Title
Genome mining on fungal polyketide natural products for discovery of novel enzymology

Author
HUNG, YIU-SUN

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Genome mining on fungal polyketide natural products for discovery of novel enzymology

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

by

Yiu-Sun Hung

2020

ABSTRACT OF THE DISSERTATION

Genome mining on fungal polyketide natural products for discovery of novel enzymology

by

Yiu-Sun Hung

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2020

Professor Yi Tang, Chair

Genome mining is a promising approach to elucidate the biosynthetic origins of natural products. Their structural complexity often originated from intriguing enzymatic reactions, in which understanding in biosynthesis could lead to discoveries of novel enzymology, which can be further utilized as biocatalysts for various biotechnological applications. Throughout my doctoral research, I have applied this methodology in biosynthetic studies for three fungal polyketide natural products of interest, where their synthetically challenging chemical moieties were resolved by enzyme catalysis in a highly regio- and stereo-specific manner. The newly characterized enzymes have diverse function, including isomerase, halogenase, and radical oxidase. Apart from their potential to be further developed into biocatalysts to conduct challenging chemical transformations, new drug analogs can also be developed from their bioactive parenting molecules. As a whole, this document not only aims to illustrate the theoretical and experimental details to arrive to these observations and conclusions, it can hopefully serve as motivation for future optimization and application of this promising concept of genome mining with extended applications, coupled with the rapid technological advancement.

The dissertation of Yiu-Sun Hung is approved.

Yvonne Y. Chen

Neil K. Garg

Junyoung O. Park

Yi Tang, Committee Chair

University of California, Los Angeles

2020

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Sections 1 and 2 contains material written by Hung, Y. S. from the following publication:

Hung, Y. S., Tang, Y. "Genome Mining of Fungi" *Comprehensive Natural Products III: Chemistry and Biology (Volume 6)* **2020**

Section 3 contains material written by Hung, Y. S. from the following publication:

Liu, N., Hung, Y. S., Gao, S-S., Hang, L., Zou, Y., Tang, Y. "Identification and Heterologous Production of a Benzoyl- Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid Biosynthetic Pathway." *Org. Lett.* **2017**

I would like to first acknowledge my doctoral advisor Professor Yi Tang for his conscientious guidance and encouragement throughout my doctoral research. He truly serves as a role model of scientific researcher with passion and professionalism. Also, I would like to acknowledge my committee members for their valuable advices and inspirations. I want to recognize the support from National Institute of Health for my doctoral research. (NIH Predoctoral Training T32 GM067555)

Moreover, I want to express my gratefulness for the assistances from my lab members and friends, with special thanks to Dr. John Billingsley and Dr. Nicholas Liu for their insightful comments on my dissertation. Last but not least, the endless encouragement and unconditional support from my parents and my wife Shang Gao help me to overcome the hardship during my graduate studies and continue to accomplish in life.

VITA

2012 – 2015	University of California, Los Angeles B.S. in Chemical Engineering *magna cum laude* Los Angeles, CA
2017 – 2019	NIH Biotech Training in Biomedical Sciences and Engineering Fellowship University of California, Los Angeles Los Angeles, CA

PRESENTATIONS

Hung, Y. S., Tang, Y., “Targeted genome mining for brefeldin A biosynthesis and resistance.” 2019 UCLA Biotechnology Symposium, 2019, UCLA. Presentation.

Hung, Y. S., Tang, Y., “Elucidating the biosynthesis of brefeldin A with CRIPSR-Cas9.” 2019 UCLA CBE 4th Year Graduate Students Symposium, 2019, UCLA. Presentation.

Hung, Y. S., Tang, Y., “Elucidation of the Biosynthesis of Wortmannilactones: NADH-fumarate reductase inhibitors as potential antiparasitics.” 2018 UCLA Biotechnology Symposium, 2018, UCLA. Presentation.

PUBLICATIONS

Hung, Y. S., Tang, Y. "Genome Mining of Fungi" *Comprehensive Natural Products III: Chemistry and Biology (Volume 6)* **2020**

Yee, D. A., Kakule, T., Cheng, W., Chen, M., Chong, C., Hai, Y., Hang, L., Hung, Y. S., Liu, N., Ohashi, M., Okorafor, I., Song, Y., Tang, M-C., Zhang, Z., Tang, Y. "Genome Mining of Alkaloidal Terpenoids from a Hybrid Terpene and Nonribosomal Peptide Biosynthetic Pathway." *J. Am. Chem. Soc.* **2020**

Zhao, Z-Y., Yin, Y., Hung, Y. S., Tang, Y. "Genome Mining Reveals Neurospora crassa can Produce the Salicylaldehyde Sordarial." *J. Nat. Prod.*, **2019**

Gao, S-S., Zhang, T., Garcia-Borras, M., Hung, Y. S., Billingsley, J. M., Houk, K. N., Hu, Y., Tang, Y. “Biosynthesis of Heptacyclic Duclauxins Requires Extensive Redox Modifications of the Phenalenone Aromatic Polyketide” *J. Am. Chem. Soc.* **2018**

Liu, N., Hung, Y. S., Gao, S-S., Hang, L., Zou, Y., Tang, Y. "Identification and Heterologous Production of a Benzoyl- Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid Biosynthetic Pathway." *Org. Lett.* **2017**

Hang, L., Tang, M-C., Harvey, C., Page, C., Li, J., Hung, Y. S., Liu, N., Hillenmeyer, M., Tang, Y. "Reversible Product Release and Recapture by Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain." *Angew. Chem. Int. Ed.* **2017**

Sato, M., Dander, J. E., Sato, C., Hung, Y. S., Gao, S-S, Tang, M-C, Hang, L., Winter, J. M., Garg, N. K., Watanabe, K., Tang, Y., "Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasynthases." *J. Am. Chem. Soc.* **2017**

Li, L. Yu, P., Tang, M-C., Zou, Y., Gao, S-S., Hung, Y. S., Zhao, M., Watanabe, K., Houk, K. N., Tang, Y. "Biochemical Characterization of a Eukaryotic Decalin-Forming Diels- Alderase." *J. Am. Chem. Soc.* **2016**

Darling, N. J.; Hung, Y. S.; Sharma, S.; Segura, T. "Controlling the kinetics of thiol-maleimide Michael-type addition gelation kinetics for the generation of homogenous poly(ethylene glycol) hydrogels." *Biomaterials* **2016**

1. INTRODUCTION

Genome mining is becoming a common approach for discovering new natural products in the post-genomic era. Rapid development of bioinformatic tools for comparative genomics, along with the innovation in genome editing, facilitate examination of previously uncharacterized biosynthetic gene clusters. Through deciphering the linkage between natural products and their corresponding genetic information, novel enzymology for the construction of these bioactive and often complex molecules can be revealed. Consequently, industrially applicable biocatalysts can be developed from these discoveries upon optimization, involving either solely enzymatic or chemo-enzymatic cascades. The utilization of biocatalysts overcomes key traditional synthetic bottlenecks by ensuring the regio- and stereo-selectivity, and bio-based materials can also perform optimally in mild and non-toxic environment. With the advancement in genome sequencing technology, fungal genome mining is an emerging methodology with the potential to unleash the discoveries of new enzymology embedded in fungal genomics. In this dissertation, the emerging advancements in biocatalysis application in the industry, as well as the methodology of fungal genome mining will be introduced. Subsequently, our engineering approach in platform development for heterologous expression, and its application in the three examples of biosynthetic studies of fungal polyketide natural products will be discussed, with an emphasis placed on the resulting novel enzymology discoveries. Lastly, the potential of these discoveries and development as industrially useful biocatalysts will be examined. Overall, this work serves as a record in utilizing cutting-edge technological development to unleash the enzymological potential inspired by fungal natural products, which can be further translated to various industrial applications as powerful biocatalysts. Ultimately, this work can motivate future researchers in the field to achieve breakthroughs in enzymology discoveries and fungal genome mining.

2. BACKGROUND OF FUNGAL GENOME MINING AND BIOCATALYSTS DEVELOPMENT

2.1 Recent approaches used in fungal genome mining

Fungi have been a rich source of medicinally and agriculturally relevant natural products for the past century.[1] The structural complexity of fungal natural products, evidenced by multiple ring systems and stereocenters, gives rise to their rich and diverse bioactivities.[2] However, with the declining rate of discovering new natural products from fungi in recent years, the question arises whether the pool of fungal natural products is nearing saturation, and consequently whether additional efforts devoted to discovery of new fungal natural products would be worthwhile. The recent advances in genomics have addressed such concerns and suggest that there is still immense potential for the discovery of new fungal natural products. For instance, with the use of a molecular approach for fungal species delimitation, the most recent estimate of the total number of fungal species is between 2.2 to 3.8 million, in contrast to less than 10% of the fungal species (about 120,000) that are currently characterized.[3] As a result, many of the uncharacterized fungal species, such as marine or endophytic fungi, have great potential for the discovery of new fungal natural products.

Another concern regarding the discovery of fungal natural products is the increasing difficulties associated with finding new chemical scaffolds and bioactivities. Prior to genomic approaches, fungal natural products were mostly discovered through traditional natural product isolation approaches, including screening for bioactive compounds from various fungal extract libraries and subsequently performing bioassay-guided isolation of the compounds of interest. However, such bioactivity-guided screenings have become less effective and more labor intensive, as the re-discovery rate of known natural products is high due to the limited bioassay screenings available.[4] Consequently, since the 1990s, there has been a gradual decline of interest from pharmaceutical companies in continual investment in such research programs, as there was

a perception that most of the medicinally and agriculturally relevant fungal natural products had already been discovered. Instead, these companies redirected their focus on screening for bioactive compounds from synthetic chemical libraries.[5] Around the same time, the initialization of the human genome sequencing project accelerated the advancement of genome sequencing technologies,[6] leading to drastic decrease in sequencing cost and enabled individual labs to sequence any organism of interest. This has led to an exponential increase in the number of fully sequenced fungal genomes. Efforts from the Joint Genome Institute and the Broad Institute have sled to sequencing and annotation of diverse fungi, of which the data is publicly accessible.[7-8] Since the first report of the complete genome sequence for the model fungus *Neurospora crassa* in 2003,[9] such abundant genomic information has brought the spotlight back to these fungi as rich producers of metabolites.

Genome based approaches have arguably brought forth a renaissance of natural product discovery and biosynthesis. Bioinformatic analysis of the fungal genome sequences suggested a tremendous untapped biosynthetic potential of filamentous fungi. Using anchor biosynthetic enzymes as beacons for discovery, such analyses have revealed a large number of biosynthetic gene clusters potentially encoding a variety of classes of metabolites, including polyketides (PKs), non-ribosomal peptides (NRPs) and terpenes.[10] The number of putative biosynthetic gene clusters well exceeds the number of natural products discovered to date from fungi, suggesting that as many as 97% of all gene clusters are cryptic under the standard laboratory culturing conditions.[11-12] Therefore, there is a need for technological development to systematically "decipher" these uncharacterized gene clusters presented in the fungal genome. These efforts are collectively referred to as "genome mining", a new paradigm of exploiting genomic information to guide discovery of new microbial natural products and uncovering the chemical secrets behind the genetic "dark matter."[13]

2.2 Discoveries of novel enzymology and the development of biocatalyst

Biocatalysis can be broadly defined as the use of biological sources, such as living cells or cellular components, to accelerate chemical transformations.[14-15] Humans have been utilizing biocatalysis for various application for centuries, from ethanol brewing to pharmaceutical production. The discovery of the enzyme, the biological unit that catalyzes these transformations, perhaps is one of most important breakthroughs in the history of biological research. Enzymatic catalysis provides benefits that include unparalleled selectivity, increased atom economy, and improved safety, which are the traditional challenges in chemical syntheses.[16] In addition to the eons of evolution that enzymes have experienced, scientists have altered the function of these natural catalysts with synthetic biology, such as directed evolution, to further expand the substrate scope and enhance the catalytic efficiency.[17-18] Moreover, scientists have observed that ordinary elements, such as radiation with light, can significant revise the enzymatic function by altering the energetics of their associated cofactors.[19] Combining two or more enzymatic steps into biocatalytic cascade sequences has also been shown to diversify the mode of catalysis, with the elimination of intermediate purification.[20-22] Nonetheless, these applications and optimizations are based on the discoveries of enzymology, which are encoded in the genome sequence of living organisms. With the advancement of genome sequencing technology, deciphering such genetic cryptography is more achievable than ever. The resulting methodology, referred to as genome mining, not only expands the chemical landscape but also the potential biological toolbox for biocatalysis.

One classical example to illustrate the usefulness of biocatalysts in drug development is the chemo-enzymatic synthesis simvastatin, a major cholesterol-lowering medication. By identifying the dedicated acyltransferase LovD in the natural fungal biosynthetic pathway, Tang and coworkers collaborated with Codexis to perform many iterations of *in vitro* evolution, creating libraries of LovD variants with improved activity for scale-up process. The approximately 1000-

fold improved enzyme and the new process pushed the reaction to completion at high substrate loading and minimized the amounts of acyl donor and of solvents for extraction and product separation. Simvastatin can therefore be produced efficiently from a renewable feedstock, thereby reducing the use of toxic and hazardous substances like tert-butyl dimethyl silane chloride, with an astonishing yield of 97% simvastatin **(Figure 1)**. Altogether, these examples illustrate the potential of utilizing the biosynthetic logic from nature and incorporating them into drug development, where genome mining can play in a major role for novel enzymology discoveries from microbes.

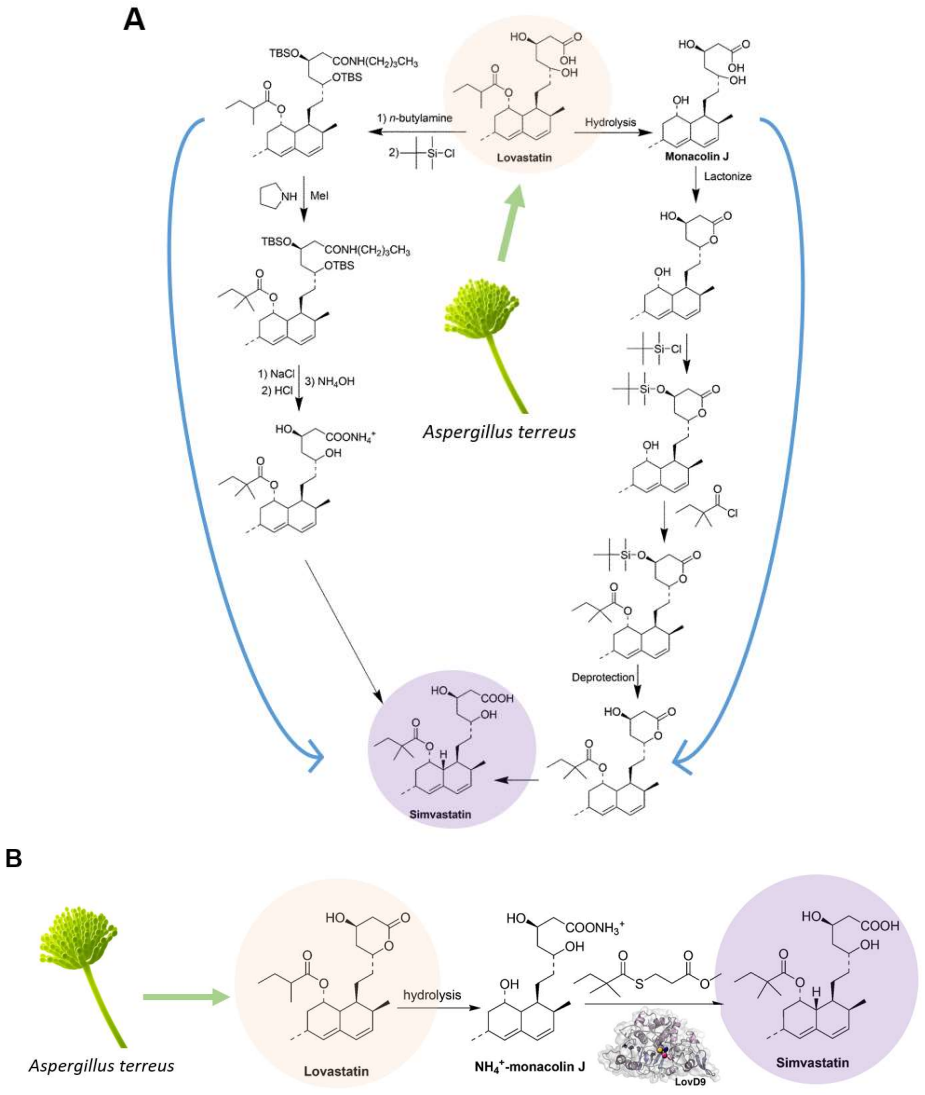


Figure 1. Synthetic methods for cholesterol-lowering agent simvastatin. (A) Traditional chemical syntheses. (B) Chemo-enzymatic synthesis with LovD mutants.

Overall, this chapter serves as an update and summary of the various strategies employed in fungal genome mining, as well as selected examples reported in the last ten years to highlight these strategies. With the incorporation of some of the most breakthrough technologies to date, such as CRISPR/Cas9 gene editing and low-cost DNA synthesis, many of the previous challenges in genetic manipulations of natural hosts and heterologous reconstitution in model hosts can be overcome. This chapter aims to illustrate fungal genome mining as a promising technique to uncover the full biosynthetic potential of filamentous fungi for drug discovery in the next decade and beyond.

2.2.1 Bioinformatic analysis

Next-generation sequencing has provided an affordable and high throughput technology for complete genome sequencing of microbial species, including many fungal species in the last 10 years.[23] As a result, there is a rapid increase in the number of fungal genome sequences available in the database, as exemplified by the recent completion of the "1000 Fungal Genomes Project" initiated by the Joint Genome Institute (JGI).[24] With this growing amount of fungal genome sequences, development of high-throughput analytical tools to process the genomic data are needed for genome mining purposes. A number of useful computational algorithms were developed in recent years, many of which are tailored to scan through microbial genome sequences and predict biosynthetic gene clusters. Structural prediction, while still very preliminary, can be done by comparing to clusters with known natural products. The biosynthetic knowledge and logic of the anchoring enzymes in the gene clusters, such as polyketide synthases (PKSs), non-ribosomal peptide synthetases (NRPSs) and terpene synthases (TSs), etc can also guide and predict the structure of the natural product.[25] The significance of these computational tools is to rapidly provide preliminary information to researchers, enabling *in silico* dereplication of biosynthetic clusters by establishing similarity and synteny between clusters from different organisms. Such analysis has become invaluable for researchers to decide whether a particular

biosynthetic gene cluster is worth experimental investigation, and if so, the type of genome mining approach that is most suitable to use.

Among all the commonly used bioinformatic tools to date, the central algorithms all revolve around BLAST (Basic Local Alignment Search Tool), the pioneering algorithms for DNA or protein sequence comparison based on local similarity and statistical significance.[23] BLAST can be used to infer functional and evolutionary relationships between sequences, as well as to help identify members of gene families. Many useful applications are built upon BLAST, including the commonly used "CD (Conserved Domain)-Search" for locating conserved domains in the query sequence based on databases of domains compiled from previously characterized genes. CD-Search is particularly useful for predicting domain architecture for magasynthases such as PKSs or NRPSs.[26] Other algorithms that can perform *ab initio* gene prediction for eukaryotic genomes are useful for analysis of fungal genomes. One example is AUGUSTUS,[27] which is a Hidden Markov Model (HMM) based algorithm to model the intron length distribution in order to accurately predict the size and location of introns from the query sequences. AUGUSTUS is also capable of predicting alternative splicing, as well as the location of 5'UTR (untranslated region) and 3'UTR, which are critical pieces of information for approaches such as heterologous expression and transcription factor overexpression. The web-based program 2nd Find uses AUGUSTUS to search for secondary metabolite biosynthetic gene cluster from bacterial or fungal genome sequences. In addition, 2ndFind also uses Pfam (Protein Family)[28] database to predict and assign putative protein identity to query sequence. This makes 2ndFind a powerful tool that can display the putative functions of genes in a given cluster. Such information is useful in determining the putative boundary of a biosynthetic gene cluster for fungal genome mining purposes.

In addition to the algorithms that analyze relatively short DNA sequences or individual clusters, global genome analysis programs such as antiSMASH (antibiotics & Secondary Metabolite Analysis Shell) are widely used for rapid identification and annotation of biosynthetic

gene clusters from both bacterial and eukaryotic genomes.[29] There are also computational algorithms that are more specific than antiSMASH, such as SMURF (Secondary Metabolite Unique Regions Finder)[30] that specializes in annotation of the gene clusters containing PKSs and NRPSs in particular. PRISM (Prediction Informatics for Secondary Metabolomes)[31] primarily focuses on the prediction of chemical structures for the natural products associated with the biosynthetic pathways. The most recent version of antiSMASH (antiSMASH 4.0) incorporates many of these algorithms in its software, including CASSIS (Cluster Assignment by Islands of Sites),[32] which enables the prediction of the boundary of the biosynthetic gene cluster and the likelihood of genes surrounding the anchoring megasynthase to participate in the biosynthetic pathway. Cluster boundary can be predicted by comparing sequence similarity of promoters for each gene in the vicinity of the core biosynthetic enzyme. This feature can lead to the identification genes that are not typically predicted to be biosynthetic, including those that are routinely categorized to be associated with primary metabolism. Identification and characterization of such predicted genes could lead to novel enzymology or potential resistance genes towards the encoded natural products. The advancements in computational prediction algorithms for application to natural product discovery are summarized in a review by *Blin et al.* A list of web tools and databases for mining secondary metabolite biosynthetic gene clusters discussed in this review is adopted from the review and listed in **Table 1**.[33]

Tool	Functions	URL	Reference
antiSMASH 4	Genome mining BGC analysis Domain analysis	http://antismash.secondarymetabolites.org	29
antiSMASH database	BGC database	http://antismash-db.secondarymetabolites.org	29
CASSIS	BGC boundary prediction	https://sbi.hki-jena.de/cassis/cassis.php	32
MIBiG	BGC database Reference data	http://mibig.secondarymetabolites.org	34
PRISM 3	Genome mining BGC analysis Domain analysis	http://magarveylab.ca/prism	31
SMURF	Genome mining	http://www.jcvi.org/smurf	30

Table 1. List of selected web tools and databases for mining secondary metabolite biosynthetic gene clusters.[33]

To facilitate a systematic deposition and retrieval of biosynthetic gene cluster (BGC) data, a new data standard named “Minimum Information about a Biosynthetic Gene cluster (MIBiG)”[34] was developed, which is based on the “Minimum Information about any Sequence (MIxS)”[35] initiated by the Genomic Standards Consortium (GSC).[36] The compiled experimental information including gene knockout phenotypes and verified gene functions for specific BGCs are included in the MIBiG database, along with class-specific checklists for gene clusters encoding pathways that produce PKs, NRPs, terpenes, alkaloids, etc. For instance, the substrate specificities of acyltransferase (AT) domains and starter units for polyketide BGCs, adenylation domain substrate specificities for non-ribosomal peptides BGCs, as well as releasing/cyclization mechanisms are included in this database. Currently, the MIBiG database includes more than 1700 experimentally characterized gene clusters (about 250 of them are fungal gene clusters), which represents a significant advancement in coverage of the field and exceeds previously established platforms, such as ClusterMine360[37] (288 BGCs) and DoBISCUIT[38] (103 BGCs). With voluntary user input, the growing database of MIBiG will provide the entire field with a centralized and readily accessible platform for comparative analysis and *in silico* dereplication of BGCs of interest.

2.2.2 Epigenetics modulation and microbial cocultivation

Chromosomal location and histone modification are known to be important factors for gene transcription for a variety of organisms. Early genome sequencing in *Aspergillus* illustrated that secondary metabolite gene clusters tend to localize sub-telomerically, a region where chromatin modifiers impact transcription of these clustered genes the most.[39] Moreover, post-translational modifications such as acetylation and methylation of histones were shown to be important in modulating the expression of biosynthetic genes in secondary metabolism. Therefore, altering the level of these modification which can lead to changes in packing of the heterochromatin, can potentially activate cryptic fungal gene clusters. In an overly simplified scheme, hyperacetylation and methylation on H3K4 (histone 3 lysine residue 4) are associated with gene activation, while

hypoacetylation and methylation on H3K9 (histone 3 lysine residue 9) are associated with gene silencing.[40] Previous studies showed that deletion of *hdaA*, encoding an *Aspergillus nidulans* histone deacetylase (HDAC), causes transcriptional activation of two telomere-proximal gene clusters, but not of a telomere-distal cluster.[41] Deletion of three genes that are responsible for histone modification, *hdaA*, *hepA* (heterochromatin protein 1) and *clrD* (H3K9 methyltransferase) resulted in the increased production of sterigmatocystin and penicillin.[41-43] Treatment of fungi with HDAC inhibitors, such as suberoylanilide hydroxamic acid (SAHA), resulted in overproduction of several other metabolites,[44-47] suggesting a conserved mechanism of HDAC repression of secondary-metabolite gene clusters. On the other hand, deletion of *cclA*, which encodes for an enzyme that is involved in methylation of H3K4 in *A. nidulans*, led to the production of antimicrobial monodictyphenone and anti-osteoporosis F9775.[37] Although these bioactive compounds were reported in other fungi before this finding, with monodictyphenone and F9775 isolated from *Monodictys putredinis*[48] and *Paecilomyces carneus*[49] respectively, this work revealed the potential of the model *A. nidulans* to produce these natural products. Consequently, the biosynthesis of these compounds were elucidated using the *ΔcclA* mutant of *A. nidulans*,[49] the key finding being the precursor of F9775 is orsellinic acid synthesized by a non-reducing polyketide synthase (NRPKS). Epigenetic modulation has been used in other organisms, including inactivation of a histone 3 deacetylase that led to pleiotropic activation and overexpression of more than 75% of the biosynthetic genes in the mushroom-endophytic fungus *Calcarisporium arbuscula;*[46] and the production of pestaloficiols and ficiolides from plant endophytic fungus *Pestalotiopsis fici*.[50]

In addition to epigenetic modulation, untargeted exploration of the metabolic potential of filamentous fungi through cocultivation with microbes has also been successful. The co-culturing concept originates from the observation that many of these fungal natural products exhibit antimicrobial activities. Physical contacts and/or environmental signals from other microbes in the vicinity of fungi may be necessary to activate the production of the natural products for self-

defense purposes. Examples including the well-studied fungal-cyanobacterial symbiosis for lichen formation[51] and bacterial-fungal cytosolic endosymbiosis for mycotoxin production.[52] Recreating such multispecies interactions and communications in laboratory culturing conditions may therefore elicit the expression of previously silent gene clusters in these fungi. For example, *A. nidulans* was co-cultivated with different actinomyces, and the resulting expression of various biosynthetic genes were monitored through microarray experiments.[53] The co-cultivation led to the production of orsellinic acid and its derivatives due to the upregulation of these gene clusters. The much sought-after PKS responsible for the biosynthesis of these compounds was revealed through knockout studies,[53] and this PKS was surprisingly found to be widespread in many other fungal species, including mycobionts of lichens. Further characterization of the fungal-bacterial interactions including electron microscopy confirmed that intimate physical interaction between the fungi and bacteria are necessary, instead of the previous hypothesis that diffusible signals are sufficient. In this case, production of new metabolites through co-cultivation is due to bacterial-triggered Saga/Ada mediated histone acetylation.[54] Chromatin immunoprecipitation showed that the Saga/Ada-dependent increase of acetylation H3K9 and H3K14 occurred during interaction between fungus and bacterium.

Other examples of fungal-bacterial coculture mediated natural product discovery include the production of pestalone, a potent antibiotic against methicillin-resistant *Staphylococcus aureus* (MRSA), resulting from the co-culture of a marine-derived gram-negative bacterium of the genus *Thalassopia sp.* (CNJ-328) and the marine fungus *Pestalotia*.[55] In addition, co-cultivation of the same bacterium with the fungus *Libertella sp*. led to the production of cyctotoxic libertellenones by the fungus. In this case, such diterpenoids were neither produced in a *Libertella* monoculture nor by adding supernatant or extract of the bacterial culture, which suggested again that physical contacts between the microbes are necessary.[56] *A. fumigatus* was also shown to be involved in a microbial interaction from coal mine drainages , and co-cultures of a *Sphingomonas*

strain with *A. fumigatus* led to the emergence of a new diketopiperazine glionitrin A, which has antibiotic activity against both MRSA as well as cytotoxic activities against four human cancer cell lines.[57] Apart from fungal-bacterial co-culturing, fungus-fungus cocultivation has also been used for the discovery of new natural products, including acremostatins isolated from a mixed culture of *Acremonium sp.* and *Mycogone rosea*,[58] aspergicin discovered from a culture of two *Aspergillus* species,[59] and cyclo-(L-leucyl-*trans*-4-hydroxy-L-prolyl-D-leucyl-*trans*-4-hydroxy-L-proline) produced in the co-culture broth of two mangrove fungi *Phomopsis* sp. K38 and *Alternaria* sp. E33.[60] These examples from fungal-microbial cocultivation illustrate the potential of using non-genetic methods for the discovery of new natural products.

2.2.3 Fungal genome editing using CRISPR/Cas9 system

CRISPR/Cas9 (Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9) is a genome editing breakthrough developed from the bacteria immune system in *Streptococcus pyogenes*.[61] It includes two major components: Cas9 is an endonuclease that catalyzes a double stranded DNA break, while single-guided RNA (sgRNA) acts as the guidance for Cas9 based on sequence complementarity to the targeted DNA. Genome editing can be performed through the formation of the ribonucleoprotein (RNP) complex between sgRNA and Cas9, with higher efficiency compared with traditional methods such as homologous recombination alone. The CRISPR system has been optimized to perform genome editing in various organisms including *Saccharomyces cerevisiae* (yeast),[62] mammalian cells,[63] and plants.[64] Inspired by the success of CRISPR editing in yeast, a CRISPR system has been developed for *Aspergilli,* with the first example of genome editing in *Aspergillus nidulans*.[65] Since then, the CRISPR system has been applied to a wide range of filamentous fungi to test its utility in fungal genome editing, including in *Neurospora*,[66] *Trichoderma*,[67] and *Penicillium*.[68] In some cases, the CRISPR system enabled genome editing in fungal species with difficult or intractable genetic systems. This makes CRISPR a valuable tool for fungal natural product genome mining.

An example of successful genome editing using CRISPR/Cas9 whereas there were no previous reports of genetic modification is with the fungus *Talaromyces atroroseus*.[69] Gene knockout was successfully performed on this strain with CRISPR, which led to the identification of the gene responsible for the production of ZG-1494α, as well as its derivative talaroconvolutin A, an inhibitor of platelet-activating factor (PAF) acetyltransferase.[70] By co-transforming the CRIPSR plasmid with constitutive expression of CRIPSR-Cas9 along with the homology plasmid with antibiotic selection to the target fungus, this study successfully confirmed the involvement of a PKS-NRPS enzyme in synthesizing ZG-1494α and derivatives, and enabled combinatorial biosynthesis in in *A. nidulans* as a heterologous host.[71] From a genetic tool development perspective, this example illustrates the versatility of the CRISPR system developed for *Aspergillus* can be used without significant modification for genome editing in other less-well characterized fungal species.

CRISPR/Cas9 system can be further modified to modulate gene expression through deactivation of the DNA cleavage activity of Cas9 and further tethering to a transcription activator or repressor. This so called CRISPR/deactivated Cas9 (dCas9) system has not been studied in filamentous fungus but has been applied to yeast for controlling expression of gene of interest.[72] Studies have shown successful transcriptional regulation upon the optimization of few parameters, such as the binding location of the sgRNA and regulatory elements fused to dCas9. This technique can be particularly useful for studying the biosynthesis of natural products produced by filamentous fungi as a genome mining tool. As mentioned previously, many fungal biosynthetic gene clusters do not have pathway-specific transcription factors, thereby precluding the use of transcription factor overexpression for pathway activation. However, the CRISPR/dCas9 system can be engineered to bypass this obstacle and activate the gene of interest directly. Indeed, CRISPR/dCas9 activation can be applied to multiple genes at the same time, as shown in other organisms.[73] Lastly, by tuning the parameters for transcription regulation

using the CRISPR/dCas9 system, one can potentially modulate the expression of individual pathway genes at different levels, which can be an effective method to identify biosynthetic intermediates or perform precursor-directed biosynthesis. Therefore, CRISPR/Cas9 is a powerful tool for genome editing in filamentous fungi and has enormous potential in fungal genome mining, most of which are waiting to be realized.

2.2.4 Heterologous expression in Saccharomyces cerevisiae

The tools discussed above are all developed for directly working with the fungi of interest, and to activate the expression of a biosynthetic pathways in the original host. An equally powerful method of genome mining is the use of genetically well-established model organisms as heterologous production hosts. This is especially applicable for working with clusters from genetically intractable or unculturable fungi, from environmental DNA, or studying fungal hosts from which the cluster of interest is not available. Among the heterologous hosts available, *Saccharomyces cerevisiae* (Baker's yeast) is a widely used host for genome mining of fungal natural products. Yeast is arguably the most well-characterized and the most well-understood fungus. It is highly efficient in homologous recombination which enables rapid assembly of heterologous pathways from oligonucleotides and PCR fragments. Due to the ease of genetic manipulation and a long history of strain engineering, there are many engineered yeast strains available that are tailored for the production of fungal natural products. One example is *S. cerevisiae* BJ5464-NpgA, which is a derivative of yeast strain BJ5464 that contains a chromosomally integrated copy of the phosphopantetheinyl transferase (PPase) gene *npgA* from *A. nidulans*.[74] PPase converts the inactive *apo*-thiolation domains in PKSs and NRPSs to the active *holo*-forms by attaching the phosphopantetheine to the serine residue.[75] The strain BJ5464 itself has been optimized for expression of high molecular weight proteins such as PKS and NRPS through deletion of the vacuolar proteases PEP4 and PRB1. Fungal PKS and NRPS mega-synthases in excess of 250 kDa can be expressed intact at mg/L levels in the BJ5464 strain.

Yeast also has a collection of constitutive and inducible promoters for expression of foreign genes. A frequently used promoter in the BJ5464 expression host is the strong and inducible promoter ADH2.[76] The ADH2 promoter is repressed in the presence of glucose, and is active upon depletion of glucose in the media. The use of this promoter separates yeast growth and metabolite production and minimizes toxic effects of foreign metabolites on yeast growth. A suite of episomal vectors, including both high copy (2 micron) and single copy (centromeric origins) carrying the ADH2 promoter has been developed for use of pathway reconstitution and mining in yeast.[76] A further useful derivative of BJ5464-NpgA is the strain RC01, which contains a genome integrated copy of the cytochrome P450 reductase (CPR) from *A. terreus*.[77] Integration of the CPR allows heterologous expression of P450 enzymes from biosynthetic pathways, which is common in many fungal BGCs.

Among the numerous successful examples of using yeast heterologous expression in fungal natural product reconstitution and mining, an early example is the complete reconstitution of the PKS enzymes involved in lovastatin biosynthesis.[78-79] The use of an engineered *Saccharomyces cerevisiae* strain BJ5464-NpgA led to the successful purification of lovastatin nonaketide synthase LovB (335 kDa) at a level of 4 mg/L. The chromosomally expressed NpgA completely modified the ACP domain. LovB expressed from yeast is fully functional and can execute more than 35 catalytic steps accurately to produce the first intermediate dihydromonacolin L. The same yeast system has been used for the reconstitution of many fungal biosynthetic pathways, including several from the resorcylic acid lactone family,[80] multimodular NRPS enzymes[81] and indole diterpenes.[77] Multiple cryptic biosynthetic pathways have also been introduced into the yeast strain and led to the successful mining of new natural products.[82-83]

The recently reported HEx (Heterologous Expression) platform was developed to integrate high-throughput assembly of gene clusters, yeast cultivation and untargeted metabolomics. This platform relies entirely on synthetic genes, which makes it independent of the original fungal

species. A success rate of producing new metabolites was observed in ~50% of gene clusters assembled and assayed in the HEx pipeline.[84] One important component of this method is a panel of strong and inducible promoters to drive synchronized expression of the genes in a pathway. The promoters include the yeast ADH2 promoter, and promoters that have the same inducible features and strengths from yeast and other *Saccharomyces* species. The need to find these similar but sequentially dissimilar promoters is to enable multi-fragment, multi-gene DNA assembly in yeast. Using a panel of four promoters, a four-gene cluster each with its own promoter can be assembled in a single yeast recombination step. Other engineered components of the HEx platform include a faster growing yeast strain and a cost-effective procedure that allows direct sequence verification from yeast.

Some pathways can also be directly reconstituted in the bacterial host *E. coli.* Successful examples of expressing fungal biosynthetic genes in *E. coli* are abundant. Both full pathway reconstitution and expression of partial pathways for bioconversion have been reported. The expression of 6-methylsalicylic acid synthase (6-MAS) led to the production of 6-methylsalicylic acid in titers of 75mg/L.[85] The complete reconstitution of the antitumor agent terrequinone A biosynthetic pathway was accomplished by expression of five enzymes(TdiA-TdiE) from *E. coli*.[86] This led to the discovery of novel enzymatic functions, including the iterative prenylation by the prenyltransferase (TidB), as well as the “head-to-tail” dimerization mechanism of tethered α-ketoacyl-NRPS intermediates by the NRPS (TdiA) to set up the subsequent tetra-substituted benzoquinone. Certain NRPKSs such as PKS4 and PKS13 from the bikaverin[87] and zearalenone[88] biosynthetic pathways, respectively, can be highly expressed in *E. coli*, leading to the formation of the correct PKS products. Using the fungal NRPKS as scaffold generating enzymes, and combining it with bacterial aromatic polyketide components, bacterial aromatic polyketides were produced in *E. coli* for the first time.[89] For most pathways involving mega-synthases however, the translation and protein folding machinery of *E. coli* simply cannot handle

these large enzymes, leading to truncated or insoluble enzymes. For these reasons, large PKS or NRPS containing pathways are usually directly expressed in yeast instead of *E. coli.*

2.2.5 Heterologous expression in filamentous fungi

Model filamentous fungi are also important heterologous hosts for reconstitution of fungal BGCs and genome mining. A principle reason is their ability to recognize and splice fungal introns. As a result, genomic DNA of genes of interest can be cloned directly into fungal expression plasmids and expressed in fungal hosts. In contrast, the intron-free cDNA (complementary DNA) are needed for yeast expression since yeast cannot splice the fungal intron despite the presence of an intact spliceosome machinery. Obtaining the intron-free coding region can be accomplished if the fungal gene is transcribed in the original host, or through manual assembly of exon pieces if the transcript is not available. The latter approach relies on precise intron prediction software, and despite significant advances in predictive ability, it is prone to mis-annotation of introns. For this reason heterologous reconstitution in a more closely related filamentous fungus is much more feasible. Additional advantages of using filamentous fungi over yeast include a more similar intracellular condition for protein expression, the presence of post-translational modification systems if required, presence of accessory enzymes for protein function, etc.

However, as opposed to the well-established genomic tools that are available for yeast, generating an auxotroph of a filamentous fungus for using selectable markers is challenging. The knockout of homologs of KU70 in filamentous fungi (*nkuA*[90] and *kusA*[91] in *A. nidulans* and *A. niger*, respectively), which is responsible for non-homologous end jointing, has led to a much higher knockout efficiency by homologous recombination. Auxotrophic strains can be generated more readily in these *ku70* homologs deficient strains by knocking out the selective markers via homologous recombination. This has led to the development of several fungal expression platforms that are discussed subsequently.

One of the early examples of heterologous expression in filamentous fungi is the reconstitution of the BGC for spirotryprostatins in *A. niger*, a tryptophan-containing nonribosomal peptide isolated from *Aspergillus fumigatus*. The reconstitution work revealed a biochemical crosstalk between the related fumitremorgin biosynthetic pathway and unrelated fumiquinazoline biosynthetic pathway.[92] Another example of using the *A. nidulans* expression platform is reconstitution of the leporin B pathway from *Aspergillus flavus* in *A. nidulans*.[93] This work revealed one of the biosynthetic enzyme in the pathway, LepI, a *S*-adenosyl-L-methionine (SAM)-dependent enzyme that can catalyze stereoselective dehydration as well as three pericyclic transformations. This finding represents the first report of enzyme catalyzed retro-Claisen rearrangement and illustrates the evolution of a methyltransferase enzyme into one that catalyzes pericyclic reactions. A recent genome mining example is the discovery of natural products produced by a unusual PKS-C (C represents the condensation domain from NRPS enzymes) hybrid from *Talaromyces wortmanii*.[94] Heterologous expression of the gene encoding PKS-C along with other tailoring genes in *A. nidulans* led to production of wortmanamides, which are new compounds synthesized from reduced long-chain polyketides amidated with a specific ω-amino acid 5- aminopentanoic acid (5PA). These compounds are similar to known GPCR antagonists isolated from bacteria.

Aspergillus oryzae is another common heterologous host used for pathway reconstitution and genome mining. A major advantage of this host is that it possesses a relatively clean endogenous metabolic background that allow for straightforward identification and purification of natural products synthesized from exogenous pathways. Representative examples of using *A. oryzae* as a heterologous host include the reconstitution and characterization of the biosynthesis of the indole diterpene paxilline through a stepwise introduction of the six genes in the biosynthetic pathway.[95] Abe and coworkers have used this host extensively in the heterologous production of polyketide-terpene meroterpenoids such as andrastin A,[96] terretonin,[97] and anditomin.[98] In

studying these pathways, the authors found that *A. oryzea* host can consistently afford titers of metabolites at 30-50 mg/L, and nearly all introduced enzymes can be functionally expressed. This host enabled Abe and coworkers to fully map out the modular nature of related biosynthetic pathways. Also using *A. oryzae*, the biosynthetic pathway of meroterpenoid xenovulene A was elucidated and led to the discovery of three previously unrecognized classes of biosynthetic enzymes including a potential hetero Diels-Alderase.[99]

Fusarium heterosporum was developed to be the first non-*Aspergillus* heterologous host for the production of the antituberculosis agent pyrrolocin A and its derivative with an exceptionally high yield of ~ 1 g/L.[100] Such high production of the compound of interest is achieved by utilizing the transcriptional regulatory elements from the natively high-producing biosynthetic pathway of equisetin, which can achieve >2 g/L titer in a controllable manner.[101] Although this *Fusarium*-based heterologous system can achieve high yield compared with other *Aspergillus* systems, the current system is limited with the expression of only 4 genes in maximum (2 genes under each bi-directional *eqxS* promoter for 2 selective markers), where continual development of this system would be beneficial for the investigation of larger size BGCs.

However, a limitation for the current fungal heterologous expression system is that these hosts only allow the investigation of at most one or two gene clusters at a time, while the construction of the expression vectors can be labor- and time-intensive. This led to the development of a high-throughput heterologous platform termed FAC-MS (fungal artificial chromosomes and metabolomic scoring) shown in **Figure 2**.[102] The advantage of this platform compared with the approaches described previously is the use of the fungal artificial chromosomes to directly clone fragments of the genomic DNA from the fungal host of interest. Each of these FACs allows the capture of the DNA insert up to 300 kb in size, which is more than

sufficient to contain the full-length gene cluster including the corresponding accessory and regulatory elements. The concept of using FAC for fungal heterologous expression was first reported in the discovery of terezine D in *Aspergillus terreus.*[103] In FAC-MS, combining the use of FAC-based cloning with metabolomic scoring allows the assignment of scores to compounds

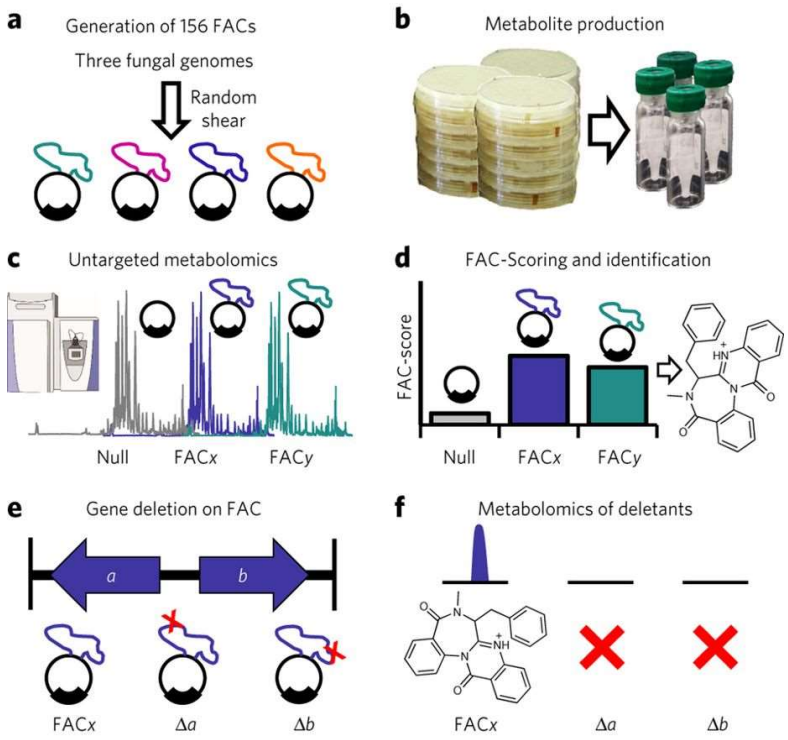


Figure 2. Schematics of the FAC-MS platform.[102] (a) Fungal artificial chromosomes (FACs) are generated by cloning randomly sheared fungal genomic DNA into fungal-bacterial shuttle vectors. (b) Metabolites are obtained from the extract of *A. nidulans* heterologous expression of the FACs. (c) Metabolites are subjected to untargeted mass spectrometry analysis. (d) FAC-score is assigned to each metabolite detected in order to separate it from the host metabolic background. (e) Predicted core genes for the targeted metabolites are deleted. (f) The metabolites are subjected for detection from the fungal extract with the newly constructed deletant FACs expressed to confirm the elimination of the predicted metabolites. [Figure reprinted with permission from Clevenger et al, *Nature Chemical Biology* 13, 896, (**2017**).]

observed from LC-MS analysis based on its appearance in different FAC strains. High scores are assigned to compounds that are detected from only one or few FAC strains, while the compounds that appears in many FAC strains will get lower scores and regarded as background metabolites from the heterologous host. The establishment of this scoring system is necessary because there are as many as 5000 signals observed from previous FAC experiments. FAC-MS was applied to three *Aspergilli* (*A. terreus*, *A. wentii*, and *A. aculeatus*), and a library of 56 FACs harboring uncharacterized fungal BGCs were screened, leading to the discovery of 15 new metabolites, including the macrolactone valactamide A. After sequencing the FACs, biosynthetic gene clusters were assigned to the new metabolites. While preliminary success of this approach shows FAC-MS is a promising high-throughput alternative to methods that focus on one cluster at a time, it remains unclear how silent BGCs can become transcriptionally active when cloned into FAC and introduced to the new host without complete refactoring of the transcriptional elements.

2.2.6 Resistance-gene-directed genome mining

While the approaches described above have been successful in activating cryptic BGCs and identifying the associated natural product, the true biological activities of the produced compound are typically unknown. Compared to more classical natural product discovery, in which a phenotypical assay guides the purification of a compound, the genomic approaches are not activity-guided. Given the large number of BGCs available, it is essential that genome driven discovery of natural products to be prioritized by biological activity. This begs the question: how can we predict the activity of an NP based on BGC sequence? The answer to this question can unlock the true potential of the tens of thousands of unexplored BGCs. Recent examples of resistance-guided genome mining offer an approach to connect genome data to biological activity. It has been known for some time now that in order for the producing organism not to be harmed by the natural product it produces, self-resistance mechanisms must be in place when a toxic natural product is produced.[104] Several of these self-defense mechanisms are known, including

the use of efflux pumps to actively transport the natural products to the extracellular space, the use of antidotal proteins to bind and detoxify the natural product, the use of specific enzymes to modify the target of the natural product as a means to evade inhibition, and the use of a functionally equivalent resistance enzyme to compensate for the inhibition of the housekeeping enzyme by natural product. This last mechanism is especially intriguing as the resistance enzyme can be a second and homologous copy to the housekeeping enzyme that is targeted by the natural product. This second copy can perform the same catalytic function but is sufficiently mutated to be insensitive to the natural product. The resistance gene is also typically co-clustered with the BGC and is co-transcribed with the rest of the cluster. The self-resistance gene can therefore serve as a predictive window to find new natural products of desired biological activity in the following work flow: 1) after identifying a desired target enzyme that is also present in microorganisms, one can search through genome databases for BGCs carrying duplicate copies of the target that is located close to a biosynthetic anchoring enzyme; 2) different synthetic biology approaches can be applied to produce the natural product encoded in the cluster; 3) the natural product is isolated and the structure is elucidated using NMR spectroscopy; and 4) inhibition of the target, and insensitivity towards the resistance gene are validated biochemically or genetically.

Indeed, the resistance gene hypothesis is supported by several prominent fungal natural products pathways. For instance, there exists a second copy of the HMG-CoA (3-hydroxy-3-methyl-glutaryl-coenzyme A) reductase within the lovastatin gene cluster in *A. terreus* in addition to the housekeeping copy.[105] This is consistent with lovastatin targeting the HMG-CoA reductase in the cholesterol biosynthetic pathway. Another example is the immunosuppressant mycophenolic acid, of which its BGC contains an extra copy of inosine-5'-monophosphate dehydrogenase (IMPDH), the target of mycophenolic acid.[106] Zaragozic acids, or squalestatin, which target squalene synthase in the cholesterol biosynthetic pathway, have been considered as potential cholesterol-lowering drug. In the BGCs of these compounds, a second copy of

squalene synthase is found in addition to the housekeeping copy encoded elsewhere in the chromosome.[107-108] The chemical structure of the natural products along with the schematic of their BGCs are shown in **Figure 3**.

We recently reported the resistance-guided rediscovery of the sesquiterpenoid aspterric acid as a potential herbicide through the inhibition of dihydroxyacid dehydratase (DHAD), which is one of the key enzymes involved in the branched-chain amino acid (BCAA) biosynthetic pathway that is vital for plant growth.[109] Using DHAD as a potential target, we found a conserved four-gene cluster in several fungi, including the well-characterized *A. terreus*. Three of the four genes are biosynthetic, including a terpene synthase and two P450 enzymes. The fourth conserved gene has ~60% sequence homology to housekeeping DHADs well conserved in fungi. This additional copy suggests this may be a resistance gene, and the enzyme encoded can confer resistance to the natural product encoded in the cluster. Heterologous expression in yeast using the aforementioned host/vector system showed that the three biosynthetic enzymes synthesize the sesquiterpene aspterric acid. Aspterric acid can inhibit the housekeeping DHAD from *A. terreus* and the DHAD from *Arabidopsis thaliana* potently (~300 nM IC50). However, the resistance enzyme in the cluster is insensitive to aspterric acid at the solubility limit of 8 mM. To illustrate the potential that this resistance DHAD and aspterric acid can be further developed for herbicidal purposes, a transgenic plant expressing the resistance DHAD was shown to be insensitive to aspterric acid, while growth of a normal plant is significantly inhibited when aspterric acid is applied. This example showcases the usefulness of the resistance-gene targeted genome mining for discovering natural products with a desired biological activity. Although a known natural product

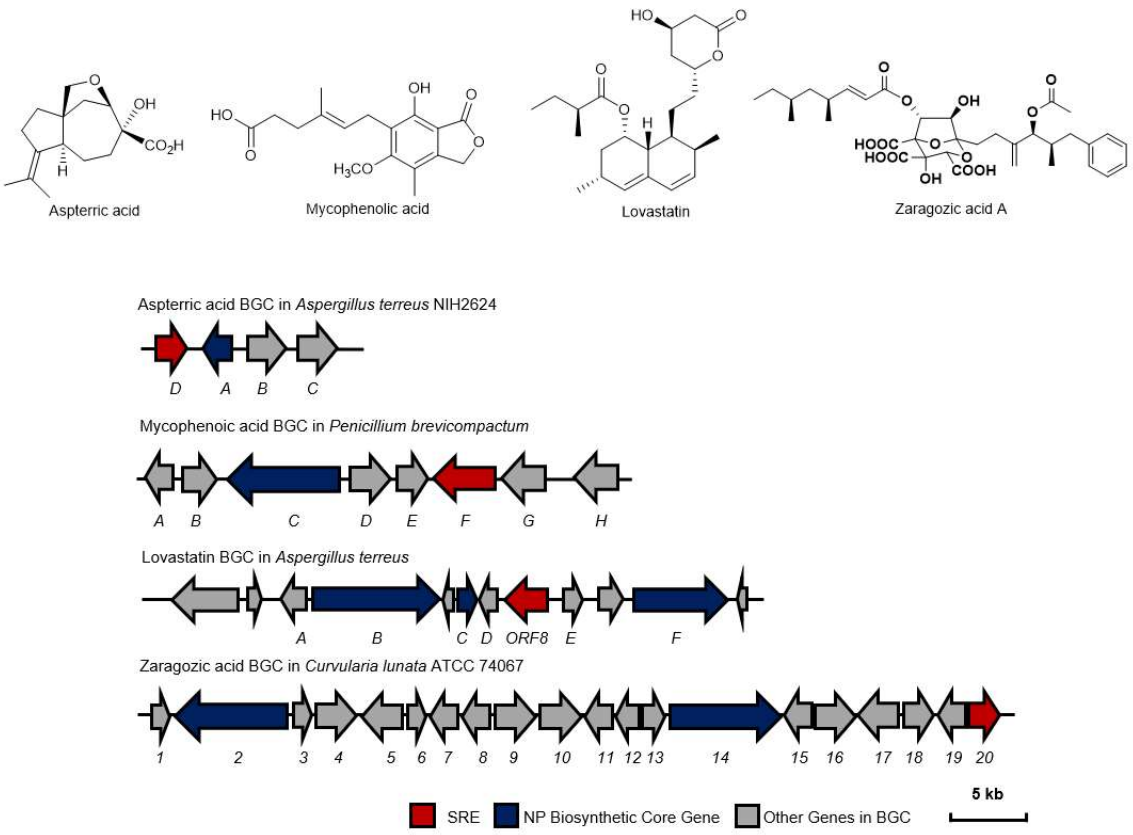


Figure 3. Schematic of BGCs with potential self-resistance enzymes (SREs) – with chemical structures for aspterric acid, lovastatin, mycophenolic acid, and mycophenolic acid.

was rediscovered in this work, the logic of this approach enabled assignment of molecular target of aspterric acid, which otherwise would have been difficult. This approach can also be used to locate biosynthetic gene clusters of compounds of known biological activity. In the case of fellutamide B, which is a potent inhibitor of proteasome, a gene cluster was identified in *A. nidulans* to encode a potential resistance copy of the target, and serial promoter replacement of the genes in the cluster led to activation of this cluster and production of fellutamide B.[110]

Collectively, these works illustrate the usefulness of using a targeted-guided approach to identify and prioritize BGCs. While it is true that not all BGCs have an encoded resistance enzyme,

there are a sufficient number of BGCs that do, and powerful data mining algorithms may be needed to search through the genomes. Lastly, studying the mechanisms of resistance genes available in BGCs can also provide interesting insights into how Nature evolves a housekeeping enzyme to become resistant to its own evolved inhibitor, while retaining catalytic function. This information can offer useful insights into the emergence of resistance against antibiotics, herbicides and other useful natural products.

* Sections 2.2.1-2.2.6 are part of my review article “Genome Mining in Fungi”.[111]

2.3 Conclusion

The advancements in genome sequencing and molecular biology have brought forth a renaissance of natural products discovery. Genome mining of fungi uses several cutting-edge tools in bioinformatic analysis and synthetic biology, starting from powerful computational algorithms for identification of novel biosynthetic gene clusters, followed by manipulation and activation of gene cluster of interests with *in vivo* approaches. Many new natural products previously unknown to science have been discovered using the suite of genome mining approaches described here. Natural product discovery has evolved from a traditional isolation chemistry-dominant discipline to a multidisciplinary effort, combining skillsets from many fields. As a result, the relevance of natural product discovery has been restored as evidenced by the plethora of startup companies using genome mining as a central platform for drug discovery. It is important to note, however, that natural products with beautiful structures and impressive biological activities, will always require rigorous chemical knowledge to determine their structures and improve their biological function. These skills are often overlooked in the age of data-driven science but are absolutely essential to keep natural products discovery a rigorous and fruitful discipline.

