Development of a modular platform for the tailor-made biosynthesis of lignans: From monolignols towards (-)-podophyllotoxin

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Davide Decembrino

from Aosta, Italy

Düsseldorf, September 2021

From the Institute of Biochemistry, Chair II
at the Heinrich Heine University Düsseldorf

Printed by permission of the Faculty of Mathematics and Natural Sciences of
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Vlada B. Urlacher
Co-supervisor: Prof. Dr. Ilka Maria Axmann

Date of the oral examination: 27/09/2021

Affidavit

I declare under oath that I have produced my thesis independently and without any undue assistance by third parties under consideration of the 'Principles for the Safeguarding of Good Scientific Practice at Heinrich Heine University Düsseldorf'.
This dissertation was not submitted to any other faculty, neither in this nor in a similar form. I have not undertaken any unsuccessful or successful attempt at a doctorate yet.

Düsseldorf, 27/10/2021

(Davide Decembrino)

“Never say never. ‘Cause – you know – limits, just like fears, are often just an illusion”

Michael Jeffrey Jordan

“Il pessimismo della ragione, l’ottimismo della volontà”

Romain Rolland, Antonio Gramsci

Table of contents

Publications and conference contributions

Publications

- **Decembrino D.**, Ricklefs E., Wohlgemuth S., Girhard M., Schullehner K., Jach G., Urlacher V.B. **2020**, Assembly of plant enzymes in *E. coli* for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide, *ACS Synthetic biology*, 9 (11), 3091-3103, https://doi.org/10.1021/acssynbio.0c00354
- **Decembrino D.**, Girhard M. and Urlacher V.B. **2021**, Use of copper as a trigger for *in vivo* activity of *E. coli* laccase CueO: a simple tool for biosynthetic purposes, *ChemBioChem*, 22 (8), 1470-1479, https://doi.org/10.1002/cbic.202000775
- Hilberath T., Windeln L.M., **Decembrino D.**, Le-Huu P., Bilsing F.L., Urlacher V.B. **2020**, Two-step screening for identification of drug-metabolizing bacterial cytochromes P450 with diversified selectivity, *ChemCatChem*, 12 (6), 1710-1719, https://doi.org/10.1002/cctc.201901967
- **Decembrino D.**, Raffaele A., Hilberath T., Girhard M., Urlacher V. B. **2021**, Synthesis of (-)-deoxypodophyllotoxin, (-)-podophyllotoxin and (-)-epipodophyllotoxin via multi-enzyme cascade reactions in *E. coli* - Manuscript submitted to *Microbial Cell Factories*

Conference contributions

- Poster at the 9th BioCat international conference in Hamburg (Germany), 26-30 August 2018. Title: “Application of promiscuous cytochrome P450s from actinomycetes for the production of drug metabolites”
- Speaker at the 13th International CeBiTec Symposium in Multi-Step Syntheses in Biology & Chemistry in Bielefeld (Germany), 2-4 December 2019. Title: “Reconstitution of high-lignans biosynthetic pathway for the production of podophyllotoxin valuable precursor (-)-pluviatolide in *E. coli* whole-cell biocatalyst”

Abstract

Plants secondary metabolic pathways are source of a broad plethora of bioactive compounds that have been widely employed in both modern and traditional medicine. In this regard, lignans are prominent examples, as they have long been attributed to several health promoting effects on the human body. Traditional means of production involve the extraction of lignans from *Forsythia*, *Linum*, *Sesamum* and *Podophyllum* plant species, however the increasing commercial demand determined overexploitation and environmental endangerment of the sources. Total-synthesis routes have been developed as alternatives; however, their feasibility is hampered by the high complexity and stereochemistry of the compounds. In search for more sustainable approaches, intensive genome mining led to the disclosure of many enzymes composing the lignans secondary metabolic network *in planta*. Complementary, advances in metabolic engineering and synthetic biology allowed the assembly of such pathways in recombinant bacteria and yeast.

In this thesis, *E. coli* was used as host to harbour the biosynthetic route starting from coniferyl alcohol and proceeding towards the microtubules depolymerizer (-)-podophyllotoxin. In total, ten enzymes were assembled to develop combinable, independent modular units targeting the biosynthesis of diverse lignans such as (±)-pinoresinol, (-)-matairesinol, (-)-pluviatolide, and (-)-deoxypodophyllotoxin at need. Given the fundamental role within plant secondary metabolism, special attention was given to the recombinant expression of plant cytochrome P450 monooxygenases because reconstituting their activity in prokaryotic host is not trivial. Within this work, three P450 enzymes originating from *Sinopodophyllum hexandrum* were successfully expressed in active state in *E. coli* for the first time.

In the first developmental step within this work, the biotransformation of the monolignol coniferyl alcohol to the lignan (-)-matairesinol was achieved. *E. coli*'s endogenous multi-copper oxidase CueO was employed after copper-mediated activation to furnish (±)-pinoresinol via coniferyl alcohol dimerization. The kinetic resolution of (±)-pinoresinol represents the first branching point in lignan biosynthesis and it was here performed by the

pinoresinol-lariciresinol reductase from *Forsythia intermedia* (FiPLR). The secoisolariciresinol dehydrogenase from *Podophyllum peltatum* (PpSDH) was implemented as a further step to produce (-)-matairesinol.

In the second step, the biotransformation of (+)-pinoresinol to the non-commercially available lignan (-)-pluviatolide was achieved at a preparative scale. To produce (-)-pluviatolide, (-)-matairesinol is oxidized by the action of a methylenedioxy bridge-forming P450 from *Sinopodophyllum hexandrum* (CYP719A23) with high regio- and enantioselectivity. To assemble a 4-step multi-enzyme cascade, FiPLR and PpSDH were coexpressed with CYP719A23, whose activity was supported and optimized by employing the NADPH-dependent reductase ATR2 from *Arabidopsis thaliana*. After the optimization of this setup, the efficient conversion of (+)-pinoresinol was achieved leading to the isolation of 137 mg/L product with 76% isolated yield and high purity (>99% UV/vis, >94% MS, ee ≥99%).

In the third step, the previous multi-enzyme cascade was prolonged by assembling an effective 5-steps biotransformation of (-)-matairesinol to (-)-deoxypodophyllotoxin (98% yield), further extendable to (-)-epipodophyllotoxin. In this setup, among the seven enzymes involved, activity of three plant P450s, CYP719A23, CYP71CU1, and CYP82D61 could be reconstituted in *E. coli*. To date, a physiological putative (-)-podophyllotoxin synthase from *S. hexandrum* performing the last step – the hydroxylation of (-)-deoxypodophyllotoxin to (-)-podophyllotoxin - remains unidentified. Alternatively, two potential candidates from the P450 library of our group (CYP107Z and CYP105D from *Streptomyces platensis*) were evaluated for their ability to hydroxylate (-)-deoxypodophyllotoxin. Both P450s were found active, and (-)-podophyllotoxin and (-)-epipodophyllotoxin were detected as reaction products.

Abbreviations

2-ODD	2-oxoglutarate/Fe(II) dependent dioxygenase
2,6-DMP	2,6-dimethoxyphenol
5-Ala	5-aminolevulinic acid
A. thaliana	*Arabidopsis thaliana*
ADH	Alcohol dehydrogenase
Ala	Alanine
Amp	Ampicillin
APCI	Atmospheric pressure chemical ionization
API	Active pharmaceutical ingredient
AU	Arbitrary unit
B. megaterium	*Bacillus megaterium*
B. subtilis	*Bacillus subtilis*
C. glutamicum	*Corynebacterium glutamicum*
CE	Crude extract (soluble protein fraction)
CFU	Colony forming units
CO	Carbon monoxide
CPR	Cytochrome P450 reductase
cww	Cell wet weight
CYP	Cytochrome P450
Da	Dalton
ddH_2O	Double deionized water
DIR	Dirigent protein
DMSO	Dimethyl sulphoxide
DNA	Deoxy ribonucleic acid
dNTPs	Deoxynucleotides triphosphate
D-PTOX	(-)-deoxypodophyllotoxin
δ	Chemical shift
E. coli	*Escherichia coli*
EC	Enzyme class

ee	Enantiomeric excess
Epi-PTOX	(-)-epipodophyllotoxin
ER	Endoplasmic reticulum
ESI	Electron spray ionization
EtOH	Ethanol
EV	Empty vector (negative control)
F. intermedia	*Forsythia intermedia*
FAD	Flavin adenine dinucleotide
FdR	Flavodoxin reductase from *E. coli*
FMN	Flavin mononucleotide
GC	Gas chromatography
GDH	Glucose dehydrogenase
HPLC	High Performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IS	Internal standard
Kan	Kanamycin
L. usitatissimum	*Linum usitatissimum*
LB	Lysogeny broth (Luria Bertani medium)
LC/MS	Liquid chromatography / mass spectrometry
m/z	Mass-to-charge ratio
MCS	Multiple cloning site
MDB	Methylenedioxy-bridge
MeOH	Methanol
MIC	Minimum inhibitory substrate concentration
MS	Mass spectrometry
N. benthamiana	*Nicotiana benthamiana*
NAD(P)H	Nicotinamide dinucleotide (phosphate)
nat	Native codon
NMR	Nuclear magnetic resonance
N-terminus	Amino terminus
oc	Codon optimised

OD_{600}	Optical density measured at 600 nm
OMT	O-methyltransferase
P	Pellet (Protein insoluble fraction)
P. peltatum	*Podophyllum peltatum*
P450	Cytochrome P450
PDA	Photo diode array
PLR	Pinoresinol-lariciresinol reductase
PMSF	phenylmethylsulfonylfluoride
ppm	Parts-per-million
PTMs	Post translational modifications
PTOX	(-)-podophyllotoxin
rbs	Ribosome binding site
ROS	Oxygen reactive species
rpm	Rounds per minute
RT	Retention time
S. cerevisiae	*Saccharomyces cerevisiae*
S. hexandrum	*Sinopodophyllum hexandrum*
S. platensis	*Streptomyces platensis*
SDH	Secoisolariciresinol reductase
SIM	Single (selected) ion monitoring
Strp	Streptomycin
T. plicata	*Thuja plicata*
TB	Terrific broth
TCA	Tricarboxylic acid cycle
U	Unit
U/mL	Volumetric activity
Uv/Vis	Ultraviolet / visible light
WHO	World Health Organization
wt	Wild type
YkuN	Flavodoxin from *B. subtilis*
ε	Molar absorption extinction coefficient

1 Introduction

1.1 Green chemistry, synthetic biology and biocatalysis

Over the last decades, sustainable development became subject matter for the interplay between economy, technology and environment.[1] Starting in the mid of the 20th century, rising environmental awareness fuelled the debates not only among scholars but also in the public sphere about the environmental sustainability of industrial processes. In fact, the pursuit of progress caused unintended harm to the planet and consequently to biodiversity and civilization.[2] In the light of these aspects, the necessity of a paradigm change in the chemical industry led to the editing of the "12 Principles of Green Chemistry", which outline how to design and develop novel chemical processing to avoid or minimize waste, usage, and production of toxic or generally hazardous compounds, as well as preventing overexploitation of natural sources (Figure 1.1).[3]

Classical chemical routes towards fine chemicals or pharmaceuticals are often multi-step processes involving the application of inorganic catalysts in combination with harsh conditions of temperature, pressure and pH. Differently, the exploitation of enzymes as biocatalysts – either in isolated form or as whole-cell catalysts – generally allows to achieve energy efficient reactions, with minor waste production and lower reagents toxicity than inorganic catalysts.[4] Enzyme-catalysed reactions offer further advantages. Optimal activities are often achieved under mild, biologically compatible reaction conditions, which allow lower energy consumption and reduced recycling and disposal costs, as most of them are applied in aqueous solutions.[5] Not only, billions of years of evolution optimized the activities of an incredibly vast number of enzymes, leading to the production of numerous diverse molecules.[6] Specifically, this fact endorses enzymes as powerful candidates for the fine tailoring of highly complex chemicals. Thus - under the big umbrella of Green Chemistry - nature becomes an inspiration source, making synthetic biology and biocatalysis irreplaceable approaches to target sustainable chemical processes.[4, 6]

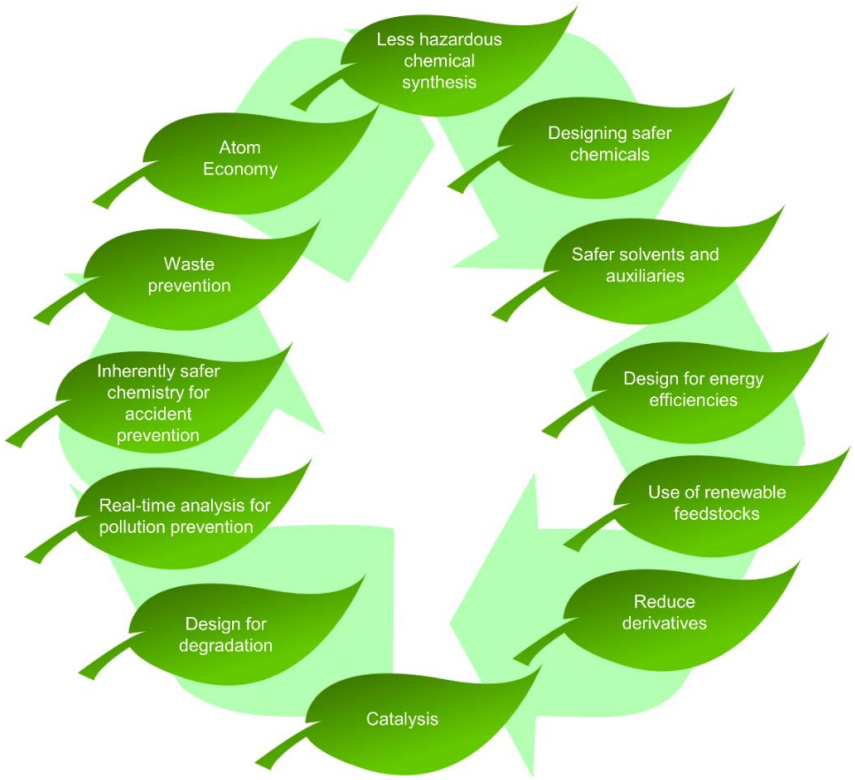


Figure 1.1. Overview of the twelve principles of Green Chemistry.[3]

As DNA technology cost dropped over the years, genome sequencing together with custom gene synthesis and protein engineering allowed more and more enzymes to be available for biosynthetic purposes and led to the improvement of enzymatic features such as selectivity, stability and substrate spectrum.[7, 8] Consequently, biocatalytic processes could be specifically developed to be applied in food and beverages industry, as well as pharmaceuticals production and waste or plastic degradation.[5, 9, 10] Concerning the synthesis of natural products, the plant kingdom represents an extremely vast source of bioactive specialized chemicals deriving from secondary metabolic pathways. More than 200,000 of plant natural compounds have been identified and eventually exploited in modern and traditional medicine as well as in industry as fragrances, nutrients and building blocks.[11-13] In the attempt of sustainable production, studies in the fields of metabolic engineering and synthetic biology promoted the manipulation of microorganisms for enhanced metabolic properties and/or to insert unnatural, novel molecular parts for desired purposes.[14, 15] Specifically, these

approaches provided the chance to engineer multi-step enzymatic cascades to rebuild complex plant pathways in bacteria and yeasts for the *de-novo* synthesis of highly valuable compounds from cheap starting material such as sugars or amino acids.[12, 16]

1.1.1 Development of whole-cell biocatalysts

Given the premises described above, the growing availability of biocatalysts allowed the development of more and more complex combinations among them, generating *de novo* synthetic cascades for *in vitro*, *in vivo* or hybrid approaches.[17] Among these three ways, the building of multi-enzyme cascades within living cells offers several advantages compared to the approaches with isolated enzymes. First, whole-cell biocatalysts can perform biotransformations with little manipulation during or after cell cultivation. Intuitively, their usage overcomes tedious prepping steps such as cell lysis, enzymatic purification and immobilization, which are typical of *in vitro* approaches. This simplification inherently makes whole-cell systems the cheapest possible catalyst formulation,[18] which is a relevant argument for the industrial chemical manufacturing decision-makers. In fact, the more steps are needed to prepare a biocatalyst, the more equipment/resources are required, resulting in higher costs. Along the same line, the scalability of whole-cell processes to implement fermentation processes is additional benefit.[19] Another advantage over *in vitro* setups relies on the possibility to harness the host cells metabolism to directly generate cofactors or precursors *in loco*.[18] For instance, to sustain reactions involving oxidoreductases, expensive co-substrates such as NAD(P)H can be supplied via cell metabolism starting from cheap energy sources like glucose and glycerol.[18, 20] Moreover, benefits of the compartmentalization of one or multiple enzymes within a single cell may provide improved biocatalysts stability and efficiency.[21]

Despite all advantages, some drawbacks need to be considered when developing whole-cell biocatalysts as well. When the components of a cascade are placed into a living system, the efficiency of the catalysts may be affected by the host organism itself. Among others, mass transfer limitations across the cell membrane represent the most prominent disadvantages. As an example, considering bacteria, the presence of cell wall may interfere with the

substrates uptake and thus the efficiency of the involved enzymes. For this reason, permeabilization techniques or coexpression of membrane transporters are often applied to overcome this issue.[22-24] In some cases, once the desired compound has entered the cell, native enzymes may compete for the substrate and transform it into undesired side-products. In this case, gene knock-out/down may be implemented to prevent interferences with the host's enzymes.[25] Additionally, the building of non-native complex pathways by the inclusion of multiple enzymes of diverse origin may compromise not only cell viability and growth rate, but provoke gene expression imbalances as well, resulting in significant metabolic burden. Overall, all these factors substantially concur in influencing the efficiency of the biocatalyst. Such bottlenecks can be relieved in a case-by-case scenario by a combinatorial approach involving the core elements of the gene-expression machinery such as promoters, ribosome bindings sites (RBS), the genes themselves etc.[21]

From these aspects, it seems clear that within the first development attempts, whole-cell biocatalysts often perform in suboptimal conditions which need to be optimized stepwise in order to fully disclose their potential as "green" chemical manufacturing devices.

1.2 Cytochrome P450 monooxygenases

1.2.1 General features

Among all enzymes involved in secondary metabolism, cytochrome P450 monooxygenases, commonly abbreviated as P450s or CYPs, play a key role. They represent an array of enzymes capable of leading to an unrivalled diversity of specialized compounds by the performance of incredibly diverse reactions.[26, 27] For this reason, such enzymes became significantly attractive to target chemically challenging transformations for biosynthetic purposes.[28-31]

Cytochromes P450 are heme-thiolate enzymes which belong to the enzyme class of oxidoreductases (E.C: 1.14.-.-) and are widely distributed all over natural kingdoms.[32, 33] Their nomenclature is owed to a unique spectral property: A distinct absorption maximum at 450 nm rises when carbon monoxide (CO) is bound to the heme-group in the catalytic centre, where the iron atom is reduced from Fe^{III} to Fe^{II}.[34] Specifically, the iron is coordinated by the four nitrogens of the porphyrin ring (Figure 1.2) with a fifth axial bond made by a conserved cysteine which allow the formation of a sixth coordination bond with oxygen or CO.[33] P450s act as protein terminals in electron transfer chains where molecular oxygen (O_2) is the final acceptor of electrons. Electrons are delivered to the heme by nicotinamide adenine dinucleotide or nicotinamide adenine dinucleotide phosphate (NAD(P)H) cofactors via redox partner proteins. Specifically, P450s catalyse the introduction of an oxygen atom into a non-activated C-H bond, whereas the second oxygen atom is reduced to water as by-product, as summarized in the equation below.[35]

$$\mathbf{R\text{-}H + O_2 + NADP(H) + H^+ \rightarrow R\text{-}OH + NADP^+ + H_2O}$$

After the first hydroxylation, P450s may carry out a wide range of diverse common and uncommon reactions such as carbon-carbon cleavage or coupling, O-, N- and S-dealkylation, but also N- and S-oxidation, aromatic coupling and other.[36, 37]

1.2.2 Catalytic cycle

For most P450s a common catalytic cycle has been postulated (Figure 1.2). At first, the catalytic centre is at resting state, meaning the ferric heme-iron (Fe^{III}) is bound to water, which serves as sixth ligand (**1**). Upon substrate binding, water is displaced determining the heme-iron to undergo to a spin-state shift from low-spin to high-spin. This event results at increasing the redox potential, allowing Fe^{III} to accept the first electron coming from NAD(P)H via redox partner proteins and being reduced to the ferrous state (Fe^{II}) (**2 - 3**). Afterwards, molecular oxygen (O_2) is bound the Fe^{II} forming a dioxygen adduct (**4**). At this step, a second electron can be delivered to the complex to form the ferric-peroxo intermediate (**5**) which becomes the ferric-hydroperoxo intermediate after subsequent protonation (**6**). The hydrogen addition leads to the heterolytic cleavage of the dioxygen bond with the consequent release of water and the generation of the highly reactive intermediate named **Compound I**, a ferryl species, in which the radical cation is delocalized over the porphyrin ring and the thiolate ligand (**6-7**). **Compound I** abstracts a hydrogen atom from the substrate via the so-called rebound mechanism, generating a ferryl-hydroxo species named **Compound II** and a substrate radical (**8**). The radical rebound to the hydroxyl group of Compound II allows the generation of a mono-oxygenated product, whereas water can coordinate to the catalytic centre restoring the resting state (**9**).[37]

Alternatively to the common cycle, some P450s can go through the peroxide shunt pathway, exploiting H_2O_2 as sole source of electrons and protons to generate the ferric hydroperoxo intermediate (**6**).[38] However, other shunt reactions may occur as well leading to P450 self-inactivation due to the uncoupling of substrate oxidation and electron consumption reactions and the generation of reactive oxygen species which may modify the apoenzyme or bleach the heme.[39]

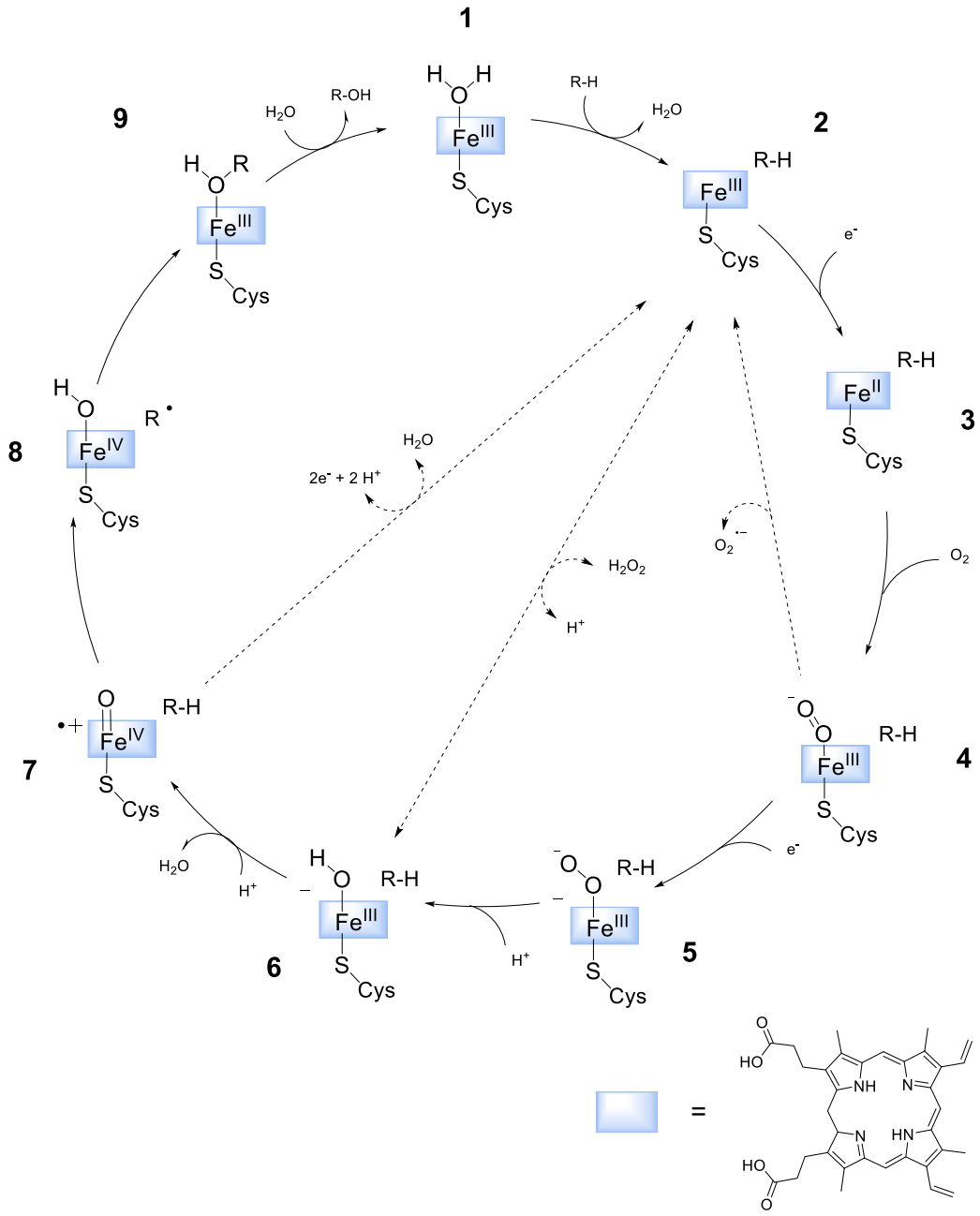


Figure 1.2. The P450 catalytic cycle. The central iron has six possible coordination bonds. Four are formed with the protoporphyrin ring (blue box) surrounding the iron. A fifth axial bond is made by the thiol group of a cysteine. By this conformation, a sixth axial bond is possible with oxygen or other molecules of higher affinity such as carbon or nitrogen monoxides. 1: resting state, Fe^{III}-water complex; 2: penta-coordinated ferric-Fe complex; 3: penta-coordinated ferrous-Fe complex; 4: dioxygen adduct; 5: ferric-peroxo intermediate; 6: ferric-hydroperoxo intermediate; 7: Compound I; 8: Compound II; 9: free Fe^{III}-complex. Dotted lines represent alternative shunt pathways. Adapted from Li *et al*., 2020.[40]

1.2.3 Redox partner systems

By definition, to achieve the activation of molecular oxygen and perform substrate hydroxylation, P450s act as terminal electron acceptors receiving electrons from NAD(P)H via redox partner proteins.[40] According to the identified in nature electron transfer systems, up to ten P450 classes have been defined. Among them, Class I and Class II represent the two major and more common systems in nature (Figure 1.3). Class I redox systems are typical for prokaryotic and mitochondrial P450s and are composed of three proteins involving two redox partners and a P450. Required electrons are provided from NAD(P)H and shuttled firstly to a ferredoxin reductase (**FdR**) containing flavin adenine dinucleotide (**FAD**) as cofactor. From FdR, electrons are transferred to a small iron-sulphur protein ferredoxin from which they are transferred eventually to the heme group of the P450. Whereas in bacterial systems all proteins are soluble, mitochondrial systems are characterized by the association of both reductase and P450 to the inner mitochondrial membrane.[33]

Class II redox systems exploits a two-protein system in which both proteins are typically bound to the endoplasmic reticulum (ER) of all eukaryotic organisms. Here, a single cytochrome P450 reductase (CPR) holding both FAD and flavin mononucleotide (FMN) domains is recruited to shuttle electrons from NADPH to the P450.[41] In addition to Class I and II, it is worthy to mention a single-protein system (Class VIII) found in prokaryotes and lowers eukaryotes whose major feature is the natural fusion occurring between the eukaryotic-like FAD/FMN reductase domain and the P450 domain in a single polypeptide (Figure 1.3).[33, 42] The prokaryotic members of this class, for instance CYP102A1 from *Bacillus megaterium*, known as P450 BM3, are cytosolic soluble proteins whereas the eukaryotic members such as CYP505 from *Fusarium oxysporum* are often ER-bound.[33, 43]

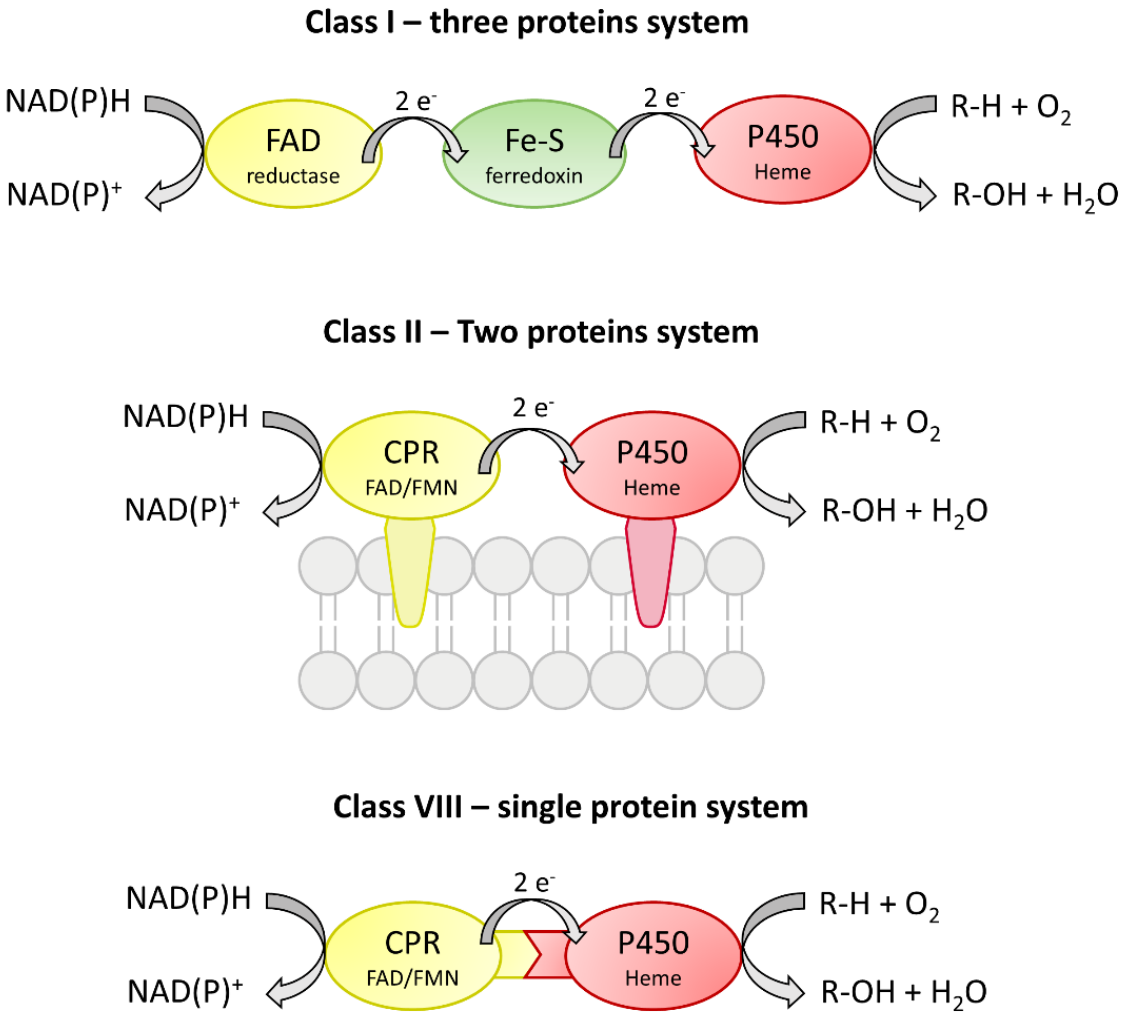


Figure 1.3. Overview on various redox partners systems. Class I: example of a soluble bacterial system. Class II: eukaryotic microsomal system. Both P450 and CPR are anchored to the endoplasmic reticulum. Class VIII: bacterial CPR-P450 fusion system. Adapted from Hannemann *et al.*, 2007.[33]

Concerning the application of plant P450s, whose natural redox partners systems usually belongs to Class II, often the lack of knowledge about the physiologic cognate CPR limited the reconstitution of their activity for biotechnological purposes. For this reason, redox systems from Class I or the generation of fusion constructs mimicking Class VIII systems have been often applied to reconstitute P450 activity.[44-47]

1.2.4 P450s as oxidative tailors of complex molecules

The whole range of diverse reactions catalysed by P450s are involved in either degradation or in biosynthesis. Human hepatic P450s – among which well-studied examples are CYP3A4 and CYP2D6 – are mainly involved in drugs and xenobiotics detoxification, whereas steroidogenic enzymes like CYP17A or CYP21A in steroid metabolism.[48] Microbial P450s participate for C-source assimilation, fatty acids hydroxylation, as well as in the synthesis of secondary metabolites, appointing such enzymes as valuable tools for biosynthetic and/or bioremediation purposes.[49] To date, *Streptomyces* species are being kept in the spotlight as the source of many antibiotic, antifungal, antiparasitic, anticancer and immunosuppressant compounds which became commonly used pharmaceuticals.[50] Moreover, because of the wide substrate scope which may mimic human P450s activities, P450s from *Streptomyces* have been applied for the generation of drug metabolites from a vast array of chemically and structurally different compounds.[51]

Concerning the plant kingdom, P450s represent the largest gene class, with little less than 200,000 (~172,000) entries collected into databases and more than 16.000 properly classified.[52] The incredible variety of structurally and chemically diverse metabolites produced by plants reckon P450s as key actors in secondary metabolism and their role in plants defence, adaptation, and development.[53] The continuous finding of more and more secondary metabolites becomes causally related to the progressive genetic evolution of P450s, amplifying the range of possible catalytic mechanisms which lead to the production of sophisticated chemicals with diverse biological activities.[53] To sum up, all these metabolites can be gathered into three major classes: terpenoids, phenylpropanoids, and nitrogen-containing compounds. The latter group covers alkaloids, cyanogenic glucosides, and glucosinolates. All these natural products constitute not only a chemical reservoir with essential ecological functions, but also a valuable source of novel pharmaceuticals, nutraceuticals, cosmetics, flavours and pesticides.[13]

1.2.5 Role of P450s in *de novo* assembly of plant metabolic pathways in *E. coli*

As the market demand of plant-derived products is constantly growing, the interest towards the application of plant P450s rises exponentially. Thus, the attempts to engineer microbial cell factories to produce desired plant metabolites.[54] Understanding the chemical logic behind Nature's strategies to achieve high diversity becomes fundamental to build effective bio-bricks able to replicate complex biosynthetic routes.[55] Following this logic, understanding the role of P450s is crucial, as the oxygenation they perform may act as branching point driving further modifications as well as a final chemical trimming.[26, 56] Keeping an eye on industrial scalability, microorganisms such as *Saccharomyces cerevisiae* and *Escherichia coli* that can achieve high cell density, represent undoubtedly the organisms of choice for the heterologous reconstitution of plant metabolic pathways. In particular, the *de novo* synthesis of artemisinic acid, the precursor of the anti-malaria drug artemisinin, in *S. cerevisiae,* and taxa-4,11-diene, the precursor of the anticancer drug taxol, in *E. coli* are some of the most prominent examples.[28, 31] Even though the expected ramp up in artemisinin production did not happen and taxol still cannot be synthesized in *E. coli*, both examples illustrate the power of synthetic biology and the role of P450 enzymes as essential bio-bricks.

Regarding the development of *E. coli* as a microbial chassis to harbour multiple recombinant proteins*,* this organism does not possess any P450 genes in its genome, it can be easily genetically manipulated, easily cultured, and its physiology can thrive in various growth conditions, allowing the engineering of new phenotypes.[20] However, since most of eukaryotic P450s are membrane-anchored, the absence of the well-developed system of inner membranes in the bacterium may represent a limitation for their expression. Similarly, the narrow range of post-translational modification typical for prokaryotes may represent another issue.[57] Additionally, the translational efficiency of foreign genes may be critical to achieve the heterologous expression of desired target genes. In this regard, techniques of codon optimization and harmonization have been developed to minimize events limiting gene translation - such as the formation of mRNA secondary structures - or influencing the protein

folding, harming the final tertiary structure.[58, 59] Such events may also complicate the successful production of catalytically active eukaryotic proteins in *E. coli*.

Concerning the improvement of recombinant expression of cytochromes P450 in *E. coli*, several techniques have been gradually developed following the growing knowledge of P450s biology and biochemistry. Together with the just described codon usage optimization techniques, other factors such as the variation of cell growth conditions, the selection of different *E. coli* strains and promoters, and the modification of certain genetic features without altering the functional properties of the protein of interest, have been endeavoured and described to positively influence heterologous expression. Amongst these, the manipulation of the N-terminal membrane associated sequence typical for eukaryotic P450s including complete or partial truncation, or substitution of some amino acids, or addition of alanine as the second amino acid have been broadly applied and revealed to be often beneficial.[60]

P450s involved into secondary metabolic pathways often possess high substrate- and product-specificity as a result of a millenary natural optimization process.[61] This fact has to be considered as an added value, since the selective oxygenation of complex molecules can be performed without cumbersome procedures or conditions typical of synthetic chemistry. However, the extensive utilization of plant P450s in industrial processes is often limited by their low activity and stability,[62] implying further efforts in protein engineering towards the enzymes themselves and their interdependency with redox partner proteins.[54]

1.3 Lignans

Derived from the shikimate pathway, lignans are phenylpropanoid secondary metabolites widespread all over the plant kingdom, with *Forsythia*, *Linum*, *Sesamum* and *Podophyllum* identified as model species among flowering plants, whereas *Thuya*, *Juniperus*, *Callitris*, *Tsuga* and *Picea* among gymnosperms.[63] These compounds are formed from two different phenyl propane units linked by a β-β' bond with different oxidation and substitution patterns.[64] Within plants physiology, the role of lignans is similar to others plant secondary metabolites like alkaloids and thus related to plant defence against pathogens and herbivores.[64] Based on the skeleton composition, the cyclization pattern, and oxygen position within the skeleton, lignans are divided between eight groups: arylnaphthalene, aryltetralin, dibenzocyclooctadiene, dibenzylbutane, dibenzylbutyrolactol, dibenzylbutyrolactone, furan, and furofuran (Figure 1.4).[65]

Lignans, have been proven to possess various biological activities. Among others, anti-inflammatory and antioxidant effects were described to be associated to lignan-rich diets, as well as prevention of prostate and breast cancers and cardiovascular diseases.[66] Lignans such as pinoresinol or lariciresinol, but especially secoisolariciresinol and matairesinol are identified as phytoestrogens and eventually transformed by the human intestine microflora to the so-called enterolignans named enterodiol and enterolactone (Figure 1.5).[67, 68] Specifically, their presence in biological fluids has been associated with the beneficial effects described above. However, among all lignans, (-)-podophyllotoxin (PTOX) represents a peculiar compound, and it is used in the treatment of viral infections (e.g. venereal warts). Moreover, this molecule has been appointed as a potent anti-neoplastic agent, acting by preventing the microtubules assembly within cells, subsequently leading to apoptosis. Not only, direct derivatives of PTOX such as teniposide and etoposide are widely used within several chemotherapy regimens because of their activity as topoisomerase inhibitors,[69, 70] with the latter inserted in the list of essential medicines by the World Health Organisation.[71] For all these reasons, in a world in which the demand for novel pharmaceuticals and dietary intakes with lifestyle diseases

reducing features, lignans represent valuable compounds for the potential development of drugs, functional foods and dietary supplements.[72]

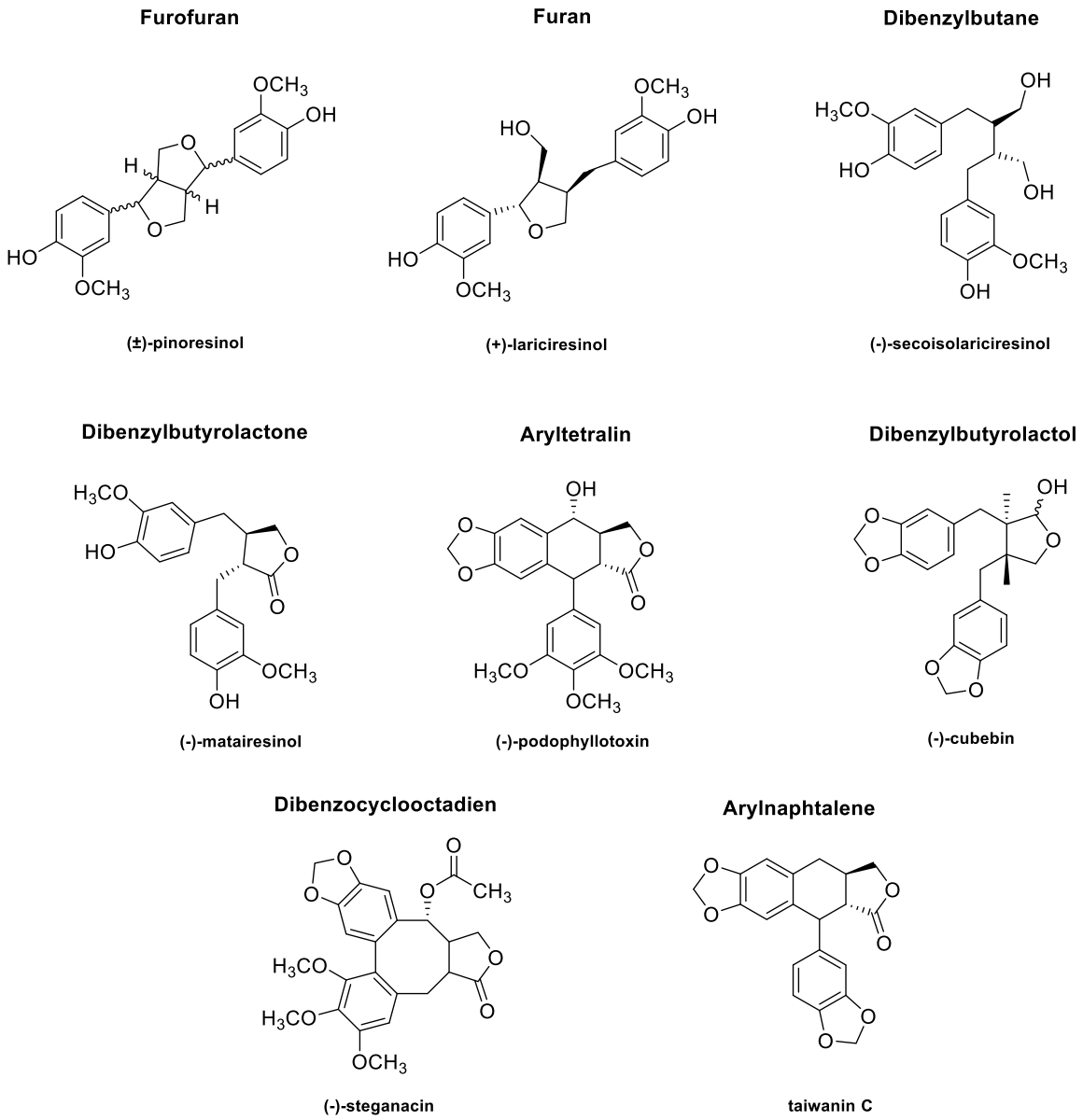


Figure 1.4. Exemplary chemical structures representing the eight lignan families. Adapted from Satake *et al.,* 2015.[65]

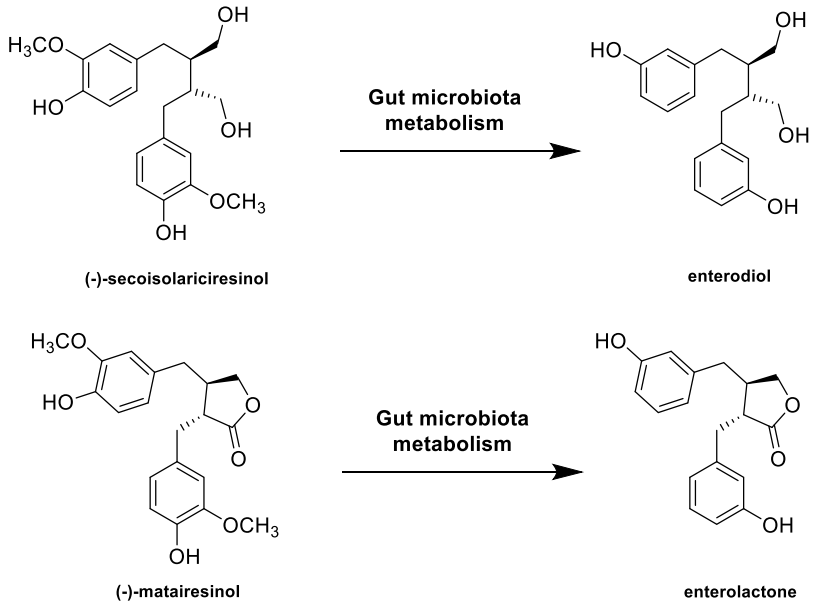


Figure 1.5. The synthesis of enterolignans in human intestines. The transformations are carried out by bacterial communities composed among others by *Peptostreptococcucus* and *Eubacterium* species which perform dehydroxylations and demethylation on (-)-matairesinol and secoisolariciresinol which activate their biological activities.[67]

1.3.1 Cascade reactions for (-)-podophyllotoxin biosynthesis

Historically, the aryltetralin lignan (-)-podophyllotoxin has been widely employed in traditional oriental medicine through the application of *Podophyllum* and *Sinopodophyllum* plants and other related *Berberidaceace* species. Nowadays, the increasing focus towards traditional means for healthcare and medicine, plus the utilization of its derivatives as chemotherapeutic agents notably raised PTOX commercial demand, determining serious endangerment of the plant sources. In other words, the conventional ways of isolation became environmentally unbearable and its commercialization hampered due to the high costs of production.[73] To overcome these limitations, mixed approaches combining organometallic chemical catalysts with single or multiple enzymatic steps have been proven successful, easing PTOX commercialization.[74] Specifically, the preparation of PTOX, as well as its deoxygenated precursor or its epimer - which are all etoposide and teniposide precursors by themselves - has been successfully carried out via stereodivergent biotransformation and biocatalytic kinetic resolution of the corresponding dibenzylbutyrolactones with a 2-oxoglutarate-dependent dioxygenase.[75] However, because of increasing environmental awareness,

researchers are keen on disclosing the potential of multi-enzymatic cascades avoiding cumbersome although elegant chemo-synthetic steps. As introduced in the previous sections, the production of highly specialized chemicals could be targeted and developed in more feasible and environmentally friendly ways.[76] Thinking of nature as a model, metabolic and enzymatic networks represent highly efficient systems of reaction cascades to churn out – for instance- a vast array of alcohols, aldehydes, acids, esters, ketones, in prokaryotic and eukaryotic cells.[77]

Over the years, the biosynthesis of lignans in plants has been widely investigated, and the enzymatic network leading to dibenzylbutanes like (-)-secoisolariciresinol and dibenzylbutyrolactones like (-)-matairesinol from coniferyl alcohol has been confirmed.[78, 79] Nevertheless, only the rise of next generation sequencing technology allowed to cast light on the further steps necessary to elucidate the biosynthetic route to aryltetralin lignans in *Podophyllum* genus.[70, 80] To date, no enzymes have been discovered to hydroxylate the last precursor (-)-deoxypodophyllotoxin (D-PTOX) to yield PTOX with the desired stereoselectivity. Instead, the enzymatic steps responsible for the synthesis of the epimer (-)-epipodophyllotoxin (Epi-PTOX) and the analogue (-)-desmethyl-epipodophyllotoxin were identified.[70] As mentioned previously, both compounds are direct precursors to the antineoplastic compounds etoposide and teniposide, meaning that the discovered enzymes provide another significant gear to develop a biocatalytic clockwork for the sustainable production of (-)-podophyllotoxin-derived pharmaceuticals (Figure 1.6).

1.3.1.1 Coniferyl alcohol to (+)-pinoresinol

Biosynthetic pathways towards lignans are typical for vascular plants, in which they start with the dimerization of the monolignol coniferyl alcohol yielding the common precursors (+)- or (-)-pinoresinol to further follow enantiocomplementary or other divergent routes.[64, 81, 82] As a result of the activity of an oxidizing enzyme like laccase or peroxidase, coniferyl alcohol undergoes one electron oxidation, generating a radical intermediate which then couples to a second radicalized coniferyl alcohol unit.[83]

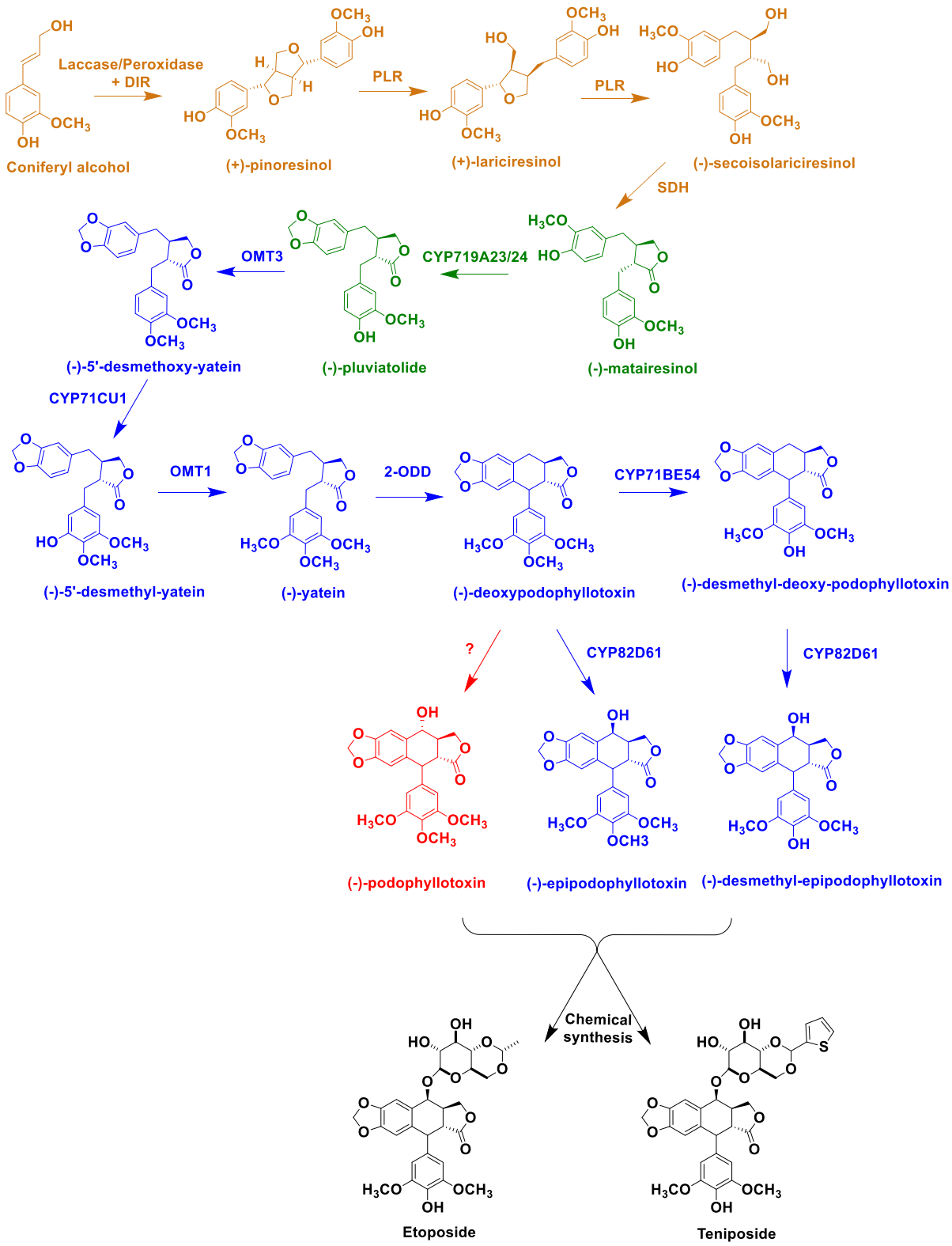


Figure 1.6. Overview on the metabolic pathway leading to aryltetralin lignans. In orange, the transformation of coniferyl alcohol towards dibenzylbutane and dibenzylbutyrolactone lignans (-)-secoisolariciresinol and (-)-matairesinol, as described by Davin *et al*., 1997[83], Ford *et al*.,2001[78], and Xia *et al*., 2001.[79] In green, the transformation of (-)-matairesinol to (-)-pluviatolide is drawn as elucidated by Marques *et al*., 2014.[80] In blue the cascade elongation is shown as discovered by Lau and Sattely, 2015.[70] The unknown enzymatic step leading to (-)-podophyllotoxin is highlighted in red, whereas in black the structures of etoposide and teniposide drugs are shown.

To gain stereospecificity of such reactions, the so-called Dirigent proteins (DIR) functionally act as navigators to drive the bimolecular phenoxy radical coupling mechanism towards regio- and stereo-selective product formation, preventing “random” dimerization and promoting enantiomeric pure pinoresinol synthesis (Figure 1.7).[81] In regard to the pathway leading towards aryltetralin lignans, a (+)-pinoresinol-forming DIR is believed to orientate coniferyl alcohol radicals, guiding the covalent bond to occur between C8 and C8’ of two coniferyl alcohol units. [82, 84]

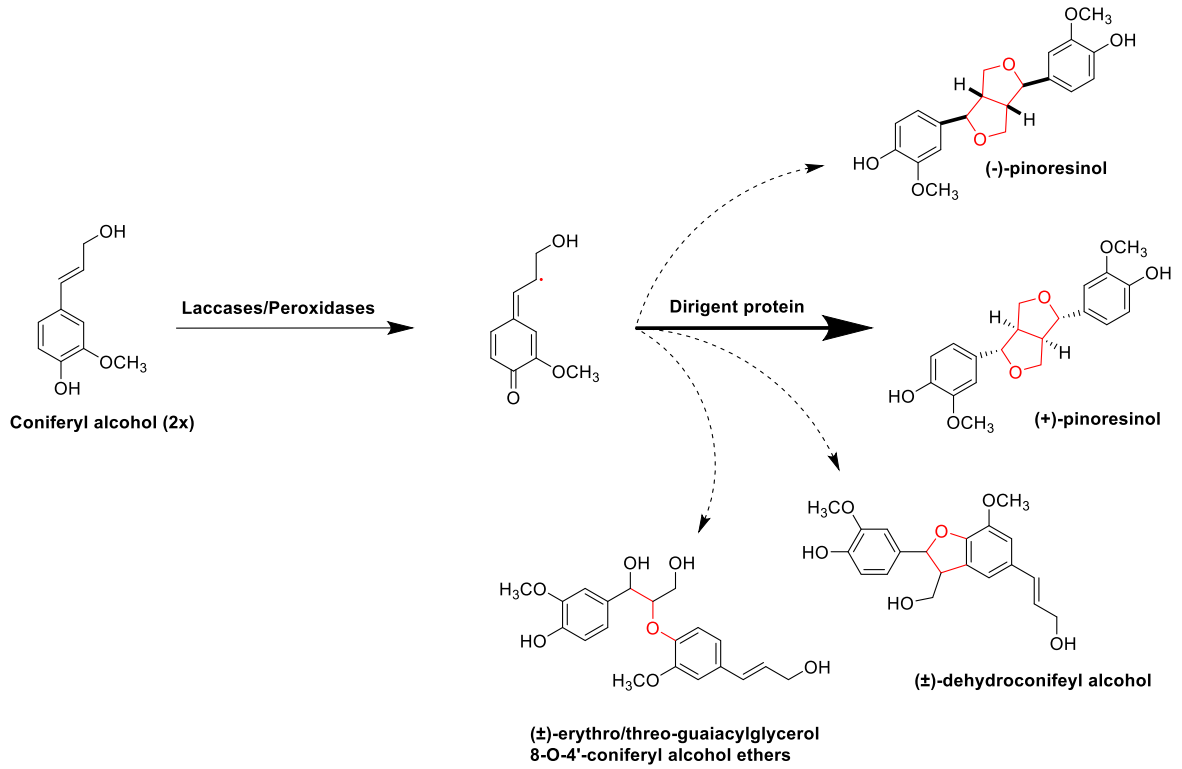


Figure 1.7. The one-electron oxidation of coniferyl alcohol leading to diverse dimerization products is depicted. In the presence of a (+)-Dirigent protein (DIR),[82] the coupling is guided to form (+)-pinoresinol.

Since no evidence has been reported describing DIR to possess catalytic activity, the oxidation of coniferyl alcohol has been achieved *in vitro* and *in vivo* by the application of oxidants such as ammonium peroxydisulphate, or oxidases like laccases (E.C: 1.10.3.2) and peroxidases (E.C: 1.11.1.-).[84-86] *In planta*, DIR are secreted in the apoplastic space and undergo post-translational modifications (PTMs) such as N-terminal cleavage and glycosylation. Especially the latter was hypothesised to be crucial for proper protein folding, influencing as well

enzymatic activity, solubility and stability.[81] Overall, PTMs may represent a crucial bottleneck for the application of DIR for biosynthetic purposes in prokaryotic hosts. More precisely, the heterologous expression in *E. coli* did not yield in active DIR so far.[81] This means that, if pondering the development of an *E. coli* based biocatalyst for the synthesis of lignans, the production of undesired dimers as a result of coniferyl alcohol coupling should be taken into account.

1.3.1.2 (+)-Pinoresinol to (-)-secoisolariciresinol

After pinoresinol formation, a pinoresinol-lariciresinol reductase (PLRs) catalyse two consecutive reduction steps to produce lariciresinol and secoisolariciresinol. This reaction represents a crucial step, as it drives the metabolic pathway towards the synthesis of dibenzylbutane, dibenzylbutyrolactone and aryltetralin lignans in an enantioselective manner.[87] PLRs are NADPH-dependent reductases, isolated and characterized mainly from *Forsythia*, *Thuja,* and *Linum* species.[87-89] On this subject, extensive investigation upon different PLRs revealed the presence of a conserved "GXXGXXG" sequence at the N-terminal end. Moreover, the characterization of PLRs catalytic activity gave proof of distinct, diverse enantioselective preferences among these enzymes, depending not only on the species but also on the plant organ or the developmental stage.[63, 90] For instance, both FiPLRs identified from *Forsythia intermedia* have been appointed with distinct preference towards (+)-pinoresinol and (+)-lariciresinol to produce (-)-secoisolariciresinol.[87] Differently, two PLRs from *Linum usitatissimum* - namely LuPLR1 and LuPLR2 - have been isolated and described to possess opposite enantiospecificity (Figure 1.8), divergent expression regulation and tropism, as they have been found in seed or leaves respectively.[89] A similar scenario has been described for TpPLR1 and TpPLR2 from *Thuja plicata*, with TpPLR1 showing preference for (-)-enantiomers and TpPLR2 for (+)-enantiomers. Structure-function analysis has been performed after solving the crystal structures of these enzymes in the attempt of clarifying the origin of the different enantiospecificity. Few amino acid residues appeared to remark a different local environment within the substrate-binding site that could generate the

enantiospecificity of TpPLRs.[91, 92] However, in spite of the extensive knowledge gathered around PLRs, more has to be done to fully understand the catalytic preferences, roles and regulation mechanisms of these enzymes.[63]

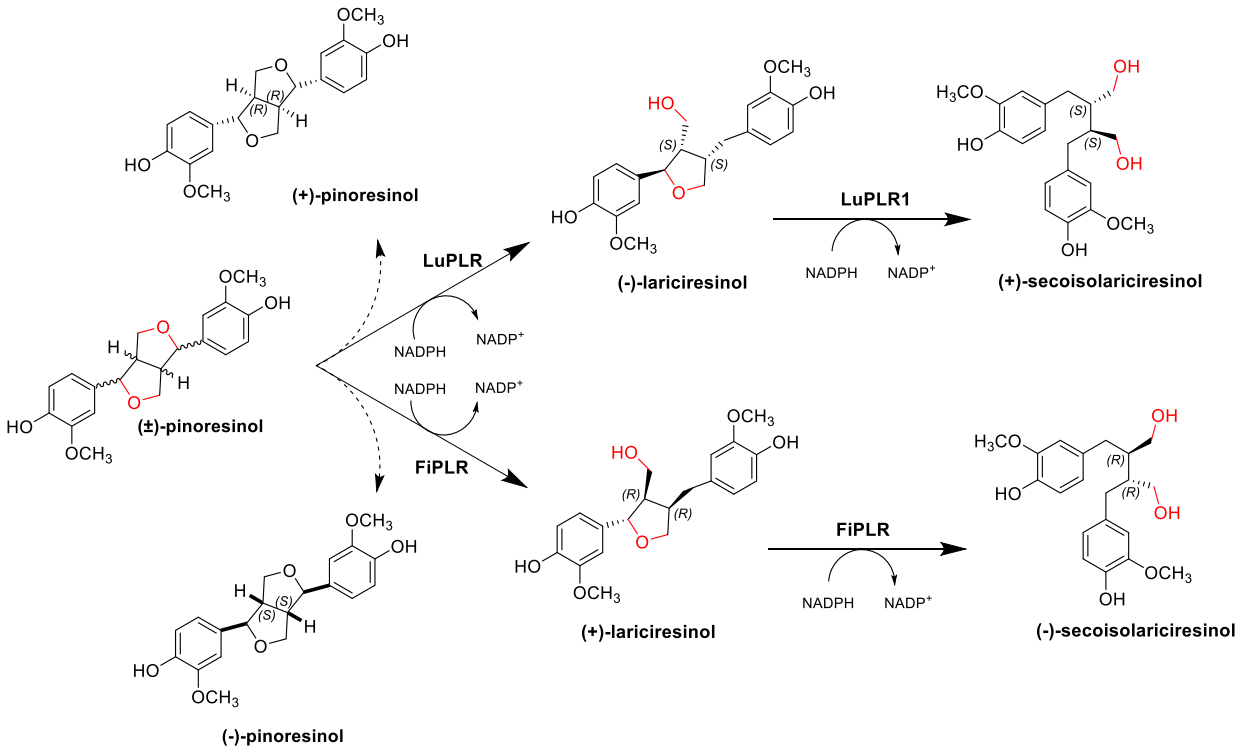
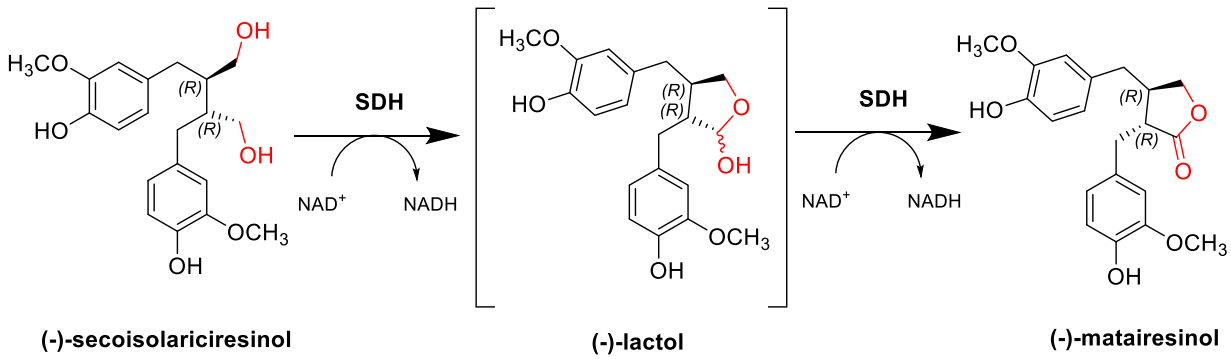


Figure 1.8. Pinoresinol-lariciresinol reductases of *Linum usitatissimum* and *Forsythia intermedia* are taken as examples to show the different enantiospecificity of the reaction depending on the organism the enzymes originate from.

1.3.1.3 (-)-Secoisolariciresinol to (-)-matairesinol

Proceeding along the path towards aryltetralin lignans, after the reduction of pinoresinol mediated by PLRs, the enantiospecific deprotonation of (-)-secoisolariciresinol promotes the formation of the dibenzylbutyrolactone lignan (-)-matairesinol via the (-)-lactol intermediate (Figure 1.9). This reaction is catalysed by secoisolariciresinol dehydrogenases (SDH), which are NAD^+-dependent enzymes belonging to the family of short-chain dehydrogenases/reductases (SDR). The isolation and purification of homologues from *Forsythia intermedia* and *Podophyllum peltatum*, namely SDH_Fi321 and SDH_Pp7 respectively, allowed the characterization of the two enzymes, which exhibited strong enantiospecificity towards (-)-secoisolariciresinol to produce (-)-matairesinol. Moreover, the presence of a highly conserved NAD-binding site, composed by the triad Ser^{153}, Tyr^{167} and Lys^{171}, was identified by sequence alignment among SDH_Fi321, SDH_Pp7 and other SDRs like CPRD12 from *Vigna unguiculata*, adNt1 from *Nicotiana benthamiana* and adAt1 from *Arabidopsis thaliana.*[79] Further, deeper studies exploiting site-directed mutagenesis of the highly conserved catalytic triad of SDH_Pp7 allowed to elucidate the mode of hydride abstraction from (-)-secoisolariciresinol and its subsequent oxidation. In the light of this discovery, a catalytic model for the enantiospecific synthesis of (-)-matairesinol was proposed by Moinuddin *et al.*[93]



(-)-secoisolariciresinol **(-)-lactol** **(-)-matairesinol**

Figure 1.9. The double oxidation of (-)-secoisolariciresinol generating (-)-matairesinol via the formation of the intermediate (-)-lactol is depicted. Adapted from Moinuddin *et al.*, 2006.[93]

1.3.1.4 (-)-Matairesinol to (-)-pluviatolide

Only recently, the combination of next generation sequencing technologies, bioinformatics and metabolomic analysis, enabled the reconstitution of the next steps aryltetralin lignans pathway. Specifically, the transcriptome analysis performed on *Sinopodophyllum hexandrum* and *Podophyllum peltatum* led to the discovery of two P450s performing the conversion of (-)-matairesinol to (-)-pluviatolide via methylenedioxy-bridge formation (MDB).[70, 80] In plants, the occurrence of MDB structures is often found within several specialized metabolites belonging to lignans, alkaloids and isoflavones, upon which the oxidative formation of MDB moieties is attributed to the action of two cytochrome P450 families, CYP81Q and CYP719A.[94] To discover homologues enzymes to appoint as (-)-pluviatolide synthases, Marques and co-workers performed sequence alignment of the assembled transcriptomes of *P. hexandrum* and *P. peltatum* against known MDB-forming P450s genes such as CYP81Q1 and CYP81Q2 from *Sesamum indicum,*[95] and CYP719A1 from *Coptis japonica.*[29] Eventually, two enzymes were identified as putative (-)-pluviatolide synthases, namely CYP719A23 and CYP719A24 respectively from *S. hexandrum* and *P. peltatum*, and successfully expressed in *S. cerevisiae*. Both exhibited catalytic activity against (-)-matairesinol and showed ~68% homology with CYP719A1, with ~98% identity between each other.[80] Later, the confirmation of CYP719A23/A24 as (-)-pluviatolide synthases, and of (-)-pluviatolide as the direct precursor of (-)-podophyllotoxin, came by the work of Lau and Sattely.[70] Although the way of MDB formation within the CYP719A family is still unclear, the reaction mechanism proposed for the CYP81Q family could be adapted instead. Indeed, Noguchi and co-workers noted that both families lack presence of a distal threonine, which is normally conserved in P450s (e.g. P450BM3 and P450cam from *Pseudomonas putida),*[96] that is replaced by an alanine and by a serine in the CYP81Q and CYP719A classes respectively. According to their work, the distal threonine would hamper the formation of the MDB by altering the substrate-participated hydrogen-bonding network that is hypothesised to trigger the intramolecular nucleophilic cyclization characterizing the catalytic mechanism (Figure 1.10).[94] Thus far, the proposed reaction mechanism highlights MDB-forming P450s as performers of a unique form of P450-

mediated oxygenation, as no oxygen is integrated into the substrate to give the final reaction product.

