UCLA
UCLA Electronic Theses and Dissertations

Title

Natural products: metabolic engineering and discovery of novel compounds and enzymes

Permalink

https://escholarship.org/uc/item/7k58j6vm

Author

Yee, Danielle

Publication Date

2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Natural products: metabolic engineering and discovery of novel compounds and enzymes

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

by

Danielle Ashley Yee

2022

© Copyright by

Danielle Ashley Yee

2022

ABSTRACT OF THE DISSERTATION

Natural products: metabolic engineering and discovery of novel compounds and enzymes

by

Danielle Ashley Yee

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2022

Professor Yi Tang, Chair

Natural products from plants, fungi, and bacteria are an invaluable source of bioactive compounds with numerous applications in the pharmaceutical industry and agriculture. Metabolic engineering can be utilized to increase the yield of these valuable compounds and their precursors. Meanwhile, genome mining can be employed to discover novel natural product scaffolds and unique functional groups that can confer various biological activities to the resulting compound. During my doctoral research, I have performed work in both of these fields to engineer yeast for precursor production of the pharmaceutically relevant monoterpene indole alkaloids as well as elucidate the biosynthesis of new natural products from several fungal species. Strategies for yeast metabolic engineering included targeting the production of monoterpene precursors to the mitochondria and further mitochondrial engineering through deletion of transporter genes and transcription factor overexpression. Genome mining efforts led to the discovery of new compounds with a unique hybrid terpene and amino acid derived

scaffold as well as a new family of biosynthetic core genes that form arginine containing cyclodipeptides. Study of these pathways also resulted in the identification of enzymes that can perform novel reactions to create greater chemical diversity. These genes can expand the query database to mine for more compounds with novel structures.

The dissertation of Danielle Ashley Yee is approved.

Yvonne Y. Chen

Neil K. Garg

Junyoung O. Park

Yi Tang, Committee Chair

University of California, Los Angeles

2022

TABLE OF CONTENTS

TABLE OF FIGURES

ACKNOWLEDGMENTS

Sections 2 contains material written by Yee, D. A. from the following publication:

Yee, D. A., Tang, Y. “Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: *Aspergillus nidulans* as a Heterologous Host” *Engineering Natural Product Biosynthesis*. **2022**, 41-52.

Sections 3 contains material written by Yee, D. A. from the following publication:

Yee, D. A.; DeNicola, A. B.; Billingsley, J. M.; Creso, J. G.; Subrahmanyam, V.; Tang, Y. Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metabolic Eng.* **2019**, *55*, 76-84.

Section 6 contains material written by Yee, D. A. from the following publication:

Yee, D. A.*; Kakule, T. B.*; Cheng, W.*; Chen, M.; Chong, C. T.; Hai, Y.; Hang, L. F.; Hung, Y. S.; Liu, N.; Ohashi, M.; Okorafor, I. C.; 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**, *142*, 710-714.

I would like to thank my doctoral advisor Professor Yi Tang for his guidance and mentorship throughout my doctoral research. He has given a tremendous amount of help throughout this work, and his passion for science has been extremely inspiring. Also, I would like thank my committee members for their inspiration and advice. I would like to thank the support from UCLA’s dissertation fellowship for my doctoral research.

In addition, I would like to thank my lab members and friends for all of their assistance, with special thanks to my mentors Dr. Anthony DeNicola and Dr. Mancheng Tang, as well as Dr. Wei Cheng and Dr. Bruno Perlatti, who performed the majority of the NMR analysis, and Dr. John Billingsley and Dr. Undraa Bat-Erdene for all of their support. Last but not least, I would like to thank my parents for all of their support and encouragement during my graduate studies and beyond.

VITA

2012 – 2016 Princeton University
B.S. in Chemical and Biomolecular Engineering
Certificates in Materials Science, Engineering Biology, Quantitative Computational Biology, Applied Mathematics
Princeton, NJ

2020 – 2021 UCLA Dissertation Year Fellowship
University of California, Los Angeles
Los Angeles, CA

PRESENTATIONS

Keystone Symposia on Natural Products and Cellular Biology: Natural Products and Synthetic Biology: Parts and Pathways (01/2018)
Poster: Yee, Danielle; DeNicola, Anthony; Billingsley, John; Tang, Yi.
Compartmentalization for geraniol production in Saccharomyces cerevisiae

PUBLICATIONS

Yee, D. A.*; Kakule, T. B.*; Cheng, W.*; Chen, M.; Chong, C. T.; Hai, Y.; Hang, L. F.; Hung, Y. S.; Liu, N.; Ohashi, M.; Okorafor, I. C.; 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**, *142*, 710-714.

Yee, D. A.; DeNicola, A. B.; Billingsley, J. M.; Creso, J. G.; Subrahmanyam, V.; Tang, Y. Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metabolic Eng.* **2019**, *55*, 76-84.

Yee, D. A., Tang, Y. “Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: *Aspergillus nidulans* as a Heterologous Host” *Engineering Natural Product Biosynthesis*. **2022**, 41-52.

Tang, M. C.*; Zou, Y.*; Yee, D. A.; Tang, Y. Identification of the pyranonigrin A biosynthetic gene cluster by genome mining in *Penicillium thymicola* IBT 5891. *AIChE Journal* **2018**, *64*, 4182-4186.

1. INTRODUCTION

The term “natural products” refers to any chemical moiety or other material that comes from a living organism. In the field of natural product biosynthesis, the definition of natural products is often restricted to organic molecules produced from natural sources. Natural products can be classified into two broad categories: primary and secondary metabolites. Primary metabolites are small molecules required for basic metabolic pathways that are essential for survival of the host organism. On the other hand, secondary metabolites are not necessary for survival but can play countless roles in defense mechanisms and interspecies communication amongst various organisms.[1] Because of their widespread use in nature, many secondary metabolites have rich biological activity and can serve as inspiration for pharmaceutical drugs, herbicides, insecticides, etc. It should be noted that primary metabolites can also have powerful biological activity and should be considered for these applications as well. Withstanding this fact, “natural products” will refer to organic compounds from secondary metabolism throughout the rest of this work. This introductory chapter will give an overview of the enzymes commonly involved in the biosynthesis of these molecules.

Natural products are found in nearly all kingdoms of life, including bacteria, fungi, plants, and animals. Some notable examples with potent bioactivity include rapamycin, an immunosuppressant from the bacteria *Streptomyces hygroscopicus*,[2] taxol, an anticancer drug from Pacific yew *Taxus brevifolia*,[3] aspterric acid, a terpene herbicide from the filamentous fungus *Aspergillus terreus*,[4] and artemisinin, an antimalarial drug from the plant *Artemisia annua*[5] (Figure 1.1). Natural product biosynthesis often begins with formation of the compound’s scaffold by a core enzyme. The main classes of biosynthetic core genes include

polyketide synthases, nonribosomal peptide synthetases, and terpene cyclases. Polyketide synthases (PKSs) have many similarities to fatty acid synthases and condense units of malonyl-CoA to form carbon chains, often with subsequent modifications to the ketone group. The minimal domains of the PKS include the β-ketoacyl synthase (KS), the acyl carrier protein (ACP), and the malonyl-CoA ACP transacylase (MAT). Additional modification domains include the ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER).

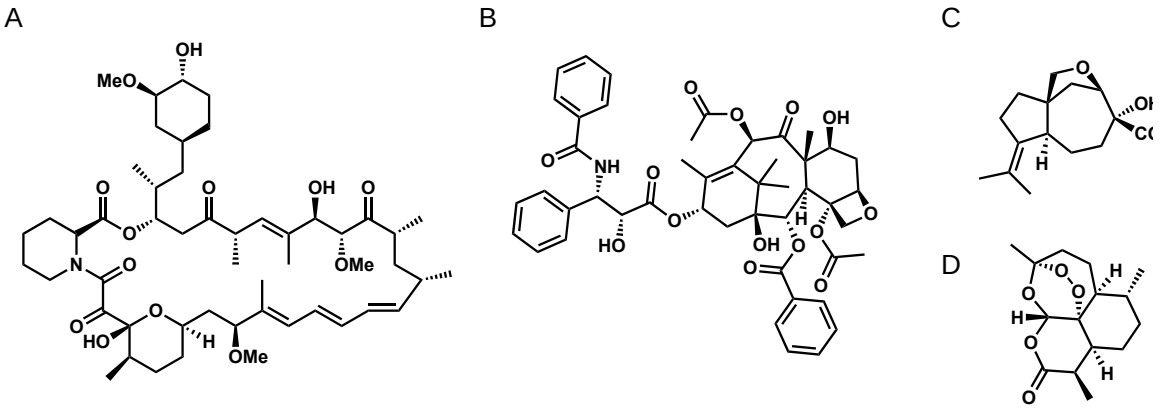


Figure 1.1 Structures of natural products with notable bioactivities: (**A**) rapamycin, an immunosuppressant derived from a PKS, (**B**) taxol, a terpene anticancer drug, (**C**) aspterric acid, a terpene herbicide, and (**D**) artemisinin, a terpene antimalarial drug.

Type I PKSs have two subcategories: iterative PKSs composed of a single module with multiple domains that perform iterative rounds of catalysis (most commonly found in fungi), and multi-modular PKSs in which each module performs a single iteration of chain extension and optional modifications (predominantly found in bacteria).[6] Type II PKS are complexes of multiple enzymes that only perform one round of catalysis each and are mostly found in bacteria.[7] Type III PKS perform iterative condensation of malonyl-CoA but do not contain an ACP domain.[8] One subclass of type III PKSs is the chalcone synthase superfamily, which condense the starter molecule with units of malonyl-CoA to form numerous aromatic

compounds, such as flavonoids, stilbenes, circuminoids, and pyrones.[8] Some notable examples of bioactive molecules formed by this enzyme superfamily include quercetin, resveratrol, genistein, catechins, and anthocyanins.[9]

Nonribosomal peptide synthetases (NRPSs) catalyze peptide and ester bonds to condense amino acids and carboxylic acids and are commonly found in fungi and bacteria. These enzymes produce many pharmaceutically valuable compounds such as cyclosporine A (immunosuppressant), cephalosporin and penicillin (antibiotic), and echinocandin (antifungal). NRPSs can add greater structural diversity to the resulting compound compared to ribosomally-produced peptide products because they can incorporate non-proteinogenic amino acids and carboxylic acids. Furthermore, NRPS products can undergo cyclization during chain release in contrast to linear peptide products formed by the ribosome.[10] NRPSs are multimodular enzymes in which the growing peptide chain is passed down through each module, which performs one round of catalysis, before chain release.[11] The minimal domains of the NRPS include the adenylation (A) domain, thiolation (T) domain, and condensation (C) domain. The A domain selectively activates the amino acid or carboxylic acid substrate using ATP, which is then loaded onto the phosphopantetheine (PPant) arm of the T domain. Optional tailoring domains can epimerize or methylate the tethered substrate.[10] The C domain then catalyzes peptide or ester bond formation between two substrates tethered to adjacent modules, or in the case of a single modular NRPS, between the tethered substrate and the second substrate that is not covalently bound to the enzyme. The poplypeptide chain is released from the NRPS through the thioesterase (TE) domain, which results in a macrocyclic or linear product. Cyclization may confer resistance to proteolytic degradation.[11]

Terpene cyclases/synthases (TCs/TSs) cyclize or hydrolyze isoprenoid diphosphates to afford diverse carbon scaffolds. Although thousands of terpenoid compounds have been isolated, these products are formed from one of five isoprenoid precursors, which demonstrates the immense structural diversity that can be achieved by these enzymes.[12] Because of this, terpene cyclization is often referred to as “chemical wizardry”.[13] All terpene natural products are derived from the five-carbon building blocks dimethylallyl diphosphate and isopentenyl diphosphate produced by the mevalonate pathway. TCs/TSs utilize geranyl diphosphate to produce monoterpenes, farnesyl diphosphate to produce sesquiterpenes, geranyl geranyl diphosphate to produce diterpenes, geranyl farnesyl diphosphate to produce sesterterpenes, and C30 diphosphates to produce triterpenes.[12] Class I TCs remove the substrate’s diphosphate group to afford an allylic cation using metal clusters, usually Mg^{2+}, coordinated by polar and acidic amino acids. Class II TCs protonate a double bond in the isoprenoid substrate to form the initial carbocation. In both cases, formation of the reactive cationic species results in subsequent cyclization to afford the terpene scaffold.[13] Well-studied examples of TCs include trichodiene synthase, aristolochene synthase, taxadiene synthase, and squalene-hopene cyclase.[12] Another diverse class of molecules include the meroterpenoids, which are terpene-polyketide hybrids formed by membrane-bound meroterpenoid cyclases. Several meroterpenoids have potent bioactivity, such as pyripyropene A (inhibitor of sterol O-acyltransferase 2), mycophenolic acid (immunosuppressant, inhibitor of IMPDH), and subglutinol A (immunosuppressant, more potent and less toxic than cyclosporine A). The first meroterpenoid cyclase to be discovered was Pyr4 from the pyripyropene biosynthetic pathway from *Aspergillus fumigatus*. After

formation of the polyketide moiety, a prenyltransferase transfers one unit of farnesyl diphosphate (FPP) to the scaffold and following further oxidation, Pyr4 catalyzes cyclization of the prenyl group. The cyclized intermediate is further modified by additional tailoring genes to arrive at the final product, pyripyropene A.[14]

Another less well-studied class of natural products is RiPPs, which are ribosomally synthesized and post-translationally modified peptides. Almost all RiPPs are biosynthesized by a precursor peptide, which usually contains a leader sequence at the *N* or *C* terminus that is often used for substrate export and recognition by post-translational modification enzymes. Common post-translational modifications include Ser/Thr dehydratases, radical SAM methyltransferases, head-to-tail cyclization, and conversion of cysteine to disulfides, thioethers, sulfoxides, and additional sulfur-containing functional groups. Leader peptides can be removed in one cleavage step or in several proteolytic steps. Using heterologous expression to study RiPPs biosynthesis is advantageous because the biosynthetic pathways are often short and production is not limited by precursor supply. Major classes of RiPPs include lanthipeptides, linaridins, proteusins, linear azole-containing peptides, cyanobactins, thiopeptides, bottromycins, microcins, lassopeptides, microviridins, etc. Several of these classes of molecules have strong bioactivities, such as cytotoxic, anticancer, and antibacterial properties.[15]

An additional class of non-canonical biosynthetic enzymes that can produce natural product scaffolds is the family of NRPS-independent siderophore (NIS) synthetases. Siderophores are ferric iron chelators excreted by numerous microbial organisms that allow these species to uptake adequate amounts of iron. After binding to ferric iron outside the cell,

siderophores are transported back inside the cell, and the iron is released from the siderophores by a reductase enzyme. Siderophores can be produced by the traditional NRPS or an NIS synthetase, which usually catalyzes condensation of citric acid with an amino acid or alcohol. The majority of characterized NIS enzymes are from bacteria with few examples in fungi. Despite this fact, bioinformatics analysis has predicted many putative NIS sequences in fungal genomes. All NIS synthetases contain an iron uptake chelate (Iuc) conserved domain at the *N*-terminus. NIS synthetases activate citric acid through adenylation, but the adenylation domain is distinct from that of the NRPS and aminoacyl-tRNA synthetases. Achromobactin is a siderophore produced from *Dickeya dadantii*, *Pseudomonas syringae*, and *Sodalis glossinidius* synthesized by NIS enzymes AcsD, AcsC, and AcsA, which condense citric acid with serine, 2,4 diaminobutyrate, and α-ketoglutarate. Another example is the siderophore aerobactin, which is produced in numerous *Vibrio* and *Yersinia* species. In this pathway, lysine is first oxidized by IucD and acetylated by IucB to form ah-Lys. Then the NIS IucA condenses citric acid with ah-Lys, and a second NIS IucC condenses citryl-ah-Lys with an additional unit of ah-Lys to afford aerobactin.[16]

Natural products scaffolds produced by core enzymes are often modified by tailoring genes. Examples of tailoring enzymes that catalyze redox reactions include the cytochrome P450, the flavin-dependent monooxygenase (FMO), the α-ketoglutarate oxygenase (αKG), and the short-chain dehydrogenase/reductase (SDR). Cytochrome P450s are typically membrane-bound enzymes containing heme targeted to the endoplasmic reticulum and utilize molecular oxygen to carry out numerous oxidative reactions such as hydroxylation, desaturation, epoxidation, and aromatic ring coupling. The first step in the reaction is electron

transfer from the partner reductase enzyme through consumption of NADPH or NADH to reduce the Fe(III) center to Fe(II). O_2 can then bind to form a Fe(III)-oxygen bond. A second electron is transferred to afford the iron(III)-peroxo complex, which is quickly protonated. A second proton is transferred to form water and the Fe(V)-oxo species, which can then be used to catalyze oxidation of the substrate. Hydroxylation can occur through a concerted or oxygen rebound mechanism. The rebound mechanism generates the Fe(IV)-OH species through H-abstraction of the substrate and can result in lack of stereo-specificity due to formation of the radical substrate intermediate. Epoxidation of olefins can occur through conversion of Fe(V) to the Fe·(IV)-O· radical species followed by radical rearrangement. Transfer of the iron bound oxygen to the substrate allows for regeneration of the Fe(III) species.[17]

Although there are many other tailoring enzymes not mentioned here, additional examples include methyltransferases, PLP-dependent enzymes, decarboxylases, glucosyltransferases, acetyltransferases, Diels-alderases, and halogenases. PLP-dependent enzymes utilize the cofactor pyridoxal-5'-phosphate (PLP) to catalyze myriad reactions, such as racemization, transamination, alpha elimination, beta elimination, gamma elimination, beta addition, etc.[18] Decarboxylases can form amine substrates that can be used in the biosynthesis of numerous alkaloid compounds, including tryptophan decarboxylase in the strictosidine pathway from the plant *Catharanthus roseus*[19] and dimethyl-lysine decarboxylase in the flavunoidine pathway from the filamentous fungus *Aspergillus flavus*.[20]

2. INVESTIGATING FUNGAL BIOSYNTHETIC PATHWAYS USING HETEROLOGOUS GENE EXPRESSION IN *ASPERGILLUS NIDULANS*

Fungal natural products encompass an important source of pharmaceutically relevant molecules. Heterologous expression of biosynthetic pathways in chassis strains enables the discovery of new secondary metabolites and characterization of pathway enzymes. In our laboratory, biosynthetic genes in a clustered pathway have been refactored in engineered heterologous hosts such as *Aspergillus nidulan*s. Here we describe the assembly of heterologous expression vectors, transformation into *A. nidulans*, and detection of new compounds in the transformant strains.

2.1 Background on heterologous expression in *A. nidulans*

Secondary metabolites from fungi and their derivatives comprise a broad class of natural products with rich structural diversity and powerful biological activity.[21] Recent genome sequencing efforts have revealed many natural product gene clusters in fungi are not expressed under standard laboratory culturing conditions.[22] To access silent gene clusters, pathway genes can be reconstituted in well-established heterologous hosts such as *Aspergillus nidulans* and *Aspergillus oryzae*.[23–28] Use of these platforms has allowed for the discovery of countless natural products and characterization of their biosynthetic pathways.[29,30] Here we describe an episomal heterologous expression system in *A. nidulans* A1145, which can accommodate up to 12 genes on three plasmids.[31] For a cleaner background during compound detection, the modified strain *A. nidulans* A1145ΔSTΔEM can be used, in which endogenous production of sterigmatocystin[32] and emericellamide[33] has been abolished.[34] The general workflow for this system is as follows: first, heterologous

expression plasmids for *A. nidulans* are assembled through yeast homologous recombination. Following this, the plasmids are transformed into *A. nidulans* protoplasts. Individual transformant colonies are then assayed for compound production. RT-PCR (reverse transcription-polymerase chain reaction) can be performed to verify genes are successfully expressed in the transformant strains.

2.2 Materials

Strains

1. *A. nidulans* A1145[31] (genotype: *pyrG89; pyroA4; nkuA::argB; riboB2*) or *A. nidulans* A1145ΔSTΔEM[34]
2. *S. cerevisiae* JHY651[35] or other yeast strain auxotrophic for uracil
3. *E. coli* TOP10 (Invitrogen), DH10b (Thermo Scientific), etc.
4. Original fungal host strain with target gene cluster (if genomic DNA needs to be acquired)

Plasmids

1. pYTR (see Note 1)
2. pYTP (see Note 2)
3. pYTU (see Note 3)

Reagents for cloning

Many of the commercial kits are optional. If following other protocols, additional materials are required.

1. Restriction enzymes PacI, NotI, SwaI, BamHI, PshAI
2. Agarose gel for DNA electrophoresis

3. Gel DNA Recovery Kit

4. *Quick*-DNA Fungal/Bacterial Microprep Kit (Zymo Inc. USA)

5. Primers to amplify genes and promoters

6. Proof-reading DNA polymerase

7. Solid YPEG medium: 5 g bacto-peptone, 2.5 g yeast extract, 7.5 mL glycerol, 5 g agar, 250 mL ddH_2O. Autoclave and let solution cool to around 55 ^oC. Add 10 mL of 100% ethanol. Mix thoroughly and pour plates.

8. Liquid YPD medium: 5 g bacto-peptone, 2.5 g yeast extract, 5 g dextrose, 250 mL ddH_2O. Autoclave and store at room temperature.

9. Frozen-EZ Yeast Transformation II Kit (Zymo Inc. USA)

10. Solid SDC_{AA}(-U) medium: 5 g dextrose, 5 g agar, 200 mL ddH_2O. Autoclave. Prepare supplements: 1.25 g casamino acids, 1.7 g yeast nitrogen base without amino acids, 10 mg adenine, 10 mg tryptophan, 50 mL ddH_2O. Filter sterilize. Add supplements to melted agar, mix thoroughly, and pour plates.

11. Zymoprep™ Yeast Plasmid Miniprep I (Zymo Inc. USA)

12. Isopropanol

13. 1000X carbenicillin (50 mg/mL): Dissolve 0.5 g of carbenicillin in 10 mL ddH_2O. Filter sterilize. Prepare 1 mL aliquots and store at -20 ^oC.

14. 1000X ampicillin (100 mg/mL): Dissolve 1 g of ampicillin in 10 mL ddH_2O. Filter sterilize. Prepare 1 mL aliquots and store at -20 ^oC.

15. Solid LB medium: 6.25 g LB powder, 5 g agar, 250 mL ddH_2O. Autoclave. To prepare LB plates with carbenicillin, let melted agar cool to around 55 ^oC, then add 250 µL of 1000X carbenicillin. Mix thoroughly and pour plates.

16. Liquid LB medium: 6.25 g LB powder, 250 mL ddH_2O. Autoclave and store at room temperature. Aliquot and add ampicillin as needed.

17. Plasmid Miniprep Kit

18. Yeast RNA Isolation Kit

19. SuperScript III First-Strand Synthesis System (Invitrogen)

20. Oligo-dT primer

21. GoTaq Master Mix (Promega)

Media and solutions for *A. nidulans* transformation and culturing

1. 100X uracil (500 mM): 2.24 g, 40 mL of ddH_2O. Add 10N NaOH until fully dissolved. Filter sterilize.

2. 100X uridine (1 M): 9.76 g, 40 mL of ddH_2O. Filter sterilize.

3. 1000X pyridoxine (0.5 mg/mL): 20 mg, 40 mL of ddH_2O. Filter sterilize.

4. 1000X riboflavin (0.125 mg/mL): 5 mg, 40 mL of ddH_2O. Filter sterilize.

5. 20X nitrate salts: 120 g $NaNO_3$, 10.4 g KCl, 10.4 g $MgSO_4 \cdot 7H_2O$, 30.4 g KH_2PO_4. Add ddH_2O. up to 1 liter, stir until dissolved, and store at room temperature.

6. Trace elements: 2.20 g $ZnSO_4 \cdot 7H_2O$**,** 1.10 g H_3BO_3**,** 0.50 g $MnCl_2 \cdot 4H_2O$,0.16 g $FeSO_4 \cdot 7H_2O$**,** 0.16 g $CoCl_2 \cdot 5H_2O$**,** 0.16 g $CuSO_4 \cdot 5H_2O$**,** 0.11 g $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$. Add each to 80 mL of ddH_2O in the order shown. Add more ddH_2O to bring the total volume to 100 mL. Adjust pH to 6.5 with 1 N KOH.

7. Solid CD-sorbitol medium (1 L): 10 g glucose, 50 mL 20X nitrate salts, 1 mL trace elements, 218.6 g sorbitol (1.2 M), 20 g agar. Autoclave then pour plates as needed. Add the appropriate supplements to individual empty petri dishes before pouring in the agar media. If not making plates immediately, solid media can be stored at room temperature and re-melted in the microwave.

8. Solid CD medium (1 L): Prepare as described above for solid CD-sorbitol medium, omitting the sorbitol.

9. Liquid CD medium (1 L): 10 g glucose, 50 mL 20x nitrate salts, 1 mL trace elements. Autoclave. Aliquot as needed and add the appropriate supplements.

10. Liquid CD-ST medium (1 L): 20 g starch, 20 g peptone (acidic digest) or casamino acids (acidic digest), 50 mL 20x nitrate salts, 1 mL trace elements. Add starch to 100 mL of ddH_2O and mix with a stir bar for 10 min. Add boiling ddH_2O to bring the total volume to 950 mL. Add the remaining ingredients. Continue to mix until everything is dissolved. Autoclave. Aliquot as needed and add the appropriate supplements.

11. Solid CD-ST medium (1 L): Prepare as described above for liquid CD-ST medium. Add 20 g of agar after everything is dissolved. Autoclave then prepare plates with the appropriate supplements.

12. Osmotic medium (500 mL): 147.9 g $MgSO_4$ (1.2 M), 10mM sodium phosphate buffer (NaPB) (can be made from a 2M NaPB stock: 90.9 g Na_2HPO_4 and 163.4 g NaH_2PO_4 or 187.9 g NaH_2PO_4per liter, pH 6.5). Adjust pH to 5.8 with 1M Na_2HPO_4 (about 25 mL). Filter sterilize and store at 4^oC. Tip: start with 450 mL of water, then the total volume will end up to be about 500 mL.

13. Trapping Buffer (1 L): 109.3 g sorbitol (0.6 M), 0.1 M Tris-HCI, pH 7.0 (can be made using 100 mL 1 M Tris). Autoclave and store at 4°C.

14. STC buffer (1 L): 218.6 g sorbitol (1.2 M), 1.47 g $CaC1_2$ (10mM), 10 mM Tris-HCI, pH 7.5 (can be made using 10 mL 1 M Tris). Autoclave and store at 4°C.

15. PEG solution (100 mL): 60% PEG 4000 (BDH), 50 mM $CaCl_2$, 50 mM Tris-HCI, pH 7.5 (can be made using 50 mL 1 M Tris). Autoclave and store at room temperature.

16. Lysing enzyme, yatalase, 0.2 μm syringe filters

Other materials for *A. nidulans* transformation, culturing, and analysis

1. Two sterile 125-mL flasks
2. Sterile 30 mL Corex tube or sterile 50 mL Falcon tube
3. Sterile cell strainer (Fisher, Cat No. 22363547, optional)
4. Solvents e.g. ethyl acetate, acetone, methanol

2.3 Methods

Cloning of plasmids for *A. nidulans* expression

The plasmids for heterologous expression in *A. nidulans* A1145 are denoted as pYTU, pYTP, and pYTR[34]. Their features include auxotrophic markers for uracil (*pyrG*), pyridoxine (*pyroA*), and riboflavin (*riboB*), respectively, as well as the AMA1 origin of replication for *Aspergillus* (Fig. 1). For cloning purposes, these vectors also contain the uracil auxotrophic marker (*URA3*) and 2-μm origin for *S. cerevisiae*, and the ampicillin resistance marker (*ampR*) and ColE1 origin for *E. coli*. pYTU contains the starch inducible promoter P_{glaA}, pYTP contains the starch inducible promoter P_{amyB}, and pYTR contains the constitutive promoter P_{gpdA} from *Aspergillus niger*. Additional commonly used promoters include the constitutive P_{gpdA} promoters from *Penicillium oxalicum* (P_{POgpdA}) and *Penicillium expansum* (P_{PEgpdA}), as well as

the constitutive P_{coxA} promoter from *A. niger*. Typically, one to four genes expressed under different promoters can be inserted into each plasmid.

1. If using the original promoters included in the pYT vectors, digest pYTU with the restriction enzymes PacI/NotI and pYTP and pYTR with PacI/SwaI or BamHI/SwaI following the manufacturer's instructions. If using different promoters, digest pYTU with PshAI/NotI, or PshAI/PacI, and pYTP and pYTR with NotI/PacI, NotI/BamHI, or NotI/SwaI (see Note 1). Run digestion reactions on an agarose gel and recover DNA using a commercial kit such as Zymoclean™ Gel DNA Recovery Kit.

2. To obtain template DNA for amplification of target genes, culture the original fungal host on solid or liquid media such as potato dextrose agar or broth. Isolate the strain's genomic DNA using a commercial kit such as *Quick*-DNA Fungal/Bacterial Microprep Kit, phenol-chloroform extraction,[36] or microwave total genomic DNA extraction.[37]

3. Amplify the genes of interest and their native terminators (300-500 bp downstream from the stop codon) by PCR using the genomic DNA of the original host as the template. A high-fidelity polymerase such as Q5® High-Fidelity DNA Polymerase (NEB), AccuPrime Pfx DNA Polymerase (Invitrogen), or Phusion (NEB) should be used. If replacing the original promoters or if more than one gene is to be inserted into each plasmid, amplify the promoters to be used by PCR. Use overhang primers to introduce 25-40 bp of homology between adjacent fragments. Run PCR products on an agarose gel and recover DNA using a commercial kit such as Zymoclean™ Gel DNA Recovery Kit.

4. Co-transform the overlapping DNA fragments and their corresponding digested vectors into *S. cerevisiae* JHY651[3] or other yeast strain auxotrophic for uracil to assemble the expression plasmids *in vivo* by yeast homologous recombination. Before preparing competent cells, yeast strains should be grown on YPEG agar plates and cultured in YPD liquid media. Yeast competent cells can be prepared using the PEG-lithium acetate method[38] or a commercial kit such as Frozen-EZ Yeast Transformation II Kit (Zymo Inc. USA).

Transformants should be plated on SDC_{AA}(-U) agar plates. Yeast colonies will appear after 1-3 days of incubation at 28°C.

5. Extract the assembled plasmids from yeast using a commercial kit such as Zymoprep™ Yeast Plasmid Miniprep I (Zymo Inc. USA), and transform into *E. coli* TOP10, DH10b, etc. by electroporation to isolate single clones. Plate cells on LB agar plates with 50 µg/mL of carbenicillin. *E. coli* colonies will appear after 12-16 hours of incubation at 37 °C.

6. Culture single *E. coli* colonies in 3-5 mL of LB medium with 100 µg/mL of ampicillin for 12-16 hours, shaking at 250 rpm, 37°C. Extract the plasmids from the overnight cultures using a commercial kit such as Zyppy™ Plasmid Miniprep Kit (Zymo Inc. USA) or QIAprep Spin Miniprep Kit (Qiagen). Confirm correct assembly by performing digestion checks with restriction enzymes followed by DNA sequencing.

Procedure for transformation of plasmids into *A. nidulans*

Culturing of germlings

1. Grow *A. nidulans* A1145[11] or A1145ΔSTΔEM[34] on CD agar supplemented with uracil, uridine, pyridoxine, and riboflavin for 3-5 days at 37°C. If using cells from a frozen protoplast stock, streak the strain on CD-sorbitol agar with the appropriate supplements.

2. When green spores appear, inoculate 25 mL of liquid CD medium containing the appropriate supplements in a sterile 125-mL flask with fresh spores (2×10^6 spores/mL, about 3 cm^2 of spores) using a sterile cotton-tipped applicator and shake at 28°C, 250 rpm for 16-20 hours. A proper culture should have an abundance of young germlings in small aggregates.

Digestion of germlings

1. Harvest the culture by centrifugation at 4300 x *g* for 20 min at 20°C.

2. Remove the supernatant and resuspend the pellet in 10 mL of osmotic medium. Spin down by centrifugation at 4300 x *g* for 20 min at 20°C.

3. Dissolve 30 mg of lysing enzymes from *Trichoderma harzianum* and 20 mg of Yatalase in 10 mL of osmotic medium and sterilize with a 0.2 µm syringe filter.

4. Remove the supernatant from the centrifuged cells and resuspend the pellet in the sterilized lysing enzyme mixture. Transfer directly into a sterile 125-mL flask.

5. Digest cells by shaking at 80 rpm for 4-6 hours at 28^{o}C. Protoplasts are thin-walled and about 2 times as large as spores.

Harvesting digested cells

1. Pour cells directly into an autoclaved 30 mL Corex tube and gently overlay with 10 mL of trapping buffer. Centrifuge in Beckmann equivalent rotor at 5200 x *g* for 20 min at 4^{o}C (see Note 5).

2. After centrifugation, protoplasts will accumulate in the cloudy layer at the interface of the two buffers. Remove the protoplasts from the interface with a pipet and transfer them to a sterile 15 mL falcon tube.

3. Add 2 volumes of STC buffer and centrifuge at 4300 x *g* for 20 min at 4^{o}C (see Note 6).

4. Decant the supernatant. Resuspend the protoplasts in STC buffer at a concentration of 10^{8}- 10^{9} (usually 1 mL) with minimal pipetting, which can damage the protoplasts. Aliquot 100 µL of protoplasts each in sterile 1.5 mL centrifuge tubes. If desired, store aliquots for future use at -80^{o}C (see Note 7).

Transformation

1. For each transformation, add 3 µL of each plasmid to 100 µL of protoplasts, and incubate on ice for one hour. As a negative control, include one transformation with the appropriate empty pYT vectors. Plasmid stock DNA concentrations should be at least 100 ng/µL (see Note 8).

2. Add 600 µL of PEG solution to each tube. Mix gently by turning the tube on its side and rotating it. Incubate at room temperature for 20 min.

3. Using a pipet, drop the PEG mixture on CD-Sorbitol agar plates with the appropriate supplements if transforming only one or two plasmids. Using a spreader is not necessary and

can damage the protoplasts. Incubate at 37 ^{o}C right side up to let dry. Colonies will appear after 2-4 days.

Procedure for production of compounds and biotransformation

Production of compounds

1. For expression of genes under starch inducible promoters, transformants must be cultured in CD-ST media. For expression of genes under constitutive promoters, CD or CD-ST media may be used (see Note 9). For liquid cultures, inoculate the spores from selected transformants in 10 mL of liquid media with appropriate supplements using a sterile cotton-tipped applicator. Shake cultures at 250 rpm at 28°C. For solid cultures, use a sterile cotton-tipped applicator to streak selected transformants onto CD or CD-ST agar plates. Incubate at 28°C.

2. Check cultures for compound production every other day, beginning after two days of growth. For liquid cultures, use a pipet with autoclaved cut tips to transfer 500 µL of culture to a 1.5 mL microcentrifuge tube (see Note 10). Centrifuge at maximum speed for 5 min. Transfer the supernatant to a new microcentrifuge tube. Add 500 µL ethyl acetate to the supernatant and vortex for 1 min. Add 500 µL acetone or methanol (see Note 11) to the cell pellet and vortex for 15 min. For solid cultures, cut out a 1 cm^2 piece of agar and transfer to a microcentrifuge tube with 500 µL of acetone or methanol and vortex for 15 min.

3. Centrifuge the samples at max speed for 5 min. Transfer the organic layers to new microcentrifuge tubes and dry in a speed vacuum.

4. Resuspend the dried extracts in 100 µL of methanol. Centrifuge at maximum speed for 5 min. and transfer 50 µL of the extract to a LC-MS vial. Inject 20 µL of extract on the LC-MS for compound detection.

Biotransformation

1. Use the same culturing conditions as described previously for compound production. Substrates for the biotransformation assay can be added at the beginning of culturing or after the cells have grown for a few days. Add the substrate to the culturing medium to a final

concentration of 200-500 μM. If adding substrates at the beginning of a solid culture, allow melted agar to cool before adding the substrate. If feeding to a solid culture after cells have grown, the substrate can be layered on top of the plate.

2. To check cultures for biotransformation products, follow the same procedure for metabolite analysis as described in steps 2-4 above.

Procedure for RT-PCR (reverse transcription-polymerase chain reaction) to verify gene expression

1. Use the same culturing conditions as described previously for compound production and biotransformation. Cells can be harvested for RNA extraction when compounds are expected to be produced.

2. Extract RNA from the cells using RiboPure™ Yeast RNA Isolation Kit (Ambion) following the manufacturer's instructions. Digest residual genomic DNA in the RNA extracts with DNase I (provided in the kit) at 37 °C for 4-6 hours. Follow the manufacturer's instructions to inactivate the DNase.

3. Use SuperScript III First-Strand Synthesis System (Invitrogen) for cDNA synthesis with oligo-dT primers following directions from the user manual.

4. Using the synthesized cDNA as the template, set up PCR reactions to amplify fragments of the genes for heterologous expression using GoTaq Master Mix (Promega) following the manufacturer's instructions. To check for gene expression, it is not necessary to amplify the entire open reading frames of the genes. If possible, design primers flanking the introns in the target genes so the smaller cDNA product can be separated from possible bands from genomic DNA contamination. As a positive control, include a PCR reaction to amplify a region of *actA* (actin gene for *A. nidulans*) flanking its introns to assess the quality of the cDNA. Run the PCR reactions on an agarose gel. Bands with the correct cDNA product size will indicate the genes are expressed (see Note 12).

2.4 Notes

1. Selection is by *riboB*, final concentration 0.125 μg/mL riboflavin.

2. Selection is by *pyroA*, final concentration 0.5 µg/mL pyridoxine HCl.

3. Selection is by *pyrG*, final concentration 10 mM uridine and 5 mM uracil.

4. If using SwaI, perform the digestions with each enzyme separately. First, digest with only SwaI using NEB buffer 3.1 at 25 °C, then column purify the reaction with a commercial kit such as DNA Clean & Concentrator (Zymo Inc. USA). Next, digest the recovered DNA with the second restriction enzyme using NEB Cutsmart® buffer at 37°C.

5. Alternatively, pour cells into a sterile 50 mL Falcon tube and gently overlay with 10 mL of trapping buffer. Centrifuge at 4300 xg for 30 min at 4ºC.

6. As another alternative to steps 1-3, pour cells through a sterile cell strainer (Fisher, Cat No. 22363547) into a sterile 50 mL Falcon tube. Add an equal volume of STC buffer and gently invert the tube to mix. Centrifuge at 4300 x *g* for 20 min at 4ºC. Decant the supernatant and gently resuspend the pellet in 10 mL of STC buffer with minimal pipetting to wash the cells. Centrifuge at 4300 xg for 20 min at 4ºC.

7. Fresh protoplasts have the highest transformation efficiency. Frozen protoplasts can be used up to three weeks after storage but will rapidly lose competency.

8. It is optimal to add less than 10 µL of DNA solution to maintain the concentration of the STC buffer components. If plasmid DNA concentrations are low, plasmid stocks eluted in water can be lyophilized and resuspended in a smaller volume of STC buffer to increase DNA concentration.

9. Strains grow faster and to a higher cell density in CD-ST, which is a richer medium compared to CD. However, the background metabolite profile is generally cleaner in CD compared to CD-ST.

10. Uncut pipet tips may be used in early stages of culturing. It is necessary to use cut tips once the mycelium has grown too thick to be aspirated by an uncut pipet tip.

11. Depending on their solubility, some compounds may only be detected by extracting with certain solvents but not others. Therefore, it may be advantageous to test different solvents for extraction if target compounds cannot be detected.

12. To confirm genes are properly spliced, the entire open reading frames of the genes can be amplified by PCR using a high-fidelity polymerase. The resulting PCR products can be purified and sent for DNA sequencing.

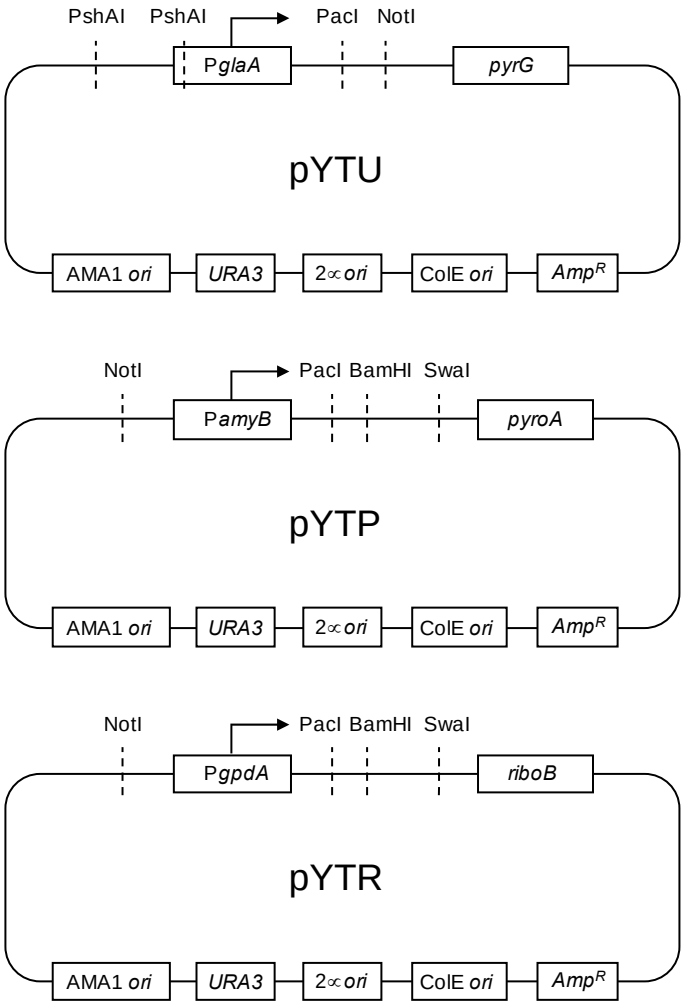


Figure 2.1 Vectors for heterologous expression in *A. nidulans*: pYTU, pYTP, and pYTR.

* Section 2 is part of my submitted methods article “Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: *Aspergillus nidulans* as a Heterologous Host”.

3. ENGINEERED MITOCHONDRIAL PRODUCTION OF MONOTERPENES IN *SACCHAROMYCES CEREVISIAE*

Monoterpene indole alkaloids (MIAs) from plants encompass a broad class of structurally complex and medicinally valuable natural products. MIAs are biologically derived from the universal precursor strictosidine. Although the strictosidine biosynthetic pathway has been identified and reconstituted, extensive work is required to optimize production of strictosidine and its precursors in yeast. In this study, we engineered a fully integrated and plasmid-free yeast strain with enhanced production of the monoterpene precursor geraniol. The geraniol biosynthetic pathway was targeted to the mitochondria to protect the GPP pool from consumption by the cytosolic ergosterol pathway. The mitochondrial geraniol producer showed a 6-fold increase in geraniol production compared to cytosolic producing strains. We further engineered the monoterpene-producing strain to synthesize the next intermediates in the strictosidine pathway: 8-hydroxygeraniol and nepetalactol. Integration of geraniol hydroxylase (G8H) from *Catharanthus roseus* led to essentially quantitative conversion of geraniol to 8-hydroxygeraniol at a titer of 227 mg/L in a fed-batch fermentation. Further introduction of geraniol oxidoreductase (GOR) and iridoid synthase (ISY) from *C. roseus* and tuning of the relative expression levels resulted in the first *de novo* nepetalactol production. The strategies developed in this work can facilitate future strain engineering for yeast production of later intermediates in the strictosidine biosynthetic pathway.

3.1 Background on MIAs and microbial production of precursors

Natural products from bacteria, fungi, and plants and their derivatives remain an

important source of pharmaceutical compounds.[21] Monoterpene indole alkaloids (MIAs) from plants comprise a diverse array of natural products with high structural complexity and useful biological activity, with properties ranging from antimalarial (quinine) to antihypertensive (ajmalicine) to anticancer (camptothecin, vinblastine, and vincristine).[39] Native plant producers remain the sole commercial source for many of these medicinally valuable natural products. MIAs such as vinblastine and vincristine are costly due to low yields from the plant extraction process, yet there is an increasing clinical demand for these pharmaceutical compounds.[40] Consequently, it is desirable to reconstitute plant MIA pathways in engineered microbial hosts for more economical and reliable production of MIAs.[41] Thousands of MIAs are derived from the universal precursor strictosidine, which is produced from the monoterpene precursor geraniol through eleven enzymatically catalyzed modifications.[42] The biosynthetic pathway for strictosidine was reconstituted in *Saccharomyces cerevisiae* upon an impressive number of genetic modifications.[43] Notwithstanding this feat, further strain engineering work is needed to optimize the yeast performance, both in terms of strain robustness and monoterpene titer. To this end, constructing a yeast strain that produces high amounts of monoterpene precursors is a necessary prerequisite.

Geraniol is natively produced in many plant species including *Ocimum basilicum*[44] and *Valeriana officinalis*[45] via hydrolysis of geranyl pyrophosphate (GPP) by geraniol synthase (GES). In MIA producing plants such as *Catharanthus roseus*, geraniol is hydroxylated by the cytochrome P450 geraniol hydroxylase (G8H) to form 8-hydroxygeraniol[46], which undergoes two oxidation steps by geraniol oxidoreductase (GOR) to form 8-oxogeranial[42]. 8-oxogeranial is cyclized by iridoid synthase (ISY) to form nepetalactol[47] (Fig. 3.1). Six additional enzymes

are required to further modify nepetalactol through oxidation, glucosylation, and methylation followed by condensation with tryptamine to form strictosidine.[42,48–51]

Industrial brewing yeast have recently been engineered for precise monoterpene production. Titers on the order of ~1 mg/L geraniol resulted in enhanced “hoppy” flavor of beer with minimal metabolic perturbation.[52] Engineering yeast for high titer geraniol production, however, has proven to be challenging due to the low supply of intracellular GPP.[53,54] The native farnesyl pyrophosphate (FPP) synthase in yeast, ERG20, produces GPP from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) and subsequently converts most of that GPP to FPP for biosynthesis of ergosterol and other primary metabolites.[55] Although previous strain engineering has significantly improved geraniol production in yeast,[53,54,56,57] most of these efforts utilize plasmid-based systems for overexpression of native cytosolic pathway genes. In this work, we aimed to develop a fully integrated strain for geraniol production to facilitate future optimization of later steps in the strictosidine pathway. A completely integrated strain would also allow for more stable expression of genes in the geraniol pathway and greater flexibility with culturing conditions compared to plasmid based systems.[58] Previous works have shown rapid conversion of cytosolic GPP to FPP,[54–56] prompting us to compartmentalize the geraniol biosynthetic pathway to the mitochondria. Using a fully integrated platform strain, we further extended the pathway to achieve the highest reported titers for *de novo* 8-hydroxygeraniol and nepetalactol production.

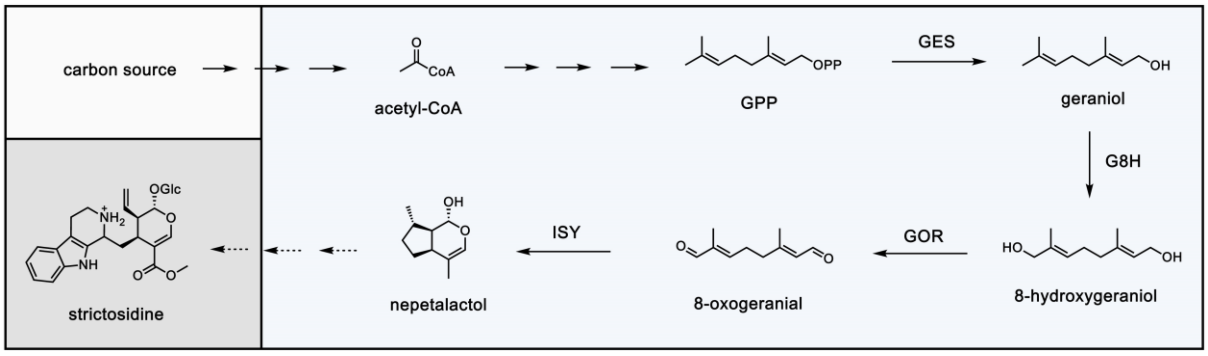


Fig. 3.1 Biosynthesis of strictosidine and its precursors. GPP is produced from acetyl-CoA through the mevalonate pathway and is then hydrolyzed by GES to form geraniol. In plant MIA producers, geraniol is hydroxylated by the cytochrome P450 G8H to afford 8-hydroxygeraniol, which is oxidized twice by GOR to form the dialdehyde 8-oxogeranial. 8-oxogeranial is reductively cyclized by ISY to form nepetalactol, which undergoes further modifications to yield strictosidine.

3.2 Selection of a yeast strain background and bioprospecting GPPS/GES

We began construction of a platform monoterpene yeast strain by screening different strain backgrounds for geraniol production: JHY651, CEN.PK2-1C, and X303-1B. JHY651 is descended from BY4741/BY4742 and has improved respiratory growth and increased mitochondrial stability.[35] CEN.PK2-1C is frequently used in industry and metabolic engineering applications.[59] X303-1B is derived from the W303 strain, which was selected for good sporulation and high transformation efficiency.[60] A low copy plasmid with geraniol synthase from *C. roseus* (CrGES) was transformed into these strains. JHY651 and X303-1B showed higher geraniol production (~ 0.3 mg/L) than CEN.PK2-1C (Fig. 3.2A). Given recent successes in refactoring natural product biosynthesis gene clusters into JHY-based strains,[35,61] we decided to use JHY651 as our starting strain. To facilitate rapid strain construction, iCas9 from *Streptococcus pyogenes* with the mutations D147Y and P411T[62] was integrated into JHY651 to create strain **S1**. Subsequent genome modifications were

performed through transformation of plasmids with target gRNA. CrGES was integrated into **S1** to create strain **S2**, which produced 0.32 mg/L geraniol (Fig. 3.2B).

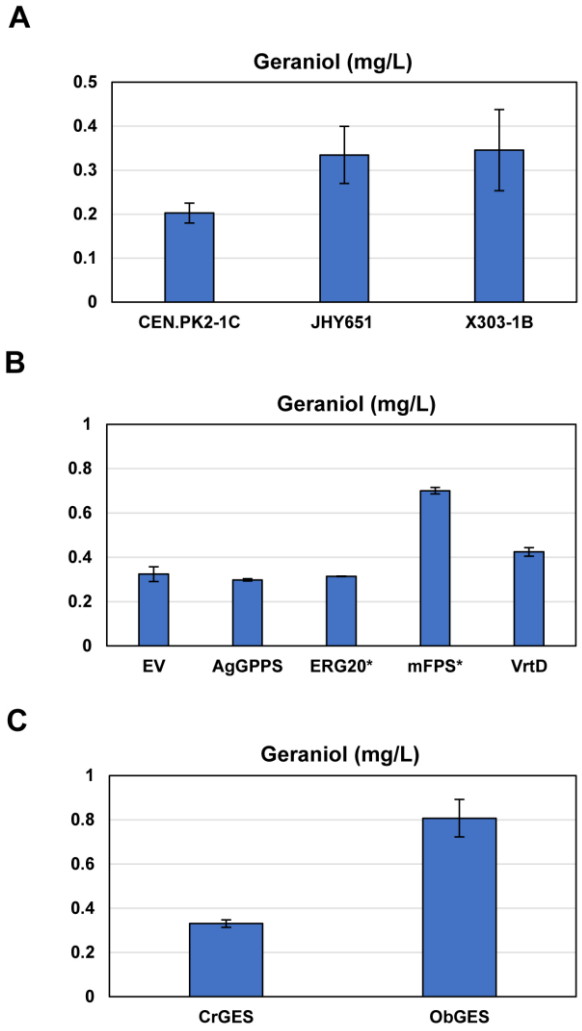


Fig. 3.2 Initial efforts to set up geraniol production platform strain. Plasmids transformed into each strain are designated as (p#). (**A**) Geraniol production in CEN.PK2-1C, JHY651, and X303-1B with episomal CrGES expression (p4). (**B**) Geraniol production in **S2** with an empty vector (EV, p1) and plasmids containing expression cassettes for *A. grandis* GPPS (AgGPPS, p6), *S. cerevisiae* ERG20*K197G (ERG20*, p7), *G. gallus* mFPS*N144W (mFPS*, p8), and *P. aethiopicum* VrtD (VrtD, p9). (**C**) Geraniol production in JHY651 with episomal *C. roseus* GES (CrGES, p4) and *O. basilicum* GES (ObGES, p5) expression.

Several GPP synthases or mutant FPP synthases that preferentially produce more GPP than FPP have been used to enrich yeast's intracellular GPP supply.[43,53,54,63] We compared the performance of GPP synthases from the plant *Abies grandis* (AgGPPS) and the

filamentous fungus *Penicillium aethiopicum* (VrtD)[64] as well as mutant FPP synthases from *S. cerevisiae* (ERG20*K197G) and the bird *Gallus gallus* (mFPS*N144W). AgGPPS, ERG20*, and mFPS have been utilized in engineered yeast strains for monoterpene production,[43,53,63] but no direct comparison has been made between all of them. Plasmids containing expression cassettes for these GPP synthases were transformed into **S2**. mFPS from *G. gallus* showed the highest increase in geraniol titer (Fig. 3.2B) and is used in most of our subsequent strain constructions. Contrary to previously reported results,[53,54] AgGPPS and ERG20* did not improve geraniol production, which can be attributed to the low geraniol production level of our starting strain. Other bottlenecks in the pathway such as precursor supply and GES expression level may obscure the effect of these GPP synthases with weaker activity. We also compared the performance of geraniol synthases from *C. roseus* (CrGES) and *Ocimum basilicum* (ObGES).[44,57,65] Plasmids containing CrGES or ObGES with their plastid targeting sequences removed were transformed into JHY651. ObGES showed a 2.4-fold increase in geraniol production compared to CrGES (Fig. 3.2C). Since GES-mutant ERG20 fusion constructs have been shown to improve geraniol production,[54,56,57,66] an ObGES-ERG20*K197G fusion protein was integrated in **S2** to construct strain **S3**. With two integrated copies of GES (CrGES and ObGES-ERG20*K197G), **S3** produced 1.26 mg/L geraniol (Fig. 3.3B).

3.3 Constructing an integrated platform yeast strain with cytosolic geraniol production

We performed a series of combinatorial genomic integrations on **S3** to incorporate additional strain modifications that increased yeast production of geraniol or improved metabolic flux to its precursors.

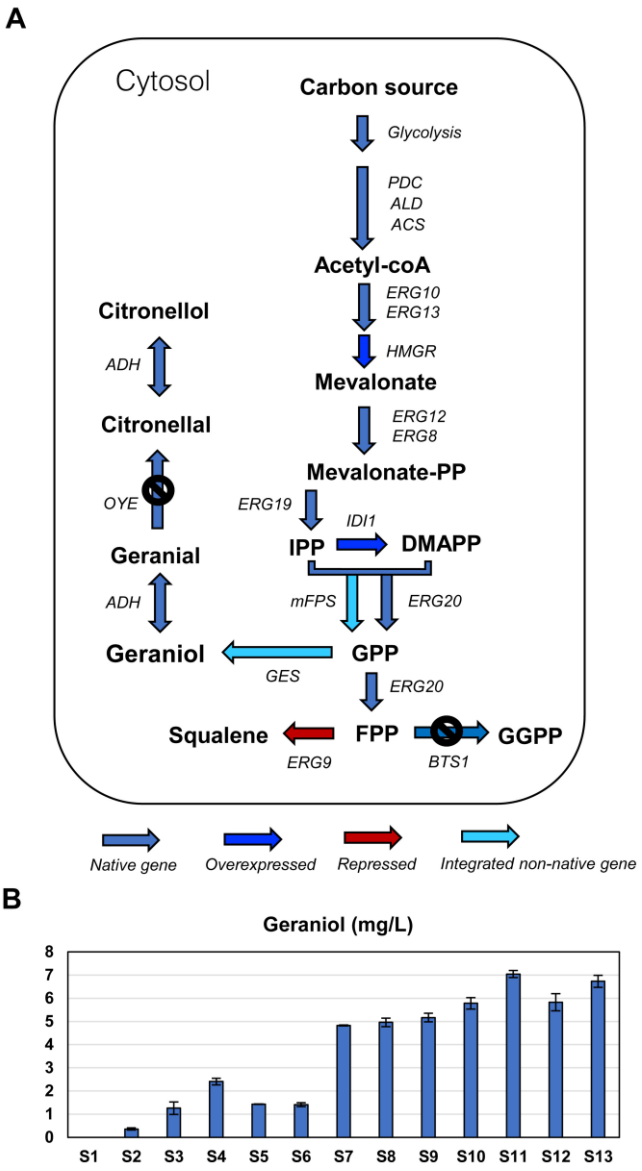


Fig. 3.3 Combinatorial strain modification to improve cytosolic geraniol titer. (**A**) Scheme of most strain modifications in cytosolic engineered strains **S4**-**S13**. The strategies shown include overexpression of HMGR and IDI1, truncation of the ERG9 promoter, introduction of mFPS and GES, and deletion of OYE2, OYE3, and BTS1. (**B**) Geraniol production in strains **S1**-**S13**.

For example, overexpression of IDI1 and truncated HGM1 (tHMG1) or HMG2 with a K6R stabilizing mutation (HMG2*) improved production of geraniol and other terpene-derived compounds.[53,54,67] Deletion of OYE2[68] decreased conversion of geraniol to the shunt product citronellol. Furthermore, deletion of ROX1, YPL062W, and YJL064W and truncation of the

ERG9 promoter led to increased production of mevalonate, a precursor to geraniol.[69] Using CRISPR-Cas9, we compiled these modifications in **S4**-**S13** to test their combined effect in a fully integrated and plasmid-free strain (Fig. 3.3A). Of all the strain variants, **S11** showed the highest geraniol production at 7 mg/L. A 5.6-fold increase in geraniol titer was observed from **S3** to **S11** (Fig. 3.3B). Overexpressing both tHMG1 and HMG2* in **S12** and integrating an additional copy of IDI1 in **S13** did not increase titer compared to **S11**. Furthermore, attempts to improve geraniol production by diverting flux through the mevalonate pathway were compromised by increased production of FPP. Overexpression of tHMG1 in **S10** and HMG2* in **S11** led to more than a 2-fold increase in farnesol production compared to **S7** (Fig. S1). FPP is converted to farnesol by yeast phosphatases LPP1 and DPP1.[70–72] The higher levels of farnesol in **S10** and **S11** indicate increased conversion of GPP to FPP, highlighting the need to protect the GPP pool from further chain extension.

3.4 Compartmentalizing the geraniol biosynthetic pathway to the mitochondria

To circumvent nonproductive consumption of GPP and other geraniol pathway intermediates by competing cytosolic pathways, we targeted the geraniol biosynthetic pathway to the mitochondria. We hypothesized mitochondrial compartmentalization would shelter the diphosphate intermediates since yeast has no known mitochondrial FPP synthase. In addition, the mitochondria have a rich supply of acetyl-CoA, the main precursor for the geraniol biosynthetic pathway. It is estimated that the concentration of acetyl-CoA is 20-30 times higher in the mitochondria compared to the cytosol.[73] Therefore, targeting geraniol biosynthesis to the mitochondria would leverage the mitochondrial acetyl-CoA pool and simultaneously protect GPP from conversion to FPP for increased geraniol production.

N-terminal cox4 tags were used to target all of the geraniol biosynthetic enzymes to the mitochondria as described previously.[74–76] All geraniol pathway genes were expressed under the bidirectional galactose-inducible promoter GAL1-10p to prevent stress from pathway expression until induction.[77,78] Four sequential integrations were performed to introduce a mitochondrial GPP biosynthetic pathway containing ERG13, ERG10, ERG12, tHMG1, ERG19, ERG8, IDI1, and *G. gallus* mFPS to the base strain **S1**. Three intermediate strains **S14**-**S16** were constructed along with the final strain **S17**, which has the complete pathway converting acetyl-CoA to GPP in the mitochondria (Fig. 3.4A). High copy plasmids with either mitochondrial (**M**)- or cytosolic (**C**)-targeted ObGES were transformed into strains **S1** and **S14**-**S17** to investigate the effect of these modifications on the GPP supply in the mitochondria and cytosol, respectively.

Geraniol production in the base strain and intermediate strains with the mitochondrial GES plasmid (**S1-M**, **S14-M** to **S16-M**) was more than 60% lower than that of the base strain with the cytosolic GES plasmid (**S1-C**) (Fig. 3.4B and 3.4C). This suggests *N*-cox4-GES is primarily targeted to the mitochondria since these strains lack a complete mitochondrial GPP pathway. **S1-C** showed higher geraniol production than the previous strains **S2** and **S3**, attributed to high copy episomal expression of cytosolic GES. An 11.5-fold increase in geraniol production was observed from **S16-M** (2.2 mg/L) to **S17-M** (25.2 mg/L) (Fig. 3.4B). The large titer increase upon completion of the mitochondrial GPP pathway indicates geraniol biosynthesis was successfully compartmentalized to the mitochondria in **S17-M**.

In mitochondria-engineered strains expressing cytosolic GES, unexpected increases in geraniol production were observed from **S14-C** to **S15-C** and **S16-C** to **S17-C** (Fig. 3.4C).

We hypothesized that the titer increases in **S15-C** and **S17-C** may be due to leakage of intermediate metabolites from the mitochondria to the cytosol or incomplete enzyme localization to the mitochondria, resulting in increased accumulation of geraniol in the cytosol. Previous western blot analysis showed the first 7 enzymes from ERG13 to IDI1 with *N*-terminal cox4 tags primarily localize to the mitochondria.[75] We performed a western blot to check the localization of mFPS with an *N*-terminal cox4 tag and showed mFPS is completely localized to the mitochondria (Fig. S2). These results suggest precursors to geraniol may be transported into the cytosol. Since mitochondrial ERG12 and tHMG1 are introduced in **S15-C** and mitochondrial IDI1 and mFPS are introduced in **S17-C**, we hypothesize that mevalonate, DMAPP, or GPP may cross through the mitochondrial membrane. It is unlikely that HMG-CoA can be transported due to the large size of the CoA group, although an unidentified mitochondrial transporter may facilitate its export. After exiting the mitochondria, these intermediates can be assimilated into the cytosolic GPP pathway, leading to increased geraniol production in the presence of cytosolic GES.

To construct fully integrated strains, mitochondrial ObGES was integrated into **S17** with simultaneous removal of the Cas9 expression cassette to construct **S18**. In parallel, cytosolic ObGES was integrated into **S17** to construct **S19** for comparison purposes. Integration of mitochondrial GES (**S18**) led to a 78% increase in geraniol titer (43.3 mg/L) compared to expression of mitochondrial GES on a high copy plasmid (**S17-M**) (Fig. 3.4D). Since all pathway genes are under GAL1-10p, high copy episomal expression of mitochondrial GES may lead to suboptimal expression levels of GPP pathway enzymes, which are only present as single copy integrations.

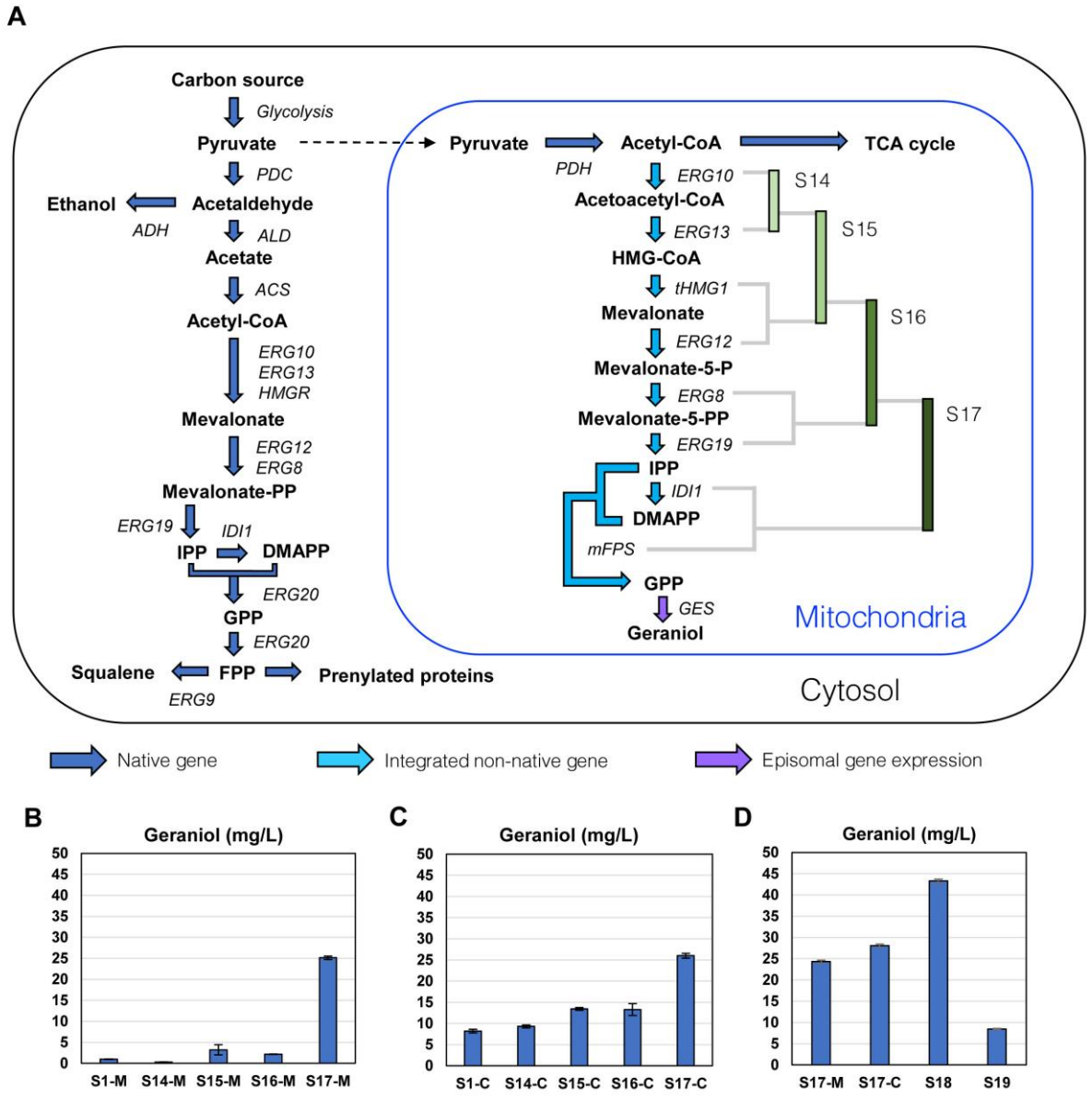


Fig. 3.4 Mitochondrial geraniol production. Plasmids transformed into each strain are designated as (p#). (**A**) Scheme of strain modifications in mitochondrial engineered strains **S14-S17** with episomal expression of mitochondrial GES (p10). These strains contain sequential integrations of mitochondrial ERG10 and ERG13 (**S14**), tHMG1 and ERG12 (**S15**), ERG8 and ERG19 (**S16**), and IDI and mFPS (**S17**). (**B**) Geraniol production in **S1** and **S14-S17** with episomal expression of mitochondrial ObGES (M, p10). (**C**) Geraniol production in **S1** and **S14-S17** with episomal expression of cytosolic ObGES (C, p11). (**D**) Geraniol production in **S17** with episomal expression of mitochondrial ObGES (p10), **S17** with episomal expression of cytosolic ObGES (p11), **S18** (mitochondrial ObGES integration) with an empty vector (p2) and **S19** (cytosolic ObGES integration) with an empty vector (p2).

On the other hand, high copy expression of mitochondrial GES may be detrimental to cell health with high amounts of GES being shuttled into the mitochondria. This was confirmed by growth curve measurements which showed **S17-M** had impaired growth compared to **S18** with control plasmids (Fig. S3, Table S1). In contrast to mitochondrial geraniol production, cytosolic geraniol production was limited by the expression level of cytosolic GES. Integrating a single copy of cytosolic GES led to a 63% decrease in geraniol titer (**S19**) compared to expression of cytosolic GES on a high copy plasmid (**S17-C**) (Fig. 3.4D). This indicates higher expression of GES in the cytosol is required to convert more of the cytosolic GPP into geraniol. Most importantly, the significant increase in geraniol titer from **S19** (cytosolic GES integration) to **S18** (mitochondrial GES integration) demonstrates targeting the GES to the mitochondria leads to greater capture of the mitochondrial GPP pool for geraniol production.

The relative farnesol levels of the fully integrated geraniol producing strains further suggest mitochondrial compartmentalization effectively protects the engineered GPP pool from competing cytosolic pathways. **S18** showed nearly a 50% decrease in farnesol production compared to **S19**, indicating reduced consumption of GPP by ERG20 (Fig. S4). Moreover, the geraniol titer of **S18** was 6-fold higher than that of our best performing cytosolic integrated strain (**S11**, 7 mg/L). This demonstrates mitochondrial compartmentalization is a more effective strategy compared to our previous efforts to optimize cytosolic geraniol production through conventional means.

3.5 Strain engineering for *de novo* production of 8-hydroxygeraniol

We then turned our attention towards production of 8-hydroxygeraniol, the next intermediate in the strictosidine pathway. Geraniol is converted to 8-hydroxygeraniol by the

membrane-bound cytochrome P450 G8H in plant MIA producers.[79] The ER-targeted form of G8H from *C. roseus* was integrated into our final mitochondrial geraniol producing strain **S18**. To decrease shunt product formation by yeast's native ene-reductases OYE2 and OYE3[61] (Fig. S5), G8H was integrated into the OYE2 or OYE3 locus for simultaneous deletion of these genes. For each locus, we integrated G8H using the native OYE promoter (**S20**, **S22**) and GAL1p (**S21**, **S23**). Due to evidence that OYE expression can be induced by geraniol,[61,80,81] we hypothesized timing of 8-hydroxygeraniol production may be optimized through use of substrate inducible promoters to drive on-demand G8H expression.

These constructs with single OYE deletions showed complete conversion of geraniol to 8-hydroxygeraniol with substantial accumulation of 8-hydroxy-2,3-dihydrogeraniol (Fig. S6), a by-product of OYE catalyzed reduction (with ADH). Consequently, we deleted the remaining OYE in the single OYE deletion strains, which abolished shunt product formation (**S24**-**S27**, Fig. S7). At 24 hours, **S25** and **S27** with GAL1p driving expression of G8H showed higher production of 8-hydroxygeraniol than **S24** and **S26** with the native OYE promoters, which demonstrates addition of GAL1p results in faster induction of G8H expression (Fig. 3.5A). 8-hydroxygeraniol production in **S24** with OYE2p driving G8H expression increased substantially from 24 to 48 hours, whereas **S25** and **S27** reached their peak 8-hydroxygeraniol titer at 24 hours. While using the native OYE2 promoter led to similar 8-hydroxygeraniol titers compared to GAL1p, these results suggest OYE2p can be used for other pathway genes in which pathway metabolite induced expression is required. 8-hydroxygeraniol production decreased from 24 to 48 hours in **S25**, most likely due to gradual oxidation of 8-hydroxygeraniol to 8-oxogeraniol by G8H.[82] Low production of 8-hydroxygeraniol in **S26**

indicates OYE3p is not strongly induced under these geraniol production levels. Out of all the double OYE knockout strains, **S25** (G8H under GAL1p in OYE2 locus) at 24 hours had the highest 8-hydroxygeraniol titer, producing 52 mg/L of 8-hydroxygeraniol. This is essentially quantitative conversion of geraniol to 8-hydroxygeraniol with a titer approximately 10-fold higher than the previous highest titer (5.3 mg/L).[63] This is also the first report of *de novo* 8-hydroxygeraniol production from a fully integrated and plasmid-free yeast strain.

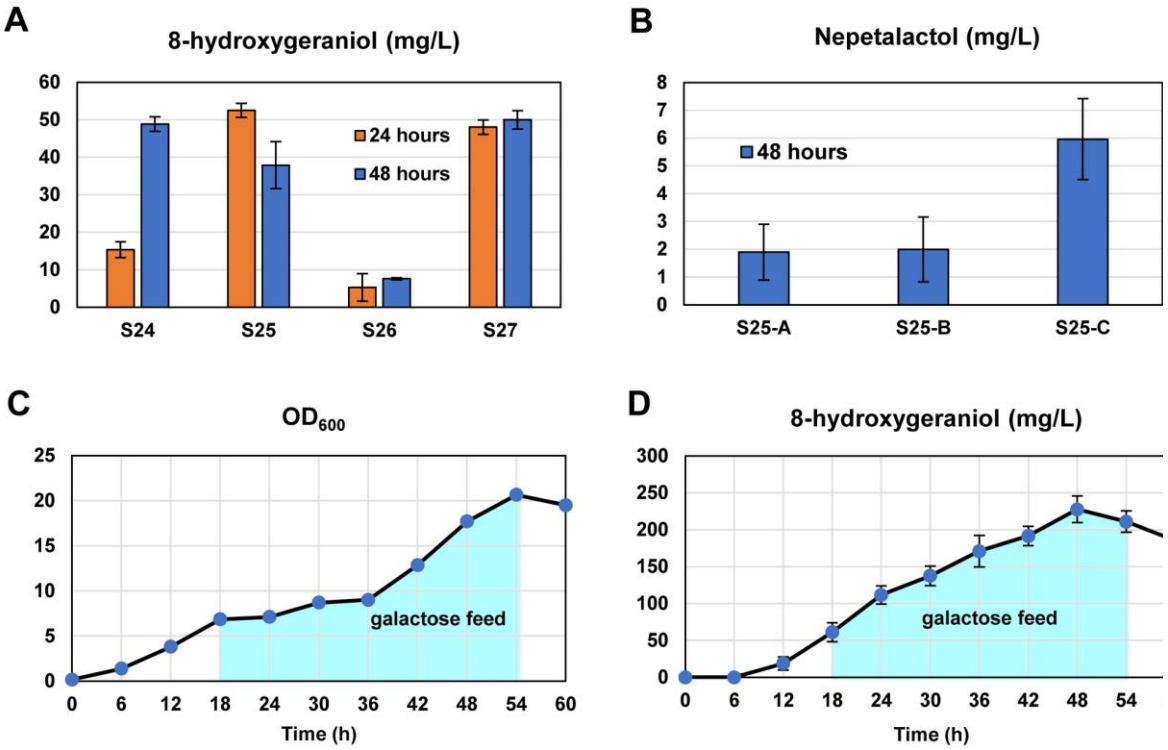


Fig. 3.5 *De novo* production of 8-hydroxygeraniol and nepetalactol. Plasmids transformed into each strain are designated as (p#). (**A**) 8-hydroxygeraniol production in strains **S24-S27** after 24 and 48 hours of culturing. (**B**) Nepetalactol production in **S25-G8H-GOR-ISY (S25-A)** with high copy G8H plasmid (p13) and low copy GOR/ISY plasmid (p12), **S25-B** with high copy G8H plasmid (p13) and high copy GOR/ISY plasmid (p16), and **S25-C** with high copy G8H/GOR plasmid (p17) and low copy ISY plasmid (p18) after 48 hours of culturing. (**C**) OD_{600} vs. time during fed-batch fermentation of **S25**. (**D**) 8-hydroxygeraniol titer vs. time during fed-batch fermentation of **S25**.

A fed-batch fermentation was performed to evaluate the scalability of our best performing 8-hydroxygeraniol producing strain. Upon inoculation of **S25** in YP media

containing glucose and galactose, the OD_{600} steadily rose throughout the initial glucose fermentation and subsequent galactose consumption phases, reaching a maximum of 20.7 at 54 hours (Fig. 3.5C). Galactose was fed continuously from 18 to 54 hours to maintain pathway induction. 8-hydroxygeraniol production continued to increase during the galactose feeding phase, reaching a maximum titer of 227 mg/L at 48 hours (Fig. 3.5D). A maximum productivity of 4.7 mg L^{-1} h^{-1} was observed after 36 hours of fermentation.

3.6 Strain engineering for *de novo* production of nepetalactol

We further extended the pathway to demonstrate *de novo* production of nepetalactol in **S25**. To convert 8-hydroxygeraniol to nepetalactol, a low copy plasmid with cytosolic GOR and ISY from *C. roseus* was transformed into **S25** to afford **S25-GOR-ISY**, but no detectable amount of nepetalactol was produced. Instead, geraniol and 8-hydroxygeraniol were fully converted to the shunt products citronellol and 8-hydroxytetrahydrogeraniol, indicating the interference of endogenous redox metabolism[61,63] (Fig. S5, S8, S10). To channel more flux towards geraniol hydroxylation instead of conversion to citronellol, we transformed a high copy plasmid with G8H alongside the GOR/ISY expression plasmid into **S25** (**S25-G8H-GOR-ISY**).

S25-G8H-GOR-ISY produced 2.0 mg/L of nepetalactol after 48 hours of culturing (Fig. 3.5B) while still accumulating a significant amount of 8-hydroxytetrahydrogeraniol. Citronellol production was nearly abolished, which indicates inclusion of additional G8H copies indeed directed more flux towards geraniol hydroxylation rather than reduction by ISY via geranial (Fig. S8). This is consistent with previous work, which showed higher G8H expression is required for *de novo* strictosidine production in yeast.[43] We further optimized expression levels of GOR and ISY by varying the plasmid copy number of each gene. Expressing G8H and

GOR on a high copy plasmid and ISY on a low copy plasmid resulted in 5.9 mg/L nepetalactol titer, which is 11% conversion from 8-hydroxygeraniol (Fig. 3.5B). We hypothesize more copies of G8H and GOR allow for faster oxidation of 8-hydroxygeraniol to 8-oxogeranial and fewer copies of ISY decrease flux through the ADH/ISY shunt pathway, leading to higher production of nepetalactol. Since conversion from 8-hydroxygeraniol to nepetalactol remains low, future work is required to further optimize nepetalactol production and decrease accumulation of off-pathway shunt products. Possible strategies include deletion of yeast ADHs,[61] additional tuning of G8H, GOR, and ISY expression levels, and optimization of culturing conditions. In addition, nepetalactol-related short-chain dehydrogenase enzymes (NEPS) from catmint (*Nepeta mussinii*) can be introduced to catalyze formation of stereospecific nepetalactol isomers.[83]

3.7 Conclusions

In this work, we achieved a significant increase in *de novo* geraniol production titer in *S. cerevisiae* by compartmentalizing the entire mevalonate pathway, GPPS and GES into the mitochondria in a fully genome integrated strain. The integration of G8H from *C. roseus* into the geraniol producing strain resulted in 8-hydroxygeraniol titers of 52 mg/L in small-scale cultures and 227 mg/L in a fed-batch fermentation, providing the first benchmark for scalable 8-hydroxygeraniol production. Addition of cytosolic GOR and ISY from *C. roseus* and tuning of relative expression levels resulted in 5.9 mg/L nepetalactol titer, which is the highest *de novo* nepetalactol titer reported to date. The strategies developed here can be used in future research to improve production of iridoids, strictosidine and MIAs in *S. cerevisiae*.

* Section 3 is part of my research article “Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*”.

4. ENGINEERING A BIOSENSOR FOR NEPETALACTOL DETECTION IN *S. CEREVISIAE*

4.1 RNA-seq analysis to identify genes induced by nepetalactol

To begin engineering a nepetalactol biosensor for high throughput strain engineering, our first aim was to identify specific nepetalactol-inducible promoters that have no response to the precursor 8-hydroxygeraniol. We performed RNA-seq analysis of yeast in response to nepetalactol and 8-hydroxygeraniol. The yeast strain YJB051[61] with an empty plasmid was inoculated in uracil dropout media and subcultured in YPD media. After 24 hours of growth in YPD, the cells were treated with nepetalactol, 8-hydroxygeraniol, or DMSO as a negative control. 6 biological replicates were performed for each substrate treatment. Total RNA was extracted from the strains 45 minutes after substrate addition, followed by preparation of barcoded cDNA libraries. The samples were sequenced by Novogene with 10 million reads per sample and differential gene expression analysis was performed.

A

gene name	log2Fold Change	adjusted p-value	total reads, nepetalactol	total reads, dmso
PDR5	7.01	0	26242	204
YLR346C	5.42	0	1791	42
RSB1	4.69	0	1779	70
YGR035C	4.66	0	1991	78
ICT1	4.60	0	4572	188

B

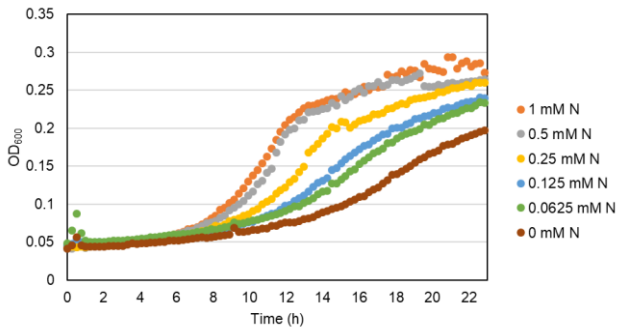


Figure 4.1 (A) The top 5 genes with the highest fold changes in expression upon addition of nepetalactol. (B) Growth of YJB051 with YLR346Cp HIS3 plasmid in histidine dropout media.

The RNA-seq results showed many genes were upregulated with the addition of nepetalactol but few genes had changes in expression levels upon addition of 8-hydroxygeraniol. Many of the genes upregulated by nepetalactol are involved in yeast's stress response. The top five genes with the highest fold changes in expression include PDR5 (multidrug transporter),[84] YLR346C (protein of unknown function, regulated by transcription factors involved in pleiotropic drug resistance),[85] RSB1 (putative sphingoid long-chain base efflux transporter),[86] YGR035C (putative protein of unknown function),[87] and ICT1 (lysophosphatidic acid acyltransferase, induced by organic solvent stress)[88] (Figure 4.1A). PDR5 showed a log2 fold increase of 7.0 and YLR346C showed a log2 fold increase of 5.4. The promoters of these genes are strong candidates for use in reporter assays with biosensor applications since their expression was significantly upregulated in the presence of nepetalactol but not 8-hydroxygeraniol, with p-values less than 0.001.

4.2 Growth and fluorescence assays using the nepetalactol-inducible YLR346C/PDR5 promoters

We then validated YLR346C's response to nepetalactol by performing a growth assay on YJB051 with a plasmid containing HIS3 under the YLR346C promoter in histidine dropout media. Increasing concentrations of nepetalactol led to improved growth (Figure 4.1B), further demonstrating YLR346C can be induced by nepetalactol. Despite this fact, we were unable to achieve similar results for strains with *de novo* production of nepetalactol using the YLR346Cp growth-based sensor. As a result, we switched to a fluorescence based assay using the PDR5 promoter. We integrated the fluorescence reporter mCherry into the PDR5 locus in S35 (S25 ΔOYE3::ADH2p_GOR;PCK1p_ISY;MLS1p_MLPL ΔCIS1::GAL2p_MLPL)

to construct a nepetalactol biosensor strain. Upon media optimization of glucose and galactose concentrations, the de novo nepetalactol producing strain S35 showed significantly higher fluorescence compared to the 8-hydroxygeraniol producing strain S25 with a single copy of mCherry in the PDR5 locus (Fig. 4.2A). This suggests fluorescence can function as a marker for nepetalactol production.

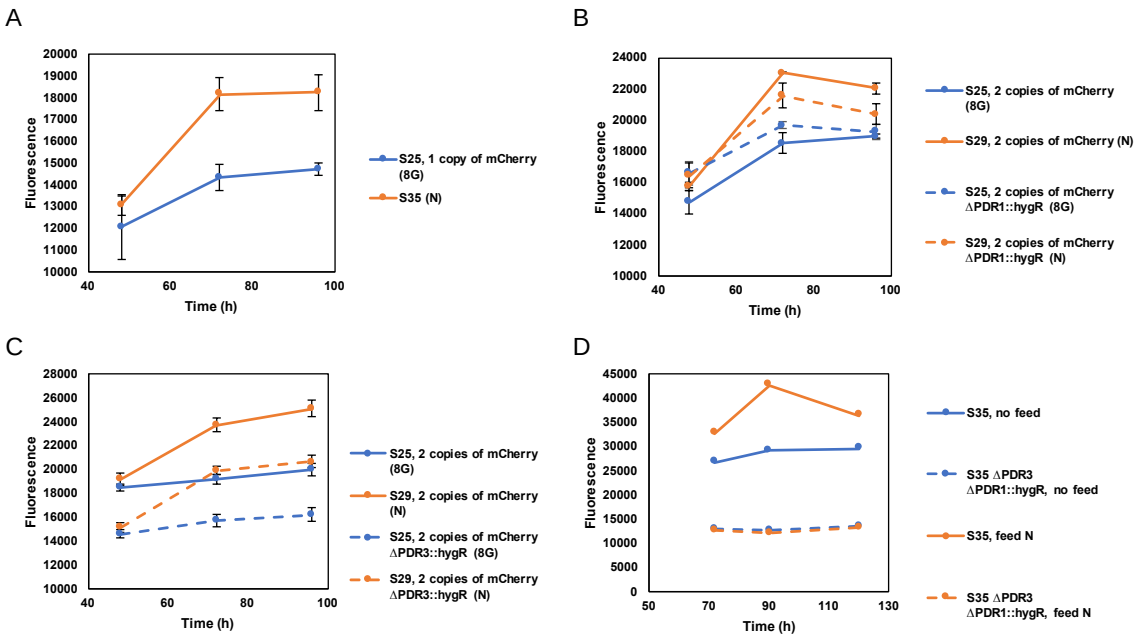


Figure 4.2 (A) Fluorescence of the de novo 8-hydroxygeraniol producer S25 with a single copy of mCherry in the PDR5 locus and the de novo nepetalactol producer S35 (ΔPDR5::mCherry) using optimized media conditions with YP 0.5% glucose 3.8% galactose. (B) Fluorescence of S25 with two copies of PDR5p_mCherry, the de novo nepetalactol producer S29 with two copies of PDR5p_mCherry, and parallel strains with knockout of PDR1. (C) Fluorescence of S25 with two copies of PDR5p_mCherry, S29 with two copies of PDR5p_mCherry, and parallel strains with knockout of PDR3. (D) Fluorescence of S35 and S35 ΔPDR3 ΔPDR1::hygR with no feeding and feeding 0.25 mM nepetalactol.

4.3 Identification of the activator responsible for PDR5 induction by nepetalactol

To identify additional directions for sensor optimization, we performed assays to elucidate the mechanism by which PDR5 is induced by nepetalactol. Deletion of the transcriptional activator PDR1[89] significantly decreased the fluorescence signal from

nepetalactol (Fig. 4.2B). To contrast, deletion of the paralog PDR3[89] in S35 decreased the background fluorescence but did not change the magnitude of the signal (Fig. 4.2C). This suggests PDR1 plays a greater role in the induction of PDR5 by nepetalactol, although PDR3 can rescue the response in a PDR1 knockout strain. Deletion of both PDR1 and PDR3 abolished the response to nepetalactol, demonstrating one copy of PDR1/PDR3 is required for PDR5 induction (Fig. 4.2D). These results indicate the potential for further engineering of PDR1/PDR3 or their DNA binding domains in the PDR5 promoter by mutagenesis to increase the biosensor's sensitivity to nepetalactol.

5. ENGINEERING *S. CEREVISIAE* FOR *DE NOVO* PRODUCTION OF NEPETALACTOL

5.1 CRISPRi screening

Recently, O'Connor et al. discovered a major latex protein (MLPL) from *Nepeta mussinii* that facilitates cyclization of 8-oxogeranial (after reduction by ISY) to form nepetalactol.[90] We tested the effect of MLPL on *de novo* nepetalactol production in yeast by integrating GOR, ISY, and MLPL under ADH2-like promoters in the OYE3 locus of the 8-hydroxygeraniol producing strain S25. The resulting strain S35 produced 18 mg/L nepetalactol, a 3-fold increase from the previous titer without MLPL. Plasmid based CRISPRi and CRISPRa screening was performed to identify mitochondrial gene targets leading to higher monoterpene/iridoid production upon repression or activation. The CRISPRi and CRISPRa plasmids were developed and cloned by Stanford Genome Technology Center.

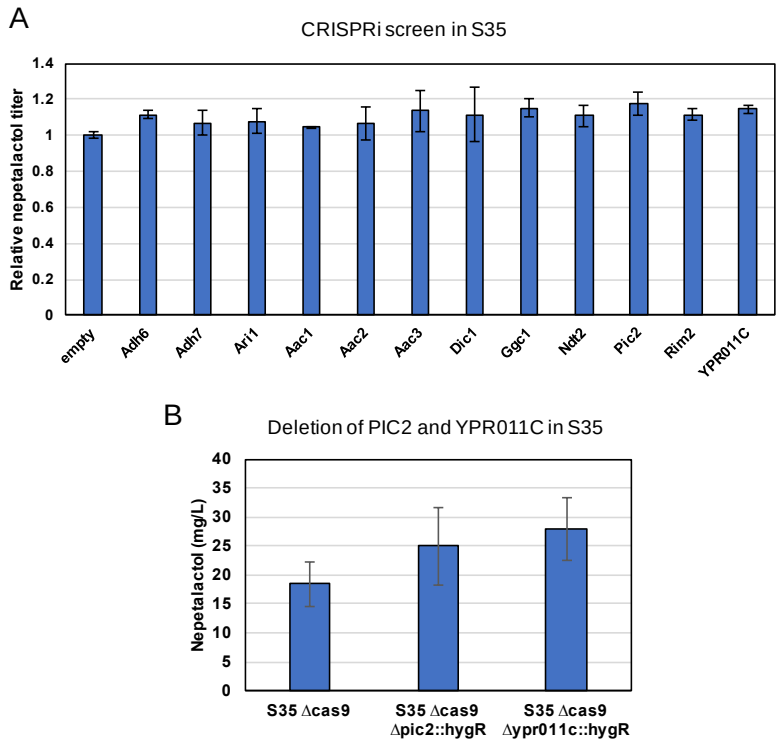


Figure 5.1 (A) Nepetalactol production from the plasmid based CRISPRi screen in strain S35. (B) Nepetalactol production in the Pic2 and YPR011C deletion strains.

The CRISPRi targets included mitochondrial transporters that may shuttle diphosphate intermediates out the mitochondria and alcohol dehydrogenases (ADHs) previously shown to decrease shunt product formation as a control (Billingsley 2017). CRISPRi repression of yeast ADHs increased nepetalactol production, as expected. Repression of the mitochondrial transporters Pic2 (copper/phosphate)[91] and YPR011C (APS - adenosine 5'-phosphosulfate)[92] led to a 14% and 17% increase in nepetalactol titer compared to the empty vector control (Figure 5.1A). qRT-PCR analysis showed YPR011C and Pic2 were repressed 4 and 5-fold by CRISPRi. Following this, we performed individual deletions of Pic2 (S36) and YPR011C (S37), which led to a 35% and 50% increase in nepetalactol production, respectively (Figure 5.1B).

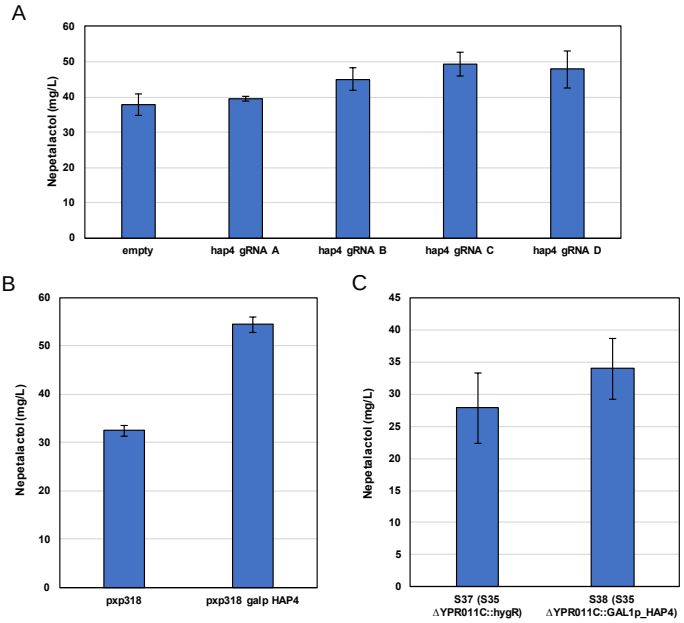


Figure 5.2 (A) Nepetalactol production from the plasmid based CRISPRa screen to activate Hap4 in strain S37. (B) Nepetalactol production in strains with Hap4 overexpression under GAL1p. (C) Nepetalactol production with integration of the Hap4 overexpression cassette.

5.2 CRISPRa screening

The strain S37 was subsequently used to perform CRISPRa screening. We tested the effect of activating the transcription factor Hap4, which was shown to increase cell respiration and mitochondrial biogenesis.[93] After testing 4 guide RNAs, we observed guide RNA C resulted in a 30% increase in nepetalactol production (Figure 5.2A). To verify the effect of Hap4 overexpression, Hap4 was expressed under the GAL1 promoter on a low copy plasmid in S37, leading to a 67% increase in nepetalactol titer (Figure 5.2B). The overexpression cassette for Hap4 was then integrated into the YPR011C locus in S35, which led to a 22% increase in nepetalactol titer compared to the parent strain (S38) (Figure 5.2C).

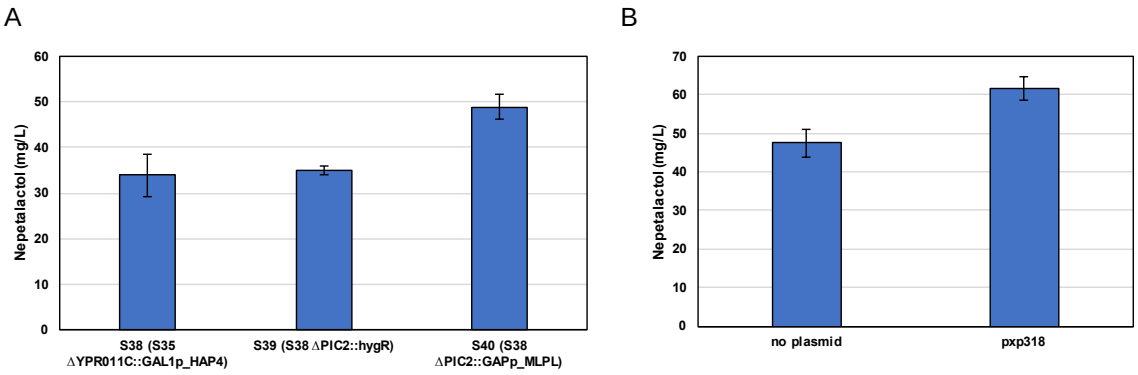


Figure 5.3 (A) Nepetalactol production in strains S38-S40 (B) Nepetalactol production in S40 without and with an empty URA3 plasmid.

Following this, Pic2 was deleted in this strain (S39), but unexpectedly, no increase in nepetalactol titer was observed (Figure 5.3A). An additional copy of MLPL under the constitutive GAP promoter was integrated into the Pic2 locus (S40), which led to a 40% increase in nepetalactol production (Figure 5.3A). Throughout these assays, we noticed strains containing plasmids with the URA3 marker produced higher levels of nepetalactol. To test this directly, we compared nepetalactol production in S40 with no plasmid and with an

empty low copy plasmid containing URA3. The plasmid containing strain showed a 30% increase in nepetalactol titer (62 mg/L) (Figure 5.3B). This is a 10.5-fold increase in nepetalactol production compared to the un-optimized strain.

6. GENOME MINING OF ALKALOIDAL TERPENOIDS FROM A HYBRID TERPENE AND NONRIBOSOMAL PEPTIDE BIOSYNTHETIC PATHWAY

Biosynthetic pathways containing multiple core enzymes have potential to produce structurally complex natural products. Here we mined a fungal gene cluster that contains two predicted terpene cyclases (TCs) and a nonribosomal peptide synthetase (NRPS). We showed the *flv* pathway produces flavunoidine **1**, an alkaloidal terpenoid. The core of **1** is a tetracyclic, cage-like and oxygenated sesquiterpene that is connected to dimethylcadaverine via a C-N bond, and is acylated with 5,5-dimethyl-l-pipecolate. The roles of all *flv* enzymes are established based on metabolite analysis from heterologous expression.

6.1 Mining for a TC-NRPS hybrid BGC

Structural complexities of natural products (NPs) are generated by enzymes in the biosynthetic pathways.[94] Scaffolds assembled by core enzymes such as polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPSs) or terpene cyclase (TCs), etc., can be morphed into complex NPs by accessory enzymes including transferases[95,96] and oxidoreductases,[97–99] etc. In fungi, the combinations of different core enzymes in the same biosynthetic pathway, such as PKS/PKS,[100,101] PKS/NRPS,[102] PKS/TC[103–106] (Figure 6.1A), can result in complex hybrid NPs unachievable with a single core enzyme. In contrast, biosynthetic pathways containing the combination of NRPS and TC have not been well-studied. While many metabolites are derived from prenylation of peptidyl cores via prenyltransferases,[107–109] the use of a TC to generate a terpene core that is decorated by a NRPS is rare. However, bioinformatic scanning of sequenced fungal genomes suggests

TC/NRPS hybrid clusters are common (Figure S1). Isolation of fungal aminoacylated terpenoids also suggests such hybrid molecules can be synthesized by fungi (Figure 6.1B).[110,111] Recent characterization of the aculene A biosynthetic pathway demonstrates such collaboration between the TC and a single module NRPS.[112] Based on these evidences, we believe there is significant potential in mining fungal TC/NRPS pathways for new NPs.

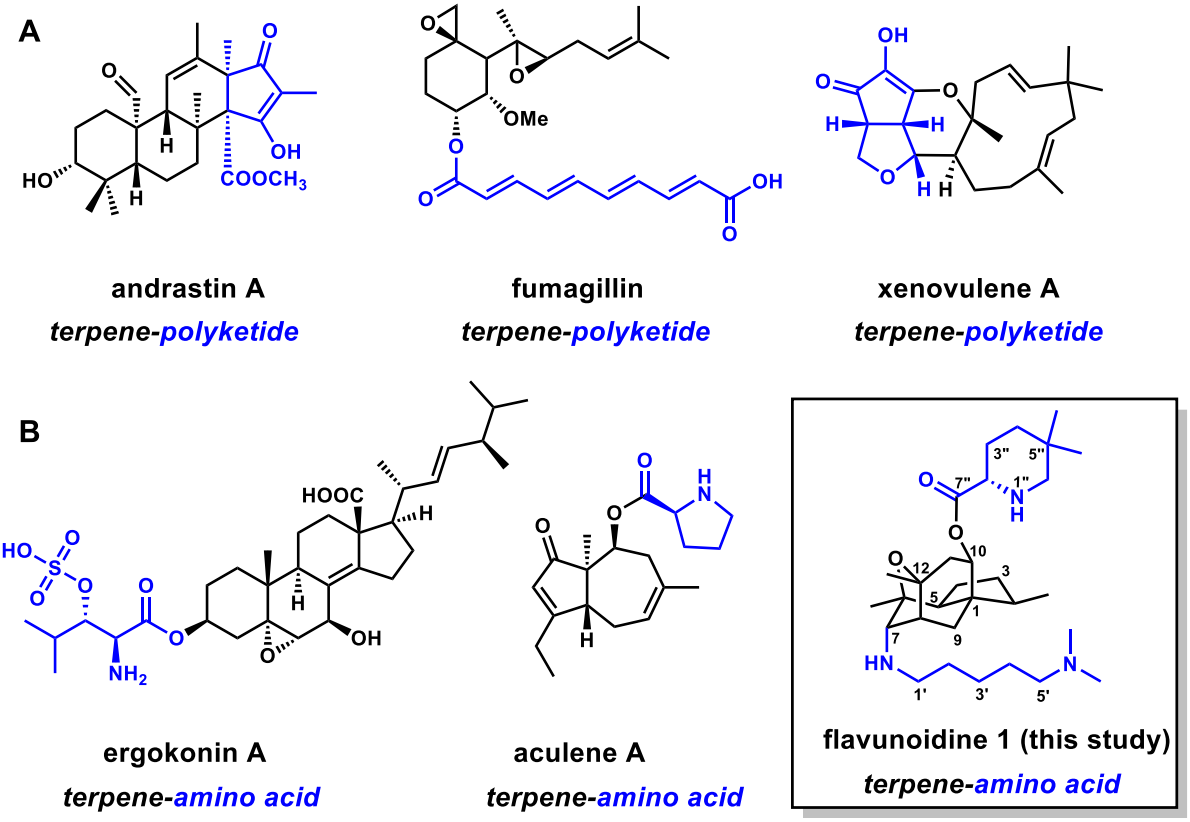


Figure 6.1 Structures of fungal polyketides synthesized by collaborative efforts of core enzymes. (**A**) meroterpenoids derived from TC/PKS,[6b-d] etc.; (**B**) compounds derived from TC and NRPS enzymes include ergokonin A[8a] (proposed), aculene A,[9] and flavunoidine **1** (this study).

Among the predicted gene clusters containing both TC and NRPS, we selected an uncharacterized nine-gene cluster conserved in several well-studied fungal species (Figure 6.2A, Figure S2, Table S1). The *flv* cluster in *Aspergillus flavus* encodes two TCs: FlvE and

FlvF, with sequence similarity to trichodiene synthase[113] and ophiobolin synthase,[114] respectively; and a single-module NRPS FlvI. This cluster also contains redox enzymes, including FlvB (short chain reductase, SDR), FlvC and FlvD (P450s), and a didomain enzyme FlvA. The *N*-terminal half of FlvA is predicted to be a PLP-dependent lyase,[115–117] while the *C*-terminal half is predicted to be a non-heme Fe, α-ketoglutarate dependent oxygenase (α-KG) (Figures S3-S4).[118–120] The remaining enzymes encoded in the gene cluster include FlvG, which is a homolog of ornithine decarboxylase; and FlvH, which has sequence similarity to histone lysine *N*-methyltransferase (*N*-MT).[121] To investigate the metabolite that can be biosynthesized from the *flv* cluster, we used *Aspergillus nidulans* as a heterologous host to mine the pathway.

6.2 Results from heterologous expression and *in vitro* assays

When the entire *flv* gene cluster was introduced into *A. nidulans*, we detected and isolated a new metabolite **1** with molecular weight (MW) of 503 (Figure 6.2B, i) (1.2 mg/L). The structure of **1** was solved to be flavunoidine shown in Figure 6.1B (Table S5 and Figures S5, S17-22). The tetracyclic cage was only previously found in the phospholipase C inhibitor hispidospermidin **4** (Figure 6.2C).[122,123] The C7 axial trimethyl-spermidine substituent in **4** is replaced by *N,N*-dimethylcadaverine in **1**, while the C10 position is hydroxylated and acylated with 5,5-dimethyl-L-pipecolate. **4** was the only NP with the same tetracyclic core as in **1**, and has been the subject of total synthesis by Danishefsky,[124,125] Overman[126] and Sorenson,[127,128] etc. **1** does not display notable cytotoxicity, nor is it antifungal or antibacterial. No production of **1** can be found from *A. flavus* grown under typical laboratory conditions (Figure S6).

We first investigated formation of the dimethylpipecolate in **1**. To probe the role of the NRPS FlvI, we expressed the other eight genes FlvA-H. This led to the absence of **1** but the emergence of **2** (MW 157) (0.7 mg/L) and **3** (MW 364) (0.9 mg/L) (Figure 6.2B, ii). NMR analysis showed that **2** is 5,5-dimethyl-L-pipecolic acid (Table S6 and Figures S23-27), while **3** is the unacylated precursor of **1** (Table S7 and Figures S28-33). The accumulation of these separate building blocks suggests the NRPS FlvI is responsible for esterifying **2** and **3** (Figure 6.2D). Indeed, feeding of **2** and **3** to *A. nidulans* expressing FlvI led to the biotransformation to **1** (Figure 6.2B, iii).

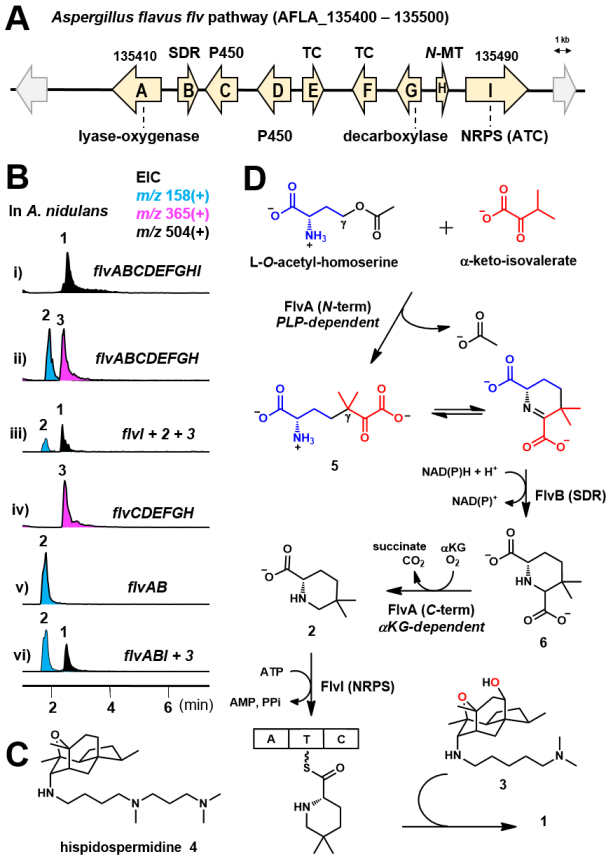


Figure 6.2 Heterologous expression of *flv* pathway. (**A**) The *flv* gene cluster. (**B**) LC/MS analysis of extracts from *A. nidulans* expressing different combinations of *flv* genes. (**C**) The structure of hispidospermidin **4**. (**D**) Proposed biosynthesis of **1** and **2** (For details see Figure S7).

Pipecolate biosynthesis from lysine has been shown to involve a PLP-dependent enzyme and reductase,[129] which led us to propose that the didomain enzyme FlvA and SDR FlvB may be involved in biosynthesis of **2.** When FlvC-H were expressed, we only observed accumulation of **3** (Figure 6.2B, iv), while coexpression of FlvA and FlvB separately resulted in formation of **2** (Figure 6.2B, v). Individual expression of either FlvA or FlvB did not result in formation of **2**. In addition, coexpression of FlvA, FlvB and FlvI, accompanied by feeding of **3**, led to the production of **1** (Figure 6.2B, vi). These results implicate FlvA and FlvB in the biosynthesis of **2.** We propose the PLP-dependent lyase[115–117] domain of FlvA catalyzes a γ-replacement reaction as shown in Figure 6.2D (detailed mechanism shown in Figure S7). L-*O*-acetyl-homoserine can bind and form the aldimine, which can undergo two proton abstraction steps to eliminate acetate and form the vinyl glycine quinonoid. This species can be attacked by α-keto-isovalerate to form a new C-C bond, which upon protonation and transaldimination can lead to release of the ketone **5** that can cyclize intramolecularly to yield the imine. The imine can be reduced by FlvB to yield the 6-carboxylated pipecolate **6**. The *C*-terminal α-KG dependent oxygenase domain of FlvA is then proposed to catalyze the decarboxylation of **6** to yield **2,**[130–133] with a possible mechanism shown in Figure S7.

We next investigated formation of the tetracyclic core and the unusual *trans* diaxial nitrogen and oxygen functionality in **3**. We have established that coexpression of six enzymes FlvC-H can synthesize **3** (Figure 6.2A, iv). To examine the roles of the TCs (FlvE and FlvF), we removed either gene and analyzed the resulting metabolic profiles. Removing *flvE* abolished all related metabolites (Figure 6.3A, iii), suggesting its involvement in core synthesis; while removing *flvF* led to accumulation of a pair of metabolites **7a** and **7b** with the

same MW (Figure 6.3, ii). **7a** was purified and structurally determined (Figure 6.3C, Tables S8, Figures S34-39) to contain the same core as **3,** but substituted at C7 by ethanolamine via an axial C-N bond. Based on data below for the stereoisomer pair **9a** and **9b**, we propose **7b** is the C7 equatorial stereoisomer of **7a.** Formation of both axial and equatorial isomers indicates the C-N bonds formed in **7a** and **7b** may be uncatalyzed. This also hints FlvF may be responsible for stereoselectively forming the C-N bond in **3.**

To analyze the function of core TC FlvE, we expressed the enzyme in *Saccharomyces cerevisiae* JHY651.[35] GCMS analysis revealed a sesquiterpene product **8** (Figure S8). Purified FlvE also synthesized **8** using farnesyl diphosphate (FPP) (Figures S9 and S10). Isolation and characterization confirmed **8** to be (*1R,4R,5S*)-(+)-acoradiene.[134] The (-)-enantiomer was previously isolated from plants.[135] We then coexpressed the P450s, FlvD and FlvC, with FlvE to determine if oxidative modifications of **8** can generate the core in **3**. Coexpression of both P450s with FlvE in *A. nidulans* resulted in formation of **7a** and **7b** (Figure 6.3A, v), while coexpression of only FlvD and FlvE led to the C7-stereoisomers **9a** (axial) and **9b** (equatorial) (Figure 6.3a, iv) (Table S10 and Figures S42-53), which do not contain the C10 hydroxyl. Coexpression of FlvD-H without FlvC in *A. nidulans* resulted in the formation of **12** (Figure 6.3a, vi), which is the C10- deshydroxy version of **3**. Collectively, these results implicate FlvC as the C10 hydroxylase (Figure 6.3C), while FlvD is solely responsible for oxidative conversion of **8** into the tetracyclic cage present in the final product **1**.