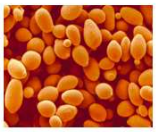
3. HETEROLOGOUS EXPRESSION PLATFORM DEVELOPMENT

3.1 Challenges in fungal heterologous expression system

As discussed in the previous chapter, heterologous expression is one of the key approaches in genome mining. While there are various hosts available for performing heterologous expression, choosing the appropriate host often has a significant impact on the outcome of the examination. In particular, the pros and cons for using *S. cerevisiae* versus filamentous fungi as heterologous host are outlined in **Figure 4**. For instance, when examining the function of genes originating from filamentous fungus, using a fungal host such as *Aspergillus nidulans* could be advantageous, most likely due to the availability of corresponding accessory proteins for natural product biosynthesis. Moreover, with compatible intron processing machineries, only genomic DNA is needed for gene expression in fungal hosts, as opposed to the intron-free complementary DNA (cDNA) required for hosts like *Escherichia coli* and *Saccharomyces cerevisiae*.[112-113] Despite the convenience of fungal hosts in heterologous expression, they often suffer from drawbacks such as a crowded metabolic background or production of toxins, which should be addressed to increase their usefulness.

Although fungal hosts can often splice the exogenous intron correctly, researchers have begun to observe cases with introns mis-spliced by the host, when studying fungi that are phylogenetically distant from the host. For instance, in the heterologous expression of a cluster from *Magnaporthe oryzae* in the heterologous host *Aspergillus oryzae*, one of the three introns was not spliced properly, as demonstrated by fluorescent studies.[114] This is problematic because incorrect splicing often leads to faulty translation and abolishes the function of the resulting enzyme. However, there has yet to be a reported case of intron mis-splicing during heterologous expression of genes within the same genus. In fact, there are only a few choices of heterologous host for filamentous fungi, compared with the vast numbers of fungal genera available. This

limitation is mainly due to the difficulties in generating auxotrophic strains using conventional gene editing methods. For instance, one of the established strategies involves screening hundreds of colonies for a few positive colonies due to random mutation.[115] Another common method utilizes knocking out the gene via replacement with an antibiotic-resistant gene, followed by subsequent recycling of the marker with toxicity selection.[116] Therefore, with the advancement of genome editing technology such as CRSIPR/Cas9, the challenges in optimization of existing fungal hosts and development of new hosts in diverse gerera should be addressed with the new technologies.



	Yeast (*Saccharomyces cerevisiae*)	Filamentous Fungi
Advantages ✓	• **Well-established** genomic tools • **Less** likely to **interfere** with exogenous fungal pathways • Generally regarded as **safe**	• **Can splice** fungal introns (only gDNA is needed) • With **accessory proteins** (PPase and CPR) for fungal natural product production
Disadvantages ✗	• **Cannot splice** fungal introns • Need genes encoded for the accessory proteins to be **co-express**	• Relatively **crowded** metabolic background • May produce toxic compounds such as **mycotoxins**

** **PPTase**: Phosphopantetheinyl transferase
CPR: Cytochrome P450 reductase

Figure 4. Comparison between *S. cerevisiae* and filamentous fungi as a heterologous host.

3.2 Optimizing am existing Aspergillus platform host for heterologous expression

A commonly used heterologous host in our lab is *A. nidulans A1145* that is capable of expressing up to 12 genes using three plasmids.[117] A detailed discussion of its development and examples of use in a biosynthetic study can be found in Chapter 2. Despite its application in fungal genome mining, a major drawback of this host stems from is its crowded metabolic background,

which could hinder the discovery of the targeted natural products during heterologous expression. Using the fungal CRISPR system optimized for *Aspergillus*, “marker-less” genome editing was shown to be achieved in *A. nidulans* readily.[65] By examining the metabolic profile of *A. nidulans A1145* under common expression conditions, sterigmatocystin (ST)[118] and emericellamides (EM)[119] were found to be the major contributors to the rich metabolic background. Therefore, the scaffold PKSs corresponding to these metabolites were knocked-out with CRISPR sequentially, generating *A. nidulans ΔST and A. nidulans ΔSTΔEM* strains. **Figure 5** illustrates the successful application of the engineered hosts in elucidating the biosynthesis of a cholesterol-lowering agent zaragozic acid, allowing for the emergence of the targeted metabolites from the double knock-out strain.[107] This strain is currently wildly used in our lab, with advantages in ease of discovery of new natural products and downstream purification for structural elucidation.

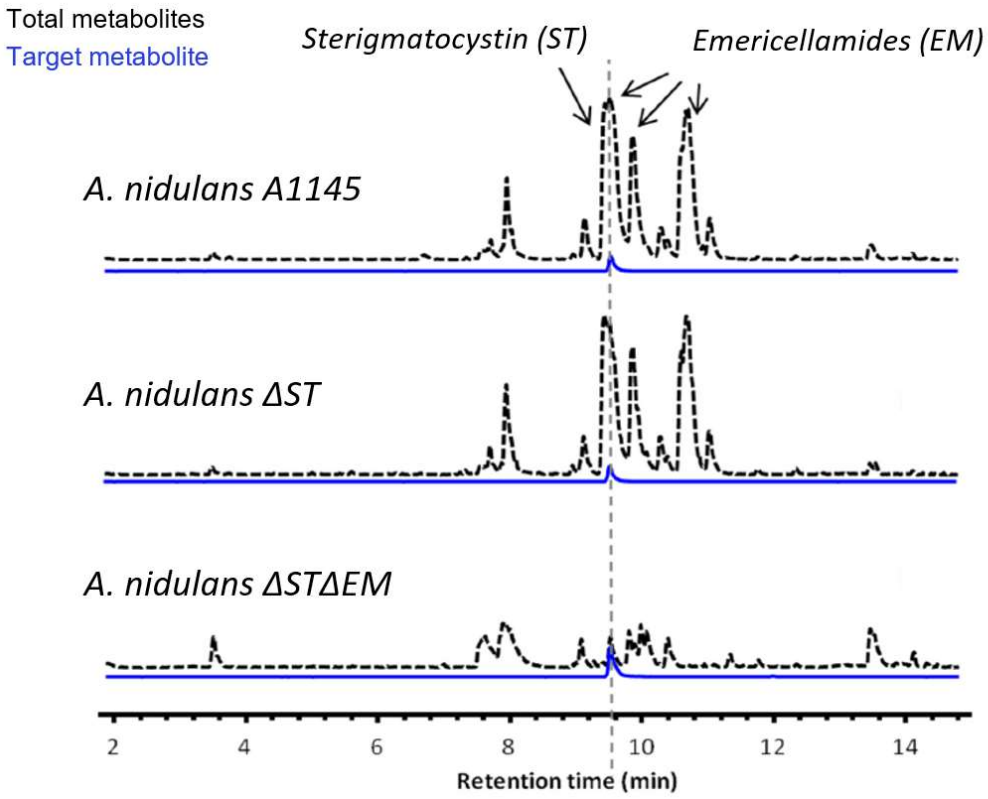


Figure 5. CRISPR engineered *A. nidulans* strains (*A. nidulans ΔST and A. nidulans ΔSTΔEM*) and application in fungal genome mining.

3.3 Developing a new Penicillium platform host for heterologous expression

As the model organism of the *Penicillium* genus, *Penicillium chrysogenum* seems to be the most probable candidate as a new heterologous host.[120] Not only it is the most well-characterized *Penicillium* strain, *Penicillium chrysogenum* is also fast growing and non-toxic under most laboratory culturing conditions, which are important criteria for a heterologous host.[121] However, *Penicillium chrysogenum* was reported to be a prolific producer of metabolites under various laboratory culturing conditions, which is problematic as the crowded metabolic background may hinder the discovery of compounds of interest during heterologous expression.[122] On the other hand, as a fast growing fungus with a minimal metabolic background under the culturing condition of interest,[123] *Penicillium expansum* is perhaps an even better candidate for heterologous host development.

Once establishing *Penicillium expansum* as the choice for heterologous host development, we sought to generate the auxotroph strains by knocking out desired markers. Common markers for a filamentous fungal host include the genes responsible for the biosynthesis of uracil (*pyrG*),[124] pyridoxine (*pyroA*),[125] and riboflavin (*riboB*).[126] As these nutrients are essential for the growth of most filamentous fungi, knocking out these markers followed by the supplementation of an exogenous marker along with the gene of interest is a common technique for fungal heterologous expression. Using CRISPR-facilitated homologous recombination as opposed to homologous recombination alone, the uracil auxotroph in *P. expansum* was generated readily, followed by the selection with 5-Fluoroorotic acid (5-FOA). 5-FOA is commonly used as selection for uracil autotrophy as the orotidine 5'-monophosphate decarboxylase encoded by *pyrG* can convert 5-FOA to a toxic compound to kill the cells, while cells that live in the presence of 5-FOA indicate the disruption of *pyrG*.[127] To increase the expression capacity of the newly engineered host, we decided to generate more auxotrophs, similar to *Aspergillus nidulans* A1145 and other existing heterologous hosts. Since it was not feasible to use antibiotic-resistant genes for genomic

disruption due to limited antibiotic choices, we adopted the strategy of using CRISPR to induce random mutation through non-homologous end joining (NHEJ) to disrupt the function of the genes of interest. Briefly, the CRISPR plasmid targeting the marker is transformed into *Penicillium expansum*. Other auxotrophs were generated subsequently using the workflow outlined in Figure 6A, in which CRIPSR transformants were grown in the presence of supplement in the media, followed by selection in the supplement drop-out media. The final engineered strain *P. expansum (-UPR)* was generated readily after three successive knockouts as show in **Figure 6B**. While knockout efficiency ranged from 10-50% in a locus-dependent manner, the fact that auxotrophic strains were generated readily in a "traceless" fashion demonstrates the immense utility of using a similar methodology in generating more hosts for various fungal genera.

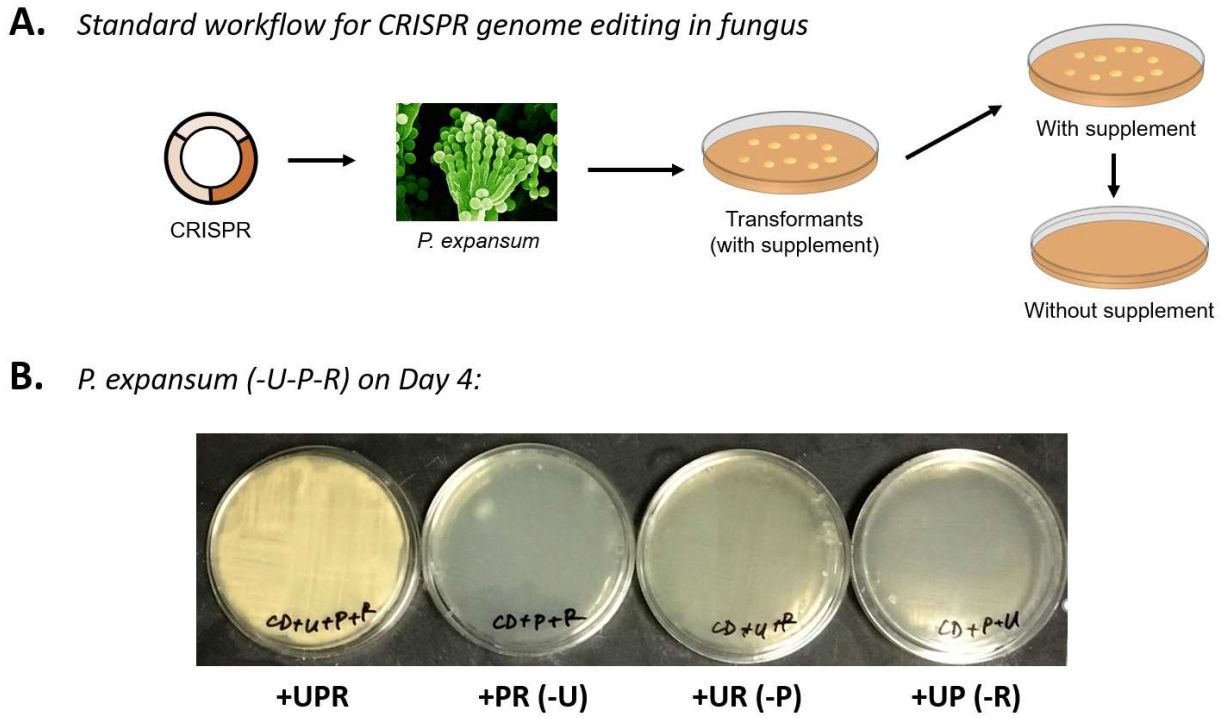


Figure 6. CRIPSR engineered *P. expansum*. (A) Standard workflow for CRIPSR genome editing in fungus. (B) *P. expansum's* selective growth under synthetic media.

After the successful generation of the auxotrophic strains of *P. expansum* with CRISPR, the fastest growing engineered strain *P. expansum(-U)* was examined for its ability to produce a targeted natural product through heterologous expression. We chose to express the PKS gene from the wortmannilactone gene cluster, which is one of three biosynthetic projects presented here and will be discussed in detail in Chapter 4. Briefly, expressing this gene in both *A. nidulans* and P. *expansum* yields the compound with mass of 395(+), as shown in **Figure 7**. However, the metabolic background of *P. expansum* expression was significantly cleaner than that of *A. nidulans*. Also, there is an increase in the production of the PKS product in *P. expansum* compared with *A. nidulans*, which could be due to the increase in precursor supply due to the minimization of the background, or enhancement of enzymatic activity as the cellular machineries are more closely related in *P. expansum*, given the producing host of wortmannilactone is *Penicillium variabile*. Nonetheless, this result suggested that the newly created *P. expansum* host is functional and its performance is comparable with the existing host *A. nidulans*.

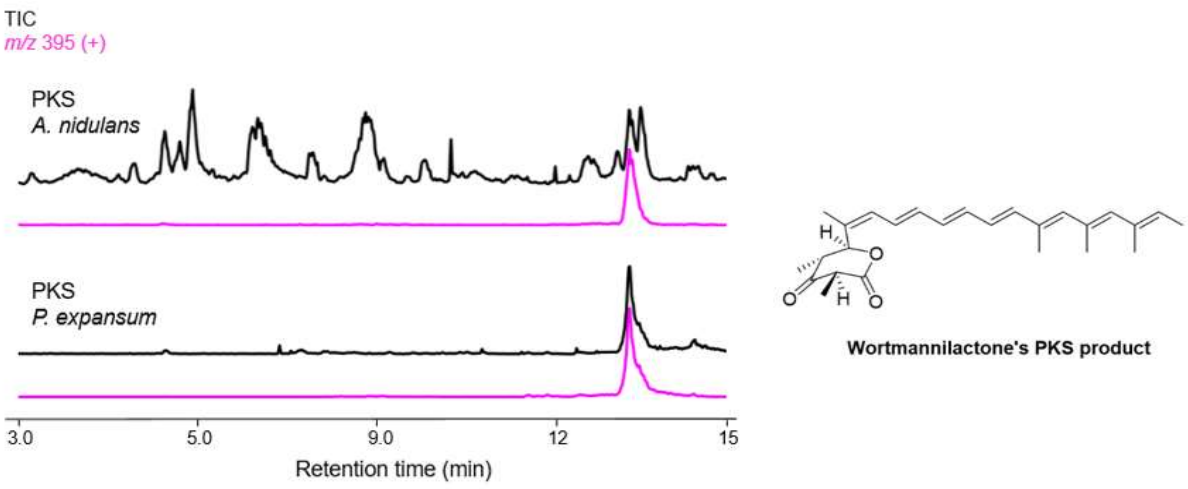


Figure 7. Comparison between heterologous expression in *A. nidulans* and *P. expansum*. TIC: Total ion count; 395 (+) was the filtered-mass corresponding to the molecular weight of the PKS product of wortmannilactone (structure shown on the right)

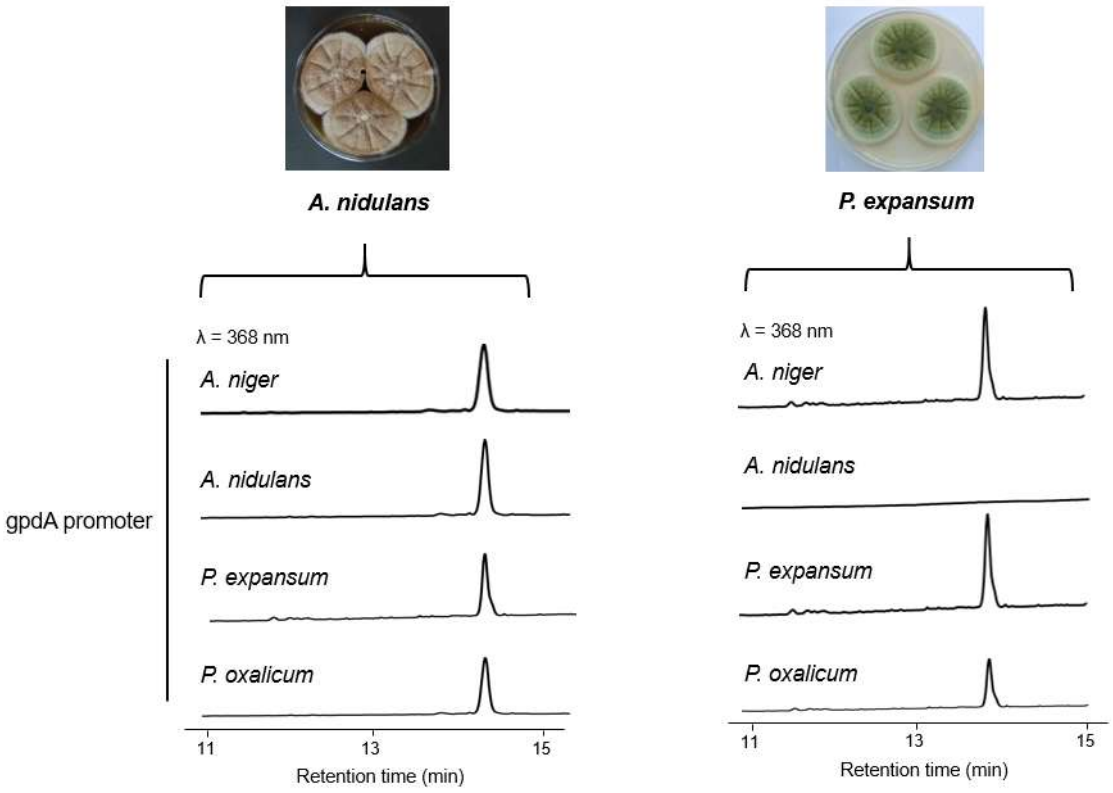


Figure 8. Promoters studies for *A. nidulans* and *P. expansum*. Various sources are used for heterologous expression in *A. nidulans* and *P. expansum* - peak corresponding to the production of wortmannilactone's PKS product.

Being able to generate the auxotroph is only one of many factors for creating a heterologous expression system. Establishing a set of promoters that are suitable to drive the expression of exogenous genes in *Penicillium expansum* is vital.[128] In addition, the production must be stable over the time span of interest. The set of promoters from the *A. nidulans* system have proven to be robust so were a good starting point for testing. Among these, only the gpdA promoter from *Aspergillus niger* worked well in *Penicillium expansum*, but fortunately the gene expression is high and stable up to four days after inoculation. This indicated that the origin of replication is compatible between *Aspergillus nidulans* and *Penicillium expansum*. Therefore, the next step was to establish a new set of promoters for the new heterologous host. Three gpdA promoters were identified from *Aspergillus nidulans*, *Penicillium expansum*, and *Penicillium*

oxalicum with the aim of using the selected promoters to be part of a common set of promoters for a Penicillium expression system. This would ideally increase the chance for successful heterologous expression. The results in **Figure 8** show that except for the gpdA promoter from *Aspergillus nidulans*, all other three promoters worked in both hosts. Therefore, with the new set of common promoters established, the engineered host can now express three genes with one marker and three promoters.

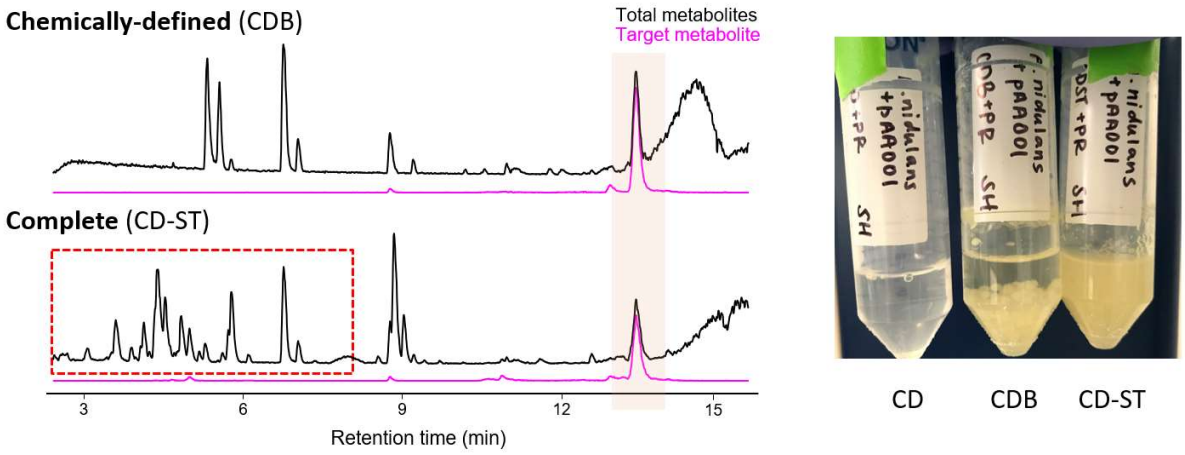


Figure 9. Media studies on *A. nidulans*. (A) Metabolic profiles of *A. nidulans* under different media. (B) Growth morphology of *A. nidulans* under different media.

Moreover, media studies by comparing the commonly used complete media CD-ST with other chemically-defined media suggested using chemically-defined media, such as CDB, could be beneficial in yielding a cleaner metabolic background, most likely to due to the weaker expression of endogenous metabolic pathways under a nutrient-poor environment **(Figure 9)**. The subsequent deviation in growth behavior, using *A. nidulans* as the example host, could also be advantageous depending on the application. For instance, the cells of *A. nidulans* phase separatel and sediment to the bottom phase when growing under CDB, which could be beneficial for downstream purification by saving effort in separating media and cells with the conventional

filtering method **(Figure 9)**. Similar growth morphology and metabolic background were observed in the engineered *P. expansum* host as well (data not shown), and similar media studies should be conducted with other available fungal hosts in order to tailor fungal characteristics for various applications.

3.4 Conclusion

The successful generation of the engineered host *A. nidulans ΔSTΔEM* and *P. expansum (-UPR)*, mediated by the fungal optimized CRIPSR genome editing technology, were shown to be beneficial in performing heterologous expression compared with the previously existing host. They provided a cleaner background, enabling the discovery of natural product targets which were neglected due to low yield and coelution with endogenous metabolites. Expanded choices of promoters also allow the researcher to express more genes at a given time, which is often a bottleneck when examining biosynthetic gene clusters with larger sizes. Lastly, the data regarding the culturing media and growth morphology of the fungal host could also be beneficial parameters to adjust, depending on the purpose of a specific experiment. Overall, the availability of these additional engineered hosts proves to be beneficial, as they can serve as an alternative to currently available choices for fungal heterologous expression, which is an essential approach for fungal genome mining.

4. BIOSYNTHESIS OF WORTMANNILACTONES

4.1 Significance in studying the biosynthesis of wortmannilactones

Wortmannilactones are anti-parasitic polyketides produced by the fungus *Penicillium variabile*[129] and *Talaromyces wortmannii*.[130] Wortmannilactones possess several interesting structural features, namely the 2-oxabicyclo[2.2.1]heptane linked to a dihydropyran by a polyene chain, which were shown to be essential for their inhibitory activities against NADH-fumarate reductase (NFRD).[131] Other natural products featuring the 2-oxabicyclo[2.2.1]heptane skeleton have been isolated from various organisms, including algae[132] and fungi.[133-136] Cymathere ethers,[132] discovered in the early 1990s, were the first natural products identified with a 2-oxabicyclo[2.2.1]heptane skeleton, and were followed by shimalactone A,[133] coccidiostatin A,[134] prugosene A1,[135] and ukulactones A[136] **(Figure 10A)**.

2-oxabicyclo[2.2.1]heptane is an interesting chemical moiety to be studied, as it features a synthetically challenging C-C bond formation between two quaternary stereocenters. Recent biosynthetic studies suggested the origin of the 2-oxabicyclo[2.2.1]heptane. Fujii et al. have shown that a highly-reducing polyketide synthase (HRPKS) ShmA biosynthesizes the polyketide backbone pre-shimalactone, followed by epoxidation by an FAD-dependent monooxygenase (FMO) ShmB to form the pre-shimalactone epoxide. The bicyclo-rings formation follows non-enzymatically from the pre-shimalactone epoxide to form shimalactones, as observed both *in vivo* and *in vitro* **(Figure 10B)**.[137] Moreover, the involvement of the HRPKS and FMO in the formation of the 2-oxabicyclo[2.2.1]heptane moiety was also reported from the biosynthetic studies of aspernidgulenes in *Aspergillus nidulans*.[138] However, as in the case of aspernidgulenes, only the ring-open form aspernidgulenes A1 was observed **(Figure 11B)**, whereas 2-oxabicyclo[2.2.1]heptane was observed in shimalactone biosynthesis. Therefore, by characterizing the biosynthesis of wortmannilactones, we saught to gain mechanistic insights

regarding the formation and cleavage of the 2-oxabicyclo[2.2.1]heptane moiety, as well as to discover previously uncharacterized enzymes involved in these biosyntheses.

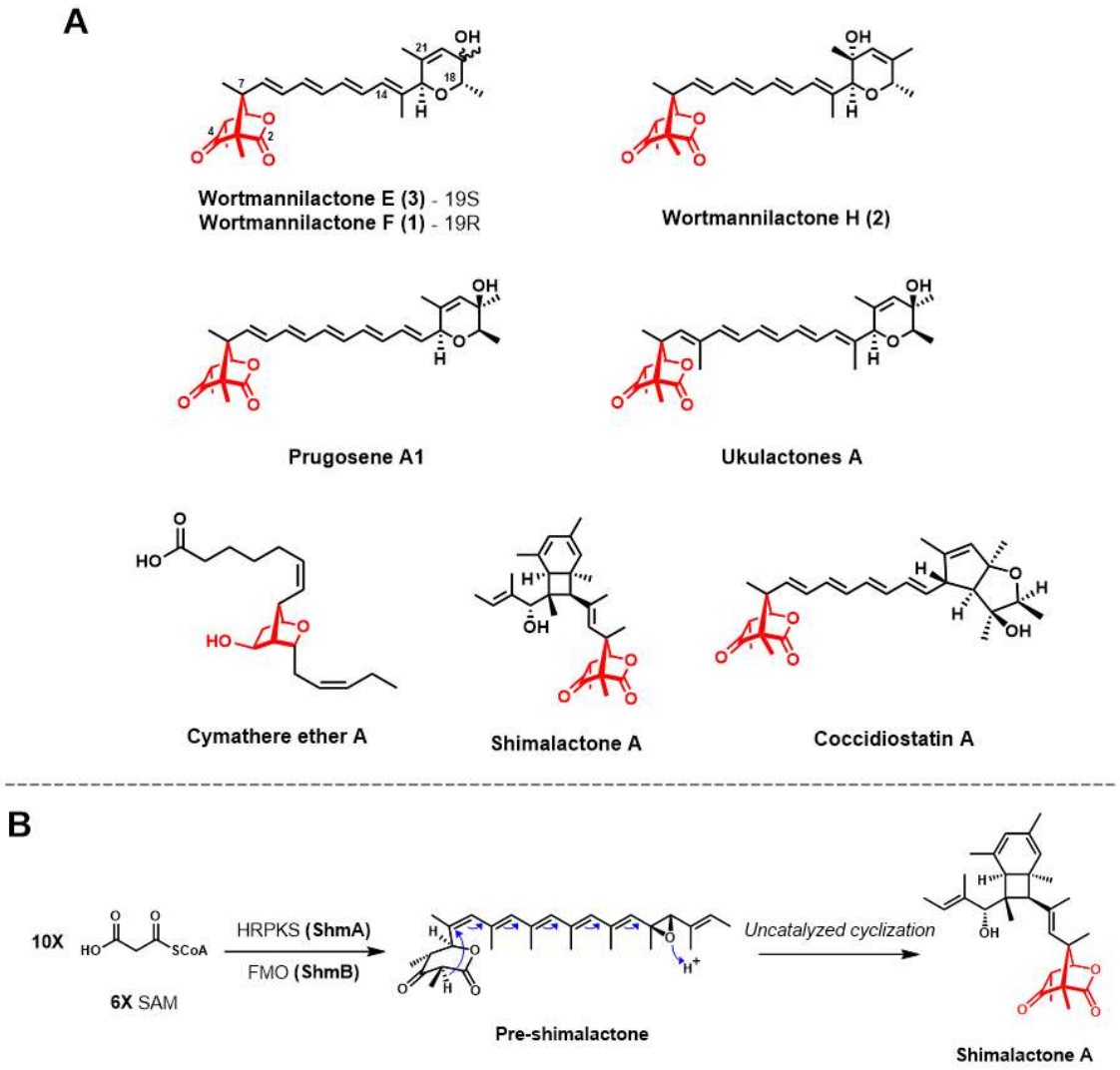
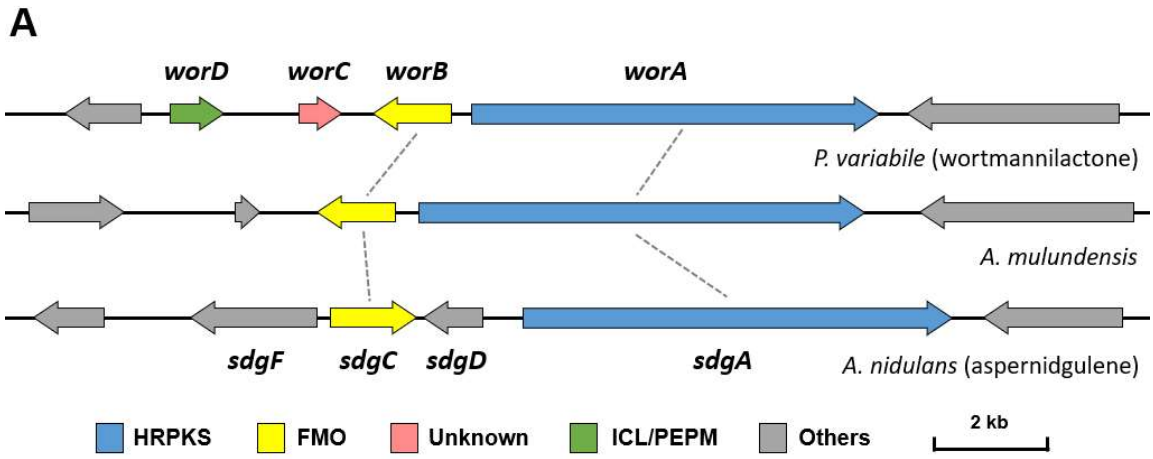


Figure 10. Natural products that consist of the 2-oxabicyclo[2.2.1]heptane moiety. (A) Wortmannilactones and structurally related natural products (with 2-oxabicyclo[2.2.1]heptane moiety highlighted in red). (B) Previous biosynthetic studies of shimalactones.

4.2 Discovering the biosynthetic origin of wortmannilactones

We sequenced the previously reported wortmannilactone producing strain *P. variabile HXQ-H-1*[129] and revealed a total of twenty-six putative HR-PKSs in its genome. With the structural

similarities between aspernidgulenes and wortmannilactones, the gene *sdgA* encoding the corresponding HRPKS from aspernidgulene biosynthesis was used as a query. Comparative genomic analyses suggested a candidate gene in contig 257, termed as *worA*, which shared about 50% identity with *sdgA* and putatively encoded for the HRPKS **(Figure 11)**. Moreover, similar to *sdgA* and *shmA*, *worA* also contains a methyltransferase (MT) domain while lacking an enoyl-reductase (ER) domain. This supports the polyketide backbone of wortmannilactones to be



P. variable (wor)	Putative function	*Homolog in A. nidulans*	*Homolog in A. mulundensis*
worA	Highly reducing polyketide synthase (HRPKS)	*sdgA* (50%)	DSM5745_08393 (50%)
worB	Flavin-dependent monooxygenase (FMO)	*sdgB* (30%)	DSM5745_08394 (30%)
worC	Unknown	N/A	N/A
worD	Isocitrate lyase (ICL)/phosphoenolpyruvate mutase (PEPM) superfamily	N/A	N/A

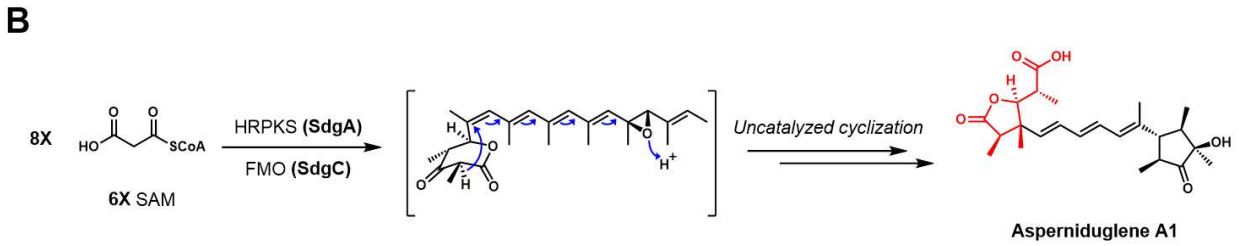


Figure 11. Comparative genomics for *wor* biosynthetic gene cluster. (A) Alignment between *wor* and other homologous gene clusters. (B) Proposed biosynthetic pathway for aspernidulene A1.

the proposed polyene heptaenoyl-β-ketolactone (**5**), which could be further tailored into the final product by an epoxidase and possibly other biosynthetic enzymes **(Figure 14)**. Production of wortmannilactones E **(3)**, F **(1)**, and H **(2)** was confirmed by cultivating *P. variabile HXQ-H-1* in potato-based medium containing the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) as reported previously **(Figure 12A, i-ii)**.[139] According to transcriptional analyses, *worA* is the only candidate that was expressed under the wortmannilactone production condition in *P. variabile*, while containing a MT domain and lacking an ER domain **(Figure S1)**. Transcriptional analyses also suggested that three other genes were expressed along with *worA*: the FMO encoded by *worB* (homologs also present in other homologous gene clusters), an uncharacterized protein encoded by *worC*, and an isocitrate lyase-liked enzyme by *word*; in total just four genes were predicted to be in the *wor* biosynthetic gene cluster for wortmannilactone biosynthesis **(Figures 11A and S2)**.

4.3 Dissecting the biosynthetic pathway with heterologous expression in S. cerevisiae

Heterologous expression of the *wor* gene cluster in *S. cerevisiae BJ5464-NpgA*[140] with *worABC*, either with or without *worD*, led to the same metabolic profile with the production of **1** and **4**, suggesting *worD* is likely to be not involved in wortmannilactone biosynthesis **(Figure 12A, iii)**. This confirmed that the *wor* gene cluster is indeed responsible for the biosynthesis of wortmannilactones. Although expressing *worABC* led to the production of **1**, which was the major product observed in the original producing host *P. variabile*, **4** was also observed from yeast expression and possessed similar UV absorption patterns as **1**. The structure of **4** was determined by spectroscopic methods to be the 2-oxabicyclo moiety opening form of **1**, which resembled the observation of asperniduglene A1 in asperniduglene biosynthesis[138] and prugosenes B1 in prugosene biosynthesis.[135] Previous biomimetic studies on natural products with the 2-oxabicyclo moiety suggested the interconversion between the ring-closed and ring-opened form in a pH dependent manner.[135, 141] Indeed, cleavage of the 2-oxabicyclo moiety was observed upon the

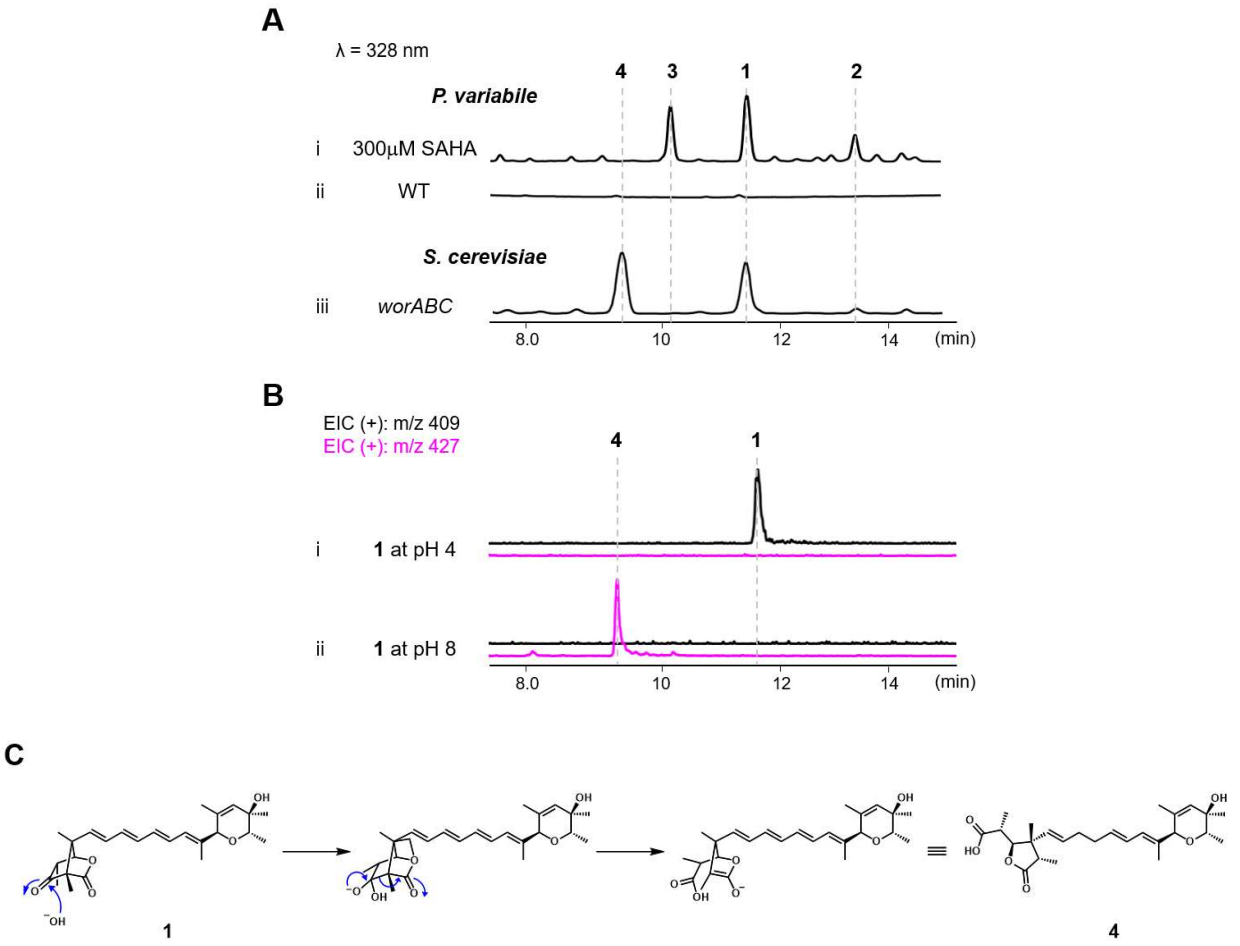


Figure 12. Biosynthetic studies of wortmannilactones. (A) *P. variabile's* induction with SAHA and heterologous expression of *worABC* in *S. cerevisiae*. (B) *In vitro* studies of **1** under pH 4 or pH 8 conditions. (C) Proposed mechanism of base-catalyzed hydrolysis of **1** to **4**.

addition of strong based such as NaOH.[131] These results inspired our testing of wortmannilactones stability under various pHs, specifically in buffers with pH 4 and 8, which were the culturing pH for *P. variabile* and yeast respectively. Complete conversion to **4** was observed by incubating **1** under buffer with pH8 for overnight, while a similar conversion is not observed for buffer with pH 4 **(Figure 12B)**. Therefore, our hypothesis of 2-oxabicyclo[2.2.1]heptane cleavage under basic condition (even in mildly basic condition) was validated from *in vitro* studies, and the proposed base-catalyzed hydrolysis mechanism from **1** to **4** is shown in **Figure 12C**.

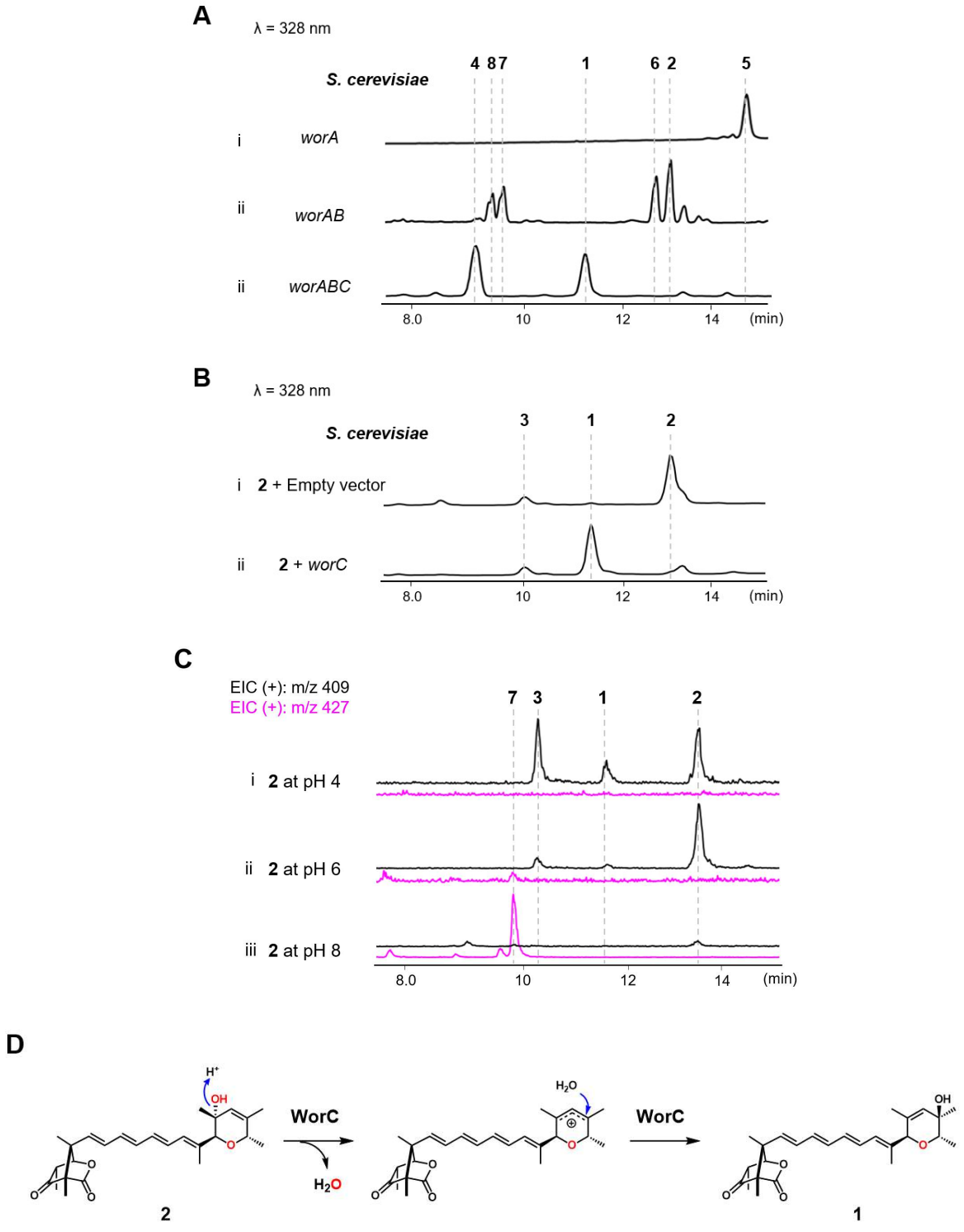


Figure 13. Characterization of the *wor* biosynthetic gene cluster. (A) Expression of *worABC* combinatorically in *S. cerevisiae*. (B) Biotransformation of **2** in *S. cerevisiae* expressing *worC*. (C) Incubation of **2** under different pH conditions. (D) Proposed mechanism of WorC under acidic condition.

To further dissect the biosynthesis of wortmannilactones in order confirm the functions of each gene in the *wor* gene cluster, *worABC* were expressed in *S. cerevisiae* combinatorially. Expression of *worA* in *S. cerevisiae* yielded compound **5**, which was confirmed to be a heptaenoyl-β-ketolactone as proposed. Moreover, co-expressing *worAB* in *S. cerevisiae* led to the production of compounds **2**, **6**, **7**, and **8 (Figure 13A, ii)**. NMR characterization confirmed the structure of **6** as the furan analogue of **2**, an *en route* product derived from a 5-*exo* instead of *6-endo* cyclization. While **7** and **8** were the ring-opened forms of **2** and **6**, respectively **(Figure 15)**.

4.4 Characterization of WorC as potential isomerase

Based on the heterologous expression results in *S. cerevisiae*, WorC was suggested to be responsible to for the formation of **1** from **2**. To further validate the function of WorC, biotransformation studies were carried in *S. cerevisiae* with **2** as substrate. Indeed, almost complete conversion of **2** to **1** was achieved with overnight incubation of yeast expressing *worC*, in contrast with the trace amount of **1** that is observed with yeast expressing the empty vector as a control **(Figure 13B)**. As **1** and **2** are isomers with the hydroxyl group attached to different carbons on the pyran, these results suggest that WorC is an isomerase that could catalyze intramolecular transfer of a hydroxyl group between allylic carbons. In addition, the increase in conversion to **1** from the *in vitro* incubation of **2** with decreasing pH **(Figure 13C)**, supports the hypothesis that WorC employs an acid catalyzed mechanism for the isomerization between **1** and **2**, with the reversibility depending on the pH of the assay conditions **(Figure 13D)**.

4.5 Conclusion

Although there are quite a few homologs with relatively high identity to WorC present in fungal genome databases, none of the candidates have been previously characterized. In addition, they are typically not clustered with a megasynthase, such as a PKS or NRPS, as in the case of WorC. Therefore, WorC is suggested to be the first isomerase clustered with a polyketide

gene cluster that acts on a polyketide scaffold for hydroxyl transfer. Based on the prediction from TMHMM,[142] WorC is a transmembrane protein with six transmembrane helices **(Figure S3)**. However, it does not have any conserved domain. Therefore, as shown by the biotransformation of **2** in yeast, we propose that WorC is a new class of isomerase that could catalyze regio- and stereo-selective isomerization, employing **2** as the substrate.

Even though the amino sequence of WorC does not affiliate with any previous characterized enzymes, WorC is most analogous to linalool isomerase (Lis) from *Thauera linaloolentis* (EC 5.4.4.8),[143] based on the allylic rearrangement reactions catalyzed by both enzymes. Lis belongs to the enzyme family of intramolecular hydroxyl group transferases with the function of 3,1- hydroxyl-Δ^1-Δ^2-mutase, which catalyzes the conversion of linalool to geraniol. A similar reaction was described for the bifunctional enzyme linalool dehydratase/isomerase from *Castellaniella defragrans* (EC 5.4.4.4), in which the enzyme catalyzes the reversible hydration of β-myrcene to (S)-linalool and its isomerization to geraniol.[144] While cofactor-independent enzymes are known for allylic rearrangements that are catalyzed by acid-base mechanisms,[145] isomerization of geraniol to linalool requires a protonation of the hydroxyl group to leave as water, and a shift in electron density leading to a tertiary carbocation intermediate which may be attacked by water or a hydroxyl ion, resulting in the formation of linalool. This proposed mechanism is similar to our proposed mechanism for WorC **(Figure 13D)**. As both Lis and WorC are transmembrane protein with four and six trans-membrane helices respectively, this could indicate the importance of membrane anchoring for proper functioning of these enzymes. Moreover, Lis is also active in the reduced state and sensitive towards detergents, which could guide the further characterization of WorC. Nonetheless, WorC appears to be the first fungal enzyme in the family of intramolecular hydroxyl group transferases.

In conclusion, this work led to the confirmation of the enzymes responsible for the biosynthesis of wortmannilactones, a family potential anta-parasitic drugs. With only three

enzymes, the HRPKS (WorA), FMO (WorB), and a newly characterized isomerase (WorC), different forms of wortmannilactones can be biosynthesized, which illustrates the conciseness and diversity in biosynthesis of rather complex molecules **(Figure 5)**. Discovery of a novel isomerase WorC provides the opportunity of further biocatalytic development and exemplifies the potential of genome mining in novel enzymology discoveries. Lastly, extended studies regarding formation and cleavage of the 2-oxabicyclo moiety could facilitate the generation of natural product analogs for performing structure–activity relationships,[131] which could facilitate drug discoveries.

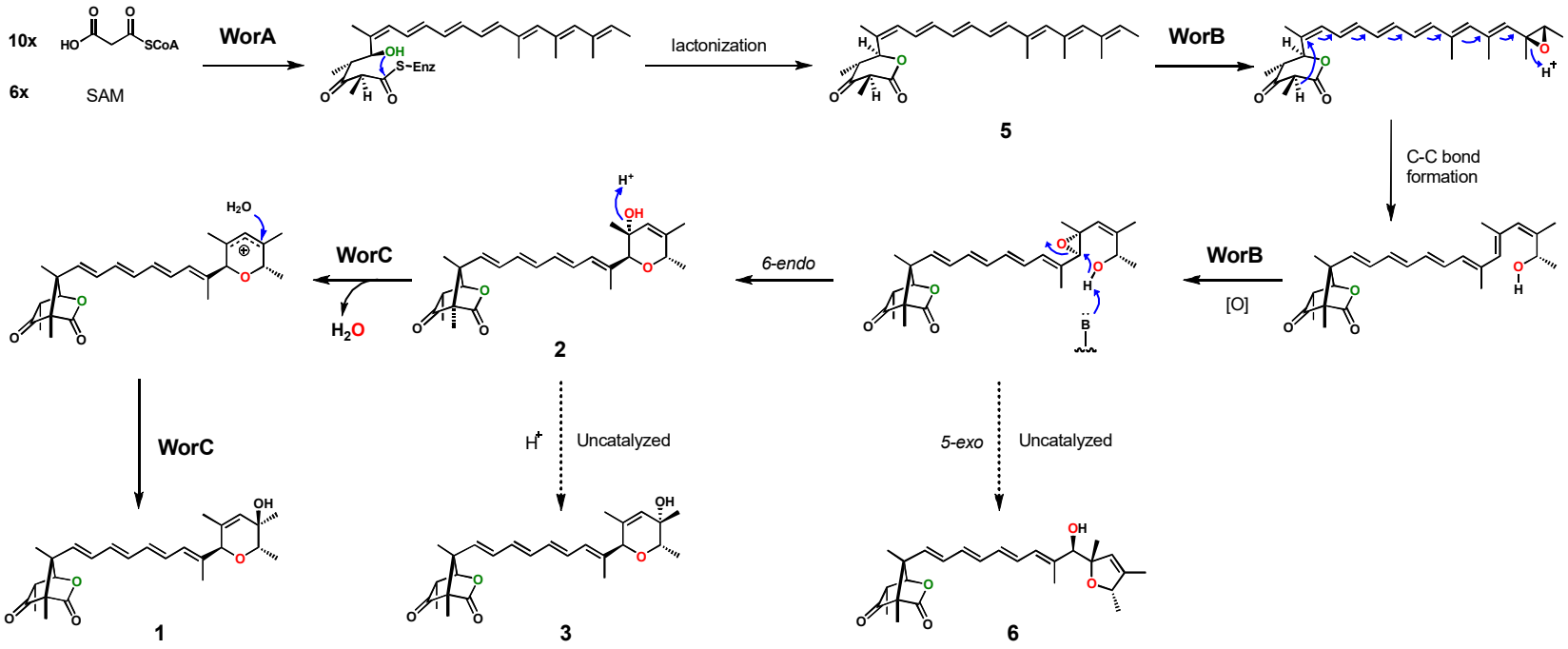


Figure 14. Proposed biosynthetic pathway of wortmannilactones - based on biochemical evidence present in this work.

5. BIOSYNTHESIS OF DICHLORODIAPORTHIN

5.1 Background on enzyme catalyzed aliphatic halogenation

Halogenation has a significant impact on small-molecule pharmaceuticals, whether in enhancing biological activities[146] or generating versatile precursors for a variety of synthetic transformations.[147] In particular, synthetic chemists have developed various methodologies to achieve regioselective halogenation on unactivated sp^3 carbon despite the challenges in maintaining the balance between regioselectivity and chemical yield.[148] While most of the strategies developed employ a relatively harsh and non-ecological conditions, nature, on the other hand, has evolved halogenases as an economically attractive and sustainable solution to achieve these transformation. For instance, halogenases can catalyze direct halogenation of C–H bonds with inorganic halides, such as NaCl, under mild condition. However, these alkyl halides mainly belong to the non-heme Fe^{II}/α-ketoglutarate (Fe^{II}/αKG) family, where a radical substrate is generated by using high-valent metal–oxo intermediate to catalyze the subsequent halogenation on unactivated carbon.[149] This family of Fe^{II}/αKG-dependent alkyl halides mainly acts on the carrier-protein-tethered substrates,[150-152] with a few exceptions on standalone substrates.[153-155]

Flavin dependent halogenases (FDHs), in contrast, have a lot less reported examples of catalyzing halogenation on unactivated carbons, mainly due to its electrophilic aromatic substitution mechanism and utilization of either a bound hypochlorous acid or chloramine species instead of radicals.[156] Examples for FDHs that are proposed to catalyze aliphatic halogenation include CmlS for chloramphenicol[157] and PloK for 3-chloro-6-hydroxymellein[158] **(Figure 15A)**. Despite this, the halogenation mechanism by such FDHs remained elusive, as the conventional electrophilic mechanism cannot satisfactorily explain such unactivated halogenation. However, dichlorination of chlorotonil A catalyzed by CtoA suggested that such halogenation could occur

on the α-carbon of a diketo moiety **(Figure 15A)**. This type of catalysis would be compatible with the conventional mechanism of FDH due its significantly lower pKa compared with the terminal methyl carbon; a subsequent modification, such as deacetylation, could give rise to halogenation on the terminal carbon as in chloramphenicol.[159] To date, the only characterized FDH that catalyzes aliphatic halogenation is AoiQ for the biosynthesis of dichlorodiaporthin, where AoiQ was reported to catalyze dichlorination on diaporthin **7** to form dichlorodiaporthin **1 (Figure 15B)**. Although the authors suggested that AoiQ may employ a novel mechanism to achieve such aliphatic halogenation,[160] the low turnover observed from *in vitro* studies suggested that AoiQ may act on a different substrate. Indeed, our investigation, both *in vivo* and *in vitro*, suggested that the substrate of AoiQ is **2** instead of **7**, which dichlorination occurs on the diketo moiety of **2** followed by subsequent deacetylation and reduction to yield final product **1**.

