Identification and characterization of novel cytochromes P450 from *actinomycetes*

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Thomas Hilberath

aus Kottenborn

Düsseldorf, August 2020

Aus dem Institut für Biochemie, Lehrstuhl II
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referentin: Prof. Dr. Vlada B. Urlacher
Korreferentin: Prof. Dr. Martina Pohl
Tag der mündlichen Prüfung: 27. Oktober 2020

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Diese Dissertation wurde weder in dieser noch in ähnlicher Form bei einer anderen Fakultät eingereicht. Ich habe bisher keinen erfolglosen oder erfolgreichen Promotionsversuch unternommen.

Düsseldorf, _____ _____

(Thomas Hilberath)

„Viel geschrieben ist hier nieder,

die Studiererei, es spiegelt wieder.

Den kleinen Vers zu Ehren ich widme dir,

in deinem Stile, am Anfang dieses Werkes hier.“

In dedication to my grandfather Klaus Wissen

Table of contents

Publications and conference contributions

Publications resulting from this work:

Hilberath T, Windeln LM, Decembrino D, Le-Huu P, Bilsing FL, Urlacher VB, **2020**, Two-step screening for identification of drug-metabolizing bacterial cytochromes P450 with diversified selectivity, *ChemCatChem*, **12**(6), 1710-1719.

Grumaz C, Vainshtein Y, Kirstahler P, Luetz S, Kittelmann M, Schroer K, Eggimann F, Czaja R, Vogel A, **Hilberath T**, Worsch A, Girhard M, Urlacher VB, Sandberg M, Sohn K, **2017**, Draft genome sequences of three *Actinobacteria* strains presenting new candidate organisms with high potentials for specific P450 cytochromes, *Genome Announcements* **5**(28):e00532-17.

Publications not related to this work:

Dreßen A, **Hilberath T**, Mackfeld U, Billmeier A, Rudat J, Pohl M, **2017**, Phenylalanine ammonia lyase from *Arabidopsis thaliana* (*At*PAL2): A potent MIO-enzyme for the synthesis of non-canonical aromatic alpha-amino acids: Part I: Comparative characterization to the enzymes from *Petroselinum crispum* (*Pc*PAL1) and *Rhodosporidium toruloides* (*Rt*PAL), *Journal of biotechnology* **258**: 148-157.

Dreßen A, **Hilberath T**, Mackfeld U, Rudat J, Pohl M, **2017**, Phenylalanine ammonia lyase from *Arabidopsis thaliana* (*At*PAL2): A potent MIO-enzyme for the synthesis of non-canonical aromatic alpha-amino acids: Part II: Application in different reactor concepts for the production of (*S*)-2-chloro-phenylalanine, *Journal of biotechnology* **258**: 158-166.

Conference contributions:

Hilberath T, Decembrino D, Le-Huu P, Worsch A, Girhard M, Urlacher VB, Efficient oxidation of chemically diverse scaffolds by promiscuous cytochrome P450s from *actinomycetes*; *20th International Conference on Cytochrome P450 (20th ICCP450)*, **2017**, Düsseldorf, Germany. (Poster)

Hilberath T, Windeln LM, Decembrino D, Le-Huu P, Bilsing FL, Urlacher VB, Exploring novel promiscuous cytochrome P450 monooxygenases for syntheses of diverse drug metabolites; *6th Novel Enzymes Conference*, **2018**, Darmstadt, Germany. (Short talk and poster)

Hilberath T, Windeln LM, Decembrino D, Le-Huu P, Bilsing FL, Urlacher VB, Systematic screening of promiscuous P450 monooxygenases to access chemically different drug metabolites; *14th International Symposium on Biocatalysis and Biotransformations (BioTrans 2019)*, **2019**, Groningen, The Netherlands. (Poster)

Abstract

Cytochromes P450 (CYP or P450) are heme-b containing oxidoreductases that catalyze the reductive cleavage of molecular oxygen and oxidation of non-activated and activated C-H bonds of a vast variety of organic molecules. One oxygen atom is introduced into the substrate while the second one is reduced to water. The electrons required for catalysis are mainly transferred from NAD(P)H via redox partner proteins to the heme iron of the P450. P450s play an important role in the metabolism of xenobiotics and drugs as well as in the biosynthesis of secondary metabolites. Consequently, these enzymes are of biotechnological relevance as biocatalysts for the production of drug metabolites and other pharmaceutically relevant active ingredients. In this context, bacterial P450s from *actinomycetes* are of particular interest as they are either directly involved in the biosynthesis of pharmaceutically active compounds or can be used for the production of human drug metabolites.

In this thesis new P450s from *actinomycetes* were identified and characterized. Specifically, P450s from *Streptomyces platensis* and *Pseudonorcardia autotrophica* were investigated, as these bacteria are used by pharmaceutical companies for the production of drug metabolites. Based on the sequenced genomes, 70 putative P450 genes were identified in both strains. Sixteen of these P450 genes were heterologously expressed in *Escherichia coli* and subsequently tested for their activity against a set of drug compounds. A two-step screening was developed to identify promising enzyme candidates with a broad substrate spectrum and diversified product selectivity. By testing with the three structurally different model drugs ritonavir, testosterone, and amitriptyline, CYP105D and CYP107Z from *S. platensis* were identified as the most promising P450s in the first step. In the second step, structurally related molecules and derivatives of the model drugs were tested as substrates for CYP105D and CYP107Z. Thus, a correlation between the chemo- and regioselectivity of the P450 enzymes and the chemical structure of the target drugs could be shown.

In the next step, the cytochrome P450 complement (CYPome) of *S. platensis* was investigated. The CYPome of *S platensis* consists of 39 P450 genes, which were differentiated with respect to their phylogenetic families, possible functions in the secondary metabolism and heterologous expression in *E. coli*. Comparative activity studies with chemically diverse

drug substrates revealed a panel of cytochromes P450 from *S. platensis* with high potential for the production of drug metabolites.

In a further study, CYP105D from *S. platensis* was identified as a promising biocatalyst to oxidize bile acids. Bile acids such as chenodesoxycholic acid are converted with lower activity and selectivity compared to other steroid substrates like testosterone. The type and position of polar groups in the molecule seem to have a decisive influence as it was shown by testing bile acids with different functional groups. Oxidation of chenodeoxycholic acid has not been reported for any other bacterial P450 before. NMR analysis of the products provided information on the structure of new bile acid metabolites with a modified side chain.

Finally, the applicability of recombinant lyophilized *E. coli* cells expressing P450 enzymes and redox partners for biocatalysis was demonstrated. For this purpose, a whole-cell catalyst was constructed which allows the oxidation of testosterone by CYP105D. Subsequently, the regeneration of the cofactor NADH by a co-expressed alcohol dehydrogenase from *Rhodococcus erythropolis* (Re-ADH) was analyzed. The cofactor regeneration by ADH was found crucial for activity of lyophilized cells. The possibility to use P450s in lyophilized whole cell catalysts opens up new possibilities for a broader application of these enzymes for biocatalysis.

Zusammenfassung

Cytochrome P450 (CYP oder P450) sind Häm-b-haltige Oxidoreduktasen, die die reduktive Spaltung von molekularem Sauerstoff und die Oxidation von nicht aktivierten und aktivierten C-H Bindungen von einer Vielzahl organischer Moleküle katalysieren. Dabei wird ein Sauerstoffatom in das Substrat eingeführt, während das zweite zu Wasser reduziert wird. Die zur Katalyse benötigten Elektronen werden zumeist von NAD(P)H über Redoxpartnerproteine auf das Häm-Eisen der P450 übertragen. P450 spielen eine bedeutende Rolle bei der Verstoffwechselung von Xenobiotika und Arzneistoffen sowie in der Biosynthese von Sekundärmetaboliten. Folglich sind diese Enzyme als Biokatalysatoren von biotechnologischer Relevanz für die Herstellung von Arzneistoffmetaboliten und anderen pharmazeutisch relevanten Wirkstoffen. In diesem Zusammenhang sind bakterielle P450 aus *Actinomyceten* besonders interessant, da diese entweder direkt an der Biosynthese pharmazeutischer Wirkstoffe beteiligt sind oder für die Herstellung von humanen Arzneistoffmetaboliten eingesetzt werden können.

Im Rahmen dieser Arbeit wurden neue P450 aus *Actinomyceten* identifiziert und charakterisiert. Konkret wurden P450 aus *Streptomyces platensis* und *Pseudonorcardia autotrophica* untersucht, da diese Bakterien von pharmazeutischen Unternehmen zur Herstellung von Arzneistoffmetaboliten genutzt werden. Ausgehend von den sequenzierten Genomen wurden 70 putative P450 Gene in beiden Stämmen identifiziert. Sechzehn dieser P450 Gene wurden heterolog in *Escherichia coli* exprimiert und anschließend hinsichtlich ihrer Aktivität gegenüber einer Reihe von Arzneistoffen getestet. Um besonders vielversprechende Enzymkandidaten mit einem breiten Substratspektrum und unterschiedlicher Produktselektivität zu identifizieren, wurde ein zweischrittiges Screeningverfahren entwickelt. Durch Testung mit den drei strukturell verschiedenen Modellarzneistoffen Ritonavir, Testosteron und Amitriptylin wurden CYP105D und CYP107Z aus *S. platensis* im ersten Schritt als vielversprechendste P450 identifiziert. Im zweiten Schritt wurden strukturell ähnliche Moleküle und Derivate der Modelarzneistoffe als Substrate für CYP105D und CYP107Z getestet. Dadurch konnte ein Zusammenhang zwischen der Chemo-und Regioselektivität der P450 Enzyme und chemischer Struktur der Arzneistoffe gezeigt werden.

Im nächsten Schritt wurde das Cytochrom P450 Komplement (CYPome) von *S. platensis* weiter untersucht. Das CYPome besteht aus insgesamt 39 P450 Genen, die hinsichtlich ihrer phylogenetischen Familien, möglichen Funktionen im Sekundärstoffwechsel und heterologer Expression in *E. coli* differenziert wurden. Vergleichende Aktivitätsstudien mit chemisch unterschiedlichen Arzneistoffsubstraten zeigten, dass verschiedene Cytochrome P450 aus *S. platensis* ein hohes Potential zur Produktion von Arzneistoffmetaboliten besitzen.

In einer weiteren Studie wurde CYP105D aus *S. platensis* als vielversprechender Biokatalysator zur Oxidation von Gallensäuren identifiziert. Gallensäuren, wie Chenodesoxycholsäure, wurden im Vergleich zu anderen Steroidsubstraten, wie Testosteron, mit geringerer Aktivität und Selektivität umgesetzt. Einen entscheidenden Einfluss scheinen dabei die Art und die Position polarer Gruppen im Molekül zu haben, wie durch Testung von Gallensäuren mit unterschiedlichen funktionellen Gruppen gezeigt wurde. Die Oxidation von Chenodeoxycholsäure wurde bisher für keine andere bakterielle P450 nachgewiesen. Die NMR-Analyse der Produkte lieferte Hinweise auf die Struktur neuer Metaboliten von Gallensäuren, die an der Seitenkette modifiziert wurden.

Abschließend wurde die Anwendbarkeit von rekombinanten lyophilisierten *E. coli*-Zellen, die P450-Enzyme und Redoxpartner exprimieren, für die Biokatalyse demonstriert. Hierzu wurde zunächst ein Ganzzellkatalysator zur Oxidation von Testosteron durch CYP105D konstruiert. Anschließend wurde die Regenerierung des Kofaktors NADH durch eine koexprimierte Alkoholdehydrogenase aus *Rhodococcus erythropolis* (Re-ADH) analysiert. Es wurde festgestellt, dass die Kofaktorregeneration durch ADH entscheidend für die Aktivität lyophilisierter Zellen ist. Die Möglichkeit, P450 in lyophilisierten Ganzzellkatalysatoren einzusetzen, eröffnet neue Möglichkeiten für eine breitere Anwendung dieser Enzyme für die Biokatalyse.

Abbreviations

5-ALA	5-aminolevulinic acid	IS	internal standard
7α-HSDH	7α-hydroxysteroid dehydrogenase	kb	kilobases
A. annua	*Artemisia annua*	k_{cat}	turnover number
ACN	acetonitrile	K_D	dissociation constant
ADH	alcohol dehydrogenase	kDa	kilodalton
AlkL	membrane protein from *Pseudomonas putida*	KPi	potassium phosphate
antiSMASH	antibiotics and secondary metabolites analysis shell	LB	Lysogeny broth
APCI	atmospheric pressure chemical ionization	LC/MS	liquid chromatography/mass spectrometry
B. megaterium	*Bacillus megaterium*	MeOH	methanol
B. subtilis	*Bacillus subtilis*	MS	mass spectrometry
bp	base pair	m/z	mass-to-charge-ratio
BSA	*bovine* serum albumine	NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
cdw	cell dry weight	n. d.	not determined
CO	carbon monoxide	NMR	nuclear magnetic resonance
CPR	cytochrome P450 reductase	OD_{600}	optical density at 600 nm
CYP	cytochrome P450	*P. autotrophica*	*Pseudonocardia autotrophica*
CYPED	CYtochrome P450 Engineering Database	*P. putida*	*Pseudomonas putida*
CYPome	cytochrome P450 complement	P450	cytochrome P450
ddH_2O	double deionized water	P450 BM3	CYP102A1 from *Bacillus megaterium*
dH_2O	deionized water	$P450_{cam}$	CYP101A1 from *Pseudomonas putida*
DMSO	dimethyl sulfoxide	$P450_{terf}$	CYP107L from *Streptomyces platensis*
DNA	deoxyribonucleic acid	PAGE	polyacrylamide gel electrophoresis
dNTPs	deoxynucleotide triphosphates	PCR	polymerase chain reaction
E. coli	*Escherichia coli*	Pdr	putidaredoxin reductase from *Pseudomonas putida*
EC	enzyme class	Pdx	putidaredoxin from *Pseudomonas putida*
EDTA	ethylenediaminetetraacetic acid	PikC	CYP107L1 from *Streptomyces venezuelae*
ER	endoplasmatic reticulum	PMSF	phenylmethylsulfonyl fluoride
EtOH	ethanol	Re-ADH	ADH from *Rhodococcus erythropolis*
ESI	electro spray ionization	rev	reverse
EtOH	ethanol	rpm	rounds per minute
FAD	flavin adenine dinucleotide	RT	room temperature
Fdr	flavodoxin reductase from *Escherichia coli*	*S. platensis*	*Streptomyces platensis*
FhuA	ferric hydroxamate uptake protein from *Escherichia coli*	SDS	sodium dodecyl sulfate
FMN	flavin mononucleotide	SIM	selected ion monitoring
fwd	forward	*sp.*	*species*
FXR	farnesoid x receptor	Tris	tris(hydroxymethyl)aminomethane
GC/MS	gas chromatography/mass spectrometry	TTN	total turnover number
GDH	glucose dehydrogenase	U	unit
GPCR	G-protein-coupled receptors	UV	ultraviolet
HIV	human immunodeficiency virus	wt	wild type
HMBC	heteronuclear multiple bond correlation		
HSQC	heteronuclear single quantum coherence	YkuN	flavodoxin from *Bacillus subtilis*
IMAC	immobilized metal ion affinity chromatography	δ	chemical shift
IPTG	isopropyl β-D-1-thiogalactopyranoside	ε	molar absorption coefficient

1. Introduction

1.1 Biocatalysis

1.1.1 Definition and characteristics

Biocatalysis deals with the usage of enzymes and microbes for the synthesis of organic compounds.[1] The use of enzymes as 'green' biocatalysts compared to classical chemical catalysts is particularly attractive in view of growing demands for environmentally friendly alternatives in today's society.[1-2] Enzymes can accept a broad range of natural and non-natural substrates transforming them into valuable products, often with high chemo-, regio- and stereoselectivity. The reactions are typically catalyzed with high efficiency under environmental friendly conditions.[3] Hence, biocatalysts are used for various processes such as wastewater treatment or the production of food, polymers, fragrances, cosmetics, bulk chemicals and pharmaceuticals.[4] As many of these industrially relevant molecules differ from the natural substrate range of enzymes and are built up in a diverse and increasingly complex manner[5], biocatalytic processes have to be optimized for application as it is summarized by the biocatalysis cycle (Figure 1.1).[6] The biocatalytic cycle illustrates the relevant aspects of each phase for the establishment of a biocatalytic process in industry. Starting from the identification of a suitable enzyme candidate (biocatalyst selection), the respective biocatalyst has to be characterized and optimized in order to develop process-relevant parameters such as a suitable concept for application and product recovery. The ultimate goal is the implementation of a fast, economical and environmentally friendly process which requires a detailed understanding of each step within this cycle.[6] Especially in the pharmaceutical industry, biocatalysts are used in drug discovery and development, which are described in more detail in the next chapter.[7]

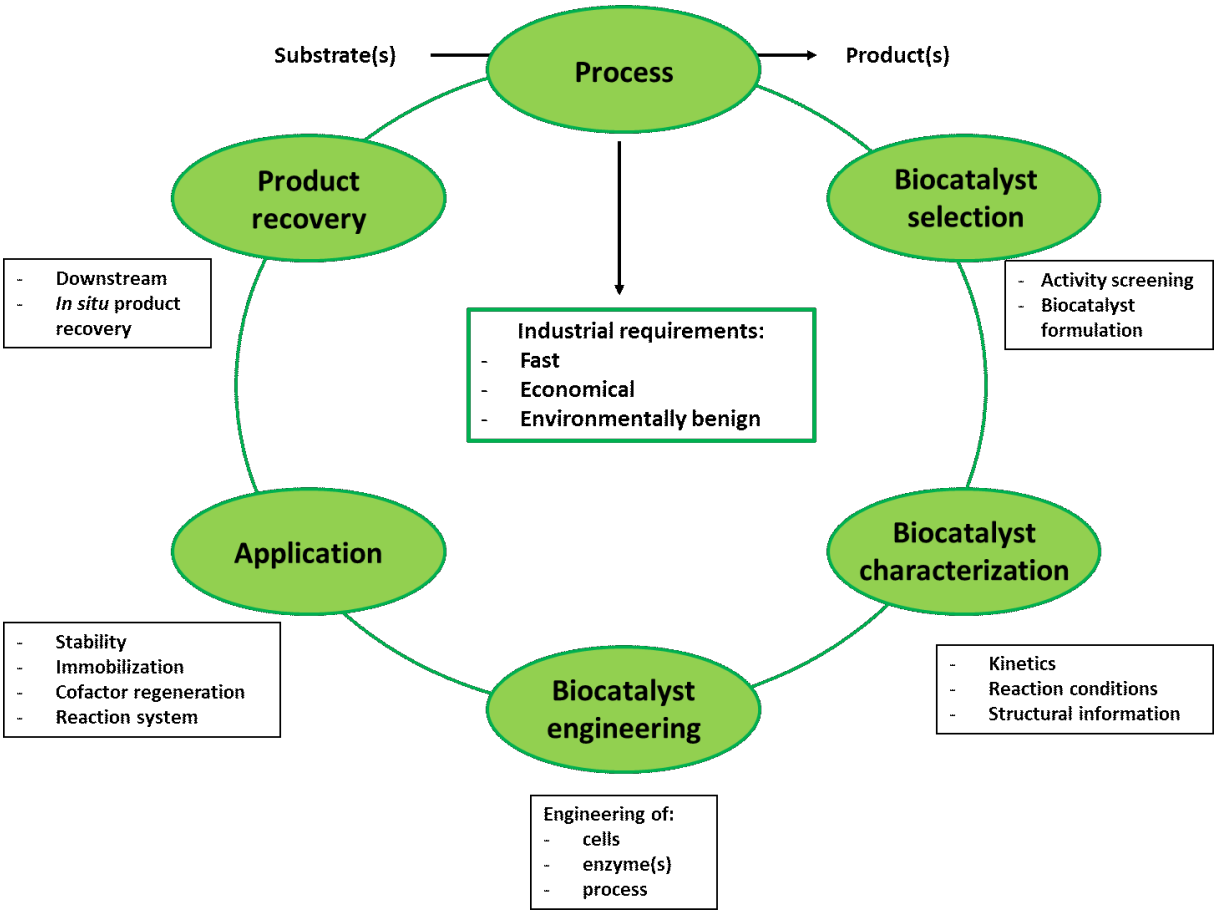


Figure 1.1: The biocatalysis cycle was originally proposed by A. Schmid, et al. [6] and updated based on recent literature.[1, 5] The biocatalysis cycle consists of six phases and highlights the different parameters which have to be optimized to establish a biocatalytic process.

