Investigation on Cyclodipeptide Formation and Metabolism in Actinomycetes and Application of Fungal Prenyltransferases in the Chemoenzymatic Synthesis

Untersuchungen zur Entstehung und Metabolismus von zyklischen Dipeptiden in Aktinomyzeten sowie Anwendungen von pilzlichen Prenyltransferazen in der chemoenzymatischen Synthese

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

dem Fachbereich Pharmazie
der Philipps-Universität Marburg

vorgelegt von

Huili Yu
aus Tai'an, China

Marburg an der Lahn, 2018

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

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

Eingereicht am 16. August 2018

Tag der mündlichen Prüfung: 02. Oktober 2018

Hochschulkennziffer: 1180

Dedicated to my family

(献给我的家人)

Table of Contents

List of publications

1. Liu, J.*, **Yu, H.***, & Li, S.-M. (2018). Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains. *Appl. Microbiol. Biotechnol.*, *102* (10), 4435-4444. [*equal contribution]

   Liu, J*., **Yu, H*.**, & Li, S.-M. (2018). Correction to: Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains. *Appl. Microbiol. Biotechnol.*, *102*, 5787-5789. [*equal contribution]
2. **Yu, H**., Xie, X., & Li, S.-M. (2018). Coupling of guanine with *cyclo*-L-Trp-L-Trp mediated by a cytochrome P450 homologue from *Streptomyces purpureus*. *Org. Lett*., *20* (16), 4921-4925.
3. **Yu, H.**, Liebhold, M., Xie, X., & Li, S.-M. (2015). Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate. *Appl. Microbiol. Biotechnol*., *99* (17), 7115-7124.

Share of author contributions

Publications	**Authors**	**Estimated equity ratio [%]**	**Status**
Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate. (*Appl. Microbiol. Biotechnol*.)	Yu, H.; Liebhold, M.; Xie, X. and Li, S.-M.	60	published
Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains. (*Appl. Microbiol. Biotechnol*.)	Liu, J.*; Yu, H.* and Li, S.-M.	40	published
Coupling of guanine with *cyclo*-L-Trp-L-Trp mediated by a cytochrome P450 homologue from *Streptomyces purpureus*. (*Org. Lett*.)	Yu, H.; Xie, X. and Li, S.-M.	65	published

*: These authors contributed equally to this work.

Doctoral candidate: Supervisor:

Academic activities

Yu, H., Liebhold, M., Xie, X., & Li, S. M.

Enzymatic *N*- and *O*-alkylation/benzylation of tyrosine and derivatives

Poster presentation, Annual Conference of the Association for General and Applied Microbiology (VAAM), 01–03. March 2015, Marburg

Abbreviations

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

CDP	cyclodipeptide
CDPS	cyclodipeptide synthase
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthetase
DKP	diketopiperazine
E. coli	*Escherichia coli*
HPLC	high performance liquid chromatography
LC-MS	liquid chromatography–mass spectrometry
HR-EIMS	high-resolution electron ionization mass spectrometry
NMR	nuclear magnetic resonance spectroscopy
cWX	tryptophan-containing cyclodipeptide
cWW	*cyclo*-L-Trp-L-Trp
cWA	*cyclo*-L-Trp-L-Ala
cWL	*cyclo*-L-Trp-L-Leu
cWP	*cyclo*-L-Trp-L-Pro
cWY	*cyclo*-L-Trp-L-Tyr
DMATS	dimethylallyltryptophan synthase
DMAPP	dimethylallyl diphosphate
GPP	geranyl diphosphate
FPP	farnesyl diphosphate
MAPP	methylallyl diphosphate
2-pen-PP	2-pentenyl diphosphate
benzyl-PP	benzyl diphosphate

GC-content	guanine-cytosine content
A domain	adenylation domain
T domain	thiolation domain
C domain	condensation domain
TE domain	thioesterase domain
aa-tRNA	aminoacyl-tRNA
ncAA	non-canonical amino acid
α-KGD	α-ketoglutarate/Fe^{II}-dependent dioxygenase
CDO	cyclic dipeptide oxidase
cWF	*cyclo*-L-Trp-L-Phe
cWH	*cyclo*-L-Trp-L-His
cFS	*cyclo*-L-Phe-L-Ser
cFL	*cyclo*-L-Phe-L-Leu
cLL	*cyclo*-L-Leu-L-Leu
cYY	*cyclo*-L-Tyr-L-Tyr
cFY	*cyclo*-L-Phe-L-Tyr
cFF	*cyclo*-L-Phe-L-Phe
cIL	*cyclo*-L-Ile-L-Leu
cFH	*cyclo*-L-Phe-L-His
P450	cytochrome P450
P450cam	cytochrome P450 metabolizing camphor in *Pseudomonas putida*
PDB	protein data bank
NADH	reduced nicotinamide adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate

Cpd 0	compound 0
Cpd I	compound I
Cpd II	compound II
S. coelicolor	*Streptomyces coelicolor*
PT	prenyltransferase
m/z	mass-to-charge ratio
Ni-NTA	nickel-nitrilotriacetic acid
kDa	kilo Dalton
K_M	Michaelis-Menten constant
k_{cat}	turnover number
UV-Vis	ultraviolet-visible
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
OD_{600}	optical density at 600 nm
TB	Terrific-Broth
DNA	deoxyribonucleic acid
gDNA	genomic DNA
PCR	polymerase chain reaction
PPi	inorganic pyrophosphate
ppm	parts per million
J	coupling constant
δ_C	chemical shift of ^{13}C
δ_H	chemical shift of ^{1}H
MHz	mega hertz
s	singlet
d	doublet

t	triplet
q	quartet
m	multiplet
APT	attached proton test
COSY	correlation spectroscopy
DQF-COSY	double-quantum filtered COSY
HSQC	heteronuclear single quantum coherence
HMBC	heteronuclear multiple bond correlation
D_2O	deuterium oxide
$CDCl_3$	deuterated chloroform
CD_3OD	deuterated methanol
DMSO-*d*6	deuterated dimethyl sulfoxid

Summary

Secondary metabolites originated from plants, bacteria and fungi constitute a large group of compounds, which are not directly related to the growth, development and reproduction of the organism. However, the producer can benefit from these metabolites to defend itself against natural enemies, competitors or environmental pressure. Furthermore, they often exhibit useful and interesting biological and pharmaceutical activities, which provide benefits to human health and life. For example, cyclodipeptides (CDPs) and their derivatives have gained increasing attention for their interesting and diverse pharmaceutical activities. Natural CDPs are mainly formed by nonribosomal peptide synthetases (NRPSs) or cyclodipeptide synthases (CDPSs). The CDP cores resulted from the NRPS or CDPS pathways are often further modified by varieties of tailoring enzymes leading to the formation of sophisticated natural products. The small size of CDPSs (200−300 residues) makes them genetically easily manipulated in synthetic biology.

In the first project of this thesis, a cooperation project with Jing Liu, nine new CDPS genes from eight *Streptomyces* strains were functionally characterized after cloning into pET28a vector and expressing in *E. coli*. The products were analyzed on LC-MS and then isolated for NMR analysis. The products were identified as tryptophan-containing cyclodipeptides (cWXs) that are the most common precursors of diketopiperazines (DKPs) with pharmaceutical interest, due to the various modification possibilities on the indole ring. The nine identified CDPSs include six specific CDP synthases consisting of one cWL synthase, two cWP synthases and three new cWW synthases as well as three unspecific CDPSs producing several products with cWA or cWY as the major cyclodipeptide. Total product yields of CDPs were calculated based on HPLC analysis to be 46 to 211 mg/L *E. coli* culture. Our findings represent rare examples of cWXs derived from CDPS pathways in actinobacteria. This study also provides a valid experimental basis for further combination of CDPSs with other tailoring enzymes in biosynthesis of interesting compounds.

The three identified cWW synthases from *Streptomyces purpureus* NRRL B-5737, *Streptomyces lavendulae* NRRL B-2774, and *Streptomyces xanthophaeus* NRRL B-5414 are genetically associated with a putative cytochrome P450. The three P450 homologues share sequence identities of more than 75% with each other on the amino acid level, indicating same or very similar function. As a representative, the function of $P450_{NB5737}$ from *S. purpureus* was identified by heterologous expression in *Streptomyces coelicolor* and *in vitro* assays with the recombinant protein produced in *E. coli*. Isolation and structure elucidation confirmed that $P450_{NB5737}$ catalyzes the transfer of a guanine moiety onto the indole ring of cWW *via* a C–N bond, leading to unprecedented adducts of cyclodipeptide with a nucleobase. This represents one additional example of rare P450 enzymes catalyzing the coupling of two different substrates.

Besides CDPs, prenylated aromatic products often possess various biological and pharmacological activities, in comparison to their non-prenylated precursors. The prenyltransferases of the DMATS superfamily show broad flexibility toward aromatic substrates, but relatively strict prenyl donor tolerance. They usually accept only DMAPP, but not GPP or other prenyl donors with a longer chain length. To broaden the prenyl donor spectra, the acceptance of tyrosine *O*-prenyltransferases TyrPT and SirD towards unnatural DMAPP analogs such as MAPP, 2-pen-PP, and benzyl-PP was tested in the presence of tyrosine and its derivatives. The results revealed that TyrPT and SirD also use the unnatural donors and exhibit similar behavior as their natural prenyl donor DMAPP. Alkylated or benzylated products at the same positions of that with DMAPP have therefore been identified. This study increased significantly the diversity of prenylated products by chemoenzymatic synthesis.

The results presented in this thesis contributed significantly to the formation and metabolism of tryptophan-containing CDPs in nature and expanded the potential application of prenyltransferases. They provide a solid basis for applications in the synthetic biology in the future.

Zusammenfassung

Sekundärmetabolite aus Pflanzen, Bakterien und Pilzen stellen eine große Substanzklasse dar, die nicht direkt mit Wachstum, Entwicklung und Reproduktion des Organismus zusammenhängen. Jedoch kann der Produzent sich diese Metaboliten zu Nutze machen, um sich vor natürlichen Feinden, Konkurrenten oder Umwelteinflüssen zu schützen. Darüber hinaus weisen sie oft nützliche und interessante biologische und pharmazeutische Aktivitäten auf, wovon die menschliche Gesundheit und das Leben profitieren können. Zum Beispiel haben zyklische Dipeptide (CDPs) und ihre Derivate an zunehmender Aufmerksamkeit in Bezug auf ihre interessanten und vielfältigen pharmazeutischen Aktivitäten gewonnen. Natürlich vorkommende CDPs werden in der Regel von nichtribosomalen Peptidsynthetasen (NRPSs) oder Cyclodipeptidsynthasen (CDPSs) gebildet. Die von der NRPS- oder CDPS-Synthesewege gebildeten CDP-Grundstrukturen werden häufig mithilfe von verschiedenen „Tailoring enyzmes“ weiter modifiziert, sodass komplexere Naturstoffe entstehen. Wegen ihrer geringen Größe (200–300 Aminosäuren) können CDPSs gentechnisch leicht manipuliert werden.

In dem ersten Projekt dieser Arbeit, einem Kooperationsprojekt mit Jing Liu, wurden neun neue CDPS-Gene aus acht *Streptomyceten*-Stämmen jeweils in einen pET28a Vektor kloniert und in *E. coli* exprimiert. Die Produkte wurden mittels LC-MS analysiert und anschließend für die Strukturaufklärung isoliert. Die Produkte wurden als zyklische Dipeptide mit einer Tryptophaneinheit (cWXs) identifiziert. Diese sind pharmazeutisch besonders interessant, da sie vielseitig am Indolring modifiziert werden können. Sechs der neun identifizierten CDPSs sind spezifische CDP-Synthasen, davon eine cWL Synthase, zwei cWP Synthasen und drei neuen cWW Synthasen, sowie drei unspezifische CDPSs, die cWA oder cWY als Hauptprodukt, bilden.

Die Gesamtmenge an produzierten CDPs von 46 bis 211 mg/L in *E. coli* Kultur wurde mittels HPLC berechnet. Unsere Ergebnisse repräsentieren seltene Beispiele an cWXs, die aus den CDPS-Synthesewegen von Actinobakterien hervorgehen. Diese Arbeit liefert auch eine valide experimentelle Grundlage für weitere Kombinationen von CDPSs mit anderen modifizierende Enzymen in der synthetischen Biologie.

Die drei identifizierten cWW-Synthasen aus *Streptomyces purpureus* NRRL B-5737, *Streptomyces lavendulae* NRRL B-2774 und *Streptomyce xanthophaeus* NRRL B-5414

sind genetisch mit putativen Cytochrom P450 Enzymen assoziiert. Die drei P450 Homologe weisen untereinander eine Identität der Aminosäuresequenz von mehr als 75% auf, was auf die gleiche oder eine sehr ähnliche Funktion hindeutet.

Als repräsentativer Vertreter der drei P450s wurde die Funktion von $P450_{NB5737}$ aus *S. purpureus* durch heterologe Expression in *Streptomyces coelicolor*, sowie mithilfe von *in vitro* Assays mit dem rekombinanten Protein untersucht. Die Übertragung einer Guanineinheit auf den Indolring von cWW mit Ausbildung einer C–N-Bindung durch $P450_{NB5737}$ wurde nach Isolierung und anschließender Strukturaufklärung bestätigt. Diese Verknüpfung stellt die Katalyse eines noch nie nachgewesenen Addukts zwischen einem zyklisches Dipeptid mit einer Nukleinbase dar. Dies ist ein weiteres Beispiel von seltenen P450 Enzymen, welche die Verknüpfungen zwei unterschiedlicher Substrate katalysierten.

Im Vergleich zu ihren nicht prenylierten Vorläufern, zeigen prenylierte aromatische Produkte oft diverse biologische und pharmazeutische Aktivitäten. Die Prenyltransferasen der DMATS-Superfamilie sind ihren aromatischen Substraten gegenüber sehr flexibel, allerdings weisen sie ihrem Prenyldonor gegenüber nur eine geringe Toleranz auf. In der Regel akzeptieren sie nur DMAPP, jedoch nicht GPP oder andere Prenyldonoren mit einer noch längeren Kettenlänge. Um das Spektrum des Prenyldonors zu erweitern, wurde die Akzeptanz von zwei Tyrosin-*O*-Prenyltransferasen TyrPT und SirD gegenüber unnatürlichen DMAPP Analoga wie MAPP, 2-pen-PP und Benzyl-PP in Anwesenheit von Tyrosin und Derivaten getestet. Die Ergebnisse zeigten, dass TyrPT und SirD auch die Übertragung der unnatürlichen Donoren katalysieren und dabei ein ähnliches Verhalten wie mit dem natürlichen Prenyldonor DMAPP beibehalten. Produkte, die die Alkyl- bzw. Benzylgruppe an denselben Positionen wie mit DMAPP tragen, wurden daher identifiziert. Diese Arbeit trägt erheblich zur Vielfalt von prenylierten Produkten durch chemoenzymatische Synthese bei.

Die Ergebnisse dieser Dissertation tragen maßgeblich zur Bildung und zum Metabolismus von tryptophanhaltigen CDPs in der Natur bei, sowie die Erweiterung potenzieller Anwendungen von Prenyltransferasen. Darüber hinaus bilden sie eine solide Grundlage für zukünftige Anwendungen in der synthetischen Biologie.

1 Introduction

1.1 Actinobacteria

Actinobacteria is one of the largest phyla of Domain Bacteria. It contains gram-positive bacteria with high guanine-cytosine (GC) content from 51% to over 70% in their DNA, with some exceptions of less than 50% G+C[1,2]. They are widely distributed in both terrestrial and aquatic habitats[3]. Containing most members of economic or medical significance, the order of Actinomycetales often grows extensive mycelia and can be reproduced by spores or hyphae. Behaving like fungi as decomposer in soil, they exhibit economic importance for agriculture and forests. Actinomycetes also play important roles in drug discovery and contribute to 61% of all microorganism-derived bioactive compounds so far discovered[4]. Two-thirds of known antibiotics in current use are originated from actinomycetes, with *Streptomyces* as the main source[4,5]. Since the discovery of streptomycin from *Streptomyces griseus* in 1943, numerous antibiotics and their derivatives within different chemical classes such as β-lactams, tetracyclines, aminoglycosides, macrolides, and glycopeptides have been discovered in actinomycetes and used clinically in modern medicine for treatment of different diseases[6]. However, the increasing occurrence of antibiotic-resistant pathogenic organisms[7] and the remarkable decline in discovery of novel antimicrobial agents[8] make the public heath confront with great challenges. To solve the antibiotic-resistance problem and identify more bioactive compounds, intensive researches on novel approaches for natural product discovery have been performed in recent decades. One strategy is to activate silent or cryptic gene clusters of secondary metabolites by the induction *in vivo* or heterologous expression in engineered *Streptomyces* host with clear background[9]. Other efforts include exploiting rare actinomycetes in extreme habitats, metagenomics analysis of uncultured environmental samples, improvement of analytical methods[10], and elicitation of secondary metabolite with biological, chemical and molecular strategies[11].

1.2 Cyclodipeptides

Cyclodipeptides (CDPs) and their derivatives constitute a large class of secondary metabolites with potential biological and pharmacological activities[12,13]. For instance, *cyclo*-L-His-L-Tyr (**1**) and *cyclo*-L-His-L-Phe (**2**) could be employed as antiarrhythmic agents[14]. Compound **2** and its derivative phenylahistin (**3**) show significant antitumor activity[14,15]. *cyclo*-L-Phe-L-Pro (**4**) and *cyclo*(L-Phe-*trans*-4-OH-L-Pro) (**5**) act as antifungal agents[16]. *cyclo*-L-Leu-L-Pro (**6**) and echinulin (**7**) inhibit mycobacteria[17,18]. *cyclo*-L-His-L-Pro (**8**) and *cyclo*(dehydroala-L-Leu) (**9**) represent antihyperglycaemic agents[19,20] and **8** also exerts anti-inflammatory effects (Figure 1)[21]. The main structure of CDPs are generated by the condensation of two α-amino acids, which can be achieved through total chemical synthesis or biosynthesis catalyzed by two groups of enzymes, non-ribosomal peptide synthetases (NRPSs) and cyclodipeptide synthases (CDPSs), in nature[13].

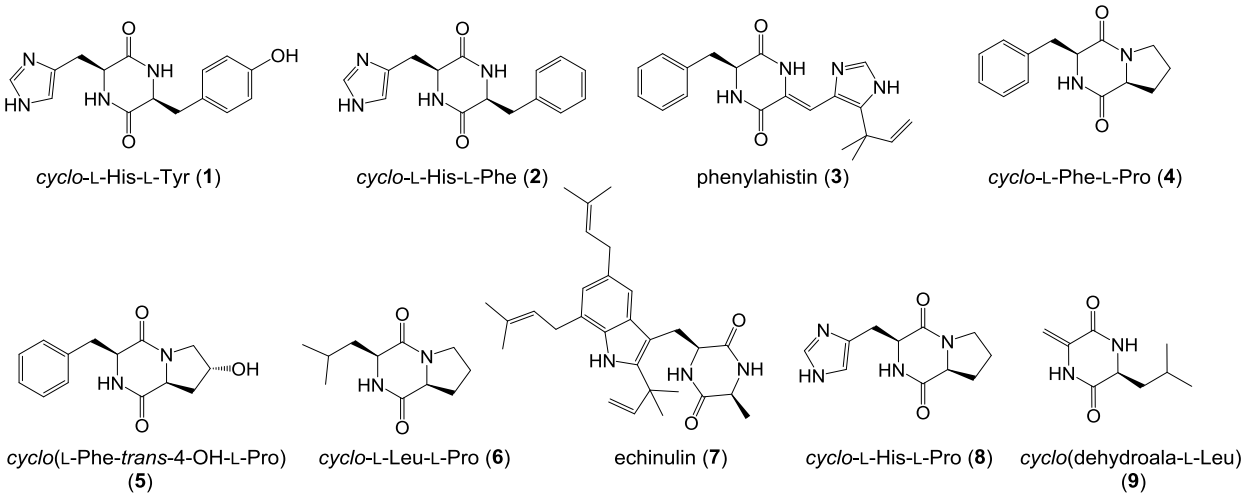


Figure 1. Selected examples of cyclodipeptides and derivatives with pharmacological activities.

1.2.1 CDPs from NRPS pathways

Non-ribosomal peptide synthetases (NRPSs) are large multimodular enzymes that assemble complex peptides with diverse structures and biological activities in microorganisms such as bacteria and fungi[22]. In the single polypeptide chain of NRPS lies an array of large, repetitive modules and each module comprises several catalytic domains. As shown in Figure 2, one elongation module contains at least three domains, adenylation (A), thiolation (T), and condensation (C) domain. The initiation module has no C domain and the termination module has an additional thioesterase (TE) domain. The non-ribosomal peptides (NRPs) assembly initiates with the activation of the amino acid substrate resulting in an aminoacyl-AMP intermediate under the catalysis of A domain. The first activated intermediate is then transferred to the thiol group tethered of T domain and condensed with the second intermediate bound to the adjacent T domain to form the peptide bond, which is catalyzed by C domain. After condensation with the last amino acid, the final product is released by hydrolysis or cyclization catalyzed by TE domain[23]. Different to the normal ribosomal-dependent synthesis

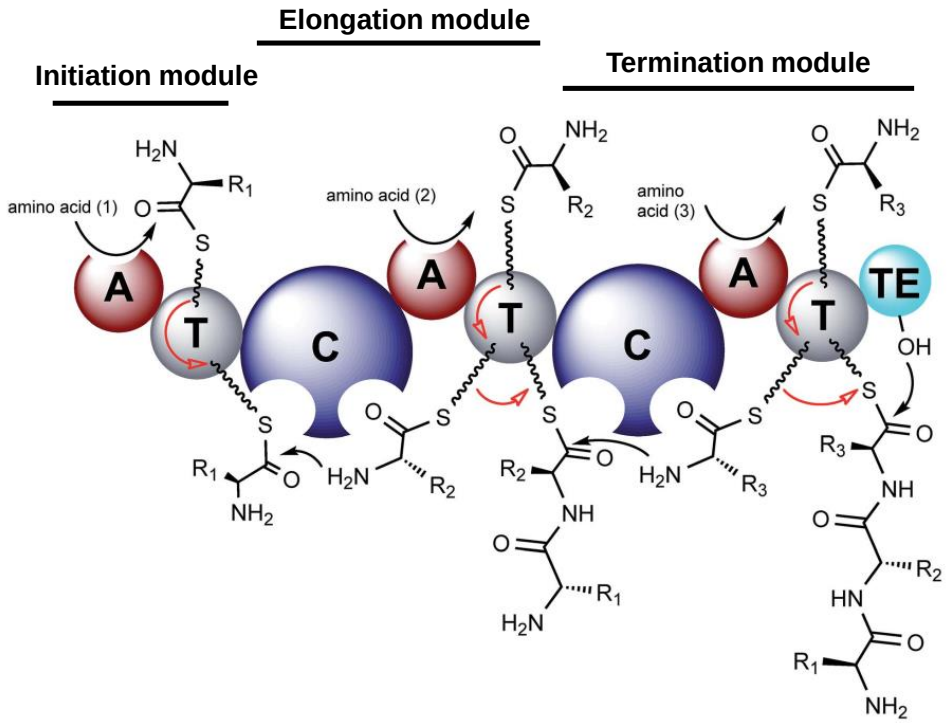


Figure 2. Model of NRPs biosynthesis. (Modified from review [23])

machinery that incorporates only the proteinogenic amino acids, NRPSs can catalyze additionally the assembly of hundreds of other building blocks like D- and β-amino acids, α-keto- and α-hydroxy acids and even heterocycles and fatty acids[24], which increases significantly the structural diversity of natural products. The biosynthetic origins of many CDP-containing natural products in fungi and some bacteria are catalyzed by NRPSs, through dedicated biosynthetic pathways (bimodular NRPS) or through premature release of dipeptidyl intermediates as side CDP products during the biosynthesis of a larger peptide by NRPSs[25-27]. Several bioactive compounds such as thaxtomin (**10**)[25,28,29], roquefortine C (**11**)/meleagrin (**12**)[30], gliotoxin (**13**)[31], fumitremorgin (**14**)[32], acetylaszonalenin (**15**)[33], and ditryptophenalaine (**16**)[34] (Figure 3) are products of NRPS pathways, and the core CDPs are catalyzed by bimodular NRPSs.

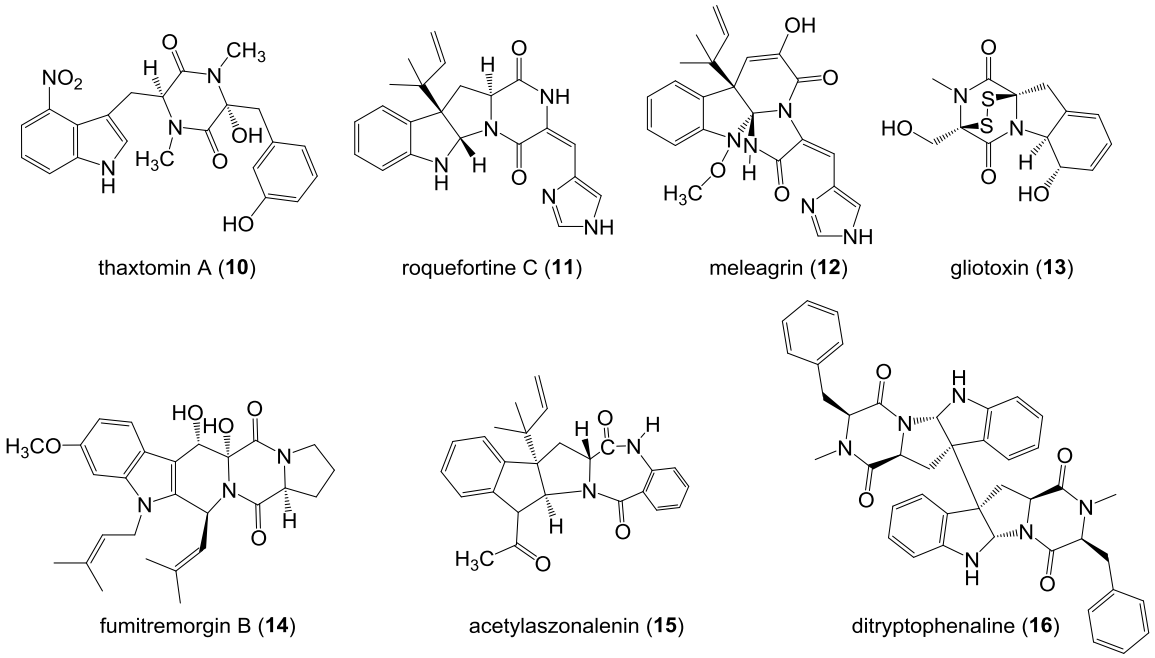


Figure 3. Selected examples of CDP-containing bioactive compounds generated by NRPS pathways.

1.2.2 CDPs from CDPS pathways

In addition to NRPSs, the characterization of AlbC in 2002, an NRPS-unrelated enzyme able to specifically form CDPs by hijacking aminoacyl-tRNAs (aa-tRNA) as substrates in the albonoursin pathway in *Streptomyces noursei*, revealed a new enzyme family

named cyclodipeptide synthases (CDPSs), which started a new era in the investigation on biosynthesis of cyclodipeptides[35,36]. Subsequently, several CDPSs from different organisms were identified to form varieties of CDPs[37]. Benefiting from the steadily increasing number of complete or partial genome sequences available since 2005[38], more than one hundred CDPSs have been experimentally characterized so far, incorporating 18 of the 20 proteinogenic amino acids into cyclodipeptides, with the exception for Asp and Lys[37,39-42]. In contrast to the large NRPSs, CDPSs are small peptide bond–forming enzymes with typical polypeptides of 200–300 amino acid residues[39]. They were found to be mostly distributed in the Actinobacteria, Proteobacteria and Firmicutes, with only a few were identified in eukaryotes[37,39,43].

To investigate the catalytic mechanism of CDPSs, crystal structures of three CDPSs, AlbC from *S. noursei*, Rv2275 from *Mycobacterium tuberculosis* and YvmC from *Bacillus licheniformis*, have been solved[44-47]. It was proposed that CDPSs used a ping-pong mechanism to catalyze the formation of CDPs by hijacking aa-tRNAs from primary metabolism as substrates (Figure 4)[37]. The reaction begins with binding of the first aa-tRNA, followed by the transfer of its aminoacyl moiety onto the conserved S37 (AlbC numbering) of the enzyme pocket. The resulted aminoacyl enzyme intermediate subsequently reacts with the aminoacyl moiety of the second aa-tRNA leading to formation of the dipeptidyl enzyme intermediate. Then the second peptide bond is formed by intramolecular cyclization thus generating the final CDP[37].

During the formation of CDP, the catalytic residues S37, N40, Y178, E182, Y202 and H203 play important roles in catalysis *via* hydrogen bond formation or deprotonation of ammonium group for nucleophilic attack[47]. Based on the distinguished conserved residues at position 40 and 203 and the highly conserved Y202 within the active sites, CDPSs are classified into three subfamilies, NYH, SYQ and XYP subfamiliy[39]. CDPSs can also incorporate non-canonical amino acids (ncAAs) into CDPs by loading the ncAAs to tRNAs using the promiscuity of aminoacyl-tRNA synthetases, which breaks the limitation of potential precursors to the 20 proteinogenic amino acids and diversifies CDPS products significantly[48]. Until now, several bioactive compounds derived from the CDPS pathways have been described, for instance, albonoursin (**17**)[35,36], cWW-Me_2

(**18**)[49], pulcherriminic acid (**19**)[45,50], mycocyclosin (**20**)[51,52], nocazine E (**21**)/neihumicin (**22**)[53], bicyclomycin (**23**)[54,55], and dehydrophenylahistin (**24**)[56] (Figure 5).

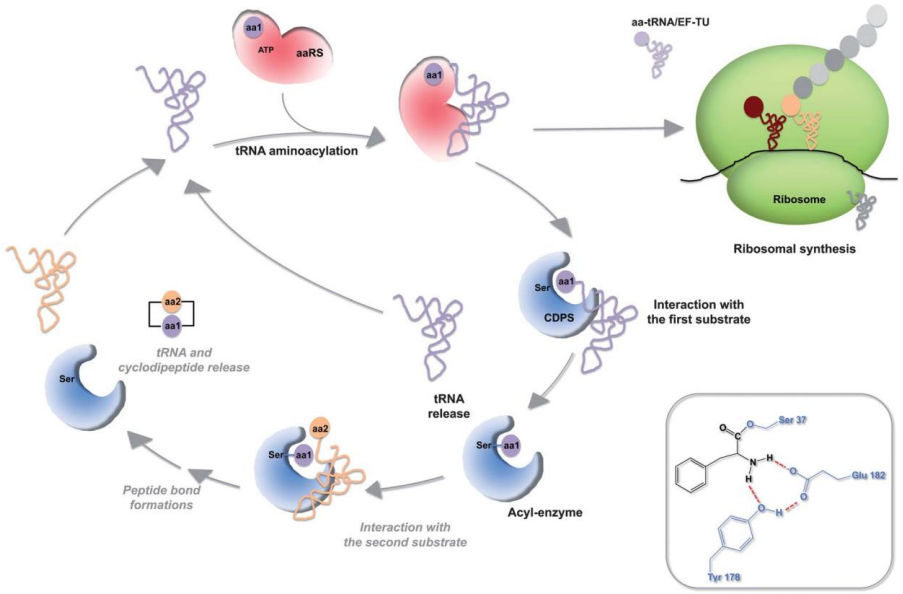


Figure 4. Proposed mechanism for cyclodipeptide biosynthesis by CDPS[37].

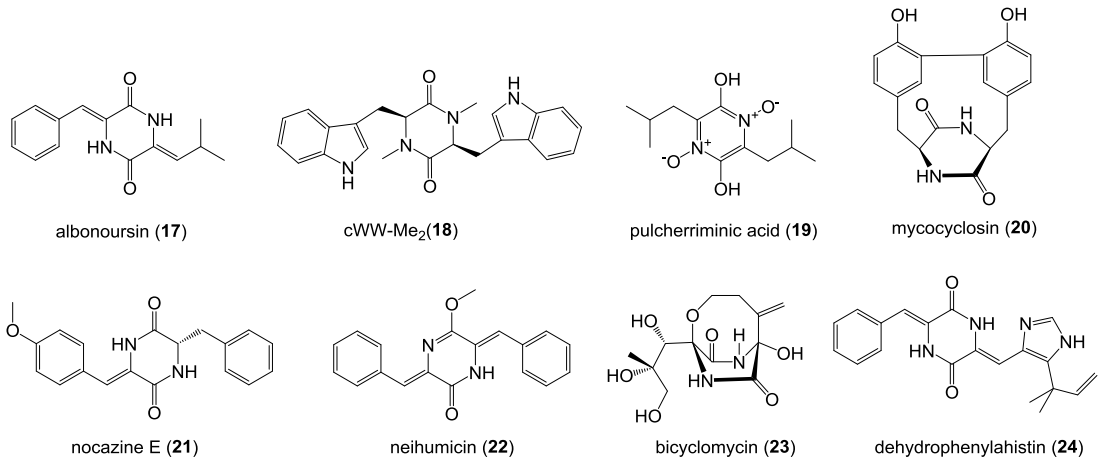


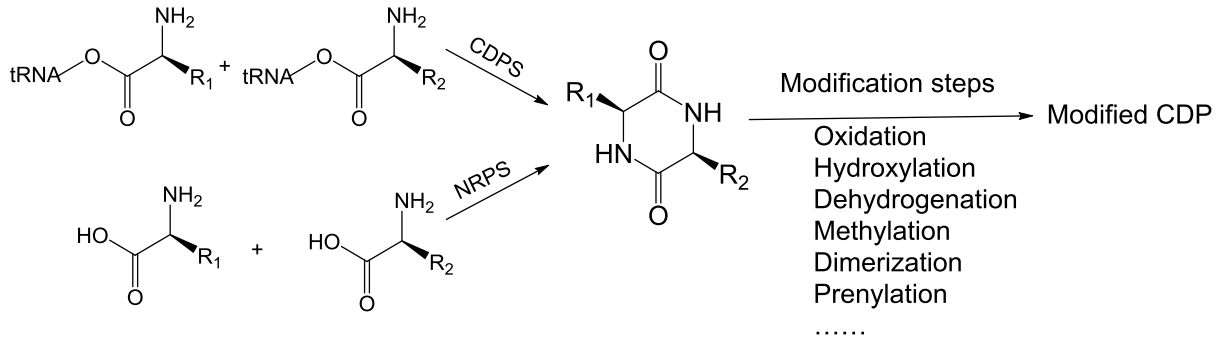
Figure 5. Selected examples of CDP-containing bioactive compounds generated by CDPS pathways.

1.2.3 CDP tailoring enzymes

In the CDP-containing natural product biosynthesis, the initially assembled CDPs by CDPSs or NRPSs are usually further modified by the associated tailoring enzymes

(Scheme 1). In the CDPS pathways, a variety of putative enzyme classes can be found genetically associated with respective CDPS gene. These include numerous oxidative enzymes of different types of cytochromes P450, α-ketoglutarate/Fe^{II}-dependent dioxygenases (α-KGDs) and flavin-containing monooxygenases. A large number of different *O*-, *N*-, and *C*-methyltransferases, α/β-hydrolases, peptide ligases, and acyl-CoA transferases also belong to the CDP tailoring enzymes[57]. Among the modifying enzymes, cyclic dipeptide oxidases (CDOs), a family of novel flavin-dependent amino acyl α,β-dehydrogenases, are widely distributed in the CDPS pathways and show extensive substrate specificity[35,36,53]. NRPS-dependent pathways contain similar types of modification enzymes, with the distinguishing feature of wide presence of a range of prenyltransferases catalyzing regular or reverse prenylation at different positions of the CDP scaffold[58,59]. The tailoring enzymes involved in the biosynthesis of selected CDP-containing natural products from NRPS and CDPS pathways are summarized in Table 1 and Table 2, respectively.

By using synthetic biology, combinatorial expression of tailoring enzymes originating from unrelated biosynthetic pathways provides the possibility to produce newly modified CDPs with diverse structures. For example, dehydrophenylahistin (**24**) is transformed by CDO-containing cell-free extract of *Streptomyces albulus* KO-23 from the fungal metabolite phenylahistin (**3**), resulting in more than 1000 times higher cytotoxicity than phenylahistin[56]. Among the CDP modifying enzymes, cytochromes P450 play important roles in natural product biosynthesis *via* diverse oxidative reactions.



Scheme 1. Different kinds of modifications of CDP generated by CDPS or NRPS.

Table 1. Overview of different tailoring enzymes associated within gene clusters of the NRPS pathways.

Compounds (Producer)	**CDP Core**	**Tailoring Enzymes (Function)**
thaxtomin A (**10**) (*Streptomyces scabies*)	cWF	cytochrome P450 (hydroxylation)
roquefortine C(**11**)/ meleagrin (**12**) (*Penicillium chrysogenum*)	cWH	prenyltransferase (*C*3-reverse prenylation) methyltransferase (*O*-methylation) cytochrome P450 (oxidative cyclization) N-hydroxylase (*N*-hydroxylation) oxidoreductase (α,β-dehydrogenatioin)
gliotoxin (**13**) (*Aspergillus fumigatus*)	cFS	two cytochromes P450 (oxidation) glutathione S-transferase (sulfurization) aminocyclopropane carboxylic acid synthase (cyclopropane-formation) two methyltransferase (*O*- /*N*-methylation)
fumitremorgin B (**14**) (*Aspergillus fumigatus, Neosartorya fischeri*)	cWP	two prenyltransferases (*N*- /*C*2-prenylation) three cytochrome P450 (oxidative cyclization, hydroxylation) methyltransferase (*O*-methylation)
acetylaszonalenin (**15**) (*Neosartorya fischeri*)	cWF	prenyltransferase (*C*3-reverse prenylation) acetyltransferase (*N*-acetylation)
ditryptophenalaine (**16**) (*Aspergillus fumigatus*)	cWF	cytochrome P450 (dimerization)

Table 2. Overview of different tailoring enzymes associated within gene clusters of the CDPS pathways.

Compounds	CDP Core	Tailoring Enzymes (Function)
albonoursin (**17**) (*Streptomyces noursei*)	cFL	CDO (α,β-dehydrogenation)
cWW-Me_2 (**18**) (*Actinosynnema mirum*)	cWW	methyltransferase (*N*-methylation)
pulcherriminic acid (**19**) (*Bacillus subtilis*)	cLL	cytochrome P450 (*N*-oxide formation and aromatization of the DKP-ring)
mycocyclosin (**20**) (*Mycobacterium tuberculosis*)	cYY	cytochrome P450 (intramolecular C─C coupling)
nocazine E (**21**)/ neihumicin (**22**) (*Nocardiopsis dassonvillei*)	cFY cFF	CDO (α,β-dehydrogenation) methyltransferase (*O*-methylation)
bicyclomycin (**23**) (*Streptomyces sapporonensis*)	cIL	α-KGD (epoxidation) three α-KGDs (hydroxylation) α-KGD (dehydrogenation) cytochrome P450 (hydroxylation)
dehydrophenylahistin (**24**)	cFH	CDO (α,β-dehydrogenation)

1.3 Cytochromes P450

Cytochromes P450 (P450s) constitute a large superfamily of heme-containing monooxygenases distributed in all life kingdoms. They catalyze diverse oxidative reactions with a wide range of substrates[60]. The terminology P450 is based on their typically maximum absorbance at 450 nm in reduced form when bound to carbon monoxide[61]. P450s are the most versatile biological catalysts in nature for their contribution to biosynthesis of steroids, fatty acids, fat-soluble vitamins and bile acids. They are also involved in biotransfomation of xenobiotics, metabolism of drugs, as well as conversion of alkanes, terpenes and aromatic compounds[62]. With extraordinarily broad substrate spectra and regio- and stereoselectivity, they play important roles in the biosynthesis of natural products[63,64].

1.3.1 Catalytic mechanism of P450s

Since the determination of the first crystal structure of $P450_{cam}$, a monooxygenase metabolizing camphor in *Pseudomonas putida*, in the mid-1980s[65,66], more than 800 entries with “P450” can be found in the Protein Data Bank (PDB) to date. Many of them are involved in natural product biosynthesis[67]. In all members of P450 family, the only absolutely conserved and catalytically most important residue is cysteine (Cys), which is situated at the N-terminus of the L-helix and acts as the fifth axial ligand of the heme iron coordinating to the Fe atom *via* a thiolate bond[63]. The vast majority of P450s catalyze the chemical reactions depending on the transfer of electrons to the heme iron *via* an electron transport chain. Based on the components of redox partners, different classes of electron transfer chains were described by Hannemann et al.[62] Most bacterial P450s belong to class I redox system consisting of three separate and soluble proteins, *i.e.* ferredoxin, ferredoxin reductase and P450[62]. Usually, bacterial P450s, especially those from *Streptomyces*, also accept electrons from heterologous redox partners[64]. In the P450 catalytic cycle, the electrons are transferred from NAD(P)H to iron of the P450, resulting in the incorporation of one oxygen atom into the substrate *via* complex multiple steps (Figure 6). The catalytic cycle begins with the binding of substrate to displace the iron coordinated water. The ferric heme iron is reduced by the first electron. Molecular oxygen binding is followed by the second reduction with electron and two successive

protonations of the distal oxygen, leading to the formation of the intermediate Compound 0 (Cpd 0, Fe^{III}-OOH) and reactive ferryl-oxo intermediate Compound I (Cpd I, Fe^{IV}=O, porphyrin π-cation radical). The substrate radical is formed through abstraction of one hydrogen atom by Cpd I, and followed by the recombination with hydroxyl radical from ferryl-hydroxo Compound II (Cpd II) to give the hydroxylated product. Finally, the hydroxylated product is released from the active site and water rebinds to restore the enzyme to its resting state.

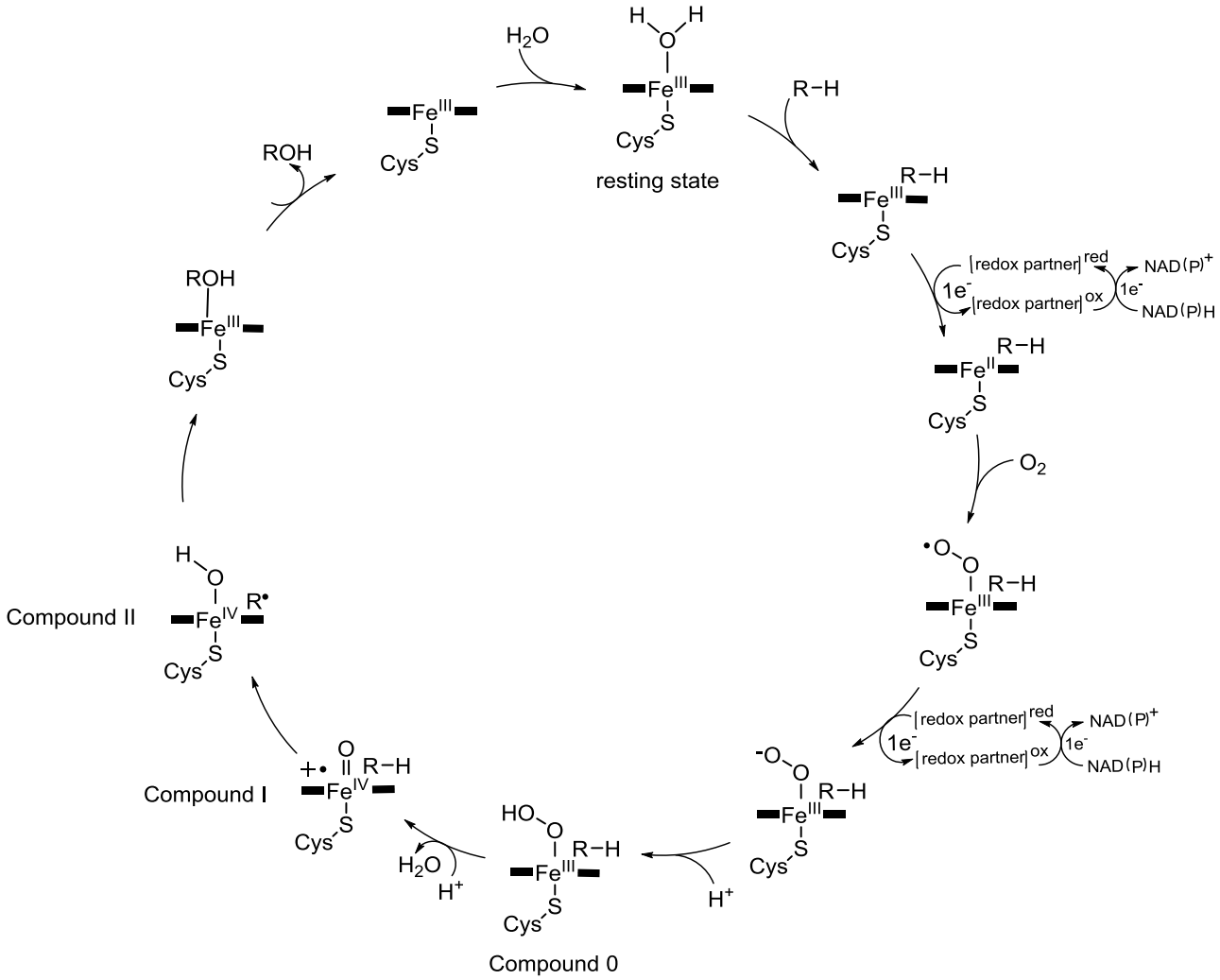
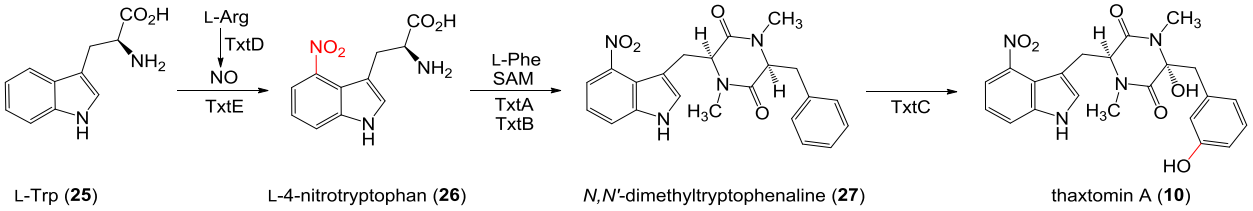


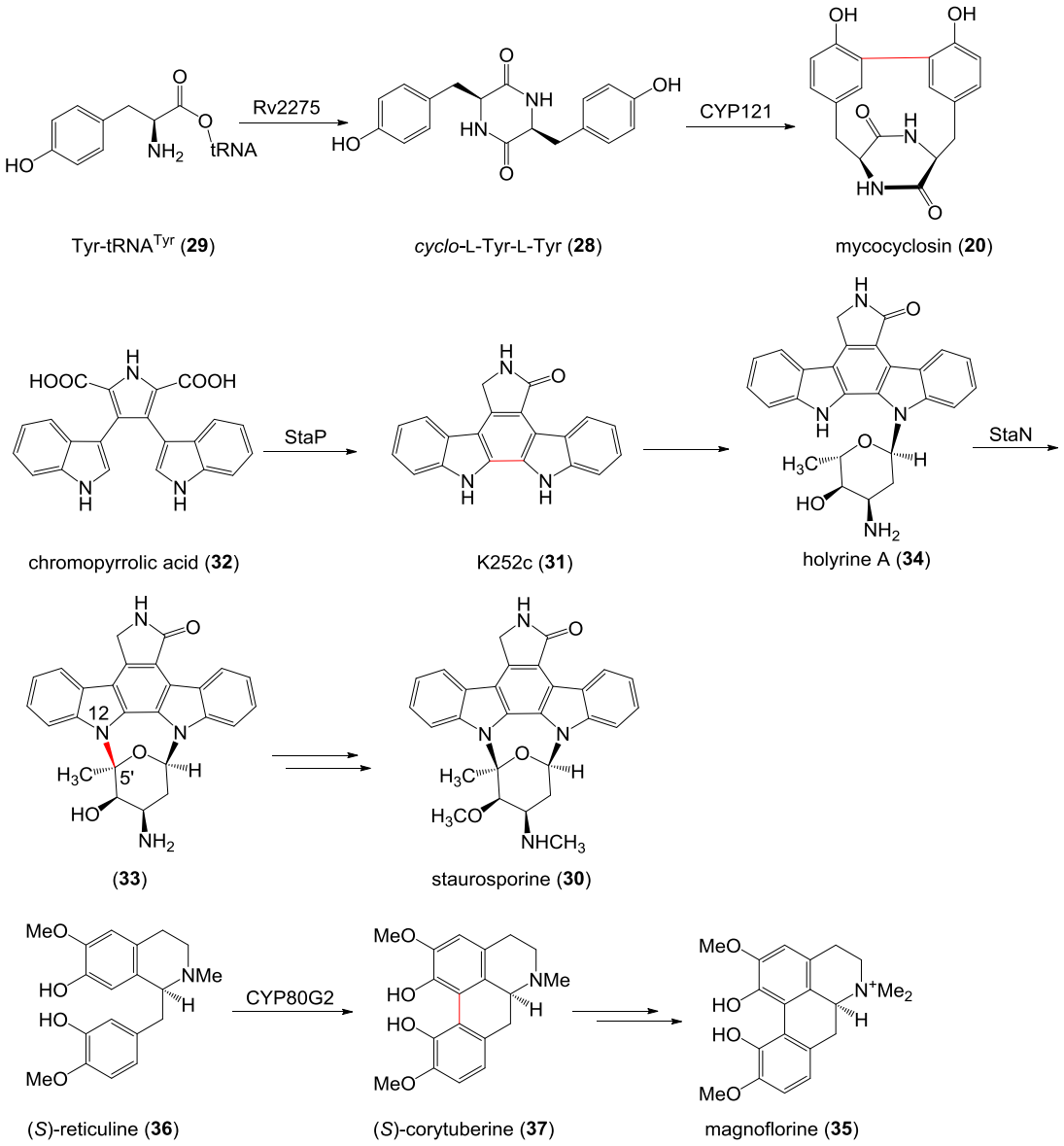
Figure 6. P450 catalytic cycle with hydroxylation of the substrate RH to ROH as an example.

1.3.2 Aryl–aryl coupling catalyzed by P450s in natural product biosynthesis

As ubiquitous catalysts in nature, P450s catalyze diverse reactions in natural product biosynthesis. Among them, hydroxylation of C–H or N–H bond and epoxidation of C=C bond are regarded as common P450 reactions, leading to only slight changes in the substrate skeleton[68]. These typical P450s often exhibit multifunction and catalyze consecutive hydroxylation/epoxidation during the biosynthesis of interesting molecules, such as the MycG involved in the formation of macrolide mycinamicin I, II, V from *Micromonospora griseorubida*[69,70] and TamI from *Streptomyces* sp. 307-9 catalyzing the production of diverse tirandamycins[71,72]. In addition to the common P450 reactions, unusual reactions, such as aromatic/phenolic coupling, bond cleavage or migration and ring opening or rearrangement, have gained more and more attention in recent years. One well-known example is the P450 TxtE in the biosynthesis of phytotoxin thaxtomin A (**10**), which catalyzes the unusual nitration of L-tryptophan (**25**) resulting in the formation of L-4-nitrotryptophanthen (**26**). The P450 enzyme TxtC performs the typical phenylalanyl hydroxylation after the CDP assembly (**27**)[25,28,29] (Scheme 2.). Additionally, intra-/intermolecular aryl–aryl coupling reactions have been found in many natural product biosynthesis, as illustrated in Scheme 3 and Scheme 4, respectively. For example, the CDP tailoring P450 CYP121 from *Mycobacterium tuberculosis* catalyzes the intermolecular C–C coupling of the tyrosine side chain of *cyclo*- L-Tyr-L-Tyr (**28**) to form mycocyclosin (**20**), which acts as a causative agent of tuberculosis[52]. In the biosynthesis of staurosporine (**30**) in *Streptomyces* sp. TP-A0274, two intermediates are formed by P450-catalyzed intramolecular C–C and C–N coupling. The indolocarbazole skeleton K252c (**31**) is formed by intramolecular C–C bond connection of chromopyrrolic acid (**32**) by StaP[73]. StaN catalyzes the C–N bond (C5'-N12) formation of the intermediate (**33**) from holyrine A (**34**)[74]. A similar aromatic coupling reaction also occurs during the biosynthesis of magnoflorine (**35**) in the plant *Coptis japonica*. The transformation of (*S*)-reticuline (**36**) to (*S*)-corytuberine (**37**) is catalyzed by CYP80G2[75].