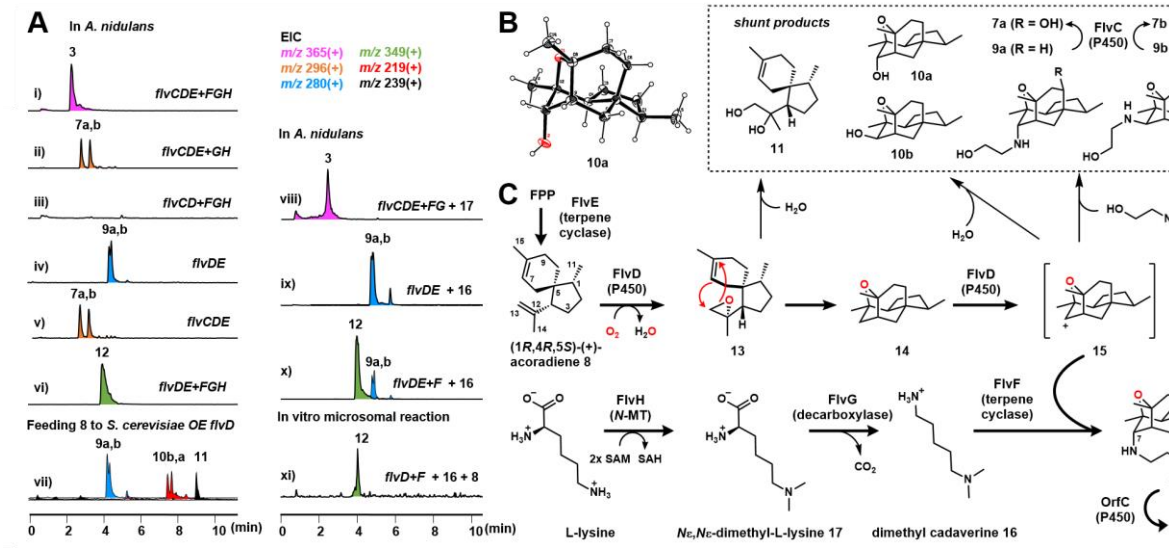


Figure 6.3 Biosynthesis of the core of **1**. (**A**) LC-MS analysis of metabolites *in vivo* and *in vitro* assays; (**B**) Crystal structure of **10a**; (**C**) Proposed biosynthetic pathway of **3**. Observed shunt products are shown in the dashed box.

Feeding **8** to *A. nidulans* expressing FlvD-H, but without the TC FlvE, restored the otherwise abolished production of **12** (Figure S11), confirming **8** is a precursor in the pathway. To analyze the function of FlvD in morphing **8**, we overexpressed the enzyme in yeast and fed **8** to analyze biotransformation products. In addition to **9a** and **9b**, new metabolites **10a, 10b** and **11** were detected (Figures 6.3a, vii). **11** retains the carbon scaffold in **8**, but with the C12, C13 diol (Table S12, Figures S60-65). **10a** is substituted with an axial hydroxyl group at C7 (Figure 6.3C) (Table S11 and Figure S54-59). We obtained an X-ray crystal structure of **10a** (Figure 6.3B), which confirmed the tetracyclic structure and allowed assignment of absolute stereochemistry of compounds. Based on the structures of **9a** and **9b**, we propose **10b** is the equatorial stereoisomer at C7. Purified microsomes from yeast expressing FlvD converted **8** to **10a** and **10b** (Figure S12). No **9a** or **9b** was detected since no ethanolamine was present in the reaction. Performing the microsomal assay in $H_2{}^{18}O$ led to incorporation of labeled ^{18}O

into **10a** and **10b**, indicating the C7 hydroxyl groups are derived from water (Figure S13). We verified **9** and **10** are shunt products, as feeding these compounds to *A. nidulans* expressing FlvD-H without FlvE did not restore biosynthesis of **12** (Figure S11)**.**

We propose **11** is derived from the nonenzymatic epoxide opening of **13**, which can be formed stereoselectively from **8** by FlvD. **13** can undergo intramolecular [3+2]-cycloaddition between the olefin and epoxide to directly forge the tetracyclic core **14.** This reaction may be assisted by an active site Lewis acid in FlvD and proceed in a step-wise mechanism.[136] This represents a very concise way to morph the terpene **8** into the caged core. The intermediate **14**, which was not isolated *in vivo* nor *in vitro*, may be further oxidized at C7 by FlvD to yield a carbocation **15**. **15** can then be quenched by nucleophiles such as water to yield **10a** and **10b**. Ethanolamine, which is biosynthesized by both yeast and *A. nidulans*,[137–139] may enter the active site of FlvD and quench **15** to yield **9a** and **9b**.

When the TC homolog FlvF is coexpressed, the nonenzymatic quenching is suppressed and dimethylcadaverine **16** can stereoselectively quench **15** to afford **12** (Figure 6.3A, vi). To test this, we coexpressed FlvD and FlvE and fed the strain with **16.** This strain still synthesized only **9a** and **9b,** and no **12** was formed (Figure 6.3A, ix). Upon coexpression of FlvF with FlvD and FlvE, feeding of **16** led to formation of **12** (Figure 6.3A, x). We then performed an *in vitro* assay using yeast microsomes containing FlvD and purified FlvF (Figure S10) in the presence of **8** and **16**. This reaction produced **12** and no shunt products were detected (Figure 6.3A, xi). Excluding FlvF from this assay led to formation of only **10a** and **10b**, even in the presence of **16** (Figure S14). Directly adding **10** and **16** to FlvD and FlvF did

not lead to formation of **12**, confirming **10** is a shunt product from nonenzymatic quenching. The mechanism by which FlvF can enable the stereoselective C-N bond formation in **12** is unexpected for an enzyme annotated as TC. Moore reported an algal TC that can catalyze *N*-geranylation of L-Glu.[140] However, the mechanism here is different since the terpene substrate is not pyrophosphorylated. We propose FlvF may form a complex with FlvD and deliver **16** to the active site where **15** is generated. The mechanism of this reaction is under investigation.

The two remaining enzymes in the pathway, FlvH (*N*-MT) and FlvG (decarboxylase) are proposed to synthesize **16** from L-lysine in a two-step reaction, in which FlvH performs methylation to give **17**, and is decarboxylated by FlvG to afford **16** (Figure 6.3C)**.** When *flvH* was removed from *A. nidulans* that produces **3**, trace amounts of **3** were formed (Figure S15)**.** The titer of **3** can be restored upon feeding of **17** (Figure 6.3A, viii)**.** When **17** was fed to the same strain without *flvG*, biosynthesis of **3** was abolished and only **7a** and **7b** were observed. Feeding of **16** restored production of **3**, establishing FlvG catalyzes decarboxylation of **17** to **16** (Figure S16).

6.3 Conclusions

In summary, we mined a fungal biosynthetic gene cluster that contained both TC and NRPS core enzymes, and discovered a new alkaloidal terpenoid **1**. The tetracyclic core of **1** is synthesized by the TC FlvE and P450 FlvD, while a second TC FlvF is required for attachment of the C7 axial dimethylcadaverine. The NRPS acylates the terpenoid core with dimethylpipecolate. The unexpected structural features of **1** highlight potential of fungal genome mining using combinations of core biosynthetic enzymes as a criterion.

* Section 6 is part of my research article “Genome Mining of Alkaloidal Terpenoids from a Hybrid Terpene and Nonribosomal Peptide Biosynthetic Pathway”.

7. DISCOVERY OF THE ARGININE CYCLODIPEPTIDE SYNTHASE ENZYME FAMILY

7.1 Background on genes of unknown function

Natural products continue to be an immense source of inspiration for bioactive molecules with therapeutic and industrial applications[141]. The majority of natural product scaffolds are synthesized by core enzymes such as polyketide synthases (PKSs)[6], nonribosomal peptide synthetases (NRPSs)[11], terpene cyclases (TCs)[12], or ribosomally synthesized and post-translationally modified peptides (RIPPS)[142], although new scaffolds have been recently discovered.[143] However, we hypothesized new natural product scaffolds could be unearthed through the discovery of additional core enzymes that have not been characterized before. Genome mining results often include clusters of tailoring genes without a conventional core gene in close proximity. Furthermore, genes of unknown function are frequently co-localized with these tailoring genes. We hypothesized such hypothetical proteins may produce novel scaffolds that are subsequently decorated by the neighboring tailoring enzymes. Study of these clusters may lead to the discovery of new scaffolds that are difficult to predict as well as novel core biosynthetic enzymes.

Here we analyzed the *ank* biosynthetic gene cluster (BGC) from the filamentous fungus *Aspergillus thermomutatus* that contains numerous biosynthetic tailoring enzymes co-localized with a hypothetical protein (HP) but no canonical core genes. Study of the HP AnkA and homologous genes from other fungi resulted in the discovery of a new family of fungal core enzymes, arginine cyclodipeptide synthases (aCDPS), that produce a diketopiperazine (DKP) scaffold through cyclization of arginine with other amino acids. Moreover, probing of an additional BGC from *Aspergillus versicolor* dI-29 with the aCDPS AvaA clustered with

tailoring genes resulted in the characterization of new cyclo-Arg-Trp analogs with novel structural motifs and the discovery of unique tailoring enzymes that catalyze new reactions. This class of biosynthetic core genes can be used to mine for other biosynthetic gene clusters that produce new cyclodipeptides and further derivatives.

7.2 Study of the *A. thermomutatus* cluster

We decided to initiate analysis of the *ava* BGC from *A. thermomutatus* because the cluster contains a diverse range of tailoring enzymes and potential biosynthetic enzymes but lacks canonical core genes (Figure 7.1a). The BGC includes a cytochrome P450 (a*nkB*), a FAD-dependent enzyme (a*nkC*), a PLP-dependent enzyme (a*nkD*), a NRPS-independent siderophore (NIS) synthetase (a*nkE*), and an O-methyltransferase (OMT) (a*nkF*). (Table S1) These tailoring genes flank a hypothetical protein with unknown function, a*nkA*. Bioinformatics analysis with NCBI Blast and conserved domain searches using the NCBI and Pfam databases did not yield strong predictions for the function of a*nkA*. However, we hypothesized a*nkA* may play a role in the biosynthesis of the BGC's product due to its proximity to the remaining tailoring genes. Moreover, the inclusion of a*nkA* is conserved in several homologous clusters from other fungal strains (Figure S1), suggesting AnkA's involvement in the pathway.

To test this hypothesis and identify potential metabolites produced by the cluster, the genes a*nkA-F* were expressed in the heterologous host *Aspergillus nidulans* A1145 ΔEMΔST.[34] This resulted in accumulation of the major product **1** with molecular weight (MW) of 622 (Figure 7.1b, 7.1c) as well as other new compounds **2**, **3**, **4**, and **7** with MW of 580, 319, 317, and 608, respectively. **1** was solved by NMR analysis to be an analog of the p300

HAT inhibitor NK13650B[144] with an additional methyl group on the C15 hydroxyl group. **1** is comprised of numerous building blocks from primary metabolism: a cyclo-Arg-Tyr core, a homoserine group, and a citric acid moiety (Figure 7.1d). The minor product **2** was co-purified with **1**, and NMR analysis showed **2** is a derivative of **1** with arginine replaced by ornithine (Figure 7.1d).

We hypothesized *A. nidulans*' native arginase from the urea cycle may hydrolyze the guanidine group of **1** to afford **2**.[145] **3** was solved by NMR analysis to be the undecorated cyclodipeptide scaffold containing arginine and tyrosine (Figure 7.1d). We reasoned the biosynthesis of **1** may begin with the formation of **3**. Upon testing different combinations of cluster genes in *A. nidulans*, expression of a*nkA* alone resulted in production of **3.** (Figure 7.1c) Expression of *ankA* in *S. cerevisiae* JHY651[35] led to production of **3** as well (Figure S7). Since a*nkA* contains no homology to bacterial cyclodipeptide synthases (CDPSs)[146] or the canonical adenylation (A), thiolation (T), or condensation (C) domains found in nonribosomal peptide synthetases (NRPSs), it was unclear how AnkA synthesizes the cyclodipeptide core. However, we decided to finish elucidation of the a*nk* pathway first and investigate the mechanism of AnkA and potential homologs at a later point.

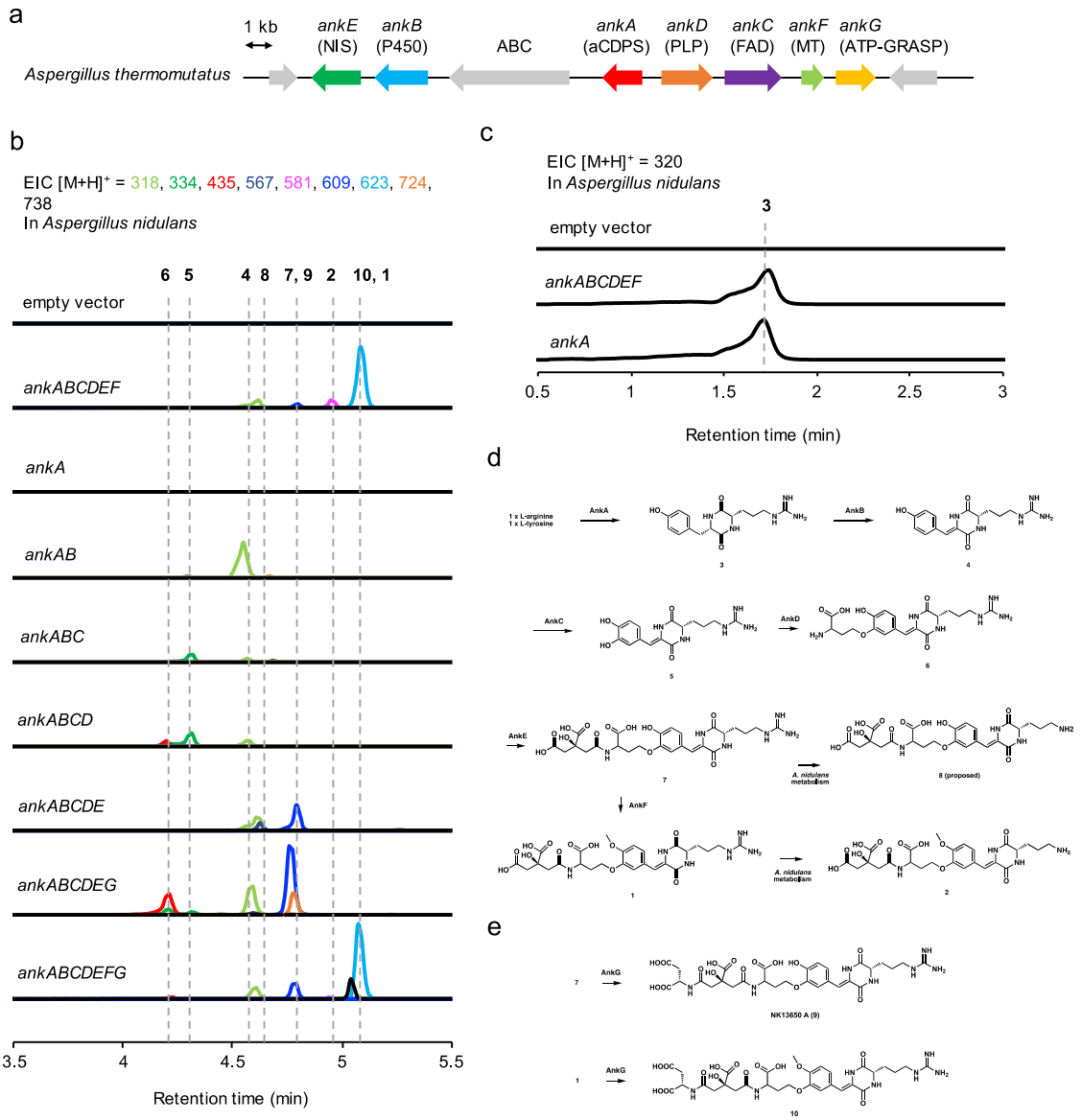


Fig. 7.1 Elucidation of the *ank* pathway from *A. thermomutatus*. A) The *ank* BGC containing the hypothetical protein *ankA* flanked by tailoring genes. B) QTOF analysis of extracts from expression of *ank* pathway genes in *A. nidulans*, retention time 3.5-5.5 min. C) QTOF analysis of extracts from expression of *ank* pathway genes in *A. nidulans*, retention time 0.5-3.0 min. D) The biosynthetic pathway of compound 1, starting from arginine and tyrosine. E) The proposed reaction by AnkG to convert **7** and **1** to **9** and **10**, respectively.

Coexpression of a*nkA* and the P450 a*nkB* resulted in production of compound **4** (Figure 7.1b). The 1D NMR spectra of **4** matched that of cyclo-(L-Arg-dehydrotyrosine)[147],

which suggests AnkB desaturates the tyrosine group of **2** (Figure 7.1d). Addition of the FAD-dependent enzyme a*nkC* (41% identity to the phenol hydroxylase Hkm7[148], Table S1) led to formation of **5**, which was solved to be a derivative of **4** with an additional hydroxyl group at C14 (Figure 7.1b, 7.1d). a*nkD* has 50% identity to the PLP-dependent enzymes CndF and FlvA (Table S1), which catalyze C-C bond formation via γ-elimination of O-acetyl-L-homoserine.[149, 20] We hypothesized AnkD may function through a similar mechanism to install the homoserine moiety onto **5** (Figure S5). Coexpression of a*nkD* with a*nkA-C* in *A. nidulans* resulted in production of a new compound **6**, which was structurally determined to contain homoserine linked to the C14 ether of **5** (Figure 7.1d). In addition, coexpression of a*nkD* with a*nkAB* in *A. nidulans* resulted in trace amounts of a new compound **4b**. Although the structure of **4b** was not confirmed by NMR, we propose **4b** may contain homoserine linked to C15 based on its mass, which suggests AnkD can also attach homoserine to **4**. (Figure S8) Purified AnkD from *E. coli* converted **4** to **4b** and **5** to **6** *in vitro* (Figure S9). It should be noted that longer incubation of **5** as well as **6** with AnkD resulted in significantly lower amounts of **6** compared to the shorter reaction time and the appearance of possible shunt product peaks (Figure S10). This suggests AnkD can further react with **6**, leading to degradation over longer time periods. Despite this fact, this is the first report of C-O bond formation by a PLP-dependent enzyme to the best of our knowledge. Addition of the NIS a*nkE* alongside a*nkA-D* in *A. nidulans* led to formation of compound **7** and a small amount of compound **8** (Figure 7.1b). The 1D NMR spectra of **7** matched that of NK13650 B[144], suggesting a*nkE* is responsible for installation of the citrate group onto **6** (Figure 7.1d). MS-MS analysis suggested **8** is a derivative of **7**. Based on the -42 decrease in mass of the major ion fragments in **8** compared to **7**, we propose **8**

may be a shunt product from **7** with hydrolysis of the guanidine group (Figure S17). Coexpression of the OMT a*nkF* with a*nkA-D* did not yield the O-methylated analog of **6**, which suggests O-methylation occurs after addition of citrate to generate **1**.

We speculated if additional genes could install the aspartic acid group to afford NK13650 A, the major product co-isolated with NK13650 B (**7**) from *Penicillium* sp. NF13650[144]. Secondary structure prediction by Phyre2[150] indicated the neighboring gene a*nkG* has homology to the ATP-GRASP enzyme glutathione synthetase[151], which catalyzes amide bond formation. Further analysis showed a*nkG* is conserved in homologous BGCs (Figure S1). We hypothesized AnkG may install the aspartic acid group onto **1** and **7**. Coexpression of a*nkG* with introns intact alongside a*nkA-F* in *A. nidulans* led to trace amounts of compounds **9** and **10** with MW of 723 and 737, respectively. Parallel strains containing *ankG* with the predicted introns removed showed increased production of **9** and **10**, suggesting *A. nidulans* may not splice this gene efficiently. (Figures 7.1b, S11, and S12) The 1D NMR spectra of **9** matched that of NK13650 A[144], and **10** was solved to be an analog of **1** with the addition of aspartic acid at C27 (Figure 7.1e). These results suggest AnkG catalyzes amide bond formation to link aspartic acid to the citric acid moiety, completing the biosynthesis of the NK13650 compounds and their methylated analogs.

Understanding this pathway's biosynthetic logic may facilitate mining of other BGCs that could encode for potential NK13650 derivatives as well as larger scale microbial production to make these compounds more accessible for research and engineering applications.

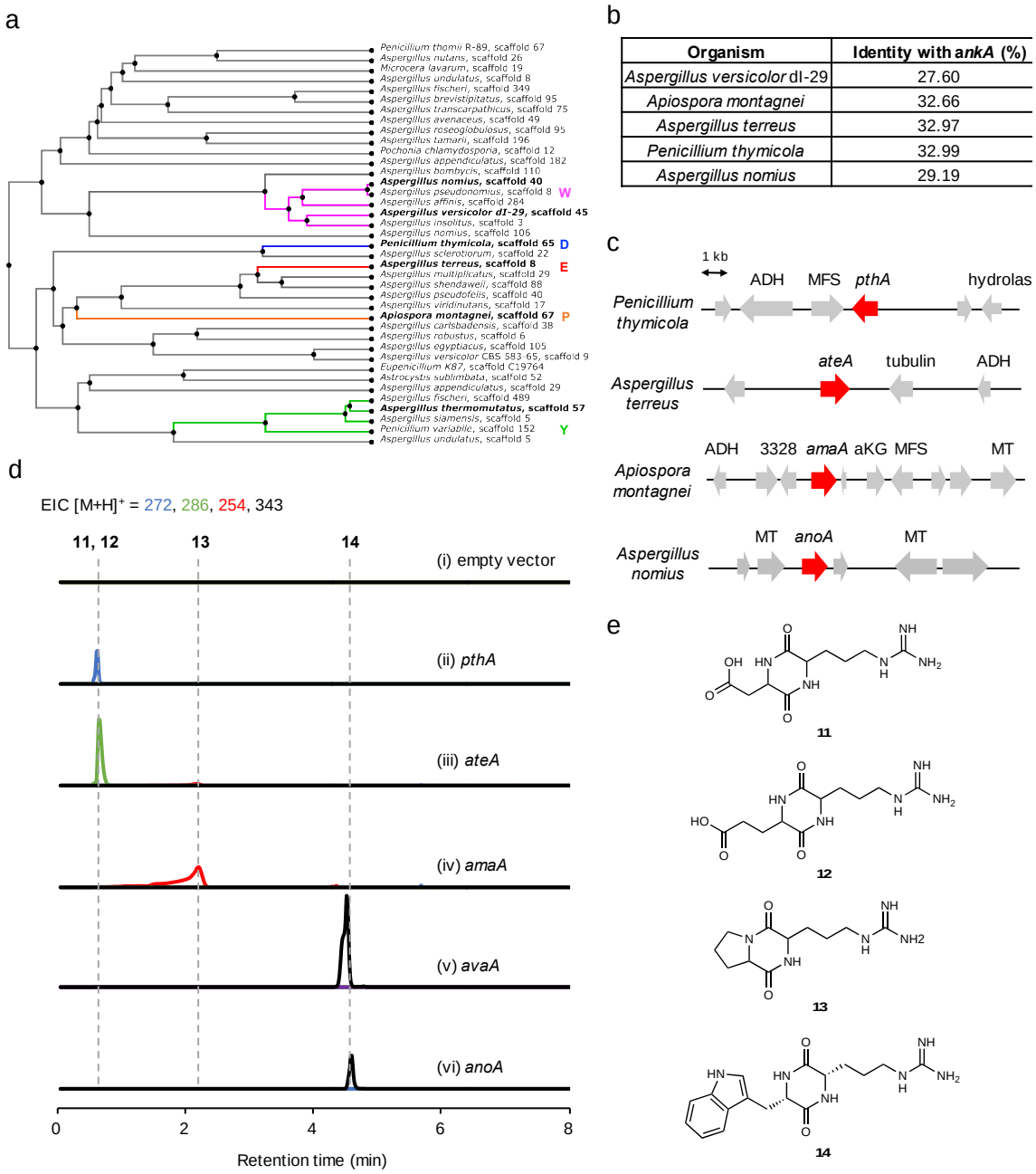


Fig. 7.2 Discovery of AnkA homologs from other fungal strains. A) MAFFT phylogenetic tree of select putative AnkA homologs. B) Table with identity of characterized AnkA homologs with *ankA*. C) The gene clusters of characterized AnkA homologs. D) QTOF analysis of extracts from expression of *ankA* homologs in *A. nidulans*. E) Structures of Arg-containing cyclodipeptides isolated from *A. nidulans* heterologous expression.

7.3 Study of AnkA homologs

To search for AnkA homologs that may catalyze similar reactions, we performed a tBlastn search using a*nkA* as a query against the JGI ascomycete genome database, NCBI nucleotide collection, as well as in-house fungal genome sequences and identified nearly 100 hits for homologs with greater than 20% identity to *ankA* (Table S6). Selected sequences were used to generate a phylogenetic tree using MAFFT software[152] (Figure 7.2a). Although the identity of most hits was low, we hypothesized these putative homologs may produce similar DKP scaffolds. Many of the putative homologs were co-localized with other tailoring genes that might further modify the core cyclodipeptide structure (Figures 7.2c, S3).

Based on the strains available in our lab, we performed heterologous expression of AnkA homologs from *Penicillium thymicola* (*pthA*), *Aspergillus terreus* (*ateA*), *Apiospora montagnei* (*amaA*), *Aspergillus versicolor* dI-29 (*avaA*), *Aspergillus nomius* (*anoA*), and *Penicillium subrubescens* (*psbA*) in *A. nidulans* (Figure 7.2b). Expression of *pthA* resulted in accumulation of compound **11** with MW of 271. The structure of **11** was solved by NMR to be cyclo-Arg-Asp. Expression of *ateA* led to production of compound **12** with MW of 285, and the structure was solved to be a new compound, cyclo-Arg-Glu. Expression of *amaA* led to production of compound **13** with MW of 253, and the 1D NMR spectra matched that of cyclo-Arg-Pro.[153] Expression of *amaA* in yeast also resulted in production of **13** (Figure S13). Expression of both *avaA* and *anoA* resulted in production of compound **14** with MW of 342, and the 1D NMR spectra matched that of cyclo-Arg-Trp.[154] (Figures 7.2d, 7.2e) The homolog from *P. subrubescens* did not yield any new compounds compared to the empty vector control. It is interesting that all active AnkA homologs utilize arginine as a substrate, since no Arg-DKP

producing CDPS or NRPS has been characterized except for the CDPS that synthesizes cyclo-Arg-Pro.[155, 156] As all of the cyclodipeptides produced by AnkA and its homologs contain arginine, we renamed this enzyme family as arginine cyclodipeptide synthases (aCDPS).

7.4 *In vitro* activity of AvaA and *in vivo* mutation assays

Since we were unable to purify AnkA from *E. coli* and yeast, the other homologs were considered for further characterization. AvaA activity was confirmed in *E. coli* and yeast, as expression in both hosts led to formation of **14** (Figure S15, S20). To determine if AvaA follows a similar mechanism to that of bacterial CDPSs, we first investigated if the reaction is RNA-dependent. Addition of Trp-*d5* to yeast lysates containing AvaA resulted in production of the labeled cyclo-Arg-Trp product with MW of 348 (**14**-*d5*). RNase A treatment of the lysate before substrate addition abolished product formation. In lysates treated with RNase A followed by addition of RNasin, an inhibitor of RNase A, formation of **14**-*d5* was only observed after addition of total yeast RNA, indicating the reaction is RNA dependent. (Figure 7.3a) Therefore, the reaction may follow a mechanism similar to that of bacterial CDPSs, which catalyze cyclodipeptide formation through cyclization of aminoacylated tRNAs.[156] In assays with desalted lysates, addition of ATP and $MgCl_2$ was required for product formation (Figure S14). We reasoned these cofactors are utilized by yeast tRNA synthetases to generate the aminoacylated substrates for AvaA. This is consistent with previous work that established amino acid tRNA synthetases require ATP and $MgCl_2$ to catalyze amino-acylation.[157]

We then performed *in vitro* reactions with AvaA and yeast Arg and Trp tRNA synthetases (ArgRS and TrpRS) purified from *E. coli* to further confirm that cyclodipeptide production requires aminoacylated tRNA. Complete reconstitution with AvaA, ArgRS, TrpRS,

Trp-*d5*, L-Arg, deacylated total yeast RNA, ATP, and $MgCl_2$ resulted in production of **14**-*d5*. Omitting ArgRS and TrpRS together and separately abolished cyclodipeptide formation, which suggests these enzymes generate the aminoacylated tRNA substrates to be utilized by AvaA. Omission of Trp-*d5* abolished product formation, whereas exclusion of L-Arg resulted in trace amounts of **14**-*d5* being produced. This background activity may be due to incomplete deacylation of Arg-tRNA in the total RNA prep, which can then be utilized by AvaA to afford the cyclodipeptide product. Exclusion of ATP and $MgCl_2$ led to loss of product formation, consistent with the results from the desalted lysate assays. (Figure 7.3b) Overall, these results are consistent with our proposal that ArgRS and TrpRS first generate the activated amino acid substrate using tRNA, followed by subsequent cyclization by AvaA (Figure 7.3e). Following this, we isolated tRNA from yeast total RNA, (Figure S21), confirmed the band's identity through RT-PCR (Figure S22), and performed *in vitro* assays with purified AvaA, ArgRS, TrpRS, unlabeled substrates, and cofactors. Addition of purified tRNA in place of total RNA still led to formation of **14** (Figure 7.3c), which indicates the reaction is tRNA-dependent and supports our proposed mechanism (Figure 7.3e).

We tested various methoxy, methyl, fluoro, bromo, chloro, nitro, cyano, and hydroxyl tryptophan analogs to determine the scope of modified cyclo-Arg-Trp variants that can be generated (Supplementary Methods, section 6). Addition of 4-fluoro Trp, 5-fluoro Trp, 6-fluoro Trp, 6-methyl Trp, and 7-methyl Trp to the reaction resulted in incorporation of these analogs into the cyclodipeptide scaffold (Figure 7.3d). The incompatibility of the other substrates may be due to inability of yeast TrpRS to activate these analogs, or AvaA may not be able to cyclize the activated derivatives. However, in the future, it may be possible to expand the library of

cyclo-Arg-Trp analogs by introducing variants of Trp and Arg tRNA synthetases that have been engineered to activate non-canonical amino acids.[158–160]

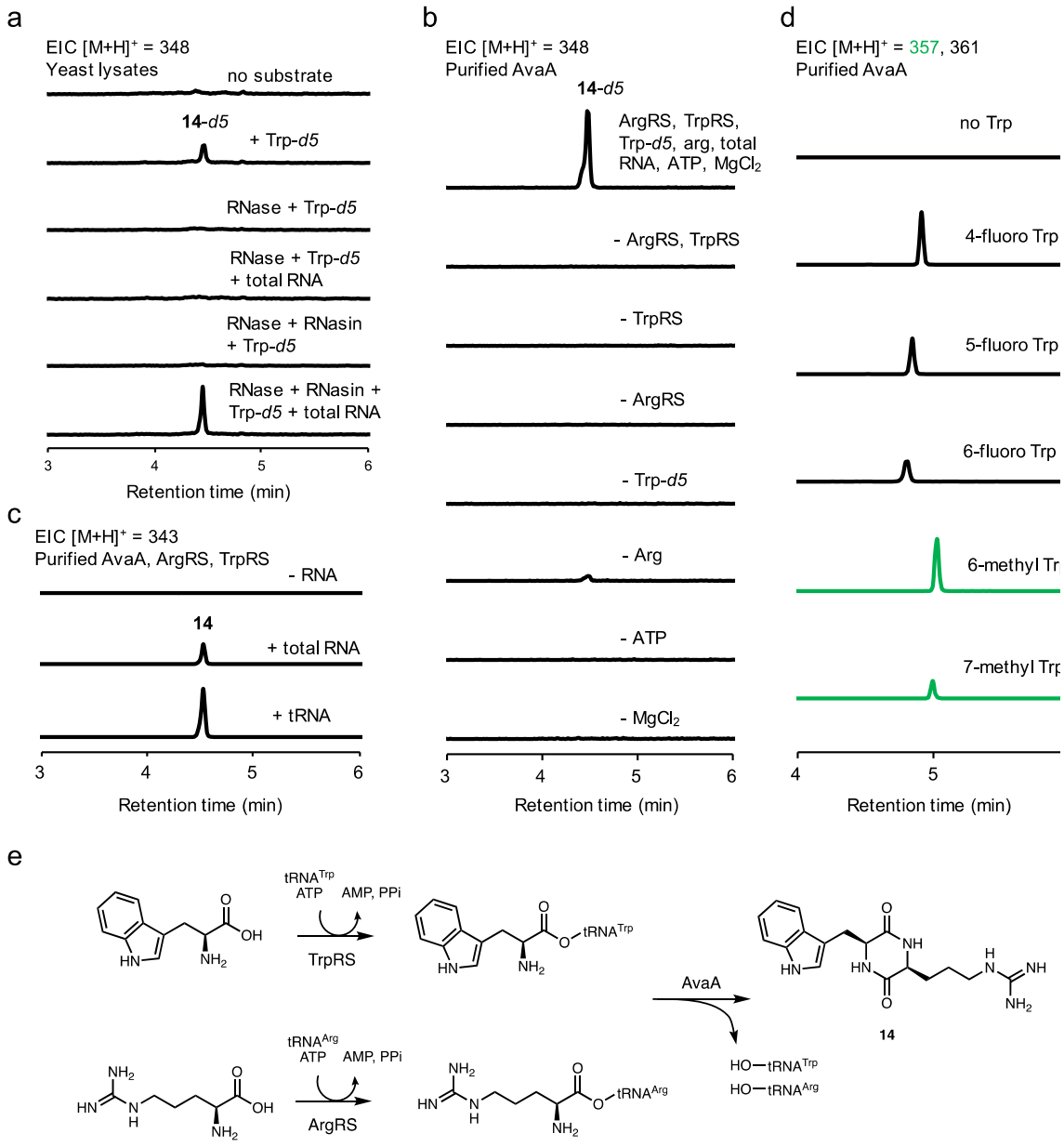


Fig. 7.3 Biochemical characterization of AvaA. a) QTOF analysis of yeast lysate assays with RNase treatment and addition of yeast total RNA. b) *In vitro* assays with purified AnkA, ArgRS, TrpRS, yeast total RNA, and cofactors. c) *In vitro* assays with purified AnkA, ArgRS, TrpRS, purified yeast tRNA, unlabeled substrates, and cofactors. d) *In vitro* assays with purified AnkA, ArgRS, TrpRS, Trp analogs, L-Arg, and cofactors. e) The proposed mechanism for cyclodipeptide formation by TrpRS, ArgRS, and AvaA.

We aligned the sequences for a*vaA* and its five validated homologs. Conserved

residues were identified at only 23 positions and a DDXXE motif was found (Figure S4). Although a*vaA* showed no alignment with bacterial CDPSs, conserved glutamic acid and tyrosine residues were detected amongst the a*vaA* homologs. We hypothesized these residues may perform a similar function to the catalytic Glu and Tyr residues of *AlbC* and other CDPSs that deprotonate the amino group to facilitate peptide bond formation.[161] The AvaA mutants E315A, E432A, Y151A, and Y515F were expressed in yeast alongside wildtype AvaA. The E315A mutation resulted in decreased activity compared to the wildtype enzyme, whereas E432A and Y515A abolished cyclodipeptide formation. The Y515F mutation led to trace amounts of **14**, so it is unclear if Y515 has a catalytic or structural function. (Figure S15) Overall, these results suggest E432 and Y515 are important for catalytic activity of the enzyme. Conservation of these residues should be considered when mining for additional putative aCDPS homologs in future studies.

7.5 Further study of tailoring genes in the *ava* cluster

Further analysis was performed on the tailoring genes of the *ava* BGC from *A. versicolor*. The BGC includes a FAD-dependent monooxygenase (FMO) (*avaB*), a kynurenine formamidase (KFA) (*avaC*), an *N*-acetyltransferase (*avaD*), as well as other putative tailoring enzymes, including five cytochrome P450s (*avaF*, *avaH*, *avaI*, *avaJ*, *avaL*), a beta-lactamase (*avaE*), a lysophospholipase (*avaG*), an epimerase (*avaK*), a dioxygenase (*avaM*), and a glycosyltransferase (*avaN*) (Figure 7.4a, Table S2). Coexpression of *avaA* and *avaB* in *A. nidulans* resulted in production of three new compounds **15**, **17**, and **18** with MW of 358, 346, and 344 after two days of growth and an additional compound **16** with MW of 316 after four days of growth on solid media. Addition of *avaC* did not lead to accumulation of other

compounds, but production of **17** and **18** increased slightly. (Figure 7.4b) NMR analysis showed **15** is a derivative of **14** with the tryptophan group morphed to a 6-5-5 ring system and hydroxylation at C9b, similar to the modified Trp moiety in brevianamide[162] and notamide.[163] **16** was solved to have a similar structure to **15** with hydrolysis of the guanidine group, similar to that observed in **1** and **2**. (Figure 7.4c) We hypothesize *A. nidulans*' primary metabolism may also be responsible for conversion of **15** to **16**.[145] The structure of compound **17** was solved to be an analog of **14** with the tryptophan group replaced by kynurenine. The structure of **18** was solved to be a derivative of **14** with modification of tryptophan to form an anthranil group with a 10 e^- conjugated system and N-O five-membered ring. (Figure 7.4c)

Based on the structures of **15**-**18**, we propose the FMO AvaB may hydroxylate **14** at C9b. The nitrogen in the DKP core may then attack C4b to form the 6-5-5 ring system in the putative shunt products **15** and **16**. Similar FMO-driven reactions were observed in the brevianamide[162] and notamide[163] pathways, although no sequence alignment was detected between AvaB and these enzymes. However, AvaB may also perform further oxidation on the transient hydroxylated species to generate the N-formyl kynurenine intermediate. AvaC or the kynurenine formamidase from *A. nidulans*' primary metabolism may cleave the formate group, resulting in formation of **17**. We propose AvaB may then hydroxylate the NH_2 group on C5a. The hydroxyl group may attack the C9b carbonyl followed by spontaneous hydrolysis to afford the anthranil moiety in **18**. (proposed mechanism in Figure S19) To the best of our knowledge, this is the first report of an anthranil functional group found in natural products. The mechanism of the reaction by AvaB and potentially AvaC is under investigation.

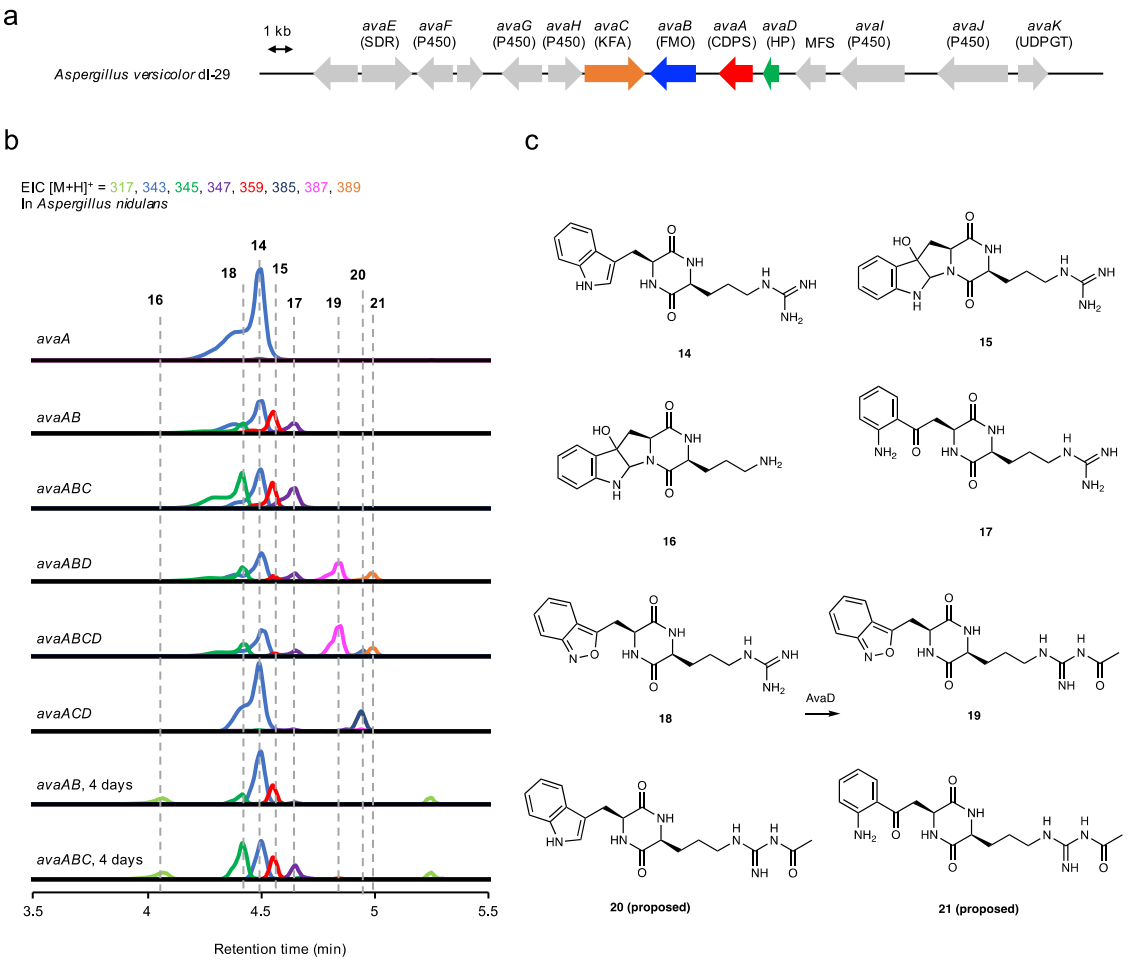


Fig. 7.4 Heterologous expression of *ava* tailoring enzymes from *A. versicolor* dI-29. A) The *ava* BGC containing the aCDPS *avaA* flanked by tailoring genes. B) QTOF analysis of extracts from expression of *ava* pathway genes in *A. nidulans*. C) Structures of cyclo-Arg-Trp analogs modified by AvaB and AvaD isolated from *A. nidulans* heterologous expression.

Coexpression of the *N*-acetyltransferase *avaD* with *avaAC* or *avaABC* in *A. nidulans* resulted in the production of three new compounds **19**, **20**, and **21** with MW of 386, 384, and 388, respectively (Figure 7.4b). **19** was solved to be an analog of **18** with acetylation on N18 of the arginine group (Figure 7.4c). The MS-MS fragment profiles of **14**, **20** and **17**, **21** were very similar, suggesting these compounds are structurally related (Figure S18). Based on the mass increase of 42, we propose **20** and **21** may be acetylated-Arg derivatives of **14** and **17**,

respectively (Figure 7.4c). Although a few natural products containing acetylated arginine have been isolated previously,[164,165] this is the first report of an arginine acetyltransferase with *in vivo* activity to the best of our knowledge. The sequences of *avaB* and *avaD* can be used as queries to mine for new natural products with related modifications and resulting functional groups. The roles of the remaining tailoring genes *avaE-N* are currently under investigation.

7.6 Conclusions

In this work, we discovered a new family of enzymes termed "aCDPS" from fungi that produce Arg-containing cyclodipeptides through cyclization of aminoacylated tRNAs. Motivated by the potential of studying HPs co-localized with tailoring genes, we performed in-depth analysis of the *ank* cluster, resulting in the discovery of the aCDPS AnkA and elucidation of the biosynthetic pathway of the NK13650 compounds and additional analogs. Five aCDPS homologs were characterized with incorporation of Tyr, Trp, Glu, Asp, and Pro. By using aCDPS as a query to mine for additional BGCs, novel compounds derived from the modified Arg-Trp scaffold were discovered from the *ava* cluster. Study of the tailoring genes in these pathways led to the discovery of enzymes that catalyze novel reactions, including a PLP-dependent enzyme that catalyzes C-O bond formation, an FMO that forms an anthranil moiety, and an arginine acetyltransferase. The aCDPS constitute a new class of core genes in natural product biosynthesis, and together with the other tailoring enzymes found here, can significantly broaden opportunities in future genome mining endeavors for novel natural products.

8. MATERIALS AND METHODS

8.1 Plasmids and strains for engineering *S. cerevisiae* for monoterpene and iridoid production

All restriction enzymes used in this work were purchased from NEB. Q5® High-Fidelity DNA Polymerase (NEB), AccuPrimeTM Pfx DNA Polymerase (InvitrogenTM), and Phusion (NEB) were used for PCR. Oligonucleotides were ordered from IDT. Plasmids were assembled through restriction-digestion cloning, NEB Hifi assembly, or yeast homologous recombination. Plasmids were maintained and propagated in *E. coli* strains DH10b, TOP10, and XL-1. Plasmids used for metabolite production and western blotting are listed in Table 1. The plasmids pδBLE2.0m-ERG13/ERG10, pδBLE2.0m-ERG12/tHMG1, pδBLE2.0m-ERG19/ERG8, and pδBLE2.0m-IDI1/ERG20 containing expression cassettes for mevalonate pathway genes targeted to the mitochondria were gifts from Dr. Jifeng Yuan. pCRCT was a gift from Huimin Zhao (Addgene plasmid # 60,621) (Bao 2014). The GES from *C. roseus* was codon optimized and cloned into the plasmid p4 by Gen9.

The *S. cerevisiae* strains JHY651, X303, and CEN.PK2-1C were compared for production of geraniol. Gene integrations and deletions were performed through CRISPR-Cas9 unless otherwise noted using the protocol described in mode iv from Horwitz et al. (Horwitz 2015). Guide RNA sequences were designed with the CRISPy Cas9 target finder (Jakočiūnas 2015). Plasmids with gRNA sequences contained *hygR* or *kanMX* antibiotic resistance markers. CRISPR-Cas9 was integrated into the YPRCTy1-2 locus in JHY651 to construct **S1** using LEU2 marker mediated integration. For construction of strains **S2**-**S17**, the gRNA plasmid was transformed as 2 linear pieces with the plasmid backbone and a

homology arm containing the gRNA target sequence. To construct **S18** and **S19**, the integrated CRISPR-Cas9 cassette was replaced with a cassette containing the URA3 marker flanked by *Sce*I cut sites. A second transformation mediated by CRISPR-Cas9 and *Sce*I was performed to replace URA3 with mitochondrial or cytosolic ObGES. For construction of strains **S18**-**S27**, the gRNA plasmid was transformed as 3 linear pieces: one piece with the Cas9 cassette and two other pieces with the plasmid backbone and a homology arm containing the gRNA target sequence. All strain modifications were verified by sequencing. The strains used in this study are listed in Table 1.

The primers used in this study are listed in Table S2. Gene sequences and their sources are listed in Table S3. Other plasmids used for cloning are listed in Table S4. Yeast transformations were performed through the LiOAc/ssDNA method (Gietz and Schiestl, 2007). Transformants were selected using uracil dropout media, leucine dropout media, histidine dropout media, uracil and histidine dropout media, or YPD plates supplemented with 200 mg/L hygromycin or G418.

Table 1. Strains and plasmids used in this study.

Strain	Parent	Genome modifications to parent	Reference
BY4742	S288C	*MATα his3Δ1 leu2Δ0 ura3Δ0 lys2Δ0*	(Brachmann et al., 1998)
DHY214	BY4742	*SAL1^{+} CAT5(91M) MIP1(661T) MKT1(30G) RME1(INS-308A) TAO3(1493Q) HAP1+*	(Harvey et al., 2018)
JHY651	DHY214	*MATα prb1Δ pep4Δ*	*ibid.*
X303-1B	W303	*MATα ADE2 TRP1 ura3Δ0 leu2-3,-112 his3-11,-15 CAN1 MIP1(661T) SSD1+*	*ibid.*
CEN.PK2-1C	n/a	*MATa; his3D1; leu2-3_112; ura3-52; trp1-289; MAL2-8c; SUC2*	(Entian and Kötter 2007)
S1	JHY651	*YPRCTy1–2::iCas9::LEU2*	*This study*
S2	S1	*ura3Δ::TEF1p-CrGES-CYC1t*	*This study*
S3	S2	*rox1Δ::TEF1p-ERG20*(f)ObGES-CYC1t*	*This study*
S4	S3	*oye2Δ::TEF1p-mFPS-CYC1t*	*This study*
S5	S4	*erg9p truncation*	*This study*
S6	S5	*bts1Δ*	*This study*
S7	S5	*bts1Δ::TEF1p-IDI1-CYC1t*	*This study*

S8	S7	*ypl062wΔ*	*This study*
S9	S7	*yjl064wΔ*	*This study*
S10	S7	*ypl062wΔ::GPDp-tHMG1-ADH1t*	*This study*
S11	S7	*yjl064wΔ::TEF1p-HMG2*-CYC1t*	*This study*
S12	S11	*ypl062wΔ::GPDp-tHMG1-ADH1t*	*This study*
S13	S12	*oye3Δ::TEF1p-IDI1-CYC1t*	*This study*
S14	S1	*ura3Δ::GAL10p-nCox4-ERG13-ADH1t;GAL1p-nCox4-ERG10-CYC1t*	*This study*
S15	S14	*X-2Δ::GAL10p-nCox4-ERG12-ADH1t;GAL1p-nCox4-tHMG1-CYC1t*	*This study*
S16	S15	*HOΔ::GAL10p-nCox4-ERG19-ADH1t;GAL1p-nCox4-ERG8-CYC1t*	*This study*
S17	S15	*rox1Δ::GAL10p-nCox4-IDI1-ADH1t;GAL1p-nCox4-mFPS-CYC1t*	*This study*
S18	S17	*YPRCTy1–2Δ::GAL1p-nCox4-ObGES-CYC1t*	*This study*
S19	S17	*YPRCTy1–2Δ::GAL10p-ObGES-CYC1t*	*This study*
S20	S18	*oye2Δ::OYE2p-CrG8H-OYE2t*	*This study*
S21	S18	*oye2Δ::GAL1p-CrG8H-CYC1t*	*This study*
S22	S18	*oye3Δ::OYE3p-CrG8H-OYE3t*	*This study*
S23	S18	*oye3Δ::GAL1p-CrG8H-CYC1t*	*This study*
S24	S20	*oye3Δ*	*This study*
S25	S21	*oye3Δ*	*This study*
S26	S22	*oye2Δ*	*This study*
S27	S23	*oye2Δ*	*This study*

Plasmid	**Description**	**Reference**
pxp218 (p1)	2μ; URA3; AmpR (Addgene plasmid #26831)	(Fang et al., 2010)
pxp318 (p2)	CEN/ARS; URA3; AmpR (Addgene plasmid #26837)	(Fang et al., 2010)
pxp320 (p3)	CEN/ARS; HIS3; AmpR (Addgene plasmid #26838)	(Fang et al., 2010)
p4	CEN/ARS; URA3; AmpR; *TEF1p-CrGES-CYC1t*	*This study*
p5	CEN/ARS; URA3; AmpR; *TEF1p-ObGES-CYC1t*	*This study*
p6	2μ; URA3; AmpR; *PGK1p-AgGPPS-CYC1t*	*This study*
p7	2μ; URA3; AmpR; *PGK1p-ERG20*K197G-CYC1t*	*This study*
p8	2μ; URA3; AmpR; *PGK1p-mFPS-CYC1t*	*This study*
p9	2μ; URA3; AmpR; *PGK1p-VrtD-CYC1t*	*This study*
p10	2μ; URA3; AmpR; *GAL1p-nCox4-ObGES-CYC1t*	*This study*
p11	2μ; URA3; AmpR; *GAL1p-ObGES-CYC1t*	*This study*
p12	CEN/ARS; URA3; AmpR; *GAL1p-CrGOR-PRM9t; GAL10p-CrISY-CPS1t*	*This study*
p13	2μ; HIS3; AmpR; *GAL1p-CrG8H-CYC1t*	*This study*
p14	CEN/ARS; URA3; AmpR; *GAL1p-nCox4-mFPS-cFLAG-CYC1t*	*This study*
p15	CEN/ARS; HIS3; AmpR; *GAL1p-nCox4-GAPDH-cFLAG-CYC1t; GAL1p-COX4-cFLAG-CYC1t*	*This study*
p16	2μ; URA3; AmpR; *GAL1p-GOR-PRM9t; GAL10p-CrISY-CPS1t*	*This study*
p17	2μ; URA3; AmpR; *GAL10p-CrG8H-CYC1t; GAL1p-CrGOR-PRM9t*	*This study*
p18	CEN/ARS; HIS3; AmpR; *GAL10p-CrISY-CYC1t*	*This study*

8.2 Yeast culturing for metabolite production assays

For fully integrated strains, single colonies of each construct were inoculated in 1 mL YPD. For plasmid bearing strains, single colony transformants were inoculated in 1 mL

synthetic defined (SD) 2% glucose media with the appropriate dropouts. Starter cultures were shaken at 28°C and 250 rpm for 16-24 hours. Culture tubes containing 3 mL of fresh YPD, YP 0.2% glucose 1.8% galactose, SD 2% glucose, or SD 0.2% glucose 1.8% galactose were inoculated with 100 μL of starter culture. The medium used for each strain's subcultures is listed in Table S5. For strains being tested for geraniol production, 400 μL of dodecane was layered on top of each subculture. Subcultures were shaken at 28°C and 250 rpm for 48 hours unless otherwise noted.

8.3 Culture extraction and quantification

For quantification of geraniol production, the two-phase subcultures were centrifuged at 4300 rpm at 20°C for 6 minutes to separate the dodecane and aqueous layers. 8-hydroxygeraniol and nepetalactol production was measured by extracting 700 μL subculture with 300 μL of an organic phase consisting of 25% acetone and 75% ethyl acetate. The samples were vortexed for 1 minute then centrifuged for 10 minutes. The organic layers were analyzed on an Agilent Technologies GC-MS 6890/5973 equipped with a DB-FFAP column. An inlet temperature of 220°C and constant pressure of 4.2 psi were used. The oven temperature was held at 60°C for 5 min and then ramped at 60°C/min for 1.5 min, followed by a ramp of 15°C/min for 16 min and a hold for 10 min. All measurements were taken in biological triplicate unless otherwise noted. Standard curves were generated using pure geraniol, 8-hydroxygeraniol, and nepetalactol.

8.4 Isolation of mitochondria and western blotting

Low copy plasmids containing expression cassettes for *N*-cox4-mFPS from *Gallus gallus* (p14) and COX4 and GAPDH from *S. cerevisiae* (p15) with *C*-terminal FLAG tags under

the GAL1-10 promoter were transformed into the strain **S16**. A single colony transformant was inoculated in 10 mL SD 2% glucose uracil and histidine dropout media. The starter culture was shaken at 28°C and 250 rpm for 24 hours. Two flasks each containing 500 mL YP 0.2% glucose 1.8% galactose were inoculated with 5 mL starter culture. The large-scale cultures were shaken at 28°C and 250 rpm for 24 hours. To harvest the cells, the cultures were centrifuged and the cell pellet was collected. The crude mitochondrial and cytosolic fractions were isolated as described previously (Gregg 2009). The isolated mitochondria were then lysed by freezing and thawing 10 times. Total protein in each subcellular fraction was measured with the Bradford assay. 10 μg of total protein from the mitochondrial and cytosolic fractions were loaded on a polyacrylamide gel. Standard electrophoresis was used to fractionate the protein samples, which were then transferred to polyvinylidene difluoride membranes. The membranes were incubated with the following primary and secondary antibodies (Sigma-Monoclonal anti-FLAG M2; Goat anti-Mouse IgG-Fc Fragment HRP conjugated, Bethyl A90-131P). Proteins were visualized by enhanced chemiluminescence using SuperSignal™ West Dura Extended Duration Substrate purchased from Thermo Scientific.

8.5 Growth curve assay

Single colony transformants were inoculated in 3 mL SD 2% glucose uracil dropout media. Starter cultures were shaken at 28°C and 250 rpm for 16 hours. 100 □L of SD 0.2% glucose 1.8% galactose uracil dropout media was inoculated with starter culture to an initial OD_{600} of 0.02 in a Corning96fc UV transparent plate. The plate was covered with a sealing film (Microseal® 'B' PCR Plate Sealing Film, adhesive, optical #msb1001 from BioRad).

Growth was measured on a Tecan Infinite M200 Pro plate reader. The program temperature was set to 28.5°C. The program was run for 102 cycles, with each cycle consisting of 840 s of orbital shaking (3 mm amplitude) and 60 s of orbital shaking (1.5 mm amplitude). Absorbance at 600 nm was measured after each cycle.

8.6 Fed-batch fermentation and production of 8-hydroxygeraniol

A New Brunswick BioFlo/CelliGen 115 fermenter equipped with a 2 L vessel was used for the continuous fed-batch fermentation. A single colony from **S25** was inoculated in 2 mL of liquid YPD media, which was shaken at 28°C and 250 rpm for 16 h. The starter culture was then transferred to 50 mL of liquid YPD and was shaken at 28°C and 250 rpm for 24 h. The culture was centrifuged at 800 g for 5 min and resuspended in 20 mL of fresh YPD. The fermenter working volume was 1.5 L. The media consisted of YP 1.8% galactose 0.2% glucose. After inoculation to an OD_{600} of 0.16, the fermenter temperature, pH, and dissolved oxygen were automatically regulated. Temperature was maintained at 28°C using a heating jacket and recirculated cooling water, which was also used to cool the exhaust condenser. Fermentation pH was maintained at ~ 5 through addition of 2 M NaOH or 2 M HCl. The dissolved oxygen set point was programmed to 25% by varying the agitation rate. After 18 hours of culturing, a 1 M galactose solution was continuously fed to the fermenter at a flow rate of 3 mL/h to maintain induction of pathway genes. Galactose feeding was stopped at 54 hours and the fermentation was ended at 60 hours. Extractions were performed in technical triplicate to quantify 8-hydroxygeraniol production every 6 hours.

8.7 Strains and culture conditions for genome mining of alkaloidal terpene compounds and cyclodipeptide derivatives

Aspergillus flavus NRRL3357 was 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. *Aspergillus thermomutatus*, *Aspergillus versicolor* dI-29, *Apiospora montagnei*, *Aspergillus terreus*, *Penicillium thymicola*, *Aspergillus nomius*, and *Penicillium subrubescens* 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. *Aspergillus nidulan*s A1145[31] 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 $NaNO_3$, 10.4 g of KCl, 10.4 g of $MgSO_4 \cdot 7H_2O$, 30.4 g of KH_2PO_4 were dissolved in 1 L of double distilled water. For preparation of the trace element solution, 2.20 g of $ZnSO_4 \cdot 7H_2O$, 1.10 g of H_3BO_3, 0.50 g of $MnCl_2 \cdot 4H_2O$, 0.16 g of $FeSO_4 \cdot 7H_2O$, 0.16 g of $CoCl_2 \cdot 5H_2O$, 0.16 g of $CuSO_4 \cdot 5H_2O$, and 0.11 g of $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ were dissolved in 100 mL of double-distilled water, and the pH was adjusted to 6.5.[166] 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.

8.8 General DNA manipulation techniques for heterologous expression of biosynthetic gene clusters and protein expression in *E. coli*

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* strain JHY651[35] was used as the yeast host for *in vivo* homologous recombination to construct the *A. nidulans* expression plasmids. *S. cerevisiae* RC01[167] was used for cytochrome P450 expression and related biotransformations. JHY651[35] was used for expression of FlvE.

For isolation of RNA from *A. nidulans* transformants containing the *flv* gene cluster, 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.

8.9 Heterologous expression of the *flv* gene cluster, the *ank* gene cluster, the *ava* gene cluster, and aCDPS homologs in *A. nidulans*

To construct plasmids for heterologous expression in *A. nidulans* A1145, the plasmids

pYTU, pYTP, and pYTR[4] with auxotrophic markers for uracil (*pyrG*), pyridoxine (*pyroA*), and riboflavin (*riboB*), respectively, were used as backbones to insert genes. Genes from the *flv* cluster and their native terminators were amplified by PCR with overhang primers using the genomic DNA (gDNA) of *A. flavus* as the template. Genes from the *ank* and *ava* gene clusters as well as AnkA homologs from *A. montagnei*, *A. terreus*, *P. thymicola*, *A. nomius*, and *P. subrubescens* and their native terminators were amplified by PCR with overhang primers using the genomic DNA (gDNA) from the native hosts as the template. Constitutive gpdA promoters from *A. niger* (gpdAp), *Penicillium oxalicum* (POgpdAp), and *Penicillium expansum* (PEgpdAp) as well as coxAp from *A. niger* 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* JHY651[3] 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* A1145 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* A1145 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 $MgSO_4$, 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-HCI, 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-HCI, pH 7.5). The protoplasts were then resuspended in 1 mL of STC buffer.[6]

For each transformation, plasmids were added to 100 µl of the *A. nidulans* A1145 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.

8.10 Chemical analysis and isolation of flavunoidines and precursors

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 FlvDE and FlvCDE were grown on CD-ST agar for 3 days at 28 °C and then extracted with acetone. For small scale analysis of **9**-**11** in yeast, *S. cerevisiae* RC01 expressing FlvDE was inoculated in 1 mL of dropout media for 24 hours. 100 µL 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 small scale analysis of **8**, *S. cerevisiae* JHY651 expressing FlvE (codon optimized) was inoculated in 1 mL of dropout media for 24 hours. 100 µL of starter culture was used to inoculate 3 mL of YPD with a 400 µL overlay of dodecane. The cells were grown at 28 °C, 250 rpm for 48 hours. The two-phase subcultures were centrifuged at 4300 rpm at 20 °C for 6 minutes to separate the organic and aqueous layers. The organic layers were analyzed on an Agilent Technologies GC-MS 6890/5973 equipped with a DB-FFAP column. An inlet temperature of 220°C and constant pressure of 4.2 psi were used. The oven temperature was held at 60°C for 5 min and then ramped at 60°C/min for 1.5 min, followed by a ramp of 15°C/min for 16 min and a hold for 10 min.

For isolation of **1**, **2**, **3**, and **12**, transformants of *A. nidulans* strains were grown on 4 L of solid CD media for 7 days at 28 °C and then extracted with acetone. The extracts were concentrated by rotary evaporator and partitioned between ethyl acetate and H_2O four to eight times, followed by evaporation of the organic phase for **1**, **3**, and **12**. For **2**, the acetone extract was dried and dissolved in methanol. 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 5% B; 5-30 min 5-100% B; 30-40 min 100% B). Fractions containing the target compounds were combined and used for further purification by HPLC with a semi-preparative reverse-phase column (Agilent, C18, 5 µm, 100 Å, 10 × 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-3 min 5% B; 3-23 min 5-40% B; 23-30 min 95% B; 30-35 min 5% B for compounds **1**-**3** and 0-3 min 5% B; 3-23 min 5-50% B; 23-30 min 95% B; 30-35 min 5% B for compound **12**.

To prepare **7a**, **9a**, **9b**, and **10a**, transformants of *A. nidulans* strains were grown in 4 L of liquid CD-ST media for 3-4 days at 28 °C, 250 rpm. For isolation of **11**, *S. cerevisiae* RC01 expressing FlvDE was inoculated in 5 mL of dropout media for 24 hours. The starter culture was used to inoculate 1 L of YPD, and the cells were shaken at 28 °C, 250 rpm for 48 hours. 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 for **7a** and 0-8 min 30-80% B; 8-9 min 80% B; 9-10 min 30% B for **9a**, **9b**, **10a**, and **11**.

For isolation of **8**, *S. cerevisiae* JHY651 expressing FlvE (codon optimized) was

inoculated in 5 mL of dropout media for 24 hours. The starter culture was used to inoculate 1 L of YPD, and the cells were shaken at 28 °C, 250 rpm for 36 hours. The cell pellet was harvested by centrifugation and extracted with acetone. The acetone extract was dried and resuspended in acetonitrile, followed by extraction with hexane. The hexane extract was absorbed with 3 g of silica, which was purified with the CombiFlash system (Teledyne) using normal phase column chromatography with hexane as the mobile phase.

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.

8.11 Structure elucidation of compound flavunoidine 1

Compound **1** was obtained as a white powder. Its HRESIMS (*m/z* 504.4214 $[M + H]^+$) was in agreement with a molecular formula of $C_{30}H_{53}N_3O_3$, requiring six degrees of unsaturation. The ^{13}C NMR spectra showed 30 carbon signals, including 7 methyl, 12 methylene, 6 methine, and 5 quaternary carbons by assignment of protons to corresponding carbons in the HSQC spectrum. Further 2D NMR (COSY and HMBC) data established the linkages of substructures. In the HMBC spectrum, correlations from H_3-13 (δ_H 1.09) to C-8 (δ_C 44.0), C-11 (δ_C 42.9) and C-12 (δ_C 79.5), from H_3-14 (δ_H 1.09) to C-1 (δ_C 46.9), C-2 (δ_C 43.2) and C-3 (δ_C 31.9), from H_3-15 (δ_H 1.09) to C-5 (δ_C 52.2), C-6 (δ_C 81.7) and C-7 (δ_C 67.0), as

well as correlations from H-10 (δ_H 4.92, d, J = 8.0 Hz) to C-1, C-2, C-5, C-9 (δ_C 28.4) and C-12 were observed. In association with COSY relationships from H-2 (δ_H 1.65, m) to H-5 (δ_H 2.20, m), H-7 (δ_H 2.72, d, J = 4.3 Hz) to H_2-9 (δ_H 1.44 and 1.20, m), and between H-10 and H-11 (δ_H 2.00, dd, J = 15.0, 8.0 Hz; 1.42, d, J = 15.0 Hz), a tricyclic sesquiterpene core was determined. A spin system from H-1′ (δ_H 2.46, m) to H-5′ (δ_H 2.29, t, J = 8.0 Hz) together with chemical shifts of C-1′ (δ_C 48.6) and C-5′ (δ_C 59.1), indicated a cadavarine moiety, which has two methyl groups located on N-5′ from correlations between H_6 singlet at δ_H 2.18 and C-5′. The other end of the cadavarine chain was attached to C-7 through a C-N bond according to HMBC correlation from H-7 to C-1′. The remaining NMR resonances were attributed to an unusual 5,5-dimethyl pipecolate moiety, which was identified by HMBC correlations from H-2″ to C-4″ (δ_C 37.3) and C-6″ (δ_C 56.9) in addition to correlations from two methyl groups at δ_H 0.88 (H-8″) and 0.80 (H-9″) to C-4″, C-5″ (δ_C 29.6) and C-6″. HMBC evidence from H-10 to the carbonyl carbon C-7″ (δ_C 173.5) showed the precursor 5,5-dimethyl pipecolic acid condensed with the terpene core though ester bond formation. Thus, the basic skeleton of compound **1** was established. The molecular weight and degrees of unsaturation indicated an oxygen bridge between C-6 and C-12, which lends a cage structure to the terpene moiety. Therefore, the planar structure of **1** was elucidated. The stereochemistry of compound **1** was solved by NOE data analysis in association with the crystal structure of precursor **10a**, in which the absolute configurations of C-1, C-5, C-6, C-7, C-8 and C-12 were unambiguously determined by X-ray diffraction. In the NOESY spectrum, interaction between H-7 and H_3-13 indicated the opposite orientation of the C-N bond to the oxygen bridge. The NOE interactions from H-10 to H-9b (δ_H 1.20, m) and to H_3-14 suggested spatial approximation of H-10 and *Me*-14. The

absolute stereochemistry of the terpene core in **1** was finally elucidated as 1*R*,2*R*,5*S*,6*S*,7*S*,8*R*,10*R*,12*R*. Moreover, the configuration of C-2″ was determined by $[\alpha]_D^{24}$ values of its biosynthetic building block **2**, which gave a $[\alpha]_D^{24}$ = -52.0 (*c* 0.2, H_2O), in accordance with that reported for L-(-)-pipecolic acid, indicating a *S* configuration of C-2″.

8.12 Expression and purification of FlvE from *S. cerevisiae* and *in vitro* assays

FlvE was cloned into the yeast expression vector XW55[167] with a *C*-terminal His-tag. In this plasmid, expression of *flvE* was under the control of the *ADH2* promoter. Overexpression and subsequent protein purification of FlvE were performed as follows:[168] the yeast strain JHY651 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 FlvE using affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified FlvE 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 FlvE was analyzed by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to calculate protein concentration.

Enzyme assays of FlvE were performed in 50 mM Tris-HCl (pH 7.5) buffer with a final volume of 100 µL. The assay contained 10 mM of $MgCl_2$, 0.1 mM of $MnCl_2$, 100 µM of *E,E*-

FPP, and 56.5 μM of recombinant FlvE (intact or boiled). The reactions were incubated at 30 °C for 1 hour and extracted with 100 μL hexanes. Samples were centrifuged and the organic layers were analyzed by GC-MS as described previously.

8.13 Expression and purification of FlvF from *E. coli* BL21(DE3)

The *flvF* gene was amplified with overhang primers from the cDNA of *A. nidulans* expressing all nine genes in the *flv* cluster. The PCR product and expression vector pET28a were digested with *NdeI*/*XhoI* and ligated with T4 ligase (Invitrogen). The sequence of the resulting plasmid p2018 (Table S2) was confirmed by DNA sequencing. Overexpression and subsequent protein purification of FlvF were performed as follows:[169] 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 (OD_{600}) 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 His_6-tagged FlvF using affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified FlvF 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 FlvF was analyzed by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to measure protein concentration. Aliquots of purified FlvF were flash frozen and stored at -80 °C.

8.14 Expression and purification of microsomes containing FlvD from *S. cerevisiae* and *in vitro* assays

The *flvD* gene was amplified with overhang primers from the cDNA of *A. nidulans* expressing all nine genes in the *flv* cluster. Amplified DNA was inserted into *SpeI*/*PmlI*-digested vector pXW02[4] by yeast recombination to generate p2019 (Table S2). In this plasmid, expression of *flvD* was under the control of the *ADH2* promoter. The sequence of the resulting plasmid p2019 was confirmed by DNA sequencing. *Saccharomyces cerevisiae* RC01 harboring p2019 was inoculated in 3 mL of selective dropout media for 24 hours. 40 mL of YPD was inoculated with 3 mL of starter culture and grown for 48 hours. The cells were spun down and resuspended in 4 mL of TEGM buffer (50 mM of Tris HCl, 1 mM of EDTA, 20% glycerol, 1.5 mM of 2-mercaptoethanol, pH 7.4, dissolved with one protease inhibitor tablet, Pierce). Cells were disrupted by beads (0.5 mm in diameter, Biospec) for 10 min with 30 s shaking intervals on ice. The lysate was centrifuged at 4 °C, 10,000 *g* for 10 minutes to pellet cell debris. The supernatant was centrifuged at 4 °C, 20,000 *g* for 1 hour to obtain the microsomal fractions, which were resuspended in 1 mL of storage buffer (50 mM of Tris-HCl, 10% glycerol), aliquoted, and stored at -80 °C.[170]

Enzyme assays of FlvD were performed in 50 mM Tris-HCl (pH 7.4) buffer with a final volume of 500 µL. The assay contained 400 µM of compound **8**, 800 µM of NADPH, and 20 µL of microsomal fractions containing FlvD. A reaction mixture without the substrate (**8**) was prepared in parallel as a negative control. The reaction was shaken in a 10 mL culture tube at 250 rpm, 28 °C for 12 hours and extracted with 2 volumes of ethyl acetate. The organic layers were dried and dissolved in methanol for LC-MS analysis. For assays with H_2O-18, reactions

were performed in a total volume of 100 µL with 85% consisting of H_2O-18 (Sigma). The reactions were shaken at 250 rpm at 28 °C for 6 hours and were prepared for LC-MS analysis as described previously.

Combination assays with FlvF and FlvD were performed in 50 mM Tris-HCl (pH 7.4) buffer with a final volume of 500 µL. The assays contained 400 µM of compound **8**, 800 □M of NADPH, 15 mM of $MgCl_2$, 100 µM of dimethyl cadaverine, 100 µM of recombinant FlvF, and 20 µL of microsomal fractions containing FlvD. Reaction mixtures without the substrate (**8**) or FlvF were prepared in parallel as negative controls. A reaction using **10a** instead of **8** as the substrate was also prepared. The reactions were shaken at 250 rpm, 28 °C for 15 hours and were extracted with 2 volumes of ethyl acetate. The organic layers were dried and dissolved in methanol for LC-MS analysis.

8.15 Biotransformation assays

To verify the function of FlvD (P450) function by yeast biotransformation, *Saccharomyces cerevisiae* RC01 transformed with p2019 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. **8** 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.

For the biotransformation assays in *A. nidulans*, all fed compounds except **8** were dissolved in CD agar at a final concentration of 50 µM. Transformants of *A. nidulans* strains

were grown on the prepared plates for 3 days at 28 °C and extracted with acetone. The organic phase was dried and dissolved in methanol for LC-MS analysis. To feed **8**, transformants of *A. nidulans* strains were grown on CD agar for 3 days at 28 °C. 20 µL of a hexane solution containing 2 mg/ml of **8** was layered on top of the cells. The cells were incubated at 28 °C for 12 hours and prepared for LC-MS analysis as described previously.

8.16 Chemical analysis and isolation of compounds from the *ank* and *ava* clusters and aCDPS homologs

For small scale metabolite analysis in *A. nidulans*, transformants were grown on CD agar for 2-6 days at 28 °C and then extracted with methanol. *A. nidulans* strains expressing AvaABD, AvaA-D, and AvaACD were grown on CD agar for 2 days at 28 °C and then extracted with acetone. For small scale analysis of production of compounds **3**, **13** and **14** in yeast, *S. cerevisiae* JHY651 strains expressing AnkA, AmaA, and AvaA (wildtype and mutants) were inoculated in 1 mL of dropout media for 24 hours. 100 µL of starter culture was used to inoculate 3 mL of YPD. The cells were grown at 28 °C, 250 rpm for 48 hours and the cell pellets were extracted with acetone for **3** and **13** and methanol for **14**, respectively. The organic phases were dried and dissolved in 10% water 90% 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.

QTOF and MS-MS analyses were performed on an Agilent Quadrupole Time of Flight LC/MS (6545 LC/Q-TOF).

For isolation of **1** and **2**, an *A. nidulans* transformant expressing AnkA-F was grown on 4 L of solid CD media for 5 days at 28 °C and then extracted with methanol. For isolation of **7**, an *A. nidulans* transformant expressing AnkA-E was grown on 4 L of solid CD media for 5 days at 28 °C and then extracted with methanol. For isolation of **9**, an *A. nidulans* transformant expressing AnkABCDEG was grown on 4 L of solid CD media for 3 days at 28 °C and then extracted with methanol. For isolation of **10**, an *A. nidulans* transformant expressing AnkA-G was grown on 4 L of solid CD media for 3 days at 28 °C and then extracted with methanol. The extracts were concentrated and the methanol was removed by a rotary evaporator. The pH of the aqueous extracts was lowered to 4, and the extracts were mixed with HP20 resin for 1 hour. The resin mixtures were then poured through a column, and the flowthrough was discarded. The resin was washed with water and 25% methanol with 0.1% formic acid. The compounds were eluted from the resin with 50% methanol. The eluent was concentrated and extracted with chloroform to remove impurities. The resulting aqueous extracts were used for purification by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and methanol (B) with 0.1% formic acid using a gradient of 0-20 min 25% B; 25-32 min 100% B; 32-39 min 25% B. For **2,** fractions containing the target compound were combined and further purified by HPLC with a semi-preparative reverse-phase column (Kinetex 5 um C18 100 Å, 250 x 10.0 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-83 min 5-10% B; 83-90 min 100% B; 90-97 min 5% B. For **7** and **9,** fractions containing the target compounds were combined

and further purified by HPLC with a semi-preparative reverse-phase column (Kinetex 5 um C18 100 Å, 250 x 10.0 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-60 min 5-8.5% B; 60-67 min 100% B; 67-74 min 5% B. For **10,** fractions containing the target compound were combined and further purified by HPLC with a semi-preparative reverse-phase column (Kinetex 5 um C18 100 Å, 250 x 10.0 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-85 min 5-10% B; 85-92 min 100% B; 92-99 min 5% B.

For isolation of **3** and **4**, *A. nidulans* transformants expressing AnkA and AnkAB, respectively, were grown on 4 L of solid CD media for 4 days at 28 °C and then extracted with acetone. For isolation of **13**, *amaA* was cloned into the yeast expression vector XW55[4] through yeast homologous recombination. In this plasmid, expression of *amaA* was under the control of the *ADH2* promoter. *S. cerevisiae* JHY651 expressing AmaA was inoculated in 80 mL of dropout media for 24 hours. The starter culture was used to inoculate 4 L of YPD, and the cells were shaken at 28 °C, 250 rpm for 48 hours. The cell pellet was harvested by centrifugation and extracted with acetone. The crude extracts were absorbed with 3 g of Celite, which was purified with the CombiFlash system (Teledyne) using a 100 g HP C18 column (RediSepR_f) and reverse phase gradient elution with water (A) and methanol (B) using a gradient of 0-5 min 5% B; 5-125 min 5-100% B; 125-135 min 100% B. Fractions containing **4** were subject to purification by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-17 min 20% B; 17-24 min 100% B; 24-31 min 20% B. A second HPLC step was performed with a semi-preparative reverse-phase column (Cosmosil PBr,

10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 0% B; 5-65 min 0-20% B; 65-72 min 100% B; 72-79 min 0% B. For **13**, fractions containing the target compound from the CombiFlash were combined and purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-25 min 5-95% B; 25-32 min 95% B; 32-39 min 5% B.

For isolation of **11** and **12**, *A. nidulans* transformants expressing PthA and AteA, respectively, were grown on 4 L of solid CD media for 4 days at 28 °C and then extracted with methanol. The crude extracts were absorbed with 3 g of Celite, which was purified with the CombiFlash system (Teledyne) using a 100 g HP C18 column (RediSepR_f) and reverse phase gradient elution with water (A) and methanol (B) using a gradient of 0-40 min 5% B; 40-55 min 100% B. For **11**, fractions containing the target compound were combined and purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 0% B; 5-20 min 0-1% B; 20-27 min 100% B; 27-34 min 0% B. The compound was purified with two more rounds of HPLC with the same column and solvent system using a gradient of 0-20 min 0% B; 20-27 min 100% B; 27-34 min 0% B. For **12**, fractions containing the target compound from the CombiFlash were combined and purified by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-8 min 5% B; 8-15 min 100% B; 15-22 min 5% B. The compound was further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using

a gradient of 0-5 min 0% B; 5-20 min 0-4% B; 20-27 min 100% B; 27-34 min 0% B.

For isolation of **5** and **6**, an *A. nidulans* transformant expressing AnkABCDF was grown on 4 L of solid CD media for 3 days at 28 °C and then extracted with methanol. The extracts were concentrated and the methanol was removed by a rotary evaporator. The pH of the aqueous extracts was lowered to 4, and the extracts were mixed with HP20 resin for 1 hour. The resin mixture was then poured through a column, and the flowthrough was discarded. The resin was washed with water and 10% methanol with 0.1% formic acid. The compounds were eluted from the resin with 30% methanol. The eluent was concentrated and extracted with chloroform to remove impurities. The resulting aqueous extract was used for purification by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-40 min 5-13.8% B; 40-47 min 100% B; 47-54 min 5% B. Fractions containing **5** were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 2% B; 5-105 min 2-12% B; 105-112 min 100% B; 112-117 min 2% B. Fractions containing **6** were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 2% B; 5-50 min 2-13.7% B; 50-57 min 100% B; 57-64 min 2% B.

For isolation of **14**, an *A. nidulans* transformant expressing AvaA was grown on 4 L of solid CD media for 4 days at 28 °C and then extracted with methanol. For isolation of **15** and **16**, an *A. nidulans* transformant expressing AvaABC was grown on 4 L of solid CD media for

3 and 6 days, respectively, at 28 °C and then extracted with methanol. For isolation of **17** and **18**, an *A. nidulans* transformant expressing AvaA-D was grown on 4 L of solid CD media for 3 days at 28 °C and then extracted with methanol. The extracts were concentrated by a rotary evaporator. The aqueous extracts were mixed with HP20 resin for 1 hour. The resin mixtures were then poured through a column, and the flowthrough was discarded. The resin was washed with water and 25% and 50% methanol. The compounds were eluted from the resin with 75% methanol. The eluent was concentrated and extracted with chloroform to remove impurities. The resulting aqueous extracts were used for purification by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and methanol (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-45 min 5-19% B; 45-52 min 100% B; 52-59 min 5% B. For **14** and **15**, fractions containing the target compounds were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 6% B; 5-65 min 6-13% B; 65-72 min 100% B; 72-79 min 6% B. For **16**, fractions containing the target compound were purified by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-25 min 5% B; 25-32 min 100% B; 32-39 min 5% B. For **17**, fractions containing the target compound were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 2% B; 5-65 min 2-20% B; 65-72 min 100% B; 72-79 min 2% B. For **18**, fractions containing the target compound were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID

x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 2% B; 5-65 min 2-14% B; 65-72 min 100% B; 72-79 min 2% B.

The stereochemistry of **14** was determined by Marfey's analysis and proton NMR spectroscopy. For Marfey's analysis, 0.2 mg of **14** was hydrolyzed with 500 uL of 6 M HCl with the addition of 1% beta-mercaptoethanol (1-s2.0-0003269787901321-main) to prevent degradation of tryptophan for 1 hour. Standards of L-Trp and D-Trp and hydrolyzed **14** were derivatized with Marfey's reagent (acs.jnatprod.1c00414.pdf). The samples were analyzed on the QTOF (Figure S), indicating **14** contains L-Trp. Since the proton NMR spectra of **14** matched that of L-L or D-D cyclo-Arg-Trp (cite), we concluded **14** is composed of L-Arg and L-Trp.

For isolation of **19**, an *A. nidulans* transformant expressing AvaA-D was grown on 8 L of solid CD media for 2 days at 28 °C and then extracted with acetone. The extracts were concentrated by a rotary evaporator. The aqueous extract was mixed with HP20 resin for 1 hour. The resin mixture was then poured through a column, and the flowthrough was discarded. The resin was washed with water and the target compound was eluted with 25% acetonitrile. The eluent was concentrated and extracted with chloroform to remove impurities. The resulting aqueous extract was used for purification by HPLC with a semi-preparative reverse-phase column (Cosmosil 5C18-AR-II, 10ID x 250 mm) with water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 5% B; 5-45 min 5-15% B; 45-52 min 100% B; 52-59 min 5% B. Fractions containing the target compound were further purified by HPLC with a semi-preparative reverse-phase column (Cosmosil PBr, 10ID x 250 mm) with

water (A) and acetonitrile (B) with 0.1% formic acid using a gradient of 0-5 min 1% B; 5-125 min 1-15% B; 125-132 min 100% B; 132-139 min 1% 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).

8.17 Preparation of yeast lysates containing AvaA and *in vitro* lysate assays

AvaA was cloned into the yeast expression vector XW55[167] through yeast homologous recombination. In this plasmid, expression of *avaA* was under the control of the *ADH2* promoter. Overexpression of AvaA was performed as follows:[168] the yeast strain JHY651 harboring the expression plasmid was grown overnight in 3 x 1 mL cultures of uracil dropout medium at 28 °C. The starter culture was used to inoculate 24 x 3 mL cultures of YPD. The cells were grown at 28 °C, 250 rpm for 24 hours. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM of Tris-HCl, 150 mM NaCl, pH 8.0, 10% glycerol). Cells were disrupted by beads (zirconia beads from RiboPure^{TM}-Yeast Kit) for 2 min with 30 s shaking intervals on ice. The lysate was centrifuged at 4 °C, 15,000 *g* for 10 minutes to pellet cell debris.

Lysate assays were performed in 50 mM Tris-HCl, 150 mM NaCl, pH 8.0 with a final volume of 50 µL. For the assays in Figure S, the lysates were first desalted using a spin desalting column (Zeba, 40K MWCO). The assays contained 0-2 mM ATP, 0-10 mM $MgCl_2$, 5 mM L-Arg, and 5 mM Trp-d5. For the assays in Figure 3A, lysates with the addition of 6 mM DTT with and without 0.3 µg/mL RNase A (Qiagen) were incubated at room temperature for 20 min. RNasin (Promega) was added to the RNase treated samples at a final concentration

of 20 U/mL, followed by incubation at room temperature for 5 min. The assays contained the treated lysates, 0.5 mM ATP, 10 mM $MgCl_2$, 5 mM L-Arg, and 5 mM Trp-d5, with and without 400 ng/μL of total RNA purified from yeast.[171] The reactions were incubated at room temperature for 18 hours, and an equal volume of methanol was added to quench the reactions. Samples were centrifuged and the supernatant was analyzed by QTOF as described previously.

8.18 Expression and purification of AnkD, AvaA, ArgRS, and TrpRS from *E. coli* BL21(DE3) and *in vitro* assays

The *ankD* gene was amplified with overhang primers from the cDNA of *A. nidulans* expressing AnkA-F. The *avaA* gene was amplified with overhang primers from the genomic DNA of *Aspergillus versicolor* dI-29. The *argRS* (yeast YDR341C) and *trpRS* (yeast WRS1) genes were amplified with overhang primers from the genomic DNA of *S. cerevisiae* JHY651. The expression vector pET28a was digested with NdeI/XhoI (NEB) for *N*-His tag expression and NcoI/XhoI (NEB) for *C*-His tag expression. The plasmids p2001 (*N*-His_6-*ankD*), p2002 (N-His_6-*avaA*), p2003 (argRS-C-His_6), and p2004 (trpRS-C-His_6) were constructed through Hifi assembly (NEB). The sequences of the assembled plasmids were confirmed by DNA sequencing.

Overexpression and subsequent protein purification of AnkD, AvaA, ArgRS, and TrpRS were performed as follows:[169] BL21(DE3) harboring the expression plasmid was grown overnight in 16 x 5 mL of LB medium with 50 μg/mL of kanamycin at 37 °C. The starter cultures were used to inoculate 4 × 1 L of fresh LB medium and shaken at 37 °C until the optical density

at 600 nm (OD_{600}) reached 0.8. 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 Na_2HPO_4, 150 mM of NaCl, 10% glycerol, pH 8.0), and lysed on ice by sonication. The lysate was centrifuged at 14,000 *g* for 15 min at 4 °C to remove cellular debris. Purification of the recombinant His_6-tagged proteins using affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified proteins were concentrated and exchanged into storage buffer (50 mM Na_2HPO_4, 150 mM NaCl, 10% glycerol, pH 8.0) by using Centriprep filters (Amicon). Purified proteins were analyzed by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to measure protein concentration. Aliquots of purified enzymes were flash frozen and stored at -80 °C. For the assays in Figure 3D, AvaA and ArgRS were further purified by FPLC with a size exclusion column (HiLoad™ 16/600 Superdex™ 75 pg) with a flow rate of 1 mL/min for 120 and 240 min, respectively.

Enzyme assays of AnkD were performed in 50 mM Tris-HCl (pH 8.0) buffer with a final volume of 50 µL. The assays contained ~100 µM of compound **4**, **5**, or **6**, 100 µM PLP, 1 mM O-acetyl homoserine, and 4 µM of purified AnkD. The reactions were incubated at room temperature for 30 min to 4 hours, and two volumes of methanol were added to quench the reactions. Samples were centrifuged and the supernatant was analyzed by QTOF as described previously.

For the assays in Figure 3B, total yeast RNA was deacylated by incubating 160 µg of total RNA in 50 µL of 1 M Tris (pH 8.8) with RNasin (20 U/µL) at 37 °C for 3 hours.[172]

Deacylated RNA was recovered by ethanol precipitation, and the pellet was washed with 75% ethanol before resuspension in 50 mM Na_2HPO_4, 150 mM NaCl buffer (pH 8.0). Enzyme assays with AvaA were performed in 50 mM Na_2HPO_4, 150 mM NaCl buffer with a final volume of 20 µL. The assays contained 3 mM $MgCl_2$, 2.5 mM ATP, 0.8 mM L-Arg, 0.8 mM Trp-d5, 400 ng/µL deacylated total yeast RNA, 0.3 µM TrpRS, 0.4 µM ArgRS, and 1 µM AvaA. The reactions were incubated at room temperature for 18 hours, and an equal volume of methanol was added to quench the reactions. Samples were centrifuged and the supernatant was analyzed by QTOF as described previously.

For the assays in Figure 3C, yeast tRNA was purified from yeast total RNA by neutral RNA polyacrylamide gel electrophoresis.[173] The tRNA band was cut out of the gel and crushed followed by overnight incubation in 3 volumes of 0.3 M NaCl at 4 °C with constant agitation to elute the tRNA. The samples were centrifuged, and the supernatant was collected. The tRNA was recovered by ethanol precipitation, and the pellet was resuspended in 50 mM Na_2HPO_4, 150 mM NaCl buffer (pH 8.0). To confirm the identity of the tRNA band, RT-PCR was performed using Superscript III (Invitrogen) and specific primers that anneal to a Trp tRNA containing an intron. The Trp tRNA cDNA product was amplified by two rounds of overhang PCR, and the spliced sequence was confirmed by Sanger sequencing. Enzyme assays with AvaA were performed in 50 mM Na_2HPO_4, 150 mM NaCl buffer with a final volume of 35 µL. The assays contained 3 mM $MgCl_2$, 2.5 mM ATP, 0.8 mM L-Arg, 0.8 mM L-Trp, 50 ng/µL total yeast RNA or 16 ng/ µL purified yeast tRNA, and 3 µM each of purified TrpRS, ArgRS, and AvaA. For the assays in Figure 3D, L-Trp was replaced with 0.8 mM of 5-methoxy L-Trp, 2-methyl DL-Trp, 4-methyl DL-Trp, 5-methyl-DL Trp, 7-methyl DL-Trp, 4-fluoro DL-Trp, 5-fluoro

DL-Trp, 6-fluoro DL-Trp, 4-bromo L-Trp, 5-bromo DL-Trp, 6-bromo DL-Trp, 7-bromo DL-Trp, 5-chloro DL-Trp, 6-chloro DL-Trp, 5-nitro DL-Trp, 5-cyano DL-Trp, or 5-hydroxy L-Trp, and 30 ng/μL total yeast RNA was used in each reaction. The reactions were incubated at room temperature for 18 hours, and an equal volume of methanol was added to quench the reactions. Samples were centrifuged and the supernatant was analyzed by QTOF as described previously.

9. APPENDICES

9.1 Supplementary information for Section 3

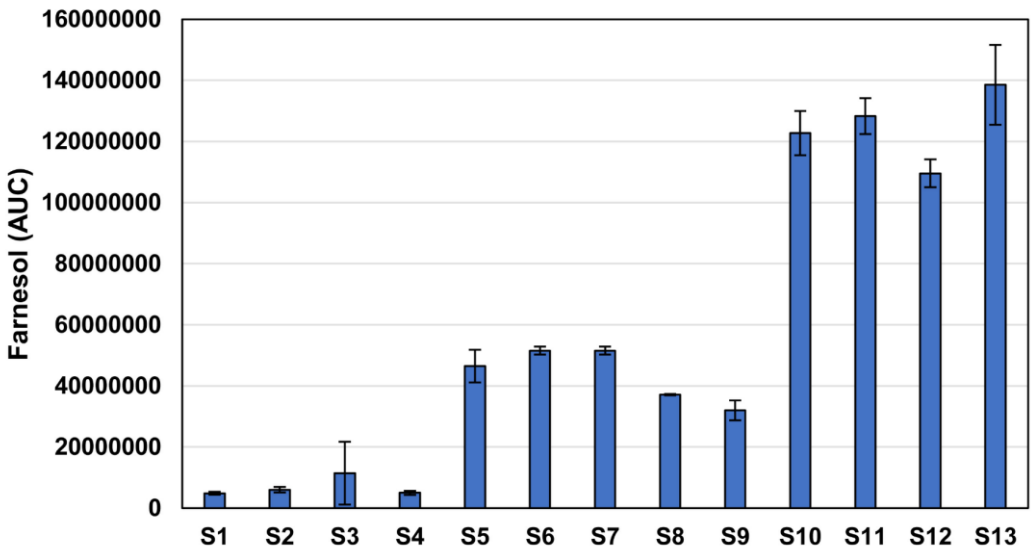


Figure S1. Relative farnesol production in **S1**-**S13** with combinatorial strain modifications to improve cytosolic geraniol titer. Samples were analyzed by GC-MS after 48 hours of culturing. Relative farnesol production was measured by area under the curve (AUC).

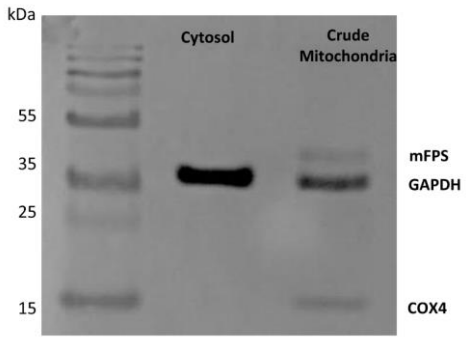


Figure S2. Western blot of cytosolic and mitochondrial fractions of **S16** harboring plasmids with mFPS with an *N*-terminal cox4 tag and *C*-terminal FLAG tag (p14) and GAPDH and COX4 with *C*-terminal FLAG tags (p15). The COX4 band (mitochondrial control) is only present in the mitochondrial fraction, which shows the mitochondria were successfully isolated. The GAPDH band (cytosolic control) is present in both the cytosolic and mitochondrial fractions, indicating some cytosolic contamination in the mitochondrial fraction. This may be due to the fact that only the crude mitochondria were isolated without further purification. However, the mFPS band is only present in the mitochondrial fraction, which demonstrates the mFPS completely localized to the mitochondria.

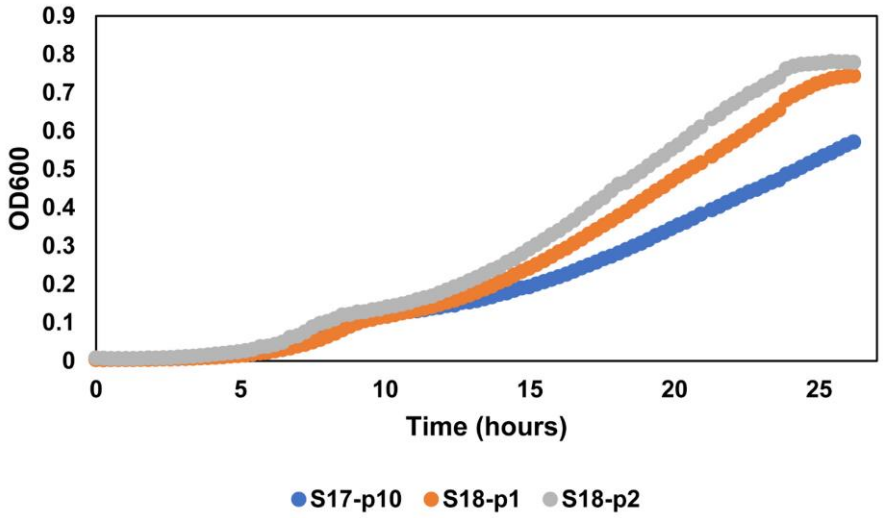


Figure S3. Growth curves of **S17** with a 2 μm mitochondrial GES plasmid (p10) (**S17-M**), **S18** with an empty 2 μm plasmid (p1), and **S18** with an empty CEN/ARS plasmid (p2) in SD 0.2% glucose 1.8% galactose uracil dropout media. Expression of mitochondrial GES on a 2 μm plasmid (**S17-p10**) led to growth defects compared to strains with the mitochondrial GES integration (**S18-p1**, **S18-p2**), suggesting high copy episomal expression of mitochondrial GES leads to cell stress.

	S17-p10	**S18-p1**	**S18-p2**

Average growth rate ($ln[OD_{600}]/h$)	0.108 +/- 0.002	0.14 +/- 0.001	0.156 +/- 0.006
Average doubling time (h)	6.33 +/- 0.13	4.84 +/- 0.02	4.43 +/- 0.18

Table S1. Average growth rates and doubling times of **S17-p10**, **S18-p1**, and **S18-p2**. Strains were grown as described in Fig. S3. Growth rates were determined in the exponential region of the galactose consumption phase. **S17-p10** had a slower growth rate and longer doubling time compared to **S18-p1** and **S18-p2**. This further demonstrates high copy episomal expression of mitochondrial GES results in slower growth compared to strains with the mitochondrial GES integration. Growth measurements were taken in biological triplicate.

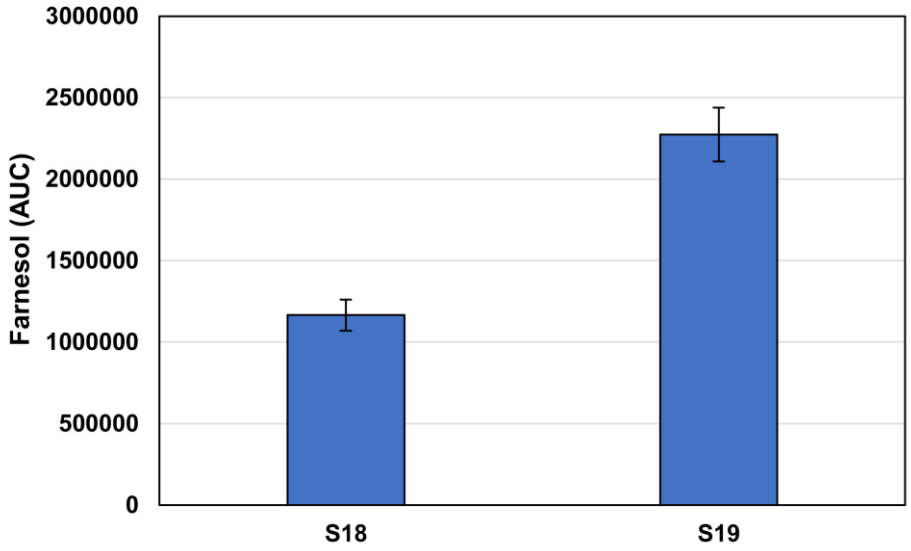


Figure S4. Relative farnesol production in **S18** and **S19** (mitochondrial and cytosolic GES integrations). Samples were analyzed by GC-MS after 48 hours of culturing. Relative farnesol production was measured by area under the curve (AUC).

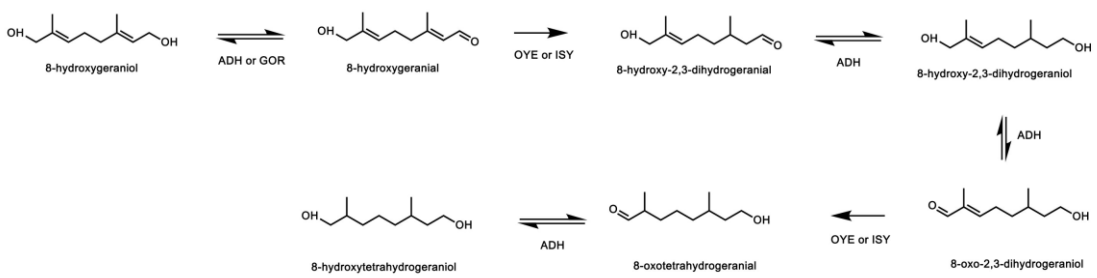


Figure S5. Shunt pathway from 8-hydroxygeraniol to 8-hydroxytetrahydrogeraniol by yeast's native ADHs or heterologous expression of GOR coupled with endogenous OYEs or heterologous expression of ISY.

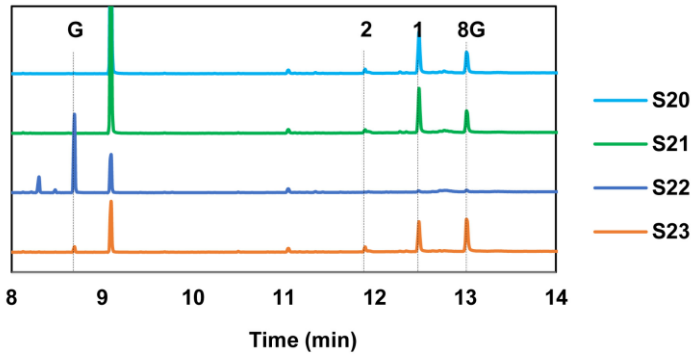


Figure S6. 8-hydroxygeraniol production in **S20**-**S23** (CrG8H integrations with single OYE deletion). Samples were analyzed by GC-MS after 48 hours of culturing. The total ion count (TIC) is shown. Peaks are labeled as G: geraniol; 8G: 8-hydroxygeraniol; 1: 8-hydroxy-2,3-dihydrogeraniol; 2: 8-hydroxytetrahydrogeraniol.

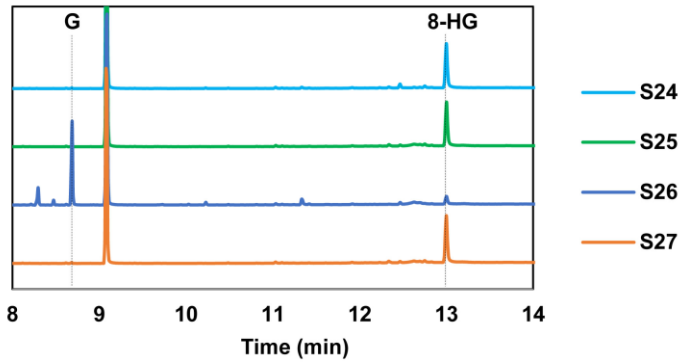


Figure S7. 8-hydroxygeraniol production in **S24**-**S27** (CrG8H integrations with double OYE knockouts). Samples were analyzed by GC-MS after 48 hours of culturing. The total ion count (TIC) is shown. Peaks are labeled as G: geraniol; 8-HG: 8-hydroxygeraniol.

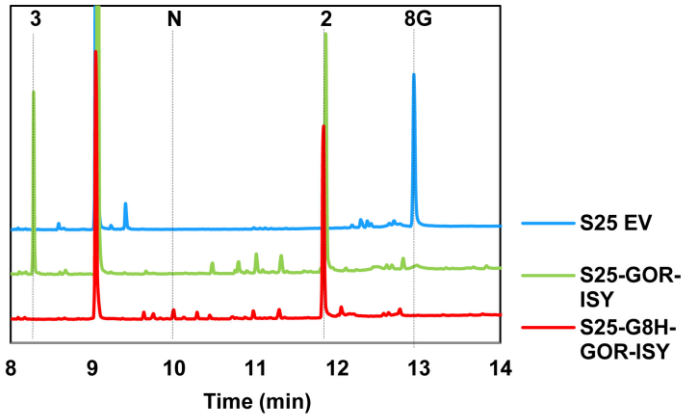


Figure S8. Shunt product accumulation and nepetalactol production in **S25** with empty vectors (p2 and p3), **S25** with GOR/ISY plasmid (p12), and **S25** with G8H plasmid (p13) and GOR/ISY plasmid (p12). Samples were analyzed by GC-MS after 48 hours of culturing. The total ion count (TIC) is shown. Peaks are labeled as 8G: 8-hydroxygeraniol; N: nepetalactol; 2: 8-hydroxytetrahydrogeraniol; 3: citronellol.

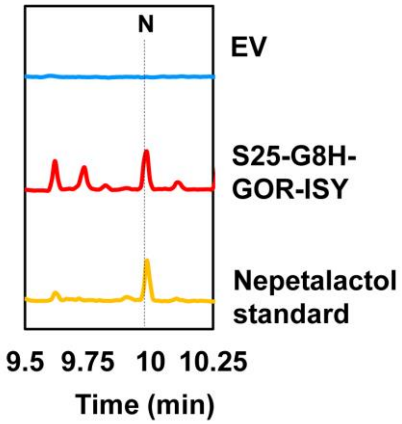


Figure S9. Close up of GC-MS traces of S25 with empty vectors (p2 and p3), S25 with G8H plasmid (p13) and GOR/ISY plasmid (p12) after 48 hours of culturing, and nepetalactol standard. The total ion count (TIC) is shown. The nepetalactol peak is labeled as N.

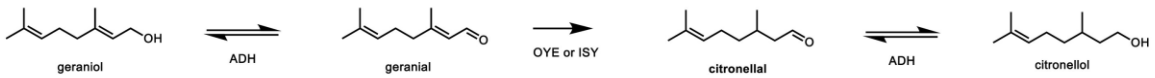


Figure S10. Shunt pathway from geraniol to citronellol by yeast's native ADHs coupled with endogenous OYEs or heterologous expression of ISY.

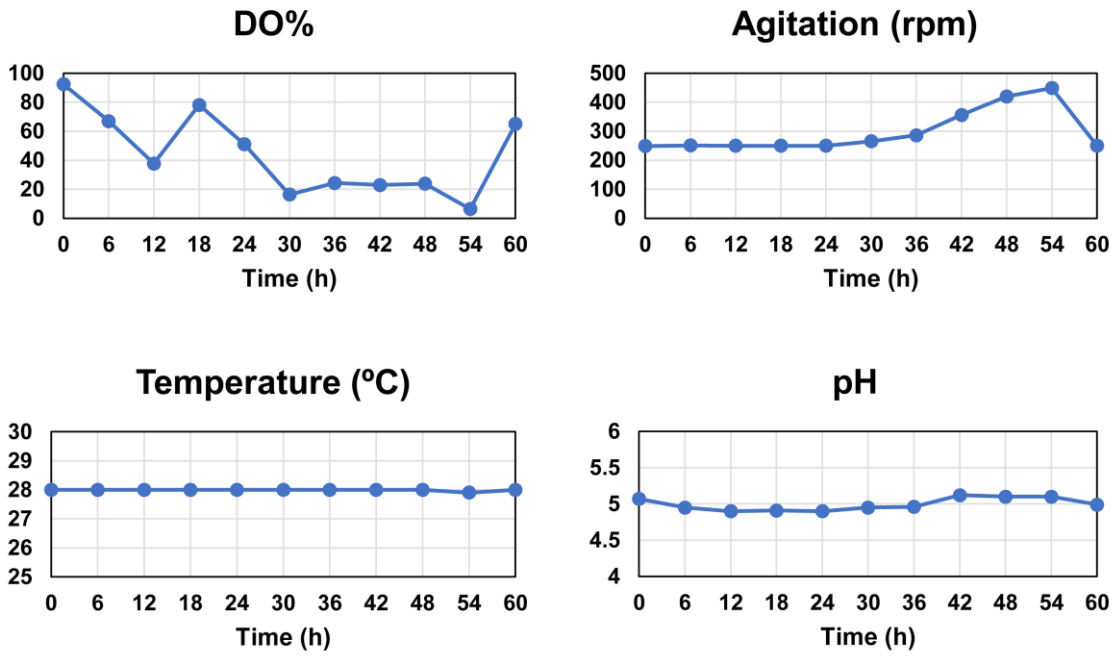


Figure S11. Additional data for fed-batch fermentation of **S25**.

Table S2. Primers used in this work.