1.1.2 Biocatalysis in drug discovery and development

All drug candidates must undergo a process commonly known as drug development pipeline prior to their commercial acquisition.[8] This process includes the initial identification of a desired drug compound, characterization of its biological action, pharmacokinetics, toxicity, metabolic fate, interaction with other drugs, optimization of its synthesis, and finally the preclinical and clinical evaluation (Figure 1.2).[7-8] The implementation of biocatalysts and biocatalytic concepts can be beneficial at various stages in drug discovery and drug development because the use of biocatalysts can reduce time and costs.[9] Once the drug target against a certain disease is defined, lead compounds have to be identified that show the desired biological activity towards the drug target. To this end, different strategies can be pursued such as the high-throughput screening of libraries obtained by combinatorial chemistry or natural product chemistry.[7] These processes are often laborious and expensive as many compounds have to be tested.

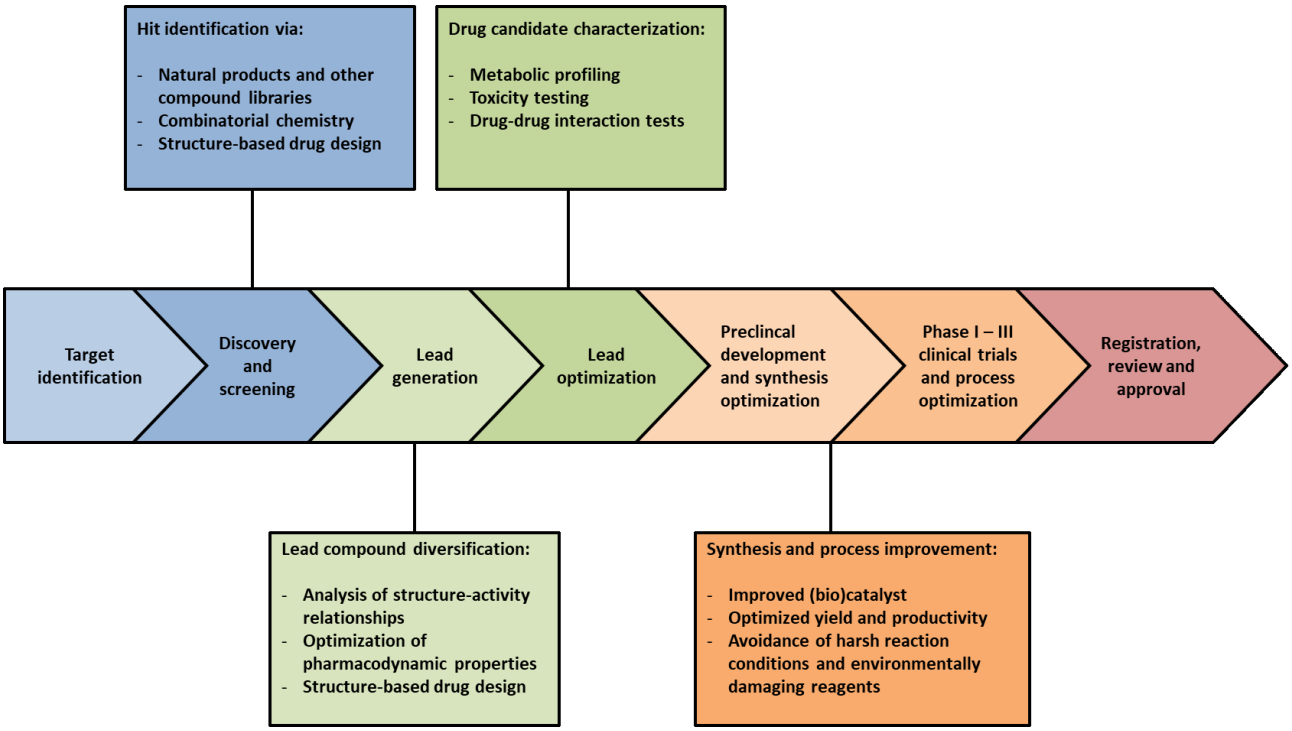


Figure 1.2: Drug discovery and development pipeline highlighting stages in which biocatalysis could be implemented.[7]

Initially identified lead structures are often not optimized with regard to their therapeutic efficacy, as they often exhibit side effects such as toxicity or undesired pharmacokinetics.[7] Therefore, lead structures are modified by altering certain functional groups. Biocatalysts have a big advantage to also convert derivatives of originally identified substrates maintaining their chemo-, regioselectivity and stereoselectivity. Because of this 'substrate promiscuity', biocatalysts can also be used for a shortened and cost efficient production of these derivatives.[9-10] Another bottleneck in medical chemistry is due to human enzymes metabolizing drugs to specific drug metabolites. These drug metabolites have to be produced in significant quantities for structural characterization, toxicological evaluation and drug-drug interactions tests.[7, 11] Although it is usually difficult to obtain sufficient quantities, biocatalytic approaches are still advantageous since the formed products are distinct and difficult to produce via chemical synthesis.[12]

Considering the general factors for optimizing a biocatalytic reaction, many different applications for enzymes as biocatalysts in the pharmaceutical industry can be defined. In

this respect, particular attention has been subjected to cytochromes P450. These are the most important enzymes in phase I metabolism given their catalytic versatility and substrate promiscuity.[7, 13]

1.2 Cytochromes P450

1.2.1 Characteristics and catalyzed reactions

Cytochromes P450 (CYP or P450) belong to the enzyme class of oxidoreductases (E.C: 1.14.-.-). Their name is derived from a broad absorption band at 450 nm, which results from a reduced carbon monoxide (CO)-bound form. This is caused by coordination of the CO to the heme b-iron in the active site and a conserved cysteine residue axially bound to the heme.[14]

Cytochromes P450 catalyze the introduction of one atom of oxygen into activated or non-activated C-H bonds. For the reduction of molecular oxygen (O_2) electrons are required that are delivered by redox partners from the cofactors nicotinamide adenine dinucleotide or nicotinamide adenine dinucleotide phosphate (NAD(P)H). The second atom of molecular oxygen is reduced to water, which is produced as a by-product[15] :

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

CYPs are capable of catalyzing a broad range of reactions in addition to the common hydroxylation of aliphatic and aromatic C-H bonds or C=C bond epoxidation. Depending on the chemical environment of the oxidized atom, rearrangements of oxidized products can lead to reactions such as C-C bond cleavage, dehalogenations, and N-, S- or O-dealkylations.[16] Amongst others, P450 enzymes and variants thereof can also catalyze heteroatom oxidations, dehydrogenation, cyclopropanation and intramolecular C-H amination.[16b, 17]

1.2.2 P450 nomenclature

Cytochromes P450 are found in all domains of life and some viruses.[16a, 18] Around 350.000 P450 sequences are annotated in databases from which more than 41.000 sequences are classified.[19] Because of the continuously growing number of P450 sequences, a systematic and standardized nomenclature for P450 genes was established, although for some prominent P450 enzymes including P450 BM3 (CYP102A1) or $P450_{cam}$ (CYP101A1) trivial names are still common.[20] This nomenclature by David Nelson is based on amino acid sequence similarities and consists of four parts for the specific assignment of a P450 (Figure 1.3). After the general abbreviation for cytochrome p450 (CYP), an Arabic number denotes the P450 family (*e.g.* CYP107). Members of a P450 family have a sequence identity of more than 40% in common. A sequence identity higher than 55% is the threshold for P450s which belong to the same subfamily indicated by a letter following the number of the family. The final number at the end of the combination names the individual P450 isoform within a subfamily and is incremented sequentially according to the order of discovery.[21]

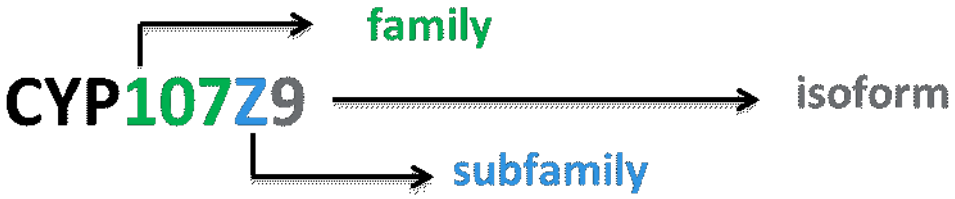


Figure 1.3: Nomenclature of cytochromes P450 according to David Nelson.[20]

1.2.3 P450 redox partner systems

In order to achieve reduction of inert molecular oxygen for substrate oxidation, cytochromes P450 rely on reducing equivalents which are usually delivered from NAD(P)H via redox partner proteins.[22] Cytochromes P450 can be divided into up to ten classes, depending on the organization of the electron transport chain and the redox partners involved.[14a] The three major redox partner systems are shown in Figure 1.4.

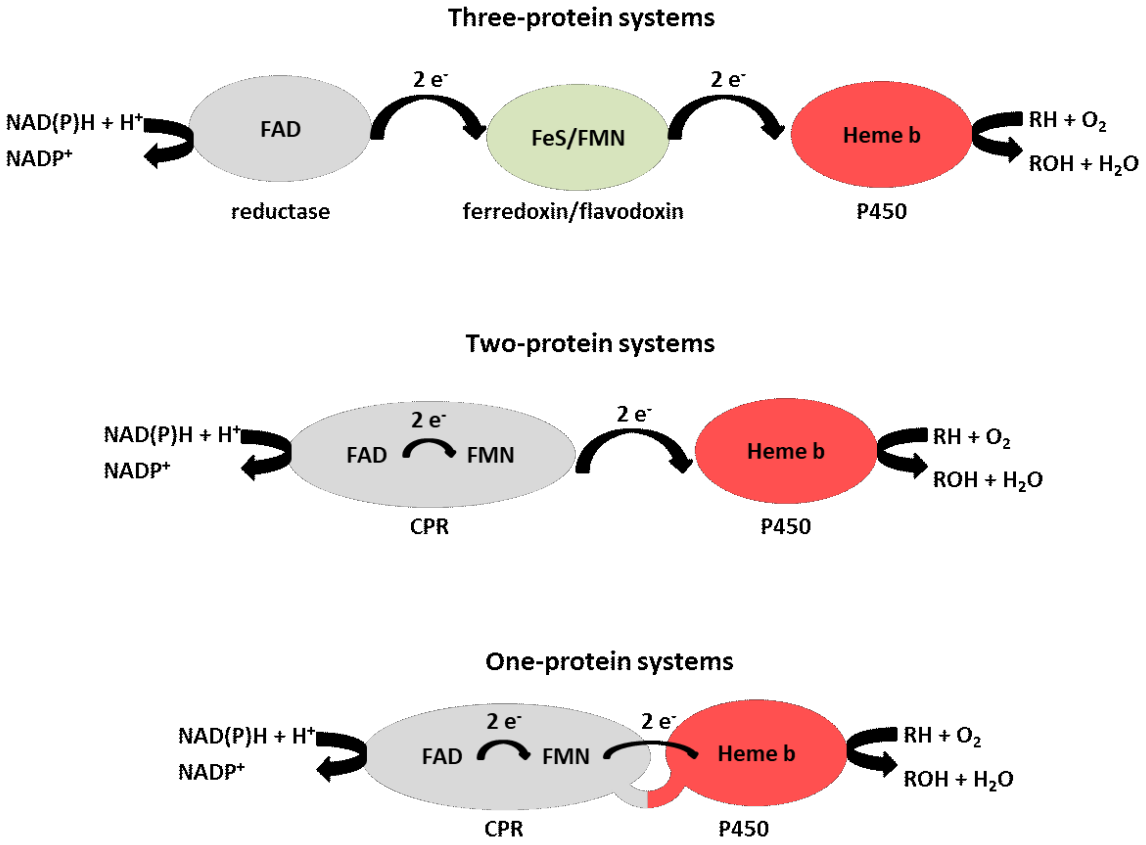


Figure 1.4: Schematic organization of cytochrome P450 redox partner systems. Three-protein systems consist of a reductase, a ferredoxin or flavodoxin and the P450. Two protein systems consist of a NADPH-dependent cytochrome P450 reductase (CPR) and the P450. One-protein systems are fusion proteins between a CPR-like reductase and the P450. Modified according to Urlacher and Girhard.[23]

Most bacterial and mitochondrial P450s belong to the three-protein system and require two additional redox partner proteins. Starting with NAD(P)H, reduction equivalents in form of a hydride are first transferred to a flavin adenine dinucleotide (FAD)-containing reductase. In the next step, the reductase delivers electrons to either a ferredoxin with an iron-sulfur cluster or a flavodoxin with flavin mononucleotide (FMN) as a cofactor. From the FMN or iron-sulfur cluster electrons are transferred to the heme b of the P450.[22] In bacterial three-protein systems, all components are localized soluble in the cytosol, while in mitochondrial systems the ferredoxin is soluble and P450 and the reductase are membrane-bound or membrane-associated.[14a]

Two-protein systems are typically found in the endoplasmic reticulum (ER) of eukaryotes. This system is comprised of the P450 and a FAD- and FMN- containing cytochrome P450 reductase, which are often bound to the inner membrane of the ER via a transmembrane

anchor. In prokaryotes and lower eukaryotes P450s belonging to the one-protein system can be found.[24] The characteristic of this class is the natural fusion of a FMN/FAD-containing reductase via a short linker region with the P450 in a single polypeptide chain. Most members of this class such as CYP102A1 (P450 BM3) from *Bacillus megaterium* are soluble, cytosolic proteins, whereas some eukaryotic members such as CYP505A1 from *Fusarium oxysporum* are membrane-bound.[14a, 25]

1.2.4 P450 catalytic mechanism

The common catalytic mechanism of P450s is described by formation of reactive intermediates and complexes (Figure 1.5).[26] In the resting state, water is bound as a sixth ligand to the heme iron present in its oxidation state Fe^{III} (**1**). Upon substrate binding, structural changes lead to a dissociation of the water ligand and a spin-state shift of the heme iron from low-spin to high-spin (**2**). The high spin Fe^{III} possesses a higher redox potential and can consequently accept the first electron from redox partner proteins to get reduced to Fe^{II} (**3**).[27]

In the next step, molecular oxygen (O_2) is bound to form the iron-(III)-superoxide complex (**4**). The delivery of the second electron from redox partner proteins results in another reduction to generate the iron-(III)-peroxo intermediate (**5**), which is then protonated to form the iron-(III)-hydroperoxo intermediate (**6**).[26b]

Protonation of the iron-(III)-hydroperoxo intermediate leads to a heterolytic cleavage of the dioxygen bond and the release of water. A highly reactive iron-(IV)-oxo species, known as compound I is formed (**7**), in which the radical cation is delocalized via the porphyrin ring and the thiolate ligand.[28] Starting from **compound I** the substrate gets oxidized by the so-called rebound mechanism.[29] After proton abstraction of the C-H bond, the substrate radical combines directly by 'rebound' with an iron-(IV)-hydroxyl species named **compound II** (**8**) to produce the oxygenated product and the one-electron reduced Fe^{III} complex (**9**).[29-30] Next to the dissociation of the product, water can coordinate again to the heme iron to regenerate the resting state (**1**). In addition to the described intermediates shunt reactions (dotted lines) can occur which uncouple the consumption of reducing equivalents and substrate oxidation, meaning no product is formed. In shunt reactions water

or reactive oxygen species such as hydrogen peroxide (H_2O_2) and the superoxide anion ($O_2^{\bullet -}$) are produced which can destabilize the P450 by bleaching of the heme or modification of the apoenzyme.[31]

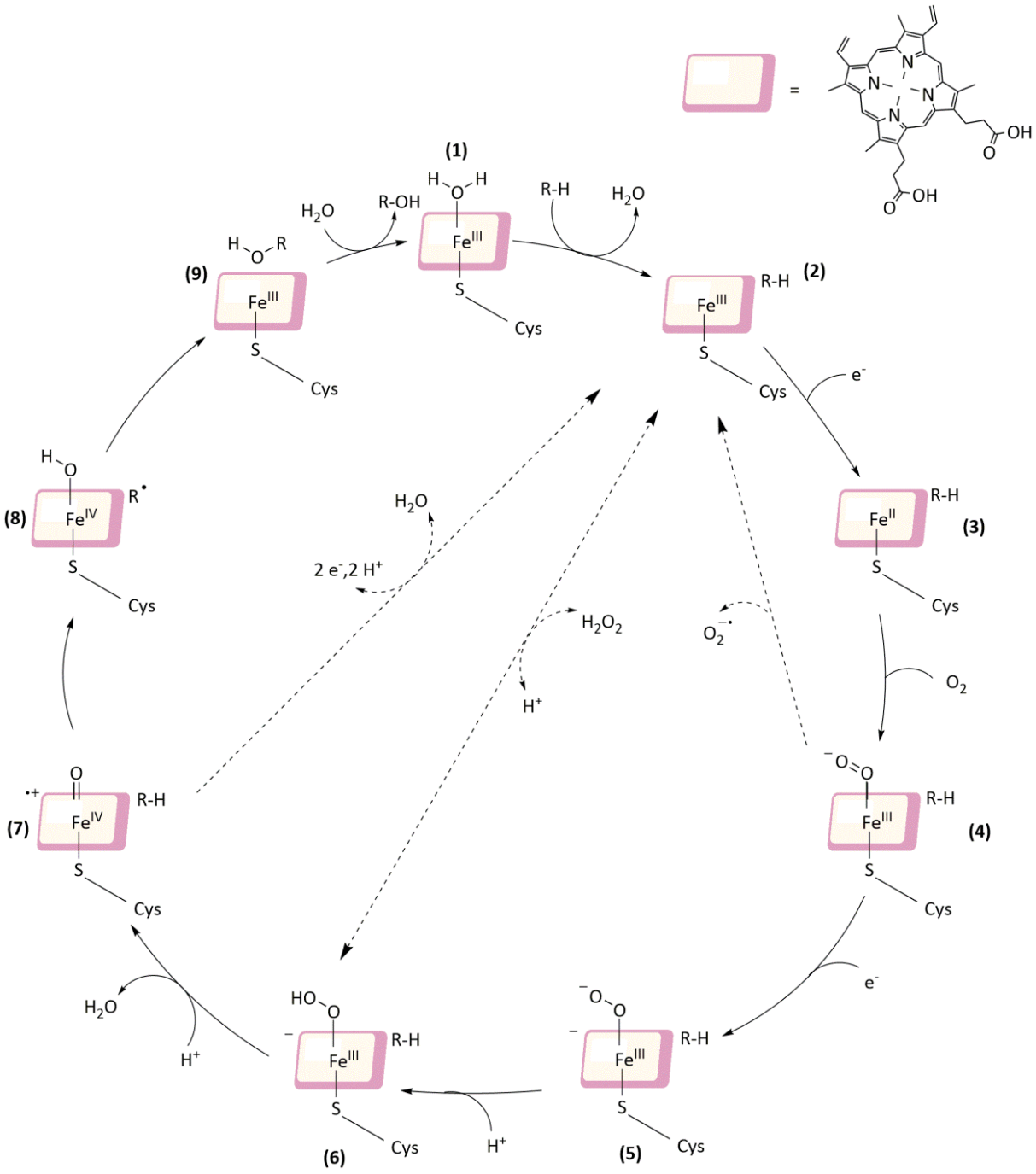


Figure 1.5: The P450 catalytic cycle. The heme iron (pink box) is axially linked to the protein via the sulfur atome of a cysteinate. (1) iron(III)-aqua-complex, (2) free iron(III)-complex, (3) iron(II)-(high spin)-complex, (4) iron(III)-superoxide-complex, (5) iron-(III)-peroxo intermediate, (6) iron-(III)-hydroperoxo intermediate, (7) compound I, (8) compound II (9) free iron(III)-complex. Shunt pathways are indicated with dotted lines. Modified according to Li *et al.*[22]

1.2.5 Industrial applications of P450s

P450s are involved in the metabolism of drugs and xenobiotics as well as in the synthesis of secondary metabolites and hormones.[19b, 22] Accordingly, these enzymes are especially interesting as biocatalysts for the pharmaceutical industry.[7] The hydroxylation of steroids by microbial strains expressing P450 enzymes was one of the first biocatalytic processes established in industry starting in the late 1940s (Figure 1.6).

