

**Investigation on the biosynthesis of polyketide products in
Aspergillus ustus and cyclodipeptide derivatives in
Streptomyces strains**

**Untersuchung zur Biosynthese von Polyketid-Produkten in
Aspergillus ustus und Cyclodipeptid-Derivaten in
Streptomyces-Stämmen**

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Pharmazie
der Philipps-Universität Marburg

vorgelegt von

Yiling Yang
aus Yuxi, China

Marburg an der Lahn, 2023

Erstgutachter: **Prof. Dr. Shu-Ming Li**

Zweitgutachter: **Prof. Dr. Michael Keusgen**

Eingereicht am 14. Feb. 2023

Tag der mündlichen Prüfung: 28. March 2023

Hochschulkennziffer: 1180

Dedicated to my family

Table of contents

| | |
|--|-----|
| List of publications..... | III |
| Abbreviations..... | VII |
| Summary..... | 1 |
| Zusammenfassung..... | 3 |
| 1 Introduction..... | 5 |
| 1.1 Natural products..... | 5 |
| 1.2 Strategy for NP-based drug discovery..... | 6 |
| 1.3 Categories of natural products..... | 7 |
| 1.3.1 Polyketides..... | 7 |
| 1.3.2 Peptides: 2,5-Diketopiperazines..... | 8 |
| 1.4 Biosynthesis of NPs..... | 9 |
| 1.4.1 Backbone enzymes of NPs..... | 9 |
| 1.4.2 Tailoring enzymes..... | 14 |
| 2 Aims of this thesis..... | 21 |
| 3 Results and discussion..... | 23 |
| 3.1 Biosynthesis of ustethylins in <i>Aspergillus ustus</i> | 23 |
| 3.2 Expanding structural diversity of prenylated CDPs and elucidation of the streptoazine biosynthetic pathway in <i>Streptomyces aurantiacus</i> | 28 |
| 3.3 Biosynthesis of various C2-guaninylated guanitrypmycin analogs by a <i>Streptomyces</i> cytochrome P450 enzyme..... | 33 |
| 4 Publications..... | 39 |
| 4.1 Ustethylin biosynthesis implies phenethyl derivative formation in <i>Aspergillus ustus</i> ... | 40 |
| 4.2 Elucidation of the streptoazine biosynthetic pathway in <i>Streptomyces aurantiacus</i> reveals the presence of a promiscuous prenyltransferase/cyclase..... | 104 |
| 4.3 A <i>Streptomyces</i> cytochrome P450 enzyme catalyzes regiospecific C2-guaninylation for the synthesis of diverse guanitrypmycin analogs..... | 157 |
| 5 Conclusions and future prospects..... | 191 |
| 6 References..... | 193 |

TABLE OF CONTENTS

Statutory Declaration..... 201
Acknowledgements..... 203
Curriculum Vitae..... 205

List of publications

1. Liujuan Zheng*, **Yiling Yang***, Haowen Wang, Aili Fan, Liping Zhang, and Shu-Ming Li (2020). Ustethylin biosynthesis implies phenethyl derivative formation in *Aspergillus ustus*. *Organic Letters*, 22, 7837-7841, DOI:10.1021/acs.orglett.0c02719. (*equal contribution)
2. Jing Liu*, **Yiling Yang***, Lauritz Harken, and Shu-Ming Li (2021). Elucidation of the streptoazine biosynthetic pathway in *Streptomyces aurantiacus* reveals the presence of a promiscuous prenyltransferase/cyclase. *Journal of Natural Products*, 84, 3100–3109. DOI: 10.1021/acs.jnatprod.1c00844. (*equal contribution)
3. Jing Liu*, **Yiling Yang***, Xiulan Xie, and Shu-Ming Li (2023). A *Streptomyces* cytochrome P450 enzyme catalyzes regiospecific C2-guaninylation for the synthesis of diverse guanitrypmycin analogs. *Journal of Natural Products*, 86, 94 – 102. DOI: 10.1021/acs.jnatprod.2c00787. (*equal contribution)
4. Jing Liu*, Lauritz Harken*, **Yiling Yang**, Xiulan Xie, and Shu-Ming Li (2022). Widely distributed bifunctional bacterial cytochrome P450 enzymes catalyze both intramolecular C—C bond formation in *cyclo*-L-Tyr-L-Tyr and its coupling with nucleobases. *Angewandte Chemie International Edition*, 2022, 61, e202200377. DOI: 10.1002/anie.202200377. (*equal contribution)
5. Marlies Peter, **Yiling Yang**, and Shu-Ming Li (2022). A terpene cyclase from *Aspergillus ustus* is involved in the biosynthesis of geosmin precursor germacradienol. *RSC Advances*, 2022, 12, 28171–28177, DOI: 10.1039/d2ra05033a.

Erklärung zum Eigenanteil

| Titel der Publikation und Journal incl. Jahr, Heft, Seitenzahl + doi O: Originalarbeit Ü: Übersichtartikel/Review | Autoren | geschätzter Eigenanteil in % | Bitte angeben: angenommen/ eingereicht |
|---|---|---|---|
| Ustethylin biosynthesis implies phenethyl derivative formation in <i>Aspergillus ustus</i> . <i>Organic Letters</i> , 2020, 22, 7837–7841 DOI: 10.1021/acs.orglett.0c02719. Originalarbeit | Liujuan Zheng*, Yiling Yang* , Haowen Wang, Aili Fan, Liping Zhang, and Shu-Ming Li | 32 | angenommen |
| Elucidation of the streptoazine biosynthetic pathway in <i>Streptomyces aurantiacus</i> reveals the presence of a promiscuous prenyltransferase/cyclase. <i>Journal of Natural Products</i> , 2021, 84, 3100–3109. DOI: 10.1021/acs.jnatprod.1c00844. Originalarbeit | Jing Liu*, Yiling Yang* , Lauritz Harken, and Shu-Ming Li | 40 | angenommen |
| A <i>Streptomyces</i> cytochrome P450 enzyme catalyzes regiospecific C2-guaninylation for the synthesis of diverse guanitrypmycin analogs. <i>Journal of Natural Products</i> , 2023, 86, 94–102. DOI: 10.1021/acs.jnatprod.2c00787. Originalarbeit | Jing Liu*, Yiling Yang* , Xiulan Xie, and Shu-Ming Li | 40 | angenommen |
| Widely distributed bifunctional bacterial cytochrome P450 enzymes catalyze both intramolecular C – C bond formation in <i>cyclo</i> -L-Tyr-L-Tyr and its coupling with nucleobases. <i>Angewandte Chemie International Edition</i> , 2022, 61, e202200377. DOI: 10.1002/anie.202200377. Originalarbeit | Jing Liu*, Lauritz Harken*, Yiling Yang , Xiulan Xie, and Shu-Ming Li | 8 | angenommen |
| A terpene cyclase from <i>Aspergillus ustus</i> is involved in the biosynthesis of geosmin precursor germacradienol. <i>RSC Advances</i> , 2022, 12, 28171–28177 DOI: 10.1039/d2ra05033a. Originalarbeit | Marlies Peter, Yiling Yang , and Shu-Ming Li | 10 | angenommen |

*: These authors contributed equally to this work.

Kandidat(in)

Unterschrift Betreuer(in)

Abbreviations

The international system of units and units derived thereof have been used.

| | |
|--------------------|-------------------------------------|
| aa | amino acid |
| aa-tRNA | aminoacyl tRNA |
| aaRSs | Aminoacyl tRNA synthetases |
| A domain | adenylation domain |
| ACP domain | acyl carrier protein domain |
| bp | base pair |
| BGC | biosynthetic gene cluster |
| CD | circular dichroism |
| CD ₃ OD | deuterated methanol |
| CDCl ₃ | deuterated chloroform |
| CDP | cyclodipeptide |
| CDPS | cyclodipeptide synthase |
| cDNA | complementary deoxyribonucleic acid |
| CDO | cyclodipeptide oxidase |
| CLC | Claisen-like cyclase |
| CLF | Chain length factor |
| CoA | coenzyme A |
| COSY | correlation spectroscopy |
| cFP | <i>cyclo</i> -(L-Phe-L-Pro) |
| cFL | <i>cyclo</i> -(L-Phe-L-Leu) |
| cWA | <i>cyclo</i> -(L-Trp-L-Ala) |
| cWF | <i>cyclo</i> -(L-Trp-L-Phe) |
| cWΔF | <i>cyclo</i> -(L-Trp-ΔPhe) |
| cWH | <i>cyclo</i> -(L-Trp-L-His) |
| cWL | <i>cyclo</i> -(L-Trp-L-Leu) |
| cWΔL | <i>cyclo</i> -(L-Trp-ΔLeu) |
| cWM | <i>cyclo</i> -(L-Trp-L-Met) |
| cWΔM | <i>cyclo</i> -(L-Trp-ΔMet) |
| cWP | <i>cyclo</i> -(L-Trp-L-Pro) |

ABBREVIATIONS

| | |
|-----------------------------|---|
| cWW | <i>cyclo</i> -(L-Trp-L-Trp) |
| cW Δ W | <i>cyclo</i> -(L-Trp- Δ Trp) |
| cWY | <i>cyclo</i> -(L-Trp-L-Tyr) |
| cW Δ Y | <i>cyclo</i> -(L-Trp- Δ Tyr) |
| d | doublet |
| D ₂ O | deuterium oxide |
| Da | dalton |
| dd | double doublet |
| ddd | double double doublet |
| DH domain | dehydratase domain |
| DKP | diketopiperazine |
| DMA | dimethylallyl |
| DMAPP | dimethylallyl diphosphate |
| DMATS | dimethylallyltryptophan synthase |
| DMSO- <i>d</i> ₆ | deuterated dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| dq | double quartet |
| dt | double triplet |
| <i>e.g.</i> | <i>exempli gratia</i> |
| EIC | extracted ion chromatogram |
| ER domain | enoyl reductase domain |
| ESI | electrospray ionization |
| EtOAc | ethyl acetate |
| FDA | food and drug administration |
| FPP | farnesyl diphosphate |
| gDNA | genomic deoxyribonucleic acid |
| GMM | glucose minimal medium |
| GPP | geranyl diphosphate |
| HMBC | heteronuclear multiple bond correlation |
| HPLC | high performance liquid chromatography |
| HR-MS | high resolution-mass spectrometry |
| HR-PKS | highly reducing polyketide synthase |

ABBREVIATIONS

| | |
|-------------|--|
| HSQC | heteronuclear single quantum coherence |
| Hz | hertz |
| <i>i.e.</i> | id est |
| <i>J</i> | coupling constant |
| kbp | kilo base pairs |
| <i>kcat</i> | turnover number |
| kDa | kilodalton |
| K_M | Michaelis-Menten constant |
| KR domain | ketoreductase domain |
| KS domain | ketosynthase domain β -ketoacyl synthase |
| LC-MS | liquid chromatography-mass spectrometry |
| m | multiplet |
| MAT domain | malonyl-CoA-ACP transacylase domain |
| mAU | Milli absorbance unit |
| Mb | mega base pairs |
| MeOH | methanol |
| MeT domain | methyltransferase domain |
| MHz | mega hertz |
| mRNA | messenger ribonucleic acid |
| MTs | methyltransferase |
| multi | multiplicity |
| <i>m/z</i> | mass-to-charge ratio |
| NADH | nicotinamide adenine dinucleotide (reduced form) |
| NADPH | nicotinamide adenine dinucleotide phosphate (reduced form) |
| NMR | nuclear magnetic resonance |
| NP | natural product |
| NR-PKS | Nonreducing polyketide synthase |
| NRPS | nonribosomal peptide synthetase |
| OH | hydroxylate |
| P450 | cytochrome P450 |
| PCP domain | peptidyl carrier protein domain |
| PCR | polymerase chain reaction |

ABBREVIATIONS

| | |
|------------|---|
| PD | potato dextrose |
| PDB | potato dextrose broth |
| PEG | polyethylene glycol |
| Phe | phenylalanine |
| PKS | polyketide synthase |
| PKS-NRPS | polyketide synthase-nonribosomal peptide synthetase |
| ppm | parts per million |
| PR-PKS | partially reducing polyketide synthase |
| PT | prenyltransferase |
| PT domain | product template domain |
| P1 | pocket 1 |
| P2 | pocket 2 |
| q | quartet |
| QM | quinone methide |
| RNA | ribonucleic acid |
| rpm | revolutions per minute |
| s | singlet |
| SAM | S-adenosyl-L-methionine |
| SAT domain | starter unit acyltransferase domain |
| SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SM | secondary metabolite |
| t | triplet |
| TB | terrific broth |
| TC | terpene cyclase |
| T domain | thiolation domain |
| td | triple doublet |
| TE domain | thioesterase domain |
| Tris | tris(hydroxymethyl)aminomethane |
| UV | ultraviolet |
| v/v | volume per volume |
| w/v | weight per volume |
| WT | wild-type |

ABBREVIATIONS

| | |
|-------------------------|-----------------------------------|
| δ_C | chemical shift of ^{13}C |
| δ_H | chemical shift of ^1H |
| $[\text{M}+\text{H}]^+$ | molecular ion plus proton |
| $[\text{M}-\text{H}]^-$ | molecular ion minus proton |
| $\times g$ | gravitational acceleration |
| 2-OG | 2-oxoglutarate |
| 2,5-DKP | 2,5-diketopiperazines |

Summary

Natural products have high structural diversity with various pharmacological or biological activities, which are of great significance to our life and drug research. Millions of natural products with versatile structural diversity have been found in nature. In recent years, a large number of microbial genome sequences have been released in public databases and revealed many silent or cryptic secondary metabolite gene clusters hidden in their genomes. This shows the great potential for discovering new metabolites. Advances in sequencing technology and bioinformatics analysis also provide great advantages for studying the biosynthesis and structural diversity of these metabolites. Structural differentiation of natural products begins with the formation of basic scaffolds using basic building blocks derived from primary metabolism catalyzed by different backbone enzymes. The structural complexity of natural products mainly arises from tailoring enzymes to highly functionalize the skeletons with a set of chemical transformations. The well-studied modification enzymes range from different types of oxidoreductases, cytochrome P450 enzymes, to various prenyltransferases (PTs) and methyltransferases (MTs). In addition, nonenzymatic events have also contributed to the formation of final products with vast diversity and complexity. Therefore, fully exploring these unexplored gene clusters and the substrate promiscuity of enzymatic and non-enzymatic reactions for the natural product formation may be a promising way and a new strategy to explore the metabolite diversity.

In a cooperation study with Dr. Liujuan Zheng, the biosynthesis of a highly oxygenated phenethyl derivative ustethylin A, isolated from *Aspergillus ustus*, was elucidated. Due to the instability of ustethylin A, it was acetylated before isolation and structure elucidation. Gene deletion and heterologous expression proved that the phenethyl core structure is assembled by a polyketide synthase (UttA) harboring a methyltransferase domain. Isotopic labelling experiments proved that the backbone of ustethylin A is derived from malonyl-CoA and the methyl groups, also in the phenethyl residue, are from L-methionine. Modifications on the core structure by an aryl acid reductase (UttJ), a putative nonheme Fe^{II}/2-oxoglutarate dependent oxygenase (UttH), a cytochrome P450 enzyme (UttC) and a O-methyltransferase (UttF) led to the final product ustethylin A. This study is the first report on the biosynthetic pathway of a phenethyl-containing natural product.

In cooperation with Dr. Jing Liu, the biosynthesis of streptoazine C and guanitrypmycin D1 was elucidated. Firstly, a three-gene cluster coding for a cyclodipeptide synthase, a prenyltransferase, and a methyltransferase was identified in *Streptomyces aurantiacus* by genome mining. Heterologous expression and precursor incubation experiments led to the elucidation of the biosynthetic steps of streptoazine C. *In vivo* biotransformation experiments proved the high flexibility of the prenyltransferase SasB toward tryptophan-containing cyclodipeptides and their

SUMMARY

dehydroderivatives for regular C-3-prenylation. This study provides an enzyme with a high substrate promiscuity from the less explored prenyltransferase group in cyclodipeptide synthase-related pathways.

Afterwards, a two-gene cluster coding for a CDPS and a cytochrome P450 was identified in *Streptomyces* sp. NRRL S-1521 by phylogenetic analysis. Heterologous expression and structural elucidation of the isolated products proved that the cytochrome P450 GutD₁₅₂₁ catalyzes the regiospecific transfer of guanine to C-2 of the indole ring of *cyclo*-(L-Trp-L-Tyr) *via* a C-C linkage, which represents a new chemical transformation within this enzyme class. Precursor incubation experiments revealed that GutD₁₅₂₁ efficiently accepts several other tryptophan-containing cyclodipeptides or derivatives for regiospecific coupling with guanine, thus generating different guanitrypmycin analogs. This study provides a biocatalyst for a new linkage pattern between an indole ring and a guanine moiety and expands the functional spectrum of P450s as tailoring enzymes.

Zusammenfassung

Naturstoffe haben eine hohe strukturelle Vielfalt mit verschiedenen pharmakologischen und biologischen Aktivitäten, die für unsere Lebens- und Arzneimittelforschung von großer Bedeutung sind. Millionen von Naturstoffen mit vielseitiger Strukturvielfalt wurden in der Natur gefunden. In den letzten Jahren wurde eine große Anzahl mikrobieller Genomsequenzen in öffentlichen Datenbanken publiziert, welches viele stille oder kryptische Gencluster von Sekundärmetaboliten enthüllte, die in den Genomen der Mikroorganismen verborgen sind. Dies verdeutlicht das große Potenzial zur Entdeckung neuer Stoffwechselprodukte. Fortschritte in der Sequenzierungstechnologie und der bioinformatischen Analyse bieten auch große Vorteile für die Untersuchung der Biosynthese und der strukturellen Vielfalt dieser Metaboliten. Die strukturelle Differenzierung von Naturstoffen beginnt mit der Bildung von Grundgerüsten unter Verwendung von Grundbausteinen aus dem Primärstoffwechsel, die durch verschiedene Rückgratenzyme katalysiert werden. Die strukturelle Komplexität von Naturstoffen entsteht hauptsächlich durch sogenannte Tailoring Enzyme, welche die Grundgerüste durch eine Reihe chemischer Umwandlungen hochgradig funktionalisieren. Die am besten untersuchten Modifikationsenzyme reichen von verschiedenen Arten von Oxidoreduktasen, Cytochrom P450 Enzymen, bis hin zu verschiedenen Prenyltransferasen (PTs) und Methyltransferasen (MTs). Darüber hinaus tragen auch die nicht enzymatische Ereignisse zu der großen Vielfalt und Komplexität der Endprodukte bei. Die Erforschung unbekannter Gencluster und die Nutzung der Substratpromiskuität von enzymatischen und nichtenzymatischen Reaktionen für die Naturstoff-Biosynthese sind somit ein vielversprechender Weg und eine neue Strategie zur Erforschung der Naturstoffvielfalt.

In einer Kooperationsstudie mit Dr. Liujuan Zheng wurde die Biosynthese eines aus *Aspergillus ustus* isolierten hoch oxygenierten Phenethylderivates Ustethylin A aufgeklärt. Hierbei wurden Probleme bei der Isolierung und Strukturaufklärung instabiler Verbindungen mittels Acetylierung überwunden. Gendeletion und heterologe Expression bewiesen, dass die Phenethyl-Grundstruktur von einer Polyketid-Synthase (UttA) zusammengesetzt wird, welche eine Methyltransferase-Domäne besitzt. Isotopenmarkierungsexperimente bewiesen, dass das Rückgrat von Ustethylin A von Malonyl-CoA, die Methylgruppe im Phenethylrest und die O-Methylgruppe aus L-Methionin abgeleitet sind. Modifikationen an der Grundstruktur durch eine Arylsäurereduktase (UttJ), eine mutmaßliche Nighthäm-Fe^{II}/2-Oxoglutarat-abhängige Oxygenase (UttH), ein Cytochrom-P450-Enzym (UttC) und eine O-Methyltransferase (UttF) führen zum Endprodukt Ustethylin A. Diese Studie ist der erste Bericht über den Biosyntheseweg eines phenethyl-haltigen Naturstoffs.

In Zusammenarbeit mit Dr. Jing Liu wurde die Biosynthese von Streptoazin C und Guanitrypmycin D1 aufgeklärt. Zunächst wurde durch Genome Mining in *Streptomyces aurantiacus* ein Cluster mit

ZUSAMMENFASSUNG

drei Genen identifiziert, das für eine Cyclodipeptid-Synthase, eine Prenyltransferase und eine Methyltransferase kodiert. Heterologe Expressions- und Zufütterungsexperimente dienten der Aufklärung der Biosyntheseschritte von Streptoazin C. *In-vivo*-Biotransformationsexperimente bewiesen die hohe Flexibilität der Prenyltransferase SasB gegenüber Tryptophan-haltiger Cyclodipeptide und deren Dehydroderivate für die reguläre C-3-Prenylierung. Diese Studie beschreibt ein Enzym mit einer hohen Substratpromiskuität aus der wenig erforschten Gruppe der Prenyltransferasen in Stoffwechselwegen mit Cyclodipeptid-Synthasen.

Anschließend wurde in *Streptomyces* sp. NRRL S-1521 durch phylogenetische Analysen, heterologe Expression und Strukturaufklärung bewiesen, dass das Cytochrom P450 Enzym GutD₁₅₂₁ den regiospezifischen Transfer von Guanin auf C-2 des Indolrings von *cyclo*-(L-Trp-L-Tyr) über eine C-C-Verknüpfung katalysiert, welches eine neue chemische Umwandlung innerhalb dieser Enzymklasse repräsentiert. Zufütterungsexperimente zeigten, dass GutD₁₅₂₁ auch weitere tryptophanhaltige Cyclodipeptide und Derivate davon effizient für eine regiospezifische Kopplung mit Guanin akzeptiert, wodurch verschiedene Guanitrypmycin-Analoga erzeugt wurden. Diese Studie beschreibt einen Biokatalysator für ein neues Bindungsmuster zwischen einem Indolring und einer Guanineinheit und erweitert das funktionelle Spektrum von Cytochrom P450 Oxidasen als Tailoring Enzyme.

1 Introduction

1.1 Natural products

Natural products (NPs) are chemical substances produced by living organisms in nature.¹ Although there is no consensus on more restrictive definitions of NPs,² they can be classified according to their biological function, biosynthetic pathway or source. In the field of organic chemistry, NPs are usually defined as primary and secondary metabolites. In the fields of medicinal chemistry and pharmacognosy, more specific definitions are often used, limiting NPs to secondary metabolites (SMs).² Primary metabolites are organic molecules with an intrinsic function that is essential to the survival of the organism that produces them. In contrast, SMs are organic molecules that typically have an extrinsic function that primarily affects other organisms besides the producer. SMs are not essential for survival but increase an organism's competitiveness in its environment.²⁻¹²

Most naturally occurring compounds are end products of secondary metabolism, which are unique compounds for particular organisms or classes of organisms.² Most of the NPs possess a high degree of structural diversity and unique pharmacological or biological activities due to thousands of years of natural selection and evolutionary processes. In fact, the structural diversity of NPs far exceeds the capabilities of synthetic organic chemists in the laboratory. In the past years, NPs have been widely used in traditional and modern medicine to treat diseases. Currently, NPs are often used as starting points for drug discovery, which are then synthetically modified to help reduce side effects and increase bioactivity. Nearly half of all drugs approved by the U.S. Food and Drug Administration (FDA) are derived from NPs.¹³ In addition to pharmaceuticals, NPs and their derivatives are often used as food additives in the form of spices and herbs, antimicrobials and antioxidants to preserve the freshness and longevity of food. Furthermore, natural and organic products permeate nearly every aspect of our lives, from the clothes we wear to plastic and rubber products, health and beauty products, and even the energy we use to power our cars.

NPs can be directly extracted from microorganisms, plants and animals.^{2, 14-16} A crude extract from any one of these sources contains a range of structurally diverse chemical compounds. Chemical diversity in nature is based on biological diversity. Usually, a natural extract has some form of biological activity that can be detected and attributed to a single compound or a set of related compounds produced by the organism. These active compounds can be used in drug discovery and development directly as they are, or they may be synthetically modified to enhance biological properties or reduce side effects. In the last decades, great progress has been made in isolation and chemical characterization of NPs from microorganisms.¹⁷⁻¹⁹ Microbial metabolites are one of the most important constitution of reported NPs.

1.2 Strategy for NP-based drug discovery

The classical strategy for discovering new bioactive compounds is based on bioactivity screenings, which have made great contributions to new NP-based drug discovery in the past century.²⁰ The process begins with extraction of NPs from organisms, followed by the identification of a crude extract with promising pharmacological activities, then the next step is (often multiple) consecutive bioactivity-guided fractionation until the pure bioactive compounds are isolated (Figure 1).

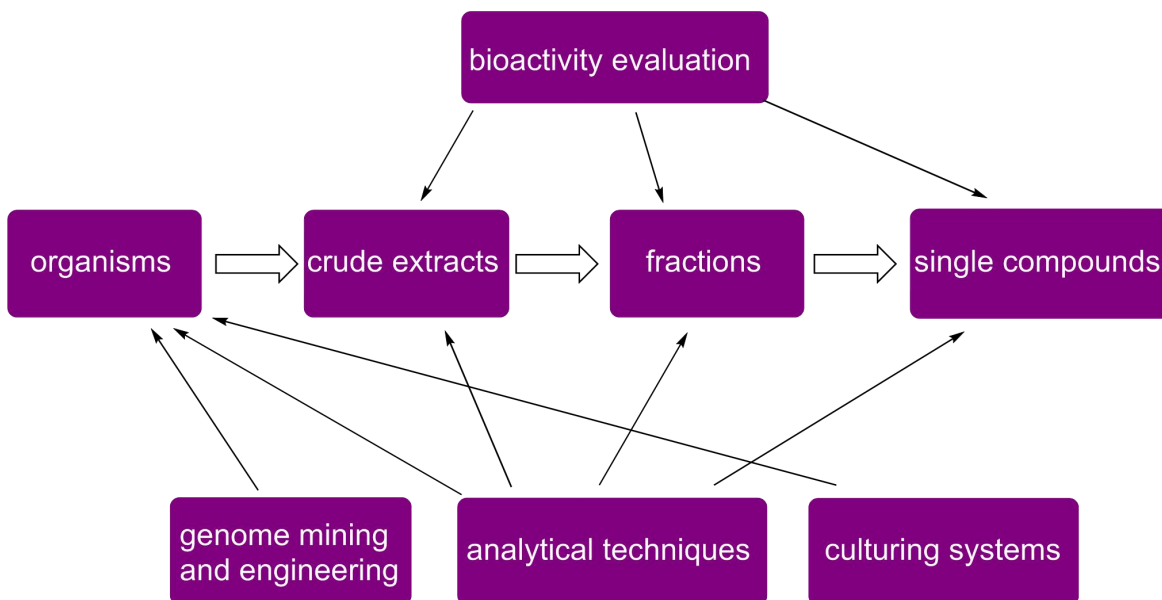


Figure 1. Strategies for NP-based drug discovery.

However, discovering new compounds with this strategy has become more difficult due to the repeated isolation of already known ones. In addition, obtaining a sufficient amount of biological material to isolate and characterize a new bioactive NP, and identifying the bioactive compounds of interest can be challenging. Fortunately, there have been substantial advances both in the development of screening assays and strategies to identify the modes of action of active compounds.²¹⁻²⁴ Application of analytical techniques, genome mining and engineering, and advances in microbial culturing systems help to overcome challenges in NP-based drug discovery. Many advances discussed above are supported by computational tools including databases such as genomic, chemical or spectral analysis data and NP databases. There are increasing tools that enable the analysis of genetic information, the prediction of chemical structures and pharmacological activities, the integration of data sets with diverse information and machine learning applications.²⁵⁻²⁷ In conclusion, advances in microbiology, biochemistry, genome sequencing and bioinformatics provide unlimited possibilities to enrich the NP library and expand the pharmaceutical repertoire.

1.3 Categories of natural products

The majority of so far discovered NPs can be roughly classified into polyketides, peptides, terpenoids, and alkaloids, etc.

1.3.1 Polyketides

Polyketides are a large group of structurally diverse and therapeutically important NPs.²⁸ The fundamental chemical aspect in the biosynthesis of polyketides is the Claisen condensation reaction, where the polyketide backbone is formed by the condensation of starter units such as acetyl-CoA with extender units such as malonyl-CoA. On the one hand, structural variety is achieved by PKS inherent factors such as the use of various starter and extender units, the difference in polyketide length, degree of reduction and methylation, as well as by different release and cyclization mechanisms. On the other hand, further diversification can be achieved by polyketide tailoring enzymes for oxidation, reduction, rearrangement, and transfer reactions.²⁹⁻³⁰

Some polyketides and their derivatives are important drugs for clinical use (Figure 2). Among them, lovastatin, also known as mevinolin, was isolated from *Aspergillus terreus* in 1978 and used as a cholesterol-lowering agent.³¹⁻³⁴ Griseofulvin, isolated from *Penicillium* sp., has antifungal properties.³⁵ Tetracyclines isolated from *Streptomyces aureofaciens* and erythromycins obtained from *Saccharopolyspora erythraea* are used as macrolide antibiotics.³⁶⁻³⁷ Doxorubicin was obtained from *Streptomyces peucetius* as a chemotherapeutic agent in the treatment of cancer.³⁸ In addition to the potential within the development of new drugs, polyketides can also be used for production of biofuel in the chemical industry, as well as pigments in the textile industry.³⁹⁻⁴⁰

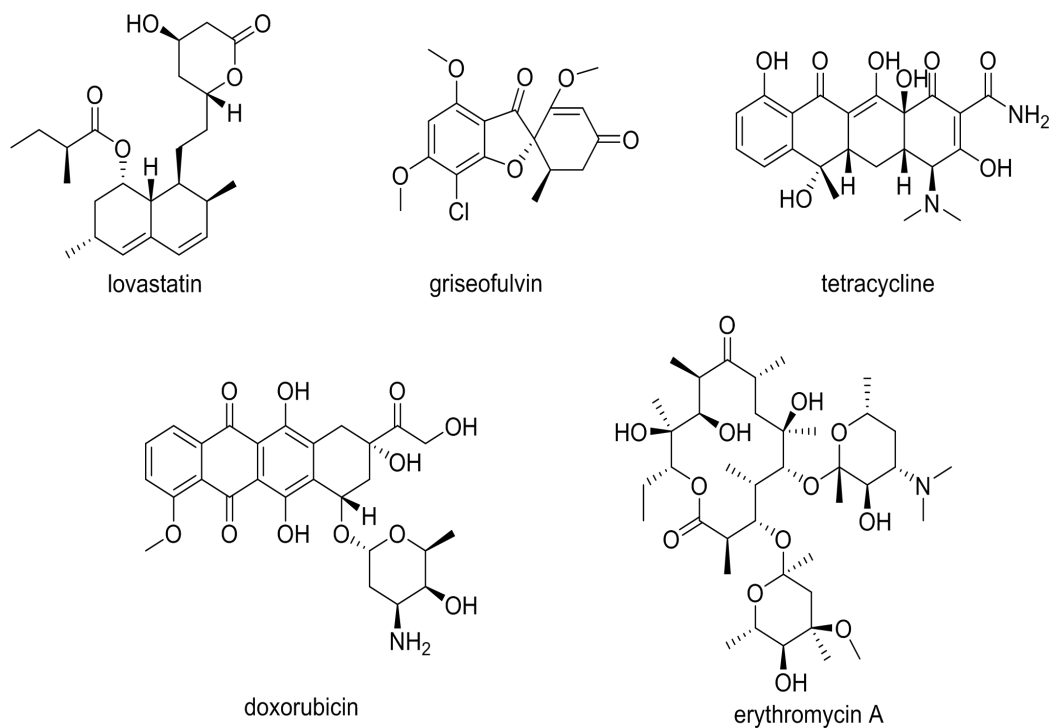


Figure 2. Representatives of polyketides.

1.3.2 Peptides: 2,5-Diketopiperazines

2,5-diketopiperazines (2,5-DKPs), the smallest class of cyclic peptides, are achieved by the condensation of two α -amino acids.⁴¹⁻⁴² They are heterocyclic compounds and characterized by a central diketopiperazine (DKP) ring. The general core of 2,5-DKPs is shown in Figure 3. Different substitution of side chain groups R_1 and R_2 , depending on the incorporation of different amino acids, will generate the simplest cyclodipeptides (CDPs). Their central scaffold, the six membered ring, can be then modified by various substitutions and different stereochemistry.

The DKP scaffolds can be easily obtained from α -amino acids by conventional methodology.⁴³ In recent years, the synthesis of 2,5-DKPs *via* solid-phase intramolecular cyclization has been the most utilized method, which is useful for the construction of chemical libraries for drug lead discovery.⁴⁴ In nature, the 2,5-DKP scaffolds are synthesized by two different types of enzymes, the nonribosomal peptide synthetases (NRPSs) and the cyclodipeptide synthases (CDPSs). Furthermore, the tailoring enzymes introduce specific modifications to the DKP cores and (or) the side chains to generate more complex DKP-containing NPs.

2,5-DKPs are ubiquitous in nature and often found as side products of polypeptides, especially during the production process of food and beverages.⁴⁵ In recent years, CDPs and their derivatives

have attracted an increasing interest due to their important and diverse biological and potential pharmacological properties, including antibacterial, antifungal, antiviral, antitumor, and immunosuppressive effects.^{42, 46} Prominent representatives are *cyclo*-(L-Phe-L-Pro) (cFP) and *cyclo*-(L-Phe-*trans*-4-OH-L-Pro), which exhibit antifungal activities.⁴⁷ Phenylahistin, shows an inhibitory effect on the cell cycle progression.⁴⁸ Plinabulin (BPI-2358) is used for non-small cell lung cancer treatment.⁴⁹ Gliotoxin is used as an immunosuppressive cytotoxin (Figure 3).⁵⁰

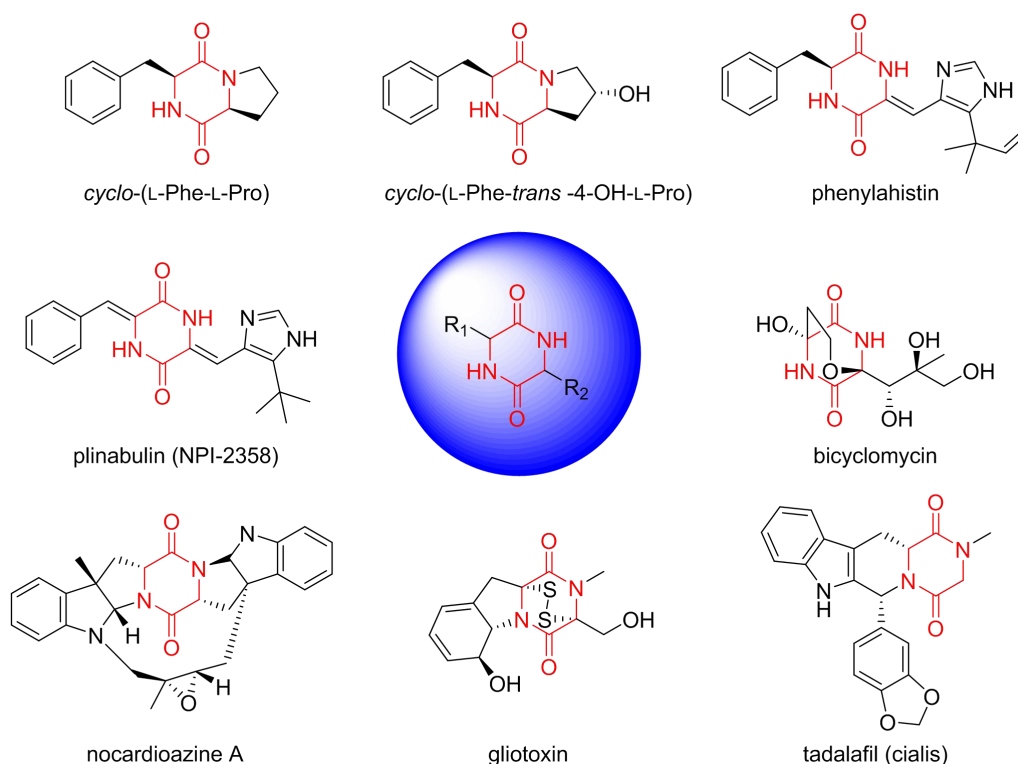


Figure 3. General structure of 2,5-DKPs and examples of bioactive DKPs and derivatives.

1.4 Biosynthesis of NPs

1.4.1 Backbone enzymes of NPs

The biosynthetic genes for the formation of a given NP are often closely located in the microbial genome and form a so-called gene cluster. Such clusters usually consist of one or more backbone gene(s) as well as several genes for modifications.⁵¹⁻⁵² The backbone enzymes involved in the biosynthesis of NPs generally include polyketide synthase (PKS), NRPS, CDPS, terpene cyclase (TC), and so on. These enzymes usually catalyze the first step of the biosynthesis to form scaffolds, which are afterwards modified by tailoring enzymes in multiple biosynthetic steps to produce the final products. PKSs and CDPSs will be discussed in this section.

1.4.1.1 Polyketide synthases

PKSs are one of the most abundant enzyme classes attributed in microorganisms for NP biosynthesis. On the basis of their structural composition, PKSs can basically be divided into three types.^{29, 53-54}

Type I PKSs are large multidomain megaenzymes with distinct modules and can be further divided into modular and iterative type I PKSs.²⁹ In modular type I PKSs, each module consists of different domains and is used only once during polyketide biosynthesis. Most prominent representative of the modular type I PKSs is 6-deoxyerythronolide B synthase (DEBS), participating in the biosynthesis of erythromycin A (Figure 4).⁵⁵ In contrary, in iterative type I PKSs, domains are clustered in a single module which is used repeatedly for polyketide formation. Depending on the domain architecture, the type I iterative PKSs can be further subdivided into nonreducing PKSs (NR-PKSs), partially reducing PKSs (PR-PKSs), and highly reducing PKSs (HR-PKSs).⁵⁵⁻⁵⁷ Their representative examples are 5-methylorsellinic acid synthase (MpaC), 6-methylsalicylic acid synthase (6-MSAS) and lovastatin nonaketide synthase (LovB) (Figure 4).^{31, 58-61} As mentioned above, modular type I PKSs don't work iteratively. Since each module is used only once, it is easier to predict the formed polyketide compared to iterative type I PKS.

Type II PKSs are enzyme complexes with mono functional proteins. Fungal type II PKSs have not been described yet to date, therefore type II PKSs are currently limited to the bacterial kingdom, especially to *Streptomyces* species.^{53, 62-65} The minimal set-up of a type II PKS displays the Ketosynthase domain (KS), catalyzing the Claisen-like condensation of acetyl- and malonyl-CoA units in an iterative fashion. The acyl carrier protein domain (ACP) anchors the polyketide to the enzyme complex and the chain length factor (CLF) shows very high similarity to the KS and determines the length of the growing polyketide. Additionally, in some cases, the CLF domain can catalyze the decarboxylation of malonyl-CoA to acetyl-CoA to generate the starter unit.⁶⁶ Another part of the enzyme complex is the malonyl-CoA:ACP transferase (MCAT) responsible for loading malonyl-CoA to the ACP domain. A well-studied example for type II PKS is TcmKLM involved in the formation of the antibiotic tetracenomycin C (Figure 5).⁶⁷

INTRODUCTION

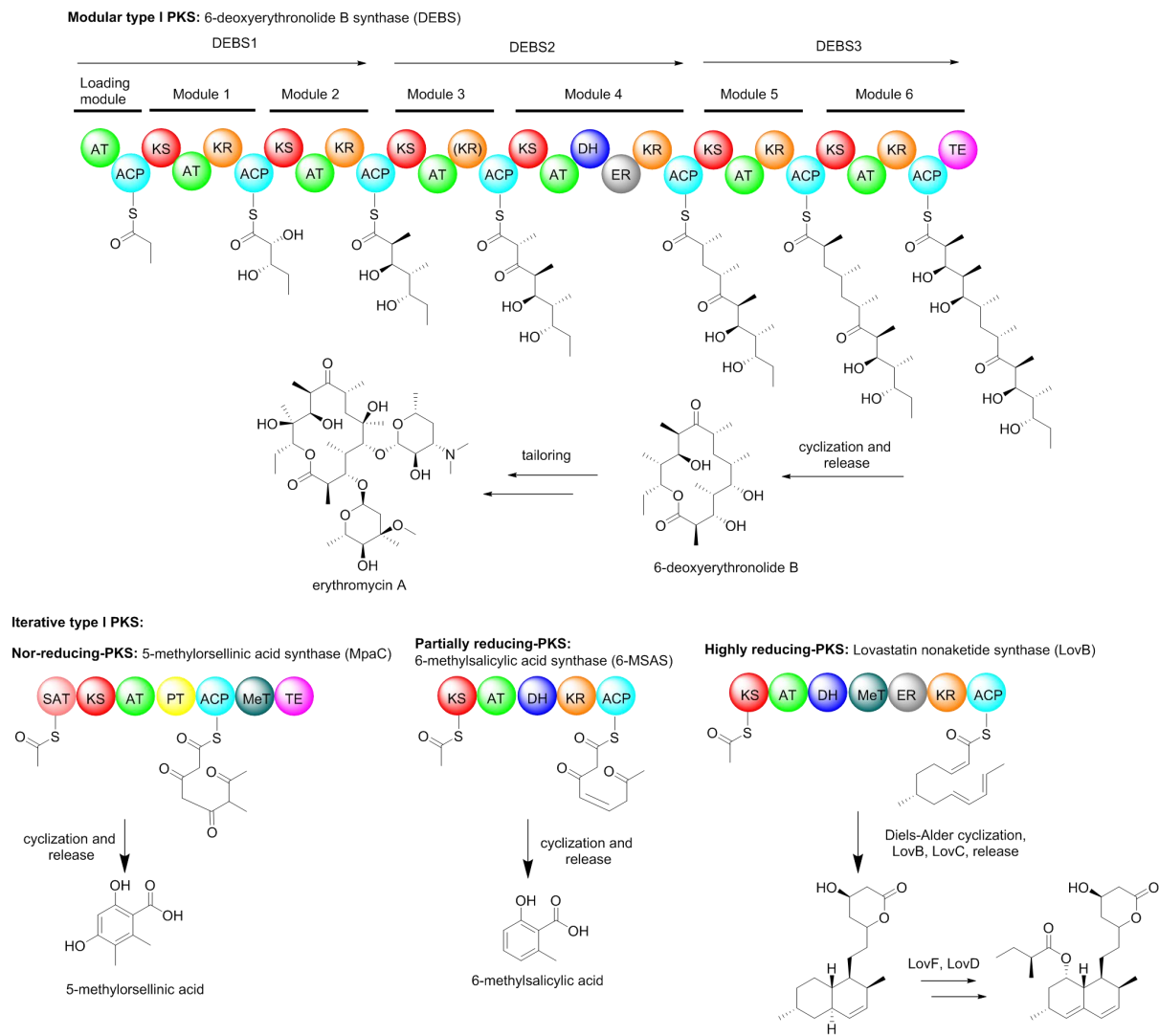


Figure 4. Representatives of the type I PKSs, their domain architecture and reaction products.

Type II PKS:

Tetracenomycin polyketide synthase complex TcmKLMN:

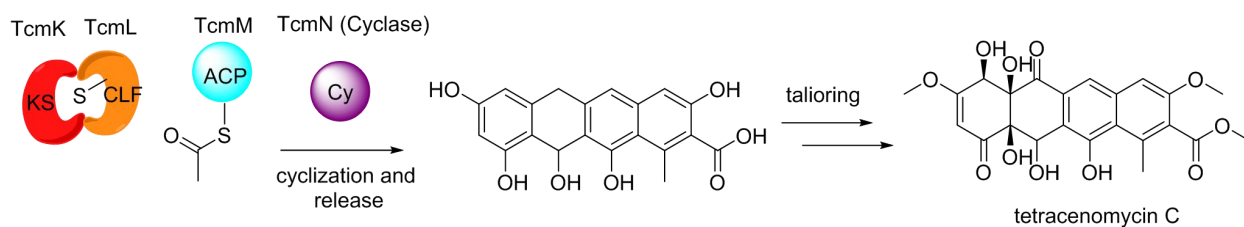


Figure 5. Representative of the type II PKS, their domain architecture and their reaction products

Distinct from the above mentioned type I and type II PKSs, type III PKSs have no acyl carrier protein domain. Type III PKSs are the smallest types of PKS known up to this point with a common size of

80 – 90 kDa.⁶⁸ Type III PKSs were not only firstly identified in plants but are also widely distributed over the plant kingdom and were initially thought to be exclusive to plants.⁶⁹ Later, type III PKSs were also commonly discovered in bacteria.⁷⁰ Fungal type III PKSs are rare, but they are more frequently identified in recent years.⁷¹ Type III PKSs catalyze the biosynthesis of small aromatic polyketides. Typically, circular-form and linear-form acyl-CoAs are used as starter units and are extended by the condensation of malonyl-CoA. For cyclization and product release, various types of cyclization reactions are used like Claisen condensation, aldol or lactone formation.⁶⁸ The best-studied type III PKS belongs to the family of chalcone synthases catalyzing the formation of naringenin (Figure 6).⁷²⁻⁷³

Type III PKS:

Naringenin chalcone synthase (CHS):

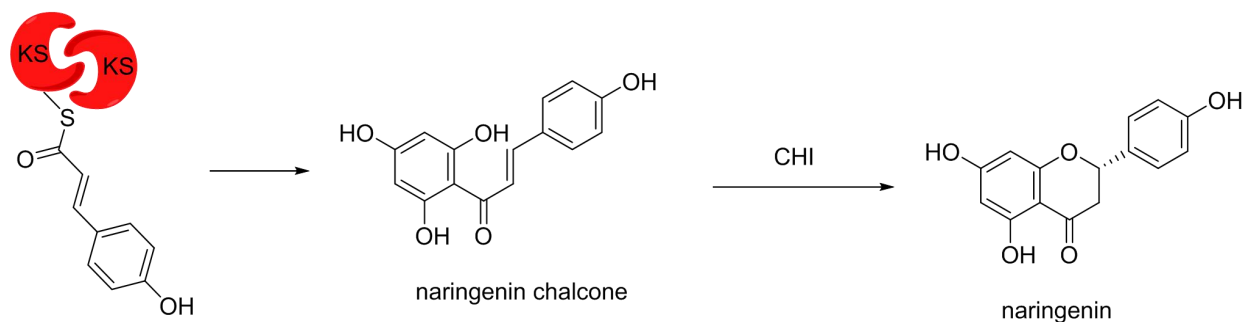


Figure 6. Representative of the type III PKS, their domain architecture and their reaction product

The biosynthesis of polyketides proceeds in three phases, (1) starter unit loading, (2) chain elongation and reduction, (3) polyketide cyclization and release. All these steps are catalyzed by the concerted action of the different domains. The basic domains required for polyketide elongation include KS domains for catalyzing the decarboxylative Claisen condensation to extend the polyketide chain, acyltransferases (AT) for selection and recognition of the starter unit and extender unit, and ACP domains which shuttle growing polyketides between the active sites of the PKS. In addition to the minimal domain architecture of KS, AT, and ACP domains, there are further accessory domains for polyketide chain modification. These include ketoreductase (KR) domain for reduction of the β -keto to a hydroxyl group, dehydratase (DH) domain for dehydration to generate an α,β -unsaturated thioester, and enoylreductase (ER) domain for further reduction of the double bond to a saturated moiety.⁷⁴⁻⁷⁵

The vast diversity and complexity of polyketides can be ascribed to the following strategies that are utilized by PKSs during the assembly process. Firstly, utilization of different starter and extender units by PKSs leads to the variation of polyketide skeletons.⁷⁵⁻⁷⁶ Secondly, PKSs employ different reduction domains to form polyketide chains with different degrees of unsaturation.⁷⁷ Thirdly,

different cyclization mechanisms contribute to the polyketide variety.⁷⁸ In recent years, combinatorial biosynthesis was used as a new approach to generate large libraries of new compounds. Such strategies include engineering modular PKSs by swapping and replacing PKS single domains or entire modules.⁷⁹

1.4.1.2 Cyclodipeptide synthases

Cyclodipeptide synthases (CDPSs) are enzymes that directly use the aminoacyl-tRNAs (aa-tRNAs) from the primary metabolism as substrates to form the DKP scaffolds. They are small molecular proteins (~30 kDa) typically with 200 – 300 amino acid residues.⁸⁰ Since the first description of a CDPS enzyme, AlbC from *Streptomyces noursei* in 2002 responsible for the formation of *cyclo*-(L-Phe-L-Leu) (cFL), more than 120 CDPSs have been characterized.⁸¹ Over 75 different cyclodipeptides are assembled by CDPSs, consisting of 18 of the 20 proteinogenic amino acids. Very recently, CDPSs have been also demonstrated to incorporate non-canonical amino acids (ncAAs) to produce noncanonical 2,5-DKPs. Functionally characterized CDPSs are mainly from three bacteria phyla of *Actinobacteria*, *Firmicutes*, and *Proteobacteria*.⁸²⁻⁸³ According to specific catalytic residues, they fall into two main phylogenetically distinct subfamilies, namely NYH and XYP.⁸²⁻⁸⁵ As the number of identified CDPSs keeps increasing, the classification of its subfamilies is constantly adjusted and improved.

The earliest resolved crystallographic structures of three CDPSs, AlbC (PDB 3OQV), Rv2275 (PDB 2X9Q) and YvmC (PDB 3OQH), revealed the catalytic mechanism of the CDP formation.⁸⁶⁻⁸⁸ Their monomeric protein possesses a common compact α/β fold and a conserved Rossmann-fold domain.⁸⁷ Although showing only about 15% sequence similarities, the three CDPSs mentioned above share a high degree of structural similarity with the catalytic domains of class-Ic of aminoacyl-tRNA synthetases (aaRSs), *i.e.*, the Rossmann-fold subdomain and a helical connective polypeptide 1 (CP1) subdomain.⁴⁶ Furthermore, all CDPSs possess two surface-accessible pockets for the substrate selection and catalysis: pocket 1 (P1), corresponding to the aminoacyl binding pocket in class-Ic aaRSs, and pocket 2 (P2), missing in the aaRSs.⁸¹ The specificity of the first aa-tRNA depends on its aminoacyl moiety, conversely that of the second aa-tRNA lies on both the aminoacyl moiety and its tRNA sequence. It has been suggested that CDPSs use a sequential ping-pong mechanism to achieve the synthesis of cyclodipeptides (Figure 7).⁸⁹ After recognition of the first substrate, the catalytic step begins with the binding of the first aa-tRNA to the CDPS and the subsequent transfer of the aminoacyl group to the conserved serine residue of P1 to form an acyl-enzyme intermediate. Then, the resulting intermediate reacts with the aminoacyl moiety of the second aa-tRNA to form a dipeptidyl intermediate, which will further undergo intramolecular cyclization, leading to the formation of the second peptide bond and the yield of final CDP product.⁴⁶

It is noteworthy that most of the CDPSs exhibit some promiscuity in recognizing of the aa-tRNA substrates, resulting in a mixture of CDP products. Subsequently, tailoring enzymes encoded by the genes from CDPS BGC can catalyze various modifications on the CDP product.^{46, 90}

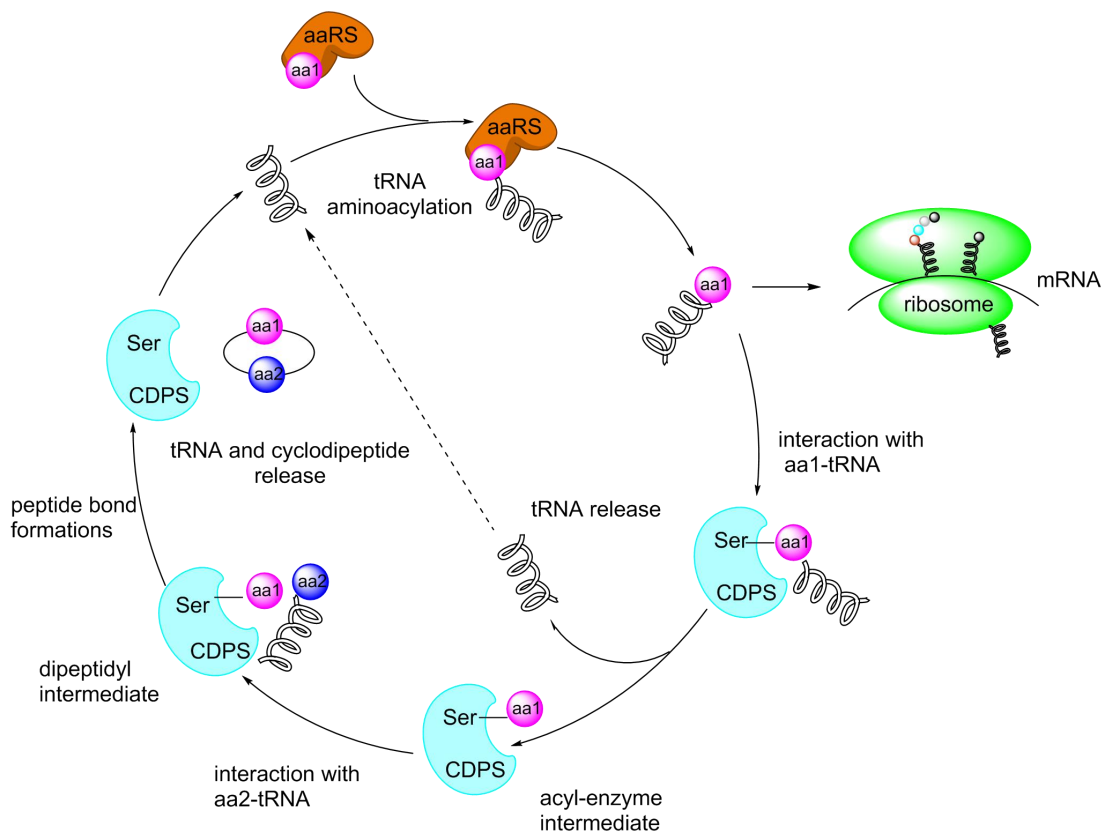


Figure 7. The proposed catalytic mechanism of CDPSs for cyclodipeptide biosynthesis.⁴⁶

1.4.2 Tailoring enzymes

With the rapid increasing works on the NP biosynthesis, many tailoring enzymes for modifications of backbone structures have been identified, such as prenyltransferases (PTs), cytochrome P450 enzymes, methyltransferases (MTs), nonheme Fe^{II}/2-oxoglutarate (Fe^{II}/2-OG)-dependent oxygenases, and flavin-containing oxidoreductases. Once the scaffold has been synthesized by a backbone enzyme, it can be further diversified through subsequent oxidation, reduction, rearrangement, and transfer reactions. In this section, prenyltransferases (PTs) and cytochrome P450 enzymes will be introduced in detail.

1.4.2.1 Prenyltransferases

Prenyltransferases are one of the most important modifying enzymes of NPs, catalyzing the transfer reactions of different prenyl units ($n \times C_5$, $n=1, 2, 3, 4$ etc.) from prenyl diphosphates to various

aliphatic and aromatic acceptors, which greatly increases the diversity of NPs.⁹¹⁻⁹⁴ Depending on enzyme inherent factors, they can be further classified into UbiA-type, CloQ/NphB-type, and dimethylallyltryptophan synthase (DMATS)-type groups.⁹⁵⁻⁹⁶

UbiA-type PTs are membrane-bound enzymes and possess one or more conserved (N/D)DXXD motifs. The bacterial UbiA and its eukaryotic homolog COQ2 are involved in the biosynthesis of ubiquinones and menaquinones. They require metal ions, such as Mg²⁺ for their catalytic activity. In addition to participation in primary metabolism, members of the UbiA family also play an important role in the biosynthesis of SMs.⁹⁷

PTs of the CloQ/NphB-type and DMATS-type are soluble proteins with a characteristic $\alpha\beta\beta\alpha$ fold, also known as PT-barrel. Most of them can function ion-independently.⁹⁸ The eponym for the CloQ/NphB group is due to the first enzyme CloQ⁹⁹ identified in *Streptomyces roseochromogenes*, which is involved in the biosynthesis of clorobiocin, and later NphB¹⁰⁰ identified in *Streptomyces* sp. CL190, involved in the biosynthesis of the naphterpin derivatives.

During the past decades, PTs belonging to the DMATS superfamily have been intensively investigated in biochemistry, molecular biology, and structural biology.⁹⁸ To date, more than 60 representatives of this family have been identified and characterized. For this reason, they are the best-studied group of prenyltransferases.¹⁰¹ In addition to biochemical characterization, the crystal structures of some prenyltransferases such as FtmPT1, AtaPT, FgaPT2, and CdpNPT has also been elucidated.¹⁰²⁻¹⁰⁵ The elucidation of the crystal structures also enabled the clarification of the reaction mechanism. Formation of a dimethylallylic cation by cleavage of the diphosphate group from the prenyl donor initiates the Friedel-Crafts alkylation. The dimethylallylic cation is then nucleophilically attacked by the substrate to be prenylated, for example by the indole nucleus. Rearomatization finally leads to the release of the prenylated product.¹⁰²

One of the characteristics of DMATS enzymes is their substrate flexibility towards a broad spectrum of aromatic compounds (Figure 8). Most DMATSS are specific for their prenyl donors and use DMAPP as a prenyl donor. However, there are also examples of DMATS using FPP and GPP for their reactions.¹⁰¹ The first member of the DMATS superfamily is DmaW (4-DMATS) identified in the ergot alkaloid gene cluster in *Claviceps fusiformis*.¹⁰⁶ It catalyzed the regular transfer of the dimethylallyl (DMA) moiety from DMAPP to the C-4 position of tryptophan.¹⁰⁷ The dimethylallyl moieties can be attached to N-1, C-2, C-3, C-4, C-5, C-6, or C-7 of the indole ring in a regular or reverse manner. Among them FgaPT2, DmaW-Cs, and MaPT are for C4-prenylation, 5-DMATS for C5-prenylation, 6-DMATS and CdpC7PT for C7-prenylation, NotF,¹⁰⁸ BrePT,¹⁰⁹ CdpC2PT,¹¹⁰ TdiB,¹¹¹ and FtmPT1 for C2-prenylation; CdpNPT,¹¹² AnaPT,¹¹³ CdpC3PT,¹¹⁴ and Sas B for C3-prenylation; and CTrpPT for both N1- and C7-prenylation (Figure 8). For the regular prenylation, DMAPP was

proposed to form a dimethylallyl cation/pyrophosphate ion pair. The primary center of DMAPP is attacked by the electron-rich aromatic ring with a concerted displacement of pyrophosphate to form the arenium ion intermediate, which re-aromatizes by deprotonation to form the final product.¹¹⁵⁻¹¹⁶ For the reverse prenylation, the nucleophilic attack takes place at the tertiary center instead of the primary center of the dimethylallyl carbocation. In addition, SirD¹¹⁷ from *Leptosphaeria maculans* and TyrPT¹¹⁸ from *Aspergillus niger* are examples of known tyrosine O-prenyltransferases. Furthermore, there are also some members of the DMATS group that transfer dimethylallyl moieties onto other structural skeletons. Examples given in Figure 8 include PaxD,¹¹⁹ XptB,¹²⁰ and NscD.¹²¹

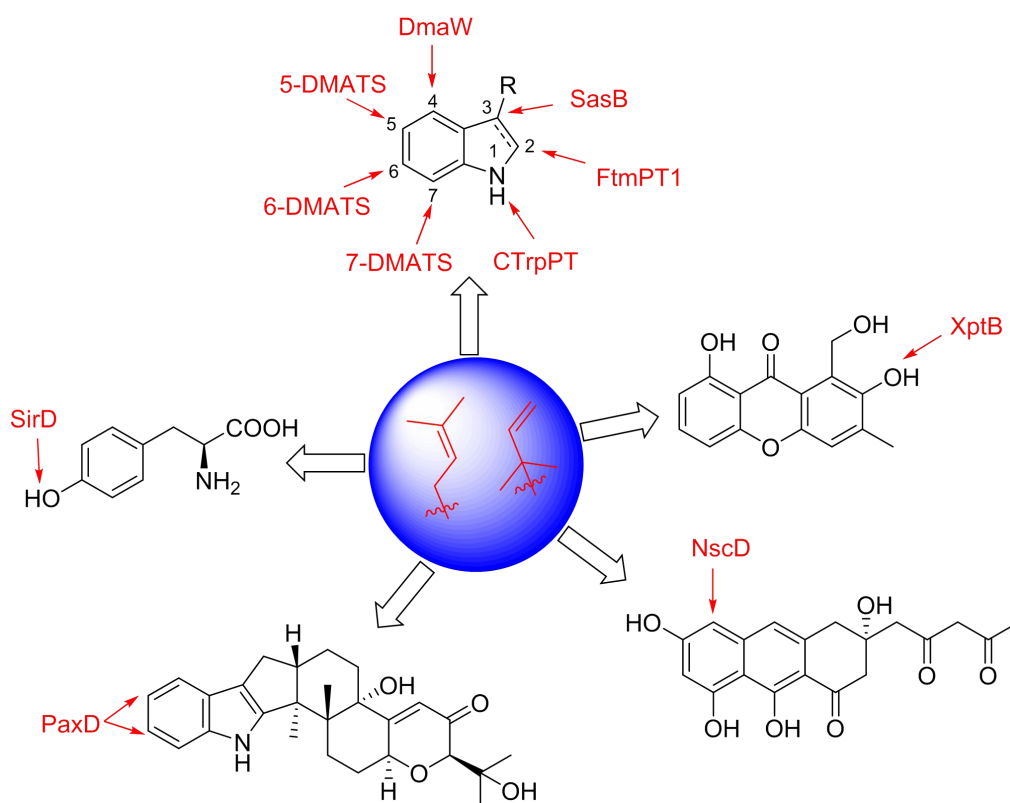


Figure 8. Examples of substrates and prenylation positions of different PTs.

In general, conversion of indole derivatives, cyclodipeptides, tyrosine derivatives, flavonoids, xanthenes and hydroxynaphthalenes to their prenylated derivatives can be catalyzed by PTs of the DMATS type.¹²² Despite the extremely high abundance of prenylated products in nature, their diversity can be further expanded by using engineered PTs for chemoenzymatic synthesis.

1.4.2.2 Cytochrome P450 enzymes

Cytochrome P450 enzymes (P450s) are a large superfamily of heme-dependent monooxygenases. They are widely described as monooxygenases because they are able to catalyze versatile

reactions that introduce oxygen into a vast range of molecules. ¹²³The term "P450" is derived from the spectrophotometric peak at the wavelength of the absorption maximum of the enzyme (450 nm) when it is in the reduced state and complexed with carbon monoxide. ¹²⁴P450s are the most versatile biocatalysts in nature and widely distributed throughout all life kingdoms. At present, more than 370,000 P450 sequences have been demonstrated (UniProt), which are distributed in all kingdoms of life such as bacteria, plants, and yeast. P450s can catalyze a vast variety of reactions, such as regio- and stereoselective oxidations of C–C and C–H bonds with oxygen under atmospheric conditions, ¹²⁵⁻¹³⁰ which is used for detoxification of xenobiotics, ¹³¹⁻¹³⁶ drug metabolism and biosynthesis of steroids.^{123,137-142}

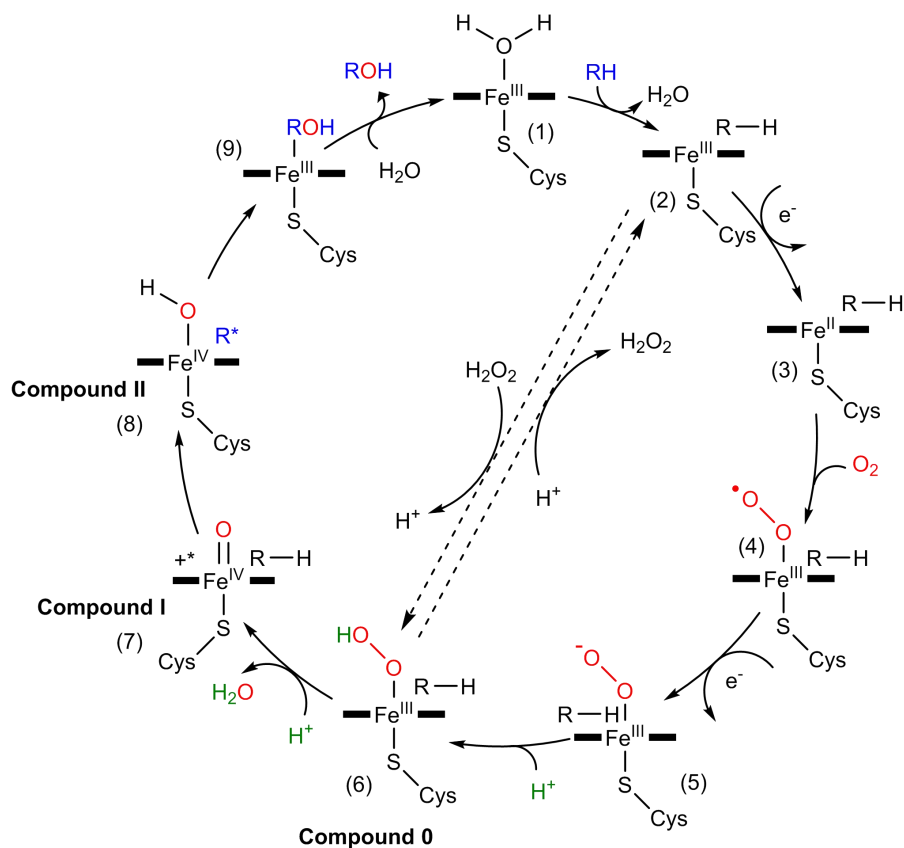


Figure 9. The P450 catalytic cycle with hydroxylation as an example.¹⁴³

The active site of cytochrome P450 enzymes contains a heme-iron center. The iron is tethered to the protein via a cysteine thiolate ligand. The typical catalytic system of P450s involves the substrate, electron shuttle carriers, NAD(P)H as electron donor and O₂ as oxidant.¹⁴⁴⁻¹⁴⁹ It employs a sophisticated, multi-step catalytic cycle involving a range of transient intermediates (Figure 9). (1) One water molecule is coordinated to the ferric heme-iron (Fe^{III}) as the sixth ligand. (2) The substrate (R-H) binds to the active site and displaces the water ligand, resulting in Fe^{III} to form high-spin Fe^{III}-RH. (3) A single electron is transferred from a NADPH (P450 reductase) and reduces the ferric (Fe^{III})

state to the ferrous (Fe^{II}) state. (4) One molecular oxygen binds to the ferrous heme iron (Fe^{II}) to form the ferrous dioxy [$\text{Fe}^{\text{II}}\text{-O}_2$] complex. (5) The second electron reduction event generates a peroxy-ferric [$\text{Fe}^{\text{II}}\text{-OO}^{2-}$] intermediate. (6) This intermediate is protonated to form the ferric hydroperoxy [$\text{Fe}^{\text{III}}\text{-OOH}$] complex (compound 0). (7) The second protonation generates a transient intermediate ($\text{Fe}^{\text{III}}\text{-OOH}_2$) and further heterolytic cleavage of the O–O bond with concurrent release of a water molecule gives rise to the transient and highly reactive ferryl-oxo intermediate [$\text{Fe}^{\text{IV}}\text{=O}$] (compound I). (8) Compound I abstracts a hydrogen atom from the substrate to form the ferryl-hydroxo compound II with a substrate radical. (9) Compound II rebounds with hydroxyl radicals to form the hydroxylated product (R-OH). The dissociation of the monooxygenated product (R-OH) from the active site and the rebound of a water molecule as the sixth heme ligand results in the regeneration of the resting state of the P450 enzyme, thus completing the catalytic cycle. Specially, some substrate-P450 complexes can directly convert into compound 0 by using H_2O_2 as the sole electron and proton donor, termed the peroxide shunt pathway.

P450 are one of the most commonly used and versatile enzymes, catalyzing multiple reactions to modify different NP scaffolds.¹⁴³ The types of P450 transformations known to occur on structurally diverse NP backbones can be classified into the following categories (Figure 10).¹⁴³ (1) Oxygenation: Hydroxylation of an aliphatic carbon is the prototypical transformation catalyzed by P450s, but P450s can also oxygenate carbons by epoxidation and aromatic hydroxylation. These P450 oxidative reactions are very often stereo- and regioselective, thereby preserving configuration.¹⁴³ (2) Dehydrogenation: P450s can catalyze the transfer of two or more electrons, converting an sp^3 hybridized carbon into an sp^2 or sp hybridization state. It also catalyzes the dehydro coupling of molecules. (3) Other transformations: P450s can catalyze rare rearrangements, the installation of oxygen on unactivated carbons, the formation of ether linkages with pre-installed oxygens. They are also able to form C - N and C - S bonds and oxidative decarboxylation.

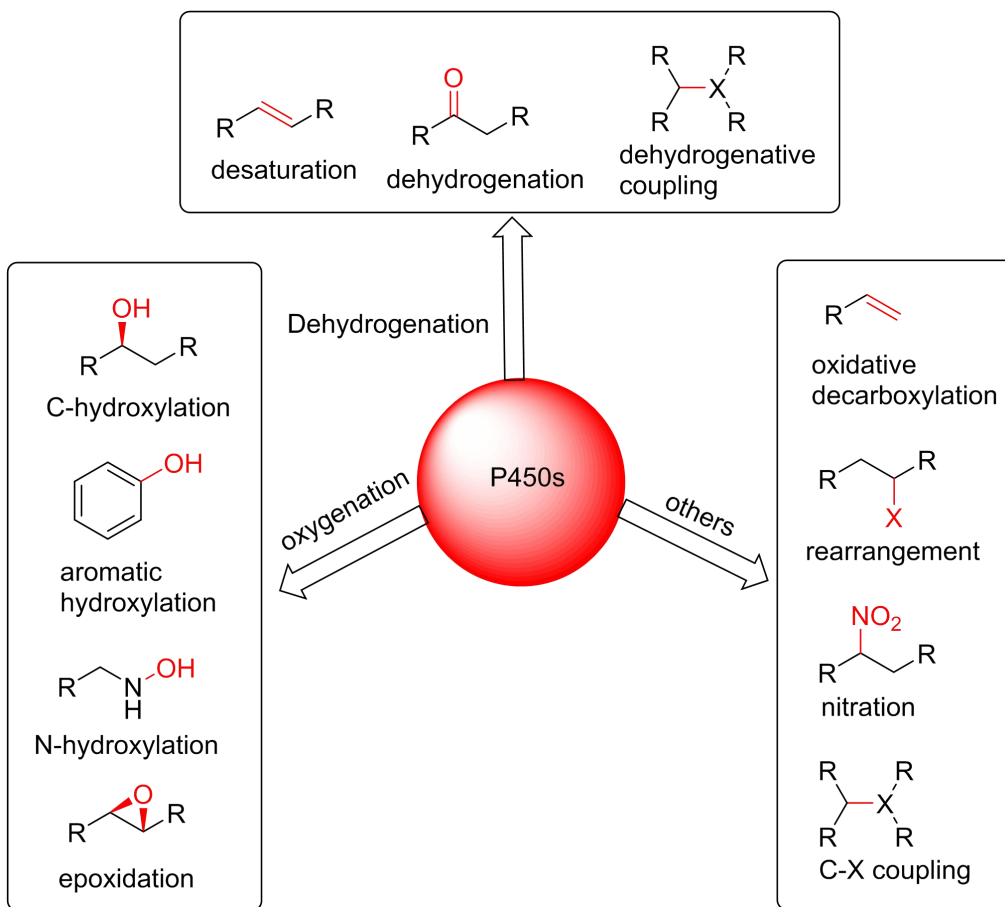


Figure 10. Diverse transformations catalyzed by P450s in NP biosynthesis.

2 Aims of this thesis

In this thesis, the following issues have been addressed:

Biosynthesis of ustethylins in *Aspergillus ustus*

Phenethyl-containing NPs are rare microbial metabolites. The ethyl group in these compounds usually originates from propionate in bacteria, while L-methionine has been implicated in the formation of some fungal metabolites. However, the responsible genes/enzymes for the methylation have not been reported prior to this study. Hence, the aim of this project is to elucidate the biosynthetic pathway of a phenethyl derivative and to identify the enzymes responsible for the methylation. The following experiments were carried out in cooperation with Dr. Liujuan Zheng:

- Identification and bioinformatic analysis of the *utt* gene cluster
- Deletion of the genes in the ustethylin cluster (*uttA-J*) by using the split marker strategy
- Heterologous expression of the PKS genes *uttA*, and NRPS-like gene *uttJ* in *Aspergillus nidulans* LO8030
- Isolation and structure elucidation of ustethylin derivatives from the wild-type (WT) strain, deletion mutants, and heterologous expression mutants
- Feeding with ¹³C-labeled precursors in WT strain
- Feeding biosynthetic precursors in $\Delta uttA$, $\Delta uttD$, and $\Delta uttJ$ mutants

Expanding structural diversity of prenylated CDPs and elucidation of the streptoazine biosynthetic pathway in *Streptomyces aurantiacus*

2,5-Diketopiperazine (DKP) alkaloids with an indole or indoline ring and isoprenoid moieties are usually derived from tryptophan-containing cyclodipeptides (CDPs). They represent an important class of hybrid natural products and display diverse biological and pharmacological activities, including antibacterial, anti-tumor, anti-inflammatory and insecticidal effects. Prenylation by prenyltransferases (PTs) at different positions of the indole ring of tryptophan-containing CDPs plays a key role for structural diversification of indole alkaloids and is involved in the biosynthesis of a large number of CDP derivatives. PTs from bacteria and fungi are usually highly permissive and can use structurally distinct compounds for prenylation. Until now, only two PTs from CDPS-dependent pathways have been described. To investigate more PTs from CDPS-dependent pathways, the following experiments were carried out in cooperation with Dr. Jing Liu:

- Bioinformatic and phylogenetic analysis to identify the putative *sas* gene cluster

- Functional proof of the gene cluster for the biosynthesis of streptoazine
- Large scale fermentation and isolation of the streptoazine
- Substrate promiscuity of SasB and generation of diverse streptoazine derivatives
- Structural elucidation of the streptoazine derivatives by 1D and 2D NMR
- Incubation in deuterium-enriched conditions to prove the non-enzymatic formation of streptoazine

Biosynthesis of various C2-guaninylated guanitrypmycin analogs by a *Streptomyces* cytochrome P450 enzyme

Cytochrome P450 enzymes have attracted significant attention in recent years, because they can catalyze a wide range of interesting chemical transformations. In the featured CDPS-related biosynthetic pathways, P450s were identified to catalyze various intriguing reactions, including intramolecular C – C bond formation, different types of dimerization, aromatization of the DKP ring, as well as nucleobase transfer reactions. Bioinformatic and phylogenetic analysis revealed the presence of a *cdps-p450* gene cluster, named *gut₁₅₂₁* from *Streptomyces* sp. NRRL S-1521. The cytochrome P450 enzyme GutD₁₅₂₁ shows high identities (approximate 51–63%) to the known GutDs and GtmD that function as nucleobase transferases. However, this candidate was located in a separate subclade in the phylogenetic tree based on the characterized P450s. Therefore, it could be involved in the biosynthesis of novel DKP derivatives. The following experiments were carried out in cooperation with Dr. Jing Liu:

- Identification and bioinformatic analysis of the *gut₁₅₂₁* gene cluster
- Heterologous expression of CDPS gene *gutA₁₅₂₁* in *Escherichia coli* BL21 (DE3)
- Heterologous expression of *gut₁₅₂₁* gene cluster in *Streptomyces albus* J1074
- Big-scale cultivation and isolation of new DKP products
- Generation of diverse guanitrypmycin analogs by biotransformation
- Cultivation of the *S. albus* transformants in media containing ¹⁵NH₄Cl
- Isolation and structural elucidation of guanitrypmycin D analogs
- Antibacterial assays of generated guanitrypmycin analogs

3 Results and discussion

3.1 Biosynthesis of ustethylins in *Aspergillus ustus*

Although the SMs of *Aspergillus ustus* are abundant, the biosynthesis of only a few of them has been reported, including phenethyl-containing derivatives. The phenethyl units in NPs are products of polyketide synthases (PKSs), which was proved in some cases by feeding experiments and genetic studies. The ethyl groups in the phenethyl residue of bacterial metabolites are mostly originated from propionate as starter unit of PKSs.¹⁵⁰⁻¹⁵¹ In fungi, it can be derived from acetate, which was confirmed by feeding with [1, 2-¹³C] acetate.¹⁵² However, most of the methyl groups of the phenethyl residue in fungal metabolites are derived from S-adenosyl L-methionine (SAM), which has been proven by feeding experiments with [methyl-¹³C]-L-methionine.¹⁵³⁻¹⁵⁴ The responsible enzymes for the methylation and the biosynthetic pathways for such metabolites have not been reported prior to this study. In this study, we elucidated the biosynthetic pathway of phenethyl derivatives in *A. ustus* and identified the enzymes responsible for the methylation by gene deletion, isotope-labeling experiments, and heterologous expression.

Ustethylin A (**1**) was the major product detected by HPLC analysis of an EtOAc extract from *A. ustus* cultures in PD medium. During the purification, the amount of **1** was significantly reduced, which suggests that **1** might be instable. Dissolution of the final isolated sample in DMSO-*d*₆ resulted in an immediate precipitation. The ¹H NMR spectrum of the sample supernatant was very complex and thus uninterpretable. To overcome this instability, we used acetylation to immediately convert **1** in the fungal extract to its triacetate **2** for structural elucidation (Figure 11).

To elucidate the origin of **1**, we carried out feeding experiments with isotope-labeled precursors in *A. ustus*. In the ¹³C NMR spectra of the acetylated product **2** after feeding with sodium [1,2-¹³C] acetate, four signal pairs of coupling carbons, C-1/C-7, C-2/C-3, C-4/C-5, and C-6/C-9, were detected, proving unequivocally the incorporation of four intact acetate units. After feeding with [2-¹³C] acetate and [2-¹³C] malonic acid, significantly increased intensities were observed for the signals of C-2, C-4, C-6 and C-7. Feeding sodium [1-¹³C] acetate clearly increased intensities were observed for the signals of C-1, C-3, C-5 and C-9. All these proved the incorporation of four acetate units. To determine whether *A. ustus* utilizes propionate as a starter unit, sodium [2-¹³C] propionate was fed into the culture. To our surprise, the labeling pattern of **2** is very similar to that of feeding with [1-¹³C] acetate. These results proved unequivocally that sodium [2-¹³C] propionate was not directly utilized for incorporation, but was degraded to acetyl-CoA likely *via* pyruvate by α -oxidation.¹⁵⁵ Acetyl-CoA was converted to malonyl-CoA and incorporated in **1**. After feeding with [methyl-¹³C]-L-methionine,

RESULTS AND DISCUSSION

the three signals of C-8, C-10, and C-11 were enhanced, proving that the methyl group of the phenethyl residue is derived from SAM (Figure 12).

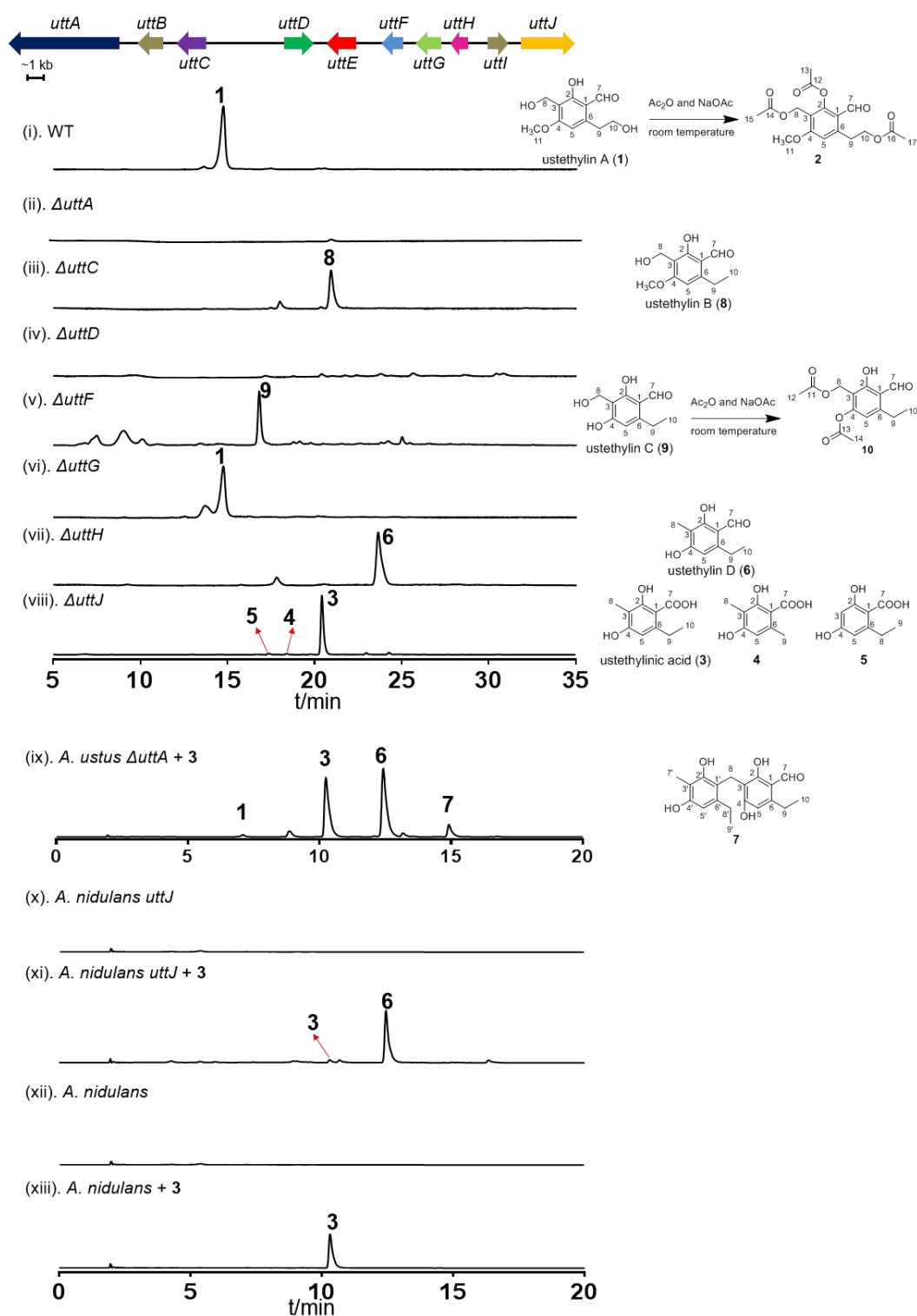


Figure 11. The *utt* gene cluster and HPLC analysis of SMs from *A. ustus*.

RESULTS AND DISCUSSION

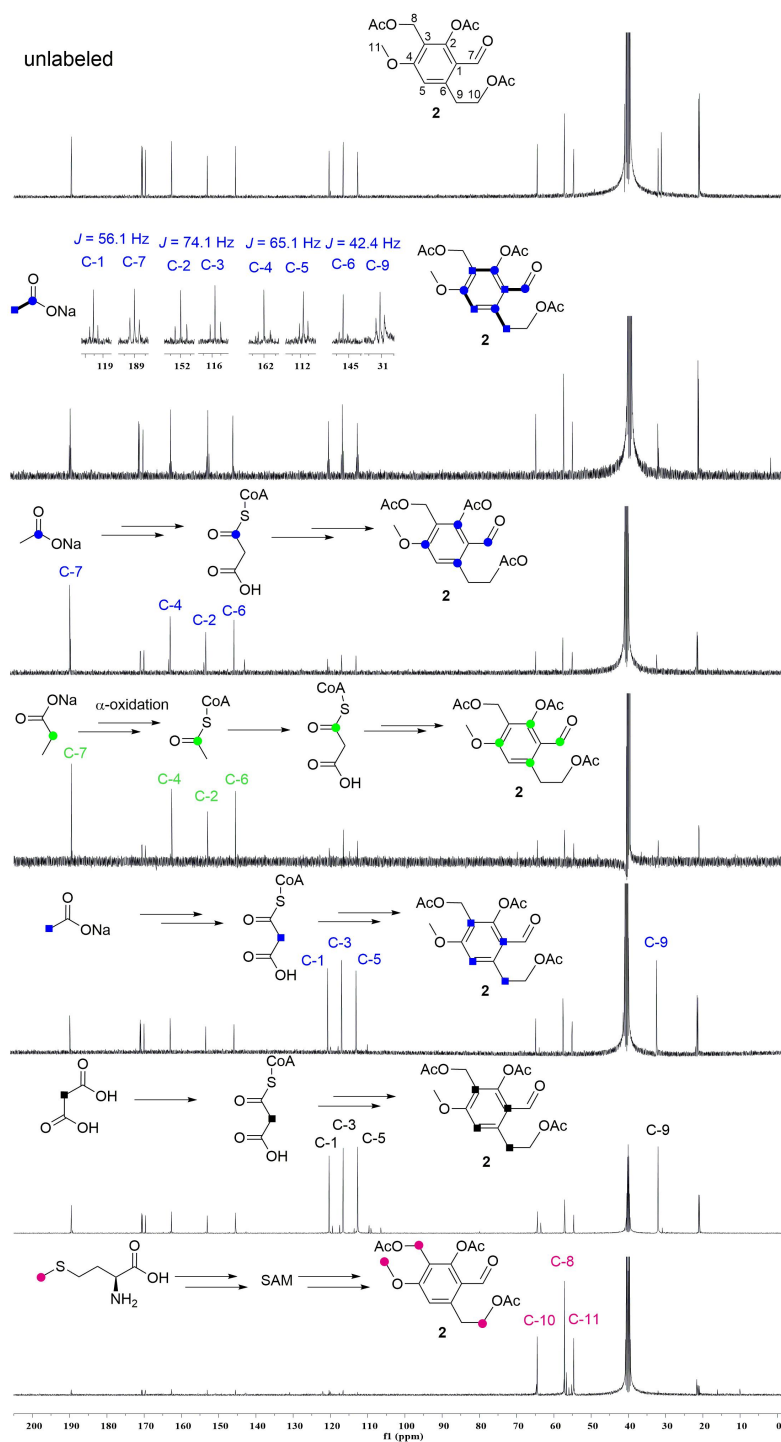


Figure 12. ¹³C NMR spectra of the labeled and unlabeled 2.

The labeled carbons after feeding with sodium [1-¹³C] acetate are highlighted with filled blue circles, sodium [2-¹³C] propionate with filled green circles, sodium [2-¹³C] acetate with filled blue squares, [2-¹³C] malonic acid with filled black squares, sodium [1,2-¹³C] acetate with bold bonds for intact acetate unit, and [methyl-¹³C]-L-methionine with filled pink circles.

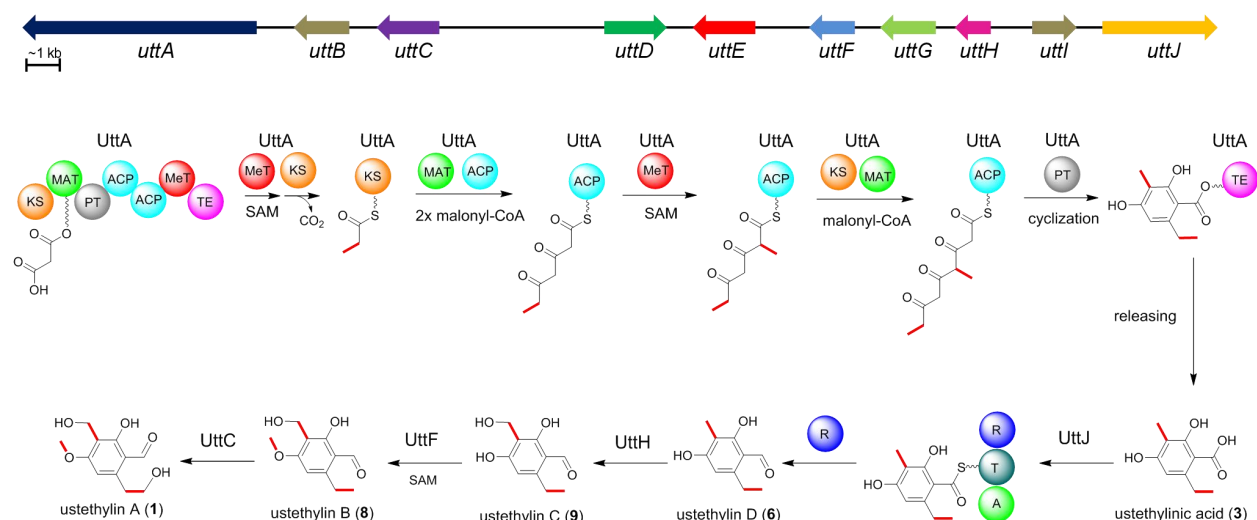
From the structure of **1**, it can be deduced that a PKS would be responsible for the formation of its backbone. Transcriptome analysis indicated that the PKS gene coding for KIA75596, termed UttA in this study, was one of the eighty best expressed genes under our culture conditions.

To prove its function, *uttA* was replaced with a hygromycin B resistance cassette by using a split marker gene replacement protocol. HPLC analysis of the culture extract of a Δ *uttA* mutant showed complete loss of **1** production. Afterward, *uttA* was heterologously expressed in *A. nidulans* LO8030. In comparison to the secondary metabolite profile of the negative control, three additional products **3** – **5** were detected in the *uttA* overexpression transformant with **3** as the predominant peak (98%). They differ from each other only in the methyl group at C-3 and the ethyl group at C-6 of the benzene ring, indicating multimethylation steps during the formation of **3** (Figure 11).

Feeding **3** into the *A. ustus* Δ *uttA* mutant led to detection of one additional major peak **6** and one minor peak **7**. Structure elucidation confirmed **6** to be the corresponding aldehyde of **3** (ustethylin D) and the minor peak **7** as a dimerization product (Figure 11, Scheme 1). Trace amounts of **1** were also detected in this culture, proving **3** to be a precursor of **1**. It can be speculated that the metabolism of **6** is a limited step in the biosynthesis. Interestingly, **3** was not detected in the Δ *uttA* mutant after feeding with **4**, excluding the direct methylation of the C6-methyl group in **4**.

1 differs structurally from **3** in the oxidation states of the functional groups at C-1, C-3, and C-6, as well as O-methylation at OH-4. The conversion of **3** to **1** would require three oxidoreductases and an O-MeT. Deletion of *uttJ* abolished the formation of **1** and led to accumulation of **3** – **5**, which resembles the product profile of the *A. nidulans* *uttA* expression strain. This unambiguously proved its role in the reduction of the carboxyl group to an aldehyde. Feeding **3** into the *A. nidulans* *uttJ* overexpression strain led to the detection of **6**, proving UttJ as an aryl acid reductase (Figure 11). Further sequence comparison and analysis revealed UttH to be a putative nonheme Fe^{II}/2-oxoglutarate dependent oxygenase. Deletion of *uttH* led to the accumulation of UttJ product **6**, proving the reaction order of both enzymes. Deletion of *uttC* coding for a cytochrome P450 enzyme abolished the formation of **1** and production of **8** (ustethylin B), which differs from **1** just in the oxidation state of the ethyl group (Figure 11). This proved that UttC catalyzed the last step in the biosynthesis of **1**. Detection of **8** with a methoxy group in Δ *uttC* mutant indicates that the methylation of the C4-hydroxyl group occurs before UttC and after UttH reactions and could be catalyzed by the putative O-MeT UttF. Similar to **1**, **9** was also found to be unstable and could not be obtained in pure form for structure elucidation. However, its structure can be elucidated after conversion to its diacetylated derivative **10** (Figure 11).

Gene deletion results revealed the reaction sequence of the tailoring enzymes for the conversion of **3** to **1**. Extracted ion chromatograms of the culture proved the presence of **1** as almost the only pathway product, indicating the high efficiency of the involved enzymes in the wildtype *A. ustus*. The *utt* cluster is positively regulated by a DNA binding enzyme UttD. Deletion of *uttD* completely abolished product formation. Even feeding **3** to the Δ *uttD* deletion mutant did not lead to any conversion. Deletion of *uttG* coding for an MFS transporter reduced production of **1** to 30.8% of that of the wildtype *A. ustus*. **1** was still detected in the deletion mutants of the two oxidoreductase genes *uttB* and *uttE*. They very likely are not involved in the formation of ustethylin A (Scheme 1).



Scheme 1. Biosynthetic pathway of ustethylins in *A. ustus* and *utt* gene cluster.

In summary, we have identified the biosynthetic gene cluster of the highly oxygenated arylaldehyde derivative ustethylin A in this study and elucidated its biosynthetic pathway by gene deletion, expression, and isotopic labeling experiments. The PKS UttA is responsible for the formation of the phenethyl core structure with methylation as a key reaction. Consecutive and coordinated modifications by three different types of oxidoreductases and one O-MeT led to the final product. To the best of our knowledge, this is the first report on the biosynthetic pathway of a phenethyl-containing fungal metabolite.

[For details on this work, please see the publications \(sections 4.1\)](#)

Liujuan Zheng*, **Yiling Yang***, Haowen Wang, Aili Fan, Liping Zhang, and Shu-Ming Li (2020). Ustethylin biosynthesis implies phenethyl derivative formation in *Aspergillus ustus*. *Organic Letters*. 22, 7837-7841, DOI:10.1021/acs.orglett.0c02719. (*equal contribution)

3.2 Expanding structural diversity of prenylated CDPs and elucidation of the streptoazine biosynthetic pathway in *Streptomyces aurantiacus*

Significant progress has recently been achieved regarding the understanding of the biosynthesis of prenylated CDPs and derivatives thereof, especially of those from fungi of the genera *Penicillium* and *Aspergillus*.^{91, 95, 156-158} 2,5-Diketopiperazine (DKP) alkaloids with an indole or indoline ring and isoprenoid moieties are usually derived from tryptophan-containing cyclodipeptides (CDPs).^{42, 94, 159} They represent an important class of hybrid NPs and display diverse biological and pharmacological activities, including antibacterial, anti-tumor, anti-inflammatory and insecticidal effects.^{42, 159-160} Prenylation by prenyltransferases (PTs) at different positions of the indole ring of tryptophan-containing CDPs plays a key role for structural diversification of indole alkaloids and is involved in the biosynthesis of a large number of CDP derivatives.^{94, 156} PTs from bacteria and fungi are usually highly permissive and can use structurally distinct compounds for prenylation.^{95, 161}

To investigate more PTs from CDPS-dependent pathways, we analyzed a wide range of *cdps*-containing clusters by using characterized proteins as probes and identified a candidate from *S. aurantiacus* NRRL ISP-5412. The cluster of interest, termed the *sas* cluster, consists of three open reading frames coding for a putative CDPS (SasA) and two tailoring enzymes, a PT (SasB) and a MT (SasC). To verify their functions, we first cloned the *cdps* gene *sasA* from *S. aurantiacus* into expression vector pPWW50A¹⁶² and expressed it in *Streptomyces albus* J1074. One major product **11**, was detected with a $[M + H]^+$ ion at m/z 373.1659 compared to the host strain J1074 carrying pPWW50A (Figures 13i – 13ii). Compound **11** was identified as *cyclo*-(L-Trp-L-Trp) (cWW) by comparison with an authentic standard, proving SasA to be a cWW synthase. Then, the whole gene cluster comprising *sasABC* was cloned into pPWW50A and overexpressed in J1074 as described above. In addition to the predominant **11**, four new additional compounds were observed (Figure 13iii). The second dominant product **12** was detected with a $[M + H]^+$ ion at m/z 537.3224, 164 Da larger than that of cWW, indicating the attachment of two prenyl and two methyl groups to **11**. The three minor compounds **13**, **14**, and **15** with $[M + H]^+$ ions at 441.2285, 509.2911, and 523.3068, are 68, 136, and 150 Da larger than **11**, implying one prenyl, two prenyl, and two prenyl moieties plus one methyl group in their structures, respectively. Compound **12** was then isolated by semi-preparative HPLC after large scale fermentation. Comprehensive interpretation and comparison of the ¹H NMR data as well as the ECD spectrum with those reported in the literature¹⁶³ confirmed **12** to be streptoazine C (Scheme 2). These data strongly support the function of SasB as a regular C-3-prenyltransferase and SasC as an indoline N-methyltransferase. Due to the low product yields, **13** – **15** could not be isolated from the *sasABC* transformant for structure elucidation by NMR analysis.

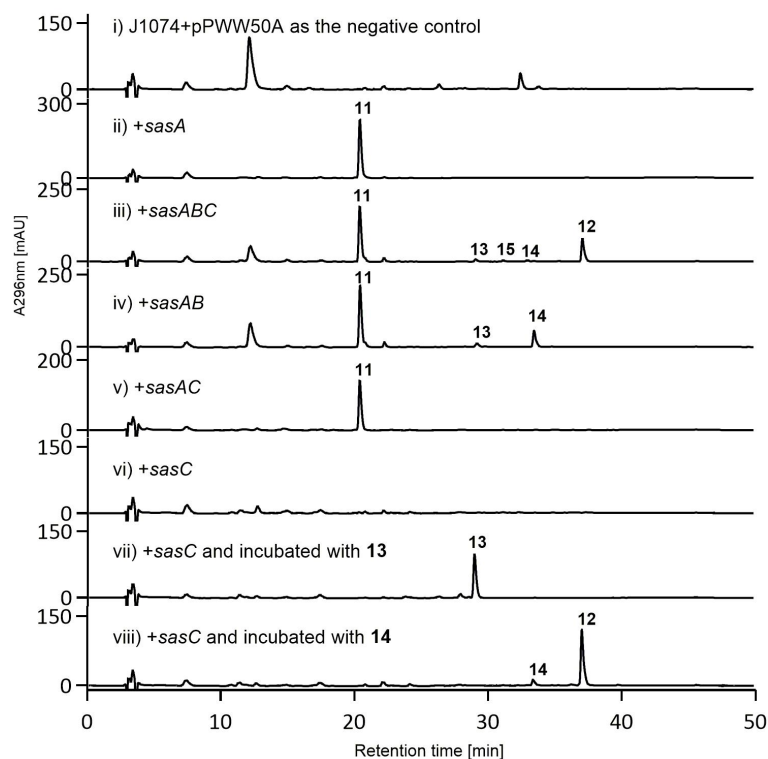
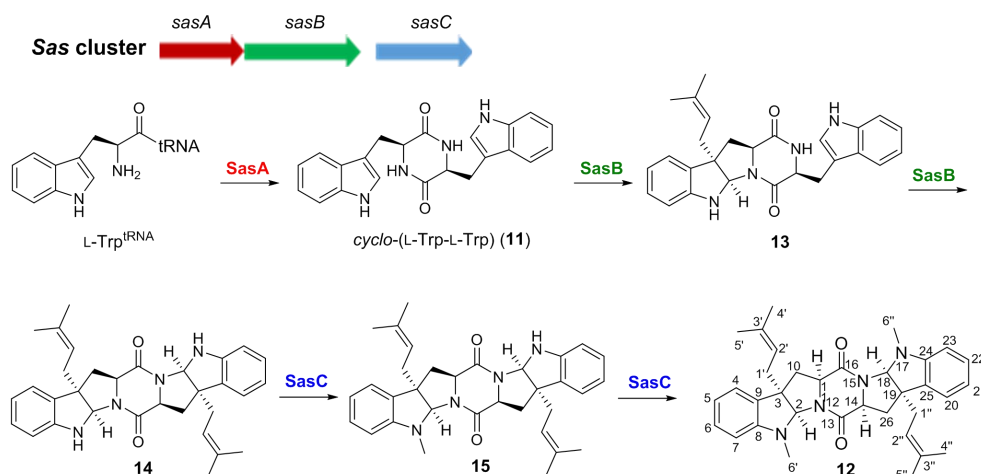


Figure 13. LC-MS analysis of *S. albus* J1074 transformants with and without precursor incubation.