Category	Primer name	Sequence	Amplicon
p5 assembly	TEF1p_obGES F	TAGCAATCTAATCTAAGTTTTAATTACAAAACTAGTATGCCATTATCTTCAACTCCTTTG	ObGES
p5 assembly	obGES cyc1t R	ATGTAAGCGTGACATAACTAATTACATGACTCGAGTTATTGAGTGAAAAACAATGCATCG	ObGES
p6 assembly	PGK_AgGPPS2_F	AAGTAATTATCTACTTTTTACAACAAATATATGTTTGATTTCAACAAGTACATGGATTCC	AgGPPS
p6 assembly	AgGPPS2_cyc_R	AATGTAAGCGTGACATAACTAATTACATGATCAATTTTGTCTGAATGCCACGTAATCTGC	AgGPPS
p6 assembly	PGK1p front R	ATATTTGTTGTAAAAAGTAGATAATTACTTCC	PGK1p URA3 2um CYC1t
p6 assembly	cyc1 F	TCATGTAATTAGTTATGTCACGCTTAC	PGK1p URA3 2um CYC1t
p7 assembly	pgk1 erg20* F	TCAAGGAAGTAATTATCTACTTTTTACAACAAATATATGGCTTCAGAAAAAGAAATTAGG	ERG20*K197G
p7 assembly	erg20* cyc1 R	TGAATGTAAGCGTGACATAACTAATTACATGACTATTTGCTTCTCTTGTAAACTTTGTTC	ERG20*K197G
p8 assembly	pgk_mFPS_F	AAGTAATTATCTACTTTTTACAACAAATATATGCATAAATTTACTGGTGTCAATGCCAAG	mFPS*N144W
p8 assembly	mFPS_cyc_R	AATGTAAGCGTGACATAACTAATTACATGATCATTTCTGGCGTTTGTAGATCTTCTGTGC	mFPS*N144W
p9 assembly	pgk_vrtD_1_F	ATCATCAAGGAAGTAATTATCTACTTTTTACAACAAATATATGGCCACTAGCACTACCAC	vrtD
p9 assembly	vrtD_1_2_R	CGGTTGCACTTGCCACCGAGGGTATTGTGATTTAAAGTACTTTCAAACCGATCCCACACG	vrtD
p9 assembly	vrtD_2_F	ACTTTAAATCACAATACCCTCGGTGGCAAGTG	vrtD
p9 assembly	vrtD_2R	GTAAATGGCCCGTCCCGCGAAATACTTC	vrtD
p9 assembly	vrtD_2_3F	TGCTGAAGAAGTATTTCGCGGGACGGGCCATTTACC	vrtD

		TCCCTGTGGTCGACTTATTCCATG	
p9 assembly	vrtD_3_cyc_R	AGGGCGTGAATGTAAGCGTGACATAACTAATTACATGACTACTTTCGCCGCTGATAGATC	vrtD
p10 assembly	cox4_GES_F	ACTTTGTGTAGCTCTAGATATCTGCTTCAGGGATCCATGCCATTATCTTCAACTCCTTTG	GES
p10 assembly	GES_CYC1_R	AATGTAAGCGTGACATAACTAATTACATGACTCGAG	GES
p10 assembly	pxp218_gal1_F	GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTCGGATTAGAAGCCGCCGAG	GAL1p NCox4
p10 assembly	cox4_rev	GGATCCCTGAAGCAGATATC	GAL1p NCox4
p10 assembly	pxp218_R	ATATTATTGAAGCATTTATCAGGGTTATTG	URA3 2um CYC1t
p10 assembly	cyc1 F	TCATGTAATTAGTTATGTCACGCTTAC	URA3 2um CYC1t
p11 assembly	gal_GES_F	ATATACCTCTATACTTTAACGTCAAGGAGAAAAAACATGCCATTATCTTCAACTCCTTTG	GES
p11 assembly	gal_R	GTTTTTTCTCCTTGACGTTAAAGTATAG	GAL1p
p12 assembly	GAL VS7 R	TTTCAAAAATTCTTACTTTTTTTTTTGGATGG	GAL1-10p
p12 assembly	ISY-GAL VS7 F	CTTTGCGTCCATCCAAAAAAAAAGTAAGAATTTTTGAAAATGTCCTGGTGGTGGAAAAGG	ISY CPS1t
p12 assembly	CPS1 VS7 R	AGATGCTTCGTTCATAGCCATGCCTTCACATATAGTATTTGACACTTGATTTGACACTTC	ISY CPS1t
p12 assembly	2m VS7 F	CAAATCAAGTGTCAAATACTATATGTGAAGGCATGGCTATGAACGAAGCATCTGTGCTTC	2um URA3
p12 assembly	ori VS7 R	CTTCGGAAAATACGATGTTGAAAATGCTTGGACATCTTCGCGTTTTTCCATAGGCTCCGC	2um URA3
p12 assembly	PRM9 VS7 F	CGGAGCCTATGGAAAAACGCGAAGATGTCCAAGCATTTTCAACATCGTATTTTCCGAAGC	GOR PRM9t
p12 assembly	GOR-GAL VS7 R	ATATACCTCTATACTTTAACGTCAAGGAGAAAAAACA	GOR PRM9t

		TGACTAAAACTAATTCTCCAGCC	
p12 assembly	cps1t CEN/ARS F	TCAAATCAAGTGTCAAATACTATATGTGAAGGCATGGCTATGGTAACTTACACGCGCCTC	CEN/ARS
p12 assembly	CEN/ARS ura3 R	GATGCGGCCAGCAAAACTAAAAAACTGTATTATAAGTAAGATTACGAATTCATCACGTGC	CEN/ARS
p13 assembly	galp_G8H F	TTAACGTCAAGGAGAAAAAACCCCGGATCTAAAACAATGGACTACCTGACCATTATTTTG	G8H
p13 assembly	G8H cyc1 R	CGGATCTTAGCTAGCCGCGGTACCAAGCTTACTCGAGTTACAAAGTAGATGGAACAGCTC	G8H
p14 assembly	pxp318 mFPS F	TGAATACTCATACTCTTCCTTTTTCAATATTATTGAATTCGAATTTTCAAAAATTCTTAC	GAL1p mFPS
p14 assembly	mFPS flag R	TCAGATCTTATCGTCGTCATCCTTGTAATCTTTCTGGCGTTTGTAGATCTTC	GAL1p mFPS
p14 assembly	flag cyc1 F	ACGCCAGAAAGATTACAAGGATGACGACGATAAGATCTGACTCGAGTAAGCTTGGTACCG	cFLAG CYC1t
p14 assembly	cyc1 pxp318 R	CATGAGACAATAACCCTGATAAATGCTTCAATAATATTGGCCCTTCGAGCGTCCCAAAAC	cFLAG CYC1t
p14 assembly	pxp318 F	GGCCAATATTATTGAAGCATTTATCAG	CEN/ARS URA3 CYC1t cFLAG
p14 assembly	pxp318 R	TAATATTGAAAAAGGAAGAGTATGAGTATTCAAC	CEN/ARS URA3 CYC1t cFLAG
p15 assembly	gal gapdh F	CTTTAACGTCAAGGAGAAAAAACCCCGGATCTAAAACAATGATCAGAATTGCTATTAACG	GAPDH
p15 assembly	gapdh flag R	AGCTTACTCGAGTCAGATCTTATCGTCGTCATCCTTGTAATCAGCCTTGGCAACATATTC	GAPDH
p15 assembly	cyc1 pxp322 F	AAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGGCCAATATTATTGAAGCATTTATC	LEU2
p16 assembly	pxp322 SpeI R	TAACTTCGTATAATGTATGCTATACGAAGTTATACTAGTGGATTTTCTTAACTTCTTCGG	LEU2

p16 assembly	pxp322 NotI F	GTATAACTTCGTATAGCATACATTATACGAAGTTATGCGGCCGCcccGGGTACCGAGCTC	CEN/ARS GAL1p
p15 assembly	gal1_10 R	TGTTTTAGATCCGGGGTTTTTTCTC	CEN/ARS GAL1p
p15 assembly	gal_cox4 F	TACTTTAACGTCAAGGAGAAAAAACCCCGGATCTAAAACAATGCTTTCACTACGTCAATC	COX4
p15 assembly	cox4_flag R	CTTACTCGAGTCAGATCTTATCGTCGTCATCCTTGTAATCGTGATGGTGGTCATCATTTG	COX4
p15 assembly	SpeI cyc1 F	ATATCAAACTAGTCTTCGAGCGTCCCAAAAC	GAL1p COX4 cFLAG CYC1t
p15 assembly	galp_NotI R	AAATATAGCGGCCGCTTTCAAAAATTCTTACTTTTTTTTTGGATG	GAL1p COX4 cFLAG CYC1t
S1 construction	YPRC Ty1-2 – pXP F	CGAGAGAACTTCTAGTATATCTGTATACATAATATGATAGCCTTTACCAACAATGGACAGTGCCAAGCTTGCATGCC	TEF1p Cas9 ADH2t LEU2
S1 construction	YPRCTy1-2 ADH2t R	CCGGTATTTATTTCTTTGCAACCAAAATATGGATATCGAGATGTATTTGATGAATAATTATGAGAAATATCGAGGGACTCG	TEF1p Cas9 ADH2t LEU2
S2 construction	URA-TEF1 F	ATGGTGAAGGATAAGTTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCAGGGTCCCCGCGAATCCTTACATCACACC	TEF1p CrGES CYC1t
S2 construction	URA3-CYC1t R	TTTTCTTTCCAATTTTTTTTTTTTCGTCATTATAGAAATCATTACGACCGAGATTCCCGGTGGCCGCAAATTAAAGCCTTCG	TEF1p CrGES CYC1t
S3 construction	ROX1 - TEF1p F	CATTATTCCAGAAAATACTAATACTTCTTCACACAAAAGAACGCAGTTAGACAATCAACACCGCGAATCCTTACATCACACC	TEF1p ERG20*(f)ObGES CYC1t
S3 construction	ROX1-CYC1t R	TTAGTTAAAGGGAATATAGTATAATATAATATAACGGAAAGAAGAAATGGAAAAAAAAAAGGCCGCAAATTAAAGCCTTCG	TEF1p ERG20*(f)ObGES CYC1t
S3 construction	ROX1 gRNA R (2)	GCACCACCGACTCGGTG	gRNA ROX1
S3 construction	ROX1-gRNA F	TGCGCATGTTTCGGCGTTC	gRNA ROX1

S4 construction	OYE2_mFPS_DonorFor	TCCAGATATAGAATAAATCATCATATTAAGCTAAATATAGACGATAATATAGTATCGATACCGCGAATCCTTACATCACACC	TEF1p mFPS CYC1t
S4 construction	OYE2_mFPS_DonorRev	TTCATTAATTATATAAATTAGAAGAAAAAGAAATGGTGCTACAAAGTACGGTTAACACTATGGCCGCAAATTAAAGCCTTCG	TEF1p mFPS CYC1t
S5 construction	Erg9 Cassette F	ACCGGATTCAGTCGTCACTCATGGTGATTTCTCAC	donor DNA for ERG9p truncation
S5 construction	Erg9 Cassette R	GCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGC	donor DNA for ERG9p truncation
S5 construction	Erg9 Target F	TGATCTTTTCCACTGCACTTTGCATGTTTTAGAGCTAGAAATAGCAAG	gRNA ER9p
S5 construction	Erg9 Target R	AAACATGCAAAGTGCAGTGGAAAAGATCATTTATCTTTCACTGCG	gRNA ER9p
S6 construction	BTS1 del 1	GAACAAAGCGTTTACGAGTCTGGAAAATCAAATAAATTGATCAATCAAATTAGTGGAGGA	donor DNA for BTS1 deletion
S6 construction	BTS1 del 2	TCCTCCACTAATTTGATTGATCAATTTATTTGATTTTCCAGACTCGTAAACGCTTTGTTC	donor DNA for BTS1 deletion
S6 construction	gBTS1-pCB32 4 v2	GTGAAAGATAAATGATCCAGATTAACGTTGAGACTCAGTTTTAGAGCTAGAAATAGCAAG	gRNA-BTS1 2um
S6 construction	pCB32 3	GGAGGTACTAGACTACCTTTCATCCTACATAAATAGACG	gRNA-BTS1 2um
S6 construction	pCB32 1	GAGGTCGAGTTTAGATGCAAGTTCAAGGAG	HygR gRNA-BTS1 2um
S6 construction	gBTS1-pCB32 2 v2	ACTGAGTCTCAACGTTAATCTGGATCATTTATCTTTCACTGCGGAGAAGTTTCGAACGCC	HygR gRNA-BTS1 2um
S7 construction	BTS1u-M F	CTACTAATAGAAAGAGAACAAAGCGTTTACGAGTCTGGAAAATAGCGTTTCTGGGTGAGC	TEF1p IDI1 CYC1t
S7 construction	BTS1d-CYC1t R	TCTGACTATCTTCCTCCACTAATTTGATTGATCAATTTATTCCGCAAATTAAAGCCTTCG	TEF1p IDI1 CYC1t
S8 construction	YPL062W del 1	CACATACGACACTGCCCCTCACGTAAGGGCCACCGA	donor DNA for YPL062W deletion

		CCATGTGGGCAAATTCGTAATAAA	
S8 construction	YPL062W del 2	TTTATTACGAATTTGCCCACATGGTCGGTGGCCCTTACGTGAGGGGCAGTGTCGTATGTG	donor DNA for YPL062W deletion
S8 construction	gYPL062W-pCB32 2	TTTCTAGCTCTAAAACTGGGGATGAGGCGAGCGCCGGATCATTTATCTTTCACTGCGGAG	KanMX gRNA-YPL062W 2um
S8 construction	gYPL062W-pCB32 4	TGATCCGGCGCTCGCCTCATCCCCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	gRNA-YPL062W 2um
S9 construction	YJL064W del 1	ACATAGGCGGAGTAAACTTCATTAGGGGGCTCACATCGAGCGAGCTATCCTGACGCTATT	donor DNA for YJL064W deletion
S9 construction	YJL064W del 2	AATAGCGTCAGGATAGCTCGCTCGATGTGAGCCCCCTAATGAAGTTTACTCCGCCTATGT	donor DNA for YJL064W deletion
S9 construction	gYJL064W-pCB30 2	TTTCTAGCTCTAAAACAGATGAACTCACGCTGTCGTGATCATTTATCTTTCACTGCGGAG	KanMX gRNA-YJL064W 2um
S9 construction	gYJL064W-pCB30 4	TGATCACGACAGCGTGAGTTCATCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	gRNA-YJL064W 2um
S10 construction	YPL062W tHMG1 u	GGAACTGCCGTCACATACGACACTGCCCCTCACGTAAGGGCCACGCTTTTTCAGTTCGAG	GPDp tHMG1 ADH1t
S10 construction	YPL062W tHMG1 d	GGAACTGCCGTCACATACGACACTGCCCCTCACGTAAGGGCCACGCTTTTTCAGTTCGAG	GPDp tHMG1 ADH1t
S11 construction	YJL064 HMG2 u	ATCGTTCACCACATAGGCGGAGTAAACTTCATTAGGGGGCACCGCGAATCCTTACATCAC	TEF1p HGM2* CYC1t
S11 construction	YJL064W HMG2 d	AACAAGAGAGAATAGCGTCAGGATAGCTCGCTCGATGTGAGCCGCAAATTAAAGCCTTCG	TEF1p HGM2* CYC1t
S13 construction	OYE3_TEF1 F	ACAACTGTAGTTCAGTATAGCGAAGTTTAAATTTAGAAGCCGCGAATCCTTACATCACAC	TEF1p IDI1 CYC1t
S13 construction	OYE3_CYC1 R	TATGAAAAATACATAACATCAATGTCTTTATTCATGATTGCAAATTAAAGCCTTCGAGCG	TEF1p IDI1 CYC1t

S13 construction	gOYE3 F	GTGAAAGATAAATGATCG GTATAACCTTCCGCGGAC AGTTTTAGAGCTAGAAATA GCAAG	gRNA-OYE3 2um
S13 construction	gOYE3 R	TTT CTA GCT CTA AAA CTG TCC GCG GAA GGT TAT ACC GAT CAT TTA TCT TTC ACT GCG GAG	HygR gRNA-OYE3 2um
S14 construction	URA3_M1_F2	CCTTTTTTGCGAGGCATA TTTATGGTGAAGGATAAG TTTTGCTTCGAGCGTCCC AAAACC	CYC1t ERG10 NCox4 GAL1-10p NCox4 ERG13 ADH1t
S14 construction	URA3_M1_R2	TTTTCGTCATTATAGAAAT CATTACGACCGAGATTCC CGGGAGCGACCTCATGCT ATACC	CYC1t ERG10 NCox4 GAL1-10p NCox4 ERG13 ADH1t
S15 construction	x2_cyc1_F	AAAGATCCTCTCATACCAT ATTAAGTAAATTGCCTCCA TTTCTTCGAGCGTCCCAA AACC	CYC1t tHMG1 NCox4 GAL1-10p NCox4 ERG12 ADH1t
S15 construction	adh1_x2_R	CTGAAAAAAGAAACGAGC GGAGGAATAGTATGATAA ATCTGAGCGACCTCATGC TATACC	CYC1t tHMG1 NCox4 GAL1-10p NCox4 ERG12 ADH1t
S15 construction	CB0121	ATGATCCCTGCGAGTTTC TCTGCCCGGTTTTAGAGC TAGAAATAGCAAGTTAAAA TAAGG	gRNA-X-2 2um
S15 construction	CB0122	CTCTAAAACCGGGCAGAG AAACTCGCAGGGATCATT TATCTTTCACTGCGGAG	HygR gRNA-X-2 2um
S16 construction	HO_cyc1	TGAAGCATGATGAAGCGT TCTAAACGCACTATTCATC ATTACTTCGAGCGTCCCA AAACC	CYC1t ERG10 NCox4 GAL1-10p NCox4 ERG13 ADH1t
S16 construction	adh1_HO	TTGAGATGGCGTATTTCT ACTCCAGCATTCTAGTTAA GAAGAGCGACCTCATGCT ATACC	CYC1t ERG10 NCox4 GAL1-10p NCox4 ERG13 ADH1t
S16 construction	HO gRNA 4	TGATCGGTCCAGTACTCG CAGGAAAGTTTTAGAGCT AGAAATAGCAAGTTAAAAT AAGGC	gRNA-HO 2um
S16 construction	HO gRNA 2	TTTCTAGCTCTAAAACTTT CCTGCGAGTACTGGACCG ATCATTTATCTTTCACTGC GGAG	HygR gRNA-HO 2um
S17 construction	rox1 cyc1 F	AATACTTCTTCACACAAAA GAACGCAGTTAGACAATC AACACTTCGAGCGTCCCA AAACC	CYC1t mFPS NCox4 GAL1-10p NCox4 IDI1 ADH1t
S17 construction	adh1 rox1 R	GAATATAGTATAATATAAT ATAACGGAAAGAAGAAAT	CYC1t mFPS NCox4 GAL1-10p

		GGAGAGCGACCTCATGCTATACC	NCox4 IDI1 ADH1t
S18 construction	YPRCTy1-2 SceI F URA3 gRNA	CCTTTACCAACAATGGAACCATCAAAGAAGGTTAATGTGGGAAGGTCATGAGTGCCAATG	URA3
S18 construction	SceI YPRCTy1-2 R2	AACCAAAATATGGATATCGAGATGTATTTGATGAATAATTTTTGTCGTCGCTGTTACACC	URA3
S18 construction	YPRCTy1-2 F2 URA3 gRNA	CGAGAGAACTTCTAGTATATCTGTATACATAATATGATAGCCTTTACCAACAATGGAACC	URA3
S18 construction	YPRCTy1-2 R	ACATCTGAATCATATCTCCGGTATTTATTTCTTTGCAACCAAAATATGGATATCGAGATG	URA3
S18 construction	iCas9-pCB30 1	TGAATCGAGTCCCTCGATATTTCTCATACTAGTTCTAGAGCGCGCCAACAAATATATTGC	KanMX gRNA-URA3
S18 construction	ori R	TTGAGCGTCGATTTTTGTGATGC	KanMX gRNA-URA3
S18 construction	pJB097 pCB30 F	TTTTGATGGGGCAGGGCGGAAATAGAGGATAGGATAAGCCGACATGGAGGCCCAGAATAC	gRNA-URA3 2um
S18 construction	iCas9-pCB30 3	AAGAGTAAAAAAGGAGTAGAAACATTTTGAAGCTATGCAGACGAAAGGGCCTCGTG	gRNA-URA3 2um
S18 construction	iCas9ec F	CTGCATAGCTTCAAAATGTTTCTACTCC	Cas9
S18 construction	iCas9ec R	TCTAGAACTAGTATGAGAAATATCGAG	Cas9
S18 construction	M13F*	GTTGTAAAACGACGGCCAGTG	SceI
S18 construction	M13R*	TCACACAGGAAACAGCTATGACC	SceI
S18 construction	YPRCTy1-2 gal1 F	ATATCTGTATACATAATATGATAGCCTTTACCAACAATGGACGGATTAGAAGCCGCCGAG	GAL1p NCox4 GES CYC1t
S18 construction	cyc1 YPRCTy1-2 R	CAAAATATGGATATCGAGATGTATTTGATGAATAATTGCAAATTAAAGCCTTCGAGCGTC	GAL1p NCox4 GES CYC1t
S18 construction	YPRCTy1-2 F	TTCTTATATTCTCGAGAGAACTTCTAGTATATCTGTATACATAATATGATAGCCTTTACC	GAL1p NCox4 GES CYC1t

S20 construction	OYE2 G8H F	ATTAAGCTAAATATAGACG ATAATATAGTATCGATAAT GGACTACCTGACCATTAT TTTG	G8H
S20 construction	G8H OYE2 R	GAAAAAGAAATGGTGCTA CAAAGTACGGTTAACACT ATTACAAAGTAGATGGAA CAGCTC	G8H
S20 construction	gOYE2 F	GTGAAAGATAAATGATCT GTTCTTCGGACCAGATAC CGTTTTAGAGCTAGAAAT AGCAAG	gRNA-OYE2 2um
S20 construction	gOYE2 R	TTTCTAGCTCTAAAACGGT ATCTGGTCCGAAGAACAG ATCATTTATCTTTCACTGC GGAG	HygR gRNA-OYE2
S21 construction	OYE2 galp F	CTAAATATAGACGATAATA TAGTATCGATATTTCAAAA ATTCTTACTTTTTTTTTGG ATG	GAL1p G8H CYC1t
S21 construction	cyc1t OYE2 R	TAGAAGAAAAAGAAATGG TGCTACAAAGTACGGTTA ACACTACTTCGAGCGTCC CAAAAC	GAL1p G8H CYC1t
S22 construction	OYE3 G8H F	ACTGTAGTTCAGTATAGC GAAGTTTAAATTTAGAAGA TGGACTACCTGACCATTA TTTTG	G8H
S22 construction	G8H OYE3 R	TGAAAAATACATAACATCA ATGTCTTTATTCATGATTT TACAAAGTAGATGGAACA GCTC	G8H
S23 construction	OYE3 galp F	GTTCAGTATAGCGAAGTT TAAATTTAGAAGTTTCAAA AATTCTTACTTTTTTTTTG GATG	GAL1p G8H CYC1t
S23 construction	cyc1t OYE3 R	GAATATGAAAAATACATAA CATCAATGTCTTTATTCAT GATTCTTCGAGCGTCCCA AAAC	GAL1p G8H CYC1t
S24 and S25 construction	OYE3 deletion F	GTTCAGTATAGCGAAGTT TAAATTTAGAAGAATCATG AATAAAGACATTGATGTTA TGTA	Donor DNA for OYE3 deletion
S24 and S25 construction	OYE3 deletion R	TACATAACATCAATGTCTT TATTCATGATTCTTCTAAA TTTAAACTTCGCTATACTG AAC	Donor DNA for OYE3 deletion
S26 and S27 construction	OYE2 deletion F	CTAAATATAGACGATAATA TAGTATCGATATAGTGTTA ACCGTACTTTGTAGCACC ATTT	Donor DNA for OYE2 deletion
S26 and S27 construction	OYE2 deletion R	AAATGGTGCTACAAAGTA CGGTTAACACTATATCGA	Donor DNA for OYE2 deletion

		TACTATATTATCGTCTATA TTTAG	

Table S3. Gene sequences

Gene name	Source	DNA sequence
CrGES	*Catharanthus roseus*	ATGTCTTTGCCACTGGCTACTCCATTGATCAAAGAC AATGAGTCGTTGATCAAGTTCTTGAGACAACCATTG GTCCTGCCACATGAAGTTGACGATTCTACCAAGAGA AGGGAGTTGTTGGAGAGAACTCGCAAAGAACTGGA ACTGAACGCTGAAAAACCCTTGGAGGCTCTGAAGAT GATCGATATTATTCAGCGCCTGGGTTTGTCCTACCA TTTCGAAGACGATATCAACTCCATTTTGACTGGTTTC AGCAACATCTCTTCCCAGACCCATGAGGACTTGTTG ACTGCTTCTTTGTGCTTCAGACTGTTGAGGCATAAT GGCCATAAGATCAACCCCGACATCTTCCAGAAGTTT ATGGACAACAATGGCAAGTTCAAAGACTCCTTGAAG GATGACACTTTGGGCATGCTGTCTTTGTACGAAGCC TCTTATTTGGGTGCTAATGGTGAAGAGATCTTGATG GAAGCCCAAGAATTCACCAAAACTCATCTGAAGAAC TCCTTGCCAGCTATGGCTCCATCTTTGTCCAAGAAA GTCTCTCAAGCTTTGGAGCAACCAAGGCATAGGAGA ATGTTGAGATTGGAAGCTAGAAGATTCATTGAGGAG TACGGTGCCGAAAATGATCACAACCCAGATTTGTTG GAACTGGCCAAGCTGGACTACAATAAGGTCCAATCT TTGCATCAGATGGAATTGTCTGAAATCACTCGCTGG TGGAAACAATTGGGACTGGTCGATAAATTGACCTTC GCTAGGGATCGTCCATTGGAATGTTTCTTGTGGACT GTTGGTTTGTTGCCAGAGCCAAAGTACTCTGGTTGC AGGATTGAGTTGGCTAAGACCATTGCTATTCTGTTG GTCATTGACGACATTTTCGACACTCACGGTACCCTG GATGAATTGTTGCTGTTTACCAACGCCATTAAGAGA TGGGATTTGGAGGCCATGGAGGACCTACCAGAATA CATGCGCATTTGCTATATGGCCTTGTACAACACCAC TAACGAAATCTGCTATAAGGTTCTGAAAGAGAACGG CTGGTCTGTCTTGCCATACCTGAAGGCTACCTGGAT CGACATGATCGAAGGCTTCATGGTTGAAGCCGAGT GGTTCAACTCTGATTATGTTCCCAATATGGAAGAATA TGTTGAAAACGGTGTCAGAACCGCTGGTTCTTACAT GGCCCTGGTCCACTTGTTCTTCCTGATTGGCCAAGG TGTTACCGAGGACAATGTCAAGTTGTTGATTAAACC CTACCCCAAATTGTTCTCTTCCTCTGGTCGCATCTTG AGATTATGGGACGACTTGGGCACTGCTAAAGAGGA ACAAGAAAGGGGTGACTTGGCTTCCTCCATTCAGTT GTTCATGAGGGAAAAGGAAATTAAATCCGAAGAGGA GGGTAGGAAGGGCATCTTGGAAATCATCGAAAACCT GTGGAAGGAGCTAAACGGTGAATTGGTTTACAGGG AAGAAATGCCATTGGCGATTATTAAGACCGCCTTCA ACATGGCTAGGGCCTCTCAGGTCGTTTACCAGCATG AAGAAGATACCTACTTTTCCTCCGTCGATAATTACGT TAAGGCTCTGTTCTTTACTCCATGCTTTTAA

ObGES	*Ocimum basilicum*	ATGCCATTATCTTCAACTCCTTTGATAAATGGTGACA ACTCACAAAGAAAGAATACTAGACAACACATGGAAG AATCCAGTTCTAAGAGAAGAGAATACTTGTTAGAAG AAACTACAAGAAAATTGCAAAGAAACGACACAGAAT CCGTAGAAAAATTGAAGTTAATCGATAACATCCAACA ATTGGGTATCGGTTACTACTTCGAAGACGCCATTAA TGCTGTTTTGAGAAGTCCATTCTCTACAGGTGAAGA AGATTTGTTTACTGCTGCATTGAGATTCAGATTGTTA AGACACAACGGTATCGAAATCTCACCTGAAATTTTCT TGAAATTCAAGGATGAACGTGGTAAATTCGACGAAT CCGATACTTTGGGTTTGTTATCATTGTACGAAGCATC TAATTTGGGTGTTGCAGGTGAAGAAATTTTGGAAGA AGCTATGGAATTTGCCGAAGCTAGATTGAGAAGATC ATTATCCGAACCAGCCGCTCCTTTACATGGTGAAGT TGCACAAGCCTTGGATGTACCAAGACACTTAAGAAT GGCAAGATTGGAAGCCAGAAGATTCATTGAACAATA CGGTAAACAATCTGATCATGACGGTGACTTGTTGGA ATTGGCTATCTTAGATTATAACCAAGTCCAAGCACAA CACCAAAGTGAATTGACTGAAATCATCAGATGGTGG AAAGAATTGGGTTTAGTTGACAAGTTGTCTTTTGGTA GAGATAGACCTTTGGAATGTTTCTTATGGACAGTCG GTTTGTTACCAGAACCTAAATACTCATCCGTTAGAAT AGAATTGGCTAAGGCAATCTCAATCTTGTTAGTTATA GATGACATCTTTGATACTTATGGTGAAATGGATGACT TGATATTATTCACCGACGCCATAAGAAGATGGGATT TGGAAGCTATGGAAGGTTTGCCAGAATACATGAAAA TCTGTTACATGGCTTTGTACAATACCACTAACGAAGT CTGCTATAAGGTTTTAAGAGACACTGGTAGAATCGT TTTGTTGAATTTGAAATCTACATGGATCGATATGATC GAAGGTTTTATGGAAGAAGCAAAATGGTTCAATGGT GGTTCAGCCCCTAAGTTGGAAGAATACATTGAAAAC GGTGTTTCCACAGCCGGTGCTTATATGGCATTTGCC CATATCTTTTTCTTAATAGGTGAAGGTGTTACTCACC AAAACAGTCAATTGTTTACACAAAAGCCATACCCTAA GGTATTCTCTGCAGCCGGTAGAATTTTGAGATTATG GGATGACTTAGGTACAGCCAAGGAAGAACAAGAAC GTGGTGACTTGGCATCATGCGTTCAATTGTTCATGA AGGAAAAGTCATTAACCGAAGAAGAAGCAAGATCCA GAATATTGGAAGAAATCAAAGGTTTGTGGAGAGATT TGAACGGTGAATTGGTATACAATAAGAACTTGCCAT TGAGTATCATTAAGGTCGCTTTGAACATGGCTAGAG CATCTCAAGTTGTATACAAGCATGACCAAGATACCT ACTTTAGTTCTGTAGACAATTATGTCGATGCATTGTT TTTCACTCAATAA
ERG20*K197G	*Saccharomyces cerevisiae*	ATGGCTTCAGAAAAAGAAATTAGGAGAGAGAGATTC TTGAACGTTTTCCCTAAATTAGTAGAGGAATTGAACG CATCGCTTTTGGCTTACGGTATGCCTAAGGAAGCAT GTGACTGGTATGCCCACTCATTGAACTACAACACTC CAGGCGGTAAGCTAAATAGAGGTTTGTCCGTTGTGG ACACGTATGCTATTCTCTCCAACAAGACCGTTGAAC AATTGGGGCAAGAAGAATACGAAAAGGTTGCCATTC TAGGTTGGTGCATTGAGTTGTTGCAGGCTTACTTCT TGGTCGCCGATGATATGATGGACAAGTCCATTACCA GAAGAGGCCAACCATGTTGGTACAAGGTTCCTGAA

		GTTGGGGAAATTGCCATCAATGACGCATTCATGTTA GAGGCTGCTATCTACAAGCTTTTGAAATCTCACTTCA GAAACGAAAAATACTACATAGATATCACCGAATTGTT CCATGAGGTCACCTTCCAAACCGAATTGGGCCAATT GATGGACTTAATCACTGCACCTGAAGACAAAGTCGA CTTGAGTAAGTTCTCCCTAAAGAAGCACTCCTTCATA GTTACTTTCGGTACTGCTTACTATTCTTTCTACTTGC CTGTCGCATTGGCCATGTACGTTGCCGGTATCACG GATGAAAAGGATTTGAAACAAGCCAGAGATGTCTTG ATTCCATTGGGTGAATACTTCCAAATTCAAGATGACT ACTTAGACTGCTTCGGTACCCCAGAACAGATCGGTA AGATCGGTACAGATATCCAAGATAACAAATGTTCTT GGGTAATCAACAAGGCATTGGAACTTGCTTCCGCAG AACAAAGAAAGACTTTAGACGAAAATTACGGTAAGA AGGACTCAGTCGCAGAAGCCAAATGCAAAAAGATTT TCAATGACTTGAAAATTGAACAGCTATACCACGAATA TGAAGAGTCTATTGCCAAGGATTTGAAGGCCAAAAT TTCTCAGGTCGATGAGTCTCGTGGCTTCAAAGCTGA TGTCTTAACTGCGTTCTTGAACAAAGTTTACAAGAGA AGCAAA
mFPS*N144W	*Gallus gallus*	ATGCATAAATTTACTGGTGTCAATGCCAAGTTTCAGC AACCCGCGTTGAGGAACCTCAGCCCCGTGGTGGTT GAGAGGGAGAGGGAGGAGTTCGTGGGGTTCTTCCC GCAGATCGTCCGCGATCTGACCGAGGACGGCATCG GACACCCGGAGGTGGGCGACGCTGTGGCGCGGCT GAAGGAGGTGCTGCAATACAACGCTCCCGGTGGGA AATGCAACCGTGGGCTGACGGTGGTGGCTGCGTAC CGGGAGCTGTCGGGGCCGGGGCAGAAGGATGCTG AGAGCCTGCGGTGCGCGCTGGCCGTGGGTTGGTG CATCGAGTTGTTCCAGGCCTTCTTCCTGGTGGCTGA TGATATCATGGATCAGTCCCTCACGCGCCGGGGGC AGCTGTGTTGGTATAAGAAGGAGGGGGTCGGTTTG GATGCCATCTGGGACTCCTTCCTCCTCGAATCCTCT GTGTACAGAGTGCTGAAGAAGTACTGCAGGCAGCG GCCGTATTACGTGCATCTGTTGGAGCTCTTCCTGCA GACCGCCTACCAGACTGAGCTCGGGCAGATGCTGG ACCTCATCACAGCTCCCGTCTCCAAAGTGGATTTGA GTCACTTCAGTGAGGAGAGGTACAAAGCCATCGTTA AGTACAAGACTGCCTTCTACTCCTTCTACCTACCCG TGGCTGCTGCCATGTATATGGTTGGGATCGACAGTA AGGAAGAACACGAGAATGCCAAAGCCATCCTGCTG GAGATGGGGGAATACTTCCAGATCCAGGATGATTAC CTGGACTGCTTTGGGGACCCGGCGCTCACGGGGAA GGTGGGCACCGACATCCAGGACAATAAATGCAGCT GGCTCGTGGTGCAGTGCCTGCAGCGCGTCACGCC GGAGCAGCGGCAGCTCCTGGAGGACAACTACGGC CGTAAGGAGCCCGAGAAGGTGGCGAAGGTGAAGG AGCTGTATGAGGCCGTGGGGATGAGGGCTGCGTTC CAGCAGTACGAGGAGAGCAGCTACCGGCGCCTGCA GGAACTGATAGAGAAGCACTCGAACCGCCTCCCGA AGGAGATCTTCCTCGGCCTGGCACAGAAGATCTACA AACGCCAGAAATGA
VrtD	*Penicillium aethiopicum*	ATGGCCACTAGCACTACCACTAGCCTGAAGGAATTT TTGTCCGTCTTCCCCCAACTGGTGGCCGACCTCCGT

		GCCCTGTGTTTGGAGGAATATCAGCTTCCCGCGTG CGTGTGGGATCGGTTTGAAAGTACTTTAAATCACAA TACCCTCGGTGGCAAGTGCAACCGCGGCCTCAGTG TCATTGACTCGGTGAGGCTGTTGCGTGACGGCCTC GAGCTGAGCCCAGCTGAGTACTTTGACGCCGCAGT GCTAGGCTGGCTGGTCGAGTTGCTACAGGCGACGA TGCTGGTTCTAGACGATATCATGGACGGATCACCCA CGCGGCGTGGCAAGCCCAGCTGGTACCGTGTACCA GGGGTGGGAATGGCCGCCGTGAACGACGCGACCA TGCTCGAGAGCGCGATATACATGTTGCTGAAGAAGT ATTTCGCGGGACGGGCCATTTACCTCCCTGTGGTC GACTTATTCCATGAAACTGCTCTTCAAATCGAACTCG GCCAGGCCTTTGACATGCTGATCGCCAACGAGGGC ACCCCTGACCTCACGACCTTCGTTCCCGCCACCTAC TCTCAGATTGTTACCTACAAGACTGCCTTCTACAGTT TCTACCTTCCAGTTGCCCTGGCCCTGCACGCCGTTG ATGCTGCCACCCCAACCAACCTAGCTGCCGCCCGC GCCATCCTTGTCCCCATGGGTGAGTATTTCCAGGTC CAAGATGACTATCTAGACTGCTTCGCCGATCCAACC GTCCTGGGTAAGGTTGGCACAGACATTATTGAGGG CAAGTGTTCGTGGCTCGTCGTGCAGGCCCTACAGC GTGCGTCCACAGACCAGGCTCAACTGTTGGCGGAG AATTACGGTTCTGCCAGTGGCGAATCCTCCGTGAAG GCGCTGTACAGCGAGCTGGACCTCGAGAGCGTGTA TCGGACGTTCGAGGAACAGCGCGTTGCTGAGTTGC GCACTCTGATCACCGGACTGGATGAGAGCCAGGGG CTGCGGAAGAGCGTTTTCGAGGAGTTGCTGGGAAA GATCTATCAGCGGCGAAAGTAG
G8H	*Catharanthus roseus*	ATGGACTACCTGACCATTATTTTGACCTTGTTGTTTG CTCTAACCTTGTACGAGGCTTTCTCTTACTTGTCCC GTCGCACCAAGAATTTGCCACCAGGTCCATCTCCAT TGCCATTCATTGGTTCTTTGCATTTGTTGGGCGACC AACCCCATAAAAGCTTGGCTAAATTGTCCAAGAAGC ACGGTCCAATCATGTCCTTGAAGTTGGGTCAGATCA CTACTATCGTCATCTCCTCCTCTACCATGGCTAAAG AAGTCTTGCAAAAACAGGACCTGGCCTTTTCTTCTA GGTCTGTTCCAAACGCCCTGCATGCTCACAACCAAT TCAAATTTTCTGTCGTCTGGTTGCCCGTTGCTTCCA GATGGAGATCTTTGCGCAAGGTCTTGAATTCCAACA TTTTTTCCGGCAACCGCTTGGACGCTAACCAGCATT TGAGGACTCGCAAAGTTCAGGAATTGATCGCTTACT GTAGGAAGAATTCCCAATCTGGTGAAGCTGTCGATG TCGGTAGAGCTGCTTTTAGGACTTCCCTAAACCTGC TGTCTAACTTGATCTTTTCCAAGGACTTGACTGACCC CTATTCTGACTCTGCCAAGGAGTTCAAGGATCTAGT CTGGAACATTATGGTCGAAGCTGGTAAACCAAACTT GGTCGACTTTTTTCCCTTGCTGGAAAAGGTCGATCC ACAAGGCATCAGGCATAGAATGACCATCCATTTTGG CGAGGTTTTGAAGCTATTTGGCGGCTTGGTCAACGA AAGATTGGAACAAAGGAGGTCTAAGGGTGAAAAGAA CGACGTCTTGGACGTTCTGTTGACAACCTCCCAAGA GTCTCCAGAAGAAATTGATAGGACCCATATCGAAAG AATGTGTTTGGACCTGTTCGTTGCTGGTACCGATAC TACCTCCTCAACTTTGGAATGGGCCATGTCTGAGAT

		GTTGAAGAACCCAGATAAGATGAAGAAGACTCAGGA TGAATTGGCCCAAGTTATCGGCAGGGGTAAGACTAT CGAGGAGTCCGACATTAATAGACTGCCATATTTGAG GTGCGTCATGAAGGAAACATTGCGCATTCATCCACC AGTCCCATTTTTGATCCCAAGGAAGGTCGAGCAATC TGTTGAGGTTTGCGGTTATAACGTTCCAAAGGGTTC TCAAGTCTTAGTCAATGCTTGGGCTATTGGTAGGGA TGAAACTGTCTGGGATGACGCCTTGGCTTTCAAACC AGAAAGATTCATGGAATCTGAACTGGACATTAGAGG TAGGGACTTTGAACTGATTCCCTTTGGTGCTGGCAG ACGCATTTGTCCAGGTTTGCCATTGGCTTTGAGAAC AGTTCCATTGATGTTGGGTTCCTTGCTGAACAGCTT TAATTGGAAATTGGAAGGCGGTATGGCCCCCAAGG ATTTGGATATGGAGGAAAAGTTCGGCATTACTTTGC AAAAGGCTCATCCATTGAGAGCTGTTCCATCTACTTT GTAA
GOR	*Catharanthus roseus*	ATGACTAAAACTAATTCTCCAGCCCCATCTGTCATTA CTTGCAAGGCTGCTGTCGTTTGGAAATCCGGTGAAC CACCAAAGGTCGAAGAGATCCAAGTTGATCCACCCA AGGCTTCTGAAGTTCGCATTAAGATGTTGTGTGCTT CCTTGTGCCACACCGATTTCTTGGCTTGTAATGGTC TGCCAGTTCCATTGTTTCCCAGAATTCCAGGTCACG AAGGTGTTGGTATGATCGAATCTGTCGGTGAAAACG TCACCAACTTGAAGGAAGGTGACATTGTCATGCCAT TGTACTTGGGTGAGTGTGGCGAATGCTTGAATTGCA AGTCCGGCAGGACTAACTTGTGTCATAAGTATCCGT TGGGTTTTTCTGGCCTGTTGTTGGATGGCACTTCCA GGATGAGCATTGGCGAACAAAAAGTCTACCACCACT TCTCTTGTTCCACCTGGTCTGAATACATTGTTATTGA GGCCGCCTACGCAGTTAAAGTTGACCCAAGGGTTA GCTTGCCACATGCTTCTTTCCTGTGTTGCGGTTTTA CTACTGGCTTTGGCGCCACTTGGAGAGATGTTAATG TTGTCAAAGGCTCTACTGTCGCTGTTTTGGGTTTAG GTGCCGTCGGTTTGGGTGCTGTTCAAGGCGCTAAA TCTCAAGGTGCCTCCAGGATCATTGGTTTAGACATT AACGATAAGAAGAGGGAGAAAGGCGAAGCTTTCGG CATGACCGAATTCATCAACCCCAAGGGCTCCAATAA GTCCATCTCCGAATTGATCAACGAAGCTACTGGTGG TCTAGGTTTGGACTACGTTTATGAATGCACTGGTGT CCCAGCTCTGTTGAACGAAGCCATTGAGTCCTCTAA AGTTGGTCTGGGTACTGCCGTCTTGATTGGTGCTGG TCTAGAAACCTCTGGTGAAATCAAATTCATTCCCCTG TTGTGCGGCAGAACTGTTAAAGGTTCCATTTACGGT GGTGTTAGGCCAAAGTCCGACTTGCCAACTCTGATT GAGAAGTGCATTAACAAGGAGATTCCAATGGACGAG CTGATGACCCATGAGGTGTCTCTGTCCGAGATCAAC AAGGGTTTCGAGTACTTGAAGCACCCAGACTGTGTC AAAGTTGTTATTAAGTTCTAA
ISY	*Catharanthus roseus*	ATGTCCTGGTGGTGGAAAAGGTCTATTGGTGCTGG CAAAAACTTGCCAAACCAAAACAAGGAAAACGGTGT CTGCAAGTCTTACAAATCTGTCGCCTTGGTCGTCGG TGTTACTGGTATTGTTGGTTCTTCTCTGGCTGAGGTT TTGAAGTTGCCAGATACTCCAGGTGGTCCATGGAAA GTTTATGGTGTTGCTAGAAGACCATGTCCAGTCTGG

		TTGGCTAAGAAGCCAGTCGAGTACATCCAGTGTGAC GTCTCCAATAACCAAGAAACCATTTCTAAGCTGTCTC CCCTGAAAGACATCACTCACATCTTCTATGTCTCCT GGATTGGCTCTGAGGATTGCCAGACTAATGCCACCA TGTTCAAGAACATCTTGAACTCCGTTATCCCAAATGC TTCCAACTTGCAGCACGTCTGCCTACAAACCGGCAT TAAGCATTACTTCGGCATTTTCGAAGAGGGTTCCAA AGTCGTTCCACATGATTCCCCCTTTACCGAAGATTT GCCACGCTTGAACGTCCCAAACTTTTATCACGACCT GGAAGACATTTTGTACGAGGAGACAGGCAAAAATAA CCTAACCTGGTCCGTTCACAGGCCAGCTTTGGTTTT CGGTTTTTCCCCATGCTCCATGATGAATATCGTCTCT ACTCTGTGCGTCTACGCTACTATTTGCAAGCATGAG AACAAGGCTCTGGTTTACCCAGGTTCCAAGAATTCC TGGAATTGCTATGCTGATGCTGTCGATGCTGACTTG GTTGCTGAGCATGAAATTTGGGCTGCTGTTGATCCA AAGGCCAAAAACCAGGTTCTGAATTGCAACAACGGC GACGTCTTCAAATGGAAACATATCTGGAAGAAGCTG GCTGAAGAGTTTGGTATCGAGATGGTCGGTTATGTT GAAGGCAAAGAACAGGTCAGCCTGGCCGAATTGAT GAAAGATAAGGATCAAGTCTGGGACGAAATCGTCAA GAAAAACAACCTGGTGCCAACTAAGTTGAAGGAGAT TGCCGCCTTCTGGTTTGCCGATATCGCCTTTTGCTC TGAAAACTTGATCTCTTCCATGAACAAGTCCAAGGA GCTGGGTTTCCTAGGCTTCAGGAACTCTATGAAGTC TTTCGTCTCCTGTATCGACAAGATGAGAGACTACAG ATTCATTCCATAA

Table S4. Plasmids used for cloning

Plasmid	Description	Purpose	Reference
pCB30	2µ; AmpR; KanMX; URA3_gRNA_cassette	PCR template for fragments to be assembled into CRISPR-Cas9 plasmid, contains guide RNA for URA3 locus	Bond 2018[174]
pCB32	2µ; AmpR; HygR; Nte1_gRNA_cassette	PCR template for fragments to be assembled into CRISPR-Cas9 plasmid	Bond 2018[174]
pCRCT	pRS426; TEF1p_iCas9; RPR1p_tracrRNA	PCR template for fragments to be assembled into CRISPR-Cas9 plasmid	Bao 2014[62]
pδBLE2.0m-ERG13/ERG10	pδBLE2.0m::P_{GAL10}-*NCox4-ERG13*-T_{ADH1}; P_{GAL1}-*NCox4-ERG10*-T_{CYC1}	PCR template for donor DNA for mitochondrial ERG13 and ERG10	Yuan 2016[76]

pδBLE2.0m-ERG12/tHMG1	pδBLE2.0m::P_{GAL10}-*NCox4-ERG12*-T_{ADH1}; P_{GAL1}-*NCox4-tHMG1*-T_{CYC1}	PCR template for donor DNA for mitochondrial ERG12 and tHMG1	Yuan 2016[76]
pδBLE2.0m-ERG19/ERG8	pδBLE2.0m::P_{GAL10}-*NCox4-ERG19*-T_{ADH1}; P_{GAL1}-*NCox4-ERG8*-T_{CYC1}	PCR template for donor DNA for mitochondrial ERG19 and ERG8	Yuan 2016[76]
pδBLE2.0m-IDI1/mFPS	pδBLE2.0m::P_{GAL10}-*NCox4-IDI1*-T_{ADH1}; P_{GAL1}-*NCox4-ERG20*-T_{CYC1}	PCR template for donor DNA for mitochondrial IDI1 and ERG10	This study

Table S5. Media used for yeast culturing

Strain	Starter culture media	Subculture media
CEN.PK1-2C to X303 (Fig. 2A)	SD 2% glucose –uracil	YPD
JHY651 CrGES and JHY651 ObGES (Fig. 2B)	SD 2% glucose –uracil	SD 2% glucose –uracil
JHY651 EV to JHY651 VrtD (Fig. 2C)	SD 2% glucose –uracil	SD 2% glucose –uracil
S1 to S13 (Fig. 3B, Fig. S1)	YPD	YPD
S1-M to S17-M (Fig. 4B)	SD 2% glucose –uracil	SD 0.2% glucose 1.8% galactose –uracil
S1-C to S17-C (Fig. 4C)	SD 2% glucose –uracil	SD 0.2% glucose 1.8% galactose –uracil
S17-M to S19 (Fig. 4D, Fig. S3, Fig. S4)	SD 2% glucose –uracil	SD 0.2% glucose 1.8% galactose –uracil
S20 to S27 (Fig. 5A, Fig. S6, Fig. S7)	YPD	YP 0.2% glucose 1.8% galactose
S25 EV to S25-G8H-GOR-ISY (Fig. 5B, Fig. S8, Fig. S9)	SD 2% glucose –uracil –histidine	YP 0.2% glucose 1.8% galactose

9.2 Supplementary information for Section 6

Table S1. Bioinformatics analysis of the *A. flavus* TC-NRPS gene cluster

Protein	**Size (aa)**	**Proposed function**	**Homologs (ident/pos)**	**Strains**
FlvA *N*-terminal domain (AFLA_135410)	349	pyridoxal-phosphate dependent enzyme	XP_015407945.1 (50%/70%)	*Aspergillus nomius* NRRL 13137
FlvA *C*-terminal domain (AFLA_135410)	213	□KG-Fe dependent dioxygenase	KJK65917.1 (52%/ 71%)	*Aspergillus parasiticus* SU-1

FlvB (AFLA_135420)	308	short chain reductase	KMK58192.1 (59%/73%)	*Aspergillus fumigatus* Z5
FlvC (AFLA_135430)	538	cytochrome P450	XP_018069302.1 (34%/51%)	*Phialocephala scopiformis*
FlvD (AFLA_135440)	511	cytochrome P450	PQE19030.1 (51%/66%)	*Rutstroemia sp.* NJR-2017a WRK4
FlvE (AFLA_135450)	390	trichodiene synthase	O59947.1 (45%/65%)	*Stachybotrys chartarum*
FlvF (AFLA_135460)	366	ophiobolin F synthase	GCB22886.1 (24%/44%)	*Aspergillus awamori*
FlvG (AFLA_135470)	425	ornithine decarboxylase	OOQ91244.1 (60%/74%)	*Penicillium brasilianum*
FlvH (AFLA_135480)	178	histone-lysine N-methyltransferase	PNS15775.1 (51%/69%)	*Sphaceloma murrayae*
FlvI (AFLA_135490)	1071	nonribosomal peptide synthetase	RJE18795.1 (40%/57%)	*Aspergillus sclerotialis*

Table S2. Plasmids used in this study

Plasmid	Vector	Genes
p2001	pYTU	n/a
p2002	pYTR	n/a
p2003	pYTP	n/a
p2004	pYTU	PEgpdAp-FlvD; gpdAp-FlvE; POgpdAp-FlvF
p2005	pYTR	gpdAp-FlvC; POgpdAp-FlvG; PEgpdAp-FlvH
p2006	pYTP	gpdAp-FlvA; POgpdAp-FlvB; PEgpdAp-FlvI
p2007	pYTR	gpdAp-FlvC
p2008	pYTP	POgpdAp-FlvG
p2009	pYTP	gpdAp-FlvA
p2010	pYTP	gpdAp-FlvB
p2011	pYTP	gpdAp-FlvA; POgpdAp-FlvB
p2012	pYTU	PEgpdAp-FlvD; gpdAp-FlvE
p2013	pYTU	PEgpdAp-FlvD; POgpdAp-FlvF
p2014	pYTR	POgpdAp-FlvG; PEgpdAp-FlvH
p2015	pYTU	glaAp-FlvE
p2016	pYTP	amyBp-FlvD
p2017	pXW55	ADH2p-FlvE(codon optimized for *S. cerevisiae*)-ADH2t
p2018	pET28a	T7p-*N*-his-FlvF-T7t
p2019	pXW02	ADH2p-FlvD-ADH2t
p2020	pXW55	ADH2p-FlvE-*C*-his-ADH2t

Table S3. Codon optimization of FlvE for *S. cerevisiae*

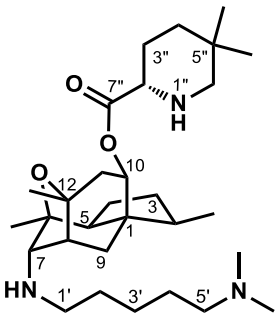
FlvE	ATGCCAGGTAAGCAATTCCCATTGAAGGAATACATCGCTGCTTTGGCTAAGTTC TTGGACACTATCGAATACCAAGACGACAACTTCTCTCACGAACAAAGAGTTGAA TCTTTGAGATACGTTTACCAACACACTGCTAAGCACTTCGACCAACCAATCGAA AAGGCTGCTGTTACTGTTTCTCCAAAGAGATTGCAAGCTGTTATGAGAACTTCT ACTTTGGTTACTGTTTACTGTTGGGTTAAGTGTCCATTGGACGTTATGGTTGGT GTTTCTATCTACTTCGCTTACATCATCATGTTGGACGACTCTTCTGACACTCCA ACTACTGAAATGAAGACTTTCTGTGAAGACTTGATCAAGGGTAGACCACAAAAG CACTTGTTCTGGCAAAGAATGAACGCTCACTTGACTAACTTCTTGAGATACTAC GACGGTTTCTGTGCTATCACTATCTTCAGATCTACTTTGGACTTCTTCCAAGGT TGTTGGATCGAACAAAACAACTTCGGTGGTTTCCCAGGTTCTTCTTACTTCCCA CACTTCTTGAGAAGATTGAACGGTTTGGGTGGTATCTCTTCTGCTACTTTGTTC CCAAGATCTGAATTCGACGAAGGTACTGTTTTCGAAGAAATCGTTACTGCTATC GCTCAAATCGAACCACAATTGACTTTGTGTAACGACTTGATCTCTTTCTACAAG GAATACGACTCTCCAAGAGACCAAATCAACTTGGTTTCTAACTTGGCTCACTGT AACGGTGTTTCTTGGGAAATCGCTTTCGAAGAATTGACTAGAGACACTATCTTG TACTGTGAACAATTGGTTACTGTTTTCAAGGGTAAGGACCCAAAGGTTGAAGCT ACTGTTAGAGCTTTCGTTCACGGTTACGTTACTTGGCACTTGTGTGACCCAAGA TTCAGAATGCAAGAAGTTTACGAACAAGCTGGTCAATCTGAAGCTGACTTGAA GTTCAGACACTTCTACGAACAAGCTACTTCTATCGGTGTTATCGACTTCAAGTT GTGGGCTTCTCCATCTAGATTGTCTTCTGACAAGAGAAAGCACGAACAAGCTTT CGGTGACGACCCACAAAACGGTAAGACTAGAGTTTTGGAATCTATCGGTCAAG CTAACGCTTCTGAAGCTGTTGCTTTGGCTCCATTGGCTTAA

Table S4. Primers used in this study

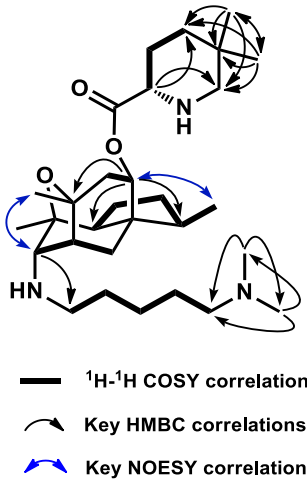
Primer	Sequence (5'-3')
gpdA_TRI F	CTTGACTAACCATTACCCCGCCACATAGACACATCTAAACAA TGCCCGGAAAACAATTCC
TRI_POgpdA R	GAAAGTAGAATCAGTAAGCTCACATGTATTCCTGGAGCAAAA TACAGTCAGTCTGGTGCG
POgpdA_IPPS F	TACAAGTGCATACAGAACACTTCAAACAATCGCAAAAATGGA AGGACTCAGAAGACATTC
IPPS_PEgpdA R	GGGTCCCCAATATTCCAACCTTGGGAAGCCCTGGACGAATC CATTGACGCGTAAGTTCGG
PEgpdA_p450-2 F	CCTCTTATACATGATCTAACAACTTCTAGTAAACCGCAATCAT GCTGGACACCATTTCAG
p450-2_pYTU R	GGGGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGGA ATAATACTAGTGTTGGGAC
gpdA_p450-1 F	ACCATTACCCCGCCACATAGACACATCTAAACATTAATTAAAT GGACTTTGGTGGTATCC
p450-1_POgpdA R	AAAGTAGAATCAGTAAGCTCACATGTATTCCTGGAGCAAAGT AGATGTGCATAGAATGGC
POgpdA_Orn F	AAGCTTACAAGTGCATACAGAACACTTCAAACAATCGCAAAA ATGCCATACGCGACAGAG
orn_PEgpdA R	TGGGGTCCCCAATATTCCAACCTTGGGAAGCCCTGGACGAA TCAAGCCTGGGGGGTCTTG
PEgpdA_HP F	TCTTATACATGATCTAACAACTTCTAGTAAACCGCAATCATGT CTGAATTCATCCCTACC
HP_pYTR R	GCTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAAATGC AGCTAGAAGGATATTAGG
gpdA_OG-Fe F	GCTTGACTAACCATTACCCCGCCACATAGACACATCTAAACA ATGACACCCAAACCCTTG
Og-Fe_POgpdA R	TCGAAAGTAGAATCAGTAAGCTCACATGTATTCCTGGAGCAA ACACTGACGGCCCTCTTG
POgpdA_SDR F	CTTACAAGTGCATACAGAACACTTCAAACAATCGCAAAAATG AGATACCTCATAACAGGG
SDR_PEgpdA R	GGGTCCCCAATATTCCAACCTTGGGAAGCCCTGGACGAATC ACACAAACCCGGCCATATC
PEgpdA_NRPS F	CCCTCTTATACATGATCTAACAACTTCTAGTAAACCGCAATCA TGGCGAAGGATCAGTCC
NRPS_pYTP R	TGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAACAA AGATCTCAAAGTTGTTGG
POgpdA F	TTTGCTCCAGGAATACATGTG
PO_gpdA R	TTTTGCGATTGTTTGAAGTGTTC
PE_gpdA F	GATTCGTCCAGGGCTTCC
PEgpdA R	GATTGCGGTTTACTAGAAGTTGTTAG
p450-1_pYTR R	GCTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAAATGT AGATGTGCATAGAATGGC
pYTR_POgpdA F	CGCGGGTGTTCTTGACGATGGCATCCTGCGGCCGCTTTGCT CCAGGAATACATGTGAGC
orn_pYTR R	TCTGCTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAAA TAAGCCTGGGGGGTCTTG
pYTP_gpdA F	AACCTCGCGGGTGTTCTTGACGATGGCATCCTGCGGCCGCA CTCCGGTGAATTGATTTGG
gpdA SDR F	TGACTAACCATTACCCCGCCACATAGACACATCTAAACAATG AGATACCTCATAACAGGG
SDR pYTP R	CTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAA CACAAACCCGGCCATATC
2OgFe_pYTP R	AGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTA ACACTGACGGCCCTCTTG

TC_PEgpdA R	GGGTCCCCAATATTCCAACCTTGGGAAGCCCTGGACGAATC ATACAGTCAGTCTGGTGCG
pYTU POgpdA F	AGTCATTTACTCAGCACACTCGCGCTGACGCTCGTCTTTGCT CCAGGAATACATGTGAGC
pET28a AflaIPPS F	ATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA TATGGAAGGACTCAGAAGAC
AflaIPPS pET28a R	GTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCG AGTCAGTTTGCGTACCGAAC
Afl-782tp-F	AGCATCATTACACCTCAGCAATGCCCGGAAAACAATTCCCC
Afl-782tp-R	AGAGCGGCCGCCTTAATCTAAGCCAGAGGAGCTAACG
Afl-780cyt-F	CACATAGACACATCTAAACAATGGACTTTGGTGGTATCCTC
Afl-780cyt-R	CCCGGGGGATCCTTAATCCATCGAATCAAAGCGATATGG
Afl-781cyt-F	AACCCCACAGAAGGCATTTATGCTGGACACCATTTCAGAAAG
Afl-781cyt-R	CGAGGTCGACGGTATCGATATTACAGGATCTTTTCCGTAACT TTC
XW55-782tp-F	CTATTAACTATATCGTAATACCATATGCCCGGAAAACAATTCC CC
XW55-782tp-R	GATGGTGATGGTGATGCACTTACTAAGCCAGAGGAGCTAAC G
XW02-781cyt-F	CTATTAACTATATCGTAATACCATATGCTGGACACCATTTCAG AAAG
XW02-781cyt-R	CTATAAATCGTGAAGGCATGTTTACAGGATCTTTTCCGTAAC TTTC
Afla RT 2OG-Fe F	TCCTTGCCCAGGTGATGG
Afla RT 2OG-Fe R	CTTGACGACGAATGCTGAG
Afla RT SDR F	GTAATCGACGTCGAAAAGGTC
Afla RT SDR R	CTCGAATCCACCCTGGATC
Afla RT p450-1 F	ACAACTTCCCTTCCTGGTTC
Afla RT p450-1 R	CTTGGGGAAGTCTCTGACG
Afla RT p450-2 F	CCGTAACTTTCACCTTCACCTC
Afla RT p450-2 R	CTTGAACAACAAGACCATCCTG
Afla RT Tri F	ATGCCCGGAAAACAATTCC
Afla RT Tri R	GAAGGCACGAACGGTAGC
Afla RT IPPS F	CCTTTCAGCCTGTTCAGAAG
Afla RT IPPS R	GGCCTATCTCAATGAAGTGTTG
Afla RT ODC F	CTATAGATAGATGTCCGCCGCA
Afla RT ODC R	GTGGCCCTTCCAAAGTCAC
Afla RT HP F	ATGTCTGAATTCATCCCTACC
Afla RT HP R	TCAAGCTTTGAGGGAAATCTCC
Afla RT NRPS F	ATGGCGAAGGATCAGTCC
Afla RT NRPS R	GGCGACTAAAGCTTGGTG
Afla RT actin F	CCCCCCGTGCTGTCTTC
Afla RT actin R	GAAAGTGTAACCGCGCTCAG

Table S5: Spectroscopic data of **1**



1



^{1}H-^{1}H COSY correlations

Key HMBC correlations

Key NOESY correlations

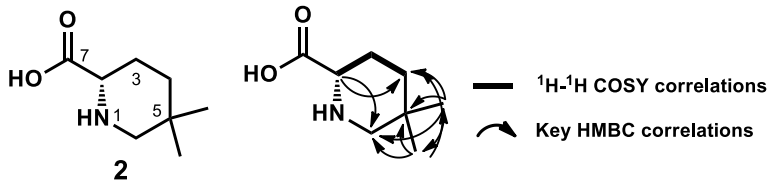
HRMS (ESI, M+H^{+}) calculated for $C_{30}H_{54}N_3O_3$ 504.4165; found 504.4214

$[\alpha]_D^{24}$ + 10 (*c* 0.1, H_2O)

Position	δ_H (mult., $J_{H\text{-}H}$ in Hz)	δc
1		46.9
2	1.65 (1H, m)	43.2
3	1.77 (1H, m) 1.07 (1H, m)	31.9
4	1.80 (1H, m) 1.26 (1H, m)	22.7
5	2.20 (1H, m)	52.2
6		81.7
7	2.72 (1H, d, *J* = 4.3 Hz)	67.0
8	1.95 (1H, t, *J* = 4.3 Hz)	44.0
9	1.44 (1H, m) 1.20 (1H, m)	28.4
10	4.92 (1H, d, *J* = 8.0 Hz)	72.4
11	2.00 (1H, dd, *J* = 15.0, 8.0 Hz) 1.42 (1H, d, *J* = 15.0 Hz)	42.9
12		79.5
13	1.09 (3H, s)	28.7
14	0.76 (3H, d, *J* = 6.6 Hz)	15.1
15	1.03 (3H, s)	21.4
1′	2.46 (2H, m)	48.6
2′	1.41 (2H, m)	30.2
3′	1.27 (2H, m)	25.0
4′	1.40 (2H, m)	26.8
5′	2.29 (2H, t, *J* = 8.0 Hz)	59.1
2″	3.08 (1H, dd, *J* = 10.0, 2.3 Hz)	58.8
3″	1.63 (1H, m) 1.48 (1H, m)	24.8
4″	1.35 (1H, m) 1.22 (1H, m)	37.3
5″		29.6
6″	2.51 (1H, d, *J* = 10.0 Hz) 2.29 (1H, d, *J* = 10.0 Hz)	56.9
7″		173.5
8″	0.88 (3H, s)	24.9
9″	0.80 (3H, s)	28.8
(*Me*)$_2$N-5′	2.18 (6H, s)	45.0

In DMSO-d_6, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

Table S6: Spectroscopic data of **2**



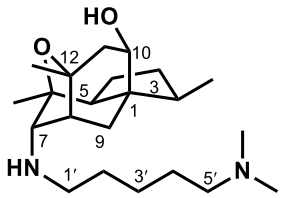
HRMS (ESI, M+H$^+$) calculated for $C_8H_{16}NO_2$ 158.1181; found 158.1232

Absolute stereochemistry was determined as L-(-)-5, 5-dimethyl pipecolic acid according to $[\alpha]_D^{24}$ = -52.0 (*c* 0.2, H_2O), in accordance with that of reported L-(-)-pipecolic acid.

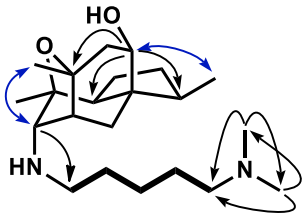
Position	δ_H (mult., J_{H-H} in Hz)	δc
2	3.44 (1H, dd, *J* = 11.2, 3.6 Hz)	58.5
3	2.01 (1H, m) 1.73 (1H, m)	23.0
4	1.49 (1H, m) 1.39 (1H, m)	34.5
5		28.3
6	2.95 (1H, d, *J* = 12.8 Hz) 2.72 (1H, d, *J* = 12.8 Hz)	52.9
7		174.2
8	0.93 (3H, s)	22.7
9	0.90 (3H, s)	27.0

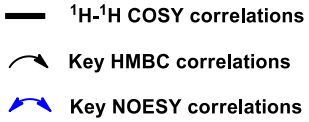
In D_2O, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

Table S7: Spectroscopic data of **3**



3





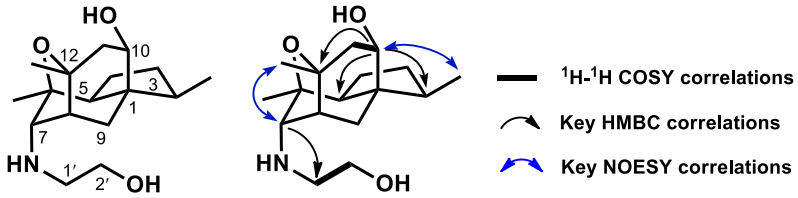
HRMS (ESI, $M+H^+$) calculated for $C_{22}H_{41}N_2O_2$ 365.3168; found 365.3152

$[\alpha]_D^{24}$ - 14 (*c* 0.1, H_2O)

Position	δ_H (mult., J_{H-H} in Hz)	δc
1		47.7
2	1.64 (1H, m)	43.0
3	1.80 (1H, m)	31.2
	1.29 (1H, m)	
4	1.66 (1H, m)	22.5
	1.35 (1H, m)	
5	1.78 (1H, dd, *J* = 12.0, 8.0 Hz)	52.8
6		81.9
7	3.07 (1H, d, *J* = 4.3 Hz)	65.4
8	2.19 (1H, t, *J* = 4.3 Hz)	42.3
9	1.35 (1H, dd, *J* = 14.0, 4.3 Hz)	28.0
	1.03 (1H, d, *J* = 14.0 Hz)	
10	3.92 (1H, d, *J* = 7.8 Hz)	69.0
11	2.02 (1H, dd, *J* = 15.2, 7.8 Hz)	43,5
	1.61 (1H, d, *J* = 15.2 Hz)	
12		82.5
13	1.16 (3H, s)	26.8
14	0.83 (3H, d, *J* = 6.9 Hz)	13.9
15	1.15 (3H, s)	19.9
1′	2.74 (2H, t, *J* = 7.1 Hz)	47.4
2′	1.54 (2H, m)	26.9
3′	1.27 (2H, m)	23.2
4′	1.62 (2H, m)	23.7
5′	3.00 (2H, t, *J* = 8.0 Hz)	57.5
(*Me*)$_2$N-5′	2.74 (6H, s)	42.6

In D_2O, 500 MHz for ^1H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H-^1H COSY, HMBC and HSQC correlation.

Table S8: Spectroscopic data of **7a**

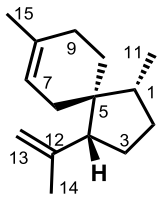


7a

Position	δ_H (mult., J_{H-H} in Hz)	δc
1		48.0
2	1.57 (1H, m)	44.0
3	1.70 (1H, m)	32.4
	1.41 (1H, m)	
4	1.83 (1H, m)	23.2
	1.17 (1H, m)	
5	2.10 (1H, dd, *J* = 13.5, 8.5 Hz)	52.4
6		82.1
7	2.73 (1H, d, *J* = 4.5 Hz)	67.0
8	1.89 (1H, t, *J* = 4.5 Hz)	44.2
9	1.23 (1H, d, *J* = 12.5 Hz)	28.6
	1.09 (1H, dd, *J* = 12.5, 4.5 Hz)	
10	3.75 (1H, d, *J* = 8.0 Hz)	68.8
11	1.88 (1H, dd, *J* = 14.0, 8.0 Hz)	45.5
	1.57 (1H, d, *J* = 14.0 Hz)	
12		80.7
13	1.12 (3H, s)	29.0
14	0.88 (3H, d, *J* = 7.0 Hz)	15.1
15	1.03 (3H, s)	21.4
1′	2.55 (2H, t, *J* = 5.0 Hz)	51.2
2′	3.43 (2H, t, *J* = 5.0 Hz)	61.1

In DMSO-d_6, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

Table S9: Spectroscopic data of **8**



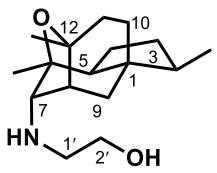
8

$[\alpha]_D^{24.4}$ +2.22 (*c* 0.1, hexane). The optical rotation of (*1R,4R,5S*)-(+)-acoradiene was reported to be: $[\alpha]_D^{23}$ +2.96 (*c* 0.75, hexane)

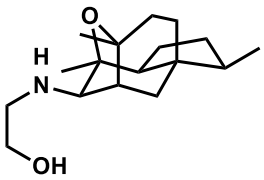
Position	(*1R,4R,5S*)-(+)-acoradiene in $CDCl_3$ (500MHz) *δc*	Reported (*1R,4R,5S*)-(+)-acoradiene in $CDCl_3$ (400MHz) *δc*
1	44.0	44.0
2	31.4	31.4
3	27.8	27.8
4	57.9	57.9
5	44.6	44.6
6	39.1	39.1
7	120.9	120.9
8	133.8	133.8
9	28.7	28.7
10	24.1	24.1
11	15.8	15.8
12	147.1	147.1
13	111.8	111.8
14	24.1	24.1
15	23.3	23.3

In $CDCl_3$, 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm.

Table S10: Spectroscopic data of **9a** and **9b**



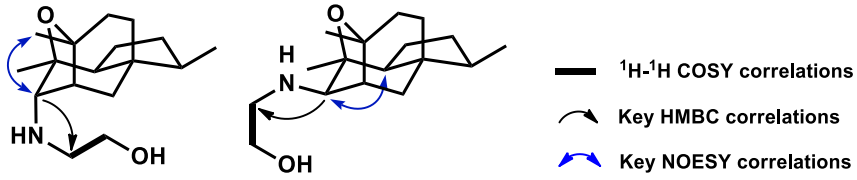
9a



9b

9a HRMS (ESI, M+H$^+$) calculated for $C_{17}H_{30}NO_2$ 280.2277; found 280.2288 $[\alpha]_D^{24}$ -41 (*c* 0.1, MeOH)

9b HRMS (ESI, M+H$^+$) calculated for $C_{17}H_{30}NO_2$ 280.2277; found 280.2235 $[\alpha]_D^{24}$ -131 (*c* 0.1, MeOH)



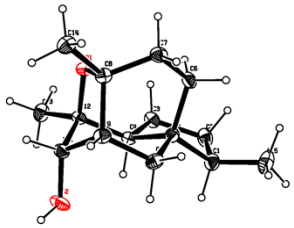
9a and **9b** are C-7 isomers as shown above based on the NOE interactions from H-7 to H_3-13 in **9a** and from H-7 to H-5 in **9b**. The corresponding interactions can be found in figures S47 and S53.

	9a		**9b**	
Position	δ_H (mult., $J_{H\text{-}H}$ in Hz)	δc	δ_H (mult., $J_{H\text{-}H}$ in Hz)	δc
1		43.2		43.9
2	1.47 (1H, m)	43.0	1.44 (1H, m)	43.0
3	1.66 (1H, m) 1.02 (1H, m)	29.6	1.72 (1H, m) 1.10 (1H, m)	30.3
4	1.39 (1H, m) 1.30 (1H, m)	20.5	1.44 (1H, m) 1.38 (1H, m)	20.9
5	1.77 (1H, dd, *J* = 12.8, 7.7 Hz)	51.6	1.48 (1H, m)	60.1
6		81.7		83.9
7	2.67 (1H, d, *J* = 5.0 Hz)	66.6	2.53 (1H, s)	73.8
8	1.95 (1H, t, *J* = 5.0 Hz)	44.5	2.13 (1H, d, *J* = 5.5 Hz)	46.2
9	1.32 (2H, dd, *J* = 14.0, 5.0 Hz)	26.9	1.67 (1H, dd, *J* = 12.5, 5.5 Hz) 0.94 (1H, d, *J* = 12.5 Hz)	34.7
10	1.25 (1H, m) 1.10 (1H, m)	18.3	1.33 (1H, m) 1.13 (1H, m)	18.1
11	1.54 (1H, m) 1.41 (1H, m)	32.4	1.48 (1H, m) 1.37 (1H, m)	32.8
12		79.7		81.9
13	1.09 (3H, s)	29.3	1.19 (3H, s)	29.1
14	0.76 (3H, d, *J* = 6.8 Hz)	14.6	0.75 (3H, d, *J* = 6.8 Hz)	14.6
15	0.97 (3H, s)	21.5	1.02 (3H, s)	17.5
1′	2.54 (2H, m)	51.0	2.66 (1H, m) 2.50 (1H, m)	50.1
2′	3.43 (2H, m)	61.1	3.45 (2H, m)	61.2

In DMSO-d_6, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

Table S11: Spectroscopic data of **10a**

$[\alpha]_D^{24}$ - 45 (*c* 0.1, $CHCl_3$)



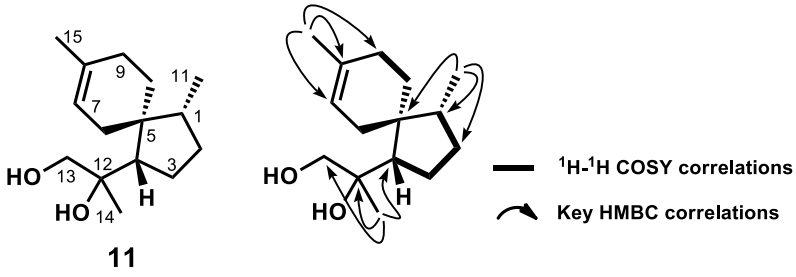
Crystallographic data for the structure has been deposited with the Cambridge Crystallographic Data Center as deposition number CCDC-1963931. Copies of the data can be obtained, free of charge, on application to CCDC.

Computer-generated ORTEP structure

Position	δ_H (mult., $J_{H\text{-}H}$ in Hz)	δc
1		42.9
2	1.56 (1H, m)	43.3
3	1.76 (1H, m) 1.14 (1H, m)	29.5
4	1.59 (1H, m) 1.45 (1H, m)	20.4
5	1.85 (dd, 1H, *J* = 12.8, 7.7 Hz)	51.0
6		80.9
7	3.88 (d, 1H, *J* = 5.4 Hz)	76.5
8	2.04 (t, 1H, *J* = 5.4 Hz)	46.0
9	1.48 (dd, 2H, *J* = 14.7, 5.4 Hz)	27.0
10	1.43 (1H, m) 1.26 (1H, m)	18.0
11	1.71 (1H, m) 1.47 (1H, m)	32.2
12		79.6
13	1.21 (3H, s)	28.9
14	0.83 (d, 3H, *J* = 6.9 Hz)	14.1
15	1.14 (3H, s)	20.1

In $CDCl_3$, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation. The absolute stereochemistry was determined by single-crystal X-ray diffraction analysis.

Table S12: Spectroscopic data of **11**



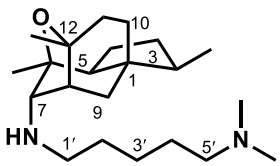
11

$[\alpha]_D^{24}$ - 48 (*c* 0.1, $CHCl_3$)

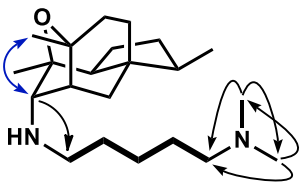
Position	δ_H (mult., J_{H-H} in Hz)	δc
1	1.77 (1H, m)	42.8
2	1.71 (1H, m)	30.5
	1.30 (1H, m)	
3	1.65 (1H, m)	26.0
	1.48 (1H, m)	
4	1.85 (1H, m)	56.1
5		44.6
6	2.00 (2H, s)	40.3
7	5.36 (1H, br s)	121.4
8		134.4
9	2.01 (2H, m)	28.9
10	1.87 (1H, m)	24.2
	1.79 (1H, m)	
11	0.91 (d, 3H, *J* = 6.9 Hz)	15.5
12		75.5
13	3.43 (d, 1H, *J* = 10.0 Hz)	70.7
	3.30 (d, 1H, *J* = 10.0 Hz)	
14	1.23 (3H, s)	22.1
15	1.65 (3H, s)	23.3

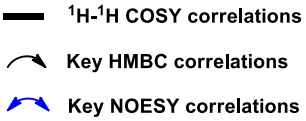
In $CDCl_3$, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

Table S13: Spectroscopic data of **12**



12





HRMS (ESI, M+H$^+$) calculated for $C_{22}H_{41}N_2O$ 349.3219; found 349.3203

$[\alpha]_D^{24}$ - 29 (*c* 0.1, MeOH)

Position	δ_H (mult., $J_{H\text{-}H}$ in Hz)	δc
1		43.2
2	1.49 (1H, m)	43.0
3	1.66 (1H, m) 1.02 (1H, m)	29.6
4	1.39 (1H, m) 1.30 (1H, m)	20.5
5	1.79 (1H, dd, *J* = 12.7, 7.8 Hz)	51.5
6		81.6
7	2.65 (1H, d, *J* = 4.6 Hz)	66.7
8	1.94 (1H, m)	44.3
9	1.32 (2H, m)	26.9
10	1.24 (1H, m) 1.11 (1H, m)	18.3
11	1.52 (1H, m) 1.41 (1H, m)	32.3
12		79.4
13	1.08 (3H, s)	29.3
14	0.75 (3H, d, *J* = 6.8 Hz)	14.5
15	0.97 (3H, s)	21.5
1′	2.45 (2H, m)	48.5
2′	1.41 (2H, m)	30.1
3′	1.27 (2H, m)	24.9
4′	1.42 (2H, m)	26.4
5′	2.37 (2H, m)	58.8
(*Me*)$_2$N-5′	2.24 (6H, s)	44.6

In DMSO-d_6, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

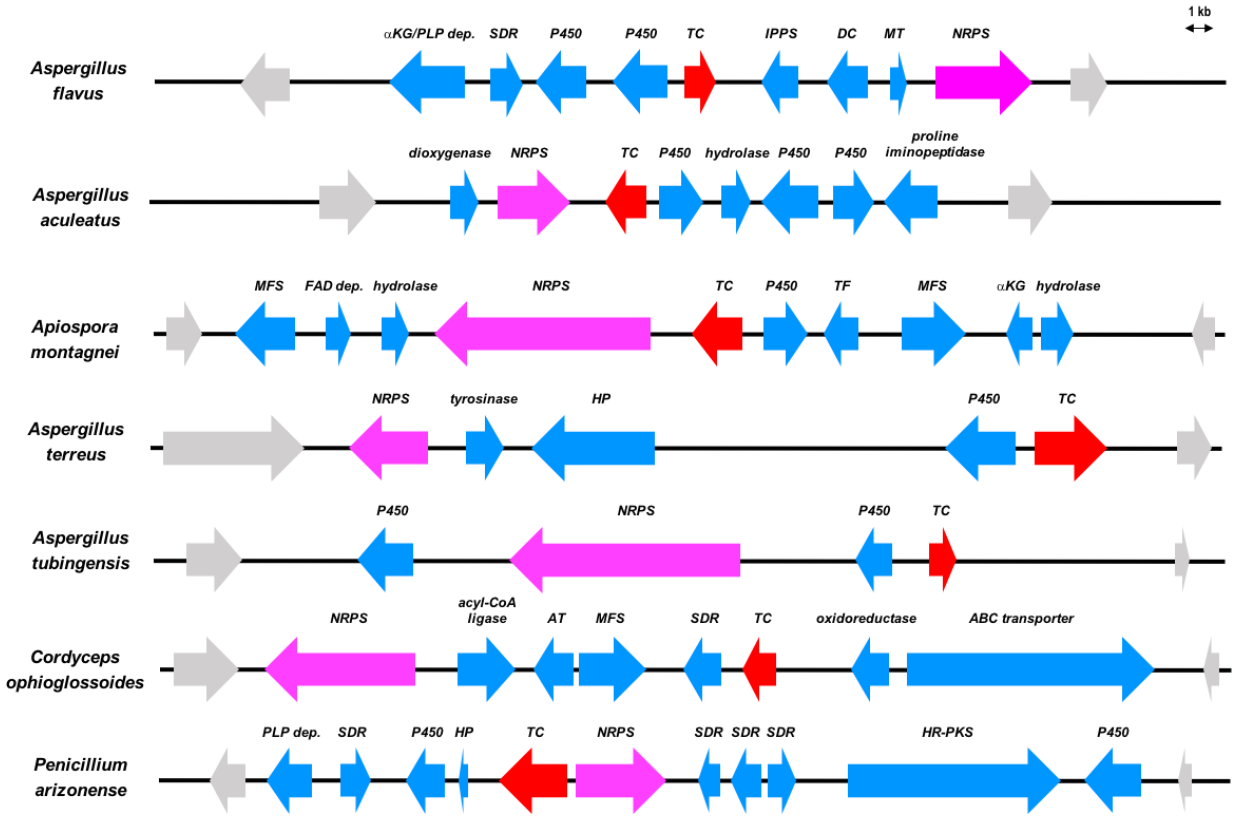


Figure S1. Natural product gene clusters containing a TC and NRPS in fungi. The cluster from *A. flavus* was further investigated in this study. The cluster from *A. aculeatus* encodes the pathway for aculene A.[10] The remaining clusters are putative.

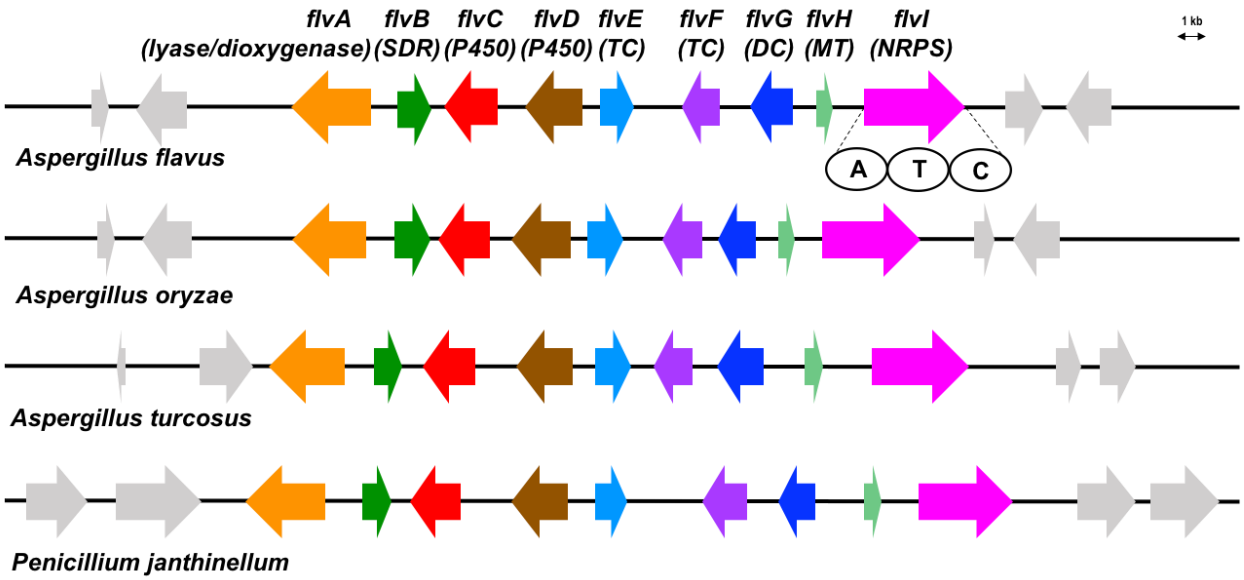


Figure S2. Homologous clusters in *A. flavus*, *Aspergillus turcosus*, and *Penicillium janthinellum*. Nine genes from *flvA* to *flvI* are conserved in all of the clusters.

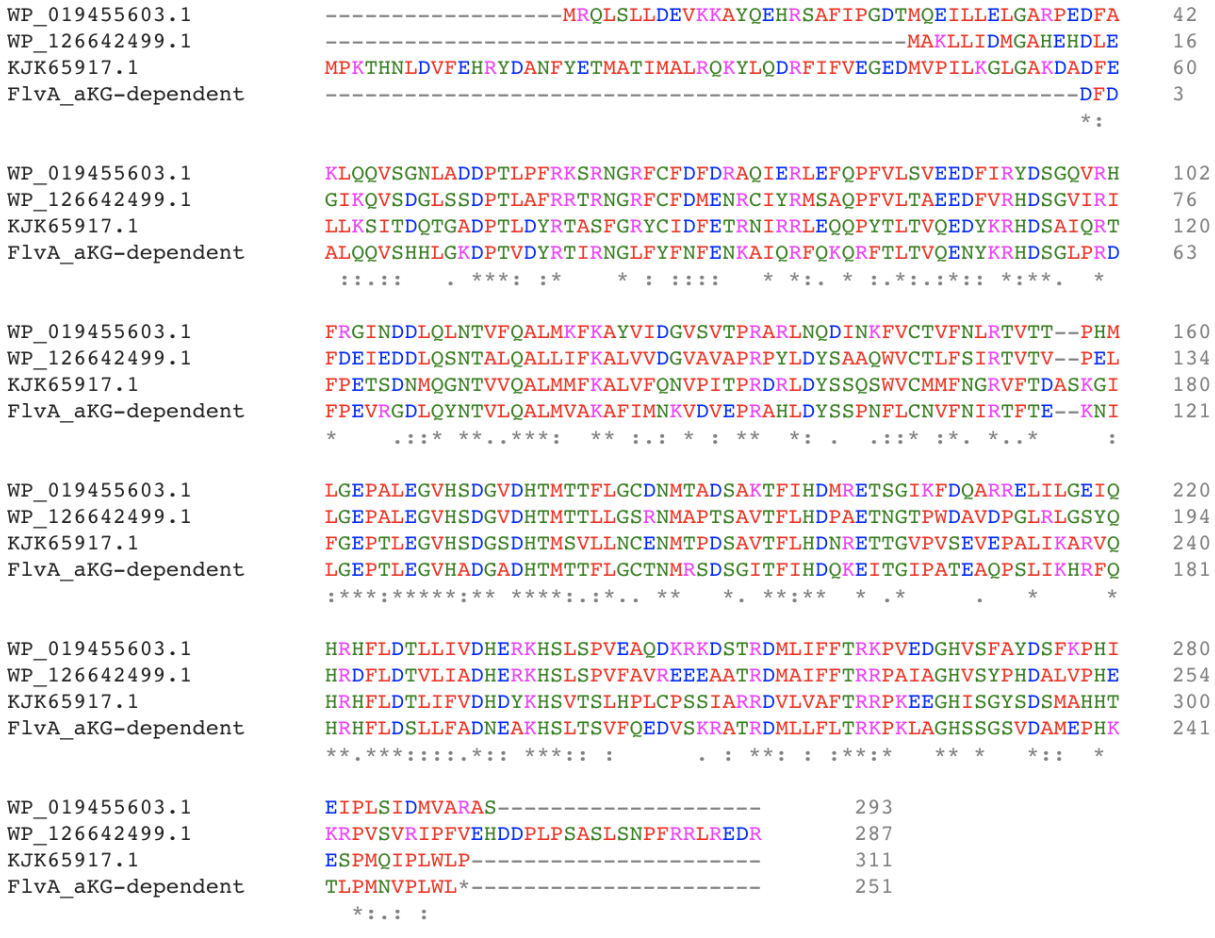


Figure S3. Clustal omega (1.2.4) multiple sequence alignment of the *C*-terminal αKG-dependent domain of FlvA with other genes annotated as αKG dioxygenases: WP_019455603.1 from *Serratia marcescens* (51.4% identity), WP_126642499.1 from *Streptomyces hyalinus* (49.4% identity), and KJK65917.1 from *Aspergillus parasiticus* SU-1 (51.8% identity).

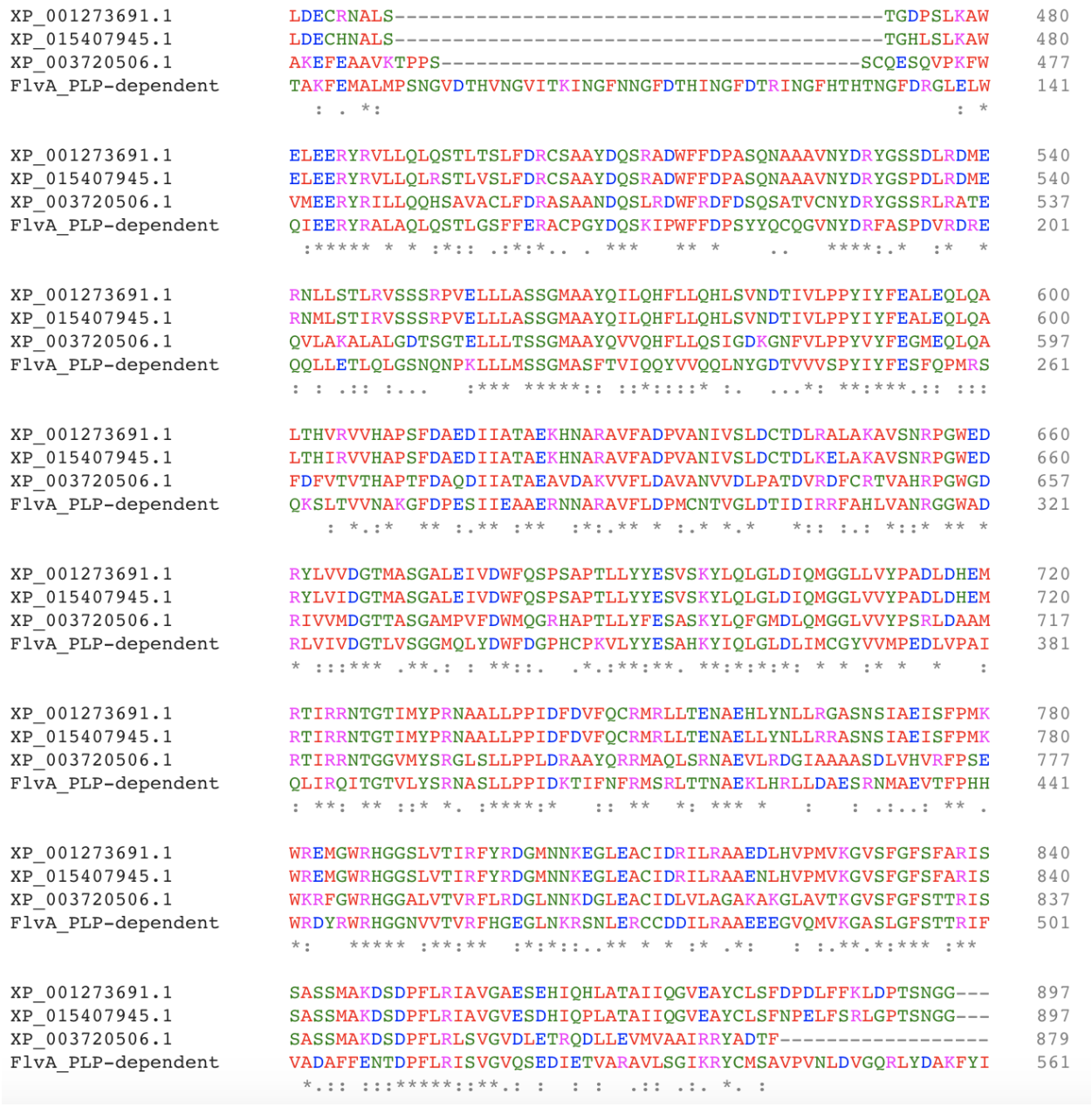


Figure S4. Clustal omega (1.2.4) multiple sequence alignment of the *N*-terminal PLP-dependent domain of FlvA with other genes annotated as PLP-dependent enzymes: XP_001273691.1 from *Aspergillus clavatus* NRRL 1 (50.1% identity), XP_015407945.1 from *Aspergillus nomius* NRRL 13137 (49.6% identity), and XP_003720506.1 from *Pyricularia oryzae* 70-15 (45.5% identity).

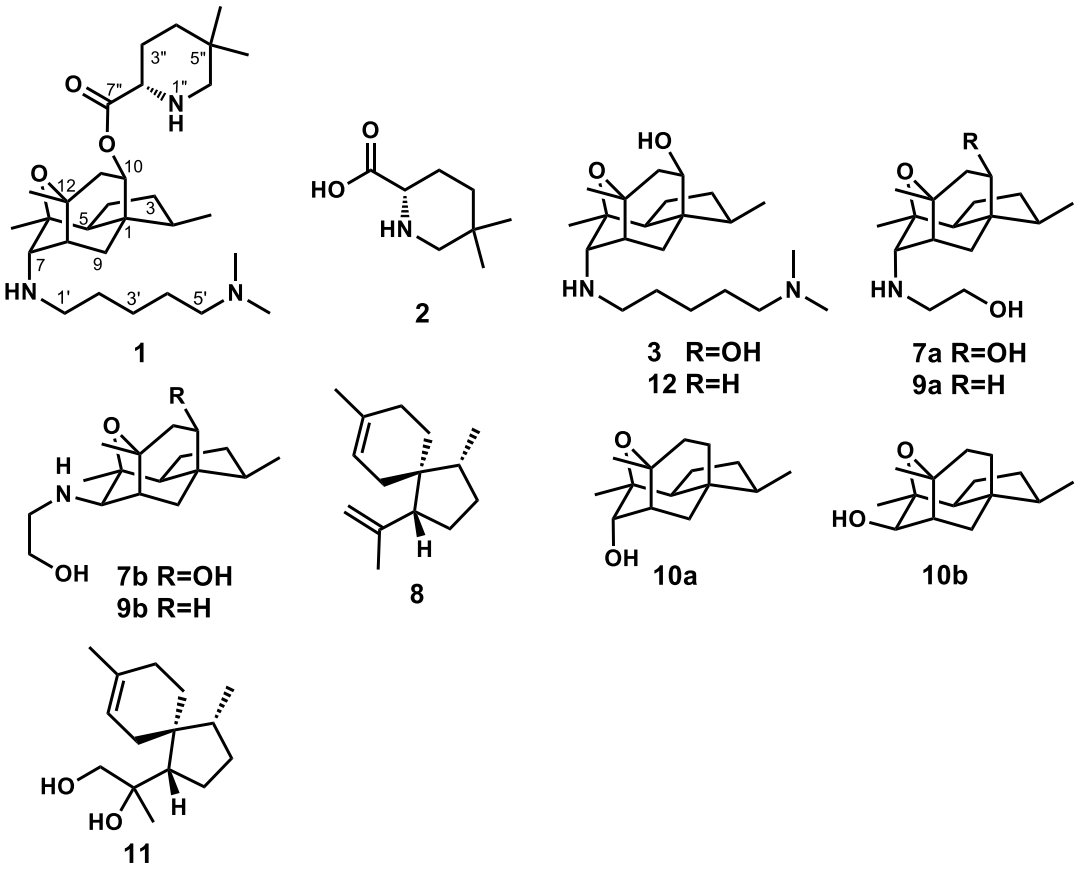


Figure S5. Structures of compounds **1-3, 7**-**12**

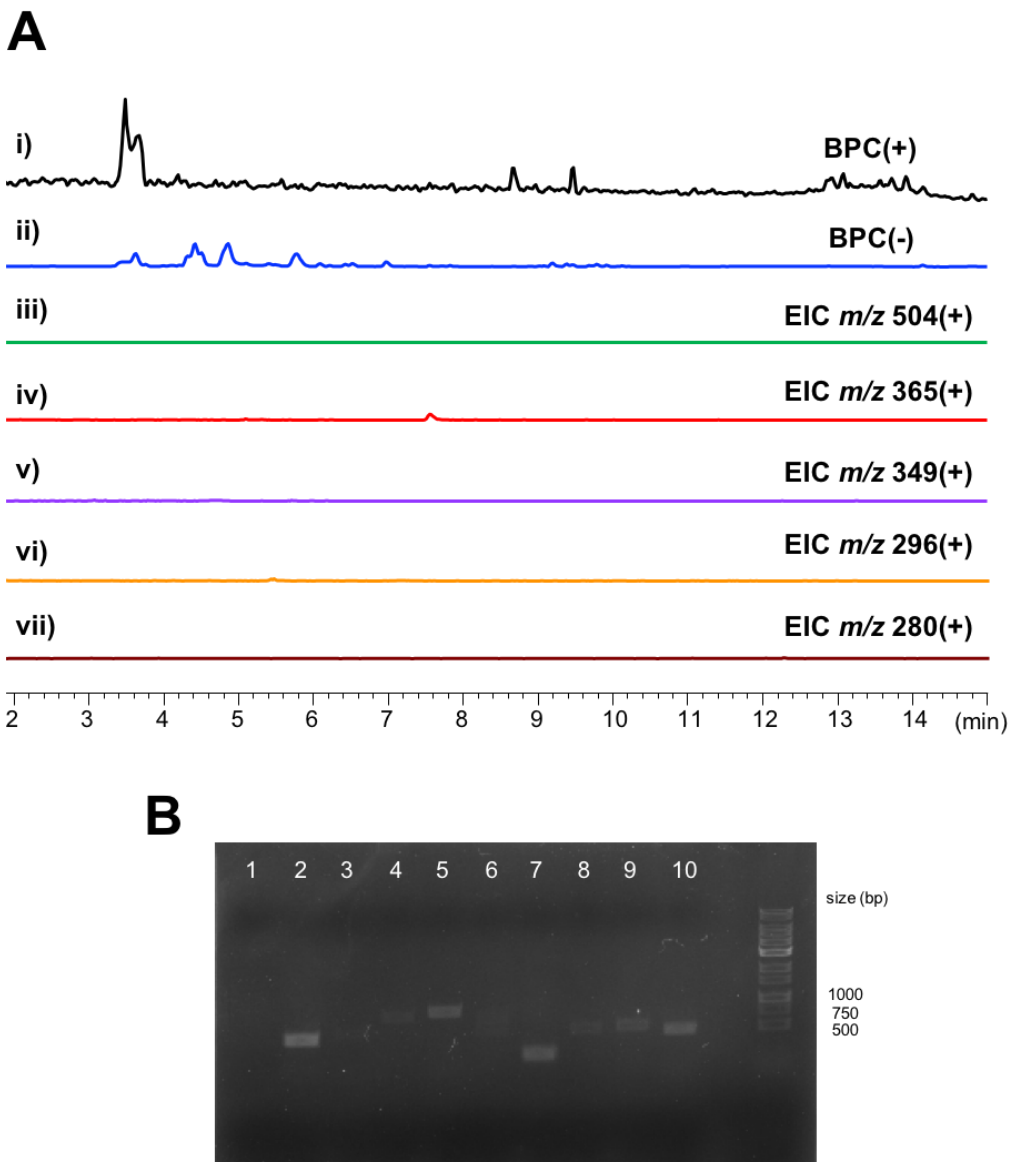


Figure S6. (A) LC/MS analysis of the metabolic extracts from *A. flavus* grown on PDA for 5 days. No flavunoidines were detected. BPC(+) (i), BPC(-) (ii), and extracted LC traces corresponding to the m/z^+ for **1** (m/z^+ = 504) (iii), **3** (m/z^+ = 365) (iv), **12** (m/z^+ = 349) (v), **7a,b** (m/z^+ = 296) (vi), and **9a,b** (m/z^+ = 280) (vii) are shown. (B) RT-PCR of *flv* cluster genes from *A. flavus* grown on PDA for 5 days. Lanes 1-10 correspond to *flvA-I* and actin as a positive control, respectively. *flvA* was not expressed and the PCR bands for many other genes were faint. This qualitatively suggests weak to no expression of the cluster.

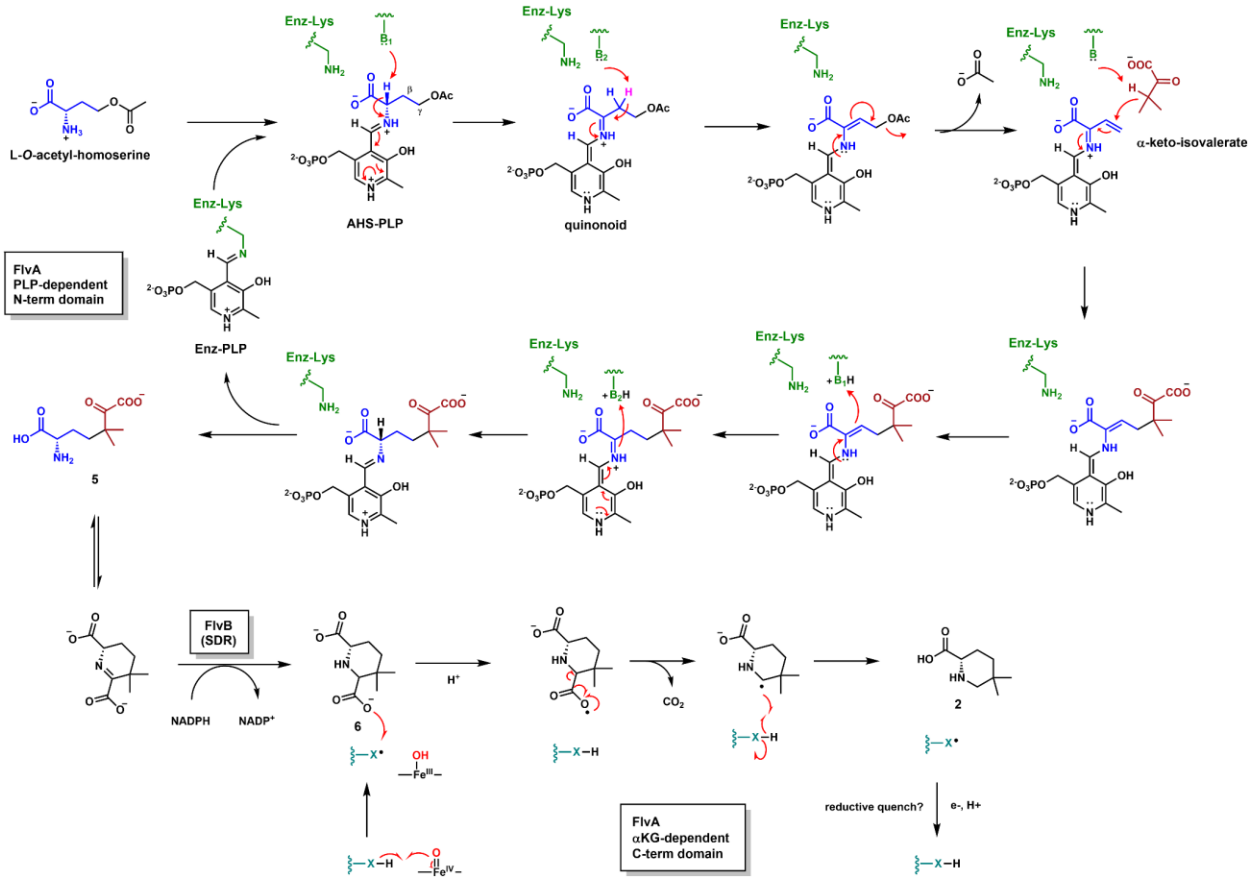


Figure S7. Proposed pathway for the biosynthesis of **2** from L-*O*-acetyl-homoserine and α–keto-isovalerate by FlvA and FlvB. The mechanism of decarboxylation of **6** to **2** can involve a radical decarboxylative route facilitated by an active site tyrosyl[11, 12] or thiyl[13] radical generated from the Fe(IV)=O. Upon decarboxylation to generate the C4 radical, hydrogen delivery from the active residue forms **2**, and reductive quenching of the radical carrier regenerates the enzyme.

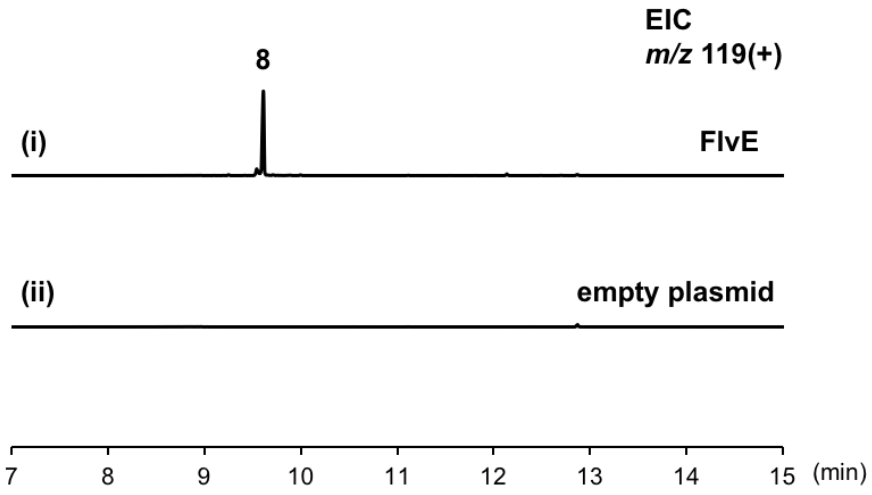


Figure S8. GC/MS analysis of the metabolic extracts from the culture of (i) *S. cerevisiae* expressing FlvE and (ii) *S. cerevisiae* with an empty vector (negative control). Extracted GC traces corresponding to the m/z^+ for **8** (m/z^+ = 119).

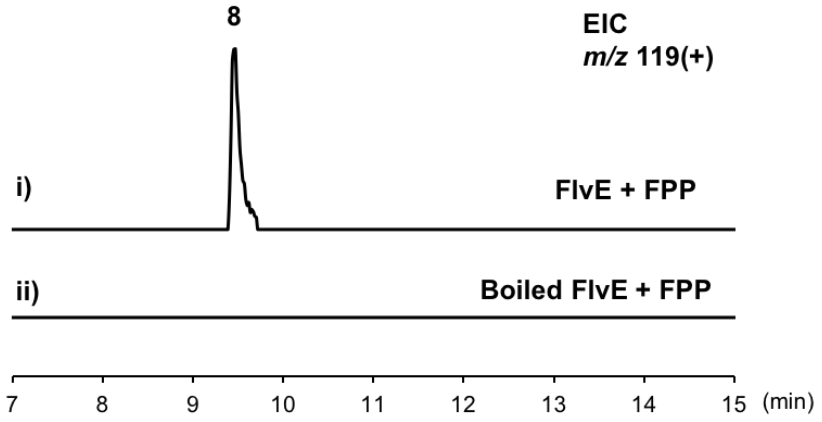


Figure S9. GC/MS analysis of the extracts from an *in vitro* reaction with (i) recombinant FlvE purified from *E. coli* and FPP and (ii) boiled recombinant FlvE purified from *E. coli* and FPP (negative control). Extracted GC traces corresponding to the m/z^+ for **8** (m/z^+ = 119).

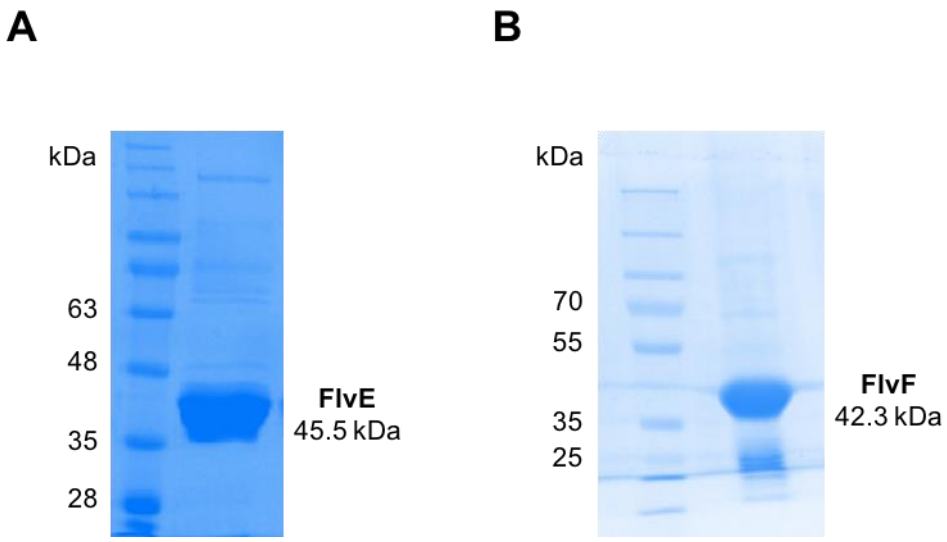


Figure S10. SDS-PAGE (12%) analysis of (A) FlvE (codon-optimized) and (B) FlvF expressed and purified from *S. cerevisiae* and *E. coli*, respectively

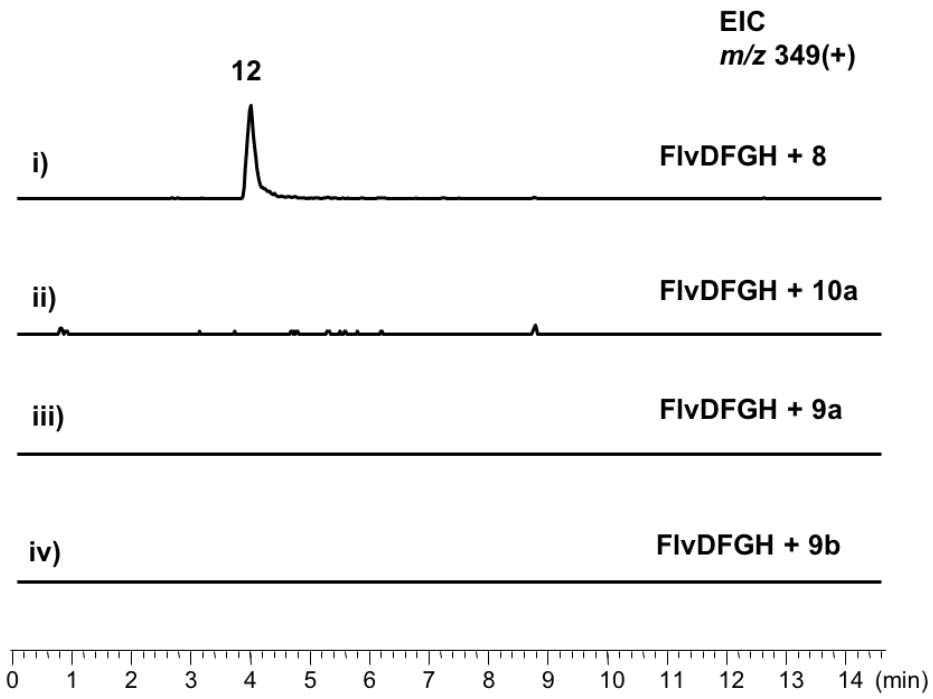


Figure S11. LC/MS analysis of the metabolic extracts from the culture of *A. nidulans* expressing FlvDFGH feeding (i) **8**, (ii) **10a**, (iii) **9a**, and (iv) **9b**. Extracted LC traces corresponding to the m/z^+ for **12** (m/z^+ = 349).