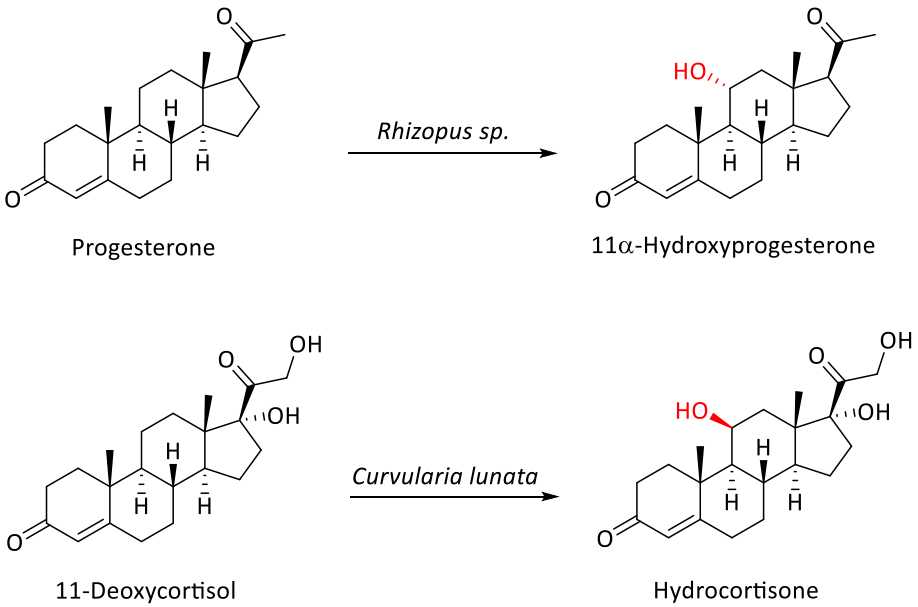


Figure 1.6: Microbial steroid oxidations used in industry. The biosynthesis of hydrocortisone from 11-deoxycortisol was applied by Schering AG (now Bayer); 11α-hydroxylation of progesterone to 11α-hydroxyprogesteron was commercialized by Pharmacia & Upjohn (now Pfizer). The inserted hydroxyl group is highlighted in red.

Steroidal compounds possess anabolic, anti-hormonal, anti-inflammatory, anti-rheumatic, contraceptive or sedative properties and represent one of the largest sectors in the pharmaceutical industry with global markets in the region of US$10 billion.[32] The microbial oxidation of steroids offered alternatives to classical chemical syntheses as reaction steps could be reduced and syntheses were performed more cost-efficiently and ecologically friendly.[33] Two microbial steroid oxidations (Figure 1.6) were commercialized namely the biosynthesis of hydrocortisone from 11-deoxycortisol with the fungus *Curvularia lunata* (used by Schering AG, now Bayer)[34] or the 11α-hydroxylation of progesterone to 11α-hydroxyprogesteron with the molds *Rhizopus arrhizus* and *Rhizopus nigricans* (used by

Pharmacia & Upjohn, now Pfizer).[35] Besides steroids, secondary metabolites from plants and microorganisms are sources for new drug compounds. In many biosynthetic pathways, P450s are involved in the diversification of respective secondary metabolites by catalyzing their oxidation with a high chemo-, regio-, and stereoselectivity.[36] A prominent example is the potent anti-malarial drug artemisinin from the plant *Artemisia annua*. In the biosynthesis of artemisinin its precursor artemisinic acid is formed by a three-step oxidation catalyzed by the P450 CYP71AV1 (Figure 1.7).[37] CYP71AV1 was implemented in a recombinant yeast strain used by Sanofi for the production of artemisinic acid on the ton scale.[37-38]

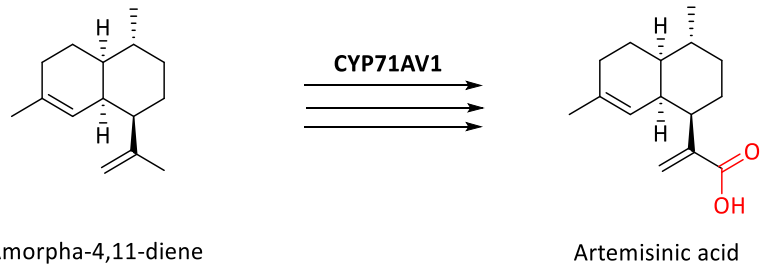


Figure 1.7: Oxidation of amorpha-4,11-diene to artemisinic acid catalyzed by the P450 CYP71AV1 from *A. annua*. Artemisinic acid is a precursor of the anti-malarial drug artemisinin and is produced by Sanofi in recombinant yeast.[37-38] The inserted carboxyl group is marked in red.

The evaluation of pharmacological effects and toxicity for drugs and their metabolites is crucial for drug development (chapter 1.1.2). P450s are the major drug-metabolizing enzymes in the human body because they oxidize about two-thirds of all drugs to make them more water soluble.[13] Consequently, the metabolic fate of numerous drugs was analyzed by preparing metabolites in P450-mediated biotransformations prior to industrial commercialization. The investigated drugs are chemically diverse and are used to treat *e.g.* HIV,[39] hypercholesterolemia,[40] major depressive disorders,[41] hypertension,[42] pain,[43] gastrointestinal complaints,[44] impotence[45] and cancer.[46] Selected examples leading to annual sales of more than 1 billion US dollars (so-called 'blockbuster drugs') are highlighted in Figure 1.8. [47]

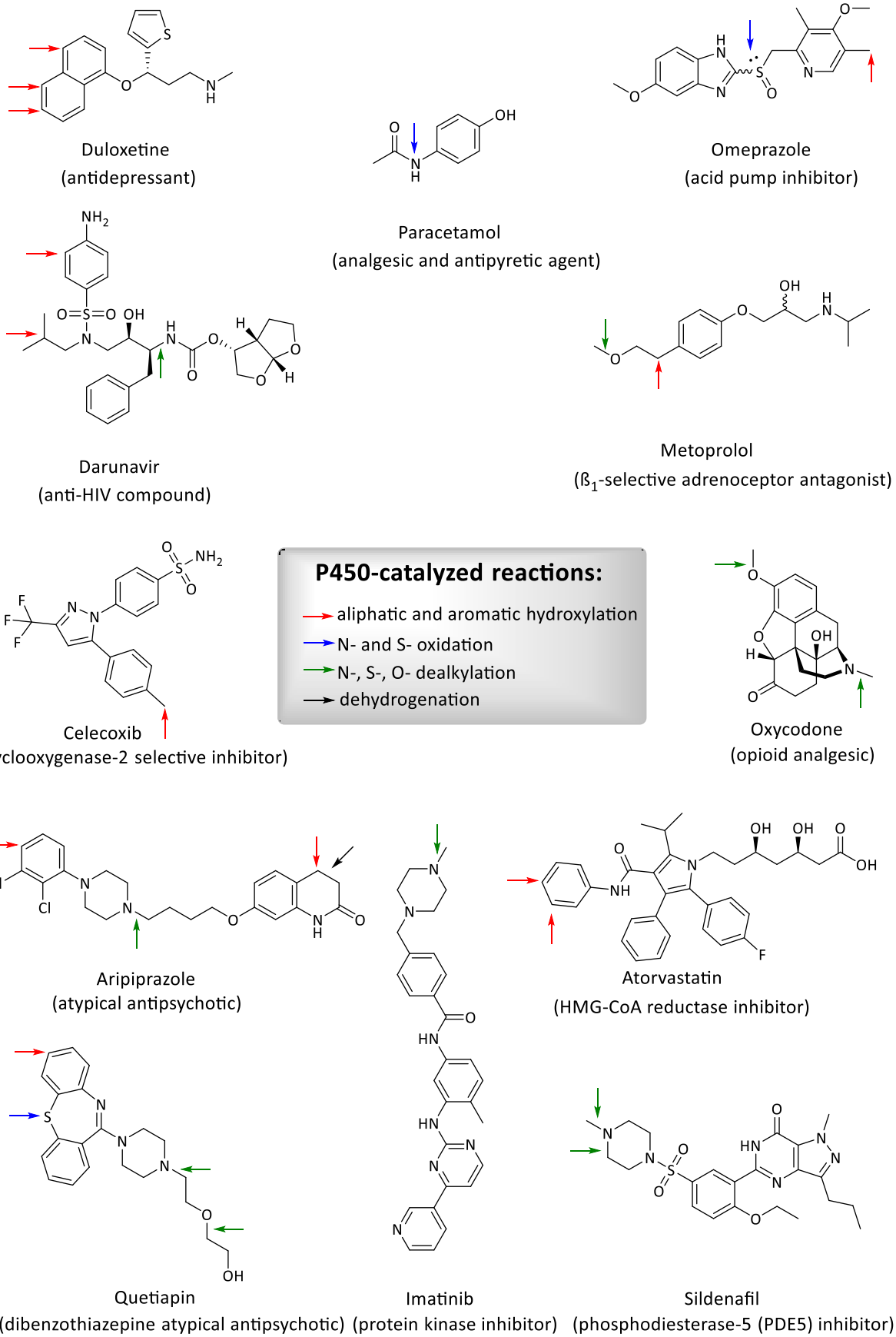


Figure 1.8: Selected examples of blockbuster pharmaceuticals, which are metabolized by human cytochromes P450.[39-46] The color code for each reaction type is displayed in the grey box. Major oxidation sites of human P450s are indicated by arrows.

1.2.6 Limitations of P450s as biocatalysts

Despite their enormous potential for the pharmaceutical industry, the application of P450s as biocatalysts is hampered by several challenges.[15, 48] These bottlenecks can be targeted by strategies to improve (1) the activity and selectivity of the P450, (2) the effectiveness of the electron transfer chain, and (3) general biocatalytic parameters such as the choice of the reaction system or the biocatalyst formulation (Figure 1.1):[15, 48-49]

(1) Improving the activity and selectivity of P450s

In comparison to other biocatalysts like lipases or alcohol dehydrogenases, most cytochromes P450 have low activities resulting in k_{cat} numbers in the range of 1 - 300 min^{-1}.[15, 50] The highest turnover numbers of 5,000 – 17,100 min^{-1} were measured for the bacterial P450 CYP102A1 (P450 BM3) in the oxidation of fatty acids.[51] However, the conversion of pharmaceutically interesting substrates was often only possible with eukaryotic P450s and not with bacterial enzymes like CYP102A1 because of their limited substrate spectrum.[15, 52] These observations led to the development of two basic strategies, both based on improvement of P450 properties through protein engineering. Most eukaryotic P450s are membrane-bound via their N-terminus which limits their expression and activity in heterologous hosts. The replacement or the truncation of the N-terminal sequence became a common procedure for optimized productivities of eukaryotic P450s such as the human-drug-metabolizing P450 CYP3A4[53] or the plant P450 CYP71AV1 mentioned above for artemisinic acid production.[49, 54] In addition other factors, such as their stability or their codon usage, are often not ideal, which makes tedious optimizations of eukaryotic P450s necessary before application.[55]

In contrast to eukaryotic enzymes, prokaryotic P450s are soluble, cytosolic enzymes, which can be easily produced in heterologous hosts. Variants of bacterial enzymes were designed either via rational protein design or directed evolution enabling oxidation of drugs and other pharmaceutical relevant compounds. The most frequently investigated bacterial P450 is CYP102A1, whose variants carry up to 20 mutations to achieve selective drug oxidations.[56] An alternative to engineering of known P450s is the exploration of new P450 sequences from sequence databases, which steadily increase by the availability of newly sequenced

genomes.[57] For their fast identification and characterization, the development of efficient screening methods is compulsory (chapter 1.2.7).

(2) Improving the electron transfer chain

Activation of molecular oxygen by P450s requires the transfer of two electrons naturally delivered by redox partner proteins. The inefficiency of the electron transfer chain leading to uncoupling reactions is a major issue for P450s (chapter 1.2.3). This problem is further aggravated when the natural redox partner proteins are not available due to lack of annotation, insufficient expression of the respective genes or low activity. Strategies circumventing the usage of additional redox partner proteins imply the application of alternative electron sources such as hydrogen peroxide, light or voltage.[58] However, these alternative methods suffer from other problems such as increased enzyme destabilization when using peroxides and electrochemical reduction or the need to use expensive, rare and toxic chemicals such as ruthenium-based photosensitizers.[15, 59] The construction of artificial fusion proteins is another method for improving electron transfer. Natural one-protein systems, including CYP102A1 (chapter 1.2.3), exhibit higher turnover numbers compared to non-fused P450 systems. Artificial fusions of P450s and redox partners have been constructed for improved and simplified enzyme performance. Examples described in the literature include fusions of CYP101A1 from *Pseudomonas putida* with its natural redox partner proteins Pdx and Pdr[60] or CYP51 from *Saccharomyces cerevisiae* with its native cytochrome P450 reductase (CPR).[61] The turnover numbers of such fusions do not substantially exceed those of non-fused proteins, since other factors, such as the linkage between redox partner and P450 or the intermolecular stoichiometry of redox partner proteins, also influence the activity.[62] Consequently, the testing of heterologous non-fused redox partners is still common although this strategy does not always result in high activity.[15]

(3) Reaction parameter optimization and cell engineering

The application of whole-cell biocatalysts represents the cheapest catalyst formulation.[63] Accordingly, whole-cell systems are the preferred biocatalyst form for P450s because they provide intracellular regeneration of NAD(P)H, possible co-expression of redox partner proteins and a protected environment to enhance enzyme stability.[64] Nevertheless, whole-cell systems can have several limitations which can be targeted by cell engineering or optimization of reaction parameters (Table 1.1).

Table 1.1: General and specific challenges in the application and process design of P450 based whole-cell biocatalysts. Modified according to Bernhardt and Urlacher.[15]

Limitation	Possible solution
General for P450s	
NAD(P)H requirement	Enzymatic cofactor regeneration [65]
Product degradation	Two-liquid phase systems[66]
Low substrate solubility	Two-liquid phase systems[67] Addition of co-solvents[68]
Specific for whole-cell biocatalysts	
Substrate or product toxicity	Alternative hosts with altered uptake systems[69]
Mass transfer across the cell membrane	Co-expression of transporters and uptake systems[70] Cell membrane permeabilization[71]

The intracellular level of NAD(P)H can become rate limiting, especially when high reaction rates are aimed for.[15] Co-expression of NAD(P)H regenerating enzymes such as glucose dehydrogenase or alcohol dehydrogenases can help to increase activity in P450 based whole-cell systems.[65] In addition, the typically hydrophobic substrates of P450s may cause further challenges, reflected *e.g.* by low substrate solubility, limited transfer of substrates and products across the cell membrane or substrate and product toxicity.

The testing of organic solvents either as co-solvent or as additional second liquid phase might improve substrate solubility, independent whether whole cells or isolated enzyme are

used.[67-68] In case of two-phase liquid systems product degradation can also be avoided as was demonstrated in conversions of α-pinene and limonene.[66] Substrate or product toxicity can be overcome by alternative hosts with altered uptake systems as it was for example shown in P450-mediated conversions for alkanes or monoterpenoids.[69] Cellular uptake systems for hydrophobic compounds can be further optimized by additional co-expression of transporter proteins such as AlkL or FhuA.[70] Alternatively, cells can be permeabilized with water-miscible organic solvents or detergents to improve substrate and product transfer across the cell membrane.[71]

In addition any biocatalyst engineering strategy must go hand in hand with reaction engineering. Especially in upscaling of biocatalytic reactions, parameters including pH, temperature, media, oxygen supply, and carbon source need to be investigated to achieve optimized productivity and stability of the process.[49, 72]

1.2.7 Screening for P450 activity

The development of screening methods is required for a fast identification and characterization of novel P450 enzymes. A selection of different screening approaches is displayed in Figure 1.9.

In silico methods such as molecular docking, cluster analysis or sequence alignments (Figure 1.9a) have the advantage that they can make rapid predictions for a large set of compounds in a high-throughput mode.[73] In bacteria and fungi, the genes responsible for a biosynthetic pathway or other biological function may be located closely together in gene clusters. This could help to narrow down the prediction of function and substrate of a P450 if the function of neighboring genes is known.[74] The comparison of amino acid sequences via alignments would be another approach to identify putative substrates for P450s. However, high levels of sequence identity are compulsory for this approach.[75] In case of available three-dimensional structures, molecular docking might be useful to identify a P450 substrate. Since many P450s exhibit major structural changes between the unbound and the substrate-bound form and the position for hydroxylation is not always close to the active site, this approach is often of limited use.[15, 75]

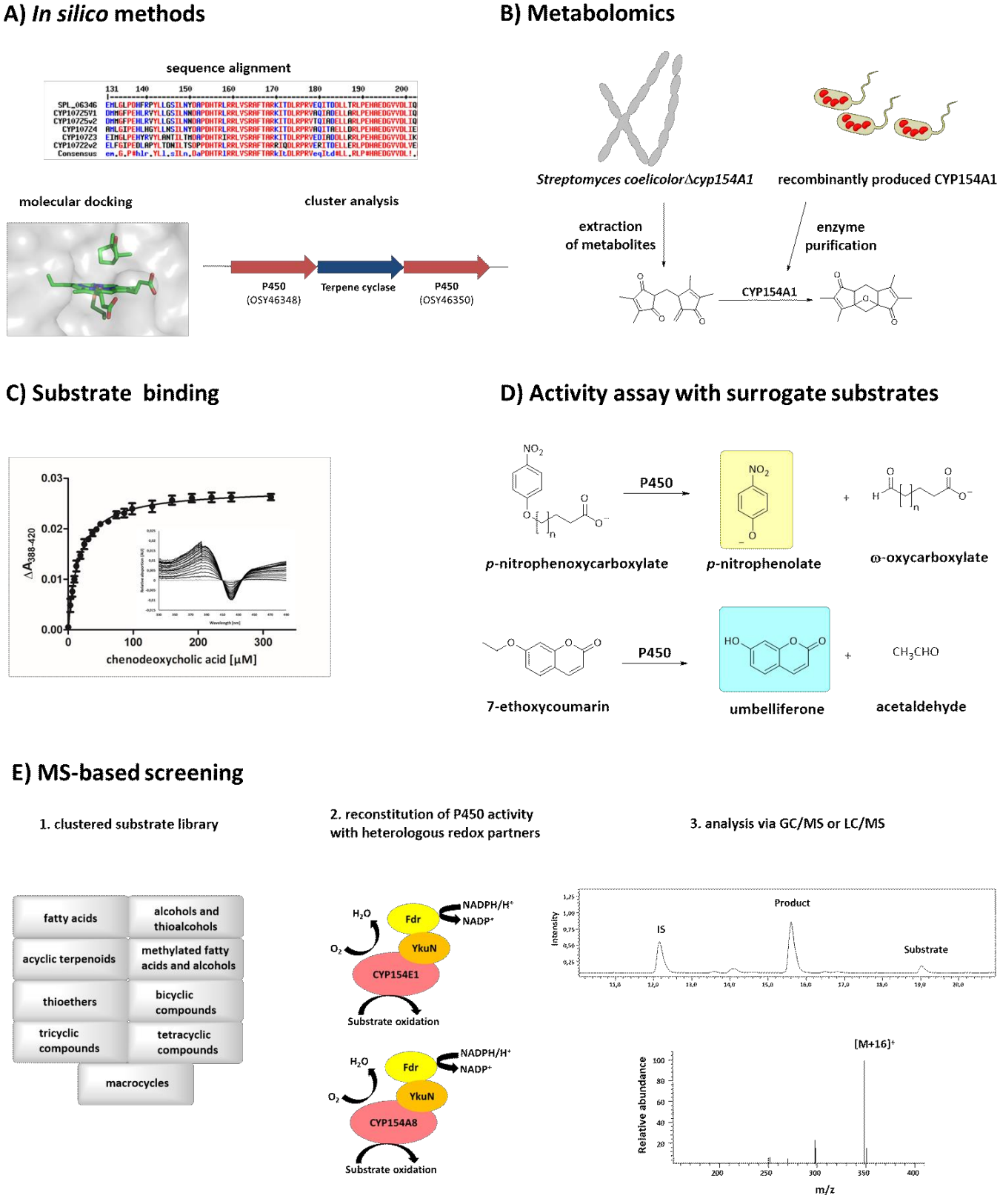


Figure 1.9: Selected examples of screening methods for P450 activities. A) *In silico* methods , B) Metabolomics C) Substrate binding indicates spectral changes D) Colorimetric activity assays E) MS-based screening approaches; Fdr: flavodoxin reductase from *E. coli*, YkuN: flavodoxin from *Bacillus subtilis*, CYP154E1: P450 154E1 from *Thermobifida fusca*, CYP154A8: P450 154A8 from *Nocardia farcinica*.