To confirm the functions of SasB and SasC and to figure out the reaction order, we co-expressed *sasA* with *sasB* and *sasC* separately to yield *sasAB* and *sasAC*. In addition to the CDPS product **11**, two additional compounds **13** and **14** were detected in the fermentation culture of the *sasAB* transformant (Figure 13iv). Isolation and structure elucidation by MS and ^1H NMR analyses as well as comparison with the data of known compounds¹⁶³ confirmed **13** and **14** to be regularly C3-monoprenylated cWW and streptoazine A, respectively (Scheme 2). In contrast, only **11** was observed in the culture of the *sasAC* transformant (Figure 13v). Neither mono-, nor dimethylated **11** was detected in the *sasABC* transformant, even in the sensitive extracted ion chromatograms (EICs). These results supported that **11** cannot be methylated by the methyltransferase SasC and prenylation takes place before methylation. Incubating the *sasC* transformant with **14** led to the clear detection of **12** (Figure 13vi), whereas no new peaks were observed in the culture after incubation with **13** (Figure 13vii). This demonstrated that methylation proceeds only after the attachment of two prenyl moieties (Scheme 2). This reaction order is the same as for the recently reported two-gene cluster responsible for streptoazine C biosynthesis.¹⁶³



Scheme 2. The *sas* gene cluster and biosynthetic pathway of streptoazine C in *S. aurantiacus*.

To further verify that the formation of **13** and **14** are catalyzed by SasB, its coding sequence was cloned into pPWW50A and expressed in J1074. Neither **11** nor other additional metabolites were observed in the *sasB* transformant (Figure 14Aiii). **13** and **14** were identified from incubating the *sasB* transformant with **11** (100 μ M) and cultivation for 5 days (Figure 14Aiv), whereas no consumption of **11** was found in the control culture (Figure 14Aii). These results demonstrated that SasB is able to catalyze the regular C-3 prenylation of **11**.

In addition, we investigated the substrate specificity of SasB. *In vitro* testing of the acceptance of tryptophan-containing CDPs by SasB was not possible, as recombinant proteins could not be obtained after heterologous expression in both *E. coli* and *Streptomyces*. Thus, we supplied 12 CDPs to the J1074 transformant with *sasB* and monitored their consumption by LC-MS analysis. As shown in Figures 14Av – 14Aviii, *cyclo*-(L-Trp-L-Phe) (cWF), *cyclo*-(L-Trp-L-Tyr) (cWY), *cyclo*-(L-Trp-L-Leu) (cWL), and *cyclo*-(L-Trp-L-Met) (cWM) were efficiently converted by SasB. The $[M + H]^+$ ions of the products **16** – **19** are 68 daltons larger than those of the precursors, indicating the attachment of one dimethylallyl moiety to the substrates. Products **16** – **19** were obtained by large-scale fermentation and subsequent separation by preparative HPLC, and their structures were analyzed by NMR, including ^1H , ^{13}C , COSY, HSQC, HMBC, and NOESY. The typical signals of a regular C-3-prenyl residue in the ^1H NMR spectra are found in the ranges of δ_{H} 2.44 – 2.46 (d, 6.9 – 7.2 Hz, H-1'), 5.01 – 5.03 (t, 6.9 – 7.2 Hz, H-2'), 1.57 – 1.58 (s, H-4'), and 1.63 – 1.64 (s, H-5'). The signals of the five carbons are detected in the ^{13}C spectra at about δ_{C} 34 (C-1'), 120 (C-2'), 134 (C-3'), 18 (C-4'), and 26 (C-5'). Prenylation at C-3 abolishes the aromatic character of the indole system and causes a shielded shift of the H-2 signal to δ_{H} 5.24 – 5.26 as well as those of C-2 and C-3 to δ_{C} 80 and 55, respectively. The configuration of the products was determined based on the correlations

between H-1' and H-11, H-1' and H-2 as well as H-2 and H-11 in the NOESY spectra. Comparison of their ECD spectra provided additional evidence for their configurations. All the obtained data confirmed that **16** – **19** are C-3-prenylated derivatives of the corresponding CDPs (Figure 9B). Low conversions to prenylated derivatives were also detected by LC-MS analysis for *cyclo*-(L-Trp-L-Ala) (cWA), *cyclo*-(L-Trp-D-Ala), *cyclo*-(D-Trp-L-Ala), *cyclo*-(D-Trp-D-Ala), *cyclo*-(L-Trp-L-Pro) (cWP), *cyclo*-(L-Trp-D-Pro), *cyclo*-(D-Trp-L-Pro), and *cyclo*-(D-Trp-D-Pro). Due to the low product yields, the structures of these products could not be elucidated in this study. These results suggest a more flexible substrate specificity of SasB from *S. aurantiacus* than that of SazB from *S. leeuwenhoekii*.¹⁶³ It was reported that cWF, cWY, cWA, and cWP were not accepted by SazB.¹⁶³ In our case, all of these four CDPs were prenylated by SasB with high conversions for cWF and cWY (Figure 14).

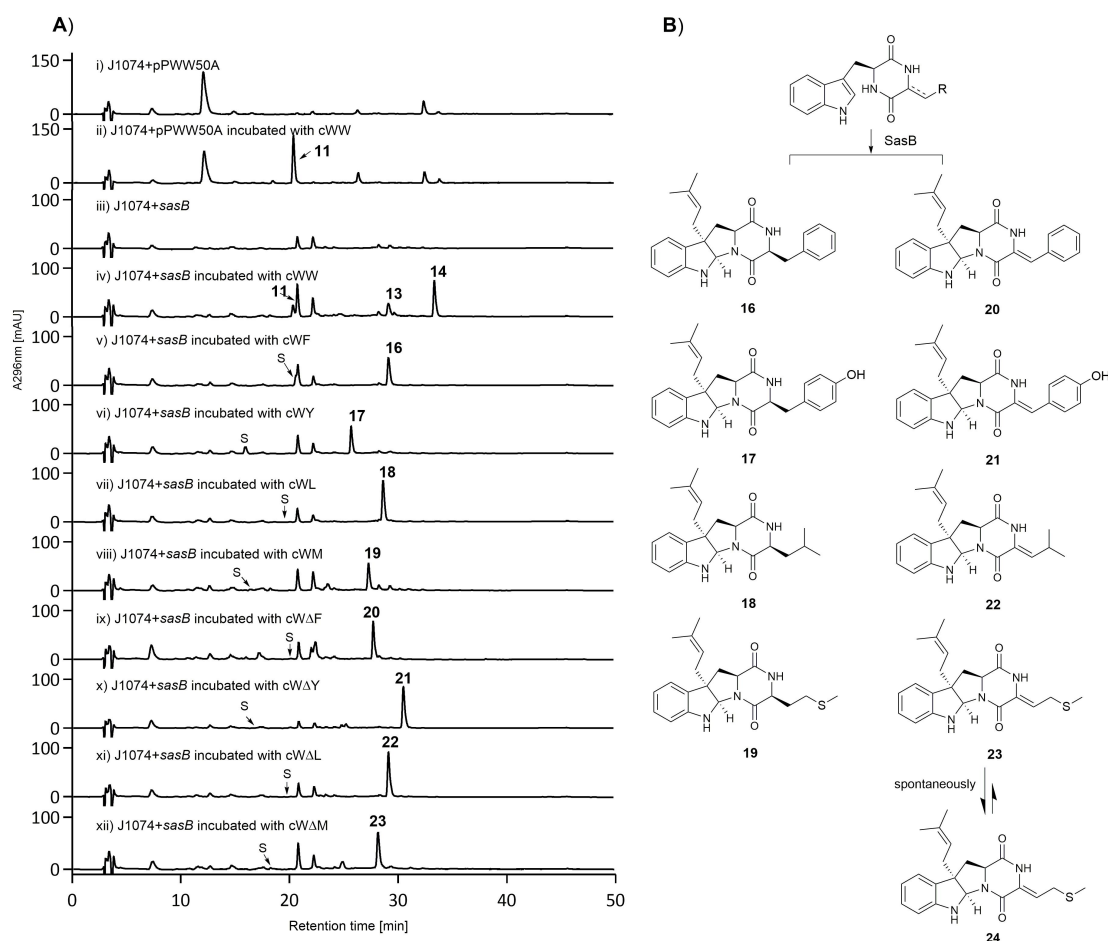


Figure 14. A) HPLC analysis of the *sasB* transformant with and without precursors and B) prenylated products of SasB. S: substrate.

Cyclodipeptide oxidases (CDOs) are frequently found in the CDPs-related pathways and install *exo* double bonds at the DKP ring.¹⁶⁴ For combinatorial application of SasB with these oxidases, we

tested its acceptance of the dehydrogenated forms of the four efficiently converted CDPs, *i.e.* *cyclo*-(L-Trp- Δ Phe) (cW Δ F), *cyclo*-(L-Trp- Δ Tyr) (cW Δ Y), *cyclo*-(L-Trp- Δ Leu) (cW Δ L), and *cyclo*-(L-Trp- Δ Met) (cW Δ M), by incubation experiments in the *sasB* transformant. LC-MS analysis showed that all of these compounds were good substrates for SasB and were completely converted to their prenylated products (Figures 14Aix – 14Axii). The products **20** – **23** were subsequently isolated and their structures were confirmed to be regularly C-3 prenylated derivatives at the indoline ring (Figure 14B) by detailed interpretation of their NMR data and the comparison with the data of **16** – **19**. Observation of the interaction between NH-15 and H-19/H-23 in the NOESY spectrum of **21** as well as NH-15 and H-18 in that of **22** supported the *Z*-configuration of the *exo* double bonds in their structures.

During the isolation procedure, we observed the conversion of the cW Δ M product **23** to a new compound **24**. Isolation by using a chiral-phase HPLC column and structure elucidation by NMR analysis including interpretation of the NOESY data and comparison of its ECD spectrum with that of **23** confirmed the epimerization at the C-11 position. As the nonenzymatic epimerization *via* keto-enol tautomerism was already observed for the guanitrypmycins,¹⁶⁵ we speculated a similar mechanism may explain the conversion of **23** to **24** (Figure 14). Incubation of **23** in CD₃OD/D₂O (1:1) at pH 9 and 12 for 14 h and LC-MS analysis confirmed indeed the conversion of **23** to **24** and incorporation of one deuterium in both molecules. This supported the epimerization at C-11 *via* keto-enol tautomerism.

In conclusion, we elucidated the biosynthetic pathway of streptoazine C from *S. aurantiacus* by heterologous gene expression in *S. albus* and precursor incubation experiments. More importantly, the prenyltransferase SasB displays a remarkably high substrate tolerance and accepts not only a number of tryptophan-containing CDPs, but also their dehydrogenated derivatives for rare regular C3 α -prenylation at the indole ring. This study provides an enzyme with a high substrate promiscuity from the less explored prenyltransferase group in cyclodipeptide synthase-related pathways and provides more details on their biochemical properties.

[For details on this work, please see the publication \(section 4.2\)](#)

Jing Liu*, **Yiling Yang***, Lauritz Harken, and Shu-Ming Li (2021). Elucidation of the streptoazine biosynthetic pathway in *Streptomyces aurantiacus* reveals the presence of a promiscuous prenyltransferase/cyclase. *Journal of Natural Products*, 84, 3100–3109. DOI: 10.1021/acs.jnatprod.1c00844. (*equal contribution)

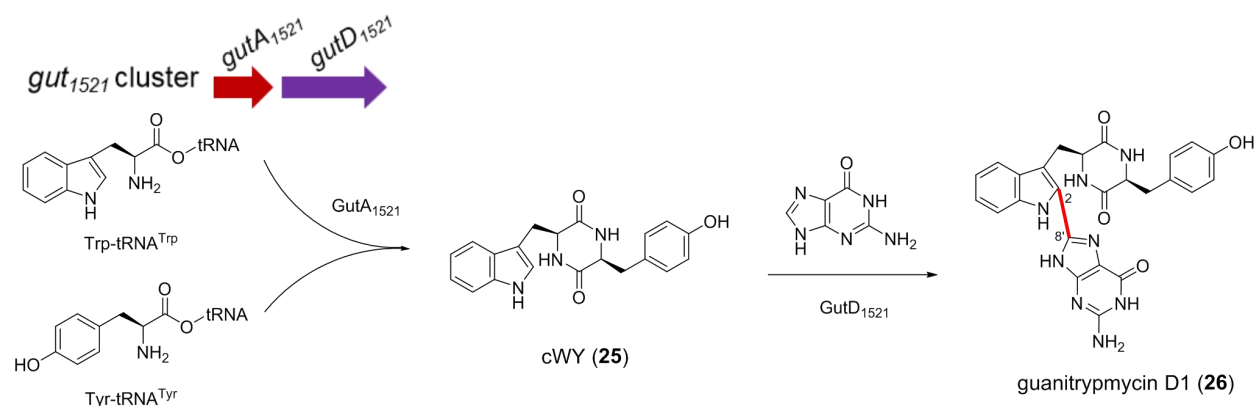
3.3 Biosynthesis of various C2-guaninylated guanitrypmycin analogs by a *Streptomyces* cytochrome P450 enzyme

Cytochrome P450s were found as the most common modification enzymes in the characterized *cdps*-related gene clusters. P450 enzymes from the featured biosynthetic pathways catalyze a wide range of interesting chemical transformations, such as intramolecular C—C bond formation, different types of dimerization, aromatization of the DKP ring, as well as nucleobase transfer reactions.^{164, 166-167} Seven different types of cyclodipeptide-nucleobase linkages have been characterized from CDPS-P450-related nucleobase transfer pathways, including C—C, C—N and C—O bonds.^{165, 168-171} Therefore, we aimed to find more CDPS related modification enzymes to expand the spectrum of CDP derivatives and obtain more novel biocatalysts.

In the previous studies, we took the functionally characterized CDPSs and P450s as probes to search and identify their putative homologs in the public databases. Subsequent phylogenetic analysis led to the identification of plenty of uncharacterized *cdps-p450* gene clusters. Based on the phylogenetic information, one *cdps-p450* gene cluster from *Streptomyces* sp. NRRL S-1521 attracted our attention. Following the nomenclature of the known clusters,^{165, 169} we named *gutA*₁₅₂₁ and *gutD*₁₅₂₁ for the *cdps* and *P450* genes, respectively.

A sole peak for **25** was identified by HPLC analysis in the *E. coli* transformant harboring *gutA*₁₅₂₁, which was not detected in the mutant with the empty vector pET28a (+) as the negative control (Figures 15i — 15ii). Compared to an authentic standard, compound **25** was characterized as cWY, which was also confirmed by its ¹H and ¹³C NMR data. This proved that the CDPS GutA₁₅₂₁ functions as a cWY synthase (Scheme 3). Based on the LC-MS data, expression of the candidate gene cluster *gut(AD)*₁₅₂₁ resulted in the production of two compounds **25** and **26**. Neither was found in the negative control with pPWW50A (Figures 15iii — 15iv). Compound **26** exhibited a [M + H]⁺ ion at *m/z* 499.1843, which is 149 Da larger than that of cWY. Thus, we deduced that an additional guanine residue was attached to cWY. Subsequently, compound **26** was isolated from a large-scale fermentation culture and its structure was further elucidated by detailed NMR analysis. Inspection of the NMR data of compound **26** revealed the presence of three characteristic ¹H signals of a guaninyl residue at δ 10.70 (H-1'), 12.98 (H-9'), and 6.33 (H-10') with five corresponding ¹³C signals at δ 151.6 (C-2'), 146.0 (C-4'), 106.0 (C-5'), 159.3 (C-6'), and 140.5 (C-8'). Although no clear correlation in the HMBC spectrum was observed between the cWY skeleton and the guaninyl moiety, the missing ¹H signal for H-2 of the cWY part and that for H-8' of guanine indicated the new C—C bond between C-2 and C-8' of the two moieties (Scheme 3). In addition, the signal of C-3 in the ¹³C spectrum of compound **26** is deshielded by 5 ppm in comparison to that in compound **25**, whereas

signals for other carbons like C-5 – C-7 are only deshielded approximately 3 ppm. As compound **26** features tryptophanyl and guaninyl residues, we named it guanitrypmycin D1. Cultivation of *S. albus* carrying *gut(AD)₁₅₂₁* in ¹⁵NH₄Cl-containing medium revealed incorporation of three and eight ¹⁵N atoms in **25** and **26**, respectively, providing additional evidence for the guanitrypmycin D1 structure. The above results implied GutD₁₅₂₁ as a new nucleobase transferase for the specific C-2 – C-8' connection between the indole and guaninyl units, differing from the previous reported P450s.^{165, 168-171}



Scheme 3. The *gut*₁₅₂₁ gene cluster and biosynthesis of guanitrypmycin D1 in *Streptomyces* sp. NRRL S-1521.

In order to confirm the GutD₁₅₂₁ function, we intended to carry out biochemical characterizations with a recombinant protein overproduced in *E. coli*. Unfortunately, no soluble GutD₁₅₂₁ was obtained. Therefore, we cloned it into pPWW50A for expression in *S. albus* J1074, followed by precursor incubation experiments. Supplementation of compound **25** to the *gutD*₁₅₂₁ transformant led to the production of **26**, while no additional metabolites were detected in J1074 harboring pPWW50A (Figures 15v – vii). These data proved unequivocally GutD₁₅₂₁ as a specific C-2 – C-8' guaninyl transferase, and being responsible for the generation of a new guaninylated DKP (Scheme 3). Feeding of *cyclo*-(D-Trp-L-Tyr) to the *gutD*₁₅₂₁ transformant did not lead to any conversion, proving the importance of the L-configuration of the tryptophanyl moiety for acceptance by GutD₁₅₂₁.

RESULTS AND DISCUSSION

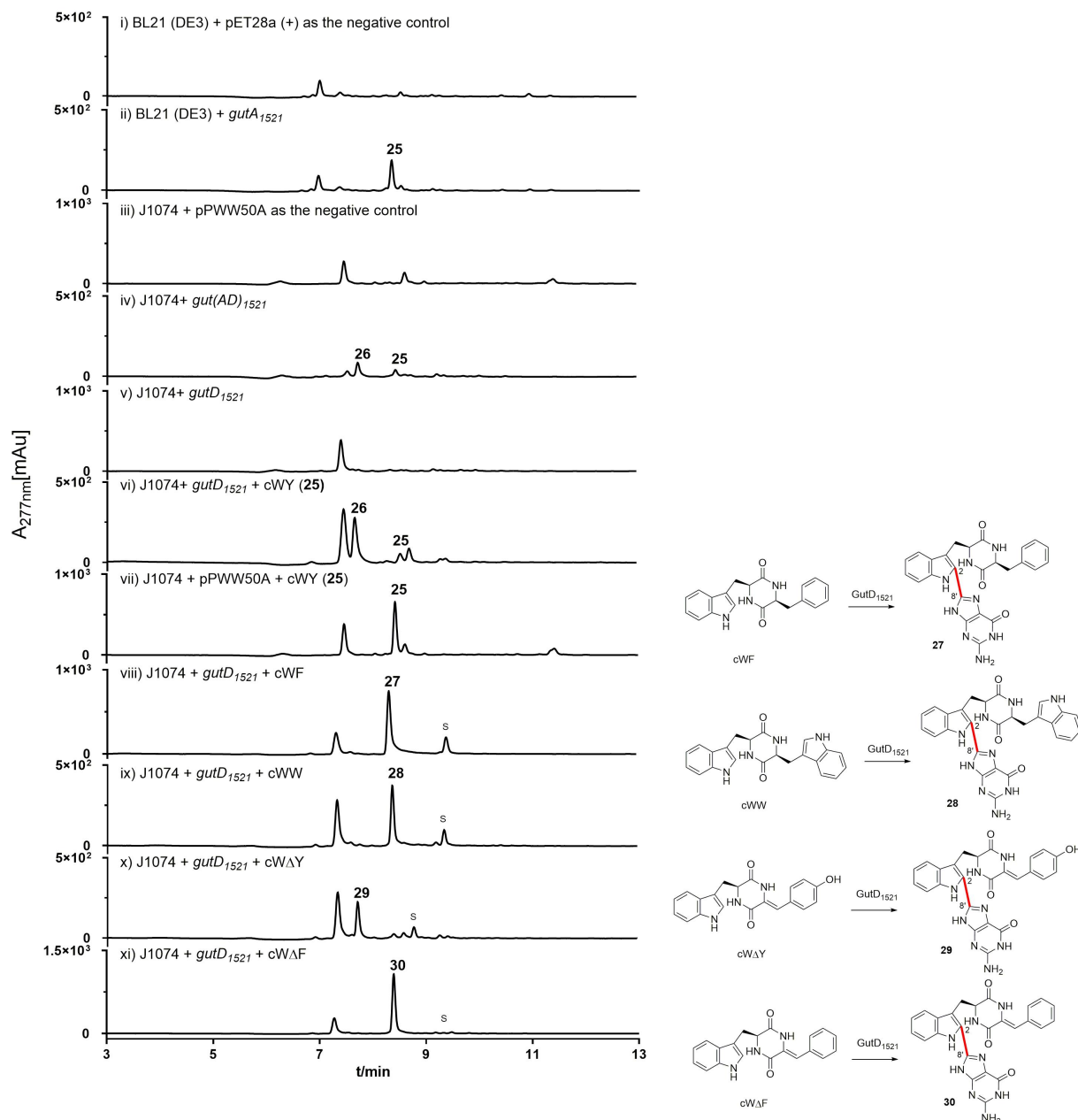


Figure 15. HPLC analysis of the generated transformants and the *gutD*₁₅₂₁ transformant with/without precursor feeding. S: substrate.

Based on the GutD₁₅₂₁ function, we investigated its substrate specificity toward other tryptophan-containing CDPs including cWA, cWF, *cyclo*-(L-Trp-L-His) (cWH), cWL, cWM, cWP, and cWW. After supplying these CDPs into the J1074 transformant harboring *gutD*₁₅₂₁, the 5 day-old cultures were monitored for their substrate conversion by LC-MS. As shown in Figure 15, cWF and cWW were efficiently transformed to the products **27** and **28** (Figures 15viii – 15ix). Their [M+H]⁺ ions are 149 daltons larger than those of the corresponding precursors, indicating the attachment of a guanylyl

residue. In contrast, other CDPs like cWA, cWH, cWL, cWM, and cWP cannot be efficiently converted by GutD₁₅₂₁. Compounds **27** and **28** were subsequently isolated from the large-scale cultures and their structures were elucidated based on NMR data. The typical signals of the guaninyl moiety were clearly observed in their ¹H NMR spectra. For compound **27**, these signals are found at δ_{H} 10.63 (br s, H-1'), 12.94 (br s, H-9'), along with 6.27 (br s, H-10'). For compound **28**, they are at δ_{H} 10.80 (br s, H-1'), 12.91 (br s, H-9'), and 6.42 (br s, H-10'). Compared with compound **26**, very similar values can be assigned for the five carbons of the guanine residue in the ¹³C spectra as well. Similar to that of compound **26**, the key correlation between C-2 — C-8' was absent in the HMBC spectra. Nevertheless, the absence of the corresponding ¹H signals supported them to be C-2-guaninylated cWF (**27**, guanitrypmycin D2) and cWW (**28**, guanitrypmycin D3), respectively (Figure 15).

Because GutD₁₅₂₁ can use cWF, cWY, and cWW as substrates for coupling with guanine, we further tested its acceptance for the dehydrogenated forms of these well converted CDPs. Due to difficulties in obtaining *cyclo*-(L-Trp- Δ Trp) (cW Δ W), only cW Δ Y and cW Δ F were prepared by large-scale enzyme assays with the functionally characterized CDO Ndas_1146/1147 for biotransformation with the *gutD*₁₅₂₁ transformant.¹⁷² After incubation for 7 days, the two dehydrogenated CDPs were converted to peaks **29** and **30** (Figures 15x — 15xi). Structure elucidation by detailed interpretation of NMR data confirmed both compounds as guaninylated derivatives at the C-2 of the indole ring (Figure 15), *i.e.* guanitrypmycin D4 (**29**) from cW Δ Y and guanitrypmycin D5 (**30**) from cW Δ F. Addition of cWF, cWW, cW Δ Y, and cW Δ F to cultures of *S. albus* harboring *gutD*₁₅₂₁ containing ¹⁵NH₄Cl led to detection of compounds **27** — **30** with incorporation of five ¹⁵N atoms, respectively, further supporting their structures suggested by NMR and ECD analyses.

Additionally, we tested the antibacterial activity of the isolated compounds. Compounds **26** — **30** were subsequently tested for their antibacterial activities against *E. coli* ATCC 25922 and DH5 α , *Enterococcus faecalis* DSM2570, *Klebsiella pneumoniae* DSM26371, *Bacillus subtilis* NCIB 3610 and BSB 01, *Bacillus circulans* NRRL B-380, NRRL B-14032, and NRRL NRS-1108, *Staphylococcus aureus* ATCC 29213, *Staphylococcus delphini* DSM20771 as well as *Pseudomonas aeruginosa* ATCC 27853. Unfortunately, no inhibitory activity was observed.

Taking the results together, a two-gene cluster coding for a CDPS and a P450 was identified in *Streptomyces* sp. NRRL S-1521 by phylogenetic analysis. Heterologous expression of the gene cluster led to the identification of a new guaninylated DKP guanitrypmycin D1. Biotransformation experiments demonstrated that GutD₁₅₂₁ catalyzes the transfer of a guanine moiety onto C-2 of the indole ring of cWY *via* a C — C bond formation. Precursor incubation experiments revealed that

GutD₁₅₂₁ can also utilize other tryptophan-containing CDPs as well as their dehydrogenated forms as substrates for the synthesis of different guanitrypmycin analogs. Therefore, this study provides a biocatalyst for a new linkage pattern between a DKP indole ring and a guanine moiety and expands the functional spectrum of P450s as tailoring enzymes.

[For details on this work, please see the publication \(section 4.3\)](#)

Jing Liu*, **Yiling Yang***, Xiulan Xie, and Shu-Ming Li (2023). A *Streptomyces* cytochrome P450 enzyme catalyzes regiospecific C2-guaninylation for the synthesis of diverse guanitrypmycin analogs. *Journal of Natural Products*, 2023, 86, 94–102. DOI: 10.1021/acs.jnatprod.2c00787. (*equal contribution)

4 Publications

4.1 Ustethylin biosynthesis implies phenethyl derivative formation in *Aspergillus ustus*

Ustethylin Biosynthesis Implies Phenethyl Derivative Formation in *Aspergillus ustus*

Liujuan Zheng,[▽] Yiling Yang,[▽] Haowen Wang, Aili Fan, Liping Zhang, and Shu-Ming Li*



Cite This: *Org. Lett.* 2020, 22, 7837–7841



Read Online

ACCESS |



Metrics & More

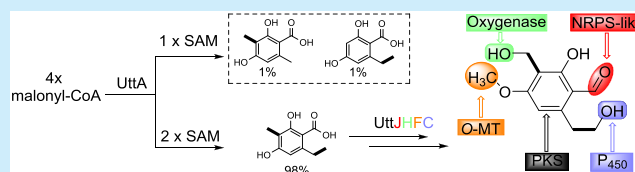


Article Recommendations



Supporting Information

ABSTRACT: A highly oxygenated phenethyl derivative ustethylin A was isolated from *Aspergillus ustus*. Gene deletion, isotope labeling, and heterologous expression proved that the phenethyl core structure is assembled from malonyl-CoA by a polyketide synthase harboring a methyltransferase domain. Propionate was converted via acetyl-CoA to malonyl-CoA and incorporated into the molecule. Modifications on the core structure by three different oxidoreductases and one *O*-methyltransferase lead to the final product, ustethylin A.



Phenethyl-containing natural products are common microbial metabolites. Barnol¹ and marilone A² are examples from fungi, while gilvocarcin E³ and tiacumicin B⁴ occur in *Streptomyces* (Figure 1). Feeding experiments and genetic

feeding experiments with [methyl-¹³C]-L-methionine.^{2,8} The responsible enzymes for the methylation and the biosynthetic pathways for such metabolites have not been reported prior to this study.

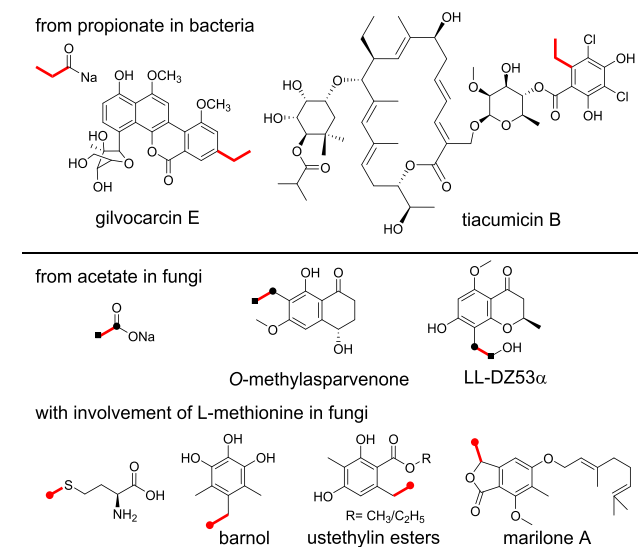


Figure 1. Origins of ethyl groups in phenethyl-containing natural products.

studies proved that the phenethyl units are products of polyketide synthases (PKSs).^{3–5} The ethyl groups in phenethyl residue of bacterial metabolites are mostly originated from propionate as starter unit of PKSs.^{3,4} In fungi, it can be derived from acetate, as in the cases of LL-DZ53α⁶ and *O*-methylasparvenone,⁷ which was confirmed by feeding with [1,2-¹³C] acetate. However, the most methyl groups of the phenethyl residue in fungal metabolites are derived from *S*-adenosyl *L*-methionine (SAM), which has been proven by

HPLC analysis of the EtOAc extract of an *A. ustus* culture in PD media revealed the presence of one predominant peak **1** (Figure 2B(i)) with a [M + Na]⁺ ion at *m/z* 249.0732 and a deduced molecular formula of C₁₁H₁₄O₅ (see Figure S7 in the Supporting Information). Attempts to get interpretable ¹H NMR spectrum for **1** from a large-scale fermentation failed, although it was almost the only product peak in the HPLC chromatogram. During isolation, the amount of **1** decreased evidently. Dissolving the finally isolated 7 mg sample in DMSO-*d*₆ led to precipitation immediately. The ¹H NMR spectrum of the supernatant was very complex, so that an interpretation was impossible (data not shown). To overcome the instability, **1** in the fungal extract was immediately converted to its triacetate **2** for structural elucidation (see Table S5, Figure S19, and Figures S21–S25 in the Supporting Information), which confirmed **1** to be 2-hydroxy-3-hydroxymethyl-4-methoxy-6-hydroxyethylbenzaldehyde, termed ustethylin A (Scheme 1).

To elucidate the origin of **1**, we performed a feeding experiment with sodium [1,2-¹³C] acetate in *A. ustus*. In the ¹³C NMR spectrum of the acetylated product **2** (Figure 3), four signal pairs of coupling carbons, C-1/C-7, C-2/C-3, C-4/C-5, and C-6/C-9, were detected, proving unequivocally the incorporation of four intact acetate units. Feeding with sodium

Received: August 14, 2020

Published: October 2, 2020

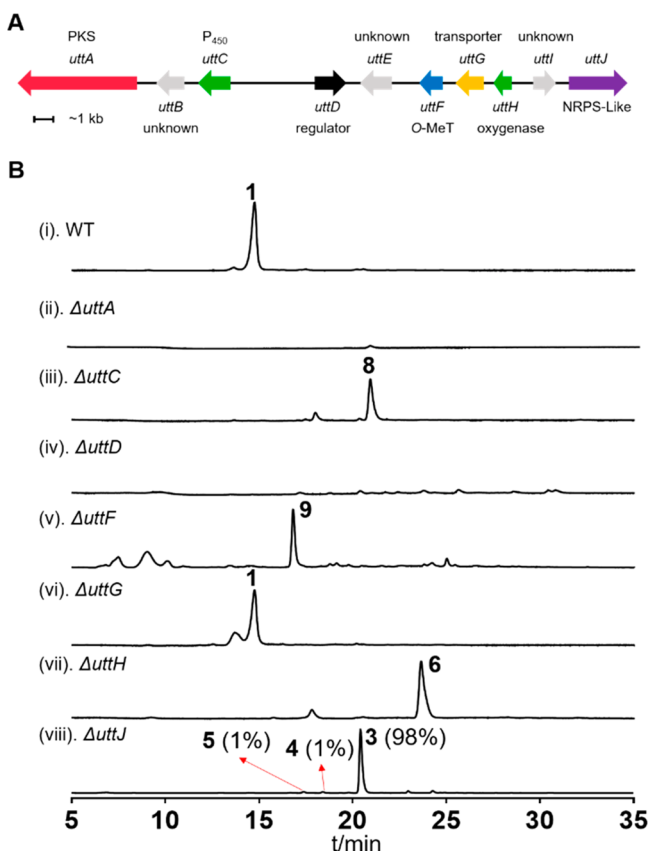


Figure 2. (A) Schematic representation of the *utt* cluster in *A. ustus* and (B) HPLC analysis of the fungal extracts.

[1-¹³C] acetate revealed that C-2 at δ_C 152.5, C-4 at 162.1, C-6 at 144.9, and C-7 at 189.0 ppm are from the carbonyl group of acetate, with 3.8–6.1-fold enrichments (see Table S11 in the Supporting Information). Correspondingly, significantly increased intensity was observed for the signals at δ_C 119.8 (C-1), 116.0 (C-3), 112.1 (C-5), and 31.5 ppm (C-9), with 4.9–8.9-fold of those of the unlabeled **2** after feeding with sodium [2-¹³C] acetate and [2-¹³C] malonic acid (see Table S11). This confirmed the methyl/methylene group of acetate/malonate as their origin.

To determine whether *A. ustus* utilizes propionate as a starter unit, sodium [2-¹³C] propionate was fed into the culture. To our surprise, the labeling pattern of **2** is very similar to that with [1-¹³C] acetate (Figure 3), with 3.6–6.1-fold

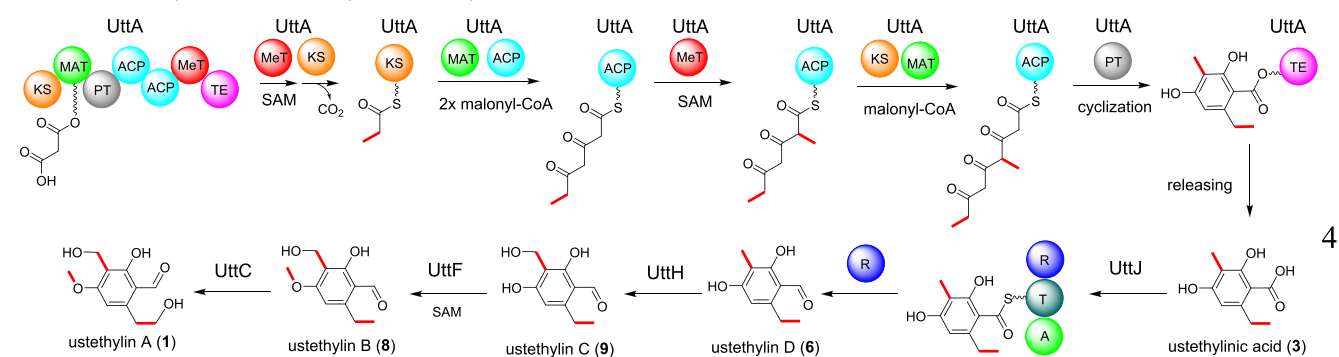
enrichments for C-2, C-4, C-6, and C-7 (see Table S11). No enrichment for C-9 at 31.5 ppm was observed. These results proved unequivocally that sodium [2-¹³C] propionate was not directly utilized for incorporation, but was degraded to acetyl-CoA likely via pyruvate by α -oxidation.⁹ Acetyl-CoA was converted to malonyl-CoA and incorporated in **1**.

In the ¹³C NMR spectrum of **2** after feeding with [methyl-¹³C]-L-methionine (Figure 3), the three signals at δ_C 54.1 (C-8), 63.9 (C-10), and 56.6 ppm (C-11) were enhanced to 13.1–15.9-fold of those of the unlabeled **2** (see Table S11), proving that the methyl group of the phenethyl residue is also from SAM.

The genome of *A. ustus* 3.3904 was sequenced and published in 2015.¹⁰ For our biosynthetic studies, we resequenced it and used both sequences for prediction of putative gene clusters by AntiSMASH.¹¹ Our sequence correlated very well with the published data, at least for the cluster described in this study. From the structure of **1**, it can be deduced that a PKS would be responsible for the formation of its backbone.¹² Genome mining indicated the presence of more than 20 putative PKS genes. Transcriptome analysis revealed that the PKS gene coding for KIA75596, termed *uttA* in this study (recall Figure 2A), was one of the 80 best expressed genes under our culture conditions (data not included).

To prove its function, *uttA* was replaced with a hygromycin B resistance cassette by using a split marker gene replacement protocol.¹³ The potential mutants were verified by PCR (Figure S1 in the Supporting Information) and cultivated in PD medium for secondary metabolite production. HPLC analysis of the culture extract of a $\Delta uttA$ mutant showed complete loss of **1** production (recall Figure 2B(ii)). Afterward, *uttA* was cloned into pYH-*gpdA*-*pyrG* via homologous recombination in yeast¹⁴ for heterologous expression.¹⁵ The obtained plasmid pLZ51 was linearized by *Sma*I and integrated into the genome of *A. nidulans* LO8030 (Figure S2 in the Supporting Information).¹⁶ In comparison to that of the negative control, three additional products **3**–**5** were detected in the *uttA* overexpression transformant with **3** as the predominant peak (98%) (see Figure 4B). The *UttA* products **3**–**5** were identified as benzoic acid derivatives by NMR analysis and comparison with published data^{17,18} (see Tables S6 and S7 in the Supporting Information, as well as Figure 4C and Figures S26–S31 in the Supporting Information). They differ from each other only in the methyl group at C-3 and the ethyl group at C-6 of the benzene ring (Figure 4C), indicating multimethylation steps during the formation of **3**.

Scheme 1. Biosynthetic Pathway of Ustethylins in *A. ustus*^a



^aThe red bonds indicate SAM-associated methylation.

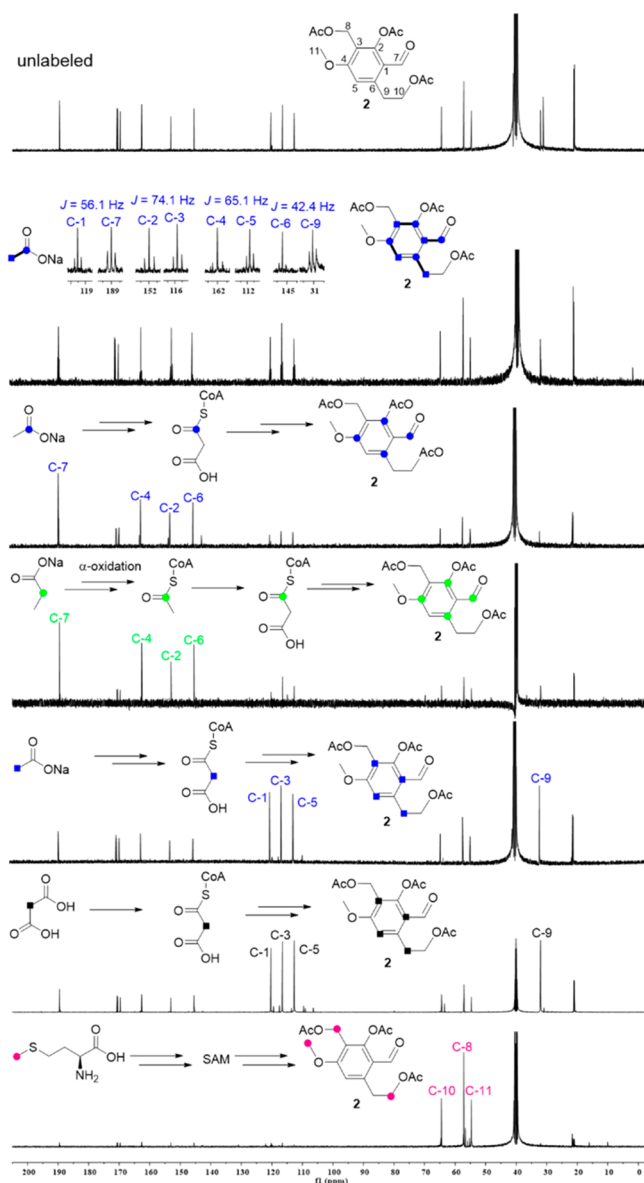


Figure 3. ^{13}C NMR spectra of the labeled and unlabeled **2**. [Legend: solid blue circles represent the labeled carbons after feeding with sodium $[1-^{13}\text{C}]$ acetate, solid green circles represent sodium $[2-^{13}\text{C}]$ propionate, solid blue squares represent sodium $[2-^{13}\text{C}]$ acetate, solid black squares represent $[2-^{13}\text{C}]$ malonic acid, bold bonds represent sodium $[1,2-^{13}\text{C}]$ acetate for the intact acetate unit, and solid pink circles represent $[\text{methyl-}^{13}\text{C}]\text{-L-methionine}$.]

Feeding **3** into the ΔuttA mutant led to detection of one additional major peak **6** and one minor peak **7** (see Figure 5A). Structure elucidation confirmed **6** to be the corresponding aldehyde of **3** (ustethylin D; see Scheme 1, as well as Table S7 and Figure S32 in the Supporting Information) and the minor peak **7** as a dimerization product (see Table S8 and Figures S33–S37 in the Supporting Information). Trace amounts of **1** were also detected in this culture (Figure 5A), proving **3** to be a precursor of **1**. It can be speculated that the metabolism of **6** is a limited step in the biosynthesis. Interestingly, **3** was not detected in the ΔuttA mutant after feeding with **4** (Figure S18 in the Supporting Information), excluding the direct methylation of the C6-methyl group in **4**.

Bioinformatics analysis revealed that UttA is a nonreducing PKS with a domain architecture of KS-MAT-PT-ACP-ACP-

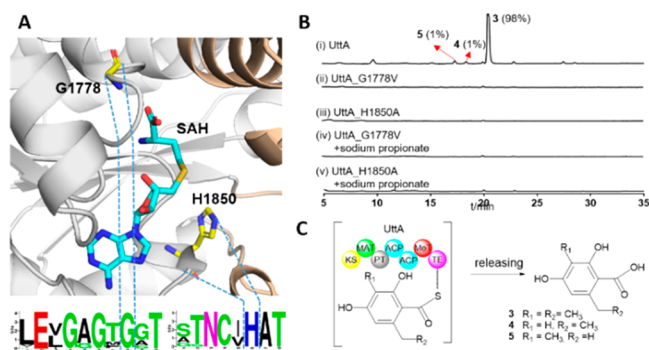


Figure 4. (A) Structural model of UttA-MeT with conserved motifs by alignments of 96 MeT domains in PKSs; (B) HPLC results of *A. nidulans* heterologous expression strains; and (C) the reactions catalyzed by UttA.

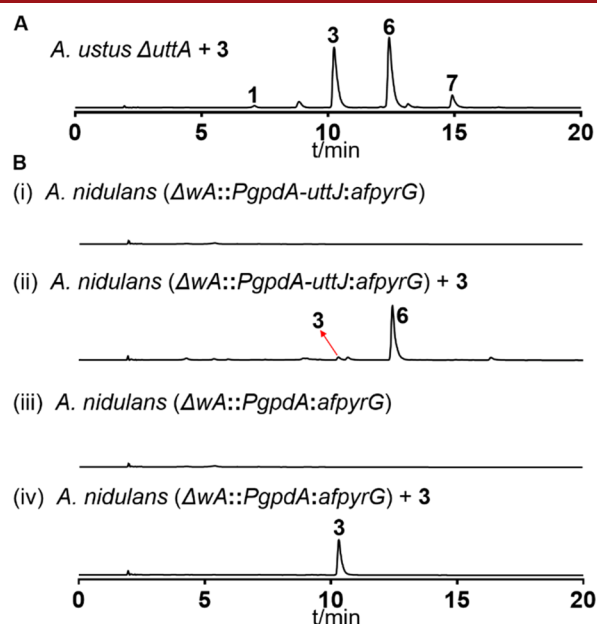


Figure 5. HPLC analysis of culture extracts of (A) *A. ustus* ΔuttA and (B) *A. nidulans* uttJ overexpression mutant after feeding with **3**.

MeT-TE (β -ketoacyl synthase, KS; malonyl-CoA-ACP transacylase, MAT; product template, PT; acyl carrier protein, ACP; methyltransferase, MeT; thiolesterase, TE) (Figure 4C).¹⁹ The PT domain was deduced by phylogenetic analysis with 30 known PKS PTs (see Figure S3 in the Supporting Information).²⁰ It can be speculated that the methyl groups in **3** are transferred by the MeT domain from S-adenosyl L-methionine. UttA shares a sequence identity of 37.8% with the citrinin synthase PksCT from *Monascus purpureus*.²¹ The sequence identity of their MeT domains is found to be 39.7%. Therefore, a model of UttA-MeT was built, using PksCT-MeT as a template with SWISS-MODEL.²² Meanwhile, 96 MeT domains in PKSs were analyzed and presented with Weblogo (Figure 4A).^{23,24} The highly conserved residues of the SAM binding motif ExGxGxGx were identified at residues 1772–1779 in UttA. His1850 acting as a key catalytic residue for enolization is also highly conserved (see Figure S4 in the Supporting Information). To delete MeT activity from UttA, the UttA_G1778 V and UttA_H1850A mutants were constructed and expressed in *A. nidulans* LO8030. HPLC results showed the abolishment of the PKS products **3**–**5**

without accumulation of any other products, indicating the essential role of MeT domain for the functionality of UttA. Feeding propionate to the two mutants did not result in any detectable product formation, confirming that UttA cannot directly utilize propionate as starter as described above (Figure 4B(ii)–(v)). These results indicated that the methylation step is essential for the polyketide assembling by UttA. We propose that malonyl-CoA is loaded onto the MAT domain and transferred to the ACP domain. The propionyl-ACP complex is formed by methylation via SAM and decarboxylation. After condensation with two malonyl-CoA molecules, the MeT domain attaches the second methyl group. Condensation with another malonyl-CoA molecule led to the production of a polyketide chain, which is subsequent cyclized by the PT domain and released by the TE domain, resulting in the formation of the predominant product **3** (98%; see Figures 2B and 4B). However, it cannot be excluded that mutation at G1778 and H1850 had influence on the transcription, translation process, or protein stability. Attempts to get recombination protein of the MeT domain failed, so that no in vitro study was possible. The reaction catalyzed by UttA with involvement of two methylation steps is closely related to that of 3-methylcinaldehyde synthase MOS from *Acremonium strictum*. However, MOS contains a terminal reductase domain for direct release of an aldehyde,²⁵ while the UttA product, an aryl acid, is afterward reduced to an aldehyde.

1 differs structurally from **3** in oxidation states of the functional groups at C-1, C-3, and C-6, as well as *O*-methylation at OH-4. The conversion of **3** to **1** would require three oxidoreductases and an *O*-MeT. Inspection of the genomic neighborhood of *uttA* in *A. ustus* revealed the presence of a putative biosynthetic gene cluster (*uttA*–*uttJ*), coding for the putative proteins KIA75596–KIA75587 in the database; see Figure 2A) containing such genes as well as those coding for regulator and transporter (Table S1 in the Supporting Information). *UttJ* coding for an NRPS-like enzyme with an A-T-R domain architecture is 1 of the 10 highly expressed genes in *A. ustus* under our culture conditions (data not included). Given such enzymes from different fungal strains for aryl acid reduction to aldehydes,^{26,27} *UttJ* is likely involved in the conversion of **3** to **6**. Indeed, deletion of *uttJ* abolished the formation of **1** and accumulation of **3**–**5** with almost the same product profile of the *A. nidulans* *uttA* overexpression strain (see Figures 2B(viii) and 4B(i), as well as Figure S16 in the Supporting Information). This unambiguously proved its role in the reduction of the carboxyl group to an aldehyde. Feeding **3** into the *A. nidulans* *uttJ* overexpression strain led to the detection of **6** (see Figure 5B(ii)), proving *UttJ* as an aryl acid reductase (Scheme 1). Further sequence comparison and analysis revealed *UttH* to be a putative nonheme Fe^{II}/2-oxoglutarate dependent oxygenase and shares 58% and 52% sequence identities with CitB from *Monascus ruber*²⁸ and ClaD from *Penicillium crustosum*,²⁹ respectively (see Table S1 and Figure S5 in the Supporting Information). Both known enzymes catalyze hydroxylations of aryl methyl groups. Deletion of *uttH* led to the accumulation of *UttJ* product **6**, proving the reaction order of both enzymes (see Figure 2B(vii), as well as Figure S15 in the Supporting Information). Deletion of *uttC* coding for a cytochrome P₄₅₀ enzyme abolished the formation of **1** and production of **8** (ustethylin B; see Table S9 in the Supporting Information, Figure 2B(iii), and Figures S9 and S38–S41 in the Supporting Information), which differs from **1** just in the oxidation state of

the ethyl group. This proved that *UttC* catalyzed the last step in the biosynthesis of **1**. Bioinformatics analysis showed that *UttC* contains the conserved motifs ExxR (EAGR, 349–352) and CxG (CLG, 434–436) of P₄₅₀ enzymes (see Figure S6 in the Supporting Information).³⁰ Usually, the hydroxylation of phenethyl group occurs at the α -position (–CH₂–), e.g., in the biosynthesis of marilone A.³¹ Here, we present an unusual β -hydroxylation of the phenethyl group by the cytochrome P₄₅₀ *UttC*. Detection of **8** with a methoxy group in Δ *uttC* mutant indicates that the methylation of the C4-hydroxyl group occurs before *UttC* and after *UttH* reactions and could be catalyzed by the putative *O*-MeT *UttF*. Indeed, one predominant peak **9** (ustethylin C) with a [M + Na]⁺ ion at *m/z* 219.0626, which is 14 Da less than that of **8**, was observed in the Δ *uttF* mutant (see Figure 2B(v), as well as Figure S13 in the Supporting Information). Similar to **1**, **9** was also found to be unstable and could not be obtained in pure form for structure elucidation. However, its structure can be elucidated after conversion to its diacetylated derivative **10** (see Table S10 and Figures S42–S45 in the Supporting Information).

Gene deletion results revealed the reaction sequence of the tailoring enzymes for **3** conversion to **1**. Extracted ion chromatograms of the culture proved the presence of **1** as almost the only pathway product (see Figure S7 in the Supporting Information), indicating the high efficiency of the involved enzymes in wildtype *A. ustus*. The *utt* cluster is positively regulated by a DNA binding enzyme *UttD*. Deletion of *uttD* completely abolished product formation (see Figure 2B(iv), as well as Figure S11 in the Supporting Information). Even feeding **3** to the Δ *uttD* deletion mutant did not lead to any conversion (see Figure S17 in the Supporting Information). Deletion of *uttG* coding for an MFS transporter reduced **1** production to 30.8% of that of the wildtype *A. ustus* (see Figures 2B(vi), as well as Figure S14 in the Supporting Information). **1** was still detected in the deletion mutants of the two oxidoreductase genes *uttB* and *utte* (see Figures S9 and S12 in the Supporting Information). They very likely are not involved in the formation of ustethylin A.

In summary, in this study, we have identified the biosynthetic gene cluster of the highly oxygenated aryl-aldehyde derivative ustethylin A and elucidated its biosynthetic pathway by transcriptome analysis, gene deletion, and expression, as well as isotopic labeling experiments. The PKS *UttA*, as a key enzyme, is responsible for the formation of the phenethyl core structure with methylation as key reactions. Consecutive and coordinated modifications by three different types of oxidoreductases and one *O*-MeT lead to the final product. To the best of our knowledge, this is the first report on the biosynthetic pathway of a phenethyl-containing fungal metabolite.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.orglett.0c02719>.

Supplementary methods, tables, and figures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Shu-Ming Li – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität

Marburg, 35037 Marburg, Germany; orcid.org/0000-0003-4583-2655; Email: shuming.li@staff.uni-marburg.de

Authors

Liujuan Zheng – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Yiling Yang – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Haowen Wang – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Aili Fan – College of Life Science and Technology, Beijing University of Chemical Technology, Chaoyang District 100029 Beijing, China

Liping Zhang – Key Laboratory of Tropical Marine Bio-resources, South China Sea Institute of Oceanology Chinese Academy of Sciences, Guangzhou 510301, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.orglett.0c02719>

Author Contributions

[†]These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Rixa Kraut, Lena Ludwig-Radtke, and Stefan Newel (Fachbereich Pharmazie, Philipps-Universität Marburg) for taking MS and NMR spectra. This project was financially funded in part by the Deutsche Forschungsgemeinschaft (DFG) (No. INST 160/620-1). L.Z. (No. 201604910536) and Y.Y. (No. 201808530447) are scholarship recipients from the China Scholarship Council.

REFERENCES

- (1) Better, J.; Gatenbeck, S.; Hornfeldt, A.-B.; Thorstad, O.; Undheim, K.; Berg, J.-E.; Pilotti, A.-M. *Acta Chem. Scand.* **1977**, *31*, 391.
- (2) El Maddah, F.; Eguereva, E.; Kehraus, S.; König, G. M. *Org. Biomol. Chem.* **2019**, *17*, 2747.
- (3) Shepherd, M. D.; Kharel, M. K.; Zhu, L. L.; Van Lanen, S. G.; Rohr, J. A. *Org. Biomol. Chem.* **2010**, *8*, 3851.
- (4) Yu, Z.; Zhang, H.; Yuan, C.; Zhang, Q.; Khan, I.; Zhu, Y.; Zhang, C. *Org. Lett.* **2019**, *21*, 7679.
- (5) Yeh, H. H.; Chang, S. L.; Chiang, Y. M.; Bruno, K. S.; Oakley, B. R.; Wu, T. K.; Wang, C. C. *Org. Lett.* **2013**, *15*, 756.
- (6) McIntyre, C. R.; Simpson, T. J.; Trimble, L. A.; Vederas, J. C. *J. Chem. Soc., Chem. Commun.* **1984**, 706.
- (7) Simpson, T. J.; Stenzel, D. J. *J. Chem. Soc., Chem. Commun.* **1981**, 239.
- (8) De Jesus, A. E.; Horak, R. M.; Steyn, P. S.; Vleggaar, R. J. *J. Chem. Soc., Perkin Trans. 1* **1987**, 2253.
- (9) Brock, M.; Fischer, R.; Linder, D.; Buckel, W. *Mol. Microbiol.* **2000**, *35*, 961.
- (10) Pi, B.; Yu, D.; Dai, F.; Song, X.; Zhu, C.; Li, H.; Yu, Y. *PLoS One* **2015**, *10*, No. e0116089.
- (11) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, *47*, W81.
- (12) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.
- (13) Goswami, R. S. *Methods Mol. Biol.* **2012**, *835*, 255.
- (14) Joska, T. M.; Mashruwala, A.; Boyd, J. M.; Belden, W. J. *J. Microbiol. Methods* **2014**, *100*, 46.

(15) Zhang, P.; Wang, X.; Fan, A.; Zheng, Y.; Liu, X.; Wang, S.; Zou, H.; Oakley, B. R.; Keller, N. P.; Yin, W. B. *Mol. Microbiol.* **2017**, *105*, 469.

(16) Chiang, Y. M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C.; Oakley, B. R. *Angew. Chem., Int. Ed.* **2016**, *55*, 1662.

(17) Durrani, A. A.; Tyman, J. H. P. *J. Chem. Soc., Perkin Trans. 1* **1980**, 1658.

(18) Culberson, C. F.; Culberson, W. L. *Exp. Mycol.* **1978**, *2*, 245.

(19) Chooi, Y. H.; Tang, Y. J. *Org. Chem.* **2012**, *77*, 9933.

(20) Li, Y.; Xu, W.; Tang, Y. J. *Biol. Chem.* **2010**, *285*, 22764.

(21) Storm, P. A.; Herbst, D. A.; Maier, T.; Townsend, C. A. *Cell Chem. Biol.* **2017**, *24*, 316.

(22) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. *Nucleic Acids Res.* **2018**, *46*, W296.

(23) Crooks, G. E.; Hon, G.; Chandonia, J. M.; Brenner, S. E. *Genome Res.* **2004**, *14*, 1188.

(24) Schneider, T. D.; Stephens, R. M. *Nucleic Acids Res.* **1990**, *18*, 6097.

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

(26) Li, C.; Matsuda, Y.; Gao, H.; Hu, D.; Yao, X. S.; Abe, I. *ChemBioChem* **2016**, *17*, 904.

(27) Wang, M.; Beissner, M.; Zhao, H. *Chem. Biol.* **2014**, *21*, 257.

(28) He, Y.; Cox, R. J. *Chem. Sci.* **2016**, *7*, 2119.

(29) Fan, J.; Liao, G.; Kindinger, F.; Ludwig-Radtke, L.; Yin, W.-B.; Li, S.-M. *J. Am. Chem. Soc.* **2019**, *141*, 4225.

(30) Syed, K.; Mashele, S. S. *PLoS One* **2014**, *9*, e95616–1.

(31) Du, L.; Dong, S.; Zhang, X.; Jiang, C.; Chen, J.; Yao, L.; Wang, X.; Wan, X.; Liu, X.; Wang, X.; Huang, S.; Cui, Q.; Feng, Y.; Liu, S. J.; Li, S. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, No. E5129.

Supporting Information

Ustethylin biosynthesis implies phenethyl derivative formation in *Aspergillus ustus*

Liujuan Zheng,^{‡a} Yiling Yang,^{‡a} Haowen Wang,^a Aili Fan,^b Liping Zhang^c and Shu-Ming Li^{*a}

^a Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, Robert-Koch Straße 4, 35037 Marburg (Germany)

^b College of Life Science and Technology, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, 100029 Beijing (China)

^c Key Laboratory of Tropical Marine Bio-resources, South China Sea Institute of Oceanology Chinese Academy of Sciences 164 West Xingang Road, Guangzhou 510301 (China)

SUPPORTING INFORMATION

Table of Contents

| | |
|---|----|
| Experimental Procedures | 5 |
| 1. Chemicals | 5 |
| 2. Strains, media, and growth conditions | 5 |
| 3. Genomic DNA isolation | 5 |
| 4. Genome sequencing and sequence analysis | 5 |
| 5. PCR amplification, gene cloning and plasmid construction | 6 |
| 6. Molecular modeling for UttA_MeT | 6 |
| 7. Genetic manipulation in <i>A. ustus</i> 3.3904 and cultivation of deletion mutants | 6 |
| 8. Heterologous expression in <i>A. nidulans</i> | 7 |
| 9. Site-directed mutagenesis of UttA | 7 |
| 10. HPLC equipment for analysis and metabolite isolation | 7 |
| 11. Large-scale fermentation, extraction and isolation of secondary metabolites | 7 |
| 12. Feeding experiments | 9 |
| 13. LC-MS analysis | 10 |
| 14. NMR analysis | 10 |
| 15. Physicochemical properties of the compounds described in this study | 10 |
| 16. Structural elucidation | 11 |
| Table S1. Putative functions of the genes from the ustethylin (<i>uti</i>) gene cluster | 13 |
| Table S2. Strains used in this study | 14 |
| Table S3. Plasmids used and constructed in this study | 15 |
| Table S4. Primers used in this study | 16 |
| Table S5. NMR data of compound 2 in DMSO-<i>d</i>₆ (500 MHz, δ in ppm, <i>J</i> in Hz) | 20 |
| Table S6. NMR data of compound 3 in DMSO-<i>d</i>₆ (500 MHz, δ in ppm, <i>J</i> in Hz) | 21 |
| Table S7. ¹H NMR data of compounds 4 – 6 in DMSO-<i>d</i>₆ (500 MHz, δ in ppm, <i>J</i> in Hz) | 22 |
| Table S8. NMR data of compound 7 in DMSO-<i>d</i>₆ (500 MHz, δ in ppm, <i>J</i> in Hz) | 23 |
| Table S9. NMR data of compound 8 in CDCl₃ (500 MHz, δ in ppm, <i>J</i> in Hz) | 24 |
| Table S10. NMR data of compound 10 in DMSO-<i>d</i>₆ (500 MHz, δ in ppm, <i>J</i> in Hz) | 25 |
| Table S11. Enrichments in 2 after feeding with ¹³C labeled precursors | 26 |
| Figure S1. PCR verification of deletion mutants of <i>A. ustus</i> 3.3904. | 27 |

SUPPORTING INFORMATION

| | |
|--|----|
| Figure S2. PCR verification of heterologous expression transformants. | 28 |
| Figure S3. Phylogenetic analysis of UttA_PT domain with 30 known PT domains from fungi. ... | 29 |
| Figure S4. Point mutation in UttA and analysis of the obtained mutants. | 30 |
| Figure S5. Sequence alignments of 2-OG-dependent oxygenases..... | 31 |
| Figure S6. Weblogo illustration for the conserved ExxR and CxG motifs in UttC by using 96 P ₄₅₀ enzyme in fungi..... | 32 |
| Figure S7. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> wild type..... | 33 |
| Figure S8. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttA mutant | 34 |
| Figure S9. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttB mutant | 35 |
| Figure S10. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttC mutant | 36 |
| Figure S11. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttD mutant..... | 37 |
| Figure S12. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttE mutant | 38 |
| Figure S13. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttF mutant | 39 |
| Figure S14. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttG mutant..... | 40 |
| Figure S15. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttH mutant..... | 41 |
| Figure S16. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttJ mutant | 42 |
| Figure S17. LC-MS analysis of the metabolite profile of the <i>A.ustus</i> Δ uttD after feeding with 3 .. | 43 |
| Figure S18. LC-MS analysis of the metabolite profile of the <i>A. ustus</i> Δ uttA after feeding with 4 . | 44 |
| Figure S19. LC-MS analysis of the acetylated EtOAc extract from <i>A. ustus</i> | 45 |
| Figure S20. LC-MS analysis of the acetylated EtOAc extract from Δ uttF of <i>A. ustus</i> | 45 |
| Figure S21. UV spectra of the compounds identified in this study..... | 46 |
| Figure S22. ¹ H NMR spectrum of compound 2 in DMSO- <i>d</i> ₆ (500 MHz)..... | 47 |
| Figure S23. ¹³ C{ ¹ H} NMR spectrum of compound 2 in DMSO- <i>d</i> ₆ (125 MHz)..... | 47 |
| Figure S24. HSQC spectrum of compound 2 in DMSO- <i>d</i> ₆ | 48 |
| Figure S25. HMBC spectrum of compound 2 in DMSO- <i>d</i> ₆ | 48 |
| Figure S26. ¹ H NMR spectrum of compound 3 in DMSO- <i>d</i> ₆ (500 MHz)..... | 49 |
| Figure S27. ¹³ C{ ¹ H} NMR spectrum of compound 3 in DMSO- <i>d</i> ₆ (125 MHz)..... | 49 |
| Figure S28. HSQC spectrum of compound 3 in DMSO- <i>d</i> ₆ | 50 |
| Figure S29. HMBC spectrum of compound 3 in DMSO- <i>d</i> ₆ | 50 |
| Figure S30. ¹ H NMR spectrum of compound 4 in DMSO- <i>d</i> ₆ (500 MHz)..... | 51 |
| Figure S31. ¹ H NMR spectrum of compound 5 in DMSO- <i>d</i> ₆ (500 MHz)..... | 51 |

SUPPORTING INFORMATION

| | |
|---|----|
| Figure S32. ^1H NMR spectrum of compound 6 in $\text{DMSO-}d_6$ (500 MHz)..... | 52 |
| Figure S33. ^1H NMR spectrum of compound 7 in $\text{DMSO-}d_6$ (500 MHz)..... | 52 |
| Figure S34. $^{13}\text{C}\{^1\text{H}\}$ NMR spectrum of compound 7 in $\text{DMSO-}d_6$ (125 MHz)..... | 53 |
| Figure S35. HSQC spectrum of compound 7 in $\text{DMSO-}d_6$ | 53 |
| Figure S36. $^1\text{H-}^1\text{H}$ COSY spectrum of compound 7 in $\text{DMSO-}d_6$ | 54 |
| Figure S37. HMBC spectrum of compound 7 in $\text{DMSO-}d_6$ | 54 |
| Figure S38. ^1H NMR spectrum of compound 8 in CDCl_3 (500 MHz) | 55 |
| Figure S39. $^{13}\text{C}\{^1\text{H}\}$ NMR spectrum of compound 8 in CDCl_3 (125 MHz) | 55 |
| Figure S40. HSQC spectrum of compound 8 in CDCl_3 | 56 |
| Figure S41. HMBC spectrum of compound 8 in CDCl_3 | 56 |
| Figure S42. ^1H NMR spectrum of compound 10 in $\text{DMSO-}d_6$ (500 MHz)..... | 57 |
| Figure S43. $^{13}\text{C}\{^1\text{H}\}$ NMR spectrum of compound 10 in $\text{DMSO-}d_6$ (125 MHz)..... | 57 |
| Figure S44. HSQC spectrum of compound 10 in $\text{DMSO-}d_6$ | 58 |
| Figure S45. HMBC spectrum of compound 10 in $\text{DMSO-}d_6$ | 58 |
| References | 59 |

Experimental Procedures

1. Chemicals

Sodium [1-¹³C] acetate, sodium [2-¹³C] acetate, and sodium [2-¹³C] propionate were purchased from Cambridge Isotope Laboratories. Sodium [1,2-¹³C] acetate, [2-¹³C] malonic acid, and [methyl-¹³C]-L-methionine were obtained from Sigma-Aldrich. Other reagents were from Fisher scientific, VWR or Sigma-Aldrich.

2. Strains, media, and growth conditions

Escherichia coli DH5 α cells were grown in LB medium (1% NaCl, 1% tryptone, and 0.5% yeast extract). 50 mg/mL ampicillin were supplemented for cultivation of recombinant *E. coli* strains.

Saccharomyces cerevisiae HOD114-2B cells were grown in YPD medium (1% yeast extract, 2% peptone and 2% glucose). 1.5% agarose was used for plates. The SC-uracil medium (6.7 g/L yeast nitrogen base with ammonium sulfate, 650 mg/L CSM-His-Leu-Ura (MP Biomedicals), 20 mg/L His and 60 mg/L Leu, 2% glucose, pH 6.2 – 6.3, 1.5% agarose was used for plates) was used for selection.

Fungal strains used in this study are summarized in Table S2. *Aspergillus ustus* (*A. ustus*) 3.3904 was purchased from China General Microbiological Culture Collection Center (Beijing, China) and cultivated in PD (potato dextrose broth, Sigma) or ISP3 (6% oat) medium at 230 rpm and 30 °C for secondary metabolite (SM) production.

Aspergillus nidulans strains were grown at 37 °C on GMM medium (1.0% glucose, 50 mL/L salt solution, 1 mL/L trace element solution, 1.6% agar) for sporulation and transformation with appropriate nutrition as required. The salt solution comprises (w/v) 12% NaNO₃, 1.04% KCl, 1.04% MgSO₄·7H₂O, and 3.04% KH₂PO₄. The trace element solution contains (w/v) 2.2% ZnSO₄·7H₂O, 1.1% H₃BO₃, 0.5% MnCl₂·4H₂O, 0.16% FeSO₄·7H₂O, 0.16% CoCl₂·5H₂O, 0.16% CuSO₄·5H₂O, 0.11% (NH₄)₆Mo₇O₂₄·4H₂O, and 5% Na₄EDTA.

3. Genomic DNA isolation

The mycelia of *A. ustus* 3.3904 and *A. nidulans* were dried on filter paper and collected in 2 mL Eppendorf tubes. Four glass beads (2.85 mm in diameter) and 400 μ L of LETS buffer (10 mM Tris-HCl pH 8.0, 20 mM EDTA pH 8.0, 0.5% SDS, and 0.1 M LiCl) were added to the tubes. After vigorous mixing for 4 min, 300 μ L LETS buffer were added. The solution was then treated with 700 μ L phenol: chloroform: isoamyl alcohol (25: 24: 1). Genomic DNA was precipitated by addition of 900 μ L of absolute EtOH. After centrifugation at 13,000 rpm for 30 min and washing with 70% EtOH, the obtained DNA was dissolved in 50 – 100 μ L distilled H₂O.

4. Genome sequencing and sequence analysis

The genome of *A. ustus* 3.3904 was sequenced by Genewiz (Suzhou, China) using Nova-seq6000/X-ten (Illumina). Initial prediction and analysis of biosynthetic gene clusters were carried out by using AntiSMASH.¹

SUPPORTING INFORMATION

Prediction of the enzyme function was performed with the online BLAST tools (<http://blast.ncbi.nlm.nih.gov>). The genes *uttA-J* in the *utt* cluster are summarized in Table S1. The genomic DNA sequence of the *utt* cluster reported in this study corresponds to that depicted at GenBank under accession numbers JOMC01000079.1.

5. PCR amplification, gene cloning and plasmid construction

Plasmids and primers used in this study are listed in Table S3 and Table S4, respectively. Primers were synthesized by SeqLab GmbH (Göttingen, Germany). PCR amplification was carried out by using Phusion® High-Fidelity DNA polymerase from New England Biolabs (NEB) on a T100™ Thermal cycler from Bio-Rad. PCR thermal profiles were set as recommended by the manufacturer's instruction. The plasmids for gene deletion and heterologous expression were constructed *via* homologous recombination in *Escherichia coli* DH5 α or *Saccharomyces cerevisiae* HOD114-2B by using primers listed Table S4.

6. Molecular modeling for UttA_MeT

Homolog modelling for 300 amino acids of UttA_MeT was carried out by using SWISS-MODEL.² S-Adenosyl-L-homocysteine (SAH) was manually positioned by using Coot.³ The illustration was created with Pymol (DeLano Scientific LLC, Version 1.3.x.).

7. Genetic manipulation in *A. ustus* 3.3904 and cultivation of deletion mutants

Fresh spores of *A. ustus* 3.3904 were inoculated into 50 mL LMM medium in 250 mL flask and incubated at 230 rpm and 30 °C for germination. The germlings were harvested after 11 h by centrifugation at 5,000 rpm and 4 °C for 10 min and washed with distilled H₂O. The mycelia were then transferred into a 25 mL flask with 10 mL of osmotic buffer (1.2 M MgSO₄ in 10 mM sodium phosphate, pH 5.8) containing 40 mg lysing enzyme from *Trichoderma harzianum* (Sigma) and 30 mg yatalase from *Corynebacterium sp.* OZ-21 (OZEKI Co., Ltd.). After shaking at 100 rpm and 30 °C for 10 h, the mixture was transferred into a 50 mL falcon tube and overlaid gently with 10 mL of trapping buffer (0.6 M sorbitol in 0.1 M Tris-HCl, pH 7.0). After centrifugation at 5,000 rpm and 4 °C for 10 min, the protoplasts were collected from the interface of the two buffer systems. The collected protoplasts were then transferred to a sterile 15 mL falcon tube and resuspended in 100 μ L of STC buffer (1.2 M sorbitol, 10 mM CaCl₂ in 10 mM Tris-HCl, pH 7.5) for transformation.

The DNA fragments (2 – 3 μ g in 8 – 10 μ L) were mixed with 100 μ L of the protoplasts and incubated for 50 min on ice. 1.25 mL of PEG solution (60% polyethylene glycol 4000, 50 mM CaCl₂, 50 mM Tris-HCl, pH 7.5) was then added and gently mixed. After incubation at room temperature for 30 min, 5 mL STC buffer was added into the mixture and spread on plates with SMM bottom medium (1.0% glucose, 50 mL/L salt solution, 1 mL/L trace element solution, 1.2 M sorbitol, and 1.6% agar) containing 100 μ g/mL hygromycin B. SMM top medium (1.0% glucose, 50 mL/L salt solution, 1 mL/L trace element solution, 1.2 M sorbitol, and 0.8% agar) containing 50 μ g/mL hygromycin B was overlaid softly on the plates. 3 – 4 days later, the transformants were transferred onto fresh PDA

SUPPORTING INFORMATION

plates (PD medium with 3% agar) containing 100 $\mu\text{g}/\text{mL}$ hygromycin B for selection. The obtained transformants were inoculated in PD medium for isolation of genomic DNA to verify the integrity *via* PCR amplification (Figure S2). After cultivation in PD liquid medium at 230 rpm and 30 °C for 7 days, the cultures were extracted with EtOAc, dissolved in DMSO and subjected to HPLC and LC-MS for analysis.

8. Heterologous expression in *A. nidulans*

A. nidulans LO8030⁴ was used as the recipient host. The protoplast preparation and transformation were performed as described previously.⁵ PLZ51 – 54 containing the PKS gene *uttA*, *uttA-G1778V* and *uttA-H1850A* as well as the NRPS-like gene *uttJ* were transformed into *A. nidulans* LO8030 to create expression strain LZ51, LZ52, LZ53 and LZ54, respectively. The transformants were verified by PCR (Figures S2 and S4).

9. Site-directed mutagenesis of *UttA*

The fragments containing the point mutation were constructed *via* fusion PCR. The mutated and non-mutated fragments of *uttA* were then integrated into pYH-*gpdA-pyrG* following the same procedure for pLZ51 to produce pLZ53 (*UttA_G1778V*) and pLZ54 (*UttA_H1850A*). The primers used for plasmids constructing were listed in Table S3 – S4.

10. HPLC equipment for analysis and metabolite isolation

EtOAc extracts of fungal strains were analyzed on an Agilent HPLC series 1200 (Agilent Technologies) equipped with an Agilent Eclipse XDB-C18 column (5 μm , 4.6 \times 150 mm). A linear gradient from 10 to 90% ACN in H₂O in 20 min was used. The column was then washed with 100% ACN for 5 min and equilibrated with 10% ACN in H₂O for another 5 min. Detection was carried out with a photodiode array detector from 190 to 400 nm.

The same HPLC system was also used for product isolation with a Multospher 120 RP-18 column (5 μm , 10 \times 250 mm). The products were eluted with different solvent gradients of ACN in H₂O, with or without HCOOH, at a flow rate of 2 mL/min.

11. Large-scale fermentation, extraction and isolation of secondary metabolites

For metabolite extraction after large-scale fermentation, the supernatant was separated from mycelia by filtration and extracted with equal volume of EtOAc for three times. The mycelia were extracted with acetone and concentrated under reduced pressure to afford an aqueous solution and then extracted with EtOAc for three times. Both EtOAc extracts were evaporated under reduced pressure to afford the crude extracts for further purification.

To isolate compound **1**, *A. ustus* spores were cultivated in 10 x 250 mL flasks containing 50 mL PDB liquid medium for 2 days, then transferred to 10 x 2 L flasks containing 500 mL PDB liquid medium each. The cultures were maintained on a rotary shaker at 230 rpm and 30 °C for 9 days. The cultures were harvested and extracted as

SUPPORTING INFORMATION

mentioned above to give 0.4 g crude extract. The crude extract was subjected to silica gel column chromatography by using stepwise gradient elution with mixtures of petroleum ether/EtOAc (20:1 to 0:1, v/v) to give five fractions (1 – 5). Fraction 2 was further purified on the HPLC system mentioned above by using ACN/H₂O (40:60) as elution solvents, resulting in 7.0mg of **1**.

To identify the structure of **1**, we used the previously published acetylation method.⁶ *A. ustus* spores were inoculated into 40 x 250 mL flasks containing 50 mL PDB liquid medium each and incubated on a rotary shaker at 230 rpm and 30 °C for 6 days. 400 mg of the obtained crude extracts were immediately acetylated with acetic anhydride (21.24 mmol) and NaOAc·3H₂O (0.3 mmol) at room temperature for 16 h. The mixture was extracted with 15 mL EtOAc and washed with 15 mL saturated solution of NaHCO₃ for three times. After evaporation of the solvent, 11.4 mg of **2** were isolated by isocratic elution with ACN/H₂O (55:45) on the aforementioned HPLC system for MS and NMR analyses.

To isolate compounds **3** – **5** from the *A. nidulans*-pYH-*gpdA-uttA-pryG* transformant, the strain was cultivated in ISP3 medium at 30 °C for 8 days. After extraction, 4.7 g crude extract was obtained from 5 L culture and subjected to silica gel column chromatography. Petroleum ether/EtOAc (50:1 to 0:1, v/v) were used as elution solvents to give 17 fractions (1 – 17). Fraction 6 was purified on the HPLC by isocratic elution with ACN/H₂O (40:60, 0.1% HCOOH) to get **3** (26.0 mg). **4** (14.6 mg) was obtained from fraction 8, which was purified on the HPLC system by isocratic elution with ACN/H₂O (25:75) containing 0.1% HCOOH. **5** (10.6 mg) was obtained from fraction 9 under the same conditions as for **4**. **3** – **5** were also isolated from the *A. ustus* Δ *uttJ* mutant in a similar procedure.

To isolate compound **6**, the Δ *uttH* mutant was cultivated in PDB medium at 30 °C for 7 days. After extraction, 0.1 g crude extract was obtained from 5 L culture and subjected to silica gel column chromatography by using petroleum ether/EtOAc (50:1 to 0:1, v/v) as elution solvents to give 11 fractions (1 – 11). **6** (1.0 mg) was obtained from fraction 2 after purification on HPLC using isocratic elution with ACN/H₂O (53:47) containing 0.1% HCOOH).

To isolate compound **8** from Δ *uttC* mutant, the mutant was cultivated in ISP3 medium at 30 °C for 8 days. After extraction, 8.2 g crude extract was obtained from 5 L culture and subjected to silica gel column chromatography by using petroleum ether/EtOAc (50:1 to 0:1, v/v) as elution solvents to give 13 fractions (1 – 13). Fraction 5 was purified on the HPLC with a linear gradient from 10 to 100% ACN containing 0.1% HCOOH in H₂O containing 0.1% HCOOH in 22 min. The column was then washed with 100% ACN containing 0.1% HCOOH for 5 min, followed by 5 min equilibration with 10% ACN containing 0.1% HCOOH. 9.3 mg of **8** were obtained for MS and NMR analyses.

To identify compound **9**, spores of Δ *uttF* mutant were inoculated into 40 x 250 mL flasks containing 50 mL PDB medium each and incubated on a rotary shaker at 230 rpm and 30 °C for 6 days. After extraction, the crude extract (175.0 mg) was also immediately acetylated *via* the same procedure described for compound **1**. Finally, 5.9 mg of

10 was obtained after purification on the HPLC by isocratic elution with ACN/H₂O (46:54) for MS and NMR analyses.

12. Feeding experiments

Feeding with ¹³C-labeled precursors

For labeling experiments, appropriate amounts of *A. ustus* spores were transferred from plates into 250 mL flasks containing 50 mL PDB medium and cultivated on a rotary shaker at 230 rpm and 30 °C. Aqueous stock solution of the respective precursor was fed after 30 h cultivation, followed by a second feeding 24 h later. After cultivation for another 60 h, the fungal cultures were extracted with EtOAc for three times. The EtOAc extracts were evaporated at 30 °C to dryness and acetylated as mentioned above. The acetylated product was further purified on the HPLC system and subjected to NMR analysis. The culture size, precursor amounts and product yields are given below.

250 and 125 mg of sodium [1-¹³C] acetate were used for the first and second feeding of the cultures in 25 flasks, leading to 0.5 mg of **2**. The feeding experiments with sodium [2-¹³C] acetate and sodium [1,2-¹³C] acetate followed the same procedure, resulting in 1.5 and 1.0 mg of **2**, respectively.

293 and 147 mg of [2-¹³C] malonic acid were used for the first and second feeding of the cultures in 22 flasks, yielding 2.5 mg of **2**.

200 and 100 mg of sodium [2-¹³C] propionate were used for the first and second feeding of the cultures in 12 flasks, giving 0.6 mg of **2**.

180 and 120 mg of [methyl-¹³C]-L-methionin were used for the first and second feeding of the cultures in 40 flasks, leading to 6.0 mg of **2**.

Feeding experiments in *A. nidulans* expression mutants with UttA_G1778V and UttA_H1850A

A. nidulans with empty vector, UttA_G1778V or UttA_H1850A mutant was cultivated as duplicate in 250 mL flask containing 50 mL PDB medium at 230 rpm and 30 °C. 3 mL propionic acid (3 g) were diluted with NaOH solution to 6 ml stock solution (pH 7). 650 μL (0.325 g propionic acid) of this solution were added into each of the two days-old cultures. Another 350 μL (0.175 g propionic acid) solution were added one day later. After cultivation in PD liquid medium at 230 rpm and 30 °C for 4 days, the cultures were extracted with EtOAc, dissolved in ACN and subjected to LC-MS analysis.

Precursor feeding in Δ*uttA*, Δ*uttD* and *uttJ* overproduction mutants

SUPPORTING INFORMATION

Compound **3** was dissolved in DMSO to a concentration of 18 mg/mL. 277 μ L (5 mg) of this solution was added into *A. ustus* Δ uttA cultures in 250 mL flasks containing 50 mL PDB medium each after fermentation at 230 rpm and 30 °C for three days. 2 L total culture in 40 flasks were used for this experiment and harvested after 7 days. 1.0 mg of **7** and 1.0 mg of **8** were obtained and subjected to NMR and MS analyses. Feeding compound **3** into *A. nidulans*-pYH-gpdA-uttJ-pyrG transformant led to the isolation of 12.4 mg of **6** for NMR and MS analyses. Compound **3** was also fed to *A. ustus* Δ uttD mutant.

1.0 mg (100 μ L) of **4** was administered to a 100 mL flask containing 10 mL *A. ustus* Δ uttA culture after 3 days fermentation at 230 rpm and 30 °C. 1 mL culture was extracted for LC-MS analysis.

13. LC-MS analysis

LC-MS analysis was carried out on an Agilent HPLC 1260 series system equipped with a Bruker microTOF QIII mass spectrometer by using an Agilent Eclipse XDB C18 column (5 μ m, 4.6 \times 150 mm). Separation was performed at a flow rate of 0.25 mL/min with a 40 min linear gradient from 5 to 100% ACN in H₂O, both containing 0.1% (v/v) HCOOH. The column was then washed with 100% ACN for 5 min and equilibrated for 5 min. The parameters of the spectrometer were set as the following: electrospray positive ion mode for ionization, capillary voltage with 4.5 kV, collision energy with 8.0 eV. Sodium formate was used in each run for mass calibration. The masses were scanned in the range of m/z 100 – 1500. Data were evaluated with the Compass DataAnalysis 4.2 software (Bruker Daltonik, Bremen, Germany).

14. NMR analysis

NMR spectra of the isolated products were recorded at room temperature on a JEOL ECA-500 spectrometer (JEOL, Akishima, Tokyo, Japan). The samples were dissolved in DMSO- d_6 or CDCl₃. All spectra were processed with MestReNov.9.0.0 (Mestrelab Research, Santiago de Compostella, Spain).

The ¹³C enrichments were calculated by comparison of the integrals of the ¹³C signals in the ¹³C NMR spectra of **2**. The integrals of the C-12 signal at δ_C 169.1 ppm in both labeled and unlabeled samples were chosen as reference and set as 1.0. The integrals of other signals were normalized and expressed as relative values to this signal. For a given carbon, the enrichment is the ratio of the normalized value of the labeled to unlabeled sample.

15. Physicochemical properties of the compounds described in this study

2: yellow powder; ¹H and ¹³C NMR data given in Table S5; UV spectrum in Figure S21; HRMS (ESI) m/z : [M + Na]⁺ calcd. for C₁₇H₂₀NaO₈ 375.1050; found 375.1057.

3: white powder; ¹H and ¹³C NMR data given in Table S6; UV spectrum in Figure S21; HRMS (ESI) m/z : [M + H]⁺ calcd. for C₁₀H₁₃O₄ 197.0808; found 197.0810 .

SUPPORTING INFORMATION

4: white powder; ^1H NMR data given in Table S7; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd. for $\text{C}_9\text{H}_{11}\text{O}_4$ 183.0652; found 183.0652.

5: white powder; ^1H NMR data given in Table S7; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd. for $\text{C}_9\text{H}_{11}\text{O}_4$ 183.0652; found 183.0656.

6: white powder; ^1H NMR data given in Table S7; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd. for $\text{C}_{10}\text{H}_{13}\text{O}_3$ 181.0859; found 181.0862.

7: yellow powder; ^1H and ^{13}C NMR data given in Table S8; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ calcd. for $\text{C}_{19}\text{H}_{23}\text{O}_5$ 331.1540; found 331.1544.

8: yellow powder; ^1H and ^{13}C NMR data given in Table S9; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{Na}]^+$ calcd. for $\text{C}_{11}\text{H}_{14}\text{NaO}_4$ 233.0784; found 233.0784.

10: yellow oil; ^1H and ^{13}C NMR data given in Table S10; UV spectrum in Figure S21; HRMS (ESI) m/z : $[\text{M} + \text{Na}]^+$ calcd. for $\text{C}_{14}\text{H}_{16}\text{NaO}_6$ 303.0839; found 303.0845.

16. Structural elucidation

The structures of the isolated products were elucidated by comprehensive interpretation of their MS and NMR data (Figures S21 – S45). By comparison with the literature data, **4**,⁷ **5**⁸ and **6**⁹ were identified as known compounds.

The molecular formula of **2** was deduced from its HRMS as $\text{C}_{17}\text{H}_{20}\text{O}_8$. **2** was obtained as a triacetylated derivative of **1** after acetylation of the crude extract from the wild type, which was confirmed by detection of the difference of their $[\text{M} + \text{Na}]^+$ ions and by the presence of the corresponding signals for three acetyl groups in the NMR spectra of **2** (δ_{C} 169.1 ppm, C-12, δ_{H} 2.32 and δ_{C} 20.4 ppm, CH_3 -13; δ_{C} 170.1 ppm, C-14, δ_{H} 1.96 and δ_{C} 20.6 ppm, CH_3 -15; δ_{C} 170.0 ppm, C-16, δ_{H} 1.96 and δ_{C} 20.4 ppm, CH_3 -17, Table S5). The ^1H NMR data of **2** suggested also the presence of an aldehyde group (δ_{H} 10.10 ppm), a five-substituted benzene ring (δ_{H} 7.05 ppm), an aromatic methoxy group (δ_{H} 3.94 ppm), and three methylene groups (δ_{H} 5.02, 3.33 and 4.22 ppm). This was also proven by interpretation of its ^{13}C NMR data. HMBC correlations revealed that two of the methylene groups are connected to each other (Table S5 and Figure S25). Elucidation of the structure of **2** proved **1** to be 2-hydroxy-6-(2-hydroxyethyl)-3-(hydroxymethyl)-4-methoxybenzaldehyde.

The molecular formula of **3** was deduced from its HRMS data to be $\text{C}_{10}\text{H}_{12}\text{O}_4$. Interpretation of its NMR spectra including ^1H , ^{13}C , HSQC, and HMBC (Table S6 and Figures S26 – S29) revealed its structure to be 6-ethyl-2,4-dihydroxy-3-methylbenzoic acid.

SUPPORTING INFORMATION

The molecular formula of **8** was deduced from the HRMS data to be $C_{11}H_{14}O_4$ with five degrees of unsaturation. Comparison of its NMR data with those of **2** revealed the presence of signals for two instead of three methylene groups. One of them couples with a methyl group, *i.e.* as an ethyl group as observed in **3** and **5**. The signals of an aromatic methoxy group are still detectable (δ_H 3.93 and δ_C 56.1 ppm). This proved the structure of **8** to be 6-ethyl-2-hydroxy-3-(hydroxymethyl)-4-methoxybenzaldehyde.

9 was unstable and its structure was elucidated after acetylation. The molecular formula of the diacetylated derivative **10** was deduced from the HRMS data to be $C_{14}H_{16}O_6$, one oxygen more than that of **6**. The existence of OH at C-2 was verified by the signal at δ_H 12.62 ppm. Comparison of the 1H NMR spectrum of **10** with that of **6** indicated the hydroxylation of the methyl group at C-3. Differing from those in **2** and **8**, no signal for an aromatic methoxy group was detected in the spectra of **6** and **10**. This proved **9** as the hydroxylation product of C3-methyl group in **6**.