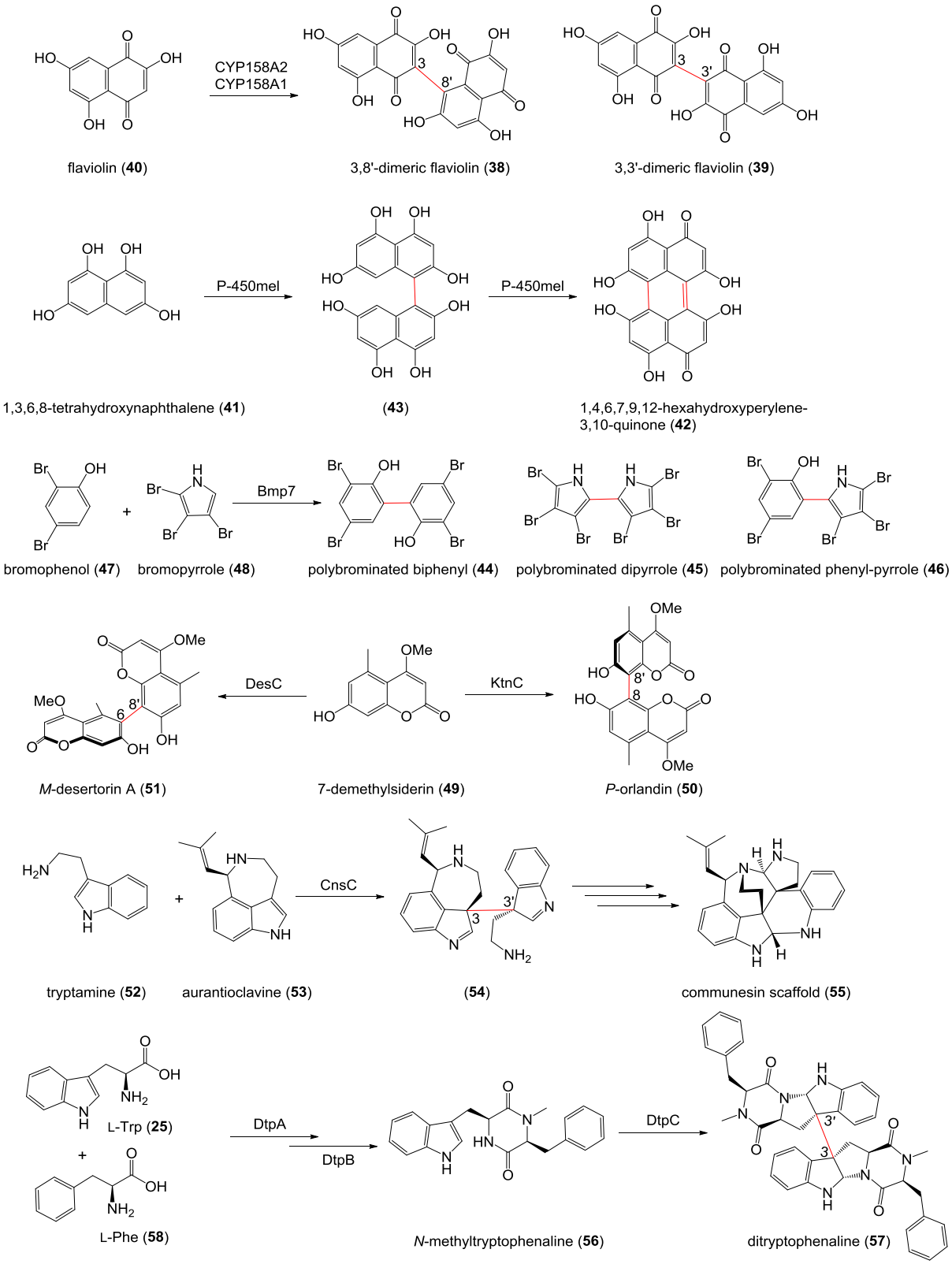
Scheme 2. Biosynthetic pathway of thaxtomin A.



Scheme 3. Examples of intramolecular aryl–aryl coupling reactions catalyzed by P450s.

The dimeric compounds often show useful and interesting biological activities. One such example are the pigments dimeric flaviolins (**38** and **39**) formed *via* intermolecular C–C

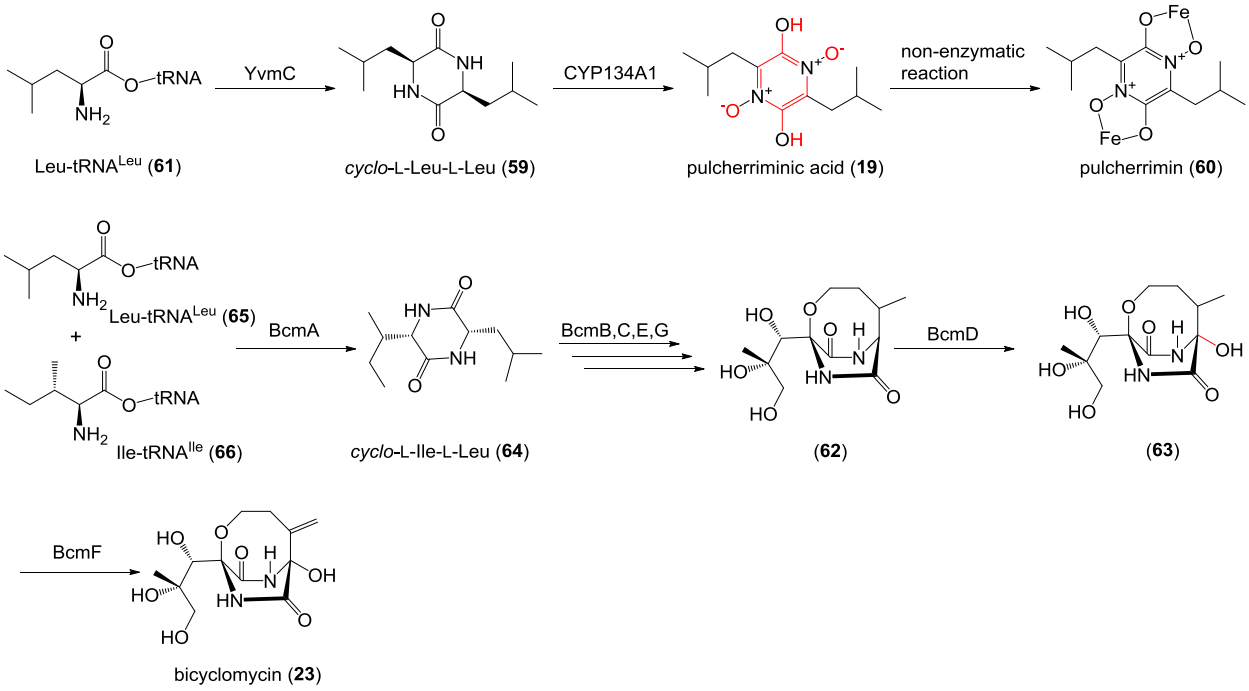
coupling of flaviolin (**40**) by two P450s (CYP158A1 and CYP158A2) in *S. coelicolor* A3 (2). These pigments are thought to protect microbes from UV radiation[76,77]. Similarly, intermolecular dimerization of 1,3,6,8-tetrahydroxynaphthalene (**41**) to generate 1,4,6,7,9,12-hexahydroxyperylene-3,10-quinone (**42**) *via* intermediate (**43**) was achieved by P-450mel in *Streptomyces griseus*[78]. In marine bacteria *Pseudoalteromonas luteoviolacea* 2ta16 and *P. phenolica* O-BC30, the P450 Bmp7 was found to catalyze the formation of polybrominated aromatic compounds (**44−46**) *via* homo or heterodimerization of their monomers (**47**, **48**)[79]. In addition to bacteria, P450s from fungi also catalyze intermolecular aryl–aryl coupling. Two homologous fungal P450 enzymes, KtnC from *A. niger* and DesC from *Emericella desertorum*, were identified to mediate aryl–aryl coupling of 7-demethylsiderin (**49**) at different positions to form regioisomers *P*-orlandin (**50**) and *M*-desertorin A (**51**), respectively[80]. In the biosynthesis of communesins in *Penicillium expansum*, the P450 enzyme CnsC is responsible for heterodimeric coupling of two different indole moieties, tryptamine (**52**) and aurantioclavine (**53**), resulting in the formation of heterodimer (**54**), a precursor of communesin scaffold (**55**)[81]. The CDP tailoring P450 in the NRPS pathway from *A. flavus*, DtpC, catalyzes the homodimerization of *N*-methyltryptophenaline (**56**) to form ditryptophenaline (**57**)[34]. The enzymes Bmp7 and CnsC represent rare examples of P450s that are able to perform aryl–aryl coupling of two different substrates.



Scheme 4. Examples of intermolecular aryl–aryl coupling reactions catalyzed by P450s.

1.3.3 P450s involved in the CDPS pathways

Although several gene clusters of natural products derived from CDPS pathways have been identified (Table 2), only three associated P450s were functionally characterized: i) the CYP121 from *M. tuberculosis*[52] described in Scheme 3; ii) CYP134A1 from *B. subtilis* catalyzing double N-oxide formation and aromatization of the DKP-ring of *cyclo*-L-Leu-L-Leu (**59**) leading to the formation of pulcherriminic acid (**19**). Further transformation to pulcherrimin pigment (**60**) happens *via* a non-enzymatic reaction[50]; iii) P450 BcmD, performing the hydroxylation of intermediate **62** in late modification step during the biosynthesis of bicyclomycin (**23**) in *Streptomyces sapporonensis*, leading to the formation of intermediate (**63**), which is further modified by α-KGD BcmF[55]. The reactions catalyzed by the latter two P450s are illustrated in Scheme 5.



Scheme 5. P450s involved in the CDPS pathways.

1.4 Prenylated aromatic compounds in nature

As mentioned under **1.2.3**, prenyltransferases (PTs) participate in many prenylated secondary metabolite biosynthesis and catalyze regular (**67**) or reverse prenylation (**68**) at different positions of the CDP scaffold derived from the NRPS pathway. Besides the CDP core, PTs also catalyze the attachment of one or more prenyl moieties (**69**) to other aromatic cores in natural product biosynthesis, such as indole (**70**), tyrosine (**71**), xanthone (**72**), flavone (**73**), and hydroxynaphthalene (**74**)[82-85] (Figure 7). PTs are found in many living organisms including plants, bacteria, and fungi and can be classified in different subgroups based on their amino acid sequences structures, and biochemical properties[86-88]. The resulted prenylated products often show biological and pharmacological activities distinct from their non-prenylated precursors, which makes these enzymes to be attractive biocatalysts for structure modifications[58,89].

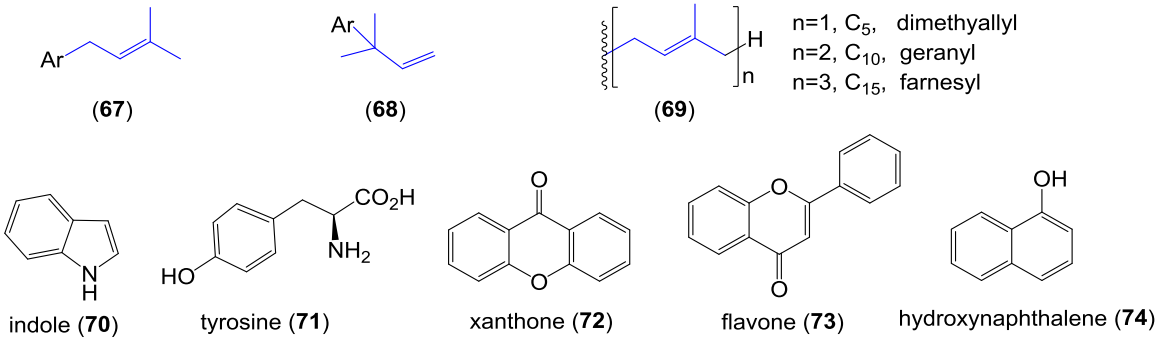


Figure 7. Regular and reverse prenyl pattern, different lengths of prenyl moieties andaromatic scalfolds as PT acceptors.

1.4.1 Prenylated tyrosine derivatives

Tyrosine is one of the 20 proteinogenic amino acids that are used by cells to synthesize proteins in nature. It is produced *via* shikimate pathway in plants and most microorganisms and from the essential amino acid phenylalanine (derived from food) in mammals[90]. Benefiting from the *para*-hydroxyl group, tyrosine is readily to take chemical reactions and signal transductions as well as the post-translational modifications such as phosphorylation, sulfation, nitration, oxidation, halogenation,

glycosylation, AMPylation or cross-linking[90]. Several simple prenylated tyrosine derivatives were found to be *O*- or *C3*-prenylated, such as the 5-HT receptor ligand (**75**) isolated from *Pithomyces Ellis*[91], stachyline A (**76**) found in *Stachylidium* sp. harboring a unique terminal oxime group[92] and a *C3*-prenylated tyrosine (**77**) isolated from *Streptomyces* sp. IFM 10937[93]. Until now, two tyrosine PTs have been identified and characterized biochemically. The first one, SirD, catalyzes the *O*-prenylation of L-tyrosine (**78**) resulting in **79**, which is the precursor of the phytotoxin sirodesmin PL (**80**) in *Leptosphaeria maculans*[94,95]. Another tyrosine PT, TyrPT, from an unknown biosynthetic pathway in *A. niger* shows similar biochemical features as SirD (Figure 8)[96].

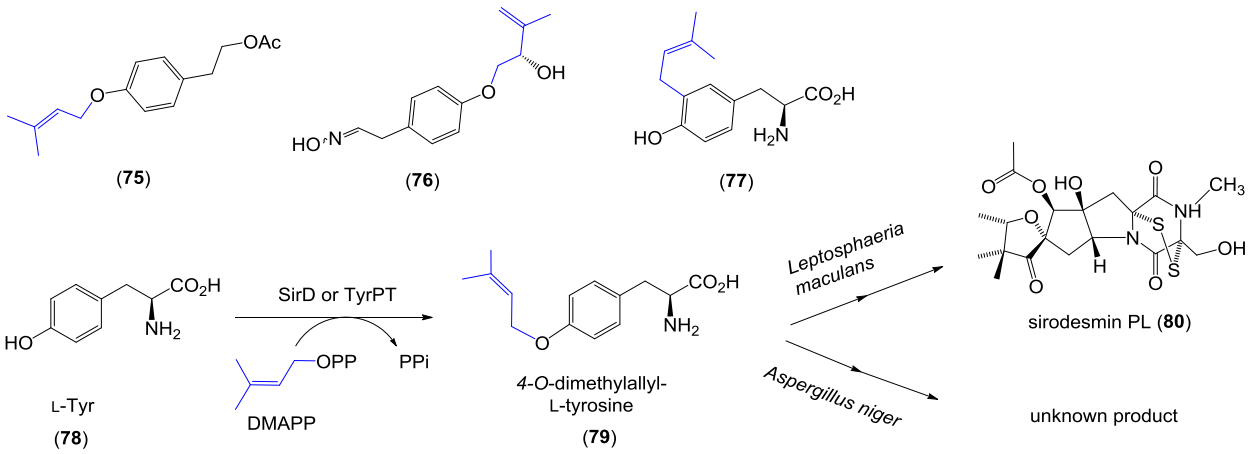


Figure 8. Examples of simple prenylated tyrosine derivatives and the function of the two identified tyrosine PTs.

1.4.2 Chemoenzymatic synthesis of prenylated aromatic compounds with DMAPP

As described previously, many prenylated secondary metabolites exhibit different and efficient activities comparing to their non-prenylated precursors[58,97]. Therefore, the diversity of prenylated compounds has gained increasing attention in pharmaceutical study. With the advantage of broad aromatic substrate spectra, high regio- and stereoselectivity of the reactions and mild reaction conditions in comparison to the chemical synthesis, chemoenzymatic synthesis of unnatural prenylated compounds is of interest for current research. Such compounds can be obtained by using aromatic PTs

with diverse substrates. For example, the regiospecific regular *C*4-, *C*5- and *C*7-prenylated simple indole derivatives (**81−86** in Figure 9) were generated *via* chemoenzymatic synthesis by the respective L-tryptophan PTs FgaPT2, 5-DMATS and 7-DMATS in the presence of DMAPP[59,98,99]. In addition to tryptophan derivatives, the indole PTs also use other aromatic substrates with distinct structures such as hydroxynaphthalenes, flavonoids, or xanthones[100-102]. FgaPT2 and 7-DMATS have been demonstrated to accept tyrosine and its derivatives as substrates besides the simple indole derivatives[96,103]. Similarly, the tyrosine PTs SirD and TyrPT catalyze both the *O*- or *N*-prenylation of a variety of tyrosine derivatives and the prenylation of tryptophan and derivatives[95,96,104,105].

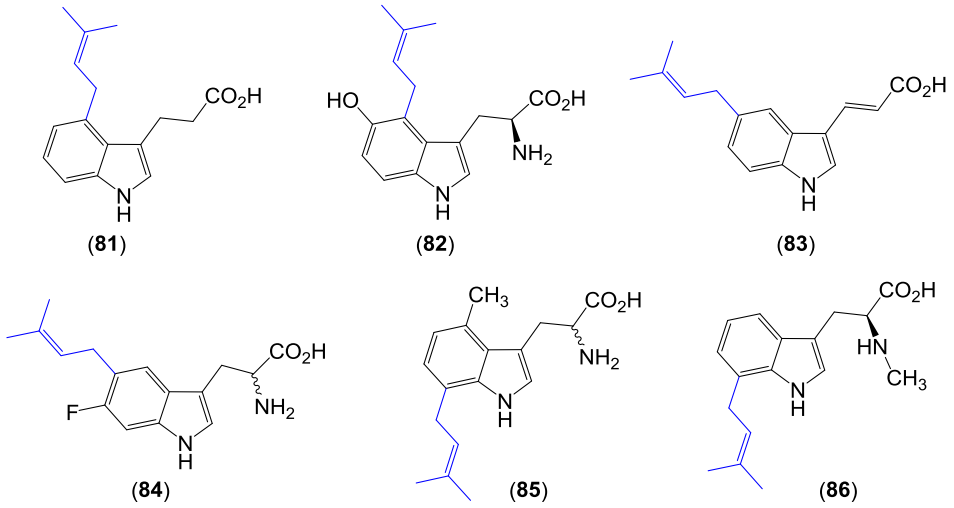


Figure 9. Examples of prenylated indole derivatives.

1.4.3 Chemoenzymatic synthesis of prenylated aromatic compounds with unnatural prenyl analogs

As mentioned above, the PTs of DMATS superfamily utilize a large number of aromatic substrates. In contrast to their flexible aromatic substrate spectra, most of the DMATS members show a restricted specificity toward their prenyl donors, accepting usually DMAPP as prenyl donor[88], with a few exceptions that use GPP, or both DMAPP and GPP, or DMAPP, GPP, and FPP as prenyl donors[102,106-108]. To broaden their specificity

for alkyl donors, two tryptophan PTs FgaPT2 and 5-DMATS and five cyclic dipeptide prenyltransferases BrePT, FtmPT1, AnaPT, CdpNPT, and CdpC3PT have been demonstrated prior to this study to accept the unnatural prenyl analogs methylallyl diphosphate (MAPP) (**87**), 2-pentenyl diphosphate (2-pen-PP) (**88**), and benzyl diphosphate (benzyl-PP) (**89**), with partial or complete shifts of alkylation or benzylation positions (**90**-**92**) comparing to the natural prenyl donor (Figure 10)[109-111].

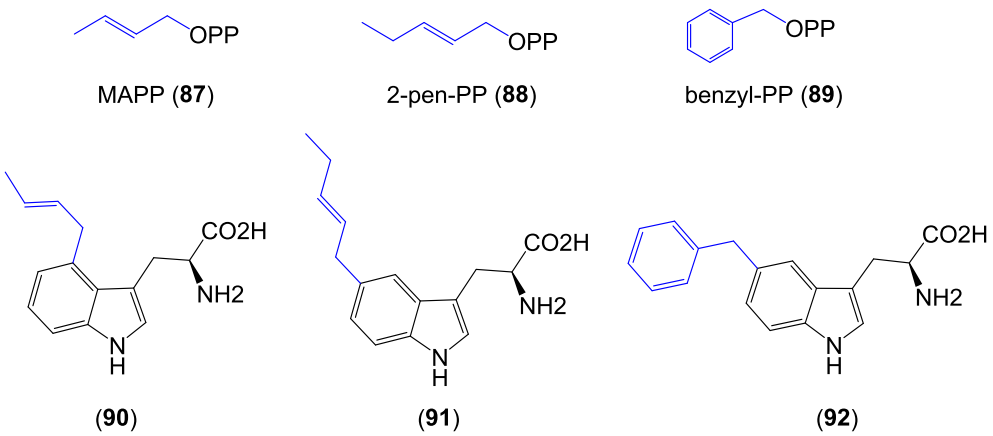


Figure 10. Examples of unnatural prenyl analogs and alkylated or benzylated L-tryptophan derivatives.

2 Aims of this thesis

The following issues have been addressed in this thesis:

1) Identification of CDPSs that catalyze the formation of tryptophan-containing cyclodipeptides from *Streptomyces*

CDPSs catalyze the formation of CDPs by using the loaded tRNAs as substrates. Tryptophan-containing CDPs have received an increasing interest in recent years due to their promising scaffolds for structural modification. Although over ten specific cWW synthases have been functionally characterized, only one CDPS originating from the sea anemone *Nematostella vectensis* was demonstrated to produce several cWXs prior to this study. The product yields are however very low with only 0.8–1.0 mg/L in *E. coli*. The aim of this project is to find new CDPSs for the formation of CDPs comprising tryptophan and another amino acid. The following experiments were carried out:

- Search for CDPS homologues in database by blast with the CDPSs that form cWW
- Phylogenetic analysis of the CDPS candidates with the known CDPSs
- Cloning of the CDPS genes into pET28a vector and expression in *E. coli*
- LC-MS analysis of the *E. coli* cultures harboring the CDPS expressing constructs
- Isolation and structure elucidation of CDPS products by HR-EIMS and NMR analyses
- Determination of the CDPS product yields on HPLC

2) Functional proof of a cytochrome P450 homologue from *Streptomyces purpureus* NRRL B-5737

In nature, the CDP cores synthesized by NRPSs or CDPSs are usually further modified by diverse tailoring enzymes to form sophisticated compounds. In the first project, several CDPSs that catalyze the formation of cWXs have been identified in different *Streptomyces* strains, including cWW synthases from *S. pupureus*, *S. lavendulae*, and

S. xanthophaeus. In these three strains, the cWW-forming CDPS coding gene is genetically associated with a gene encoding putative cytochrome P450 in the downstream with the same orientation. The three P450 homologues share sequence identities of more than 75% on the amino acid level with each other. Therefore it could be speculated that these three clusters might synthesize similar CDP derivatives. The aim of this project is to identify the function of the three tailoring P450 homologues by taking $P450_{NB5737}$ as a representative. The following experiments were carried out:

- Phylogenetic analysis of the three P450 homologues with structurally defined natural product P450s from bacteria
- Phylogenetic analysis of P450s from this study and functionally characterized P450s from bacteria
- Cloning of $CDPS_{NB5737}$ alone or together with $P450_{NB5737}$ into integrative plasmid pSET152 or replicative vector pPWW50 and expression in *S. coelicolor* M1146
- Confirmation of the cWW metabolism by $P450_{NB5737}$ *via* biotransformation experiment
- Isolation of the $P450_{NB5737}$ product and structure elucidation by HR-EIMS and NMR analyses
- Production of the recombinant $P450_{NB5737}$ in *E. coli* and performing *in vitro* assays of $P450_{NB5737}$
- Postulation of the catalytic mechanism of $P450_{NB5737}$

3) Biochemical investigations on the acceptance of unnatural DMAPP analogs by tyrosine *O*-prenyltransferases TyrPT and SirD

Prenyltransferases of the DMATS superfamily show broad substrate spectra, but strict prenyl donor specificity. The aim of this thesis was to elucidate the behavior of *O*-prenyltransferases TyrPT and SirD towards chemically synthesized DMAPP analogs, MAPP, 2-pen-PP and benzyl-PP. The following experiments were carried out:

- Overproduction of TyrPT and SirD in *E. coli* and subsequent purification

- Incubation of assays containing L-tyrosine and derivatives and the DMAPP analogs in the presence of TyrPT and SirD, respectivey
- Analysis of the enzyme activities on HPLC and isolation of the enzyme products
- Structure elucidation of the products by HR-EIMS and NMR analyses
- Determination of the kinetic parameters of the corresponding reactions

3 Results and discussion

3.1 Identification of nine tryptophan-containing cyclcodipeptide synthases from *Streptomyces* strains

Prior to this study, dozens of CDPSs that are responsible for cyclodipeptide formation have been identified[37,39-42]. Overview the CDPS products, 17 of 20 proteinogenic amino acids have been incorporated into cyclodipeptides[39]. Although over ten specific cWW synthases have been characterized[39,49,112], CDPSs synthesizing other tryptophan-containing cyclodipeptides have not been reported for microorganisms.

The working flow of this project is illustrated in Figure 11. Firstly, to get the CDPSs for tryptophan-containing cyclodipeptides, a BlastP search was carried out using sequences of known cWW-forming CDPSs. To this end, thirteen candidate CDPSs from eleven *Streptomyces* strains were selected as investigation objectives.

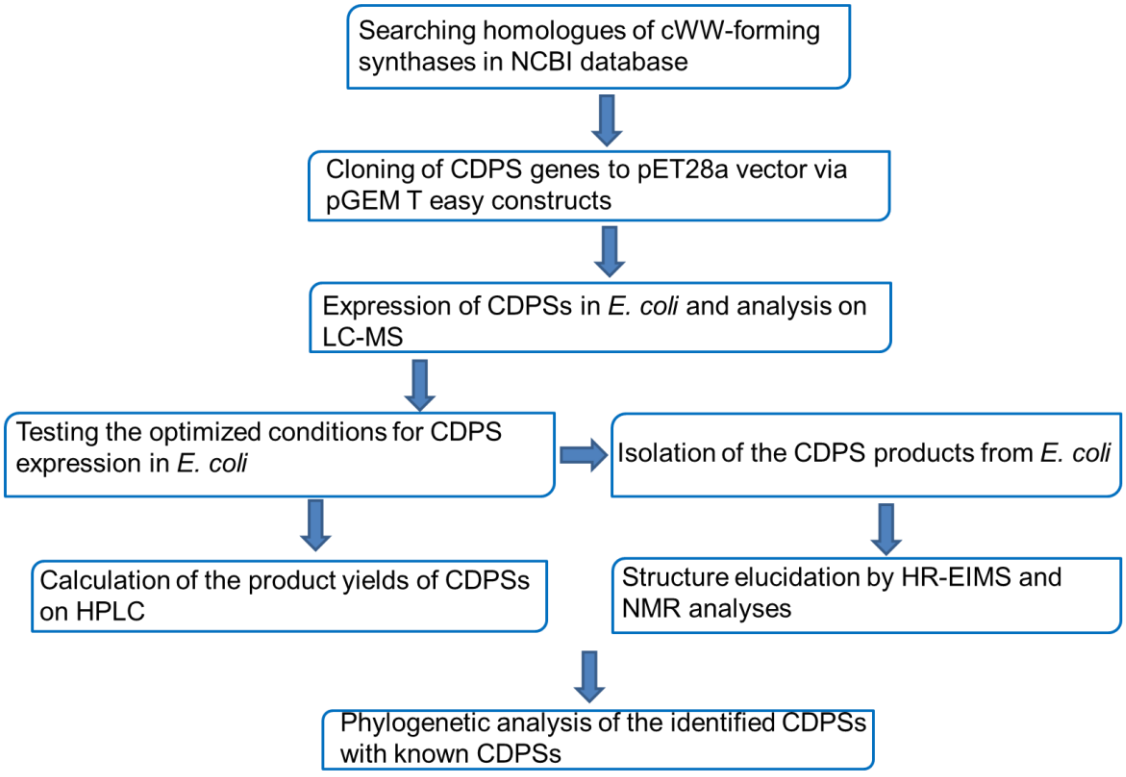


Figure 11. Working flow of identification of cWX-forming CDPSs

For functional proof, the coding genes of CDPSs were amplified by PCR with their respective gDNA as templates. Subsequently, genes encoding CDPSswere ligated into pET28a vector *via* pGEM T easy constructs and transformed in *E. coli* SoluBL21. The production of CDP was induced with appropriate concentration of IPTG when an absorption of OD_{600} at 0.6 was reached. The expression system was tested with two known cWW-forming CDPSs WP_014141671 and WP_014140974 from *Streptomyces cattleya* DSM 46488. The expected product cWW was confirmed by LC-MS and NMR analyses. The expression constructs of other eleven CDPSs were also expressed in *E. coli* SoluBL21. The optimized expression condition for each CDPS was tested with different IPTG concentrations (0.2 mM and 0.5 mM) and incubation temperatures (20 °C, 30 °C and 37 °C). The cultures of *E. coli* transformant harboring each CDPS expression construct were analyzed on LC-MS, with the empty vector pET28a as negative control. Then CDPSs products were isolated and sent for NMR analysis. Structure elucidation revealed nine cWX-forming CDPSs including one cWL, two cWP, and three cWW synthases as well as three unspecific CDPSs producing several products with cWA or cWY as the major cyclodipeptide (Figure 12). The other two CDPSs WP_031028810 from *Streptomyces* sp. NRRL F-5639 and BAU83478 from *Streptomyces laurentii* DSM 41684 were proven to catalyze the formation of cFL.

Total product yields of CDPs were calculated based on HPLC analysis to be 46 to 211 mg/L *E. coli* culture after cultivation for 36 h under their optimized conditions. Phylogenetic analysis of these CDPS homologues with the known CDPSs from actinomycetes revealed that tryptophan-containing cyclodipeptide synthases from actinobacteria are located in two clades: clade I comprising only cWX synthases with the exception for cLL and clade II including not only cWX synthases but also cFX and cYX synthases.

The results presented in this project expanded significantly the spectrum of tryptophan-containing cyclodipeptide (**93**-**100**) synthases (Scheme 6). And the high product yields of cWXs provided a solid basis for further combination of these CDPS genes with other modification genes in synthetic biology.

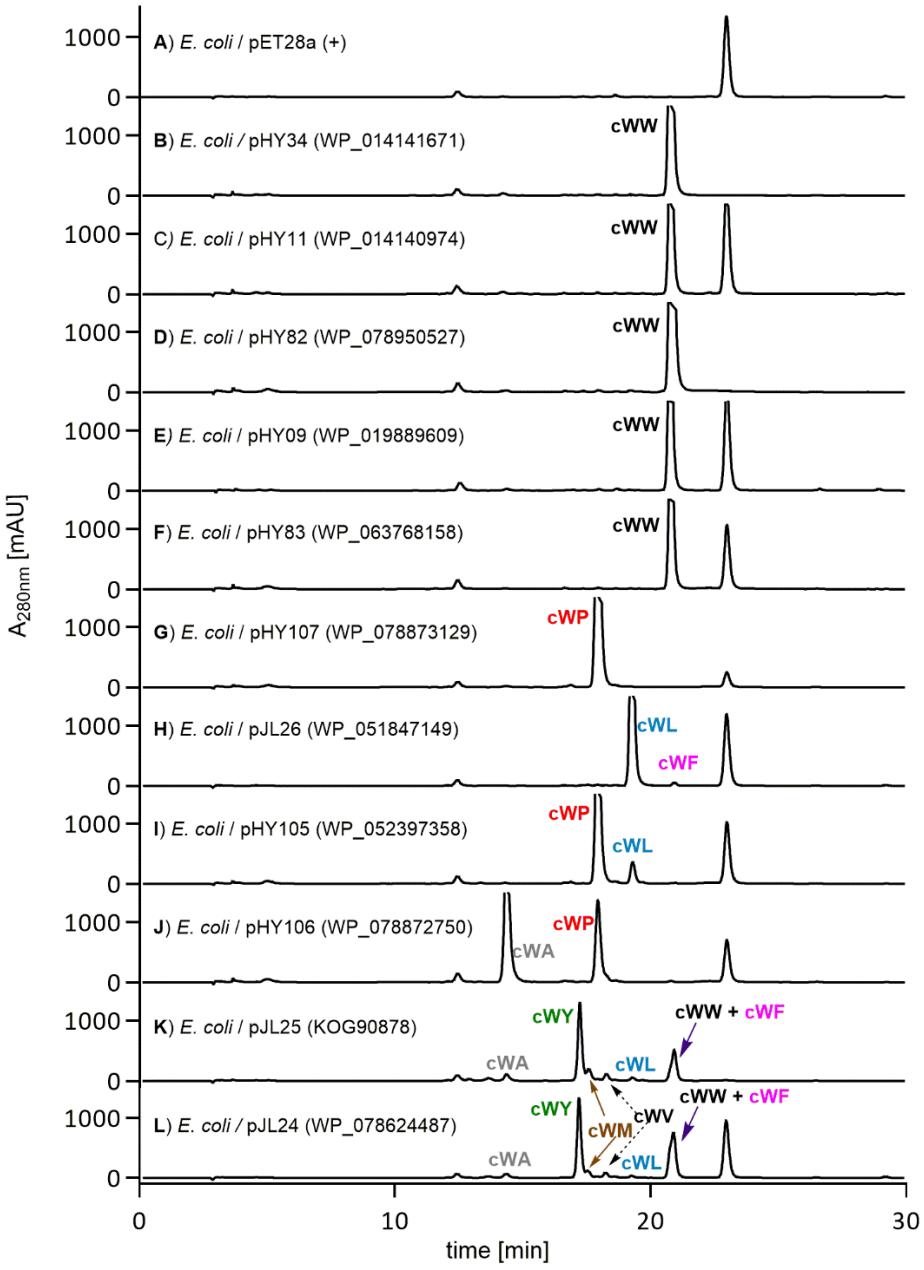
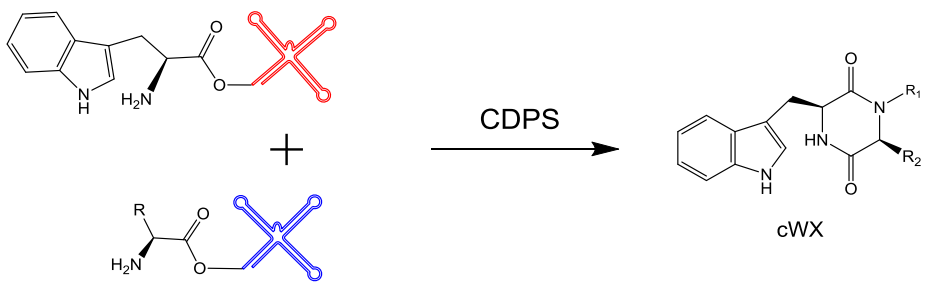


Figure 12. HPLC analysis of recombinant *E. coli* strains with empty vector (a) or with CDPS genes (b–l)[40].



(**93**) X= W, R1= H, R2= indolyl (**94**) X= F, R1=H, R2= phenyl (**95**) X= L, R1=H, R2= $-CH_2CH(CH_3)_2$

(**96**) X= V, R1=H, R2= $-CH(CH_3)_2$ (**97**) X= A, R1= H, R2= $-CH_3$ (**98**) X= P, R1-R2= $-CH_2CH_2CH_2-$

(**99**) X= M, R1= H, R2= $-CH_2CH_2SCH_3$ (**100**) X=Y, R1=H, R2= 4OH phenyl

Scheme 6. The CDPSs catalyzing the formation of cWX in this study.

This project was cooperated with Jing Liu. The expression constructs of CDPSs with names “pHY××” (Figure 12 B-G, I, J) were prepared by the author of this thesis.

For details on this work, please see the publication (section 4.1)

Liu, J.*, **Yu, H.***, & Li, S.-M. (2018). Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains. *Appl. Microbiol. Biotechnol*., *102* (10), 4435-4444. [*equal contribution]

Liu, J*., **Yu, H*.**, & Li, S.-M. (2018). Correction to: Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains. *Appl. Microbiol. Biotechnol.*, *102*, 5787-5789. [*equal contribution]

3.2 Coupling of *cyclo*-L-Trp-L-Trp and guanine by a P450 homologue from *Streptomyces purpureus*

Cytochrome P450s are common CDP tailoring enzymes that participate in the modification of CDP cores to form sophisticated natural products. However, only three P450s involved in CDPS pathways have been identified prior to this study. They are CYP121 in the biosynthesis of mycocyclosin from *M. tuberculosis*[52], CYP134A1 in the pulcherrimin pathway from *B. subtilis*[50], and the P450 BcmD in the bicyclomycin pathway[55]. Inspection of the CDPS genes identified in the first project revealed that the cWW-forming CDPS genes from *S. pupureus*, S*. lavendulae*, and *S. xanthophaeus* are respectively associated with a downstream putative cytochrome P450 gene with the same orientation (Figure 13).

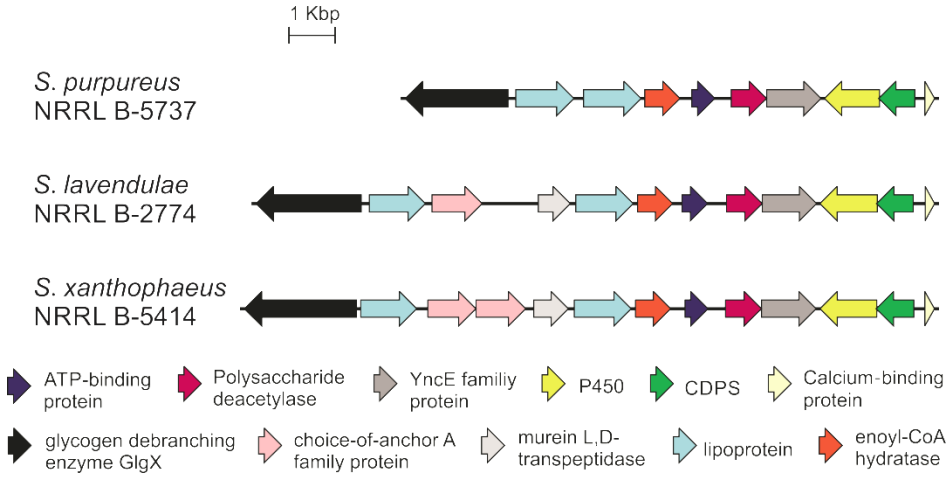


Figure 13. *CDPS*- and *P450*-containing gene clusters from three *Streptomyces* strains.

As is shown in Figure 14, to characterize the three CDP tailoring P450s from *S. pupureus*, S*. lavendulae*, and *S. xanthophaeus*, the bioinformatic analysis was made firstly. The gene and amino acid sequences of the three P450 homologues were downloaded from the NCBI database. Alignments of the three P450 genes with the structurally defined P450s from bacteria revealed three featured motifs that are conserved in the bacterial P450s: (i) the highly conserved A/G^{241}G^{242}XXT^{245} motif with Thr as a critical residue for the H-bonding network (this residue is replaced by a Ser in

the CDP tailoring P450s; EryF numbering); (ii) $E^{280}XXR^{283}$ motif found in the K helix; (iii) $G^{347}XXXC^{351}$ motif in the heme-binding loop harboring the invariant and catalytically most important Cys. Phylogenetic analysis of the CDP tailoring P450s with functionally characterized P450s from bacteria indicated that the CDP tailoring P450s of interest are closely related to CYP121 as well as P-450mel, CYP158A1 and CYP158A2, which catalyze the intramolecular C–C coupling of cYY and dimerization of hydroxynaphthalenes/-naphthoquinones *via* aryl–aryl coupling in *Streptomyces*[76-78].

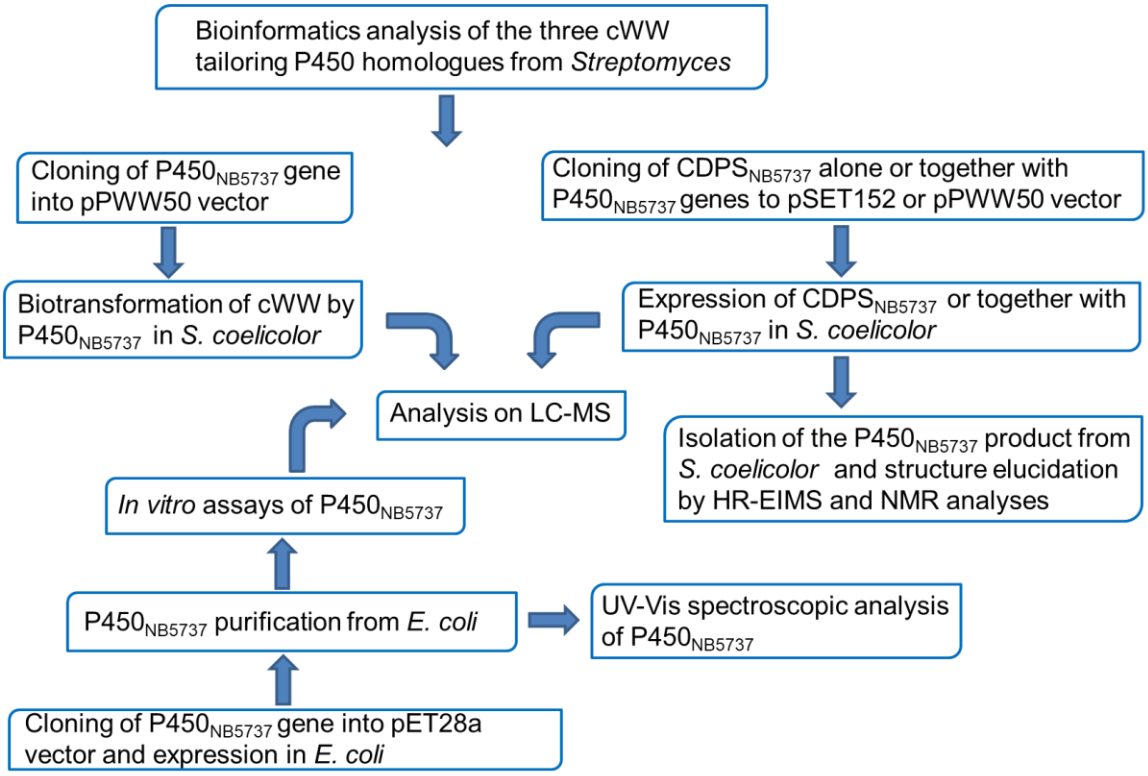
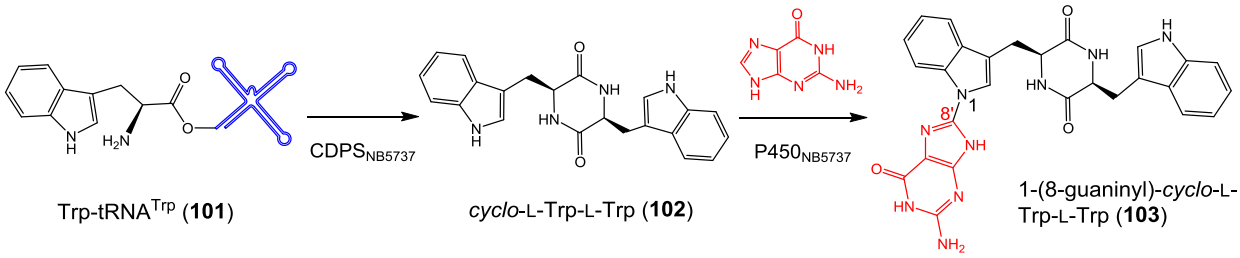


Figure 14. Working flow of identification of a cWW tailoring P450.

Due to the high identity of over 75% on the amino acid sequence level among the three P450 homologues, the $P450_{NB5737}$ was investigated as a representative. *CDPS*$_{NB5737}$ alone or together with *P450*$_{NB5737}$ were cloned into the integrative plasmid pSET152 or replicative vector pPWW50 and expressed in *S. coelicolor* M1146. LC-MS analysis of cultures of M1146 transformants carrying the coexpression constructs of *CDPS*$_{NB5737}$ with *P450*$_{NB5737}$ revealed the additional product with $[M+H]^+$ ions at *m/z* 522.1997 comparing to the *CDPS*$_{NB5737}$ transformant. The cWW metabolism by $P450_{NB5737}$ was

confirmed by feeding 80 µM cWW to two days old culture of *CDPS*$_{NB5737}$ transformant. For structure elucidation, 13 mg of P450$_{NB5737}$ product was isolated from 4 L GYM cultures of M1146 transformant carrying the expression construct of *CDPS*$_{NB5737}$ and *P450*$_{NB5737}$ and sent to ^{1}H, ^{13}C APT, DQF-COSY, HSQC, and HMBC NMR analyses. The structure elucidation indicated that the P450$_{NB5737}$ catalyze the transfer of a guanine moiety to the indole ring of the cWW (**102**) *via* a C–N bond, leading to formation of **103** (Scheme 7).



Scheme 7. Biosynthesis and metabolism of cWW in *S. purpureus* NRRL B-5737 (modified from publication[113])

To investigate the activity of P450$_{NB5737}$ *in vitro*, the P450$_{NB5737}$ gene was cloned into pET28a vector and expressed in *E. coli* Rosetta (DE3). P450$_{NB5737}$ production was induced with 0.5 mM IPTG at 18 °C for 20 h in TB media. Approximate 3.5 mg/L culture P450$_{NB5737}$ was obtained after purification with Ni-NTA agarose. SDS-PAGE analysis revealed a dominant band at approximate 44.0 kDa (Figure 15). The enzyme assays were carried out by using spinach ferredoxin and ferredoxin-NADP^{+} reductase for electron transport. Incubation of the purified P450$_{NB5737}$ with cWW and guanine in the presence of electron transport composition and NADPH resulted in the formation of same product as from the *CDPS*$_{NB5737}$ and *P450*$_{NB5737}$ transformant. No product was detected in negative controls with heat-inactivated protein, without cWW or guanine. Unfortunately, the conversion yield was extremely low, with only around 1% was observed from *in vitro* assays. Increasing the concentration of ferredoxin, ferredoxin reductase, or both to 5-fold did not change the conversion yield. UV-Vis spectra analysis of ferrous CO complex of P450$_{NB5737}$ after reduction with sodium dithionite revealed strong absorption at 415 nm of an inactive form and only a very weak shoulder peak around 446 nm of the active form. This could be caused by the inappropriate

folding or non-incorporation of heme group into $P450_{NB5737}$ produced in *E. coli*. The incorrect state of the protein could also explain the low activity *in vitro*.

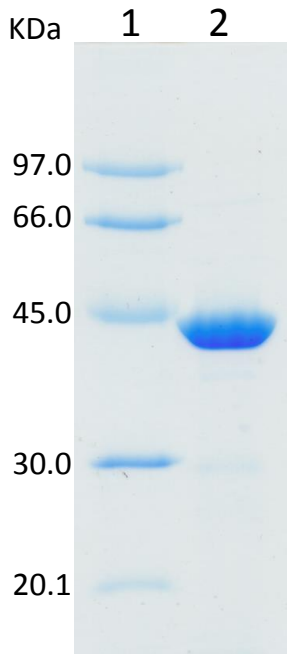


Figure 15. SDS-PAGE analysis of the recombinant $P450_{NB5737}$.

Finally, the catalytic mechanism of $P450_{NB5737}$ was proposed (Figure 16). In this hypothesis, two radicals (compound II) are formed parallel *via* oxidation of FeII *via* FeIII and ferryl-oxo intermediate (Fe^{IV}=O, compound I). Coupling of the two radicals leads to the formation of cWW adducts with guanine.

In conclusion, we identified in this study a P450 enzyme from *Streptomyces* that catalyzes an unusual coupling reaction of a cyclodipeptide with a nucleobase, leading to unprecedented adducts of cyclodipeptide with guanine. This represents one additional rare atypical P450 example catalyzing the coupling of two different substrates. Coupling of guanine with cWW significantly increases the CDP diversity and opens new prospective pathways in natural product discovery. To better understand the catalytic mechanism of this P450, the crystal structure is necessary to be solved in the future.

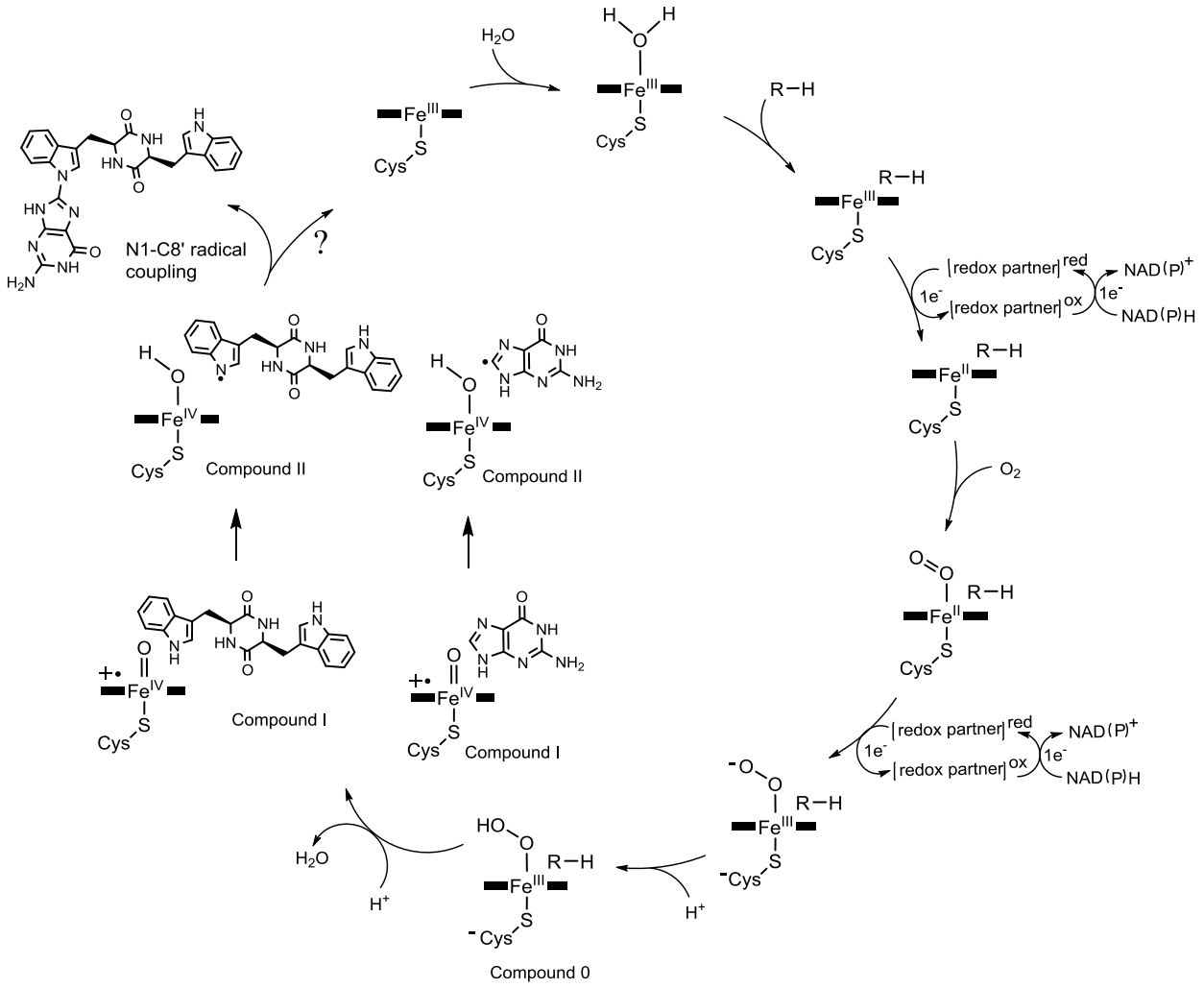


Figure 16. The proposed catalytic mechanism of P450$_{NB5737}$.

For details on this work, please see the publication (section 4.2)

Yu, H., Xie, X., & Li, S. M. (2018). Coupling of guanine with *cyclo*-L-Trp-L-Trp mediated by a cytochrome P450 homologue from *Streptomyces purpureus*. *Org. Lett*. *20* (16), 4921-4925.

3.3 Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate

As described previously, prenyltransferases of the DMATS superfamily are versatile enzymes accepting a large number of aromatic substrates with relatively strict prenyl donor spectra[88]. To broaden the prenyl donor spectra, the acceptance of the unnatural prenyl analogs such as MAPP, 2-pen-PP, and benzyl-PP have been tested in the presence of two tryptophan prenyltransferases FgaPT2 and 5-DMATS and five cyclic dipeptide prenyltransferases, BrePT, FtmPT1, AnaPT, CdpNPT, and CdpC3PT in our laboratory prior to this study[109-111]. The results indicated that the unnatural DMAPP analogs were successfully utilized by these PTs. The structure elucidation of the products revealed partial or complete shifts of alkylation or benzylation positions comparing to the natural prenyl donor DMAPP[109-111]. This inspired us to test these DMAPP analogues towards tyrosine *O*-prenyltransferases TyrPT and SirD. The working flow was described in Figure 17.

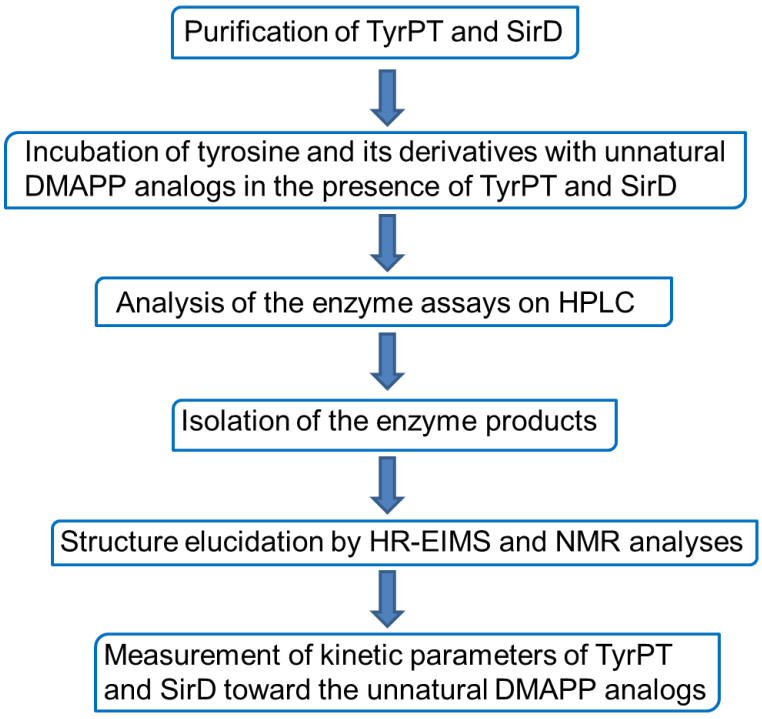


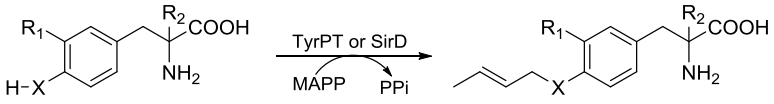
Figure 17. Working flow of investigation on the acceptance of TyrPT and SirD toward unnatural DMAPP analogs.

The enzyme assays towards the unnatural substrates often require relatively high amount of proteins in comparison to their natural substrates. Therefore the expression vector of SirD was changed from pQE70 (pAK2) to pET28a (pHY07). With the help of the strong T7 promoter and nutrient-rich TB media, the protein yield was improved from 2 to 4 mg/L *E. coli* SoluBL21 culture. TyrPT (pHC 16) with yield of 4–6 mg/L was obtained in *E. coli* SoluBL21.

At first, the DMAPP analogs MAPP, 2-pen-PP and benzyl-PP were incubated in the presence of TyrPT or SirD together with tyrosine and 12 derivatives. The conversion yields were measured on HPLC. Nineteen of the enzyme assays showed product yields of more than 10%. Only one substrate, 4-amino-L-phenylalanine (**106**), accepted all three DMAPP analogs with high conversions (34%–67%) for both TyrPT and SirD.

For structure elucidation, nine of the products were isolated on HPLC using reverse phase C18 column and subjected to NMR and MS analyses. It was shown that both TyrPT and SirD catalyzed the *O*-/*N*-alkylation/benzylation of tyrosine and derivatives (**104**-**117**) at the same position as in the presence of their natural prenyl donor DMAPP (Scheme 8).