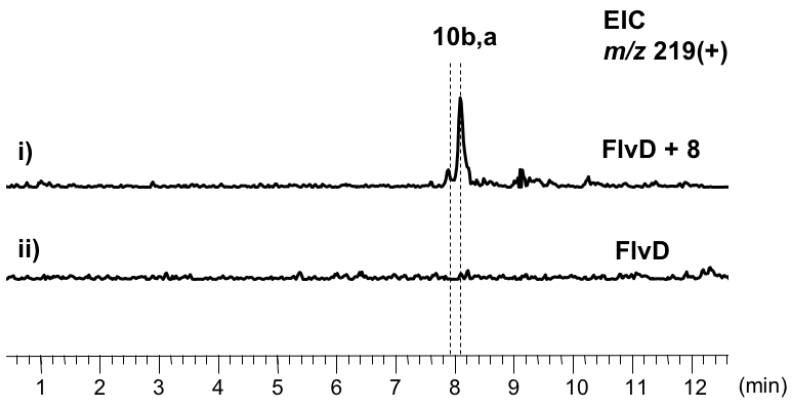


Figure S12. LC/MS analysis of the extracts from an *in vitro* reaction with (i) purified microsomes from yeast expressing FlvD and **8** and (ii) purified microsomes from yeast expressing FlvD with no substrate (negative control). Extracted LC traces corresponding to the m/z^+ for **10a,b** (m/z^+ = 219).

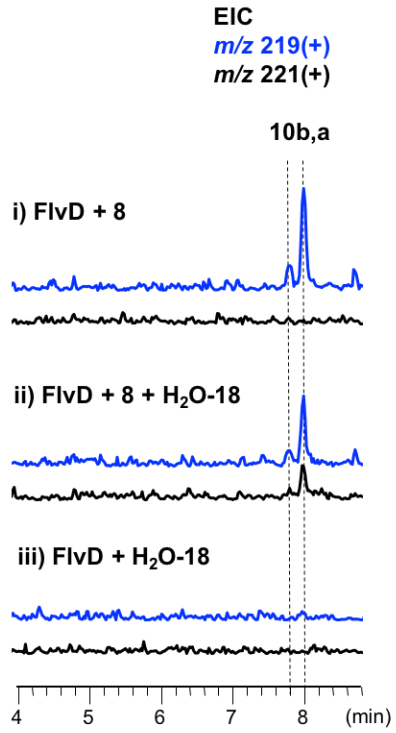


Figure S13. LC/MS analysis of the extracts from an *in vitro* reaction with (i) purified microsomes from yeast expressing FlvD and **8**, (ii) purified microsomes from yeast expressing FlvD and **8** in 85% H_2O-18, and (iii) purified microsomes from yeast expressing FlvD with no substrate in 85% H_2O-18. Extracted LC traces corresponding to the m/z^+ for **10a,b** with water incorporated (m/z^+ = 219) and labeled water incorporated (m/z^+ = 221).

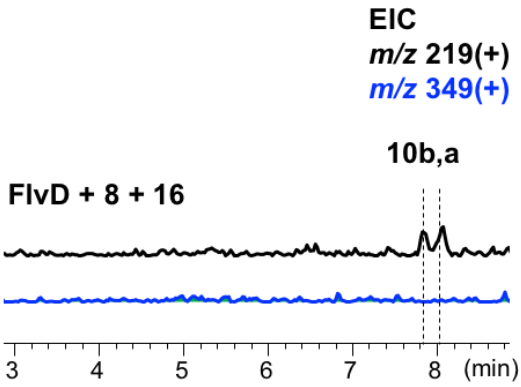


Figure S14. LC/MS analysis of the extracts from an *in vitro* reaction with purified microsomes from yeast expressing FlvD, **8**, and dimethylcadaverine (**16**) without recombinant FlvF. Extracted LC traces corresponding to the m/z^+ for **10a,b** (m/z^+ = 219) and the m/z$^+$ for **12** (m/z^+ = 349).

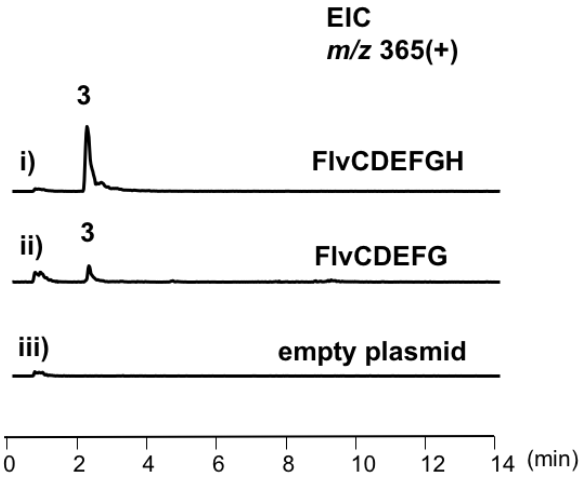


Figure S15. LC/MS analysis of the metabolic extracts from the culture of (i) *A. nidulans* (*An*) expressing FlvC-H, (ii) *An* expressing FlvC-G, and (iii) *An* with an empty vector (negative control). Extracted LC traces corresponding to the m/z^+ for **3** (m/z^+ = 365).

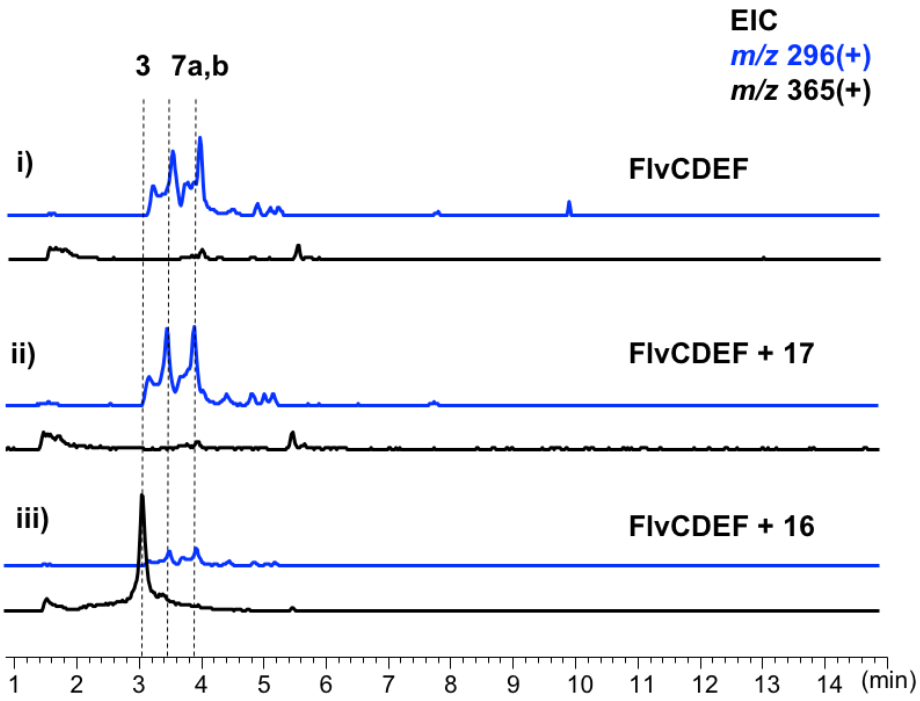


Figure S16. LC/MS analysis of the metabolic extracts from the culture of (i) *A. nidulans* (*An*) expressing FlvC-F, (ii) feeding dimethyl lysine (**17**) to *An* expressing FlvC-F, and (iii) feeding dimethyl cadaverine (**16**) to *An* expressing FlvC-F. Extracted LC traces corresponding to the m/z^+ for **3** (m/z^+ = 365) and **7a,b** (m/z^+ = 296).

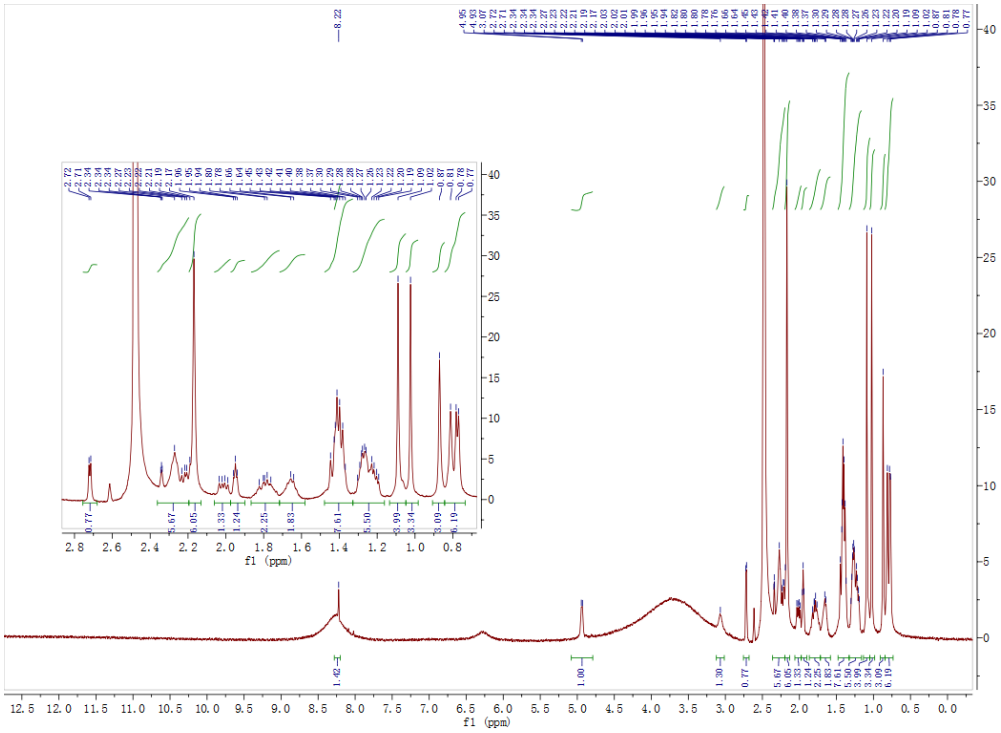


Figure S17. The ^{1}H NMR spectrum of compound **1** in DMSO-d_6 (500 MHz)

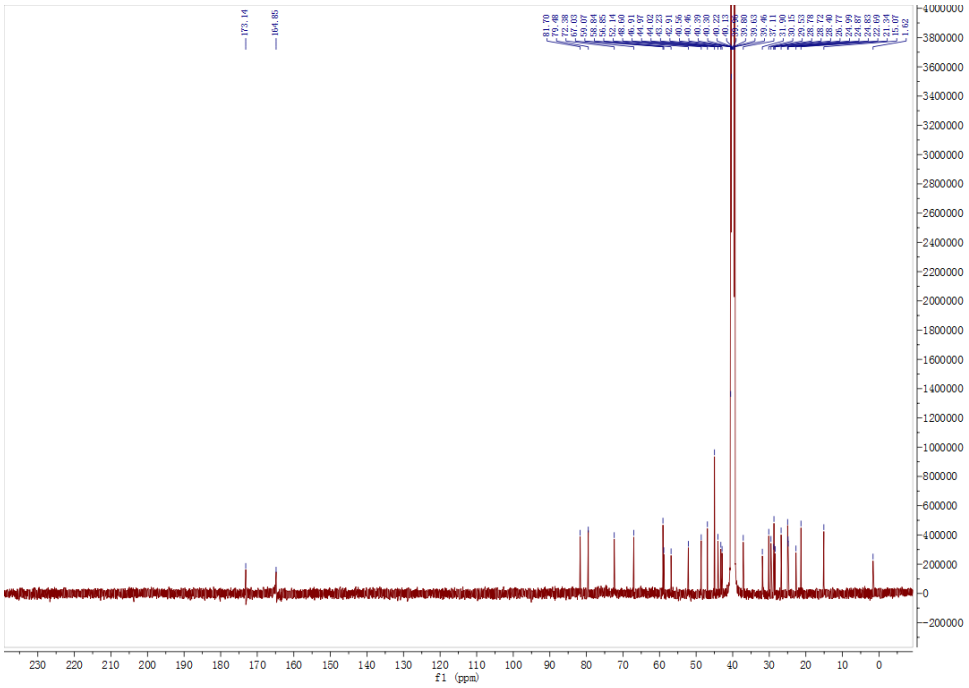


Figure S18. The ^{13}C NMR spectrum of compound **1** in DMSO-d_6 (125 MHz)

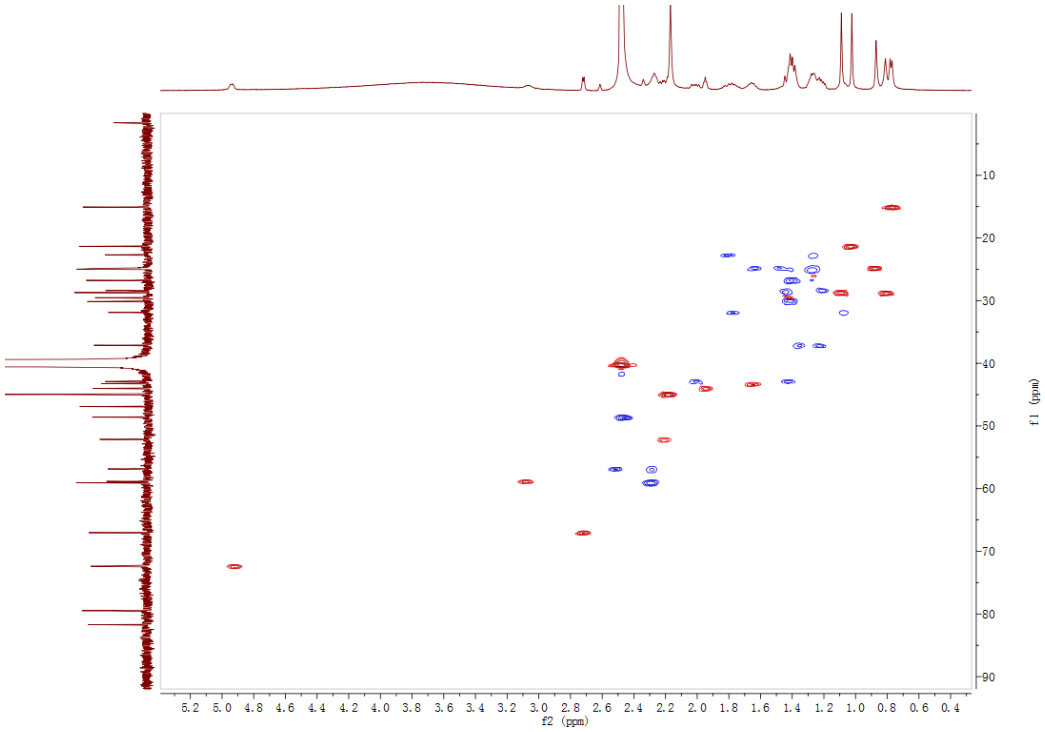


Figure S19. The HSQC spectrum of compound **1** in DMSO-d_6

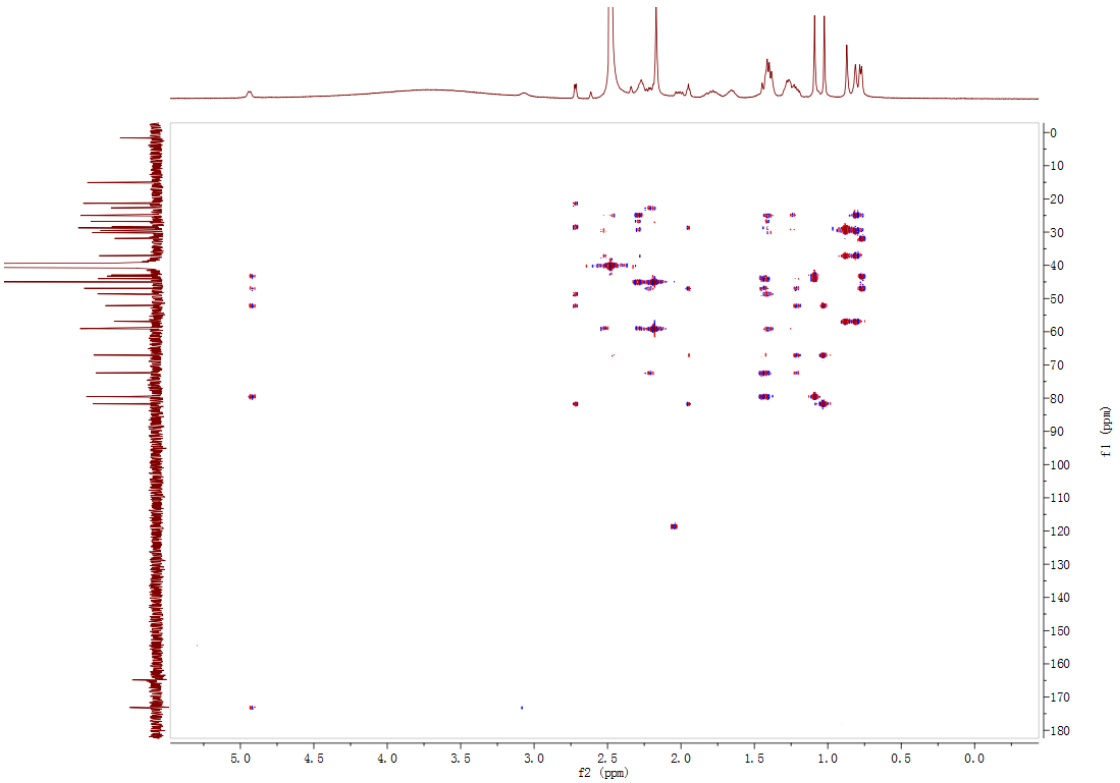


Figure S20. The HMBC spectrum of compound **1** in DMSO-d_6

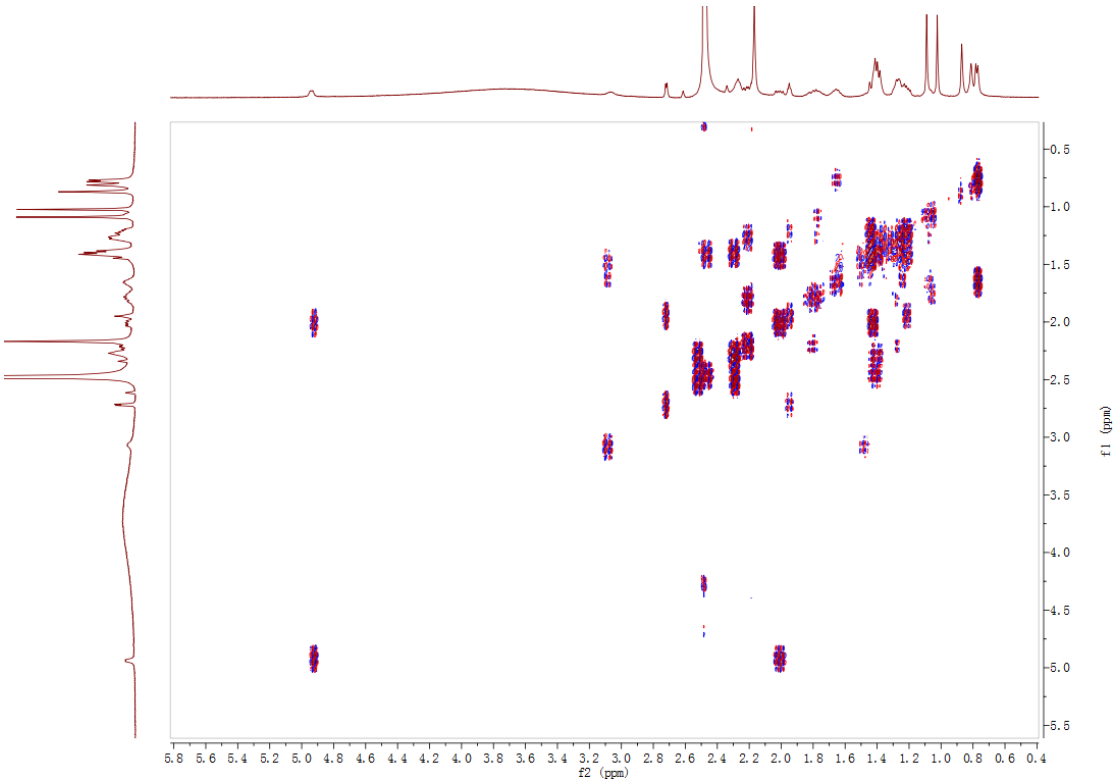


Figure S21. The ^{1}H-^{1}H COSY spectrum of compound **1** in DMSO-d_6

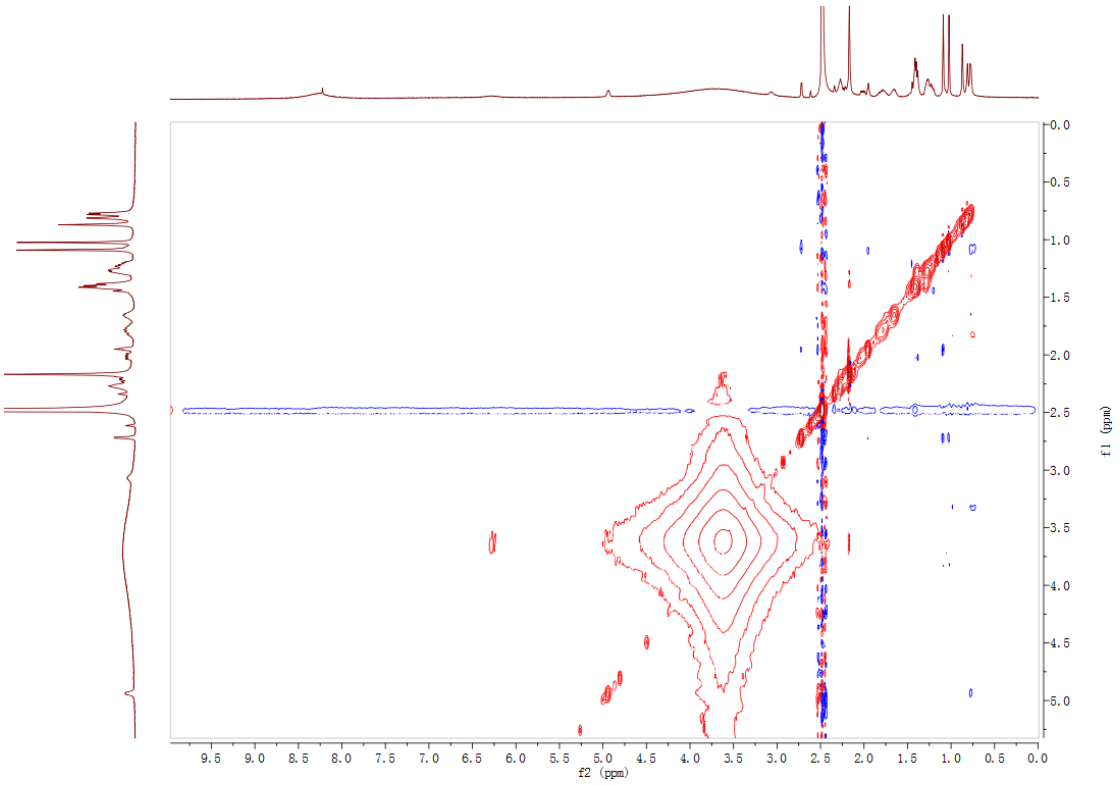


Figure S22. The NOESY spectrum of compound **1** in DMSO-*d*

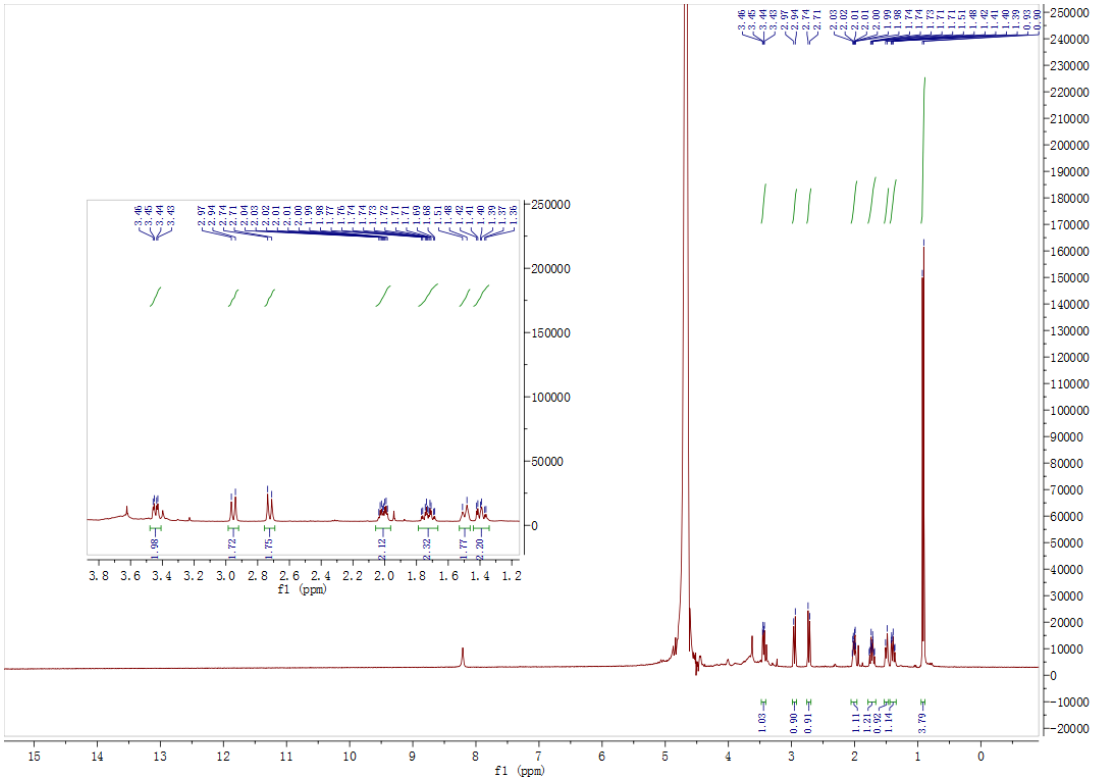


Figure S23. The ^{1}H NMR spectrum of compound **2** in D_2O (500 MHz)

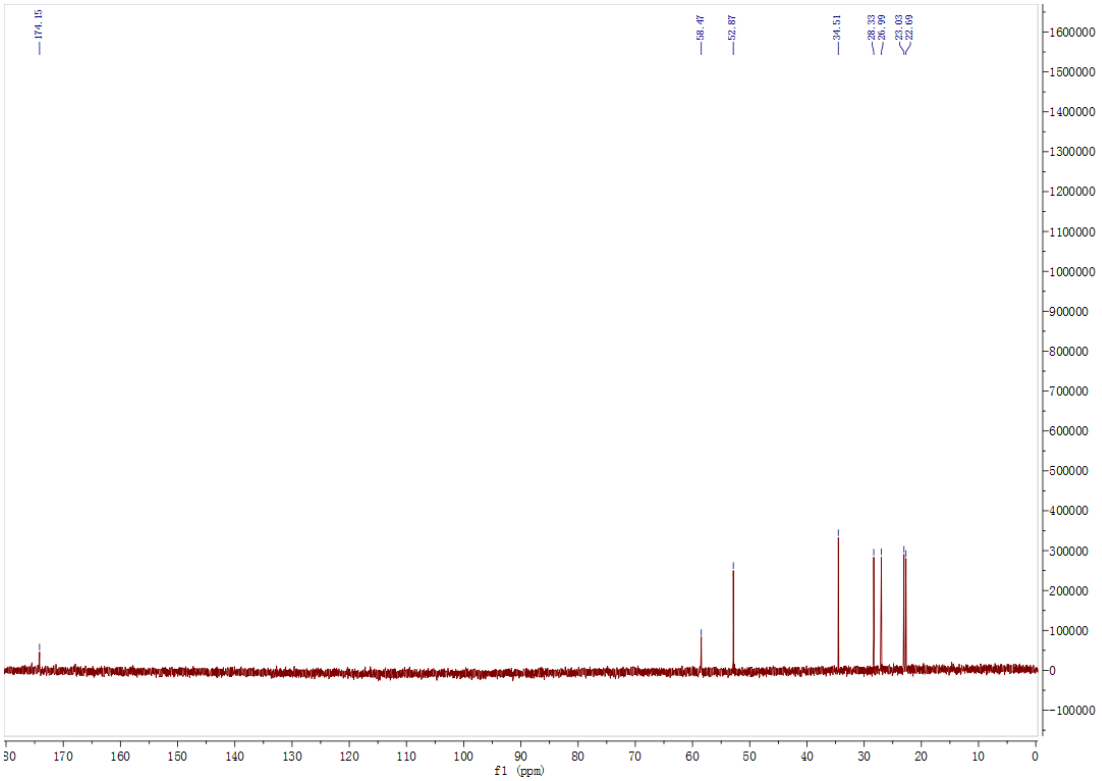


Figure S24. The ^{13}C NMR spectrum of compound **2** in D_2O (125 MHz)

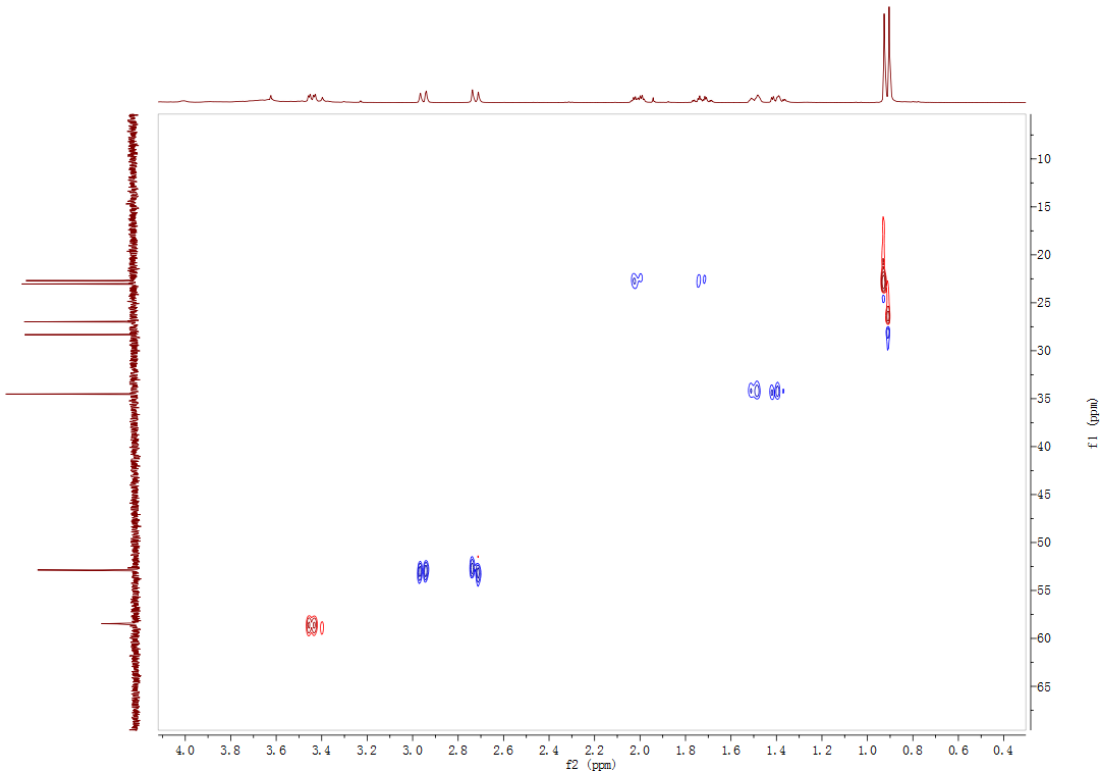


Figure S25. The HSQC spectrum of compound **2** in D_2O

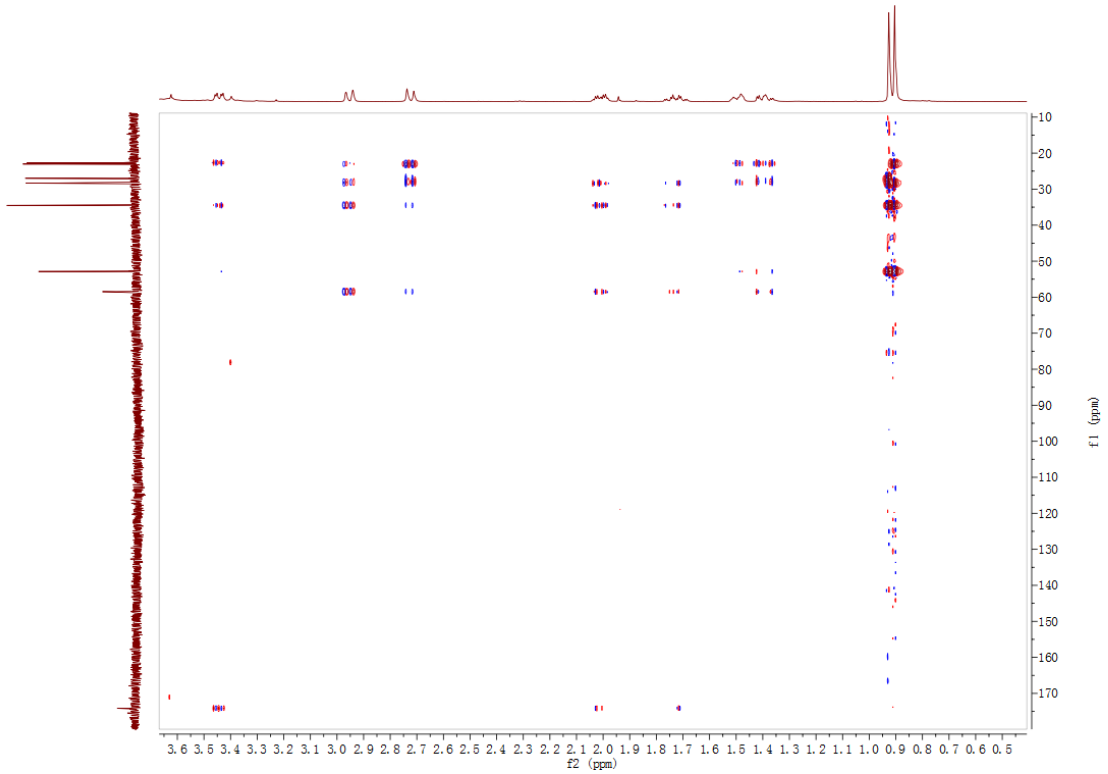


Figure S26. The HMBC spectrum of compound **2** in D_2O

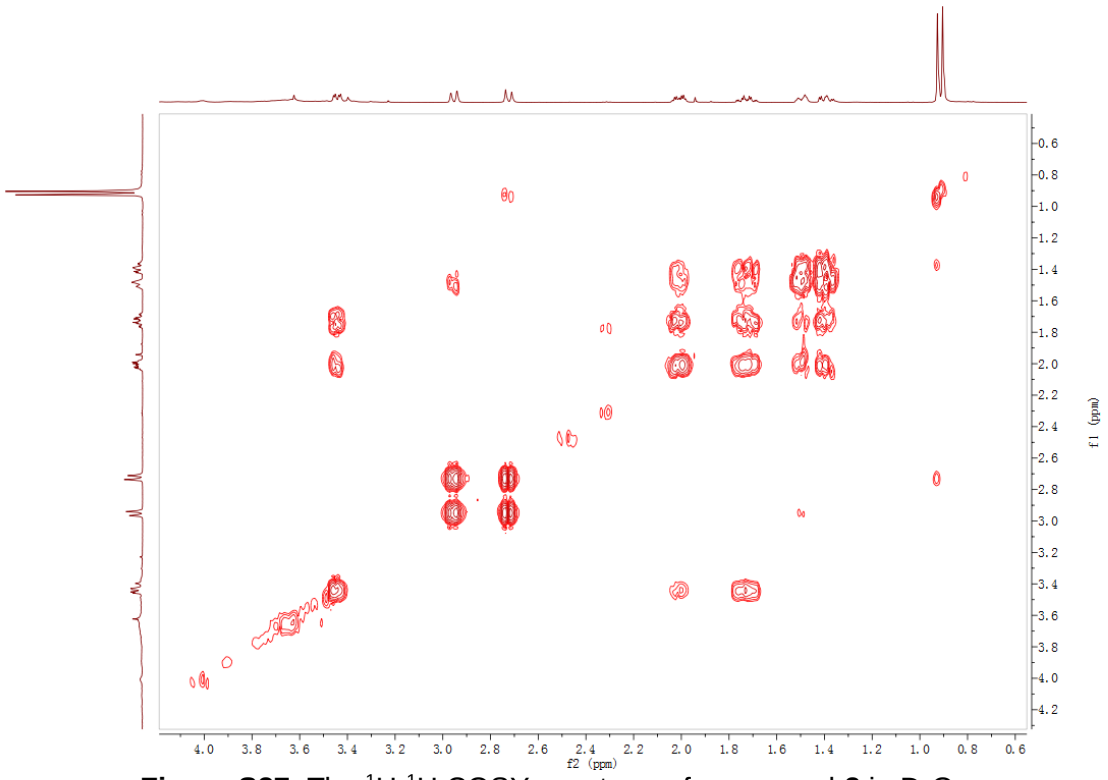


Figure S27. The ^{1}H-^{1}H COSY spectrum of compound **2** in D_2O

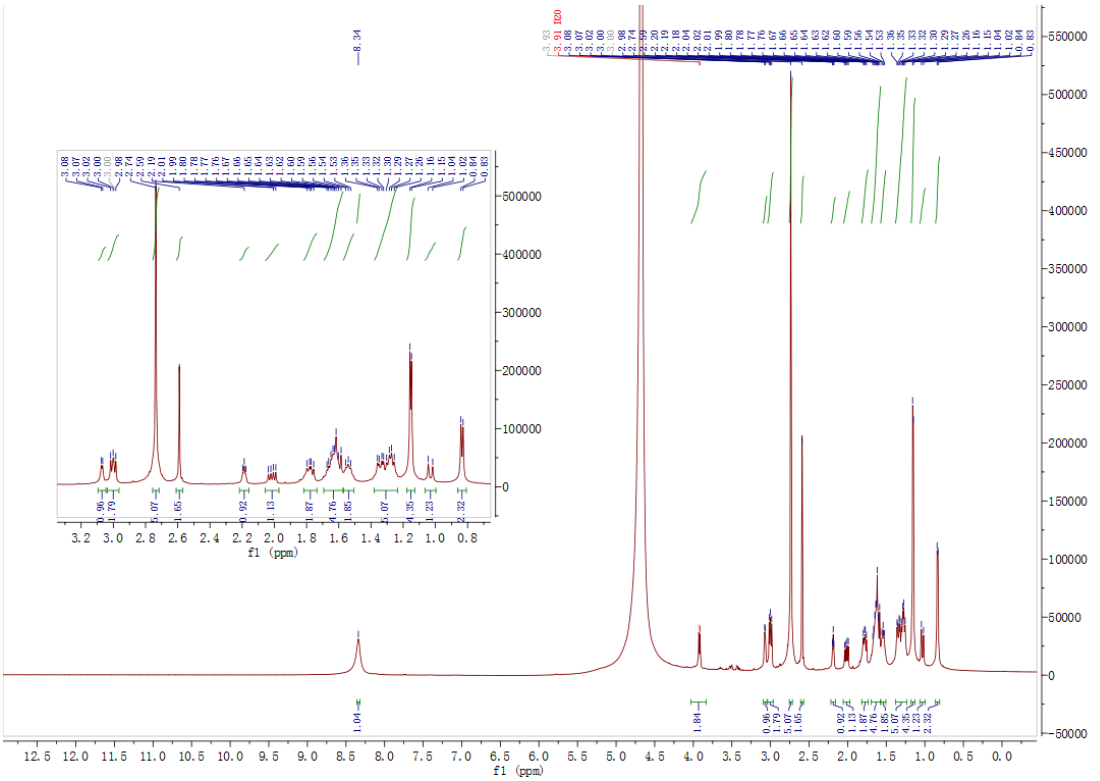


Figure S28. The ^{1}H NMR spectrum of compound **3** in D_2O (500 MHz)

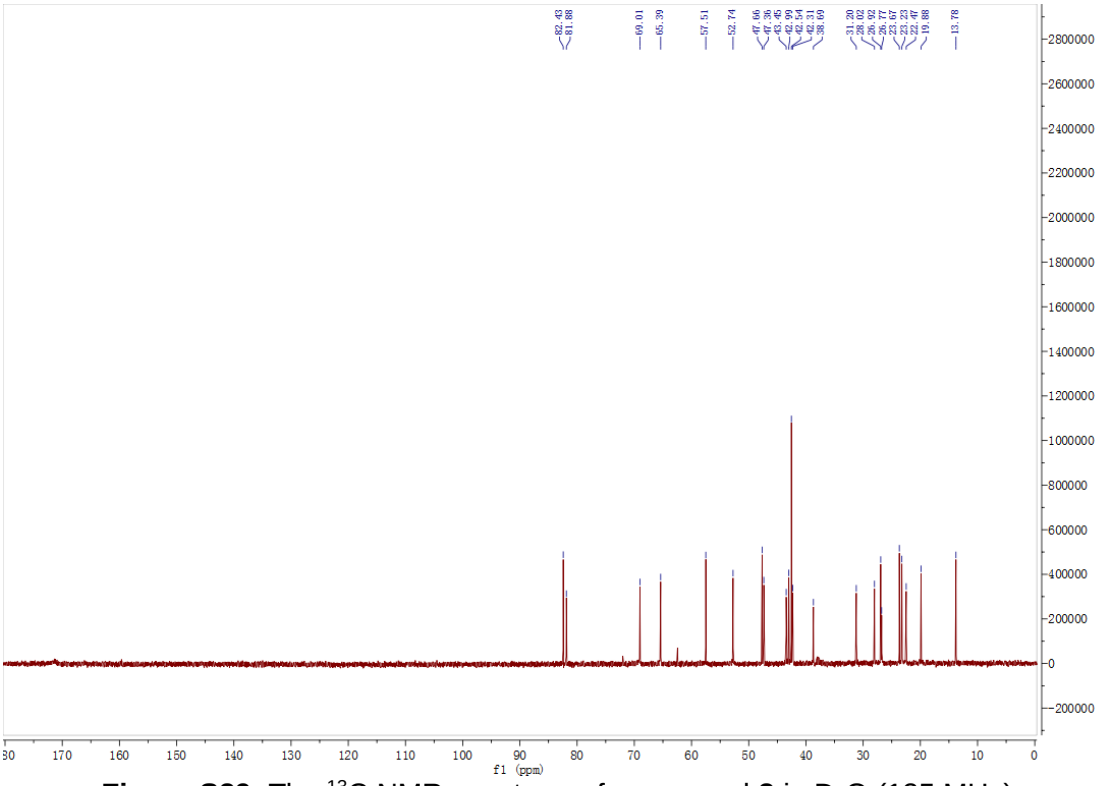


Figure S29. The ^{13}C NMR spectrum of compound **3** in D_2O (125 MHz)

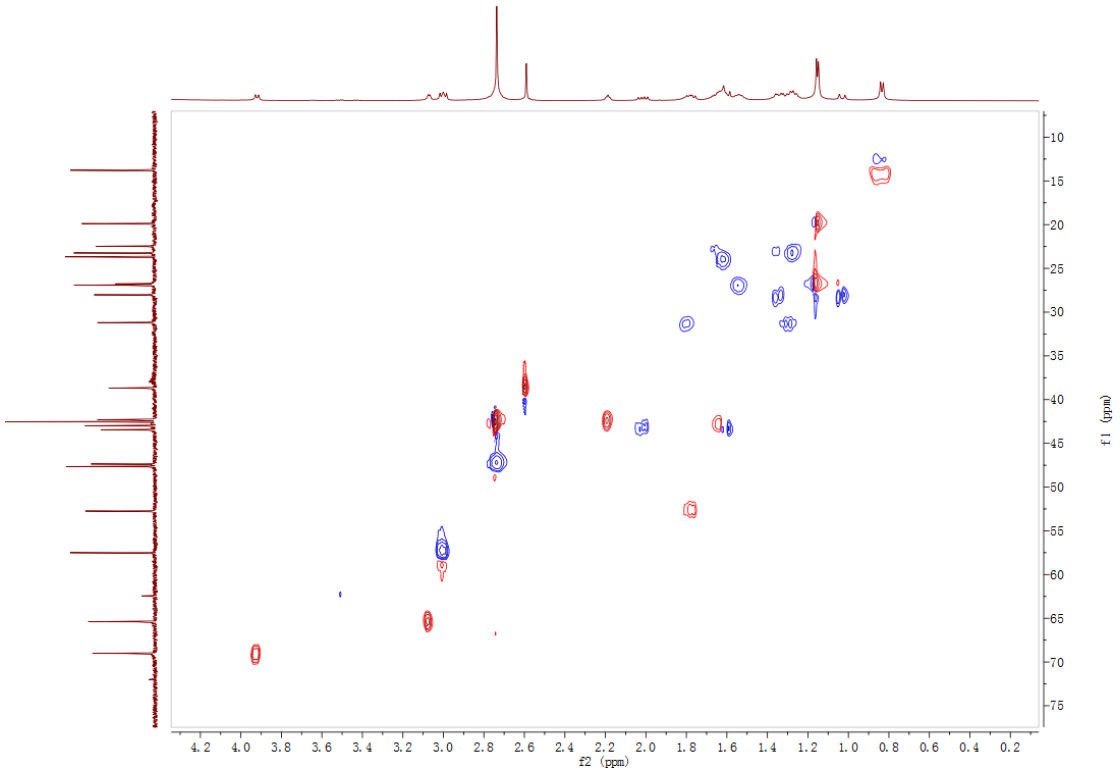


Figure S30. The HSQC spectrum of compound **3** in D_2O

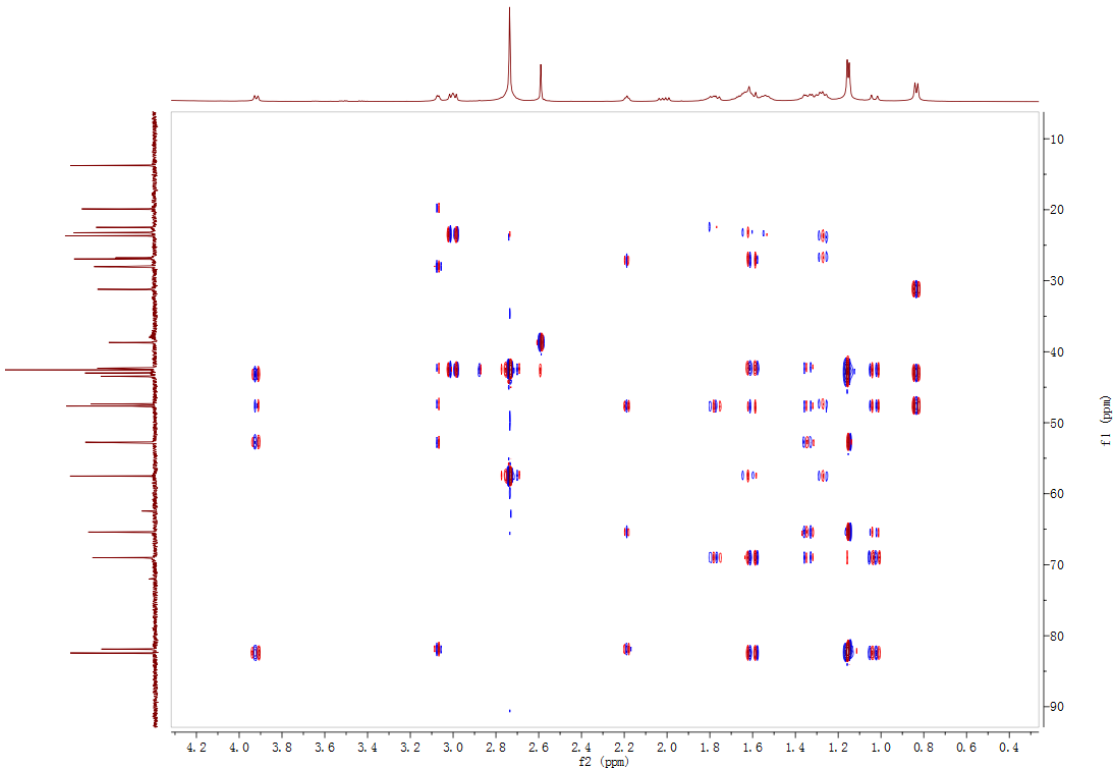


Figure S31. The HMBC spectrum of compound **3** in D_2O

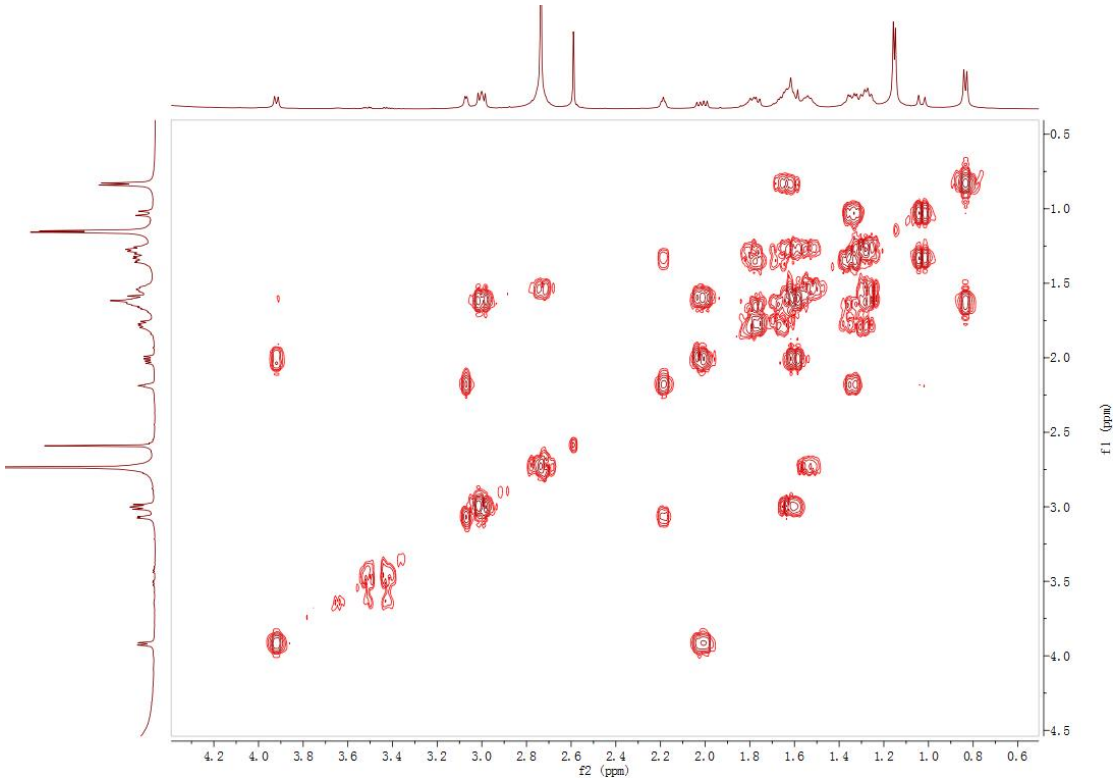


Figure S32. The ^{1}H-^{1}H COSY spectrum of compound **3** in D_2O

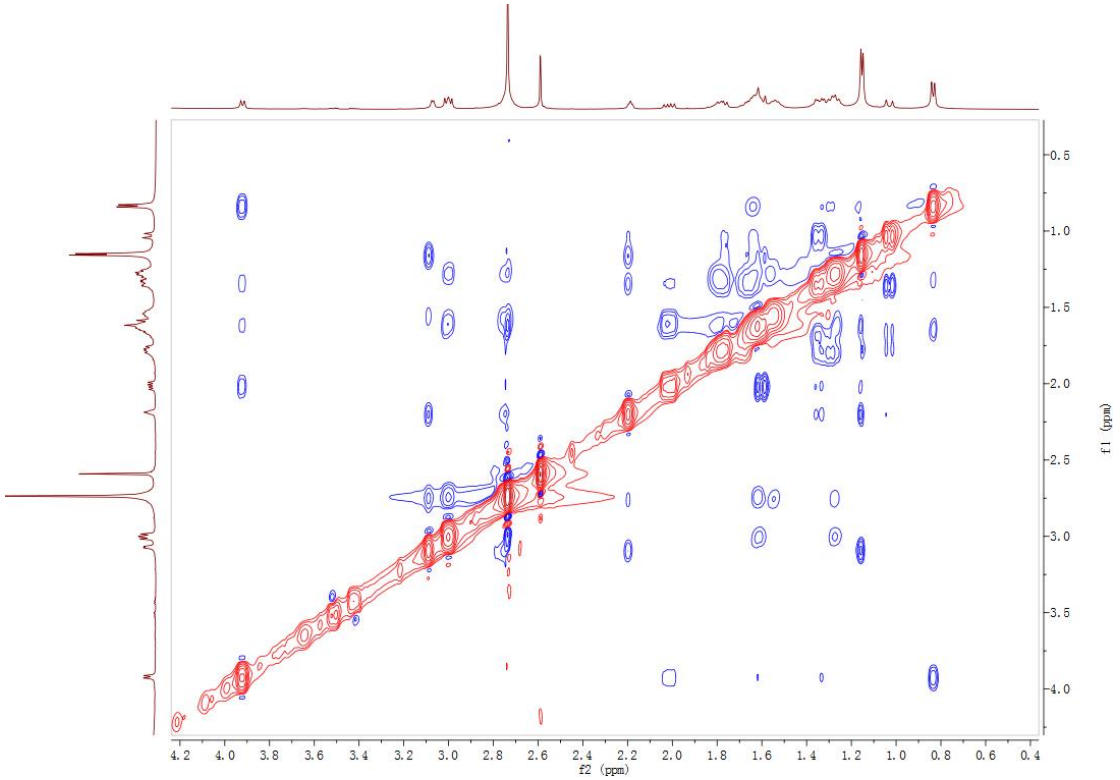


Figure S33. The NOESY spectrum of compound **3** in D_2O

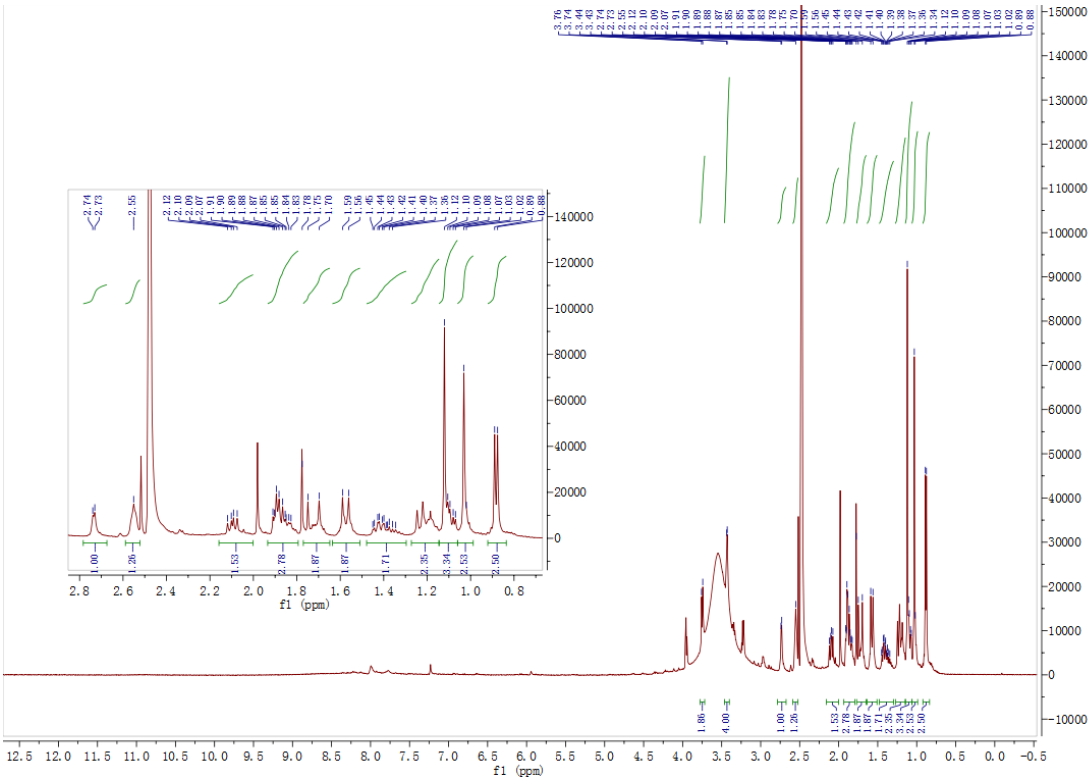


Figure S34. The ^{1}H NMR spectrum of compound **7a** in DMSO-d_6 (500 MHz)

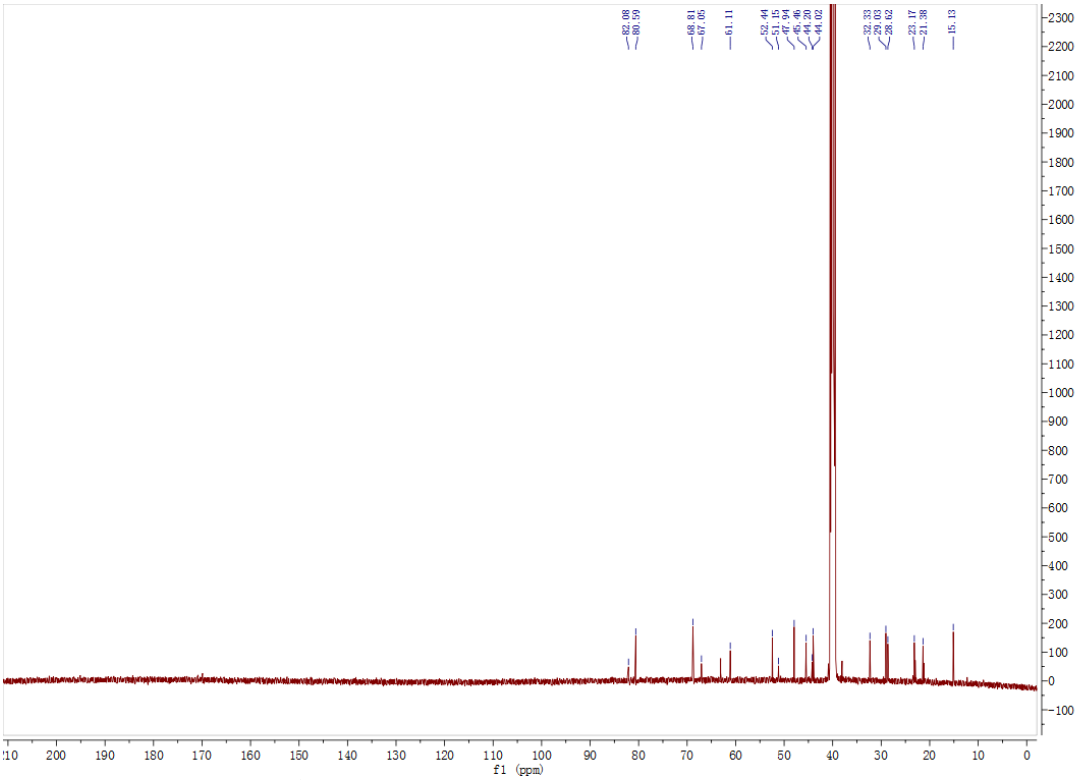


Figure S35. The ^{13}C NMR spectrum of compound **7a** in DMSO-d_6 (125 MHz)

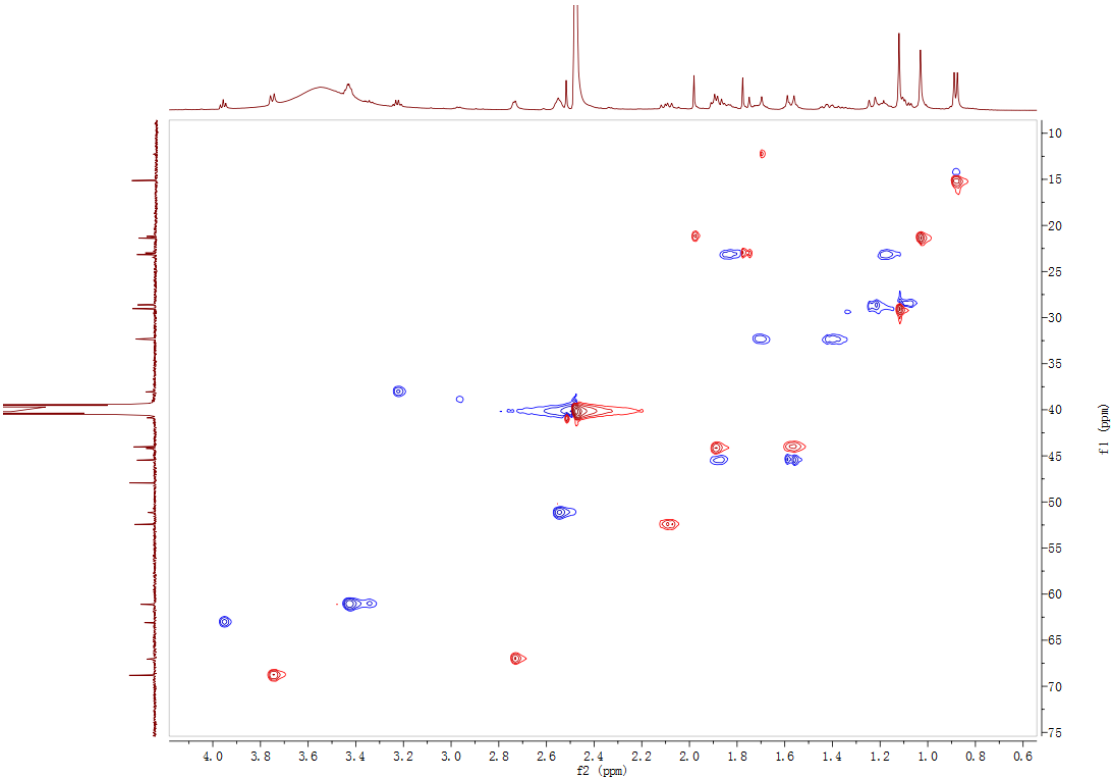


Figure S36. The HSQC spectrum of compound **7a** in DMSO-d_6

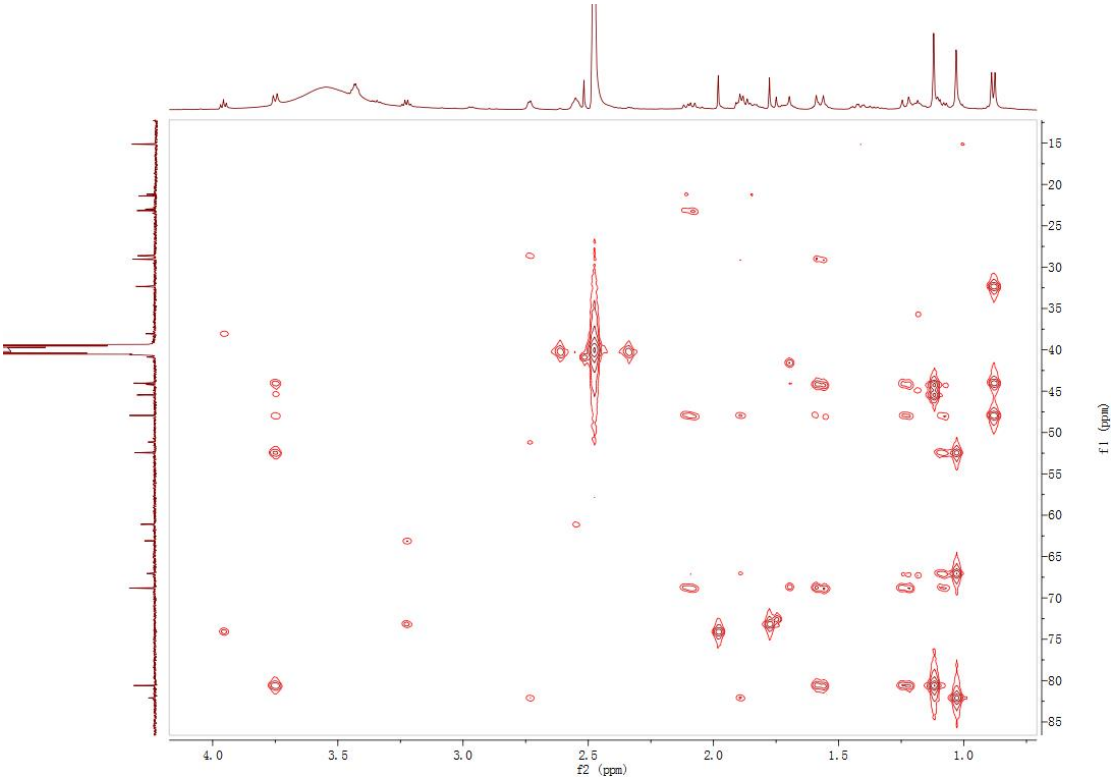


Figure S37. The HMBC spectrum of compound **7a** in DMSO-d_6

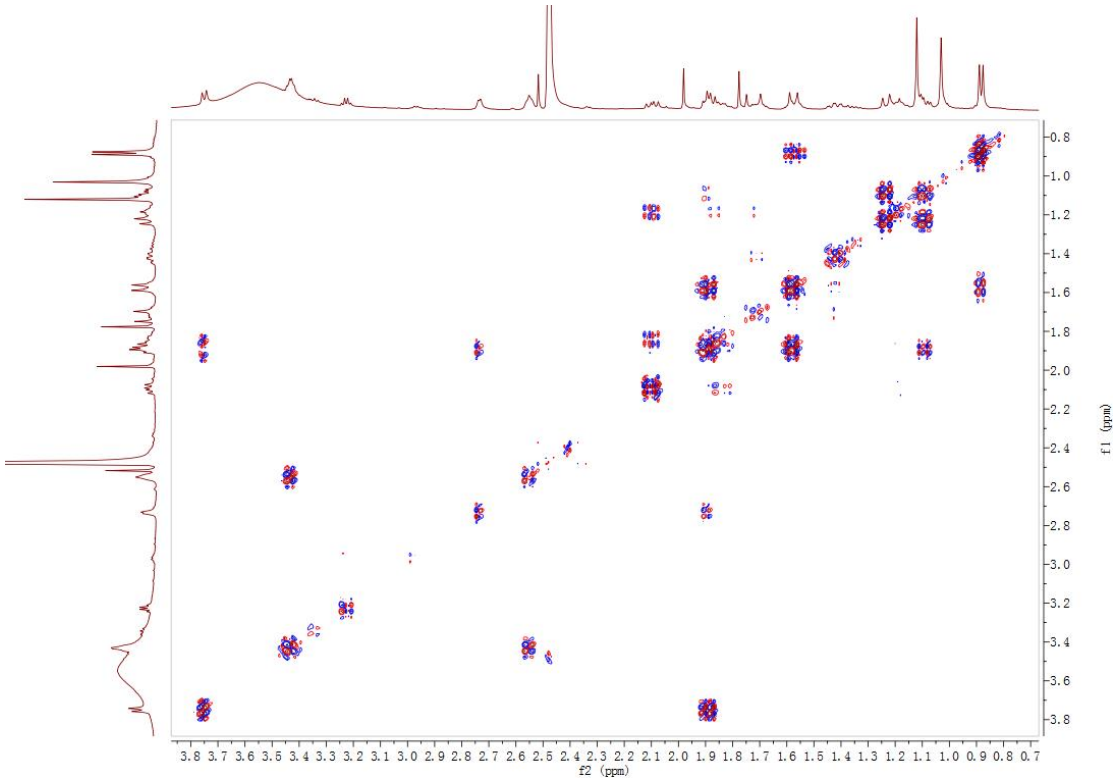


Figure S38. The ^{1}H-^{1}H COSY spectrum of compound **7a** in DMSO-d_6

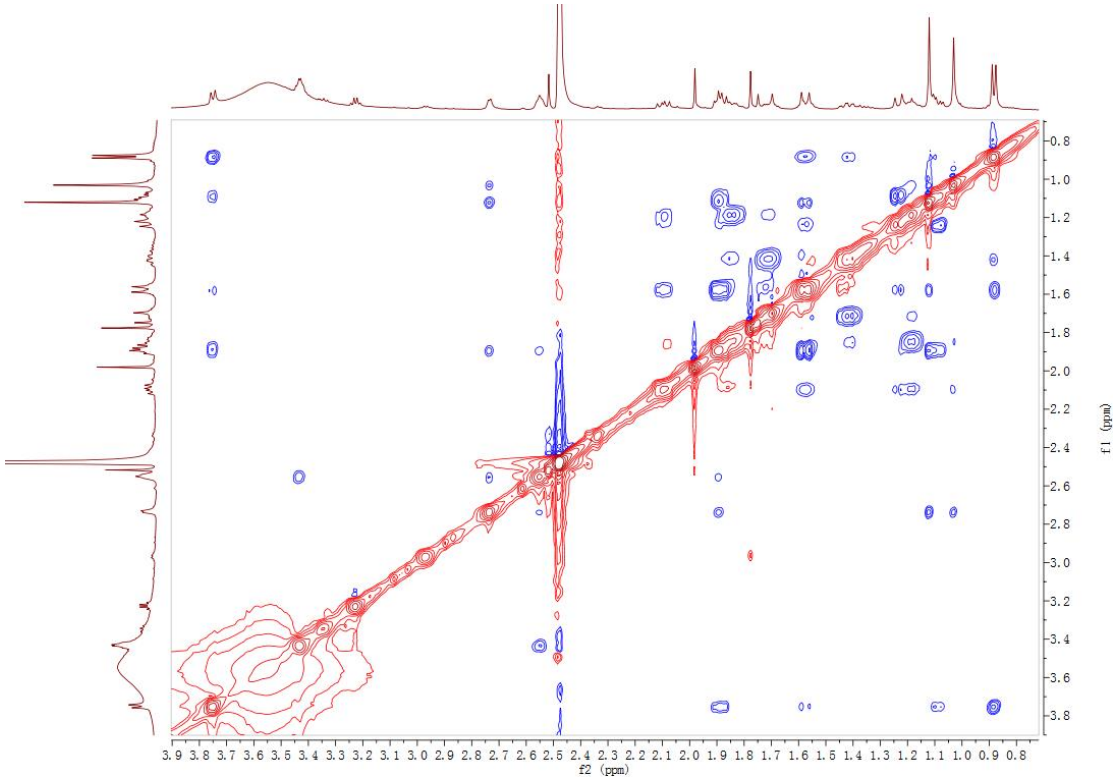


Figure S39. The NOESY spectrum of compound **7a** in DMSO-d_6

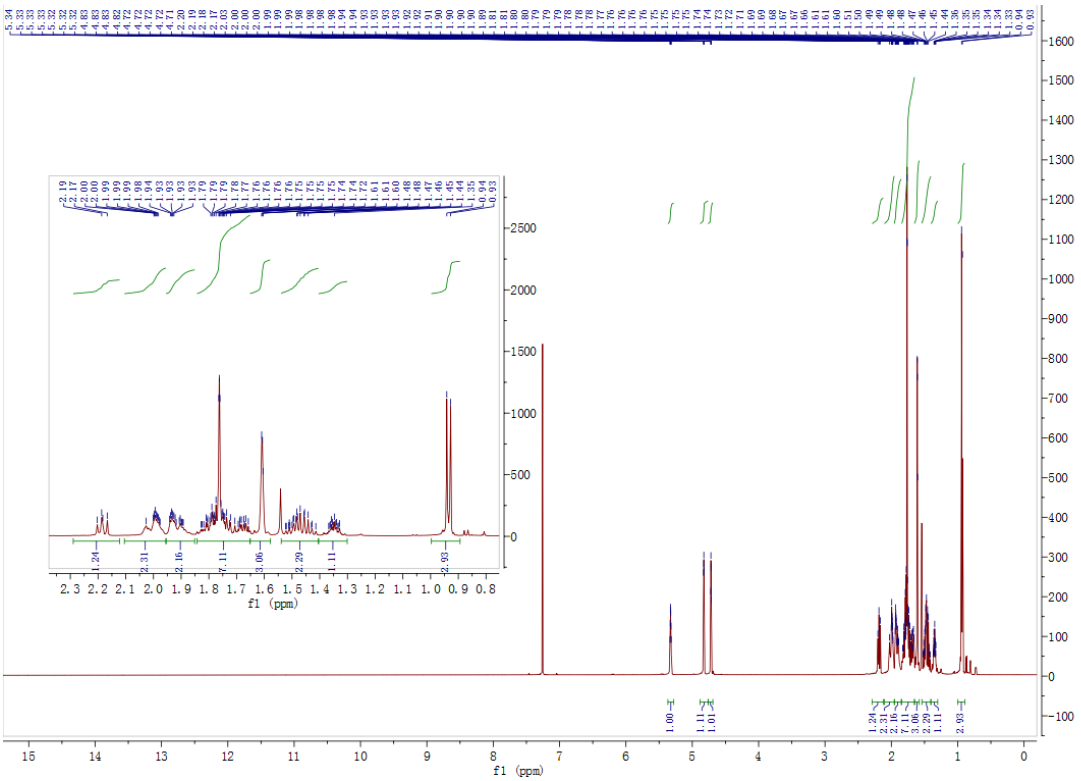


Figure S40. The ^{1}H NMR spectrum of compound **8** in $CDCl_3$ (500 MHz)

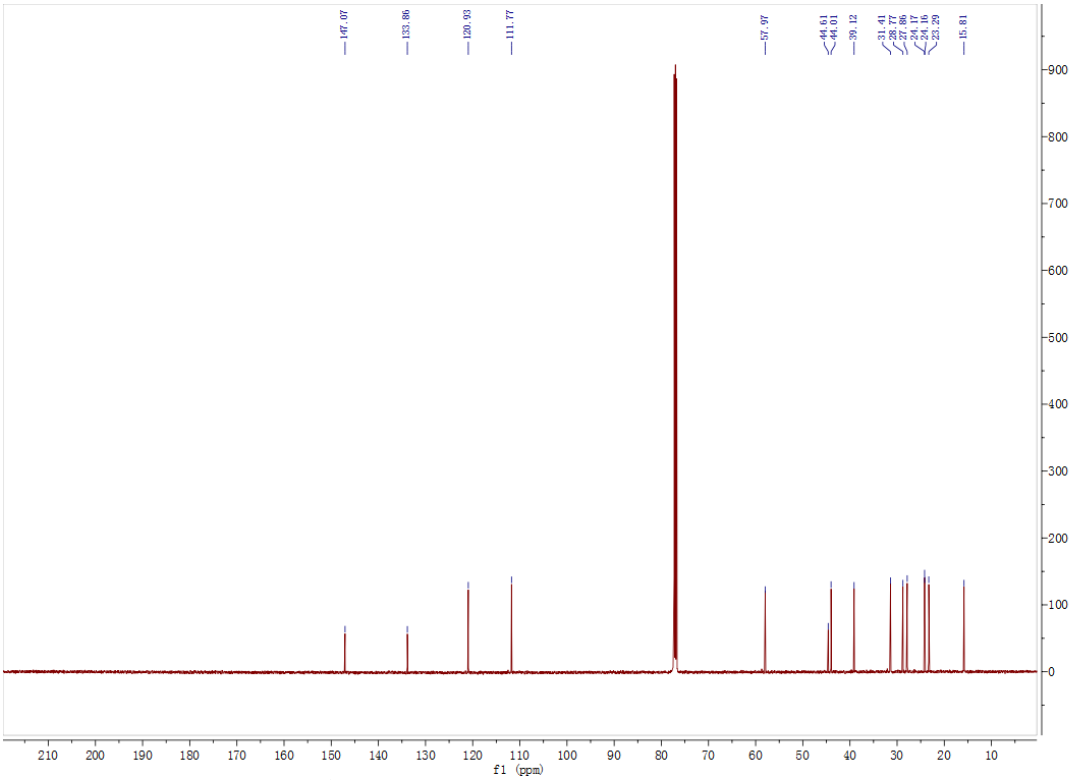


Figure S41. The ^{13}C NMR spectrum of compound **8** in $CDCl_3$ (125 MHz)

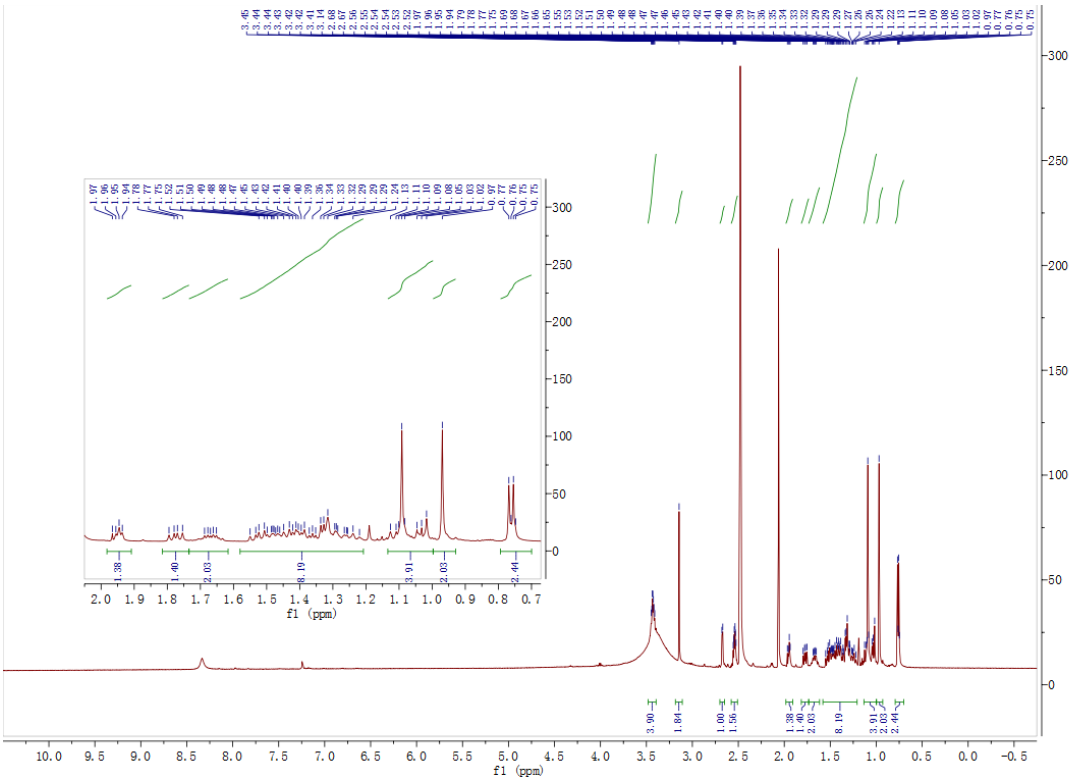


Figure S42. The ^{1}H NMR spectrum of compound **9a** in DMSO-d_6 (500 MHz)

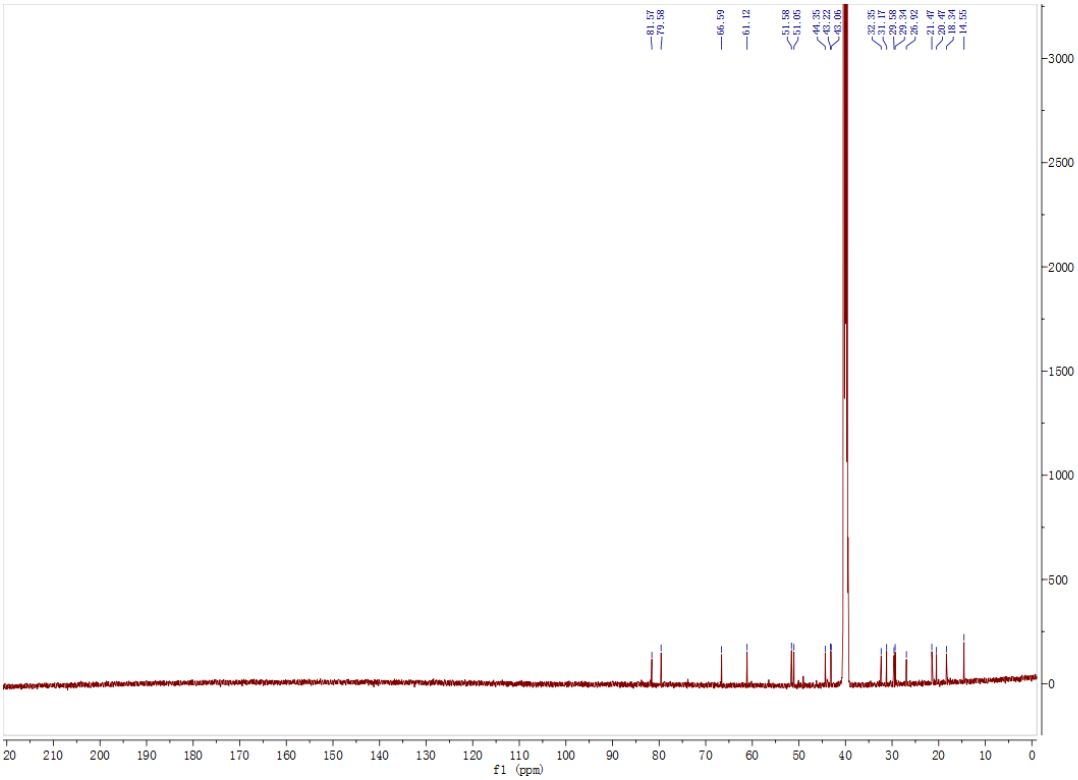


Figure S43. The ^{13}C NMR spectrum of compound **9a** in DMSO-d_6 (125 MHz)

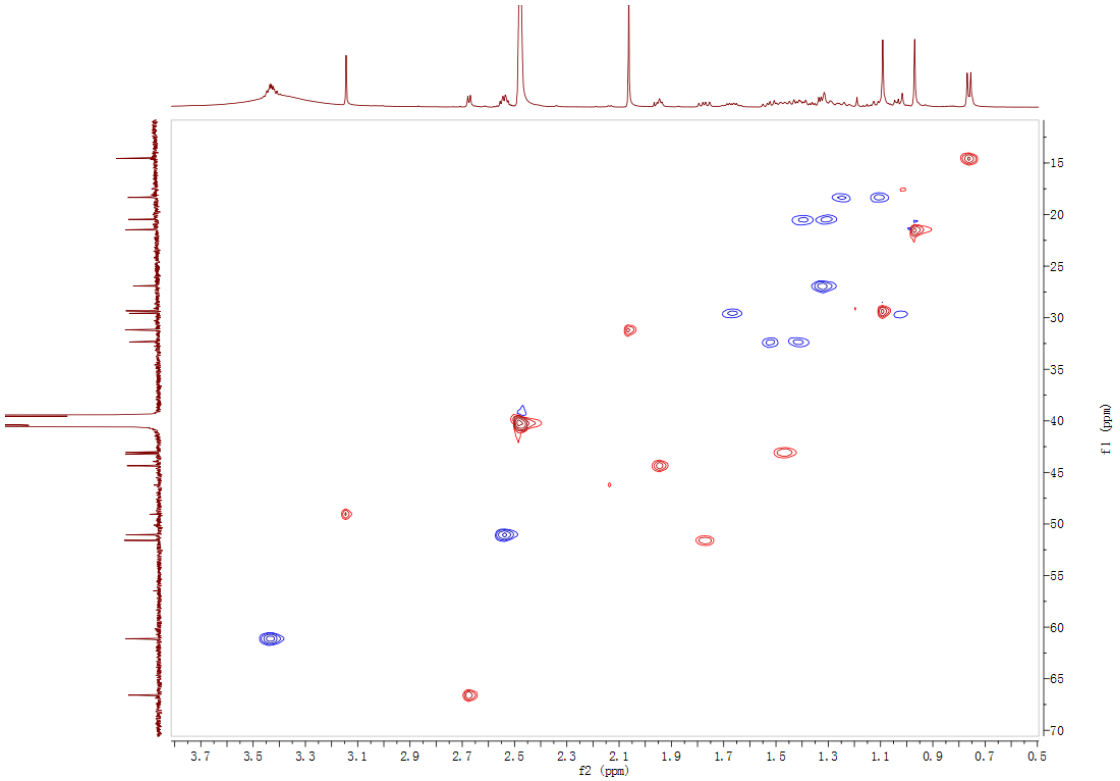


Figure S44. The HSQC spectrum of compound **9a** in DMSO-d_6

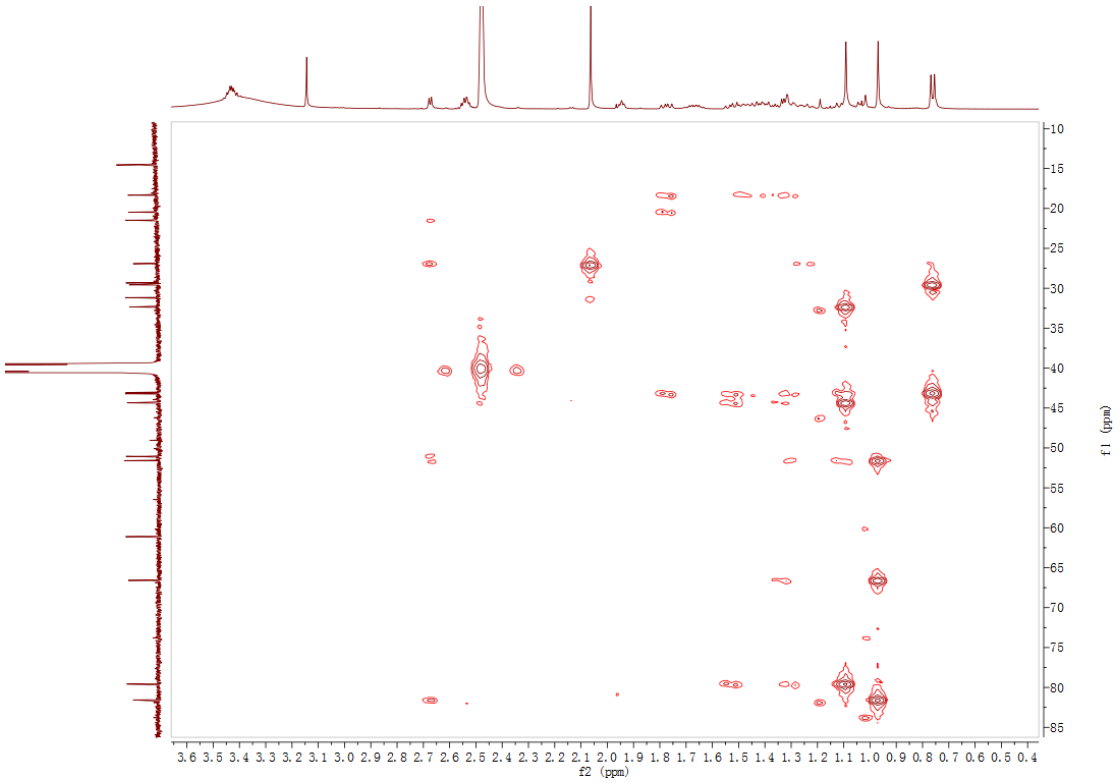


Figure S45. The HMBC spectrum of compound **9a** in DMSO-d_6

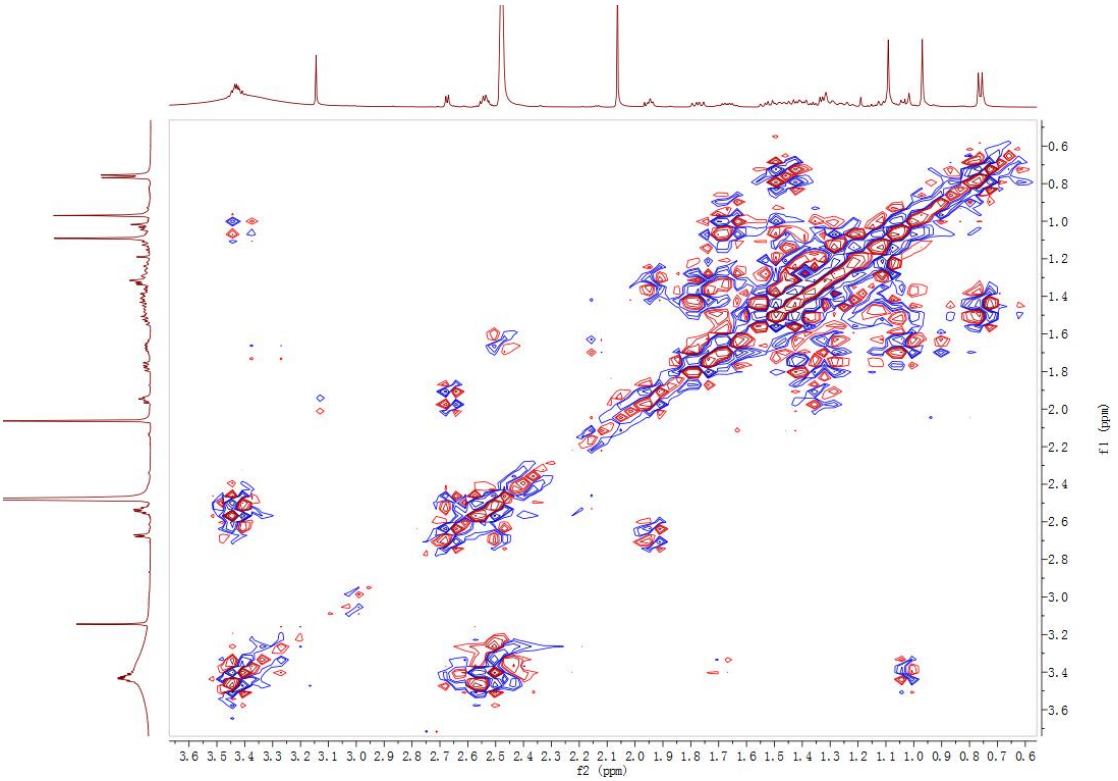


Figure S46. The ^{1}H-^{1}H COSY spectrum of compound **9a** in DMSO-d_6

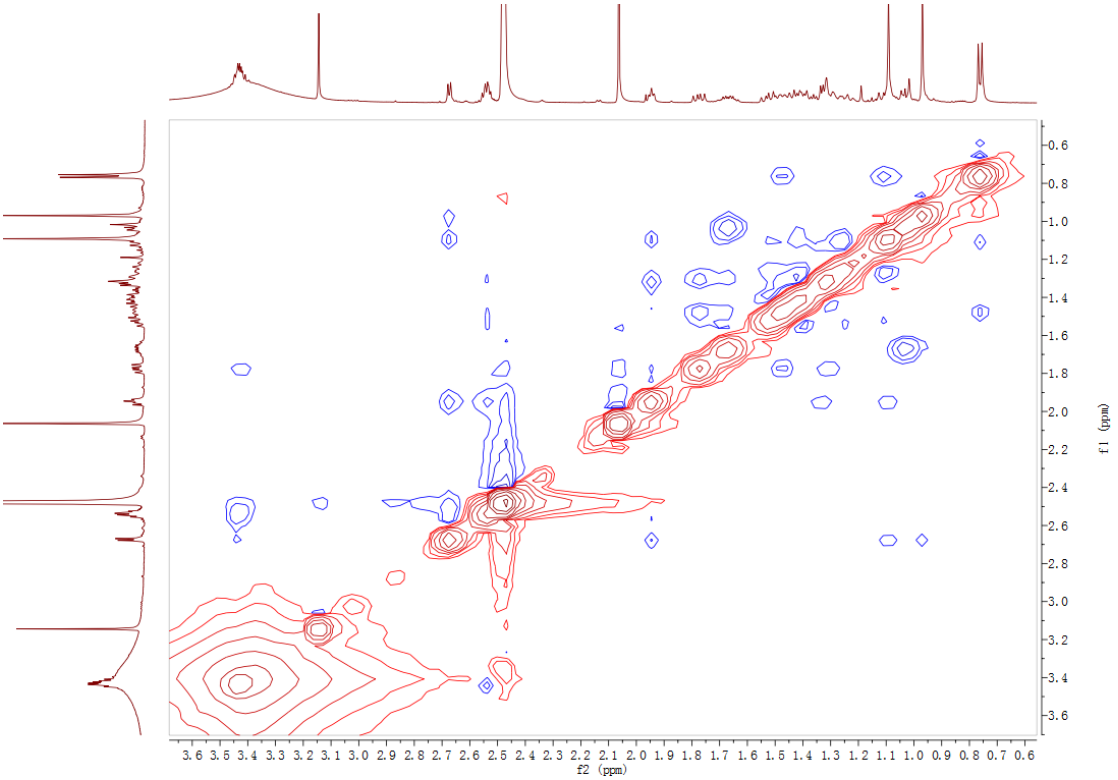


Figure S47. The NOESY spectrum of compound **9a** in DMSO-d_6

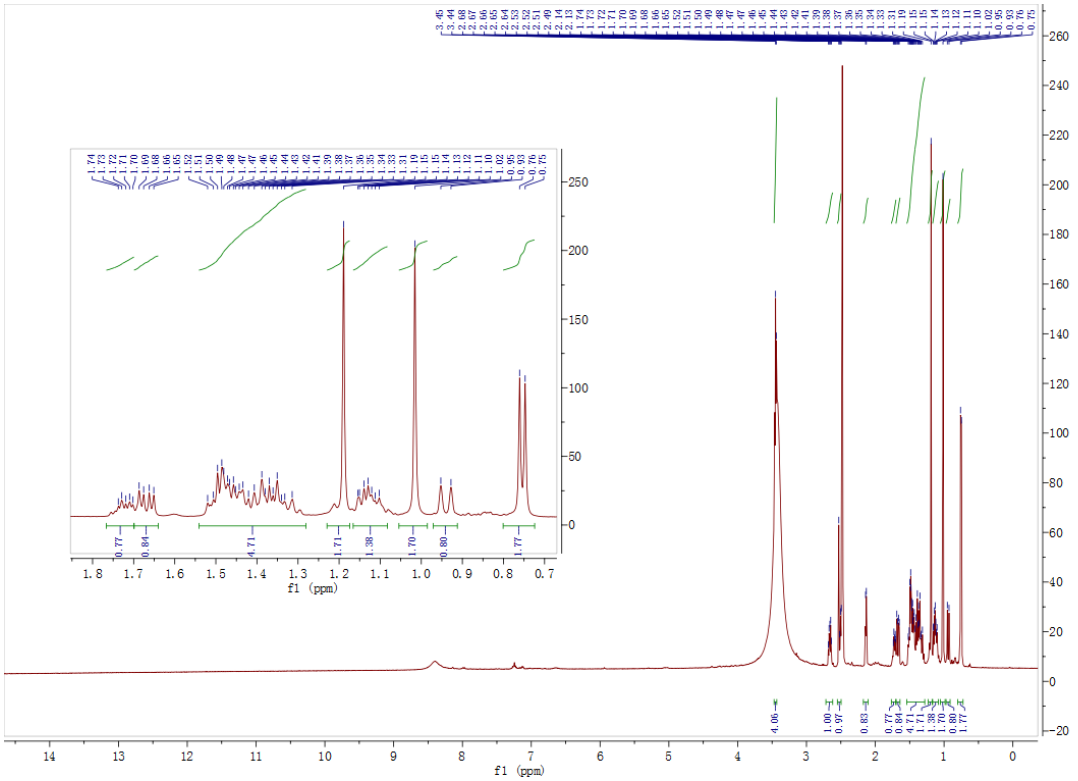


Figure S48. The ^{1}H NMR spectrum of compound **9b** in DMSO-d_6 (500 MHz)

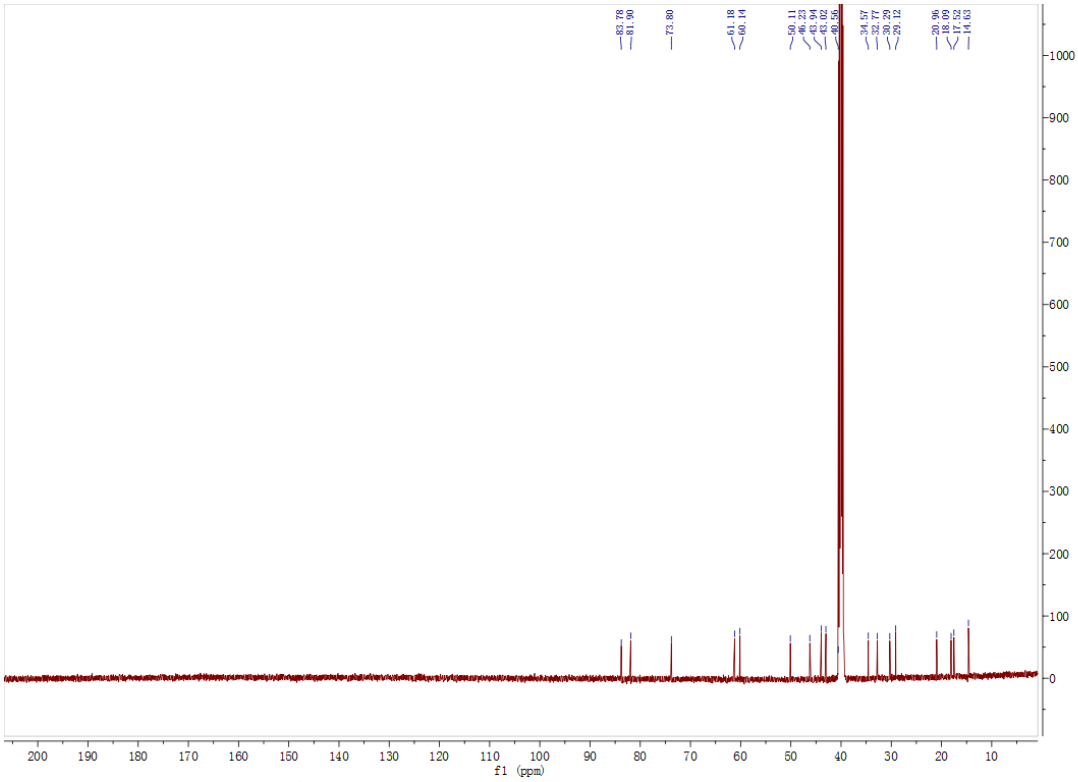


Figure S49. The ^{13}C NMR spectrum of compound **9b** in DMSO-d_6 (125 MHz)

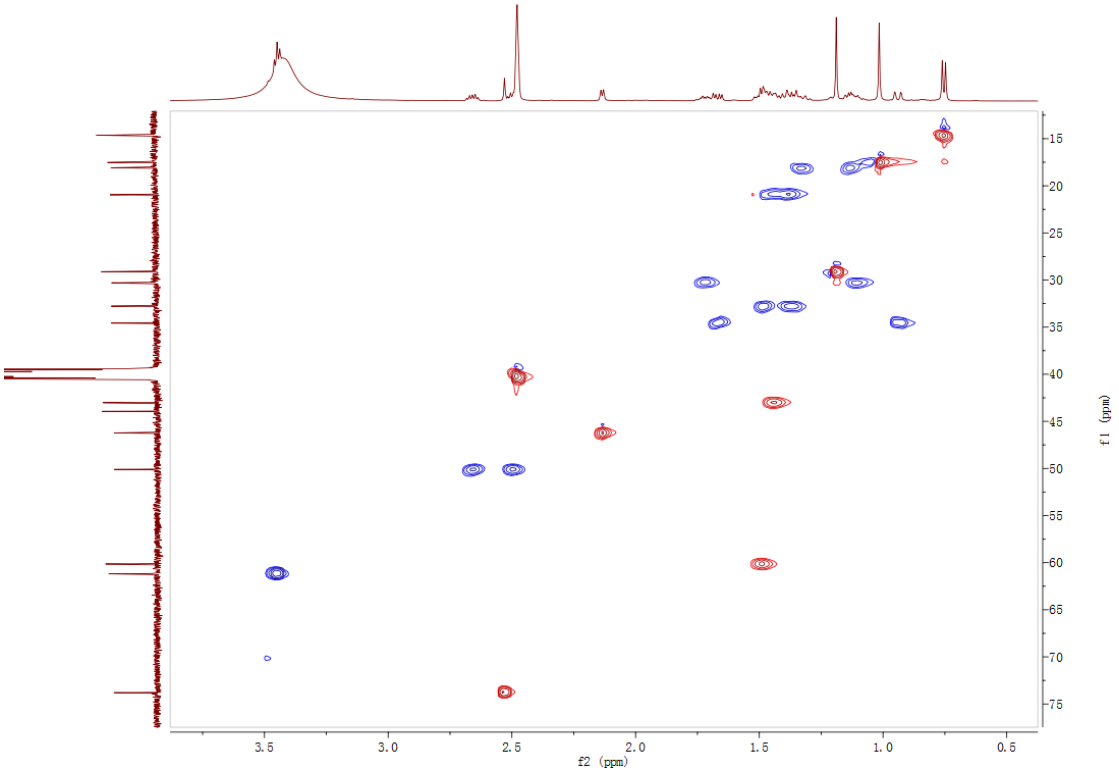


Figure S50. The HSQC spectrum of compound **9b** in DMSO-d_6

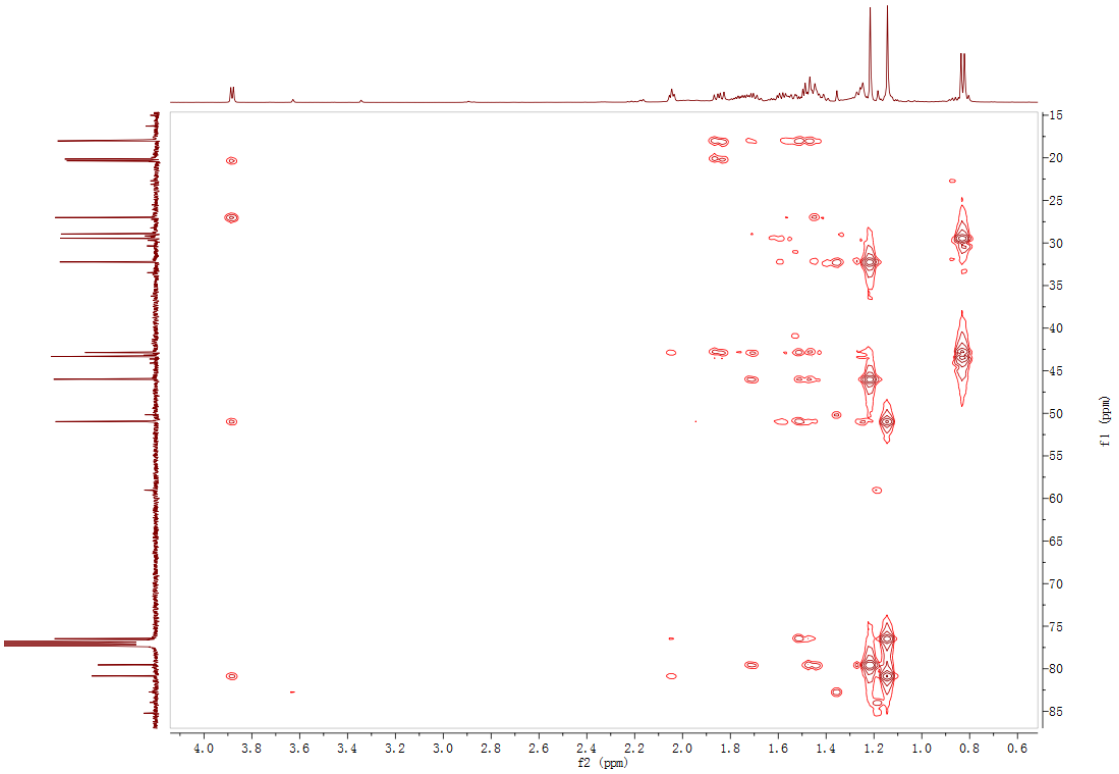


Figure S51. The HMBC spectrum of compound **9b** in DMSO-d_6

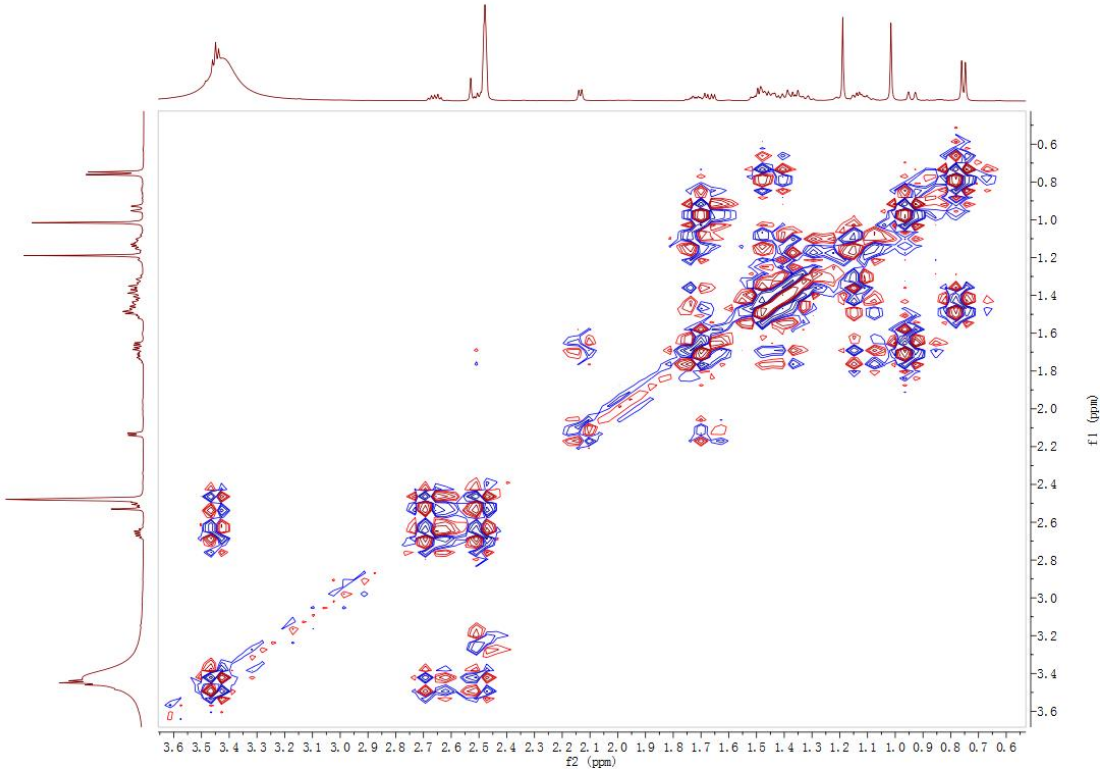


Figure S52. The ^{1}H-^{1}H COSY spectrum of compound **9b** in DMSO-d_6

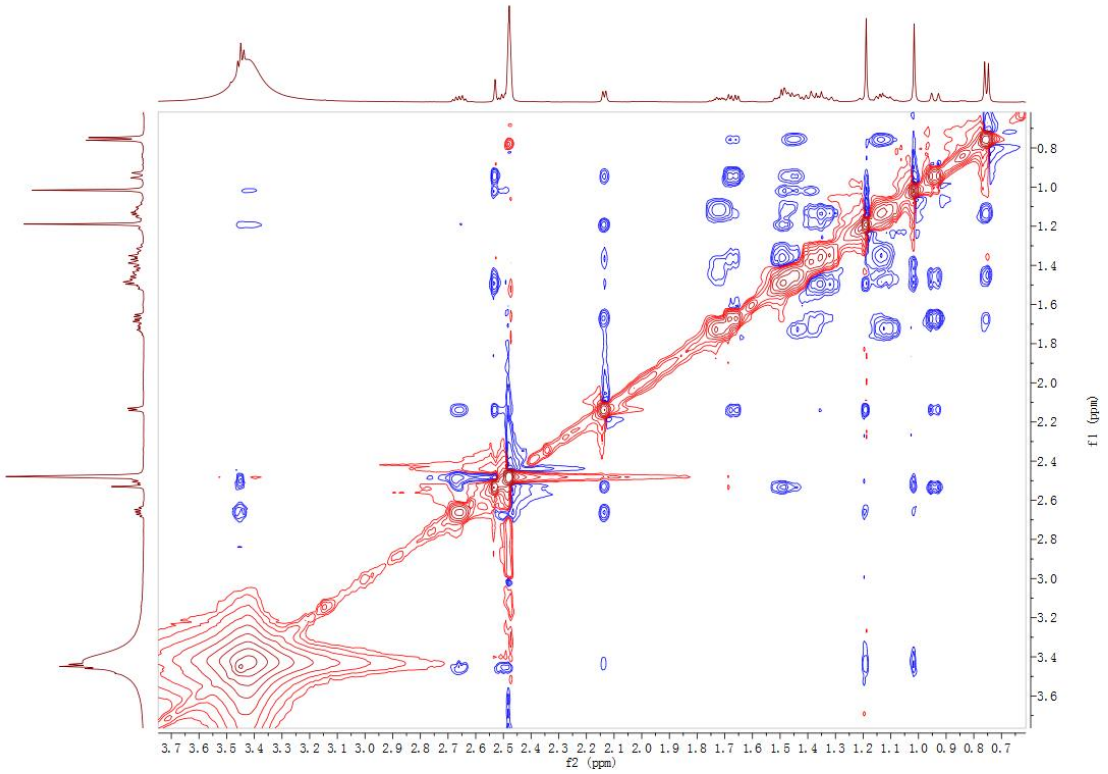


Figure S53. The NOESY spectrum of compound **9b** in DMSO-d_6

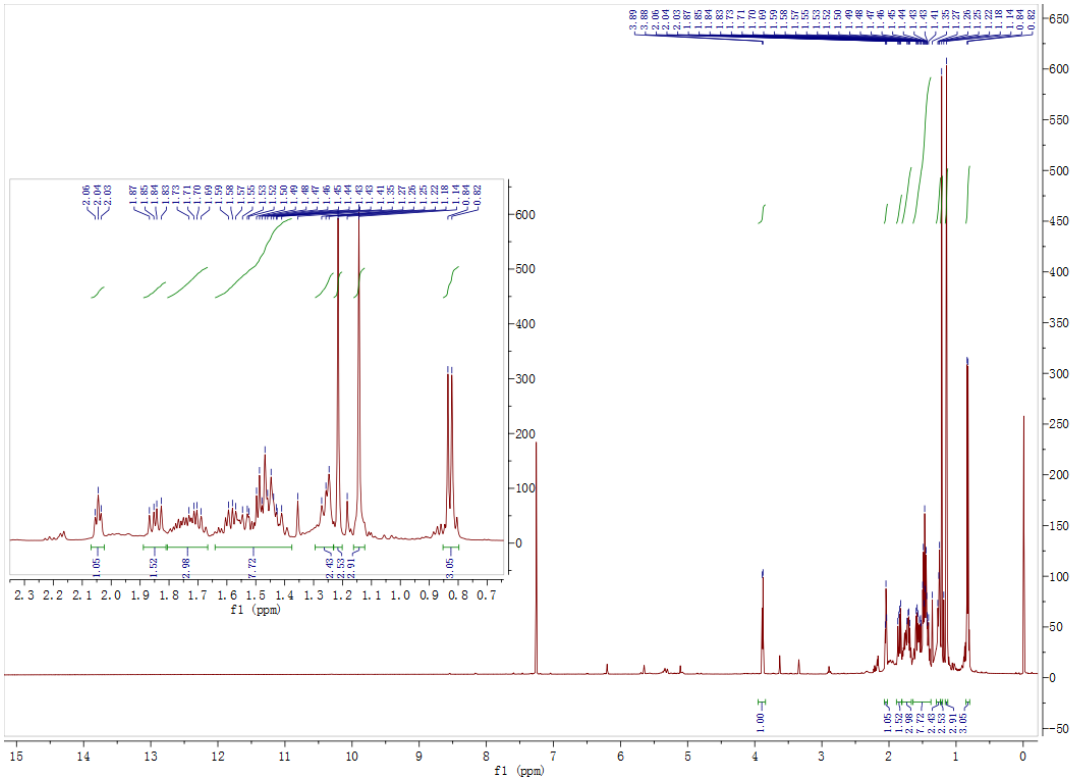


Figure S54. The ^{1}H NMR spectrum of compound **10a** in $CDCl_3$ (500 MHz)