SUPPORTING INFORMATION

Table S1. Putative functions of the genes from the ustethylin (*utt*) gene cluster

| Protein | No. | Size (aa) | cover/identity, homologous protein, organism | Putative function |
|---------|----------|-----------|---|---|
| UttA | KIA75596 | 2318 | 99/46, polyketide synthase PkbA, AN6448.4, <i>Aspergillus nidulans</i> FGSC A4 | Ustethylinic acid synthase |
| UttB | KIA75595 | 486 | 95/43, oxidoreductase AzaL, G3XMD0.2, <i>Aspergillus niger</i> ATCC 1015 | FAD-dependent oxidase |
| UttC | KIA75594 | 506 | 93/48, cytochrome P ₄₅₀ Cich, AN6449.2, <i>Aspergillus nidulans</i> FGSC A4 | P ₄₅₀ , ethyl hydroxylase |
| UttD | KIA75593 | 490 | 100/55, myb-related protein B, KFX41786.1, <i>Talaromyces marneffe</i> PM1 | DNA-binding protein, positive regulator |
| UttE | KIA75592 | 503 | 95/33, bifunctional solanapyrone synthase Sol5, CEL54807.1, <i>Rhizoctonia solani</i> AG-1 IB | FAD-dependent oxidase |
| UttF | KIA75591 | 417 | 100/83, O-methyltransferase FtmD, KFX41536.1, <i>Talaromyces marneffe</i> PM1 | O-methyltransferase |
| UttG | KIA75590 | 532 | 95/49, MFS drug efflux transporter, PLN86962.1, <i>Aspergillus taichungensis</i> | MFS transporter |
| UttH | KIA75589 | 344 | 92/58, 2-oxoglutarate-dependent dioxygenase CitB, A0A159BP93.1, <i>Monascus ruber</i> 92/52, clavatul oxidase ClaD, QBK15042.1, <i>Penicillium crustosum</i> | Phenylmethyl oxidase |
| UttI | KIA75588 | 340 | 100/82, putative oxidoreductase, CEN59745.1, <i>Aspergillus calidoustus</i> | Oxidoreductase |
| UttJ | KIA75587 | 1120 | 97/42, NRPS-like CicB, AN6444.4, <i>Aspergillus nidulans</i> FGSC A4 | aryl acid reductase |

SUPPORTING INFORMATION

Table S2. Strains used in this study

| Stains | Genotype |
|------------------------------|--|
| Wild type | <i>A. ustus</i> 3.3904 |
| Δ uttA | Δ uttA::hph in <i>A. ustus</i> 3.3904 |
| Δ uttB | Δ uttB::hph in <i>A. ustus</i> 3.3904 |
| Δ uttC | Δ uttC::hph in <i>A. ustus</i> 3.3904 |
| Δ uttD | Δ uttD::hph in <i>A. ustus</i> 3.3904 |
| Δ uttE | Δ uttE::hph in <i>A. ustus</i> 3.3904 |
| Δ uttF | Δ uttF::hph in <i>A. ustus</i> 3.3904 |
| Δ uttG | Δ uttG::hph in <i>A. ustus</i> 3.3904 |
| Δ uttH | Δ uttH::hph in <i>A. ustus</i> 3.3904 |
| Δ uttJ | Δ uttJ::hph in <i>A. ustus</i> 3.3904 |
| <i>A. nidulans</i> LO8030 | <i>pyroA4, riboB2, pyrG89, nkuA::argB</i> sterigmatocystin cluster (<i>AN7804-AN7825</i>) Δ , emericeamide cluster (<i>AN2545-AN2549</i>) Δ , asperfuranone cluster (<i>AN1039-AN1029</i>) Δ , monodictyphenone cluster (<i>AN10023-AN10021</i>) Δ , terrequinone cluster (<i>AN8512-AN8520</i>) Δ , austinol cluster part 1 (<i>AN8379-AN8384</i>) Δ , austinol cluster part 2 (<i>AN9246-AN9259</i>) Δ , F9775 cluster (<i>AN7906-AN7915</i>) Δ , asperthecin cluster (<i>AN6000-AN6002</i>) Δ |
| LZ51 | <i>gpdA::uttA::AfpYrG</i> in <i>A. nidulans</i> LO8030 |
| LZ52 | <i>gpdA:: uttJ::AfpYrG</i> in <i>A. nidulans</i> LO8030 |
| LZ53 | <i>gpdA:: uttA_G1778V::AfpYrG</i> in <i>A. nidulans</i> LO8030 |
| LZ54 | <i>gpdA:: uttA_H1850A::AfpYrG</i> in <i>A. nidulans</i> LO8030 |

SUPPORTING INFORMATION

Table S3. Plasmids used and constructed in this study

| Plasmids | Description |
|--------------------------------|--|
| p5HY | Two-third of the <i>hph</i> resistance gene at the 5'-end, originated from the pUChph and inserted into pESC-URA. For gene replacement using <i>hph</i> as selection marker. |
| p3YG | Two-third of the <i>hph</i> resistance gene at the 3'-end, originated from the pUChph and inserted into pESC-URA. For gene replacement using <i>hph</i> as selection marker. |
| pLZ101(p5HY- <i>uttA</i>) | a 1171 bp US PCR fragment of <i>uttA</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ102(p3YG- <i>uttA</i>) | a 1170 bp DS PCR fragment of <i>uttA</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ103(p5HY- <i>uttB</i>) | a 1535 bp US PCR fragment of <i>uttB</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ104(p3YG- <i>uttB</i>) | a 1486 bp DS PCR fragment of <i>uttB</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ105(p5HY- <i>uttC</i>) | a 1476 bp US PCR fragment of <i>uttC</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ106(p3YG- <i>uttC</i>) | a 1463 bp DS PCR fragment of <i>uttC</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ107(p5HY- <i>uttD</i>) | a 1477 bp US PCR fragment of <i>uttD</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ108(p3YG- <i>uttD</i>) | a 1492 bp DS PCR fragment of <i>uttD</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ109(p5HY- <i>uttE</i>) | a 1527 bp US PCR fragment of <i>uttE</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ110(p3YG- <i>uttE</i>) | a 1410 bp DS PCR fragment of <i>uttE</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ111(p5HY- <i>uttF</i>) | a 1580 bp US PCR fragment of <i>uttF</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ112(p3YG- <i>uttF</i>) | a 1523 bp DS PCR fragment of <i>uttF</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ113(p5HY- <i>uttG</i>) | a 1480 bp US PCR fragment of <i>uttG</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ114(p3YG- <i>uttG</i>) | a 1449 bp DS PCR fragment of <i>uttG</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ115(p5HY- <i>uttH</i>) | a 1479 bp US PCR fragment of <i>uttH</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ116(p3YG- <i>uttH</i>) | a 1426 bp DS PCR fragment of <i>uttH</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pLZ117(p5HY- <i>uttJ</i>) | a 1416 bp US PCR fragment of <i>uttJ</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p5HY. |
| pLZ118(p3YG- <i>uttJ</i>) | a 1505 bp DS PCR fragment of <i>uttJ</i> from genomic DNA of <i>A. ustus</i> 3.3904 inserted in p3YG. |
| pYH- <i>gpdA</i> - <i>pyrG</i> | <i>URA3</i> , <i>wA</i> flanking, <i>gpdA</i> , <i>AfpyrG</i> , <i>Amp</i> |
| pLZ51 | <i>pYH-gpdA-uttA-pyrG</i> ; a 7836 bp fragment of <i>uttA</i> with its terminator from genomic DNA of <i>A. ustus</i> 3.3904 inserted in <i>pYH-gpdA-pyrG</i> |
| pLZ52 | <i>pYH-gpdA-uttJ-dMeT-pyrG</i> ; a 177 bp fragment of <i>uttA</i> was removed in <i>pYH-gpdA-pyrG</i> . |
| pLZ53 | <i>pYH-gpdA-uttA_G1778V-pyrG</i> ; mutation at Gly1778 to Val in pLZ51. |
| pLZ54 | <i>pYH-gpdA-uttA_H1850A-pyrG</i> ; mutation at His1778 to Ala in pLZ51. |

US: upstream; DS: downstream

SUPPORTING INFORMATION

Table S4. Primers used in this study

| Primers | Sequence 5'-3' | Targeted amplification |
|-----------|--|---|
| P5HY | CAAGACCAATGCGGAGCATATAC | 2/3 of the <i>hph</i> resistance gene at the 5'-end from pUChph to construct p5HY |
| P3YG | GAATTGATTCCGGAAGTGCTTGAC | 2/3 of the <i>hph</i> resistance gene at the 3'-end from pUChph to construct p3YG |
| p5HY-R | GCTGAAGTCGATTTGAGTCCAC | US of <i>hph</i> to verify 5F of <i>A. ustus</i> 3.3904 mutant |
| p3YG-F | GCATTAATGCATTGGACCTCGC | DS of <i>hph</i> to verify 3F of <i>A. ustus</i> 3.3904 mutant |
| uttA-U-F | AAGAATTGTTAATTAAGAGCTCAGATCAAGAAGTGGGATCCGAAGGG | 1171bp US fragment of <i>uttA</i> to construct pLZ101 |
| uttA-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGTTGAAGCGTGCGGAAAGAG | |
| uttA-D-F | ACTCACTATAGGGCCCGGGCGTTCGAGGATGGATGGATGAGCTGGAT | 1170 bp DS fragment of <i>uttA</i> to construct pLZ102 |
| uttA-D-R | TAGCCGCGGTACCAAGCTTACTCGACTATATCTGCGTACTGGTGCG | |
| uttA-F | CTTGAAATCCTTCGGGAGCAAC | 1594 bp partial fragment of <i>uttA</i> |
| uttA-R | GTTGAACACCTTGACACGAGC | |
| uttA-5F-F | CCCCTGCAATTTTGATCGAC | US of <i>hph</i> to verify Δ <i>uttA</i> mutant |
| uttA-3F-R | CTTGAAATCCTTCGGGAGCAAC | DS of <i>hph</i> to verify Δ <i>uttA</i> mutant |
| uttB-U-F | AAGAATTGTTAATTAAGAGCTCAGATCAACGTAACAGCAGGAAGCGA | 1535 bp US fragment of <i>uttB</i> to construct pLZ103 |
| uttB-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGTCCACCGAACCGAGGAAAAGA | |
| uttB-D-F | ACTCACTATAGGGCCCGGGCGTTCGAGTACTTCGCAATGAGGGGGA | 1486 bp DS fragment of <i>uttB</i> to construct pLZ104 |
| uttB-D-R | TAGCCGCGGTACCAAGCTTACTCGAACGCACAAACACCGACATAG | |
| uttB-F | GCAAGCTTGTCGACGGAGCTCGAATTCCTAAACAACCGGCAACCCATTA | 1754 bp partial fragment of <i>uttB</i> |
| uttB-R | GGACAGCAAATGGGTCGCGGATCCATGGTTTCGTTCCCTTCGATTCACAC | |
| uttB-5F-F | CAACGAAAGACTCGAAGAGCTG | US of <i>hph</i> to verify Δ <i>uttB</i> mutant |
| uttB-3F-R | GCTAGAATTGCATTGCAGGCTG | DS of <i>hph</i> to verify Δ <i>uttB</i> mutant |
| uttC-U-F | AAGAATTGTTAATTAAGAGCTCAGATCAGCTAGCTAGCTAGCAAGGT | 1476 bp US fragment of <i>uttC</i> to construct pLZ105 |
| uttC-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGTCGGTCGTTCTTTCGTTTCG | |

SUPPORTING INFORMATION

Table S4. (continued)

| | | |
|-----------|---|--|
| uttC-D-F | ACTCACTATAGGGCCCGGGCGTTCGAATTTGCACCCATCCAGCTAG | 1463 bp DS fragment of <i>uttC</i> to construct pLZ106 |
| uttC-D-R | TAGCCGCGGTACCAAGCTTACTCGATGGAGACGGTGATCAGGTTC | |
| uttC-F | TCTACACAAGCATCGCACTGAC | 1555 bp partial fragment of <i>uttC</i> |
| uttC-R | AGTAAGAAGTGCCCTCCCA | |
| uttC-5F-F | TATCTGCTGAAACGCCTCCT | US of <i>hph</i> to verify Δ <i>uttC</i> mutant |
| uttC-3F-R | CATTGAACGAAGCCAGCGTC | DS of <i>hph</i> to verify Δ <i>uttC</i> mutant |
| uttD-U-F | AAGAATTGTTAATTAAGAGCTCAGATCAGCGGATCGTATCGGAGAAG | 1477 bp US fragment of <i>uttD</i> to construct pLZ107 |
| uttD-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGTGAAGAATGGTTGCGGGGAT | |
| uttD-D-F | ACTCACTATAGGGCCCGGGCGTTCGATACCTTCAAGGGTATCTGGCG | 1456 bp DS fragment of <i>uttD</i> to construct pLZ108 |
| uttD-D-R | TAGCCGCGGTACCAAGCTTACTCGAGAACAGGGCAGTGGAATCTTC | |
| uttD-F | CATCAATGGGCGTATTCCACG | 1492 bp partial fragment of <i>uttD</i> |
| uttD-R | CGGTGGATCAAGCTGGATAGT | |
| uttD-5F-F | CACACCACTGCACAAGTACTAG | US of <i>hph</i> to verify Δ <i>uttD</i> mutant |
| uttD-3F-R | GTCGATGATGCCTCACCCAT | DS of <i>hph</i> to verify Δ <i>uttD</i> mutant |
| uttE-U-F | AAGAATTGTTAATTAAGAGCTCAGATCTGCTTGGGCCACTAGATACAG | 1527 bp US fragment of <i>uttE</i> to construct pLZ109 |
| uttE-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGAGATACCTCACCATCTTGCCC | |
| uttE-D-F | ACTCACTATAGGGCCCGGGCGTTCGAGGACACAAGGAGCACATGTTTG | 1410 bp DS fragment of <i>uttE</i> to construct pLZ110 |
| uttE-D-R | TAGCCGCGGTACCAAGCTTACTCGACACACCAATCTCCACTTCCG | |
| uttE-F | CCGCAAGCTTGTGACGGAGCTCGAATTCCTACCGCCGAGGGAGCTTTT | 1909 bp partial fragment of <i>uttE</i> |
| uttE-R | GGTGGACAGCAAATGGGTGCGGGATCCATGCGCGCAACAACTGCTTCAA | |
| uttE-5F-F | TCGGACTGGAAGTCGCTCTTT | US of <i>hph</i> to verify Δ <i>uttE</i> mutant |
| uttE-3F-R | GAAGAATGACGGCTACAACAGC | DS of <i>hph</i> to verify Δ <i>uttE</i> mutant |

SUPPORTING INFORMATION

Table S4. (continued)

| | | |
|-----------|---|--|
| uttF-U-F | AAGAATTGTTAATTAAGAGCTCAGATCTGCTTGGGCCACTAGATACAG | 1580 bp US fragment of <i>uttF</i> to construct pLZ111 |
| uttF-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGAGATACCTCACCATCTTGCCC | |
| uttF-D-F | ACTCACTATAGGGCCCCGGGCGTTCGAGGACACAAGGAGCACATGTTTG | 1523 bp DS fragment of <i>uttF</i> to construct pLZ112 |
| uttF-D-R | TAGCCGCGGTACCAAGCTTACTCGACACACCAATCTCCACTTCCG | |
| uttF-F | TTGCTGATCGCAGTCTTGACTG | 1223 bp partial fragment of <i>uttF</i> |
| uttF-R | TGGACTTTAATTGTGCGGGGTG | |
| uttF-5F-F | TCGGACTGGAAGTCGCTCTTT | US of <i>hph</i> to verify Δ <i>uttF</i> mutant |
| uttF-3F-R | GAAGAATGACGGCTACAACAGC | DS of <i>hph</i> to verify Δ <i>uttF</i> mutant |
| uttG-U-F | AAGAATTGTTAATTAAGAGCTCAGATCTCCCACAGTGGACATATCCG | 1480 bp US fragment of <i>uttG</i> to construct pLZ113 |
| uttG-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGCCCTAACACGTAACAACCTCGC | |
| uttG-D-F | ACTCACTATAGGGCCCCGGGCGTTCGAGGGGAAGGAAGGAATGGGTTA | 1449 bp DS fragment of <i>uttG</i> to construct pLZ114 |
| uttG-D-R | TAGCCGCGGTACCAAGCTTACTCGAGATCTCGCGGTAGAACGAGT | |
| uttG-F | TGCTGGCCATCCTCACTTCAA | 1451 bp partial fragment of <i>uttG</i> |
| uttG-R | CAGACTATCCGCAATCGTGCT | |
| uttG-5F-F | GCCTGACTTCAAGAGTGAGACT | US of <i>hph</i> to verify Δ <i>uttG</i> mutant |
| uttG-3F-R | TGCGCCGTGAAGAAGTCATG | DS of <i>hph</i> to verify Δ <i>uttG</i> mutant |
| uttH-U-F | AAGAATTGTTAATTAAGAGCTCAGATCCATACCACCATCAGCAGAAACC | 1479 bp US fragment of <i>uttH</i> to construct pLZ115 |
| uttH-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGGATGAAGGTGGTGATGATCGTG | |
| uttH-D-F | ACTCACTATAGGGCCCCGGGCGTTCGATTGGTTACCGGATGCGGTTG | 1426 bp DS fragment of <i>uttH</i> to construct pLZ116 |
| uttH-D-R | TAGCCGCGGTACCAAGCTTACTCGAGAAGGCGATTGTTAGTACGCC | |
| uttH-F | GCAAAAACCGCACCGACTCAA | 975 bp partial fragment of <i>uttH</i> |
| uttH-R | ATAACTCTCCGCAACCCTCC | |

SUPPORTING INFORMATION

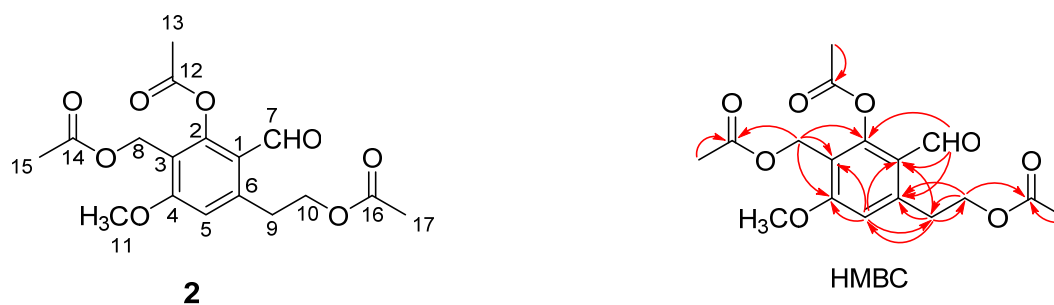
Table S4. (continued)

| | | |
|----------------|---|---|
| uttH-5F-F | ATGTCACGACCACTCGCTTGA | US of <i>hph</i> to verify Δ <i>uttH</i> mutant |
| uttH-3F-R | CGATTGTGACGACGAGAAACAC | DS of <i>hph</i> to verify Δ <i>uttH</i> mutant |
| uttJ-U-F | AAGAATTGTTAATTAAGAGCTCAGATCGTCTTCGCTGTAAGTCCACAG | 1416 bp US fragment of <i>uttJ</i> to construct pLZ119 |
| uttJ-U-R | ACCCTCACTAAAGGGCGGCCGCACTAGTCGGGTAAAGACACTAGGATGG | |
| uttJ-D-F | ACTCACTATAGGGCCCGGGCGTCAATGGAGTTCCAGGGTCTCTCT | 1505 bp DS fragment of <i>uttJ</i> to construct pLZ120 |
| uttJ-D-R | TAGCCGCGGTACCAAGCTTACTCGAGCCTCATCCTTCACATCATCCA | |
| uttJ-F | ACGAATATCCGGAAGATACCCC | 1529 bp partial fragment of <i>uttJ</i> |
| uttJ-R | ACGATCTGATCCGTCTAGCGA | |
| uttJ-5F-F | CGCGATCCTCGACTTTTCTAGG | US of <i>hph</i> to verify Δ <i>uttJ</i> mutant |
| uttJ-3F-R | AGCACTGCTGTTCAAGGCATAC | DS of <i>hph</i> to verify Δ <i>uttJ</i> mutant |
| HE-uttA-P1-F | TCATCTTCCCATCCAAGAACCTTTAATCATGGTCGTCGAAGGGTATCCA | 4014 bp partial fragment of <i>uttA</i> from <i>A. ustus</i> 3.3904 to construct pLZ51 |
| HE-uttA-P1-R | CGGAGGTCTTTTTCGATCCCT | |
| HE-uttA-P2-F | GTTTAGTCGGGTGCTCATCTCC | 4095 bp partial fragment of <i>uttA</i> with its 624 bp terminator from <i>A. ustus</i> 3.3904 to construct pLZ51 |
| HE-uttA-P2-R | CGTCAGACACAGAATAACTCTCGCTAGGCTTCTTCCCGCTTCTGAAGT | |
| HE-uttJ-F | CTTCCCATCCAAGAACCTTTAATCATGTGCGTGATTAATGGATCTGAAG | 4229 bp of <i>uttJ</i> with its 651 bp terminator from <i>A. ustus</i> 3.3904 to construct pLZ52 |
| HE-uttJ-R | CATATTCGTCAGACACAGAATAACTCTCACGAAGTCTGAGCCCCCTCAAAA | |
| UttA-G1778V-F | GGGACAGTCTCGACGACGAAATGGGTTGTCGATGCGCT | Containing a mutation for G1778 to Val to construct pLZ53 |
| UttA-G1778V-R | CGTCGTCGAGACTGTCCCCGCGCCGAGTTCAGGATGC | |
| UttA-H1850A-F | TGCATCGCCGACGAGCAATCTGCCAACTCGCTCAC | Containing a mutation for H1850 to Ala to construct pLZ54 |
| UttA- H1850A-R | GCTCGTGGCGCGATGCAGTTGGTCGAGAGGATGGTAT | |
| UttA-muta-F | TCTGGAAGAACGTGTACCCAC | 1203 bp of <i>uttA</i> fragment including the mutated points to construct pLZ53 and pLZ54 |
| UttA-muta-R | TTGGTGAGGATAATCCCGCTG | |

US: upstream; DS: downstream

SUPPORTING INFORMATION

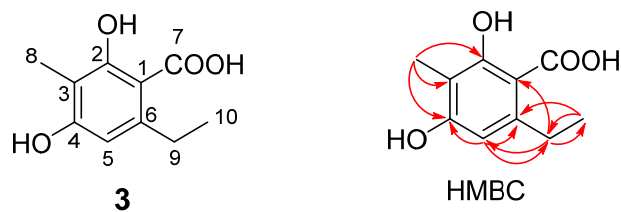
Table S5. NMR data of compound **2** in DMSO-*d*₆ (500 MHz, δ in ppm, *J* in Hz)



| Position | δ_c | δ_H (multi. <i>J</i>) | HMBC |
|----------|--------------------------|-------------------------------|---------------------|
| 1 | 119.8 (C) | – | – |
| 2 | 152.5 (C) | – | – |
| 3 | 116.0 (C) | – | – |
| 4 | 162.1 (C) | – | – |
| 5 | 112.1 (CH) | 7.05 (s) | C-1, C-3, C-4, C-9 |
| 6 | 144.9 (C) | – | – |
| 7 | 189.0 (CH) | 10.10 (s) | C-1, C-2, C-6 |
| 8 | 54.1 (CH ₂) | 5.02 (s) | C-2, C-3, C-4, C-14 |
| 9 | 31.5 (CH ₂) | 3.33 (t, <i>J</i> = 6.6 Hz) | C-1, C-5, C-6, C-10 |
| 10 | 63.9 (CH ₂) | 4.22 (t, <i>J</i> = 6.6 Hz) | C-1, C-9, C-16 |
| 11 | 56.6 (OCH ₃) | 3.94 (s) | C-4 |
| 12 | 169.1 (C) | – | – |
| 13 | 20.4 (CH ₃) | 2.32 (s) | C-12 |
| 14 | 170.1 (C) | – | – |
| 15 | 20.6 (CH ₃) | 1.96 (s) | C-14 |
| 16 | 170.0 (C) | – | – |
| 17 | 20.4 (CH ₃) | 1.96 (s) | C-16 |

SUPPORTING INFORMATION

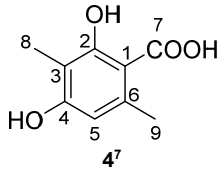
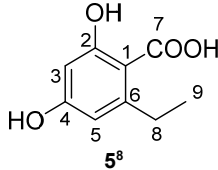
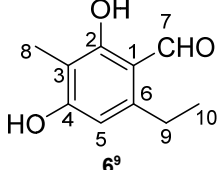
Table S6. NMR data of compound **3** in DMSO-*d*₆ (500 MHz, δ in ppm, *J* in Hz)



| Position | δ_c | δ_H (multi. <i>J</i>) | HMBC |
|----------|-------------------------|-------------------------------|--------------------|
| 1 | 103.0 (C) | – | – |
| 2 | 162.1 (C) | – | – |
| 3 | 108.0 (C) | – | – |
| 4 | 160.1 (C) | – | – |
| 5 | 108.9 (CH) | 6.28 (s) | C-1, C-4, C-6, C-9 |
| 6 | 145.8 (C) | – | – |
| 7 | 173.8 (COOH) | 9.97 (s) | – |
| 8 | 8.0 (CH ₃) | 1.93 (s) | C-2, C-3, C-4 |
| 9 | 28.8 (CH ₂) | 2.81 (q, <i>J</i> = 7.4 Hz) | C-5, C-6, C-10 |
| 10 | 16.1 (CH ₃) | 1.10 (t, <i>J</i> = 7.4 Hz) | C-6, C-9 |
| 4-OH | – | 13.33 (s) | – |
| 2-OH | – | 12.65 (s) | – |

SUPPORTING INFORMATION

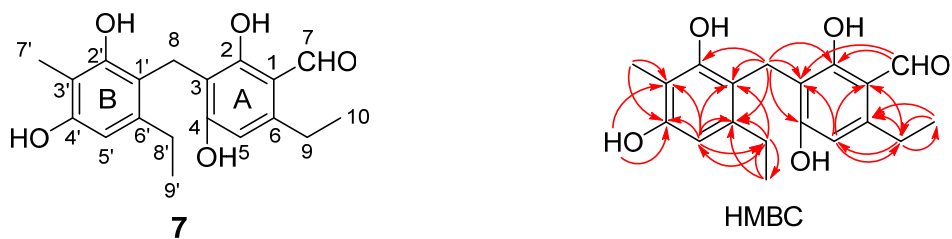
Table S7. ^1H NMR data of compounds **4** – **6** in $\text{DMSO-}d_6$ (500 MHz, δ in ppm, J in Hz)

| Compounds |  4⁷ |  5⁸ |  6⁹ |
|-----------|---|---|--|
| Position | δ_{H} (multi., J) | δ_{H} (multi., J) | δ_{H} (multi., J) |
| 3 | – | 6.13 (d, $J = 2.4$ Hz) | – |
| 5 | 6.25 (s) | 6.18 (d, $J = 2.4$ Hz) | 6.33 (s) |
| 7 | – | – | 10.63 (s) |
| 8 | 1.93 (s) | 2.78 (q, $J = 7.4$ Hz) | 1.92 (s) |
| 9 | 2.39 (s) | 1.11 (t, $J = 7.4$ Hz) | 2.83 (q, $J = 7.5$ Hz) |
| 10 | – | – | 1.17 (t, $J = 7.5$ Hz) |
| 2-OH | 12.94 (s) | 12.42 (s) | 12.77 (s) |
| 4-OH | 12.94 (s) | 12.42 (s) | 10.02 (s) |
| 7-OH | 9.99 (s) | 10.02 (s) | – |

The NMR data of **3**, **4** and **5** correspond very well to those reported previously.⁷⁻⁹

SUPPORTING INFORMATION

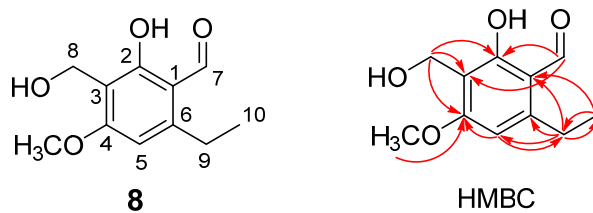
Table S8. NMR data of compound **7** in DMSO-*d*₆ (500 MHz, δ in ppm, *J* in Hz)



| Position | δ_C | δ_H (multi., <i>J</i>) | HMBC |
|----------|-------------------------|--------------------------------|---------------------------------|
| 1 | 111.5 (C) | – | – |
| 2 | 163.8 (C) | – | – |
| 3 | 112.7 (C) | – | – |
| 4 | 154.3 (C) | – | – |
| 5 | 109.4 (CH) | 6.28 (s) | C-1, C-3, C-9 |
| 6 | 148.8 (C) | – | – |
| 7 | 194.1 (CH) | 10.00 (s) | C-2, C-3 |
| 8 | 18.9 (CH ₂) | 3.71 (s) | C-2, C-3, C-4, C-1', C-2', C-6' |
| 9 | 24.3 (CH ₂) | 2.82 (q, <i>J</i> = 7.5 Hz) | C-5, C-6, C-10 |
| 10 | 17.0 (CH ₃) | 1.15 (t, <i>J</i> = 7.5 Hz) | C-6, C-9 |
| 1' | 115.6 (C) | – | – |
| 2' | 163.8 (C) | – | – |
| 3' | 108.7 (C) | – | – |
| 4' | 154.4 (C) | – | – |
| 5' | 107.5 (CH) | 6.16 (s) | C-1', C-3', C-4', C-8' |
| 6' | 141.2 (C) | – | – |
| 7' | 9.6 (CH ₃) | 1.92 (s) | C-3', C-4' |
| 8' | 25.9 (CH ₂) | 2.63 (q, <i>J</i> = 7.5 Hz) | C-1', C-5', C-6', C-9' |
| 9' | 16.1 (CH ₃) | 0.98 (t, <i>J</i> = 7.5 Hz) | C-6', C-8' |
| 4'-OH | – | 8.72 (s) | C-3', C-4' |
| 2-OH | – | 13.15 (s) | – |

SUPPORTING INFORMATION

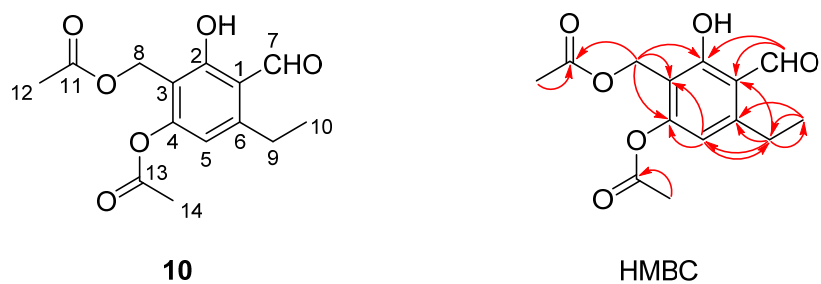
Table S9. NMR data of compound **8** in CDCl₃ (500 MHz, δ in ppm, *J* in Hz)



| Position | δ_c | δ_H (multi., <i>J</i>) | HMBC |
|--------------------|--------------------------|--------------------------------|--------------------|
| 1 | 151.2 (C) | – | – |
| 2 | 163.5 (C) | – | – |
| 3 | 114.2 (C) | – | – |
| 4 | 164.3 (C) | – | – |
| 5 | 104.0 (CH) | 6.34 (s) | C-3, C-4, C-6, C-9 |
| 6 | 112.7 (C) | – | – |
| 7 | 193.3 (CH) | 10.14 (s) | C-2, C-3 |
| 8 | 53.9 (CH ₂) | 4.76 (s) | C-2, C-3, C-4 |
| 9 | 25.7 (CH ₂) | 2.93 (q, <i>J</i> = 7.6 Hz) | C-5, C-6, C-10 |
| 10 | 17.3 (CH ₃) | 1.31 (t, <i>J</i> = 7.6 Hz) | C-1, C-9 |
| 4-OCH ₃ | 56.1 (OCH ₃) | 3.93 (s) | C-4 |
| 2-OH | – | 12.62 (s) | – |

SUPPORTING INFORMATION

Table S10. NMR data of compound **10** in DMSO-*d*₆ (500 MHz, δ in ppm, *J* in Hz)



| Position | δ_c | δ_H (multi., <i>J</i>) | HMBC |
|----------|-------------------------|--------------------------------|---------------------|
| 1 | 114.3 (C) | | |
| 2 | 162.6 (C) | | |
| 3 | 115.6 (C) | | |
| 4 | 155.9 (C) | | |
| 5 | 115.0 (CH) | 6.72 (s) | C-4, C-6, C-9 |
| 6 | 150.9 (C) | | |
| 7 | 196.1 (CH) | 10.28 (s) | C-1, C-2 |
| 8 | 53.9 (CH ₂) | 5.02 (s) | C-2, C-3, C-4, C-12 |
| 9 | 24.0 (CH ₂) | 2.99 (q, <i>J</i> = 7.6 Hz) | C-5, C-6, C-1, C-10 |
| 10 | 16.2 (CH ₂) | 1.21 (t, <i>J</i> = 7.6 Hz) | C-6, C-9 |
| 11 | 170.0 (C) | | |
| 12 | 20.4 (CH ₃) | 1.96 (s) | C-11 |
| 13 | 168.4 (C) | | |
| 14 | 20.6 (CH ₃) | 2.29 (s) | C-13 |
| 2-OH | | 12.58 (brs) | |

SUPPORTING INFORMATION

Table S11. Enrichments in **2** after feeding with ¹³C labeled precursors

| Position | δ_c | sodium [^{1-¹³C}] acetate | sodium [^{2-¹³C}] propionate | sodium [^{2-¹³C}] acetate | [^{2-¹³C}] malonic acid | [methyl- ¹³ C]-L-methionine |
|----------|--------------------------|--|---|--|--|--|
| | | | | | | |
| 1 | 119.8 (C) | 1.0 | 0.8 | 4.9 | 6.7 | 1.0 |
| 2 | 152.5 (C) | 4.5 | 3.7 | 1.1 | 1.2 | 1.1 |
| 3 | 116.0 (C) | 0.9 | 1.3 | 5.8 | 7.4 | 0.9 |
| 4 | 162.1 (C) | 3.8 | 3.6 | 1.2 | 1.0 | 1.1 |
| 5 | 112.1 (CH) | 1.0 | 1.4 | 7.9 | 8.5 | 1.2 |
| 6 | 144.9 (C) | 4.4 | 4.0 | 1.1 | 1.2 | 1.2 |
| 7 | 189.0 (CH) | 6.1 | 6.1 | 1.3 | 1.3 | 1.2 |
| 8 | 54.1 (CH ₂) | 0.9 | 1.0 | 0.9 | 1.0 | 13.1 |
| 9 | 31.5 (CH ₂) | 0.9 | 1.2 | 6.6 | 8.9 | 0.8 |
| 10 | 63.9 (CH ₂) | 0.9 | 1.0 | 1.0 | 1.1 | 15.0 |
| 11 | 56.6 (OCH ₃) | 0.8 | 0.8 | 1.1 | 1.0 | 15.9 |
| 12 | 169.1 (C) | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 13 | 20.4 (CH ₃) | 0.9 | 1.1 | 1.1 | 1.0 | 1.1 |
| 14 | 170.1 (C) | 0.9 | 0.8 | 0.9 | 1.0 | 1.0 |
| 15 | 20.6 (CH ₃) | 0.9 | 1.0 | 1.0 | 1.0 | 1.0 |
| 16 | 170.0 (C) | 1.0 | 0.9 | 1.0 | 1.0 | 1.0 |
| 17 | 20.4 (CH ₃) | 0.9 | 1.1 | 1.1 | 1.2 | 1.2 |

SUPPORTING INFORMATION

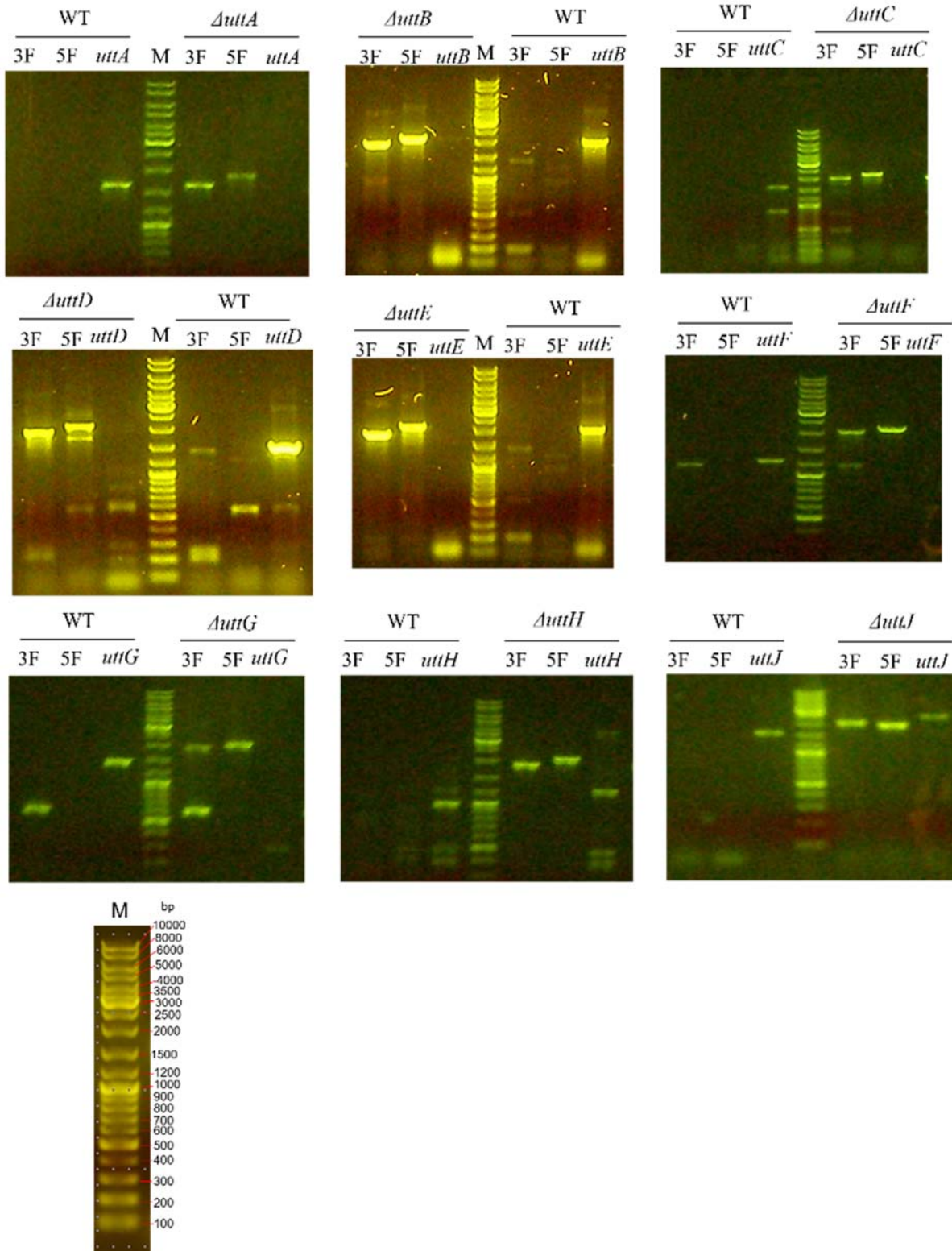


Figure S1. PCR verification of deletion mutants of *A. ustus* 3.3904.

PCR amplification for three different fragments from genomic DNA of WT and deletion mutants was used to prove the presence/absence of the target gene and the integration site of the selection marker with up- and downstream regions. The PCR primers are given in Table S4.

SUPPORTING INFORMATION

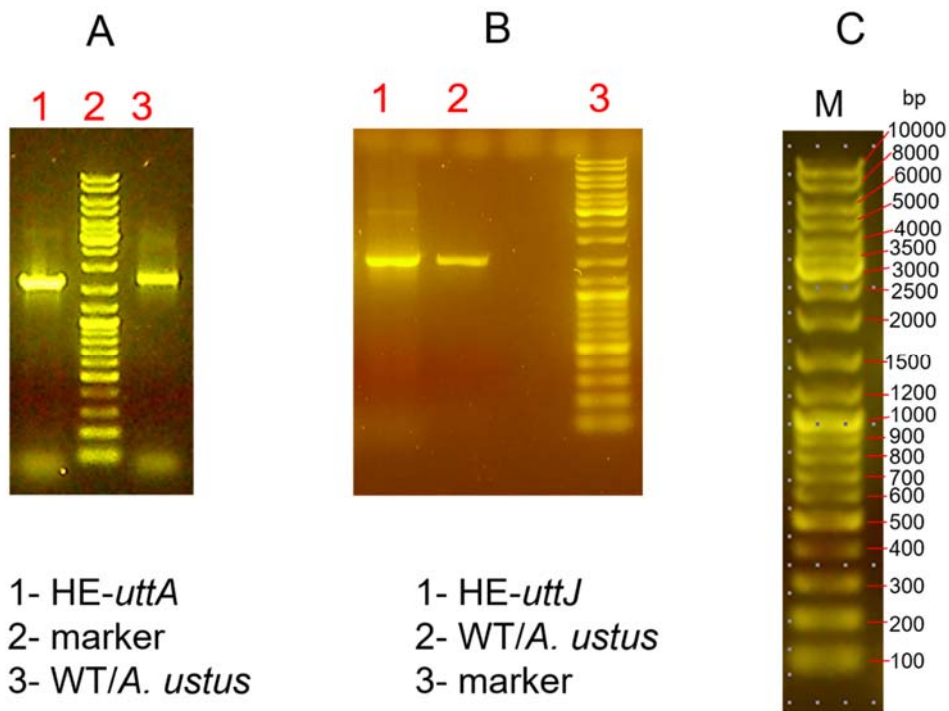


Figure S2. PCR verification of heterologous expression transformants.

A. nidulans-pYH-*gpdA-uttA-pyrG* (HE-*uttA*) (A), *A. nidulans*-pYH-*gpdA-uttJ-pyrG* (HE-*uttJ*) (B), and marker reference (C). A fragment of 1.5 kb within the target gene was amplified from the primers listed in Table S4.

SUPPORTING INFORMATION

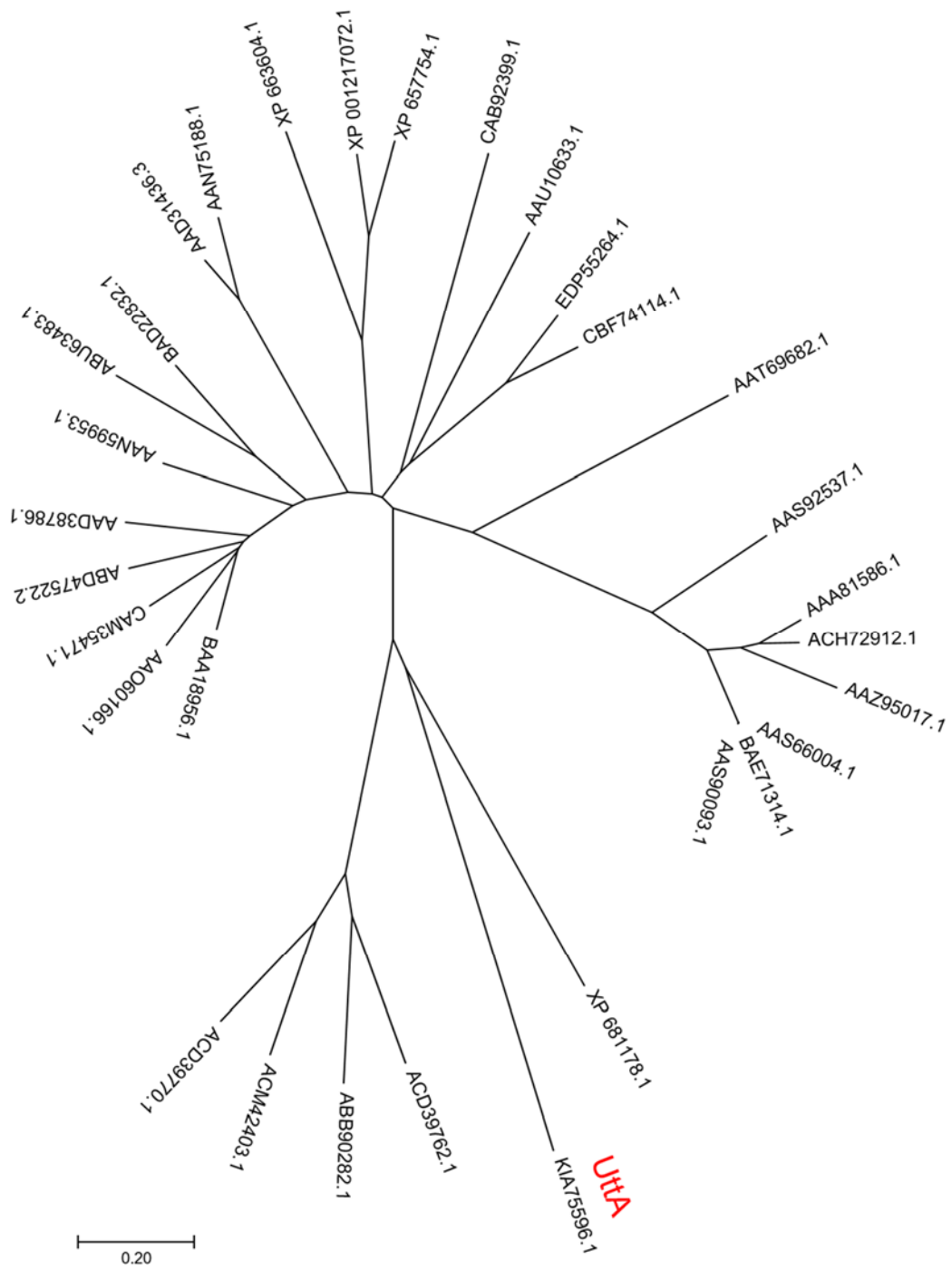


Figure S3. Phylogenetic analysis of Uta_PT domain with 30 known PT domains from fungi.

SUPPORTING INFORMATION

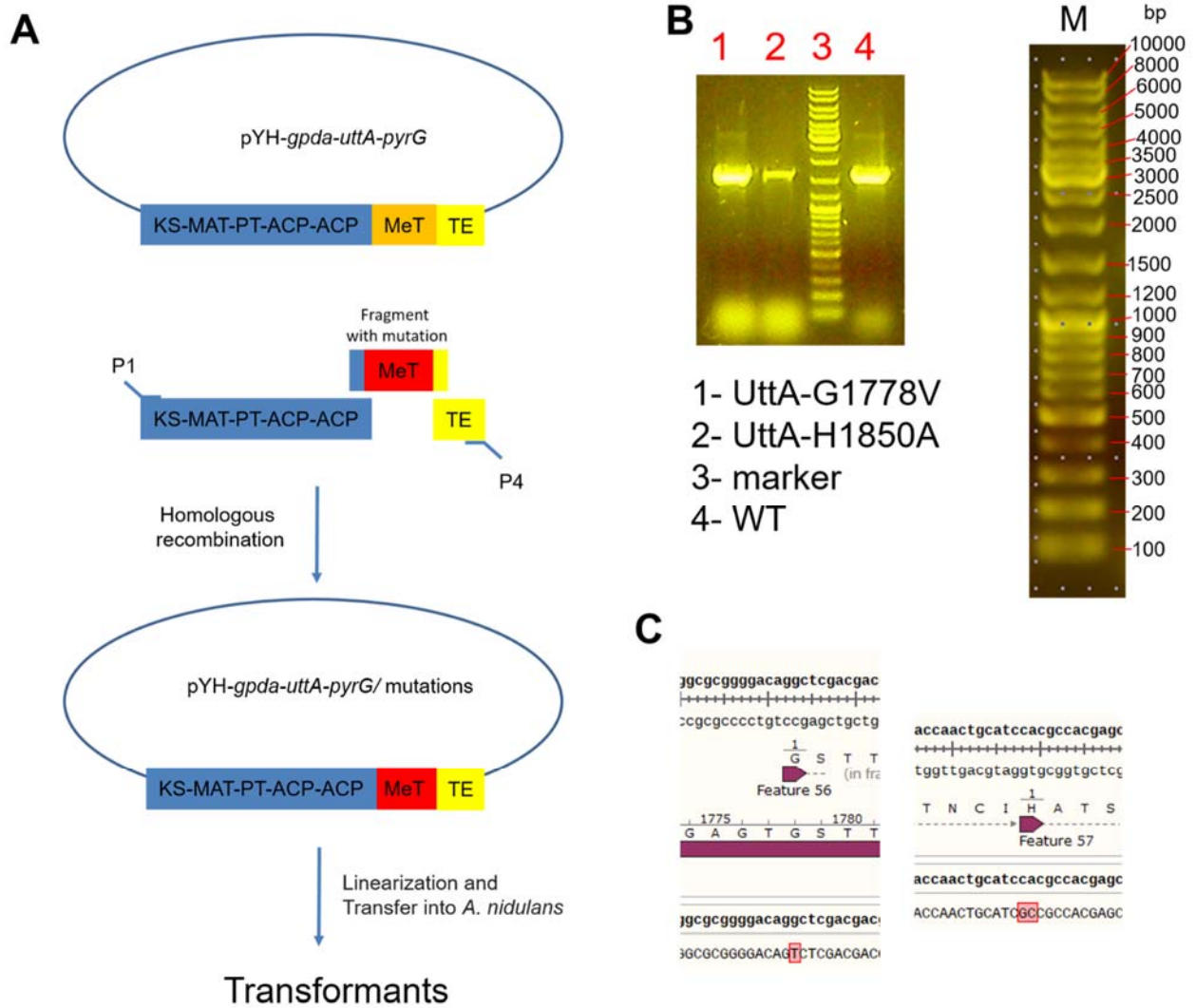


Figure S4. Point mutation in UttA and analysis of the obtained mutants.

(A) Schematic representation for the point mutation in UttA. (B) PCR verification of the mutants. (C) Sequence analysis of the UttA mutants

SUPPORTING INFORMATION

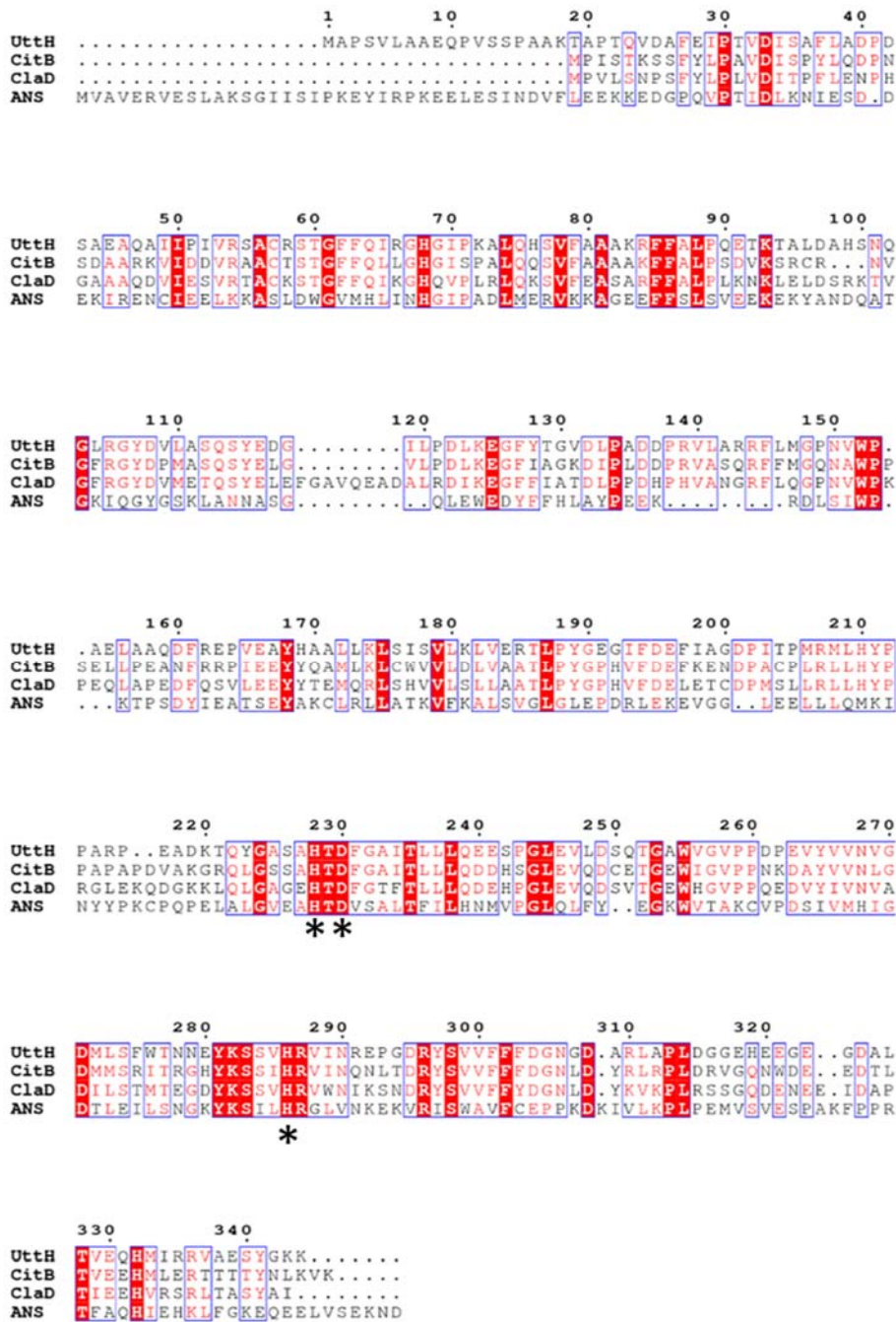


Figure S5. Sequence alignments of 2-OG-dependent oxygenases.

CitB (AL192653)¹⁰, ClaD (QBK15042), and ANS (Q96323) are from *Monascus ruber* M7, *Penicillium crustosum* and *Arabidopsis thaliana*, respectively. UttH contains the typical conserved 2-His-1-Asp ion-binding triad of non-heme Fe^{II}/2-oxoglutarate-dependent oxygenases (His228, His296 and Asp296) (marked with *) compared with the crystal structure of ANS¹¹. Protein sequence alignments were carried out by using the sequence alignment function of Multiple Sequence Alignment by CLUSTALW (<https://www.genome.jp/tools-bin/clustalw>) and visualized with ESPript 3.0 (<http://espript.ibcp.fr/ESPript/ESPript/>).

SUPPORTING INFORMATION

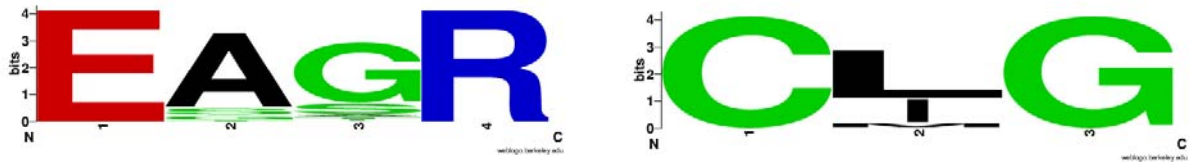


Figure S6. Weblogo illustration for the conserved ExxR and CxG motifs in UttC by using 96 P₄₅₀ enzyme in fungi. EAGR (349-352) in UttC (A); CLG (434-436) in UttC (B)

The amino acids in the following enzymes (accession number are listed) were used for analysis. **UttC (KIA75594)** is highlighted in red.

XP_664053.1_10-490, XP_033674816.1_17-507, XP_033595470.1_20-516, XP_033430050.1_21-509, XP_033391313.1_19-508, XP_026608368.1_1-495, XP_025573321.1_5-496, XP_025464315.1_5-498, XP_024692041.1_6-487, XP_016600678.1_16-516, XP_016589218.1_27-507, XP_016221951.1_24-495, XP_007806251.1_26-498, XP_007805323.1_1-429, XP_003234761.1_17-497, XP_003170503.1_31-508, XP_003071602.1_35-507, XP_002848393.1_22-522, XP_002145722.1_1-497, XP_001911463.1_5-497, XP_001905657.1_14-503, XP_001245240.2_35-507, XP_001228060.1_35-520, VBB75775.1_5-497, TVY58386.1_20-510, SLM38802.1_31-474, RYP92946.1_1-423, RYP43872.1_19-502, RVX74072.1_1-443, RMZ76241.1_762-1242, RMD39760.1_78-530, RAO64728.1_1-365, QGA14808.1_1-497, Q0CRQ3.2_13-480, PVH96859.1_19-500, POS74238.1_21-516, PLN86963.1_21-508, PCG99300.1_13-489, OXV11655.1_4-509, OTB02447.1_24-516, OTA92375.1_13-516, OTA64244.1_13-516, OSS44053.1_30-509, OJJ68595.1_1-492, OCK75412.1_40-531, KUL85074.1_3-484, KMP04844.1_28-510, KMM70176.1_35-510, KKY30492.1_16-525, KIH90520.1_27-527, **KIA75594.1_1-506**, KGO38587.1_16-495, KFY99026.1_1-431, KFY63720.1_19-514, KFH44065.1_16-507, KFA68360.1_26-512, KFA60546.1_25-483, KEY71920.1_31-483, KEY71720.1_26-512, KAF4779182.1_34-507, KAF4310737.1_17-496, KAF4228494.1_6-502, KAF3895032.1_17-525, KAF3406269.1_34-522, KAF3023069.1_9-517, KAF3012828.1_17-494, KAF2994041.1_1-422, KAF2847821.1_21-515, KAF2678715.1_16-511, KAF2258967.1_9-501, KAF2205452.1_8-503, KAF2189392.1_14-523, KAF2004773.1_15-520, KAE8553912.1_1-497, KAB8343000.1_743-1215, KAB5575480.1_31-514, KAA6412416.1_75-547, GFF93090.1_21-527, GFF54863.1_21-527, GFF23783.1_21-510, GES61355.1_13-512, GAW17174.1_5-514, GAO84805.1_21-527, GAM43297.1_34-522, EZF36180.1_17-525, EZF23312.1_17-525, ERS97303.1_27-525, EPE10457.1_18-502, EKG19934.1_1-426, EGE04287.1_17-525, EGD97508.1_17-525, EFW13271.1_28-510, CEN59739.1_4-442, CEJ60330.1_15-521, CBF69449.1_10-497

SUPPORTING INFORMATION

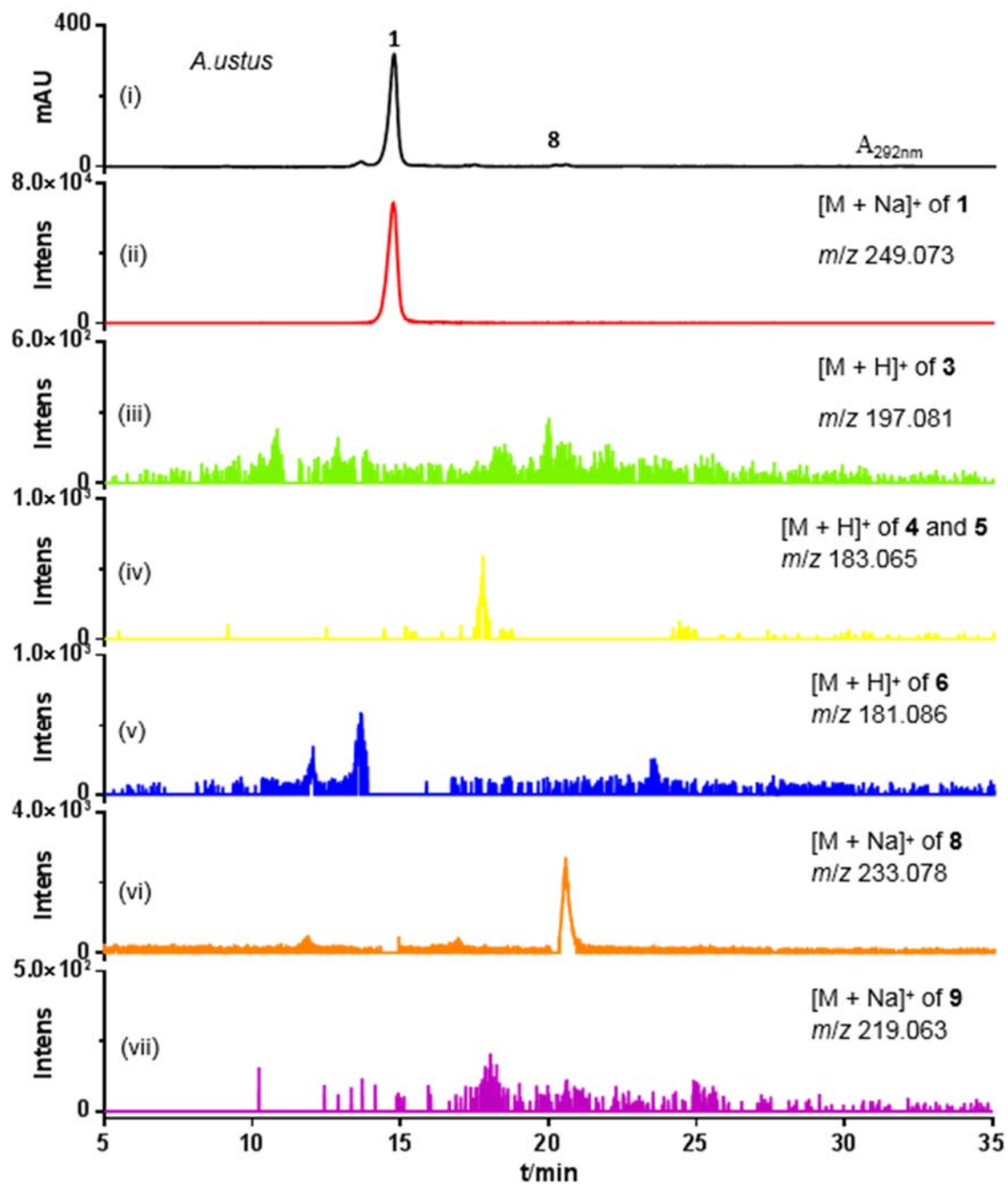


Figure S7. LC-MS analysis of the metabolite profile of the *A. ustus* wild type

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3** – **6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

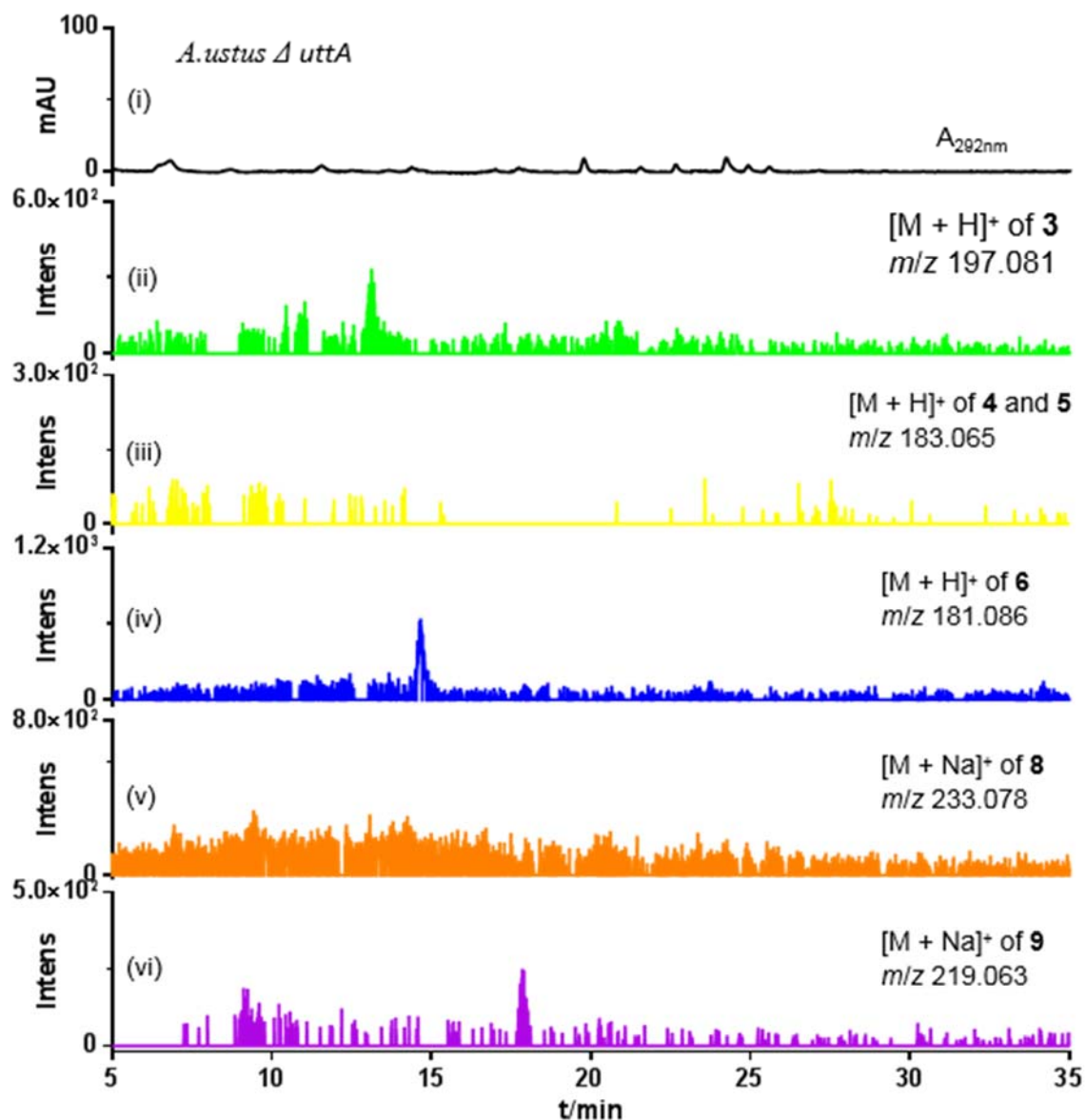


Figure S8. LC-MS analysis of the metabolite profile of the *A. ustus* Δ uttA mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **3 – 6, 8 and 9** (ii – vi).

SUPPORTING INFORMATION

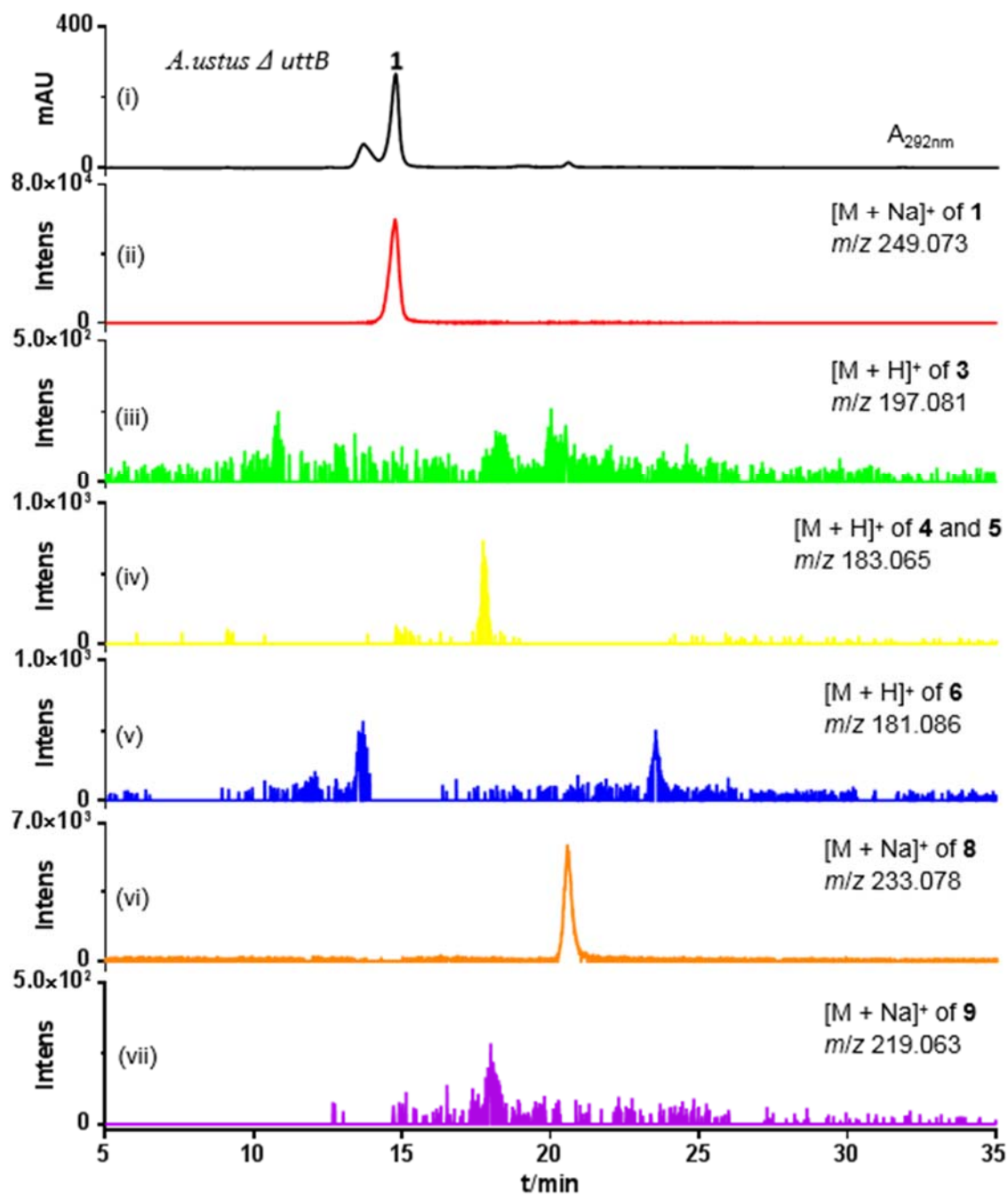


Figure S9. LC-MS analysis of the metabolite profile of the *A. ustus* Δ *uttB* mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3 – 6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

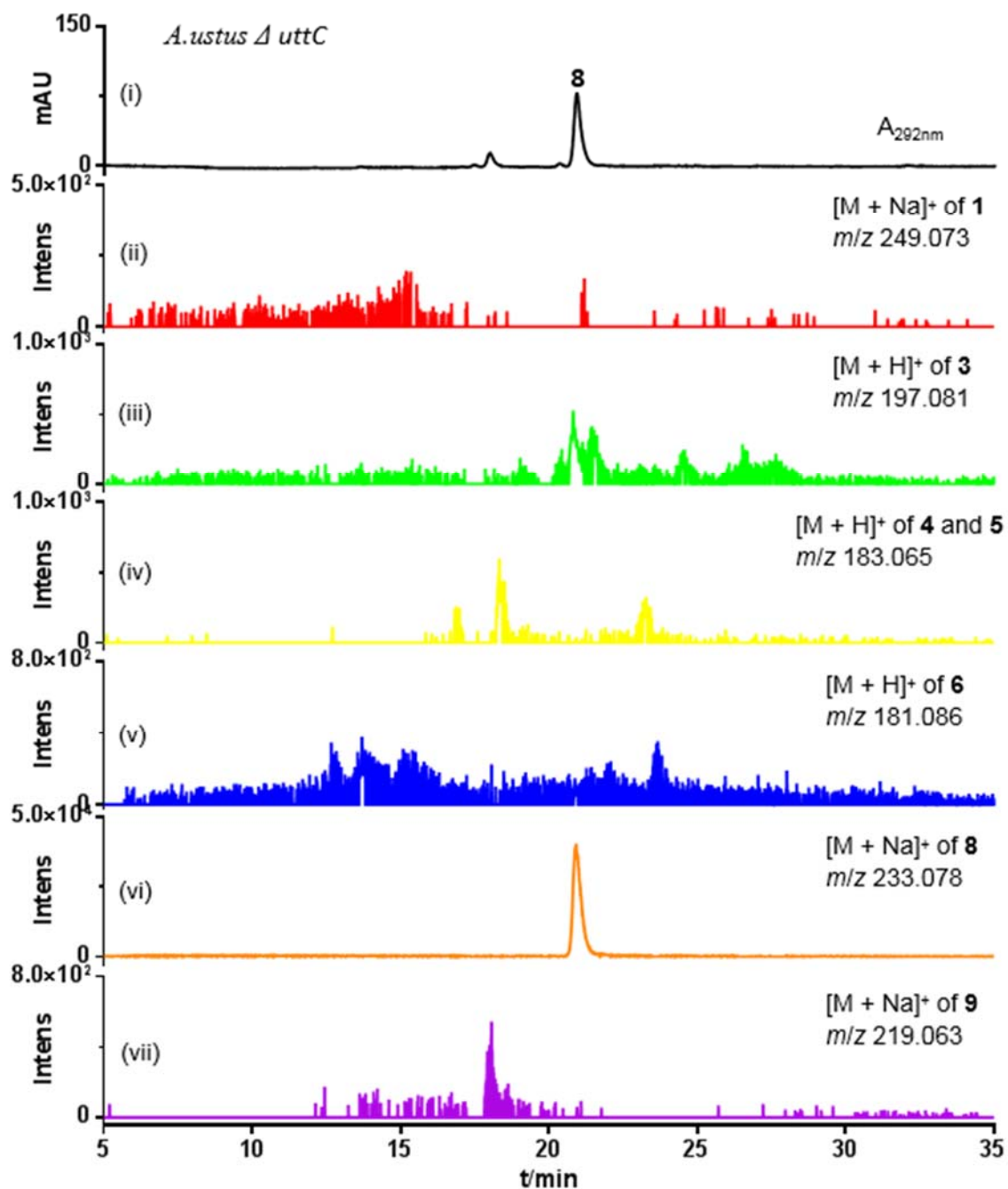


Figure S10. LC-MS analysis of the metabolite profile of the *A. ustus* Δ uttC mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3 – 6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

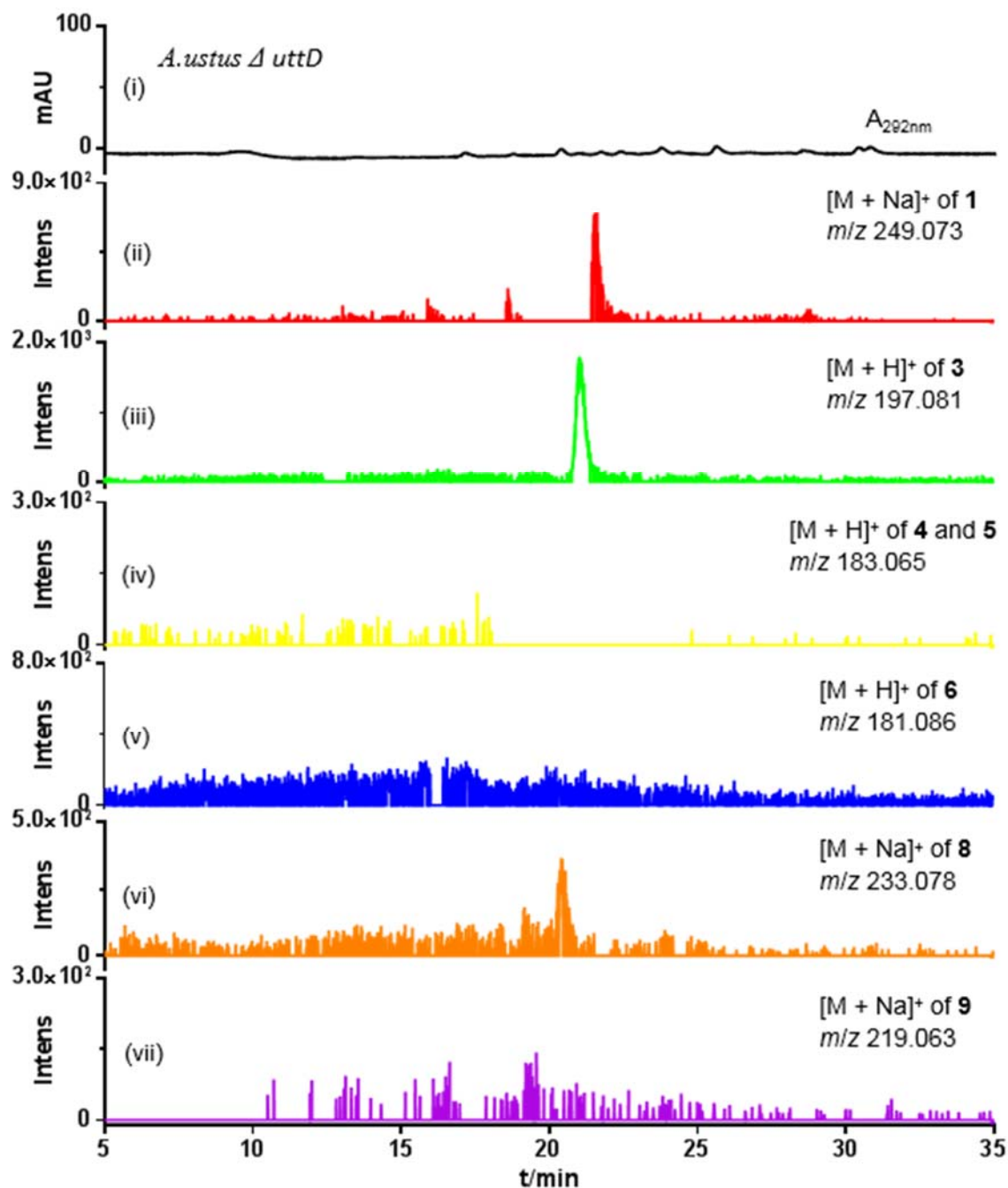


Figure S11. LC-MS analysis of the metabolite profile of the *A. ustus* Δ uttD mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3 – 6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

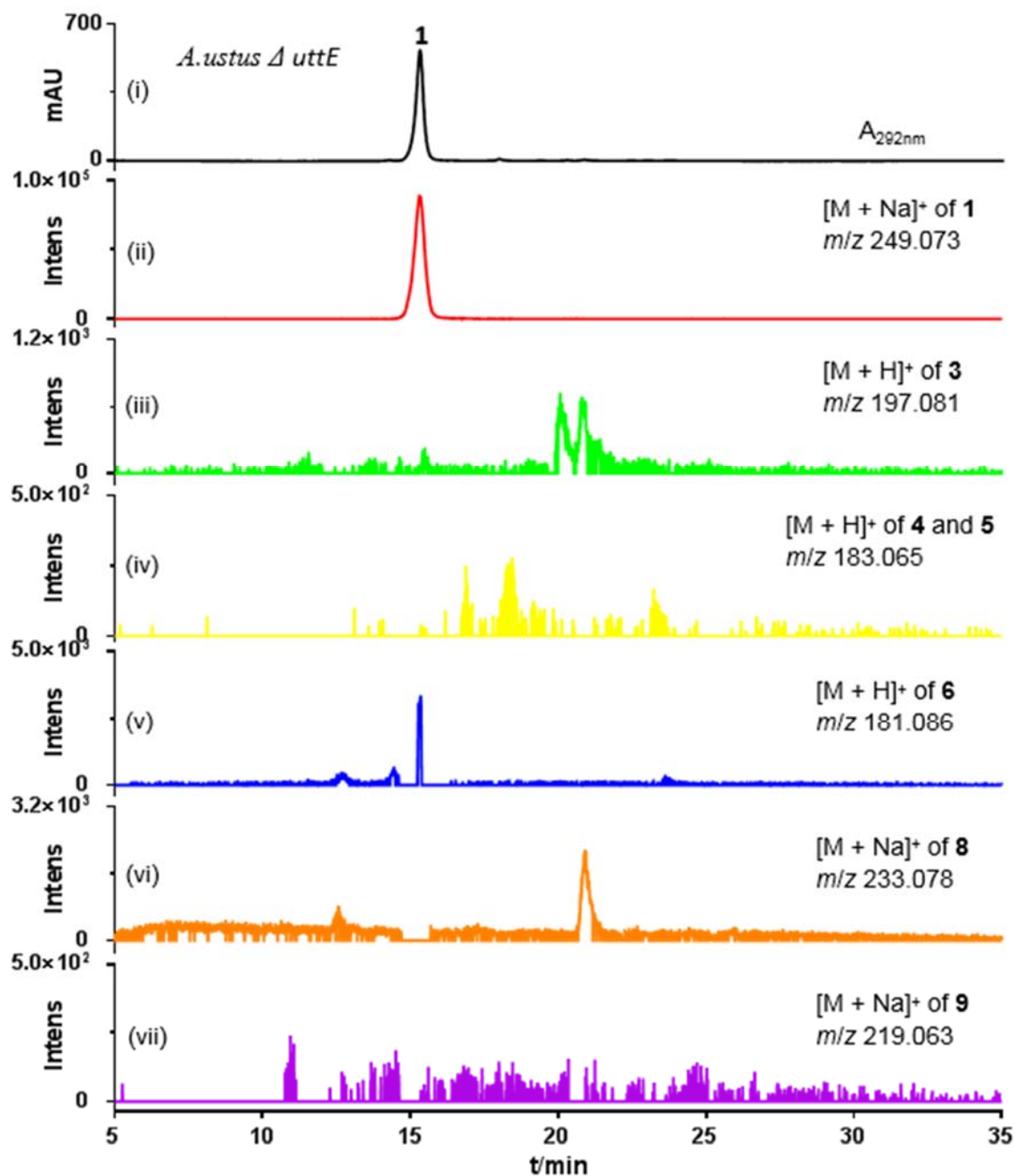


Figure S12. LC-MS analysis of the metabolite profile of the *A. ustus* Δ *uttE* mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3 – 6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

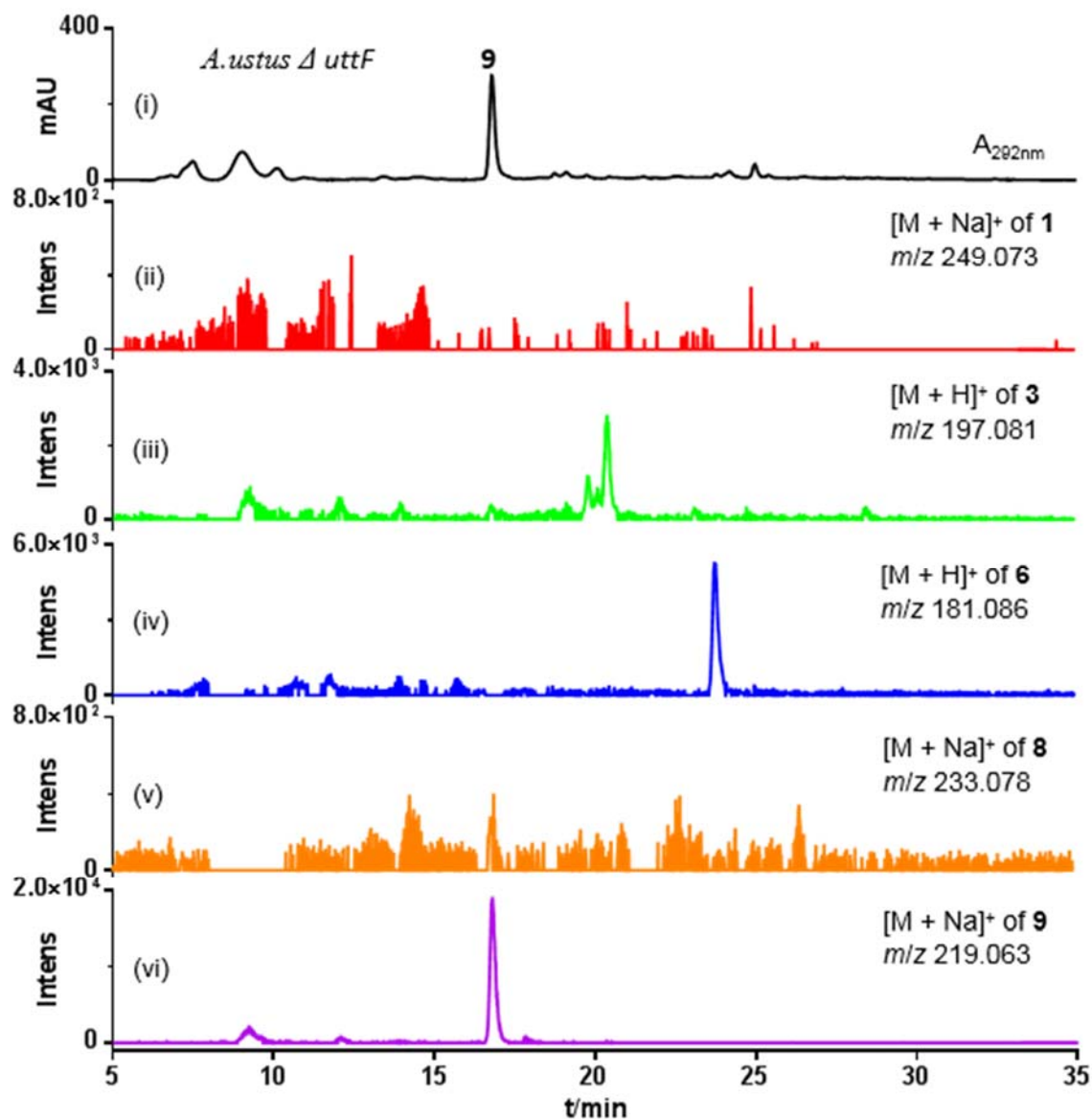


Figure S13. LC-MS analysis of the metabolite profile of the *A. ustus* Δ *uttF* mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3**, **6**, **8**, and **9** (ii – vi).

SUPPORTING INFORMATION

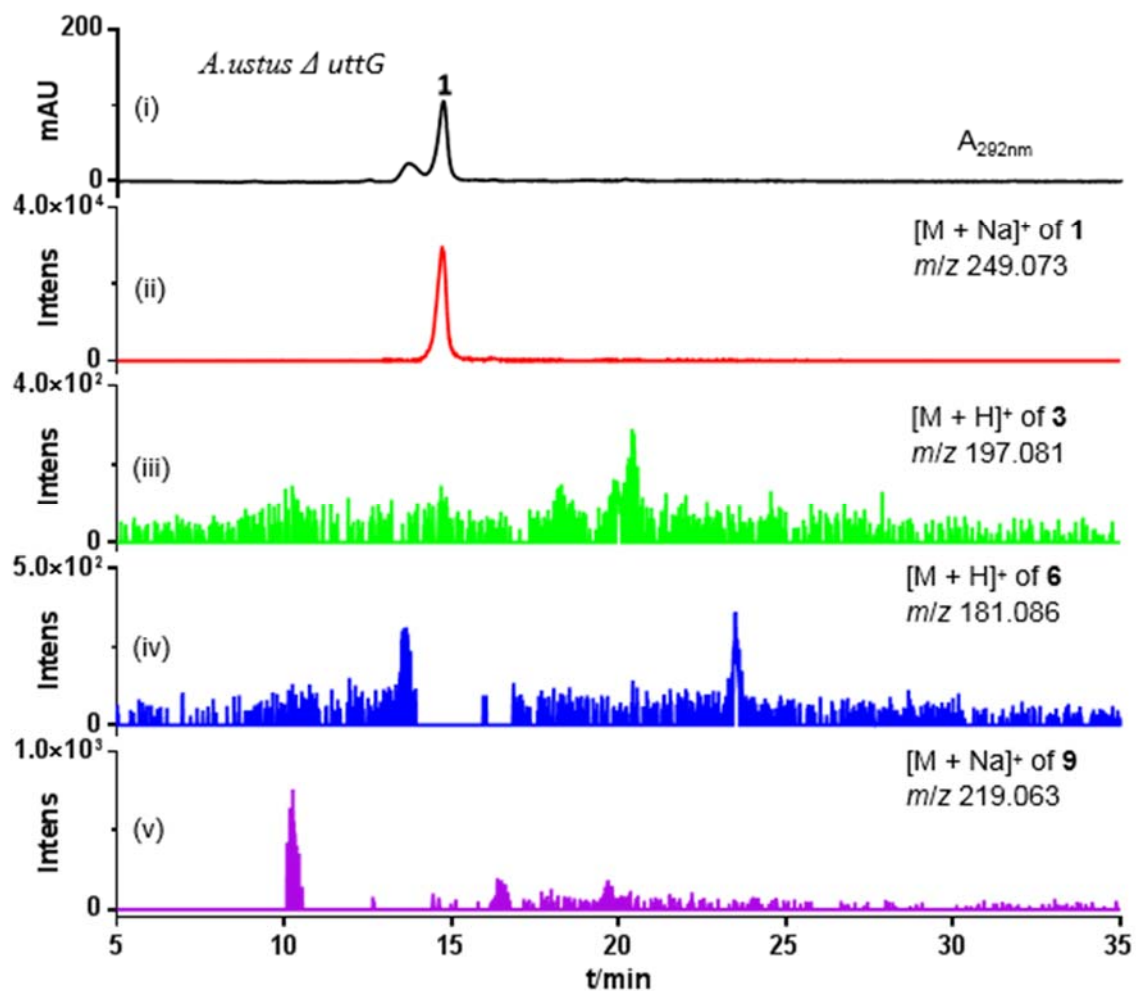


Figure S14. LC-MS analysis of the metabolite profile of the *A. ustus* Δ uttG mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3**, **6**, and **9** (ii – v).

SUPPORTING INFORMATION

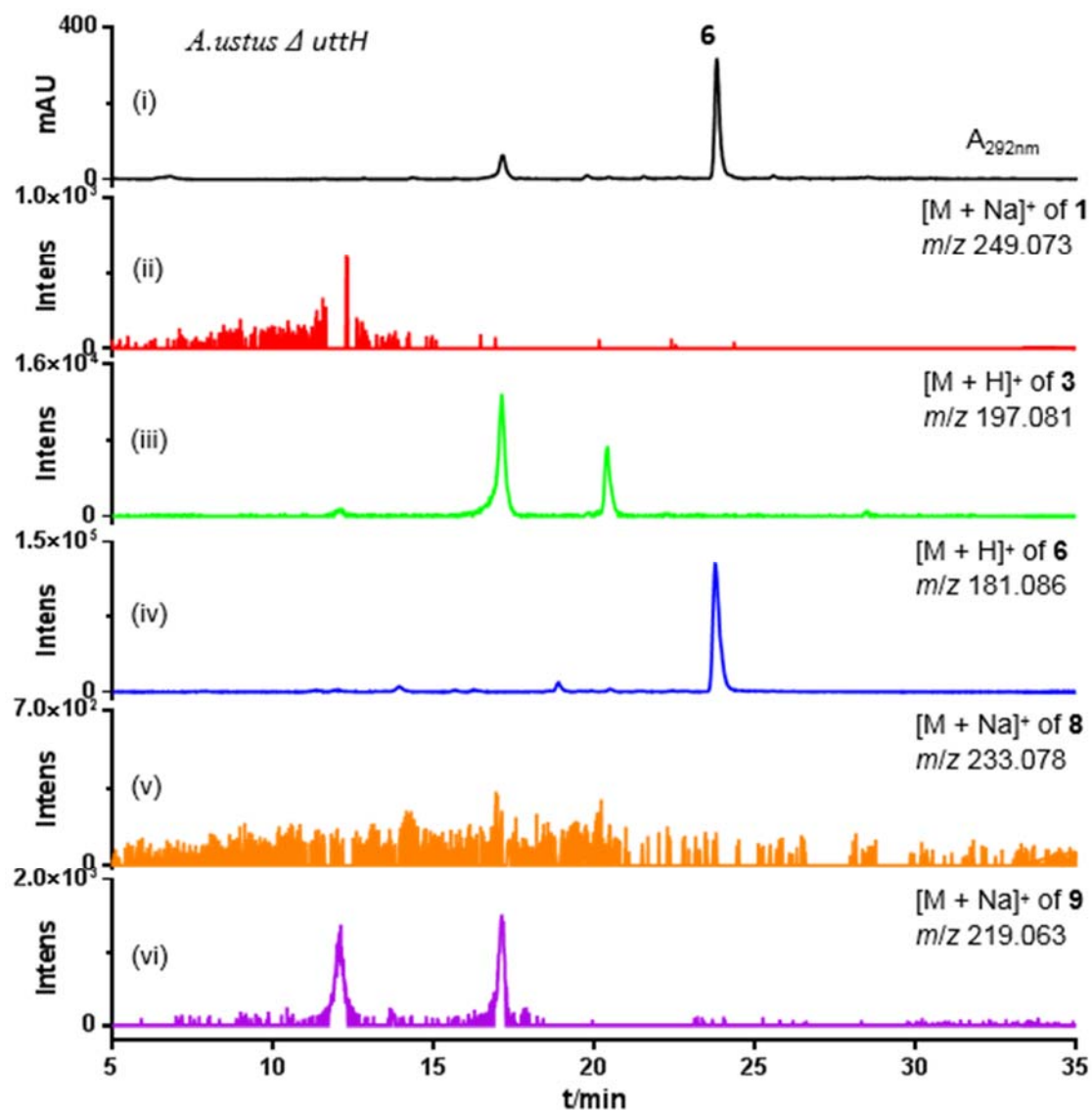


Figure S15. LC-MS analysis of the metabolite profile of the *A. ustus* Δ uttH mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3**, **6**, **8**, and **9** (ii – vi).

SUPPORTING INFORMATION

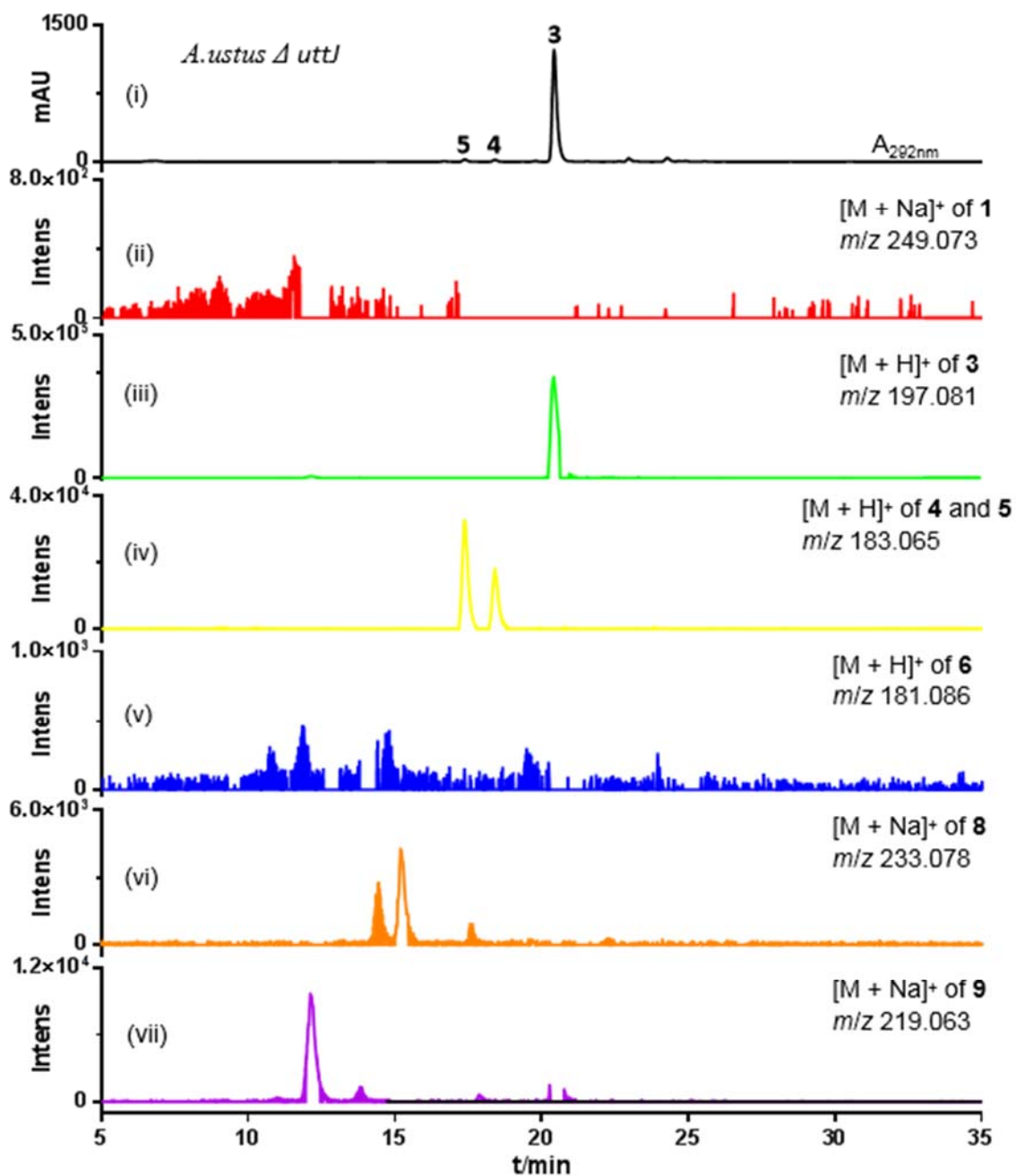


Figure S16. LC-MS analysis of the metabolite profile of the *A. ustus* Δ *uttJ* mutant

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated. (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3 – 6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

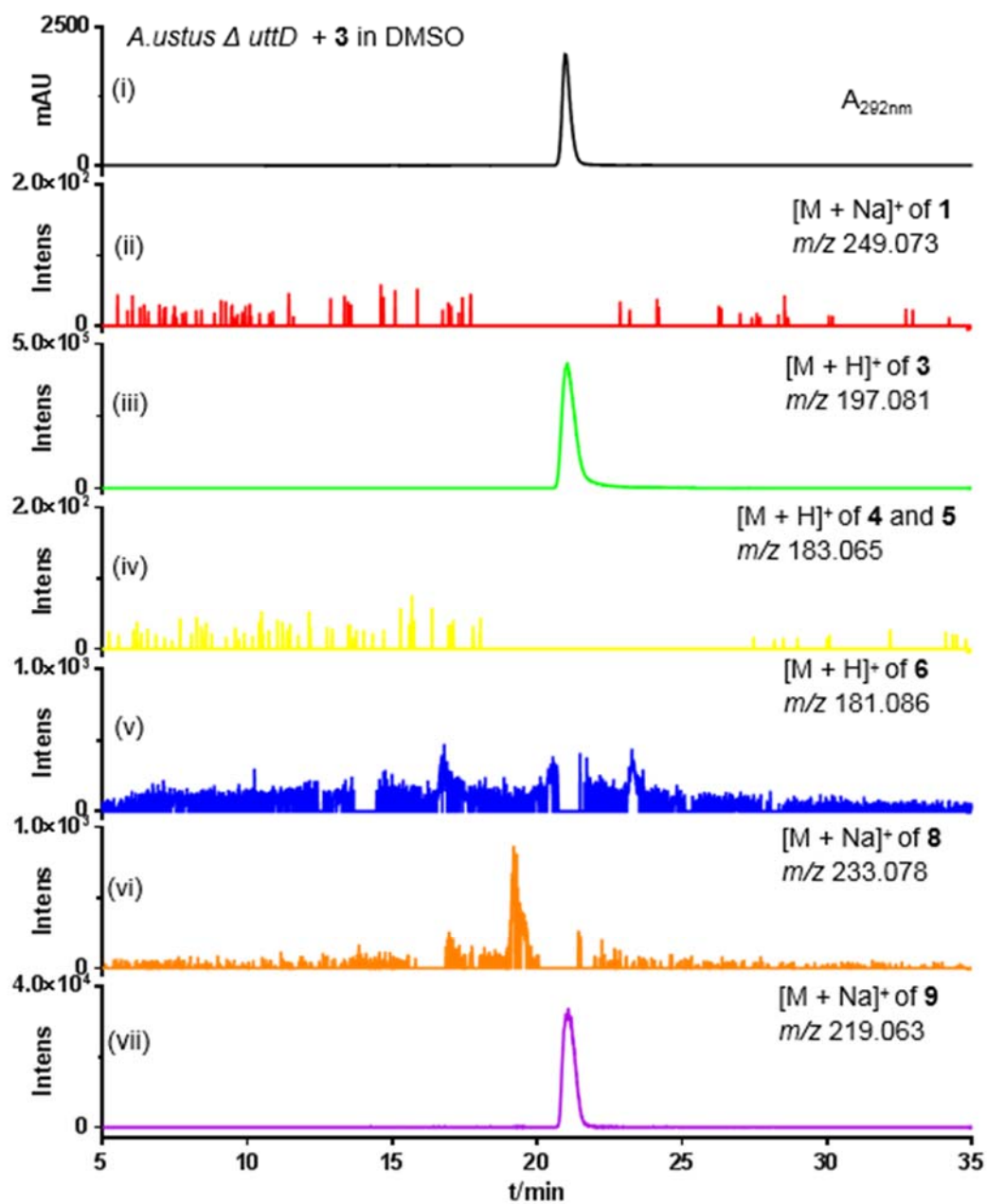


Figure S17. LC-MS analysis of the metabolite profile of the *A.ustus* Δ *uttD* after feeding with **3**.

UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ or $[M + Na]^+$ ions of **1**, **3** – **6**, **8** and **9** (ii – vii).

SUPPORTING INFORMATION

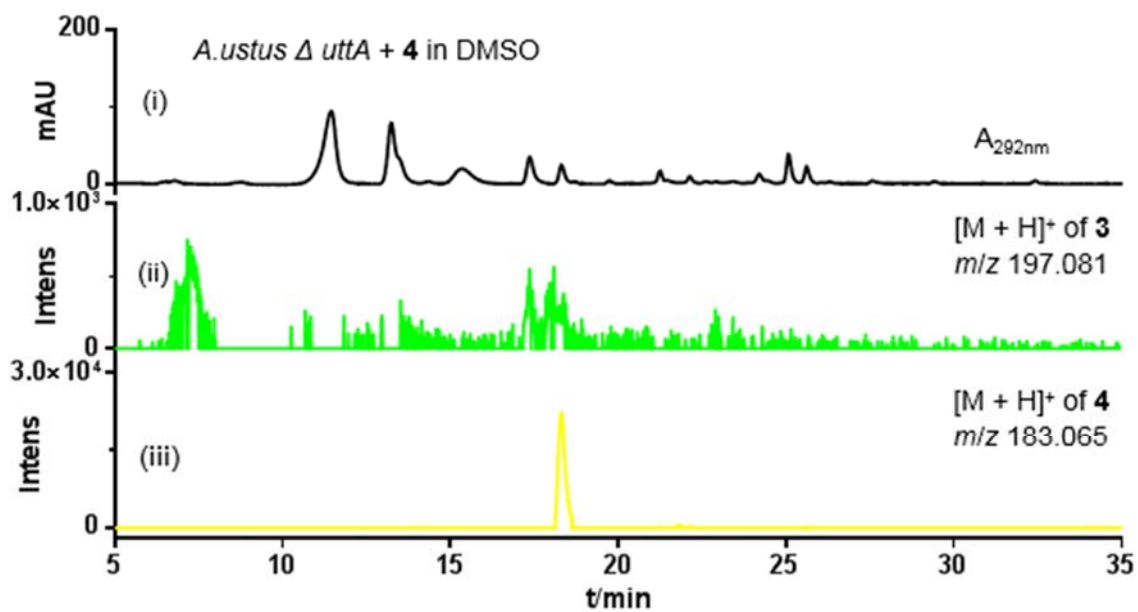


Figure S18. LC-MS analysis of the metabolite profile of the *A. ustus* Δ *uttA* after feeding with **4**. UV detection was carried out on a diode array detector and absorptions at 292 nm are illustrated (i). EICs with a tolerance range of ± 0.005 refer to $[M + H]^+$ of **3** and **4** (ii – iii).

SUPPORTING INFORMATION

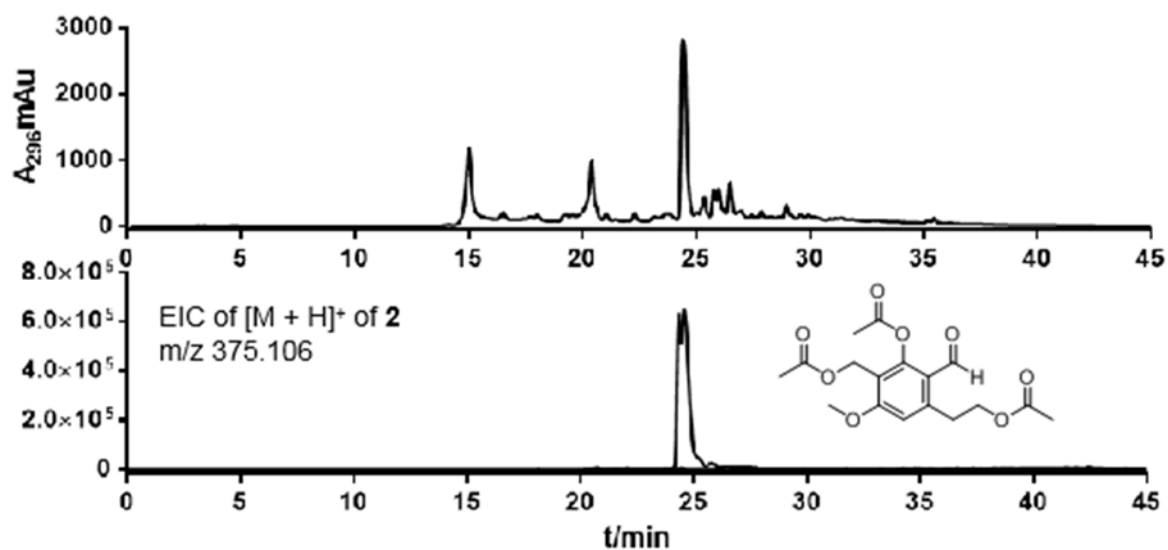


Figure S19. LC-MS analysis of the acetylated EtOAc extract from *A. ustus*.

EIC of **2** is selected with a tolerance range of ± 0.005 .

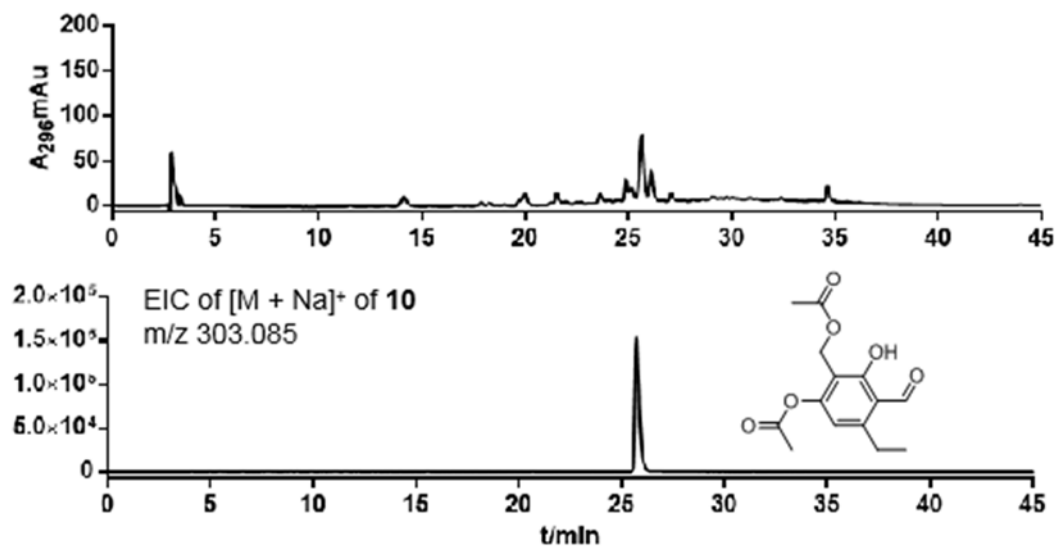


Figure S20. LC-MS analysis of the acetylated EtOAc extract from Δutf of *A. ustus*.

EIC of **10** is selected with a tolerance range of ± 0.005 .