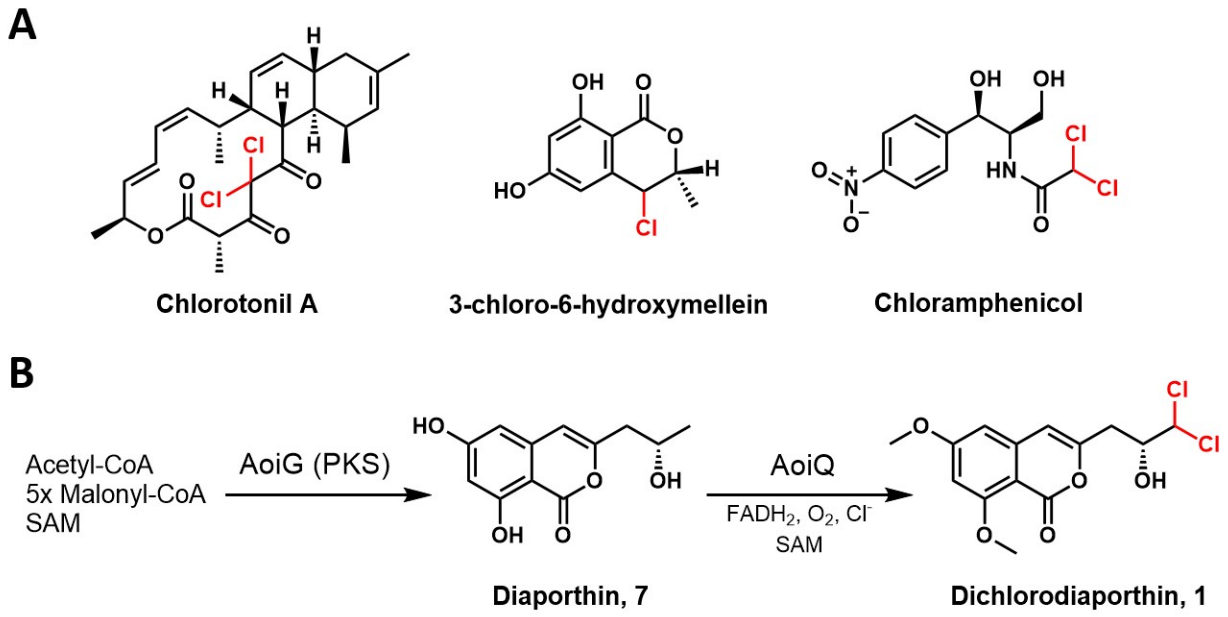


Figure 15. Natural products with aliphatic chlorination. (A) Representative natural products with chlorination on sp^3 carbon by FDHs. (B) Proposed mechanism for AoiQ to catalyze the conversion of diaporthin **7** to dichlorodiaporthin **1**.

5.2 Examining the function of AoiQ through heterologous expression

AoiQ was reported to be unclustered with the *aoi* biosynthetic gene cluster, which was responsible for the biosynthesis of the precursor **7**. As biosynthetic crosstalk in fungal biosynthetic pathways is extremely rare,[92, 161] *aoiQ* may potentially cluster with the gene that putatively encods for a non-reducing polyketide synthase (NRPKS) in the vicinity instead (termed as *diaA*).[160] By performing genome mining and comparative genomics across the fungal kingdom, results suggested that *aoiQ* is conserved in many other fungi along with *diaA*, while the rest of the *aoi* gene cluster is absent. Coincidentally, the biosynthetic gene cluster that consists of *diaA* is conserved and always colocalized with *aoiQ* for many different fungal strains, including *Aspergillus nidualns* and *Trichoderma harzianum* **(Figure 16A)**. Apart from *aoiQ*, other conserved genes in the *dia* biosynthetic gene cluster are *diaA-E*. As a result, DiaA could be responsible for the biosynthesis of the polyketide backbone as an NRPKS, while other genes in the *dia* gene cluster could potential encode for downstream tailoring enzymes, with DiaB as a beta-lactamase, DiaC as a short-chain dehydrogenases/reductases (SDR), and DiaE as a flavin-dependent monooxygenase (FMO). DiaD, with a putative function as a transcription factor, could be responsible for regulating the expression of the *dia* gene cluster.

To examine the hypothesis that *aoiQ* clusters with the *dia* biosynthetic gene cluster, heterologous expression in the *A. nidulans ΔSTΔEM* was performed.[107] When expressing *aoiG/F* and *aoiQ* in *A. nidulans*, production of **1** was not observed, while only the demethylated **7** was observed. However, expressing *diaA/B* and *aoiQ* in *A. nidulans* led to the production of **1**, which suggested that *aoiQ* clusters with the NRPKS *diaA* instead of *aoiG* **(Figure S1)**. However, since *A. nidulans* also consists of the *dia* gene cluster **(Figure 16A)**, heterologous expression in *Saccharomyces cerevisiae* would be more appropriate to further dissect the biosynthetic pathway to avoid potential cross-talk. In contrast with the expression results from *A. nidulans*, expressing

diaA/*B* and *aoiQ* in *S. cerevisiae* only yielded a trace amount of **1**, along with many degradated products as observed from the HPLC profile **(Figure 16B, trace iii)**. However, **1** was observed as a major product when expressing *diaC* in addition to these three genes, where the reduced non-chlorinated **4** and **5** were observed, either with or without methylation respectively **(Figure 16B, trace iv)**. Moreover, expression of *diaD* in addition to these four genes yield a similar metabolic profile, which suggested that DiaD is likely to be not involved in the biosynthesis of dichlorodiaporthin (data not shown).

To further investigate the potential substrate for AoiQ, expression of *diaA* alone in yeast yielded **2**-**4 (Figure 16B, traces i)**. NMR characterization suggested that **2** was the diketo polyketide product from NRPKS, which contained one ketide extra compared with **7**. Moreover, **3** and **4** were the reduced from of **2**, on the terminal and β-ketone respectively, which could result from reduction by endogenous reductases from yeast. While co-expression of *diaA/B* yielded a similar metabolic profile, the yield was increased by approximately 10 fold **(Figure 16B, trace ii)**. Since DiaB was annotated as a beta-lactamase, it could possibly assist the release of the NRPKS product **2** from DiaA, analogous to the NscA/B pairs from neosartoricin biosynthesis.[162] Based on the heterologous expression results in *S. cerevisiae,* a putative biosynthesis of **1** was proposed and shown in **Figure 16C**. DiaA/B are responsible for the synthesis and release of polyketide **2**, which can be reduced to **3** or **4** by either DiaC or endogenous reductases from yeast. The non-reduced **2** could possibly serve as substrate for AoiQ, and dichlorination could occur on the acidic α-carbon, followed by deacetylation and reduction by DiaC to yield **1**.

5.3 Dissecting the halogenation mechanism of AoiQ

Based on the expression results in *S. cerevisiae*, **2** was the most probable candidate as a substrate of AoiQ. Feeding **2** into yeast expressing *aoiQ* and *diaC* yield **1** as the major product,

along with **4** and **5 (Figure 17A, traces ii)**. However, **1** was not observed when **4** or **5** was fed to yeast expressing *aoiQ* and *diaC*, which suggested that they were not the substrate of *aoiQ* (data not shown). Moreover, only trace amount of **1** was observed when feeding **2** to yeast expressing *aoiQ* alone **(Figure 17A, traces i)**, where similar results were observed from heterologous expression. Furthermore, *in vitro* studies with purified AoiQ and DiaC also showed the conversion of **2** to **1**, where the mono-chlorinated **6** was also observed as the by-product, only in the presence of DiaC and absence of flavin reductase **(Figure 17C)**.These biotransformation results, along with similar observation from *in vitro* studies, suggested that **2** was the substrate for AoiQ instead of the previously reported **7**.

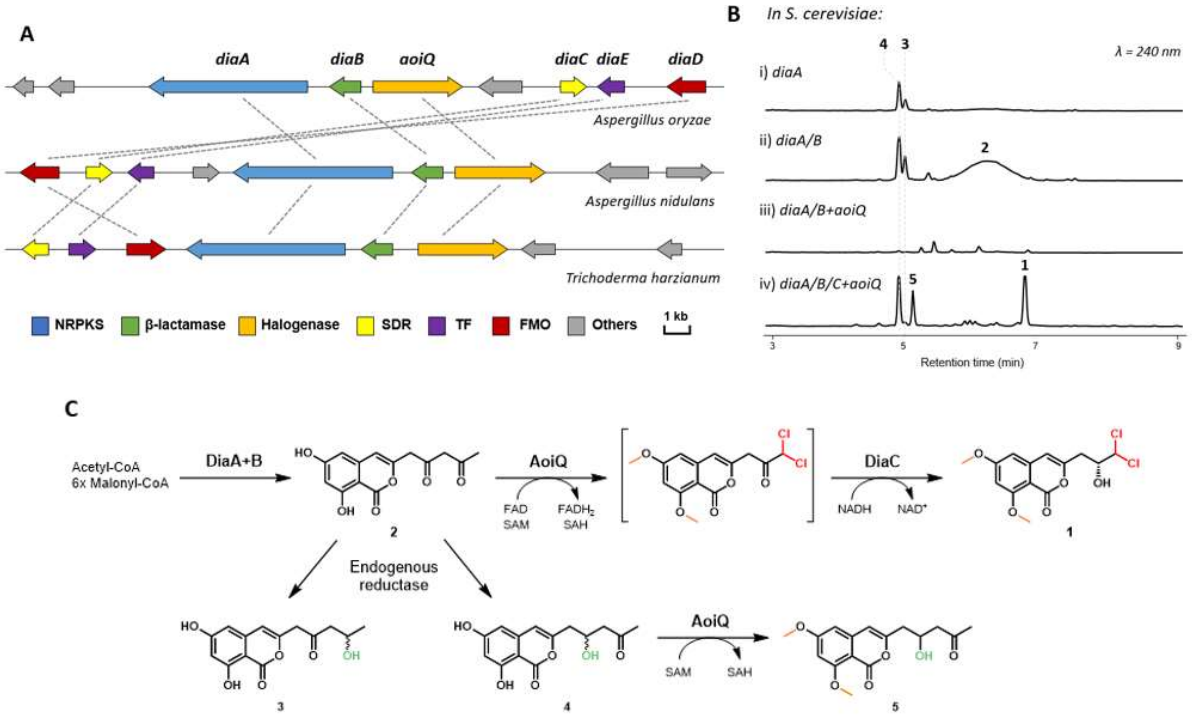


Figure 16. Characterization of the *dia* biosynthetic gene cluster. (A) Comparative genomics for the AoiQ gene clusters. (B) Yeast heterologous expression for the AoiQ gene cluster. (C) Proposed of complete biosynthetic pathway of **1**.

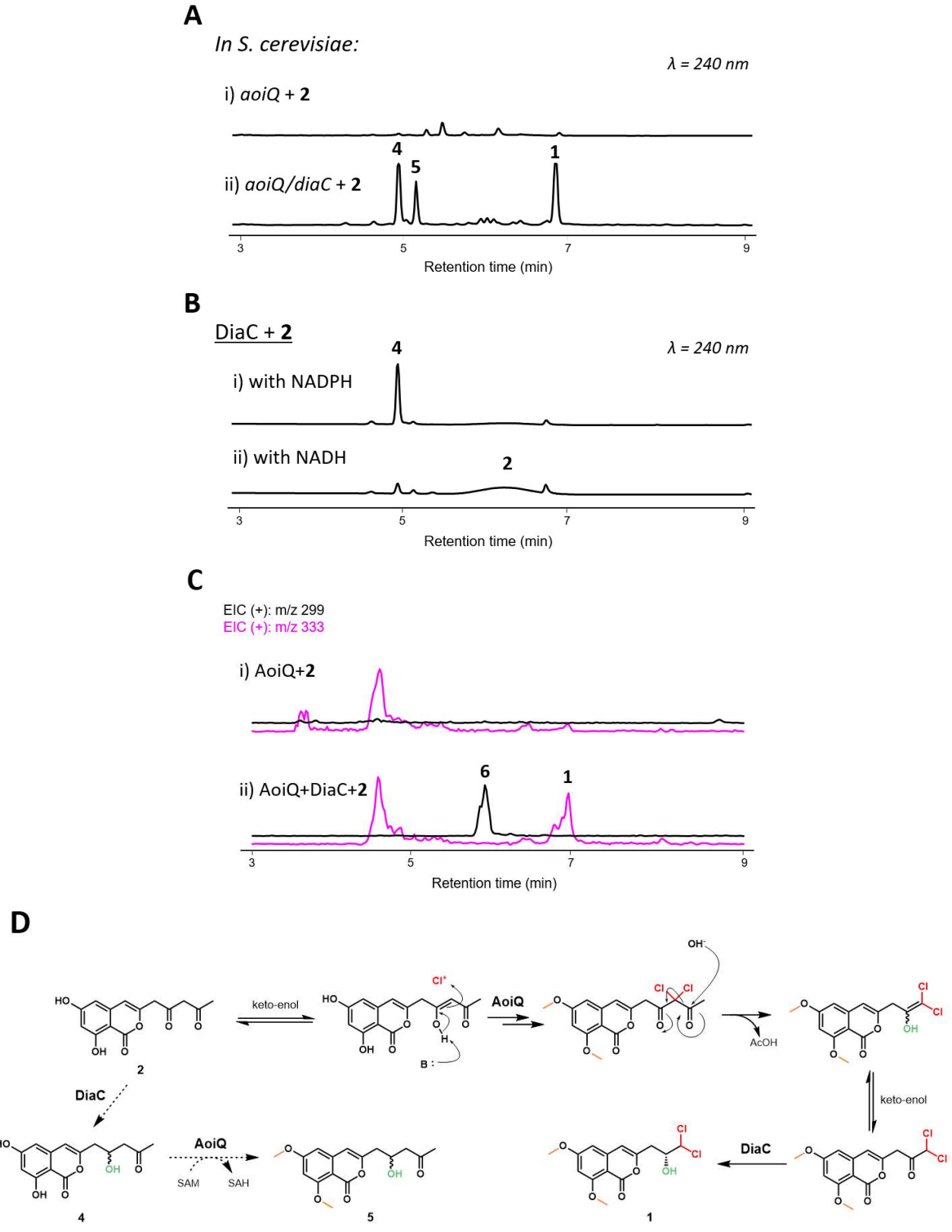


Figure 17. Characterization of AoiQ and DiaC for chlorination of **2**. (A) Biotransformation studies in yeast with **2**. (B) *In vitro* characterization of reduction of **2** by DiaC with **2**. (C) *In vitro* characterization of halogenation of **2** by AoiQ and DiaC. (D) Proposed halogenation mechanism for AoiQ and DiaC.

In addition, the role of DiaC was also examined *in vitro*, where **4** was observed to be the major product when **2** as the substrate, and NADPH was shown to be the preferred cofactor for DiaC instead of NADH **(Figure 17B)**. Although DiaC was proposed to act on the oxidized **1** as the intended substrate in the biosynthesis of dichlorodiaporthin, **2** was used as substrate instead due to the instability of oxidized **1**, suggesting DiaC reduction on the β-keto may enhance the stability of the chlorinated products. Therefore, with both *in vivo* and *in vitro* characterization, the proposed enzymatic cascade catalyzed by AoiQ and DiaC with **2** as the substrate was shown in **Figure 17D**.

5.4 Conclusion

By determining that **2** is the actual substrate of AoiQ, both *in vivo* and *in vitro*, we suggest that halogenation on unactivated sp^3 carbon can proceed via the conventional FDH mechanism, with the diketo moiety as the staring scaffold. Indeed, **2** and the previously reported substrate **7** shared a very similar polyketide skeleton, with **2** being an extra ketide longer and maintained in the oxidized form when offloaded from NRPKS. With the pKa of the alpha carbon in diketo **2** being much lower compared with the terminal carbon in alcohol **7**, this makes the abstraction of a proton to be more feasible by a basic residue, according to previous structural studies of FDHs. Alternatively, potential point mutation studies on the highly conserved lysine residue, which is responsible for the formation of the lysine chloramine species, can further support this hypothesis. Similar utilization of the diketo moiety for halogenation was also shown in organic syntheses,[163] and is consistent with the examples of natural products like chlorotonil A, where the diketo moiety remains in the absence of deacetylation.[159]

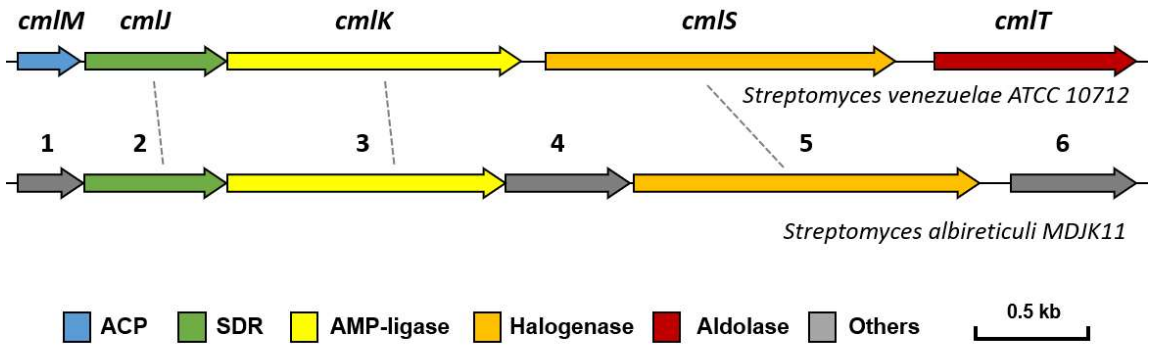
In addition, the involvement of the SDR DiaC was also unexpected, but clearly shown to be the case from both *in vivo* and *in vitro* studies. DiaC was shown to selectively reduce the β-

ketone in the diketo moiety of **2** from *in vitro* studies. However, the actual role of DiaC was expected to reduce the same location on the dichlorinated and deacetylated compound, which is oxidized form of **1**. Such reduction was likely to stabilize the final product **1**, as the oxidized form could be unstable, as exemplified by the degradated compounds upon expression of aoiQ in yeast, as well as the absence of reported cases from the natural hosts. Moreover, the premature reduction, whether without chlorination (**4**) or with mono-chlorination (**6**), were likely to be unintended and led to shunt product formation. Therefore, the timing of each catalyzing event, including methylation by AoiQ, were intricately tuned, which could also explain the relatively low turnover observed in an *in vitro* setting, with chlorination as the rate limiting steps.

Perhaps a more striking finding from this study was the observation of chlorinated compounds, such as **1** and **6**, *in vitro* in the absence of flavin reductase. Flavin reductase is required to generate the reduce cofactor $FADH_2$ or $FMNH_2$ in many flavin-dependent enzymes, and it was shown to be essential for FDHs in particular. In other studies of FDHs, exogenous flavin reductase was included along with the FDH of interest, with a few exceptions of flavin reductase included in the FDH itself. However, chlorinated compounds were not observed in the presence of the commonly used flavin reductase FRE and SsuE, which is likely due to the incompatibility of these flavin reductases with AoiQ. Searching for the compatible flavin reductase may increase the turnover of *in vitro* conversion to a similar extent to *in vivo*, which could enhance the usefulness of AoiQ to further be developed into a practical biocatalyst upon optimization for catalyzing the synthetically challenging regioselective halogenation of an unactivated carbon, as well as extending the knowledge on aliphatic halogenation by FDHs.

The findings in this study could be insightful for biosynthetic efforts towards other structurally similar compounds. Chloramphenicol is another well-known antibiotic that also contains a dichloro-acetyl moiety, and chlorination was shown to be catalyzed by the FDH

CmlS.[164] However, genetic studies also suggested that there were missing genetic element other than CmlS for chlorination.[164] Coincidentally, the SDR CmlJ, which was shown to be essential for the production of chloramphenicol,[165] could act along with CmlS similar to the AoIQ/DiaC pairs. Therefore, we focused our efforts on examining the function CmlJ, as well as its potential coupling with CmlS. Clustering of CmlJ and CmlS shown in other organisms, other than the producing host *Streptomyces venezuelae*, could serve as evidence for the potential involvement of CmlJ in chlorination for chloramphenicol **(Figure 18)**.



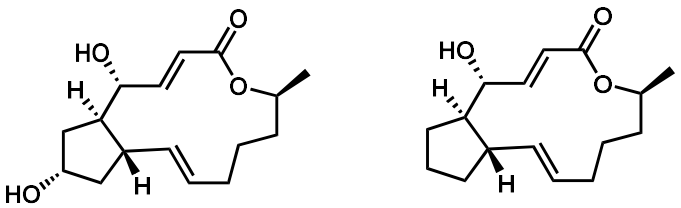
Streptomyces albireticuli (Accession #)	Putative function	*Streptomyces venezuelae's homolog* (% identity)
1 (SMD11_RS33490)	Serine protease	N/A
2 (SMD11_RS33485)	Short-Chain Dehydrogenase/Reductase	*cmlJ* (70%)
3 (SMD11_RS33480)	Acyl-CoA-ACP synthetase, AMP-ligase	*cmlK* (54%)
4 (SMD11_RS33475)	UDP-N-acetyl-bacillosamine N-acetyltransferase	N/A
5 (WP_087930012)	Flavin-dependent halogenase	*cmlS* (68%)
6 (SMD11_RS35655)	4'-phosphopantetheinyl transferase	N/A

Figure 18. Biosynthetic gene cluster of chloramphenicol *(cml)* with comparative genomics.

6. BIOSYNTHESIS OF BREFELDIN A

6.1 Significance and challenges of studying the biosynthesis of brefeldin A

Brefeldin A (**1**) was first isolated in 1958 by Singleton et al. from *Penicillium decumbens*.[166] Other related natural products, including the biogenetic precursor brefeldin C (**2**), share core structure of a fusion between a cyclopentane and 13-membered lactone **(Figure 19)**.[167-168] Apart from its outstanding bioactivities as antifungal,[169] antiviral,[170] and antimitotic,[171] brefeldin A is widely used in biomedical research for its prominence effects on intracellular protein trafficking. Upon incubation with brefeldin A, eukaryotic cells rapidly disassemble the Golgi apparatus and redistribute its constituents into the endoplasmic reticulum (ER). This massive but reversible morphological change is caused by binding of brefeldin A to a protein complex consisting of a catalytic guanine exchange factor (GEF) and the small G protein adenosine ribosylation factor 1 (ARF1), which exerts key regulatory functions for vesicle budding and transport.



1
Brefeldin A
(BFA)

2
Brefeldin C
(BFC)

Figure 19. Natural brefeldins, including **1** brefeldin A (BFA) and **2** brefeldin C (BFC).

Equally rich is the synthetic record of brefeldin A, in which than 40 different strategies in pursuit of brefeldin A or analogs have been described over the past decades, with focus including the construction of the synthetically challenging stereo-specific cyclopentane moiety. On the other hand, biosynthetic studies regarding the cyclopentane moiety, Including the reductive cyclization

catalyzed by iridoid synthase (ISY) in partnership with reductases NEPS and MLPL, to mediate hydride transfer from the NADH or NADPH cofactor and control the profile of stereoisomers formed from the spontaneous cyclization reaction.[172] In addition, a ring contraction mechanism is utilized in the biosynthesis of gibberellic acid, where a homolytic cleavage of C–H bond mediated by a P450 CYP88A yields a radical that rearranges to convert middle ring from cyclohexane to cyclopentane.[173] On the other hand, previous biosynthetic studies on **1** from our group were limited to the initial steps involved in the formation of the polyketide backbone, mostly due to the challenges in genome editing for the producing host *Penicillium brefeldianum*.[174] Therefore, with the advancement of genome editing technology such as CRISPR-Cas9, as discussed in Chapter 2, the biosynthesis of **1** was revisited, with the hope of realized analogue development upon elucidation of the biosynthesis of **1**. However, as well as resolving the formation of cyclopentane with a fungal origin.

From genome sequencing and mining of the producing strain *P. brefeldianum* ATCC 58665, previous studies have identified an HRPKS involved in the biosynthesis of BFA, followed by reconstitution of its activity in *Saccharomyces cerevisiae* and *in vitro*. The previous proposed biosynthesis is shown in **Figure 20**. In summary, the HRPKS is proposed to synthesize the precisely reduced octa-ketide precursor, which could then be directly offloaded by the thio-hydrolase enzyme followed by a P450-mediated formation of the cyclopentane ring and macrocyclization to afford the 7-deoxy BFA **2** (top scheme). Alternatively, the first ring annulation can also occur on the ACP-tethered intermediate before the thio-hydrolase release and lactonization (bottom scheme). The C7-hydroxylation is believed to be the final step in the process to obtain the final structure of **1**.

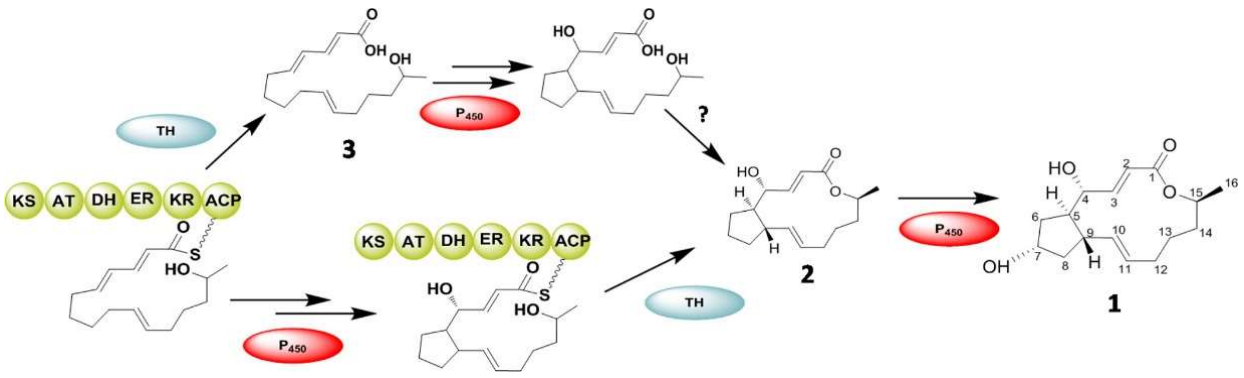
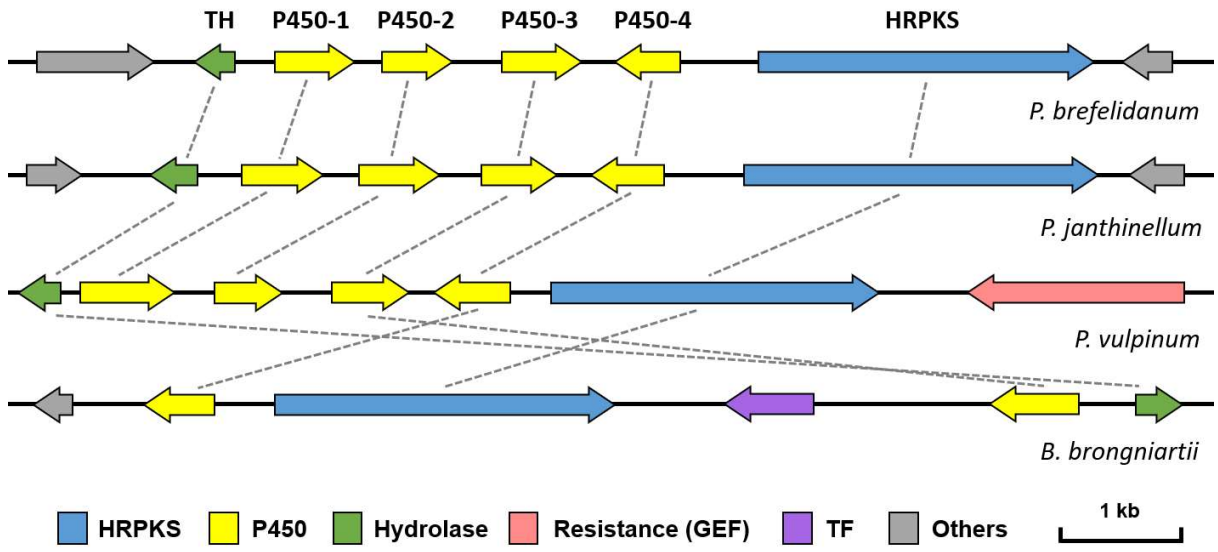


Figure 20. Putative biosynthetic pathway for **1** as proposed by previous studies.[174]

6.2 Determining the biosynthetic origin of BFA with CRISPR knockout

With the rapid increase in the number of fungal genomes being deposited in the database, revisiting the gene clusters and comparative genomics could yield valuable information. As shown in **Figure 21**, the predicted gene cluster from the previous studies aligned well with another reported host for **1** and its dimers (dibrefeldins) called *Penicillium janthinellum*,[175] with the exception of *orf8* which a gene with unknown function that was transcribed during the production of **1** in *P. brefeldianum*. Similarly, all 6 genes are aligned with high identity (more than 80%) to another fungus called *Penicillium vulpinum*, which the existence of a potential resistance copy of the Sec7 domain could present an alternative resistance mechanism of **1** from the fungus. However, only P450-3 and P450-4 are aligned when looking into the putative BFA gene cluster in *Beauveria brongniartii*, where the identity of the 4 homologous genes are also greater than 80%. Although this fungus is not yet a reported strain for production of **1**, this is the only non-*Penicillium* strain that possess the biosynthetic cluster of **1** on the database to date. This could resolve the prior debate on the presence of two excess P450s, enabling the following hypothesis around the biosynthesis of brefeldin A: the PKS and TH biosynthesize the backbone of brefeldin A with the tailoring oxidation steps from the remaining two cytochrome P450s (P450-3 and P450-

4). Since there is no report of confirmation of the biosynthetic origin of brefeldin A, we decided to the confirm the biosynthetic origin of **1**.



P. brefeldianum	Putative function	*Homolog in B. brongniartii*
HRPKS	Highly reducing polyketide synthase	BBO_07026 (82%)
TH	Thiohydrolase	BBO_07031 (87%)
P450-1	Cytochromes P450	N/A
P450-2	Cytochromes P450	N/A
P450-3	Cytochromes P450	BBO_07030 (85%)
P450-4	Cytochromes P450	BBO_07027 (86%)

Figure 21. Comparative genomics for the *bref* biosynthetic gene cluster.

P. brefeldianum ATCC 58665 is a common BFA producing strain utilized industrially. Following the similar culturing condition for the optimized production of **1** as reported,[176] production of **1** was observed readily and allowed us to the proceed in generating a BFA null mutant. However, the challenges in performing genome editing in the BFA producing strain *P. brefeldianum ATCC 58665* is due to its non-sporulating nature, which is common for filamentous

fungi. There previous attempts at genetic manipulation were futile due to multinucleated protoplasts.[174] With the emergence of the CRISPR editing technology, there were reports on successful genome editing in genetically intractable fungi, those which cannot be modified with traditional methods, with the fungal optimized CRISPR system.[65] Since this fungal CRISPR system has been successfully applied to other *Penicillium spp.*, it could potentially work on *P. brefeldianum*.[69] Using CRISPR, the gene encoding the HRPKS was knocked out successfully with relatively high efficient (1 in 4 colonies), and knock-out was confirmed by both PCR verification on the specific locus from the genome of the mutant, as well as observing the abolishment of the production of **1** in the KO strain of the scaffold gene PKS with liquid chromatography–mass spectrometry (LC-MS) **(Figure 22A)**. This allowed for confirmation of the biosynthetic origin of brefeldin A for the first time, and we decided to employ similar knockout techniques to resolve the remaining biosynthetic pathway of **1**.

Apart from the HRPKS, three other genes were knocked-out successfully, including the genes putatively encoding P450-1, P450-3, and P450-4. Knocking out the genes encoding P450-1 and P450-2 generate a similar metabolic profile as observed from the wildtype *P. brefeldianum*, which suggests either they are not involved in the biosynthesis of **1**, or there are redundant copies of these genes located somewhere else in the gene cluster or genome. We hypothesized that they are likely to be not involved in the biosynthesis of **1** because the absence of these genes from the homologous gene cluster in *B. brongniartii*, as well as the lack of a homolog of P450-2. For P450-1, interestingly, it shares a relatively a high identity with P450-4 for a 60% identity. However, we demonstrated that P450-1 is not a homolog of P450-4 with heterologous expression in *A. nidulans* (data not shown) in the later part of this discussion, which further validated our hypothesis. On the other hand, knocking out P450-3 abolished the production of **1**, without the emergence of any new metabolite compared with the wildtype. Therefore, it is likely that P450-3

is involved in the biosynthesis of **1**, but either the resulting metabolite(s) is/are too unstable to be observed via LC-MS or tethered to the carrier protein (ACP in this case) if it is involved before the release of the polyketide product from the HRPKS. We were also aware of the possibility of unintentionally expression knock-down of other the neighboring genes as reported.

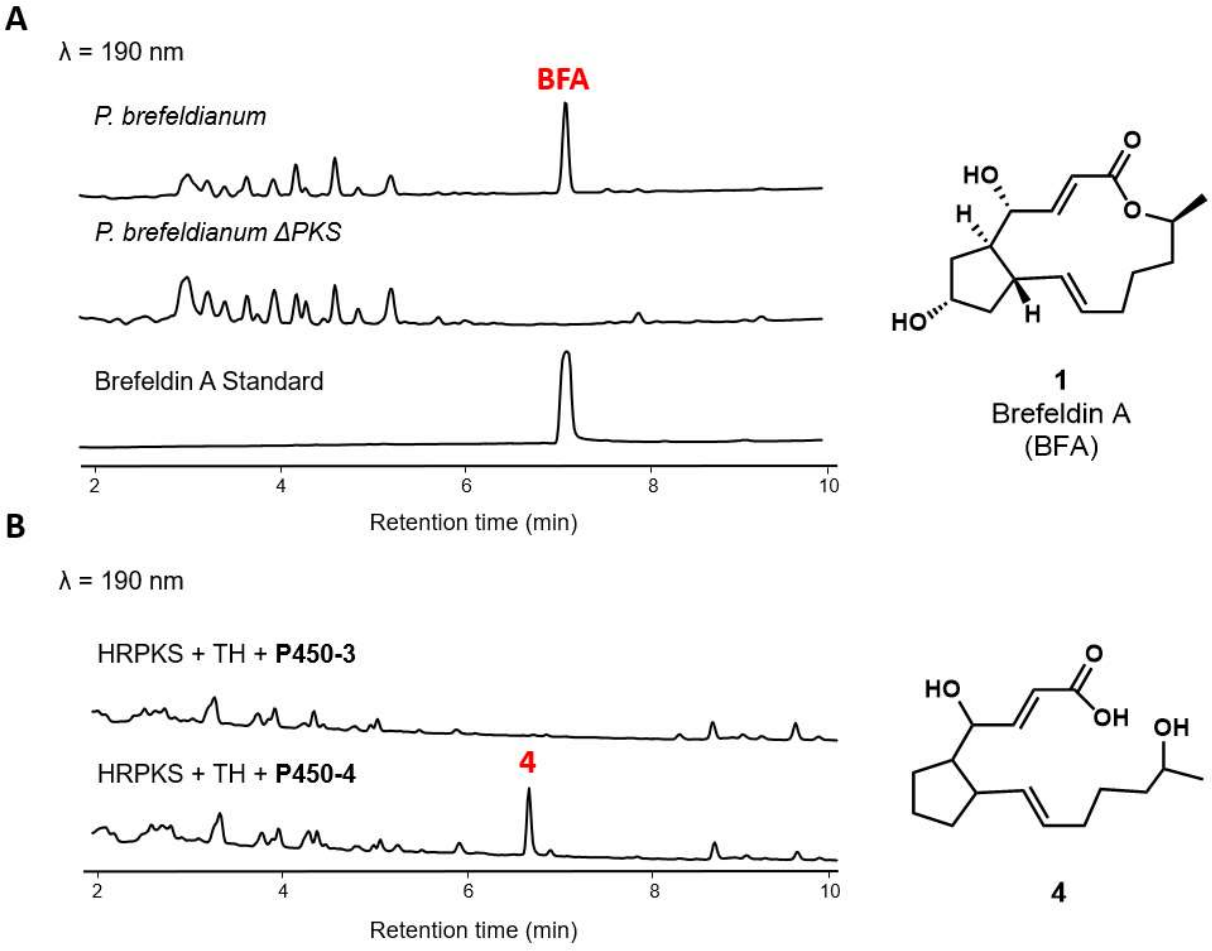


Figure 22. Characterization of the *bref* biosynthetic gene cluster. (A) Gene knockout with CRISPR. (B) heterologous expression in A. nidulans

6.3 Discovery of cyclopentane formation with heterologous expression in A. nidulans

Since the CRISPR knockout did not generate any new metabolites that are related to the biosynthetic pathway of **1**, we switched to heterologous expression in *A. nidulans*. Previous overexpression of the HRPKS and TH in *S. cerevisiae* generated the polyketide backbone **3**, as illustrated in **Figure 20**. However, subsequent expression of the P450(s) along with the HRPKS and TH did not yield any new metabolites. This could be because one or more P450s are not functional in yeast, due to the absence of compatible partnering enzymes such as the cytochrome P450 reductase (CPR). Moreover, when feeding the ring-opened intermediates (such as **3**) obtained from yeast expression, and other reduced forms of the HRPKS products to the HRPKS knockout strain of *P. brefeldianum*, the final product **1** is not observed (data not shown). This indicates that the polyketide products observed from the previous yeast expression could be shunt products instead of on-pathway intermediates.

Expression of the HRPKS and TH in *A. nidulans* did not yield **3** as reported from previous yeast expression.[174] However, expression of P450-4 along with the HRPKS and TH yield **4**, which correspond to the hydrolyzed **2** upon NMR characterization **(Figure 22)**. With the cyclopentane observed in **3**, it suggests that oxidation of the polyketide backbone catalyzed by P450-4 leads to a cascade of reactions that results in the formation of a cyclopentane, likely to be via a radical mechanism. Specifically, **Figure 23** illustrates the proposed mechanism for the formation of the cyclopentane ring. We propose that a C9 radical initiated by a P450-4 could perform the alkene addition to the C4-C5 double bond to form the cyclopentane moiety. This is followed by an oxygen rebound which hydroxylated C4 as in **4**. However, expressing P450-3 along with the HRPKS and TH, either with or without P450-4, did not generate any new metabolites. Considering the knockout results of P450-3 and comparative genomics with *B. brongniartii*, which suggests P450-3 involvement in the biosynthesis of **1**, we proposed P450-3 may strictly utilize **2** (cyclized **4**) as

a substrate, and catalyzed the hydrolyzation at the C7 position to complete the biosynthesis of **1**. This hypothesis is supported by previous labelling studies that suggested **2** is an on-pathway penultimate intermediate for the biosynthetic pathway of **1**.[167-168, 177]

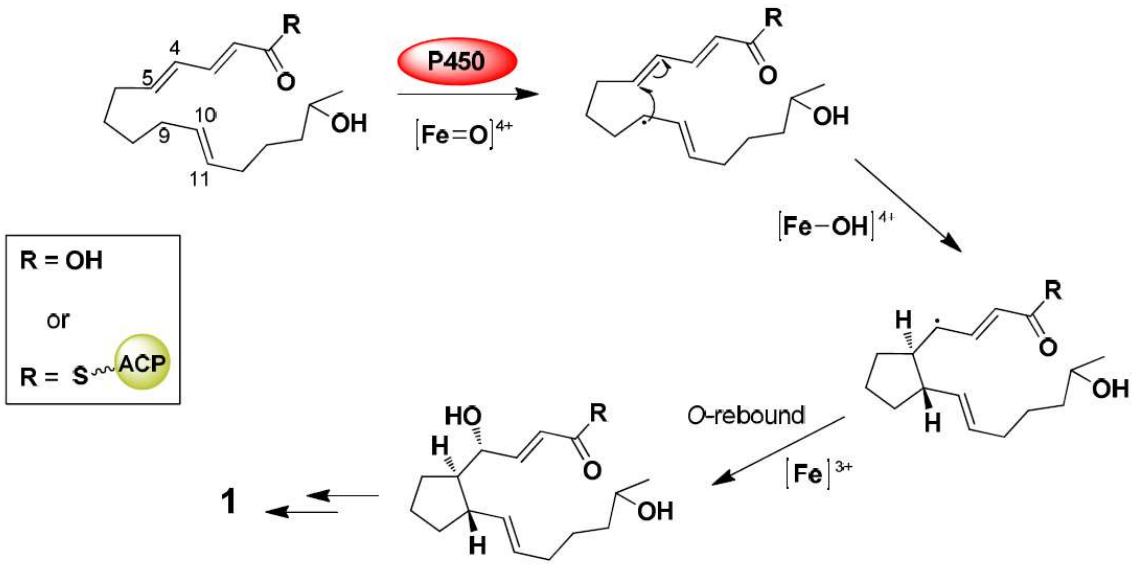


Figure 23. Proposed radical-mediated cyclization in BFA biosynthesis.[174]

To validate the hypothesis that the observation of **4** after *A. nidulans* expression was due to hydrolysis by heterologous host itself, we planned to carry out feeding studies with **2** into *A. nidulans*. However, 2 was not commercially available and synthetically challenging to be obtained due to the presence of multiple hydroxyl groups. Therefore, we decided to use 1 as a probe to validate our hypothesis. Feeding **1** (BFA) into *A. nidulans* under similar condition as heterologous expression yielded 5, which was suggested to be the hydrolyzed **1** based on the predicted mass from mass spectrometry (MS). Indeed, there was reports of a brefeldin A esterase that would hydrolyze **1** as a defense mechanism to protect the host from toxicity, and homologs of this esterase were observed in different microbes based on genetic alignment.[178] *A. nidulans* was likely to have such an esterase based on sequence alignment (AN3191), which may hydrolyze 2

unintentionally due to its structural similarity with 1. However, this hypothesis and the function of the proposed uncharacterized gene still have to be confirmed. Nevertheless, this finding can explain the inability to complete the biosynthetic studies in *A. nidulans*, for which an alternative host or *in vitro* method should be considered to examine the function of the remaining enzyme P450-3.

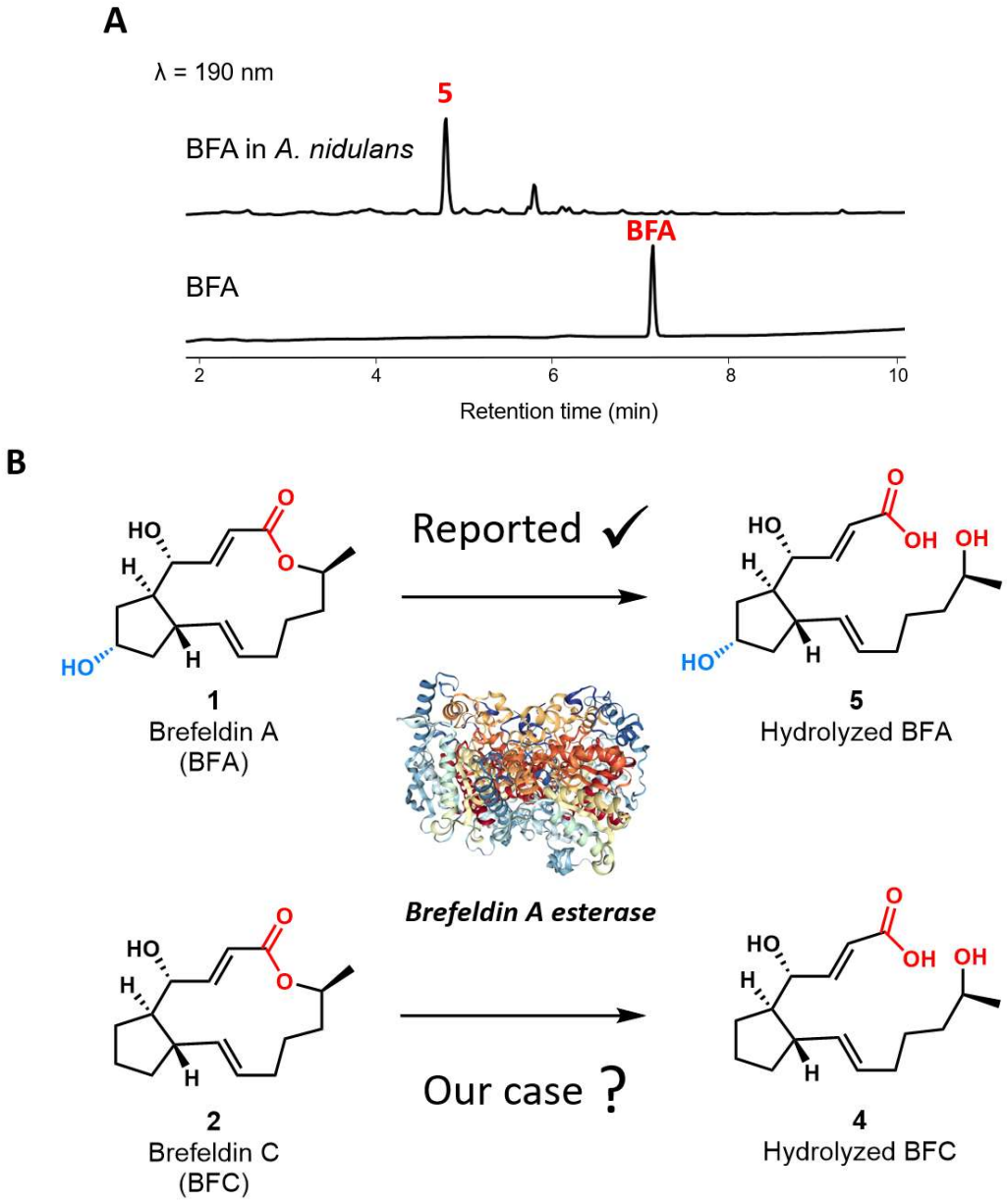


Figure 24. Endogenous resistance mechanism from *A. nidulans*. (A) Hydrolysis of BFA in *A. nidulans*. (B) Potential hydrolysis of **2** by brefeldin A esterase.

6.4 Discovery of potential fungal resistance mechanism for brefeldin A

As discussed in Chapter 2, for many biosynthetic clusters, there exists a resistance copy of the gene encoding the target of the toxin produced by the fungus. Often these resistance genes are located in the vicinity of the biosynthetic gene cluster for the concerted gene regulation, and they also have a substantial sequence identity to the housekeeping copy, which encodes for the protein with same function but sensitive to the natural product. Although there other resistance mechanisms exist, such as a specific transporter for the efflux of these toxins, the existence of these resistance gene can enhance the study of the biological application of these cytotoxic natural products and their mode of action. Specifically, with a collaboration with Hexagon Bio, we discovered a potential GEF that is resistantce to the brefeldin A, with the mutated Sec7 catalytic domain, which is the target of brefeldin A. Previous resistance studies suggested the mutation of M194L in ARNO from human or M699L in Gea1 from yeast can results with resistance to brefeldin A, and that methionine is important for interaction with BFA based on the crystal structures. Moreover, directed evolution of a bacterial strain also reported that a single mutation on the equivalent methionine to isoleucine also would enable the corresponding GEF to confer resistance to BFA. Investigation of the *Pencillium vulpinum* GEF next to the biosynthetic gene cluster revealed a 50% identity to the housekeeping copy as well as the methionine to leucine mutation. We therefore propose that this GEF is a resistance copy that allows *Pencillium vulpinum* to withstand the toxicity from BFA. To verify this hypothesis, we plan to test the efficacy of this enzyme both *in vivo* and *in vitro*. *In vivo* experiment includes the induction of the housekeeping and the potential resistance copy to yeast with an erg6 knockout, which was previously shown to facilitate sensitivity to BFA. *In vitro* experiments include the purification of the corresponding enzymes form *E. coli* or yeast in order to test the catalytic activity coupled with ARF1 in the presence of BFA. Upon the validation of the resistance mechanism, this would be the first reported

natural fungal resistance GEF and would enhance the resistance mechanism studies, with potential application to cancer therapeutic application.

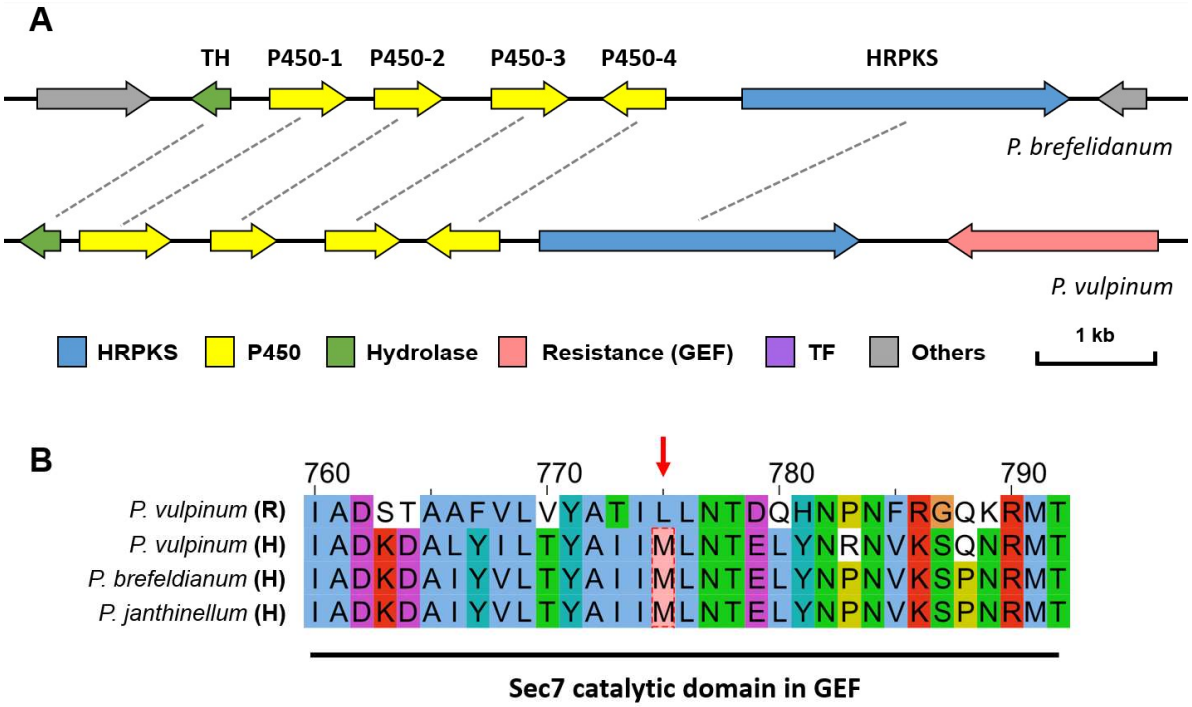


Figure 25. Potential resistance copies of Sec7 in *P. vulpinum*. (A) Gene cluster alignment between *P. brefeldianum* and *P. vulpinum*. (B) Gene alignment to identity potential mutation for resistance

6.5 Conclusion

To summarize, we confirmed the biosynthetic origin of a well-known natural product brefeldin A with CRISPR genome editing and characterized the stereospecific formation of a synthetically challenging cyclopentane catalyzed by a P450. This represents the first example of the enzymatic cyclopentane formation in a non-isoprenoid natural product. Moreover, we potentially discovered the first fungal BFA resistance copy of GEF via a targeted-genome mining approach, which enables the long-sought after investigation of brefeldin A inhibitory and resistance mechanisms.

7. SUMMARY

Natural products have long been a source of pharmaceuticals, while microbes are cell factories for many important applications. However, these two seemingly separate events have a common linkage, which is enzyme catalysis. As we began to examine microbe in a genetic level, the immense potential from their intricate metabolic network are slowly revealed, with the techniques such as genome mining. During my doctoral research, I have examined three selected natural products, motivated by their structural complexity and potent bioactivities. Biosynthetic studies for these molecules not only enhanced the understanding on how nature makes these elegant structures with the efficient use of enzyme catalysis, such as the AoiQ catalyzed regio-selective halogenation in dichlorodiaporthin and the P450 catalyzed cyclopentane formation in brefeldin A, it also elucidate the intricate cellular response and regulation involved in the production of these bioactive molecules.

Indeed, genome mining may have the potential to elevate natural product research to another golden era. Apart from discovering more bioactive compounds that can be served as pharmaceuticals, discoveries of novel enzymology can also collaborate with the advancement in synthetic chemistry to make synthesize new drug analogs in a chemo-enzymatic setting. Moreover, strategies such as targeted-genome mining can prioritize the biosynthetic gene cluster to be studied, which could result with directed drug-and-target relationship as well. With engineering, the workflow of genome mining can certainly be further optimized, whether *in vivo* or *in silico*. Overall, genome mining approach revolutionized the field of natural products. Other than the traditional natural product chemists, interdisciplinary opportunities are available for expertise from different fields to collaborate and continue optimizing the workflow that lead to more breakthroughs that are awaiting to come.

8. MATERIALS AND METHODS

8.1. Strains and culture conditions

Penicillium variabile HXQ-H-1 (wortmannilactone), *Aspergillus oryzae* RIB40 (dichlorodiaporthin), *Penicillium brefeldianum* ATCC 58665 (brefeldin A) were grown on PDA (potato dextrose agar, BD) at 28 °C for 3 days for cell proliferation or in liquid PDB medium (PDA medium without agar) for isolation of genomic DNA. *Penicillium variabile* HXQ-H-1 (wortmannilactone) was cultivated in PDB supplemented with 300 μM suberoylanilide hydroxamic acid (SAHA) at 28°C for 3 days for production of metabolites. *Penicillium brefeldianum* ATCC 58665 (brefeldin A) was cultivated in MEM (malt extract media) at 28°C for 3 days for production of metabolites.

Aspergillus nidulans ΔSTΔEM was grown at 28 °C on CD medium (1 L: 10 g of glucose, 50 mL of 20X nitrate salts, 1 mL of trace elements, pH 6.5, and 20 g/L of agar for solid cultivation); or in CD-ST media (1 L: 20 g of starch, 20 g of casamino acids, 50 mL of 20 × nitrate salts, 1 mL of trace elements, pH 6.5) for heterologous expression of the gene cluster, compound production, and RNA extraction. For preparation of 20X nitrate salts, 120 g of NaNO3, 10.4 g of KCl, 10.4 g of MgSO4•7H2O, 30.4 g of KH2PO4 were dissolved in 1 L of double distilled water. For preparation of the trace element solution, 2.20 g of ZnSO4•7H2O, 1.10 g of H3BO3, 0.50 g of MnCl2•4H2O, 0.16 g of FeSO4•7H2O, 0.16 g of CoCl2•5H2O, 0.16 g of CuSO4•5H2O, and 0.11 g of (NH4)6Mo7O24•4H2O were dissolved in 100 mL of double-distilled water, and the pH was adjusted to 6.5.

All *Escherichia coli* strains were cultured in LB media at 37 °C. Yeast strains were cultured in YPD media (yeast extract 1%, peptone 2%, glucose 2%) at 28 °C.

8.2 General DNA manipulation techniques

E. coli TOP10 and *E. coli* XL-1 were used for cloning, following standard recombinant DNA techniques. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs, NEB). PCR reactions were performed using Q5 High-Fidelity DNA Polymerase (NEB), Phusion High-Fidelity DNA Polymerase (NEB), and PFX High-Fidelity DNA Polymerase (Invitrogen). The gene-specific primers are listed in Table S2. PCR products were confirmed by DNA sequencing. *E. coli* BL21(DE3) (Novagen) was used for protein expression. The *Saccharomyces cerevisiae* BJ5464-NpgA *(MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL*) was used as the yeast host for *in vivo* homologous recombination to construct the *A. nidulans* expression plasmids.

For isolation of RNA from *A. nidulans* transformants, the strains were grown on CD agar for 6 days at 28 °C. The RNA extraction steps were performed using RiboPure™ Yeast RNA Isolation Kit (Ambion) following the manufacturer's instructions. Residual genomic DNA in the extracts was digested by DNase I (2 U/μL) (Invitrogen) at 37 °C for 4 hours. SuperScript III First-Strand Synthesis System (Invitrogen) was used for cDNA synthesis with Oligo-dT primers following directions from the user manual.

8.3 Heterologous expression in A. nidulans

To construct plasmids for heterologous expression in *A. nidulans* ΔSTΔEM, the plasmids pYTU, pYTP, and pYTR with auxotrophic markers for uracil (*pyrG*), pyridoxine (*pyroA*), and riboflavin (*riboB*), respectively, were used as backbones to insert genes. Genes from the targeted cluster and their native terminators were amplified by PCR with overhang primers using the genomic DNA (gDNA) of producing fungus as the template. gpdA promoters from *A. nidulans*, *Penicillium oxalicum* (POgpdA), and *Penicillium expansum* (PEgpdA) were amplified by PCR.

pYTP and pYTR were digested with *PacI*/*NotI* and pYTU was digested with *PshAI*/*NotI*. The overlapping DNA fragments and their corresponding digested vectors were co-transformed into *S. cerevisiae* BJ5464-NpgA to assemble the expression plasmids *in vivo* by yeast homologous recombination. The plasmids were extracted from yeast using Zymoprep™ Yeast Plasmid Miniprep I (Zymo Inc. USA), and transformed into *E. coli* TOP10 by electroporation to isolate single plasmids. After extraction from *E. coli*, the plasmids were sequenced to confirm correct assembly.