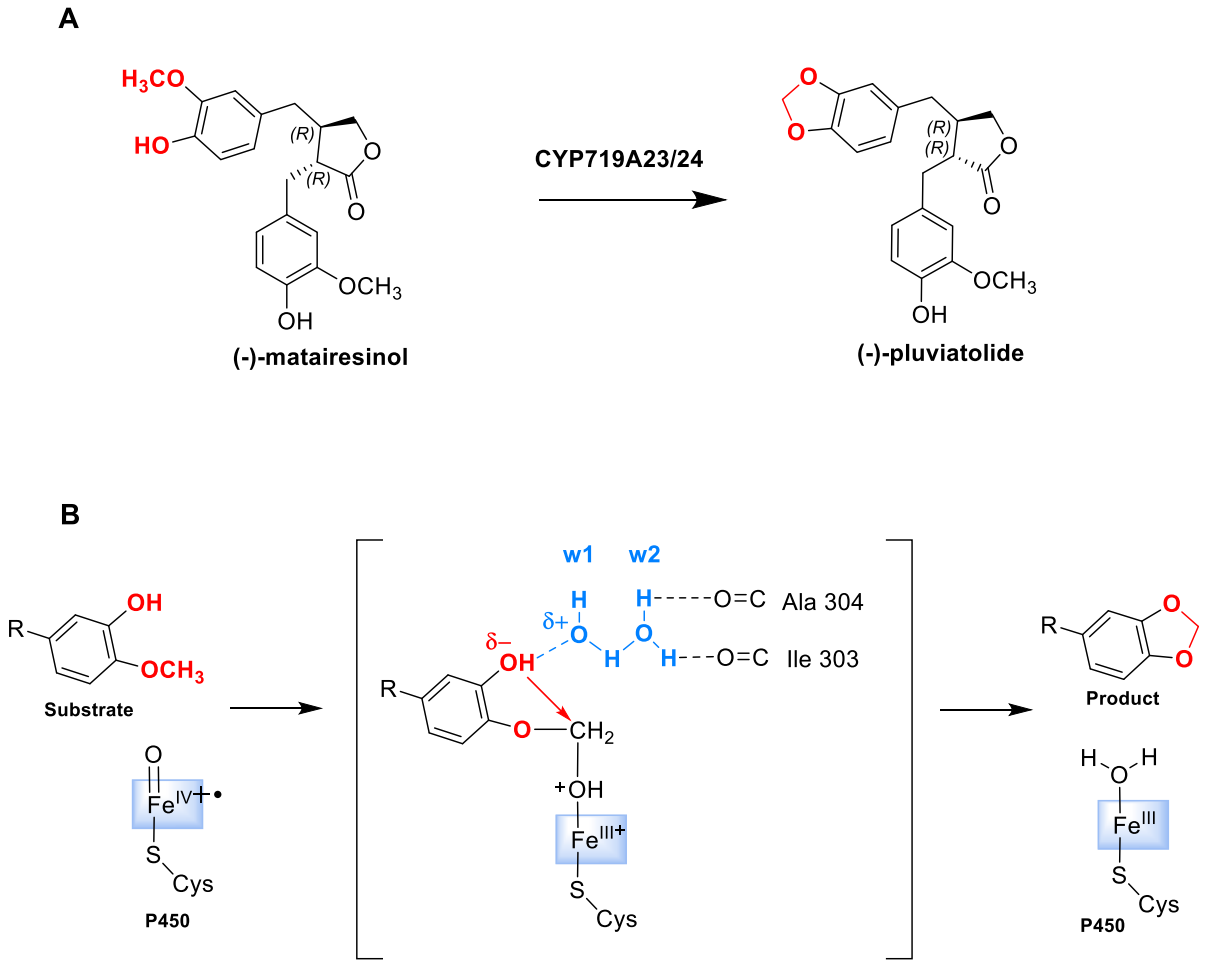


Figure 1.10. **A**: The methylenedioxy bridge (MDB) formation occurring on (-)-matairesinol by the action of CYP719A23 or A24 is depicted. **B**: The proposed catalytic mechanism for MDB-forming P450s is drawn adapted from Noguchi *et al*., 2014. The interaction among the water molecule (w2) with the carbonyl groups of the P450 apoprotein main chain – in this case Ala304 and Ile303 from CYP81Q1- starts the hydrogen bond relay towards the hydroxy group of the substrate via two water molecules (w1 and w2). The nucleophilic attack (red arrow) towards the proximal methoxy group of the substrate is provoked by the activated oxygen, resulting in the formation of the MDB when the bond between the substrate and the heme is cleaved.[94] The blue box holding the iron in different oxidation states represent the heme-group of the P450, as already depicted in Figure 1.2.

1.3.1.5 (-)-Pluviatolide to (-)-deoxypodophyllotoxin

According to Lau and Sattely,[70] the combination of qRT-PCR with hierarchical RNA-seq clustering allowed the genome mining of *S. hexandrum* to identify the expression profiles of potential candidates for the next steps after (-)-pluviatolide formation. This approach reduced the selection down to several putative O-methyltransferases (OMTs), P450s, a 2-oxoglutarate/Fe(II)-dependent dioxygenase (2-ODD), and polyphenol oxidases (PPOs). Among all, an O-methyltransferase (OMT3) was first identified, that accepts both (-)-matairesinol and (-)-pluviatolide as substrates (Figure 1.11). Although OMT3 seems not to have strong substrate specificity, the determination of kinetic parameters *in vitro* revealed efficient (-)-pluviatolide methylation [(K_m) = 1.4 μM, (k_{cat}) = 0.72 s^{-1}].[70] Based on these data, OMT3 was accepted as the next-in-line enzyme in the pathway, generating the intermediate (-)-5'-desmethoxy-yatein (Figure 1.11).

For the identification of the following step, further computational analysis based on the gene expression profiles within *S. hexandrum* tissues narrowed the selection to six putative hydroxylases for the synthesis of (-)-desmethyl-yatein. Using (-)-matairesinol as substrate, these were all singularly coexpressed with CYP719A23 and OMT3 in *Nicothiana benthamiana* leaves, identifying CYP71CU1 as the only candidate performing hydroxylation of (-)-5'-desmethoxy-yatein. The functionalization (-)-5'-desmethoxy-yatein by the action of CYP71CU1 was confirmed in isolated microsomal fraction after successful heterologous expression in *S. cerevisiae*, likely endorsing this enzyme as the next gear in the metabolic clockwork.

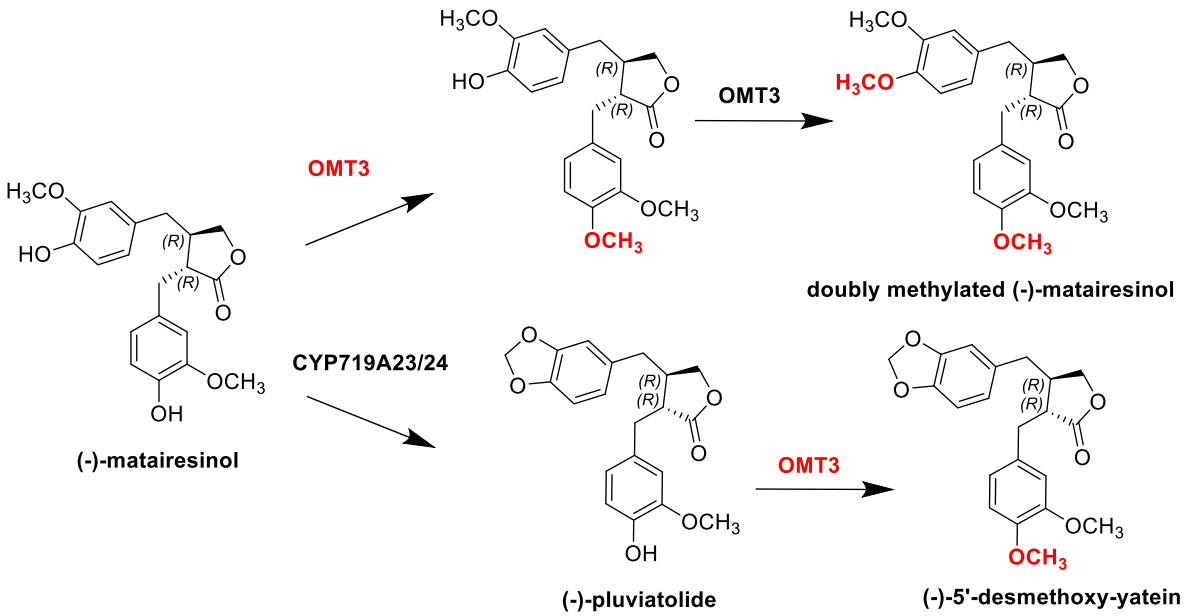


Figure 1.11. Overview on the reactions catalysed by OMT3. Below, the proposed activity as (-)-pluviatolide-O-methyltransferase is depicted. Above, OMT3 catalyses the double-methylation of (-)-matairesinol as metabolic side-products.

In the past, feeding experiments performed on *S. hexandrum* identified (-)-yatein as a close precursor of (-)-podophyllotoxin.[97] Using the same approach in *N. benthamiana* leaves infiltrated with (-)-matairesinol, Lau and Sattely identified the O-methyltransferase OMT1 as responsible for the methylation of (-)-5'-desmethyl-yatein to yield (-)-yatein (Figure 1.12).

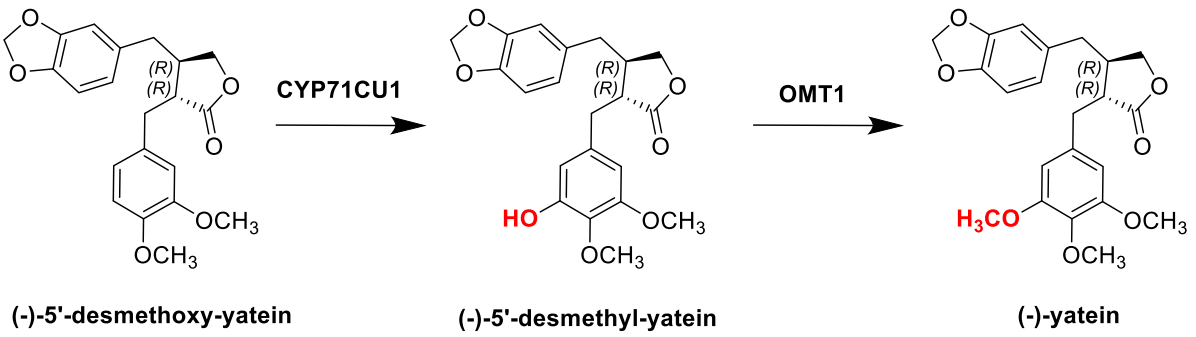


Figure 1.12. The hydroxylation with the subsequent methylation performed by CYP71CU1 and OMT1 are depicted.

Last, the step from (-)-yatein to (-)-deoxypodophyllotoxin necessarily involves the closure of the ring system by carbon-carbon bond formation to generate the aryltetralin scaffold.[97] This reaction was confirmed to be catalysed by a 2-ODD, as a result of a first hydroxylation at the 7' carbon, followed by dehydration generating a quinone methide intermediate prior to the stereoselective cyclization (Figure 1.13).[70]

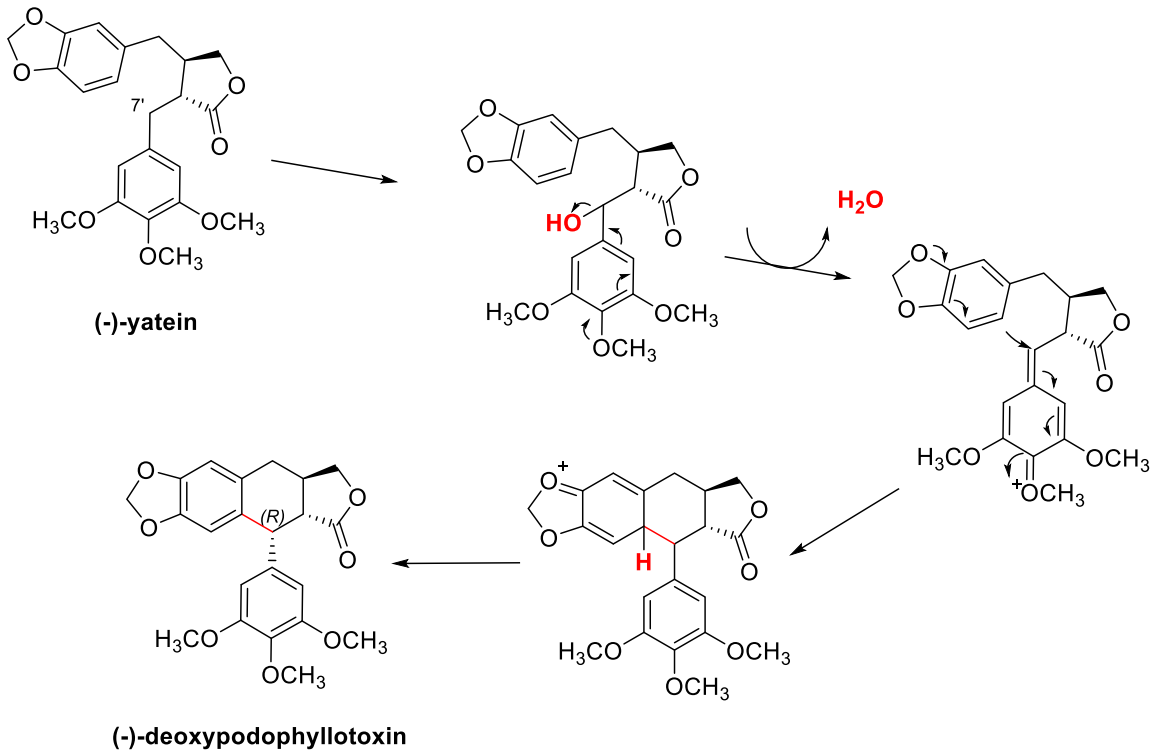


Figure 1.13. The conversion of (-)-yatein to (-)-deoxypodophyllotoxin is depicted. The reaction mechanism for the arytetralin backbone closure by 2-ODD is drawn; adapted from Lau and Sattely, 2015. [70]

Overall, without considering P450s which have been extensively described previously, all these enzymes represent important components in the biosynthesis secondary metabolites in plants. In this regard, 2-oxoglutarate/Fe(II)-dependent dioxygenases significance is often underscored. In fact, these non-heme iron-containing enzymes can be compared to CYPs both in terms of reaction versatility and role as oxidative tailors of complex molecules.[98] Moreover, they do not require any redox partner to sustain their activity or expensive nicotinamide cofactors as electron source, as 2-ODDs co-substrate is 2-oxoglutarate. In the same way, O-methyltransferases, which belong to the large family of S-adenosyl-L-

methionine-dependent methyltransferases, are widely distributed all along the metabolic networks of plants. Within this context, OMTs contribute to the diversification of metabolites and their biological functions, for instance reducing the reactivity of phenolic hydroxyl groups, altering the solubility and activity against pathogens.[99]

1.3.1.6 The unknown step: (-)-deoxypodophyllotoxin hydroxylation

To date no enzymes have been identified to perform the direct C7-hydroxylation on the aromatic ring of (-)-deoxypodophyllotoxin with the correct stereochemistry to produce (-)-podophyllotoxin. However, biosynthetic means have been proven successful in producing derivatives and epimers which can be used to chemically generate the essential pharmaceuticals etoposide and teniposide (Figure 1.6). In sought for a P450 enzyme as putative (-)-podophyllotoxin synthase within *S. hexandrum*, Lau and Sattely discovered instead two P450s, CYP71BE54 and CYP82D61, that catalyse the oxidation of (-)-deoxypodophyllotoxin yielding (-)-desmethyl-deoxy-podophyllotoxin, (-)-desmethyl-epipodophyllotoxin and (-)-epipodophyllotoxin.[70] Heterologous expression of these P450s in *N. benthamiana* has provided a proof of concept which paves the way for reconstituting the arytetralin lignans pathway even in prokaryotic hosts. Complementary to this, the interest in producing actual (-)-podophyllotoxin by biosynthetic means remains pivotal. Needless to say, that more efforts are required in scavenging plant genomes and in the understanding of the physiologic metabolic network responsible for (-)-podophyllotoxin synthesis in plants. Alternatively, screening for substituting enzymes to be engineered could represent a breakthrough. Aiming to the hydroxylation of (-)-deoxypodophyllotoxin, bacterial P450s might represent starting candidates. In fact, numerous successful protein engineering examples with P450 BM3 or other bacterial P450s have been reported targeting the on-demand compounds with optimized activity and selectivity.[100, 101]

1.4 Aim of the work

Lignans are plant secondary metabolites of particular interest. These compounds have been described as highly potent in the treatment and prevention of cancers and cardiovascular diseases, with anti-inflammatory and antioxidant activities reported as well. Prominent examples of lignans acknowledged for such health-promoting activities are (±)-pinoresinol, (-)-secoisolariciresinol, (-)-matairesinol, and (-)-podophyllotoxin. For these reasons, lignans are used as active pharmaceutical ingredients (APIs) to develop drugs and dietary supplements and encountered a growing commercial demand. Traditionally, lignans are extracted directly from plants, resulting in the endangerment of the original sources.

In this work, performed within the framework of the BMBF funded LignaSyn project, we aimed to establish an artificial biosynthetic pathway to target the biosynthesis of lignans from coniferyl alcohol in *E. coli*. A modular and versatile enzyme-based platform was developed exploiting enzymes from different plants such as *Forsythia intermedia*, *Podophyllum peltatum* and *Sinopodophyllum hexandrum*. Developmental steps were:

1. Optimization of the multi-copper oxidase-mediated coupling of coniferyl alcohol to (±)-pinoresinol in an *E. coli*-based whole-cell biocatalyst. Development of a multi-enzyme cascade for the biotransformation of coniferyl alcohol to (-)-matairesinol (2.1).
2. Development of a whole-cell multi-enzyme cascade to produce the non-commercially available lignan (-)-pluviatolide from (+)-pinoresinol at a preparative scale (2.2).
3. Reconstitution of *S. hexandrum* biosynthetic pathway from (-)-matairesinol to (-)-deoxypodophyllotoxin and (-)-epipodophyllotoxin in *E. coli*. Additionally, activity of the two bacterial P450s CYP107Z and CYP105D for the synthesis of (-)-podophyllotoxin or its analogues was evaluated (2.3)

2 Results

Within the following three chapters the results achieved are presented. Overall, the three chapters describe the development of combinable, independent *"modules"* to target the biosynthesis of various high-lignans analogues towards the potent antiviral agent (-)-podophyllotoxin. Amongst these, chapter 2.1 and 2.2 will present the published articles, whereas chapter 2.3 will disclose the unpublished results, organized in a submitted manuscript. The chapters will follow the progressive functionalization of lignans over the metabolic pathway. Starting from the oxidative coupling of the monolignol coniferyl alcohol to (+)-pinoresinol, the syntheses of (-)-matairesinol and (-)-pluviatolide are described in chapter 2.1 and 2.2. Further, the multi-enzyme cascade was extended towards (-)-podophyllotoxin-like lignans such as (-)-deoxypodophyllotoxin and (-)-epipodophyllotoxin (chapter 2.3).

Own contributions to the single parts are given in the preface to each chapter.

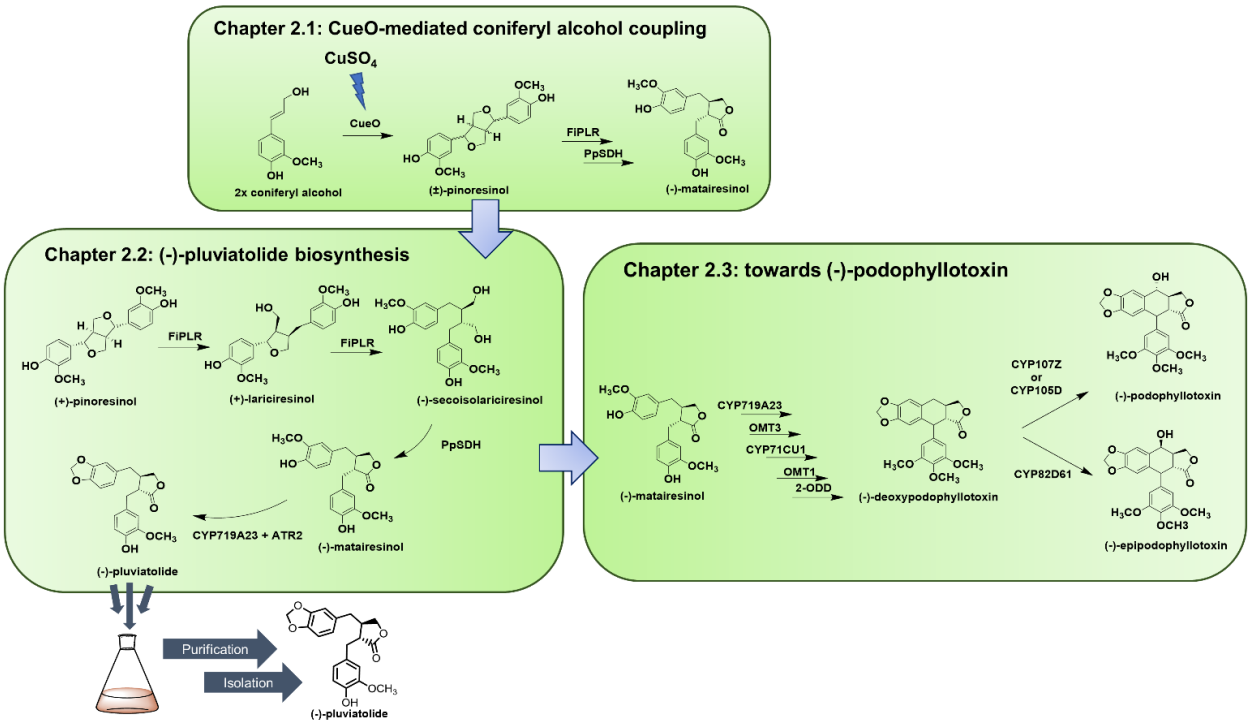


Figure 2.1. Schematic overview of the chapters composing this thesis.

2.1 Copper-mediated enzymatic coniferyl alcohol coupling

Title: *"Use of copper as a trigger for the in vivo activity of E. coli laccase CueO: a simple tool for biosynthetic purposes"*

Authors: Davide Decembrino, Marco Girhard, Vlada B. Urlacher

Published in: *ChemBioChem* 2021, *22*, 1470-1479. (Wiley-VHC GmbH)

DOI: 10.1002/cbic.202000775

Own contribution: conceptualization, design, performance of the experiments, analysis and evaluation of the data, drafting of the manuscript and the artworks. Relative contribution 90%

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

Very Important Paper

Use of Copper as a Trigger for the *in Vivo* Activity of *E. coli* Laccase CueO: A Simple Tool for Biosynthetic Purposes

Davide Decembrino,[a] Marco Girhard,[a] and Vlada B. Urlacher*[a]

Laccases are multi-copper oxidases that catalyze the oxidation of various electron-rich substrates with concomitant reduction of molecular oxygen to water. The multi-copper oxidase/laccase CueO of *Escherichia coli* is responsible for the oxidation of Cu^{+} to the less harmful Cu^{2+} in the periplasm. CueO has a relatively broad substrate spectrum as laccase, and its activity is enhanced by copper excess. The aim of this study was to trigger CueO activity *in vivo* for the use in biocatalysis. The addition of 5 mM $CuSO_4$ was proven effective in triggering CueO activity at need with minor toxic effects on *E. coli* cells. Cu-treated *E. coli* cells were able to convert several phenolic compounds to the corresponding dimers. Finally, the endogenous CueO activity was applied to a four-step cascade, in which coniferyl alcohol was converted to the valuable plant lignan (−)-matairesinol.

Introduction

Laccases belong to the superfamily of multi-copper oxidases, which are found in insects, fungi, bacteria, and plants.[1] They catalyze the one-electron oxidation of electron-rich phenolic and nonphenolic compounds with concomitant four-electron reduction of molecular oxygen to water. Generally, laccases possess a wide substrate spectrum, do not require any external cofactors, and produce water as the only by-product. These characteristics make them interesting tools for biotechnological applications ranging from wastewater bioremediation, the decolorization of industrial dyes and the degradation of pharmaceutical micropollutants to manufacturing biologically active compounds.[2] Consequently, diverse laccases from many organisms have been discovered, characterized and widely applied as biocatalysts.[3]

Bacterial laccases have recently attracted attention, as they offer several advantages in comparison to fungal or plant laccases.[4] Generally higher thermal and pH stability of bacterial laccases, as well as lower sensitivity towards inhibitors in comparison to eukaryotic ones have been reported. A number of bacterial laccases have been identified, isolated and characterized.[5] However, the expression of recombinant laccases in *Escherichia coli* might be affected by incomplete incorporation of Cu ions into the apoprotein and by interference with activity of the endogenous multi-copper oxidase CueO.[6]

For the past 20 years, the structure and function of CueO, whose former gene name was *yacK*, have been intensively investigated.[7] As part of the homeostasis system in *E. coli*, CueO is responsible for the oxidation of Cu^{+} to the less harmful Cu^{2+} in the periplasm, thereby preventing Cu^{+} from entering the cytoplasm.[8] Typically for multi-copper oxidases, CueO possesses a Type 1 copper site and a Type 2/Type 3 trinuclear copper cluster that together are responsible for substrate oxidation and oxygen reduction.[9] The binding site of CueO is hidden by a methionine-rich helical region that was suggested to prevent the access of bulky organic substrates, determining enzyme specificity as a cuprous oxidase.[10] At low environmental copper concentrations, CueO is expressed and folded into the cytoplasm and then transported to the periplasm via the twin arginine pathway as apoprotein.[8] Once in the periplasm, copper ions assemble as metal cofactors to apo-CueO determining structural stabilization and catalytic activation of the enzyme. In this regard, the addition of external Cu^{2+} ions, like copper sulfate ($CuSO_4$), has been demonstrated to readily activate the enzyme, leading to the full incorporation of four copper ions in the T1, T2 and T3 centers.[8,11] Previous studies have revealed that CueO has a rather broad substrate spectrum with low activity *in vitro*, yet the activity increased in the presence of Cu^{2+} excess.[7a] Recently, CueO was overexpressed in *E. coli* and its activity was tested *in vitro* for oxidative coupling of phenolic compounds.[12]

Within this study, the use of the *in vivo* oxidase activity of CueO triggered by the addition of external copper ions to *E. coli* cells for oxidative phenolic coupling was investigated. This approach represents a suitable solution for biocatalytic applications. For instance, in cascade reactions, laccase activity could be switched on only when needed and without host manipulation. Oxidative coupling of various phenolic compounds with *E. coli* cells was tested under copper induced stress. A major focus was set on the dimerization of coniferyl alcohol to the lignan (±)-pinoresinol, because in a previous study from our group, dimers of coniferyl alcohol were formed by exploiting an *E. coli*-based whole-cell biocatalyst, expressing the recombinant CgL1 laccase from *Corynebacterium glutamicum*.[13] As a proof of

[a] D. Decembrino, Dr. M. Girhard, Prof. V. B. Urlacher
Institute of Biochemistry
Heinrich-Heine University Düsseldorf
Universitätsstrasse 1, 40225 Düsseldorf (Germany)
E-mail: vlada.urlacher@uni-duesseldorf.de

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202000775

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

concept, we developed an *in vivo* four-step cascade to prove the applicability of this approach. Endogenous CueO activity was switched on by $CuSO_4$ addition, and combined with heterologously expressed pinoresinol-lariciresinol reductase from *Forsythia intermedia* (FiPLR) and secoisolariciresinol dehydrogenase from *Podophyllum pleianthum* (PpSDH). Utilizing this setup, up to 5 mM coniferyl alcohol were efficiently converted to the valuable lignan (−)-matairesinol.

Results and Discussion

Copper toxicity and its influence on *E. coli* BL21(DE3) cells

Copper is a fundamental element for cell life and enzymatic functions, yet its toxicity to biological systems has been widely recognized.[14] Specifically, lethal cell damaging related to environmental excess of copper has been associated with the displacement of iron from iron–sulfur proteins leading to the inactivation of enzymes involved in TCA cycle and pentose phosphate pathway, but also with the formation of oxygen reactive species (ROS).[11a,15] Although the occurrence of oxidative damage on DNA was excluded, the addition of $CuSO_4$ to *E. coli* directly correlated with superoxide generation in the cells. Yet, $CuSO_4$ is often supplemented within heterologous expression of bacterial laccases in *E. coli* to achieve complete copper loading in the target enzymes.[16] For these reasons, we first tested the effect of increasing concentrations of $CuSO_4$ on growing *E. coli* cells and their viability. Concerning aerobic growth conditions in complex media, sensibility of *E. coli* cells towards copper has been reported to occur at a millimolar concentration; specifically, in LB medium the reported minimum inhibitory concentration (MIC) was 3.5 mM $CuSO_4$.[16b,17]

Cell cultivation was carried out in TB medium in the presence of increasing $CuSO_4$ concentrations (3, 5, 10, 15, 20 or 30 mM), which revealed mild growth inhibition after addition of 3, 5, 10 and 15 mM compared to the negative control without copper added (Figure 1A). Within 48 h cell growth, the OD_{600} value of the negative control reached ~12, while under addition of up to 15 mM $CuSO_4$ OD_{600} values of ~10 were recorded. As the major difference between LB and TB media is the addition of glycerol as carbon source and phosphate buffer to temper pH variations, these results suggest that the growth of well-fed and metabolically active *E. coli* cells is indeed influenced by copper ion concentrations >3 mM, yet copper tolerance appears higher than the previously reported MIC in LB medium.

Cell viability was additionally investigated by counting the number of colony forming units (CFUs) on agar plates. Aliquots taken immediately after addition of 3, 5, or 10 mM $CuSO_4$ (t = 0 h) indicate that copper significantly affects the number of CFU as cell viability apparently drops by ~30–40% as compared to the negative control (Figure 1B). However, after 24 h incubation time, the cell viability is recovered, since similar numbers of CFU were counted for the samples supplemented with up to 10 mM $CuSO_4$ (Figure 1B). To explain this behavior, metabolic shock with following adaption of the cellular environment is hypothesized. For instance, iron–sulfur dehydratase clusters, and among these the fumarase A involved in the TCA cycle, have been reported to be inhibited by Cu ions. Nonetheless, the activity of this enzyme was shown to be protected from Cu ions if the active site was already occupied by the substrate malate, but also restored once the inhibiting copper was removed.[15] Because of this, it is possible that *E. coli* cells metabolism is harmed by $CuSO_4$ addition at first, and then recovers once copper ions are cleansed or detoxified by the homeostasis systems.

15 mM $CuSO_4$ appears as a threshold value for *E. coli* resilience: Similar OD_{600} values to those recorded with 3–10 mM $CuSO_4$ were achieved (Figure 1A), but the viability on agar plates dropped by ~90% after 24 h of incubation (Figure 1B). Differently, cell growth was strongly inhibited when 20 or 30 mM $CuSO_4$ concentrations were used (Figure 1A), and no colonies were detected on agar plates after 24 h.

For further analysis, *E. coli* cells were analyzed by microscopy. Cells from the negative control and experiments with 3 and 5 mM $CuSO_4$ appeared evenly distributed over the slide surface with no clear differences in cells morphology. Concerning 10 mM CuSO4, *E. coli* cells were evenly distributed on the

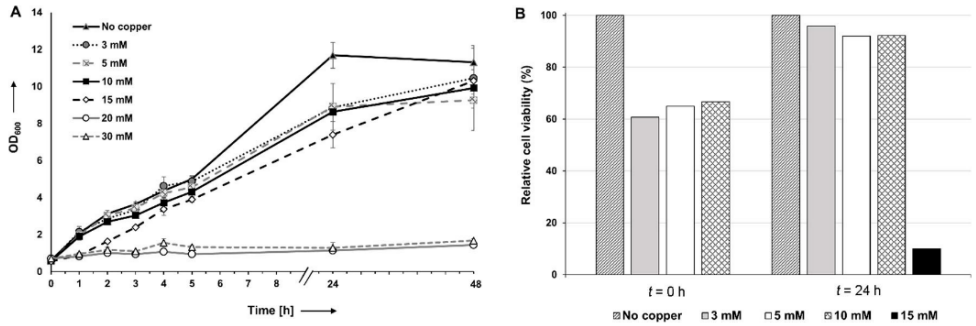


Figure 1. Effect of $CuSO_4$ on *E. coli* BL21(DE3) cell growth and viability. A) Cell growth in complex medium (terrific broth). B) *E. coli* cell viability on 3, 5, 10 and 15 mM $CuSO_4$ immediately after copper addition (time point 0 h), and after 24 h. Cell viability 0 h after addition of 15, 20, and 30 mM and 24 h after the addition of 20, and 30 mM $CuSO_4$ are not shown because no colonies were detected on agar plates. The number of CFUs from the negative control without $CuSO_4$ was set as 100%. Shown data result from two independent measurements.

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

glass slide, but the cells appeared smaller (Figure S1 in the Supporting Information).

Dimerization of phenolic compounds by $CuSO_4$-triggered CueO

The oxidase activity of CueO depends on the presence of environmental Cu ions and it has been proven to be active on diverse laccase substrates *in vitro*.[12,18] In light of this, we chose 5 mM $CuSO_4$ to trigger CueO activity in *E. coli* cells for the oxidation of phenolic compounds with subsequent dimerization. Generally, dimers of phenolic compounds have been reported to gain enhanced biological activities compared to the monomers.[19] Several compounds including 2,6-dimethoxyphenol (2,6-DMP), ferulic acid, coniferyl alcohol, tyrosol, *trans*-resveratrol and 17β-estradiol were chosen for tests for the following reasons. 2,6-DMP and ferulic acid have been described as substrates for CueO that catalyzes their oxidation *in vitro*.[12] Other compounds are reported substrates for different laccases. One-electron oxidation of phenolic compounds catalyzed by a laccase leads to the formation of phenoxy radicals which couple to dimers or/and oligomers. Dimers of 2,6-DMP were reported to possess higher antioxidant capacity than the substrate itself.[20] The reaction with 2,6-DMP (*m/z* 173 $[M+H+H_2O]^+$) oxidation turned bright yellow and three products could be detected by LC/MS. The first product with *m/z* 307 $[M+H]^+$ was identified as the desired dimer, whereas the two others were characterized by *m/z* 293 $[M+H-CH_3]^+$ and *m/z* 278 $[M+H-2CH_3]^+$, thus suggesting the loss of one or two methyl groups, respectively (Figure S2 and Table S1).[21]

Concerning ferulic acid (*m/z* 196 $[M+2H]^+$), whose dimers have been reported to increase its antioxidant activity, four products were detected. Amongst these, the analysis of the ionization patterns suggested two of them being the β–β and β–5 dimers (*m/z* of 387 $[M+H]^+$; Figure S3), coherently to those described in the literature. The MS data of the other detected products suggest the loss of CHO_2 by the dimers (*m/z* 341; Figure S3 and Table S2).[22]

Starting with coniferyl alcohol, the formation of four coupling products was observed. These compounds were identified as (±)-erythro/threo-guaiacylglycerol 8-*O*-4′-coniferyl alcohol ethers, (±)-dehydrodiconiferyl alcohol, and (±)-pinoresinol (Figure S4 and Table S3), based on the knowledge achieved from a previous work from our group.[23] Together with the observed coupling products, coniferyl aldehyde (*m/z* 179 $[M+H]^+$) was identified as resulting from the activity of endogenous dehydrogenases in *E. coli*.[13,24] Among the coupling products, the β–5 dimer dehydrodiconiferyl alcohol and the β–β dimer (±)-pinoresinol have been reported to possess interesting features: dehydrodiconiferyl alcohol was proven to exhibit anti-adipogenic, anti-oxidant and anti-inflammatory activities, and also appeared active in osteoblasts differentiation process;[25] (±)-pinoresinol is known for anti-inflammatory, anticancer and anti-microbial activities.[26]

Three products were formed during conversion of resveratrol (*m/z* 229 $[M+H]^+$). However, only one out of three showed *m/z* 455 $[M+H]^+$ corresponding to resveratrol dehydrodimers reported in the literature.[27]

Because of poor ionization, tyrosol could be identified only in the UV/Vis spectrum; here, one conversion product was observed (Figure S5). No activity towards 17β-estradiol could be detected under the tested conditions, which might be due to high reduction potential or large size of this substance which make it inappropriate as a substrate for CueO.

Optimization of endogenous CueO activity in *E. coli* for coniferyl alcohol coupling

Influence of oxygen supply and copper concentration: With the idea of further development of a *E. coli* whole cell biocatalyst featuring the multi-copper oxidase activity of CueO, we chose coniferyl alcohol as starting substrate to produce (±)-pinoresinol, which is the main precursor for higher-lignans biosynthesis.[28] As multi-copper oxidases rely on oxygen for their activity, in the first test, the conversion of coniferyl alcohol was carried out in tubes with open or sealed lids.

Coniferyl alcohol was efficiently converted (~85%) in tubes with open lid when 5 mM $CuSO_4$ were added to the cells. Contrary, in closed tubes low conversion (~4%) was observed even after $CuSO_4$ addition (Figure 2A). No substrate conversion was observed in tubes with open lids without added copper (Figure 2B). These results can be explained based on the mode of action of the two major components of copper homeostasis in *E. coli*, namely the *cue* system and the *cus* system.[11a,17] According to the literature, the *cue* system is identified as the primary homeostasis system, either in aerobic or anaerobic conditions, and its genes *copA* and *cueO* are described as basally expressed with mild-overexpression of ~12-fold enhancement in case of increasing copper stress. The efflux *cus* system is only strongly expressed by ~800-fold, if the primary *cue* system is overwhelmed. Nonetheless, in anaerobic conditions, the *cus* system is described to take over the primary role, enhancing copper efflux supported by CopA, whereas CueO is not involved in copper tolerance anymore.[11b] All in all, as multi-copper oxidase require oxygen for catalysis, this might represent a complementary explanation for the limited CueO activity which was observed under limited air availability.

Next, we tested lower and higher $CuSO_4$ concentrations to verify its influence on laccase activity of *E. coli* cells. CueO specific activity *in vitro* was reported at 3–4 mM $CuSO_4$ concentration;[8] for this reason we chose diverse copper concentrations ranging from 1–10 mM. Our results revealed that increasing copper concentrations resulted in increased coniferyl alcohol conversion with 5 mM $CuSO_4$ leading to ~70% and 10 mM $CuSO_4$ to 100% conversion (Figure 3A). Aiming to reduce copper usage and in order to be able to compare the reaction conditions, we used 5 mM $CuSO_4$ in further experiments because this concentration resulted in satisfactory substrate conversion.

In a previous work from our group, coniferyl alcohol dimerization to (±)-pinoresinol via *E. coli* biotransformation was achieved using heterologously expressed bacterial laccases

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

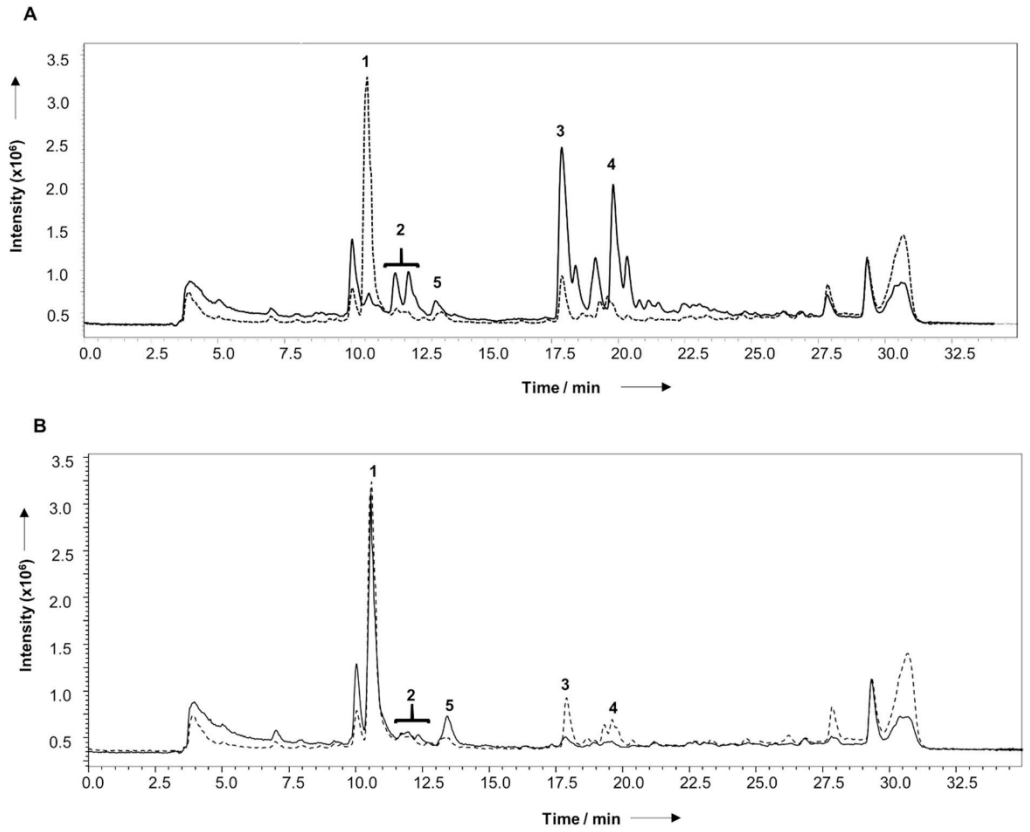


Figure 2. Comparison of coniferyl alcohol conversion catalyzed by *E. coli* cells under copper-induced stress in open and closed tubes. A) The reaction performed with 5 mM $CuSO_4$ under aeration (—) is plotted against the sample with 5 mM $CuSO_4$ in limited air availability (---). B) The sample with 5 mM $CuSO_4$ in limited air availability (---). is plotted against the control sample in an opened tube with no copper added (—). 1: coniferyl alcohol, 2: (±)-erythro/threo-guaiacylglycerol 8-*O*-4′-coniferyl alcohol ethers, 3: (±)-dehydrodiconiferyl alcohol, 4: (±)-pinoresinol, 5: coniferyl aldehyde.

including CgL1 from *C. glutamicum*.[23] For this reason, we decided to compare the endogenous CueO-system with heterologously expressed CgL1 and quantify the achieved (±)-pinoresinol concentration. Complete conversion of 200 μM coniferyl alcohol was achieved after 2 h with CueO whereas ~93% conversion was observed with CgL1, resulting in 65 and 40 μM (±)-pinoresinol, respectively. However, after 4 and 24 h the amount of the desired product (±)-pinoresinol decreased in both cases. This is no surprise, because phenolic compounds like (±)-pinoresinol are accepted by laccases as substrates as well, leading to the formation of oligo- or polymers.[29]

Complete substrate depletion by CueO was not expected, since the addition of 5 mM $CuSO_4$ in the previous experiments did not lead to full conversion. A reason for better performance might be related to cell permeability. Whereas in the previous experiments *E. coli* resting cells were used directly after cultivation, in these experiments cells were frozen after harvesting and thawed prior usage, allowing higher membrane permeability and increased substrate accessibility in the cell,[30] which likely resulted in improved conversion.

CueO led to higher (±)-pinoresinol concentration after 2 h compared to CgL1 (Figure 3B), which is probably due to the different reduction potentials of these enzymes. More precisely, the reduction potential reported for CueO is 440 mV, but 260 mV for CgL1, resulting in faster monomer oxidation by CueO, but also faster oxidation of the product (±)-pinoresinol (Figure 3B).[23,31]

In order to prove if this approach could be applied to higher substrate concentrations, a tenfold higher coniferyl alcohol load (2 mM) was tested, and 5 mM $CuSO_4$ was applied to trigger CueO activity. In this case, complete substrate conversion was achieved after 16 h, resulting in a (±)-pinoresinol concentration of 243±15 μM. This value has to be put in context with the product distribution, as the mechanism of bimolecular phenoxy radical coupling typical for multi-copper oxidases prevents regio- and stereoselectivity, leading always to a mixture of products.[32] In nature, stereoselective phenoxy radical coupling

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

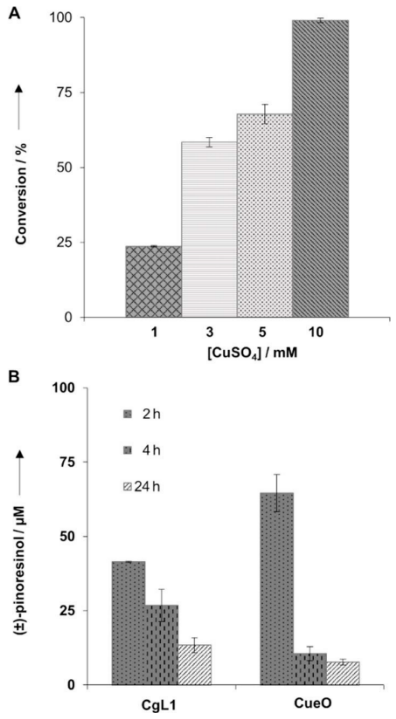


Figure 3. A) Influence of copper sulfate [mM] on CueO activity with 200 µM substrate. B) Comparison of (±)-pinoresinol concentrations measured at different time points with heterologous Cgl1 and endogenous CueO starting with 200 µM substrate.

is directed by the presence of so-called dirigent proteins, whose physiological role is to prevent random coupling of radicals occurring after the one-electron oxidation of monolignols like coniferyl alcohol performed by laccases or other oxidizing enzymes.[32] In absence of dirigent proteins, coniferyl alcohol radicals have been described to generate spontaneously (±)-dehydrodiconiferyl alcohol, (±)-erythro/threo-guaiacylglycerol 8-*O*-4′-coniferyl alcohol ethers and (±)-pinoresinol at a ~10:3:5 ratio.[33] Coherently, in our experiments, coniferyl alcohol radicals were most frequently coupled to (±)-dehydrodiconiferyl alcohol (~65%), while (±)-erythro/threo-guaiacylglycerol 8-*O*-4′-coniferyl alcohol ethers and the desired product (±)-pinoresinol set to ~8% and ~27%, respectively (Figure S6). This means that when starting with 2 mM coniferyl alcohol resulting in 1 mM of coupling products that are available after oxidative dimerization, ~250 µM (±)-pinoresinol is the theoretically possible titer. Thus, the detected (±)-pinoresinol concentration appears close to the theoretically achievable optimum.[33,34]

Validation of the roles of CueO and copper ions in coniferyl alcohol coupling: So far, it was assumed that CueO is responsible for coniferyl alcohol coupling, being readily activated by the addition of copper ions. As copper is widely used as catalyst to perform several reactions,[35] aerobic copper-mediated generation of coniferyl alcohol radicals was tested applying 3, 5 or 10 mM $CuSO_4$ with 2 mM substrate in phosphate buffer without cells or isolated enzymes. Approximately 25% depletion was observed with 3 mM $CuSO_4$ and ~35% with both 5 and 10 mM. However, only minor amounts of (±)-dehydrodiconiferyl alcohol appeared as a clearly distinguishable product (Figure S7A). In the light of these results, $CuSO_4$ could be addressed to partake in the reaction independently from CueO. However, this activity doesn't seem crucial *in vivo*, since 5 and 10 mM $CuSO_4$ applied to *E. coli* cells led to >99% substrate conversion (Figures 3A, 4 A, and S7B). Moreover, as already described before, *E. coli* possesses several homeostatis systems capable of interacting with free copper ions, which – together with specific and unspecific chelators – will immediately diminish the amount of available copper that could trigger radical formation and consequent coupling.[11a,36] In order to further validate this hypothesis, the *E. coli* BL21(DE3) strain was compared to an *E. coli* BL21(DE3) strain with deleted *cueO* gene (Δ*cueO*). While 99% conversion of 2 mM coniferyl alcohol was reached using *E. coli* BL21(DE3) resting cells after $CuSO_4$ addition, the *E. coli* BL21 (DE3) Δ*cueO* strain reached only 17% conversion, endorsing the role of CueO (Figure 4A). Low conversion observed with the Δ*cueO* strain is not surprising: apart from the above described non-enzymatic reaction it is also likely that other enzymes may take over the activity of CueO within the cell rather than free copper.[36] For instance, endogenous *E. coli* peroxidases, namely Prx01 and Prx02, have been described capable to catalyze the oxidative coupling of coniferyl alcohol to (±)-pinoresinol.[37] As copper cytotoxicity correlates with oxidative stress, the presence of H_2O_2 necessary for peroxidase activity seems plausible.

We also compared common copper sources like $CuSO_4$ and $CuCl_2$ to exclude other ions to influence the desired activity and confirm the role of copper ions as leading actor. As a matter of fact, variable responses of laccase activities have been associated to different ions. Specifically, sulfate salts have been reported in some cases to enhance laccase activity, whereas Cl^- ions often show an inhibiting effect.[38] The addition of 5 mM

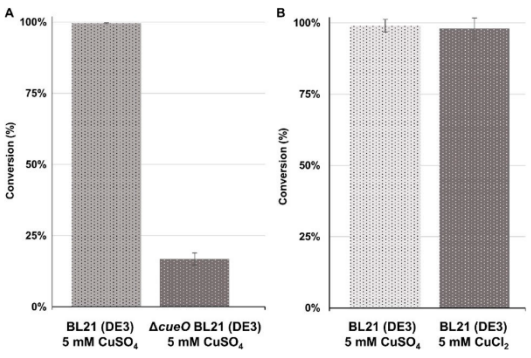


Figure 4. Conversion of coniferyl alcohol. A) Comparison of conversions catalyzed by *E. coli* BL21(DE3) and by *E. coli* BL21(DE3) Δ*cueO*; 5 mM $CuSO_4$ were added to growing cells, and conversion was performed with resting cells. B) Effect of different copper salts on coniferyl alcohol conversion by *E. coli* BL21(DE3); 5 mM $CuSO_4$ or $CuCl_2$ were added to resting cells together with 2 mM coniferyl alcohol.

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

$CuSO_4$ or $CuCl_2$ resulted in >99% conversion of 2 mM coniferyl alcohol in both cases, thus suggesting no influence of sulfate or chloride on CueO activity (Figure 4B).

Timing of copper sulfate addition: Aiming to prove copper addition to switch on CueO activity at need, $CuSO_4$ supplementation was tested at different time points of strain cultivation. Generally, $CuSO_4$ was added to growing cells at ~0.6 OD_{600}, whereas substrate conversion was performed with resting cells after cell harvesting. As an alternative, we tested substrate conversion with growing cells by adding coniferyl alcohol and $CuSO_4$ together at ~0.6 OD_{600} during cell growth. The growing cell approach led only to very low substrate conversion, and product detection was possible by mass spectrometry in single ion monitoring only (Figure S8). Prolongation of cell growth up to 72 h and lowering the incubation temperature from 30 to 25 °C did not result in any improvement, which demonstrates that resting cells are superior for the desired purpose.

In a third approach, cells were grown, harvested, and copper was added to resting cells together with the substrate. Especially with this setup, we aimed to prove that $CuSO_4$ acts indeed as the trigger inducing activity of CueO. In this case, the mechanism of copper ions loading in CueO is probably different compared to the growing cells. Nevertheless, utilizing this setup, efficient coniferyl alcohol depletion (~99%) was observed and (±)-pinoresinol concentrations were similar to those achieved when copper was added during cell growth (244 ± 53 μM).

A possible explanation for the diverse results achieved with growing and resting cells may be the interplay between oxygen availability and copper content leading to different cellular environments that influence catalytic activity or expression. Previous studies have shown that during growth phase oxygen consumption by *E. coli* cells is sensibly higher than with resting cells,[39] therefore CueO activity in growing cells might be limited because oxygen is consumed for other metabolic needs. Moreover, the heterologous overexpression of CueO in *E. coli* and its oxidizing activity *in vitro* were shown to correlate with cell growth under aerobic or micro-aerobic conditions.[12] According to that study, higher levels of CueO were achieved under fully aerobic conditions although CueO isolated from micro-aerobic cultures showed higher oxidizing activity.

Complementarily, the different outcome of conversions with growing or resting cells might be due to other physiological parameters like the cell concentrations applied for biotransformation.[40] To assess this, we performed whole-cell biotransformations of 2 mM coniferyl alcohol with resting cells normalized to cell wet weights of 70, 50, 30 and 20 g/L. These cell density values were chosen because after biotransformation with growing cells, the determined cell wet weights were usually ranging from 20–30 g/L.

After 4 h, 43 and 55% conversion was observed at cell concentrations of 70 and 50 g/L, respectively. With 30 and 20 g/L, substrate depletion achieved 36 and 17% (Figure S9). However, coniferyl aldehyde produced from coniferyl alcohol by endogenous *E. coli* dehydrogenase was detected as well, and its concentration was higher when lower cell concentrations were used. In particular, 14% coniferyl aldehyde was observed with 70 g/L, and ~20% with 50, 30 and 20 g/L, whereas 8–10% (±)-pinoresinol was detected for every tested condition. After 21 h, ~98% conversion of coniferyl alcohol was achieved regardless of the amount of cells (Figure S10). Despite the high amounts of coniferyl aldehyde spotted at 4 h, after 21 h coniferyl aldehyde is barely detected, and is probably reduced again to coniferyl alcohol, which is then dimerized. Given that different cell densities did not result in higher conversion values after 21 h, these results suggest that in resting cells a certain amount of "ready-to-go" CueO is present to efficiently convert coniferyl alcohol, whereas in the set up with growing cells, the continuous basal expression of CueO during cell growth may not be sufficient to convert the same concentration of coniferyl alcohol under the tested conditions, even with moderate overexpression as reported previously.[11b] However, metabolic engineering tools such as the usage of stronger promoters could provide useful tools for the improvement of the growing cell setup.[41] Furthermore, cell permeability could be another reason leading to the different behaviors of resting and growing cells. As already proven within the previous experiments, the integrity of the cell membrane might limit substrate diffusion and, as a consequence, the conversion.[30,42]

Application of CueO triggered activity to a cascade reaction

In order to demonstrate its usefulness, the CueO triggered activity was applied to a four-step cascade. The resting whole-cell approach was used to convert coniferyl alcohol to (−)-matairesinol. Within this cascade two additional enzymes are involved: Pinoresinol-lariciresinol reductase from *F. intermedia* (FiPLR) which catalyzes the two-step reduction of (+)-pinoresinol to (−)-secoisolariciresinol via the intermediate (+)-lariciresinol, and secoisolariciresinol dehydrogenase from *P. pleianthum* (PsSDH), which catalyzes the oxidation of (−)-secoisolariciresinol to (−)-matairesinol.

Three different setups were tested with 5 mM coniferyl alcohol. First, a "one-cell one-pot" setup was applied (Figure 5A): CueO activity was triggered by adding 5 mM $CuSO_4$ to *E. coli* C41(DE3) co-expressing FiPLR and PsSDH. After 20 h conversion, coniferyl alcohol was converted to ~89%, however, a high amount of coniferyl aldehyde was detected (~64% of the total products), while the desired compounds were poorly detectable (Figure S11).

Additionally, two mixed "two-cells one-pot" approaches were tested. *E. coli* BL21(DE3) cells with 5 mM $CuSO_4$ added during the growth phase was appointed as "*module one*", whereas *E. coli* C41(DE3) co-expressing FiPLR and PpSDH was set as "*module two*". In the first trial, the cascade was realized by mixing *modules one* and *two* at the same time immediately before substrate addition (Figure 5B). After 20 h reaction time, ~78% substrate depletion was achieved and again coniferyl aldehyde appeared as a distinct signal among the products accounting to 22%. Nevertheless, higher signals for the products of coniferyl alcohol coupling were detected, and the desired final product (−)-matairesinol (*m/z* 359 $[M+H]^+$, 341

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

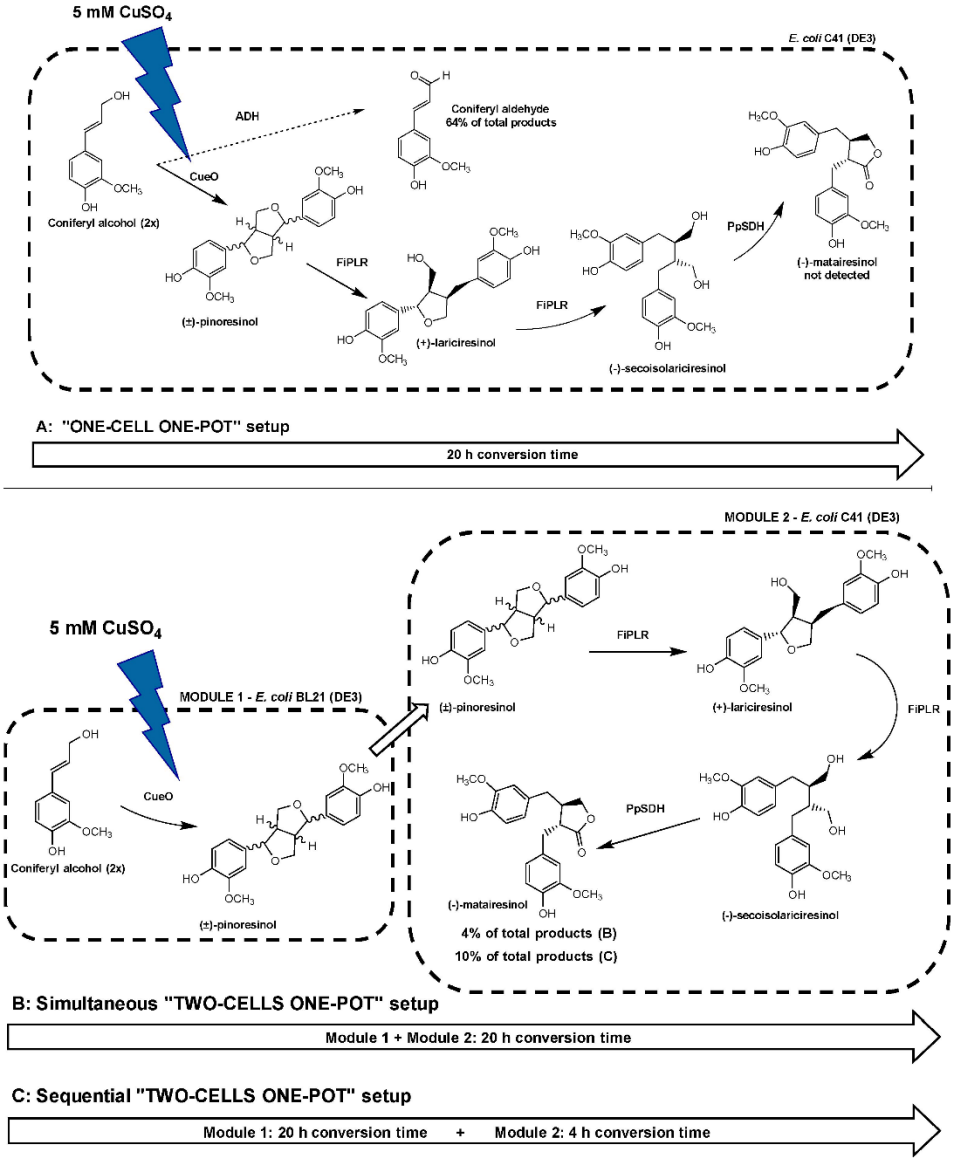


Figure 5. Overview of the three setups used for multistep transformation of coniferyl alcohol to (−)-matairesinol. A) "One-cell one pot" setup: *E. coli* C41(DE3) cells harboring FiPLR and PpSDH supplemented with 5 mM $CuSO_4$; side product coniferyl aldehyde: ~64%, no desired product (−)-matairesinol detected; B) Simultaneous "two-cells one-pot" setup: *E: coli* BL21(DE3) cells supplemented with 5 mM $CuSO_4$ were mixed with *E: coli* C41(DE3) cells harboring FiPLR and PpSDH prior to the start of the reaction; side product coniferyl aldehyde: 22%, desired product (−)-matairesinol: 4% of the total products. C) Sequential "two-cells one-pot" setup: *E: coli* BL21(DE3) cells supplemented with 5 mM $CuSO_4$ were incubated for 20 h alone to convert coniferyl alcohol to (±)-pinoresinol, afterwards *E. coli* C41(DE3) cells harboring FiPLR and PpSDH were added, and the incubation time was prolonged for 4 h. With this setup, production of coniferyl aldehyde was reduced to 2%, whereas (−)-matairesinol accounted for 10% of the total products, corresponding to the theoretically expected value for this cascade.

$[M+H-H_2O]^-$) accounted to 4% of the total products (Figure S12).

In the sequential "two-cells one-pot" approach, the conversion of 5 mM coniferyl alcohol with *module one* run for 20 h; afterwards *module two* was added and conversion was then prolonged for 4 h (Figure 5C). Overall, this approach appeared to be most promising, as coniferyl alcohol was efficiently converted within the first 20 h (97%) with only marginal

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

production of coniferyl aldehyde (2%) and the distribution of the coupling products according to those of the previous experiments, with (±)-pinoresinol representing 20% (Figure S12). After addition of *module two*, approximately 50% of (±)-pinoresinol were consumed, indicating that only one enantiomer was converted by FiPLR, as expected based on a previous report.[13] A few residues of the intermediate lariciresinol (*m/z* 383 $[M+Na]^+$) were detected (~2%), but no (−)-secoisolariciresinol fragments (*m/z* 363 $[M+H]^+$, 345 $[M+H-H_2O]^+$). The final product (−)-matairesinol corresponded to ~10%. As discussed above starting with 5 mM substrate concentration, the phenolic coupling occurring on coniferyl alcohol will result in 2.5 mM concentration of the products, approximately 500 μM of which is racemic (±)-pinoresinol. This means that the observed 10% (−)-matairesinol would correspond to ~250 μM. Indeed, the final product was quantified via internal standard calibration, resulting in 247±11 μM (−)-matairesinol, which corresponds well to the theoretically expected value. Altogether, the sequential "two-cells one-pot" approach seems to be the most suitable setup for further investigation.

Overall, the balance between coniferyl alcohol and coniferyl aldehyde obviously influences *E. coli* cell performance in various setups. Indeed, the more aldehyde was observed the less coniferyl alcohol coupling products were detected. Endogenous *E. coli* dehydrogenases which might be responsible for these oxidation/reduction reactions use nicotinamide species as cofactors,[43] meaning that their function is influenced by the redox and energetic state of the cell. On the other hand, the ratio between coniferyl alcohol and coniferyl aldehyde was affected by the presence of heterologous FiPLR and PpSDH, which are known to use NADPH and NADH as cofactors, respectively.[44,45] It means that both enzymes exploit cell metabolism for their function in parallel with endogenous dehydrogenases, altering the energetic state and probably competing for the same cofactors. This might explain a high ratio of coniferyl aldehyde observed in the "one-cell one-pot" setup. When endogenous ADHs were separated from the heterologous FiPLR and PpSDH in the "two-cells one-pot" setups, the balance was shifted towards coniferyl alcohol, which was converted to pinoresinol in the *module one*. In the *module two* pinoresinol was converted in three steps to (−)-matairesinol (Figure 5B and C). The lowest coniferyl aldehyde concentration and the highest (−)-matairesinol concentration could be achieved when the two *modules* were separated not only spatially but also added at different time points. This observation supports the above suggestion that in the *E. coli* cells expressing FiPLR and PpSDH coniferyl alcohol is easier oxidized to coniferyl aldehyde and thus remains inaccessible for the cascade.

Conclusion

In conclusion, the *in vivo* activity of the *E. coli* endogenous multi-copper oxidase CueO was triggered by the addition of copper salts. Treated *E. coli* cells were proven effective to oxidize several phenolic compounds leading to the corresponding dimers. In these experiments it was possible to switch on CueO activity arbitrarily at need, with the further aim to include host endogenous activity into a whole-cell biocatalyst featuring a four-step cascade to produce (−)-matairesinol. Starting with 5 mM coniferyl alcohol 247 μM final product was formed, which corresponds to the theoretically expected value.

Notwithstanding its limitations, this approach offers a simplification in the landscape of the whole-cell biocatalysis for multi-enzyme cascades and is generally an interesting approach to develop powerful microbial cell factories.

Experimental Section

Bacterial strains, Cultivation, and Expression

E. coli BL21(DE3) cells were purchased from Novagen (Merck), whereas *E. coli* BL21(DE3) with *cueO* gene deletion (Δ*cueO*) were generated via TargetTron mutagenesis kit (Sigma–Aldrich) as described previously[6b] by introducing a kanamycin resistance gene. As *cueO* is an endogenous *E. coli* gene, cells were transformed with pET16a or pET24b vectors (Merck) with no foreign genes by heat shock procedure to ensure antibiotic resistance as selection marker preventing contamination. Heterologous expression of *Corynebacterium glutamicum* laccase CgL1, cloned in pET16b was performed as described previously.[13] More details are provided in Tables S4 and S5.

Pre-cultures were prepared in 5 mL LB medium supplemented with the appropriate antibiotic and inoculated with one colony from LB-agar plates or from a preserved cryo-stock, always in biological duplicate at least. Cells were then grown overnight (O/N) at 37 °C, 180 rpm. From these pre-cultures, 500 μL were inoculated in 50 mL fresh Terrific broth medium (TB) supplemented with an appropriate antibiotic and cells grew at 37 °C, 180 rpm until an OD_{600} value of ~0.6. The heterologous expression of foreign genes was induced by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the cultures. Cells were then incubated at 30 °C for 21–22 h, 140 rpm (Multitron, Infors HT, Switzerland). For the proof-of-concept biotransformation of coniferyl alcohol to (−)-matairesinol, *E. coli* C41 (DE3) OverExpress (Lucigen) cells were co-transformed with the vector pCDFDuet-1, harboring pinoresinol-lariciresinol reductase from *F. intermedia* (GenBank AAC49608; FiPLR) and secoisolariciresinol dehydrogenase from *P. pleianthum/Dysosma pleiantha* (GenBank AHB18702; PpSDH). Protein expression was induced with 0.5 mM IPTG and carried out at 25 °C for 48 h, 120 rpm (Multitron, Infors HT, Switzerland).

Harvest, Normalization of Cell Density and Whole-Cell Biotransformation Setup

After cell growth and gene expression, cultures were harvested via centrifugation for 30 min, 4 °C, 3220*g* and the resulting pellets either directly used or stored at −20 °C. In the latter case, prior to further experiments, cells were thawed and their density normalized to 70 g/L (unless stated otherwise) using 50 mM phosphate buffer, KPi (80% K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5) supplemented with 500 mM D-glucose and 0.1 mM IPTG. Whole-cell biotransformation were performed in 500 μL reaction volume composed the normalized cell suspension and substrate solution in 2% (*v/v*) DMSO. For screening purposes, 200 μM substrate concentration was used and samples incubated in 1.5 mL Eppendorf tubes with open lids at 25 °C, 1500 rpm for 24 h unless stated otherwise.

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

Substrates tested with *E. coli* with induced CueO are listed in Table S6. For further investigation and product quantification, substrate concentration was increased to 2 mM and conversion after 2, 4, 16 and 24 h were analyzed. Conversion of coniferyl alcohol by CueO was investigated using both growing and resting cells. Concerning conversions with growing cells, the setup was the same as the expression experiments; once the cultures reached an OD_{600} of ~0.6, 5 mM $CuSO_4$ and 2 mM coniferyl alcohol–2% (*v*/*v*) DMSO- were added. Incubation was performed at 25 or 30 °C, for 21 to 72 h at 140 rpm (Multitron, Infors HT, Switzerland).

Three setups were tested for the proof-of-concept biotransformation of coniferyl alcohol to (−)-matairesinol using the resting cells approach. In every case, cells were resuspended in 50 mM phosphate buffer KPi (80% K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5) supplemented with 500 mM D-glucose and 0.1 mM IPTG with cell wet weight normalized to 70 g/L. A "one-cell one pot" was prepared consisting of 10 mL of *E. coli* C41(DE3) cell suspension co-expressing FiPLR and PpSDH and 5 mM coniferyl alcohol added together with 5 mM $CuSO_4$ to trigger CueO activity. Conversions were carried out in 100 mL Erlenmeyer flasks at 25 °C, 200 rpm for 20 h.

Alternatively, a mixed "two-cells one pot" was applied. Concerning the simultaneous setup, 5 mL of *E. coli* BL21(DE3) supplemented with 5 mM $CuSO_4$ during growth phase were mixed with 5 mL *E. coli* C41(DE3) co-expressing FiPLR and PpSDH. 5 mM coniferyl alcohol was added and the conversions was carried out in 100 mL Erlenmeyer flasks at 25 °C, 200 rpm for 20 h.

In a sequential setup, the first step of the reaction was performed with 10 mL *E. coli* BL21(DE3) supplemented with 5 mM $CuSO_4$ during growth phase (*module one*), with 5 mM coniferyl in 100 mL Erlenmeyer flasks at 25 °C, 200 rpm for 20 h. Subsequently, 10 mL of *E. coli* C41(DE3) with co-expressed FiPLR and PpSDH were added as *module two* and the reaction run for 4 h.

Copper Addition

The addition of $CuSO_4$ as a trigger for enzymatic activity was performed during cell growth phase, or in case of resting cells after harvest but prior substrate addition. $CuSO_4$ concentrations of 1, 3, 5, and 10 mM were tested. The same $CuSO_4$ concentrations were used to test copper mediated coniferyl alcohol oxidation and subsequent radical coupling. Reactions were performed in 50 mM phosphate buffer KPi (80% K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5), 500 μL reaction volume with 2 mM substrate in 2% DMSO, *v*/*v*). Experiments were performed in technical triplicate. Samples were incubated at 25 °C, 1500 rpm for 24 h.