SUPPORTING INFORMATION

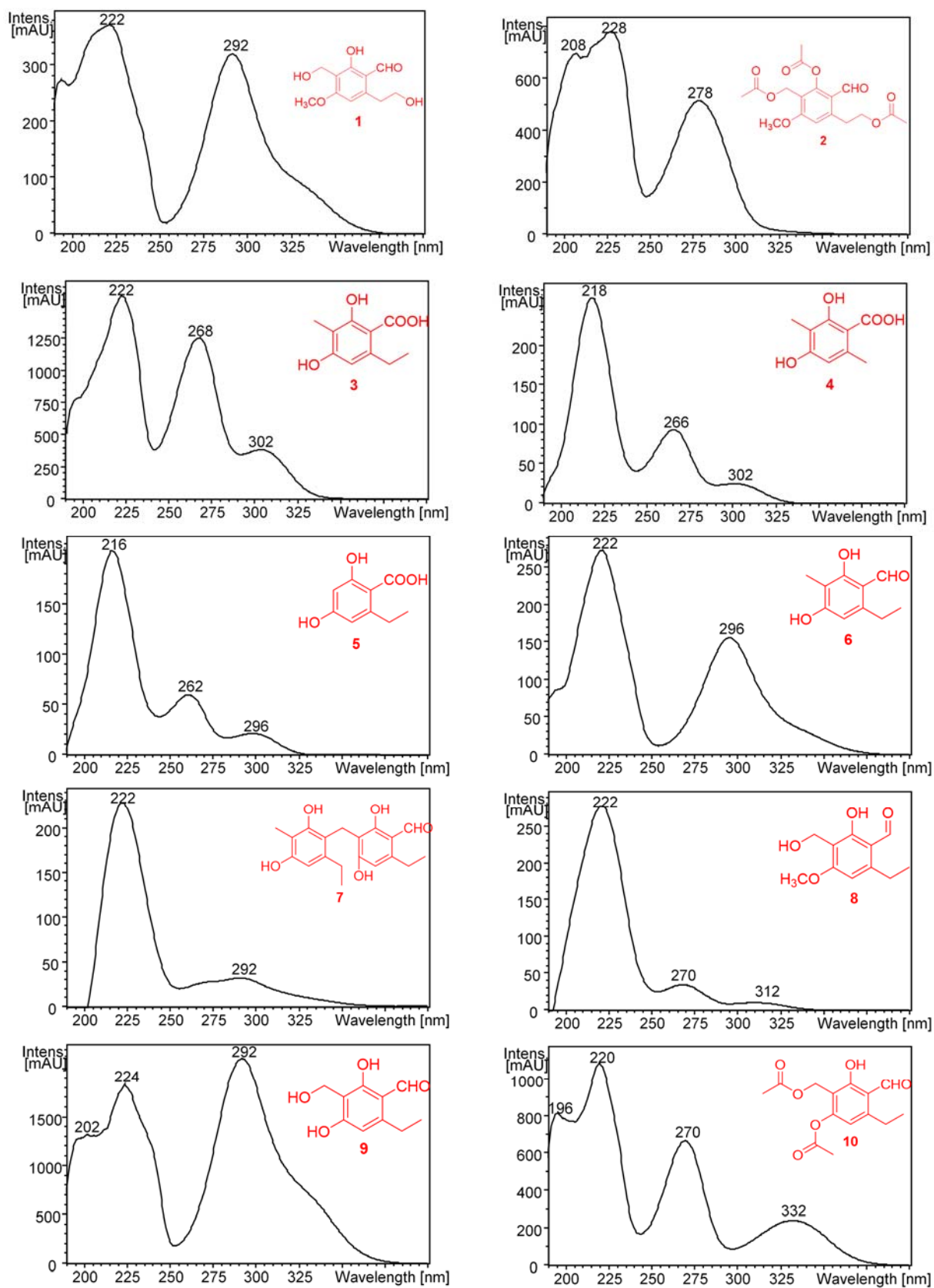


Figure S21. UV spectra of the compounds identified in this study

SUPPORTING INFORMATION

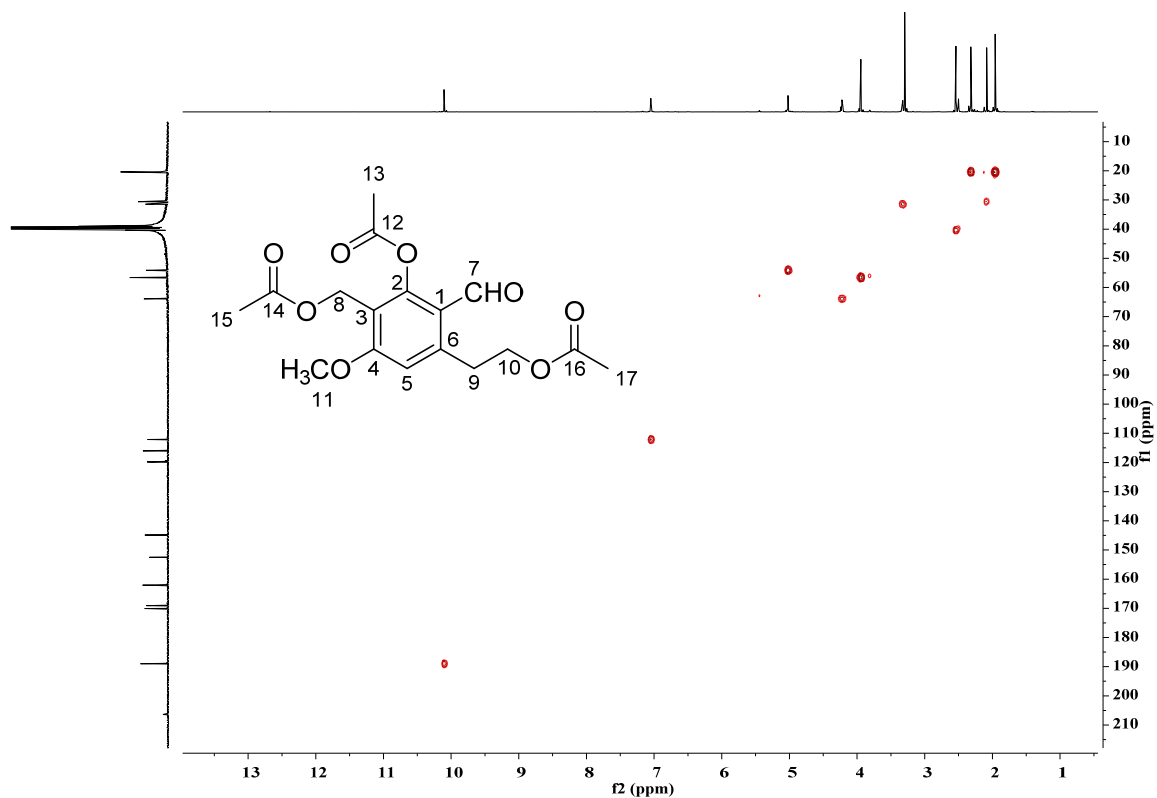


Figure S24. HSQC spectrum of compound **2** in DMSO- d_6

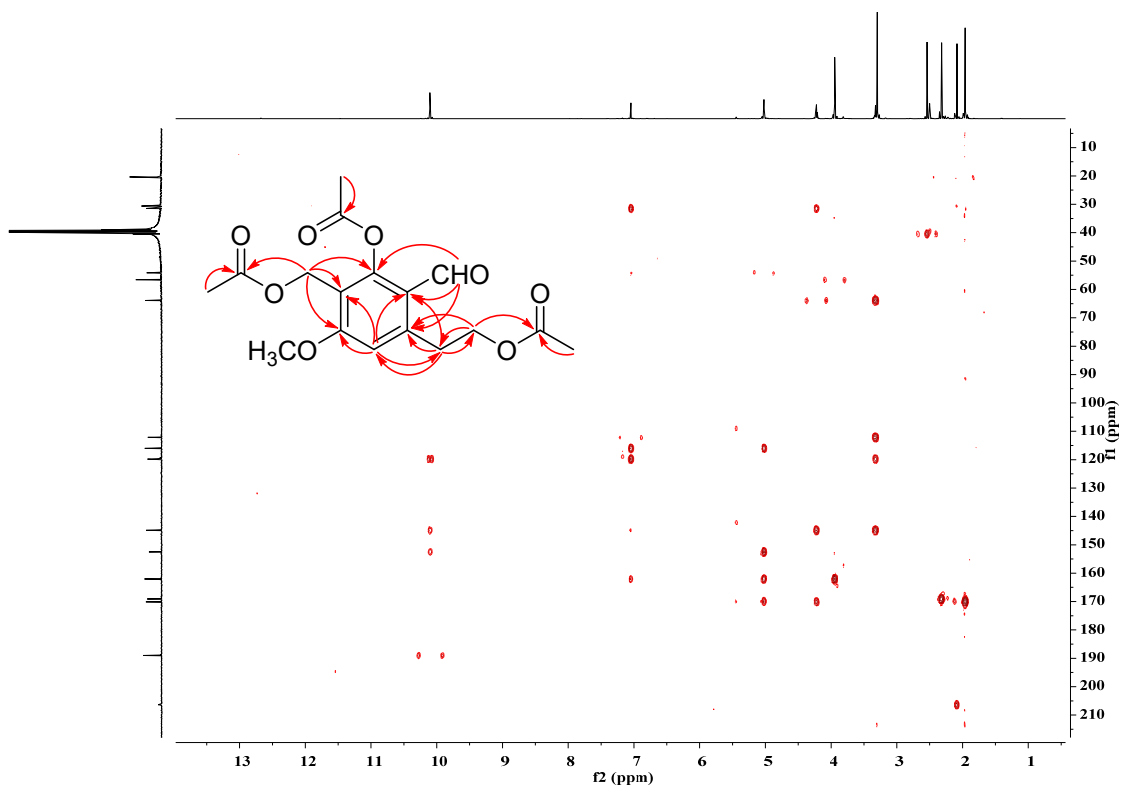
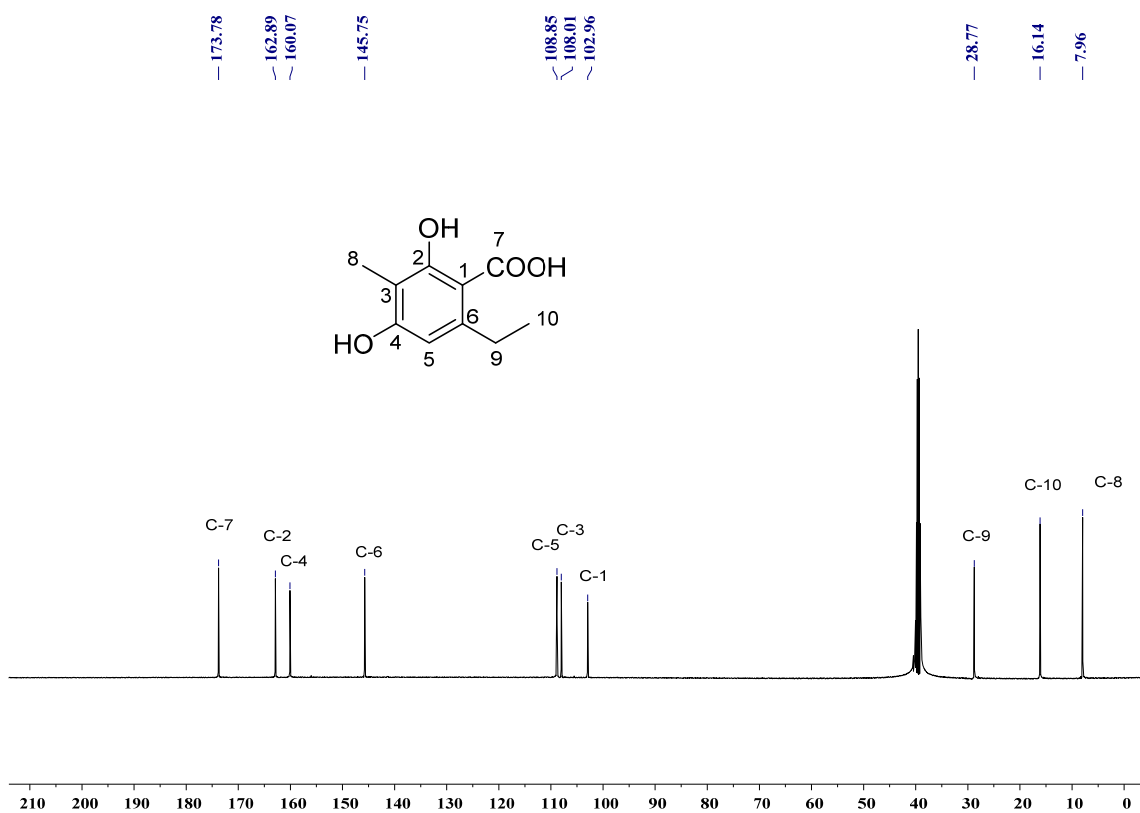
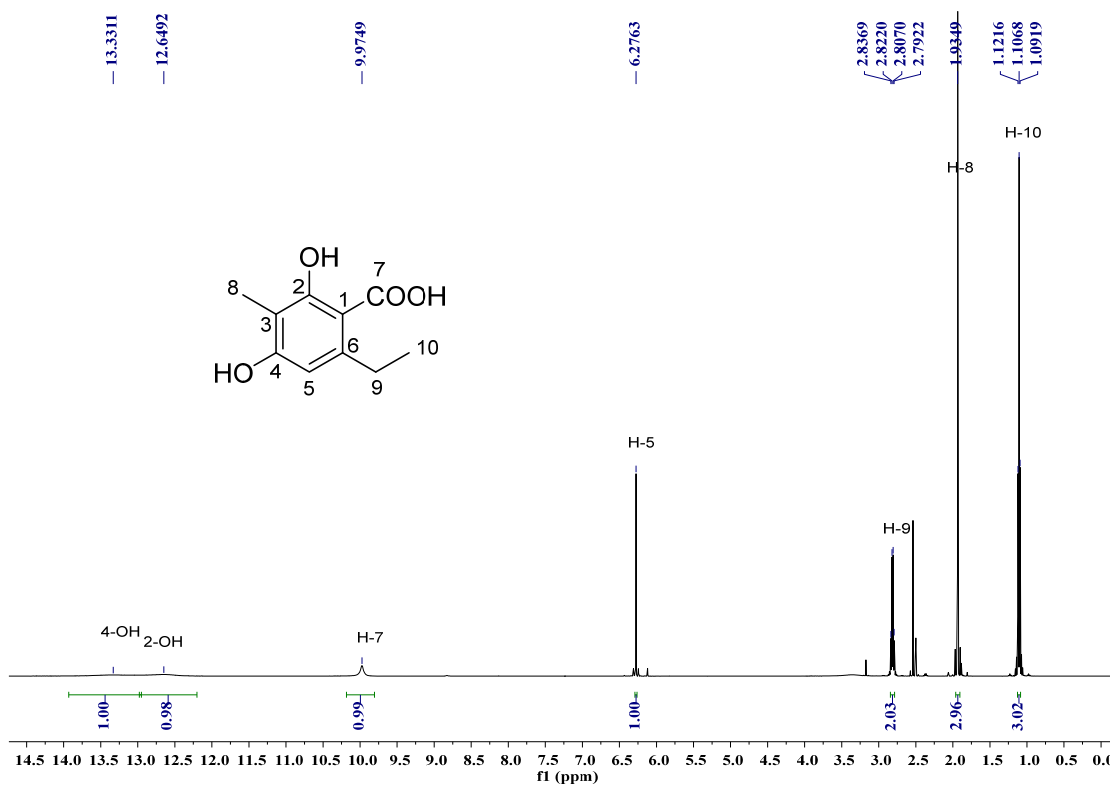


Figure S25. HMBC spectrum of compound **2** in DMSO- d_6

SUPPORTING INFORMATION



SUPPORTING INFORMATION

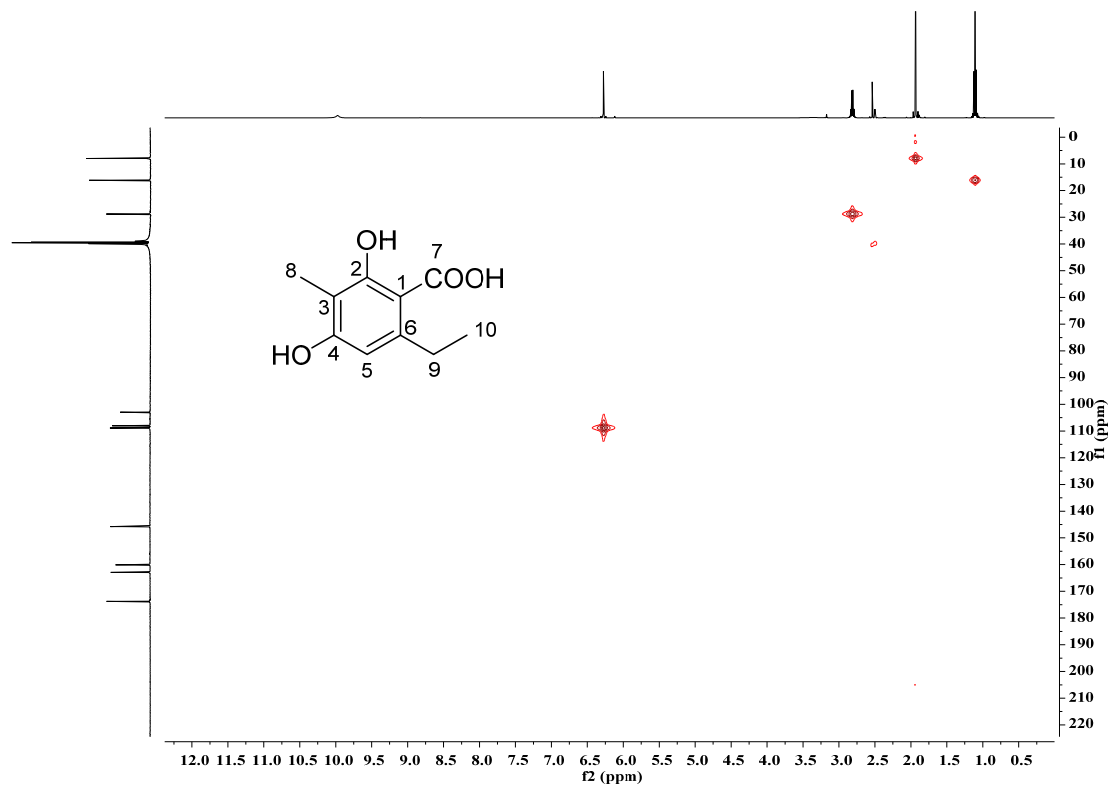


Figure S28. HSQC spectrum of compound **3** in DMSO- d_6

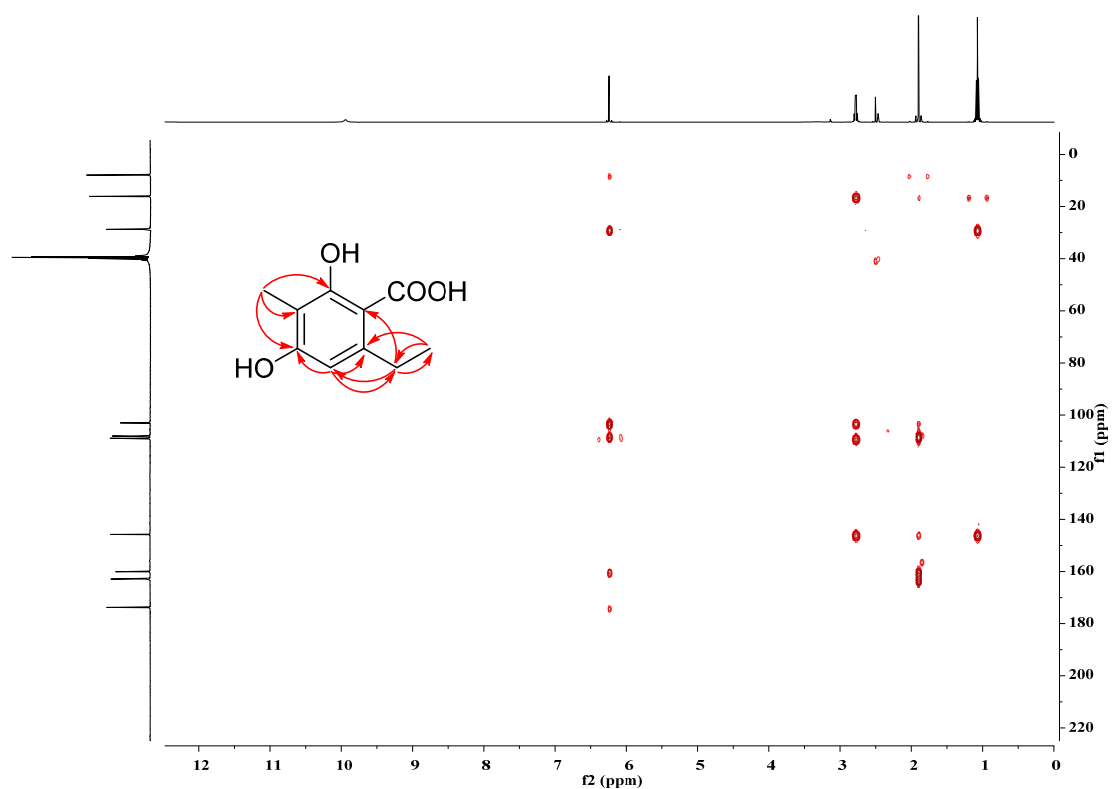


Figure S29. HMBC spectrum of compound **3** in DMSO- d_6

SUPPORTING INFORMATION

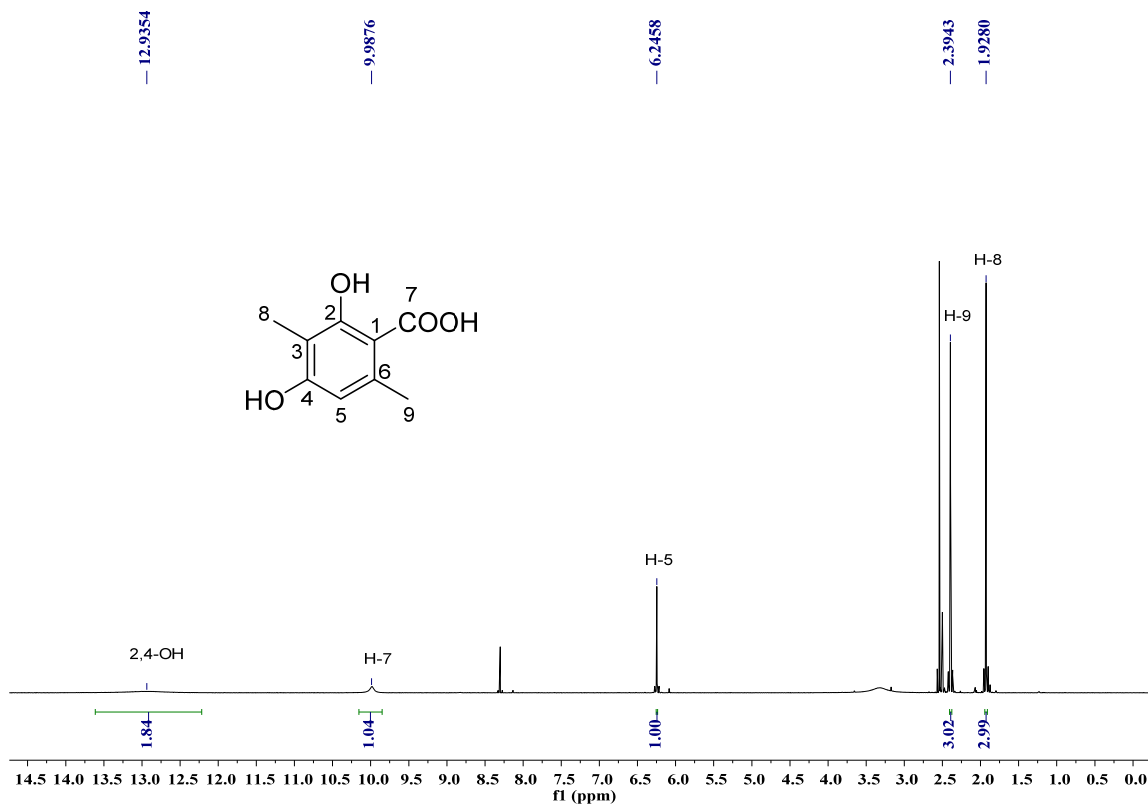


Figure S30. ^1H NMR spectrum of compound **4** in $\text{DMSO-}d_6$ (500 MHz)

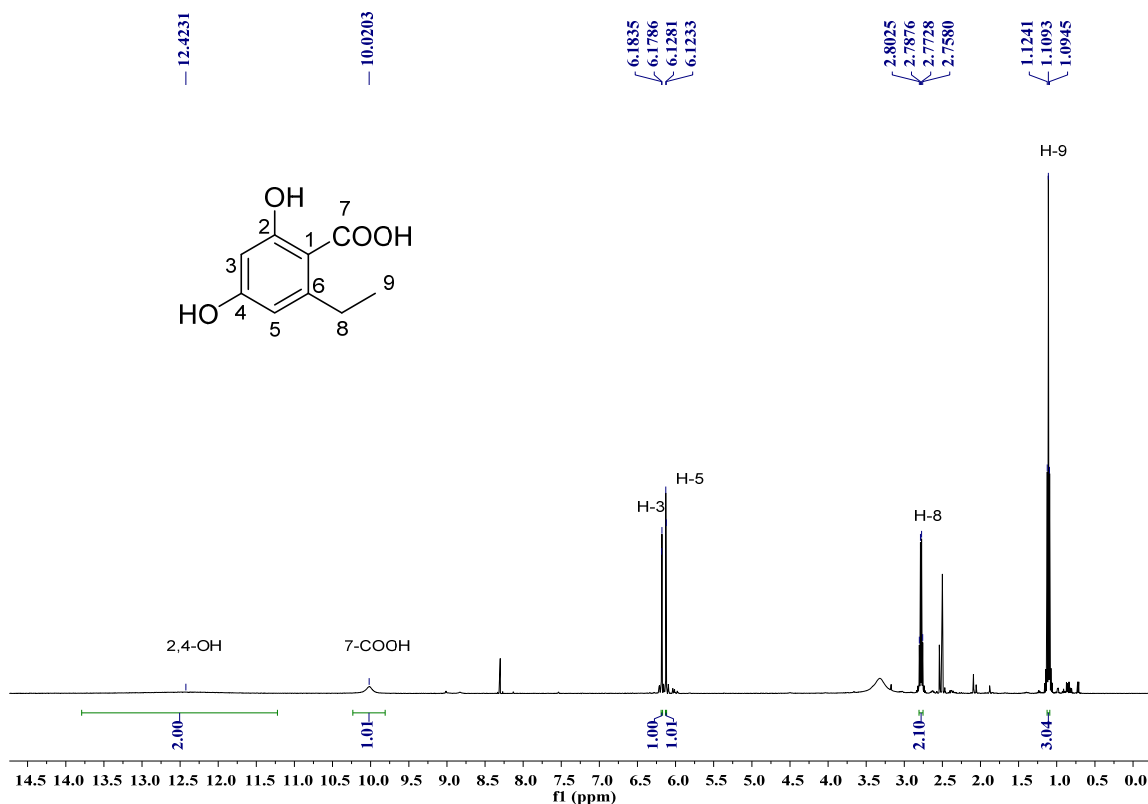


Figure S31. ^1H NMR spectrum of compound **5** in $\text{DMSO-}d_6$ (500 MHz)

SUPPORTING INFORMATION

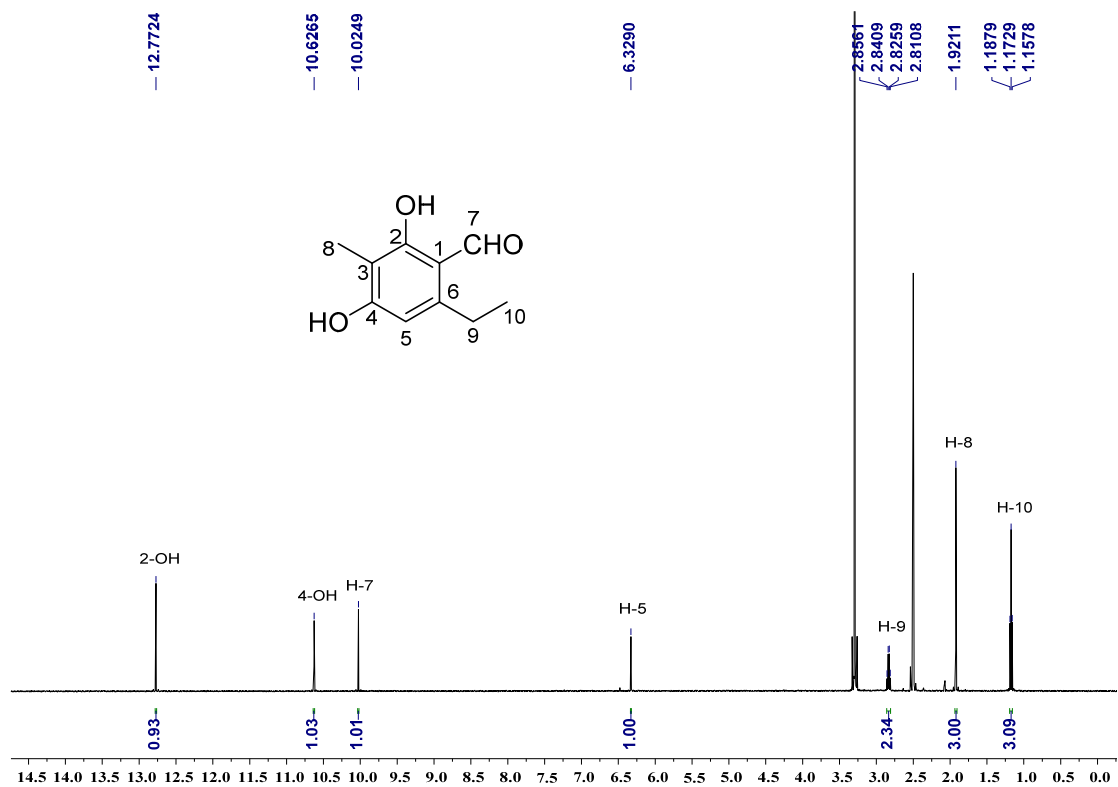


Figure S32. ^1H NMR spectrum of compound **6** in $\text{DMSO}-d_6$ (500 MHz)

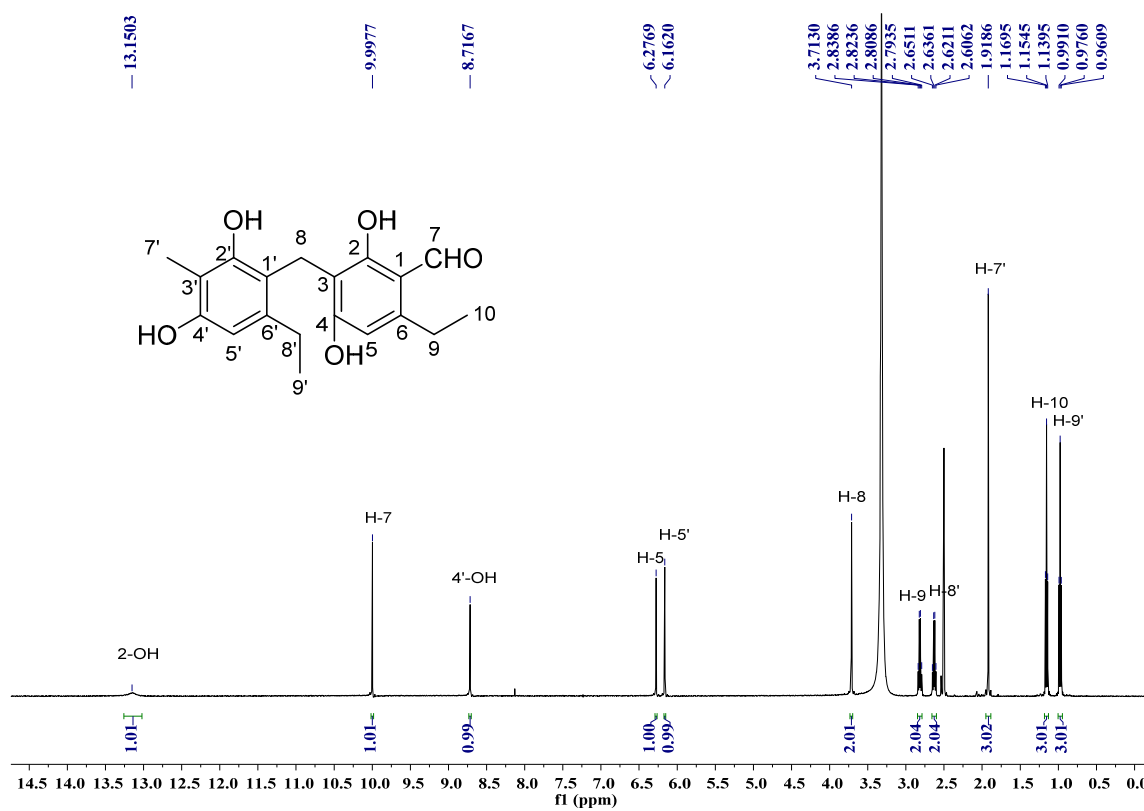


Figure S33. ^1H NMR spectrum of compound **7** in $\text{DMSO}-d_6$ (500 MHz)

SUPPORTING INFORMATION

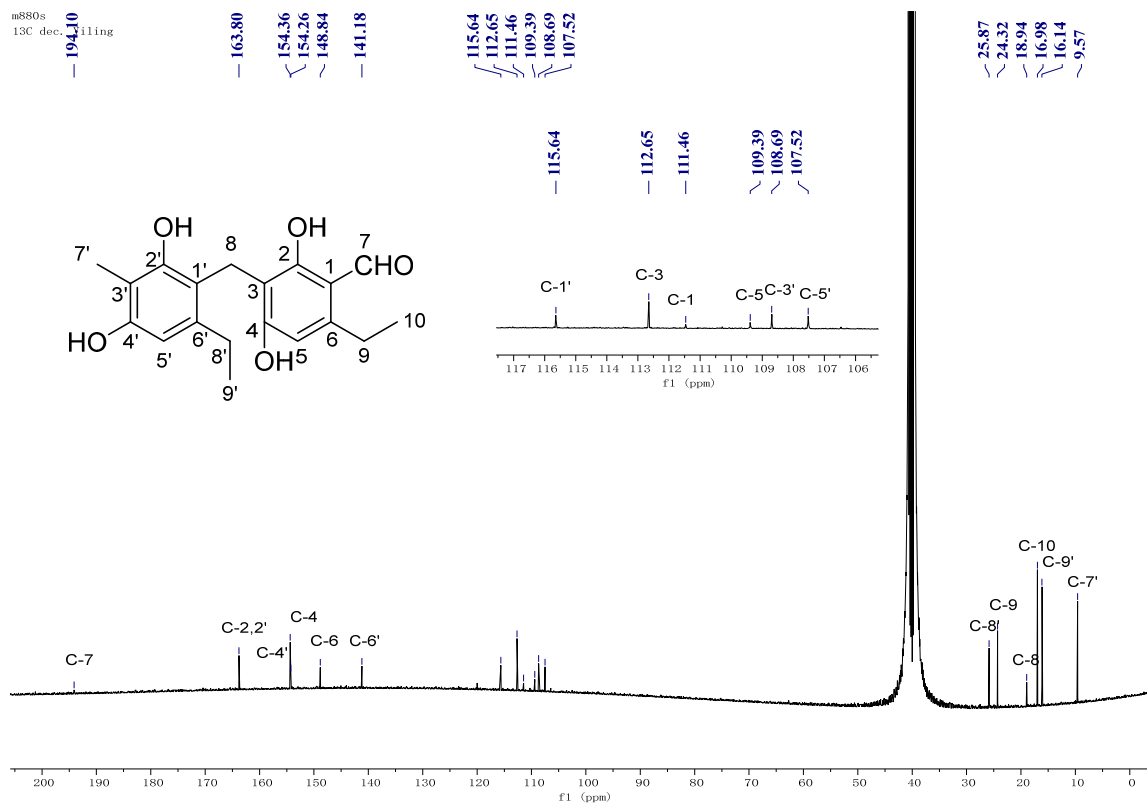


Figure S34. $^{13}\text{C}\{^1\text{H}\}$ NMR spectrum of compound **7** in $\text{DMSO-}d_6$ (125 MHz)

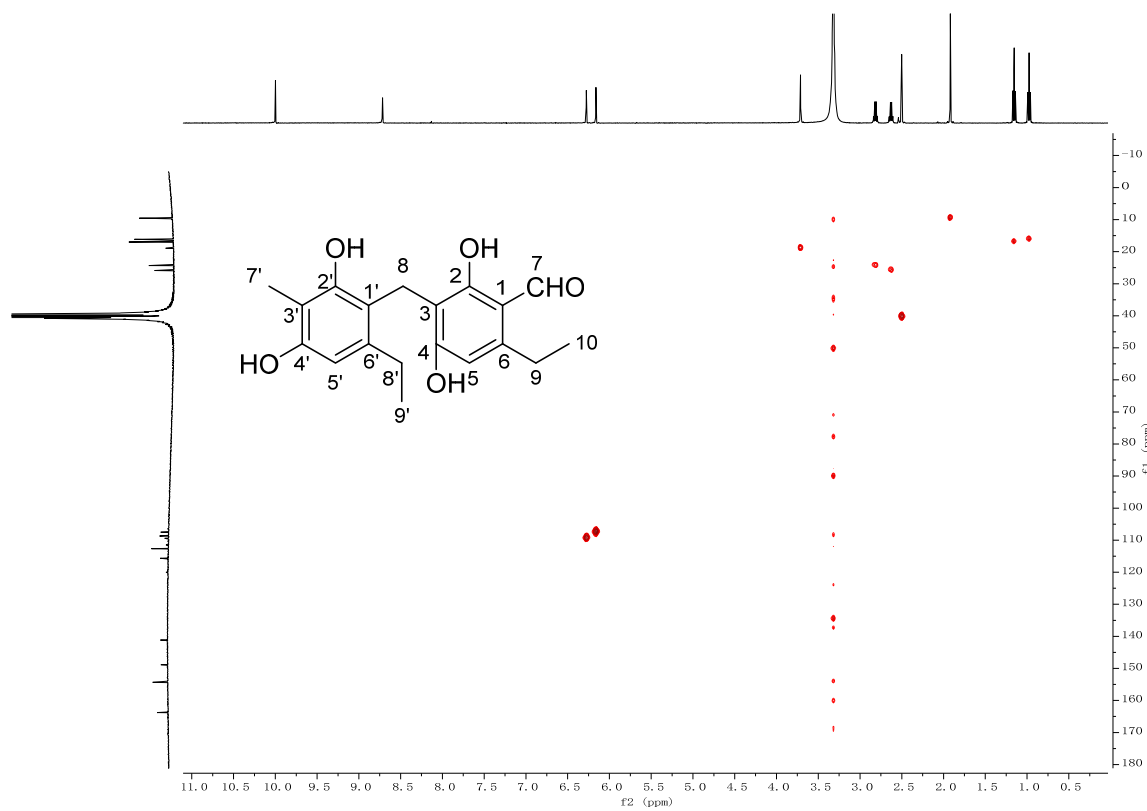


Figure S35. HSQC spectrum of compound **7** in $\text{DMSO-}d_6$

SUPPORTING INFORMATION

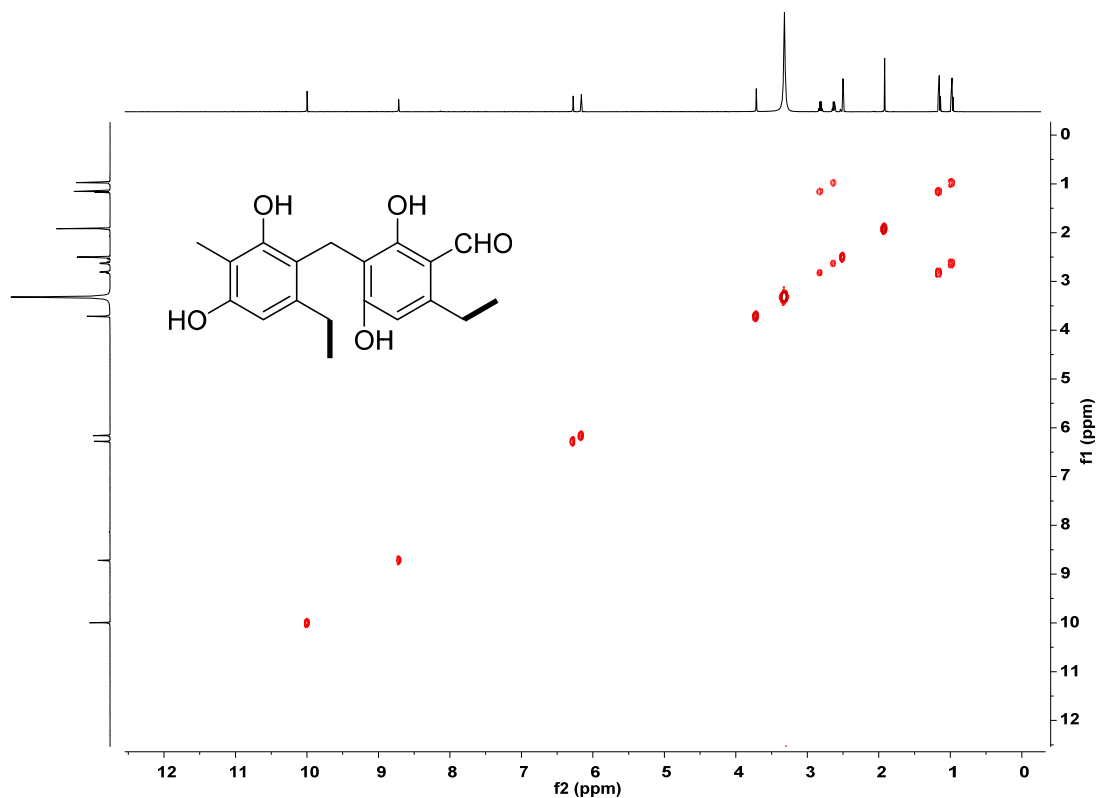


Figure S36. ¹H-¹H COSY spectrum of compound 7 in DMSO-*d*₆

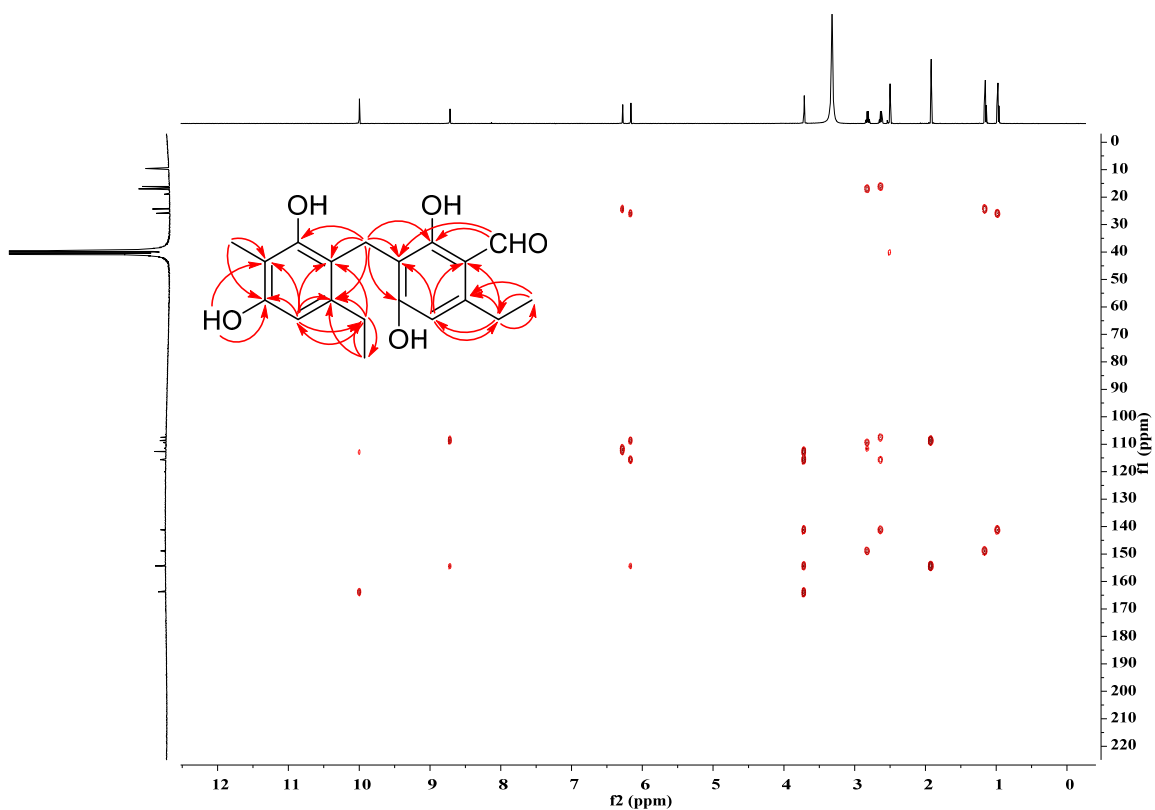
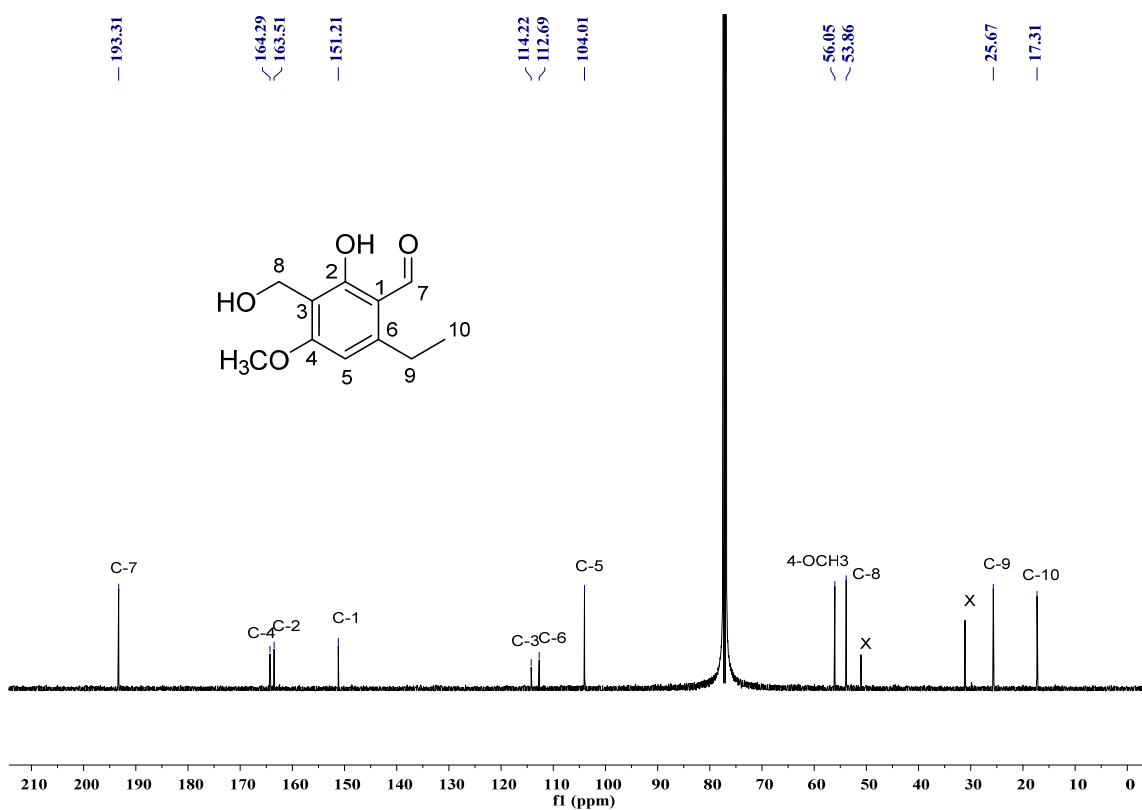
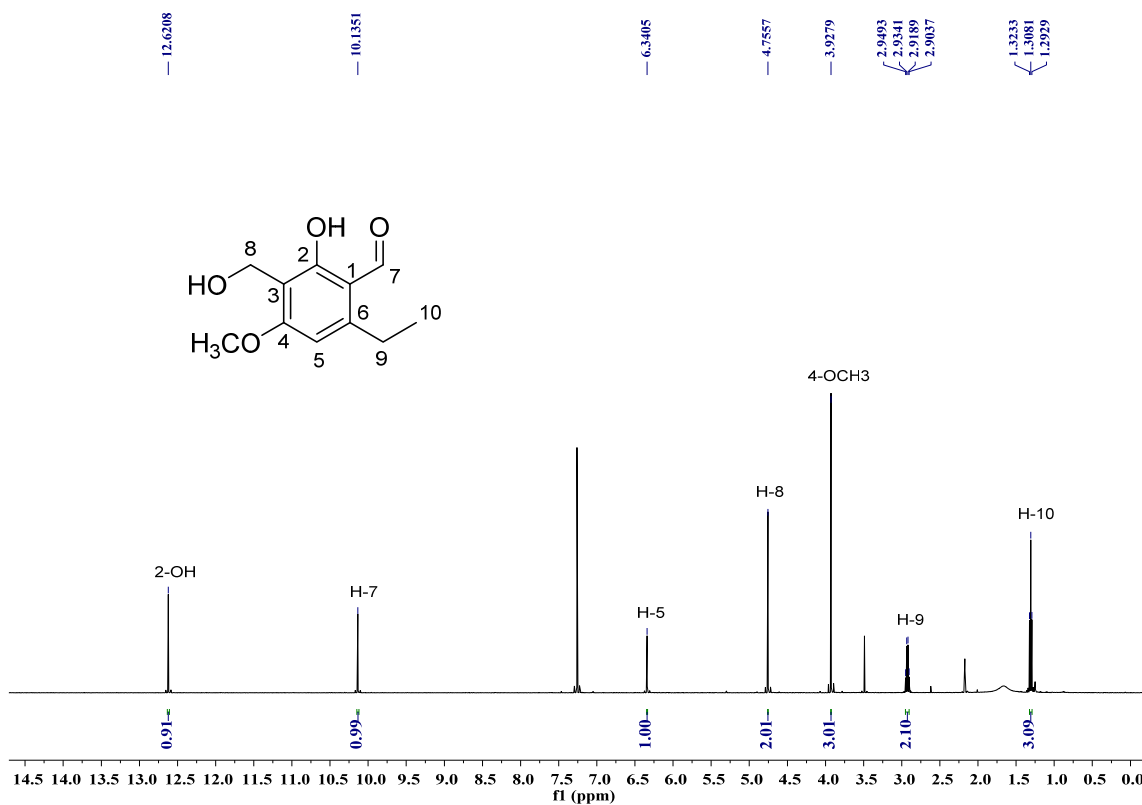


Figure S37. HMBC spectrum of compound 7 in DMSO-*d*₆

SUPPORTING INFORMATION



SUPPORTING INFORMATION

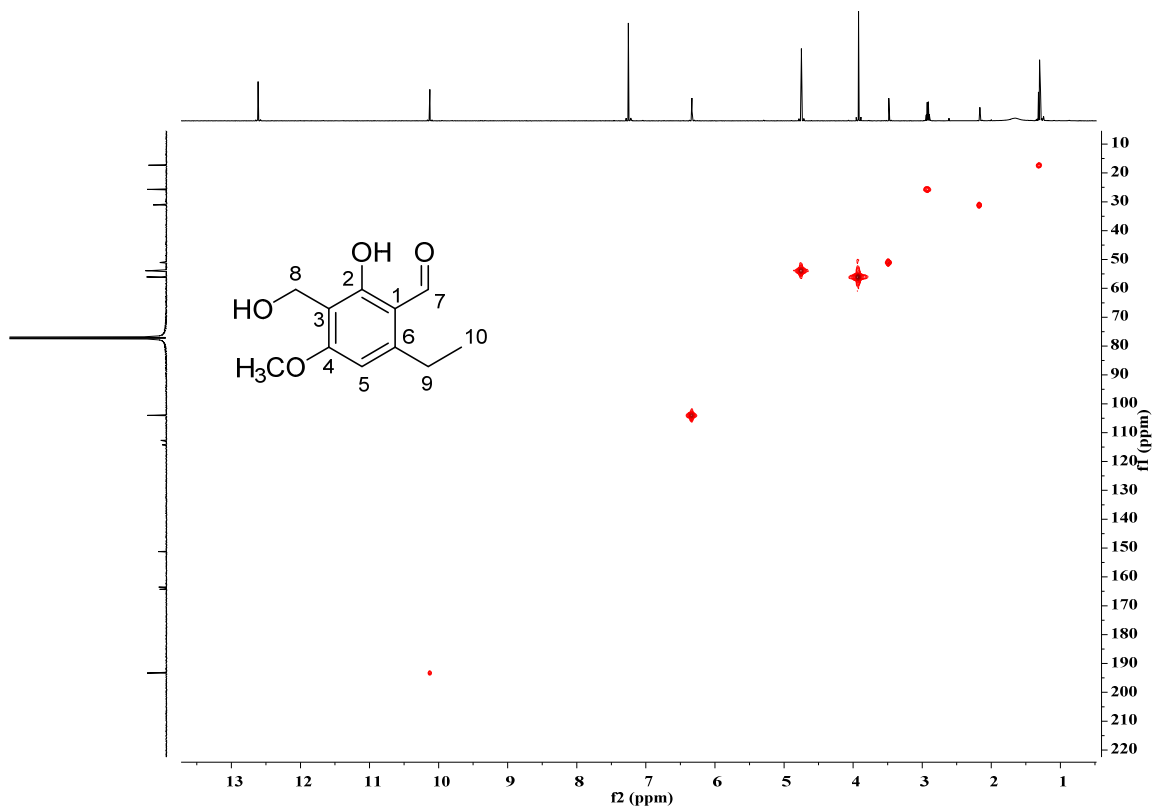


Figure S40. HSQC spectrum of compound **8** in CDCl_3

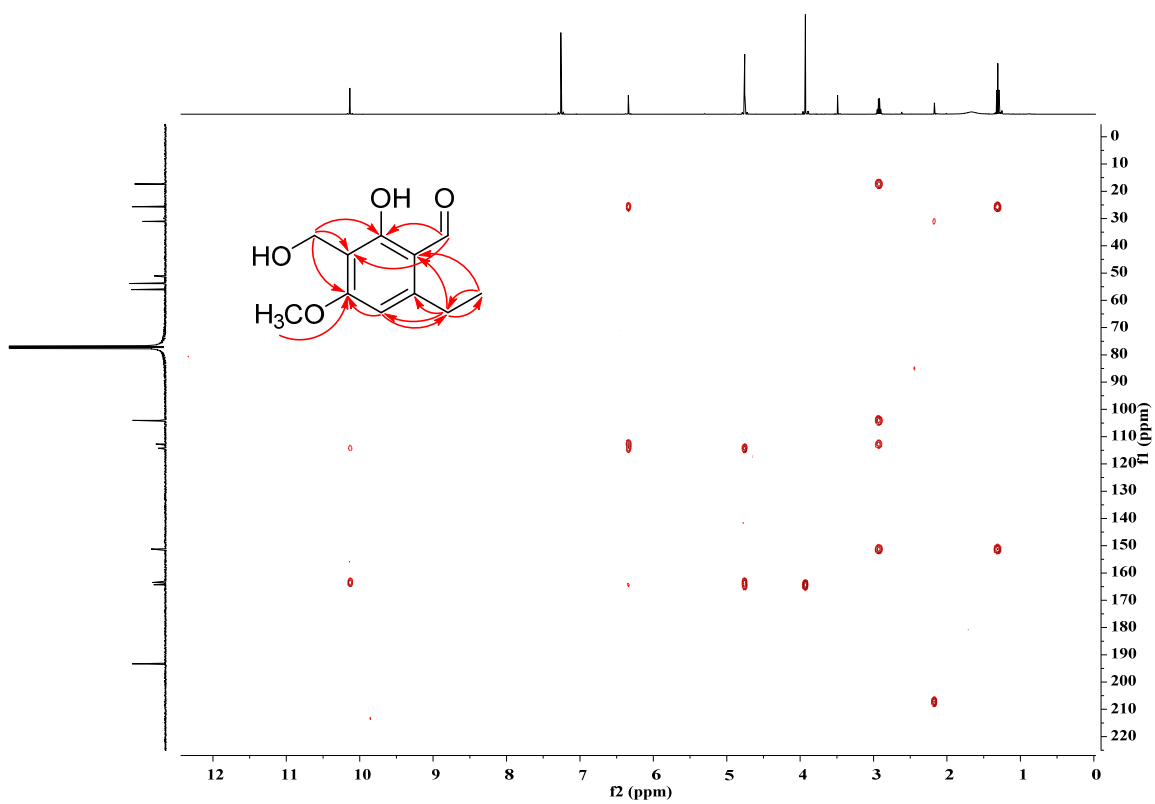
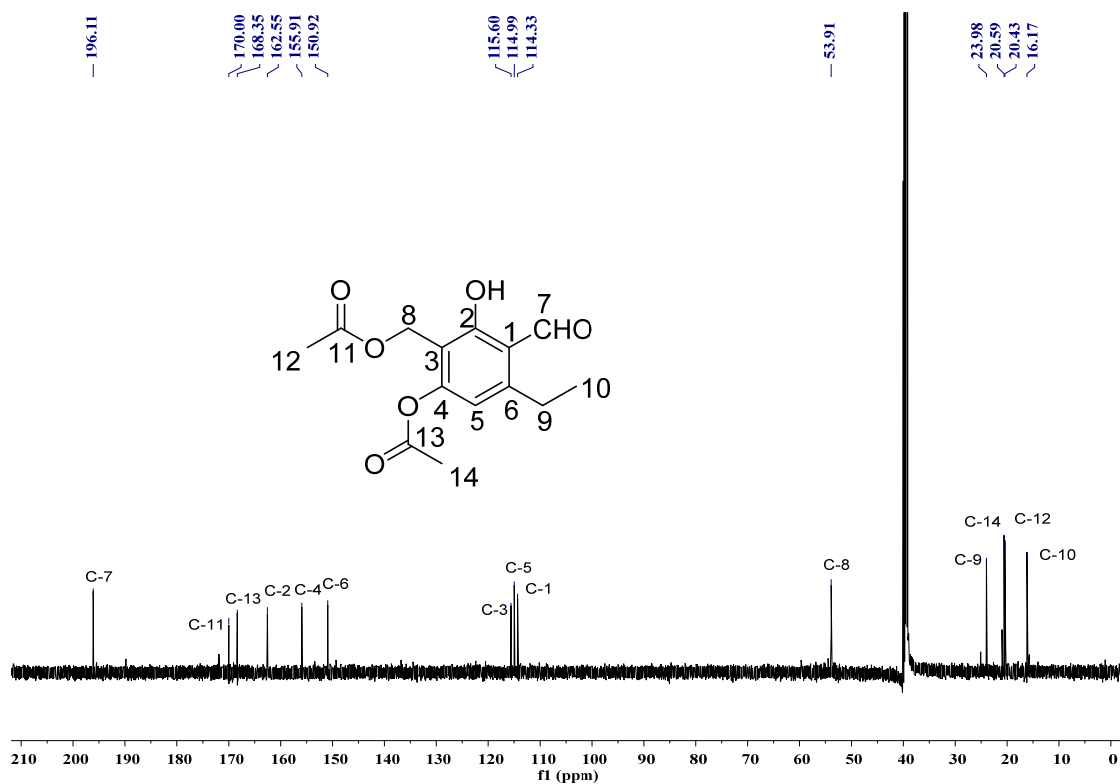
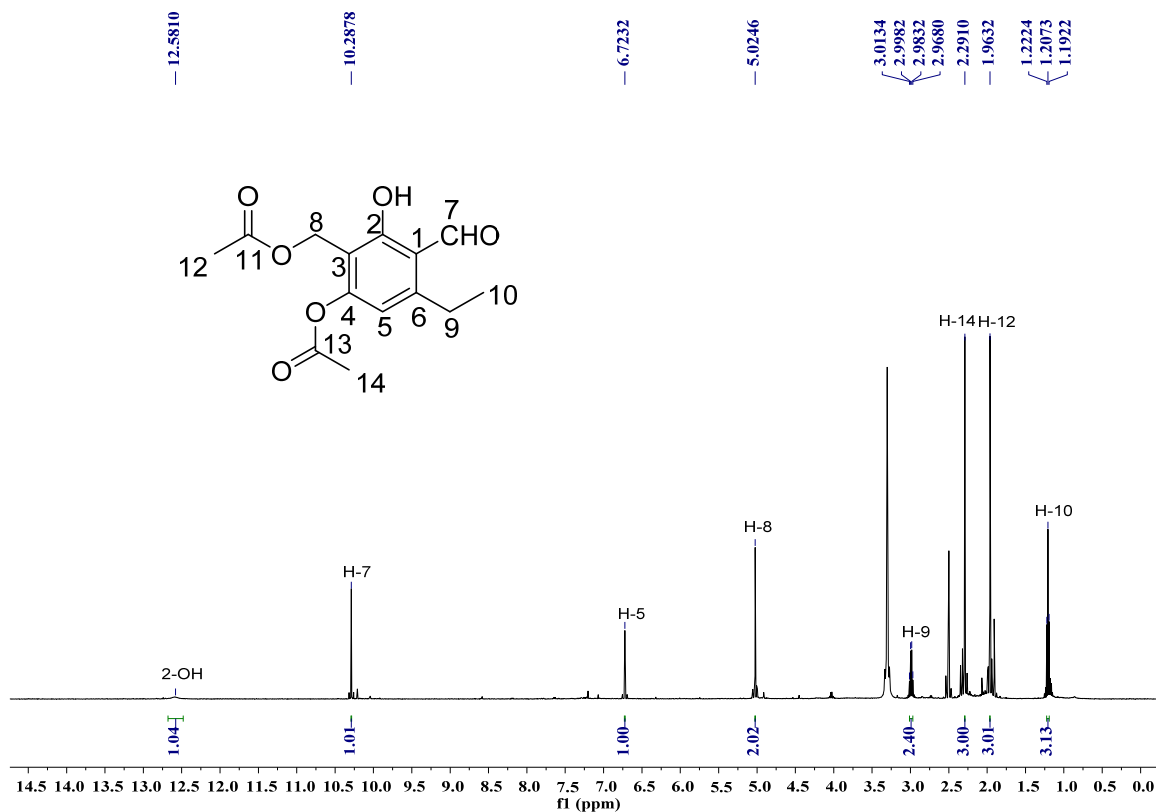


Figure S41. HMBC spectrum of compound **8** in CDCl_3

SUPPORTING INFORMATION



SUPPORTING INFORMATION

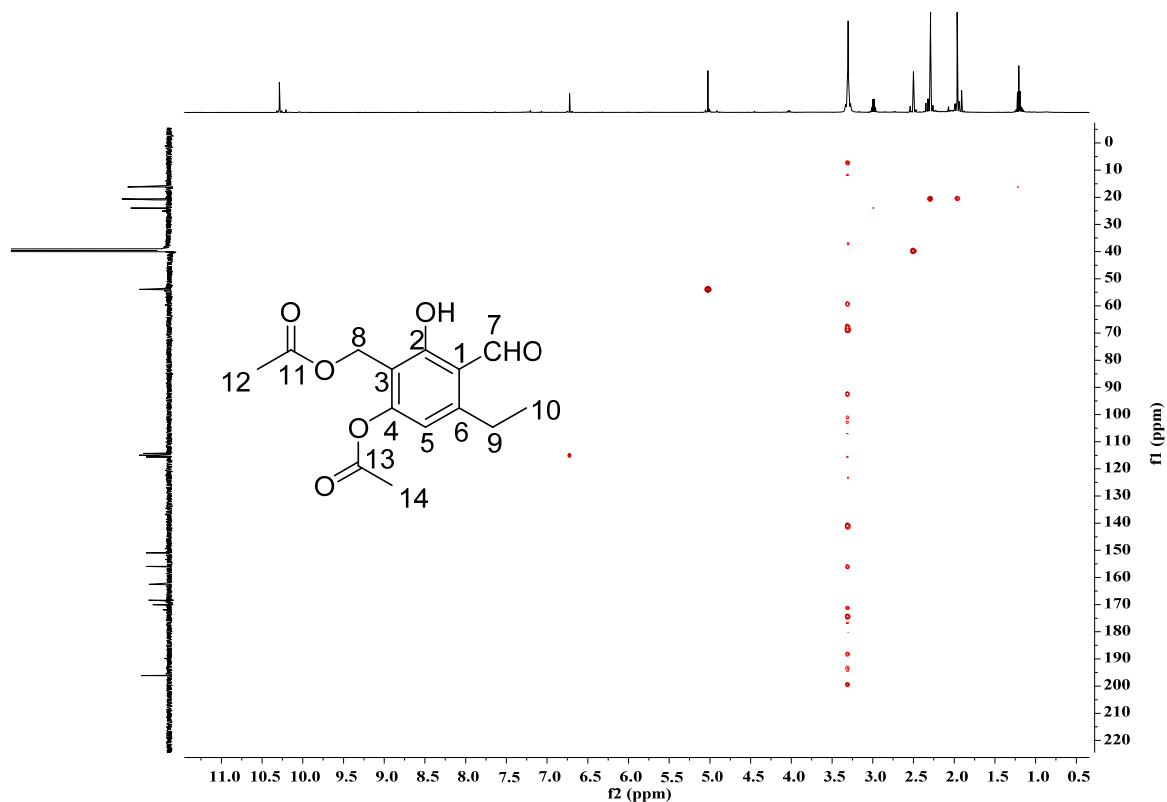


Figure S44. HSQC spectrum of compound **10** in DMSO- d_6

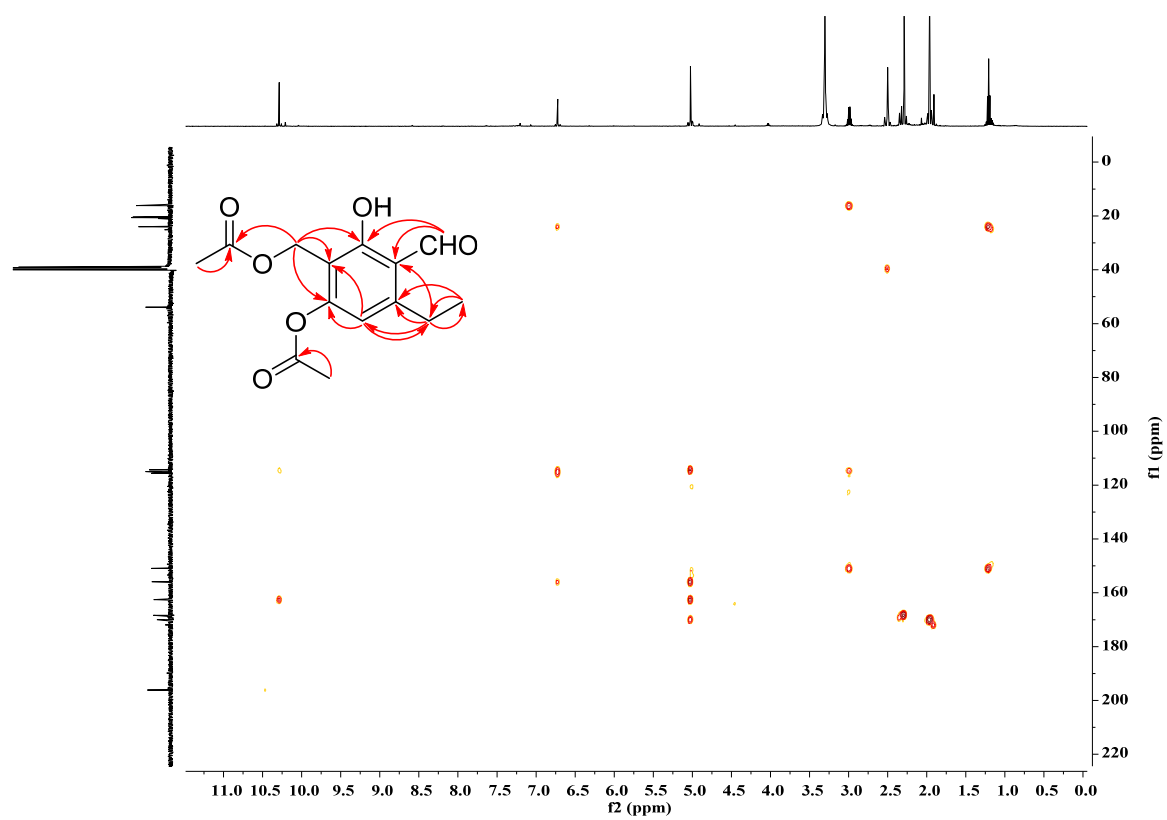


Figure S45. HMBC spectrum of compound **10** in DMSO- d_6

References

- (1) Weber, T.; Charusanti, P.; Musiol-Kroll, E. M.; Jiang, X.; Tong, Y.; Kim, H. U.; Lee, S. Y. Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol.* **2015**, *33*, 15.
- (2) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296.
- (3) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 486.
- (4) Chiang, Y. M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C.; Oakley, B. R. Development of genetic dereplication strains in *Aspergillus nidulans* results in the discovery of aspercryptin. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 1662.
- (5) Fan, J.; Liao, G.; Kindinger, F.; Ludwig-Radtke, L.; Yin, W.-B.; Li, S.-M. Peniphenone and penilactone formation in *Penicillium crustosum* via 1,4-Michael additions of *ortho*-quinone methide from hydroxyclovatol to *g*-butyrolactones from crustosic acid. *J. Am. Chem. Soc.* **2019**, *141*, 4225.
- (6) Mojtahedi, M. M. and Samadian, S. Efficient and rapid solvent-free acetylation of alcohols, phenols, and thiols using catalytic amounts of sodium acetate trihydrate. *J. Chem.* **2013**, *2013*, Article ID 642479.
- (7) Durrani, A. A. and Tyman, J. H. P. Long-chain phenols. Part 16. A novel synthesis of homologous orsellinic acids and their methyl ethers. *J. Chem. Soc., Perkin Trans. 1* **1980**, 1658.
- (8) Culberson, C. F. and Culberson, W. L. β -Orcinol derivatives in lichens: Biogenetic evidence from *Oropogon loxensis*. *Experim. Mycol.* **1978**, *2*, 245.
- (9) Shao, C.; Wang, C.; Wei, M.; Jia, Z.; She, Z.; Lin, Y. Two new benzaldehyde derivatives from mangrove endophytic fungus (No. ZZF 32). *Chem. Nat. Compd.* **2009**, *45*, 779.
- (10) He, Y. and Cox, R. J. The molecular steps of citrinin biosynthesis in fungi. *Chem. Sci.* **2016**, *7*, 2119.
- (11) Hagel, J. M. and Facchini, P. J. Expanding the roles for 2-oxoglutarate-dependent oxygenases in plant metabolism. *Nat. Prod. Rep.* **2018**, *35*, 721.

4.2 Elucidation of the streptoazine biosynthetic pathway in *Streptomyces aurantiacus* reveals the presence of a promiscuous prenyltransferase/cyclase

Elucidation of the Streptoazine Biosynthetic Pathway in *Streptomyces aurantiacus* Reveals the Presence of a Promiscuous Prenyltransferase/Cyclase

Jing Liu,[§] Yiling Yang,[§] Lauritz Harken, and Shu-Ming Li*



Cite This: *J. Nat. Prod.* 2021, 84, 3100–3109



Read Online

ACCESS |



Metrics & More

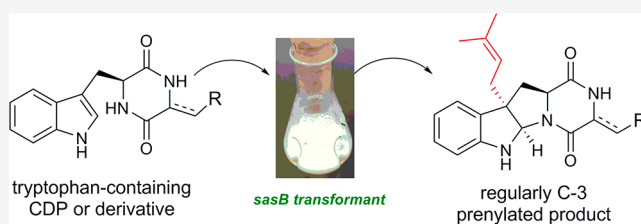


Article Recommendations



Supporting Information

ABSTRACT: Heterologous expression of a three-gene cluster from *Streptomyces aurantiacus* coding for a cyclodipeptide synthase, a prenyltransferase, and a methyltransferase led to the elucidation of the biosynthetic steps of streptoazine C (**2**). *In vivo* biotransformation experiments proved the high flexibility of the prenyltransferase SasB toward tryptophan-containing cyclodipeptides for regular C-3-prenylation. Furthermore, their corresponding dehydrogenated derivatives prepared by using cyclodipeptide oxidases were also used for prenylation. This study provides an enzyme with high substrate promiscuity from a less explored group of prenyltransferases for potential use to generate prenylated derivatives.



2,5-Diketopiperazine (DKP) alkaloids with an indole or indoline ring and isoprenoid moieties are derived from tryptophan-containing cyclodipeptides (CDPs).^{1–3} They represent an important class of hybrid natural products and display diverse biological and pharmacological activities, including antibacterial, antitumor, anti-inflammatory, and insecticidal effects.^{2–4} Representatives of tryptophan-containing CDP derivatives with various amino acids and one or more dimethylallyl (C₅) moieties at different positions of the indole or indoline ring are shown in Chart 1. Okaramin C,^{5,6} fellutanine D,^{7,8} fructigenine A,⁹ fumitremorgin B,¹⁰ roquefortine E,¹¹ and echinulin¹² are examples of a large number of fungal products. In comparison, only a limited number of prenylated DKP derivatives, such as nocardiozine A,¹³ drimentine G,¹⁴ and streptoazine C (**2**),¹⁵ are bacterial metabolites.

Significant progress has been achieved recently regarding the understanding of the biosynthesis of prenylated CDPs and derivatives thereof, especially of those from fungi of the genera *Penicillium* and *Aspergillus*.^{16–20} In nature, the 2,5-DKP scaffolds are usually biosynthesized by two distinct enzyme types, either by the nonribosomal peptide synthetases (NRPSs)^{18,21} or by the cyclodipeptide synthases (CDPSs).^{22,23} NRPSs are modular multidomain enzyme complexes and incorporate free amino acids to form the final peptide products.²⁴ Bimodular NRPSs are responsible for the formation of CDPs.^{16,24} In contrast, CDPSs, mostly of bacterial origin, directly hijack aminoacyl-tRNAs from the protein biosynthesis as substrates to form the DKP scaffolds.²⁵ The DKP scaffolds can be further modified by diverse tailoring enzymes including prenyltransferases (PTs), methyltransferases (MTs), and cytochrome P450 enzymes.^{16,17,25,26}

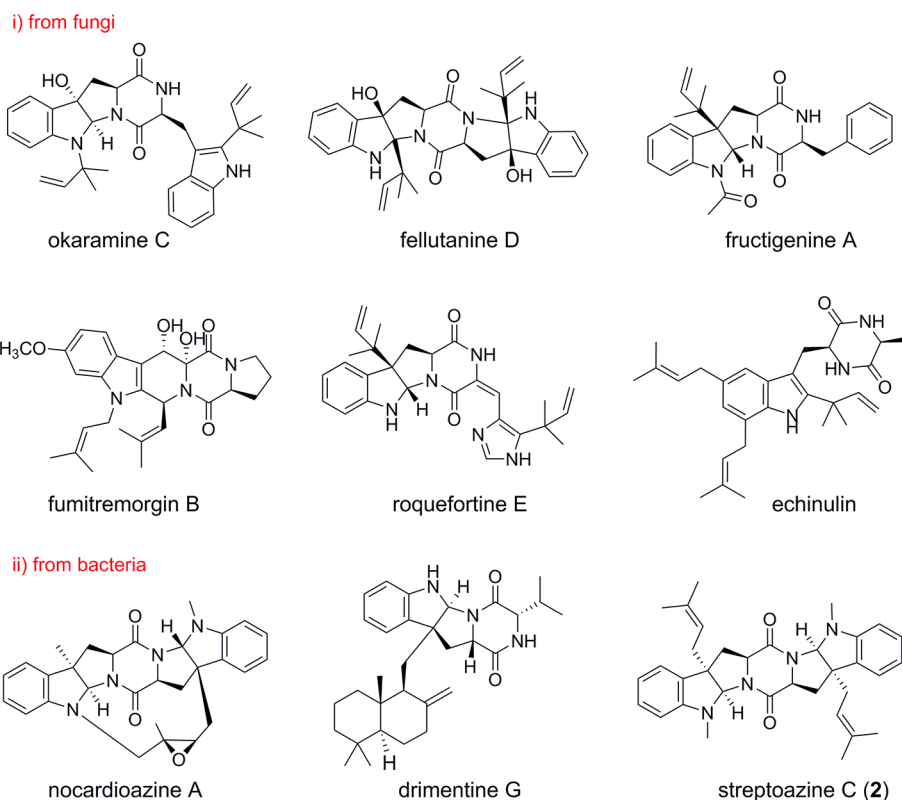
Prenylation by PTs at different positions of the indole ring of tryptophan-containing CDPs plays a key role for structural diversification of indole alkaloids and is involved in the biosynthesis of a large number of CDP derivatives.^{1,16} So far, most of the prenylated CDP alkaloids have been identified from NRPS-dependent pathways in fungi, and at least 13 fungal CDP PTs for regular (normal) or reverse prenylation were characterized biochemically.^{17,27} PTs from bacteria and fungi are usually highly permissive and can use structurally distinct compounds for prenylation.^{19,28} For example, a fungal PT for a given CDP shows high flexibility toward not only CDPs but also hydroxynaphthalenes and flavonoids.²⁸ Substrate and catalytic promiscuity were frequently reported for bacterial PTs as well. NphB in the biosynthesis of the naphterpin catalyzes a C-prenylation of hydroxynaphthalenes and can accept some simple phenols, phenylpropanoids, flavonoids, and stilbenes for O- and C-prenylation as well.^{29,30} Until now, only two PTs from CDPS-dependent pathways have been described. Zhang et al.¹⁵ recently reported the identification of a two-gene cluster from *Streptomyces leeuwenhoekii*, being responsible for the biosynthesis of streptoazine C (**2**) (Chart 1). In this pathway, *cyclo*-(L-Trp-L-Trp) (cWW), the product of the CDPS SazA, was further prenylated and methylated by the bifunctional enzyme SazB containing both PT and MT domains. It was reported that

Received: September 5, 2021

Published: November 30, 2021



Chart 1. Examples of Prenylated Diketopiperazine Derivatives



SazB-PT showed strict substrate specificity and accepted only dimethylallyl diphosphate and cWW as substrates.¹⁵ The prenyltransferase DmtC1 from *Streptomyces youssoufiensis* is involved in the biosynthesis of drimენტines and catalyzes the C-3-farnesylation of *cyclo*-(L-Trp-L-Pro), *cyclo*-(L-Trp-L-Val), *cyclo*-(L-Trp-L-Leu), and *cyclo*-(L-Trp-L-Ile).¹⁴ However, all of these compounds are CDPS products of the drimentine pathway, i.e., the natural substrates of DmtC1. The sharp contrast between the high flexibility of fungal CDP PTs and the strict substrate specificity of the bacterial SazB-PT encouraged us to investigate more PTs from CDPS-dependent pathways.

In this study, we identified by genome mining a putative *cdps*-containing gene cluster with one gene for a PT and an additional gene for an MT. Heterologous expression provided evidence for their roles in the biosynthesis of streptoazine C (2). Furthermore, we proved via biotransformation that the prenyltransferase SasB was able to prenylate diverse tryptophan-containing cyclodipeptides and their dehydrogenated derivatives, which highlights its potential as a useful biocatalyst to generate diverse prenylated DKPs.

RESULTS AND DISCUSSION

Identification and Analysis of the Putative *sas* Gene Cluster. Genome mining and heterologous expression in a well-characterized host have been proven to be an efficient strategy to explore the silent/cryptic biosynthetic potential for natural product production.^{31–33} Using this strategy, we have successfully identified several new metabolites from different CDPS-associated biosynthetic pathways and characterized intriguing chemical reactions thereof, including the novel nucleobase-containing alkaloid guanitrypmycins and several dimeric DKPs with distinct linkage patterns.²⁶ In analogy, we

analyzed a wide range of *cdps*-containing clusters by using characterized proteins as probes and identified a candidate from *S. aurantiacus* NRRL ISP-5412. The cluster of interest, termed the *sas* cluster, consists of three open reading frames coding for a putative CDPS (SasA, WP_079103588.1) and two tailoring enzymes, SasB (WP_121505431.1) and SasC (WP_054413754.1) (Table S1). SasA with a polypeptide chain length of 252 amino acids shares a sequence identity of 82% on the amino acid level with SazA mentioned above (Table S1). SasB comprising 347 amino acids displays sequence identities of 85% and 38% with the known SazB-PT and DmtC1, respectively, indicating its role as a prenyltransferase. Phylogenetic analysis with functionally characterized PTs showed that SasB and SazB-PT are closely located to each other (Figure S1). The 290 amino acid bearing SasC has a high sequence identity of 82% with the MT domain of SazB. All of these data indicate that the two clusters probably evolved from the same ancestor and underwent diversification during the evolutionary process.

Functional Proof of the Gene Cluster in the Biosynthesis of Streptoazine. To verify their functions, we first cloned the *cdps* gene *sasA* from *S. aurantiacus* into pPWW50A³⁴ and expressed it in *Streptomyces albus* J1074 (Tables S2 and S3).³⁵ The obtained transformant harboring *sasA* was cultivated in modified RS media at 28 °C for 7 days. The bacterial culture was subsequently extracted with EtOAc and analyzed by LC-MS. In comparison to the host strain J1074 harboring pPWW50A (Figure 1-i), one predominant product (1) bearing a $[M + H]^+$ ion at m/z 373.1659 was detected (Figure 1-ii). Compound 1 was identified as cWW by comparison with an authentic standard, proving SasA to be a cWW synthase. Afterward, the whole gene cluster comprising *sasABC* was cloned into pPWW50A and overexpressed in

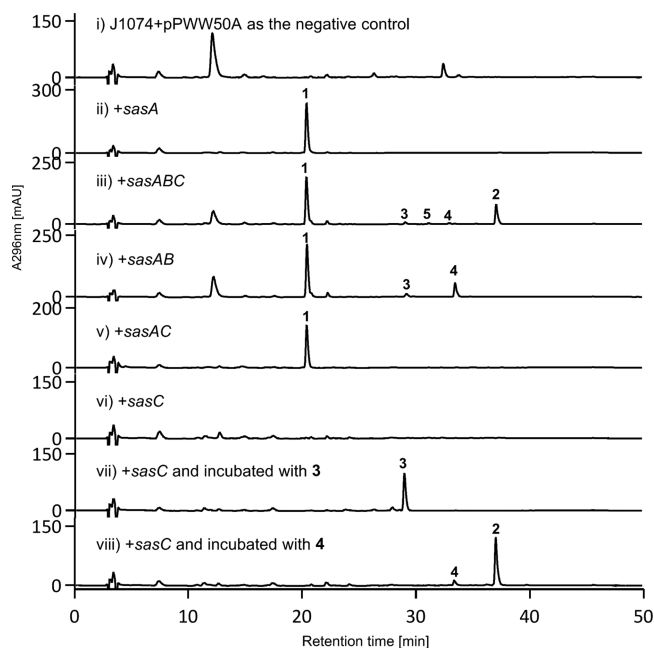


Figure 1. LC-MS analysis of *S. albus* J1074 transformants with and without precursors. Absorptions at UV 296 nm are illustrated. $[M + H]^+$ ions with a tolerance range of ± 0.005 were detected at m/z 373.166 for 1, 537.322 for 2, 441.229 for 3, 509.291 for 4, and 523.307 for 5, respectively. The $[M + H]^+$ ion of the peak at 32.5 min in the negative control differs from that of 4.

J1074 as described above. In addition to the predominant 1, four new additional compounds were observed (Figure 1-iii). The second dominant product (2) was detected with a $[M + H]^+$ ion at m/z 537.3224, 164 Da larger than that of cWW, indicating the attachment of two prenyl and two methyl groups to 1. The three minor compounds 3, 4, and 5 with $[M + H]^+$ ions at 441.2285, 509.2911, and 523.3068 are 68, 136, and 150 Da larger than 1, implying one prenyl, two prenyl, and two prenyl moieties plus one methyl group in their structures, respectively. Compound 2 was then isolated by semi-preparative HPLC after large-scale fermentation. Comprehensive interpretation and comparison of the ^1H NMR data as well as the ECD spectrum (Table S4, Figures S2 and S59) with those reported in the literature¹⁵ confirmed 2 to be

streptoazine C (Scheme 1). These data strongly support the function of SasB as a regular C-3-prenyltransferase and SasC as an indoline *N*-methyltransferase. Due to the low product yields, 3–5 could not be isolated from the *sasABC* transformant for structure elucidation by NMR analysis.

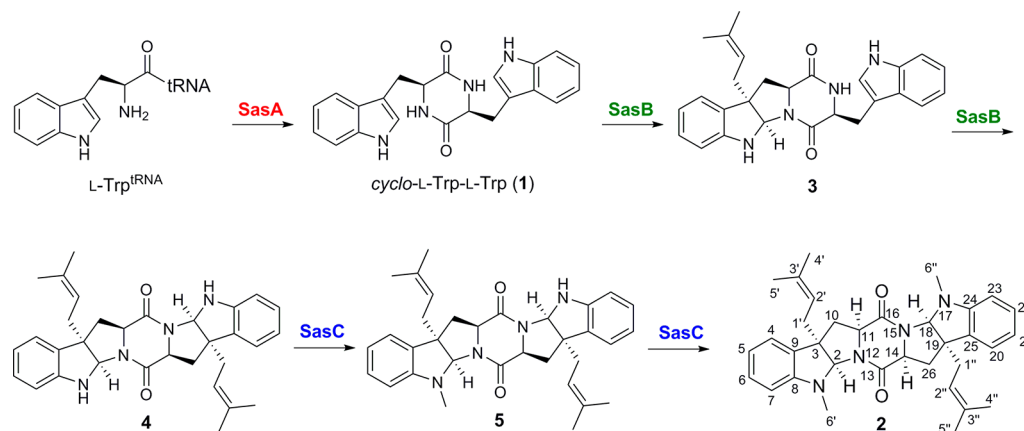
To confirm the SasB and SasC functions and figure out the reaction order, we performed the coexpression of *sasA* with *sasB* and *sasC* separately, that is, *sasAB* and *sasAC*. In addition to the CDPS product 1, two additional compounds, 3 and 4, were detected in the fermentation culture of the *sasAB* transformant (Figure 1-iv). Isolation and structure elucidation by MS and ^1H NMR analyses as well as comparison with the data of known compounds¹⁵ confirmed 3 and 4 to be regularly C-3 monoprenylated cWW and streptoazine A, respectively (Scheme 1, Table S4, and Figures S3 and S4). In contrast, only 1 was observed in the culture of the *sasAC* transformant (Figure 1-v). Neither mono- nor dimethylated 1 was detected in the *sasABC* transformant, even in the sensitive EIC chromatogram (data not shown). These results supported that 1 cannot be methylated by the methyltransferase SasC and prenylation and cyclization take place before methylation.

Incubating 4 with the *sasC* transformant led to the clear detection of 2 (Figure 1-vi), whereas no new peaks were observed in the culture after incubation with 3 (Figure 1-vii). This demonstrated that methylation proceeds only after the attachment of two prenyl moieties (Scheme 1). This is also the reason for the absence of methylated monoprenylated 1 in the *sasABC* transformant (Figure 1-iii). This order of reactions is the same as that recently reported for the two-gene cluster responsible for streptoazine C biosynthesis.¹⁵

To further verify that the formation of 3 and 4 is catalyzed by SasB, its coding sequence was cloned into pPWW50A and expressed in J1074. In comparison to J1074 harboring the empty vector, neither 1 nor other additional metabolites were observed in the *sasB* transformant (Figure 2A-iii). Incubating the *sasB* transformant with 1 (100 μM) and cultivation for 5 days led to the identification of 3 and 4 (Figure 2A-iv), whereas no consumption of 1 was found in the control culture (Figure 2A-ii). These results demonstrated that SasB is able to catalyze the regular C-3 prenylation of 1.

Substrate Promiscuity of SasB and Generation of Diverse Prenylated Tryptophan-Containing DKP Derivatives. SasB acts as a C-3-prenyltransferase and complements

Scheme 1. Biosynthetic Pathway of Streptoazine in *S. aurantiacus*^a



^aThe structures of streptoazine C (2), compound 3, and streptoazine A (4) were confirmed by NMR analysis.

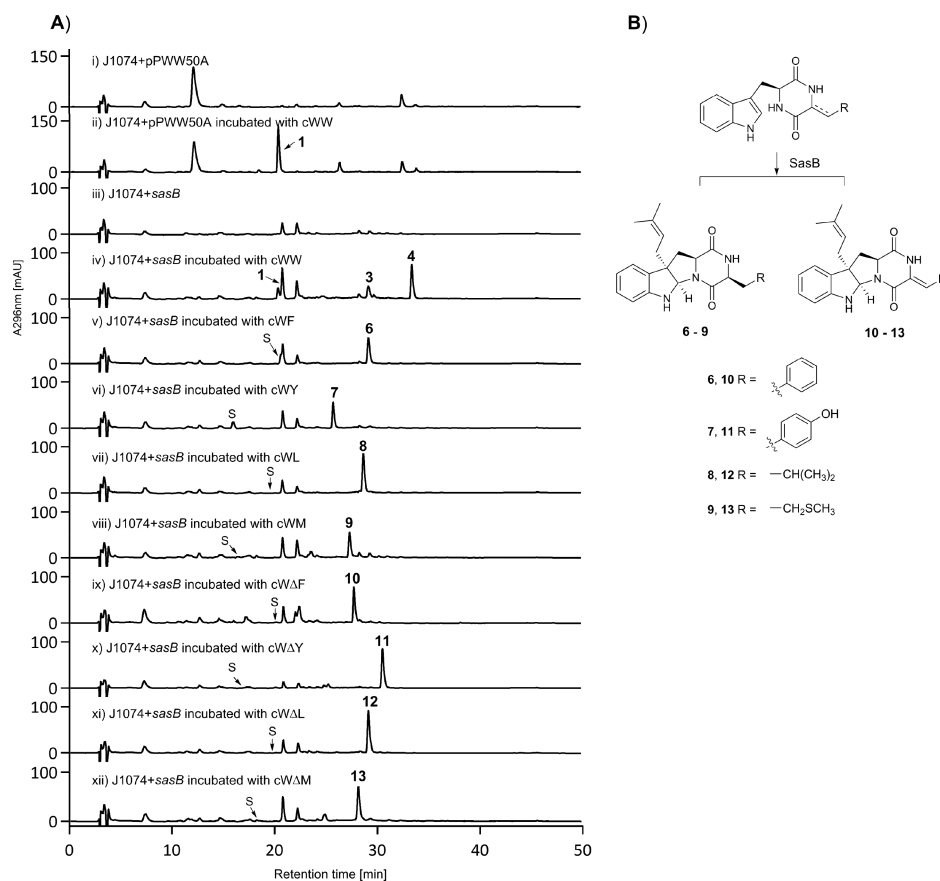


Figure 2. (A) HPLC analysis of the *sasB* transformant with and without precursors and (B) prenylated products of SasB. The $[M + H]^+$ ions of the peaks at 26.3 and 32.5 min in the negative control differ clearly from those of 7 and 4, respectively. S: substrate.

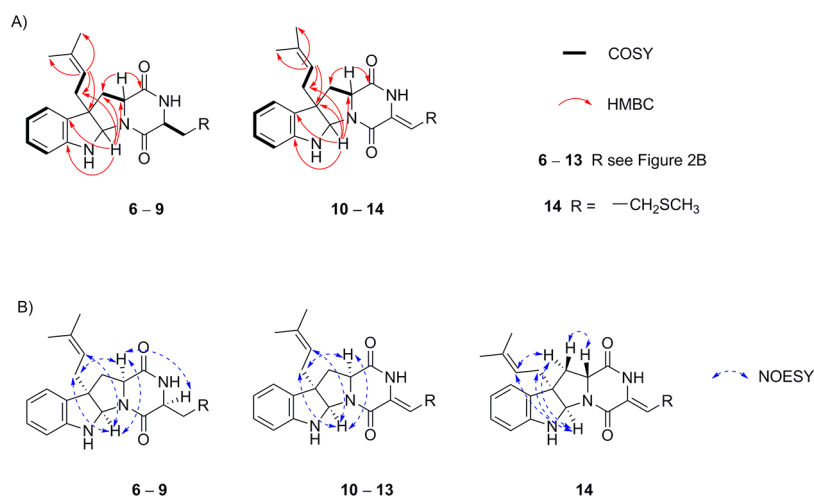


Figure 3. (A) Key COSY, and HMBC and (B) NOESY correlations in 6–14.

the reverse C-3-prenylation of fungal CDP PTs such as AnaPT, CdpNPT, and CdpC3PT¹⁷ and the regular C-3-prenylation feature of FtmPT1.³⁶ Such an enzyme with flexible substrate specificity is welcome for its potential use in the production of prenylated DKPs.³⁷ We therefore investigated the substrate specificity of SasB. Because no recombinant protein was obtained after heterologous expression in *E. coli* and *Streptomyces* (data not shown), *in vitro* testing for the acceptance of tryptophan-containing CDPs by SasB cannot be achieved. Thus, we supplied 12 CDPs to the J1074

transformant with *sasB* and monitored their consumption by LC-MS analysis. As shown in Figure 2A–v–viii, cWF, cWY, cWL, and cWM were efficiently converted by SasB. The $[M + H]^+$ ions of the products 6–9 are 68 Da larger than those of the precursors, indicating the attachment of one dimethylallyl moiety to the substrates.

Large-scale fermentation and subsequent isolation via preparative HPLC resulted in the products 6–9 of high purity for NMR analysis including ¹H, ¹³C, COSY, HSQC, HMBC, and NOESY (Figures 3 and S5–S28). The typical signals of a

Table 1. NMR Data of Compounds 6–9 in DMSO- d_6

| position | 6 | | 7 | | 8 | | 9 | |
|-------------|-----------------------|-------------------------------|-----------------------|-------------------------------|-----------------------|-------------------------------|-----------------------|-------------------------------|
| | δ_C , type | δ_H , multi. (J in Hz) | δ_C , type | δ_H , multi. (J in Hz) | δ_C , type | δ_H , multi. (J in Hz) | δ_C , type | δ_H , multi. (J in Hz) |
| 1 | | 6.53, d (2.5) | | 6.53, d (3.0) | | 6.50, d (3.0) | | 6.49, d (3.0) |
| 2 | 79.9, CH | 5.24, d (2.5) | 79.8, CH | 5.24, d (3.0) | 79.8, CH | 5.26, d (3.0) | 79.8, CH | 5.26, d (3.0) |
| 3 | 55.0, C | | 55.0, C | | 55.1, C | | 55.0, C | |
| 4 | 122.4, CH | 7.04, d (7.3) | 122.4, CH | 7.05, d (7.3) | 122.4, CH | 7.07, d (7.4) | 122.3, CH | 7.07, d (7.3) |
| 5 | 117.4, CH | 6.59, t (7.4) | 117.4, CH | 6.58, td (7.3, 1.2) | 117.5, CH | 6.59, dd (7.4, 0.9) | 117.5, CH | 6.59, td (7.3, 1.0) |
| 6 | 127.7, CH | 6.96, t (7.6) | 127.7, CH | 6.96, td (7.7, 1.2) | 127.7, CH | 6.95, dd (7.7, 0.9) | 127.7, CH | 6.95, td (7.8, 1.0) |
| 7 | 108.7, CH | 6.53, t (7.8) | 108.7, CH | 6.51, d (7.7) | 108.7, CH | 6.49, d (7.7) | 108.7, CH | 6.50, d (7.8) |
| 8 | 148.5, C | | 148.7, C | | 148.5, C | | 148.5, C | |
| 9 | 132.6, C | | 132.6, C | | 132.6, C | | 132.6, C | |
| 10 α | 39.1, CH ₂ | 2.32, dd (12.8, 7.6) | 39.8, CH ₂ | 2.31, dd (12.9, 8.0) | 38.9, CH ₂ | 2.34, dd (13.0, 7.8) | 39.3, CH ₂ | 2.35, dd (13.0, 7.6) |
| 10 β | | 1.90, dd (12.8, 10.3) | | 1.91, dd (12.9, 10.1) | | 2.05, dd (13.0, 9.8) | | 2.01, dd (13.0, 10.2) |
| 11 | 57.2, CH | 4.45, t (8.7) | 57.2, CH | 4.44, t (8.8) | 57.4, CH | 4.49, t (8.8) | 57.3, CH | 4.48, br t (8.7) |
| 13 | 167.3, C | | 167.5, C | | 168.4, C | | 167.5, C | |
| 14 | 55.8, CH | 4.37, t (5.2) | 56.0, CH | 4.27, t (5.3) | 52.9, CH | 4.04, dd (7.8, 5.1) | 53.6, CH | 4.18, br t (5.1) |
| 15 | | 7.87, s | | 7.71, s | | 7.96, s | | 8.14, s |
| 16 | 169.6, C | | 169.6, C | | 170.3, C | | 170.2, C | |
| 17 α | 34.5, CH ₂ | 3.16, dd (14.5, 5.1) | 33.6, CH ₂ | 3.03, dd (14.5, 5.2) | 37.6, CH ₂ | 1.79, ddd (13.6, 8.5, 4.9) | 28.7, CH ₂ | 2.10, m |
| 17 β | | 3.01, dd (14.5, 5.4) | | 2.89, dd (14.5, 5.6) | | 1.39, ddd (13.6, 7.8, 5.9) | | 1.89, m |
| 18 | 137.3, C | | 127.3, C | | 24.0, CH | 1.90, m | 28.9, CH ₂ | 2.58, t (8.5) |
| 19 | 129.5, CH | 7.31, d (7.5) | 130.5, CH | 7.10, d (8.5) | 22.8, CH ₃ | 0.87, d (6.3) | | |
| 20 | 127.9, CH | 7.20, t (7.4) | 114.8, CH | 6.59, dd (8.5, 1.9) | 21.8, CH ₃ | 0.88, d (6.3) | 14.4, CH ₃ | 2.05, s |
| 21 | 126.2, CH | 7.14, t (6.9) | 155.7, C | | | | | |
| OH-21 | | | | 9.11, s | | | | |
| 22 | 127.9, CH | 7.20, t (7.4) | 114.8, CH | 6.59, dd (8.5, 1.9) | | | | |
| 23 | 129.5, CH | 7.31, d (7.5) | 130.5, CH | 7.10, d (8.5) | | | | |
| 1' | 34.4, CH ₂ | 2.44, d (6.9) | 34.3, CH ₂ | 2.44, d (7.2) | 34.2, CH ₂ | 2.46, d (7.2) | 34.2, CH ₂ | 2.46, d (7.2) |
| 2' | 119.6, CH | 5.02, t (6.9) | 119.6, CH | 5.01, br t (7.2) | 119.7, CH | 5.03, t (7.2) | 119.7, CH | 5.02, t (7.2) |
| 3' | 133.9, C | | 133.9, C | | 133.9, C | | 133.9, C | |
| 4' | 17.9, CH ₃ | 1.57, s | 17.9, CH ₃ | 1.57, s | 17.9, CH ₃ | 1.58, s | 17.9, CH ₃ | 1.58, s |
| 5' | 25.6, CH ₃ | 1.63, s | 25.7, CH ₃ | 1.63, s | 25.6, CH ₃ | 1.64, s | 25.6, CH ₃ | 1.64, s |

^aSee Supporting Information for structure numbering.