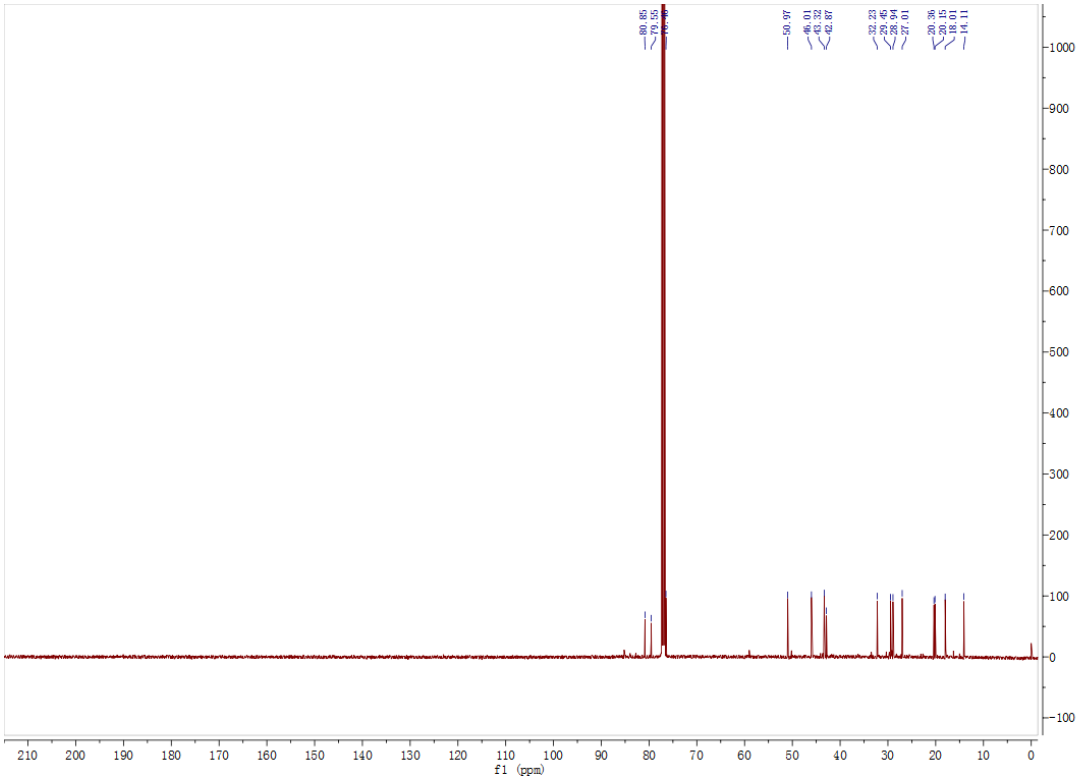


Figure S55. The ^{13}C NMR spectrum of compound **10a** in $CDCl_3$ (125 MHz)

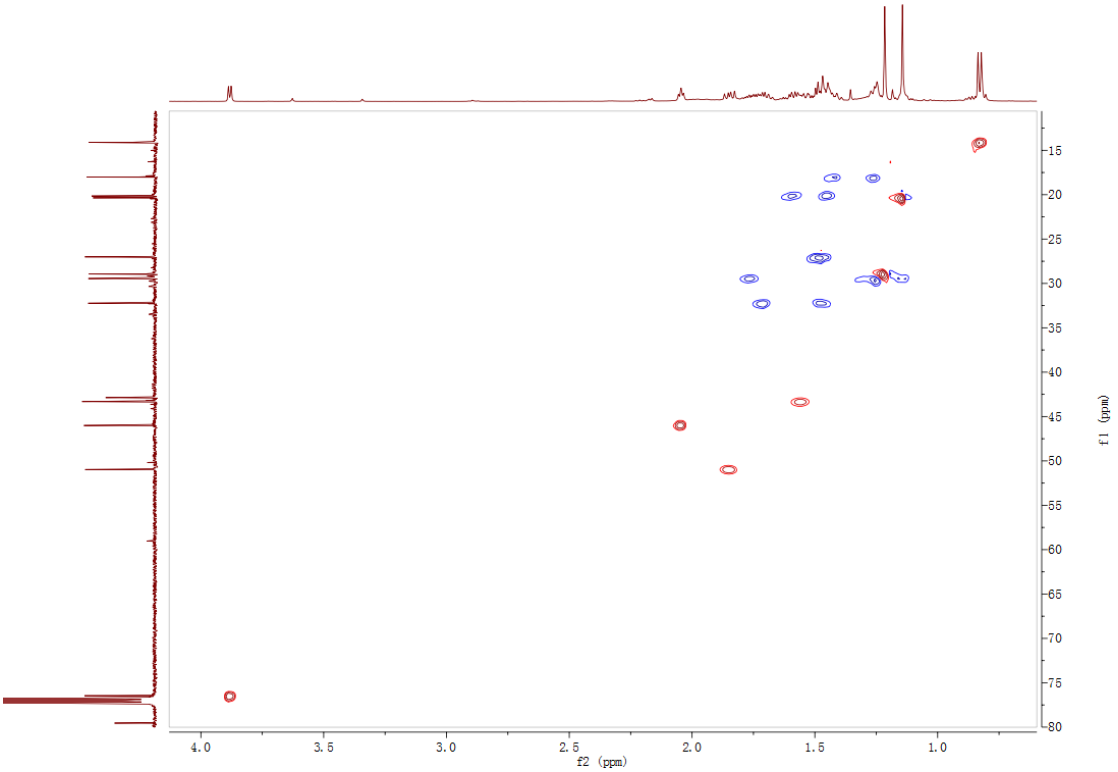


Figure S56. The HSQC spectrum of compound **10a** in $CDCl_3$

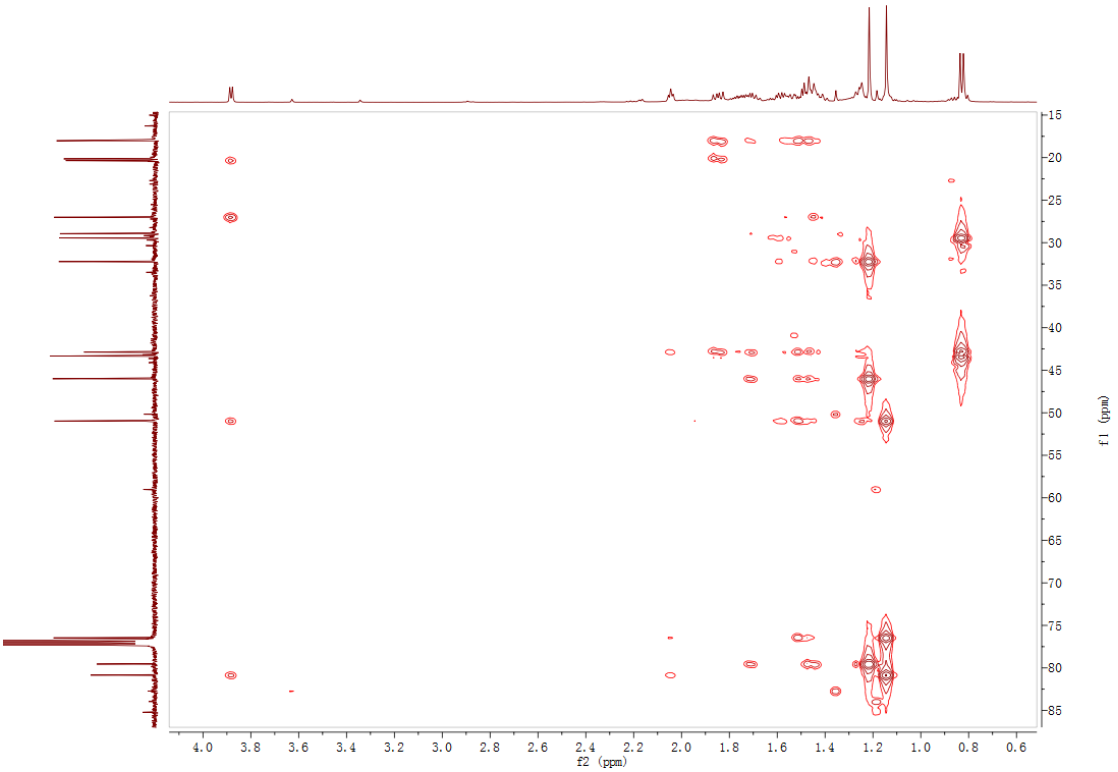


Figure S57. The HMBC spectrum of compound **10a** in $CDCl_3$

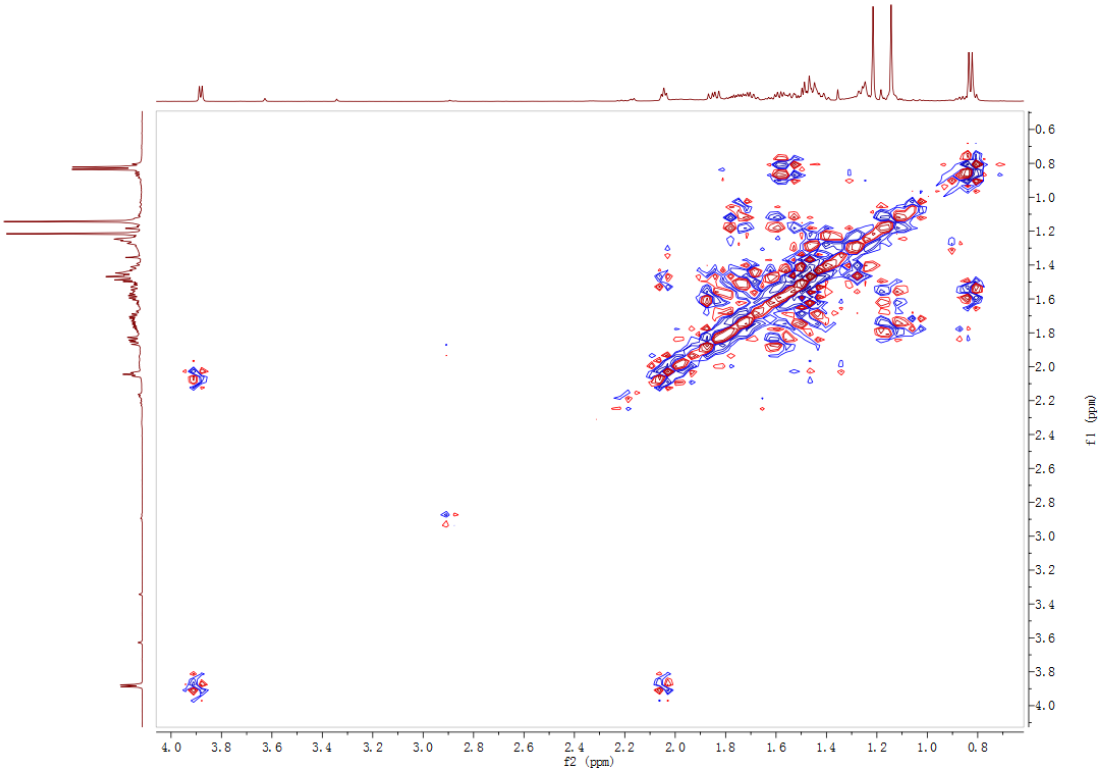


Figure S58. The ^{1}H-^{1}H COSY spectrum of compound **10a** in $CDCl_3$

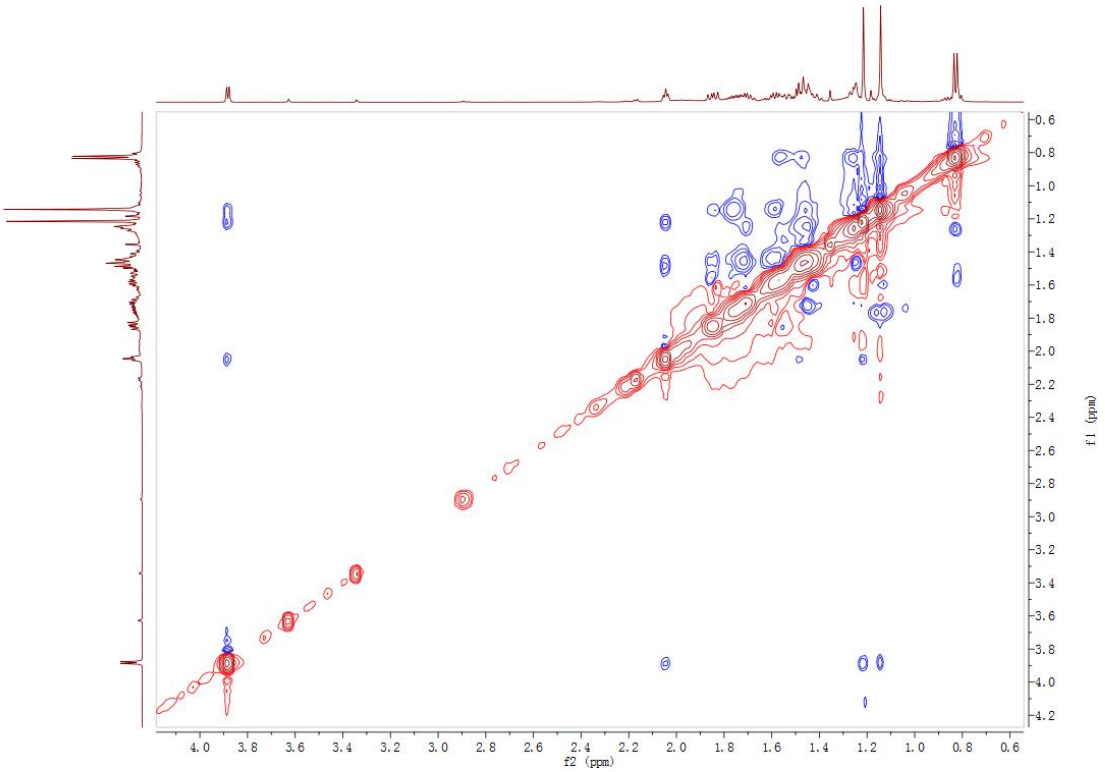


Figure S59. The NOESY Spectrum of Compound **10a** in $CDCl_3$

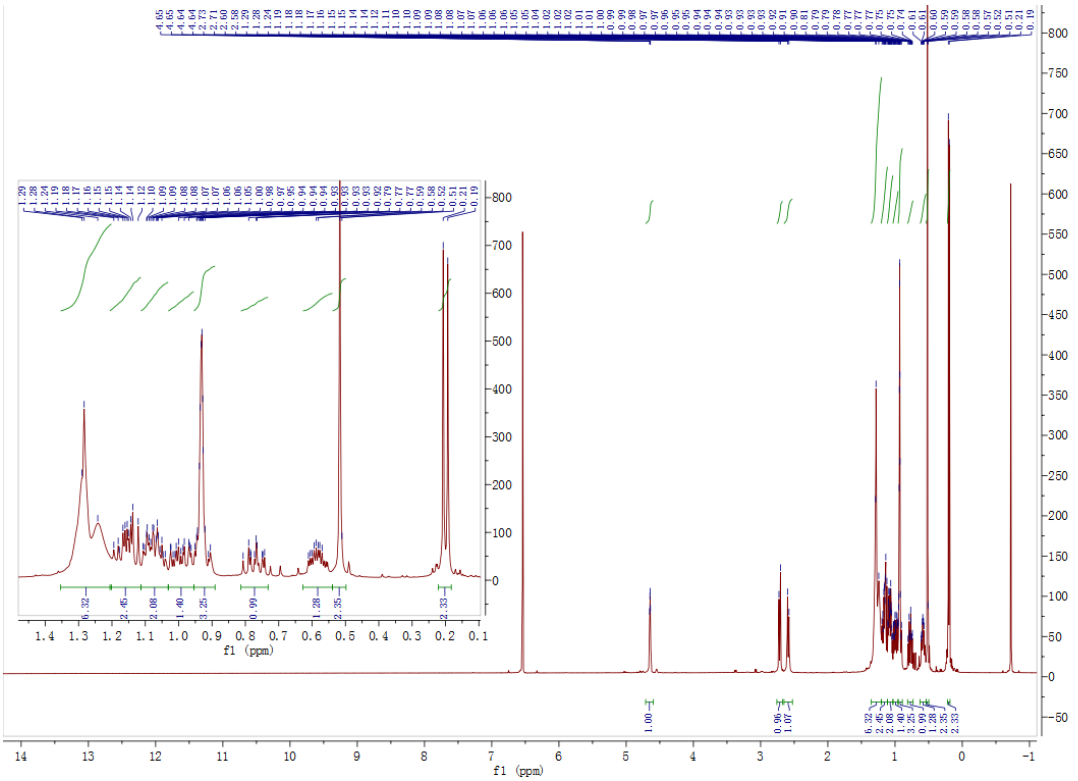


Figure S60. The ^{1}H NMR spectrum of compound **11** in $CDCl_3$ (500 HMz)

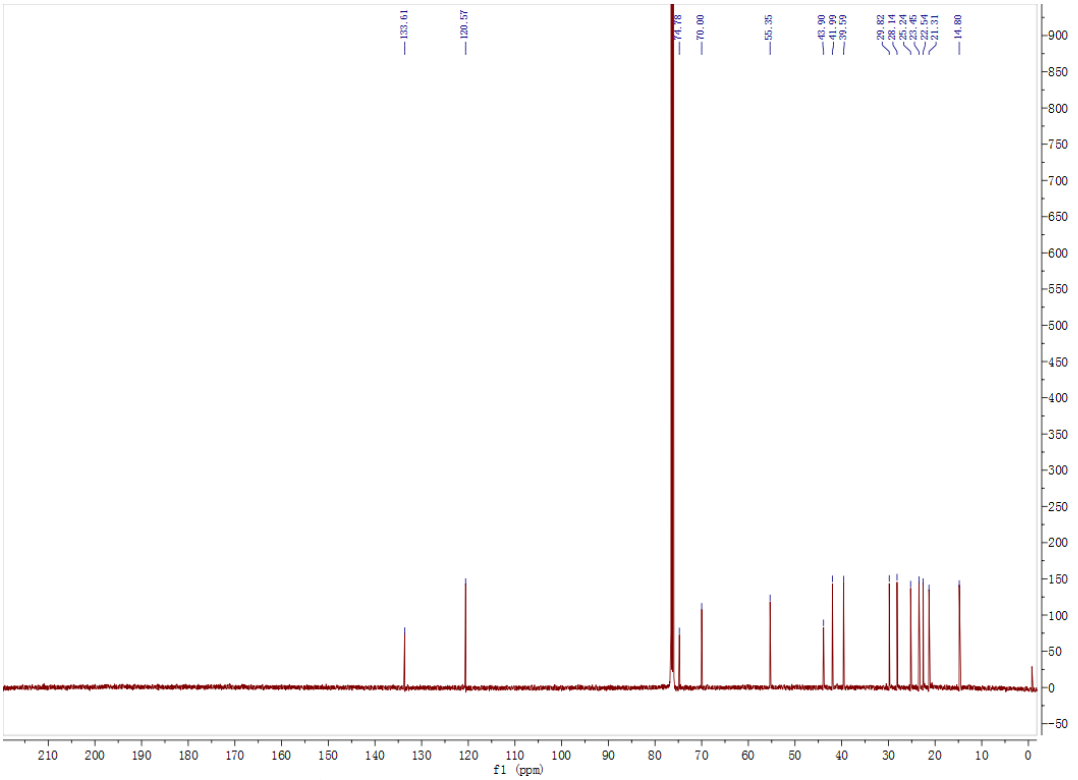


Figure S61. The ^{13}C NMR spectrum of compound **11** in $CDCl_3$ (125 MHz)

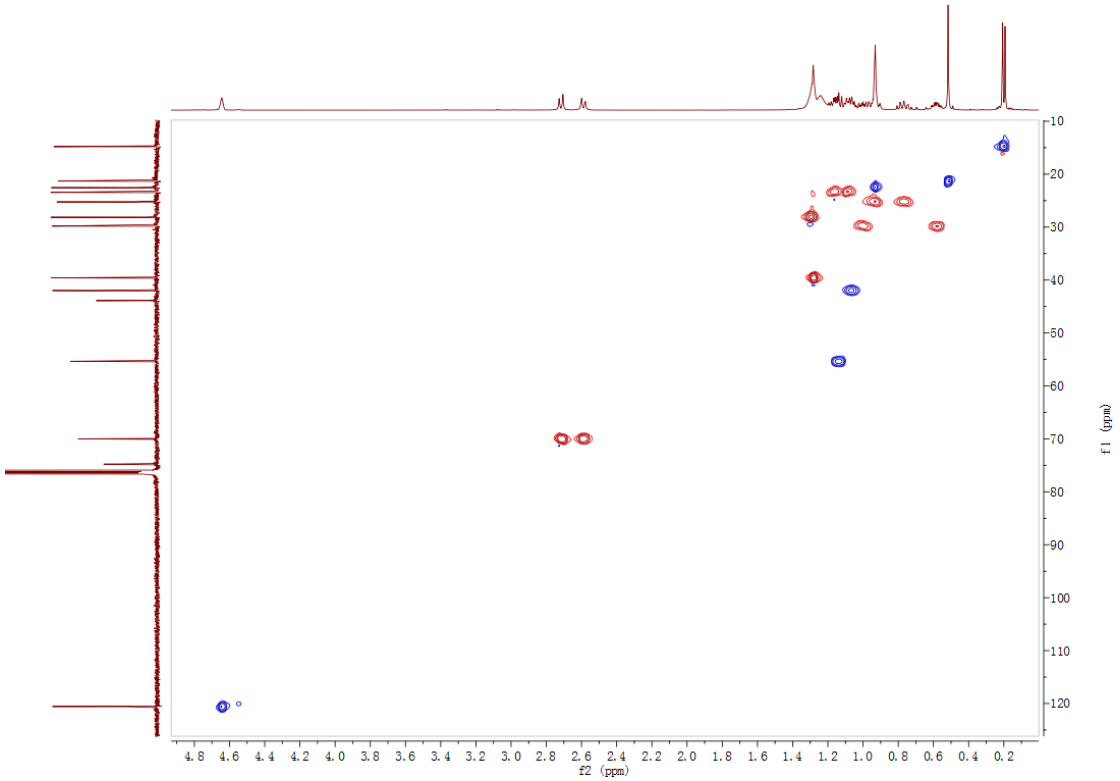


Figure S62. The HSQC spectrum of compound **11** in $CDCl_3$

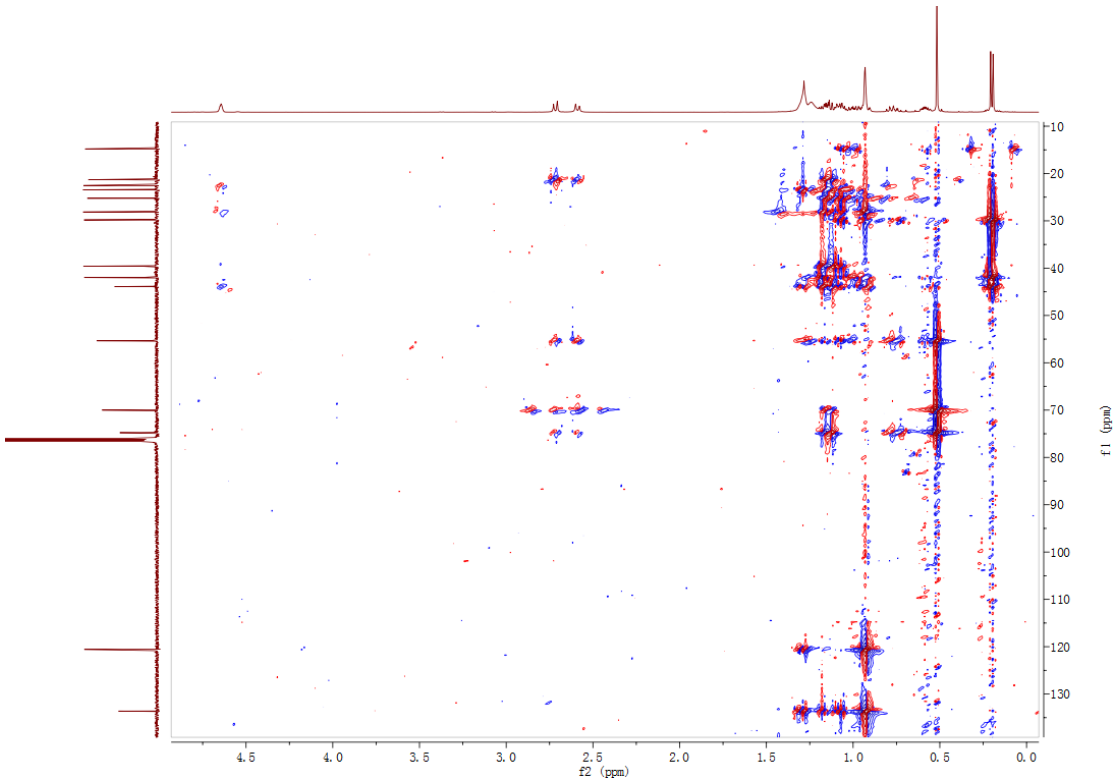


Figure S63. The HMBC spectrum of compound **11** in $CDCl_3$

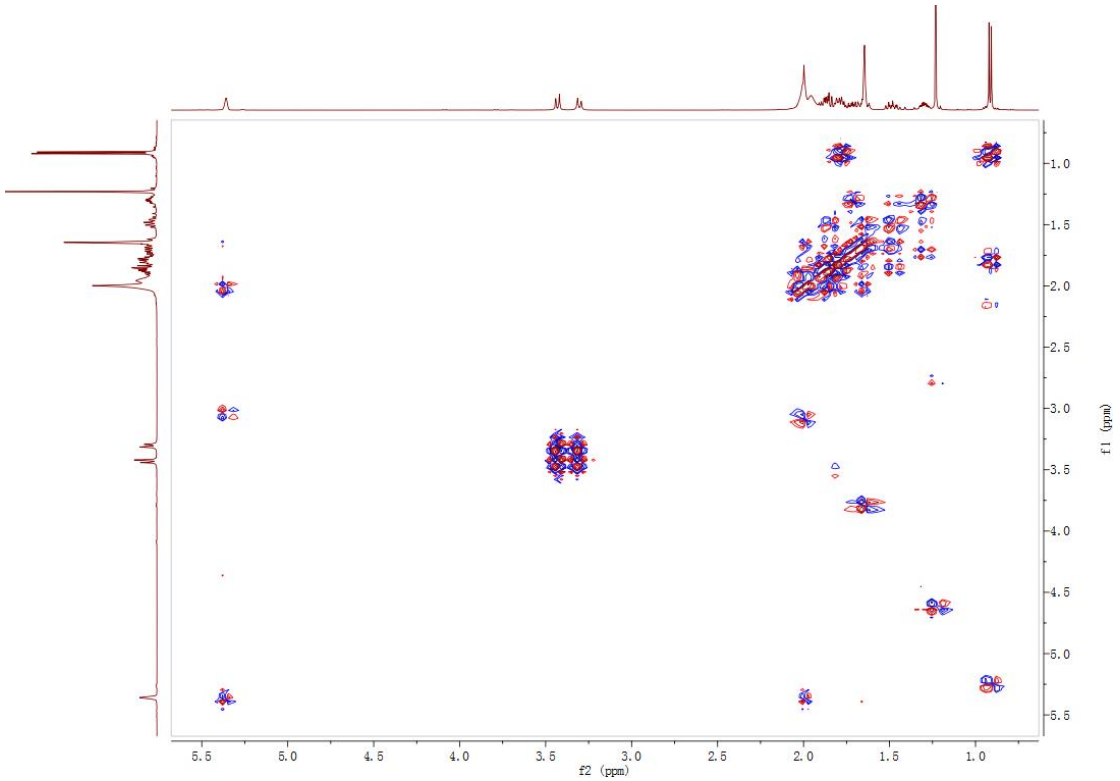


Figure S64. The ^{1}H-^{1}H COSY spectrum of compound **11** in $CDCl_3$

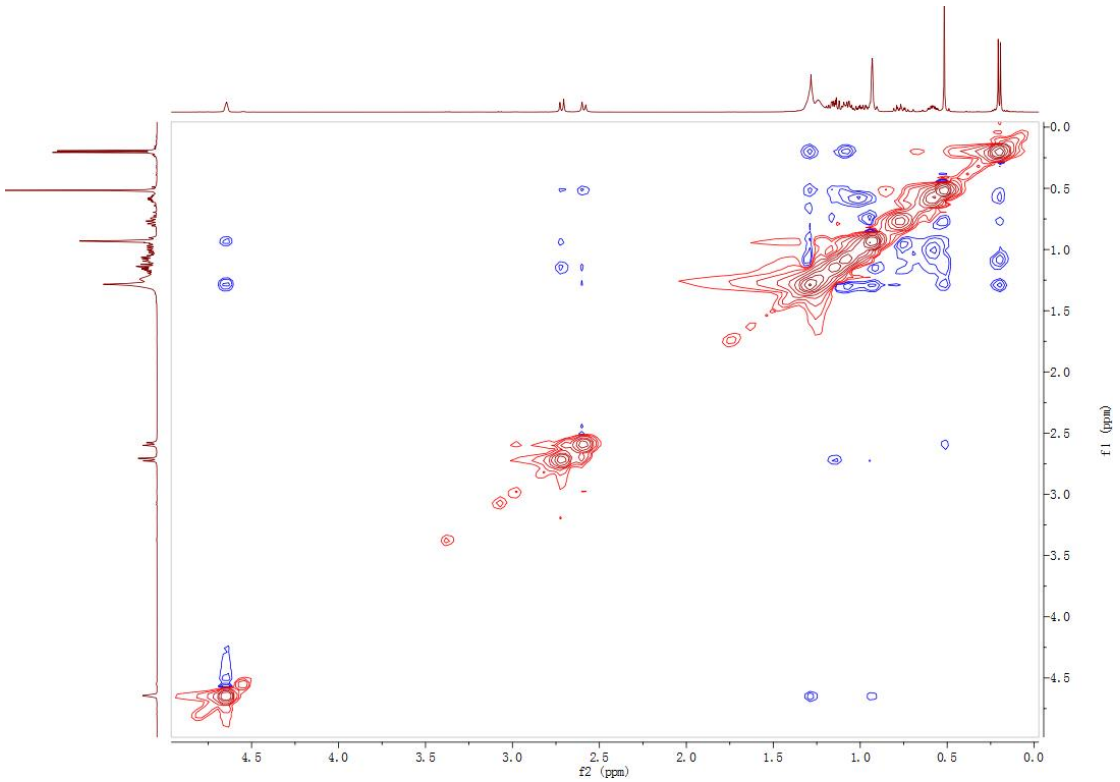


Figure S65. The NOESY spectrum of compound **11** in $CDCl_3$

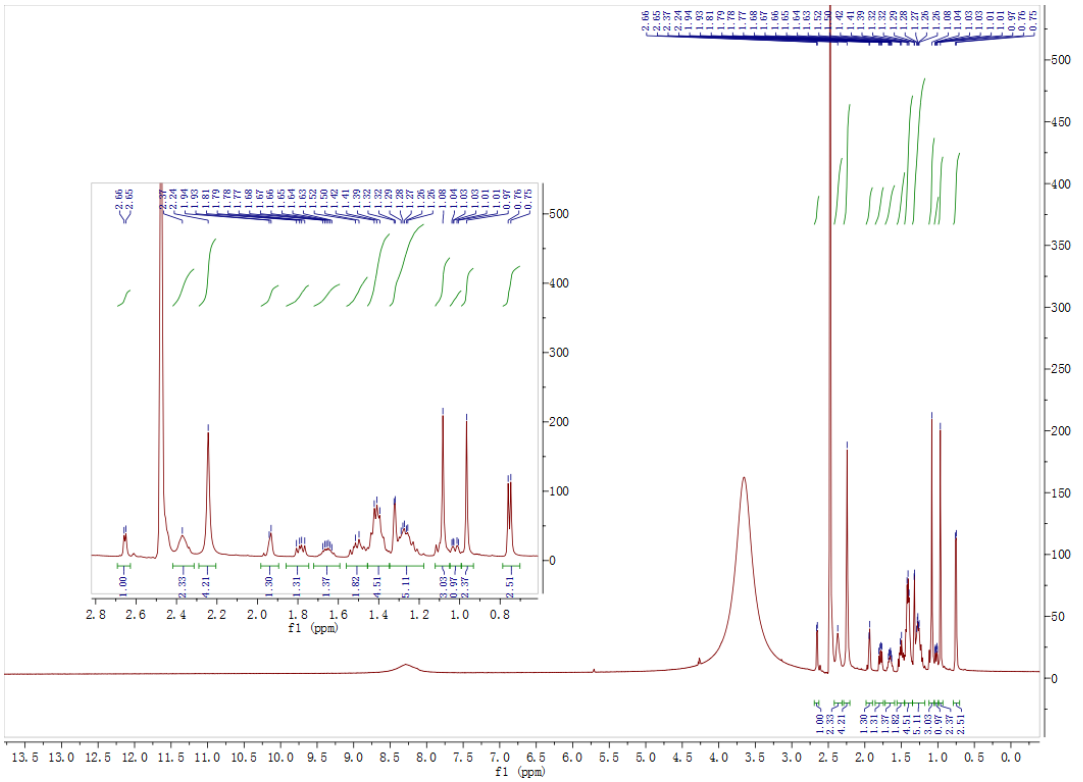


Figure S66. The ^{1}H NMR spectrum of compound **12** in DMSO-d_6 (500 MHz)

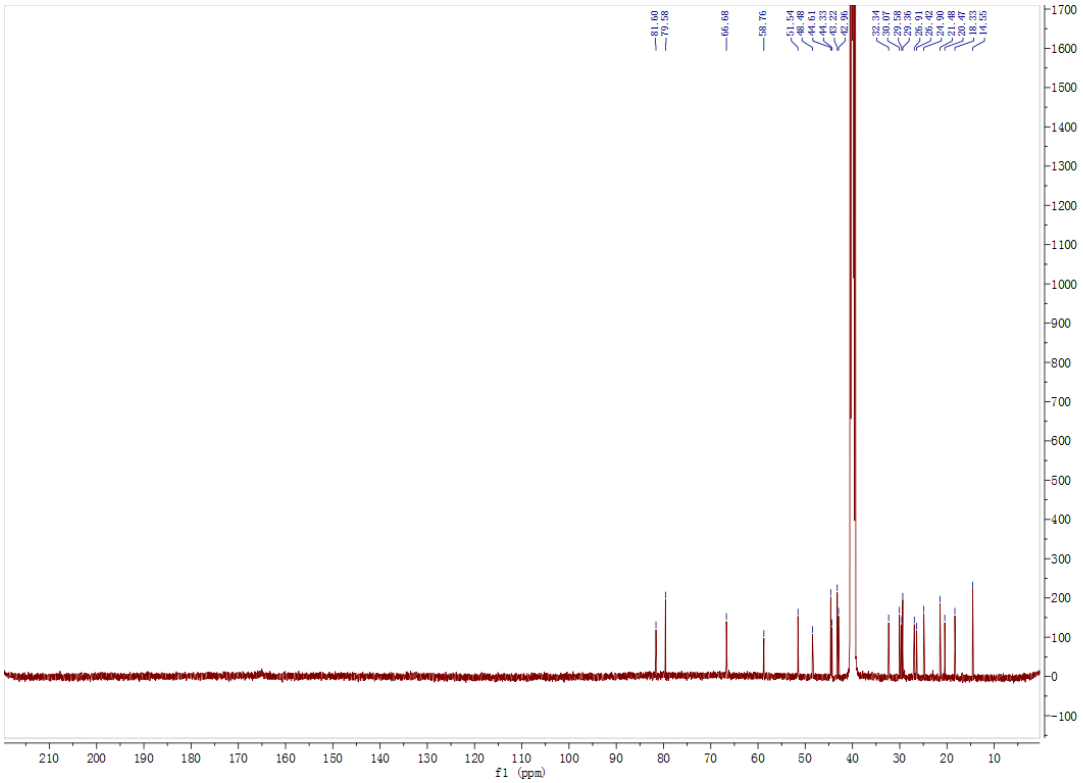


Figure S67. The ^{13}C NMR spectrum of compound **12** in DMSO-d_6 (125 MHz)

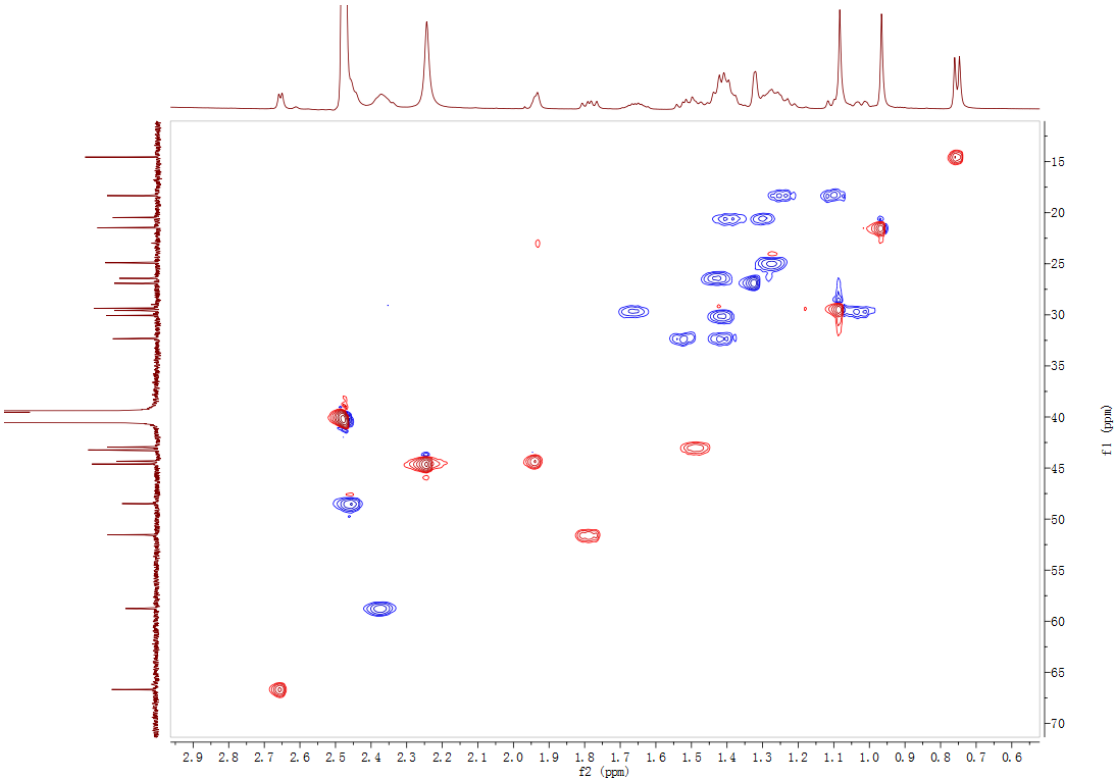


Figure S68. The HSQC spectrum of compound **12** in DMSO-d_6

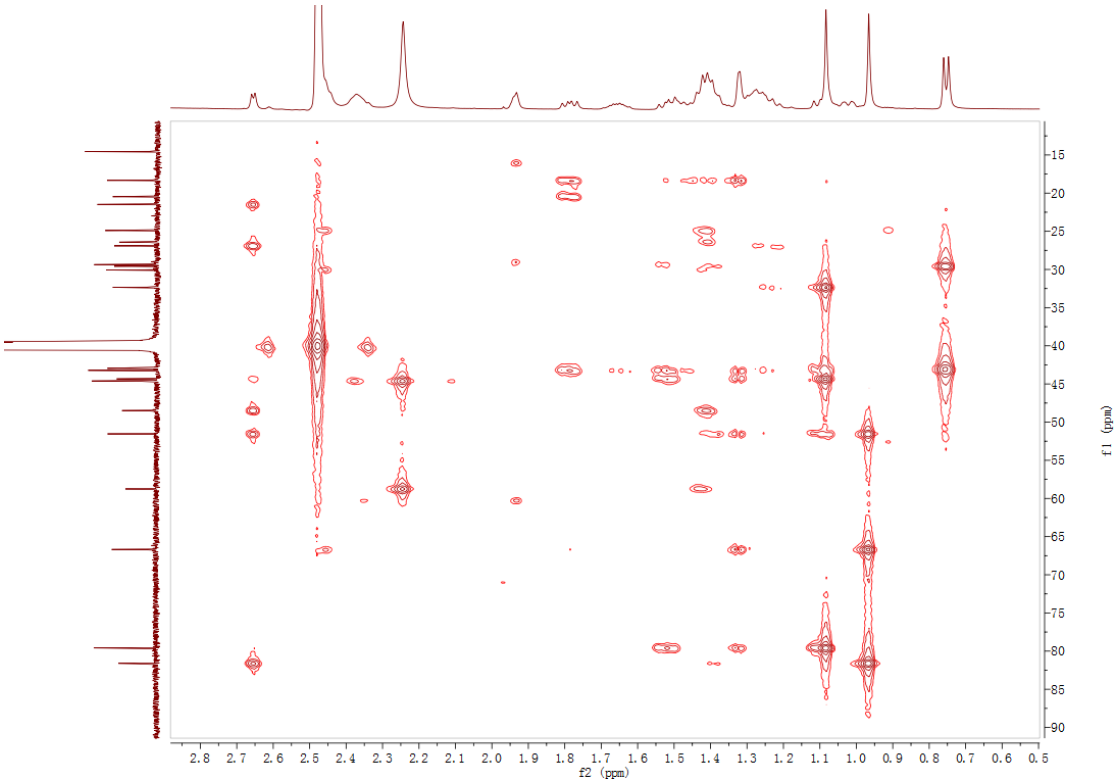


Figure S69. The HMBC spectrum of compound **12** in DMSO-d_6

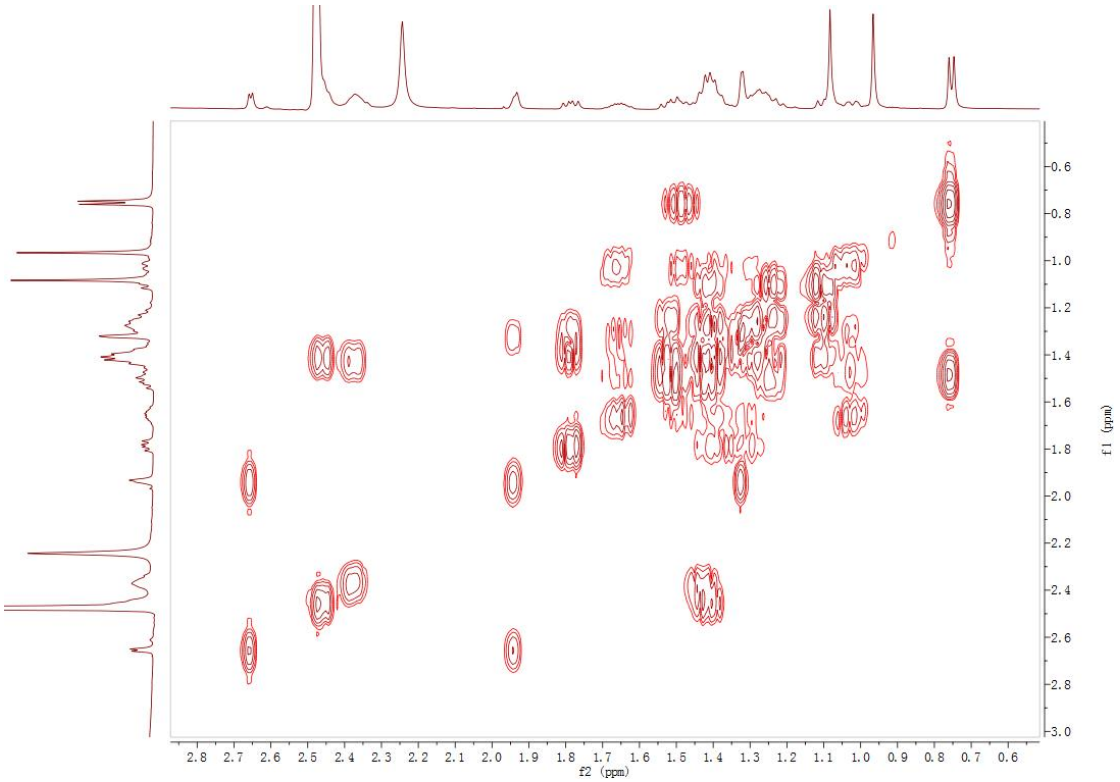


Figure S70. The ^{1}H-^{1}H COSY spectrum of compound **12** in DMSO-d_6

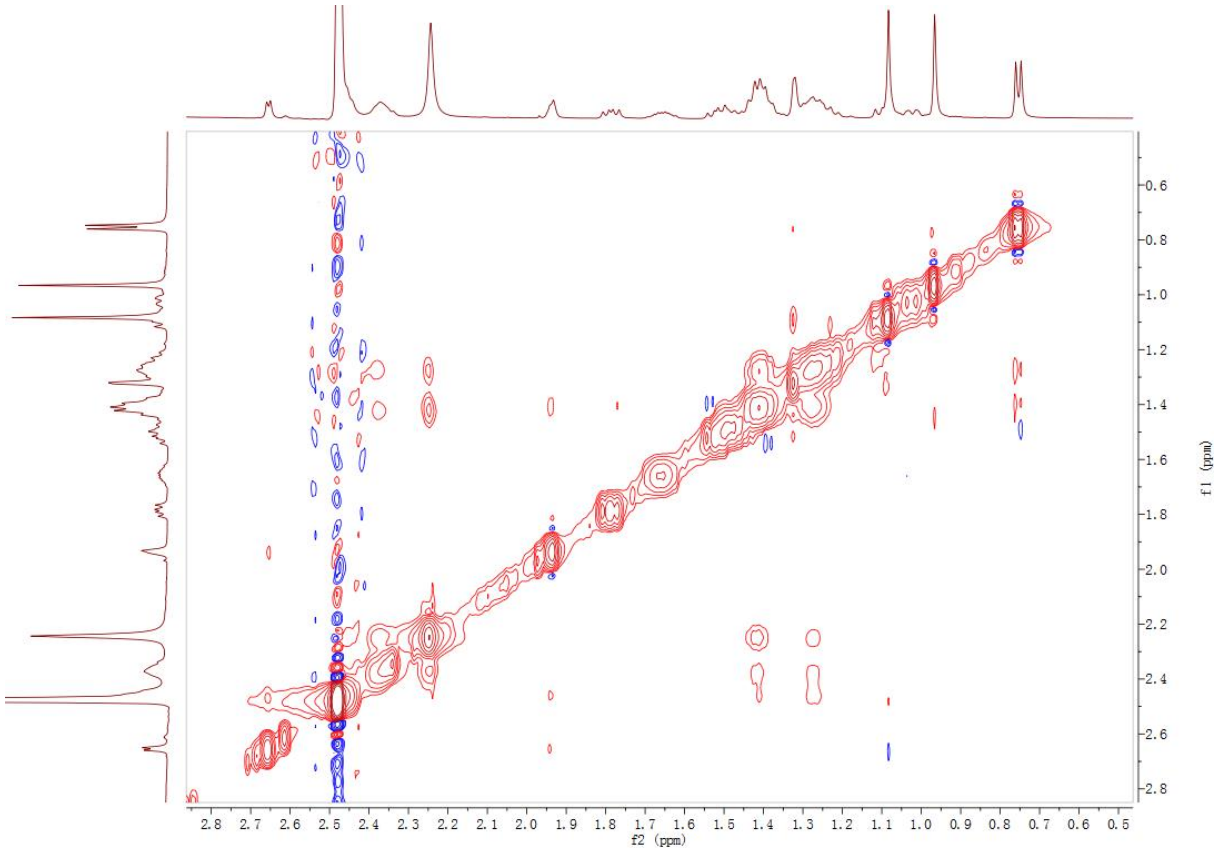


Figure S71. The NOESY spectrum of compound **12** in DMSO-d_6

9.3 Supplementary information for Section 7

Table S1. Bioinformatics analysis of the *ank* gene cluster from *A. thermomutatus* using NCBI BlastP search

Protein	**Size (aa)**	**Proposed function**	**Homologs (ident/coverage)**	**Strains**
AnkA (XP_026617165.1)	518	hypothetical protein	no hits	no hits
AnkB (XP_026617163.1)	524	cytochrome P450	KAG9240506.1 (60%/ 90%)	*Calycina marina*
AnkC (XP_026617167.1)	587	FAD-dependent monooxygenase	P0DUL5.1 (41%/89%)	*Aspergillus hancockii*
AnkD (XP_026617166.1)	563	pyridoxal-phosphate dependent enzyme	KAF7618911.1 (50%/93%)	*Aspergillus flavus* NRRL3357
AnkE (XP_026617162.1)	604	NRPS independent siderophore synthetase (NIS)	CDM31866.1 (48%/93%)	*Penicillium roqueforti* FM164
AnkF (XP_026617168.1)	239	O-methyltransferase	XP_015700694.1 (41%/100%)	*Paracoccidioides lutzii* Pb01
AnkG (XP_026617169.1)	480	ATP-grasp protein	KAF7349305.1 (35%/99%)	*Mycena sanguinolenta*

Table S2. Bioinformatics analysis of the *ava* gene cluster from *A. versicolor* dI-29 using NCBI BlastP search

Protein	Size (aa)	Proposed function	Homologs (ident/coverage)	Strains
AvaA	541	hypothetical protein	no hits	no hits
AvaB	709	FAD-dependent monooxygenase	TVY66699.1 (28%/ 98%)	*Fusarium oxysporum* f. sp. cubense
AvaC	793	kynurenine formamidase	A5DNX8.2 (28%/34%)	*Meyerozyma guilliermondii* ATCC 6260
AvaD	290	GNAT family N-acetyltransferase	MBE0409255.1 (38%/85%)	*Anaerolineales bacterium*
AvaE	336	beta-lactamase	KAH6869491.1 (33%/83%)	*Thelonectria olida*
AvaF	537	cytochrome P450	KAH6680460.1 (34%/87%)	*Halenospora varia*
AvaG	319	lysophospholipase catalytic domain family protein	TPR05408.1 (26%/97%)	*Aspergillus niger*
AvaH	492	cytochrome P450	KAF2091838.1 (31%/70%)	*Saccharata proteae* CBS 121410
AvaI	437	cytochrome P450	KAH8801390.1 (28%/84%)	*Xylogone* sp. PMI_703
AvaJ	561	cytochrome P450	KAF8155335.1 (32%/83%)	*Mycena galopus* ATCC 62051
AvaK	255	NAD-dependent epimerase/dehydratase	TQW01081.1 (45%/92%)	*Cordyceps javanica*
AvaL	362	cytochrome P450	XP_033685598.1 (32%/90%)	*Trematosphaeria pertusa*
AvaM	338	dioxygenase	XP_025392866.1 (64%/100%)	*Aspergillus eucalypticola* CBS 122712
AvaN	453	glycosyltransferase	KAE9372740.1 (39%/97%)	*Chalara longipes* BDJ

Table S3. Plasmids used in this study

Plasmid	Vector	Genes
p3001	pYTU	n/a
p3002	pYTR	n/a
p3003	pYTP	n/a
p3004	pYTR	gpdAp-AnkD; POgpdAp-AnkA
p3005	pYTU	gpdAp-AnkE; POgpdAp-AnkB
p3006	pYTP	gpdAp-AnkC; POgpdAp-AnkF
p3007	pYTR	POgpdAp-AnkA
p3008	pYTU	POgpdAp-AnkB
p3009	pYTP	gpdAp-AnkC
p3010	pYTU	gpdAp-AnkE; POgpdAp-AnkB; PEgpdAp-AnkG
p3011	pYTP	gpdAp-AnkE; POgpdAp-AnkB; PEgpdAp-AnkG(spliced)
p3012	pYTU	gpdAp-PthA
p3013	pYTU	gpdAp-AteA
p3014	pYTU	gpdAp-AmaA
p3015	pYTR	gpdAp-AvaA
p3016	pYTU	gpdAp-AnoA
p3017	pYTU	gpdAp-PsbA
p3018	pXW55	ADH2p-AmaA-ADH2t
p3019	pXW55	ADH2p-*N*-his-AvaA-ADH2t
p3020	pXW55	ADH2p-*N*-his-AvaA(E315A)-ADH2t
p3021	pXW55	ADH2p-*N*-his-AvaA(E432A)-ADH2t
p3022	pXW55	ADH2p-*N*-his-AvaA(Y515A)-ADH2t
p3023	pXW55	ADH2p-*N*-his-AvaA(Y515F)-ADH2t
p3024	pET28a	T7p-*N*-his-AnkD-T7t
p3025	pET28a	T7p-*N*-his-AvaA-T7t
p3026	pET28a	T7p-TrpRS-*C*-his-T7t
p3027	pET28a	T7p-ArgRS-*C*-his-T7t
p3028	pYTU	coxAp-AvaB
p3029	pYTU	coxAp-AvaB; PEgpdAp-AvaC
p3030	pYTP	POgpdA-AvaD

Table S4. cDNA sequence of *ankG*

ankG	ATGTATCAAATTTCTCTAAAGGCGACGAAGAGCGCTGCAGAACCCACCTCTTCAACCGAC GCATCTCACGACGACCGTCAAGTAGAGCGCGATTCATACGACACCGCGATGCTGCGCTT CATCAGCGACATGGAACTGTCCTCAATGTACGCGGAGAAGACGTCACCATACCCCATTC TCATGCCGAGAAGCTTCCTTGAGGACCTGAAGAATTTCCAAGATCTGCTATTTGTTGCTG TATCAAACATCCTGGATCGATGGTGGGAGGACAGGGAAGCCGACTTTCCCCGTCGCATG CCATTGGAACCTCACGAGGAGAGCGTATTGAAGCATTACAATGAGCGCCCGCTACACTG GAGGCCGGACATGCTCCTCCCAGCAGCCGGCGATCCAAACACCAACCTACCTTTCAAAA TTTGCGAGATCAACGCTCGTTCCCCCTTCAACTCAATGATTAAGAGCATCTGTATGTTTCA GGCAGCGGCCGCAAGCAAAACAGCCCTCCCCGACGGGTTAGAACTCGCTTCGACTGCG GATAGCCTGGTGGATAGTCTGGTCTCGCTGTTCAATCCGGACCTCCCTCTTCATGTTATT TGGCATGAGGGTATCACTGACCCCATCGATGCGTTCAGTTCCTTCTACAAAAAGCGAACA GGCAAGATACCGCGAGTAATTCGCGCGACAGACCTGCGTCTCGCGCCGGATTCCTCATC CCCGACGGGTCGGATTCTGTGCTGCGTTGCCTCCGCCGCGTCCGCAGACCATTCTAAC GGCCAGGCAATTGTATCCGAGACAGGCGAACCCCTGGAACGTATCTACCAGGTCGGACT CCAGATGTCCCATCGGGATTACAGTGAGCTCTCTTCGGAAGTGCTGCAGCAGCTCGCCG TGGACGGAATCTGCGACCTACGCAACATCTTCCTGGTCAGCGATAAACGCATGCTCGGC GTAATCTGGCAAGAGCTGGACTCTCTCGTGCACAAGCACCACGTCCTCACGGCTGAGCA GGCCGAGATCCTCCGACAGGGCATCGTACATACCATTTTGCCCGGCTCCGAGGATATGG AACGCCTCTTGCGGCAGACTCGCGAGGGCTCCGTATCCAAGGATAGCTATCTTCTCAAA CCTGCCCGTGGACACCGGGGAATGGGAATATTGCTTGGGAAAGACTTGGGGCAGGAGG AGTTTGAGGGTCTCCTCGAGGAGCTGGCGGACCCGTTACTGCCTGCTGACAGGAGATAT GTTGTGCAGCCTTTTATTGAGCAGGCACTATTCGGTTTGAGGCTGTATGATGATAGCGAG CCGCAACAATGCCAGATGACGGGGACATATCATGCCATTGGTGGGTGCTTTGCTGGGTT GGGCGTCTGGAGAGCTGATAGCGAGAGAATTTGCTCTCGATTCCACGGTGCCTTTAGTA TTCCAGCGATTGTTCCCCGTTAG

Table S5. Primers used in this study

Primer	Sequence (5'-3')
A thermomutatus NRPS check R	ACTGGGGAAGTGGCGAAAG
A thermomutatus NRPS F	ATGACTGTGCCACCAATAAGC
A thermomutatus PLP check F	CGTTCTGGCGAGATTCAAGC
A thermomutatus PLP check R	GTGATCTCGCAGAACAGCAC
A therm NRPS POgpdA F	CCTTTTATCAATCGGGGACTTCGTCGTAGTTGACATCTTTGCTCCAGGAATACATGTGAG
POgpdA Atherm P450 R	GGCCAGAAATGGCCCAGGTTGAGTTGAATTGTGCCATTTTTGCGATTGTTTGAAGTGTTC
Atherm P450 pYTU R	CGGGGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATCAGTCAGGATGCCACTACTC
gpdA atherm NRPS R	CGAACGAGGCTCGTTCGCTTATTGGTGGCACAGTCATTGTTTAGATGTGTCTATGTGGCG
gpdA atherm PLP R	TTCCATGTTGAGCTGATGCAGGGCCCAGTTCACCCATTGTTTAGATGTGTCTATGTGGCG
atherm PLP POgpdA F	CATCAAAGTCTAACAGAGCCAGCAAAGTAATCCAAACTTTGCTCCAGGAATACATGTGAG
POgpdA atherm HP R	TGCTGGAGGTCGGCAACTGGAATCCTGTCTTCCATTTTTGCGATTGTTTGAAGTGTTCTG
atherm HP F	ATGGAAGACAGGATTCCAGTTGC
atherm HP pYTR R	CGAATTCCTGCAGCCCGGGGGATCCTTAATTAACACAATATGTAACTATCGCTTTGTTGG
gpdA atherm hydroxylase R	TAATCAACACATCAACCGCAGACTCTTTCTTGGTCATTGTTTAGATGTGTCTATGTGGCG
atherm hydroxylase F	ATGACCAAGAAAGAGTCTGCG
atherm hydroxylase POgpdA F	GCTGGCGGTCCATCCTCAAGCAGTGTAAGTAAAATTCTTTGCTCCAGGAATACATGTGAG
atherm hydroxylase R	GAATTTTACTTACACTGCTTGAGGATGG
POgpdA atherm MT R	GTTTCATCGGGGAAGGGAATGATGCTTTCGTACATTTTTGCGATTGTTTGAAGTGTTCTG
atherm MT F	ATGTACGAAAGCATCATTCCCTTC
atherm MT pYTP R	GCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGGTTCTGCAGCGCTCTTC
atherm PLP pYTR R	TATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGTTTGGATTACTTTGCTGGCTCTG
atherm PLP F3	GGCACCAATTTCTCGATTTCCATC
atherm PLP F4	GGAGGTGGAACTGAGAGAGAAG
atherm PLP F5	TGCTGGTTTCAGAAACTCGATTC
atherm PLP R3	CTATCCCATCTTCTCTCTCAGTTCC
atherm HP R2	GTCCAGTTTTCGACGGTGAG
atherm hydroxylase R2	CATCTTGAATCCCAGTTGCTTCC
atherm MT R2	CAAGAGGACGTCGAGCTG

atherm P450 R2	CTCTCGAATGAAGGGGCTAGC
atherm nrps F2	CCCTGTTTAATCGTCCGTCAG
atherm NRPS F3	CATGCGGAGATTGTCTTTGGTC
atherm P450 R3	AGATCCTGGCCGTTGTCG
atherm hydroxylase F2	CTGTGTCTGCCTCACTTGTTATC
atherm hydroxylase R3	GTGCCACAGGTATTGAACCTG
atherm P450 PEgpdA F	CTCGTGACATCTTTTGACTTGAGTAGTGGCATCCTGACTGAGATTCGTCCAGGGCTTCCC
PEgpdA atherm g20 R	CGCTCTTCGTCGCCTTTAGAGAAATTTGATACATGATTGCGGTTTACTAGAAGTTGTTAG
atherm g20 F	ATGTATCAAATTTCTCTAAAGGCGACGAAG
atherm g20 pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATAAAGATCACTACTCACACCAACTCAC
atherm hydroxylase check F	AAGACTACAGAGACTGCTGAGG
atherm hydroxylase check R	TGCTTTTTGGCCTCGAAAATCG
atherm MT check F	CCTTTGTCACTATTCTCAGGTCAG
atherm hp check F	ACCCTGTGAAGGGCATGC
atherm hp check R	AGATACACCTCGAGTAGATGAGG
gpdA apimo HP_cdps F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGGTTTCAATCTCCAAAAGCCTC
apimo hp_cdps pYTU R	GGGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGAACGTTAATGTTTCGGGAGAAC
gpdA av hp_cdps F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGTCTATACCAACCGATATCGC
av hp_cdps pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATGCCAGGTCTGATGTTTGATTTCTC
av HP F	ATGTCTATACCAACCGATATCGCC
av HP F2	CCTTCGTGGAAGCCATCAGAG
av HP R2	CGCAGCTGGATTTGTCCTTG
av HP R	TGCCAGGTCTGATGTTTGATTTCTC
ADH2p apimo HP F	ACAATCAACTATCAACTATTAACTATATCGTAATACATGGTTTCAATCTCCAAAAGCCTC
apimo HP xw55 R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTAAACTACGTATGTCCTGAAACATTGACG
apimo HP seq R	AAGGCAGATGCCACGAAG
gpdAp av in house HP F	AAGCTTGACTAACCATTACCCCGCCACATAGACACATCTAAACAATGACGGCTACCGTGC
av in house HP term pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGGAAAGATTTCTGTATCTACTCGGTTC
gpdAp pthy HP F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGAGTACACTGTATCTGCTTGC
pthy HP pYTU R	CACTAGTTCTAGAGCGGCCGCCTTAATTAATTTTAGTGGTAGTTGTTTAGAGGATCAACG
gpdAp aspnom HP F	CTTGACTAACCATTACCCCGCCACATAGACACATCTAAACAATGATGGGTGCCGTACAAC

aspnom HP pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATCCTTGAAAGATTTTTACT ACTGGCAG
gpdAp aspte HP F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGGATGTGCAATAT GTACTCGC
aspte HP pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGGATCTGGCAGCATTG AATATCTCC
pET28a YDR341C F	ATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGCTAGCACAGCAAAT ATGATTTC
YDR341C pET28a R	GCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCATTCTTTCTACGGGAGT TAAACCTAAC
YDR341C check F	GGATTATGGTTCTTGCAAATTGGTTG
YDR341C check R	ACAGTTTTGCGGCAGGTTC
av in house HP pYTR R	GCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGACATTCAG CCCGCTGTG
adh2p av HP in house F	AAAAGCATACAATCAACTATCAACTATTAACTATATCGTAATACATGACGGCT ACCGTGC
av HP in house xw55 R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTGACGATCTAGTCTAATTCGA ATTCATCC
av in house HP seq R	GGGCCTTGGTCCAAGTAGC
n-his av in house HP F	ATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGACG GCTACCGTGC
POgpdA av HP3 R	GACTGATCCCCAGTGACTGTAGATCAATGCCTGCCATTTTTGCGATTGTTTG AAGTGTTC
av HP3 F	ATGGCAGGCATTGATCTACAG
av HP3 PEgpdA F	TAGCAATATTCAGATTCCCAGACTCCTCTTTTACCCGCCCAGATTCGTCCAG GGCTTCCC
av HP3 R	TGGGCGGGTAAAAGAGGAG
PEgpdA av KFA R	GTCTCTCATCTTCAGAGTGTTCGGGCGAAGTCATGATTGCGGTTTACTAGA AGTTGTTAG
av KFA F	ATGACTTCGCCCGAACACTC
av KFA F2	CTGGTTCGCATGATGATATTTGGAG
av FMO F2	ACATGCTTCTCCGAGATCTGG
av FMO cox4p F	GAATATTGCATACCAGAACCAACAATGACCACTCGCGGCATTGTCCTGGTG GGTGGGTTG
av FMO F	ATGCCGCGAGTGGTCATTG
cox4p pYTR R	TGCTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAATTAAAATGCCTGA TCCAGCCC
atherm hydroxylase pYTP R	GAATTCCTGCAGCCCGGGGGATCCTTAATTAAGAATTTTACTTACACTGCTT GAGGATGG
pYTU coxAp F	GTAGTGAGTCATTTACTCAGCACACTCGCGCTGACGCTCGTCAATGCCTGA TCCAGCCCC
av fmo pYTU R	CCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACGACCTGAG GAGACTTCTG
av KFA pYTU R	CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCTTTGTGGCTCT GGGCTTTGTC
av hp3 pYTP R	CTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAATGGGCGGGTA AAAGAGGAG
n-his atherm plp F	TCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGGTGAACTGG GCCCTGCATC

atherm plp pet28A R	CGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCTATCCCATCTTCTCTCTCAGTTCCAC
trp tRNA F	CGGTGGCTCAATGGTAGAG
trp tRNA R	AACGGACAGGAATTGAACCTG
m13 rev trp tRNA F	CAGGAAACAGCTATGACGTCTAGCGCTTTACGGAAGACAATCGGTGGCTCAATGGTAGAG
trp tRNA m13 fwd R	GTAAAACGACGGCCAGTCAAGGCGATTAAGTTGGGTAACAACGGACAGGAATTGAACCTG
m13 rev trp tRNA F	CAGGAAACAGCTATGACGTCTAGCGCTTTACGGAAGACAATCGGTGGCTCAATGGTAGAG
trp tRNA m13 fwd R	GTAAAACGACGGCCAGTCAAGGCGATTAAGTTGGGTAACAACGGACAGGAATTGAACCTG
atherm g20 spliced OH R	GTCCGGCCTCCAGTGTAGCGGGCGCTCATTGTAATGCTTCAATACGCTCTCCTCGTGAGG
atherm g20 spliced F	AGCATTACAATGAGCGCCCGCTAC
gpdA pensub HP F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGCTTGGTCCAATTTCGTTACAC
pensub HP pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAACAAGCCTTCAGTGGAAGAAGGAATG
av HP Y515A F	CATTGGTCAGATCGCAGGAGCGGAGGTTC
av HP Y515A R	CTCCGCTCCTGCGATCTGACCAATG
av HP E315A F	CACCATTATCCACGCAGTACGTCGACG
av HP E315A R	CGTCGACGTACTGCGTGGATAATGGTG
av HP E432A	GTGGACGATATCGCCGCATGGAGAATCTTTGAC
av HP E432A R	GTCAAAGATTCTCCATGCGGCGATATCGTCCAC
av HP Y515F F	CATTGGTCAGATCTTCGGAGCGGAGGTTC
av HP Y515F R	CTCCGCTCCGAAGATCTGACCAATG

Table S6. AnkA homologs

Strain	AA size of homolog	% identity to AnkA	e-value
Aspergillus udagawae CBS 114217	523	83.779	0
Aspergillus siamensis CBS 137452	523	84.895	0
Penicillium variabile	515	73.552	0
Fusarium tricinctum MPI-SDFR-AT-0068	524	42.857	7.40E-137
Astrocystis sublimbata CBS 130006	501	40.076	1.49E-125
Xylaria curta CBS 114988	501	40.571	4.94E-124
Aspergillus appendiculatus CBS 374.75	501	41.078	6.90E-124
Xylaria berteri	501	39.619	6.40E-120
Hypoxylon argillaceum CBS 527.63	500	38.783	6.69E-120
Aspergillus megasporus DTO 048-I3	524	40.49	1.04E-119
Nemania diffusa NC0034 v1.0	500	38.593	1.12E-119
Xylaria venustula FL0490	498	39.313	2.88E-119
Xylogone sp. PMI_703	297	61.074	1.22E-118
Eurotium rubrum	501	40.338	4.31E-117
Aspergillus cumulatus DTO 311-F5	501	40.49	6.64E-117
Aspergillus penicilloides CBS 540.65	490	43.75	1.55E-116
Xylariaceae sp. FL1272	501	38.692	6.82E-116
Aspergillus endophyticus DTO 354-I2	501	41.055	3.75E-114
Aspergillus endophyticus DTO 354-I2	501	37.121	2.67E-112
Poronia punctata CBS 180.79	507	38.015	1.00E-110
Podosordaria jugoyasan CBS 670.77	516	39.074	2.26E-110
Aspergillus teporis DTO058-E5	485	39.279	3.47E-105
Atropellis piniphila CBS 197.64	524	38.095	6.57E-99

Alternaria carthami BMP1963	519	35.019	1.06E-94
Nemania sp. FL0916 v1.0	452	34.63	1.82E-94
Monascus ruber NRRL 1597	584	59.211	4.85E-23
Monascus purpureus	589	60.526	3.50E-23
Aspergillus caperatus DTO 337-E6	410	40.541	3.81E-81
Apiospora montagnei NRRL 25634	543	32.299	4.23E-78
Aspergillus chevalieri CBS 522.65	405	39.394	0.015
Aspergillus multiplicatus CBS 646958	544	32.673	1.19E-70
Aspergillus carlsbadensis CBS 123894	524	32.24	1.60E-69
Pseudogymnoascus sp. VKM F-103	548	28	6.35E-66
Zalerion varium MPI-CAGE-AT-0135	519	31.216	1.82E-65
Leptosphaeria microscopica UNIPAMPA013	327	39.45	4.98E-65
Pseudogymnoascus verrucosus UAMH 10579	548	27.826	8.70E-65
Arthroderma benhamiae CBS 112371	509	33.69	3.54E-57
Trichophyton tonsurans CBS 112818	471	32.8	8.43E-56
Trichophyton interdigitale MR816	509	32.353	2.14E-54
Aspergillus egyptiacus CBS 656.73	508	33.245	6.00E-53
Thelebolus microsporus ATCC 90970	393	32.836	1.2E-52
Thelebolus globosus UNIPAMPA015	393	32.587	1.83E-51
Trichophyton equinum CBS 127.97	330	32.584	2.08E-51
Trichophyton verrucosum HKI 0517	509	33.422	7.53E-51
Aspergillus spiritus DTO 241-G7	305	32.031	1.42E-49
Endocarpon pusillum Z07020	489	29.197	1.04E-48
Aspergillus versicolor	508	32.62	2.65E-48
Nannizzia gypsea CBS 118893	509	31.818	2.89E-48
Aspergillus robustus CBS 428.77	310	36.928	3.78E-48

Aspergillus terreus NIH 2624	431	32.967	1.26E-47
Aspergillus shendaweii IBT 34197	400	32.817	1.96E-47
Aspergillus pseudofelis IBT 34107	334	31.622	1.19E-46
Aspergillus parafelis IBT 34187	402	31.635	5.93E-46
Antarctomyces sp. UNIPAMPA016	387	33.036	2.03E-44
Cucurbitaria berberidis CBS 394.84	387	31.549	6.57E-44
Thelebolus stercoreus	384	30.667	6.50E-43
Penicillium thymicola DAOMC 180753	494	28.571	1.62E-40
Aspergillus insolitus CBS 384.61	528	29.867	2.42E-39
Aspergillus pseudonomius CBS 119388	528	29.459	4.25E-37
Aspergillus nomius NRRL 13137	528	29.459	4.25E-37
Aspergillus versicolor dI-29	532	27.604	8.21E-37
Byssochlamys spectabilis No. 5	528	25.094	9.57E-35
Aspergillus affinis CBS 129190	527	27.968	1.81E-34
Aspergillus sclerotiorum CBS 549.65	436	27.331	2.15E-33
Aspergillus roseoglobulosus CBS112800	427	25.278	1.85E-31
Aspergillus tamarii CBS 117626	449	26.75	3.30E-31
Microcera lavarum	528	26.35	1.24E-30
Aspergillus nutans CBS 121.56	465	23.333	1.39E-30
Rhizina undulata CBS 300.56	497	23.475	3.46E-30
Lomentospora prolificans JHH-5317	546	22.922	1.11E-29
Aspergillus nomius IBT 12657	483	25.145	1.07E-28
Aspergillus pseudonomius CBS 119388	483	25.723	1.14E-28
Neofusicoccum parvum UCRNP2	519	26.05	2.58E-28
Tolypocladium paradoxum NRBC 100945	496	26.269	3.05E-28
Gyromitra infula GyrinfSpk-SM18	518	26.861	5.92E-28

Geopyxis carbonaria CBS 144460	497	27.393	6.47E-28
Tricharina praecox CBS 144465	476	30.345	1.64E-27
Aspergillus undulatus CBS 261.88	501	24.933	5.38E-27
Aspergillus siamensis CBS 137452	492	23.16	1.34E-26
Aspergillus pseudofelis IBT 34107	492	23.116	1.38E-25
Aspergillus appendiculatus CBS 374.75	491	23.497	2.66E-25
Aspergillus neowarcupii IBT 29024	1003	24.149	5.39E-25
Aspergillus brevistipitatus	491	24.923	7.78E-25
Aspergillus avenaceus IBT 18842	480	25.895	2.36E-23
Aspergillus parafelis IBT 34187	492	22.613	6.27E-23
Penicillium thomii	498	24.747	2.42E-22
Aspergillus shendaweii IBT 34197	491	21.833	5.99E-22
Aspergillus waksmanii IBT 31900	480	21.981	3.57E-21
Hirsutella minnesotensis 3608	689	24.541	7.14E-21
Aspergillus transcarpathicus CBS 423.68	490	24.069	7.73E-21
Pochonia chlamydosporia 170	509	20.179	1.44E-20
Scedosporium apiospermum IHEM 14462	499	22.684	2.60E-18
Lobaria pulmonaria	528	22.363	2.27E-17
Drechslerella stenobrocha 248	345	26.038	2.19E-16
Monacrosporium haptotylum CBS 200.50	541	22.892	2.02E-15
Metarhizium robertsii ARSEF 23	379	23.009	1.05E-14
Botryosphaeria dothidea	442	26.432	6.73E-14
Stanjemonium grisellum CBS 655.79	460	22.174	2.14E-13

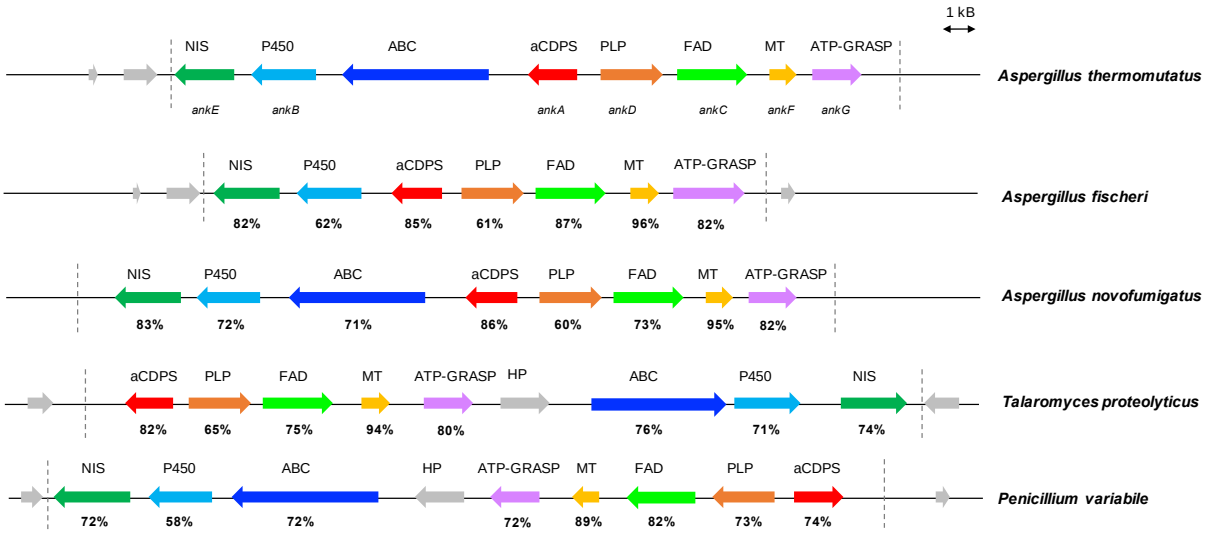


Figure S1. Homologous clusters in *Aspergillus thermomutatus, Aspergillus fischeri*, *Aspergillus novofumigatus*, *Talaromyces proteolyticus*, and *Penicillium variable*. Seven genes from *ankA* to *ankG* are conserved in all of the clusters.

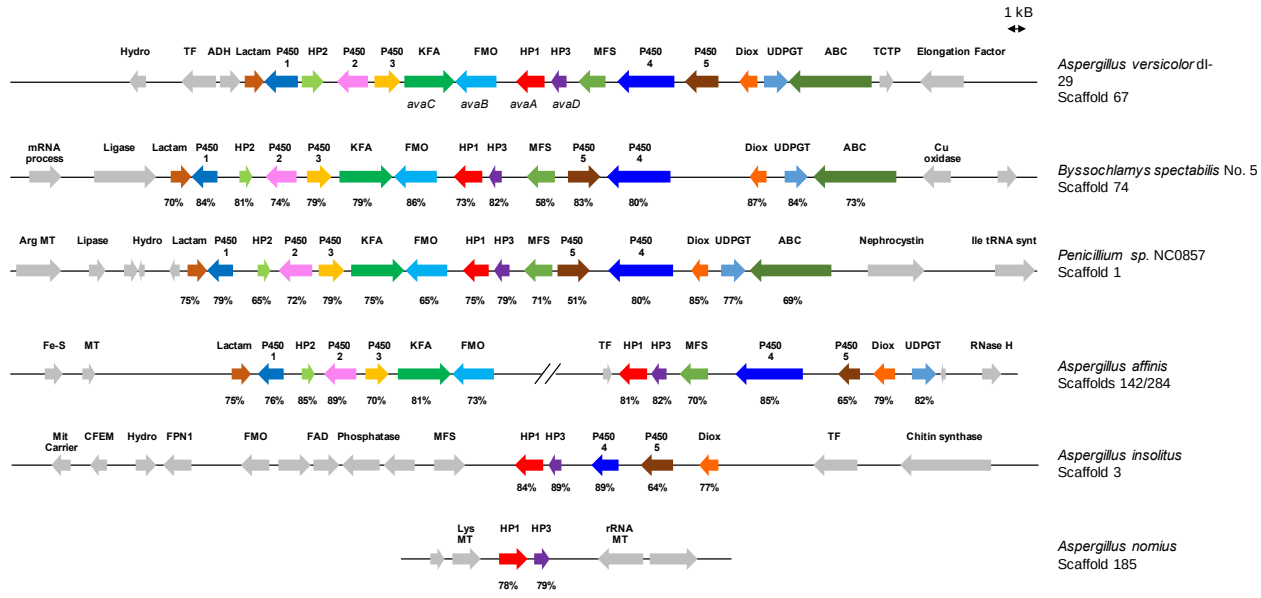


Figure S2. Homologous clusters in *Aspergillus versicolor* dI-29, *Byssochlamys spectabilis* No. 5, *Penicillium sp.* NC0857, *Aspergillus affinis*, *Aspergillus insolitus*, and *Aspergillus nomius*. Nearly all biosynthetic cluster genes are conserved in *A. versicolor* dI-29, *B. spectabilis* No. 5, *P. sp.* NC0857, and *A. affinis*.

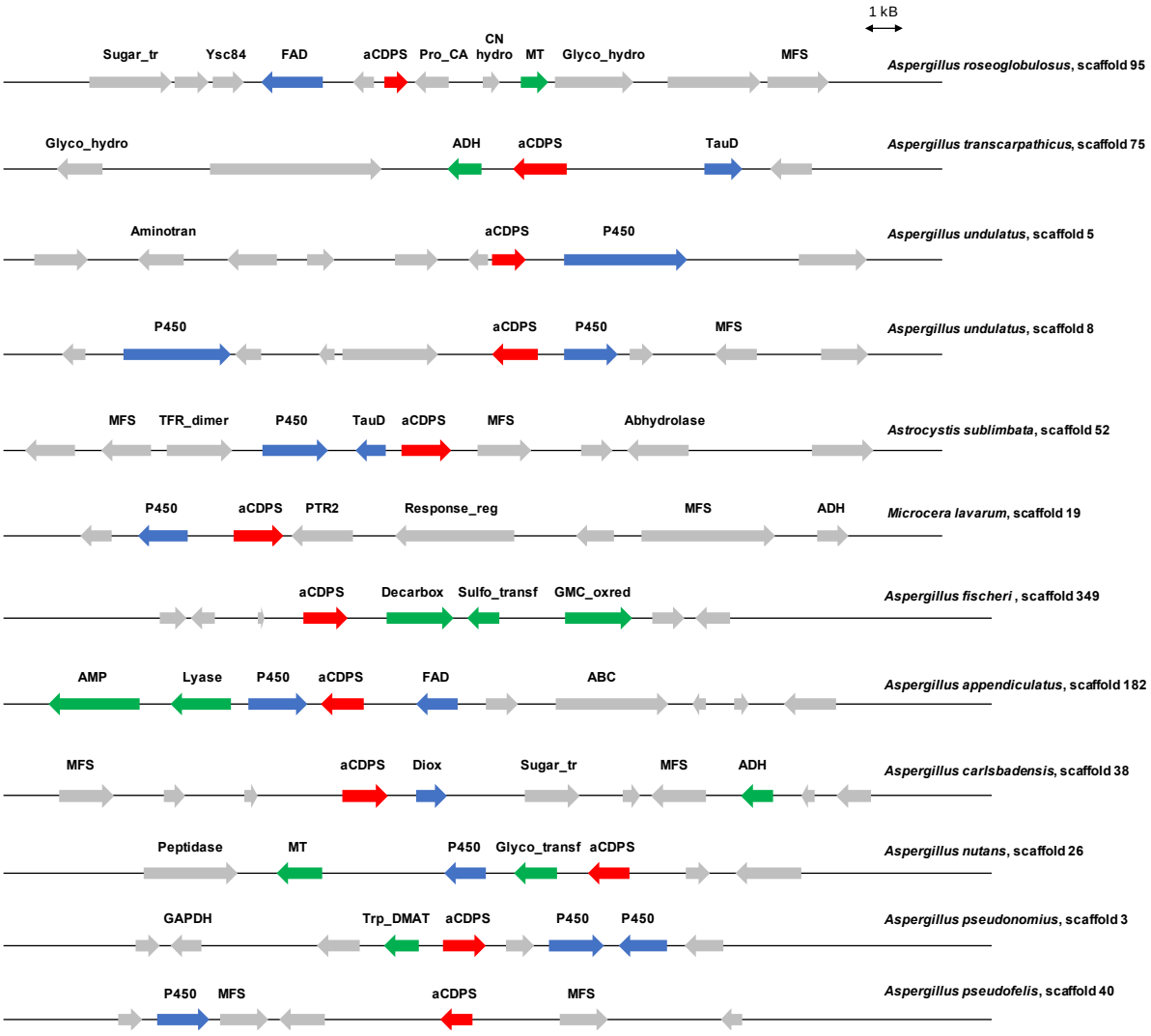


Figure S3. aCDPS homologs (red) clustered with biosynthetic tailoring genes. The clusters are putative. Potential oxidative enzymes are colored blue and other potential tailoring enzymes are colored green.

```
CLUSTAL O(1.2.4) multiple sequence alignment


A_versicolor_dI-29      MTATVQHTVKTEVCLRGDVASQPISDSKPPTW-VAVLQQHEYQSSVTYHKLDDNRLIPDV    59
A_nomius                MMGAVQLPVETEACLHDDAASRPISDSNT-VR-VAVLRQHQYHSSVTYHRLDDNRLIPGV    58
P_thymicola             --------MSTQYLLANL-RNRP--DGRPSIDFGSLSARE----SDPVK------MPHEI    39
A_terreus               --------MAPFYV--------------------------------VSIE------RQHGL    15
A_thermomutatus         --------MEDRIPVAD---------LQHSAGSLAAIAFHNI-EHCTQK------TSSAI    36
A_montagnei             -MVSISKSLVADICVGS---------SPPTVRFVTLLENHDI-SNTTFP------SLSGV    43
                                :                                               :


A_versicolor_dI-29      FGYSYLRDIADQGKKQSTG-------KQHARVIQAYSKIHSLLSPPWKSSSIDSN-----    107
A_nomius                VGYSYLRDIADEGKKQSPS-------KQHARVIQAYSKIHSLLSPPREPGPIDSE-----    106
P_thymicola             TRLFSSDYCTSTS-TTLEAEESEVYWKALSRNKRAAGILEHFLNGEQAAAHPSVYL----    94
A_terreus               SGLVSIVPCSTTPTIHYDEKDKKAYQKAWSRGRRFAALFENLLNGGKMATKTSCSL----    71
A_thermomutatus         LGYYHLPHIGSEHSSSDNADEAAKYTKAFIKGRKVAVKVANLINQGK-------------    83
A_montagnei             LAQYGSSLQGPAGLQHVGLAQDQKYQHALTKGLKAVEKLQHFLDGPAPMSGDPVVLRCSS    103
                                              :   :  :    .  ::.


A_versicolor_dI-29      --TKTRENKGSPVILEDKRCAFIERLEPPPNSKADIVSTVFAQVNLQTPTGPP-------    158
A_nomius                --TKTREKKSSPVIVHDKRCAFIERLEPPPNSKADIVNTVFAQVNLRTPVGPP-------    157
P_thymicola             -QDRARLRLCEEAILVGAPRGTATRT--------REGIASYDKMEFLFPESR------QR    139
A_terreus               -QQTPYLQLRDEALIVGSPLGTPTRT--------REVERIAGL-----------------    105
A_thermomutatus         TIHPVPLRVTGSGLVGGPPCEGHARL--------IEGITWFEKTQFHHQSLRP--EGNHS    133
A_montagnei             ARSSTLLEITDETITSGPPCAGAARE--------REGITSFHRMKFHLPKHVGCVQDVHE    155
                              .     :  .       *


A_versicolor_dI-29      -LE--QFLSCRASNIKANDLKN-AAEGVRRPITISTGICLFSSRLLGFIPTNGLSVSHDA    214
A_nomius                -LD--QFVNCRSSNIKINDLKD-SANGLCRPITISTGICLFSSRLLGFIPTNGLSTTDGA    213
P_thymicola             KSIQMLHVSCPSLQPGATLLDS-HSGSVHDSITILSGMCLFTRKILNSTT-TNGCLD---    194
A_terreus               --------SSAIADIAALVLSR-YDGKCHKPVNII-------------------------    131
A_thermomutatus         TASHYLHFSTPLAQIIPGAEGRTQPMDVVQDITLISGMCLWASRIIQSSMPRGQRLD---    190
A_montagnei             TSSQMLHLSCDGLRSEDTLFDR-FGRWSRDHVTIVSGMCLFSDKIMRDAVQGDGKLD---    211
                              .                       :.:


A_versicolor_dI-29      TETVVPPLPYTTVATFYELETCTRMAMTIAGLAVTA-STGGTAGSRPIVVRLDVPNLQYY    273
A_nomius                TETVIPPLPYSADATFYELETCARMAMTIAGLAVTA-GTGGTAGSRPIVVRLDVPCLQYY    272
P_thymicola             -----GQKLRHMSPTLYEMECIARLSAII----------------------ADVTTLTQS    227
A_terreus               ---------------------------------------------------LDLPTWHYY    140
A_thermomutatus         -----VSRLLSSKITLYEVEYLARAVPVIADLAAAL-SANNPYGPTPLNIRLDIPSFHYY    244
A_montagnei             -----VSKLSSQEPTLYELEIIARSSSAIADVASHILGKNGAAKPIPLEIVLDAPSWHYF    266
                                                                              *


A_versicolor_dI-29      CY--------PLELLEAGLVSWEYVEEWFRLVDRRHRQVATLLKDTIIHEVRRRNCDVQV    325
A_nomius                CY--------PLELLQAGLVSWKYVQEWFRLVDRRHRQVAGLLKDTITHEVLRRGGDIQV    324
P_thymicola             NFGEDNHPAVNIDCIKSDACSLEEAIDWTEAIKLRRQQLASVLKKAVWHELGQRQVAHKS    287
A_terreus               QS--------VEDQLQHRGYKPCEVMDWIEAITLRSQQLAGLLKSAVLHELDRRGVSSSQ    192
A_thermomutatus         HS--------LEERLRDGCCTFPEALQWMHAVEKRHHQLSRVFCRLIDHELSRRWVGTPH    296
A_montagnei             QV--------VHGNLASGHCTPAEALDWLQAVELRCEQVTTVFENSVRHEMGLRGVPAGF    318
                                  :     .   . :* . :  * .*:: ::   : **:  *


A_versicolor_dI-29      ------DVTSGTIAATQLLRLCVLGRRKIPSVNDMLFVLSWIGPYQAAWREFLAIVDDCQ    379
A_nomius                ------EVTAGTIAATQLLRLSVFDRMELPSVDDILFVLKWVGPYQAAWREFLDILDDCQ    378
P_thymicola             -TLKAIQISPESTVVDELIK-ETLQRGHQPRLDDILHAL--SGTPHKI------------    331
A_terreus               -RLYDIQISPGSALVDDAFR-ETLKYENLPCLDNILEAL--SGSEDGSWQQFYSLLPERD    248
A_thermomutatus         RRKLDVQVSPLADLVFQLIC-DSLANSVLPDVDDIL-QI--VQTEDTTWVRFYSLVPENE    352
A_montagnei             YHIL---AAPGTAGVGTSIR-QALTSGMVPDIADCMDAI--CEVEGERWAMFYSLIPEKD    372
                               :  :  .   :    :     * : : :  :
```

```
A_versicolor_dI-29      RPKDLRSLALMAYVFEVMYPALQQATTKTKTPH----------------------GKG   415
A_nomius                RPKDLRSVALMAYVFEVMYPALHQIATKS--LH----------------------GNG   412
P_thymicola             -----SGFS----MYFKFRPALLKAIPCQEAPS--CKA-------PACEQPNLRR--GQR   371
A_terreus               RPQSIKDLSYLFYIFEVVRPAVLAGIQSGNQQTPTQSADMDSTVSHRQQQPASSRSYTSK   308
A_thermomutatus         KVTDFRALSYLFYVYQVLRPALEQTHVNTTEG-------------------------   384
A_montagnei             RPCDFRSLGNLFYIYEVVRPALAAKNGSTAAELEVSGPNSALDDSSSSGHSDILDSALHL   432
                                ..    ::  . **:

A_versicolor_dI-29      EESSGRPLLIQVDDIAEWRIFDRAEMLLKRFKHR--QHGLDPLLVGVFPSPRIFTSEDQG   473
A_nomius                -EKSGRPLLIQVDDIAEWRIFDHAEKLLKRFKER--QHGFYPLLVGVFPSPRIFTSEDQG   469
P_thymicola             QCLPQLPLIISLDDRAERKIYSQAHSLLLRLSRS-SNQLVNPTLVQLYMLRRVYIDGNKD   430
A_terreus               QNQMPRPLVISVDDPSERKLYSKAHSFFLRLPKN-PIYPADPTLVQVYTTRRVFVDGNRK   367
A_thermomutatus         -DEATSKLLISVDDAFERRIYSRSQKLLKKLRASLPASAAVPHLLEVYLCRRIFINSNET   443
A_montagnei             KANVERPLIMSIDDRAERKIYSKAQAFIKKIRRA-PQIPSHPLLLELYSARRVFINGNTG   491
                               *::.:**  * :::.::. :: ::          * *: ::   *:: . :

A_versicolor_dI-29      RSTLFLHDPGLKISQTRPPSSG--------------GDSEDHSCVVGPLDIIGQIYGAEV   519
A_nomius                RSTLFLHDPGLKILQTRSPSSN----------------SEEGSCVVRPLDIIGQIYGREV   513
P_thymicola             GQRLYWHDPSPVLPLLEGTKEQQ---------RRDARNHDHHRRELQQTDILTELYGREC   481
A_terreus               GERLYRNDPSPVMPELSNGKLYG---------DGGDAGYSKYARELQQMDFIAKLYGAEC   418
A_thermomutatus         GSNLYLDDPSPEPFVLRLCQGLAPEGKQELVRAQDRHNTRPETIKLDAFDVVEGLYGSHI   503
A_montagnei             GDSLYWNDSSPHPLRMNTTA--------------------GSPELEAFGVIRNLYGNDC   530
                         . *: .* .                                :   ..:  :** .

A_versicolor_dI-29      QDTLMRLIIERGLSPEDEFELD  541
A_nomius                QDTLGQLTRKHGLSAADGPESD  535
P_thymicola             SANIQGWFKEAGLC--------  495
A_terreus               AANIQRWSKEVGLC--------  432
A_thermomutatus         AEVLKDLFAEVGLGA-------  518
A_montagnei             AQNLQRWFSEAGVNVSGHT---  549
                           :     : *:
```

Figure S4. Clustal omega (1.2.4) multiple sequence alignment of aCDPS homologs with activity verified *in vivo*.

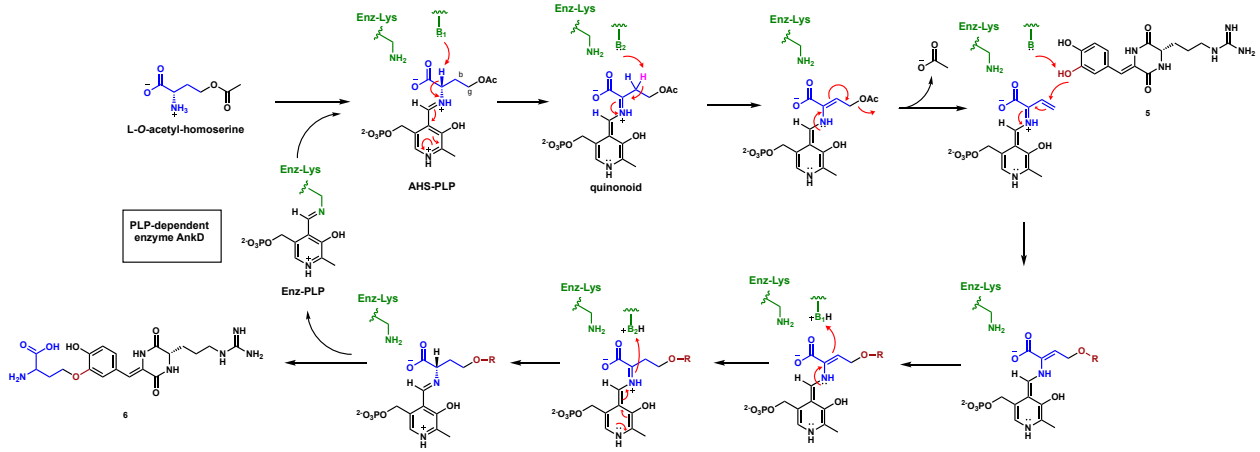


Figure S5. Proposed pathway for the biosynthesis of **6** from L-*O*-acetyl-homoserine and **5** by PLP-dependent enzyme AnkD.

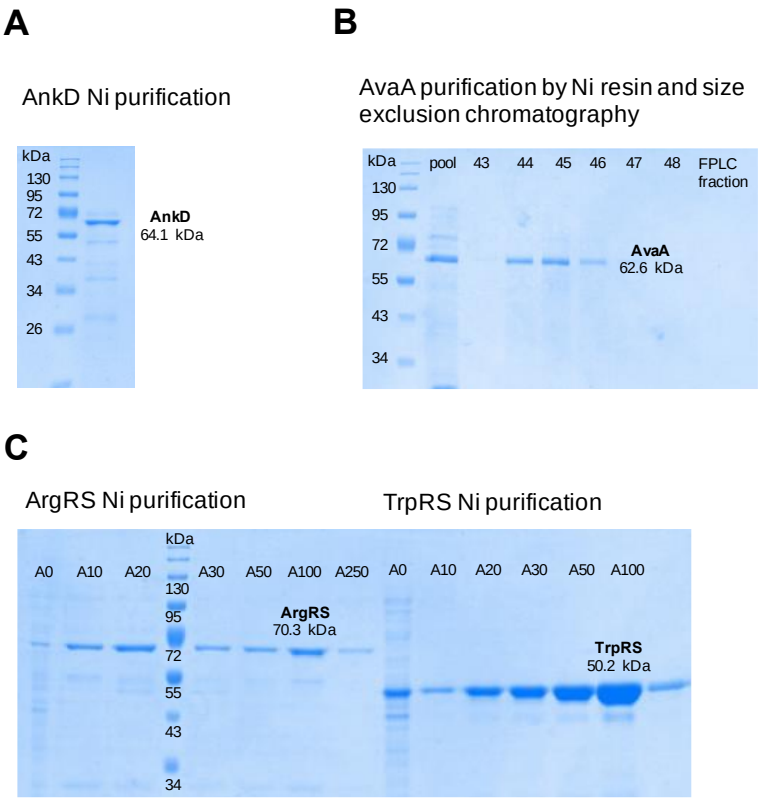


Figure S6. SDS-PAGE (12%) analysis of (A) AnkD, (B) AvaA, (C) ArgRS and TrpRS expressed and purified from *E. coli* BL21(DE3).

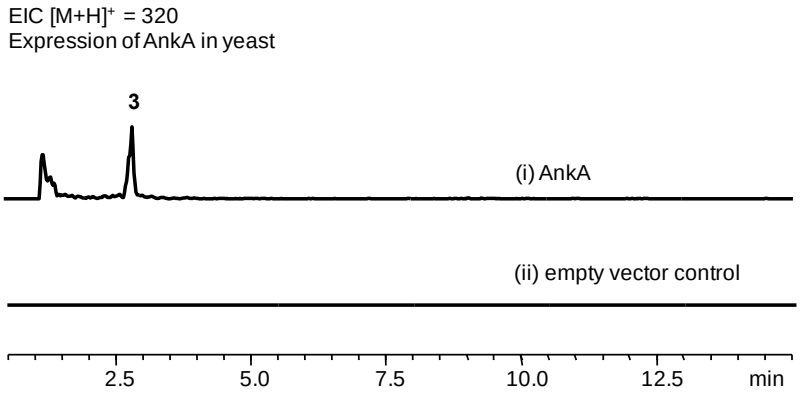


Figure S7. LC-MS analysis of extracts from expression of AnkA in yeast (i) compared to the empty vector control (ii). Extracted QTOF traces correspond to the $[M+H]^+$ for **3** ($[M+H]^+$ = 320).

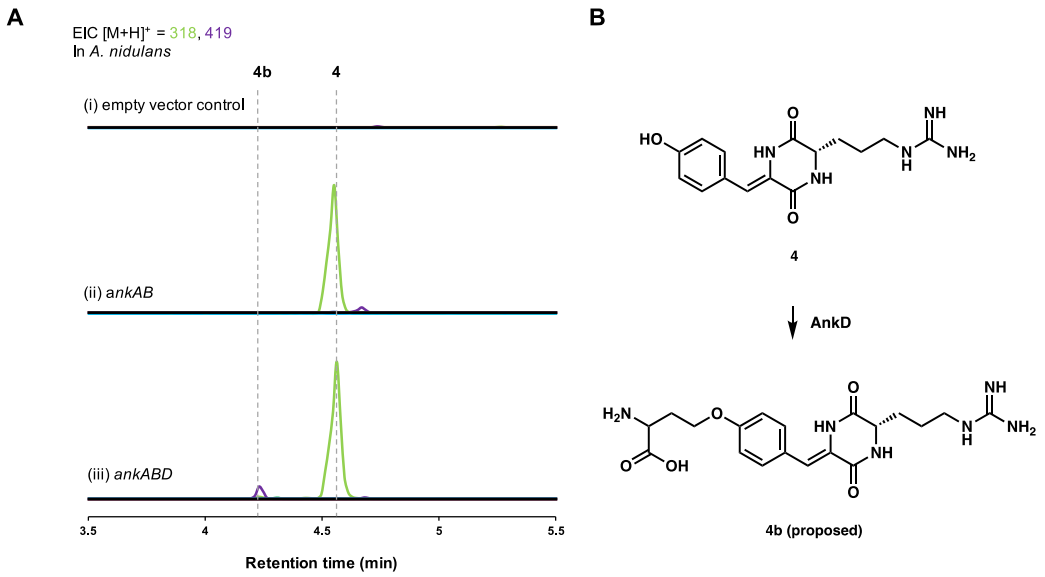


Figure S8. (A) QTOF analysis of the extracts from *in vivo* production of **4** and **4b** in *A. nidulans* with (i) empty vector control, (ii) expression of AnkAB, and (iii) expression of AnkABD. Extracted QTOF traces correspond to the $[M+H]^+$ for **4** ($[M+H]^+$ = 318) and **4b** ($[M+H]^+$ = 419). (B) The proposed reaction by AnkD from **4** to **4b**.

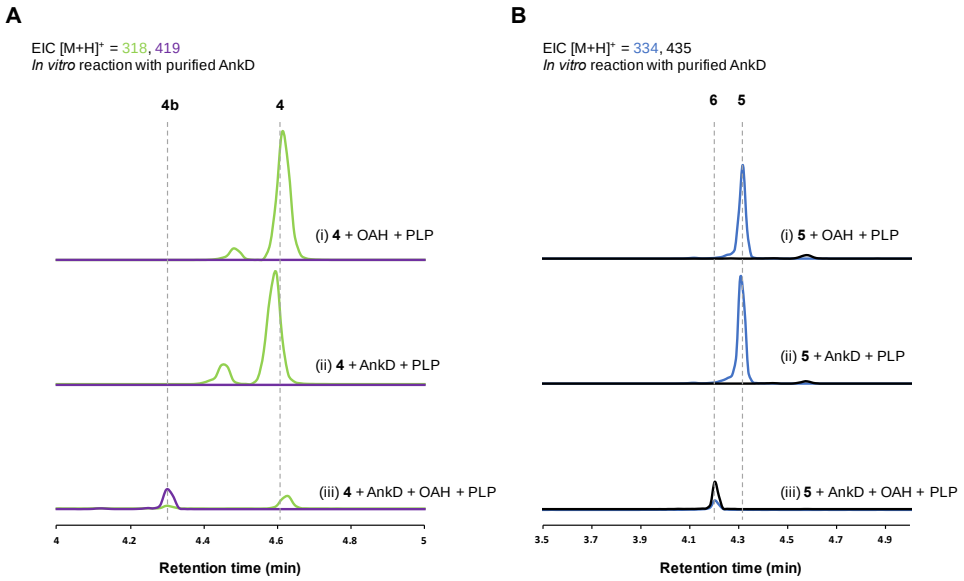


Figure S9. QTOF analysis of the extracts from *in vitro* reactions with (A) **4** and (B) **5** with (i) OAH and PLP, (ii) purified AnkD and PLP, and (iii) purified AnkD, OAH, and PLP. Extracted QTOF traces correspond to the $[M+H]^+$ for **4** ($[M+H]^+$ = 318), **4b** ($[M+H]^+$ = 419), **5** ($[M+H]^+$ = 334), and **6** ($[M+H]^+$ = 435).

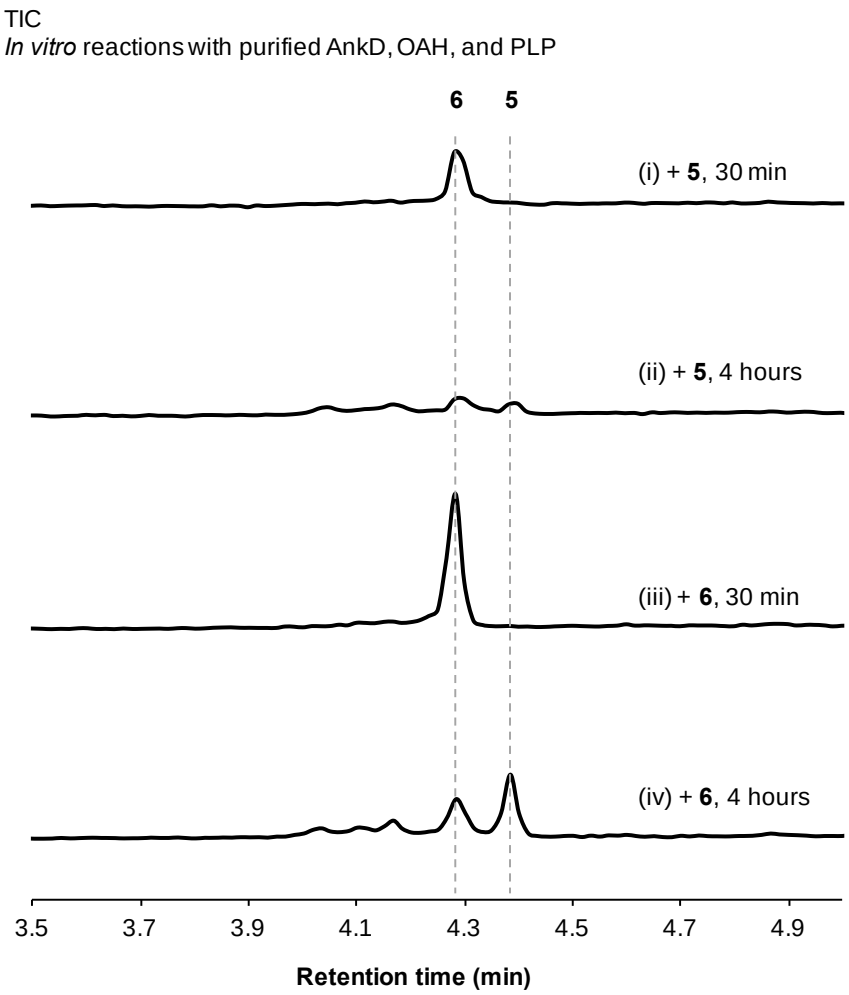


Figure S10. QTOF analysis of the extracts from *in vitro* reactions with purified AnkD, OAH, and PLP at different time points with addition of (i) **5**, 30 min incubation, (ii) **5**, 4-hour incubation, (iii) **6**, 30 min incubation, (iv) **6**, 4-hour incubation. The longer 4-hour incubation of AnkD with **6** resulted in the degradation of the compound, the re-emergence of compound **5**, and the appearance of new peaks with earlier retention times, which may be degradation products. QTOF traces correspond to the total ion count (TIC).