MAPP:

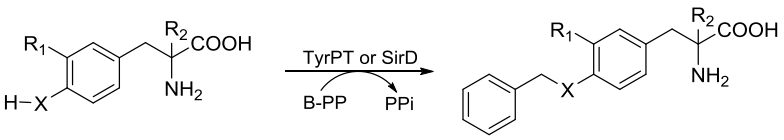


104 X=O; R_1=H; R_2= **105**
106 X=NH; R_1=H; R_2= **107**
108 X=O; R_1=F; R_2= **109**

2-pen-PP:

104 X=O; R_1=H; R_2= **110**
106 X=NH; R_1=H; R_2= **111**
108 X=O; R_1=F; R_2= **112**
113 X=O; R_1=OH; R_2= **114**
115 X=O; R_1=H; R_2= **116**

benzyl-PP:



106 X=NH; R_1=H; R_2= **117**

Scheme 8. The reactions of tyrosine and derivatives and unnatural DMAPP analogs catalyzed by TyrPT and SirD, respectively.

Finally, Michaelis-Menten constants (K_M) values and turnover number (k_{cat}) of TyrPT and SirD were calculated for unnatural alkyl or benzyl donors in the presence of L-tyrosine or 4-amino- L-phenylalanine. The results indicated that MAPP was accepted by both enzymes with lower K_M values than those of 2-pen-PP, benzyl-PP, and DMAPP. All three unnatural donors showed much lower k_{cat} values than those of DMAPP.

In conclusion, tyrosine *O*-prenyltransferases TyrPT and SirD were demonstrated to accept MAPP, 2-pen-PP and benzyl-PP in the presence of tyrosine and derivatives by exhibiting similar behavior as their natural prenyl donor DMAPP. This study expanded the prenyl donor spectra of TyrPT and SirD and increased significantly the diversity of prenylated products by chemoenzymatic synthesis.

For details on this work, please see the publication (section 4.3)

Yu, H., Liebhold, M., Xie, X., & Li, S. M. (2015). Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate. *Appl. Microbiol. Biotechnol.*, *99* (17), 7115-7124.

4 Publications

4.1 Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains.

Applied Microbiology and Biotechnology (2018) 102:4435–4444
https://doi.org/10.1007/s00253-018-8908-6

Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains

Jing Liu[1] · Huili Yu[1] · Shu-Ming Li[1]

Received: 17 January 2018 / Accepted: 28 February 2018 / Published online: 24 March 2018

Abstract

Cyclodipeptide synthases (CDPSs) comprise normally 200–300 amino acid residues and are mainly found in bacteria. They hijack aminoacyl-tRNAs from the ribosomal machinery for cyclodipeptide formation. In this study, nine new CDPS genes from eight *Streptomyces* strains were cloned into pET28a vector and expressed in *Escherichia coli*. Structural elucidation of the isolated products led to the identification of one *cyclo*-L-Trp-L-Leu, two *cyclo*-L-Trp-L-Pro, and three *cyclo*-L-Trp-L-Trp synthases. Other three CDPSs produce *cyclo*-L-Trp-L-Ala or *cyclo*-L-Trp-L-Tyr as the major cyclodipeptide. Total product yields of 46 to 211 mg/L *E. coli* culture were obtained. Our findings represent rare examples of CDPS family derived from actinobacteria that form various tryptophan-containing cyclodipeptides. Furthermore, this study highlights the potential of the microbial machinery for tryptophan-containing cyclodipeptide biosynthesis and provides valid experimental basis for further combination of these CDPS genes with other modification genes in synthetic biology.

Keywords Aminoacyl t-RNA · Diketopiperazine · Cyclodipeptide synthase · *Streptomyces* · Tryptophan-containing cyclodipeptide

Introduction

Cyclodipeptides (CDPs) with a 2,5-diketopiperazine (DKP) skeleton are chemically condensation products of two amino acids and represent a class of secondary metabolites with simple scaffold, but different modification possibilities (Borthwick 2012). A large number of DKP derivatives show a broad range of biological and pharmacological activities such as antibacterial, antifungal, antiviral, and immunosuppressive activities (Borthwick 2012; Giessen et al. 2013), which make them attractive molecules for drug discovery and development. So far, DKP derivatives are mainly isolated from microorganism including bacteria and fungi. The DKP scaffold in fungi is usually assembled by bimodular nonribosomal peptide synthetases (NRPSs) with a typical polypeptide chain length of 2300 amino acids (Maiya et al. 2006; Walsh 2016; Xu et al. 2014), whereas in bacteria mainly by tRNA-dependent cyclodipeptide synthases (CDPSs) comprising normally 200–300 amino acid residues (Giessen and Marahiel 2014; Gondry et al. 2009; James et al. 2015; Moutiez et al. 2017). NRPSs can use both proteinogenic and non-proteinogenic free amino acids as substrates, while CDPSs directly hijack the activated aminoacyl-tRNAs (aa-tRNAs) from the ribosomal machinery for cyclodipeptide formation (Huang et al. 2010).

The DKP scaffold is frequently modified by methylation, hydroxylation, prenylation, dimerization, and further cyclization (Borthwick 2012; Giessen et al. 2013; Li 2010; Xu et al. 2014). Among the DKPs, tryptophan-containing CDPs have received an increasing interest in recent years due to their promising scaffolds for structural modification (Li 2010; Xu et al. 2014). The electron-rich indole ring in tryptophanyl moiety can undergo different enzymatic and spontaneous modifications and rearrangement like prenylation, oxidation, methylation, and dehydrogenation to generate increased chemical complexity (Alkhalaf and Ryan 2015; Walsh 2014).

Since 2005, we have intensively studied prenyltranferases (PTs) involved in the biosynthesis of indole alkaloids in microorganisms, especially those for the prenylation of

Jing Liu and Huili Yu contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00253-018-8908-6) contains supplementary material, which is available to authorized users.

✉ Shu-Ming Li
shuming.li@staff.uni–marburg.de

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

tryptophan and tryptophan-containing CDPs (Winkelblech et al. 2015). The tryptophan-containing CDP PTs catalyzed regio- and stereospecific prenylations at the indole ring. PTs for specific prenylation at nearly all the positions of the indole ring have been characterized (Winkelblech et al. 2015; Wohlgemuth et al. 2017). Most of these PTs are from ascomycetous fungi and are involved in the modification of CDPs assembled by NRPSs. In the view of synthetic biology, it would be of essential importance to combine such PT genes with the small CDPS genes forming tryptophan-containing CDPs, especially those for *cyclo*-L-Trp-L-Pro (cWP), *cyclo*-L-Trp-L-Ala (cWA), *cyclo*-L-Trp-L-Leu (cWL), and *cyclo*-L-Trp-L-Trp (cWW). All these CDPs serve as precursors in the biosynthesis of diverse pharmacologically active compounds and are very well accepted by microbial PTs (Fan et al. 2015; Li 2010; Winkelblech et al. 2015).

By the end of July 2016, about 450 putative CDPS genes were identified by sequence homology search and their number is increasing steadily (Moutiez et al. 2017). Sixty-five of these genes have been proven to be responsible for the formation of different DKPs and about 30 are from actinomycetes (Brockmeyer and Li 2017; Jacques et al. 2015; James et al. 2015; Moutiez et al. 2017). At least ten of the identified CDPSs from eight actinomycetes catalyze the formation of cWW as the sole product (Fig. S1). These include Amir_4627 from *Actinosynnema mirum*, NozA, and NcdA from *Nocardiopsis* sp. CMB-M0232, whose products had been isolated and identified (Giessen et al. 2013; James et al. 2015). The products of other known cWW synthases were identified by LC-MS analysis (Jacques et al. 2015). CDPSs synthesizing other tryptophan-containing cyclodipeptides have not been reported in microorganisms. Until now, only one CDPS originating from the sea anemone *Nematostella vectensis* was demonstrated to produce a number of tryptophan-containing cyclodipeptides like cWL, cWA, *cyclo*-L-Trp-L-Phe (cWF), *cyclo*-L-Trp-L-Met (cWM), and *cyclo*-L-Trp-Gly (cWG) with very low product yields of 0.8–1.0 mg/L in *Escherichia coli* (Seguin et al. 2011). It is therefore absolutely essential to find new CDPSs for the formation of DKPs comprising tryptophan and another amino acid.

Materials and methods

Computer-assisted sequence analysis

Protein sequences listed in Figs. 1 and S1 were taken from the NCBI database (http://www.ncbi.nlm.nih.gov/protein) and compared with each other by using BLAST programs (http://blast.ncbi.nlm.nih.gov/). Multiple sequence alignments were carried out by using the program ClustalW and visualized with ESPript 3.0 (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) to identify strictly conserved amino acid residues. The phylogenetic trees in Figs. 1 and S1 were created by MEGA version 5.0 (http://www.megasoftware.net).

Bacterial strains, plasmids, and culture conditions

pGEM-T Easy used for cloning and pET28a (+) for expression experiments were obtained from Promega (Mannheim, Germany) and Novagen (Darmstadt, Germany), respectively. *E. coli* DH5α (Invitrogen) was used as host for cloning and SoluBL21 (Genlantis) for expression experiments. They were grown at 37 °C in liquid Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl)) or on solid LB medium with 1.5% agar (*w*/*v*). For selection of recombinant *E. coli* strains, 100 μg/mL of ampicillin or 50 μg/mL of kanamycin were supplied in the media.

Streptomyces strains listed in Table S1 were kindly provided by ARS Culture Collection (NRRL) or purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). They were maintained at 28 °C in GYM (4.0 g/L glucose, 4.0 g/L yeast extract, 10.0 g/L malt extract, 15.0 g/L agar, and pH 7.2) or ISP4 medium (10.0 g/L soluble starch, 1.0 g/L K_2HPO_4, 1.0 g/L $MgSO_4$ $7H_2O$, 1.0 g/L NaCl, 2.0 g/L $(NH_4)_2SO_4$, 1.0 g/L $CaCO_3$, 1.0 mL/L trace element solution, 15.0 g/L agar, and pH 7.2).

Plasmid construction and gene expression in *E. coli*

Genetic manipulation in *E. coli* was carried out according to the protocol by Sambrook and Russell (2001). Isolation of genomic DNA from *Streptomyces* was performed as described previously (Kieser et al. 2000). The genes encoding CDPSs were amplified by PCR from genomic DNA by using primers listed in Table 1 with Phusion® High-Fidelity DNA Polymerase from New England Biolabs (NEB). The generated PCR fragments were cloned into pGEM-T Easy vector and sequenced by SEQLAB (Göttingen, Germany) to confirm sequence integrity. After sequencing, the fragments were released with the appropriate restriction endonucleases from pGEM-T Easy vector and ligated into pET28a (+) vector, which was digested with the same enzymes previously. The generated plasmids (Table 1) were transformed into *E. coli* SoluBL21 for gene expression.

For CDPS overproduction, 0.5 mL of 16-h cultures of different expression constructs was used to inoculate 50 mL LB medium containing 50 μg/mL kanamycin. The cultures were maintained at 230 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 to a final concentration of 0.2 or 0.5 mM. The induction was kept at 20, 30, or 37 °C for 20 h for condition optimization (Table S2) or for 36 h for quantification. One milliliter of culture was extracted with one volume of ethyl acetate for three times. The organic phases were combined and evaporated, and the residues were afterward

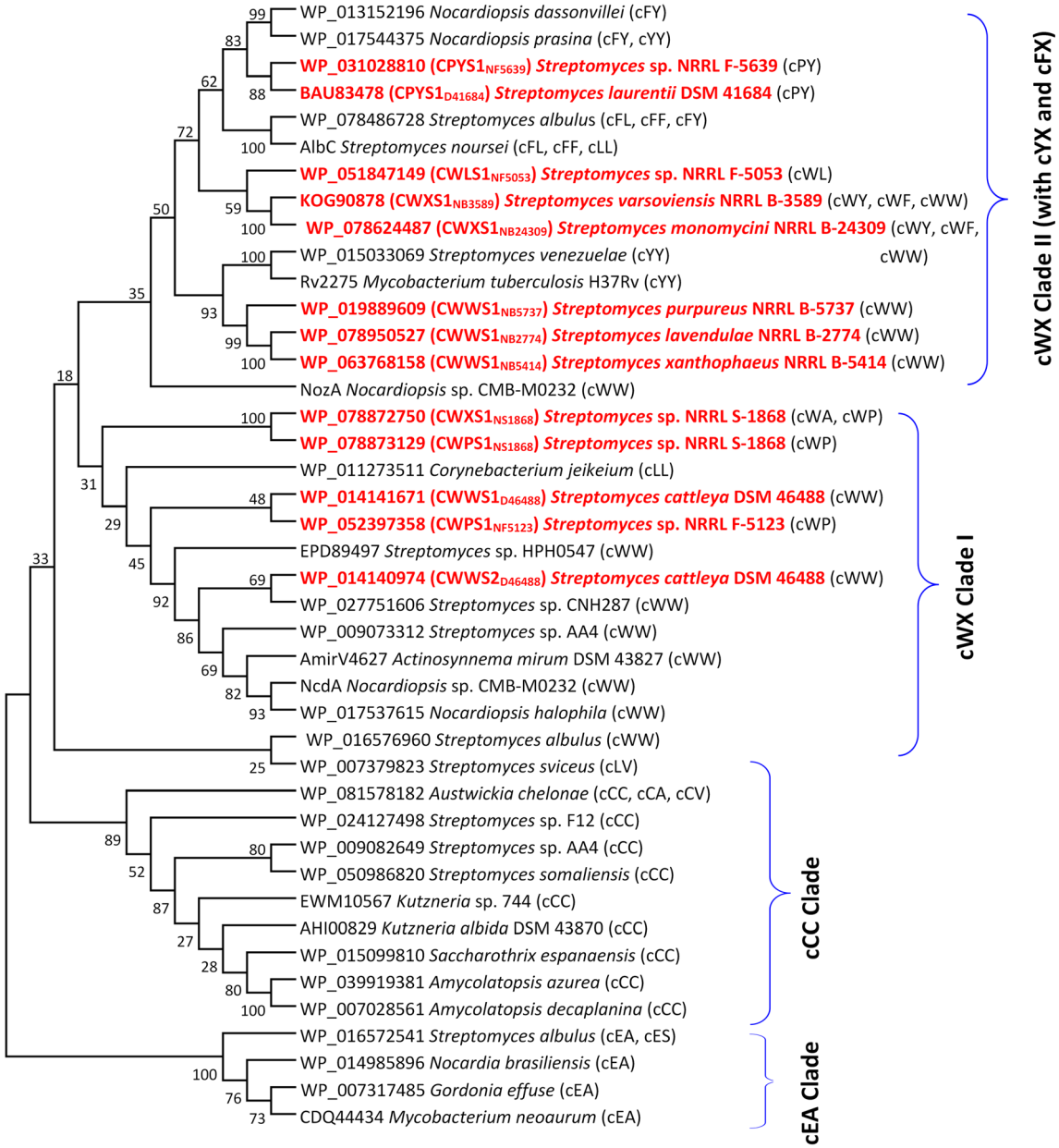


Fig. 1 Phylogenetic analysis of CDPSs from actinobacteria. The main CDPS products are given in parenthesis. Enzymes investigated in this study are highlighted in red

dissolved in 40 μL of methanol. Five microliters of such samples were taken for LC-MS analysis.

HPLC and LC-MS analyses

The ethyl acetate extracts were analyzed on an Agilent HPLC 1260 series equipped with a photo diode array detector and a Bruker microTOF QIII mass spectrometer by using the Agilent Eclipse XDB C18 column (5 μm, 4.6 × 150 mm). A linear gradient of 5–100% acetonitrile in water, both containing 0.1% formic acid, in 40 min and a flow rate at 0.25 mL/min were used. The column was then washed with 100% acetonitrile containing 0.1% formic acid for 5 min and equilibrated with 5% acetonitrile in water for 5 min. The parameters of the spectrometer were set as following: electrospray positive ion mode for ionization, capillary voltage with 4.5 kV, collision energy with 8.0 eV.

For quantification, an Agilent HPLC 1200 series equipped with a photo diode array detector and an identical column were used. A linear gradient of 10 to 100% acetonitrile in water without acid in 40 min was followed by 100% acetonitrile for 5 min and 10% acetonitrile in water for 5 min. The flow rate was 0.5 mL/min. The absorption at 280 nm was used for quantification.

Table 1 Information on CDPSs identified in this study

CDPS	Accession number	Protein length (aa)	Primer sequences (5′-3′)	Cloning constructs in pGEM-T easy	Cloning sites in pET28a	Expression constructs
$CWWS1_{D46488}$	WP_014141671	278	AACATATGCACGACAACGGTCATCGGCCC TTGGATCCTTACGTCACCGGCT	pHY02	*Nde*I *Bam*HI	pHY34
$CWWS2_{D46488}$	WP_014140974	234	GCCATATGCTGCATCGAACGTCCTT TAGGATCCTCACGGCTCGGCGGGCAGTT	pHY01	*Nde*I *Bam*HI	pHY11
$CWWS1_{NB2774}$	WP_078950527	265	CATATGACGATCACAGCTGACGCATCATTC GGATCCTCAAGCTGCTCGACGCTCAT	pHY72	*Nde*I *Bam*HI	pHY82
$CWWS1_{NB5737}$	WP_019889609	239	GCCATATGACTCTCATCGAAGACAC GGATCCTCAAGCGGCTCGACGGTCAT	pHY08	*Nde*I *Bam*HI	pHY09
$CWWS1_{NB5414}$	WP_063768158	273	CATATGAGGGCGATCACACAGGTGAC GGATCCTCAAGCTGCTCGACGCTCAT	pHY77	*Nde*I *Bam*HI	pHY83
$CWPS1_{NS1868}$	WP_078873129	241	CATATGAACACTTCCCTCGCTGC GAATTCTCAGCGTTCGGCCGCCCGGTC	pHY103	*Nde*I *Eco*RI	pHY107
$CWLS1_{NF5053}$	WP_051847149	243	CATATGTGCGAGGGCGCCGATGTGC GGATCCTCAGGATTCGTCCACCGG	pJL03	*Nde*I *Bam*HI	pJL26[a]
$CWPS1_{NF5123}$	WP_052397358	263	CATATGACCAGCAGGACCGAAAC GAATTCTCACGGAAGCAGCCGGGGG	pHY100	*Nde*I *Eco*RI	pHY105
$CWXS1_{NS1868}$	WP_078872750	267	CATATGGCCACACACGCCTCCGC GAATTCTCACTGCTGCGTCACGTGGTC	pHY101	*Nde*I *Eco*RI	pHY106
$CWXS1_{NB3589}$	KOG90878	255	CATATGGGGGCCCCGCAGCCC GGATCCTCACGTCAAGTCCCTTTCTCC	pJL02	*Nde*I *Bam*HI	pJL25
$CWXS1_{NB24309}$	WP_078624487	282	GCGCATATGAGTGCATCGCAGGCTGCTG GGATCCTCACGTCACGTCCCCTGC	pJL01	*Nde*I *Bam*HI	pJL24[b]
$CPYS1_{NF5639}$	WP_031028810	278	CATATGAACCGCCGCTGTTCCTTCG GGATCCTCATGCTCCTGGGGCACTG	pJL04	*Nde*I *Bam*HI	pJL29
$CPYS1_{D41684}$	BAU83478	224	CGCATATGAACCAGTTCGACGTGCTGCC GGATCCCCGGACCACGAGGAAGCC	pJL05	*Nde*I *Bam*HI	pJL30[c]

Restriction sites for cloning are underlined in primer sequences

[a] The 12 base pairs at the 5′-end of the sequence coding for WP_051847149 were not included in the expression construct pJL26

[b] The 96 base pairs at the 5′-end of the sequence coding for WP_078624487 were not included in the expression construct pJL24

[c] According to the annotation in database, BAU83478 comprises 484 amino acids and consists of CDPS sequence at its N-terminus and that for a putative methyltransferase at its C-terminus. The sequence coding for amino acids 1-224 of BAU83478 was cloned into pET28a generating pJL30

Isolation and structural elucidation of generated cyclodipeptides

For structural elucidation of the accumulated CDPs, *E. coli* SoluBL21 cultures carrying the respective expression constructs were prepared under the best tested conditions (Table S2) and extracted with ethyl acetate. The CDPs were then isolated on an Agilent HPLC 1260 series equipped with a photo diode array detector by using a semi-preparative Multospher 120 RP-18 column (10 × 250 mm, 5 μm) with the same solvents and gradient as for quantification. The flow rate was set to 2.5 mL/min. The obtained products were purified on the same column with slightly improved gradients, if necessary. In this study, 5, 3, and 4 mg of cWW were isolated as white powders from 300 mL *E. coli* cultures with pHY34, pHY11, and pHY09, respectively. Thirty milligrams of cWP and 12 mg of cWL were isolated as white powders from 1000 mL *E. coli*/pHY107 and *E. coli*/pJL26, respectively. From 2000 mL *E. coli* culture with pJL25, 1.5 mg of cWA, 10 mg of *cyclo*-L-Trp-L-Tyr (cWY), 5 mg of cWM, 2.5 mg of *cyclo*-L-Trp-L-Val (cWV), and 2.5 mg of cWF were obtained as white powders.

The isolated compounds were subjected to NMR and MS analyses. ^{1}H NMR spectra (Figs. S2−S9) were recorded at room temperature on an ECX-500 spectrometer (JEOL, Tokyo, Japan). All spectra were processed with MestReNova 5.2.2 (Metrelab Research, S5 Santiago de Compostella, Spain). MS and NMR data of the cyclodipeptides are provided as Tables S3 and S4. Interpretation of the NMR and MS data of the isolated products and comparison with those of known compounds led to the undoubted identification of the cyclodipeptides.

Results

Phylogenetic analysis of CDPSs

Phylogenetic analysis of known CDPSs from actinomycetes (Fig. S1) revealed that the known cWW synthases are located close together, with the exception for WP_016576960 from *Streptomyces albulus*. CDPSs for DKPs consisting of aliphatic amino acids like *cyclo*-L-Cys-L-Cys (cCC) and *cyclo*-L-Glu-L-Ala (cEA) and those

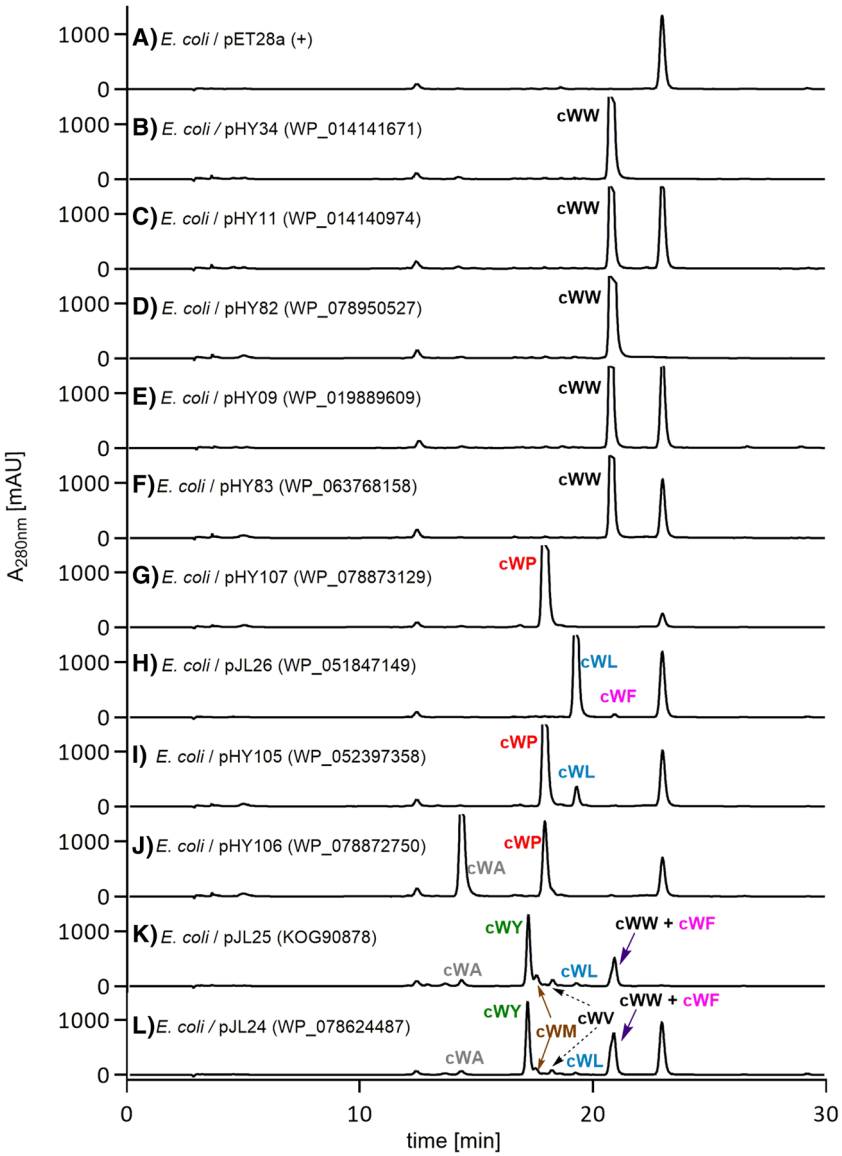


Fig. 2 LC-MS analysis of recombinant *E. coli* strains with empty vector (**a**) or with CDPSs (**b**–**l**). The ethyl acetate extracts were analyzed on an Agilent HPLC series 1260 with UV and mass detections. Absorptions at 280 nm are illustrated in this figure. The products are predicted by inspection of their exact $[M + H]^+$ ions and proven by NMR analysis after isolation

comprising aromatic amino acids such as *cyclo*-L-Tyr-L-Tyr (cYY) and *cyclo*-L-Tyr-L-X (cYX) build their own groups. To get CDPSs for other tryptophan-containing DKPs rather than cWW, we searched in databases for homologs of the actinobacterial cWW synthases and integrated them into Fig. S1 generating Fig. 1. Thirteen candidate CDPSs close to cWW and cYY synthases from 11 *Streptomyces* strains (highlighted in red in Fig. 1) were selected as investigation objectives.

Overproduction of CDPSs in *E. coli* and identification of the products

For functional proof, we established our expression system in *E. coli* SoluBL21 by using the two known CDPSs WP_014141671 and WP_014140974 from *Streptomyces cattleya* DSM 46488, termed $CWWS1_{D46488}$ (cWW synthase 1 from strain DSM 46488) and $CWWS2_{D46488}$ in this study, respectively (Fig. 1). After optimization of expression conditions (Table S2), the 20-h-old cultures after induction at 30 °C with 0.5 mM IPTG for $CWWS1_{D46488}$ and 0.2 mM for $CWWS2_{D46488}$ were extracted with ethyl acetate and then analyzed on LC-MS. In comparison to the culture of *E. coli* harboring the empty vector (Fig. 2a), one additional predominant peak each with the expected $[M + H]^+$ ion at m/z 373.1636 for cWW was detected in *E. coli* transformants (Fig. 2b, c). Isolation and structure elucidation of both peaks with the help of NMR confirmed their integrity as cWW (Tables S3 and S4, Fig. S2).

In analogy, the other nine candidate genes were individually cloned into pET28a (+) (Table 1) and expressed in *E. coli* SoluBL21 under different optimized conditions (Table S2). The ethyl acetate extracts of the 20-h cultures were analyzed on LC-MS. In comparison to that of negative control (Fig. 2a),

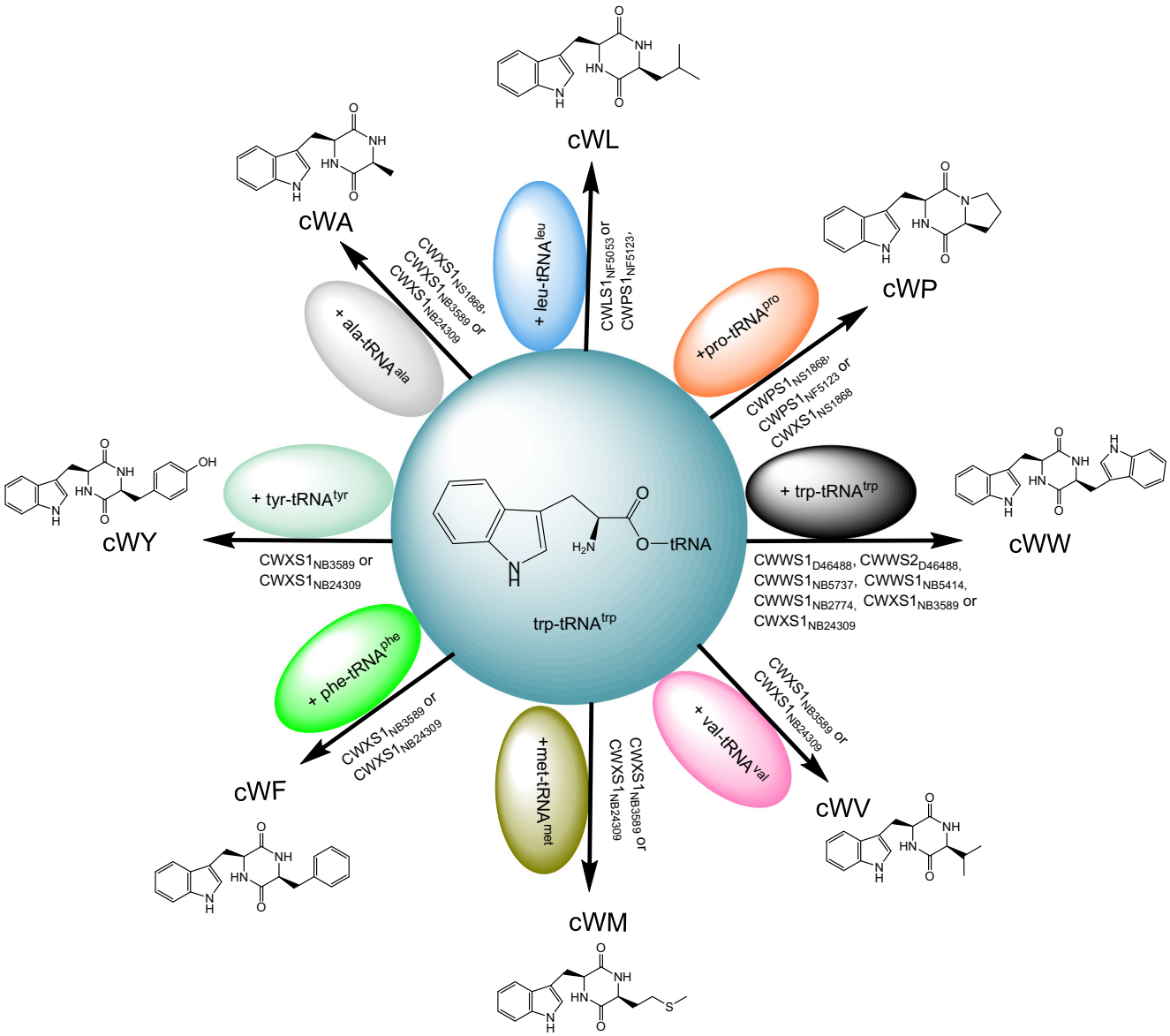


Fig. 3 Formation of tryptophan-containing cyclodipeptides by CDPSs

the chromatograms of five CDPS transformants showed one predominant product peak each (Fig. 2d–h). These include WP_078950527 from *Streptomyces lavendulae* NRRL B-2774 (Fig. 2d), WP_019889609 from *Streptomyces purpureus* NRRL B-5737 (Fig. 2e), WP_063768158 from *Streptomyces xanthophaeus* NRRL B-5414 (Fig. 2f), WP_078873129 from *Streptomyces* sp. NRRL S-1868 (Fig. 2g), and WP_051847149 from *Streptomyces* sp. NRRL F-5053 (Fig. 2h). The products of the first three transformants (Fig. 2d–f) share the same retention time and $[M+H]^+$ ion with those of the two known $CWWS1_{D46488}$ and $CWWS2_{D46488}$ (Fig. 2b, c), indicating the presence of cWW. Interpretation of the NMR and MS data of the isolated peak from the culture with pHY09 (Tables S3 and S4, Fig. S2) confirmed these products to be cWW (Giessen et al. 2013; Lu et al. 2017). Therefore, these three enzymes were termed $CWWS1_{NB2774}$, $CWWS1_{NB5737}$, and $CWWS1_{NB5414}$, respectively (Fig. 3).

For structure elucidation, the products of WP_078873129 (Fig. 2g) and WP_051847149 (Fig. 2h) were isolated from the bacterial cultures and subjected to NMR and MS analyses (Tables S3 and S4; Figs. S3 and S4), which proved unequivocally cWP (Grundmann and Li 2005; He et al. 2013; Kumar et al. 2014) as the sole product of WP_078873129, termed hereafter $CWPS1_{NS1868}$ (Fig. 3). The two products of WP_051847149 with a ratio of 120:1 were identified as cWL (Kumar et al. 2014) and cWF (Chu et al. 2011),

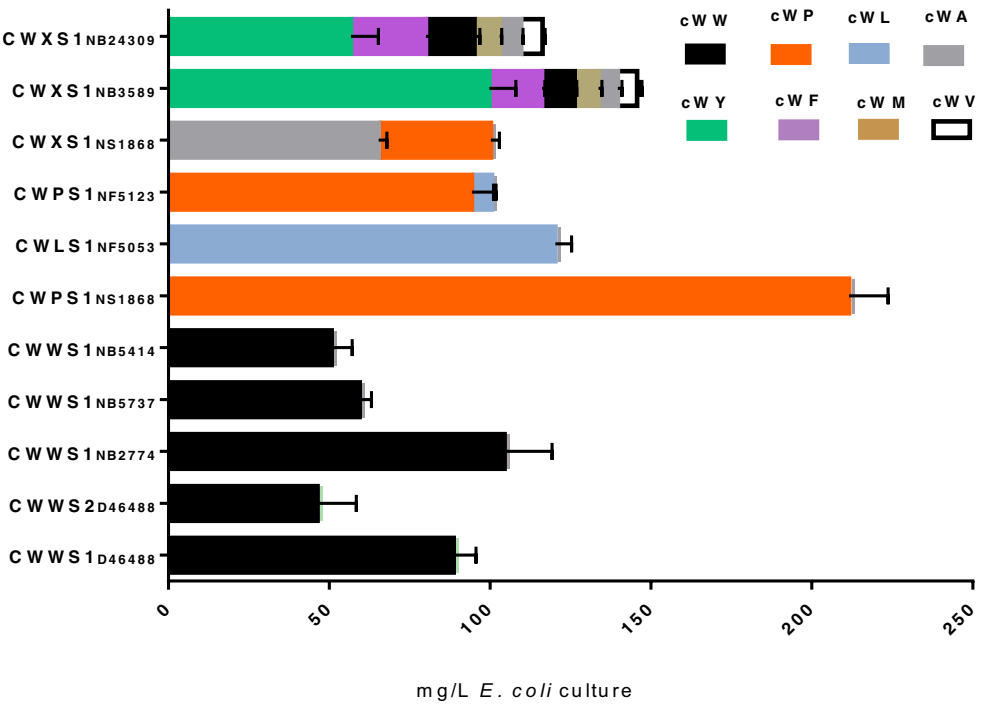


Fig. 4 Product yields of CDPSs. The data were obtained from three independent 36-h old cultures

respectively. Due to the significant dominance of cWL, this enzyme was named as $CWLS1_{NF5053}$ (Fig. 3).

Two product peaks each were detected in the transformants with WP_052397358 from *Streptomyces* sp. NRRL F-5123 (Fig. 2i) and another CDPS WP_078872750 from *Streptomyces* sp. NRRL S-1868 (Fig. 2j). One of them was found in both extracts and identified as cWP by comparison of its retention time and $[M+H]^+$ ion with those of $CWPS1_{NS1868}$ product (Fig. 2g). cWP in the transformant with WP_052397358 contributes about 94% to the total CDPs, and this enzyme was therefore termed $CWPS1_{NF5123}$. The minor product of $CWPS1_{NF5123}$ showed a $[M+H]^+$ of cWL. Another main product in *E. coli* with pHY106 (Fig. 2j) had a $[M+H]^+$ of cWA, which was identified by comparison with the product from KOG90878 of *Streptomyces varsoviensis* NRRL B-3589 (Fig. 2k). WP_078872750 was therefore termed $CWXS1_{NS1868}$.

The transformant with KOG90878 had a similar and complex product spectrum as WP_078624487 from *Streptomyces monomycini* NRRL B-24309 (Fig. 2l). Two predominant and at least four minor peaks were detected for both cultures at 280 nm. They differ slightly from each other in the product yields and ratios, especially those of the two dominant peaks. Product isolation from the culture with KOG90878 and structural elucidation proved the main peak at 15.4 min to be cWY and the second major peak as a mixture of cWF and cWW. The four minor peaks were identified as cWA, cWM, cWV, and cWL, respectively (Tables S3 and S4; Figs. S5−S9). KOG90878 and WP_078624487 are therefore responsible for the formation of at least seven tryptophan-containing cyclodipeptides and termed $CWXS1_{NB3589}$ and $CWXS1_{NB24309}$, respectively (Fig. 3).

Product yields of CDPs

Generally, high product yield is a prerequisite for potential application in the biotechnology. The product yields of cWW from *E. coli* culture at 29 and 42 mg/L were reported for the known cWW synthases Amir_4627 and NozA after codon optimization, respectively (Giessen et al. 2013; James et al. 2015). To prove the productivity of the recombinant strains, we determined the CDP contents of 36-h-old cultures by HPLC analysis using the isolated products as standards. As shown in Fig. 4, the product yields achieved in this study for several enzymes are much higher than those mentioned above. The product yields of the five cWW synthases were detected between 46 ± 10 and 104 ± 10 mg/L with the best one for $CWWS1_{NB2774}$. The highest product yield of 211 ± 12 mg cWP per liter culture was observed for the culture with $CWPS1_{NS1868}$, followed by 120 ± 5 mg cWL in the culture with $CWLS1_{NF5053}$. Total product yields between 100 and 140 mg/L were calculated for other four CDPSs. These data provide a solid basis for further combination of these CDPS genes with other modification genes in synthetic biology.

Table 2 Amino acid residues of selected CDPSs constituting the two binding pockets

	Amino acids constituting binding pocket P1								Amino acids constituting binding pocket P2							Reported product	Predicted product*	Identified in this study
AlbC	L_{33}	G_{35}	V_{65}	V_{67}	L_{119}	L_{185}	F_{186}	L_{200}	M_{152}	A_{155}	V_{156}	N_{159}	I_{204}	T_{206}	P_{207}	cFL, cFF, cFY		
Amir_4627	V_{33}	G_{35}	V_{65}	P_{67}	M_{119}	F_{184}	F_{185}	N_{199}	M_{152}	E_{155}	V_{156}	L_{159}	R_{203}	L_{205}	P_{206}	cWW (Giessen et al. 2013)		
NozA	L_{32}	G_{34}	V_{64}	A_{66}	L_{118}	F_{183}	F_{184}	S_{198}	M_{151}	A_{154}	V_{155}	S_{158}	M_{202}	L_{204}	P_{205}	cWW (James et al. 2015)		
NcdA	V_{36}	G_{38}	V_{58}	P_{60}	M_{112}	F_{177}	F_{178}	N_{192}	M_{145}	D_{148}	V_{149}	A_{152}	R_{196}	L_{198}	P_{199}	cWW (James et al. 2015)		
WP_016576960	L_{53}	G_{55}	I_{85}	P_{87}	L_{139}	L_{204}	C_{205}	H_{219}	A_{17}	K_{175}	A_{176}	S_{179}	Q_{223}	I_{225}	P_{226}	cWW (Jacques et al. 2015)		
WP_009073312	V_{26}	G_{28}	I_{58}	P_{60}	M_{112}	F_{177}	F_{178}	N_{192}	M_{145}	E_{148}	V_{149}	A_{152}	Q_{196}	L_{198}	P_{197}	cWW (Jacques et al. 2015)		
EPD89497	V_{29}	G_{31}	I_{61}	P_{63}	M_{115}	F_{180}	F_{181}	C_{195}	A_{148}	T_{151}	V_{152}	A_{155}	R_{199}	L_{201}	P_{198}	cWW (Jacques et al. 2015)		
WP_027751606	V_{25}	G_{27}	I_{57}	P_{59}	M_{111}	F_{176}	F_{177}	N_{191}	M_{144}	D_{147}	V_{148}	A_{151}	Q_{195}	L_{197}	P_{208}	cWW (Jacques et al. 2015)		
WP_017537615	V_{35}	G_{37}	I_{67}	P_{69}	M_{121}	F_{186}	F_{187}	N_{201}	M_{154}	E_{157}	V_{159}	A_{171}	R_{205}	L_{207}	P_{197}	cWW (Jacques et al. 2015)		
$CWWS1_{D46488}$	V_{77}	G_{79}	V_{109}	P_{111}	L_{164}	F_{229}	F_{230}	C_{242}	M_{197}	A_{200}	A_{201}	S_{204}	K_{246}	V_{248}	P_{249}	cWW (Jacques et al. 2015)		cWW
$CWWS2_{D46488}$	V_{27}	G_{29}	V_{59}	P_{61}	M_{113}	F_{178}	F_{179}	C_{193}	M_{146}	E_{149}	A_{150}	A_{153}	R_{197}	I_{199}	P_{200}	cWW (Jacques et al. 2015)		cWW
$CWWS1_{NB2774}$	L_{29}	G_{31}	V_{61}	A_{63}	L_{115}	F_{181}	F_{182}	A_{196}	M_{148}	H_{151}	F_{152}	S_{155}	V_{200}	M_{202}	P_{203}		*cFX*	cWW
$CWWS1_{NB5737}$	L_{28}	G_{30}	V_{60}	A_{62}	L_{114}	F_{180}	F_{181}	S_{195}	M_{147}	T_{150}	F_{151}	P_{154}	Q_{199}	I_{201}	P_{202}		**cFX**	cWW
$CWWS1_{NB5414}$	L_{35}	G_{37}	V_{67}	A_{69}	L_{121}	F_{187}	F_{188}	A_{200}	M_{154}	H_{157}	F_{158}	S_{161}	V_{204}	M_{206}	P_{207}		*cFX*	cWW
$CWPS1_{NS1868}$	L_{32}	G_{34}	I_{64}	P_{66}	L_{118}	F_{189}	F_{190}	V_{204}	A_{151}	R_{154}	A_{155}	V_{158}	M_{208}	V_{210}	P_{211}		*cFX*	cWP
$CWLS1_{NF5053}$	L_{36}	G_{38}	V_{68}	A_{70}	L_{122}	F_{188}	L_{189}	S_{203}	M_{155}	S_{158}	F_{159}	G_{162}	M_{207}	A_{209}	Q_{210}		**cFX**	cWL
$CWPS1_{NF5123}$	L_{42}	G_{44}	V_{74}	P_{76}	L_{128}	L_{193}	L_{194}	A_{208}	A_{161}	S_{164}	V_{165}	G_{168}	K_{212}	I_{214}	D_{215}		*cFX*	cWP
$CWXS1_{NS1868}$	L_{43}	G_{45}	I_{75}	P_{77}	L_{129}	F_{205}	F_{206}	A_{220}	L_{162}	R_{165}	A_{166}	V_{169}	A_{224}	V_{226}	G_{227}		*cFX*	cWA, cWP
$CWXS1_{NB3589}$	L_{30}	G_{32}	I_{62}	A_{64}	L_{116}	F_{193}	F_{194}	S_{208}	M_{149}	R_{152}	H_{153}	G_{156}	M_{212}	T_{214}	P_{215}		**cFX**	cWY, cWF, cWW
$CWXS1_{NB24309}$	–	–	I_{30}	A_{32}	L_{84}	F_{177}	F_{178}	S_{192}	M_{117}	I_{120}	H_{121}	G_{124}	L_{196}	T_{198}	P_{199}		**cFX**	cWY, cWF, cWW

*The products of the CDPSs of interest were predicted in a previous report (in bold) (Jacques et al. 2015) and in this study (in italic) according to the conserved residues suggested by Jacques et al. (Jacques et al. 2015)

Discussion

To find candidates for the formation of tryptophan-containing CDPs, we searched the database by using known cWW synthases as probes and carried out phylogenetic analysis. Gene expression and identification of CDP products led to identification of nine new members for the biosynthesis of tryptophan-containing CDPs. Inspection the phylogenetic tree (Fig. 1) revealed that the tryptophan-containing cyclodipeptide synthases from actinobacteria are located in two clades. With the exception for *cyclo*-L-Leu-L-Leu (cLL), the clade I comprises only cWX synthases, while the members of clade II are responsible for the formation of not only cWX, but also cYX and cFX. To prove the diversity of cWX clade II, we took two genes close to cYX synthases from this clade and expressed them in *E. coli*. As shown in Fig. S10, WP_031028810 from *Streptomyces* sp. NRRL F-5639 and BAU83478 from *Streptomyces laurentii* DSM 41684 were proven to be responsible for the formation of *cyclo*-L-Pro-L-Tyr (cPY). Together with the CDPSs for tryptophan-containing cyclodipeptide biosynthesis (Figs. 2, 3, and 4), these results provide evidence that phylogenetic analysis could be successfully used for prediction and finding of desired CDPS candidates. However, exact CDPS product cannot be predicted by just phylogenetic analysis. For example, the 13 specific cWW synthases are distributed in two cWX clades. One of the cWP synthases, $CWPS1_{NF5123}$ is closer to $CWWS1_{D46488}$ than to the second one $CWPS1_{NS1868}$, which is again closer to $CWXS1_{NS1868}$. $CWPS1_{NS1868}$ and $CWXS1_{NS1868}$ share a sequence identity of 62% on the amino acid level, and the latter produces a mixture of cWP and cWA with a ratio of 15:7 (Figs. 2 and 4).

Based on crystal structure and mutagenesis experiments, Moutiez et al. proposed that two pockets of CDPSs are for aminoacyl-tRNA binding (Jacques et al. 2015; Moutiez et al. 2014a; Moutiez et al. 2014b). It was postulated that P1 is more

specific for binding the first aminoacyl-tRNA, and the wider pocket P2 is responsible for the less specific binding of the second and often variable aminoacyl-tRNA (Moutiez et al. 2014a; Moutiez et al. 2014b). The amino acid residues of the two pockets were used for prediction of CDPS products (Jacques et al. 2015). With the help of the residues in the two pockets, the functions of the nine tryptophan-containing CDPSs identified in this study were predicted to be responsible for the formation of cFX (Table 2) (Jacques et al. 2015). However, as shown in Figs. 2, 3, and 4, only $CWLS1_{NF5053}$, $CWXS1_{NB3589}$, and $CWXS1_{NB24309}$ produce cWF as minor or second major product, with 1, 11, and 22% of total CDPs, respectively. No cFX was detected in other CDPS transformants. Comparing the amino acid residues in pockets 1 and 2 of tryptophan-containing cyclodipeptide synthases with each other and with those of *cyclo*-L-Phe-L-Leu (cFL) synthase AlbC (Table 2) revealed no clear indications for cWX synthases. In a previous study, we demonstrated that mutation of the related residues in the pocket P1 did not lead to the desired changes of product spectrum, but their product yields (Brockmeyer and Li 2017). This indicates that these prediction tools still need to be optimized by identification of more CDPS structures and amino acid residues in the binding pockets.

In summary, after confirmation of two known cWW synthases, we identified three new cWW, one cWL, and two cWP synthases from *Streptomyces*, which catalyze the formation of one predominant DKP. Three additional CDPSs produce one DKP like cWA or cWY as the major metabolite (Figs. 2 and 4). These CDPSs are thus ideal candidates for production of tryptophan-containing CDPs for further modification by tailoring enzymes like PTs. It should be also mentioned that cWV was not reported as CDPS product prior to this study. Therefore, our results expand significantly the product spectrum of tryptophan-containing cyclodipeptide synthases and raise hopes to find other members of this important enzyme group.

Acknowledgements We thank ARS Culture Collection (NRRL) for providing *Streptomyces* strains, S. Newel, and R. Kraut (University Marburg) for taking NMR and MS spectra.

Funding information The Bruker micrOTOF QIII mass spectrometer was financially supported in part by a grant from the Deutsche Forschungsgemeinschaft (INST 160/620-1 to S.-M. L.). J.L. and H.Y are scholarship recipients of China Scholarship Council (201608310118 and 201306220024).

Compliance with ethical standards

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

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

Alkhalaf LM, Ryan KS (2015) Biosynthetic manipulation of tryptophan in bacteria: pathways and mechanisms. Chem Biol 22:317–328

Borthwick AD (2012) 2,5-diketopiperazines: synthesis, reactions, medicinal chemistry, and bioactive natural products. Chem Rev 112:3641–3716

Brockmeyer K, Li S-M (2017) Mutations of residues in pocket P1 of a cyclodipeptide synthase strongly increase product formation. J Nat Prod 80:2917–2922

Chu D, Peng C, Ding B, Liu F, Zhang F, Lin H, Li Z (2011) Biological active metabolite cyclo (L-Trp-L-Phe) produced by South China Sea sponge *Holoxea* sp. associated fungus *Aspergillus versicolor* strain TS08. Bioprocess Biosyst Eng 34:223–229

Fan A, Winkelblech J, Li S-M (2015) Impacts and perspectives of prenyltransferases of the DMATS superfamily for use in biotechnology. Appl Microbiol Biotechnol 99:7399–7415

Giessen TW, Marahiel MA (2014) The tRNA-dependent biosynthesis of modified cyclic dipeptides. Int J Mol Sci 15:14610–14631

Giessen TW, von Tesmar AM, Marahiel MA (2013) A tRNA-dependent two-enzyme pathway for the generation of singly and doubly methylated ditryptophan 2,5-diketopiperazines. Biochemistry 52:4274–4283

Gondry M, Sauguet L, Belin P, Thai R, Amouroux R, Tellier C, Tuphile K, Jacquet M, Braud S, Courcon M, Masson C, Dubois S, Lautru S, Lecoq A, Hashimoto S, Genet R, Pernodet JL (2009) Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. Nat Chem Biol 5:414–420

Grundmann A, Li S-M (2005) Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from *Aspergillus fumigatus*. Microbiology 151:2199–2207

He F, Bao J, Zhang XY, Tu ZC, Shi YM, Qi SH (2013) Asperterrestide A, a cytotoxic cyclic tetrapeptide from the marine-derived fungus *Aspergillus terreus* SCSGAF0162. J Nat Prod 76:1182–1186

Huang R, Zhou X, Xu T, Yang X, Liu Y (2010) Diketopiperazines from marine organisms. Chem Biodivers 7:2809–2829

Jacques IB, Moutiez M, Witwinowski J, Darbon E, Martel C, Seguin J, Favry E, Thai R, Lecoq A, Dubois S, Pernodet JL, Gondry M, Belin P (2015) Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity. Nat Chem Biol 11:721–727

James ED, Knuckley B, Alqahtani N, Porwal S, Ban J, Karty JA, Viswanathan R, Lane AL (2015) Two distinct cyclodipeptide synthases from a marine actinomycete catalyze biosynthesis of the same diketopiperazine natural product. ACS Synth Biol 5:547–553

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical Streptomyces genetics. John Innes Foundation, Norwich

Kumar SN, Mohandas C, Nambisan B (2014) Purification, structural elucidation and bioactivity of tryptophan containing diketopiperazines, from *Comamonas testosteroni* associated with a rhabditid entomopathogenic nematode against major human-pathogenic bacteria. Peptides 53:48–58

Li S-M (2010) Prenylated indole derivatives from fungi: structure diversity, biological activities, biosynthesis and chemoenzymatic synthesis. Nat Prod Rep 27:57–78

Lu C, Xie F, Shan C, Shen Y (2017) Two novel cyclic hexapeptides from the genetically engineered *Actinosynnema pretiosum*. Appl Microbiol Biotechnol 101:2273–2279

Maiya S, Grundmann A, Li S-M, Turner G (2006) The fumitremorgin gene cluster of *Aspergillus fumigatus*: identification of a gene encoding brevianamide F synthetase. Chembiochem 7:1062–1069

Moutiez M, Belin P, Gondry M (2017) Aminoacyl-tRNA-utilizing enzymes in natural product biosynthesis. Chem Rev 117:5578–5618

Moutiez M, Schmitt E, Seguin J, Thai R, Favry E, Belin P, Mechulam Y, Gondry M (2014a) Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. Nat Commun 5:5141

Moutiez M, Seguin J, Fonvielle M, Belin P, Jacques IB, Favry E, Arthur M, Gondry M (2014b) Specificity determinants for the two tRNA substrates of the cyclodipeptide synthase AlbC from *Streptomyces noursei*. Nucleic Acids Res 42:7247–7258

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Seguin J, Moutiez M, Li Y, Belin P, Lecoq A, Fonvielle M, Charbonnier JB, Pernodet JL, Gondry M (2011) Nonribosomal peptide synthesis in animals: the cyclodipeptide synthase of *Nematostella*. Chem Biol 18:1362–1368

Walsh CT (2014) Biological matching of chemical reactivity: pairing indole nucleophilicity with electrophilic isoprenoids. ACS Chem Biol 9:2718–2728

Walsh CT (2016) Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. Nat Prod Rep 33:127–135

Winkelblech J, Fan A, Li S-M (2015) Prenyltransferases as key enzymes in primary and secondary metabolism. Appl Microbiol Biotechnol 99:7379–7397

Wohlgemuth V, Kindinger F, Xie X, Wang BG, Li S-M (2017) Two prenyltransferases govern a consecutive prenylation cascade in the biosynthesis of echinulin and neoechinulin. Org Lett 19:5928–5931

Xu W, Gavia DJ, Tang Y (2014) Biosynthesis of fungal indole alkaloids. Nat Prod Rep 31:1474–1487

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-9074-6

CORRECTION

Correction to: Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains

Jing Liu[1] · Huili Yu[1] · Shu-Ming Li[1]

Correction to: Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-8908-6

The original version of this article contained mistake. After careful re-examination of the LC-MS data, the products of CDPSs "WP_031028810" and "BAU83478" should be cFL instead of cPY. We apologize for any inconvenience that this may have caused.