Table 2. NMR Data of Compounds 10 and 11 in DMSO-*d*₆

| position | 10 | | 11 | |
|-------------|-----------------------|-------------------------------|-----------------------|-------------------------------|
| | δ_C , type | δ_H , multi. (J in Hz) | δ_C , type | δ_H , multi. (J in Hz) |
| 1 | | 6.64, d (2.8) | | 6.60, d (2.8) |
| 2 | 79.9, CH | 5.32, d (2.8) | 79.8, CH | 5.29, d (2.8) |
| 3 | 54.7, C | | 54.7, C | |
| 4 | 122.5, CH | 7.11, d (7.1) | 122.5, CH | 7.10, d (7.5) |
| 5 | 117.6, CH | 6.62, t (7.7) | 117.6, CH | 6.61, td (7.5, 1.3) |
| 6 | 127.8, CH | 6.97, br t (7.6) | 127.8, CH | 6.96, td (7.9, 1.3) |
| 7 | 108.8, CH | 6.52, d (7.8) | 108.7, CH | 6.51, d (7.9) |
| 8 | 148.5, C | | 148.5, C | |
| 9 | 132.7, C | | 132.8, C | |
| 10 α | 39.8, CH ₂ | 2.46 ^a | 39.5, CH ₂ | 2.44, dd (12.8, 7.4) |
| 10 β | | 2.04, dd (12.8, 10.5) | | 2.04, dd (12.8, 10.5) |
| 11 | 56.6, CH | 4.72, dd (10.5, 7.4) | 56.5, CH | 4.65, dd (10.5, 7.4) |
| 13 | 160.4, C | | 161.0, C | |
| 14 | 128.3, C | | 125.7, C | |
| 15 | | 9.97, br s | | 9.86, br s |
| 16 | 167.3, C | | 167.2, C | |
| 17 | 115.9, CH | 6.77, s | 116.8, CH | 6.70, s |
| 18 | 133.3, C | | 124.1, C | |
| 19/23 | 129.4, CH | 7.56, d (7.6) | 131.2, CH | 7.42, d (8.6) |
| 20/22 | 128.6, CH | 7.41, t (7.4) | 115.5, CH | 6.80, d (8.6) |
| 21 | 128.2, CH | 7.32, t (7.4) | 157.8, C | |
| OH-21 | | | | 9.78, br s |
| 1' | 34.6, CH ₂ | 2.46 ^a | 34.6, CH ₂ | 2.48, d (7.2) |
| 2' | 119.6, CH | 5.02, t (7.3) | 119.6, CH | 5.02, t (7.2) |
| 3' | 134.1, C | | 134.0, C | |
| 4' | 18.0, CH ₃ | 1.59, s | 18.0, CH ₃ | 1.58, s |
| 5' | 25.7, CH ₃ | 1.64, s | 25.7, CH ₃ | 1.64, s |

^aSignals overlapping with each other. See Supporting Information for structure numbering.

regular C-3-prenyl residue in the ¹H NMR spectra are found in the ranges of δ_H 2.44–2.46 (d, 6.9–7.2 Hz, H-1'), 5.01–5.03 (t, 6.9–7.2 Hz, H-2'), 1.57–1.58 (s, H-4'), and 1.63–1.64 (s, H-5') (Table 1). The signals of the five carbons are detected in the ¹³C spectra at about δ_C 34 (C-1'), 120 (C-2'), 134 (C-3'), 18 (C-4'), and 26 (C-5') (Table 1). Prenylation at C-3 destroys the aromatic character of the indole system and causes a shielded shift of the H-2 signal to δ_H 5.24–5.26 as well as those of C-2 and C-3 to δ_C 80 and 55, respectively. The configuration of the products was determined based on the correlations between H-1' and H-11, H-1', and H-2 as well as H-2 and H-11 in the NOESY spectra. Comparison of their ECD spectra provided additional evidence for their configurations (Figure S60). All the obtained data confirmed that 6–9 are C-3-prenylated derivatives of the corresponding CDPs (Figure 2B). More fascinatingly, SasB performed also very effective conversions of these four substrates, at least in our experiments.

Low conversions to prenylated derivatives were also detected by LC-MS analysis for *cyclo*-(L-Trp-L-Ala), *cyclo*-(L-Trp-D-Ala), *cyclo*-(D-Trp-L-Ala), *cyclo*-(D-Trp-D-Ala), *cyclo*-(L-Trp-L-Pro), *cyclo*-(L-Trp-D-Pro), *cyclo*-(D-Trp-L-Pro), and *cyclo*-(D-Trp-D-Pro) (Figures S63 and S64). Due to the low product yields, the structures of these products could not be elucidated in this study. These results suggest a more flexible substrate specificity of SasB from *S. aurantiacus* than that of SazB from *S. leeuwenhoekii*.¹⁵ It was reported that cWF, cWY, cWA, and cWP were not accepted by SazB.¹⁵ In our case, all of these four CDPs were prenylated by SasB with high conversion for cWF and cWY (Figures 2, S63, and S64).

Cyclodipeptide oxidases (CDOs) are frequently found in the CDPS-related pathways and install *exo* double bonds at the DKP ring.²⁶ For combinatorial application of SasB with these oxidases, we tested its acceptance of the dehydrogenated forms of the four efficiently converted CDPs, i.e., cWΔF, cWΔY, cWΔL, and cWΔM, by incubation experiments in the *sasB* transformant. LC-MS analysis showed that all of these compounds were good substrates for SasB and were completely converted to their prenylated products (Figure 2A-ix-xii). The products 10–13 were subsequently isolated, and their structures confirmed to be regularly C-3-prenylated derivatives at the indoline ring (Figure 2B) by detailed interpretation of their NMR data and the comparison with the data of 6–9 (Tables 2 and 3, and Figures 3, S29–S52, and S61). Observation of the interaction between NH-15 and H-19/H-23 in the NOESY spectrum of 11 as well as NH-15 and H-18 in that of 12 supported the *Z*-configuration of the *exo* double bonds in their structures (Figures S40 and S46).

During the isolation procedure, we observed the conversion of the cWΔM product 13 to the new compound 14. Isolation by using a chiral-phase HPLC column (Figure S65) and structure elucidation by NMR analysis including interpretation of the NOESY data and comparison of its ECD spectrum with that of 13 (Table 3 and Figures 3, S53–S58, and S62) confirmed the epimerization at the C-11 position. As the nonenzymatic epimerization via keto–enol tautomerism was already observed for the guanitrypmysins,³⁸ we speculated a similar mechanism may explain the conversion of 13 to 14. Incubation of 13 in CD₃OD/D₂O (1:1) at pH 9 and 12 for 14 h and LC-MS analysis confirmed indeed the conversion of 13

Table 3. NMR Data of Compounds 12–14 in DMSO-*d*₆

| position | 12 | | δ_C | 13 | | δ_C , type | δ_H , multi. (<i>J</i> in Hz) |
|-------------|-----------------------|---------------------------------------|-----------------------|---|-----------------------|-------------------|---------------------------------------|
| | δ_C , type | δ_H , multi. (<i>J</i> in Hz) | | δ_H , multi. (<i>J</i> in Hz) | δ_C , type | | |
| 1 | | 6.55, d (2.8) | | 6.60, d (2.9) | | | 6.51, d (1.0) |
| 2 | 79.8, CH | 5.23, d (2.8) | 79.9, CH | 5.26, d (2.9) | 79.0, CH | | 5.33, br s |
| 3 | 54.4, C | | 54.4, C | | 54.4, C | | |
| 4 | 122.5, CH | 7.08, d (7.2) | 122.5, CH | 7.08, dd (7.4, 1.1) | 123.1, CH | | 7.12, d (7.2) |
| 5 | 117.5, CH | 6.60, td (7.2, 1.0) | 117.5, CH | 6.60, td (7.4, 1.1) | 117.6, CH | | 6.61, td (7.2, 1.0) |
| 6 | 127.8, CH | 6.95, td (7.7, 1.0) | 127.8, CH | 6.96, td (7.8, 1.3) | 128.1, CH | | 6.97, td (7.4, 1.0) |
| 7 | 108.6, CH | 6.49, d (7.7) | 108.6, CH | 6.50, d (7.8) | 108.2, CH | | 6.51, d (7.4) |
| 8 | 148.5, C | | 148.5, C | | 150.2, C | | |
| 9 | 132.8, C | | 132.7, C | | 130.9, C | | |
| 10 α | 40.1, CH ₂ | 2.42, dd (12.8, 7.1) | 40.1, CH ₂ | 2.44, dd (12.7, 7.1) | 39.8, CH ₂ | | 2.20, t (11.9) |
| 10 β | | 1.94, dd (12.8, 10.8) | | 1.95, dd (12.7, 11.0) | | | 2.46, dd (11.9, 5.9) |
| 11 | 56.3, CH | 4.57, dd (10.8, 7.1) | 56.4, CH | 4.62, dd (11.0, 7.1) | 57.8, CH | | 4.02, dd (11.9, 5.9) |
| 13 | 159.7, C | | 159.1, C | | 157.0, C | | |
| 14 | 127.0, C | | 130.1, C | | 129.7, C | | |
| 15 | | 9.92, s | | 10.12, s | | | 10.10, s |
| 16 | 166.9, C | | 166.9, C | | 165.9, C | | |
| 17 | 125.5, CH | 5.68, d (10.5) | 113.6, CH | 5.84, t (8.7) | 112.8, CH | | 5.77, t (8.7) |
| 18 | 24.4, CH | 2.82, m | 27.3, CH ₂ | 3.41, dd (14.0, 8.7) 3.29 ^a | 27.2, CH ₂ | | 3.39, dd (14.0, 8.7) |
| 19 | 22.3, CH ₃ | 0.96, d (6.5) | | | | | |
| 20 | 22.1, CH ₃ | 0.96, d (6.5) | 14.0, CH ₃ | 2.01, s | 14.0, CH ₃ | | 1.99, s |
| 1' | 34.6, CH ₂ | 2.46, d (7.3) | 34.6, CH ₂ | 2.47, d (7.2) | 36.0, CH ₂ | | 2.40, d (7.3) |
| 2' | 119.6, CH | 4.99, t (7.3) | 119.6, CH | 4.99, t (7.2) | 119.3, CH | | 5.08, t (7.3) |
| 3' | 134.0, C | | 134.0, C | | 133.8, C | | |
| 4' | 18.0, CH ₃ | 1.58, s | 18.0, CH ₃ | 1.58, s | 18.0, CH ₃ | | 1.51, s |
| 5' | 25.7, CH ₃ | 1.63, s | 25.7, CH ₃ | 1.64, s | 25.7, CH ₃ | | 1.64, s |

^aSignals overlapping with that of water. See Supporting Information for structure numbering.

to 14 and incorporation of one deuterium in both molecules (Figure S66). This supported the epimerization at C-11 via keto–enol tautomerism.

Taking the above results together, a putative *cdps*-containing gene cluster with a prenyltransferase and a methyltransferase gene was identified from *S. aurantiacus* and successfully expressed in *S. albus* J1074. The diprenylated and dimethylated diketopiperazine indole alkaloid streptoazine C (2) was identified as the cluster end product. Heterologous expression of different gene combinations and precursor incubation experiments enabled us to evaluate the order of the biosynthetic steps to 2. The cWW formation catalyzed by SasA is followed by two regular prenylation steps with SasB and final *N*-methylations with SasC. It is noteworthy that the same order of reaction steps occurs in the biosynthesis of the same product in *S. leuwenhoekii*.¹⁵ However, differing from *sazB* from *S. leuwenhoekii* coding for both PT and MT activities, the two independent genes *sasB* and *sasC* are located in the streptoazine cluster in *S. aurantiacus* described in this study. More importantly, SasB displays a remarkably high substrate tolerance and can accept not only a number of tryptophan-containing CDPs but also their dehydrogenated derivatives for prenyl decoration. The successful production of dehydrogenated and prenylated CDPs by combination of cyclodipeptide oxidases and the prenyltransferases SasB will provide an excellent example for accessing diversified natural products in the combinatorial biology field, which also inspires us to explore the combination of SasB with other modification genes in order to access more diverse DKPs. This is the first report on the substrate flexibility of a CDPS-related PT toward non-natural substrates. Identification of additional PTs from

this less explored enzyme group will provide more details on their biochemical properties.

EXPERIMENTAL SECTION

General Experimental Procedures. The optical rotation was measured with an A KRÜSS P3000 polarimeter at 20 °C using the D-line of the sodium lamp at $\lambda = 589.3$ nm. Prior to the measurement, the polarimeter was calibrated with chloroform as solvent. Circular dichroism spectra were taken on a J-1500 CD spectrometer (Jasco Deutschland GmbH). The samples were dissolved in MeOH and measured in the range of 200–400 nm by using a 1 mm path length quartz cuvette. The NMR spectra of the purified compounds were recorded on a JEOL ECA-500 MHz spectrometer in DMSO-*d*₆, and all spectra were processed with MestReNova 9.0.0 (Metrelab). Chemical shifts are referred to those of DMSO-*d*₆ (δ_H 2.50 and δ_C 39.5). High-resolution mass spectrometric analysis was performed on an Agilent 1260 HPLC system equipped with a microTOF-Q III spectrometer (Bruker) using a Multospher 120 RP-18 column (250 × 2 mm, 5 μ m) (CS-Chromatographie Service GmbH) or a Multohigh Chiral AM-RP HPLC column (150 × 4.6 mm, CS-Chromatographie Service GmbH). Electrospray positive ionization mode was selected for determination of the exact masses. The capillary voltage was set to 4.5 kV and the collision energy to 8.0 eV. Sodium formate was used in each run for mass calibration. The masses were scanned in the range of *m/z* 100–1500. Data were evaluated with the Compass DataAnalysis 4.2 software (Bruker Daltonik). Semipreparative HPLC was performed on the same HPLC equipment with an Agilent ZORBAX Eclipse XDB C18 HPLC column (250 × 9.4 mm, 5 μ m) and a Multohigh Chiral AM-RP HPLC column (250 × 10 mm, CS-Chromatographie Service GmbH). Sephadex LH-20 (Merck) was used for column chromatography, and HPLC for detection of the desired substance in the fractions.

Chemicals. *cyclo*-(L-Trp-L-Trp), *cyclo*-(L-Trp-L-Phe), *cyclo*-(L-Trp-L-Tyr), and *cyclo*-(L-Trp-L-Leu) were purchased from Bachem. *cyclo*-(L-Trp-L-Trp), *cyclo*-(L-Trp-L-Phe), *cyclo*-(L-Trp-L-Tyr), and *cyclo*-(L-Trp-L-Leu) were purchased from Bachem.

(L-Trp-L-Met) was isolated from the strain described previously.²² The stereoisomers of *cyclo*-Trp-Ala and *cyclo*-Trp-Pro were synthesized according to the method published previously.³⁹ The dehydrogenated CDP substrates used in this study were prepared by using cyclodipeptide oxidases according to the method described in a previous work.⁴⁰

Bacterial Strains, Plasmids, and Growth Conditions.

Streptomyces aurantiacus NRRL ISP-5412 was obtained from the ARS Culture Collection (NRRL). *Streptomyces albus* J1074³⁵ was kindly gifted by Prof. Dr. Andriy Luzhetskyy (Saarland University). *S. albus* J1074 and the generated exconjugants were maintained on MS plates (mannitol 20.0 g/L, soya flour 20.0 g/L, agar 20.0 g/L) at 28 °C for sporulation. For secondary metabolite production, *S. albus* J1074 transformants were cultivated in liquid modified R5 medium (sucrose 103.0 g/L, glucose 10.0 g/L, yeast extract 5.0 g/L, MgCl₂·6H₂O 10.12 g/L, K₂SO₄ 0.25 g/L, Difco casamino acids 0.1 g/L, MOPS 21.0 g/L, trace element solution 2 mL/L, pH 7.2) at 28 °C for 7 days.

Computer-Assisted Sequence Analysis. Nucleotide and amino acid sequences used in this study were obtained from NCBI databases (<http://www.ncbi.nlm.nih.gov>). Comparison of protein sequences was carried out by using the BLASTP program (<http://blast.ncbi.nlm.nih.gov/>). The phylogenetic tree of PTs (Figure S1) was created by MEGA version 7.0 (<http://www.megasoftware.net>).

Genetic Manipulation, PCR Amplification, and Gene Cloning. Strains and plasmids used in this study are listed in Table S2 and Table S3, respectively. Recombinant *E. coli* strains were cultivated in liquid or on solid lysogeny-broth (LB) with 100 µg/mL ampicillin, 50 µg/mL kanamycin, 50 µg/mL apramycin, or 25 µg/mL chloramphenicol when necessary.

Genetic manipulation in *E. coli* was performed according to the protocol by Green and Sambrook.⁴¹ Genomic DNA isolation from *Streptomyces* was performed as described in the literature.⁴²

Gene regions were amplified by PCR from genomic DNA of *S. aurantiacus* by using primers listed in Table S3 and Phusion high-fidelity DNA polymerase from New England Biolabs. The generated PCR fragments were cloned into pGEM-T Easy vector (Promega), and the sequence integrity was confirmed by sequencing. Subsequently, the fragments were released with restriction endonucleases from pGEM-T Easy and ligated into pPWW50A,³⁴ which was digested with the same enzymes, previously. The generated constructs (Table S3) were used for heterologous expression in *S. albus* J1074.

Heterologous Gene Expression in *Streptomyces albus* J1074 and Cultivation for Secondary Metabolite Production. The constructed plasmids harboring different genes or the gene cluster were transformed into nonmethylated *E. coli* ET12567/pUZ8002, then conjugated with *S. albus* J1074. The positive conjugants were selected by the phenotype showing apramycin resistance and further confirmed by PCR. The spores of the *S. albus* J1074 transformants were inoculated into 50 mL of modified R5 liquid media supplied with 50 µg/mL of apramycin in 250 mL baffled flasks and cultured at 28 °C and 200 rpm for 7 days. This culture (1 mL) was extracted with the same volume of EtOAc three times. After that, the organic phases were combined and evaporated, and the dried residues were dissolved in 200 µL of MeOH. Samples (5 µL) were subjected to LC-MS for analysis.

Biotransformation for the Generation of Various Prenylated DKPs. The *S. albus* J1074 transformant harboring *sasB* was incubated in modified R5 medium at 28 °C, 200 rpm for 2 days. Tryptophan-containing CDPs or dehydrogenated derivatives were separately added to 10 mL of these precultures to final concentrations of 100 µM. After cultivation at 28 °C for an additional 5 days, the metabolites were extracted with EtOAc and analyzed by LC-MS.

LC-MS Analysis. For secondary metabolite analysis, a linear gradient of 5–100% MeCN in H₂O, both containing 0.1% HCOOH, in 40 min and a flow rate at 0.25 mL/min were used. The column was then washed with 100% MeCN containing 0.1% HCOOH for 5 min and equilibrated with 5% MeCN in H₂O for 5 min. For analysis of the samples after incubation of the prenylated dehydrogenated DKPs in CD₃OD/D₂O, a Multohigh Chiral AM-RP HPLC column (150 × 4.6

mm, CS-Chromatographie Service GmbH) was used. Separation was carried out with a linear gradient of 50–100% MeCN in H₂O in 30 min and a flow rate of 0.5 mL/min.

Extraction and Isolation of Secondary Metabolites. For structure elucidation of the accumulated products, *S. albus* J1074 transformants harboring *sasABC* and *sasAB* were fermented in modified R5 medium on large scales (8 L) at 28 °C for 7 days. The culture supernatants were collected and extracted with an equal volume of EtOAc three times. The EtOAc phases were evaporated to dryness, dissolved in MeOH, and applied to chromatography on a Sephadex LH-20 column with MeOH as eluent. The fractions containing the target products were further purified on an Agilent HPLC 1260 series by using a semipreparative Agilent ZORBAX Eclipse XDB C18 HPLC column (250 × 9.4 mm, 5 µm). The flow rate was set to 2.0 mL/min. Compounds 2, 3, and 4 were purified with 95%, 70%, and 85% MeCN in H₂O, respectively.

For the prenylated DKP derivatives generated by biotransformation, the extracts were obtained by extraction with EtOAc as described above. Compounds 6–12 were further purified on an Agilent HPLC 1260 series with 60–65% MeCN in H₂O. Compounds 13 and 14 were separated on an Agilent HPLC 1260 series by using a semipreparative Multohigh Chiral AM-RP HPLC column (250 × 10 mm, CS-Chromatographie Service GmbH) with 80% MeCN in H₂O. The flow rate was set to 2.0 mL/min.

Streptoazine C (2): yellow oil; [α]_D²⁰ +180 (c 0.33, CHCl₃); ECD (0.29 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 292 (+3.71), 277 (+2.64), 255 (+9.86), 229 (−3.08), 211 (+10.51) nm; ¹H NMR data, Table S4; HRESIMS *m/z*: 537.3245 [M + H]⁺ (calcd for C₃₄H₄₁N₄O₂, 537.3224).

Compound 3: yellow oil; [α]_D²⁰ +45 (c 0.40, CHCl₃); ECD (0.27 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 290 (+3.09), 265 (+1.59), 244 (+10.82), 222 (−4.78), 214 (−4.60) nm; ¹H NMR data, Table S4; HRESIMS *m/z* 441.2288 [M + H]⁺ (calcd for C₂₇H₂₉N₄O₂, 441.2285).

Streptoazine A (4): yellow oil; [α]_D²⁰ +340 (c 0.26, CHCl₃); ECD (0.30 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 291 (+2.69), 265 (+1.08), 245 (+11.14), 219 (−1.46), 208 (+25.47) nm; ¹H NMR data, Table S4; HRESIMS *m/z* 509.2911 [M + H]⁺ (calcd for C₃₂H₃₇N₄O₂, 509.2911).

Compound 6: yellow oil; [α]_D²⁰ +40.1 (c 1.87, CHCl₃); ECD (0.70 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 294 (+3.70), 270 (+0.43), 243 (+9.02), 218 (−1.24) nm; ¹H and ¹³C NMR data, Table 1; HRESIMS *m/z* 402.2182 [M + H]⁺ (calcd for C₂₅H₂₈N₃O₂, 402.2176).

Compound 7: yellow oil; [α]_D²⁰ +25 (c 0.71, CHCl₃); ECD (0.51 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 287 (+2.23), 267 (+0.55), 243 (+12.14), 225 (−1.80), 213 (+1.14) nm; ¹H and ¹³C NMR data, Table 1; HRESIMS *m/z* 418.2127 [M + H]⁺ (calcd for C₂₅H₂₈N₃O₃, 418.2125).

Compound 8: yellow oil; [α]_D²⁰ +113 (c 1.34, CHCl₃); ECD (0.55 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 293 (+3.10), 264 (+0.83), 242 (+10.11), 223 (−0.97), 210 (+8.14) nm; ¹H and ¹³C NMR data, Table 1; HRESIMS *m/z* 368.2330 [M + H]⁺ (calcd for C₂₂H₃₀N₃O₂, 368.2333).

Compound 9: yellow oil; [α]_D²⁰ +14.2 (c 2.19, CHCl₃); ECD (0.68 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 291 (+2.28), 268 (+0.57), 243 (+8.54), 224 (+1.23), 210 (+7.29) nm; ¹H and ¹³C NMR data, Table 1; HRESIMS *m/z* 368.1906 [M + H]⁺ (calcd for C₂₁H₂₈N₃O₂S, 368.1897).

Compound 10: yellow oil; [α]_D²⁰ +150 (c 0.89, CHCl₃); ECD (0.34 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 293 (+8.48), 254 (+1.11), 229 (+12.71), 215 (+12.83) nm; ¹H and ¹³C NMR data, Table 2; HRESIMS *m/z* 400.2020 [M + H]⁺ (calcd for C₂₅H₂₆N₃O₂, 400.2020).

Compound 11: yellow oil; [α]_D²⁰ +25 (c 0.85, CHCl₃); ECD (0.31 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 328 (+7.90), 315 (+7.98), 301 (+10.74), 258 (−2.19), 221 (+20.25) nm; ¹H and ¹³C NMR data, Table 2; HRESIMS *m/z* 416.1968 [M + H]⁺ (calcd for C₂₅H₂₆N₃O₃, 416.1969).

Compound 12: yellow oil; [α]_D²⁰ +24 (c 0.67, CHCl₃); ECD (0.28 mM, MeOH) λ_{\max} ($\Delta\epsilon$) 306 (+2.94), 276 (+0.26), 248 (+13.56), 237

(+9.72), 219 (+26.54) nm; ^1H and ^{13}C NMR, Table 3; HRESIMS m/z 366.2178 $[\text{M} + \text{H}]^+$ (calcd for $\text{C}_{22}\text{H}_{28}\text{N}_3\text{O}_2$, 366.2176).

Compound 13: yellow oil; $[\alpha]_{\text{D}}^{20}$ +227 (c 1.29, CHCl_3); ECD (0.40 mM, MeOH) λ_{max} ($\Delta\epsilon$) 306 (+2.48), 287 (+1.71), 247 (+10.40), 246 (+10.25), 220 (+29.36) nm; ^1H and ^{13}C NMR data, Table 3; HRESIMS m/z 384.1746 $[\text{M} + \text{H}]^+$ (calcd for $\text{C}_{21}\text{H}_{26}\text{N}_3\text{O}_2\text{S}$, 384.1740).

Compound 14: yellow oil; $[\alpha]_{\text{D}}^{20}$ +170 (c 0.61, CHCl_3); ECD (0.48 mM, MeOH) λ_{max} ($\Delta\epsilon$) 302 (+3.88), 287 (+3.11), 249 (+12.83), 220 (−2.94), 207 (−13.33) nm; ^1H and ^{13}C NMR data, Table 3; HRESIMS m/z 384.1737 $[\text{M} + \text{H}]^+$ (calcd for $\text{C}_{21}\text{H}_{26}\text{N}_3\text{O}_2\text{S}$, 384.1740).

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jnatprod.1c00844>.

Strains, plasmids, primers, cluster information, NMR data of compounds 2–4, NMR spectra of compounds 2–14, experimental ECD spectra, HPLC, and LC-MS chromatograms (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Shu-Ming Li – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany; orcid.org/0000-0003-4583-2655; Email: shuming.li@staff.uni-marburg.de

Authors

Jing Liu – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Yiling Yang – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Lauritz Harken – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jnatprod.1c00844>

Author Contributions

§J.L. and Y.Y. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the ARS Culture Collection (NRRL) for providing *Streptomyces aurantiacus* strain and R. Kraut, L. Ludwig-Radtke, and S. Newel (Fachbereich Pharmazie, University of Marburg) for LC-MS and NMR analyses. This project was funded in part by the Deutsche Forschungsgemeinschaft (DFG, Li844/14-1 and INST 160/620-1). J.L. (201608310118) and Y.Y. (201808530447) are the scholarship recipients from the China Scholarship Council.

■ REFERENCES

- (1) Li, S.-M. *Nat. Prod. Rep.* **2010**, *27*, 57–78.
- (2) Wang, X.; Li, Y.; Zhang, X.; Lai, D.; Zhou, L. *Molecules* **2017**, *22*, 2026.
- (3) Borthwick, A. D. *Chem. Rev.* **2012**, *112*, 3641–3716.

(4) Mishra, A. K.; Choi, J.; Choi, S. J.; Baek, K. H. *Molecules* **2017**, *22*, E1796.

(5) Hayashi, H.; Takiuchi, K.; Murao, S.; Arai, M. *Agric. Biol. Chem.* **1988**, *52*, 2131–2133.

(6) Murao, S.; Hayashi, H.; Akiuchi, K.; Arai, M. *Agric. Biol. Chem.* **1988**, *52*, 885–886.

(7) Kozlovsky, A. G.; Vinokurova, N. G.; Adanin, V. M.; Burkhardt, G.; Dahse, H. M.; Gräfe, U. *J. Nat. Prod.* **2000**, *63*, 698–700.

(8) Kozlovsky, A. G.; Vinokurova, N. G.; Adanin, V. M.; Burkhardt, G.; Dahse, H.-M.; Gräfe, U. *J. Nat. Prod.* **2001**, *64*, 553–554.

(9) Arai, K.; Kimura, K.; Mushiroda, T.; Yamamoto, Y. *Chem. Pharm. Bull.* **1989**, *37*, 2937–2939.

(10) Yamazaki, M.; Sasago, K.; Miyaki, K. *J. Chem. Soc., Chem. Commun.* **1974**, 408–409.

(11) Clark, B.; Capon, R. J.; Lacey, E.; Tennant, S.; Gill, J. H. *J. Nat. Prod.* **2005**, *68*, 1661–1664.

(12) Cardani, C.; Casnati, G.; Piozzi, F.; Quilico, A. *Tetrahedron Lett.* **1959**, *1*, 1–8.

(13) Raju, R.; Piggott, A. M.; Huang, X. C.; Capon, R. J. *Org. Lett.* **2011**, *13*, 2770–2773.

(14) Yao, T.; Liu, J.; Liu, Z.; Li, T.; Li, H.; Che, Q.; Zhu, T.; Li, D.; Gu, Q.; Li, W. *Nat. Commun.* **2018**, *9*, 4091.

(15) Zhang, Y.; Yao, T.; Jiang, Y.; Li, H.; Yuan, W.; Li, W. *Appl. Environ. Microbiol.* **2021**, *87*, e02525–20.

(16) Xu, W.; Gavia, D. J.; Tang, Y. *Nat. Prod. Rep.* **2014**, *31*, 1474–1487.

(17) Winkelblech, J.; Fan, A.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 7379–7397.

(18) Li, S.-M. *J. Antibiot.* **2011**, *64*, 45–49.

(19) Mori, T. *J. Nat. Med.* **2020**, *74*, 501–502.

(20) Fraley, A. E.; Sherman, D. H. *FEBS J.* **2020**, *287*, 1381–1402.

(21) Nies, J.; Li, S.-M. *ACS Chem. Biol.* **2021**, *16*, 185–192.

(22) Liu, J.; Yu, H.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 4435–4444.

(23) Jacques, I. B.; Moutiez, M.; Witwinowski, J.; Darbon, E.; Martel, C.; Seguini, J.; Favry, E.; Thai, R.; Lecoq, A.; Dubois, S.; et al. *Nat. Chem. Biol.* **2015**, *11*, 721–727.

(24) Süßmuth, R. D.; Mainz, A. *Angew. Chem., Int. Ed.* **2017**, *56*, 3770–3821.

(25) Canu, N.; Moutiez, M.; Belin, P.; Gondry, M. *Nat. Prod. Rep.* **2020**, *37*, 312–321.

(26) Harken, L.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2021**, *105*, 2277–2285.

(27) Wohlgemuth, V.; Kindinger, F.; Xie, X.; Wang, B. G.; Li, S.-M. *Org. Lett.* **2017**, *19*, 5928–5931.

(28) Fan, A.; Winkelblech, J.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 7399–7415.

(29) Kumano, T.; Richard, S. B.; Noel, J. P.; Nishiyama, M.; Kuzuyama, T. *Bioorg. Med. Chem.* **2008**, *16*, 8117–8126.

(30) Kuzuyama, T.; Noel, J. P.; Richard, S. B. *Nature* **2005**, *435*, 983–987.

(31) Medema, M. H.; de Rond, T.; Moore, B. S. *Nat. Rev. Genet.* **2021**, *22*, 553.

(32) Covington, B. C.; Xu, F.; Seyedsayamdost, M. R. *Annu. Rev. Biochem.* **2021**, *90*, 763–788.

(33) Skinnider, M. A.; Johnston, C. W.; Gunabalasingam, M.; Merwin, N. J.; Kieliszek, A. M.; MacLellan, R. J.; Li, H.; Ranieri, M. R. M.; Webster, A. L. H.; Cao, M. P. T.; et al. *Nat. Commun.* **2020**, *11*, 6058.

(34) Zhu, Y.; Fu, P.; Lin, Q.; Zhang, G.; Zhang, H.; Li, S.; Ju, J.; Zhu, W.; Zhang, C. *Org. Lett.* **2012**, *14*, 2666–2669.

(35) Ziburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. *BMC Genomics* **2014**, *15*, 97.

(36) Wollinsky, B.; Ludwig, L.; Xie, X.; Li, S.-M. *Org. Biomol. Chem.* **2012**, *10*, 9262–9270.

(37) Dubois, P.; Correia, I.; Le Chevalier, F.; Dubois, S.; Jacques, I.; Canu, N.; Moutiez, M.; Thai, R.; Gondry, M.; Lequin, O.; et al. *Sci. Rep.* **2019**, *9*, 9208.

- (38) Liu, J.; Xie, X.; Li, S.-M. *Angew. Chem., Int. Ed.* **2019**, *58*, 11534–11540.
- (39) Yin, S.; Yu, X.; Wang, Q.; Liu, X. Q.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1649–1660.
- (40) Mikulski, L.; Schäfer, J.; Brockmeyer, K.; Kraut, R.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 2523–2536.
- (41) Green, M. R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2012.
- (42) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*, 2nd ed.; John Innes Foundation: Norwich, UK, 2000.

Supporting Information

Elucidation of the Streptoazine Biosynthetic Pathway in *Streptomyces aurantiacus* Reveals the Presence of a Promiscuous Prenyltransferase/Cyclase

Jing Liu^{† §}, Yiling Yang^{† §}, Lauritz Harken[†], and Shu-Ming Li^{†*}

[†] Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, Robert-Koch-Straße 4, 35037 Marburg, Germany

Table of Contents

| | |
|---|----------|
| Supplementary Tables | 5 |
| Table S1. Comparison of <i>sas</i> genes with known entries | 5 |
| Table S2. Bacterial strains used in this study | 5 |
| Table S3. Cloning and expression constructs used in this study | 5 |
| Table S4. ¹ H NMR data of compounds 2 — 4 in DMSO- <i>d</i> ₆ | 6 |
| Supplementary Figures | 7 |
| Figure S1. Phylogenetic analysis of PTs | 7 |
| Figure S2. ¹ H NMR spectrum of compound 2 in DMSO- <i>d</i> ₆ (500 MHz) | 8 |
| Figure S3. ¹ H NMR spectrum of compound 3 in DMSO- <i>d</i> ₆ (500 MHz) | 8 |
| Figure S4. ¹ H NMR spectrum of compound 4 in DMSO- <i>d</i> ₆ (500 MHz) | 9 |
| Figure S5. ¹ H NMR spectrum of compound 6 in DMSO- <i>d</i> ₆ (500 MHz) | 9 |
| Figure S6. ¹³ C NMR spectrum of compound 6 in DMSO- <i>d</i> ₆ (125 MHz) | 10 |
| Figure S7. HSQC spectrum of compound 6 in DMSO- <i>d</i> ₆ | 10 |
| Figure S8. COSY spectrum of compound 6 in DMSO- <i>d</i> ₆ | 11 |
| Figure S9. HMBC spectrum of compound 6 in DMSO- <i>d</i> ₆ | 11 |
| Figure S10. NOESY spectrum of compound 6 in DMSO- <i>d</i> ₆ | 12 |
| Figure S11. ¹ H NMR spectrum of compound 7 in DMSO- <i>d</i> ₆ (500 MHz) | 12 |
| Figure S12. ¹³ C NMR spectrum of compound 7 in DMSO- <i>d</i> ₆ (125 MHz) | 13 |
| Figure S13. HSQC spectrum of compound 7 in DMSO- <i>d</i> ₆ | 13 |
| Figure S14. COSY spectrum of compound 7 in DMSO- <i>d</i> ₆ | 14 |
| Figure S15. HMBC spectrum of compound 7 in DMSO- <i>d</i> ₆ | 14 |
| Figure S16. NOESY spectrum of compound 7 in DMSO- <i>d</i> ₆ | 15 |
| Figure S17. ¹ H NMR spectrum of compound 8 in DMSO- <i>d</i> ₆ (500 MHz) | 15 |
| Figure S18. ¹³ C NMR spectrum of compound 8 in DMSO- <i>d</i> ₆ (125 MHz) | 16 |
| Figure S19. HSQC spectrum of compound 8 in DMSO- <i>d</i> ₆ | 16 |
| Figure S20. COSY spectrum of compound 8 in DMSO- <i>d</i> ₆ | 17 |
| Figure S21. HMBC spectrum of compound 8 in DMSO- <i>d</i> ₆ | 17 |
| Figure S22. NOESY spectrum of compound 8 in DMSO- <i>d</i> ₆ | 18 |
| Figure S23. ¹ H NMR spectrum of compound 9 in DMSO- <i>d</i> ₆ (500 MHz) | 18 |
| Figure S24. ¹³ C NMR spectrum of compound 9 in DMSO- <i>d</i> ₆ (125 MHz) | 19 |

| | |
|---|----|
| Figure S25. HSQC spectrum of compound 9 in DMSO- <i>d</i> ₆ | 19 |
| Figure S26. COSY spectrum of compound 9 in DMSO- <i>d</i> ₆ | 20 |
| Figure S27. HMBC spectrum of compound 9 in DMSO- <i>d</i> ₆ | 20 |
| Figure S28. NOESY spectrum of compound 9 in DMSO- <i>d</i> ₆ | 21 |
| Figure S29. ¹ H NMR spectrum of compound 10 in DMSO- <i>d</i> ₆ (500 MHz)..... | 21 |
| Figure S30. ¹³ C NMR spectrum of compound 10 in DMSO- <i>d</i> ₆ (125 MHz)..... | 22 |
| Figure S31. HSQC spectrum of compound 10 in DMSO- <i>d</i> ₆ | 22 |
| Figure S32. COSY spectrum of compound 10 in DMSO- <i>d</i> ₆ | 23 |
| Figure S33. HMBC spectrum of compound 10 in DMSO- <i>d</i> ₆ | 23 |
| Figure S34. NOESY spectrum of compound 10 in DMSO- <i>d</i> ₆ | 24 |
| Figure S35. ¹ H NMR spectrum of compound 11 in DMSO- <i>d</i> ₆ (500 MHz)..... | 24 |
| Figure S36. ¹³ C NMR spectrum of compound 11 in DMSO- <i>d</i> ₆ (125 MHz)..... | 25 |
| Figure S37. HSQC spectrum of compound 11 in DMSO- <i>d</i> ₆ | 25 |
| Figure S38. COSY spectrum of compound 11 in DMSO- <i>d</i> ₆ | 26 |
| Figure S39. HMBC spectrum of compound 11 in DMSO- <i>d</i> ₆ | 26 |
| Figure S40. NOESY spectrum of compound 11 in DMSO- <i>d</i> ₆ | 27 |
| Figure S41. ¹ H NMR spectrum of compound 12 in DMSO- <i>d</i> ₆ (500 MHz)..... | 27 |
| Figure S42. ¹³ C NMR spectrum of compound 12 in DMSO- <i>d</i> ₆ (125 MHz)..... | 28 |
| Figure S43. HSQC spectrum of compound 12 in DMSO- <i>d</i> ₆ | 28 |
| Figure S44. COSY spectrum of compound 12 in DMSO- <i>d</i> ₆ | 29 |
| Figure S45. HMBC spectrum of compound 12 in DMSO- <i>d</i> ₆ | 29 |
| Figure S46. NOESY spectrum of compound 12 in DMSO- <i>d</i> ₆ | 30 |
| Figure S47. ¹ H NMR spectrum of compound 13 in DMSO- <i>d</i> ₆ (500 MHz)..... | 30 |
| Figure S48. ¹³ C NMR spectrum of compound 13 in DMSO- <i>d</i> ₆ (125 MHz)..... | 31 |
| Figure S49. HSQC spectrum of compound 13 in DMSO- <i>d</i> ₆ | 31 |
| Figure S50. COSY spectrum of compound 13 in DMSO- <i>d</i> ₆ | 32 |
| Figure S51. HMBC spectrum of compound 13 in DMSO- <i>d</i> ₆ | 32 |
| Figure S52. NOESY spectrum of compound 13 in DMSO- <i>d</i> ₆ | 33 |
| Figure S53. ¹ H NMR spectrum of compound 14 in DMSO- <i>d</i> ₆ (500 MHz)..... | 33 |
| Figure S54. ¹³ C NMR spectrum of compound 14 in DMSO- <i>d</i> ₆ (125 MHz)..... | 34 |
| Figure S55. HSQC spectrum of compound 14 in DMSO- <i>d</i> ₆ | 34 |
| Figure S56. COSY spectrum of compound 14 in DMSO- <i>d</i> ₆ | 35 |

| | |
|---|-----------|
| Figure S57. HMBC spectrum of compound 14 in DMSO- <i>d</i> ₆ | 35 |
| Figure S58. NOESY spectrum of compound 14 in DMSO- <i>d</i> ₆ | 36 |
| Figure S59. Experimental ECD spectra of 2 — 4 in MeOH..... | 37 |
| Figure S60. Experimental ECD spectra of 6 — 9 in MeOH..... | 37 |
| Figure S61. Experimental ECD spectra of 10 — 13 in MeOH..... | 37 |
| Figure S62. Experimental ECD spectra of 13 — 14 in MeOH..... | 37 |
| Figure S63. LC-MS analysis of <i>sasB</i> transformant after incubating with <i>cyclo</i> -Trp-Ala isomers..... | 38 |
| Figure S64. LC-MS analysis of <i>sasB</i> transformant after incubating with <i>cyclo</i> -Trp-Pro isomers..... | 39 |
| Figure S65. Purification of 13 and 14 on the Multohigh Chiral AM-RP column. | 40 |
| Figure S66. LC-MS analysis of 13 after incubation in CD ₃ OD/D ₂ O (1:1) for 14 h. | 41 |
| Reference | 42 |

Supplementary Tables

Table S1. Comparison of *sas* genes with known entries



| Protein | Accession No. | Length (aa) | Sequence identity (length in aa) in % | |
|---------|----------------|-------------|---|---|
| | | | <i>S. leeuwenhoekii</i> NRRL B-24963 ¹ | <i>S. youssoufiensis</i> OUC6819 ² |
| SasA | WP_079103588.1 | 252 | SazA (252) 82 | DmtA1 (233), 41 |
| SasB | WP_121505431.1 | 347 | SazB-PT domain (314) 85 | DmtC1 (311), 38 |
| SasC | WP_055513754.1 | 290 | SazB-MT domain (276) 82 | |

Table S2. Bacterial strains used in this study

| Strain | Source | Cultivation media |
|---|------------|-------------------|
| <i>E. coli</i> DH5α | Invitrogen | LB |
| <i>E. coli</i> ET12567/pUZ8002 | 3 | LB |
| <i>Streptomyces albus</i> J1074 | 4 | MS |
| <i>Streptomyces aurantiacus</i> NRRL ISP-5412 | NRRL | modified R5 |

NRRL: ARS Culture Collection

LB medium: tryptone 10.0 g/L, yeast extract 5.0 g/L, NaCl 10.0 g/L.

MS medium: mannitol 20.0 g/L, soya flour 20.0 g/L, agar 20.0 g/L.

Modified R5 medium: sucrose 103.0 g/L, glucose 10.0 g/L, yeast extract 5.0 g/L, MgCl₂·6H₂O 10.12 g/L, K₂SO₄ 0.25 g/L, Difco casaminoacids 0.1 g/L, MOPS 21.0 g/L, trace element solution 2 mL/L, pH 7.2.

Table S3. Cloning and expression constructs used in this study

| Gene | Primer sequences (5'-3') | Cloning construct | Expression vector | Cloning sites | Expression constructs |
|---------------|---|-------------------|-------------------|---------------------------------------|-----------------------|
| <i>sasA</i> | <u>CATATG</u> TCCAGCAAGGACGTCGAC TCTAGACTATGTGCGGTGACTTCCTTC | pJL81 | pPWW50 | <i>NdeI/XbaI</i> | pJL87 |
| <i>sasABC</i> | <u>CATATG</u> TCCAGCAAGGACGTCGAC ACTAGTCTGCGTTCACCGGGTCG | pJL82 | pPWW50 | <i>NdeI/SpeI</i> | pJL88 |
| <i>sasAB</i> | <u>CATATG</u> TCCAGCAAGGACGTCGAC ACTAGTTCACCGGTCCGTCTCCGC | pJL83 | pPWW50 | <i>NdeI/SpeI</i> | pJL89 |
| <i>sasAC</i> | <u>CATATG</u> TCCAGCAAGGACGTCGAC <u>AGATCT</u> TGTCAGCAAGGACGTCGAC <u>ACTAGT</u> ATGTATCAGTCCGGGACCCGTTTC | pJL84 pJL86 | pPWW50 | <i>NdeI/BglII</i> <i>SpeI/XbaI</i> | pJL90 |
| <i>sasB</i> | TCTAGATCACCGGGTCGGACCGCTG <u>CATATG</u> AGCCAGCGAAGTCAACCG ACTAGTTCACCGGTCCGTCTCCGC | pJL85 | pPWW50 | <i>NdeI/SpeI</i> | pJL91 |
| <i>sasC</i> | <u>ACTAGT</u> ATGTATCAGTCCGGGACCCGTTTC C TCTAGATCACCGGGTCGGACCGCTG | pJL86 | pPWW50 | <i>SpeI/XbaI</i> | pJL92 |

Restriction sites for cloning are underlined in the primer sequences. Cloning constructs are based on pGEM T EASY vector.

Table S4. ^1H NMR data of compounds **2** – **4** in $\text{DMSO-}d_6$

| Position | 2 | 3 | 4 |
|-------------|--|--|--|
| | δ_{H} , multi. (<i>J</i> in Hz) | δ_{H} , multi. (<i>J</i> in Hz) | δ_{H} , multi. (<i>J</i> in Hz) |
| 1 | | 6.53, d (3.0) | 6.49, d (3.4) |
| 2 | 5.10, s | 5.28, d (3.0) | 5.28, d (3.4) |
| 4 | 7.11, dd (7.4, 0.9) | 7.05, d (7.4) | 7.08, d (7.3) |
| 5 | 6.70, td (7.4, 0.9) | 6.59, td (7.4, 1.0) | 6.59, t (7.3) |
| 6 | 7.08, td (7.8, 1.3) | 6.98, td (7.8, 1.0) | 6.94, td (7.8, 0.9) |
| 7 | 6.47, d (7.8) | 6.49, d (7.8) | 6.44, d (7.8) |
| 10 α | 2.38, dd (12.9, 7.2) | 2.33, dd (13.0, 7.7) | 2.41, dd (13.1, 7.7) |
| 10 β | 1.96, dd (12.9, 11.1) | 1.98, dd (13.0, 10.0) | 1.99, dd (13.1, 10.7) |
| 11 | 4.76, dd (11.1, 7.2) | 4.50, t (9.0) | 4.66, dd (10.7, 7.7) |
| 14 | 4.76, dd (11.1, 7.2) | 4.37, t (5.5) | 4.66, dd (10.7, 7.7) |
| 15 | | 7.59, s | |
| 17 | | 10.81, s | |
| 18 | 5.10, s | 7.22, d (2.3) | 5.28, d (3.4) |
| 20 | 7.11, dd (7.4, 0.9) | 7.57, d (7.8) | 7.08, d (7.3) |
| 21 | 6.70, td (7.4, 0.9) | 6.95, td (7.8, 1.0) | 6.59, t (7.3) |
| 22 | 7.08, td (7.8, 1.3) | 7.07, td (8.1, 1.0) | 6.94, td (7.8, 0.9) |
| 23 | 6.47, d (7.8) | 7.32, d (8.1) | 6.44, d (7.8) |
| 26 α | 2.38, dd (12.9, 7.2) | 3.37, dd (15.2, 5.5) | 2.41, dd (13.1, 7.7) |
| 26 β | 1.96, dd (12.9, 11.2) | 3.00, dd (15.2, 6.7) | 1.99, dd (13.1, 10.7) |
| 1' / 1'' | 2.54, d (7.4) | 2.45, d (7.1) | 2.51, d (7.3) |
| 2' / 2'' | 4.99, br t (7.3) | 5.02, br t (7.3) | 5.03, t (7.1) |
| 4' / 4'' | 1.61, s | 1.58, s | 1.60, s |
| 5' / 5'' | 1.64, s | 1.64, s | 1.64, s |
| 6' / 6'' | 2.87, s | | |

The NMR data of **2** – **4** correspond well to those reported previously.¹

Supplementary Figures

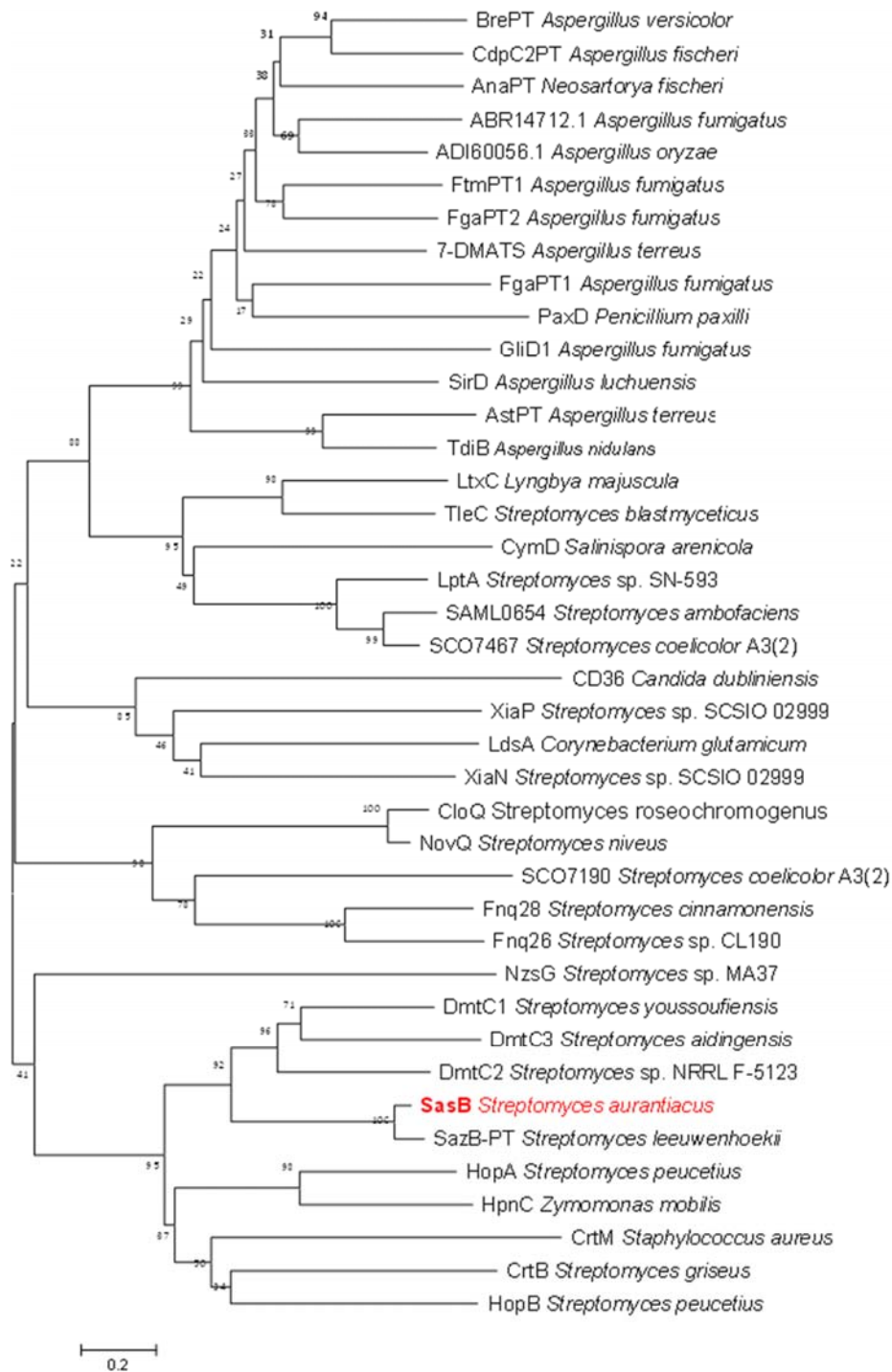


Figure S1. Phylogenetic analysis of PTs

Investigated in this study (in bold red) and functionally characterized PTs from bacteria and fungi. The protein sequences were downloaded from NCBI database.

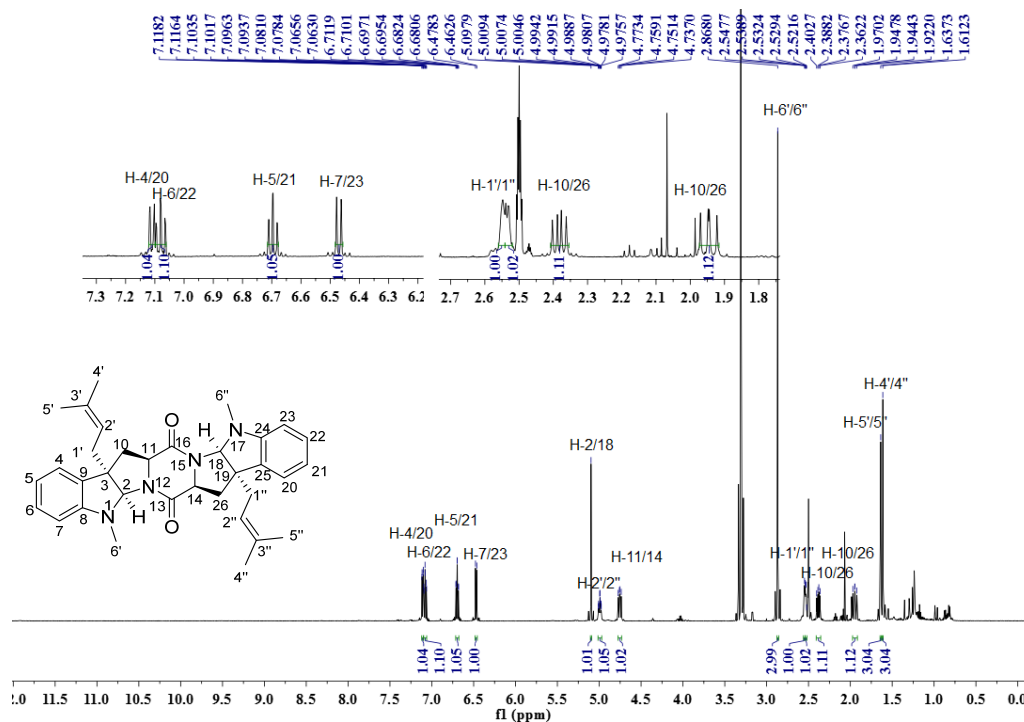


Figure S2. ^1H NMR spectrum of compound **2** in $\text{DMSO-}d_6$ (500 MHz)

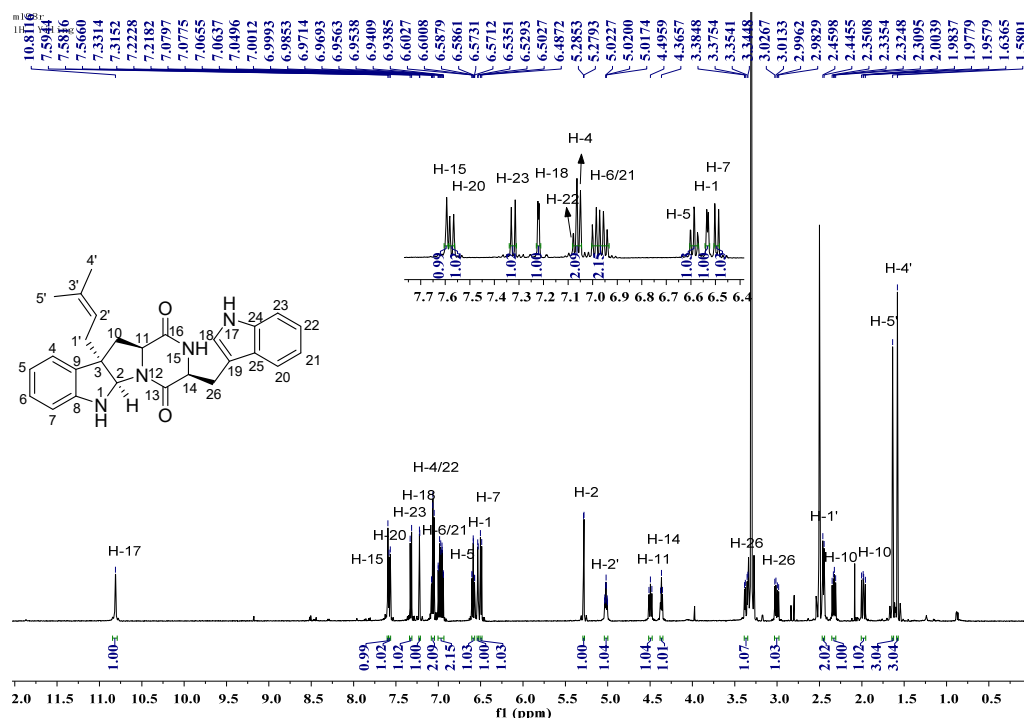
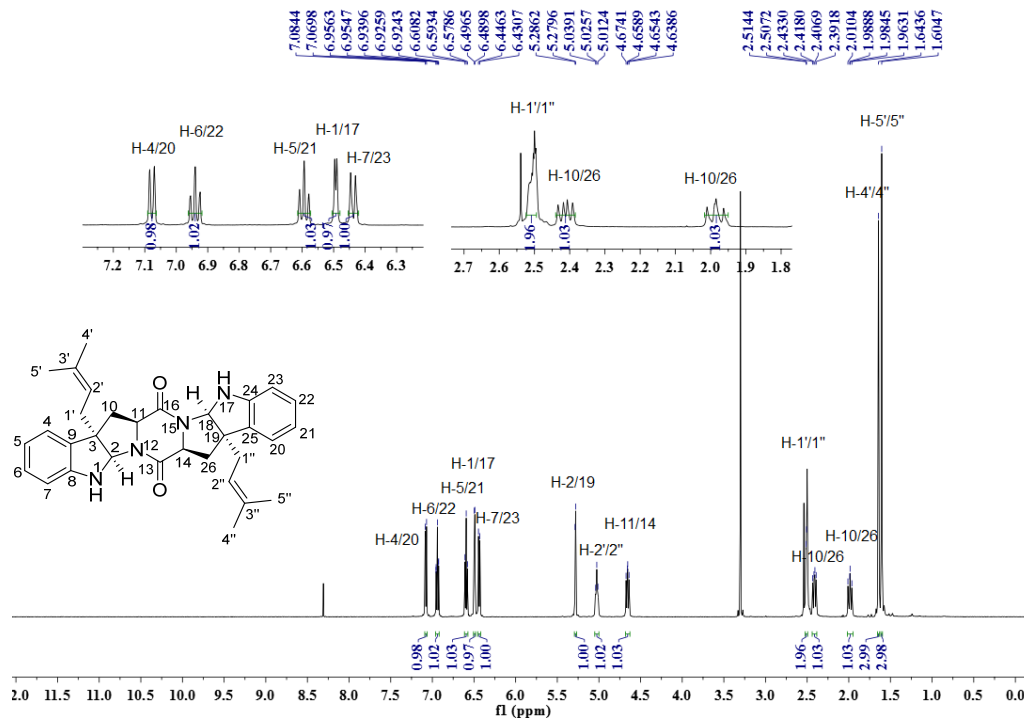


Figure S3. ^1H NMR spectrum of compound **3** in $\text{DMSO-}d_6$ (500 MHz)



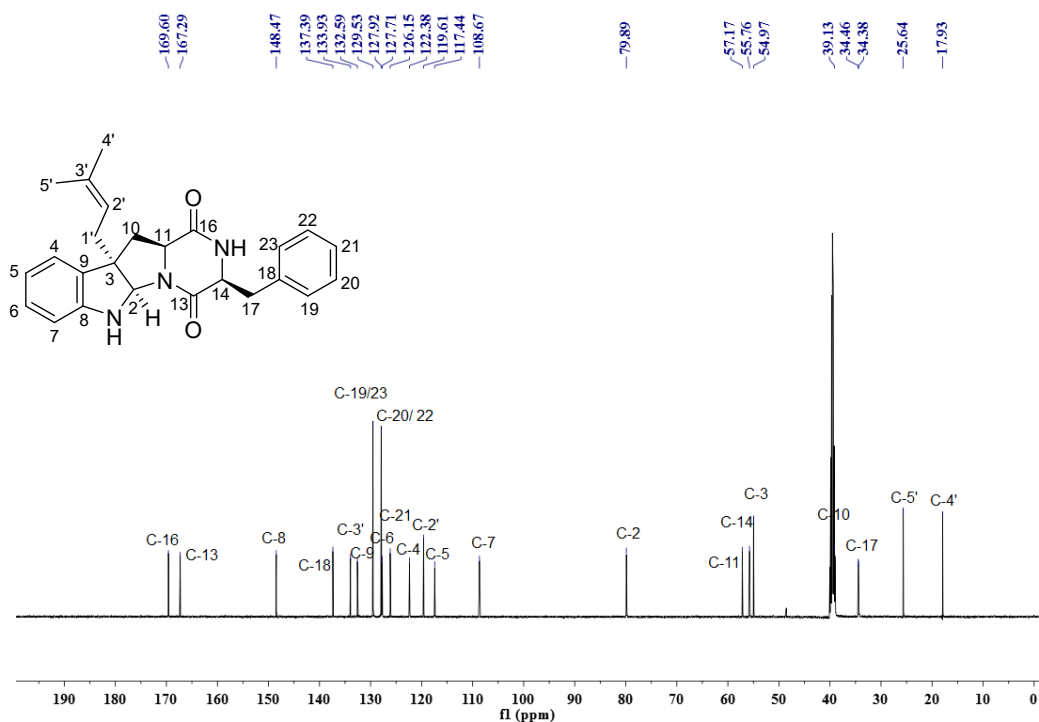


Figure S6. ^{13}C NMR spectrum of compound **6** in $\text{DMSO-}d_6$ (125 MHz)

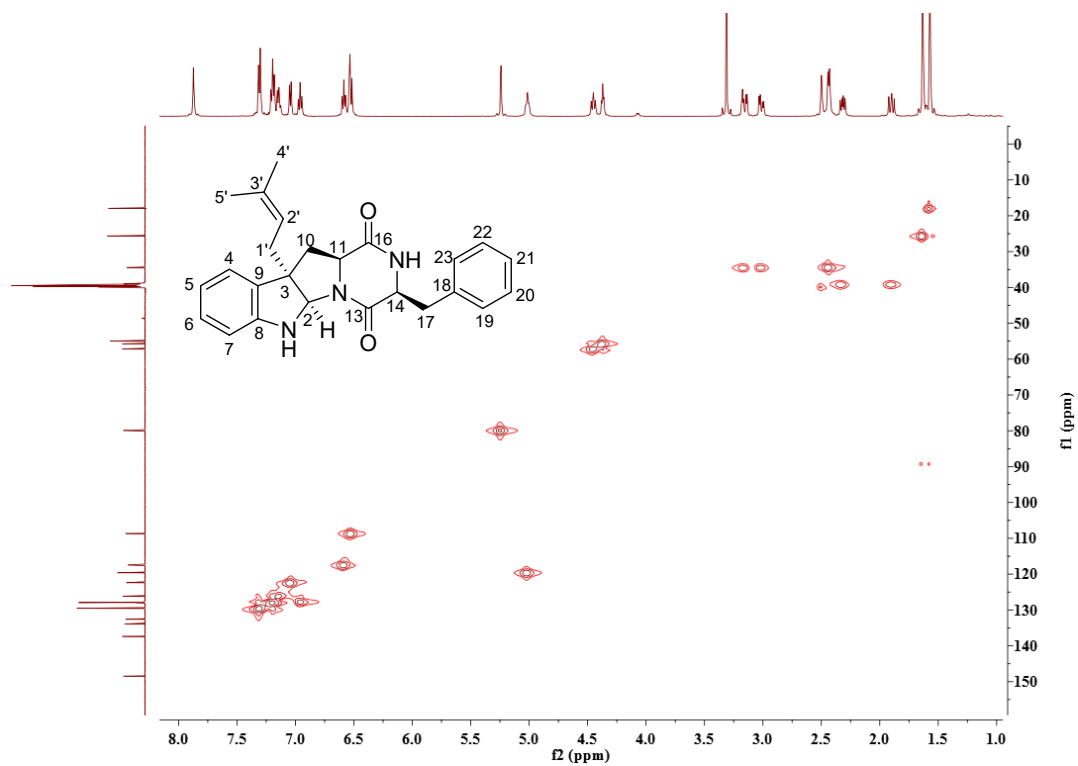


Figure S7. HSQC spectrum of compound **6** in $\text{DMSO-}d_6$

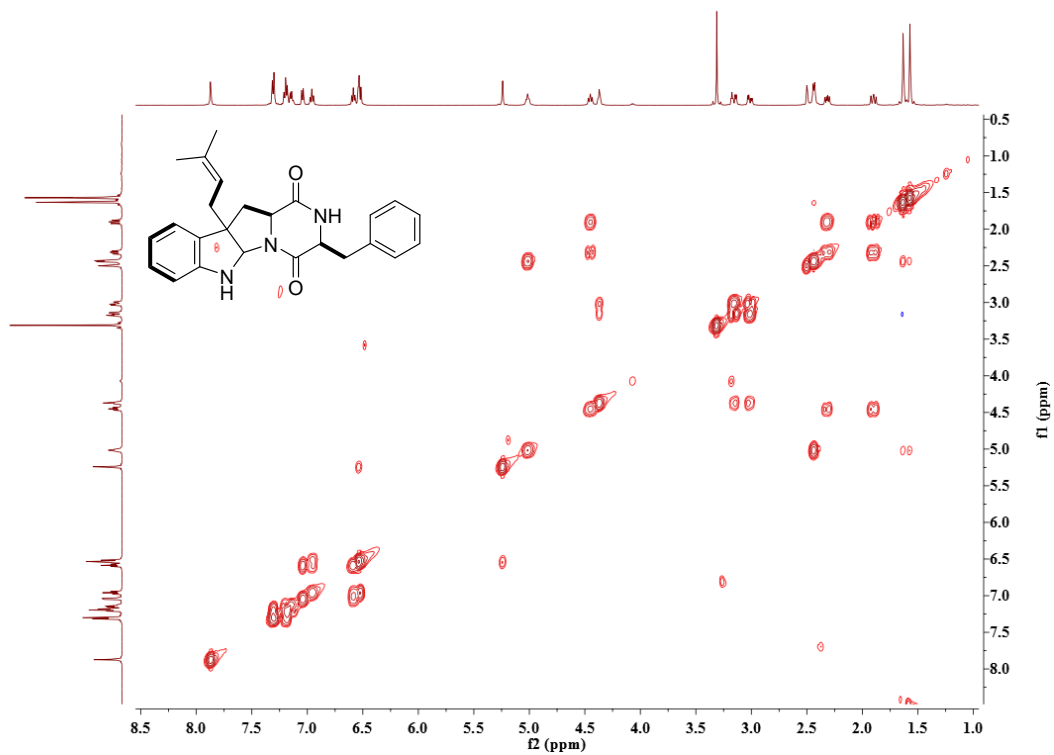


Figure S8. COSY spectrum of compound **6** in DMSO-*d*₆

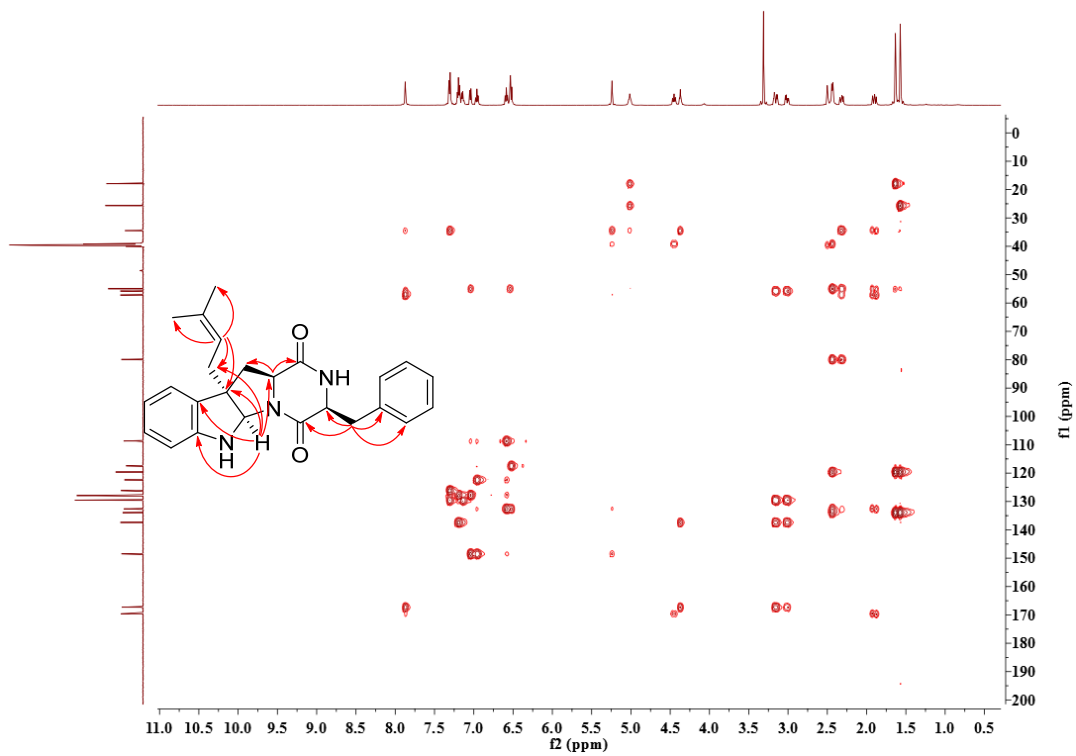


Figure S9. HMBC spectrum of compound **6** in DMSO-*d*₆

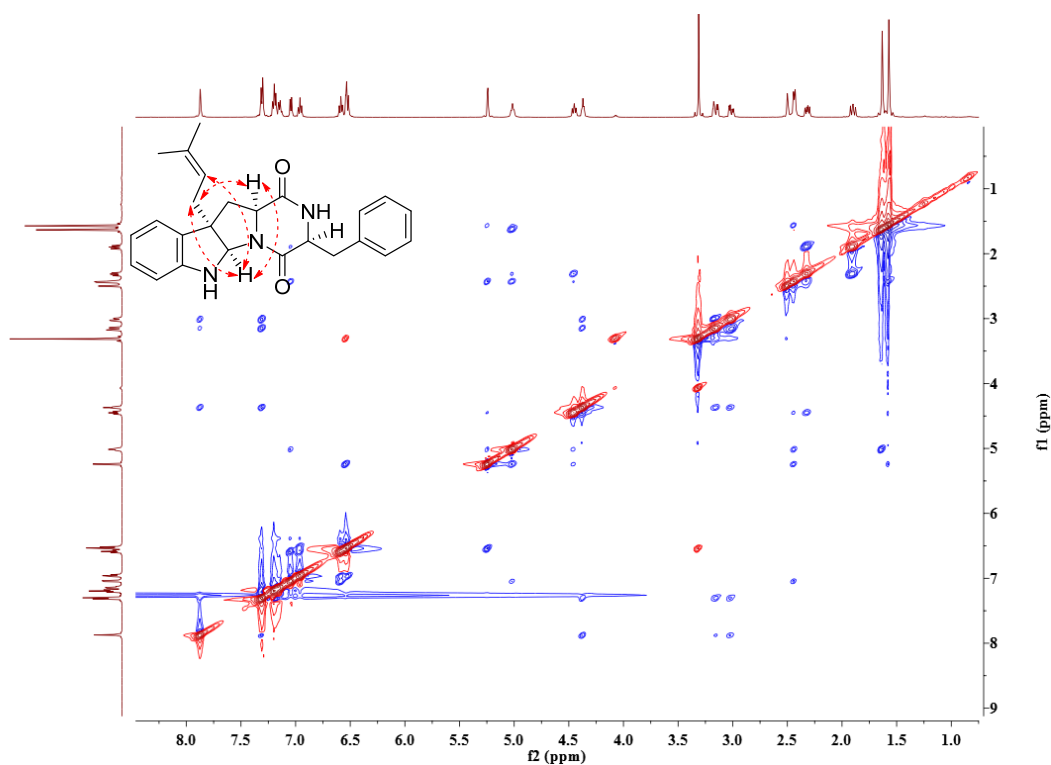


Figure S10. NOESY spectrum of compound 6 in DMSO-*d*₆

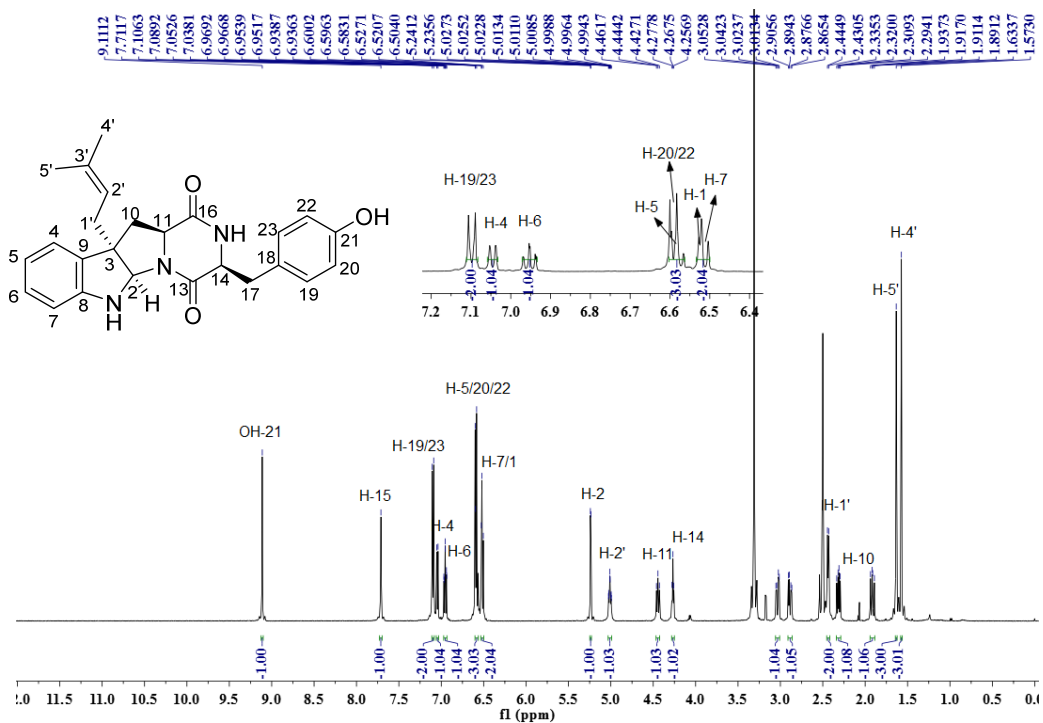


Figure S11. ¹H NMR spectrum of compound 7 in DMSO-*d*₆ (500 MHz)

S12

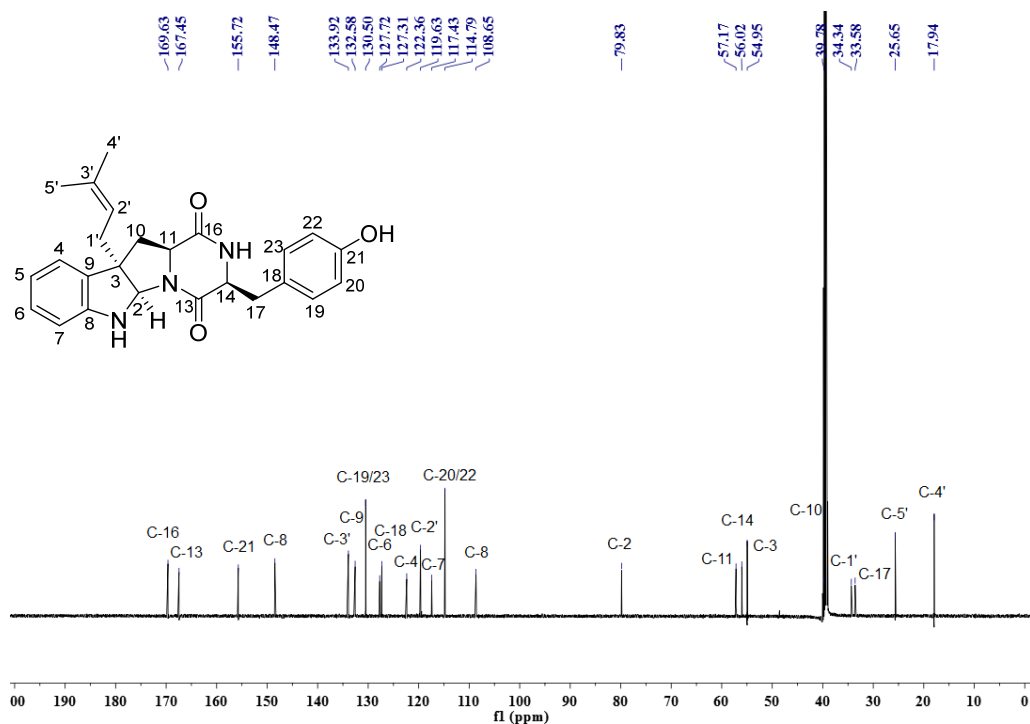


Figure S12. ^{13}C NMR spectrum of compound 7 in $\text{DMSO-}d_6$ (125 MHz)

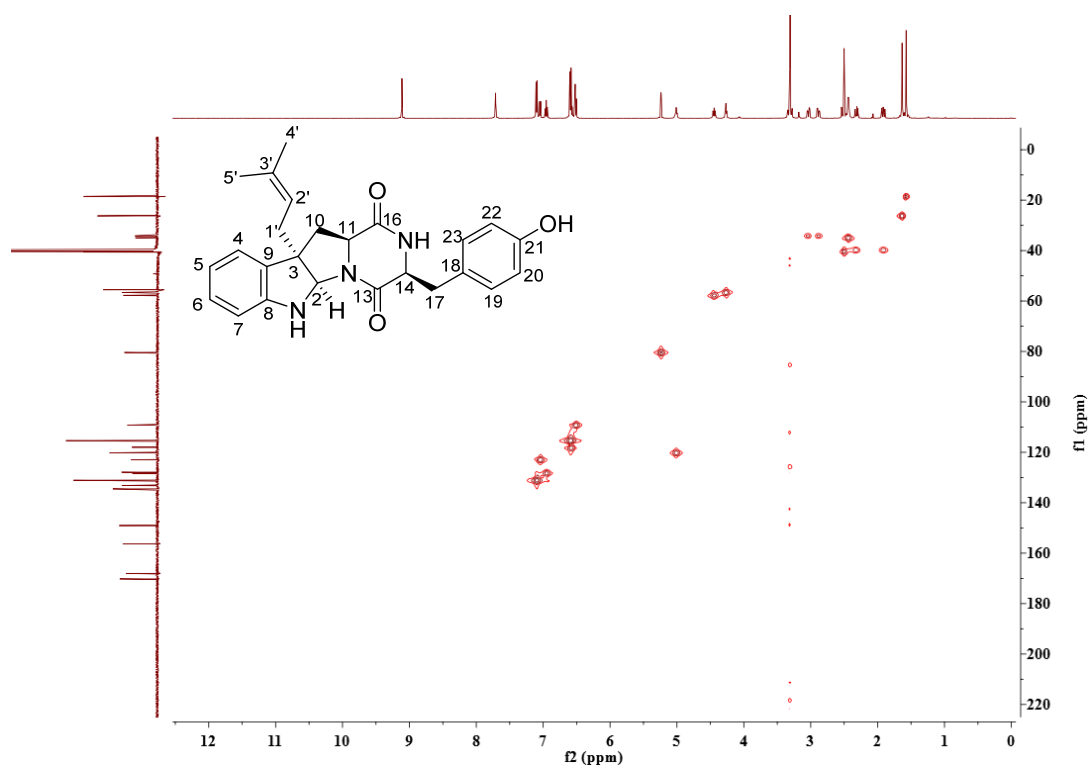


Figure S13. HSQC spectrum of compound 7 in $\text{DMSO-}d_6$

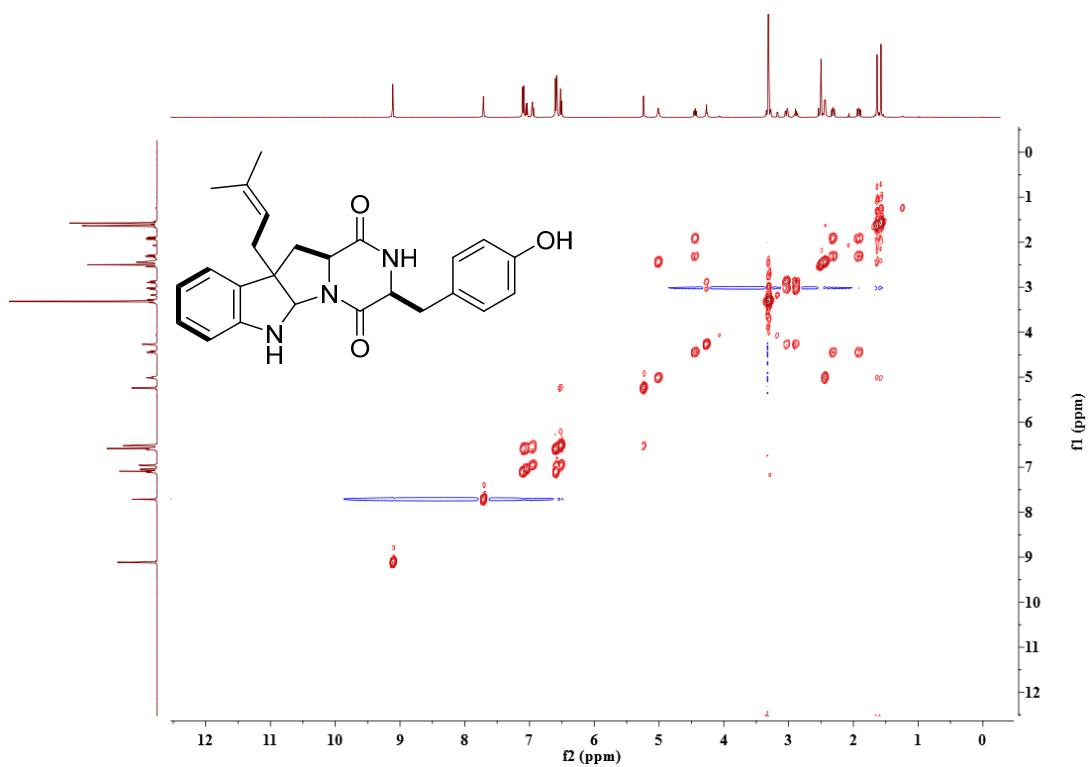


Figure S14. COSY spectrum of compound **7** in DMSO-*d*₆

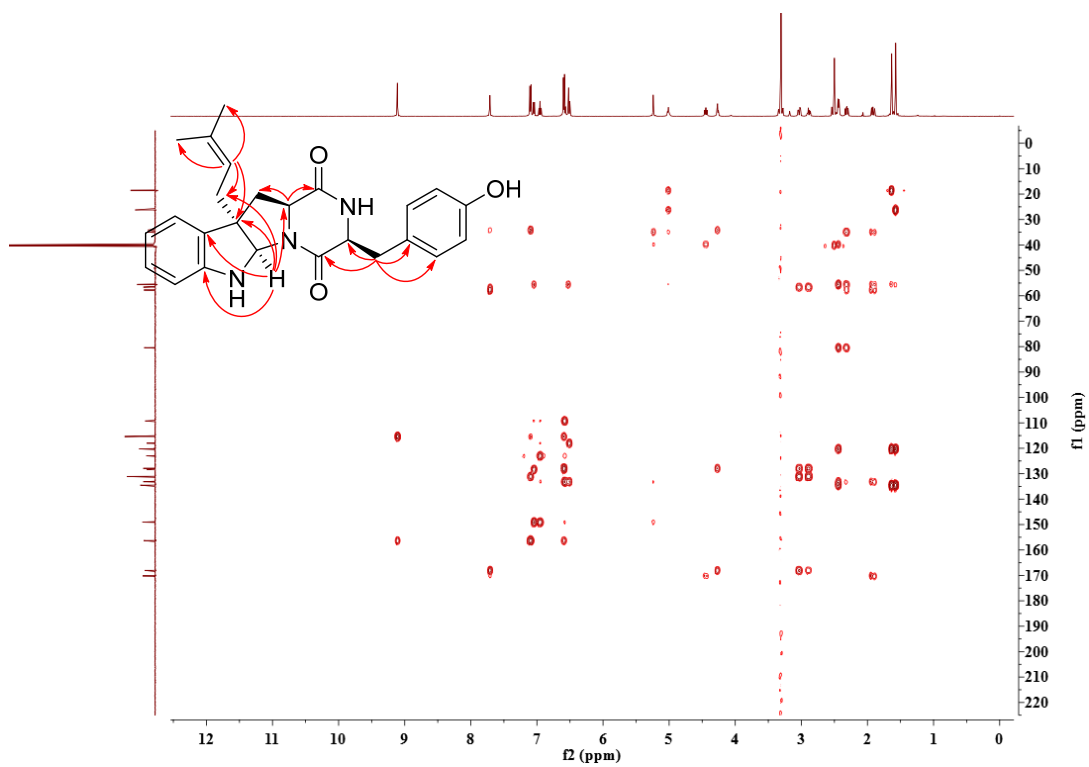


Figure S15. HMBC spectrum of compound **7** in DMSO-*d*₆

S14

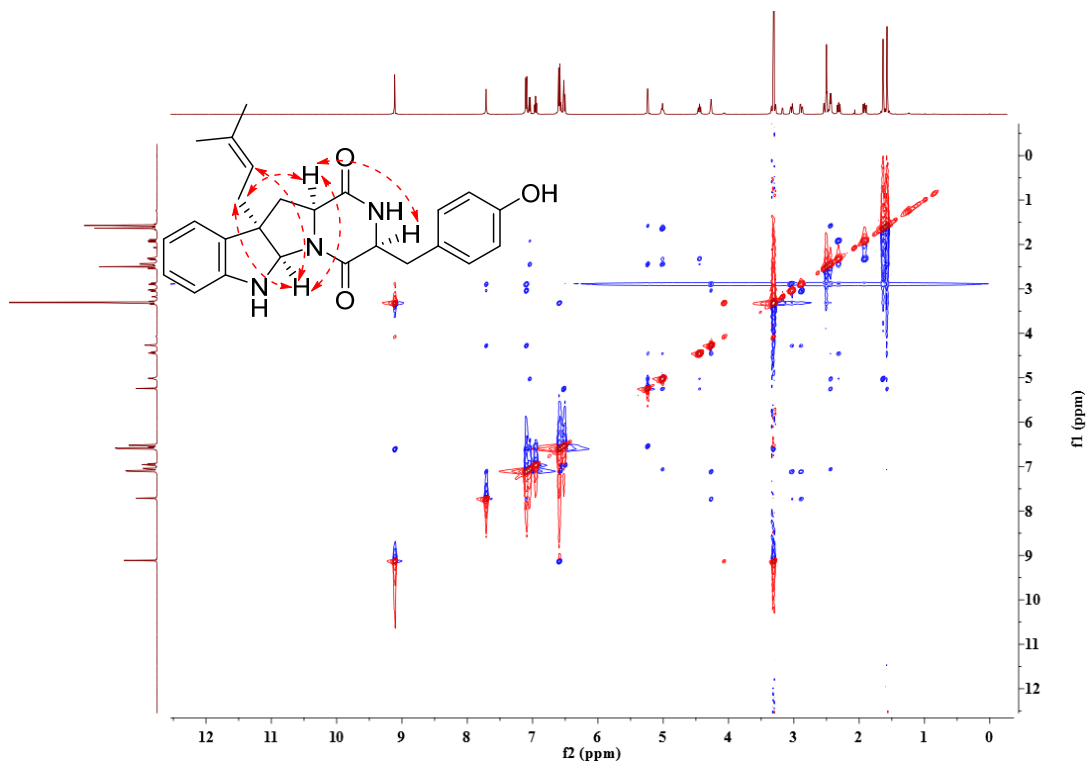


Figure S16. NOESY spectrum of compound **7** in DMSO-*d*₆

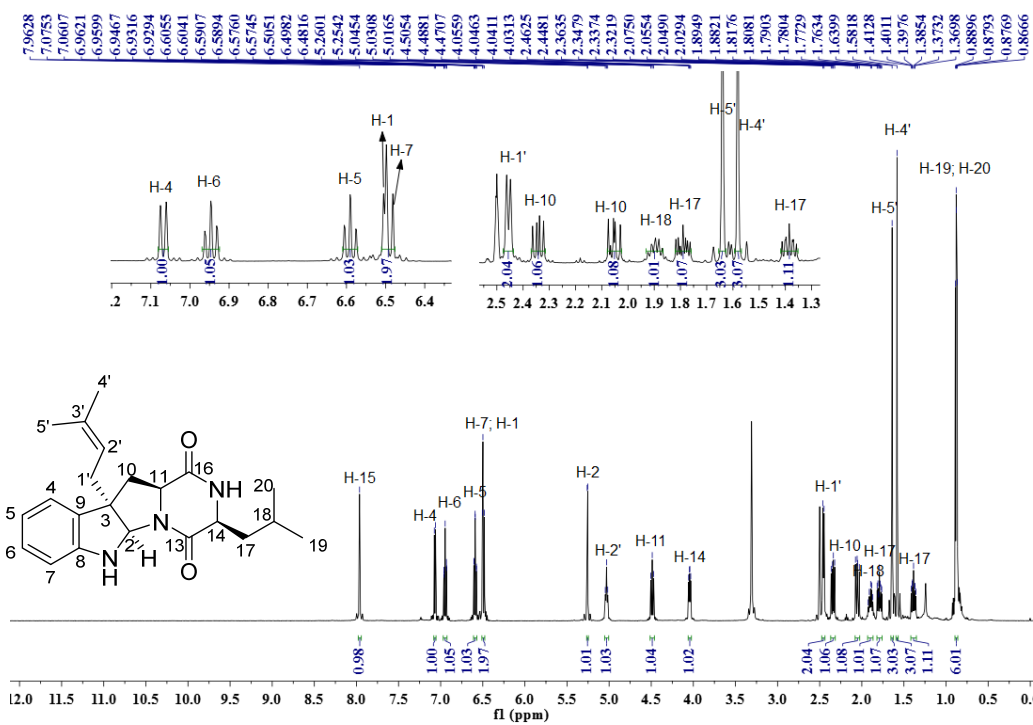


Figure S17. ¹H NMR spectrum of compound **8** in DMSO-*d*₆ (500 MHz)

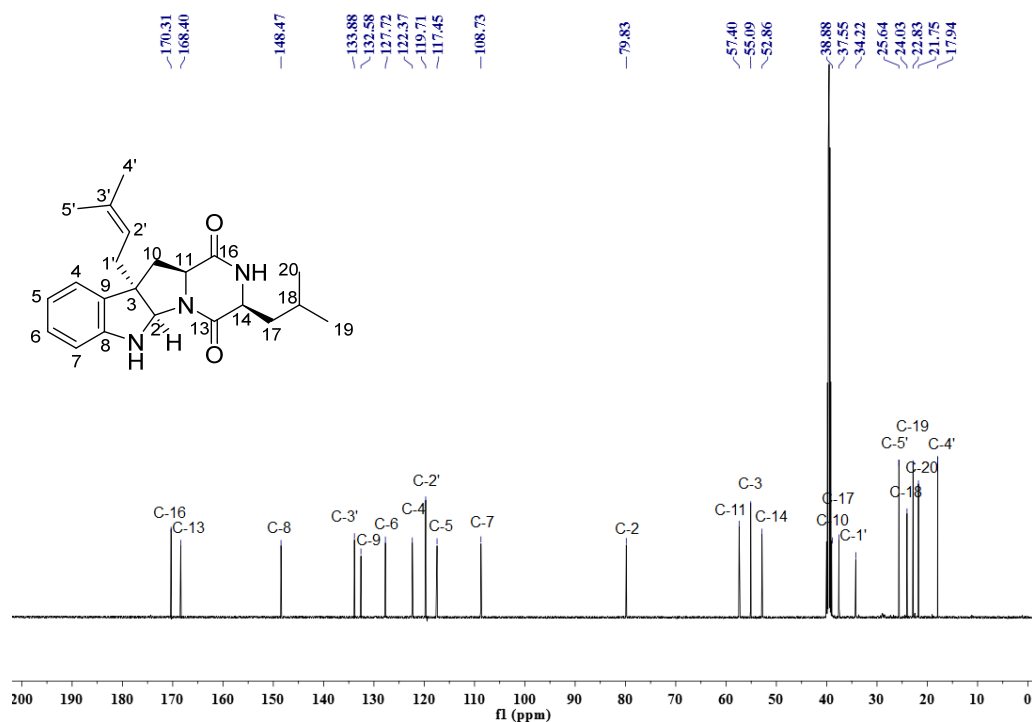


Figure S18. ^{13}C NMR spectrum of compound **8** in $\text{DMSO-}d_6$ (125 MHz)

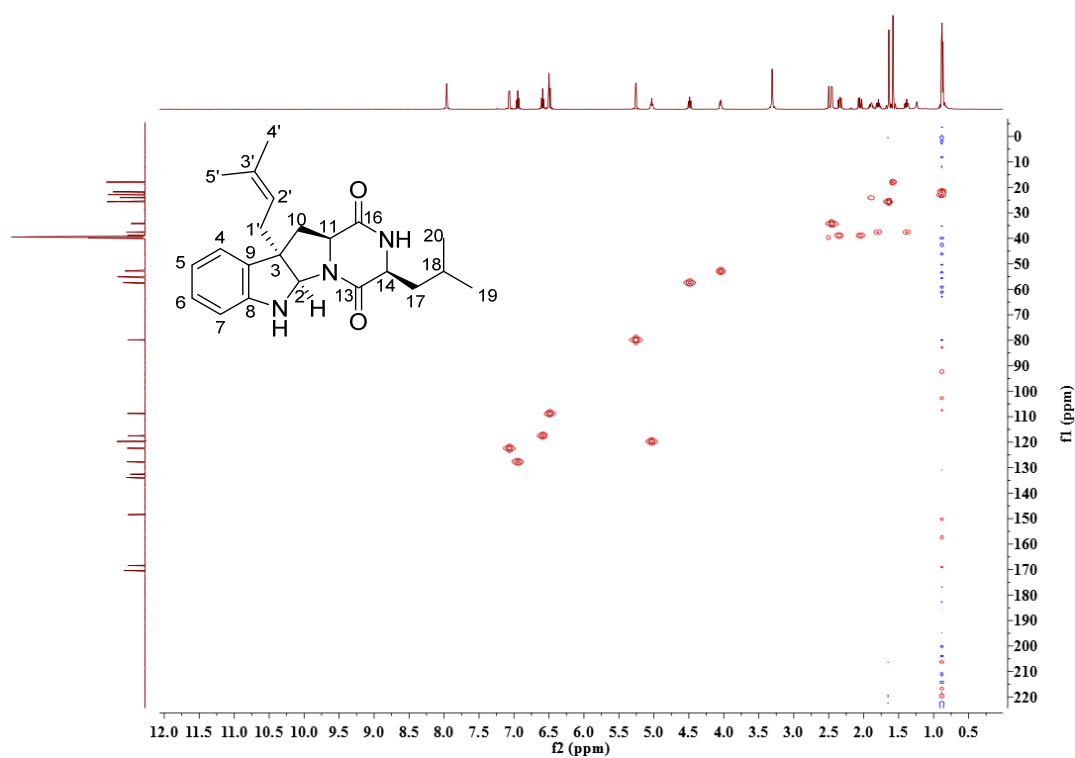


Figure S19. HSQC spectrum of compound **8** in $\text{DMSO-}d_6$

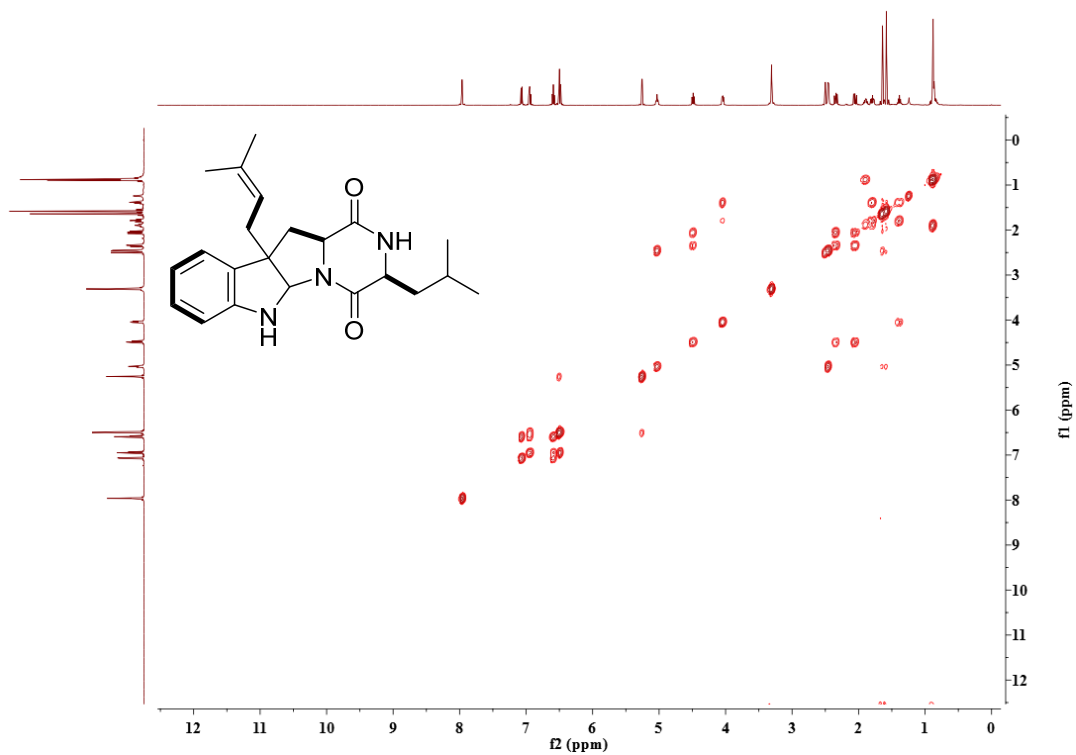


Figure S20. COSY spectrum of compound **8** in DMSO-*d*₆

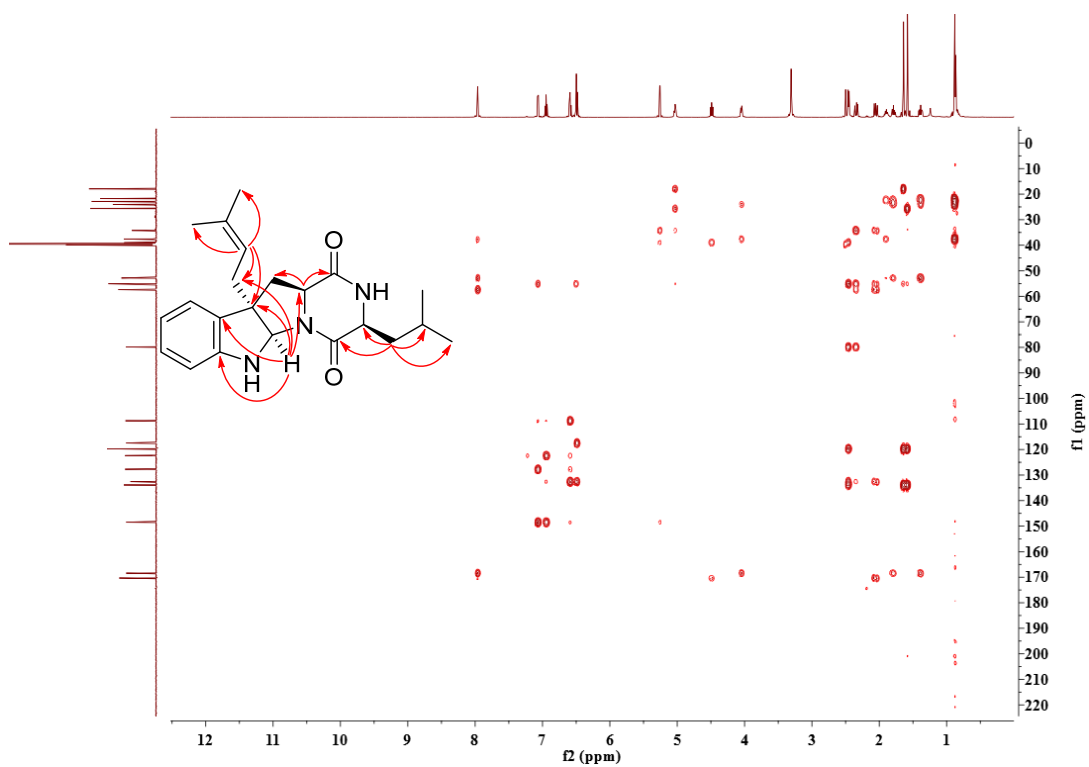


Figure S21. HMBC spectrum of compound **8** in DMSO-*d*₆

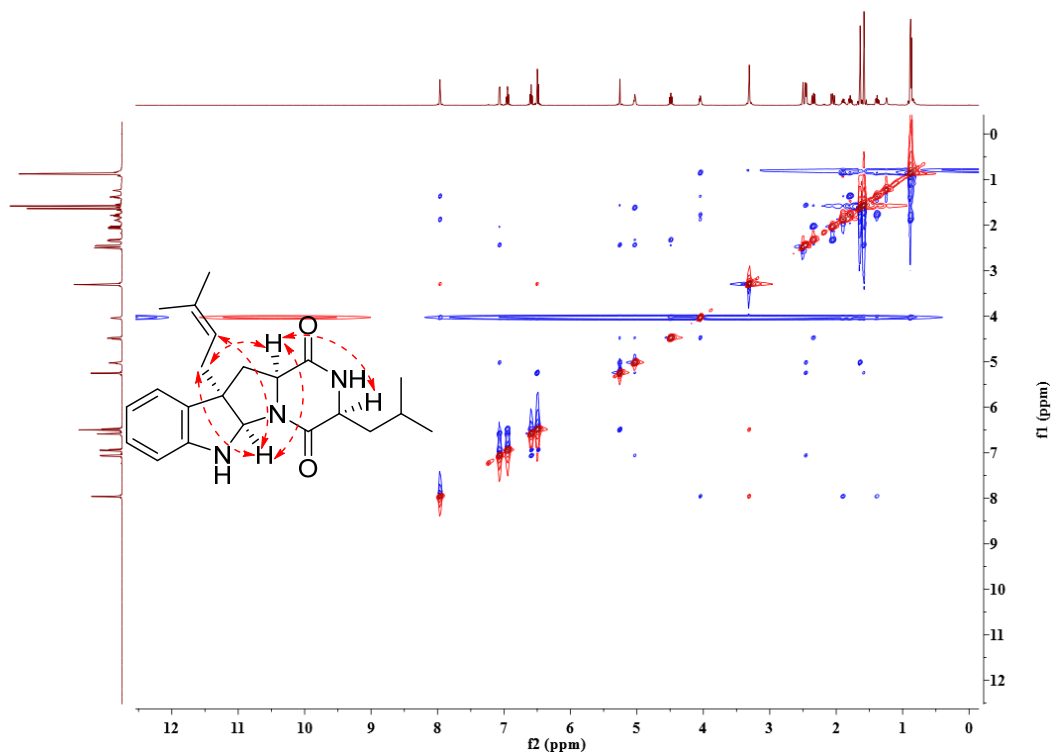


Figure S22. NOESY spectrum of compound **8** in DMSO-*d*₆

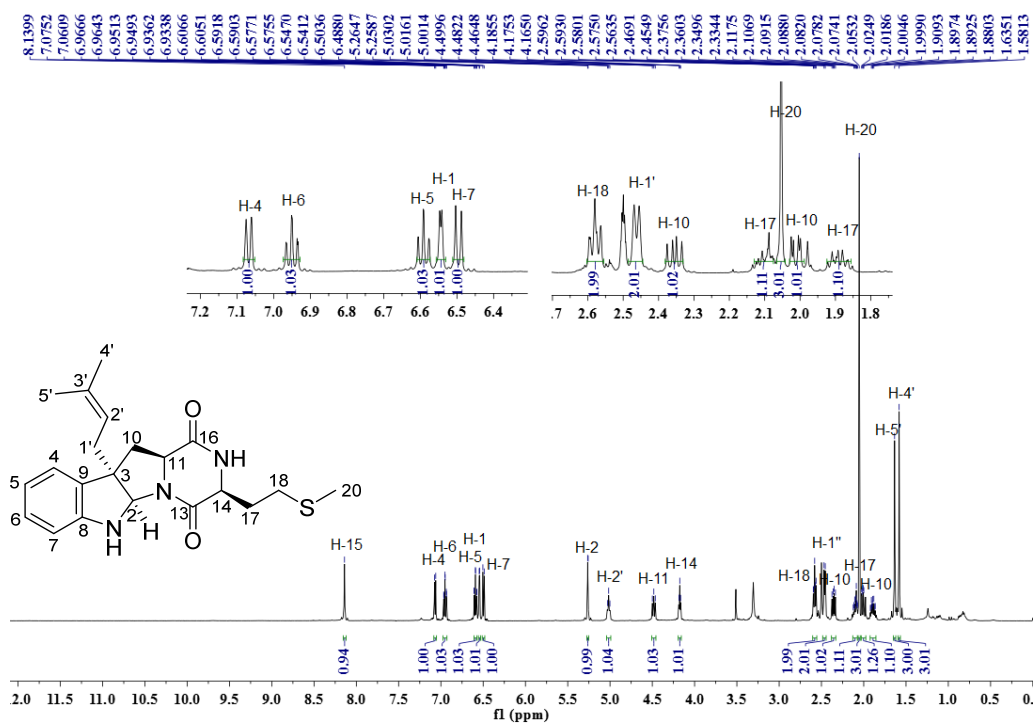


Figure S23. ¹H NMR spectrum of compound **9** in DMSO-*d*₆ (500 MHz)

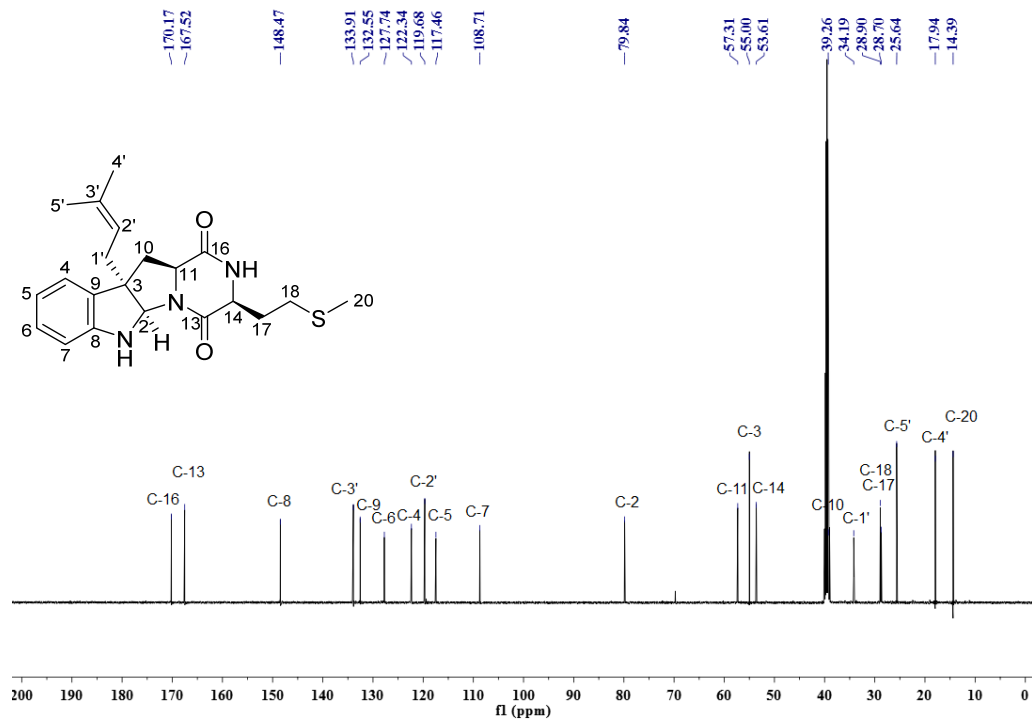


Figure S24. ^{13}C NMR spectrum of compound **9** in $\text{DMSO-}d_6$ (125 MHz)

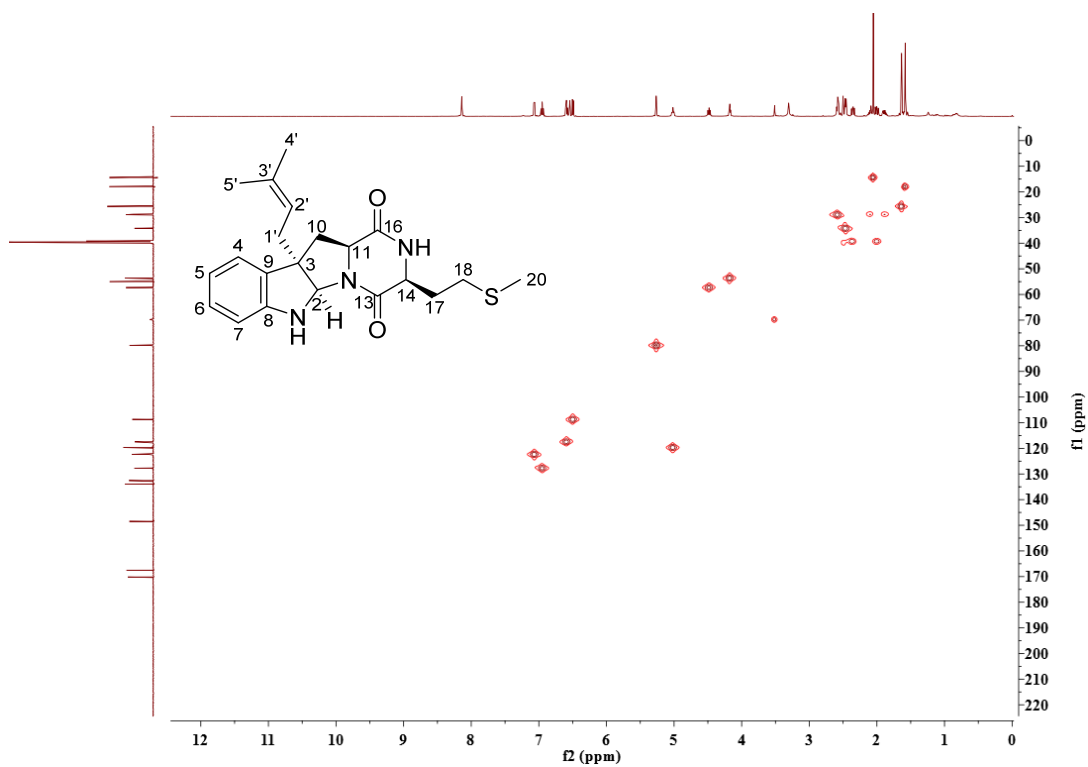


Figure S25. HSQC spectrum of compound **9** in $\text{DMSO-}d_6$

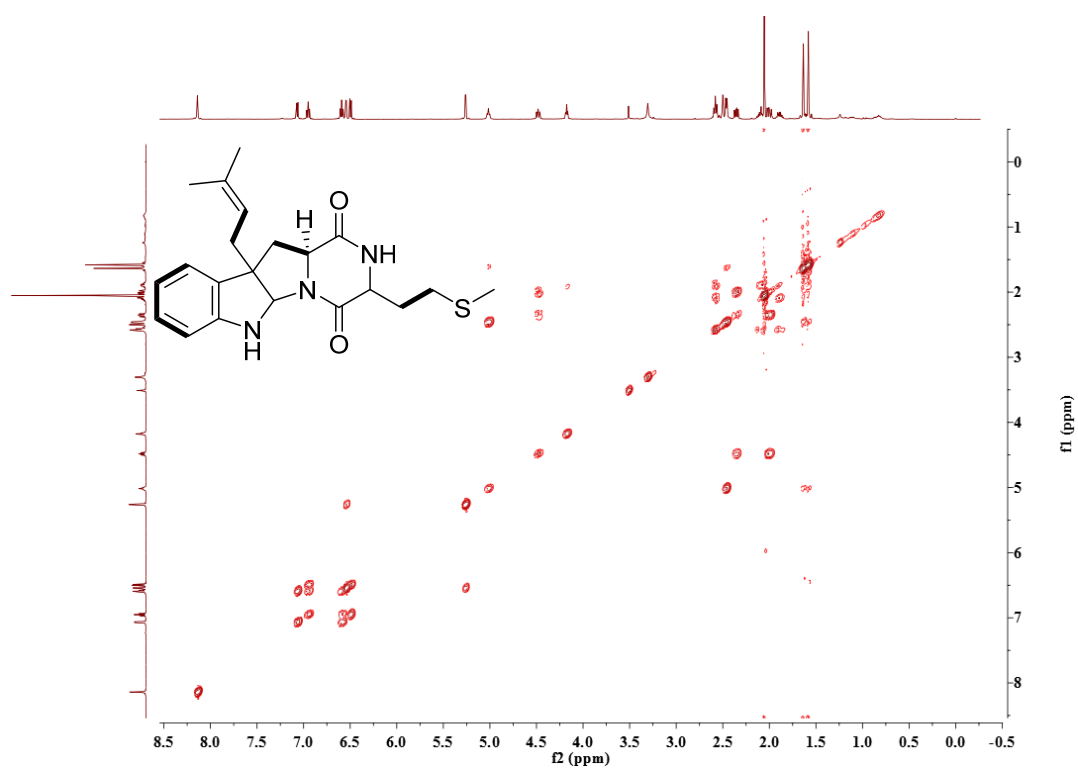


Figure S26. COSY spectrum of compound **9** in DMSO- d_6 .