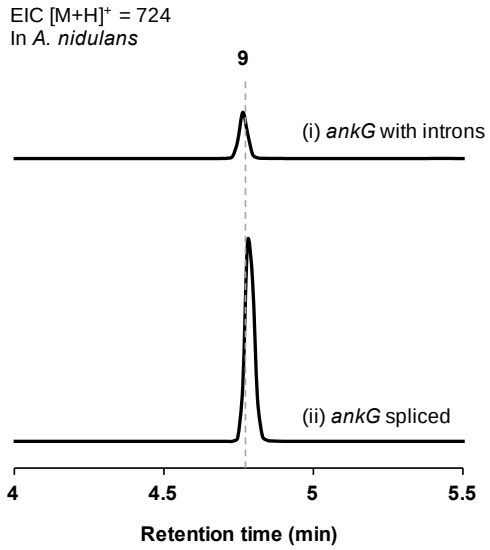


Figure S11. QTOF analysis of the extracts from *in vivo* production of **9** in *A. nidulans* transformed with *ankA-E* alongside (i) *ankG* containing introns or (ii) the spliced sequence of *ankG*. Extracted QTOF traces correspond to the $[M+H]^+$ for **9** ($[M+H]^+$ = 724).

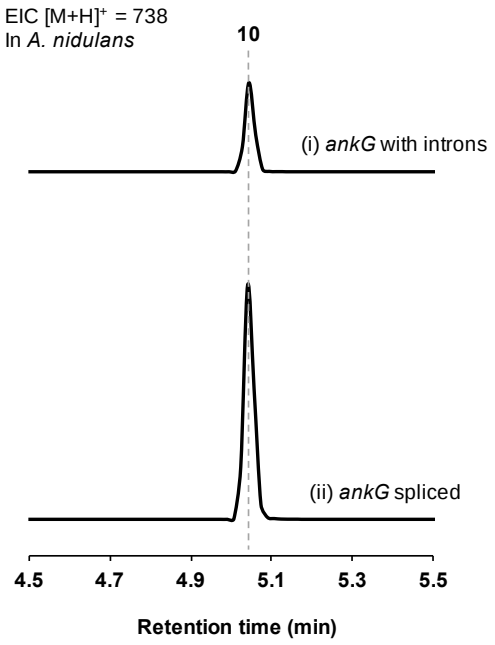


Figure S12. QTOF analysis of the extracts from *in vivo* production of **10** in *A. nidulans* transformed with *ankA-F* alongside (i) *ankG* containing introns or (ii) the spliced sequence of *ankG*. Extracted QTOF traces correspond to the $[M+H]^+$ for **10** ($[M+H]^+$ = 738).

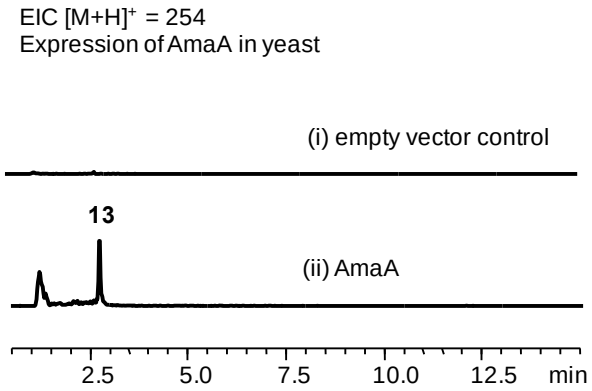


Figure S13. LC-MS analysis of extracts from expression of AmaA in yeast (ii) compared to the empty vector control (i). Extracted LC-MS traces correspond to the $[M+H]^+$ for 1**3** ($[M+H]^+$ = 254).

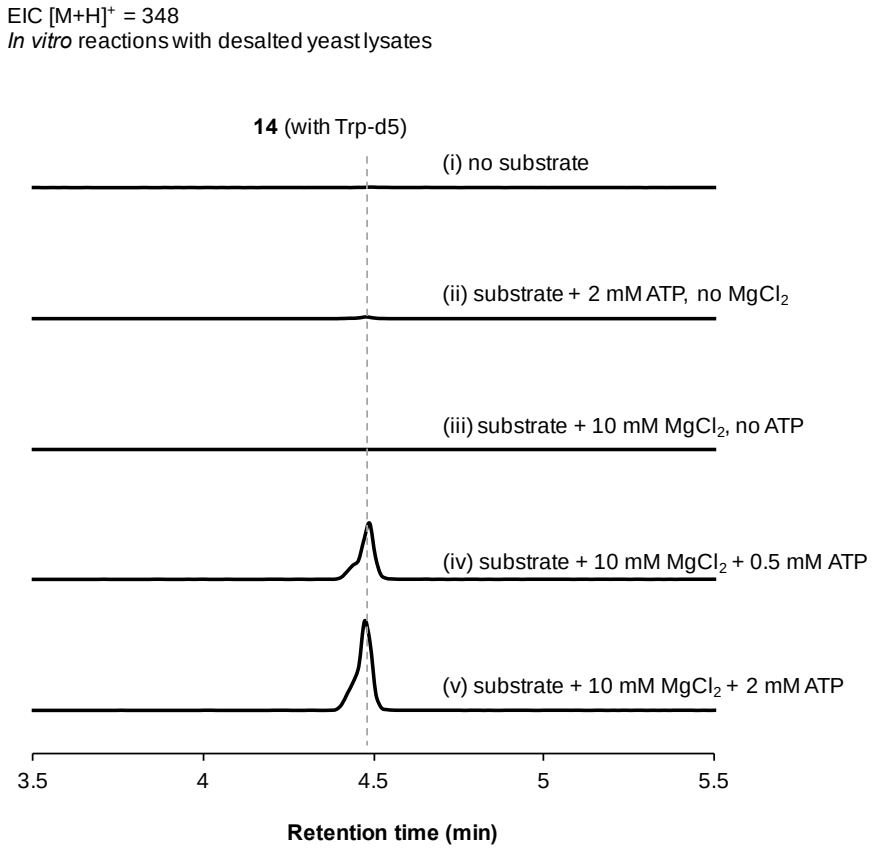


Figure S14. QTOF analysis of the extracts from *in vitro* reactions with desalted yeast lysates with (i) no substrate or cofactors, (ii) 5 mM L-Arg, 5 mM Trp-d5, 2 mM ATP, no $MgCl_2$, (iii) 5 mM L-Arg, 5 mM Trp-d5, 10 mM $MgCl_2$, no ATP, (iv) 5 mM L-Arg, 5 mM Trp-d5, 0.5 mM ATP, 10 mM $MgCl_2$, and (v) 5 mM L-Arg, 5 mM Trp-d5, 2 mM ATP, 10 mM $MgCl_2$. Extracted QTOF traces correspond to the $[M+H]^+$ for 1**4** with incorporation of Trp-d5 ($[M+H]^+$ = 348).

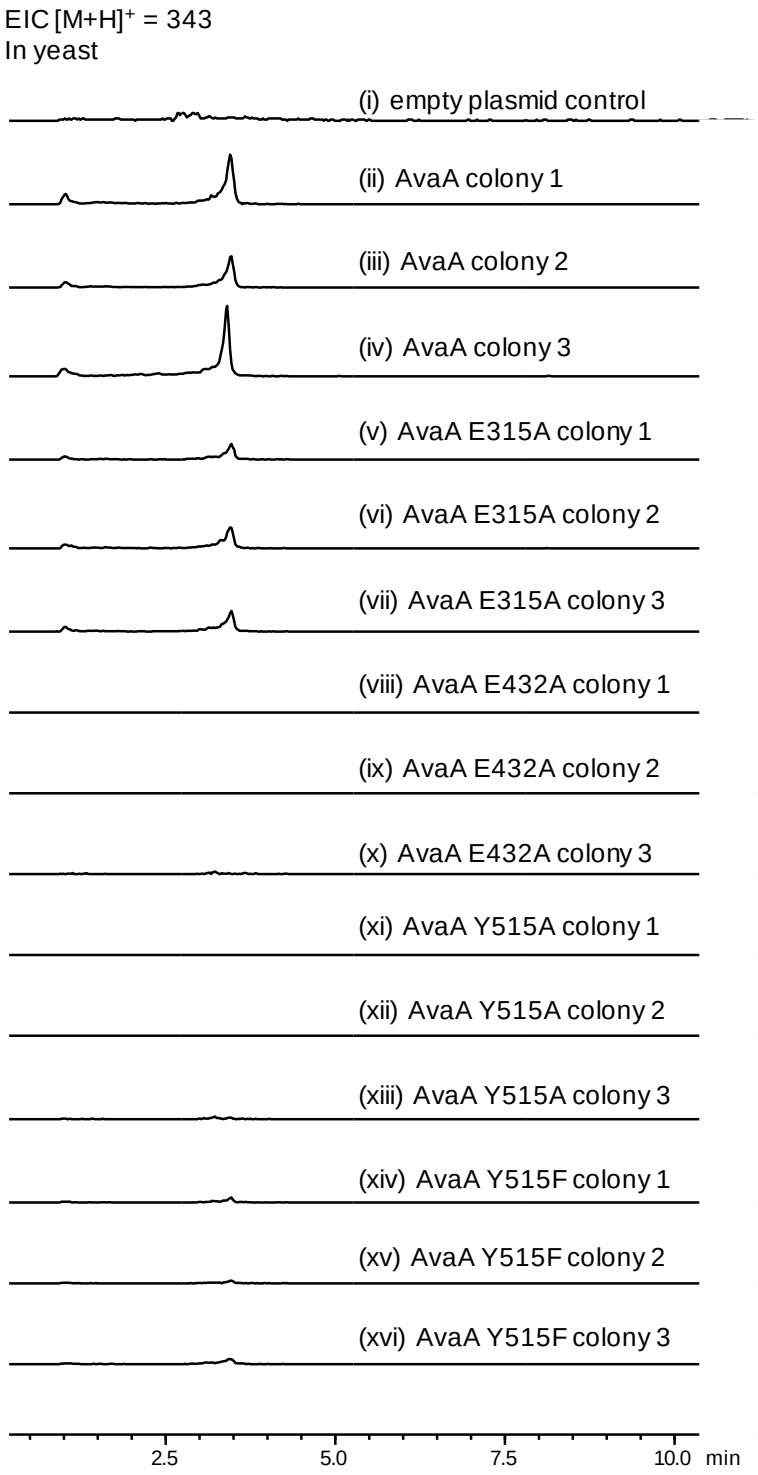


Figure S15. LC-MS analysis of the extracts from yeast expression of AvaA with (i) empty vector control, (ii-iv) wildtype AvaA in biological triplicate, and (v-xvi) AvaA mutants E315A, E432A, Y515A, and Y515F in biological triplicate. Extracted LC-MS traces correspond to the $[M+H]^+$ for 1**4** ($[M+H]^+ = 343$).

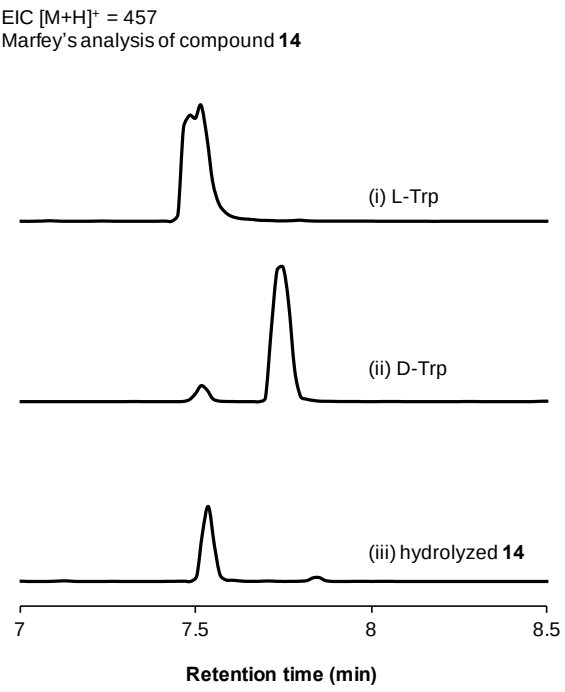
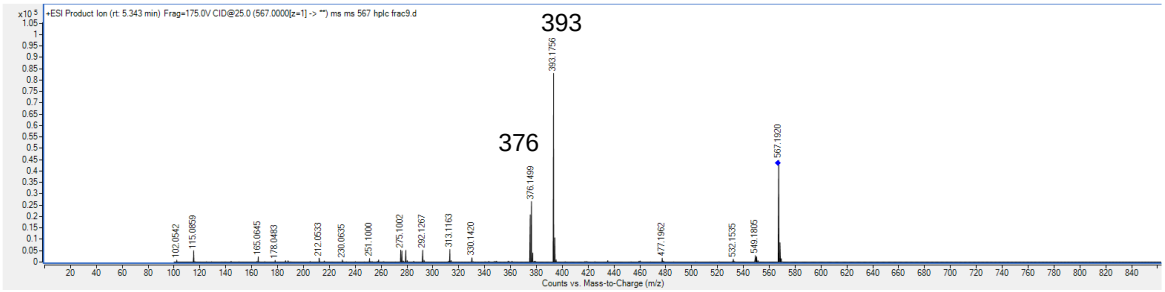
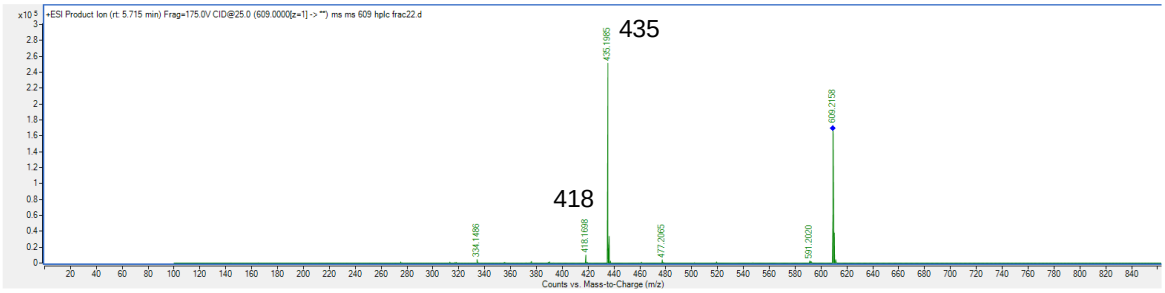


Figure S16. Marfey's analysis to determine the stereochemistry of Trp in compound **14**. Marfey's reagent was used to derivatize (i) L-Trp, (ii) D-Trp, and (iii) hydrolyzed **14**, and the reaction products were analyzed by QTOF, indicating **14** contains L-Trp. Extracted QTOF traces correspond to the $[M+H]^+$ for derivatized Trp ($[M+H]^+ = 457$).

A



B



C

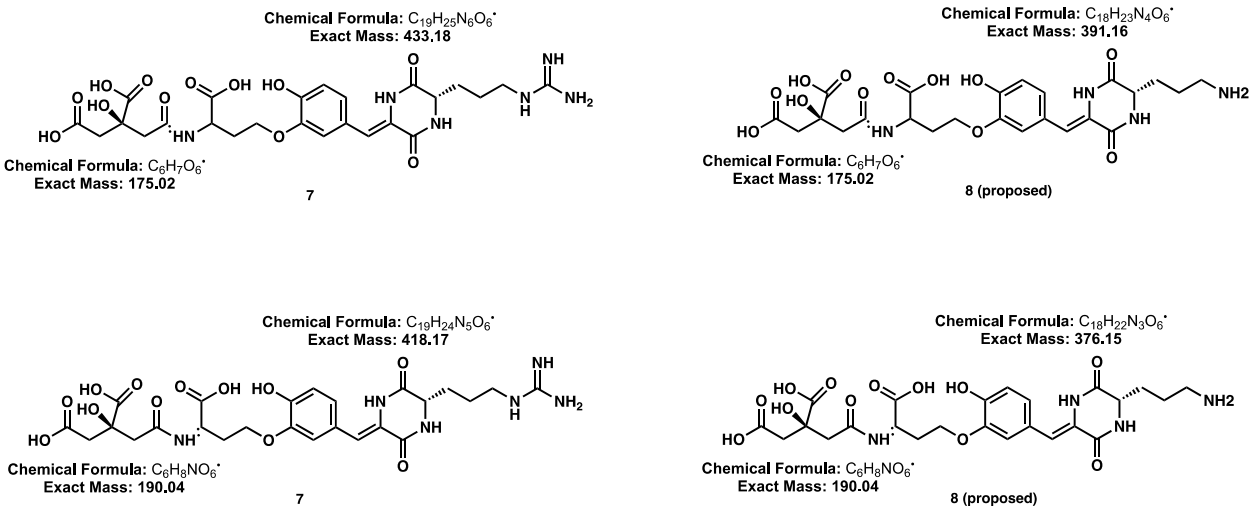
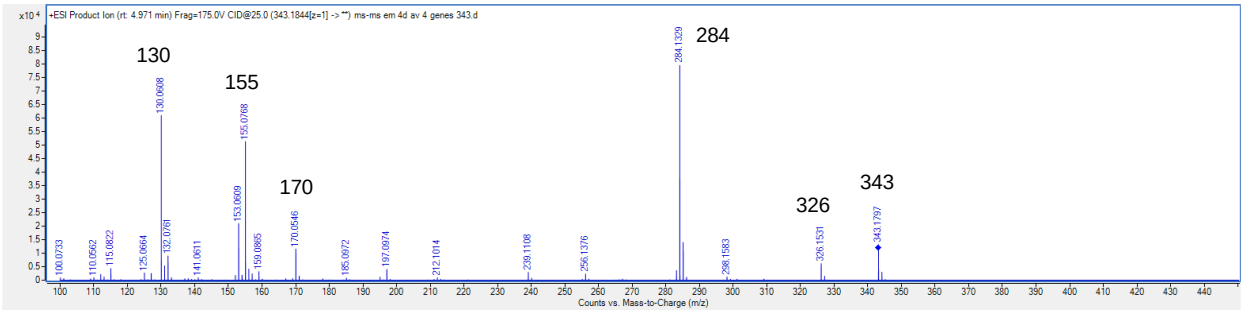
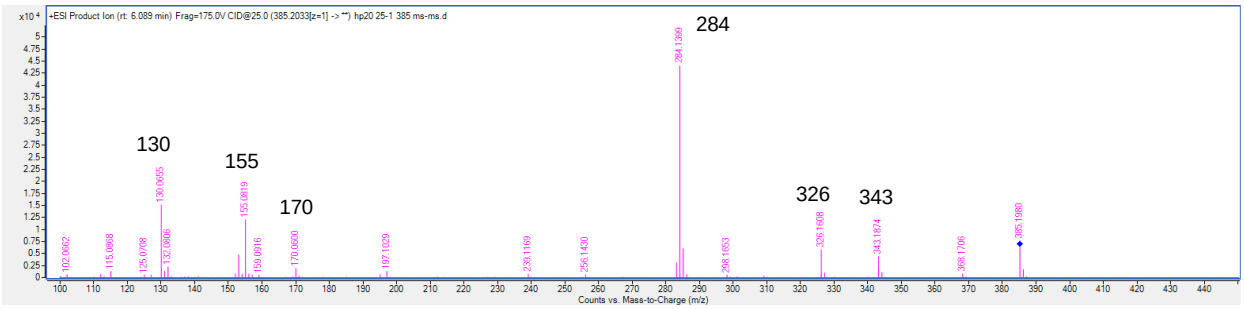


Figure S17. QTOF MS-MS analysis of compounds (A) **8** and (B) **7**, which suggests **8** is a derivative of **7**, as a -42 shift in mass is observed for the major ion fragments 435 and 418. The 42-mass decrease is in accordance with the proposed hydrolyzed structure. The predicted ion fragments are shown in (C).

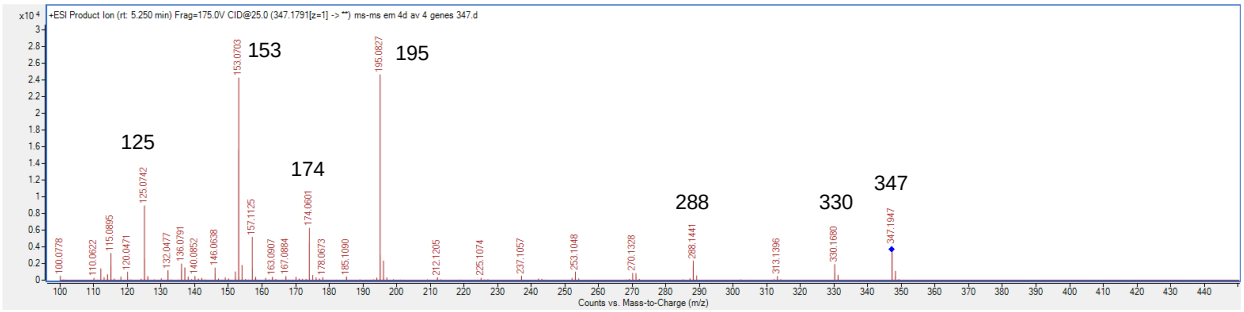
A



B



C



D

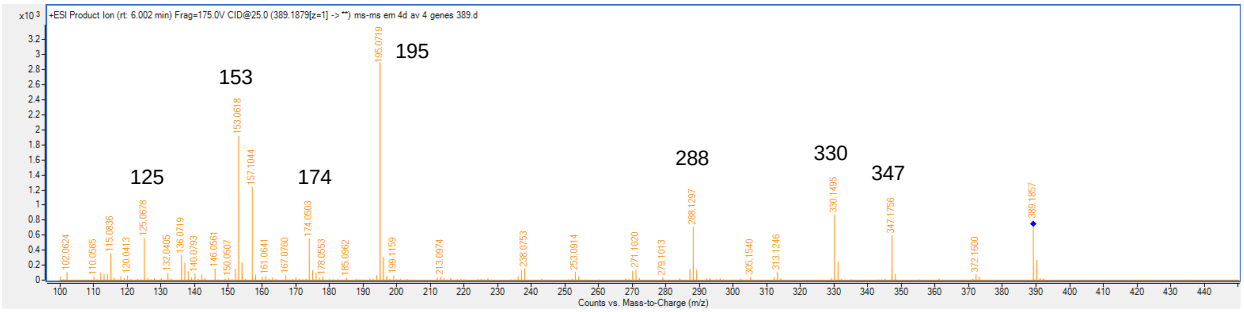


Figure S18. QTOF MS-MS analysis of compounds (A) **14**, (B) **20**, (C) **17**, and (D) **21**. The fragmentation patterns of **14**, **20** and **17**, **21** are very similar, suggesting **20** is an analog of **14** and **21** is an analog of **17**.

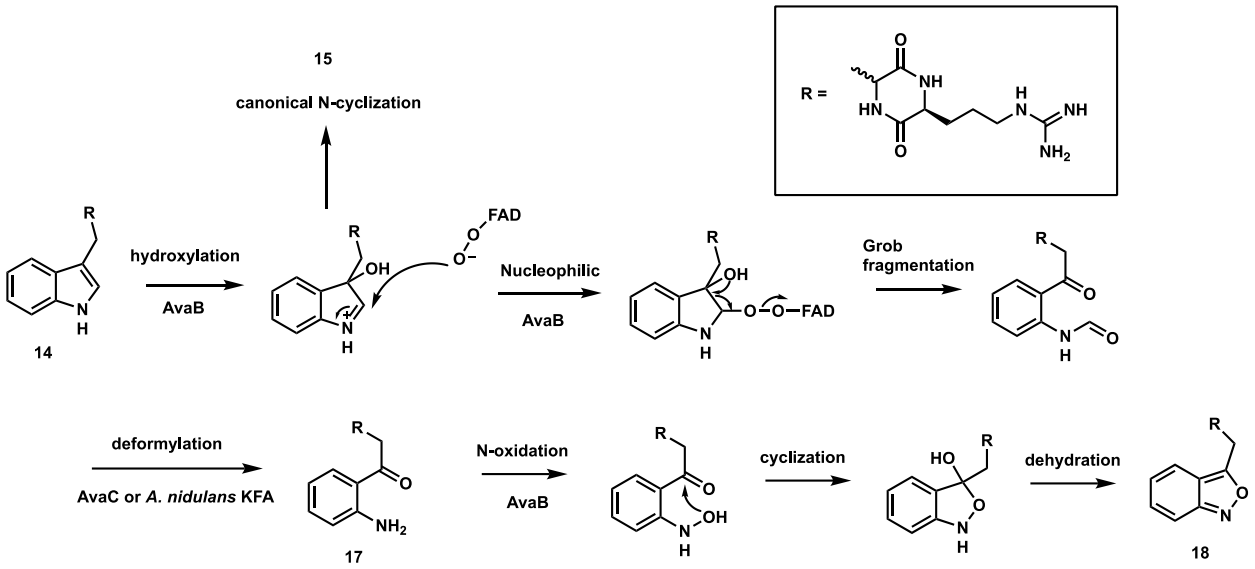


Figure S19. The proposed mechanism for the FMO AvaB. AvaB first hydroxylates **14**, and the resulting intermediate can undergo spontaneous N-cyclization to afford the 6-5-5 ring system in **15**. However, AvaB can further oxidize the intermediate through nucleophilic attack to generate the peroxide bound species, which can undergo Grob fragmentation followed by deformylation by AvaC or *A. nidulans* kynurenine formamidase to generate **17**. We propose AvaB next catalyzes N-oxidation, and the hydroxylated intermediate undergoes spontaneous cyclization and dehydration to afford **18**.

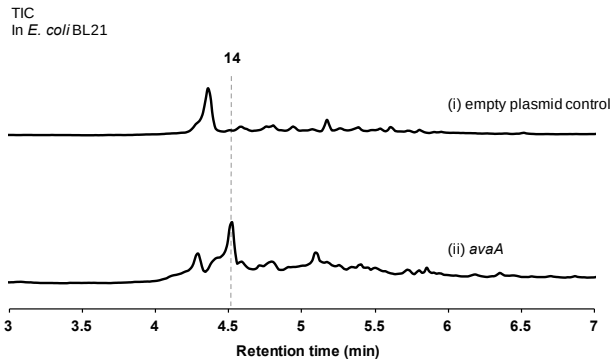


Figure S20. QTOF analysis of the extracts from (i) expression of *avaA* and (ii) the empty plasmid control in *E. coli* BL21.

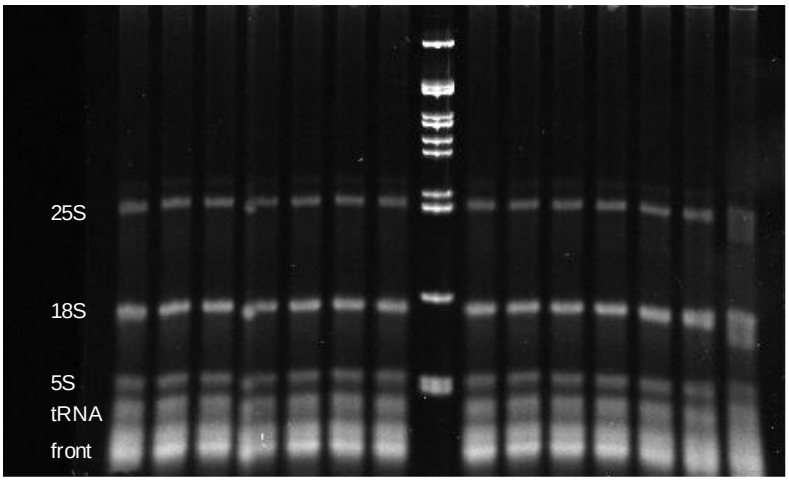


Figure S21. RNA-PAGE purification of total yeast RNA for isolation of tRNA.

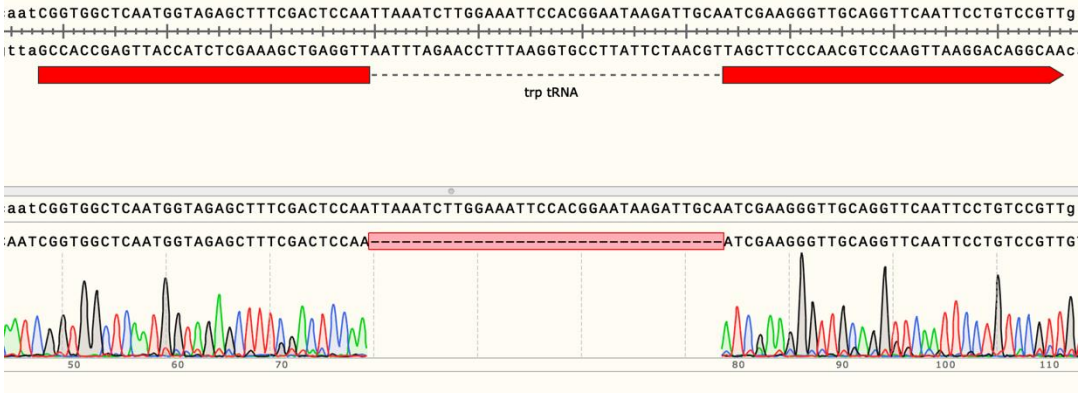
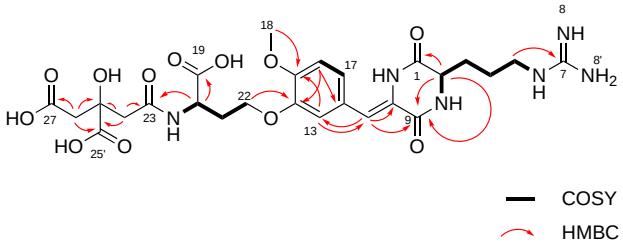


Figure S22. Sanger sequencing of RT-PCR product of Trp tRNA from the isolated tRNA band. The sequencing results showed the intron was removed, indicating the transcript was not amplified from genomic DNA contamination.

Table S7: Spectroscopic data of **1**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	168.6, C			
2	54.8, CH	4.32, t (5.1)	1.98, 1.89	168.6, 162.4, 30.8, 22.8
3	30.8, CH2	Ha: 1.98, m; Hb: 1.89, m	4.32, 1.71, 1.63	54.8
4	22.8, CH2	Ha: 1.71, m ; Hb: 1.63, m	3.21, 1.98, 1.89	40.4
5	40.4, CH2	3.21, t (7.0)	1.71, 1.63	156.7, 30.8, 22.8
6	NH			
7	156.7, C			
8	NH			
9	162.4, C			
10	123.6, C			
11	118.6, CH	6.87, s		162.4, 123.6, 114.0
12	125.3, C			
13	114.0, CH	7.05, s		149.3, 123.0, 118.6
14	147.3, C			
15	149.3, C			
16	112.3, CH	7.08, d (8.6)	7.12	147.3,123.0
17	123.0, CH	7.12, d (8.6)	7.08	149.3, 114.0
18	55.7, CH3	3.88, s		149.3
19	174.9, C			
20	50.0, CH	4.64, dd (4.4, 8.9)	2.40, 2.19	174.9, 170.9, 65.3
21	30.1, CH2	Ha: 2.40, m; Hb: 2.19, m	4.64, 4.21, 4.13	
22	65.3, CH2	Ha: 4.21, m; Hb: 4.13, m	2.40,2.19	147.2, 50.0
23	170.9, C			
24	44.2, CH2	Ha: 2.83, d (14.4); Hb: 2.74, d (14.4)	Ha: 2.74; Hb: 2.83	176.4, 170.9, 73.4, 42.8
25	73.4, C			
26	42.8, CH2	Ha: 2.97, d (16.6); Hb: 2.76, d (16.6)	Ha: 2.76; Hb: 2.97	176.4, 173.3, 73.4, 44.2
27	173.3, C			
25'	176.4, C			
8'	NH2			
	112.78, 115.1, 117.4, 119.7 - TFA			



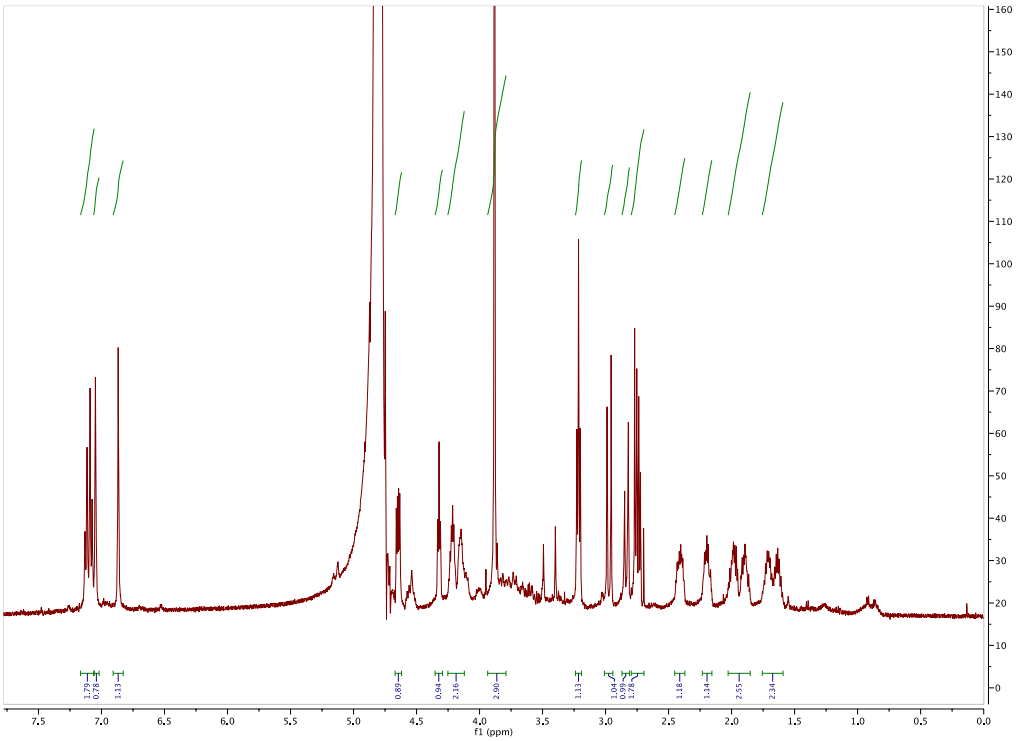


Figure S23. The ^{1}H NMR spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz)

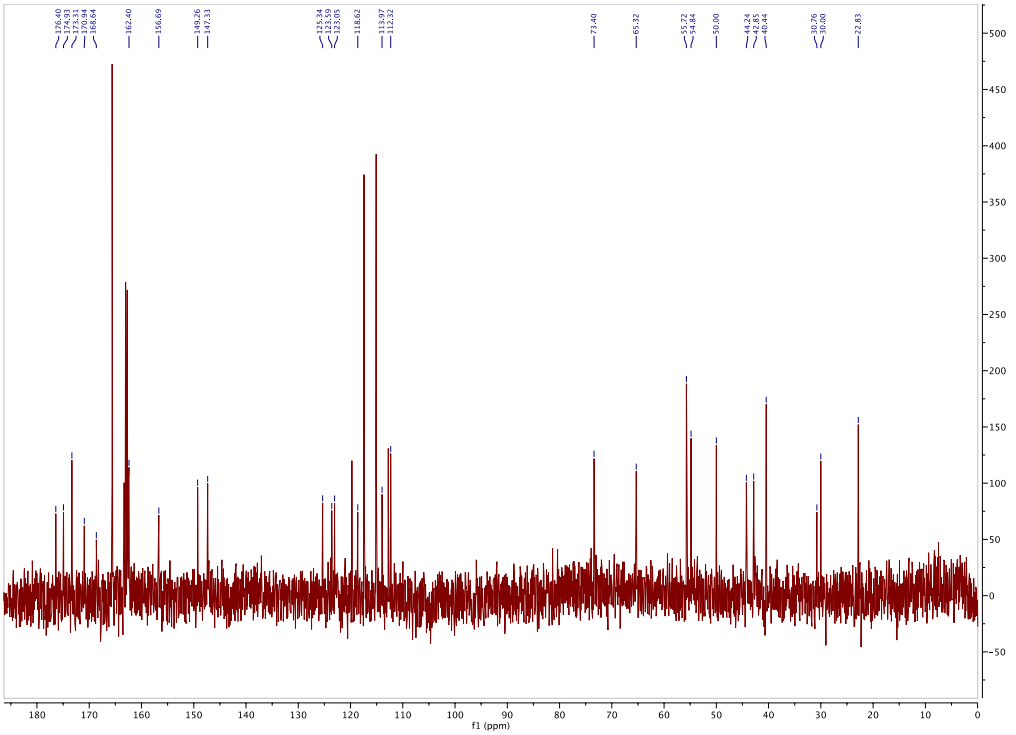


Figure S24. The ^{13}C NMR spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz)

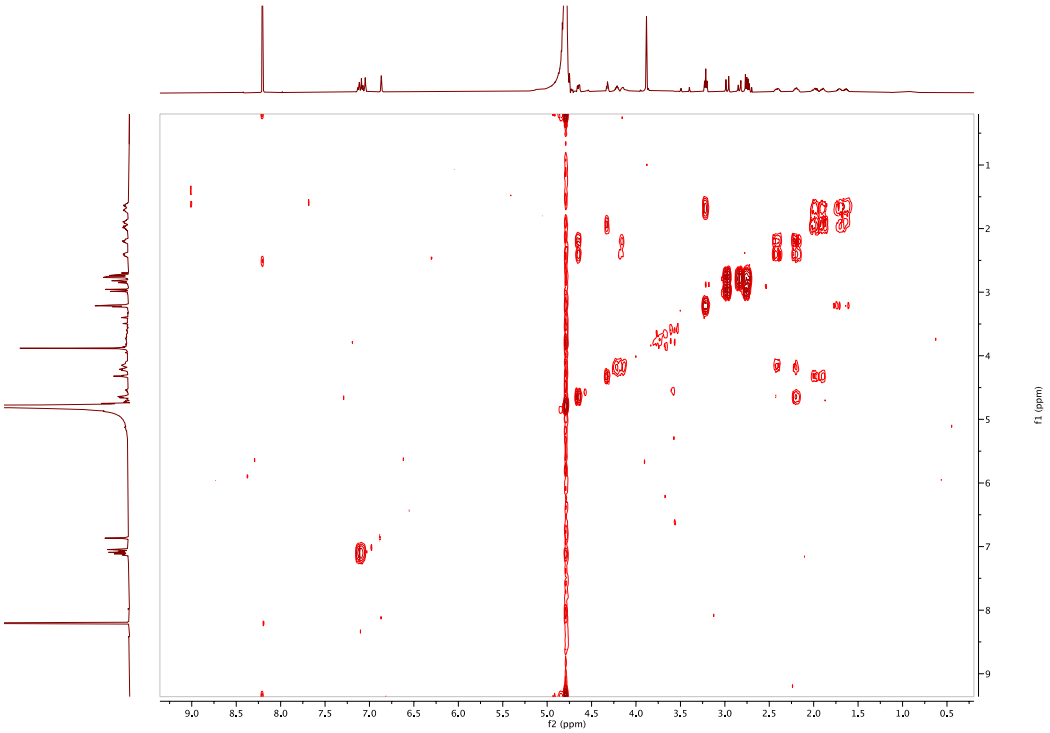


Figure S25. The COSY spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz)

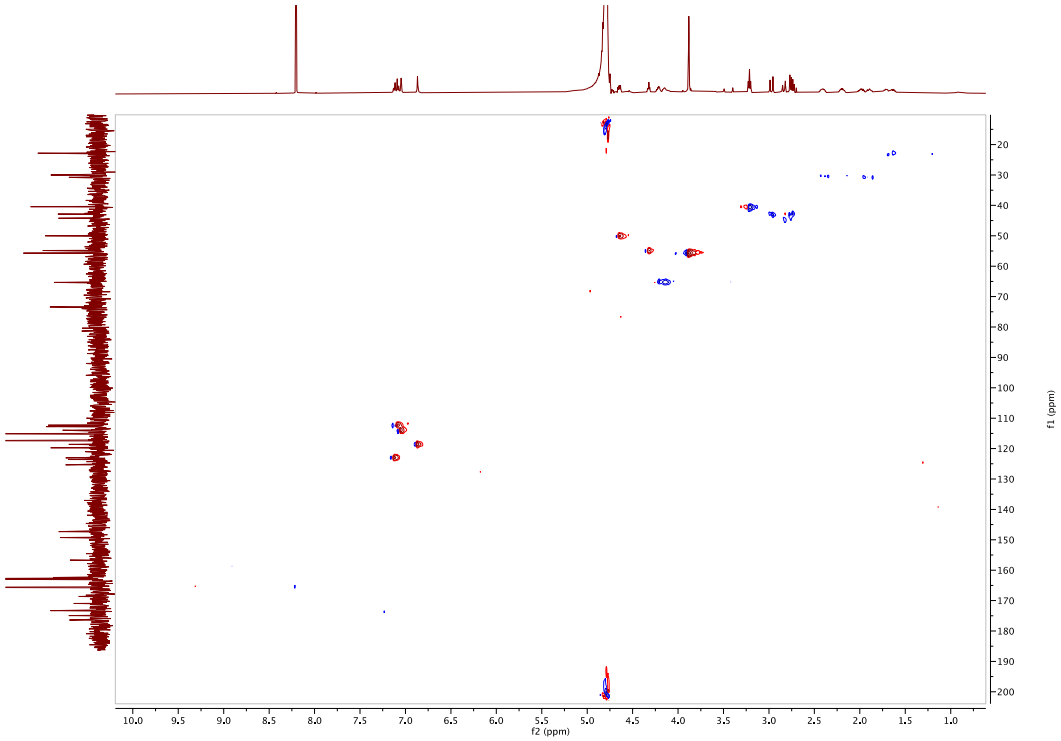


Figure S26. The HSQC spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz)

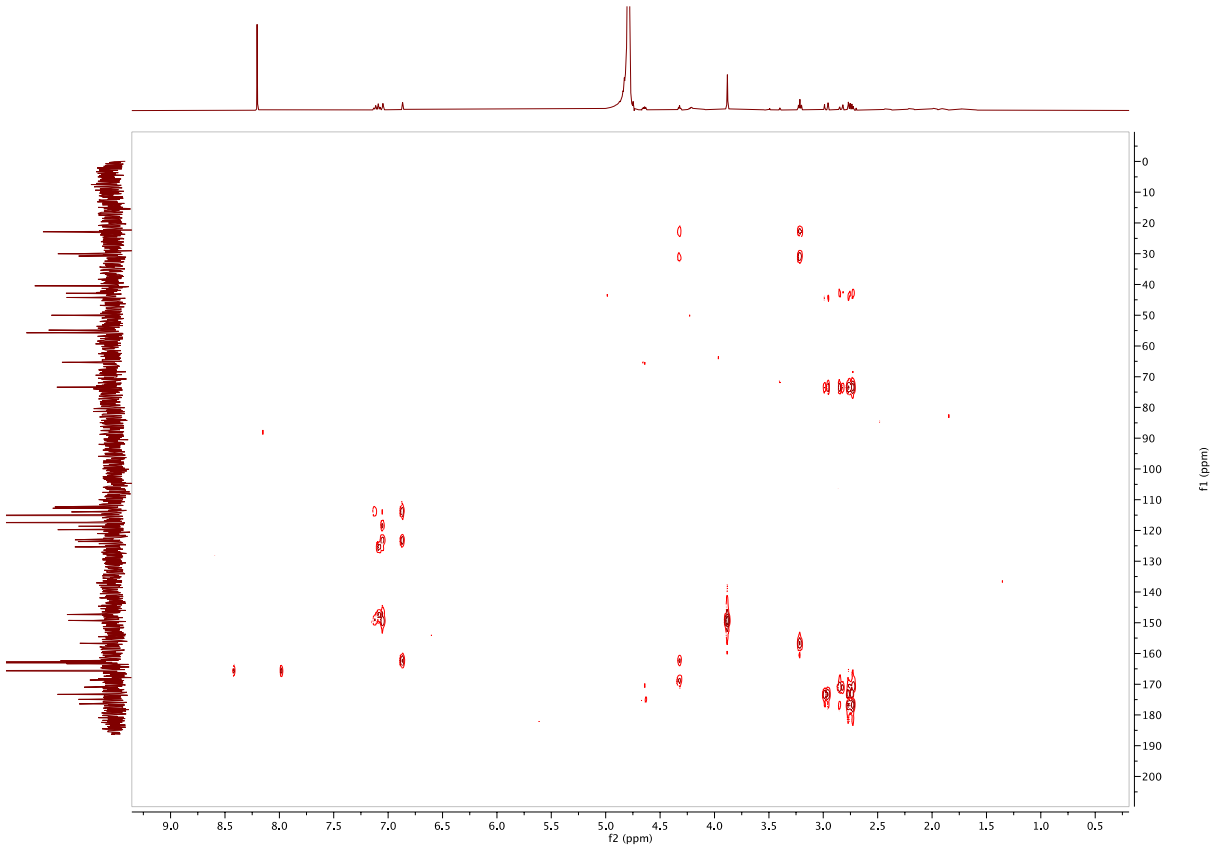
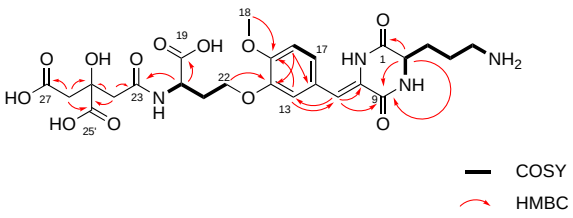


Figure S27. The HMBC spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz)

Table S8: Spectroscopic data of **2**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	168.2, C			
2	54.6, CH	4.32, t (5.1)	1.98, 1.89	168.2, 162.4, 30.5, 21.7
3	30.5, CH2	Ha: 1.96, m; Hb: 1.87, m	4.32, 1.71, 1.63	38.8, 54.6, 168.2
4	21.7, CH2	Ha: 1.80, m ; Hb: 1.74, m	3.21, 1.98, 1.89	30.5, 38.8, 54.6
5	38.8, CH2	3.04, t (7.0)	1.71, 1.63	21.7, 30.5
6	NH			
8	NH			
9	162.4, C			
10	123.6, C			
11	118.6, CH	6.87, s		162.4, 123.6, 114.0
12	125.3, C			
13	114.0, CH	7.05, s		149.3, 123.0, 118.6
14	147.3, C			
15	149.3, C			
16	112.3, CH	7.08, d (8.6)	7.12	147.3,123.0
17	123.0, CH	7.12, d (8.6)	7.08	149.3, 114.0
18	55.7, CH3	3.88, s		149.3
19	174.9, C			
20	50.0, CH	4.64, dd (4.4, 8.9)	2.40, 2.19	174.9, 170.9, 65.3, 30.1
21	30.1, CH2	Ha: 2.40, m; Hb: 2.19, m	4.64, 4.21, 4.13	50.0, 65.3, 174.9
22	65.3, CH2	Ha: 4.21, m; Hb: 4.13, m	2.40,2.19	147.2, 50.0
23	170.9, C			
24	44.9, CH2	Ha: 2.83, d (14.4); Hb: 2.74, d (14.4)	Ha: 2.74; Hb: 2.83	176.4, 170.9, 73.4, 42.8
25	73.4, C			
26	42.8, CH2	Ha: 2.97, d (16.6); Hb: 2.76, d (16.6)	Ha: 2.76; Hb: 2.97	176.4, 173.3, 73.4, 44.2
27	173.3, C			
25'	176.4, C			
8'	NH2			
	112.78, 115.1, 117.4, 119.7 - TFA			



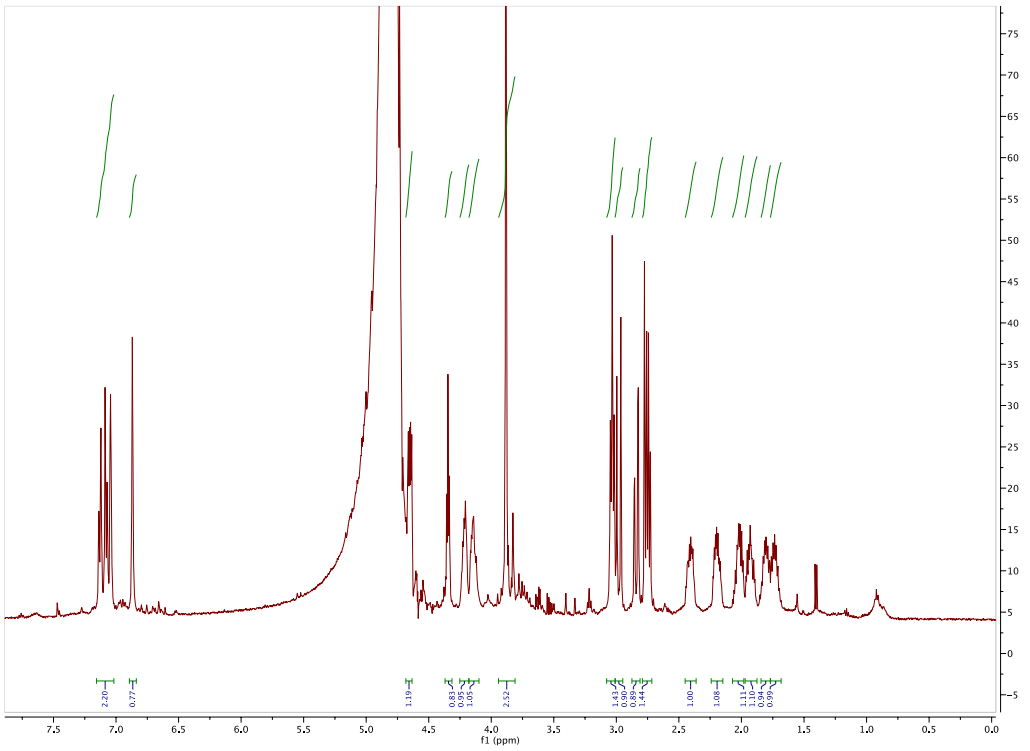


Figure S28. The ^{1}H NMR spectrum of compound **2** in D_2O 0.5% TFA-*d* (500 MHz)

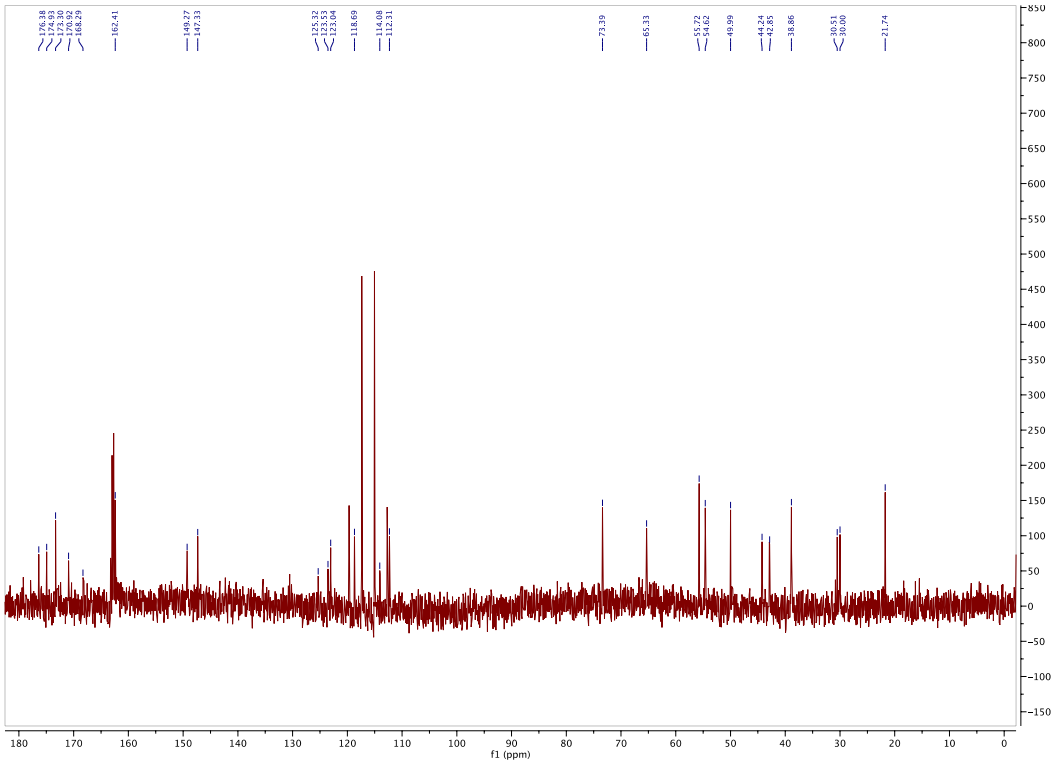


Figure S29. The ^{13}C NMR spectrum of compound **2** in D_2O 0.5% TFA-*d* (500 MHz)

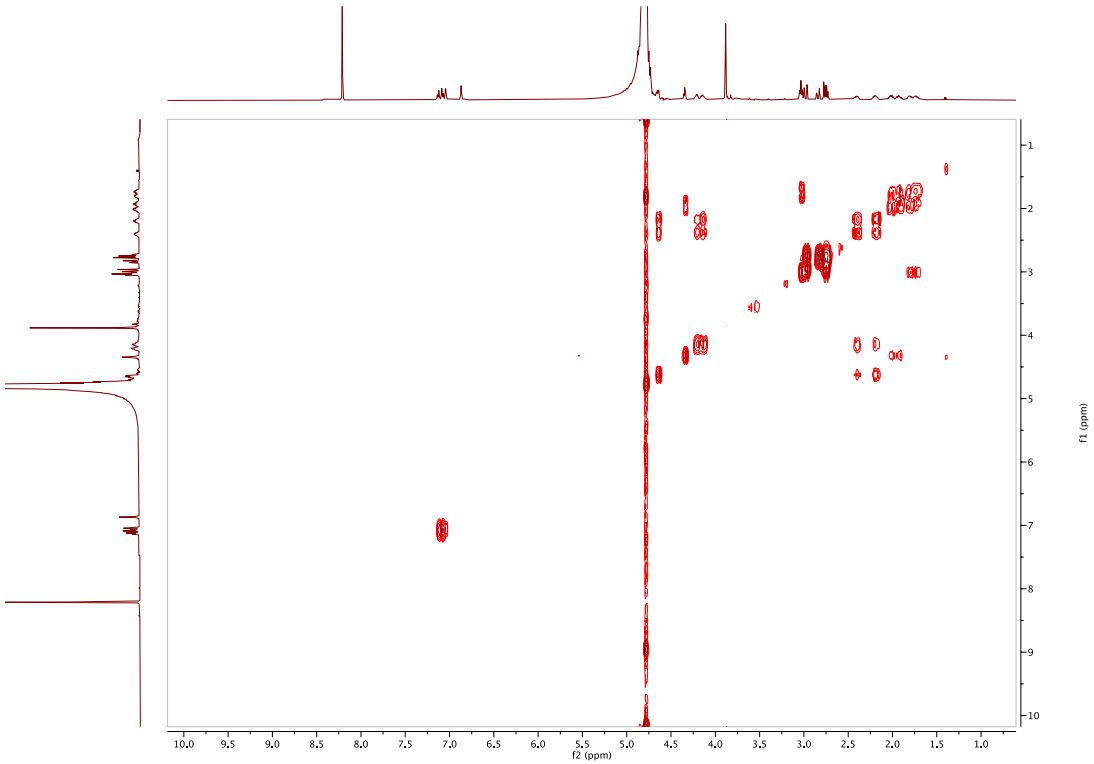


Figure S30. The COSY spectrum of compound **2** in D_2O 0.5% TFA-*d* (500 MHz)

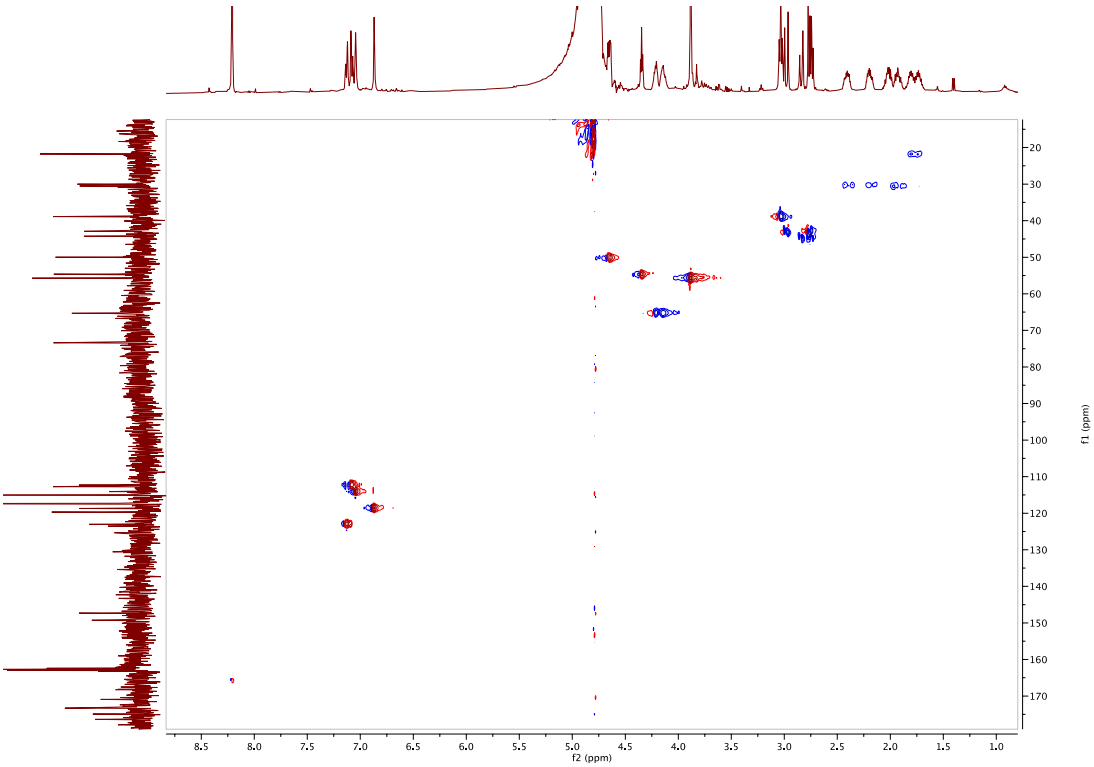


Figure S31. The HSQC spectrum of compound **2** in D_2O 0.5% TFA-*d* (500 MHz)

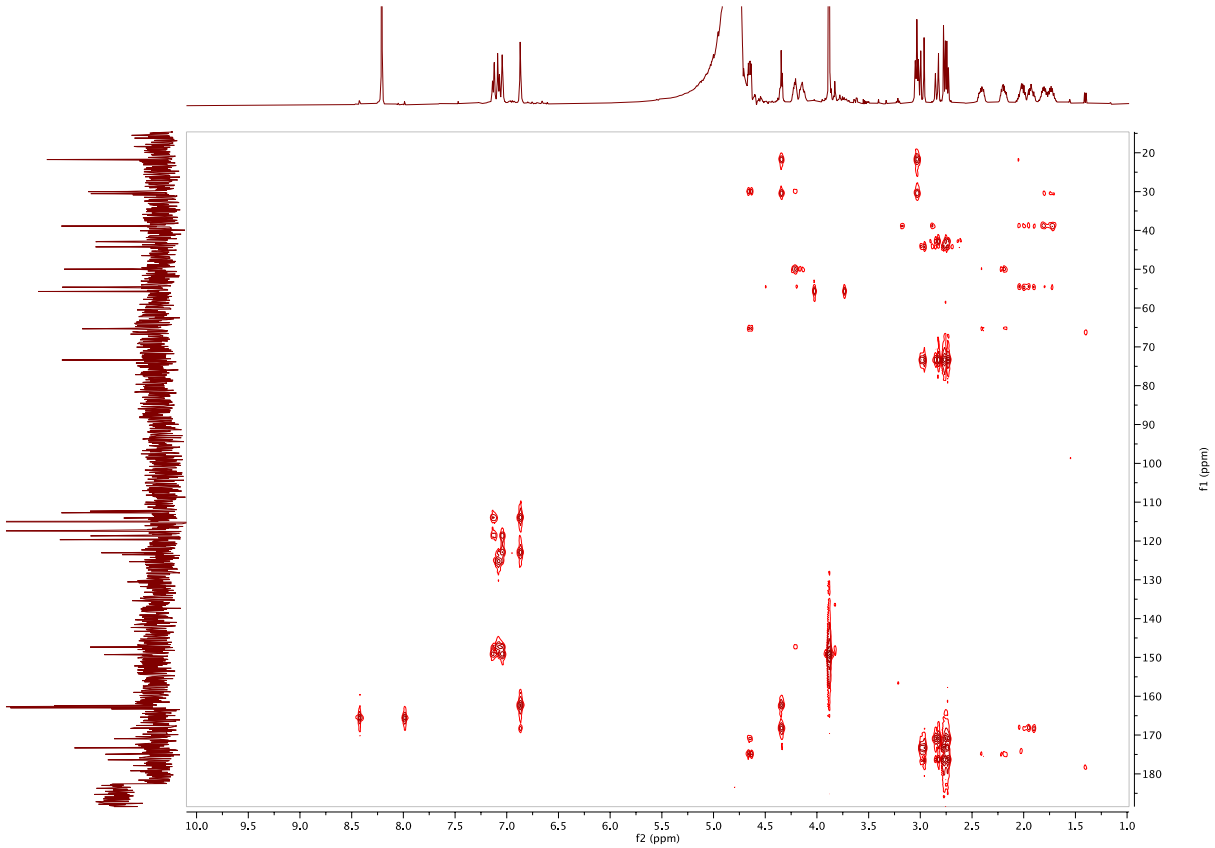
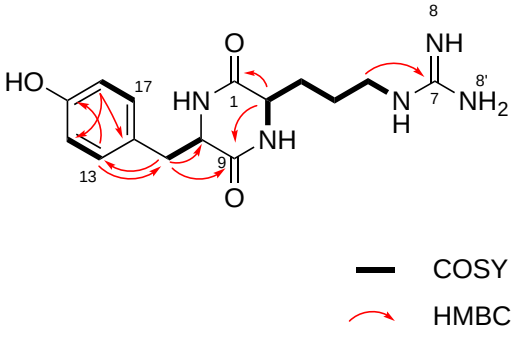


Figure S32. The HMBC spectrum of compound **2** in D_2O 0.5% TFA-*d* (500 MHz)

Table S9: Spectroscopic data of **3**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	168.3, C			
2	53.8, CH	3.74, t (5.1)	0.70, 1.12	23.7, 30.8, 167.7, 168.3
3	30.8, CH2	Ha: 1.10 m; Hb: 0.70, m	1.12, 1.20, 3.74	23.7, 40.4, 53.8, 168.3
4	23.7, CH2	Ha: 1.20, m; Hb: 1.12, m	0.70, 1.10, 2.93	30.8, 40.4, 53.8
5	40.4, CH2	2.93, m	1.12, 1.20	23.7, 30.8, 157.1
6	NH			
7	157.1, C			
8	NH			
9	167.7, C			
10	56.1, CH	4.26, m	2.83, 3.23	37.7, 126.2, 167.7, 168.3
11	37.7, CH2	Ha: 3.23, dd (3.5, 14.0); Hb: 2.83, dd (4.6, 14.0).	4.26	56.1, 126.2, 131.5, 167.7
12	126.2, C			
13, 17	131.5, CH	7.05, d (8.6)	6.73	37.7, 56.1, 115.0, 131.5, 156.4
14, 16	115.0, CH	6.73, d (8.6)	7.05	115.0, 126.2, 131.5, 156.4
15	156.4, C			



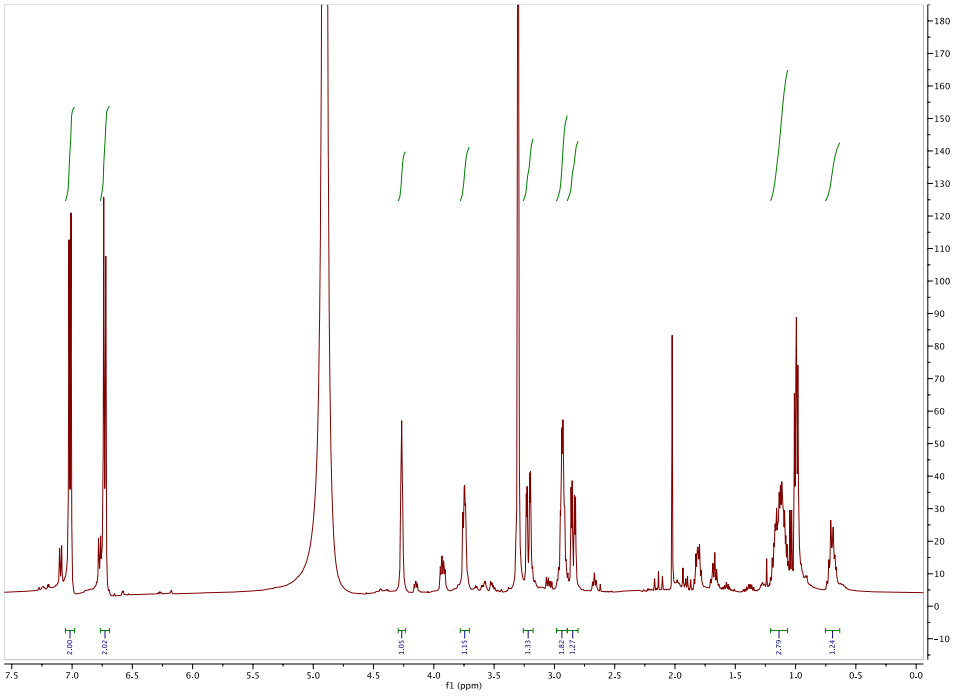


Figure S33. The ^{1}H NMR spectrum of compound **3** in MeOH-*d4* (500 MHz)

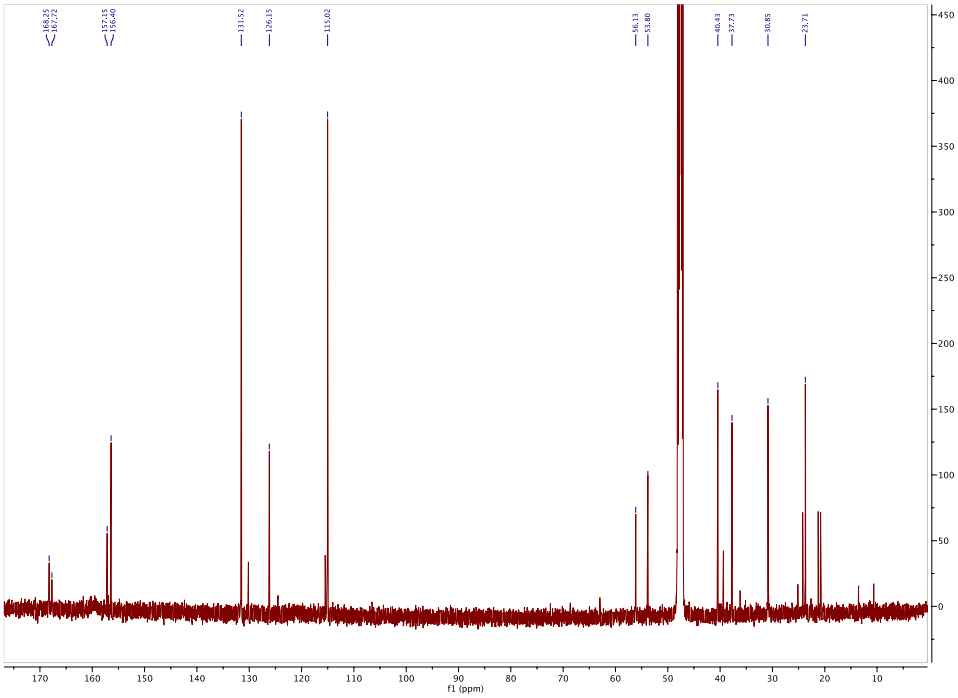


Figure S34. The ^{13}C NMR spectrum of compound **3** in MeOH-*d4* (500 MHz)

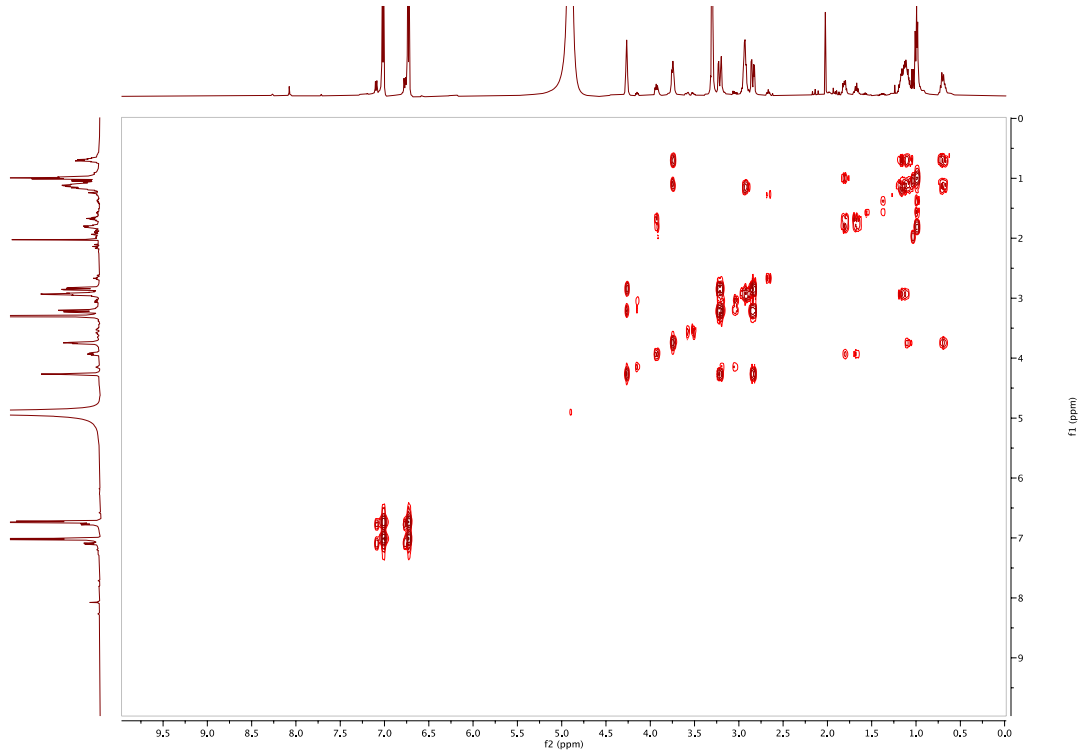


Figure S35. The ^{1}H-^{1}H COSY spectrum of compound **3** in MeOH-*d4* (500 MHz)

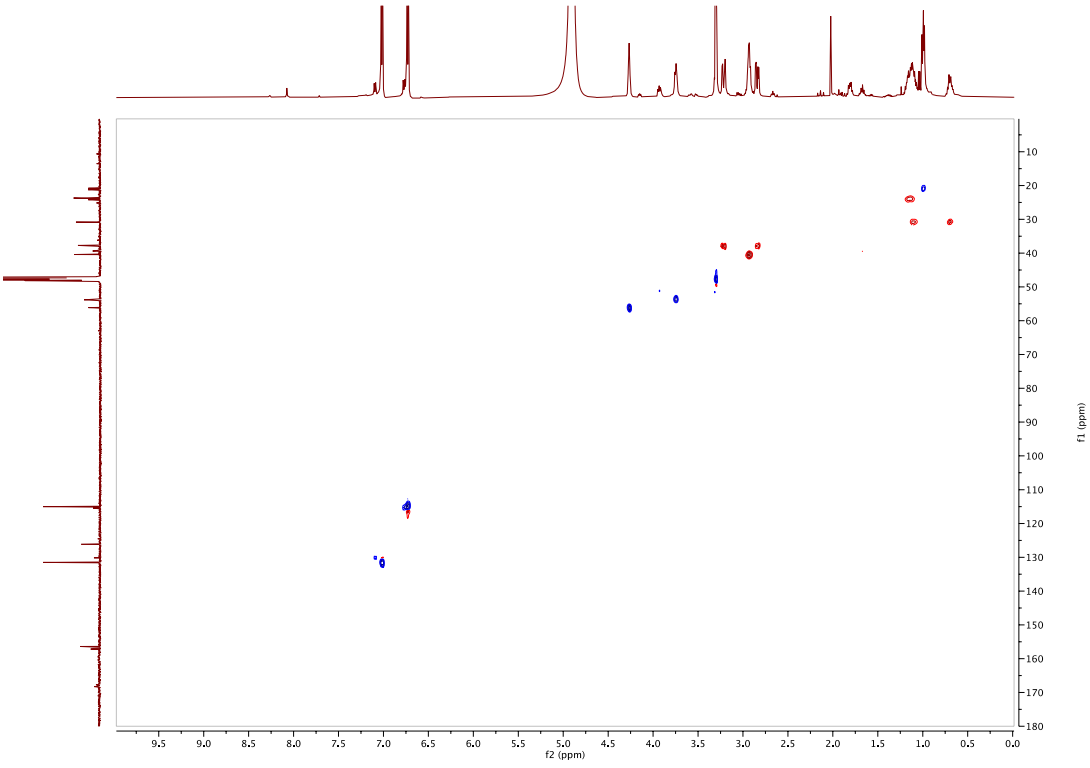


Figure S36. The HSQC spectrum of compound **3** in MeOH-*d4* (500 MHz)

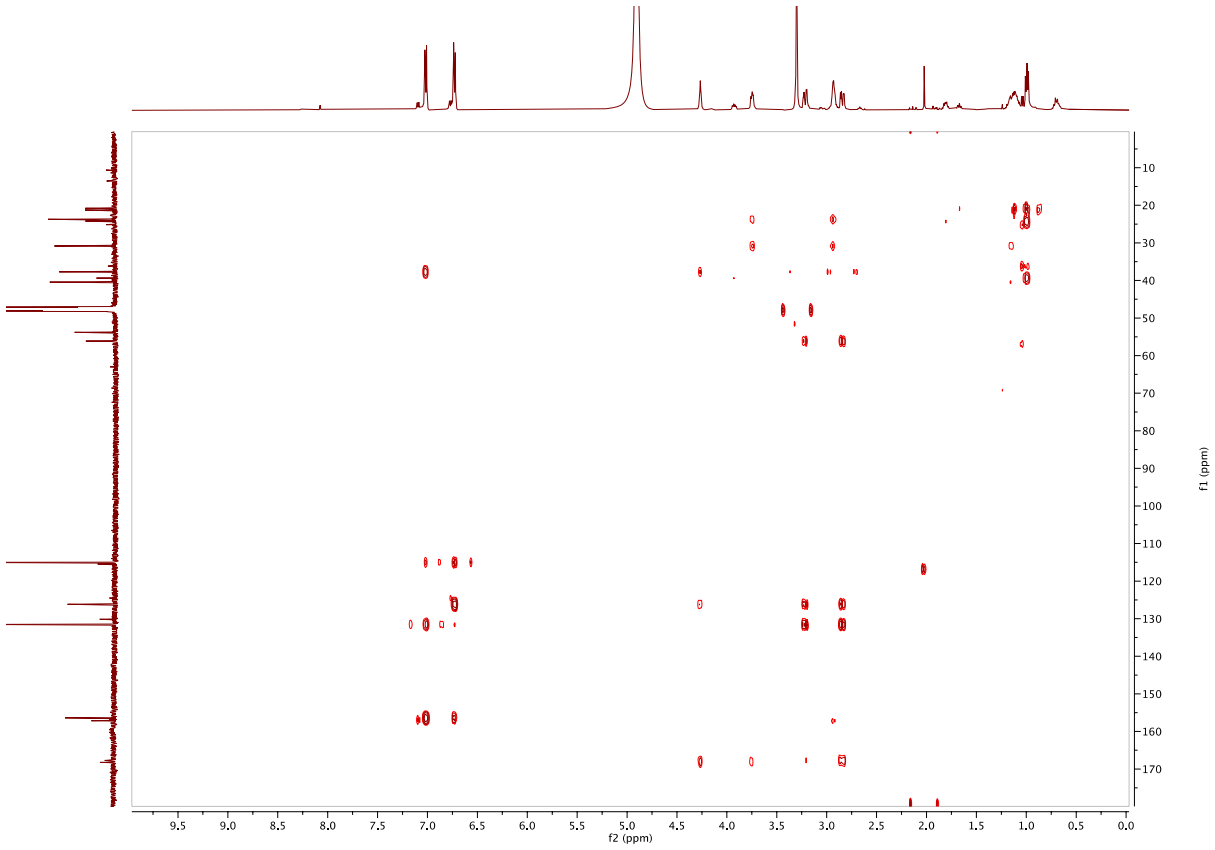
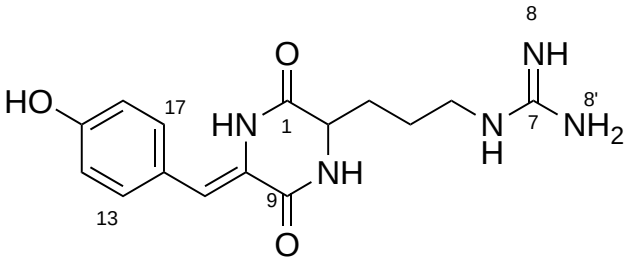


Figure S37. The HMBC spectrum of compound **3** in MeOH-*d4* (500 MHz)

Table S10: Spectroscopic data of **4**

Position	**δC, mult**	**δH, mult (J in Hz)**	**Reported δC, mult**[147]	**Reported δH, mult (J in Hz)**[147]
1	167.1, C		166.7, s	
2	54.5, CH	3.97	54.5, d	3.99, m
3	31.1, CH2	1.71	31.1, t	1.72, m
4	23.9, CH2	1.51	23.9, t	1.53, m
5	Under solvent peak	3.04	40.3, t	3.10, m
6	NH			
7	157.4, C		156.6, s	
8	NH			
9	160.9, C		160.8, s	
10	124.0, C		124.0, s	
11	115.2, CH	6.58	115.1, d	6.61, s
12	124.2, C		124.2, s	
13, 17	131.0, CH	7.32	130.9, d	7.34, d
14, 16	115.7, CH	6.77	115.5, d	6.78, d
15	157.8, C		157.5, s	



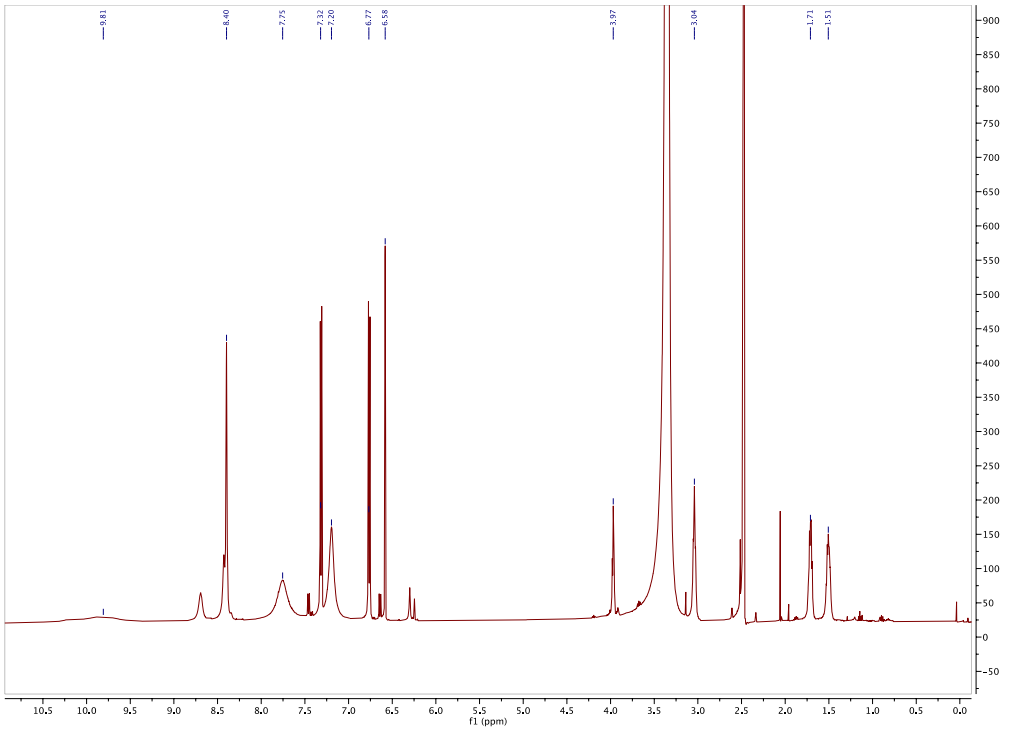


Figure S38. The ^{1}H NMR spectrum of compound **4** in DMSO-*d6* (500 MHz)

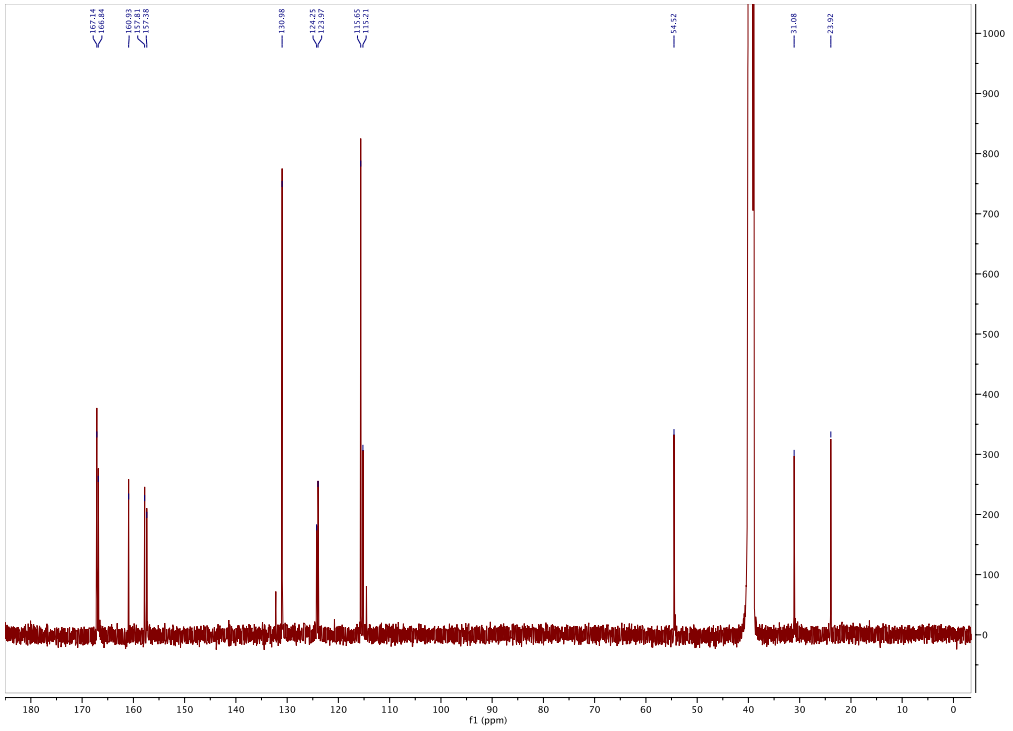
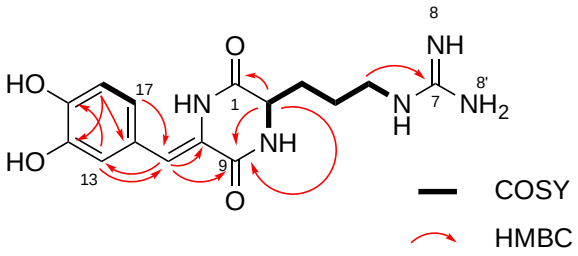


Figure S39. The ^{13}C NMR spectrum of compound **4** in DMSO-*d6* (500 MHz)

Table S11: Spectroscopic data of **5**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	169.0, C=O			
2	55.2, CH	4.3, t (5.0)	1.98, 1.88	23.1, 31.1, 162.9, 169.0
3	31.1, CH2	Ha: 1.98; Hb: 1.88	1.63, 1.71, 4.3	55.2
4	23.1, CH2	Ha: 1.71; Hb: 1.63	1.63, 1.71, 3.2	40.7
5	40.7, CH2	3.2, t (6.9)	1.63, 1.71	23.1, 31.1, 157.0
6	NH			
7	157.0, C			
8	NH			
8'	NH2			
9	162.9, C=O			
10	123.5, C			
11	119.2, CH	6.81		116.9, 122.7, 123.5, 162.9
12	125.3, C			
13	116.9, CH	7.00		119.2, 122.7, 144.5, 145.4
14	144.5, C			
15	145.4, C			
16	116.8, CH	6.95		125.3, 144.5
17	122.7, CH	6.96		116.9, 119.2, 125.3, 145.4



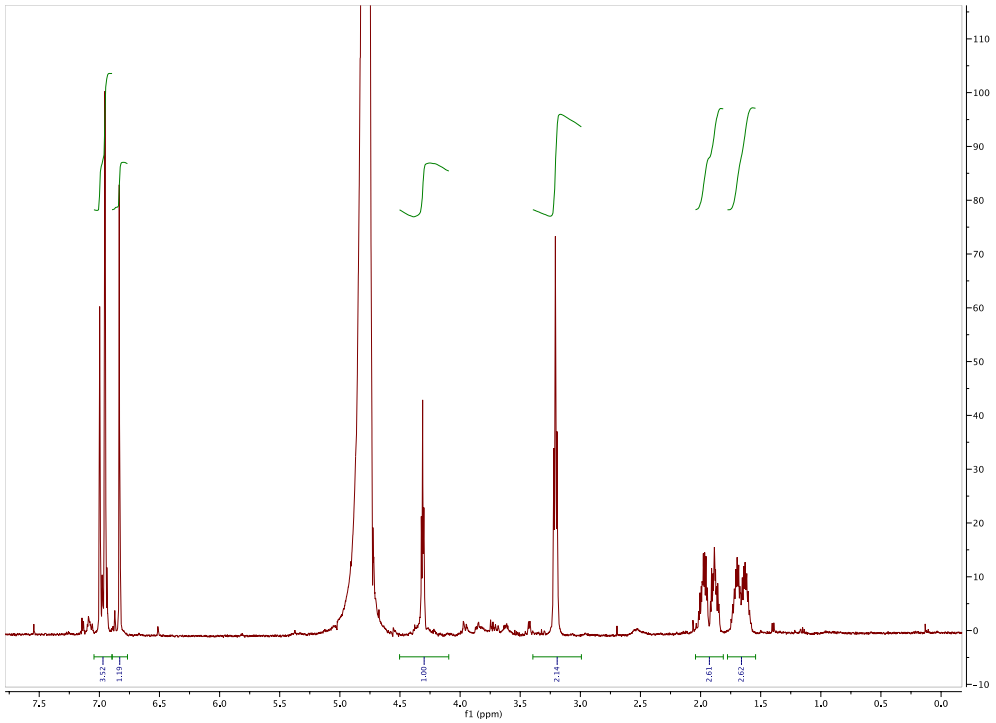


Figure S40. The ^{1}H NMR spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz)

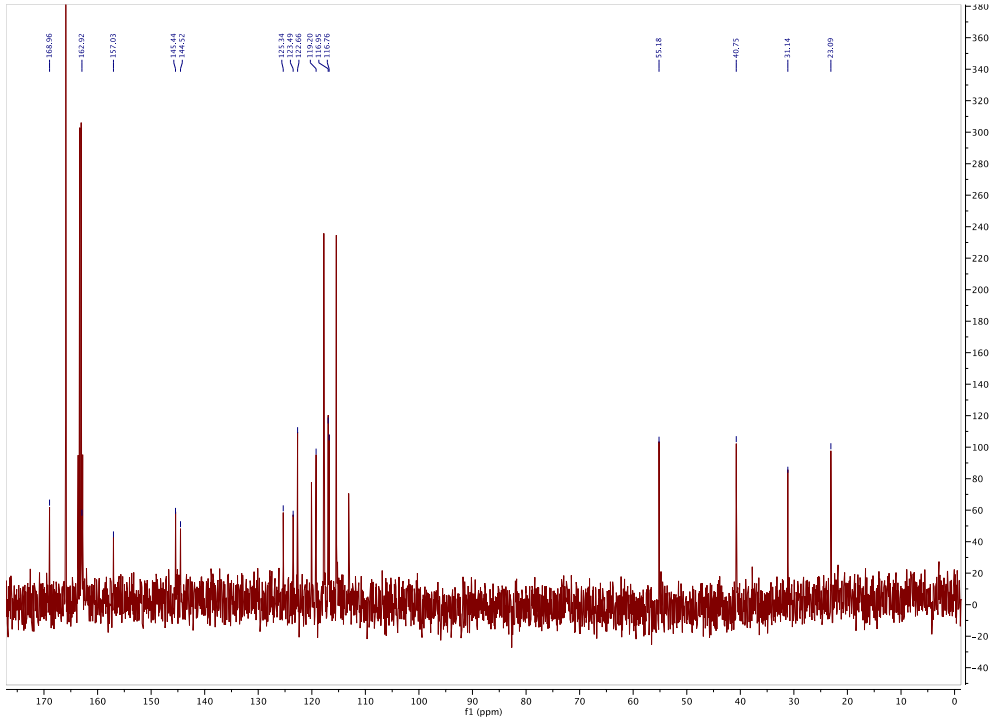


Figure S41. The ^{13}C NMR spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz)

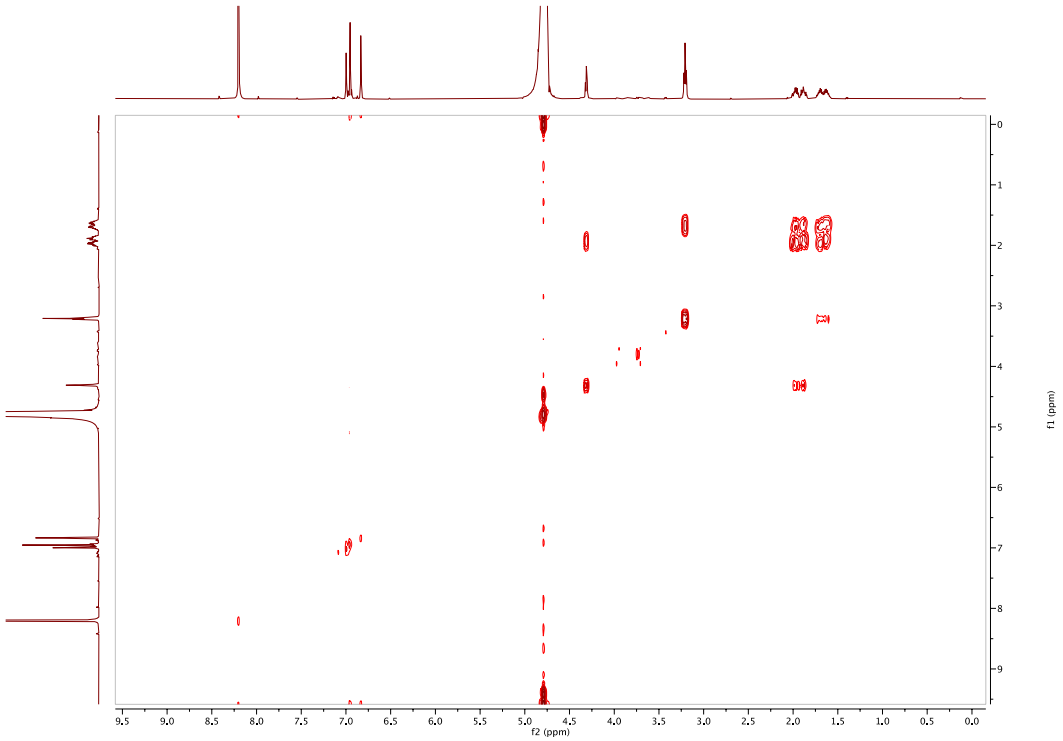


Figure S42. The ^{1}H-^{1}H COSY spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz)

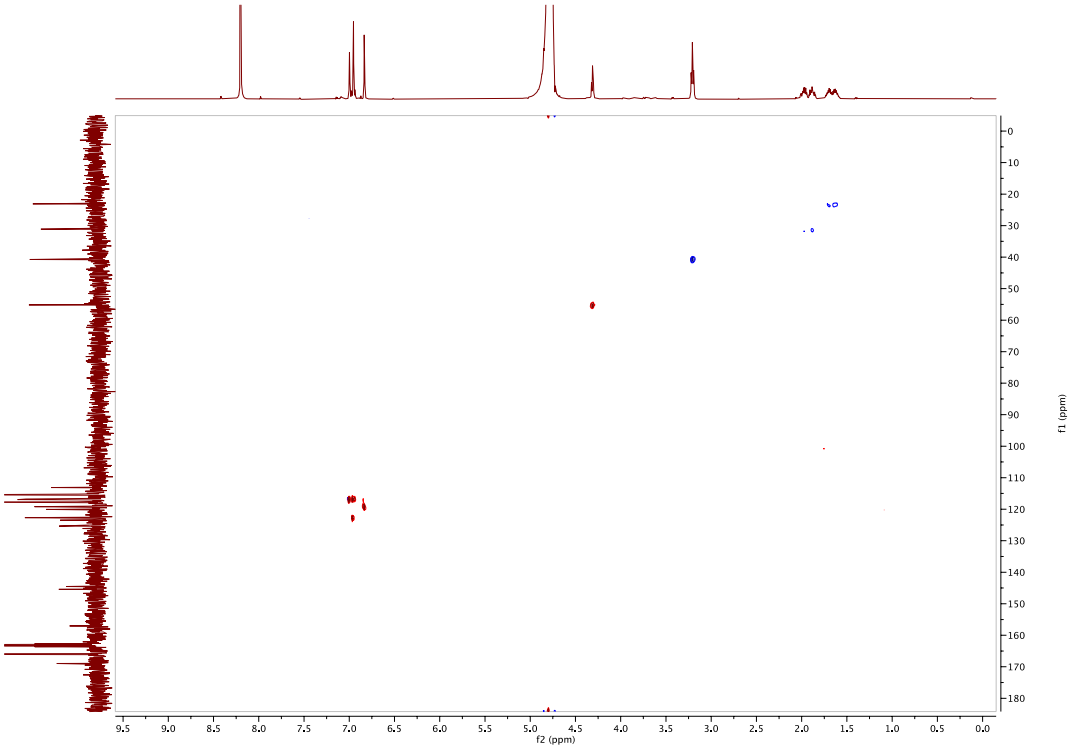


Figure S43. The HSQC spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz)

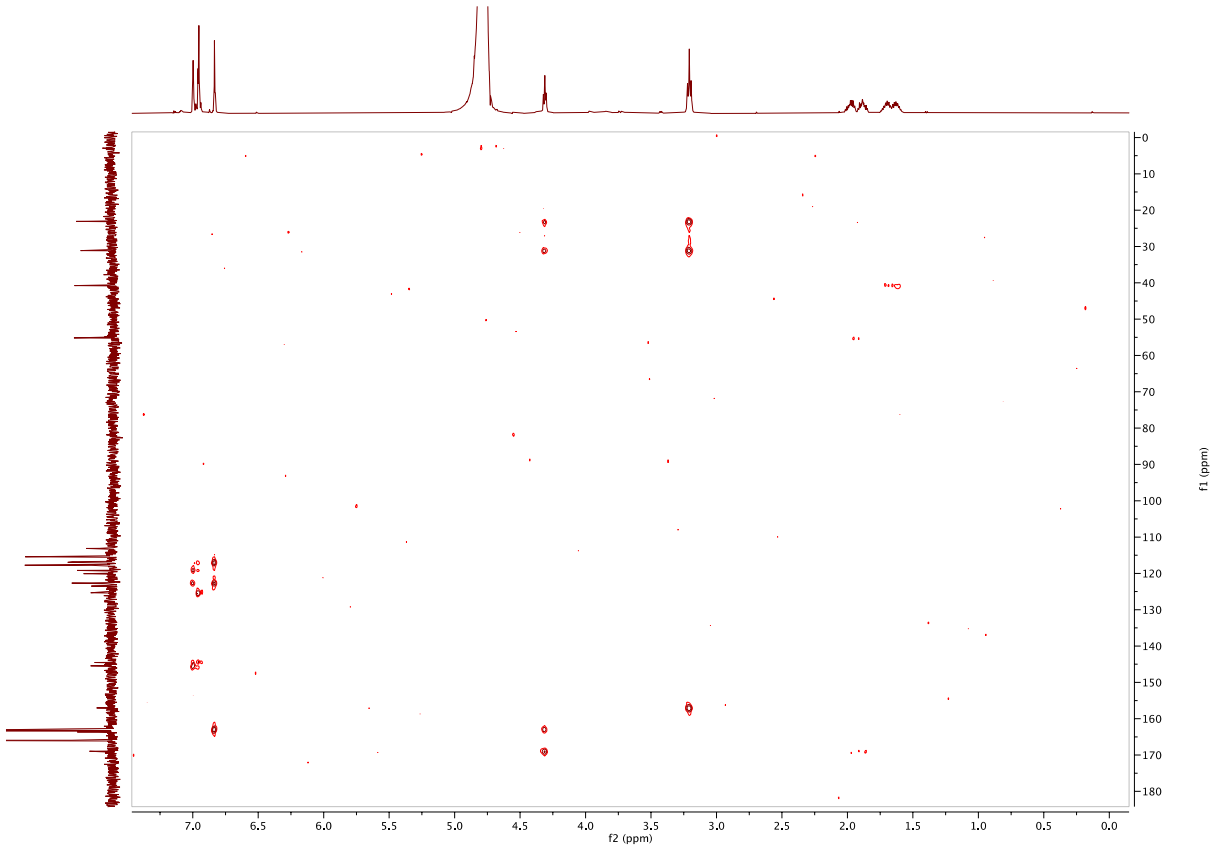
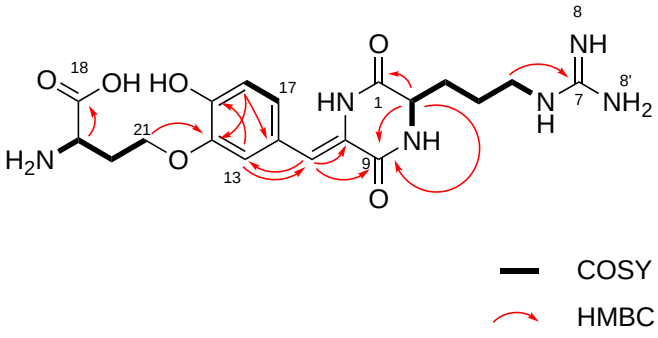


Figure S44. The HMBC spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz)

Table S12: Spectroscopic data of **6**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	168.6, C			
2	54.9, CH	4.32, t	1.97, 1.89	168.6, 162.4, 30.8, 22.9
3	30.8, CH2	Ha: 1.97, m; Hb: 1.89, m	4.32, 1.70, 1.62	54.8
4	22.9, CH2	Ha: 1.70, m; Hb: 1.62, m	3.23, 1.97, 1.89	40.4
5	40.5, CH2	3.23, t	1.70, 1.62	156.7, 30.8, 22.9
6	NH			
7	156.7, C			
8	NH			
9	162.4, C			
10	123.4[a], C			
11	118.7, CH	6.85, s		162.4, 123.3-123.4[a], 114.8
12	125.0, C			
13	114.8, CH	7.06, s		146.1-146.4[b], 123.3-123.4[a], 118.7
14	146.1, C			
15	146.4, C			
16	116.1, CH	6.98, d	7.07	146.1-146.4[b], 125.0
17	123.3[a], CH	7.07, d	6.98	146.1-146.4[b], 118.7
18	172.3, C			
19	51.8, CH	4.27, dd	2.49, 2.42	172.3, 65.8, 29.4
20	29.4, CH2	Ha: 2.49, m; Hb: 2.42, m	4.30, 4.27	
21	65.8, CH2	4.30	2.49, 2.42	146.1

[a,b] Carbon data for positions 10/17 and 14/15 may be interchangeable.