Metabolite Extraction and Analysis

Prior to the extraction, 200 μM of internal standard ferulic acid or sesamin were added to the samples appointed for quantitative analysis. Metabolites were extracted twice using 1 mL of ethyl acetate, the resulting organic phase was then evaporated under reduced pressure. After evaporation, samples were resuspended in methanol (MeOH, 99.9% LC/MS grade, Fischer Scientific) for LC/MS analysis.

Both qualitative and quantitative analysis were performed by liquid chromatography coupled with mass spectrometry (LC/MS) measurements on LCMS-2020 system (Shimadzu, Tokyo, Japan) equipped with a Chromolith® Performance RP-18e column (100× 4.6 mm, Merck). More details are provided in Tables S7 and S8). Samples appointed for quantitative analysis were made in technical and biological duplicate at least.

Copper Toxicity and Cell Viability Tests

Increasing concentrations of $CuSO_4$ were added to investigate copper toxicity on cells growth and viability. 400 mL of fresh TB medium were inoculated with 4 mL of overnight culture of *E. coli* BL21(DE3) and cells grew to an OD_{600} of 0.6 at 37 °C, 180 rpm (Multitron, Infors HT, Switzerland). At the OD_{600} of 0.6, cells culture was divided in 100 mL aliquots, $CuSO_4$ at concentrations of 3, 5, 10, 15, 20, and 30 mM was added, and cells were further incubated at 37 °C, 180 rpm. Growth was followed over 24 h by measuring OD_{600}. Viability of *E. coli* cells after copper addition was investigated by taking 40 μL of cell suspension straight after copper addition (t= 0 h) and/or after 24 h incubation time (t=24 h). Aliquots were diluted to get the same cell amount, corresponding to an OD_{600} value of 0.5, and spread on LB agar plates supplemented with 30 μg/mL kanamycin.

Acknowledgements

This work was funded by the Federal Ministry of Education and Research (Germany) to Heinrich-Heine University Düsseldorf [grant no. 031B0362A] as part of the "Nationale Forschungsstrategie BioÖkonomie 2030" project name "LignaSyn". Open access funding enabled and organized by Projekt DEAL.

Conflicts of Interests

The authors declare no conflicts of interests.

Keywords: biotransformation · CueO · enzymes · laccase · phenolic coupling

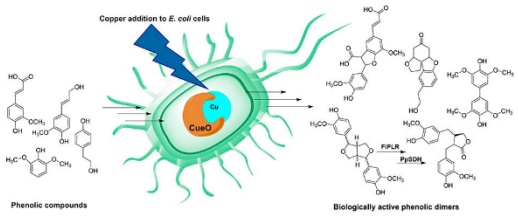
[1] G. Janusz, A. Pawlik, U. Swiderska-Burek, J. Polak, J. Sulej, A. Jarosz-Wilkolazka, A. Paszczynski, *Int. J. Mol. Sci.* **2020**, *21*, 1–25.
[2] a) R. Chandra, P. Chowdhary, *Environ. Sci. Process. Impacts* **2015**, *17*, 326–342; b) F. de Salas, P. Aza, J. F. Gilabert, G. Santiago, S. Kilic, M. E. Sener, J. Vind, V. Guallar, A. T. Martinez, S. Camarero, *Green Chem.* **2019**, *21*, 5374–5385; c) J. Yang, W. Li, T. B. Ng, X. Deng, J. Lin, X. Ye, *Front. Microbiol.* **2017**, *8*, 832.
[3] D. M. Mate, M. Alcalde, *Microb. Biotechnol.* **2017**, *10*, 1457–1467.
[4] Z. B. Guan, Q. Luo, H. R. Wang, Y. Chen, X. R. Liao, *Cell. Mol. Life Sci.* **2018**, *75*, 3569–3592.
[5] a) S. Brander, J. D. Mikkelsen, K. P. Kepp, *PLoS One* **2014**, *9*, e99402; b) M. Gunne, V. B. Urlacher, *PLoS One* **2012**, *7*, e52360; c) M. F. Hullo, I. Moszer, A. Danchin, I. Martin-Verstraete, *J. Bacteriol.* **2001**, *183*, 5426–5430; d) K. Koschorreck, S. M. Richter, A. B. Ene, E. Roduner, R. D. Schmid, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2008**, *79*, 217–224; e) M. C. Machczynski, E. Vijgenboom, B. Samyn, G. W. Canters, *Protein Sci.* **2004**, *13*, 2388–2397; f) R. Reiss, J. Ihssen, L. Thony-Meyer, *BMC Biotechnol.* **2011**, *11*, 9.
[6] a) P. Durao, Z. Chen, A. T. Fernandes, P. Hildebrandt, D. H. Murgida, S. Todorovic, M. M. Pereira, E. P. Melo, L. O. Martins, *J. Biol. Inorg. Chem.* **2008**, *13*, 183–193; b) M. Gunne, D. Al-Sultani, V. B. Urlacher, *J. Biotechnol.* **2013**, *168*, 252–255; c) Y. Li, W. Zuo, Y. Li, X. Wang, *Adv. Biol. Chem.* **2012**, *2*, 248–255; d) L. O. Martins, C. M. Soares, M. M. Pereira, M. Teixeira, T. Costa, G. H. Jones, A. O. Henriques, *J. Biol. Chem.* **2002**, *277*, 18849–18859.
[7] a) C. Kim, W. W. Lorenz, J. T. Hoopes, J. F. Dean, *J. Bacteriol.* **2001**, *183*, 4866–4875; b) G. Grass, C. Rensing, *Biochem. Biophys. Res. Commun.* **2001**, *286*, 902–908.
[8] P. Stolle, B. Hou, T. Bruser, *J. Biol. Chem.* **2016**, *291*, 13520–13528.
[9] Y. Ueki, M. Inoue, S. Kurose, K. Kataoka, T. Sakurai, *FEBS Lett.* **2006**, *580*, 4069–4072.

ChemBioChem
Full Papers
doi.org/10.1002/cbic.202000775
Chemistry Europe
European Chemical Societies Publishing

[10] K. Kataoka, H. Komori, Y. Ueki, Y. Konno, Y. Kamitaka, S. Kurose, S. Tsujimura, Y. Higuchi, K. Kano, D. Seo, T. Sakurai, *J. Mol. Biol.* **2007**, *373*, 141–152.
[11] a) M. Solioz in *Copper and Bacteria: Evolution, Homeostasis and Toxicity*, Springer, Cham, **2018**, pp. 49–80; b) F. W. Outten, D. L. Huffman, J. A. Hale, T. V. O'Halloran, *J. Biol. Chem.* **2001**, *276*, 30670–30677.
[12] T. Hiraishi, K. Tachibana, N. Asakura, H. Abe, M. Maeda, *Polym. Degrad. Stab.* **2019**, *164*, 1–8.
[13] E. Ricklefs, M. Girhard, V. B. Urlacher, *Microb. Cell Fact.* **2016**, *15*, 78.
[14] a) Z. Zhu, R. McKendry, C. L. Chavez in *Environmental Stressors and Gene Responses, Vol. 1* (Eds.: J. M. Storey, K. B. Store), Elsevier, Amsterdam, **2000**, pp. 293–300; b) C. L. Dupont, G. Grass, C. Rensing, *Metallomics* **2011**, *3*, 1109–1118.
[15] L. Macomber, J. A. Imlay, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8344–8349.
[16] a) T. Kimura, H. Nishioka, *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **1997**, *389*, 237–242; b) L. Macomber, C. Rensing, J. A. Imlay, *J. Bacteriol.* **2007**, *189*, 1616–1626.
[17] G. Grass, C. Rensing, *J. Bacteriol.* **2001**, *183*, 2145–2147.
[18] L. Zhang, H. Cui, G. V. Dhoke, Z. Zou, D. F. Sauer, M. D. Davari, U. Schwaneberg, *Chem. Eur. J.* **2020**, *26*, 4974–4979.
[19] T. Kudanga, B. Nemadziva, M. Le Roes-Hill, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 13–33.
[20] O. E. Adelakun, T. Kudanga, I. R. Green, M. Le Roes-Hill, S. G. Burton, *Process Biochem.* **2012**, *47*, 1926–1932.
[21] S. M. Bolicke, W. Ternes, *Meat Sci.* **2016**, *118*, 108–116.
[22] O. E. Adelakun, T. Kudanga, A. Parker, I. R. Green, M. Le Roes-Hill, S. G. Burton, *J. Mol. Catal. B* **2012**, *74*, 29–35.
[23] E. Ricklefs, M. Girhard, K. Koschorreck, M. S. Smit, V. B. Urlacher, *ChemCatChem* **2015**, *7*, 1857–1864.
[24] a) A. M. Kunjapur, Y. Tarasova, K. L. Prather, *J. Am. Chem. Soc.* **2014**, *136*, 11644–11654; b) C. A. Holland-Staley, K. Lee, D. P. Clark, P. R. Cunningham, *J. Bacteriol.* **2000**, *182*, 6049–6054.
[25] a) J. Lee, J. Choi, W. Lee, K. Ko, S. Kim, *Mol. Immunol.* **2015**, *68*, 434–444; b) W. Lee, K. R. Ko, H. K. Kim, S. Lim, S. Kim, *Biochem. Biophys. Res. Commun.* **2018**, *495*, 2242–2248.
[26] a) D. E. Sok, H. S. Cui, M. R. Kim, *Recent Pat. Food Nutr. Agric.* **2009**, *1*, 87–95; b) H. Zhou, J. Ren, Z. Li, *Food Control* **2017**, *79*, 192–199.
[27] S. Nicotra, M. R. Cramarossa, A. Mucci, U. M. Pagnoni, S. Riva, L. Forti, *Tetrahedron* **2004**, *60*, 595–600.
[28] R. B. Teponno, S. Kusari, M. Spiteller, *Nat. Prod. Rep.* **2016**, *33*, 1044–1092.
[29] J. R. Jeon, P. Baldrian, K. Murugesan, Y. S. Chang, *Microb. Biotechnol.* **2012**, *5*, 318–332.
[30] P. H. Calcott, R. A. MacLeod, *Can. J. Microbiol.* **1975**, *21*, 1724–1732.
[31] J. Zeng, Q. Zhu, Y. Wu, X. Lin, *Chemosphere* **2016**, *148*, 1–7.
[32] L. B. Davin, H. Wang, A. L. Crowell, D. L. Bedgar, D. M. Martin, S. Sarkanen, N. G. Lewis, *Science* **1997**, *275*, 362–367.
[33] S. C. Halls, L. B. Davin, D. M. Kramer, N. G. Lewis, *Biochemistry* **2004**, *43*, 2587–2595.
[34] B. Pickel, A. Schaller, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 8427–8438.
[35] S. E. Allen, R. R. Walvoord, R. Padilla-Salinas, M. C. Kozlowski, *Chem. Rev.* **2013**, *113*, 6234–6458.
[36] J. J. Tree, S. P. Kidd, M. P. Jennings, A. G. McEwan, *Biochem. Biophys. Res. Commun.* **2005**, *328*, 1205–1210.
[37] Y. Lv, X. Cheng, G. Du, J. Zhou, J. Chen, *Biotechnol. Bioeng.* **2017**, *114*, 2066–2074.
[38] a) A. M. Farnet, G. Gil, E. Ferre, *Chemosphere* **2008**, *70*, 895–900; b) K. P. Kepp, *Inorg. Chem.* **2015**, *54*, 476–483.
[39] T. E. Riedel, W. M. Berelson, K. H. Nealson, S. E. Finkel, *Appl. Environ. Microbiol.* **2013**, *79*, 4921–4931.
[40] Y. Yang, M.-Z. Liu, Y.-S. Cao, C.-K. Li, W. Wang, *Catalysts* **2019**, *9*, 970.
[41] D. Yang, S. Y. Park, Y. S. Park, H. Eun, S. Y. Lee, *Trends Biotechnol.* **2020**, *38*, 745–765.
[42] B. E. White, C. J. Fenner, M. S. Smit, S. T. L. Harrison, *Microb. Cell Fact.* **2017**, *16*, 156.
[43] S. Elleuche, G. Antranikian, *OA Alcohol* **2013**, *1*, 3.
[44] L. Markulin, C. Corbin, S. Renouard, S. Drouet, L. Gutierrez, I. Mateljak, D. Auguin, C. Hano, E. Fuss, E. Laine, *Planta* **2019**, *249*, 1695–1714.
[45] Z. Q. Xia, M. A. Costa, H. C. Pelissier, L. B. Davin, N. G. Lewis, *J. Biol. Chem.* **2001**, *276*, 12614–12623.

Manuscript received: November 16, 2020
Revised manuscript received: December 15, 2020
Accepted manuscript online: December 17, 2020
Version of record online: ■■, ■■■

FULL PAPERS



*D. Decembrino, Dr. M. Girhard, Prof. V. B. Urlacher**

Simple and efficient: The endogenous *E. coli* laccase CueO was activated by the addition of copper ions to cells, which were subsequently used for oxidative phenolic coupling. As a proof of concept, CueO-catalyzed dimerization of coniferyl alcohol was integrated into a multi-enzyme cascade in *E. coli* to produce the plant lignan (−)-matairesinol.

2.1.1 Supporting Information

ChemBioChem

Supporting Information

Use of Copper as a Trigger for the *in Vivo* Activity of *E. coli* Laccase CueO: A Simple Tool for Biosynthetic Purposes

Davide Decembrino, Marco Girhard, and Vlada B. Urlacher*

Supporting results

Copper toxicity: Microscopy

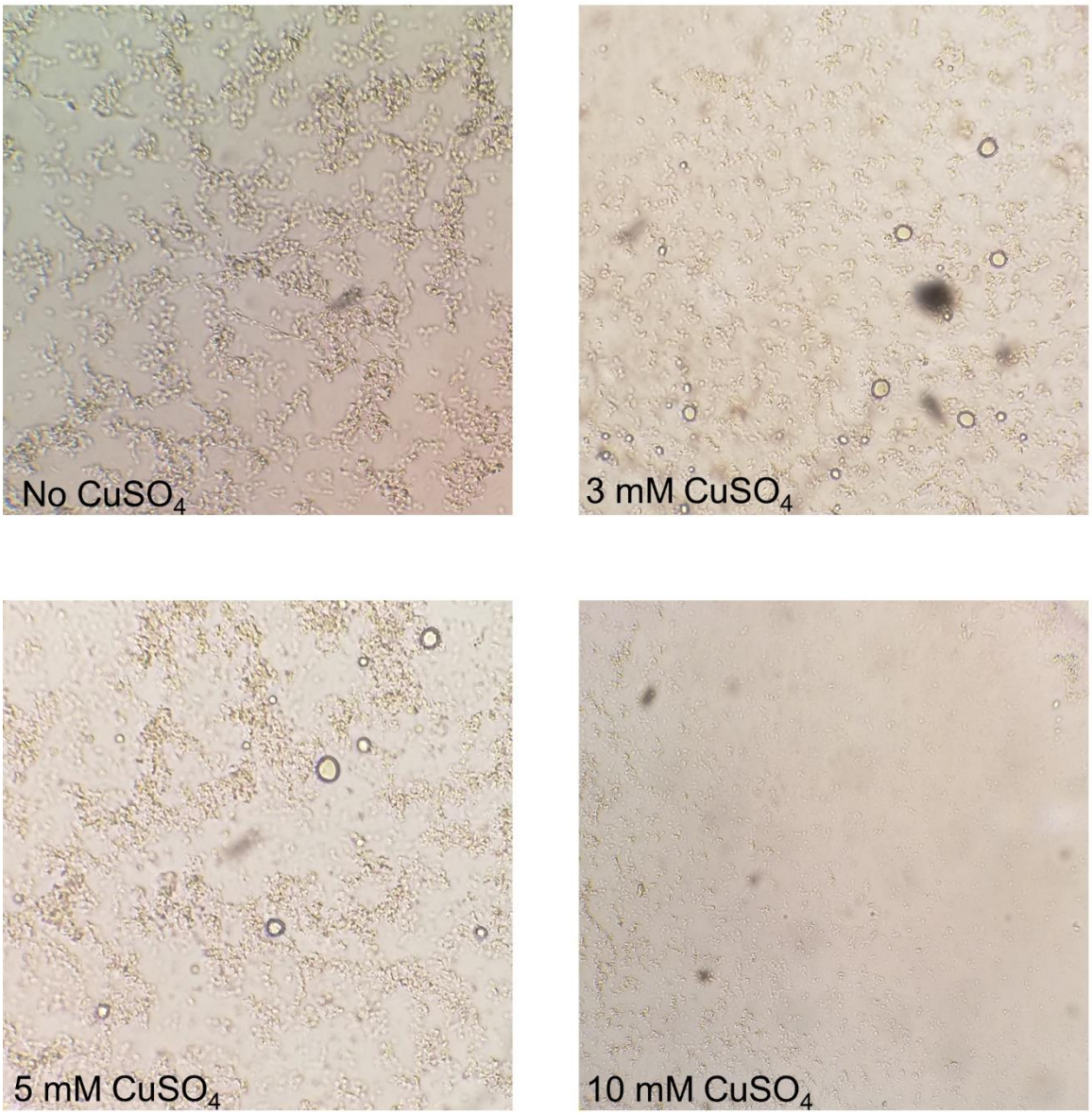


Figure S1. Effect of increasing $CuSO_4$ concentrations on *E. coli* cells. Optical microscopy 40x magnification

Substrate screening

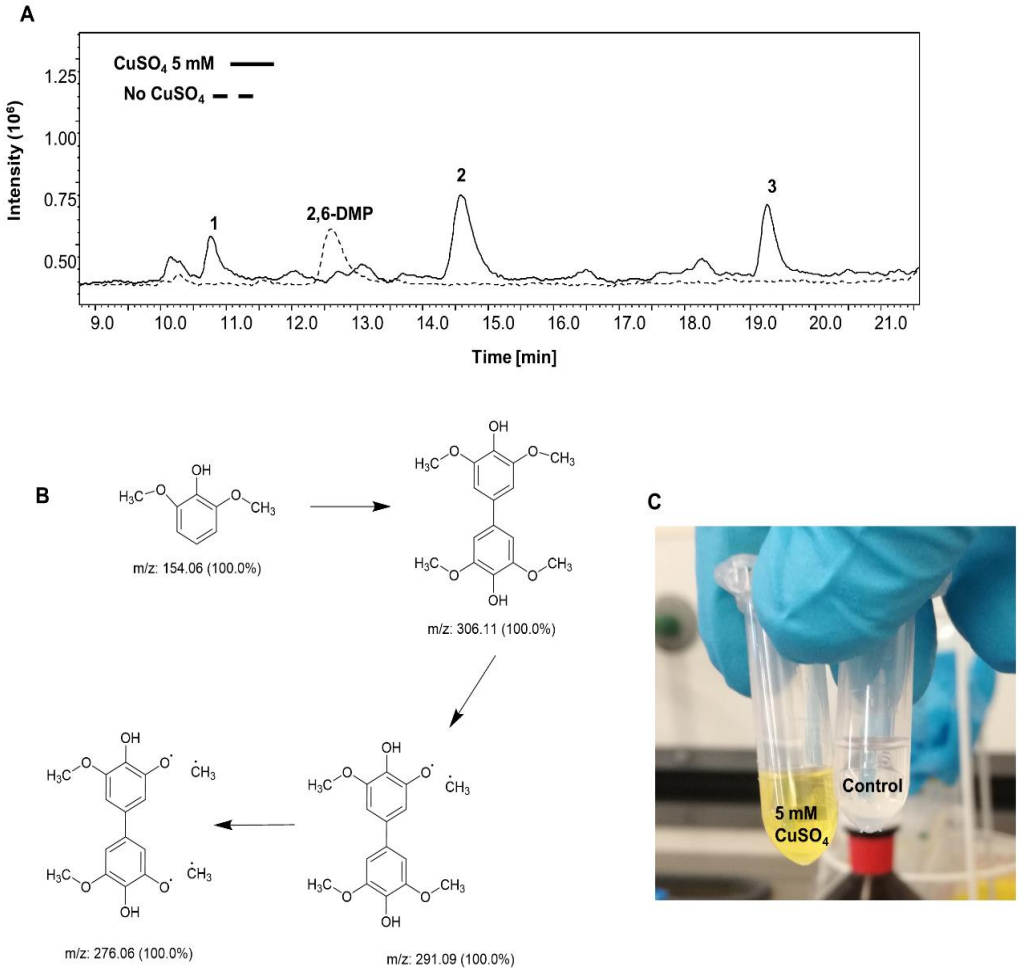


Figure S2. Oxidation of 2,6-DMP. A: LC/MS analysis. B: Proposed coupling and degradation products; m/z predicted by ChemDraw. C: Reaction sample and negative control after reaction. Products numbers have to be associated to the values listed in Table S1.

Table S1. Features of 2,6-DMP and coupling products

Compound	Theoretical m/z	Observed m/z fragments and proposed adducts	Retention time [min]
2,6-DMP	154	173 $[M+H+H_2O]^+$	12.6
Product 1	306	277 $[M+H-2CH_3]^+$	10.7
Product 2	306	292 $[M+H-CH_3]^+$	14.6
Product 3	306	307 $[M+H]^+$	19.4

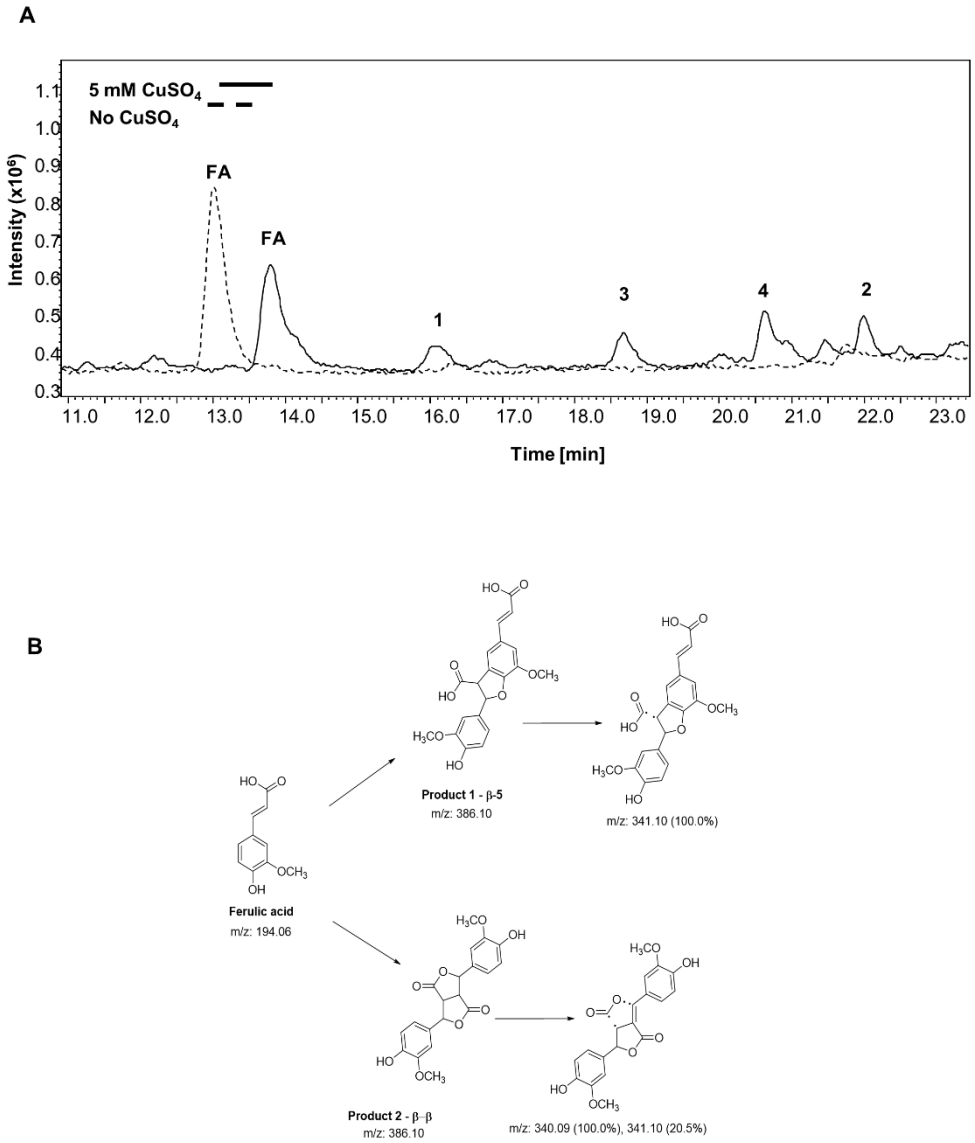


Figure S3. Oxidation of ferulic acid. A: LC/MS analysis. B: Proposed coupling and degradation products; m/z predicted by ChemDraw. Products numbers have to be associated to the values listed in Table S2.

Table S2. Features of ferulic acid and coupling products.

Compound	Theoretical m/z	Observed m/z fragments and proposed adducts	Retention time [min]
Ferulic acid	194	196 $[M+2H]^+$	13-14
Product 1	386	387 $[M+H]^+$	16.1
Product 2	386	387 $[M+H]^+$	22
Product 3-4	386	341 $[M+H\text{-}CHO_2^{\cdot}]^+$	18.6 – 20.6

Table S3. Features of coniferyl alcohol and its products

	Coniferyl alcohol	Coniferyl aldehyde	(±)-erythro/threo-guaiacylglycerol 8-O-4'-coniferyl alcohol ether	(±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ether	(±)-dehydrodiconiferyl alcohol	(±)-pinoresinol
MW	180.2	178.2	376.15	376.15	358.39	358,39
Retention Time (RT) [min]	**10.5**	**13**	**11.6**	**12.11**	**17.5**	**19.5**
Characteristic m/z fragments	164	179	341	341	341	341- 359

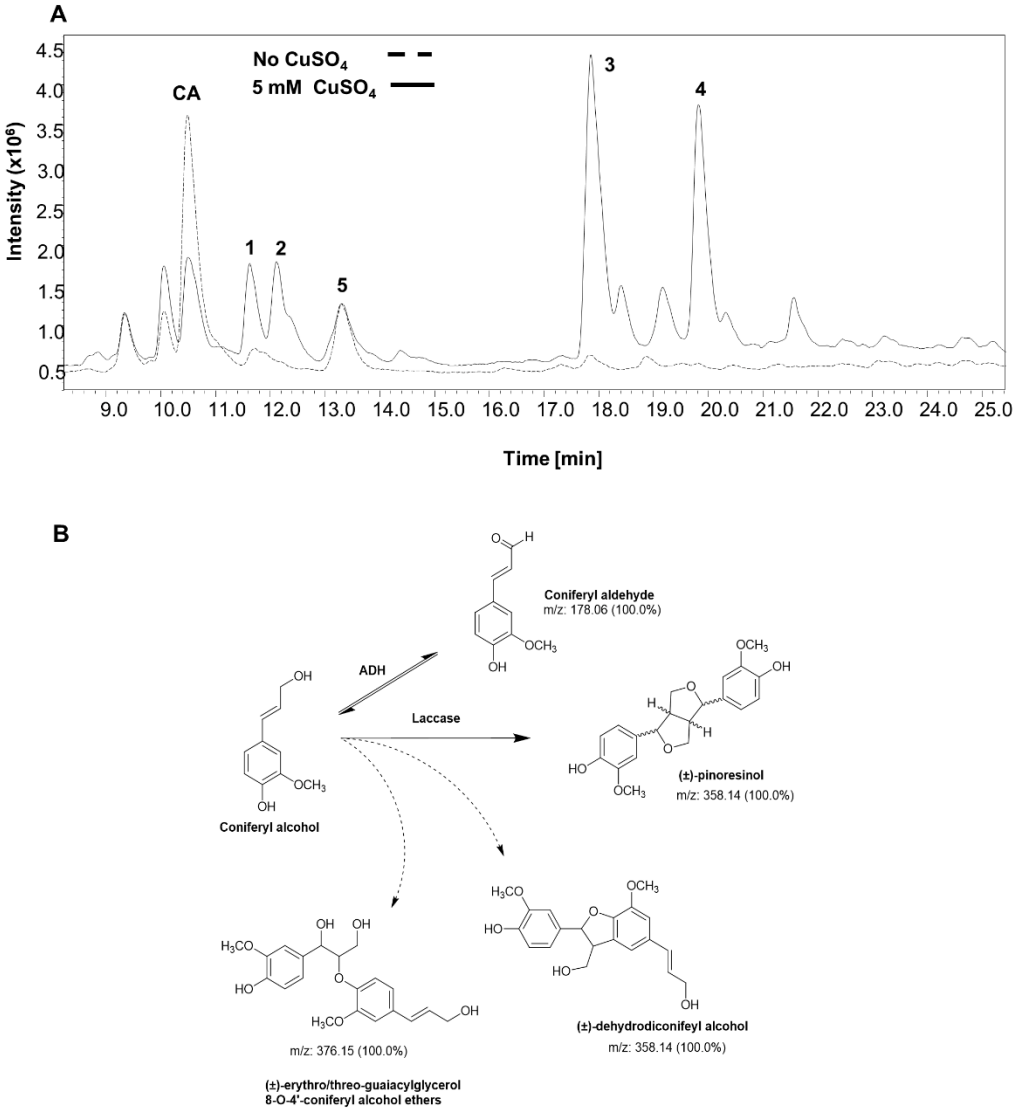


Figure S4. Oxidation of coniferyl alcohol. A: LC/MS analysis. B: Reaction products. Listed clockwise from coniferyl alcohol: coniferyl aldehyde, (±)-pinoresinol, (±)-dehydrodiconiferyl alcohol 2: (±)-erythro/threo-guaiacylglycerol 8-O-4'-coniferyl alcohol ethers

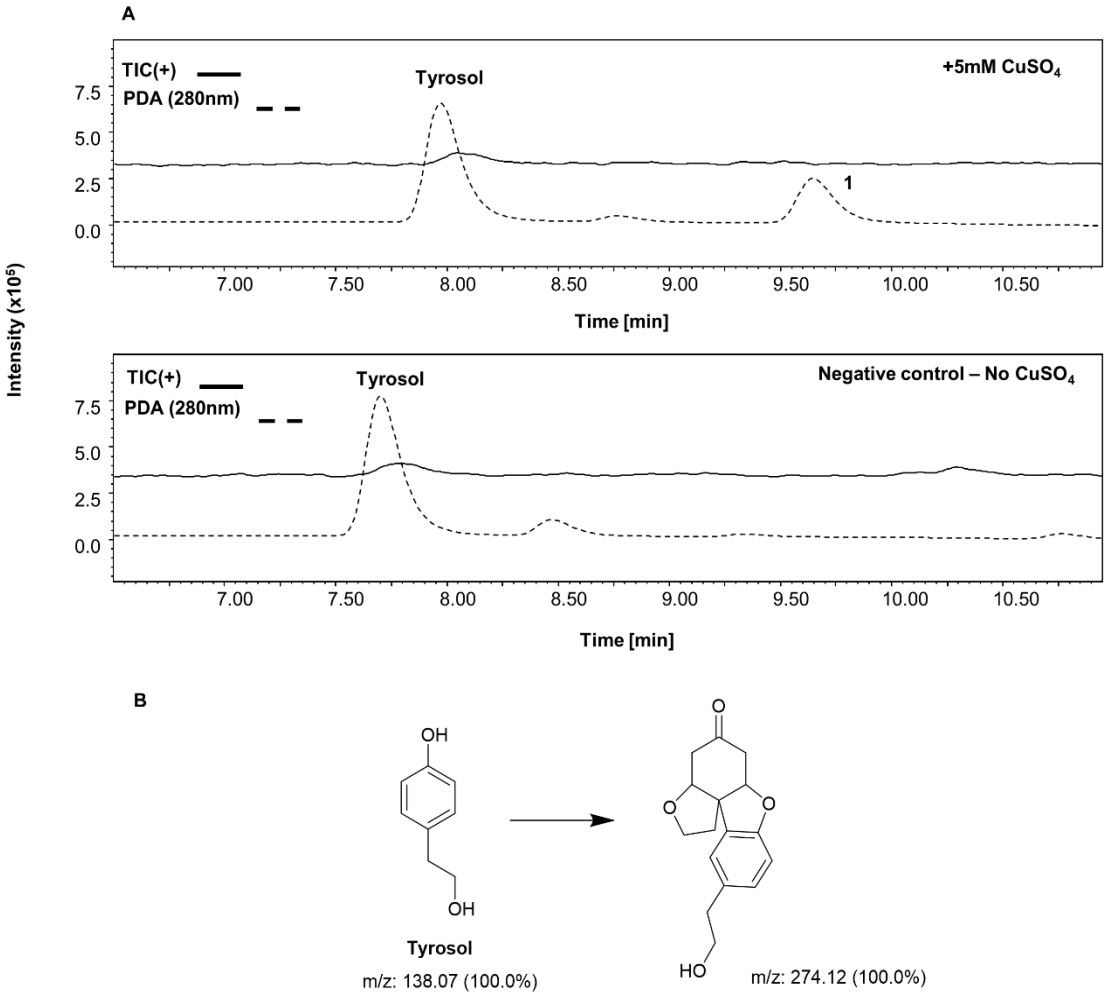


Figure S5. Reaction with tyrosol. A: LC/MS analysis, UV/Vis chromatogram 280 nm. B: Proposed coupling product; m/z predicted by ChemDraw.

Coniferyl alcohol coupling

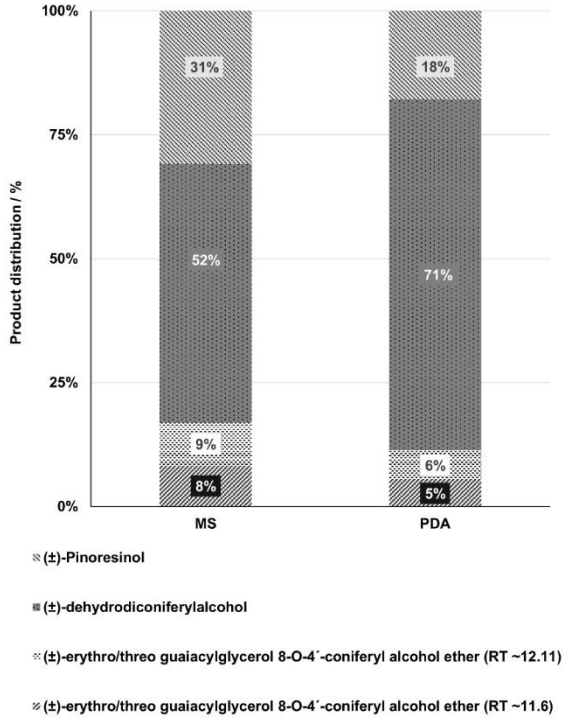


Figure S6. Distribution of coniferyl alcohol coupling products. Plotted values are calculated based on the areas detected via mass spectrometry (MS) or UV/Vis (PDA). Average values are given in the main manuscript.

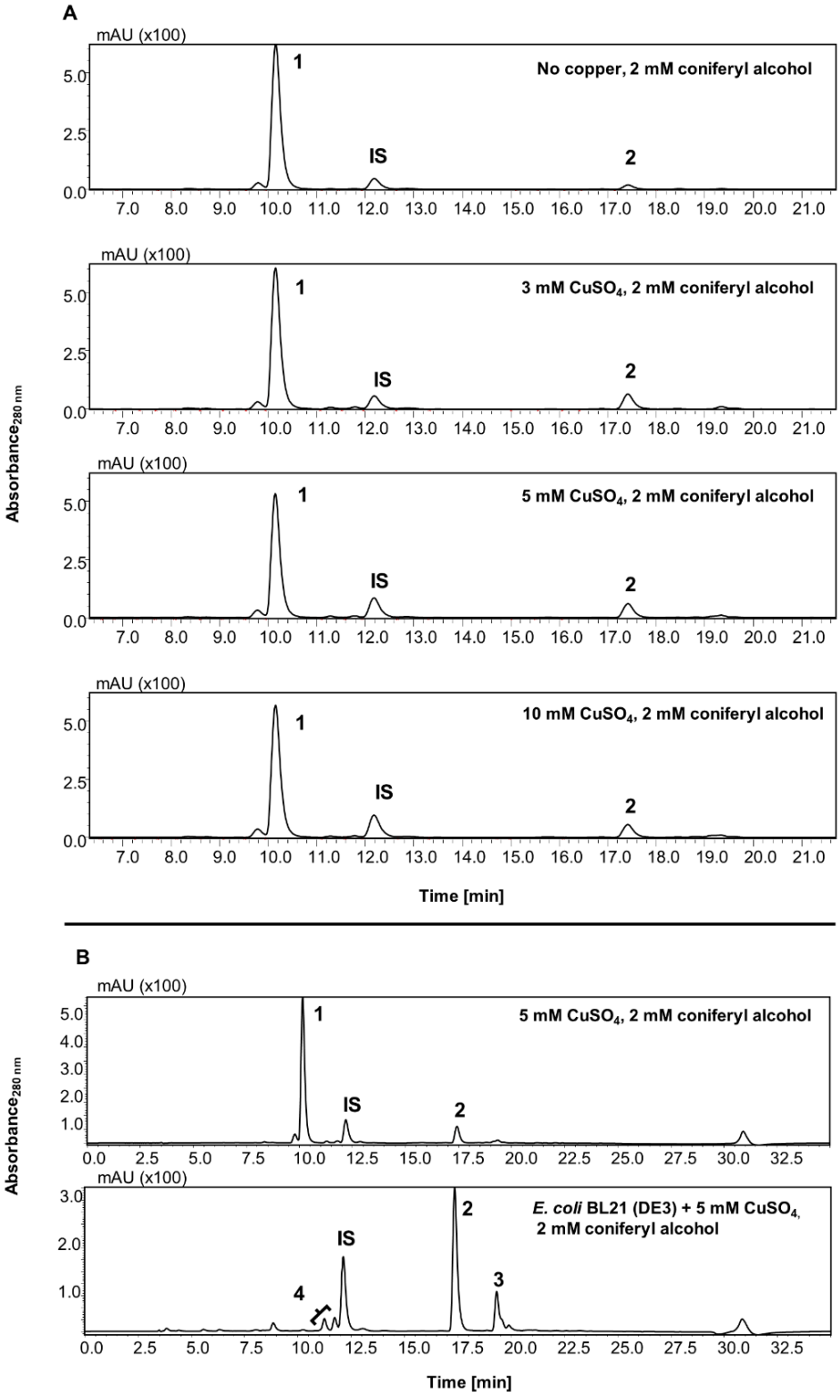


Figure S7. Effect of $CuSO_4$ on the conversion of 2 mM coniferyl alcohol. UV/Vis chromatograms at 280 nm are shown, peaks are identified as follows: 1: coniferyl alcohol; 2: (±)-dehydrodiconiferyl alcohol; 3: (±)-pinoresinol. 4: (±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ethers; IS: ferulic acid (internal standard). A: Effect of increasing $CuSO_4$ concentrations on coniferyl alcohol coupling in aqueous solution (phosphate buffer KPi, 50 mM pH 7.5). A minor fragment of product 2 appeared as detectable in the "No copper" control as well; B: Qualitative comparison of coniferyl alcohol coupling by CueO in E. coli cells supplemented with 5 mM $CuSO_4$,and 5 mM $CuSO_4$-mediated conversion in aqueous solution (phosphate buffer KPi, 50 mM pH 7.5).

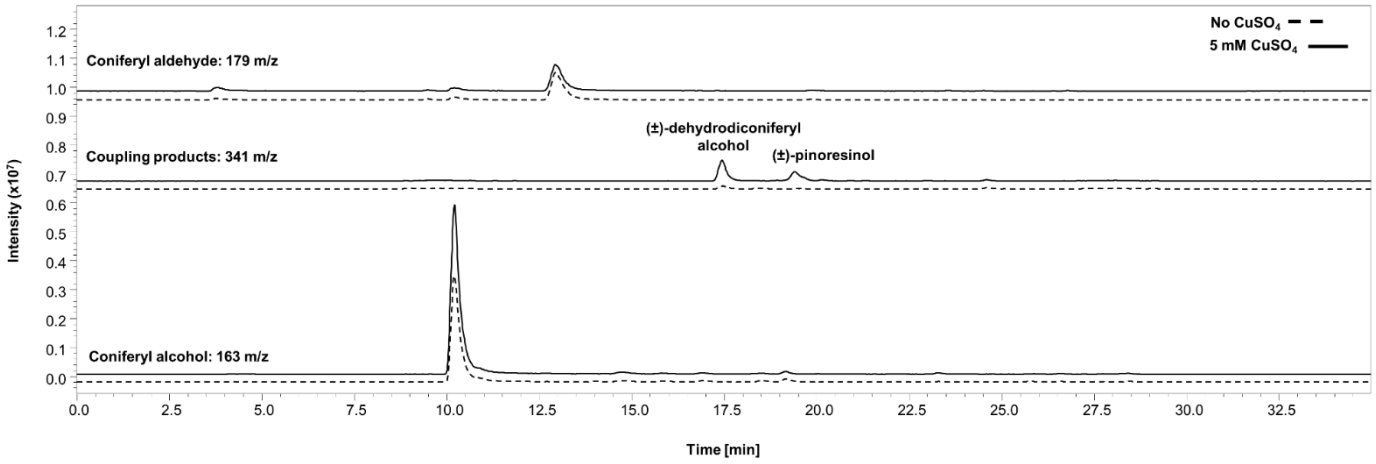


Figure S8. LC/MS analysis of coniferyl alcohol conversion with growing cells. Single ion monitoring was used to detect characteristic m/z fragments belonging to the desired product or to the substrate.

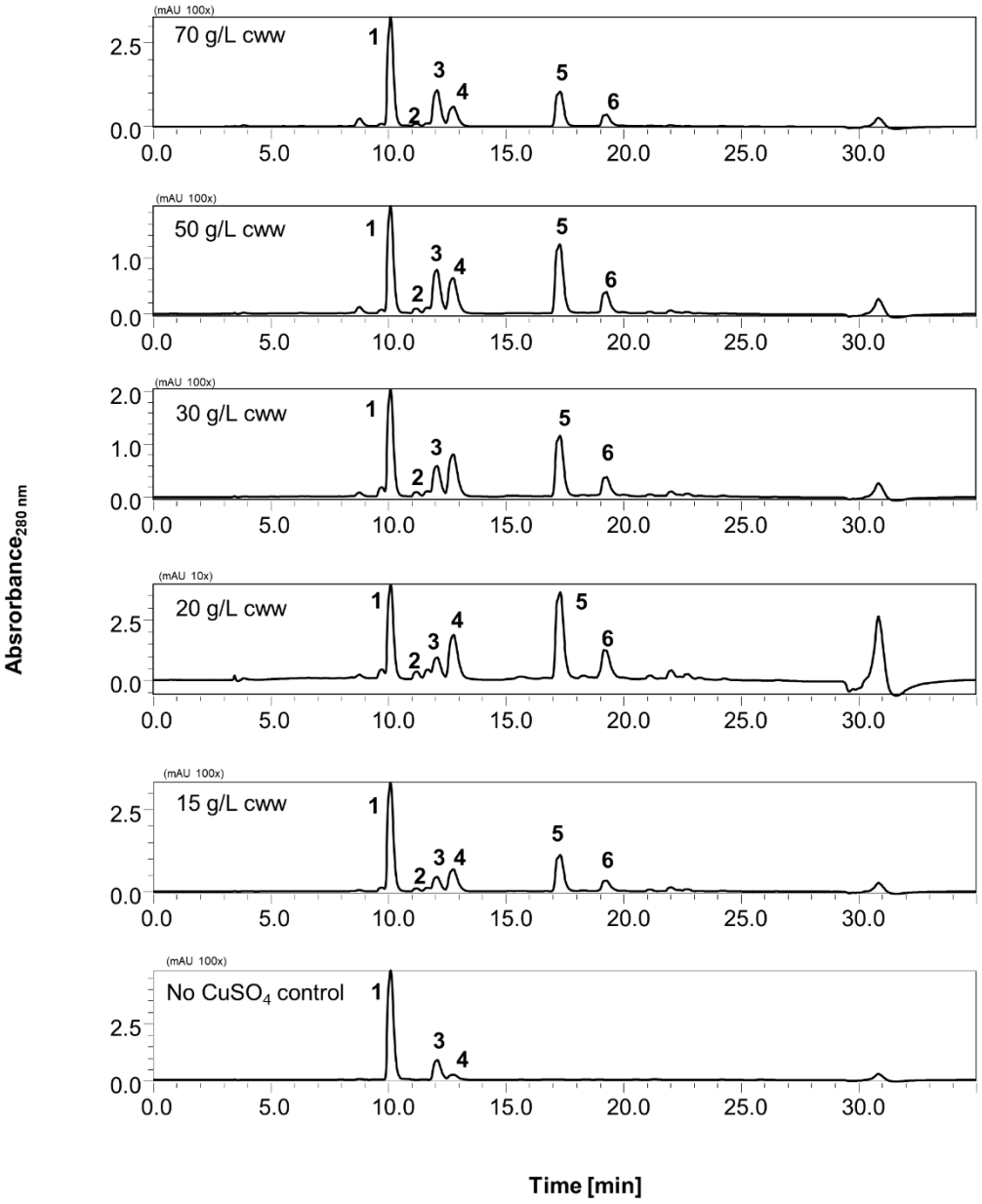


Figure S9. LC/MS analysis of coniferyl alcohol conversion with different cell densities after 4 h reaction time. UV/Vis chromatograms at 280 nm are shown. 1: coniferyl alcohol; 2: (±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ethers; 3: ferulic acid (internal standard); 4: coniferyl aldehyde; 5: (±)-dehydrodiconiferyl alcohol; 6: (±)-pinoresinol.

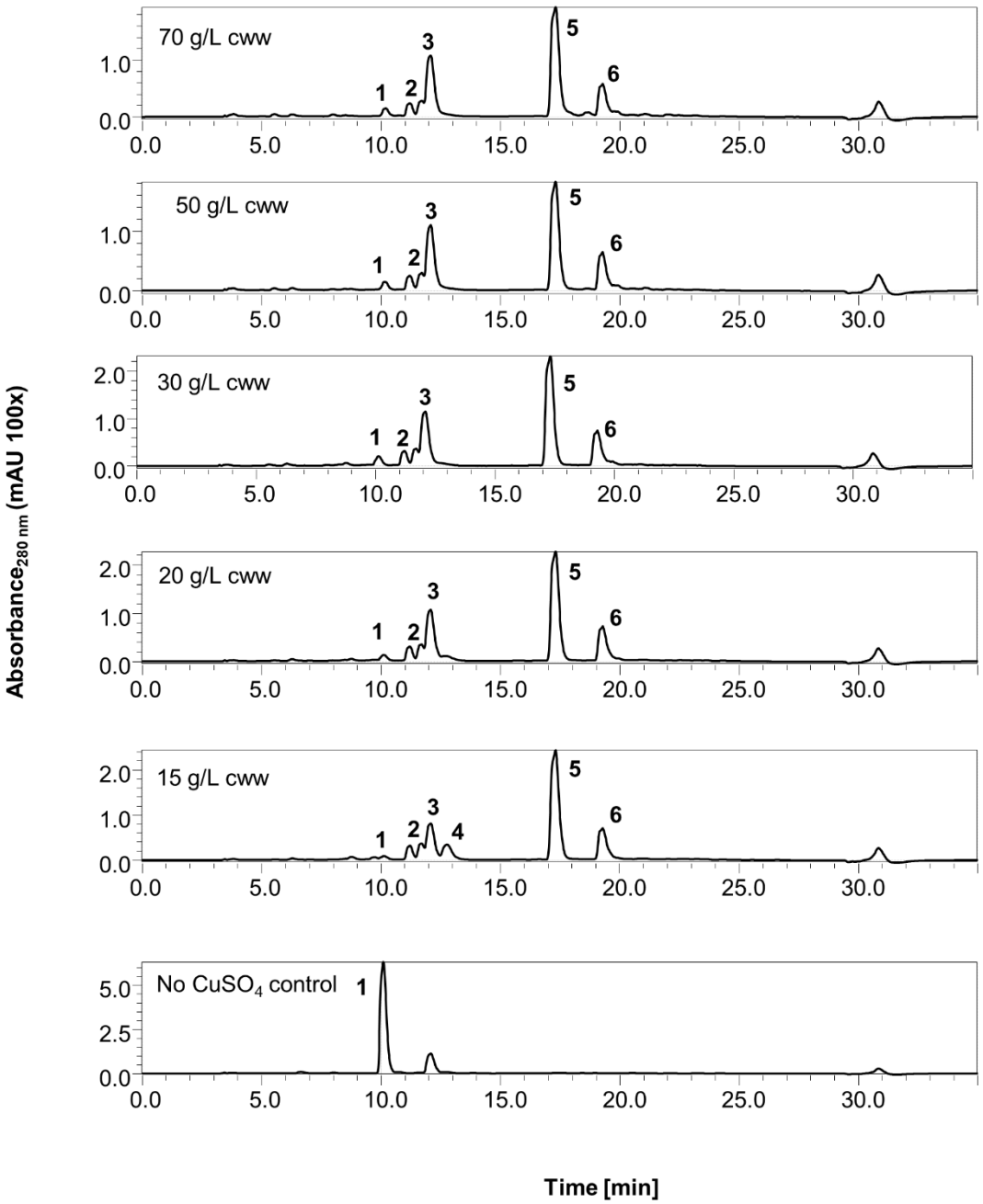


Figure S10. LC/MS analysis of coniferyl alcohol conversion with different cell densities after 21 h reaction time. UV/Vis chromatograms at 280 nm are shown. 1: coniferyl alcohol; 2: (±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ethers; 3: ferulic acid (internal standard); 4: coniferyl aldehyde; 5: (±)-dehydrodiconiferyl alcohol; 6: (±)-pinoresinol.

Copper driven cascade reactions

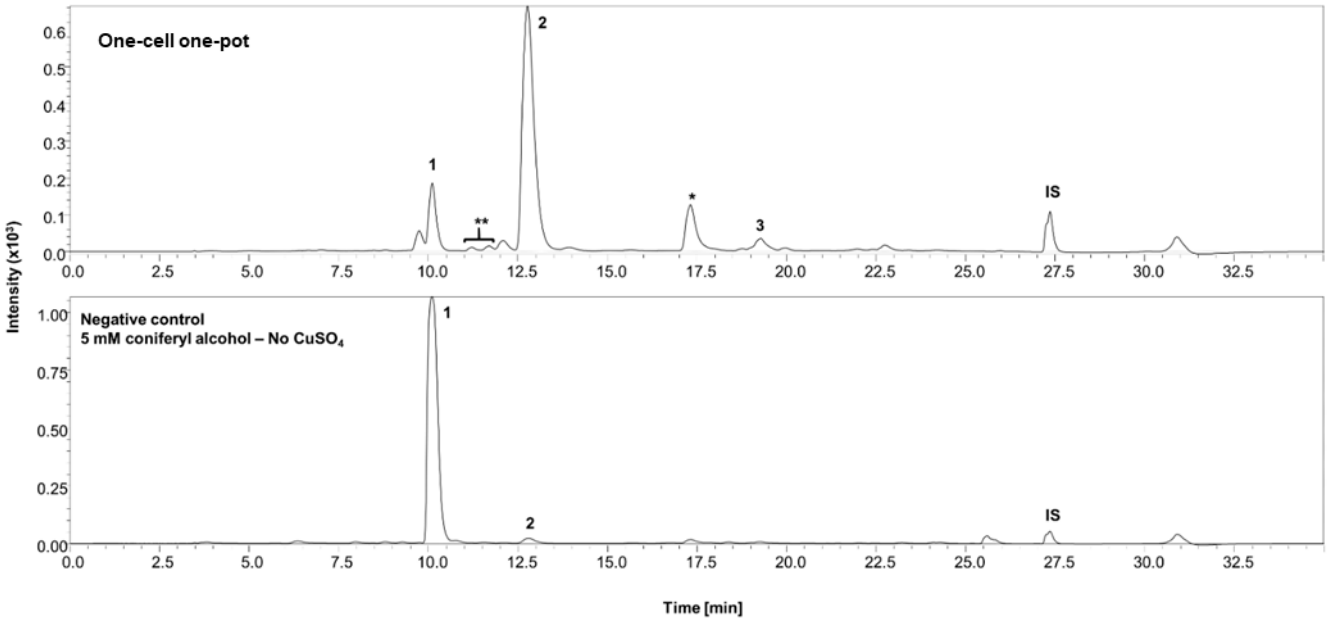


Figure S11. LC/MS analysis after coniferyl alcohol conversion to (-)-matairesinol. UV/Vis chromatograms at 280 nm are shown. “One-cell one-pot” setup. 1: coniferyl alcohol; 2: coniferyl aldehyde; 3: (±)-pinoresinol; **: (±)-erythro/threo-guaiacylglycerol 8-O-4'-coniferyl alcohol ethers; *: (±)-dehydrodiconiferyl alcohol; IS: internal standard sesamin.

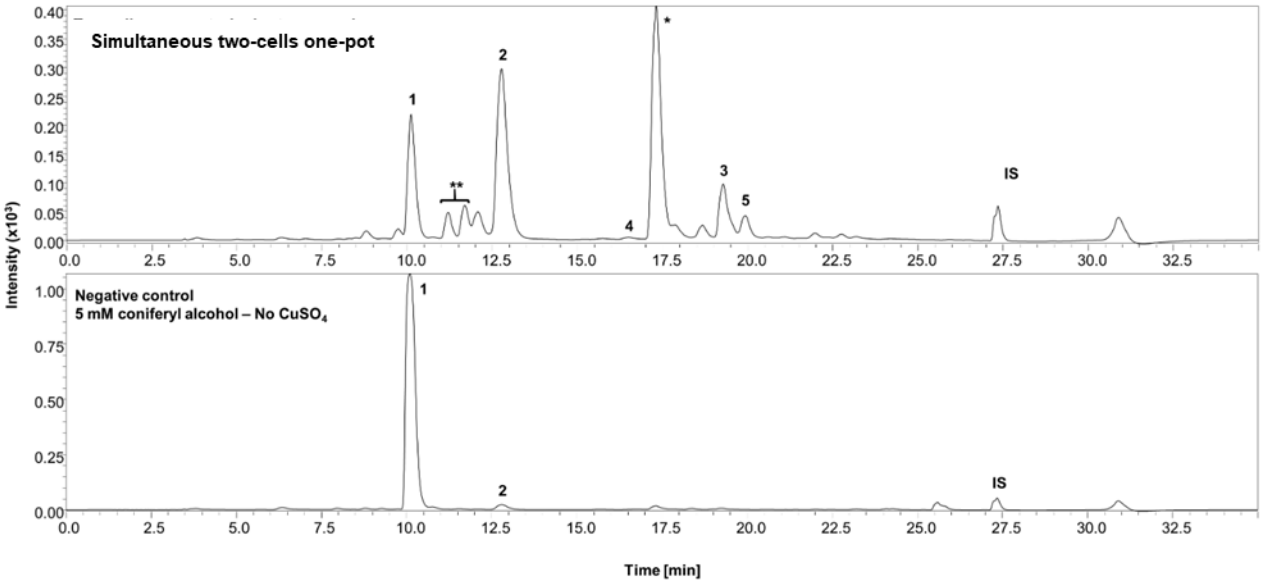


Figure S12. LC/MS analysis of coniferyl alcohol conversion to (-)-matairesinol. UV/Vis chromatograms at 280 nm are shown. Simultaneous “two-cells one-pot” setup. 1: coniferyl alcohol; 2: coniferyl aldehyde; 3: (±)-pinoresinol; 4: (+)-lariciresinol; 5: (-)-matairesinol; **: (±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ethers; *: (±)-dehydrodiconiferyl alcohol; IS: internal standard sesamin.

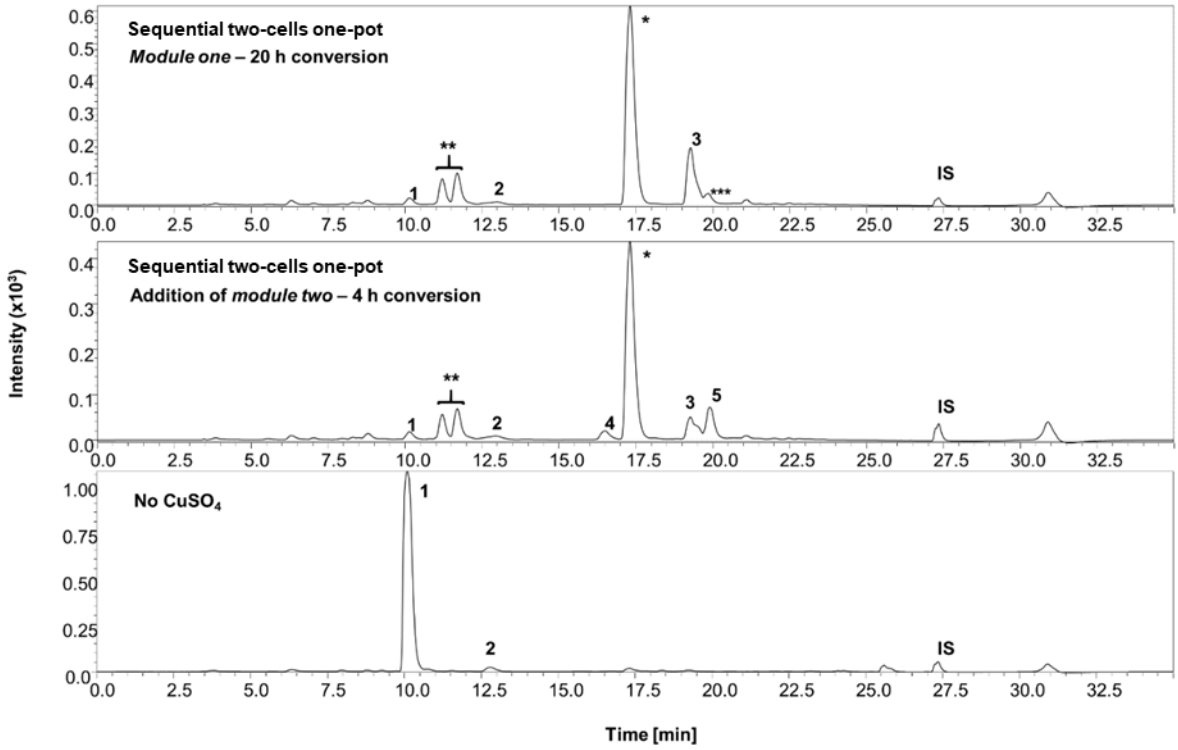


Figure S13. LC/MS analysis of coniferyl alcohol conversion to (-)-matairesinol. UV/Vis chromatograms at 280 nm are shown. Sequential "two-cells one-pot" setup. 1: coniferyl alcohol; 2: coniferyl aldehyde; 3: (±)-pinoresinol; 4: (+)-lariciresinol; 5: (-)-matairesinol; **: (±)-erythro/threo-guaiacylglycerol 8-O-4´-coniferyl alcohol ethers; *: (±)-dehydrodiconiferyl alcohol; ***: undefined product; IS: internal standard sesamin.

Supporting materials and methods

Bacterial strains, plasmids and reference compounds

Table S4. Features of bacterial strains used in this work

Bacterial strains	Genomic features
***E. coli* BL21 (DE3)**	F^- ompT gal dcm lon $hsdS_B(r_B^-m_B^-)$ λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) $[malB^+]_{K\text{-}12}(\lambda^S)$
E. coli* BL21 (DE3) Δ*cueO	F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) $[malB+]_{K\text{-}12}(\lambda^S)$ Δ(cueO)
***E. coli* C41 (DE3) Overexpress**	F – *ompT* $hsdS_B$ (r_B- m_B-) *gal dcm* (DE3)

Table S5. Features of plasmids used in this work

Plasmids	Features	Copy number
pET16b_cgl1	ColE1, P_{T7}, *lacI,* Amp^R	~40
pET16b	ColE1, P_{T7}, *lacI,* Amp^R	~40
pET24b	ColE1, P_{T7}, *lacI,* Kan^R	~40
pCDFDuet_*ppsdh_fiplr*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20-40

Table S6. Compound references for substrates or products used in this work and relative supplier

Substrate	Manufacturer	Molecular weight
Coniferyl alcohol	Sigma Aldrich	180.2
Ferulic acid	Sigma Aldrich	194.18
Tyrosol	Sigma Aldrich	138.16
2,6-DMP	Sigma Aldrich	154.17
Resveratrol	Sigma Aldrich	228.24
(+)-pinoresinol	Sigma Aldrich	358.39
(+)-lariciresinol	Sigma Aldrich	360.41
(±)-secoisolariciresinol	Sigma Aldrich	362.42
(-)-matairesinol	Phytolab	358.39
(+)-sesamin	TCI	354.35

LC/MS analysis

Table S7. Details of solvent gradient used in this work

Time [min]	ddH_2O + 0.1% Formic acid [%]
0.01	80
5.00	65
10.00	65
25.00	25
25.01	10 (100% MeOH)
26.00	10 (100% MeOH)
26.01	80
35.00	80

The column temperature was kept at 30°C and 1 µL samples were injected for the analysis. A mobile phase gradient constituted of methanol and double deionized water (ddH_2O) with 0.1% formic acid with 0.5 mL/min flow rate was used to separate metabolites. The separated compounds were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI); desolvation and block temperatures were kept at 275°C and 400°C, respectively. Nebulization gas flow was set at 1.5 L/min whereas the drying gas flow was fixed at 15 L/min. Analytes were monitored both based on their mass-to-charge ratio (m/z) in the positive ion mode between 159-1000 m/z and via UV/Vis signals at 280 nm. Substrate conversion was determined by comparison with a negative control, and products identified by retention times and plausible m/z fragmentation patterns from the literary sources and compared with ChemDraw data, Perkin Elmer).

Quantitative analysis

Quantitative analysis was performed for coniferyl alcohol depletion and its product pinoresinol formation. Pinoresinol was identified among different coupling products by comparing with a commercially available reference. Substrate conversion values and product distribution were calculated based on the corresponding peak areas in UV/Vis at 280 nm and by MS. The quantification of the produced racemic pinoresinol and (-)-matairesinol was done using the corresponding calibration curves. To this end, different pinoresinol concentrations ranging from 10 to 300 µM were mixed with 200 µM ferulic acid as internal standard, while 50 to 500 µM (-)-matairesinol solutions were mixed with 200 µM sesamin as internal standard. Mixtures were extracted and analyzed as described in the main experimental section. Ratios of the peak areas of the substrate to the internal standard (Sub/IS) were plotted against the according compound concentrations to give a straight-line calibration plot.

Table S8. Equations used for LC/MS data analysis

Conversion [%] (normalized to internal standard) *	$= 1- (Sub_{sample}/IS_{sample})/(Sub_{control}/IS_{control})$
Conversion [%]	$= \Sigma(P_{area}) / \Sigma(S_{area}+P_{area}) * 100$
Product distribution [%]	$= P_{area} / \Sigma(S_{area}+P_{area}) * 100$

*IS: internal standard, Sub: substrate

2.2 Synthesis of (-)-pluviatolide from (+)-pinoresinol

Title: *"Assembly of plant enzymes in E. coli for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide"*

Authors: Davide Decembrino, Esther Ricklefs, Stefan Wolhgemuth, Marco Girhard, Katrin Schullehner, Guido Jach and Vlada B. Urlacher

Published in: *ACS Synthetic Biology* 2020, 9, 11, 3091–3103

DOI: 10.1021/acssynbio.0c00354

Own contribution: conceived and conducted most of the experiments, analysed the data, designed the artwork, and prepared the original draft. Relative contribution 70%

ACS SyntheticBiology

pubs.acs.org/synthbio Research Article

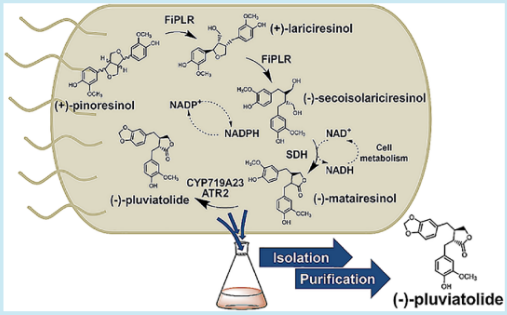
Assembly of Plant Enzymes in *E. coli* for the Production of the Valuable (−)-Podophyllotoxin Precursor (−)-Pluviatolide

Davide Decembrino, Esther Ricklefs, Stefan Wohlgemuth, Marco Girhard, Katrin Schullehner, Guido Jach, and Vlada B. Urlacher*

Cite This: https://dx.doi.org/10.1021/acssynbio.0c00354 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Lignans are plant secondary metabolites with a wide range of reported health-promoting bioactivities. Traditional routes toward these natural products involve, among others, the extraction from plant sources and chemical synthesis. However, the availability of the sources and the complex chemical structures of lignans often limit the feasibility of these approaches. In this work, we introduce a newly assembled biosynthetic route in *E. coli* for the efficient conversion of the common higher-lignan precursor (+)-pinoresinol to the noncommercially available (−)-pluviatolide *via* three intermediates. (−)-Pluviatolide is considered a crossroad compound in lignan biosynthesis, because the methylenedioxy bridge in its structure, resulting from the oxidation of (−)-matairesinol, channels the biosynthetic pathway toward the microtubule depolymerizer (−)-podophyllotoxin. This oxidation reaction is catalyzed with high regio- and enantioselectivity by a cytochrome P450 monooxygenase from *Sinopodophyllum hexandrum* (CYP719A23), which was expressed and optimized regarding redox partners in *E. coli*. Pinoresinol-lariciresinol reductase from *Forsythia intermedia* (FiPLR), secoisolariciresinol dehydrogenase from *Podophyllum pleianthum* (PpSDH), and CYP719A23 were coexpressed together with a suitable NADPH-dependent reductase to ensure P450 activity, allowing for four sequential biotransformations without intermediate isolation. By using an *E. coli* strain coexpressing the enzymes originating from four plants, (+)-pinoresinol was efficiently converted, allowing the isolation of enantiopure (−)-pluviatolide at a concentration of 137 mg/L (ee ≥99% with 76% isolated yield).



