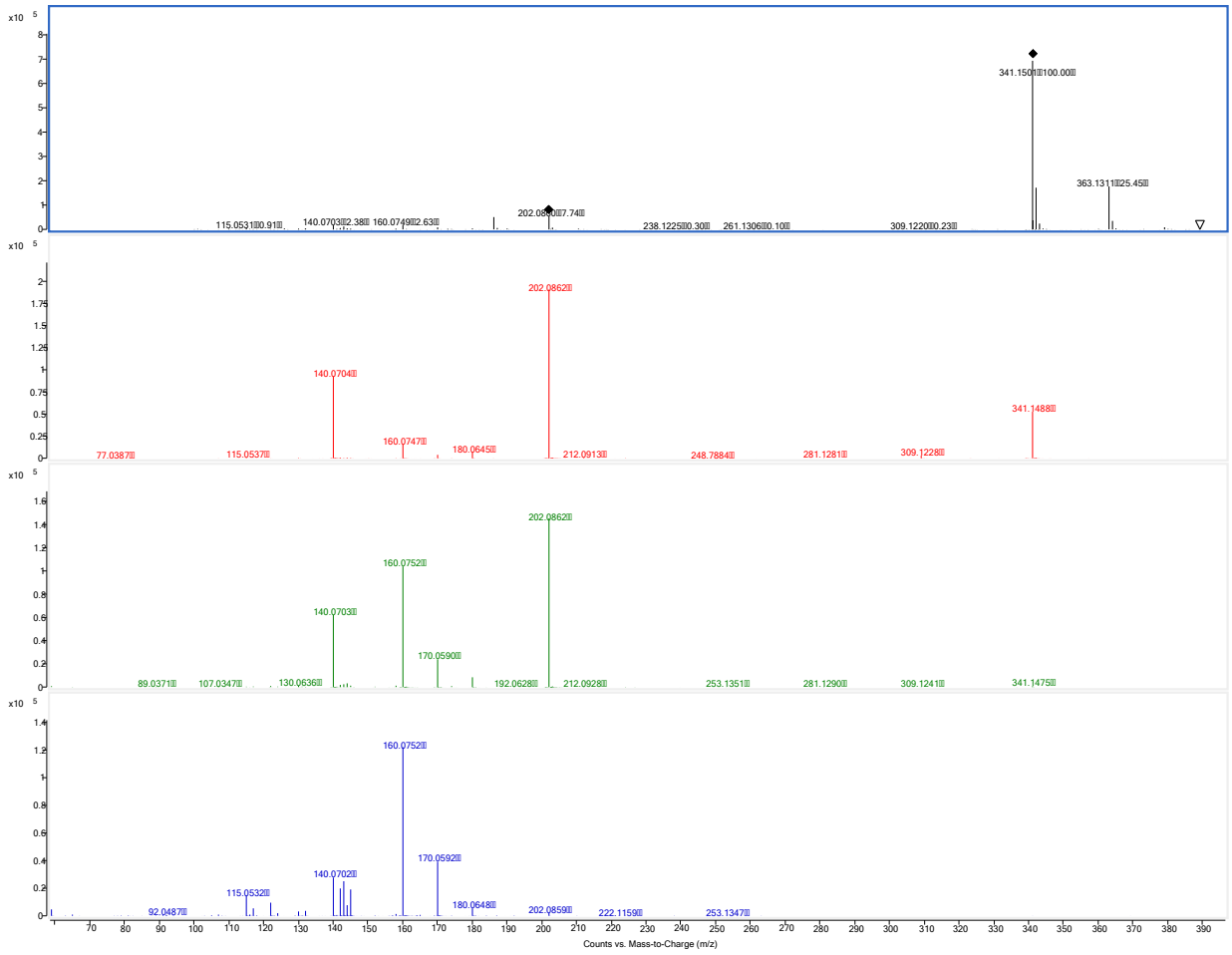


Figure S1. HRESIMS and MSMS (10, 20 and 40 ev) spectra of **1**.

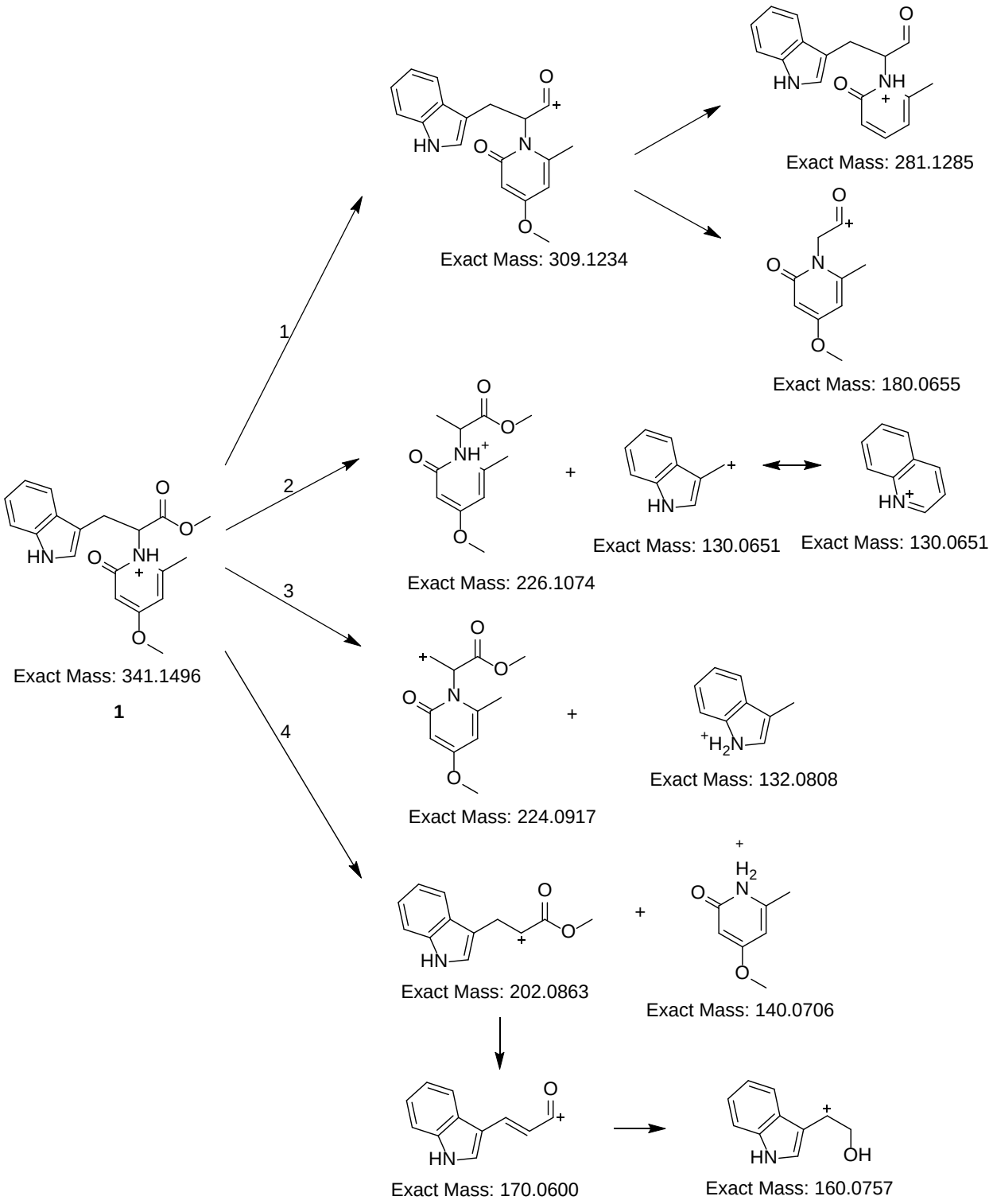


Figure S2. Proposed MSMS fragmentation of **1**.

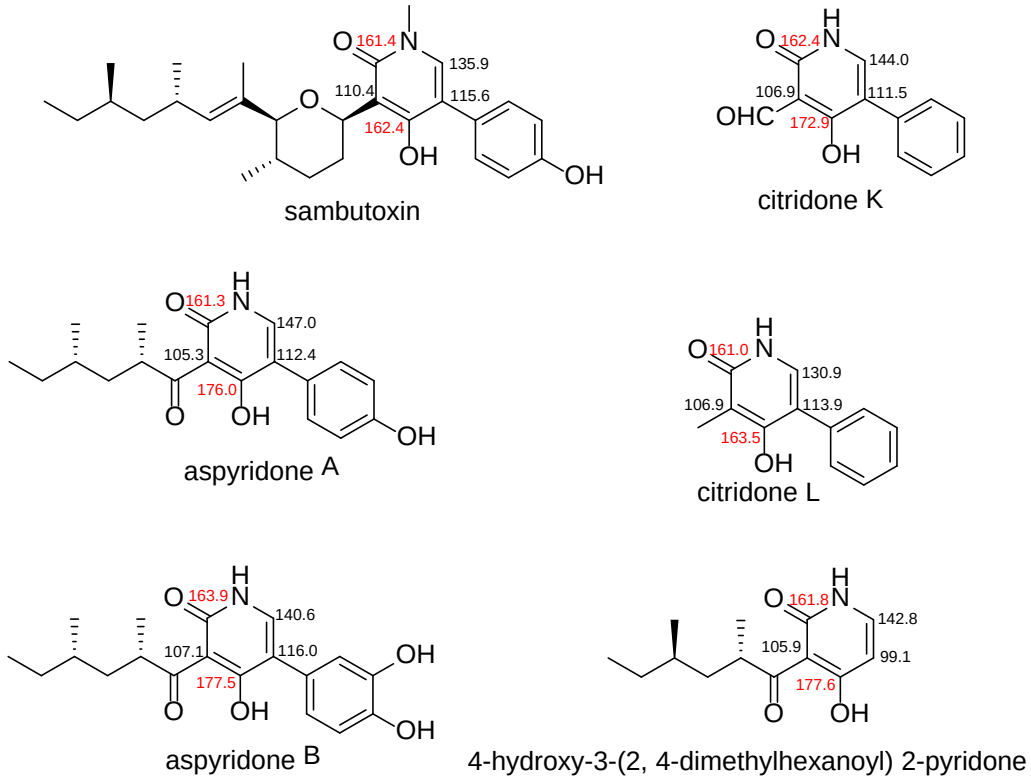


Figure S3. Carbon chemical shifts of 2-pyridone analogues.

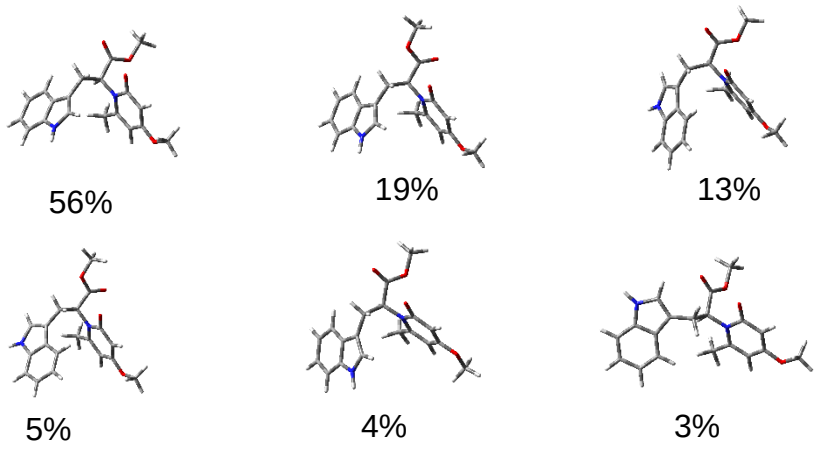


Figure S4. Boltzmann distribution of predominant conformers of 7*S*-**1**.

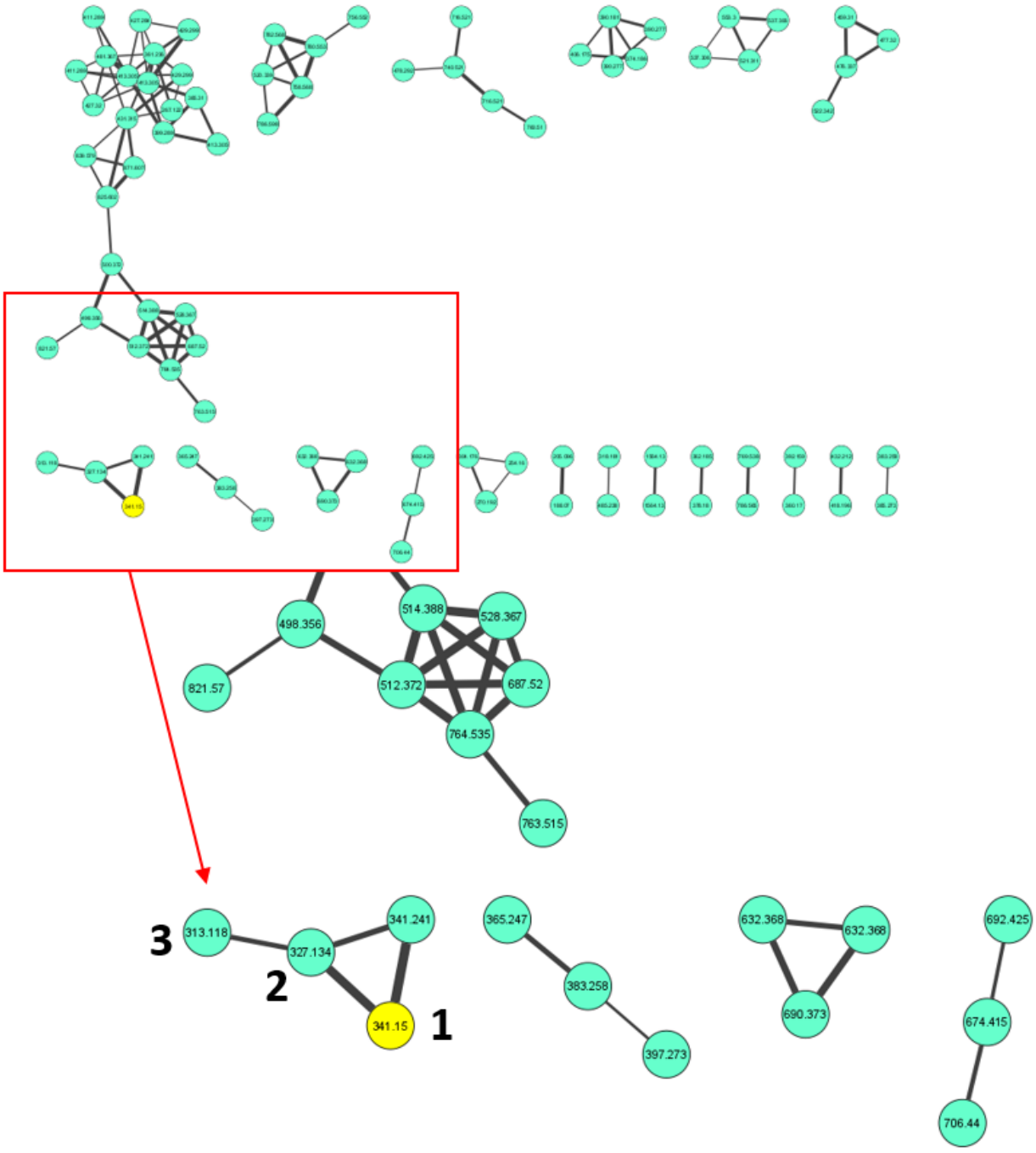


Figure S5. Molecular networking based discovery of **2** and **3**.

(Ion *m/z* 341.241 was proposed to be as **1** as they both had the identical MSMS fragments, and the mass inaccuracy was possibly caused by ion saturation)

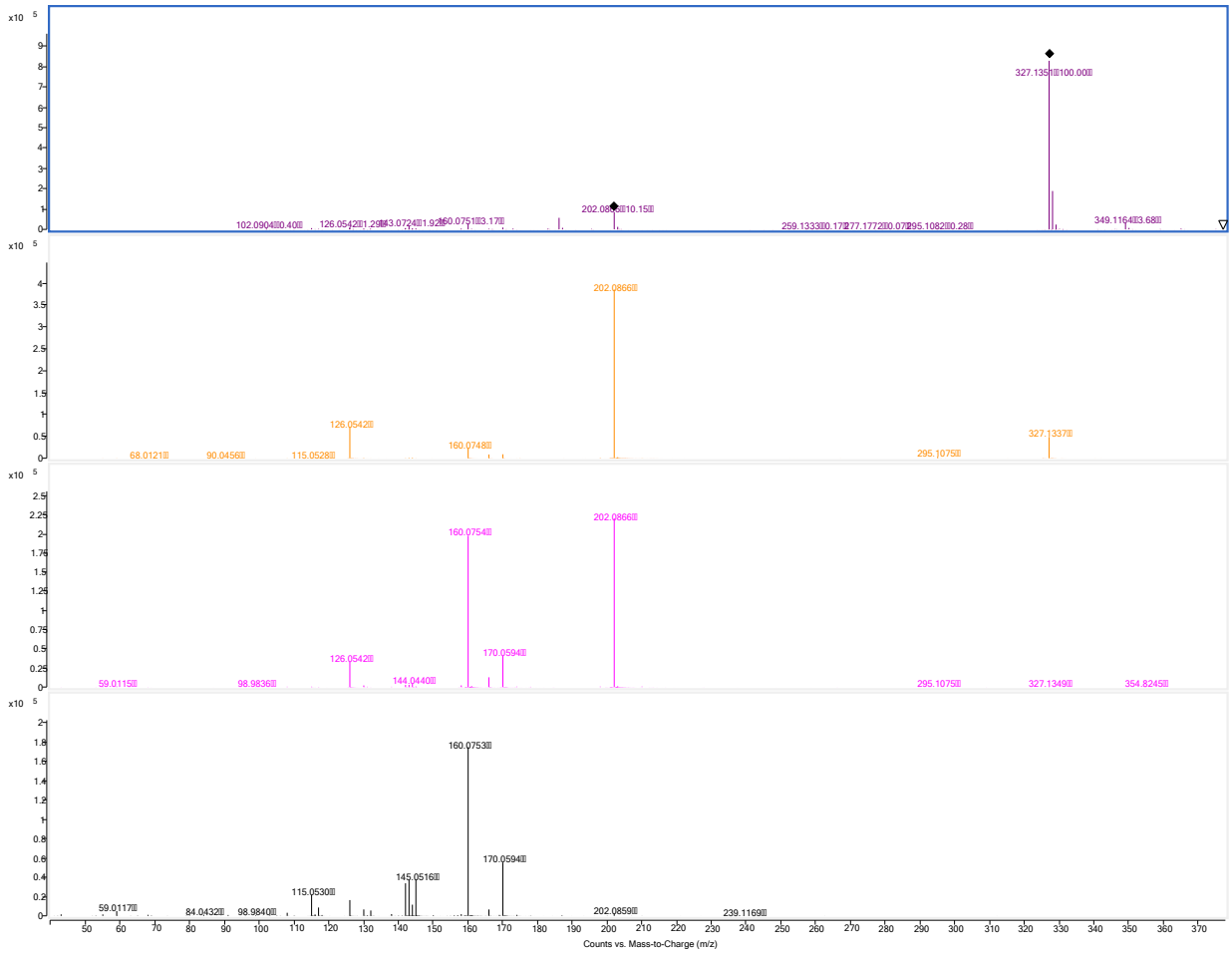


Figure S6. HRESIMS and MSMS (10, 20 and 40 ev) spectra of **2**.

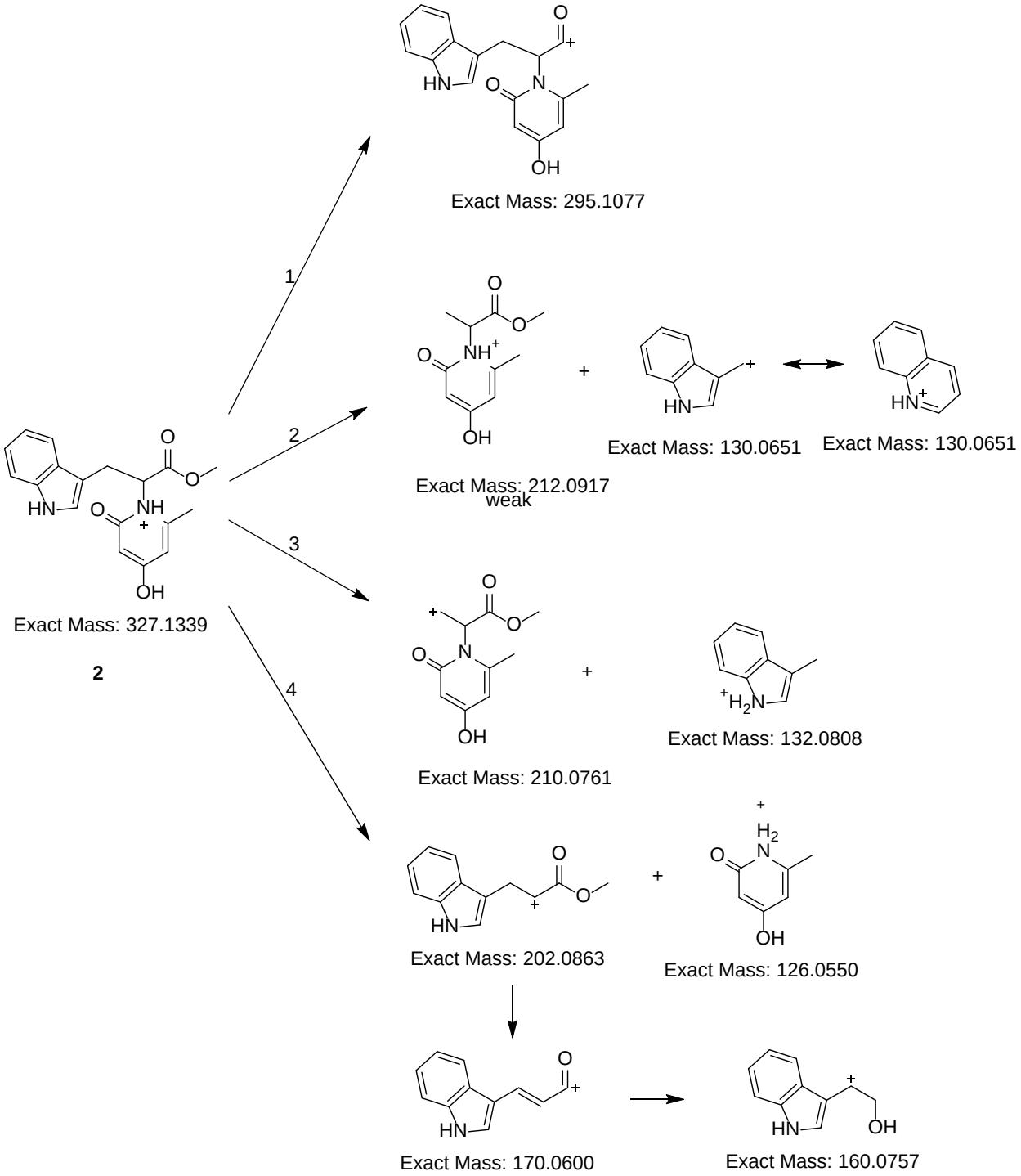


Figure S7. Proposed MMS fragmentation of **2**.