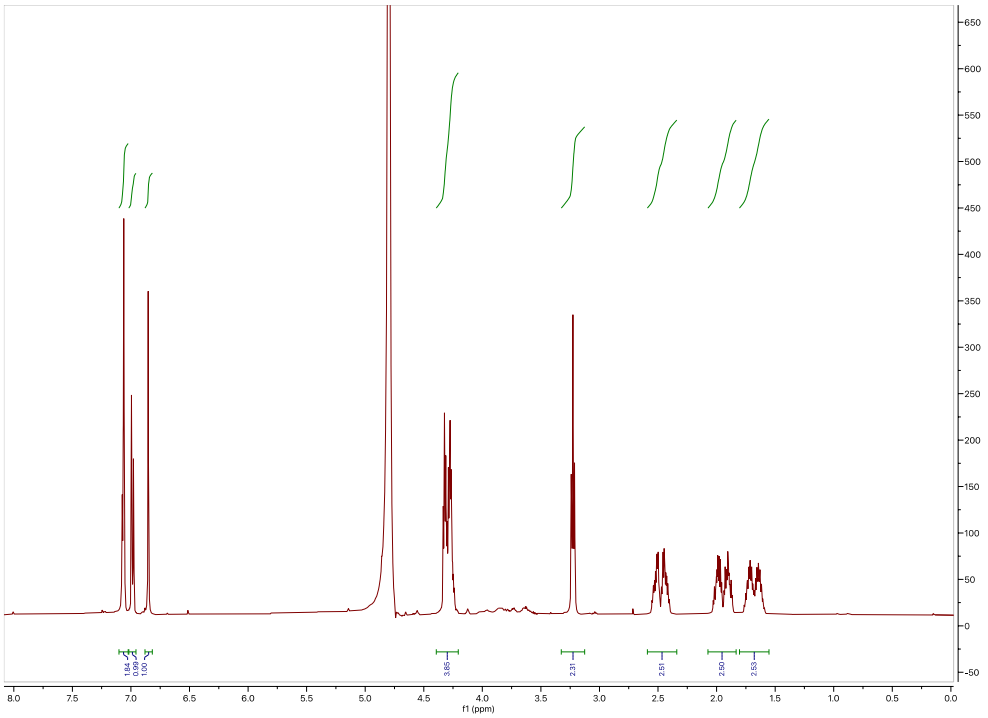


Figure S45. The ^{1}H NMR spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz)

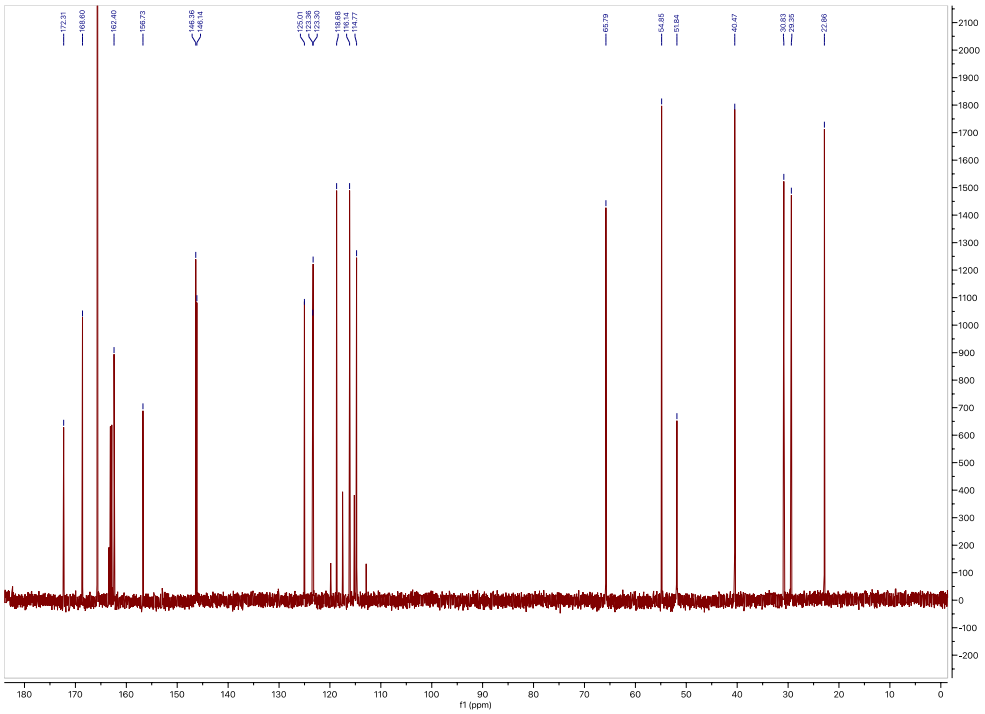


Figure S46. The ^{13}C NMR spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz)

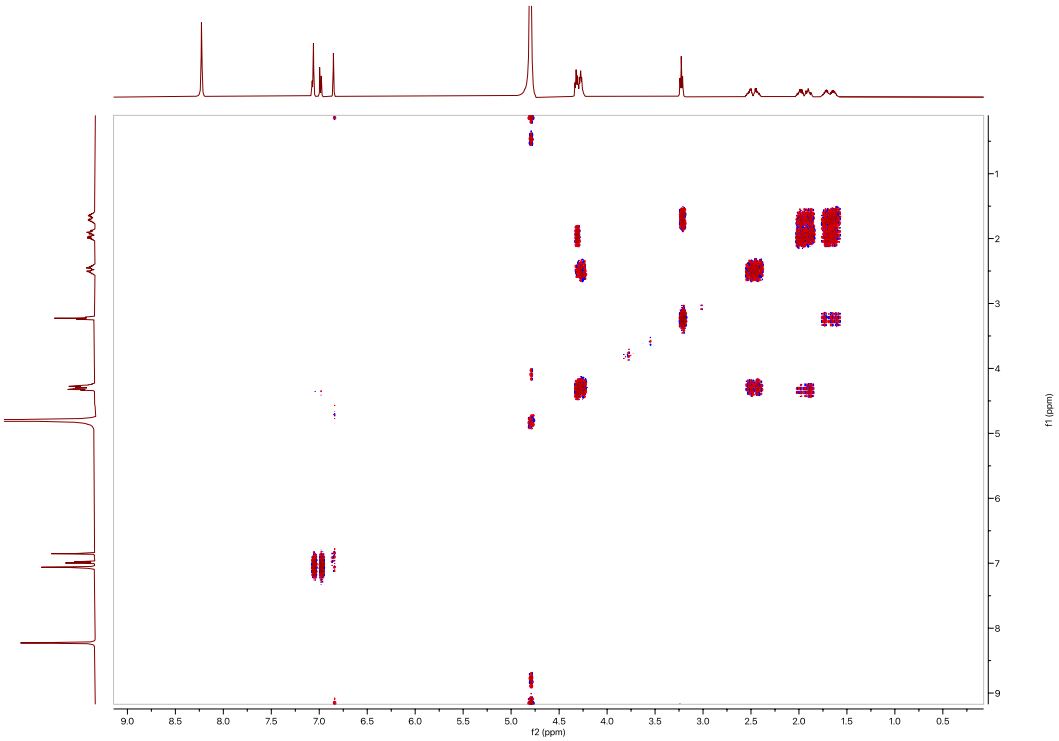


Figure S47. The COSY spectrum of compound 6 in D_2O 0.5% TFA-*d* (500 MHz)

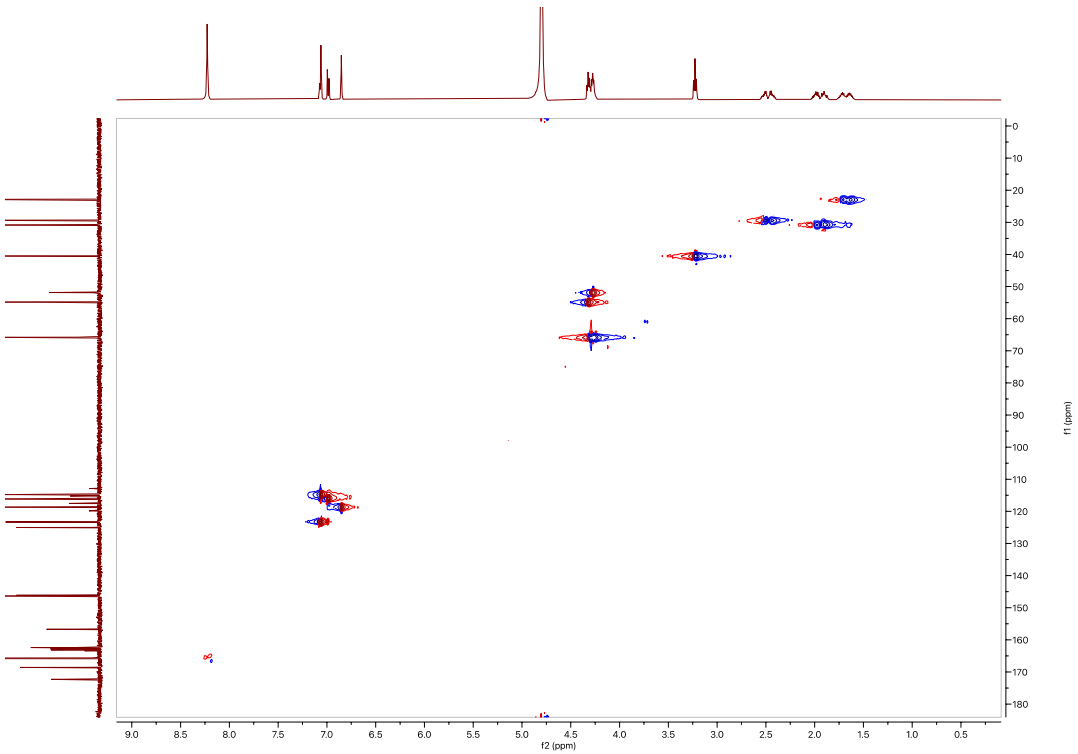


Figure S48. The HSQC spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz)

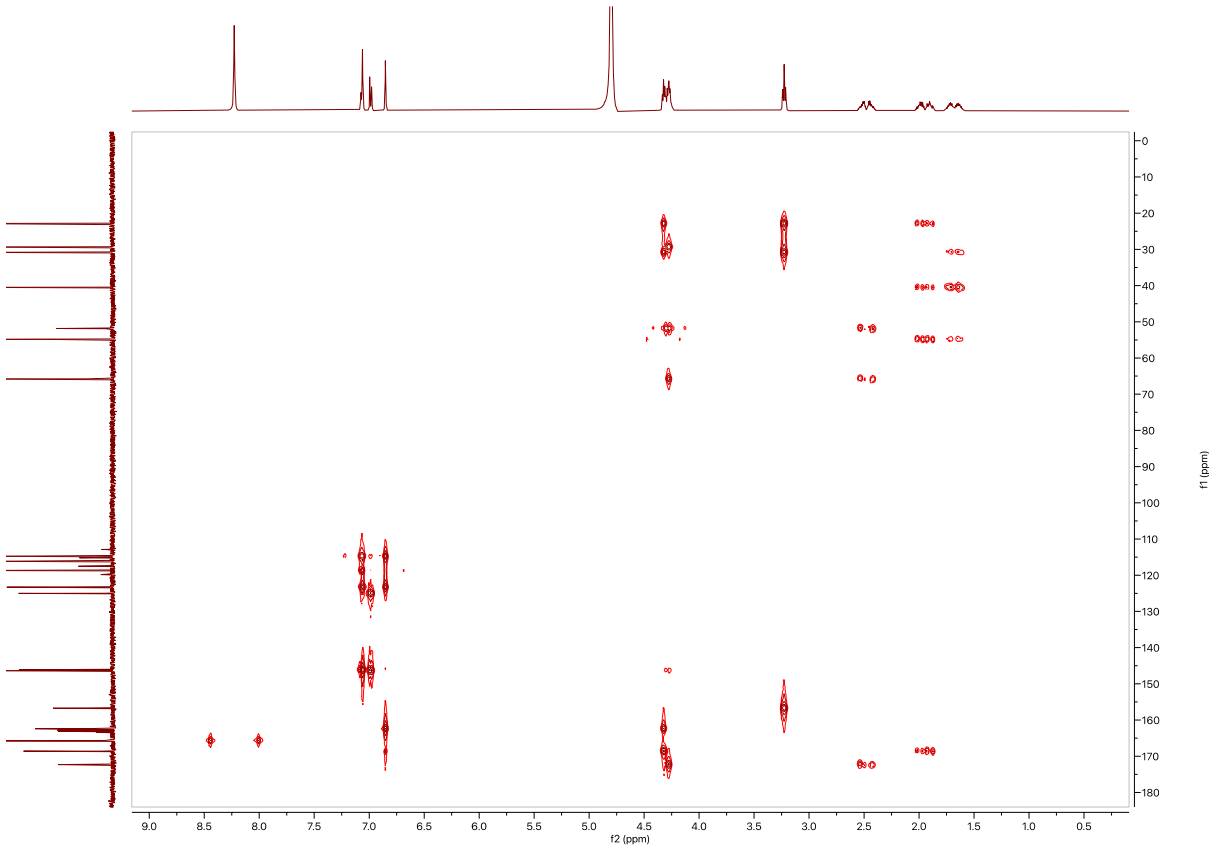
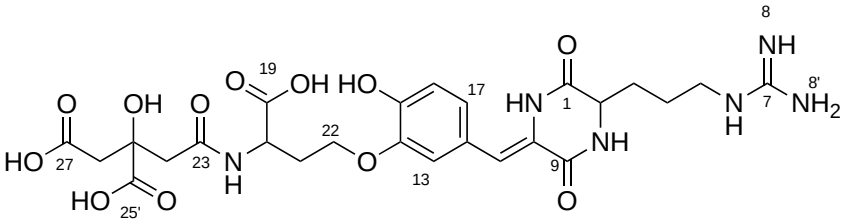


Figure S49. The HMBC spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz)

Table S13: Spectroscopic data of **7**

Position	**δC, mult (D_2O)**	**δH, mult (J in Hz) (D_2O)**	**Reported δC, mult (DMSO)**[144]	**Reported δH, mult (J in Hz) (DMSO)**[144]
1	168.5, C		166.9	
2	54.8, CH	4.31, t	54.7	4.01
3	30.7, CH2	Ha: 1.98, m; Hb: 1.89, m	31.3	1.72
4	22.8, CH2	Ha: 1.71, m ; Hb: 1.63, m	24.2	1.53
5	40.4, CH2	3.22, t	40.5	3.10
6	NH			
7	156.6, C		156.8	
8	NH			
9	162.4, C		160.9	
10	124.8, C		124.6	
11	118.8, CH	6.85, s	115.5	6.62
12	124.8, C		124.7	
13	114.3, CH	7.04, s	115.7	7.01
14	146.1, C		146.6	
15	146.4, C		147.6	
16	116.0, CH	6.97, d	116.0	6.81
17	123.1, CH	7.02, d	123.1	6.99
19	175.0, C		173.4	
20	49.9, CH	4.69, dd	49.5	
21	30.1, CH2	Ha: 2.44, m; Hb: 2.20, m	30.9	2.00, 2.19
22	65.2, CH2	Ha: 4.22, m; Hb: 4.16, m	65.5	4.05, 4.00
23	171.0, C		169.7	
24	44.3, CH2	Ha: 2.85, d; Hb: 2.75, d	43.3	2.55-2.67
25	73.4, C		73.0	
26	42.9, CH2	Ha: 3.00, d; Hb: 2.78, d	43.1	2.55-2.67
27	173.3, C		171.5	
25'	176.4, C		174.8	
8'	NH2			
	112.78, 115.1, 117.4, 119.7 - TFA			



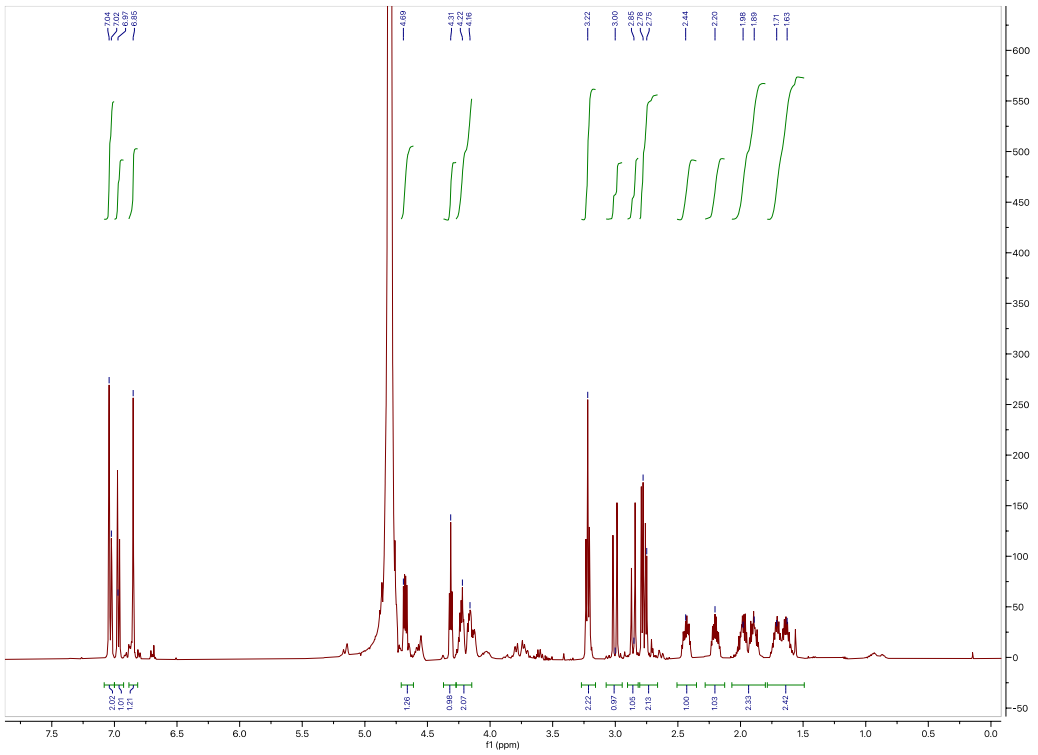


Figure S50. The ^{1}H NMR spectrum of compound **7** in D_2O 0.5% TFA-*d* (500 MHz)

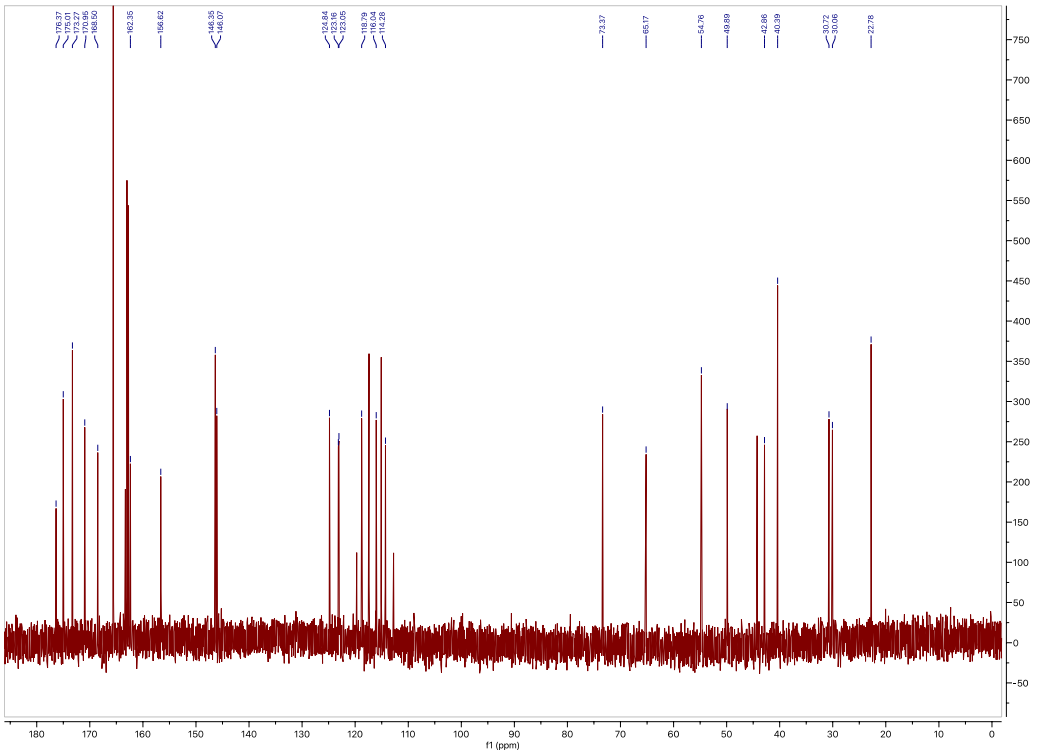
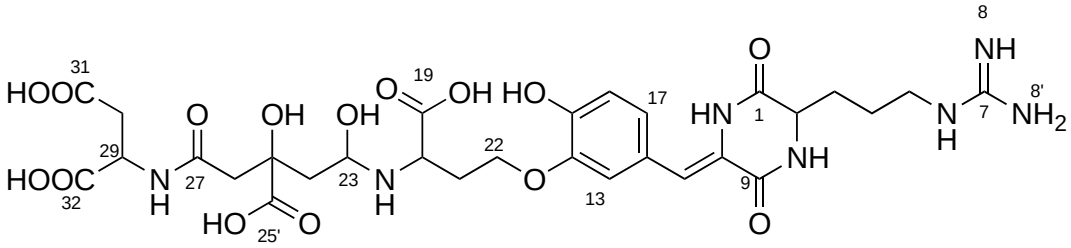


Figure S51. The ^{13}C NMR spectrum of compound **7** in D_2O 0.5% TFA-*d* (500 MHz)

Table S14: Spectroscopic data of **9**

Position	δC, mult (D_2O)	δH, mult (J in Hz, D_2O)	Reported δC, mult (DMSO)[144]	Reported δH, mult (J in Hz, DMSO)[144]
1	168.9, C		166.9	
2	55.2, CH	4.33, t	54.7	4.01
3	31.1, CH2	Ha: 1.99, m; Hb: 1.90, m	31.3	1.72
4	23.2, CH2	Ha: 1.72, m; Hb: 1.64, m	24.2	1.53
5	40.8, CH2	3.22, t	40.5	3.11
6	NH			
7	157.0, C		156.8	
8	NH			
9	missing[a]		160.9	
10	123.6, C		124.6	
11	119.2, CH	6.88, s	115.5	6.62
12	125.3, C		124.7	
13	114.8, CH	7.08, s	115.7	7.01
14	146.5, C		146.6	
15	146.8, C		147.6	
16	116.4, CH	6.98, d	116.0	6.81
17	123.4, CH	7.05, d	123.1	7.01
19	175.4, C		173.4	
20	50.3, CH	4.68, dd	49.5	4.41
21	30.4, CH2	Ha: 2.43, m; Hb: 2.21, m	30.9	2.00, 2.18
22	65.6, CH2	Ha: 4.24, m; Hb: 4.17, m	65.5	3.94-4.07
23	171.5, C		169.7	
24	44.3, CH2	2.70-2.90	43.2	2.55-2.67
25	74.1, C		73.5	
26	44.0, CH2	2.70-2.90	43.2	2.55-2.67
27	171.2, C		169.5	
25'	176.6, C		174.9	
28	NH			8.26
29	49.2, CH	4.71, m	48.6	4.51
30	35.8, CH2	2.70-2.90	36.3	2.55-2.67
31	174.5, C		171.8	
32	174.1, C		172.4	
8'	NH_2			
	112.78, 115.1, 117.4, 119.7 - TFA			

[a]The carbon peak for C9 should be ~163 ppm and may be under one of the four TFA peaks nearby.



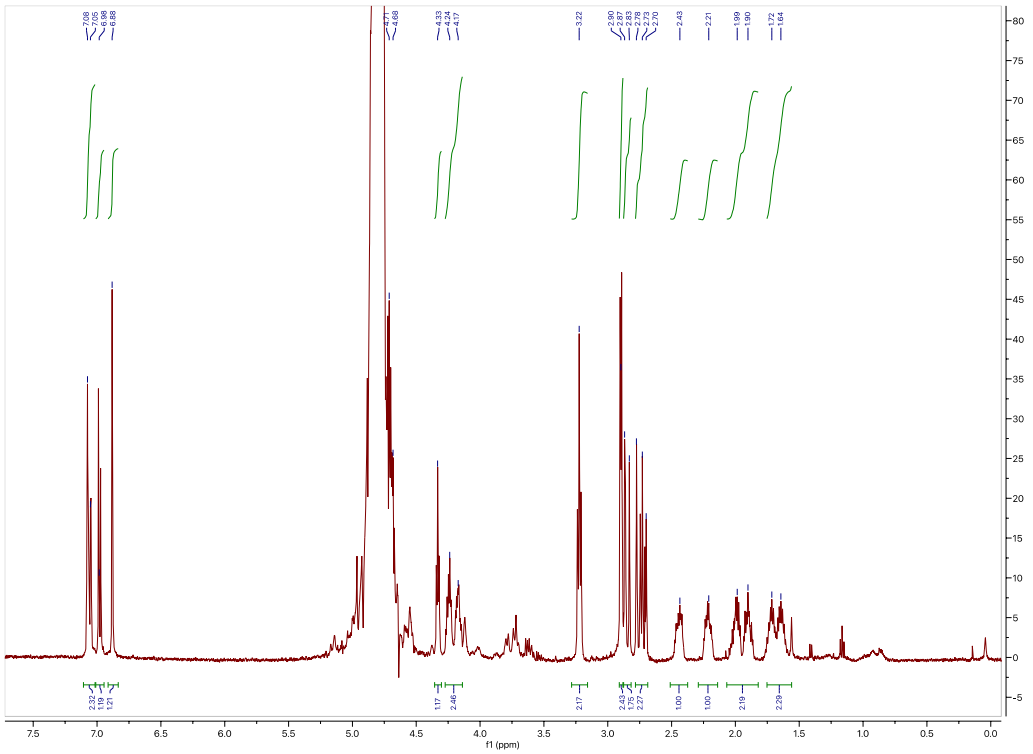


Figure S52. The ^{1}H NMR spectrum of compound **9** in D_2O 0.5% TFA-*d* (500 MHz)

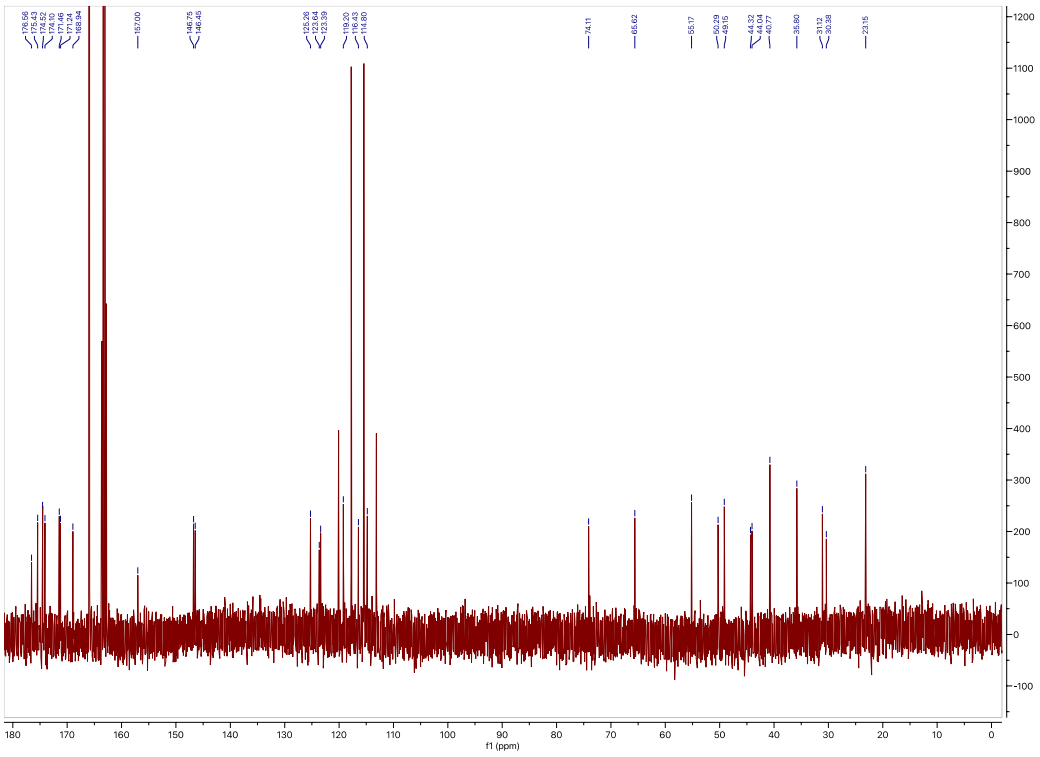
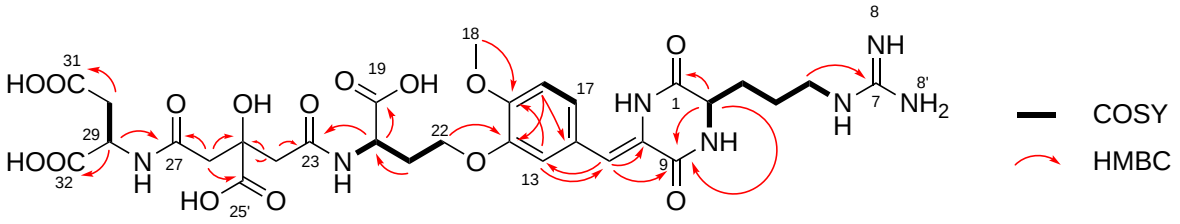


Figure S53. The ^{13}C NMR spectrum of compound **9** in D_2O 0.5% TFA-*d* (500 MHz)

Table S15: Spectroscopic data of **10**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	168.9, C			
2	55.2, CH	4.32, t (5.1)	1.98, 1.89	168.9, 162.7, 31.2, 23.2
3	31.2, CH2	Ha: 1.98, m; Hb: 1.89, m	4.32, 1.71, 1.63	55.2
4	23.2, CH2	Ha: 1.71, m; Hb: 1.63, m	3.21, 1.98, 1.89	40.8
5	40.8, CH2	3.21, t (7.0)	1.71, 1.63	157.1, 31.2, 23.2
6	NH			
7	157.1, C			
8	NH			
9	162.7, C			
10	123.9, C			
11	118.9, CH	6.87, s		162.7, 123.9, 114.4
12	125.7, C			
13	114.4, CH	7.05, s		149.6, 123.3, 118.9
14	147.7, C			
15	149.6, C			
16	112.6, CH	7.08, d (8.6)	7.12	147.7, 123.3
17	123.3, CH	7.12, d (8.6)	7.08	149.6, 114.4
18	56.1, CH3	3.88, s		149.6
19	175.3, C			
20	50.3, CH	4.64, dd (4.4, 8.9)	2.40, 2.19	174.5, 171.4, 65.7
21	30.3, CH2	Ha: 2.40, m; Hb: 2.19, m	4.64, 4.21, 4.13	
22	65.7, CH2	Ha: 4.21, m; Hb: 4.13, m	2.40, 2.19	147.7, 50.3
23	171.4, C			
24	44.3, CH2	Ha: 2.83, d (14.4); Hb: 2.74, d (14.4)	Ha: 2.74; Hb: 2.83	176.6, 171.4, 74.1, 44.0
25	74.1, C			
26	44.0, CH2	Ha: 2.86, d (16.6); Hb: 2.76, d (16.6)	Ha: 2.76; Hb: 2.97	176.6, 171.3, 74.1, 44.3
27	171.3, C			
25'	176.6, C			
28	NH			
29	49.2, CH	4.68, m		35.9, 171.3, 174.1
30	35.9, CH2	2.88, m		49.2, 171.3, 174.5
31	174.5, C			
32	174.1, C			
8'	NH2			
	112.78, 115.1, 117.4, 119.7 - TFA			



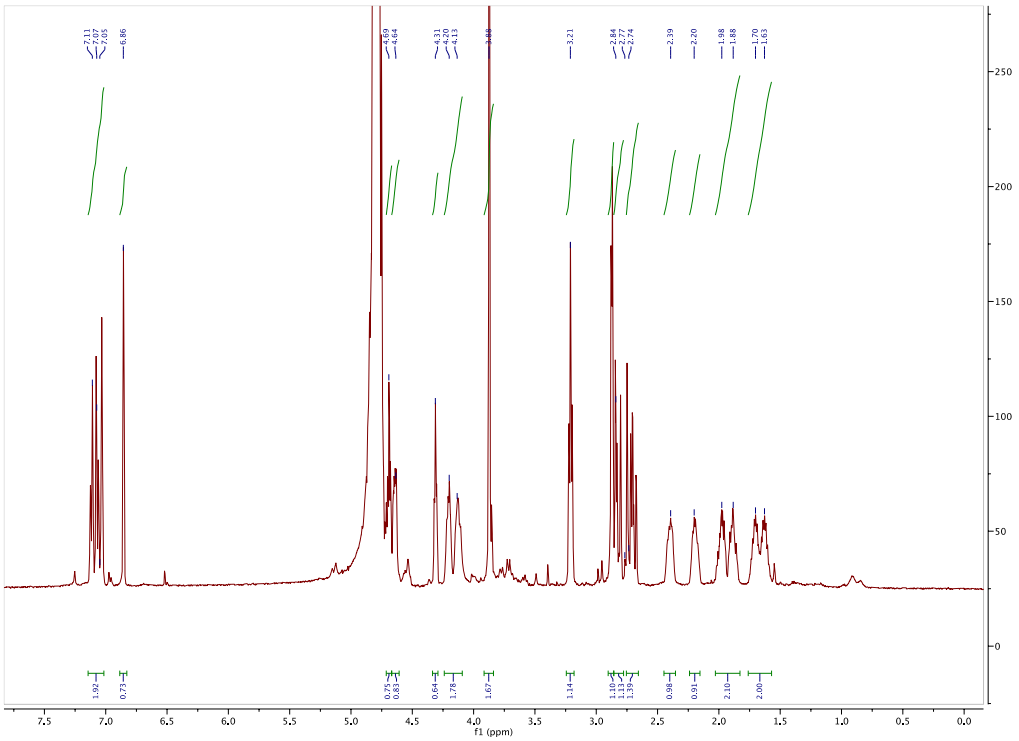


Figure S54. The ^{1}H NMR spectrum of compound **10** in D_2O 0.5% TFA-*d* (500 MHz)

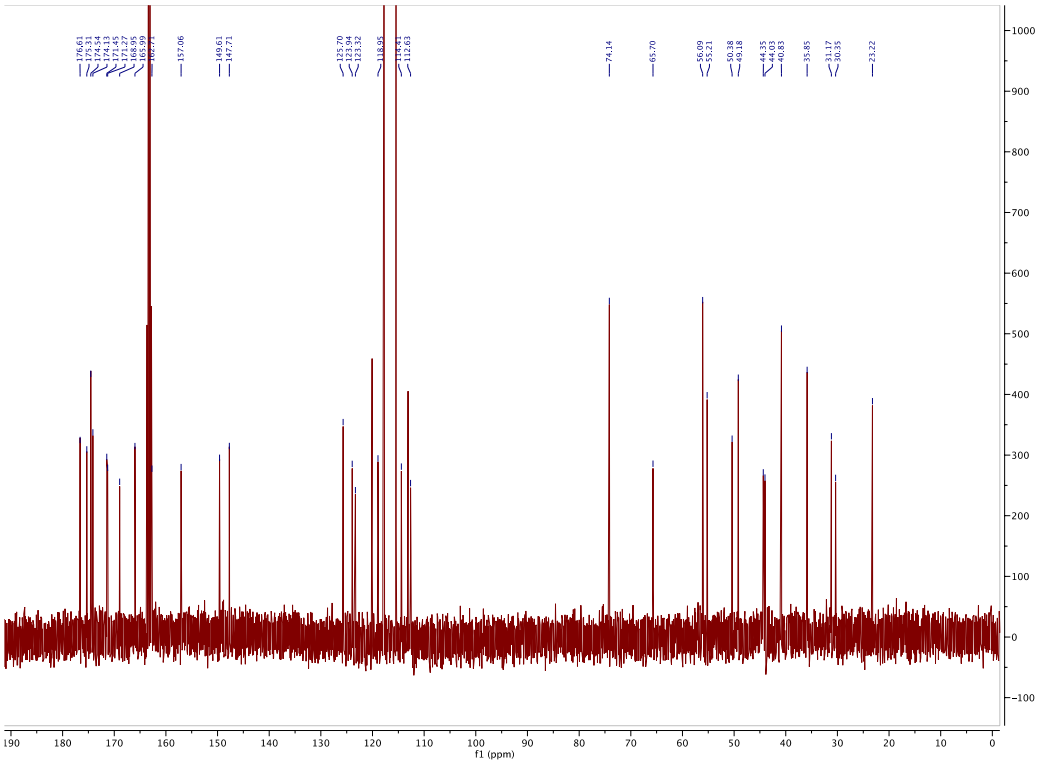


Figure S55. The ^{13}C NMR spectrum of compound **10** in D_2O 0.5% TFA-*d* (500 MHz)

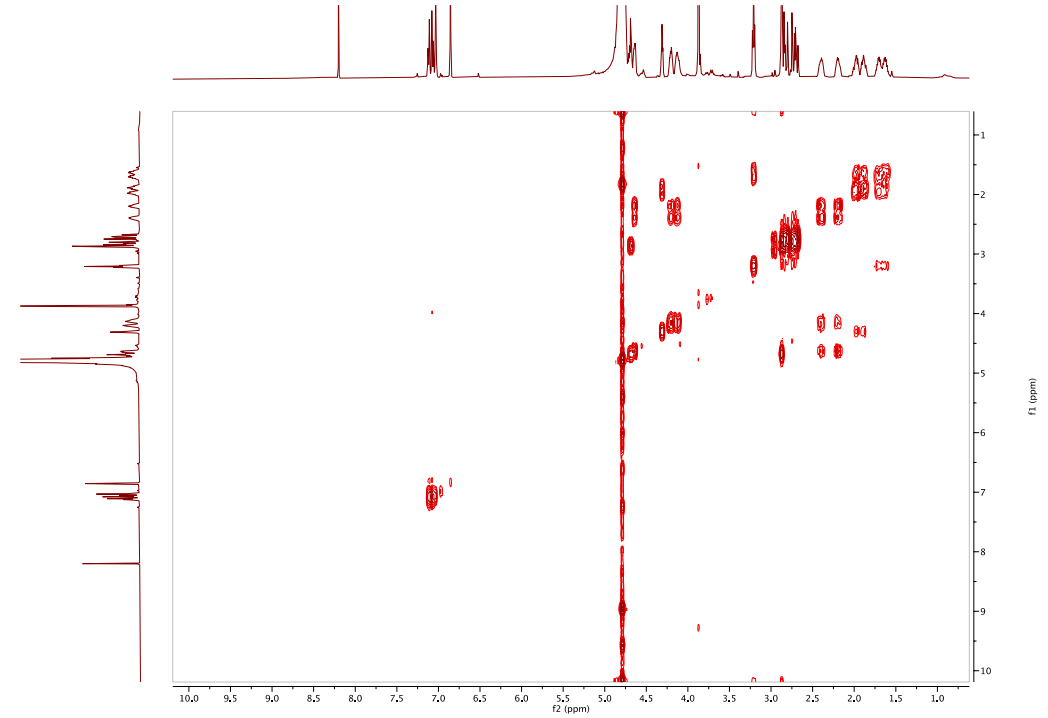


Figure S56. The COSY spectrum of compound **10** in D_2O 0.5% TFA-*d* (500 MHz)

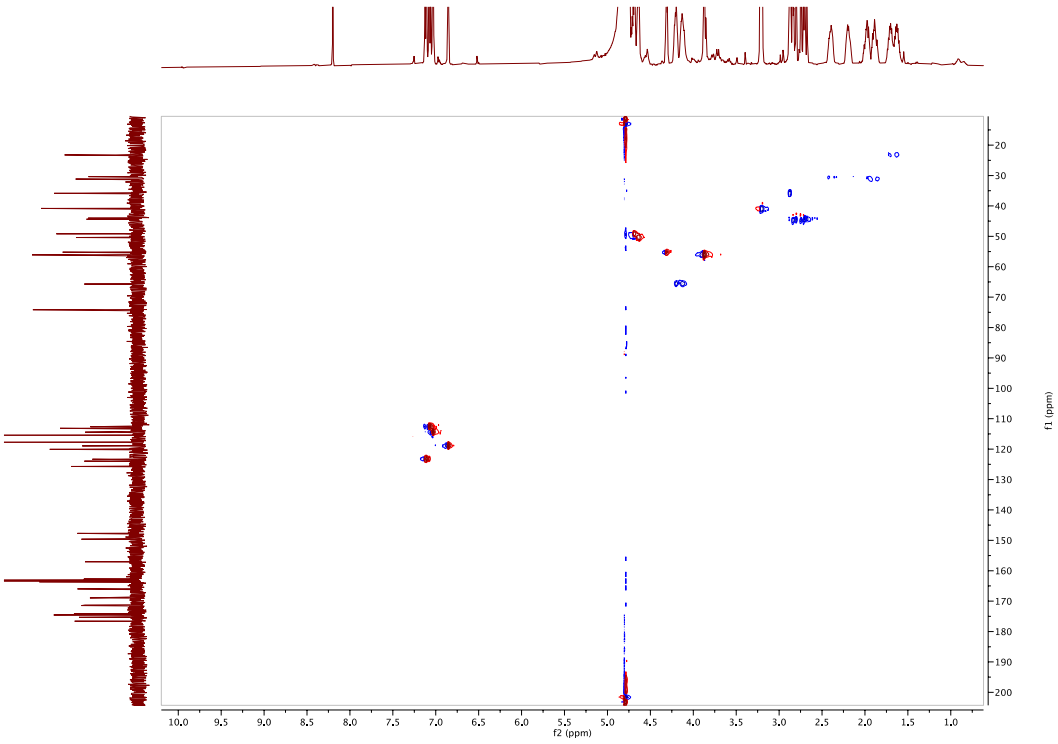


Figure S57. The HSQC spectrum of compound **10** in D_2O 0.5% TFA-*d* (500 MHz)

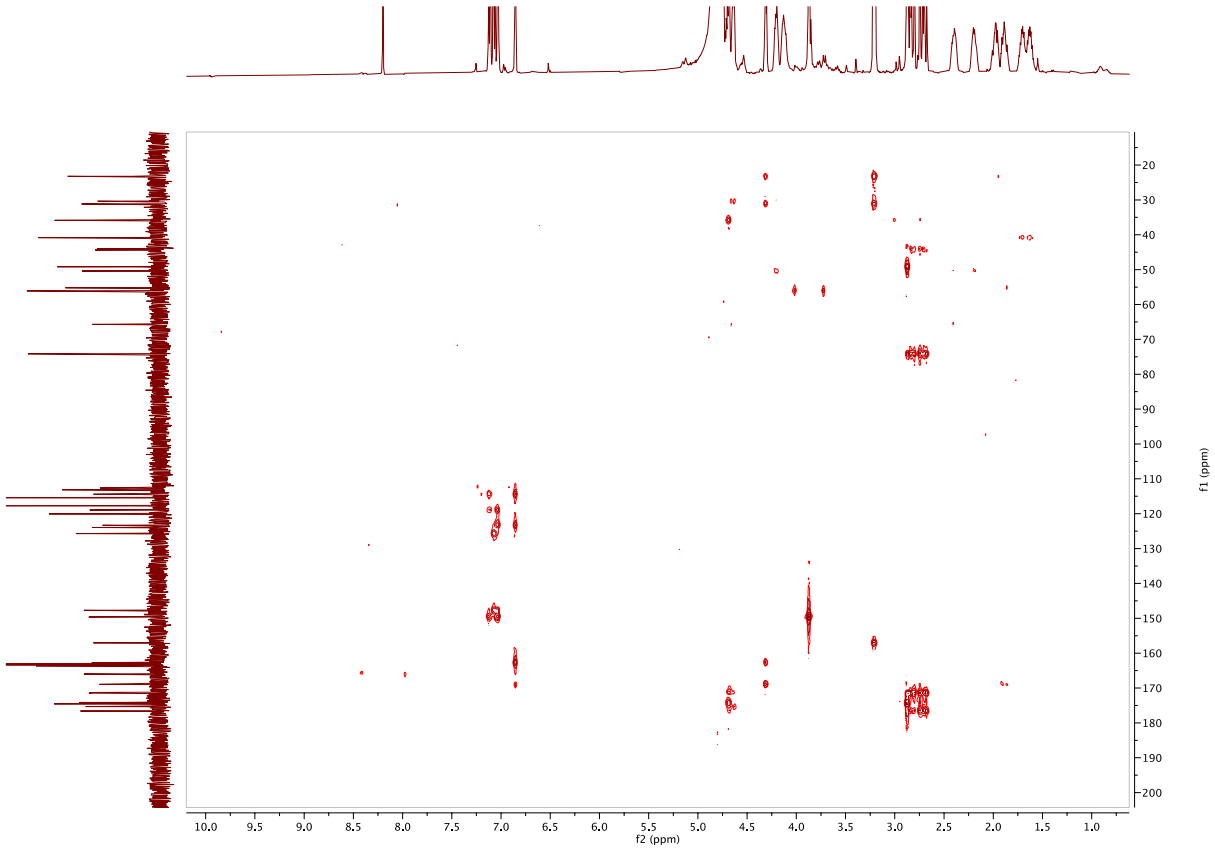
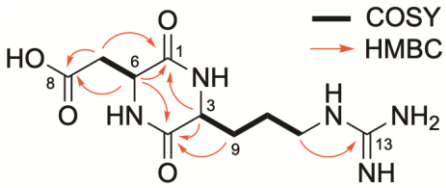


Figure S58. The HMBC spectrum of compound **10** in D_2O 0.5% TFA-*d* (500 MHz)

Table S16: Spectroscopic data of **11**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	169.8, C			
2	N			
3	54.1, CH	4.07, t (6.0)	1.70-1.87	23.3, 30.3, 169.7, 169.8
4	169.7, C			
5	N			
6	52.4, CH	4.23	2.55, 2.63	40.9, 169.7, 169.8, 177.1
7	40.9, CH2	Ha: 2.63, dd (16.3, 4.1); Hb: 2.55, dd (16.3, 7.7)	4.23	52.4, 169.8, 177.1
8	177.1, C			
9	30.3, CH2	1.70-1.87, m	1.48-1.66, 4.07	23.3, 40.7, 54.1, 169.7
10	23.3, CH2	1.48-1.66, m	1.70-1.87, 3.11	30.3, 40.5, 54.1
11	40.7, CH	3.11, t (6.9)	1.48-1.66	23.3, 30.3, 156.7
12	N			
13	156.7, C			



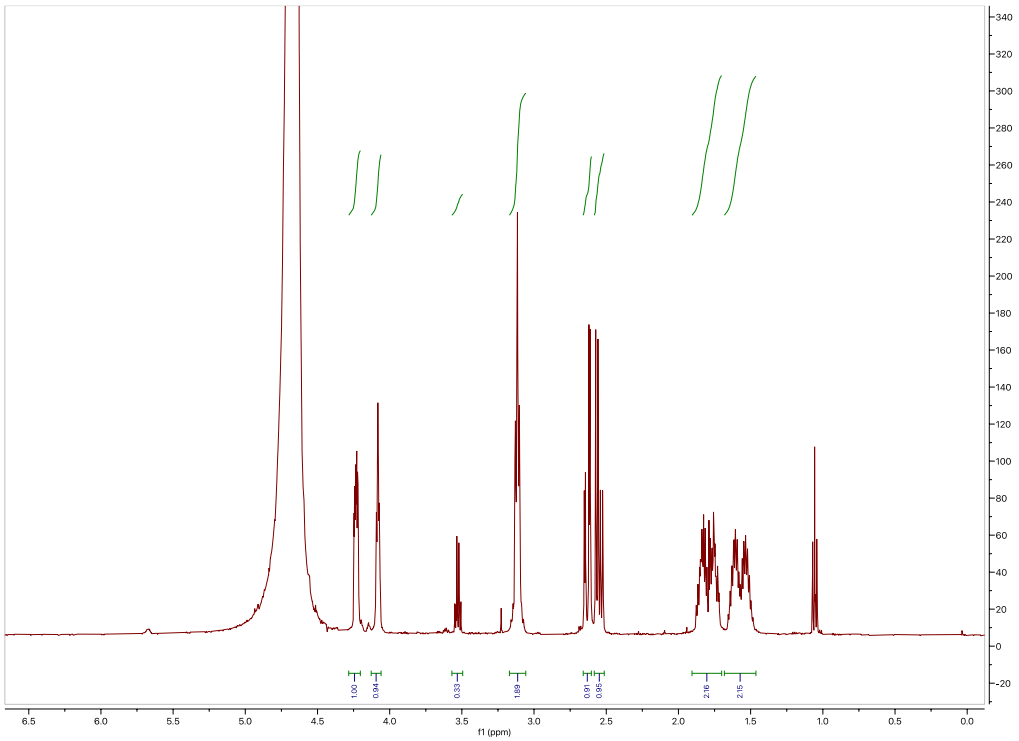


Figure S59. The ^{1}H NMR spectrum of compound **11** in D_2O (500 MHz)

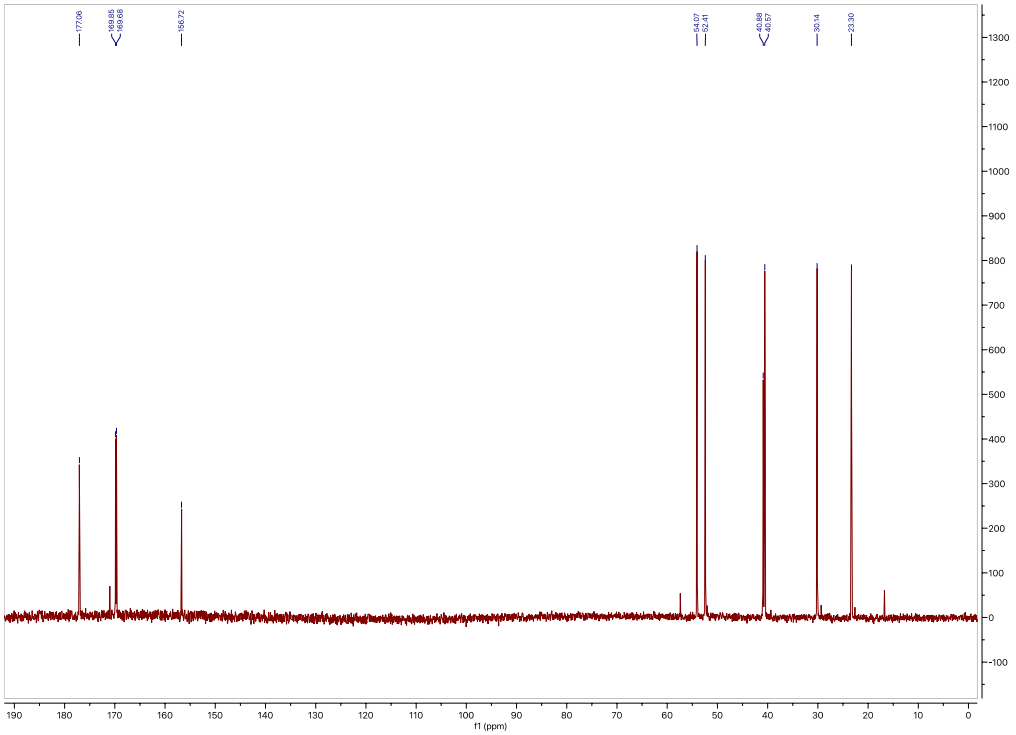


Figure S60. The ^{13}C NMR spectrum of compound **11** in D_2O (500 MHz)

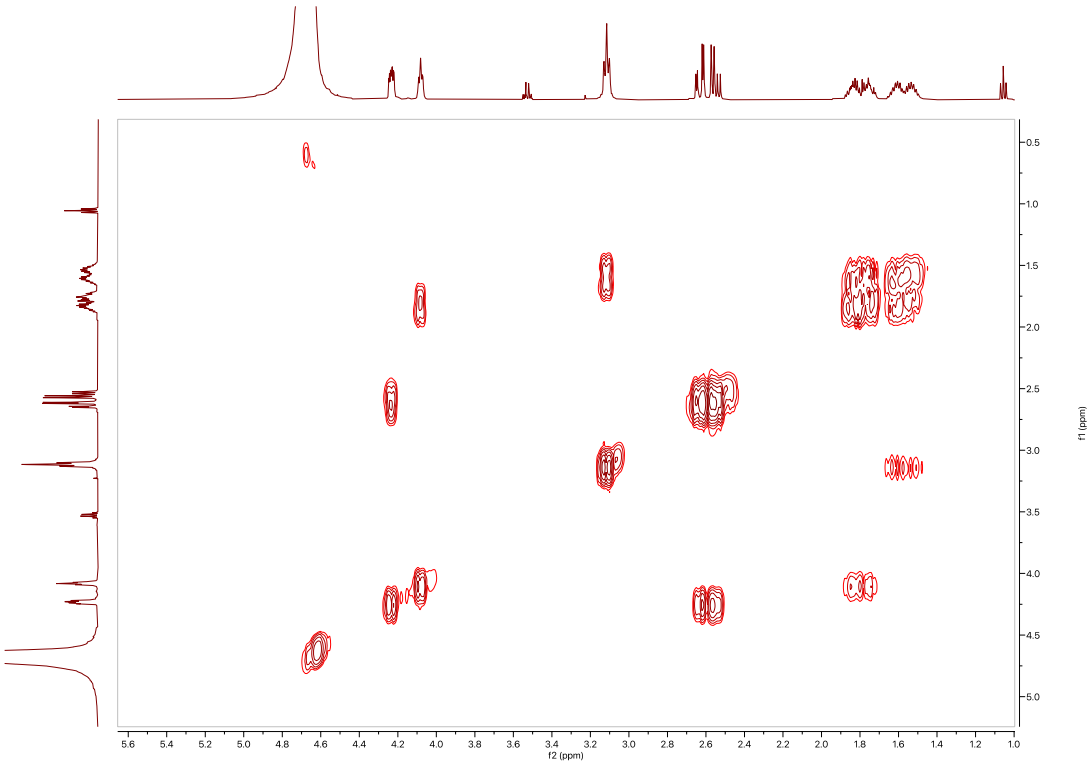


Figure S61. The COSY spectrum of compound **11** in D_2O (500 MHz)

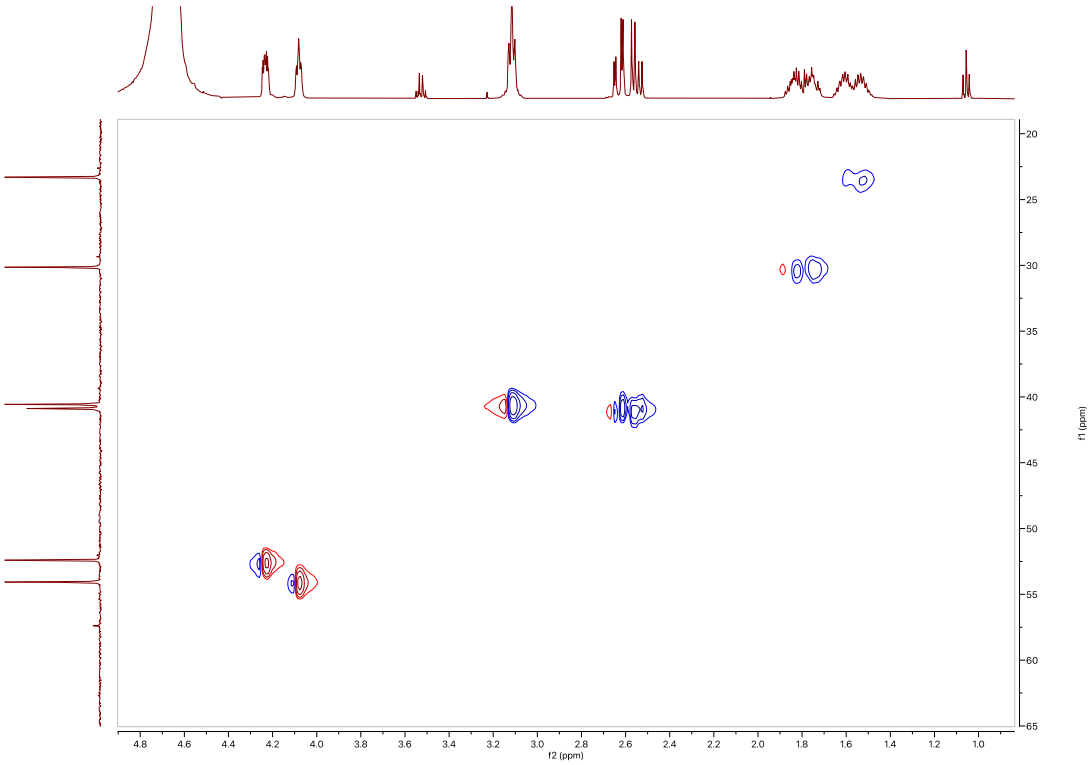


Figure S62. The HSQC spectrum of compound **11** in D_2O (500 MHz)

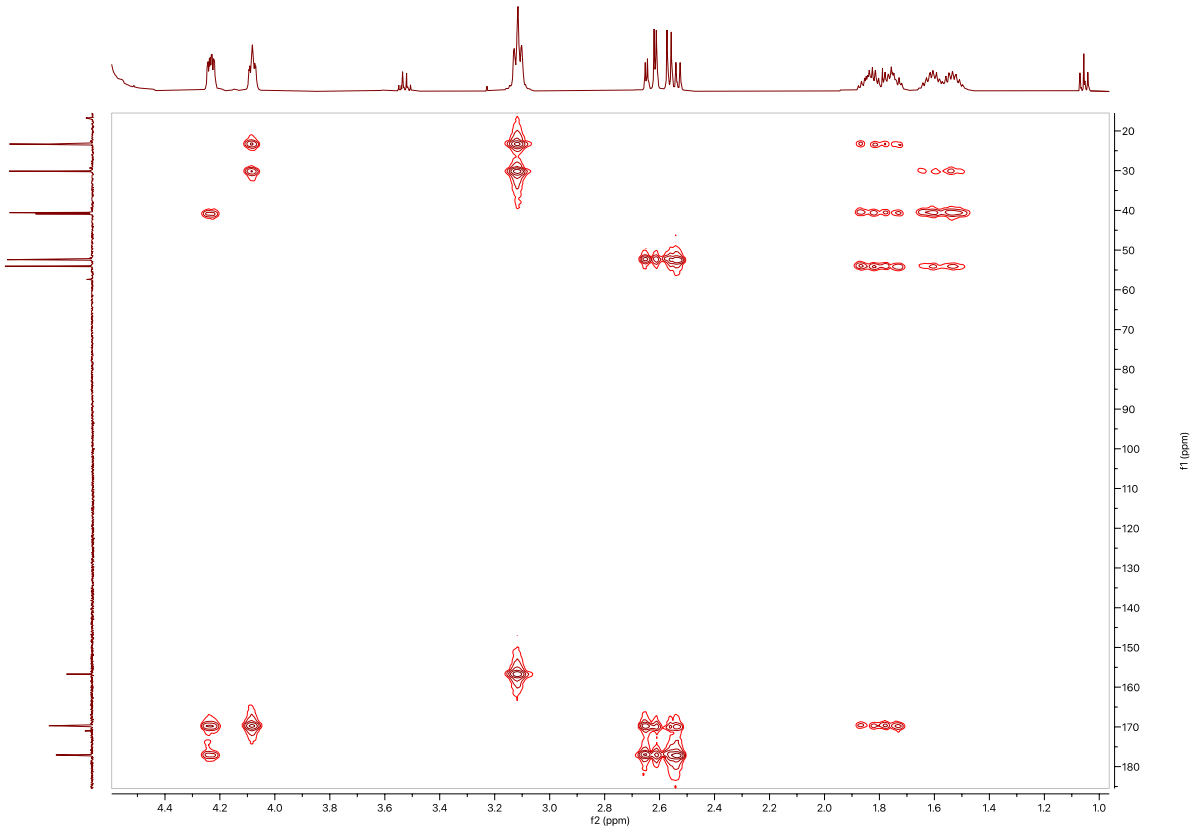
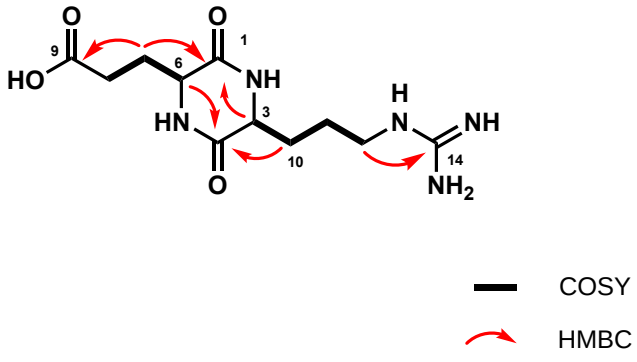


Figure S63. The HMBC spectrum of compound **11** in D_2O (500 MHz)

Table S17: Spectroscopic data of **12**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	169.8, C			
2	N			
3	54.1, CH	4.07, m	1.72-1.83	23.3, 30.2, 169.8
4	169.8, C			
5	N			
6	54.2, CH	4.08, m	1.97	30.3, 32.6, 169.8
7	30.3, CH2	1.97, m	2.06-2.20, 4.08	32.6, 54.2, 169.8, 181.2
8	32.6, CH2	2.06-2.20, m	1.97	30.3, 54.2, 181.2
9	181.2, C			
10	30.2, CH2	1.72-1.83, m	1.44-1.64, 4.07	23.3, 40.6, 54.1, 168.9
11	23.3, CH2	1.46-1.64, m	1.72-1.83, 3.13	30.2, 40.6, 54.1
12	40.6, CH	3.13, t (6.9)	1.46-1.64	23.3, 30.2, 156.7
13	N			
14	156.7, C			



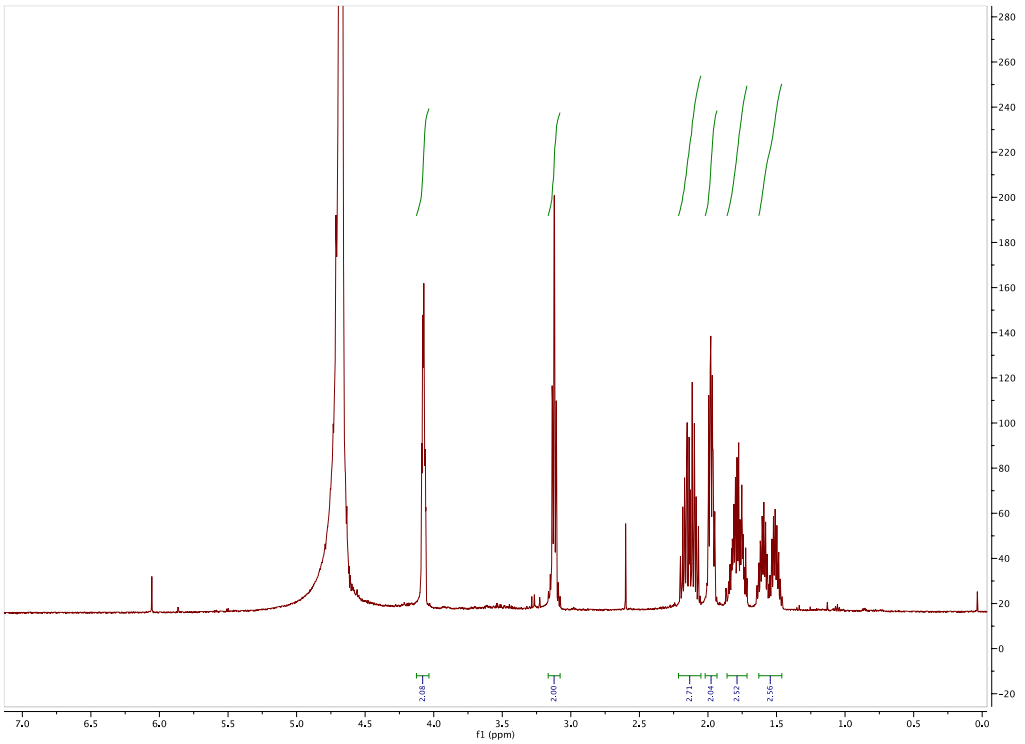


Figure S64. The ^{1}H NMR spectrum of compound **12** in D_2O (500 MHz)

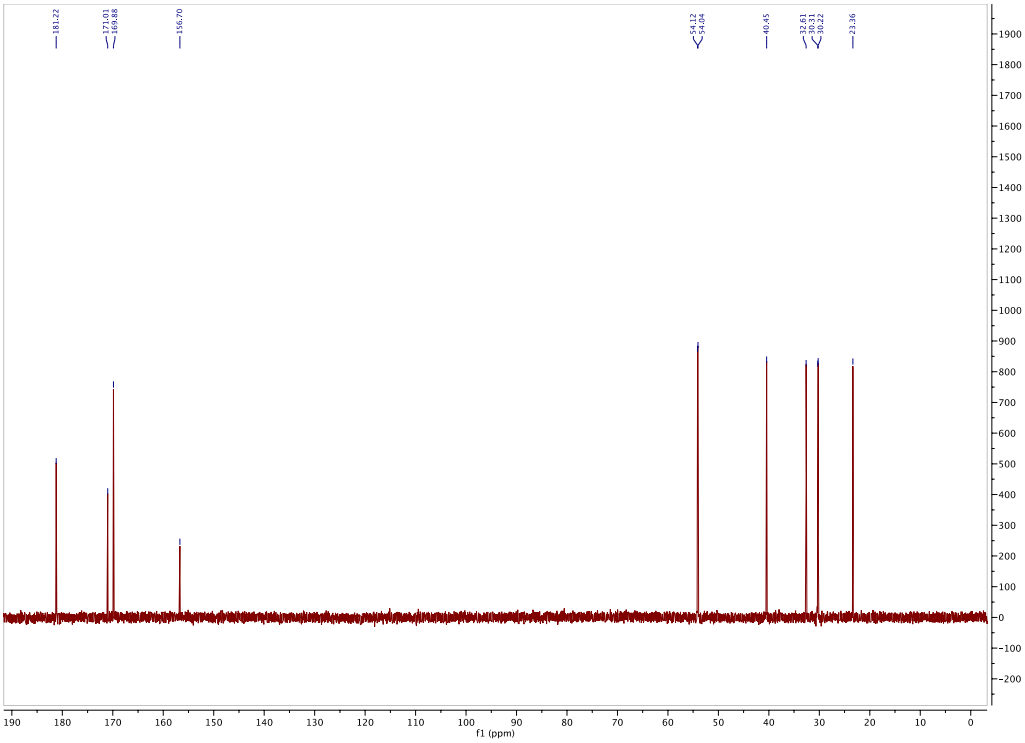


Figure S65. The ^{13}C NMR spectrum of compound **12** in D_2O (500 MHz)

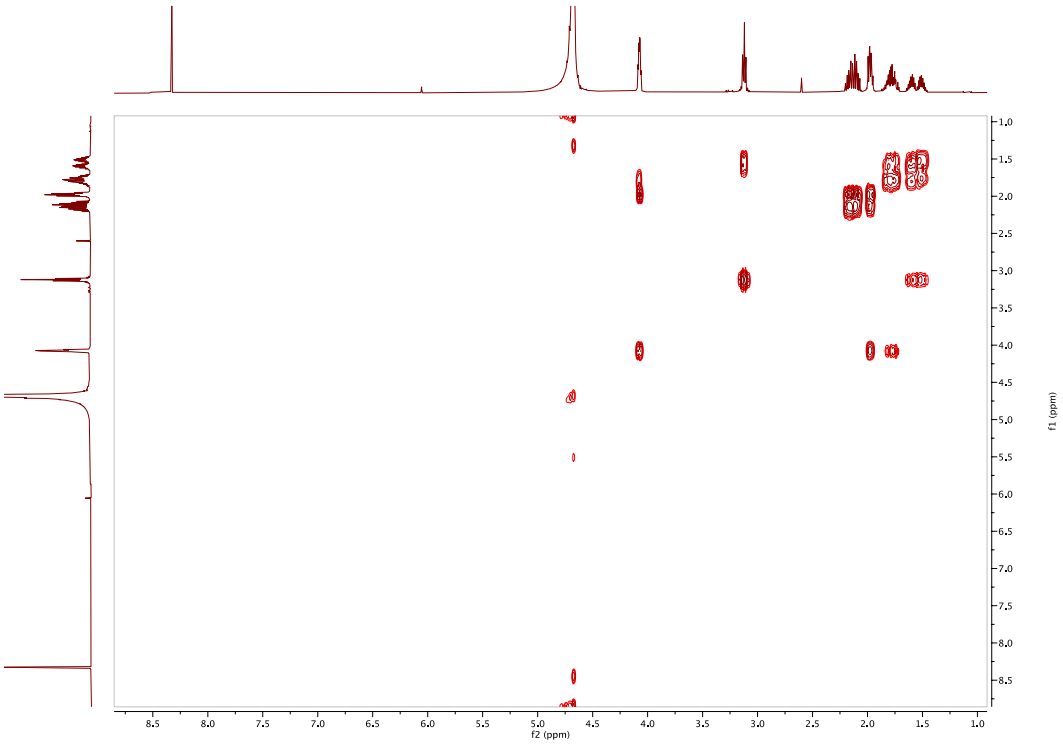


Figure S66. The COSY spectrum of compound **12** in D_2O (500 MHz)

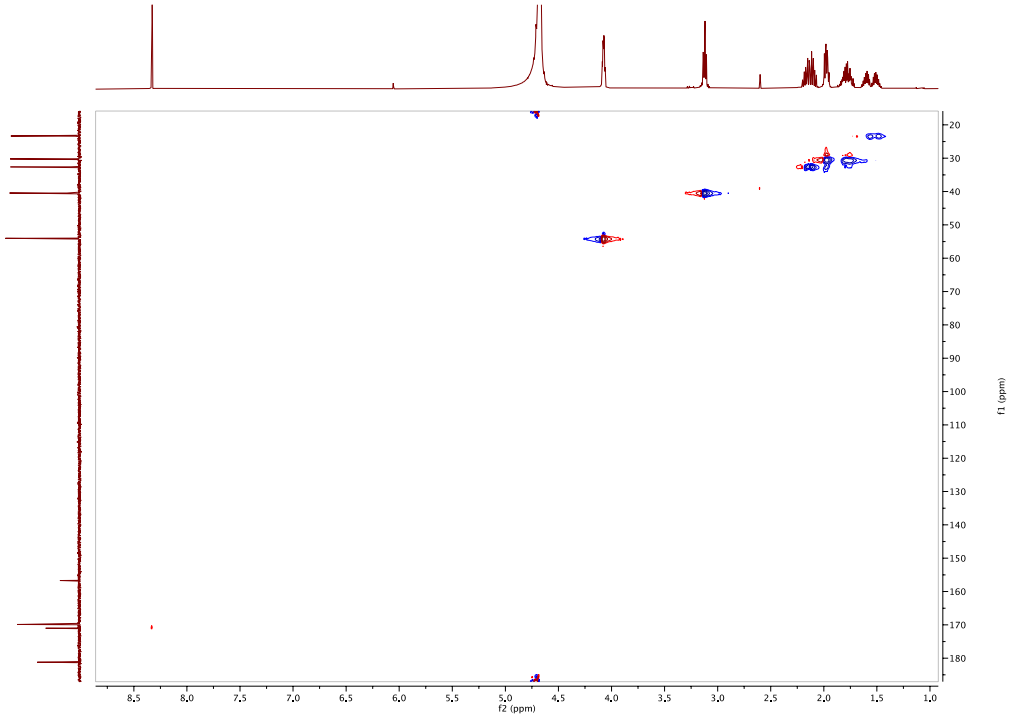
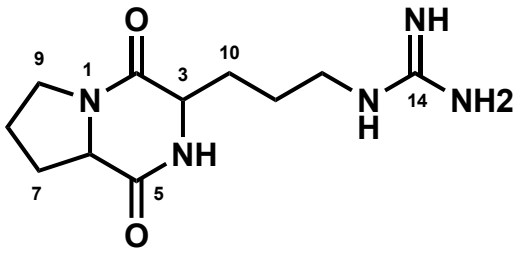


Figure S67. The HSQC spectrum of compound **12** in D_2O (500 MHz)

Figure S68. The HMBC spectrum of compound **12** in D_2O (500 MHz)

Table S18: Spectroscopic data of **13**

Position	**δC, mult (D_2O)**	**δH, mult (J in Hz) (D_2O)**	**Reported δC, mult (D_2O)[153]**	**Reported δH, mult (J in Hz) (D_2O)[153]**
1	N			
2	169.6, C		169.6	
3	57.2, CH	4.36	57.2	4.41
4	NH			
5	174.9, C		174.9	
6	CH	4.34	61.6	4.38
7	30.4, CH2	2.07, 2.34	30.4	2.03, 2.41
8	24.3, CH2	1.85-2.01	24.3	2.03, 2.13
9	47.9, CH2	3.57	47.9	3.62
10	28.8, CH2	1.85-2.01	28.8	1.97, 2.03
11	25.4, CH2	1.57, 1.68	25.5	1.63, 1.76
12	43.3, CH2	3.23	43.3	3.29
13	NH			
14	159.3, C		159.3	



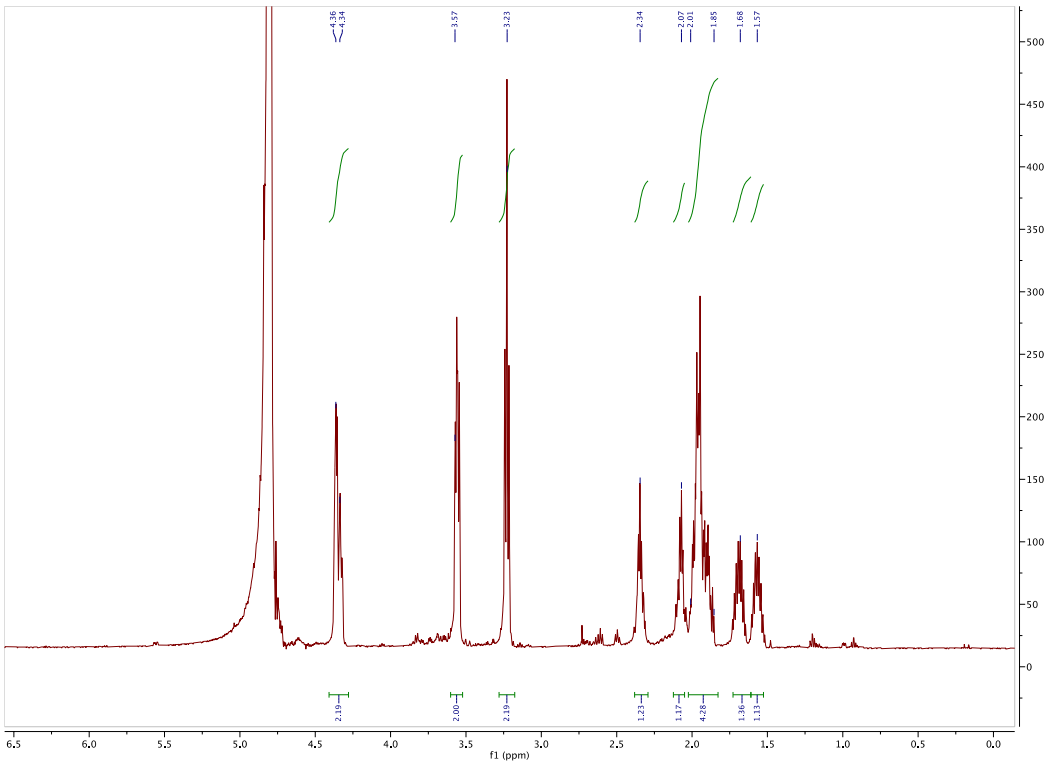


Figure S69. The ^{1}H NMR spectrum of compound **13** in D_2O (500 MHz)

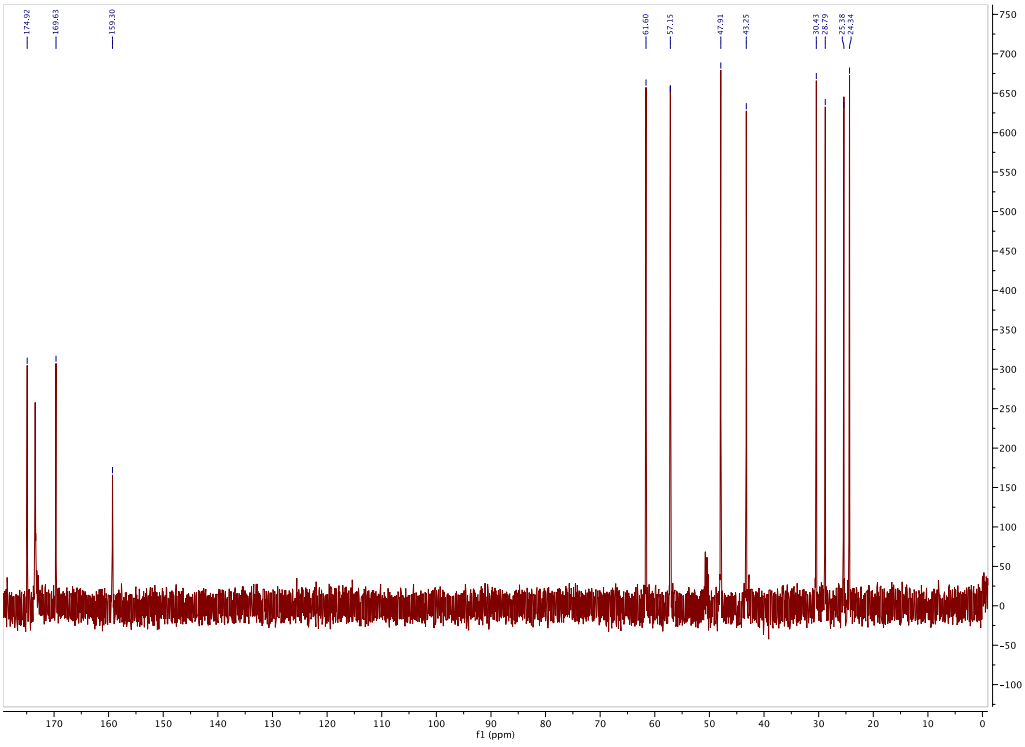
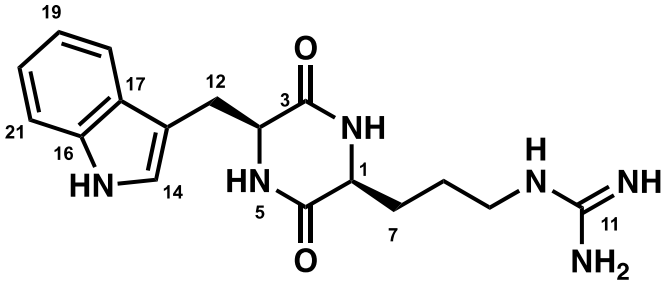


Figure S70. The ^{13}C NMR spectrum of compound **13** in D_2O (500 MHz)

Table S19: Spectroscopic data of **14**

Position	**δC, mult (D_2O)**	**δH, mult (J in Hz) (D_2O)**	**Reported δC, mult (MeOH-*d4*)**[154]	**Reported δH, mult (J in Hz) (MeOH-*d4*)**[154]
1	53.55, CH	3.84	55.17	3.66
2	NH			
3	169.49, C		169.87	
4	55.76, CH	4.48	57.49	4.31
5	NH			
6	169.17, C		169.48	
7	29.72, CH2	0.39	32.06	0.49
8	22.12, CH2	0.63, 0.94	24.53	0.87
9	40.49, CH2	2.63	41.72	2.62
10	NH			
11	156.37, C		158.44	
12	28.47, CH2	3.20, 3.53	30.48	3.14, 3.49
13	107.66, C		109.66	
14	125.44, CH	7.19	126.07	7.08
15	NH			
16	135.91, C		137.80	
17	127.57, C		129.38	
18	119.45, CH	7.64	120.28	7.63
19	118.82, CH	7.16	120.22	7.01
20	121.80, CH	7.22	122.54	7.09
21	111.56, CH	7.49	112.18	7.34



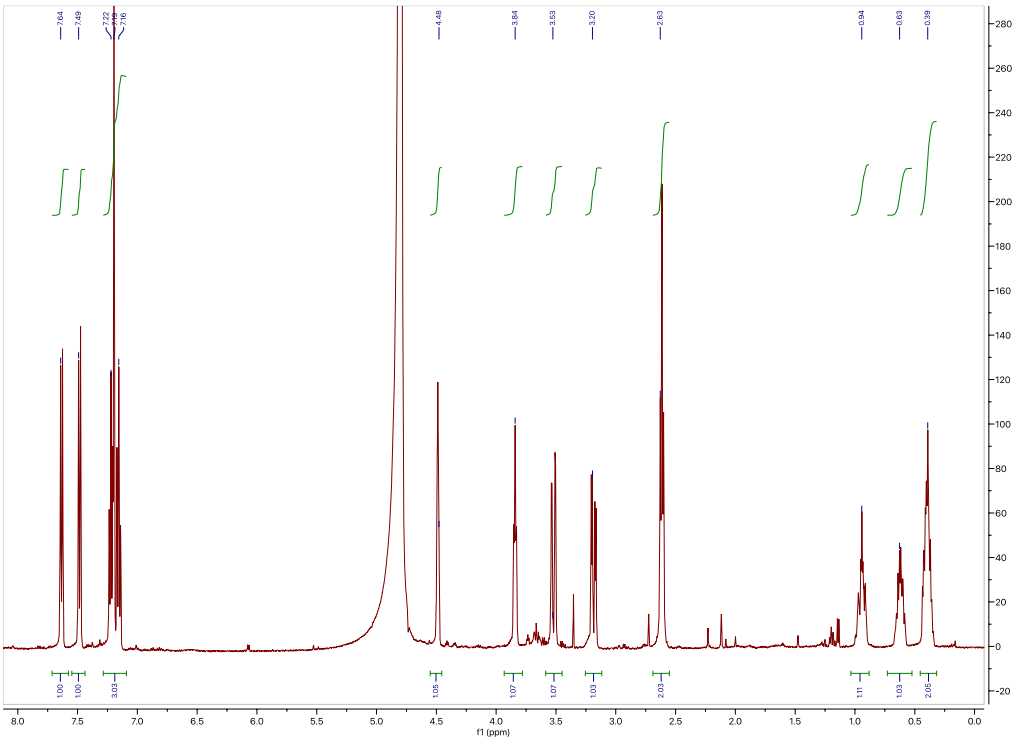


Figure S71. The ^{1}H NMR spectrum of compound **14** in D_2O (500 MHz)

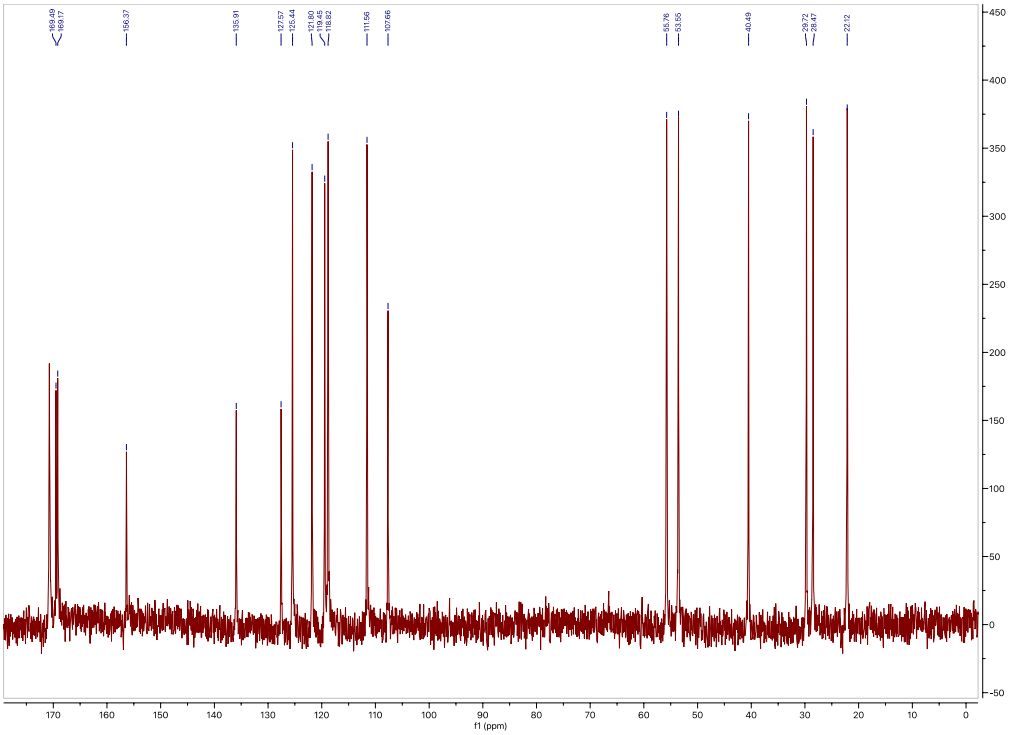
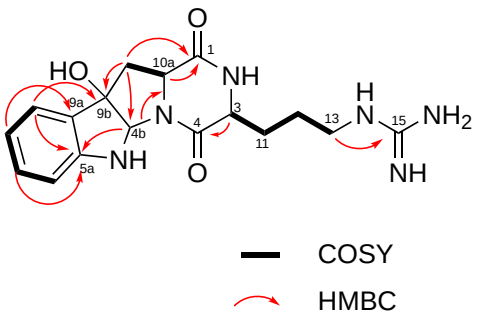


Figure S72. The ^{13}C NMR spectrum of compound **14** in D_2O (500 MHz)

Table S20: Spectroscopic data of **15**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	170.5, C			
2	N			
3	54.5, CH	4.35, m	2.02, 1.90	
4	167.2, C			
4a	N			
4b	79.7, CH	5.53, s		149.4, 128.7, 57.9
5	N			
5a	149.4, C			
6	111.3, CH	6.87, d (7.3)	7.35	128.7, 121.2
7	131.1, CH	7.35, t (7.3)	7.04, 6.87	149.3, 124.3
8	121.2, CH	7.04, t (7.3)	7.49, 7.35	128.7, 111.3
9	124.3, CH	7.49, d (7.3)	7.04	149.4, 131.1
9a	128.7, C			
9b	85.6, C			
10	40.3, CH2	Ha: 2.96, dd (11.8, 6.3); 2.54, dd (12.0, 11.8)	Ha: 4.23, 2.54; Hb: 4.23, 2.96	Ha: 128.7, 85.6, 79.7, 57.9; Hb: 170.5, 128.7, 85.6, 57.9
10a	57.9, CH	4.23, dd (12.0, 6.3)	2.96, 2.54	170.5
11	26.5, CH2	Ha: 2.02, m; Hb: 1.90, m	Ha: 4.35, 1.90, 1.73, 1,60; Hb: 4.35, 2.02, 1.73, 1.60	Hb: 167.2, 54.5, 22.6
12	22.6, CH2	Ha: 1.73, m; Hb: 1.60, m	Ha: 3.26, 2.02, 1.90, 1.60; Hb: 3.26, 2.02, 1.90, 1.73	Ha and Hb: 40.6, 40.3
13	40.6, CH2	3.26	1.73, 1.60	156.7, 26.5, 22.6
14	N			
15	156.7, C			



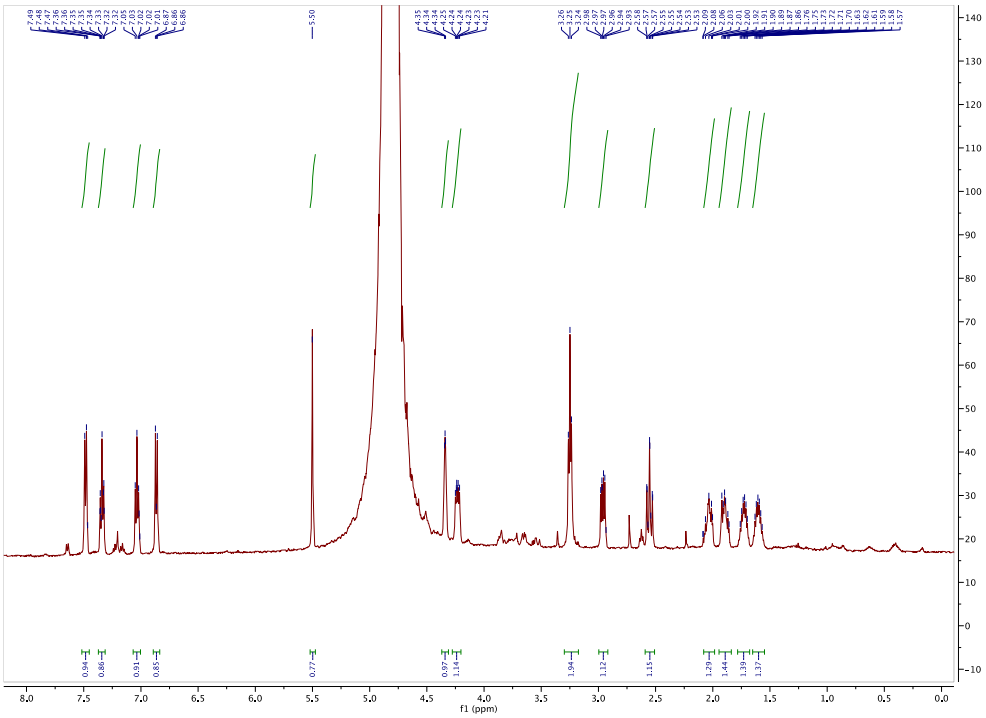


Figure S73. The ^{1}H NMR spectrum of compound **15** in D_2O (500 MHz)

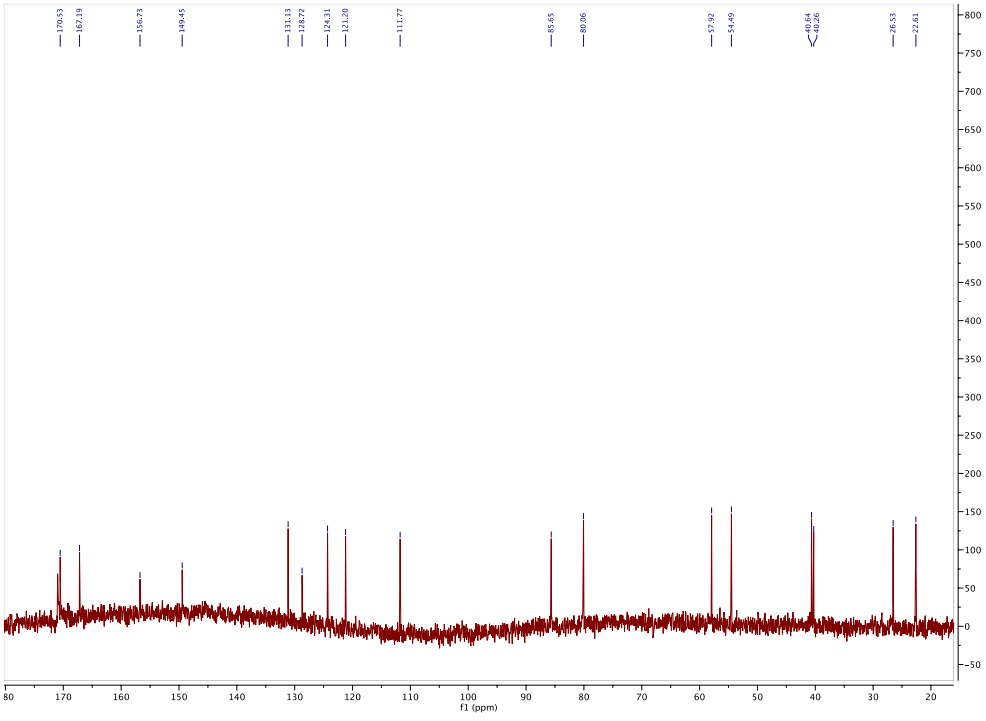


Figure S74. The ^{13}C NMR spectrum of compound **15** in D_2O (500 MHz)

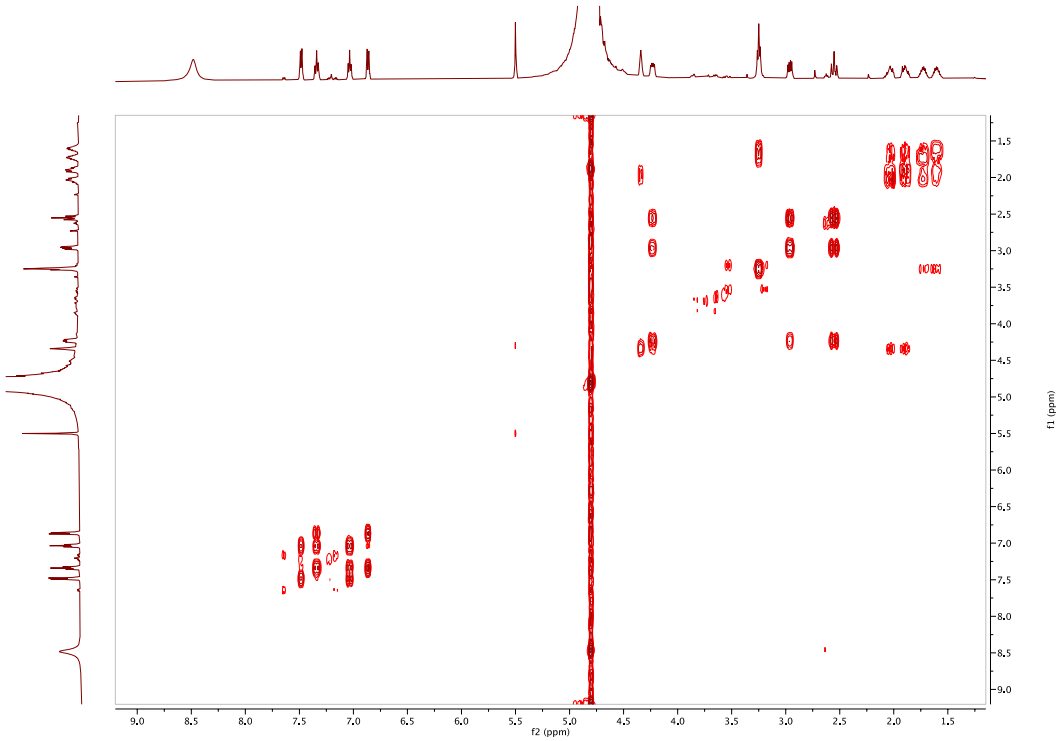


Figure S75. The COSY spectrum of compound **15** in D_2O (500 MHz)

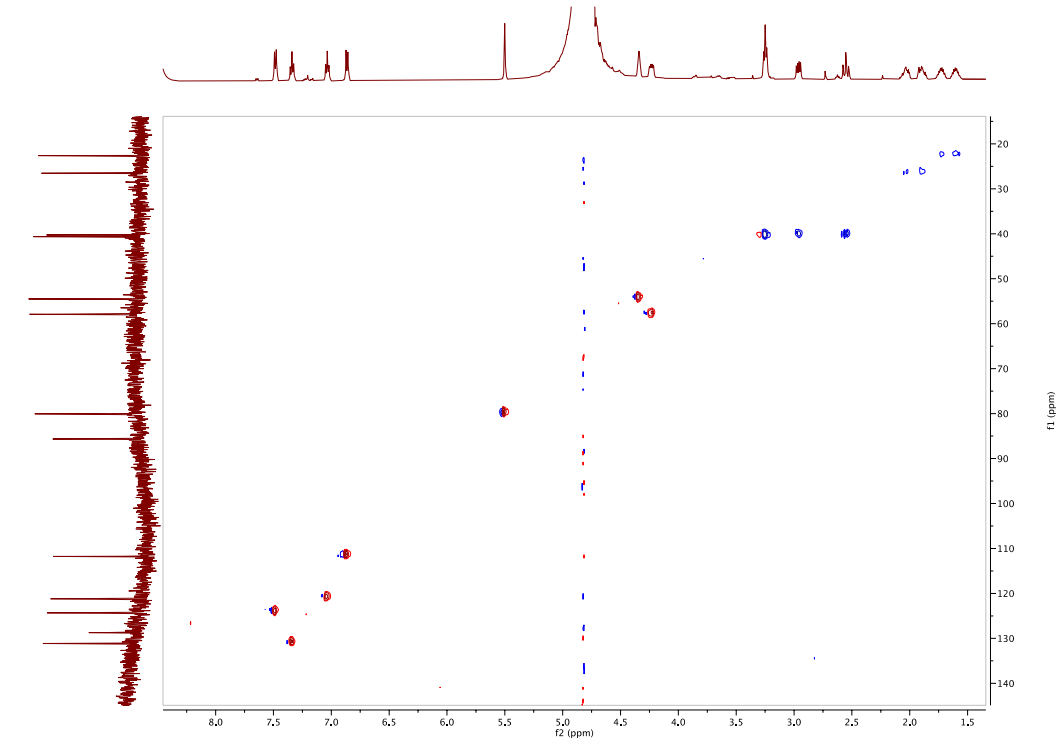


Figure S76. The HSQC spectrum of compound **15** in D_2O (500 MHz)

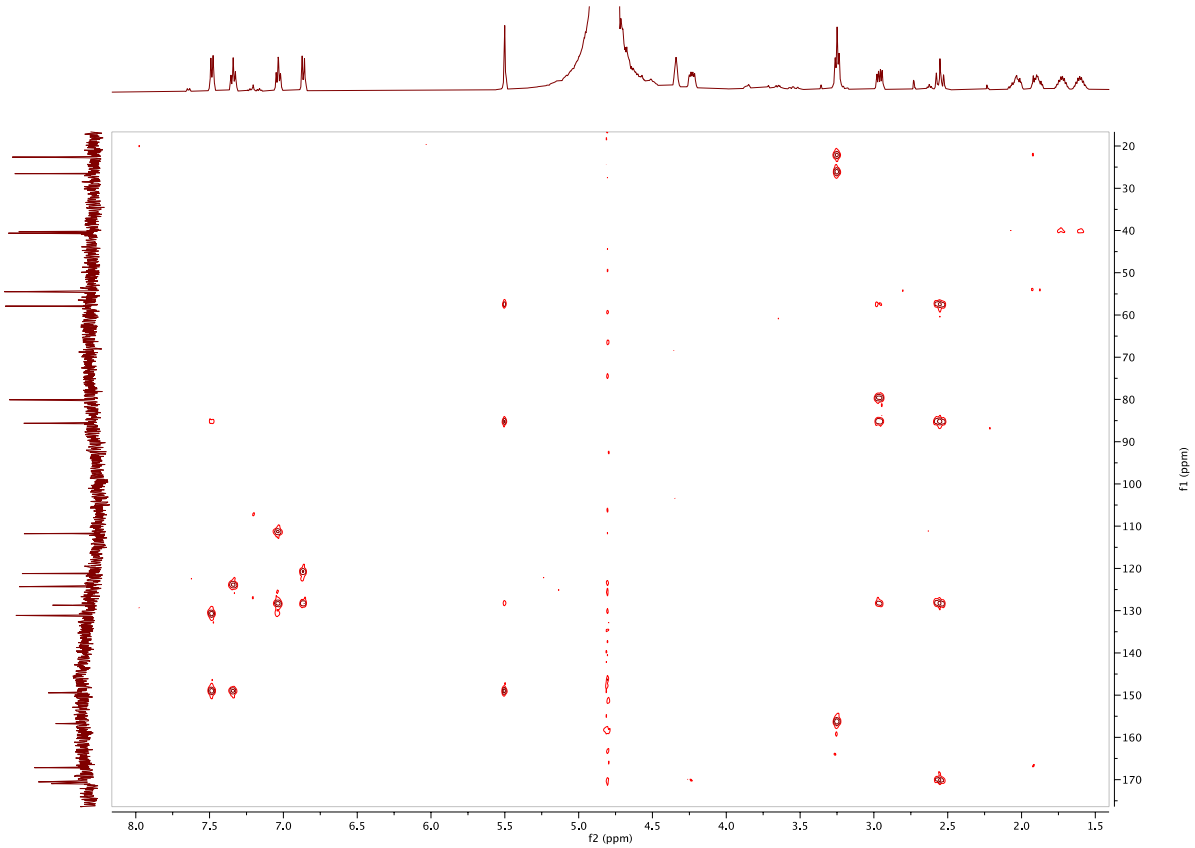
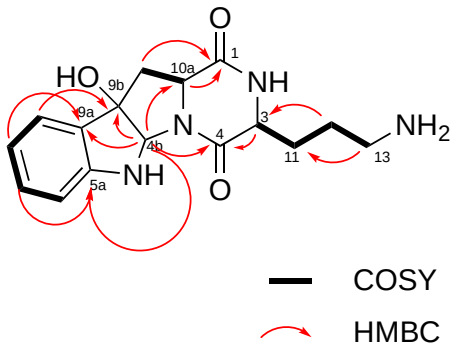


Figure S77. The HMBC spectrum of compound **15** in D_2O (500 MHz)

Table S21: Spectroscopic data of **16**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	170.7, C			
2				
3	54.2, CH	4.34, m	2.05, 1.91	167.0
4	167.0, C			
4a				
4b	80.0, CH	5.49, s		167.0, 149.4, 128.7, 85.6, 58.0
5				
5a	149.4, C			
6	111.8, CH	6.86, dd (7.6, 2.2)	7.33, 7.02	128.7, 121.2
7	131.1, CH	7.33, ddd (7.6, 7.4, 1.6)	7.47, 7.02, 6.86	149.4, 124.3, 121.1, 111.9
8	121.2, CH	7.02, ddd (7.4, 7.1, 2.2)	7.47, 7.33, 6.86	131.0, 128.7, 111.8
9	124.3, CH	7.47, dd (7.1, 1.6)	7.33, 7.02	149.4, 131.1, 111.8, 85.6
9a	128.7, C			
9b	85.6, C			
10	40.1, CH2	Ha: 2.95, dd (12.3, 6.2); Hb: 2.56, dd (12.3, 12.3)	Ha: 4.23, 2.56; Hb: 4.23, 2.95	Ha: 128.7, 85.6, 80.0, 58.0; Hb: 170.7, 128.7, 85.6, 58.0
10a	58.0, CH	4.23, dd (12.3, 6.2)	2.95, 2.56	170.7, 40.1
11	26.2, CH2	Ha: 2.05, m; Hb: 1.91, m	Ha: 4.34, 1.91, 1.81, 1.69; Hb: 4.34, 2.05, 1.81, 1.69	
12	21.5, CH2	Ha: 1.81, m; Hb: 1.69, m	Ha: 3.04, 2.05, 1.91, 1.69; Hb: 3.04, 2.05, 1.91, 1.81	
13	39.0, CH2	3.04, t (7.6)	1.81, 1.69	26.2, 21.5



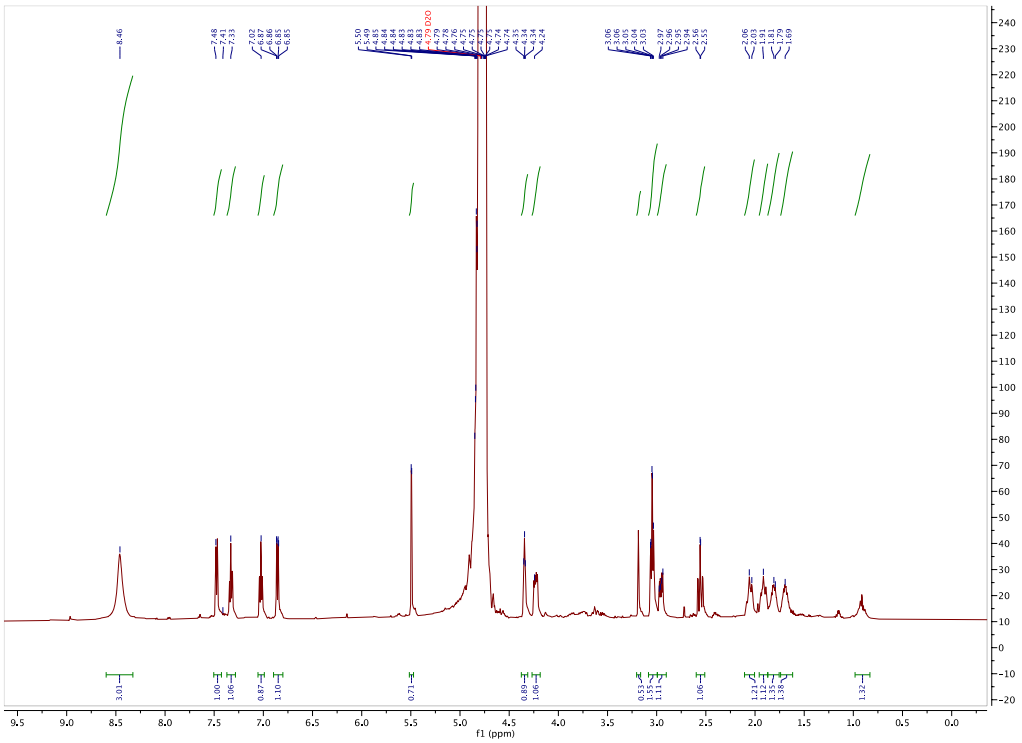


Figure S78. The ^{1}H NMR spectrum of compound **16** in D_2O (500 MHz)

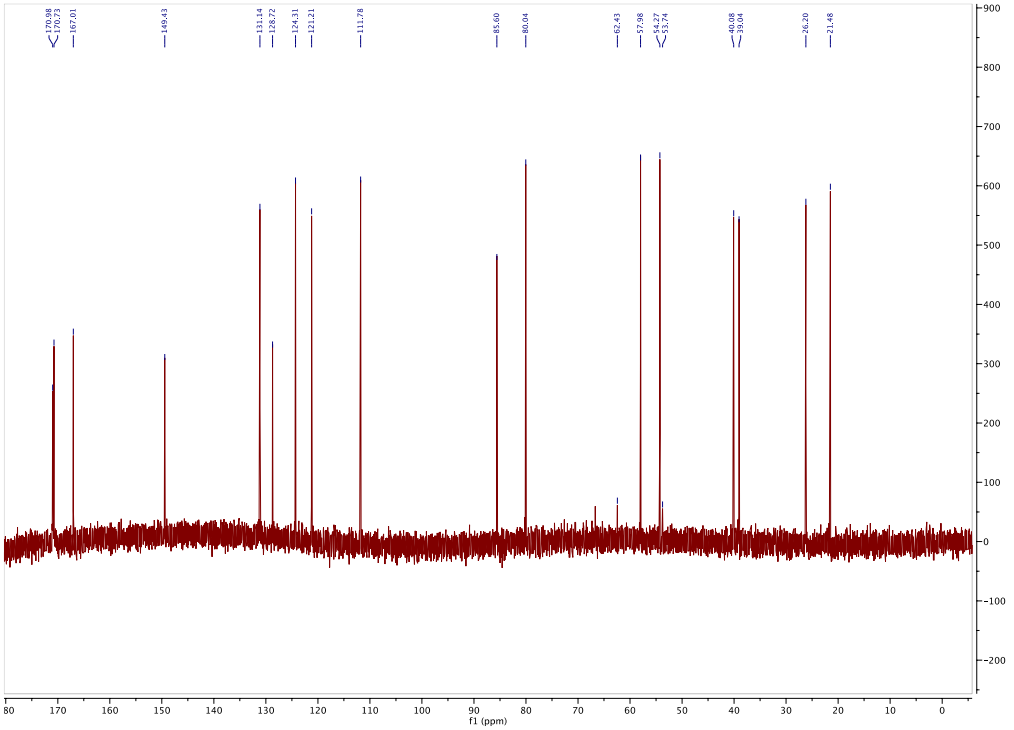


Figure S79. The ^{13}C NMR spectrum of compound **16** in D_2O (500 MHz)

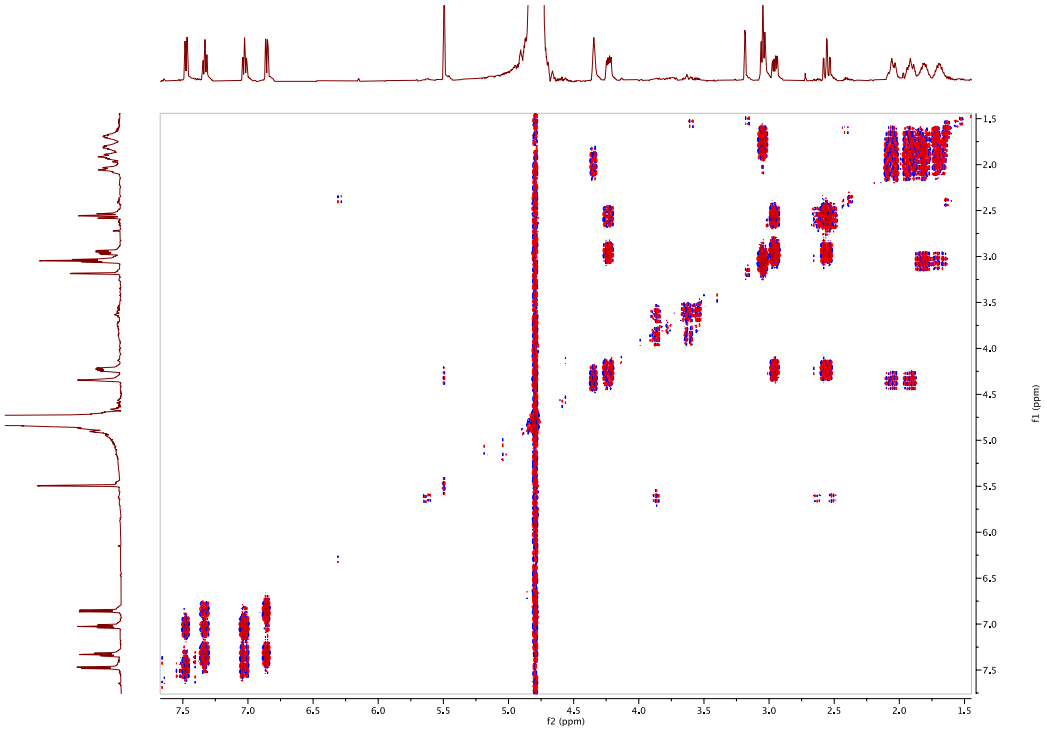


Figure S80. The COSY spectrum of compound **16** in D_2O (500 MHz)

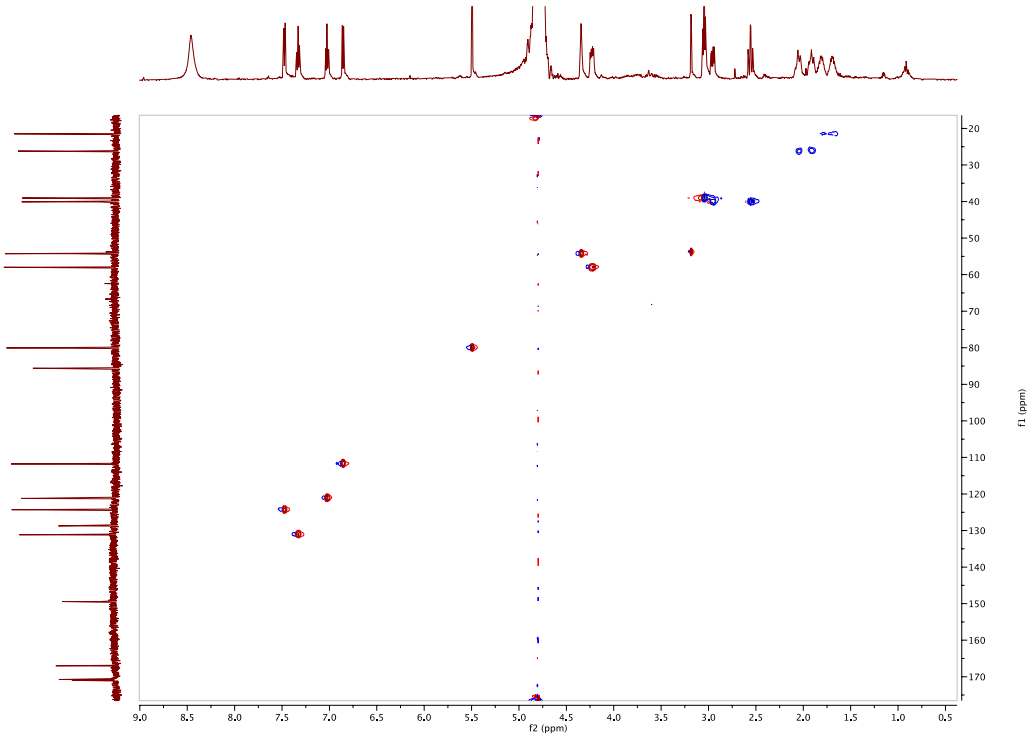


Figure S81. The HSQC spectrum of compound **16** in D_2O (500 MHz)

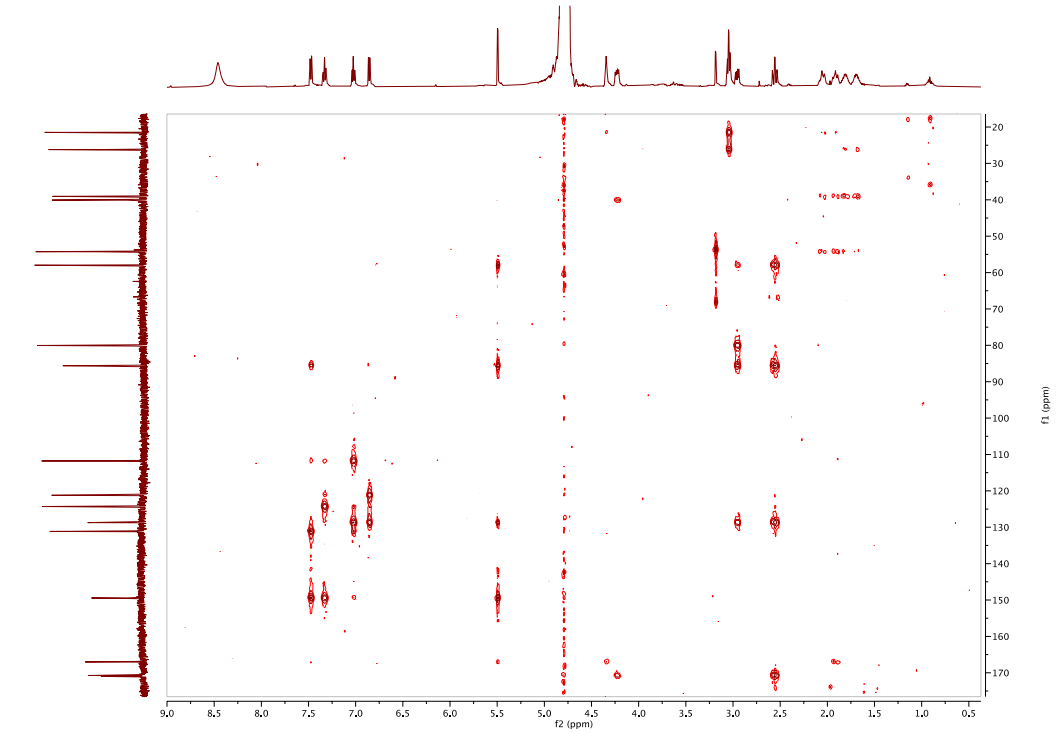


Figure S82. The HMBC spectrum of compound **16** in D_2O (500 MHz)

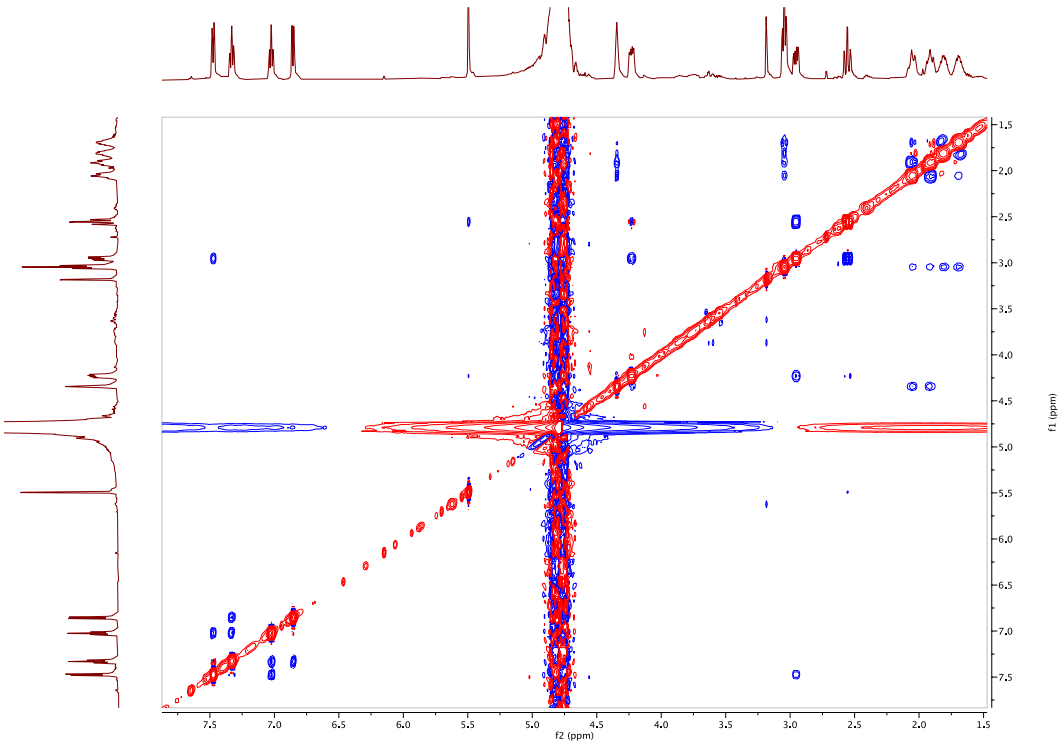
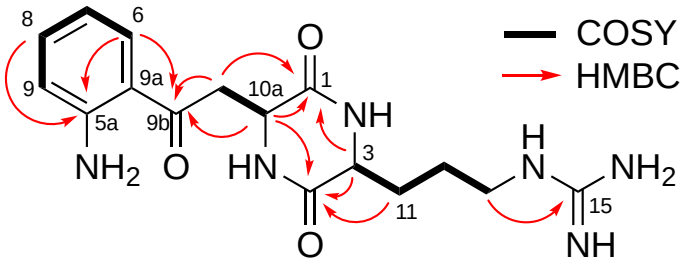


Figure S83. The NOESY spectrum of compound **16** in D_2O (500 MHz)