To prepare protoplasts*, A. nidulans* was initially grown on CD agar plates supplemented with 10 mM of uridine, 5 mM of uracil, 0.5 µg/ml of pyridoxine HCl and 2.5 µg/ml of riboflavin at 30 °C for 5 days. Fresh spores of *A. nidulans* were inoculated into 50 mL of liquid CD media in a 250 mL flask and germinated at 30 °C, 250 rpm for 16 h. Mycelia were harvested by centrifugation at 3,500 rpm for 10 min and washed with 10 mL of osmotic buffer (1.2 M of MgSO4, 10 mM of sodium phosphate, pH 5.8). The mycelia were transferred into 10 mL of osmotic buffer containing 30 mg of lysing enzymes from *Trichoderma* and 20 mg of Yatalase in a 125 mL flask. The cells were digested for 16 hours at 30 °C, 80 rpm. Cells were collected in a 30 mL Corex tube and overlaid gently by 10 mL of trapping buffer (0.6 M of sorbitol, 0.1 M of Tris-HCl, pH 7.0). After centrifugation at 3,500 rpm for 15 min at 4°C, protoplasts were collected in the interface of the two buffers. The protoplasts were transferred to a sterile 15 mL falcon tube and washed with 10 mL of STC buffer (1.2 M of sorbitol, 10 mM of $CaCl_2$, 10 mM of Tris-HCl, pH 7.5). The protoplasts were then resuspended in 1 mL of STC buffer.

For each transformation, plasmids were added to 100 µl of the *A. nidulans* protoplast suspension prepared above, and the mixture was incubated for 60 min on ice. Then 600 µl of PEG solution (60% PEG, 50 mM of calcium chloride, and 50 mM of Tris-HCl, pH 7.5) was added to the protoplast mixture, followed by additional incubation at room temperature for 20 min. The

mixture was spread on the regeneration medium (CD solid medium with 1.2 M of sorbitol and the appropriate supplements: 10 mM of uridine, 5 mM of uracil, 0.5 μg/mL of pyridoxine HCl, and/or 2.5 μg/mL of riboflavin according to the markers in the transformed plasmids) and incubated at 37 °C for 2-3 days.

8.4 Chemical analysis and compound isolation

For small scale metabolite analysis in *A. nidulans*, transformants were grown on CD agar for 5-6 days at 28 °C and then extracted with acetone. *A. nidulans* strains expressing the gene(s) of interest were grown on CD-ST agar for 3 days at 28 °C and then extracted with acetone. For small scale analysis of in yeast, *S. cerevisiae* expressing the gene(s) of interest was inoculated in 1 mL of dropout media for 24 hours. 100μL of starter culture was used to inoculate 3 mL of YPD. The cells were grown at 28 °C, 250 rpm for 48 hours and extracted with an organic phase consisting of 75% EtOAc and 25% acetone. The organic phases were dried and dissolved in methanol for analysis. LC-MS analyses were performed on a Shimadzu 2020 EV LC-MS with a reverse-phase column (Phenomenex Kinetex, C18, 1.7 μm, 100 Å, 2.1 × 100 mm) using positive- and negative-mode electrospray ionization with a linear gradient of 5-95% acetonitrile-H_2O (containing 0.1% formic acid) in 15 min followed by 95% acetonitrile for 3 min with a flow rate of 0.3 ml/min.

For compound isolation, transformants of *A. nidulans* strains were grown in 4 L of liquid CD-ST media for 3-4 days at 28 °C, 250 rpm. Cell pellets from the large-scale cultures were extracted with acetone and the supernatant was extracted with ethyl acetate. After evaporation of the organic phases, the crude extracts were absorbed with 3 g of Celite, which was purified with the CombiFlash system (Teledyne) using reverse phase gradient elution with water (A) and acetonitrile (B) (0-5 min 10% B; 5-45 min 10-100% B; 45-50 min 100% B). Fractions containing

the target compounds were combined and used for further purification by HPLC with a semi-preparative reverse-phase column (Phenomenex Kinetics, C18, 5 μm, 100 Å, 10 × 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-8 min 10-30% B; 8-15 min 50% B; 15-16 min 10% B.

NMR spectra were obtained with a Bruker AV500 spectrometer with a 5-mm dual cryoprobe at the UCLA Molecular Instrumentation Center. (1H 500 MHz, 13C 125 MHz). X-ray crystallography were conducted in the UCLA Chemical and Biochemical Department. High resolution mass spectra were obtained from Thermo Fisher Scientific Exactive Plus with IonSense ID-CUBE DART source at the UCLA Molecular Instrumentation Center. Optical rotations were measured on a Rudolph Research Analytical Autopol III Automatic Polarimeter.

8.5 Expression and purification of AoiQ from S. cerevisiae and in vitro assays

AoiQ was cloned into the yeast expression vector XW55 with a *C*-terminal His-tag. In this plasmid, expression of *aoiQ* was under the control of the *ADH2* promoter. Overexpression and subsequent protein purification of AoiQ were performed as follows: the yeast strain BJ5464-NpgA harboring the expression plasmid was grown overnight in 2 x 5 mL cultures of uracil dropout medium at 28 °C. The starter cultures were used to inoculate 2 × 1 L of YPD medium, which were shaken at 28 °C, 250 rpm for 48h. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM of Tris-HCl, 500 mM of NaCl, pH 8.0), and lysed on ice by sonication. The lysate was centrifuged at 15,000 *g* for 30 min at 4 °C to remove the cellular debris. Purification of the recombinant His6-tagged AoiQ using affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified AoiQ was concentrated and exchanged into storage buffer (50 mM of Tris-HCl, 100 mM of NaCl, 10% glycerol, pH 8.0) with

Centriprep filters (Amicon). The purified AoiQ was analyzed by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to calculate protein concentration.

8.6 Expression and purification of DiaC from E. coli BL21(DE3)

The *diaC* gene was amplified with overhang primers from the cDNA of *A. nidulans* transformant. The PCR product and expression vector pET28a were digested with *NdeI*/*XhoI* and ligated with T4 ligase (Invitrogen). Overexpression and subsequent protein purification of DiaC were performed as follows: BL21(DE3) harboring the expression plasmid was grown overnight in 2 x 5 mL of LB medium with 50 μg/mL of kanamycin at 37 °C. The starter cultures were used to inoculate 2 × 1 L of fresh LB medium and shaken at 37 °C until the optical density at 600 nm (OD600) reached 0.6. Then expression of the gene was induced with 0.1 mM of isopropylthio-*β*-D-galactoside (IPTG) at 16 °C. After 20 hours, cells were harvested by centrifugation, resuspended in lysis buffer (50 mM of Tris-HCl, 500 mM of NaCl, pH 8.0), and lysed on ice by sonication. The lysate was centrifuged at 15,000 *g* for 30 min at 4 °C to remove cellular debris. Purification of the recombinant His6-tagged DiaC using affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified DiaC was concentrated and exchanged into storage buffer (50 mM of Tris-HCl, 100 mM of NaCl, pH 8.0) by using Centriprep filters (Amicon). Purified DiaC was analyzed by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to measure protein concentration. Aliquots of purified DiaC were flash frozen and stored at -80 °C.

8.7 In vitro assays for AoiQ and DiaC (need editing)

Enzyme assays of AoiQ and DiaC were performed in 50 mM potassium phosphate (pH 6.5) buffer with a final volume of 100 μL. The assay contained 0.1 mM of $MgCl_2$, 0.1 mM of $MnCl_2$, 1mM of FAD, 1mM of SAM, 1mM of NADH/NADPH, 150mM of NaCl and 10 μM of recombinant

AoiQ/DiaC (intact or boiled). Enzyme assays of DiaC were performed in 50 mM potassium phosphate (pH 6.5) buffer with a final volume of 100 µL. The assay contained 1mM of NADH/NADPH, and 10 µM of recombinant DiaC (intact or boiled). The reactions were incubated at 25 °C for 2 hours and extracted with 100 µL of EtOAc. Samples were centrifuged and the organic layers were analyzed by GC-MS as described previously.

8.8 Biotransformation assays in S. cerevisiae

To verify the function of enzyme(s) of interest by yeast biotransformation, *Saccharomyces cerevisiae* transformed with expression plasmid was inoculated in 1 mL of selective dropout media for 24 hours. 3 mL of YPD was inoculated with 100 µL of starter culture and grown for 24 hours. Substrate was fed to 500 µL of yeast culture to a final concentration of 400 µM and the cultures were shaken at 250 rpm at 28 °C for 12 hours. The cultures were extracted two times with 500 µL of an organic phase consisting of 75% EtOAc and 25% acetone. The organic layers were dried and dissolved in methanol for LC-MS analysis.

9. APPENDICES

9.1 Supplementary information for Section 4

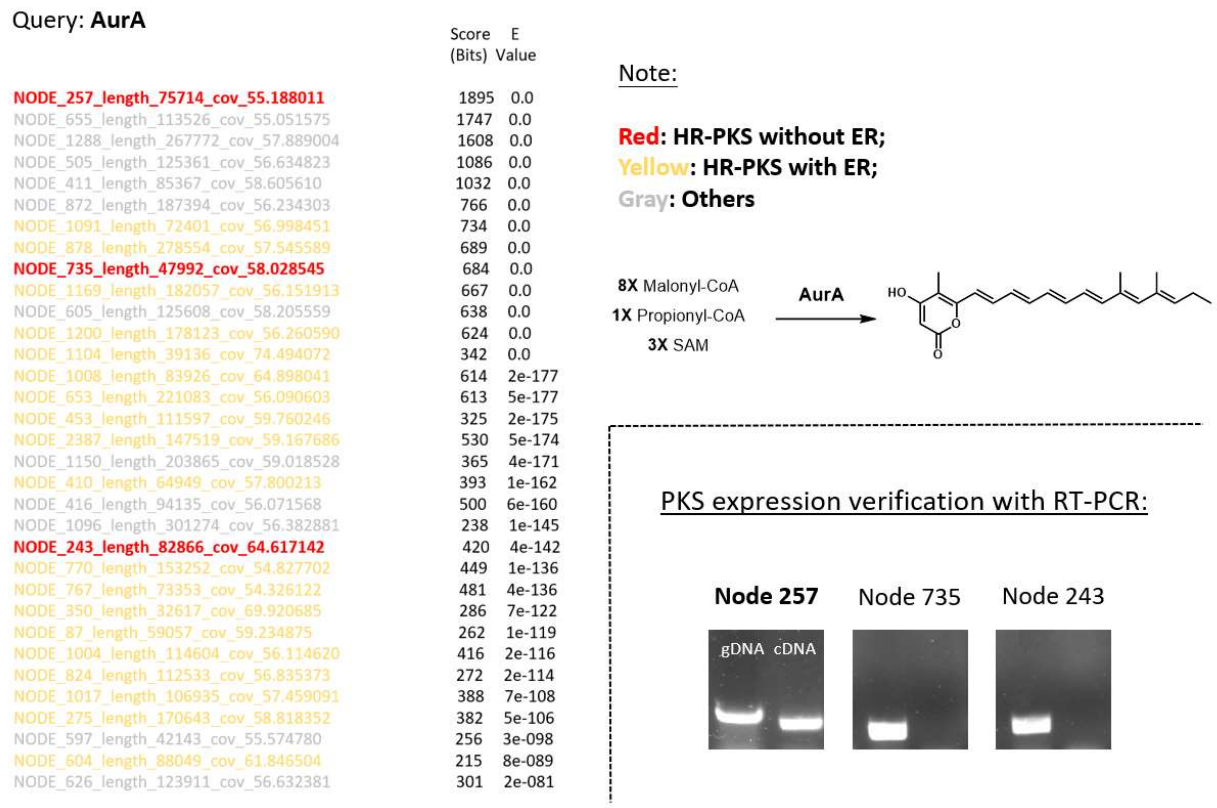
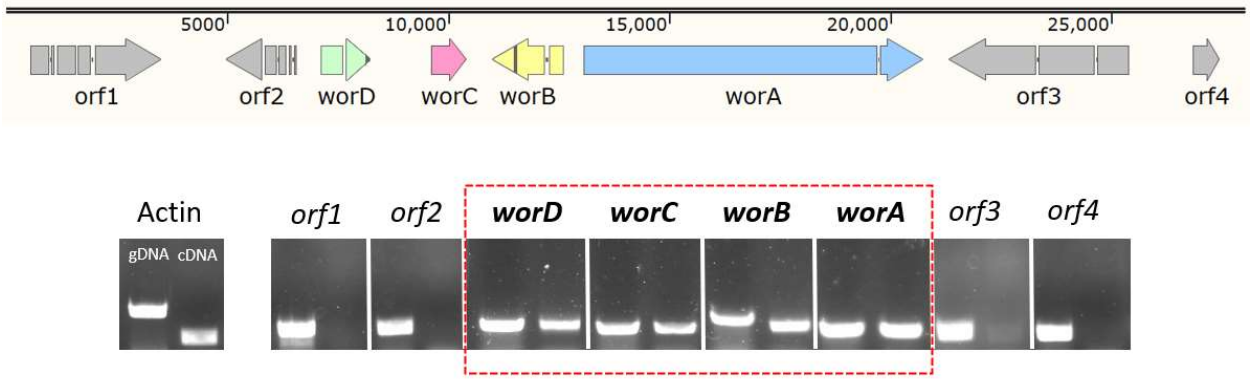


Figure S1. Bioinformatics and transcriptional analyses for PKSs in *P. variabile HXQ-H-1* (with HRPKS AurA from aurovertins' biosynthesis as query)



Gene	Size (bp/aa)	Accession #	Identity (%)	Putative function	*E*- value
orf1	2961/903	Q9J4Z5.1	29	Ankyrin repeat protein FPV245	0.41
orf2	1601/459	Q9P6J0.1	29	Uncharacterized transporter C1683.12	2.0E-62
worD	1122/329	P11435.3	39	Carboxyvinyl-carboxyphosphonate phosphorylmutase	2.0E-49
worC	798/265	N/A	N/A	N/A	N/A
worB	1608/485	Q9C447.1	34	FAD-dependent monooxygenase	2.0E-81
worA	7694/2545	Q0C9L7.1	43	Highly reducing polyketide synthase	0
orf3	4082/1325	O59810.2	26	Vigilin	6.0E-122
orf4	606/201	Q12436.1	27	Zinc-regulated transporter	3.9

Figure S2. Transcriptional analyses and annotations for the *wor* biosynthetic gene cluster

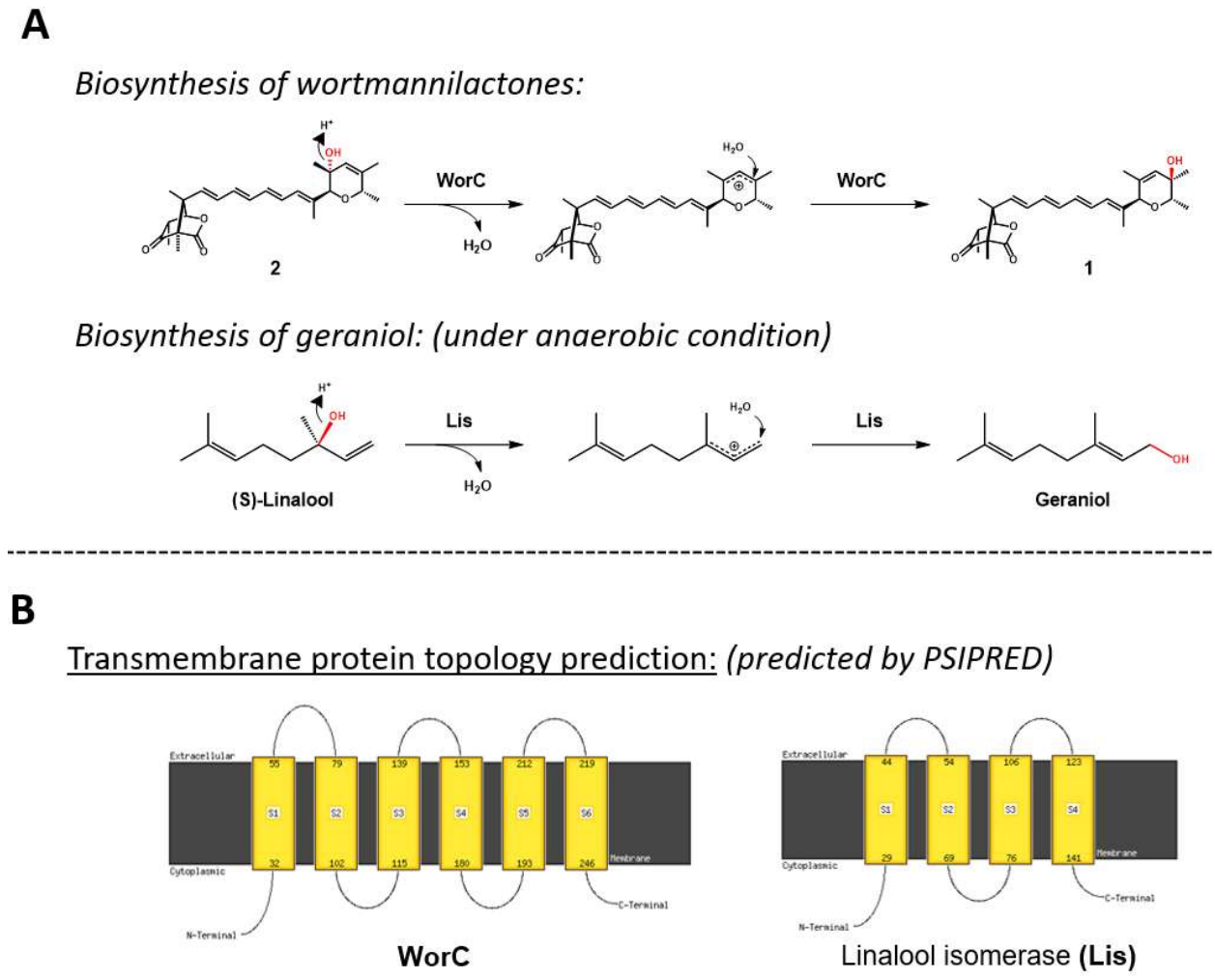


Figure S3. Analogy between WorC and Lis. (A) Proposed mechanism for WorC and Lis. (B) Transmembrane topology prediction of WorC and Lis with PSIPRED.

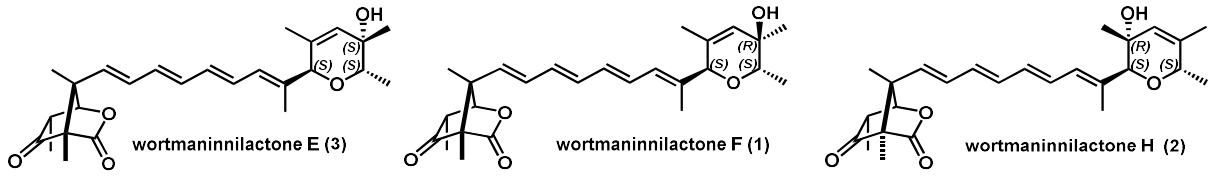


Table S1. NMR data for **1**, **2** and **3** in CD_3OD (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

position	**3**		**1**		**2**	
	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C
2	-	173.4	-	173.5	-	173.5
3	-	70.8	-	70.8	-	70.8
4	-	209.5	-	209.6	-	209.6
5	2.81, dq (7.0, 2.5)	44.5	2.88, dq (7.0, 2.0)	44.5	2.88, dq (7.0, 2.5)	44.5
6	5.03, d (2.5)	86.3	5.03, d (2.0)	86.4	5.02, d (2.0)	86.4
7	-	59.5	-	59.5	-	59.5
8	5.79, d (15.0)	131.3	5.78, d (15.0)	131.3	5.75, d (15.0)	130.6
9	6.46, dd (15.0,	135.5	6.46, dd (15.0,	135.6	6.44, dd (15.0,	136.0
10	6.31, dd (15.0,	133.6	6.32, dd (15.0,	133.5	6.28, dd (15.0,	133.2
11	6.44, dd (15.0,	134.9	6.44, dd (15.0,	134.9	6.44, dd (15.0,	135.1
12	6.31, dd (15.0,	134.8	6.29, dd (15.0,	134.8	6.28, dd (15.0,	132.5
13	6.58, dd (15.0,	130.4	6.58, dd (15.0,	130.8	6.61, dd (15.0,	131.1
14	5.92, d (11.0)	130.4	6.04, d (11.0)	130.6	6.21, d (11.0)	128.0
15		136.0		136.4	-	138.1
16	4.34, s	81.5	4.29, s	81.9	4.05, s	78.9
17	-	135.6	-	134.3	-	71.2
18	5.61, d (1.5)	130.9	5.58, (br s)	132.4	5.36, d (1.5)	130.8
19	-	68.1	-	69.8	-	137.3
20	3.55, q (6.5)	73.2	3.65, q (6.5)	72.6	4.18, q (7.0)	74.0
21	1.08, s	5.0	1.08, s	5.0	1.08, s	5.0
22	1.20, d (7.0)	11.7	1.19, d (7.0)	11.7	1.19, d (7.0)	11.7
23	1.17, s	16.9	1.16, s	16.9	1.16, s	16.9
24	1.87, s	15.7	1.88, s	15.8	1.90, brs	16.0
25	1.62, s	20.1	1.58, s	19.9	1.03, s	22.8
26	1.15, s	25.6	1.11, s	22.1	1.64, s	19.2
27	1.16, d (6.0)	14.4	1.08, d (6.0)	14.7	1.30, d (6.0)	17.5

The NMR data of compounds **1**-**3** are identical to the data reported in *Helv. Chim. Acta* **2009**, 92, 567.

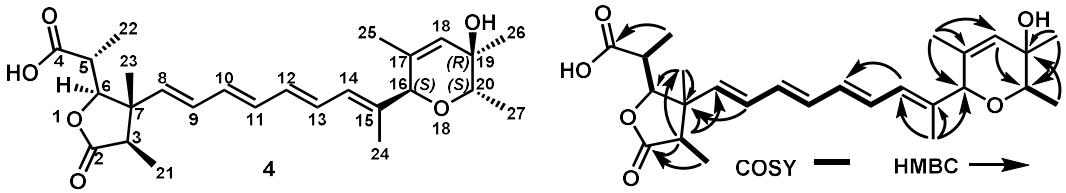
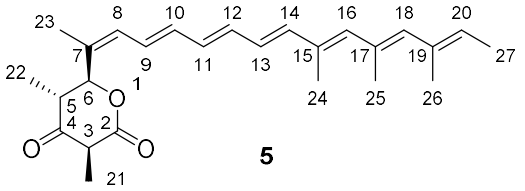


Table S2. NMR data for **4** in CD_3OD (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

No	δ_C	δ_H (J in Hz)	NOE correlations
2	179.5	-	
3	48.1	2.80, q (7.0)	H-6, H-8
4	179.5	-	
5	43.4	2.65, dq (11.0, 7.0)	H_3-23
6	87.9	4.41, d (10.5)	H-8
7	49.5	-	
8	137.7	5.82, d (14.0)	
9	132.4	6.29, dd (15.0, 11.0)*	
10	133.6	6.32, dd (15.0, 11.0)*	
11	134.6	6.40 dd (15.0, 11.0)	
12	134.9	6.29, dd (15.0, 11.0)*	
13	130.1	6.56, dd (15.0, 11.0)	
14	130.8	6.03, d (11.0)	
15	136.1	-	
16	81.9	4.28, s	H_3-26, H_3-27
17	134.3	-	
18	132.2	5.58, brs	
19	69.8	-	
20	72.6	3.65, q (6.5)	
21	7.7	0.97, d (7.0)	
22	14.7	1.08, d (6.5)	
23	12.0	1.05, s	
24	15.8	1.88, brs	
25	19.8	1.58, s	
26	22.1	1.11, s	
27	14.7	1.11, d (6.0)	

HRMS: found *m/z* 427.2491 $[M + H\text{-}H_2O]^+$, Cacld. for $C_{26}H_{35}O_5^+$ 409.2484

Optical rotation: $[\alpha]^{25}_D$ = +178.15 (*c* = 0.5, CH_3OH)



5

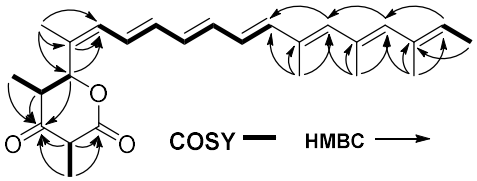


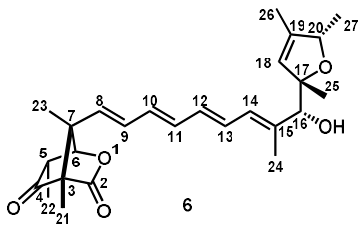
Table S3. NMR data for **5** in $CDCl_3$ (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

No	δ_C	δ_H (*J* in Hz)	Key NOE correlations
2	169.8	-	
3	50.4	3.55, q (6.5)	
4	204.5	-	
5	44.6	2.53, dq (10.5, 7.0)	H_3-23
6	85.8	4.75, d (11.0)	H-3, H_3-22
7	130.0	-	
8	132.4	6.19, d (10.0)	
9	126.6	6.42, m	
10	127.5$^{\Delta}$	6.30-6.36, m*	
11	131.6$^{\Delta}$	6.30-6.36, m*	
12	135.7$^{\Delta}$	6.30-6.36, m*	
13	136.2$^{\Delta}$	6.30-6.36, m*	
14	140.5	6.34, m	
15	133.7	-	
16	138.8	6.03, s	H-14, H-18
17	132.3	-	
18	136.2	5.88, s	H_3-24
19	133.7	-	
20	125.9	5.48, q (7.0)	H-18, H_3-25
21	8.4	1.39, d (6.5)	
22	11.4	1.08, d (7.0)	
23	12.0	1.84, s	
24	14.3	1.98, s	
25	19.2	1.96, s	
26	16.9	1.78, s	
27	14.1	1.72, d (7.0)	

$^{\Delta}$ interchangeable. * overlapped.

HRMS: found *m/z* 395.2588 $[M + H]^+$, Cacld. for $C_{26}H_{35}O_3^+$ 395.2586;

Optical rotation: $[\alpha]^{25}_D$ = +16.62 (*c* = 0.3, $CHCl_3$)



6

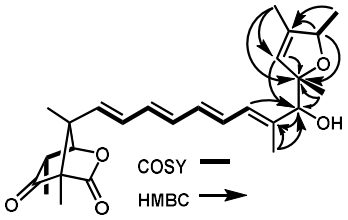


Table S4. NMR data for **6** in CD_3OD (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

No	δ_C	δ_H (*J* in Hz)	Key NOE correlations
2	173.5	-	
3	70.8	-	
4	209.6	-	
5	44.5	2.89, dq (7.0, 2.5)	
6	86.4	5.02, d (3.5)	
7	59.5	-	
8	130.7	5.76, d (15.0)	
9	135.0	6.45, dd (15.0, 11.0)	
10	133.4	6.27, dd (15.0, 11.0)	
11	135.9	6.43, dd (15.0, 11.0)	
12	132.7	6.27, dd (15.0, 11.0)	
13	131.0	6.57, dd (15.0, 11.0)	
14	128.7	6.16, d (11.0)	
15	140.5	-	
16	83.1	3.94, s	H_3-25
17	93.1	-	
18	127.3	5.48, s	
19	140.8	-	
20	84.3	4.70, q (6.0)	H_3-25
21	5.0	1.08, s	
22	11.7	1.19, d (7.0)	
23	16.9	1.16, s	
24	15.5	1.85, s	
25	23.0	1.15, s	
26	12.3	1.69, s	
27	21.0	1.26, d (6.0)	

HRMS: found *m/z* 409.2363 $[M + H\text{-}H_2O]^+$, Cacld. for $C_{26}H_{33}O_4^+$ 409.2378;

Optical rotation: $[\alpha]^{25}_D$ = +21.0 (*c* = 0.2, CH_3OH)

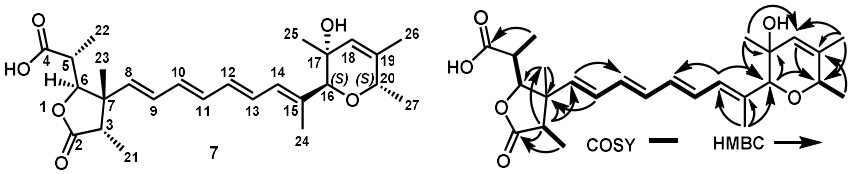


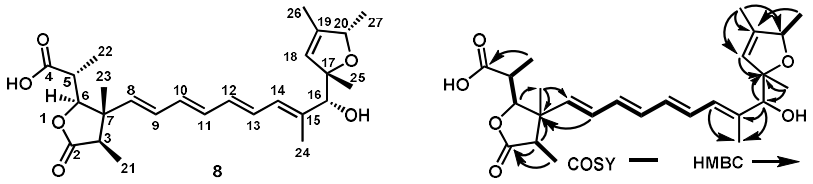
Table S5. NMR data for **7** in CD_3OD (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

No	δ_C	δ_H (*J* in Hz)	NOE correlations
2	179.8	-	
3	48.3	2.77, q (7.5)	H-6, H-8
4	181.1$^{\Delta}$	-	
5	44.5	2.60, m	H_3-23
6	88.6	4.41, d (10.0)	H-8
7	49.6	-	
8	137.6	5.79, d (14.0)	
9	132.8	6.22-6.30, m*	
10	132.1	6.22-6.30, m*	
11	134.8	6.37, m	
12	133.4	6.22-6.30, m*	
13	130.5	6.58, dd (14.5, 11.5)	
14	128.1	6.19, d (11.5)	H_3-25
15	137.5	-	
16	78.9	4.05, s	H_3-27
17	71.2	-	
18	130.8	5.35, brs	
19	137.3	-	
20	74.0	4.18, q (6.5)	
21	7.7	0.96, d (7.5)	
22	15.0	1.06, d (6.5)	
23	12.1	1.03, s	
24	16.0	1.89, brs	
25	22.8	1.03, s	
26	19.2	1.64, s	
27	17.5	1.30, d (6.5)	

$^{\Delta}$observed in HMBC spectrum. * overlapped.

HRMS: found *m/z* 427.2483 $[M + H\text{-}H_2O]^+$, Cacld. for $C_{26}H_{35}O_5^+$ 409.2484

Optical rotation: $[\alpha]^{25}_D$ = +15.4 (*c* = 1.0, CH_3OH)



8

Table S6. NMR data for **8** in CD_3OD (500 and 125 MHz for ^1H and ^{13}C NMR, resp)

No	δ_C	δ_H (*J* in Hz)	NOE correlations
2	179.3	-	
3	48.2	2.79, q (7.0)	H-6, H-8
4	179.7	-	
5	44.3	2.60, dq (11.0, 7.0)	H_3-23
6	88.4	4.40, d (11.0)	H-8
7	49.3	-	
8	137.5	5.79, d (14.0)	
9	132.9	6.28, dd (14.5, 11.5)*	
10	133.6	6.28, dd (14.5, 11.5)*	
11	134.7	6.37 dd (14.5, 11.5)	
12	132.1	6.28, dd (14.5, 11.5)*	
13	130.4	6.54, dd (14.5, 11.5)	
14	128.8	6.15, d (11.0)	
15	140.0	-	
16	83.1	3.94, s	
17	93.1	-	
18	127.3	5.48, brs	
19	140.8	-	
20	84.2	4.69, q (6.0)	H_3-25
21	7.7	0.96, d (7.0)	
22	14.9	1.07, d (7.0)	
23	12.0	1.04, s	
24	15.4	1.84, brs	
25	23.1	1.15, s	
26	12.3	1.69, brs	
27	21.0	1.26, d (6.5)	

* overlapped

HRMS: found *m/z* 427.2481 $[M + H\text{-}H_2O]^+$, Cacld. for $C_{26}H_{35}O_5^+$ 409.2484

Optical rotation: $[\alpha]^{25}_D$ = +26.5 (*c* = 0.7, CH_3OH)

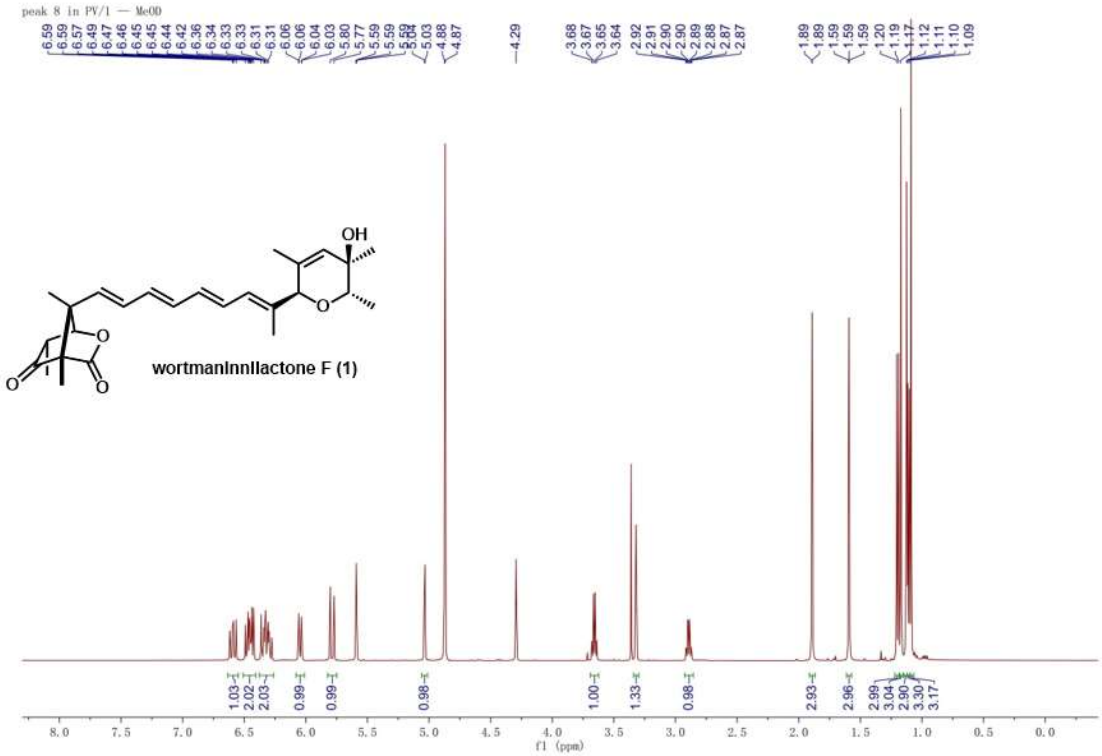


Figure S4. ^{1}H NMR spectrum of compound **1** in CD_3OD.

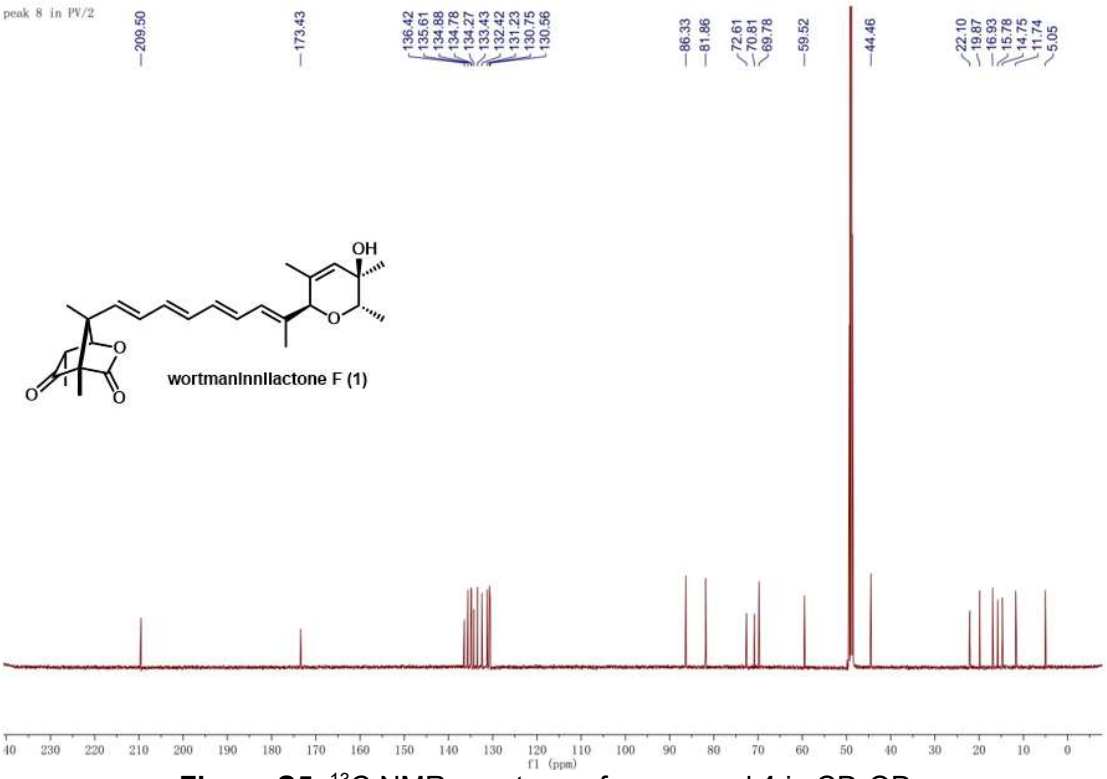


Figure S5. ^{13}C NMR spectrum of compound **1** in CD_3OD.

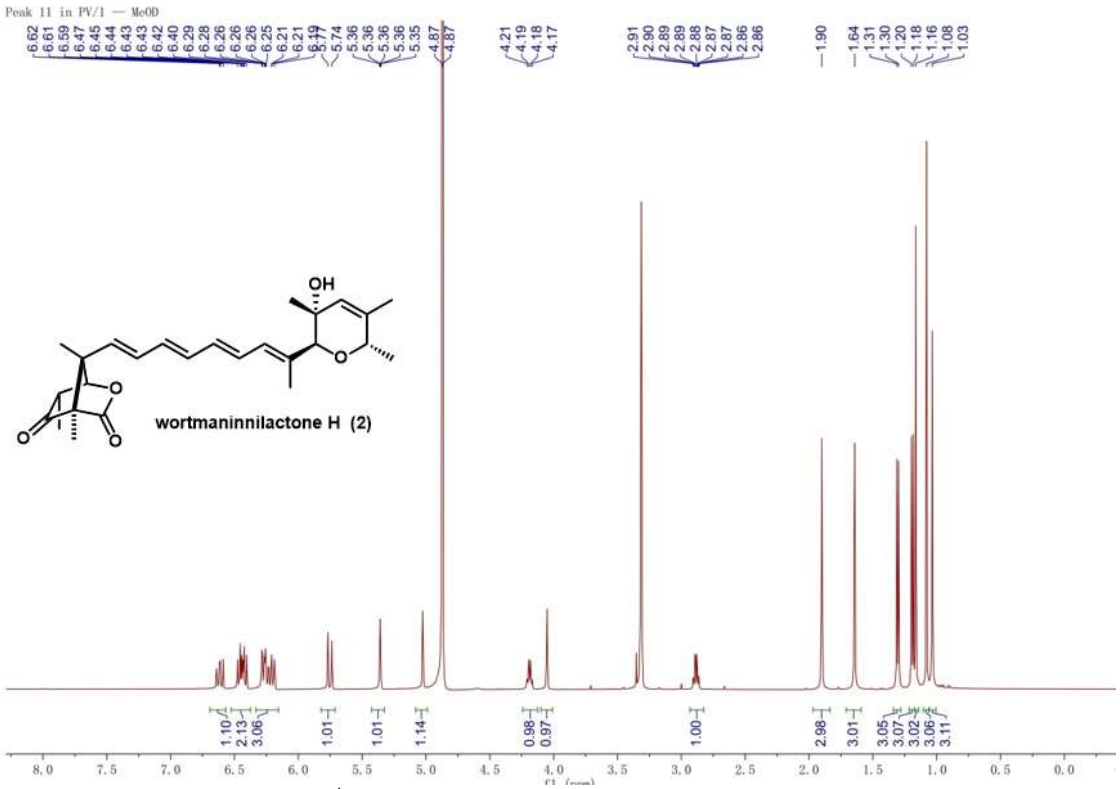


Figure S6. ^{1}H NMR spectrum of compound **2** in CD_3OD.

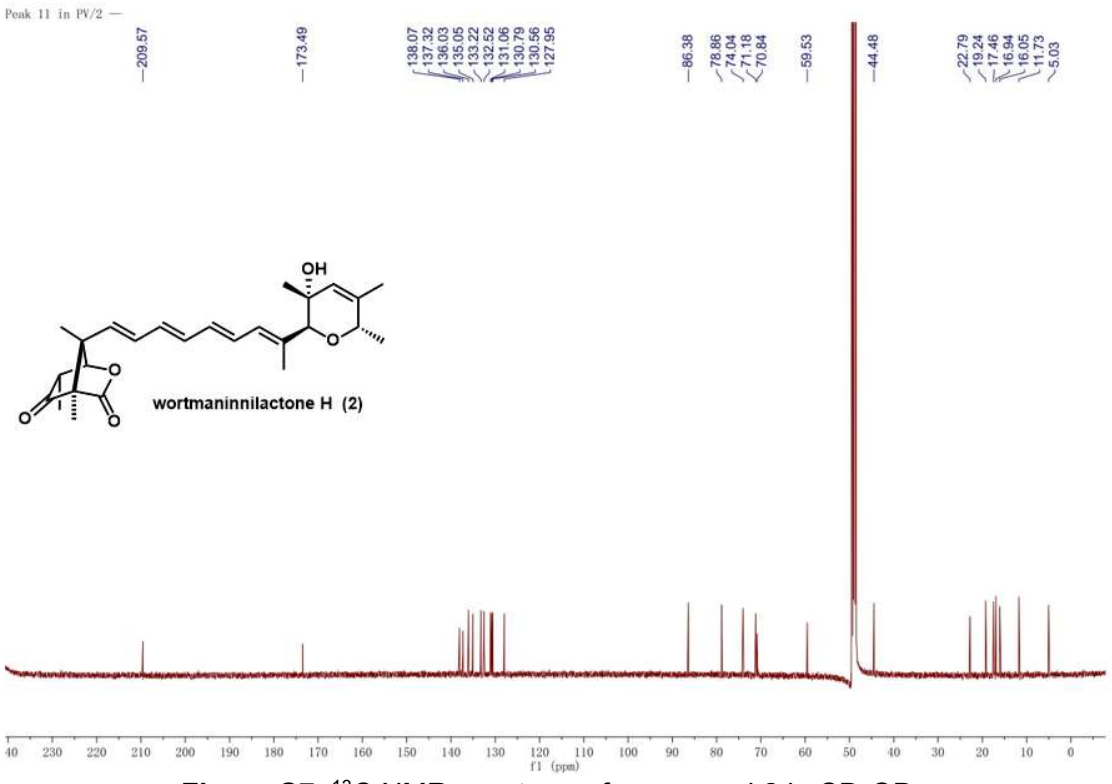


Figure S7. ^{13}C NMR spectrum of compound **2** in CD_3OD.

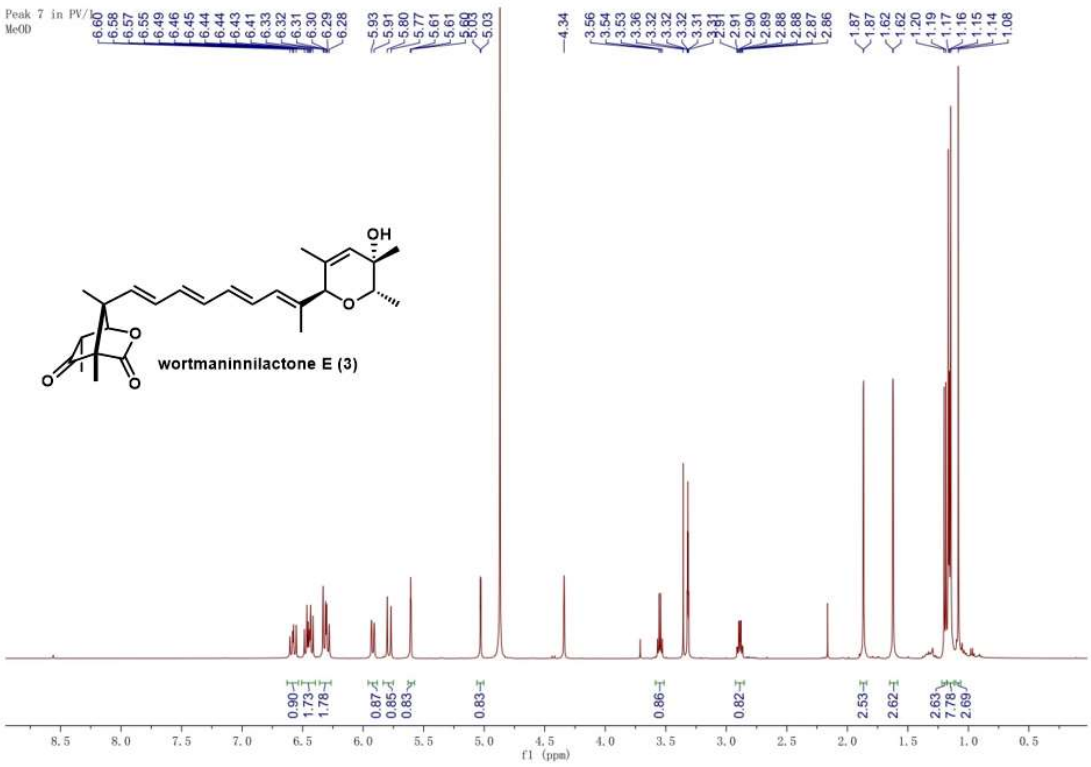


Figure S8. ^{1}H NMR spectrum of compound **3** in CD_3OD.

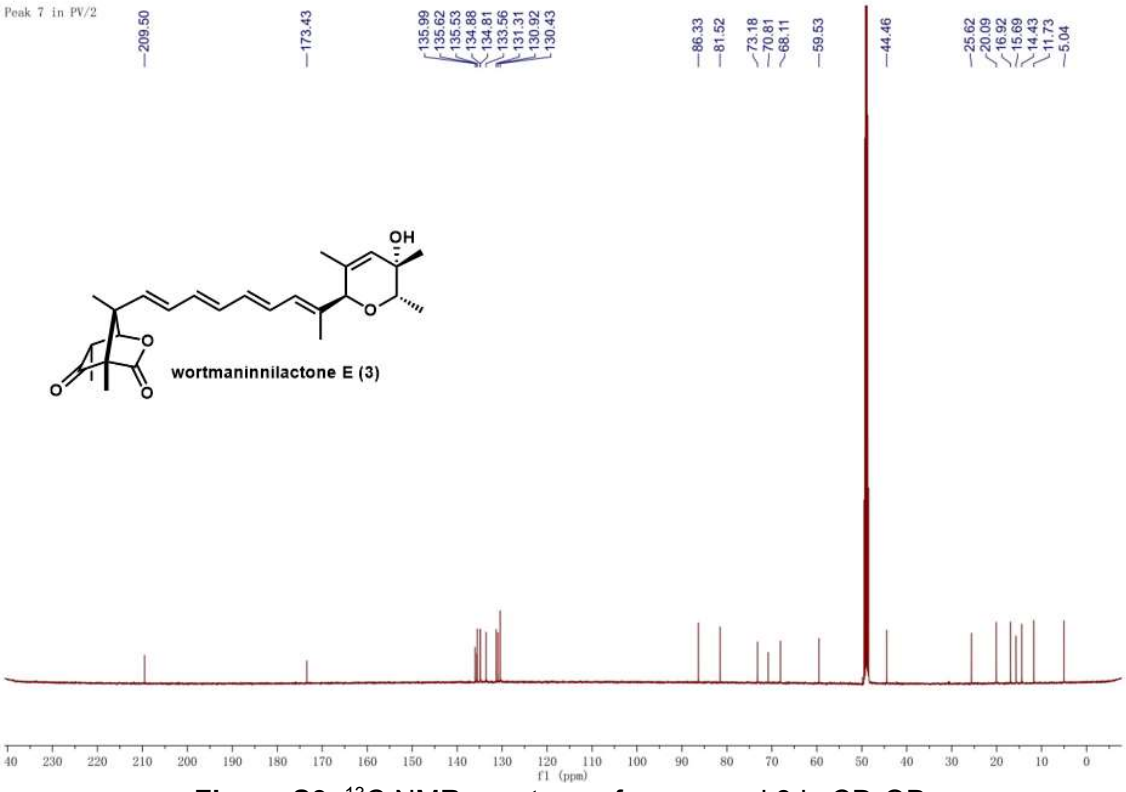


Figure S9. ^{13}C NMR spectrum of compound **3** in CD_3OD.

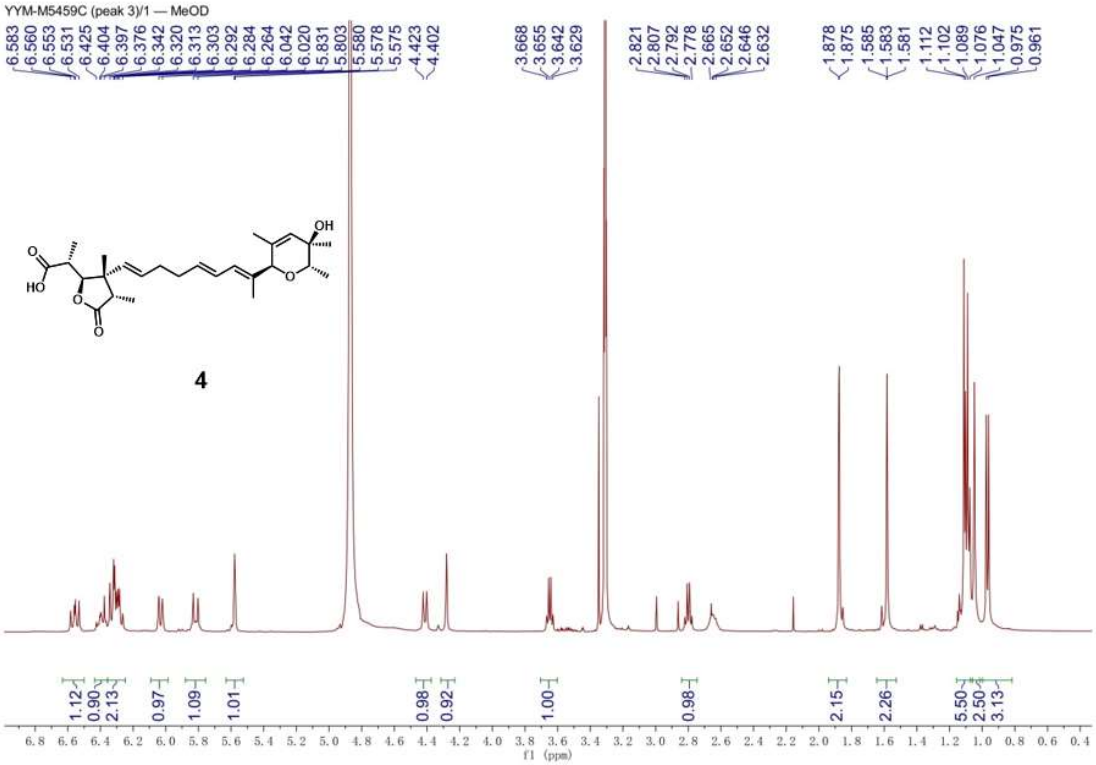


Figure S10. ^{1}H NMR spectrum of compound **4** in CD_3OD.

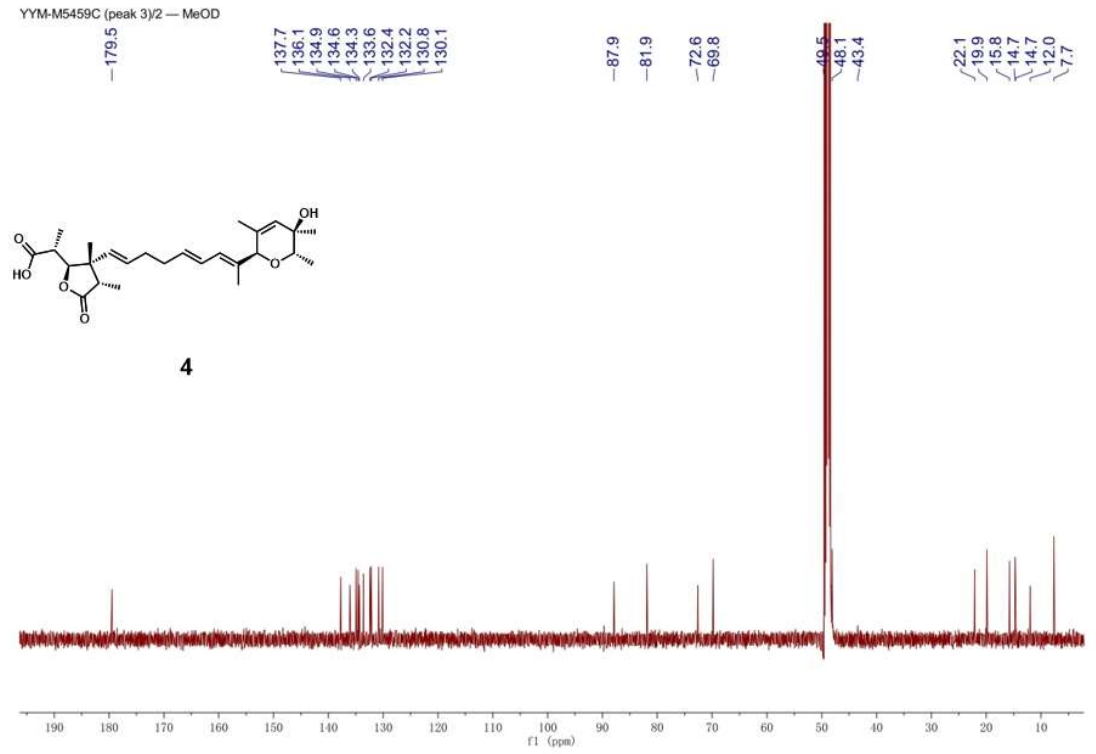


Figure S11. ^{13}C NMR spectrum of compound **4** in CD_3OD.

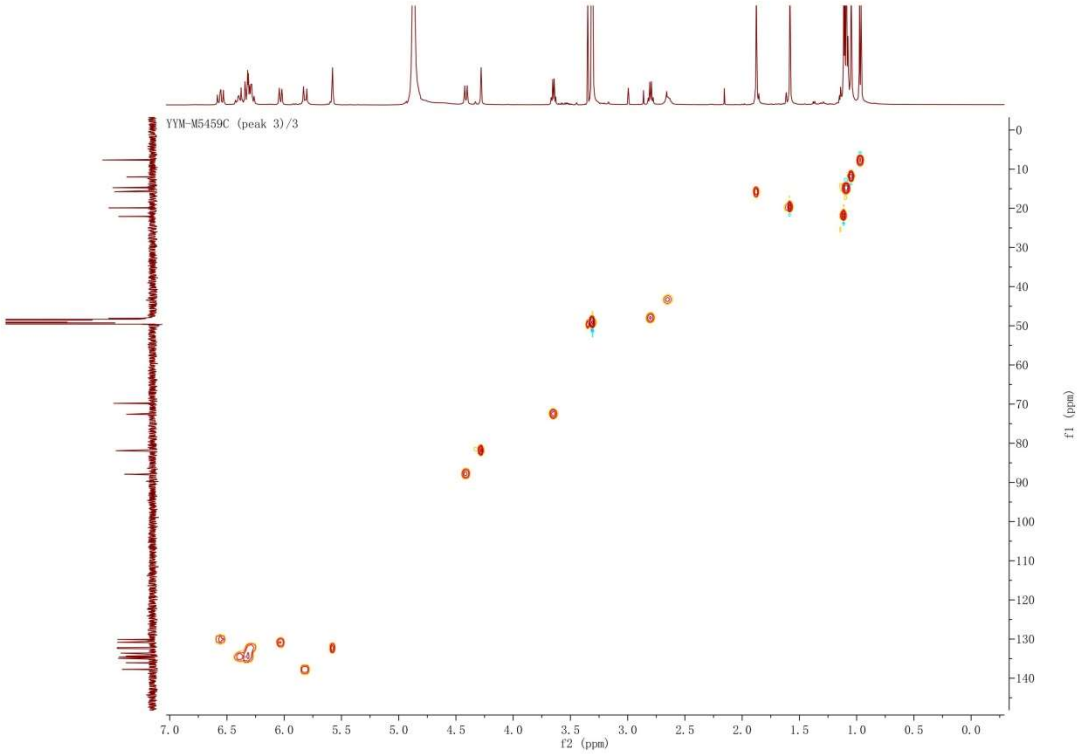


Figure S12. HSQC spectrum of compound **4** in CD_3OD.