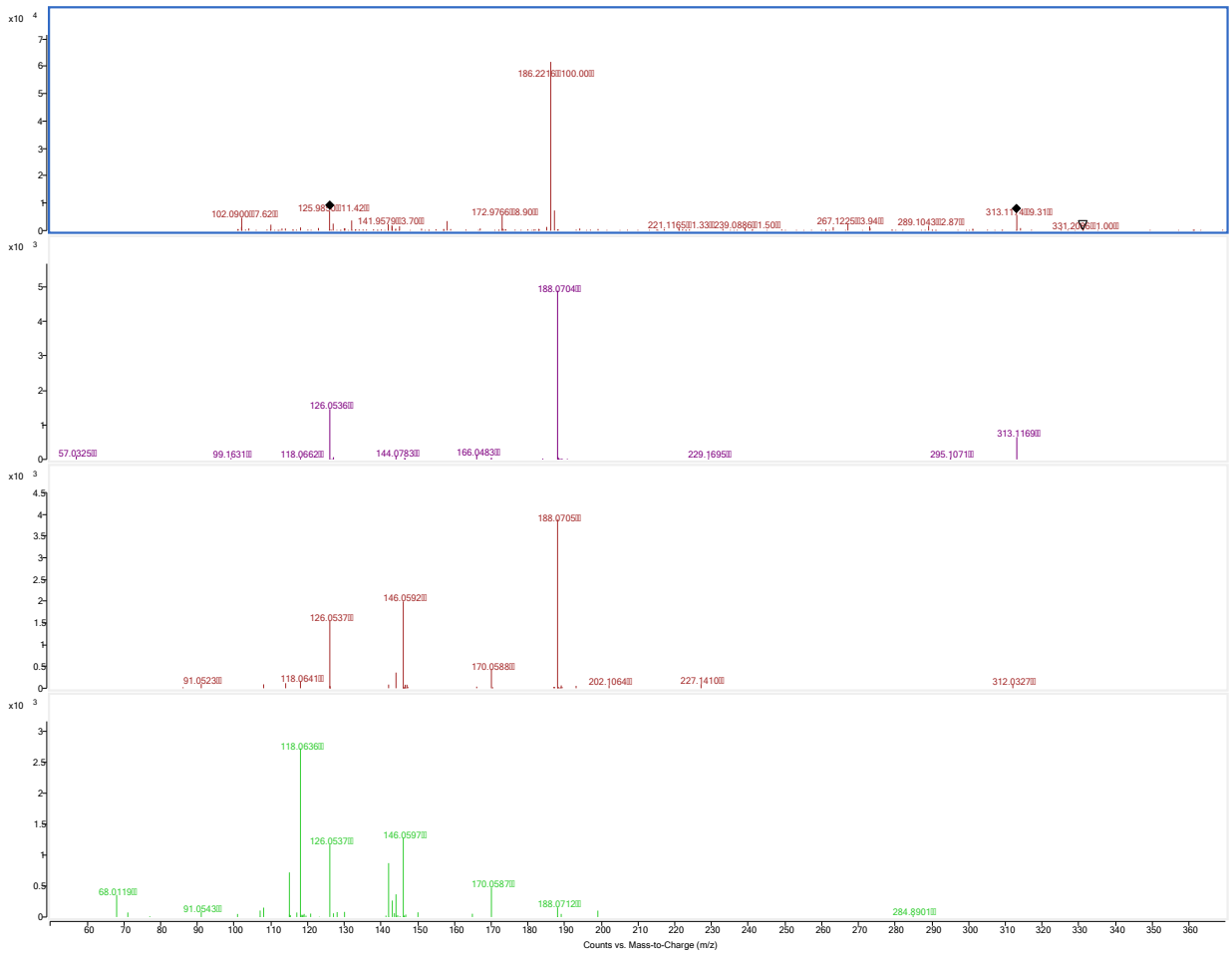


Figure S8. HRESIMS and MSMS (10, 20 and 40 ev) spectra of **3**.

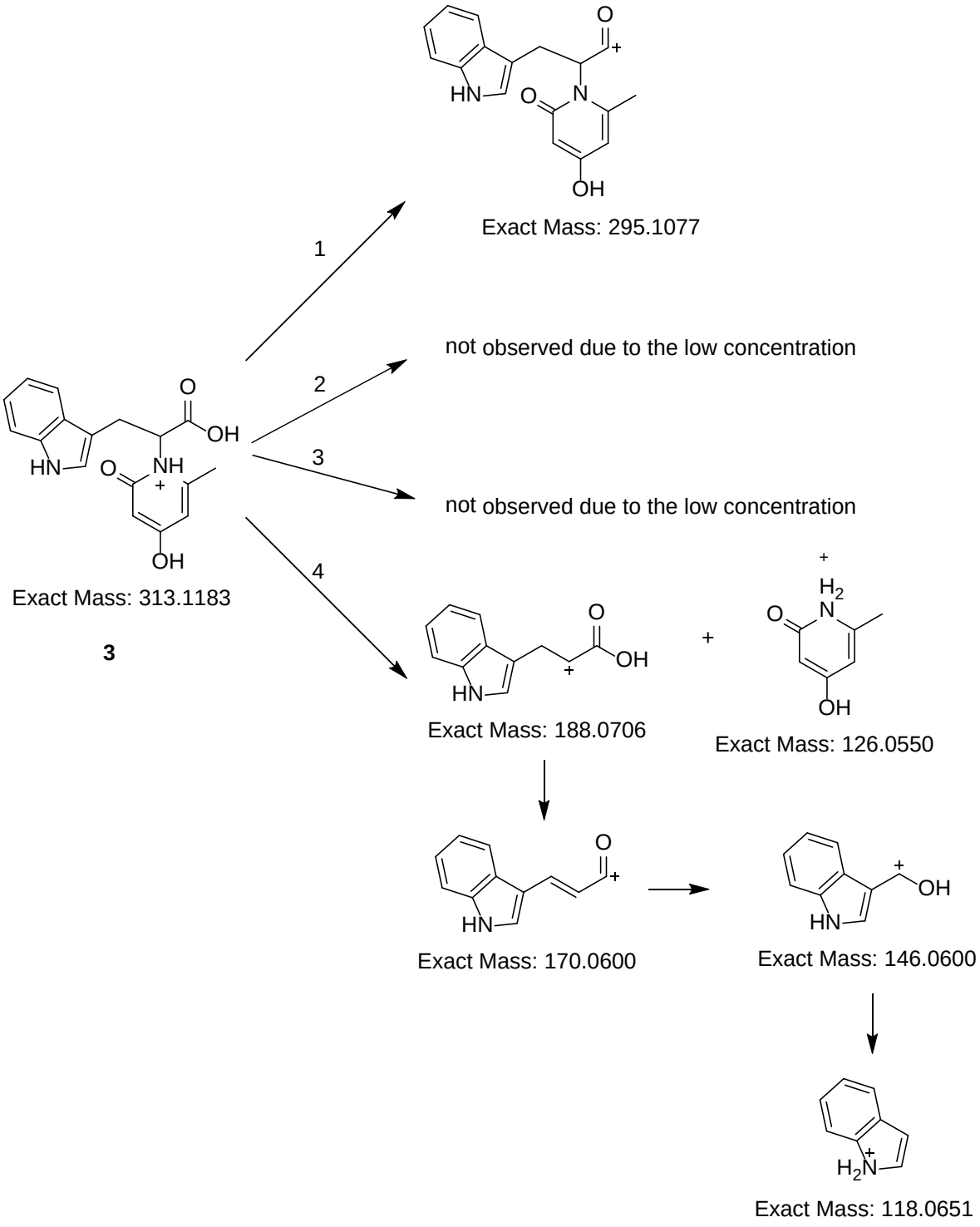


Figure S9. Proposed MMS fragmentation of **3**.

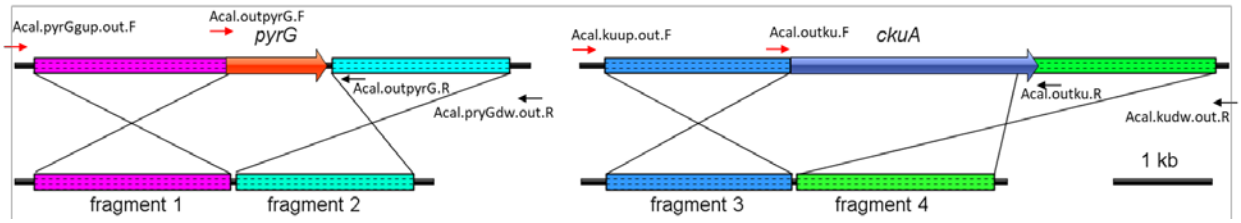


Figure S10. Illustrations of deleting *pyrG* and c*kuA* in *A. californicus*, and primer binding sites in PCR reactions.

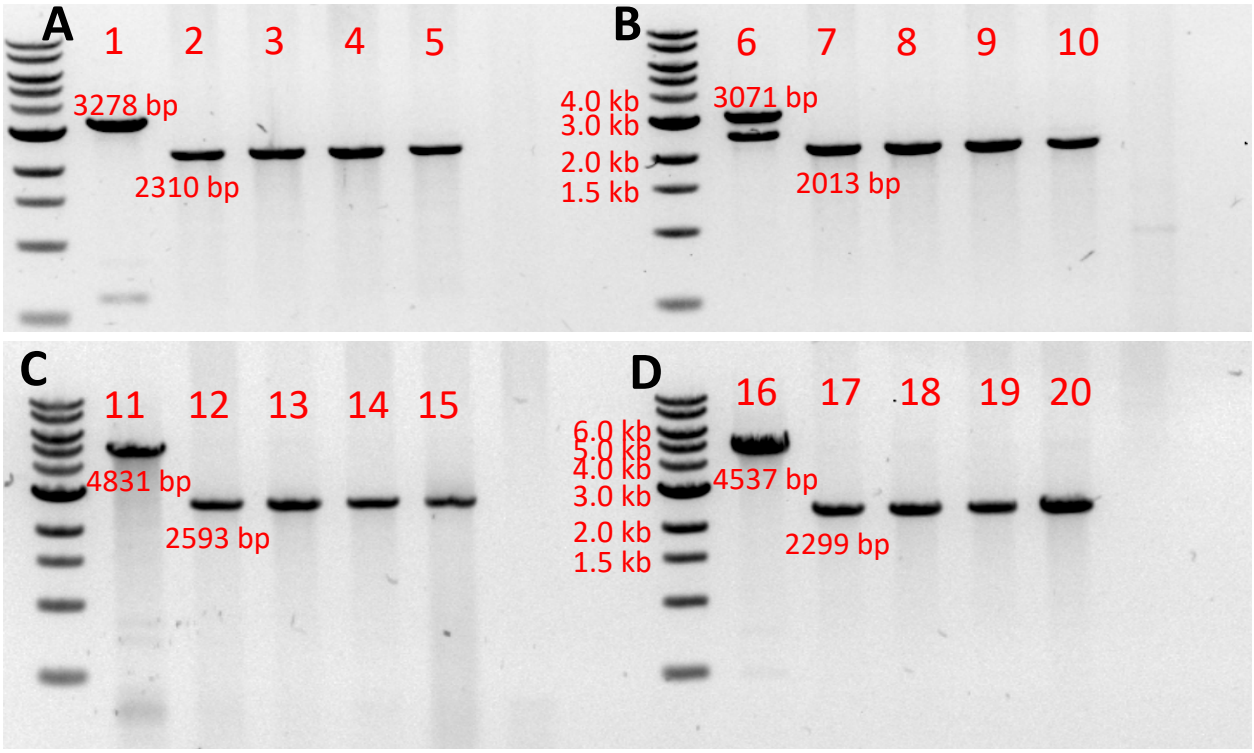


Figure S11. PCR verification of mutants (*pyrG*Δ and c*kuA*Δ) after monosporic purification.

(Lanes 1, 6, 11, and 16 are WT. **A**. Primers for amplification Acal.pyrGup.out.F/Acal.outpyrG.R, lanes 2-5 mutants; **B**. Primers for amplification Acal.outpyrG.F/Acal.pryGdw.out.R, lanes 7-10 mutants, the lower band in lane 6 was from unspecific PCR amplification from scaffold 105; **C**. Primers for amplification Acal.kuup.out.F/Acal.outku.R, lanes 12-15 mutants; **D**. Primers for amplification Acal.outku.F/Acal.kudw.out.R, lanes 17-20 mutants)

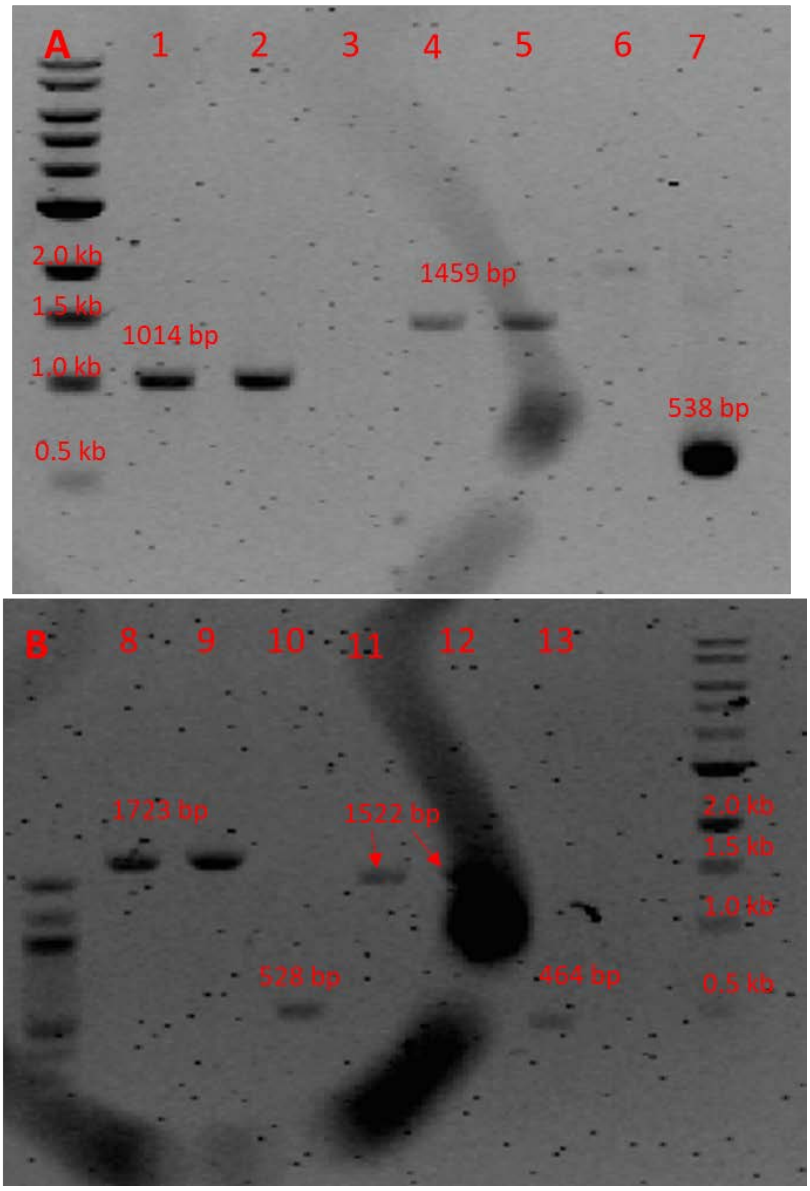


Figure S12. PCR verification of biosynthetic gene deletions.

(Lanes 1, 4, 8 and 11 are WT, lanes 2, 5, 9 and 12 are CAL001. **A**. *cpdA*Δ verification, primers for lanes 1-3 Acal.pyr255550.F1/Acal.pyr255550.R1, Primers for lanes 4-6 Acal.pyr255550.F2/Acal.pyr255550.R2, Primers for lane 7 Acal.pyr255550.F1/Acal.pyr255550.R2; Primer binding pattern was similar to the PCR verification of *pyrG* deletion. **B**. *cpdC*Δ and *cpdB*Δ verification, primers for lanes 8-10 Acal.pyr222318.out.F/Acal.pyr222318.out.R, primers for lanes 11-13, Acal.pyr222319.out.F/ Acal.pyr222319.out.R. Primers bind to 150-350 bp outside both genes in PCR verification)

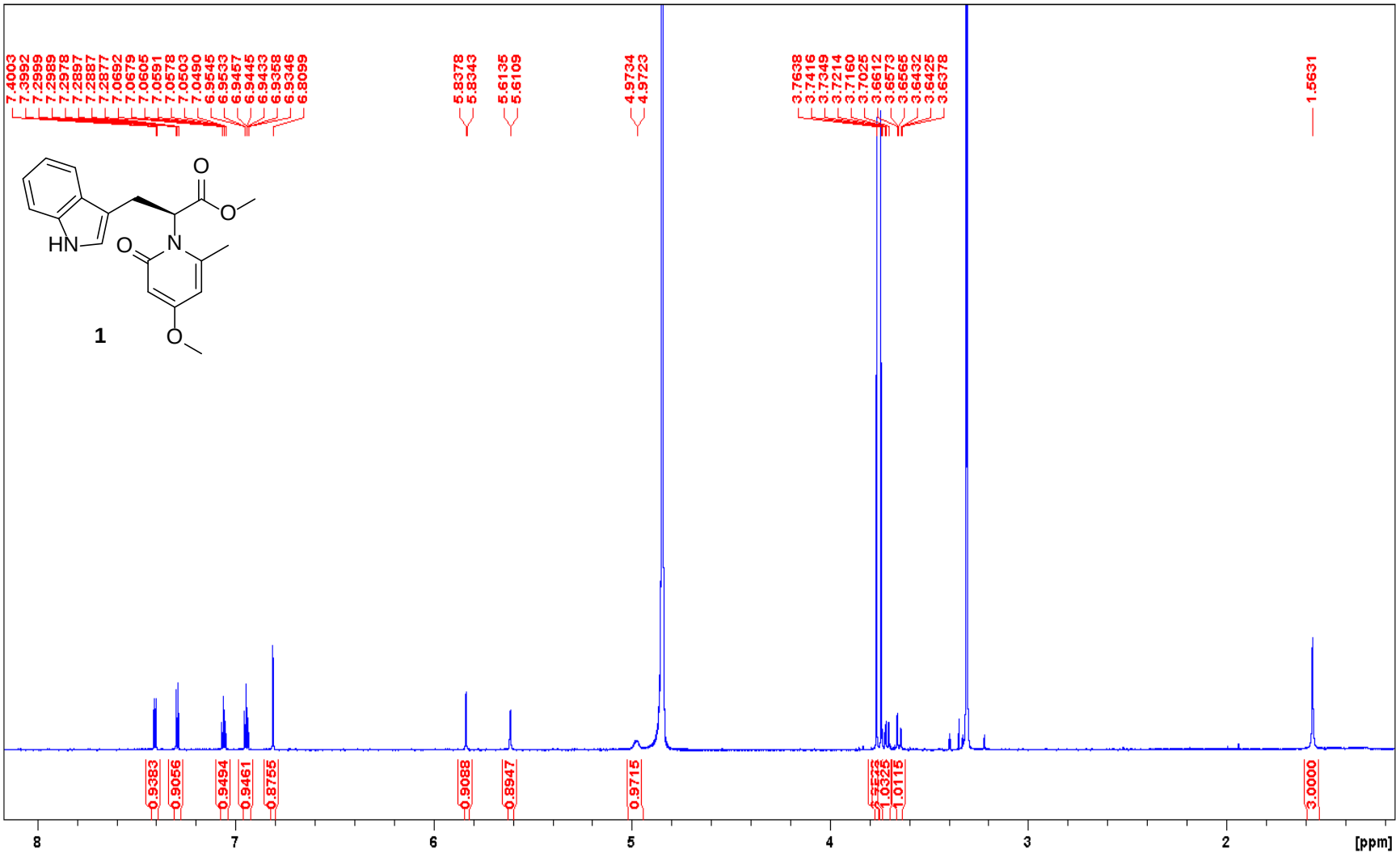


Figure S13. ^{1}H NMR spectrum of **1** (800 MHz, CD_3OD).

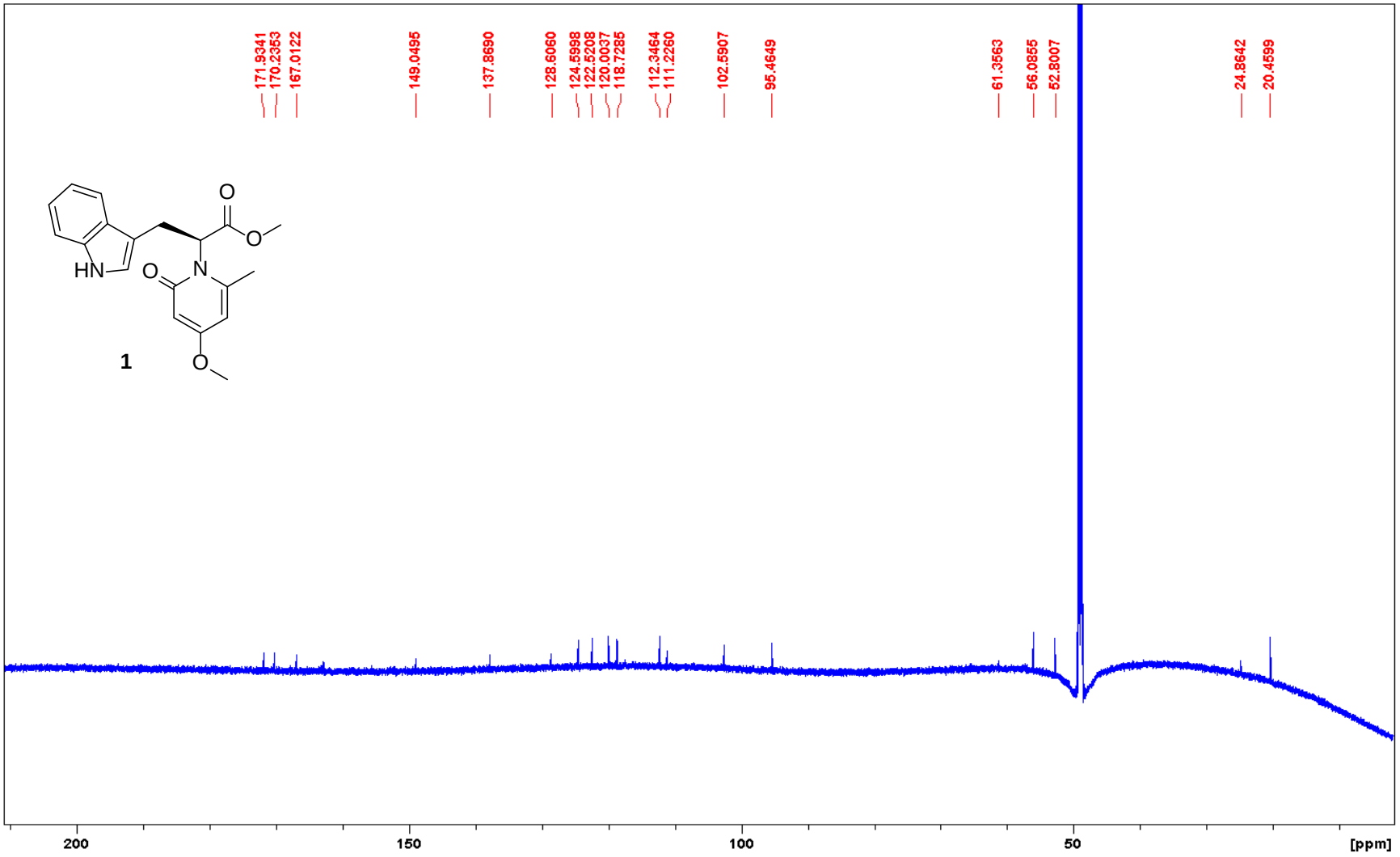


Figure S14. ^{13}C NMR spectrum of **1** (200 MHz, CD_3OD).