A method for screening of endogenous P450 substrates based on metabolomics was described by Guengerich and co-workers (Figure 1.9b).[76] In order to identify a physiological substrate of CYP154A1 from *Streptomyces coelicolor*, a mutant of *S. coelicolor* was designed in which the gene for *cyp154A1* was knocked out. After extraction of all metabolites of this knockout strain, the extract was incubated in an *in vitro* reaction with recombinantly produced CYP154A1. The analysis of the metabolic profile before and after the incubation identified a dipentaenone as substrate of CYP154A1, which undergoes an unexpected intramolecular cyclization to a Paternò-Büchi-like product.[75]

Upon substrate binding, cytochromes P450 undergo a spin-state shift of the heme iron from low-spin to high-spin (chapter 1.2.3). This spin-state shift is accompanied by an absorption shift from 418 nm (low spin substrate free form) to 390 nm (high spin substrate bound form), which can be used to record of the so-called type I binding spectra (Figure 1.9c). Based on substrate binding spectra compound libraries can be screened for P450 substrates as it was done for the bacterial P450s CYP106A2[77] and CYP260A1.[78] Nevertheless, this method is limited because, on one hand, not every substrate induces a type I spectrum and, on the other hand, compounds that show substrate binding are not necessarily converted by the P450.[79]

Surrogate substrates such as *p*-nitrophenoxycarboxylates or 7-ethoxycoumarine allow high-throughput screening for P450 activity.[80] After or during P450 oxidation, a chromophore (*e.g. p*-nitrophenolate)[81] or a fluorophore (*e.g.* umbelliferon)[82] is formed, which can be measured via UV/VS- and fluorescence spectroscopy, respectively (Figure 1.9d). The limitation of surrogate substrates is that the activity against the surrogate substrate is not necessarily representative for the actual substrate and that these kinds of substrates are not available for every compound of interest.[83]

Gas chromatography/mass spectrometry (GC/MS) or liquid chromatography/mass spectrometry (LC/MS) based methods have a lower throughput than colorimetric assays or *in silico* methods but allow direct analysis of P450 substrates for product formation and distribution. One particular screening method named ‘cluster screening’ (Figure 1.9e) was used for the characterization of CYP154E1 and CYP154A8.[79a] A substrate library of 51 compounds was organized into nine groups according to their chemical properties. The

differences in chemical structure, size and functional groups made it possible to investigate the structural requirements necessary for conversion by P450. Hence, detailed information about the substrate spectrum and chemo- and regioselectivity could be obtained.[79a]

1.3 *Actinomycetes* as promising source for P450 enzymes

Actinomycetes are an order of the class *Actinobacteria*.[84] *Actinomycetes* represent one of the largest and most diverse bacteria found in soil, compost or marine habitats.[85] Most of these Gram-positive bacteria have a high GC-content ranging from 51% to over 70% and are known for their unsurpassed capacity to produce secondary metabolites with diverse biological activities.[85b, 86] As soil microbes, *actinomycetes* can also degrade xenobiotics such as pharmaceuticals, agro-chemicals, and environmental pollutants.[85b, 87] Cytochromes P450 of *actinomycetes* are involved both in the synthesis of secondary metabolites and in xenobiotic catabolism. In the biosynthesis of natural products P450s catalyze a broad range of physiologically important oxidative reactions with high chemo-, regio-, and stereoselectivity (Figure 1.10).[88]

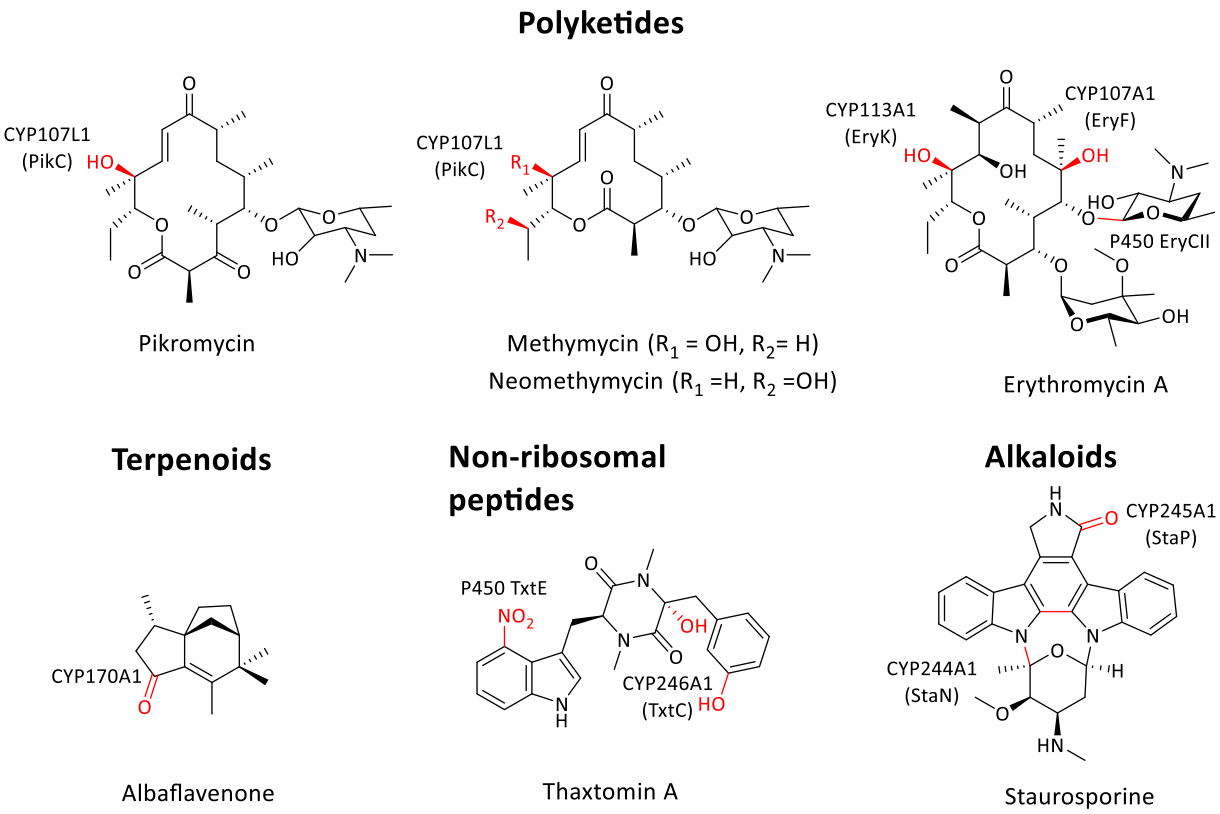


Figure 1.10: Examples of secondary metabolites from *actinomycetes*, which are oxidized by cytochrome P450s: CYP107L1 (PikC) from *Streptomyces venezuelae* ATCC 15439,[89] CYP113A1 (EryK), CYP107A1 (EryF) and EryCII from *Saccharopolyspora erythraea* NRRL 2338,[90] CYP170A1 from *Streptomyces coelicolor* A3(2),[91]CYP246A1 (TxtC) and TxtE from *Streptomyces scabies* 87.22,[92] CYP244A1 (StaN) and CYP245A1 (StaP) from *Streptomyces longisporoflavus* DSM 10189[93] and *Streptomyces* sp. TP-A0274[94]. The P450-catalysed reactions are highlighted in red.[88]

A well investigated example is CYP107L1 (PikC) from *Streptomyces venezuelae.* This P450 demonstrates unique substrate flexibility and catalyzes the hydroxylation of polyketides with high selectivity to produce antibiotics such as pikromycin, methmycin, and neomethmyin.[89] Polyketides and other natural products can also be modified by several P450s as it is the case during erythromycin A biosynthesis.[90] Further, P450s from *actinomycetes* are also often associated with other the biosynthesis of other secondary metabolites such as terpenoids,[91] non-ribosomal peptides,[92] and alkaloids.[94]

The investigation of alternative degradation pathways in *actinomycetes* revealed the production of metabolites that are equal to human P450 metabolites (Figure 1.11).

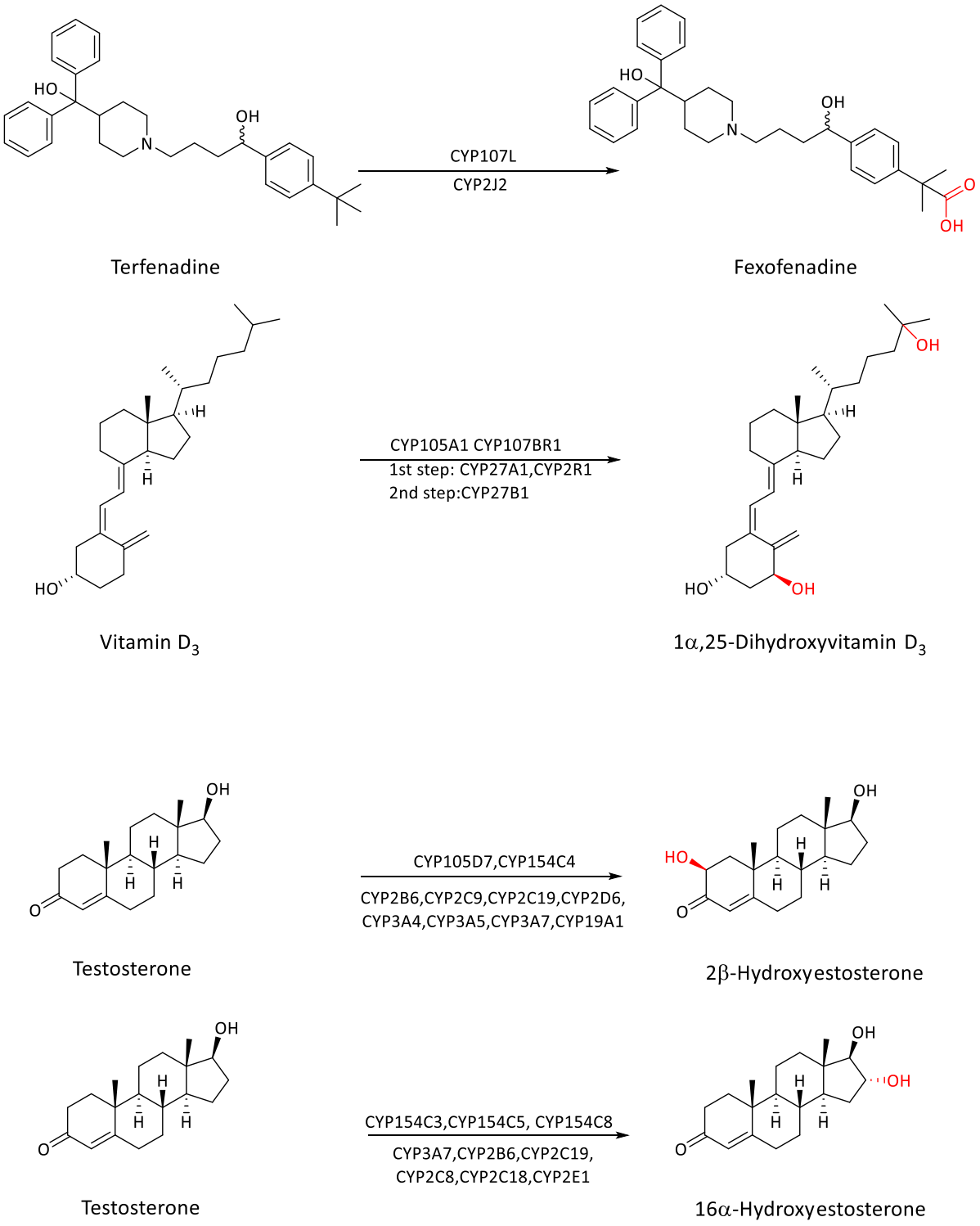


Figure 1.11: Comparison of drug oxidation by human P450s and P450s from *actinomycetes*. 1) Oxidation of terfenadine to fexofenadine by CYP107L from *Streptomyces platensis* NRRL2364 or human CYP2J2.[95] 2) Oxidation of vitamin D_3 to 1α,25-dihydroxyvitamin D_3 by CYP105A1 from *Streptomyces griseolus* ATCC 11796,[96] CYP107BR1 from *Pseudonocardia autotrophica* NBRC 12743[97] or human CYP27A1, CYP2R1 and CYP27B1.[97] 3) Oxidation of testosterone to 2β-hydroxytestosterone by CYP105D7 from *Streptomyces avermitilis* MA4680,[98] CYP154C4 from *Streptomyces sp.*[99] or human CYP2B6, CYP2C9, CYP2C19,CYP2D6,CYP3A4,CYP3A5,CYP3A7 and CYP19A1.[100] 4) Oxidation of testosterone to 16α-hydroxytestosterone by CYP154C3 from *Streptomyces griseus* IFO13350,[101] CYP154C5 from *Nocardia farcinica* IFM 10152,[102] CYP154C8 from *Streptomyces sp.* W2233-SM[103] or human CYP3A7, CYP2B6,CYP2C19,CYP2C8,CYP2C18 and CYP2E1.[100] The oxidation sites are highlighted in red.

This includes compounds like fexofenadine, which is the pharmacologically active metabolite of the anti-histamine drug terfenadine. $P450_{terf}$ (CYP107L) from *Streptomyces platensis* catalyzes the multi-step oxidation of terfenadine to the corresponding carboxylic acid fexofenadine, mimicking the activity of the human P450 CYP2J2.[95] Furthermore P450s from *actinomycetes* allow reduction of steps used for the synthesis of drug metabolites as shown for the prohormone vitamin D_3. Vitamin D_3 is hydroxylated at two positions by at least two different human P450s to produce the active form 1α,25-dihydroxyvitamin D_3. CYP105A1 from *S. griseolus* and CYP107BR1 from *Pseudonocardia autotrophica* are both able to oxidize vitamin D_3 directly to 1α,25-dihydroxyvitamin D_3 avoiding the usage of an additional enzyme compared to the human metabolism.[96-97] Drugs can also be oxidized at different positions of the molecule, because P450s have different regio- and stereoselectivities. This leads to metabolites with possible different biological activities. A very well investigated substrate in this regard is testosterone. Steroids including testosterone were tested with different P450 enzymes from *actinomycetes* providing hydroxylated steroid metabolites such as 2β-hydroxytestosterone or 16α-hydroxytesterone among others.[98-99, 101-103]

The majority of P450s present in *actinomycetes* belong to the CYP105 and CYP107-subfamilies, which members catalyze the oxidation of chemically diverse substrates (see Figures 1.10 and 1.11).[88a] Because of their low sequence and structural similarities, the identification of drug metabolizing P450s is the major hurdle for biotechnological applications and consequently an efficient screening system has to be provided (1.2.7).[104]

1.4 Aim of this work

Cytochromes P450 are attractive enzymes for the preparation of human drug metabolites, hydroxylated synthetic precursors, and natural products with interesting pharmaceutical activities (1.1.2 and 1.2.5).[19b, 105] Applications with bacterial P450 enzymes are especially attractive as they can be easily produced in heterologous hosts and generally possess a higher activity compared to eukaryotic P450s (1.2.6).[12b] P450s from *actinomycetes* are especially interesting as they can be used for the production of natural compounds and human drug metabolites (1.3). Increasing numbers of sequenced genomes of *actinomycetes* revealed a large number of genes encoding putative cytochrome P450s.[104] Although these enzymes might be promising candidates for biotechnological application, their substrate scope, product selectivities, and biochemical properties often remain unknown.

To gain access to those novel putative P450 enzymes a detailed analysis and characterization besides genome sequence identity is needed. Therefore the objective of this thesis was to identify P450 enzymes from *actinomycetes* and characterize them in order to pave the way towards biocatalytic applications. On the basis of drug-metabolizing *actinomycetes*, genomes of which were sequenced in the framework Era-IB project 'Integrative Approach to Promote Hydroxylations with Novel P450 Enzymes for Industrial Processes' (HyPerIn), the aims of this study were:

1) to identify P450 genes in genomes of drug-metabolizing *actinomycetes* (2.1)

2) to develop a screening method to identify P450s with drug-metabolizing activities and complementary selectivities (2.2)

3) to analyze all P450s of one drug-metabolizing *actinomycete* to compare (heterologous) expression, substrate scope and product selectivity (2.3)

4) to investigate bile acids as substrates regarding the production of new metabolites (2.4)

5) to demonstrate the applicability of recombinant lyophilized *E. coli* cells expressing P450 enzymes and redox partners for biocatalysis (2.5)

2. Results

The following five chapters describe the results of this thesis. Among these, the results presented in chapters 2.1 and 2.2 are already published, whereas the other chapters contain currently unpublished results. The chapters are presented in a sequential manner starting with the identification of drug-metabolizing cytochrome P450s (2.1 and 2.2), continuing with their characterization (2.2, 2.3 and 2.4) and finally their application (2.5). The own contribution to each study is given at the beginning of each chapter.

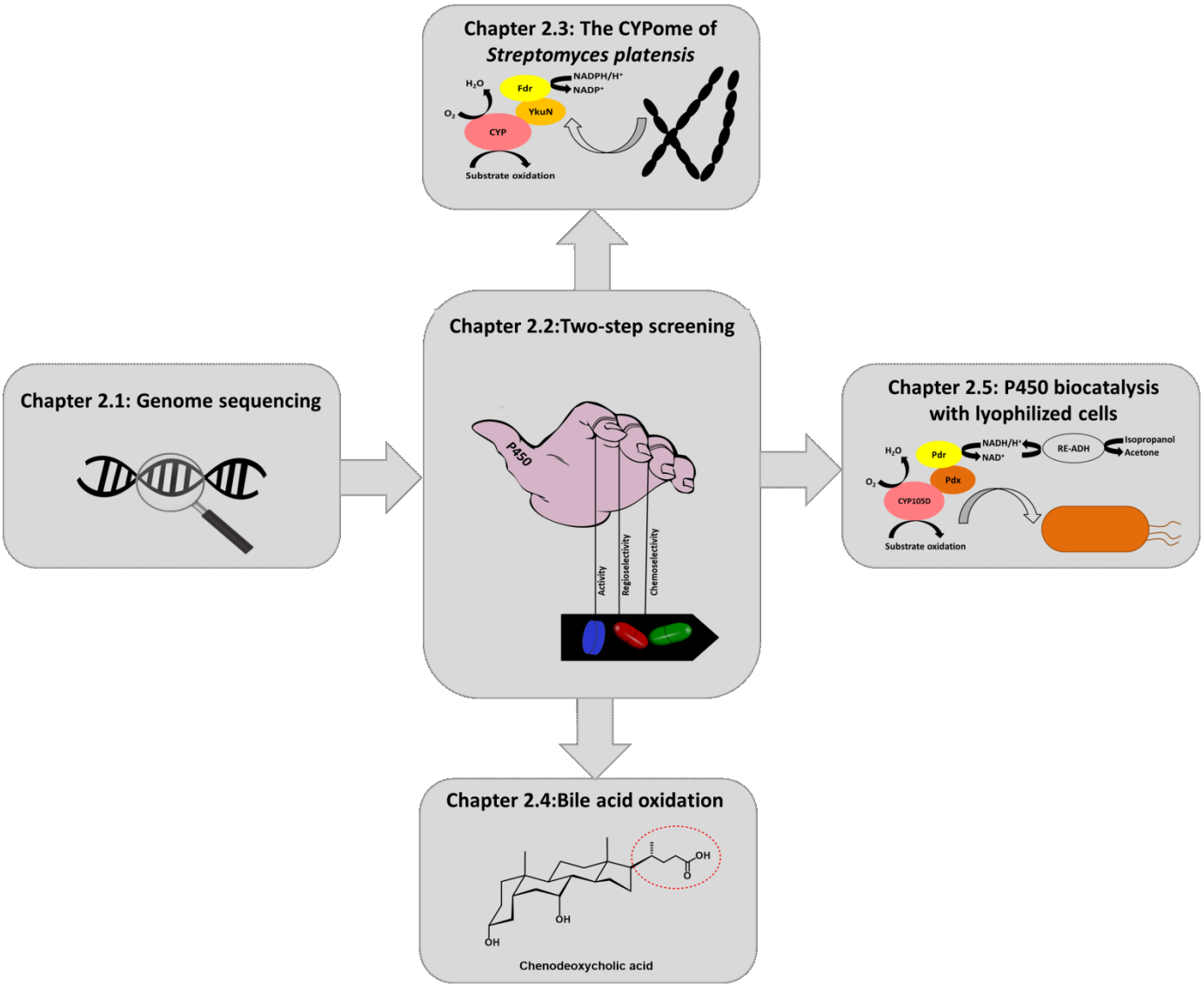


Figure 2.0.1: Schematic overview of this thesis.

2.1 Genome sequencing

Title: Draft genome sequences of three *Actinobacteria* strains presenting new candidate organisms with high potentials for specific P450 cytochromes

Authors: Christian Grumaz*, Yevhen Vainshtein, Philipp Kirstahler, Stephan Luetz, Matthias Kittelmann, Kirsten Schroer, Fabian K. Eggimann, Rico Czaja, Andreas Vogel, **Thomas Hilberath**, Anne Worsch, Marco Girhard, Vlada B. Urlacher, Marcel Sandberg, Kai Sohn*

* corresponding authors

published in: *Genome Announcements,* **2017***,* 5, e00532-00517
DOI: 10.1128/genomeA.00532-17

Own contribution: Identification of all P450 genes in the three genome sequences. Relative contribution: 10%.