Correct presentations given below:

The sixth sentence of the first paragraph in the **Discussion Section**:

"As shown in Fig. S10, WP_031028810 from *Streptomyces* sp. NRRL F-5639 and BAU83478 from *Streptomyces laurentii* DSM 41684 were proven to be responsible for the formation of *cyclo*-L-Pro-L-Tyr (cPY)."

should be changed to

"As shown in Fig. S10, WP_031028810 from *Streptomyces* sp. NRRL F-5639 and BAU83478 from *Streptomyces laurentii* DSM 41684 were proven to be responsible for the formation of *cyclo*-L-Phe-L-Leu (cFL)."

Jing Liu and Huili Yu contributed equally to this work.

The online version of the original article can be found at https://doi.org/10.1007/s00253-018-8908-6

✉ Shu-Ming Li
shuming.li@staff.uni-marburg.de

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

Published online: 16 May 2018

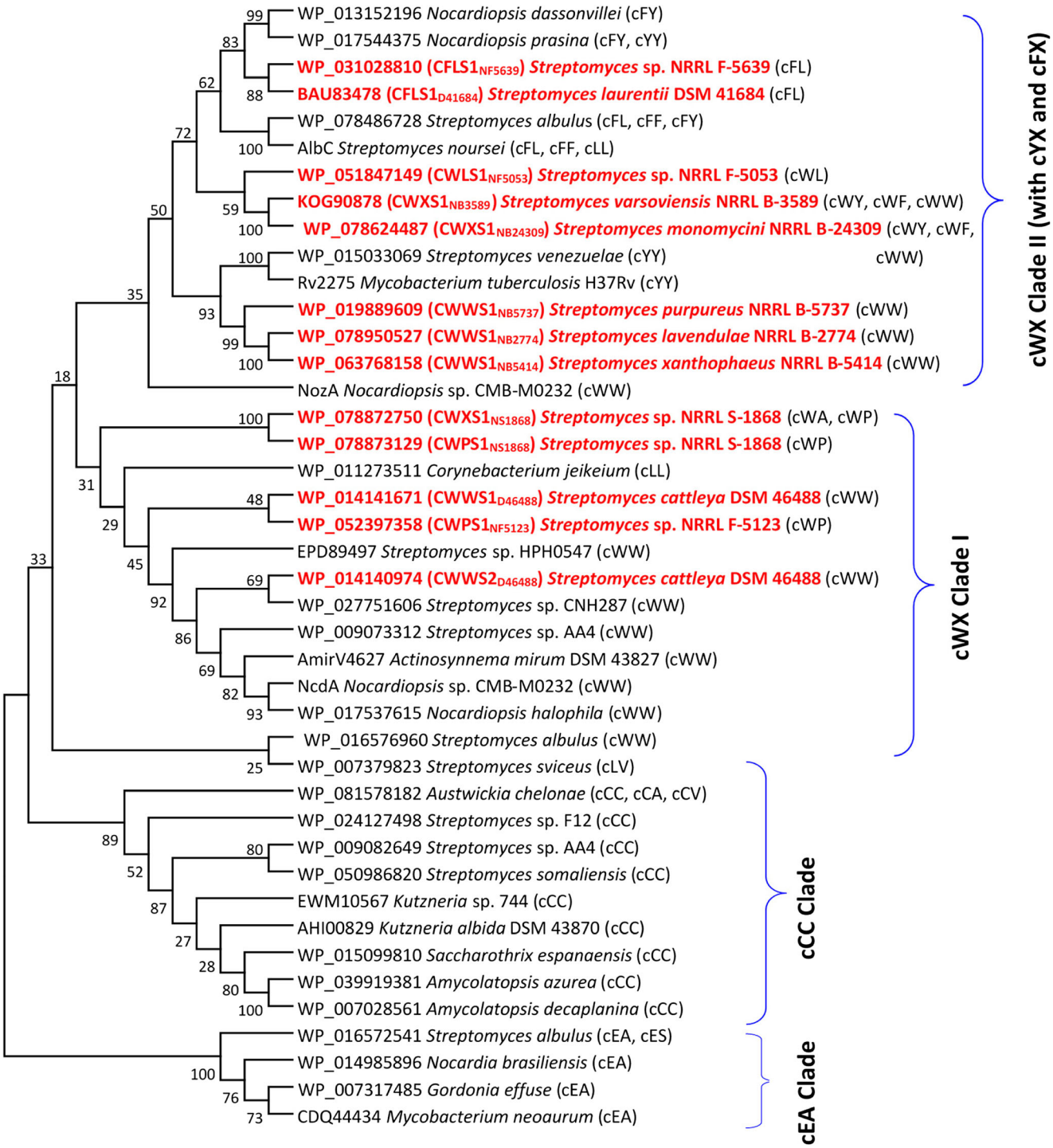


Fig. 1 Phylogenetic analysis of CDPSs from actinobacteria. The main CDPS products are given in parenthesis. Enzymes investigated in this study are highlighted in red

Table 1 Information on CDPSs identified in this study

CDPS	Accession number	Protein length (aa)	Primer sequences (5'-3')	Cloning constructs in pGEM-T easy	Cloning sites in pET28a	Expression constructs
$CWWS1_{D46488}$	WP_014141671	278	AACATATGCACGACAACGGTCATCGGCCC TTGGATCCTTACGTCACCGGCT	pHY02	*Nde*I *Bam*HI	pHY34
$CWWS2_{D46488}$	WP_014140974	234	GCCATATGCTGCATCGAACGTCCTT TAGGATCCTCACGGCTCGGCGGGCAGTT	pHY01	*Nde*I *Bam*HI	pHY11
$CWWS1_{NB2774}$	WP_078950527	265	CATATGACGATCACAGCTGACGCATCATTC GGATCCTCAAGCTGCTCGACGCTCAT	pHY72	*Nde*I *Bam*HI	pHY82
$CWWS1_{NB5737}$	WP_019889609	239	GCCATATGACTCTCATCGAAGACAC GGATCCTCAAGCGGCTCGACGGTCAT	pHY08	*Nde*I *Bam*HI	pHY09
$CWWS1_{NB5414}$	WP_063768158	273	CATATGAGGGCGATCACACAGGTGAC GGATCCTCAAGCTGCTCGACGCTCAT	pHY77	*Nde*I *Bam*HI	pHY83
$CWPS1_{NS1868}$	WP_078873129	241	CATATGAACACTTCCCTCGCTGC GAATTCTCAGCGTTCGGCCGCCCGGTC	pHY103	*Nde*I *Eco*RI	pHY107
$CWLS1_{NF5053}$	WP_051847149	243	CATATGTGCGAGGGCGCCGATGTGC GGATCCTCAGGATTCGTCCACCGG	pJL03	*Nde*I *Bam*HI	pJL26[a]
$CWPS1_{NF5123}$	WP_052397358	263	CATATGACCAGCAGGACCGAAAC GAATTCTCACGGAAGCAGCCGGGGG	pHY100	*Nde*I *Eco*RI	pHY105
$CWXS1_{NS1868}$	WP_078872750	267	CATATGGCCACACACGCCTCCGC GAATTCTCACTGCTGCGTCACGTGGTC	pHY101	*Nde*I *Eco*RI	pHY106
$CWXS1_{NB3589}$	KOG90878	255	CATATGGGGGCCCCGCAGCCC GGATCCTCACGTCAAGTCCCTTTCTCC	pJL02	*Nde*I *Bam*HI	pJL25
$CWXS1_{NB24309}$	WP_078624487	282	GCGCATATGAGTGCATCGCAGGCTGCTG GGATCCTCACGTCACGTCCCCTGC	pJL01	*Nde*I *Bam*HI	pJL24[b]
$CFLS1_{NF5639}$	WP_031028810	278	CATATGAACCGCCGCTGTTCCTTCG GGATCCTCATGCTCCTGGGGCACTG	pJL04	*Nde*I *Bam*HI	pJL29
$CFLS1_{D41684}$	BAU83478	224	CGCATATGAACCAGTTCGACGTGCTGCC GGATCCCCGGACCACGAGGAAGCC	pJL05	*Nde*I *Bam*HI	pJL30[c]

Restriction sites for cloning are underlined in primer sequences

a The 12 base pairs at the 5′-end of the sequence coding for WP_051847149 were not included in the expression construct pJL26

b The 96 base pairs at the 5′-end of the sequence coding for WP_078624487 were not included in the expression construct pJL24

c According to the annotation in database, BAU83478 comprises 484 amino acids and consists of CDPS sequence at its N-terminus and that for a putative methyltransferase at its C-terminus. The sequence coding for amino acids 1 - 224 of BAU83478 was cloned into pET28a generating pJL30

Applied Microbiology and Biotechnology

Supplementary Materials

Expanding Tryptophan-containing Cyclodipeptide Synthase Spectrum by Identification of Nine Members from *Streptomyces* Strains

Jing Liu[1], Huili Yu[1], and Shu-Ming Li[1]

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

*Corresponding author

Email: shuming.li@staff.uni-marburg.de

Tel/Fax:+49 6421 2822461/+49 6421 2825365

Content

Table S1 *Streptomyces* strains used in this study

Strains	Source	Cultivation Media
Streptomyces cattleya DSM 46488	DSMZ	GYM
Streptomyces lavendulae NRRL B-2774	NRRL	GYM
Streptomyces purpureus NRRL B-5737	NRRL	ISP4
Streptomyces xanthophaeus NRRL B-5414	NRRL	ISP4
Streptomyces sp. NRRL S-1868	NRRL	ISP4
Streptomyces sp. NRRL F-5053	NRRL	ISP4
Streptomyces sp. NRRL F-5123	NRRL	GYM
Streptomyces varsoviensis NRRL B-3589	NRRL	GYM
Streptomyces monomycini NRRL B-24309	NRRL	GYM
Streptomyces sp. NRRL F-5639	NRRL	ISP4
Streptomyces laurentii DSM 41684	DSMZ	GYM

DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen

NRRL: ARS Culture Collection

Table S2 The yields (mg/L culture) of the main products of the 11 CDPSs under six different conditions

CDPSs / Temperature -IPTG final conc.	$CWWS1_{D46488}$ (cWW)	$CWWS2_{D46488}$ (cWW)	$CWWS1_{NB2774}$ (cWW)	$CWWS1_{NB5737}$ (cWW)	$CWWS1_{NB5414}$ (cWW)
20 ℃ - 0.2 mM	77.0	33.7	69.2	18.8	55.1
20 ℃ - 0.5 mM	88.8	33.7	85.8	24.4	48.8
30 ℃ - 0.2 mM	88.2	59.1	95.8	41.2	34.3
30 ℃ - 0.5 mM	95.5	51.8	98.6	63.7	34.5
37 ℃ - 0.2 mM	81.9	19.5	82.6	9.0	3.4
37 ℃ - 0.5 mM	89.7	28.5	65.0	11.3	-

CDPSs / Temperature -IPTG final conc.	$CWPS1_{NS1868}$ (cWP)	$CWPS1_{NF5123}$ (cWP)	$CWXS1_{NS1868}$ (cWA)	$CWLS1_{NF5053}$ (cWL)	$CWXS1_{NB3589}$ (cWY)	$CWXS1_{NB24309}$ (cWY)
20 ℃ - 0.2 mM	144.9	69.2	69.3	93.9	141.4	51.7
20 ℃ - 0.5 mM	173.8	85.8	75.5	96.4	111.8	64.7
30 ℃ - 0.2 mM	131.4	95.8	68.1	104.9	25.8	55.8
30 ℃ - 0.5 mM	198.2	98.6	74.7	104.2	44.9	54.9
37 ℃ - 0.2 mM	146.3	82.6	14.7	104.9	94.6	15.5
37 ℃ - 0.5 mM	128.4	65.0	12.1	84.7	106.1	13.4

Table S3 HR-EI-MS data of the CDPS products

Compound	Chemical Formula	$[M+H]^+$		Deviation (ppm)
		Calculated	Measured	
cWA	$C_{14}H_{15}N_3O_2$	258.1237	258.1224	5.0
cWY	$C_{20}H_{19}N_3O_3$	350.1499	350.1495	1.1
cWM	$C_{16}H_{19}N_3O_2S$	318.1271	318.1289	-5.7
cWV	$C_{16}H_{19}N_3O_2$	286.1550	286.1536	4.9
cWW	$C_{22}H_{20}N_4O_2$	373.1659	373.1636	6.2
cWF	$C_{20}H_{19}N_3O_2$	334.1550	334.1545	1.5
cWL	$C_{17}H_{21}N_3O_2$	300.1707	300.1709	-0.7
cWP	$C_{16}H_{17}N_3O_2$	284.1394	284.1417	-8.1
cFL	$C_{15}H_{20}N_2O_2$	261.1598	261.1619	-8.0

Table S4 ^{1}H NMR data of the isolated cyclodipeptides

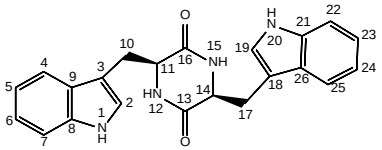
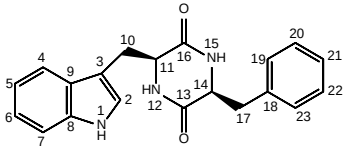
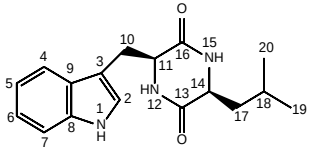
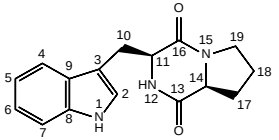
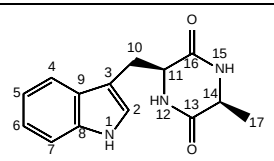
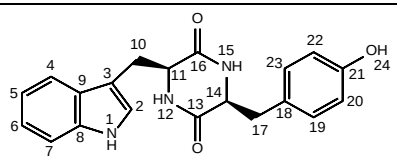
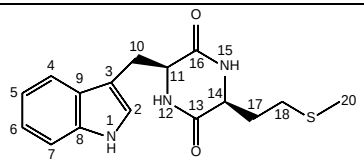
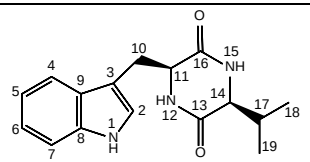
Comp.	 *cyclo*-(L-Trp-L-Trp) (DMSO-d6)	 *cyclo*-(L-Trp-L-Phe) (DMSO-d6)	 *cyclo*-(L-Trp-L-Leu) ($CDCl_3$)	 *cyclo*-(L-Trp-L-Pro) ($CDCl_3$)
Pos.	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]
1	10.81, s	10.90, s	8.15, s	8.55, s
2	6.61, d, 2.3	7.00, d, 2.3	7.11, d, 2.3	7.06, d, 2.1
4	7.35, d, 8.0	7.52, dd, 8.0, 1.1	7.63, dd, 8.0, 1.0	7.58, dd, 8.0, 1.0
5	6.95, ddd, 8.0, 7.0, 1.1	7.02, ddd, 8.0, 7.0, 1.1	7.15, ddd, 8.0, 7.1, 1.0	7.13, ddd, 8.0, 7.0, 1.0
6	7.04, ddd, 8.1, 7.0, 1.1	7.10, ddd, 8.1, 7.0, 1.1	7.23, ddd, 8.2, 7.1, 1.0	7.22, ddd, 8.1, 7.0, 1.1
7	7.28, dt, 8.1, 1.1	7.35, dt, 8.1, 1.1	7.39, dt, 8.2, 1.0	7.38, dt, 8.1, 1.1
10	2.71, dd, 14.3, 4.2	2.84, dd, 14.5, 5.9	3.49, ddd, 14,7, 3.1, 0.7	3.75, ddd, 15.1, 3.8, 1.0
	2.20, dd, 14.3, 6.7	2.57, m	3.21, dd, 14.7, 8.4	2.97, dd, 15.1, 10.7
11	3.87, m	4.00, m	4.32, m	4.37, ddd, 10.7, 3.8, 1.0
12	7.66, d, 2.5	7.90, d, 2.4	5.88, s	5.83, s
14	3.87, m	3.89, m	3.89, m	4.06, bt, 7.4
15	7.66, d, 2.5	7.71, d, 2.5	5.86, s	-
17	2.71, dd, 14.3, 4.2	2.50, m	1.51, m	2.31, m
	2.20, dd, 14.3, 6.7	1.92, dd, 13.6, 7.1	1.48, m	2.01, m
18	-	-	0.98, m	1.96, m
	-	-	-	1.89, m
19	6.61, d, 2.3	6.76, dd, 8.0, 2.0	0.84, d, 6.5	3.65, m
	-	-		3.58, m
20	10.81, s	7.20[a]	0.81, d, 6.5	-
21	-	7.20[a]	-	-
22	7.28, dt, 8.1, 1.1	7.20[a]	-	-
23	7.04, ddd, 8.1, 7.0, 1.1	6.76, dd, 8.0, 2.0	-	-
24	6.95, ddd, 8.0, 7.0, 1.1	-	-	-
25	7.35, d, 8.0	-	-	-

Table S4 (continued)

Comp.	 *cyclo*-(L-Trp-L-Ala) (CD_3OD)	 *cyclo*-(L-Trp-L-Tyr) (DMSO-d6)	 *cyclo*-(L-Trp-L-Met) (CD_3OD)	 *cyclo*-(L-Trp-L-Val) (DMSO-d6)
Pos.	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]	δ_H, multi., *J* [Hz]
1	-	10.90, d, 1.5	-	10.85, s
2	7.05, s	7.00, d, 1.5	7.04, s	7.11, d, 2.4
4	7.58, dt, 8.0, 1.0	7.51, dd, 7.9, 1.0	7.59, dt, 8.0, 1.0	7.63, d, 7.9
5	6.97, ddd, 8.0, 7.1, 1.0	7.02, ddd, 7.9, 7.0, 1.0	6.96, ddd, 8.0, 7.0, 1.0	6.97, ddd, 7.9, 7.1, 1.0
6	7.05, ddd, 8.1, 7.1, 1.0	7.09, ddd, 8.1, 7.0, 1.0	7.05, ddd, 8.1, 7.0, 1.1	7.05, ddd, 8.1, 7.1, 1.0
7	7.30, dt, 8.1, 1.0	7.34, dt, 8.1, 1.0	7.30, dt, 8.1, 1.1	7.32, d, 8.1
10	3.44, dd, 14.6, 4.1	2.83, dd, 14.5, 4.2	3.47, dd, 14.7, 3.5	3.24, dd, 14.5, 5.0
	3.12, dd, 14.6, 4.5	2.49[b]	3.09, dd, 14.7, 4.6	3.11, dd, 14.5, 4.5
11	4.25, ddd, 4.5, 4.1, 1.3	3.97, m	4.27, ddd, 4.6, 3.5, 1.4	4.17, m
12	-	7.80, d, 2.5	-	7.96, s
14	3.67, qd, 7.1, 1.2	3.81, m	3.77, ddd, 7.4, 4.7, 1.3	3.52, m
15	-	7.62, d, 2.6	-	7.87, s
17	0.32, d, 7.1	2.45[b]	1.17, m	1.68, m
	-	1.86, dd, 13.6, 4.5	0.60, m	-
18	-	-	1.63, m	0.64, d, 7.1
19	-	6.61, dt, 8.6, 2.0	-	0.25, d, 6.8
20	-	6.56, dt, 8.6, 2.0	1.73, s	-
22	-	6.56, dt, 8.6, 2.0	-	-
23	-	6.61, dt, 8.6, 2.0	-	-
24	-	9.18, s	-	

[a, b]: Signals with the same letter overlapping with each other

The data correspond well to those reported for cWA (Caballero et al. 1998), cWY (Kumar et al. 2014), cWV (He et al. 2013), cWL (Kumar et al. 2014), cWW (Giessen et al. 2013a; Lu et al. 2017), cWF (Kumar et al. 2014), and cWP(Grundmann and Li 2005), respectively.

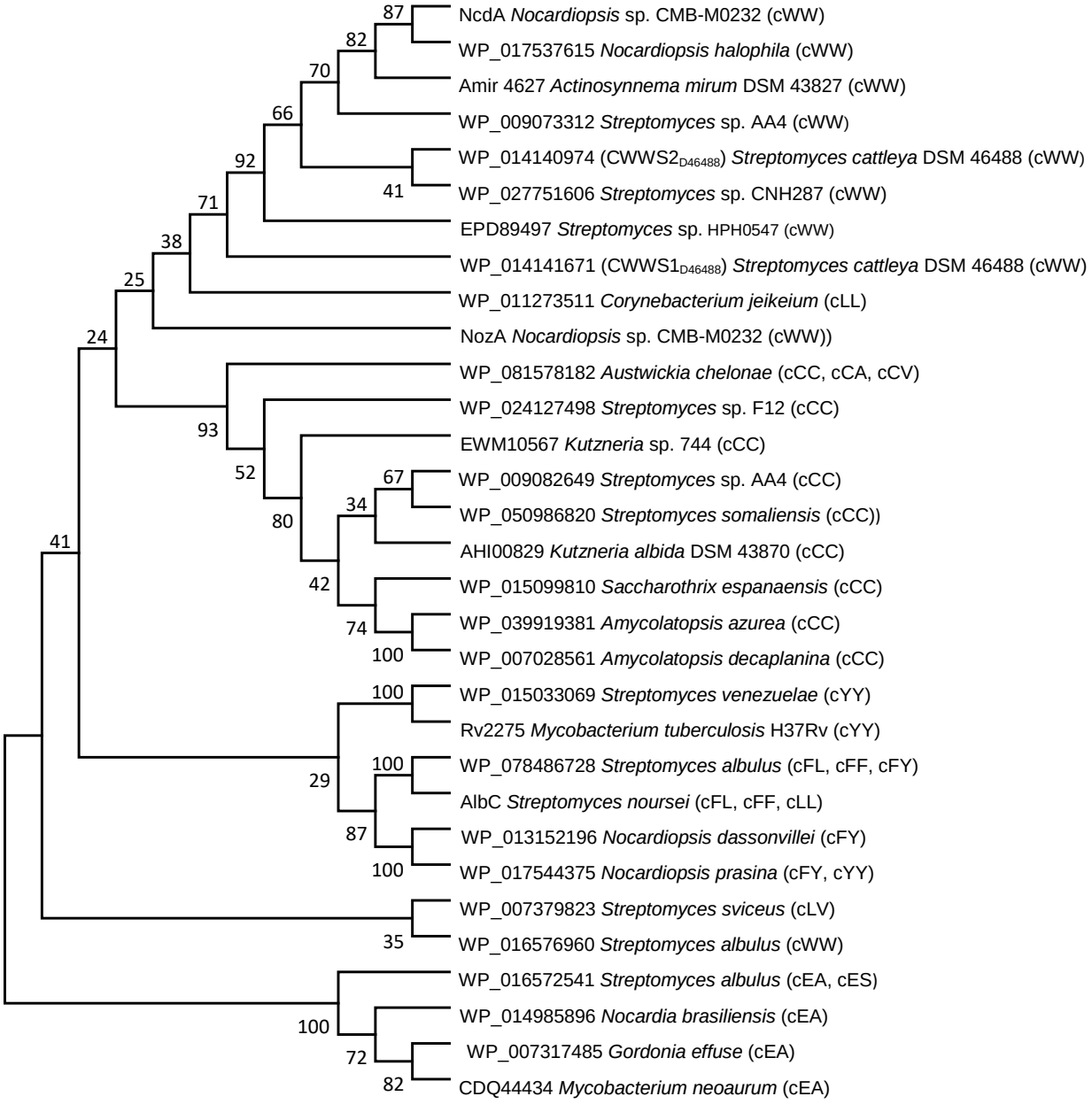


Fig. S1 Phylogenetic tree of known CDPSs from actinobacteria prior to this study (Brockmeyer and Li 2017; Giessen et al. 2013a; Giessen et al. 2013b; Jacques et al. 2015; James et al. 2015)

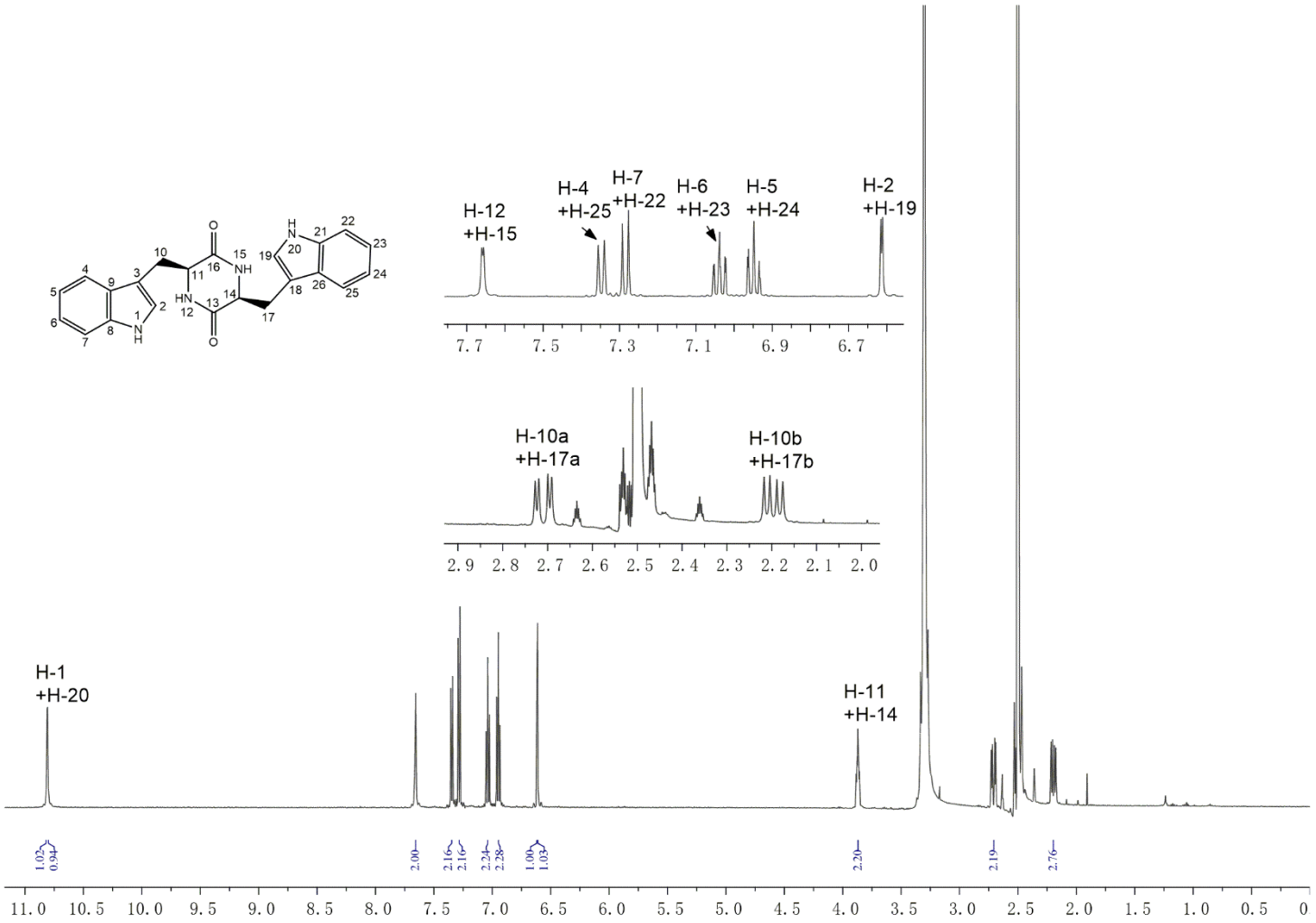


Fig. S2 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Trp) in DMSO-d6

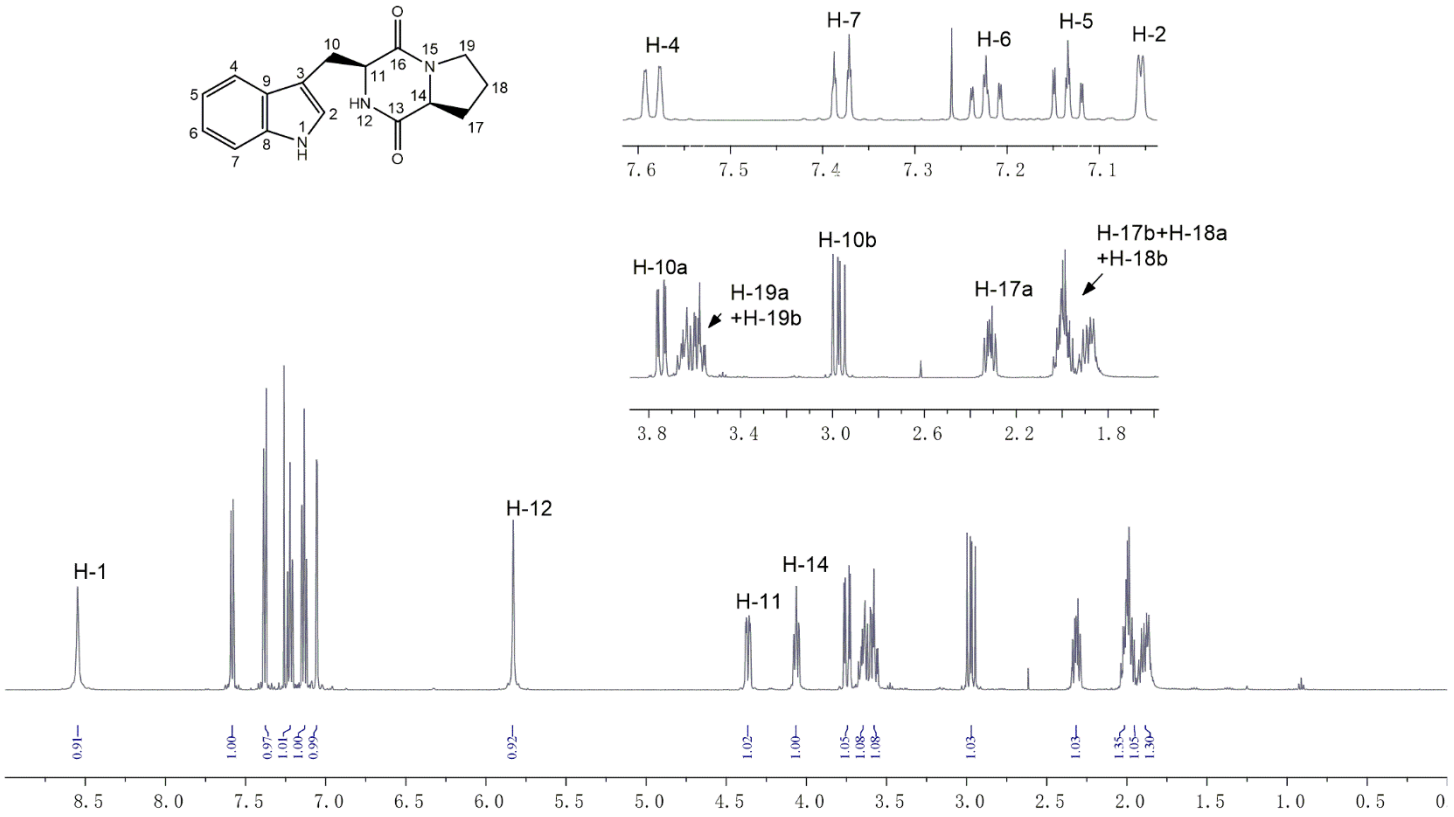


Fig. S3 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Pro) in $CDCl_3$

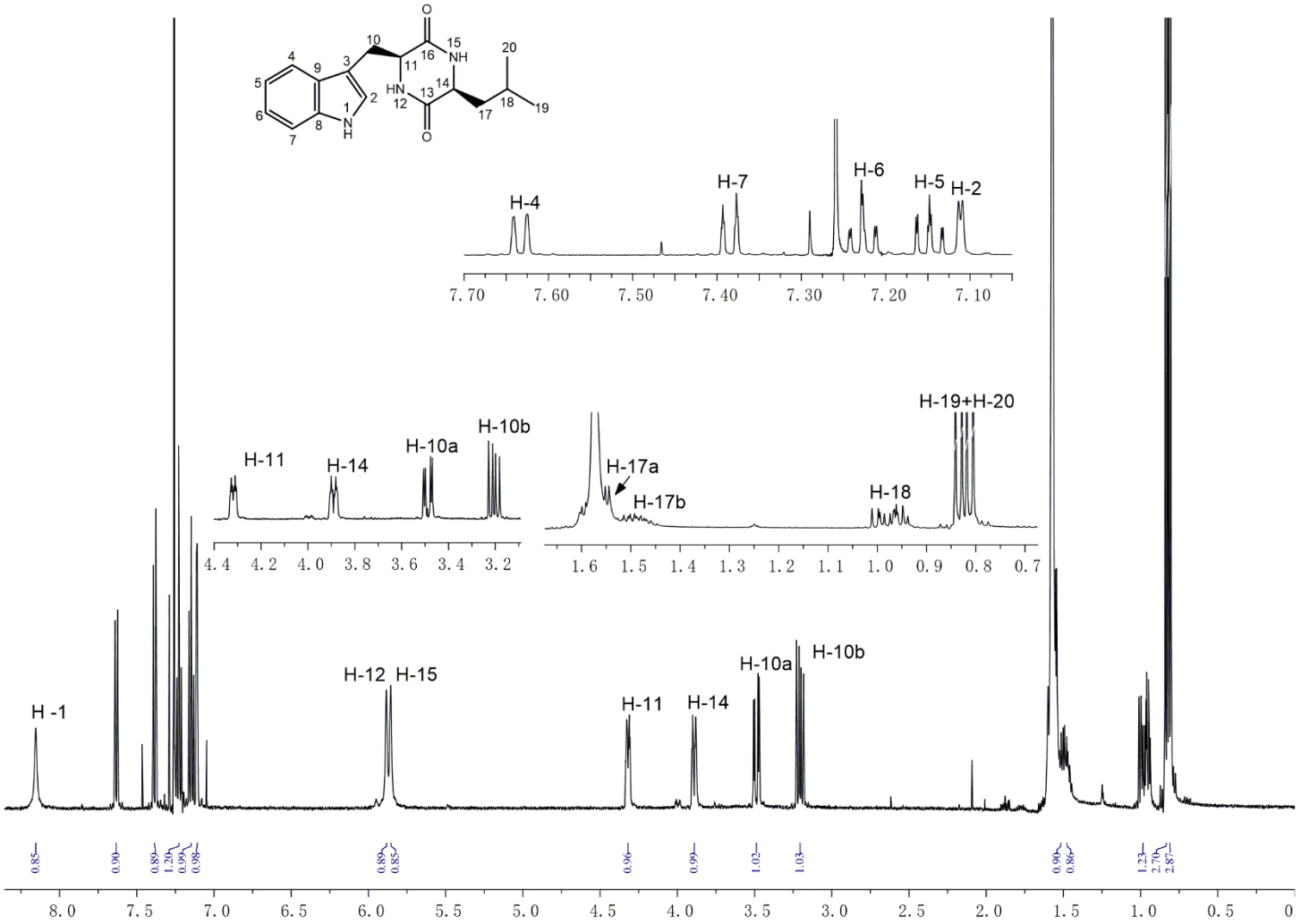


Fig. S4 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Leu) in $CDCl_3$

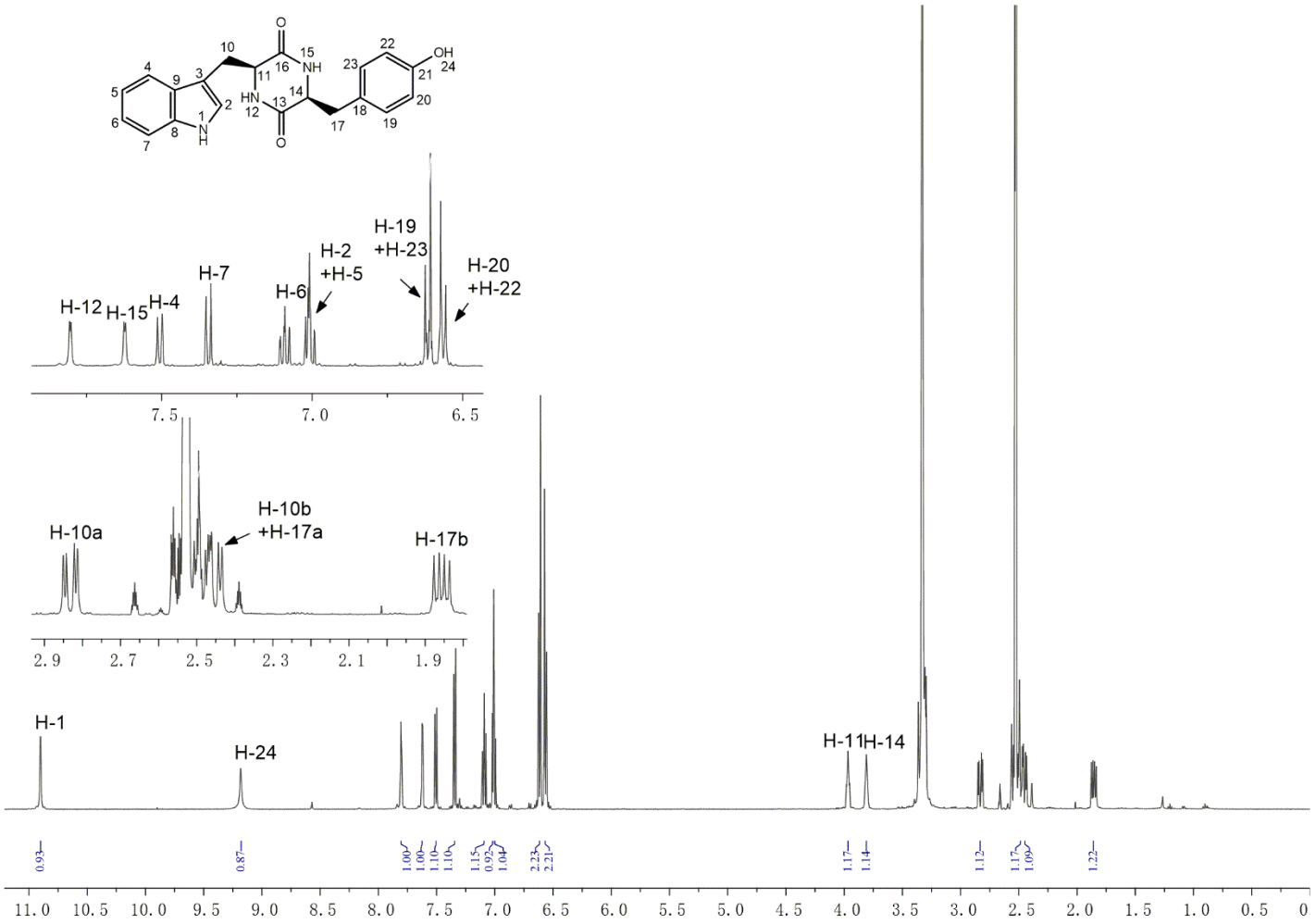


Fig. S5 ^{1}H NMR spectrum of in *cyclo*-(L-Trp-L-Tyr) in DMSO-d6

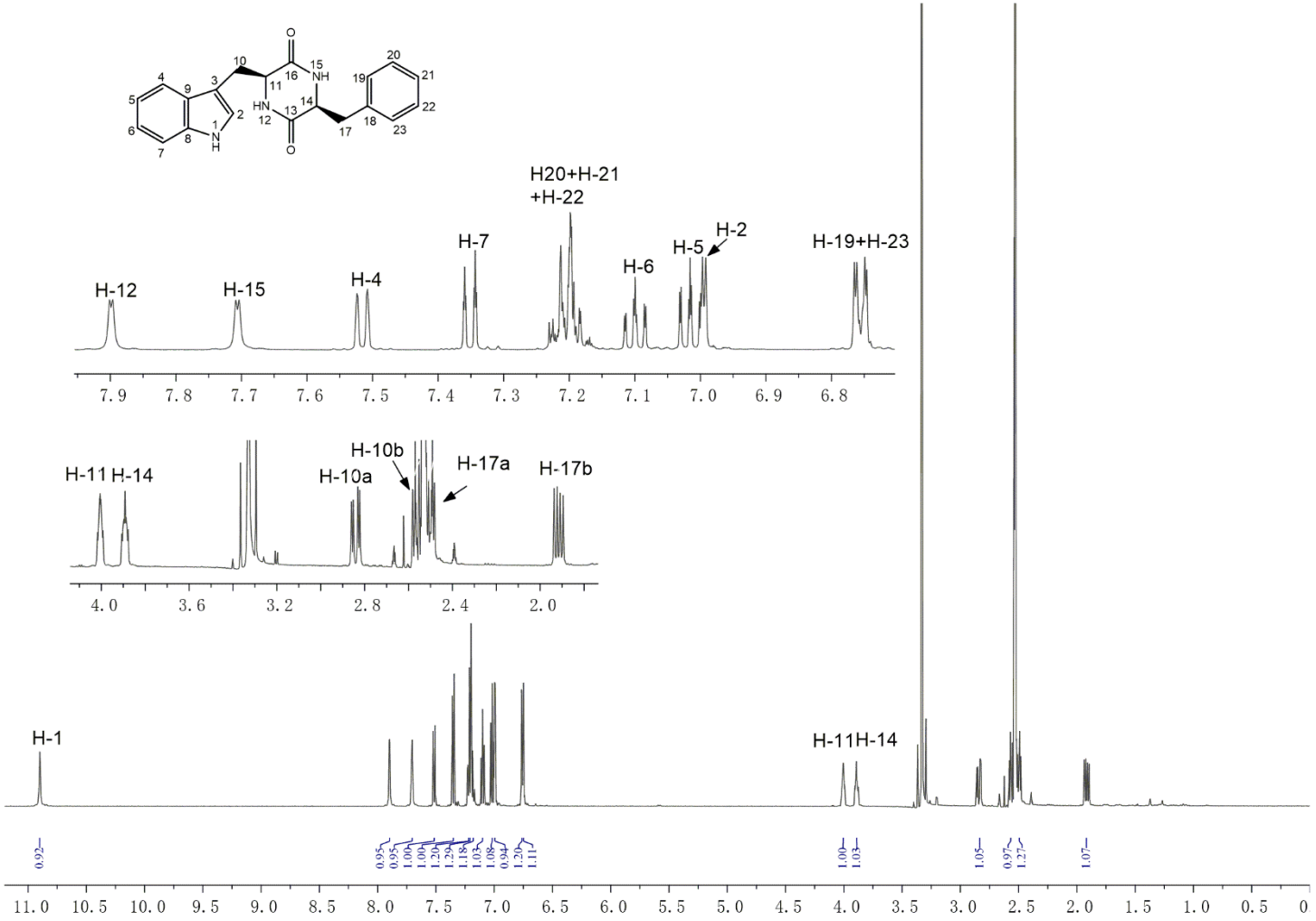


Fig. S6 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Phe) in DMSO-d6

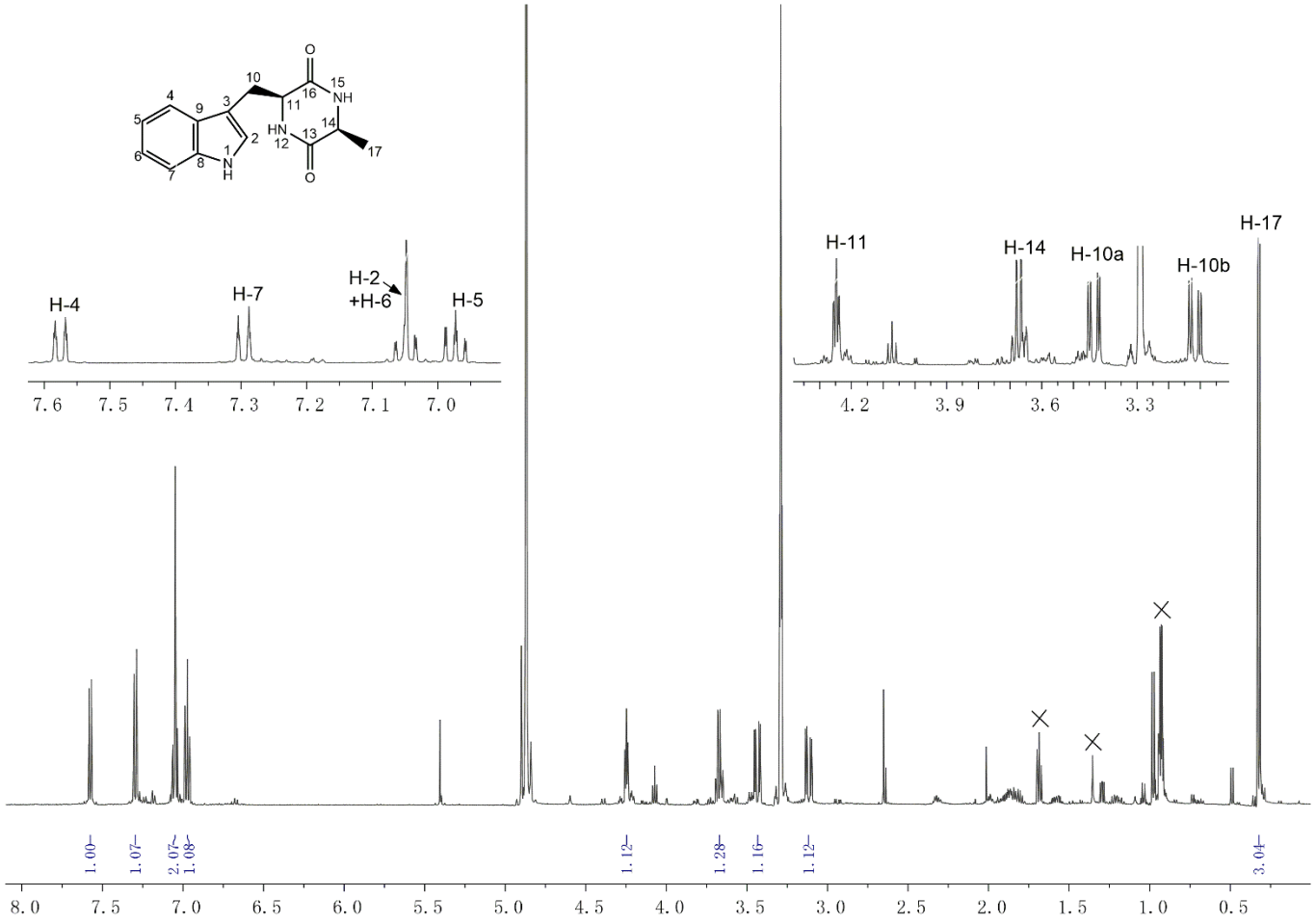


Fig. S7 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Ala) in CD_3OD

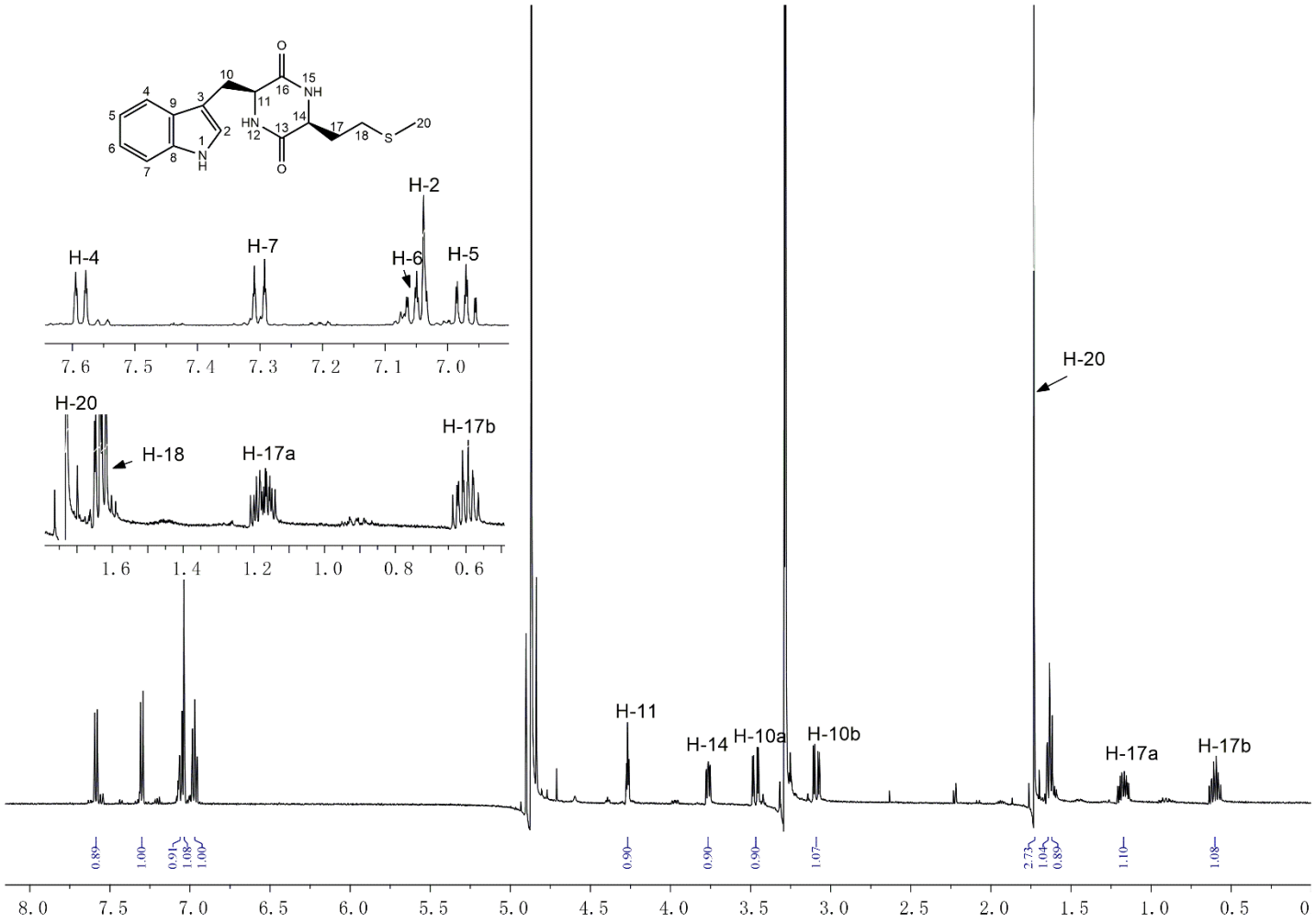


Fig. S8 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Met) in CD_3OD

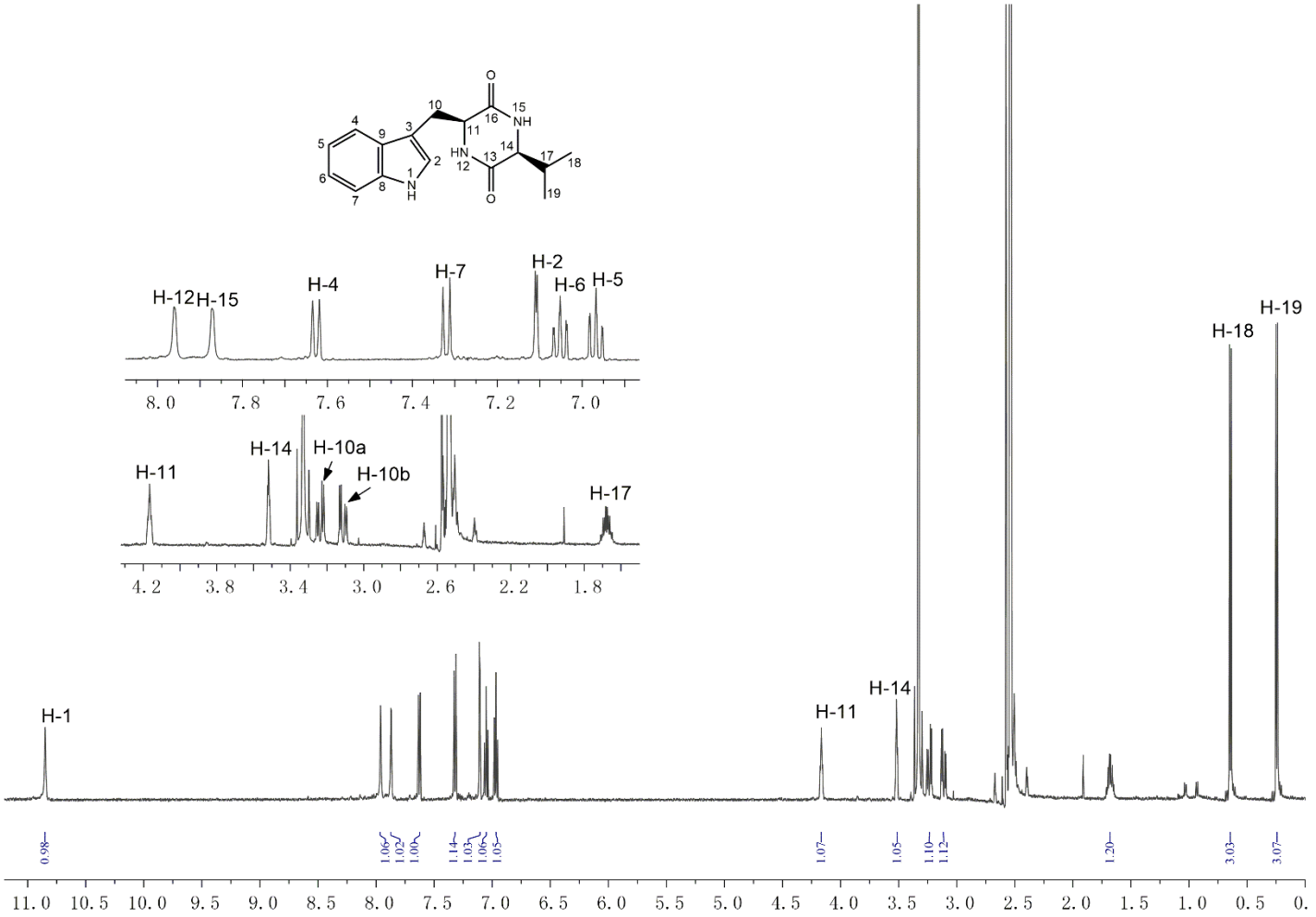


Fig. S9 ^{1}H NMR spectrum of *cyclo*-(L-Trp-L-Val) in DMSO-d6

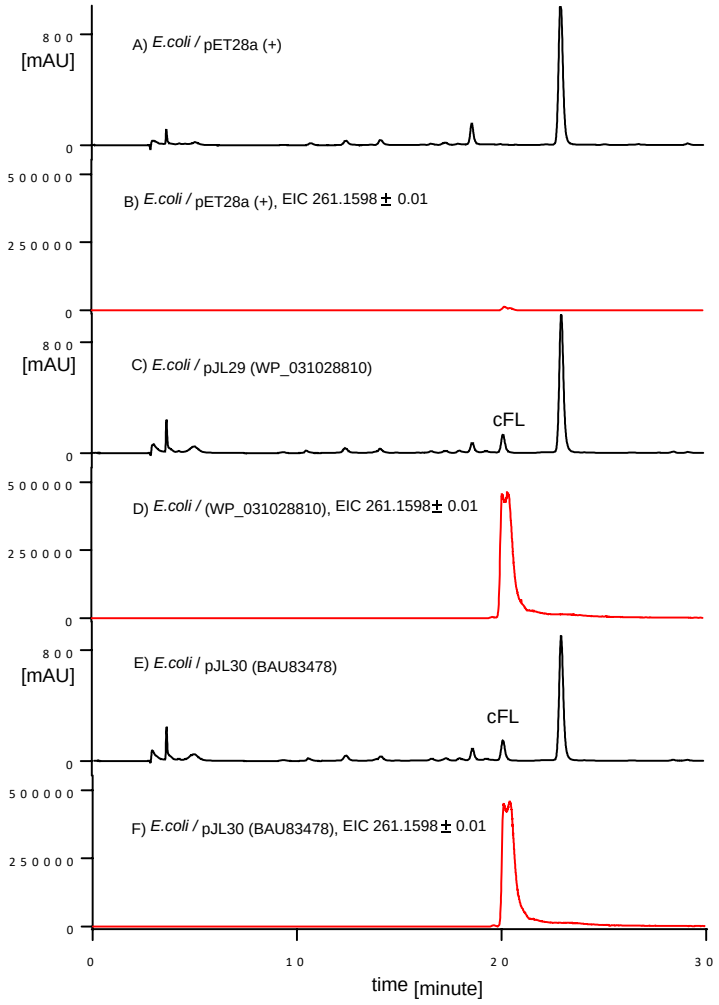


Fig. S10 LC-MS analysis of recombinant *E. coli* strains with empty vector (A and B) and expression constructs (C-F). Black lines are UV absorptions at 258 nm and the red lines refer to extracted positive ion chromatograms (EIC).