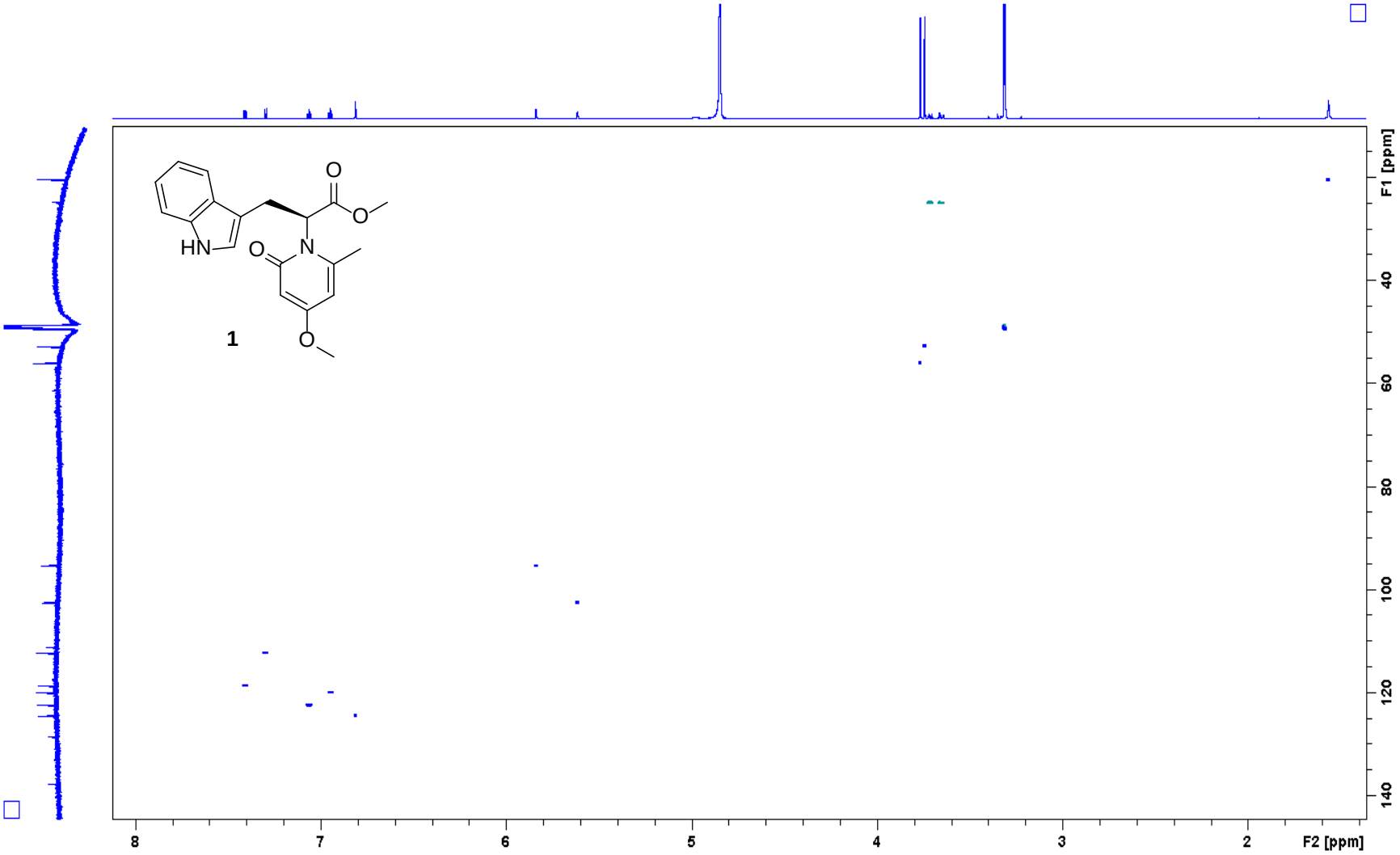


Figure S15. HSQC spectrum of **1**.

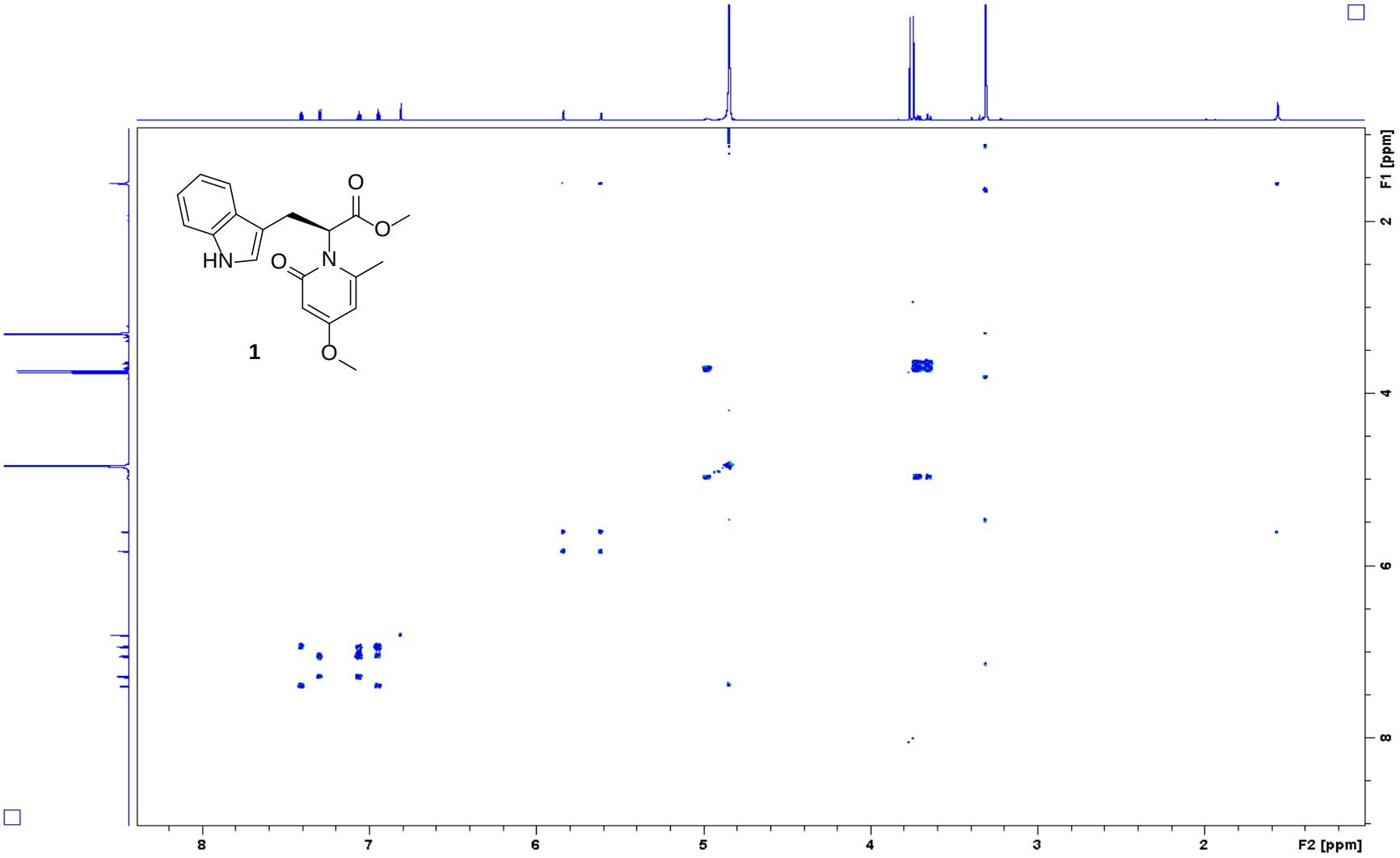


Figure S16. COSY spectrum of **1**.

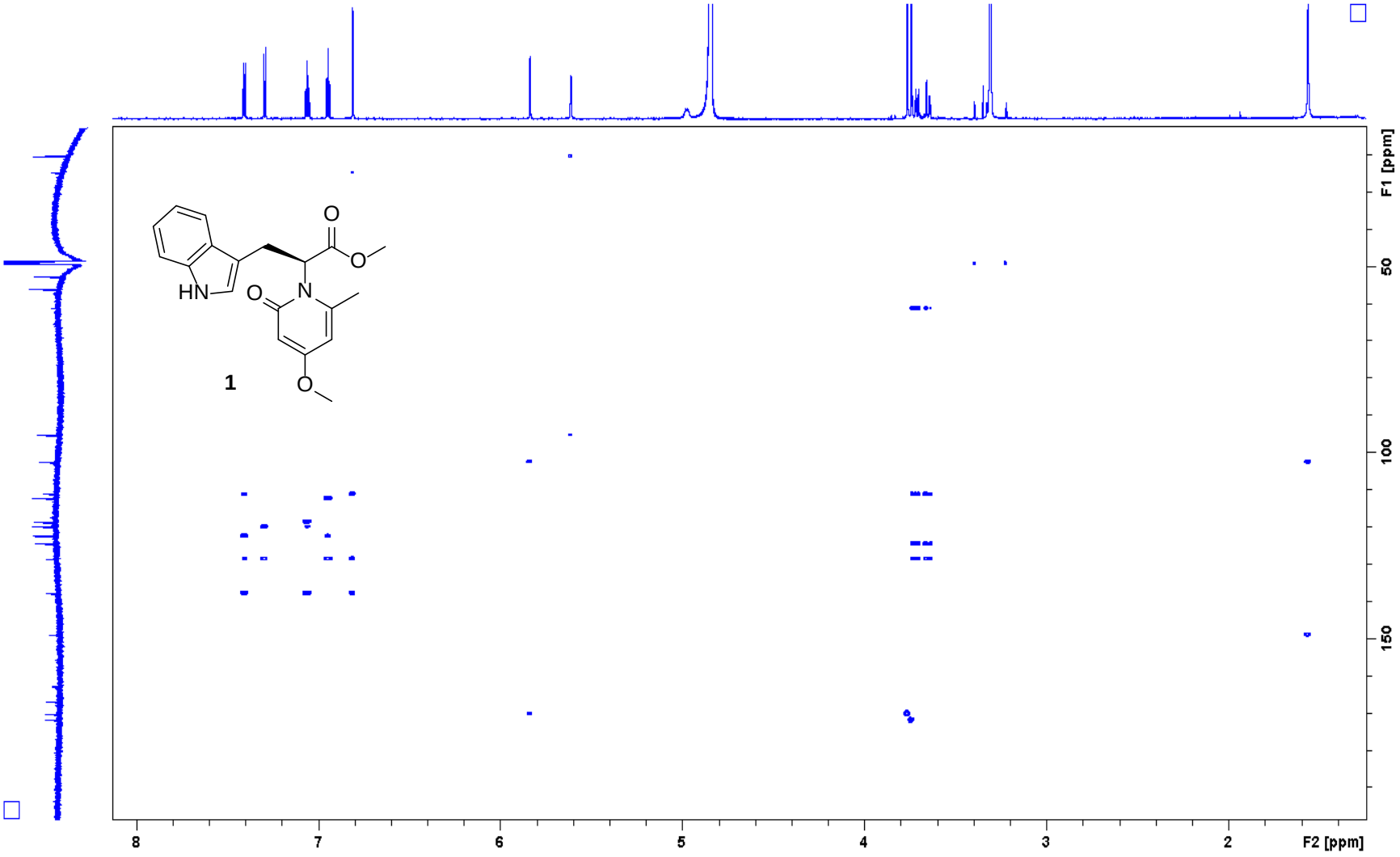


Figure S17. HMBC spectrum of **1**.

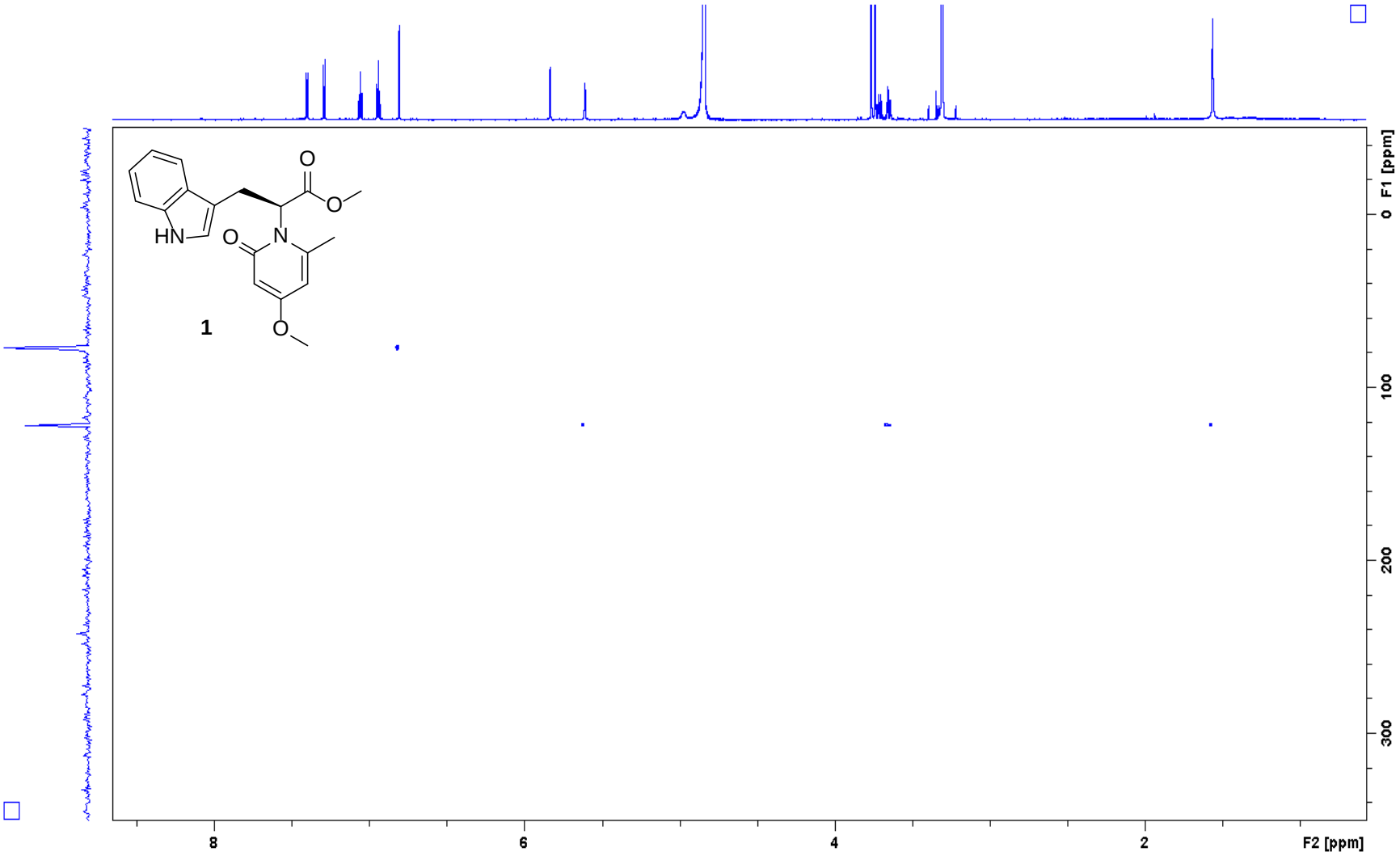


Figure S18. ^{1}H-^{15}N HMBC spectrum of **1**.

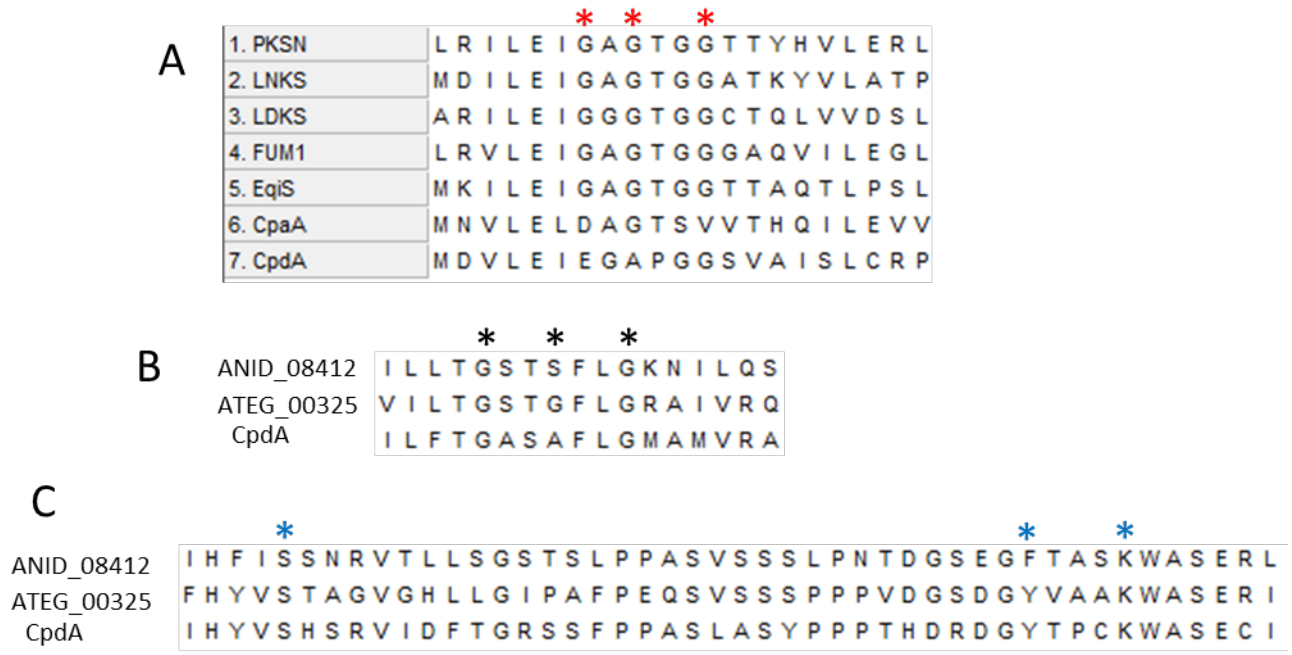


Figure S19. Alignment of the domain sequences of CpdA with other fungal PKS-NRPS.

(Sequences were aligned using ClustalW algorithm with MEGA-X V.10.1.8. All the sequences used for alignment were collected from ref.[15,16] A: MT domain SAM binding site conserved amino acids indicated as red star. B. TD domain NADH/NADPH binding site conserved amino acids as black star. C. TD domain catalytic triad serine-tyrosine-lysine as blue star).

Cartesian coordinates (Å) of the most populated conformers of 7*S*-**1**

conf. 1				conf. 2				conf. 3			
C	0.1877	1.3416	0.0517	C	0.1872	1.3246	-0.2031	C	-1.0501	-1.5116	-0.1302
N	1.1175	0.2220	-0.1733	N	1.1611	0.2268	-0.3183	N	-1.0428	-0.0488	-0.3081
O	1.9237	2.7563	-0.7453	O	0.3839	3.6342	0.3790	O	-3.4182	-1.3683	-0.2877
C	0.9266	2.6813	0.1444	C	0.9244	2.6620	-0.3591	C	-2.4232	-2.0166	0.3298
C	-0.7485	1.1153	1.2597	C	-0.7163	1.1990	1.0469	C	0.0850	-2.0371	0.7748
C	-1.6998	-0.0284	1.0867	C	-1.5776	-0.0271	1.0462	C	1.4667	-1.8091	0.2352
C	-2.9466	-0.0202	0.3632	C	-2.8344	-0.2061	0.3632	C	2.3513	-0.7076	0.5191
C	-3.5007	-1.3252	0.4664	C	-3.2811	-1.5262	0.6442	C	3.5483	-0.9296	-0.2155
N	-2.6275	-2.0793	1.2156	N	-2.3371	-2.1098	1.4568	N	3.3878	-2.1071	-0.9076
C	-1.5550	-1.2974	1.5854	C	-1.3236	-1.2077	1.6957	C	2.1455	-2.6320	-0.6280
C	-3.6536	0.9592	-0.3520	C	-3.6310	0.6180	-0.4473	C	2.2640	0.4315	1.3352
C	-4.8650	0.6244	-0.9366	C	-4.8226	0.1201	-0.9507	C	3.3388	1.3040	1.3937
C	-5.3922	-0.6757	-0.8232	C	-5.2423	-1.1920	-0.6614	C	4.5112	1.0663	0.6507
C	-4.7201	-1.6650	-0.1213	C	-4.4805	-2.0301	0.1388	C	4.6319	-0.0510	-0.1607
O	0.5793	3.6007	0.8481	O	1.8269	2.8336	-1.1464	O	-2.5880	-2.9613	1.0657
C	2.6633	3.9961	-0.7844	C	0.9554	4.9530	0.2287	C	-4.7614	-1.7946	0.0281
C	0.9709	-0.6530	-1.2307	C	1.0505	-0.7574	-1.2794	C	-0.6693	0.5449	-1.4972
C	2.1619	0.0935	0.7713	C	2.2159	0.2447	0.6243	C	-1.4550	0.7069	0.8134
C	1.8268	-1.7059	-1.3609	C	1.9541	-1.7775	-1.3100	C	-0.6529	1.9040	-1.5981
C	2.8670	-1.9073	-0.4182	C	3.0094	-1.8316	-0.3641	C	-1.0231	2.7119	-0.4928
C	3.0309	-1.0278	0.6205	C	3.1369	-0.8443	0.5784	C	-1.4158	2.1266	0.6829
C	-0.1162	-0.4287	-2.2406	C	-0.0521	-0.6901	-2.2953	C	-0.2991	-0.3074	-2.6760
O	3.6377	-2.9897	-0.6472	O	3.8306	-2.8934	-0.4914	O	-0.9574	4.0400	-0.7144
C	4.7176	-3.2686	0.2536	C	4.9306	-3.0255	0.4185	C	-1.3145	4.9335	0.3485
H	-1.3023	2.0439	1.4059	H	-1.3437	2.0903	1.0784	H	-0.0902	-3.1064	0.8954
H	-0.1346	0.9729	2.1472	H	-0.0860	1.2228	1.9334	H	-0.0219	-1.5899	1.7621
H	-0.7626	-1.7103	2.1889	H	-0.4900	-1.4795	2.3235	H	1.8391	-3.5678	-1.0682
H	-2.7550	-3.0446	1.4712	H	-2.3824	-3.0417	1.8345	H	4.0760	-2.5362	-1.5038
H	-3.2631	1.9658	-0.4458	H	-3.3241	1.6319	-0.6763	H	1.3708	0.6273	1.9157
H	-5.4188	1.3734	-1.4896	H	-5.4447	0.7485	-1.5764	H	3.2803	2.1849	2.0213
H	-5.1276	-2.6646	-0.0316	H	-4.8056	-3.0385	0.3644	H	5.5352	-0.2376	-0.7288
H	-6.3413	-0.9071	-1.2911	H	-6.1785	-1.5528	-1.0695	H	5.3340	1.7677	0.7166
H	-0.4247	1.4329	-0.8424	H	-0.4520	1.2788	-1.0819	H	-0.9186	-1.9520	-1.1159
H	2.0027	4.8247	-1.0351	H	2.0028	4.9442	0.5259	H	-4.9023	-2.8391	-0.2456
H	3.4138	3.8604	-1.5571	H	0.3781	5.5937	0.8882	H	-5.4130	-1.1557	-0.5599
H	3.1375	4.1813	0.1783	H	0.8681	5.2884	-0.8034	H	-4.9573	-1.6629	1.0913
H	1.7187	-2.3941	-2.1856	H	1.8738	-2.5509	-2.0588	H	-0.3592	2.3754	-2.5237
H	3.8135	-1.1249	1.3563	H	3.9281	-0.8287	1.3113	H	-1.7185	2.6921	1.5501
H	-1.1063	-0.4405	-1.7824	H	-1.0379	-0.7049	-1.8278	H	0.5395	-0.9680	-2.4514
H	-0.0781	-1.2184	-2.9877	H	0.0213	-1.5485	-2.9593	H	-0.0140	0.3356	-3.5056
H	0.0075	0.5266	-2.7562	H	0.0205	0.2128	-2.9060	H	-1.1391	-0.9245	-3.0040
H	5.4423	-2.4522	0.2567	H	5.6126	-2.1777	0.3301	H	-2.3550	4.7888	0.6454
H	5.1866	-4.1735	-0.1219	H	5.4423	-3.9392	0.1295	H	-1.1838	5.9342	-0.0531
H	4.3458	-3.4389	1.2657	H	4.5761	-3.1113	1.4474	H	-0.6593	4.7937	1.2103
O	2.2701	0.9500	1.6632	O	2.2878	1.1894	1.4260	O	-1.8270	0.1063	1.8343