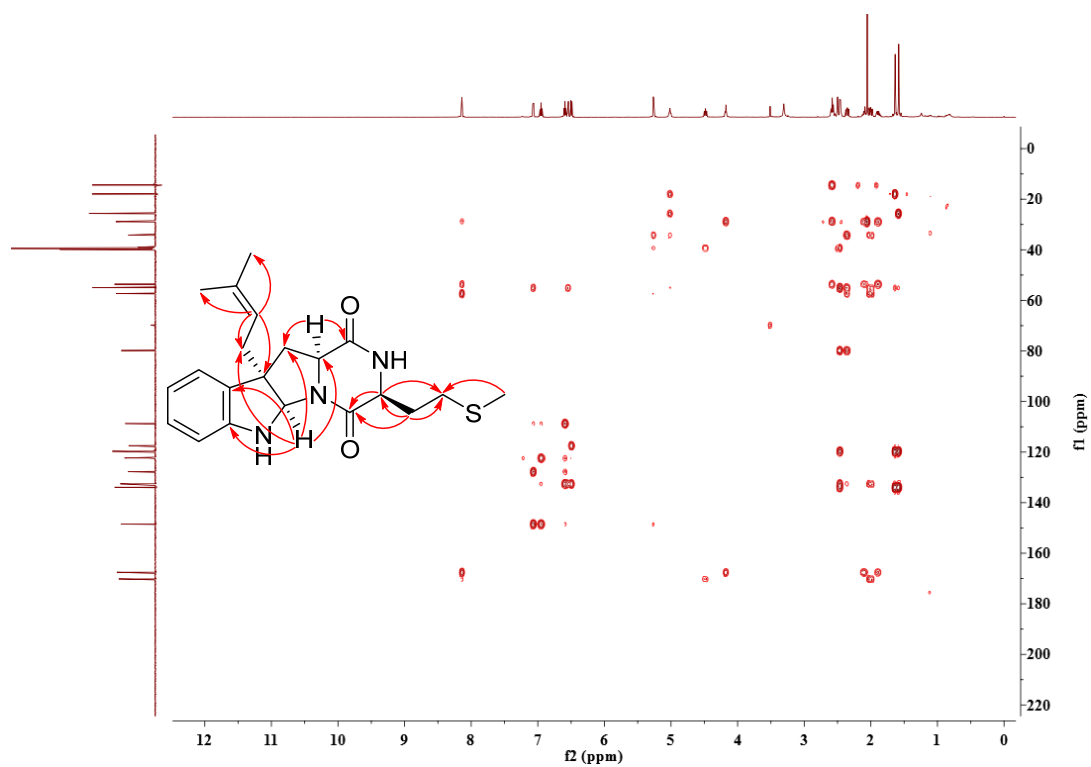


Figure S27. HMBC spectrum of compound **9** in DMSO- d_6 .

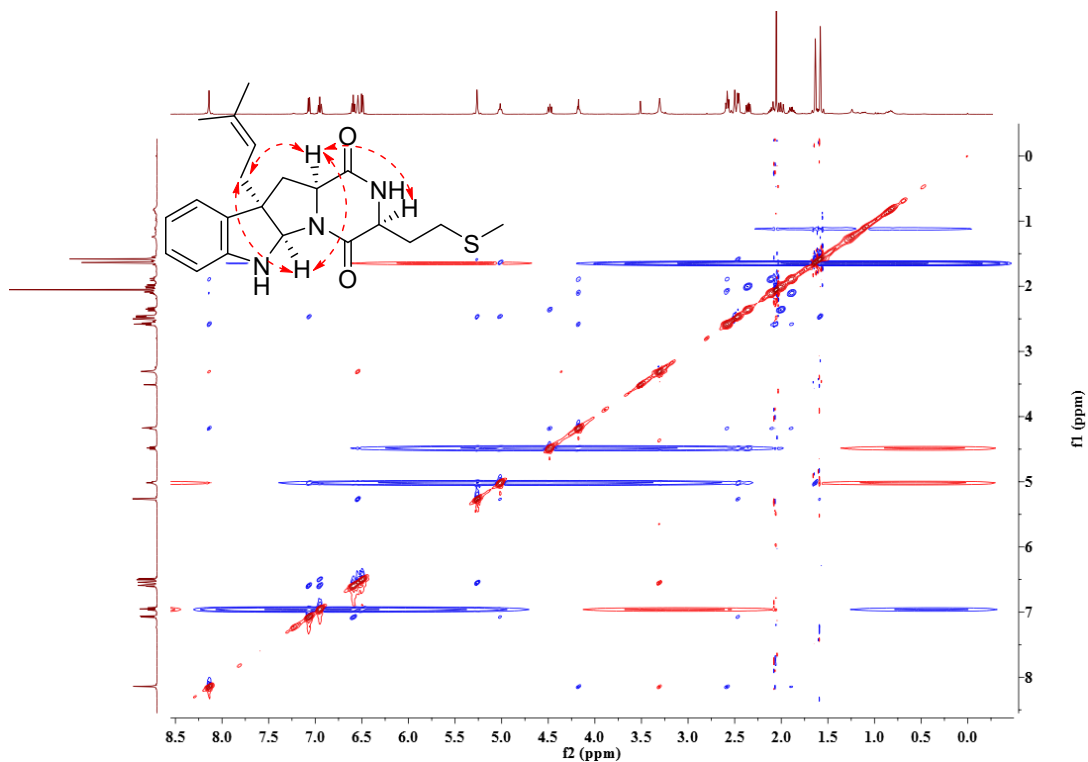


Figure S28. NOESY spectrum of compound **9** in DMSO- d_6

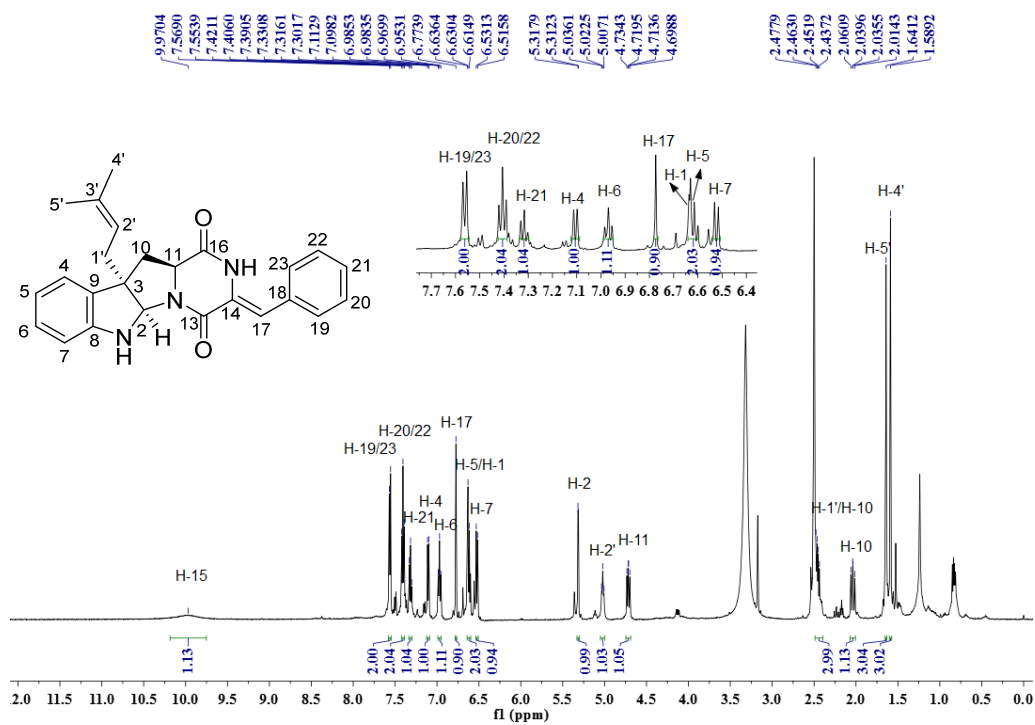


Figure S29. ^1H NMR spectrum of compound **10** in DMSO- d_6 (500 MHz)

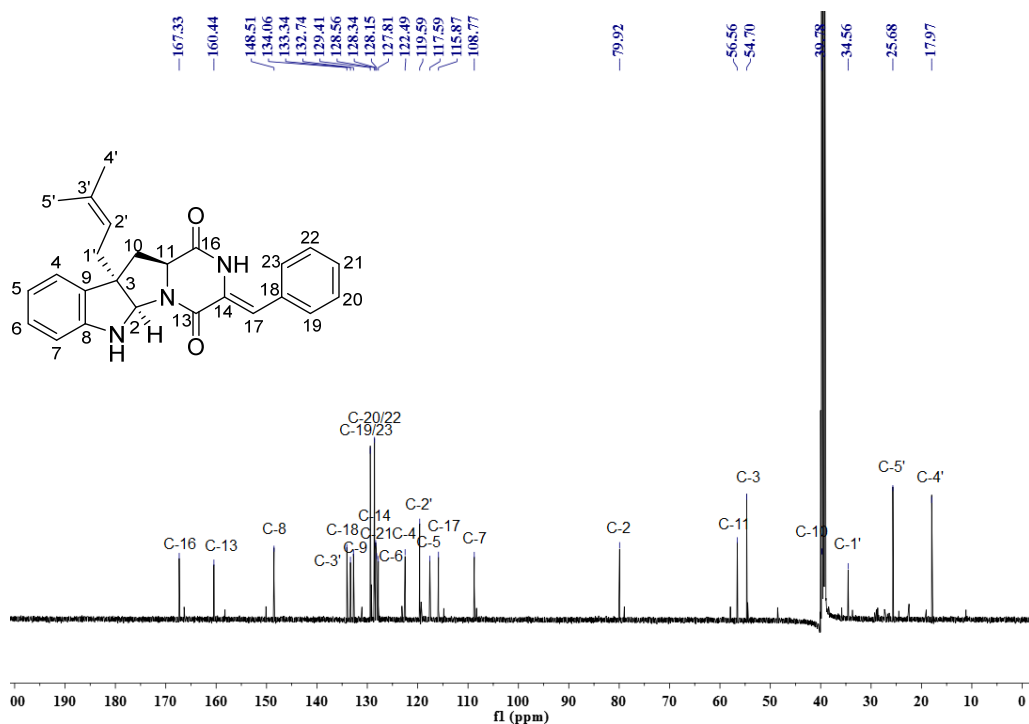


Figure S30. ^{13}C NMR spectrum of compound **10** in $\text{DMSO-}d_6$ (125 MHz)

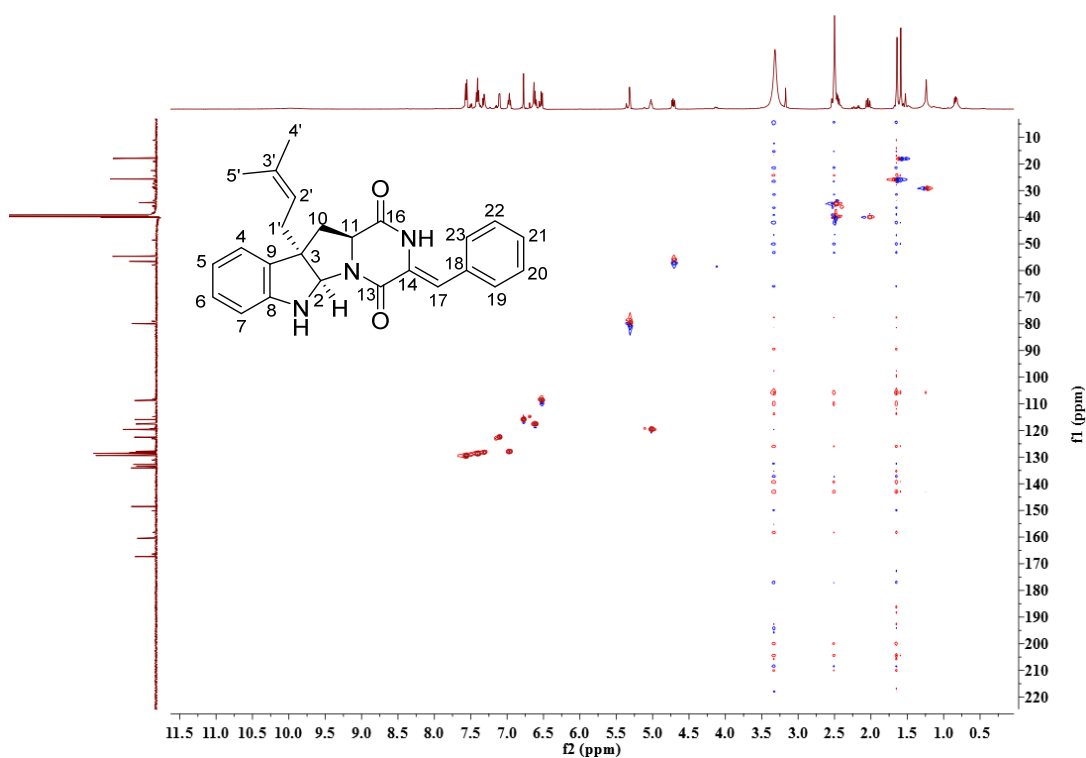


Figure S31. HSQC spectrum of compound **10** in $\text{DMSO-}d_6$

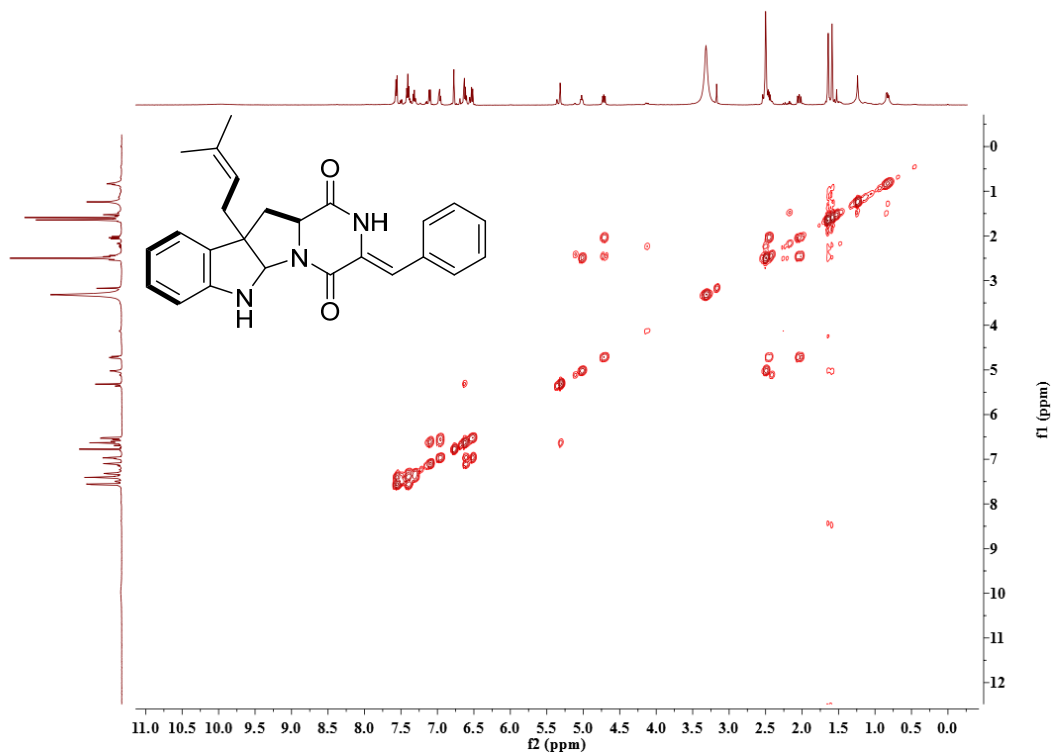


Figure S32. COSY spectrum of compound **10** in DMSO- d_6

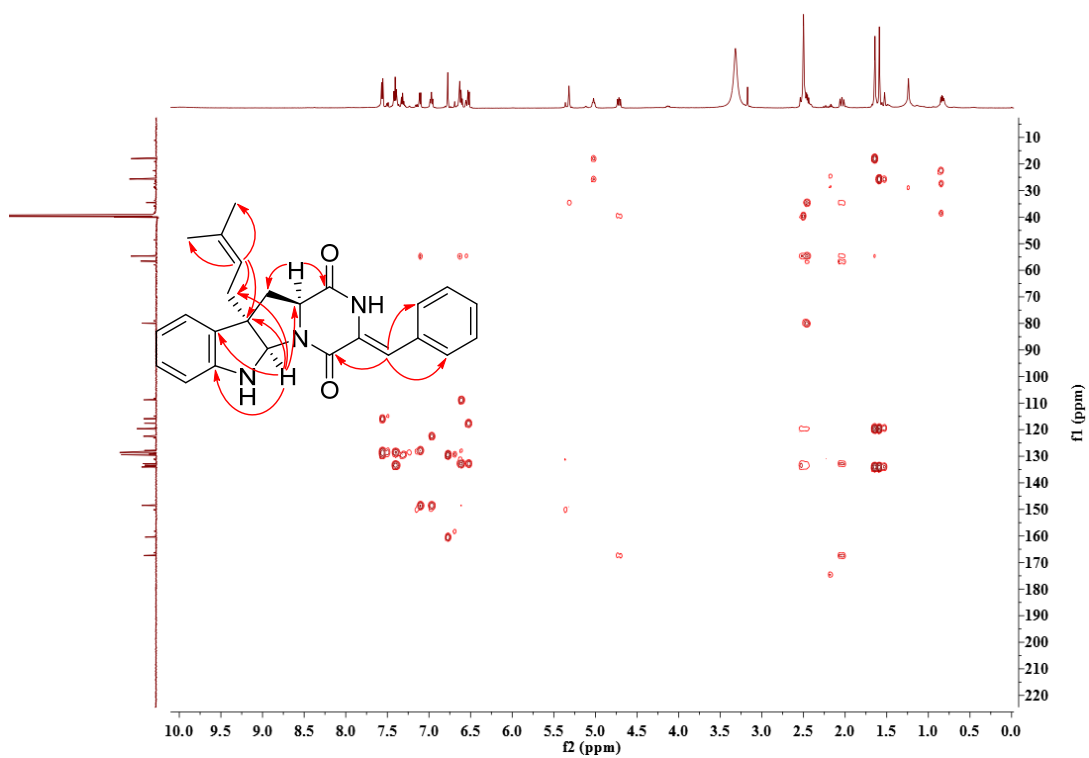


Figure S33. HMBC spectrum of compound **10** in DMSO- d_6

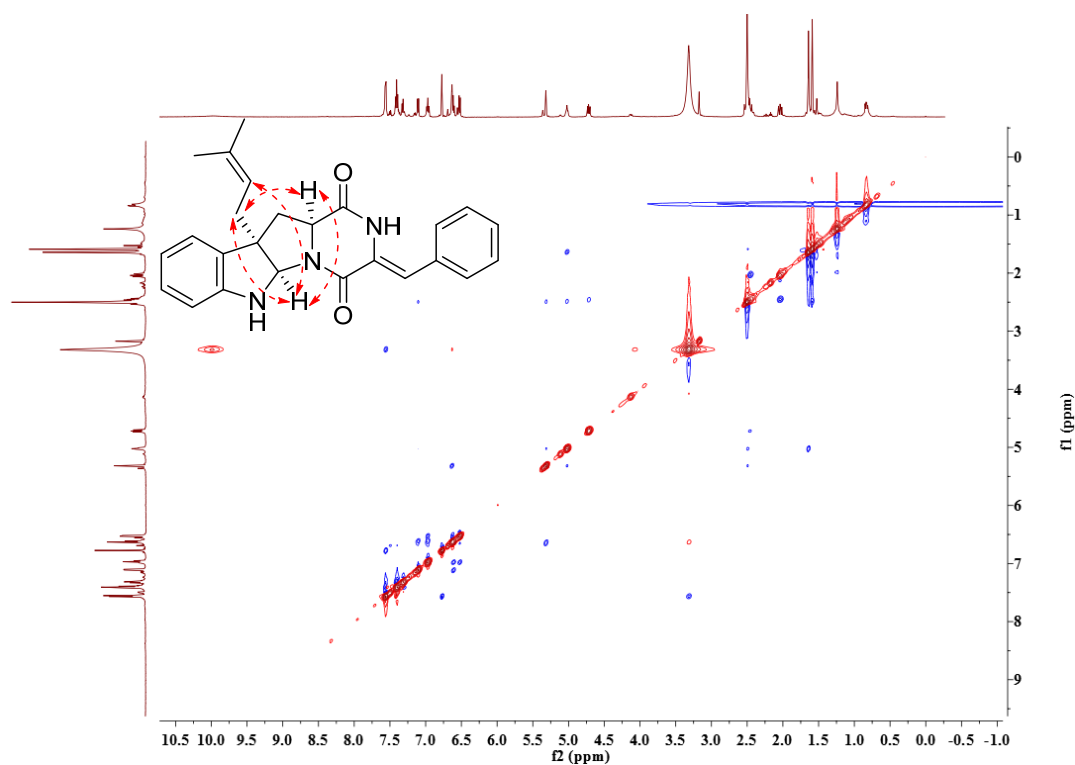


Figure S34. NOESY spectrum of compound **10** in DMSO-*d*₆

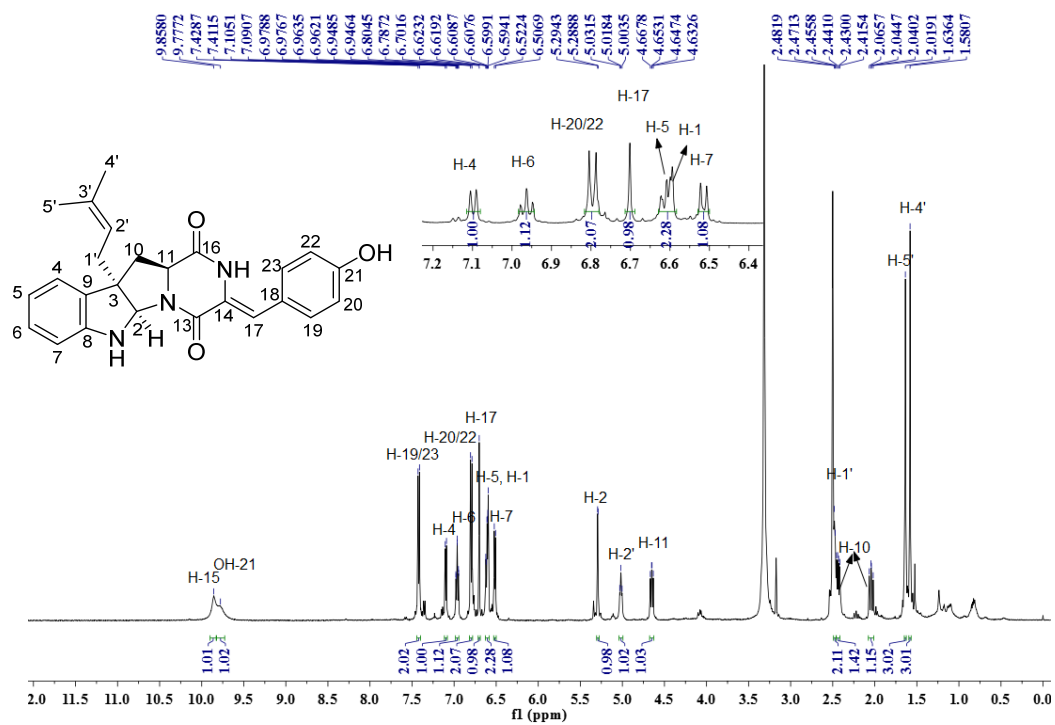


Figure S35. ¹H NMR spectrum of compound **11** in DMSO-*d*₆ (500 MHz)

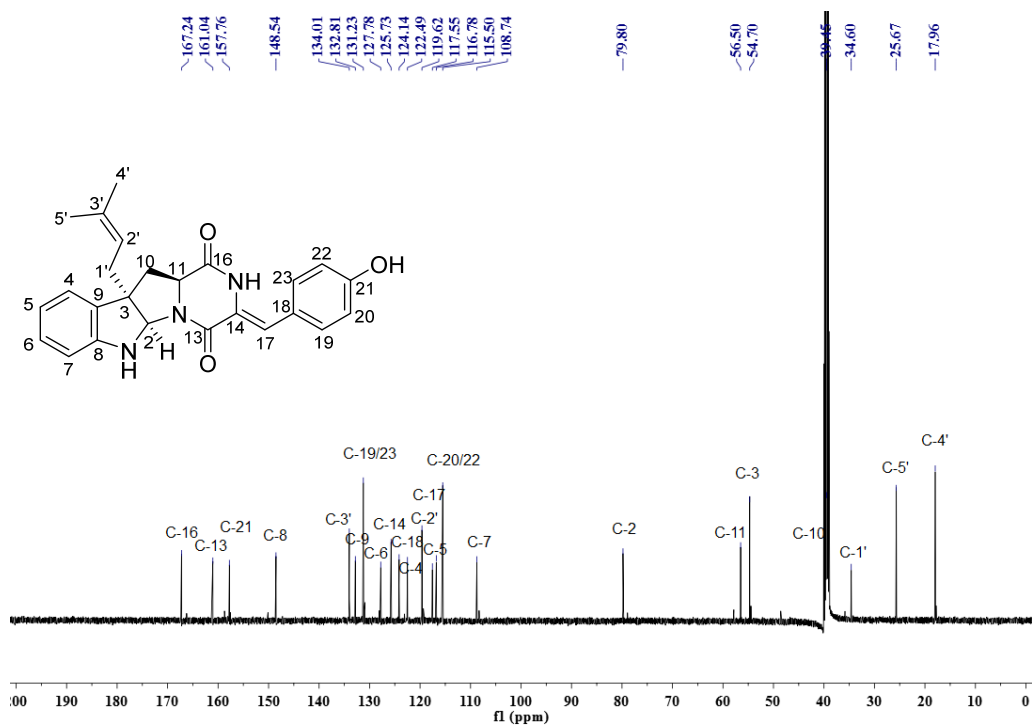


Figure S36. ^{13}C NMR spectrum of compound **11** in $\text{DMSO-}d_6$ (125 MHz)

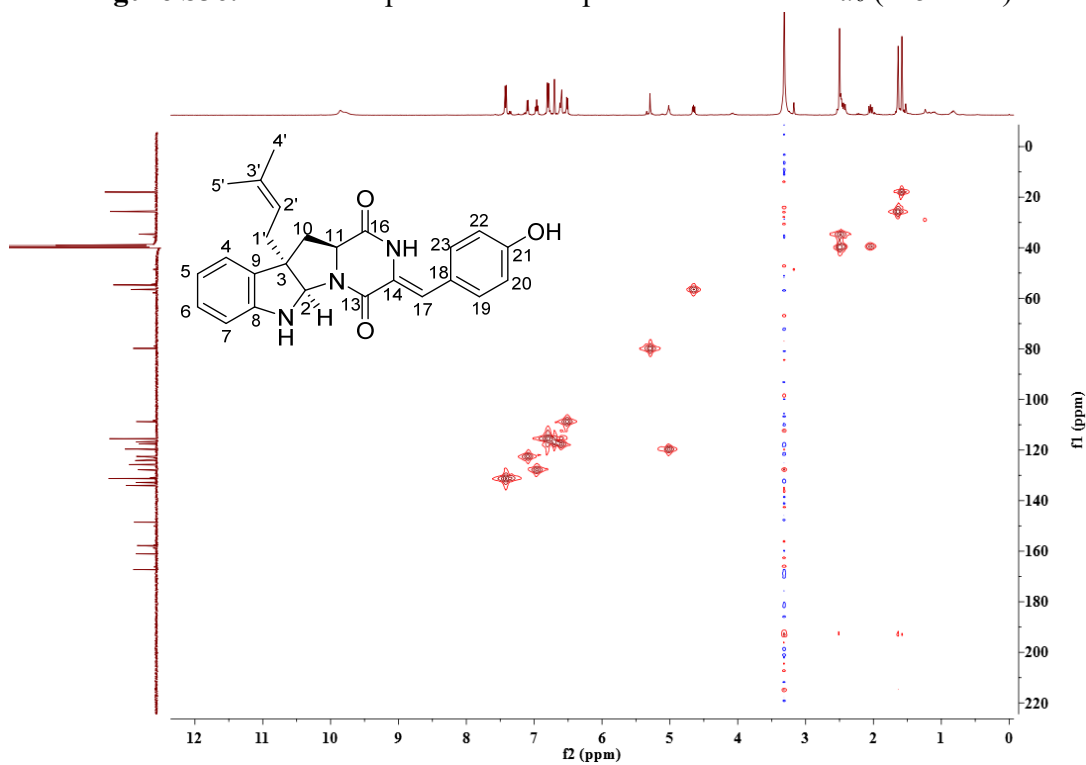


Figure S37. HSQC spectrum of compound **11** in $\text{DMSO-}d_6$

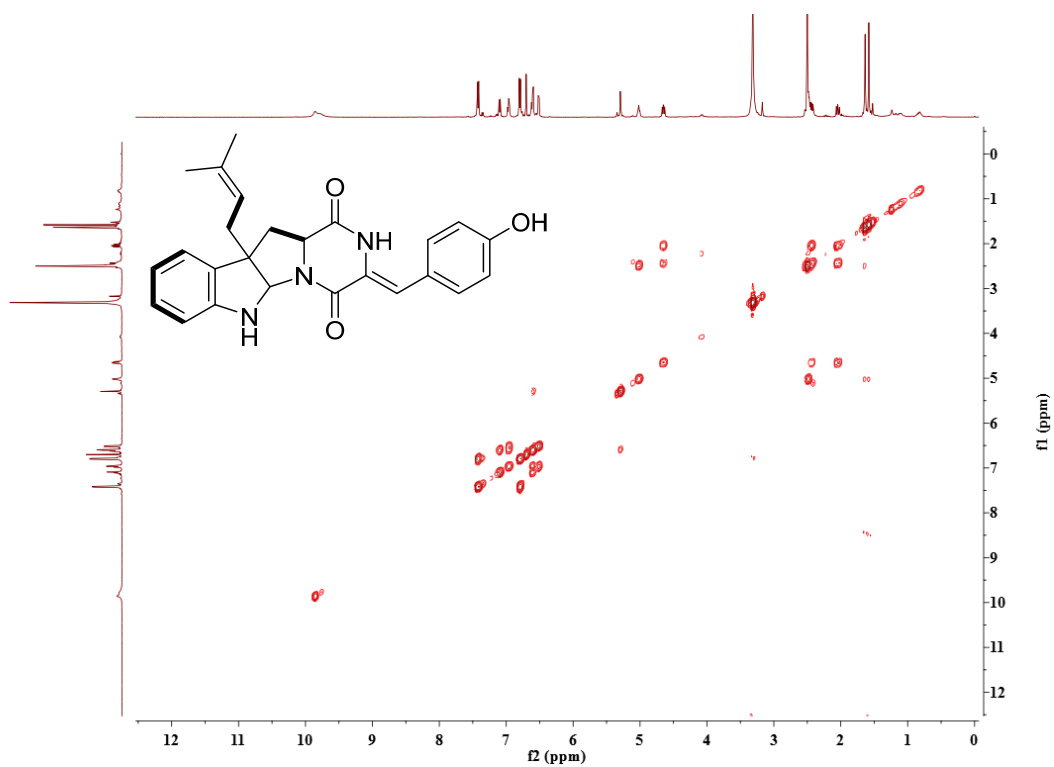


Figure S38. COSY spectrum of compound **11** in DMSO-*d*₆

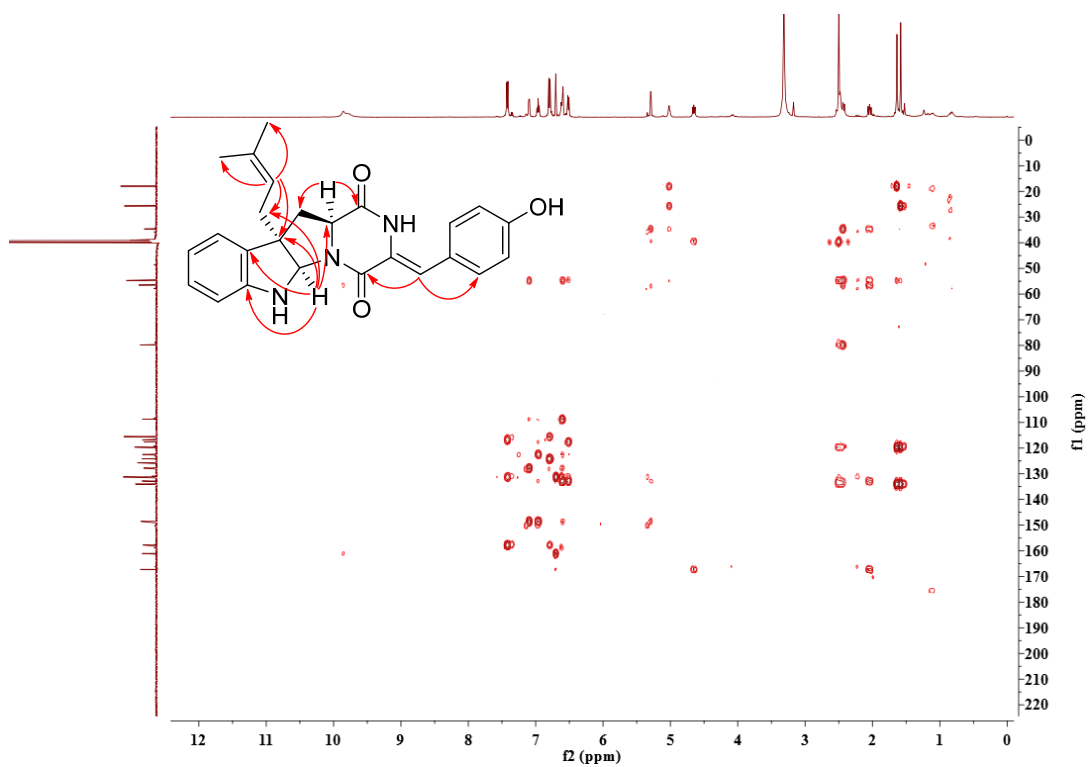


Figure S39. HMBC spectrum of compound **11** in DMSO-*d*₆

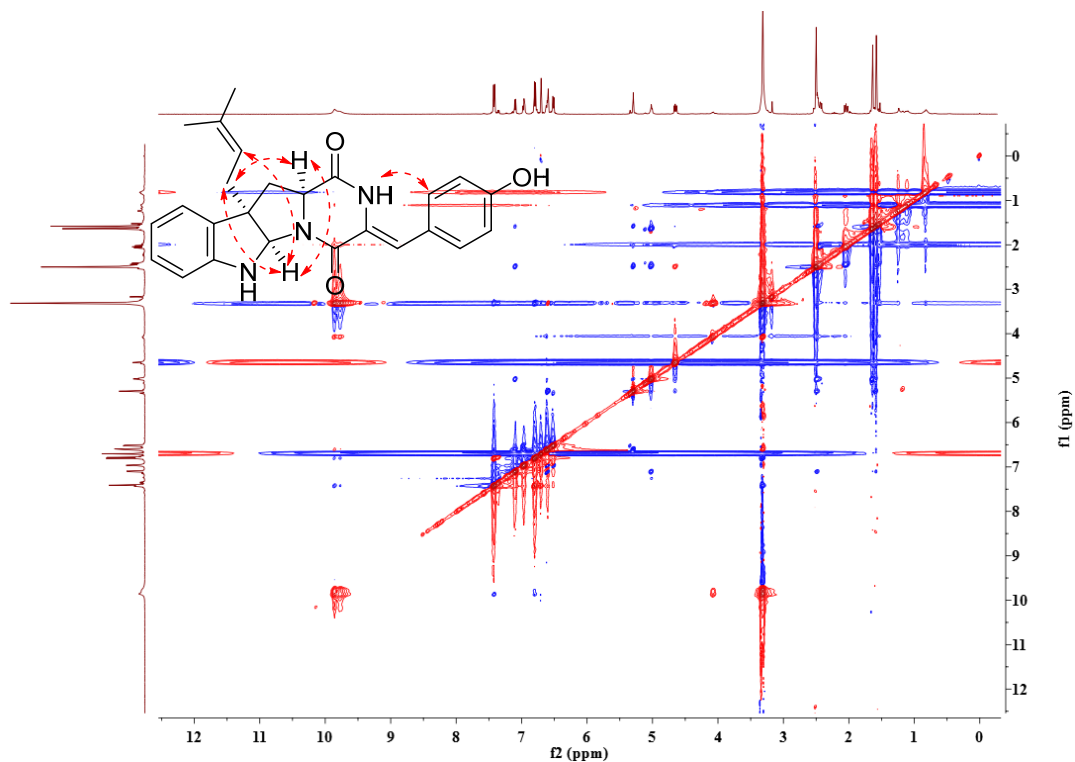


Figure S40. NOESY spectrum of compound **11** in DMSO-*d*₆

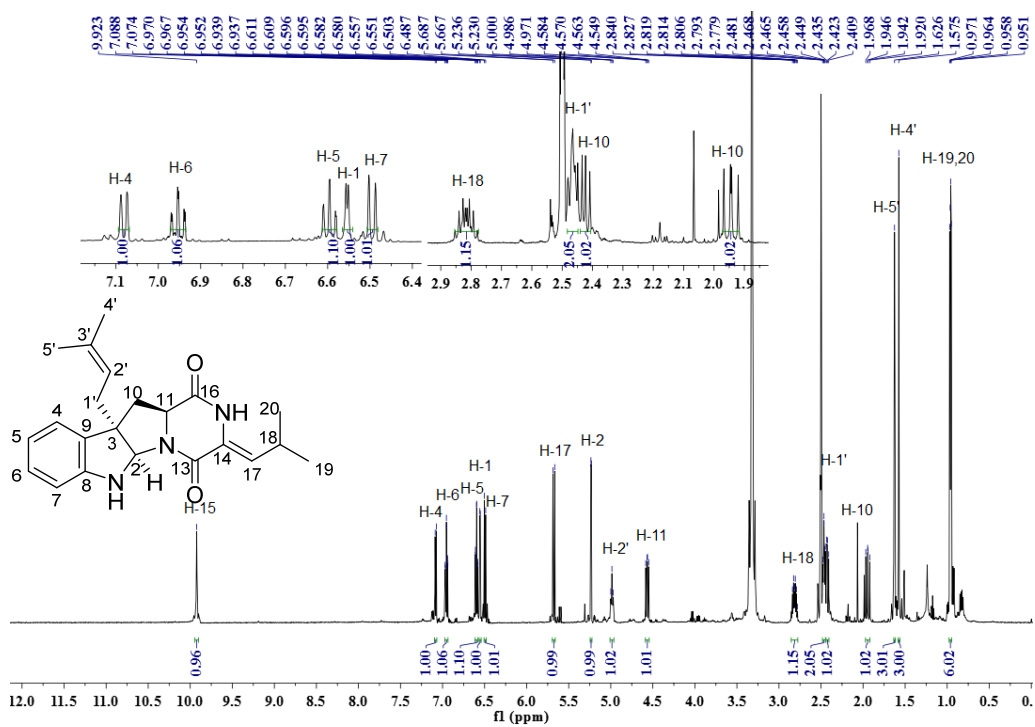


Figure S41. ¹H NMR spectrum of compound **12** in DMSO-*d*₆ (500 MHz)

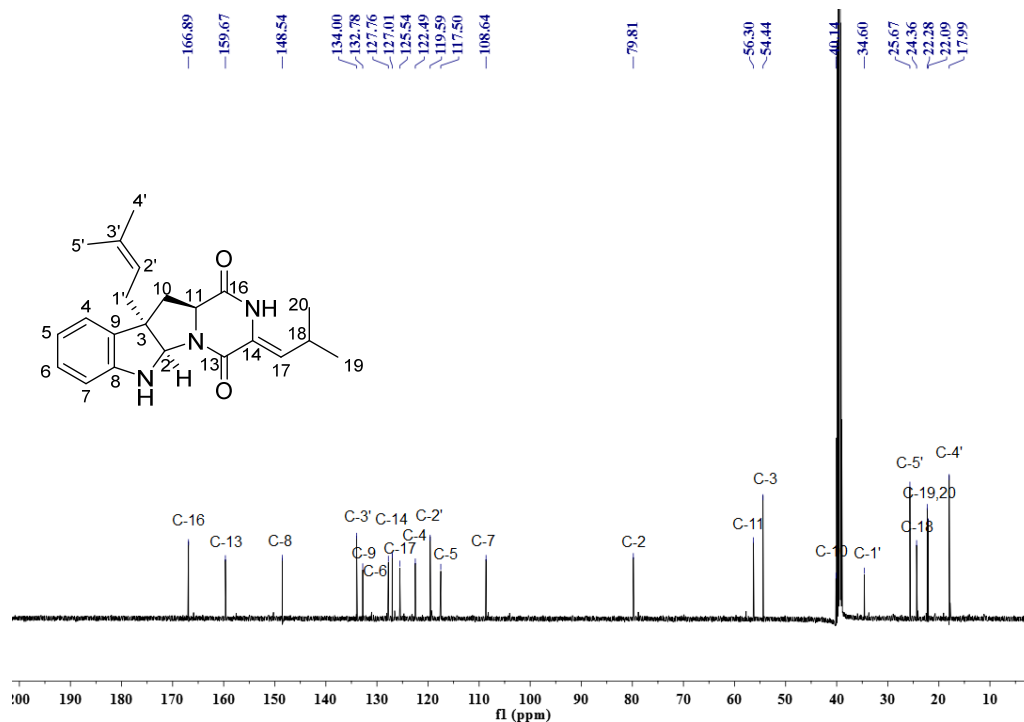


Figure S42. ^{13}C NMR spectrum of compound **12** in $\text{DMSO-}d_6$ (125 MHz).

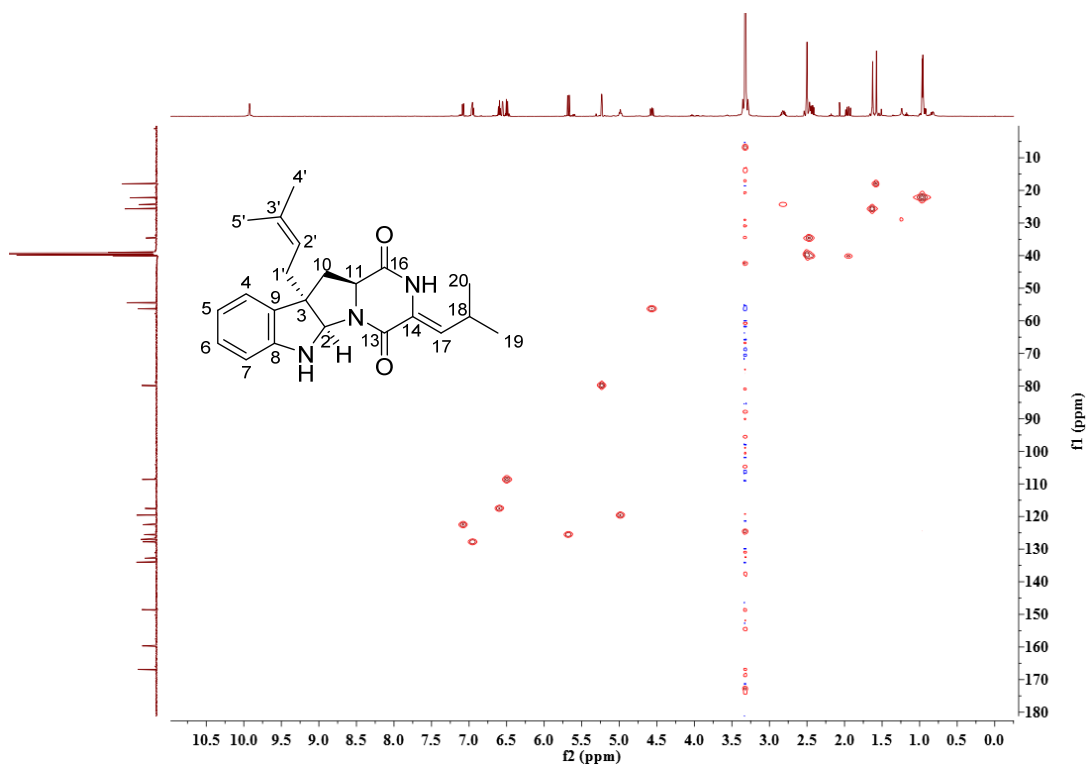


Figure S43. HSQC spectrum of compound **12** in $\text{DMSO-}d_6$

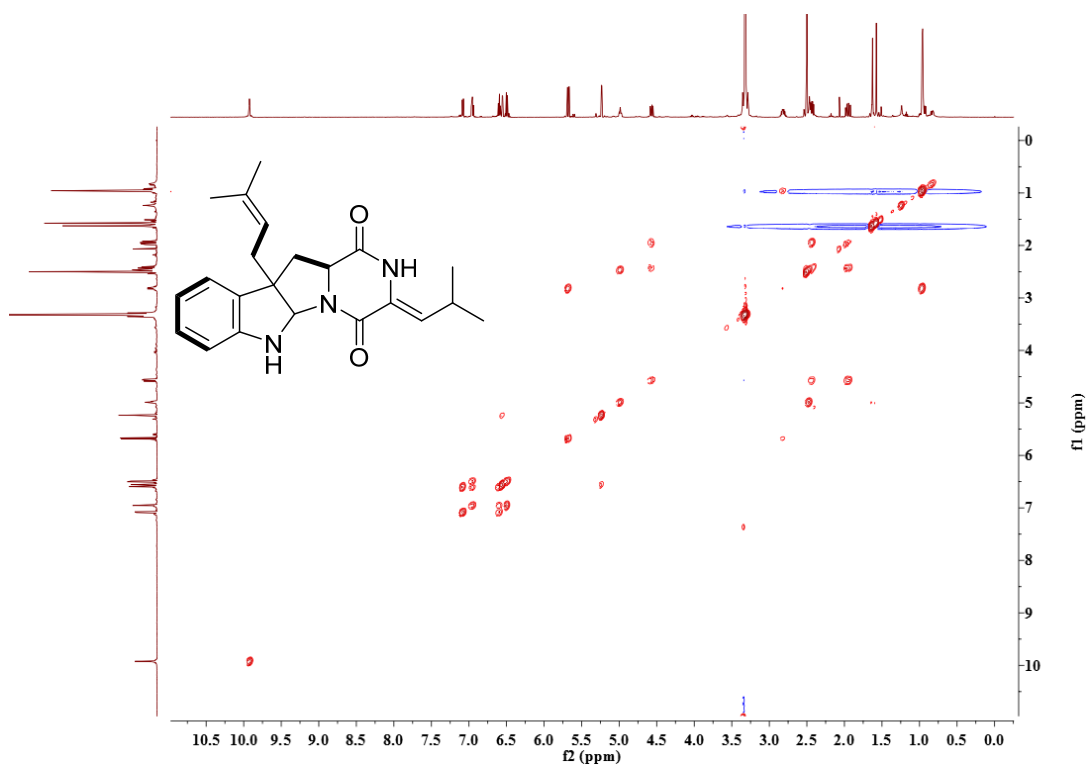


Figure S44. COSY spectrum of compound **12** in DMSO-*d*₆

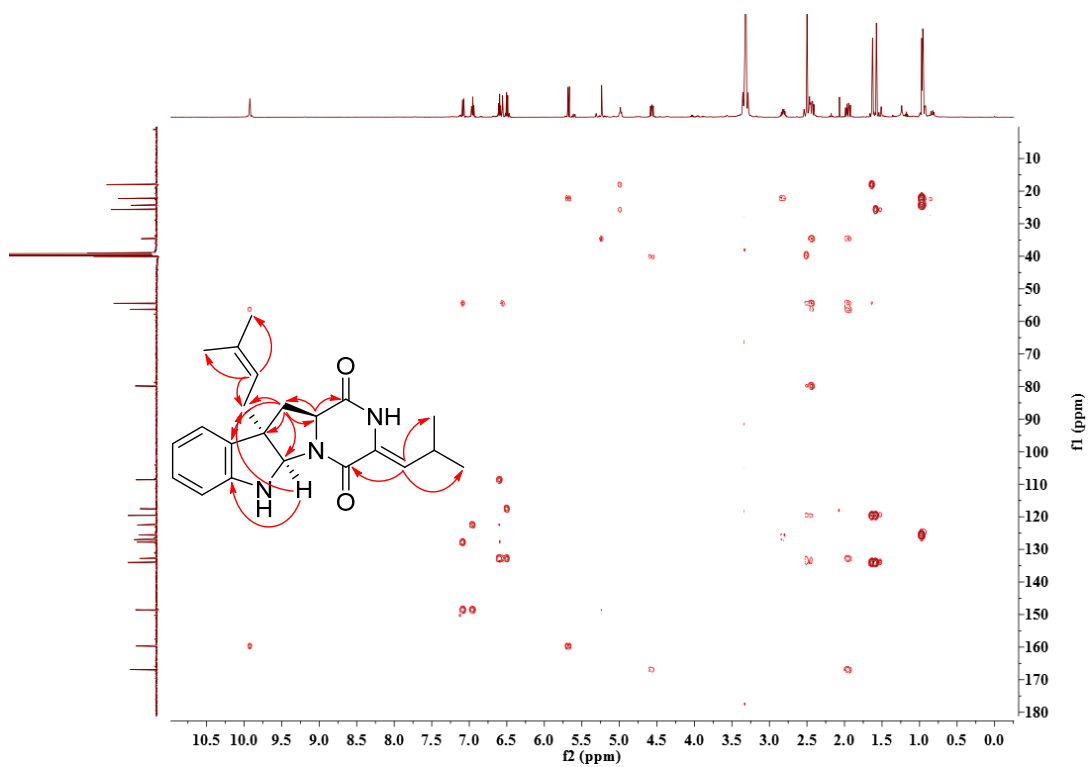


Figure S45. HMBC spectrum of compound **12** in DMSO-*d*₆

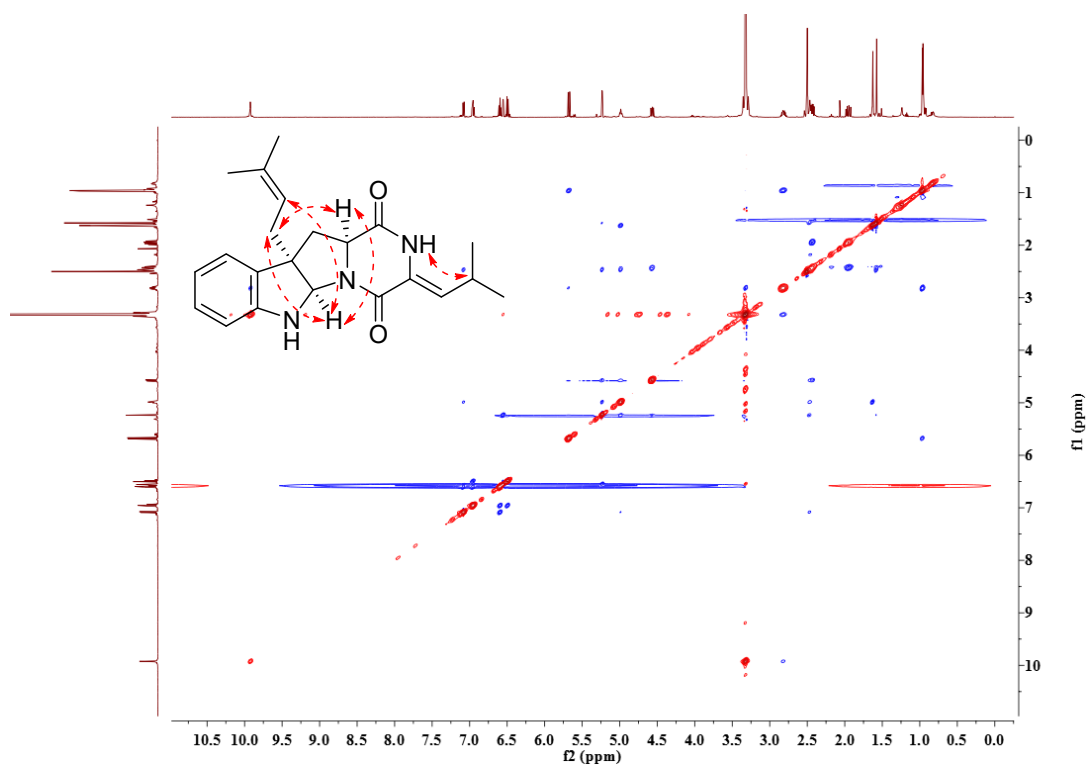


Figure S46. NOESY spectrum of compound **12** in DMSO-*d*₆

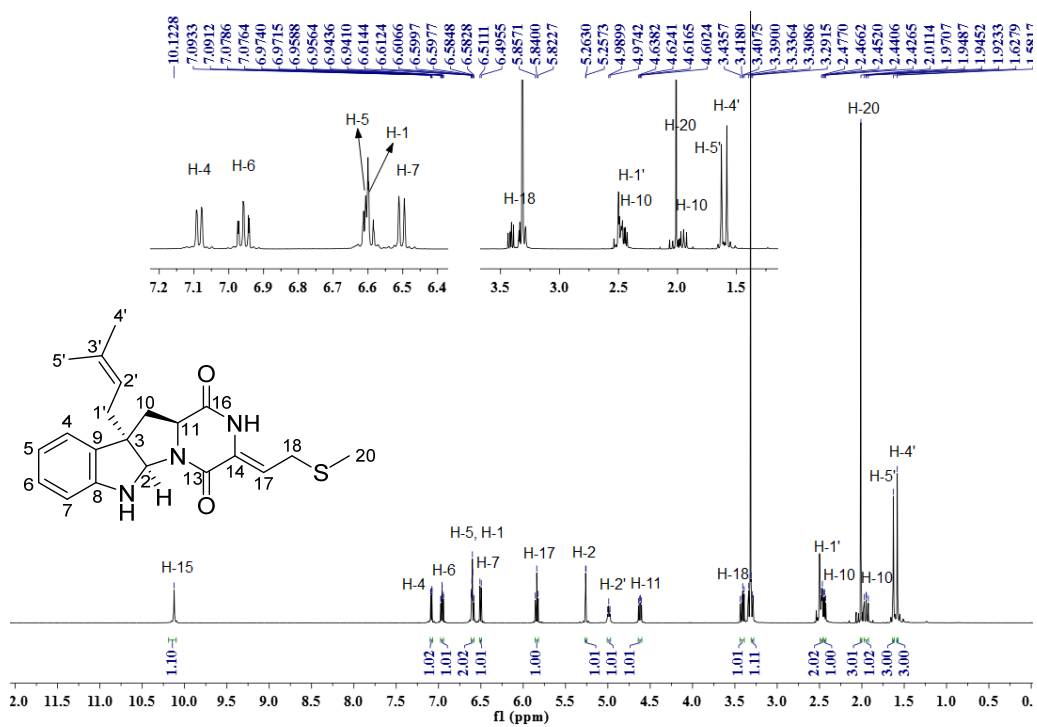


Figure S47. ¹H NMR spectrum of compound **13** in DMSO-*d*₆ (500 MHz)

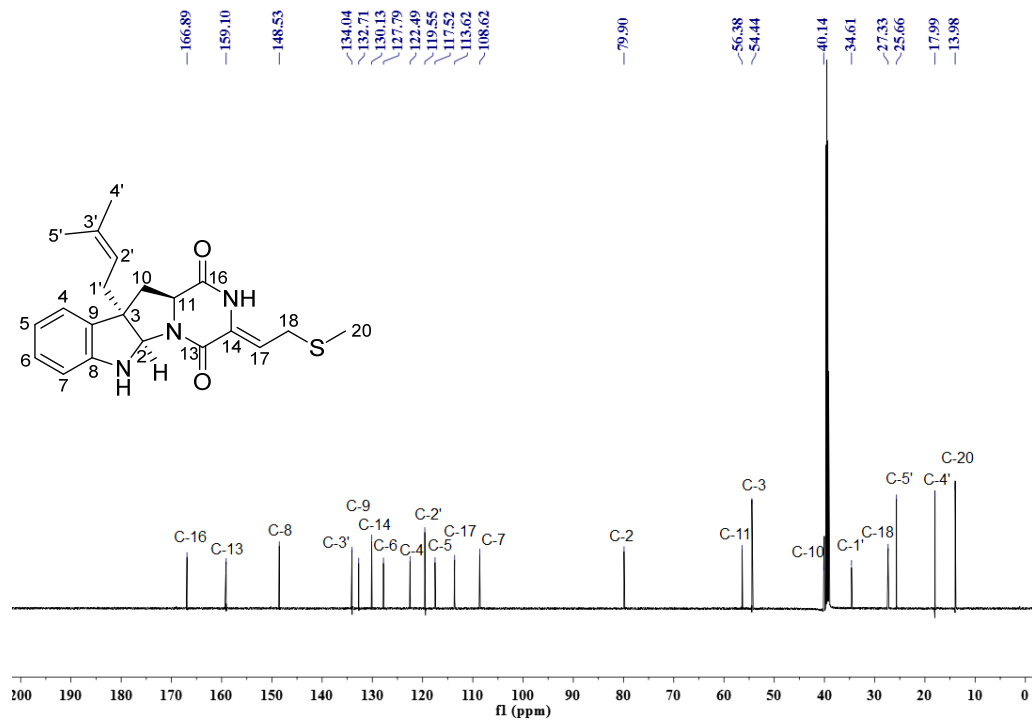


Figure S48. ^{13}C NMR spectrum of compound **13** in $\text{DMSO-}d_6$ (125 MHz)

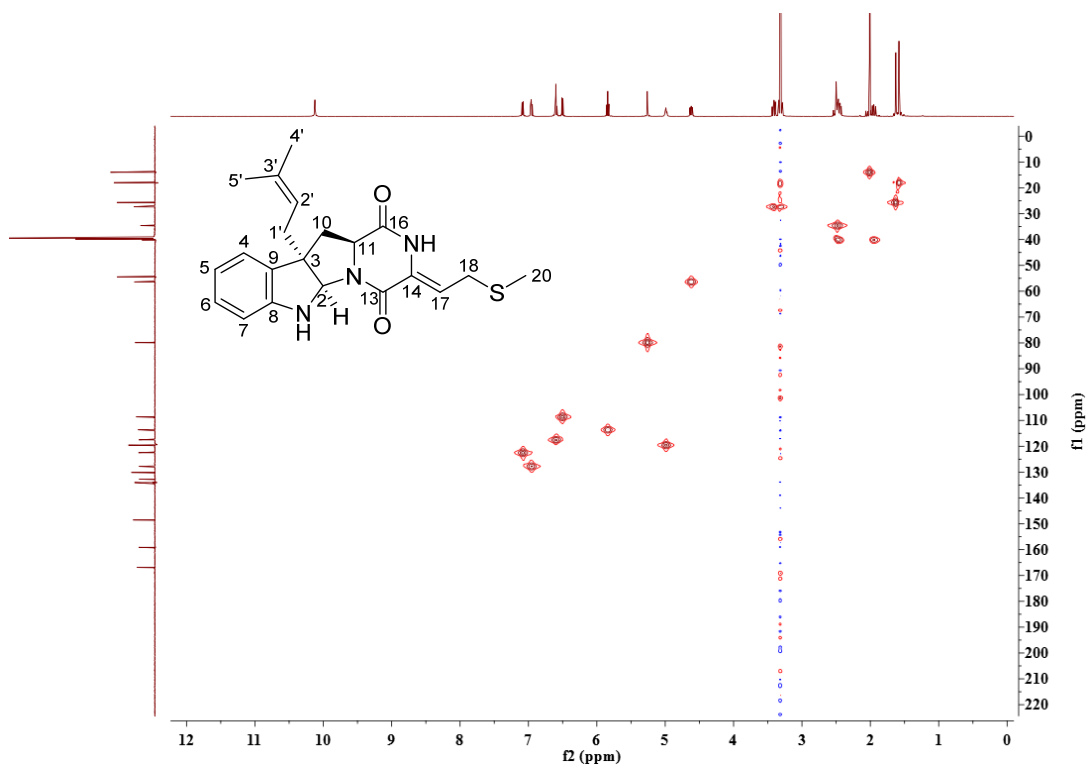


Figure S49. HSQC spectrum of compound **13** in $\text{DMSO-}d_6$

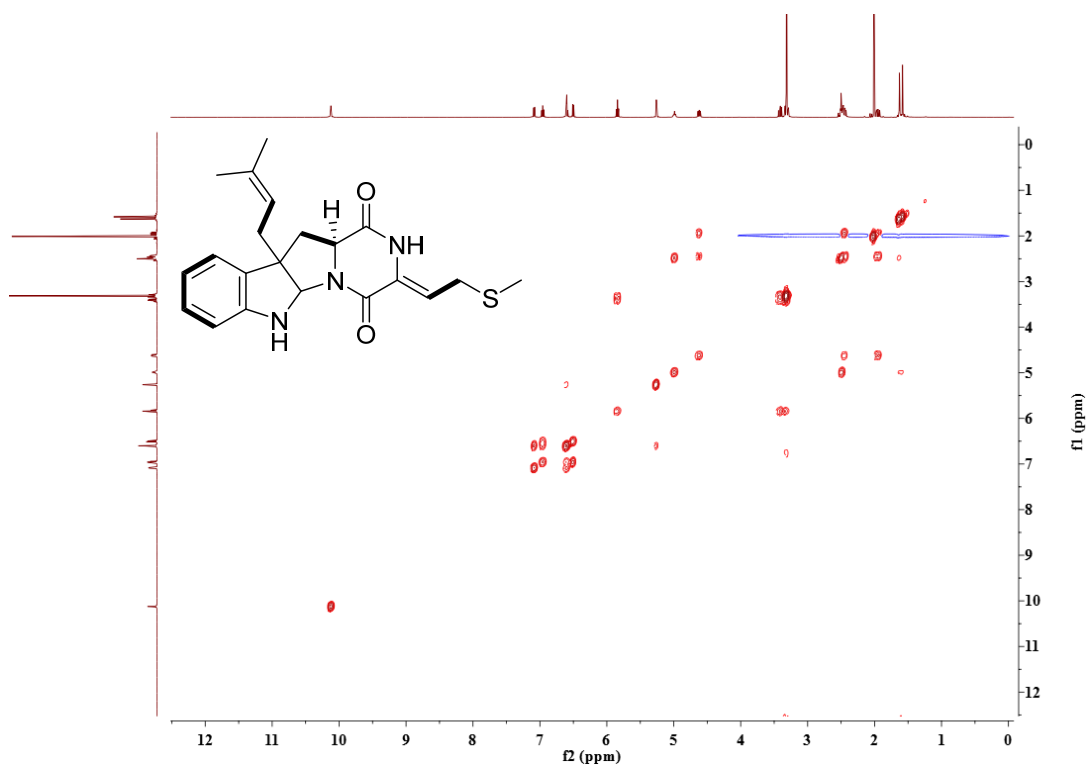


Figure S50. COSY spectrum of compound **13** in DMSO- d_6

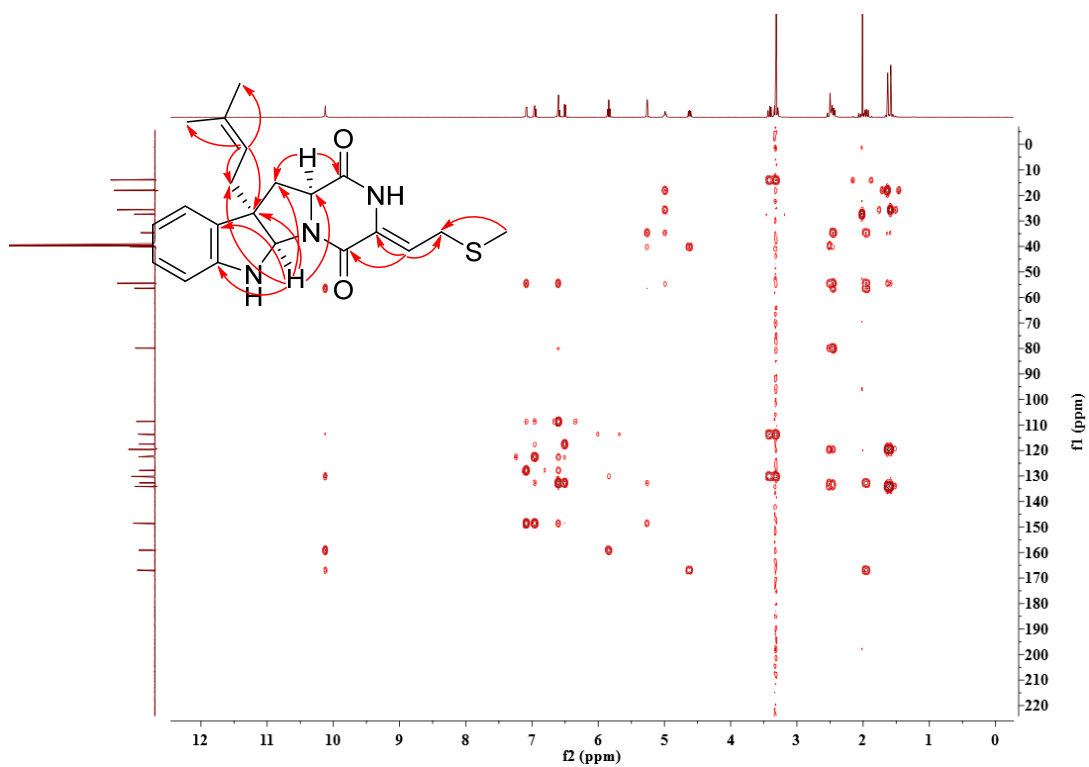


Figure S51. HMBC spectrum of compound **13** in DMSO- d_6

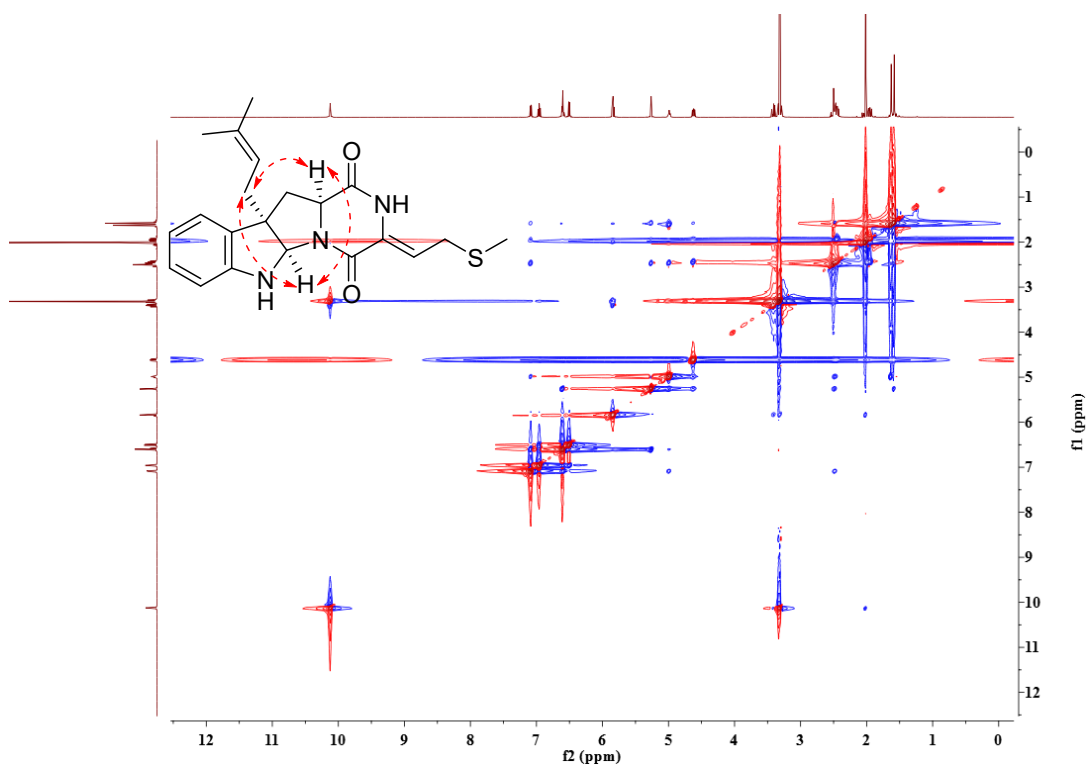


Figure S52. NOESY spectrum of compound 13 in DMSO-*d*₆.

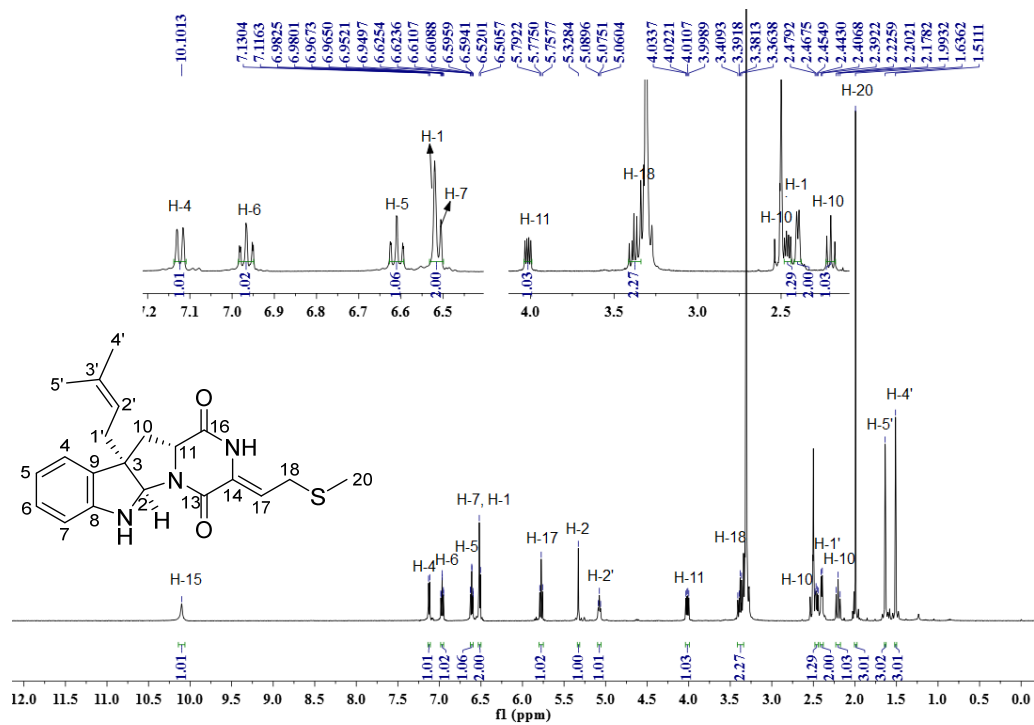


Figure S53. ¹H NMR spectrum of compound 14 in DMSO-*d*₆ (500 MHz)

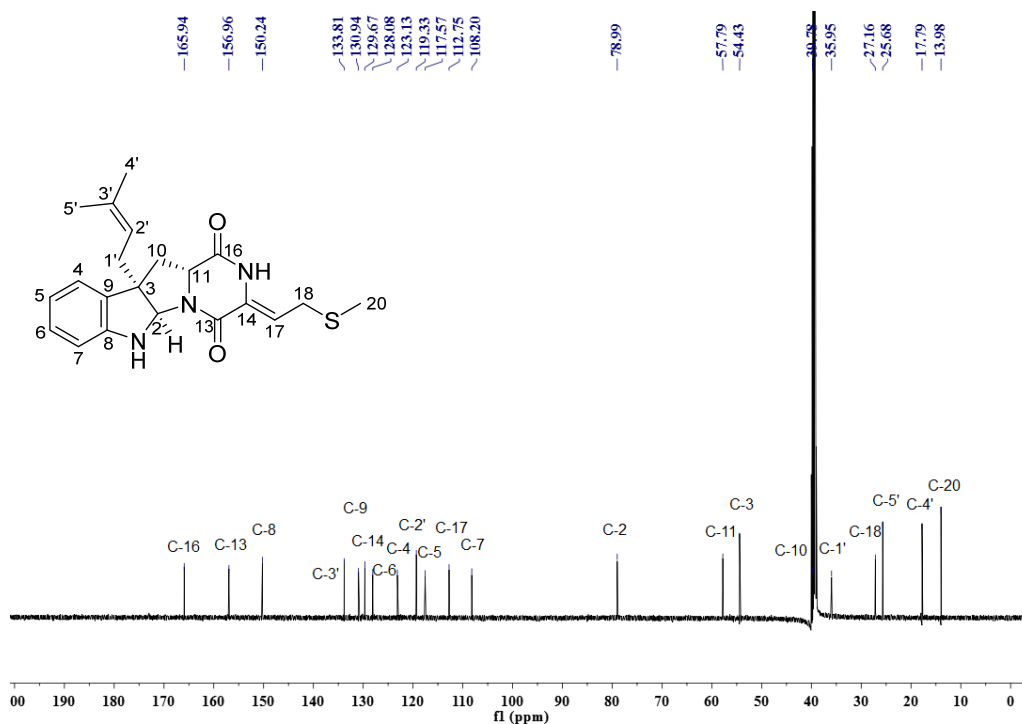


Figure S54. ^{13}C NMR spectrum of compound **14** in $\text{DMSO-}d_6$ (125 MHz)

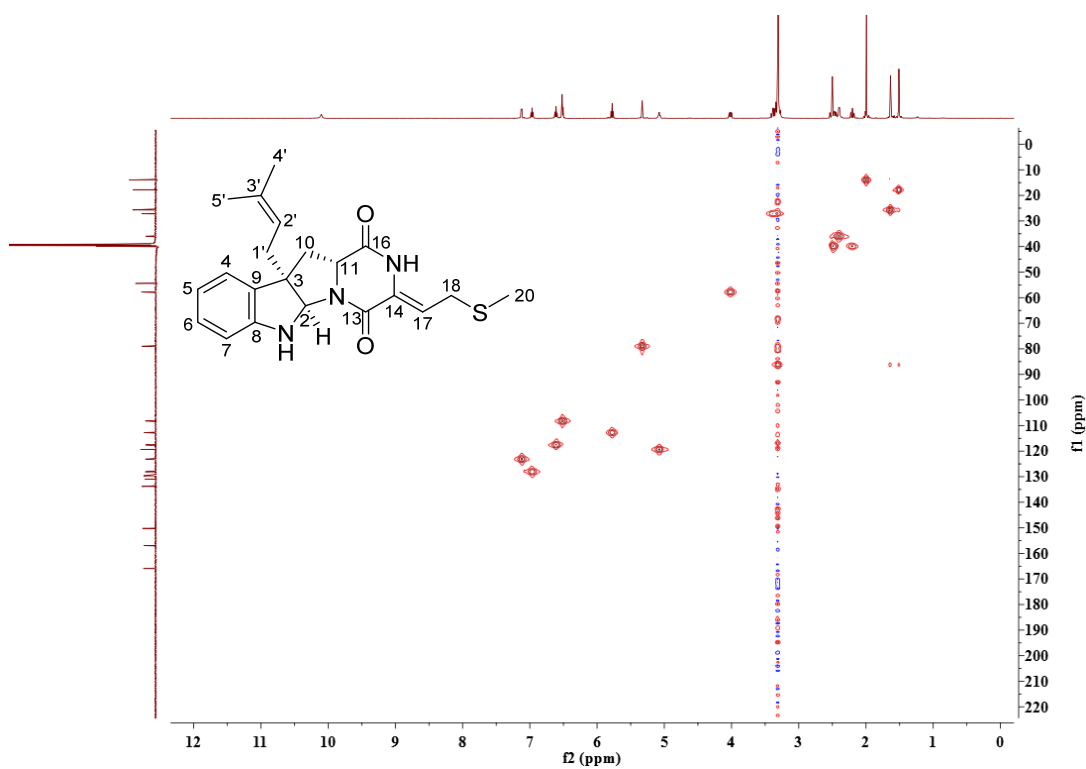


Figure S55. HSQC spectrum of compound **14** in $\text{DMSO-}d_6$

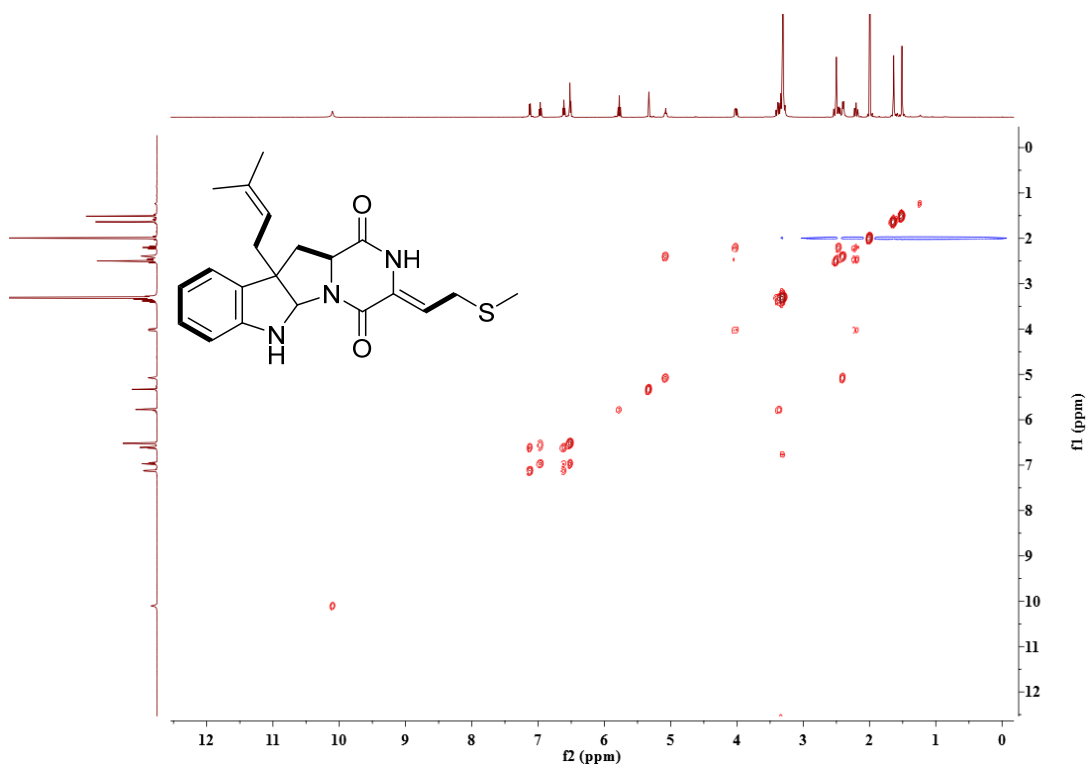


Figure S56. COSY spectrum of compound **14** in DMSO-*d*₆

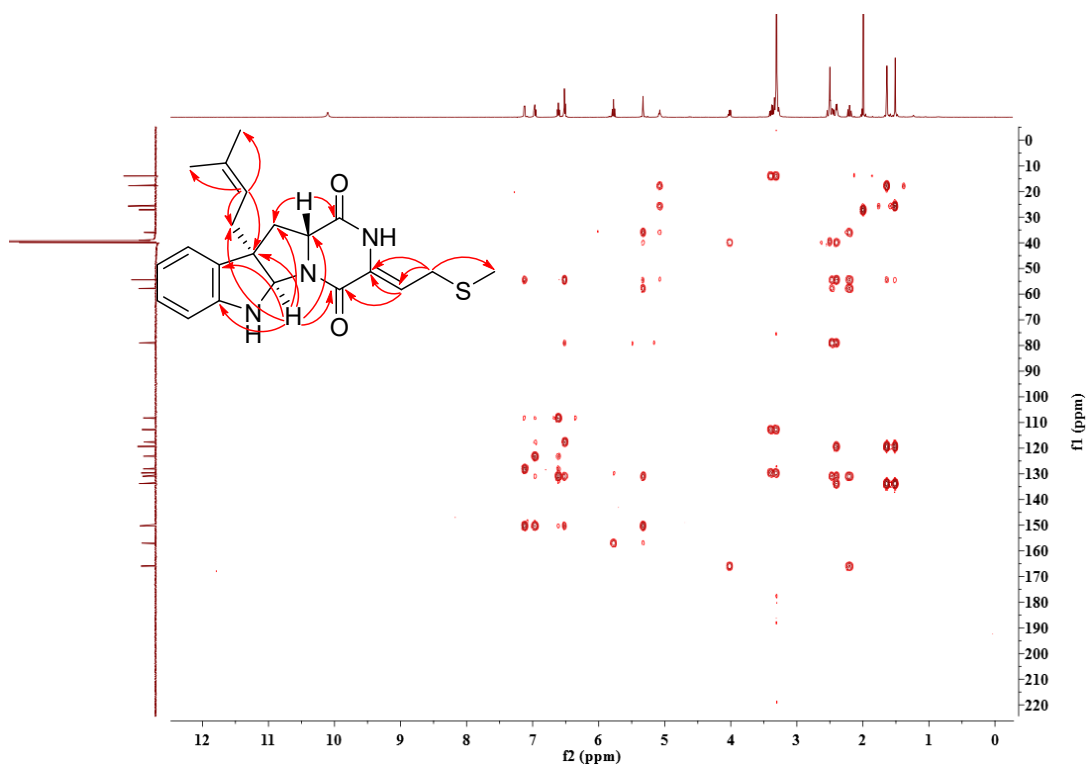


Figure S57. HMBC spectrum of compound **14** in DMSO-*d*₆

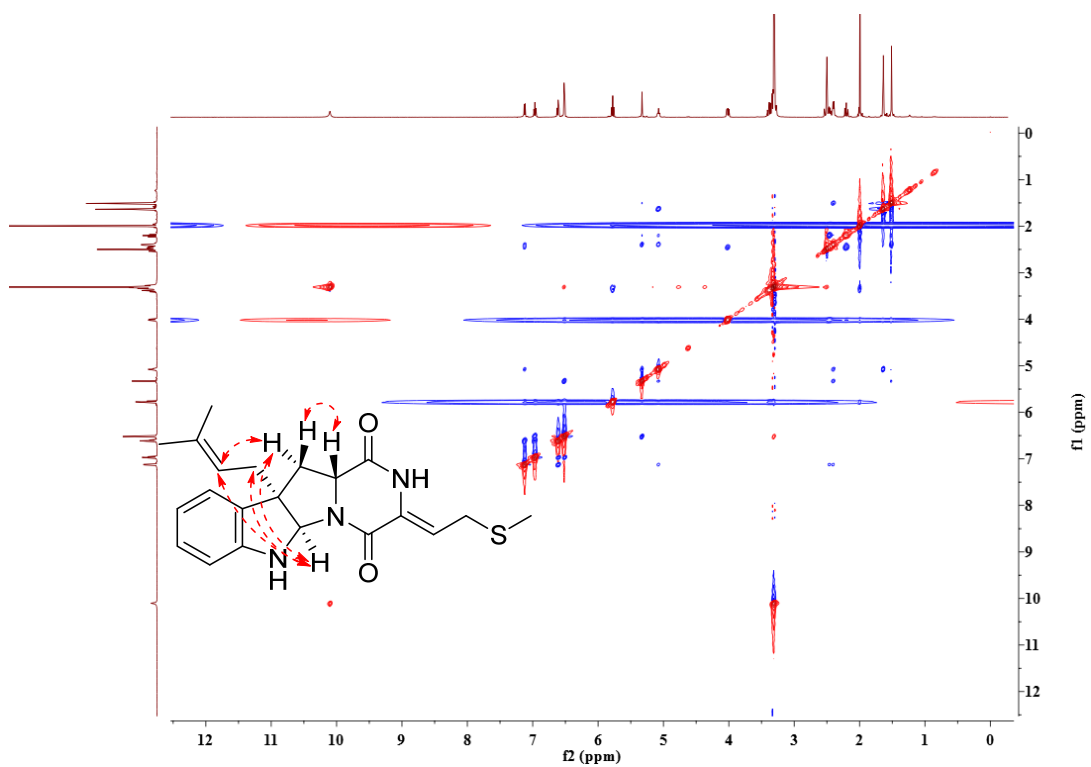


Figure S58. NOESY spectrum of compound **14** in DMSO-*d*₆

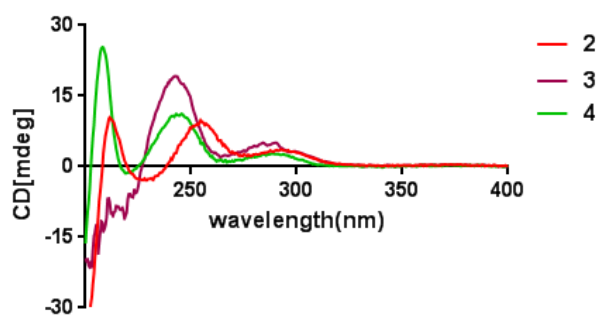


Figure S59. Experimental ECD spectra of 2 – 4 in MeOH

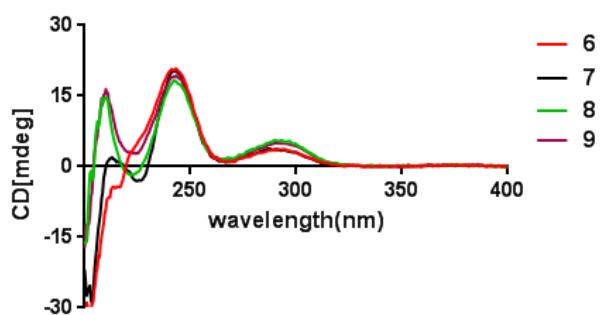


Figure S60. Experimental ECD spectra of 6 – 9 in MeOH

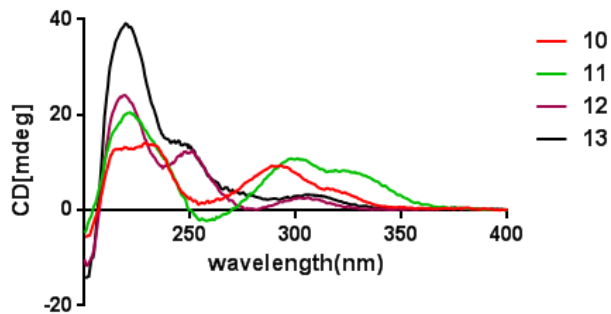


Figure S61. Experimental ECD spectra of 10 – 13 in MeOH

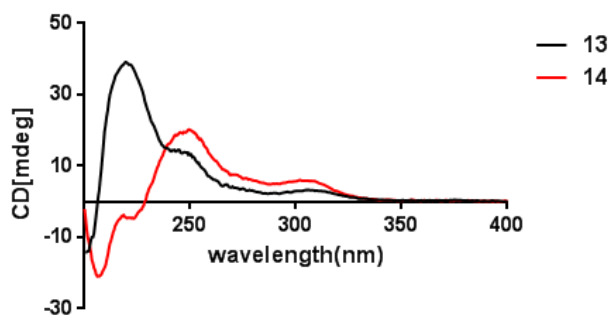


Figure S62. Experimental ECD spectra of 13 – 14 in MeOH.