References

Brockmeyer K, Li S-M (2017) Mutations of residues in pocket P1 of a cyclodipeptide synthase strongly increase product formation. J Nat Prod 80:2917-2922

Caballero E, Avendaño C, Menéndez JC (1998) Stereochemical issues related to the synthesis and reactivity of pyrazino[2',1'-5,1]pyrrolo[2,3-b]indole-1,4-diones. Tetrahedron: Asymmetry 9:967-981

Giessen TW, von Tesmar AM, Marahiel MA (2013a) A tRNA-dependent two-enzyme pathway for the generation of singly and doubly methylated ditryptophan 2,5-diketopiperazines. Biochemistry 52:4274-4283

Giessen TW, von Tesmar AM, Marahiel MA (2013b) Insights into the generation of structural diversity in a tRNA-dependent pathway for highly modified bioactive cyclic dipeptides. Chem Biol 20:828-838

Grundmann A, Li S-M (2005) Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from *Aspergillus fumigatus*. Microbiology 151:2199-2207

He R, Wang B, Wakimoto T, Wang M, Zhu L, Abe I (2013) Cyclodipeptides from metagenomic library of a Japanese marine sponge. J Braz Chem Soc 24:1926-1932

Jacques IB, Moutiez M, Witwinowski J, Darbon E, Martel C, Seguin J, Favry E, Thai R, Lecoq A, Dubois S, Pernodet JL, Gondry M, Belin P (2015) Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity. Nat Chem Biol 11:721-727

James ED, Knuckley B, Alqahtani N, Porwal S, Ban J, Karty JA, Viswanathan R, Lane AL (2015) Two distinct cyclodipeptide synthases from a marine actinomycete catalyze biosynthesis of the same diketopiperazine natural product. ACS Synth Biol 5:547-553

Kumar SN, Mohandas C, Nambisan B (2014) Purification, structural elucidation and bioactivity of tryptophan containing diketopiperazines, from *Comamonas testosteroni* associated with a rhabditid entomopathogenic nematode against major human-pathogenic bacteria. Peptides 53:48-58

Lu C, Xie F, Shan C, Shen Y (2017) Two novel cyclic hexapeptides from the genetically engineered *Actinosynnema pretiosum*. Appl Microbiol Biotechnol 101:2273-2279

4.2 Coupling of guanine with *cyclo*-L-Trp-L-Trp mediated by a cytochrome P450 homologue from *Streptomyces purpureus*.

Organic LETTERS
Letter
Cite This: Org. Lett. 2018, 20, 4921−4925
pubs.acs.org/OrgLett

Coupling of Guanine with *cyclo*-L-Trp-L-Trp Mediated by a Cytochrome P450 Homologue from *Streptomyces purpureus*

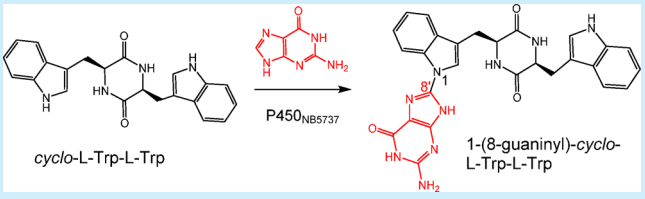
Huili Yu,[†] Xiulan Xie,[‡] and Shu-Ming Li*[,†]

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

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

Supporting Information

ABSTRACT: A cyclo-L-Trp-L-Trp tailoring P450 with novel function from *Streptomyces purpureus* was identified by heterologous expression in *S. coelicolor* and in vitro assays the recombinant protein. Structural elucidation revealed that this enzyme catalyzes the transfer of a guanine moiety to the indole ring of the cyclodipeptide via a C−N bond. Adduct products of CDP and guanine are unprecedented in nature, and CDP modification by coupling with guanine has not been reported prior to this study.



2,5-Diketopiperazine (DKP) scaffold-containing natural products have gained increasing attention in recent years for their diverse and interesting biological and pharmacological activities ranging from antibacterial,[1,2] antifungal,[3,4] antitumor,[5,6] and immunosuppressive[7] to anti-inflammatory[8] activities. These compounds were prepared by chemical synthesis[9,10] or isolated from natural sources. In nature, cyclodipeptides (CDPs) are biosynthesized by large multimodular nonribosomal peptide synthetases (NRPSs) or small cyclodipeptide synthases (CDPSs).[11−13] CDPSs with typical polypeptide chains of 200−300 residues hijack aminoacyl-tRNAs from the ribosomal protein synthetic machinery and use them as substrates to form CDPs.[14] The CDPSs are genetically associated with a variety of tailoring enzymes, such as cytochromes P450 (P450s), methyltransferases, and cyclodipeptide oxidases, which has drawn remarkable attention for identification of new modified CDP derivatives and pathways.[11] Tryptophan-containing cyclodipeptides (cWXs) are the richest precursors of DKPs with pharmaceutical interest due to the various modification possibilities on the indole ring.[9,15] Very recently, several CDPSs that are responsible for the formation of cWXs have been identified in different *Streptomyces* strains, including cWW synthases in *S. purpureus*, *S. lavendulae*, and *S. xanthophaeus*.[16−18] In these strains, the cWW-forming CDPS gene is located directly upstream of a putative cytochrome P450 gene with the same orientation. Both genes are flanked by a putative calcium-binding and YncE family protein (Figure 1). The three P450 homologues share sequence identities of more than 75% on the amino acid level with each other (Table S1). It could be therefore speculated that these three clusters encode biosynthetic pathways of very similar DKP derivatives.

P450s constitute a superfamily of ubiquitous hemoproteins that catalyze a large variety of reactions via a complex multistep mechanism.[19] They are the major biocatalysts in the biosynthesis of steroids and bile acids as well as in the metabolism of drugs and xenobiotics.[20,21] With extraordinarily broad substrate spectra and regio- and stereoselectivity, they also play key roles in the biosynthesis of natural products.[19,22] In addition to typical P450 reactions such as hydroxylation and epoxidation, they also catalyze unusual reactions such as aryl−aryl coupling, bond cleavage or migration and ring opening, nitration, or rearrangement.[23,24] Until now, several CDP tailoring P450s have been functionally characterized in the biosynthesis of mycocyclosin,[25] pulcherriminic acid,[26] and (−)-ditryptophenaline.[27] They catalyze intramolecular C−C bond formation of cYY, double N-oxide formation and aromatization of the DKP ring in cLL, and dimerization of cWF, respectively.[25−27]

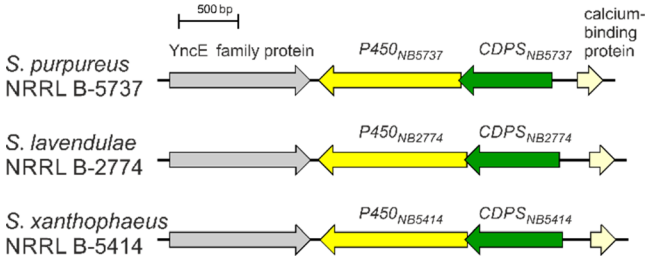


Figure 1. *CDPS*- and *P450*-containing gene clusters from three *Streptomyces* strains.

After identification of cWW as the sole product of the three CDPSs deposited in Figure 1, we were interested in the function of the P450 homologues in these clusters, termed in this study $P450_{NB5737}$, $P450_{NB2774}$, and $P450_{NB5414}$, respectively.

Received: June 30, 2018
Published: August 3, 2018

DOI: 10.1021/acs.orglett.8b02051
Org. Lett. 2018, 20, 4921−4925

Alignments of their sequences with those of known P450s from bacteria revealed the presence of the three P450 featuring motifs (Figure S1): (i) the highly conserved A/G^{241}G^{242}XXT^{245} motif with Thr as a critical residue for the H-bonding network[22] (this residue is replaced by a Ser in the CDP tailoring P450s),[25] (ii) E^{280}XXR^{283} motif found in the K helix with the two invariant Glu and Arg to ensure correct tertiary structure and stabilization of the heme within the protein,[19] and (iii) G^{347}XXXC^{351} motif in the heme-binding loop harboring the invariant and catalytically most important Cys.[22] Phylogenetic analysis of functionally characterized P450s from bacteria (Figure S2) indicated that the CDP tailoring P450s of interest distribute in the same subclade as CYP121 from *Mycobacterium tuberculosis*, which catalyzes the intramolecular aryl−aryl coupling of cYY.[25] They are also closely related to P-450mel, CYP158A1, and CYP158A2, which catalyze dimerization of hydroxynaphthalenes/naphthoquinones via aryl−aryl coupling in *Streptomyces*.[28−30] Based on these bioinformatics results, we speculated initially that P450$_{NB5737}$, P450$_{NB2774}$, and P450$_{NB5414}$ could also be responsible for intramolecular or intermolecular aryl−aryl coupling of cWW.

For functional proof, we chose the cluster from *S. purpureus* NRRL B5737 for expression. First, *CDPS*$_{NB5737}$ was cloned alone or together with *P450*$_{NB5737}$ into the integrative vector pSET152[31] (Table S2) and expressed in *S. coelicolor* M1146.[32] The transformants were cultivated in GYM liquid media at 28 °C for 7 days. The ethyl acetate extracts of the cell cultures were subjected to LC−MS analysis (Figure 2). In comparison to that of the wild-type strain M1146 (Figure 2i), *CDPS*$_{NB5737}$ transformant showed the sole product peak with the expected $[M + H]^+$ ion for cWW (**1**) at *m*/*z* 373.1659 (Figure 2ii), which is consistent with the results obtained from *E. coli* transformants (Scheme 1).[16] In the *CDPS*$_{NB5737}$ and *P450*$_{NB5737}$ coexpressing strain, one additional product peak **2** was detected by UV 277 nm at 18.5 min, and two additional product peaks for **2** and **2′** were detected by MS with $[M + H]^+$ ions at *m*/*z* 522.1997 (Figure 2iii). Expression of *CDPS*$_{NB5737}$ alone or together with *P450*$_{NB5737}$ in *S. coelicolor* M1146 with the replicative vector pPWW50[33] led to nearly identical results (Figure 2iv,v).

To confirm the metabolism of **1** by P450$_{NB5737}$, we cloned its coding sequence in pPWW50 and expressed it in *S. coelicolor* M1146. As shown in Figure 2vii, neither **1** nor **2** or **2′** was detected in the transformant with *P450*$_{NB5737}$. A 2-day-old *P450*$_{NB5737}$ transformant was then fed with 80 μM of **1** and cultivated further for 7 days. LC−MS analysis indicated that **1** was completely consumed, and **2** and **2′** with a similar ratio as in the coexpression transformant of *CDPS*$_{NB5737}$ and *P450*$_{NB5737}$ were detected as products (Figure 2iii,viii). In comparison, feeding **1** into M1146 did not lead to any conversion to **2** or **2′** (Figure 2vi). These results prove that cWW produced by CDPS$_{NB5737}$ was metabolized by P450$_{NB5737}$ to adduct products with an unknown reaction partner rather than to intra- or intermolecular coupling products within one or between two cWW molecules. This was deduced from the $[M + H]^+$ ion of **1** at *m*/*z* 373.1659 and those of **2** and **2′** at *m*/*z* 522.1997.

For structure elucidation, approximate 13 mg of **2** was isolated from 4 L 7 day-old culture of M1146 transformant harboring pHY48 for expression of *CDPS*$_{NB5737}$ and *P450*$_{NB5737}$ in GYM and subjected to NMR analysis including ^{1}H, ^{13}C APT, DQF-COSY, HSQC, and HMBC (Table S3, Figures S3−S8). Interpretation of the NMR spectra revealed that a guanine moiety was attached via its C8′ to N1 of **1**, resulting in the downfield shifting of H2 to δ 7.75 and H7 to δ 8.63 ppm, respectively. In comparison to those of **1**, five additional ^{13}C signals were detected at δ 156.8 (C2′), 150.0 (C4′), 116.4 (C5′), 161.1 (C6′), and δ 153.1 (C8′), which can be assigned to those of a guanine residue (Table S3).[34,35] The key correlation between H2 and C8′ in HMBC spectrum confirmed the connection of the two moieties via N1−C8′ (Scheme 1). The production yield of **2** of the 7 day-old culture was calculated to be 8 mg/L. Due to the low quantity, the

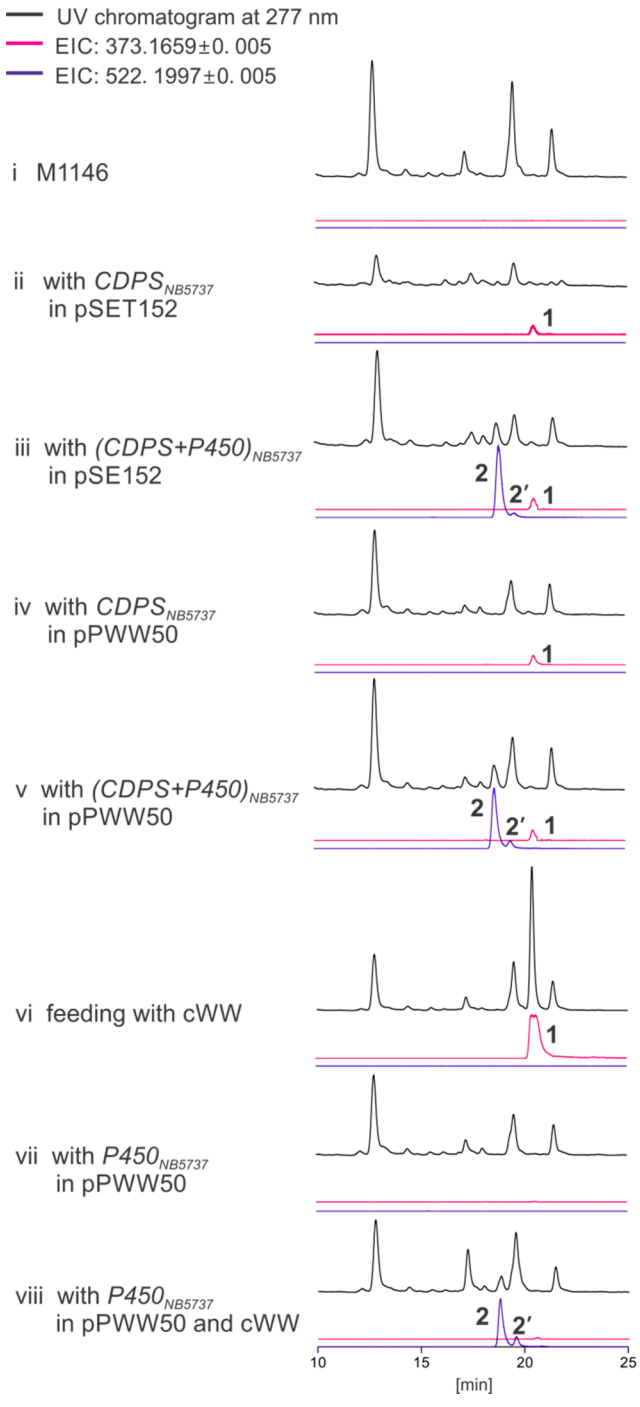
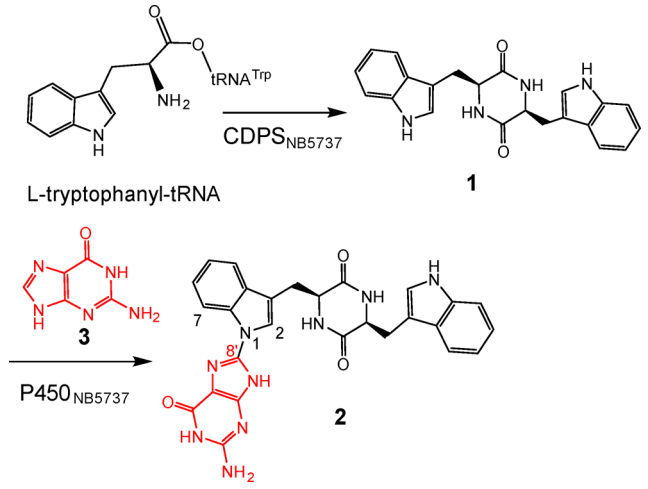


Figure 2. LC−MS analysis of the expression of *CDPS* and *P450* from *S. purpureus* in *S. coelicolor* M1146. Detection was carried out at UV 277 nm and EIC for $[M + H]^+$ ions of **1** and **2**.

DOI: 10.1021/acs.orglett.8b02051
Org. Lett. 2018, 20, 4921−4925

Scheme 1. Biosynthesis and Metabolism of cWW in *S. purpureus* NRRL B-5737



structure of **2′** cannot be determined in this study. From the HR-EIMS data, it could be, however, speculated that **2′** is also a coupling product of guanine with **1**.

Taking the structure information on **2** and the results obtained from heterologous expression and biotransformation into consideration, we postulated that $P450_{NB5737}$ catalyzes a transfer reaction of guanine (**3**) to **1** (Scheme 1). To prove this hypothesis, the coding sequence of $P450_{NB5737}$ was cloned into pET28a (+) (Table S2) and expressed in *E. coli* Rosetta (DE3). Induction with 0.5 mM IPTG at 18 °C for 20 h and purification with Ni-NTA agarose resulted in an apparently homologous protein with a yield of 3.5 mg/L culture (Figure S9).

Most bacterial P450s belong to a class I redox system consisting of three separate and soluble proteins, i.e., ferredoxin, ferredoxin reductase, and P450.[36] Usually, bacterial P450s, especially those from *Streptomyces*, also accept electrons from heterologous redox partners.[19] We investigated therefore the $P450_{NB5737}$ activity by using commercially available spinach ferredoxin and ferredoxin-$NADP^+$ reductase for electron transport. Incubation of the purified $P450_{NB5737}$ with **1** and **3** in the presence of these redox partners and NADPH led to the detection of **2** (Figure 3i), which was identified by comparison of its retention time, UV spectrum, $[M + H]^+$ ion, fragmentation pattern in MS^2 mode with those of the isolated product described above.

No product formation was observed in negative controls with heat-inactivated protein, without **1** or **3** (Figures 3ii–iv). However, only low conversion of approximate 1% was achieved under the tested condition. Increasing the concentration of ferredoxin, ferredoxin reductase, or both to 5-fold did not change the conversion yield (data not shown). Using GTP instead of guanine as donor, a trace of **2** was also detected (Figure 3v), but only approximately 2.3% of that with guanine (Figure 3i). It cannot be excluded that this amount was also a product of **1** with **3** as impurity or hydrolytic product of GTP.

UV–vis spectroscopic analysis of $P450_{NB5737}$ in the presence of sodium dithionite and carbon monoxide[37−39] revealed strong absorption at 415 nm of an inactive form and only a very weak shoulder peak around 446 nm of the active form (Figure S10). Addition of DTT to the mixture did not increase the intensity of the absorption at 446 nm (data not shown). In

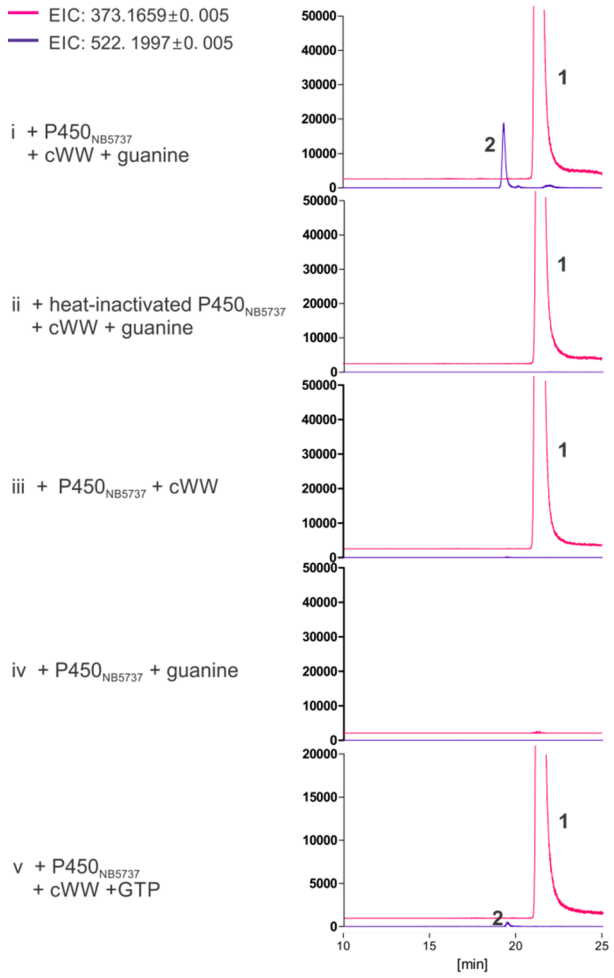


Figure 3. In vitro enzyme assays of recombinant $P450_{NB5737}$.

the case of TxtE, an atypical P450 catalyzing L-tryptophan nitration in the biosynthesis of thaxtomin A, the absorption at 447 nm strongly benefited from the added DTT.[37] All of these results indicate the presence of only a small portion of the active $P450_{NB5737}$ in the purified protein.

As aforementioned, P450s also catalyze intra- and intermolecular aryl–aryl coupling via radial intermediates.[25,27−30,40] In analogy, we proposed that the radicals of **1** at N1 and **3** at C8′are key intermediates (Figure 4). In this hypothesis, the two radicals (compound II) are formed parallel via oxidation of Fe^{II} via Fe^{III} and ferryl-oxo intermediate (Fe^{IV}=O, compound I).[19] Coupling of the two radicals led to the formation of **2**.

In conclusion, we identified in this study a P450 enzyme from *Streptomyces* that catalyzes an unusual coupling reaction of a cyclodipeptide with a nucleobase, leading to unprecedented adducts of cyclodipeptide with guanine. This represents one additional rare atypical P450 example catalyzing the coupling of two different substrates.[40,41] Intramolecular carbon–nitrogen bond formation between a sugar and its aglycone core was reported for StaN, a cytochrome P450 in the biosynthesis of staurosporine.[24,42] Coupling of guanine with *cyclo*-L-Trp-L-Trp increases significantly the CDP diversity and opens new prospective pathway in natural product discovery. To the best of our knowledge, modification of peptides by a nucleobase has not been reported prior to this study. Although $P450_{NB5737}$ is phylogenetically located closely to CYP121[25] and

DOI: 10.1021/acs.orglett.8b02051
Org. Lett. 2018, 20, 4921–4925

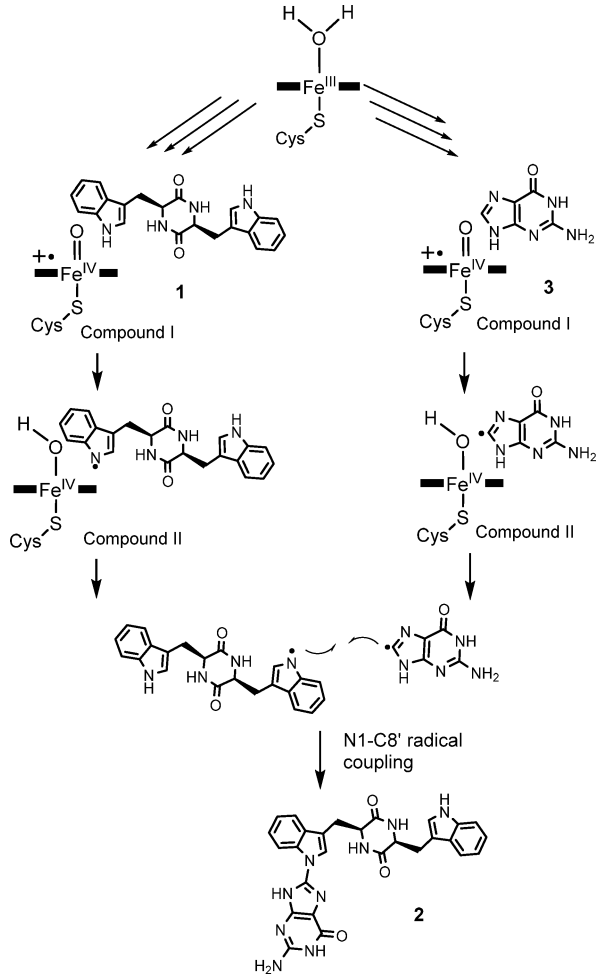


Figure 4. Proposed mechanism for the $P450_{NB5737}$-mediated intermolecular coupling reaction of cWW and guanine.

near to P-450mel, CYP158A1, and CYP158A2,[28−30] it uses clearly diverse substrates and catalyzes a distinct reaction as these enzymes. More importantly, this special P450 enzyme is obviously able to bind two significantly different structures as substrates and transfer them to radical intermediates for a coupling reaction. This indicates the presence of novel active site(s) differing from those of known P450s, which should be investigated in the future.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.8b02051.

Experimental details, spectroscopic data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: shuming.li@staff.uni-marburg.de. Tel/Fax: + 49-6421-28-22461/25365.

ORCID

Shu-Ming Li: 0000-0003-4583-2655

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the ARS Culture Collection (NRRL) for providing *Streptomyces* strains, R. Kraut (University Marburg) for LC−MS analysis, and Prof. Wenhan Lin (Peking University) for comments on the manuscript. The Bruker micrOTOF QIII mass spectrometer was financially supported in part by a grant from the Deutsche Forschungsgemeinschaft (INST 160/620-1 to S.-M.L.). H.Y. is a scholarship recipient of the China Scholarship Council (201306220024).

■ REFERENCES

(1) Cain, C. C.; Lee, D.; Waldo, R. H., III; Henry, A. T.; Casida, E. J., Jr.; Wani, M. C.; Wall, M. E.; Oberlies, N. H.; Falkinham, J. O., III *Antimicrob. Agents Chemother.* **2003**, *47*, 2113.
(2) Magyar, A.; Zhang, X.; Kohn, H.; Widger, W. R. *J. Biol. Chem.* **1996**, *271*, 25369.
(3) Ström, K.; Sjögren, J.; Broberg, A.; Schnürer, J. *Appl. Environ. Microbiol.* **2002**, *68*, 4322.
(4) Musetti, R.; Polizzotto, R.; Vecchione, A.; Borselli, S.; Zulini, L.; D'Ambrosio, M.; di Toppi, L. S.; Pertot, I. *Micron* **2007**, *38*, 643.
(5) Yamazaki, Y.; Tanaka, K.; Nicholson, B.; Deyanat-Yazdi, G.; Potts, B.; Yoshida, T.; Oda, A.; Kitagawa, T.; Orikasa, S.; Kiso, Y.; Yasui, H.; Akamatsu, M.; Chinen, T.; Usui, T.; Shinozaki, Y.; Yakushiji, F.; Miller, B. R.; Neuteboom, S.; Palladino, M.; Kanoh, K.; Lloyd, G. K.; Hayashi, Y. *J. Med. Chem.* **2012**, *55*, 1056.
(6) Kanzaki, H.; Yanagisawa, S.; Kanoh, K.; Nitoda, T. *J. Antibiot.* **2002**, *55*, 1042.
(7) Waring, P.; Beaver, J. *Gen. Pharmacol.* **1996**, *27*, 1311.
(8) Minelli, A.; Grottelli, S.; Mierla, A.; Pinnen, F.; Cacciatore, I.; Bellezza, I. *Int. J. Biochem. Cell Biol.* **2012**, *44*, 525.
(9) Borthwick, A. D. *Chem. Rev.* **2012**, *112*, 3641.
(10) González, J. F.; Ortín, I.; de la Cuesta, E.; Menéndez, J. C. *Chem. Soc. Rev.* **2012**, *41*, 6902.
(11) Giessen, T. W.; Marahiel, M. A. *Int. J. Mol. Sci.* **2014**, *15*, 14610.
(12) Belin, P.; Moutiez, M.; Lautru, S.; Seguin, J.; Pernodet, J. L.; Gondry, M. *Nat. Prod. Rep.* **2012**, *29*, 961.
(13) Mishra, A. K.; Choi, J.; Choi, S. J.; Baek, K. H. *Molecules* **2017**, *22*, E1796.
(14) 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. *Nat. Chem. Biol.* **2015**, *11*, 721.
(15) Li, S. M. *Nat. Prod. Rep.* **2010**, *27*, 57.
(16) Liu, J.; Yu, H.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 4435.
(17) Liu, J.; Yu, H.; Li, S.-M. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 5787.
(18) 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.
(19) Rudolf, J. D.; Chang, C. Y.; Ma, M.; Shen, B. *Nat. Prod. Rep.* **2017**, *34*, 1141.
(20) Guengerich, F. P. *Chem. Res. Toxicol.* **2001**, *14*, 611.
(21) Kelly, S. L.; Kelly, D. E. *Philos. Trans. R. Soc., B* **2013**, *368*, 20120476.
(22) Podust, L. M.; Sherman, D. H. *Nat. Prod. Rep.* **2012**, *29*, 1251.
(23) Zhang, X.; Li, S. *Nat. Prod. Rep.* **2017**, *34*, 1061.
(24) Greule, A.; Stok, J. E.; De Voss, J. J.; Cryle, M. J. *Nat. Prod. Rep.* **2018**, DOI: 10.1039/C7NP00063D.
(25) Belin, P.; Le Du, M. H.; Fielding, A.; Lequin, O.; Jacquet, M.; Charbonnier, J. B.; Lecoq, A.; Thai, R.; Courcon, M.; Masson, C.; Dugave, C.; Genet, R.; Pernodet, J. L.; Gondry, M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 7426.
(26) Cryle, M. J.; Bell, S. G.; Schlichting, I. *Biochemistry* **2010**, *49*, 7282.
(27) Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K. *ChemBioChem* **2014**, *15*, 656.

DOI: 10.1021/acs.orglett.8b02051
Org. Lett. 2018, 20, 4921−4925

(28) Funa, N.; Funabashi, M.; Ohnishi, Y.; Horinouchi, S. *J. Bacteriol.* **2005**, *187*, 8149.
(29) Zhao, B.; Guengerich, F. P.; Bellamine, A.; Lamb, D. C.; Izumikawa, M.; Lei, L.; Podust, L. M.; Sundaramoorthy, M.; Kalaitzis, J. A.; Reddy, L. M.; Kelly, S. L.; Moore, B. S.; Stec, D.; Voehler, M.; Falck, J. R.; Shimada, T.; Waterman, M. R. *J. Biol. Chem.* **2005**, *280*, 11599.
(30) Zhao, B.; Lamb, D. C.; Lei, L.; Kelly, S. L.; Yuan, H.; Hachey, D. L.; Waterman, M. R. *Biochemistry* **2007**, *46*, 8725.
(31) 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.
(32) Gomez-Escribano, J. P.; Bibb, M. J. *Microb. Biotechnol.* **2011**, *4*, 207.
(33) Doumith, M.; Weingarten, P.; Wehmeier, U. F.; Salah-Bey, K.; Benhamou, B.; Capdevila, C.; Michel, J.-M.; Piepersberg, W.; Raynal, M.-C. *Mol. Gen. Genet.* **2000**, *264*, 477.
(34) Marzilli, L. D.; de Castro, B.; Solorzano, C. *J. Am. Chem. Soc.* **1982**, *104*, 461.
(35) Teijeira, M.; Santana, L.; Uriarte, E. *Magn. Reson. Chem.* **1997**, *35*, 806.
(36) Hannemann, F.; Bichet, A.; Ewen, K. M.; Bernhardt, R. *Biochim. Biophys. Acta, Gen. Subj.* **2007**, *1770*, 330.
(37) Barry, S. M.; Kers, J. A.; Johnson, E. G.; Song, L.; Aston, P. R.; Patel, B.; Krasnoff, S. B.; Crane, B. R.; Gibson, D. M.; Loria, R.; Challis, G. L. *Nat. Chem. Biol.* **2012**, *8*, 814.
(38) Omura, T.; Sato, R. *J. Biol. Chem.* **1964**, *239*, 2370.
(39) Shenkman, J. B.; Jansson, I. Spectral analyses of cytochrome P450. In *Cytochrome P450 Protocols*, 2nd ed.; Phillips, I. R., Shephard, E. A., Eds.; Humana Press Inc: Totowa, NJ, 2006; pp 11−18.
(40) Lin, H.-C.; McMahon, T. C.; Patel, A.; Corsello, M.; Simon, A.; Xu, W.; Zhao, M.; Houk, K. N.; Garg, N. K.; Tang, Y. *J. Am. Chem. Soc.* **2016**, *138*, 4002.
(41) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. *Nat. Chem. Biol.* **2014**, *10*, 640.
(42) Onaka, H.; Asamizu, S.; Igarashi, Y.; Yoshida, R.; Furumai, T. *Biosci., Biotechnol., Biochem.* **2005**, *69*, 1753.

DOI: 10.1021/acs.orglett.8b02051
Org. Lett. 2018, 20, 4921−4925

Supporting information

Coupling of Guanine with *cyclo*-L-Trp-L-Trp Mediated by a Cytochrome P450 Homologue from *Streptomyces purpureus*

Huili Yu[†], Xiulan Xie[‡], and Shu-Ming Li[†*]

[†] Institut für Pharmazeutische Biologie und Biotechnologie, 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

Computer-assisted sequence analysis

Nucleotide and protein sequences used in this study were downloaded from NCBI databases (http://www.ncbi.nlm.nih.gov). Protein sequences were compared with each other by using BLAST programs (http://blast.ncbi.nlm.nih.gov/). Multiple sequence alignments were performed by using the program ClustalW and visualized with ESPript 3.0 (http://endscript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) to identify strictly conserved amino acid residues (Figure S1). The phylogenetic tree in Figure S2 was created by MEGA version 5.0 (http://www.megasoftware.net).

Plasmids, bacterial strains, and culture conditions

Plasmids and constructs used in this study are listed in Tables S2. *E. coli* DH5α was used as host for cloning and *E. coli* Rosetta (DE3) for P450 production. *Streptomyces* strains mentioned in Figure 1 and Table S1 were obtained from the ARS Culture Collection (NRRL). *S. coelicolor* M1146[1] was kindly gifted by Prof. Bibb (John Innes Centre). Recombinant *E. coli* strains were cultivated in liquid/solid LB or TB media containing 100 μg/mL ampicillin, 50 μg/mL apramycin or 50 μg/mL kanamycin. The *Streptomyces* strains were maintained at 28 °C in GYM (glucose 4.0 g/L, yeast extract 4.0 g/L, malt extract 10.0 g/L, agar 15.0 g/L, pH 7.3) or ISP4 media (soluble starch 10.0 g/L, K_2HPO_4 1.0 g/L, $MgSO_4 \cdot 7H_2O$ 1.0 g/L, NaCl 1.0 g/L, $(NH_4)_2SO_4$ 2.0 g/L, $CaCO_3$ 1.0 g/L, trace element solution[2] 1.0 mL/L, agar 15.0 g/L, pH 7.2.).

Genetic manipulation, PCR amplification, and gene cloning

Genetic manipulation in *E. coli* was carried out according to the protocol by Sambrook and Russell.[3] Isolation of genomic DNA from *Streptomyces* was performed as described previously.[2] The genes encoding CDPS and P450 were amplified by PCR from genomic DNA by using primers listed in Table S2 and Phusion® High-Fidelity DNA Polymerase from New England Biolabs (NEB). The generated PCR fragments were cloned into pGEM-T Easy vector and the sequence integrity was confirmed by sequencing. After sequencing, the fragments were cut off with the appropriate restriction endonucleases from pGEM-T Easy and ligated into pSET152, pPWW50 or pET28a (+) vector. They were digested with the same enzymes, previously. The generated plasmids (Table S2) were transformed into *S. coelicolor* M1146 and *E. coli* Rosetta (DE3) for gene expression and P450 protein production, respectively.

Gene expression

For gene expression, the constructs of CDPSs and P450s were transformed into *S. coelicolor* M1146 by conjugation.[2] The exconjugants were selected with 1 mg apramycin and 0.5 mg trimethoprim in 1 mL H_2O. The spores of the *S. coelicolor* M1146 transformants were inoculated to 30 mL of GYM liquid media supplied with 30 µg/mL of apramycin in 250 mL baffled flask and cultured at 28 °C, 180 rpm for 7 days. In the biotransformation experiments, the spores of M1146 harboring the P450 expression plasmid pHY44 were inoculated into GYM liquid media (30 µg/mL of apramycin) and cultivated for 2 days. After addition of 80 µM of cWW, the bacterial cultures were maintained further for 7 days. 1 mL of culture was extracted with one volume of ethyl acetate for three times. The organic phases were combined, evaporated, and the residues were afterward dissolved in 30 µL of methanol. 5 µL of such samples were taken for LC-MS analysis.

LC-MS analysis

The ethyl acetate extracts were analyzed on an Agilent HPLC 1260 series equipped with a photo diode array detector and a Bruker microTOF QIII mass spectrometer by using an Agilent Eclipse XDB C18 column (5 µm, 4.6 x 150 mm). A linear gradient of 5 -100% acetonitrile in water, both containing 0.1% formic acid, in 40 min and a flow rate at 0.25 mL/min were used. The column was then washed with 100% acetonitrile containing 0.1% formic acid for 5 min and equilibrated with 5% acetonitrile in water for 5 min. The parameters of the spectrometer were set as following: electrospray positive ion mode for ionization, capillary voltage with 4.5 kV, collision energy with 8.0 eV.

Isolation and structural elucidation of P450 products

For isolation of the P450 products, the CDPS and P450 coexpression transformant was cultivated on a large scale (4 L) at 28 °C for 7 days. The culture supernatant was extracted with equal volume of ethyl acetate for three times. The organic phases were combined, evaporated, and dried in a freezing dryer. The extracts were loaded on a silica gel column (45×3.5 cm) filled with 120 g silica gel 60 (230-400 mesh, Carl Roth GmbH, Karlsruhe, Germany) and eluted with mixtures of CH_2Cl_2 : MeOH in ratios of 100:2, 100:3, 100:5, 100:10 (1000 mL each). The products were detected in the fraction eluted with the mixture of 100:10. Fractions containing product **2** was further separated on an Agilent HPLC 1260 series equipped with a photo diode array detector by using a semipreparative Multospher 120 RP-18 column (10 x 250 mm, 5 µm)

with 30% ACN in water. The flow rate was set to 2.5 mL/min. Finally, 13 mg of **2** was obtained as pale yellow powder and subjected to NMR analysis.

NMR experiments were performed at room temperature on a Bruker AVIII spectrometer equipped with a 5 mm BBO cryo probe Prodigy with z-gradient. All spectra were processed with MestReNova 5.2.2 (Metrelab Research, S5 Santiago de Compostella, Spain). NMR data and spectra of the product **2** were provided as Table S3 and Figures S3-S8, respectively.

Determination of production yield of compound 2 in *Streptomyces*

For quantification, an Agilent HPLC 1200 series equipped with a photo diode array detector and the same column with that used in LC-MS was used. A linear gradient of 10 to 100% acetonitrile in water without acid in 40 min was followed by 100% acetonitrile for 5 min and 10% acetonitrile in water for 5 min. The flow rate was 0.5 mL/min. The absorption at 277 nm was used for quantification. 3 mg of compound **2** was dissolved in 250 µL DMSO and used as standard. 2 mL culture of the M1146 transformant carrying pHY48 for $CDPS_{NB5737}$ and $P450_{NB5737}$ expression was extracted with 1 mL ethyl acetate for 3 times. The organic fractions were combined and dried. The residue was dissolved in 50 µL of methanol and analyzed on HPLC. The production yield was calculated to be 8 mg/L in *S. coelicolor* M1146.

Production and purification of $P450_{NB5737}$ in *E. coli*

E. coli Rosetta (DE3) cells harboring pHY99 were cultivated in 50 mL LB media (50 µg/mL kanamycin) for 16 h as preculture. Five milliliters of the preculture were transferred to 500 mL TB media (50 µg/mL kanamycin) in 2 L-Erlenmeyer flask and grown at 37 °C, 230 rpm to an absorption of 0.6 at 600 nm. The gene expression was induced with 0.5 mM IPTG at 18 °C for 20 h. The bacterial cultures were harvested by centrifugation (4500 rpm, 15 min, 4 °C) and the cells were resuspended in lysis buffer (10 mM imidazole, 50 mM NaH_2PO4, 300 mM NaCl, pH 8.0) with 2–5 mL/g wet weight. Lysozyme from the chicken egg white was added at a final concentration of 1 mg/mL and incubated on ice for 30 min. The cells were lysed by sonication on ice. Cell debris was removed by centrifugation at 13,000 rpm and 4 °C for 30 min. One-step purification of the recombinant His_6-tagged protein was performed by using Ni-NTA agarose (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. The storage buffer was changed to 50 mM Tris-HCl (pH 7.5) containing 15% (v/v) glycerol through a NAP-5 column (GE Healthcare, Freiburg, Germany), which had been equilibrated with the same buffer. The obtained protein was stored frozen at −80 °C. Protein concentration was determined on

Nanodrop C 2000 (Thermo Scientific, Braunschweig, Germany) and 3.5 mg/L culture of $P450_{NB5737}$ was obtained. The purity of obtained protein was proven with 12% (w/v) SDS-PAGE (Figure S9).

UV-Vis spectroscopic analysis of $P450_{NB5737}$

To measure the typical absorbance of P450 ferrous CO complex after reduction, carbon monoxide gas was bubbled into the $P450_{NB5737}$ solution (10.5 µM in 0.1 M potassium phosphate buffer containing 20% glycerol, pH 7.4) for 2 min. After addition of 0.1 mg/mL of sodium dithionite, a UV-Vis spectrum between 350 and 550 nm was taken on a Multiskan™ GO Microplate Spectrophotometer (Thermo Scientific, Dreieich, Germany) by using Greiner UV-Star® 96 well plates (Greiner Bio-One GmbH, Frickenhausen, Germany). UV-Vis spectra of $P450_{NB5737}$ without treatment and after addition of sodium dithionite were also taken as controls. The spectra are shown in Figure S10.

In vitro assays of $P450_{NB5737}$

$P450_{NB5737}$ activities toward cWW and guanine or GTP were measured in 50 µL system containing 5 µM purified $P450_{NB5737}$, 1 mM cWW, 0.5 mM guanine or GTP, 1 µM spinach ferredoxin (Sigma-Aldrich), 0.1 unit/mL spinach ferredoxin-$NADP^+$ reductase (Sigma-Aldrich), and 1 mM NADPH in 50 mM Tris-HCl at pH 7.5. The reaction mixtures were incubated at 30 °C for 16 h. The negative controls with heat-inactivated $P450_{NB5737}$, without cWW or guanine were performed under the same conditions. The reactions were stopped by addition of 200 µL MeOH each. The supernatants after centrifugation were dried in Speedvac and redissolved in 30 µL MeOH. 5 µL of these samples were taken for LC-MS analysis.

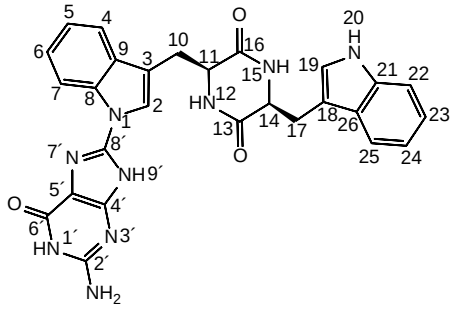
Table S1. Comparison of CDPSs and P450s in three *Streptomyces* strains.

Protein (length in aa) of *Strptomyces purpureus* NRRL B-5737	Identity to proteins (length in aa) of (%)			Protein function
	Streptomyces lavendulae NRRL B-2774		*Streptomyces xanthophaeus* NRRL B-5414	
WP_019889609 (239)	WP_078950527 (265) 58		WP_063768158 (273) 58	CDPS (cWW synthase)
WP_019889608 (398)	WP_051841251 (416) 75		WP_031150378 (412) 75	Putative P450

Table S2. Cloning and expression constructs used in this study.

Protein	Accession number	Primer sequences (5'-3')	pGEM-T easy construct	Cloning sites	Expression vector	Expression construct
$CDPS_{NB5737}$	WP_ 019889609	GCCATATGACTCTCATCGAAGACAC TCTAGACTAAGCGGCTCGACGGTCAT	pHY40	*Nde*I/ *Xba*I *Nde*I/*Xba*I	pPWW50 pSET152	pHY42 pHY43
$P450_{NB5737}$	WP_ 019889608	CATATGACCGTCGAGCCGCTTGA TCTAGACTACCACAACACCGGCAGCC	pHY41	*Nde*I/*Xba*I	pPWW50	pHY44
		CATATGACCGTCGAGCCGCTTGA GAATTCTCACCACAACACCGGCAGC	pHY98	*Nde*I/*Eco*RI	pET28a(+)	pHY99
$(CDPS+ P450)_{NB5737}$	-	GCCATATGACTCTCATCGAAGACAC TCTAGACTACCACAACACCGGCAGCC	pHY46	*Nde*I/*Xba*I *Nde*I/*Xba*I	pPWW50 pSET152	pHY47 pHY48

Table S3. ^{1}H (500 MHz) and ^{13}C NMR (125 MHz) data of the isolated compound **2** in DMSO-*d*6

		
Pos.	δ_H, multi., *J* [Hz]	δ_C
1	-	-
2	7.75, s	125.6
3	-	109.6
4	7.38, d, 7.9	118.5
5	7.00, dd, 7.6, 2.9	119.1
6	7.12, t, 7.6	121.3
7	8.63, d, 8.3	114.8
8	-	134.5
9	-	128.9
10 a	2.80, dd, 14.4, 4.4	29.0
10 b	2.56, dd, 14.4, 5.4	
11	4.04, t, 4.4	55.2
12	7.99, d, 1.7	-
13	-	167.09
14	3.99, t, 5.4	55.0
15	7.79, d, 1.9	-
16	-	167.10
17 a	2.77, dd, 14.4, 4.4	29.1
17 b	2.46, dd, 14.4, 6.1	
18	-	108.6
19	6.59, d, 1.9	124.1
20	11.12, br	-
21	-	136.0
22	7.29, d, 8.1	111.3
23	6.99, dd, 7.6, 2.9	120.7
24	6.90, t, 7.5	118.2
25	7.35, d, 8.0	118.3
26	-	127.5
1'	10.00, br	-
NH_2-2'	5.59, s	156.8
3'	-	-
4'	-	150.0
5'	-	116.4
6'	-	161.1
7'	-	-
8'	-	153.1
9'	Not observed	-

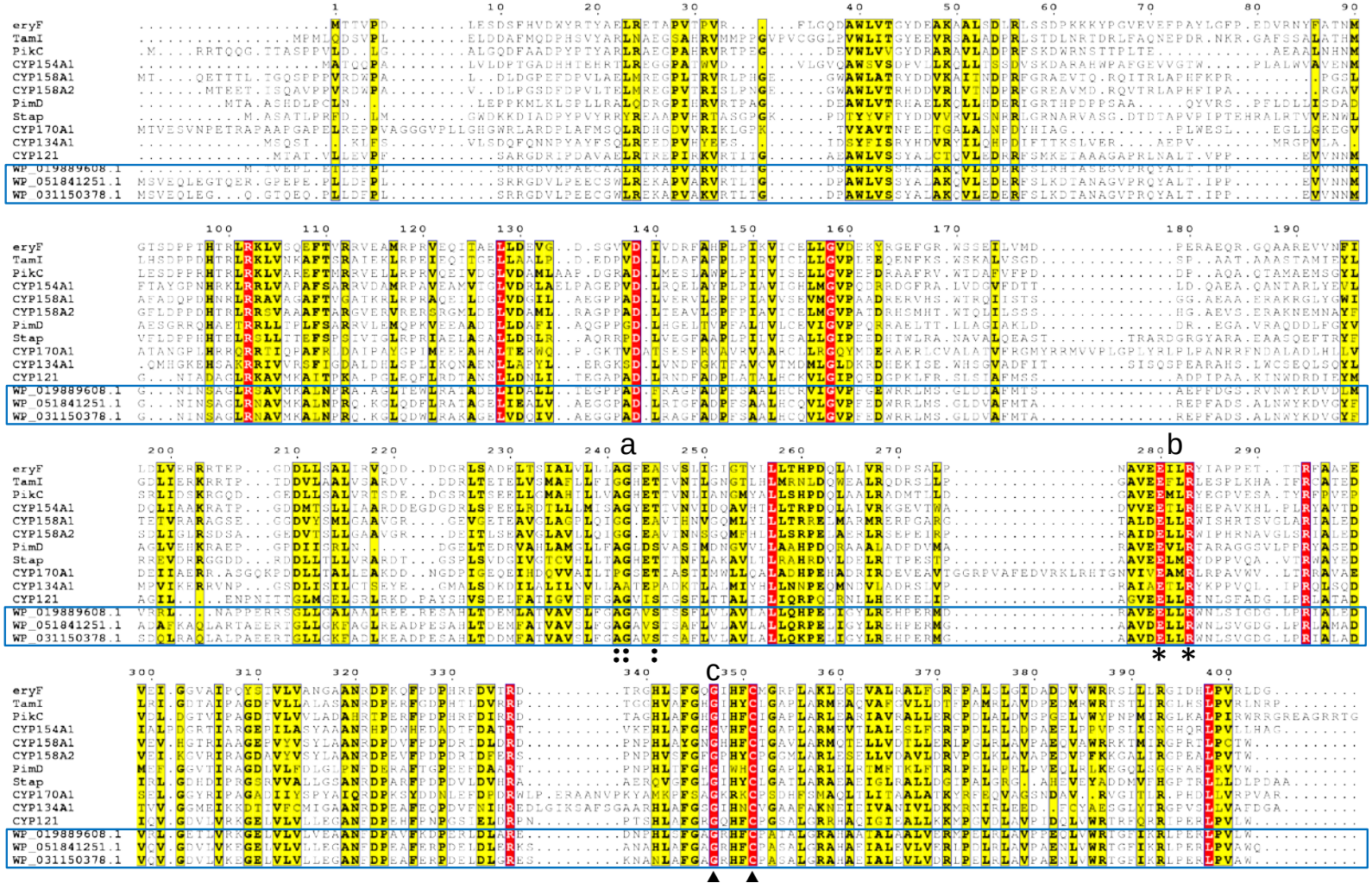


Figure S1. Alignments of P450s from this study (blue box) and structurally defined natural product P450s from bacteria. Residues are numbered according to the sequence of EryF.[4] a: A/G^{241}G^{242}XXT^{245} motif, the highly conserved residue A/G241, G242 and T245 are marked with "**:**". b: E^{280}XXR^{283} motif, the strictly conserved residues E280 and R283 are marked with "*****". c: G^{347}XXXC^{351} motif, the strictly conserved residues G347 and C351 are marked with "▲".