conf. 4

C	-1.4437	-1.0624	-0.3651
N	-1.0110	0.3432	-0.4447
O	-3.2833	-2.1440	0.7145
C	-2.9452	-1.0990	-0.0440
C	-0.5481	-1.9405	0.5392
C	0.8799	-2.0361	0.0849
C	1.9979	-1.2410	0.5255
C	3.1446	-1.7019	-0.1774
N	2.7333	-2.7240	-1.0004
C	1.3808	-2.9244	-0.8333
C	2.1524	-0.2024	1.4576
C	3.4101	0.3439	1.6571
C	4.5294	-0.1241	0.9424
C	4.4124	-1.1514	0.0190
O	-3.7427	-0.3226	-0.5183
C	-4.6916	-2.3110	0.9929
C	-0.4500	0.8771	-1.5874
C	-1.2306	1.1167	0.7186
C	-0.0496	2.1798	-1.5987
C	-0.2156	2.9932	-0.4487
C	-0.7935	2.4742	0.6806
C	-0.2976	0.0262	-2.8145
O	0.2243	4.2604	-0.5831
C	0.0916	5.1566	0.5280
H	-0.9934	-2.9352	0.5464
H	-0.6013	-1.5634	1.5588
H	0.8739	-3.6999	-1.3856
H	3.3267	-3.2615	-1.6105
H	1.3020	0.1669	2.0177
H	3.5385	1.1450	2.3749
H	5.2734	-1.5165	-0.5277
H	5.4992	0.3245	1.1195
H	-1.3862	-1.4697	-1.3719
H	-5.0720	-1.4464	1.5346
H	-4.7598	-3.2045	1.6057
H	-5.2482	-2.4392	0.0659
H	0.3927	2.6036	-2.4875
H	-0.9546	3.0505	1.5780
H	0.3203	-0.8525	-2.6241
H	0.1760	0.6125	-3.5987
H	-1.2660	-0.3140	-3.1889
H	-0.9580	5.2916	0.7956
H	0.5093	6.1020	0.1938
H	0.6511	4.7912	1.3913
O	-1.7834	0.5860	1.6954

conf. 5

C	-0.2536	1.3800	0.0040
N	-1.1781	0.2338	0.0502
O	-2.0507	2.6793	0.8587
C	-1.0072	2.7146	0.0210
C	0.7478	1.2974	-1.1694
C	1.7035	0.1482	-1.0761
C	2.9119	0.0832	-0.2925
C	3.4884	-1.1968	-0.5154
N	2.6646	-1.8668	-1.3898
C	1.6019	-1.0560	-1.7236
C	3.5687	0.9799	0.5650
C	4.7539	0.5899	1.1688
C	5.3041	-0.6844	0.9347
C	4.6816	-1.5920	0.0909
O	-0.6362	3.7110	-0.5539
C	-2.8103	3.8997	0.9970
C	-1.0776	-0.7502	1.0028
C	-2.1705	0.2065	-0.9661
C	-1.9248	-1.8260	0.9707
C	-2.9068	-1.9242	-0.0444
C	-3.0199	-0.9293	-0.9846
C	-0.0511	-0.6362	2.0922
O	-3.7665	-2.9611	-0.1424
C	-3.7063	-4.0259	0.8171
H	1.2963	2.2405	-1.1802
H	0.1825	1.2518	-2.0983
H	0.8473	-1.4029	-2.4114
H	2.8197	-2.7943	-1.7490
H	3.1603	1.9661	0.7525
H	5.2691	1.2749	1.8312
H	5.1074	-2.5711	-0.0919
H	6.2319	-0.9602	1.4208
H	0.3089	1.3759	0.9349
H	-2.1769	4.7002	1.3764
H	-3.5992	3.6700	1.7066
H	-3.2345	4.1892	0.0367
H	-1.8244	-2.5817	1.7319
H	-3.7733	-0.9805	-1.7575
H	0.9623	-0.5837	1.6920
H	-0.1156	-1.5081	2.7392
H	-0.2192	0.2506	2.7076
H	-2.7377	-4.5280	0.7880
H	-4.4850	-4.7241	0.5232
H	-3.9081	-3.6591	1.8250
O	-2.2363	1.1626	-1.7568

conf. 6

C	0.1360	0.7061	-0.1425
N	1.4207	-0.0196	-0.1395
O	1.2830	2.4799	-1.2252
C	0.3523	2.2160	-0.2983
C	-0.7586	0.3599	1.0744
C	-2.2317	0.4356	0.7924
C	-3.0844	-0.6561	0.3924
C	-4.3961	-0.1300	0.2350
N	-4.3312	1.2120	0.5272
C	-3.0368	1.5440	0.8607
C	-2.8788	-2.0267	0.1689
C	-3.9521	-2.8196	-0.2065
C	-5.2398	-2.2733	-0.3641
C	-5.4786	-0.9249	-0.1442
O	-0.3006	3.0638	0.2613
C	1.5461	3.8734	-1.4976
C	1.6794	-1.0533	-1.0172
C	2.3588	0.3967	0.8322
C	2.8598	-1.7299	-0.9353
C	3.8256	-1.3749	0.0412
C	3.5817	-0.3352	0.9009
C	0.6689	-1.4151	-2.0666
O	4.9470	-2.1229	0.0262
C	5.9825	-1.8299	0.9737
H	-0.4873	1.0116	1.9017
H	-0.5064	-0.6581	1.3770
H	-2.7784	2.5581	1.1146
H	-5.1053	1.8552	0.5075
H	-1.8960	-2.4663	0.2929
H	-3.8016	-3.8781	-0.3808
H	-6.4701	-0.5043	-0.2601
H	-6.0587	-2.9178	-0.6592
H	-0.3910	0.4084	-1.0453
H	0.6461	4.3645	-1.8649
H	2.3198	3.8781	-2.2593
H	1.8965	4.3735	-0.5959
H	3.0697	-2.5396	-1.6176
H	4.2853	-0.0158	1.6535
H	-0.2850	-1.7214	-1.6338
H	1.0501	-2.2452	-2.6571
H	0.4782	-0.5804	-2.7449
H	6.3653	-0.8177	0.8305
H	6.7710	-2.5511	0.7785
H	5.6192	-1.9487	1.9962
O	2.0764	1.3633	1.5570

Supplementary references

(1) Nødvig, C. S.; Nielsen, J. B.; Kogle, M. E.; Mortensen, U. H. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLoS One* **2015**, *10* (7), 1–18.

(2) Robert A Samson, Jos Houbraken, Ulf Thrane, Jens C Frisvad, B. A. *Food and Indoor Fungi*; CBS-KNAW Fungal Biodiversity Centre, 2010.

(3) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105* (8), 2999–3093.

(4) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, D. J. Gaussian 16. Gaussian, Inc.: Wallingford CT 2016.

(5) GENNARO PESCITELLI, T. B. Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra. *Chirality* **2016**, *28* (6), 466–474.

(6) Warnke, I.; Furche, F. Circular Dichroism: Electronic. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 150–166.

(7) Adusumilli R., M. P. Data Conversion with ProteoWizard MsConvert. In: Comai L., Katz J., Mallick P. (Eds) Proteomics. Methods in Molecular Biology. In *Proteomics. Methods in Molecular Biology , vol 1550*; Comai L., Katz J., M. P., Ed.; Humana Press: New York, NY., 2017; pp 339–368.

(8) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837.

(9) Shannon, P.; Markiel, A.; Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, D. R.; Amin, N.; Benno Schwikowski, and T. I. Cytoscape: A Software Environment for Integrated Models. *Genome Res.* **2003**, *13* (22), 2498–2504.

(10) Nour-Eldin, H. H.; Geu-Flores, F.; Halkier, B. A. *USER Cloning and USER Fusion: The*

Ideal Cloning Techniques for Small and Big Laboratories; 2010; Vol. 643.

(11) Nørholm, M. H. H. A Mutant Pfu DNA Polymerase Designed for Advanced Uracil-Excision DNA Engineering. *BMC Biotechnol.* **2010**, *10*.

(12) Nødvig, C. S.; Hoof, J. B.; Kogle, M. E.; Jarczynska, Z. D.; Lehmbeck, J.; Klitgaard, D. K.; Mortensen, U. H. Efficient Oligo Nucleotide Mediated CRISPR-Cas9 Gene Editing in Aspergilli. *Fungal Genetics and Biology*. 2018, pp 78–89.

(13) Nielsen, M. L.; Albertsen, L.; Lettier, G.; Nielsen, J. B.; Mortensen, U. H. Efficient PCR-Based Gene Targeting with a Recyclable Marker for *Aspergillus nidulans*. *Fungal Genet. Biol.* **2006**, *43* (1), 54–64.

(14) Nødvig, C. S.; Hoof, J. B.; Kogle, M. E.; Jarczynska, Z. D.; Lehmbeck, J.; Klitgaard, D. K.; Mortensen, U. H. Efficient Oligo Nucleotide Mediated CRISPR-Cas9 Gene Editing in Aspergilli. *Fungal Genet. Biol.* **2018**, *115* (January), 78–89.

(15) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Functional Expression of the *Aspergillus flavus* PKS-NRPS Hybrid CpaA Involved in the Biosynthesis of Cyclopiazonic Acid. *Bioorganic Med. Chem. Lett.* **2009**, *19* (12), 3288–3292.

(16) Gressler, M.; Zaehle, C.; Scherlach, K.; Hertweck, C.; Brock, M. Multifactorial Induction of an Orphan PKS-NRPS Gene Cluster in *Aspergillus terreus*. *Chem. Biol.* **2011**, *18* (2), 198–209.

Genetic origin of homopyrones, a rare type of hybrid phenylpropanoid and polyketide derived yellow pigments from *Aspergillus homomorphus*

Malgorzata E. Futyma,[1,4,5] Yaojie Guo,[1,5] Casper Hoeck,[2,3] Jakob B. Hoof,[1] Charlotte H. Gotfredsen,[2] Uffe H. Mortensen,[1*] Thomas O. Larsen,[1*]

*Corresponding authors, um@bio.dtu.dk, tol@bio.dtu.dk

Abstract

In recent years there has been an increasing demand for the replacement of synthetic food colorants with naturally derived alternatives. Filamentous fungi are prolific producers of secondary metabolites including polyketide derived pigments, many of which have not been fully characterized yet. During our ongoing investigations of black aspergilli, we noticed that *Aspergillus homomorphus* turned yellow when cultivated on malt extract agar plates. Chemical discovery guided by UV and MS led to the isolation of two novel yellow natural products and their structures were elucidated as aromatic α-pyrones homopyrones A (**1**) and B (**2**) by HRMS and NMR. Combined investigations including retro-biosynthesis, genome mining, and gene deletions successfully linked both compounds to their related biosynthetic gene clusters. This demonstrated that homopyrones are biosynthesized by using cinnamoyl-CoA as the starter unit, followed by extension with three malonyl-CoA units, reduction of one keto group, to give the core hybrid backbone structure after lactonization. The polyketide synthase includes a *C*-methylation domain, which however seems to be promiscuous, since only homopyrone B is *C*-methylated. Altogether, the homopyrones represent a rare case of a hybrid phenylpropanoid and polyketide derived natural products in filamentous fungi.

Keywords: *Aspergillus homomorphus*, homopyrone, pigments, CRISPR-Cas9, gene deletion, biosynthesis

[1] Department of Biotechnology and Biomedicine, Søltofts Plads, Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark
[2] Department of Chemistry, Kemitorvet B207, Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark
[3] Novo Nordisk A/S, Smørmosevej 10-12, 2880 Bagsværd, Denmark
[4] CHRETO, Lejrvej 17, 3500 Værløse, Denmark
[5] The authors contributed equally to this work

Introduction

Aspergilli are well known for their ability to produce valuable secondary metabolites (SMs) some of which are used in the pharmaceutical industry including echinocandin-type antimicrobials, cholesterol-lowering statins, and immunosuppressive cyclosporins, but on the other hand, they are also known to produce harmful mycotoxins like aflatoxins, fumonisins, and ochratoxins that cause human disease and food spoilage (Frisvad and Larsen 2015; Bills and Gloer 2016). The chemical space of SMs produced by the *Aspergillus* genus is vast and includes compounds where the pharmaceutical or food industrial potential remains to be discovered. This view is strongly supported by the recent sequencing of a large number of *Aspergillus* genomes showing that the total number of biosynthetic gene clusters (BGCs) of SMs is even larger than what was expected based on the metabolome analyses (Vesth et al. 2018). To uncover and exploit the promising *Aspergillus* SM repertoire, there is an increasing interest in linking *Aspergillus* SMs to their BGCs as this will facilitate SM discovery and set the stage for the construction of efficient cell factories for valuable SM production. Recent examples in black aspergilli include the linking of the structures of azanigerones (Zabala et al. 2012), malformins (Theobald et al. 2018), acurin A (Wolff et al. 2020), and pyrophen (Hai et al. 2020) to their related BGCs.

In recent years there has been an increasing demand for the replacement of synthetic food colorants with alternatives derived from natural sources. Filamentous fungi are good natural producers of various colorants like genera *Monascus*, *Penicillium*, *Talaromyces, Fusarium*, and *Aspergillus*, and some species have been used as cell factories for the production of pigments such as carminic acid and azophilones (Frisvad et al. 2013; Akilandeswari and Pradeep 2016; Christiana 2016; Frandsen et al. 2018; Isbrandt et al. 2020; Tolborg et al. 2020). As part of our ongoing interest and research in the discovery of natural pigments from Aspergilli, we noted that *Aspergillus homomorphus*, a genetically uncharacterized uniseriate species belonging to section *Nigri* series *Homomorphi* (Houbraken et al. 2020), produces ample amounts of yellow compounds on malt extract agar (MEA) plates. Therefore, we set out to elucidate the chemical structures and biosynthesis of these yellow compounds which turned out to be aromatic α-pyrones, not isolated from Nature before. Herein, we

report our work on the characterization of the two yellow compounds homopyrones A (**1**) and B (**2**) (Figure 1) and the linking of both compounds to their related BGC.

Materials and methods

Strains, media, and instruments

A. homomorphus IBT 21893 (= ITEM 7556, CBS 101889) was used for isolation of the compounds and genetic manipulation. *Escherichia coli* DH5α was used for cloning and plasmid amplification in liquid or solid Luria-Bertani (LB) medium supplemented with ampicillin 100 μg/ml (Wolff et al. 2020). The media used for cultivation of *A. homomorphus* in this study include minimal medium (MM), transformation medium (TM), and Malt Extract Agar (MEA) supplemented with uridine (Uri) (10 mM) when necessary (Samson et al. 2010; Hoof et al. 2018). LC-MS data were acquired on an Agilent Infinity 1290 UHPLC-DAD-6545 QTOFMS system using electrospray ionization with an Agilent Poroshell 120 Phenyl-Hexyl column (2.1*250 mm, 2.7 μm) at 60 °C. The HPLC program was H_2O as mobile phase A and acetonitrile as mobile phase B, both buffered with 20mM formic acid, at a flow rate of 0.35 mL/min and the following gradient: 0-15 min 10%-100% mobile phase B; 15-17min 100% mobile phase B, 17.1-20 min 100%-10% mobile phase B. All NMR spectra were acquired using standard pulse sequences on a Varian Unity Inova 500 at 298 K (500 MHz for ^{1}H, 125 MHz for ^{13}C). The NMR solvent used for both compounds was DMSO-d_6 which was also used as the internal standard reference of chemical shifts (δ_H = 2.50 ppm, δ_C = 39.52 ppm).

Cultivation, isolation, and structural characterization of the homopyrones A and B.

A. homomorphus IBT 21893 was inoculated as three-point stabs on 40 MEA plates and incubated in the dark at 30 °C for 5 days. The fungi were harvested and extracted twice with ethyl acetate yielding a crude extract of 217.1 mg which was dissolved in MeOH:H_2O (9:1) and washed with heptane. Water was added to dilute the solvent to MeOH:H_2O (1:1) and compounds were extracted with dichloromethane (DCM) followed by concentration in *vacuo* to offer a DCM extract of 36.3 mg. Next, the DCM extract was purified on a semi-preparative HPLC Waters 600 Controller coupled to a Waters 996 Photodiode Array Detector with a flow rate

of 5 mL/min through a Phenomenex Luna II C_{18} column (5 µm, 250 × 10 mm). An isocratic run was performed with 55% acetonitrile in 20 min, and **1** (2.5 mg) and **2** (1.7 mg) were collected at 13.4 and 15.8 min, respectively.

Vector construction

Primers including protospacers (PSs) were designed based on the whole-genome sequence of *A. homomorphus* CBS 101889 v1.0 from JGI Mycocosm (Grigoriev et al. 2014; Vesth et al. 2018). Primes and protospacers (PSs) are listed in Tables S2 and S3, respectively. For creating a point mutation in AHOM_468628 the putative *pyrG* homolog in *A. homomorphus*, a single sgRNA splicing cassette strategy was chosen (Nødvig et al. 2018). However, to increase the gene-targeting efficiency a multiple sgRNA splicing cassette with three PSs was designed (Figure 3) (Nødvig et al. 2018): PS1 in the proximity of start codon, preferentially upstream, PS2 around 100-200 bp downstream of the start codon, and PS3 in the proximity of the stop codon. The plasmids with three PSs were accompagnied by GE-Oligonucleotides. PCR fragments were amplified using PfuX7 polymerase (Nørholm 2010) or PhusionU Hot start PCR Master Mix (Life Technologies Europe BV) with primers purchased from IDT Integrated DNA Technologies, Belgium. PCR was performed according to the manufacturer's protocol. Promoter and terminator (*A. fumigatus* U3), *A. nidulans* glycine tRNA, gRNA fragments were amplified from vector pFC902 (Nødvig et al. 2018). CRISPR plasmids were constructed by Uracil-Specific Excision Reagent (USER) cloning into vector pFC330 or pFC332 depending the selection strategy, see Table S4 (Nødvig et al. 2015). After USER cloning, the plasmid DNA was transformed into chemically competent *E. coli* DH5α. Plasmids were purified with the GenElute™ plasmid purification kit (Sigma-Aldrich, Germany) and analyzed by restriction analysis (*BspEI,* NEB) and sequencing (StarSEQ GmbH, Germany) using primers MF233 and MF416, respectively. Plasmids used in this study are shown in Table S4. To support CRISPR induced double-strand breaks repair, oligonucleotides of 60 nt (MF787 and MF788, homologous to 30 nt sequence in the proximity of PS1 and PS3 aligning to the sequences outside the targeted gene) were designed and purchased from IDT (Nødvig et al. 2018).

Strain construction and verification

A. homomorphus protoplastation was performed according to the protocols developed for *A. nidulans* (Nielsen et al. 2006). For each transformation approximate 3 μg of CRISPR plasmid was used for 100 μL of protoplasts (Table S4), and for transformations employing multiple splicing cassettes 15 μL of 100 nM 60 nt corresponding oligonucleotide repair substrate (Table S2) were also added.. Subsequently, 150 μL of PCT solution (50% w/v PEG 8000, 50 mM $CaCl_2$, 20 mM Tris, 0.6 M KCl) was added and gently mixed with protoplasts, followed by incubation for 10 min at room temperature. 250 μL of STC solution (1M sorbitol, 50 mM Tris pH 8, 50 mM $CaCl_2$) was added and the entire volume was plated on TM agar plates. Incubation was performed in the darkness at 30°C until colonies appeared (5-6 days). The obtained colonies were re-streaked on MM + Uri plates and screened by tissue PCR with MyTaq™ Plant-PCR kit (Bioline Reagents Ltd, UK). The candidates with desired gene-deletion bands were further dissected with a micromanipulator and a microscope. Colonies from single-spores were screened again by tissue PCR using the primer combinations shown in Figures S2 and S3, and the correct ones were further verified by PCR with the genomic DNA extracted with FastDNA™ SPIN Kit (MP Biomedicals, USA).

Plug extraction

Plug extraction was performed with validated *A. homomorphus* mutants and reference strains HOM2 (*pyrG*$^-$) and HOM1 (wild type, WT). The strains were inoculated as three-point stabs on MEA + Uri plates and incubated in the darkness at 30°C for seven days. Six plugs were harvested from the colonies and transferred into a 2 ml tube, and each plate was extracted in duplicates. 800 μL of ethyl acetate:isopropanol (3:1) solvent with 1% formic acid was added to the tube and extracted in an ultra-sonication bath for 1 hour. The liquid fraction was transferred into a new tube and evaporated in a nitrogen stream at 42°C. Dried samples were re-suspended with 300 μL of methanol and ultrasonicated for 30 min. Subsequently, the samples were centrifuged (20 000g, 5-10 min), and 150 μL supernatant was transferred to HPLC vials for chemical analysis.

Results

Characterization of the structures of two yellow compounds

During our recent investigations of the chemistry of black aspergilli we noted by visual inspection that *A. homomorphus* IBT 21893 apparently produces large amounts of yellow compounds especially when cultivated on OAT meal agar based media. Analysis based on plug extraction followed by HPLC-DAD-HRMS (data not shown) immediately drew our attention to the two major peaks in the chromatogram, with UV-VIS spectra indicating yellow compounds. Dereplication based on HRMS next revealed that the two unknown compounds had the elemental compositions $C_{16}H_{14}O_3$ and $C_{16}H_{14}O_3$, which did not match any likely compounds when searched against e.g. ACD Antibase. The compounds (**1**) and (**2**) were therefore isolated and purified from cultivation and extraction of 40 Petri dishes followed by NMR analysis. Compounds **1** and **2** were elucidated as aromatic α-pyrones by HRMS and NMR (Figure S1, Table S1). The structure of **1** turned to be the same as the synthetic compound 4-methoxy-6-((1E,3E)-4-phenylbuta-1,3-dien-1-yl)-2*H*-pyran-2-one (Li et al. 2018), and **2** was the same to a synthetic product 3-methyl-4-methyloxy-6-(2,4-butadien-2-phenyl)-yl-2*H*-2pyran-2-one (Oliver et al. 2001). The pyrones had not previously been isolated from Nature before, so we named them homopyrones A (**1**) and B (**2**).