KEYWORDS: *E. coli, lignans, CYP719A23, (−)-pluviatolide, (−)-podophyllotoxin, biosynthesis*

For thousands of years, plants and plant-derived products have been representing irreplaceable components in human diet and inexhaustible resources of traditional medicines.[1] As life expectancy grew together with health consciousness, the demand for natural dietary supplements, functional food, and novel drugs has notably increased nowadays. Thus, novel approaches for the production of useful plant-derived compounds *via* "green" and sustainable ways are of high demand.[2] In this regard, lignans are particularly interesting secondary metabolites, as they have been found useful for the human body already in the late 1950s.[3] Common for these compounds is a broad range of biological activities that they have been proven to possess. Among others, anti-inflammatory and antioxidant activities as well as preventive effect of lignans against prostate and breast cancers and cardiovascular diseases have been reported.[4,5] Lignans like pinoresinol, lariciresinol, secoisolariciresinol, and matairesinol have been identified as phytoestrogens and are eventually transformed by the human intestinal microflora to the so-called enterolignans named enterodiol and enterolactone.[6] Furthermore, (−)-podophyllotoxin is an effective antiviral compound.[7] Glycosylated derivatives of (−)-podophyllotoxin like teniposide and etoposide have been used as chemotherapeutics for more than 20 years, with the latter one introduced into the World Health Organization list of essential medicines and health products.[8]

Despite the widespread occurrence of lignans in more than 60 species of vascular plants, stable and efficient strategies for their supply are limited.[9] For example, flaxseeds represent one of the major sources of dietary lignans and especially secoisolariciresinol. However, toxic cyanogenic glycosides are present in seeds as well; thus, their removal during the extraction procedure reduces the overall lignan titers.[10] (−)-Podophyllotoxin is currently isolated from the roots and rhizomes of *Podophyllum hexandrum*, which is distributed in very limited regions and is now endangered due to unregulated plant uprooting.[11]

Received: July 2, 2020

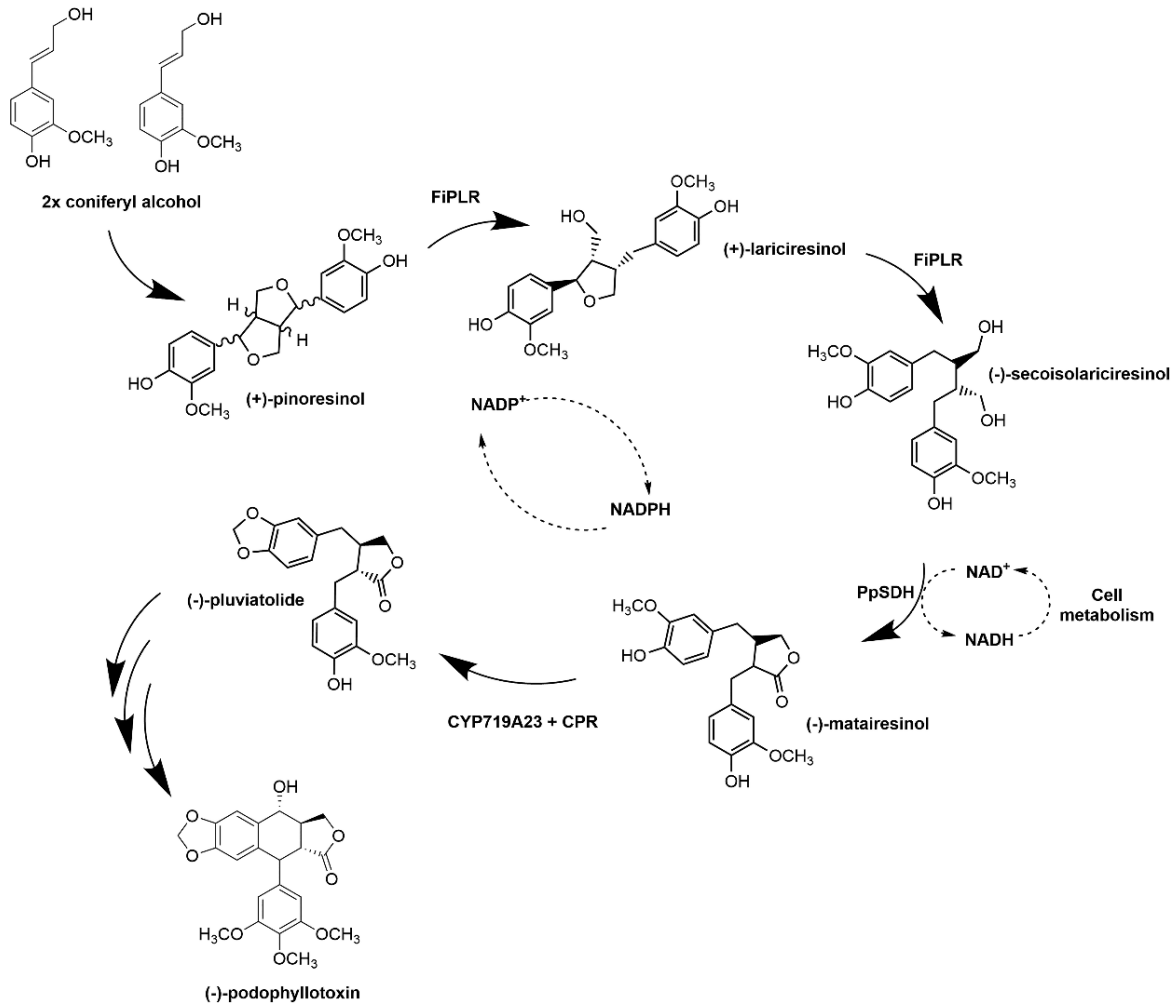


Figure 1. Schematic view of the enzymatic cascade for (−)-podophyllotoxin biosynthesis. The assembled enzymes used for (+)-pinoresinol biotransformation to (−)-pluviatolide are displayed.

Over the years, different approaches have been developed and applied for the production of lignans on a preparative scale. In most cases, these methods rely on high amounts of starting biomass and include cumbersome extraction and purification steps. Furthermore, the environmental impact of large-scale plant cultivation and the unregulated uprooting of some plant species have to be considered. Chemical total syntheses, as well as semisynthetic or chemo-enzymatic routes have also been developed, for instance, for the production of dibenzylbutane- (*e.g.*, secoisolariciresinol), dibenzylbutyrolactone- (*e.g.*, matairesinol), and aryltetralin-type lignans (*e.g.*, podophyllotoxin).[12−16] However, these approaches often involve the use of expensive or/and toxic reagents and catalysts. Despite belonging to the same lignan subtype of dibenzylbutyrolactones, (−)-pluviatolide has been investigated less extensively than its precursor (−)-matairesinol, even though its cytotoxic and anti-spasmodic effects have been reported.[17] These circumstances highlight why the development of alternative solutions is required. In this regard, recent advances in deciphering of biosynthetic pathways in plants combined with establishing of effective molecular tools for synthetic biology have accelerated the reconstitution of plant biosynthetic pathways or their parts in heterologous microorganisms.[18] In the past, metabolic engineering has been defined as the manipulation of specific biochemical reactions to improve or tune cellular properties, as well as the introduction of new ones by exploiting recombinant DNA technology.[19] However, despite the recent advances, the reconstitution of specific complex eukaryotic metabolic pathways in prokaryotic organisms is still an ongoing challenge.

The chemical structure of lignans consist of two phenylpropane units, like coniferyl alcohol, linked by a $\beta-\beta'$ bond. Previous reports described the possibility to biosynthesize such lignan precursors in *E. coli*[20−22] or other microorganisms.[23] Phenolic coupling of coniferyl alcohol leads to (±)-pinoresinol, which has been successfully produced in *E. coli*, as described in our previous work.[24] Given these premises, in this study, a biosynthetic pathway was introduced into *E. coli* to generate the noncommercially available lignan (−)-pluviatolide. Overall, (−)-pluviatolide represents a crucial intermediate in lignan

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

ACS Synthetic Biology pubs.acs.org/synthbio Research Article

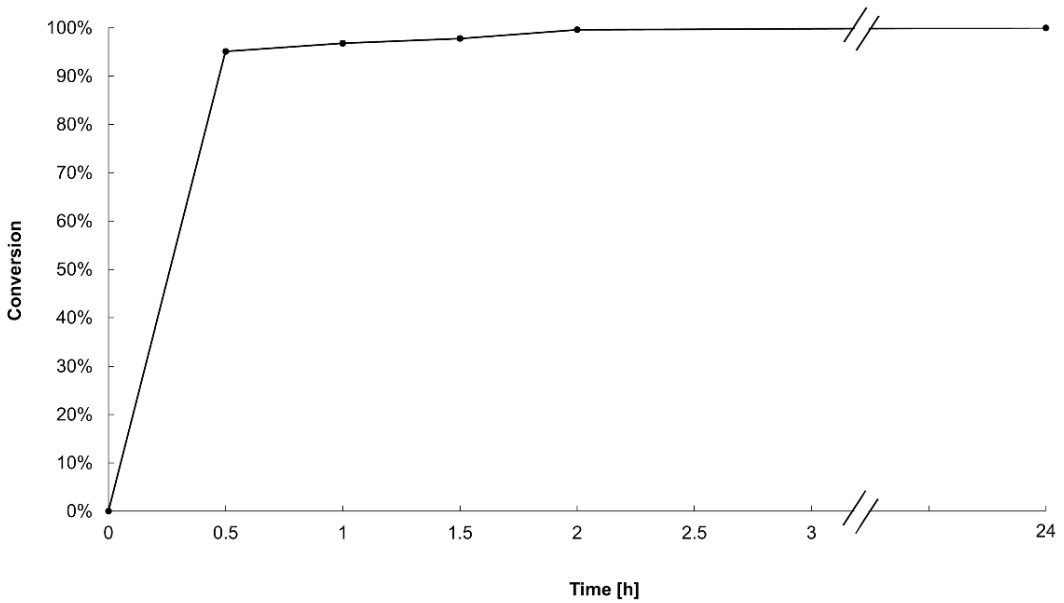


Figure 2. Time course of the conversion of 0.2 mM (±)-pinoresinol (77% ee of the (+)-enantiomer) by individually expressed FiPLR.

biosynthesis in plants: the methylenedioxy ring closure channels the biosynthesis toward (−)-podophyllotoxin instead of other lignans (Figure 1).[5] Specifically, in *Sinopodophyllum hexandrum*, this reaction is catalyzed with high regio- and enantioselectivity by a cytochrome P450 monooxygenase, CYP719A23, to produce (−)-pluviatolide. Up to now, besides *Nicotiana benthamiana* (tobacco) leaves, the heterologous expression of CYP719A23 has been reported only in *Saccharomyces cerevisiae*, where activity of this enzyme toward (−)-matairesinol was confirmed.[25,26]

The major sources of (−)-pluviatolide and (−)-podophyllotoxin are *S. hexandrum* and *Podophyllum pleianthum*. Yet, some enzymes involved in this metabolic route are present not only in these but also in other plants, like the pinoresinol-lariciresinol reductase from *Forsythia intermedia* (FiPLR)[27] used in this study. In the first step, (+)-pinoresinol is enantioselectively reduced to (+)-lariciresinol, which is reduced again in the second step to form (−)-secoisolariciresinol. Both reduction steps are catalyzed by FiPLR. In the third reaction, catalyzed by the secoisolariciresinol dehydrogenase from *P. pleianthum* (PpSDH), (−)-secoisolariciresinol is oxidized to (−)-matairesinol. Finally, (−)-matairesinol is oxidized by CYP719A23 to (−)-pluviatolide (Figure 1). In this study, plant CYP719A23 from *S. hexandrum* was functionally expressed in *E. coli* for the first time and its activity was tuned using different redox partners. To produce (−)-pluviatolide starting from (+)-pinoresinol, a modular two-plasmid *E. coli*-based system was designed. Using the optimized *E. coli* system, (−)-pluviatolide with an ee ≥99 and an overall isolated yield of 76% was produced at a concentration of 137 mg/L. The product structure and purity were confirmed by MS fragmentation patterns and NMR.

■ RESULTS AND DISCUSSION

Determination of Activity and Enantioselectivity of FiPLR and PpSDH. For the pinoresinol-lariciresinol reductase FiPLR expressed in *E. coli*, 95% conversion of 0.2 mM (±)-pinoresinol (77% ee of (+)-enantiomer, Supporting Information Figure S1) was observed after 30 min and 99% after 2 h (Figure 2). Complementary to this observation, the products distribution sets the percentage of secoisolariciresinol at 88% at any considered time point, whereas the intermediate product lariciresinol reached 6% after 30 min and increased to 11% after 1.5−2 h, with no significant variation after 24 h (11%). Considering the presence of 11.5% of (−)-pinoresinol and 88.5% of (+)-pinoresinol in the substrate mixture, complete conversion can result only from enzymatic activity of FiPLR toward both enantiomers. In order to clarify this aspect, the FiPLR catalyzed conversion was investigated also *via* chiral HPLC. To this end, either enantiopure (+)-lariciresinol or enantiopure (+)-pinoresinol (≥96% ee of (+)-enantiomer, Figure S1) at final concentration of 0.2 mM were used as substrates. In both cases, chiral HPLC analysis at 280 nm revealed the appearance of a single peak eluting at 13 min that was identified as (−)-secoisolariciresinol by the overlap in retention time with the reference compound. Conversion of (±)-pinoresinol (77% ee of (+)-enantiomer) was analyzed *via* chiral HPLC too, revealing a second minor peak eluting at 26 min, which was suggested to be (−)-lariciresinol (Figure 3). As this compound is noncommercially available, we compared the minor peak (Figure 3B) and its MS data with those of (−)-lariciresinol, which was produced by the (−)-pinoresinol selective reductase from *Arabidopsis thaliana* (Figure 3F) and isolated in our previous work.[24] Up to now, a strong preference toward (+)-pinoresinol and (+)-lariciresinol has been described for FiPLR in both subsequent reduction steps, leading to the enantiospecific synthesis of (−)-secoisolariciresinol.[28] However, although the preference of FiPLR toward (+)-pinoresinol remains clear, the minor amounts of (−)-pinoresinol present in the stock solutions seem to be metabolized by FiPLR in the first reduction step as well, leading to (−)-lariciresinol. The latter product is not reduced by the same enzyme in the second step, which explains the complete conversion of (±)-pinoresinol (ee 77%) accom-

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

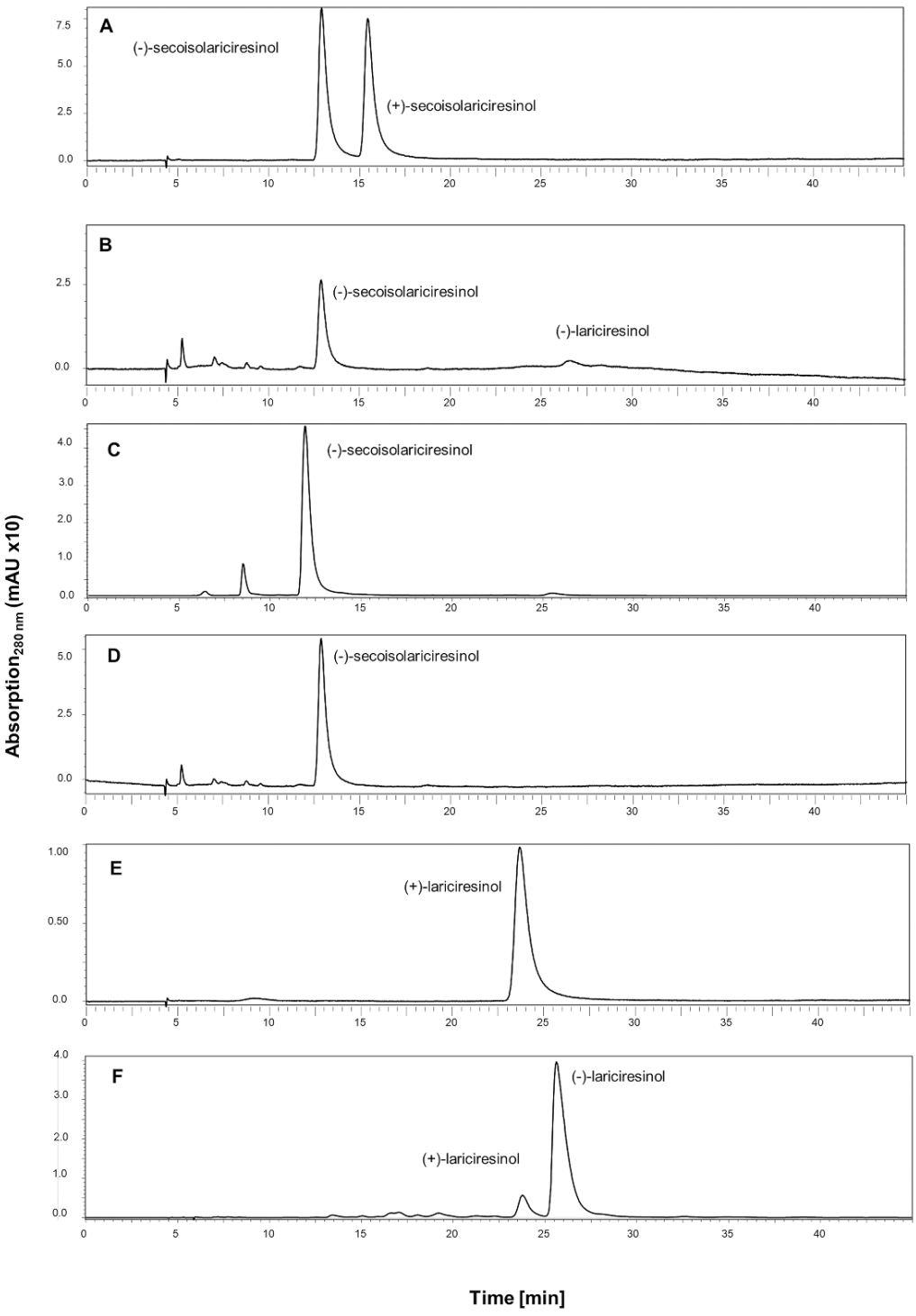


Figure 3. Chiral HPLC analysis of FiPLR activity. (A) Enantiomeric composition of (±)-secoisolariciresinol reference. (B) FiPLR activity with (±)-pinoresinol (77% ee of (+)-enantiomer). (C) FiPLR activity with (+)-pinoresinol (≥96% ee of (+)-enantiomer). (D) FiPLR activity with the intermediate (+)-lariciresinol as the substrate. (E) Authentic (+)-lariciresinol reference. (F) Enantiomeric composition of isolated (−)-lariciresinol taken from our previous work.[24] The separation of all substances including (−)-lariciresinol was performed using method 2 (Table S9).

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

panied by the 2-fold increase of the lariciresinol signal over time with no further conversion.

A similar scenario took place when FiPLR and the secoisolariciresinol dehydrogenase PpSDH were coexpressed in *E. coli*. Within this reaction cascade, (±)-pinoresinol (77% ee of (+)-enantiomer) was converted *via* two consecutive reduction reactions catalyzed by FiPLR, first to (+)-lariciresinol and then to (−)-secoisolariciresinol. These two reductions were subsequently followed by the oxidation of (−)-secoisolariciresinol to (−)-matairesinol catalyzed by PpSDH, whereas (−)-lariciresinol remained unconverted (Figure S2). Chiral HPLC analysis revealed the enantiomeric purity of formed (−)-matairesinol of >99.9% (Figure S3D). *E. coli* cells expressing PpSDH alone were also tested with racemic secoisolariciresinol. Chiral HPLC analysis revealed clear depletion for the (−)-enantiomer of the racemic secoisolariciresinol, which was converted to (−)-matairesinol with high enantioselectivity (Figure S3A). Thus, our observations confirmed the results from the previous study describing PpSDH as selective for (−)-secoisolariciresinol.[29]

Expression of N-Terminal-Modified CYP719A23 Variants in *E. coli*. FiPLR and PpSDH have cytoplasmic localization in the plant cells with no membrane association; thus, they could be expressed in soluble form in *E. coli*. Differently, CYP719A23 is an integral membrane protein bound to the membrane of the endoplasmic reticulum. Cytochromes P450 (P450, CYPs) use molecular oxygen as an oxidant and require for their activity electrons that are ultimately derived from the reduced nicotine amide cofactors. Electrons derived from NADPH are transferred to most eukaryotic P450s, including the plant ones, *via* a membrane bound cytochrome P450 reductase (CPR).[30] Therefore, (−)-pluviatolide production in *E. coli* is dependent on the one hand on functional expression of CYP719A23 and on the other hand on proper interactions between CYP719A23 and an appropriate CPR in the recombinant cell.

Cytochromes P450 are involved in various biosynthetic pathways and have become important biobricks for synthetic biology, acting as keystones in the synthesis of highly complex molecules.[31,32] In this regard, the heterologous expression of membrane-bound P450 enzymes in bacterial hosts represents the crucial step. However, production of plant membrane-bound P450s in *E. coli* is generally a challenging task due to the lack of a well-developed inner membrane.[33] As expected, the wildtype *cyp719a23* gene could not be expressed in *E. coli* under the conditions described in the experimental section, neither utilizing the native nor the codon-optimized sequence. Hence, the N-terminal membrane-associated sequence was manipulated, since successful expression of plant P450s in *E. coli* has been achieved using this strategy previously.[34−36] The first 26 amino acids of native CYP719A23 were predicted to interact with membrane; thus, *cyp719a23* variants with the truncated first 26 amino acids (Δ1−26) were created in both native (*cyp719a23nat_t*) or codon-optimized (*cyp719a23oc_t*) forms. Alternatively, functional expression was attempted targeting the bacterial inner plasma membrane by replacing the native N-terminal region with the sequences originating from the N-termini of the bovine CYP17A1 (LLLAVFL) or the rabbit CYP2C3 (KKTSSKGK),[37,38] again with both native (*cyp719a23nat_bov* and *cyp719a23nat_rab*) and codon-optimized sequences (*cyp719a23oc_bov* and *cyp719a23oc_rab*). In all cases, the triplet encoding for alanine was added as the second codon of the gene, since this strategy was proven effective to enhance eukaryotic P450 expression in bacterial hosts[39,40] (Tables S4−S5). Remarkably, all *cyp719a23* variants with manipulated N-terminal sequence led to the production of the correctly folded P450 in *E. coli*, confirmed by CO-difference spectra, which is used as a measure for functional expression of P450s as well as for their quantification.

Among all variants, the Δ1−26 truncated form resulted in the highest expression levels with native and codon-optimized sequences of 1012 ± 41 and 1457 ± 120 $\mu g/g_{cww}$, respectively (Table S12). Since the expression of codon-optimized sequence in *E. coli* was higher, the variant *cyp719a23oc_t* (from now on referred to as CYP719A23), was chosen for the implementation into the reaction cascade.

Generally, the expression temperature and duration substantially influenced CYP719A23 expression (Table S13). Although the highest expression of CYP719A23 was achieved at 20 °C, 25 °C was chosen as a compromise for establishing the reaction cascade with regard to the optimal expression conditions for FiPLR and PpSDH reported elsewhere[41] and confirmed in our previous study.[24]

Effect of Redox Partners and CYP719A23 Expression on (−)-Matairesinol Conversion. Since potential physiological redox partners of CYP719A23, CPR from *S. hexandrum* are unknown to date, various nonphysiological redox partner proteins were tested for their ability to reconstitute P450 activity. The NADPH-dependent cytochrome P450 reductase 2 from *Arabidopsis thaliana* (ATR2) was selected as potential candidate, because it has been intensively characterized and used to reconstitute activity of several plant P450s.[35,42,43] The bacterial redox couple comprising the flavin reductase FdR from *E. coli* and the flavodoxin YkuN from *Bacillus subtilis* was chosen because it has also been reported to support various bacterial and eukaryotic P450s.[35,44] In addition, CYP719A23 was fused to ATR2, since fusion proteins of plant CYP−CPR, mimicking the natural bacterial chimera P450 BM3 from *Bacillus megaterium*, have been described to enhance yields of high-value plant P450-derived products in *E. coli*.[34,45]

First, CYP719A23 activity was investigated upon conversion of enantiopure (−)-matairesinol with *E. coli* coexpressing CYP719A23 and the respective redox partners from a single pETDuet-1 expression vector. The fastest conversion was observed for *E. coli* coexpressing CYP719A23 with ATR2; LC/MS analysis revealed 50% conversion of 0.2 mM enantiopure (−)-matairesinol after ∼15 min. Almost complete substrate depletion (99%) was observed after 1 h, and no m/z fragments related to (−)-matairesinol could be detected after 1.5 h. For the CYP-ATR2 fusion, 50% substrate depletion was observed after 45 min and 93% after 3 h. With FdR/YkuN, 50% conversion of enantiopure (−)-matairesinol was reached after 2.5−3 h, increasing to 82% after 24 h (Figure S4).

In an attempt to explain these differences in substrate conversions with various CYP-redox partner combinations, we investigated P450 concentration in the corresponding recombinant *E. coli* cells. Interestingly, given the same conditions of cell growth and cloning backbone, only in the cells coexpressing CYP719A23 and ATR2 was a clearly detectable absorbance peak at 450 nm observed, indicating a higher P450 concentration in this construct (170 ± 4 $\mu g/g_{cww}$) compared to those of the two others (Figure S4 and Table S14).

The observations *in vivo* indicate that (−)-matairesinol consumption correlated with the P450 concentration in the cell. In order to better elucidate the role of P450 concentration

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

ACS Synthetic Biology pubs.acs.org/synthbio Research Article

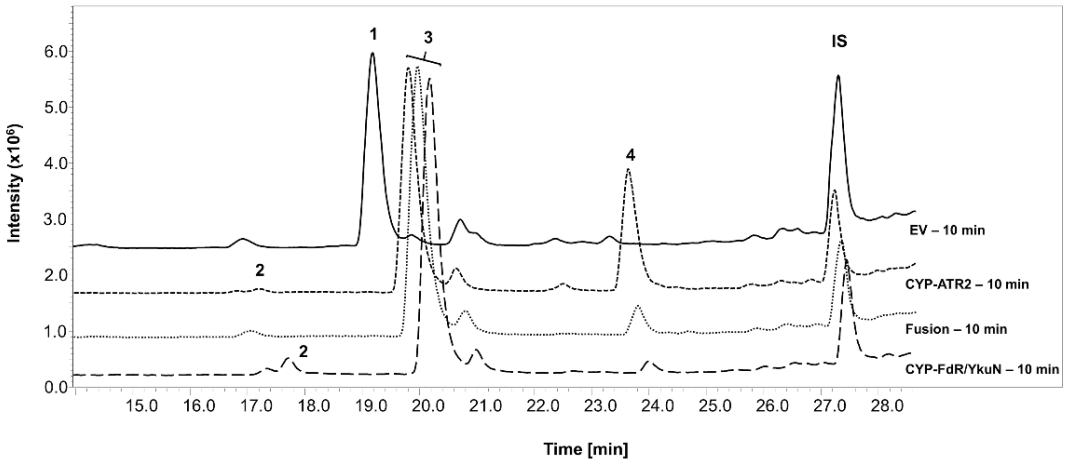


Figure 4. (+)-Pinoresinol conversion after a 10 min reaction time with *E. coli* coexpressing FiPLR, PpSDH, CYP719A23, and redox partners. Negative control (EV) with (+)-pinoresinol (≥96% ee) is represented with solid line. (1) (+)-Pinoresinol, retention time (RT) 19 min (m/z 359 [M + H$^+$], 341 [M + H − H_2O]). (2) (−)-Secoisolariciresinol, RT 17.5 min (m/z 345 [M + H − H_2O], 327 [M − 2OH$^-$], 363 [M + H$^+$]). (3) (−)-Matairesinol, RT 20 min (m/z 359 [M + H$^+$], 341 [M − OH$^-$], 376 [M + H_3O^+]). (4) (−)-Pluviatolide, RT 24.7 min (m/z 357 [M + H$^+$], 339 [M + H − H_2O], 375 [M + H_3O^+]). IS, internal standard (+)-sesamin. Conversion was performed in 500 μL of reaction volume in 1 mL Eppendorf tubes with open lids, at 25 °C, 1500 rpm, 0.2 mM substrate, and 2% DMSO.

and redox partners interactions on productivity of the three systems, we performed (−)-matairesinol conversion to (−)-pluviatolide *in vitro*. As the stoichiometry between CYP and CPR in the fusion construct is 1:1, first, this fixed ratio was used to investigate the P450 reaction with ATR2 and FdR/YkuN. Using this setup, after 4 h, ∼17% of (−)-matairesinol was converted by CYP719A23 supported with ATR2 and only ∼4% by CYP719A23 with FdR/YkuN. Differently, the CYP−CPR fusion led to 66% substrate depletion, suggesting that the physical linkage makes the interaction between CYP719A23 and ATR2 more efficient. Complementarily, as the use of an excess of redox partners over P450 often improves *in vitro* activity in reconstituted systems, we applied a 1:10 ratio to the CYP719A23/ATR2 couple and a 1:1:10 ratio to the CYP719A23/FdR/YkuN system. As expected, P450 activity increased, leading to higher conversion values with both redox systems: up to ∼64% with CYP719A23 combined with ATR2 and ∼17% with CYP719A23 supported by FdR/YkuN (Figure S5). In all cases, the prolongation of reaction time to 24 h did not result in complete (−)-matairesinol depletion. As a matter of fact, the best performances *in vitro* were recorded for the CYP−CPR fusion and for CYP719A23 supported by ATR2 at a ratio of 1:10 with ∼69% and ∼76% conversion, respectively (Figure S5). Thus, the CYP−CPR fusion (with 1:1 ratio) appears to be the most effective system, and comparable activity of the individual CYP719A23 *in vitro* was achieved only with a 10-fold excess of ATR2. The use FdR−YkuN was recognized as the weakest strategy to sustain CYP719A23 activity both *in vivo* and *in vitro*.

The highest (−)-pluviatolide concentration achieved *in vivo* correlated with the highest P450 concentration measured in the *E. coli* cells, in which CYP719A23 was coexpressed with ATR2. As all three CYP719A23/redox partner gene combinations were identically cloned into the same pET-vector with similar plasmid copy numbers and under control of the IPTG-inducible T7 promoter, comparison of the expression background is possible. It can be stated that the choice of redox partners has affected the P450 expression and thus the *E. coli* productivity. Despite the superior *in vitro* activity of the CYP−CPR fusion, its poor expression in *E. coli* counterbalances the situation, rendering *E. coli* coexpressing CYP719A23 and ATR2 more effective for conversion *in vivo*.

Further, both the interaction between P450 and redox partners and the concentration of redox partners in the cell might play roles as well. In this regard, some studies report an optimal esteemed ratio of P450 to CPR of ∼15:1 or higher for natural plant systems;[46,47] however, this is relevant to physiological interaction between P450s and cognate CPRs, where the electron transfer is supposed to be optimal. In this sense, the use of a native CPR may represent an effective strategy to improve P450 catalytic activity. As mentioned before, native CPRs from *S. hexandrum* are unknown to date. Thus, the discovery of physiological cognate redox partner CPR among several isoforms and the eventual production of active CPR are factors, which should not be underestimated.[46] Furthermore, although CYP−CPR fusions may secure the electron shuttling to the target P450,[34] the 1:1 ratio of fusion partners may not always represent the most feasible and effective solution in *E. coli*[48] or the linkage may act as physical constraint.[49]

Whereas some reports describe that higher CPR ratios resulted in better P450 efficacy in recombinant microbial hosts, others argue that increased redox partners' levels may have a negative effect on cell viability and also on P450 expression and activity, possibly due to increased burden, resource competition, or inefficient electron transfer.[49−51] In our study, this seems to be the case for FdR/YkuN, as their coexpression negatively affected P450 expression and caused inefficient electron transfer (shown *in vitro*). In any case, engineering of both P450 and CPR represents a valuable starting point for further improvement and extension of the cascade.

Enantioselectivity of CYP719A23. Next, the enantiospecificity of CYP719A23 was investigated using 0.2 mM racemic

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

ACS Synthetic Biology pubs.acs.org/synthbio Research Article

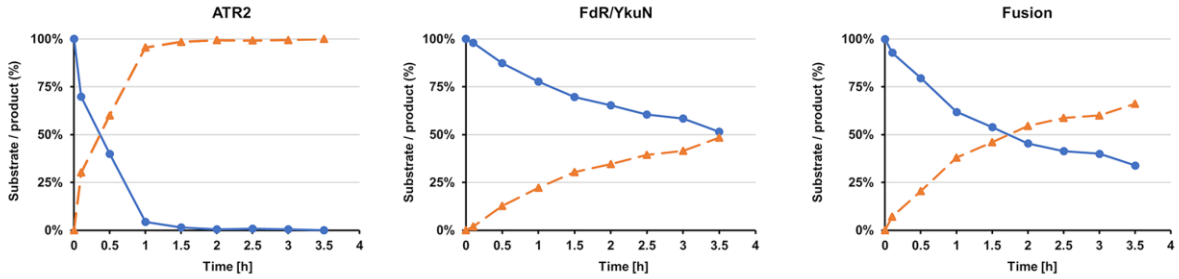


Figure 5. (−)-Pluviatolide biosynthesis over time with FiPLR, PpSDH, and CYP719A23/redox partners coexpressed in one *E. coli* strain. As the substrate (+)-pinoresinol is fully consumed after 10 min reaction, percentages of depleted (−)-matairesinol (dots, solid line) against (−)-pluviatolide formation (diamonds, dashed line) are depicted.

(±)-matairesinol as the substrate in whole-cell conversions. Prior to the conversion, chiral HPLC analysis of the stock revealed an ee-value of 1% (Figure S1), meaning the 0.2 mM substrate load was roughly composed of 50% (+)- and 50% (−)-matairesinol. In this case, LC/MS analysis revealed 50% conversion of (±)-matairesinol within 3.5 h, with no further change over 24 h for all strains, implying that only one enantiomer was consumed. Specifically, 50% conversion was observed after ∼10 min with CYP719A23 coexpressed with ATR2, after ∼1.5−2 h with CYP719A23 supported by FdR/YkuN and after ∼30 min applying the CYP−CPR fusion (Figure S6). Unambiguously, these values are coherent to those observed for the conversion of enantiopure (−)-matairesinol. To confirm the enantiospecificity of CYP719A23, the oxidation of (±)-matairesinol in comparison to that of enantiopure (−)-matairesinol catalyzed by isolated CYP719A23 *in vitro* was analyzed *via* chiral HPLC. It revealed, as expected, the depletion of the (−)-enantiomer only (Figure S7). The previous literature reports described that this P450 converts only (−)-matairesinol, but no experimental evidence existed that (+)-matairesinol is not accepted as substrate.[25,26]

Production of (−)-Pluviatolide from Pinoresinol in *E. coli*. Although the highest (−)-matairesinol conversion was achieved with *E. coli* coexpressing CYP719A23 and ATR2, this scenario may change during the reconstitution of the complete cascade (Figure 1), because the coexpression of multiple enzymes may influence the overall metabolic stoichiometry. For this reason, *E. coli* was cotransformed with two plasmids, generating a modular coexpression system: a pCDFDuet-1 vector harboring the genes encoding for FiPLR and PpSDH (FiPLR-PpSDH module) and either a pETDuet-1 or pET28a-(+) vector harboring P450-redox partner combinations (P450-module), as shown in Figure S8.

Again, the focus was put on the (−)-pluviatolide production with regard to the chosen CYP719A23/redox-partner system. Time course analysis revealed full depletion of 0.2 mM (+)-pinoresinol (≥96% ee, Figure S1) after 10 min at 25 °C in all samples (Figure 4). At this reaction time, only minor residues of the intermediate (−)-secoisolariciresinol were detected, while (−)-matairesinol was only partially converted to (−)-pluviatolide. This indicates high activity of FiPLR, satisfactory activity of PpSDH, and rather low activity of CYP719A23 being a bottleneck of this artificial pathway. Regarding the distribution between (−)-matairesinol and (−)-pluviatolide after 10 min reaction, roughly 30% of the produced (−)-matairesinol was depleted when CYP719A23 was combined with ATR2, whereas only <5% and <10% depletion was achieved with FdR/YkuN and CYP−CPR fusion, respectively (Figure 4). This demonstrates that the efficacy of the combination CYP719A23 and ATR2 remained the highest also when incorporated into the cascade.

Overall, the biosynthesis of (−)-pluviatolide in the constructed recombinant *E. coli* strains seems not to be significantly affected by the colocalization of the two plasmids (Figure 5 and Tables S14 and S15). Only the CYP−CPR fusion seems to be negatively influenced by the coexpression with FiPLR and PpSDH, as 50% of (−)-matairesinol was detected after 2 h within the cascade (Figure S6 and Table S15), whereas the same conversion was achieved after 45 min with the fusion expressed alone (Figure S4 and Table S14).

It is no surprise that the biotransformation of (−)-matairesinol to (−)-pluviatolide was identified as a bottleneck of the artificial biosynthesis in *E. coli*. Cytochrome P450-catalyzed reactions have often been identified as rate-limiting in reconstituted biosynthetic cascades due to several reasons. Along with issues related to membrane association and incorrect folding, host growth inhibition and alteration of microbial energy metabolism have also been reported to negatively affect heterologous expression of eukaryotic P450s.[18,52] These factors together with the requirement for redox partners may be more critical to achieve optimal metabolic flux within artificial cascades, compared to other involved enzymes. In fact, in this work, the productivity was influenced more by the P450-module rather than by the FiPLR−PpSDH module or the overall metabolic burden.

In addition to these considerations, the formation of a multienzyme complex can be hypothesized, which could support the overall performance of the cascade too. Over the years, more and more enzymes belonging to the same biosynthetic pathways have been reported to build the so-called metabolons. Moreover, examples of artificial metabolons established using physiologically unrelated proteins are present in the literature.[53,54] In our work, the interaction among the enzymes that are involved in the developed cascade seems likely to take place. Even if not originating from the same plant but belonging to the same biosynthetic pathway, the pinoresinol-lariciresinol reductase from *F. intermedia* combined with the secoisolariciresinol dehydrogenase from *P. pleianthum* successfully produced high levels of (−)-matairesinol. However, different post-translational modifications and the removal and modification of the membrane anchor sequence in the recombinant enzymes result in the altered protein−protein interactions as compared with the native enzymes in plants, which limits the verification of such hypotheses.

Biosynthesis of (−)-Pluviatolide at a Preparative Scale. Enzyme proximity is also important for efficient

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

ACS Synthetic Biology pubs.acs.org/synthbio Research Article

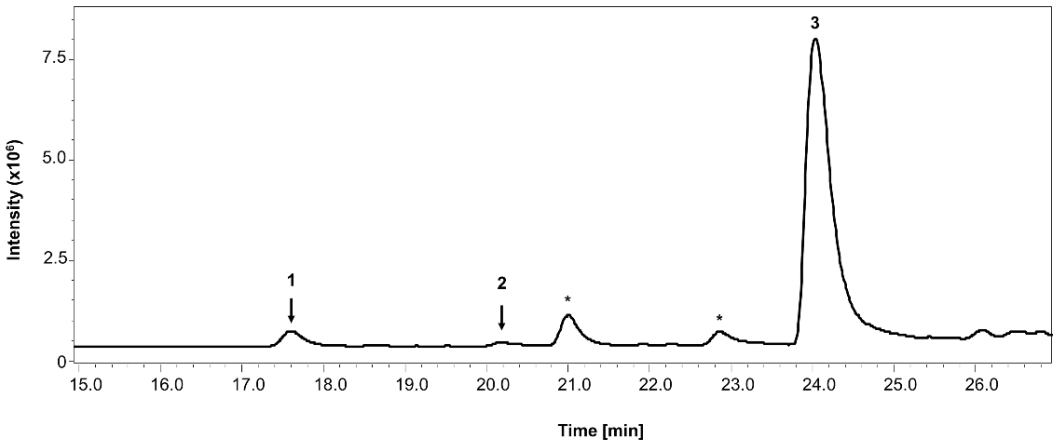


Figure 6. Conversion of (+)-pinoresinol for the isolation of (−)-pluviatolide in the optimized setup after 5 h. 100% substrate conversion. (1) Secoisolariciresinol, 3.2%, RT 17.5 min (m/z 345 [M + H − H_2O], 327 [M − $2OH^-$], 363 [M + H^+]). (2) (−)-Matairesinol, 0.6%, RT 20 min (m/z 359 [M + H^+], 341 [M + H − H_2O], 376 [M + H_3O^+]). (3) (−)-Pluviatolide, 96.2%, RT 24.7 min (m/z 357 [M + H^+], 339 [M + H − H_2O], 375 [M + H_3O^+]).

cofactor recycling. NADPH utilized in the reactions catalyzed by FiPLR and CYP719A23-ATR2, and NAD^+ used by PpSDH, have to be regenerated via cell metabolism to ensure enzyme activity (Figure 1). Moreover, uncoupling reactions may occur within CYP−CPR interactions. Thus, the addition of a carbon source to sustain *E. coli* metabolism and to avoid resource competition between heterologous and endogenous reactions was considered. Glucose was chosen for ease of use considering that its metabolism in *E. coli* provides both NADPH and NADH *via* pentose phosphate pathway, glycolysis, and TCA cycle.[55] Moreover, a beneficial effect of glucose on *E. coli* resting cells carrying P450 enzymes have been described in the literature.[56,57]

Utilizing the *E. coli* strain coexpressing FiPLR, PpSDH, CYP719A23, and ATR2 as chassis, preparative biotransformations were performed with (+)-pinoresinol at a concentration of 180 mg/L. Similar to experiments at low scale, (+)-pinoresinol was completely converted after 10 min. As in the above-described experiments, the reaction step catalyzed by CYP719A23 was revealed to be the bottleneck. Specifically, when using Erlenmeyer flasks at a shaking speed of 180 rpm, only 60% of (−)-matairesinol was converted to (−)-pluviatolide after 24 h, with stagnation after 48 h. As P450 enzymes require molecular oxygen for their activity, proper oxygen supply becomes sometimes critical for the system.[58] Thus, we increased the shaking speed up to 250 rpm, which enhanced (−)-matairesinol conversion to 80% after 24 h. Aiming at improved cells aeration, we switched from nonbaffled to baffled Erlenmeyer flasks at 250 rpm shaking speed. The increased oxygen supply was found to be crucial to enhance P450 activity and the efficiency of (−)-pluviatolide biosynthesis, leading to <1% detectable (−)-matairesinol after 5 h (Figure 6), along with 100% of (+)-pinoresinol converted and >96% of (−)-pluviatolide synthesized.

Chiral analysis of samples from the individual time points revealed that (−)-pluviatolide was the sole product formed from (−)-matairesinol, which confirmed high enantiospecificity of CYP719A23 also at a higher scale (Figure S9). All things considered, proper aeration was recognized as a crucial matter for further upscaling experiments. For the same reason, in this study, we chose resting cells over growing cells, as oxygen consumption by the host has been reported to be sensibly higher during exponential phase rather than in the stationary phase.[59] In this way, the potential competition between cell metabolism and enzymatic activity for oxygen utilization should be minimized.

(−)-Pluviatolide was successfully isolated and purified, resulting in 76% isolated yield with >99% UV/vis purity and >94% MS purity. The structure of the isolated product was analyzed and confirmed *via* ^1H and ^{13}C NMR (Tables S16 and S17 and Figures S11−S13). Previously, CYP719A23 was expressed in *S. cerevisiae* to confirm enzyme activity.[26] Starting with 36 mg/L (−)-matairesinol, 5 mg of (−)-pluviatolide was isolated after one-step oxidation catalyzed by CYP719A23, which means a 13% isolated yield, whereas the herein constructed *E. coli* strain allowed us to isolate 137 mg/L (−)-pluviatolide starting from 180 mg/L (+)-pinoresinol. A reason for this can be the use of a more appropriate redox partner for CYP719A23, while in the previously reported experiment, CYP719A23 activity was supported by *S. cerevisiae*'s endogenous CPRs. Most importantly though is that this result was achieved by combining four enzymes from different plants catalyzing four consecutive biotransformations.

Further improvement of the production rate and the final concentration of (−)-pluviatolide should involve various factors. Cloning of the orthologous/homologous genes identified in other plants like CYP719A24 from *P. peltatum*,[25] pinoresinol-lariciresinol reductase 2 from *Thuja plicata*,[60] and (−)-secoisolariciresinol dehydrogenase from *F. intermedia*[29] with the same enantioselectivity might lead to higher concentrations of the final product due to their possibly better expression in *E. coli* and/or higher catalytic activities. Higher heterologous expression of CYP719A23 or the orthologous-membrane-associated P450s and CPRs might be achieved in *E. coli* strains with altered lipid composition of cell membranes.[61] Protein engineering can also be applied to enhance activity of CYP719A23 or to improve its interaction with redox partners because it has been broadly used for other P450s, including those from plants.[62] Understanding the interaction and interdependency between P450s and their redox partners in

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

a heterologous host is fundamental to develop functional efficient microbial factories. Thus, engineering of the redox partner would require estimation of the optimal P450/CPR ratio in order to tune the relative expression titer by using different promoters or gene copy numbers. Genome mining efforts should be made to identify the cognate CPR for CYP719A23. Furthermore, metabolic engineering of *E. coli* for more efficient NADPH recycling or the introduction of the genes from the shikimic acid pathway would enable the production of lignans from amino acids or glucose with high efficiency. Finally, bioengineering approaches including high cell density cultivation in a bioreactor and substrate feeding could also contribute to process optimization.

■ CONCLUSIONS

Herein, we report the development of a functional recombinant *E. coli* strain in which four enzymes originating from different plants catalyzed four sequential biotransformations with high enantioselectivity to efficiently synthesize a highly complex lignan (−)-pluviatolide with multiple stereocenters. Racemic or (+)-pinoresinol were used as substrates. Basically, both forms can be isolated from plant material.[63] However, low amounts of this compound in plants reflected by high prices for its recovery and notable environmental impact due to the long plant growth period has prompted researchers to look for other sources or to develop alternative synthetic routes. For example, recently endophytic fungi have been identified as an alternative source of (+)-pinoresinol.[64] A number of chemical synthetic approaches have been suggested to afford racemic pinoresinol starting with simple low-value, low-price compounds like methyl acetoacetate.[65] Furthermore, in our previous studies, we explored an *in vitro* enzymatic cascade as well as *E. coli* whole-cell biocatalysts to produce (±)-pinoresinol from the inexpensive and abundant precursor eugenol with following kinetic resolution catalyzed by pinoresinol-lariciresinol reductases.[24,66] Thus, this system can be combined with the strain developed within this study to produce (−)-pluviatolide starting with the cheap natural product eugenol.

The production of (−)-pluviatolide itself can be considered from multiple perspectives. First, the herein proposed biosynthesis of a noncommercially available plant compound in a prokaryotic organism may lay the foundation for its production at a high scale and elucidation of its biological activities. Second, (−)-pluviatolide is a starting compound for the introduction of further modules toward downstream high-value lignans such as (−)-yatein or (−)-deoxypodophyllotoxin, in order to get one step closer to the biosynthesis of (−)-podophyllotoxin in bacteria. Complementarily, the modularity of this two-plasmid system offers the possibility to synthesize (−)-secoisolariciresinol or (−)-matairesinol at need. Although further metabolic engineering efforts have to be implemented, this study provides a step forward toward a reliable, feasible alternative for the production of higher-lignans.

■ METHODS

Bacterial Strains, Enzymes and Chemicals. For cloning purposes, *E. coli* DH5α (Clonetech) was used, while *E. coli* OverExpress C41 (DE3) (Lucigen) was utilized for gene expression. Enzymes for cloning procedures (Phusion High Fidelity DNA polymerase, endonucleases *Nco*I, *Hind*III, *Nde*I, *Xho*I) originated from Thermo Scientific, while T5 exonuclease and Taq-DNA Ligase were from New England BioLabs. Reference compounds used in this work are described in the Supporting Information (Table S1).

Plasmid Design and Reaction Cascade Construction. Synthetic genes were ordered from Eurofins genomics with codon optimization for *E. coli*. Pinoresinol-lariciresinol reductase from *F. intermedia* (GenBank AAC49608; FiPLR) and secoisolariciresinol dehydrogenase from *P. pleianthum*/*Dysosma pleiantha* (GenBank AHB18702; PpSDH) were cloned into the pCDFDuet-1 plasmid (Sigma-Aldrich) into multiple cloning sites I and II, respectively (Tables S2 and S3). The *cyp719a23* gene from *S. hexandrum* (Gene bank AGC29953.1) was ordered both as native and codon optimized sequences. Both gene sequences were cloned into pETDuet-1 *via* sequence- and ligation-independent cloning (SLIC).[67] As potential physiological cytochrome P450 reductases (CPR) from *S. hexandrum* are unknown to date, CYP719A23 activity within the cascade was ensured by the coexpression of various nonphysiological redox partner proteins, namely, the NADPH-cytochrome P450 reductase 2 from *Arabidopsis thaliana* (Protein ID NP_194750.1) ATR2, chosen because of proven effective expression in *E. coli*[35,42] (Tables S2 and S3), or a bicistronic construct composed by *E. coli* NADPH-dependent reductase (FdR) and *Bacillus subtilis* flavodoxin (YkuN) as described elsewhere[68] (Tables S2 and S3). The gene sequences of these redox partner proteins were inserted into pETDuet-1 within multiple cloning site I. In addition, a CYP−CPR fusion was generated by fusing the C-terminal domain of CYP719A23 (variant *cyp_719a23_oc_t*) with ATR2 (Tables S2 and S3). The fusion construct was cloned into pET28a(+) (Merck).

The putative membrane associated N-terminal sequence of CYP719A23 was identified by the TMHMM software (version 2.0; http://www.cbs.dtu.dk/services/TMHMM/), and CYP719A23 truncated variants were created and screened for expression in *E. coli*, as summarized in Tables S4 and S5.

For *in vitro* P450 activity reconstitution, CYP719A23, ATR2,[35] FdR, and YkuN[69]—as well as glucose dehydrogenase (GDH) from *Bacillus megaterium* (GenBank accession no. D10626)—were subcloned in pET-vectors, as summarized in the Supporting Information (Table S3).

The correct insertion of all genes was verified by automated Sanger DNA-sequencing (Eurofins Genomics). The CYP719A23 variant that achieved the highest expression level was implemented into the whole-cell reaction cascade and its expression was optimized (see the Supplementary Results in the Supporting Information). Generally, the cascade design was based on the colocalization of two plasmids and coexpression of three, four, or five enzymes, as summarized in Figure S8.

Expression in *E. coli*. The expression of target genes was carried out in 50 mL of Terrific broth (TB) media supplemented with required antibiotics (streptomycin 50 μg/mL and/or ampicillin 100 μg/mL and/or kanamycin 30 μg/mL) and inoculated with 500 μL from an overnight culture in Luria−Bertani (LB) medium. Cells were grown at 37 °C, 180 rpm to an OD_{600} of 0.6. The expression was induced with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and 0.5 mM 5-aminolevulinic acid (5-ALA) and 0.1 mM $FeSO_4$ were added at the same time point if a P450 was coexpressed. Thereafter, the cultures were incubated for 48 h at 25 °C, 120 rpm. Single expressions of GDH, ATR2, and

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

YkuN were carried out for 21 h at 25 °C, 140 rpm, whereas 20 °C, 180 rpm were set for FdR production.

Cell Density Normalization and *in Vivo* Reaction Setup. *E. coli* cultures were harvested by centrifugation (30 min, 3220*g*, 4 °C), and the cell pellets were stored at −20 °C. Prior to biotransformations or cell disruption for expression analysis, cell pellets were thawed and the cell wet weight (cww) was determined and normalized to 70 g/L (corresponding to ∼15 g/L cell dry weight) in resuspension buffer (80% 50 mM K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5, 500 mM D-glucose).

Ten milliliter aliquots from these cell suspensions were supplemented with 0.1 mM IPTG prior to the reaction start. Biotransformations at analytical scale were set at 500 μL of reaction volume, 0.2 mM substrate concentration, 2% DMSO in 2 mL reaction tubes at 25 °C, 1500 rpm in a Thermomixer C (Eppendorf).

The activity of single-expressed FiPLR was investigated *via* time course sampling every 30 min over 2 h as well as after 24 h *via* LC/MS. For the comparison on CYP/redox partners' combinations activity, a time course sampling of 10 min, 30 min, and then every 30 min for 3.5 h and after 24 h was set to follow either single reactions or the complete cascade under identical reaction conditions.

Preparative flask-scale reactions were set for (−)-pluviatolide isolation as follows: after cell wet weight normalization, 10 mL of reaction buffer containing (+)-pinoresinol (0.5 mM or 180 mg/L) and 5% DMSO were poured in 100 mL Erlenmeyer flasks or 100 mL baffled Erlenmeyer flasks. Biotransformation was performed at 25 °C in an orbital shaker Multitron (Infors HT) at 180 rpm as starting conditions. The shaking speed was raised up to 250 rpm for reaction optimization. Five hundred microliter aliquots were taken to follow the reaction over time during reaction optimization. A control reaction composed of *E. coli* C41(DE3) supplemented with substrate but without the desired enzymes was used for comparison.

Cell Disruption and Enzymatic Assays. Prior to cell disruption, the cww was normalized to 70 g/L in case of P450 concentration estimation for *in vivo* experiments, whereas for *in vitro* assays, harvested cells were resuspended in 5 mL of potassium phosphate buffer (80% 50 mM K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5). Samples were then supplemented with 100 μM phenylmethylsulfonylfluorid (PMSF). Cell disruption was performed *via* sonication in Sonifier SFX250 (Branson Ultrasonics). Soluble and insoluble protein fractions were separated by ultracentrifugation (40 000*g*, 25 min, 4 °C).

CO-difference spectra were recorded for the estimation of P450 concentration as described elsewhere.[70] From every biological replicate, three measurements were performed. A calculation was done using the extinction coefficient $\varepsilon_{450\ nm}$ = 91 mM^{-1} cm^{-1}.[70]

Prior to the estimation of concentration, redox partner proteins were purified as described elsewhere.[35,69] Concentrations were determined spectrophotometrically referring to previously reported extinction coefficients: ATR2, $\varepsilon_{453\ nm}$ = 13.7 mM^{-1} cm^{-1} and $\varepsilon_{381\ nm}$ = 12.6 mM^{-1} cm^{-1}; FdR, $\varepsilon_{456\ nm}$ = 7.01 mM^{-1} cm^{-1}, YkuN, $\varepsilon_{461\ nm}$ = 10.01 mM^{-1} cm^{-1}. The GDH activity was determined in cell lysate by measuring the generation of NADPH during glucose oxidation at 340 nm. The extinction coefficient $\varepsilon_{340\ nm}$ = 6.22 mM^{-1} cm^{-1} was used to calculate volumetric activity (U/mL).

Determination of P450 Activity *in Vitro*. After expression, cell disruption, and protein fractions separation, P450 concentration was estimated in the clarified cell lysate containing soluble CYP719A23 or CYP−CPR fusion. *In vitro* activity determination was performed using 0.5 μM P450. In the fusion construct, the ratio of CYP and CPR is 1:1; thus, to maintain the same proportion of P450/FAD/FMN, 0.5 μM ATR2 or 0.5 μM FdR and 0.5 μM YkuN were used in the respective reactions. Additionally, 5 μM ATR2 resulting in a P450/ATR2 ratio of 1:10 or 5 μM YkuN resulting in a P450/FdR/YkuN ratio of 1:1:10 were used. As a cofactor, 0.2 mM NADPH was applied together with a cofactor regeneration system composed of glucose-6-phosphate dehydrogenase (GDH) 5 U/mL and 20 mM D-glucose. Bovine liver catalase (600 U/mL) were added to reduce the reactive oxygen species potentially generated due to uncoupling reactions. *In vitro* experiments were performed at an analytical scale in 250 μL of reaction volume, 0.2 mM (−)-matairesinol as substrate, 2% DMSO in 2 mL reaction tubes at 25 °C, 1500 rpm in a Thermomixer C (Eppendorf). The reaction was performed in technical triplicate, and 4 and 24 h were taken as conversion time checkpoints.

Metabolite Extraction. For analytical purposes, 0.2 mM of an internal standard (IS) was added to the solution. The lignan (+)-sesamin was chosen as the IS due to its structural similarity to the substrate (+)-pinoresinol. Metabolites were extracted twice with ethyl acetate (1 mL for analytical, 20 mL for preparative scale reactions). The resulting organic phases were collected and evaporated under reduced pressure. Thereafter, samples were resuspended either in methanol (99.9% LC/MS grade, Fisher Scientific)—for LC/MS analysis—or ethanol (HPLC grade, Merck) for chiral HPLC or semipreparative HPLC.

LC/MS and Chiral HPLC Analysis. The conversion of (+)-pinoresinol to (−)-pluviatolide *via* the intermediates (+)-lariciresinol, (−)-secoisolariciresinol, and (−)-matairesinol was analyzed *via* liquid chromatography coupled with mass spectrometry (LC/MS) on an LCMS-2020 system (Shimadzu) equipped with a Chromolith Performance RP-18e column (100 mm × 4.6 mm, Merck). The column temperature was kept at 30 °C, and 1 μL of sample was injected. Separation was carried out with a mobile phase gradient constituting of methanol and double deionized water (ddH_2O) with 0.1% formic acid and a flow rate of 0.5 mL/min (Table S6). The separated compounds were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI); desolvation and block temperatures were kept at 275 and 400 °C, respectively. The nebulization gas flow was set at 1.5 L/min, whereas the drying gas flow was set at 15 L/min. Analytes were monitored both by their mass-to-charge ratio (m/z) in the positive ion mode between 159 and 1000 m/z and UV/vis at 280 nm. From LC/MS measurements, substrate conversion values were calculated with the product distribution resulting from the peak areas of reaction products and substrates in both MS and UV/vis (see Table S7). The analytes were identified by their retention times and typical fragmentation patterns in MS in comparison to commercially available reference compounds (Table S8). All values represent the average of measurements from biological and technical duplicates at least.

The enantiomeric composition of the synthesized compounds was investigated *via* chiral HPLC (Shimadzu) equipped with a CHIRALPACK IB column 250 mm × 4.6 mm (Daicel Chiral Technologies). The column was kept at 30 °C, and the separation was carried out with a solvent mixture of ethanol and *n*-hexane (95% HPLC, Chemsolute) under

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

isocratic mode with 0.7 mL/min flow rate (Table S9). ee% values were calculated from peak areas at 280 nm (Table S10).

Product Isolation and Analysis. (−)-Pluviatolide was isolated *via* semipreparative HPLC (Shimadzu) equipped with a MultoHigh 100 Si-10 μ column (250 mm × 10 mm, pore size 100 Å, 10 μm particle size, CS-Chromatographie Service) with a solvent gradient of *n*-heptane (99% HPLC, Roth) and ethanol (Table S11). The column temperature was set to 15 °C with a 7.5 mL/min flow rate, and the elution was monitored by UV/vis in a range of 190−800 nm. Fractions containing (−)-pluviatolide were collected, evaporated under reduced pressure, and then resuspended in methanol for LC/MS purity check. Methanol was evaporated, and the isolated (−)-pluviatolide underwent freeze-drying in order to remove any solvent residues. The desired product was identified by ^{1}H, ^{13}C, and ^{1}H−^{13}C-HSQC NMR analysis in comparison to those of previously published literature references.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00354.

Tables of compounds used for LC/MS and HPLC and relative suppliers, codon optimized sequences, overview on the genes and features of plasmids used in this work, names and features for the generated *cyp719a23* gene variants, SLIC primers used for *cyp719a23* gene amplification, sequence substitution or truncation, details for the LC/MS method used, equations used for LC/MS and chiral HPLC data analysis, overview on compounds observed features, details for the chiral HPLC and preparative HPLC methods used, overview on the expression levels achieved and optimization, correlation of (−)-matairesinol conversion and detected P450 concentration, and comparison of ^{1}H and ^{13}C spectroscopic assignments, figures of chiral HPLC analysis, LC/MS analysis, ratio of substrate depletion and product synthesis, CO-difference spectra, diverse ratios of CYP719A23-redox partners, conversion tendencies, overview on the modular two-plasmid system, and ^{1}H and ^{13}C NMR spectra, and discussions of screening of CYP719A23 variants, disclaimer on (−)-pluviatolide availability, and NMR analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vlada B. Urlacher − *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany;* orcid.org/0000-0003-1312-4574; Email: vlada.urlacher@uni-duesseldorf.de

Authors

Davide Decembrino − *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Esther Ricklefs − *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Stefan Wohlgemuth − *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Marco Girhard − *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Katrin Schullehner − *Phytowelt Green Technologies GmbH, 41334 Nettetal, Germany*

Guido Jach − *Phytowelt Green Technologies GmbH, 41334 Nettetal, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00354

Author Contributions

D.D. conceived and conducted most of the experiments, analyzed the data, designed the artwork, and prepared the original draft. E.R. and S.W. contributed to conceiving preliminary experiments on FiPRL and PpSDH characterization. M.G. contributed to results validation and manuscript writing and editing. K.S. and G.J. constructed and evaluated the CYP−CPR fusion. V.B.U. contributed to conceiving experiments, interpretation of data, and writing of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the Federal Ministry of Education and Research, Germany to Heinrich-Heine University Düsseldorf [grant number 031B0362A] and to Phytowelt GreenTechnologies GmbH [grant number 031B0362D] under the umbrella of the "Nationale Forschungsstrategie BioÖkonomie 2030" project "LignaSyn".

■ ABBREVIATIONS

CYP, cytochrome P450; CPR, NADPH-cytochrome P450 reductase; FdR, flavin reductase from *E. coli*; YkuN, flavodoxin from *B. subtilis*; ATR2, NADPH-cytochrome P450 reductase 2 from *Arabidopsis thaliana*; FiPLR, pinoresinol-lariciresinol reductase from *Forsythia intermedia*; PpSDH, secoisolariciresinol dehydrogenase from *Podophyllum pleianthum*; LC, liquid chromatography; MS, mass spectrometry; HPLC, high-pressure liquid chromatography

■ REFERENCES

(1) Pan, S. Y., Litscher, G., Gao, S. H., Zhou, S. F., Yu, Z. L., Chen, H. Q., Zhang, S. F., Tang, M. K., Sun, J. N., and Ko, K. M. (2014) Historical perspective of traditional indigenous medical practices: the current renaissance and conservation of herbal resources. *Evid. Based Complement. Alternat. Med. 2014*, 525340.

(2) Satake, H., Shiraishi, A., Koyama, T., Matsumoto, E., Morimoto, K., Bahabadi, S. E., Ono, E., and Murata, J. (2017) Lignan biosynthesis for food bioengineering. In *Handbook of Food Bioengineering* (Grumezescu, A., and Holban, A. M., Eds.) 1st ed., pp 351−379, Academic Press, New York.

(3) Kelly, M. G., and Hartwell, J. L. (1954) The biological effects and the chemical composition of podophyllin. A review. *J. Natl. Cancer Inst. 14*, 967−1010.

(4) Zalesak, F., Bon, D. J. D., and Pospisil, J. (2019) Lignans and neolignans: plant secondary metabolites as a reservoir of biologically active substances. *Pharmacol. Res. 146*, 104284.

(5) Teponno, R. B., Kusari, S., and Spiteller, M. (2016) Recent advances in research on lignans and neolignans. *Nat. Prod. Rep. 33*, 1044−1092.

(6) Landete, J. M., Arques, J., Medina, M., Gaya, P., de Las Rivas, B., and Munoz, R. (2016) Bioactivation of phytoestrogens: intestinal bacteria and health. *Crit. Rev. Food Sci. Nutr. 56*, 1826−1843.

(7) Yu, X., Che, Z., and Xu, H. (2017) Recent advances in the chemistry and biology of podophyllotoxins. *Chem. - Eur. J. 23*, 4467−4526.

(8) World Health Organization (2017) *WHO model list of essential medicines*, 20th list.

(9) Gordaliza, M., García, P. A., del Corral, J. M., Castro, M. A., and Gómez-Zurita, M. A. (2004) Podophyllotoxin: distribution, sources, applications and new cytotoxic derivatives. *Toxicon 44*, 441−459.
(10) Sok, D., Cui, S. H., and Kim, M. R. (2009) Isolation and bioactivities of furfuran type lignan compounds from edible plants. *Recent Pat. Food Nutr. Agric. 1*, 87−95.
(11) Nadeem, M., Palni, L. M. S., Purohit, A. N., Pandey, H., and Nandi, S. K. (2000) Propagation and conservation of *Podophyllum hexandrum* Royle an important medicinal herb. *Biol. Conserv. 92*, 121−129.
(12) Eklund, P., Lindholm, A., Mikkola, J. P., Smeds, A., Lehtila, R., and Sjoholm, R. (2003) Synthesis of (−)-matairesinol, (−)-enterolactone, and (−)-enterodiol from the natural lignan hydroxymatairesinol. *Org. Lett. 5*, 491−493.
(13) Stadler, D., and Bach, T. (2008) Concise stereoselective synthesis of (−)-podophyllotoxin by an intermolecular iron(III)-catalyzed Friedel-Crafts alkylation. *Angew. Chem., Int. Ed. 47*, 7557−7559.
(14) Jeffries, D. E., and Lindsley, C. W. (2019) Asymmetric synthesis of natural and unnatural dibenzylbutane lignans from a common intermediate. *J. Org. Chem. 84*, 5974−5979.
(15) Lazzarotto, M., Hammerer, L., Hetmann, M., Borg, A., Schmermund, L., Steiner, L., Hartmann, P., Belaj, F., Kroutil, W., Gruber, K., and Fuchs, M. (2019) Chemoenzymatic total synthesis of deoxy-, epi-, and podophyllotoxin and a biocatalytic kinetic resolution of dibenzylbutyrolactones. *Angew. Chem., Int. Ed. 58*, 8226−8230.
(16) Li, J., Zhang, X., and Renata, H. (2019) Asymmetric chemoenzymatic synthesis of (−)-podophyllotoxin and related aryltetralin lignans. *Angew. Chem., Int. Ed. 58*, 11657−11660.
(17) Zhang, G., Shimokawa, S., Mochizuki, M., Kumamoto, T., Nakanishi, W., Watanabe, T., Ishikawa, T., Matsumoto, K., Tashima, K., Horie, S., Higuchi, Y., and Dominguez, O. P. (2008) Chemical constituents of *Aristolochia constricta* antispasmodic effects of its constituents in guinea-pig Ileum and isolation of a diterpeno-lignan hybrid. *J. Nat. Prod. 71*, 1167−1172.
(18) Chemler, J. A., and Koffas, M. A. (2008) Metabolic engineering for plant natural product biosynthesis in microbes. *Curr. Opin. Biotechnol. 19*, 597−605.
(19) Stephanopoulos, G. (1999) Metabolic fluxes and metabolic engineering. *Metab. Eng. 1*, 1−11.
(20) Aschenbrenner, J., Marx, P., Pietruszka, J., and Marienhagen, J. (2019) Microbial production of natural and unnatural monolignols with *Escherichia coli*. *ChemBioChem 20*, 949−954.
(21) Chen, Z., Sun, X., Li, Y., Yan, Y., and Yuan, Q. (2017) Metabolic engineering of *Escherichia coli* for microbial synthesis of monolignols. *Metab. Eng. 39*, 102−109.
(22) Jansen, F., Gillessen, B., Mueller, F., Commandeur, U., Fischer, R., and Kreuzaler, F. (2014) Metabolic engineering for p-coumaryl alcohol production in *Escherichia coli* by introducing an artificial phenylpropanoid pathway. *Biotechnol. Appl. Biochem. 61*, 646−654.
(23) Yang, J., Liang, J., Shao, L., Liu, L., Gao, K., Zhang, J. L., Sun, Z., Xu, W., Lin, P., Yu, R., and Zi, J. (2020) Green production of silybin and isosilybin by merging metabolic engineering approaches and enzymatic catalysis. *Metab. Eng. 59*, 44−52.
(24) Ricklefs, E., Girhard, M., and Urlacher, V. B. (2016) Three-steps in one-pot: whole-cell biocatalytic synthesis of enantiopure (+)- and (−)-pinoresinol via kinetic resolution. *Microb. Cell Fact. 15*, 78.
(25) Marques, J. V., Kim, K. W., Lee, C., Costa, M. A., May, G. D., Crow, J. A., Davin, L. B., and Lewis, N. G. (2013) Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis. *J. Biol. Chem. 288*, 466−479.
(26) Lau, W., and Sattely, E. S. (2015) Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science 349*, 1224−1228.
(27) Markulin, L., Corbin, C., Renouard, S., Drouet, S., Gutierrez, L., Mateljak, I., Auguin, D., Hano, C., Fuss, E., and Laine, E. (2019) Pinoresinol-lariciresinol reductases, key to the lignan synthesis in plants. *Planta 249*, 1695−1714.
(28) Dinkova-Kostova, A. T., Gang, D. R., Davin, L. B., Bedgar, D. L., Chu, A., and Lewis, N. G. (1996) (+)-Pinoresinol/(+)-lariciresinol reductase from *Forsythia intermedia*. *J. Biol. Chem. 271*, 29473−29482.
(29) Xia, Z. Q., Costa, M. A., Pelissier, H. C., Davin, L. B., and Lewis, N. G. (2001) Secoisolariciresinol dehydrogenase purification, cloning, and functional expression. Implications for human health protection. *J. Biol. Chem. 276*, 12614−12623.
(30) Hannemann, F., Bichet, A., Ewen, K. M., and Bernhardt, R. (2007) Cytochrome P450 systems - biological variations of electron transport chains. *Biochim. Biophys. Acta, Gen. Subj. 1770*, 330−344.
(31) Girvan, H. M., and Munro, A. W. (2016) Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Curr. Opin. Chem. Biol. 31*, 136−145.
(32) Renault, H., Bassard, J. E., Hamberger, B., and Werck-Reichhart, D. (2014) Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Curr. Opin. Plant Biol. 19*, 27−34.
(33) Schuler, M. A., and Werck-Reichhart, D. (2003) Functional genomics of P450s. *Annu. Rev. Plant Biol. 54*, 629−667.
(34) Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science 330*, 70−74.
(35) Kranz-Finger, S., Mahmoud, O., Ricklefs, E., Ditz, N., Bakkes, P. J., and Urlacher, V. B. (2018) Insights into the functional properties of the marneral oxidase CYP71A16 from *Arabidopsis thaliana*. *Biochim. Biophys. Acta, Proteins Proteomics 1866*, 2−10.
(36) Morrone, D., Chen, X., Coates, R. M., and Peters, R. J. (2010) Characterization of the kaurene oxidase CYP701A3, a multifunctional cytochrome P450 from gibberellin biosynthesis. *Biochem. J. 431*, 337−344.
(37) Barnes, H. J., Arlotto, M. P., and Waterman, M. R. (1991) Expression and enzymatic activity of recombinant cytochrome P450 17α-hydroxylase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A. 88*, 5597−5601.
(38) von Wachenfeldt, C., Richardson, T. H., Cosme, J., and Johnson, E. F. (1997) Microsomal P450 2C3 is expressed as a soluble dimer in *Escherichia coli* following modifications of its N-terminus. *Arch. Biochem. Biophys. 339*, 107−114.
(39) Zelasko, S., Palaria, A., and Das, A. (2013) Optimizations to achieve high-level expression of cytochrome P450 proteins using *Escherichia coli* expression systems. *Protein Expression Purif. 92*, 77−87.
(40) Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A., and van Knippenberg, P. H. (1987) Influence of the codon following the AUG initiation codon on the expression of a modified lacZ gene in *Escherichia coli*. *EMBO J. 6*, 2489−2492.
(41) Kuo, H. J., Wei, Z. Y., Lu, P. C., Huang, P. L., and Lee, K. T. (2014) Bioconversion of pinoresinol into matairesinol by use of recombinant *Escherichia coli*. *Appl. Environ. Microbiol. 80*, 2687−2692.
(42) Hull, A. K., and Celenza, J. L. (2000) Bacterial expression and purification of the *Arabidopsis* NADPH-cytochrome P450 reductase ATR2. *Protein Expression Purif. 18*, 310−315.
(43) Schuckel, J., Rylott, E. L., Grogan, G., and Bruce, N. C. (2012) A gene-fusion approach to enabling plant cytochromes P450 for biocatalysis. *ChemBioChem 13*, 2758−2763.
(44) Jenkins, C. M., and Waterman, M. R. (1994) Flavodoxin and NADPH-flavodoxin reductase from *Escherichia coli* support bovine cytochrome P450c17 hydroxylase activities. *J. Biol. Chem. 269*, 27401−27408.
(45) Leonard, E., and Koffas, M. A. (2007) Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. *Appl. Environ. Microbiol. 73*, 7246−7251.
(46) Jensen, K., and Møller, B. L. (2010) Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry 71*, 132−141.
(47) Bassard, J.-E., Møller, B. L., and Laursen, T. (2017) Assembly of dynamic P450-mediated metabolons - Order versus chaos. *Curr. Mol. Biol. Rep. 3*, 37−51.