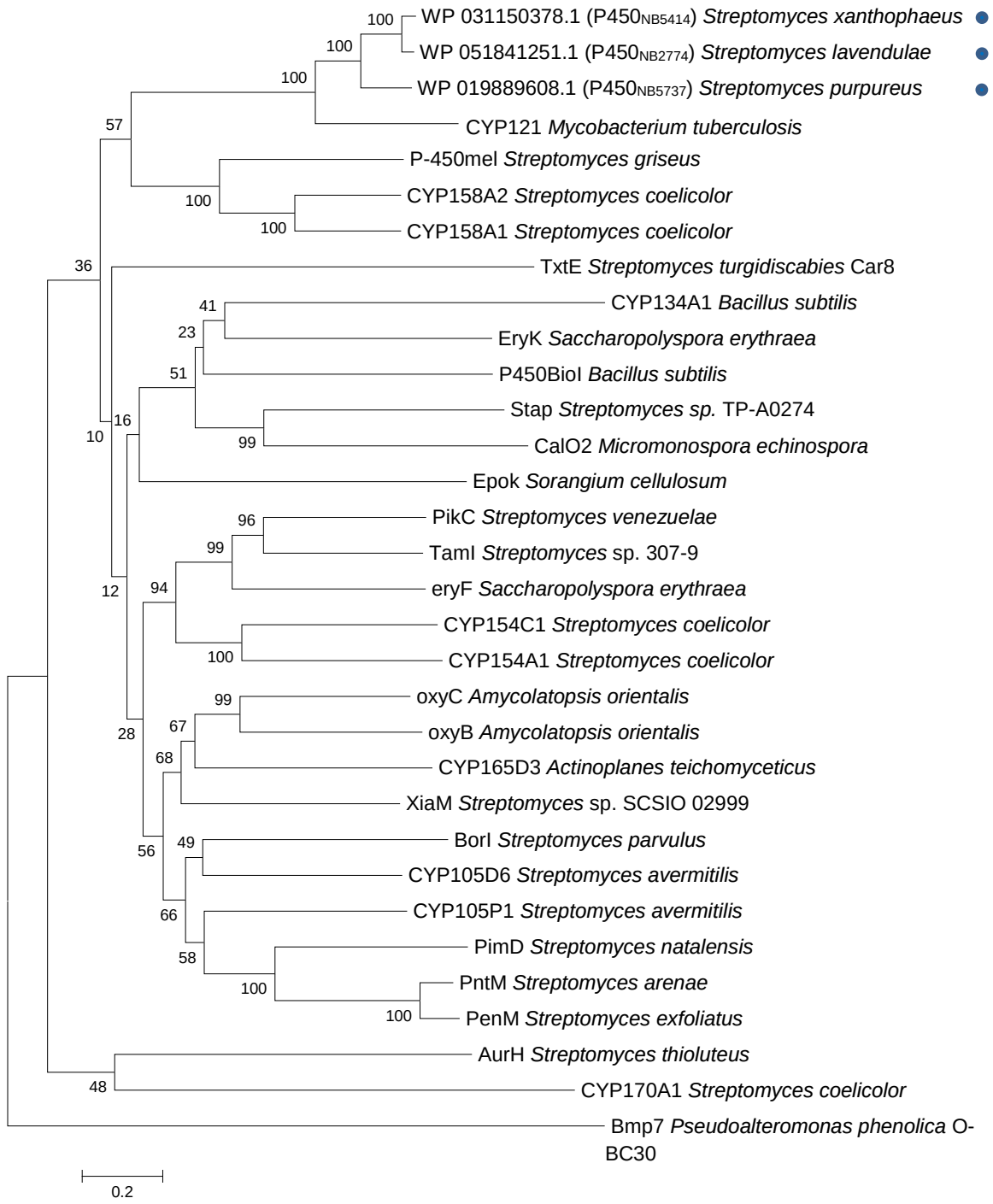


Figure S2. Phylogenetic analysis of P450s from this study (●) and functionally characterized P450s from bacteria. Scale bar, 0.2 substitutions per site. The protein sequneces were downloaded from NCBI database.

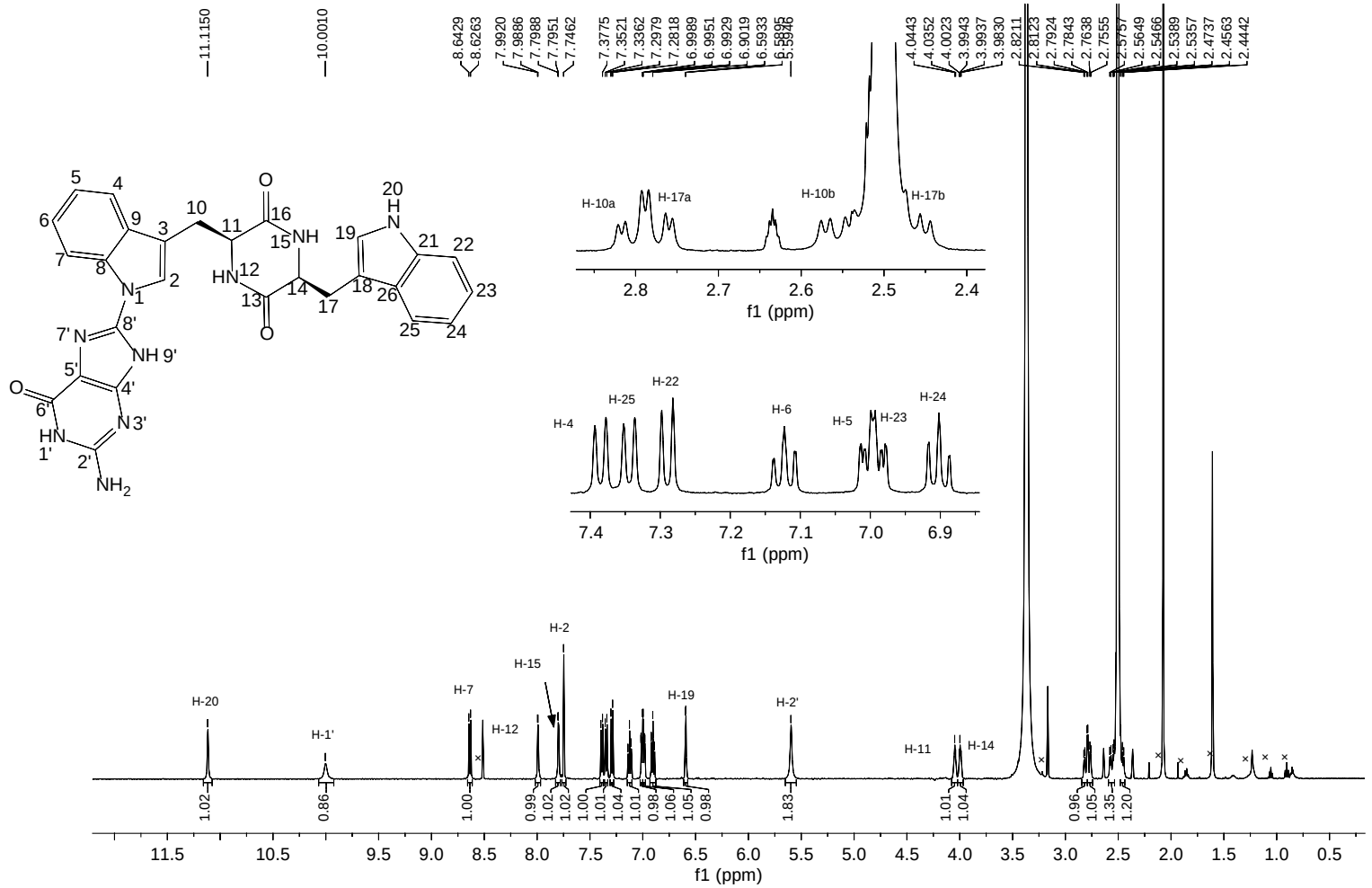


Figure S3. ^{1}H NMR spectrum of compound **2** (DMSO-*d*6, 500 MHz).

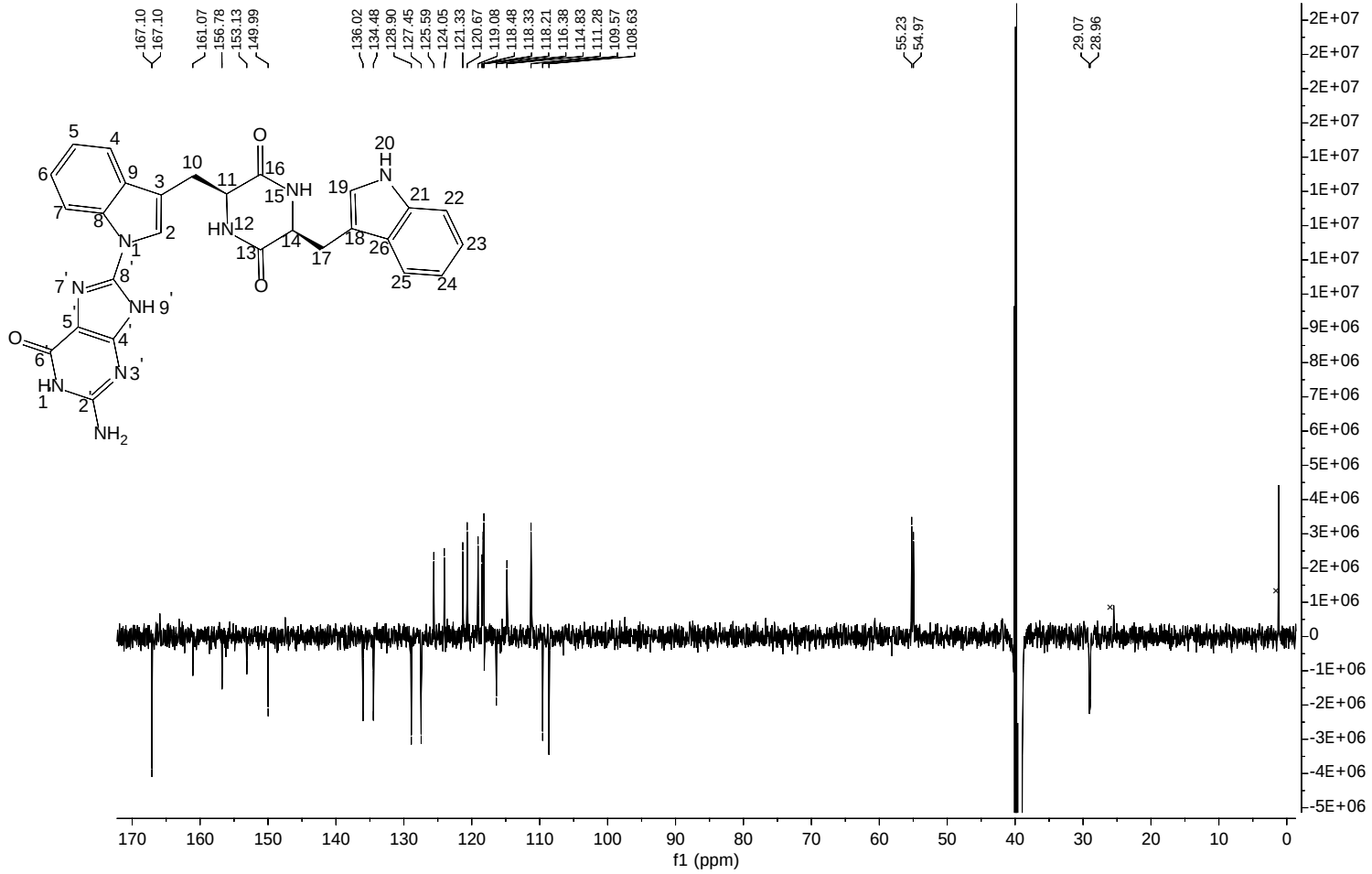


Figure S4. ^{13}C APT spectrum of compound **2** (DMSO-*d*6, 125 MHz).

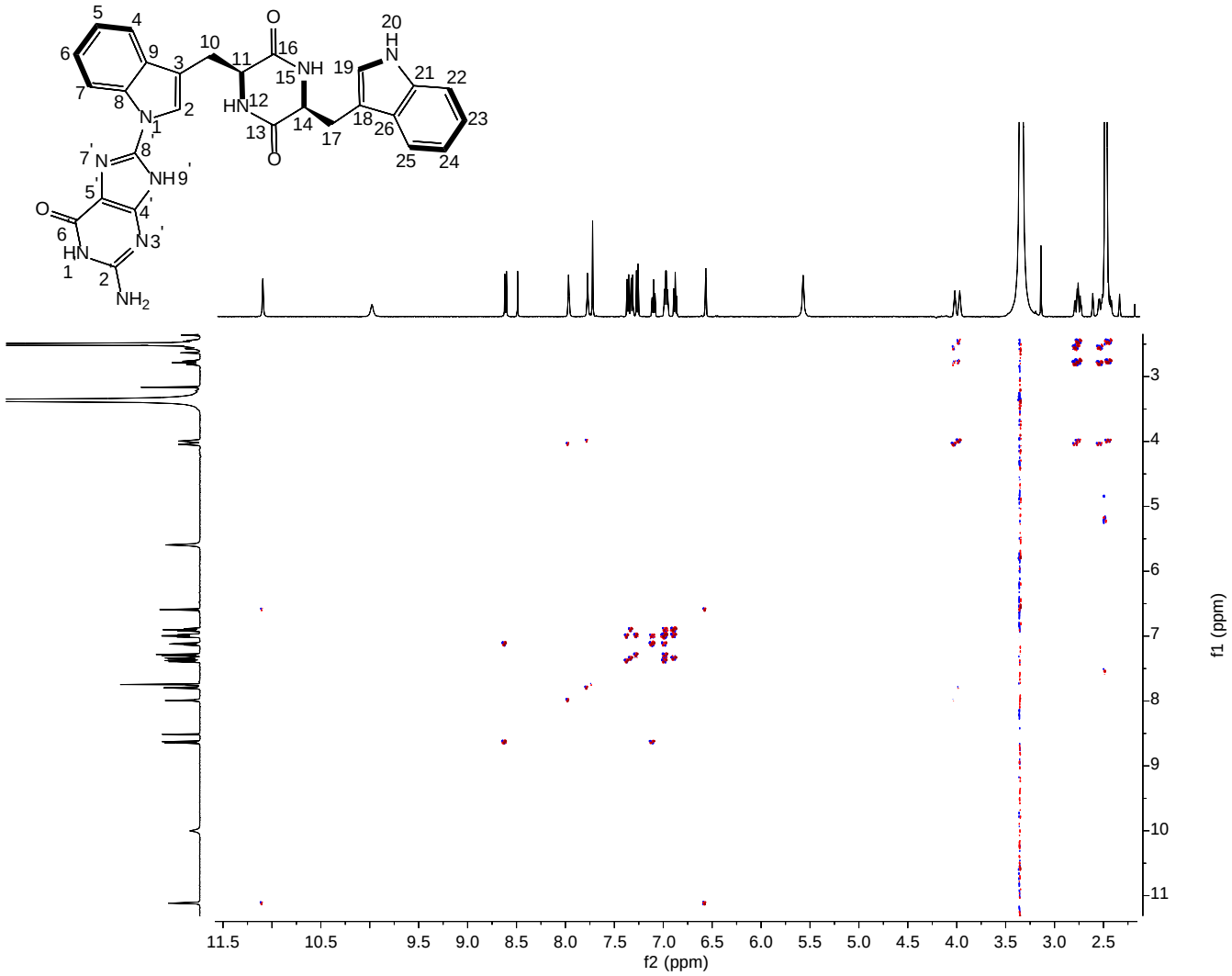


Figure S5. DQF-COSY spectrum of compound **2** (DMSO-*d*6, 500 MHz, 500 MHz).

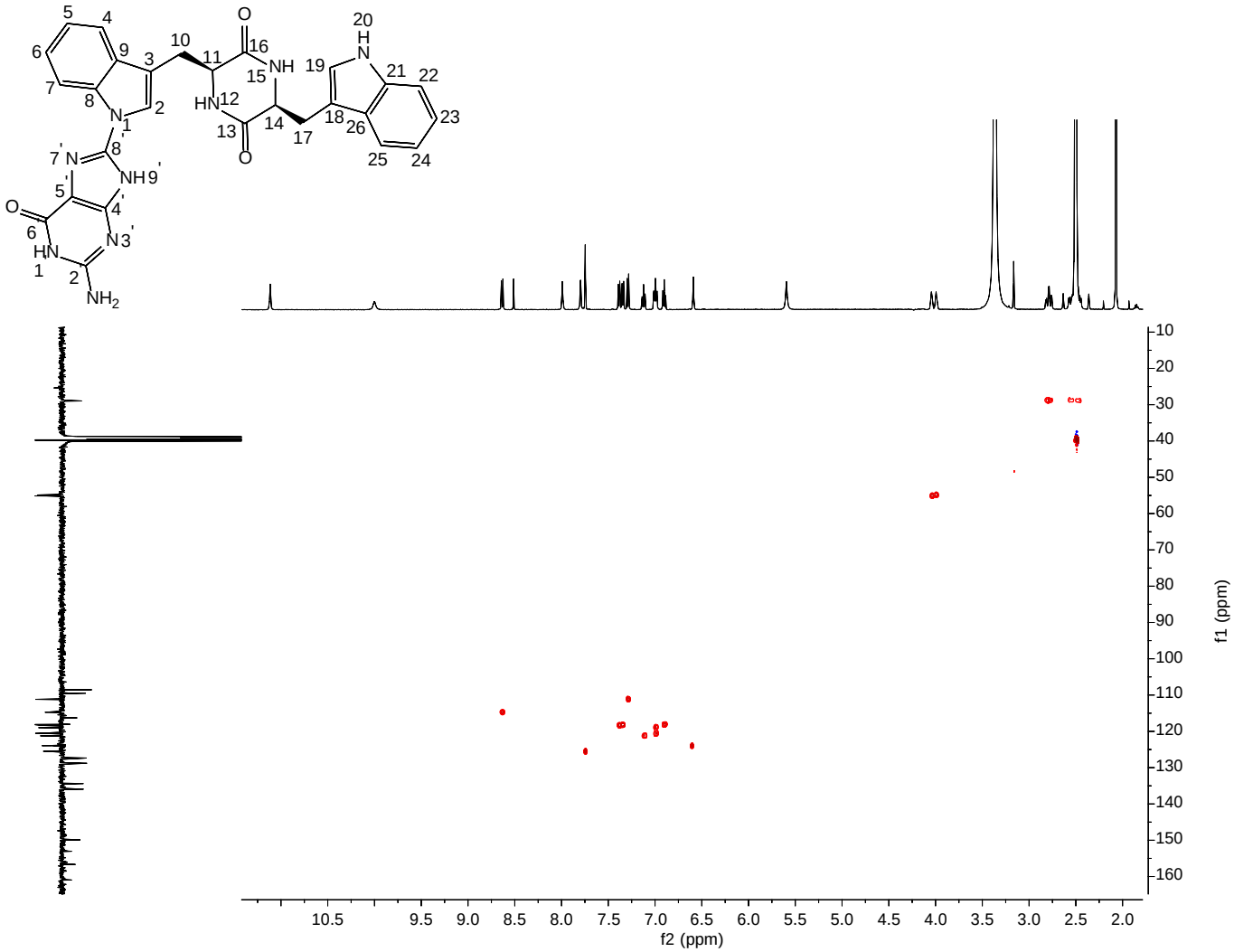


Figure S6. HSQC spectrum of compound **2** (DMSO-*d*6, 500 MHz, 125MHz).

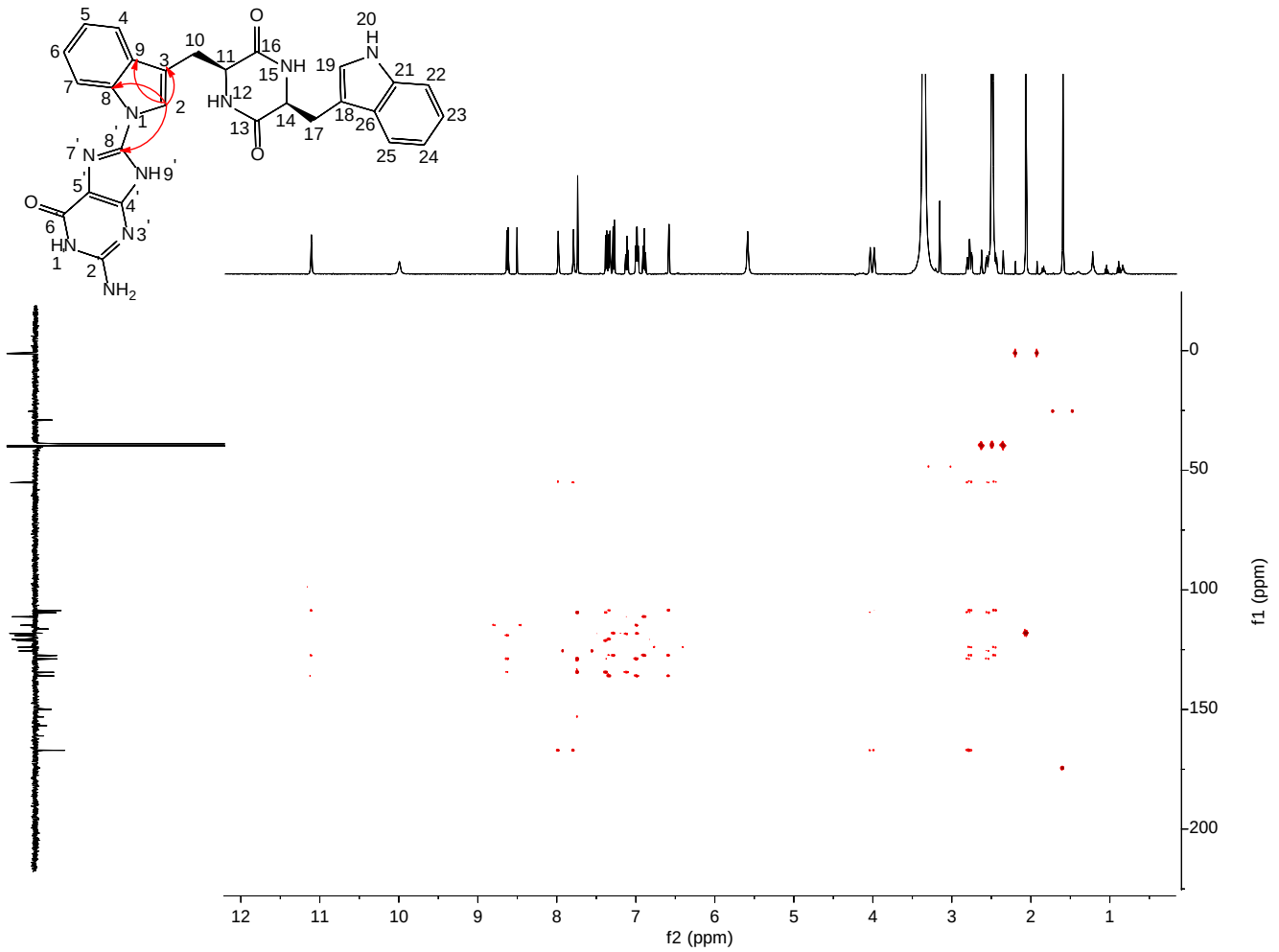


Figure S7. HMBC spectrum of compound **2** (DMSO-*d*6, 500 MHz, 125MHz).

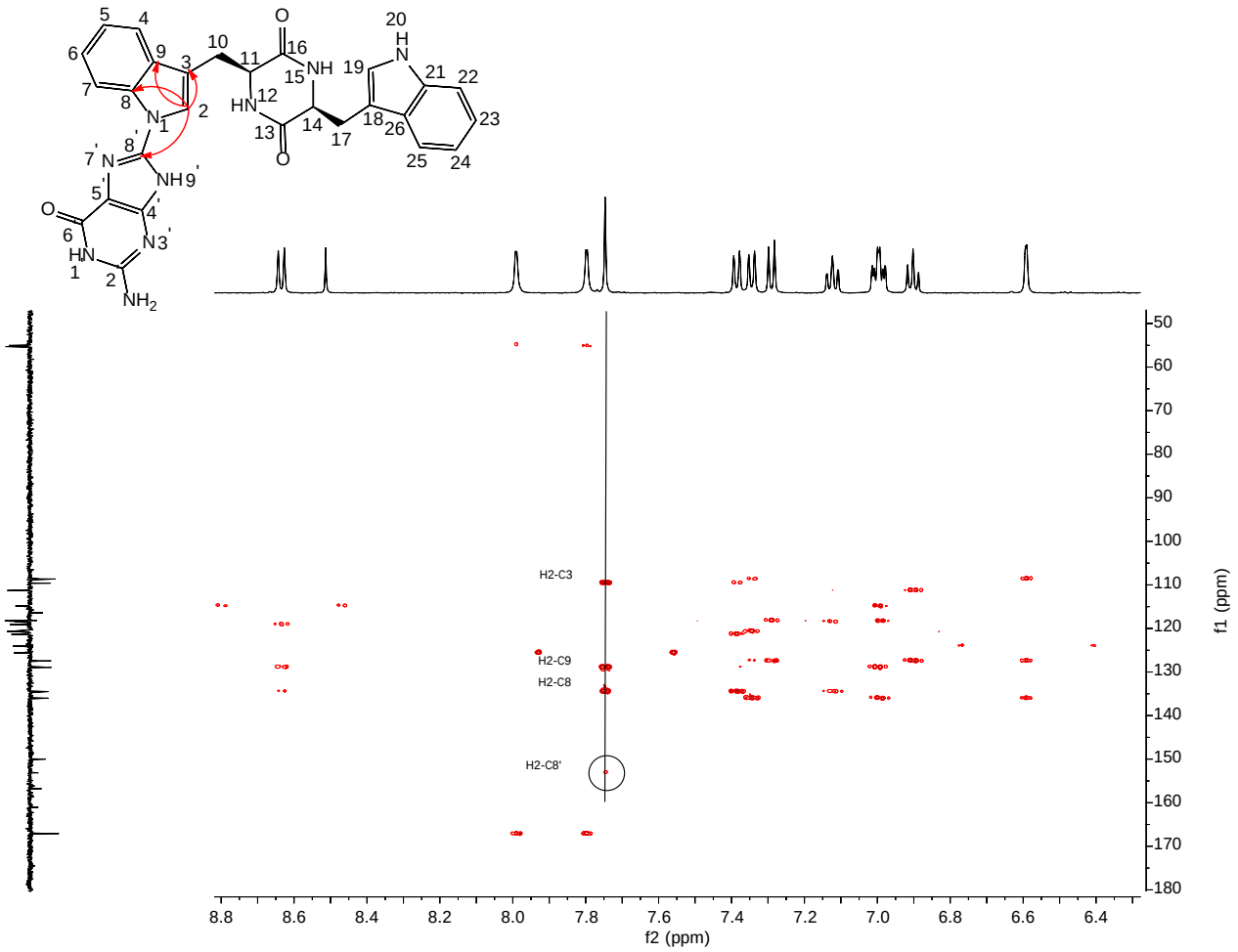


Figure S8. Selected region of HMBC spectrum of compound **2** (DMSO-*d*6, 500 MHz, 125MHz).

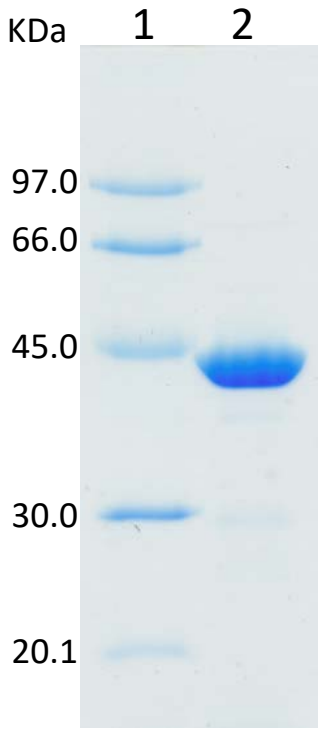


Figure S9. SDS-PAGE analysis of the purified $P450_{NB5737}$. The proteins were separated on a 12% polyacrylamide gel and stained with Coomassie brilliant blue R-250. Lane 1: Protein marker, lane 2: purified $P450_{NB5737}$.

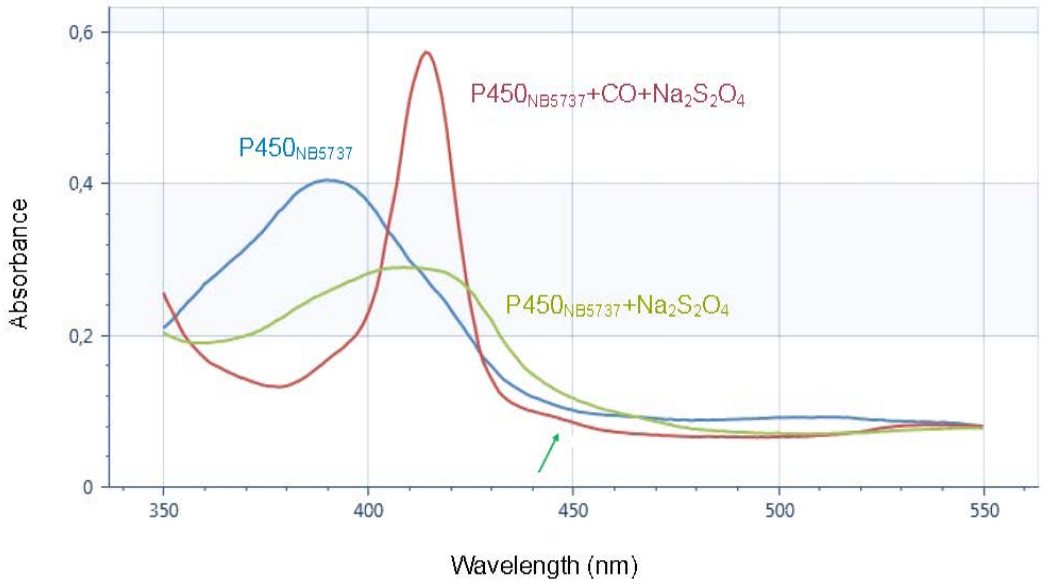


Figure S10. Spectroscopic analysis of $P450_{NB5737}$.

References:

(1) Gomez-Escribano, J. P. and Bibb, M. J. *Microb.Biotechnol.* **2011,** *4,* 207.
(2) 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.
(3) Sambrook, J.; Russell, D. W. *Molecular cloning: a laboratory manual;* 3rd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor: New York, 2001.
(4) Cupp-Vickery, J. R. and Poulos, T. L. *Nat.Struct.Biol.* **1995,** *2,* 144.

4.3 Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate.

Appl Microbiol Biotechnol (2015) 99:7115–7124
DOI 10.1007/s00253-015-6452-1

Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior toward unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate

Huili Yu · Mike Liebhold · Xiulan Xie · Shu-Ming Li

Received: 16 December 2014 / Revised: 23 January 2015 / Accepted: 31 January 2015 / Published online: 24 February 2015

Abstract Prenyltransferases of the dimethylallyltryptophan synthase (DMATS) superfamily are involved in the biosynthesis of secondary metabolites and contribute as modification enzymes significantly to structural diversity of natural products. They show usually broad specificity toward their aromatic substrates with regiospecific prenylations on aromatic rings. However, most members of this superfamily exhibit a high specificity toward their prenyl donors and usually accept exclusively dimethylallyl diphosphate (DMAPP). Recently, several indole prenyltransferases from this family were also demonstrated to accept unnatural DMAPP analogs such as methylallyl, 2-pentenyl and benzyl diphosphate for alkylation, or benzylation of the indole ring. Partial or complete shift of the substitution position was observed for these enzymes. In this study, we report the acceptance of these DMAPP analogs by two tyrosine *O*-prenyltransferases TyrPT from *Aspergillus niger* and SirD from *Leptosphaeria maculans* for alkylation or benzylation of tyrosine and derivatives. NMR and mass spectrometry (MS) analyses of nine isolated enzyme products confirmed the regiospecific *O*- or *N*-alkylation or benzylation at position C-4 of the aromatic ring, which is the same prenylation position of these enzymes in the presence of DMAPP.

Keywords DMATS superfamily · Unnatural alkyl donors · Tyrosine and derivatives · Enzyme catalysis

Introduction

Attachment of a prenyl moiety to an aromatic core plays an important role for creation of structural diversity of secondary metabolites (Heide 2009). Prenyltransferases catalyzing such attachments are found in many living organisms including plants, bacteria, and fungi and can be classified in different subgroups based on their amino acid sequences and biochemical properties (Heide 2009; Yazaki et al. 2009; Yu and Li 2012). The members of the dimethylallyltryptophan synthase (DMATS) superfamily catalyze the transfer reactions of prenyl moiety from prenyl donors, usually dimethylallyl diphosphate (DMAPP), to diverse aromatic substrates (Chooi et al. 2012; Yu and Li 2012). The resulted prenylated products often show interesting biological and pharmacological activities distinct from their non-prenylated precursors, which makes these enzymes to be interesting biocatalysts for pharmaceutical and medicinal research (Li 2010; Wollinsky et al. 2012). So far, more than 40 such enzymes have been characterized biochemically by using heterologously overproduced and purified proteins (Fan et al. 2014; Pockrandt et al. 2014; Wunsch et al. 2015; Yu and Li 2012). The majority of these enzymes accepts indole derivatives (Yu and Li 2012), while several members use tyrosine or nitrogen-free aromatic compounds as natural substrates (Fan et al. 2014; Kremer and Li 2010; Pockrandt et al. 2012, 2014). Until now, two L-tyrosine prenyltransferases have been identified and characterized biochemically. The first one, SirD, is involved in the biosynthesis of sirodesmin PL in *Leptosphaeria maculans* (Fig. 1) (Gardiner et al. 2004; Kremer and Li 2010). The very recently identified TyrPT from an unknown biosynthetic pathway in *Aspergillus niger* (Fig. 1) showed similar biochemical features

Electronic supplementary material The online version of this article (doi:10.1007/s00253-015-6452-1) contains supplementary material, which is available to authorized users.

H. Yu · M. Liebhold · S.-M. Li (✉)
Institut für Pharmazeutische Biologie und Biotechnologie, Philipps-Universität Marburg, Deutschhausstrasse 17A, 35037 Marburg, Germany
e-mail: shuming.li@staff.uni-marburg.de

X. Xie
Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Straße 4, 35032 Marburg, Germany

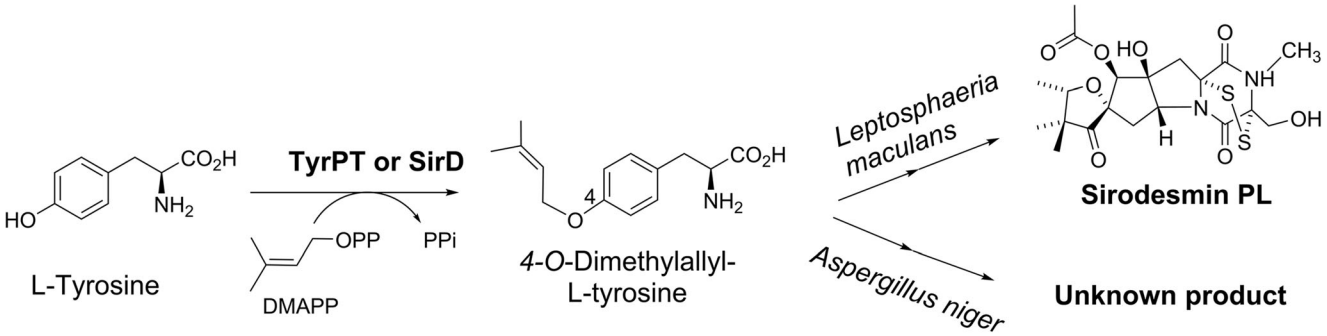


Fig. 1 Prenylations catalyzed by TyrPT and SirD and their role in nature

as SirD (Fan et al. 2014). Both enzymes catalyze the same *O*-prenylation of the hydroxyl group at the benzene ring (Fan et al. 2014; Kremer and Li 2010).

One common feature of the DMATS superfamily enzymes is their high flexibility toward aromatic substrates. In addition to tryptophan derivatives, indole prenyltransferases also accept distinct structures such as hydroxynaphthalenes (Yu et al. 2011), flavonoids (Yu and Li 2011), or xanthones (Tarcz et al. 2014) as prenyl acceptors. The tryptophan *C4*-(FgaPT2) and *C7*-prenyltransferase (7-DMATS) have also been demonstrated to accept tyrosine and derivatives as substrates (Fan et al. 2015; Fan and Li 2014). Similar to the behavior of indole prenyltransferases, the tyrosine prenyltransferases SirD and TyrPT also accept a number of tyrosine derivatives and catalyze the *O*-prenylation of the hydroxyl or *N*-prenylation of the amino group at the aromatic rings (Fan et al. 2014; Kremer and Li 2010; Zou et al. 2011). Furthermore, they also catalyze prenylation reactions of tryptophan and derivatives (Fan et al. 2014; Kremer and Li 2010; Rudolf and Poulter 2013).

In contrast to their high substrate flexibility for acceptors, most of the DMATS enzymes show a restricted specificity toward their prenyl donors, accepting usually DMAPP as prenyl donor (Yu and Li 2012). Only a few enzymes use GPP, DMAPP and GPP, or DMAPP, GPP, and FPP as prenyl donors (Chooi et al. 2010; Pockrandt et al. 2014; Pockrandt and Li 2013; Tarcz et al. 2014). This feature prohibits their application as biocatalysts in chemoenzymatic synthesis. To broaden their specificity for alkyl donors, the acceptance of the unnatural analogs, methylallyl diphosphate (MAPP), 2-pentenyl diphosphate (2-pen-PP), and benzyl diphosphate (benzyl-PP), was tested with two tryptophan prenyltransferases FgaPT2 and 5-DMATS as well as five cyclic dipeptide prenyltransferases BrePT, FtmPT1, AnaPT, CdpNPT, and CdpC3PT in our laboratory (Liebhold et al. 2012, 2013; Liebhold and Li 2013). Our results demonstrated that the tested enzymes also used such donors for alkylation or benzylation. However, the restricted regioselectivity observed with their natural prenyl donor DMAPP has been broken. Partial or complete shifts of alkylation or benzylation were detected in these reactions (Liebhold et al. 2012, 2013; Liebhold and Li 2013). After availability of SirD and TyrPT, we decided to test the behavior of tyrosine *O*-prenyltransferases toward MAPP, 2-pen-PP, and benzyl-PP.

Materials and methods

Chemicals

MAPP, 2-pen-PP, and benzyl-PP were synthesized (Liebhold et al. 2012; Liebhold and Li 2013) following the protocol described for geranyl diphosphate (Woodside et al. 1988). Tyrosine and derivatives used for the enzyme assays were purchased from Bachem (Bubendorf, Switzerland), Sigma-Aldrich (Steinheim, Germany), Alfa Aesar (Karlsruhe, Germany), and TCI (Zwijindrecht, Belgium).

Bacterial strains, plasmids, culture conditions, overproduction, and purification of His-tagged TyrPT and SirD

pGEM-T easy used for cloning and pET28a for expression experiments were obtained from Promega (Mannheim, Germany) and Novagen (Darmstadt, Germany), respectively. *Escherichia coli* XL1 Blue MRF' (Stratagene, Amsterdam, the Netherlands) was used as host for cloning and SoluBL21 (Novagen) for expression experiments. They were grown at 37 °C in liquid Luria-Bertani (LB) or Terrific Broth (TB) medium or on solid LB medium with 1.5 % agar (*w*/*v*). For selection of recombinant *E. coli* strains, 100 μg mL^{-1} of ampicillin or 25 μg mL^{-1} of kanamycin were used.

In this study, SirD was cloned into pET28a and expressed under the control of a strong T7 promoter. Taking the former expression plasmid in pQE70 (pAK2) (Kremer and Li 2010) as template, the coding region of *sirD* was amplified with Expand High Fidelity PCR Kit (Roche Diagnostic, Mannheim, Germany) by using primers SirD_F (5-CC*ATGGCG*ATGCAGACAGCTCGTCTCTTCC-3) at the 5′-end and SirD_R (5-*GCGGCCGC*CTGTCTGTAGCGATTTGGAT-3) at the 3′-end. The underlined letters represent the restriction enzyme sites *Nco*I and *Not*I, respectively, and

italic letters indicate the inserted mutations in comparison to the original gene sequence. PCR amplification was carried out on an iCycler thermal cycler (BioRad, Munich, Germany) with annealing temperature at 56 °C. The PCR fragment of 1360 bp was cloned into pGEM-T easy vector and subsequently sequenced by Eurofins Genomics (Ebersberg, Germany) to confirm the sequence, resulting in the plasmid pHY06. The coding region was then released with *Nco*I and *Not*I and ligated into pET28a vector, which had been digested with the same enzymes previously. The resulted plasmid pHY07 was used for overproduction of SirD. pHC16 described in a previous publication (Fan et al. 2014) was used for production of TyrPT.

For gene expression, *E. coli* SoluBL21 cells harboring pHY07 or pHC16 were cultivated in a 2000-mL cylindrical flask containing 1000 mL liquid TB medium supplemented with kanamycin (25 μg mL^{-1}). After growing at 37 °C to an absorption of 0.6 at 600 nm, the gene expression was induced with isopropyl thiogalactoside at a final concentration of 0.8 mM for TyrPT and 0.5 mM for SirD at 30 °C for 16 h. The bacterial cultures were harvested by centrifugation and resuspended in lysis buffer (10 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) with 2–5 mL g^{-1} wet weight. Lysozyme from the chicken egg white was added at a final concentration of 1 mg mL^{-1}. After incubation on ice for 30 min, the cells were sonicated six times for 10 s each at 200 W. To separate the cellular debris from the soluble proteins, the lysate was centrifuged at 13,000 rpm and 4 °C for 30 min. One-step purification of the recombinant His_6-tagged fusion protein by affinity chromatography with Ni-NTA agarose resin (Macherey-Nagel, Düren, Germany) was carried out according to the manufacturer's instruction. The protein was eluted with 250 mM imidazole in 50 mM NaH_2PO_4 and 300 mM NaCl, pH 8.0. In order to remove imidazole, the protein fraction was passed through a NAP-5 column (GE Healthcare, Freiburg, Germany), which had been equilibrated with 50 mM Tris-HCl (pH 7.5) containing 15 % (*v*/*v*) glycerol. The purified TyrPT and SirD were eluted with the same buffer as for column equilibration and stored at −80 °C for enzyme assays. The protein concentration was determined on Nanodrop C 2000 (Thermo Scientific, Braunschweig, Germany) and then analyzed on 12 % (*w*/*v*) SDS-PAGE (Laemmli 1970) and stained with Coomassie Brilliant Blue G-250.

Enzyme assays for TyrPT and SirD

For determination of the TyrPT and SirD activity, the reaction mixtures (100 μL) contained 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 1 mM aromatic substrate (Table 1), 1 mM DMAPP or analogs, 1.5–5 % (*v*/*v*) glycerol, 0–5 % dimethyl sulfoxide (DMSO), and 20 μg of the purified recombinant protein. After incubation at 37 °C for 16 h, the reactions were terminated by addition of 100 μL methanol. Protein was removed by centrifugation at 13,000 rpm for 20 min before injection on HPLC. Two to four independent assays were carried out routinely. For determination of kinetic parameters of DMAPP analogs, 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 40 μg purified recombinant protein, 1 mM L-tyrosine or 4-amino-L-phenylalanine, and DMAPP analogs at final concentration of 0.01, 0.02, 0.05, 0.08, 0.1, 0.2, 0.5, 0.8, 1.0, 2.0, and 5.0 mM were used in enzyme assays. The reaction mixtures were incubated at 37 °C for 2 h.

Preparative synthesis of prenylated products for structural elucidation

The enzyme assays for isolation of enzymatic products contained 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 1 mM aromatic substrate, 1 mM DMAPP analog, and 2 mg of purified recombinant protein, in a total volume of 10 mL. The reaction mixtures were incubated in glass vials at 37 °C for 16 h, and the reactions were terminated by addition of 1 volume of methanol. After centrifugation at 6000 rpm for 30 min to precipitate protein, the supernatants were concentrated on a rotary evaporator at 35 °C to 1–2 mL and used as samples for separation on HPLC.

HPLC analysis and isolation of enzyme products for structure elucidation

The enzyme assays of TyrPT and SirD were routinely analyzed on an Agilent HPLC series 1200 (Agilent, Böblingen, Germany) by using a Multospher 120 RP-18 column (250×4 mm, 5 μm C + S Chromatographie Service, Langerwehe, Germany) at a flow rate of 1 mL min^{-1}. Double-distilled water (solvent A) and methanol (solvent B) were used as solvents. To analyze the enzyme products, a linear gradient of 30–100 % (*v*/*v*) solvent B in 20 min was used. The column was then washed with 100 % (*v*/*v*) solvent B for 5 min and equilibrated with 30 % (*v*/*v*) solvent B for another 5 min. Detection was carried out by a Photo Diode Array detector and illustrated for absorption at 277 nm in this study.

For structure elucidation, the products were isolated on a Multospher 120 RP 18 column (250×10 mm, 5 μm) with a linear gradient of 60–100 % (*v*/*v*) of methanol (solvent B) in water (solvent A) in 60–80 min and a flow rate of 2.5 mL min^{-1}. After each run, the column was then washed with 100 % (*v*/*v*) solvent B for 10 min and equilibrated with 60 % (*v*/*v*) solvent B for 10 min. The collected fractions were evaporated to dryness and subjected to NMR and mass spectrometry (MS) analyses.

Table 1 Product yields of TyrPT and SirD reactions with tyrosine and derivatives in the presence of DMAPP and analogs

Substrate	DMAPP		MAPP		2-pen-PP		benzyl-PP	
	Conversion (%)		Conversion (%)		Conversion (%)		Conversion (%)	
	TyrPT	SirD	TyrPT	SirD	TyrPT	SirD	TyrPT	SirD
L-Tyrosine (1a)	89.1±6.20	99.3±0.06	21.2±0.72	28.0±2.7	56.6±6.21	44.3±2.90	2.2±0.27	1.3±0.24
4-Amino-L-phenylalanine (2a)	78.1±0.36	66.1±1.38	57.7±1.45	34.8±0.20	46.4±0.52	42.6±2.57	67.2±3.84	54.6±1.54
3-Fluoro-DL-tyrosine (3a)	50.1±0.15	66.5±3.24	13.1±1.15	8.9±0.66	43.9±4.69	20.5±1.23	1.5±0.40	0.4±0.05
3,4-Dihydroxy-L-phenyl-alanine (4a)	100±0.0	88.2±1.28	5.8±0.15	7.3±0.23	30.2±0.18	14.3±0.77	0.4±0.02	0.3±0.10
α-Methyl-L-tyrosine (5a)	98.3±0.03	89.4±0.44	2.9±0.06	16.0±2.11	10.3±0.44	33.5±2.23	0.1±0.02	0.5±0.01
3-Iodo-L-tyrosine (6a)	100±0.0	100±0.0	4.9±0.37	1.1±0.21	12.0±0.89	3.0±0.08	0.2±0.03	0.2±0.04
3,5-Dibromo-L-tyrosine (7a)	20.6±0.30	53.3±4.26	2.8±0.09	0.7±0.15	3.9±0.18	1.4±0.17	0.3±0.08	0.2±0.05
D-Tyrosine (8a)	59.5±0.30	57.6±4.10	0.1±0.04	0.2±0.09	1.7±0.15	0.8±0.08	0.1±0.02	0.2±0.02
3-Amino-3-(4-hydroxyphenyl)-propanoic acid (9a)	10.1±0.18	8.7±0.04	1.7±0.06	1.6±0.02	1.8±0.14	1.8±0.12	2.0±0.04	2.0±0.23
3-(3,4-Dihydroxyphenyl)-DL-serine (10a)	16.1±0.54	29.9±1.75	0.6±0.04	0.5±0.03	7.1±0.64	1.3±0.38	0.1±0.03	<0.01
DL-m-Tyrosine (11a)	1.4±0.02	0.3±0.02	0.7±0.04	0.6±0.02	0.7±0.06	0.6±0.01	0.8±0.01	0.6±0.02
3,5-Diiodo-L-tyrosine (12a)	2.2±0.14	2.2±0.01	0.1±0.03	0.1±0.01	0.1±0.04	0.2±0.07	0.1±0.01	0.1±0.01
3-Nitro-L-tyrosine (13a)	6.6±0.03	44.9±1.71	0.1±0.01	<0.01	0.1±0.02	<0.01	<0.01	<0.01

DMAPP dimethylallyl diphosphate, *MAPP* methylallyl diphosphate, *2-pen-PP* 2-pentenyl diphosphate, *benzyl-PP* benzyl diphosphate

NMR and MS analyses

The isolated enzyme products were dissolved in D_2O or CD_3OD. ^1H NMR spectra were recorded at room temperature on a JEOL ECA-500 spectometer (JEOL, Tokyo, Japan). Two-dimensional spectra HSQC and HMBC were recorded on a Bruker Avance 500 MHz spectrometer (Bruker, Karlsruhe, Germany). All of the spectra were processed with MestReNova 6.0.2 (Metrelab Research, Santiago de Compostella, Spain), and chemical shifts were referenced to the signal of D_2O at 4.79 ppm and that of CD_3OD at 3.30 ppm. The NMR data of isolated products are given in Table 2 and spectra in Figs. S1–S12 in electronic supplementary material.

The isolated compounds were also analyzed by high-resolution electron impact mass spectrometry (HR-EI-MS) on a Micromass Auto Spec spectrometer (Waters, Milford, MA). The obtained MS data are given in Table 3.

Results

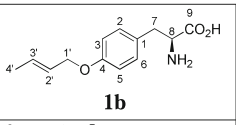
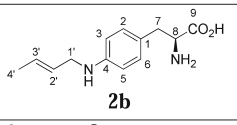
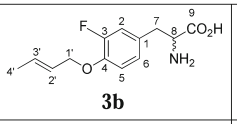
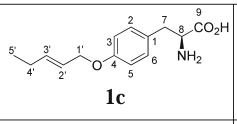
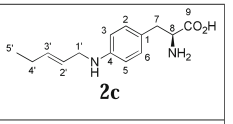
Overproduction and purification of SirD and TyrPT

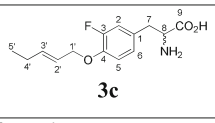
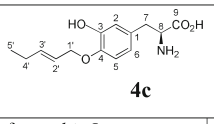
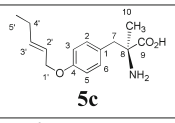
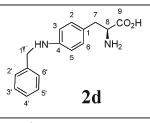
Expression of *sirD* was carried out with the newly constructed plasmid pHY07 in pET28a. Purification of the overproduced SirD led to detection of a predominant protein band on SDS-PAGE (Fig. 2a). With the help of the strong T7 promoter and cultivation in nutrient-rich TB media, the protein yield was improved from 2 to 4 mg L^{-1} bacterial culture. By incubation of L-tyrosine with DMAPP and 2-pen-PP, the activity of SirD from the new construct was proven to be nearly the same of that from the previous construct (data not shown). Therefore, SirD for enzyme assays presented in this study was obtained by using pHY07 as expression vector. TyrPT with a protein yield of 4–6 mg L^{-1} bacterial culture was overproduced and purified as described previously (Fan et al. 2014) (Fig. 2b).

Acceptance of the unnatural DMAPP analogs by tyrosine prenyltransferases

The purified SirD and TyrPT were firstly assayed with DMAPP in the presence of tyrosine (1a) and 12 derivatives (2a–13a). As shown in Table 1, the behaviors of both enzymes for these substrates were similar to those observed previously (Fan et al. 2014; Zou et al. 2011). In comparison to previous works with 3 (Zou et al. 2011) or 8.6 μg proteins per 100 μL assay (Fan et al. 2014), 20 μg were used in this study. As expected, significantly higher conversion yields were obtained for most substrates. The acceptance of 1a–13a by TyrPT and SirD was then tested in the presence of MAPP, 2-pen-PP, and benzyl-PP as alkyl or benzyl acceptors. After incubation with TyrPT and SirD under the same conditions, the reaction mixtures were analyzed on HPLC (Materials and methods). In comparison to negative controls with heat-inactivated proteins (data not shown), product formation was detected in 73 of 78 incubation mixtures (Table 1). Conversion yields of more than 10 % were calculated for 19 enzyme reactions after incubation

Table 2 NMR data of the isolated products 1b–3b, 1c–5c, and 2d

Compound	 **1b**	 **2b**	 **3b**	 **1c**	 **2c**
Position	δ_H, *multi, J*	δ_H, *multi, J*	δ_H, *multi, J*	δ_H, *multi, J*	δ_H, *multi, J*
1	-	-	-	-	-
2	7.19, dd, 8.7, 2.1	7.22, d, 8.5	7.04, dd, 12.5, 2.0	7.19, dd, 8.7, 2.0	7.22, d, 8.4
3	6.88, dd, 8.7, 2.1	6.90, d, 8.5	-	6.88, dd, 8.7, 2.1	6.90, d, 8.4
4	-	-	-	-	-
5	6.88, dd, 8.7, 2.1	6.90, d, 8.5	7.01, t, 8.0	6.88, dd, 8.7, 2.1	6.90, d, 8.4
6	7.19, dd, 8.7, 2.1	7.22, d, 8.5	6.98, dd, 8.5, 1.8	7.19, dd, 8.7, 2.0	7.22, d, 8.4
7	3.23, dd, 14.7, 4.3 2.93, dd, 14.7, 8.8	3.25, dd, 14.7, 5.1 3.09, dd,14.7, 7.9	3.19, dd, 14.7, 4.5 2.94, dd, 14.7, 8.4	3.22, dd, 14.7, 4.3 2.93, dd, 14.7, 8.8	3.21, dd, 14.6, 5.1 3.05, dd, 14.6, 7.9
8	3.71, dd, 8.8, 4.3	3.99, dd, 7.9, 5.1	3.74, dd, 8.4, 4.5	3.70, dd, 8.8, 4.3	3.93, dd, 7.9, 5.1
1’	4.44, dq, 6.0, 1.2	approx. 3.75*	4.49, d, 6.1	4.46, dd, 6.0, 1.2	3.75, d, 6.1
2’	5.68, dtq, 15.3, 6.0, 1.6	5.66, dtq, 15.4, 6.1, 1.7	5.68, dtq, 15.3, 6.1, 1.7	5.87, dtt, 15.4, 6.4, 1.3	5.89, dbrt, 15.5, 6.2
3’	5.84, dqt, 15.3, 6.5, 1.4	5.83, dqt, 15.4, 6.5, 1.4	5.85, dqt, 15.3, 6.5, 1.3	5.66, dtt, 15.4, 6.0, 1.6	5.64, dbrt, 15.5, 6.2
4’	1.72, dd, 6.5, 1.3	1.73, dd, 6.5, 1.4	1.71, dd, 6.5, 1.2	2.09, dqq, 7.2, 6.9, 1.2	2.10, dq, 7.5, 7.4
5’	-	-	-	1.01, t, 7.5	1.01, t, 7.5
Solvent	CD_3OD	D_2O	CD_3OD	CD_3OD	D_2O

Compound	 **3c**	 **4c**		 **5c**	 **2d**
Position	δ_H, *multi, J*	δ_H, *multi, J*	δ_C	δ_H, *multi, J*	δ_H, *multi, J*
1	-	-	128.9	-	-
2	7.03, dd, 12.4, 2.0	6.86, brs	116.6	7.14, d, 8.6	7.18, d, 8.4
3	-	-	145.8	6.83, d, 8.7	6.90, d, 8.5
4	-	-	145.4	-	-
5	7.00, t, 8.0	7.06, d, 8.3	115.3	6.83, d, 8.7	6.90, d, 8.5
6	6.97, dd,8.4, 1.9	6.82, d, 8.3	121.3	7.14, d, 8.6	7.18, d, 8.4
7	3.13, dd, 14.6, 4.7 2.88, dd, 14.6, 7.9	3.19, dd, 14.6, 5.0 3.01, dd, 14.6, 7.9	35.8	3.16, d, 14.2 2.77, d, 14.2	3.22, dd, 14.7, 5.1 3.04, dd, 14.7, 8.0
8	3.63, dd, 7.9, 4.7	3.93, dd, 7.6, 5.1	55.9	-	3.95, dd, 8.0, 5.1
9	-	-	174.5	-	-
10	-	-	-	1.42, s	-
1’	4.49, dd, 6.1, 1.1	4.61, d, 6.3	70.3	4.42, d, 5.9	-
2’	5.87, dtt, 15.4, 6.4, 1.3	6.00, d brt, 15.4, 6.2	139.2	5.84, d brt, 15.4, 6.2	7.49, d, 7.4
3’	5.65, dtt, 15.4, 6.1, 1.6	5.73, d brt, 15.4, 6.2	123.1	5.63, d brt, 15,4, 6.0	7.46, t, 7.2
4’	2.07, dq, 7.6, 7.5	2.11, dq, 7.2, 7.0	24.6	2.05, dq, 7.5, 7.4	7.38, t, 7.2
5’	0.99, t, 7.5	1.01, t, 7.5	12.2	0.98, t, 7.5	7.46, t, 7.2
6’	-	-	-	-	7.49, d, 7.4
7’	-	-	-	-	4.43, s
Solvent	CD_3OD	D_2O	D_2O	CD_3OD	D_2O

The spectra were taken on a JEOL ECA-500 spectrometer. Chemical shifts (δ) are given in ppm and coupling constants in hertz

brt broad triplet

* Signals overlapping with ethanol (impurity)

with 20 μg protein at 37 °C for 16 h. Inspection of the HPLC chromatograms of the reaction mixtures revealed the presence of one product peak each. HPLC chromatograms of nine selected enzyme assays are illustrated in Fig. 3.