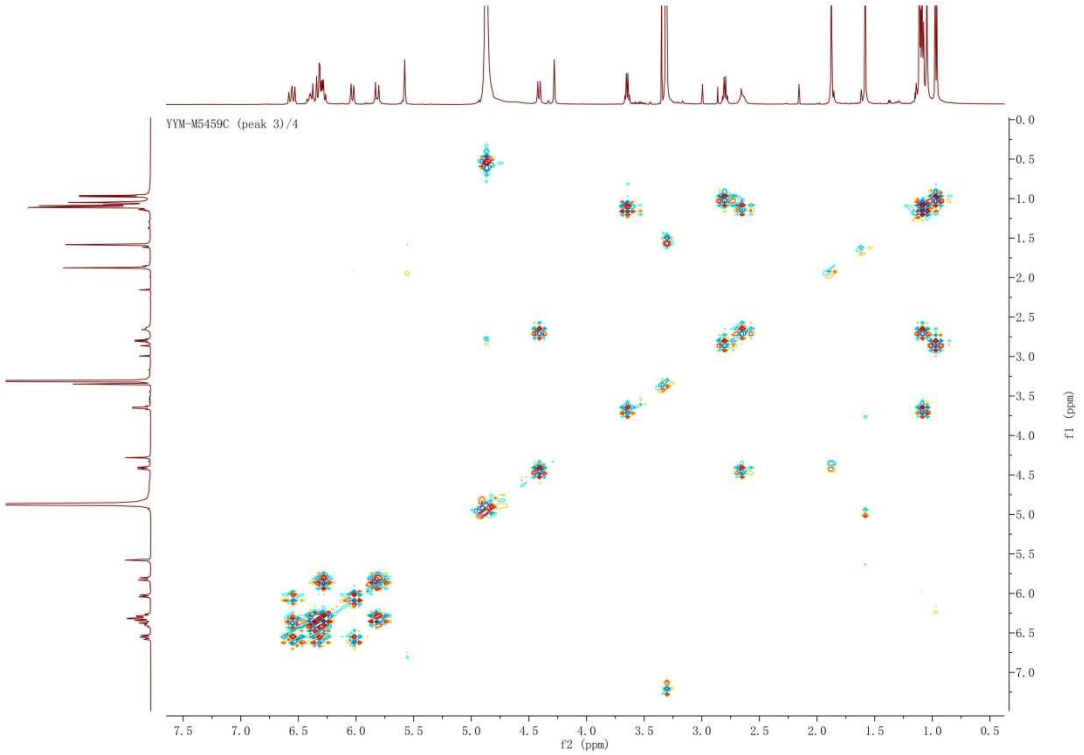


Figure S13. COSY spectrum of compound **4** in CD_3OD.

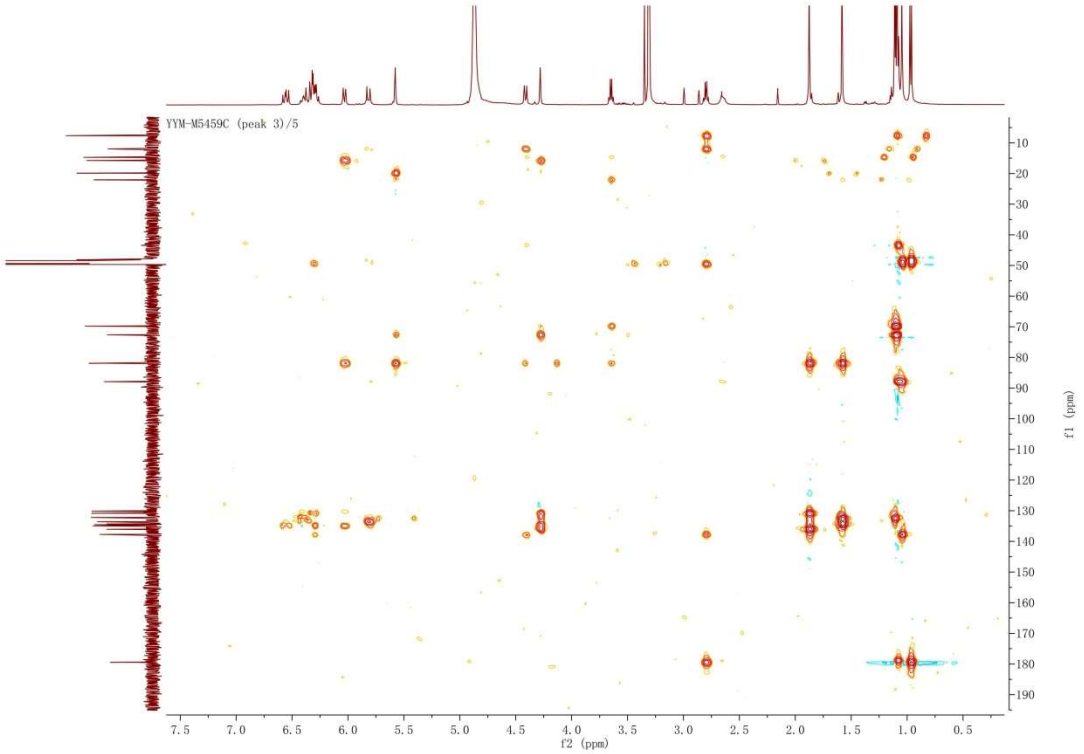


Figure S14. HMBC spectrum of compound **4** in CD_3OD.

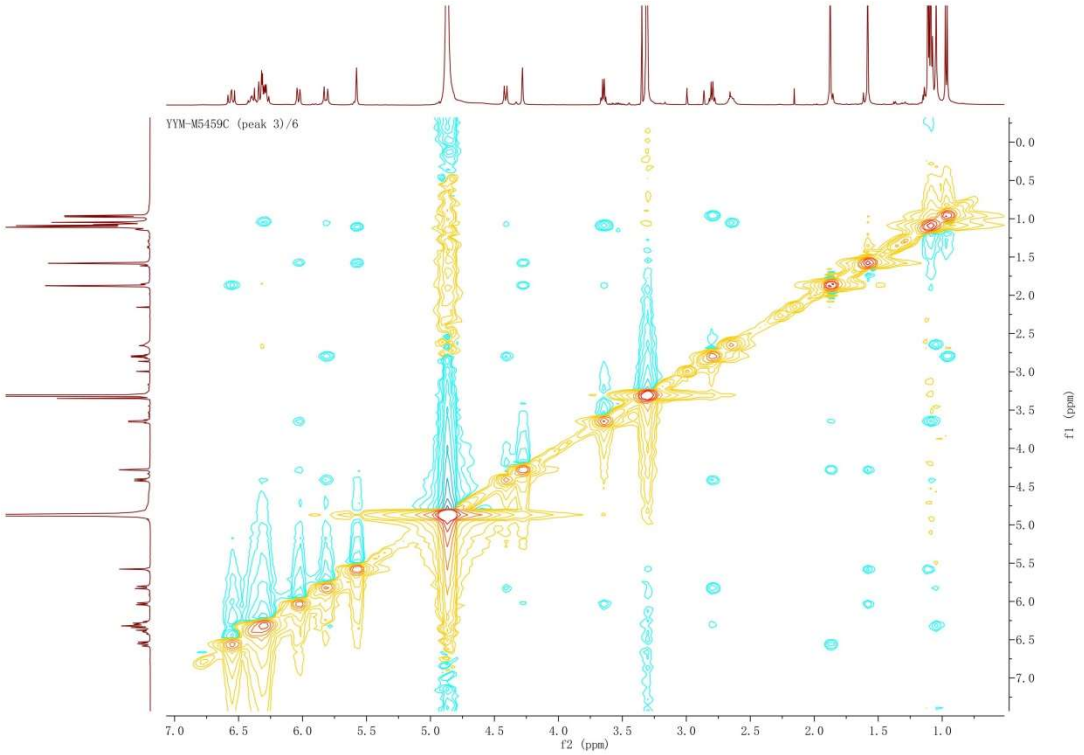


Figure S15. NOESY spectrum of compound **4** in CD_3OD.

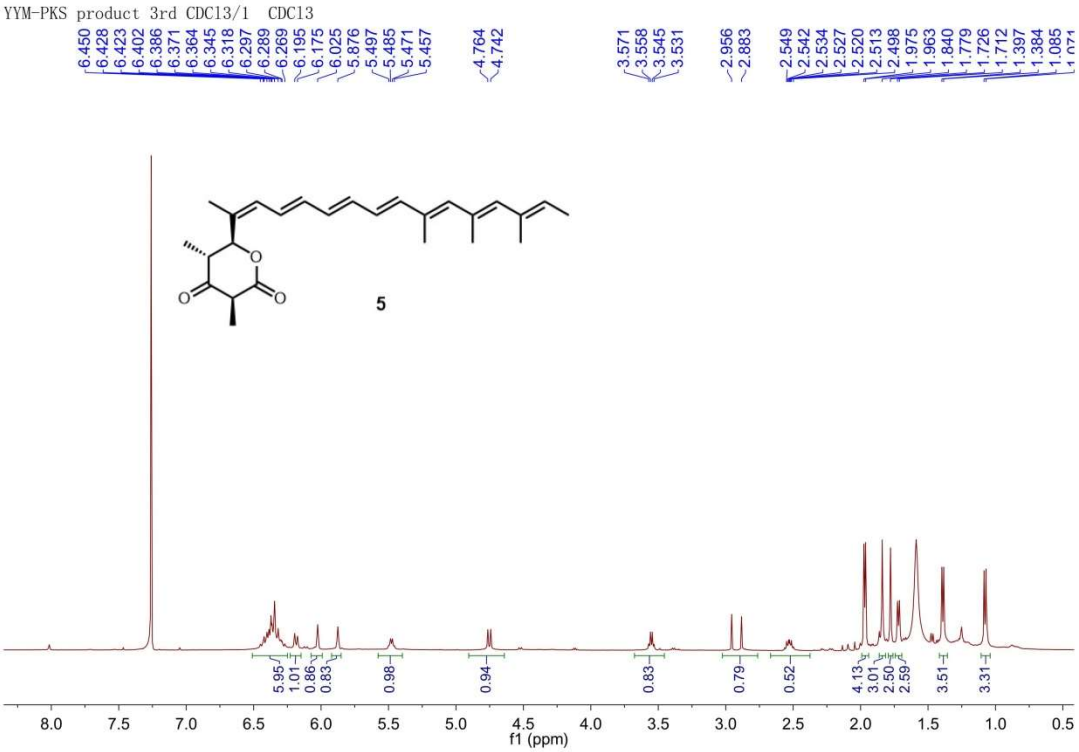


Figure S16. ^{1}H NMR spectrum of compound **5** in $CDCl_3$.

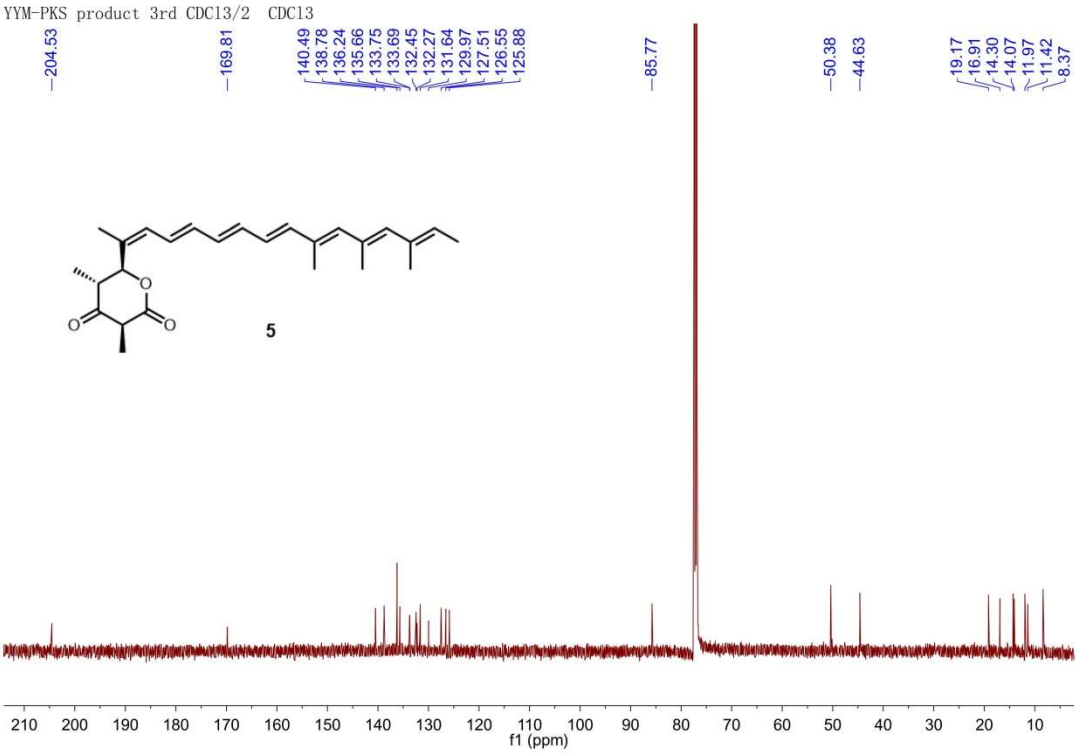


Figure S17. ^{13}C NMR spectrum of compound **5** in $CDCl_3$.

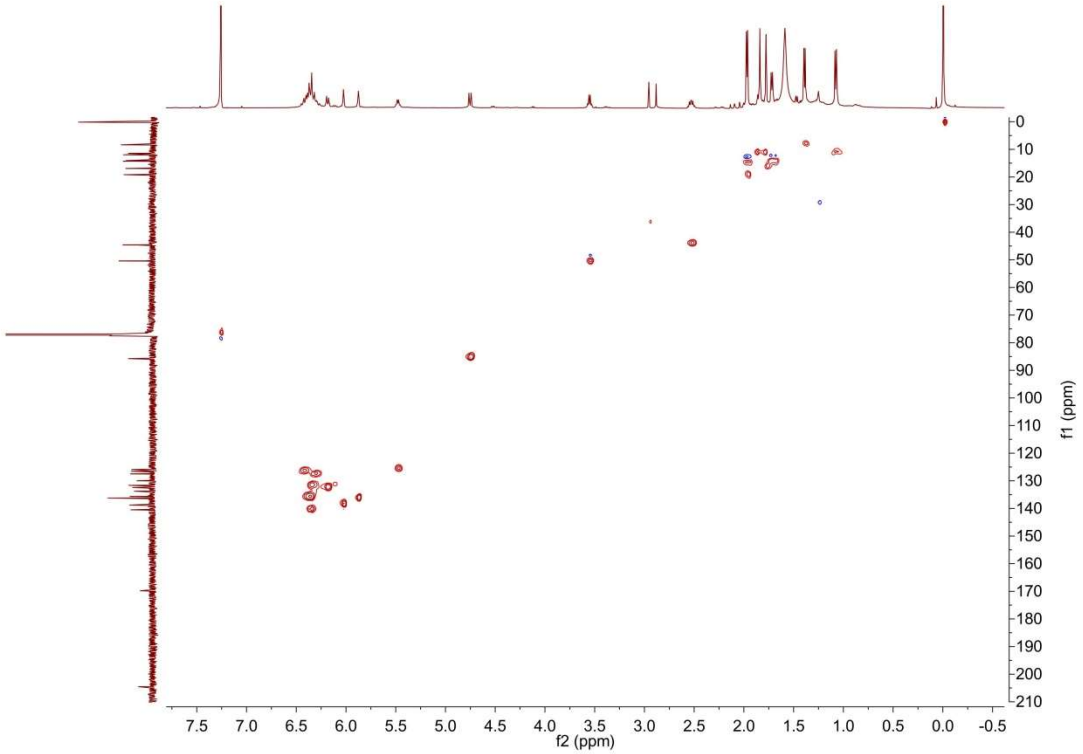


Figure S18. HSQC spectrum of compound **5** in $CDCl_3$.

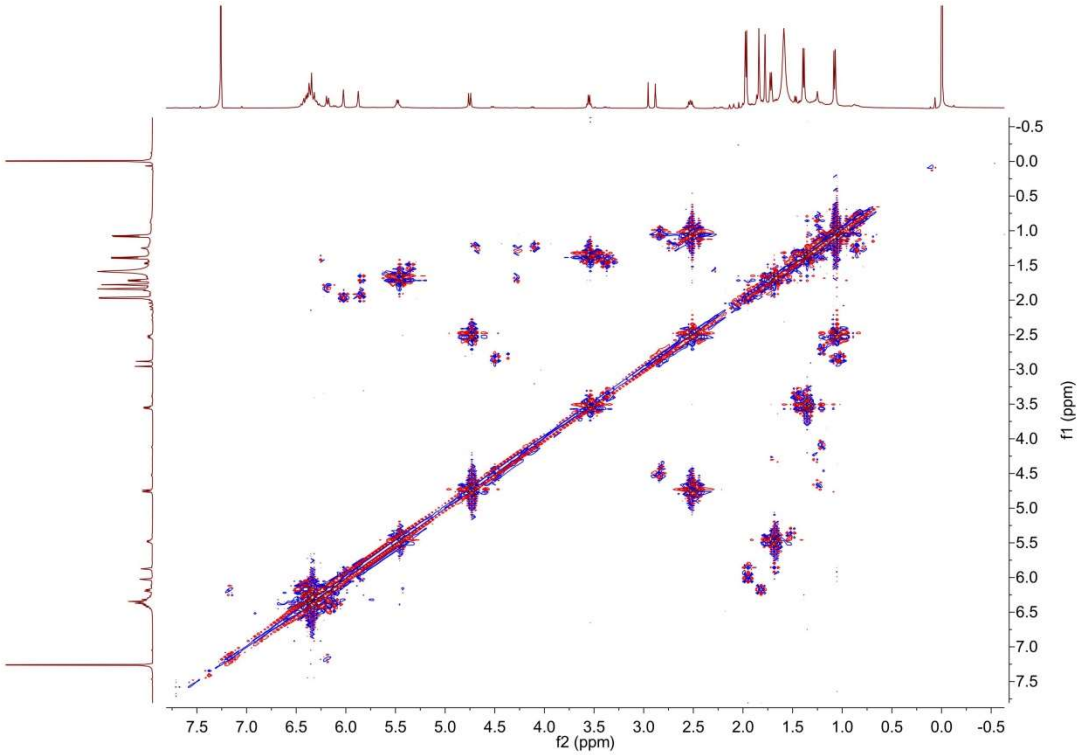


Figure S19. COSY spectrum of compound **5** in $CDCl_3$.

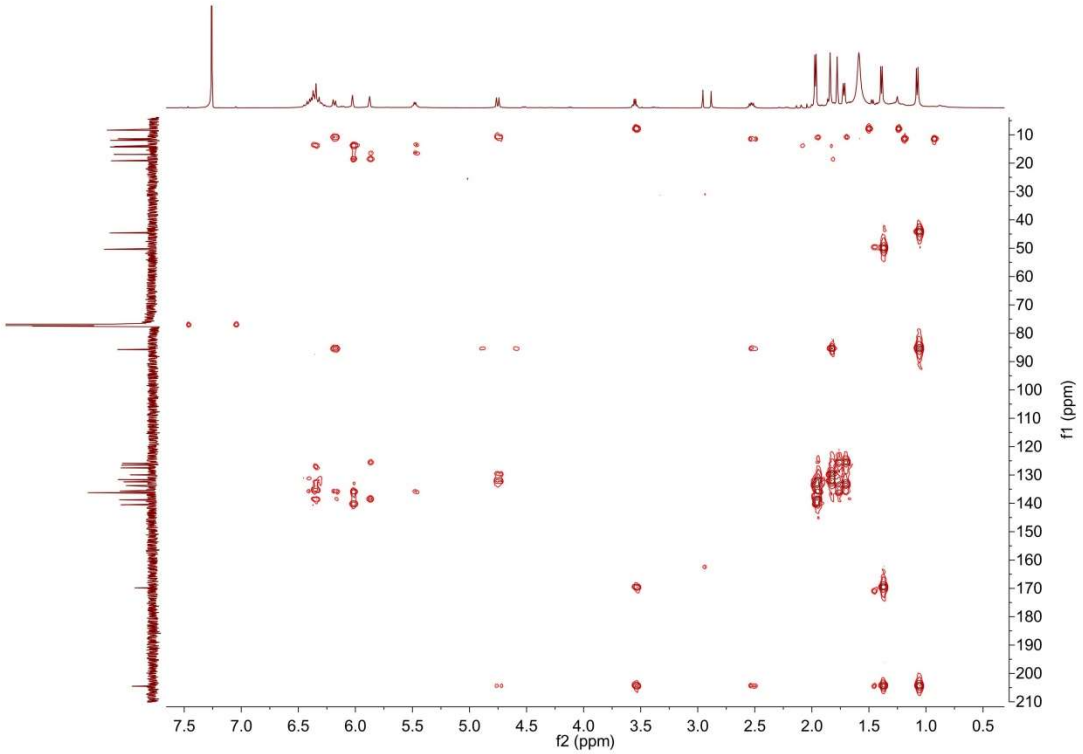


Figure S20. HMBC spectrum of compound **5** in $CDCl_3$.

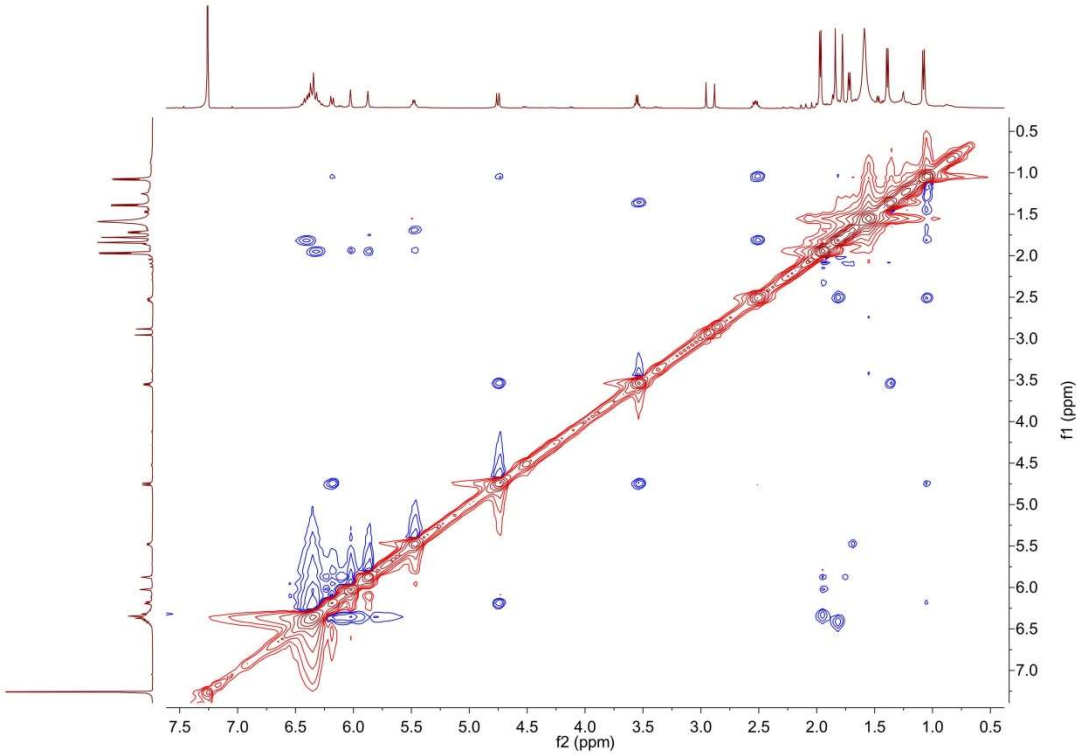


Figure S21. NOESY spectrum of compound **5** in $CDCl_3$.

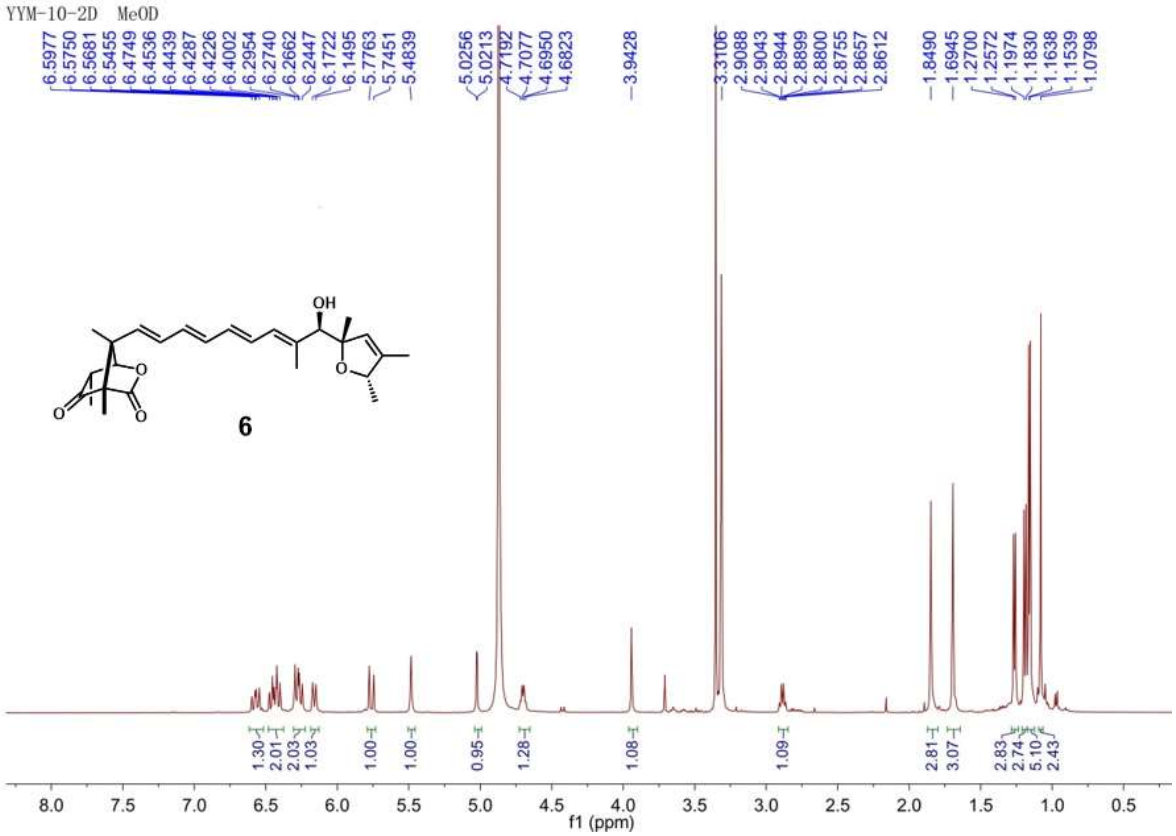


Figure S22. ^{1}H NMR spectrum of compound **6** in CD_3OD.

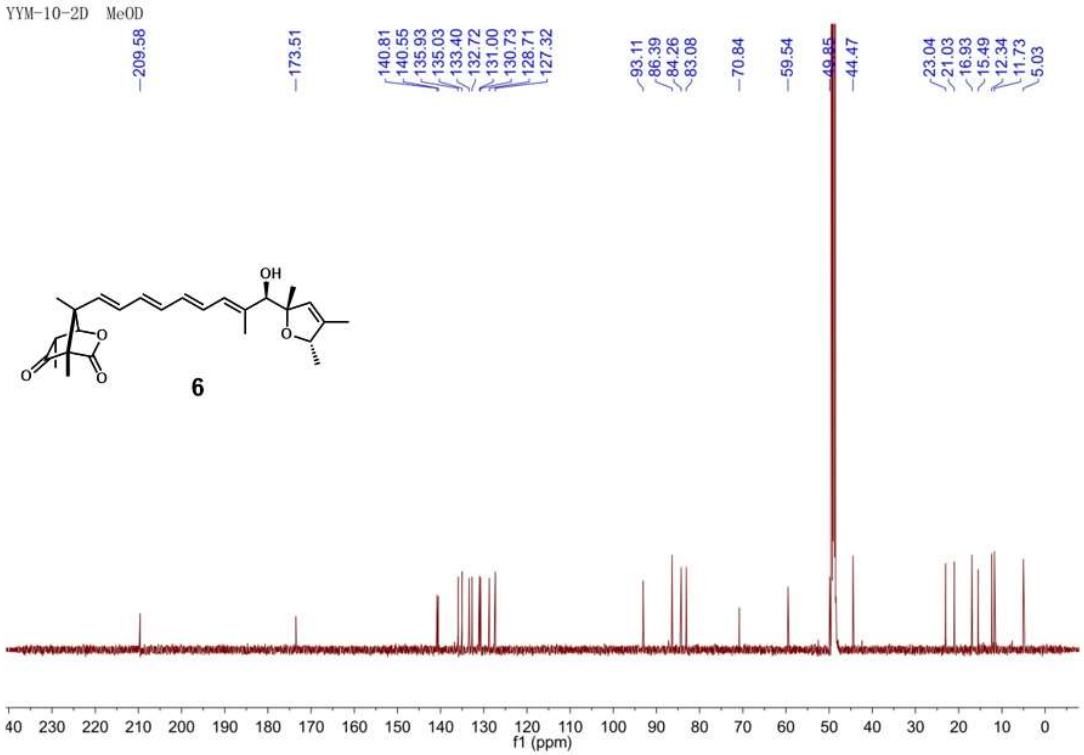


Figure S23. ^{13}C NMR spectrum of compound **6** in CD_3OD.

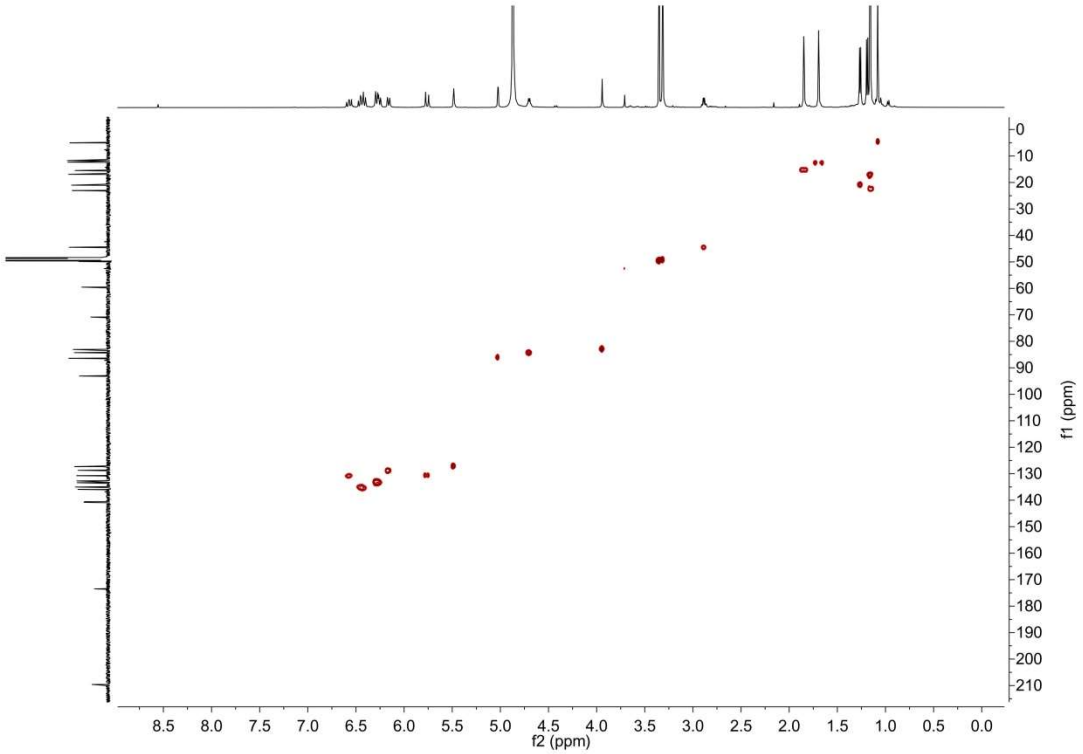


Figure S24. HSQC spectrum of compound **6** in CD_3OD.

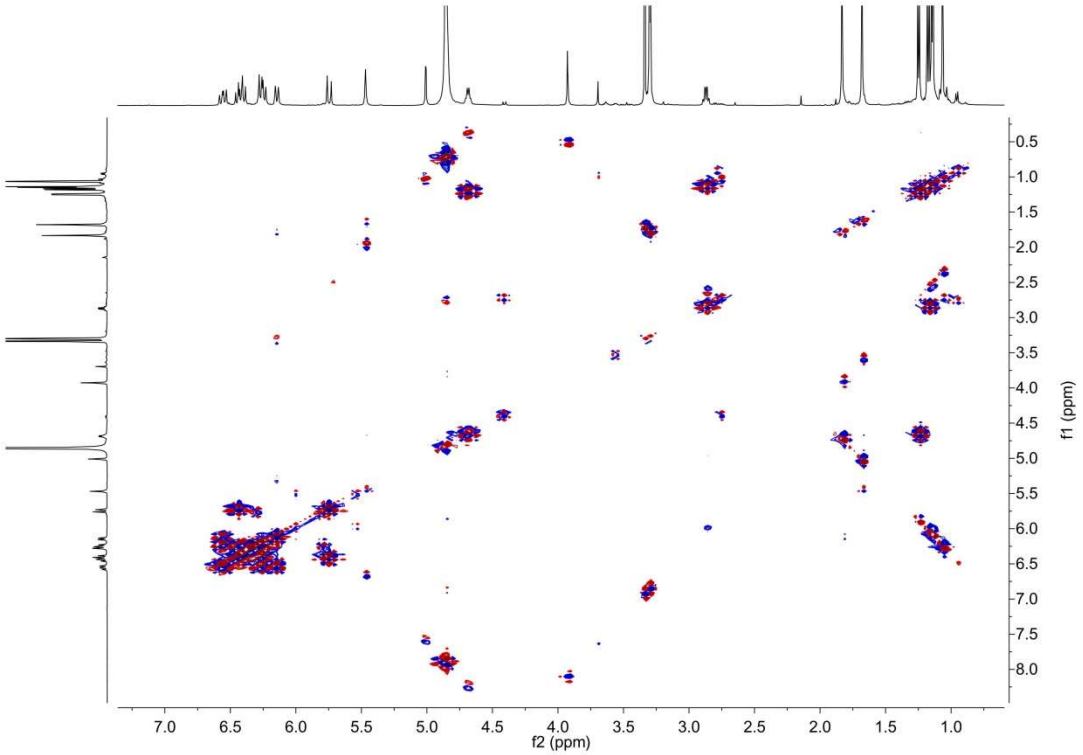


Figure S25. COSY spectrum of compound **6** in CD_3OD.

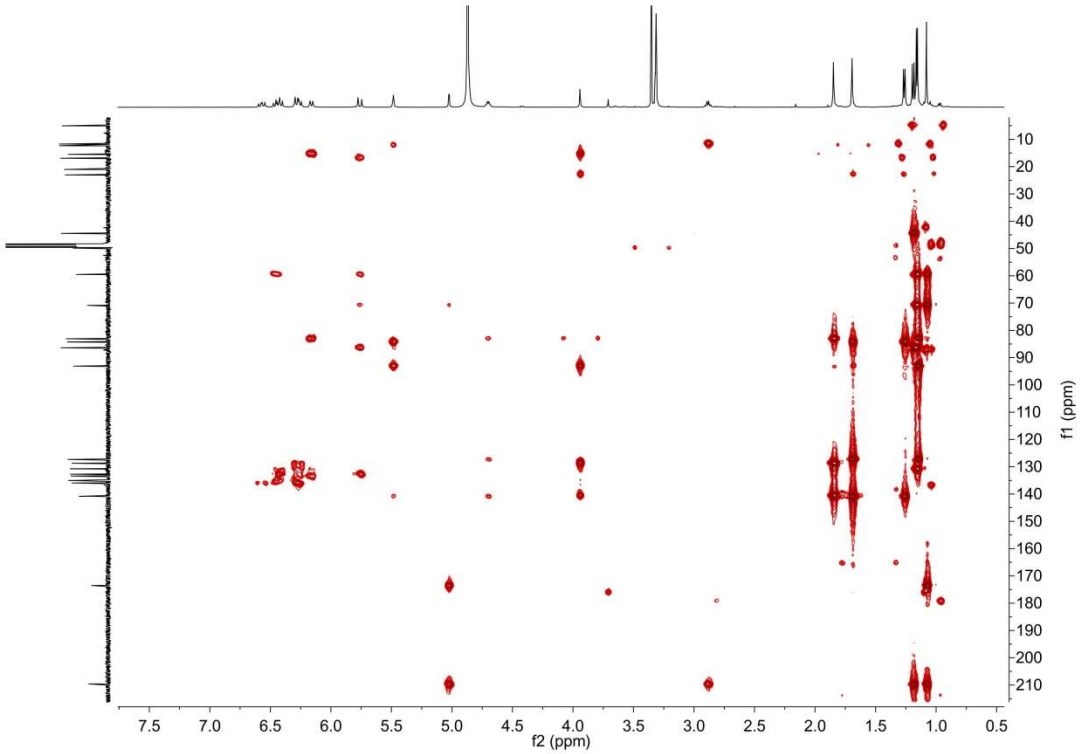


Figure S26. HMBC spectrum of compound **6** in CD_3OD.

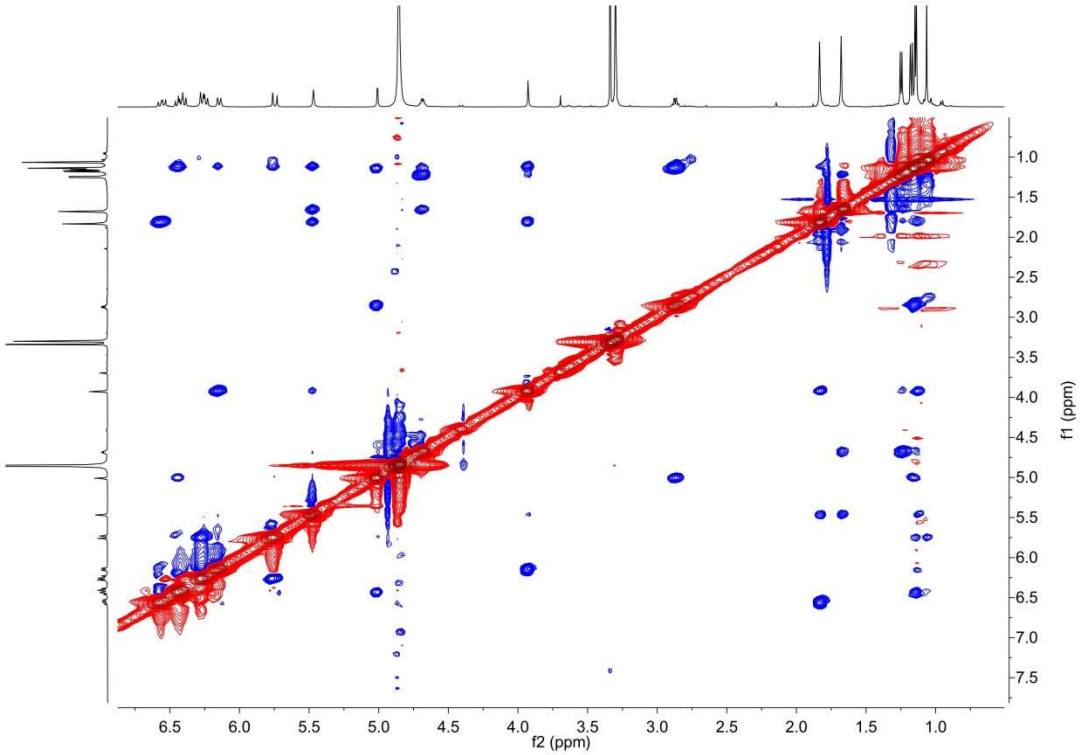


Figure S27. NOESY spectrum of compound **6** in CD_3OD.

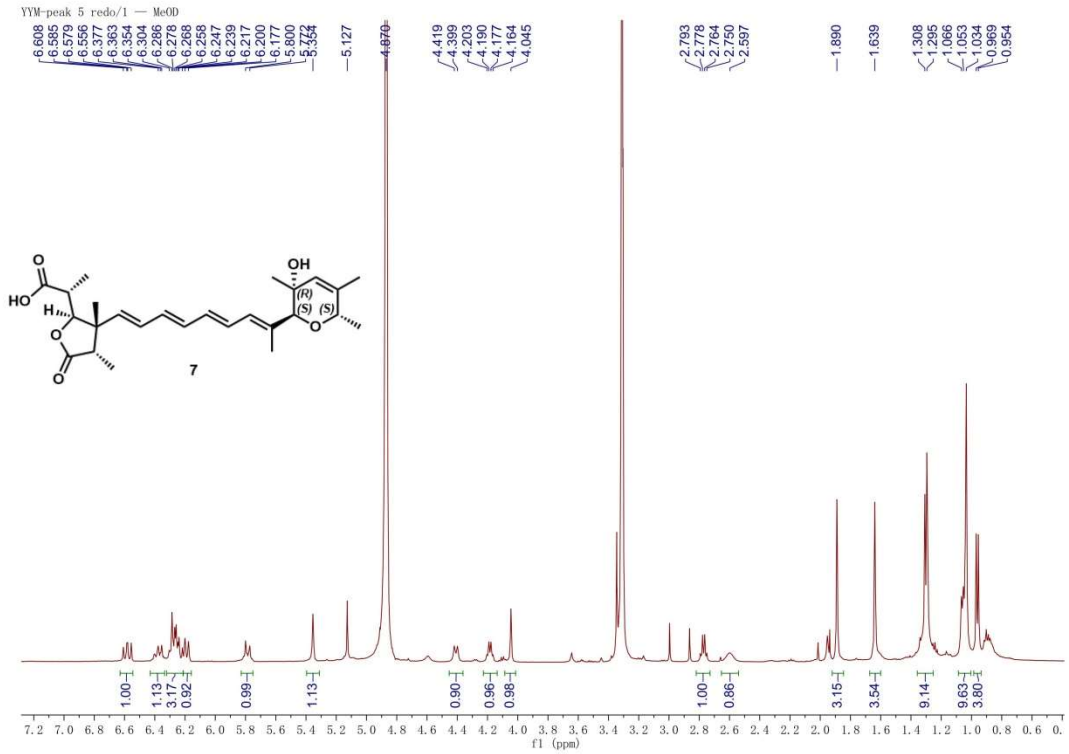


Figure S28. ^{1}H NMR spectrum of compound **7** in CD_3OD.

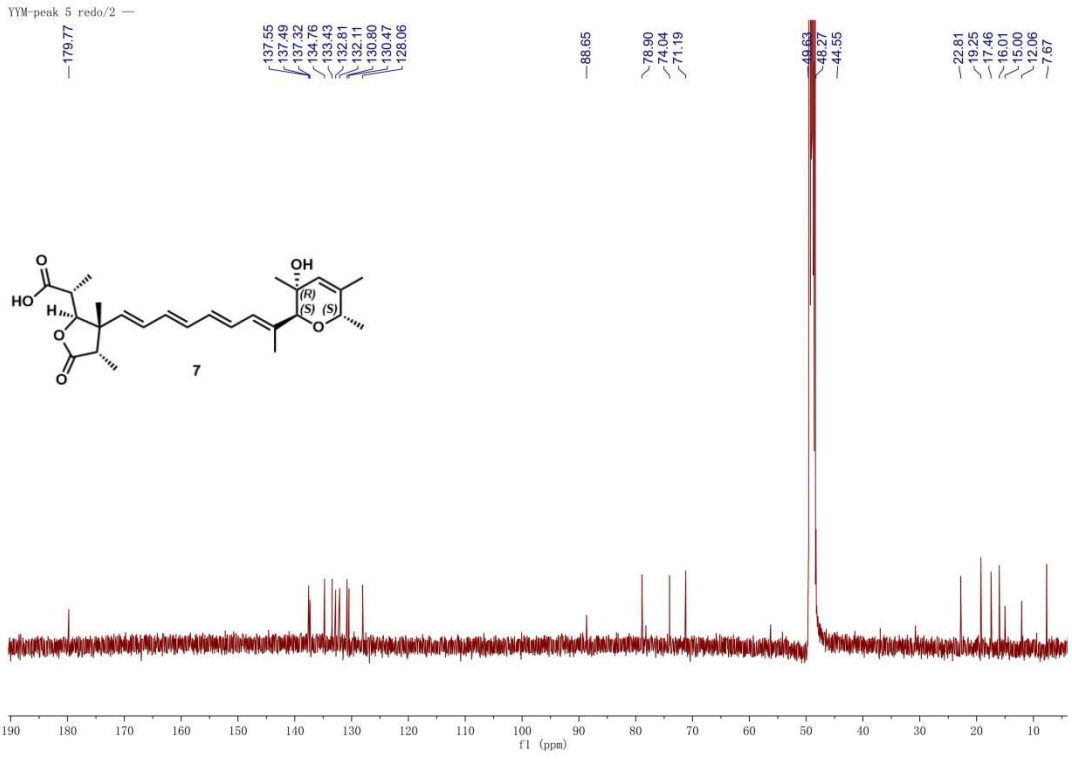


Figure S29. ^{13}C NMR spectrum of compound **7** in CD_3OD.

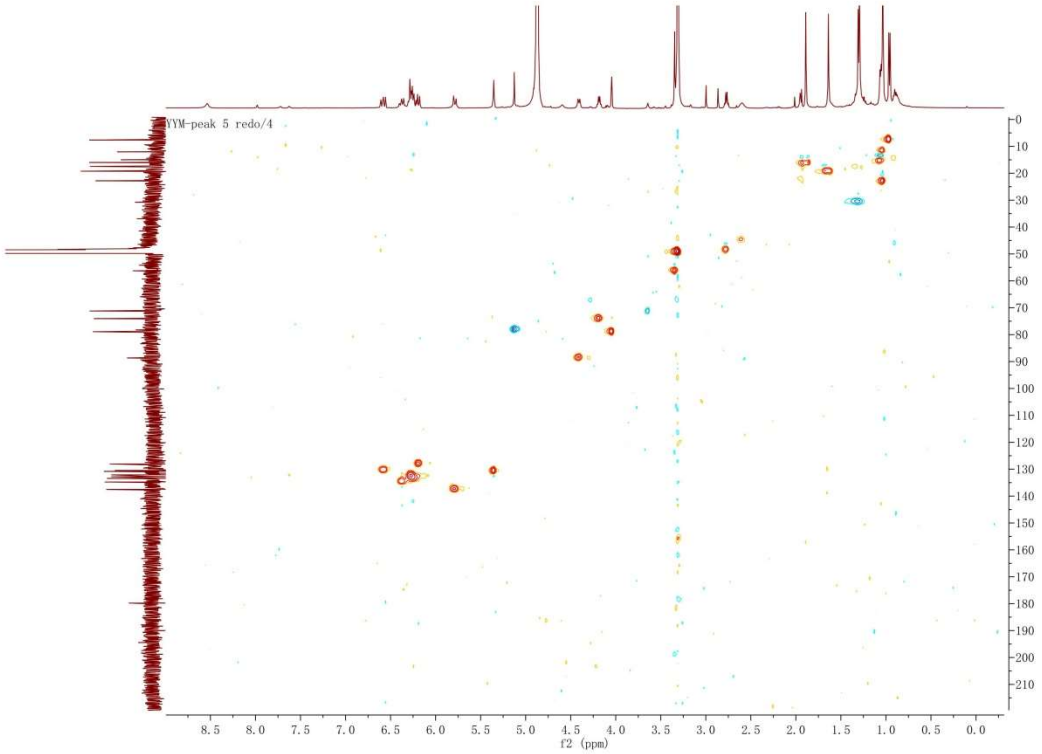


Figure S30. HSQC spectrum of compound **7** in CD_3OD.

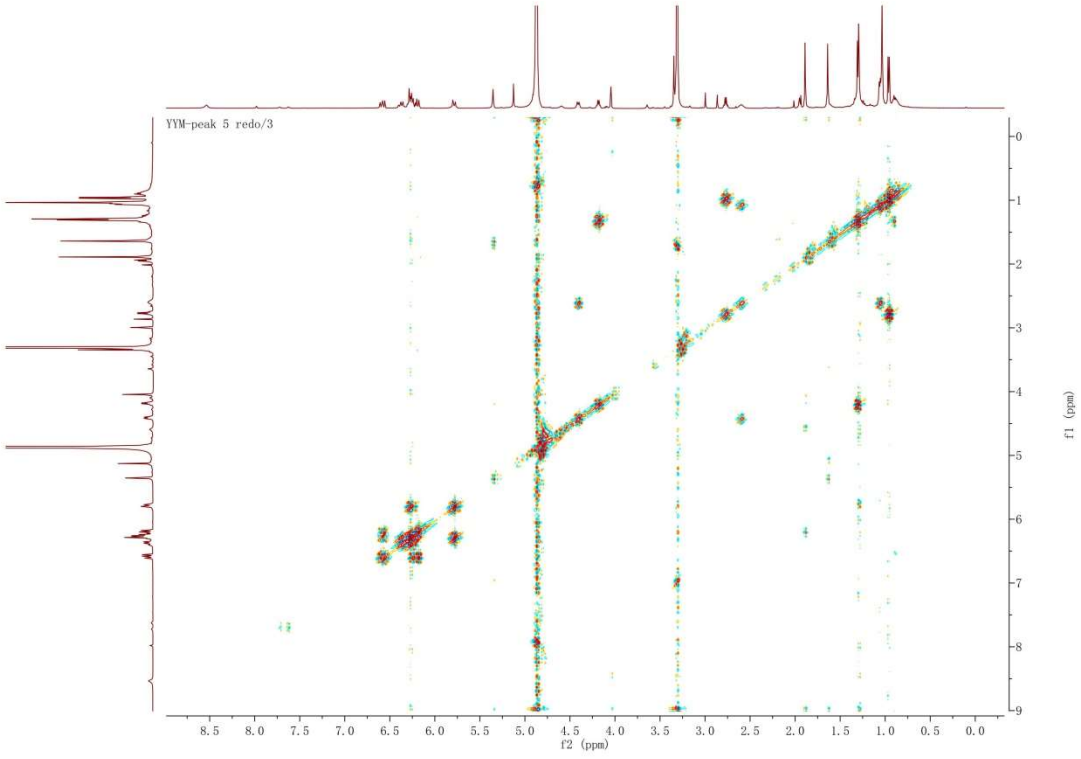


Figure S31. COSY spectrum of compound **7** in CD_3OD.

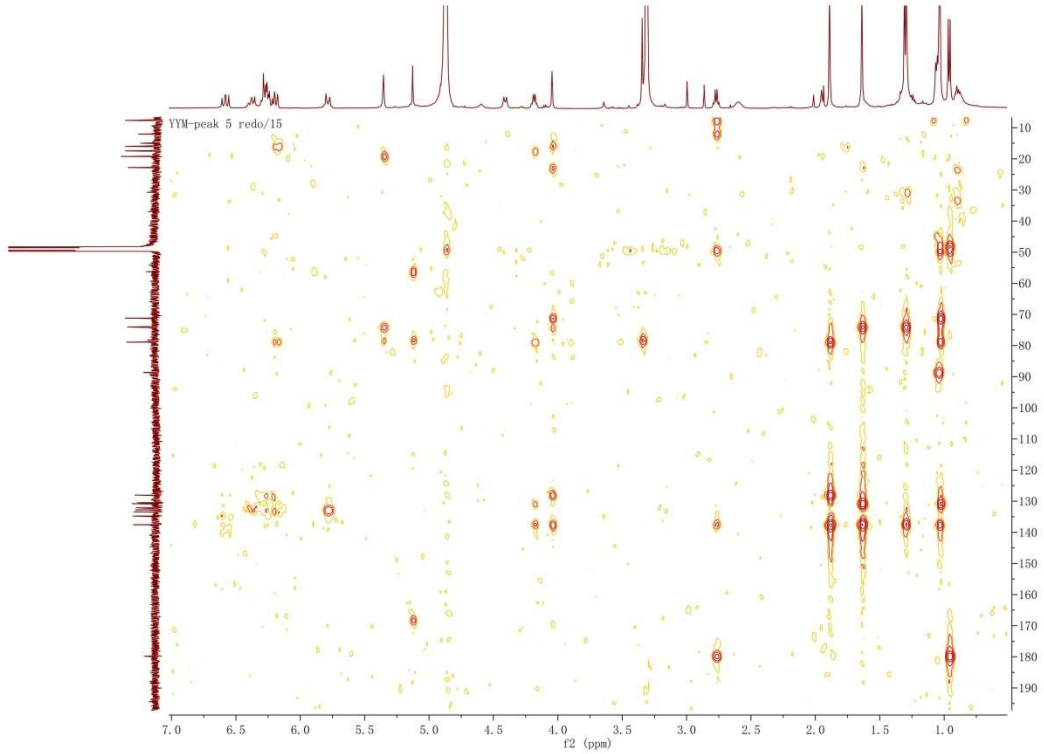


Figure S32. HMBC spectrum of compound **7** in CD_3OD.

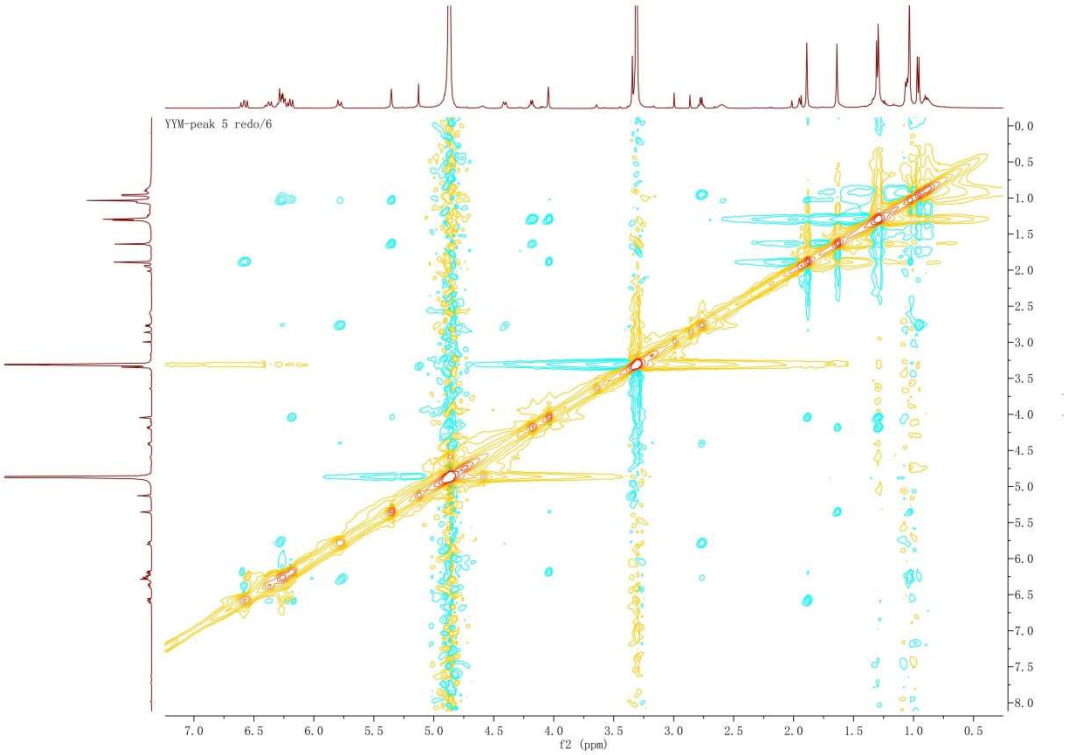


Figure S33. NOESY spectrum of compound **7** in CD_3OD.