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(48) Yang, H., Liu, F., Li, Y., and Yu, B. (2018) Reconstructing biosynthetic pathway of the plant-derived cancer chemopreventive precursor glucoraphanin in *Escherichia coli*. *ACS Synth. Biol. 7*, 121−131.
(49) Biggs, B. W., Lim, C. G., Sagliani, K., Shankar, S., Stephanopoulos, G., De Mey, M., and Ajikumar, P. K. (2016) Overcoming heterologous protein interdependency to optimize P450-mediated taxol precursor synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A. 113*, 3209−3214.
(50) Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D., and Dumas, B. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat. Biotechnol. 21*, 143−149.
(51) Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., and Newman, J. D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature 496*, 528−532.
(52) Wagner, S., Baars, L., Ytterberg, A. J., Klussmeier, A., Wagner, C. S., Nord, O., Nygren, P. A., van Wijk, K. J., and de Gier, J. W. (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol. Cell. Proteomics 6*, 1527−1550.
(53) Singleton, C., Howard, T. P., and Smirnoff, N. (2014) Synthetic metabolons for metabolic engineering. *J. Exp. Bot. 65*, 1947−1954.
(54) Thomik, T., Wittig, I., Choe, J. Y., Boles, E., and Oreb, M. (2017) An artificial transport metabolon facilitates improved substrate utilization in yeast. *Nat. Chem. Biol. 13*, 1158−1163.
(55) Pontrelli, S., Chiu, T. Y., Lan, E. I., Chen, F. Y., Chang, P., and Liao, J. C. (2018) *Escherichia coli* as a host for metabolic engineering. *Metab. Eng. 50*, 16−46.
(56) Tieves, F., Erenburg, I. N., Mahmoud, O., and Urlacher, V. B. (2016) Synthesis of chiral 2-alkanols from n-alkanes by a *P. putida* whole-cell biocatalyst. *Biotechnol. Bioeng. 113*, 1845−1852.
(57) Walton, A. Z., and Stewart, J. D. (2004) Understanding and improving NADPH-dependent reactions by nongrowing *Escherichia coli* cells. *Biotechnol. Prog. 20*, 403−411.
(58) Kaluzna, I., Schmitges, T., Straatman, H., van Tegelen, D., Müller, M., Schürmann, M., and Mink, D. (2016) Enabling selective and sustainable P450 oxygenation technology. Production of 4-hydroxy-α-isophorone on kilogram scale. *Org. Process Res. Dev. 20*, 814−819.
(59) Riedel, T. E., Berelson, W. M., Nealson, K. H., and Finkel, S. E. (2013) Oxygen consumption rates of bacteria under nutrient-limited conditions. *Appl. Environ. Microbiol. 79*, 4921−4931.
(60) Hwang, J. K., Moinuddin, S. G. A., Davin, L. B., and Lewis, N. G. (2020) Pinoresinol-lariciresinol reductase: Substrate versatility, enantiospecificity, and kinetic properties. *Chirality 32*, 770−789.
(61) Kanonenberg, K., Royes, J., Kedrov, A., Poschmann, G., Angius, F., Solgadi, A., Spitz, O., Kleinschrodt, D., Stuhler, K., Miroux, B., and Schmitt, L. (2019) Shaping the lipid composition of bacterial membranes for membrane protein production. *Microb. Cell Fact. 18*, 131.
(62) Liu, X., Zhu, X., Wang, H., Liu, T., Cheng, J., and Jiang, H. (2020) Discovery and modification of cytochrome P450 for plant natural products biosynthesis. *Synth. Syst. Biotechnol. 5*, 187−199.
(63) Milder, I. E. J., Arts, I. C. W., Putte, B. v. d., Venema, D. P., and Hollman, P. C. H. (2005) Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. *Br. J. Nutr. 93*, 393−402.
(64) Zhu, J., Yan, L., Xu, X., Zhang, Y., Shi, J., Jiang, C., and Shao, D. (2018) Strategies to enhance the production of pinoresinol and its glucosides by endophytic fungus (*Phomopsis sp.* XP-8) isolated from Tu-chung bark. *AMB Express 8*, 55.
(65) Maruyama, J., Kobayashi, M., Miyashita, M., Kouno, I., and Irie, H. (1994) A synthesis of (±)-pinoresinol and its related compound using potassium persulfate ($K_2S_2O_8$) oxidation of benzoylacetates. *Heterocycles 37*, 839−845.
(66) Ricklefs, E., Girhard, M., Koschorreck, K., Smit, M. S., and Urlacher, V. B. (2015) Two-step one-pot synthesis of pinoresinol from eugenol in an enzymatic cascade. *ChemCatChem 7*, 1857−1864.
(67) Zhang, Y., Werling, U., and Edelmann, W. (2012) SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res. 40*, e55.
(68) Worsch, A., Eggimann, F. K., Girhard, M., Bühler, C. J., Tieves, F., Czaja, R., Vogel, A., Grumaz, C., Sohn, K., Lutz, S., Kittelmann, M., and Urlacher, V. B. (2018) A novel cytochrome P450 monooxygenase from *Streptomyces platensis* resembles activities of human drug metabolizing P450s. *Biotechnol. Bioeng. 115*, 2156−2166.
(69) Girhard, M., Klaus, T., Khatri, Y., Bernhardt, R., and Urlacher, V. B. (2010) Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl. Microbiol. Biotechnol. 87*, 595−607.
(70) Omura, T., and Sato, R. (1964) The carbon monoxide-binding pigment of liver microsomes I. Evidence for its hemoprotein nature. *J. Biol. Chem. 239*, 2370−2378.

https://dx.doi.org/10.1021/acssynbio.0c00354
ACS Synth. Biol. XXXX, XXX, XXX−XXX

2.2.1 Supporting Information

Supplementary materials

Table S1. Reference compounds for LC/MS and HPLC.

Substrate	**Manufacturer**	**Purity**
(±)-pinoresinol [a]	Sigma Aldrich	≥ 95% (HPLC)
(+)-pinoresinol [b]	Phytolab	≥ 95% (HPLC)
(+)-lariciresinol	Sigma Aldrich	≥ 95% (HPLC)
(±)-secoisolariciresinol	Sigma Aldrich	≥ 95% (HPLC)
(±)-matairesinol [c]	Cayman chemical	≥ 95% (HPLC)
(-)-matairesinol	Phytolab	≥ 99% (HPLC)
(+)-sesamin	TCI	> 98% (GC)

[a] 77% ee of (+)-enantiomer, [b] ≥ 96% ee of (+)-enantiomer, [c] 1% ee of (+)-enantiomer

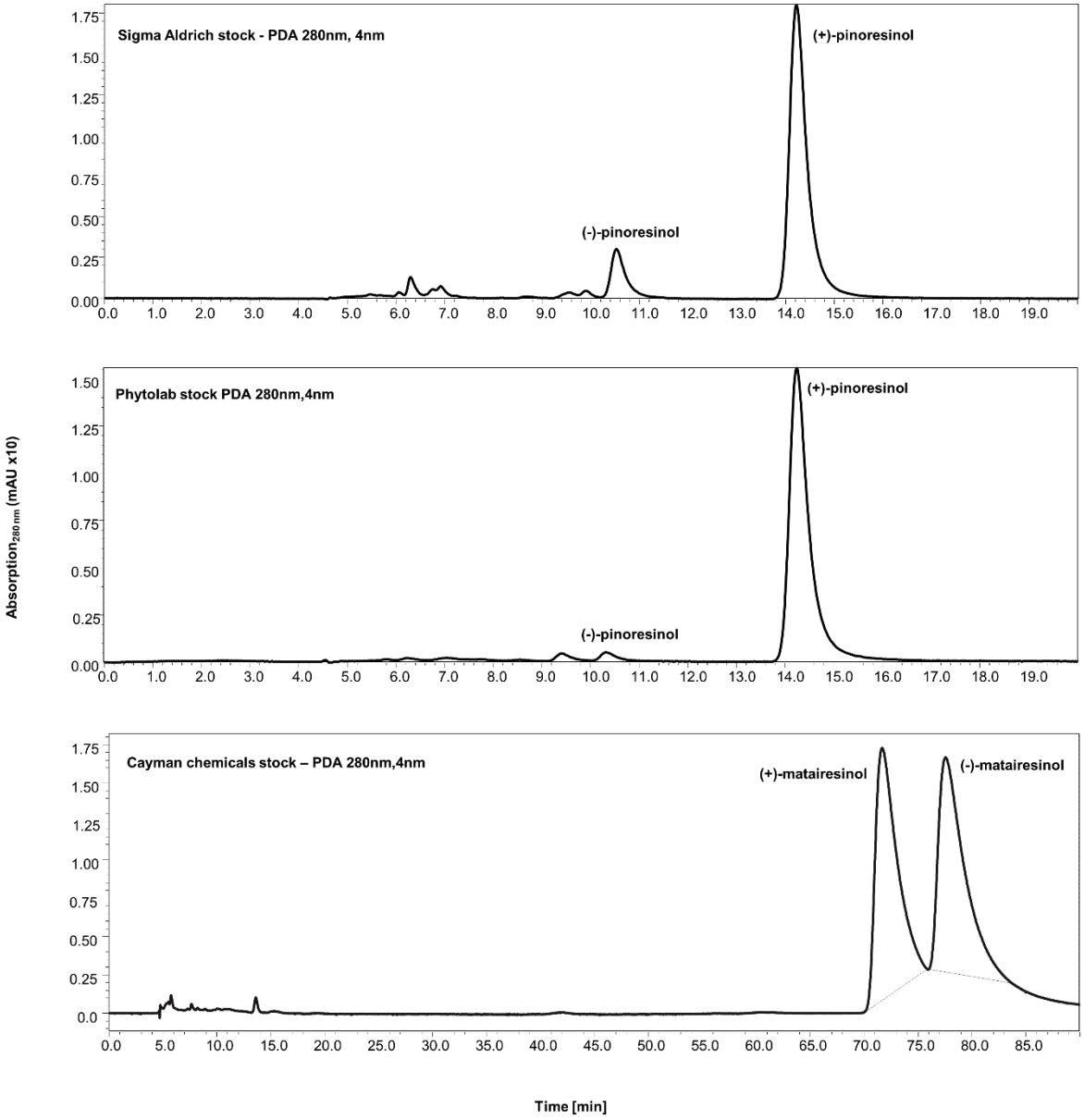


Figure S1. Chiral HPLC analysis of the reference compounds of pinoresinol and matairesinol (Table S1). Separation: Method 1 was used for pinoresinol and method 3 was used for matairesinol (Table S9).

Table S2. Overview on the genes used in this work. Sequences of codon-optimized genes and the fusion construct are listed below.

Gene	Complete name	Organism	Plasmid	MCS	Reference
fiplr	pinoresinol-lariciresinol reductase	*Forsythia intermedia*	pCDFDuet-1	2	[102]
ppsdh	secoisolariciresinol dehydrogenase	*Podophyllum pleianthum / Dysosma pleiantha*	pCDFDuet-1	1	This work
cyp719a23oc_t	pluviatolide synthase (truncated, codon optimized)	*Sinopodophyllum hexandrum*	pETDuet-1	-	This work
atr2	NADPH-cytochrome P450 reductase 2 (ATR2)	*Arabidopsis thaliana*	pETDuet-1	1	[103]
ykun	Flavodoxin YkuN	*Bacillus subtilis*	pETDuet-1	1	[104]
fdr	NADPH dependent flavodoxin reductase	*Escherichia coli JM109*			
fusion	CYP719A23-CPR	-	pET28a(+)	1	Phytowelt

fiplr (GenBank: AAC49608), codon-optimized sequence

ATGCTGATTAGCTTCAAAATGCAAGGTGCGCATCTGGTGAGTGGCTCCTTTAAAGACTTCAA
TAGCCTCGTTGAAGCGGTAAAACTGGTAGATGTGGTTATTTCGGCAATCAGTGGCGTGCATA
TCCGCAGCCATCAGATTCTGTTACAGCTGAAACTTGTGGAAGCGATCAAAGAAGCTGGTAAC
GTTAAGCGCTTCTTACCGTCTGAATTCGGGATGGATCCTGCTAAATTCATGGATACCGCTAT
GGAACCCGGGAAAGTCACGTTAGACGAAAAGATGGTGGTACGCAAAGCCATTGAGAAAGCAG
GAATCCCGTTTACCTATGTGTCAGCCAATTGCTTTGCCGGCTATTTTCTGGGAGGTCTGTGT
CAGTTTGGGAAGATCCTCCCATCACGGGATTTCGTCATCATTCATGGCGATGGGAACAAGAA
AGCGATTTACAATAACGAGGATGACATTGCAACTTACGCGATCAAAACGATCAATGATCCGC
GTACTTTGAACAAAACCATCTACATTAGCCCACCTAAGAACATCCTTAGCCAACGCGAAGTT
GTGCAGACATGGGAGAAACTGATTGGCAAAGAATTGCAGAAAATTACGTTGTCGAAGGAGGA
CTTTCTGGCATCGGTGAAGGAACTGGAGTATGCGCAGCAAGTGGGTCTGAGTCACTATCATG
ATGTCAACTACCAGGGATGCTTAACCTCCTTTGAAATTGGCGATGAAGAGGAAGCCTCTAAA
CTGTATCCGGAAGTGAAATACACCTCTGTTGAAGAATACCTTAAACGCTATGTTGGCGGTCA
TCACCATCATCACCACTGA

ppsdh (GenBank: AHB18702), codon-optimized sequence

```
ATGGGATCCACTTCTACGCCTGCAAGTTCCACTAACCGCTTACAAGACAAGGTAGCGATCAT
TACTGGAGGTGCAGGTGGCATTGGTGAGACGACGGCAAAATTGTTTATCCGTTACGGCGCTA
AAGTGGTTATAGCTGACATTTCTGATGATCATGGGCAGAAAGTCTGTAAAAATATTGGTTCT
CCGGATGTGATTTCATTTGTCCACTGCGACGTCACAAAAGATGAAGACGTTCGCAATCTTGT
TGATACCACCATTGCCAAGCATGGCAAGTTGGATATTATGTTCGGCAACGTTGGCGTATTAA
GTACGACACCGTATTCCATTCTGGAGGCGGGTAATGAGGATTTCAAACGGGTCATGGACATT
AACGTGTATGGCGCTTTCCTTGTGGCGAAGCACGCGGCACGAGTGATGATCCCAGCCAAAAA
AGGTAGTATCGTATTTACCGCTTCAATCTCATCGTTTACTGCCGGGGAAGGTGTTTCGCACG
TATATACCGCTACGAAACATGCCGTCTTAGGACTTACCACCAGCTTGTGTACCGAACTGGGC
CAGTACGGGGTTCGTGTAAACTGCGTTTCGCCGTATATCGTTGCCAGCCCCCTGCTGACAGA
CGTCTTCGGCGTGGACAGCAGCCGTGTGGAAGAACTTGCACATCAGGCGGCGAATCTGAAAG
GTACCCTTCTGCGCGCCGAAGATGTGGCCGATGCCGTGGCATACCTGGCGGGCGATGAAAGC
AAATACGTGAGCGGGCTGAACCTGGTGATCGATGGCGGTTATACGCGCACCAACCCAGCGTT
TCCGACCGCGCTGAAACATGGCCTGGCG
```

cyp719a23 (GenBank: AGC29953.1), codon-optimized sequence

```
ATGGAAATGGAAATGTCCGTTCTGGCTATGTCGTCAACCTTGATTCTCGCGTTGGCTATGGC
GCTGATCTTCCTGTTCAAAGCGAAAAGTTCCTCCGCAATCAAATGGCCTCCGGGTCCGAAAA
CGTTGCCAATCATCGGGAATCTGCACCAACTGGGTGGTGATGAACTGCATATTGTACTCGCA
AAACTTGCACGCGTGCATGGCGCGATCATGACGATTTGGATGGCCAAGAAACCGGTTATTGT
CGTATCAGACGTCAATAGCGTGTGGGAGGTGTTAGTCTCGAAATCTAGCGACTATGCAGCCC
GTGATGCAGCGGAAATCTCGAAAATTGTGTCTGCGAGCTCGCATAGTATCAACACCTCTGAC
AGTGGCCCATATTGGCAGACGCTGCGTCGTGGCCTTACGCATGGGCCTCTTGGTCCGCTGAA
CATTTCAGCTCAGATTCCGATCCAACAGCGCGATATGCAACGTGTGATTCGGGAAATGCAGC
AGGATGCTGCGGCTAATGGCGGCATTATCAAGCCACTGGATCATCTGAAACGCTCCTCTACG
CGCTTGGTTTCGCGCCTCATCTTCGGGGATACGTTTGACAACGACCCCTACAACGACTCCAT
GCACGAAGTGGTGCAAGATCTGAATCGTTTTGGCGGAATTGCCCTGTTAGAACAAGCATTCA
GCTTTGCCAAACATCTCCCCAGCTACAAACGCGGTGTGAAAGAGTTCCACATCCACAAACGG
AAAATTGACGATCTTGTTCGTCCTGTGGTAGCCAGCGCGAATCCACCGAGTAACAGCTATCT
GGGCTTTCTCCAGAGCCAGAACTATTCCGAAGAGATTATTATTGCGTGCATCTTCGAGCTGT
ACTTACTGGCGATGGATAGCTCGGCGTCAACCGCGACTTGGGCGTTAGCCTTTATGATCCGT
GATCAGCAAGTCCAGGAAAAGCTGTACCAGGACATCAAACGCGTTATTGGTGATGGAGTGGA
CCTGGTTAAAGCCGAAGATCTGAGCAAGATGCACTATCTTCAGGCAGTTGTGAAAGAAACCA
TGCGCATGAAGCCAATTGCACCGTTGGCCATTCCGCATAAAACAGCCATTGACACCACCGTC
ATGGGTACCAAAGTTCCGAAGGGTACTTGCGTGATGGTGAATCTGTACGCGTTGCACCATGA
TGAAAGTGTATGGGCGAAACCGTATACCTTCATGCCGGAACGCTTTCTGCAAGGCGAGGATG
GTAAAAGCGTCACAGAACAGGCCTTTCTGCCCTTTGGAGCTGGCATGCGCATTTGTGGCGGG
ATGGAAGTTGGCAAACTGCAGTTTTCTCTGGCACTGGCCAACCTGGTAAACGCATTTAAGTG
GACTAGTGCCGCTGAGGGTAAACTGCCTGATATGTCAGATGAGTTACAGTTCATTACCGTCA
TGAAAACACCGTTAGAAGCGCGCATCATTCCGCGTAATCCG
```

***Fusion* (*cyp719a23_oc_t-cpr*)**, CPR sequence underlined

ATGGCGAAAGCGAAAAGTTCCTCCGCAATCAAATGGCCTCCGGGTCCGAAAACGTTGCCAATCAT
CGGGAATCTGCACCAACTGGGTGGTGATGAACTGCATATTGTACTCGCAAAACTTGCACGCGTGC
ATGGCGCGATCATGACGATTTGGATGGCCAAGAAACCGGTTATTGTCGTATCAGACGTCAATAGC
GTGTGGGAGGTGTTAGTCTCGAAATCTAGCGACTATGCAGCCCGTGATGCAGCGGAAATCTCGAA
AATTGTGTCTGCGAGCTCGCATAGTATCAACACCTCTGACAGTGGCCCATATTGGCAGACGCTGC
GTCGTGGCCTTACGCATGGGCCTCTTGGTCCGCTGAACATTTCAGCTCAGATTCCGATCCAACAG
CGCGATATGCAACGTGTGATTCGGGAAATGCAGCAGGATGCTGCGGCTAATGGCGGCATTATCAA
GCCACTGGATCATCTGAAACGCTCCTCTACGCGCTTGGTTTCGCGCCTCATCTTCGGGGATACGT
TTGACAACGACCCCTACAACGACTCCATGCACGAAGTGGTGCAAGATCTGAATCGTTTTGGCGGA
ATTGCCCTGTTAGAACAAGCATTCAGCTTTGCCAAACATCTCCCCAGCTACAAACGCGGTGTGAA
AGAGTTCCACATCCACAAACGGAAAATTGACGATCTTGTTCGTCCTGTGGTAGCCAGCGCGAATC
CACCGAGTAACAGCTATCTGGGCTTTCTCCAGAGCCAGAACTATTCCGAAGAGATTATTATTGCG
TGCATCTTCGAGCTGTACTTACTGGCGATGGATAGCTCGGCGTCAACCGCGACTTGGGCGTTAGC
CTTTATGATCCGTGATCAGCAAGTCCAGGAAAAGCTGTACCAGGACATCAAACGCGTTATTGGTG
ATGGAGTGGACCTGGTTAAAGCCGAAGATCTGAGCAAGATGCACTATCTTCAGGCAGTTGTGAAA
GAAACCATGCGCATGAAGCCAATTGCACCGTTGGCCATTCCGCATAAAACAGCCATTGACACCAC
CGTCATGGGTACCAAAGTTCCGAAGGGTACTTGCGTGATGGTGAATCTGTACGCGTTGCACCATG
ATGAAAGTGTATGGGCGAAACCGTATACCTTCATGCCGGAACGCTTTCTGCAAGGCGAGGATGGT
AAAAGCGTCACAGAACAGGCCTTTCTGCCCTTTGGAGCTGGCATGCGCATTTGTGGCGGGATGGA
AGTTGGCAAACTGCAGTTTTCTCTGGCACTGGCCAACCTGGTAAACGCATTTAAGTGGACTAGTG
CCGCTGAGGGTAAACTGCCTGATATGTCAGATGAGTTACAGTTCATTACCGTCATGAAAACACCG
TTAGAAGCGCGCATCATTCCGCGTAATCCGATGTCCGGTTCTGGGAATTCAAAACGTGTCGAGCC
TCTTAAGCCTTTGGTTATTAAGCCTCGTGAGGAAGAGATTGATGATGGGCGTAAGAAAGTTACCA
TCTTTTTCGGTACACAAACTGGTACTGCTGAAGGTTTTGCAAAGGCTTTAGGAGAAGAAGCTAAA
GCAAGATATGAAAAGACCAGATTCAAAATCGTTGATTTGGATGATTACGCGGCTGATGATGATGA
GTATGAGGAGAAATTGAAGAAAGAGGATGTGGCTTTCTTCTTCTTAGCCACATATGGAGATGGTG
AGCCTACCGACAATGCAGCGAGATTCTACAAATGGTTCACCGAGGGGAATGACAGAGGAGAATGG
CTTAAGAACTTGAAGTATGGAGTGTTTGGATTAGGAAACAGACAATATGAGCATTTTAATAAGGT
TGCCAAAGTTGTAGATGACATTCTTGTCGAACAAGGTGCACAGCGTCTTGTACAAGTTGGTCTTG
GAGATGATGACCAGTGTATTGAAGATGACTTTACCGCTTGGCGAGAAGCATTGTGGCCCGAGCTT
GATACAATACTGAGGGAAGAAGGGGATACAGCTGTTGCCACACCATACACTGCAGCTGTGTTAGA
ATACAGAGTTTCTATTCACGACTCTGAAGATGCCAAATTCAATGATATAAACATGGCAAATGGGA
ATGGTTACACTGTGTTTGATGCTCAACATCCTTACAAAGCAAATGTCGCTGTTAAAAGGGAGCTT
CATACTCCCGAGTCTGATCGTTCTTGTATCCATTTGGAATTTGACATTGCTGGAAGTGGACTTAC
GTATGAAACTGGAGATCATGTTGGTGTACTTTGTGATAACTTAAGTGAAACTGTAGATGAAGCTC
TTAGATTGCTGGATATGTCACCTGATACTTATTTCTCACTTCACGCTGAAAAAGAAGACGGCACA
CCAATCAGCAGCTCACTGCCTCCTCCCTTCCCACCTTGCAACTTGAGAACAGCGCTTACACGATA
TGCATGTCTTTTGAGTTCTCCAAAGAAGTCTGCTTTAGTTGCGTTGGCTGCTCATGCATCTGATC
CTACCGAAGCAGAACGATTAAAACACCTTGCTTCACCTGCTGGAAAGGATGAATATTCAAAGTGG
GTAGTAGAGAGTCAAAGAAGTCTACTTGAGGTGATGGCCGAGTTTCCTTCAGCCAAGCCACCACT
TGGTGTCTTCTTCGCTGGAGTTGCTCCAAGGTTGCAGCCTAGGTTCTATTCGATATCATCATCGC
CCAAGATTGCTGAAACTAGAATTCACGTCACATGTGCACTGGTTTATGAGAAAATGCCAACTGGC
AGGATTCATAAGGGAGTGTGTTCCACTTGGATGAAGAATGCTGTGCCTTACGAGAAGAGTGAAAA
CTGTTCCTCGGCGCCGATATTTGTTAGGCAATCCAACTTCAAGCTTCCTTCTGATTCTAAGGTAC
CGATCATCATGATCGGTCCAGGGACTGGATTAGCTCCATTCAGAGGATTCCTTCAGGAAAGACTA
GCGTTGGTAGAATCTGGTGTTGAACTTGGGCCATCAGTTTTGTTCTTTGGATGCAGAAACCGTAG
AATGGATTTCATCTACGAGGAAGAGCTCCAGCGATTTGTTGAGACAGGCGCCCTTTCTGAGTTGA
TTGTAGCCTTCTCCCGTGAGGGTCCAACTAAACAATATGTGCAACATAAAATGACTGAGAAAGCA
ACGGAACTTTGGAATATCATCTCCCAAGGCGGATATGTATACGTGTGTGGAGATGCTAAGGGCAT
GGCTAGAGATGTTCACAGAGTTCTTCATACTATTGCTCAAGAGCAGGGAGGTATGGATAGCTCCA
AAACAGAAAGCTTCATCAAGAGTTTGCAAATGGAAGGGAGATACCTAAGAGATGTATGG

Table S3. Plasmids used in this work.

Plasmid	Features	Copy number	Reference
pCDFDuet_*ppsdh_fiplr*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20-40	This work
pETDuet_*atr2_cyp719a23oc_t*	ColE1, P_{T7}, *lacI,* Amp^R	~40	This work
pETDuet_*ykun_fdr_cyp719a23oc_t*	ColE1, P_{T7}, *lacI,* Amp^R	~40	This work
pET28a_*cyp719a23_t_cpr**	ColE1, P_{T7}, *lacI,* Kan^R	~40	This work
pET28a_*cyp719a23oc_t* [a]	ColE1, P_{T7}, *lacI,* Kan^R	~40	This work
pET22a_*atr2* [a]	ColE1, P_{T7}, *lacI,* Amp^R	~40	103
pET16b-*ykuN*(N10His) [a]	ColE1, P_{T7}, *lacI,* Amp^R	~40	105
pET16b-*fdr*(N6His) [a]	ColE1, P_{T7}, *lacI,* Amp^R	~40	105
pET22b_*gdh* [a]	ColE1, P_{T7}, *lacI,* Amp^R	~40	106

[a] Plasmids used for expression of single enzymes to reconstitute *cyp719a23* activity *in vitro*

cyp719a23 variants and oligonucleotides

Table S4. Names and features for the generated *cyp719a23* gene variants.

Variant	Original sequence	Inserted modification (amino acids)	Truncated amino acids
cyp719a23nat_bov	Native	Ala + N-terminal sequence from bovine CYP17A1 (ALLLAVFL)	1-14
cyp719a23nat_rab	Native	Ala + N-terminal sequence from rabbit CYP2C3 (AKKTSSKGK)	1-35
cyp719a23nat_tr	Native	Ala	1-26
cyp719a23oc_bov	Codon optimized	N-terminal sequence from bovine CYP17A1(ALLLAVFL)	1-14
cyp719a23oc_rab	Codon optimized	N-terminal sequence from rabbit CYP2C3 (KKTSSKGK)	1-35
cyp719a23oc_tr	Codon optimized	Ala	1-26

Table S5. SLIC primers for *cyp719a23* gene manipulations.

Primer name	Oligonucleotide sequence (5' → 3') [a]
cyp719a23nat_bov_fwd	AGTTAAGTATAAGAAGGAGATATACATATGGCATGGCTCT GTTATTAGCAGTTTTT***TTAGCCTTGGCTATGGCACTG***
cyp719a23nat_bov_rev	CAGCAGCGGTTTCTTTACCAGACTCGAG***TCAAGGATTGCG AGGAATGATCC***
cyp719a23nat_rab_fwd	TTAAGTATAAGAAGGAGATATACATATGGCATGGCTAAGA AAACGAGCTCTAAAGGGAAG***AAATGGCCTCCAGGGCCAAA AACATTACCC***
cyp719a23nat_rab_rev	CAGCAGCGGTTTCTTTACCAGACTCGAG***TCAAGGATTGCG AGGAATGATCC***
cyp719a23nat_tr_fwd	AGTTAAGTATAAGAAGGAGATATACATATGGCG***AAAGCAA AGTCGTCTTCTGCAATTAAATGGCCTCC***
cyp719a23nat_tr_rev	CAGCAGCGGTTTCTTTACCAGACTCGAG***TCAAGGATTGCG AGGAATGATCC***
cyp719a23oc_bov_fwd	AGTTAAGTATAAGAAGGAGATATACATATGGCATGGCTCT GTTATTAGCAGTTTTT***CTCGCGTTGGCTATGGCGCTGATC TTCCTG***
cyp719a23oc_bov_rev	CAGCAGCGGTTTCTTTACCAGACTCGAGTTACGGATTACG CGGAATGATGC
cyp719a23oc_rab_fwd	TTAAGTATAAGAAGGAGATATACATATGGCATGGCTAAGA AAACGAGCTCTAAAGGGAAG***TGGCCTCCGGGTCCGAAAAC GTTGCC***
cyp719a23oc_rab_rev	CAGCAGCGGTTTCTTTACCAGACTCGAGTTACGGATTACG CGGAATGATGC
cyp719a23oc_tr_fwd	AGTTAAGTATAAGAAGGAGATATACATATGGCG***AAAGCGA AAAGTTCCTCCGCAATCAAATGGCC***
cyp719a23oc_tr_rev	CAGCAGCGGTTTCTTTACCAGACTCGAGTTACGGATTACG CGGAATGATGC

[a] Underlined: Complementary sequence to the vector pETDuet-1. Normal: Sequence of N-terminal modification to be inserted. ***Bold-italics***: Complementary sequence to the corresponding genes.

Screening of CYP719A23 variants

The expression was induced at OD_{600} of 0.6 by the addition of 0.5 mM IPTG, with 0.5 mM 5-aminolevulinic acid and 0.1 mM $FeSO_4$ as supplements, and it was carried out at 20 °C, 140 rpm for 16 – 21 h. From the expression of CYP719A23 variants, cells were harvested by centrifugation (3.220 x g, 4°C, 20 min) and resuspended in Potassium phosphate buffer (80% 50 mM K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5, 100 µM PMSF). Cell suspension was supplemented with 0.5 % Triton X-100 before cell disruption by sonication. P450 concentration was determined in the disrupted cell suspension by CO difference spectroscopy according to reference [107].

Expressions of truncated CYP719A23 were incubated for 44 h with temperatures of 20°C - 30°C. Cells were harvested and resuspended as described above, except that no Triton X-100 was added. The soluble protein fraction was recovered by centrifugation (12,300 x g, 30 min, 4 °C) and CO-difference spectroscopy was performed with the soluble protein fraction only. Results for initial and optimized expression conditions are shown in Tables S11 - S12.

Analytical methods

Table S6. LC/MS method. Solvent A methanol, solvent B ddH_2O + 0.1% formic acid.

Time [min]	ddH_2O + 0.1% formic acid [%]
0.01	80
5.00	65
10.00	65
20.00	38
25.00	0 (100% MeOH)
26.00	0 (100% MeOH)
26.01	80
35.00	80

Table S7. Equations used for LC/MS data analysis

Conversion [%]	$= \Sigma(P_{area}) / \Sigma(S_{area}+P_{area}) * 100$
Product distribution [%]	$= P_{area} / \Sigma(S_{area}+P_{area}) * 100$

Table S8. Overview on compounds observed features. Commercially available substances were used as references. (-)-pluviatolide was identified via the typical fragmentation pattern and the elution order as described elsewhere.[70, 80] *The most common observed retention time is stated for the compounds.

	(+)-pinoresinol	**(+)-lariciresinol**	**(-)-secoiso-lariciresinol**	**(-)-matai-resinol**	**(-)-pluviatolide**
Molecular weight	358.39	360.61	362.41	358.38	356
Retention Time [min]*	19	17	18 (17.5)	20	24
Characteristic m/z fragments	359 [M+H$^+$] 341 [M+H-H_2O]	220 [M-$C_7H_7O_2^-$]	345 [M+H-H_2O] 327 [M-2OH^-] 363 [M+H$^+$]	359 [M+H$^+$] 341 [M+H-H_2O] 376 [M+H_3O^+]	357 [M+H$^+$] 339 [M+H-H_2O] 375 [M+H_3O^+]

Table S9. Chiral HPLC methods used in this study, isocratic elution.

Method number	***n*-hexane [%]**	**ethanol [%]**	**Run time [min]**
1	50	50	20
2	80	20	45
3	90	10	90

Table S10. Equation used for chiral HPLC data analysis

$$\text{Enantiomeric excess (ee) [\%]} = (E1_{area} - E2_{area}) / (E1_{area} + E2_{area})$$

Table S11. Preparative HPLC method used for (-)-pluviatolide isolation. Solvent A *n*-heptane, solvent B ethanol.

Time [min]	**ethanol [%]**
0.01	10
8.00	25
8.01	90
9.50	90
9.51	10
13.00	10

Supplementary results

Table S12. Overview on the expression levels achieved for CYP719A23 variants in *E. coli* C41(DE3).

Variant	Inserted modification	Truncated amino acids	P450 concentration	
			[µg/g_{cww}]	[nmol/l]
cyp719a23nat	-	-	nd	nd
cyp719a23nat_bov	Ala + N-terminal sequence from bovine CYP17A1	1-14	374 ± 71	227 ± 43
cyp719a23nat_rab	Ala + N-terminal sequence from rabbit CYP2C3	1-35	907 ± 55	541 ± 33
cyp719a23nat_tr	Ala	1-26	1012 ± 41	633 ± 26
cyp719a23oc	-	-	nd	nd
cyp719a23oc_bov	Ala + N-terminal sequence from bovine CYP17A1	1-14	464 ± 190	279 ± 114
cyp719a23oc_rab	Ala + N-terminal sequence from rabbit CYP2C3	1-35	1267 ± 67	763 ± 40.5
cyp719a23oc_tr	Ala	1-26	1457 ± 120	827 ± 68

nd = not detectable

Table S13. Overview of expression optimization for *cyp719a23oc_t*. Despite the highest P450 concentration was detected at 20°C, 25°C was chosen as compromise, given the expression conditions of fiPLR and ppSDH reported elsewhere.[102, 108]

	P450 concentration			
	[µg/gcww]		[nmol/l]	
Temperature	**20h**	**44h**	**20h**	**44h**
20°C	350 ± 54	673 ± 5	256 ± 4	664 ± 5
25°C	573 ± 12	499 ± 8	333 ± 7	442 ± 7
30°C	49 ± 1	39 ± 4	27 ± 1	28 ± 3

FiPLR, PpSDH and CYP719A23 activities and enantioselectivities

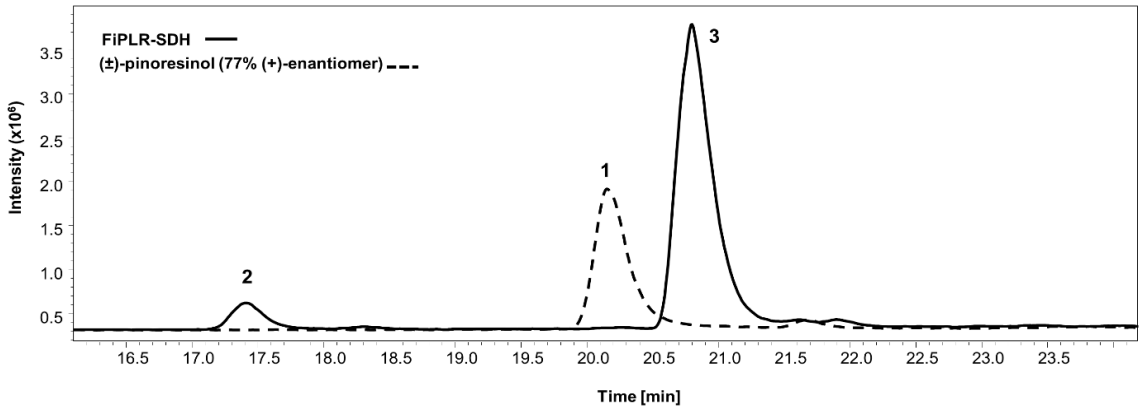


Figure S2. LC/MS analysis of (±)-pinoresinol (77% ee of (+)-enantiomer) biotransformation to (-)-matairesinol. 1: (±)-pinoresinol RT, 19.7 min, 359 [$M+H^+$] 341[$M+H-H_2O$]. 2: (-)-lariciresinol, RT 17.4 min, 220 [$M-C_7H_7O_2-OH^-$]. 3: (-)-matairesinol, RT 20.7 min, 359 [$M+H^+$] 341 [$M+H-H_2O$] 376 [$M+H_3O^+$].

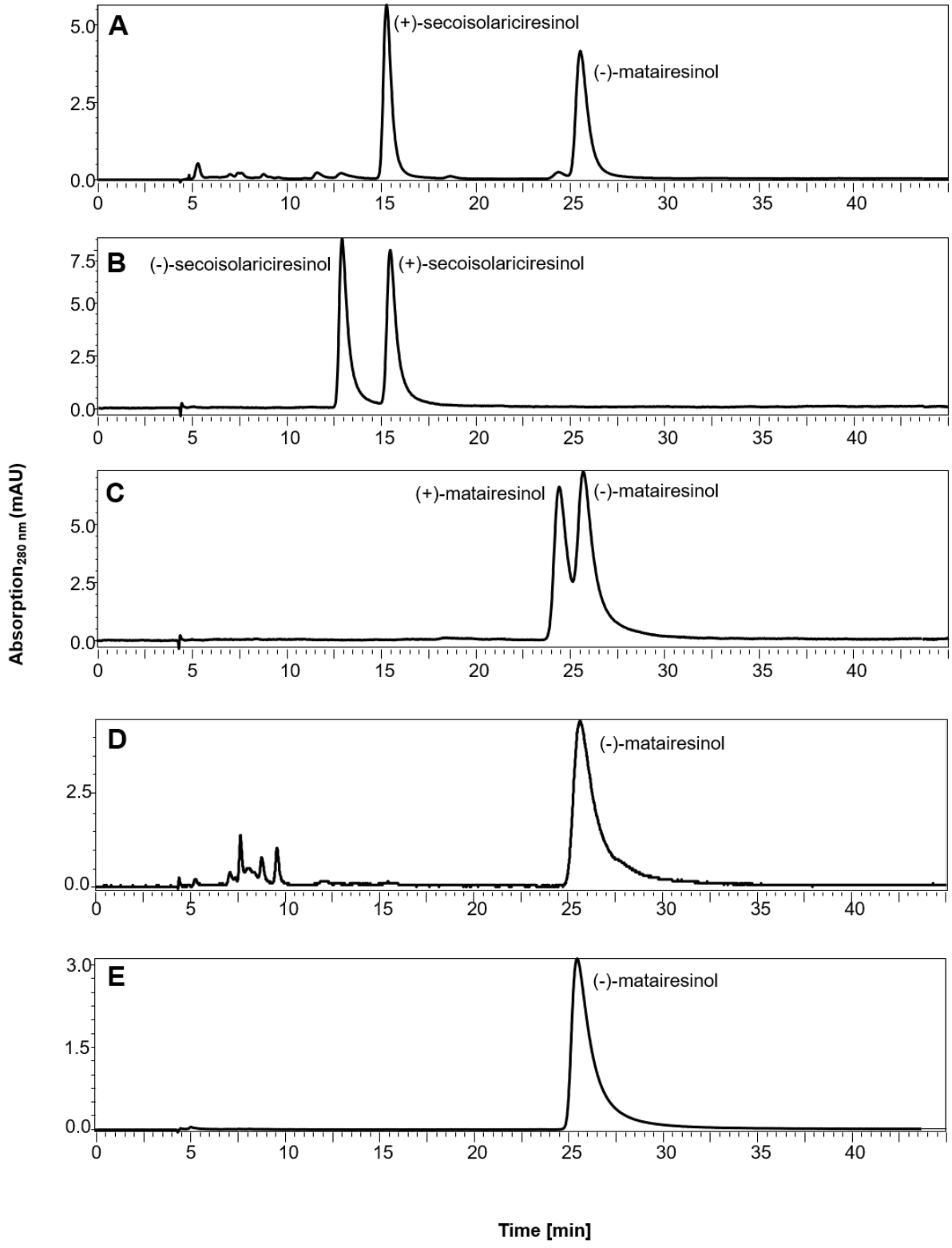


Figure S3. Chiral HPLC analysis of (-)-matairesinol synthesis. A: SDH activity on (±)-secoisolariciresinol. B: Reference for (±)-secoisolariciresinol. C: Reference for (±)-matairesinol. D: Activity of *E. coli* co-expressing FiPLR–SDH. E: Reference for (-)-matairesinol. All substances were analysed with method 2 (Table S9).

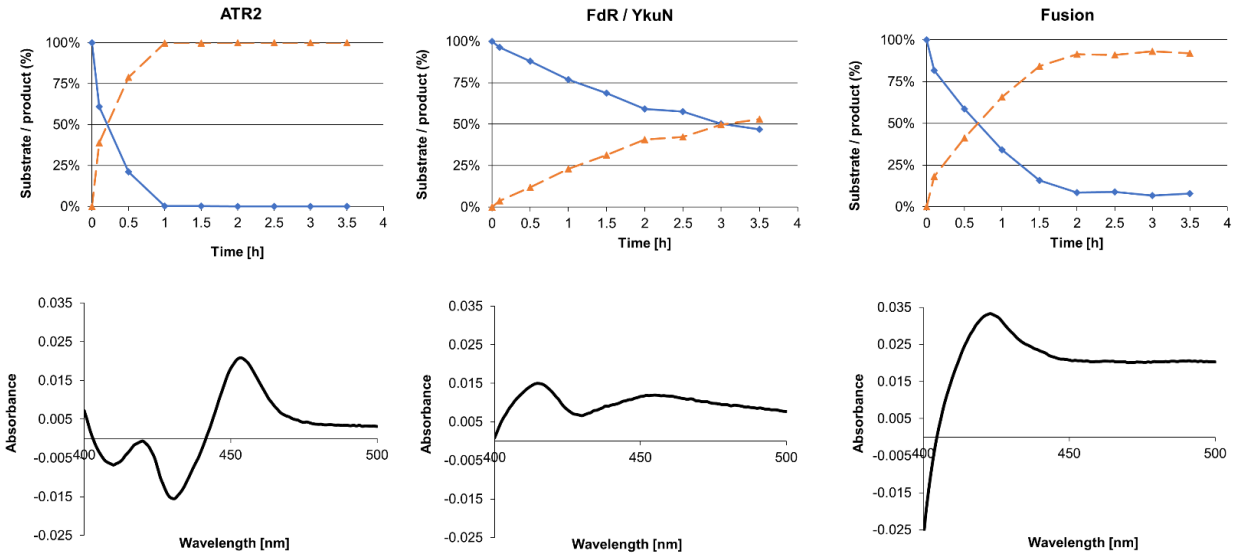


Figure S4. Effect of redox partners on (-)-matairesinol depletion and P450 concentration. Upper pictures show (-)-matairesinol depletion (blue, diamond, solid line) and (-)-pluviatolide formation (orange, triangle, dashed line). Bottom pictures show CO-difference spectra measured within the corresponding experiment.

Table S14. (-)-matairesinol conversion compared to detected P450 concentration. Co-expression of CYP719A23 and suitable redox partners.

Combination	~50% conversion time	~100% conversion time	P450 concentration [µg/g_{cww}]
P450 + ATR2	~15 min	1 h > 99%	170 ± 40
P450 + FdR/YkuN	3 h	24 h – 82%	nd
P450-CPR fusion	~45 min	3 h – 93%	nd

nd = not detectable: P450 concentration below detection limit.

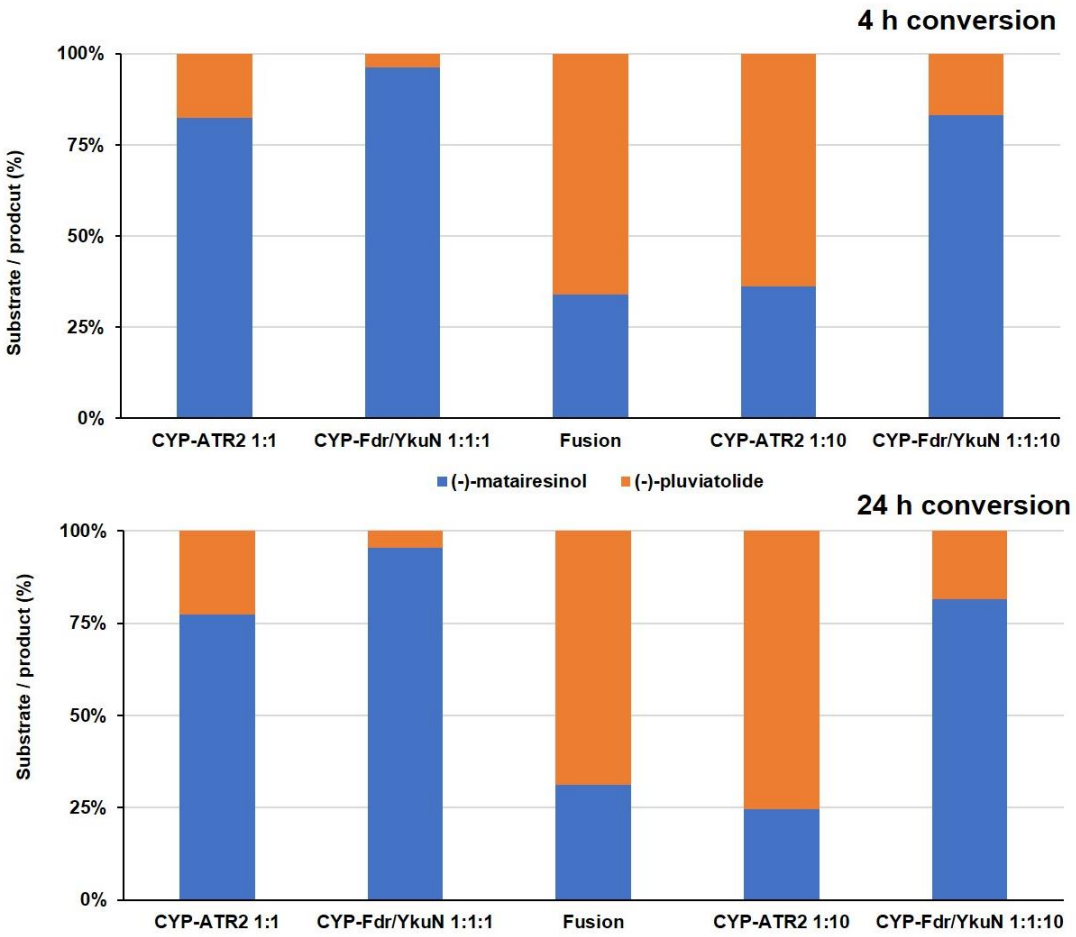


Figure S5. *In vitro* activity of CYP719A23 supported by different redox partners' combinations. Different ratios of CYP719A23-redox partners are compared to the fusion construct.

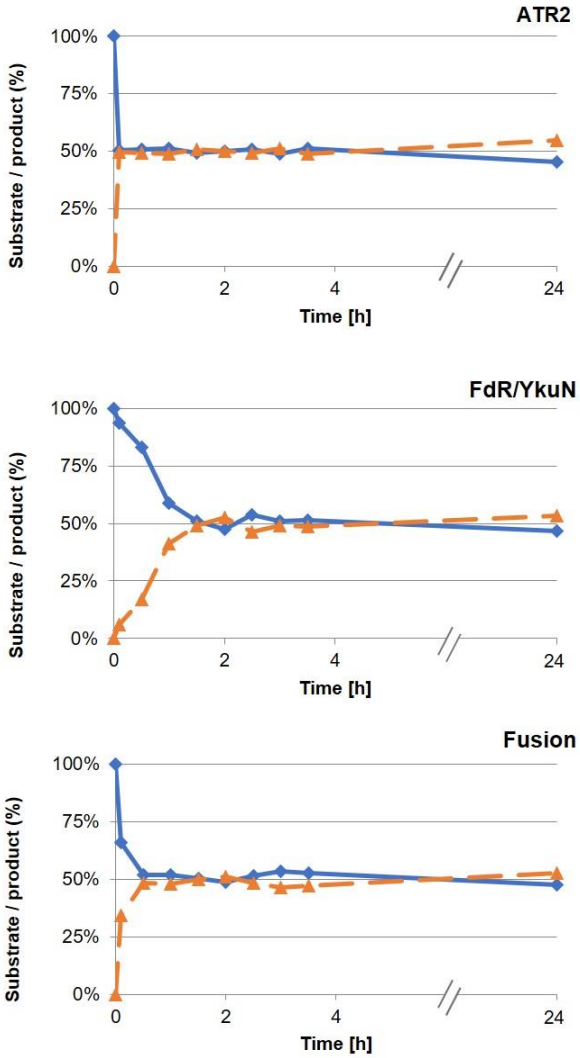


Figure S6. Depletion of racemic matairesinol (blue, diamond, solid line) and detected (-)-pluviatolide (orange, triangle, dashed line) for biotransformation of (±)-matairesinol in the final *E. coli* strain. Top to bottom: Biotransformation with co-expressed P450 and ATR2, co-expressed P450, FdR/YkuN, and P450-CPR fusion.

Table S15. (-)-Matairesinol conversion compared to detected P450 concentration. FiPLR, ppSDH, CYP719A23-redox partners were co-expressed together. (+)-pinoresinol was used as substrate.

Combination	~50% conversion time	~100% conversion time	P450 concentration [µg/g_{cww}]
P450 + ATR2	~20 min	2 h > 99%	190 ± 70
P450 + FdR/YkuN	3.5 h	24 h – 99%	nd*
P450-CPR fusion	2 h	3 h – 90%	nd*

nd = not detectable: P450 concentration below detection limit.

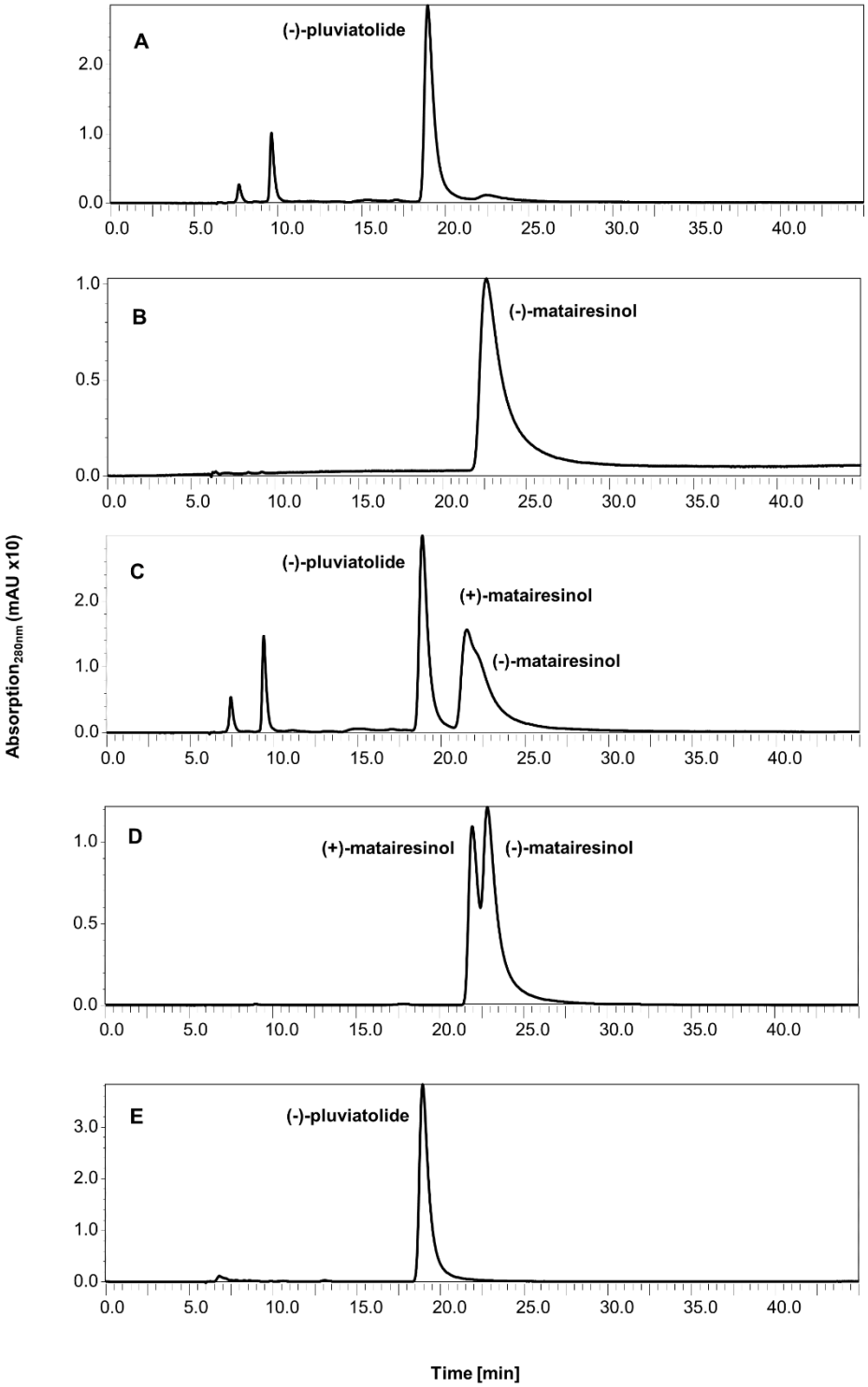


Figure S7. Chiral HPLC analysis of (-)-matairesinol and (±)-matairesinol conversion catalyzed by CYP719A23. A: Conversion of 100 µM (-)-matairesinol (-)-pluviatolide; B: reference for (-)-matairesinol, 100 µM; C: (±)-matairesinol conversion to (-)-pluviatolide; D: reference for (±)-matairesinol; E: isolated (-)-pluviatolide. Conversions were carried in 1 mL Eppendorf tubes with open lids, at 25°C, 11500 rpm, 4 h reaction time. All substances were analysed with method 2 (Table S9).

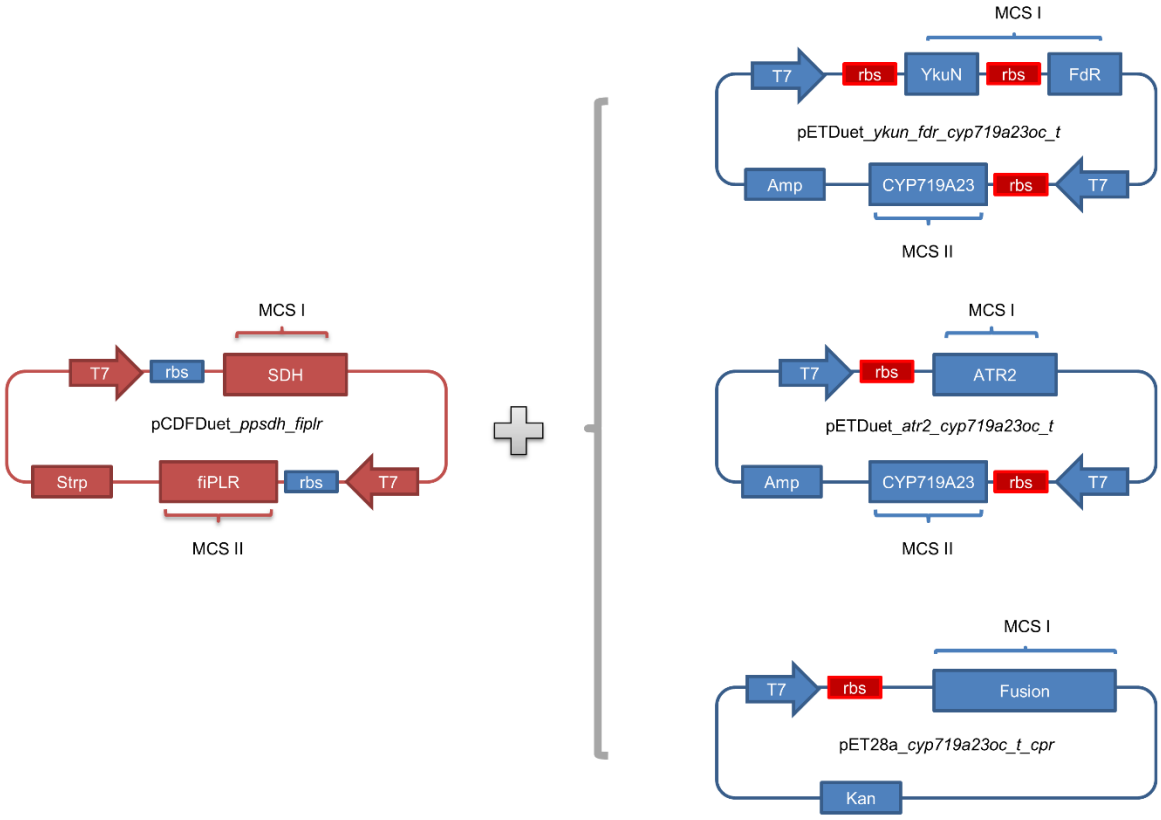


Figure S8. Overview on the modular two-plasmid system for enzyme co-expression in *E. coli*. pCDFDuet harboring *ppsdh* and *fiplr* genes was combined with the corresponding pET-vector harboring the desired cyp719a23 gene variant and redox partner genes, *atr2*, *ykun*, *fdr*, or the *cyp719a23ac-t-cpr-fusion*. T7: T7 promoter; rbs: ribosome binding site; Amp: ampicillin resistance gene; Strp: streptomycin resistance gene; Kan: kanamycin resistance gene. MCS I/II: multiple cloning site I or II.

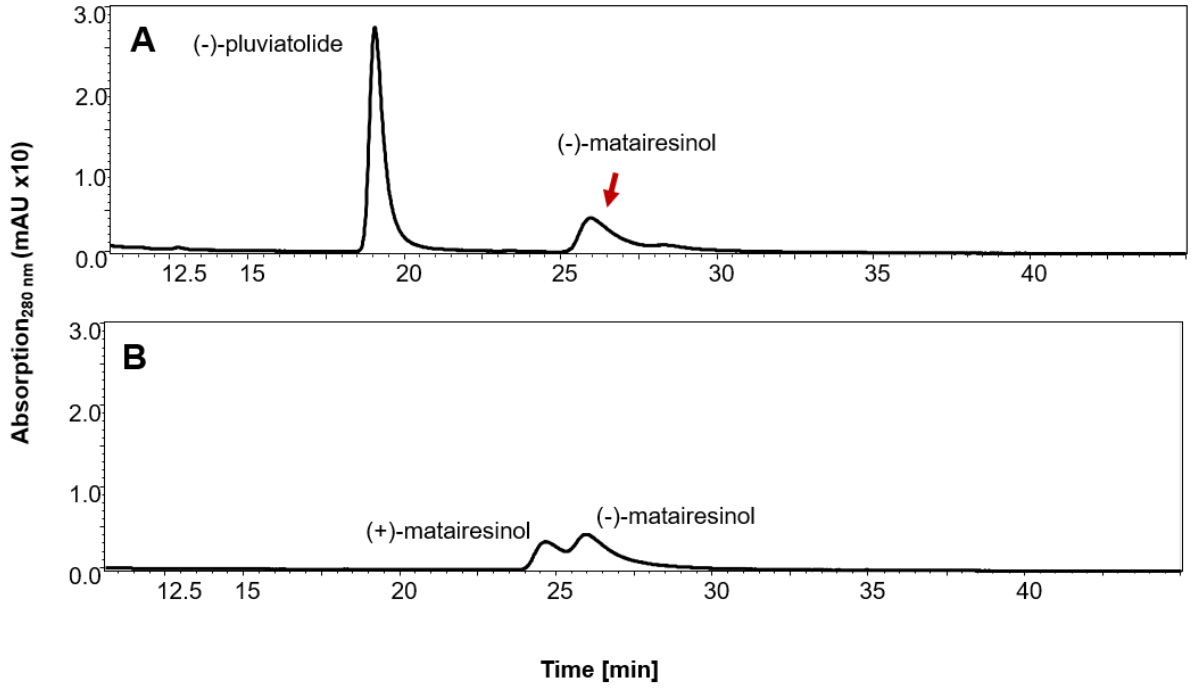


Figure S9. Chiral HPLC analysis from upscaling experiment (non-optimized trials). Whole-cell biotransformation of 500 µM (+)-pinoresinol, 25°C, 180 rpm in 10 mL reaction buffer, 24 h reaction time. A: Co-expression of FiPLR, PpSDH, CYP719A23 and ATR2; B: Reference compound of (±)-matairesinol. Sample was analysed with method 2 (Table S9).

NMR-data

Disclaimer on (-)-pluviatolide availability

Reference compound of (-)-pluviatolide (CAS No. 28115-68-6) was purchased from ChemFaces (China). According to the supplier, these stocks are extracted from the plant *Macaranga tanarius*, from which the only reported lignan among the extracted secondary metabolites is (+)-pinoresinol diglucoside [109].

The LC/MS analysis on the compound revealed a major m/z fragment of 357 [$M+H^+$] relatable to (-)-pluviatolide, however the observed retention time overlaps completely with (+)-pinoresinol (m/z fragment 341 [$M+H-H_2O$]), dramatically differing from the retention time for (-)-pluviatolide reported in the literature and observed in this work (Table S8, Figure S9).

On the other hand the ^{13}C-NMR data provided by the chemical supplier are consistent with the values recorded in this work for the desired product (Table S17). It is thus possible that a diastereomer of (-)-pluviatolide was isolated and sold instead of the claimed compound.

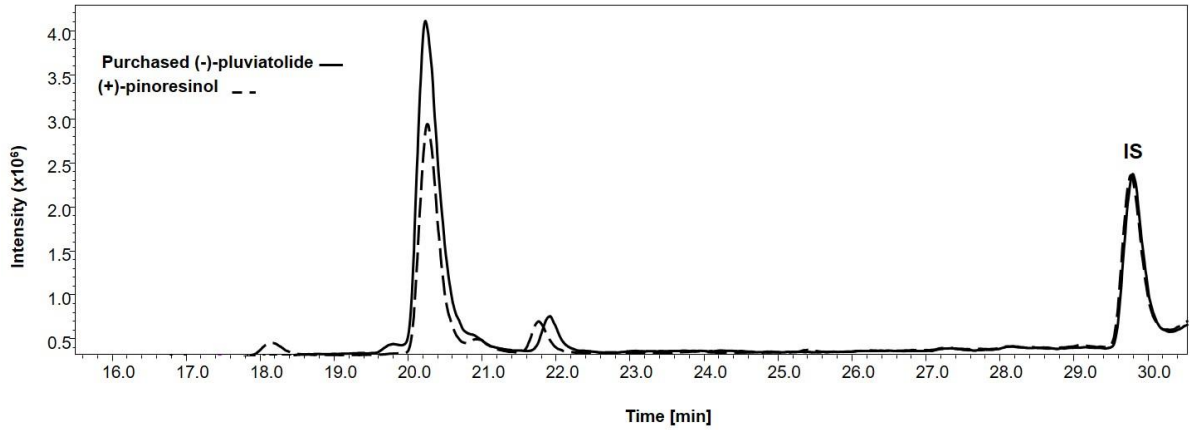


Figure S10. LC/MS analysis of (+)-pinoresinol (Table S1) and purchased (-)-pluviatolide stock (extracted from *Macaranga tanarius*). IS: internal standard (+)-sesamin.

NMR analysis – ^{1}H NMR (-)-pluviatolide ($CDCl_3$):

δ 6.84 (d, *J* = 8.1 Hz, 1H), 6.69 (d, *J* = 7.7 Hz, 1H), 6.67 (d, *J* = 1.9 Hz, 1H), 6.63 (dd, *J* = 8.1, 1.9 Hz, 1H), 6.48 – 6.42 (m, 2H), 5.93 (d, *J* = 1.5 Hz, 1H), 5.93 (d, *J* = 1.4 Hz, 1H), 4.11 (dd, *J* = 9.2, 6.9 Hz, 1H), 3.88 – 3.85 (m, 1H), 3.85 (s), 2.96 (dd, *J* = 14.1, 5.1 Hz, 1H), 2.89 (dd, *J* = 14.1, 7.0 Hz, 1H), 2.62 – 2.46 (m, 4H).

Table S16. Comparison of ^{1}H spectroscopic assignments from literature references and the observed δ ^{1}H for isolated (-)-pluviatolide. Data was recorded on a Brucker Avance 600 MHz spectrometer with $CDCl_3$ as solvent.