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Draft Genome Sequences of Three *Actinobacteria* Strains Presenting New Candidate Organisms with High Potentials for Specific P450 Cytochromes

Christian Grumaz,[a] Yevhen Vainshtein,[a] Philipp Kirstahler,[a] Stephan Luetz,[b]* Matthias Kittelmann,[b] Kirsten Schroer,[b] Fabian K. Eggimann,[b] Rico Czaja,[c] Andreas Vogel,[c] Thomas Hilberath,[d] Anne Worsch,[d] Marco Girhard,[d] Vlada B. Urlacher,[d] Marcel Sandberg,[e] Kai Sohn[a]

Department of Molecular Biotechnology, Fraunhofer Institute for Interfacial Engineering and Biotechnology, Stuttgart, Germany[a]; Novartis Pharma AG, Novartis Institutes of Biomedical Research, Basel, Switzerland[b]; c-LEcta GmbH, Leipzig, Germany[c]; Heinrich Heine University, Institute of Biochemistry, Düsseldorf, Germany[d]; Kappa Bioscience, Oslo, Norway[e]

ABSTRACT The three *Actinobacteria* strains *Streptomyces platensis* DSM 40041, *Pseudonocardia autotrophica* DSM 535, and *Streptomyces fradiae* DSM 40063 were described to selectively oxyfunctionalize several drugs. Here, we present their draft genomes to unravel their gene sets encoding promising cytochrome P450 monooxygenases associated with the generation of drug metabolites.

Received 1 May 2017 **Accepted** 15 May 2017 **Published** 13 July 2017

Citation Grumaz C, Vainshtein Y, Kirstahler P, Luetz S, Kittelmann M, Schroer K, Eggimann FK, Czaja R, Vogel A, Hilberath T, Worsch A, Girhard M, Urlacher VB, Sandberg M, Sohn K. 2017. Draft genome sequences of three *Actinobacteria* strains presenting new candidate organisms with high potentials for specific P450 cytochromes. Genome Announc 5:e00532-17. https://doi.org/10.1128/genomeA.00532-17.

Address correspondence to Christian Grumaz, christian.grumaz@igb.fraunhofer, or Kai Sohn, kai.sohn@igb.fraunhofer.de.

* Present address: Stephan Luetz, TU Dortmund University, Department of Biochemical and Chemical Engineering, Dortmund, Germany.

Hydroxylation of C-H bonds can lead directly to the formation of high-value chiral compounds in demand as specialty chemicals and pharmaceutical synthons. In this context, cytochrome P450 monooxygenases (CYPs) remain unsurpassed in their targeted specificity and scope (1–3). Consequently, the application of CYPs in synthetic organic chemistry is considered to be "potentially the most useful of all biotransformations" (4). Thus, the discovery of novel target activities by isolation and characterization of new enzymes obviously plays a crucial role in the development of CYP-based biooxidation applications (5, 6). Comprehensive genetic screening of promising strains will widely expand the repertoire of specific CYPs, enabling not only access to a range of completely new compounds but also increasing the efficiency of already-established reactions.

In a screen of conspicuous *Actinobacteria* with certain drug compounds, three strains were identified to catalyze several new biooxidation products: *Streptomyces platensis* DSM 40041, *Pseudonocardia autotrophica* DSM 535, and *Streptomyces fradiae* DSM 40063 (data not shown). For this reason, we screened the genomic contents of each of these strains for novel *cyp* genes. Here, we describe their draft genome sequences, with a special focus on the identification of putative CYPs. For this purpose, extracted DNA was prepared for Illumina HiSeq 2500 with the Nextera DNA kit using the standard protocol. Sequencing was performed in paired-end mode with 2 × 250 cycles for DSM 40041 and DSM 535 and 2 × 150 cycles for DSM 40063. Illumina reads were removed for contaminations, adapters, and low-quality sequences with BBDuk from the BBMap package version 34.41 (http://sourceforge.net/projects/bbmap/), resulting in 662 Mb (DSM 40041), 1,111 Mb (DSM 535), and 2,619 Mb (DSM 40063) of trimmed reads used for assembly either with GS *de novo* Assembler version 2.9 (DSM 40041 and DSM 535) or with ABySS version 1.5.2 (DSM 40063). Gene annotation was carried out using Prokka 1.11 (7). The properties of the draft genomes are summarized in Table 1.

TABLE 1 Properties of the draft genomes for DSM 40041, DSM 535, and DSM 40063

Species name	Strain name	Bioproject sample no.	Accession no.	Coverage (×)	No. of contigs	N_{50} (kb)	Size (Mb)	G+C content (%)	No. of proteins
Streptomyces platensis	DSM 40041	SAMN05722965	MIGA00000000	79	127	115	8.4	71.2	7,302
Amycolata autotrophica	DSM 535	SAMN05722966	MIGB00000000	151	117	146	7.4	73.0	6,860
Streptomyces fradiae	DSM 40063	SAMN05722967	MIFZ00000000	390	352	31	6.7	72.2	5,799

To unravel putative CYPs, we first searched for oxidoreductases assigned to E.C. 1.14.-.-, which are related to oxyfunctionalization of C-H bonds, and then checked them for CYP family relationships using CYPED (https://cyped.biocatnet.de/). Out of the overall 294 found oxidoreductases, we were able to assign 90 targets to certain CYP families defining them as bona fide novel CYPs. The identified gene sequences will be heterologously expressed and characterized and may then be used for further oxyfunctionalization of different compounds in a variety of biocatalytic applications.

Accession number(s). The whole-genome shotgun projects have been deposited in GenBank under the accession numbers specified in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

The financial support for this work carried out under the umbrella of the ERA-IB2 3rd call project "hyperin" (project number EIB.12.026) was provided by the Federal Ministry of Education and Research to Fraunhofer Stuttgart (grant number 031A223B) and Heinrich Heine University Düsseldorf (grant number 031A223A).

REFERENCES

1. Schroer K, Kittelmann M, Lütz S. 2010. Recombinant human cytochrome P450 monooxygenases for drug metabolite synthesis. Biotechnol Bioeng 106:699–706. https://doi.org/10.1002/bit.22775.
2. Martinez CA, Rupashinghe SG. 2013. Cytochrome P450 bioreactors in the pharmaceutical industry: challenges and opportunities. Curr Top Med Chem 13:1470–1490. https://doi.org/10.2174/15680266113139990111.
3. Schulz S, Girhard M, Urlacher VB. 2012. Biocatalysis: key to selective oxidations. ChemCatChem 4:1889–1895. https://doi.org/10.1002/cctc.201200533.
4. Leak DJ, Sheldon RA, Woodley JM, Adlercreutz P. 2009. Biocatalysts for selective introduction of oxygen. Biocatal Biotransform 27:1–26. https://doi.org/10.1080/10242420802393519.
5. Furuya T, Kino K. 2010. Genome mining approach for the discovery of novel cytochrome P450 biocatalysts. Appl Microbiol Biotechnol 86:991–1002. https://doi.org/10.1007/s00253-010-2450-5.
6. Urlacher VB, Girhard M. 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. Trends Biotechnol 30:26–36. https://doi.org/10.1016/j.tibtech.2011.06.012.
7. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

2.2 Screening for identification of drug-metabolizing cytochromes P450

Title: Two-step screening for identification of drug-metabolizing bacterial cytochromes P450 with diversified selectivity

Authors: **Thomas Hilberath**, Leonie M. Windeln, Davide Decembrino, Priska Le-Huu, Florestan L. Bilsing, Vlada B. Urlacher*

* corresponding author

Published in: *ChemCatChem,* **2020***,* 12, 1710- 1719

DOI: 10.1002/cctc.201901967

Own contribution: Conceptualization, design and conduction of most of the experiments, analysis and interpretation of all data, supervision of the master theses of Leonie Windeln and Davide Decembrino, drafting of the manuscript. Relative contribution: 75%.

ChemPubSoc Europe

DOI: 10.1002/cctc.201901967

CHEMCATCHEM Full Papers

Check for updates

Two-step Screening for Identification of Drug-metabolizing Bacterial Cytochromes P450 with Diversified Selectivity

Thomas Hilberath,[a] Leonie M. Windeln,[a] Davide Decembrino,[a] Priska Le-Huu,[a] Florestan L. Bilsing,[a] and Vlada B. Urlacher*[a]

The evaluation of drug metabolites is compulsory during drug development. Since recently, bacterial cytochromes P450 and their mutated variants have attracted considerable interest as an alternative to hepatic P450s for the synthesis of human drug metabolites. Thus, straightforward screening approaches are required that enable rapid identification and evaluation of drug-metabolizing bacterial P450s with different product selectivities. Herein, we report a two-step screening method for discovery and characterization of new P450s from *actinomycetes* that enable oxidation of various drugs. In the first step, substrate profiling with three structurally different model drugs, ritonavir, testosterone, amitriptyline, allowed us to select CYP105D and CYP107Z from *Streptomyces platensis* DSM 40041 that accepted all model substrates and produced human-like drug metabolites. In the second step, activity tests with an array of 25 structurally-related molecules and derivatives of the three model compounds revealed a correlation between structural variations in the target drugs and the enzyme chemo- and regioselectivity.

Introduction

Low efficiency of some existing therapies, population ageing, and the occurrence of new diseases make the development of new medicines indispensable.[1] Assessment of the efficacy, safety and therapeutic effect of a new drug includes the elucidation of its metabolic fate.[2] Thus, production of drug metabolites for their structural characterization, toxicological evaluation and drug-drug interactions tests remains an important task in the drug development process. Traditionally, hepatocytes, liver microsomes and recombinant human drug-metabolizing enzymes have been used for this purpose.[2c] The major drug-metabolizing enzymes in the human body are heme-containing cytochrome P450 monooxygenases (CYP or P450), which have been reported to oxidize about two-third of all drugs.[3]

In the presence of molecular oxygen, P450s catalyze a variety of oxidation reactions on chemically diverse compounds including aliphatic and aromatic hydroxylation, C=C bond epoxidation, N-and S-oxidation, N-and S-dealkylation and many others.[4] For their activity P450s require electrons that are derived from the nicotine amide cofactors NADH or NADPH and are transferred via redox partner proteins to the heme iron. Generally, human drug-metabolizing P450s possess an extraordinary broad substrate spectrum and form distinct oxidation products, which are difficult to access via chemical synthesis.[5] However, application of membrane bound human P450s is somewhat limited by their low expression titers, activity and process stability.[6] For this reason, bacterial P450s with high activity and stability, easily expressible in recombinant hosts, have been subjected to rational protein design and/or directed evolution to achieve production of drug metabolites. As a result, engineered variants of the well-studied P450 BM3 from *Bacillus megaterium* (CYP102A1) were constructed which are not only able to oxidize various drugs but also form human drug metabolites.[2b,7] Nevertheless, the correlation between metabolite patterns produced by human and engineered bacterial P450s still remains a challenge for protein engineering, which often relies on extensive screening rounds.

Complementary to recombinant human P450 based platforms and engineered bacterial P450s, pharmaceutical companies use microbial biotransformation of drugs, especially when larger amounts of metabolites must be prepared.[8] Microbial strain collections used for this purpose usually include many strains from the phylum actinobacteria. Recent genome sequencing projects have revealed large collections of P450 genes in those strains.[9] P450s from *actinomycetes* have been found to often mimic activities of human drug-metabolizing P450s and thus are considered as worthwhile candidates for metabolite production.[10] For example, an array consisting of 250 recombinant P450s from *actinomycetes* has been established in *E. coli* and successfully applied for drug biotransformation in a 96-well format.[11] Many P450s from *actinomycetes* belong to the largest bacterial subfamilies CYP105 and CYP107, which are known to accept and oxidize a variety of chemically diverse scaffolds e.g. for the diversification of natural products or the degradation of xenobiotics.[9] The examples include among others CYP105D7 from *Streptomyces avermitilis*, which

[a] *T. Hilberath, L. M. Windeln, D. Decembrino, Dr. P. Le-Huu, F. L. Bilsing, Prof. V. B. Urlacher*
Institute of Biochemistry
Heinrich-Heine University Düsseldorf
Universitätsstrasse 1
Düsseldorf 40225 (Germany)
E-mail: Vlada.Urlacher@uni-duesseldorf.de

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cctc.201901967

ChemPubSoc Europe
CHEMCATCHEM Full Papers

catalyzes oxidation of the natural sesquiterpenoid pentalenic acid[12] and of the non-physiological substrate daidzein;[13] CYP107L from *Streptomyces platensis*, which combines activities against the drugs amodiaquine, ritonavir, amitriptyline and thioridazine;[14] as well as CYP107 (PldB) and CYP105 (PsmA) that perform selective hydroxylations during the pladieniolide D synthesis in recombinant *Streptomyces platensis*.[15]

Nevertheless, the identification of activities similar to human P450s cannot easily be deduced from the sequence or even structural similarity of P450s. Consequently, extensive screenings are required to identify a desired P450. Testing P450s towards one or few chemically related drug compounds like steroids[11a,16] or tricyclic antidepressants[17] reveals a correlation between the enzyme activity/selectivity and the particular drug compound's structure. This focused approach has a limitation as it does not necessarily distinguish between P450s with narrow and broad substrate spectra. A randomized screening with various drug compounds facilitates discovery of P450s with broad substrate spectra as for example described for CYP267 from *S. cellulosum* or P450 BM3 variants.[18] But in this case the effect of small changes in drug structure on the enzyme activity and selectivity cannot be deduced.

In this study we combined the advantages of both screening methods aiming at identification and characterization of drug-metabolizing P450 enzymes from *Streptomyces platensis* DSM 40041 and *Pseudonocardia autotrophica* DSM 535. Both strains are utilized by pharmaceutical companies for the oxidation of various drugs.[14,19] Using substrate profiling with three model drugs namely ritonavir, testosterone and amitriptyline in the first step, two P450s from *S. platensis* with broad substrate spectra and complementary selectivities were selected. In the second step, these two P450s classified as CYP105D and CYP107Z were subsequently evaluated against an array of structurally-related drugs and derivatives of the model compounds. Besides their extraordinary broad substrate promiscuity, the newly identified P450s produced human drug metabolites. Following this two-step screening approach, we recognized how small variations in structure of the target drugs influence chemo- and regioselectivity of the selected P450s. This correlation will help to reduce future screening efforts for the identification of a P450 to furnish a certain drug metabolite.

Results and Discussion

Identification of drug-metabolizing P450s

The recently published genomes of *Pseudonocardia autotrophica* DSM 535 and *Streptomyces platensis* DSM4004129[19] contain 70 *cyp* genes. 33 of them were found in *P. autotrophica* and 37 *cyp* genes in *S. platensis*. According to the online P450 database CYPED[20] 31 genes belong to the CYP105 and the CYP107 families. Fourteen of those genes (13 of *S. platensis* and 1 of *P. autotrophica*) have been investigated in previous studies.[14,21] One *cyp* gene in *P. autotrophica* has a length of only 378 bp and was not considered further. The remaining 16 so far uncharacterized *cyp* genes were cloned and expressed in *E. coli* C43 (DE3) (Table S4). Judging from the CO-difference spectra that is used as indicator for functional P450 enzymes, all P450s could be produced in soluble form at concentrations varying from 13 to 1036 $nmol_{P450}/g_{CDW}$ (corresponding to 0.6–48 mg_{P450}/g_{CDW}, Figures S1 and S2).

Thirteen P450s expressed in *E. coli* at concentrations of above 100 $nmol_{P450}/g_{CDW}$ were tested *in vitro* against three model drugs: the anti-HIV agent ritonavir (**1**), the steroid testosterone (**7**) and the tricyclic antidepressant amitriptyline (**16**) (Figure 1). These drugs of various size and shape belong to different chemical classes. P450 mediated oxidation of the model drugs was supported by heterologous redox partners, either flavodoxin YkuN from *B. subtilis* and flavodoxin reductase from *E. coli* (Fdr) or putidaredoxin (Pdx) and putidaredoxin reductase (Pdr) from *P. putida* (Figures S3 and S4). Oxidized products were identified by LC/MS analysis (Figures S5–S36; Tables S7–S31).

Ritonavir (**1**) is the drug with the highest molecular weight (720.9 g/mol). This peptide analogue contains many sites for

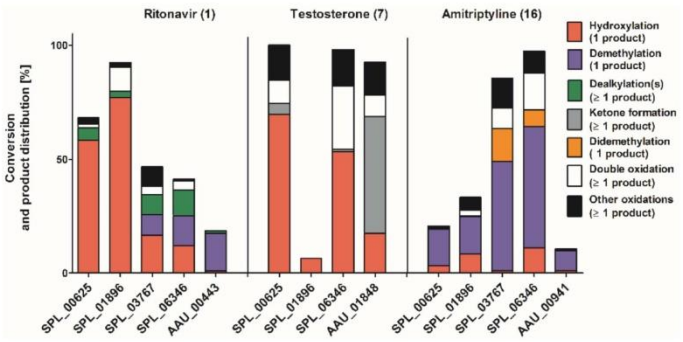
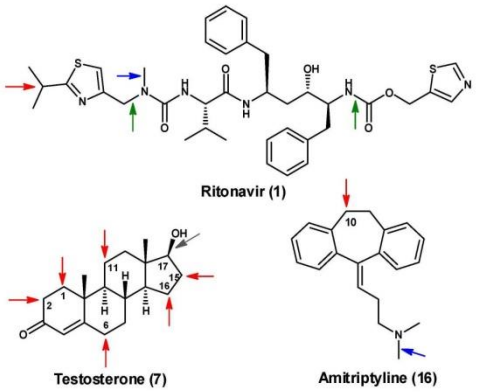


Figure 1. Conversions of the model drugs **1**, **7** and **16** and product patterns observed with active P450s. Only P450s showing activities against at least one drug are shown. The hydroxylation products of substrate **7** are different for different P450s.