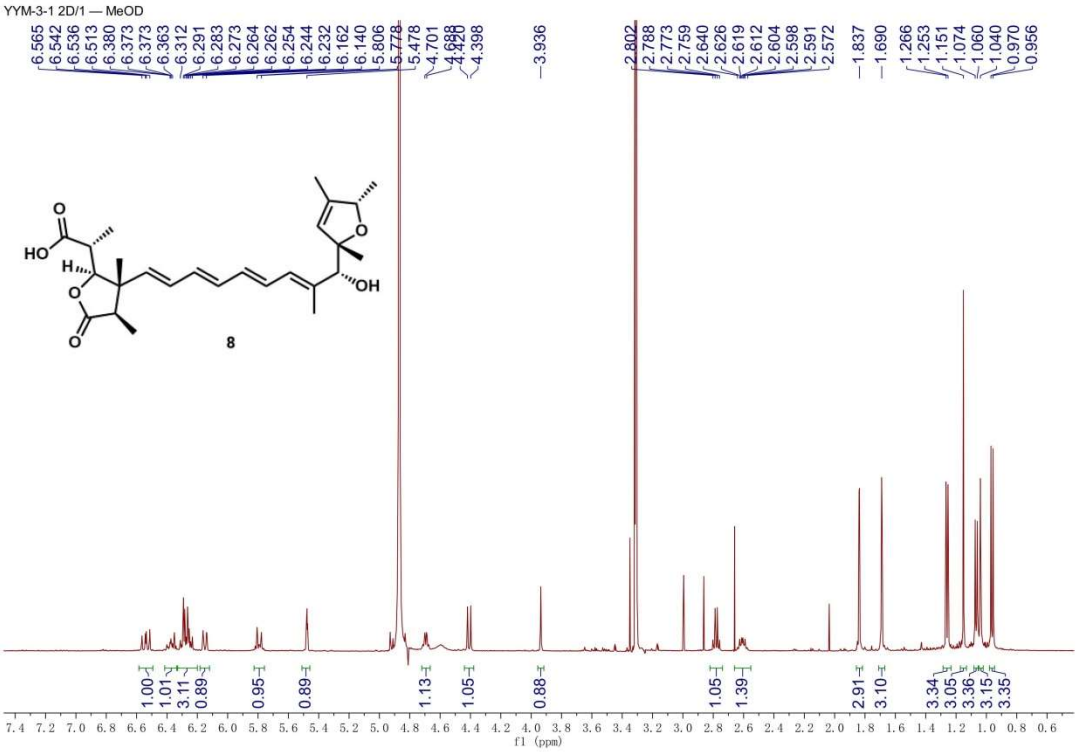


Figure S34. ^{1}H NMR spectrum of compound **8** in CD_3OD.

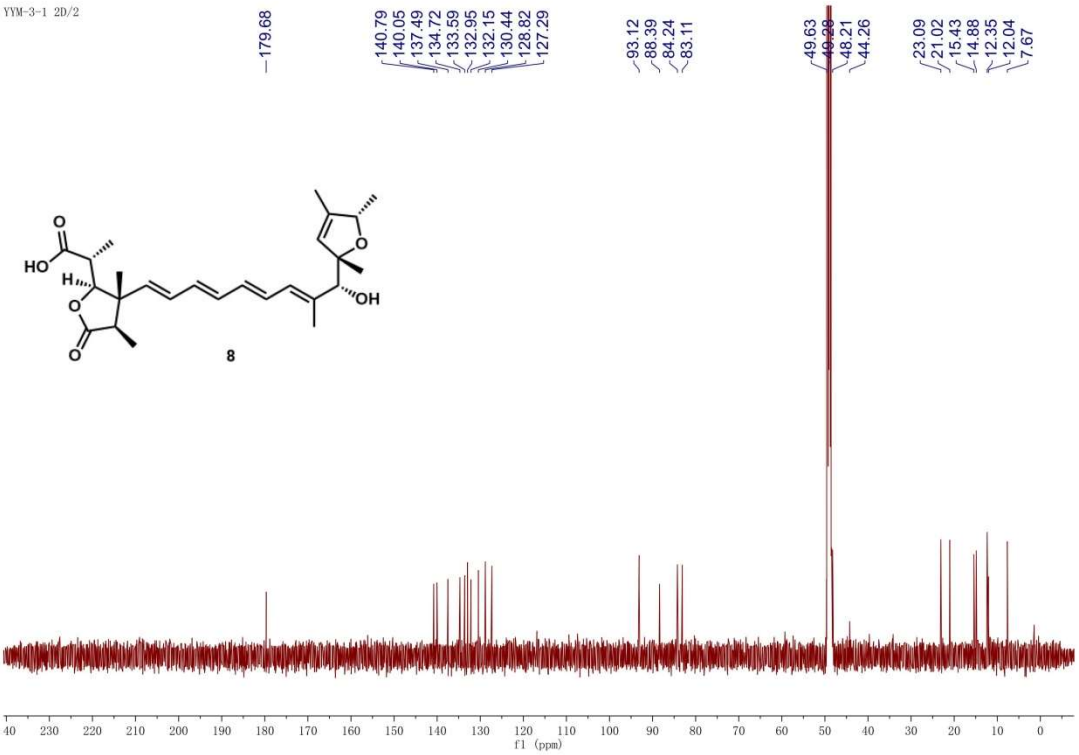


Figure S35. ^{13}C NMR spectrum of compound **8** in CD_3OD.

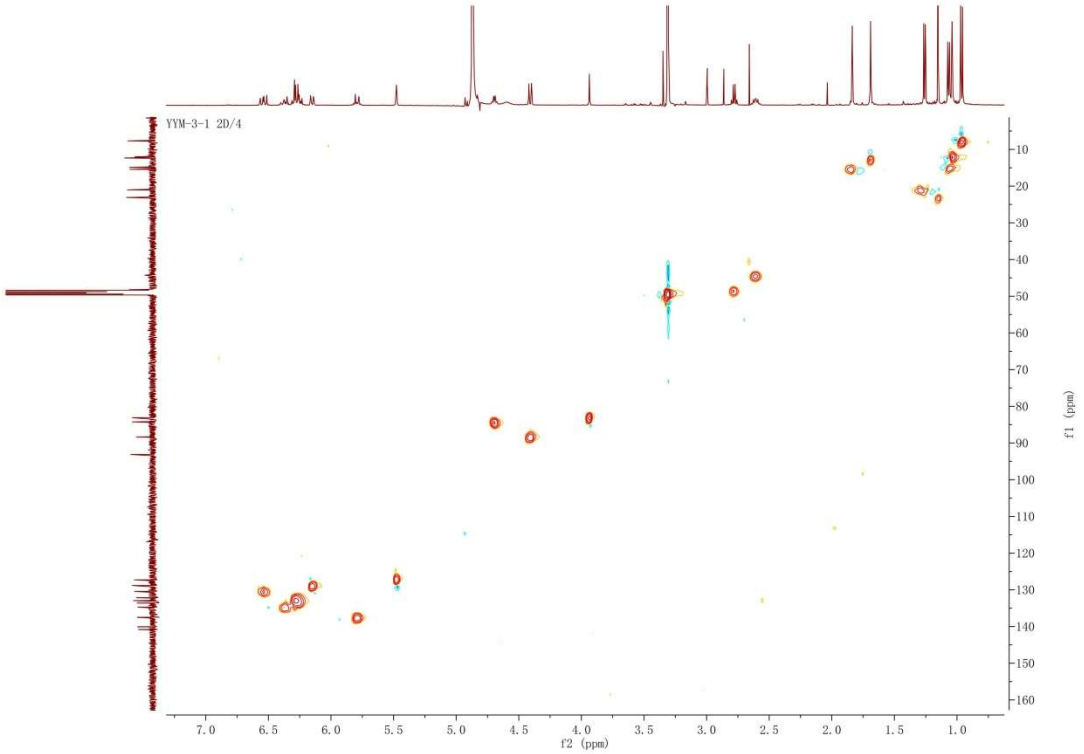


Figure S36. HSQC spectrum of compound **8** in CD_3OD.

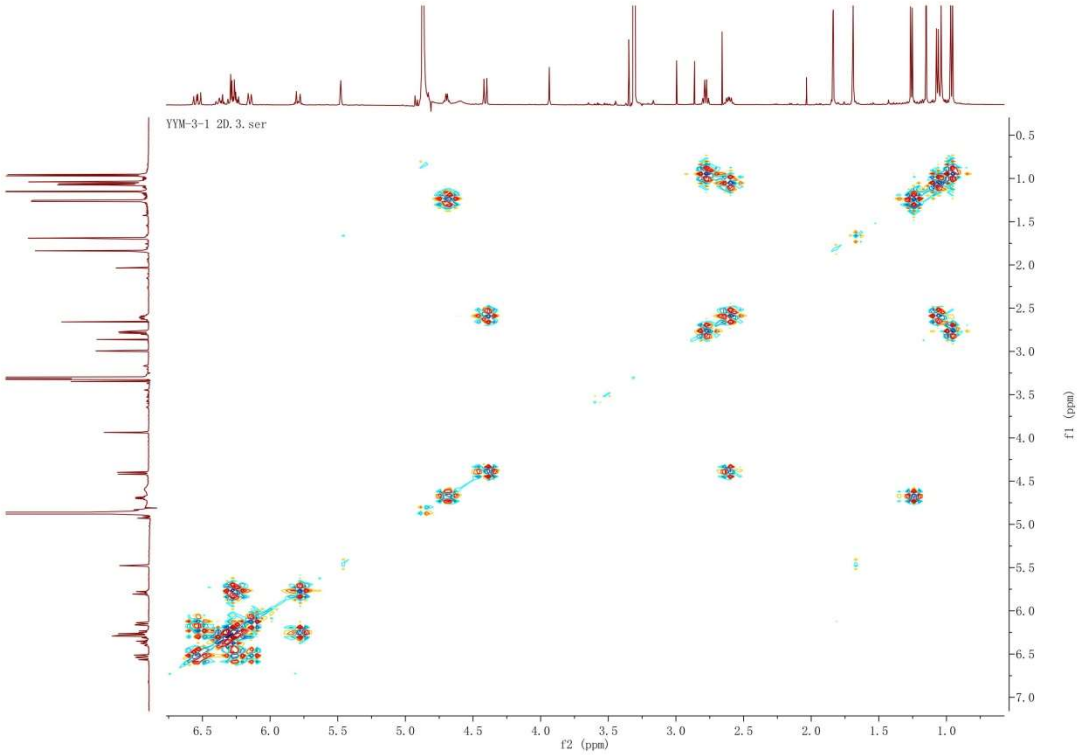


Figure S37. COSY spectrum of compound **8** in CD_3OD.

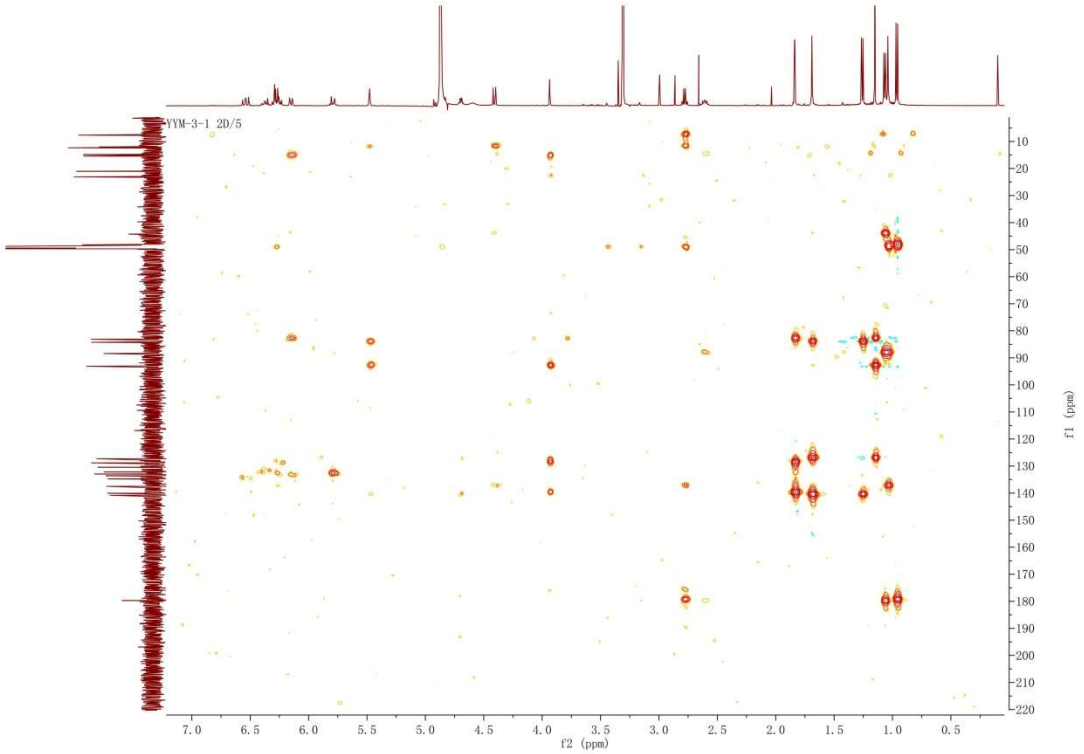


Figure S38. HMBC spectrum of compound **8** in CD_3OD.

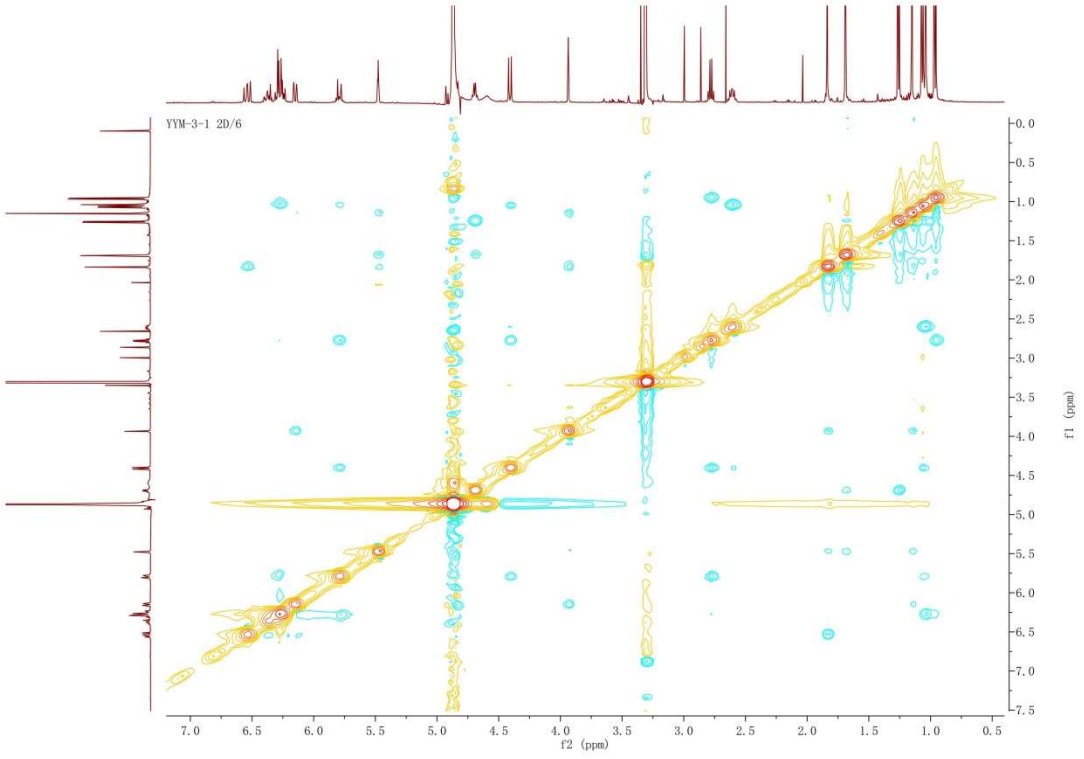


Figure S39. NOESY spectrum of compound **8** in CD_3OD.

9.2 Supplementary information for Section 5

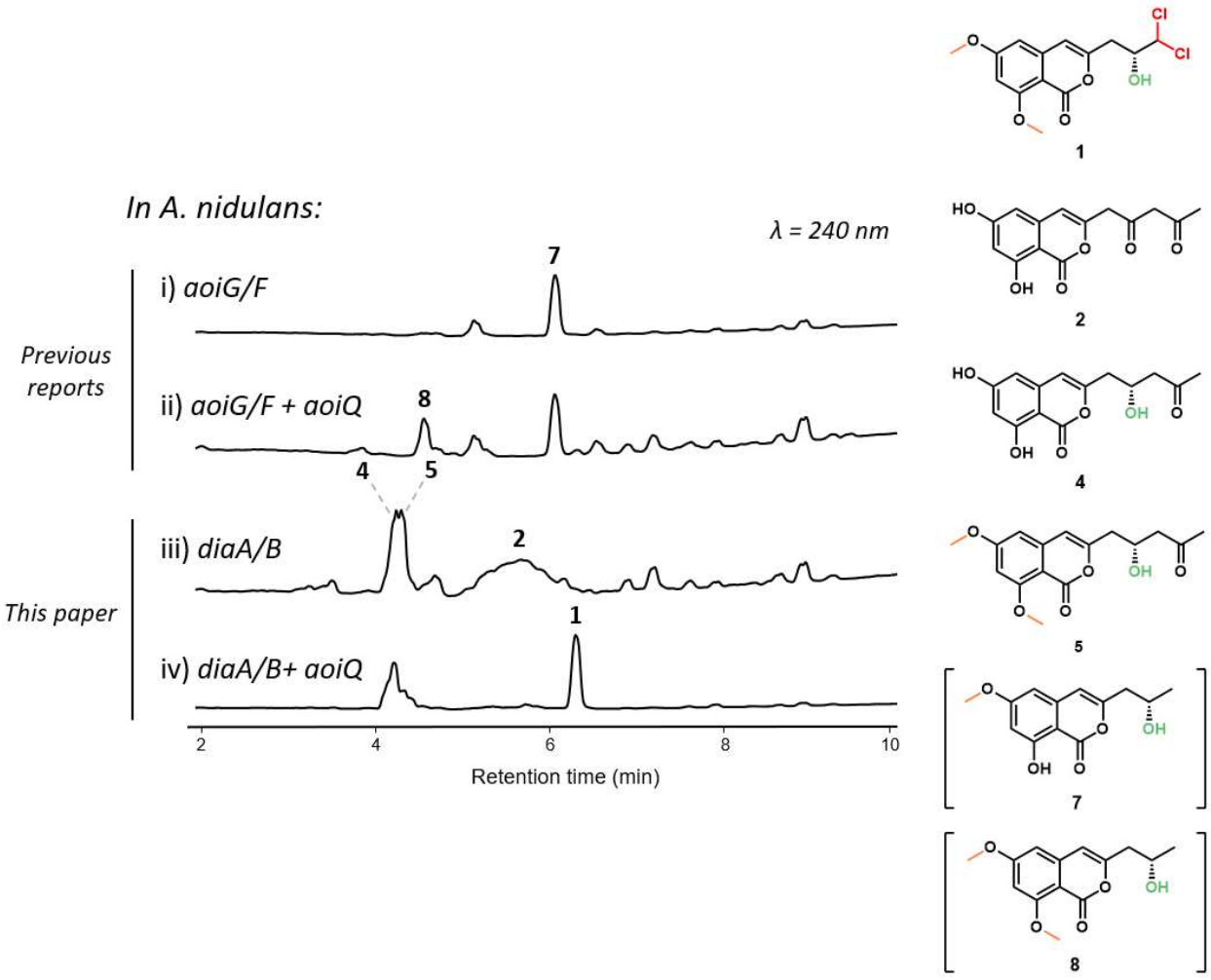


Figure S40. Heterologous expression of *aoi* and *dia* gene clusters in *A. nidulans*

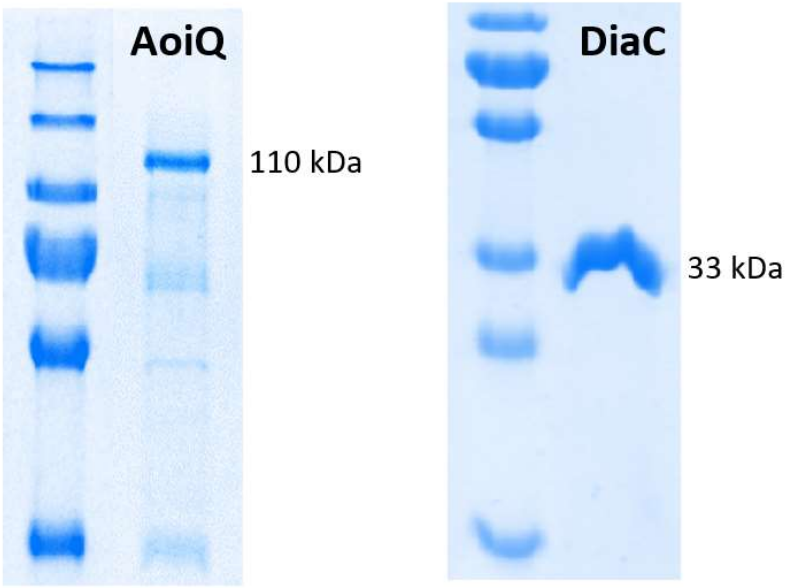


Figure S41. SDS-PAGE protein gel for AoiQ (from yeast) and DiaC (from *E. coli*)

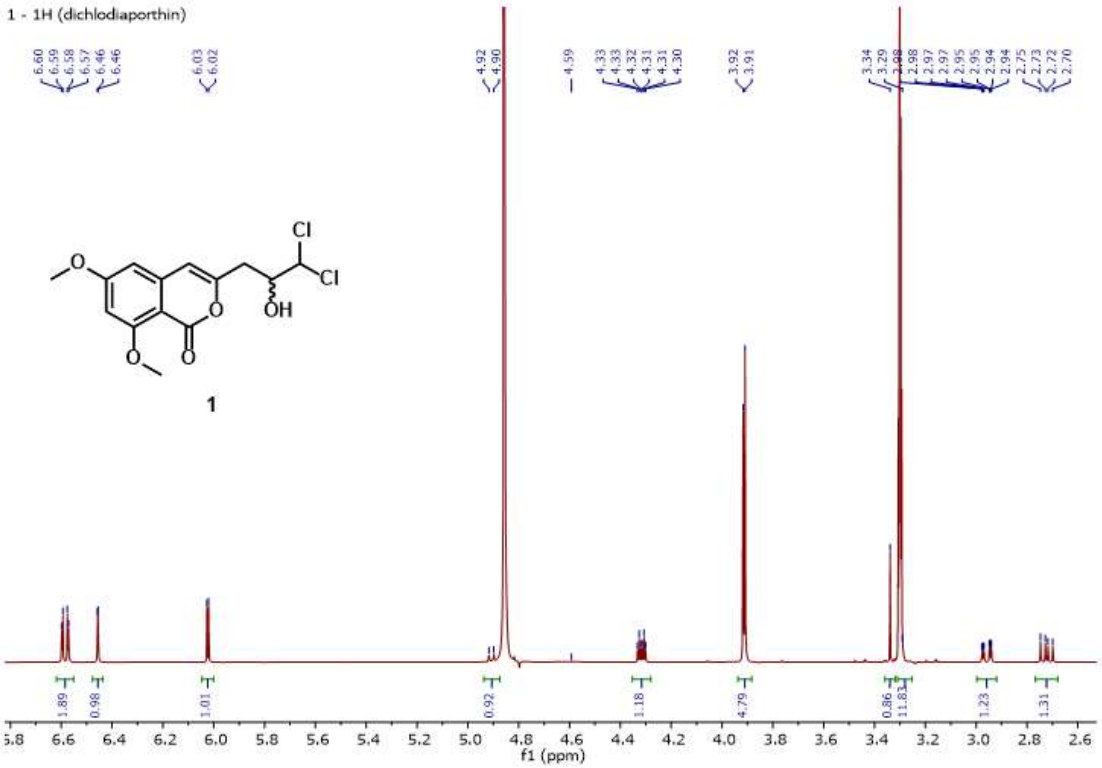


Figure S42. ^{1}H NMR spectrum of compound **1** in CD_3OD.

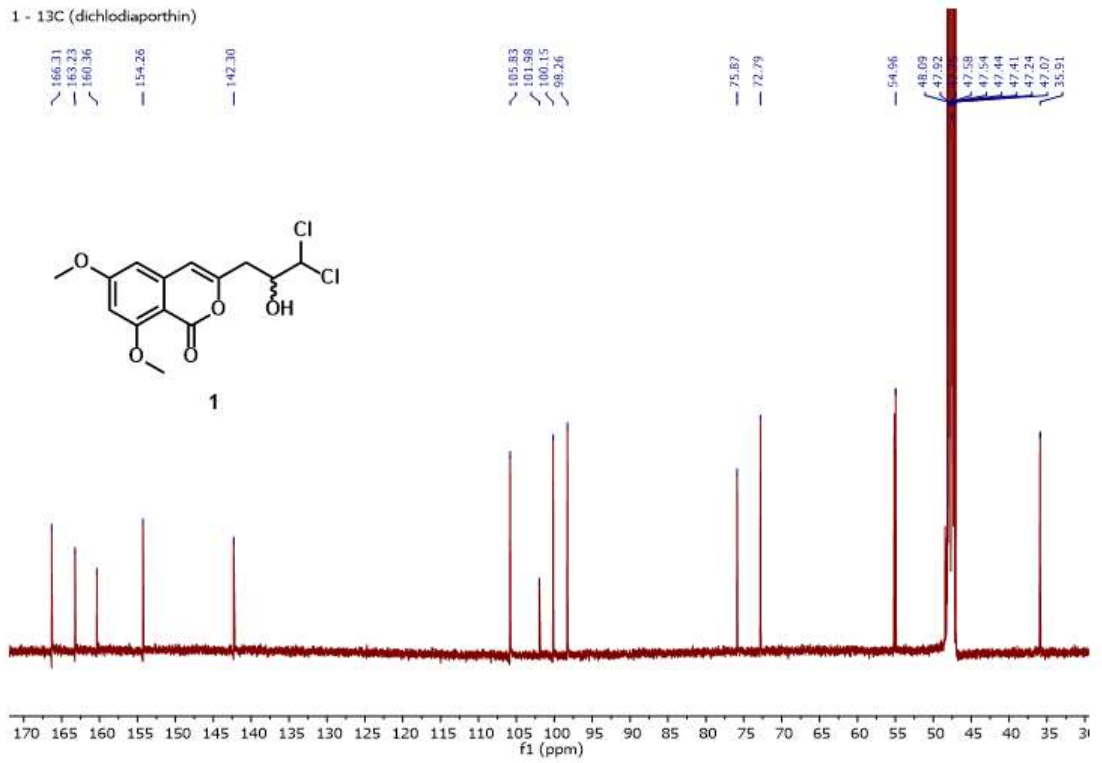


Figure S43. ^{13}C NMR spectrum of compound **1** in CD_3OD.

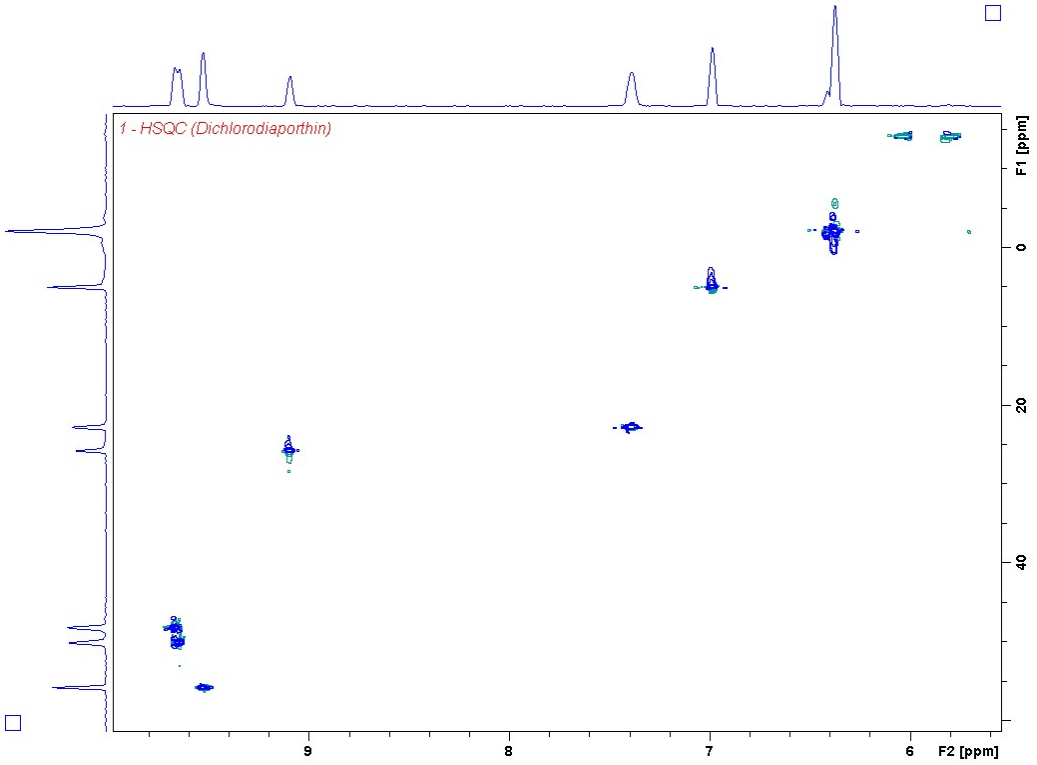


Figure S44. HSQC spectrum of compound **1** in CD_3OD.

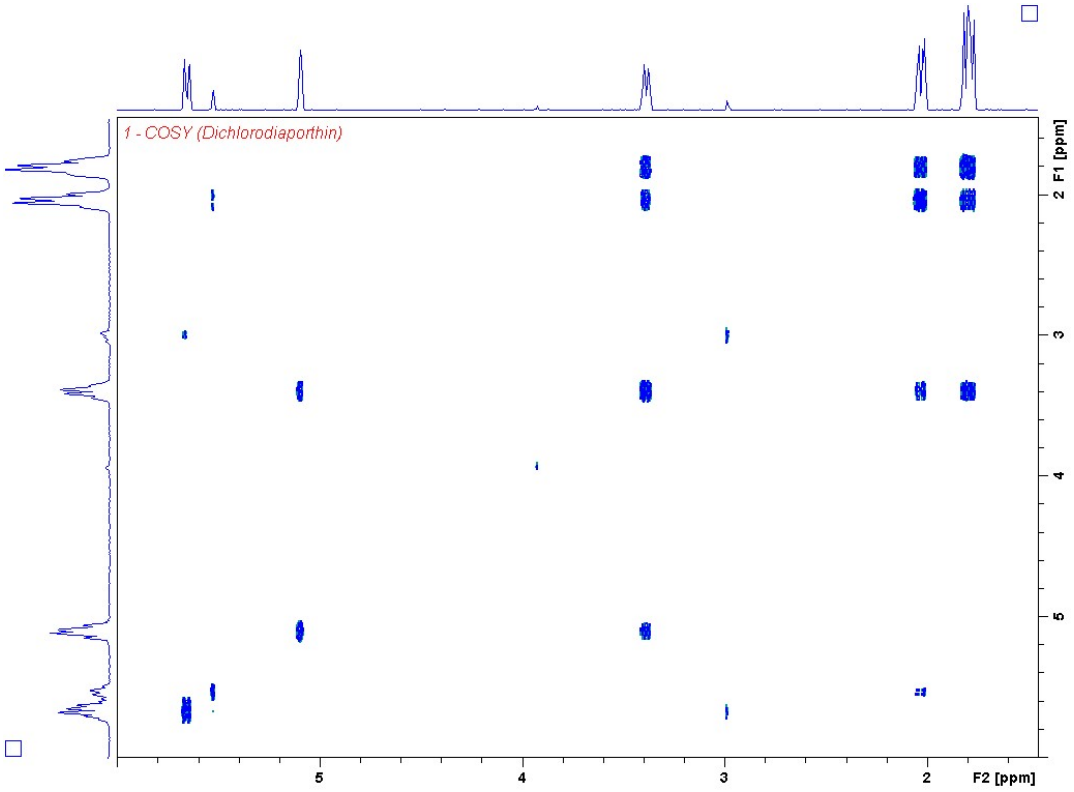


Figure S45. COSY spectrum of compound **1** in CD_3OD.

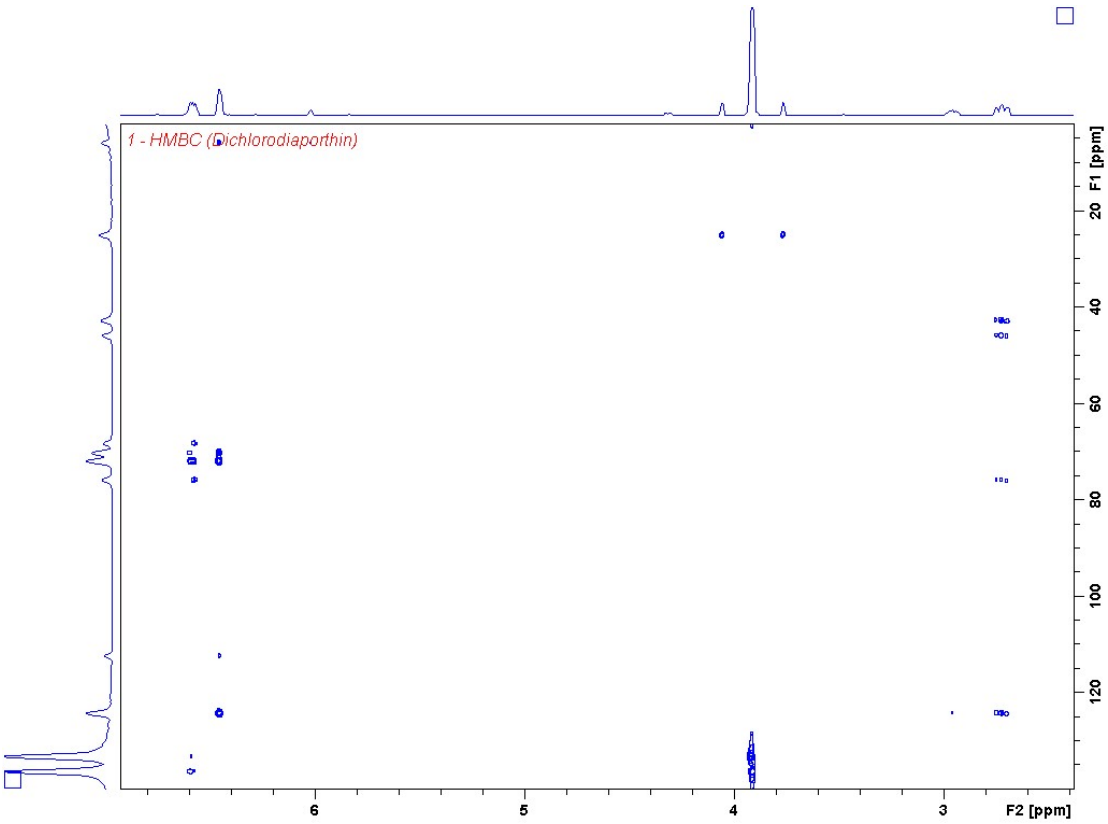


Figure S46. HMBC spectrum of compound **1** in CD_3OD.

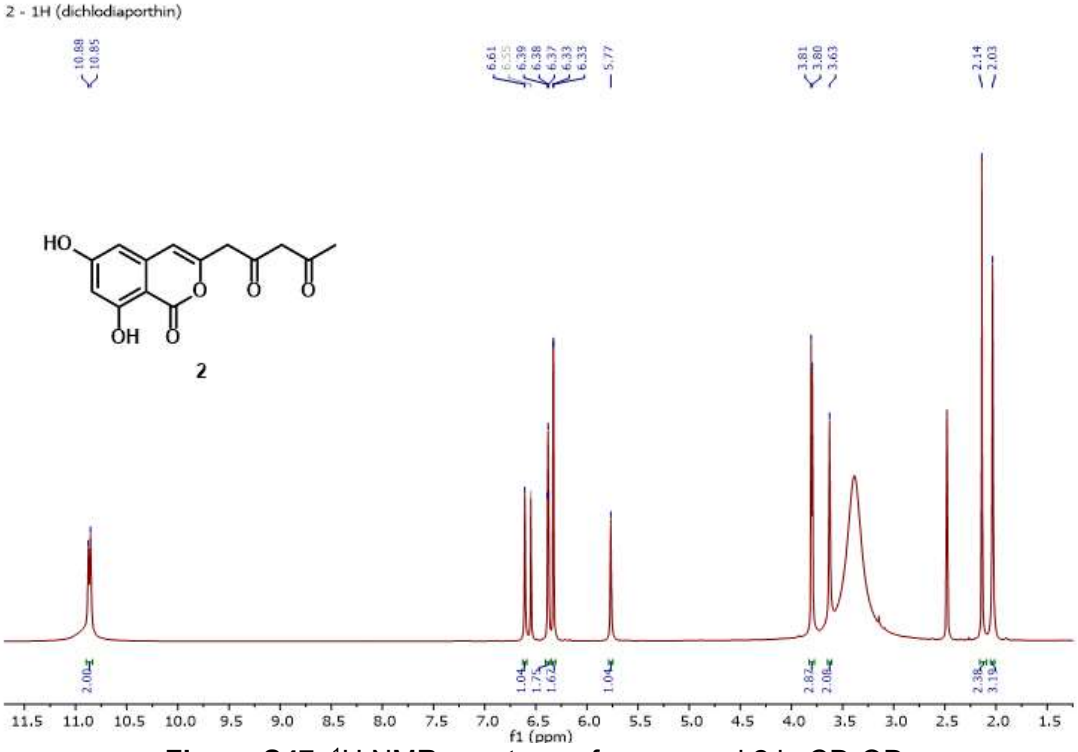


Figure S47. ^1H NMR spectrum of compound **2** in CD_3OD.

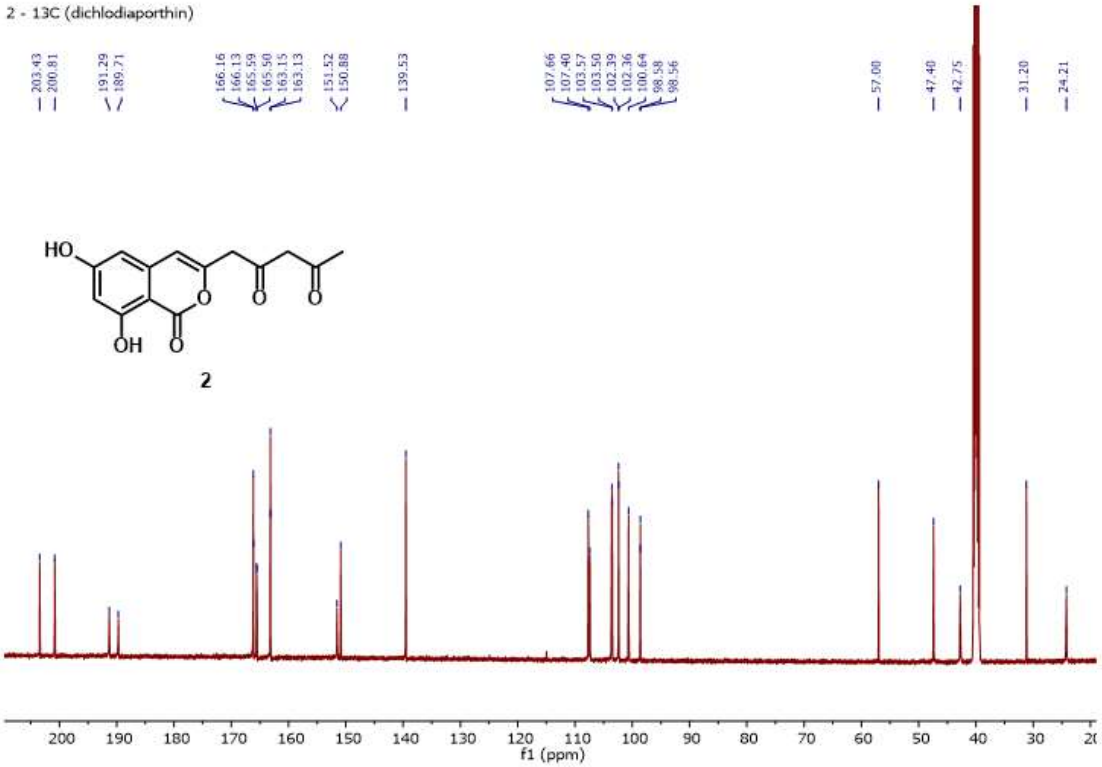


Figure S48. ^{13}C NMR spectrum of compound **2** in CD_3OD.

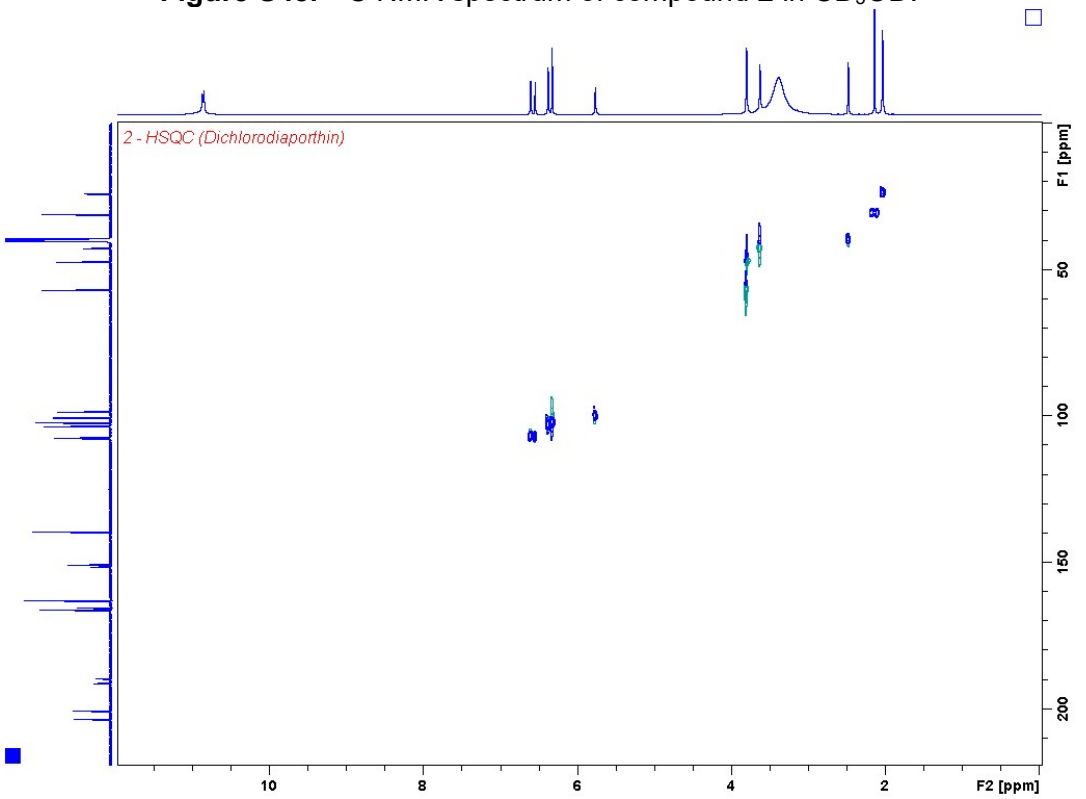


Figure S49. HSQC spectrum of compound **2** in CD_3OD.

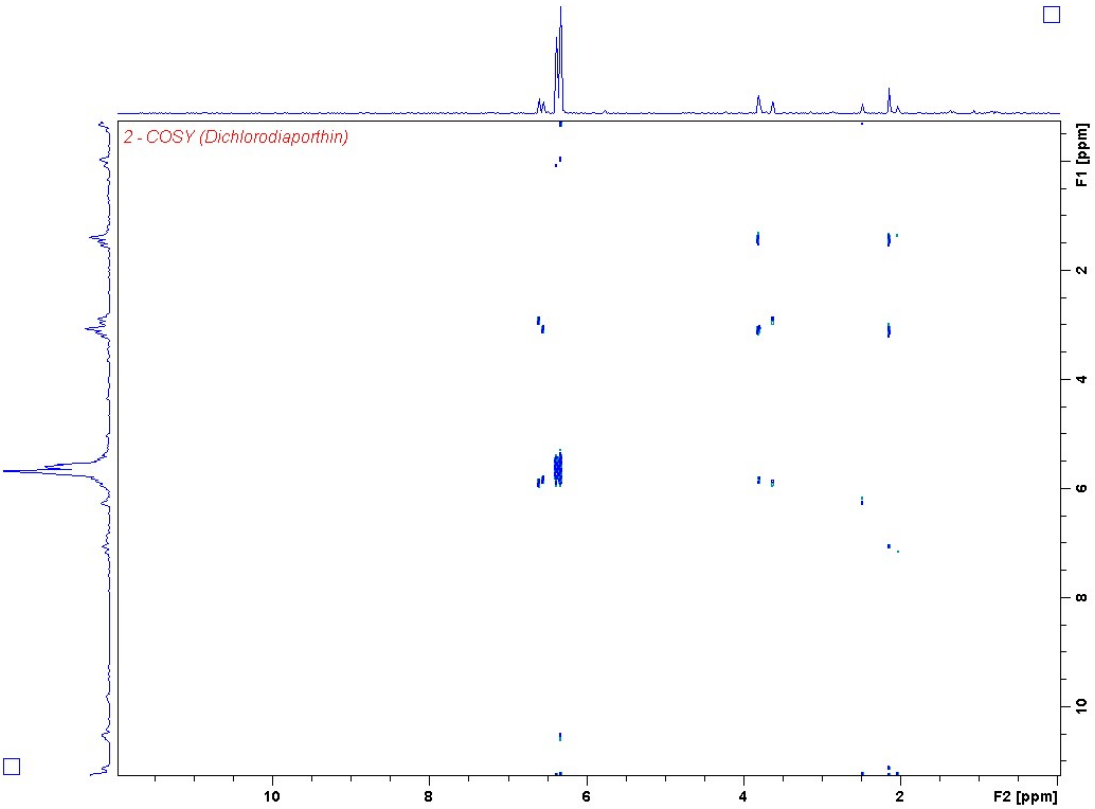


Figure S50. COSY spectrum of compound **2** in CD_3OD.

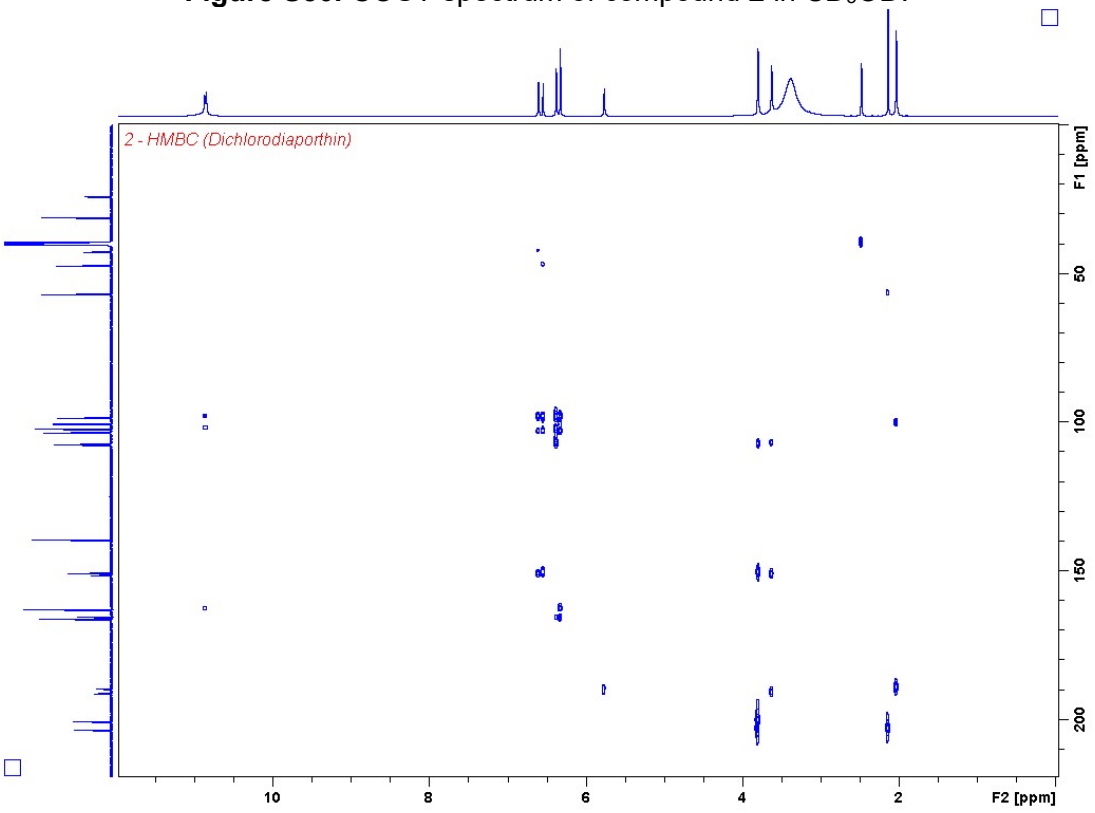


Figure S51. HMBC spectrum of compound **2** in CD_3OD.

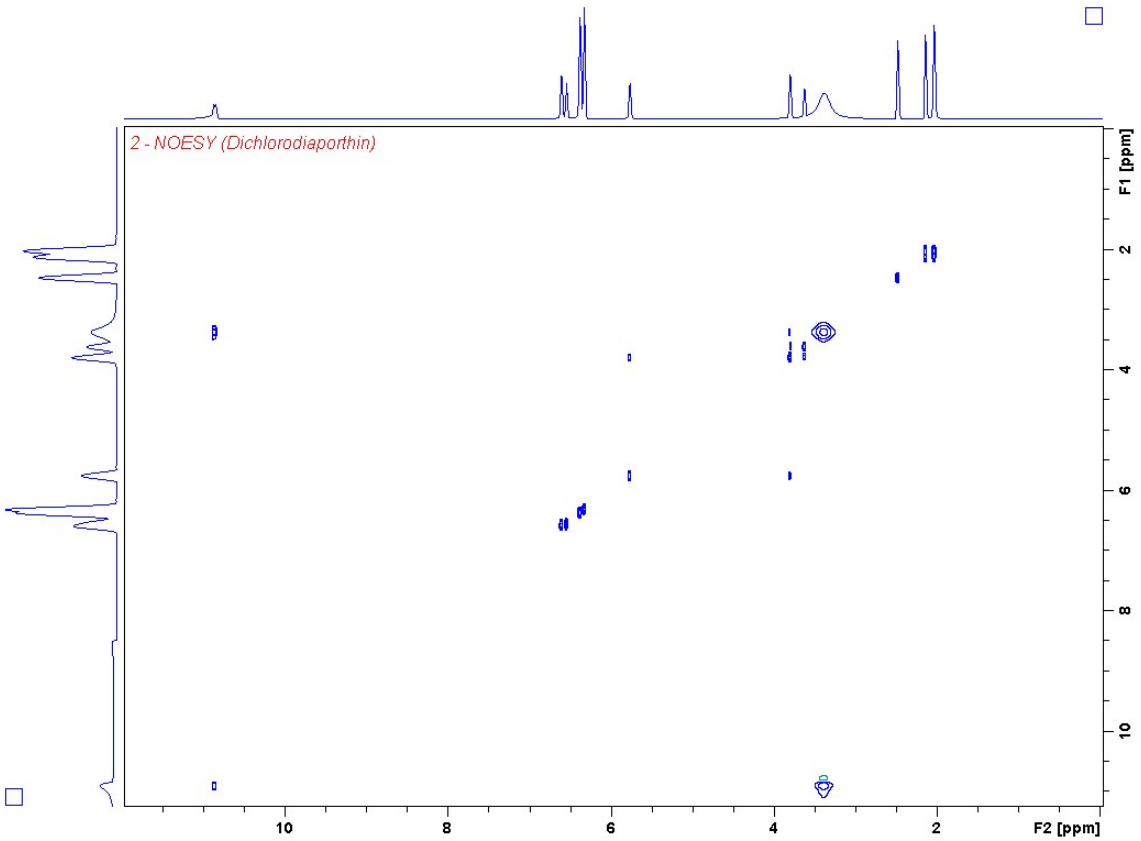


Figure S52. NOESY spectrum of compound **2** in CD_3OD.

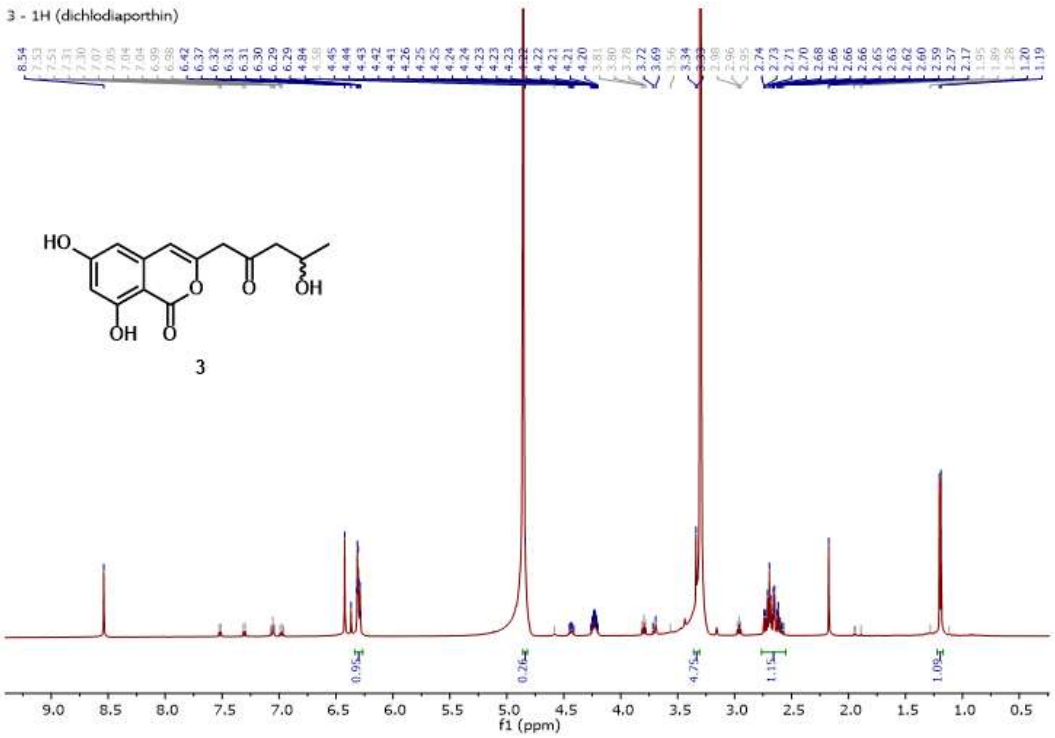


Figure S53. ^1H NMR spectrum of compound **3** in CD_3OD.