Carbon number	**Observed δ ^{1}H (ppm)**	**δ ^{1}H (ppm) [80]**	**δ ^{1}H (ppm) [70]**	**δ ^{1}H (ppm) [110]**
2, 3, 6 2', 5, 6'	6.42 – 6.48 (m, 2H) 6.63 (dd, *J* = 8.1, 1.9 Hz, 1H) 6.67 (d, *J* = 1.9 Hz, 1H) 6.69 (d, *J* = 7.7 Hz, 1H) 6.84 (d, *J* = 8.1 Hz, 1H)	6.44 – 6.47 (m) 6.63 (dd, *J* = 7.9, 1.8 Hz) 6.67 (d, *J* = 1.8 Hz) 6.69 (d, *J* = 7.7 Hz) 6.84 (d, *J* = 8 Hz)	6.41 – 6.48 (m, 2H) 6.58 – 6.71 (m, 3H) 6.83 (d, *J* = 8.0 Hz, 1H)	6.44 – 6.47 (m, 2H) 6.63 (dd, *J* = 7.9, 1.8 Hz, 1H) 6.67 (d, *J* = 1.8 Hz, 1H) 6.69 (d, *J* =7.7 Hz, 1H) 6.84 (d *J* = 8 Hz, 1H)
O-CH_2-O	5.93 (d, *J* = 1.5 Hz, 1H) 5.93 (d, *J* = 1.4 Hz, 1H)	5.93 (d, *J* = 1.4 Hz) 5.94 (d, *J* = 1.4 Hz)	5.92 (s, 2H)	5.93 (d, *J* = 1.4 Hz, 1H) 5.94 (d, *J* = 1.4 Hz, 1H)
-OCH_3	3.85 (s)	3.85 (s)	3.84 (s, 3H)	3.85 (s, 3H)
7, 8 ,8'	2.46 – 2.62 (m, 4H)	2.45 – 2.62 (m)	2.39 – 2.65 (m, 4H)	2.45 – 2.62 (m, 4H)
7'	2.89 (dd, *J* = 14.1, 7.0 Hz, 1H) 2.96 (dd, *J* = 14.1, 5.1 Hz, 1H)	2.89 (dd, *J* = 14.1, 7.0 Hz) 2.96 (dd, *J* = 14.0, 5.2 Hz)	2.88 (dd, *J* = 14.1, 6.8 Hz, 1H) 2.95 (dd, *J* =14.0, 5.0 Hz, 1H)	2.89 (dd, *J* = 14.1, 7.0 Hz, 1H) 2.96 (dd, *J* = 14.0 Hz, 5.2, 1H)
9	3.88 – 3.85 (m, 1H) 4.11 (dd, *J* = 9.2, 6.9 Hz, 1H)	3.86 (dd, *J* =7.4, 9.1 Hz) 4.11 (dd, *J* =7.1, 9.2 Hz)	3.85 – 3.91 (m, 1H) 4.06 – 4.15 (m, 1H)	3.86 (dd, *J* = 9.1, 7.4 Hz, 1H) 4.11 (dd, *J* = 9.2, 7.1 Hz, 1H)

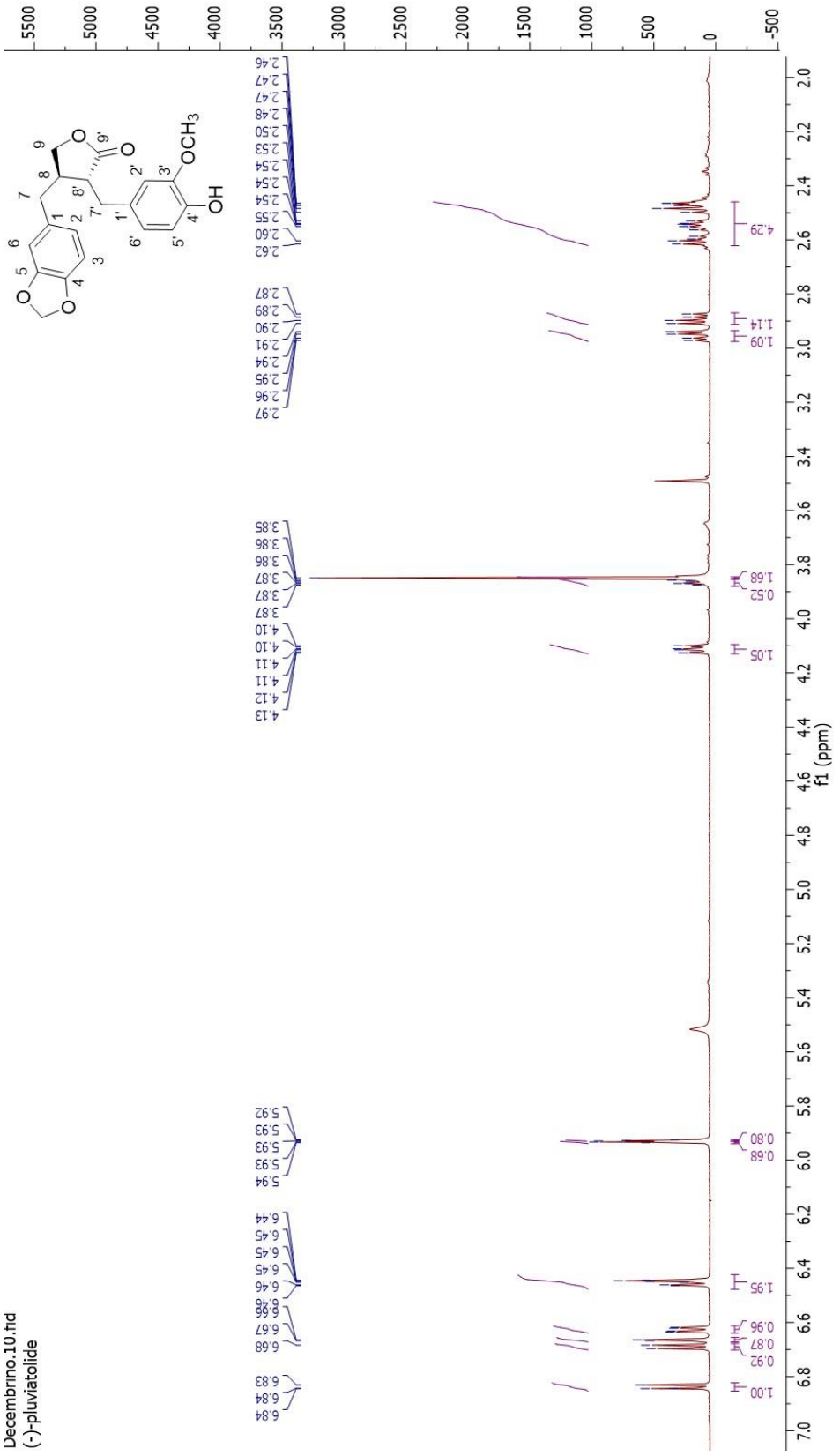


Figure S11. ^{1}H NMR spectrum of (-)-pluviatolide. Data was recorded on a Brucker Avance 600 MHz spectrometer with (-)-pluviatolide in $CDCl_3$

NMR analysis – ^{13}C NMR (-)-pluviatolide ($CDCl_3$):

δ 178.80, 148.02, 146.82, 146.49, 144.69, 131.75, 129.59, 122.23, 121.71, 114.40, 111.67, 108.96, 108.47, 101.19, 71.35, 56.04, 46.76, 41.16, 38.47, 34.78.

Table S17. Comparison of ^{13}C spectroscopic assignments from literature reference and the observed $\delta^{13}C$ for isolated (-)-pluviatolide. Data was recorded on a Brucker Avance 600 MHz spectrometer using $CDCl_3$ as solvent.

Carbon number	Observed δ ^{13}C (ppm)	δ ^{13}C [80]
1’	129.59	129.43
2’	111.67	111.48
3’	146.49	146.32
4’	144.69	144.52
5’	114.40	114.22
6’	121.71	121.55
7’	34.78	34.62
8’	46.76	46.59
9’	178.80	178.64
1	131.75	131.59
2	108.96	108.79
3	148.02	147.85
4	146.82	146.65
5	108.47	108.31
6	122.23	122.07
7	38.47	38.30
8	41.16	41.00
9	71.35	71.19
O-CH_2-O	101.19	101.04
-OCH_3	56.04	55.87

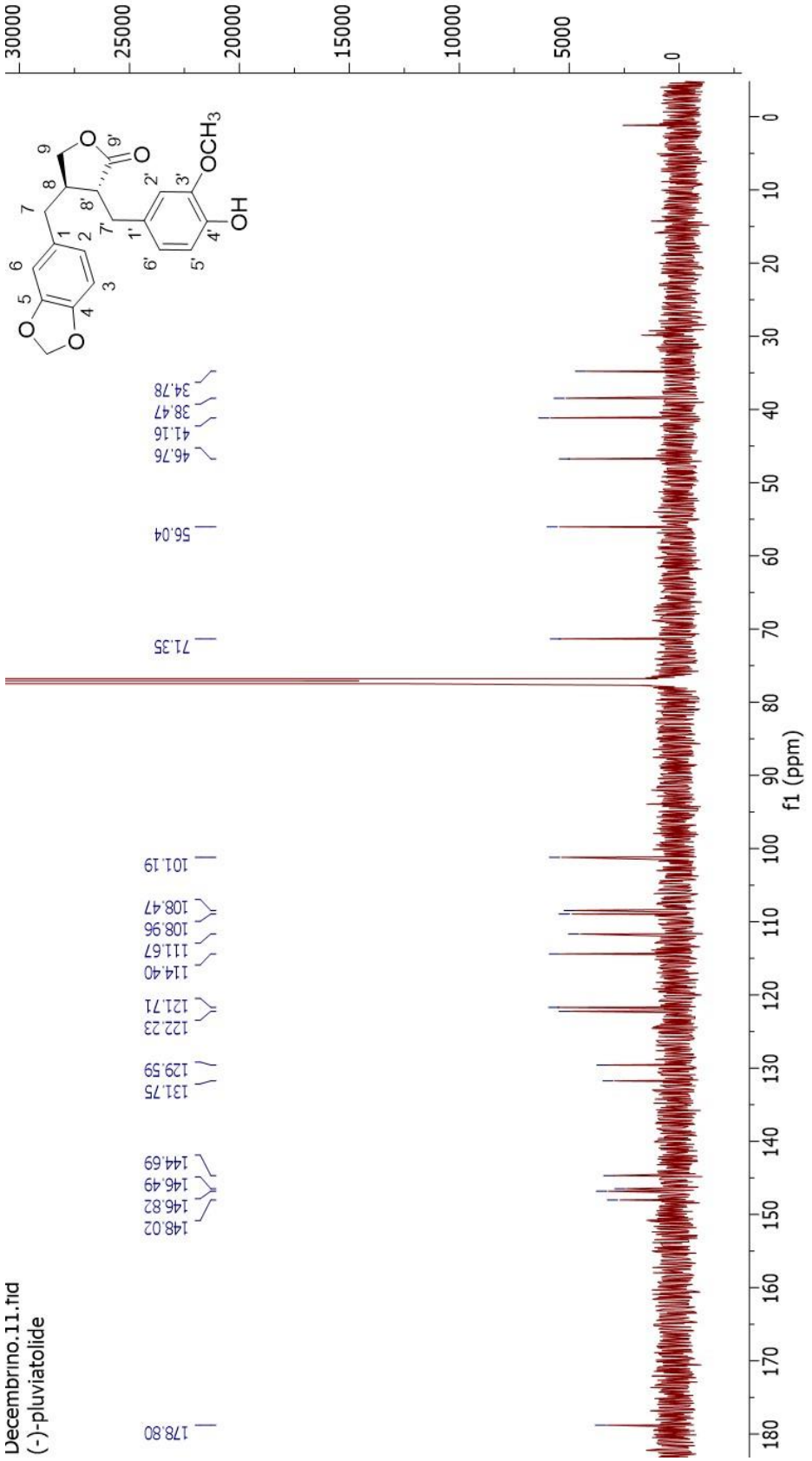


Figure S12. ^{13}C NMR spectrum of (-)-pluviatolide. Data was recorded on a Brucker Avance 600 MHz spectrometer with (-)-pluviatolide in $CDCl_3$

NMR analysis ^{1}H-^{13}C-HSQC spectrum for isolated (-)-pluviatolide

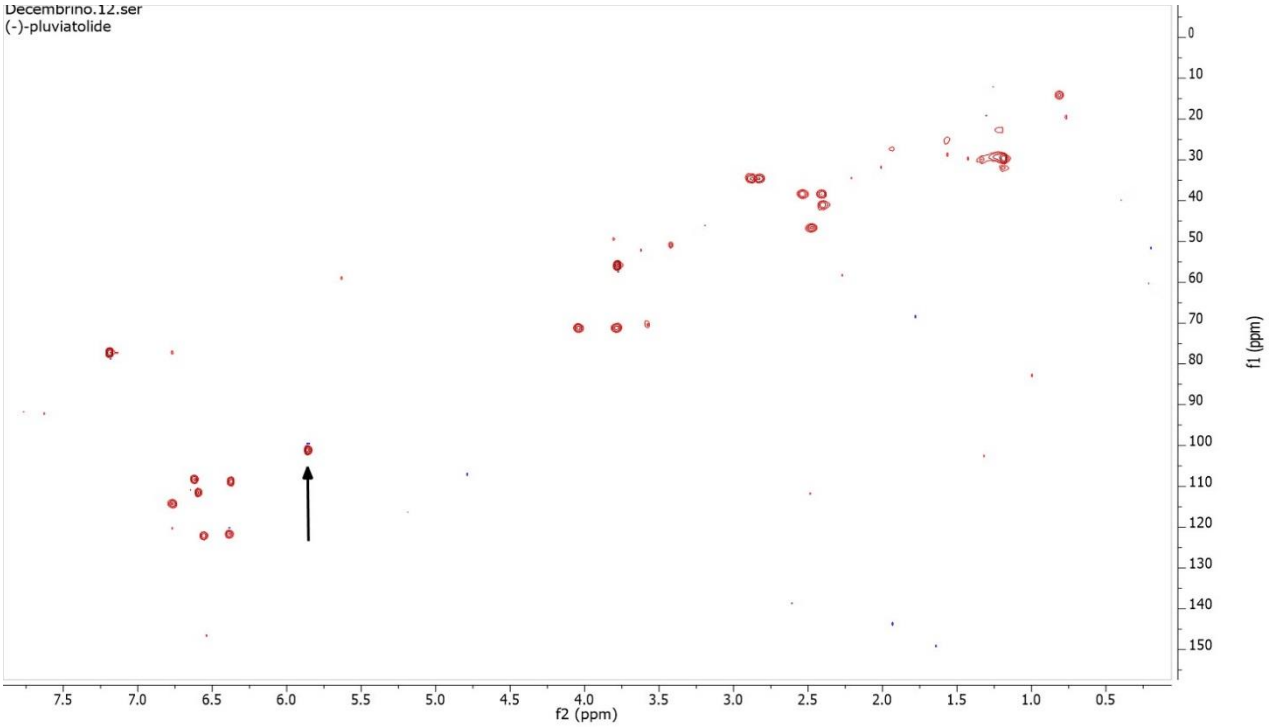


Figure S13. Data was recorded on a Brucker Avance 600 MHz spectrometer, with (-)-pluviatolide in $CDCl_3$. Plot generated with ^{1}H on the x-axis (f2) and ^{13}C on the y-axis (f1). Highlighted (black arrow) the HSQC pulse generated by the characteristic peaks of the methylenedioxy bridge O-CH_2-O at δ 5.9 (^{1}H) and δ 101 (^{13}C), in agreement to previously reported data (Marques *et al.*, 2013)

2.3 From (-) matairesinol towards (-)-podophyllotoxin

Title: *"Synthesis of (-)-deoxypodophyllotoxin, (-)-podophyllotoxin and (-)-epipodophyllotoxin via multi-enzyme cascade reactions in E. coli"*

Authors: Davide Decembrino, Alessandra Raffaele, Thomas Hilberath, Marco Girhard and Vlada B. Urlacher*

Manuscript submitted to Microbial Cell Factories

Own contribution: conceptualization, design, performance of the experiments, analysis and evaluation of the data, drafting of the manuscript and the artworks. Relative contribution 90%

Synthesis of (-)-deoxypodophyllotoxin, (-)-podophyllotoxin and (-)-epipodophyllotoxin via multi-enzyme cascade reactions in *E. coli*

Davide Decembrino[1], Alessandra Raffaele[1], Thomas Hilberath[1], Marco Girhard[1] and Vlada B. Urlacher[1*]

[1] *Institute of Biochemistry, Heinrich-Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany*

**Corresponding author: vlada.urlacher@uni-duesseldorf.de*

Abstract

The aryltetralin lignan (-)-podophyllotoxin is a potent cytotoxic, antiviral, and anti-neoplastic compound that is mainly found in *Podophyllum* and *Sinopodophyllum* plant species. Over the years, the commercial demand for this compound rose notably because of the high clinical importance of its semi-synthetic chemotherapeutic derivatives etoposide and teniposide. To satisfy this demand, (-)-podophyllotoxin is conventionally isolated from the roots and rhizomes of *Podophyllum hexandrum*, which can only grow in a few regions and is now endangered by overexploitation and environmental damage. For these reasons, the development of novel, sustainable routes are of high importance. Herein, we report the reconstitution of the biosynthetic pathway of *Sinopodophyllum hexandrum* as a multi-enzyme cascade in *E. coli* to produce (-)-deoxypodophyllotoxin and (-)-epipodophyllotoxin. The pathway involves *inter alia* three plant cytochrome P450 monooxygenases, with two of them being functionally expressed in *E. coli* for the first time. Recently, we reported a four-step synthesis of the precursor (-)-matairesinol starting from (+)-pinoresinol in *E. coli*. Herein, a five-step conversion of (-)-matairesinol to (-)-deoxypodophyllotoxin was proven effective (98% yield). As a putative (-)-podophyllotoxin synthase from *S. hexandrum* responsible for the final reaction remains unknown to date, two potential candidates CYP107Z and CYP105D from *Streptomyces platensis* were evaluated for their ability to catalyse the hydroxylation of (-)-deoxypodophyllotoxin. Both P450s accepted and converted this substrate, and (-)-podophyllotoxin and (-)-epipodophyllotoxin were detected as reaction products.

Introduction

Herbs and herbal-derived products have a long history as essential components of traditional medical treatments all over the world. Since people health awareness and life expectancy is increasing globally, the demand for effective medicines is also increasing. Nowadays, modern drug development rediscovered plants as a rich source of potent bioactive compounds.[11] In this regard, lignans represent a group of peculiar secondary metabolites with multiple biological activities. In particular, lignans have been described as antioxidant, anti-inflammatory, and powerful cytotoxic compounds, as well as agents preventing the development of breast and prostate cancers or cardiovascular diseases.[64, 111] The aryltetralin lignan (-)-podophyllotoxin has gained much attention due to its potent anti-neoplastic and antiviral properties. There are evidences for its medicinal use through the application of *Podophyllum* species starting back to 2,500 BC.[112] Podophyllotoxin is known for blocking tubulin polymerization and is currently used as antiviral topical agent. Semisynthetic (-)-podophyllotoxin derivatives like teniposide and etoposide with higher solubility, activity, and lower cytotoxicity has been traditionally used as chemotherapeutics for the treatment of different cancers,[73] with the latter being introduced to the list of essential medicines by the World Health Organization (WHO).[71, 113]

Conventional isolation of (-)-podophyllotoxin from natural sources like *Podophyllum* and *Sinopodophyllum* plants and other related *Berberidaceace* species.[69, 114] This practise has become environmentally unbearable due to unregulated plant uprooting within the few regions where these plants can be cultivated.[115] To overcome these limitations, various organometallic chemical catalysts have been successfully combined with single or multiple enzymatic steps, although such approaches often involve the usage of expensive/toxic reagents and harsh reaction conditions.[75, 116] In addition, because of increasing environmental awareness, researchers are keen on putting aside such cumbersome - though elegant - chemo-synthetic approaches and disclosing the potential of enzymes as biocatalysts. Either alone or combined in multi-enzyme cascades, *in vitro* or *in vivo*, biocatalysts generally allow to achieve reactions under mild, biologically compatible reaction conditions, with minor waste production and lower reagents toxicity in comparison to chemical catalysts.[4, 5, 76]

Recent advances in metabolic engineering and synthetic biology boosted the production of plant secondary metabolites in recombinant microorganisms, in which partial or entire plant pathways were reconstituted.[117-119] Over the years, the lignan biosynthetic pathways in plants have been explored, and the enzymes involved in the biotransformation of the phenylpropanoid (*E*)-coniferyl alcohol to the intermediate compound (-)-matairesinol have been disclosed.[63, 78, 79] Yet, only recently light has been shed on the genome of *Sinopodophyllum hexandrum*, which allowed the elucidation of the subsequent steps in the biosynthetic route to (-)-deoxypodophyllotoxin.[70, 80]

Against this background, within this study a part of the pathway of *S. hexandrum* was reconstituted in *E. coli* to allow the biosynthesis of (-)-deoxypodophyllotoxin **6** and (-)-epipodophyllotoxin **7** from the precursor (-)-matairesinol **1** (Figure 1). In two previous studies we reported on efficient four-step multi-enzyme cascades starting from coniferyl alcohol or (+)-pinoresinol to (-)-matairesinol **1** in *E. coli*.[120, 121] In the subsequent steps, (-)-matairesinol **1** is first converted to (-)-pluviatolide **2** by the cytochrome P450 monooxygenase (P450) CYP719A23. The following methylation of (-)-pluviatolide **2** is catalysed by the (-)-pluviatolide-O-methyltransferase (OMT3) yielding (-)-5'-desmethoxy-yatein **3**, which is in turn hydroxylated in the next reaction catalysed by CYP71CU1. Afterwards, a second methylation step is executed by the 5'-desmethyl-yatein O-methyltransferase (OMT1) to produce (-)-yatein **5**. Finally, 2-oxoglutarate/Fe(II)-dependent dioxygenase (2-ODD) catalyses the oxidation of (-)-yatein to furnish (-)-deoxypodophyllotoxin **6**. During this reaction a C-C bond is formed leading to ring closure, which is characteristic for the aryltetralin scaffold of the (-)-podophyllotoxin-like lignans.[70] In this study, we combined these enzymes in a modular manner which allowed to achieve the efficient 5-step conversion of (-)-matairesinol **1** to (-)-deoxypodophyllotoxin **6** with 98% yield. Whereas the heterologous expression and the reconstitution of CYP719A23 activity in *E. coli* had been described in our previous report,[120] CYP71CU1 has been expressed in *E. coli* for the first time. To date, the final stereo- and regioselective hydroxylation of (-)-deoxypodophyllotoxin **6** to (-)-podophyllotoxin **8** remains elusive, and the physiological enzyme of *S. hexandrum* catalysing this reaction has yet to be identified.[122] Within this study we therefore sought to achieve the hydroxylation of (-)-deoxypodophyllotoxin **6** by P450s identified in the genome of *S. hexandrum* or available from previous studies at our group. In particular, CYP82D61 from *S. hexandrum* and CYP107Z and

CYP105D from *Streptomyces platensis*,[51] were tested. Finally, a recombinant *E. coli* strain for the biosynthesis of (-)-epipodophyllotoxin **7** from (-)-matairesinol **1** was developed harbouring a six-step enzymatic cascade involving three P450s. Whereas CYP82D61 led to the formation of the epimer (-)-epipodophyllotoxin **7** as expected based on the work of Lau and Sattely,[70] CYP107Z and CYP105D produced mixtures of (-)-podophyllotoxin **8** and (-)-epipodophyllotoxin **7**, with the latter one being the major product.

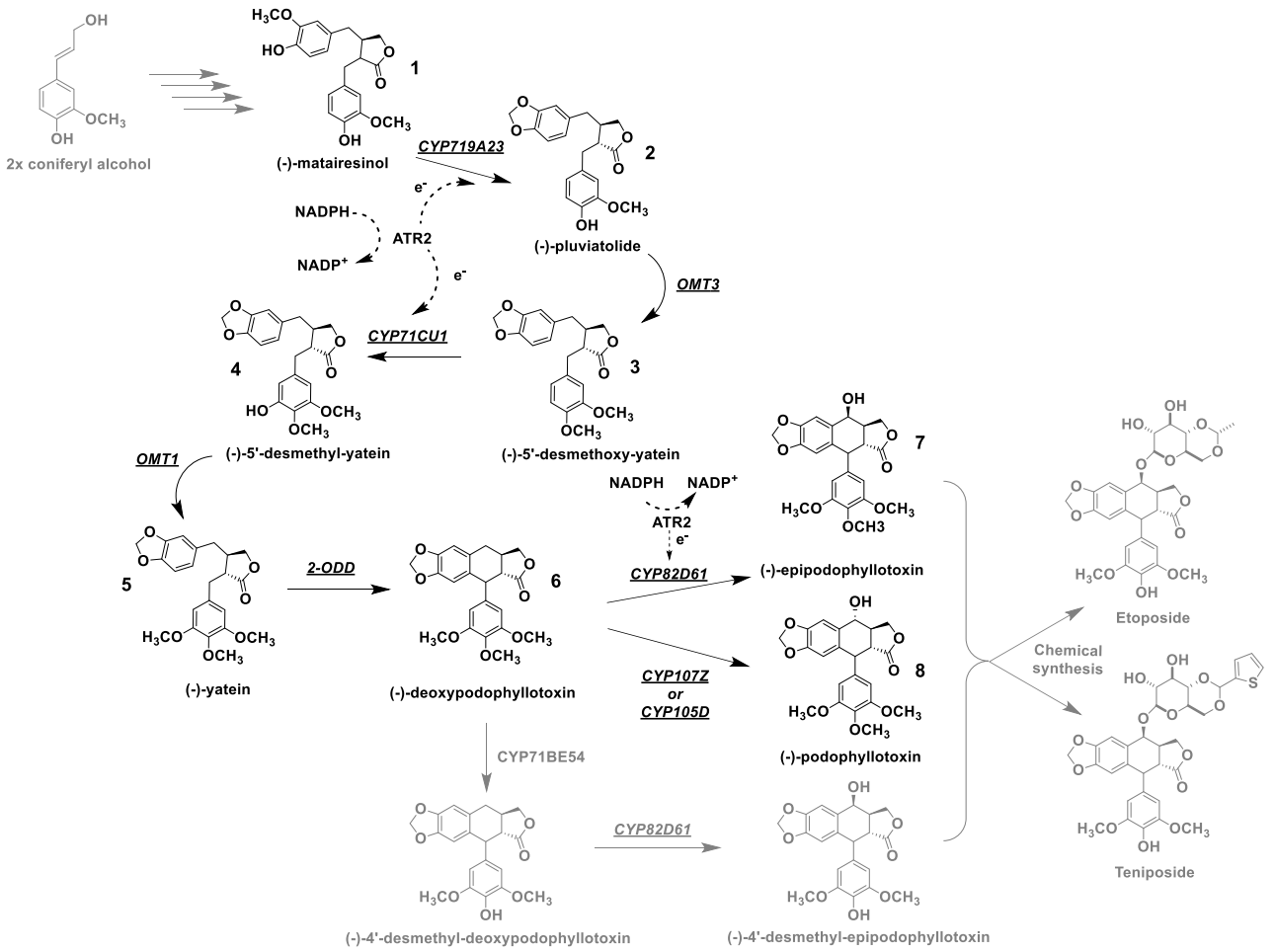


Figure 1. Schematic view for the biosynthesis of (-)-podophyllotoxin and some semi-synthetic derivatives thereof. The six enzymes from *S. hexandrum* that were assembled in this work and *S. platensis* CYPs are underlined. Individual compounds involved into the development of this work are highòighted by a **1** to **8** numbering. Compounds **2**, **3**, **4**, and **7** are not commercially available. Compound 2, (-)-pluviatolide, was isolated and identified within our previous work.[120]

Results and discussion

Screening of *E. coli* strains to express OMT3, OMT1, CYP71CU1 and 2-ODD

Following the genome mining efforts which led to their identification,[70] the successful heterologous expression in *E. coli* of the individual genes of the methyltransferases OMT3, OMT1 and the dioxygenase 2-ODD has previously been reported in literature.[70] Differently, the expression of CYP71CU1 was attempted in *Saccharomyces cerevisiae* only.[70] Within this work, we first selected the most appropriate *E. coli* strain for coexpression of all enzymes to assemble a functional multi-enzyme cascade. To achieve this, we first compared the individual expression levels of *omt3*, *omt1*, and 2-odd genes in the *E. coli* strains BL21 (DE3), C41 (DE3) and C43 (DE3). In summary, the production of OMT3 (41 kDa), OMT1 (38 kDa) and 2-ODD (35 kDa) using their native gene sequences was successful in all strains as judged by SDS-PAGE (Supplementary Figures S1, S2 and S3).

The heterologous expression of CYP71CU1 was endeavoured as well. The physiological bond of the N-terminal moiety to the endoplasmic reticulum membrane, which is a typical feature of eukaryotic P450s, is often a drawback to achieve functional expression in prokaryotic hosts because of the lack of organelles and well-developed inner membranes.[123] Nevertheless, after the expression of the complete *E. coli* codon optimized sequence, named *cyp71cu1oc_wt*, the SDS-PAGE revealed a band in the range of 56 kDa in comparison to the respective negative control (Supplementary Figure S4). *E. coli* C41 (DE3) appeared as the most suitable strain for CYP71CU1 expression. For a more detailed investigation, CO-difference spectra were recorded to estimate the amount of soluble P450. After expression, in every strain, a peak at 420 nm was seen along with a peak at 450 nm (Supplementary Figure S5). Unambiguously, ~ 25 $\mu g_{P450}/g_{cww}$ and ~ 65 $\mu g_{P450}/g_{cww}$ were calculated for *E. coli* BL21 (DE3) and C43 (DE3) respectively, in comparison to ~ 100 $\mu g_{P450}/g_{cww}$ for C41 (DE3). Based on these results, *E. coli* C41 (DE3) was chosen as chassis for all following experiments.

Optimization of CYP71CU1 expression

Since the intrinsic activity of P450s is generally low in comparison to many other enzymes, P450s involved in multi-enzyme cascades have often been reported as a bottleneck.[124] Therefore improving and optimizing of *cyp71cu1* expression and activity represented a crucial step for the development of an efficient multi-enzyme cascade. We optimized several expression parameters: By prolonging the cultivation time from 20 to 48 h an increase of the expression level (160 ± 15 $\mu g_{P450}/g_{cww}$ vs. ~ 100 $\mu g_{P450}/g_{cww}$) was achieved. In addition, the *E. coli* codon optimized gene sequence was compared to the *cyp71cu1* native sequence. After 48 h expression 263 ± 16 $\mu g_{P450}/g_{cww}$ was recorded for the native sequence, which is ~ 1.5-fold higher than the expression of the codon optimized gene (Supplementary Figure S6 and Supplementary Table S8).

Given the successful detection of functional CYP71CU1, further optimization was attempted by manipulating its ER associated N-terminal domain. The partial or total truncation and/or manipulation by changing the sequence of the putative N-terminal membrane-associated region has often been proven successful in raising the expressions levels of plant P450s in *E. coli*.[60] A proline-rich sequence was identified that followed the first 26 amino acids of CYP71CU1, this is a typical feature of microsomal P450s which is proposed to be crucial for the correct protein folding.[125] Variants with deletion of the first 20 amino acids (Δ20) were generated. In addition, a DNA-triplet encoding for the amino acid alanine was introduced as the second codon within the truncated gene sequence, since the translation efficiency was proposed to be enhanced by this strategy.[60, 126]

This manipulation remarkably influenced the amount of detectable soluble CYP71CU1; a 3-fold improvement was observed as compared with the expression levels of the non-truncated version (Supplementary Table S8). Specifically, 813 ± 163 $\mu g_{P450}/g_{cww}$ were achieved using the manipulated native gene (*cyp71cu1Δ20nat*) and 465 ± 138 $\mu g_{P450}/g_{cww}$ with the manipulated codon optimized sequence (*cyp71cu1Δ20oc*). Finally, the expression temperature was varied between 20°C and 30°C; the optimum for the expression of CYP71CU1 variants was 25°C (Supplementary Table S9, Supplementary Figure S7).

Plasmid construction, initial enzyme activity tests, and initial evaluation of coexpression

Since most of the substrates for the individual enzymes are not commercially available, only 2-ODD activity could be tested prior to the multi-enzyme cascade assembly. Using recombinant *E. coli* resting cells, complete conversion of 200 µM (-)-yatein **5** was observed after 3 h. Similar retention times and m/z fragments of the substrate (-)-yatein **5** (m/z 401 $[M+H]^+$) and the product (-)-deoxypodophyllotoxin **6** (m/z 399 $[M+H]^+$) were recorded (Figure 2A). Due to these similar signals in the mass spectrum as well as partially overlapping peaks during chromatographic resolution, the quantification of these compounds could not be unambiguous even by direct comparison to the respective authentic standards. To prove the goodness of the reaction, (-)-yatein **5** depletion was followed over 30 min to confirm the activity of 2-ODD. Indeed, substrate **5** and product **6** were concomitantly detected as twin-peaks after 5 min, regardless of the almost identical features of both compounds (Figure 2B).

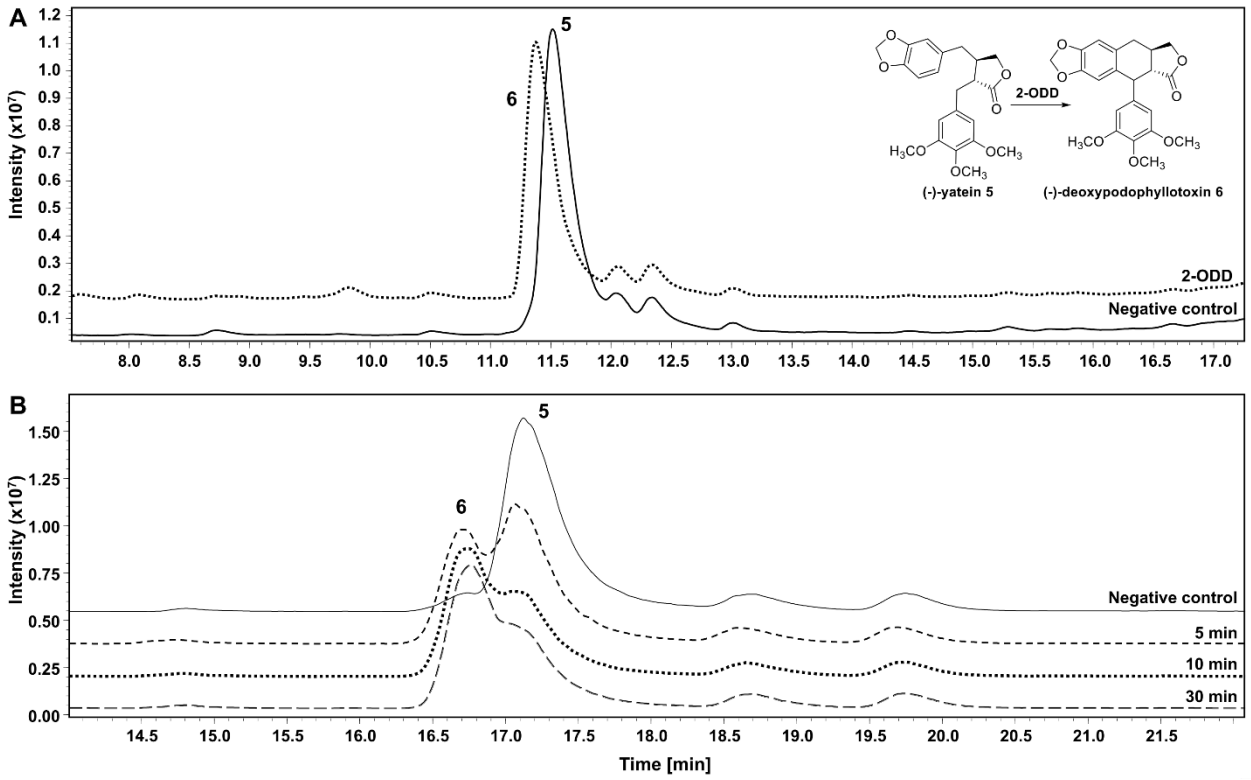


Figure 2. LC/MS analysis to validate 2-ODD activity with (-)-yatein **5**. A: Analysis with method 1 (Table S5), 3 h conversion. B: analysis with method 3, substrate depletion was followed over 30 min. Reaction conditions: 25°C, 1500 rpm in Eppendorf ThermoMixer C, in 2 mL reaction tubes with open lids, 200 µM (-)-yatein **5**.

Next, the activity of OMT3 was proven. In our previous work, the OMT3's substrate (-)-pluviatolide **2** was successfully isolated after production in *E. coli* from (-)-matairesinol **1** by CYP719A23 from *S. hexandrum* supported by the cytochrome P450 reductase ATR2 from *Arabidopsis thaliana.*[120] To perform the two-step biotransformation of (-)-matairesinol **1** to (-)-5'-desmethoxy-yatein **3**, CYP719A23 and ATR2 (plasmid pETDuet_atr2_cyp719aΔ23oc, Supplementary Table S3) were coexpressed with OMT3 (plasmid pCDFDuet_omt3). Utilizing 200 µM (-)-matairesinol **1** as substrate, three major peaks were detected after 30 min. One peak was recognized as (-)-matairesinol **1** (m/z 341 $[M+H-H_2O]^+$, 360 $[M+2H]^{2+}$) and one as (-)-pluviatolide **2** (m/z 339 $[M+H-H_2O]^+$, 358 $[M+2H]^{2+}$). The predominant fragment ions of the third peak (m/z 372 $[M+2H]^{2+}$, 354 $[M+2H-H_2O]^{2+}$) suggest a mass difference (Δm) of +14 g/mol compared to (-)-pluviatolide **2**, pointing to a methylated product. After 3 h the characteristic m/z fragments of this product distinctly increased concomitant with the full conversion of (-)-pluviatolide **2**, endorsing OMT3 activity (Supplementary Figure S8). The observed m/z fragments correspond to those associated to the isolated (-)-5'-desmethoxy-yatein **3** reported elsewhere (Supplementary Table S6).[70]

CYP71CU1-mediated hydroxylation of (-)-5'-desmethoxy-yatein **3** produced by OMT3 represents the next reaction step of the cascade. Therefore, we next evaluated the coexpression of CYP71CU1 and OMT3 (plasmid pCDFDuet_omt3_cyp71cu1Δ20nat). Under the established optimized expression conditions for CYP71CU1, when coexpressed with OMT3 in the same *E. coli* cell, CYP71CU1 concentration achieved 240 ± 15 $\mu g_{P450}/g_{cww}$ which is 3-fold lower in comparison to the single expressed CYP71CU1 (Supplementary Figure S9A). This is not surprising because in our previous work we observed a significant reduction of a P450 concentration when multiple genes were coexpressed from the same expression vector.[120] In this regard, it is worth mentioning that P450s rely on electrons which are ultimately delivered by reduced nicotine amide cofactors NAD(P)H via redox partner proteins.[35] Concerning eukaryotic organisms, the redox equivalents from NAD(P)H are transferred by cytochrome P450 reductases (CPR).[33] Intuitively, the combination of redox proteins belonging to the same organism represent the method of choice, however, CPRs from *S. hexandrum* are unknown to date. In our previous study, ATR2 was found to support CYP719A23 activity both *in vitro* and *in vivo*.[120] In this regard, in plant natural systems a P450 to CPR ratio of

~ 15:1 has been reported[127, 128]; in other words, the activity of multiple P450s should be sufficiently supported by significantly lower amounts of CPR.

Based on these considerations, we assumed that the expression level of ATR2 should be sufficient to sustain the activity of both CYP719A23 and CYP71CU1 upon coexpression with OMT3 in one *E. coli* cell (coexpression of plasmids pETDuet_atr2_cyp719aΔ23oc and pCDFDuet_omt3_cyp71cu1Δ20nat). Under the above-described optimized conditions, when coexpressed with OMT3 in the same *E. coli* cell, CYP71CU1 concentration achieved 240 ± 15 $\mu g_{P450}/g_{cww}$ which is 3 times lower than the single expressed enzyme (Supplementary Figure S9A). This is not surprising; as a matter of fact previous experience suggested the likelihood of significant reduction of P450 amounts when multiple genes are harboured within the same expression vector.[120] The inclusion of OMT1 and 2-ODD catalysing the follow-up reaction steps finished assembly of the desired multi-enzyme cascade. 2-ODD and OMT1 were coexpressed utilizing plasmid pCOLADuet_2-odd_omt1 (Supplementary Table S3 and Supplementary Figure S9B). In summary, three plasmids were designed in a modular fashion to allow the coexpression of all necessary genes and to target the biosynthesis of (-)-deoxypodophyllotoxin in *E. coli* (Supplementary Figure S10).

Implementation of the cascade from (-)-matairesinol 1 to (-)-deoxypodophyllotoxin 6 in *E. coli*

One-cell approach

As a first try to achieve the 5-step biotransformation of (-)-matairesinol **1** to (-)-deoxypodophyllotoxin **6**, all necessary pathway enzymes were coexpressed in a single cell harbouring three plasmids. This "one-cell" approach was evaluated in reaction flasks in 10 mL using 200 µM (-)-matairesinol **1** (Figure 3). Following the reaction over time, substrate depletion, intermediates formation and consumption, as well as the final product (-)-deoxypodophyllotoxin **6** were monitored. In the absence of commercially available authentic references, the identification of the intermediates (-)-5'-desmethoxy-yatein **3** and (-)-5'-demsethyl-yatein **4** was possible relying on the work of Lau and Sattely (Supplementary Table S6).[70]

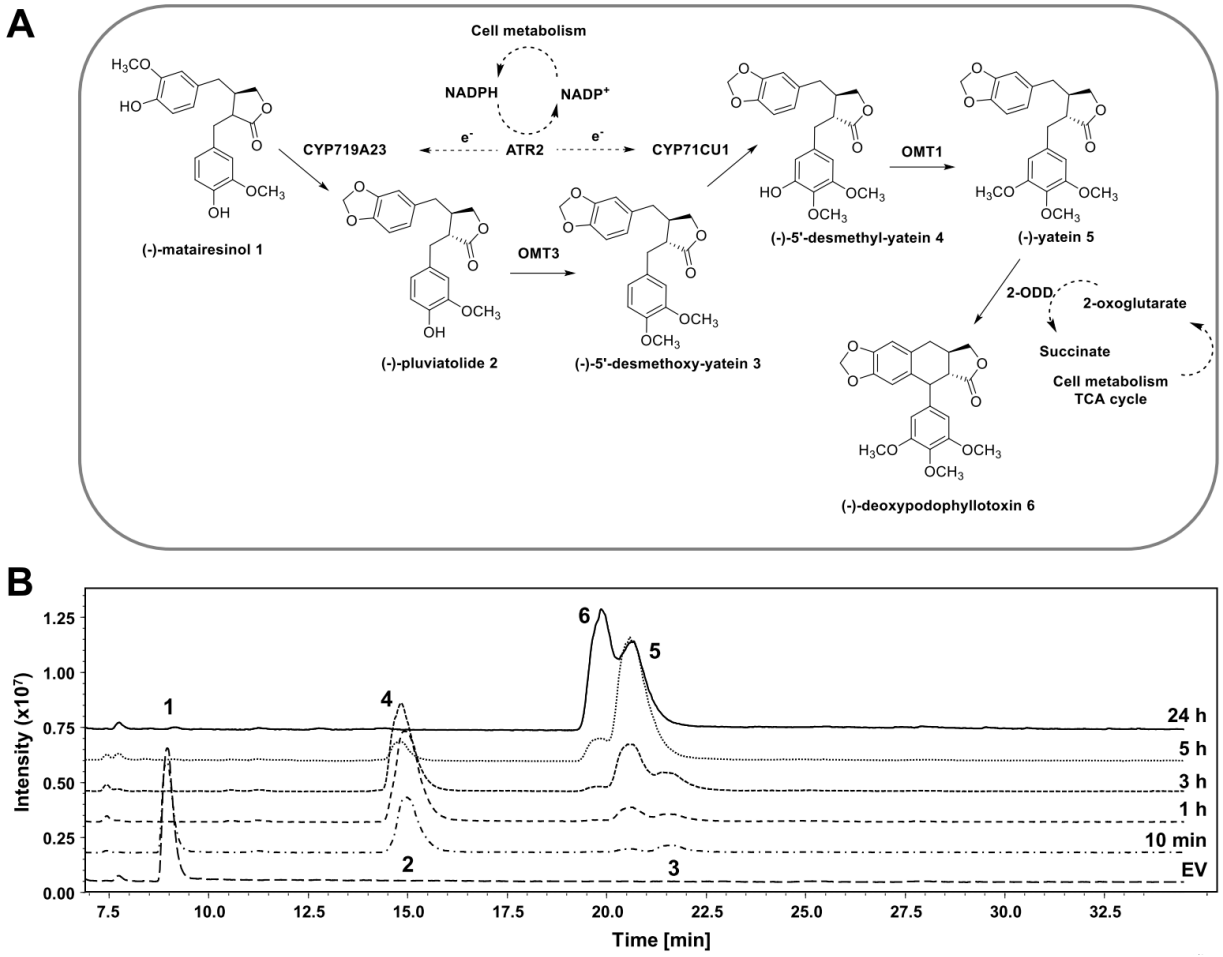


Figure 3. **A**: Schematic picture of the one-cell biotransformation of (-)-matairesinol **1** to (-)-deoxypodophyllotoxin **6**. **B**: LC/MS analysis performed with method 4 (Supplementary Table S5). Compounds were identified via MS fragmentation by comparison to either commercially available authentic standards or literature references (Supplementary Tables S6). Reaction conducted at 25°C, 250 rpm in baffled Erlenmeyer flasks in an orbital shaker; 10 mL cell suspension, 200 µM substrate.

The activity of CYP719A23 in combination with ATR2 was proven by ~50% consumption of (-)-matairesinol **1** (m/z 341 $[M+H-H_2O]^+$, 359 $[M+H]^+$) within the first 10 min, and formation of the intermediate (-)-pluviatolide **2** (m/z 339 $[M+H-H_2O]^+$, 358 $[M+H]^+$) which accounted for 86% of all detected intermediates. Complete (-)-matairesinol **1** depletion was observed after 1 h; at that time point the presence of the intermediates (-)-5'-desmethoxy-yatein **3** (~7%, m/z 353 $[M+H-H_2O]^+$, 371 $[M+H]^+$) and (-)-yatein **5** (~ 13%, m/z 383 $[M+H-H_2O]^+$, m/z 401 $[M+H]^+$) was observed, confirming the activities of OMT3 and OMT1, respectively. The activity of CYP71CU1 can also be concluded, given that this enzyme performes the hydroxylation step between the steps catalysed by OMT3 and OMT1. After 3 h, (-)-5'-desmethyl-yatein **4** formed in the CYP71CU1-catalysed reaction could be

better evaluated. The structural similarity of (-)-5'-desmethyl-yatein **4** to (-)-pluviatolide **2** generates overlapping peaks, however distinctive fragmentation patterns allow to discern the compounds qualitatively. MS signals of (-)-pluviatolide **2** (m/z 339 $[M+H-H_2O]^+$ and 358 $[M+H]^+$) were poorly detectable, while the intensity of (-)-5'-desmethyl-yatein **4** fragments (m/z 369 $[M+H-H_2O]^+$ and m/z 387 $[M+H]^+$) fragments increased. Notwithstanding the difficult quantification, this finding demonstrates that CYP71CU1 is catalytically active in *E. coli* and it also implies that ATR2 is performing as adequate electron shuttle to both CYP719A23 and CYP71CU1, in the current reaction setup. After 5 h (-)-5'-desmethyl-yatein **4** (~7%) was detected together with ~ 84% of the intermediate (-)-yatein **5** (m/z 383 $[M+H-H_2O]^+$, m/z 401 $[M+H]^+$). At the same time point, (-)-deoxypodophyllotoxin **6** (m/z 399 $[M+H]^+$, 421 $[M+Na]^+$) accounts for ~ 10%. After prolongation of the reaction to 24 h (-)-yatein **5** and (-)-deoxypodophyllotoxin **6** were detected in a 50:50 ratio, which revealed 2-ODD-catalysed reaction as an apparent bottleneck (Figure 3). This outcome was somewhat unexpected, since as described in the previous chapter, single expressed 2-ODD was able to fully convert 200 µM (-)-yatein within 3 h.

The increased metabolic burden determined by the coexpression of multiple enzymes from three plasmids within a single cell may represent a straightforward explanation for this result. The presence of several redox enzymes in the cascade is also likely to generate competition for cofactors and co-substrates within the cell.[17] Although glucose was added as carbon source to support *E. coli* metabolism for TCA cycle or pentose phosphate pathway, its provision likely needs optimization to ensure a good balance between the host primary metabolic machinery and the heterologous one. For instance, 2-ODD's co-substrate 2-oxoglutarate is highly required within TCA cycle.[20] It is therefore likely that a lack of co-substrate is restricting the performance of this enzyme.

Two-cells approach

In the attempt to reduce the metabolic burden hypothesised for the one-cell approach, the cascade was divided between two *E. coli* cells, that from now on will be referred to as modules. Module one contained CYP719A23 and ATR2 together with OMT3 and CYP71CU1 covering (-)-matairesinol **1** biotransformation to (-)-5'-desmethyl-yatein **4**, while module two harboured OMT1 and 2-ODD to finalize (-)-deoxypodophyllotoxin **6** synthesis (Figure 4).

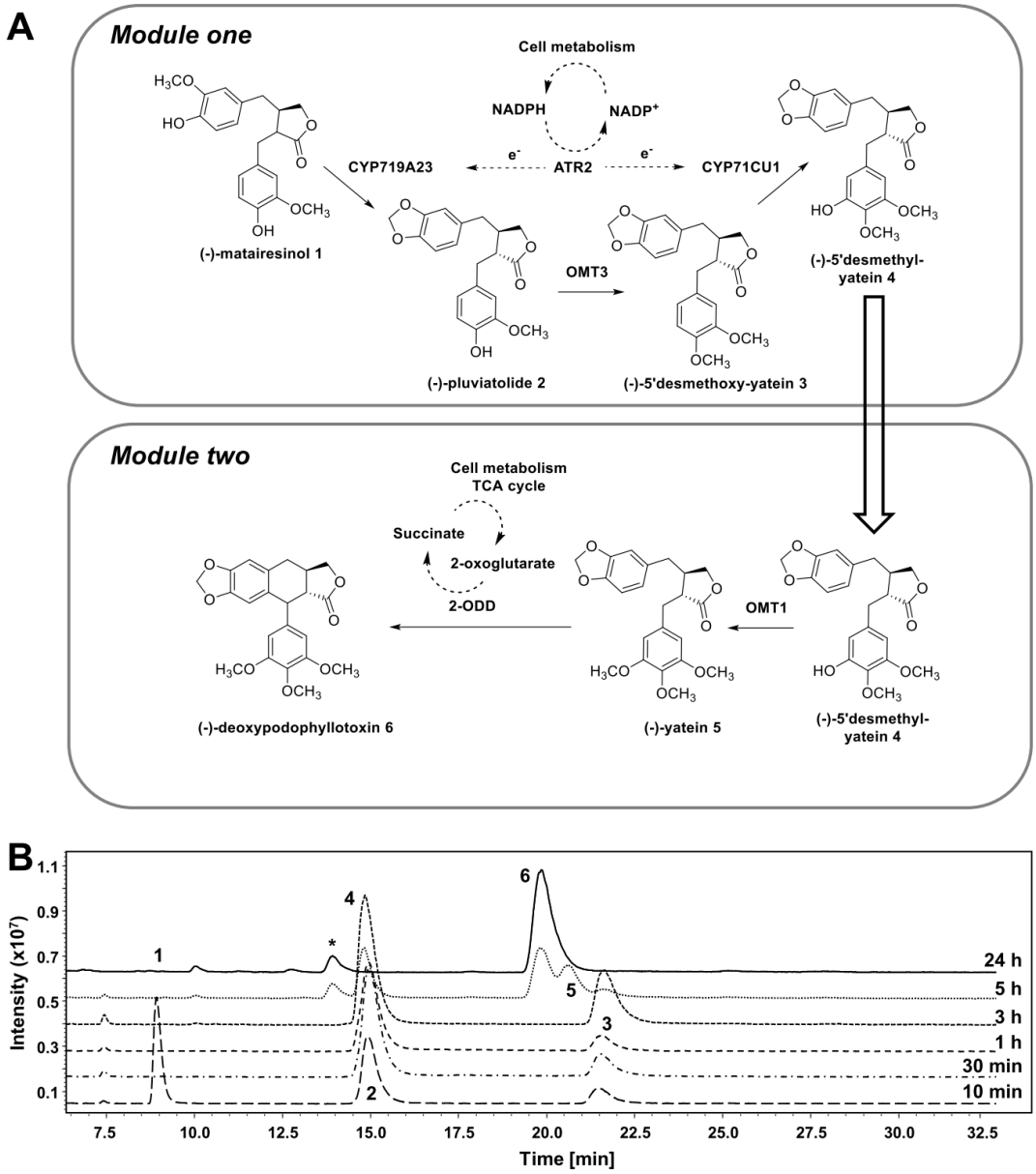


Figure 4. **A**: Schematic picture of the sequential two-cell biotransformation of (-)-matairesinol **1** to (-)-deoxypodophyllotoxin **6**. After 3 h conversion by module one, module two was added to the reaction. **B**: LC/MS analysis performed with method 4 (Supplementary Table S5). Compounds were identified via MS fragmentation by comparison to either commercially available authentic standards or literature references (Supplementary Tables S6). Reaction conducted at 25°C, 250 rpm in baffled Erlenmeyer flasks in an orbital shaker; 2x10 mL cell suspension, 200 µM (-)-matairesinol **1**.

The two modules were added in a sequential manner: 3 h after conversion of 200 µM (-)-matairesinol **1** was started using module one, module two was added. Similar to the one-cell approach, ~ 60% substrate depletion was observed within 10 min, and full (-)-matairesinol **1** conversion after 1 h. At this time, (-)-pluviatolide **2** (85%) and (-)-5'-desmethoxy-yatein **3** (15%) were the only detectable compounds. Prior to the addition of module two (after 3 h) no m/z fragments related to (-)-pluviatolide **2** were observed; instead, m/z 369 $[M+H-H_2O]^+$ and m/z 387 $[M+H]^+$ were recorded at a retention time (RT) 14.9 min, identifying (-)-5'-desmethyl-yatein **4** (the final product of module one) as the most prominent product (~ 65%).

Already 2 h after module two addition (corresponding to 5 h total reaction time), the two-cells approach seemed more efficient than the one-cell approach. At this time point, (-)-deoxypodophyllotoxin **6** and (-)-yatein **5** corresponded to ~ 40% and ~ 23% of all detected compounds respectively, and after 24 h (-)-yatein **5** was fully converted to (-)-deoxypodophyllotoxin **6**. Noticeably, 2-ODD performance was substantially enhanced in the sequential two-cells approach.

It should be noted that a distinct though minor side product peak was detected after the addition of module two at ~ 13.5 min (marked with an asterisk within Figure 4B). The mass fragments (m/z 385 $[M+H]^+$ and 407 $[M+Na]^+$) suggest that aryltetralin ring closure may happen consequentially to the 2-ODD-catalysed oxidation. As (-)-5'-desmethyl-yatein **4** molecular weight is 386 g/mol, the observed mass difference is coherent to a compound with a predicted molecular weight of 384 g/mol, which corresponds to (-)-5'-desmethyl-deoxypodophyllotoxin (Supplementary Figure S11).

Application of two-cells was remarkably effective: 196 µM of (-)-deoxypodophyllotoxin were quantified as a result of this 5-step enzymatic reaction cascade, corresponding to 98% yield unambiguously demonstrating the advantage of this setup in comparison to the one-cell approach.

Evaluation of (-)-deoxypodophyllotoxin hydroxylation

CYP82D61 expression and activity evaluation in E. coli

In a previous study by Lau and Sattely, the accumulation of (-)-epipodophyllotoxin **7** was described after (-)-deoxypodophyllotoxin **6** transformation catalysed by CYP82D61 from *S. hexandrum.* Within the same work, this compound was isolated and identified after heterologous expression of CYP82D61 in *N. benthamiana* leaves.[70] Therefore, we attempted the expression and activity evaluation of CYP82D61 in *E. coli*. Similar to the optimization strategy applied for CYP71CU1, the N-terminal membrane-associated sequence of CYP82D61 was manipulated by generating a truncated variant where the first 23 amino acids were deleted (Δ23) and the second codon was substituted by alanine. Using *E. coli* C41(DE3) as host, the expression of the *E. coli* codon optimized gene sequence *cyp82d61oc_wt* and the manipulated gene variant *cyp82d61Δ23oc* was conducted for 48 h at temperatures between 20° to 30°C. Expression of the *wt* gene was successful only at 30°C and reached 250 ± 14 $\mu g_{P450}/g_{cww}$. The N-terminal manipulation was beneficial and improved the expression of CYP82D61 at 30°C resulting in 1395 ± 220 $\mu g_{P450}/g_{cww}$ (compared to 936 ± 34 $\mu g_{P450}/g_{cww}$ at 25°C, no detectable expression at 20°C; Supplementary Table S10). Regarding the expression temperature, this outcome was somehow unexpected since multiple physiological features including gene expression have been reported to be enhanced in *S. hexandrum* at temperatures lower than 25°C.[122, 129, 130] CYP82D61 activity was further evaluated using recombinant *E. coli* cells coexpressing ATR2 (plasmid pETDuet_atr2_cyp82d61Δ23oc; Supplementary Table S3). (-)-Deoxypodophyllotoxin **6** was converted up to ~ 30% after 24 h (Figure 5A). The m/z fragments of the product (m/z 415 $[M+H]^+$ and m/z 437 $[M+Na]^+$) match the features of (-)-epipodophyllotoxin **7** as described previously by Lau and Sattely (Supplementary Table S6).[70]

In vitro activity evaluation of CYP107Z and CYP105D from S. platensis

Along with CYP82D61, two bacterial P450s were evaluated to perform the last hydroxylation step, namely CYP107Z and CYP105D from *Streptomyces platensis*. These P450s were chosen as they were found to oxidize a broad range of chemical compounds of different size.[51] Activity of both P450s against (-)-deoxypodophyllotoxin **6** was evaluated *in vitro*, using the three-component system comprising besides a P450, the flavodoxin YkuN from *Bacillus subtilis* and the flavodoxin reductase FdR from *E. coli* that have been established in previous studies.[45, 51, 101, 105] Substrate conversion achieved 32% and 17% with CYP107Z and CYP105D, respectively, with similar product distributions observed in both cases. Among the four identified peaks, two were recognized with distinct fragments of m/z 415 $[M+H]^+$, corresponding to putatively hydroxylated products given the mass difference (Δm) of +16 g/mol compared to (-)-deoxypodophyllotoxin **6**.

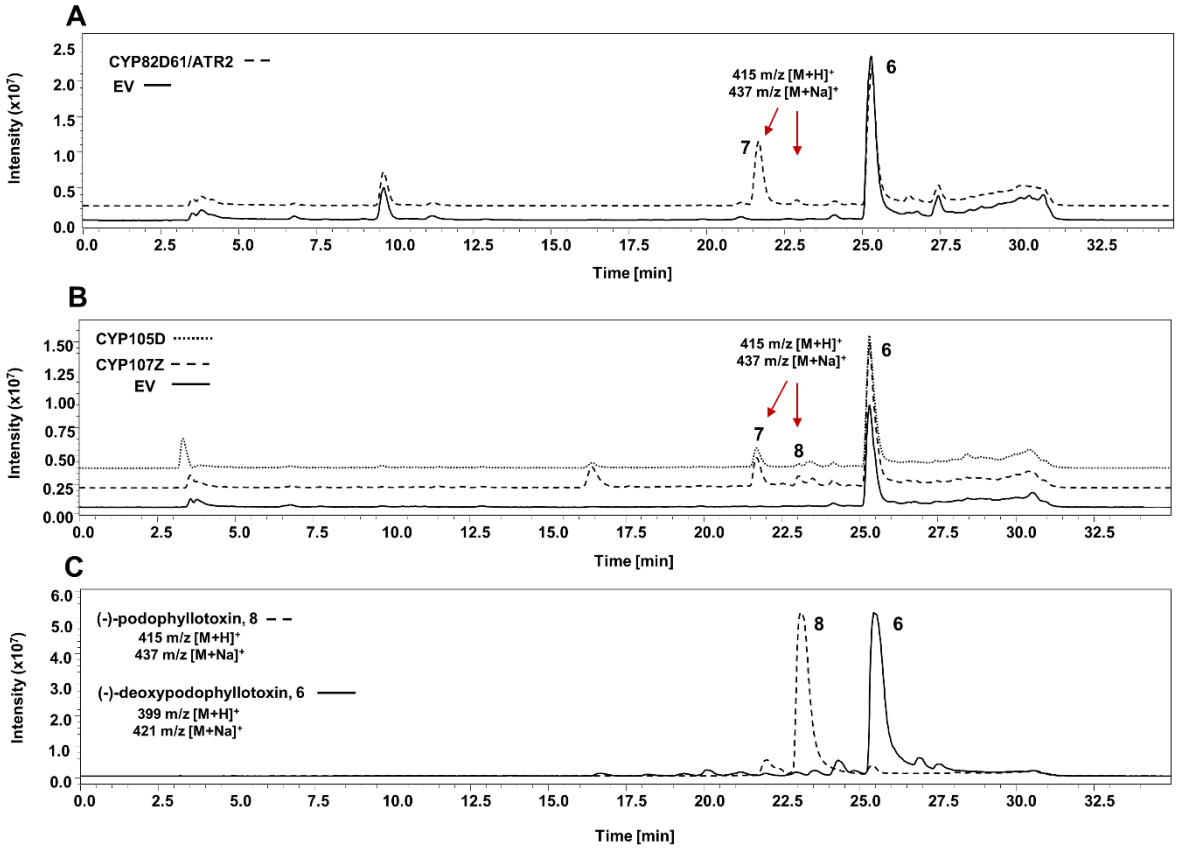


Figure 5. P450 activity screening for (-)-deoxypodophyllotoxin **6** hydroxylation. **A:** CYP82D61 from *S. hexandrum*. **B**: CYP107Z and CYP105D from *S. platensis*. **C**: authentic standards for (-)-deoxypodophyllotoxin **6** and (-)-podophyllotoxin **8**. Analysis with method 2 (Supplementary Table S5). Reaction conducted at 25°C, 1500 rpm in ThermoMixer, open lids for 24 h in 0.5 mL total volume. 200 µM substrate, 2% (v/v) DMSO.

In case of CYP107Z these two putative hydroxylated compounds represented ~10% and ~40% of the four products; the first one had corresponding RT and major m/z fragments (415 $[M+H]^+$ and 437 $[M+Na]^+$) to the (-)-podophyllotoxin **8** authentic standard. The second peak is again attributed to be the (-)-podophyllotoxin epimer **7** (Figure 5B) in light of the results achieved with CYP82D61 described in the previous chapter.

Extension of the multi-enzyme cascade toward (-)-epipodophyllotoxin 7

The established two-cells approach allowed a straightforward extension. For the six-step multi-enzyme cascade from (-)-matairesinol **1** to (-)-epipodophyllotoxin **7**, module one harboured CYP719A23, ATR2, OMT3, and CYP71CU1, whereas module two contained OMT1, 2-ODD, CYP82D61, and ATR2 (Supplementary Figure S12). Under the same conditions as described before, (-)-matairesinol **1** was readily consumed by CYP719A23 within the first 10 min, with (-)-pluviatolide **2** accumulating to its maximum concentration after 1 h and being completely depleted after 3 h. Concomitantly, (-)-5'-desmethoxy-yatein **3** was efficiently hydroxylated by CYP71CU1 resulting in (-)-5'-desmethyl-yatein **4** accumulation up to ~70% after 3 h. After 5 h total reaction time (2 h after module two addition), (-)-5'-desmethyl-yatein **4** was depleted by 50% through OMT1 yielding (-)-yatein **5**, which soared as the major intermediate at this time point (~ 55%). In comparison to the previous five-step cascade, the inclusion of CYP82D61 and ATR2 into module two appears to slightly influence the efficacy of 2-ODD, since the turnover of (-)-yatein **5** was slowed down. As hypothesised for the one-cell approach, the addition of redox enzymes may lead to resources competition that reduces enzymatic performance. Nevertheless, full depletion of (-)-yatein **5** by 2-ODD was observed after 24 h. At the same time 160 µM (-)-deoxypodophyllotoxin **6** were quantified being the most abundant product (> 85%). Additionally (-)-epipodophyllotoxin **7** (~ 10%) was observed as the product of (-)-deoxypodophyllotoxin **6** hydroxylation catalysed by CYP82D61 (Figure 6).

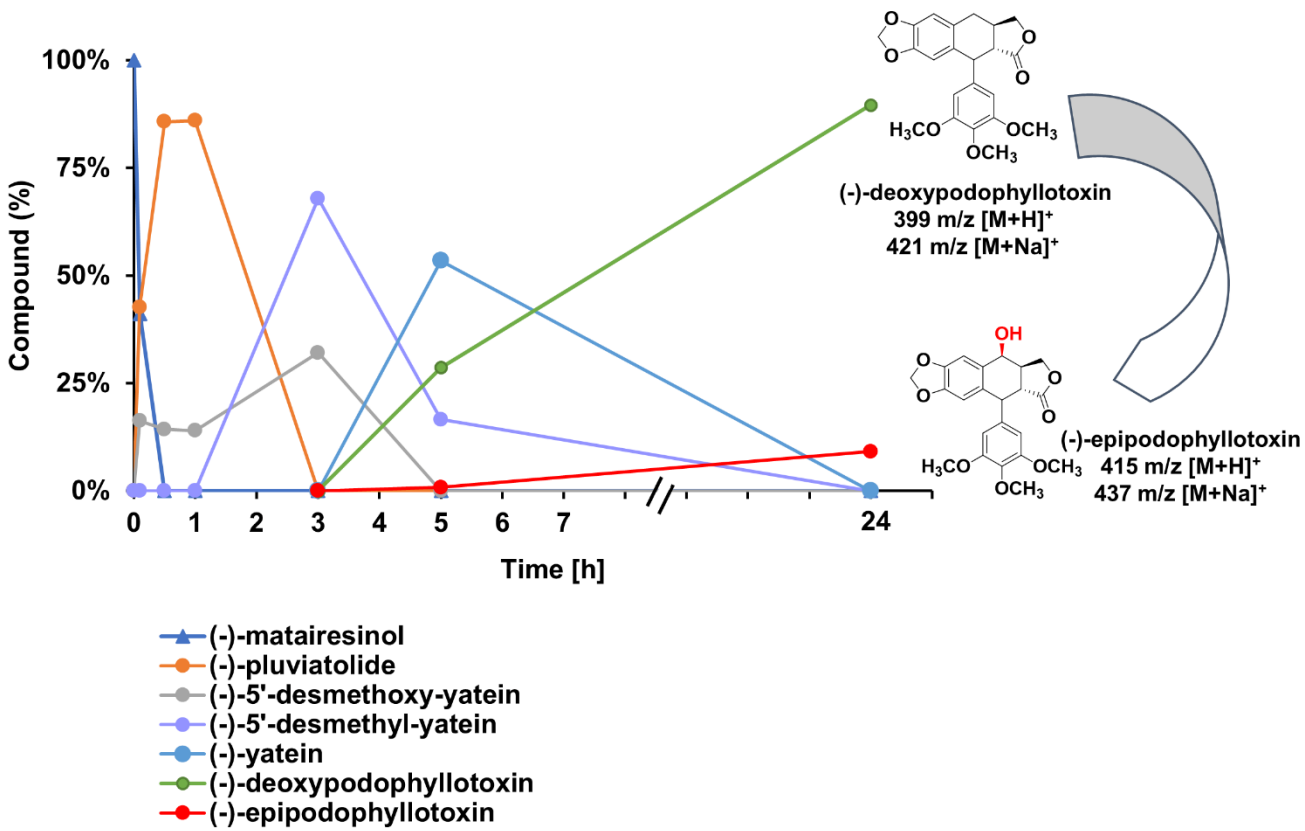


Figure 6. Two-cells biotransformation of (-)-matairesinol **1** to (-)-epipodophyllotoxin **7**. After 3 h conversion by module one, module two were added. LC/MS analysis performed with method 4 (Supplementary Table S5). Reaction conducted at 25°C, 250 rpm in baffled Erlenmeyer flasks in orbital shaker; 20 mL total volume (10 mL of each module), 200 µM substrate, 2% (v/v) DMSO.

Summary and conclusion

Within this study, eight enzymes were successfully combined to produce the direct precursor of (-)-podophyllotoxin **8** – (-)-deoxypodophyllotoxin **6** – as well as (-)-epipodophyllotoxin **7** in *E. coli* for the first time. The described results demonstrate the advantages of a sequential *two-cells* approach. The hypothesised metabolic burden resulting from the combination of multiple redox enzymes in one *E. coli* cell was relieved in a pragmatic manner, as no cumbersome host manipulation was required. Although P450-mediated reactions are often reported as rate-limiting step in reconstituted biosynthetic pathways, in the described cascade their activity does not represent a bottleneck (with only exception for the final hydroxylation step, catalysed by non-physiological promiscuous enzymes).

Due to the modular nature of the developed multi-enzyme cascade, it can straightforward be extended towards another precursor of etoposide and teniposide by introducing CYP71BE54 (Figure 1).[70] All in all, the introduction of CYP71BE54 into the established reaction cascade should be relatively straightforward due to the modular nature of the developed strategy and the well-established know-how to achieve functional expression of eukaryotic P450s in *E. coli*. Regarding CYP107Z and CYP105D, the use of prokaryotic P450s that mimic the activity of eukaryotic ones is generally of particular interest. Taking CYP107Z and CYP105D as models, the natural cytosolic localization of these enzymes simplifies the heterologous expression in *E. coli*, since extensive N-terminal manipulations are not required. Moreover, numerous examples of protein engineering upon bacterial P450s are reported regarding the tuning of their catalytic activity.[101, 131, 132] Needless to say that comprehensive protein engineering efforts could also be applied to CYP107Z and CYP105D to enhance substrate turnover, regio- and stereoselectivity.

To ensure the utilization of cheaper substrates than (-)-matairesinol **1**, the cascade can also be extended backwards and coupled with, for instance, a four-step cascade from (+)-pinoresinol to (-)-pluviatolide **2**[120] and even further with a two-step conversion of the cheap precursor eugenol to pinoresinol that both have been established earlier in our group [86, 102]

In conclusion, the present study describes the first example of a completely biosynthetic route to (-)-deoxypodophyllotoxin **6** and (-)-epipodophyllotoxin **7** and represents a solid step forward to the development of a more sustainable supply for these direct precursors to etoposide and teniposide.