As given in Table 1, TyrPT and SirD displayed similar preference for tyrosine and derivatives in the presence of all three DMAPP analogs. Higher conversion yields were detected for incubations of TyrPT and SirD in the presence of 2-pen-PP than MAPP, corresponding to those observed with indole prenyltransferases (Liebhold et al. 2012, 2013). 1a and 2a were well accepted by TyrPT and SirD in the presence of MAPP with conversion yields of more than 20 %, while 1a–5a were good acceptors for 2-pen-PP. Conversion yields of more than 20 % were calculated for incubations of 1a–4a with TyrPT and 1a–3a and 5a with SirD. In comparison to MAPP and 2-pen-PP, benzyl-PP was a poor donor for both enzymes. Conversion yields of 67.2 and 54.6 % were detected for 2a with TyrPT and SirD, respectively. Much lower conversion yields of less than 3 % were found for other substrates. As given in Table 1, 2a was accepted as the best substrate by TyrPT and SirD in the presence of MAPP and benzyl-PP. Comparable acceptance was observed for 1a and 2a by both enzymes in the presence of 2-pen-PP. Overall, 2a was the best or one of the best substrates for both TyrPT and SirD in the presence DMAPP analogs, which differs clearly from their preference toward aromatic substrates in the presence of DMAPP. Using DMAPP as prenyl donor, TyrPT showed higher activities toward 4a–6a than 2a and SirD accepted 6a

Table 3 HR-EI-MS data of the isolated enzyme products

Product	Formula	Calculated (M_r)	Measured $[M]^+$	Deviation (ppm)
1b	$C_{13}H_{17}NO_3$	235.1208	235.1215	−3.0
2b	$C_{13}H_{18}N_2O_2$	234.1368	234.1379	−4.5
3b	$C_{13}H_{16}FNO_3$	253.1114	253.1109	2.0
1c	$C_{14}H_{19}NO_3$	249.1365	249.1378	−5.2
2c	$C_{14}H_{20}N_2O_2$	248.1525	248.1530	−2.0
3c	$C_{14}H_{18}FNO_3$	267.1271	267.1285	−5.2
4c	$C_{14}H_{19}NO_4$	265.1314	265.1323	−3.4
5c	$C_{15}H_{21}NO_3$	263.1521	263.1533	−4.6
2d	$C_{16}H_{18}N_2O_2$	270.1368	270.1384	−5.9

much better than 2a (Table 1). It seems that the 4-amino group of 2a had a better interaction with the enzymes in the presence of the unnatural donors, so that the cation intermediates (Liebhold et al. 2012) are more stable than in the cases of other aromatic substrates.

In a previous study, TyrPT and SirD were demonstrated to be low stereoselective in the presence of DMAPP as prenyl donor (Fan et al. 2014). D-Tyrosine (8a) was accepted by TyrPT and SirD with relative activities of 50 and 30 % of those of 1a, respectively, which was also confirmed in this study (Table 1). As given in Table 1, 8a was a very poor substrate for both enzymes in the presence of the three DMAPP analogs, and conversion yields of less than 2 % were detected in this study. It can be speculated that in the presence of the unnatural alkyl or benzyl donors, both enzymes had changed

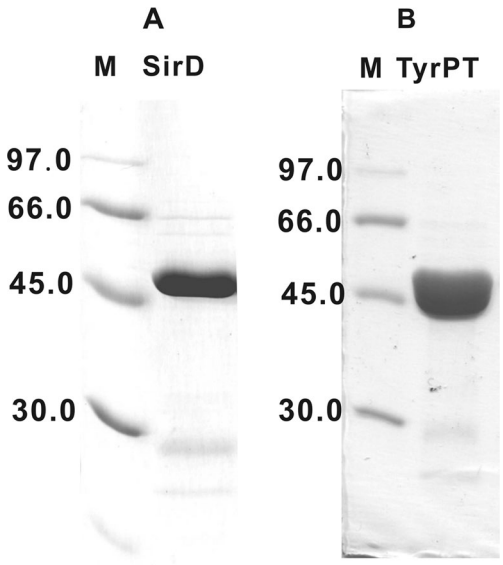


Fig. 2 SDS-PAGE analysis of purified TyrPT-His_6 and SirD-His_6. The proteins were separated on a 12 % SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue G-250. **a** Purified SirD and **b** purified TyrPT; *M* molecular mass standard

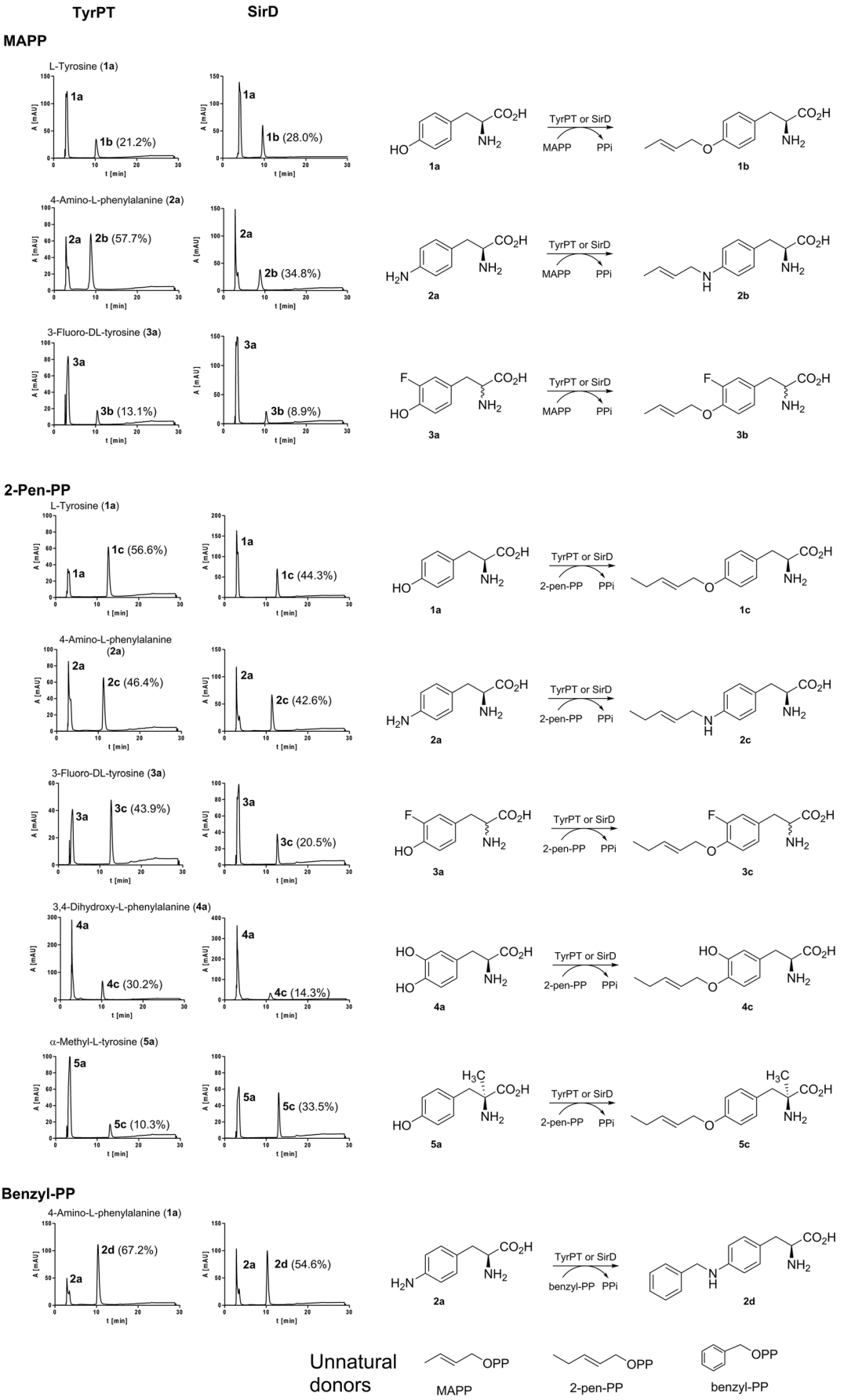
Fig. 3 HPLC analysis of the incubation mixtures of selected substrates with TyrPT (*left column*) and SirD (*middle column*) as well as the transfer reactions catalyzed by TyrPT and SirD (*right column*). The reaction mixtures contained 20 μg TyrPT or SirD, 1.0 mM of aromatic substrate, 1.0 mM DMAPP analog, 10 mM $CaCl_2$, and 50 mM Tris-HCl (pH 7.5) and were incubated at 37 °C for 16 h. The illustrated chromatograms were recorded at 277 nm. Product yields are given in *parentheses* after the product numbers

their conformations, so that 8a was much poorly fixed into the reaction cavities of both enzymes than 1a.

TyrPT and SirD catalyze the same *O*-/*N*-alkylation/benzylation of tyrosine and derivatives

For structure elucidation, nine enzyme products were isolated on preparative HPLC from the incubation mixtures of TyrPT and SirD in different combinations of tyrosine and derivatives with the three DMAPP analogs. 1b–3b were obtained from the incubation mixtures of 1a–3a with MAPP, respectively, 1c–5c from those of 1a–5a with 2-pen-PP, and 2d from that of 2a with benzyl-PP. The obtained products were subjected to MS and NMR analyses. Identical ^{1}H NMR spectra of the enzyme products were obtained from incubation mixtures of TyrPT and SirD with a given substrate illustrated in Fig. 3 (data not shown), proving the same function of both enzymes in the presence of unnatural DMAPP analogs.

Molecular formula deduced from high-resolution electron impact mass spectrometry (Table 3) proved the nine isolated products to be monoprenylated or benzylated derivatives of the respective substrate. Inspection of the ^{1}H NMR spectra (Figs. S1–S12 in the Supplementary Material) of 1b–3b, 1c–5c, and 2d revealed the presence of only one product each. The coupling pattern of the signals for the methylallyl and 2-pentenyl residues are comparable to those of previously reported compounds with a regular alkyl moiety (Liebhold et al. 2012), proving their attachment in a regular pattern. In comparison to those of the respective substrate, the chemical shifts, number, and coupling pattern of the signals for protons at the aromatic ring and the side chain of the isolated products were nearly unchanged and are similar to those of the prenylated derivatives with DMAPP (Fan et al. 2014; Kremer and Li 2010; Zou et al. 2011). All these observations suggested the alkylation at the hydroxyl or amino group of the aromatic ring. The chemical shifts of H-1' of the alkyl moieties for products 1b, 1c, 3b, 3c, and 5c between 4.44 and 4.49 ppm in CD_3OD and for 4c at 4.61 ppm in D_2O proved unequivocally their attachment to *O*-atoms (Fan et al. 2014; Kremer and Li 2010; Zou et al. 2011). In the case of 3,4-dihydroxy-L-phenylalanine (4a), *O*-alkylation at position C-3 or C-4 would be possible. The correlations found between H-1' and C-4 in the HMBC spectra of 4c (Fig. 4 and Figs. S9 and S10 in the Supplementary Material) revealed that the 2-pentenyl moiety was attached to *O*-atom at position C-4


TyrPT
SirD
MAPP
L-Tyrosine (1a)
1a
1b (21.2%)
1b (28.0%)
4-Amino-L-phenylalanine (2a)
2a
2b (57.7%)
2b (34.8%)
3-Fluoro-DL-tyrosine (3a)
3a
3b (13.1%)
3b (8.9%)
2-Pen-PP
1c (56.6%)
1c (44.3%)
2c (46.4%)
2c (42.6%)
3c (43.9%)
3c (20.5%)
3,4-Dihydroxy-L-phenylalanine (4a)
4a
4c (30.2%)
4c (14.3%)
α-Methyl-L-tyrosine (5a)
5a
5c (10.3%)
5c (33.5%)
Benzyl-PP
2d (67.2%)
2d (54.6%)
TyrPT or SirD
MAPP
PPi
2-pen-PP
benzyl-PP
t [min]
A [mAU]
Unnatural donors
OPP
MAPP
2-pen-PP
benzyl-PP

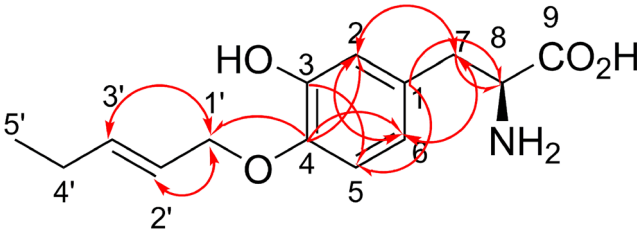


Fig. 4 Selected HMBC correlations observed in 4c

of the aromatic ring. The chemical shifts of H-1' of the alkyl moiety at 3.75 ppm were observed in the spectra of the products 2b and 2c from the incubation mixtures of 2a with MAPP and 2-pen-PP, proving an *N*-alkylation in their structures (Fan et al. 2014; Zou et al. 2011). In the spectrum of 2d from the incubation mixture of 2a with benzyl-PP, the signals for both benzene rings have been nearly unchanged and the signal of the methylene group was detected at 4.43 ppm. These data proved unequivocally the benzylation of the amino group at the aromatic ring. In conclusion, both TyrPT and SirD accepted unnatural alkyl/benzyl donors and catalyzed the *O*-/*N*-alkylation/benzylation of tyrosine and derivatives at the same position as in the presence of their natural prenyl donor DMAPP.

Kinetic parameters of TyrPT and SirD toward the unnatural DMAPP analogs

To compare the catalytic efficiencies of TyrPT and SirD in the presence of unnatural alkyl or benzyl donors with those of their natural prenyl donor DMAPP, kinetic parameters including Michaelis-Menten constant (K_M) and turnover number (k_{cat}) were determined for MAPP and 2-pen-PP by using 1a as substrate (Table 4 and Figs. S13–S16 in the Supplementary Material). Due to the low conversions of benzyl-PP with 1a, its kinetic parameters were obtained from incubation mixtures with 2a (Table 4 and Figs. S17 and S18 in the Supplementary Material). MAPP was accepted by both enzymes with very small K_M values, lower than those of 2-pen-PP, benzyl-PP, and DMAPP. All three unnatural donors showed much lower turnover numbers than those of DMAPP.

Discussion

As aforementioned, TyrPT and SirD belong to the prenyltransferases of the DMATS superfamily. Similar to other members of this family, TyrPT and SirD showed high substrate flexibility toward tyrosine and derivatives and catalyzed regioselective transfer reactions, i.e., the 4-*O*- or 4-*N*-prenylation (Fan et al. 2014; Kremer and Li 2010; Zou et al. 2011). Both enzymes also accepted tryptophan and derivatives as substrates and mainly catalyzed the *C7*-prenylation at the indole ring (Fan et al. 2014; Kremer and Li 2010; Rudolf and Poulter 2013). In comparison, the tryptophan *C4*-prenyltransferase FgaPT2 and *C7*-prenyltransferase 7-DMATS also accepted L-tyrosine and a few number of its derivatives as prenylation substrates, demonstrating the complementary substrate promiscuity of tryptophan and tyrosine prenyltransferases. However, differing from the two tyrosine *O*-prenyltransferases TyrPT and SirD with the same *C7*-prenylated product from L-tryptophan, FgaPT2 catalyzed a *C3*- and 7-DMATS an *O*-prenylation of tyrosine, respectively (Fan et al. 2015; Fan and Li 2014). Similar to most of the DMATS enzymes, these two enzymes showed strict substrate specificity toward DMAPP and did not accepted geranyl diphosphate as prenyl donor (Kremer and Li 2010) (data for TyrPT not shown). Acceptance of MAPP, 2-pen-PP, and benzyl-PP by TyrPT and SirD presented in this study provided experimental evidence on their substrate promiscuity for prenyl donor and expanded their potential usage for increasing the structural diversity of alkylated compounds in the chemoenzymatic synthesis.

In previous studies, we have demonstrated that several indole prenyltransferases including tryptophan and tryptophan-containing cyclic dipeptide prenyltransferases used MAPP and 2-pen-PP as alkyl donors and the tryptophan *C4*-

Table 4 Comparison of kinetic parameters of TyrPT and SirD toward alkyl/benzyl donors

Donor	TyrPT			SirD		
	K_M (mM)	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1} mM^{-1})	K_M (mM)	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1} mM^{-1})
DMAPP[a]	0.71	41.4	58.3	0.17	60	352.9
MAPP[b]	0.029±0.0036	0.12±0.013	4.14	0.13±0.025	0.16±0.03	1.23
2-pen-PP[b]	0.23±0.052	0.20±0.022	0.87	0.22±0.053	0.13±0.004	0.59
benzyl-PP[c]	0.52±0.022	0.96±0.003	1.85	0.75±0.0021	1.68±0.01	2.24

DMAPP dimethylallyl diphosphate, *MAPP* methylallyl diphosphate, *2-pen-PP* 2-pentenyl diphosphate, *benzyl-PP* benzyl diphosphate

[a] Data for DMAPP were adopted from previous studies (Fan et al. 2014; Kremer and Li 2010)

[b] Data were obtained from four independent assays by using 1a as aromatic substrate

[c] Data were obtained from two independent assays by using 2a as aromatic substrate

prenyltransferase FgaPT2 even used benzyl-PP for benzylation (Liebhold et al. 2012, 2013; Liebhold and Li 2013). These enzymes catalyzed regiospecific and stereospecific prenyl transfer reactions from DMAPP onto indole rings (Grundmann and Li 2005; Unsöld and Li 2005; Yin et al. 2009, 2010, 2013; Yu and Li 2012). In the presence of the unnatural DMAPP analogs, the alkylation/benzylation positions were partially or completely shifted to the neighboring position, e.g., from exclusively at C-2 with *C2*-prenyltransferases BrePT and FtmPT1 to both C-2 and C-3 (Liebhold et al. 2013). *C2*- and *C3*-alkylations with MAPP and 2-pen-PP were also observed for the three *C3*-prenyltransferases AnaPT, CdpNPT, and CdpC3PT (Liebhold et al. 2013). The alkylation positions of the two tryptophan prenyltransferases FgaPT2 and 5-DMATS were partially (in the presence of MAPP) or completely (in the presence of 2-pen-PP) shifted (Liebhold et al. 2012). The tryptophan *C4*-prenyltransferase FgaPT2 catalyzed the benzylation at C-5 of the indole ring (Liebhold and Li 2013). As shown in Fig. 3, only one enzyme product each was detected in the incubation mixtures of selected well accepted substrates and identified as *4-O*- or *4-N*-alkylated or benzylated derivative of the respective substrate. This means that the TyrPT and SirD products with the unnatural DMAPP analogs carry the alkyl/benzyl moieties at the same position as their products with the natural donor DMAPP, although alkylation at other positions of tyrosine and derivatives such as C-3 would be possible. Together with N-acetyl-3-prenyl-L-tyrosine, 3-dimethylallyl-L-tyrosine was isolated from *Streptomyces* sp. IFM 10937 (Ahmed et al. 2008). Members of the DMATS superfamily are also able to catalyze a *C3*-prenylation of tyrosine or derivatives, as demonstrated very recently with FgaPT2 and especially with its mutant FgaPT2_K174F by using L-tyrosine (1a) and 4-amino-L-phenylalanine (2a) as substrates (Fan et al. 2015). This high regioselectivity of TyrPT and SirD toward tyrosine and derivatives with different alkyl donors would be of essential importance for controlled structure modification.

As mentioned in the "Introduction" section, SirD is involved in the biosynthesis of sirodesmin (Fig. 1) (Gardiner et al. 2004; Kremer and Li 2010) and the genetic context of TyrPT in *A. niger* is unknown (Fig. 1) (Fan et al. 2014). As shown in this and in the previous studies (Fan et al. 2014; Kremer and Li 2010; Zou et al. 2011), both enzymes display similar substrate specificities toward donors and acceptors. It can be speculated that the structures of these two enzymes are highly conserved. Furthermore, both DMAPP and the tested DMAPP analogs are placed in their reaction chambers in similar positions and undergo similar reactions.

Determination of the kinetic parameters revealed that TyrPT and SirD showed high affinity to MAPP and 2-pen-PP (Table 4). However, the turnover numbers of the TyrPT and SirD reactions with the three unnatural DMAPP analogs were found between 0.1 and 1.0 min^{-1}, comparable to those of the cyclic dipeptide prenyltransferases (Liebhold et al. 2013) and lower than those of tryptophan prenyltransferases with these DMAPP analogs (Liebhold et al. 2012). The low catalytic efficiency of TyrPT and SirD toward MAPP, 2-pen-PP, and benzyl-PP should be improved in the future by site-directed mutagenesis experiments after availability of the structures of TyrPT or/and SirD.

Acknowledgments We thank Nina Zitzer and Stefan Newel for taking MS and NMR spectra, respectively. This work was supported in part by grants from the Deutsche Forschungsgemeinschaft (Li844/4-1 to S.-M L.). Huili Yu is a recipient of a scholarship from the China Scholarship Council.

References

Ahmed F, Ohtsuki T, Aida W, Ishibashi M (2008) Tyrosine derivatives isolated from *Streptomyces* sp. IFM 10937 in a screening program for TRAIL-resistance-overcoming activity. J Nat Prod 71:1963–1966

Chooi YH, Cacho R, Tang Y (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chem Biol 17:483–494

Chooi YH, Wang P, Fang J, Li Y, Wu K, Wang P, Tang Y (2012) Discovery and characterization of a group of fungal polycyclic polyketide prenyltransferases. J Am Chem Soc 134:9428–9437

Fan A, Li S-M (2014) Prenylation of tyrosine and derivatives by a tryptophan *C*7-prenyltransferase. Tetrahedron Lett 55:5199–5202

Fan A, Chen H, Wu R, Xu H, Li S-M (2014) A new member of the DMATS superfamily from *Aspergillus niger* catalyzes prenylations of both tyrosine and tryptophan derivatives. Appl Microbiol Biotechnol 98:10119–10129

Fan A, Zocher G, Stec E, Stehle T, Li S-M (2015) Site-directed mutagenesis switching a dimethylallyl tryptophan synthase to a specific tyrosine C3-prenylating enzyme. J Biol Chem 290:1364–1373

Gardiner DM, Cozijnsen AJ, Wilson LM, Pedras MS, Howlett BJ (2004) The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. Mol Microbiol 53:1307–1318

Grundmann A, Li S-M (2005) Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from *Aspergillus fumigatus*. Microbiology 151:2199–2207

Heide L (2009) Prenyl transfer to aromatic substrates: genetics and enzymology. Curr Opin Chem Biol 13:171–179

Kremer A, Li S-M (2010) A tyrosine *O*-prenyltransferase catalyses the first pathway-specific step in the biosynthesis of sirodesmin PL. Microbiology 156:278–286

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Li S-M (2010) Prenylated indole derivatives from fungi: structure diversity, biological activities, biosynthesis and chemoenzymatic synthesis. Nat Prod Rep 27:57–78

Liebhold M, Li S-M (2013) Regiospecific benzylation of tryptophan and derivatives catalyzed by a fungal dimethylallyl transferase. Org Lett 15:5834–5837

Liebhold M, Xie X, Li S-M (2012) Expansion of enzymatic Friedel-Crafts alkylation on indoles: acceptance of unnatural beta-unsaturated allyl diphospates by dimethylallyl-tryptophan synthases. Org Lett 14:4884–4885

Liebhold M, Xie X, Li S-M (2013) Breaking cyclic dipeptide prenyltransferase regioselectivity by unnatural alkyl donors. Org Lett 15:3062–3065

Pockrandt D, Li S-M (2013) Geranylation of cyclic dipeptides by the dimethylallyl transferase AnaPT resulting in a shift of prenylation position on the indole ring. Chembiochem 14:2023–2028

Pockrandt D, Ludwig L, Fan A, König GM, Li S-M (2012) New insights into the biosynthesis of prenylated xanthones: XptB from *Aspergillus nidulans* catalyses an *O*-prenylation of xanthones. Chembiochem 13:2764–2771

Pockrandt D, Sack C, Kosiol T, Li S-M (2014) A promiscuous prenyltransferase from *Aspergillus oryzae* catalyses C-prenylations of hydroxynaphthalenes in the presence of different prenyl donors. Appl Microbiol Biotechnol 98:4987–4994

Rudolf JD, Poulter CD (2013) Tyrosine *O*-prenyltransferase SirD catalyzes *S*-, *C*-, and *N*-prenylations on tyrosine and tryptophan derivatives. ACS Chem Biol 8:2707–2714

Tarcz S, Xie X, Li S-M (2014) Substrate and catalytic promiscuity of secondary metabolite enzymes: *O*-prenylation of hydroxyxanthones with different prenyl donors by a bisindolyl benzoquinone *C*- and *N*-prenyltransferase. RSC Adv 4:17986–17992

Unsöld IA, Li S-M (2005) Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from *Aspergillus fumigatus*. Microbiology 151:1499–1505

Wollinsky B, Ludwig L, Hamacher A, Yu X, Kassack MU, Li S-M (2012) Prenylation at the indole ring leads to a significant increase of cytotoxicity of tryptophan-containing cyclic dipeptides. Bioorg Med Chem Lett 22:3866–3869

Woodside AB, Huang Z, Poulter CD (1988) Trisammonium geranyl diphosphate. Org Synth 66:211–215

Wunsch C, Zou HX, Linne U, Li S-M (2015) *C7*-prenylation of tryptophanyl and *O*-prenylation of tyrosyl residues in dipeptides by an *Aspergillus terreus* prenyltransferase. Appl Microbiol Biotechnol 99:1719–1730

Yazaki K, Sasaki K, Tsurumaru Y (2009) Prenylation of aromatic compounds, a key diversification of plant secondary metabolites. Phytochemistry 70:1739–1745

Yin W-B, Grundmann A, Cheng J, Li S-M (2009) 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 284:100–109

Yin W-B, Yu X, Xie X-L, Li S-M (2010) Preparation of pyrrolo[2,3-*b*]indoles carrying a ß-configured reverse *C3*-dimethylallyl moiety by using a recombinant prenyltransferase CdpC3PT. Org Biomol Chem 8:2430–2438

Yin S, Yu X, Wang Q, Liu XQ, Li S-M (2013) Identification of a brevianamide F reverse prenyltransferase BrePT from *Aspergillus versicolor* with a broad substrate specificity towards tryptophan-containing cyclic dipeptides. Appl Microbiol Biotechnol 97:1649–1660

Yu X, Li S-M (2011) Prenylation of flavonoids by using a dimethylallyltryptophan synthase 7-DMATS from *Aspergillus fumigatus*. Chembiochem 12:2280–2283

Yu X, Li S-M (2012) Prenyltransferases of the dimethylallyltryptophan synthase superfamily. Methods Enzymol 516:259–278

Yu X, Xie X, Li S-M (2011) Substrate promiscuity of secondary metabolite enzymes: prenylation of hydroxynaphthalenes by fungal indole prenyltransferases. Appl Microbiol Biotechnol 92:737–748

Zou H-X, Xie X, Zheng X-D, Li S-M (2011) The tyrosine *O*-prenyltransferase SirD catalyzes *O*-, *N*-, and *C*-prenylations. Appl Microbiol Biotechnol 89:1443–1451

Electronic supplementary material for:

Tyrosine *O*-prenyltransferases TyrPT and SirD displaying similar behavior towards unnatural alkyl or benzyl diphosphate as their natural prenyl donor dimethylallyl diphosphate

Huili Yu[a], Mike Liebhold[a], Xiulan Xie[b], and Shu-Ming Li[a]*

[a]Institut für Pharmazeutische Biologie und Biotechnologie, Philipps-Universität Marburg, Deutschhausstrasse 17A, 35037 Marburg, Germany

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

*Correspondence to Shu-Ming Li: shuming.li@staff.uni-marburg.de;

Telefon: 49(0) 6421 282-2461; Fax: 49(0)6421 282-5365

Content

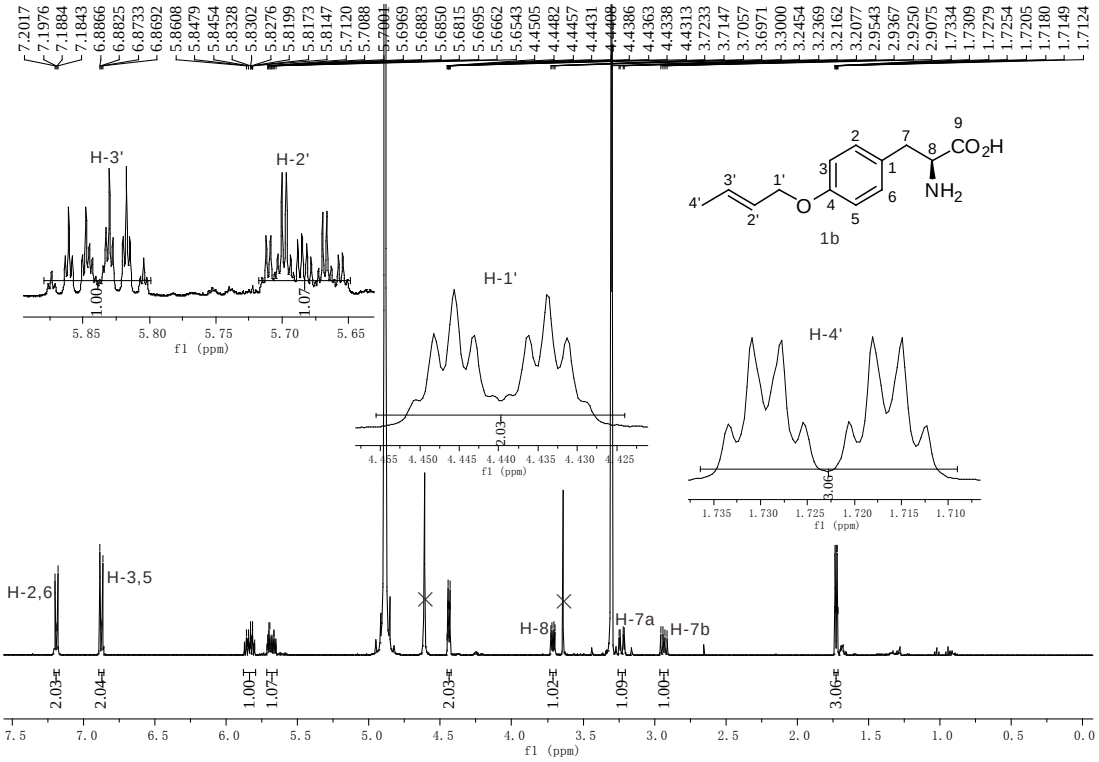


Fig. S1 ^{1}H NMR spectrum of **1b** in CD_3OD.

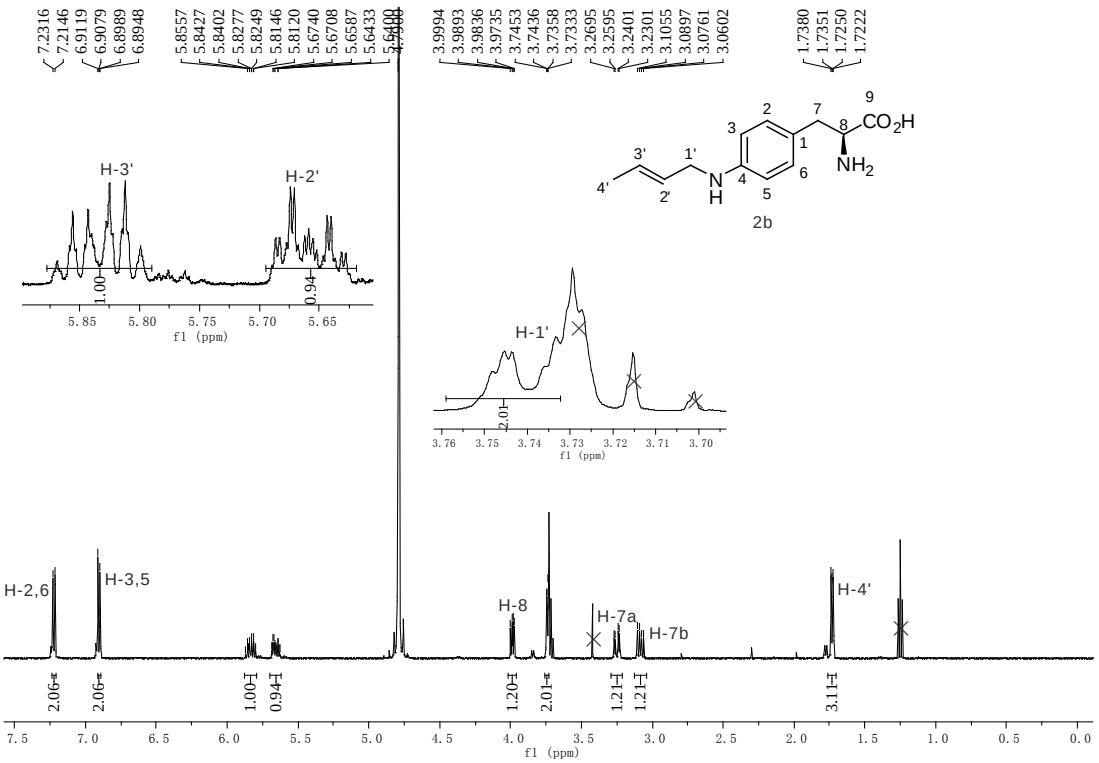


Fig. S2 ^{1}H NMR spectrum of **2b** in D_2O.

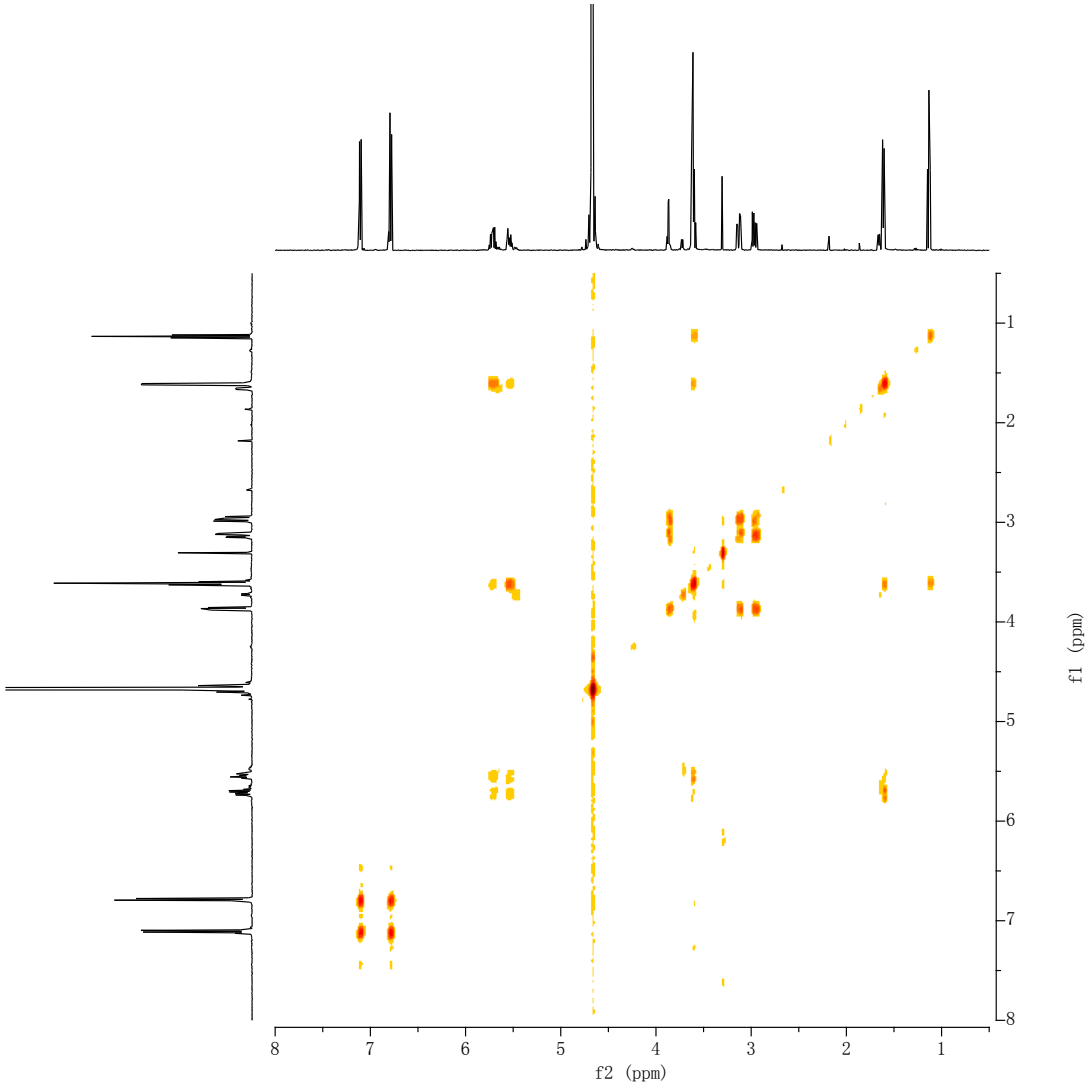


Fig. S3 H-H COSY spectrum of **2b** in D_2O.

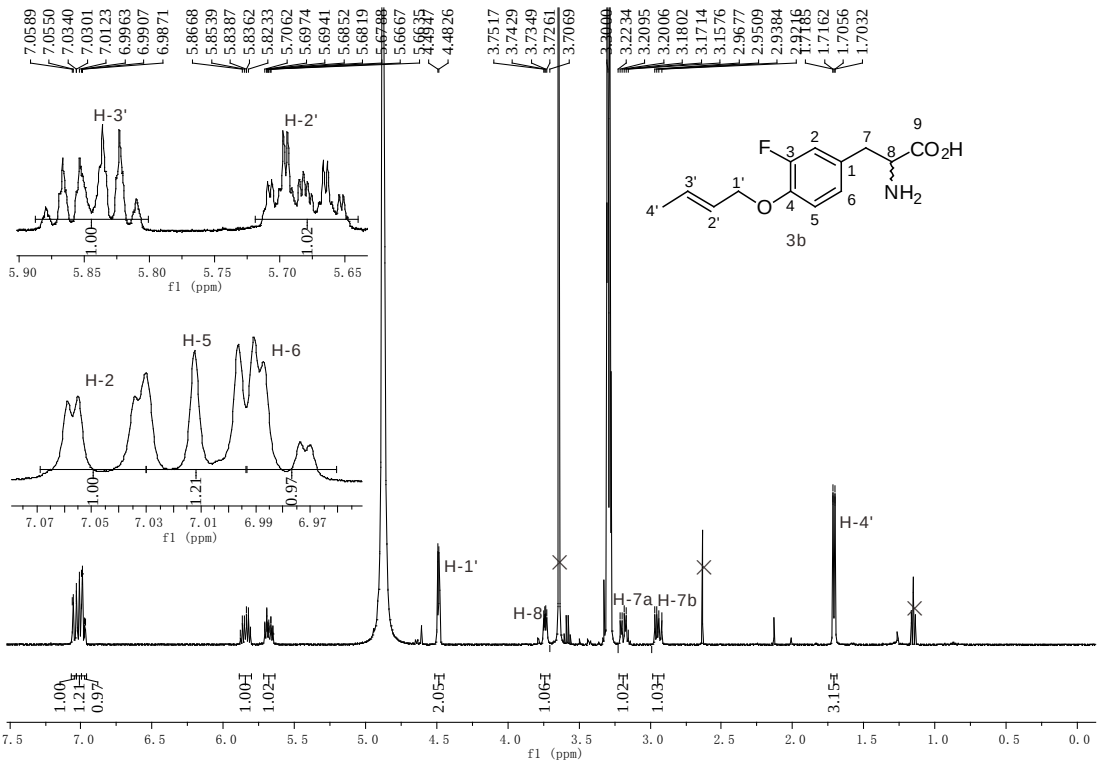


Fig. S4 ^{1}H NMR spectrum of **3b** in CD_3OD.

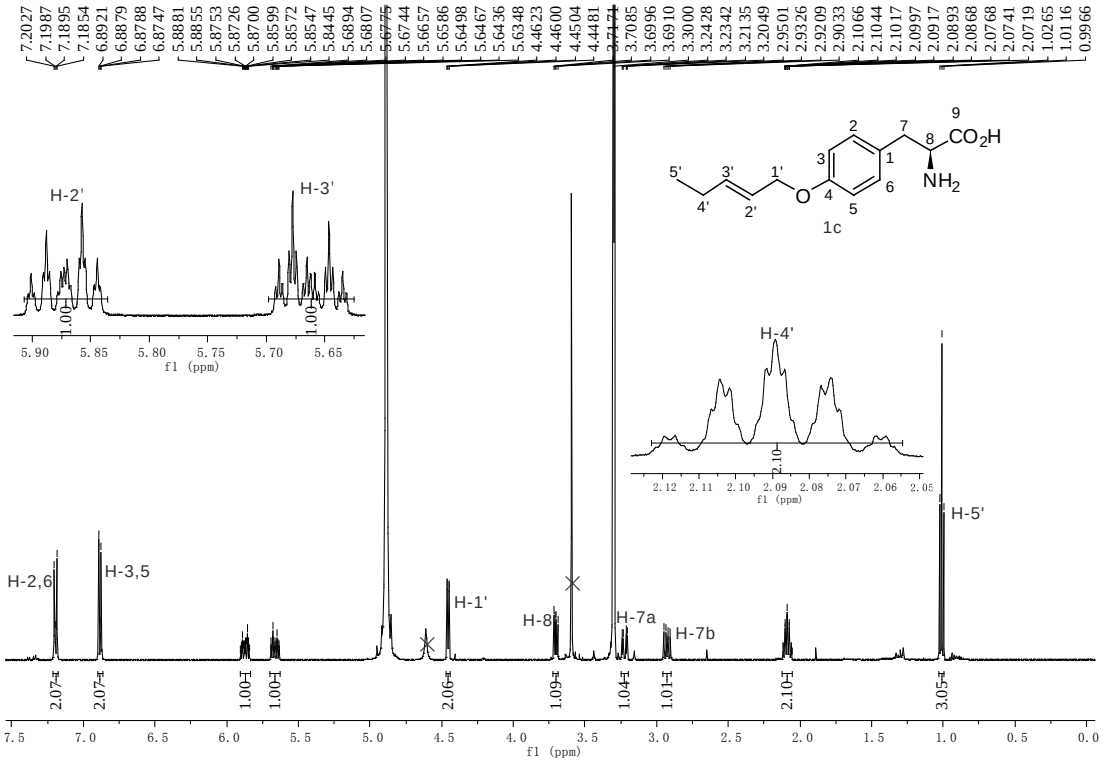


Fig. S5 ^{1}H NMR spectrum of **1c** in CD_3OD.

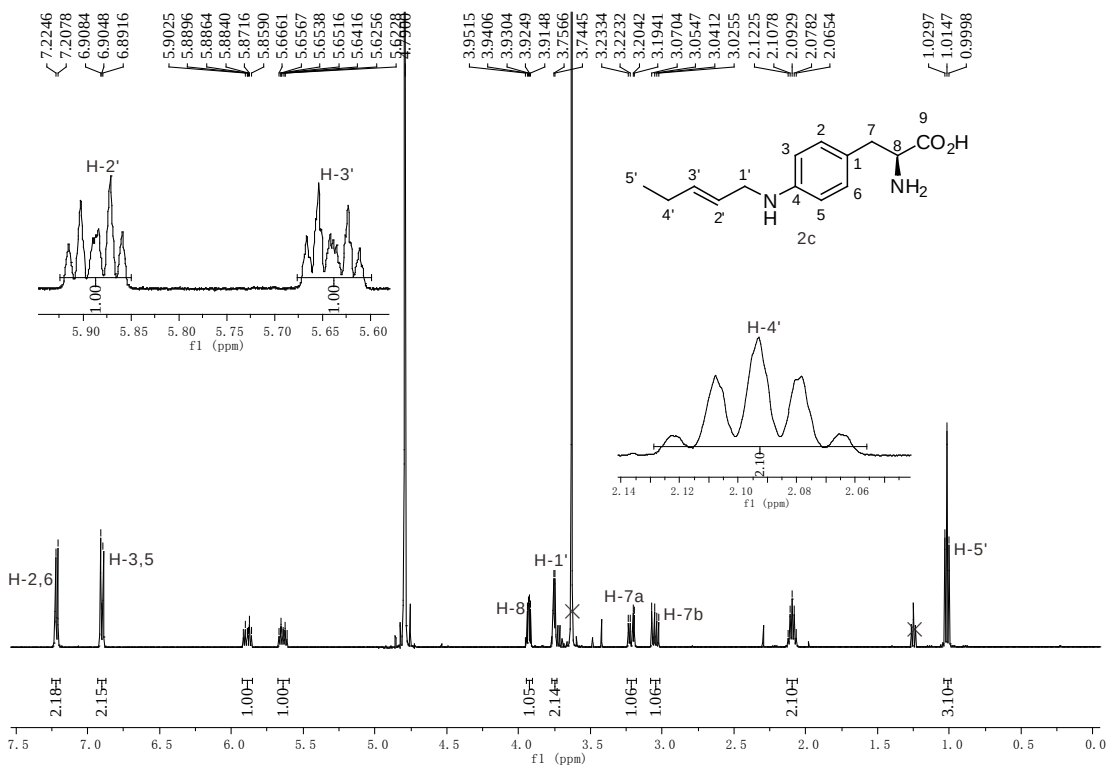


Fig. S6 ^{1}H NMR spectrum of **2c** in D_2O.

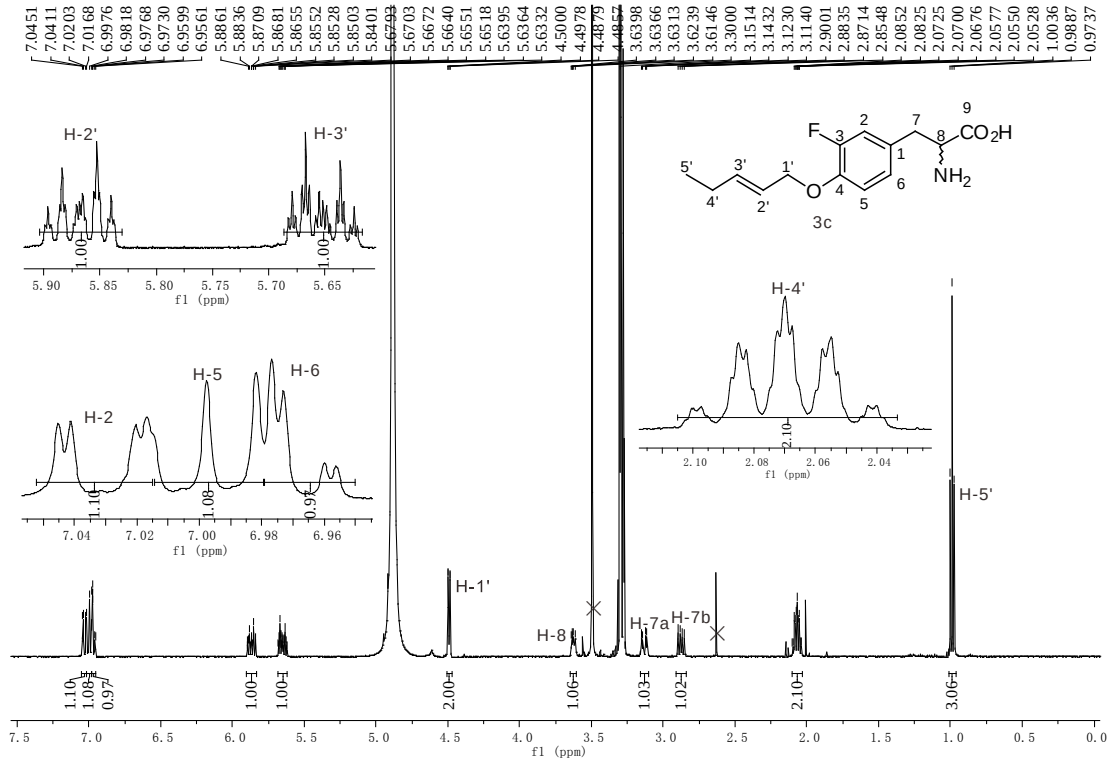


Fig. S7 ^{1}H NMR spectrum of **3c** in CD_3OD.

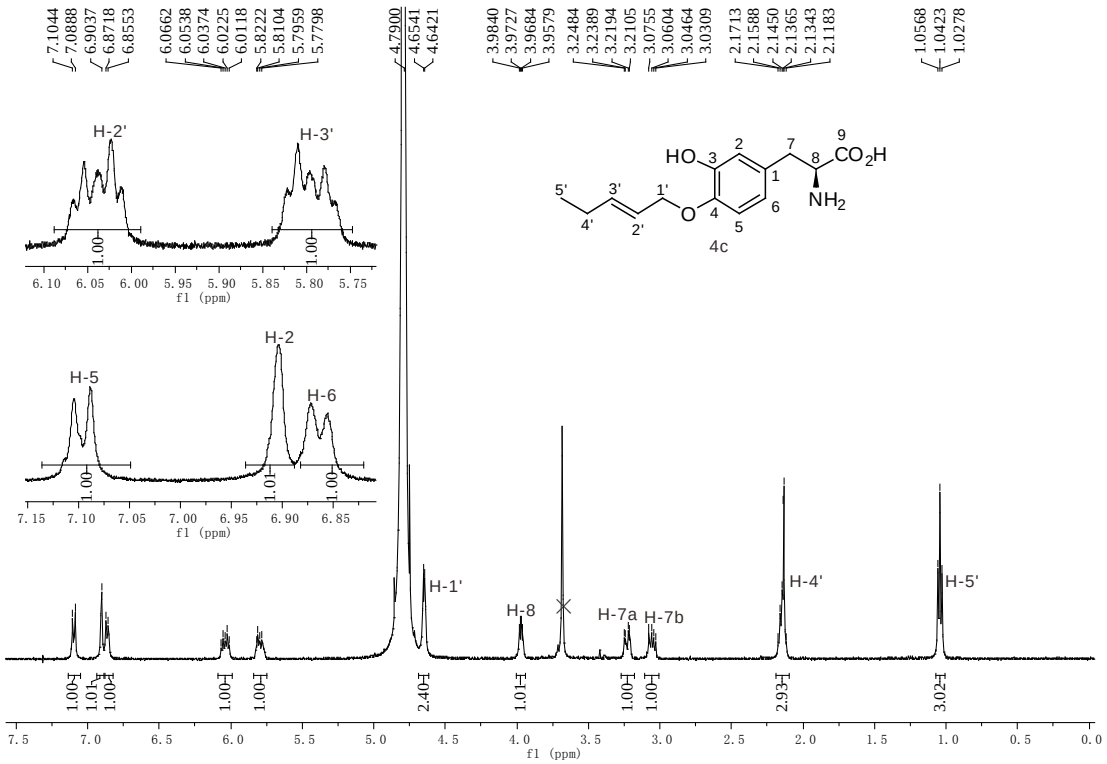


Fig. S8 ^{1}H NMR spectrum of **4c** in D_2O.

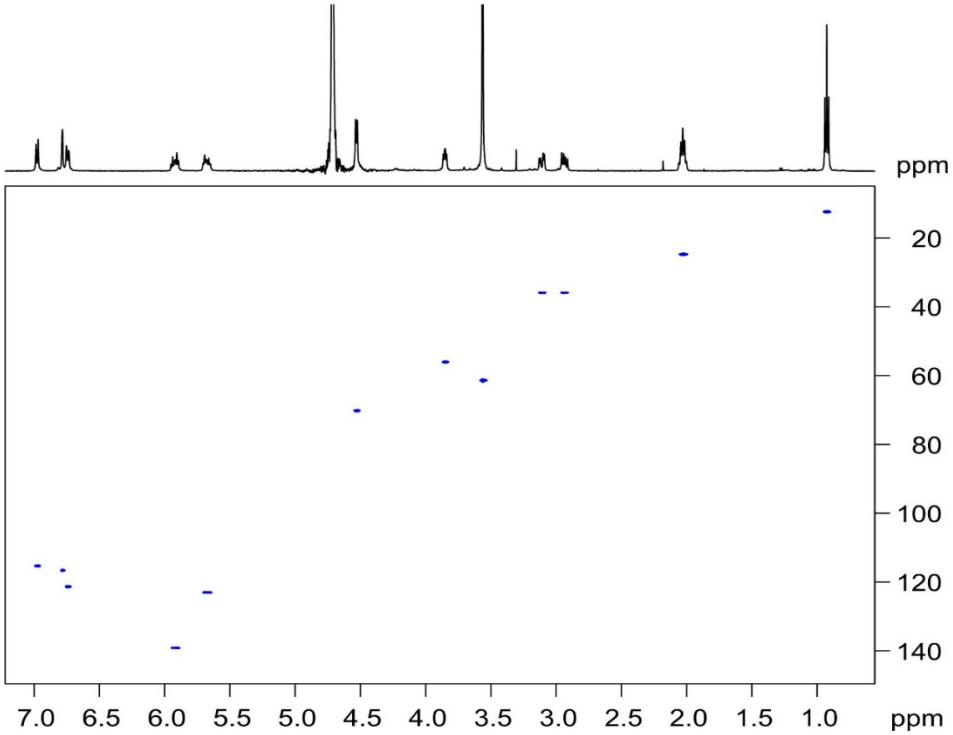


Fig. S9 HSQC spectrum of **4c** in D_2O.