Identification of a putative BGC for the formation of homopyrones

To identify the BGC responsible for the biosynthesis of the structurally related **1** and **2**, the whole-genome sequence of *A. homomorphus* was subjected to the antiSMASH platform to predict all BGCs resulting in 79 hypothetical BGCs (referred as ´regions´ on antiSMASH) (Nordberg et al. 2014; Blin et al. 2019). Next, we applied a retro-biosynthetic approach to identify the possible BGC candidates that contain genes encoding the enzymatic activities necessary for the biosynthesis (Kjærbølling et al. 2019). Both compounds were proposed to be polyketides with an unusual starter unit cinnamoyl-CoA that was elongated with three malonyl-CoA extender units. The cinnamoyl-CoA was hypothesized to be derived from the deamination of phenylalanine and typically exploited by the plant type III polyketide synthases (PKSs) (Abe and Morita 2010). A first-round searching of the BGCs for the PKS genes yielded 20 hits which were further filtered through searching for phenylalanine ammonia-lyase (PAL) encoding gene(s). BGC 68.2 (Figure 2) was the only hit containing both a PKS gene (AHOM_463923) and a putative PAL gene (AHOM_463919) by BLASTP (Boratyn et al. 2013) and was

therefore selected as the most promising gene cluster related to the biosynthesis of **1** and **2**. This selected BGC was located on scaffold 68 and most of the surrounding genes were poorly annotated. The putative functions of each protein encoded by the potential BGC genes were performed with InterPro (Mitchell et al. 2019) and BLASTP NCBI (Boratyn et al. 2013) to identify known protein domains and search for known homologs.

Introducing CRISPR technology in *A. homomorphus*.

Linking SMs to BGCs is typically done by reverse genetics followed by metabolome analysis. However, in the case of *A. homomorphous*, no genetic engineering has previously been performed. To facilitate gene targeting in species where no genetic toolbox exists, we have recently developed CRISPR technology, which appears to work in a broad range of fungi (Nødvig et al. 2015; Nielsen et al. 2017; Nødvig et al. 2018).To demonstrate that our CRISPR system works in a new species, we typically mutate a gene causing an easily scorable phenotype (Hoof et al. 2018). Regarding editing *A. homomorphus*, we decided to mutate AHOM_468628, which is a homolog of *A. nidulans pyrG* encoding orotidine 5'-phosphate decarboxylase (ODC). Importantly, *pyrG* mutant strains, unlike wild-type strains, can grow in the presence of the antimetabolite 5-fluororotic acid (5-FOA) as 5-FOA is not metabolized into toxic compounds due to missing ODC activity (Boeke et al. 1984). *A. homomorphus* was transformed with a *hph*-AMA1 based CRISPR plasmid encoding Cas9 and a single-guide RNA splicing cassette matching a protospacer in the AHOM_*pyrG* gene (Nødvig et al. 2018). If the Cas9/AHOM_*pyrG*-gRNA introduces a DNA double-strand breaks into AHOM_468628, subsequent non-conservative repair of the break by the non-homologous end-joining pathway may produce a non-functional allele. The Cas9/AHOM_*pyrG*-gRNA activity appeared high as subsequent plating on solid MM medium containing 5-FOA readily generated single colonies from the background of poorly growing transformants. A corresponding experiment using a control CRISPR vector, which does not encode any gRNA, did not generate any equivalent single transformants on this medium. Next, we demonstrated that 5-FOA resistant transformants grew on MM + Uri plates, but not on MM plates, in agreement with AHOM_468628 being a *pyrG* homolog. Finally, the *pyrG* locus of one of the *pyrG* mutant strains was sequenced and a mutation was

observed right at the Cas9/AHOM_*pyrG*-gRNA cleavage site explaining its resistance to 5-FOA. The resulting *pyrG*$^{-}$ strain, HOM2, was used as a basis for the gene deletion experiments described below.

The deletion of the putative PKS gene *AhpA* eliminated the production of both compounds.

The putative PKS gene *AhpA* (Asphom1_463923, 80366 - 87858 on scaffold 68) in BGC 68.2 was deleted using CRISPR-Cas9-vector employing the multiple gRNA splicing cassette developed by (Nødvig et al. 2018). Single spores from potential correct transformants were rigorously analyzed by PCR as shown in Figure S2, and homokaryotic deletion mutants were verified. Chemical analysis of the mutant strain HOM32 (*pyrG*$^{-}$ and *Ahp*AΔ), and reference strains HOM1 (WT) and HOM2 (*pyrG*$^{-}$) indicating the abolishment of the production of **1** and **2**, which strongly suggested the essential role of the AhpA in their biosynthesis (Figure 4).

Truncation of the putative PAL gene *AhpB* reduced the production of both compounds.

The elimination of the production of both compounds in HOM32 indicated that we have predicted the correct gene cluster. To further investigate the initiation of the polyketide formation, a putative PAL gene *AhpB* (Asphom1_463919, 75176 - 77397 on scaffold 68) encoding the phenylalanine ammonia-lyase for the generation of the starter compound cinnamic acid was targeted by strategy as for *AhpA*. PCR and subsequent sequencing results showed the truncation of the 5´ end of the gene (Figure S3). The chemical profile of the truncated mutant HOM45 (*pyrG*- and *AhpB*Δ) showed a significant decrease in the generation of two compounds compared to HOM1 and HOM2 indicating the involvement of AhpB in the biosynthesis of **1** and **2**.

Discussion

Aromatic α-pyrones **1** and **2** belong to phenylpropanoid-class of polyketides that have been mostly reported from plants, as well as from synthetic chemistry reports (McCracken et al. 2012; Kraus and Wanninayake 2015; Li et al. 2018). Several similar products with shorter bridging conjugation systems between the aromatic ring and pyrone moiety have also been reported from medicinal fungi *Phellinus* and *Inonotus* spp and *Penicillium glabrum* (Lee and Yun 2011; Hammerschmidt et al. 2012). Containing a long conjugation system, these phenylpropanoid-type compounds exhibit absorptions in the visible spectrum thus could be potentially used

191 in the food industry as natural pigments. The biosynthesis of **1** and **2** starts with a cinnamoyl-CoA unit, which
192 is very common in chalcone-related biosynthesis in higher plants but rarely in filamentous fungi. The
193 truncation of *AhpB* led to the decrease of both compounds, and thereby proved its role in the biosynthetic
194 route. However, additional PAL encoding genes were suspected to contribute to the synthesis of cinnamic acid
195 in *A. homomorphus*. This was supported by the finding of a second putative PAL (Asphom1_476219, identity
196 53.7%, coverage 90.9%) by BLASTP against the database Asphom1_GeneCatalog_proteins_20140716.aa on
197 the JGI MycoCosm using the default settings (Boratyn et al. 2013; Grigoriev et al. 2014).

198 After using cinnamic acid as the starter unit, we hypothesize that the type I iterative polyketide synthase AhpA
199 next catalyzes the biosynthesis of the core backbone of the homopyrones by the addition of three malonyl-
200 CoA extender units to offer a 15-carbon chain. To reach the final conjugated homopyrones, the keto group at
201 C-9 needs to be reduced by a ketoreductase (KR) followed by loss of water likely catalyzed by a dehydratase
202 (DH). The PKS is indeed predicted to contain domains for these reductions. Cinnamoyl-CoA here distinguishes
203 its role as an extender unit in the biosynthesis of splenocin and enterocin in *Streptomyces sp.* CNQ431 (Chang
204 et al. 2015). The *C*-methyltransferase (*C*-MT) domain was likely promiscuous as some intermediate skipped
205 the methylation process leading to the formation of **1**. We speculate that this might be due to the last glycine
206 residue of the conserved SAM binding motif GXGXXG in AhpA replaced by a serine GAGTGS (Figure S4)
207 (Seshime et al. 2009).

208 Unexpectedly, no release domain was observed in AhpA. But instead a freestanding thioesterase (TE) encoding
209 gene named as *AhpC* (Asphom1_356671, 78040 - 78931 on scaffold 68) was found around 1.4 kb upstream of
210 *AhpA*. Interestingly, its hydrolase signature GWSLG showed homology to that of canonical type I TEs GXSXG
211 (Horsman et al. 2016; Zhai et al. 2016). This discrete TE was thus proposed to be responsible for the hydrolytic
212 release and the following spontaneous lactonization. After the pyrone formation, the product was methylated
213 on C-13 by the *O*-methyltransferase encoded by the gene *AhpD* (Asphom1_418652, 79109 - 79804 on scaffold
214 68) as AhpD showed high similarity (59.47% identity, 98% coverage) with a putative *S*-adenosyl-*L*-methionine

dependent methyltransferase (accession number: XP_025385220.1) from *A. eucalypticola* CBS 122712. Further analysis downstream of *AhpA* found a putative gene *AhpE* (Asphom1_418652, 79109 - 79804) likely encoding a ligase, as it showed high similarity (56.99% identity, 99% coverage) with 4-coumarate-CoA ligase-like (4CL) enzyme sequence from *A. awamori* (accession number: GCB23709.1). AhpE might ligate the cinnamic acid to produce cinnamoyl-CoA, which is the activated starter unit in the biosynthesis (Dao et al. 2011; Bang et al. 2016).

Altogether, the biosynthetic steps of **1** and **2** can be proposed as (Figure 5): phenylalanine is de-aminized by AhpB to produce cinnamic acid which is then ligated by AhpE yielding cinnamoyl-CoA as the starter unit. Next, three malonyl-CoA units are used to extend the starter unit during which the keto group originating from the phenylalanine residue is reduced twice by the KR and DH domains of AphA, respectively. Thereafter, *C*-methylation takes place on C-14 located on the third extender unit. The efficiency of *C*-MT is relaxed as some intermediates apparently skip this process resulting in only compound **1** being methylated on C-14. In the final part of the biosynthesis of homopyrones a standalone TE catalyzes the hydrolytic release of the intermediate followed by the spontaneous generation of the pyrone, which is possibly driven by the keto-enol tautomerism of C-11. The maturation of the product is completed after the *O*-methylation on C-13 catalyzed by AhpD.

We investigated the whole-genome sequences of other *Aspergillus* species for BGCs that could produce analogs to **1** and **2**. One potential BGC was found in *A. indologenus* (Figure S5). However, *A. indologenus* is not known to produce yellow compounds (personal communication with Jens C. Frisvad). Hence, this BGC of *A. indologenus* may be silent at laboratory conditions.

Conclusions

In this study, we have successfully identified the BGC related to the biosynthesis of two phenylpropanoid-type yellow pigments **1** and **2** in *A. homomorphus* through a combination of retro-biosynthetic analysis, genome mining, and gene deletions using CRISPR-Cas9 techniques. Chemical analysis of the mutant HOM32 demonstrated the essential role of the PKS AhpA in biosynthesis, whereas the mutant HOM45 showed a

239 decrease in the production of the homopyrones. A second putative PAL enzyme was discovered by BLASTP in
240 *A. homomorphus* which might be responsible for the compensation of cinnamic acid generation in HOM45.
241 Cinnamoyl-CoA is rarely used as the starter unit in biosynthesis of polyketides in filamentous fungi, while
242 cinnamoyl-CoA and its derivative *p*-coumaroyl-CoA are widely used in the biosynthesis of flavonoids in plants
243 (Abe and Morita 2010). Moreover, the *C*-MT domain appears to be promiscuous likely due to the relaxed SAM
244 binding motif GAGTGS. Finally, our work demonstrated that oligonucleotides down to 60 nt could be used as
245 a repair template for gene deletion experiments and, to the best of our knowledge, our report represents the
246 first genetic manipulation in *A. homomorphus* and thus as a benefit sets the scene for further biochemical
247 investigations in this species.

248 ## Author contribution statement

249 All authors contributed to the study conception and design, material preparation, experiment conduction,
250 data collection and analysis, and manuscript drafting were performed by M.E.F, Y.G., C.H., and J.B.H. All
251 authors commented on the manuscript. All authors read and approved the final manuscript.

252 ## Acknowledgments

253 We thank Professor Jens C. Frisvad for sharing his expertise on metabolites of *Aspergillus homomorphus* and
254 *A. indologenus*. Y.G. is a scholarship recipient (201709110107) from the China Scholarship Council. The authors
255 thank the Novo Nordisk Foundation for their grant support (NNF15OC0016610). We thank Sara Bastrup Kreutz
256 Nielsen and Astrid Zedlitz Johansen for their valuable assistance in making the *pyrG* mutation in *A.*
257 *homomorphus*.

258 ## Compliance with ethical standards

259 This article does not contain any studies with human participants or animals.

260 ## Conflict of interest

261 The authors declares that they have no conflict of interest.

References

Abe I, Morita H (2010) Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Nat Prod Rep 27:809–838 . https://doi.org/10.1039/b909988n

Akilandeswari P, Pradeep B V. (2016) Exploration of industrially important pigments from soil fungi. Appl Microbiol Biotechnol 100:1631–1643 . https://doi.org/10.1007/s00253-015-7231-8

Bang HB, Lee YH, Kim SC, Sung CK, Jeong KJ (2016) Metabolic engineering of *Escherichia coli* for the production of cinnamaldehyde. Microb Cell Fact 15:16 . https://doi.org/10.1186/s12934-016-0415-9

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T (2019) AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47:W81–W87 . https://doi.org/10.1093/nar/gkz310

Boeke JD, La Croute F, Fink GR (1984) A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. MGG Mol Gen Genet 197:345–346 . https://doi.org/10.1007/BF00330984

Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y, Raytselis Y, Sayers EW, Tao T, Ye J, Zaretskaya I (2013) BLAST: a more efficient report with usability improvements. Nucleic Acids Res 41:W29–W33 . https://doi.org/10.1093/nar/gkt282

Chang C, Huang R, Yan Y, Ma H, Dai Z, Zhang B, Deng Z, Liu W, Qu X (2015) Uncovering the formation and selection of benzylmalonyl-CoA from the biosynthesis of splenocin and enterocin reveals a versatile way to introduce amino acids into polyketide carbon scaffolds. J Am Chem Soc 137:4183–4190 . https://doi.org/10.1021/jacs.5b00728

Christiana NO (2016) Production of food colourants by filamentous fungi. African J Microbiol Res 10:960–971 . https://doi.org/10.5897/ajmr2016.7904

Dao TTH, Linthorst HJM, Verpoorte R (2011) Chalcone synthase and its functions in plant resistance. Phytochem Rev 10:397–412 . https://doi.org/10.1007/s11101-011-9211-7

Frandsen RJN, Khorsand-Jamal P, Kongstad KT, Nafisi M, Kannangara RM, Staerk D, Okkels FT, Binderup K, Madsen B, Møller BL, Thrane U, Mortensen UH (2018) Heterologous production of the widely used natural food colorant carminic acid in *Aspergillus nidulans*. Sci Rep 8:1–10 . https://doi.org/10.1038/s41598-018-30816-9

Frisvad JC, Larsen TO (2015) Chemodiversity in the genus *Aspergillus*. Appl Microbiol Biotechnol 99:7859–7877 . https://doi.org/10.1007/s00253-015-6839-z

Frisvad JC, Yilmaz N, Thrane U, Rasmussen KB, Houbraken J, Samson R a. (2013) *Talaromyces atroroseus*, a New Species Efficiently Producing Industrially Relevant Red Pigments. PLoS One 8:e84102 . https://doi.org/10.1371/journal.pone.0084102

Gerald F. Bills. James B. Gloer (2016) Biologically Active Secondary Metabolites from the Fungi. In: The Fungal Kingdom. pp 1087–1119

Grigoriev I V., Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I (2014) MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res 42:D699–D704 . https://doi.org/10.1093/nar/gkt1183

Hai Y, Huang A, Tang Y (2020) Biosynthesis of Amino Acid Derived α-Pyrones by an NRPS-NRPKS Hybrid Megasynthetase in Fungi. J Nat Prod 83:593–600 . https://doi.org/10.1021/acs.jnatprod.9b00989

Hammerschmidt L, Wray V, Lin W, Kamilova E, Proksch P, Aly AH (2012) New styrylpyrones from the fungal endophyte *Penicillium glabrum* isolated from *Punica granatum*. Phytochem Lett 5:600–603 . https://doi.org/10.1016/j.phytol.2012.06.003

Horsman ME, Hari TPA, Boddy CN (2016) Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: Logic gate or a victim of fate? Nat Prod Rep 33:183–202 . https://doi.org/10.1039/c4np00148f

Houbraken J, Kocsubé S, Visagie CM, Yilmaz N, Wang XC, Meijer M, Kraak B, Hubka V, Bensch K, Samson RA, Frisvad JC (2020) Classification of *Aspergillus*, *Penicillium*, *Talaromyces* and related genera (*Eurotiales*):

An overview of families, genera, subgenera, sections, series and species. Stud Mycol 95:5–169 . https://doi.org/10.1016/j.simyco.2020.05.002
Isbrandt T, Tolborg G, Ødum A, Workman M, Larsen TO (2020) Atrorosins: a new subgroup of Monascus pigments from *Talaromyces atroroseus*. Appl Microbiol Biotechnol 104:615–622 . https://doi.org/10.1007/s00253-019-10216-3
Jakob B. Hoof, Christina S. Nødvig and UHM (2018) Chapter 11 Genome Editing: CRISPR-Cas9. In: Fungal Genomics. Humana Press, New York, NY., pp 119–132
Kjærbølling I, Mortensen UH, Vesth T, Andersen MR (2019) Strategies to establish the link between biosynthetic gene clusters and secondary metabolites. Fungal Genet Biol 130:107–121 . https://doi.org/10.1016/j.fgb.2019.06.001
Kraus GA, Wanninayake UK (2015) An improved aldol protocol for the preparation of 6-styrenylpyrones. Tetrahedron Lett 56:7112–7114 . https://doi.org/10.1016/j.tetlet.2015.11.021
Lee IK, Yun BS (2011) Styrylpyrone-class compounds from medicinal fungi *Phellinus* and *Inonotus* spp., and their medicinal importance. J Antibiot (Tokyo) 64:349–359 . https://doi.org/10.1038/ja.2011.2
Li C, Cheng B, Fang S, Zhou H, Gu Q, Xu J (2018) Design, syntheses and lipid accumulation inhibitory activities of novel resveratrol mimics. Eur J Med Chem 143:114–122 . https://doi.org/10.1016/j.ejmech.2017.11.017
McCracken ST, Kaiser M, Boshoff HI, Boyd PDW, Copp BR (2012) Synthesis and antimalarial and antituberculosis activities of a series of natural and unnatural 4-methoxy-6-styryl-pyran-2-ones, dihydro analogues and photo-dimers. Bioorganic Med Chem 20:1482–1493 . https://doi.org/10.1016/j.bmc.2011.12.053
Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, Gough J, Haft DR, Huang H, Letunic I, Lopez R, Luciani A, Madeira F, Marchler-Bauer A, Mi H, Natale DA, Necci M, Nuka G, Orengo C, Pandurangan AP, Paysan-Lafosse T, Pesseat S, Potter SC, Qureshi MA, Rawlings ND, Redaschi N, Richardson LJ, Rivoire C, Salazar GA, Sangrador-Vegas A, Sigrist CJA, Sillitoe I, Sutton GG, Thanki N, Thomas PD, Tosatto SCE, Yong S-Y, Finn RD (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res 47:D351–D360 . https://doi.org/10.1093/nar/gky1100
Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH (2006) Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans. Fungal Genet Biol 43:54–64 . https://doi.org/10.1016/j.fgb.2005.09.005
Nielsen ML, Isbrandt T, Rasmussen KB, Thrane U, Hoof JB, Larsen TO, Mortensen UH (2017) Genes Linked to production of secondary metabolites in *Talaromyces atroroseus* revealed using CRISPR-Cas9. PLoS One 12:1–9 . https://doi.org/10.1371/journal.pone.0169712
Nødvig CS, Hoof JB, Kogle ME, Jarczynska ZD, Lehmbeck J, Klitgaard DK, Mortensen UH (2018) Efficient oligo nucleotide mediated CRISPR-Cas9 gene editing in Aspergilli. Fungal Genet Biol 115:78–89 . https://doi.org/10.1016/j.fgb.2018.01.004
Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH (2015) A CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS One 10:1–18 . https://doi.org/10.1371/journal.pone.0133085
Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev I V., Dubchak I (2014) The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 42:26–31 . https://doi.org/10.1093/nar/gkt1069
Nørholm MH (2010) A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol 10:21 . https://doi.org/10.1186/1472-6750-10-21
Oliver Kayser, W. Ray Waters, Keith M. Woods, Steve J. Upton, Janet S. Keithly AFK (2001) Evaluation of in vitro activity of aurones and related compounds against *Cryptosporidium parvum*. Planta Med 3:722–725
Samson, Robert A. ; Houbraken, Jos ; Thrane, Ulf ; Frisvad, Jens Christian ; Andersen B (2010) Food and indoor fungi: Second Edition. CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands.
Seshime Y, Juvvadi PR, Tokuoka M, Koyama Y, Kitamoto K, Ebizuka Y, Fujii I (2009) Functional expression of the *Aspergillus flavus* PKS-NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid. Bioorganic

Med Chem Lett 19:3288–3292 . https://doi.org/10.1016/j.bmcl.2009.04.073

Theobald S, Vesth TC, Rendsvig JK, Nielsen KF, Riley R, de Abreu LM, Salamov A, Frisvad JC, Larsen TO, Andersen MR, Hoof JB (2018) Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication. Sci Rep 8:1–12 . https://doi.org/10.1038/s41598-018-36561-3

Tolborg G, Ødum ASR, Isbrandt T, Larsen TO, Workman M (2020) Unique processes yielding pure azaphilones in *Talaromyces atroroseus*. Appl Microbiol Biotechnol 104:603–613 . https://doi.org/10.1007/s00253-019-10112-w

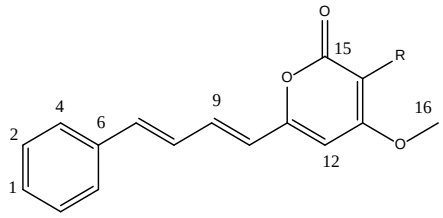
Vesth TC, Nybo JL, Theobald S, Frisvad JC, Larsen TO, Nielsen KF, Hoof JB, Brandl J, Salamov A, Riley R, Gladden JM, Phatale P, Nielsen MT, Lyhne EK, Kogle ME, Strasser K, McDonnell E, Barry K, Clum A, Chen C, LaButti K, Haridas S, Nolan M, Sandor L, Kuo A, Lipzen A, Hainaut M, Drula E, Tsang A, Magnuson JK, Henrissat B, Wiebenga A, Simmons BA, Mäkelä MR, de Vries RP, Grigoriev I V., Mortensen UH, Baker SE, Andersen MR (2018) Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. Nat Genet 50:1688–1695 . https://doi.org/10.1038/s41588-018-0246-1

Wolff PB, Nielsen ML, Slot JC, Andersen LN, Petersen LM, Isbrandt T, Holm DK, Mortensen UH, Nødvig CS, Larsen TO, Hoof JB (2020) Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*. Fungal Genet Biol 139:103378 . https://doi.org/10.1016/j.fgb.2020.103378

Zabala AO, Xu W, Chooi YH, Tang Y (2012) Characterization of a silent azaphilone gene cluster from *Aspergillus niger* ATCC 1015 reveals a hydroxylation-mediated pyran-ring formation. Chem Biol 19:1049–1059 . https://doi.org/10.1016/j.chembiol.2012.07.004

Zhai Y, Bai S, Liu J, Yang L, Han L, Huang X, He J (2016) Identification of an unusual type II thioesterase in the dithiolopyrrolone antibiotics biosynthetic pathway. Biochem Biophys Res Commun 473:329–335 . https://doi.org/10.1016/j.bbrc.2016.03.105

Figures



1: R = H
2: R = Me

Figure 1: Structures of homopyrones A (**1**) and B (**2**).