Materials and Methods

E. coli strains, enzymes and authentic reference compounds

E. coli DH5α (Clonetech, USA) was used for cloning purposes and *E. coli* BL21 (DE3), C41 (DE3) and C43 (DE3) (Lucigen, USA) served for gene expressions and biotransformations. Endonucleases *Nco*I, *Not*I, *Nde*I and *Xho*I, Phusion High Fidelity DNA polymerase, thermosensitive alkaline phosphatase (FastAP™), T4 DNA-ligase, and catalase from bovine liver originated from Thermo Scientific (Germany). T4 polynucleotide kinase and *Dpn*I were purchased from New England Biolabs (USA). Authentic reference compounds used for LC/MS identification and their respective suppliers are listed in Supplementary Table S6.

Reaction cascade design

Synthetic genes were ordered from BioCat GmbH (Germany). (-)-Pluviatolide-O-methyltransferase (GenBank KT390157.1, OMT3), 5'-desmethyl-yatein O-methyltransferase (GenBank KT390155.1, OMT1) and the 2-oxoglutarate/Fe(II)-dependent dioxygenase, deoxypodophyllotoxin synthase, (GenBank KT390173.1, 2-ODD) were purchased as native sequences. The gene encoding for CYP71CU1 (GenBank KT390172.1) was ordered as native and *E. coli* codon optimized gene sequence, whereas the gene encoding for CYP82D61 (GenBank KC110995.1) was ordered in *E. coli* codon optimized form only (Supplementary Table S1). With the aim of developing a multi-enzyme cascade, conventional cloning methods were used to insert the sequences of OMT3 and CYP71CU1 in pCDFDuet-1 (Novagen, Merck, Germany) in the multiple cloning sites (MCS) I and II, respectively (Supplementary Table S2 and S3; Supplementary Figures S10 and S12). In the same fashion, the sequences of 2-ODD and OMT1 were inserted in MCS I and II of a pCOLADuet-1 (Novagen, Merck, Germany) plasmid, whereas the CYP82D61 sequence was cloned into MCS II of a pETDuet-1 (Novagen, Merck, Germany) plasmid. The sequence coding for the NADPH-cytochrome P450 reductase 2 from *Arabidopsis thaliana* (Protein ID NP_194750.1, ATR2) was manipulated as described elsewhere,[104, 105, 133, 134] and cloned in MCS I of the pETDuet-1 vector already harbouring

cyp82D61 (Supplementary Table S2 and S3; Supplementary Figure S12). The genes encoding for CYP719A23 and ATR2 were cloned in a pETDuet-1 plasmid as described in a previous study by our group (Supplementary Table S2 and S3; Supplementary Figures S10 and S12).[120] For the expression of individual enzyme sequences coding for OMT3, CYP71CU1, OMT1 and 2-ODD were cloned in pCDFDuet-1, whereas pETDuet-1 was used for CYP82D61, as summarized in the Supplementary Table S3.

The putative membrane associated N-terminal sequences of CYP71CU1 and CYP82D61 were identified using the TMHMM software (version 2.0; http://www.cbs.dtu.dk/services/TMHMM/) in combination with the predictor of signal peptides and membrane protein topology named Spoctopus (https://octopus.cbr.su.se/index.php?about=SPOCTOPUS). Truncated variants of both enzymes were created using the oligonucleotides described in the Supplementary Tables S4 following the supplier instructions of the Q5 site-directed mutagenesis kit (New England Biolabs, USA) and screened for expression in *E. coli* in comparison to the respective wild type sequences.

CYP107Z (GenBank OSY37796) and CYP105D (GenBank OSY47991) from *Streptomyces platensis* as well as FdR, YkuN, and the glucose dehydrogenase (GDH) from *Bacillus megaterium* (GenBank D10626) for NADPH regeneration were generated as previously described (Supplementary Tables S2-S3).[51, 104-106] The correct insertion of genes within the expression vectors was verified via automated Sanger sequencing (Eurofins Genomics GmbH, Germany).

E. coli cultivation, protein expression and cell density normalization

Prior to expression of the target genes, biological duplicates of *E. coli* cells transformed with different plasmid combinations were precultivated in 5 mL Luria-Bertani (LB) medium supplemented with the appropriate antibiotics (streptomycin 50 µg/mL and/or ampicillin 100 µg/mL and/or kanamycin 30 µg/mL) and incubated for 16 h at 37°C, 180 rpm. 50 mL Terrific broth (TB) media supplemented with the appropriate antibiotics were inoculated with 500 µL from the precultures, and cells were grown at 37°C, 180 rpm to an OD_{600} of 0.6. The expression of target genes was induced with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) supplemented by 0.5 mM 5-aminolevulinic acid (5-ALA), and 0.1 mM $FeSO_4$ in case a P450 was (co-)expressed. Thereafter, the cultures were incubated for 48 h at 25°C, 120 rpm, unless stated otherwise. This general setup is a compromise of the most suitable expression conditions identified for each single expressed enzyme. The bacterial P450s CYP107Z and CYP105D, GDH, YkuN and FdR were expressed and purified as described previously. [51, 105, 133]

E. coli cells were harvested by centrifugation (30 min, 3220 x g, 4°C) and frozen at -20°C. Prior to their use as whole cell biocatalysts, the cell wet weight (cww) was adjusted to 70 g/L (corresponding to ~ 15 g/L cell dry weight (cdw)) in resuspension buffer (80% 50 mM K_2HPO_4, 20% 50 mM KH_2PO_4, pH 7.5, 500 mM D-glucose, 0.1 mM IPTG). This freeze-thaw cycle should ease mass transfer across the individual modules' membranes by increasing cell wall permeability.[24]

Enzyme activity determination

Prior the inclusion into the cascade, the activity of individual enzymes for which substrates are commercially available was tested in in 2 mL reaction tubes in a reaction volume of 500 µL composed by the 70 g/L recombinant *E. coli* cell suspension with 200 µM substrate dissolved in DMSO (final DMSO concentration 2% (v/v)). Reaction tubes were incubated with open lids at 25°C, 1500 rpm in ThermoMixer C (Eppendorf, Germany) and samples for LC/MS analysis were taken at selected time points.

Given that (-)-pluviatolide **2** is not commercially available, OMT3 activity was screened via a two-step enzymatic cascade starting from (-)-matairesinol **1**, with coexpression of ATR2 to sustain the activity of CYP719A23.

Biotransformations for synthesis of (-)-deoxypodophyllotoxin **6** and (-)-epipodophyllotoxin **7** starting from (-)-matairesinol **1** were carried out in 100 mL Erlenmeyer baffled flasks filled with 10 mL of 70 g/L *E. coli* cell suspension and 200 µM substrate, incubated at 25°C, 250 rpm in orbital shaker Multitron (Infors HT, Germany). 500 µL aliquots were taken at respective time points to determine substrate depletion, intermediate formation and consumption, and product concentration over time.

In case of the two-cells approach, 10 mL of 70 g/L *E. coli* C41 (DE3) coexpressing CYP719A23, ATR2, OMT3 and CYP71CU1 were employed as module one. After 3 h conversion, 10 mL of 70 g/L *E. coli* C41 (DE3) coexpressing OMT1 and 2-ODD (module two) were added to convert (-)-5'-desmethyl-yatein **4** to (-)-deoxypodophyllotoxin **6**. Later, module two was further enhanced by additional coexpression of CYP82D61 and ATR2 (with OMT1 and 2-ODD) to extend the reaction towards (-)-epipodophyllotoxin **7**.

Evaluation of P450s expression and CYP107Z and CYP105D activity against (-)-deoxypodophyllotoxin *in vitro*

Thawed cell pellets were resuspended in 5 mL 50 mM potassium phosphate buffer (pH 7.5) supplemented with 100 µM phenylmethylsulfonylfluoride (PMSF). Cell disruption was done via sonication (Sonifier 250, Branson, USA) and soluble and insoluble protein fractions were separated by centrifugation (40.000 x g, 25 min, 4°C).

P450 concentrations were calculated by recording CO-difference spectra as described elsewhere with extinction coefficient ε_{450nm} = 91 mM^{-1} cm^{-1}.[135] Two measurements were executed with every biological replicate.

Concentrations of purified FdR and YkuN were calculated referring to previously reported[133] extinction coefficients: FdR, ε_{456nm} = 7.01 mM^{-1} cm^{-1}; YkuN, ε_{461nm} = 10.01 mM^{-1} cm^{-1}. GDH

activity was estimated from the generation of NADPH during glucose oxidation at 340 nm. The extinction coefficient ε_{340nm} = 6.22 mM^{-1} cm^{-1} was used to calculate volumetric activity (U/mL).

The *in vitro* reaction was set with 0.5 µM P450, 0.5 µM FdR and 5 µM YkuN (P450:FAD:FMN ratio = 1:1:10). Cofactor regeneration was ensured by addition of 5 U/mL GDH and 20 mM D-glucose. 600 U/mL of bovine liver catalase (Merck, Germany) were also added to prevent the potential generation of reactive oxygen species from uncoupling reactions, *In vitro* experiments were performed in 250 µL volume with 200 µM NADPH 200 µM (-)-deoxypodophyllotoxin dissolved in DMSO (2% final DMSO concentration (v/v)) in 2 mL reaction tubes at 25°C, 1500 rpm in a Thermomixer C (Eppendorf, Germany) in triplicate.

Metabolite analysis

Prior to metabolite extraction 200 µM of (+)-sesamin was added as internal standard (IS) to the reaction solution. Metabolites were extracted twice with 1 mL ethyl acetate. The organic phases from the two extractions were combined and evaporated. The residues were resuspended in 50 µl methanol prior to analysis. The biotransformations of (-)-matairesinol **1** to (-)-deoxypodophyllotoxin or (-)-epipodophyllotoxin - through the intermediates (-)-pluviatolide **2**, (-)-5'-desmethoxy-yatein **3**, (-)-5'-desmethyl-yatein **4** and (-)-yatein **5** - were analysed via liquid chromatography coupled with mass spectrometry (LC/MS) on an LCMS-2020 system (Shimadzu, Japan) equipped with a Chromolith® Performance RP-18e column (100 x 4.6 mm, Merck, Germany). The column temperature was kept at 30°C and 1 µL sample was injected. Separation was carried out with a mobile phase gradient constituted of methanol and double deionized water (ddH_2O) with 0.1% formic acid and a flow rate of 0.5 mL/min or 0.8 mL/min depending on the respective method (Supplementary Table S5).

Once resolved, compounds were detected in the UV/Vis spectrum at 280 nm first. Further, ionization was performed by Dual Ion Source, simultaneously combining electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI) for MS generation. Desolvation and block temperatures were kept at 275°C and 400°C, respectively. Nebulization

gas flow was set at 1.5 L/min whereas the drying gas flow set at 15 L/min. The ionized masses of the analytes were monitored in the positive ion mode between 159 – 1000 m/z. The analytes were identified by their retention times (RT) and typical m/z fragmentation patterns in comparison to commercially available authentic standards or literature references (Supplementary Table S7). From LC/MS measurements, in both MS and UV/Vis detection modes, the product/intermediate distribution was estimated setting all relevant peak areas except the substrate to 100% (Supplementary Table S7). Concentrations of the final product (-)-deoxypodophyllotoxin **6** were calculated via an internal calibration using (+)-sesamin as IS (Supplementary Table S7). All values represent the average from biological and technical duplicates at least.

Acknowledgments

This work was funded by the Federal Ministry of Education and Research to Heinrich-Heine University Düsseldorf [grant number 031B0362A] under the umbrella of the “Nationale Forschungsstrategie BioÖkonomie 2030” project “LignaSyn”. The authors thank Beyza Bedel and Bernd Selting for their contributions to preliminary expression of OMT1 and 2-ODD, and OMT3 and CYP71CU1, respectively.

2.3.1 Supporting Information

omt1 **(GenBank KT390155.1), native DNA-sequence**

ATGGATACTAGGGCTGATGCTGAGATTAAAGCAATGGAGCTGATTGGTATTGGAGTACTTCCACTGGC
AATGAAGGCGATAATTGAGCTCAATGTGTTAGAGATCCTATCAAAAGCAGGACCAGATACCCAACTCA
CTGCTGCTCAAATTGTCACCGACATACCCACCACCAACCCTAACGCTGGTTTCCAACTAGATCGAATT
TTACGACTACTAGCAAGTCATTCAGTCTTGTCTAGTAGTATTACAAAATCGGGGGAGAGAGTGTATGG
GCTAACCCCTATGTGCAAATACTTTCTCCCAGATCAAGATGGAGTCTCACTAGCACCTATGGTTGTTA
CCATCCATGACAAAGTGCTGCTTCAAAGTTGGCATTATCTTAAGGACTCTGTTTTAAAACAAGGCTCT
TTGCCATTTACCGAGGCCTTTGGGATGTCACCCTTTGAGTATTCTGTCTCCGATACAAGGTTTAATAA
GGTTTTTAATGCTGGCATGTTTGACCATTCTACTCTTTGTATGAGGGATGTCCTTCAACGGTACAAAG
GATTTCAAGGCTTGGGGGAGCTGGTTGATGTTGGTGGAGGAACTGGCGGATCGTTGAAGATGATTCTT
TCTCAGTACCCCAATCTAAAGGGCATCAATTTTGATCTCCCACATGTGGTTGCCGACGCGCCTTCTTT
TCCTGGTGTGAAGCACATTGGTGGTGATATGTTTGAGAGTGTTCCCTCTGGTGATGCAATTTTCATGA
AGTGGATACTTCATGACTGGGACGATGGACGTTGCCTGACTTTGCTGAAGAATTGTTGGAATGCATTG
CCAGAGCATGGAAAGGTGATAATAGTGGAGTGGATTCTACCATCAGATGCAGCGACTGACCCAACATC
TCGCCGTGTCTTCACAGCTGATTTGATGATGTTGGCTTTCAGCGAAGGGGGTAAAGAGCGAACCTTGG
GTGACTACGGAGCACTTGCAAAGGAAGCTGGTTTTACCACTGTCAAAGATTTCCCTTGCGCAAATGGC
ATTTCAGTCATTGAGTTCCATAAGAAGTGA

omt3 **(GenBank KT390157.1), native DNA-sequence**

ATGGAAATGGCTCCAACAATGGATTTAGAGATAAGAAATGGAAATGGTTATGGTGATTCTGGAGAGGA
GCTTCTAGCAGCACAAGCTCACATATACAACCACATATTCAACTTCATAAGCTCGATGGCACTGAAAT
GTGCAGTGGAGTTAAACATACCAGAAATTCTCCACAACCATCAACCCAAAGCGGTTACTCTCTCTGAA
CTAGTACAGGCCCTTCAAATCCCCCAAGCAAAATCCGCGTGTCTGTATCGCCTGTTGCGAATACTAGT
CCATTCTGGCTTCTTTGCCATAACGAAAATACAAAGCGAGGGAGATGAAGAGGGTTATTTACCAACCC
TTTCCTCTAAACTACTACTGAAAAACCATCCCATGAGCATGTCTCCATGCTTGTTAGGACTGGTGAAT
CCTACAATGGTAGCACCCATGCATTTCTTTAGTGATTGGTTCAAGAGGAGTGATGATATGACGCCGTT
TGAGGCGACGCATGGAGCGAGCTTGTGGAAGTATTTTGGTGAAACCCCACACATGGCGGAGATATTTA
ATGAGGCAATGGGTTGTGAGACAAGGTTGGCGATGAGTGTGGTGTTGAAAGAGTGTAAGGGCAAGCTT
GAAGGAATAAGTTCGTTAGTTGATGTAGGAGGTGGTACAGGAAACGTGGGTCGGGCAATTGCTGAAGC
CTTCCCAAATGTCAAGTGCACCGTGTTAGATTTGCCACAAGTTGTTGGAAACTTGAAAGGCAGTAACA
ATTTGGAGTTTGTCAGTGGGGATATGTTTCAATTTATTCCGCCTGCAGACGTAGTTTTCTTGAAGTGG
ATATTGCATGATTGGAATGATGAGGAATGTATAAAAATCCTAAAGAGGTGCAAGGAAGCGATTCCATC
CAAGGAAGAGGGAGGGAAATTGATCATAATAGACATGGTAGTAAACGACCACAACAAGGGAAGCTATG
AGTCTACAGAAACGCAACTCTTCTATGATTTGACGCTCATGGCTCTGTTGACAGGAACAGAGAGAACC
GAAACTGAATGGAAGAAGCTCTTCGTAGCTGCTGGTTTCACAAGTTACATTATTAGCCCTGTTTTGGG
GCTCAAGTCTATCATTGAAGTGTTTCCCTAA

cyp71cu1 (GenBank KT390172.1), native DNA-sequence (*cyp71cu1nat_wt*)

N-terminally truncated part of the DNA-sequence (*CYP71CU1Δ20nat*) is highlighted in bold and underlined

ATG**GAAACATTTCAGTGCCTCACACTTTTCCTTCTCTTCATCTCCACTGTTTTCATCCTCAAG**AGAAA
ATTCTCTCATAAACCAAACCTACCACCATCACCACCAAAGCTACCAATCCTAGGCAACTTTCATCAGC
TAGGCACACTTGTCCACCGAGCTGTTACAGTTCTCGCAGCCAAATATGGCCCCCTCATGCTCTTACAC
TTTGGCAAAACTCCAGTTCTCATTGTTTCCTCACAAGAAACGGCTAAAGAGATCATGAAAACCCATGA
CCTCGCTTTGGCAAATAGACCCCTAACAACTGCTGCCAGGGCACTACTTTACGACTGCACAGATATAT
CTTTTGCACCTTATGGTGAGTACTGGAGAGAAATGAAAAAGATGGCAGTCCTAAACCTTCTTAGTATC
AAAAAAATCCAGTCGTTTCGATCCGTTAGGGAGGAATTGGCTTCTGATATGATAAAGGAGATTACTCG
TTTGTCCAAAACTGGAGCACCTGTGGATGTAACTAATATGTTATATCATTTTTCCGAAGACCTGCTTT
TTAGGTGTACTCTTGGGTTTAAACCTAAGGGACAACACAAGTTTCAGAAGCTGTCCAGGGATTTTTTG
GATCTTGTAGGAGCTTTTTGTTTTAATGATTTCTTCCCAGGGATGGCATGGATGGATGCTCTCACTGG
GCTAAACAGGAAGCTGAAGAAGGGTTCAAGAGAATTAGATGACTTTGTTGATGAACTCATAGAAGAAC
GCATTGCTATGGTAAAAGATGGCGTTGAGCCAAATGAATTTCTGGATCTTCTACTCCATACTCATAGA
GACACCACCCAGGAAATCAAACTCACCCGAGACAACGTCAAAGCAATAATATTGGACACATTTCTTGG
TGGAATTGATCTACCAGCATCAGTCATGGAATGGGCAATGGCCGAGCTTATGAGGAATCCAAGTAAGA
TGAAGATAGCTCAGGAAGAGGTCAGAAAAGTGGTTGGAAACAAAAACAAGGTAGACGAAGACGATGTG
TATCAAATGAATTTCTTGAAATCGGCTGTTAAGGAAACTCTAAGGCTACACCCACCAGCTCCTTTGCT
ATTTGCGAGAGAGTCATATACAAGTATAAACGTCGAGAACTACATCATTCCTCCTTACACTAGTGTCA
TGATCAACATCTGGCATATCCAAAGGGACCCCAAGTTATGGGACAAGGCCGAAGAGTTTATTCCAGAG
AGATTCATGAACAGCGGGATTGATTACAAATCCCATGACTATGAATTCATCCCTTTTGGATCCGGACG
AAGAGGTTGCCCTGGTATGTCGTTTGGCGTTGCAGCAGTGGAGTTTGCTGTTGCCAATCTCTTATACT
GGTTTGATTGGAAGTTTGTTGGTGATACAACTCCTGAAACACTTGATATGACTGAGGACTATTGTTTT
GCCCTCTTTAAGAAAAAACCTTTGCATTTTATTCCTATTTCTCGATCCTCTTGA

cyp71cu1, *E. coli* codon optimized DNA-sequence (c*yp71cu1oc_wt*)

N-terminally truncated part of the DNA-sequence (*cyp71cu1Δ20oc*) is highlighted in bold and underlined

ATG**GAAACCTTCCAGTGCCTGACCCTGTTCCTGCTGTTCATCTCTACCGTTTTCATCCTGAAA**CGTAA
ATTCTCTCACAAACCGAACCTGCCGCCGTCTCCGCCGAAACTGCCGATCCTGGGTAACTTCCACCAGC
TGGGTACCCTGGTTCACCGTGCTGTTACCGTTCTGGCTGCTAAATACGGTCCGCTGATGCTGCTGCAC
TTCGGTAAAACCCCGGTTCTGATCGTTTCTTCTCAGGAAACCGCTAAAGAAATCATGAAAACCCACGA
CCTGGCTCTGGCTAACCGTCCGCTGACCACCGCTGCTCGTGCTCTGCTGTACGACTGCACCGACATCT
CTTTCGCTCCGTACGGTGAATACTGGCGTGAAATGAAAAAAATGGCTGTTCTGAACCTGCTGTCTATC
AAAAAAATCCAGTCTTTCCGTTCTGTTCGTGAAGAACTGGCTTCTGACATGATCAAAGAAATCACCCG
TCTGTCTAAAACCGGTGCTCCGGTTGACGTTACCAACATGCTGTACCACTTCTCTGAAGACCTGCTGT
TCCGTTGCACCCTGGGTTTCAAACCGAAAGGTCAGCACAAATTCCAGAAACTGTCTCGTGACTTCCTG
GACCTGGTTGGTGCTTTCTGCTTCAACGACTTCTTCCCGGGTATCTTGGATGGACGCTCTGACCGGTC
TGAACCGTAAACTGAAAAAAGGTTCTCGTGAACTGGACGACTTCGTTGACGAACTGATCGAAGAACGT
ATCGCTATGGTTAAAGACGGTGTTGAACCGAACGAATTTCTGGACCTGCTGCTGCACACCCACCGTGA
CACCACCCAGGAAATCAAACTGACCCGTGACAACGTTAAAGCTATCATCCTGGACACCTTCCTGGGTG
GTATCGACCTGCCGGCTTCTGTTATGGAATGGGCTATGGCTGAACTGATGCGTAACCCGTCTAAAATG
AAAATCGCTCAGGAAGAAGTTCGTAAAGTTGTTGGTAACAAAAACAAAGTTGACGAAGACGACGTTTA
CCAGATGAACTTCCTGAAATCTGCTGTTAAAGAAACCCTGCGTCTGCACCCGCCGGCTCCGCTGCTGT
TCGCTCGTGAATCTTACACCTCTATCAACGTTGAAAACTACATCATCCCGCCGTACACCTCTGTTATG
ATCAACATCTGGCACATCCAGCGTGACCCGAAACTGTGGGACAAAGCTGAAGAATTTATCCCGGAACG
TTTCATGAACTCTGGTATCGACTACAAATCTCACGACTACGAATTTATCCCGTTCGGTTCTGGTCGTC
GTGGTTGCCCGGGTATGTCTTTCGGTGTTGCTGCTGTTGAATTTGCTGTTGCTAACCTGCTGTACTGG
TTCGACTGGAAATTCGTTGGTGACACCACCCCGGAAACCCTGGACATGACCGAAGACTACTGCTTCGC
TCTGTTCAAAAAAAAACCGCTGCACTTCATCCCGATCTCTCGTTCTTCTTAA

2-odd (GenBank KT390173.12), native DNA-sequence

ATGGGTTCTACAGCACCCCTAAGGCTTCCAGTTATAGATTTATCCATGAAGAACTTGAAGCCTGGAAC
AACTTCTTGGAACTCGGTACGCACCCAGGTACGGGAGGCACTGGAAGAATACGGTTGCTTTGAAGCTG
TGATCGATGCTGTGTCTCCAGAGCTGCAGAAGGCAGTATGTAACAAAGGACACGAGCTGCTTAATCTT
CCATTAGAAACCAAGATGTTGAACGGAAACAAACCAGAATATGATGGATTTACGTCAATACCAAACCT
CAACGAAGGCATGGGAGTCGGCAGAATAACAGATTTGGAAAAAGTTGAGAGGTTCACTAATCTTATGT
GGCCCGAGGGGAATAAGGATTTCTGTGAAACTGTGTATTCTTATGGCAAACGAATGGCGGAGGTGGAC
CACATATTGAAAATGATGGTTTTCGAGAGTTTTGGAATGGAGAAGCACTTCGACTCGTTCTGCGAATC
AACAAATTACCTTCTCCATTTCATGAGATACCAACAACCAGGGAAGGATGGACGTTCACCTGCTCTTT
CGTTGCATAAGGACAAGAGCATCTTGACCATAGTAAACCAAAATGATGTCAAGGGATTGGAATTTGAA
ACCAAGGATGGAGAATGGATTTTACCTACAGCTGACAACCATATTGTTCTTCTAGGAGACTGCTTCAT
GGCATGGAGCAATGGTAGATTACATAGTCCTCTTCACCGGGTCACGTTGGTCGCGAACCAGGCGAGGT
TATCTACATCATCGTTTTCGTTTCCAAAGGACATAATAGAGACCCCTGCAGAGCTGGTGGATGAAGAG
CATCCTTTGCTATTTAATCCCTTTGAGATAACGGAGTTGCTTGCTTACTGTTTCACAAAAGAGGGTGC
AAAGGCGGTGTGTGACCTCAAGCAATACAAGGCGTACACAGGTGCATGA

cyp82d61, *E. coli* codon optimized DNA-sequence (c*yp82d61oc_wt*)

N-terminally truncated part of the DNA-sequence (*cyp82d61*Δ*23oc*) is highlighted in bold and underlined

(for native DNA-sequence of *cyp82d61* gene see GenBank KC110995.1)

ATG**GACTCTCTGCACTGCCTGGAAACCCTGCTGCTGGGTTTCTTCGTTCTGCTGCCGTGCTTCTTCTA**
CTTCGTTTGGAAAAAACCGAACAACAAAATCAAAGAACCGCCGCAGCCGGCTGGTGCTTGGCCGATCA
TCGGTCACCTGCACCTGCTGGCTCGTGGTGACCTGCCGCACAAAATCCTGTCTTCTTTCGCTGACAAA
AACGGTCCGGTTTTCAAAATCCAGCTGGGTGTTCACCAGGCTCTGGTTGTTAACAACTCTGAAATCGC
TAAAGAATGCTTCACCACCAACGACCGTTTCTTCCTGAACCGTCCGTCTGGTGTTGCTGCTAAAATCA
TGGGTTACAACTACGTTATGCTGGGTGTTGCTCCGTACGGTCCGTACTGGCGTGACATGCGTAAAATC
ATCATGCTGGAATTCCTGTCTAACCGTCGTCTGCAGTCTCTGAAACACGTTTGGCACTCTGAAATCTC
TATCTCTTCTAAAGAACTGTACAAACTGTGGGAAACCCAGAACATCGACTTCTGCCTGGTTGACATGA
AACAGTGGCTGGCTGACCTGACCCTGAACATGTCTGTTAAAATGGTTGTTGGTAAACGTTTCTTCGGT
TCTGCTTCTGCTTCTGCTTGCGAAGAAACCGAATCTTCTAACTGCCCGAAAACCCTGCGTAACATGTT
CCGTCTGATGGGTTCTTTCGTTCTGTCTGACTACCTGCCGTACCTGCGTTGGCTGGACCTGGGTGGTC
ACGAAAAAGAAATGAAACGTACCGTTAAAGAACTGGACATCCTGTTCAAAGGTTGGCTGGACGAACAC
AAACGTAAACGTCTGTCTGGTGGTAAAGAAGACGACGACCAGGACTTCATGGACGTTATGCTGTCTAT
CCTGGAAGAATCTAAACTGGGTAACGACGTTGACACCATCAACAAAACCGCTTGCCTGGCTATCATCC
TGGGTGGTGCTGACACCACCTGGGCTACCCTGACCTGGGCTCTGTCTCTGCTGCTGAACAACCCGAAC
GCTCTGAAAAAAGCTCAGGACGAACTGGACCTGCACGTTGGTCGTGACCGTAACGTTGACGAATCTGA
CCTGGTTAAACTGACCTACATCGACGCTATCATCAAAGAAACCCTGCGTCTGTACCCGCCGGGTCCGC
TGCTGGGTCCGCGTGTTGTTACCGAAGACTGCACCATCGCTGGTTACCACGTTCGTGCTGGTACCCGT
CTGATCGTTAACGCTTGGAAAATCCAGCGTGACCCGCTGGTTTGGTCTCAGCCGCACGAATACCAGCC
GGAACGTTTCCTGGAACGTGACGTTGACATGAAAGGTCAGCACTTCGAACTGATCCCGTTCGGTTCTG
GTCGTCGTGCTTGCCCGGCTATCTCTCTGGCTCTGCAGGTTCTGCCGCTGACCCTGGCTCACATCCTG
CACGGTTTCGAACTGCGTACCCCGAACCAGAACAAAGTTGACATGACCGAAACCCCGGGTATCGTTCA
CGCTAAAGCTACCCCGCTGGAAGTTCTGGTTGCTCCGCGTATCTCTCCGAAATGCTTCGTTTAA

Expression of *omt3*, *omt1* and *2-odd* genes

Expressions of *omt3*, *omt1*, and *2-odd* were evaluated in *E. coli* BL21 (DE3), C41(DE3) and C43 (DE3). Protein expression was induced at OD_{600} = 0.6 by the addition of 0.5 mM IPTG (additional supplementation with 0.1 mM $FeSO_4$ in the case of 2-ODD), followed by incubation at 25°C, 120 rpm for 20 h.

Expression of wildtype and N-terminal truncated *cyp71cu1* and cyp82d61 gene variants

Expressions were evaluated in *E. coli* C41(DE3). Protein expression was induced at OD_{600} = 0.6 by the addition of 0.5 mM IPTG, with additional supplementation with 0.5 mM 5-aminolevulinic acid (5-ALA) and 0.1 mM $FeSO_4$. Incubation was carried out at 120 rpm for 20 h or 48 h, with temperatures of 20°C, 25°C and 30°C. Cells were harvested by centrifugation (3,220 x g, 4°C, 20 min) and the cell pellet was resuspended in 50 mM potassium phosphate buffer, pH 7.5, 100 µM PMSF. Cells were disrupted by sonication, the cell debris was sedimented by centrifugation (12,300 x g, 4°C ,30 min) and separated from the soluble protein fraction.

P450 concentration within the soluble protein fraction was determined by CO difference spectroscopy according to the method of Omura and Sato[136] to allow comparison between wildtype and truncated versions of CYP71CU1 and CYP82D61.

The results of the screening and optimized expression conditions are summarized in Tables S8, S9, and S10 within this appendix.

Table S1. Summary of generated *cyp71cu1 and cyp82d61* gene variants.

Name of variant	DNA-sequence	Inserted modification (second amino acid)	No. of N-terminal truncated amino acids
cyp71cu1Δ20nat	Native	Ala	1 - 20 (Δ20)
cyp71cu1Δ20oc	*E. coli* codon optimized	Ala	1 - 20 (Δ20)
cyp82d61Δ23oc	*E. coli* codon optimized	Ala	1 - 23 (Δ23)

Table S2. Summary of genes used for establishment of multi-enzyme cascade and (-)-deoxypodophyllotoxin hydroxylation

Gene	Complete name or purpose	Organism	Reference
cyp719aΔ23oc	(-)-pluviatolide synthase (truncated, codon optimized)	*Sinopodophyllum hexandrum*	Decembrino *et al.*, 2020 [120]
atr2	NADPH-cytochrome P450 reductase 2 (ATR2)	*Arabidopsis thaliana*	Kranz-Finger *et al.*, 2018 [133]
omt3	(-)-pluviatolide-O-methyltransferase	*Sinopodophyllum hexandrum*	This work
cyp71cu1Δ20nat	(-)-5'-desmetoxy-yatein hydroxylase (truncated, native DNA-sequence)	*Sinopodophyllum hexandrum*	This work
omt1	(-)-5'-desmethyl-yatein O-methyltransferase	*Sinopodophyllum hexandrum*	This work
2-odd	(-)-deoxypodophyllotoxin synthase	*Sinopodophyllum hexandrum*	This work
cyp82d61Δ23oc	Putative hydroxylation (truncated, codon optimized)	*Sinopodophyllum hexandrum*	This work
ykun	Flavodoxin YkuN	*Bacillus subtilis*	Worsch *et al.*, 2018[104]
fdr	NADP-dependent flavodoxin reductase	*Escherichia coli JM109*	
cyp107z	Putative hydroxylation	*Streptomyces platensis*	Hilberath *et al.*, 2020[51]
cyp105d	Putative hydroxylation	*Streptomyces platensis*	Hilberath *et al.*, 2020[51]

Table S3. Summary of plasmids used within this study.

Plasmid	Features	Copy number	Reference
pETDuet_*atr2_cyp719aΔ23oc*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~ 40	Decembrino *et al.,* 2020 [120]
pCDFDuet_*omt3_cyp71cu1Δ20nat*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_*omt3*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_c*yp71cu1Δ20nat*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_c*yp71cu1Δ20oc*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_c*yp71cu1nat_wt*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_c*yp71cu1oc_wt*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCOLADuet_*2-odd_omt1*	ColA ori, P_{T7}, *lac,* Kan^R	20 - 40	This work
pCDFDuet_*omt1*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pCDFDuet_*2-odd*	CloDf13 ori, P_{T7}, *lac,* Sm^R	20 - 40	This work
pETDuet_*atr2_cyp82d61Δ23oc*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~ 40	This work
pETDuet_cyp82d61oc_wt	ColE1 ori, P_{T7}, *lacI,* Amp^R	~ 40	This work
pETDuet_*cyp82d61Δ23oc*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~ 40	This work
pET24b_*spl06346* (CYP107Z)	ColE1 ori, P_{T7}, *lacI,* Amp^R	~40	Hilberath *et al.*, 2020[51]
pET24b_*spl00625* (CYP105D)	ColE1 ori, P_{T7}, *lacI,* Amp^R	~40	Hilberath *et al.*, 2020[51]
pET16b-*ykuN*(N_{10}His)*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~40	Girhard *et al.*, 2010[105]
pET16b-*fdr*(N_6His)*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~40	Girhard *et al.*, 2010[105]
pET22b_gdh*	ColE1 ori, P_{T7}, *lacI,* Amp^R	~40	Khatri *et al.*, 2015[106]

*Plasmids used for the expression of single enzymes for CYP107Z and CYP105D activity screening *in vitro*

Table S4. Oligonucleotides used for generation of *cyp71cu1 and cyp82d61* gene variants.

Name of variant	**Oligonucleotides 5' → 3'**
cyp71cu1Δ20nat	**forward:** CAGAAAATTCTCTCATAAACCAAAC **reverse:** GCCATATGTATATCTCCTTCTTTATAC
cyp71cu1Δ20oc	**forward:** CCGTAAATTCTCTCACAAAC **reverse:** GCCATATGTATATCTCCTTCTTATAC
cyp82d61Δ23oc	**forward:** CGTTTGGAAAAAACCGAAC **reverse:** GCCATATGTATATCTCCTTCTTATAC

Table S1. LC/MS methods used within this study. Gradients were made of methanol (solvent A) and ddH_2O + 0.1% formic acid (solvent B). A flow rate of 0.5 mL/min was employed with methods 1, 2 and 3, whereas 0.8 mL/min was used in case of method 4.

	Time [min]	**Solvent B [%]**
Method 1	0.01	57
	8.00	28
	13.00	0
	14.01	57
	20.00	57

	Time [min]	**Solvent B [%]**
Method 2:	0.01	80
	5.00	65
	10.00	65
	20.00	38
	25.00	0
	26.01	80
	35.00	80

	Time [min]	**Solvent B [%]**
Method 3	0.01	65
	5.00	50
	8.00	45
	22.00	38
	25.00	0
	26.01	80
	35.00	80

	Time [min]	**Solvent B [%]**
Method 4	0.01	80
	2.00	60
	5.00	55
	25.00	52
	32.50	0
	35.01	80
	40.00	80

Qualitative analysis

Table S6. Authentic reference compounds used for LC/MS identification of reaction products. Commercially not available compounds were identified via their characteristic m/z fragments as described elsewhere.[70]

Compound	Manufacturer or reference	Purity	Molecular weight	RT [min] (method 4)	Characteristic m/z fragments
(-)-matairesinol	Phytolab	≥99% (HPLC)	358.38	9.0	359 $[M+H]^+$ 341 $[M+H-H_2O]^+$
(-)-pluviatolide	*Decembrino et al. 2020* (isolated in-house) [120]	≥95% (HPLC)	360.61	15.0	357 $[M+H]^+$ 339 $[M+H-H_2O]^+$
(-)-yatein	BioBioPha	≥97% (HPLC)	400.40	20.6	401 $[M+H]^+$ 383 $[M+H-H_2O]^+$
(-)-deoxypodophyllotoxin	Toronto Research Chemical	≥98% (HPLC)	398.40	19.8	399 $[M+H]^+$ 421 $[M+Na]^+$
(-)-5'-desmethoxy-yatein	Commercially not available [70]	-	370.40	21.0	371 $[M+H^+]^+$ 353 $[M+H-H_2O]^+$
(-)-5'-desmethyl-yatein	Commercially not available [70]	-	386.40	14.8	387 $[M+H]^+$ 369 $[M+H-H_2O]^+$
(-)-epipodophyllotoxin	Commercially not available [70]	-	414.41	10.4	415 $[M+H]^+$ 437 $[M+Na]^+$
(-)-podophyllotoxin	BLDPharma	≥98% (HPLC)	414.41	-	415 $[M+H]^+$ 437 $[M+Na]^+$
(+)-sesamin (internal standard)	TCI	>98% (GC)	354.35	31.4	337 $[M+H-H_2O]^+$

Underlined m/z fragments correspond to those referred by Lau and Sattely [70] which were used for identification

Quantitative analysis

Table S7. For quantitative analysis, substrate conversion values were calculated by the product distribution resulting from the MS-peak areas of reaction intermediates/products and substrate. Estimation of the concentration of the final product (-)-deoxypodophyllotoxin was done via internal calibration in the range of 10 - 300 µM utilizing 200 µM (+)-sesamin as internal standard.

Conversion [%] Normalized to IS	$= 1- (S_{sample} / IS_{sample}) / (S_{control} / IS_{control})$
Conversion [%]	$= \Sigma(P_{area}) / \Sigma(S_{area} + P_{area}) * 100$
Product distribution [%]	$= P_{area} / \Sigma(S_{area} + P_{area}) * 100$

IS: internal standard; S: substrate; P = product.

Expression of OMT3, OMT1, 2-ODD and CYP71CU1

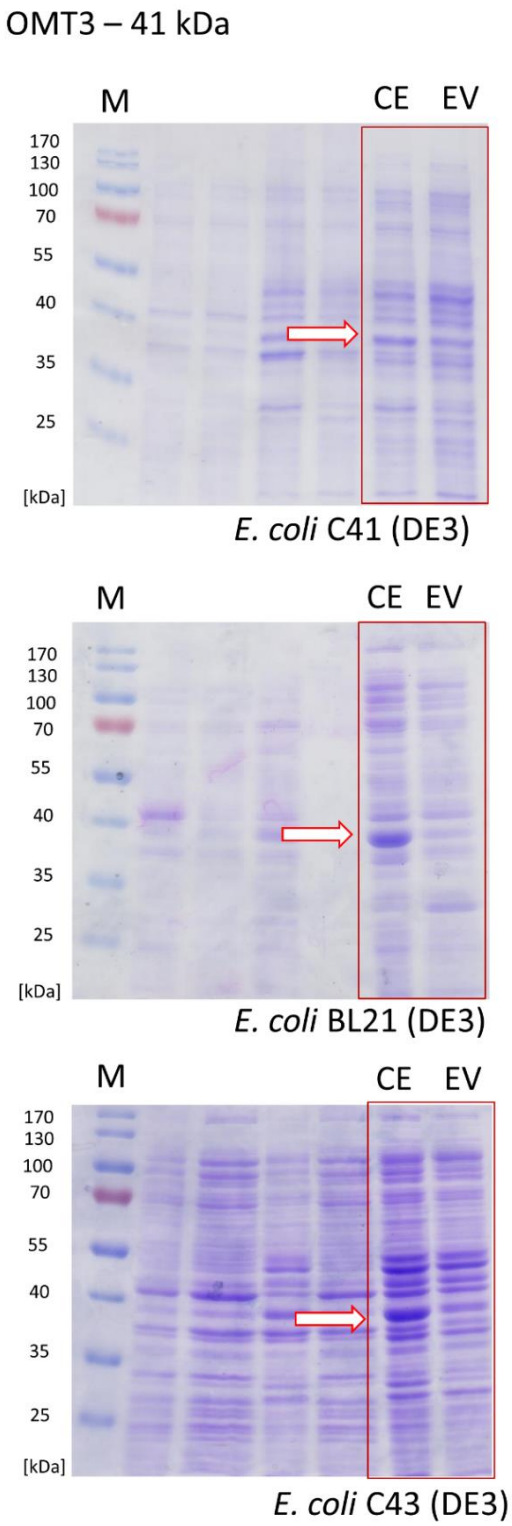


Figure S1. 12.5% SDS-gel of OMT3 (41 kDa; marked by red arrow) expression in various *E. coli* strains. M: Molecular weight size marker, EV: Vector control without *omt3* gene, CE: Soluble protein fraction after cell disruption.

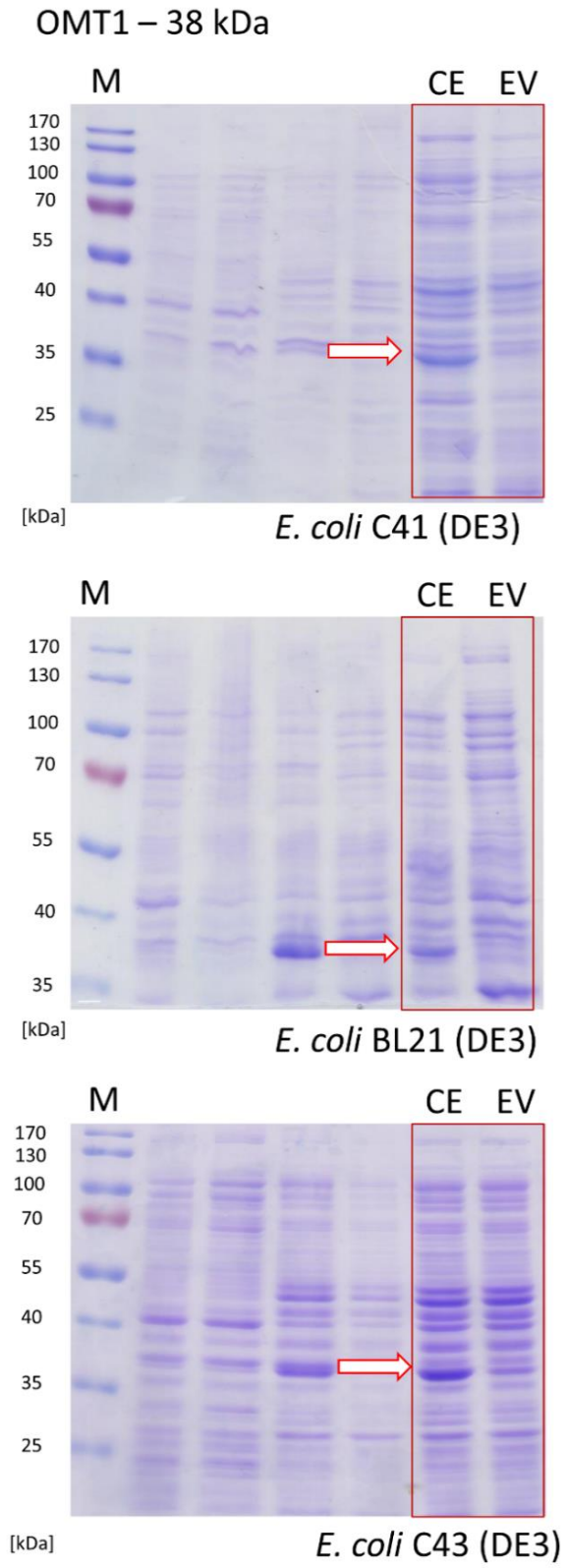


Figure S2. 12.5% SDS-gel of OMT1 (38 kDa; marked by red arrow) expression in various *E. coli* strains. M: Molecular weight size marker, EV: Vector control without *omt1* gene, CE: Soluble protein fraction after cell disruption.

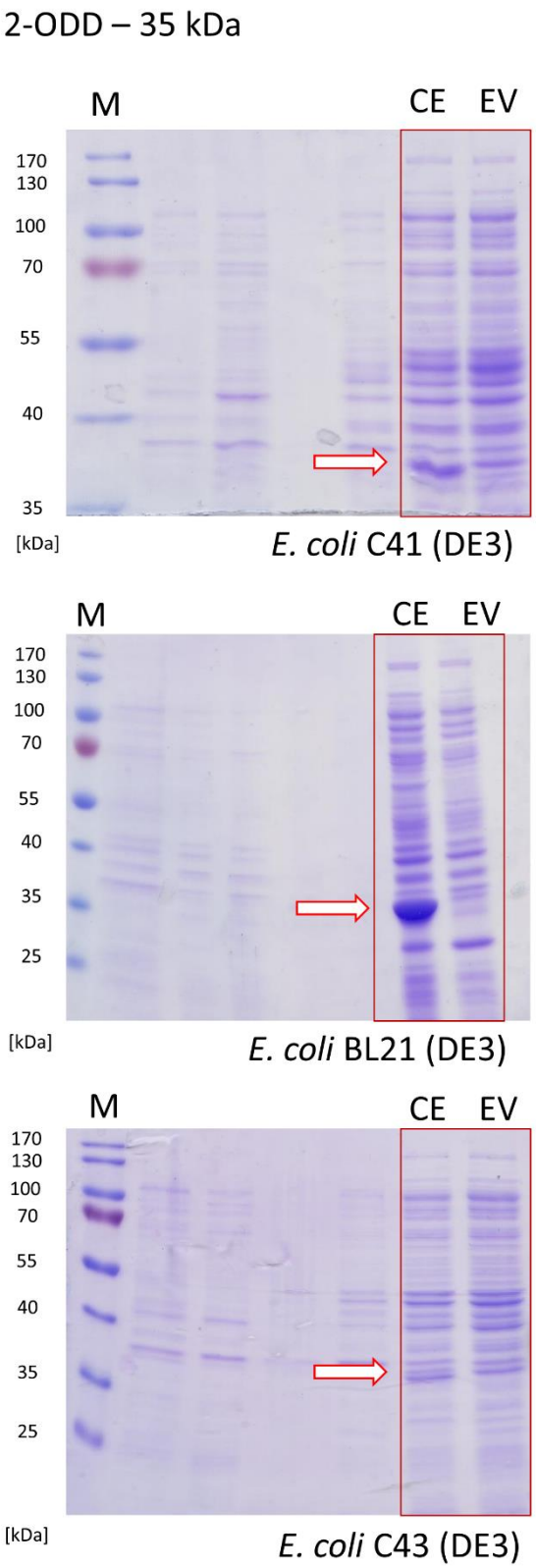


Figure S3. 12.5% SDS-gel of 2-ODD (35 kDa; marked by red arrow) expression in various *E. coli* strains. M: Molecular weight size marker, EV: Vector control without *2-odd* gene, CE: Soluble protein fraction after cell disruption.

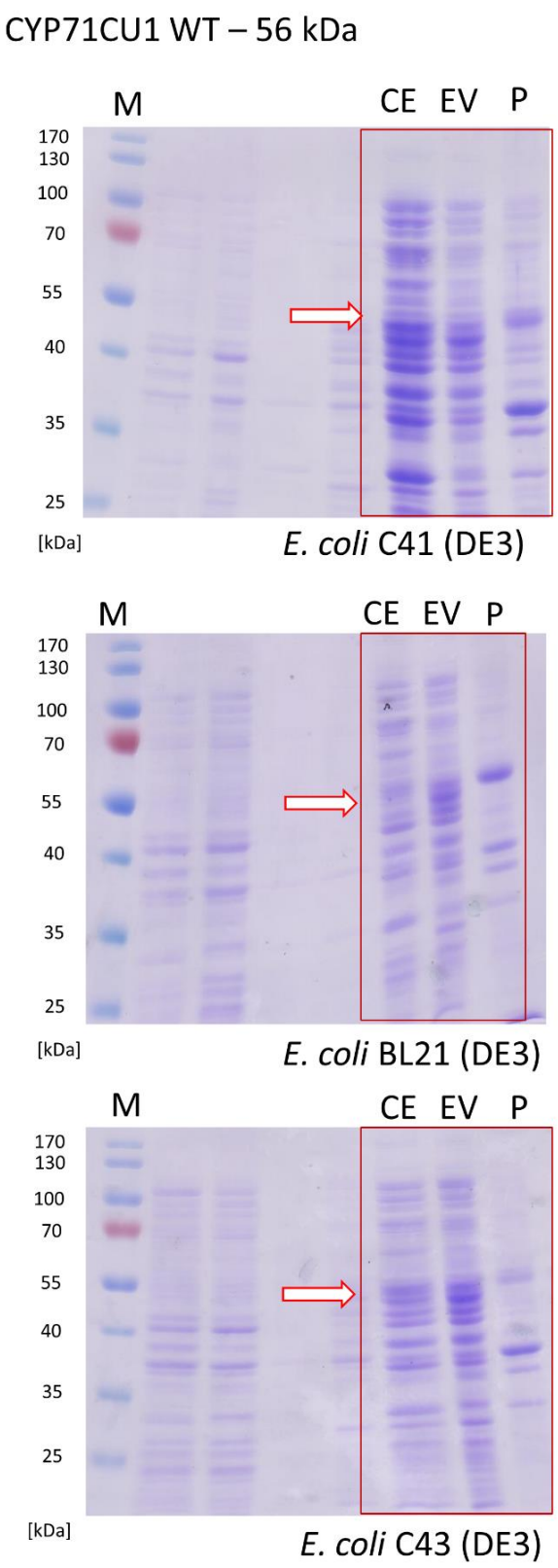


Figure S4. 12.5% SDS-gel of CYP71CU1 (56 kDa; marked by red arrow) expression in various *E. coli* strains. M: Molecular weight size marker, EV: Vector control without *cyp71cu1* gene, CE: Soluble protein fraction after cell disruption, P: Insoluble protein fraction after cell disruption.

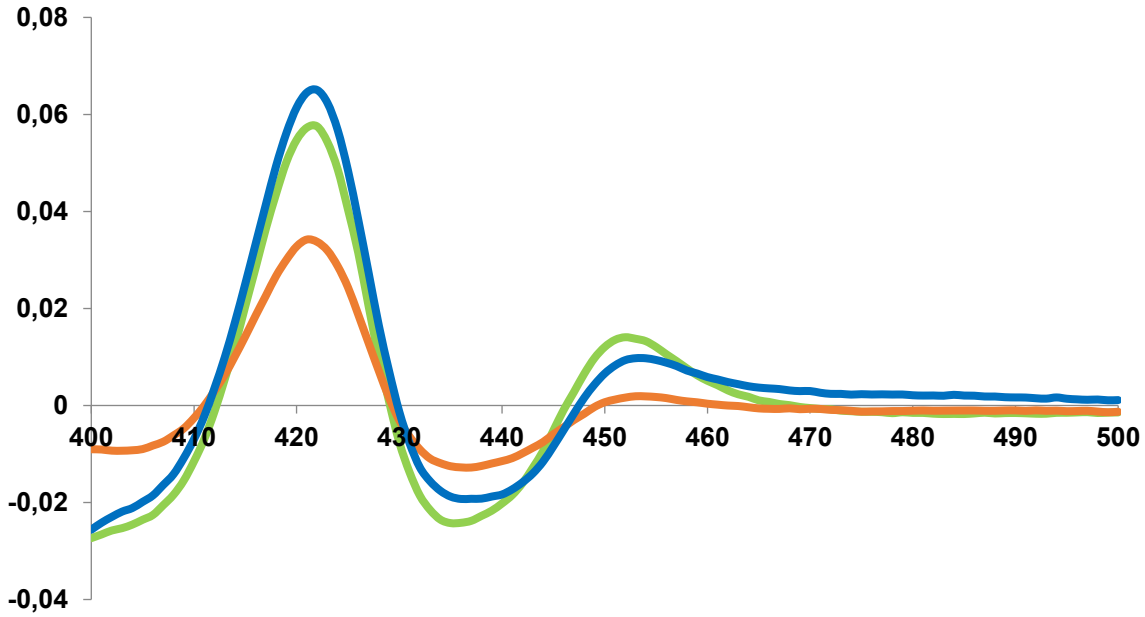


Figure S5. CO-difference spectra of soluble CYP71CU1OC_WT (codon optimized gene without N-terminal truncation) in *E. coli* BL21 (DE3) - orange -, C41 (DE3) - blue -, and C43 (DE3) - green - after 24 h expression.

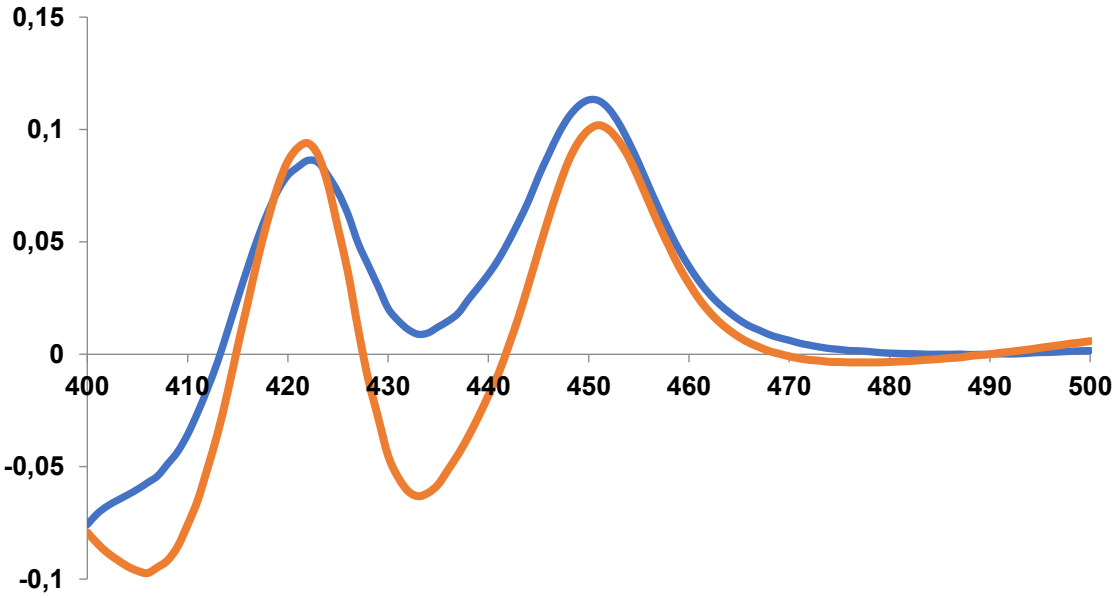


Figure S6. CO-difference spectra of soluble CYP71CU1 in native (CYP71CU1nat_WT - blue) or codon optimised (CYP71CU1OC_WT - orange) gene sequences after 48 h expression.

Table S8. Expression levels of *cyp71cu1* gene variants in *E. coli* C41(DE3) at 25°C after 48 h.

Variant	Inserted modification	Truncated amino acids	P450 concentration	
			[µg/g_{cww}]	[mg/$l_{culture\ volume}$]
cyp71cu1nat_wt	-	-	263 ± 16	8 ± 5
cyp71cu1Δ20nat	Ala	1 - 20	813 ± 163	26 ± 6
cyp71cu1oc_wt	-	-	159 ± 16	6 ± 1
cyp71cu1Δ20oc	Ala	1 - 20	465 ± 138	15 ± 4

Table S9. Expression levels of N-terminally truncated *cyp71cu1* gene variants in *E. coli* C41(DE3) at various incubation temperatures after 48 h.

Temperature	P450 concentration			
	[µg/g_{cww}]		[mg/$l_{culture\ volume}$]	
	CYP71CU1Δ20oc	**CYP71CU1Δ20nat**	**CYP71CU1Δ20oc**	**CYP71CU1Δ20nat**
20°C	282 ± 26	566 ± 39	11 ± 1	20 ± 3
25°C	465 ± 138	813 ± 163	15 ± 4	26 ± 6
30°C	112 ± 18	253 ± 108	6 ± 3	8 ± 3

Table S10. Expression levels of codon optimized WT and N-terminally truncated cyp82d61 gene variants in *E. coli* C41(DE3) at various incubation temperatures after 48 h.

Temperature	P450 concentration			
	[µg/g_{cww}]		[mg/$l_{culture\ volume}$]	
	CYP82D61oc_wt	**CYP82D61Δ23oc**	**CYP82D61oc_wt**	**CYP82D61Δ23oc**
20°C	n.d.	20 ± 17	n.d.	0.7 ± 0.6
25°C	n.d.	936 ± 34	n.d.	33 ± 2
30°C	250 ± 14	1395 ± 220	8 ± 1	42 ± 9

n.d.: Expression not detectable

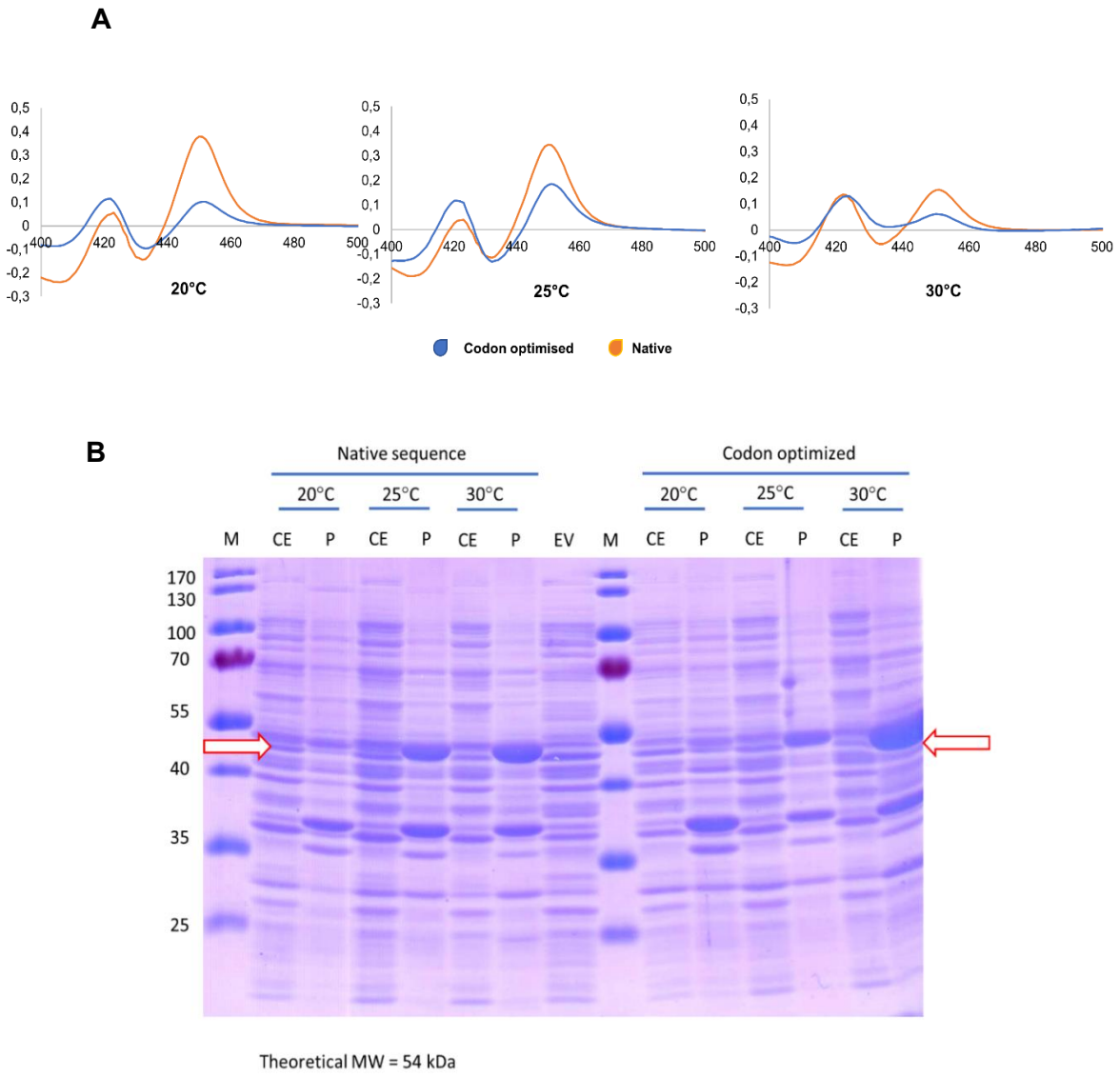


Figure S7. Expression analysis of *cyp71cu1* gene variants incubated at different temperatures. **(A)** CO-difference spectra of c*yp71cu1Δ20nat* (orange) and c*yp71cu1Δ20oc* (blue). **(B)** 12.5% SDS-gel of the expression of c*yp71cu1Δ20nat* and c*yp71cu1Δ20oc* (54 kDa; marked with a red arrow). M: Molecular weight size marker, EV: Vector control without *cyp71cu1* gene, CE: Soluble protein fraction after cell disruption, P: Insoluble protein fraction after cell disruption.

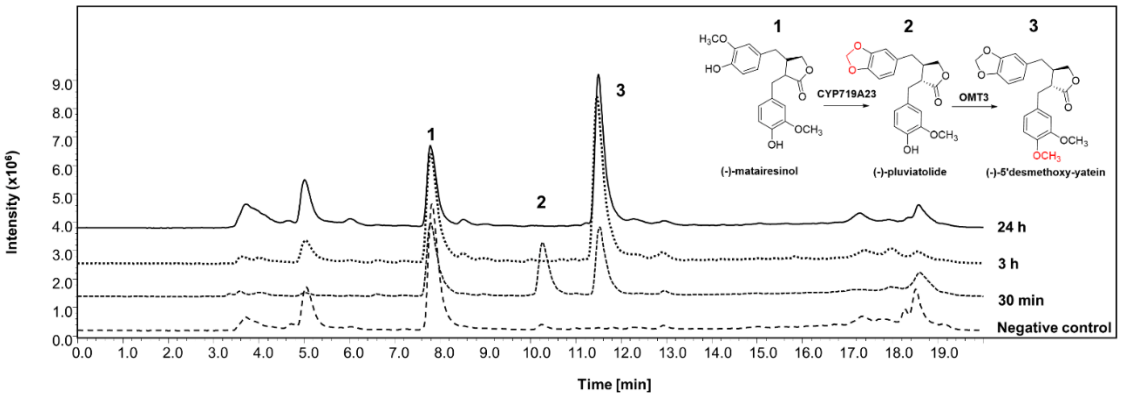


Figure S8. LC/MS analysis (method 1) of OMT3 activity in the cascade reaction with CYP719A23.

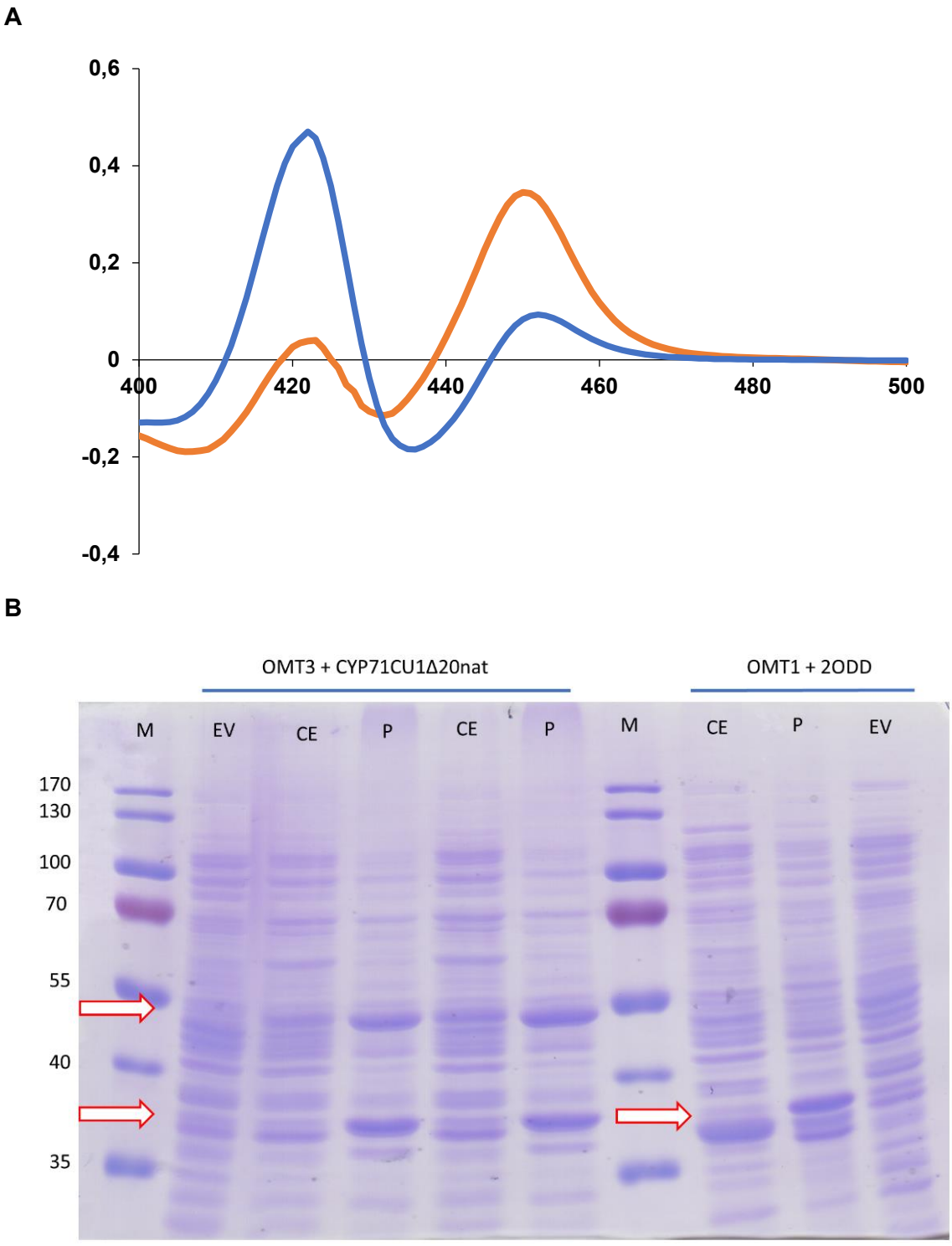


Figure S9. **(A)** CO-difference spectra of single expressed c*yp71cu1Δ20nat* (orange) and c*yp71cu1Δ20nat* coexpressed with *omt3* (blue). **(B)** 12.5% SDS-gel of the coexpressions of c*yp71cu1Δ20nat* (54 kDa) with *omt3* (41 kDa), and *omt1* (38 kDa) with *2-odd* (35 kDa). M: Molecular weight size marker, EV: Vector control without respective genes, CE: Soluble protein fraction after cell disruption, P: Insoluble protein fraction after cell disruption.

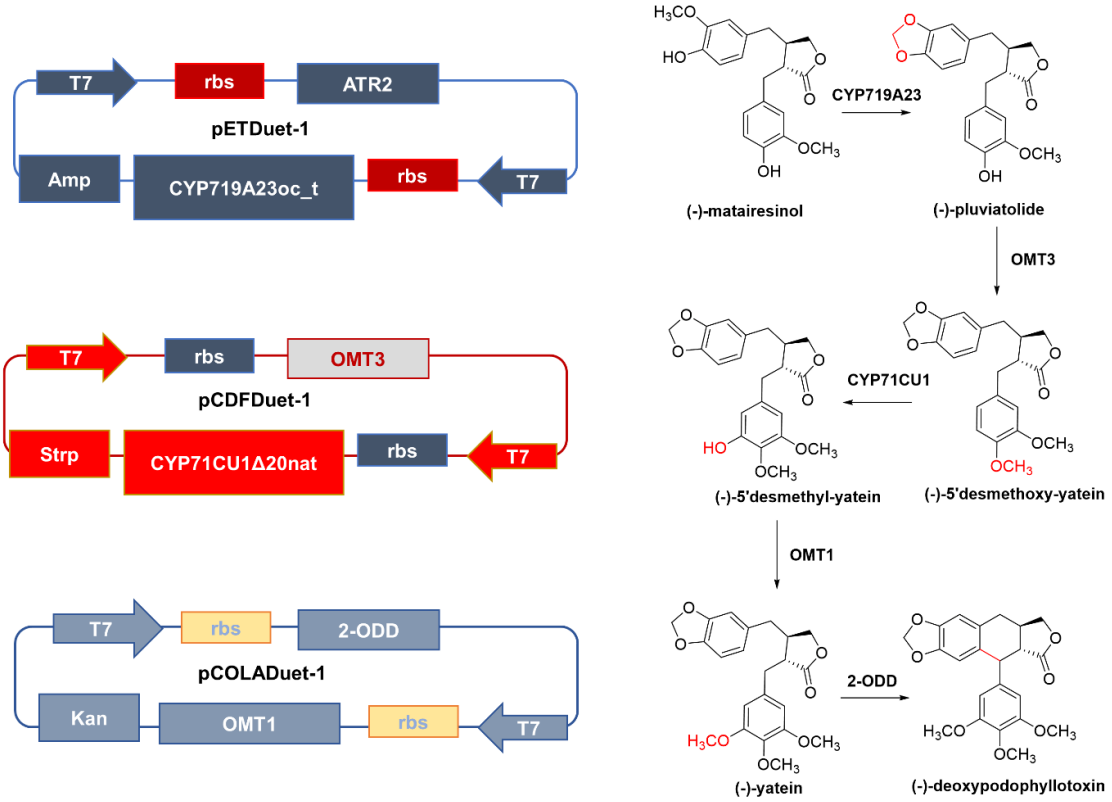


Figure S10. Schematic overview on the plasmid-based modular strategy developed to synthesize (-)-deoxypodophyllotoxin **6** from (-)-matairesinol **1** in *E. coli* via a one-pot 5-steps 6-enzyme reaction cascade.