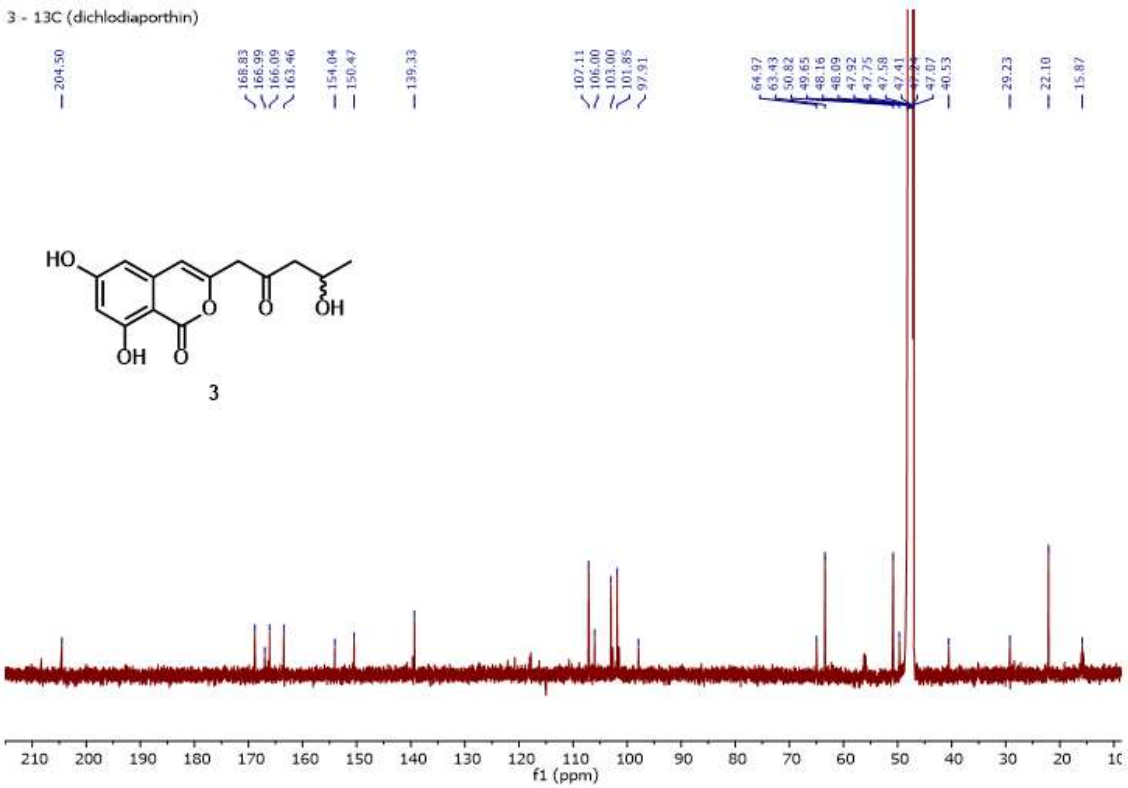


Figure S54. ^{13}C NMR spectrum of compound **3** in CD_3OD.

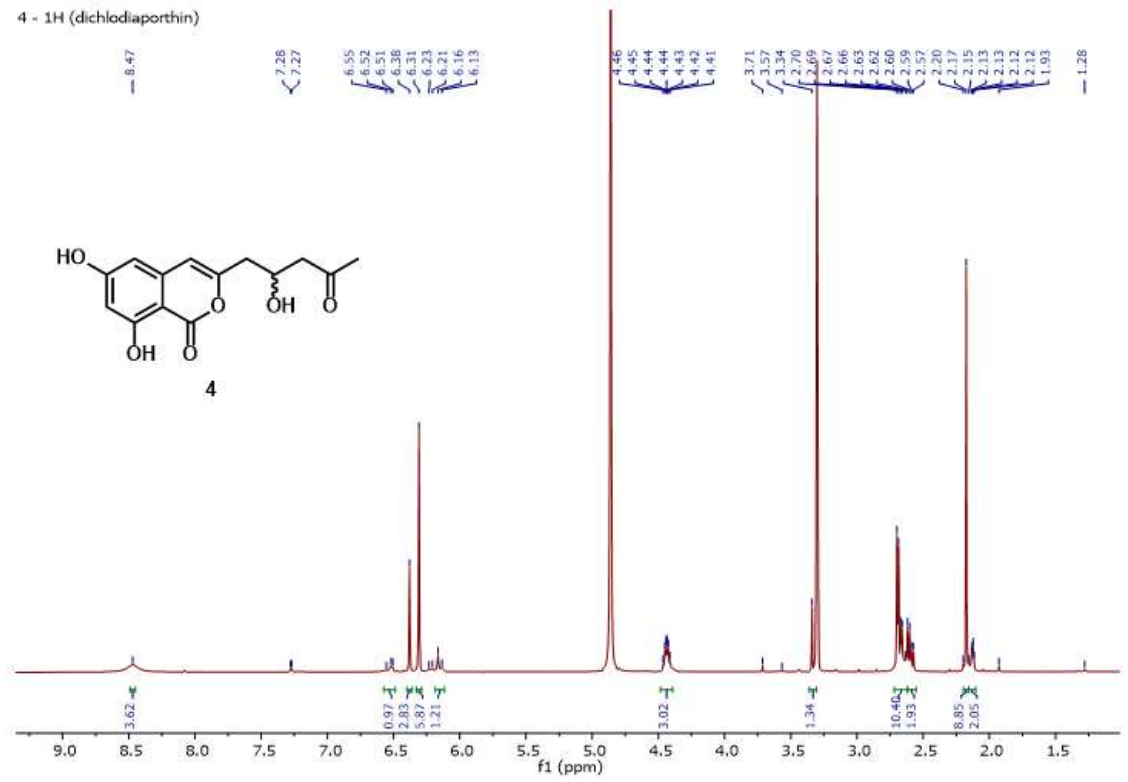


Figure S55. ^{1}H NMR spectrum of compound **4** in CD_3OD.

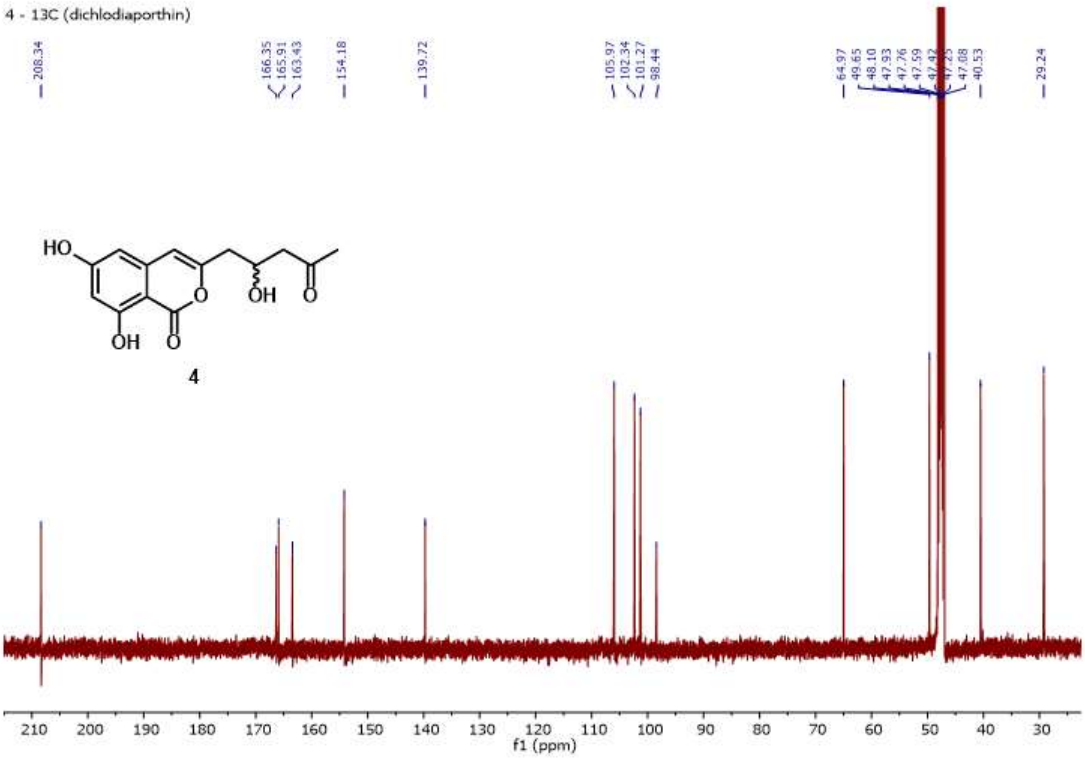


Figure S56. ^{13}C NMR spectrum of compound **4** in CD_3OD.

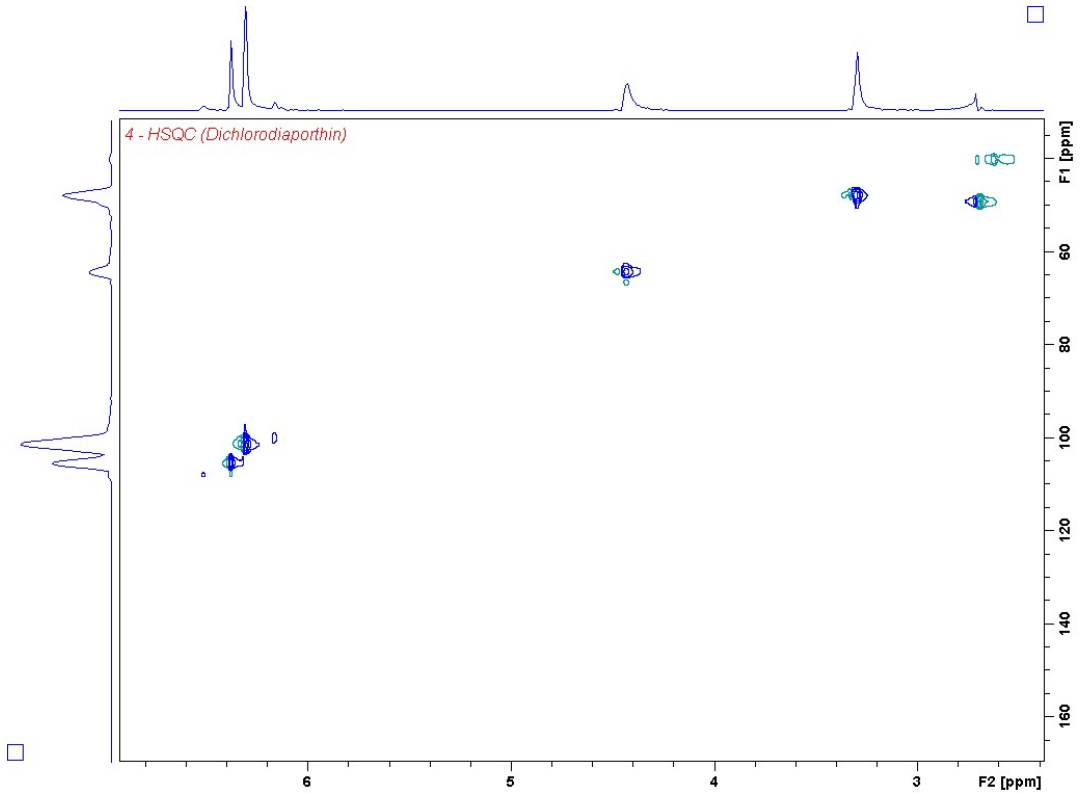


Figure S57. HSQC spectrum of compound **4** in CD_3OD.

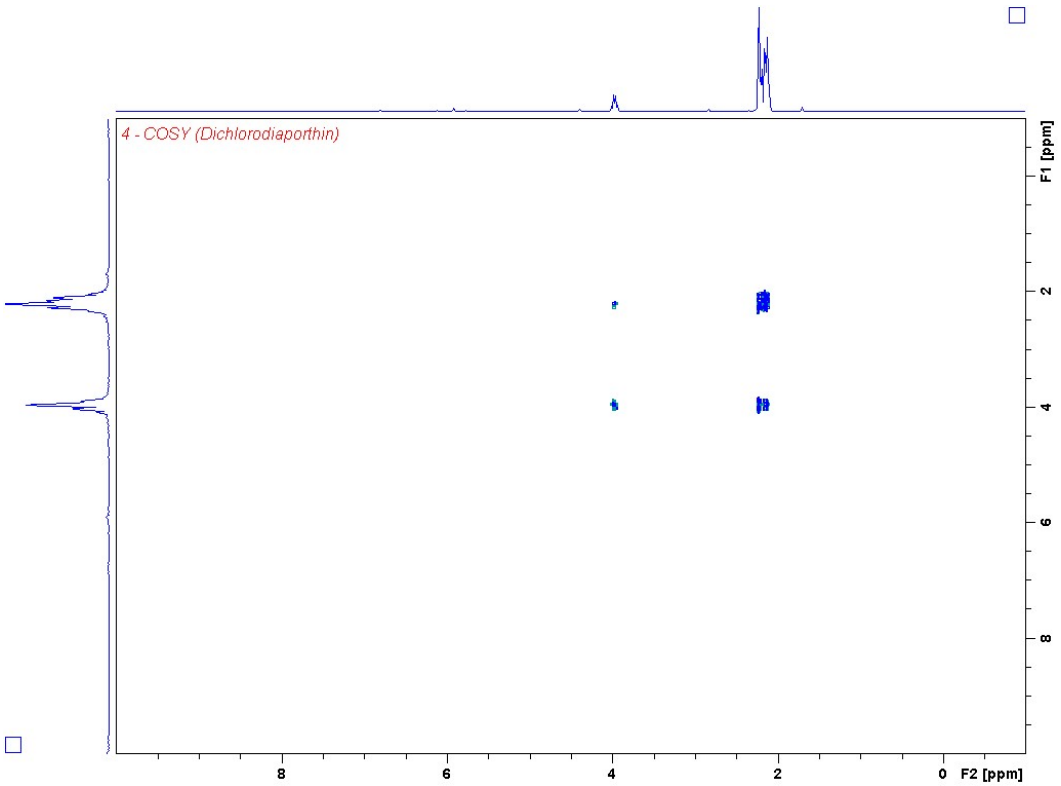


Figure S58. COSY spectrum of compound **4** in CD_3OD.

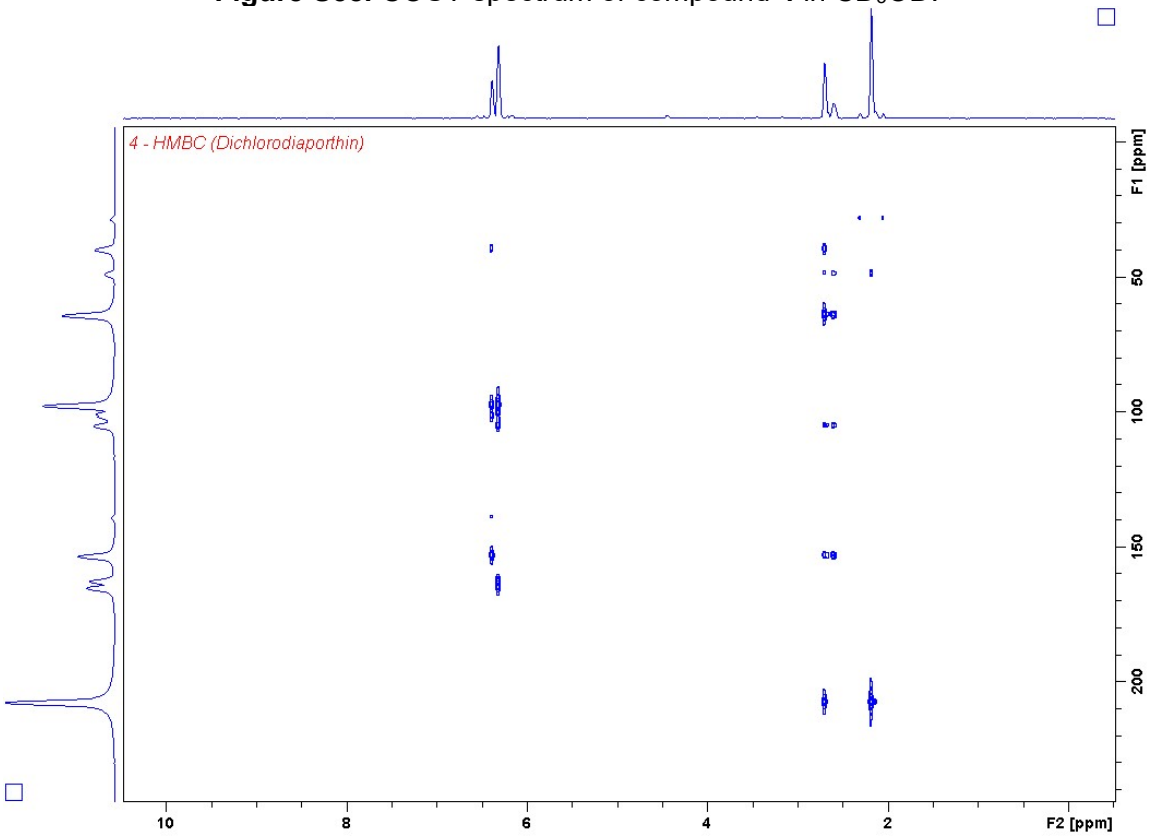


Figure S59. HMBC spectrum of compound **4** in CD_3OD.

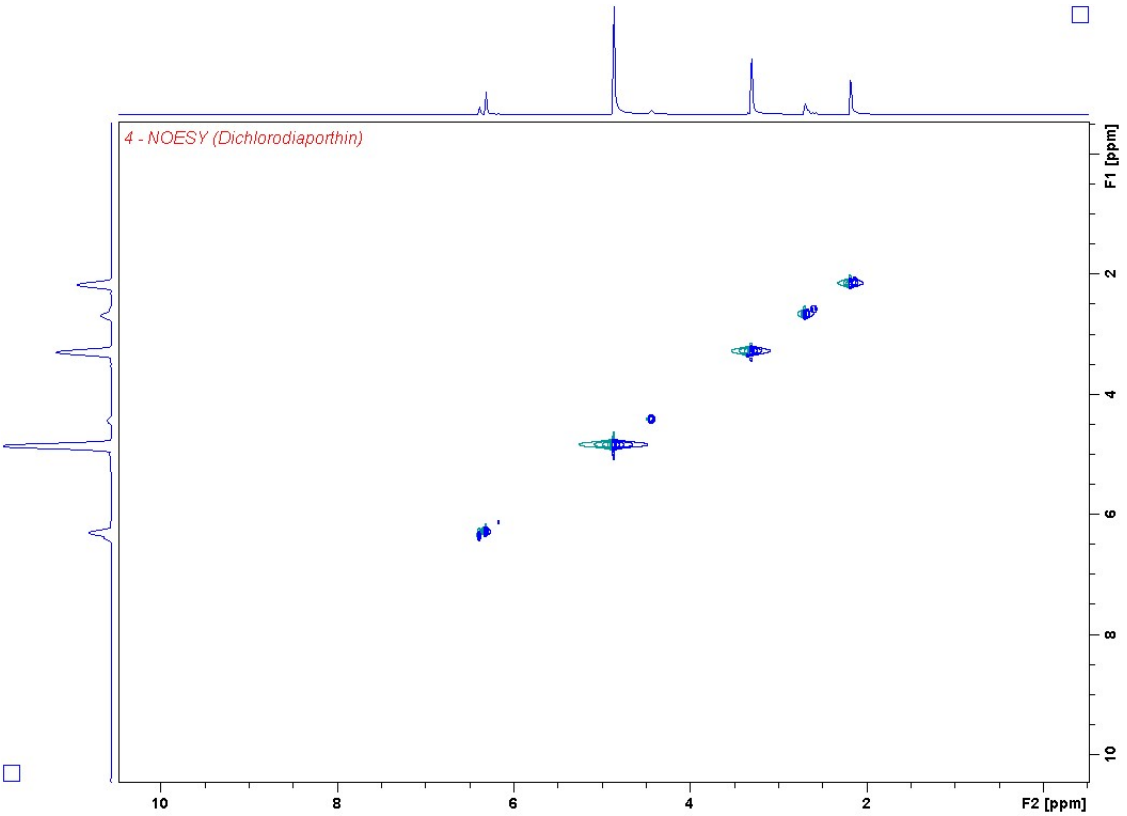


Figure S60. NOESY spectrum of compound **4** in CD_3OD.

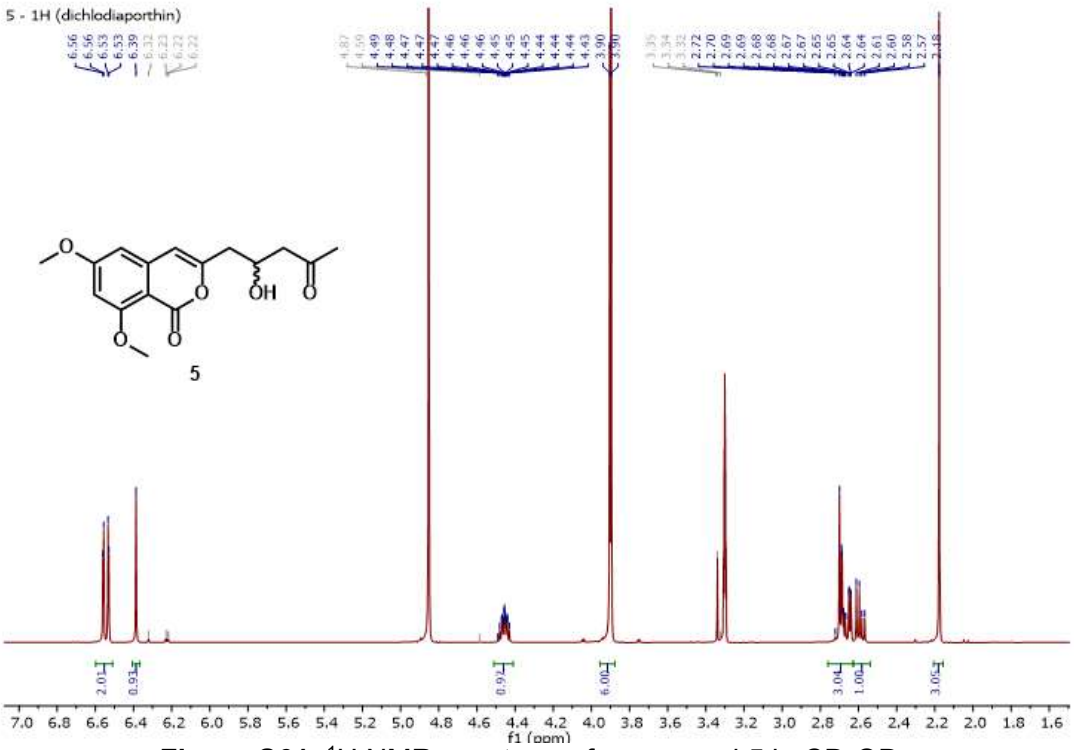


Figure S61. ^1H NMR spectrum of compound **5** in CD_3OD.

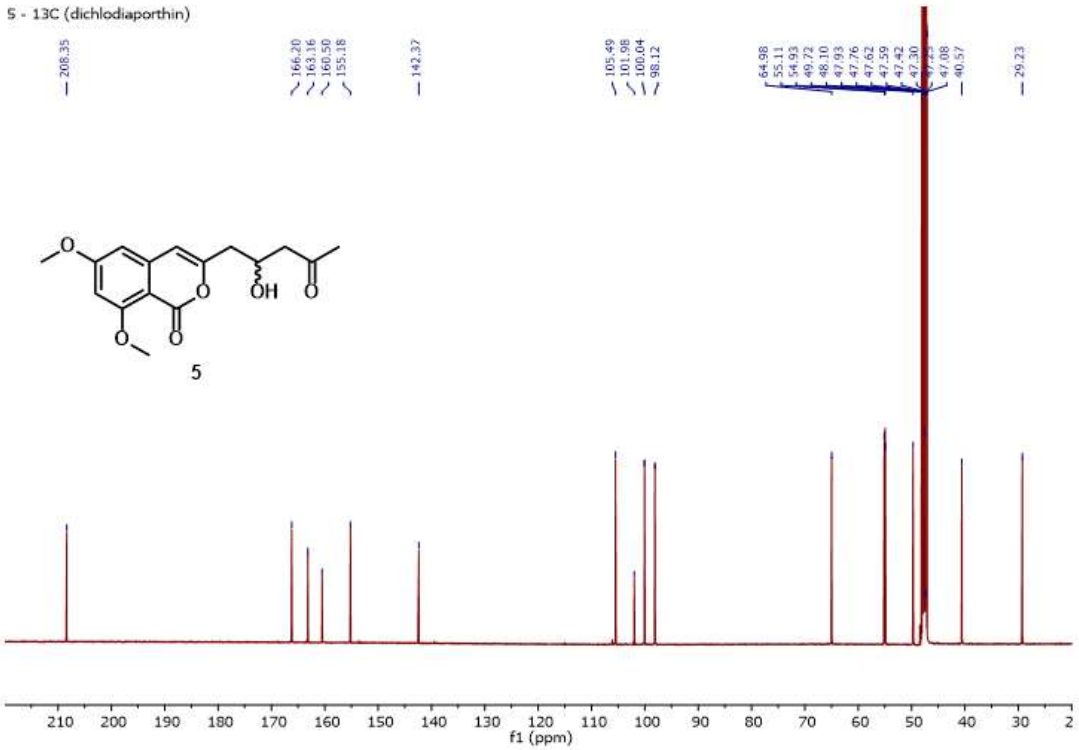


Figure S62. ^{13}C NMR spectrum of compound **5** in CD_3OD.

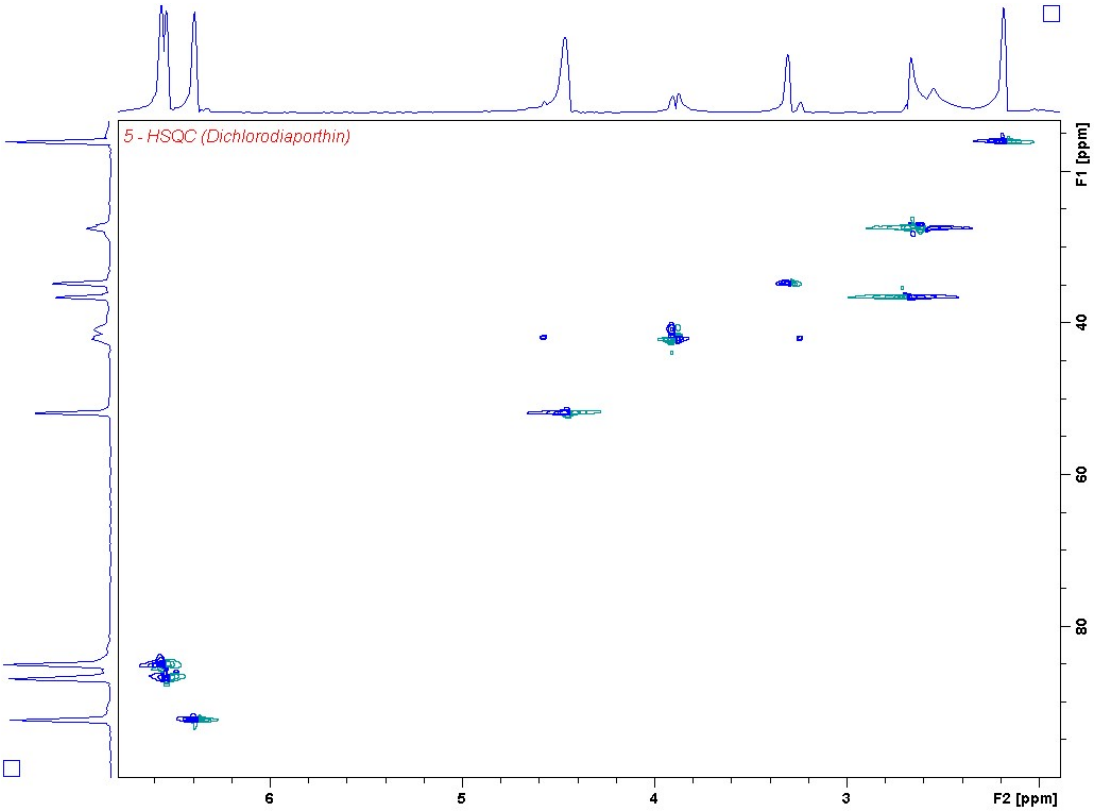


Figure S63. HSQC spectrum of compound **5** in CD_3OD.

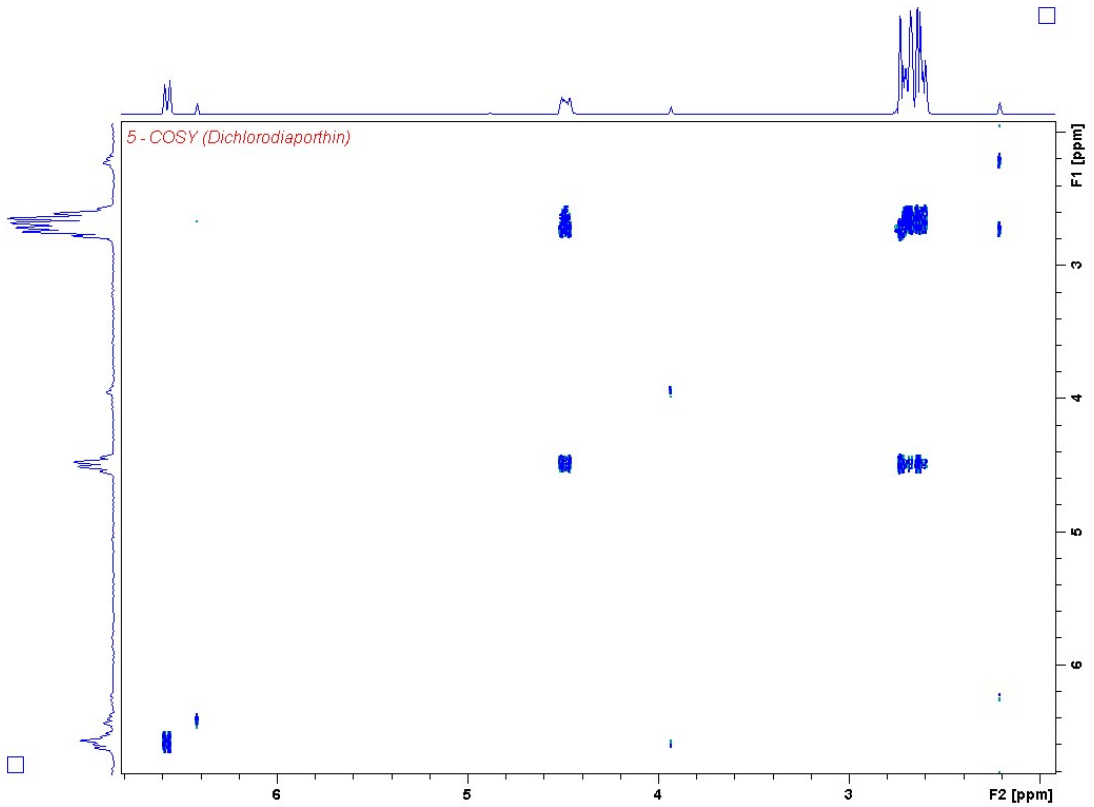


Figure S64. COSY spectrum of compound **5** in CD_3OD.

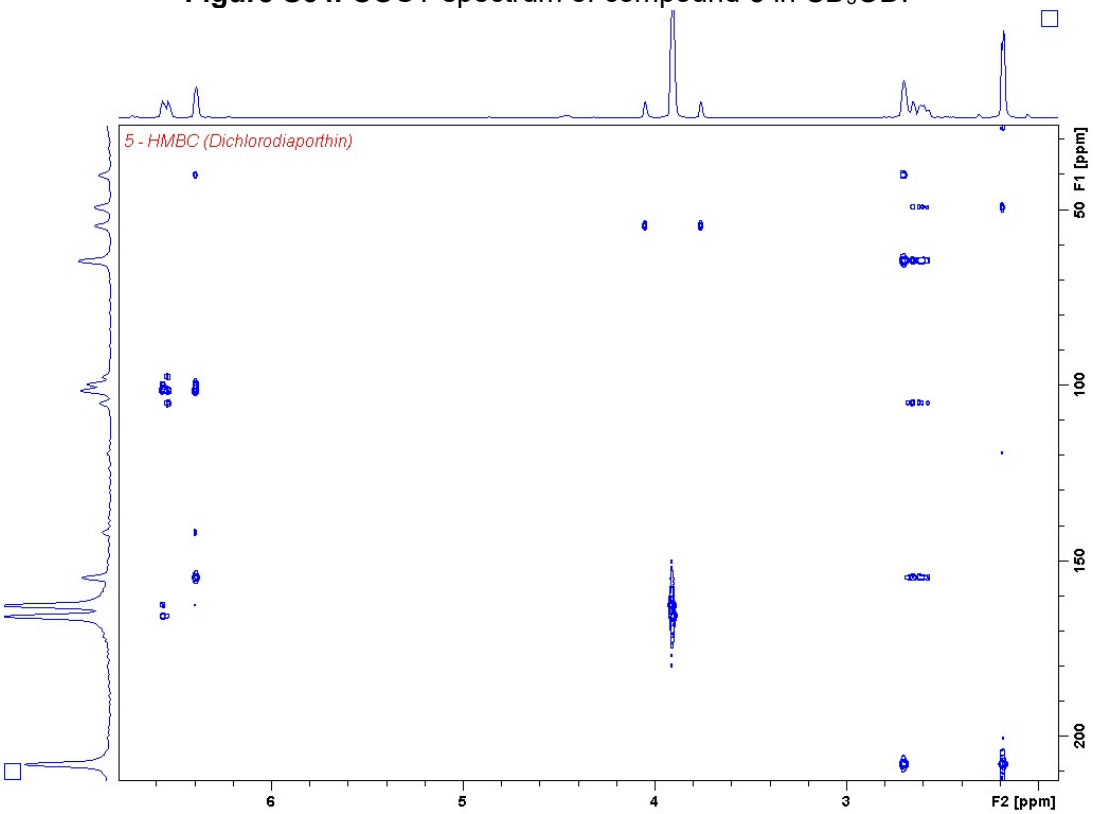


Figure S65. HMBC spectrum of compound **5** in CD_3OD.

9.3 Supplementary information for Section 6

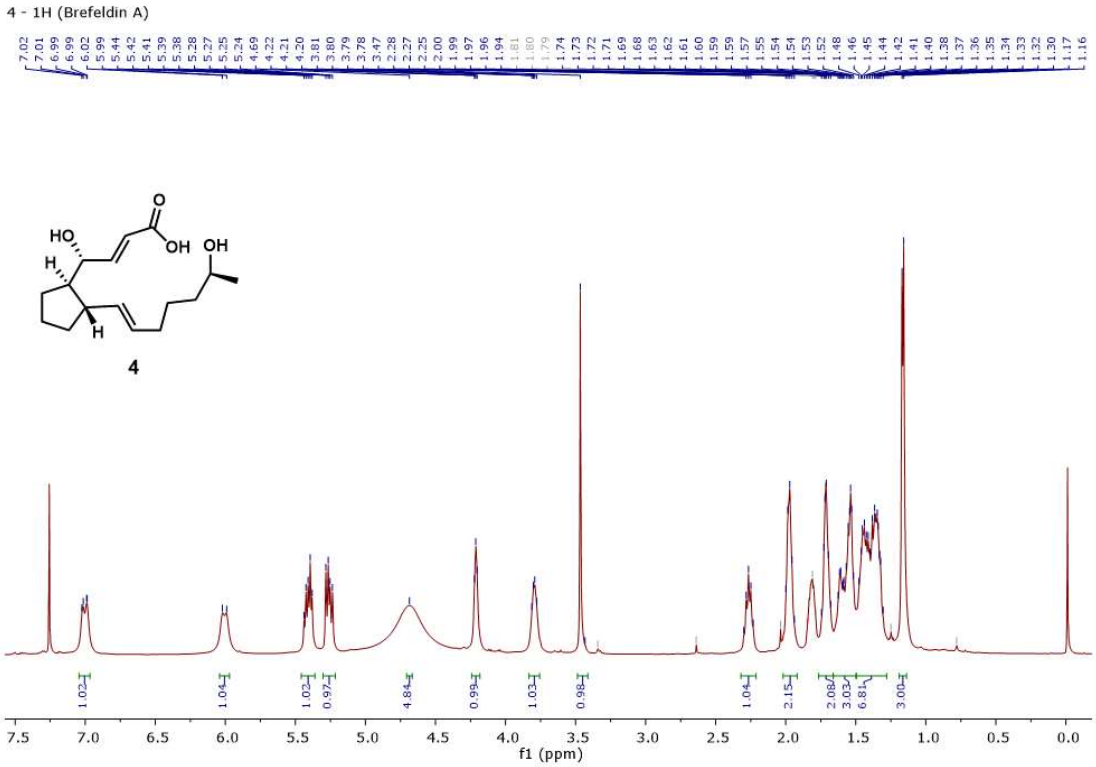


Figure S66. ^{1}H NMR spectrum of compound **4** in $CDCl_3$.

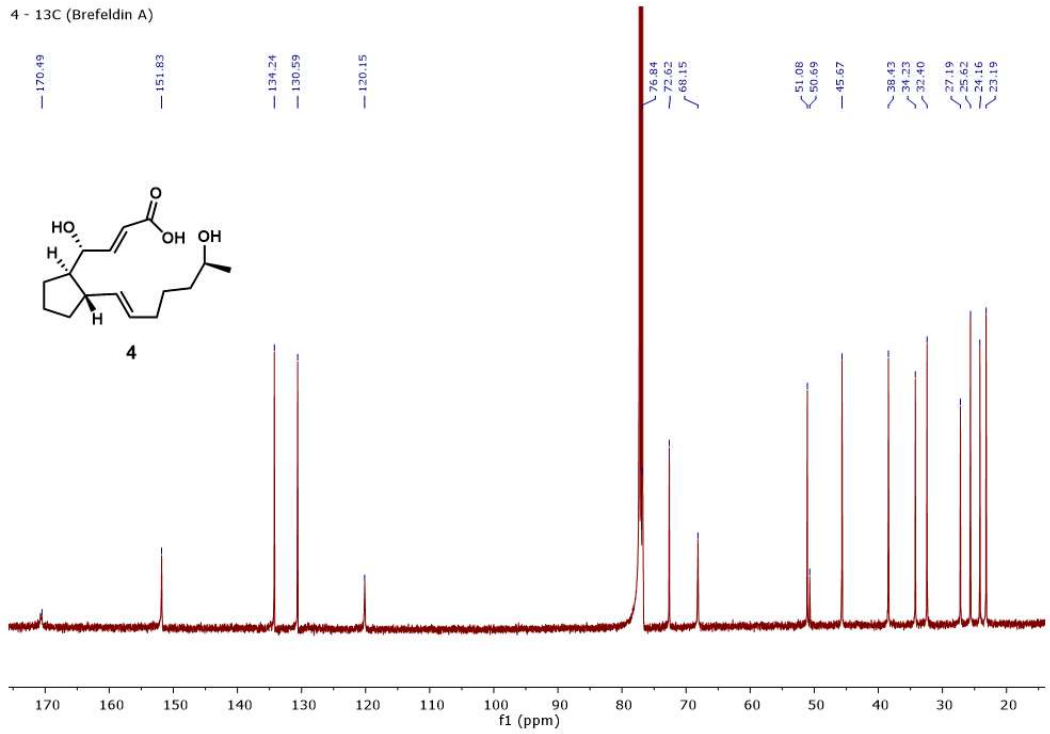


Figure S67. ^{13}C NMR spectrum of compound **4** in $CDCl_3$.

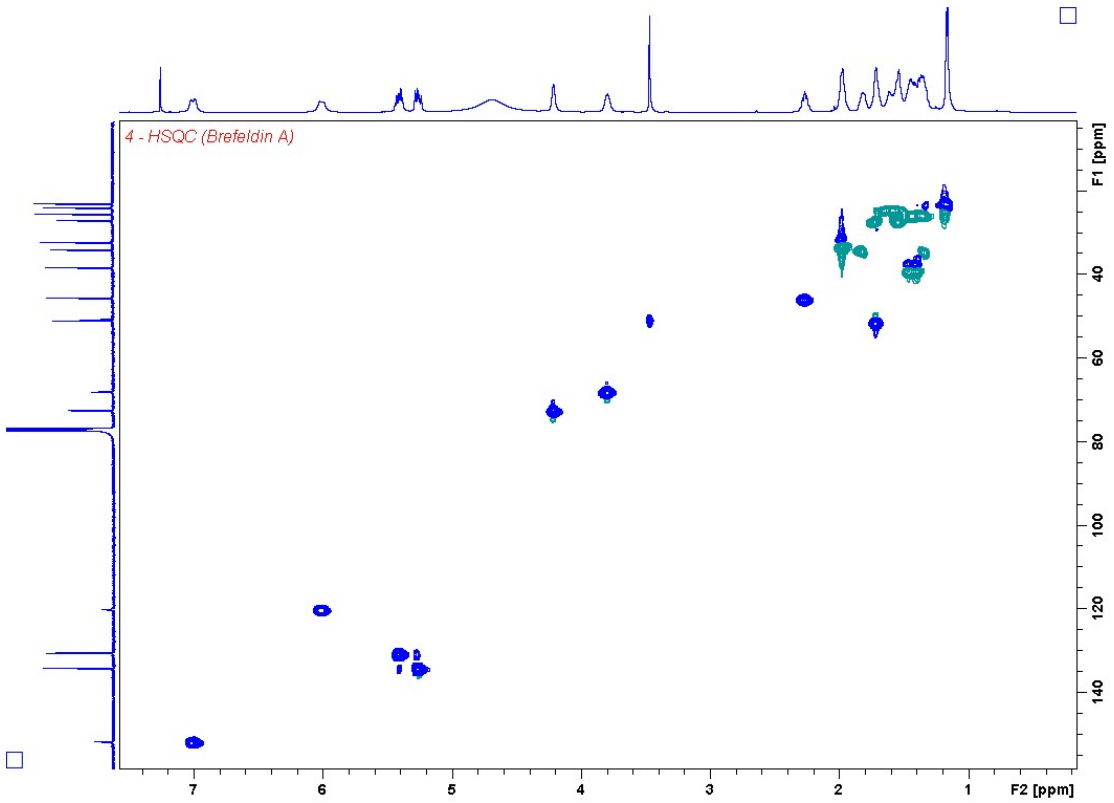


Figure S68. HSQC spectrum of compound **4** in $CDCl_3$.

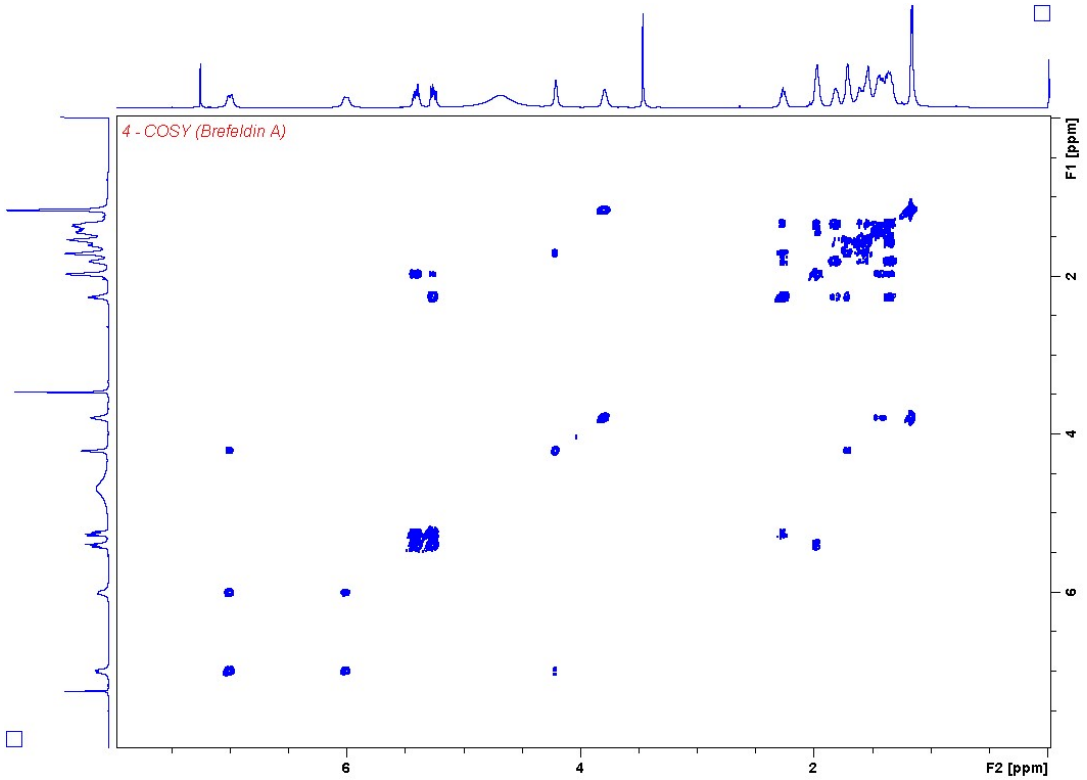


Figure S69. COSY spectrum of compound **4** in $CDCl_3$.

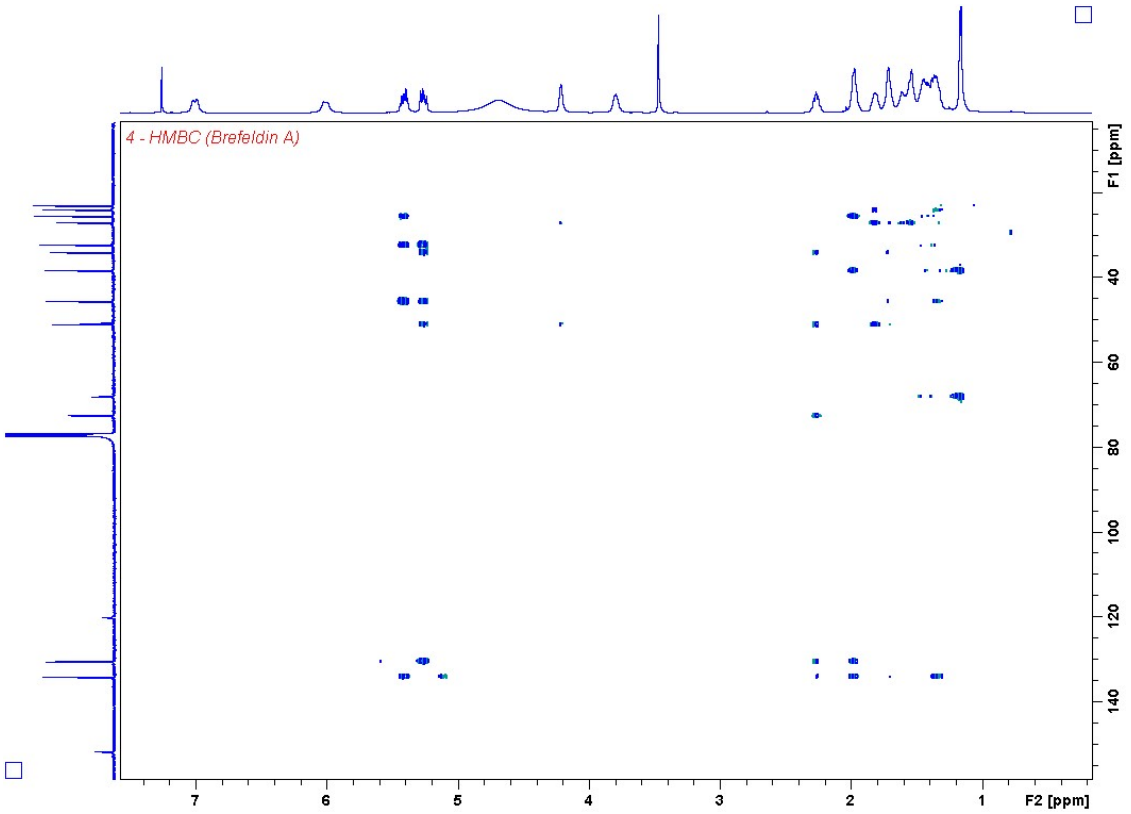


Figure S70. HMBC spectrum of compound **4** in $CDCl_3$.

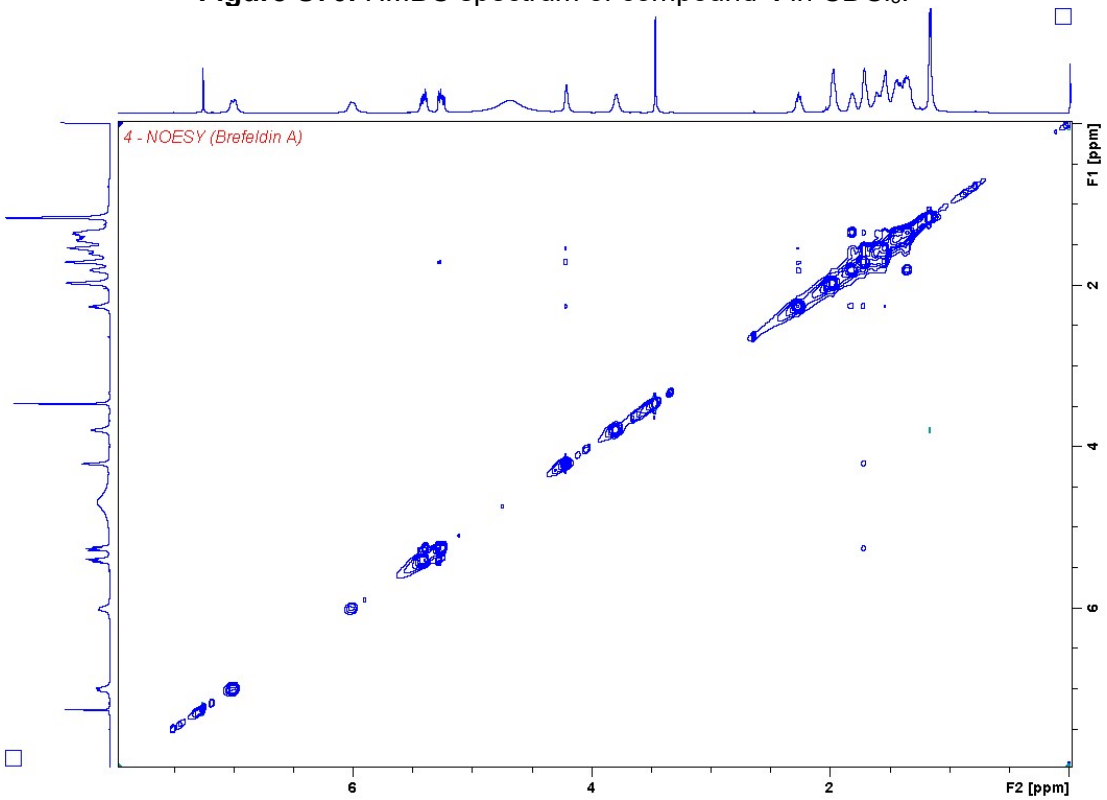


Figure S71. NOESY spectrum of compound **4** in $CDCl_3$.

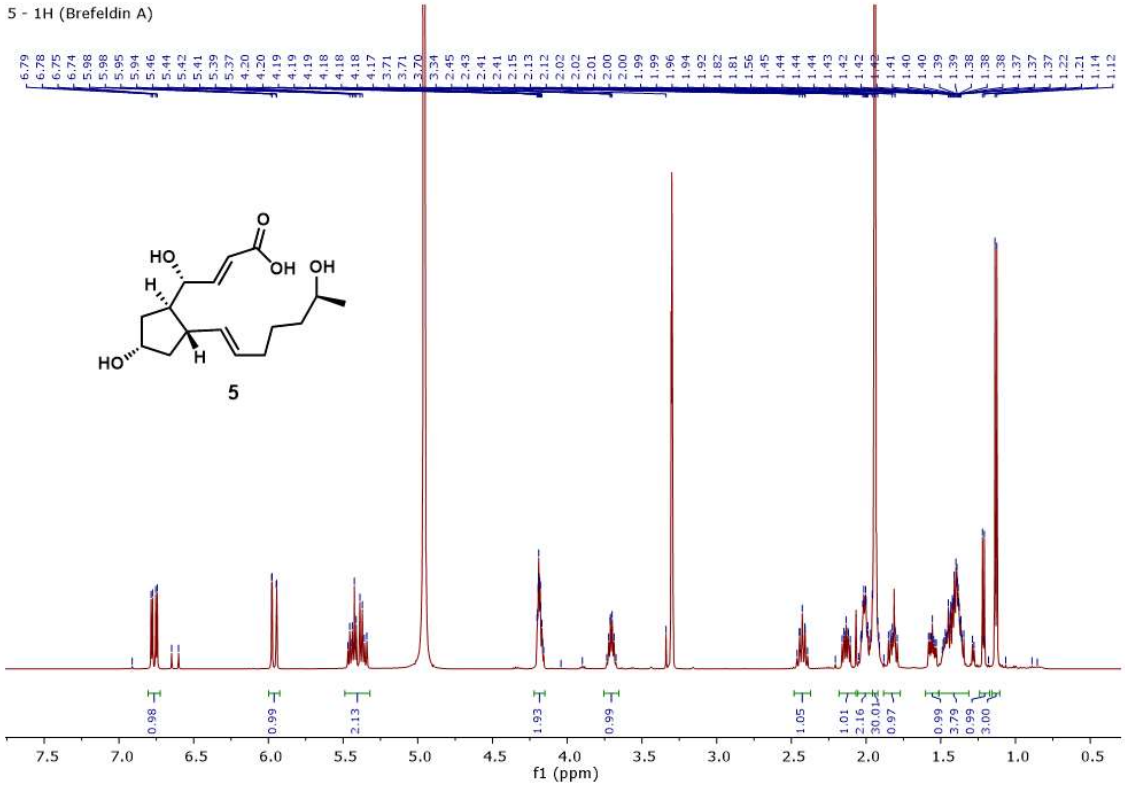


Figure S72. ^{1}H NMR spectrum of compound **5** in CD_3OD.

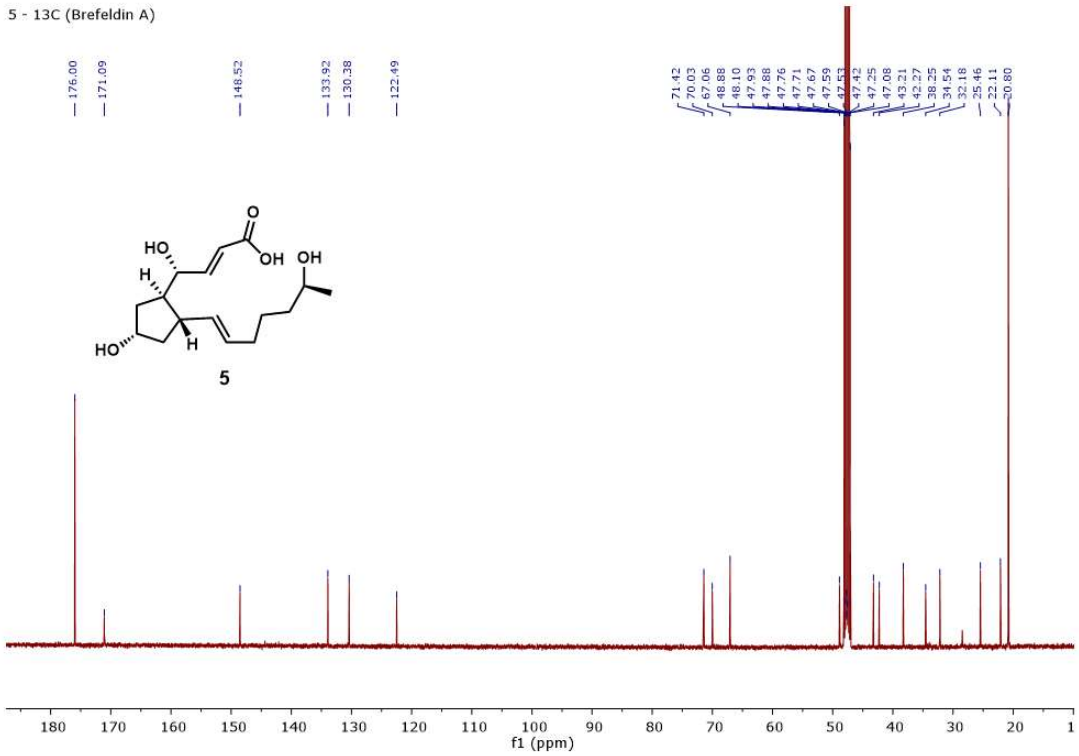


Figure S73. ^{13}C NMR spectrum of compound **5** in CD_3OD.