Figure 2: Predicted biosynthetic gene cluster for homopyrones A (**1**) and B (**2**).

(Predicted functions of each gene based on antiSMASH and BLASTP. AhpA: PKS, AhpB: phenylalanine ammonia-lyase, AhpC: thioesterase, AhpD: O-methyltransferase, AhpE: ligase. Domains of AhpA, KS: ketosynthase, AT: acyltransferase, DH: dehydratase, *C*MT: *C*-methyltransferase, KR: ketoreductase, ACP: acyl carrier protein)

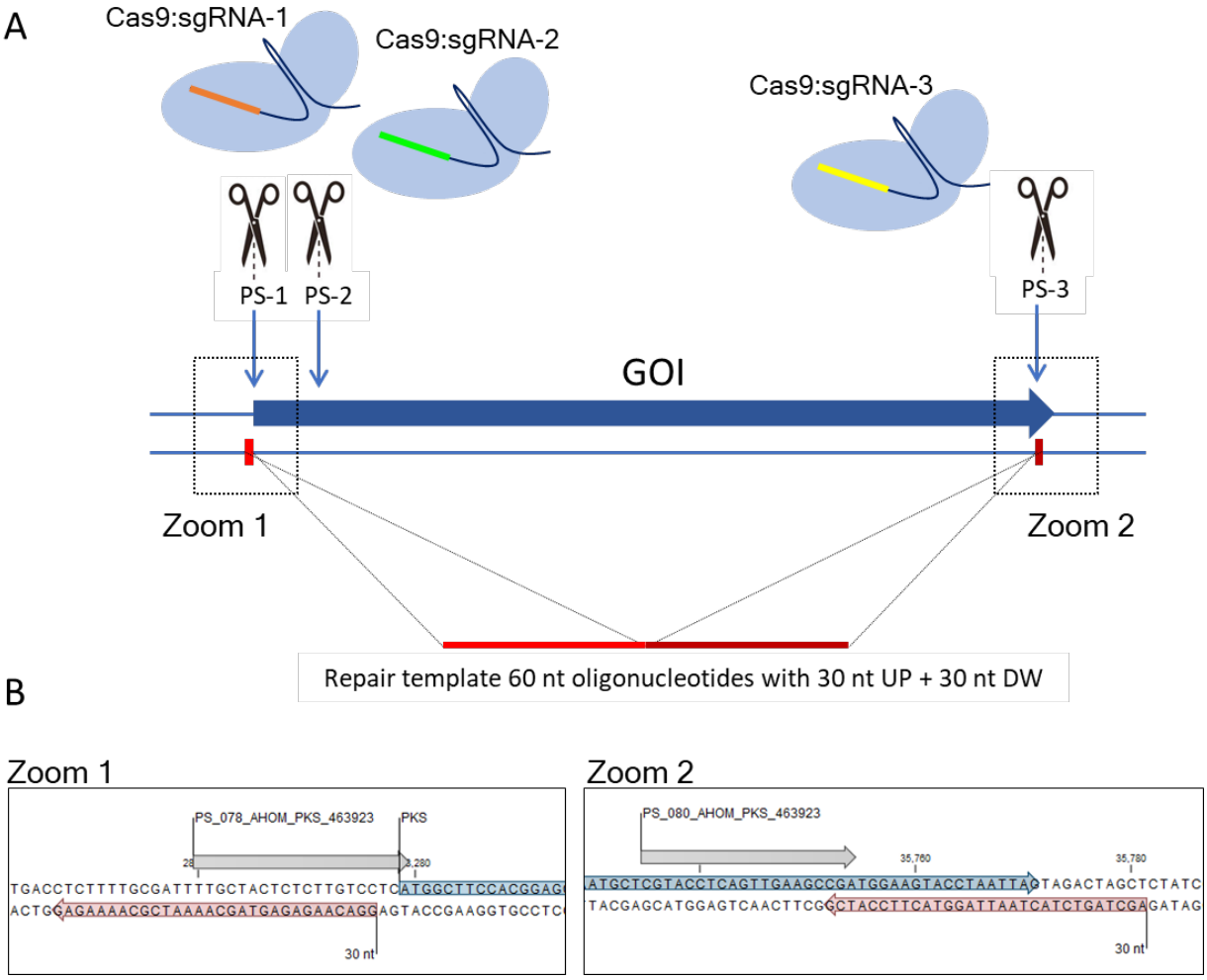
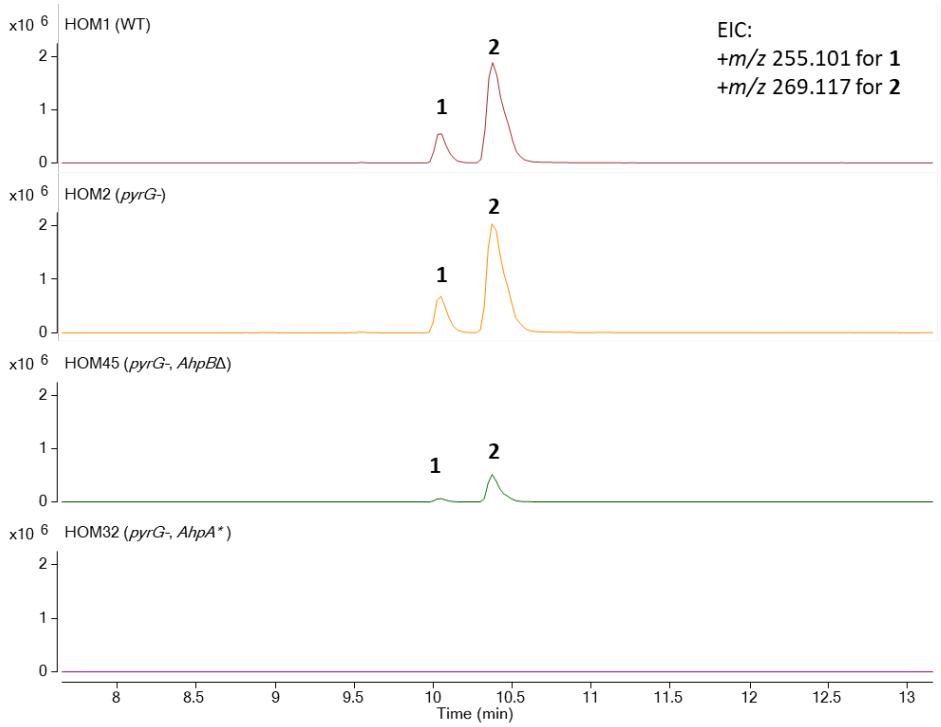


Figure 3. Strategy for CRISPR-Cas9 mediated deletions in *Aspergillus homomorphus*. (GOI, the gene of Interest; Cas9:sgRNA1-3 - the complex of Cas9 and single-guide RNA; colors to indicate three different protospacers; zooms 1 and 2 to illustrate the design of repair template 30 nt + 30 nt and position of the protospacers.)



406

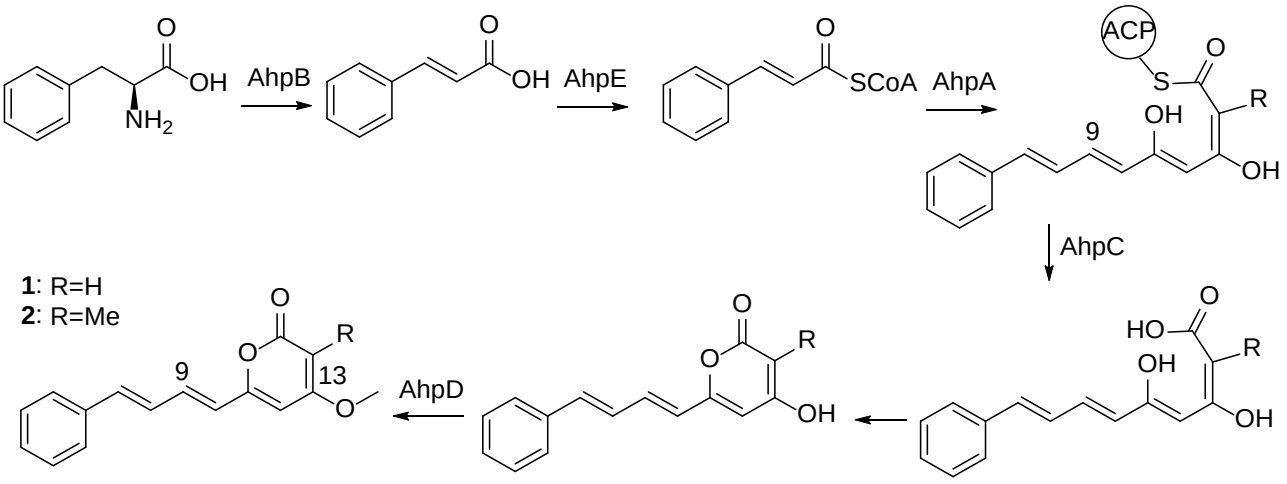
407 Figure 4. Production of homopyrones A (**1**) and B (**2**) in the mutants and reference strains.

408 *gene truncation

409

410

411



412

413 Figure 5. Proposed biosynthetic pathway for homopyrones A (**1**) and B (**2**).

414

Genetic origin of homopyrones, a rare type of hybrid phenylpropanoid and polyketide derived yellow pigment from *Aspergillus homomorphus*

Malgorzata E. Futyma,[1,4,5] Yaojie Guo,[1,5] Casper Hoeck,[2,3] Jakob B. Hoof,[1] Charlotte H. Gotfredsen,[2] Uffe H. Mortensen,[1*] Thomas O. Larsen,[1*]

*Corresponding authors, um@bio.dtu.dk, tol@bio.dtu.dk

Contents

[1] Department of Biotechnology and Biomedicine, Søltofts Plads, Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark
[2] Department of Chemistry, Kemitorvet B207, Technical University of Denmark, Kgs. Lyngby DK-2800, Denmark
[3] Novo Nordisk A/S, Smørmosevej 10-12, 2880 Bagsværd, Denmark
[4] CHRETO, Lejrvej 17, 3500 Værløse, Denmark
[5] The authors contributed equally to this work

Table S1 NMR Data (500 MHz ^{1}H; 125 MHz ^{13}C) of **1** and **2** in DMSO-d_6.

Position	**1**		**2**	
	$\delta_{C,\ type}$	δ_{H}, multi (*J* in Hz)	$\delta_{C,\ type}$	δ_{H}, multi (*J* in Hz)
1	128.3, CH	7.30, t (7.3)	128.5, CH	7.30, t (7.3)
2	128.5, CH	7.38, t (7.6)	128.8, CH	7.38, t (7.6)
3	128.5, CH	7.38, t (7.6)	128.8, CH	7.38, t (7.6)
4	126.6, CH	7.55, d (7.4)	126.9, CH	7.56, d (7.4)
5	126.6, CH	7.55, d (7.4)	126.9, CH	7.56, d (7.4)
6	136.1, C	-	136.4, C	-
7	137.4, CH	7.01, d (14.8)	137.4, CH	7.00, d (14.7)
8	127.2, CH	7.10, m	127.4, CH	7.13, m
9	134.0, CH	7.15, m	134.5, CH	7.16, m
10	122.5, CH	6.44, d (14.6)	123.0, CH	6.47, d (14.5)
11	157.7, C	-	156.7, C	-
12	100.8, C	6.31, d (2.1)	96.9, C	6.72, s
13	170.7, C	-	165.9, C	-
14	88.3, CH	5.62, d (2.1)	100.8, C	-
15	162.3, C	-	163.3, C	-
16	56.1, CH_3	3.82, s	56.6, CH_3	3.92, s
17	-	-	8.8, CH_3	1.82, s

Table S2 Primers used in this study.

Code	Description[a]	Sequence 5'-> 3'
CSN438	CSN438 Fw Afum-U3p-fwd	GGGTTTAAUGATCACATAGATGCTCGGTTGACA
MF767	PS_78_HOM_PKS_463923-R	AGAGTAGCAAUGCATCATCCGTGAATCGAAC
MF768	PS_78_HOM_PKS_463923-F	ATTGCTACTCUCTTGTCCTCAGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF769	PS_79_HOM_PKS_463923-R	ATAAGCCTCGUGCATCATCCGTGAATCGAAC
MF770	PS_79_HOM_PKS_463923-F	ACGAGGCTTAUTTCCTCGAGGGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF771	PS_80_HOM_PKS_463923-R	AACTGAGGUACGTGCATCATCCGTGAATCGAAC
MF772	PS_80_HOM_PKS_463923-F	ACCTCAGTUGAAGCCGAGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF773	UP check HOM PKS_463923-F	CGGCTAATCAGTTTGAGTTGC
MF774	UP check HOM PKS_463923-R	GTAAACTACCTCCAGCAAGAAACG
MF775	DW check HOM PKS_463923-F	GCATCTGCCTGGTGACTCG
MF776	DW check HOM PKS_463923-R	GGGCTGAATCTTGAGTGTCG
CSN790	CSN790 Rev U3-term-rv	GGTCTTAAUACCCTGAGAAGATAGATGTGAATGTG
MF777	PS_81_HOM_PAL_463919-R	ACGGCCACGAGUGCATCATCCGTGAATCGAAC
MF778	PS_81_HOM_PAL_463919-F	ACTCGTGGCCGUGAACCGTGTGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF779	PS_82_HOM_PAL_463919-R	ACAGCCATCUGCATCATCCGTGAATCGAAC
MF780	PS_82_HOM_PAL_463919-F	AGATGGCTGUCAATTGACTATGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF781	PS_83_HOM_PAL_463919-R	ATTTCAGCAAUGCATCATCCGTGAATCGAAC
MF782	PS_83_HOM_PAL_463919-F	ATTGCTGAAAUCTTTCGATTAGTTTTAGAGCTAGAAATAGCAAGTTAAA
MF783	UP check HOM PAL_463919-F	ATATAGACAGGCTCGGGGAGC
MF784	UP check HOM PAL_463919-R	GCTCTCAGCTCACCATAGATGTG
MF785	DW check HOM PAL_463919-F	CCGTCGGTCCGTCAGAGC
MF786	DW check HOM PAL_463919-R	GCCTCCTGTTCATCGAAAGC
MF787	GE-oligo to delete HOM_PKS_463923	AGCTAGTCTACTAATTAGGTACTTCCATCGGGACAAGAGAGTAGCAAAATCGCAAAAGAG
MF788	GE- oligo to delete HOM_PAL_463919	ATTCCTCTACGCAAAGCCCCTTACCCCTAACGGTTCACGGCCACGAGCATGATCCACGCC
MF233	Sequencing CRISPR-tRNA vectors-F	AGGTTCTCCTAACGCTTGGC
MF416	Sequencing CRISPR-tRNA vectors-R	AGAATTTGAGCAAACTCTGATCG
Primers for pyrG mutant		
B561	CSN791-tRNA-general-rv	ATCATCCGUGAATCGAAC
B581	sing_gRNA-HethompyrG_PS	ACGGATGAUGCAGCACAACTTCCTCATCTTCGGTTTTAGAGCTAGAAATAGCAAGTTAAA
B666	AHOMpyrG-ChkSeq-F	CAAGCACCCCAACCCTCTG
B667	AHOMpyrG-ChkSeq-R	CGACACGAAACCCATCAC

[a] F = forward, R= reverse. Primer pairs for CRISPR/Cas9-vectors: HOM_PKS_463923: CSN438+MF767, MF768+769, MF770+771, MF772+CSN790; HOM_PAL_463919: CSN438+MF777, MF778+779, MF780+781, MF782+CSN790; HOM_*pyrG*: CSN438+B561, B581+CSN790.

Table S3 Protospacers designed in this study.

Protospacer ID	Description[a]	Sequence 5'-> 3'
PS_hethom	PS_hethom_*pyrG*_468628	GCACAACTTCCTCATCTTCG
PS_078	PS_078_AHOM_PKS_463923	TTGCTACTCTCTTGTCCTCA
PS_079	PS_079_AHOM_PKS_463923	CGAGGCTTATTTCCTCGAGG
PS_080	PS_080_AHOM_PKS_463923	CGTACCTCAGTTGAAGCCGA
PS_081	PS_081_AHOM_PAL_463919	CTCGTGGCCGTGAACCGTGT
PS_082	PS_082_AHOM_PAL_463919	GATGGCTGTCAATTGACTAT
PS_083	PS_083_AHOM_PAL_463919	TTGCTGAAATCTTTCGATTA

[a]The numbers refer to protein IDs in Mycocosm/JGI

Table S4 Plasmids used in this study.

Plasmid ID	Description	Purpose
pFC902	pFC902	PCR template for promoter, terminator, tRNA, and sgRNA amplification
pFC330	pFC330	CRISPR/Cas9-vector assembly by USER-cloning with AF*pyrG* selection
pFC332	pFC332	CRISPR/Cas9-vector assembly by USER-cloning with hygromycin resistance (hph)
pHEHOP	pFC332-Het/Hom-pyrG	Targeting AHOM_pyrG468628
pAC1604	pFC330-3PSs-tRNA-HOM-PKS463923	Targeting AHOM_PKS463923
pAC1606	pFC330-3PSs-tRNA-HOM-PAL463919	Targeting AHOM_PAL463919

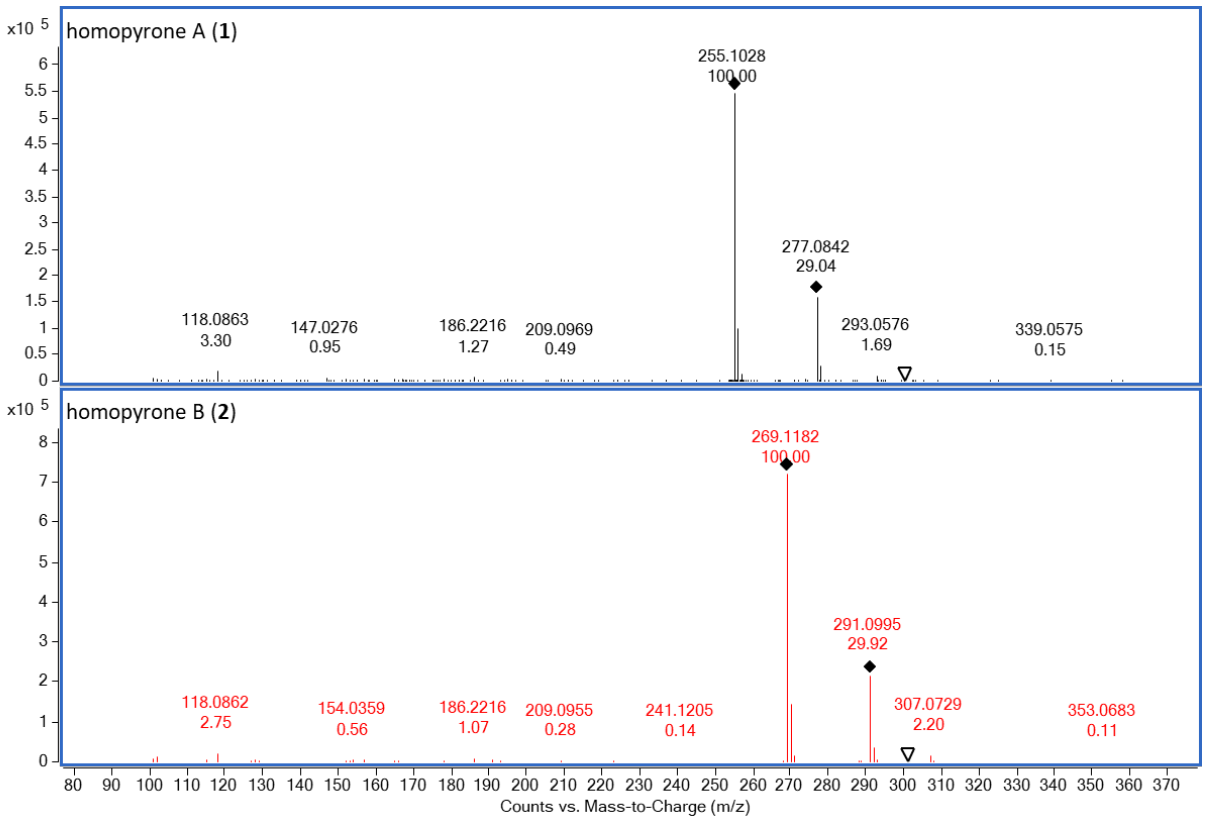


Figure S1 HRMS of homopyrones A (**1**) and B (**2**).

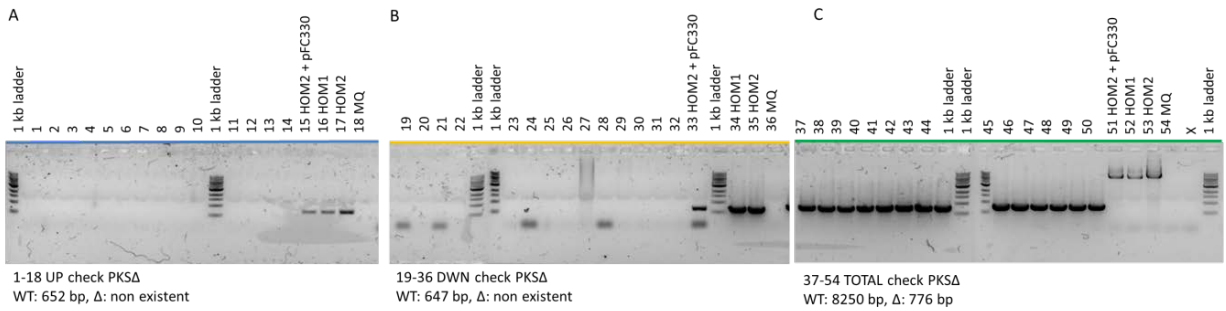


Figure S2 PCR verification of PKS gene *AhpA* deletion.