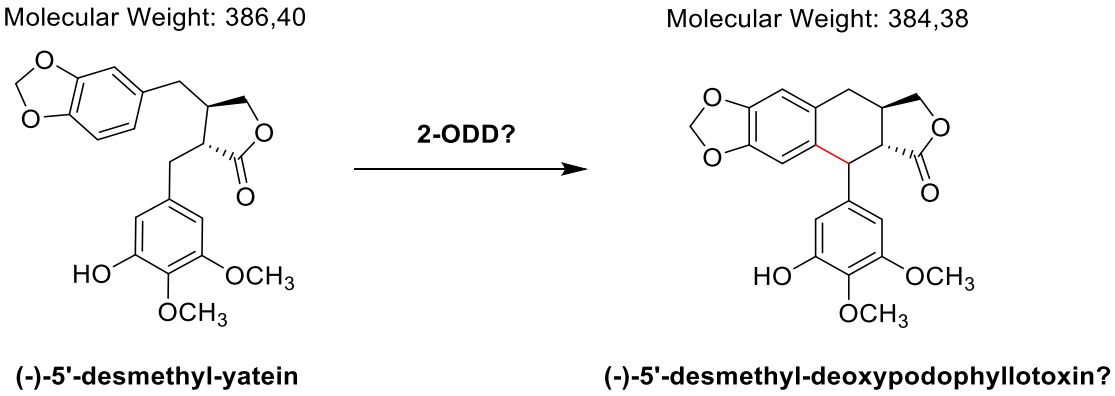


Figure S11. Hypothesised side reaction of 2-ODD occurring in the one-pot two-cells setup.

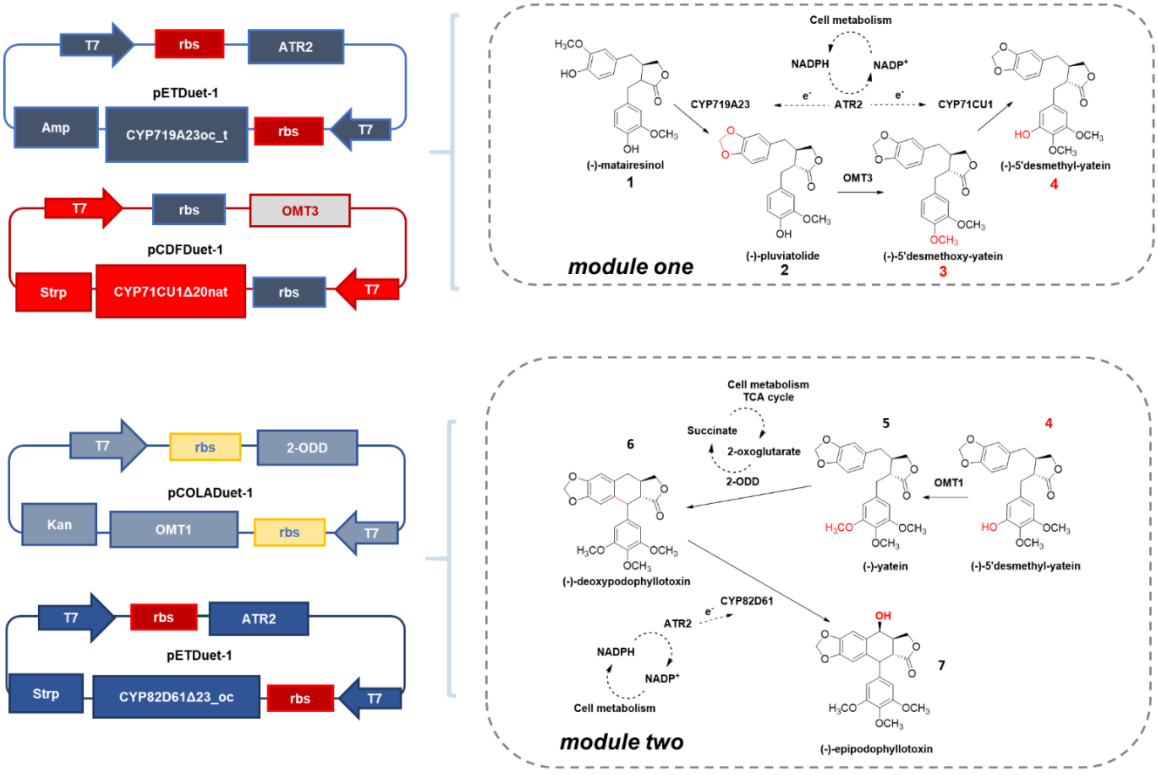


Figure S12. Schematic overview on the plasmid-based modular strategy developed to synthesize (-)-epipodophyllotoxin **7** from (-)-matairesinol **1** in *E. coli* via a one-pot two-cell 6-steps 8-enzyme reaction cascade.

3 General discussion

Within this PhD project parts of the biosynthetic pathway towards the plant secondary metabolites lignans were reconstituted in *Escherichia coli*. A modular approach was developed to target specific lignans such as (+)-pinoresinol, (-)-matairesinol, (-)-pluviatolide and (-)-deoxypodophyllotoxin. In fact, because of the numerous biological activities, some lignans are endorsed as valuable APIs, for which the development of versatile and sustainable production processes is highly valuable.

3.1 Enzymes for the reconstitution of lignans biosynthesis

The effective modular reconstitution is the result of the successful selection and implementation of appropriate enzymatic candidates. Among these, genes originating from *Sinopodophyllum hexandrum*, *Podophyllum pleianthum*, *Forsythia intermedia* and *Arabidopsis thaliana* were expressed in *E. coli* in active state. Along with plant enzymes, the *E. coli*'s endogenous multi-copper oxidase CueO was employed to catalyse coniferyl alcohol coupling.

CueO-mediated coupling

In this work, the activity of CueO was optimized regarding the coupling of coniferyl alcohol to furnish the parent lignan (±)-pinoresinol. Specifically, this enzyme was employed to substitute plant laccase or peroxidase that catalyse this reaction in plants, as well as the recombinant laccase CgL1 from *Corynebacterium glutamicum* which was employed in a previous study of our group for the same purpose.[86] CueO catalytic activity was found to be switched on *in vivo* by the simple addition of copper salts. By doing this, the host *E. coli* became active part of the multi-enzyme cascade, performing the first reaction of a 4-step biotransformation to (-)-matairesinol starting from coniferyl alcohol. Copper-triggered single-electron oxidation was proven effective in the dimerization of diverse phenolic compounds other than coniferyl alcohol. Independently from the reaction cascade, regular *E. coli* strains may be optimized as simple biosynthetic devices for the dimerization and/or polymerization of phenolic compounds

with enhanced biological properties compared to the respective monomers.[137] It is the case, for instance, of 2,6-dimethoxyphenol and ferulic acid dimers.[138, 139] Likewise, the oxidation of coniferyl alcohol determines the generation of four coupling products. Among these, (±)-dehydrodiconiferyl alcohol (β-5 dimer) and (±)-pinoresinol (β–β dimer) stand as the two major products.[84] (±)-Pinoresinol is known for its anti-inflammatory, anticancer and antimicrobial properties, as well as the common precursor to various lignans pathways.[64, 66] (±)-Dehydrodiconiferyl alcohol is a phytoestrogen associated to anti-adipogenic and antioxidant properties, which was reported effective on osteoblasts differentiation process.[140] Furthermore, it was associated to strong anti-inflammatory activity, which made it eligible as a potential therapeutic candidate for the control of inflammatory diseases such as multiple sclerosis.[141] For these reasons, (±)-dehydrodiconiferyl alcohol may represent a valuable chemical to be targeted by biosynthetic means instead of being considered an undesired side-product during lignan synthesis. Lignans biosynthetic pathways differentiate their routes via enantioselective enzymes metabolizing either (+)- or (-)-pinoresinol.[63, 95] Therefore, maximizing (±)-pinoresinol production is crucial to increase the yield of the cascade. In absence of Dirigent proteins driving the regio- and stereoselectivity of the reaction,[81] in the best-case scenario, only 15% of the starting coniferyl alcohol would be transformed to the desired pinoresinol enantiomer. In addition, another harming factor to the availability of (±)-pinoresinol is determined by the overoxidation related to laccases/multi-copper oxidases unspecific reaction mechanism.[142]

Both overoxidation and lack of stereo- and regiospecificity during phenolic coupling are undoubtedly major drawbacks pursuing optimal reaction performance. Nevertheless, in this work such issues were limited by tightly controlling the reaction time. This strategy allowed to achieve (±)-pinoresinol close to the theoretical expected value of 250 µM from 2 mM substrate and further implement CueO in a multi-enzyme cascade to (-)-matairesinol. Alternatively, other methods may be used to protect (±)-pinoresinol from laccase-mediated overoxidation. As an example a biphasic reaction system mixing water to non-miscible organic solvents has been

developed in a previous work from our group.[86] Overall, the possibility to "*switch on*" an enzyme is a valuable feature for a multi-enzyme cascade. In this case even more, as it offers a way to control the multi-copper oxidase CueO from undesired side reactions.

Biotransformation of (±)-pinoresinol

As indicated previously, diverse lignans biosynthetic pathways involve enantiocomplementary routes via kinetic resolution of (±)-pinoresinol. In this work, such step is described in both chapters 2.1 and 2.2 and performed by the pinoresinol-lariciresinol reductase from *Forsythia intermedia* to produce (-)-secoisolariciresinol via (+)-lariciresinol through two reduction steps. Further oxidation to (-)-matairesinol is catalysed by the secoisolariciresinol dehydrogenase from *Podophyllum pleianthum*. As highlighted in chapter 2.2, both enzymes resulted in the rapid turnover of the respective substrates. Particularly, (+)-pinoresinol concentrations up to 180 mg/L were transformed to (-)-matairesinol within the first 10 minutes. Strong enantioselectivity towards (-)-secoisolariciresinol was observed for PpSDH, as already annotated in previous studies.[79] Differently, for FiPLR a discrepancy was observed between the enantiospecificity of the two reduction steps. Although the preference toward (+)-pinoresinol reduction is aligned to the previous reports,[87] the depletion of minor amounts of the (-)-enantiomer was observed in our experiments as well. Interestingly, the second reduction only appears truly enantiospecific. All in all, both FiPLR and PpSDH exhibit potential features which prove them as solid gears in the metabolic clockwork designed within this work.

P450s in high lignans biosynthesis

Over time, more and more cytochromes P450 have been described as key elements in plant secondary metabolic pathways. The evolution of these enzymes has been recognized essential to achieve more and more sophisticated metabolites.[53, 62] Also P450s may act as strict branch points along the biosynthesis, channelling the pathways in precise directions with strong enantioselectivities.[143] In lignan metabolism, methylenedioxy bridge-forming enzymes CYP81Q1 and CYP719A23/24 are perfect examples for this. On the one hand, CYP81Q1 from *Sesamum indicum* produces (+)-sesamine after the double oxidation of (+)-pinoresinol

via (+)-piperitol.[95] On the other hand CYP719A23/A24, respectively from *S. hexandrum* and *P. peltatum*, address the transformation of dibenzylbutyrolactones toward aryltetralin lignans such as (-)-deoxy-, (-)-epi-, and (-)-podophyllotoxin.[70, 80] Following this specific route, three other P450s have been identified in *S. hexandrum* genome: CYP71CU1, CYP82D61, and CYP71BE54.[70]

In this work, the 'active' expression of CYP719A23, CYP71CU1 and CYP82D61 was achieved in *E. coli* for the first time. As a general matter, the expression of eukaryotic membrane-associated genes in prokaryotic hosts is not trivial, and P450s are no exception. Because of the lack of inner organelles or well-developed membranes, achieving the heterologous expression of plant P450s in bacteria is a notorious challenge.[123] In fact, eukaryotic P450s are typically membrane bound to the cytoplasm-facing side of the endoplasmic reticulum by their N-terminal domain.[123] Over time, manipulation of the latter has been presented as prompt resolution to this issue.[60] For all three P450s here considered, the analysis of the respective primary structures allowed to recognize the putative transmembrane N-terminal sequence and the presence of a proline-rich region which is reported to drive the correct folding process.[125] The generation of P450 variants where the first 20-26 amino acids were truncated, with the addition of the triplet encoding for alanine as second codon, resulted particularly effective. The expression of the wildtype genes always resulted in little to no enzyme detection, whereas correctly folded, soluble P450s were produced after N-terminus manipulation. However, codon usage was not decisive.

Although remarkable, successful heterologous expression is not the only factor to be considered. As described within the introduction to this thesis (1.2.3), the presence of redox partner proteins is fundamental for the reconstitution of P450s activity. Given that physiological CPRs from *S. hexandrum* remain unidentified to date, different non-native alternatives have been explored within this work. Using CYP719A23 as an example, the use of ATR2 from *Arabidopsis thaliana* as CPR was compared to a fusion construct P450-ATR2, and to a bacterial redox system involving the *E. coli*'s reductase FdR and the *B. subtilis*' flavodoxin

YkuN. Overall, ATR2 was determined as the most suitable candidate for CYP719A23, and it was further proven effective in sustaining the activity of all three P450s.

This evaluation highlighted the importance of appropriate interaction among a P450 and redox partners. The fusion construct was applied since it has been reported the physical linkage between P450 and CPR is likely to enhance substrate conversion.[44, 46] Although this happened to be true *in vitro*, the fusion of CYP719A23 with ATR2 significantly reduced P450 expression, negatively affecting the benefits of this strategy *in vivo*. To this end, harnessing the coexpression of P450 and CPR is crucial to achieve a suitable balance resulting in adequate activity. Although some studies suggest high redox partners ratios to be beneficial, others report detrimental effects on P450 expression levels and substrate turnover, as well as harmed cell viability.[28, 144, 145] In this work, the application of the bacterial redox system indeed affected P450 expression, significantly hampering the activity. Conversely, physiological CYP to CPR ratio are reported to be ~15:1 *in planta.*[127, 128] Thus, lower amounts of redox partners are thought to be sufficient to support the activity of multiple P450s, which seems to be the case within this work. For instance, in chapter 2.3, the sole ATR2 was used to reconstitute the activity of both CYP719A23 and CYP71CU1, resulting in efficient activity of both P450s.

In conclusion, although P450s are often addressed to be a bottleneck within metabolic pathways,[62, 143] in this work efficient implementation was reported. In chapter 2.2, 137 mg/L of (-)-pluviatolide were produced in high purity (>99% UV/vis, >94% MS) with 76% isolated yield through a 4-step cascade from 180 mg/L (+)-pinoresinol. Previously, only 13% isolated yield has been reported after the single oxidation of (-)-matairesinol by CYP719A23 in *S. cerevisiae*.[70] Also, strict enantiospecificity was observed for CYP719A23 within this work. Furthermore, the 6-steps synthesis of (-)-deoxypodophyllotoxin was achieved from (-)-matairesinol with 98% yield assembling CYP719A23, CYP71CU1, and the methyltransferases OMT3 and OMT1, plus the 2-oxoglutarate-dependent dioxygenase 2-ODD. Ultimately, CYP82D61 was introduced into the multi-enzyme cascade to achieve (-)-epipodophyllotoxin production in *E. coli* for the first time.

3.2 Modular design of the multi-enzyme cascade

When designing the assembly of complex multi-enzyme cascades, harnessing the spatial organization of their components becomes crucial for their primary implementation and further optimization.[146] In this work, the 10-steps cascade from the monolignol coniferyl alcohol to dibenzylbutyrolactone and aryltetralin lignans was implemented stepwise. Each chapter of the results section of this dissertation (2.1, 2.2, 2.3) depicts the development of a portion of the metabolic pathway. In each chapter at least one enzymatic step is shared by the previous or the following portion of the pathway. Each portion represents an independent and functional biological device which can be sequentially combined to the others. The coexpression of multiple enzymes was achieved by the design of compatible vectors. Genes were cloned into plasmids flanked by rigorously selected communal restriction enzymes to ensure the maximum combinability to the system. In this way, genes could be easily *"cut-and-pasted"* into any of the chosen expression vectors at need. The design of such units, or *modules*, can be addressed as an application of the synthetic biology principle of "standardized parts development".[147, 148] As an example, Wu and co-workers designed autonomous 2-, 3- or 4-enzyme combinable one-pot cascades to achieve oxy- and amino-functionalization of alkenes with high regio- and stereoselectivity in *E. coli.* [149]

Here, the modularity was developed not only to combine several enzymes together, but also to overcome some of the bottlenecks typical of whole-cell biocatalytic systems. In fact, enzymatic compartmentalization is often reported as effective in relieving pitfalls of metabolic nature.[17, 146, 150] Furthermore, the formation of a so-called metabolon may be hypothesised to occur for some of the enzymes involved within this work. Six of them originate from *S. hexandrum* (CYP719A23, OMT3, CYP71CU1, OMT1, 2-ODD, CYP82D61), whereas two of them (FiPLR and PpSDH) originate from *F. intermedia* and *P. peltatum* and cover the previous steps of the same biosynthetic route. Numerous examples of multi-enzymes complexes have been reported in plants secondary metabolic pathways.[151, 152] However, altered protein-protein interactions may affect such events consequently to the manipulations required for their

implementation in prokaryotic hosts. Moreover, extensive pathways involving more than 10 enzymes have been successfully split among individual host strains co-cultured together to produce salicylate 2-O-β-D-glucoside or *cis*,*cis*-muconic acid in *E. coli*. [153, 154]

In chapters 2.1 and 2.3, separate *E. coli* strains harbouring sequential parts of the pathway were mixed together and exploited as permeable compartments to allow the intermediates transfer. Such two-cells systems were proven particularly effective to overcome side-reactions and metabolic competition deriving from the host.

Considering chapter 2.1, the efficiency of coniferyl alcohol biotransformation cascade to (-)-matairesinol in a *single-cell* setup was likely hampered because of substrate seizure by endogenous alcohol dehydrogenases (ADHs).[155] Despite ~90% of substrate conversion was observed, coniferyl aldehyde soared up as the main product, limiting the formation of the expected alcohol dimers. In this scenario, (+)-pinoresinol availability is drastically reduced and so (-)-secoisolariciresinol and (-)-matairesinol formation by FiPLR and PpSDH. As well as ADHs, both enzymes are NAD(P)H dependent,[63, 79, 155] thus competition over nicotinamide cofactors could be responsible for this behaviour. Differently, notable reduction of the aldehyde, resulting in clear coniferyl alcohol dimerization, was related to the compartmentalization of the cascade in a *two-cells* system. This was especially true when the second *E. coli* strain was added after a certain amount of time. The final product (-)-matairesinol was detected in concentrations close to the theoretical optimum of 250 µM from 5 mM coniferyl alcohol. This result endorses the importance of spatial organization and appropriate reaction timing for optimal performance of multi-enzyme networks, as suggested in other recent works. [146, 156] Complementarily, the side-activity of ADHs could be targeted by metabolic engineering means such as gene knock-out and carbon source feeding-optimization as well.[20]

A similar scenario was described in chapter 2.3. There the implementation of a *two-cells* approach allowed the achievement of 98% yield from the 5-step conversion of (-)-matairesinol to (-)-deoxypodophyllotoxin. Differently, the application of a *single-cell* setup generated a

bottleneck at the last enzymatic step, blocking the conversion of (-)-yatein to (-)-deoxypodophyllotoxin. All in all, the co-localization of three plasmids into the cell is likely increasing the metabolic burden. Furthermore, the presence of redox enzymes such as two P450s and 2-ODD is likely determining resource competition between exogenous and endogenous metabolic pathways.[17]

A typical argumentative advantage of *in vivo* multi-enzyme cascades compared to their *in vitro* counterparts involves the exploitation of the host metabolism for co-substrates supplementation and regeneration.[17] However, the introduction of multiple foreign enzymes may generate the drawbacks described above. Generally, the addition of carbon sources such as glycerol or glucose is often a practical solution to support the host metabolism.[157] Specifically, in this work, glucose was provided to sustain *E. coli*'s pentose phosphate pathway and the TCA cycle for the generation of cofactors and co-substrates either. Yet, feeding strategies need to be optimized to overcome the reported metabolic side effects. Additionally, although the modular multi-cells systems offer practical advantages, other issue may arise. Frequently, one of these is related to inefficient traverse of the intermediates across the membranes. As a matter of fact, mass transfer issues are frequently reported as some of the most common drawbacks in developing efficient whole-cell biocatalysts. [17, 21] In this work, the usage of resting cells was practical to avoid such limitation. As the cells were frozen after harvest and thawed prior to usage, the membranes are physically damaged to grant substrate passage without causing cell lysis, also avoiding the usage of harmful detergents.[24]

3.3 Future perspectives

Over the last few decades, multiple findings helped casting light on *S. hexandrum* secondary metabolic pathway in order to reach the production of (-)-podophyllotoxin by sustainable biotechnological means.[70, 80, 130] Nevertheless, to date the physiological enzyme performing the final hydroxylation of (-)-deoxypodophyllotoxin to yield (-)-podophyllotoxin has yet to be

discovered.[122] In chapter 2.3, the generation of (-)-podophyllotoxin and its epimer was demonstrated via the application of two promiscuous P450s from *Streptomyces platensis,* CYP107Z and CYP105D. Following protein engineering endeavours to enhance substrate turnover, regio- and stereoselectivity,[101, 131, 132] such enzymes may become interesting candidates to patch up the lack of a known (-)-podophyllotoxin synthase. Additionally, compared to their eukaryotic counterparts, bacterial P450s are more easily expressed in recombinant *E. coli*, facilitating their successful implementation.

Complementary to this, CYP82D61 should be investigated together with CYP71BE54 to further extend the cascade. (-)-4'-desmethyl-deoxypodophyllotoxin can be produced from the demethylation of (-)-deoxypodophyllotoxin by CYP71BE54 and further hydroxylated to (-)-4'-desmethyl-epipodophyllotoxin by CYP82D61. All of these compounds are also direct precursors to etoposide and teniposide (Figure 1.6),[70] therefore their production by recombinant means is of high interest.

More genomic investigations on *S. hexandrum* should be done not only in sought for a putative (-)-podophyllotoxin synthase, but also for the identification of cognate CPRs to support the activity of the already implemented P450s. Not only, further gene scavenging should be directed to closely related plants such as *Forsythia intermedia* and *Thuja plicata*, but mainly *Podophyllum peltatum*.[79, 80, 158] The identification and subsequent evaluation of homologous and orthologous gene candidates is undoubtedly a chance to improve the current setup. As a straightforward example, CYP719A24 from *P. peltatum* should be evaluated as alternative for (-)-matairesinol oxidation to (-)-pluviatolide.[80]

Independently from these argumentations, the modules herein developed should be finally combined to produce aryltetralin lignans from coniferyl alcohol. Once done, backwards extension should be implemented to start from even cheaper sources. As a potential route, the 3-step enzymatic cascade starting from eugenol (-)-secoisolariciresinol established in a previous study of our group could be adapted as starting module.[102] Moreover, coherently to the final aim of the LignaSyn project, more efforts should be implemented to extend the

pathway backwards to even simpler precursors. L-phenylalanine, L-tyrosine or even glucose could be used as starting material exploiting genes from shikimate and phenylpropanoids pathways. [159-163]

Within this work, the importance of a balanced gene expression was highlighted concerning the direct interactions between the enzymes and their interplay with the host metabolism. To improve these aspects, gene titration studies testing weaker promoters and diverse gene copy number are valuable strategies chasing optimal metabolic flux.[164] For the same purpose, metabolic engineering efforts should be pursued on the host *E. coli* for the optimization of carbon sources utilization and limiting the competing pathways.[20] Ultimately, high cell density cultivation in bioreactors and process engineering optimization strategies should be applied for reaction upscale. In these ways, environmentally sustainable and economically feasible production of precursors to life-saving pharmaceuticals and health-improving compounds can be achieved finally.

4 References

1. Mensah J. *Sustainable development: Meaning, history, principles, pillars, and implications for human action: Literature review*. Cogent Soc Sci. **2019**; 5**:**1653531.
2. Anastas PT, Horváth IT. *Introduction: Green Chemistry*. Chem Rev. **2007**; 107**:**2167–8.
3. Anastas PT, Kirchhoff MM. *Origins, current status, and future challenges of Green Chemistry*. Acc Chem Res. **2002**; 35**:**686–94.
4. Sheldon RA, Woodley JM. *Role of biocatalysis in sustainable chemistry*. Chem Rev. **2018**; 118**:**801-38.
5. Abdelraheem EMM, Busch H, Hanefeld U, Tonin F. *Biocatalysis explained: from pharmaceutical to bulk chemical production*. React Chem Eng. **2019**; 4**:**1878-94.
6. Nielsen DR, Moon TS. *From promise to practice. The role of synthetic biology in green chemistry*. EMBO Rep. **2013**; 14**:**1034-8.
7. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. *Engineering the third wave of biocatalysis*. Nature. **2012**; 485**:**185-94.
8. Reetz MT. *Biocatalysis in organic chemistry and biotechnology: past, present, and future*. J Am Chem Soc. **2013**; 135**:**12480-96.
9. Fernandes P. *Enzymes in food processing: a condensed overview on strategies for better biocatalysts*. Enzyme Res. **2010**; 2010**:**862537.
10. Pollard DJ, Woodley JM. *Biocatalysis for pharmaceutical intermediates: the future is now*. Trends Biotechnol. **2007**; 25**:**66-73.
11. Pan SY, Litscher G, Gao SH, Zhou SF, Yu ZL, Chen HQ, Zhang SF, Tang MK, Sun JN, Ko KM. *Historical perspective of traditional indigenous medical practices: the current renaissance and conservation of herbal resources*. Evidence-Based Complementary Altern Med. **2014**; 2014**:**525340.
12. Pyne ME, Narcross L, Martin VJJ. *Engineering plant secondary metabolism in microbial systems*. Plant Physiol. **2019**; 179**:**844-61.
13. Shang Y, Huang S. *Engineering plant cytochrome P450s for enhanced synthesis of natural products: past achievements and future perspectives*. Plant Communications. **2020**; 1**:**100012.
14. Stephanopoulos G. *Metabolic fluxes and metabolic engineering*. Metab Eng. **1999**; 1**:**1-11.
15. Gross M. *What exactly is synthetic biology?* Curr Biol. **2011**; 21**:**611-4.
16. Walsh CT, Moore BS. *Enzymatic cascade reactions in biosynthesis*. Angew Chem Int Ed. **2019**; 58**:**6846-79.
17. France SP, Hepworth LJ, Turner J, Flitsch SL. *Constructing biocatalytic cascades: in vitro and in vivo approaches to de novo multi-enzyme pathways*. ACS Catal. **2016**; 7**:**710-24.
18. Wachtmeister J, Rother D. *Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale*. Curr Opin Biotechnol. **2016**; 42**:**169-77.
19. Tufvesson P, Lima-Ramos J, Nordblad M, Woodley JM. *Guidelines and cost analysis for catalyst production in biocatalytic processes*. Org Process Res Dev. **2011**; 15**:**266–74.
20. Pontrelli S, Chiu TY, Lan EI, Chen FY, Chang P, Liao JC. *Escherichia coli as a host for metabolic engineering*. Metab Eng. **2018**; 50**:**16-46.
21. Lin B, Tao Y. *Whole-cell biocatalysts by design*. Microb Cell Fact. **2017**; 16**:**106.
22. Calcott PH, MacLeod RA. *The survival of Escherichia coli from freeze-thaw damage: permeability barrier damage and viability*. Can J Microbiol. **1975**; 21**:**1724-32.
23. Grant C, Deszcz D, Wei YC, Martinez-Torres RJ, Morris P, Folliard T, Sreenivasan R, Ward J, Dalby P, Woodley JM, Baganz F. *Identification and use of an alkane transporter plug-in for applications in biocatalysis and whole-cell biosensing of alkanes*. Sci Rep. **2014**; 4**:**5844.
24. Chen RR. *Permeability issues in whole-cell bioprocesses and cellular membrane engineering*. Appl Microbiol Biotechnol. **2007**; 74**:**730-8.
25. Lin BX, Zhang ZJ, Liu WF, Dong ZY, Tao Y. *Enhanced production of N-acetyl-D-neuraminic acid by multi-approach whole-cell biocatalyst*. Appl Microbiol Biotechnol. **2013**; 97**:**4775-84.

26. Mizutani M, Sato F. *Unusual P450 reactions in plant secondary metabolism*. Arch Biochem Biophys. **2011**; 507**:**194-203.
27. Greule A, Stok JE, De Voss JJ, Cryle MJ. *Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism*. Nat Prod Rep. **2018**; 35**:**757-91.
28. Biggs BW, Lim CG, Sagliani K, Shankar S, Stephanopoulos G, De Mey M, Ajikumar PK. *Overcoming heterologous protein interdependency to optimize P450-mediated taxol precursor synthesis in Escherichia coli*. Proc Natl Acad Sci USA. **2016**; 113**:**3209-14.
29. Ikezawa N, Tanaka M, Nagayoshi M, Shinkyo R, Sakaki T, Inouye K, Sato F. *Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured Coptis japonica cells*. J Biol Chem. **2003**; 278**:**38557-65.
30. Ortiz de Montellano PR. *Hydrocarbon hydroxylation by cytochrome P450 enzymes*. Chem Rev. **2010**; 110**:**932-48.
31. Paddon CJ, Keasling JD. *Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development*. Nat Rev Microbiol. **2014**; 12**:**355-67.
32. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, et al. *P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature.* Pharmacogenetics. **1996**; 6**:**1-42.
33. Hannemann F, Bichet A, Ewen KM, Bernhardt R. *Cytochrome P450 systems--biological variations of electron transport chains*. Biochim Biophys Acta. **2007**; 1770**:**330-44.
34. Omura T, Sato R. *The carbon monoxide-binding pigment of liver microsomes*. J Biol Chem. **1964**; 239**:**2369-78.
35. Bernhardt R, Urlacher VB. *Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations*. Appl Microbiol Biotechnol. **2014**; 98**:**6185-203.
36. Guengerich FP. *Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity*. Chem Res Toxicol. **2001**; 14**:**611–50.
37. Zhang X, Li S. *Expansion of chemical space for natural products by uncommon P450 reactions*. Nat Prod Rep. **2017**; 34**:**1061-89.
38. Li Z, Jiang Y, Guengerich FP, Ma L, Li S, Zhang W. *Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications*. J Biol Chem. **2020**; 295**:**833-49.
39. Karuzina II, Archakov AI. *The oxidative inactivation of cytochrome P450 in monooxygenase reactions*. Free Radic Biol Med. **1994**; 16**:**73-97.
40. Li S, Du L, Bernhardt R. *Redox partners: Function modulators of bacterial P450 enzymes*. Trends in Microbiology. **2020**; 28**:**445-54.
41. Riddick DS, Ding X, Wolf CR, Porter TD, Pandey AV, Zhang QY, Gu J, Finn RD, Ronseaux S, McLaughlin LA, et al. *NADPH-cytochrome P450 oxidoreductase: roles in physiology, pharmacology, and toxicology*. Drug Metab Dispos. **2013**; 41**:**12-23.
42. Warman AJ, Roitel O, Neeli R, Girvan HM, Seward HE, Murray SA, McLean KJ, Joyce MG, Toogood H, Holt RA, et al. *Flavocytochrome P450 BM3: an update on structure and mechanism of a biotechnologically important enzyme*. Biochem Soc Trans. **2005**; 33**:**747-53.
43. Kitazume T, Takaya N, Nakayama N, Shoun H. *Fusarium oxysporum fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of Bacillus megaterium cytochrome P450BM3*. J Biol Chem. **2000**; 275**:**39734-40.
44. Leonard E, Koffas MA. *Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by Escherichia coli*. Appl Environ Microbiol. **2007**; 73**:**7246-51.
45. Bakkes PJ, Biemann S, Bokel A, Eickholt M, Girhard M, Urlacher VB. *Design and improvement of artificial redox modules by molecular fusion of flavodoxin and flavodoxin reductase from Escherichia coli*. Sci Rep. **2015**; 5**:**12158.
46. Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. *Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli*. Science. **2010**; 330**:**70-4.

47. Chang MC, Eachus RA, Trieu W, Ro DK, Keasling JD. *Engineering Escherichia coli for production of functionalized terpenoids using plant P450s*. Nat Chem Biol. **2007**; 3**:**274-7.
48. Nebert DW, Wikvall K, Miller WL. *Human cytochromes P450 in health and disease*. Philos Trans R Soc Lond, B, Biol Sci. **2013**; 368**:**20120431.
49. Kelly SL, Kelly DE. *Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us?* Philos Trans R Soc Lond, B, Biol Sci. **2013**; 368**:**20120476.
50. Hwang KS, Kim UK, Charusanti P, Palsson BØ, Lee SY. *Systems biology and biotechnology of Streptomyces species for the production of secondary metabolites*. Biotechnol Adv. **2014**; 32**:**255-68.
51. Hilberath T, Windeln LM, Decembrino D, Le-Huu P, Bilsing FL, Urlacher VB. *Two-step screening for identification of drug-metabolizing bacterial cytochromes P450 with diversified selectivity*. ChemCatChem. **2020**; 12**:**1710-9.
52. Nelson DR. *Cytochrome P450 diversity in the tree of life*. Biochim Biophys Acta, Proteins Proteom. **2018**; 1866**:**141-54.
53. Hamberger B, Bak S. *Plant P450s as versatile drivers for evolution of species-specific chemical diversity*. Philos Trans R Soc Lond, B, Biol Sci. **2013**; 368**:**20120426.
54. Liu X, Zhu X, Wang H, Liu T, Cheng J, Jiang H. *Discovery and modification of cytochrome P450 for plant natural products biosynthesis*. Synth Syst Biotechnol. **2020**; 5**:**187-99.
55. Anarat-Cappillino G, Sattely ES. *The chemical logic of plant natural product biosynthesis*. Curr Opin Plant Biol. **2014**; 19**:**51-8.
56. Pateraki I, Heskes AM, Hamberger B. *Cytochromes P450 for terpene functionalisation and metabolic engineering*. Adv Biochem Eng Biotechnol. **2015**; 148**:**107-39.
57. Hausjell J, Halbwirth H, Spadiut O. *Recombinant production of eukaryotic cytochrome P450s in microbial cell factories*. Biosci Rep. **2018**; 38**:**1-13.
58. Saito Y, Kitagawa W, Kumagai T, Tajima N, Nishimiya Y, Tamano K, Yasutake Y, Tamura T, Kameda T. *Developing a codon optimization method for improved expression of recombinant proteins in actinobacteria*. Sci Rep. **2019**; 9**:**8338.
59. Punde N, Kooken J, Leary D, Legler PM, Angov E. *Codon harmonization reduces amino acid misincorporation in bacterially expressed P. falciparum proteins and improves their immunogenicity*. AMB Express. **2019**; 9**:**167.
60. Zelasko S, Palaria A, Das A. *Optimizations to achieve high-level expression of cytochrome P450 proteins using Escherichia coli expression systems*. Protein Expr Purif. **2013**; 92**:**77-87.
61. Urlacher VB, Girhard M. *Cytochrome P450 Monooxygenases in Biotechnology and Synthetic Biology*. Trends Biotechnol. **2019**; 37**:**882-97.
62. Renault H, Bassard JE, Hamberger B, Werck-Reichhart D. *Cytochrome P450-mediated metabolic engineering: current progress and future challenges*. Curr Opin Plant Biol. **2014**; 19**:**27-34.
63. Markulin L, Corbin C, Renouard S, Drouet S, Gutierrez L, Mateljak I, Auguin D, Hano C, Fuss E, Laine E. *Pinoresinol-lariciresinol reductases, key to the lignan synthesis in plants*. Planta. **2019**; 249**:**1695-714.
64. Teponno RB, Kusari S, Spiteller M. *Recent advances in research on lignans and neolignans*. Nat Prod Rep. **2016**; 33**:**1044-92.
65. Satake H, Koyama T, Bahabadi SE, Matsumoto E, Ono E, Murata J. *Essences in metabolic engineering of lignan biosynthesis*. Metabolites. **2015**; 5**:**270-90.
66. Sok DE, Cui HS, Kim MR. *Isolation and bioactivities of furfuran type lignan compounds from edible plants*. Recent Pat Food, Nutr Agric. **2009**; 1**:**87-95.
67. Clavel T, Henderson G, Alpert CA, Philippe C, Rigottier-Gois L, Dore J, Blaut M. *Intestinal bacterial communities that produce active estrogen-like compounds enterodiol and enterolactone in humans*. Appl Environ Microbiol. **2005**; 71**:**6077-85.

68. Milder IE, Arts IC, van de Putte B, Venema DP, Hollman PC. *Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol*. Br J Nutr. **2005**; 93**:**393-402.
69. Ardalani H, Avan A, Ghayour-Mobarhan M. *Podophyllotoxin: a novel potential natural anticancer agent*. Avicenna J Phytomed. **2017**; 7**:**285-94.
70. Lau W, Sattely ES. *Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone*. Science. **2015**; 349**:**1224-8.
71. World Health Organization. *WHO Model List of Essential Medicines, 21th list*. **2019**.
72. Satake H, Shiraishi A, Koyama T, Matsumoto E, Morimoto K, Bahabadi SE, Ono E, Murata J: *Lignan Biosynthesis for Food Bioengineering***.** In *Food Biosynthesis. Volume* 1. Edited by Grumezescu AM, Holban AM: Elsevier; **2017**: 351-79.
73. Kumari A, Singh D, Kumar S. *Biotechnological interventions for harnessing podophyllotoxin from plant and fungal species: current status, challenges, and opportunities for its commercialization*. Crit Rev Biotechnol. **2017**; 37**:**739-53.
74. Rudroff F, Mihovilovic MD, Gröger H, Snajdrova R, Iding H, Bornscheuer UT. *Opportunities and challenges for combining chemo- and biocatalysis*. Nat Catal. **2018**; 1**:**12-22.
75. Lazzarotto M, Hammerer L, Hetmann M, Borg A, Schmermund L, Steiner L, Hartmann P, Belaj F, Kroutil W, Gruber K, Fuchs M. *Chemoenzymatic total synthesis of deoxy-, epi-, and podophyllotoxin and a biocatalytic kinetic resolution of dibenzylbutyrolactones*. Angew Chem Int Ed. **2019**; 58**:**8226-30.
76. Sperl JM, Sieber V. *Multienzyme cascade reactions—Status and recent advances*. ACS Catal. **2018**; 8**:**2385-96.
77. Song JW, Seo JH, Oh DK, Bornscheuer UT, Park JB. *Design and engineering of whole-cell biocatalytic cascades for the valorization of fatty acids*. Catal Sci Technol. **2020**; 10**:**46-64.
78. Ford JD, Huang KS, Wang HB, Davin LB, Lewis NG. *Biosynthetic pathway to the cancer chemopreventive secoisolariciresinol diglucoside-hydroxymethyl glutaryl ester-linked lignan oligomers in flax (Linum usitatissimum) seed*. J Nat Prod. **2001**; 64**:**1388-97.
79. Xia ZQ, Costa MA, Pelissier HC, Davin LB, Lewis NG. *Secoisolariciresinol dehydrogenase purification, cloning, and functional expression. Implications for human health protection*. J Biol Chem. **2001**; 276**:**12614-23.
80. Marques JV, Kim KW, Lee C, Costa MA, May GD, Crow JA, Davin LB, Lewis NG. *Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis*. J Biol Chem. **2013**; 288**:**466-79.
81. Pickel B, Schaller A. *Dirigent proteins: molecular characteristics and potential biotechnological applications*. Appl Microbiol Biotechnol. **2013**; 97**:**8427-38.
82. Dalisay DS, Kim KW, Lee C, Yang H, Rubel O, Bowen BP, Davin LB, Lewis NG. *Dirigent Protein-mediated lignan and cyanogenic glucoside formation in flax seed: integrated omics and MALDI mass spectrometry imaging*. J Nat Prod. **2015**; 78**:**1231-42.
83. Davin LB, Wang H, Crowell AL, Bedgar DL, Martin DM, Sarkanen S, Lewis NG. *Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center*. Science. **1997**; 275**:**362-7.
84. Halls SC, Davin LB, Kramer DM, Lewis NG. *Kinetic study of coniferyl alcohol radical binding to the (+)-pinoresinol forming dirigent protein*. Biochemistry. **2004**; 43**:**2587-95.
85. Lv Y, Cheng X, Du G, Zhou J, Chen J. *Engineering of an H_2O_2 auto-scavenging in vivo cascade for pinoresinol production*. Biotechnol Bioeng. **2017**; 114**:**2066-74.
86. Ricklefs E, Girhard M, Koschorreck K, Smit MS, Urlacher VB. *Two-Step One-Pot synthesis of pinoresinol from eugenol in an enzymatic cascade*. ChemCatChem. **2015**; 7**:**1857-64.
87. Dinkova-Kostova AT, Gang DR, Davin LB, Bedgar DL, Chu A, Lewis NG. *(+)-Pinoresinol/(+)-lariciresinol reductase from Forsythia intermedia*. J Biol Chem. **1996**; 46**:**29473–82.

88. Fujita M, Gang DR, Davin LB, Lewis NG. *Recombinant pinoresinol-lariciresinol reductases from western red cedar (Thuja plicata) catalyze opposite enantiospecific conversions*. J Biol Chem. **1999**; 274**:**618–27.
89. Corbin C, Drouet S, Mateljak I, Markulin L, Decourtil C, Renouard S, Lopez T, Doussot J, Lamblin F, Auguin D, et al. *Functional characterization of the pinoresinol-lariciresinol reductase-2 gene reveals its roles in yatein biosynthesis and flax defense response*. Planta. **2017**; 246**:**405-20.
90. von Heimendahl CB, Schafer KM, Eklund P, Sjoholm R, Schmidt TJ, Fuss E. *Pinoresinol-lariciresinol reductases with different stereospecificity from Linum album and Linum usitatissimum*. Phytochemistry. **2005**; 66**:**1254-63.
91. Min T, Kasahara H, Bedgar DL, Youn B, Lawrence PK, Gang DR, Halls SC, Park H, Hilsenbeck JL, Davin LB, et al. *Crystal structures of pinoresinol-lariciresinol and phenylcoumaran benzylic ether reductases and their relationship to isoflavone reductases*. J Biol Chem. **2003**; 278**:**50714-23.
92. Davin LB, Jourdes M, Patten AM, Kim KW, Vassao DG, Lewis NG. *Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis*. Nat Prod Rep. **2008**; 25**:**1015-90.
93. Moinuddin SG, Youn B, Bedgar DL, Costa MA, Helms GL, Kang C, Davin LB, Lewis NG. *Secoisolariciresinol dehydrogenase: mode of catalysis and stereospecificity of hydride transfer in Podophyllum peltatum*. Org Biomol Chem. **2006**; 4**:**808-16.
94. Noguchi A, Horikawa M, Murata J, Tera M, Kawai Y, Ishiguro M, Umezawa T, Mizutani M, Ono E. *Mode-of-action and evolution of methylenedioxy bridge forming P450s in plant specialized metabolism*. Plant Biotechnol. **2014**; 31**:**493-503.
95. Ono E, Nakai M, Fukui Y, Tomimori N, Fukuchi-Mizutani M, Saito M, Satake H, Tanaka T, Katsuta M, Umezawa T, Tanaka Y. *Formation of two methylenedioxy bridges by a Sesamum CYP81Q protein yielding a furofuran lignan, (+)-sesamin*. Proc Natl Acad Sci USA. **2006**; 103**:**10116–21.
96. Yoshigae Y, Kent UM, Hollenberg PF. *Role of the highly conserved threonine in cytochrome P450 2E1: prevention of H_2O_2-induced inactivation during electron transfer*. Biochemistry. **2013**; 52**:**4636-47.
97. Kamil WM, Dewick PM. *Biosynthetic relationship of aryltetralin lactone lignans to dibenzylbutyrolactone lignans*. Phytochemistry. **1986**; 25**:**2093-102.
98. Farrow SC, Facchini PJ. *Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism*. Front Plant Sci. **2014**; 5**:**524.
99. Lam KC, Ibrahim RK, Behdad B, Dayanandan S. *Structure, function, and evolution of plant O-methyltransferases*. Genome. **2007**; 50**:**1001-13.
100. Jung ST, Lauchli R, Arnold FH. *Cytochrome P450: taming a wild type enzyme*. Curr Opin Biotechnol. **2011**; 22**:**809-17.
101. Bokel A, Rühlmann A, Hutter MC, Urlacher VB. *Enzyme-mediated two-step regio- and stereoselective synthesis of potential rapid-acting antidepressant (2S,6S)-hydroxynorketamine*. ACS Catal. **2020**; 10**:**4151-9.
102. Ricklefs E, Girhard M, Urlacher VB. *Three-steps in one-pot: whole-cell biocatalytic synthesis of enantiopure (+)- and (-)-pinoresinol via kinetic resolution*. Microb Cell Fact. **2016**; 15**:**78.
103. Kranz-Finger S, Mahmoud O, Ricklefs E, Ditz N, Bakkes PJ, Urlacher VB. *Insights into the functional properties of the marneral oxidase CYP71A16 from Arabidopsis thaliana*. Biochim Biophys Acta. **2018**; 1866**:**2-10.
104. Worsch A, Eggimann FK, Girhard M, von Buhler CJ, Tieves F, Czaja R, Vogel A, Grumaz C, Sohn K, Lutz S, et al. *A novel cytochrome P450 mono-oxygenase from Streptomyces platensis resembles activities of human drug metabolizing P450s*. Biotechnol Bioeng. **2018**; 115**:**2156-66.

105. Girhard M, Klaus T, Khatri Y, Bernhardt R, Urlacher VB. *Characterization of the versatile monooxygenase CYP109B1 from Bacillus subtilis*. Appl Microbiol Biotechnol. **2010**; 87**:**595-607.
106. Khatri Y, Hannemann F, Girhard M, Kappl R, Hutter M, Urlacher VB, Bernhardt R. *A natural heme-signature variant of CYP267A1 from Sorangium cellulosum So ce56 executes diverse omega-hydroxylation*. FEBS Journal. **2015**; 282**:**74-88.
107. Omura TaSR. *The carbon monoxide-binding pigment of liver microsomes. Evidence for its hemoprotein nature.* J Biol Chem. **1964**; 239**:**2370–8.
108. Kuo HJ, Wei ZY, Lu PC, Huang PL, Lee KT. *Bioconversion of pinoresinol into matairesinol by use of recombinant Escherichia coli*. Applied and Environmental Microbiology. **2014**; 80**:**2687–92.
109. Magadula JJ. *Phytochemistry and pharmacology of the genus Macaranga: a review*. J Med Plant Res. **2014**; 8**:**489-503.
110. Okunishi T, Umezawa T, Shimada M. *Enantiomeric compositions and biosynthesis of Wikstroemia sikokiana lignans*. J Wood Sci. **2000**; 46**:**234-42.
111. Zalesak F, Bon DJD, Pospisil J. *Lignans and neolignans: Plant secondary metabolites as a reservoir of biologically active substances*. Pharmacol Res. **2019**; 146**:**104284.
112. Petrovska BB. *Historical review of medicinal plants' usage*. Pharmacogn Rev. **2012**; 6**:**1-5.
113. Desbène S, Giorgi-Renault S. *Drugs that inhibit tubulin polymerization: the particular case of podophyllotoxin and analogues*. Current Medicinal Chemistry - Anti-Cancer Agents **2002**; 2**:**71-90.
114. Chaurasia OP, Ballabh B, Tayade A, Kumar R, Phani Kumar G, Sing SB. *Podophyllum L.: An endergered and anticancerous medicinal plant – An overview*. Indian J Tradit Knowl. **2012**; 11**:**234-41.
115. Nadeem M, Palni LMS, Purohit AN, Pandey H, Nandi SK. *Propagation and conservation of Podophyllum hexandrum Royle animportant medicinal herb*. Biol Conserv. **2000**; 92**:**121-9.
116. Li J, Zhang X, Renata H. *Asymmetric chemoenzymatic synthesis of (-)-podophyllotoxin and related aryltetralin lignans*. Angew Chem Int Ed. **2019**; 58**:**11657-60.
117. Cravens A, Payne J, Smolke CD. *Synthetic biology strategies for microbial biosynthesis of plant natural products*. Nat Commun. **2019**; 10**:**2142.
118. Birchfield AS, McIntosh CA. *Metabolic engineering and synthetic biology of plant natural products – A minireview*. Curr Plant Biol. **2020**; 24**:**100163.
119. Chemler JA, Koffas MA. *Metabolic engineering for plant natural product biosynthesis in microbes*. Curr Opin Biotechnol. **2008**; 19**:**597-605.
120. Decembrino D, Ricklefs E, Wohlgemuth S, Girhard M, Schullehner K, Jach G, Urlacher VB. *Assembly of plant enzymes in E. coli for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide*. ACS Synth Biol. **2020**; 9**:**3091-103.
121. Decembrino D, Girhard M, Urlacher VB. *Use of Copper as a Trigger for the in Vivo Activity of E. coli Laccase CueO: A Simple Tool for Biosynthetic Purposes*. ChemBioChem. **2021**; 22**:**1470-9.
122. Li M, Lv M, Yang D, Wei J, Xing H, Paré PW. *Temperature-regulated anatomical and gene-expression changes in Sinopodophyllum hexandrum seedlings*. Ind Crop Prod. **2020**; 152**:**112479.
123. Schuler MA, Werck-Reichhart D. *Functional genomics of P450s*. Annu Rev Plant Biol. **2003**; 54**:**629-67.
124. Brown S, Clastre M, Courdavault V, O'Connor SE. *De novo production of the plant-derived alkaloid strictosidine in yeast*. Proc Natl Acad Sci USA. **2015**; 112**:**3205-10.
125. Yamazaki S, Sato K, Suhara K, Sakaguchi M, Mihara K, Omura T. *Importance of the proline-rich region following signal-anchor sequence in the formation of correct conformation of microsomal cytochrome P450s*. J Biochem. **1993**; 114**:**652–7.
126. Looman AC, Bodlaender J, Comstock LJ, Eaton D, Jhurani P, de Boer HA, van Knippenberg PH. *Influence of the codon following the AUG initiation codon on the expression of a modified lacZ gene in Escherichia coli*. EMBO J. **1987**; 6**:**2489-92.

127. Jensen K, Moller BL. *Plant NADPH-cytochrome P450 oxidoreductases*. Phytochemistry. **2010**; 71**:**132-41.
128. Bassard JE, Moller BL, Laursen T. *Assembly of dynamic P450-mediated metabolons - Order versus Chaos*. Curr Mol Biol Rep. **2017**; 3**:**37-51.
129. Kumari A, Singh HR, Jha A, Swarnkar MK, Shankar R, Kumar S. *Transcriptome sequencing of rhizome tissue of Sinopodophyllum hexandrum at two temperatures*. BMC Genom. **2014**; 15**:**1-17.
130. Li M, Sun P, Kang T, Xing H, Yang D, Zhang J, Paré PW. *Mapping podophyllotoxin biosynthesis and growth-related transcripts with high elevation in Sinopodophyllum hexandrum*. Ind Crop Prod. **2018**; 124**:**510-8.
131. Sawayama AM, Chen MM, Kulanthaivel P, Kuo MS, Hemmerle H, Arnold FH. *A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds*. Chemistry. **2009**; 15**:**11723-9.
132. Li A, Acevedo-Rocha CG, D'Amore L, Chen J, Peng Y, Garcia-Borras M, Gao C, Zhu J, Rickerby H, Osuna S, et al. *Regio- and stereoselective steroid hydroxylation at C7 by cytochrome P450 monooxygenase mutants*. Angew Chem Int Ed. **2020**; 59**:**12499-505.
133. Kranz-Finger S, Mahmoud O, Ricklefs E, Ditz N, Bakkes PJ, Urlacher VB. *Insights into the functional properties of the marneral oxidase CYP71A16 from Arabidopsis thaliana*. Biochim Biophys Acta, Proteins Proteom. **2018**; 1866**:**2-10.
134. Hull AK, Celenza JL. *Bacterial expression and purification of the Arabidopsis NADPH-cytochrome P450 reductase ATR2*. Protein Expr Purif. **2000**; 18**:**310-5.
135. Omura T, R. S. *The carbon monoxide-binding pigment of liver microsomes I. Evidence for its hemoprotein nature*. J Biol Chem. **1964**; 239**:**2370–8.
136. Omura T, Sato R. *The carbon monoxide-binding pigment of liver microsomes I. Evidence for its hemoprotein nature*. J Biol Chem. **1964**; 239**:**2370–8.
137. Kudanga T, Nemadziva B, Le Roes-Hill M. *Laccase catalysis for the synthesis of bioactive compounds*. Appl Microbiol Biotechnol. **2016**; 101**:**13-33.
138. Adelakun OE, Kudanga T, Green IR, le Roes-Hill M, Burton SG. *Enzymatic modification of 2,6-dimethoxyphenol for the synthesis of dimers with high antioxidant capacity*. Process Biochem. **2012**; 47**:**1926-32.
139. Adelakun OE, Kudanga T, Parker A, Green IR, le Roes-Hill M, Burton SG. *Laccase-catalyzed dimerization of ferulic acid amplifies antioxidant activity*. J Mol Catal B: Enzym. **2012**; 74**:**29-35.
140. Lee W, Ko KR, Kim HK, Lim S, Kim S. *Dehydrodiconiferyl alcohol promotes BMP-2-induced osteoblastogenesis through its agonistic effects on estrogen receptor*. Biochem Biophys Res Commun. **2018**; 495**:**2242-8.
141. Lee J, Choi J, Lee W, Ko K, Kim S. *Dehydrodiconiferyl alcohol (DHCA) modulates the differentiation of Th17 and Th1 cells and suppresses experimental autoimmune encephalomyelitis*. Mol Immunol. **2015**; 68**:**434-44.
142. Jeon JR, Baldrian P, Murugesan K, Chang YS. *Laccase-catalysed oxidations of naturally occurring phenols: from in vivo biosynthetic pathways to green synthetic applications*. Microb Biotechnol. **2012**; 5**:**318-32.
143. Ehlting J, Hamberger B, Million-Rousseau R, Werck-Reichhart D. *Cytochromes P450 in phenolic metabolism*. Phytochem Rev. **2006**; 5**:**239-70.
144. Szczebara FM, Chandelier C, Villeret C, Masurel A, Bourot S, Duport C, Blanchard S, Groisillier A, Testet E, Costaglioli P, et al. *Total biosynthesis of hydrocortisone from a simple carbon source in yeast*. Nat Biotechnol. **2003**; 21**:**143-9.
145. Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, et al. *High-level semi-synthetic production of the potent antimalarial artemisinin*. Nature. **2013**; 496**:**528-32.

146. Quin MB, Wallin KK, Zhang G, Schmidt-Dannert C. *Spatial organization of multi-enzyme biocatalytic cascades*. Org Biomol Chem. **2017**; 15**:**4260-71.
147. Arkin A. *Setting the standard in synthetic biology*. Nat Biotechnol. **2008**; 26.
148. Canton B, Labno A, Endy D. *Refinement and standardization of synthetic biological parts and devices*. Nat Biotechnol. **2008**; 26**:**787-93.
149. Wu S, Zhou Y, Wang T, Too HP, Wang DI, Li Z. *Highly regio- and enantioselective multiple oxy- and amino-functionalizations of alkenes by modular cascade biocatalysis*. Nat Commun. **2016**; 7**:**11917.
150. Polka JK, Hays SG, Silver PA. *Building spatial synthetic biology with compartments, scaffolds, and communities*. Cold Spring Harb Perspect Biol. **2016**; 8**:**1-16.
151. Singleton C, Howard TP, Smirnoff N. *Synthetic metabolons for metabolic engineering*. J Exp Bot. **2014**; 65**:**1947-54.
152. Jorgensen K, Rasmussen AV, Morant M, Nielsen AH, Bjarnholt N, Zagrobelny M, Bak S, Moller BL. *Metabolon formation and metabolic channeling in the biosynthesis of plant natural products*. Curr Opin Plant Biol. **2005**; 8**:**280-91.
153. Ahmadi MK, Fang L, Moscatello N, Pfeifer BA. *E. coli metabolic engineering for gram scale production of a plant-based anti-inflammatory agent*. Metab Eng. **2016**; 38**:**382-8.
154. Zhang H, Pereira B, Li Z, Stephanopoulos G. *Engineering Escherichia coli coculture systems for the production of biochemical products*. Proc Natl Acad Sci USA. **2015**; 112**:**8266-71.
155. Elleuche S, Antranikian G. *Bacterial group III alcohol dehydrogenases–function, evolution and biotechnological applications*. OA Alcohol. **2013**; 1.
156. Kuchler A, Yoshimoto M, Luginbuhl S, Mavelli F, Walde P. *Enzymatic reactions in confined environments*. Nat Nanotechnol. **2016**; 11**:**409-20.
157. Walton AZ, Stewart JD. *Understanding and improving NADPH-dependent reactions by nongrowing Escherichia coli cells*. Biotechnol Prog. **2004**; 20**:**403−11.
158. Hwang JK, Moinuddin SGA, Davin LB, Lewis NG. *Pinoresinol-lariciresinol reductase: Substrate versatility, enantiospecificity, and kinetic properties*. Chirality. **2020**; 32**:**770-89.
159. Kang SY, Choi O, Lee JK, Hwang BY, Uhm TB, Hong YS. *Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing Escherichia coli strain*. Microb Cell Fact. **2012**; 11**:**1-9.
160. Marienhagen J, Bott M. *Metabolic engineering of microorganisms for the synthesis of plant natural products*. J Biotechnol. **2013**; 163**:**166-78.
161. Munoz AJ, Hernandez-Chavez G, de Anda R, Martinez A, Bolivar F, Gosset G. *Metabolic engineering of Escherichia coli for improving L-3,4-dihydroxyphenylalanine (L-DOPA) synthesis from glucose*. J Ind Microbiol Biotechnol. **2011**; 38**:**1845-52.
162. Santos-Sánchez NF, Salas-Coronado R, Hernández-Carlos B, Villanueva-Cañongo C: *Shikimic acid pathway in biosynthesis of phenolic compounds***.** In *Plant Physiological Aspects of Phenolic Compounds.* Edited by Soto-Hernández M: IntechOpen; **2019**: 1-15.
163. Jansen F, Gillessen B, Mueller F, Commandeur U, Fischer R, Kreuzaler F. *Metabolic engineering for p-coumaryl alcohol production in Escherichia coli by introducing an artificial phenylpropanoid pathway*. Biotechnol Appl Biochem. **2014**; 61**:**646-54.
164. Pitera DJ, Paddon CJ, Newman JD, Keasling JD. *Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli*. Metab Eng. **2007**; 9**:**193-207.

5 Acknowledgments

Without the people that will be mentioned in this section, simply, I would have not been writing this dissertation, neither I would have committed to start a doctorate. Brutal truth is: I would have achieved very few things in my life without some of them.

This assay does not only represent a terrific academic achievement, but also a big step in personal consciousness and self-acceptance. Instead of fighting against windmills, like a very-less-skilled Don Quixote- it has been about starting to learn how to place the sails downwind.

When I first moved to Düsseldorf as an Erasmus student, I could have never imagined that it would have become "*Home*" for the following four years. This thesis represents thus the closure of a chapter of my life which has been of inestimable value, with its extremely low rock bottoms and incredibly rewarding climaxes. I started this PhD because it felt good, because intuition told me so. I followed blindly, even though terrified whether I could have been able to face such challenge. It was worthy to try.

The very first "*Thank you*" goes to my supervisor and first corrector **Prof. Vlada B. Urlacher**. Dear Vlada, I distinctly remember the first email exchange we had back in 2016. Addressing you in first name terms without even knowing you, and remembering Sheila remarking that was inappropriate to address a new Professor like that. I do remember how awkward it felt but, thinking backwards, I want to believe that was somehow a premise of how everything would have run smoothly. I would like to thank you for all the scientific advises and comments. Second and more important, I am extremely grateful for the trust you ensured me over the years: as an Erasmus student in the beginning, and later granting the freedom to make mistakes out of my own stubbornness and learning from them as a PhD. Even more, thank you for the open communication in every moment or circumstance of my time at the Institute. Either discussing on scientific or personal matters, especially when disagreeing, it has always been a great confrontation, growth space, and learning occasion. Herzlichen Dank!

Sincere thanks to my second supervisor **Prof. Ilka Axmann**, for the kind availability and support in both the first and the last steps of this road. Dankeschön!

Also, I really would like to express my gratitude to **Dr. Esther Ricklefs** and **Dr. Marco Girhard**. Dear Esther, I don't know how many times I've repeated how lucky it felt to receive this project from you. Your precious advises in the very beginning and your (even more) precious work have been incredibly important to let me grow confidence in what I had to deal with. Dear Marco, over the last four years your feedbacks and help have been fundamental in many ways. Besides practical and technical hints, I've really appreciated your way of pushing me towards ideas with a pragmatic attitude that was easy to start feeling mine. Vielen Dank euch beiden!

A special thought has to go to **Prof. Sheila Sadeghi**. Dear Sheila, it is very simple here. I would not have even started getting down this road without you addressing me to Düsseldorf in the very beginning. I am deeply grateful for this. Thank you!

Over the last four years I observed people getting in and out of the Institute of Biochemistry II as much as from the apartment in Bruchstraße 33. Among them, there's people I am particularly grateful to. A very special thanks goes to my very good friend **Dr. Thomas Hilberath**. Dear Thomas, we talked about this so many times that we have lost count. Somehow you're always going to be my supervisor and mentor. When we first met you started transmitting daily your passion for this world. It made me willing to learn more and more. So if I'm here writing this, having you as a supervisor played a pretty big role. Through your advises and (massive) patience things that first felt like *nuts* became ordinary. You still do that, and I extremely appreciate your point of view whichever the topic, the place or the situation. Smelling a freshly open package of coffee, having infinite Birrette/Bierchen at *Rock Am Ring*, *Schaukelstühlchen*, in front of a *Schalander's Bratkartoffeln Pfanne,* or on Zoom! Never change a running system and keep it simple, huh? Grazie mille caro, di cuore!

I want to express my heartfelt gratitude to the most complementary partner I could have ever imagined working with, **Stefan Wohlgemuth**. Dear Stefan, over the last three years and a half we've been through LC/MS issues, students, meetings, and conferences which brought us as many satisfactions as crippling doubts eventually building up a very good friendship over completely different personalities. I've immensely appreciated your help and support especially in moments when nothing was working out, and I've learned a lot from you. Our adventures at BioCat and the CeBiTec Symposium have been definitely the cherry on top of the cake! Herzlichen Dank!

Another few words go to **Alessandra Raffaele**. Ciao cara! In qualche modo il tuo arrivo a Düsseldorf ha scandito le fasi del mio dottorato e se penso a quando sono venuto a recuperarti in aeroporto ancora rido. La verità è che sei stata un'aggiunta importante in questo percorso, per le domande che mi hai portato a farmi, per i limiti che si sono evidenziati e che mi hai aiutato a smussare professionalmente e non. Non soltanto durante la supervisione della tua tesi, ma soprattutto come amica, poi. Come se non bastasse, la tua compagnia e giovialità, portando un po' di Italia tra le mura del lab, sono state di grande aiuto in momenti in cui la bussola puntava un po' dove le pareva. È proprio vero che è molto più facile accorgersi dei pregi altrui piuttosto che dei propri. Grazie di tutto, di cuore!

Thank you to all the present and former group members, PhDs, technicians and students, who shared with me the IBCII walls. I'm deeply grateful for the scientific feedbacks and the daily flair which always felt a bit more than being "just" colleagues.

Among them I would like to spend few words thanking **Anna Olbrich**, **Ansgar Bokel**, **Arsenij Kokorin**, **Fabian Schmitz**, **Leonie Windeln**, **Nina Jankowski**, **Nikolas Ditz**, **and Sebastian Hölzel**. How to sum it up? It has been a blast! Plus, it's pretty hard to separate you guys, each memory is bonded to all of you together!

Thank you, Anna, for the everyday lab-life from the very beginning of this adventure, thank you for the laughter, the talks, the shared attendings to Death Metal gigs and for making me

feel less alone as a sport junky. Thank you, Leonie, for the spirit you brought in when you joined the group, but even more for the friend you turned out to be. Ansgar, Arsenij, Fabian, Nina, Niko and Sebastian thank you guys for the insane amount of laughter, for the shootings with Nerf guns, for non-stopping me from ordering Irish Bombs for everybody and the following hangovering around Lithuania, for finding Arsenij's bike thrown into a pond, for breaking the handbrake, for squeezing in the car on the way to Rock Am Ring, for two Rock Am Ring (!), for Vikings' Chess, exploding kittens, for laminating paper sheets the wrong way. Big time people.

Thanks to all of you guys for making the effort of speaking/translating in English in any situation because of a language-lazy Italian who never felt a foreigner. Vielen, vielen, vielen Dank Leute. Es war MEGA-GEIL!

On the other side of the coin, to me, Düsseldorf is Bruchstraße 33, Flingern Nord. There's no other address. It all started with the original Flingern Family: **Alexa**, **Gaia**, **Jens** and **Kevin**. It was the perfect storm, and the time spent living together hanging around with **Ilaria, Mistico,** and **Vasilis** has been one the best times of my life (and counting)!

I also would like to thank my current roommates **Anastasia**, **Karen**, and **Dennis**. Over the last year and a half, I often found myself down in a ditch, focusing on what was going wrong instead of joying of those serenity bites that popped out from time to time. Often those bites came from you guys. Being out of that ditch now makes me realize how important you all have been in this process. For the same reason I would like to thank **Silvia M.** and **Riccardo**, as in sought after some balance you helped me more than I'd be ever able to express. Love you guys, tons.

Last but not least, there's a bunch of people which have always been there. My parents **Nora** and **Paolo** in the first place. Non ho la minima idea di come ci si senta a lasciare andare il proprio figlio alla scoperta del mondo. Credo ci vogliano tanto coraggio e tanta forza di volontà. Ciò che so è che senza di voi, senza ciò che mi avete trasmesso non sarei né fisicamente qui,

né in procinto di ultimare questa tesi. Mi avete sempre spronato ad esplorare, insegnato a vedere più in là del mio naso, lasciandomi lo spazio di sbagliare ed imparare. Mi avete insegnato a dare il meglio di me, sempre. Dirvi soltanto “Grazie” è terribilmente riduttivo.

A very last thank you goes to life-lasting friendships. **Jagielka**, **Seemo**, **Lenz**, **Stefanino, Hermete, la Doc** and **Zini**. Con voi ho alcuni dei ricordi più intensi della mia vita. Questo percorso non fa eccezione, anzi più la distanza fisica aumentava più la vostra importanza cresceva. Lo stesso vale per **Silvia C**. e **Scili**. Quando una volta dissi *“Si farà i biologi, e forse in fondo va davvero bene così”* in qualche modo l’intuito sapeva già del percorso che avremmo condiviso insieme.

They say that “Fortune favours the brave” but I disagree. Fortune happens to those which are exposing themselves to both fortune and misfortune. I may take credit for the guts of trying things out but I got tremendously lucky finding such amazing people in my life. That is what matters the most, that’s what pushed me through.

It has always been a struggle to describe people with words. Those who know me best are aware of my relationship with words and their weight. You all have been more important than I could ever be able to express. You all have been there in so many different ways, all equally important, even in moments when it was hard to recognize myself. My deepest and heartfelt gratitude, to all of you, beautiful people. You deserve a huge credit in this: *“There’s more to the picture, than meets the eye, hey hey, my my”.*

“For Those About To Rock, We Salute You!”

AC/DC