ChemPubSoc Europe
CHEMCATCHEM Full Papers

oxidation and is metabolized by human P450s including CYP3A family members mainly via N-dealkylation leading either to the loss of the thiazolyl carbamate moiety or of the isopropylthiazolylmethyl moiety, N-demethylation, and hydroxylation of the isopropyl group of the isopropylthiazolyl moiety (Scheme 1).[22] The second model drug testosterone (**7**) has a lower molecular weight (288.4 g/mol). It can be hydroxylated by human P450s like CYP3A4 or CYP2D6 at different positions often in both, α-and β-configuration (Scheme 1), and is often used for evaluation of P450 activity and selectivity.[11a] The last model compound, the tricyclic antidepressant amitriptyline (**16**) (277.4 g/mol) is primarily oxidized by human CYP2C19 at the trialkyl moiety to the demethylated product nortriptyline (**21**) followed by the hydroxylation to 10-hydroxy-amitriptyline in (*E*)-configuration catalyzed mainly by human CYP2D6 and CYP3A4 (Scheme 1). Other reactions include didemethylation to desmethylnortriptyline or N- oxidation to amitriptyline N-oxide.[23]

The *in vitro* activity screening revealed seven P450 enzymes that demonstrated activity against at least one drug compound (Figure 1). Among them, four P450s from *S. platensis* (SPL_00625, SPL_01896, SPL_03767 and SPL_06346) accepted at least two model drugs as substrates. The active P450s exhibited different product selectivities. Ritonavir (**1**) was oxidized by SPL_00625 and SPL_01896 to the same main hydroxylated product with ≥80% regioselectivity. SPL_03767 and SPL_06346 converted **1** with lower activities to form product mixtures with higher amounts of dealkylated and demethylated metabolites. Also during oxidation of testosterone (**7**) SPL_00625 formed one main hydroxylation product with a regioselectivity of 70%, while SPL_06346 was less selective and formed a product mixture (Figure S16). Both P450s were quite active and enabled almost full conversion of **7**. With amitriptyline (**16**), SPL_00625 showed again the highest regioselectivity of 80% towards nortriptyline (**21**) followed by SPL_03767, which formed besides



Scheme 1. Chemical structures of the model drugs. Major oxidation sites of human P450s are highlighted with an arrow. The color code displays either hydroxylation (red), demethylation (blue), dealkylation (green) or alcohol oxidation (grey).

the monodemethylated nortriptyline (56%) also the didemethylated product (17%). Since SPL_00625 and SPL_06346 catalyzed oxidation of all model substrates with different regioselectivities, their activity and selectivity were further evaluated in the second step using a focused substrate library. According to the CYP nomenclature by Prof. David Nelson, these P450s were classified as CYP105D (SPL_00625) and CYP107Z (SPL_06346) (Table S5).[24]

Library design for systematic substrate screening

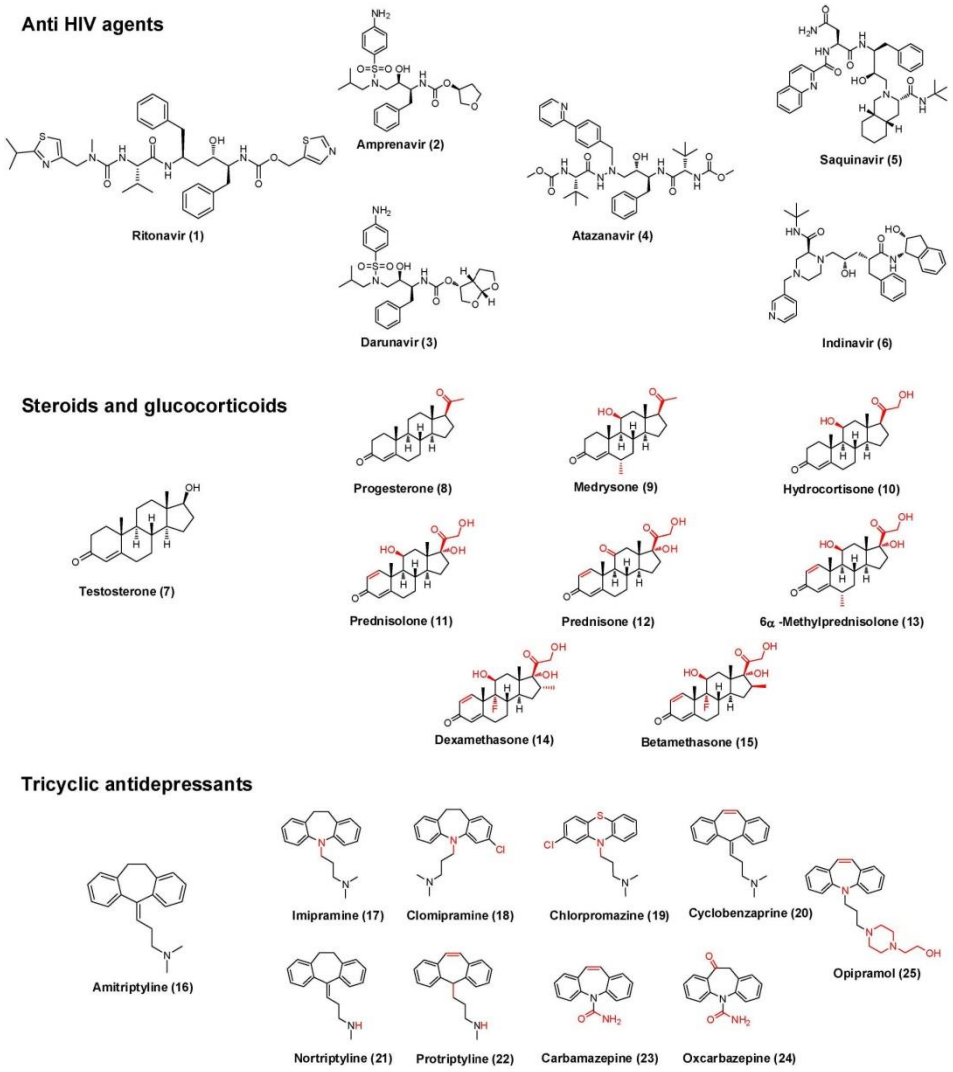
In order to elucidate how variations in structure and size of the target drug compounds influence enzymes' activity and selectivity, we followed a systematic screening with either structurally related compounds for the target drug (anti-HIV compounds **1**–**6**) or derivatives with differences in functional groups (steroids **7**–**15** and tricyclic antidepressants **16**–**25**). Overall 25 compounds were tested (Scheme 2).

Screening results for anti-HIV compounds

The tested anti-HIV agents **1**–**6** are bulky, have a high molecular weight (505.6–720.9 g/mol) and contain a hydroxyethylene group which is responsible for the inhibition of HIV protease.[25] CYP105D and CYP107Z accept all tested compounds **1**–**6** (Figure 2A). The highest conversion with both P450s was observed for saquinavir **5** (71% with CYP105D and 97% with CYP107Z). CYP105D demonstrated its lowest activity with darunavir **3** (7% conversion), while the lowest activity of CYP107Z was measured with atazanavir **4** (21% conversion). In previous studies CYP267B1 and P450 BM3 variants were reported that enabled *in vitro* oxidation of anti-HIV compounds including ritonavir **1** and saquinavir **5**. For ritonavir **1** very low conversions of 1.4% were observed with CYP267B1, while saquinavir **5** was most efficiently oxidized by the P450 BM3 variant MT35 with 35% conversion.[18a,26] Thus, the activities of CYP105D and CYP107Z are equal or at least 2–3 times higher than those of previously reported bacterial P450s. Besides, they accept a broader spectrum of anti-HIV agents.

Analysis of the product patterns revealed that CYP105D generally seemed more selective but less active than CYP107Z (Figures S5–S13). One major hydroxylated product (50–85% of the product ratio) was detected after turnover of the anti-HIV drugs (except for **2**) catalyzed by CYP105D. In contrast, CYP107Z produced with the corresponding drugs not only hydroxylated but also dealkylated metabolites at higher ratios (up to 69% of the product ratio). For example, ritonavir metabolites **1b** (Δm: −14 Da), **1c** (Δm: −139 Da) and **1d** (Δm: −141 Da) had characteristically smaller masses compared to the substrate.

Recently, our group described nine metabolites of ritonavir biotransformed by the parent organism of both P450s *Streptomyces platensis* DSM 40041. Metabolites were analyzed by LC/MS and LC/MS-MS analysis.[14] The observed mass differences and retention times for **1b**–**d** correspond to the N-dealkylated



Scheme 2. Substrate library based on derivatives and functionally related compounds of the model drugs ritonavir (**1**), testosterone (**7**) and amitriptyline (**16**). Differences in structures of the model drugs are highlighted in red.

and N-demethylated metabolites described there (Table S7). CYP107Z was able to furnish N-demethylated (13% of **1b**) and N-dealkylated (6% of **1c** and 5% of **1d**) metabolites more efficiently compared to CYP105D (Table 1). However, a single hydroxylation product was best accessed with CYP105D (58%). This product was identified as hydroxy ritonavir (**1a**) as deduced from a spiking experiment with an authentic standard (Figure S6).

In conversions with amprenavir (**2**) and darunavir (**3**) CYP107Z formed one metabolite (36% and 44%, respectively) with the same *m/z* of 339 (Table S8 and S9). By co-elution with an authentic standard, this metabolite was identified as 4-amino-N-((2*R*,3*S*)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutyl-benzenesulfonamide (**2a**) (Figure S8 and S10). This metabolite is formed through N-dealkylation.

ChemPubSoc Europe
CHEMCATCHEM Full Papers

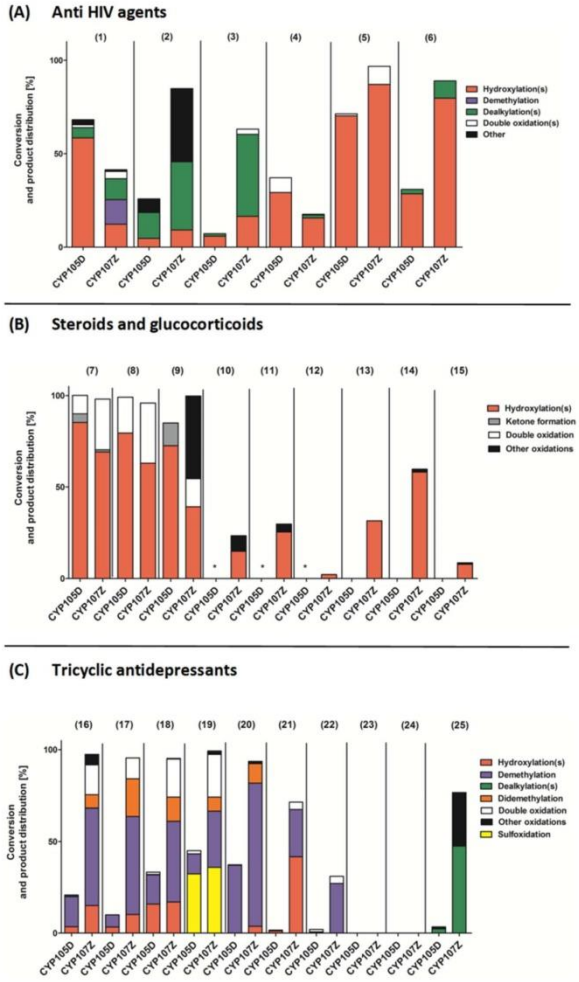


Figure 2. Overview of the focused substrate screening with CYP105D and CYP107Z. (A) Activity and product selectivity of P450s in reactions with anti-HIV drugs. (B) Activity and product selectivity of P450s in reactions with steroids. *represents traces of products, which could not be quantified. (C) Activity and product selectivity of P450s in reactions with tricyclic antidepressants. Under 'other oxidations' all products are summarized, which could not be assigned to a certain P450 reaction.

All identified metabolites are formed as major metabolites by human P450s including CYP3A isoforms and CYP2D6.[22,27]

Screening results for steroids and glucocorticoids

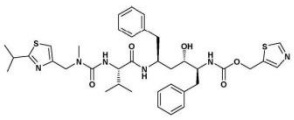
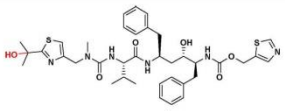
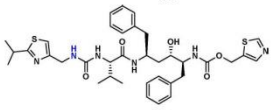
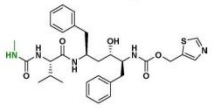
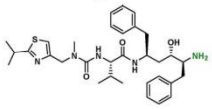
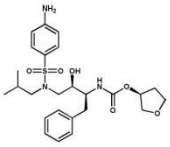
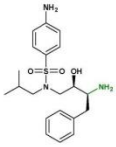
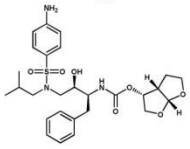
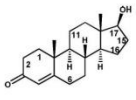
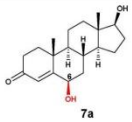
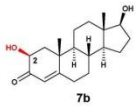
The tested steroids **7**–**15** possess a tetracyclic core structure. They differ in the number and position of hydroxy-, keto-, and methyl groups and double bonds. Synthetic glucocorticoids like **9**, **11**–**15** have more substituting groups than natural compounds like **7**, **8** or **10** which often leads to increased drug efficiency.[28] In human drug metabolism, CYP3A4 is one of the most frequently investigated P450s to form steroid metabolites.[29]

Also in case of steroids, the substrate scope of CYP107Z is generally broader compared to that of CYP105D (Figure 2B, Figures S14–S25). CYP105D converted **7**, **8** and **9** to one major product (≥66% of the product ratio) each with high activity (≥85% conversion), but barely or not accepted **10**–**15** as substrates. CYP107Z oxidized **7**, **8** and **9** nearly completely

ChemPubSoc Europe

CHEMCATCHEM Full Papers

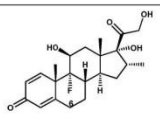
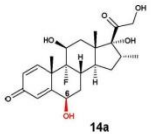
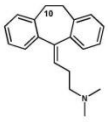
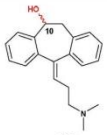
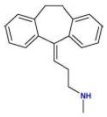
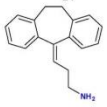
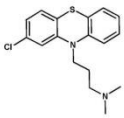
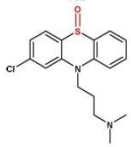
Table 1. Overview of identified metabolite structures in this study. The product percentage was calculated based on the observed peak areas related to the sum of product peaks areas. The color code highlights the type of P450 reaction depicted in the figures before.

Substrate	Reaction type	Identified metabolite structure	Product percentage [%] CYP105D	CYP107Z
Anti-HIV agents				
 Ritonavir (1)	hydroxylation **1a**: Δm: +16 Da	 **1a**	86	30
	demethylation **1b**: Δm: −14 Da	 **1b**	0	32
	dealkylation **1c**: Δm: −139 Da	 **1c**	8	14
	dealkylation **1d**: Δm: −141 Da	 **1d**	0	13
 Amprenavir (2)	dealkylation **2a**: Δm: −114 Da	 **2a**	53	43
 Darunavir (3)	dealkylation **2a**: Δm: −156 Da		18	69
Steroids and glucocorticoids				
 Testosterone (7)	hydroxylation **7a**: Δm: +16 Da	 **7a**	0	6
	hydroxylation **7b**: Δm: +16 Da	 **7b**	70	5

ChemPubSoc Europe

CHEMCATCHEM Full Papers

Table 1. continued

Substrate	Reaction type	Identified metabolite structure	Product percentage [%] CYP105D	CYP107Z
 Dexamethasone (14)	hydroxylation **14a**: Δm: +16 Da	 **14a**	0	97
Tricyclic antidepressants				
 Amitriptyline (16)	hydroxylation **16a**: Δm: +16 Da	 **16a**	15	12
	demethylation **21**: Δm: −14 Da	 **21**	79	55
	demethylation **16b**: Δm: −28 Da	 **16b**	≤1	7
 Chlorpromazine (19)	sulfoxidation **19b**: Δm: +16 Da	 **19b**	72	36

(≥96% conversion), but was less selective than CYP105D. The corresponding main product accounted for 38–54% of the total products (Figure S16–S18). In contrast to CYP105D, the substrates **10**–**13** were clearly oxidized by CYP107Z, though with low activity (2–32% conversion). Only one main product of each compound could be detected (Figure S19–S22). Remarkable is the difference in activity of CYP107Z towards dexamethasone (**14**) with a 16α-methyl group at C16 and its epimer betamethasone (**15**) carrying the methyl group at C16 in β-position. Dexamethasone (**14**) was converted to 60% with 97% selectivity to one major product. The activity against betamethasone was more than six times lower (9% conversion).

NMR analysis revealed that testosterone (**7**) was oxidized by CYP105D mainly to 2β-hydroxytestosterone (**7b**) (Table 1, Figure S37), which was confirmed later by spiking with the authentic reference (Figure S14). This product was also formed by CYP107Z, but only in traces (Figure S15). The major human metabolite of **14** is 6β-hydroxydexamethasone (**14a**) (Table 1), which is also formed by CYP107Z as it was deduced from spiking experiments (Figure S23).[29b] The preference for the 6β-position (97%) of **14** with a side chain increased significantly compared to the initial model compound **7** (only 6% 6β-hydroxytestosterone **7a**).

Other bacterial P450s including P450 BM3 variants, CYP106A1, and CYP106A2 were also tested for testosterone and dexamethasone oxidation. Dexamethasone and prednisolone were converted by CYP106A1 and CYP106A2 with different selectivities. While CYP106A2 was selective for the 15β-position, CYP106A1 oxidized these compounds at position 6β, 15β and 11.[30] Conversions between 33% and 40% were achieved with CYP106A1. Compared to the reported CYP106A1, CYP107Z oxidized dexamethasone with higher activity and selectivity for the position 6β. Considering the different selectivity of CYP106A2, the activity and selectivity of CYP107Z against dexamethasone **14** complement and expand the possibilities for drug metabolite syntheses starting with other glucocorticoids.

P450 BM3 wild type does not accept testosterone, but variants thereof carrying up to 20 mutations were reported to perform stereo- and regioselective oxidations of steroids like testosterone at position 2β, 15β and 16β.[7d,30–31]

ChemPubSoc Europe

CHEMCATCHEM Full Papers

Although significant advances have been achieved to reduce screening efforts among P450 BM3 variants,[7d] the herein described P450s might represent valuable alternatives and new targets for protein engineering as they naturally convert steroids with quite high activity and moderate selectivity.

Screening results for tricyclic antidepressants

Tricyclic antidepressants including **16**–**25** can be grouped either according to the ring arrangement or to the respective side chain.[32] Several human P450s including CYP3A4 and CYP2D6 oxidize these compounds mainly via N-demethylation and hydroxylation as summarized elsewhere.[23a]

CYP105D and CYP107Z accepted nearly all tested tricyclic antidepressants (Figure 2C). Only carbamazepine (**23**) and oxcarbazepine (**24**), both missing a long alkyl chain, were not accepted by any P450. Also during oxidation of tricyclic antidepressants CYP105D was less active than CYP107Z. Generally, trialkylated amines were better substrates than dialkylated amines. CYP107Z oxidized the trialkylated amines **16**–**20** nearly completely (≥95% conversion), while CYP105D showed 2 to 10 times lower activity (10–45% conversion) (Figures S26–S36). Dialkylated antidepressants **21** and **22** were moderately oxidized by CYP107Z (30–70% conversion), while activity of CYP105D against these substrates was almost not present (max. 2% conversion). Both, CYP105D and CYP107Z, were able to produce demethylated metabolites of **16**, **17**, **18** and **20**, which have a characteristic mass difference of −14 *m/z* compared to the substrates (Table S22–S28). CYP107Z additionally produced didemethylated products. Starting from **16**, 53% nortriptyline (**21**) and 7% desmethylnortriptyline (**16b**) were produced as it was confirmed by co-elution with authentic standards (Figure S26, Table 1).

Chlorpromazine (**19**) is an example of a diarylthioether. Besides demethylation, **19** was converted by both enzymes to another product having a mass difference of +16 *m/z* compared to the substrate (Table S25). By co-elution with an authentic standard, this product was identified as chlorpromazine sulfoxide (**19b**) (Figure S30).

Also with **16**, **17**, **18** and **21** a single oxidation product was formed with both enzymes showing a $\Delta m = +16$ Da. This product was absent after conversions of **20** and **22**. In comparison to the other tested antidepressants, cyclobenzaprine (**20**) and protriptyline (**22**) carry an additional double bond in the central ring between C10 and C11, which has been reported to hinder the oxidation there.[17] This might indicate that the hydroxylation of **16**, **17**, **18** and **21** occurs at C10, which was indeed confirmed for amitriptyline (**16**) conversion via co-elution with an authentic standard (Figure S26). The hydroxy group at C10 of (±)-(*E*)-10-Hydroxyamitriptyline (**16a**) is (*E*)-configured and matches the preferred position and configuration achieved with human P450s.[23b] The ability to perform hydroxylation, sulfoxidation, demethylation, dealkylation and didemethylation on tricyclic antidepressants has been reported for other bacterial P450s including P450 BM3 variants or CYP267B1 and CYP267A1 from *Sorangium cellulosoum* So ce56.[7f,17–18,26] The selectivities and activities to form human drug metabolites of the described enzymes there are similar to CYP105D and CYP107Z. However, the screening developed in this study helps to differentiate P450 enzymes in terms of activity, chemo- and regioselectivity, thus facilitating the preselection of an appropriate enzyme candidate for other drug compounds with different modifications of the lead structures.

Conclusions

In summary, a two-step screening approach was applied to identify P450 candidates that on the one hand accept a broad range of drugs of various size and chemical structure and on the other hand demonstrate different selectivities leading to various metabolites. The identified CYP105D and CYP107Z both from *S. platensis* DSM 40041 are promiscuous bacterial cytochrome P450 monooxygenases showing complementary activities and selectivities on chemically diverse drugs. Whereas CYP105D generally possesses a narrower substrate spectrum and lower activity than CYP107Z, it shows a higher regioselectivity and allows to produce one main hydroxylated metabolite of many drugs. For instance 2β-hydroxytestosterone was formed in the reaction with CYP105D with 70% regioselectivity after complete conversion, while hydroxy ritonavir was formed with 86% regioselectivity after a conversion of 70%. CYP107Z demonstrates a higher preference for N-dealkylation than CYP105D and thus provide an access to demethylated drug metabolites. Additionally, CYP107Z catalyzed steroid hydroxylation with a rather high regioselectivity of up to 97%, for instance for 6β-hydroxydexamethasone. The design of a focused substrate library pointed out not only the differences between both enzymes but also indicated a remarkable overlap with the substrate and product spectra of human hepatic P450s. Importantly, the screening of closely related compounds allowed us to evaluate how small variations in chemical structures of target drugs can change the respective enzyme activity and selectivity. This knowledge might reduce future screening efforts for drugs from the same chemical groups, which have not been tested with these P450s yet.

Experimental section

Bacterial strains, enzymes and chemicals: *E. coli* strains DH5α and BL21 (DE3) were purchased from Clontech, OverExpress C43(DE3) from Lucigen. Phusion High-Fidelity DNA-polymerase, restriction endonucleases (Fast Digest variants), FastAP thermosensitive alkaline phosphatase and T4 DNA-ligase were obtained from Thermo Scientific. Catalase from bovine liver was obtained from Sigma Aldrich. Unless indicated otherwise, all chemicals were of analytical grade or higher and purchased from Sigma Aldrich, VWR, AppliChem, Carl Roth, BD BioSciences and Grüssing. Suppliers of all tested drugs and respective reference metabolites are listed in Table S1.