10. REFERENCES

1. Schueffler, A.; Anke, T., Fungal natural products in research and development. *Natural Product Reports* **2014,** *31* (10), 1425-1448.

2. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Last 25 Years. *Journal of Natural Products* **2007,** *70* (3), 461-477.

3. Hawksworth, D. L.; Lucking, R., Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiology spectrum* **2017,** *5* (4).

4. Koehn, F. E.; Carter, G. T., The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* **2005,** *4* (3), 206-20.

5. Projan, S. J., Why is big Pharma getting out of antibacterial drug discovery? *Current opinion in microbiology* **2003,** *6* (5), 427-430.

6. Consortium, I. H. G. S.; Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendl, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J.-F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M.; Gibbs, R. A.; Muzny, D. M.; Scherer, S. E.; Bouck, J. B.; Sodergren, E. J.; Worley, K. C.; Rives, C. M.; Gorrell, J. H.; Metzker, M. L.; Naylor, S. L.; Kucherlapati, R. S.; Nelson, D. L.; Weinstock, G. M.; Sakaki, Y.; Fujiyama, A.; Hattori, M.; Yada, T.; Toyoda, A.; Itoh, T.; Kawagoe, C.; Watanabe, H.; Totoki, Y.; Taylor, T.; Weissenbach, J.; Heilig, R.; Saurin, W.; Artiguenave, F.; Brottier, P.; Bruls, T.; Pelletier, E.; Robert, C.; Wincker, P.; Rosenthal, A.; Platzer, M.; Nyakatura, G.; Taudien, S.; Rump, A.; Smith, D. R.; Doucette-Stamm, L.; Rubenfield, M.; Weinstock, K.; Lee, H. M.; Dubois, J.; Yang, H.; Yu, J.; Wang, J.; Huang, G.; Gu, J.; Hood, L.; Rowen, L.; Madan, A.; Qin, S.; Davis, R. W.; Federspiel, N. A.; Abola, A. P.; Proctor, M. J.; Roe, B. A.; Chen, F.; Pan, H.; Ramser, J.; Lehrach, H.; Reinhardt, R.; McCombie, W. R.; de la Bastide, M.; Dedhia, N.; Blöcker, H.; Hornischer, K.; Nordsiek, G.; Agarwala, R.; Aravind, L.; Bailey, J. A.; Bateman, A.; Batzoglou, S.; Birney, E.; Bork, P.; Brown, D. G.; Burge, C. B.; Cerutti, L.; Chen, H.-C.; Church, D.; Clamp, M.; Copley, R. R.; Doerks, T.; Eddy, S. R.; Eichler, E. E.; Furey, T. S.; Galagan, J.; Gilbert, J. G. R.; Harmon, C.; Hayashizaki, Y.; Haussler, D.; Hermjakob, H.; Hokamp, K.; Jang, W.; Johnson, L. S.; Jones, T. A.; Kasif, S.; Kaspryzk, A.; Kennedy, S.; Kent, W. J.; Kitts, P.; Koonin, E. V.; Korf, I.; Kulp, D.; Lancet, D.; Lowe, T. M.; McLysaght, A.; Mikkelsen, T.; Moran, J. V.; Mulder, N.; Pollara, V. J.; Ponting, C. P.; Schuler, G.; Schultz, J.; Slater, G.; Smit, A. F. A.; Stupka, E.; Szustakowki, J.; Thierry-

Mieg, D.; Thierry-Mieg, J.; Wagner, L.; Wallis, J.; Wheeler, R.; Williams, A.; Wolf, Y. I.; Wolfe, K. H.; Yang, S.-P.; Yeh, R.-F.; Collins, F.; Guyer, M. S.; Peterson, J.; Felsenfeld, A.; Wetterstrand, K. A.; Myers, R. M.; Schmutz, J.; Dickson, M.; Grimwood, J.; Cox, D. R.; Olson, M. V.; Kaul, R.; Raymond, C.; Shimizu, N.; Kawasaki, K.; Minoshima, S.; Evans, G. A.; Athanasiou, M.; Schultz, R.; Patrinos, A.; Morgan, M. J., Initial sequencing and analysis of the human genome. *Nature* **2001,** *409*, 860-860.

7. Martinez, D.; Larrondo, L. F.; Putnam, N.; Sollewijn Gelpke, M. D.; Huang, K.; Chapman, J.; Helfenbein, K. G.; Ramaiya, P.; Detter, J. C.; Larimer, F.; Coutinho, P. M.; Henrissat, B.; Berka, R.; Cullen, D.; Rokhsar, D., Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. *Nature Biotechnology* **2004,** *22* (6), 695-700.

8. 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 Research* **2014,** *42* (D1), 26-32.

9. Galagan, J. E.; Calvo, S. E.; Borkovich, K. A.; Selker, E. U.; Read, N. D.; Jaffe, D.; FitzHugh, W.; Ma, L.-J.; Smirnov, S.; Purcell, S.; Rehman, B.; Elkins, T.; Engels, R.; Wang, S.; Nielsen, C. B.; Butler, J.; Endrizzi, M.; Qui, D.; Ianakiev, P.; Bell-Pedersen, D.; Nelson, M. A.; Werner-Washburne, M.; Selitrennikoff, C. P.; Kinsey, J. A.; Braun, E. L.; Zelter, A.; Schulte, U.; Kothe, G. O.; Jedd, G.; Mewes, W.; Staben, C.; Marcotte, E.; Greenberg, D.; Roy, A.; Foley, K.; Naylor, J.; Stange-Thomann, N.; Barrett, R.; Gnerre, S.; Kamal, M.; Kamvysselis, M.; Mauceli, E.; Bielke, C.; Rudd, S.; Frishman, D.; Krystofova, S.; Rasmussen, C.; Metzenberg, R. L.; Perkins, D. D.; Kroken, S.; Cogoni, C.; Macino, G.; Catcheside, D.; Li, W.; Pratt, R. J.; Osmani, S. A.; DeSouza, C. P. C.; Glass, L.; Orbach, M. J.; Berglund, J. A.; Voelker, R.; Yarden, O.; Plamann, M.; Seiler, S.; Dunlap, J.; Radford, A.; Aramayo, R.; Natvig, D. O.; Alex, L. A.; Mannhaupt, G.; Ebbole, D. J.; Freitag, M.; Paulsen, I.; Sachs, M. S.; Lander, E. S.; Nusbaum, C.; Birren, B., The genome sequence of the filamentous fungus Neurospora crassa. *Nature* **2003,** *422* (6934), 859-868.

10. Keller, N. P.; Turner, G.; Bennett, J. W., Fungal secondary metabolism - from biochemistry to genomics. *Nature reviews. Microbiology* **2005,** *3* (12), 937-947.

11. Li, Y. F.; Tsai, K. J. S.; Harvey, C. J. B.; Li, J. J.; Ary, B. E.; Berlew, E. E.; Boehman, B. L.; Findley, D. M.; Friant, A. G.; Gardner, C. A.; Gould, M. P.; Ha, J. H.; Lilley, B. K.; McKinstry, E. L.; Nawal, S.; Parry, R. C.; Rothchild, K. W.; Silbert, S. D.; Tentilucci, M. D.; Thurston, A. M.; Wai, R. B.; Yoon, Y.; Aiyar, R. S.; Medema, M. H.; Hillenmeyer, M. E.; Charkoudian, L. K., Comprehensive curation and analysis of fungal biosynthetic gene clusters of published natural products. *Fungal Genetics and Biology* **2016,** *89*, 18-28.

12. Walsh, C. T.; Tang, Y., *Natural Product Biosynthesis*. The Royal Society of Chemistry: 2017.

13. Challis, G. L., Genome mining for novel natural product discovery. *Journal of medicinal chemistry* **2008,** *51* (9), 2618-2628.

14. Griengl, H.; Schwab, H.; Fechter, M., The synthesis of chiral cyanohydrins by oxynitrilases. *Trends in biotechnology* **2000,** *18* (6), 252-256.

15. Hills, G., Industrial use of lipases to produce fatty acid esters. *European journal of lipid science and technology* **2003,** *105* (10), 601-607.

16. Sheldon, R. A.; Woodley, J. M., Role of Biocatalysis in Sustainable Chemistry. *Chemical Reviews* **2018,** *118* (2), 801-838.

17. Arnold, F. H., The nature of chemical innovation: new enzymes by evolution. *Quarterly Reviews of Biophysics* **2015,** *48* (4), 404-410.

18. Arnold, F. H., Directed Evolution: Bringing New Chemistry to Life. *Angewandte Chemie International Edition* **2018,** *57* (16), 4143-4148.

19. Emmanuel, M. A.; Greenberg, N. R.; Oblinsky, D. G.; Hyster, T. K., Accessing non-natural reactivity by irradiating nicotinamide-dependent enzymes with light. *Nature* **2016,** *540* (7633), 414-417.

20. Schrittwieser, J. H.; Velikogne, S.; Hall, M.; Kroutil, W., Artificial Biocatalytic Linear Cascades for Preparation of Organic Molecules. *Chemical Reviews* **2018,** *118* (1), 270-348.

21. France, S. P.; Hepworth, L. J.; Turner, N. J.; Flitsch, S. L., Constructing Biocatalytic Cascades: In Vitro and in Vivo Approaches to de Novo Multi-Enzyme Pathways. *ACS Catalysis* **2017,** *7* (1), 710-724.

22. Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; Halsey, H. M.; Hughes, G. J.; Jo, J.; Joyce, L. A.; Kolev, J. N.; Liang, J.; Maloney, K. M.; Mann, B. F.; Marshall, N. M.; McLaughlin, M.; Moore, J. C.; Murphy, G. S.; Nawrat, C. C.; Nazor, J.; Novick, S.; Patel, N. R.; Rodriguez-Granillo, A.; Robaire, S. A.; Sherer, E. C.; Truppo, M. D.; Whittaker, A. M.; Verma, D.; Xiao, L.; Xu, Y.; Yang, H., Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **2019,** *366* (6470), 1255-1259.

23. Cacho, R. A.; Tang, Y.; Chooi, Y.-H., Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Frontiers in Microbiology* **2015,** *5* (January), 774-774.

24. 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., MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Research* **2014,** *42* (D1), 699-704.

25. Medema, M. H.; Fischbach, M. A., Computational approaches to natural product discovery. *Nature Chemical Biology* **2015,** *11* (9), 639-648.

26. Marchler-Bauer, A.; Bryant, S. H., CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research* **2004,** *32* (WEB SERVER ISS.), 327-331.

27. Stanke, M.; Steinkamp, R.; Waack, S.; Morgenstern, B., AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Research* **2004,** *32* (WEB SERVER ISS.), 309-312.

28. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D., The Pfam protein families database in 2019. *Nucleic acids research* **2018**.

29. Medema, M. H.; Blin, K.; Cimermancic, P.; De Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R., AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research* **2011,** *39* (SUPPL. 2), 339-346.

30. Khaldi, N.; Seifuddin, F. T.; Turner, G.; Haft, D.; Nierman, W. C.; Wolfe, K. H.; Fedorova, N. D., SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genetics and Biology* **2010,** *47* (9), 736-741.

31. Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A., Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Research* **2015,** *43* (20), 9645-9662.

32. Wolf, T.; Shelest, V.; Nath, N.; Shelest, E., CASSIS and SMIPS: Promoter-based prediction of secondary metabolite gene clusters in eukaryotic genomes. *Bioinformatics* **2016,** *32* (8), 1138-1143.

33. Blin, K.; Kim, H. U.; Medema, M. H.; Weber, T., Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Briefings in Bioinformatics* **2017,** (October), 1-11.

34. Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; De Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Düsterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J. N.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kötter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. A.; Moore, M.; Moss, N.; Nützmann, H. W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gómez, F.; Bechthold, A.; Bode, H. B.; Borriss, R.; Brady, S. F.; Brakhage, A. A.; Caffrey, P.; Cheng, Y. Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; Van Der Donk, W. A.; Dorrestein, P. C.; Doyle, S.; Driessen, A. J. M.; Ehling-Schulz, M.; Entian, K. D.; Fischbach, M. A.; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Höfte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kyrpides, N. C.; Kwon, H. J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linquan, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Méndez, C.; Metsä-Ketelä, M.; Micklefield, J.; Mitchell, D. A.; Moore, B. S.; Moreira, L. M.; Müller, R.; Neilan, B. A.; Nett, M.; Nielsen, J.; O'Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J. L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. A.; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Süssmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; Van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev,

S. B.; Breitling, R.; Takano, E.; Glöckner, F. O., Minimum Information about a Biosynthetic Gene cluster. *Nature Chemical Biology* **2015,** *11* (9), 625-631.

35. Yilmaz, P.; Kottmann, R.; Field, D.; Knight, R.; Cole, J. R.; Amaral-Zettler, L.; Gilbert, J. A.; Karsch-Mizrachi, I.; Johnston, A.; Cochrane, G.; Vaughan, R.; Hunter, C.; Park, J.; Morrison, N.; Rocca-Serra, P.; Sterk, P.; Arumugam, M.; Bailey, M.; Baumgartner, L.; Birren, B. W.; Blaser, M. J.; Bonazzi, V.; Booth, T.; Bork, P.; Bushman, F. D.; Buttigieg, P. L.; Chain, P. S. G.; Charlson, E.; Costello, E. K.; Huot-Creasy, H.; Dawyndt, P.; Desantis, T.; Fierer, N.; Fuhrman, J. A.; Gallery, R. E.; Gevers, D.; Gibbs, R. A.; Gil, I. S.; Gonzalez, A.; Gordon, J. I.; Guralnick, R.; Hankeln, W.; Highlander, S.; Hugenholtz, P.; Jansson, J.; Kau, A. L.; Kelley, S. T.; Kennedy, J.; Knights, D.; Koren, O.; Kuczynski, J.; Kyrpides, N.; Larsen, R.; Lauber, C. L.; Legg, T.; Ley, R. E.; Lozupone, C. A.; Ludwig, W.; Lyons, D.; Maguire, E.; Methé, B. A.; Meyer, F.; Muegge, B.; Nakielny, S.; Nelson, K. E.; Nemergut, D.; Neufeld, J. D.; Newbold, L. K.; Oliver, A. E.; Pace, N. R.; Palanisamy, G.; Peplies, J.; Petrosino, J.; Proctor, L.; Pruesse, E.; Quast, C.; Raes, J.; Ratnasingham, S.; Ravel, J.; Relman, D. A.; Assunta-Sansone, S.; Schloss, P. D.; Schriml, L.; Sinha, R.; Smith, M. I.; Sodergren, E.; Spor, A.; Stombaugh, J.; Tiedje, J. M.; Ward, D. V.; Weinstock, G. M.; Wendel, D.; White, O.; Whiteley, A.; Wilke, A.; Wortman, J. R.; Yatsunenko, T.; Glöckner, F. O., Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology* **2011,** *29* (5), 415-420.

36. Field, D.; Amaral-Zettler, L.; Cochrane, G.; Cole, J. R.; Dawyndt, P.; Garrity, G. M.; Gilbert, J.; Glöckner, F. O.; Hirschman, L.; Karsch-Mizrachi, I.; Klenk, H. P.; Knight, R.; Kottmann, R.; Kyrpides, N.; Meyer, F.; Gil, I. S.; Sansone, S. A.; Schriml, L. M.; Sterk, P.; Tatusova, T.; Ussery, D. W.; White, O.; Wooley, J., The Genomic Standards Consortium. *PLoS Biology* **2011,** *9* (6), 8-10.

37. Conway, K. R.; Boddy, C. N., ClusterMine360: a database of microbial PKS/NRPS biosynthesis. *Nucleic acids research* **2013,** *41* (Database issue), D402-7.

38. Ichikawa, N.; Sasagawa, M.; Yamamoto, M.; Komaki, H.; Yoshida, Y., DoBISCUIT : a database of secondary metabolite biosynthetic gene clusters. **2013,** *41* (November 2012), 408-414.

39. Nierman, W. C.; Pain, A.; Anderson, M. J.; Wortman, J. R.; Kim, H. S.; Arroyo, J.; Berriman, M.; Abe, K.; Archer, D. B.; Bermejo, C.; Bennett, J.; Bowyer, P.; Chen, D.; Collins, M.; Coulsen, R.; Davies, R.; Dyer, P. S.; Farman, M.; Fedorova, N.; Fedorova, N.; Feldblyum, T. V.; Fischer, R.; Fosker, N.; Fraser, A.; García, J. L.; García, M. J.; Goble, A.; Goldman, G. H.; Gomi, K.; Griffith-Jones, S.; Gwilliam, R.; Haas, B.; Haas, H.; Harris, D.; Horiuchi, H.; Huang, J.; Humphray, S.; Jiménez, J.; Keller, N.; Khouri, H.; Kitamoto, K.; Kobayashi, T.; Konzack, S.; Kulkarni, R.; Kumagai, T.; Lafton, A.; Latgé, J. P.; Li, W.; Lord, A.; Lu, C.; Majoros, W. H.; May, G. S.; Miller, B. L.; Mohamoud, Y.; Molina, M.; Monod, M.; Mouyna, I.; Mulligan, S.; Murphy, L.; O'Neil, S.; Paulsen, I.; Peñalva, M. A.; Pertea, M.; Price, C.; Pritchard, B. L.; Quail, M. A.; Rabbinowitsch, E.; Rawlins, N.; Rajandream, M. A.; Reichard, U.; Renauld, H.; Robson, G. D.; Rodriguez De Cordoba, S.; Rodríguez-Peña, J. M.; Ronning, C. M.; Rutter, S.; Salzberg, S. L.; Sanchez, M.; Sánchez-Ferrero, J. C.; Saunders, D.; Seeger, K.; Squares, R.; Squares, S.; Takeuchi, M.; Tekaia, F.; Turner, G.; Vazquez De Aldana, C. R.; Weidman, J.; White, O.; Woodward, J.; Yu, J. H.; Fraser, C.; Galagan, J. E.; Asai, K.; Machida, M.; Hall, N.; Barrell, B.; Denning, D. W.,

Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. *Nature* **2005,** *438* (7071), 1151-1156.

40. Noma, K.; Allis, C. D.; Grewal, S. I., Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science (New York, N.Y.)* **2001,** *293* (5532), 1150-1155.

41. Shwab, E. K.; Jin, W. B.; Tribus, M.; Galehr, J.; Graessle, S.; Keller, N. P., Historie deacetylase activity regulates chemical diversity in Aspergillus. *Eukaryotic Cell* **2007,** *6* (9), 1656-1664.

42. Lee, I.; Oh, J. H.; Keats Shwab, E.; Dagenais, T. R. T.; Andes, D.; Keller, N. P., HdaA, a class 2 histone deacetylase of Aspergillus fumigatus, affects germination and secondary metabolite production. *Fungal Genetics and Biology* **2009,** *46* (10), 782-790.

43. Reyes-Dominguez, Y.; Bok, J. W.; Berger, H.; Shwab, E. K.; Basheer, A.; Gallmetzer, A.; Scazzocchio, C.; Keller, N.; Strauss, J., Heterochromatic marks are associated with the repression of secondary metabolism clusters in Aspergillus nidulans. *Molecular Microbiology* **2010,** *76* (6), 1376-1386.

44. Du, L.; King, J. B.; Cichewicz, R. H., Chlorinated polyketide obtained from a Daldinia sp. treated with the epigenetic modifier suberoylanilide hydroxamic acid. *Journal of Natural Products* **2014,** *77* (11), 2454-2458.

45. Li, G.; Kusari, S.; Golz, C.; Laatsch, H.; Strohmann, C.; Spiteller, M., Epigenetic Modulation of Endophytic Eupenicillium sp. LG41 by a Histone Deacetylase Inhibitor for Production of Decalin-Containing Compounds. *Journal of Natural Products* **2017,** *80* (4), 983-988.

46. Mao, X. M.; Xu, W.; Li, D.; Yin, W. B.; Chooi, Y. H.; Li, Y. Q.; Tang, Y.; Hu, Y., Epigenetic genome mining of an endophytic fungus leads to the pleiotropic biosynthesis of natural products. *Angewandte Chemie - International Edition* **2015,** *54* (26), 7592-7596.

47. Chung, Y. M.; El-Shazly, M.; Chuang, D. W.; Hwang, T. L.; Asai, T.; Oshima, Y.; Ashour, M. L.; Wu, Y. C.; Chang, F. R., Suberoylanilide hydroxamic acid, a histone deacetylase inhibitor, induces the production of anti-inflammatory cyclodepsipeptides from Beauveria felina. *Journal of Natural Products* **2013,** *76* (7), 1260-1266.

48. Krick, A.; Kehraus, S.; Gerhäuser, C.; Klimo, K.; Nieger, M.; Maier, A.; Fiebig, H. H.; Atodiresei, I.; Raabe, G.; Fleischhauer, J.; König, G. M., Potential cancer chemopreventive in vitro activities of monomeric xanthone derivatives from the marine algicolous fungus Monodictys putredinis. *Journal of Natural Products* **2007,** *70* (3), 353-360.

49. Sanchez, J. F.; Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Ahuja, M.; Elizabeth Oakley, C.; Woo Bok, J.; Keller, N.; Oakley, B. R.; Wang, C. C. C., Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of Aspergillus nidulans. *Molecular BioSystems* **2010,** *6* (3), 587-593.

50. Wu, G.; Zhou, H.; Zhang, P.; Wang, X.; Li, W.; Zhang, W.; Liu, X.; Liu, H. W.; Keller, N. P.; An, Z.; Yin, W. B., Polyketide Production of Pestaloficiols and Macrodiolide Ficiolides

Revealed by Manipulations of Epigenetic Regulators in an Endophytic Fungus. *Organic Letters* **2016,** *18* (8), 1832-1835.

51. Barinaga, M., Origins of lichen fungi explored. *Science* **1995,** *268* (5216), 1437 LP-1437.

52. Partida-Martinez, L. P.; Hertweck, C., Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* **2005,** *437* (7060), 884-888.

53. Schroeckh, V.; Scherlach, K.; Nützmann, H.-W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A., Intimate bacterial–fungal interaction triggers biosynthesis of archetypal polyketides in <em>Aspergillus nidulans</em&gt. *Proceedings of the National Academy of Sciences* **2009,** *106* (34), 14558 LP-14563.

54. Nutzmann, H.-W.; Reyes-Dominguez, Y.; Scherlach, K.; Schroeckh, V.; Horn, F.; Gacek, A.; Schumann, J.; Hertweck, C.; Strauss, J.; Brakhage, A. A., Bacteria-induced natural product formation in the fungus Aspergillus nidulans requires Saga/Ada-mediated histone acetylation. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (34), 14282-14287.

55. Cueto, M.; Jensen, P. R.; Kauffman, C.; Fenical, W.; Lobkovsky, E.; Clardy, J., Pestalone, a new antibiotic produced by a marine fungus in response to bacterial challenge. *Journal of Natural Products* **2001,** *64* (11), 1444-1446.

56. Oh, D. C.; Jensen, P. R.; Kauffman, C. A.; Fenical, W., Libertellenones A-D: Induction of cytotoxic diterpenoid biosynthesis by marine microbial competition. *Bioorganic and Medicinal Chemistry* **2005,** *13* (17), 5267-5273.

57. Park, H. B.; Kwon, H. C.; Lee, C. H.; Yang, H. O., Glionitrin A, an antibiotic - Antitumor metabolite derived from competitive interaction between abandoned mine microbes. *Journal of Natural Products* **2009,** *72* (2), 248-252.

58. Degenkolb, T.; Heinze, S.; Schlegel, B.; Strobel, G.; Gräfe, U., Formation of new lipoaminopeptides, acremostatins A, B, and C, by co-cultivation of Acremonium sp. Tbp-5 and Mycogone rosea DSM 12973. *Bioscience, biotechnology, and biochemistry* **2002,** *66* (4), 883-886.

59. Zhu, F.; Chen, G.; Chen, X.; Huang, M.; Wan, X., Aspergicin, a new antibacterial alkaloid produced by mixed fermentation of two marine-derived mangrove epiphytic fungi. *Chemistry of Natural Compounds* **2011,** *47* (5), 767-769.

60. Li, C.; Wang, J.; Luo, C.; Ding, W.; Cox, D. G., A new cyclopeptide with antifungal activity from the co-culture broth of two marine mangrove fungi. *Natural product research* **2014,** *28* (9), 616-621.

61. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)* **2012,** *337* (6096), 816-821.

62. Dicarlo, J. E.; Norville, J. E.; Mali, P.; Rios, X.; Aach, J.; Church, G. M., Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Research* **2013,** *41* (7), 4336-4343.

63. Cho, S. W.; Kim, S.; Kim, J. M.; Kim, J. S., Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature Biotechnology* **2013,** *31* (3), 230-232.

64. Belhaj, K.; Chaparro-Garcia, A.; Kamoun, S.; Nekrasov, V., Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **2013,** *9* (1), 1-10.

65. 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.

66. Matsu-ura, T.; Baek, M.; Kwon, J.; Hong, C., Efficient gene editing in Neurospora crassa with CRISPR technology. *Fungal Biology and Biotechnology* **2015,** *2* (1), 4-4.

67. Liu, R.; Chen, L.; Jiang, Y.; Zhou, Z.; Zou, G., Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. *Cell Discovery* **2015,** *1*, 1-11.

68. Pohl, C.; Kiel, J. A. K. W.; Driessen, A. J. M.; Bovenberg, R. A. L.; Nygård, Y., CRISPR/Cas9 Based Genome Editing of Penicillium chrysogenum. *ACS Synthetic Biology* **2016,** *5* (7), 754-764.

69. 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.

70. West, R. R.; Ness, J. V.; Varming, A.-m. M.; Biggs, S.; Gasper, S.; McKernan, A.; Piggott, J.; VanNess, J.; Varming, A.-m. M.; Rassing, B.; Biggs, S.; Gasper, S.; McKernan, P. A.; Piggott, J., ZG-1494 alpha, a novel platelet-activating factor acetyltransferase inhibitor from Penicilium rubrum, isolation, structure elucidation and biological activity. *Journal of Antibiotics* **1996,** *49* (10), 967-973.

71. Nielsen, M. L.; Isbrandt, T.; Petersen, L. M.; Mortensen, U. H.; Andersen, M. R.; Hoof, J. B.; Larsen, T. O., Linker flexibility facilitates module exchange in fungal hybrid PKS-NRPS engineering. *PLoS ONE* **2016,** *11* (8), 1-18.

72. Jensen, E. D.; Ferreira, R.; Jakočiunas, T.; Arsovska, D.; Zhang, J.; Ding, L.; Smith, J. D.; David, F.; Nielsen, J.; Jensen, M. K.; Keasling, J. D., Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microbial Cell Factories* **2017,** *16* (1), 1-16.

73. Zhang, X.; Wang, J.; Wang, J.; Cheng, Q.; Zheng, X.; Zhao, G., Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discovery* **2017,** *3*, 1-9.

74. Lee, K. K. M.; Silva, N. A. D.; Kealey, J. T., Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli and Saccharomyces cerevisiae. *Analytical Biochemistry* **2009,** *394* (1), 75-80.

75. Wattanachaisaereekul, S.; Lantz, A. E.; Nielsen, M. L.; Andrésson, Ó. S.; Nielsen, J., Optimization of heterologous production of the polyketide 6-MSA in Saccharomyces cerevisiae. *Biotechnology and Bioengineering* **2007,** *97* (4), 893-900.

76. Lee, K. M.; DaSilva, N. A., Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. *Yeast* **2005,** *22* (6), 431-440.

77. Tang, M. C.; Lin, H. C.; Li, D.; Zou, Y.; Li, J.; Xu, W.; Cacho, R. A.; Hillenmeyer, M. E.; Garg, N. K.; Tang, Y., Discovery of Unclustered Fungal Indole Diterpene Biosynthetic Pathways through Combinatorial Pathway Reassembly in Engineered Yeast. *Journal of the American Chemical Society* **2015,** *137* (43), 13724-13727.

78. Karim, A. S.; Curran, K. A.; Alper, H. S., Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications. *FEMS Yeast Research* **2013,** *13* (1), 107-116.

79. Ma, S. M.; Li, J. W. H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y., Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase. *Science* **2009,** *326* (5952), 589 LP-592.

80. Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W. H.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y., Enzymatic synthesis of resorcylic acid lactones by cooperation of fungal iterative polyketide synthases involved in hypothemycin biosynthesis. *Journal of the American Chemical Society* **2010,** *132* (13), 4530-4531.

81. Gao, X.; Jiang, W.; Jiménez-Osés, G.; Choi, M. S.; Houk, K. N.; Tang, Y.; Walsh, C. T., An iterative, bimodular nonribosomal peptide synthetase that converts anthranilate and tryptophan into tetracyclic asperlicins. *Chemistry and Biology* **2013,** *20* (7), 870-878.

82. Lin, H. C.; Chooi, Y. H.; Dhingra, S.; Xu, W.; Calvo, A. M.; Tang, Y., The fumagillin biosynthetic gene cluster in Aspergillus fumigatus encodes a cryptic terpene cyclase involved in the formation of β-trans-bergamotene. *Journal of the American Chemical Society* **2013,** *135* (12), 4616-4619.

83. Wang, M.; Beissner, M.; Zhao, H., Aryl-aldehyde formation in fungal polyketides: Discovery and characterization of a distinct biosynthetic mechanism. *Chemistry and Biology* **2014,** *21* (2), 257-263.

84. Harvey, C. J. B.; Tang, M.; Schlecht, U.; Horecka, J.; Fischer, C. R.; Lin, H.-C.; Li, J.; Naughton, B.; Cherry, J.; Miranda, M.; Li, Y. F.; Chu, A. M.; Hennessy, J. R.; Vandova, G. A.; Inglis, D.; Aiyar, R. S.; Steinmetz, L. M.; Davis, R. W.; Medema, M. H.; Sattely, E.; Khosla, C.; St Onge, R. P.; Tang, Y.; Hillenmeyer, M. E., HEx: A heterologous expression platform for the discovery of fungal natural products. *Science advances* **2018,** *4* (4), eaar5459-eaar5459.

85. Kealey, J. T.; Liu, L.; Santi, D. V.; Betlach, M. C.; Barr, P. J., Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proceedings of the National Academy of Sciences* **1998,** *95* (2), 505 LP-509.

86. Balibar, C. J.; Howard-Jones, A. R.; Walsh, C. T., Terrequinone A biosynthesis through L-tryptophan oxidation, dimerization and bisprenylation. *Nature Chemical Biology* **2007,** *3* (9), 584-592.

87. Ma, S. M.; Zhan, J.; Watanabe, K.; Xie, X.; Zhang, W.; Wang, C. C.; Tang, Y., Enzymatic synthesis of aromatic polyketides using PKS4 from Gibberella fujikuroi. *Journal of the American Chemical Society* **2007,** *129* (35), 10642-10643.

88. Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y., Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. *The Journal of biological chemistry* **2010,** *285* (53), 41412-41421.

89. Zhang, W.; Li, Y.; Tang, Y., Engineered biosynthesis of bacterial aromatic polyketides in Escherichia coli. *Proceedings of the National Academy of Sciences* **2008,** *105* (52), 20683-20688.

90. Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R., A versatile and efficient gene-targeting system for Aspergillus nidulans. *Genetics* **2006,** *172* (3), 1557-1566.

91. 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. *Journal of Biotechnology* **2007,** *128* (4), 770-775.

92. Tsunematsu, Y.; Ishikawa, N.; Wakana, D.; Goda, Y.; Noguchi, H.; Moriya, H.; Hotta, K.; Watanabe, K., Distinct mechanisms for spiro-carbon formation reveal biosynthetic pathway crosstalk. *Nature Chemical Biology* **2013,** *9* (12), 818-825.

93. Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M. C.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y., SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **2017,** *549* (7673), 502-506.

94. Hai, Y.; Tang, Y., Biosynthesis of Long-Chain N-Acyl Amide by a Truncated Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrid Megasynthase in Fungi. *Journal of the American Chemical Society* **2018,** *140* (4), 1271-1274.

95. Tagami, K.; Liu, C.; Minami, A.; Noike, M.; Isaka, T.; Fueki, S.; Shichijo, Y.; Toshima, H.; Gomi, K.; Dairi, T.; Oikawa, H., Reconstitution of Biosynthetic Machinery for Indole-Diterpene Paxilline in Aspergillus oryzae. *Journal of the American Chemical Society* **2013,** *135* (4), 1260-1263.

96. Matsuda, Y.; Quan, Z.; Mitsuhashi, T.; Li, C.; Abe, I., Cytochrome P450 for Citreohybridonol Synthesis: Oxidative Derivatization of the Andrastin Scaffold. *Organic Letters* **2016,** *18* (2), 296-299.

97. Matsuda, Y.; Iwabuchi, T.; Wakimoto, T.; Awakawa, T.; Abe, I., Uncovering the unusual d-ring construction in terretonin biosynthesis by collaboration of a multifunctional cytochrome P450 and a unique isomerase. *Journal of the American Chemical Society* **2015,** *137* (9), 3393-3401.

98. Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I., Complete biosynthetic pathway of anditomin: Nature's sophisticated synthetic route to a complex fungal meroterpenoid. *Journal of the American Chemical Society* **2014,** *136* (43), 15326-15336.

99. Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J., Three previously unrecognised classes of biosynthetic enzymes revealed during the production of xenovulene A. *Nature Communications* **2018,** *9* (1).

100. Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W., Native Promoter Strategy for High-Yielding Synthesis and Engineering of Fungal Secondary Metabolites. *ACS Synthetic Biology* **2015,** *4* (5), 625-633.

101. Sims, J. W.; Fillmore, J. P.; Warner, D. D.; Schmidt, E. W., Equisetin biosynthesis in Fusarium heterosporum. *Chemical communications (Cambridge, England)* **2005,** (2), 186-188.

102. Clevenger, K. D.; Bok, J. W.; Ye, R.; Miley, G. P.; Verdan, M. H.; Velk, T.; Chen, C.; Yang, K. H.; Robey, M. T.; Gao, P.; Lamprecht, M.; Thomas, P. M.; Islam, M. N.; Palmer, J. M.; Wu, C. C.; Keller, N. P.; Kelleher, N. L., A scalable platform to identify fungal secondary metabolites and their gene clusters. *Nature Chemical Biology* **2017,** *13* (8), 895-901.

103. Bok, J. W.; Ye, R.; Clevenger, K. D.; Mead, D.; Wagner, M.; Krerowicz, A.; Albright, J. C.; Goering, A. W.; Thomas, P. M.; Kelleher, N. L.; Keller, N. P.; Wu, C. C., Fungal artificial chromosomes for mining of the fungal secondary metabolome. *BMC Genomics* **2015,** *16* (1), 1-10.

104. Almabruk, K. H.; Dinh, L. K.; Philmus, B., Self-Resistance of Natural Product Producers: Past, Present, and Future Focusing on Self-Resistant Protein Variants. *ACS Chemical Biology* **2018,** *13* (6), 1426-1437.

105. Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R., Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **1999,** *284* (5418), 1368-1372.

106. Regueira, T. B.; Kildegaard, K. R.; Hansen, B. G.; Mortensen, U. H.; Hertweck, C.; Nielsen, J., Molecular basis for mycophenolic acid biosynthesis in Penicillium brevicompactum. *Applied and Environmental Microbiology* **2011,** *77* (9), 3035-3043.

107. Liu, N.; Hung, Y.-S.; Gao, S.-S.; Hang, L.; Zou, Y.; Chooi, Y.-H.; Tang, Y., Identification and Heterologous Production of a Benzoyl-Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid A Biosynthetic Pathway. *Organic letters* **2017,** *19* (13), 3560-3563.

108. Bonsch, B.; Belt, V.; Bartel, C.; Duensing, N.; Koziol, M.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J., Identification of genes encoding squalestatin S1 biosynthesis and: In vitro production of new squalestatin analogues. *Chemical Communications* **2016,** *52* (41), 6777-6780.

109. Amorim Franco, T. M.; Blanchard, J. S., Bacterial Branched-Chain Amino Acid Biosynthesis: Structures, Mechanisms, and Drugability. *Biochemistry* **2017,** *56* (44), 5849-5865.

110. Yeh, H.-H.; Ahuja, M.; Chiang, Y.-M.; Oakley, C. E.; Moore, S.; Yoon, O.; Hajovsky, H.; Bok, J.-W.; Keller, N. P.; Wang, C. C. C.; Oakley, B. R., Resistance Gene-Guided Genome Mining: Serial Promoter Exchanges in Aspergillus nidulans Reveal the Biosynthetic Pathway for Fellutamide B, a Proteasome Inhibitor. *ACS chemical biology* **2016,** *11* (8), 2275-2284.

111. Hung, Y.-S.; Zhao, Z.; Soohoo, A. M.; Tang, Y., Genome Mining in Fungi. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier: 2019.

112. Hossain, M. A.; Johnson, T. L., Using yeast genetics to study splicing mechanisms. *Methods in molecular biology (Clifton, N.J.)* **2014,** *1126*, 285-98.

113. Houman, F.; Rho, S. B.; Zhang, J.; Shen, X.; Wang, C.-C.; Schimmel, P.; Martinis, S. A., A prokaryote and human tRNA synthetase provide an essential RNA splicing function in yeast mitochondria. *Proceedings of the National Academy of Sciences* **2000,** *97* (25), 13743-13748.

114. Song, Z.; Bakeer, W.; Marshall, J. W.; Yakasai, A. A.; Khalid, R. M.; Collemare, J.; Skellam, E.; Tharreau, D.; Lebrun, M.-H.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J., Heterologous expression of the avirulence gene ACE1 from the fungal rice pathogen Magnaporthe oryzae. *Chemical Science* **2015,** *6* (8), 4837-4845.

115. Díez, B.; Alvarez, E.; Cantoral, J. M.; Barredo, J. L.; Martín, J. F., Selection and characterization of pyrG mutants of Penicillium chrysogenum lacking orotidine-5′-phosphate decarboxylase and complementation by the pyr4 gene of Neurospora crassaa. *Current Genetics* **1987,** *12* (4), 277-282.

116. Chiang, Y. M.; Oakley, C. E.; Ahuja, M.; Entwistle, R.; Schultz, A.; Chang, S. L.; Sung, C. T.; Wang, C. C. C.; Oakley, B. R., An efficient system for heterologous expression of secondary metabolite genes in Aspergillus nidulans. *Journal of the American Chemical Society* **2013,** *135* (20), 7720-7731.

117. Sato, M.; Yagishita, F.; Mino, T.; Uchiyama, N.; Patel, A.; Chooi, Y. H.; Goda, Y.; Xu, W.; Noguchi, H.; Yamamoto, T.; Hotta, K.; Houk, K. N.; Tang, Y.; Watanabe, K., Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition. *Chembiochem* **2015,** *16* (16), 2294-8.

118. Yu, J. H.; Leonard, T. J., Sterigmatocystin biosynthesis in Aspergillus nidulans requires a novel type I polyketide synthase. *Journal of Bacteriology* **1995,** *177* (16), 4792-4800.

119. Chiang, Y. M.; Szewczyk, E.; Nayak, T.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Ho, W. Y.; Simityan, H.; Kuo, E.; Praseuth, A.; Watanabe, K.; Oakley, B. R.; Wang, C. C., Molecular genetic mining of the Aspergillus secondary metabolome: discovery of the emericellamide biosynthetic pathway. *Chemistry & biology* **2008,** *15* (6), 527-32.

120. Posch, A. E.; Spadiut, O.; Herwig, C., A novel method for fast and statistically verified morphological characterization of filamentous fungi. *Fungal genetics and biology : FG & B* **2012,** *49* (7), 499-510.

121. Kopke, K.; Hoff, B.; Bloemendal, S.; Katschorowski, A.; Kamerewerd, J.; Kück, U., Members of the Penicillium chrysogenum velvet complex play functionally opposing roles in the regulation of penicillin biosynthesis and conidiation. *Eukaryotic cell* **2013,** *12* (2), 299-310.

122. Christensen, B.; Thykaer, J.; Nielsen, J., Metabolic characterization of high- and low-yielding strains of Penicillium chrysogenum. *Appl Microbiol Biotechnol* **2000,** *54* (2), 212-7.

123. Tannous, J.; Keller, N. P.; Atoui, A.; El Khoury, A.; Lteif, R.; Oswald, I. P.; Puel, O., Secondary metabolism in Penicillium expansum: Emphasis on recent advances in patulin research. *Critical reviews in food science and nutrition* **2018,** *58* (12), 2082-2098.

124. d'Enfert, C., Selection of multiple disruption events in Aspergillus fumigatus using the orotidine-5'-decarboxylase gene, pyrG, as a unique transformation marker. *Curr Genet* **1996,** *30* (1), 76-82.

125. Osmani, A. H.; May, G. S.; Osmani, S. A., The extremely conserved pyroA gene of Aspergillus nidulans is required for pyridoxine synthesis and is required indirectly for resistance to photosensitizers. *J Biol Chem* **1999,** *274* (33), 23565-9.

126. Oakley, C. E.; Weil, C. F.; Kretz, P. L.; Oakley, B. R., Cloning of the riboB locus of Aspergillus nidulans. *Gene* **1987,** *53* (2-3), 293-8.

127. Kwon-Chung, K. J.; Varma, A.; Edman, J. C.; Bennett, J. E., Selection of ura5 and ura3 mutants from the two varieties of Cryptococcus neoformans on 5-fluoroorotic acid medium. *Journal of medical and veterinary mycology : bi-monthly publication of the International Society for Human and Animal Mycology* **1992,** *30* (1), 61-9.

128. Punt, P. J.; Dingemanse, M. A.; Kuyvenhoven, A.; Soede, R. D.; Pouwels, P. H.; van den Hondel, C. A., Functional elements in the promoter region of the Aspergillus nidulans gpdA gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene* **1990,** *93* (1), 101-9.

129. He, X.; Zhang, Z.; Che, Q.; Zhu, T.; Gu, Q.; Li, D., Varilactones and wortmannilactones produced by Penicillium variabile cultured with histone deacetylase inhibitor. *Archives of pharmacal research* **2018,** *41* (1), 57-63.

130. Dong, Y.; Lin, J.; Lu, X.; Zheng, Z.; Ren, X.; Zhang, H.; He, J.; Yang, J., Cathepsin B Inhibitory Tetraene Lactones from the Fungus Talaromyces wortmannii. *Helvetica Chimica Acta* **2009,** *92* (3), 567-574.

131. Liu, W. C.; Ren, Y. X.; Hao, A. Y.; Yu, S.; Shi, X.; Zhang, X. Q.; Xing, Y.; Xiu, Z. L.; Cui, Y.; Dong, Y. S., The activities of wortmannilactones against helminth electron transport chain enzymes, structure-activity relationships, and the effect on Trichinella spiralis infected mice. *The Journal of antibiotics* **2018,** *71* (8), 731-740.

132. Proteau, P. J.; Gerwick, W. H., Cymathere ethers A and B: Bicyclic oxylipins from the marine brown alga Cymathere triplicata. *Tetrahedron Letters* **1992,** *33* (31), 4393-4396.

133. Wei, H.; Itoh, T.; Kinoshita, M.; Kotoku, N.; Aoki, S.; Kobayashi, M., Shimalactone A, a novel polyketide, from marine-derived fungus Emericella variecolor GF10. *Tetrahedron* **2005,** *61* (33), 8054-8058.

134. Jayasuriya, H.; Guan, Z.; Dombrowski, A. W.; Bills, G. F.; Polishook, J. D.; Jenkins, R. G.; Koch, L.; Crumley, T.; Tamas; Dubois, M.; Misura, A.; Darkin-Rattray, S. J.; Gregory, L.; Singh, S. B., Isolation, Structure, and Coccidiostat Activity of Coccidiostatin A. *Journal of Natural Products* **2007,** *70* (8), 1364-1367.

135. Lang, G.; Wiese, J.; Schmaljohann, R.; Imhoff, J. F., New pentaenes from the sponge-derived marine fungus Penicillium rugulosum: structure determination and biosynthetic studies. *Tetrahedron* **2007,** *63* (48), 11844-11849.

136. Mori, M.; Morimoto, H.; Kim, Y.-P.; Ui, H.; Nonaka, K.; Masuma, R.; Sakamoto, K.; Kita, K.; Tomoda, H.; Shiomi, K.; Ōmura, S., Ukulactones A and B, new NADH-fumarate reductase inhibitors produced by Penicillium sp. FKI-3389. *Tetrahedron* **2011,** *67* (35), 6582-6586.

137. Fujii, I.; Hashimoto, M.; Konishi, K.; Unezawa, A.; Sakuraba, H.; Suzuki, K.; Tsushima, H.; Iwasaki, M.; Yoshida, S.; Kudo, A.; Fujita, R.; Hichiwa, A.; Saito, K.; Asano, T.; Ishikawa, J.; Wakana, D.; Goda, Y.; Watanabe, A.; Watanabe, M.; Masumoto, Y.; Kanazawa, J.; Sato, H.; Uchiyama, M., Shimalactone Biosynthesis Involves Spontaneous Double Bicyclo-Ring Formation with 8pi-6pi Electrocyclization. *Angew Chem Int Ed Engl* **2020**.

138. Lin, T.-S.; Chen, B.; Chiang, Y.-M.; Wang, C. C. C., Discovery and Elucidation of the Biosynthesis of Aspernidgulenes: Novel Polyenes from Aspergillus Nidulans by Using Serial Promoter Replacement. *ChemBioChem* **2019,** *20* (3), 329-334.

139. Mao, X.-M.; Zhan, Z.-J.; Grayson, M. N.; Tang, M.-C.; Xu, W.; Li, Y.-Q.; Yin, W.-B.; Lin, H.-C.; Chooi, Y.-H.; Houk, K. N.; Tang, Y., Efficient Biosynthesis of Fungal Polyketides Containing the Dioxabicyclo-octane Ring System. *Journal of the American Chemical Society* **2015,** *137* (37), 11904-11907.

140. Ma, S. M.; W-H Li, J.; Choi, J. W.; Zhou, H.; Michael Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y., SUPPORTING INFORMATION Complete Reconstitution of a Highly-Reducing Iterative Polyketide Synthase. **2009,** (October), 589-593.

141. Sofiyev, V.; Navarro, G.; Trauner, D., Biomimetic Synthesis of the Shimalactones. *Organic Letters* **2008,** *10* (1), 149-152.

142. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **2001,** *305* (3), 567-80.

143. Marmulla, R.; Šafarić, B.; Markert, S.; Schweder, T.; Harder, J., Linalool isomerase, a membrane-anchored enzyme in the anaerobic monoterpene degradation in Thauera linaloolentis 47Lol. *BMC Biochemistry* **2016,** *17* (1), 6.

144. Brodkorb, D.; Gottschall, M.; Marmulla, R.; Lüddeke, F.; Harder, J., Linalool Dehydratase-Isomerase, a Bifunctional Enzyme in the Anaerobic Degradation of Monoterpenes. *Journal of Biological Chemistry* **2010,** *285* (40), 30436-30442.

145. Schwab, J. M.; Henderson, B. S., Enzyme-catalyzed allylic rearrangements. *Chemical Reviews* **1990,** *90* (7), 1203-1245.

146. Gribble, G. W., Naturally Occurring Organohalogen Compounds. *Accounts of Chemical Research* **1998,** *31* (3), 141-152.

147. Kambe, N.; Iwasaki, T.; Terao, J., Pd-catalyzed cross-coupling reactions of alkyl halides. *Chemical Society reviews* **2011,** *40* (10), 4937-47.

148. Shilov, A. E.; Shul'pin, G. B., Activation of C−H Bonds by Metal Complexes. *Chemical Reviews* **1997,** *97* (8), 2879-2932.

149. Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L., Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **2006,** *440* (7082), 368-371.

150. Vaillancourt, F. H.; Yin, J.; Walsh, C. T., SyrB2 in syringomycin E biosynthesis is a nonheme FeII alpha-ketoglutarate- and O2-dependent halogenase. *Proc Natl Acad Sci U S A* **2005,** *102* (29), 10111-6.

151. Galonić, D. P.; Vaillancourt, F. H.; Walsh, C. T., Halogenation of Unactivated Carbon Centers in Natural Product Biosynthesis: Trichlorination of Leucine during Barbamide Biosynthesis. *Journal of the American Chemical Society* **2006,** *128* (12), 3900-3901.

152. Jiang, W.; Heemstra, J. R., Jr.; Forseth, R. R.; Neumann, C. S.; Manaviazar, S.; Schroeder, F. C.; Hale, K. J.; Walsh, C. T., Biosynthetic chlorination of the piperazate residue in kutzneride biosynthesis by KthP. *Biochemistry* **2011,** *50* (27), 6063-72.

153. Hillwig, M. L.; Liu, X., A new family of iron-dependent halogenases acts on freestanding substrates. *Nat Chem Biol* **2014,** *10* (11), 921-3.

154. Neugebauer, M. E.; Sumida, K. H.; Pelton, J. G.; McMurry, J. L.; Marchand, J. A.; Chang, M. C. Y., A family of radical halogenases for the engineering of amino-acid-based products. *Nat Chem Biol* **2019,** *15* (10), 1009-1016.

155. Zhao, C.; Yan, S.; Li, Q.; Zhu, H.; Zhong, Z.; Ye, Y.; Deng, Z.; Zhang, Y., An Fe(2+) - and alpha-Ketoglutarate-Dependent Halogenase Acts on Nucleotide Substrates. *Angew Chem Int Ed Engl* **2020**.

156. Yeh, E.; Blasiak, L. C.; Koglin, A.; Drennan, C. L.; Walsh, C. T., Chlorination by a Long-Lived Intermediate in the Mechanism of Flavin-Dependent Halogenases. *Biochemistry* **2007,** *46* (5), 1284-1292.

157. Podzelinska, K.; Latimer, R.; Bhattacharya, A.; Vining, L. C.; Zechel, D. L.; Jia, Z., Chloramphenicol biosynthesis: the structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond. *J Mol Biol* **2010,** *397* (1), 316-31.

158. Ugai, T.; Minami, A.; Tanaka, S.; Ozaki, T.; Liu, C.; Shigemori, H.; Hashimoto, M.; Oikawa, H., Biosynthetic Machinery of 6-Hydroxymellein Derivatives Leading to Cyclohelminthols and Palmaenones. *Chembiochem* **2020,** *21* (3), 360-367.

159. Jungmann, K.; Jansen, R.; Gerth, K.; Huch, V.; Krug, D.; Fenical, W.; Muller, R., Two of a Kind--The Biosynthetic Pathways of Chlorotonil and Anthracimycin. *ACS Chem Biol* **2015,** *10* (11), 2480-90.

160. Chankhamjon, P.; Tsunematsu, Y.; Ishida-Ito, M.; Sasa, Y.; Meyer, F.; Boettger-Schmidt, D.; Urbansky, B.; Menzel, K. D.; Scherlach, K.; Watanabe, K.; Hertweck, C., Regioselective Dichlorination of a Non-Activated Aliphatic Carbon Atom and Phenolic Bismethylation by a Multifunctional Fungal Flavoenzyme. *Angew Chem Int Ed Engl* **2016,** *55* (39), 11955-9.

161. Scharf, D. H.; Habel, A.; Heinekamp, T.; Brakhage, A. A.; Hertweck, C., Opposed Effects of Enzymatic Gliotoxin N- and S-Methylations. *Journal of the American Chemical Society* **2014,** *136* (33), 11674-11679.

162. Chooi, Y.-H.; Fang, J.; Liu, H.; Filler, S. G.; Wang, P.; Tang, Y., Genome Mining of a Prenylated and Immunosuppressive Polyketide from Pathogenic Fungi. *Organic Letters* **2013,** *15* (4), 780-783.

163. Ma, J.; Duan, X.; Zhou, H.; Tian, X.; Liu, J., Iodobenzene Dichloride Mediated Sequential C–Cl Bond Formation: A Safe, Convenient and Efficient Method for the Direct α,α-Dichlorination of β-Dicarbonyl Compounds. *Synthesis* **2015,** *47* (06), 777-782.

164. Piraee, M.; White, R. L.; Vining, L. C., Biosynthesis of the dichloroacetyl component of chloramphenicol in Streptomyces venezuelae ISP5230: genes required for halogenation. *Microbiology* **2004,** *150* (Pt 1), 85-94.

165. Fernandez-Martinez, L. T.; Borsetto, C.; Gomez-Escribano, J. P.; Bibb, M. J.; Al-Bassam, M. M.; Chandra, G.; Bibb, M. J., New insights into chloramphenicol biosynthesis in Streptomyces venezuelae ATCC 10712. *Antimicrob Agents Chemother* **2014,** *58* (12), 7441-50.

166. Singleton, V. L.; Bohonos, N.; Ullstrup, A. J., Decumbin, a new compound from a species of Penicillium. *Nature* **1958,** *181* (4615), 1072-3.

167. Sunagawa, M.; Ohta, T.; Nozoe, S., Biosynthesis of brefeldin A. Introduction of oxygen at the C-7 position. *The Journal of antibiotics* **1983,** *36* (1), 25-9.

168. Yamamoto, Y.; Hori, A.; Hutchinson, C. R., Biosynthesis of macrolide antibiotics. 6. Late steps in brefeldin A biosynthesis. *Journal of the American Chemical Society* **1985,** *107* (8), 2471-2474.

169. Haerri, E.; Loeffler, W.; Sigg, H.; Stähelin, H.; Tamm, C., Isolation of new metabolic products from penicillium brefeldianum. *Helv. Chim. Acta* **1963,** *46*, 1235-1243.

170. Tamura, G.; Ando, K.; Suzuki, S.; Takatsuki, A.; Arima, K., Antiviral activity of brefeldin A and verrucarin A. *The Journal of antibiotics* **1968,** *21* (2), 160-1.

171. Betina, V.; Montagier, L., Action of cyanein on the synthesis of nucleic acids and proteins in animal cells and bacterial protoplasts. *Bull. Soc. Chim. Biol* **1966,** *48*, 194-198.

172. Lichman, B. R.; Godden, G. T.; Hamilton, J. P.; Palmer, L.; Kamileen, M. O.; Zhao, D.; Vaillancourt, B.; Wood, J. C.; Sun, M.; Kinser, T. J.; Henry, L. K.; Rodriguez-Lopez, C.; Dudareva, N.; Soltis, D. E.; Soltis, P. S.; Buell, C. R.; O'Connor, S. E., The evolutionary origins of the cat attractant nepetalactone in catnip. *Science Advances* **2020,** *6* (20), eaba0721.

173. Helliwell, C. A.; Chandler, P. M.; Poole, A.; Dennis, E. S.; Peacock, W. J., The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proceedings of the National Academy of Sciences* **2001,** *98* (4), 2065-2070.

174. Zabala, A. O.; Chooi, Y. H.; Choi, M. S.; Lin, H. C.; Tang, Y., Fungal polyketide synthase product chain-length control by partnering thiohydrolase. *ACS Chem Biol* **2014,** *9* (7), 1576-86.

175. Zeng, F.; Chen, C.; Al Chnani, A. A.; Zhou, Q.; Tong, Q.; Wang, W.; Zang, Y.; Gong, J.; Wu, Z.; Liu, J.; Wang, J.; Zhu, H.; Zhang, Y., Dibrefeldins A and B, A pair of epimers representing the first brefeldin A dimers with cytotoxic activities from Penicillium janthinellum. *Bioorg Chem* **2019,** *86*, 176-182.

176. McCloud, T. G.; Burns, M. P.; Majadly, F. D.; Muschik, G. M.; Miller, D. A.; Poole, K. K.; Roach, J. M.; Ross, J. T.; Lebherz, W. B., Production of brefeldin-A. *Journal of Industrial Microbiology* **1995,** *15* (1), 5-9.

177. Gonzalez de la Parra, M.; Hutchinson, C. R., Macrolide biosynthesis: stereochemistry of the hydroxylation of brefeldin C. *The Journal of antibiotics* **1987,** *40* (8), 1170-4.

178. Wei, Y.; Contreras, J. A.; Sheffield, P.; Osterlund, T.; Derewenda, U.; Kneusel, R. E.; Matern, U.; Holm, C.; Derewenda, Z. S., Crystal structure of brefeldin A esterase, a bacterial homolog of the mammalian hormone-sensitive lipase. *Nature structural biology* **1999,** *6* (4), 340-5.