(A: UP check (MF773+774), B: DW check (MF775+776), C: total check (MF773+776); primers in parentheses listed in Table S2; all 14 colonies (1-14) from spore dissection showed clean gene deletions, colony 1 = HOM32; 1 kb ladder bottom to top: 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 kb)

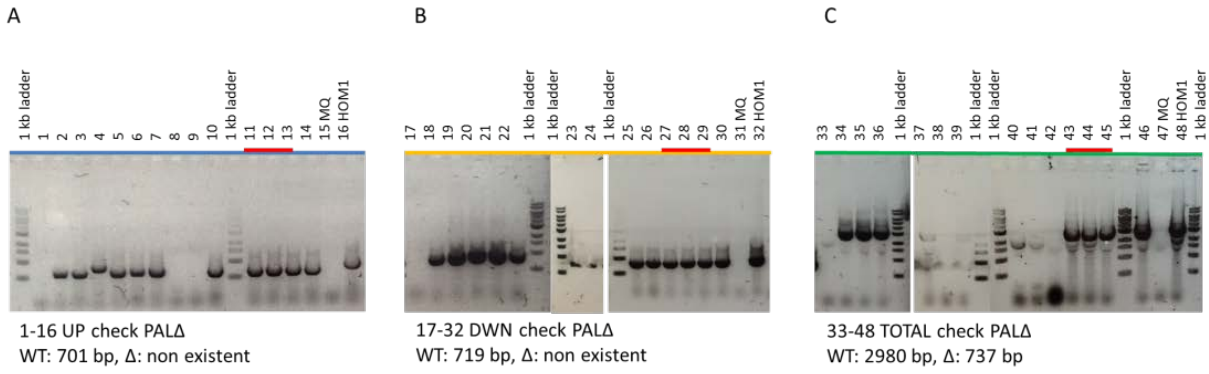


Figure S3 PCR verification of PAL gene *AhpB* truncation.

(A: UP check (MF783+784), B: DW check (MF785+786), C: total check (MF783+786);primers in parentheses listed in Table S2; 10 out of 14 colonies (1-14) from spore dissection showed truncation at 5´end including colonies 2-3, 5-7 and 10-14; the 3´ends of all 14 colonies remain unchanged; colony 11 = HOM45; 1 kb ladder bottom to top: 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 kb)

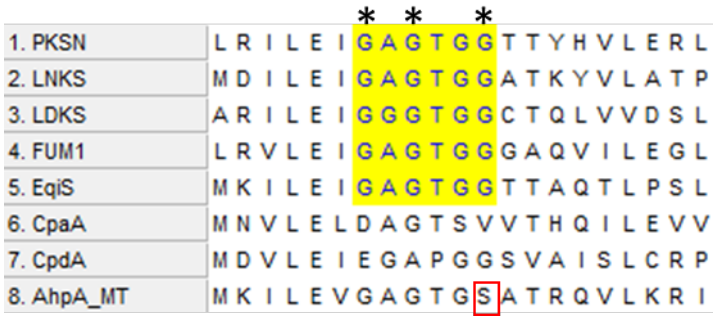


Figure S4. Alignment of the MT domain sequences of AhpA with other fungal PKSs or PKS-NRPSs.

Sequences were aligned using ClustalW algorithm with MEGA-X V.10.1.8 (Kumar et al. 2018); SAM binding site conserved amino acids GXGXXG are indicated by asterisk mark. All the sequences used for alignment were collected from (Seshime et al. 2009). CpdA sequence was from the unpublished data in our lab.

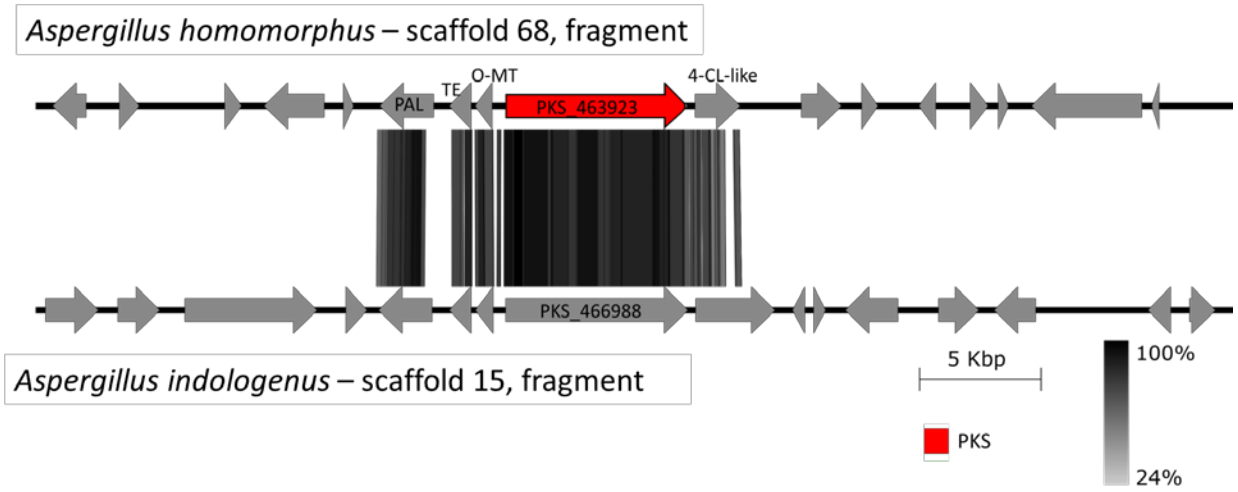


Figure S5 Synteny plot of the *Ahp* gene cluster and *A. homomorphus* and the homologous cluster in *A. indologenus*.

Sequences were aligned using Easyfig (Sullivan et al. 2011).

Supplementary references

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549 . https://doi.org/10.1093/molbev/msy096

Seshime Y, Juvvadi PR, Tokuoka M, Koyama Y, Kitamoto K, Ebizuka Y, Fujii I (2009) Functional expression of the *Aspergillus flavus* PKS-NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid. Bioorganic Med Chem Lett 19:3288–3292 . https://doi.org/10.1016/j.bmcl.2009.04.073

Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: A genome comparison visualizer. Bioinformatics 27:1009–1010 . https://doi.org/10.1093/bioinformatics/btr039

Efforts on linking oxepinamide L to its biosynthetic gene cluster and linking five genes to the cryptic metabolites in *Aspergillus californicus*

Yaojie Guo, Fabiano J. Contesini, Uffe H. Mortensen, and Thomas O. Larsen

Abstract: During our continuing discovery for new secondary metabolites from *Aspergillus californicus*, two previously undescribed nonribosomal peptides oxepinamides L and M bearing a unique oxepine-pyrimidinone-ketopiperazine (OPK) scaffold were characterized. Due to the intriguing structure of the oxepine unit, a collection of work has been conducted to characterize their biosynthetic pathway including retro-biosynthesis, genome mining, and deletions of five proposed biosynthetic genes based on the *pyrG*- and *ckuAΔ* mutant with CRISPR-Cas9 technologies. However, chemical analysis of the deletion strains indicated that we have proposed the wrong gene cluster as neither of the mutant showed changes in the production of both compounds. A new comprehensive proposal of the biosynthetic mechanisms was thus needed. Besides, efforts on discovering the compounds encoded by the five deleted genes were carried out, but these potential metabolites were likely cryptic. More work is needed to realize the biosynthetic abilities of these five genes.

Introduction

Two oxepine-pyrimidinone-ketopiperazine (OPK) metabolites oxepinamides L and M (Figure 1), were characterized from *Aspergillus californicus* as described in Appendix 3. The OPK class of nonribosomal peptides has only been discovered from fungal sources so far, and their biosynthesis remains unknown.[1] To investigate the biosynthetic pathway of oxepinamides, strategies including retro-biosynthesis, bioinformatics based biosynthetic gene cluster (BGC) prediction, and multiple gene deletion experiments were performed.

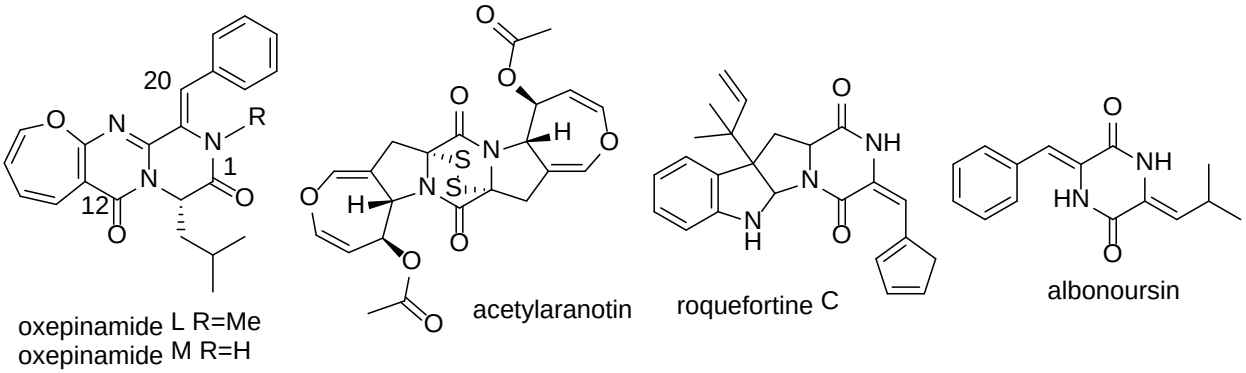


Figure 1. Structures of oxepinamides L and M, and other molecules with l similar biosynthetic steps.

Knowledge-based retro-biosynthesis

Based on the retro-biosynthesis and knowledge of the formation of fumiquinazolines (FQs) in *Aspergillus* and *Penicillium* species,[2–4] a plausible biosynthetic pathway of oxepinamide L was proposed (Figure 2). Initially, an NRPS gene with three Adenylation (A) domains is likely responsible for the incorporation of three amino acids including anthranilic acid, phenylalanine, and leucine, as well as the formation of the core quinazoline ring structure similar to the biosynthesis of fumiquinazolines. This is then followed by two or three tailoring steps including the oxepine formation, a hypothetical hydroxyl formation, and dehydration for a new double bond formation between C-3 and C-20, or the double band formation was catalyzed by some monooxygenase such as Pc21g15470 (albonoursin biosynthesis) or aminoacyl α, β-dehydrogenase with a FAD co-factor (acetylaranotin).[5,6] Generation of the oxepine moiety in natural products, to the best of our knowledge, is still unclear. A study indicated that AtaF, a cytochrome P450 monooxygenase by gene ATEG_03471.1, and AtaY, a putative p-hydroxylase encoded by ATEG_03468.1, are responsible for the dihydrooxepine

formation from a hydrolyzed benzyl ring during the biosynthesis of acetylaranotin in *A. terreus* through gene deletion experiments.[7] Benzoyl-CoA epoxidase (BoxB), a dinuclear iron enzyme catalyzing the epoxidation reaction of the aromatic ring of benzoyl-CoA, generates the oxepine structure, which is a catalytic strategy by many bacteria to cleave the aromatic pollutants.[8,9] Similar enzymes phenyl acetyl-CoA epoxidase can also generate oxepine intermediates.[10] A bacterial P450-catalyzed oxidation study in vitro showed that the equilibrium between the oxpine product and arene oxide was related to the solvent and temperature, non-polar solvent, and high temperature favoring the oxepine state (Figure 1 compounds **6**, **2a** and **4a**).[11] Though no biosynthetic P450s have been associated with the oxepine production in natural products, an oxidase was assumed essential for its generation from the aromatic ring.

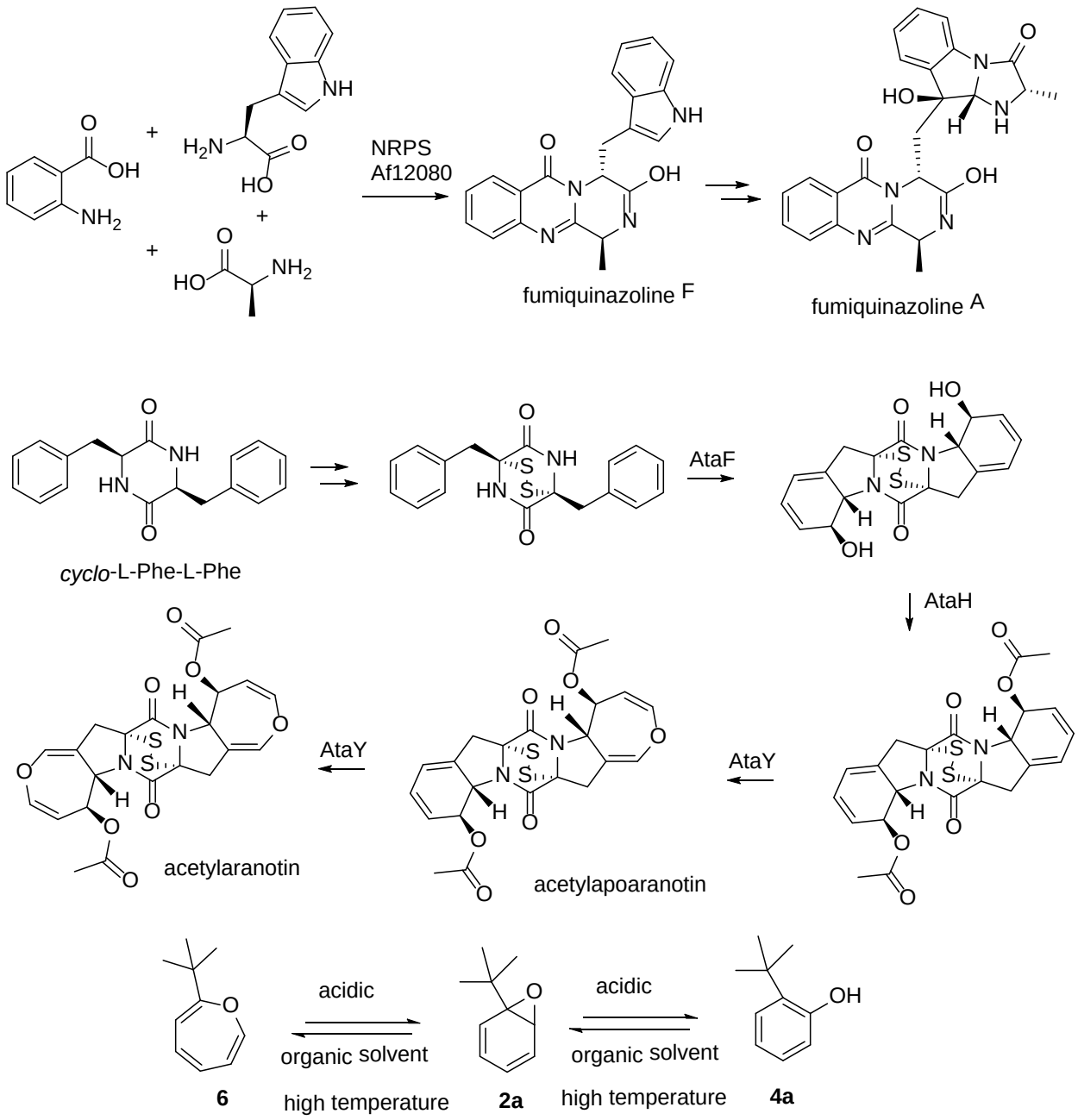


Figure 2. Generation of oxepine related products and fumiquinazolines in microorganisms.

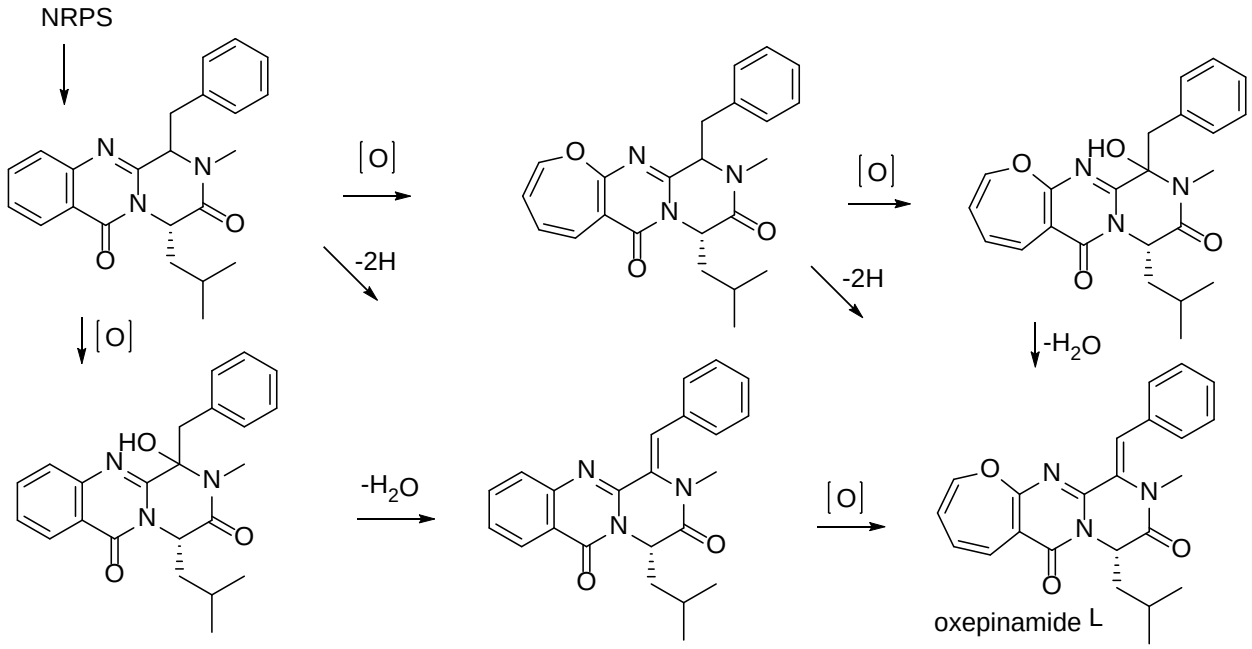


Figure 3. Proposed late-stage biosynthetic steps of oxepinamides L.

BGC prediction

The whole-genome sequence of *A. californicus* CBS 123895 (= IBT 16748) obtained from Joint Genome Institute was submitted to the antiSMASH platform to predict secondary metabolites gene clusters.[12,13] Based on the abovementioned proposal, we identified 8 BGCs (referred to as ´regions´ by antiSMASH) out of 91 that contain one NRPS gene with three A domains including BGCs 2.2, 11.2, 33.1, 55.1, 76.1, 84.1, 122.1, and 156.1. Four of these BGCs (11.2, 55.1, 76.1, and 122.1) were down prioritized due to the lack of a methyltransferase (MT) encoding gene or gene module. BGCs 33.1 and 156.1 both have an *O*-MT gene and thus also get down prioritized due to the addition of the methyl group on N-2 in oxepinamide L. BGC 84.1 contains an NRPS gene with domain organization A-PCP-C-A-(*N*-MT)-PCP-C-A-PCP-C, one gene jgi.p_Aspcalif1_261467 likely encoding FAD linked oxidase, and 15 genes with little information through antiSMASH analysis or by BLASTP.[14]

BGC 2.2 contains one NRPS gene with domain organization A-PCP-C-A-PCP-C-A-(*N*-MT)-PCP-C, two P450 encoding genes, one gene encoding short-chain dehydrogenase/reductase (SDR), one *O*-MT gene, and several other surrounding genes. This region was selected as the most promising gene cluster based on the consideration that one P450 might contribute to the oxepine formation, an SDR could catalyze the double bond formation.[15–18]

Gene deletion and chemical analysis

Five genes from BGC 2.2 were thus selected for gene deletion experiments including jgi.p_Aspcalif1_252836 (P450), jgi.p_Aspcalif1_252838 (NRPS), jgi.p_Aspcalif1_252843 (P450), jgi.p_Aspcalif1_82477 (SDR), and jgi.p_Aspcalif1_82488 (TF). However, chemical analysis of the mutants didn´t show any abolishment of the production of oxepinamide L (Figure S1), which indicated that we selected the wrong gene cluster.

Efforts on linking the five deleted genes to the cryptic metabolites

To explore what metabolites those five genes were encoding, six solid media were used to cultivate the mutants (CAL005-CAL009) including CYA, MEAox, OAT, PDA, YES, and YPD, all supplemented with 10 mM Uri+Ura.[19,20] However, the chemical profiles of mutants CAL001 and CAL005-009 were practically identical, like no phenotype differences were observed between the different mutant strains. The obvious LC-MS difference was observed between CAL000 (wild-type) and CAL001 possibly due to the stress of the excess uridine (Uri) and uracil (Ura) to mutants with *pyrG* deleted, which could be seen from the phenotype (Figure S8). The inhibition effect of excess Ura and its derivatives has also been reported on *A. nidulans* pyrG89.[21]

Discussions

Since BGC 2.2 proved not to be the responsible gene cluster for the biosynthesis of oxepinamide L, we hypothesize that BGC 84.1 with an *N*-MT module within the NRPS gene could be the most likely responsible cluster assuming that the methyl group was catalyzed by the *N*-MT domain of NRPS. But more work is needed to understand the surrounding hypothetic genes. The isolation of oxepinamide M might be a small portion of the intermediate skipping the *N*-methylation process before being released from the assembly line during the synthesis of oxepinamide L. A second explanation of the formation of oxepinamide M could be that the methylation was catalyzed by some MT after the intermediate release from an NRPS without *N*-MT domain. A recent study showed that the *N*-methylation could be catalyzed by individual genes like *pjcyA* or *pscyA* in the biosynthesis of cycloaspeptides,[22] and such methylation has only been found in primary metabolism *N*-methyltransferase before.[23] A recent research showed the *N*-methyl group was added by the

methyltransferase NanE in the biosynthesis of nanangelenin A from a novel species of Australian fungus, *Aspergillus nanangensis*.[24] In the current way Oxepinamide M would be the precursor of oxepinamide L. In the second scenario, BGC 33.1 encoding one NRPS, P450, *O*-MT, SDR, and FAD-oxidase, and BGC 156.1 encoding one NRPS, P450 and *O*-MT were also plausible to synthesize oxepinamide L to some extent. A parallel gene deletion experiments targeting three NRPS gene of these three BGCs were suggested in further study to pinpoint the responsible NRPS. Besides, considering whole genome sequences of 228 *Aspergillus* isolates are available on Joint Genome Institute,[12] comparing the genome sequences of the known OPK NRPS producers would also facilitate finding the correct BGC.[1,25]

To find the cryptic metabolites encoded by the five genes that have been deleted, more work is needed like 1) One Strain-Many Compounds (OSMAC)[26] method using liquid media or different temperatures; 2) co-culture with other microorganisms; 3) overexpression of the two transcriptional factors (Gene IDs: jgi.p_Aspcalif1_82484 and jgi.p_Aspcalif1_217404).

References

(1) Guo, Y.; Frisvad, J. C.; Larsen, T. O. Review of Oxepine-Pyrimidinone-Ketopiperazine Type Nonribosomal Peptides. *Metabolites* **2020**, *10* (6), 11–14.

(2) Ames, B. D.; Walsh, C. T. Anthranilate-Activating Modules from Fungal Nonribosomal Peptide Assembly Lines. *Biochemistry* **2010**, *49* (15), 3351–3365.

(3) Ames, B. D.; Liu, X.; Walsh, C. T. Enzymatic Processing of Fumiquinazoline F: A Tandem Oxidative-Acylation Strategy for the Generation of Multicyclic Scaffolds in Fungal Indole Alkaloid Biosynthesis. *Biochemistry* **2010**, *49* (39), 8564–8576.

(4) Gao, X.; Chooi, Y. H.; Ames, B. D.; Wang, P.; Walsh, C. T.; Tang, Y. Fungal Indole Alkaloid Biosynthesis: Genetic and Biochemical Investigation of the Tryptoquialanine Pathway in *Penicillium aethiopicum*. *J. Am. Chem. Soc.* **2011**, *133* (8), 2729–2741.