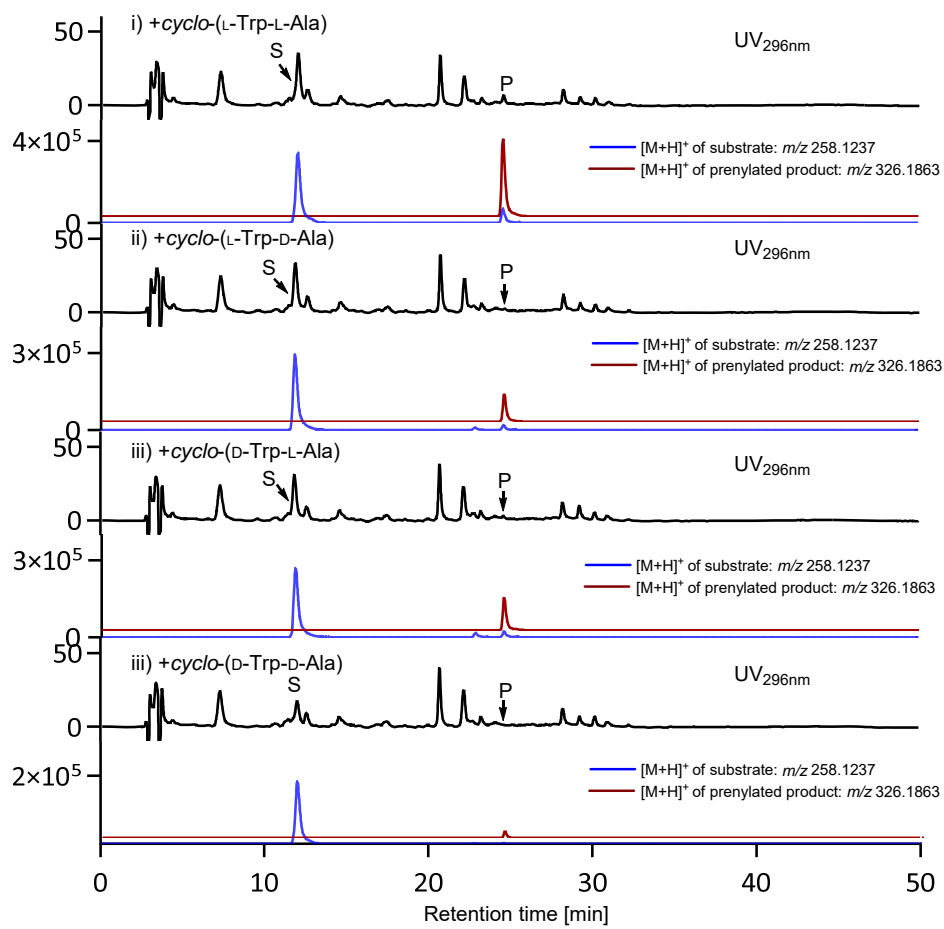


Figure S63. LC-MS analysis of *sasB* transformant after incubating with *cyclo*-Trp-Ala isomers
 S: substrate, P: product

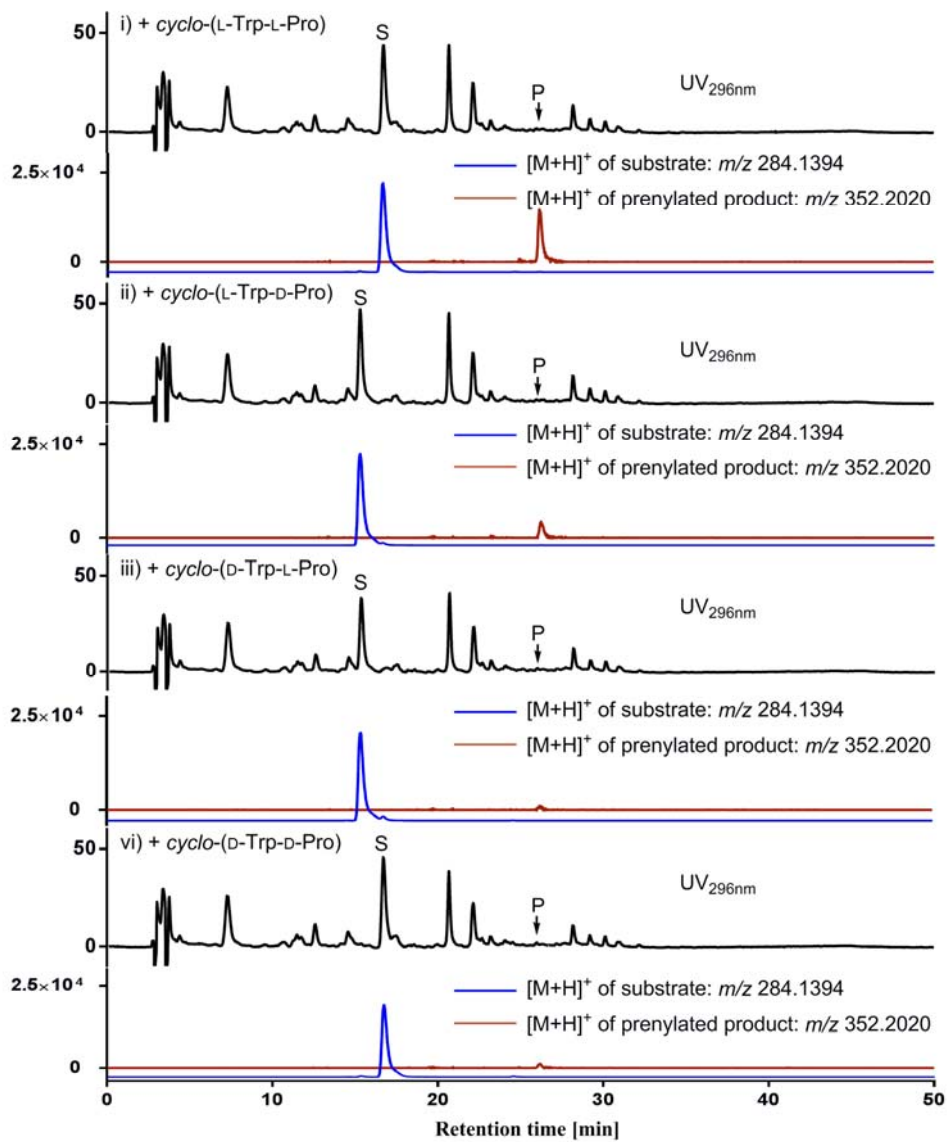


Figure S64. LC-MS analysis of *sasB* transformant after incubating with *cyclo*-Trp-Pro isomers
S: substrate, P: product

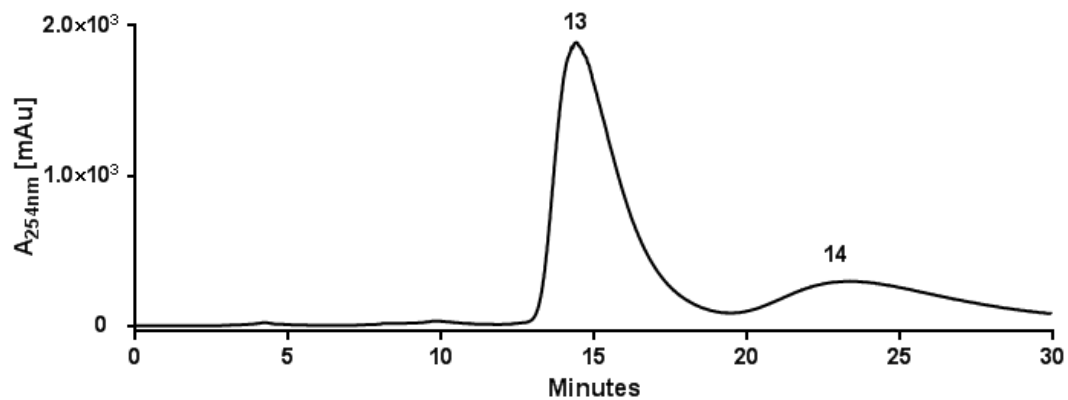


Figure S65. Purification of **13** and **14** on the Multohigh Chiral AM-RP column.

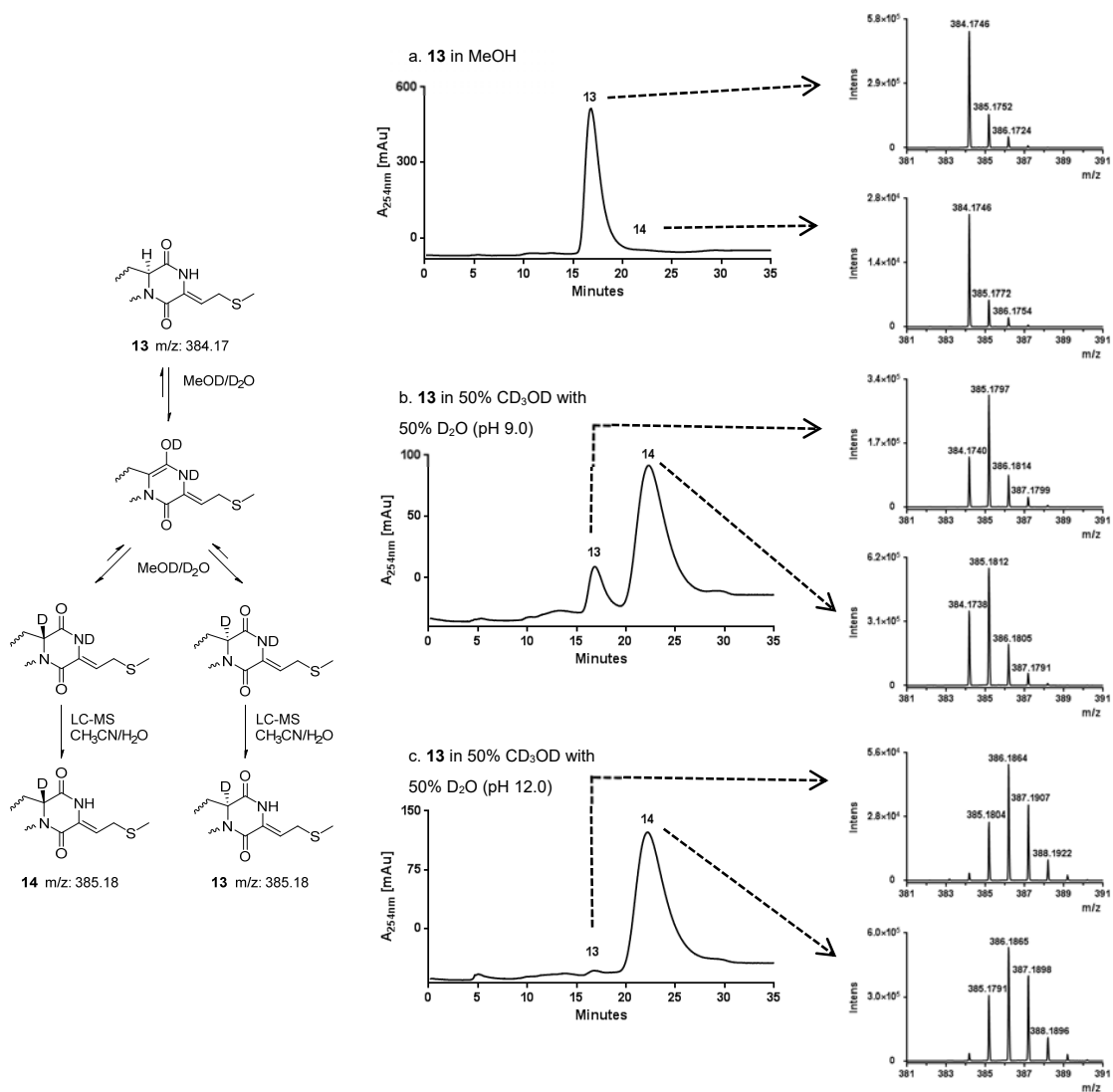


Figure S66. LC-MS analysis of **13** after incubation in CD₃OD/D₂O (1:1) for 14 h.

Reference

- (1) Zhang, Y.; Yao, T.; Jiang, Y.; Li, H.; Yuan, W.; Li, W. *Appl. Environ. Microbiol.* **2021**, *87*, e02525-20.
- (2) Yao, T.; Liu, J.; Liu, Z.; Li, T.; Li, H.; Che, Q.; Zhu, T.; Li, D.; Gu, Q.; Li, W. *Nat. Commun.* **2018**, *9*, 4091.
- (3) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U. S. A* **2003**, *100*, 1541-1546.
- (4) Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. *Bmc Genomics* **2014**, *15*, 97.

4.3 A *Streptomyces* cytochrome P450 enzyme catalyzes regiospecific C2-guaninylation for the synthesis of diverse guanitrypmycin analogs

A *Streptomyces* Cytochrome P450 Enzyme Catalyzes Regiospecific C2-Guaninylation for the Synthesis of Diverse Guanintrypmycin Analogs

Jing Liu, Yiling Yang, Xiulan Xie, and Shu-Ming Li*



Cite This: *J. Nat. Prod.* 2023, 86, 94–102



Read Online

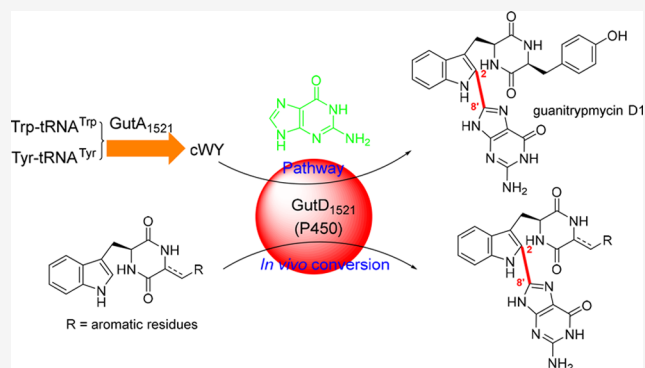
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Heterologous expression of a *cdps-p450* locus from *Streptomyces* sp. NRRL S-1521 led to the identification of guanintrypmycin D1, a new guaninylated diketopiperazine. The cytochrome P450 GutD₁₅₂₁ catalyzed the regiospecific transfer of guanine to C-2 of the indole ring of *cyclo*-(L-Trp-L-Tyr) via a C–C linkage and represents a new chemical transformation within this enzyme class. Furthermore, GutD₁₅₂₁ efficiently accepts several other tryptophan-containing cyclodipeptides or derivatives for regiospecific coupling with guanine, thus generating different guanintrypmycin analogs.



Natural products (NPs), especially those from microbes, play important roles in drug discovery and development.¹ Although the traditional methods such as the bioactivity screening strategy have made great contributions to new NP discovery in the past century, it becomes much more difficult to get new compounds because of the repeated isolation of known ones.² The rapid development of next generation sequencing (NGS) technologies led to an increase in microbial genome sequences in public databases.³ Bioinformatic analysis showed that a broad range of uncharacterized and cryptic biosynthetic gene clusters (BGCs), coding for novel metabolites, are hidden in the genomes. Different strategies such as promoter-exchange and heterologous expression have been successfully used to unveil the mystery of some of these BGCs.³ To date, a large number of NPs including 2,5-diketopiperazines (2,5-DKPs) have been discovered by exploiting these cryptic BGCs.^{4–6}

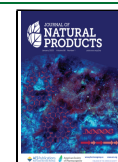
2,5-DKPs with a central diketopiperazine ring are ubiquitously distributed in nature and often found as side products of polypeptides.^{7,8} They are formed by condensation of two α -amino acids and therefore represent the smallest class of cyclic peptides.⁸ In microorganisms, they are products of two distinct enzyme families, the nonribosomal peptide synthetases (NRPSs) and cyclodipeptide synthases (CDPSs).^{9–11} NRPSs for CDP formation are bimodular enzymes and mainly found in fungi. Each module typically consists of three domains, adenylation (A) domain, peptidyl carrier protein (PCP), and the condensation (C) domain, which incorporate one amino acid into the peptide backbone.¹² Compared to NRPSs, CDPSs are small proteins first

reported from *Streptomyces noursei* and harbor very similar structures to class-I aminoacyl-tRNA synthetases.^{13,14} They use aminoacyl-tRNAs (aa-tRNAs) as substrates to synthesize the DKP scaffolds.¹⁵ Most of CDPSs in actinobacteria are clustered with genes for tailoring enzymes. To date, diverse classes of tailoring enzymes, such as prenyltransferases (PTs), methyltransferases (MTs) and cytochrome P450 enzymes, cyclodipeptide oxidases (CDOs), and 2-oxoglutarate/Fe²⁺-dependent oxygenases, have been found in *cdps*-related gene clusters.^{4,5}

It is worth mentioning that cytochrome P450s are found as the most prevalent modification enzymes in the characterized *cdps*-related gene clusters. P450 enzymes from the featured biosynthetic pathways catalyze a wide range of interesting chemical transformations, such as intramolecular C–C bond formation, different types of dimerization, aromatization of the DKP ring, and nucleobase transfer reactions.^{5,6,16} Seven different types of cyclodipeptide–nucleobase linkages have been characterized from CDPS-P450-related nucleobase transfer pathways, including C–C, C–N, and C–O bonds (Figure 1).^{17–21} Additionally, different tryptophan-containing and tyrosine-containing CDPs as well as two nucleobases, guanine

Received: September 5, 2022

Published: January 4, 2023



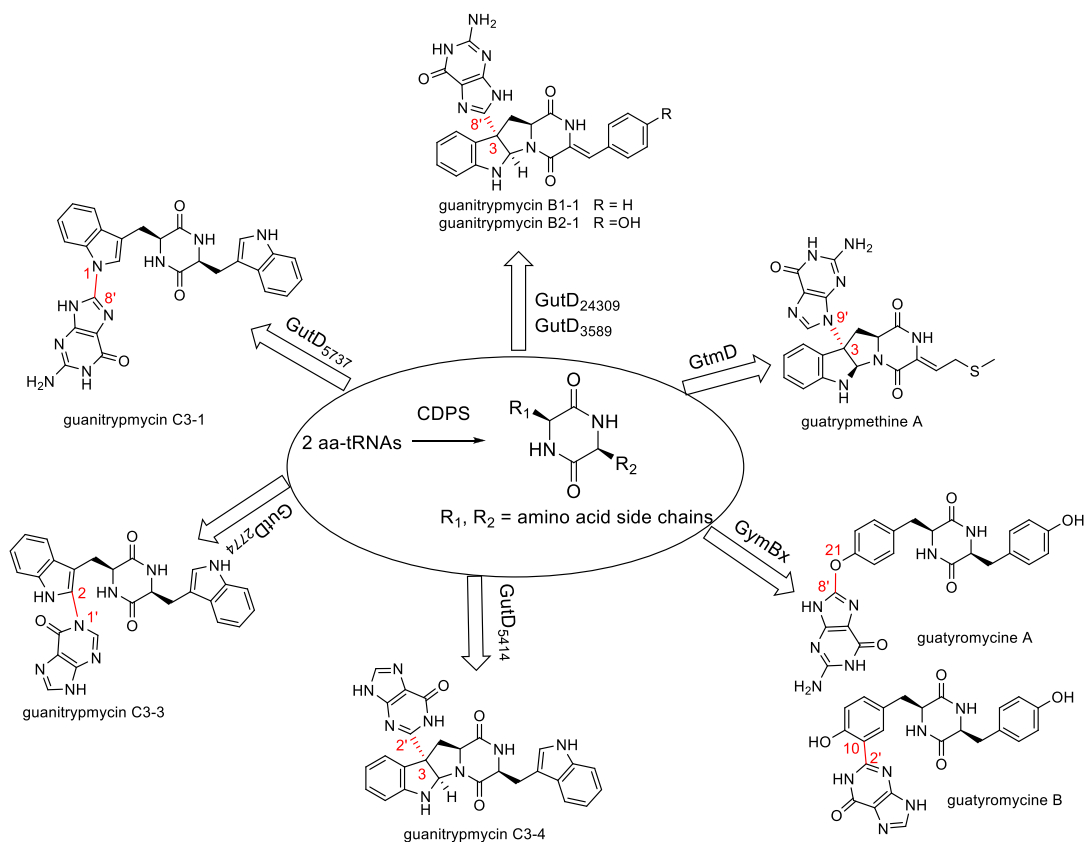


Figure 1. Examples of guanylated NPs with nucleobases attached on different positions of the DKP core.

and hypoxanthine, have been found as the substrates of these P450 enzymes.^{17–21} These findings significantly expanded the spectrum of DKP derivatives and highlight the promise of CDPS-related enzymes as unique biocatalysts for novel transformations.

In this study, a two-gene cluster coding for a CDPS and a P450 was identified in *Streptomyces* sp. NRRL S-1521 by phylogenetic analysis. Heterologous expression of the gene cluster led to the identification of a new guanylated DKP guanitrypmycin D1. Biotransformation experiments demonstrated that GutD₁₅₂₁ catalyzes the transfer of a guanine onto C-2 of the indole ring of *cyclo*-(L-Trp-L-Tyr) (cWY) via a C–C bond. Precursor incubation experiments revealed GutD₁₅₂₁ can also utilize other tryptophan-containing CDPs as well as their dehydrogenated forms as substrates, for the synthesis of different guanitrypmycin analogs. Therefore, this study provides a biocatalyst for a new linkage pattern between a DKP indole ring and a guanine moiety and expands the functional spectrum of P450s as tailoring enzymes.

RESULTS AND DISCUSSION

Identification and Analysis of the *gut*₁₅₂₁ Gene Cluster. We have identified a dozen *cdps-p450*-containing gene clusters since 2019.^{17–19,22} Most of them coded for new DKP derivatives and biosynthetic enzymes, which inspired us to explore more *cdps*-related gene clusters for novel metabolites. In the previous studies, we took the functionally characterized CDPSs and P450s as probes to search and identify their putative homologues in the public databases. Subsequent phylogenetic analysis led to the identification of plenty of uncharacterized *cdps-p450* gene clusters. Based on

the phylogenetic information, one *cdps-p450* gene cluster from *Streptomyces* sp. NRRL S-1521 attracted our attention. Following the nomenclature of the known clusters,^{18,19} we named *gutA*₁₅₂₁ and *gutD*₁₅₂₁ for the *cdps* and P450 genes, respectively. GutD₁₅₂₁ shows high identities (approximate 51–63%) to the known GutDs and GtmD that function as nucleobase transferases (Table S1). However, this candidate was located in a separate subclade in the phylogenetic tree based on the characterized P450s. Therefore, we speculated that this gene cluster could synthesize novel diketopiperazine derivative(s) (Table S1, Figure S1).

Expression of *gut*₁₅₂₁ Gene Cluster for the Production of Guanitrypmycin D1. As heterologous expression has been demonstrated to be a rapid and efficient approach to exploit the chemical potential from diverse microorganisms, we expected successful application of this method to identify the product of this gene cluster as well. First, the CDPS gene *gutA*₁₅₂₁ was directly cloned into the pET28a (+) vector and expressed in *Escherichia coli* BL21 (DE3). After induction with isopropyl β-D-1-thiogalactopyranoside (IPTG) for 20 h, the cultures were extracted with EtOAc and analyzed on LC-MS. A sole peak for **1** with a [M + H]⁺ ion at *m/z* 350.1497 was identified in the *E. coli* transformant harboring *gutA*₁₅₂₁, which was not detected in the mutant with the empty vector pET28a (+) as the negative control (Figure 2i and ii). Compared to an authentic standard, compound **1** was characterized as cWY, which was also confirmed by its ¹H and ¹³C NMR data (Table 1, Figures S2 and S3). This proved that the CDPS *gutA*₁₅₂₁ functions as a cWY synthase (Scheme 1).

After characterization of the CDPS, the two-gene cluster was amplified from genomic DNA of the strain NRRL S-1521 by

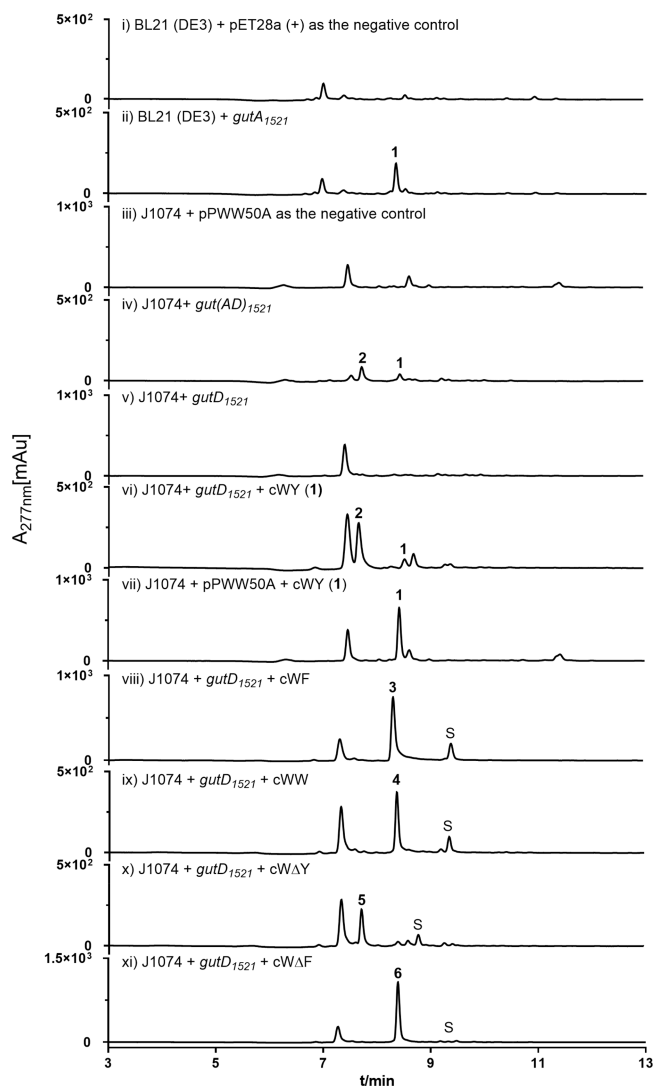


Figure 2. HPLC analysis of the generated transformants and the *gutD*₁₅₂₁ transformant with/without precursor supply. S: substrate.

PCR and cloned into pPWW50A for expression in the widely used host *Streptomyces albus* J1074 by conjugation. The obtained conjugants were cultivated in modified RS medium for 7 days and treated as mentioned above for the *gutA*₁₅₂₁ transformant. Based on the LC-MS data, expression of the candidate gene cluster *gut(AD)*₁₅₂₁ resulted in the production of two compounds, **1** and **2**. Neither was found in the negative control with pPWW50A (Figure 2iii and iv). Compound **2** exhibited a $[M + H]^+$ ion at m/z 499.1843, which is 149 Da larger than that of cWY. Thus, we deduced that an additional guanine residue was connected to cWY. Subsequently, compound **2** was isolated from a large-scale fermentation culture, and its structure was further elucidated by detailed NMR analysis.

Inspection of the NMR data of compound **2** revealed the presence of three characteristic ¹H signals of a guaninyl residue at δ 10.70 (H-1'), 12.98 (H-9'), and 6.33 (H-10') with five corresponding ¹³C signals at δ 151.6 (C-2'), 146.0 (C-4'), 106.0 (C-5'), 159.3 (C-6'), and 140.5 (C-8') (Table 1).^{18,23–25} Although no clear correlation in the HMBC spectrum was observed between the cWY skeleton and the guaninyl moiety, the missing ¹H signal for H-2 of the cWY part

and that for H-8' of guanine indicated the new C–C bond between C-2 and C-8' of the two moieties (Scheme 1, Figures S4–S8). In addition, the signal of C-3 in the ¹³C spectrum of compound **2** is deshielded by 5 ppm in comparison to that in compound **1** (Table 1, Figure S3), whereas signals for other carbons like C-5–C-7 are only deshielded approximately 3 ppm. As compound **2** features tryptophanyl and guaninyl residues, we named it guanitryptmycin D1. Cultivation of *S. albus* carrying *gut(AD)*₁₅₂₁ in ¹⁵NH₄Cl-containing medium revealed incorporation of three and eight ¹⁵N atoms in **1** and **2**, respectively, providing additional evidence for the guanitryptmycin D1 structure (Figure 3). The above results implied GutD₁₅₂₁ as a new nucleobase transferase for the specific C-2–C-8' connection between the indole and guaninyl units, differing from the previous reported P450s (Figure 1).^{17–21}

In order to confirm the GutD₁₅₂₁ function, we intended to carry out biochemical characterization with an *E. coli* overproduced recombinant protein. Unfortunately, no soluble GutD₁₅₂₁ was obtained. Therefore, we cloned it into pPWW50A for expression in *S. albus* J1074, followed by precursor incubation experiments. Supplementation of compound **1** to the *gutD*₁₅₂₁ transformant led to the production of **2**, while no additional metabolite was detected in J1074 harboring pPWW50A (Figure 2v–vii). These data proved unequivocally GutD₁₅₂₁ as a specific C-2–C-8' guaninyl transferase and being responsible for the generation of a new guaninylated DKP (Scheme 1). Addition of *cyclo*-(D-Trp-L-Tyr) to the *gutD*₁₅₂₁ transformant did not lead to any conversion (Figure S30), proving the importance of the L-configuration of the tryptophanyl moiety for acceptance by GutD₁₅₂₁.

Generation of Diverse Guanitryptmycin Analogs by Biotransformation. After proof of the GutD₁₅₂₁ function, we investigated its substrate specificity toward other tryptophan-containing CDPs including *cyclo*-(L-Trp-L-Ala) (cWA), *cyclo*-(L-Trp-L-Phe) (cWF), *cyclo*-(L-Trp-L-His) (cWH), *cyclo*-(L-Trp-L-Leu) (cWL), *cyclo*-(L-Trp-L-Met) (cWM), *cyclo*-(L-Trp-L-Pro) (cWP), and *cyclo*-(L-Trp-L-Trp) (cWW). After supplying these CDPs into the J1074 transformant harboring *gutD*₁₅₂₁, the 5-day-old cultures were monitored for their conversion by LC-MS. As shown in Figure 2, cWF and cWW were efficiently transformed to the products **3** and **4** with conversion yields of $68 \pm 2\%$ and $84 \pm 2\%$, respectively (Figure 2viii and 2ix). Their $[M + H]^+$ ions are 149 Da larger than those of the corresponding precursors, indicating the attachment of a guaninyl residue. In contrast, other CDPs like cWA, cWH, cWL, cWM, and cWP cannot be efficiently converted by GutD₁₅₂₁ (Figure S31).

Compounds **3** and **4** were subsequently isolated from the large-scale cultures, and their structures were elucidated based on NMR data (Tables 1 and 2, Figures S9–S18). The typical signals of the guaninyl moiety were clearly observed in their ¹H NMR spectra. For compound **3**, these signals are found at δ_{H} 10.63 (br s, H-1'), 12.94 (br s, H-9'), and 6.27 (br s, H-10'). For compound **4**, they are at δ_{H} 10.80 (br s, H-1'), 12.91 (br s, H-9'), and 6.42 (br s, H-10'). Compared with compound **2**, very similar values can be assigned for the five carbons of guanine residue in the ¹³C spectra as well. Similar to that of compound **2**, the key correlation between C-2–C-8' was absent in the HMBC spectra. Nevertheless, the absence of the corresponding ¹H signals supported them to be C-2-guaninylated cWF (**3**, guanitryptmycin D2) and cWW (**4**, guanitryptmycin D3), respectively (Scheme 2).

Table 1. NMR Data of Compounds 1–3 in DMSO-*d*₆

| position | <i>cyclo</i> -(1- <i>Trp</i> - <i>L</i> - <i>Tyr</i>) (1) | | guanitrypmycin D1 (2) | | guanitrypmycin D2 (3) | |
|----------|--|--|-----------------------|--|-----------------------|--|
| | δ_C , type | δ_H , mult. (J in Hz) | δ_C , type | δ_H , mult. (J in Hz) | δ_C , type | δ_H , mult. (J in Hz) |
| 1 | | 10.87, br s | | 7.72, br s | | 7.65, br s |
| 2 | 124.3, CH | 6.99, d (2.7) | 123.8, C | | 124.0, C | |
| 3 | 108.9, C | | 114.1, C | | 114.2, C | |
| 4 | 118.7, CH | 7.48, d (8.1) | 119.2, CH | 7.52, d (7.9) | 119.4, CH | 7.49, d (7.8) |
| 5 | 118.4, CH | 6.99, td (7.6, 1.1) | 121.5, CH | 7.21, td (7.6, 0.8) | 121.7, CH | 7.16 ^b |
| 6 | 120.8, CH | 7.07, td (7.6, 1.1) | 123.5, CH | 7.31, td (7.6, 1.0) | 123.7, CH | 7.28, t (7.5) |
| 7 | 111.3, CH | 7.32, d (8.1) | 114.4, CH | 8.48, d (8.0) | 114.6, CH | 8.44, d (7.4) |
| 8 | 136.1, C | | 134.7, C | | 134.9, C | |
| 9 | 127.5, C | | 129.2, C | | 129.3, C | |
| 10 | 29.9, CH ₂ | 2.81, dd (14.4, 4.0); 2.42, dd (14.4, 6.5) | 30.0, CH ₂ | 2.79, dd (14.6, 3.7); 2.10, dd (14.6, 6.3) | 29.8, CH ₂ | 2.73, dd (13.6, 3.5); 2.09, dd (13.6, 3.5) |
| 11 | 55.2, CH | 3.95, dt (6.5, 4.0) | 54.7, CH | 3.94 ^{at} | 54.9, CH | 3.93, t (3.5) |
| 12 | | 7.78, d (2.5) | | 7.95, d (2.6) | | 7.96 ^c |
| 13 | 166.2, C | | 166.2, C | | 166.3, C | |
| 14 | 55.9, CH | 3.79, dt (7.0, 4.5) | 55.9, CH | 3.94 ^{at} | 55.7, CH | 3.99, t (4.6) |
| 15 | | 7.60, d (2.5) | | 7.93, d (1.7) | | 7.96 ^c |
| 16 | 166.7, C | | 166.5, C | | 166.7, C | |
| 17 | 39.0, CH ₂ | 2.46, dd (13.6, 4.5); 1.83, dd (13.6, 7.0) | 38.9, CH ₂ | 2.58, dd (13.6, 4.6); 2.36, dd (13.6, 5.2) | 39.8, CH ₂ | 2.61, dd (13.3, 4.6); 2.42, dd (13.3, 4.6) |
| 18 | 126.4, C | | 126.2, C | | 136.3, C | |
| 19, 23 | 130.7, CH | 6.54, d (8.5) | 131.1, CH | 6.76, d (8.5) | 130.1, CH | 6.90, d (6.9) |
| 20, 22 | 114.9, CH | 6.59, d (8.5) | 115.1, CH | 6.65, d (8.5) | 128.4, CH | 7.16 ^b |
| 21 | 156.0, C | | 156.2, C | | 126.8, CH | 7.16 ^b |
| 21-OH | | 9.13, s | | 9.20, br s | | 10.63, s |
| 1' | | | | 10.70, br s | | |
| 2' | | | 151.6, C | | 156.3, C | |
| 4' | | | 146.0, C | | 146.0, C | |
| 5' | | | 106.0, C | | 106.3, C | |
| 6' | | | 159.3, C | | 159.5, C | |
| 8' | | | 140.5, C | | 141.0, C | |
| 9' | | | | 12.98, br s | | 12.94, br s |
| 10' | | | | 6.33, br s | | 6.27, br s |

^{a-c}Signals with the same letter are overlapping with each other. See spectra in the SI for numbering.

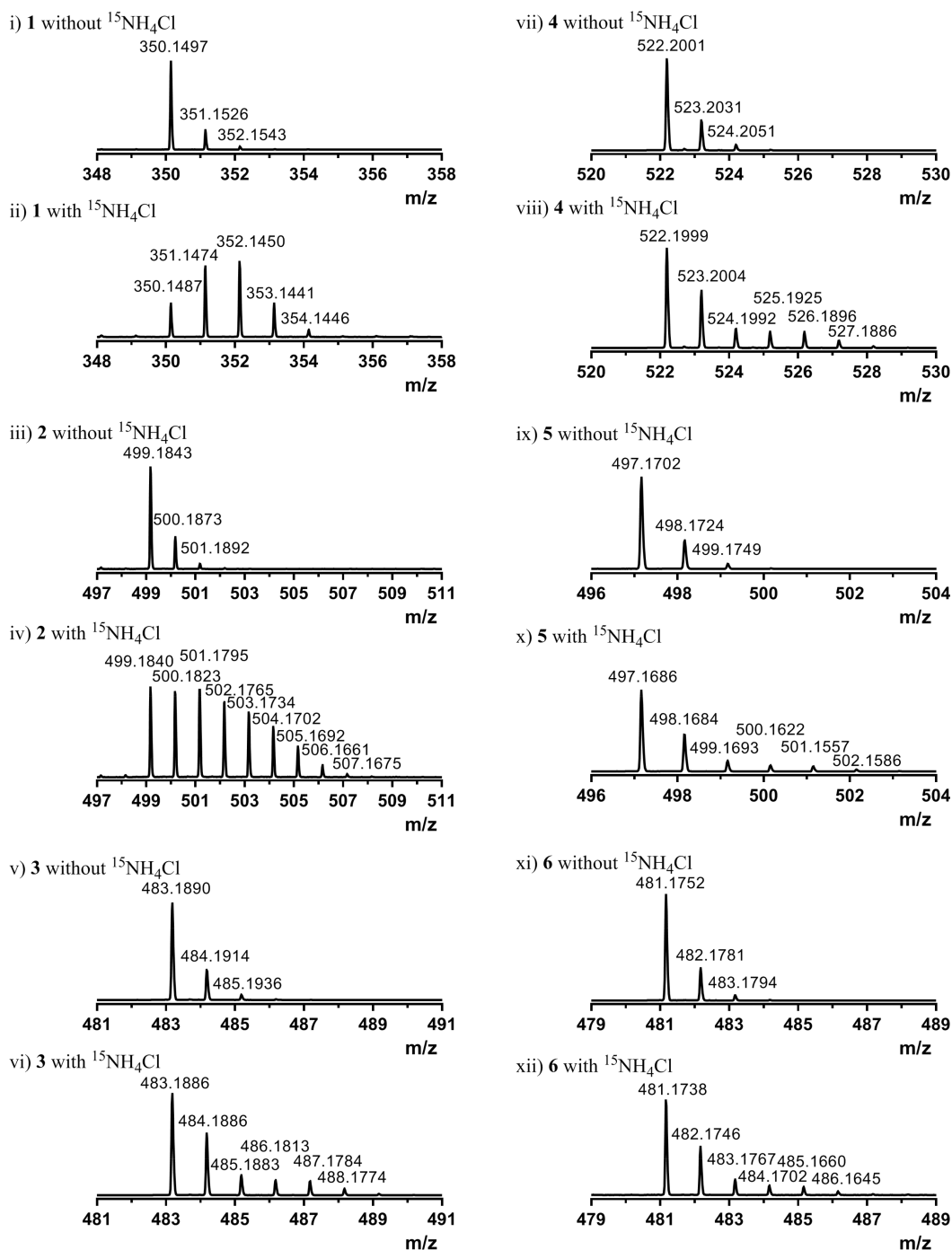
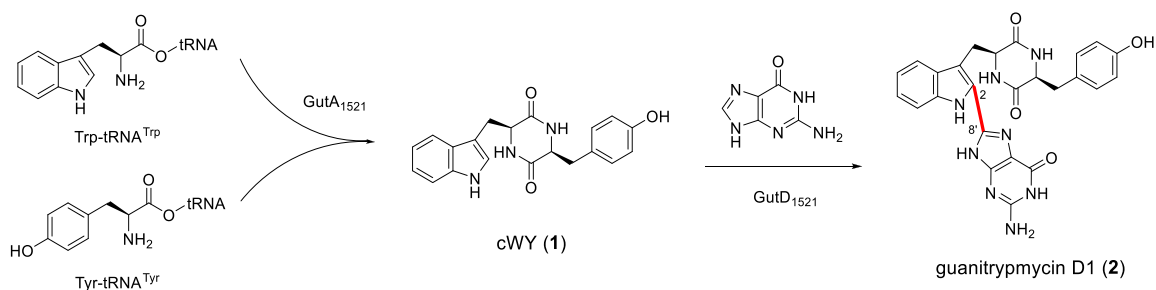
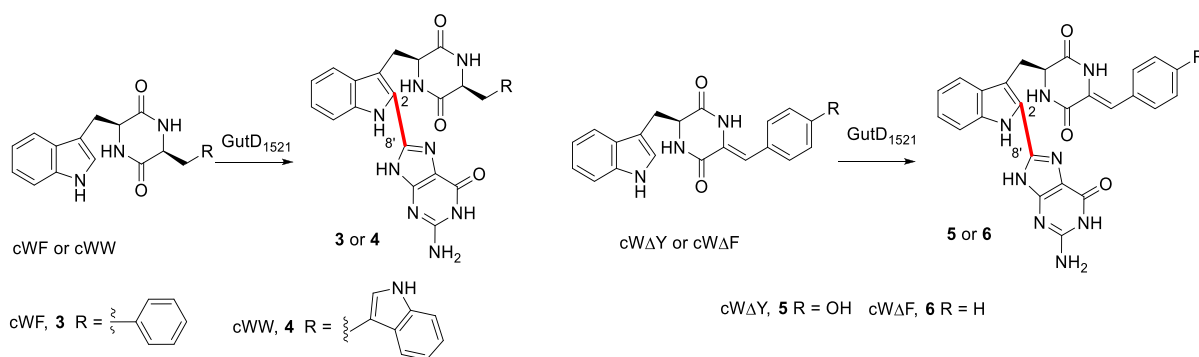
Scheme 1. Biosynthesis of Guanitrypmycin D1 in *Streptomyces* sp. NRRL S-1521Figure 3. Isotope pattern of the $[M + H]^+$ ions of products 1–6 detected in cultures with and without ¹⁵NH₄Cl.

Table 2. NMR Data of Compounds 4–6 in DMSO-*d*₆

| position | guanitarymycin D3 (4) | | guanitarymycin D4 (5) | | guanitarymycin D5 (6) | |
|----------|-----------------------|--|-----------------------|--|-----------------------|--|
| | δ_C , type | δ_H , mult. (J in Hz) | δ_C , type | δ_H , mult. (J in Hz) | δ_C , type | δ_H , mult. (J in Hz) |
| 1 | | 7.63, br s | | 7.82, br s | | 7.84, br s |
| 2 | 124.0, C | | 124.7, C | | 124.8, C | |
| 3 | 114.2, C | | 112.9, C | | 112.7, C | |
| 4 | 119.3, CH | 7.30 ^a | 119.3, CH | 7.65, d (7.8) | 119.3, CH | 7.66, d (7.8) |
| 5 | 121.6, CH | 7.16, t (7.5) | 121.5, CH | 7.21, t (7.4) | 121.5, CH | 7.23, t (7.3) |
| 6 | 123.5, CH | 7.30 ^a | 123.5, CH | 7.27, t (7.5) | 123.5, CH | 7.28, t (7.6) |
| 7 | 114.2, CH | 8.43, d (7.3) | 114.5, CH | 8.37, d (9.8) | 114.4, CH | 8.37, d (8.1) |
| 8 | 134.7, C | | 134.7, C | | 134.7, C | |
| 9 | 129.3, C | | 129.4, C | | 129.4, C | |
| 10 | 29.8, CH ₂ | 2.72, dd (14.4, 3.3); 2.11, dd (14.4, 7.0) | 30.0, CH ₂ | 3.31, dd (14.5, 4.1); 3.14, dd (14.5, 4.6) | 30.0, CH ₂ | 3.33, dd (14.6, 4.6); 3.14, dd (14.6, 4.6) |
| 11 | 55.2, CH | 3.98, td (6.0, 3.2) | 55.9, CH | 4.31, q (4.3) | 56.0, CH | 4.36, q (3.9) |
| 12 | | 8.00, d (2.6) | | 9.47, br s | | 9.66, s |
| 13 | 166.9, C | | 160.6, C | | 160.1, C | |
| 14 | 55.6, CH | 4.01, t (6.3) | 124.1, C | | 126.3, C | |
| 15 | | 8.01, d (2.7) | | 8.35, d (2.5) | | 8.48, d (2.5) |
| 16 | 166.6, C | | 166.5, C | | 166.6, C | |
| 17 | 30.1, CH ₂ | 2.84, dd (14.3, 4.0); 2.57, dd (14.3, 5.8) | 114.6, CH | 6.14, s | 113.6, CH | 6.16, s |
| 18 | | 10.89, s | 123.9, C | | 133.1, C | |
| 19 | 124.4, CH | 6.83, s | 130.5, CH | 6.73, d (8.1) | 128.8, CH | 6.82, t (3.5) |
| 20 | 108.8, C | | 115.3, CH | 6.55, d (8.4) | 128.3, CH | 7.12 ^b |
| 21 | 118.7, CH | 7.40, d (7.9) | 157.2, C | | 127.5, CH | 7.12 ^b |
| 21-OH | | | | 9.63, br s | | |
| 22 | 118.6, CH | 6.92, t (7.5) | 115.3, CH | 6.55, d (8.4) | 128.3, CH | 7.12 ^b |
| 23 | 121.1, CH | 7.03, t (7.6) | 130.5, CH | 6.73, d (8.1) | 128.8, CH | 6.82, t (3.5) |
| 24 | 111.5, CH | 7.30 ^a | | | | |
| 25 | 136.1, C | | | | | |
| 26 | 127.6, C | | | | | |
| 1' | | 10.80, br s | | 10.72, br s | | 10.76, br s |
| 2' | 151.8, C | | 153.3, C | | 153.3, C | |
| 4' | 146.0, C | | 131.1, C | | 145.1, C | |
| 5' | 106.3, C | | 115.2, C | | 111.3, C | |
| 6' | 156.3, C | | 156.3, C | | 155.3, C | |
| 8' | 140.6, C | | 126.2, C | | 140.9, C | |
| 9' | | 12.91, br s | | 12.95, br s | | 12.87, br s |
| 10' | | 6.42, br s | | 6.33, br s | | 6.36, br s |

^{a, b} Signals with the same letter are overlapping with each other. See spectra in the SI for numbering.

Scheme 2. *In Vivo* Conversion of CDPs and Derivatives by GutD₁₅₂₁

Dehydrogenation at the DKP ring catalyzed by cyclo-dipeptide oxidases (CDOs) exists widely in CDPs-related biosynthetic pathways.²⁶ Because GutD₁₅₂₁ can use cWF, cWY, and cWW as substrates for coupling with guanine, we further tested its acceptance for the dehydrogenated forms of these well-converted CDPs. Due to difficulties in obtaining *cyclo*-(L-Trp-ΔTrp) (cWΔW), only *cyclo*-(L-Trp-ΔTyr) (cWΔY) and *cyclo*-(L-Trp-ΔPhe) (cWΔF) were prepared by large-scale enzyme assays with the functionally characterized CDO Nads_1146/1147 for biotransformation with the *gutD*₁₅₂₁ transformant.²⁶ After incubation for 7 days, the two dehydrogenated CDPs were converted to peaks 5 and 6 with conversion yields of 79 ± 1% and 98 ± 1%, respectively (Figure 2x and xi). Structure elucidation by detailed interpretation of NMR data confirmed both compounds as guaninylated derivatives at the C-2 of the indole ring (Table 2, Scheme 2, and Figures S19–S28), i.e., guanitrypmycin D5 (5) from cWΔY and guanitrypmycin D5 (6) from cWΔF. Addition of cWF, cWW, cWΔY, and cWΔF to cultures of *S. albus* harboring *gutD*₁₅₂₁ containing ¹⁵NH₄Cl led to detection of compounds 3–6 with incorporation of five ¹⁵N atoms, respectively (Figure 3), further supporting their structures suggested by NMR (Figures S9–S28) and ECD analyses (Figure S29).

Compounds 2–6 were subsequently tested for their antibacterial activities against *E. coli* ATCC 25922 and DH5α, *Enterococcus faecalis* DSM2570, *Klebsiella pneumoniae* DSM26371, *Bacillus subtilis* NCIB 3610 and BSB 01, *Bacillus circulans* NRRL B-380, NRRL B-14032, and NRRL NRS-1108, *Staphylococcus aureus* ATCC 29213, *Staphylococcus delphini* DSM20771, and *Pseudomonas aeruginosa* ATCC 27853. Unfortunately, no inhibitory activity was observed.

In conclusion, a *cdps-p450*-associated gene cluster was identified in *Streptomyces* sp. NRRL S-1521 by genome mining. Introduction of this BGC into a heterologous expression host led to identification of guanitrypmycin D1 (2), a new guaninylated DKP. Biotransformation experiments confirmed that the P450 catalyzes the key step for the guanitrypmycin D1 formation, that is, the attachment of the guaninyl residue at the indole ring via the C-2–C-8' linkage. This differs from the reported guanine–DKP connections. Further precursor incubation experiments unveiled that GutD₁₅₂₁ can also accept other tryptophan-containing CDPs with a second aromatic side chain as substrates. In addition, GutD₁₅₂₁ can use the dehydrogenated forms of two CDPs for guaninyl decoration, which was achieved by a combination with a cyclodipeptide oxidase. Combination of such intriguing P450 enzymes with

other modification enzymes is a promising strategy to increase structure diversity of DKP derivatives.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were measured with an A KRÜSS P3000 polarimeter at 20 °C by using the D-line of the sodium lamp at λ = 589.3 nm. Prior to the measurement, the polarimeter was calibrated with MeOH–H₂O (1:1, v/v). UV and circular dichroism spectra were taken on a J-1500 CD spectrometer (Jasco). The samples were dissolved in MeOH–H₂O (1:1, v/v) and measured in the range of 200–400 nm by using a 1 mm path length quartz cuvette. IR spectra were acquired on a Bruker ALPHA FTIR spectrometer. NMR spectra were recorded on a JEOL ECA-500 or a Bruker AVIII 500 spectrometer. All samples were dissolved in 200 μL of DMSO-*d*₆ and filled in Wilmad 3 mm tubes (Rototec-Spintec, Bad Wildbad, Germany). The ¹³C and the ¹H–¹³C HMBC spectra were recorded on a Bruker AVIII 500 spectrometer installed with a 5 mm cryo BBO probe Prodigy, with 64 000 and 64 transients, respectively. The spectra were processed with MestReNova 9.0.0 (Metrelab). Chemical shifts are referred to those of DMSO-*d*₆ (δ_H 2.50 and δ_C 39.5). HRMS analysis was performed on an Agilent 1260 HPLC system equipped with a microTOF-Q III spectrometer (Bruker) using a VDSpher PUR 100 C18-M-SE column (150 × 2 mm, 3 μm) (VDS Optilab Chromatographie Technik). Electrospray positive ionization mode was selected for determination of the exact masses. MS conditions were set as described previously.²⁷ Semi-preparative HPLC was performed on the same equipment with a VDSpher PUR 100 C18-M-SE column (250 × 10 mm, 5 μm) (VDS Optilab Chromatographie Technik) for detection of the desired substance in the fractions.

Chemicals. *cyclo*-(L-Trp-L-Leu), *cyclo*-(L-Trp-L-Trp), *cyclo*-(L-Trp-L-Phe), *cyclo*-(L-Trp-L-Tyr), and *cyclo*-(D-Trp-L-Tyr) were purchased from Bachem. *cyclo*-(L-Trp-L-His), *cyclo*-(L-Trp-L-Pro), and *cyclo*-(L-Trp-L-Ala) were synthesized according to the methods published previously.²⁸ *cyclo*-(L-Trp-L-Met) was obtained from a previous study.²⁹ The dehydrogenated CDP derivatives used in this study were prepared by using the cyclodipeptide oxidase Ndas_1146/1147 according to the method described in a previous work.²⁶

Bacterial Strains and Growth Conditions. *Streptomyces* sp. NRRL S-1521 was kindly provided by the ARS Culture Collection (NRRL) and cultivated on MS plates (mannitol 20.0 g/L, soya flour 20.0 g/L, agar 20.0 g/L). *Streptomyces albus* J1074³⁰ was kindly gifted by Prof. Dr. Andriy Luzhetskyy (Saarland University). *S. albus* J1074 and the generated exconjugants were maintained on MS at 28 °C for sporulation. For secondary metabolite production, *S. albus* J1074 transformants were cultivated in liquid modified R5 medium (sucrose 103.0 g/L, glucose 10.0 g/L, yeast extract 5.0 g/L, MgCl₂·6H₂O 10.12 g/L, K₂SO₄ 0.25 g/L, Difco casamino acids 0.1 g/L, MOPS 21.0 g/L, trace element solution 2 mL/L, pH 7.2) at 28 °C for 7 days.

Computer-Assisted Sequence Analysis. Nucleotide and amino acid sequences used in this study were downloaded from the NCBI databases (<http://www.ncbi.nlm.nih.gov>). The phylogenetic tree of

P450s (Figure S1) was created by MEGA version 7.0 (<http://www.megasoftware.net>).

Genetic Manipulation, PCR Amplification, and Gene Cloning. Strains and plasmids used in this study are listed in Table S2 and Table S3, respectively. Recombinant *E. coli* strains were cultivated in liquid or on solid lysogeny broth (LB) with 100 $\mu\text{g}/\text{mL}$ ampicillin, 50 $\mu\text{g}/\text{mL}$ kanamycin, 50 $\mu\text{g}/\text{mL}$ apramycin, or 25 $\mu\text{g}/\text{mL}$ chloramphenicol when necessary.

Genetic manipulation in *E. coli* was performed according to the protocols by Green and Sambrook.³¹ Genomic DNA isolation from *Streptomyces* was performed as described in the literature.³²

Genetic loci of interest were amplified by PCR from the genomic DNA of NRRL S-1521 by using primers listed in Table S3 and Phusion High-Fidelity DNA Polymerase from New England Biolabs. The generated PCR fragments were cloned into pGEM-T Easy vector (Promega), and the sequence integrity was confirmed by sequencing. Subsequently, the fragments were released with restriction endonucleases (see Table S3 for details) from pGEM-T Easy and ligated into pET28a (+) or pPWW50A,³³ which had been digested with the same enzymes, previously. The generated constructs (Table S3) were used for heterologous expression in *E. coli* BL21 (DE3) or *S. albus* J1074.

Expression of the CDPS Gene *gutA*₁₅₂₁ in *E. coli*. The generated construct pJL165 was introduced into *E. coli* BL21 (DE3) by transformation. For CDPS overproduction, 0.5 mL of 16 h overnight culture was used to inoculate 50 mL of LB medium containing 50 $\mu\text{g}/\text{mL}$ kanamycin. The culture was maintained at 200 rpm and 37 °C to an absorption at 600 nm of about 0.6. Isopropyl β -D-thiogalactopyranoside (IPTG) was then added to the cultures at the final concentration of 0.2 mM. After induction at 18 °C for 20 h, 1 mL of culture was extracted with the same volume of EtOAc three times. After that, the organic phases were combined and evaporated, and the dried residues were dissolved in 200 μL of MeOH. A 5 μL amount of these samples was subjected to LC-MS analysis.

Heterologous Gene Expression in *Streptomyces albus* J1074 and Cultivation for Secondary Metabolite Production. The constructed plasmids harboring *gutD*₁₅₂₁ or the gene cluster *gut(AD)*₁₅₂₁ were transformed into the nonmethylated strain *E. coli* ET12567/pUZ8002 and then conjugated with *S. albus* J1074. The positive conjugants were selected by apramycin resistance and further confirmed by PCR amplification. The spores of the *S. albus* J1074 transformants were inoculated into 50 mL of modified R5 liquid media supplied with 50 $\mu\text{g}/\text{mL}$ of apramycin in 250 mL baffled flasks and cultured at 28 °C and 200 rpm for 7 days. Then the cultures were treated as mentioned above and sent for LC-MS analysis.

Biotransformation and Generation of Guaninylated DKPs. The *S. albus* J1074 transformant harboring *gutD*₁₅₂₁ was cultivated in modified R5 medium at 28 °C and 200 rpm for 2 days. Tryptophan-containing CDPs or dehydrogenated derivatives were separately added to 10 mL of these precultures at the concentration of 100 μM . After cultivation for an additional 5 days, the cultures were extracted with EtOAc and analyzed on LC-MS. Three independent experiments were carried out, and the conversion yields were calculated by using the isolated products as standards.

Cultivation of the *S. albus* Transformants in Media Containing ¹⁵NH₄Cl. The *S. albus* transformant harboring *gut(AD)*₁₅₂₁ was cultivated in 10 mL of modified R5 medium containing 10 mg of ¹⁵NH₄Cl at 28 °C for 7 days. The EtOAc extract was subsequently analyzed on LC-MS for detection of 1 and 2. For labeling 3–6, the *S. albus* transformant harboring *gutD*₁₅₂₁ was cultivated in the same medium under the same conditions for 30 h. After addition of *cyclo*-(L-Trp-L-Phe), *cyclo*-(L-Trp-L-Trp), *cyclo*-(L-Trp- Δ Tyr), and *cyclo*-(L-Trp- Δ Phe) separately, the cultures were maintained for additional 6 days and analyzed by LC-MS. The isotope patterns of the [M + H]⁺ ions of products 1–6 isolated from cultures with and without ¹⁵NH₄Cl are illustrated in Figure 3.

LC-MS Analysis. For secondary metabolite analysis, a linear gradient of 5–100% MeCN in H₂O, both containing 0.1% HCOOH, in 10 min and a flow rate at 0.3 mL/min were used. The column was then washed with 100% MeCN containing 0.1% HCOOH for 5 min and equilibrated with 5% MeCN in H₂O for 5 min.

Extraction and Isolation of Secondary Metabolites. For structure elucidation of the accumulated products, the *S. albus* J1074 transformant harboring *gut(AD)*₁₅₂₁ was fermented in 3 L of modified R5 medium at 28 °C for 7 days. The culture supernatants were collected and extracted with an equal volume of EtOAc three times. Then, the EtOAc phases were evaporated to dryness, dissolved in MeOH, and centrifuged. The precipitate was dispersed in MeOH, centrifuged, and repeated twice. The precipitated fractions were dissolved in DMSO and further purified on an Agilent HPLC 1260 series instrument by using a semipreparative VDSpher PUR 100 C18-M-SE column (250 \times 10 mm, 5 μm). The flow rate was set to 2.0 mL/min. Compound 2 (30 mg) was purified with 20% MeCN in H₂O.

For the guaninylated DKP derivatives generated by biotransformation, the crude extracts were obtained by extraction with EtOAc as described above. Compounds 3 (30 mg), 4 (25 mg), 5 (8 mg), and 6 (15 mg) were further purified on an Agilent HPLC 1260 series with 20–30% MeCN in H₂O.

Physicochemical Properties of the Identified Products. *Cyclo*-(L-Trp-L-Tyr) (1): white powder; ¹H and ¹³C NMR data, Table 1; HR-ESI-MS *m/z* 350.1497 [M + H]⁺ (calcd for C₂₀H₂₀N₃O₃, 350.1499).

Guanitrypmycin D1 (2): white powder; [α]_D²⁰ +280 (*c* 0.07, MeOH–H₂O, 1:1, v/v); UV (MeOH–H₂O, 1:1, v/v) λ_{max} (log ϵ) 310 (3.53), 270 (4.60), 250 (4.93) nm; ECD (1.30 mM, MeOH–H₂O, 1:1, v/v) λ_{max} ($\Delta\epsilon$) 352 (+4.72), 322 (–6.36), 293 (+6.39), 262 (–2.12), 242 (+3.74), 224 (+1.43), 208 (–0.67) nm; IR ν_{max} 3160, 2918, 1662, 1643, 1558, 1511, 1454, 1436, 1367, 1331, 1014, 772, 741 cm^{–1}; ¹H and ¹³C NMR data, Table 1; HR-ESI-MS *m/z* 499.1843 [M + H]⁺ (calcd for C₂₅H₂₃N₈O₄, 499.1837).

Guanitrypmycin D2 (3): white powder; [α]_D²⁰ +320 (*c* 0.08, MeOH–H₂O, 1:1, v/v); UV (MeOH–H₂O, 1:1, v/v) λ_{max} (log ϵ) 309 (3.74), 268 (5.01), 251 (6.29) nm; ECD (1.62 mM, MeOH–H₂O, 1:1, v/v) λ_{max} ($\Delta\epsilon$) 342 (–0.89), 330 (–0.92), 304 (+0.71), 275 (–0.12), 251 (+0.34), 217 (–5.59) nm; IR ν_{max} 3158, 3111, 2871, 2725, 1683, 1656, 1451, 1440, 1372, 1333, 745, 705 cm^{–1}; ¹H and ¹³C NMR data, Table 1; HR-ESI-MS *m/z* 483.1890 [M + H]⁺ (calcd for C₂₅H₂₃N₈O₃, 483.1888).

Guanitrypmycin D3 (4): white powder; [α]_D²⁰ +560 (*c* 0.06, MeOH–H₂O, 1:1, v/v); UV (MeOH–H₂O, 1:1, v/v) λ_{max} (log ϵ) 311 (0.87), 268 (2.44), 259 (2.57) nm; ECD (1.09 mM, MeOH–H₂O, 1:1, v/v) λ_{max} ($\Delta\epsilon$) 342 (–1.19), 330 (–1.27), 303 (+0.91), 267 (–0.02), 252 (+0.32), 217 (–8.32) nm; IR ν_{max} 3047, 2912, 1674, 1567, 1508, 1454, 1433, 1364, 1324, 1012, 741, 693 cm^{–1}; ¹H and ¹³C NMR data, Table 2; HR-ESI-MS *m/z* 522.2001 [M + H]⁺ (calcd for C₂₇H₂₄N₉O₃, 522.1997).

Guanitrypmycin D4 (5): white powder; [α]_D²⁰ –250 (*c* 0.10, MeOH–H₂O, 1:1, v/v); UV (MeOH–H₂O, 1:1, v/v) λ_{max} (log ϵ) 309 (20.45), 270 (14.14), 243 (19.10) nm; ECD (1.99 mM, MeOH–H₂O, 1:1, v/v) λ_{max} ($\Delta\epsilon$) 333 (–14.57), 308 (–0.63), 296 (–2.01), 279 (+0.41), 267 (–0.35), 251 (+0.82), 221 (–0.77) nm; IR ν_{max} 3088, 2912, 1673, 1632, 1605, 1567, 1511, 1428, 1366, 1336, 1010, 742, 683 cm^{–1}; ¹H and ¹³C NMR data, Table 2; HR-ESI-MS *m/z* 497.1702 [M + H]⁺ (calcd for C₂₃H₂₁N₈O₄, 497.1680).

Guanitrypmycin D5 (6): white powder; [α]_D²⁰ –160 (*c* 0.07, MeOH–H₂O, 1:1, v/v); UV (MeOH–H₂O, 1:1, v/v) λ_{max} (log ϵ) 303 (17.28), 270 (14.10), 248 (16.47) nm; ECD (1.44 mM, MeOH–H₂O, 1:1, v/v) λ_{max} ($\Delta\epsilon$) 321 (–12.93), 275 (–1.31), 266 (–1.64), 252 (–1.09), 230 (–5.93) nm; IR ν_{max} 3051, 2912, 1677, 1631, 1565, 1508, 1429, 1365, 1335, 1014, 742, 689 cm^{–1}; ¹H and ¹³C NMR data, Table 2; HR-ESI-MS *m/z* 481.1752 [M + H]⁺ (calcd for C₂₅H₂₁N₈O₃, 481.1731).

Antibacterial Assays. For compounds 2–6, the antibacterial activities against different bacterial strains were conducted using a disk diffusion method.³⁴ The compounds were dissolved in DMSO to prepare stock solutions at 2 mg/mL. A 1 mL overnight culture of each tested strain was mixed with 100 mL of LB medium and 1.5 g of agar to prepare the testing plates. A 5 μL amount of the stock solutions was dropped onto paper disks with a 5 mm diameter on the agar plates. After incubation at 37 °C overnight, the inhibition zones

around the disks were visually observed. Kanamycin and chloramphenicol were used as positive and DMSO as negative controls.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jnatprod.2c00787>.

Strains, plasmids, primers, cluster information, NMR and ECD spectra of compounds 1–6, HPLC, and LC-MS chromatograms (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Shu-Ming Li – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany; orcid.org/0000-0003-4583-2655; Email: shuming.li@staff.uni-marburg.de

Authors

Jing Liu – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Yiling Yang – Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, 35037 Marburg, Germany

Xiulan Xie – Fachbereich Chemie, Philipps-Universität Marburg, 35032 Marburg, Germany

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jnatprod.2c00787>

Author Contributions

J.L. and Y.Y. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the ARS Culture Collection (NRRL) for providing *Streptomyces* sp. NRRL S-1521 and L. Ludwig-Radtke and S. Newel (Fachbereich Pharmazie, University of Marburg) for LC-MS and NMR analyses, respectively. This project was funded in part by the Deutsche Forschungsgemeinschaft (DFG, Li844/14-1 and INST 160/620-1). J.L. (201608310118) and Y.Y. (201808530447) are the scholarship recipients from the China Scholarship Council.

■ REFERENCES

- (1) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T. *Nat. Rev. Drug Discovery* **2021**, *20*, 200–216.
- (2) Challis, G. L. *Microbiology* **2008**, *154*, 1555–1569.
- (3) Senges, C. H. R.; Al-Dilaimi, A.; Marchbank, D. H.; Wibberg, D.; Winkler, A.; Haltli, B.; Nowrousian, M.; Kalinowski, J.; Kerr, R. G.; Bandow, J. E. *Proc. Natl. Acad. Sci. U. S. A* **2018**, *115*, 2490–2495.
- (4) Borgman, P.; Lopez, R. D.; Lane, A. L. *Org. Biomol. Chem.* **2019**, *17*, 2305–2314.
- (5) Harken, L.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2021**, *105*, 2277–2285.
- (6) Liu, J.; Li, S.-M. *ChemBiochem* **2022**, No. e202200502.
- (7) Li, S.-M. *Nat. Prod. Rep.* **2010**, *27*, 57–78.
- (8) Borthwick, A. D. *Chem. Rev.* **2012**, *112*, 3641–3716.
- (9) Walsh, C. T. *Nat. Prod. Rep.* **2016**, *33*, 127–135.
- (10) Gondry, M.; Sauguet, L.; Belin, P.; Thai, R.; Amouroux, R.; Tellier, C.; Tiphile, K.; Jacquet, M.; Braud, S.; Courcon, M. *Nat. Chem. Biol.* **2009**, *5*, 414–420.

- (11) Gondry, M.; Jacques, I. B.; Thai, R.; Babin, M.; Canu, N.; Seguin, J.; Belin, P.; Pernodet, J. L.; Moutiez, M. *Front. Microbiol.* **2018**, *9*, 46.
- (12) Du, L.; Lou, L. *Nat. Prod. Rep.* **2010**, *27*, 255–278.
- (13) Lautru, S.; Gondry, M.; Genet, R.; Pernodet, J. L. *Chem. Biol.* **2002**, *9*, 1355–1364.
- (14) Belin, P.; Moutiez, M.; Lautru, S.; Seguin, J.; Pernodet, J. L.; Gondry, M. *Nat. Prod. Rep.* **2012**, *29*, 961–979.
- (15) Sauguet, L.; Moutiez, M.; Li, Y.; Belin, P.; Seguin, J.; Le Du, M. H.; Thai, R.; Masson, C.; Fonvielle, M.; Pernodet, J. L. *Nucleic Acids Res.* **2011**, *39*, 4475–4489.
- (16) Malit, J. J. L.; Liu, W.; Cheng, A.; Saha, S.; Liu, L. L.; Qian, P. Y. *Org. Lett.* **2021**, *23*, 6601.
- (17) Yu, H.; Xie, X.; Li, S.-M. *Org. Lett.* **2018**, *20*, 4921–4925.
- (18) Liu, J.; Xie, X.; Li, S.-M. *Angew. Chem., Int. Ed. Engl.* **2019**, *58*, 11534–11540.
- (19) Yu, H.; Xie, X.; Li, S.-M. *Org. Lett.* **2019**, *21*, 9104–9108.
- (20) Harken, L.; Liu, J.; Kreuz, O.; Berger, R.; Li, S.-M. *ACS Catal.* **2022**, *12*, 648–654.
- (21) Liu, J.; Harken, L.; Yang, Y.; Xie, X.; Li, S.-M. *Angew. Chem., Int. Ed. Engl.* **2022**, *61*, No. e202200377.
- (22) Yu, H.; Li, S.-M. *Org. Lett.* **2019**, *21*, 7094–7098.
- (23) Marzilli, L. D.; de Castro, B.; Solorzano, C. J. *Am. Chem. Soc.* **1982**, *104*, 461–466.
- (24) Teixeira, M.; Santana, L.; Uriarte, E. *Magn. Reson. Chem.* **1997**, *35*, 806–807.
- (25) Shi, J.; Xu, X.; Zhao, E. J.; Zhang, B.; Li, W.; Zhao, Y.; Jiao, R. H.; Tan, R. X.; Ge, H. M. *Org. Lett.* **2019**, *21*, 6825–6829.
- (26) Mikulski, L.; Schäfer, J.; Brockmeyer, K.; Kraut, R.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 2523–2536.
- (27) Zhou, J.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2021**, *105*, 9181–9189.
- (28) Wollinsky, B.; Ludwig, L.; Hamacher, A.; Yu, X.; Kassack, M. U.; Li, S. M. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3866–3869.
- (29) Liu, J.; Yang, Y.; Harken, L.; Li, S.-M. *J. Nat. Prod.* **2021**, *84*, 3100–3109.
- (30) Zaborannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. *Bmc Genomics* **2014**, *15*, 97.
- (31) Green, M. R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2012.
- (32) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*, 2nd ed.; John Innes Foundation: Norwich, UK, 2000.
- (33) Zhu, Y.; Fu, P.; Lin, Q.; Zhang, G.; Zhang, H.; Li, S.; Ju, J.; Zhu, W.; Zhang, C. *Org. Lett.* **2012**, *14*, 2666–2669.
- (34) Balouiri, M.; Sadiki, M.; Ibsouda, S. K. *J. Pharm. Anal.* **2016**, *6*, 71–79.

Supporting Information

A *Streptomyces* Cytochrome P450 Enzyme Catalyzes Regiospecific C2-Guaninylation for the Synthesis of Diverse Guanitrypmycin Analog

Jing Liu,^{†,§} Yiling Yang,^{†,§} Xiulan Xie[‡], and Shu-Ming Li^{†*}

[†] Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie,
Philipps-Universität Marburg, Robert-Koch Straße 4, 35037 Marburg, Germany

[‡] Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Straße 4, 35032 Marburg,
Germany

Table of Contents

| | |
|---|----------|
| Supplementary Tables | 4 |
| Table S1. Comparison of <i>cdps</i> -containing gene clusters in four <i>Streptomyces</i> strains..... | 4 |
| Table S2. Bacterial strains used in this study. | 5 |
| Table S3. Cloning and expression constructs used in this study. | 5 |
| Supplementary Figures | 6 |
| Figure S1. Phylogenetic analysis of GutD ₁₅₂₁ with putative and known P450s. | 6 |
| Figure S2. ¹ H NMR spectrum of compound 1 in DMSO- <i>d</i> ₆ (500 MHz)..... | 7 |
| Figure S3. ¹³ C NMR spectrum of compound 1 in DMSO- <i>d</i> ₆ (125 MHz)..... | 7 |
| Figure S4. ¹ H NMR spectrum of compound 2 in DMSO- <i>d</i> ₆ (500 MHz)..... | 8 |
| Figure S5. ¹³ C NMR spectrum of compound 2 in DMSO- <i>d</i> ₆ (125 MHz)..... | 8 |
| Figure S6. HSQC spectrum of compound 2 in DMSO- <i>d</i> ₆ | 9 |
| Figure S7. ¹ H- ¹ H COSY spectrum of compound 2 in DMSO- <i>d</i> ₆ | 9 |
| Figure S8. HMBC spectrum of compound 2 in DMSO- <i>d</i> ₆ | 10 |
| Figure S9. ¹ H NMR spectrum of compound 3 in DMSO- <i>d</i> ₆ (500 MHz)..... | 10 |
| Figure S10. ¹³ C NMR spectrum of compound 3 in DMSO- <i>d</i> ₆ (125 MHz) | 11 |
| Figure S11. HSQC spectrum of compound 3 in DMSO- <i>d</i> ₆ | 11 |
| Figure S12. ¹ H- ¹ H COSY spectrum of compound 3 in DMSO- <i>d</i> ₆ | 12 |
| Figure S13. HMBC spectrum of compound 3 in DMSO- <i>d</i> ₆ | 12 |
| Figure S14. ¹ H NMR spectrum of compound 4 in DMSO- <i>d</i> ₆ (500 MHz)..... | 13 |
| Figure S15. ¹³ C NMR spectrum of compound 4 in DMSO- <i>d</i> ₆ (125 MHz) | 13 |
| Figure S16. HSQC spectrum of compound 4 in DMSO- <i>d</i> ₆ | 14 |
| Figure S17. ¹ H- ¹ H COSY spectrum of compound 4 in DMSO- <i>d</i> ₆ | 14 |
| Figure S18. HMBC spectrum of compound 4 in DMSO- <i>d</i> ₆ | 15 |
| Figure S19. ¹ H NMR spectrum of compound 5 in DMSO- <i>d</i> ₆ (500 MHz)..... | 15 |
| Figure S20. ¹³ C NMR spectrum of compound 5 in DMSO- <i>d</i> ₆ (125 MHz) | 16 |
| Figure S21. HSQC spectrum of compound 5 in DMSO- <i>d</i> ₆ | 16 |
| Figure S22. ¹ H- ¹ H COSY spectrum of compound 5 in DMSO- <i>d</i> ₆ | 17 |
| Figure S23. HMBC spectrum of compound 5 in DMSO- <i>d</i> ₆ | 17 |
| Figure S24. ¹ H NMR spectrum of compound 6 in DMSO- <i>d</i> ₆ (500 MHz)..... | 18 |
| Figure S25. ¹³ C NMR spectrum of compound 6 in DMSO- <i>d</i> ₆ (125 MHz) | 18 |
| Figure S26. HSQC spectrum of compound 6 in DMSO- <i>d</i> ₆ | 19 |
| Figure S27. ¹ H- ¹ H COSY spectrum of compound 6 in DMSO- <i>d</i> ₆ | 19 |

| | |
|---|-----------|
| Figure S28. HMBC spectrum of compound 6 in DMSO- <i>d</i> ₆ | 20 |
| Figure S29. Experimental ECD spectra of compounds 2 – 6 in MeOH – H ₂ O (1:1) | 20 |
| Figure S30. LC-MS analysis of <i>gutD</i> ₁₅₂₁ transformant after addition of two <i>cyclo</i> -(Trp-Tyr) isomers | 21 |
| Figure S31. LC-MS analysis of <i>gutD</i> ₁₅₂₁ transformant supplied with CDPs | 22 |
| Reference | 23 |

Supplementary Tables

Table S1. Comparison of *cdps*-containing gene clusters in four *Streptomyces* strains.



| <i>Streptomyces</i> sp. NRRL S-1521 | | <i>S. monomycini</i> NRRL B-24309 ¹ | | <i>S. purpureus</i> NRRL B-5737 ² | | <i>S. lavendulae</i> NRRL B-2774 ³ | | <i>S. cinnamoneus</i> DSM 40646 ⁴ | |
|--|-------------|--|-----------------------|--|-----------------------|---|-----------------------|--|-----------------------|
| Protein (Accession No.) | Length (aa) | Protein (Accession No.) | Sequence identity (%) | Protein (Accession No.) | Sequence identity (%) | Protein (Accession No.) | Sequence identity (%) | Protein (Accession No.) | Sequence identity (%) |
| GutA ₁₅₂₁ (WP_079106955.1) | 236 | GutA ₂₄₃₀₉ (WP_078624487.1) | 39 | GutA ₅₇₃₇ (WP_106959855.1) | 42 | GutA ₂₇₇₄ (WP_078950527.1) | 40 | GtmA (WP_071967254.1) | 43 |
| GutD ₁₅₂₁ (WP_062768047.1) | 395 | GutD ₂₄₃₀₉ (WP_050502760.1) | 53 | GutD ₅₇₃₇ (WP_019889608.1) | 63 | GutD ₂₇₇₄ (WP_051841251.1) | 60 | GtmD (WP_079274605.1) | 51 |

Table S2. Bacterial strains used in this study.

| Strain | Source | Cultivation media |
|-------------------------------------|------------|--------------------|
| <i>E. coli</i> DH5 α | Invitrogen | LB |
| <i>E. coli</i> ET12567/pUZ8002 | 5 | LB |
| <i>S. albus</i> J1074 | 6 | MS |
| <i>Streptomyces</i> sp. NRRL S-1521 | NRRL | MS and modified R5 |

NRRL: ARS Culture Collection

LB medium: tryptone 10.0 g/L, yeast extract 5.0 g/L, NaCl 10.0 g/L.

MS medium: mannitol 20.0 g/L, soya flour 20.0 g/L, agar 20.0 g/L.

Modified R5 medium: sucrose 103.0 g/L, glucose 10.0 g/L, yeast extract 5.0 g/L, MgCl₂.6H₂O 10.12 g/L, K₂SO₄ 0.25 g/L, Difco casaminoacids 0.1 g/L, MOPS 21.0 g/L, trace element solution 2 mL/L, pH 7.2

Table S3. Cloning and expression constructs used in this study.

| Gene | Primer sequences (5'-3') | Cloning construct | Expression vector | Cloning sites | Expression constructs |
|---------------------------------|--|-------------------|-------------------|---------------|-----------------------|
| <i>gutA</i> ₁₅₂₁ | <u>CATATGACCACAGCAGTAGAACTC</u> <u>AAGCTTTCAGTCACGGCACTCGACGATC</u> | pJL164 | pET28a (+) | NdeI/HindIII | pJL165 |
| <i>gut (AD)</i> ₁₅₂₁ | <u>CATATGACCACAGCAGTAGAACTC</u> <u>ACTAGTCACCACAGCACCGGAAGC</u> | pJL166 | pPWW50A | NdeI/SpeI | pJL167 |
| <i>gutD</i> ₁₅₂₁ | <u>CATATGAATCCCGGCAGAAAGCGGAC</u> <u>GGACCTCACCCACAGCACCGGAAGCC</u> | pJL168 | pPWW50A | NdeI/BamHI | pJL170 |

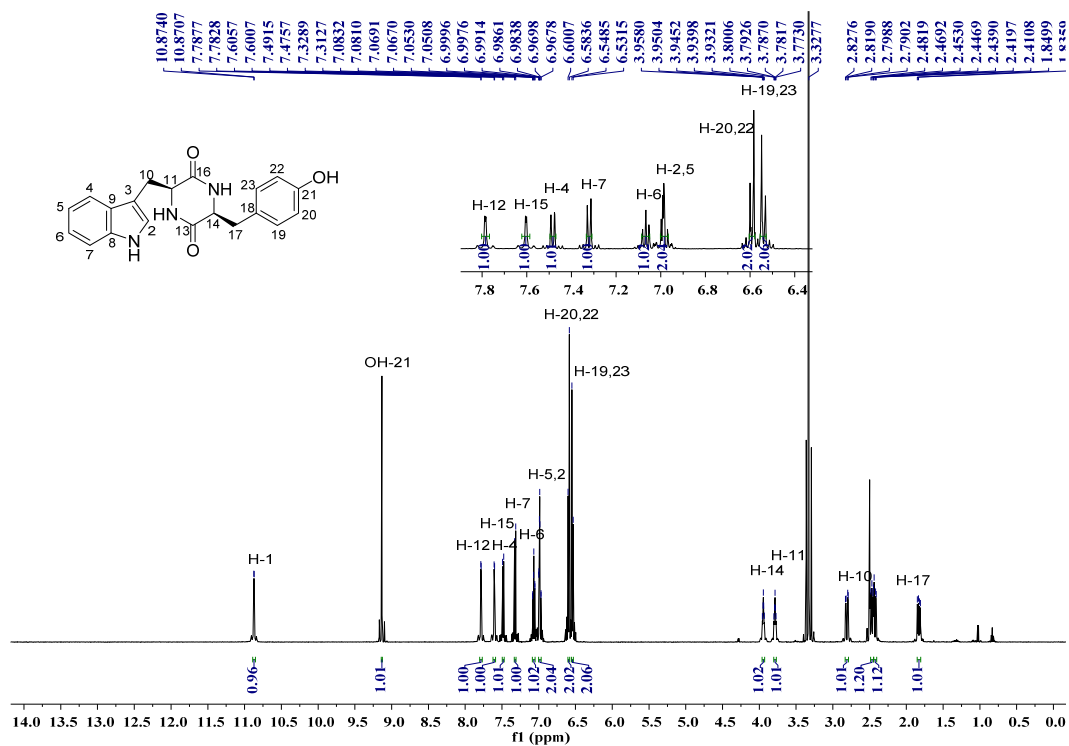


Figure S2. ^1H NMR spectrum of compound **1** in $\text{DMSO-}d_6$ (500 MHz)

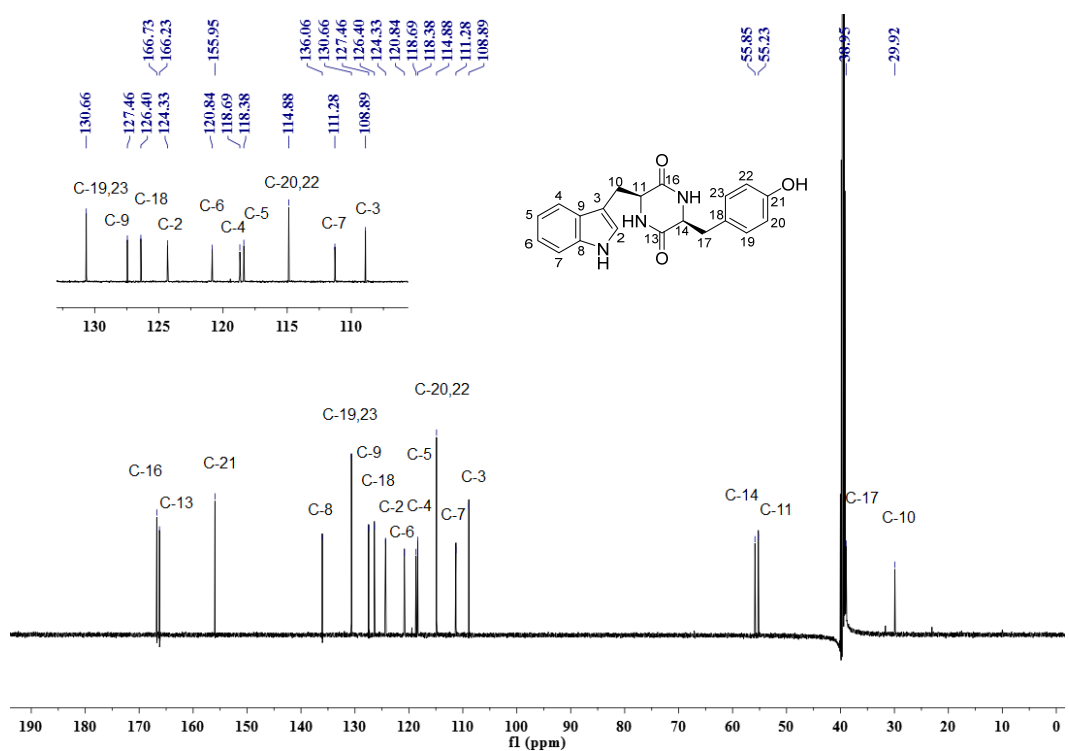


Figure S3. ^{13}C NMR spectrum of compound **1** in $\text{DMSO-}d_6$ (125 MHz)

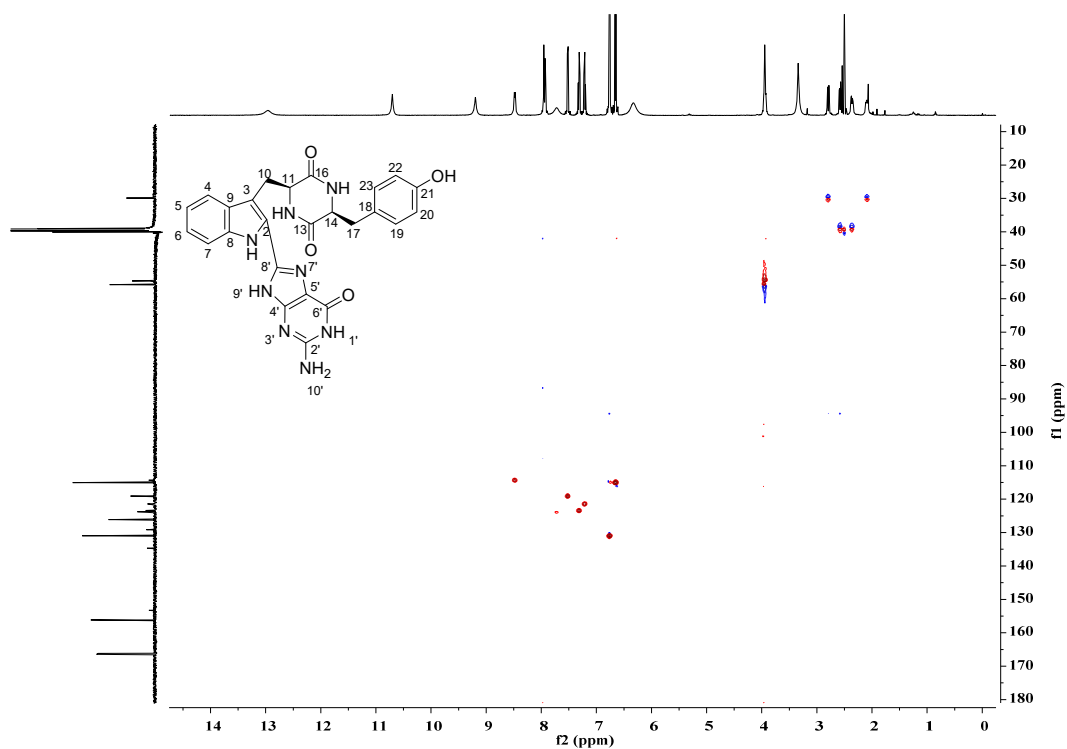


Figure S6. HSQC spectrum of compound **2** in DMSO- d_6

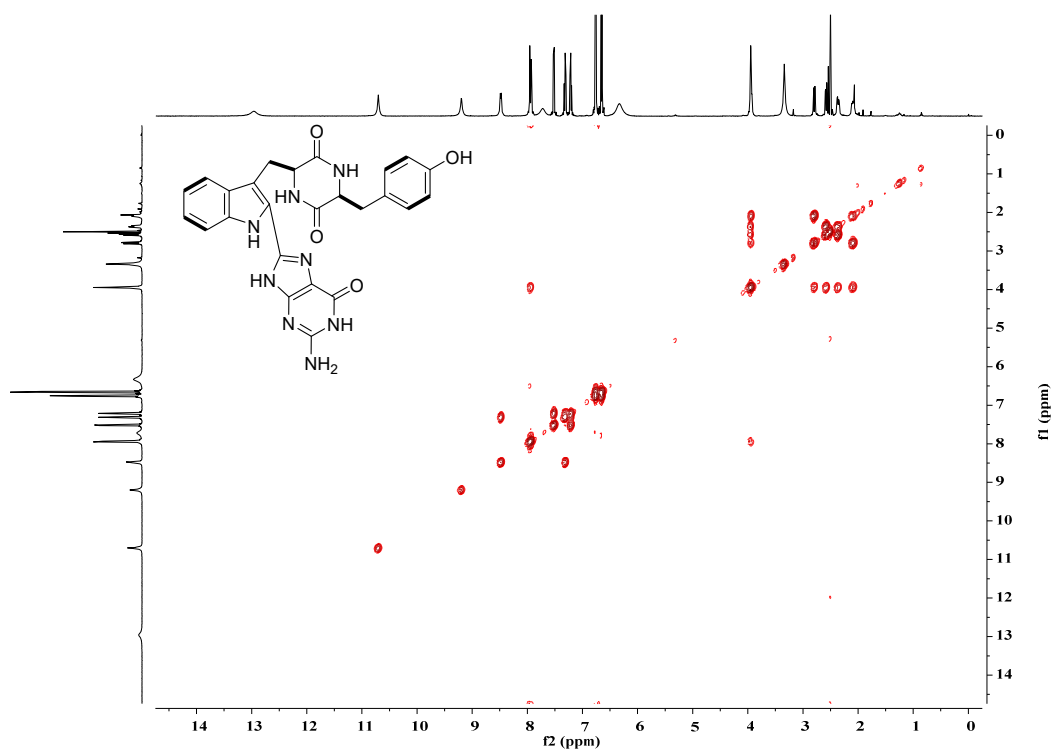


Figure S7. ^1H - ^1H COSY spectrum of compound **2** in DMSO- d_6

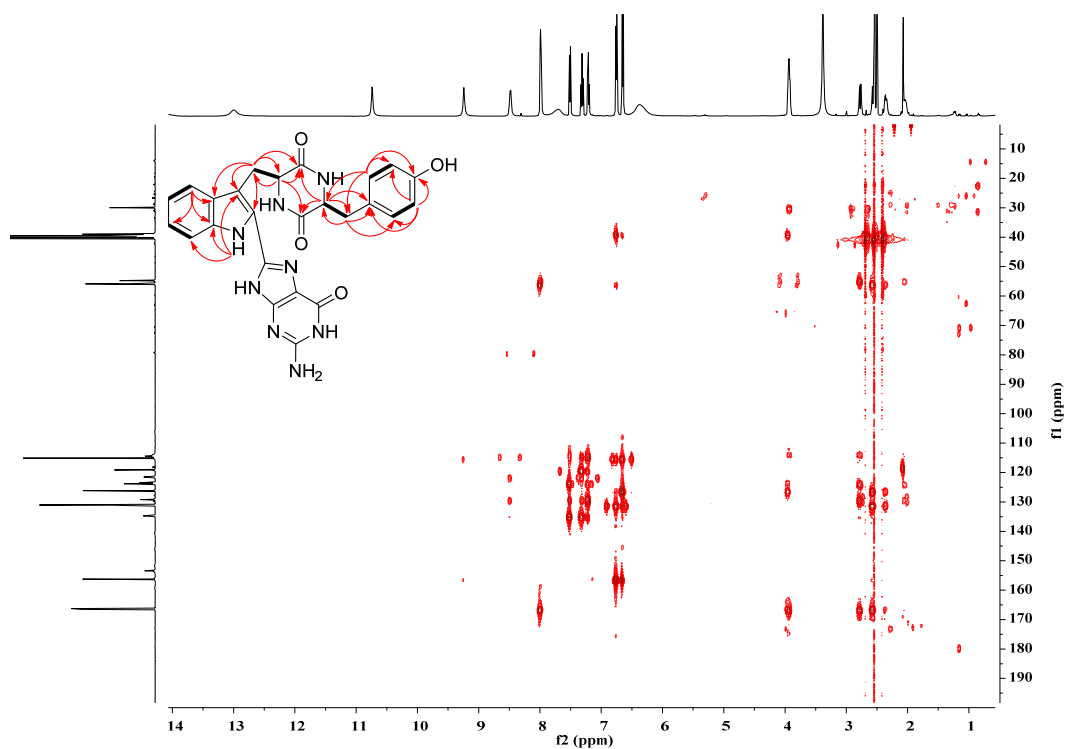


Figure S8. HMBC spectrum of compound **2** in DMSO- d_6

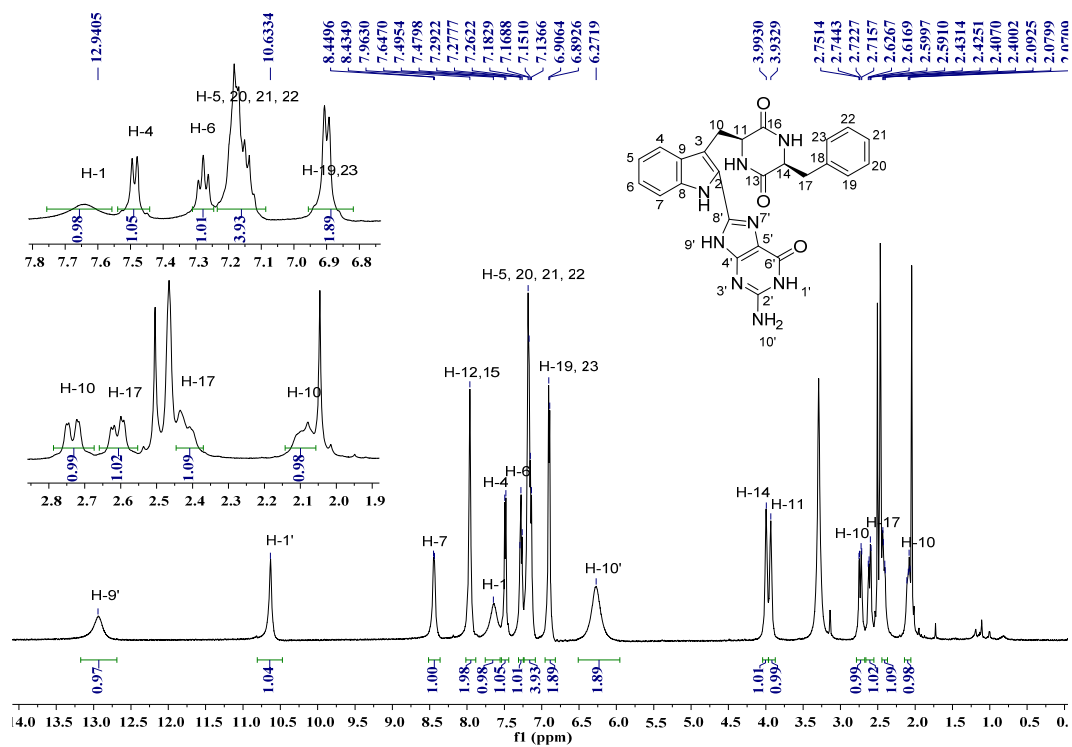


Figure S9. ^1H NMR spectrum of compound **3** in DMSO- d_6 (500 MHz)

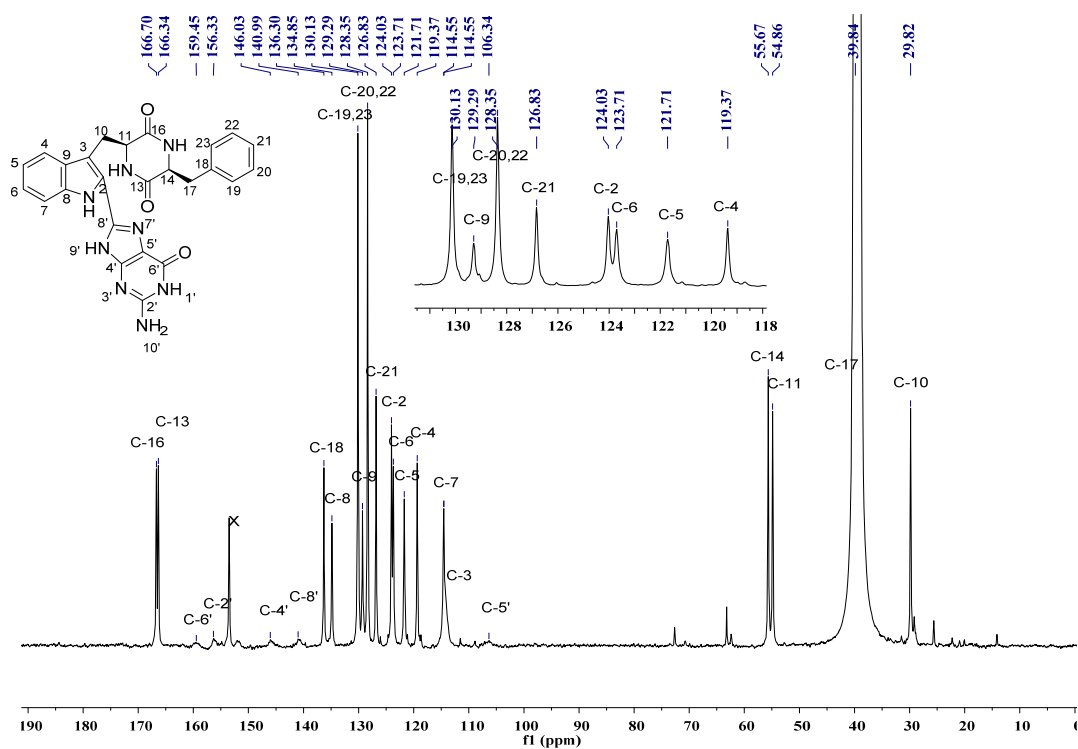


Figure S10. ^{13}C NMR spectrum of compound **3** in $\text{DMSO-}d_6$ (125 MHz)