Construction of expression vectors: Genes encoding P450s from *S. platensis* were amplified from genomic DNA and integrated with conventional cloning methods in expression vectors pET22b,

ChemPubSoc Europe

CHEMCATCHEM Full Papers

pET24b or pET28a using the primers listed in Table S2. The *cyps* from *P. autotrophica* were cloned in pET24b after amplification from genomic DNA using primers for PCR allowing vector integration via Gibson Assembly.[33] All existing GTG start codons were replaced by the start codon ATG in the primer sequences. The sequence for the N-terminal His_6-tag of *cyp105D* was inserted after the start codon following the instructions of the Q5 site-directed mutagenesis kit (New England Biolabs). The construction of pCOLADuet-PP was performed by integrating *camA* and *camB* in MSCI of pCOLADuet-1 via Gibson Assembly.[33] DNA Sequences were confirmed by sequencing (GATC Biotech, Germany).

Gene expression and enzyme preparation: All genes are integrated in vectors allowing expression using the T7-expression system. P450 genes were expressed in recombinant *E. coli* C43 (DE3), while redox partner proteins and GDHIV were produced in *E. coli* BL21 (DE3). Detailed information on expression and purification is provided in the Supporting Information (SI).

Enzyme assays: The P450 concentrations in crude cell extracts or purified fractions were calculated from CO difference spectra using the extinction coefficient $\varepsilon_{450} = 91\ mM^{-1}cm^{-1}$ as published elsewhere.[34] The concentration of tested redox partner proteins was estimated spectrophotometrically. Concentration of heterologous redox partners were determined as described in the SI.

Substrate screening: During initial screening, reactions were performed in 200 µL 50 mM KP_i, pH 7.5. The tested P450s were typically applied in form of crude cell extracts, while the tested redox partner proteins were purified before. Reaction mixtures contained 2 µM P450, 2 µM Fdr or Pdr, 20 µM YkuN or Pdx, 0.2 mM $NAD(P)^+$ and 200 µM substrate dissolved in either in 2% (v/v) DMSO (anti-HIV drugs, oxcarbazepine, opipramol and carbamazepine) or 2% (v/v) ethanol (steroids) or deionized water (tricyclic antidepressants). For NAD(P)H cofactor regeneration, 5 U/mL GDH with 20 mM glucose was used, while 600 U/mL catalase from bovine (Sigma Aldrich) was added to eliminate possibly formed hydrogen peroxide due to uncoupling. The samples were incubated at 25 °C in 2 mL reaction tubes with open lids in a thermomixer (Eppendorf) with a shaking speed of 300 min^{-1}. After 4 h or 20 h incubation, 4 µL of internal standard (10 mM stock solution in methanol, DMSO or water) was added. For conversions with tricyclic antidepressants and Anti-HIV agents, 300 µL 0.1 M sodium carbonate buffer, pH 10 was supplemented to alkalize the solution for better extraction. Then an appropriate volume of ethyl acetate was added, the samples vigorously shaken for 5 min and centrifuged for 5 min at 12300 g. The organic phase was transferred into a new reaction tube and evaporated under reduced pressure. The dried samples were resuspended in methanol (Anti-HIV drugs, steroids) or water with 20% acetonitrile and 0.1% formic acid (tricyclic antidepressants) and used for LC/MS-analysis. The given conversion values are based on substrate depletion with an internal standard referred to a control reaction with crude cell extract of *E. coli* pET24b harboring no P450. If conversions were below 5%, the sum of product and substrate peaks was used to calculate conversions. Product distribution was calculated from the sum of peak areas in the MS or the absorption at 254 nm. Due to the complex product pattern in some cases, single ion monitoring analysis was used additionally during MS analysis to distinguish the product peaks. All measurements represent mean values from technical duplicates.

Product analysis: Substrate biotransformations were analyzed by liquid chromatography coupled to mass spectrometry (LC/MS) on a Prominence/LCMS 2020 device (Shimadzu). A Chromolith® Performance RP-18e column (100×4.6 mm, Merck) was used for Anti-HIV agents and steroids whereas a Chromolith® Performance RP-8e column (100×4.6 mm, Merck) was applied for tricyclic antidepressants. Solvent A was always ddH_2O with 0.1% formic acid, while solvent B was either methanol (MeOH) or acetonitrile (ACN). 1 µL of each sample was injected and analytes were separated with a flow rate of 0.5–1 mL/min at 30 °C. The separated compounds were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI) in a dual ionization mode. Mass compounds were detected in positive scan mode in a range between 150–1000 m/z (Anti-HIV agents), 200–950 m/z (tricyclic antidepressants) and 100–500 m/z (steroids). A detailed overview about the chromatographic conditions is given in Table S6. Products were identified based on retention times, MS spectra and by spiking of authentic reference compounds to reaction solutions. NMR-analysis (1D and 2D-spectra) was performed on a Bruker Avance III-600 spectrometer (^{1}H-NMR: 600 mHz; ^{13}C-NMR: 150 mHz). Chemical shifts (δ) are given in ppm and were referred to the solvent $CDCl_3$ (^{13}C: 77.2 ppm; ^{1}H: 7.26 ppm). Coupling constants (J) are stated in Hz. The NMR data were in accordance with literature data.[35]

***2β-hydroxytestosterone* (7b)**: ^{1}H NMR (600 MHz, $CDCl_3$): δ [ppm] = 0.79 (s, 3H, 18-H), 1.05–0.95 (m, 2H, 7-H_α, 14-H), 1.13 (td, J = 12.8 Hz, 4.1 Hz, 1H, 12-H_α), 1.18 (s, 3H, 19-H), 1.31 (td,, J = 12.2 Hz, 5.8 Hz, 1H, 15 H_β), 1.39 (ddd, J = 12.4, 10.5, 3.9 Hz, 1H, 9-H), 1.48–1.42 (m, 1H, 16 H_β), 1.56–1.48 (m, 2H, 11 H_β, 1 H_β), 1.61–1.56 (m, 1H, 15-H_α), 1.70 (dtd, J = 12.0, 10.5, 4.1 Hz, 1H, 8-H), 1.83–1.75 (m, 1H, 11-H_α), 1.88 (ddd, J = 12.8, 4.0, 2.8 Hz, 1H, 12-H_β), 2.00–1.95 (m, 1H, 7 H_β), 2.12–2.03 (m, 1H, 16-H_α), 2.25 (ddd, J = 12.3, 4.3, 2.6 Hz, 1H,6-H_α), 2.47 (dd, J = 13.7, 5.5 Hz, 1H,1-H_α), 2.53 (dddd, J = 13.6, 12.3, 5.1, 1.4 Hz, 1H,6-H_β), 3.65 (t, J = 8.6 Hz, 1H, 17-H), 4.18 (dd, J = 14.0, 5.5 Hz, 1H, 2-H_α), 5.80 (d, J = 1.4 Hz, 1H, 4-H). ^{13}C NMR (151 MHz, $CDCl_3$): δ [ppm] = 11.4 (C-18), 22.7 (C-19), 23.0 (C-11), 23.5 (C-15), 30.6 (C-16), 33.2 (C-6); 34.7 (C-7), 36.1 (C-8),36.6 (C-12), 39.7 (C-1), 41.6 (C 10), 43.6 (C-13), 50.4 (C-9), 50.7 (C-14), 68.7 (C-2), 81.6 (C-17), 118.9 (C-4), 175.3 (C-5), 199.9 (C-3)

Acknowledgements

This work was performed under the umbrella of the ERA-IB23rd call project "HyPerIn" [project number EIB.12.026]. Financial support was kindly provided by the Federal Ministry of Education and Research [grant number 031 A223 A]. The authors thank all "HyPerIn"-consortium members for fruitful discussions, Jan Gebauer and Judith Wamprecht for their help in expression studies and Sebastian Hölzel for technical support.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: biocatalysis · cytochrome P450 · substrate screening · human drug metabolites · regioselectivity

[1] a) Á. Lublóy, *BMC Health Serv. Res.* **2014**, *14*, 469; b) J. Drews, *Science* **2000**, *287*, 1960–1964.
[2] a) S. S. Singh, *Curr. Drug Metab.* **2006**, *7*, 165–182; b) G. Di Nardo, G. Gilardi, *Int. J. Mol. Sci.* **2012**, *13*, 15901–15924; c) K. Schroer, M. Kittelmann, S. Lütz, *Biotechnol. Bioeng.* **2010**, *106*, 699–706.
[3] S. Rendic, F. P. Guengerich, *Chem. Res. Toxicol.* **2015**, *28*, 38–42.
[4] M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, *Chem. Rev.* **1996**, *96*, 2841–2888.

[5] a) E. M. Gillam, *Clin. Exp. Pharmacol. Physiol.* **2005**, *32*, 147–152; b) M. Winkler, M. Geier, S. P. Hanlon, B. Nidetzky, A. Glieder, *Angew. Chem. Int. Ed.* **2018**, *57*, 13406–13423.
[6] J. B. Behrendorff, E. M. Gillam, *Chem. Res. Toxicol.* **2016**, *30*, 453–468.
[7] a) M. Landwehr, L. Hochrein, C. R. Otey, A. Kasrayan, J. E. Backvall, F. H. Arnold, *J. Am. Chem. Soc.* **2006**, *128*, 6058–6059; b) A. M. Sawayama, M. M. Chen, P. Kulanthaivel, M. S. Kuo, H. Hemmerle, F. H. Arnold, *Chem. Eur. J.* **2009**, *15*, 11723–11729; c) H. Venkataraman, M. C. Verkade-Vreeker, L. Capoferri, D. P. Geerke, N. P. Vermeulen, J. N. Commandeur, *Bioorg. Med. Chem.* **2014**, *22*, 5613–5620; d) C. G. Acevedo-Rocha, C. G. Gamble, R. Lonsdale, A. Li, N. Nett, S. Hoebenreich, J. B. Lingnau, C. Wirtz, C. Fares, H. Hinrichs, *ACS Catal.* **2018**, *8*, 3395–3410; e) C. F. Butler, C. Peet, K. J. McLean, M. T. Baynham, R. T. Blankley, K. Fisher, S. E. Rigby, D. Leys, M. W. Voice, A. W. Munro, *Biochem. J.* **2014**, *460*, 247–259; f) X. Ren, J. A. Yorke, E. Taylor, T. Zhang, W. Zhou, L. L. Wong, *Chem. Eur. J.* **2015**, *21*, 15039–15047.
[8] S. P. Hanlon, T. Friedberg, C. R. Wolf, O. Ghisalba, M. Kittelmann, in *Modern biooxidation* (Eds.: R. D. Schmid, V. B. Urlacher), Wiley-VCH, Weinheim, **2007**, pp. 233–252.
[9] J. D. Rudolf, C.-Y. Chang, M. Ma, B. Shen, *Nat. Prod. Rep.* **2017**, *34*, 1141–1172.
[10] D. C. Lamb, M. R. Waterman, B. Zhao, *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 1279–1294.
[11] a) H. Agematu, N. Matsumoto, Y. Fujii, H. Kabumoto, S. Doi, K. Machida, J. Ishikawa, A. Arisawa, *Biosci. Biotechnol. Biochem.* **2006**, *70*, 307–311; b) A. Arisawa, H. Agematu, in *Modern biooxidation* (Eds.: R. D. Schmid, V. B. Urlacher), Wiley-VCH, Weinheim, **2007**, pp. 177–192.
[12] S. Takamatsu, L.-H. Xu, S. Fushinobu, H. Shoun, M. Komatsu, D. E. Cane, H. Ikeda, *J. Antibiot. (Tokyo)* **2011**, *64*, 65–71.
[13] B. P. Pandey, C. Roh, K. Y. Choi, N. Lee, E. J. Kim, S. Ko, T. Kim, H. Yun, B. G. Kim, *Biotechnol. Bioeng.* **2010**, *105*, 697–704.
[14] A. Worsch, F. K. Eggimann, M. Girhard, C. J. von Bühler, F. Tieves, R. Czaja, A. Vogel, C. Grumaz, K. Sohn, S. Lütz, M. Kittelmann, V. B. Urlacher, *Biotechnol. Bioeng.* **2018**, *115*, 2156–2166.
[15] K. Machida, Y. Aritoku, T. Nakashima, A. Arisawa, T. Tsuchida, *J. Biosci. Bioeng.* **2008**, *105*, 649–654.
[16] B. Dangi, K. H. Kim, S. H. Kang, T. J. Oh, *ChemBioChem* **2018**, *19*, 1066–1077.
[17] M. Litzenburger, F. Kern, Y. Khatri, R. Bernhardt, *Drug Metab. Dispos.* **2015**, *43*, 392–399.
[18] a) F. Kern, Y. Khatri, M. Litzenburger, R. Bernhardt, *Drug Metab. Dispos.* **2016**, *44*, 495–504; b) N. Beyer, J. K. Kulig, M. W. Fraaije, M. A. Hayes, D. B. Janssen, *ChemBioChem* **2018**, *19*, 326–337.
[19] C. Grumaz, Y. Vainshtein, P. Kirstahler, S. Luetz, M. Kittelmann, K. Schroer, F. K. Eggimann, R. Czaja, A. Vogel, T. Hilberath, A. Worsch, M. Girhard, V. B. Urlacher, M. Sandberg, K. Sohn, *Genome Announc.* **2017**, *5*, e00532–00517.
[20] CYPED can be found under https://cyped.biocatnet.de/.
[21] C.-H. Chen, J.-C. Cheng, Y.-C. Cho, W.-H. Hsu, *Biochem. Biophys. Res. Commun.* **2005**, *329*, 863–868.
[22] a) G. N. Kumar, A. D. Rodrigues, A. M. Buko, J. F. Denissen, *J. Pharmacol. Exp. Ther.* **1996**, *277*, 423–431; b) J. F. Denissen, B. A. Grabowski, M. K. Johnson, A. M. Buko, D. J. Kempf, S. B. Thomas, B. W. Surber, *Drug Metab. Dispos.* **1997**, *25*, 489–489.
[23] a) A. A. Somogyi, J. K. Coller, in *Metabolism of drugs and other xenobiotics* (Eds.: P. Anzenbacher, U. M. Zanger), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, **2012**, pp. 403–428; b) O. V. Olesen, K. Linnet, *Pharmacology* **1997**, *55*, 235–243.
[24] D. R. Nelson, *Hum. Genomics* **2009**, *4*, 59–65.
[25] C. H. Andrade, L. M. d. Freitas, V. d. Oliveira, *Braz. J. Pharm. Sci.* **2011**, *47*, 209–230.
[26] J. Reinen, J. S. van Leeuwen, Y. Li, L. Sun, P. D. Grootenhuis, C. J. Decker, J. Saunders, N. P. Vermeulen, J. N. Commandeur, *Drug Metab. Dispos.* **2011**, *39*, 1568–1576.
[27] a) M. Vermeir, S. Lachau-Durand, G. Mannens, F. Cuyckens, B. van Hoof, A. Raoof, *Drug Metab. Dispos.* **2009**, *37*, 809–820; b) A. Hsu, G. R. Granneman, R. J. Bertz, *Clin. Pharmacokinet.* **1998**, *35*, 275–291.
[28] E. de Kloet, N. O. Zacarias, O. Meijer, in *Stress: neuroendocrinology and neurobiology: Handbook of stress series* (Ed.: G. Fink), Academic Press/Elsevier Inc., **2017**, pp. 367–383.
[29] a) H. Yamazaki, T. Shimada, *Arch. Biochem. Biophys.* **1997**, *346*, 161–169; b) D. M. Gentile, E. S. Tomlinson, J. L. Maggs, B. K. Park, D. J. Back, *J. Pharmacol. Exp. Ther.* **1996**, *277*, 105–112; c) C. Ged, J. Rouillon, L. Pichard, J. Combalbert, N. Bressot, P. Bories, H. Michel, P. Beaune, P. Maurel, *Br. J. Clin. Pharmacol.* **1989**, *28*, 373–387.
[30] a) N. Putkaradze, F. M. Kiss, D. Schmitz, J. Zapp, M. C. Hutter, R. Bernhardt, *J. Biotechnol.* **2017**, *242*, 101–110; b) F. M. Kiss, Y. Khatri, J. Zapp, R. Bernhardt, *FEBS Lett.* **2015**, *589*, 2320–2326.
[31] a) S. Kille, F. E. Zilly, J. P. Acevedo, M. T. Reetz, *Nat. Chem.* **2011**, *3*, 738–743; b) D. Schmitz, S. Janocha, F. M. Kiss, R. Bernhardt, *Biochim. Biophys. Acta Proteins Proteomics* **2018**, *1866*, 11–22.
[32] M. G. Casarotto, D. J. Craik, *J. Pharm. Sci.* **2001**, *90*, 713–721.
[33] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, *Nat. Methods* **2009**, *6*, 343–345.
[34] T. Omura, R. Sato, *J. Biol. Chem.* **1964**, *239*, 2370–2378.
[35] a) D. N. Kirk, H. C. Toms, C. Douglas, K. A. White, K. E. Smith, S. Latif, R. W. Hubbard, *J. Chem. Soc., Perkin Trans. 2* **1990**, 1567–1594; b) N. E. Jacobsen, K. E. Kövér, M. B. Murataliev, R. Feyereisen, F. A. Walker, *Magn. Reson. Chem.* **2006**, *44*, 467–474.

Manuscript received: October 17, 2019
Revised manuscript received: December 17, 2019
Accepted manuscript online: December 17, 2019
Version of record online: January 28, 2020

2.2.1 Supporting information

Supplementary Material and methods

Chemicals

Table S2:Drug substrates, reference substances and respective providers.

Name	Manufacturer
substrates	
6α-Methylprednisolone; ≥ 98%	TCI
Amitriptyline HCl; ≥ 98 %	Sigma Aldrich
Amprenavir; ≥ 98%	Sigma Aldrich
Atazanavir; ≥ 98%	Sigma Aldrich
Betamethasone; ≥ 98%	TCI
Carbamazepine; ≥ 99.0%	Sigma Aldrich
Chlorpromazine HCl; ≥ 99 %	TCI
Clomipramine HCl; ≥ 98%	Sigma Aldrich
Cyclobenzaprine HCl; ≥ 98 %	TCI
Darunavir; ≥ 98%	Sigma Aldrich
Dexamethasone; ≥ 98%	Sigma Aldrich
Hydrocortisone; 98%	Alfa Cesar
Imipramine HCl; ≥99%	Sigma Aldrich
Indinavir sulfate salt hydrate; ≥ 98%	Sigma Aldrich
Medrysone; ≥ 95%	Cayman Chemical
Nortriptyline HCl; ≥98%	Sigma Aldrich
Opipramol; ≥99.0%	Fluka
Oxcarbazepine; ≥ 98 %	TCI
Prednisolone; ≥ 99%	Sigma Aldrich
Prednisone; ≥ 98%	Sigma Aldrich
Progesterone; ≥ 99 %	AppliChem
Protriptyline HCl; ≥99%	Sigma Aldrich
Ritonavir; ≥ 98 %	TCI
Saquinavir mesylate; ≥ 98%	Sigma Aldrich
Testosterone; >98 %	Sigma Aldrich

Name	Manufacturer
Reference metabolites	
16α-Hydroxytestosterone	Steraloids Inc.
16β-Hydroxytestosterone	Steraloids Inc.
2α-Hydroxytestosterone	Steraloids Inc.
2β-Hydroxytestosterone	Steraloids Inc.
4-Androstene-3,17-dione	Sigma-Aldrich
6β-Hydroxydexamethasone	Toronto Research Chemicals
6β-Hydroxytestosterone	Sigma Aldrich
Amitriptyline metabolite, (±)-E-10-hydroxylated	Sigma Aldrich
Amitriptyline N-oxide	Toronto Research Chemicals
Desmethylnortriptyline	Toronto Research Chemicals
Hydroxy Ritonavir	Toronto Research Chemicals
4-Amino-N-((2*R*,3*S*)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide	Toronto Research Chemicals

Oligonucleotides and plasmids

Table S3: Synthetic oligonucleotides for cloning. Restriction sites are underlined; inserted codons are shown in small letters.