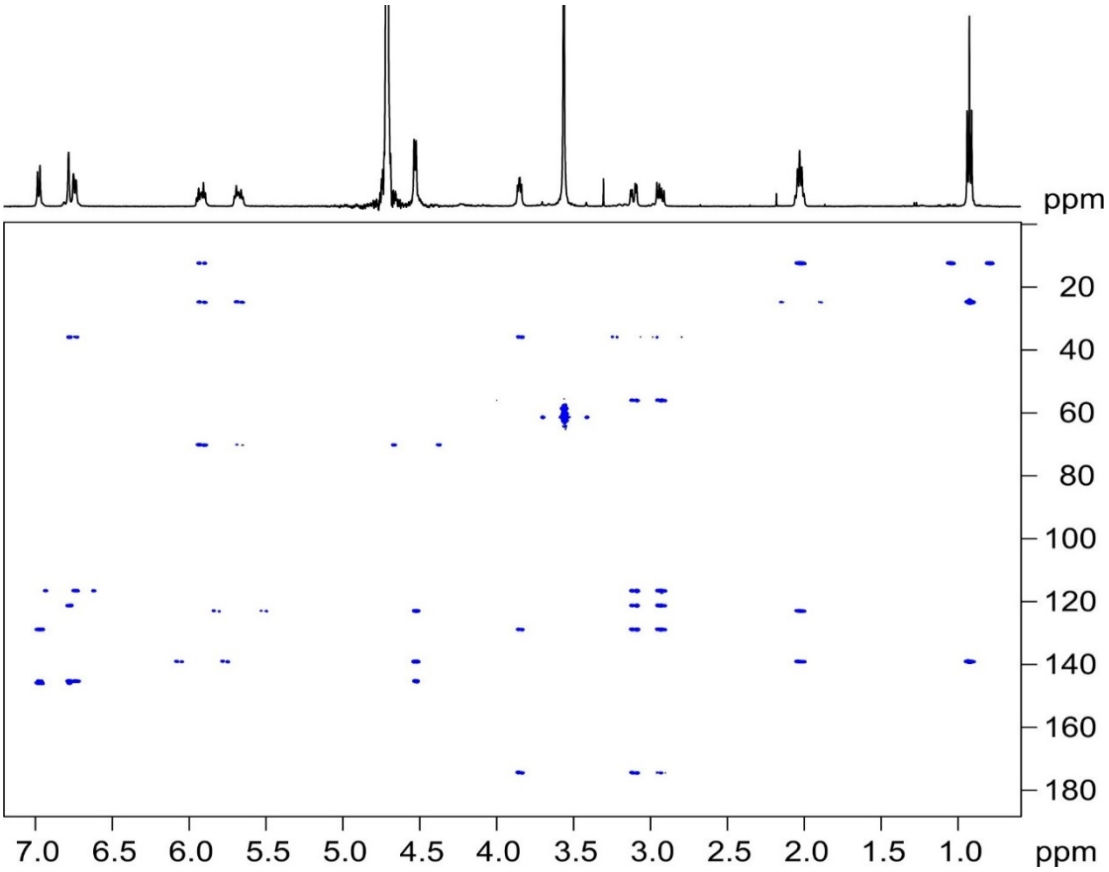


Fig. S10 HMBC spectrum of **4c** in D_2O.

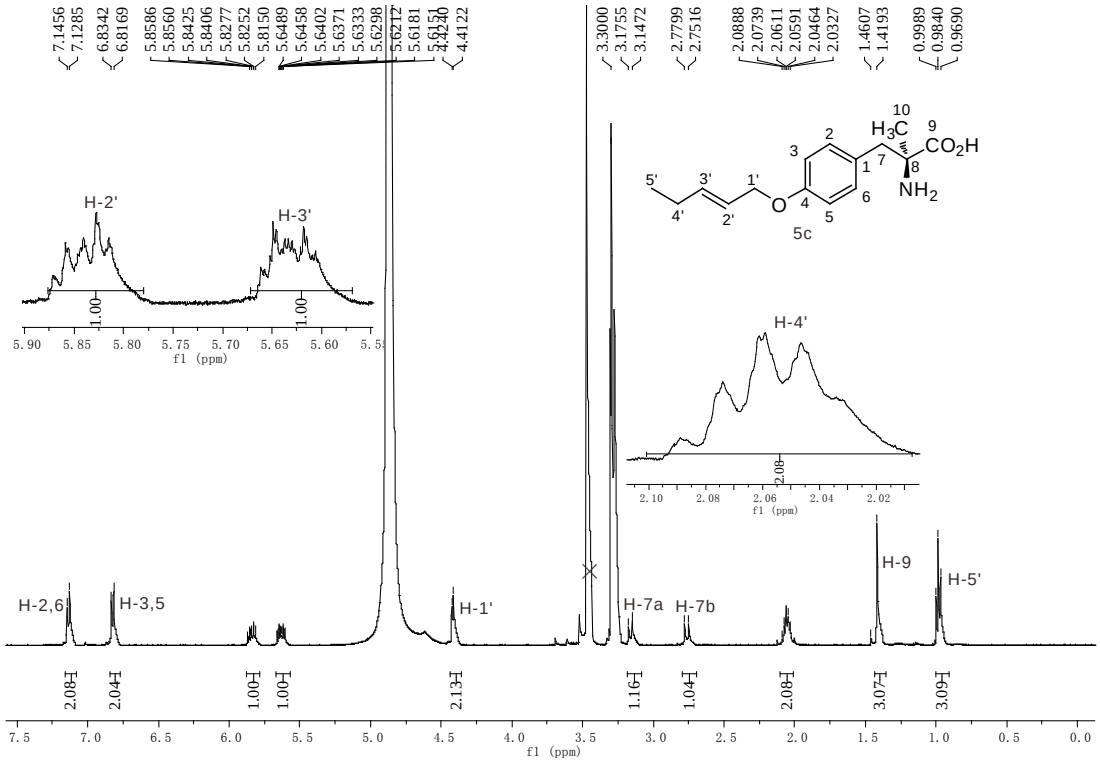


Fig. S11 ^{1}H NMR spectrum of **5c** in CD_3OD.

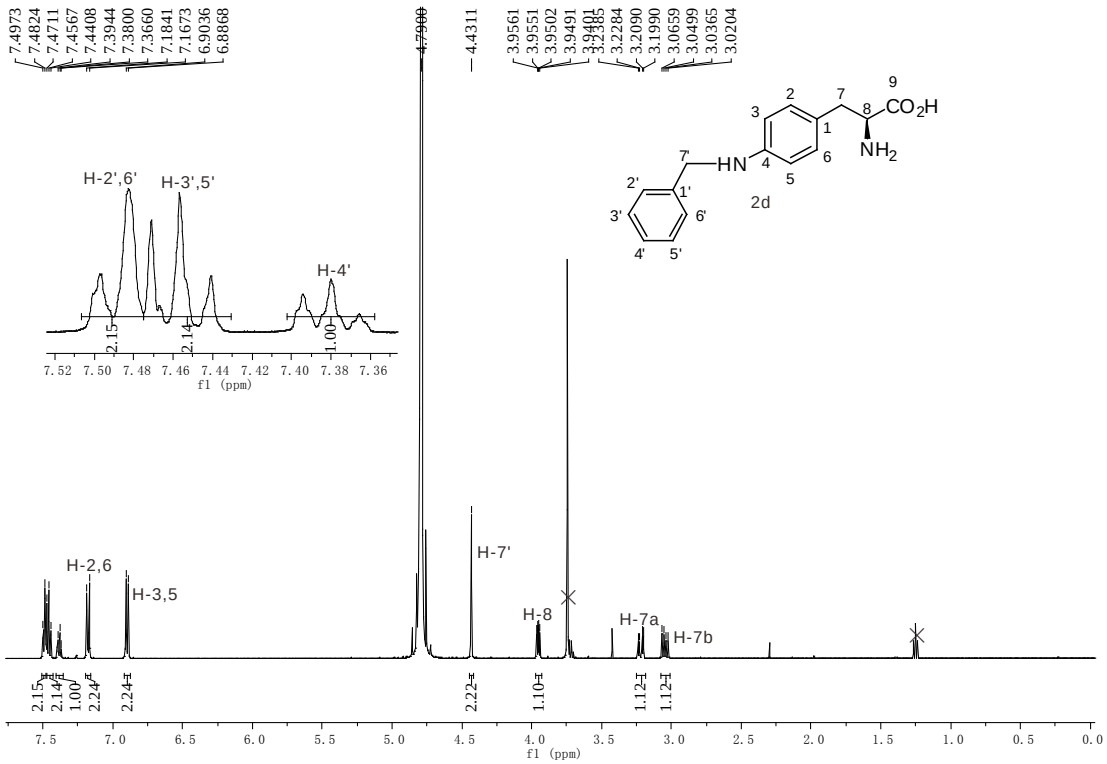


Fig. S12 ^{1}H NMR spectrum of **2d** in D_2O.

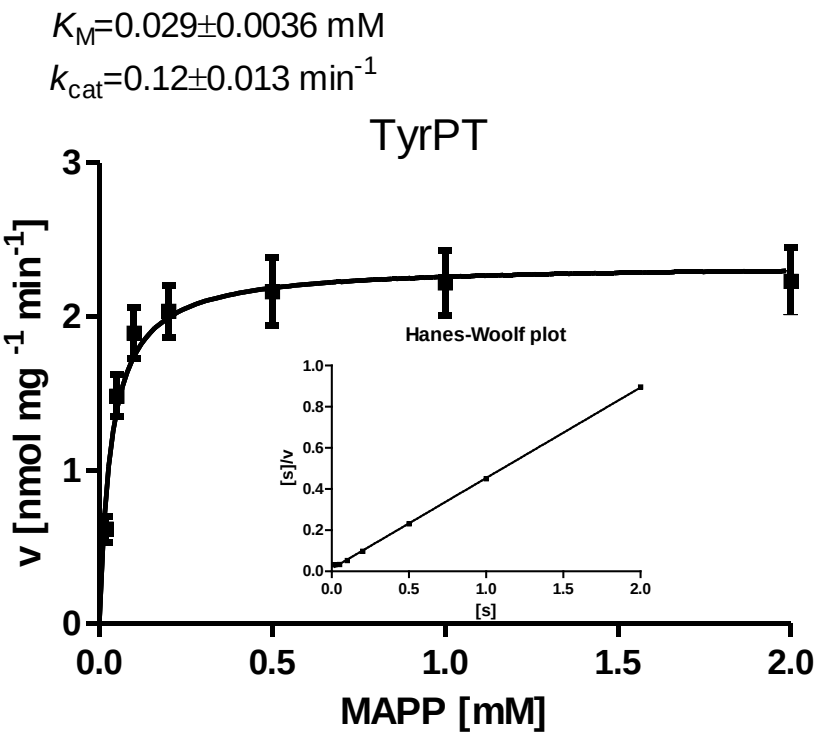


Fig. S13 Determination of kinetic parameters of TyrPT for MAPP in the presence of L-tyrosine (**1a)**.

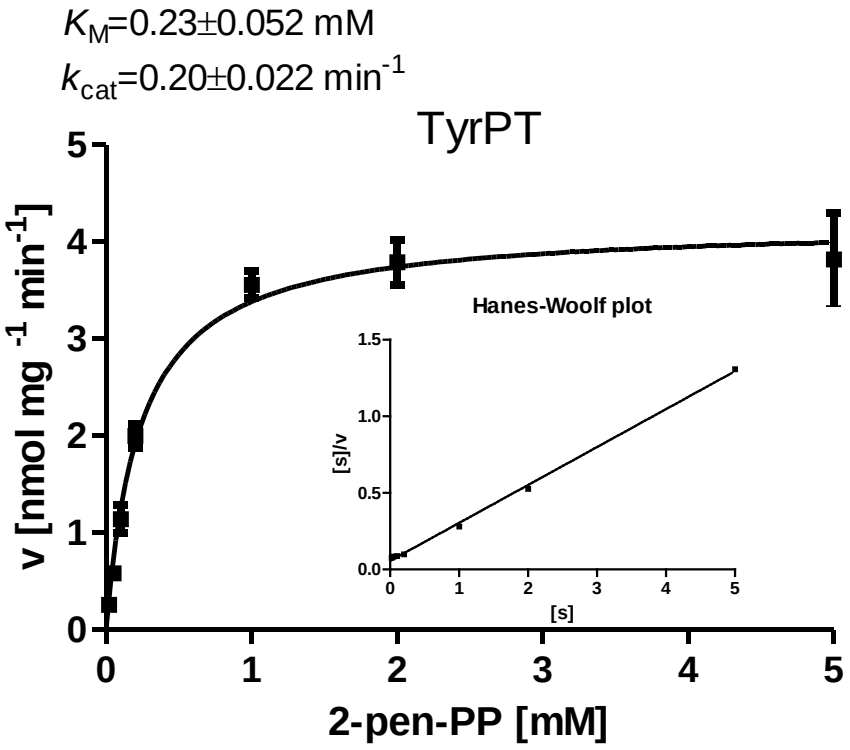


Fig. S14 Determination of kinetic parameters of TyrPT for 2-pen-PP in the presence of L-tyrosine (**1a)**.

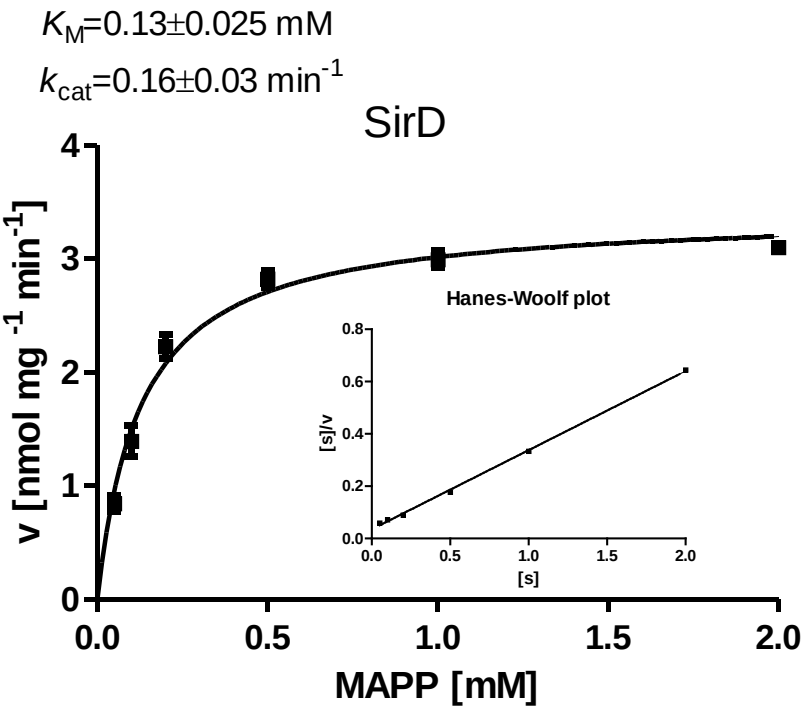


Fig. S15 Determination of kinetic parameters of SirD for MAPP in the presence of L-tyrosine (**1a)**.

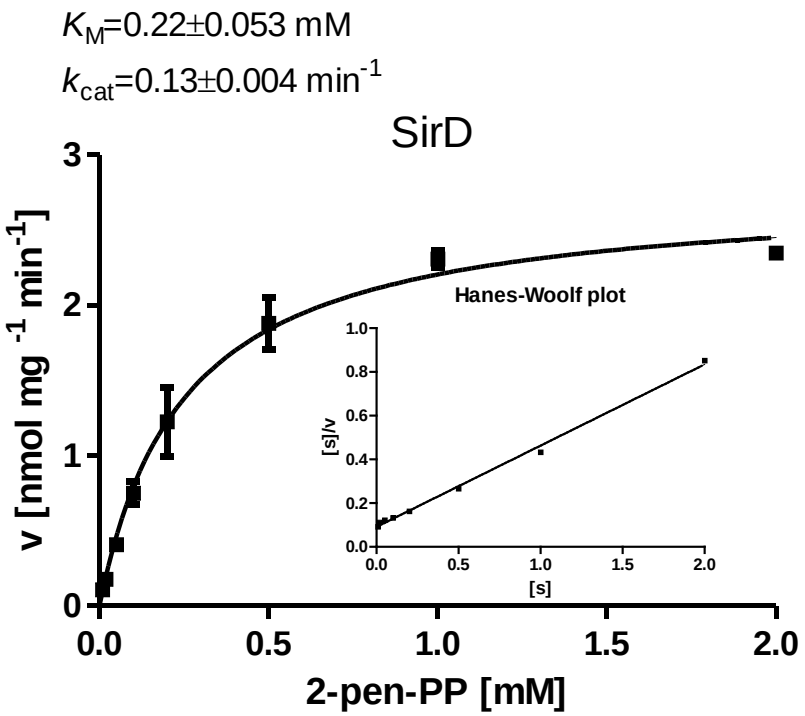


Fig. S16 Determination of kinetic parameters of SirD for 2-pen-PP in the presence of L-tyrosine (**1a)**.

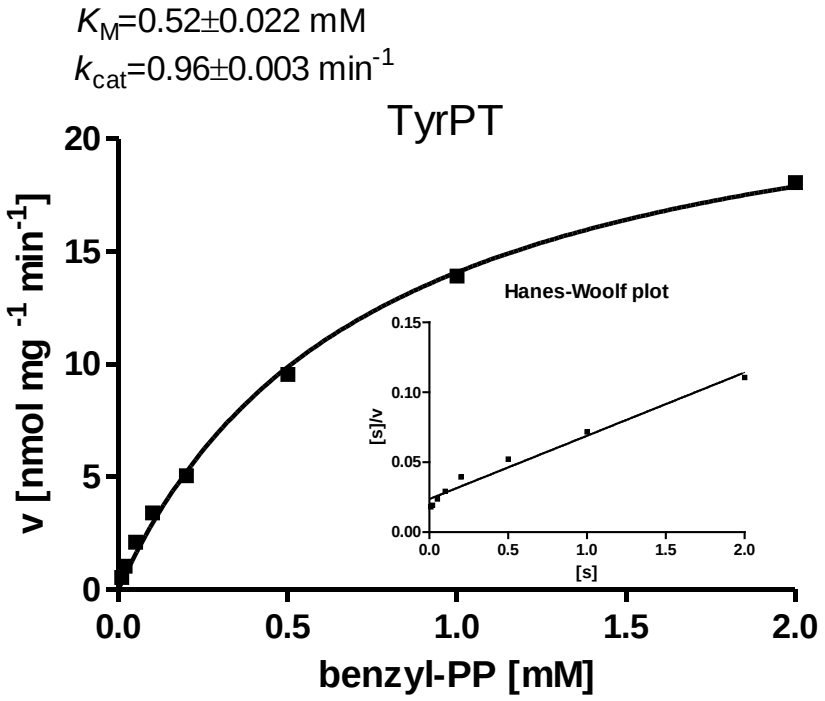


Fig. S17 Determination of kinetic parameters of TyrPT for benzyl-PP in the presence of 4-amino-L-phenylalaine (**2a**).

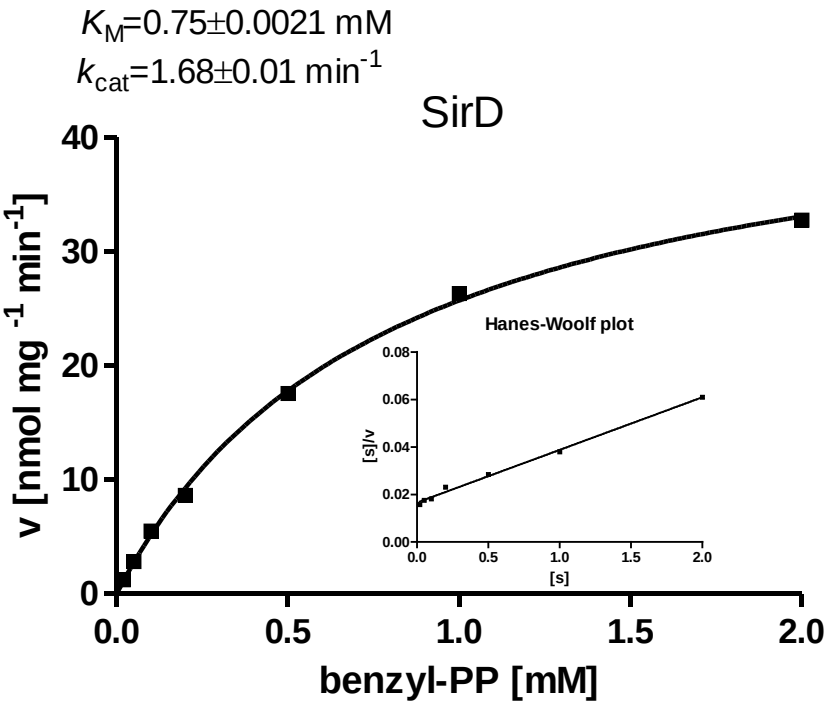


Fig. S18 Determination of kinetic parameters of SirD for benzyl-PP in the presence of 4-amino-L-phenylalaine (**2a**).

5 Conclusions and future prospects

In this thesis, the structure diversity of natural/chemoenzymatic products was significantly increased by investigation on the CDP formation and further modification in *Streptomyces* as well as by chemoenzymatic synthesis of prenylated tyrosine analogs.

Initially, thirteen CDPS homologues from eleven *Streptomyces* strains were characterized by heterologous expression in *E. coli*. Two CDPSs WP_014141671 and WP_014140974 from *Streptomyces cattleya* DSM 46488 were confirmed to form cWW, integrity with the known results. Nine new members of CDPSs were identified to catalyze the formation of cWX. With high yields of 46 to 211 mg/L *E. coli* culture, it is convenient for future application in synthetic biology. The other two CDPSs WP_031028810 from *Streptomyces* sp. NRRL F-5639 and BAU83478 from *Streptomyces laurentii* DSM 41684 were proven to catalyze the formation of cFL. Phylogenetic analysis of the CDPSs in this study with the known CDPSs revealed two clades: the relative conserved clade I containing only cWX synthases with the exception for cLL synthase and the promiscuous clade II including cWX synthases and cFX and cYX synthases.

Furthermore, one cWW tailoring P450, $P450_{NB5737}$ from *S. purpureus* NRRL-B5737, was characterized by heterologous expression in *S. coelicolor* and by *in vitro* enzyme assays with recombinant proteins. It was demonstrated that the cWW formed by $CDPS_{NB5737}$ was further modified by $P450_{NB5737,}$ resulting in an adduct of cWW and guanine *via* a C–N bond. However, the activity of $P450_{NB5737}$ is very low in the *in vitro* assays, possibly due to the incorrect state of the purified P450 from *E. coli*. The modification of peptides by a nucleobase is not reported prior to this study. $P450_{NB5737}$ is one additional example of rare atypical P450s that catalyze the coupling of two substrates with distinct structures. This indicates the presence of novel active site(s) in the identified enzymes differing from those of known P450s, which should be investigated in the future.

The spectra of prenyl donors of *O*-prenyltransferases TyprPT and SirD were also expanded by testing their acceptance toward the unnatural DMAPP analogs, such as MAPP, 2-pen-PP and benzyl-PP. The reactions were performed by incubation of

tyrosine and its 12 analogs in the presence of TyrPT and SirD. It was proven that both TyrPT and SirD exhibited similar behavior to the unnatural DMAPP analogs as their natural prenyl donor DMAPP, catalyzing the *O*-/*N*-alkylation/benzylation of tyrosine and derivatives. This study increased significantly the diversity of prenylated products by chemoenzymatic synthesis.

For future prospects, the following works can be performed:

- Coexpression of the genes of cWX-forming CDPSs with those of CDP PTs such as CdpC2PT, CdpC3PT, CdpNPT or AnaPT or even other tailoring enzymes from different synthetic pathways in engineered heterologous expression system.

- Understanding the catalytic mechanism of $P450_{NB5737}$ by solving the crystal structure of $P450_{NB5737}$.

- Identification of the extended clusters containing the $CDPS_{NB5737}$ and $P450_{NB5737}$ by heterologous expression in *Streptomyces* or gene deletion experiments.

6 References

(1) Ventura, M.; Canchaya, C.; Tauch, A.; Chandra, G.; Fitzgerald, G. F.; Chater, K. F.;van, S. D., Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum, *Microbiol.Mol.Biol.Rev.* **2007,** *71,* 495.

(2) Ghai, R.; McMahon, K. D.;Rodriguez-Valera, F., Breaking a paradigm: cosmopolitan and abundant freshwater actinobacteria are low GC, *Environ.Microbiol.Rep.* **2012,** *4,* 29.

(3) Servin, J. A.; Herbold, C. W.; Skophammer, R. G.;Lake, J. A., Evidence excluding the root of the tree of life from the actinobacteria, *Mol.Biol.Evol.* **2008,** *25,* 1.

(4) Takahashi, Y. and Nakashima, T., Actinomycetes, an inexhaustible source of naturally occurring antibiotics, *Antibiotics (Basel)* **2018,** *7.*

(5) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; et al., Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature* **2002,** *417,* 141.

(6) Mahajan, G. B. and Balachandran, L., Antibacterial agents from Actinomycetes - A Review, *Frontiers in Bioscience* **2012,** *E4,* 240.

(7) Shaw, K. and Singleton, S., "Antimicrobial resistance will kill millions--we must act now", *Nurs.Times* **2016,** *112,* 11.

(8) Baltz, R. H., Renaissance in antibacterial discovery from actinomycetes, *Curr.Opin.Pharmacol.* **2008,** *8,* 557.

(9) Gomez-Escribano, J. P. and Bibb, M. J., Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters, *Microb.Biotechnol.* **2011,** *4,* 207.

(10) Hug, J. J.; Bader, C. D.; Remskar, M.; Cirnski, K.;Muller, R., Concepts and methods to access novel antibiotics from Actinomycetes, *Antibiotics (Basel)* **2018,** *7.*

(11) Abdelmohsen, U. R.; Grkovic, T.; Balasubramanian, S.; Kamel, M. S.; Quinn, R. J.;Hentschel, U., Elicitation of secondary metabolism in actinomycetes, *Biotechnol.Adv.* **2015,** *33,* 798.

(12) Prasad, C., Bioactive cyclic dipeptides, *Peptides* **1995,** *16,* 151.

(13) Mishra, A. K.; Choi, J.; Choi, S. J.;Baek, K. H., Cyclodipeptides: An overview of their biosynthesis and biological activity, *Molecules.* **2017,** *22,* E1796.

(14) McCleland, K.; Milne, P. J.; Lucieto, F. R.; Frost, C.; Brauns, S. C.; van, d., V; Du, P. J.;Dyason, K., An investigation into the biological activity of the selected histidine-containing diketopiperazines cyclo(His-Phe) and cyclo(His-Tyr), *J.Pharm.Pharmacol.* **2004,** *56,* 1143.

(15) Kanoh, K.; Kohno, S.; Katada, J.; Takahashi, J.;Uno, I., (-)-Phenylahistin arrests cells in mitosis by inhibiting tubulin polymerization, *J.Antibiot.(Tokyo)* **1999,** *52,* 134.

(16) Ström, K.; Sjögren, J.; Broberg, A.;Schnürer, 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,* 4322.

(17) Martins, M. B. and Carvalho, I., Diketopiperazines: biological activity and synthesis, *Tetrahedron* **2007,** *63,* 9923.

(18) Kanokmedhakul, S.; Kanokmedhakul, K.; Phonkerd, N.; Soytong, K.; Kongsaeree, P.;Suksamrarn, A., Antimycobacterial anthraquinone-chromanone compound and diketopiperazine alkaloid from the fungus *Chaetomium globosum* KMITL-N0802, *Planta Med.* **2002,** *68,* 834.

(19) Hwang, I. K.; Go, V. L.; Harris, D. M.; Yip, I.; Kang, K. W.;Song, M. K., Effects of cyclo (his-pro) plus zinc on glucose metabolism in genetically diabetic obese mice, *Diabetes Obes.Metab* **2003,** *5,* 317.

(20) Kwon, O. S.; Park, S. H.; Yun, B. S.; Pyun, Y. R.;Kim, C. J., Cyclo(dehydroala-L-Leu), an alpha-glucosidase inhibitor from *Penicillium* sp. F70614, *J.Antibiot.(Tokyo)* **2000,** *53,* 954.

(21) Minelli, A.; Grottelli, S.; Mierla, A.; Pinnen, F.; Cacciatore, I.;Bellezza, I., Cyclo(His-Pro) exerts anti-inflammatory effects by modulating NF-kB and Nrf2 signalling, *Int.J.Biochem.Cell Biol.* **2012,** *44,* 525.

(22) Finking, R. and Marahiel, M. A., Biosynthesis of nonribosomal peptides, *Annu.Rev.Microbiol.* **2004,** *58,* 453.

(23) Winn, M.; Fyans, J. K.; Zhuo, Y.;Micklefield, J., Recent advances in engineering nonribosomal peptide assembly lines, *Nat.Prod.Rep.* **2016,** *33,* 317.

(24) Marahiel, M. A., A structural model for multimodular NRPS assembly lines, *Nat.Prod.Rep.* **2016,** *33,* 136.

(25) Healy, F. G.; Wach, M.; Krasnoff, S. B.; Gibson, D. M.;Loria, R., The txtAB genes of the plant pathogen *Streptomyces acidiscabies* encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity, *Mol.Microbiol.* **2000,** *38,* 794.

(26) Schultz, A. W.; Oh, D. C.; Carney, J. R.; Williamson, R. T.; Udwary, D. W.; Jensen, P. R.; Gould, S. J.; Fenical, W.;Moore, B. S., Biosynthesis and structures of cyclomarins and cyclomarazines, prenylated cyclic peptides of marine actinobacterial origin, *J Am.Chem Soc.* **2008,** *130,* 4507.

(27) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.;Marahiel, M. A., Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain, *J.Biol.Chem.* **1998,** *273,* 22773.

(28) Healy, F. G.; Krasnoff, S. B.; Wach, M.; Gibson, D. M.;Loria, R., Involvement of a cytochrome P450 monooxygenase in thaxtomin A biosynthesis by *Streptomyces acidiscabies*, *J.Bacteriol.* **2002,** *184,* 2019.

(29) Barry, S. M.; Kers, J. A.; Johnson, E. G.; Song, L.; Aston, P. R.; Patel, B.; Krasnoff, S. B.; Crane, B. R.; Gibson, D. M.; Loria, R.; et al., Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis, *Nat.Chem Biol.* **2012,** *8,* 814.

(30) García-Estrada, C.; Ullán, R. V.; Álbillos, S. M.; Fernández-Bodega, M. A.; Durek, P.; von, D. H.;Martín, J. F., A single cluster of coregulated genes encodes the biosynthesis of the mycotoxins roquefortine C and meleagrin in *Penicillium chrysogenum*, *Chem.Biol.* **2011,** *18,* 1499.

(31) Gardiner, D. M. and Howlett, B. J., Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of *Aspergillus fumigatus*, *FEMS Microbiol.Lett.* **2005,** *248,* 241.

(32) Maiya, S.; Grundmann, A.; Li, S.-M.;Turner, G., The fumitremorgin gene cluster of *Aspergillus fumigatus*: identification of a gene encoding brevianamide F synthetase, *Chembiochem.* **2006,** *7,* 1062.

(33) 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,* 100.

(34) Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.;Watanabe, K., Cytochrome P450 as dimerization catalyst in diketopiperazine alkaloid biosynthesis, *Chembiochem.* **2014,** *15,* 656.

(35) Lautru, S.; Gondry, M.; Genet, R.;Pernodet, J. L., The albonoursin gene cluster of *S. noursei*: Biosynthesis of diketopiperazine metabolites independent of nonribosomal peptide synthetases, *Chem Biol.* **2002,** *9,* 1355.

(36) Gondry, M.; Lautru, S.; Fusai, G.; Meunier, G.; Menez, A.;Genet, R., Cyclic dipeptide oxidase from *Streptomyces noursei*. Isolation, purification and partial characterization of a novel, amino acyl a,b-dehydrogenase, *Eur.J.Biochem.* **2001,** *268,* 1712.

(37) 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,* 961.

(38) Hanson, A. D.; Pribat, A.; Waller, J. C.;de Crecy-Lagard, V., 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list--and how to find it, *Biochem.J.* **2009,** *425,* 1.

(39) Jacques, I. B.; Moutiez, M.; Witwinowski, J.; Darbon, E.; Martel, C.; Seguin, J.; Favry, E.; Thai, R.; Lecoq, A.; Dubois, S.; et al., Analysis of 51 cyclodipeptide synthases reveals the basis for substrate specificity, *Nat.Chem Biol.* **2015,** *11,* 721.

(40) Liu, J.; Yu, H.;Li, S.-M., Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces* strains, *Appl.Microbiol.Biotechnol* **2018,** *102,* 4435.

(41) Liu, J.; Yu, H.;Li, S.-M., Correction to: Expanding tryptophan-containing cyclodipeptide synthase spectrum by identification of nine members from *Streptomyces s*trains, *Appl.Microbiol.Biotechnol* **2018,** *102,* 5787.

(42) Gondry, M.; Jacques, I. B.; Thai, R.; Babin, M.; Canu, N.; Seguin, J.; Belin, P.; Pernodet, J. L.;Moutiez, M., A comprehensive overview of the cyclodipeptide synthase family enriched with the characterization of 32 new enzymes, *Front Microbiol.* **2018,** *9,* 46.

(43) Seguin, J.; Moutiez, M.; Li, Y.; Belin, P.; Lecoq, A.; Fonvielle, M.; Charbonnier, J. B.; Pernodet, J. L.;Gondry, M., Nonribosomal peptide synthesis in animals: the cyclodipeptide synthase of *Nematostella*, *Chem.Biol.* **2011,** *18,* 1362.

(44) Sauguet, L.; Moutiez, M.; Li, Y.; Belin, P.; Seguin, J.; Le Du, M. H.; Thai, R.; Masson, C.; Fonvielle, M.; Pernodet, J. L.; et al., Cyclodipeptide synthases, a family of class-I aminoacyl-tRNA synthetase-like enzymes involved in non-ribosomal peptide synthesis, *Nucleic Acids Res.* **2011,** *39,* 4475.

(45) 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,* 3912.

(46) Vetting, M. W.; Hegde, S. S.;Blanchard, J. S., The structure and mechanism of the *Mycobacterium tuberculosis* cyclodityrosine synthetase, *Nat.Chem.Biol.* **2010,** *6,* 797.

(47) 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.

(48) Canu, N.; Belin, P.; Thai, R.; Correia, I.; Lequin, O.; Seguin, J.; Moutiez, M.;Gondry, M., Incorporation of non-canonical amino acids into 2,5-diketopiperazines by cyclodipeptide synthases, *Angew.Chem.Int.Ed Engl.* **2018,** *57,* 3118.

(49) Giessen, T. W.; von Tesmar, A. M.;Marahiel, M. A., A tRNA-dependent two-enzyme pathway for the generation of singly and doubly methylated ditryptophan 2,5-diketopiperazines, *Biochemistry* **2013,** *52,* 4274.

(50) Cryle, M. J.; Bell, S. G.;Schlichting, I., Structural and biochemical characterization of the cytochrome P450 CypX (CYP134A1) from *Bacillus subtilis*: a cyclo-L-leucyl-L-leucyl dipeptide oxidase, *Biochemistry* **2010,** *49,* 7282.

(51) Gondry, M.; Sauguet, L.; Belin, P.; Thai, R.; Amouroux, R.; Tellier, C.; Tuphile, K.; Jacquet, M.; Braud, S.; Courcon, M.; et al., Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes, *Nat.Chem.Biol.* **2009,** *5,* 414.

(52) Belin, P.; Le Du, M. H.; Fielding, A.; Lequin, O.; Jacquet, M.; Charbonnier, J. B.; Lecoq, A.; Thai, R.; Courcon, M.; Masson, C.; et al., Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in *Mycobacterium tuberculosis*, *Proc.Natl.Acad.Sci.U.S.A* **2009,** *106,* 7426.

(53) Giessen, T. W.; von Tesmar, A. M.;Marahiel, M. A., Insights into the generation of structural diversity in a tRNA-dependent pathway for highly modified bioactive cyclic dipeptides, *Chem.Biol.* **2013,** *20,* 828.

(54) Patteson, J. B.; Cai, W.; Johnson, R. A.; Santa Maria, K. C.;Li, B., Identification of the biosynthetic pathway for the antibiotic bicyclomycin, *Biochemistry* **2018,** *57,* 61.

(55) Meng, S.; Han, W.; Zhao, J.; Jian, X. H.; Pan, H. X.;Tang, G. L., A six-oxidase cascade for tandem C-H bond activation revealed by reconstitution of bicyclomycin biosynthesis, *Angew.Chem Int.Ed Engl.* **2018,** *57,* 719.

(56) Kanzaki, H.; Yanagisawa, S.;Nitoda, T., Enzymatic synthesis of dehydro cyclo(His-Phe)s, analogs of the potent cell cycle inhibitor, dehydrophenylahistin, and their inhibitory activities toward cell division, *Biosci.Biotechnol.Biochem.* **2004,** *68,* 2341.

(57) Giessen, T. W. and Marahiel, M. A., Rational and combinatorial tailoring of bioactive cyclic dipeptides, *Front Microbiol.* **2015,** *6,* 785.

(58) Li, S.-M., Prenylated indole derivatives from fungi: structure diversity, biological activities, biosynthesis and chemoenzymatic synthesis, *Nat.Prod.Rep.* **2010,** *27,* 57.

(59) Yu, X.; Liu, Y.; Xie, X.; Zheng, X.-D.;Li, S.-M., Biochemical characterization of indole prenyltransferases: Filling the last gap of prenylation positions by a 5-dimethylallyltryptophan synthase from *Aspergillus clavatus*, *J.Biol.Chem.* **2012,** *287,* 1371.

(60) Cook, D. J.; Finnigan, J. D.; Cook, K.; Black, G. W.;Charnock, S. J., Cytochromes P450: History, classes, catalytic mechanism, and industrial application, *Adv.Protein Chem.Struct.Biol.* **2016,** *105,* 105.

(61) Omura, T. and Sato, R., The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification, and properties, *J.Biol.Chem.* **1964,** *239,* 2379.

(62) Hannemann, F.; Bichet, A.; Ewen, K. M.;Bernhardt, R., Cytochrome P450 systems--biological variations of electron transport chains, *Biochim.Biophys.Acta* **2007,** *1770,* 330.

(63) Podust, L. M. and Sherman, D. H., Diversity of P450 enzymes in the biosynthesis of natural products, *Nat.Prod.Rep.* **2012,** *29,* 1251.

(64) 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,* 1141.

(65) Poulos, T. L.; Perez, M.;Wagner, G. C., Preliminary crystallographic data on cytochrome P-450CAM, *J.Biol.Chem.* **1982,** *257,* 10427.

(66) Poulos, T. L.; Finzel, B. C.; Gunsalus, I. C.; Wagner, G. C.;Kraut, J., The 2.6-A crystal structure of *Pseudomonas putida* cytochrome P-450, *J.Biol.Chem.* **1985,** *260,* 16122.

(67) Guengerich, F. P.; Waterman, M. R.;Egli, M., Recent structural insights into cytochrome P450 function, *Trends Pharmacol.Sci.* **2016,** *37,* 625.

(68) Zhang, X. and Li, S., Expansion of chemical space for natural products by uncommon P450 reactions, *Nat.Prod.Rep.* **2017,** *34,* 1061.

(69) Anzai, Y.; Li, S.; Chaulagain, M. R.; Kinoshita, K.; Kato, F.; Montgomery, J.;Sherman, D. H., Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics, *Chem.Biol.* **2008,** *15,* 950.

(70) Li, S.; Tietz, D. R.; Rutaganira, F. U.; Kells, P. M.; Anzai, Y.; Kato, F.; Pochapsky, T. C.; Sherman, D. H.;Podust, L. M., Substrate recognition by the multifunctional cytochrome P450 MycG in mycinamicin hydroxylation and epoxidation reactions, *J.Biol.Chem.* **2012,** *287,* 37880.

(71) Carlson, J. C.; Li, S.; Gunatilleke, S. S.; Anzai, Y.; Burr, D. A.; Podust, L. M.;Sherman, D. H., Tirandamycin biosynthesis is mediated by co-dependent oxidative enzymes, *Nat.Chem.* **2011,** *3,* 628.

(72) Mo, X.; Wang, Z.; Wang, B.; Ma, J.; Huang, H.; Tian, X.; Zhang, S.; Zhang, C.;Ju, J., Cloning and characterization of the biosynthetic gene cluster of the bacterial RNA polymerase inhibitor tirandamycin from marine-derived *Streptomyces* sp. SCSIO1666, *Biochem.Biophys.Res.Commun.* **2011,** *406,* 341.

(73) Makino, M.; Sugimoto, H.; Shiro, Y.; Asamizu, S.; Onaka, H.;Nagano, S., Crystal structures and catalytic mechanism of cytochrome P450 StaP that produces the indolocarbazole skeleton, *Proc.Natl.Acad.Sci.U.S.A* **2007,** *104,* 11591.

(74) Onaka, H.; Asamizu, S.; Igarashi, Y.; Yoshida, R.;Furumai, T., Cytochrome P450 homolog is responsible for C-N bond formation between aglycone and deoxysugar in

the staurosporine biosynthesis of *Streptomyces* sp. TP-A0274, *Biosci.Biotechnol Biochem.* **2005,** *69,* 1753.

(75) Ikezawa, N.; Iwasa, K.;Sato, F., Molecular cloning and characterization of CYP80G2, a cytochrome P450 that catalyzes an intramolecular *C-C* phenol coupling of (S)-reticuline in magnoflorine biosynthesis, from cultured *Coptis japonica* cells, *J.Biol.Chem.* **2008,** *283,* 8810.

(76) Zhao, B.; Guengerich, F. P.; Bellamine, A.; Lamb, D. C.; Izumikawa, M.; Lei, L.; Podust, L. M.; Sundaramoorthy, M.; Kalaitzis, J. A.; Reddy, L. M.; et al., Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of *Streptomyces coelicolor* A3(2) cytochrome P450 158A2, *J.Biol.Chem.* **2005,** *280,* 11599.

(77) Zhao, B.; Lamb, D. C.; Lei, L.; Kelly, S. L.; Yuan, H.; Hachey, D. L.;Waterman, M. R., Different binding modes of two flaviolin substrate molecules in cytochrome P450 158A1 (CYP158A1) compared to CYP158A2, *Biochemistry* **2007,** *46,* 8725.

(78) Funa, N.; Funabashi, M.; Ohnishi, Y.;Horinouchi, S., Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*, *J.Bacteriol.* **2005,** *187,* 8149.

(79) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.;Moore, B. S., Biosynthesis of polybrominated aromatic organic compounds by marine bacteria, *Nat.Chem.Biol.* **2014,** *10,* 640.

(80) Mazzaferro, L. S.; Huttel, W.; Fries, A.;Muller, M., Cytochrome P450-catalyzed regio- and stereoselective phenol coupling of fungal natural products, *J.Am.Chem.Soc.* **2015,** *137,* 12289.

(81) Lin, H.-C.; McMahon, T. C.; Patel, A.; Corsello, M.; Simon, A.; Xu, W.; Zhao, M.; Houk, K. N.; Garg, N. K.;Tang, Y., P450-Mediated coupling of indole fragments to forge communesin and unnatural isomers, *J.Am.Chem Soc.* **2016,** *138,* 4002.

(82) Sanchez, J. F.; Somoza, A. D.; Keller, N. P.;Wang, C. C., Advances in *Aspergillus* secondary metabolite research in the post-genomic era, *Nat.Prod.Rep.* **2012,** *29,* 351.

(83) Sunassee, S. N. and Davies-Coleman, M. T., Cytotoxic and antioxidant marine prenylated quinones and hydroquinones, *Nat.Prod.Rep.* **2012,** *29,* 513.

(84) Jo, Y. H.; Shin, B.; Liu, Q.; Lee, K. Y.; Oh, D. C.; Hwang, B. Y.;Lee, M. K., Antiproliferative prenylated xanthones and benzophenones from the roots of *Cudrania tricuspidata* in HSC-T6 cells, *J.Nat.Prod.* **2014,** *77,* 2361.

(85) Wu, S.-B.; Long, C.;Kennelly, E. J., Structural diversity and bioactivities of natural benzophenones, *Nat.Prod.Rep.* **2014,** *31,* 1158.

(86) Heide, L., Prenyl transfer to aromatic substrates: genetics and enzymology, *Curr.Opin.Chem.Biol.* **2009,** *13,* 171.

(87) Yazaki, K.; Sasaki, K.;Tsurumaru, Y., Prenylation of aromatic compounds, a key diversification of plant secondary metabolites, *Phytochemistry* **2009,** *70,* 1739.

(88) Yu, X. and Li, S.-M., Prenyltransferases of the dimethylallyltryptophan synthase superfamily, *Methods Enzymol.* **2012,** *516,* 259.

(89) Wollinsky, B.; Ludwig, L.; Hamacher, A.; Yu, X.; Kassack, M. U.;Li, S. M., Prenylation at the indole ring leads to a significant increase of cytotoxicity of tryptophan-containing cyclic dipeptides, *Bioorg.Med.Chem.Lett.* **2012,** *22,* 3866.

(90) Jones, L. H.; Narayanan, A.;Hett, E. C., Understanding and applying tyrosine biochemical diversity, *Mol.Biosyst.* **2014,** *10,* 952.

(91) Stierle, A. A.; Stierle, D. B.; Goldstein, E.; Parker, K.; Bugni, T.; Baarson, C.; Gress, J.;Blake, D., A novel 5-HT receptor ligand and related cytotoxic compounds from an acid mine waste extremophile, *J.Nat.Prod.* **2003,** *66,* 1097.

(92) Almeida, C.; Part, N.; Bouhired, S.; Kehraus, S.;Konig, G. M., Stachylines A-D from the sponge-derived fungus *Stachylidium* sp, *J.Nat.Prod.* **2011,** *74,* 21.

(93) Ahmed, F.; Ohtsuki, T.; Aida, W.;Ishibashi, M., Tyrosine derivatives isolated from *Streptomyces* sp. IFM 10937 in a screening program for TRAIL-resistance-overcoming activity, *J.Nat.Prod.* **2008,** *71,* 1963.

(94) Gardiner, D. M.; Cozijnsen, A. J.; Wilson, L. M.; Pedras, M. S.;Howlett, B. J., The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*, *Mol.Microbiol.* **2004,** *53,* 1307.

(95) Kremer, A. and Li, S.-M., A tyrosine *O*-prenyltransferase catalyses the first pathway-specific step in the biosynthesis of sirodesmin PL, *Microbiology* **2010,** *156,* 278.

(96) 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,* 10119.

(97) Botta, B.; Vitali, A.; Menendez, P.; Misiti, D.;Delle, M. G., Prenylated flavonoids: pharmacology and biotechnology, *Curr.Med.Chem.* **2005,** *12,* 717.

(98) Steffan, N.; Unsöld, I. A.;Li, S.-M., Chemoenzymatic synthesis of prenylated indole derivatives by using a 4-dimethylallyltryptophan synthase from *Aspergillus fumigatus*, *Chembiochem* **2007,** *8,* 1298.

(99) Kremer, A. and Li, S.-M., Potential of a 7-dimethylallyltryptophan synthase as a tool for production of prenylated indole derivatives, *Appl.Microbiol.Biotechnol.* **2008,** *79,* 951.

(100) Yu, X.; Xie, X.;Li, S.-M., Substrate promiscuity of secondary metabolite enzymes: prenylation of hydroxynaphthalenes by fungal indole prenyltransferases, *Appl.Microbiol.Biotechnol.* **2011,** *92,* 737.

(101) Yu, X. and Li, S.-M., Prenylation of flavonoids by using a dimethylallyltryptophan synthase 7-DMATS from *Aspergillus fumigatus*, *Chembiochem* **2011,** *12,* 2280.

(102) Tarcz, S.; Xie, X.;Li, S.-M., Substrate and catalytic promiscuity of secondary metabolite enzymes: *O*-prenylation of hydroxyxanthones with different prenyl donors by a bisindolyl benzoquinone *C*- and *N*-prenyltransferase, *RSC Adv.* **2014,** *4,* 17986.

(103) Fan, A.; Zocher, G.; Stec, E.; Stehle, T.;Li, S.-M., Site-directed mutagenesis switching a dimethylallyl tryptophan synthase to a specific tyrosine *C3*-prenylating enzyme, *J.Biol.Chem.* **2015,** *290,* 1364.

(104) Zou, H.-X.; Xie, X.; Zheng, X.-D.;Li, S.-M., The tyrosine *O*-prenyltransferase SirD catalyzes *O*-, *N*-, and *C*-prenylations, *Appl.Microbiol.Biotechnol.* **2011,** *89,* 1443.

(105) Rudolf, J. D. and Poulter, C. D., Tyrosine *O*-prenyltransferase SirD catalyzes *S*-, *C*-, and *N*-prenylations on tyrosine and tryptophan derivatives, *ACS Chem.Biol.* **2013,** *8,* 2707.

(106) Chooi, Y. H.; Cacho, R.;Tang, Y., Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*, *Chem.Biol.* **2010,** *17,* 483.

(107) Pockrandt, D.; Sack, C.; Kosiol, T.;Li, S.-M., A promiscuous prenyltransferase from *Aspergillus oryzae* catalyses *C*-prenylations of hydroxynaphthalenes in the presence of different prenyl donors, *Appl.Microbiol.Biotechnol.* **2014,** *98,* 4987.

(108) Pockrandt, D. and Li, S.-M., Geranylation of cyclic dipeptides by the dimethylallyl transferase AnaPT resulting in a shift of prenylation position on the indole ring, *Chembiochem.* **2013,** *14,* 2023.

(109) Liebhold, M.; Xie, X.;Li, S.-M., Expansion of enzymatic Friedel-Crafts alkylation on indoles: Acceptance of unnatural beta-unsaturated allyl diphospates by dimethylallyl-tryptophan synthases, *Org.Lett.* **2012,** *14,* 4884.

(110) Liebhold, M.; Xie, X.;Li, S.-M., Breaking cyclic dipeptide prenyltransferase regioselectivity by unnatural alkyl donors, *Org.Lett.* **2013,** *15,* 3062.

(111) Liebhold, M. and Li, S.-M., Regiospecific benzylation of tryptophan and derivatives catalyzed by a fungal dimethylallyl transferase, *Org.Lett.* **2013,** *15,* 5834.

(112) James, E. D.; Knuckley, B.; Alqahtani, N.; Porwal, S.; Ban, J.; Karty, J. A.; Viswanathan, R.;Lane, A. L., Two distinct cyclodipeptide synthases from a marine actinomycete catalyze biosynthesis of the same diketopiperazine natural product, *ACS Synth.Biol.* **2016,** *5,* 547.

(113) 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*, 4921-4925**.**

Statutory Declaration

Ich, Huili Yu, versichere, dass ich meine Dissertation

„Investigation on Cyclodipeptide Formation and Metabolism in Actinomycetes and Application of Fungal Prenyltransferases in the Chemoenzymatic Synthesis “

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.....

.....

Huili Yu

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Dr. Shu-Ming Li for accepting me as his PhD student and his excellent supervision during my PhD study. He is always willing to help to solve the problems in both research and life. And my personal development has greatly benefited from his rigorous approach, ambitious motivation, hard working as well as high efficiency in scientific research.

I am grateful to Prof. Dr. Michael Keusgen for acting as my second referee and examiner.

I would like to thank Dr. Xiulan Xie, Dr. Regina Ortmann and Stefan Newel for taking NMR spectra, Rixa Kraut and Nina Zitzer for taking mass spectra as well as Prof. Dr. Wenhan Lin from Peking University for his help in structure elucidation of my compounds. I also express the deep appreciation to Dr. Dieter Kreusch and Sabine Burgers for ordering chemicals and lab supplies as well as maintenance of laboratory equipment and to Nadja Braun, Buvac Drázen for preparing media and distilled water.

I would like to thank Jing Liu for her idea communication and help in the cooperation of the first project in this dissertation. I thank Lena Ludwig for her synthesis of DMAPP and its unnatural analogs in the third project. Sincere thanks to Theodor Peez and Maximilian Jung for their assistance in measurement of the UV-Vis spectra of P450 enzyme. My special gratitude goes to Lindsay Coby and Florian Kindinger for translating and proof reading the summary part of this dissertation, and Ge Liao for proof reading this dissertation.

Many thanks go to Dr. Aili Fan, Viola Wohlgemuth, Peter Mai, Kang Zhou, Linus Naumann, Jie Fan, Jinglin Wang and Sandra Ditzler for sharing the old and new offices and nice talks in the office. Also special thanks go to my “Mensa partner” Jie Fan, Huomiao Ran for their company to Mensa and excellent cooking in the kitchen.

I want to thank all my current and former colleagues in the institute including Kristin Öqvist, Elisabeth Hühner, Liujuan Zheng, Pan Xiang, Bastian Kemmerich, Jonas Nies, Dr. Katija Backhaus, Johanna Schäfer, Lena Mikulski, Alexander Frehse, Kirsten Brockmeyer, Thang Son Ta, Victoria Werner, Dr. Nina Gerhards, Dr. Julia Winkelblech,

Dr. Mike Liebhold, Dr. Kathrin Mundt, Dr. Carsten Wunsch, Dr. Daniel Pockrandt, Dr. Sylwia Tarcz, Dr. Beate Wollinsky, Dr. Jennifer Robinson, Dr. Lennart Poppe, Dr. Agus Chahyadi, Dr. Soheil Pezeshki for the great and wonderful time. I also thank the short-term communication professors and students in our institute for the great experience together.

I further express my thanks to my friends outside of the laboratory for their wonderful friendship.

I thank the China Scholarship Council (CSC) and the MArburg University Research Academy (MARA) for financial supports.

At last but not the least, I deeply thank my family for their encouragement, unconditional love and support over the past years. Special and heartful thanks to my husband Xiaowei Han for his company during this time. We always support with each other and I will remember this valuable experience forever.

Curriculum Vitae

Pages 135-136 (Curriculum vitae) contain personal data. It is therefore not part of the online-publication of this thesis.