Table S22: Spectroscopic data of **17**

Position	δC, mult	δH, mult (J in Hz)	COSY	HMBC
1	169.6, C			
2	NH			
3	54.0, CH	4.21, t (4.8)	1.95, 1.86	
4	169.9, C			
4a	NH			
5	N			
5a	132.9			
6	131.8	8.09, d (8.2)	7.52	199.6, 135.4, 132.9, 126.5
7	127.9	7.52, dd (7.2, 8.2)	8.09, 7.70	124.2, 126.5
8	135.4	7.70, dd (7.2, 8.2)	8.09, 7.52	131.8
9	124.2	7.40, d (8.2)	7.70	126.5
9a	126.5			
9b	199.6			
10	42.4	Ha: 3.78, dd (3.8, 18.2); Hb: 3.61, dd (6.8, 18.2)	Ha: 4.63, 3.61; Hb: 4.63, 3.78	Ha: 199.6, 169.6, 50.6; Hb: 199.6, 169.6, 50.6
10a	50.6	4.63, m	3.78, 3.61	199.6, 169.6-169.9, 42.4
11	29.6	Ha: 1.95, m; Hb: 1.86, m	Ha: 4.21, 1.86, 1.75, 1.63; Hb: 4.21, 1.95, 1.75, 1.63	Ha: 54.0, 23.3; Hb: 169.9, 54.0
12	23.3	Ha: 1.75, m; Hb: 1.63, m	Ha: 3.26, 1.95, 1.86, 1.63; Hb: 3.26, 1.95, 1.86, 1.75	Ha and Hb: 42.4
13	40.6	3.26, t (7.2)	1.75, 1.63	156.7, 29.6, 23.3
14	N			
15	156.7			



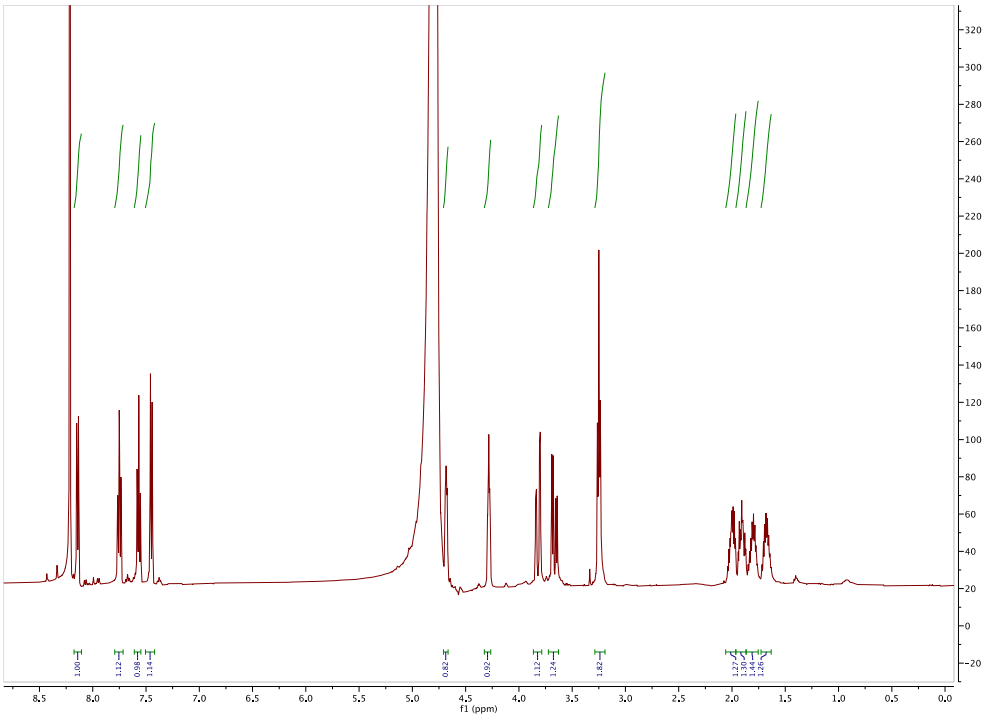


Figure S84. The ^{1}H NMR spectrum of compound **17** in D_2O (500 MHz)

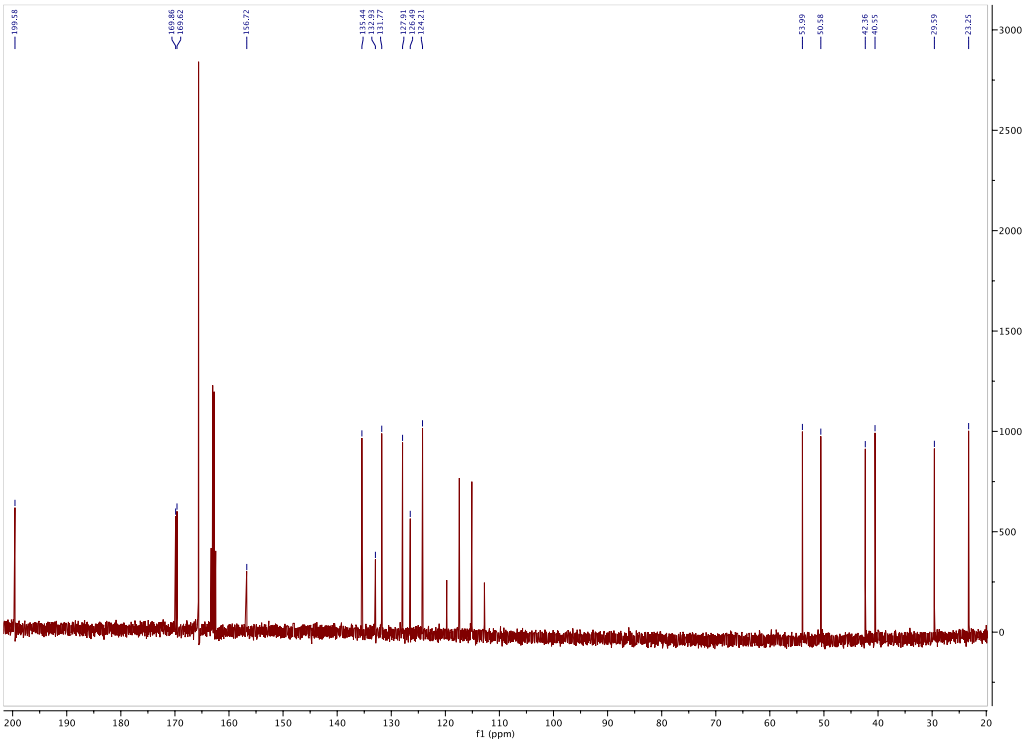


Figure S85. The ^{13}C NMR spectrum of compound **17** in D_2O (500 MHz)

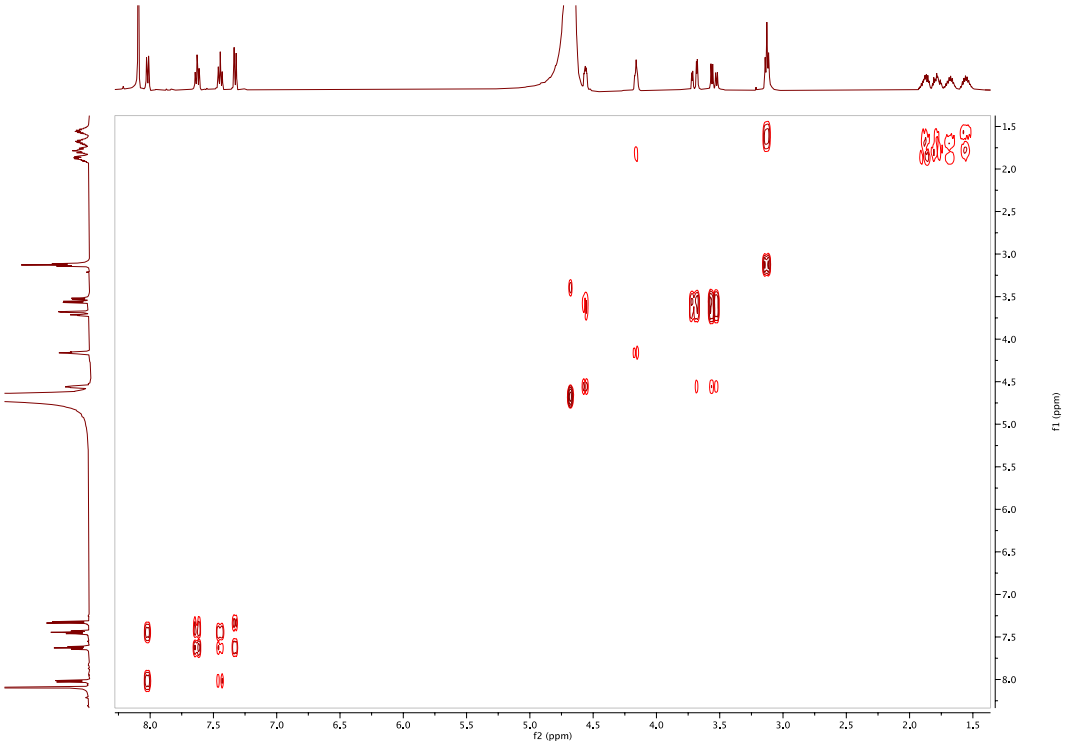


Figure S86. The COSY spectrum of compound **17** in D_2O (500 MHz)

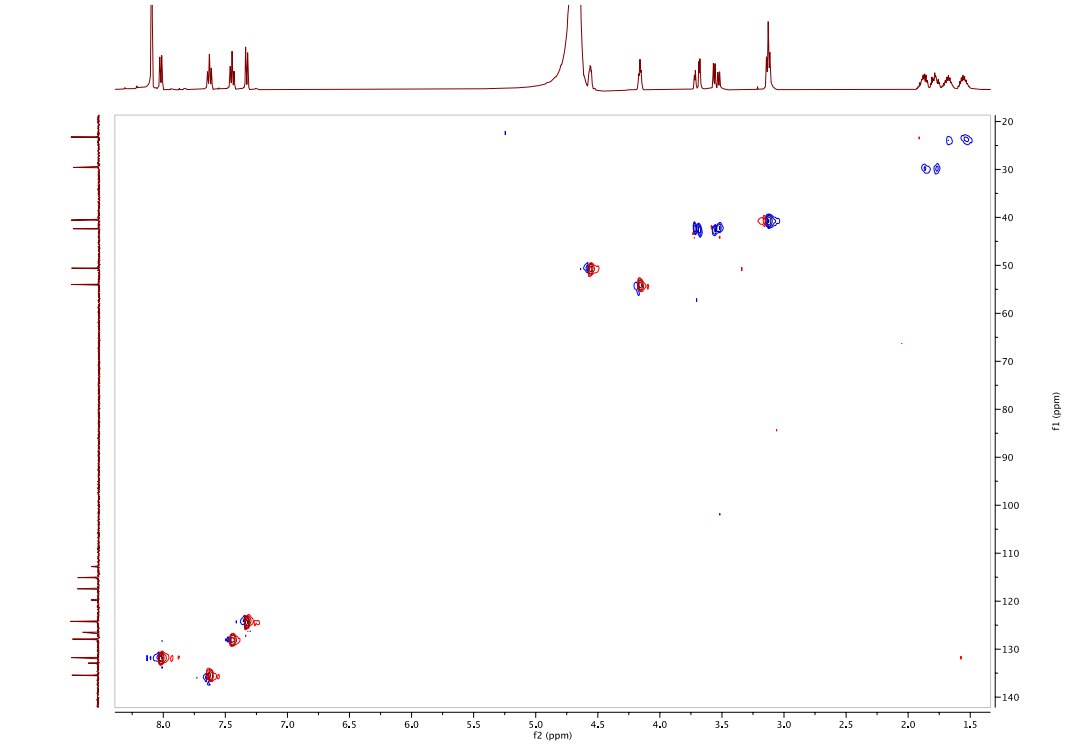


Figure S87. The HSQC spectrum of compound **17** in D_2O (500 MHz)

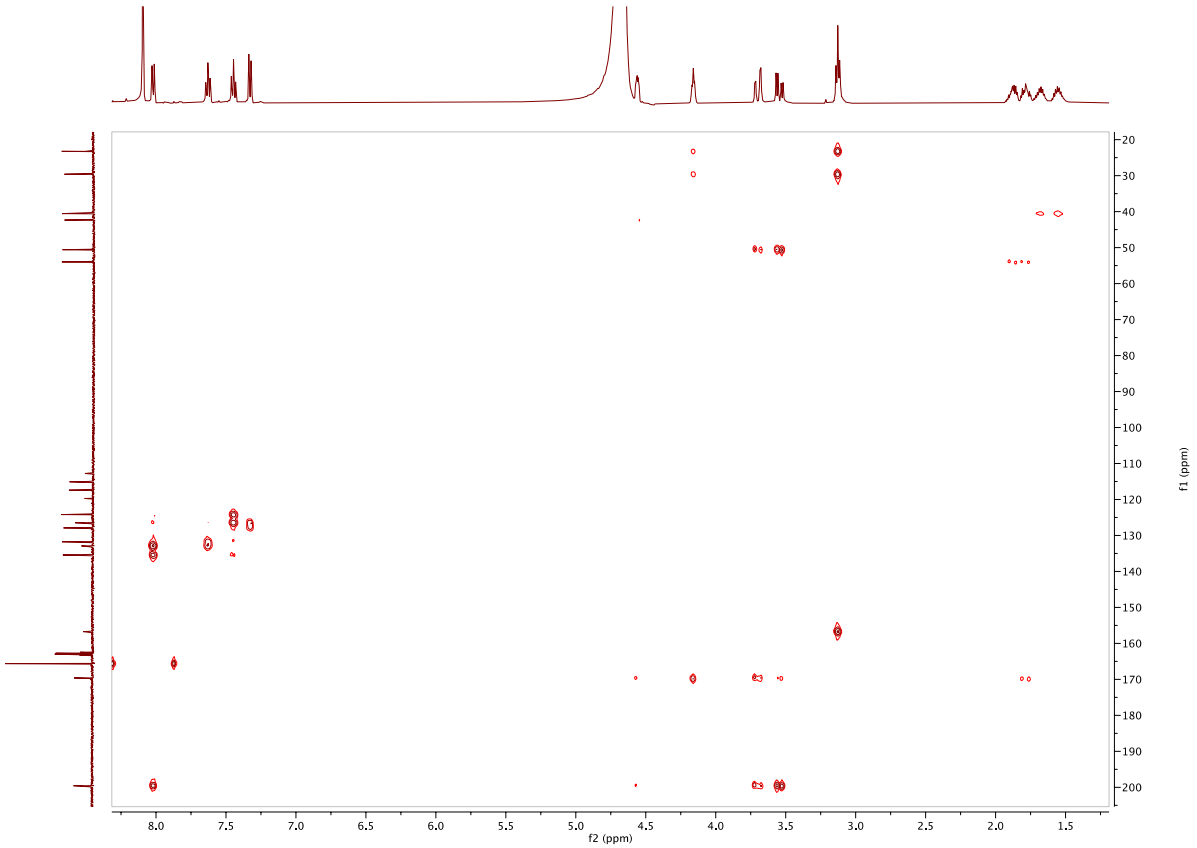
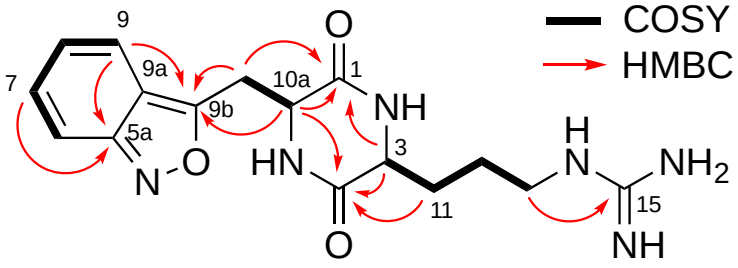


Figure S88. The HMBC spectrum of compound **17** in D_2O (500 MHz)

Table S23: Spectroscopic data of **18**

Position	**δC, mult**	**δH, mult (J in Hz)**	**COSY**	**HMBC**
1	168.1, C			
2	N			
3	53.9, CH	4.03, t (4.1)	1.18	
4	169.2, C			
4a	N			
5	N			
5a	157.3, C			
6	113.8, CH	7.48, d (8.9)	7.40	125.1, 117.7
7	133.0, CH	7.40, dd (6.3, 8.9)	7.07, 7.48	156.7, 125.1, 120.2
8	125.1, CH	7.07, dd (6.3, 8.9)	7.54, 7.40	133.0, 117.7, 113.8
9	120.2, CH	7.54, d (8.9)	7.07	156.7, 133.0, 117.7, 164.8
9a	117.7, C			
9b	164.8, C			
10	30.3, CH2	Ha: 3.96, dd (3.3, 15.5); Hb: 3.62, dd (5.1, 15.5)	Ha: 4.66, 3.62; Hb: 4.66, 3.96	Ha and Hb: 53.9 or 54.1, 117.7, 164.8, 168.1
10a	54.1, CH	4.66, m	3.96, 3.62	30.3, 164.8, 168.1
11	29.4, CH2	Ha: 1.18, m; Hb: 1.05, m	Ha: 4.03, 1.05, 0.52, 0.35; Hb: 4.03, 1.18, 0.52, 0.35	Ha: 169.2, 53.9, 40.7, 22.3; Hb: 53.9, 40.7, 22.3
12	22.3, CH2	Ha: 0.52, m; Hb: 0.35, m	Ha: 2.72, 1.05, 0.35; Hb: 2.72, 1.18, 0.52	Ha: 53.9; both: 40.7, 29.4
13	40.7 CH2	2.72, t (7.1)	0.52, 0.35	156.71, 29.4, 22.3
14	N			
15	156.7			



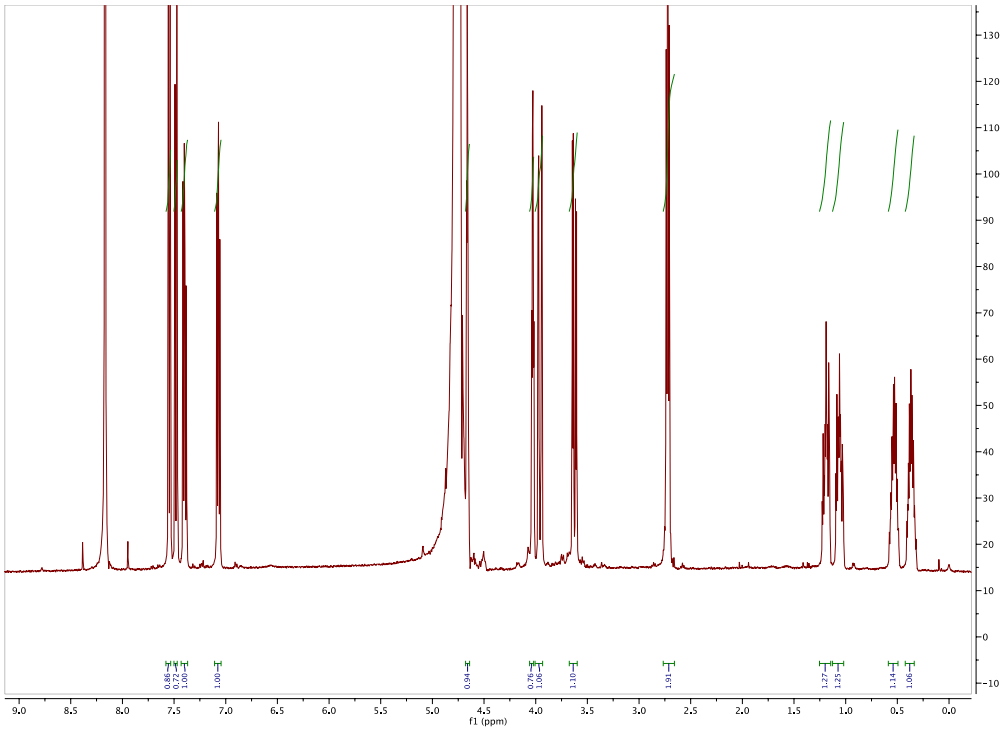


Figure S89. The ^{1}H NMR spectrum of compound **18** in D_2O 0.5% TFA-*d* (500 MHz)

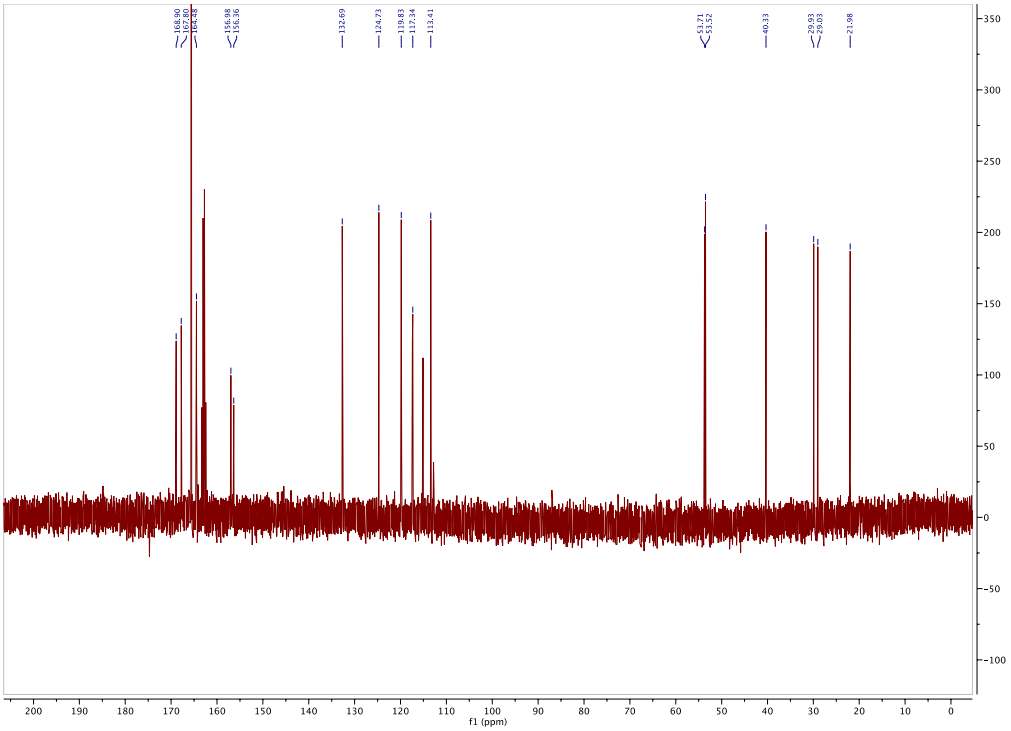


Figure S90. The ^{13}C NMR spectrum of compound 18 in D_2O 0.5% TFA-*d* (500 MHz)

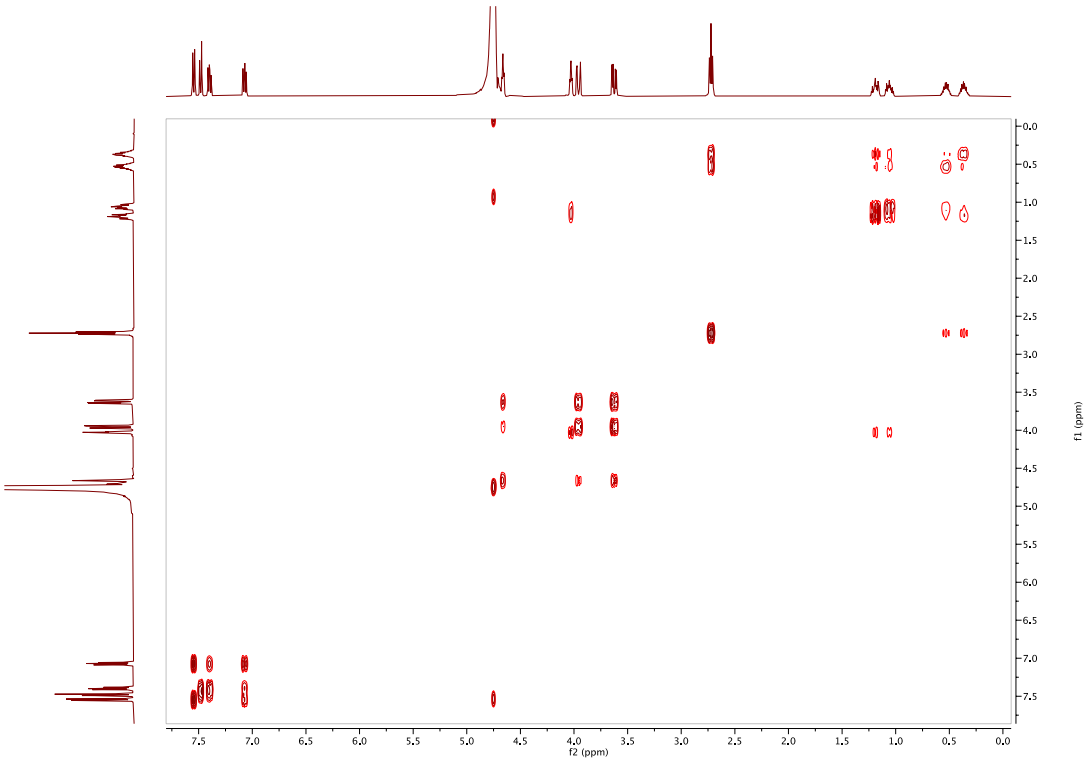


Figure S91. The COSY spectrum of compound **18** in D_2O 0.5% TFA-*d* (500 MHz)

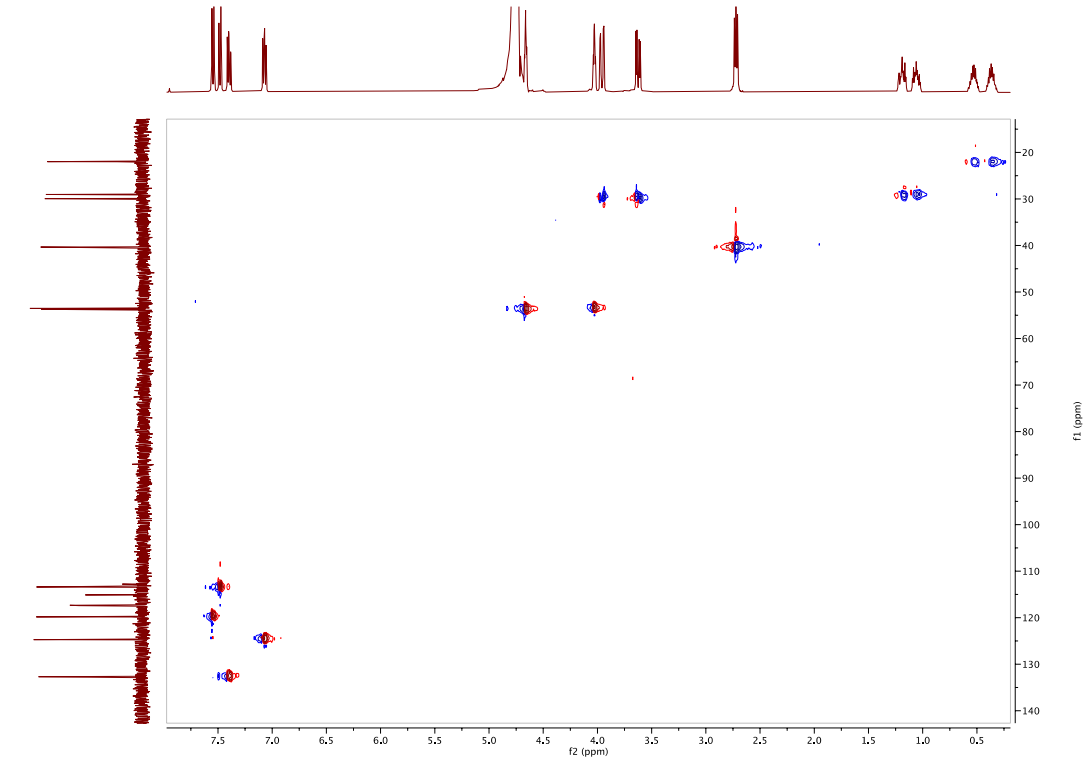


Figure S92. The HSQC spectrum of compound **18** in D_2O 0.5% TFA-*d* (500 MHz)

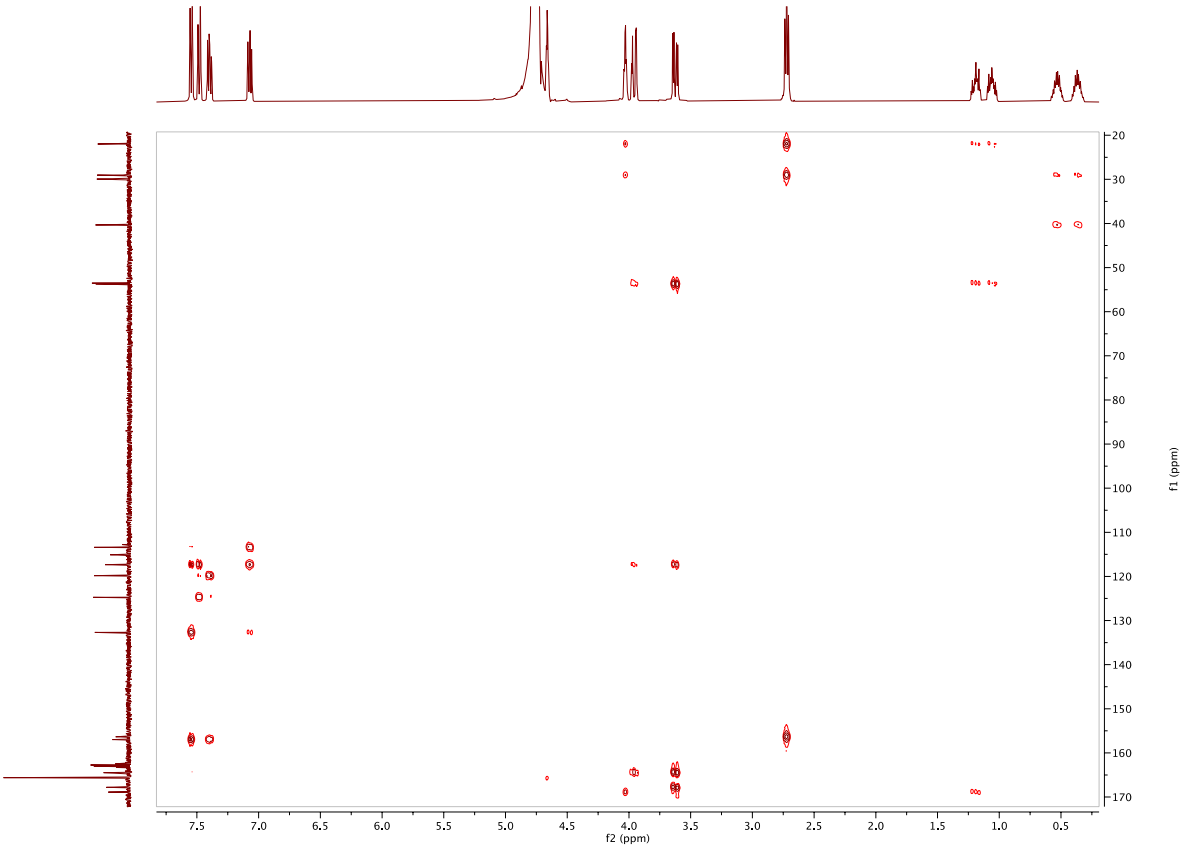
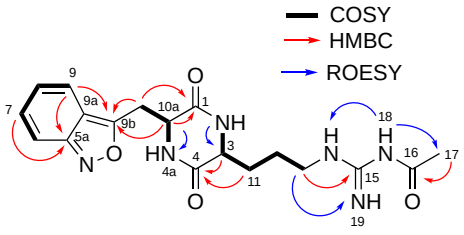


Figure S93. The HMBC spectrum of compound **18** in D_2O 0.5% TFA-*d* (500 MHz)

Table S24: Spectroscopic data of **19** in DMSO-*d6*

Pos.	δC	δH	COSY	ROESY	HMBC
1	166.0, C				
2	N	8.19	3.78	3.78, 1.28, 1.03	166.0
3	53.2, CH	3.78	8.19, 1.28, 1.10	8.19	166.8, 22.5
4	166.8, C				
4a	NH	8.42	4.42	4.42, 3.73, 3.53	166.8
5	N				
5a	156.4, C				
6	114.1, CH	7.49, d	7.31		123.3, 116.6
7	131.3, CH	7.31, t	7.49, 6.98		156.4, 123.3, 120.9
8	123.3, CH	6.98, t	7.63, 7.31		116.6, 114.1
9	120.9, CH	7.63, d	6.98		165.3, 156.4, 131.3, 116.6
9a	116.6, C				
9b	165.3, C				
10	30.2, CH2	Ha: 3.73, dd; Hb: 3.53, dd	4.42		Ha: 166.0, 116.6, 53.3; Hb: 166.0, 165.3, 116.6, 53.3
10a	53.3, CH	4.42	8.42	8.42	166.0, 165.3
11	29.7, CH2	Ha: 1.28; Hb: 1.10	Ha: 3.78, 1.03, 1.00, Hb: 3.78		Hb: 53.3
12	22.5, CH2	Ha: 1.03; Hb: 1.00	Ha: 3.01		Ha: 40.2
13	~40 (under solvent), CH2	3.01	1.03, 8.98		152.9[a]
14	NH	8.98	3.01	11.84, 3.01, 1.03	
15	152.9, C				
16	172.6, C				
17	24.2, CH3	2.15		11.84, 8.98	172.6
18	NH	11.84		2.15	
19	NH	8.66, br		3.01	

[a] HMBC correlation between positions 13 and 15 was detected in D_2O, but carbon peak for position 15 was only detected in DMSO, not D_2O.



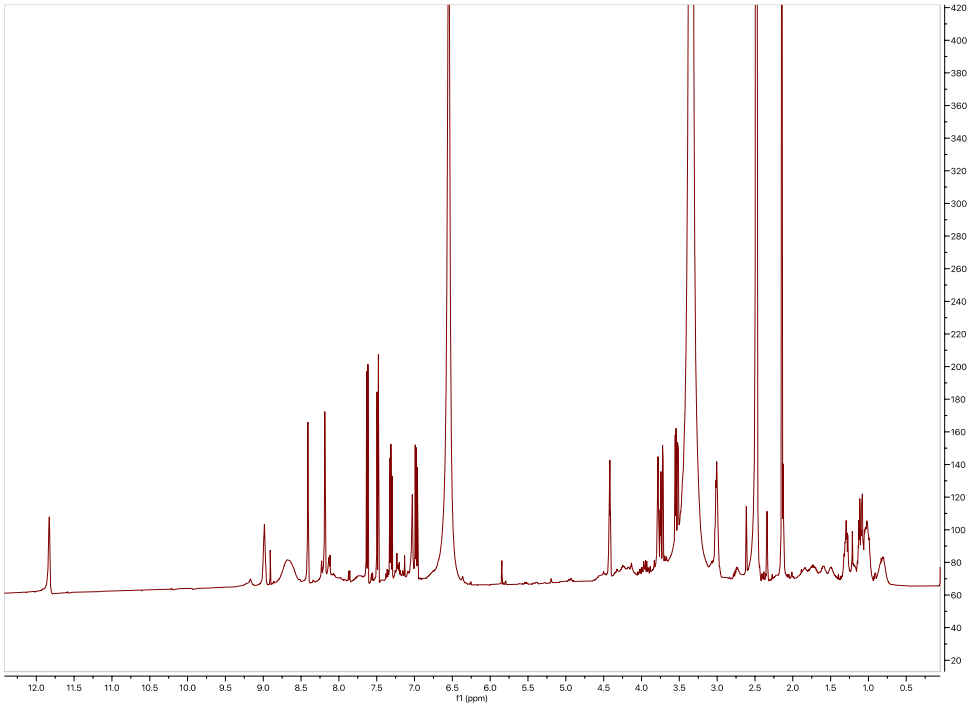


Figure S94. The ^{1}H NMR spectrum of compound **19** in DMSO-*d6* (500 MHz)

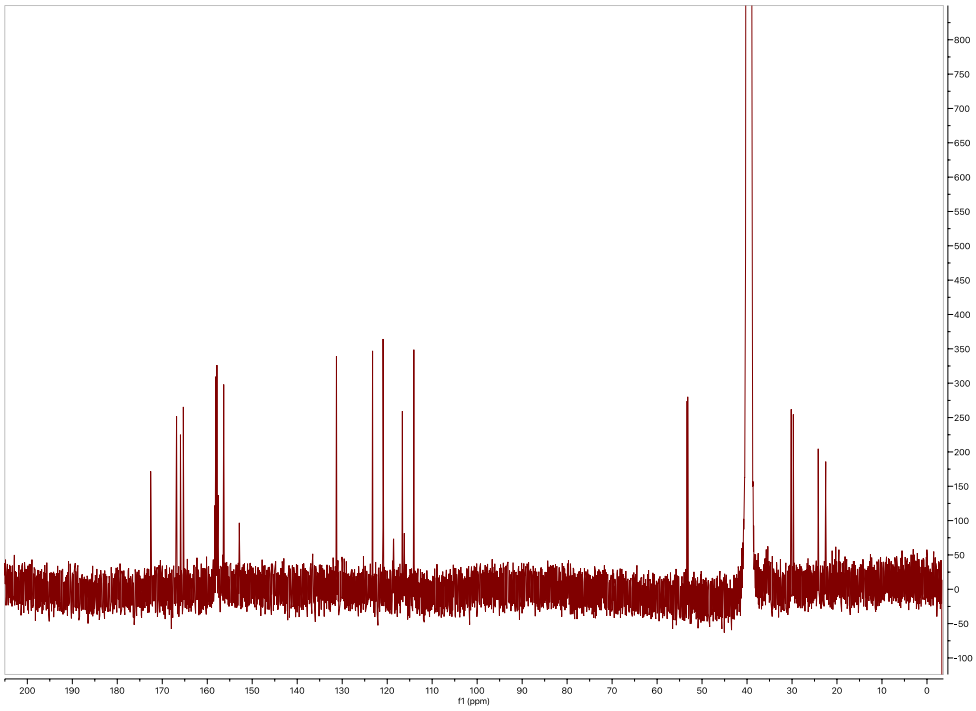


Figure S95. The ^{13}C NMR spectrum of compound **19** in DMSO-*d6* (500 MHz)

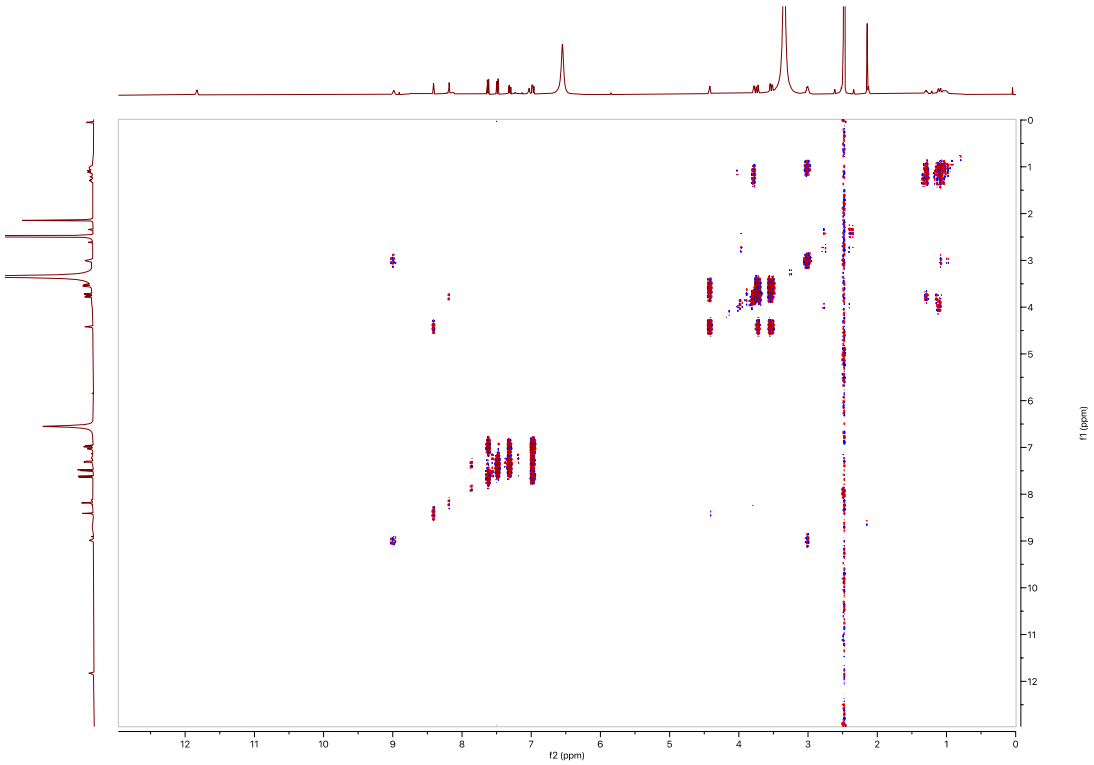


Figure S96. The COSY spectrum of compound **19** in DMSO-*d6* (500 MHz)

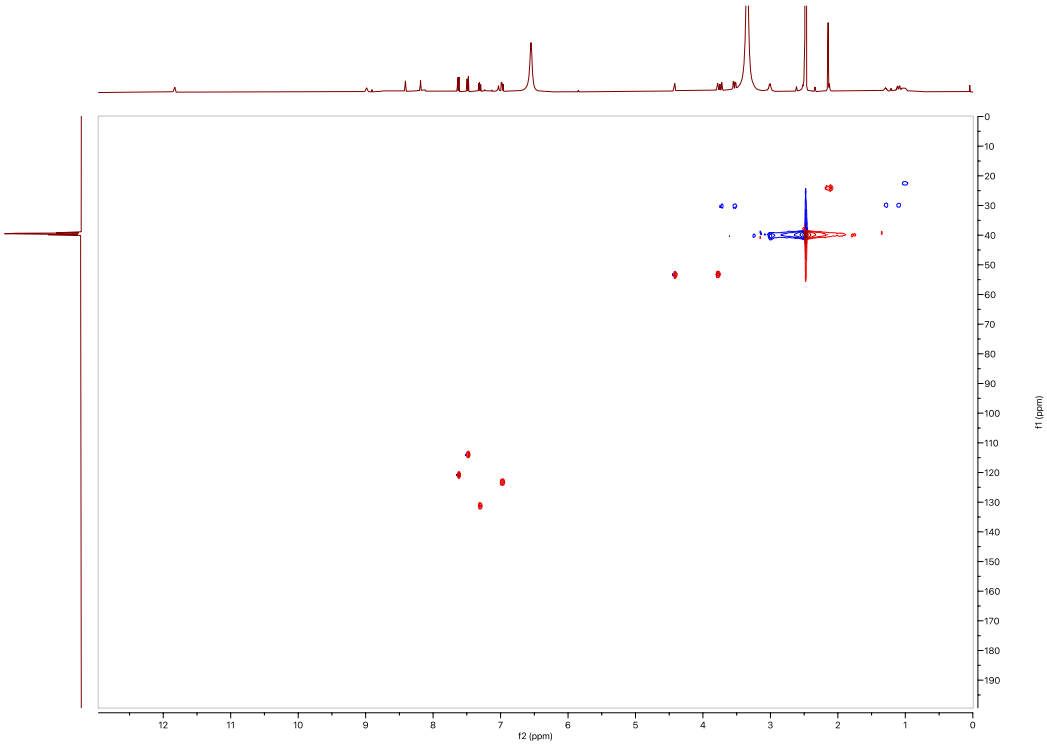


Figure S97. The HSQC spectrum of compound **19** in DMSO-*d6* (500 MHz)

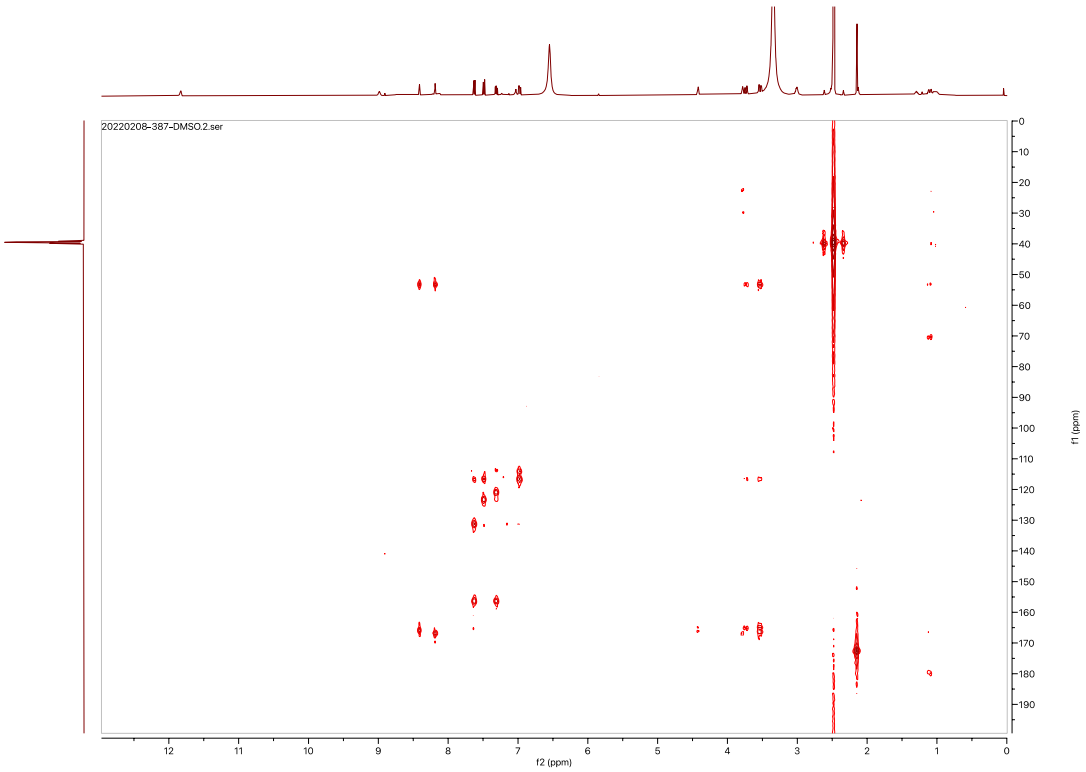


Figure S98. The HMBC spectrum of compound **19** in DMSO-*d6* (500 MHz)

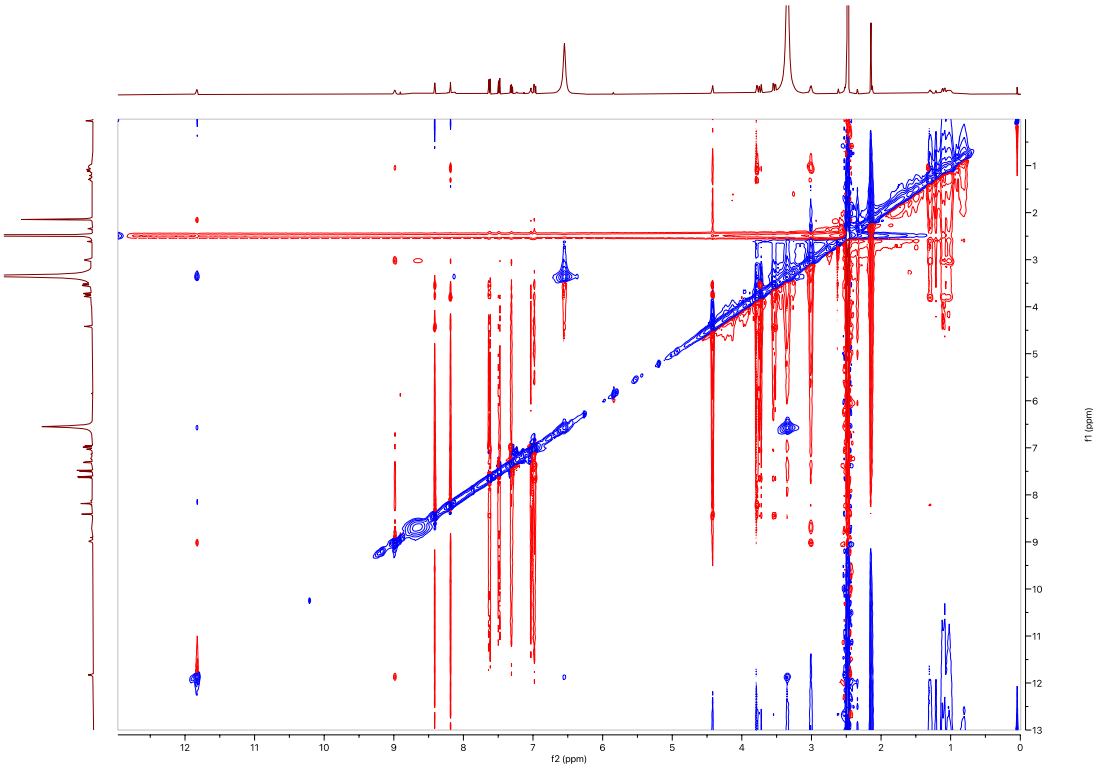


Figure S99. The ROESY spectrum of compound **19** in DMSO-*d6* (500 MHz)

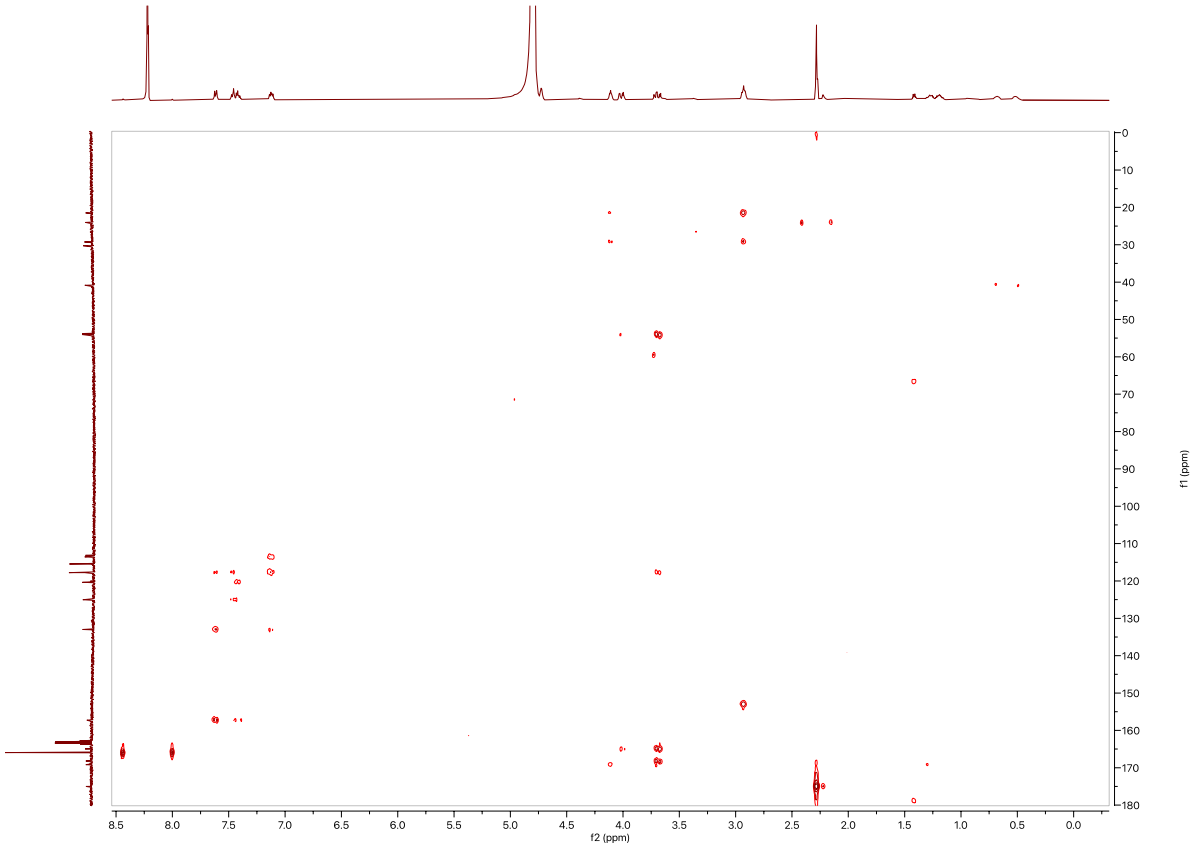


Figure S100. The HMBC spectrum of compound **19** in D_2O (500 MHz)

10. REFERENCES

(1) Gershenzon, J.; Dudareva, N. The Function of Terpene Natural Products in the Natural World. *Nat. Chem. Biol.* **2007**, *3* (7), 408–414. https://doi.org/10.1038/nchembio.2007.5.

(2) Park, S. R.; Yoo, Y. J.; Ban, Y.-H.; Yoon, Y. J. Biosynthesis of Rapamycin and Its Regulation: Past Achievements and Recent Progress. *J. Antibiot. (Tokyo)* **2010**, *63* (8), 434–441. https://doi.org/10.1038/ja.2010.71.

(3) Croteau, R.; Ketchum, R. E. B.; Long, R. M.; Kaspera, R.; Wildung, M. R. Taxol Biosynthesis and Molecular Genetics. *Phytochem. Rev.* **2006**, *5* (1), 75–97. https://doi.org/10.1007/s11101-005-3748-2.

(4) Yan, Y.; Liu, Q.; Zang, X.; Yuan, S.; Bat-Erdene, U.; Nguyen, C.; Gan, J.; Zhou, J.; Jacobsen, S. E.; Tang, Y. Resistance-Gene-Directed Discovery of a Natural-Product Herbicide with a New Mode of Action. *Nature* **2018**, *559* (7714), 415–418. https://doi.org/10.1038/s41586-018-0319-4.

(5) Wen, W.; Yu, R.; Yu, R.; Yu, R. Artemisinin Biosynthesis and Its Regulatory Enzymes: Progress and Perspective. *Pharmacogn. Rev.* **2011**, *5* (10), 189–194. https://doi.org/10.4103/0973-7847.91118.

(6) Chooi, Y.-H.; Tang, Y. Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules. *J. Org. Chem.* **2012**, *77* (22), 9933–9953. https://doi.org/10.1021/jo301592k.

(7) Wang, J.; Zhang, R.; Chen, X.; Sun, X.; Yan, Y.; Shen, X.; Yuan, Q. Biosynthesis of Aromatic Polyketides in Microorganisms Using Type II Polyketide Synthases. *Microb. Cell Factories* **2020**, *19* (1), 110. https://doi.org/10.1186/s12934-020-01367-4.

(8) Abe, I.; Morita, H. Structure and Function of the Chalcone Synthase Superfamily of Plant Type III Polyketide Synthases. *Nat. Prod. Rep.* **2010**, *27* (6), 809–838. https://doi.org/10.1039/B909988N.

(9) Panche, A. N.; Diwan, A. D.; Chandra, S. R. Flavonoids: An Overview. *J. Nutr. Sci.* **2016**, *5*, e47. https://doi.org/10.1017/jns.2016.41.

(10) Le Govic, Y.; Papon, N.; Le Gal, S.; Bouchara, J.-P.; Vandeputte, P. Non-Ribosomal Peptide Synthetase Gene Clusters in the Human Pathogenic Fungus Scedosporium Apiospermum. *Front. Microbiol.* **2019**, *10*.

(11) Finking, R.; Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annu. Rev. Microbiol.* **2004**, *58* (1), 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615.

(12) Christianson, D. W. Structural Biology and Chemistry of the Terpenoid Cyclases. *Chem. Rev.* **2006**, *106* (8), 3412–3442. https://doi.org/10.1021/cr050286w.

(13) Baunach, M.; Franke, J.; Hertweck, C. Terpenoid Biosynthesis Off the Beaten Track: Unconventional Cyclases and Their Impact on Biomimetic Synthesis. *Angew. Chem. Int. Ed.* **2015**, *54* (9), 2604–2626. https://doi.org/10.1002/anie.201407883.

(14) Barra, L.; Abe, I. Chemistry of Fungal Meroterpenoid Cyclases. *Nat. Prod. Rep.* **2021**, *38* (3), 566–585. https://doi.org/10.1039/D0NP00056F.

(15) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; Donk, W. A. van der. Ribosomally Synthesized and Post-Translationally Modified Peptide

Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2012**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(16) Carroll, C. S.; Moore, M. M. Ironing out Siderophore Biosynthesis: A Review of Non-Ribosomal Peptide Synthetase (NRPS)-Independent Siderophore Synthetases. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53* (4), 356–381. https://doi.org/10.1080/10409238.2018.1476449.

(17) Meunier, B.; de Visser, S. P.; Shaik, S. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes. *Chem. Rev.* **2004**, *104* (9), 3947–3980. https://doi.org/10.1021/cr020443g.

(18) F. Rocha, J.; F. Pina, A.; F. Sousa, S.; A. Cerqueira, N. M. F. S. PLP-Dependent Enzymes as Important Biocatalysts for the Pharmaceutical, Chemical and Food Industries: A Structural and Mechanistic Perspective. *Catal. Sci. Technol.* **2019**, *9* (18), 4864–4876. https://doi.org/10.1039/C9CY01210A.

(19) Skinner, S. E.; Walton, N. J.; Robins, R. J.; Rhodes, M. J. C. Tryptophan Decarboxylase Strictosidine Synthase and Alkaloid Production by Cinchona Ledgeriana Suspension Cultures. *Phytochemistry* **1987**, *26* (3), 721–725. https://doi.org/10.1016/S0031-9422(00)84772-X.

(20) Yee, D. A.; Kakule, T. B.; Cheng, W.; Chen, M.; Chong, C. T. Y.; Hai, Y.; Hang, L. F.; Hung, Y.-S.; Liu, N.; Ohashi, M.; Okorafor, I. C.; Song, Y.; Tang, M.; Zhang, Z.; Tang, Y. Genome Mining of Alkaloidal Terpenoids from a Hybrid Terpene and Nonribosomal Peptide Biosynthetic Pathway. *J. Am. Chem. Soc.* **2020**, *142* (2), 710–714. https://doi.org/10.1021/jacs.9b13046.

(21) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79* (3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055.

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

(23) 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. *J. Am. Chem. Soc.* **2013**, *135* (20), 7720–7731. https://doi.org/10.1021/ja401945a.

(24) Lubertozzi, D.; Keasling, J. D. Developing Aspergillus as a Host for Heterologous Expression. *Biotechnol. Adv.* **2009**, *27* (1), 53–75. https://doi.org/10.1016/j.biotechadv.2008.09.001.

(25) A. K. Pahirulzaman, K.; Williams, K.; Lazarus, C. M. Chapter Twelve - A Toolkit for Heterologous Expression of Metabolic Pathways in Aspergillus Oryzae. In *Methods in Enzymology*; Hopwood, D. A., Ed.; Natural Product Biosynthesis by Microorganisms and Plants, Part C; Academic Press, 2012; Vol. 517, pp 241–260. https://doi.org/10.1016/B978-0-12-404634-4.00012-7.

(26) Anyaogu, D. C.; Mortensen, U. H. Heterologous Production of Fungal Secondary Metabolites in Aspergilli. *Front. Microbiol.* **2015**, *6*.

(27) Sakai, K.; Kinoshita, H.; Nihira, T. Heterologous Expression System in Aspergillus Oryzae for Fungal Biosynthetic Gene Clusters of Secondary Metabolites. *Appl. Microbiol. Biotechnol.* **2012**, *93* (5), 2011–2022. https://doi.org/10.1007/s00253-011-3657-9.

(28) van Dijk, J. W. A.; Wang, C. C. C. Chapter Six - Heterologous Expression of Fungal Secondary Metabolite Pathways in the Aspergillus Nidulans Host System. In *Methods in Enzymology*; O'Connor, S. E., Ed.; Synthetic Biology and Metabolic Engineering in Plants and Microbes Part A: Metabolism in Microbes; Academic Press, 2016; Vol. 575, pp 127–142. https://doi.org/10.1016/bs.mie.2016.02.021.

(29) Qiao, Y.-M.; Yu, R.-L.; Zhu, P. Advances in Targeting and Heterologous Expression of Genes Involved in the Synthesis of Fungal Secondary Metabolites. *RSC Adv.* **2019**, *9* (60), 35124–35134. https://doi.org/10.1039/C9RA06908A.

(30) He, Y.; Wang, B.; Chen, W.; Cox, R. J.; He, J.; Chen, F. Recent Advances in Reconstructing Microbial Secondary Metabolites Biosynthesis in Aspergillus Spp. *Biotechnol. Adv.* **2018**, *36* (3), 739–783. https://doi.org/10.1016/j.biotechadv.2018.02.001.

(31) 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), 2294–2298. https://doi.org/10.1002/cbic.201500386.

(32) Yu, J. H.; Leonard, T. J. Sterigmatocystin Biosynthesis in Aspergillus Nidulans Requires a Novel Type I Polyketide Synthase. *J. Bacteriol.* **1995**, *177* (16), 4792–4800. https://doi.org/10.1128/jb.177.16.4792-4800.1995.

(33) 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. C. Molecular Genetic Mining of the Aspergillus Secondary Metabolome: Discovery of the Emericellamide Biosynthetic Pathway. *Chem. Biol.* **2008**, *15* (6), 527–532. https://doi.org/10.1016/j.chembiol.2008.05.010.

(34) Liu, N.; Hung, Y.-S.; Gao, S.-S.; Hang, L.; Zou, Y.; Chooi, Y.-H.; Tang, Y. Identification and Heterologous Production of a Benzoyl-Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid A Biosynthetic Pathway. *Org. Lett.* **2017**, *19* (13), 3560–3563. https://doi.org/10.1021/acs.orglett.7b01534.

(35) 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.; Onge, R. P. S.; Tang, Y.; Hillenmeyer, M. E. HEx: A Heterologous Expression Platform for the Discovery of Fungal Natural Products. *Sci. Adv.* **2018**, 15.

(36) Raeder, U.; Broda, P. Rapid Preparation of DNA from Filamentous Fungi. *Lett. Appl. Microbiol.* **1985**, *1* (1), 17–20. https://doi.org/10.1111/j.1472-765X.1985.tb01479.x.

(37) Goodwin, D. C.; Lee, S. B. Microwave Miniprep of Total Genomic DNA from Fungi, Plants, Protists and Animals for PCR. *BioTechniques* **1993**, *15* (3), 438, 441–442, 444.

(38) Gietz, R. D.; Schiestl, R. H. High-Efficiency Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. *Nat. Protoc.* **2007**, *2* (1), 31–34. https://doi.org/10.1038/nprot.2007.13.

(39) E. O'Connor, S.; J. Maresh, J. Chemistry and Biology of Monoterpene Indole Alkaloid Biosynthesis. *Nat. Prod. Rep.* **2006**, *23* (4), 532–547. https://doi.org/10.1039/B512615K.

(40) Mehrotra, S.; Mishra, S.; Srivastava, V. Hairy Root Cultures for Monoterpene Indole Alkaloid Pathway: Investigation and Biotechnological Production. In *Hairy Roots: An Effective Tool of Plant Biotechnology*; Srivastava, V., Mehrotra, S., Mishra, S., Eds.; Springer: Singapore, 2018; pp 95–121. https://doi.org/10.1007/978-981-13-2562-5_5.

(41) Leonard, E.; Runguphan, W.; O'Connor, S.; Prather, K. J. Opportunities in Metabolic Engineering to Facilitate Scalable Alkaloid Production. *Nat. Chem. Biol.* **2009**, *5* (5), 292–300. https://doi.org/10.1038/nchembio.160.

(42) Miettinen, K.; Dong, L.; Navrot, N.; Schneider, T.; Burlat, V.; Pollier, J.; Woittiez, L.; van der Krol, S.; Lugan, R.; Ilc, T.; Verpoorte, R.; Oksman-Caldentey, K.-M.; Martinoia, E.; Bouwmeester, H.; Goossens, A.; Memelink, J.; Werck-Reichhart, D. The Seco-Iridoid Pathway from Catharanthus Roseus. *Nat. Commun.* **2014**, *5* (1), 3606. https://doi.org/10.1038/ncomms4606.

(43) Brown, S.; Clastre, M.; Courdavault, V.; O'Connor, S. E. De Novo Production of the Plant-Derived Alkaloid Strictosidine in Yeast. *Proc. Natl. Acad. Sci.* **2015**, *112* (11), 3205–3210. https://doi.org/10.1073/pnas.1423555112.

(44) Fischer, M. J. C.; Meyer, S.; Claudel, P.; Perrin, M.; Ginglinger, J. F.; Gertz, C.; Masson, J. E.; Werck-Reinhardt, D.; Hugueney, P.; Karst, F. Specificity of Ocimum Basilicum Geraniol Synthase Modified by Its Expression in Different Heterologous Systems. *J. Biotechnol.* **2013**, *163* (1), 24–29. https://doi.org/10.1016/j.jbiotec.2012.10.012.

(45) Dong, L.; Miettinen, K.; Goedbloed, M.; Verstappen, F. W. A.; Voster, A.; Jongsma, M. A.; Memelink, J.; Krol, S. van der; Bouwmeester, H. J. Characterization of Two Geraniol Synthases from Valeriana Officinalis and Lippia Dulcis: Similar Activity but Difference in Subcellular Localization. *Metab. Eng.* **2013**, *20*, 198–211. https://doi.org/10.1016/j.ymben.2013.09.002.

(46) Collu, G.; Unver, N.; Peltenburg-Looman, A. M. G.; van der Heijden, R.; Verpoorte, R.; Memelink, J. Geraniol 10-Hydroxylase1, a Cytochrome P450 Enzyme Involved in Terpenoid Indole Alkaloid Biosynthesis. *FEBS Lett.* **2001**, *508* (2), 215–220. https://doi.org/10.1016/S0014-5793(01)03045-9.

(47) Geu-Flores, F.; Sherden, N. H.; Courdavault, V.; Burlat, V.; Glenn, W. S.; Wu, C.; Nims, E.; Cui, Y.; O'Connor, S. E. An Alternative Route to Cyclic Terpenes by Reductive Cyclization in Iridoid Biosynthesis. *Nature* **2012**, *492* (7427), 138–142. https://doi.org/10.1038/nature11692.

(48) Kutchan, T. m.; Hampp, N.; Lottspeich, F.; Beyreuther, K.; Zenk, M. h. The CDNA Clone for Strictosidine Synthase from Rauvolfia Serpentina DNA Sequence Determination and Expression in Escherichia Coli. *FEBS Lett.* **1988**, *237* (1–2), 40–44. https://doi.org/10.1016/0014-5793(88)80167-4.

(49) De Luca, V.; Marineau, C.; Brisson, N. Molecular Cloning and Analysis of CDNA Encoding a Plant Tryptophan Decarboxylase: Comparison with Animal Dopa Decarboxylases. *Proc. Natl. Acad. Sci.* **1989**, *86* (8), 2582–2586. https://doi.org/10.1073/pnas.86.8.2582.

(50) Irmler, S.; Schröder, G.; St-Pierre, B.; Crouch, N. P.; Hotze, M.; Schmidt, J.; Strack, D.; Matern, U.; Schröder, J. Indole Alkaloid Biosynthesis in Catharanthus Roseus: New

Enzyme Activities and Identification of Cytochrome P450 CYP72A1 as Secologanin Synthase. *Plant J.* **2000**, *24* (6), 797–804. https://doi.org/10.1111/j.1365-313X.2000.00922.x.

(51) Murata, J.; Roepke, J.; Gordon, H.; De Luca, V. The Leaf Epidermome of Catharanthus Roseus Reveals Its Biochemical Specialization. *Plant Cell* **2008**, *20* (3), 524–542. https://doi.org/10.1105/tpc.107.056630.

(52) Denby, C. M.; Li, R. A.; Vu, V. T.; Costello, Z.; Lin, W.; Chan, L. J. G.; Williams, J.; Donaldson, B.; Bamforth, C. W.; Petzold, C. J.; Scheller, H. V.; Martin, H. G.; Keasling, J. D. Industrial Brewing Yeast Engineered for the Production of Primary Flavor Determinants in Hopped Beer. *Nat. Commun.* **2018**, *9* (1), 965. https://doi.org/10.1038/s41467-018-03293-x.

(53) Liu, J.; Zhang, W.; Du, G.; Chen, J.; Zhou, J. Overproduction of Geraniol by Enhanced Precursor Supply in Saccharomyces Cerevisiae. *J. Biotechnol.* **2013**, *168* (4), 446–451. https://doi.org/10.1016/j.jbiotec.2013.10.017.

(54) Zhao, J.; Bao, X.; Li, C.; Shen, Y.; Hou, J. Improving Monoterpene Geraniol Production through Geranyl Diphosphate Synthesis Regulation in Saccharomyces Cerevisiae. *Appl. Microbiol. Biotechnol.* **2016**, *100* (10), 4561–4571. https://doi.org/10.1007/s00253-016-7375-1.

(55) Fischer, M. J. C.; Meyer, S.; Claudel, P.; Bergdoll, M.; Karst, F. Metabolic Engineering of Monoterpene Synthesis in Yeast. *Biotechnol. Bioeng.* **2011**, *108* (8), 1883–1892. https://doi.org/10.1002/bit.23129.

(56) Zhao, J.; Li, C.; Zhang, Y.; Shen, Y.; Hou, J.; Bao, X. Dynamic Control of ERG20 Expression Combined with Minimized Endogenous Downstream Metabolism Contributes to the Improvement of Geraniol Production in Saccharomyces Cerevisiae. *Microb. Cell Factories* **2017**, *16* (1), 17. https://doi.org/10.1186/s12934-017-0641-9.

(57) Jiang, G.-Z.; Yao, M.-D.; Wang, Y.; Zhou, L.; Song, T.-Q.; Liu, H.; Xiao, W.-H.; Yuan, Y.-J. Manipulation of GES and ERG20 for Geraniol Overproduction in Saccharomyces Cerevisiae. *Metab. Eng.* **2017**, *41*, 57–66. https://doi.org/10.1016/j.ymben.2017.03.005.

(58) Shi, S.; Liang, Y.; Zhang, M. M.; Ang, E. L.; Zhao, H. A Highly Efficient Single-Step, Markerless Strategy for Multi-Copy Chromosomal Integration of Large Biochemical Pathways in Saccharomyces Cerevisiae. *Metab. Eng.* **2016**, *33*, 19–27. https://doi.org/10.1016/j.ymben.2015.10.011.

(59) Strucko, T.; Magdenoska, O.; Mortensen, U. H. Benchmarking Two Commonly Used Saccharomyces Cerevisiae Strains for Heterologous Vanillin-β-Glucoside Production. *Metab. Eng. Commun.* **2015**, *2*, 99–108. https://doi.org/10.1016/j.meteno.2015.09.001.

(60) Matheson, K.; Parsons, L.; Gammie, A. Whole-Genome Sequence and Variant Analysis of W303, a Widely-Used Strain of Saccharomyces Cerevisiae. *G3 GenesGenomesGenetics* **2017**, *7* (7), 2219–2226. https://doi.org/10.1534/g3.117.040022.

(61) Billingsley, J. M.; DeNicola, A. B.; Barber, J. S.; Tang, M.-C.; Horecka, J.; Chu, A.; Garg, N. K.; Tang, Y. Engineering the Biocatalytic Selectivity of Iridoid Production in Saccharomyces Cerevisiae. *Metab. Eng.* **2017**, *44*, 117–125. https://doi.org/10.1016/j.ymben.2017.09.006.

(62) Bao, Z.; Xiao, H.; Liang, J.; Zhang, L.; Xiong, X.; Sun, N.; Si, T.; Zhao, H. Homology-Integrated CRISPR–Cas (HI-CRISPR) System for One-Step Multigene Disruption in Saccharomyces Cerevisiae. *ACS Synth. Biol.* **2015**, *4* (5), 585–594. https://doi.org/10.1021/sb500255k.

(63) Campbell, A.; Bauchart, P.; Gold, N. D.; Zhu, Y.; De Luca, V.; Martin, V. J. J. Engineering of a Nepetalactol-Producing Platform Strain of Saccharomyces Cerevisiae for the Production of Plant Seco-Iridoids. *ACS Synth. Biol.* **2016**, *5* (5), 405–414. https://doi.org/10.1021/acssynbio.5b00289.

(64) Chooi, Y.-H.; Cacho, R.; Tang, Y. Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from Penicillium Aethiopicum. *Chem. Biol.* **2010**, *17* (5), 483–494. https://doi.org/10.1016/j.chembiol.2010.03.015.

(65) Oswald, M.; Fischer, M.; Dirninger, N.; Karst, F. Monoterpenoid Biosynthesis in Saccharomyces Cerevisiae. *FEMS Yeast Res.* **2007**, *7* (3), 413–421. https://doi.org/10.1111/j.1567-1364.2006.00172.x.

(66) Ignea, C.; Pontini, M.; Maffei, M. E.; Makris, A. M.; Kampranis, S. C. Engineering Monoterpene Production in Yeast Using a Synthetic Dominant Negative Geranyl Diphosphate Synthase. *ACS Synth. Biol.* **2014**, *3* (5), 298–306. https://doi.org/10.1021/sb400115e.

(67) Mantzouridou, F.; Tsimidou, M. Z. Observations on Squalene Accumulation in Saccharomyces Cerevisiae Due to the Manipulation of HMG2 and ERG6. *FEMS Yeast Res.* **2010**, *10* (6), 699–707. https://doi.org/10.1111/j.1567-1364.2010.00645.x.

(68) Steyer, D.; Erny, C.; Claudel, P.; Riveill, G.; Karst, F.; Legras, J.-L. Genetic Analysis of Geraniol Metabolism during Fermentation. *Food Microbiol.* **2013**, *33* (2), 228–234. https://doi.org/10.1016/j.fm.2012.09.021.

(69) Jakočiūnas, T.; Bonde, I.; Herrgård, M.; Harrison, S. J.; Kristensen, M.; Pedersen, L. E.; Jensen, M. K.; Keasling, J. D. Multiplex Metabolic Pathway Engineering Using CRISPR/Cas9 in Saccharomyces Cerevisiae. *Metab. Eng.* **2015**, *28*, 213–222. https://doi.org/10.1016/j.ymben.2015.01.008.

(70) Faulkner, A.; Chen, X.; Rush, J.; Horazdovsky, B.; Waechter, C. J.; Carman, G. M.; Sternweis, P. C. The LPP1 and DPP1 Gene Products Account for Most of the Isoprenoid Phosphate Phosphatase Activities InSaccharomyces Cerevisiae *. *J. Biol. Chem.* **1999**, *274* (21), 14831–14837. https://doi.org/10.1074/jbc.274.21.14831.

(71) Tokuhiro, K.; Muramatsu, M.; Ohto, C.; Kawaguchi, T.; Obata, S.; Muramoto, N.; Hirai, M.; Takahashi, H.; Kondo, A.; Sakuradani, E.; Shimizu, S. Overproduction of Geranylgeraniol

by Metabolically Engineered Saccharomyces Cerevisiae. *Appl. Environ. Microbiol.* **2009**, *75* (17), 5536–5543. https://doi.org/10.1128/AEM.00277-09.

(72) Scalcinati, G.; Knuf, C.; Partow, S.; Chen, Y.; Maury, J.; Schalk, M.; Daviet, L.; Nielsen, J.; Siewers, V. Dynamic Control of Gene Expression in Saccharomyces Cerevisiae Engineered for the Production of Plant Sesquitepene α-Santalene in a Fed-Batch Mode. *Metab. Eng.* **2012**, *14* (2), 91–103. https://doi.org/10.1016/j.ymben.2012.01.007.

(73) Galdieri, L.; Zhang, T.; Rogerson, D.; Lleshi, R.; Vancura, A. Protein Acetylation and Acetyl Coenzyme A Metabolism in Budding Yeast. *Eukaryot. Cell* **2014**, *13* (12), 1472–1483. https://doi.org/10.1128/EC.00189-14.

(74) Avalos, J. L.; Fink, G. R.; Stephanopoulos, G. Compartmentalization of Metabolic Pathways in Yeast Mitochondria Improves the Production of Branched-Chain Alcohols. *Nat. Biotechnol.* **2013**, *31* (4), 335–341. https://doi.org/10.1038/nbt.2509.

(75) Lv, X.; Wang, F.; Zhou, P.; Ye, L.; Xie, W.; Xu, H.; Yu, H. Dual Regulation of Cytoplasmic and Mitochondrial Acetyl-CoA Utilization for Improved Isoprene Production in Saccharomyces Cerevisiae. *Nat. Commun.* **2016**, *7* (1), 12851. https://doi.org/10.1038/ncomms12851.

(76) Yuan, J.; Ching, C.-B. Mitochondrial Acetyl-CoA Utilization Pathway for Terpenoid Productions. *Metab. Eng.* **2016**, *38*, 303–309. https://doi.org/10.1016/j.ymben.2016.07.008.

(77) Johnston, M.; Davis, R. W. Sequences That Regulate the Divergent GAL1-GAL10 Promoter in Saccharomyces Cerevisiae. *Mol. Cell. Biol.* **1984**, *4* (8), 1440–1448. https://doi.org/10.1128/mcb.4.8.1440-1448.1984.

(78) West, R. W.; Yocum, R. R.; Ptashne, M. Saccharomyces Cerevisiae GAL1-GAL10 Divergent Promoter Region: Location and Function of the Upstream Activating Sequence UASG. *Mol. Cell. Biol.* **1984**, *4* (11), 2467–2478. https://doi.org/10.1128/mcb.4.11.2467-2478.1984.

(79) Guirimand, G.; Burlat, V.; Oudin, A.; Lanoue, A.; St-Pierre, B.; Courdavault, V. Optimization of the Transient Transformation of Catharanthus Roseus Cells by Particle Bombardment and Its Application to the Subcellular Localization of Hydroxymethylbutenyl 4-Diphosphate Synthase and Geraniol 10-Hydroxylase. *Plant Cell Rep.* **2009**, *28* (8), 1215–1234. https://doi.org/10.1007/s00299-009-0722-2.

(80) Trotter, E. W.; Collinson, E. J.; Dawes, I. W.; Grant, C. M. Old Yellow Enzymes Protect against Acrolein Toxicity in the Yeast Saccharomyces Cerevisiae. *Appl. Environ. Microbiol.* **2006**, *72* (7), 4885–4892. https://doi.org/10.1128/AEM.00526-06.

(81) Golla, U.; Bandi, G.; Tomar, R. S. Molecular Cytotoxicity Mechanisms of Allyl Alcohol (Acrolein) in Budding Yeast. *Chem. Res. Toxicol.* **2015**, *28* (6), 1246–1264. https://doi.org/10.1021/acs.chemrestox.5b00071.

(82) Höfer, R.; Dong, L.; André, F.; Ginglinger, J.-F.; Lugan, R.; Gavira, C.; Grec, S.; Lang, G.; Memelink, J.; Van Der Krol, S.; Bouwmeester, H.; Werck-Reichhart, D. Geraniol Hydroxylase and Hydroxygeraniol Oxidase Activities of the CYP76 Family of Cytochrome P450 Enzymes and Potential for Engineering the Early Steps of the (Seco)Iridoid Pathway. *Metab. Eng.* **2013**, *20*, 221–232. https://doi.org/10.1016/j.ymben.2013.08.001.

(83) Lichman, B. R.; Kamileen, M. O.; Titchiner, G. R.; Saalbach, G.; Stevenson, C. E. M.; Lawson, D. M.; O'Connor, S. E. Uncoupled Activation and Cyclization in Catmint Reductive Terpenoid Biosynthesis. *Nat. Chem. Biol.* **2019**, *15* (1), 71–79. https://doi.org/10.1038/s41589-018-0185-2.

(84) Leppert, G.; McDevitt, R.; Falco, S. C.; Van Dyk, T. K.; Ficke, M. B.; Golin, J. Cloning by Gene Amplification of Two Loci Conferring Multiple Drug Resistance in Saccharomyces. *Genetics* **1990**, *125* (1), 13–20. https://doi.org/10.1093/genetics/125.1.13.

(85) do Valle Matta, M. A.; Jonniaux, J.-L.; Balzi, E.; Goffeau, A.; van den Hazel, B. Novel Target Genes of the Yeast Regulator Pdr1p: A Contribution of the TPO1 Gene in

Resistance to Quinidine and Other Drugs. *Gene* **2001**, *272* (1), 111–119. https://doi.org/10.1016/S0378-1119(01)00558-3.

(86) Kihara, A.; Igarashi, Y. Identification and Characterization of ASaccharomyces Cerevisiae Gene, RSB1, Involved in Sphingoid Long-Chain Base Release *. *J. Biol. Chem.* **2002**, *277* (33), 30048–30054. https://doi.org/10.1074/jbc.M203385200.

(87) Ubersax, J. A.; Woodbury, E. L.; Quang, P. N.; Paraz, M.; Blethrow, J. D.; Shah, K.; Shokat, K. M.; Morgan, D. O. Targets of the Cyclin-Dependent Kinase Cdk1. *Nature* **2003**, *425* (6960), 859–864. https://doi.org/10.1038/nature02062.

(88) Ghosh, A. K.; Ramakrishnan, G.; Rajasekharan, R. YLR099C (ICT1) Encodes a Soluble Acyl-CoA-Dependent Lysophosphatidic Acid Acyltransferase Responsible for Enhanced Phospholipid Synthesis on Organic Solvent Stress in Saccharomyces Cerevisiae*. *J. Biol. Chem.* **2008**, *283* (15), 9768–9775. https://doi.org/10.1074/jbc.M708418200.

(89) Thakur, J. K.; Arthanari, H.; Yang, F.; Pan, S.-J.; Fan, X.; Breger, J.; Frueh, D. P.; Gulshan, K.; Li, D. K.; Mylonakis, E.; Struhl, K.; Moye-Rowley, W. S.; Cormack, B. P.; Wagner, G.; Näär, A. M. A Nuclear Receptor-like Pathway Regulating Multidrug Resistance in Fungi. *Nature* **2008**, *452* (7187), 604–609. https://doi.org/10.1038/nature06836.

(90) 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. *Sci. Adv. 6* (20), eaba0721. https://doi.org/10.1126/sciadv.aba0721.

(91) Vest, K. E.; Leary, S. C.; Winge, D. R.; Cobine, P. A. Copper Import into the Mitochondrial Matrix in Saccharomyces Cerevisiae Is Mediated by Pic2, a Mitochondrial Carrier Family Protein *. *J. Biol. Chem.* **2013**, *288* (33), 23884–23892. https://doi.org/10.1074/jbc.M113.470674.

(92) Todisco, S.; Di Noia, M. A.; Castegna, A.; Lasorsa, F. M.; Paradies, E.; Palmieri, F. The Saccharomyces Cerevisiae Gene YPR011c Encodes a Mitochondrial Transporter of Adenosine 5′-Phosphosulfate and 3′-Phospho-Adenosine 5′-Phosphosulfate. *Biochim. Biophys. Acta BBA - Bioenerg.* **2014**, *1837* (2), 326–334. https://doi.org/10.1016/j.bbabio.2013.11.013.

(93) Lascaris, R.; Bussemaker, H. J.; Boorsma, A.; Piper, M.; van der Spek, H.; Grivell, L.; Blom, J. Hap4p Overexpression in Glucose-Grown Saccharomyces Cerevisiae Induces Cells to Enter a Novel Metabolic State. *Genome Biol.* **2002**, *4* (1), R3. https://doi.org/10.1186/gb-2002-4-1-r3.

(94) Lin, C.-I.; McCarty, R. M.; Liu, H. The Enzymology of Organic Transformations: A Survey of Name Reactions in Biological Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (13), 3446–3489. https://doi.org/10.1002/anie.201603291.

(95) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. Modification of Post-PKS Tailoring Steps through Combinatorial Biosynthesis. *Nat. Prod. Rep.* **2002**, *19* (5), 542–580. https://doi.org/10.1039/B103920M.

(96) Olano, C.; Méndez, C.; Salas, J. A. Post-PKS Tailoring Steps in Natural Product-Producing Actinomycetes from the Perspective of Combinatorial Biosynthesis. *Nat. Prod. Rep.* **2010**, *27* (4), 571–616. https://doi.org/10.1039/B911956F.

(97) Tang, M.-C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. Oxidative Cyclization in Natural Product Biosynthesis. *Chem. Rev.* **2017**, *117* (8), 5226–5333. https://doi.org/10.1021/acs.chemrev.6b00478.

(98) Cochrane, R. V. K.; Vederas, J. C. Highly Selective but Multifunctional Oxygenases in Secondary Metabolism. *Acc. Chem. Res.* **2014**, *47* (10), 3148–3161. https://doi.org/10.1021/ar500242c.

(99) Cox, R. Oxidative Rearrangements during Fungal Biosynthesis. *Nat. Prod. Rep.* **2014**, *31* (10), 1405–1424. https://doi.org/10.1039/C4NP00059E.

(100) Cox, R. J. Polyketides, Proteins and Genes in Fungi: Programmed Nano-Machines Begin to Reveal Their Secrets. *Org. Biomol. Chem.* **2007**, *5* (13), 2010–2026. https://doi.org/10.1039/B704420H.

(101) 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**, *139* (15), 5317–5320. https://doi.org/10.1021/jacs.7b02432.

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

(103) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. Reconstitution of a Fungal Meroterpenoid Biosynthesis Reveals the Involvement of a Novel Family of Terpene Cyclases. *Nat. Chem.* **2010**, *2* (10), 858–864. https://doi.org/10.1038/nchem.764.

(104) Matsuda, Y.; Abe, I. Biosynthesis of Fungal Meroterpenoids. *Nat. Prod. Rep.* **2016**, *33* (1), 26–53. https://doi.org/10.1039/C5NP00090D.

(105) 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. *J. Am. Chem. Soc.* **2013**, *135* (12), 4616–4619. https://doi.org/10.1021/ja312503y.

(106) 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. *Nat. Commun.* **2018**, *9* (1), 1963. https://doi.org/10.1038/s41467-018-04364-9.

(107) Li, S.-M. Prenylated Indole Derivatives from Fungi: Structure Diversity, Biological Activities, Biosynthesis and Chemoenzymatic Synthesis. *Nat. Prod. Rep.* **2010**, *27* (1), 57–78. https://doi.org/10.1039/B909987P.

(108) Ding, Y.; Wet, J. R. de; Cavalcoli, J.; Li, S.; Greshock, T. J.; Miller, K. A.; Finefield, J. M.; Sunderhaus, J. D.; McAfoos, T. J.; Tsukamoto, S.; Williams, R. M.; Sherman, D. H. Genome-Based Characterization of Two Prenylation Steps in the Assembly of the Stephacidin and Notoamide Anticancer Agents in a Marine-Derived Aspergillus Sp. *J. Am. Chem. Soc.* **2010**, *132* (36), 12733–12740. https://doi.org/10.1021/ja1049302.

(109) Grundmann, A.; Kuznetsova, T.; Afiyatullov, S. Sh.; Li, S.-M. FtmPT2, an N-Prenyltransferase from Aspergillus Fumigatus, Catalyses the Last Step in the Biosynthesis of Fumitremorgin B. *ChemBioChem* **2008**, *9* (13), 2059–2063. https://doi.org/10.1002/cbic.200800240.

(110) Augustiniak, H.; Forche, E.; Reichenbach, H.; Wray, V.; Gräfe, U.; Höfle, G. Isolierung Und Strukturaufklärung von Ergokonin A Und B; Zwei Neue Antifungische Sterol-Antibiotika Aus Trichoderma Koningii. *Liebigs Ann. Chem.* **1991**, *1991* (4), 361–366. https://doi.org/10.1002/jlac.199119910163.

(111) Cóbar, O. M. Survey of 2,11-Cyclized Cembranoids from Caribbean Sources. *Nat. Prod. Res.* **2009**, *23* (1), 26–43. https://doi.org/10.1080/14786410701760797.

(112) Lee, C.-F.; Chen, L.-X.; Chiang, C.-Y.; Lai, C.-Y.; Lin, H.-C. The Biosynthesis of Norsesquiterpene Aculenes Requires Three Cytochrome P450 Enzymes to Catalyze a Stepwise Demethylation Process. *Angew. Chem. Int. Ed.* **2019**, *58* (51), 18414–18418. https://doi.org/10.1002/anie.201910200.

(113) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. Structure of Trichodiene Synthase from Fusarium Sporotrichioides Provides Mechanistic Inferences on the Terpene Cyclization Cascade. *Proc. Natl. Acad. Sci.* **2001**, *98* (24), 13543–13548. https://doi.org/10.1073/pnas.231313098.

(114) Chiba, R.; Minami, A.; Gomi, K.; Oikawa, H. Identification of Ophiobolin F Synthase by a Genome Mining Approach: A Sesterterpene Synthase from Aspergillus Clavatus. *Org. Lett.* **2013**, *15* (3), 594–597. https://doi.org/10.1021/ol303408a.

(115) Nagasawa, T.; Kanzaki, H.; Yamada, H. Cystathionine Gamma-Lyase of Streptomyces Phaeochromogenes. The Occurrence of Cystathionine Gamma-Lyase in Filamentous Bacteria and Its Purification and Characterization. *J. Biol. Chem.* **1984**, *259* (16), 10393–10403. https://doi.org/10.1016/S0021-9258(18)90978-6.

(116) Wei, Y.; Perez, L. J.; Ng, W.-L.; Semmelhack, M. F.; Bassler, B. L. Mechanism of Vibrio Cholerae Autoinducer-1 Biosynthesis. *ACS Chem. Biol.* **2011**, *6* (4), 356–365. https://doi.org/10.1021/cb1003652.

(117) Liang, J.; Han, Q.; Tan, Y.; Ding, H.; Li, J. Current Advances on Structure-Function Relationships of Pyridoxal 5′-Phosphate-Dependent Enzymes. *Front. Mol. Biosci.* **2019**, *6*.

(118) Gao, S.-S.; Naowarojna, N.; Cheng, R.; Liu, X.; Liu, P. Recent Examples of α-Ketoglutarate-Dependent Mononuclear Non-Haem Iron Enzymes in Natural Product Biosyntheses. *Nat. Prod. Rep.* **2018**, *35* (8), 792–837. https://doi.org/10.1039/C7NP00067G.

(119) Herr, C. Q.; Hausinger, R. P. Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends Biochem. Sci.* **2018**, *43* (7), 517–532. https://doi.org/10.1016/j.tibs.2018.04.002.

(120) Islam, Md. S.; Leissing, T. M.; Chowdhury, R.; Hopkinson, R. J.; Schofield, C. J. 2-Oxoglutarate-Dependent Oxygenases. *Annu. Rev. Biochem.* **2018**, *87* (1), 585–620. https://doi.org/10.1146/annurev-biochem-061516-044724.

(121) Qian, C.; Zhou, M.-M. SET Domain Protein Lysine Methyltransferases: Structure, Specificity and Catalysis. *Cell. Mol. Life Sci. CMLS* **2006**, *63* (23), 2755–2763. https://doi.org/10.1007/s00018-006-6274-5.

(122) Yanagisawa, M.; Sakal, A.; Adachi, K.; Sano, T.; Watanabe, K.; Tanaka, Y.; Okuda, T. HISPIDOSPERMIDIN, A NOVEL PHOSPHOLIPASE C INHIBITOR PRODUCED BY Chaetosphaeronema Hispidulum (Cda) Moesz NR 7127 I. SCREENING, TAXONOMY,

AND FERMENTATION. *J. Antibiot. (Tokyo)* **1994**, *47* (1), 1–5. https://doi.org/10.7164/antibiotics.47.1.

(123) Ohtsuka, T.; Itezono, Y.; Nakayama, N.; Sakai, A.; Shimma, N.; Yokose, K.; Seto, H. HISPIDOSPERMIDIN, A NOVEL PHOSPHOLIPASE C INHIBITOR PRODUCED BY Chaetosphaeronema Hispidulum (Cda) Moesz NR 7127 II. ISOLATION, CHARACTERIZATION AND STRUCTURAL ELUCIDATION. *J. Antibiot. (Tokyo)* **1994**, *47* (1), 6–15. https://doi.org/10.7164/antibiotics.47.6.

(124) Frontier, A. J.; Raghavan, S.; Danishefsky, S. J. Stereocontrolled Total Synthesis of Hispidospermidin. *J. Am. Chem. Soc.* **1997**, *119* (28), 6686–6687. https://doi.org/10.1021/ja970889s.

(125) Frontier, A. J.; Raghavan, S.; Danishefsky, S. J. A Highly Stereoselective Total Synthesis of Hispidospermidin: Derivation of a Pharmacophore Model. *J. Am. Chem. Soc.* **2000**, *122* (26), 6151–6159. https://doi.org/10.1021/ja9944960.

(126) Overman, L. E.; Tomasi, A. L. Enantioselective Total Synthesis of Hispidospermidin. *J. Am. Chem. Soc.* **1998**, *120* (16), 4039–4040. https://doi.org/10.1021/ja974361z.

(127) Tamiya, J.; Sorensen, E. J. A Concise Synthesis of (−)-Hispidospermidin Guided by a Postulated Biogenesis. *J. Am. Chem. Soc.* **2000**, *122* (39), 9556–9557. https://doi.org/10.1021/ja0026758.

(128) Tamiya, J.; Sorensen, E. J. A Spontaneous Bicyclization Facilitates a Synthesis of (−)-Hispidospermidin. *Tetrahedron* **2003**, *59* (35), 6921–6932. https://doi.org/10.1016/S0040-4020(03)00936-0.

(129) Hartmann, M.; Kim, D.; Bernsdorff, F.; Ajami-Rashidi, Z.; Scholten, N.; Schreiber, S.; Zeier, T.; Schuck, S.; Reichel-Deland, V.; Zeier, J. Biochemical Principles and Functional Aspects of Pipecolic Acid Biosynthesis in Plant Immunity. *Plant Physiol.* **2017**, *174* (1), 124–153. https://doi.org/10.1104/pp.17.00222.

(130) Huang, J.-L.; Tang, Y.; Yu, C.-P.; Sanyal, D.; Jia, X.; Liu, X.; Guo, Y.; Chang, W. Mechanistic Investigation of Oxidative Decarboxylation Catalyzed by Two Iron(II)- and 2-Oxoglutarate-Dependent Enzymes. *Biochemistry* **2018**, *57* (12), 1838–1841. https://doi.org/10.1021/acs.biochem.8b00115.

(131) Yu, C.-P.; Tang, Y.; Cha, L.; Milikisiyants, S.; Smirnova, T. I.; Smirnov, A. I.; Guo, Y.; Chang, W. Elucidating the Reaction Pathway of Decarboxylation-Assisted Olefination Catalyzed by a Mononuclear Non-Heme Iron Enzyme. *J. Am. Chem. Soc.* **2018**, *140* (45), 15190–15193. https://doi.org/10.1021/jacs.8b10077.

(132) van der Donk, W. A.; Krebs, C.; Bollinger, J. M. Substrate Activation by Iron Superoxo Intermediates. *Curr. Opin. Struct. Biol.* **2010**, *20* (6), 673–683. https://doi.org/10.1016/j.sbi.2010.08.005.

(133) Phelan, R. M.; Townsend, C. A. Mechanistic Insights into the Bifunctional Non-Heme Iron Oxygenase Carbapenem Synthase by Active Site Saturation Mutagenesis. *J. Am. Chem. Soc.* **2013**, *135* (20), 7496–7502. https://doi.org/10.1021/ja311078s.

(134) Kurosawa, S.; Bando, M.; Mori, K. Synthesis of (1R,4R,5S)-(+)-Acoradiene, the Structure Proposed for the Aggregation Pheromone of the Broad-Horned Flour Beetle. *Eur. J. Org. Chem.* **2001**, *2001* (23), 4395–4399. https://doi.org/10.1002/1099-0690(200112)2001:23<4395::AID-EJOC4395>3.0.CO;2-Q.

(135) Tomita, B.; Hirose, Y. Allo-Cedrol: A New Tricarbocyclic Sesquiterpene Alcohol. *Phytochemistry* **1973**, *12* (6), 1409–1414. https://doi.org/10.1016/0031-9422(73)80575-8.

(136) Shuler, W. G.; Combee, L. A.; Falk, I. D.; Hilinski, M. K. Intermolecular Electrophilic Addition of Epoxides to Alkenes: [3+2] Cycloadditions Catalyzed by Lewis Acids. *Eur. J. Org. Chem.* **2016**, *2016* (20), 3335–3338. https://doi.org/10.1002/ejoc.201600651.

(137) McGee, T. P.; Skinner, H. B.; Bankaitis, V. A. Functional Redundancy of CDP-Ethanolamine and CDP-Choline Pathway Enzymes in Phospholipid Biosynthesis: Ethanolamine-Dependent Effects on Steady-State Membrane Phospholipid Composition

in Saccharomyces Cerevisiae. *J. Bacteriol.* **1994**, *176* (22), 6861–6868. https://doi.org/10.1128/jb.176.22.6861-6868.1994.

(138) Fillinger, S.; Ruijter, G.; Tamás, M. J.; Visser, J.; Thevelein, J. M.; D'enfert, C. Molecular and Physiological Characterization of the NAD-Dependent Glycerol 3-Phosphate Dehydrogenase in the Filamentous Fungus Aspergillus Nidulans. *Mol. Microbiol.* **2001**, *39* (1), 145–157. https://doi.org/10.1046/j.1365-2958.2001.02223.x.

(139) Dijkema, C.; Rijcken, R. P.; Kester, H. C. M.; Visser, J. 13C-NMR Studies on the Influence of PH and Nitrogen Source on Polyol Pool Formation in Aspergillus Nidulans. *FEMS Microbiol. Lett.* **1986**, *33* (1), 125–131. https://doi.org/10.1111/j.1574-6968.1986.tb01225.x.

(140) Brunson, J. K.; McKinnie, S. M. K.; Chekan, J. R.; McCrow, J. P.; Miles, Z. D.; Bertrand, E. M.; Bielinski, V. A.; Luhavaya, H.; Oborník, M.; Smith, G. J.; Hutchins, D. A.; Allen, A. E.; Moore, B. S. Biosynthesis of the Neurotoxin Domoic Acid in a Bloom-Forming Diatom. *Science* **2018**, *361* (6409), 1356–1358. https://doi.org/10.1126/science.aau0382.

(141) Newman, D. J.; Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75* (3), 311–335. https://doi.org/10.1021/np200906s.

(142) Zhong, Z.; He, B.; Li, J.; Li, Y.-X. Challenges and Advances in Genome Mining of Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs). *Synth. Syst. Biotechnol.* **2020**, *5* (3), 155–172. https://doi.org/10.1016/j.synbio.2020.06.002.

(143) Barra, L.; Awakawa, T.; Shirai, K.; Hu, Z.; Bashiri, G.; Abe, I. β-NAD as a Building Block in Natural Product Biosynthesis. *Nature* **2021**, *600* (7890), 754–758. https://doi.org/10.1038/s41586-021-04214-7.

(144) Tohyama, S.; Tomura, A.; Ikeda, N.; Hatano, M.; Odanaka, J.; Kubota, Y.; Umekita, M.; Igarashi, M.; Sawa, R.; Morino, T. Discovery and Characterization of NK13650s, Naturally

Occurring P300-Selective Histone Acetyltransferase Inhibitors. *J. Org. Chem.* **2012**, *77* (20), 9044–9052. https://doi.org/10.1021/jo301534b.

(145) Bartnik, E.; Weglenski, P. Regulation of Arginine Catabolism in Aspergillus Nidulans. *Nature* **1974**, *250* (5467), 590–592. https://doi.org/10.1038/250590a0.

(146) Gondry, M.; Sauguet, L.; Belin, P.; Thai, R.; Amouroux, R.; Tellier, C.; Tuphile, K.; Jacquet, M.; Braud, S.; Courçon, M.; Masson, C.; Dubois, S.; Lautru, S.; Lecoq, A.; Hashimoto, S.; Genet, R.; Pernodet, J.-L. Cyclodipeptide Synthases Are a Family of TRNA-Dependent Peptide Bond–Forming Enzymes. *Nat. Chem. Biol.* **2009**, *5* (6), 414–420. https://doi.org/10.1038/nchembio.175.

(147) Tsukamoto, S.; Kato, H.; Hirota, H.; Fusetani, N. Pipecolate Derivatives, Anthosamines A and B, Inducers of Larval Metamorphosis in Ascidians, from a Marine Sponge Anthosigmella Aff. Raromicrosclera. *Tetrahedron* **1995**, *51* (24), 6687–6694. https://doi.org/10.1016/0040-4020(95)00322-Y.

(148) Li, H.; Lacey, A. E.; Shu, S.; Kalaitzis, J. A.; Vuong, D.; Crombie, A.; Hu, J.; Gilchrist, C. L. M.; Lacey, E.; Piggott, A. M.; Chooi, Y.-H. Hancockiamides: Phenylpropanoid Piperazines from Aspergillus Hancockii Are Biosynthesised by a Versatile Dual Single-Module NRPS Pathway. *Org. Biomol. Chem.* **2021**, *19* (3), 587–595. https://doi.org/10.1039/D0OB02243H.

(149) Chen, M.; Liu, C.-T.; Tang, Y. Discovery and Biocatalytic Application of a PLP-Dependent Amino Acid γ-Substitution Enzyme That Catalyzes C–C Bond Formation. *J. Am. Chem. Soc.* **2020**, *142* (23), 10506–10515. https://doi.org/10.1021/jacs.0c03535.

(150) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858. https://doi.org/10.1038/nprot.2015.053.

(151) Yamaguchi, H.; Kato, H.; Hata, Y.; Nishioka, T.; Kimura, A.; Oda, J.; Katsube, Y. Three-Dimensional Structure of the Glutathione Synthetase from Escherichia Coli B at 2·0 Å

Resolution. *J. Mol. Biol.* **1993**, *229* (4), 1083–1100. https://doi.org/10.1006/jmbi.1993.1106.

(152) Katoh, K.; Rozewicki, J.; Yamada, K. D. MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization. *Brief. Bioinform.* **2019**, *20* (4), 1160–1166. https://doi.org/10.1093/bib/bbx108.

(153) Izumida, H.; Imamura, N.; Sano, H. A Novel Chitinase Inhibitor from a Marine Bacterium, Pseudomonas Sp. *J. Antibiot. (Tokyo)* **1996**, *49* (1), 76–80. https://doi.org/10.7164/antibiotics.49.76.

(154) Li, X.; Hopmann, K. H.; Hudecová, J.; Isaksson, J.; Novotná, J.; Stensen, W.; Andrushchenko, V.; Urbanová, M.; Svendsen, J.-S.; Bouř, P.; Ruud, K. Determination of Absolute Configuration and Conformation of a Cyclic Dipeptide by NMR and Chiral Spectroscopic Methods. *J. Phys. Chem. A* **2013**, *117* (8), 1721–1736. https://doi.org/10.1021/jp311151h.

(155) Gondry, M.; Jacques, I. B.; Thai, R.; Babin, M.; Canu, N.; Seguin, J.; Belin, P.; Pernodet, J.-L.; Moutiez, M. A Comprehensive Overview of the Cyclodipeptide Synthase Family Enriched with the Characterization of 32 New Enzymes. *Front. Microbiol.* **2018**, *9*.

(156) Canu, N.; Moutiez, M.; Belin, P.; Gondry, M. Cyclodipeptide Synthases: A Promising Biotechnological Tool for the Synthesis of Diverse 2,5-Diketopiperazines. *Nat. Prod. Rep.* **2020**, *37* (3), 312–321. https://doi.org/10.1039/C9NP00036D.

(157) Ibba, M.; Söll, D. Aminoacyl-TRNA Synthesis. *Annu. Rev. Biochem.* **2000**, *69* (1), 617–650. https://doi.org/10.1146/annurev.biochem.69.1.617.

(158) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the Genetic Code of Escherichia Coli. *Science* **2001**, *292* (5516), 498–500. https://doi.org/10.1126/science.1060077.

(159) Oldach, F.; Al Toma, R.; Kuthning, A.; Caetano, T.; Mendo, S.; Budisa, N.; Süssmuth, R. D. Congeneric Lantibiotics from Ribosomal In Vivo Peptide Synthesis with Noncanonical

Amino Acids. *Angew. Chem. Int. Ed.* **2012**, *51* (2), 415–418. https://doi.org/10.1002/anie.201106154.

(160) Chatterjee, A.; Xiao, H.; Yang, P.-Y.; Soundararajan, G.; Schultz, P. G. A Tryptophanyl-TRNA Synthetase/TRNA Pair for Unnatural Amino Acid Mutagenesis in E. Coli. *Angew. Chem. Int. Ed.* **2013**, *52* (19), 5106–5109. https://doi.org/10.1002/anie.201301094.

(161) Moutiez, M.; Schmitt, E.; Seguin, J.; Thai, R.; Favry, E.; Belin, P.; Mechulam, Y.; Gondry, M. Unravelling the Mechanism of Non-Ribosomal Peptide Synthesis by Cyclodipeptide Synthases. *Nat. Commun.* **2014**, *5* (1), 5141. https://doi.org/10.1038/ncomms6141.

(162) Ye, Y.; Du, L.; Zhang, X.; Newmister, S. A.; McCauley, M.; Alegre-Requena, J. V.; Zhang, W.; Mu, S.; Minami, A.; Fraley, A. E.; Adrover-Castellano, M. L.; Carney, N. A.; Shende, V. V.; Qi, F.; Oikawa, H.; Kato, H.; Tsukamoto, S.; Paton, R. S.; Williams, R. M.; Sherman, D. H.; Li, S. Fungal-Derived Brevianamide Assembly by a Stereoselective Semipinacolase. *Nat. Catal.* **2020**, *3* (6), 497–506. https://doi.org/10.1038/s41929-020-0454-9.

(163) Li, S.; Finefield, J. M.; Sunderhaus, J. D.; McAfoos, T. J.; Williams, R. M.; Sherman, D. H. Biochemical Characterization of NotB as an FAD-Dependent Oxidase in the Biosynthesis of Notoamide Indole Alkaloids. *J. Am. Chem. Soc.* **2012**, *134* (2), 788–791. https://doi.org/10.1021/ja2093212.

(164) Arai, N.; Shiomi, K.; Yamaguchi, Y.; Masuma, R.; Iwai, Y.; Turberg, A.; Kolbl, H.; Omura, S. Argadin, a New Chitinase Inhibitor, Produced by Clonostachys Sp.FO-7314. *Chem. Pharm. Bull. (Tokyo)* **2000**, *48* (10), 1442–1446. https://doi.org/10.1248/cpb.48.1442.

(165) Houard, J.; Aumelas, A.; Noël, T.; Pages, S.; Givaudan, A.; Fitton-Ouhabi, V.; Villain-Guillot, P.; Gualtieri, M. Cabanillasin, a New Antifungal Metabolite, Produced by Entomopathogenic Xenorhabdus Cabanillasii JM26. *J. Antibiot. (Tokyo)* **2013**, *66* (10), 617–620. https://doi.org/10.1038/ja.2013.58.

(166) 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**, *138* (49), 15837–15840. https://doi.org/10.1021/jacs.6b10452.

(167) Cacho, R. A.; Tang, Y. Reconstitution of Fungal Nonribosomal Peptide Synthetases in Yeast and In Vitro. In *Nonribosomal Peptide and Polyketide Biosynthesis: Methods and Protocols*; Evans, B. S., Ed.; Methods in Molecular Biology; Springer: New York, NY, 2016; pp 103–119. https://doi.org/10.1007/978-1-4939-3375-4_7.

(168) Hang, L.; Tang, M.-C.; Harvey, C. J. B.; Page, C. G.; Li, J.; Hung, Y.-S.; Liu, N.; Hillenmeyer, M. E.; Tang, Y. Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain. *Angew. Chem.* **2017**, *129* (32), 9684–9688. https://doi.org/10.1002/ange.201705237.

(169) Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y. SAM-Dependent Enzyme-Catalysed Pericyclic Reactions in Natural Product Biosynthesis. *Nature* **2017**, *549* (7673), 502–506. https://doi.org/10.1038/nature23882.

(170) Zou, Y.; Garcia-Borràs, M.; Tang, M. C.; Hirayama, Y.; Li, D. H.; Li, L.; Watanabe, K.; Houk, K. N.; Tang, Y. Enzyme-Catalyzed Cationic Epoxide Rearrangements in Quinolone Alkaloid Biosynthesis. *Nat. Chem. Biol.* **2017**, *13* (3), 325–332. https://doi.org/10.1038/nchembio.2283.

(171) Collart, M. A.; Oliviero, S. Preparation of Yeast RNA. *Curr. Protoc. Mol. Biol.* **1993**, *23* (1), 13.12.1-13.12.5. https://doi.org/10.1002/0471142727.mb1312s23.

(172) Janssen, B. D.; Diner, E. J.; Hayes, C. S. Analysis of Aminoacyl- and Peptidyl-TRNAs by Gel Electrophoresis. *Methods Mol. Biol. Clifton NJ* **2012**, *905*, 291–309. https://doi.org/10.1007/978-1-61779-949-5_19.

(173) Petrov, A.; Tsa, A.; Puglisi, J. D. Chapter Sixteen - Analysis of RNA by Analytical Polyacrylamide Gel Electrophoresis. In *Methods in Enzymology*; Lorsch, J., Ed.; Laboratory Methods in Enzymology: RNA; Academic Press, 2013; Vol. 530, pp 301–313. https://doi.org/10.1016/B978-0-12-420037-1.00016-6.

(174) Bond, C. M.; Tang, Y. Engineering Saccharomyces Cerevisiae for Production of Simvastatin. *Metab. Eng.* **2019**, *51*, 1–8. https://doi.org/10.1016/j.ymben.2018.09.005.