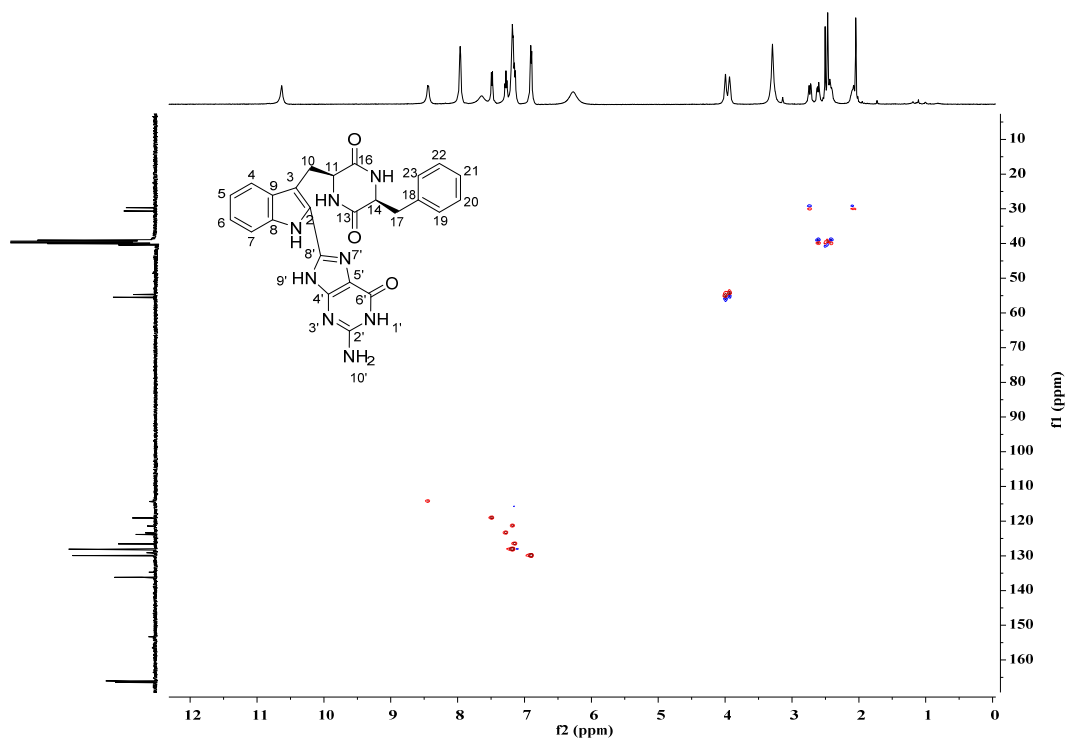


Figure S11. HSQC spectrum of compound **3** in $\text{DMSO-}d_6$

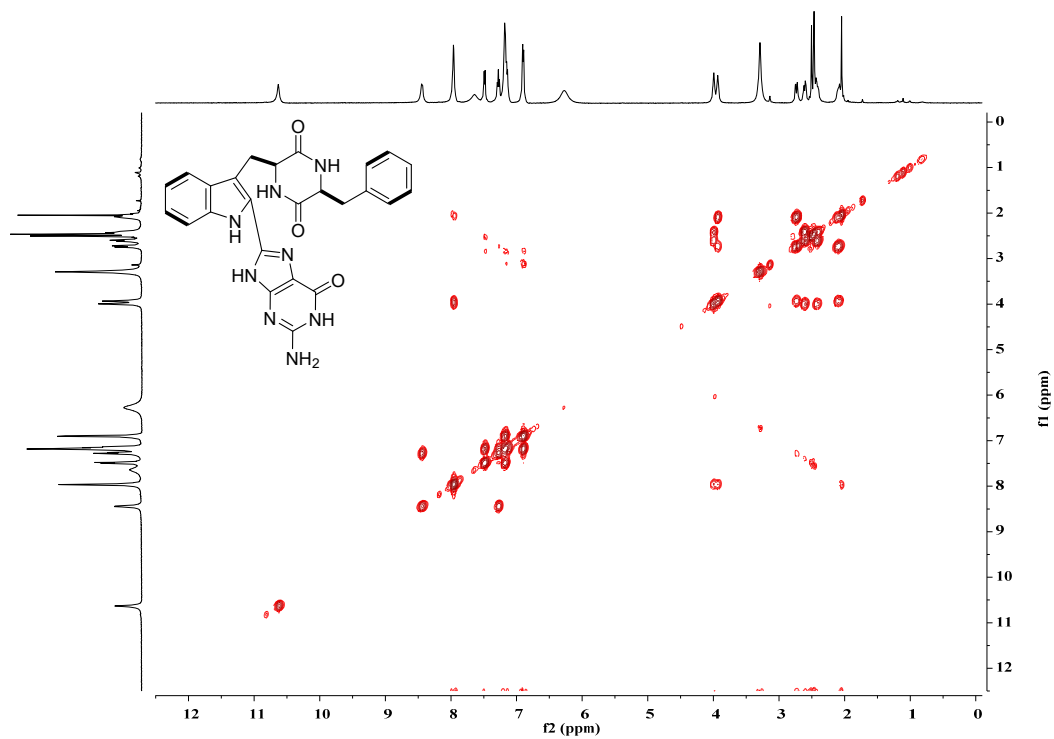


Figure S12. ^1H - ^1H COSY spectrum of compound **3** in $\text{DMSO-}d_6$

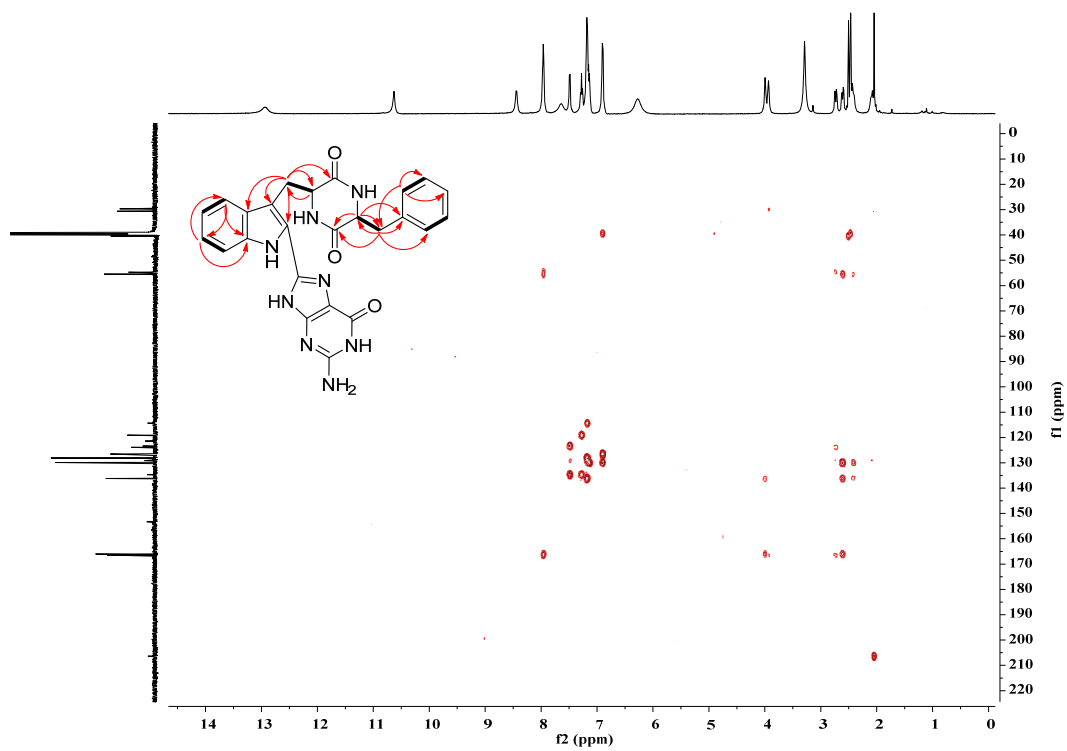


Figure S13. HMBC spectrum of compound **3** in $\text{DMSO-}d_6$

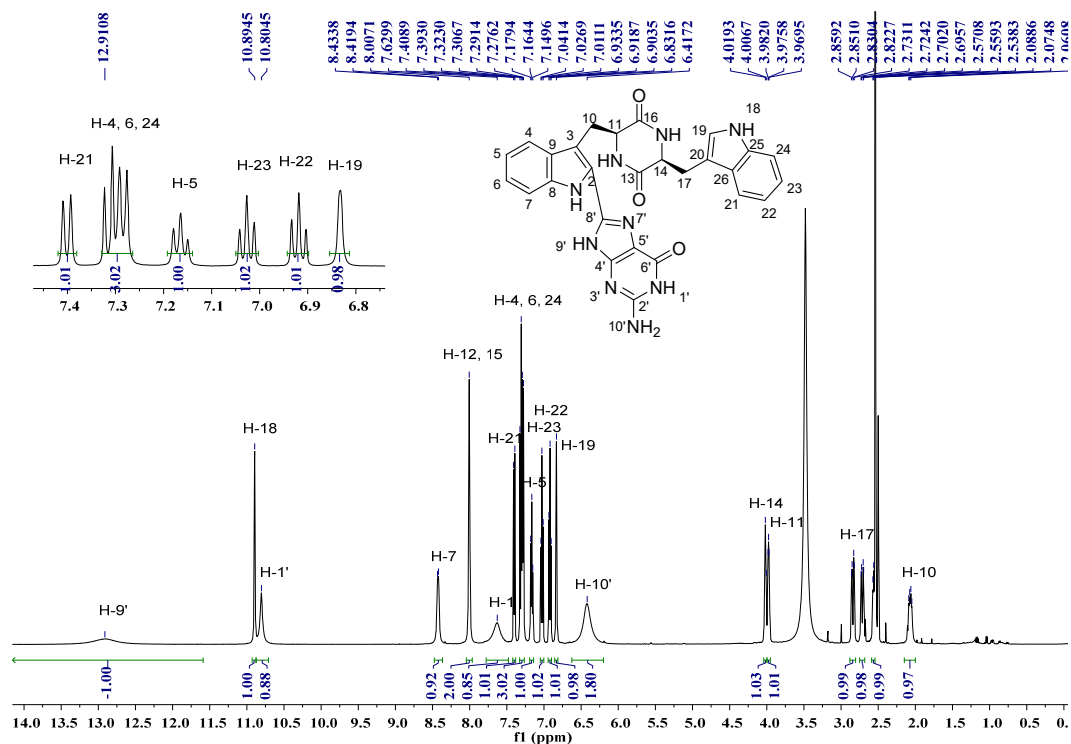


Figure S14. ^1H NMR spectrum of compound **4** in $\text{DMSO-}d_6$ (500 MHz)

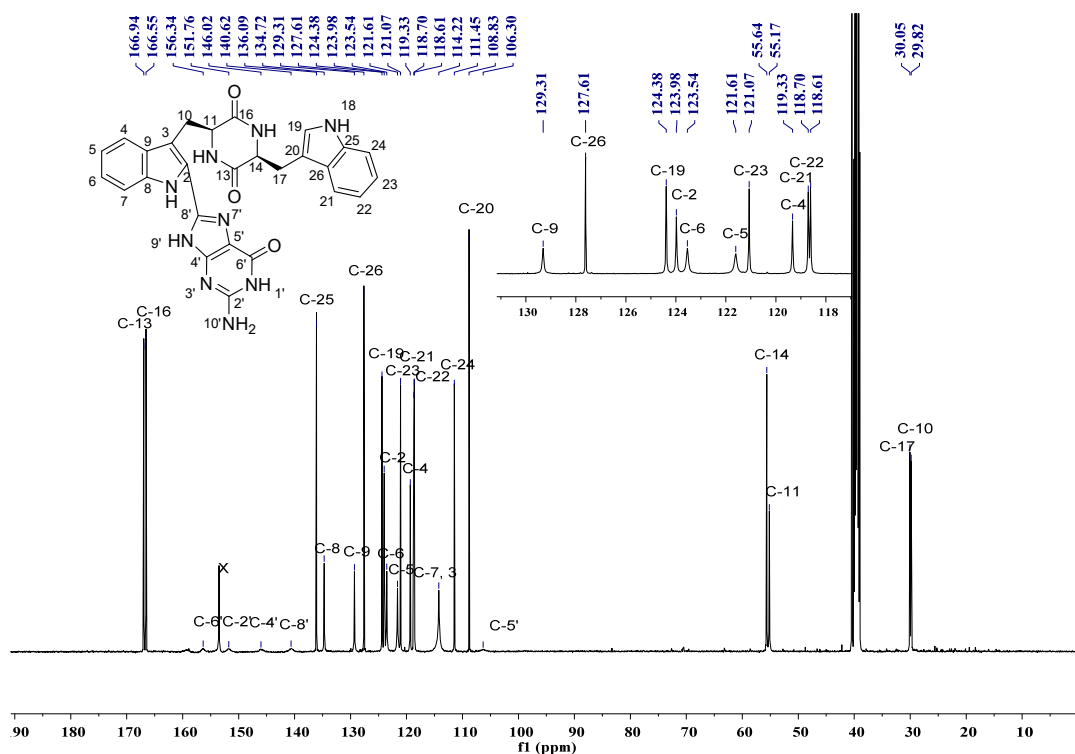


Figure S15. ^{13}C NMR spectrum of compound **4** in $\text{DMSO-}d_6$ (125 MHz)

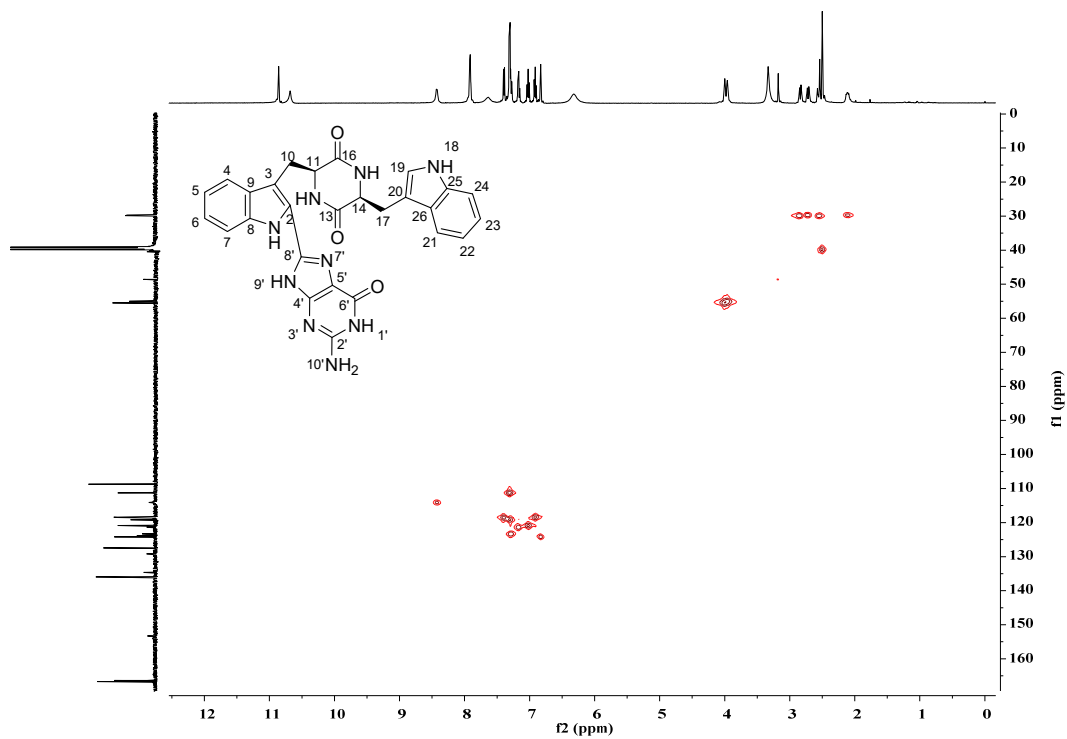


Figure S16. HSQC spectrum of compound **4** in DMSO- d_6

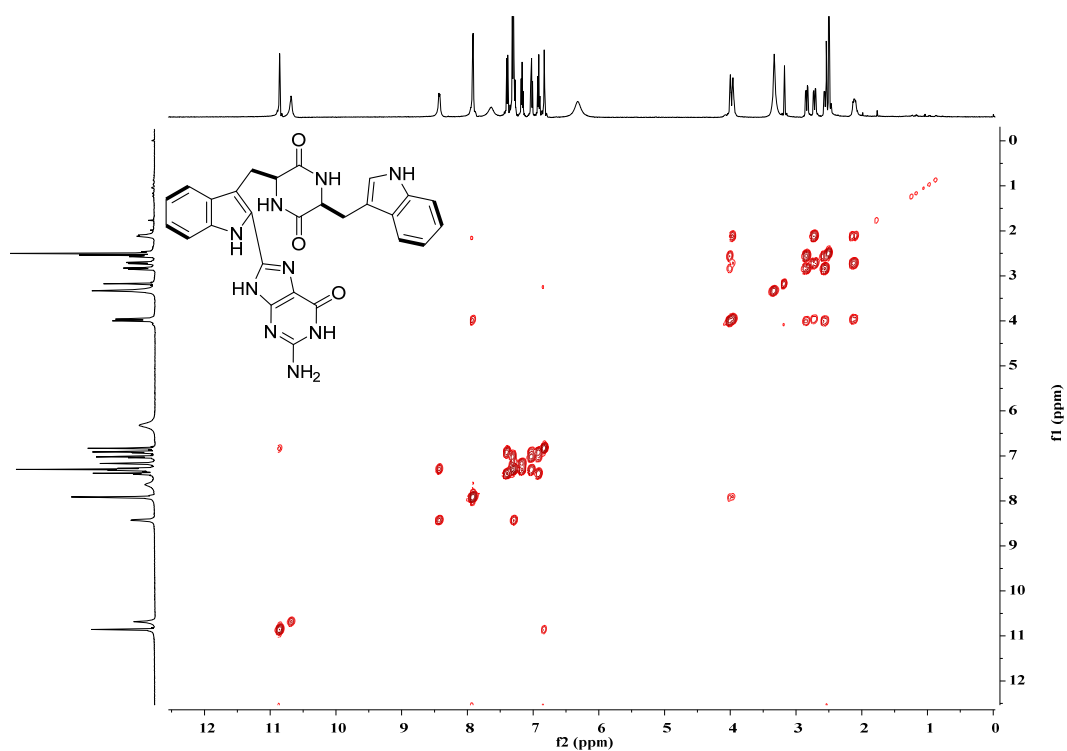


Figure S17. ^1H - ^1H COSY spectrum of compound **4** in DMSO- d_6

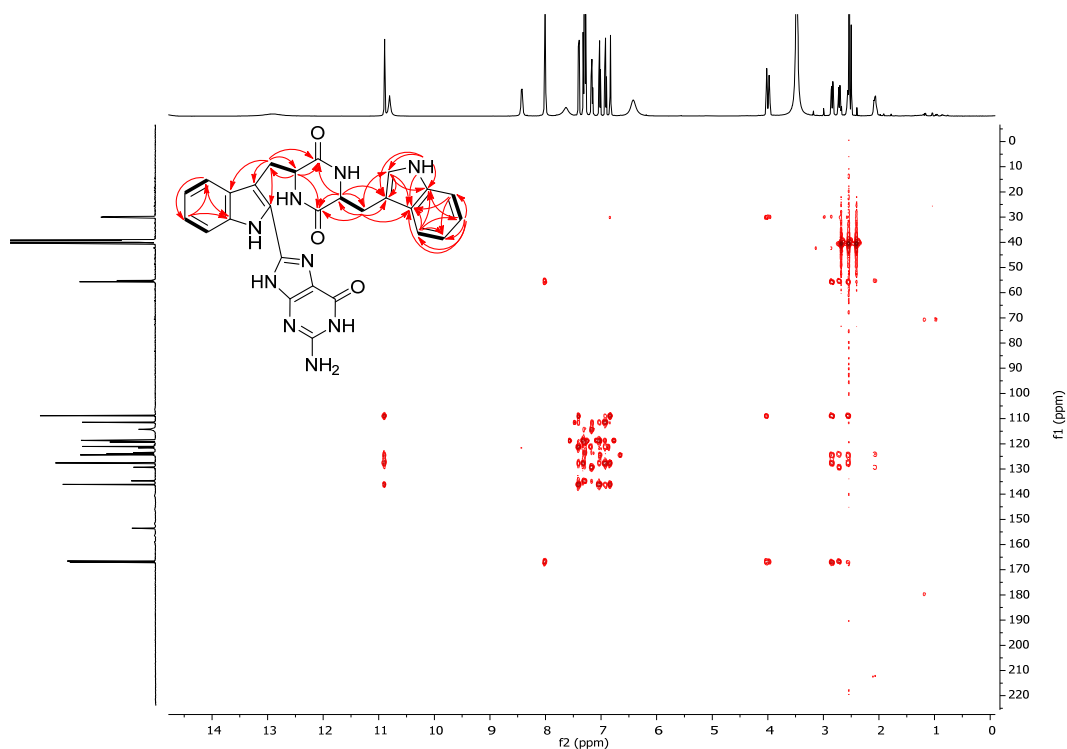


Figure S18. HMBC spectrum of compound 4 in DMSO- d_6

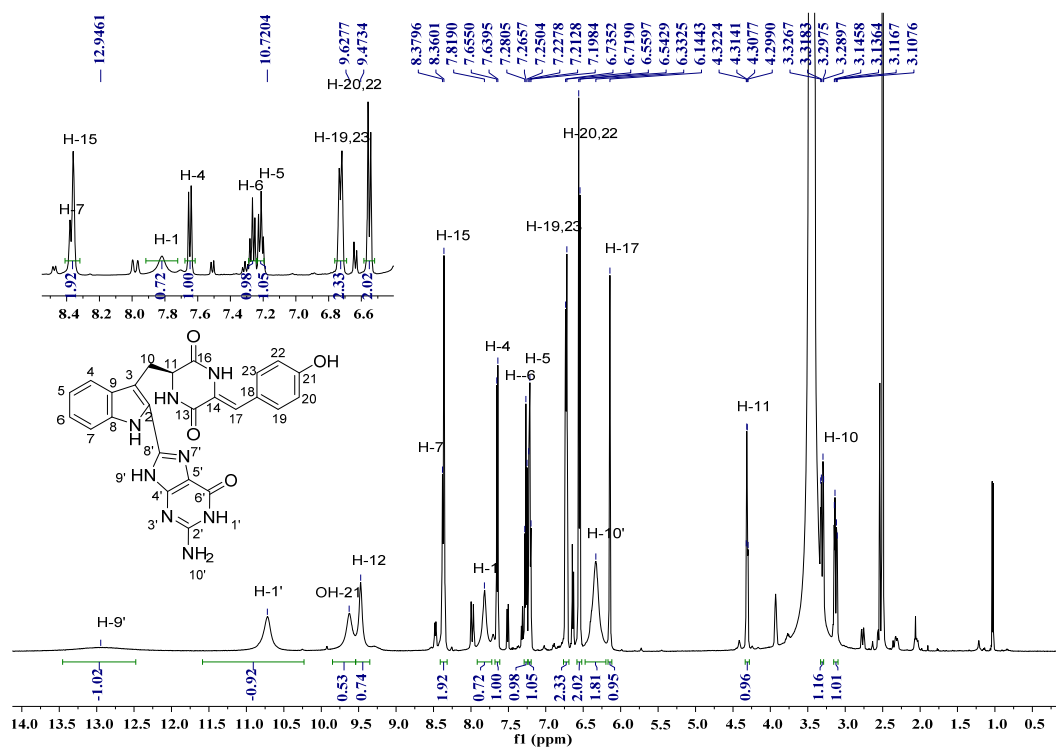


Figure S19. ^1H NMR spectrum of compound 5 in DMSO- d_6 (500 MHz)

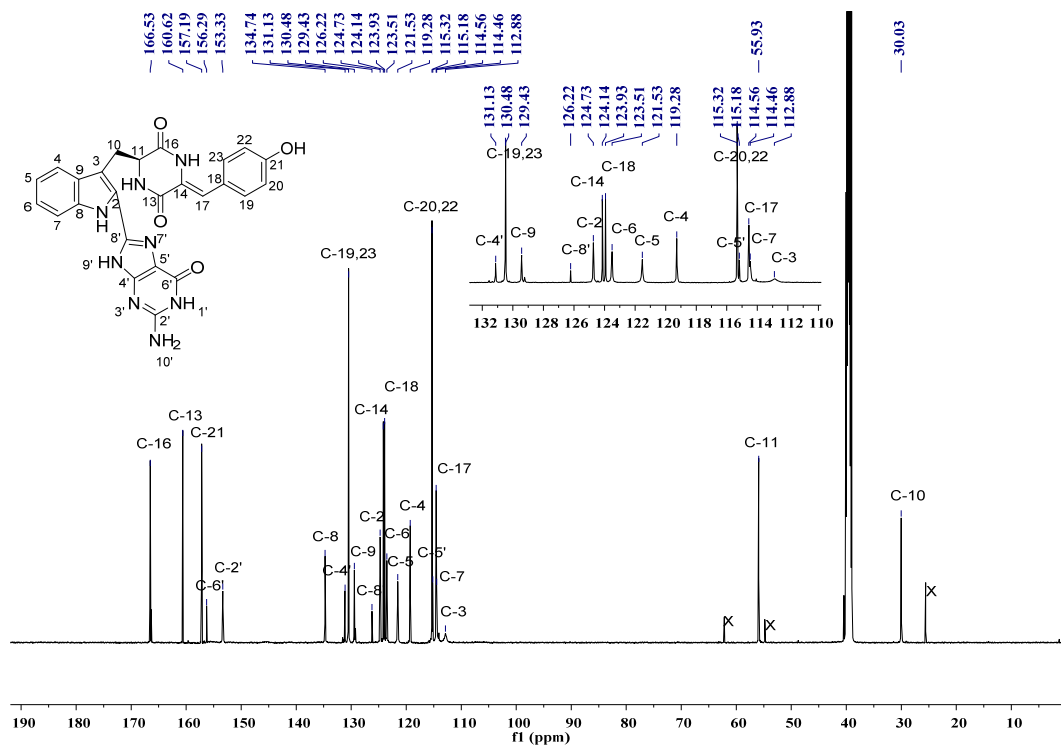


Figure S20. ¹³C NMR spectrum of compound **5** in DMSO-*d*₆ (125 MHz)

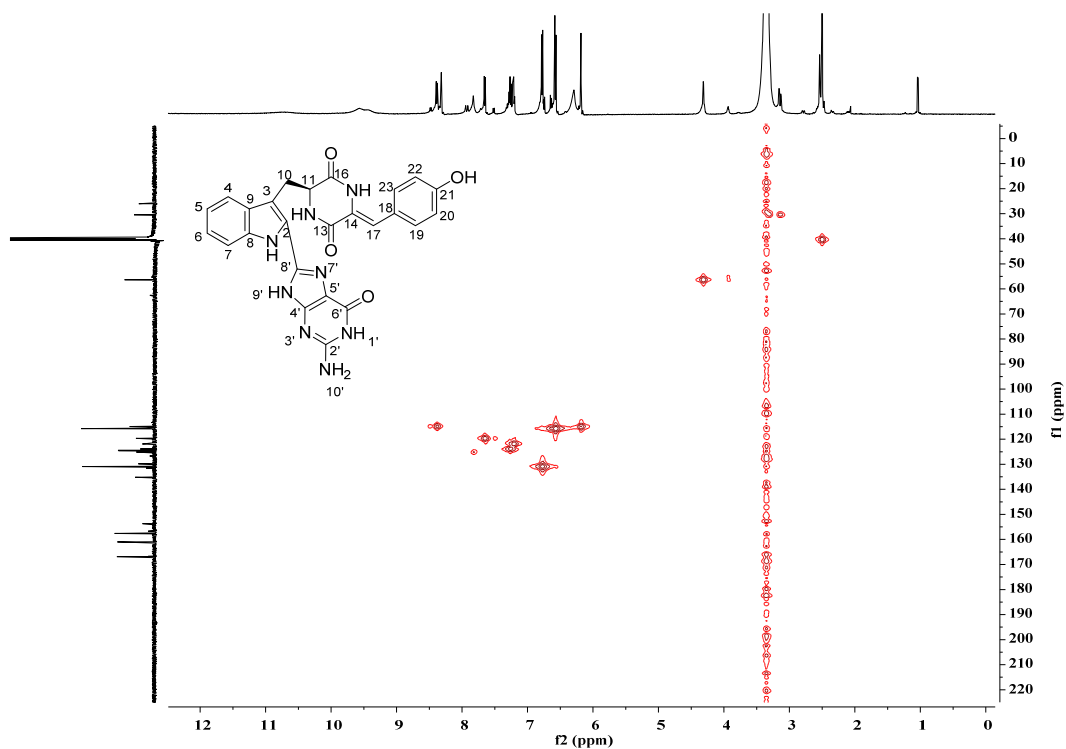


Figure S21. HSQC spectrum of compound **5** in DMSO-*d*₆

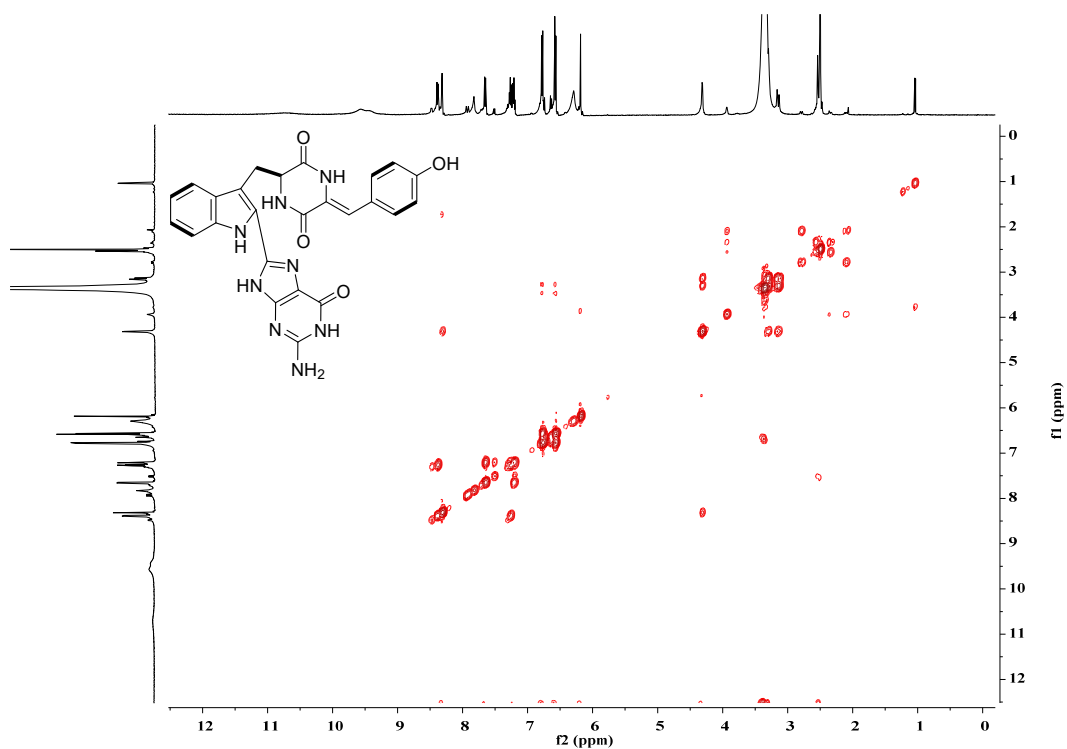


Figure S22. ^1H - ^1H COSY spectrum of compound **5** in $\text{DMSO-}d_6$.

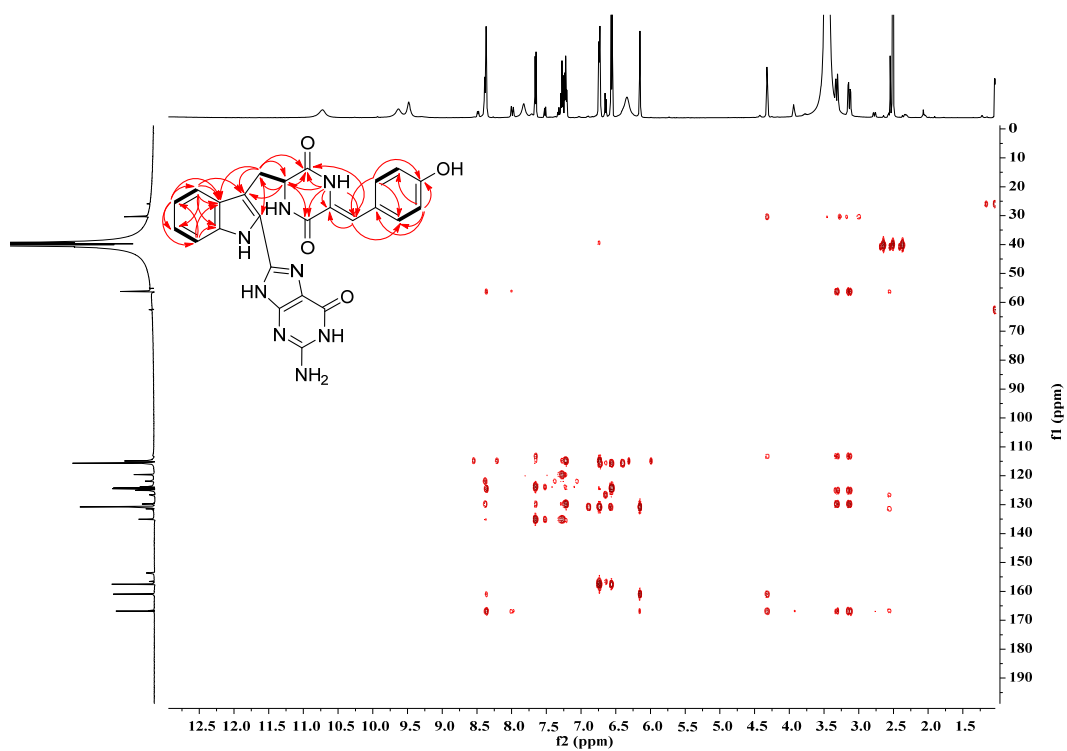


Figure S23. HMBC spectrum of compound **5** in $\text{DMSO-}d_6$

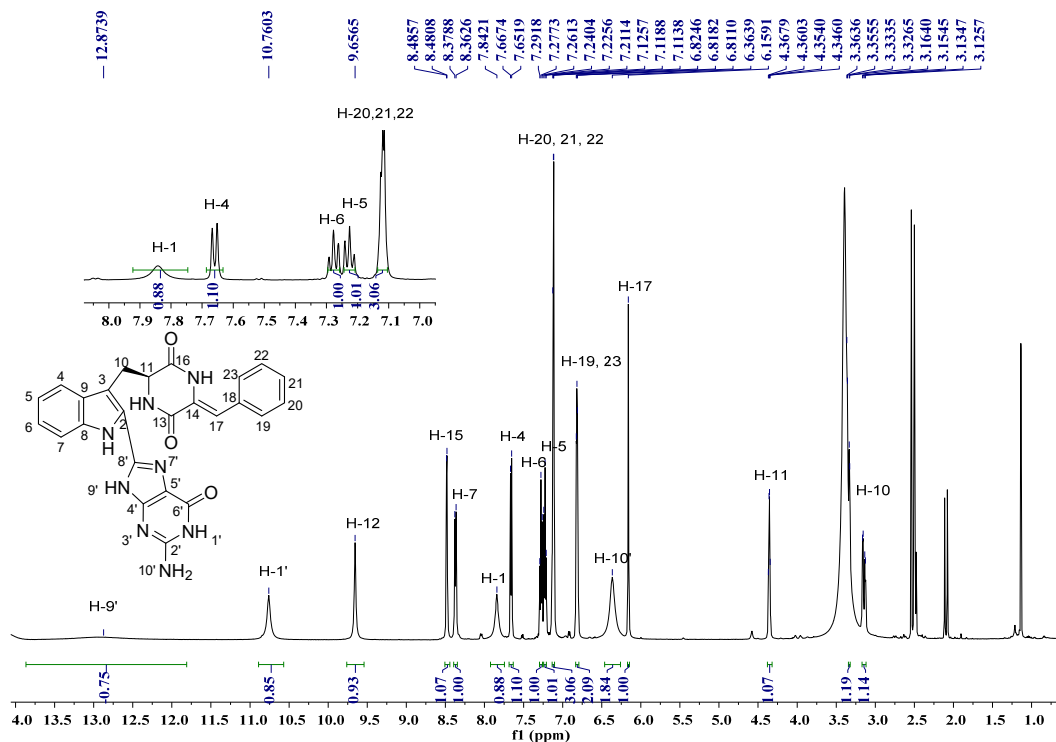


Figure S24. ^1H NMR spectrum of compound **6** in $\text{DMSO}-d_6$ (500 MHz)

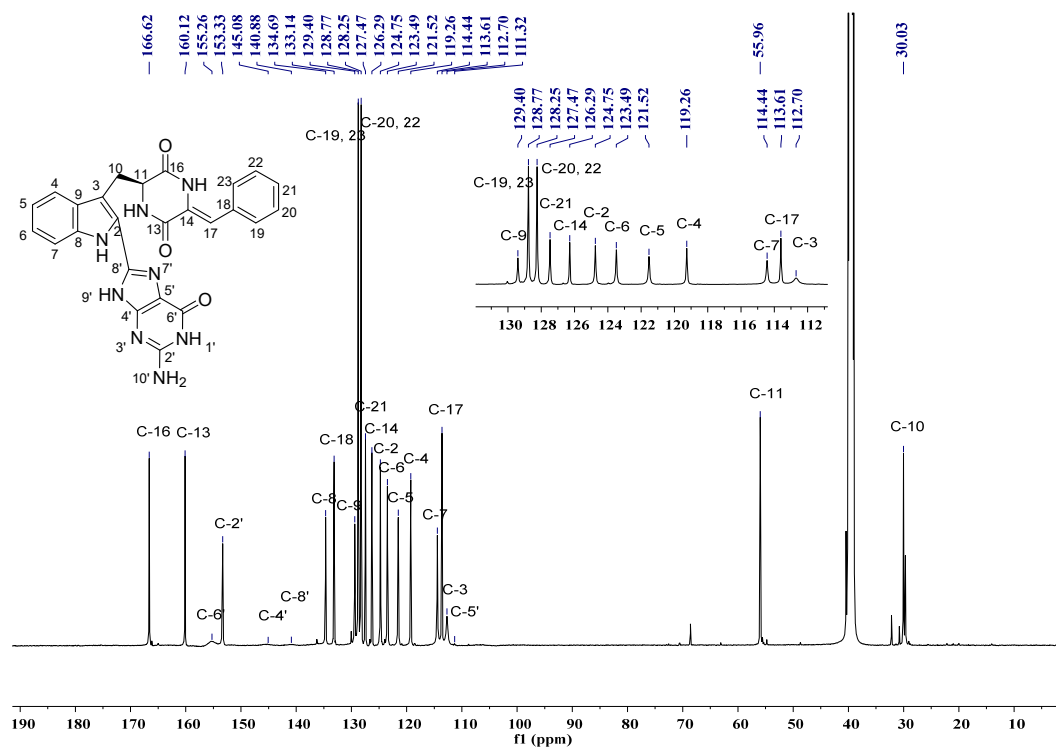


Figure S25. ^{13}C NMR spectrum of compound **6** in $\text{DMSO}-d_6$ (125 MHz)

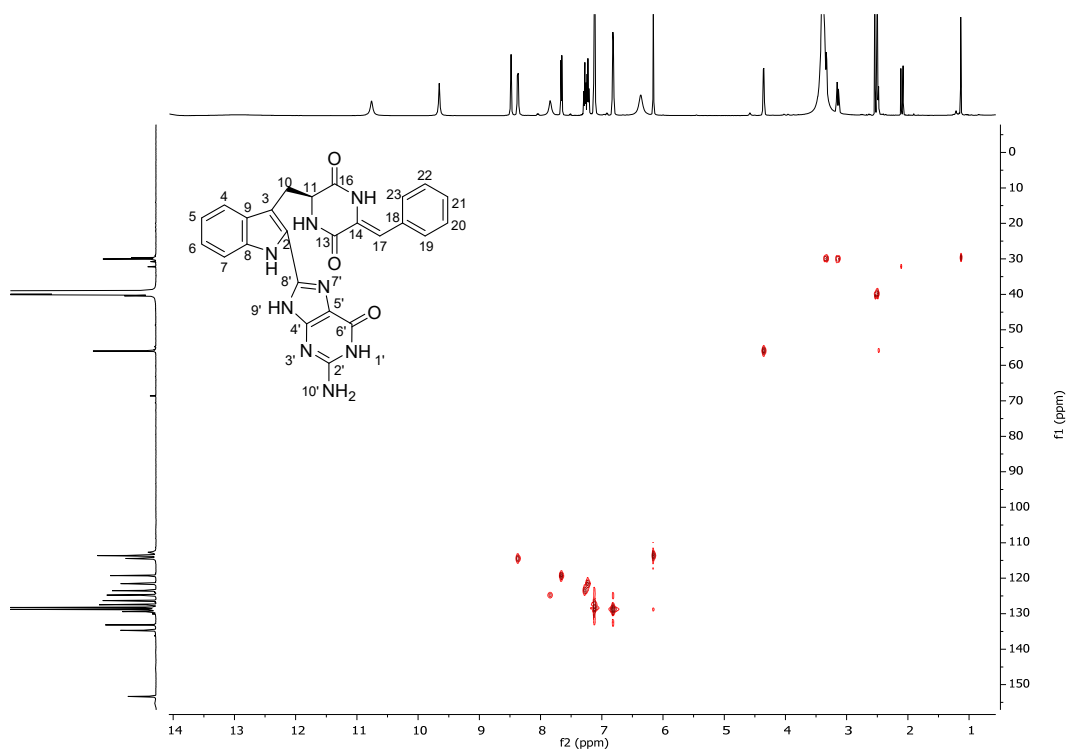


Figure S26. HSQC spectrum of compound **6** in DMSO- d_6

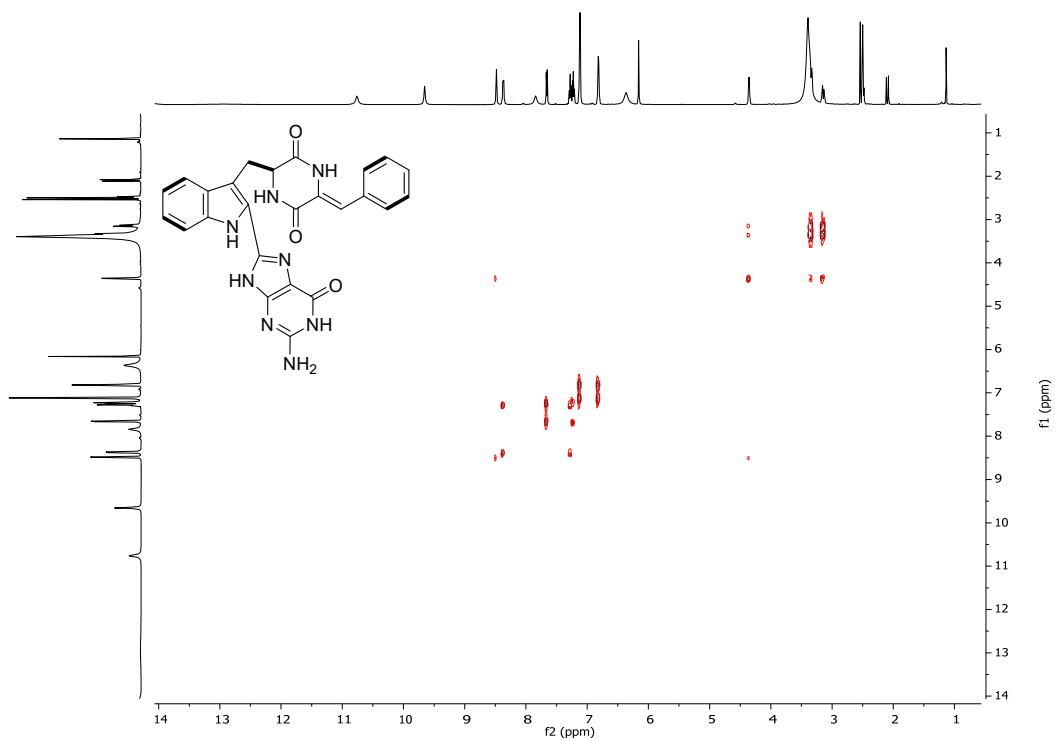


Figure S27. ^1H - ^1H COSY spectrum of compound **6** in DMSO- d_6

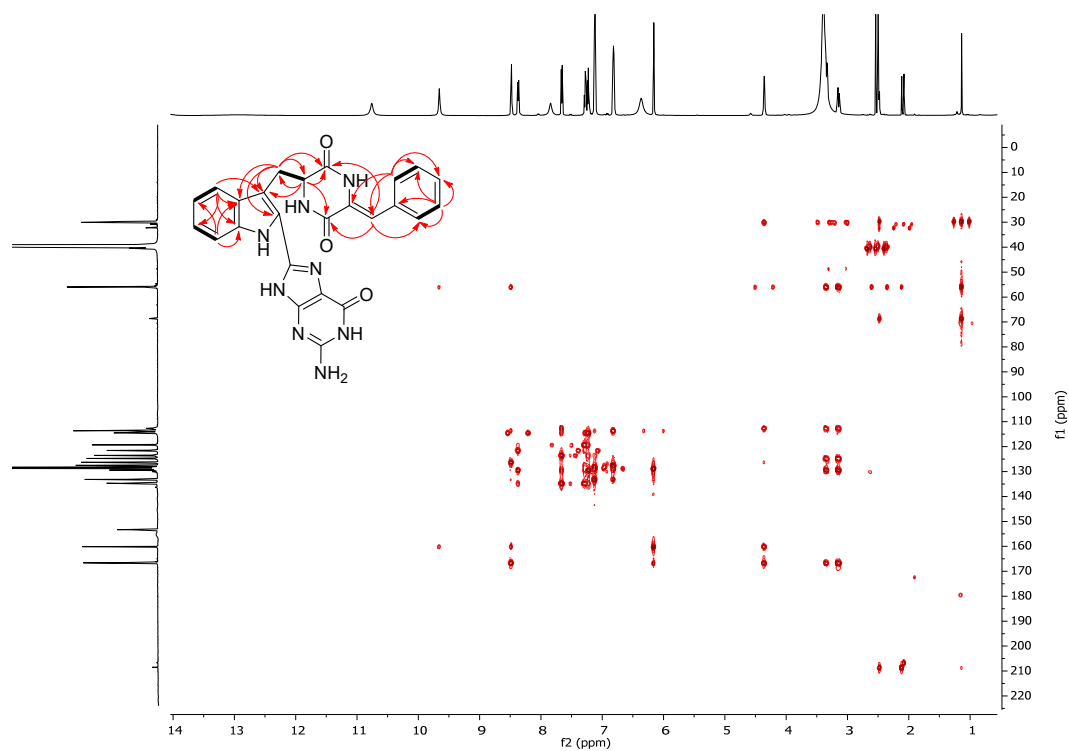


Figure S28. HMBC spectrum of compound **6** in DMSO- d_6

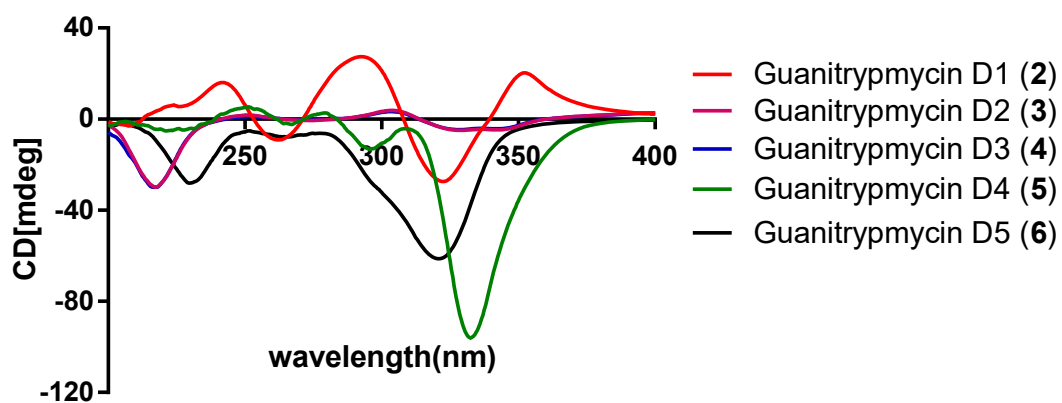


Figure S29. Experimental ECD spectra of compounds **2** – **6** in MeOH – H₂O (1:1)

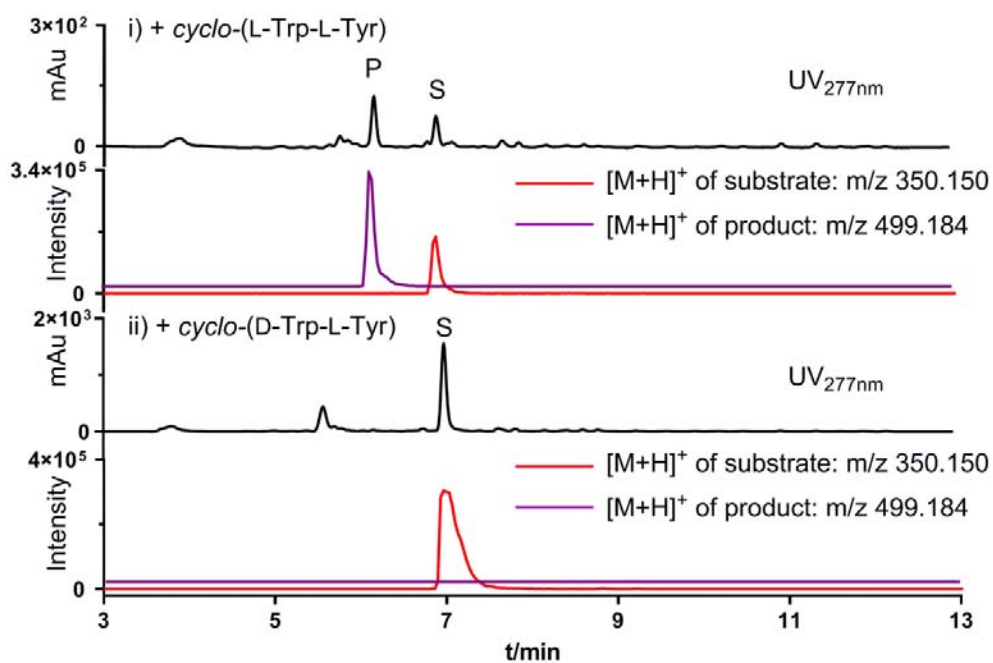


Figure S30. LC-MS analysis of *gutD*₁₅₂₁ transformant after addition of two *cyclo*-(Trp-Tyr) isomers.

A tolerance range of ± 0.005 was used for ion detection.

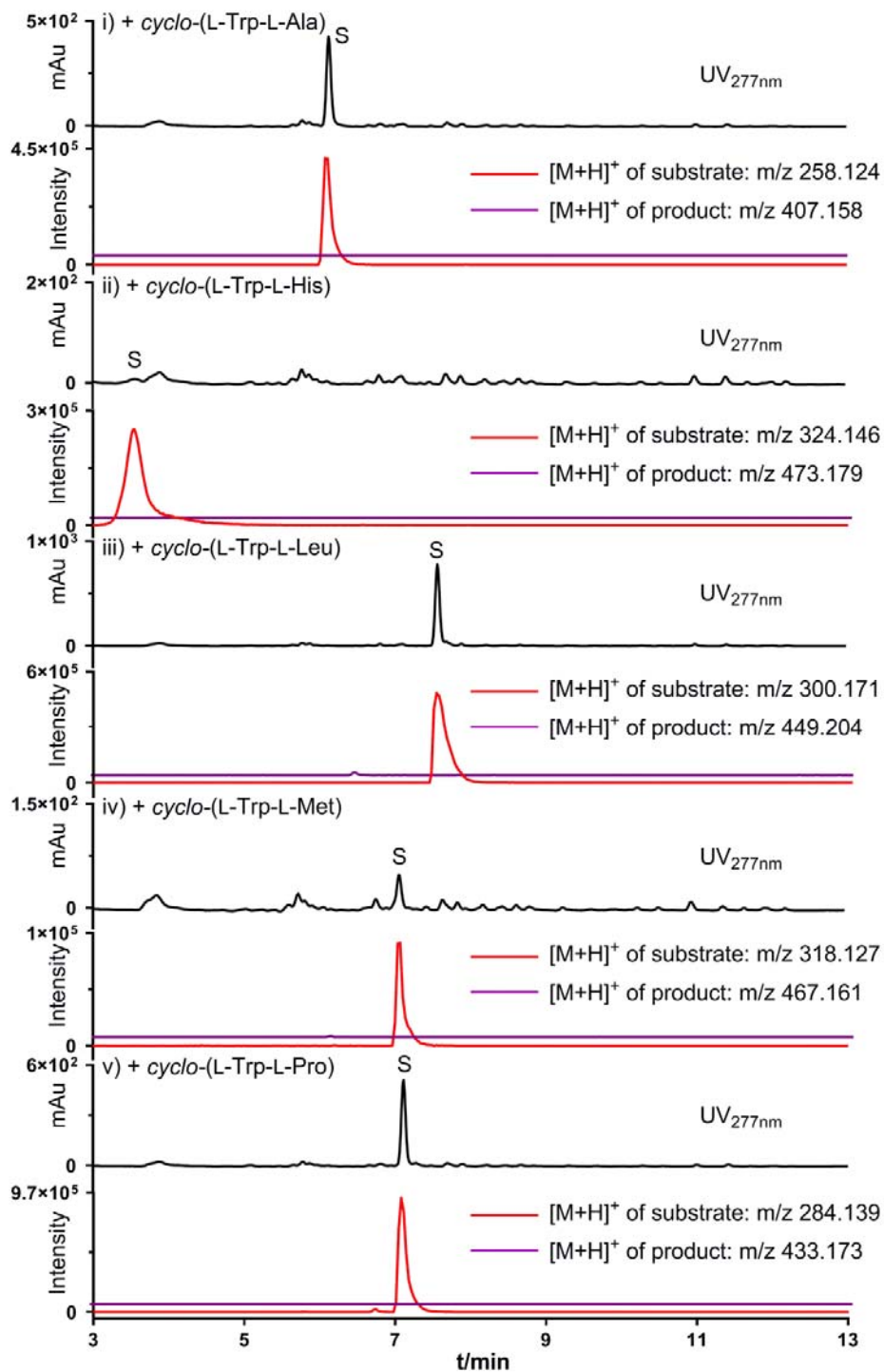


Figure S31. LC-MS analysis of *gutD*₁₅₂₁ transformant supplied with CDPs
 A tolerance range of ± 0.005 was used for ion detection.

Reference

- (1) Liu, J.; Xie, X.; Li, S.-M. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 11534-11540.
- (2) Yu, H.; Xie, X.; Li, S.-M. *Org Lett.* **2018**, *20*, 4921-4925.
- (3) Yu, H.; Xie, X.; Li, S.-M. *Org. Lett.* **2019**, *21*, 9104-9108.
- (4) Harken, L.; Liu, J.; Kreuz, O.; Berger, R.; Li, S.-M. *ACS Catal.* **2022**, *12*, 648-654.
- (5) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U. S. A* **2003**, *100*, 1541-1546.
- (6) Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. *BMC Genomics* **2014**, *15*, 97.

5 Conclusions and future prospects

In this thesis, structural diversification of NPs was achieved by using tailoring enzymes of biosynthetic pathways *via* biotransformation. Although thousands of microbial genome sequences are already available in public databases, many gene clusters remain to be explored. It is expected that advanced genome mining strategies and tools will greatly accelerate the discovery and characterization of new and interesting biocatalysts and SMs from these unexplored biosynthetic pathways.

For the biosynthesis of ustethylin A, heterologous expression and isotopic feeding experiments confirmed that the PKS UttA is responsible for assembling the phenethyl core structure with methylation as a key reaction. The *in vivo* results proved that the NRPS-like enzyme UttJ catalyzes the reduction of the aryl acid to aldehyde and the nonheme Fe^{II}/2-oxoglutarate-dependent oxygenase UttH performs the subsequent hydroxylation at the benzyl group. After methylation by the O-MeT UttF, the cytochrome P450 enzyme UttC catalyzes the hydroxylation of the phenethyl residue to form the product ustethylin A. Deletion of *uttD* coding for a regulator completely abolished product formation, proving its role in regulating the expression of the *utt* BGC. These results suggest that mining novel secondary metabolite gene clusters is a powerful tool to increase the structural diversity of natural products.

For the biosynthesis of strepazine C and its derivatives, a three-gene cluster was identified by genome mining from *Streptomyces aurantiacus*. Heterologous expression and precursor incubation experiments elucidated the biosynthetic steps of strepazine C. It demonstrated that *cyclo*-(L-Trp-L-Trp), initially assembled by the cyclodipeptide synthase SasA, serves as a DKP precursor, followed by regular C-3 prenylation catalyzed by the prenyltransferase SasB and further methylation by the methyltransferase SasC. Furthermore, *in vivo* biotransformation experiments demonstrated the high flexibility of SasB towards different tryptophan-containing cyclodipeptides, as well as their dehydro derivatives for regular C-3 prenylation. Thus, this study provides an enzyme with high substrate promiscuity to the group of prenyltransferases in the less explored cyclodipeptide synthase-related pathways and provides more details about its biochemical properties.

Furthermore, a two-gene cluster coding for a CDPS and a P450 from *Streptomyces* sp. NRRL S-1521 was identified for the biosynthesis of guanitrypmycins, which are rare and novel C2-guaninyl indole alkaloids, by phylogenetic analysis. Heterologous expression, biochemical characterization, together with structural elucidation proved that *cyclo*-(L-Trp-L-Tyr), initially assembled by the cyclodipeptide synthase GutA, serves as a DKP precursor. Subsequently, the cytochrome P450 enzyme GutD₁₅₂₁ catalyzes the regiospecific transfer of guanine to C-2 of the indole ring of *cyclo*-(L-Trp-L-Tyr) *via* a C-C linkage, which represents a new chemical transformation within this enzyme

class. Precursor incubation experiments revealed that GutD₁₅₂₁ efficiently accepts several other tryptophan-containing cyclodipeptides or derivatives for regiospecific coupling with guanine, thus resulting in different guanitrypmycin analogs. This study provides a new linkage mode between the indole ring of DKPs and a guanine moiety and expands the functional scope of P450s as tailoring enzymes.

For future prospects, the following works can be performed:

- For the biosynthesis of ustethylin A, there are still some enzymes whose functions are not clear. Therefore, the role of the unknown enzymes can be further studied by *in vivo* and *in vitro* experiments.
- For the two studies on CDP derivatives, unfortunately, neither the prenyltransferase SasB nor the cytochrome P450 GutD₁₅₂₁ could be obtained as soluble proteins. Thus, a method to obtain soluble proteins for biochemical and structural investigation still needs to be established.
- Guanitrypmycins are novel nucleobase-containing DKPs. However, antibacterial activity tests of guanitrypmycin analogs showed no obvious inhibitory activity. Therefore, further bioactivity assays toward some representative screening models are required to explore their potential biological and pharmaceutical activities.

6 References

1. Sorokina, M.; Steinbeck, C., Review on natural products databases: where to find data in 2020. *J Cheminform* **2020**, *12* (1), 20.
2. Walsh, C. T.; Tang, Y., *Natural Product Biosynthesis*. Royal Society of Chemistry: 2017.
3. Wiemann, P.; Keller, N. P., Strategies for mining fungal natural products. *J Ind Microbiol Biotechnol* **2014**, *41* (2), 301-313.
4. Brakhage, A. A., Regulation of fungal secondary metabolism. *Nat Rev Microbiol* **2013**, *11* (1), 21-32.
5. Rohlfs, M.; Churchill, A. C., Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genet Biol* **2011**, *48* (1), 23-34.
6. Rodrigues, A. P.; Carvalho, A. S.; Santos, A. S.; Alves, C. N.; do Nascimento, J. L.; Silva, E. O., Kojic acid, a secondary metabolite from *Aspergillus* sp., acts as an inducer of macrophage activation. *Cell Biol Int* **2011**, *35* (4), 335-343.
7. Dufour, N.; Rao, R. P., Secondary metabolites and other small molecules as intercellular pathogenic signals. *FEMS Microbiol Lett* **2011**, *314* (1), 10-17.
8. Fox, E. M.; Howlett, B. J., Secondary metabolism: regulation and role in fungal biology. *Curr Opin Microbiol* **2008**, *11* (6), 481-487.
9. Yim, G.; Wang, H. H.; Davies, J., Antibiotics as signalling molecules. *Philos Trans R Soc Lond B Biol Sci* **2007**, *362* (1483), 1195-1200.
10. Keller, N. P., Fungal secondary metabolism: regulation, function and drug discovery. *Nat Rev Microbiol* **2019**, *17* (3), 167-180.
11. Zhang, P.; Wang, X.; Fan, A.; Zheng, Y.; Liu, X.; Wang, S.; Zou, H.; Oakley, B. R.; Keller, N. P.; Yin, W. B., A cryptic pigment biosynthetic pathway uncovered by heterologous expression is essential for conidial development in *Pestalotiopsis fici*. *Mol Microbiol* **2017**, *105* (3), 469-483.
12. Blackwell, M., The fungi: 1, 2, 3 ... 5.1 million species? *Am J Bot* **2011**, *98* (3), 426-438.
13. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod* **2020**, *83* (3), 770-803.
14. Azmir, J.; Zaidul, I. S. M.; Rahman, M. M.; Sharif, K. M.; Mohamed, A.; Sahena, F.; Jahurul, M. H. A.; Ghafoor, K.; Norulaini, N. A. N.; Omar, A. K. M., Techniques for extraction of bioactive compounds from plant materials: a review. *J Food Eng* **2013**, *117* (4), 426-436.
15. Berdy, J., Thoughts and facts about antibiotics: where we are now and where we are heading. *J Antibiot* **2012**, *65* (8), 385-395.
16. Berdy, J., Bioactive microbial metabolites. *J Antibiot* **2005**, *58* (1), 1-26.
17. Chávez, R.; Fierro, F.; García-Rico, R. O.; Vaca, I., Filamentous fungi from extreme environments as a promising source of novel bioactive secondary metabolites. *Front Microbiol* **2015**, *6*, 903.
18. Demain, A. L.; Sanchez, S., Microbial drug discovery: 80 years of progress. *J Antibiot* **2009**, *62* (1), 5-16.
19. Hoffmeister, D.; Keller, N. P., Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep* **2007**, *24* (2), 393-416.
20. Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T.; Taskforce, I. N. P. S., Natural products in drug discovery: advances and opportunities. *Nat Rev Drug Discov* **2021**, *20* (3), 200-216.
21. Shi, Y.; Inoue, H.; Wu, J. C.; Yamanaka, S., Induced pluripotent stem cell technology: a decade of progress. *Nat Rev Drug Discov* **2017**, *16* (2), 115-130.
22. Moffat, J. G.; Vincent, F.; Lee, J. A.; Eder, J.; Prunotto, M., Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat Rev Drug Discov* **2017**, *16* (8), 531-543.
23. Fellmann, C.; Gowen, B. G.; Lin, P. C.; Doudna, J. A.; Corn, J. E., Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat Rev Drug Discov* **2017**, *16* (2), 89-100.
24. Schirle, M.; Jenkins, J. L., Identifying compound efficacy targets in phenotypic drug discovery. *Drug Discov Today* **2016**, *21* (1), 82-89.
25. Palazzotto, E.; Weber, T., Omics and multi-omics approaches to study the biosynthesis of secondary metabolites in microorganisms. *Curr Opin Microbiol* **2018**, *45*, 109-116.
26. Dias, T.; Gaudencio, S. P.; Pereira, F., A computer-driven approach to discover natural product leads for Methicillin-resistant *Staphylococcus aureus* infection therapy. *Mar Drugs* **2018**, *17* (1), 16.

REFERENCES

27. Schneider, G.; Reker, D.; Chen, T.; Hauenstein, K.; Schneider, P.; Altmann, K. H., Deorphaning the macromolecular targets of the natural anticancer compound dolicolide. *Angew Chem Int Ed Engl* **2016**, *55* (40), 12408-12411.
28. Korman, T. P.; Ames, B.; Tsai, S.-C., 1.08 - Structural enzymology of polyketide synthase: the Structure–Sequence–Function correlation. In *Comprehensive Natural Products II*, Liu, H.-W.; Mander, L., Eds. Elsevier: Oxford **2010**, pp 305-345.
29. Cox, R. J.; Skellam, E.; Williams, K., Biosynthesis of fungal polyketides. *Physiology and Genetics* **2018**, pp, 385-412.
30. Hill, A. M., The biosynthesis, molecular genetics and enzymology of the polyketide-derived metabolites. *Nat Prod Rep* **2006**, *23* (2), 256-320.
31. Campbell, C. D.; Vederas, J. C., Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **2010**, *93* (9), 755-763.
32. Tobert, J. A., Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. *Nat Rev Drug Discov* **2003**, *2* (7), 517-526.
33. Havel, R. J.; Hunninghake, D. B.; Illingworth, D. R.; Lees, R. S.; Stein, E. A.; Tobert, J. A.; Bacon, S. R.; Bolognese, J. A.; Frost, P. H.; Lamkin, G. E.; et al., Lovastatin (mevinolin) in the treatment of heterozygous familial hypercholesterolemia. A multicenter study. *Ann Intern Med* **1987**, *107* (5), 609-615.
34. Alberts, A. W.; Chen, J.; Kuron, G.; Hunt, V.; Huff, J.; Hoffman, C.; Rothrock, J.; Lopez, M.; Joshua, H.; Harris, E.; Patchett, A.; Monaghan, R.; Currie, S.; Stapley, E.; Albers-Schonberg, G.; Hensens, O.; Hirshfield, J.; Hoogsteen, K.; Liesch, J.; Springer, J., Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme a reductase and a cholesterol-lowering agent. *Proc Natl Acad Sci U S A* **1980**, *77* (7), 3957-3961.
35. Harris, C. M.; Roberson, J. S.; Harris, T. M., Biosynthesis of griseofulvin. *J Am Chem Soc* **1976**, *98* (17), 5380-5386.
36. Staunton, J.; Wilkinson, B., Biosynthesis of erythromycin and rapamycin. *Chem Rev* **1997**, *97* (7), 2611-2630.
37. Sapadin, A. N.; Fleischmajer, R., Tetracyclines: nonantibiotic properties and their clinical implications. *J Am Acad Dermatol* **2006**, *54* (2), 258-265.
38. Lomovskaya, N.; Otten, S. L.; Doi-Katayama, Y.; Fonstein, L.; Liu, X. C.; Takatsu, T.; Inventi-Solari, A.; Filippini, S.; Torti, F.; Colombo, A. L.; Hutchinson, C. R., Doxorubicin overproduction in *Streptomyces peucetius*: cloning and characterization of the *dnrU* ketoreductase and *dnrV* genes and the *doxA* cytochrome P-450 hydroxylase gene. *J Bacteriol* **1999**, *181* (1), 305-318.
39. Lebeau, J.; Venkatachalam, M.; Fouillaud, M.; Petit, T.; Vinale, F.; Dufosse, L.; Caro, Y., Production and new extraction method of polyketide red pigments produced by ascomycetous Fungi from terrestrial and marine habitats. *J Fungi* **2017**, *3* (3), 34.
40. Cai, W.; Zhang, W., Engineering modular polyketide synthases for production of biofuels and industrial chemicals. *Curr Opin Biotechnol* **2018**, *50*, 32-38.
41. Giessen, T. W.; Marahiel, M. A., The tRNA-dependent biosynthesis of modified cyclic dipeptides. *Int J Mol Sci* **2014**, *15* (8), 14610-14631.
42. Borthwick, A. D., 2,5-Diketopiperazines: synthesis, reactions, medicinal chemistry, and bioactive natural products. *Chem Rev* **2012**, *112* (7), 3641-3716.
43. Martins, M. B.; Carvalho, I., Diketopiperazines: biological activity and synthesis. *Tetrahedron* **2007**, *63* (40), 9923-9932.
44. Horton, D. A.; Bourne, G. T.; Smythe, M. L., Exploring privileged structures: the combinatorial synthesis of cyclic peptides. *J Computer-Aided Molecular Design* **2002**, *16* (5), 415-431.
45. Prasad, C., Bioactive cyclic dipeptides. *Peptides* **1995**, *16* (1), 151-164.
46. Belin, P.; Moutiez, M.; Lautru, S.; Seguin, J.; Pernodet, J.-L.; Gondry, M., The nonribosomal synthesis of diketopiperazines in tRNA-dependent cyclodipeptide synthase pathways. *Nat Prod Rep* **2012**, *29* (9), 961-979.
47. Strom, K.; Sjogren, J.; Broberg, A.; Schnurer, J., *Lactobacillus plantarum* MiLAB 393 produces the antifungal cyclic dipeptides cyclo(L-Phe-L-Pro) and cyclo(L-Phe-*trans*-4-OH-L-Pro) and 3-phenyllactic acid. *Appl Environ Microbiol* **2002**, *68* (9), 4322-4327.
48. Kanoh, K.; Kohno, S.; Asari, T.; Harada, T.; Katada, J.; Muramatsu, M.; Kawashima, H.; Sekiya, H.; Uno, I., (-)-Phenylahistin: A new mammalian cell cycle inhibitor produced by *Aspergillus ustus*. *Bioorg Med Chem Lett* **1997**, *7* (22), 2847-2852.

REFERENCES

49. Blayney, D. W.; Mohanlal, R.; Adamchuk, H.; Kirtbaya, D. V.; Chen, M.; Du, L.; Ogenstad, S.; Ginn, G.; Huang, L.; Zhang, Q., Efficacy of plinabulin vs pegfilgrastim for prevention of docetaxel-induced neutropenia in patients with solid tumors: a randomized clinical trial. *JAMA Netw Open* **2022**, *5* (1), e2145446.
50. Watanabe, A.; Kuriyama, T.; Kamei, K.; Nishimura, K.; Miyaji, M.; Sekine, T.; Waku, M., Immunosuppressive substances in *Aspergillus fumigatus* culture filtrate. *J Infect chemother* **2003**, *9* (2), 114-121.
51. Keller, N. P.; Hohn, T. M., Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology* **1997**, *21* (1), 17-29.
52. Walton, J. D., Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genetics Biology* **2000**, *30* (3), 167-171.
53. Zhang, Z.; Pan, H. X.; Tang, G. L., New insights into bacterial type II polyketide biosynthesis. *F1000Res* **2017**, *6*, 172.
54. Omura, S.; Crump, A., The life and times of ivermectin - a success story. *Nat Rev Microbiol* **2004**, *2* (12), 984-989.
55. Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E., Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem* **2007**, *76*, 195-221.
56. Chooi, Y. H.; Tang, Y., Navigating the fungal polyketide chemical space: from genes to molecules. *J Org Chem* **2012**, *77* (22), 9933-9953.
57. Crawford, J. M.; Townsend, C. A., New insights into the formation of fungal aromatic polyketides. *Nat Rev Microbiol* **2010**, *8* (12), 879-889.
58. Regueira, T. B.; Kildegaard, K. R.; Hansen, B. G.; Mortensen, U. H.; Hertweck, C.; Nielsen, J., Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. *Appl Environ Microbiol* **2011**, *77* (9), 3035-3043.
59. Holm, D. K.; Petersen, L. M.; Klitgaard, A.; Knudsen, P. B.; Jarczynska, Z. D.; Nielsen, K. F.; Goffredsen, C. H.; Larsen, T. O.; Mortensen, U. H., Molecular and chemical characterization of the biosynthesis of the 6-MSA-derived meroterpenoid yanuthone D in *Aspergillus niger*. *Chem Biol* **2014**, *21* (4), 519-529.
60. Beck, J.; Ripka, S.; Siegner, A.; Schiltz, E.; Schweizer, E., The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. *Eur J Biochem* **1990**, *192* (2), 487-498.
61. Herbst, D. A.; Townsend, C. A.; Maier, T., The architectures of iterative type I PKS and FAS. *Nat Prod Rep* **2018**, *35* (10), 1046-1069.
62. Kim, J.; Yi, G. S., PKMiner: a database for exploring type II polyketide synthases. *BMC Microbiol* **2012**, *12*, 169.
63. Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A., Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* **2007**, *24* (1), 162-190.
64. Pickens, L. B.; Tang, Y., Decoding and engineering tetracycline biosynthesis. *Metab Eng* **2009**, *11* (2), 69-75.
65. Zhang, W.; Watanabe, K.; Cai, X.; Jung, M. E.; Tang, Y.; Zhan, J., Identifying the minimal enzymes required for anhydrotetracycline biosynthesis. *J Am Chem Soc* **2008**, *130* (19), 6068-6069.
66. Bisang, C.; Long, P. F.; Cortes, J.; Westcott, J.; Crosby, J.; Matharu, A. L.; Cox, R. J.; Simpson, T. J.; Staunton, J.; Leadlay, P. F., A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* **1999**, *401* (6752), 502-505.
67. Decker, H.; Summers, R. G.; Hutchinson, C. R., Overproduction of the acyl carrier protein component of a type II polyketide synthase stimulates production of tetracenomyacin biosynthetic intermediates in *Streptomyces glaucescens*. *J Antibiot* **1994**, *47* (1), 54-63.
68. Shimizu, Y.; Ogata, H.; Goto, S., Type III polyketide synthases: functional classification and phylogenomics. *ChemBiochem* **2017**, *18* (1), 50-65.
69. Kreuzaler, F.; Hahlbrock, K., Enzymatic synthesis of aromatic compounds in higher plants: formation of naringenin (5,7,4'-trihydroxyflavanone) from p-coumaroyl coenzyme A and malonyl coenzyme A. *FEBS Lett* **1972**, *28* (1), 69-72.
70. Funai, N.; Ohnishi, Y.; Fujii, I.; Shibuya, M.; Ebizuka, Y.; Horinouchi, S., A new pathway for polyketide synthesis in microorganisms. *Nature* **1999**, *400* (6747), 897-899.
71. Hashimoto, M.; Nonaka, T.; Fujii, I., Fungal type III polyketide synthases. *Nat Prod Rep* **2014**, *31* (10), 1306-1317.

REFERENCES

72. Austin, M. B.; Noel, J. P., The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep* **2003**, *20* (1), 79-110.
73. Chemler, J. A.; Buchholz, T. J.; Geders, T. W.; Akey, D. L.; Rath, C. M.; Chlipala, G. E.; Smith, J. L.; Sherman, D. H., Biochemical and structural characterization of germicidin synthase: analysis of a type III polyketide synthase that employs acyl-ACP as a starter unit donor. *J Am Chem Soc* **2012**, *134* (17), 7359-7366.
74. Keatinge-Clay, A. T., The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10), 1050-1073.
75. Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G., Biosynthesis of polyketide synthase extender units. *Nat Prod Rep* **2009**, *26* (1), 90-114.
76. Miyanaga, A., Structure and function of polyketide biosynthetic enzymes: various strategies for production of structurally diverse polyketides. *Biosci Biotechnol Biochem* **2017**, *81* (12), 2227-2236.
77. 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.
78. Zhou, H.; Li, Y.; Tang, Y., Cyclization of aromatic polyketides from bacteria and fungi. *Nat Prod Rep* **2010**, *27* (6), 839-868.
79. Klaus, M.; Grninger, M., Engineering strategies for rational polyketide synthase design. *Nat Prod Rep* **2018**, *35* (10), 1070-1081.
80. Jacques, I. B.; Moutiez, M.; Witwinowski, J.; Darbon, E.; Martel, C.; Seguin, J.; Favry, E.; Thai, R.; Lecoq, A.; Dubois, S.; Pernodet, J. L.; Gondry, M.; Belin, P., Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity. *Nat Chem Biol* **2015**, *11* (9), 721-727.
81. 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.
82. Baysarowich, J.; Koteva, K.; Hughes, D. W.; Ejim, L.; Griffiths, E.; Zhang, K.; Junop, M.; Wright, G. D., Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proc Natl Acad Sci U S A* **2008**, *105* (12), 4886-4891.
83. Omura, S.; Imamura, N.; Oiwa, R.; Kuga, H.; Iwata, R.; Masuma, R.; Iwai, Y., Clostomicins, new antibiotics produced by *Micromonospora echinospora* subsp. *armeniaca* subsp. nov. I. Production, isolation, and physico-chemical and biological properties. *J Antibiot* **1986**, *39* (10), 1407-1412.
84. Mingeot-Leclercq, M. P.; Glupczynski, Y.; Tulkens, P. M., Aminoglycosides: activity and resistance. *Antimicrob Agents Chemother* **1999**, *43* (4), 727-737.
85. Croitoru, A.; Babin, M.; Myllykallio, H.; Gondry, M.; Aleksandrov, A., Cyclodipeptide synthases of the NYH subfamily recognize tRNA using an α -Helix enriched with positive residues. *Biochemistry* **2021**, *60* (1), 64-76.
86. Sauguet, L.; Moutiez, M.; Li, Y.; Belin, P.; Seguin, J.; Le Du, M. H.; Thai, R.; Masson, C.; Fonvielle, M.; Pernodet, J. L.; Charbonnier, J. B.; Gondry, M., Cyclodipeptide synthases, a family of class-I aminoacyl-tRNA synthetase-like enzymes involved in non-ribosomal peptide synthesis. *Nucleic Acids Res* **2011**, *39* (10), 4475-4489.
87. Bonnefond, L.; Arai, T.; Sakaguchi, Y.; Suzuki, T.; Ishitani, R.; Nureki, O., Structural basis for nonribosomal peptide synthesis by an aminoacyl-tRNA synthetase paralog. *Proc Natl Acad Sci U S A* **2011**, *108* (10), 3912-3917.
88. Vetting, M. W.; Hegde, S. S.; Blanchard, J. S., The structure and mechanism of the *Mycobacterium tuberculosis* cyclodityrosine synthetase. *Nat Chem Biol* **2010**, *6* (11), 797-799.
89. 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*, 5141.
90. Roback, P.; Beard, J.; Baumann, D.; Gille, C.; Henry, K.; Krohn, S.; Wiste, H.; Voskuil, M. I.; Rainville, C.; Rutherford, R., A predicted operon map for *Mycobacterium tuberculosis*. *Nucleic Acids Res* **2007**, *35* (15), 5085-5095.
91. Winkelblech, J.; Fan, A.; Li, S. M., Prenyltransferases as key enzymes in primary and secondary metabolism. *Appl Microbiol Biotechnol* **2015**, *99* (18), 7379-7397.
92. Alhassan, A. M.; Abdullahi, M. I.; Uba, A.; Umar, A., Prenylation of aromatic secondary metabolites: a new frontier for development of novel drugs. *Trop J Pharm Res* **2014**, *13* (2), 307-314.
93. Heide, L., Prenyl transfer to aromatic substrates: genetics and enzymology. *Curr Opin Chem Biol* **2009**, *13* (2), 171-179.

REFERENCES

94. Li, S.-M., Prenylated indole derivatives from fungi: structure diversity, biological activities, biosynthesis and chemoenzymatic synthesis. *Nat Prod Rep* **2010**, *27* (1), 57-78.
95. Mori, T., Enzymatic studies on aromatic prenyltransferases. *J Nat Med* **2020**, *74* (3), 501-512.
96. Li, W., Bringing bioactive compounds into membranes: The UbiA superfamily of intramembrane aromatic prenyltransferases. *Trends Biochem Sci* **2016**, *41* (4), 356-370.
97. Guo, R. T.; Ko, T. P.; Chen, A. P.; Kuo, C. J.; Wang, A. H.; Liang, P. H., Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: roles of the metal ion and conserved residues in catalysis. *J Biol Chem* **2005**, *280* (21), 20762-20774.
98. Winkelblech, J.; Fan, A.; Li, S. M., Prenyltransferases as key enzymes in primary and secondary metabolism. *Appl Microbiol Biotechnol* **2015**, *99* (18), 7379-7397.
99. Pojer, F.; Wemakor, E.; Kammerer, B.; Chen, H.; Walsh, C. T.; Li, S. M.; Heide, L., CloQ, a prenyltransferase involved in clorobiocin biosynthesis. *Proc Natl Acad Sci U S A* **2003**, *100* (5), 2316-2321.
100. Kuzuyama, T.; Noel, J. P.; Richard, S. B., Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **2005**, *435* (7044), 983-987.
101. Mojtahedi, M. M.; Samadian, S., Efficient and rapid solvent-free acetylation of alcohols, phenols, and thiols using catalytic amounts of sodium acetate trihydrate. *J Chem* **2013**, *2013*, 642479.
102. Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H., Heme-containing oxygenases. *Chem Rev* **1996**, *96* (7), 2841-2888.
103. Beaudry, C. M.; Malerich, J. P.; Trauner, D., Biosynthetic and biomimetic electrocyclizations. *Chem Rev* **2005**, *105* (12), 4757-4778.
104. Wiese, A.; Pietzsch, M.; Syldatk, C.; Mattes, R.; Altenbuchner, J., Hydantoin racemase from *Arthrobacter aurescens* DSM 3747: heterologous expression, purification and characterization. *J Biotechnol* **2000**, *80* (3), 217-230.
105. Vogel, E.; Günther, H., Benzene oxide - oxepin valence tautomerism. *Angew Chem Int Ed Engl* **1967**, *6*, 385-401.
106. Tsai, H. F.; Wang, H.; Gebler, J. C.; Poulter, C. D.; Schardl, C. L., The *Claviceps purpurea* gene encoding dimethylallyltryptophan synthase, the committed step for ergot alkaloid biosynthesis. *Biochem Biophys Res Commun* **1995**, *216* (1), 119-125.
107. Unsöld, I. A.; Li, S. M., Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from *Aspergillus fumigatus*. *Microbiology* **2005**, *151* (5), 1499-1505.
108. Ding, Y.; Wet, J. R.; Cavalcoli, J.; Li, S.; Greshock, T. J.; Miller, K. A.; Finfield, 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.
109. Yin, S.; Yu, X.; Wang, Q.; Liu, X. Q.; Li, S. M., Identification of a brevianamide F reverse prenyltransferase BrePT from *Aspergillus versicolor* with a broad substrate specificity towards tryptophan-containing cyclic dipeptides. *Appl Microbiol Biotechnol* **2013**, *97* (4), 1649-1660.
110. Mundt, K.; Li, S. M., CdpC2PT, a reverse prenyltransferase from *Neosartorya fischeri* with distinct substrate preference from known C2-prenyltransferases. *Microbiology* **2013**, *159* (10), 2169-2179.
111. Schneider, P.; Weber, M.; Hoffmeister, D., The *Aspergillus nidulans* enzyme TdiB catalyzes prenyltransfer to the precursor of bioactive asterriquinones. *Fung Genet Biology* **2008**, *45* (3), 302-309.
112. Schuller, J. M.; Zocher, G.; Liebhold, M.; Xie, X.; Stahl, M.; Li, S. M.; Stehle, T., Structure and catalytic mechanism of a cyclic dipeptide prenyltransferase with broad substrate promiscuity. *J Mol Biol* **2012**, *422* (1), 87-99.
113. Yin, W. B.; Grundmann, A.; Cheng, J.; Li, S. M., Acetylaszonalenin biosynthesis in *Neosartorya fischeri*: identification of the biosynthetic gene cluster by genomic mining and functional proof of the genes by biochemical investigation. *J Biol Chem* **2009**, *284* (1), 100-109.
114. Yin, W. B.; Yu, X.; Xie, X. L.; Li, S. M., Preparation of pyrrolo[2,3-b]indoles carrying a β -configured reverse C3-dimethylallyl moiety by using a recombinant prenyltransferase CdpC3PT. *Org Biomol Chem* **2010**, *8* (10), 2430-2438.
115. Metzger, U.; Schall, C.; Zocher, G.; Unsöld, I.; Stec, E.; Li, S. M.; Heide, L.; Stehle, T., The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. *Proc Natl Acad Sci U S A* **2009**, *106* (34), 14309-14314.

REFERENCES

116. Luk, L. Y. P.; Tanner, M. E., Mechanism of dimethylallyltryptophan synthase: evidence for a dimethylallyl cation intermediate in an aromatic prenyltransferase reaction. *J Am Chem Soc* **2009**, *131* (39), 13932-13933.
117. Kremer, A.; Li, S. M., A tyrosine O-prenyltransferase catalyses the first pathway-specific step in the biosynthesis of sirodesmin PL. *Microbiology* **2010**, *156* (1), 278-286.
118. Fan, A.; Chen, H.; Wu, R.; Xu, H.; Li, S. M., A new member of the DMATS superfamily from *Aspergillus niger* catalyzes prenylations of both tyrosine and tryptophan derivatives. *Appl Microbiol Biotechnol* **2014**, *98* (24), 10119-10129.
119. Liu, C.; Noike, M.; Minami, A.; Oikawa, H.; Dairi, T., Functional analysis of a prenyltransferase gene (paxD) in the paxilline biosynthetic gene cluster. *Appl Microbiol Biotechnol* **2014**, *98* (1), 199-206.
120. Pockrandt, D.; Ludwig, L.; Fan, A.; Konig, G. M.; Li, S. M., New insights into the biosynthesis of prenylated xanthenes: Xptb from *Aspergillus nidulans* catalyses an O-prenylation of xanthenes. *Chembiochem* **2012**, *13* (18), 2764-2771.
121. Chooi, Y. H.; Fang, J.; Liu, H.; Filler, S. G.; Wang, P.; Tang, Y., Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. *Org Lett* **2013**, *15* (4), 780-783.
122. Fan, A.; Winkelblech, J.; Li, S. M., Impacts and perspectives of prenyltransferases of the DMATS superfamily for use in biotechnology. *Appl Microbiol Biotechnol* **2015**, *99* (18), 7399-7415.
123. Rittle, J.; Green, M. T., Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics. *Science* **2010**, *330* (6006), 933-937.
124. Omura, T.; Sato, R., The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature. *J Biol Chem* **1964**, *239*, 2370-2378.
125. Krishnan, S.; Schenkman, J. B.; Rusling, J. F., Bioelectronic delivery of electrons to cytochrome P450 enzymes. *J Phys Chem B* **2011**, *115* (26), 8371-8380.
126. Omura, T.; Sato, R., Fractional solubilization of haemoproteins and partial purification of carbon monoxide-binding cytochrome from liver microsomes. *Biochim Biophys Acta* **1963**, *71*, 224-226.
127. Harding, B. W.; Wong, S. H.; Nelson, D. H., Carbon monoxide-combining substances in rat adrenal. *Biochim Biophys Acta* **1964**, *92*, 415-417.
128. Cooper, D. Y.; Levin, S.; Narasimhulu, S.; Rosenthal, O., Photochemical action spectrum of the terminal oxidase of mixed function oxidase systems. *Science* **1965**, *147* (3656), 400-402.
129. Cooper, D. Y.; Narasimhulu, S.; Slade, A.; Raich, W.; Foroff, O.; Rosenthal, O., Hemoprotein content and activity of solubilized steroid 11 beta-hydroxylase preparations from adrenocortical mitochondria. *Life Sci* **1965**, *4* (21), 2109-2114.
130. Katagiri, M.; Ganguli, B. N.; Gunsalus, I. C., A soluble cytochrome P450 functional in methylene hydroxylation. *J Biol Chem* **1968**, *243* (12), 3543-3546.
131. Wienkers, L. C.; Heath, T. G., Predicting in vivo drug interactions from in vitro drug discovery data. *Nat Rev Drug Discov* **2005**, *4* (10), 825-833.
132. Narhi, L. O.; Fulco, A. J., Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem* **1986**, *261* (16), 7160-7169.
133. Claude, A., The Constitution of Protoplasm. *Science* **1943**, *97* (2525), 451-456.
134. Estabrook, R. W.; Hildebrandt, A. G.; Baron, J.; Netter, K. J.; Leibman, K., A new spectral intermediate associated with cytochrome P-450 function in liver microsomes. *Biochem Biophys Res Commun* **1971**, *42* (1), 132-139.
135. Ichihara, K.; Kusunose, E.; Kusunose, M., Some properties of NADPH-cytochrome c reductase reconstitutively active in fatty-acid omega-hydroxylation. *Eur J Biochem* **1973**, *38* (3), 463-472.
136. Yamazaki, H., *Fifty years of cytochrome P450 research*. Springer: 2014.
137. Yamazaki, H.; Johnson, W. W.; Ueng, Y.-F.; Shimada, T.; Guengerich, F. P., Lack of electron transfer from cytochrome b5 in stimulation of catalytic activities of cytochrome P450 3A4. Characterization of a reconstituted cytochrome P450 3A4/NADPH-cytochrome P450 reductase system and studies with apo-cytochrome b5. *J Biol Chem* **1996**, *271* (44), 27438-27444.
138. Guengerich, F. P., Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* **2001**, *14* (6), 611-650.
139. Nebert, D. W.; Russell, D. W., Clinical importance of the cytochromes P450. *Lancet* **2002**, *360* (9340), 1155-1162.

REFERENCES

140. Dennig, A.; Kuhn, M.; Tassoti, S.; Thiessenhusen, A.; Gilch, S.; Bulter, T.; Haas, T.; Hall, M.; Faber, K., Oxidative decarboxylation of short-chain fatty acids to 1-alkenes. *Angew Chem Int Ed Engl* **2015**, *54* (30), 8819-8822.
141. Lim, H.; Jeon, H.; Hong, S.; Kim, J. H., Catalytic approach to in vivo metabolism of atractylenolide III using biomimetic iron-porphyrin complexes. *Rsc Adv* **2021**, *11* (52), 33048-33054.
142. Li, X.; Fu, J.; Li, Y.; Liu, J.; Gao, R.; Shi, Y.; Li, Y.; Sun, H.; Wang, L.; Li, Y.; Jiang, B.; Wu, L.; Hong, B., Cytochrome P450 monooxygenase for catalyzing C-42 Hydroxylation of the Glycine-derived fragment in hangtairyacin biosynthesis. *Org Lett* **2022**, *24* (6), 1388-1393.
143. Rudolf, J. D.; Chang, C. Y.; Ma, M.; Shen, B., Cytochromes P450 for natural product biosynthesis in *Streptomyces*: sequence, structure, and function. *Nat Prod Rep* **2017**, *34* (9), 1141-1172.
144. Abachi, S.; Abbott, B.; Abolins, M.; Acharya, B. S.; Adam, I. I.; Adams, D. L.; Adams, M.; Ahn, S.; Aihara, H.; Alitti, J.; Alvarez, G.; Alves, G. A.; Amidi, E.; Amos, N.; Anderson, E. W.; Aronson, S. H.; Astur, R.; Avery, R. E.; Baden, A.; Balamurali, V. V.; Balderston, J.; Baldin, B.; Bantly, J.; Bartlett, J. F.; Bazizi, K.; Bendich, J.; Beri, S. B.; Bertram, I. I.; Bezzubov, V. A.; Bhat, P. C.; Bhatnagar, V. V.; Bhattacharjee, M.; Bischoff, A.; Biswas, N.; Blazey, G.; Blessing, S.; Bloom, P.; Boehnlein, A.; Bojko, N. I.; Borchering, F.; Borders, J.; Boswell, C.; Brandt, A.; Brock, R.; Bross, A.; Buchholz, D.; Burtovoi, V. S.; Butler, J. M.; Carvalho, W.; Casey, D.; Castilla-Valdez, H.; Chakraborty, D.; Chang, S.; Chekulaev, S. V.; Chen, L.; Chen, W.; Chevalier, L.; Chopra, S.; Choudhary, B. C.; Christenson, J. H.; Chung, M.; Claes, D.; Clark, A. R.; Cobau, W. G., Studies of topological distributions of inclusive three- and four-jet events in p-barp collisions at s=1800 GeV with the D0 detector. *Phys Rev D Part Fields* **1996**, *53* (11), 6000-6016.
145. Reynald, R. L.; Sansen, S.; Stout, C. D.; Johnson, E. F., Structural characterization of human cytochrome P450 2C19: active site differences between P450s 2C8, 2C9, and 2C19. *J Biol Chem* **2012**, *287* (53), 44581-44591.
146. Fleming, B. D.; Johnson, D. L.; Bond, A. M.; Martin, L. L., Recent progress in cytochrome P450 enzyme electrochemistry. *Expert Opin Drug Metab Toxicol* **2006**, *2* (4), 581-589.
147. Kleywegt, G. J.; Jones, T. A., Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr D Biol Crystallogr* **1994**, *50* (Pt 2), 178-185.
148. Rusling, J. F.; Hvastkovs, E. G.; Hull, D. O.; Schenkman, J. B., Biochemical applications of ultrathin films of enzymes, polyions and DNA. *Chem Commun* **2008**, (2), 141-154.
149. Krishnan, S.; Wasalathanthri, D.; Zhao, L.; Schenkman, J. B.; Rusling, J. F., Efficient bioelectronic actuation of the natural catalytic pathway of human metabolic cytochrome P450s. *J Am Chem Soc* **2011**, *133* (5), 1459-1465.
150. Yu, Z.; Zhang, H.; Yuan, C.; Zhang, Q.; Khan, I.; Zhu, Y.; Zhang, C., Characterizing two cytochrome P450s in tiacumicin biosynthesis reveals reaction timing for tailoring modifications. *Org Lett* **2019**, *21* (18), 7679-7683.
151. Shepherd, M. D.; Kharel, M. K.; Zhu, L. L.; van Lanen, S. G.; Rohr, J., Delineating the earliest steps of gilvocarcin biosynthesis: role of GilP and GilQ in starter unit specificity. *Org Biomol Chem* **2010**, *8* (17), 3851-3856.
152. Simpson, T. J.; Stenzel, D. J., ¹³C and ²H nmr studies on the biosynthesis of O-methylasparvenone, a hexaketide metabolite of *Aspergillus parvulus*. *J Chem Soc, Chem Commun* **1981**, (5), 239-240.
153. El Maddah, F.; Eguereva, E.; Kehraus, S.; Konig, G. M., Biosynthetic studies of novel polyketides from the marine sponge-derived fungus *Stachylidium* sp. 293K04. *Org Biomol Chem* **2019**, *17* (10), 2747-2752.
154. De Jesus, A. E.; Horak, R. M.; Steyn, P. S.; Vlegaar, R., Metabolites of *Aspergillus ustus*. Part 4. Stable-isotope labelling studies on the biosynthesis of the australides. *J Chem Soc* **1987**, (1), 2253-2257.
155. Brock, M.; Fischer, R.; Linder, D.; Buckel, W., Methylcitrate synthase from *Aspergillus nidulans*: implications for propionate as an antifungal agent. *Mol Microbiol* **2000**, *35* (5), 961-973.
156. Xu, W.; Gavia, D. J.; Tang, Y., Biosynthesis of fungal indole alkaloids. *Nat Prod Rep* **2014**, *31* (10), 1474-1487.
157. Li, S.-M., Genome mining and biosynthesis of fumitremorgin-type alkaloids in ascomycetes. *J Antibiot* **2011**, *64* (1), 45-49.
158. Fraley, A. E.; Sherman, D. H., Enzyme evolution in fungal indole alkaloid biosynthesis. *FEBS J* **2020**, *287* (7), 1381-1402.

REFERENCES

159. Wang, X.; Li, Y.; Zhang, X.; Lai, D.; Zhou, L., Structural diversity and biological activities of the cyclodipeptides from fungi. *Molecules* **2017**, *22* (12), 2026.
160. Mishra, A. K.; Choi, J.; Choi, S.-J.; Baek, K.-H., Cyclodipeptides: an overview of their biosynthesis and biological activity. *Molecules* **2017**, *22* (10), 1796.
161. Fan, A.; Winkelblech, J.; Li, S. M., Impacts and perspectives of prenyltransferases of the DMATS superfamily for use in biotechnology. *Appl Microbiol Biotechnol* **2015**, *99* (18), 7399-7415.
162. Zhu, Y.; Fu, P.; Lin, Q.; Zhang, G.; Zhang, H.; Li, S.; Ju, J.; Zhu, W.; Zhang, C., Identification of caerulomycin A gene cluster implicates a tailoring amidohydrolase. *Org Lett* **2012**, *14* (11), 2666-2669.
163. Zhang, Y.; Yao, T.; Jiang, Y.; Li, H.; Yuan, W.; Li, W., Deciphering a cyclodipeptide synthase pathway encoding prenylated indole alkaloids in *Streptomyces leeuwenhoekii*. *Appl Environ Microbiol* **2021**, *87* (11), e02525-20.
164. Harken, L.; Li, S.-M., Modifications of diketopiperazines assembled by cyclodipeptide synthases with cytochrome P450 enzymes. *Appl Microbiol Biotechnol* **2021**, *105* (6), 2277-2285.
165. Liu, J.; Xie, X.; Li, S. M., Guanitrypmycin biosynthetic pathways imply cytochrome P450 mediated regio- and stereospecific guaninyl-transfer reactions. *Angew Chem Int Ed Engl* **2019**, *58* (33), 11534-11540.
166. Liu, J.; Li, S. M., Genomics-guided efficient identification of 2,5-Diketopiperazine derivatives from actinobacteria. *ChemBiochem* **2022**, e202200502.
167. Malit, J. J. L.; Liu, W.; Cheng, A.; Saha, S.; Liu, L.-L.; Qian, P.-Y., Global genome mining reveals a cytochrome P450-catalyzed cyclization of crownlike cyclodipeptides with neuroprotective activity. *Org Lett* **2021**, *23* (17), 6601-6605.
168. Yu, H.; Xie, X.; Li, S.-M., Coupling of Guanine with *cyclo*-L-Trp-L-Trp mediated by a cytochrome P450 homologue from *Streptomyces purpureus*. *Org Lett* **2018**, *20* (16), 4921-4925.
169. Yu, H.; Xie, X.; Li, S.-M., Coupling of *cyclo*-L-Trp-L-Trp with hypoxanthine increases the structure diversity of guanitrypmycins. *Org Lett* **2019**, *21* (22), 9104-9108.
170. Harken, L.; Liu, J.; Kreuz, O.; Berger, R.; Li, S.-M., Biosynthesis of guatrypmethine C implies two different oxidases for exo double bond installation at the diketopiperazine ring. *Acs Catal* **2022**, *12* (1), 648-654.
171. Liu, J.; Harken, L.; Yang, Y.; Xie, X.; Li, S. M., Widely distributed bifunctional bacterial cytochrome P450 enzymes catalyze both intramolecular C-C bond formation in *cyclo* - L - Tyr - L - Tyr and its coupling with nucleobases. *Angew Chem Int Ed Engl* **2022**, *61* (21), e202200377.
172. Mikulski, L.; Schafer, J.; Brockmeyer, K.; Kraut, R.; Li, S. M., Comparative studies on similarities and differences of cyclodipeptide oxidases for installation of C-C double bonds at the diketopiperazine ring. *Appl Microbiol Biotechnol* **2020**, *104* (6), 2523-2536.

Statutory Declaration

Ich, Yiling Yang, versichere, dass ich meine Dissertation

„Investigation on the biosynthesis of polyketide products in *Aspergillus ustus* and cyclodipeptide derivatives in *Streptomyces* strains“

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den.....

.....

Yiling Yang

Acknowledgements

My PhD studies at the Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, der Philipps-Universität Marburg are coming to an end. Looking back on this period of time, I have gained a lot, and I have a lot to thank!

First of all, I would like to express my sincere gratitude to Prof. Dr. Shuming Li for giving me the opportunity to study as a PhD student. I sincerely thank him for his valuable advice, guidance, support and encouragement. It is of great value for his approach to work, advice on projects, effort on manuscripts, and his help in my daily life. Everything I have learned and trained in the research group of Prof. Li will be of great benefit to my future life.

I'm very grateful to Prof. Dr. Michael Keusgen for his agreement to be my second supervisor and examiner. Furthermore, I would like to thank Prof. Dr. Cornelius Krasel and Prof. Dr. Cornelia M. Keck for their kindness to be coexaminers.

Many thanks Dr. Jing Liu and Dr. Liujuan Zheng for their help in the cooperation in this thesis and proof reading of the dissertation. I also really appreciate the efforts made by Jenny Zhou for translating the summary part and proofreading the dissertation. I would also like to thank Lena Ludwig-Radtke for measuring LC-MS samples and assisting with CD spectra, Rixa Kraut for measuring LC-MS samples, Dr. Xiulan Xie, Regina Ortmann and Stefan Newel for recording NMR spectra.

Special thanks Dr. Pan Xiang, Dr. Ge Liao, Haowen Wang, Wen Li, Yu Dai, Zhanghai Li, Xuping Zhang, Zhengxi Zhang and Dr. Jing Zhou for the wonderful time we have spent together. I also want to express my sincere gratitude to all my current and former colleagues, Dr. Huili Yu, Dr. Jie Fan, Dr. Huomiao Ran, Dr. Lauritz Harken, Dr. Lindsay Coby, Dr. Jonas Nies, Dr. Florian Kindinger, Dr. Katja Backhaus, Dr. Wei Li, Dr. Yuan Zhou, Sina Stierle, Johanna Schäfer, Marlies Peter, Daniel Jonathan Janzen, Andreas Martin, Daniel Ostendorff, Hendrik Sostmann, David Breyer, Philipp Mann, Kristin Öqvist, Huiling Wei, Meiting Wu, Xiaoling Chen, Sabine Burgers, Sonja Hiemenz and Dr. Dieter Kreuzsch.

I would like to express my special thanks to Prof. Dr. Haofu Dai, especially for introducing me into the scientific field of biosynthesis of natural products and his encouragement during these years.

I wish to thank the constant encouragement from my friends outside of the laboratory.

I would like to acknowledge the China Scholarship Council (CSC) for financial supports.

Finally, I must extend my heartfelt thanks to my family. Deeply thanks to my parents for giving me love and supports as always.