(5) García-Estrada, C.; Ullán, R. V.; Albillos, S. M.; Fernández-Bodega, M. Á.; Durek, P.; Von Döhren, H.; Martín, J. F. A Single Cluster of Coregulated Genes Encodes the Biosynthesis of the Mycotoxins Roquefortine C and Meleagrin in *Penicillium chrysogenum*. *Chem. Biol.* **2011**, *18* (11), 1499–1512.

(6) Gondry, M.; Lautru, S.; Fusai, G.; Meunier, G.; Ménez, A.; Genet, R. Cyclic Dipeptide Oxidase from *Streptomyces noursei* . *Eur. J. Biochem.* **2001**, *268* (6), 1712–1721.

(7) Guo, C. J.; Yeh, H. H.; Chiang, Y. M.; Sanchez, J. F.; Chang, S. L.; Bruno, K. S.; Wang, C. C. C. Biosynthetic Pathway for the Epipolythiodioxopiperazine Acetylaranotin in *Aspergillus terreus* Revealed by Genome-Based Deletion Analysis. *J. Am. Chem. Soc.* **2013**, *135* (19), 7205–7213.

(8) Rather, L. J.; Weinert, T.; Demmer, U.; Bill, E.; Ismail, W.; Fuchs, G.; Ermler, U. Structure and Mechanism of the Diiron Benzoyl-Coenzyme A Epoxidase BoxB. *J. Biol. Chem.* **2011**, *286* (33), 29241–29248.

(9) Liao, R. Z.; Siegbahn, P. E. M. Mechanism and Selectivity of the Dinuclear Iron Benzoyl-Coenzyme A Epoxidase BoxB. *Chem. Sci.* **2015**, *6* (5), 2754–2764.

(10) Grishin, A. M.; Ajamian, E.; Zhang, L.; Rouiller, I.; Bostina, M.; Cygler, M. Protein-Protein Interactions in the β-Oxidation Part of the Phenylacetate Utilization Pathway: Crystal Structure of the PaaF-PaaG Hydratase-Isomerase Complex. *J. Biol. Chem.* **2012**, *287* (45), 37986–37996.

(11) Stok, J. E.; Chow, S.; Krenske, E. H.; Farfan Soto, C.; Matyas, C.; Poirier, R. A.; Williams, C. M.; De Voss, J. J. Direct Observation of an Oxepin from a Bacterial Cytochrome P450-Catalyzed Oxidation. *Chem. - A Eur. J.* **2016**, *22* (13), 4408–4412.

(12) Nordberg, H.; 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* (D1), 26–31.

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

(14) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; et al. BLAST: A More Efficient Report with Usability Improvements. *Nucleic Acids Res.* **2013**, *41* (Web Server issue), W29–

W33.

(15) Loss, E. M. O.; Lee, M. K.; Wu, M. Y.; Martien, J.; Chen, W.; Amador-Noguez, D.; Jefcoate, C.; Remucal, C.; Jung, S.; Kim, S. C.; et al. Cytochrome P450 Monooxygenase-Mediated Metabolic Utilization of Benzo[a]Pyrene by *Aspergillus* Species. *MBio* **2019**, *10* (3), 1–15.

(16) Shin, J.; Kim, J. E.; Lee, Y. W.; Son, H. Fungal Cytochrome P450s and the P450 Complement (Cypome) of *Fusarium graminearum*. *Toxins (Basel).* **2018**, *10* (3), 76–91.

(17) Kavanagh, K. L.; Jörnvall, H.; Persson, B.; Oppermann, U. Medium- and Short-Chain Dehydrogenase/Reductase Gene and Protein Families: The SDR Superfamily: Functional and Structural Diversity within a Family of Metabolic and Regulatory Enzymes. *Cell. Mol. Life Sci.* **2008**, *65* (24), 3895–3906.

(18) Hussain, R.; Ahmed, M.; Khan, T. A.; Akhter, Y. Fungal P450 Monooxygenases - the Diversity in Catalysis and Their Promising Roles in Biocontrol Activity. *Appl. Microbiol. Biotechnol.* **2020**, *104* (3), 989–999.

(19) Samson, Robert A. ; Houbraken, Jos ; Thrane, Ulf ; Frisvad, Jens Christian ; Andersen, B. *Food and Indoor Fungi: Second Edition.*; CBS-KNAW Fungal Biodiversity Centre: Utrecht, The Netherlands., 2010.

(20) Jakob B. Hoof, Christina S. Nødvig, and U. H. M. Chapter 11 Genome Editing: CRISPR-Cas9. In *Fungal Genomics*; Humana Press: New York, NY., 2018; Vol. 1775, pp 119–132.

(21) Sun, X. Y.; Zhu, J. F.; Bao, L.; Hu, C. C.; Jin, C.; Harris, S. D.; Liu, H. W.; Li, S. J. PyrG Is Required for Maintaining Stable Cellular Uracil Level and Normal Sporulation Pattern under Excess Uracil Stress in *Aspergillus nidulans*. *Sci. China Life Sci.* **2013**, *56* (5), 467–475.

(22) De Mattos-Shipley, K. M. J.; Greco, C.; Heard, D. M.; Hough, G.; Mulholland, N. P.; Vincent, J. L.; Micklefield, J.; Simpson, T. J.; Willis, C. L.; Cox, R. J.; et al. The Cycloaspeptides: Uncovering a New Model for Methylated Nonribosomal Peptide Biosynthesis. *Chem. Sci.* **2018**, *9* (17), 4109–4117.

(23) Vit, A.; Misson, L.; Blankenfeldt, W.; Seebeck, F. P. Ergothioneine Biosynthetic Methyltransferase EgtD Reveals the Structural Basis of Aromatic Amino Acid Betaine Biosynthesis. *ChemBioChem* **2015**, *16* (1), 119–125.

(24) Li, H.; Gilchrist, C. L. M.; Phan, C. S.; Lacey, H. J.; Vuong, D.; Moggach, S. A.; Lacey, E.; Piggott, A. M.; Chooi, Y. H. Biosynthesis of a New Benzazepine Alkaloid Nanangelenin A from *Aspergillus nanangensis* Involves an Unusual L-Kynurenine-Incorporating NRPS Catalyzing Regioselective Lactamization. *J. Am. Chem. Soc.* **2020**, *142* (15), 7145–7152.

(25) Kjærbølling, I.; Mortensen, U. H.; Vesth, T.; Andersen, M. R. Strategies to Establish the Link between Biosynthetic Gene Clusters and Secondary Metabolites. *Fungal Genet. Biol.* **2019**, *130* (June), 107–121.

(26) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem* **2002**, *3* (7), 619–627.

Supporting information

Table of Contents

Experimental details.

All the experimental details refer to **Appendix 5**.

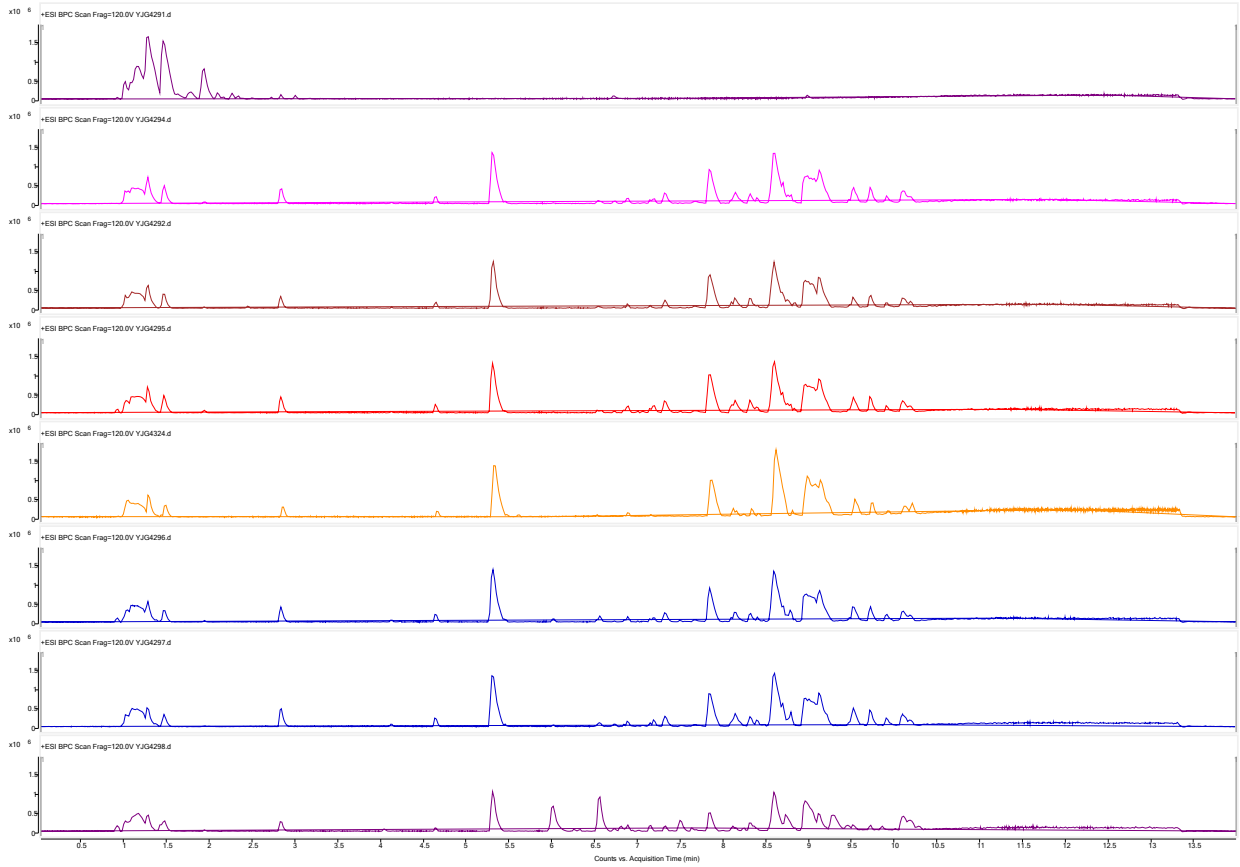
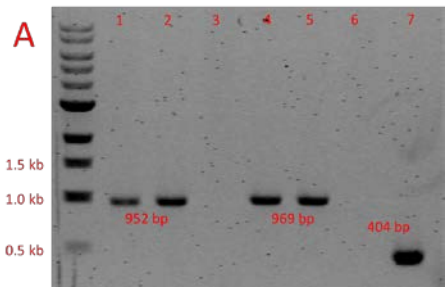


Figure S1. LC-MS profile of strains on CYA media with Uri+Ura.

(Samples from top to bottom: media blank, CAL005-009, CAL001, CAL000)



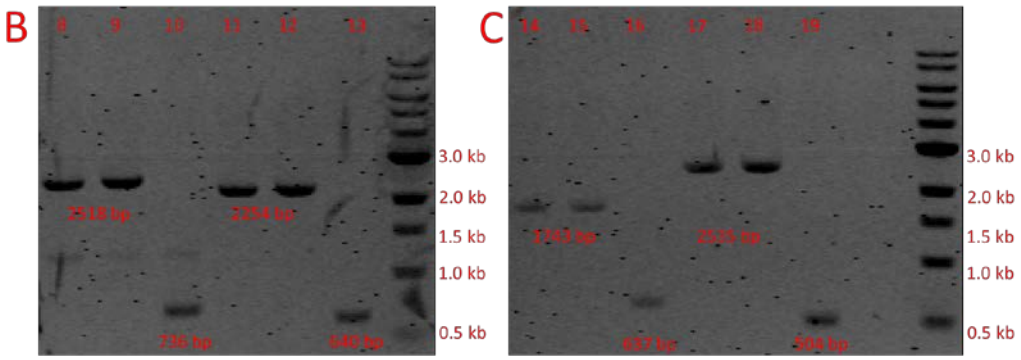


Figure S2. PCR verification of gene deletion mutants.

(Lanes 1, 4, 8, 11, 14 and 17 for WT. Lanes 2, 5, 9, 12, 15 and 18 for CAL001. **A**. CAL005 verification, primers for lanes 1-3 Acal.oxe252838.F1 and Acal.oxe252838.R1, Primers for lanes 4-6 Acal.oxe252838.F2 and Acal.oxe252838.R2, Primers for lane 7 Acal.oxe252838.F1 and Acal.oxe252838.R2; Primer binding pattern refer to PCR verification for *pyrG* deletion in appendix 5. **B**. CAL006 and CAL007 verification, primers for lanes 8-10 Acal.oxe252836.out.F and Acal.oxe252836.out.R, primers for lanes 11-13, Acal.oxe252843.out.F and Acal.oxe252843.out.R; **C**. CAL008 and CAL009 verification, primers for lanes 14-16 Acal.oxe82477.out.F and Acal.oxe82477.out.R, primers for lanes 17-19, Acal.oxe82484.out.F and Acal.oxe82484.out.R. Primers bind to 100-400 bp outside both ends of the targeting genes)

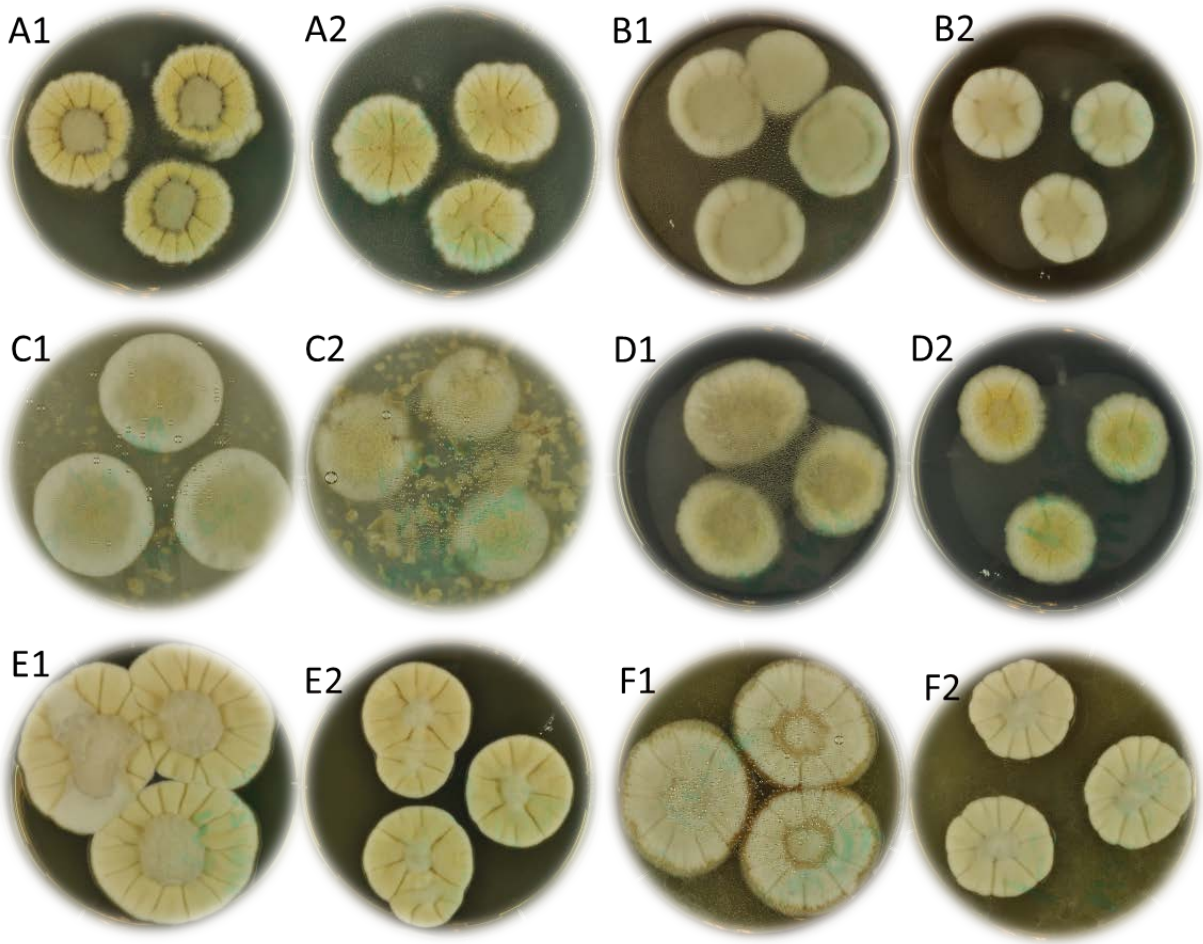


Figure S3. Pictures of CAL000 and CAL001 on six media for 11 days

(Media **A**-**F**: CYA, MEAox, OAT, PDA, YES and YPD, all supplemented with 10 mM Uri+Ura; **1**: CAL000; **2**: CAL001 except for **A2** was a picture of CAL005. Notes by Yaojie Guo: I took the picture from the bottom of CAL001 on CYA plate with Uri+Ura twice by mistake, and forgot to take a picture from the top view. However, all the mutants with *pyrG*- looked similar on the same media.)

Table S1. Primers used in this study

Primers/90-mers	Sequences (5' to 3')
Acal.oxe252838.F1	ACTCAGGTAACACGAGGCACT
Acal.oxe252838.R1	GCTATTCCAGCGTAAAACCTGCG
Acal.oxe252838.F2	GTTACCACGGAACCTGGGTG
Acal.oxe252838.R2	GTGGGCGGTTATGAGGTCTACA
Acal.oxe252838.PS1.R	ATCATCCACUAGGCTA TGCATCATCCGTGAATCGAAC
Acal.oxe252838.PS1.F	AGTGGATGAUATAG GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe252838.PS2.R	AGGTAAGCACUTGA TGCATCATCCGTGAATCGAAC
Acal.oxe252838.PS2.F	AGTGCTTACCUCTAGCA GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe252836.out.F	AGGGTGAGATATGAGGCACTGG
Acal.oxe252836.out.R	CTCCCACTTGAAGGATGACTTGTG
Acal.oxe252836.PS1.R	AGGCTTGCAGUGAGAT TGCATCATCCGTGAATCGAAC
Acal.oxe252836.PS1.F	ACTGCAAGCCUTTCC GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe252836.PS2.R	AACACGGATUGTGG TGCATCATCCGTGAATCGAAC
Acal.oxe252836.PS2.F	AATCCGTGTUAGGCGT GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe252843.out.F	TCAGGTGCCATTGGAAAGACGA
Acal.oxe252843.out.R	TCGGTGTCTTGGGGAACTCG
Acal.oxe252843.PS1.R	AAGACTACAUGCCACAAA TGCATCATCCGTGAATCGAAC
Acal.oxe252843.PS1.F	ATGTAGTCTUGC GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe252843.PS2.R	AGGCGCACCAAAUGT TGCATCATCCGTGAATCGAAC
Acal.oxe252843.PS2.F	ATTTGGTGCGCCUTCGAT GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe82477.out.F	TGGAAAACGACATCACCGAGGT
Acal.oxe82477.out.R	ATTCGGAGACTTGGGATGGCT
Acal.oxe82477.PS1.R	AGTATACTGGUCGGG TGCATCATCCGTGAATCGAAC
Acal.oxe82477.PS1.F	ACCAGTATACUCGCGC GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe82477.PS2.R	ACCCTCCAACUCCTCA TGCATCATCCGTGAATCGAAC
Acal.oxe82477.PS2.F	AGTTGGAGGGUAGGA GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe82484.out.F	TGTTGAGGAGTTGGAGGGTAGG
Acal.oxe82484.out.R	TCGTCTTTCCAATGGCACCTGA
Acal.oxe82484.PS1.R	ATTCCGCCTGGUGCTAAT TGCATCATCCGTGAATCGAAC
Acal.oxe82484.PS1.F	ACCAGGCGGAAUCT GTTTTAGAGCTAGAAATAGCAAG
Acal.oxe82484.PS2.R	AAAGACCATGTTCUACA TGCATCATCCGTGAATCGAAC
Acal.oxe82484.PS2.F	AGAACATGGTCTTUGCC GTTTTAGAGCTAGAAATAGCAAG

CSN438	GGGTTTAAUGATCACATAGATGCTCGGTTGACA
CSN790	GGTCTTAAUACCCTGAGAAGATAGATGTGAATGTG
Acal.oxe252838.90nt	TGAACTATGCGTGAAGATATACTAATAAGTCCTCAAGCGGTAGCAATTGGATACCT AGGACACGGCAGCGTCTACTTGTAATAGTCGTCA
Acal.oxe252836.90nt	GGAGCTCGCTATACAGCTGGGGATAAAGAGAATCACTCACCAGCCTTTCGATCGT CACTAGCCACCCTGCAGCGGAGCTGTAGTAACTCC
Acal.oxe82477.90nt	TGATCCACCGGCCTAAATATCTCGAGATAGGTACAAACAAGTATAGTCTGCATTCG TCTAGCATGTTCTTATTTGTTGTAGGAATATATA
Acal.oxe82484.90nt	AAGATCAGAAACCACGTTGAATGAAAGTGAAACCTAGTAATAACGTCCAACCAGC GTGGGAAATGTCGCAGTGAAGTTGGCGGAGCTCGA
Acal.oxe252843.90nt	AGACGCTTCCTTCACAGTGGCCGATATGATAGTGGCATTCTCTCCGTGTGTCTGCC GGGATCGCACGGAATCTACCTTGATACGTGGTTG

Table S2. Plasmids constructed in this study

Plasmids	Backbones	Primers for PCR amplification	Templates
pCas9-pyrg-oxe383BB	pFC330	CSN438/Acal.oxe252838.PS1.R Acal.oxe252838.PS1.F/Acal.oxe252838.PS2.R Acal.oxe252838.PS2.F/CSN790	pFC902
pCas9-pyrg-oxe363BB	pFC330	CSN438/Acal.oxe252836.PS1.R Acal.oxe252836.PS1.F/Acal.oxe252836.PS2.R Acal.oxe252836.PS2.F/CSN790	pFC902
pCas9-pyrg-oxe433BB	pFC330	CSN438/Acal.oxe252843.PS1.R Acal.oxe252843.PS1.F/Acal.oxe252843.PS2.R Acal.oxe252843.PS2.F/CSN790	pFC902
pCas9-pyrg-oxe773BB	pFC330	CSN438/Acal.oxe82477.PS1.R Acal.oxe82477.PS1.F/Acal.oxe82477.PS2.R Acal.oxe82477.PS2.F/CSN790	pFC902
pCas9-pyrg-oxe843BB	pFC330	CSN438/Acal.oxe82484.PS1.R Acal.oxe82484.PS1.F/Acal.oxe82484.PS2.R Acal.oxe82484.PS2.F/CSN790	pFC902

Table S3. Mutants constructed in this study

Strains	Genotype
CAL000	WT
CAL001	*pyrG-*, *ckuA*Δ
CAL005	*pyrG-*, *ckuA*Δ, jgi.p_Aspcalif1_252838Δ
CAL006	*pyrG-*, *ckuA*Δ, jgi.p_Aspcalif1_252836Δ
CAL007	*pyrG-*, *ckuA*Δ, jgi.p_Aspcalif1_252843Δ
CAL008	*pyrG-*, *ckuA*Δ, jgi.p_Aspcalif1_82477Δ
CAL009	*pyrG-*, *ckuA*Δ, jgi.p_Aspcalif1_82484Δ