Primer name	DNA-sequence (5'-3')	Restriction enzymes	usage
F-NcoI-SPL01100	GAGCCCATGGCCAGCACTGAAGAACCAGC	NcoI/XhoI	Amplification of SPL_01100
RC-SPL1100-XhoI	GATCCTCGAGTTACCACCTGACGGGCAGTGC		
F-NdeI-SPL02318	TATATTCCATATGCCCGTTCGCCGCCCAGCG	NdeI/XhoI	Amplification of SPL_02318
RC-SPL2318-XhoI	TAATATCTCGAGTCACTTTCGGCCTCCAGCGTTGCC		
F-NdeI-SPL00625	GATGCGCCATATGACCGAAGCCATCCCCTAC	NdeI/XhoI	Amplification of SPL_00625
RC-SPL0625-XhoI	TATATTCTCGAGCTACCAGGCCAGCGGCAG		
F-NdeI-SPL01896	GATCGACCATATGAGTGAAATGGCCGCCATC	NdeI/XhoI	Amplification of SPL_01896
RC-SPL1896-XhoI	TATATTCTCGAGCTACCGGGTCGTTTCCGTCC		
F-NdeI-SPL03767	GCTCGACCATATGCAGCACGAACAGATCGCACC	NdeI/XhoI	Amplification of SPL_03767
RC-SPL3767-XhoI	TGATCTCTCGAGCTACTCCTTGTCCCAGGTGACGG		
F-NdeI-SPL06346	GCGCGCACATATGTCGGCATTATCCAACTCCC	NdeI/XhoI	Amplification of SPL_06346
RC-SPL6346-XhoI	TATATTCTCGAGTCACCCCAGCCGCAGCGG		
pet24_AAU00443_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGACCGCACCGACCGGAGCGCCCG	NdeI/HindIII	Amplification of AAU_00443
pet24_AAU00443rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCAGGACAGCAGCACCGGCAGCTCCCC		
pet24_AAU00941_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGACCCCGACCGCCGCCGACGTCACC	NdeI/HindIII	Amplification of AAU_00941
pet24_AAU00941rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCATCCGGTGCCCGTGCGCAGGGGGTAGC		
pet24_AAU01045_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGGACGTCGACACCCGGGACGGCACG	NdeI/HindIII	Amplification of AAU_01045
pet24_AAU01045rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCACCAGGTGACGGGCAGCCGCTCCAGG		

Primer name	DNA-sequence (5'-3')	Restriction enzymes	usage
pet24_AAU01156_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGCAGCAATGCACACCGTGCAGCG	NdeI/HindIII	Amplification of AAU_01156
pet24_AAU01156rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCAGTTCCAGGTGACCGGCAACGAGTGC		
pet24_AAU01495_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATG**ATG**TCCGTCGTACCCCCCGCCGGCTCC	NdeI/HindIII	Amplification of AAU_01495
pet24_AAU01495rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCACAGGATCACCGGCCGGCTCTCCCAGC		
pet24_AAU01848_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGATCACCGCACCGCAGCTGCCCTTCG	NdeI/HindIII	Amplification of AAU_01848
pet24_AAU01848rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCACCAGGTGACCGGCAGCGCCCGC		
pet24_AAU02550_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGGCCAGTGTTGCCGCACCGCACGTCC	NdeI/HindIII	Amplification of AAU_02550
pet24_AAU02550rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCAGGTGCGGAACCGCATCGGCATTCG		
pet24_AAU03957_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGCATGCGTCCGATTCCCCAGACCCC	NdeI/HindIII	Amplification of AAU_03957
pet24_AAU03957rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCAACGCAATGAGGGCCGTTCGGCCG		
pet24_AAU05422_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGGGCGGGGCGCACGCCGGC	NdeI/HindIII	Amplification of AAU_05422
pet24_AAU05422rv	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCACCTGGCCGCCGGTCCCGTTGCG		
pet24_AAU05736_f	GAAATAATTTTGTTTAACTTTAAGAAGGAGATATA CATATGATGACGACCGTCGACGAGTTCCCGC	NdeI/HindIII	Amplification of AAU_05736
pet24_AAU05736_r	GGTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAA GCTTTCACCAGGTCACCGGAAGCTGGTGC		
F-nHis-SPL00625	caccaccacACCGAAGCCATCCCCTAC	NdeI/XhoI	Insertion of N-terminal His_6-tag for SPL_00625
R-nHis-SPL00625	atgatgatgCATATGTATATCTCCTTCTTAAAGTTAAA CAAAATTATTTC		
RC-SPL1896-cHis	TATATTCTCGAGCCGGGTCGTTTCCGTCCAGG	NdeI/XhoI	Primer without stop codon of SPL_01896, used for addition of a C-terminal His_6-tag
RC-SPL3767-cHis	TGATCTCTCGAGCTCCTTGTCCCAGGTGACGG	NdeI/XhoI	Primer without stop codon of SPL_03767, used for addition of a C-terminal His_6-tag
RC-SPL6346-cHis	TATATTCTCGAGCCCCAGCCGCAGCGGCAG	NdeI/XhoI	Primer without stop codon of SPL_06346, used for addition of a C-terminal His_6-tag
fw-PP-pCOLADuet	CAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAA TAAGGAGATATACCATGGCTTCTAAAGTAGTGTAT GTGTCACATG	NcoI/BamHI	Insertion of camA und camB in pCOLA-Duet1
Rv-PP-pCOLADuet	CAAGCTTGTCGACCTGCAGGCGCGCCGAGCTCGA ATTCGGATCCTCAGGCACTACTCAGTTCAGCTTTG		

Table S4: Plasmids used in this study.

vector (internal number)	genes with EMBL bank number	vector properties	reference/source
pET22b(+)	/	P_{T7}, *lacI,* pBR322 *ori,* Ap^R	Novagen
pET24b(+)		P_{T7}, *lacI,* pBR322 *ori,* Km^R	Novagen
pET28a(+)		P_{T7}, *lacI,*pBR322 *ori,* Km^R	Novagen
pCOLADuet-1		P_{T7}, *lacI,*pCOLA *ori,* Km^R	Novagen
pET24b-SPL_00625 (THI 09)	SPL_00625 = *cyp105D* (EMBL-Bank: OSY47991)	SPL_00625 (NdeI,XhoI) cloned in pET24b	This work
pET28a-SPL_01100 (THI 10)	SPL_01100 (EMBL-Bank: OSY47382)	SPL_01100 (NcoI,XhoI) cloned in pET28a	This work
pET24b-SPL_01896 (THI 11)	SPL_01896 = *cyp105AA* (EMBL-Bank: OSY46529)	SPL_01896 (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_02318 (THI 07)	SPL_02318 (EMBL-Bank: OSY46350)	SPL_02318 (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_03767 (THI 13)	SPL_03767 = *cyp105AC* (EMBL-Bank: OSY44745)	SPL_03767 (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_06346 (THI 16)	SPL_06346 = *cyp107Z* (EMBL-Bank: OSY37796)	SPL_06346 (NdeI,XhoI) cloned in pET24b	This work
pET24b-AAU_00443 (FLB 02)	AAU_00443 (EMBL-Bank: OSY43500)	AAU_00443 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_00941 (FLB 03)	AAU_00941 (EMBL-Bank: OSY43336)	AAU_00941 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_01045 (FLB 04)	AAU_01045 (EMBL-Bank: OSY42804)	AAU_01045 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_01156 (FLB 05)	AAU_01156 (EMBL-Bank: OSY42914)	AAU_01156 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_01495 (FLB 06)	AAU_01495 (EMBL-Bank: OSY42000)	AAU_01495 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_01848 (FLB 07)	AAU_01848 (EMBL-Bank: OSY41819)	AAU_01848 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_02550 (FLB 08)	AAU_02550 (EMBL-Bank: OSY40792)	AAU_02550 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_03957 (FLB 09)	AAU_03957 (EMBL-Bank: OSY38754)	AAU_03957 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_05422 (FLB 10)	AAU_05422 (EMBL-Bank: OSY36345)	AAU_05422 (NdeI,HindIII) cloned in pET24b	This work
pET24b-AAU_05736 (FLB 11)	AAU_05736 (EMBL-Bank: OSY35899)	AAU_05736 (NdeI,HindIII) cloned in pET24b	This work

vector (internal number)	genes with EMBL bank number	vector properties	reference/source
pET16b-YkuN-nHis$_{10}$ (THI 23)	*ykuN* with N-terminal His$_{10}$-tag (EMBL-Bank: CAA10877)	*ykuN*-nHis$_{10}$ (XhoI, BamHI) cloned in pET16b	Girhard, et al. [1]
pET11a-Fdr-nHis$_6$ (THI 25)	*fdr* with N-terminal His$_6$-tag (EMBL-Bank: L04757)	*fdr*-nHis$_6$ (NdeI, BamHI) cloned in pET11a	Bakkes, et al. [2]
pET28a-camA-nHis$_6$ (THI 01)	*camA* with N-terminal His$_6$-tag (EMBL-Bank: BAA00413)	*camA* (NdeI, EcoRI) cloned in pET28a	Girhard, et al. [3]
pET28a-camB-nHis$_6$ (THI 02)	*camB* with N-terminal His$_6$-tag (EMBL-Bank: BAA00414)	*camB* (NdeI, EcoRI) cloned in pET28a	Girhard, et al. [3]
pET22b-gdhIV (THI 41)	*gdhIV* (EMBL-Bank: D10626)	*gdhIV* (NdeI, XhoI) cloned in pET22b	Le-Huu et al. [4]
pET24b-SPL_00625-nHis$_6$ (THI31)	*cyp105D* with N-terminal His$_6$-tag	SPL_00625-nHis$_6$ (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_01896-cHis$_6$ (THI34)	*cyp105AA* with C-terminal His$_6$-tag	SPL_01896-cHis$_6$ (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_03767-cHis$_6$ (THI36)	*cyp105AC* with C-terminal His$_6$-tag	SPL_03767-cHis$_6$ (NdeI,XhoI) cloned in pET24b	This work
pET24b-SPL_06346-cHis$_6$ (THI40)	*cyp107Z* with C-terminal His$_6$-tag	SPL_06346-cHis$_6$ (NdeI,XhoI) cloned in pET24b	This work
pET22b-SPL_00625 (THI 78)	*cyp105D*	SPL_00625 (NdeI,XhoI) cloned in pET22b	This work
pCOLADuet-PP (THI 87)	*camA* and *camB*	*camA* and *camB* (MCSI: NcoI, BamHI) cloned in pCOLADuet1	This work
pETDuet-slic-PP	*camA* and *camB*	*camA* and *camB* (MCSI) cloned in modified pETDuet1	Florian Tieves

Classification of the P450 genes

Novel P450s were selected using the Cytochrome P450 Engineering database (CYPED) and classified according to the CYP nomenclature from David Nelson (**Table S4; S5**).[5] According to the nomenclature, *cyp* sequences with more than 40 % sequence identity belong to the same family, while a sequence identity higher than 55 % is the threshold for P450s belonging to the same subfamily. Genes with lower sequence identities than 40 % are referred as "CYP-like" genes. The P450 sequences were classified as accurately as possible following a two-step protocol. First, the sequences were grouped into the CYP-family using CYPED. In order to find an already existing CYP-subfamily, the sequences were then manually aligned with ClustalOmega applying the reference sequences of David Nelson's 'Cytochrome P450 homepage'. [6]

Table S5: CYP105 and CYP107-members from *S. platensis* and *P. autotrophica* with their closest homologous. Reference sequences are extracted from the 'Cytochrome P450 homepage'.

Gene name	closest sequence identity to:	identity [%]	suggested CYP-classification:
SPL_00625	CYP105D8 from *Streptomyces tubercidicus* strain I-1529	85	CYP105D
SPL_01896	CYP105AA5 from *Streptomyces lydicus*	88	CYP105AA
SPL_03767	CYP105AC1 from *Saccharopolyspora erythraea* NRRL23338	69	CYP105AC
SPL_06346	CYP107Z9 from *Streptomyces tubercidicus* strain NRAA-7027	99	CYP107Z9v2
AAU_01045	CYP107AL1 from *Streptomyces scabies*	63	CYP107AL
AAU_03957	CYP107AQ1 from *Saccharopolyspora erythraea* NRRL23338	62	CYP107AQ
AAU_05736	CYP105AB3 from *Nonomuraea recticatena*	58	CYP107AB
SPL_01100	CYP107	53	new subfamily of CYP107
SPL_02318	CYP105	52	new subfamily of CYP105
AAU_00443	CYP107	51	new subfamily of CYP107
AAU_00941	CYP107	44	new subfamily of CYP107
AAU_01156	CYP105	49	new subfamily of CYP105
AAU_05422	CYP107	51	new subfamily of CYP107
AAU_01495	CYP107-like	38	/
AAU_01848	CYP107-like	39	/
AAU_02550	CYP107-like	31	/

Table S6: Sequence identity of drug metabolizing P450s from *S. platensis* compared to members of their subfamily. Reference sequences are extracted from the 'Cytochrome P450 homepage'.

P450 name	SPL_06346	P450 name	SPL_00625	P450 name	SPL_01896	P450 name	SPL_03767
saAcmM	68	CYP105D6	57	CYP105AA1	55	CYP105AC1	69
CYP107Z1	67	CYP105D9	54	CYP105AA7	73	SPL_03767	100
CYP107Z2v2	68	CYP105D1	66	CYP105AA4	81		
CYP107Z2v1	67	CYP105D2	67	CYP105AA6	85		
CYP10Z13	74	SPL_00625	100	CYP105AA3	86		
CYP107Z3	75	CYP105D8	85	CYP105AA2	86		
CYP107Z4	81	CYP105D7	70	CYP105AA5	88		
CYP10Z75v3	82	CYP105D4	66	SPL_01896	100		
CYP107Z5v2	82	CYP105D5	66				
CYP107Z5v1	82	CYP105D3	77				
CYP107Z12	85						
CYP107Z11	84						
CYP107Z10	84						
CYP107Z6	85						
SPL_06346	100						
CYP107Z9	99						
CYP107Z8	90						
CYP107Z7	88						

Gene expression and purification

P450s

For expression of P450 genes, 50 mL TB-medium with 30 µg/mL kanamycin were inoculated to an OD_{600} of 0.05 with an overnight culture of the respective *E. coli* C43 (DE3) strain. The cultures were grown in 500 mL flasks at 37 °C and 180 rpm for 3 h. Expression of the target genes was induced by adding 500 µM isopropyl-β-D-thiogalactopyranoside (IPTG). Further, 500 µM 5-aminolevulinic acid (5-ALA) was added. Then, all cultures were incubated at 25 °C and 180 rpm for 20 h. Cells were harvested via centrifugation (5251 g, 4 °C min. 15 min) and resuspended in 5 mL 50 mM potassium phosphate buffer (KPi) pH 7.5 with 0.1 M phenylmethylsulfonyl fluoride (PMSF) and disrupted by sonication (Branson Sonfier 250; 3 x 2 min, 40 % amplitude, duty cycle 4). The soluble fraction (crude cell extract) was collected after centrifugation (40.000-50.000 g, 25 min and 4 °C) and frozen in small aliquots of 0.5 – 1 mL. In order to refer the P450 amount to the cell dry weight (CDW), 3 x 200 µL cell suspension were transferred to 1.5 mL reaction tubes before sonication and centrifuged for 2 min at 13.200 g. The supernatant was discarded and the cell pellets dried at 60 °C for at least 48 h before weighing. The determination of all P450 concentrations is based on at least two independent cultivations.

The expression of P450 mutants with a His_6-tag was performed on 400 mL scale in 2 L shaking flasks with the same conditions as described before. After centrifugation for 30 min at 2831 g and 4 °C cell pellets were resuspended in equilibration buffer (50 mM KPi, pH 8 supplemented with 500 mM NaCl and 0.1 M PMSF) and disrupted by sonication (5 x 2 min) or in 3 consecutive runs at 1.35 kbar in a high pressure homogenisator (Cell Disruption System, Constant Systems Limited). Next to the centrifugation, the crude cell extract was filtered with a 0.45 µm syringe filter (GE healthcare) and loaded on two HisTrap FF crude columns in series (5 mL, GE-Healthcare), which was connected to an ÄKTA purifier 10 liquid chromatography system (GE Healthcare). After washing steps with equilibration buffer and washing buffer containing imidazol concentrations up to 80 mM, the elution of the target protein was performed with buffer containing 300 mM imidazole. The elution fractions containing the target enzyme were pooled, concentrated to a volume of less than 3 mL (membrane cut-off 10 kDa) and then desalted using a Econo-Pac® 10DG Desalting gravity column (Bio-rad) according to the manufacturer's manual. The purified enzyme was eluted from the column with 50 mM KPi containing 5 % glycerol and 50 mM NaCl, aliquoted in portions of 100-500 µL and stored at - 20 °C until further use.

GDH and redox partner proteins

Gene expression and purification of the redox partner proteins YkuN and Fdr was carried out as described previously.[2,7] The redox partners Pdr and Pdx were expressed in recombinant *E. coli* BL21 (DE3) using plasmids described in previous studies[3].The expression of *pdx* and *pdr* was conducted in the same manner as for the P450s, while purification was carried out as described by Worsch and co-workers.[7] The gene encoding GDHIV from *B. megaterium* was expressed as described elsewhere[4] and optionally purified via salt precipitation and heat treatment [8].

Pdr (ε_{480} =8.5 mM^{-1} cm^{-1}; ε_{454} =10 mM^{-1} cm^{-1}, ε_{378} =9.7 mM^{-1} cm^{-1}) and Pdx (ε_{415} =11.1 mM^{-1} cm^{-1}, ε_{455} =10.4 mM^{-1} cm^{-1}) was calculated as the average concentration from various wavelength-specific extinction coefficients.[9] The concentration of Fdr and YkuN was based on their extinction coefficients ε_{456}= 7.1 mM^{-1} cm^{-1} and ε_{461}= 10 mM^{-1} cm^{-1}.[10]

Isolation of 2β-hydroxytestosterone for structure elucidation

Whole-cell biotransformations on a semipreparative scale were performed in 10 mL reaction volume in 100 mL flasks. As model reaction the conversion of 1 mM testosterone (in 2 % (v/v) ethanol) was conducted with *E. coli* C43 (DE3) pET22b-SPL_00625 + pCOLADuet-PP in 220 mL total volume. The genes were integrated in a two-plasmid system in *E. coli* C43 (DE3) as described previously.[7] The *cyp* genes were cloned in the pET22b vector, while the redox partner genes were integrated in pCOLADuet-1. Main cultures of the P450 whole cell

biocatalysts were prepared analogous to the expression cultures when P450s were produced singularly. After expression induction with 0.5 mM IPTG, cultures were incubated at 20 °C and 140 rpm for 20 h. Cell pellets obtained after centrifugation were washed with PSE-buffer (6.75 g/L KH_2PO_4, 85.5 g/L sucrose, 0.93 g/L EDTA-Na_2*2 H_2O, pH 7.5), adjusted to a cell wet weight (CWW) of 100 mg/mL and stored at -20 °C. After thawing, 50 mg/mL (CWW) resting cells were mixed with 1 x nutrient solution (6 mM glucose, 6 mM lactose and 12 mM citrate in PSE-buffer) and substrate. Next, the reaction mixture was incubated at 25 °C and 200 rpm shaking frequency (shaking incubator, Infors HT). Samples for LC/MS analysis were taken both at the beginning and after incubation. After 20 h, 20-30 mL reaction volume were extracted twice with 25 mL ethyl acetate, combined, dried over $MgSO_4$ and evaporated under reduced pressure. Products were isolated on a silica column via flash column chromatography (ethyl acetate: n-hexane = 4:1). 6.6 mg **7b** was isolated as white-yellow solid with ≥ 95 % purity (PDA) (10 % isolated yield).

Chromatographic conditions for LC/MS analysis

Table S7: Chromatographic conditions for the LC/MS-analysis of the drugs and their metabolites.

Compound	Internal standard	Solvent program	Reference
1	Darunavir	Gradient from 30 % B to 80 % B for 20 min, then gradient from 80 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 30 % B for 5 min Solvent B: methanol	Worsch et al., 2018[7]
2,3	Rupintrivir	Gradient from 20 % B to 75 % B for 20 min, then gradient from 75 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 20 % B for 5 min Solvent B: methanol	This work
5	Rupintrivir	Gradient from 40 % B to 60 % B for 20 min, then gradient from 60 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 40 % B for 5 min Solvent B: methanol	This work
6	Amprenavir	Gradient from 40 % B to 55 % B for 20 min, then gradient from 55 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 55 % B for 5 min Solvent B: methanol	This work
4	Ritonavir	Gradient from 40 % B to 70 % B for 20 min, then gradient from 70 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 40 % B for 5 min Solvent B: methanol	This work
7-15	Progesterone (7,9-15) Triamcinolone (8)	Gradient from 10 % B to 75 % B for 10 min, hold 75 % B for 5 min, equilibration at 10 % B for 5 min Solvent B: methanol	Modified according to von Bühler, 2013 [11]
17-25 (except 18)	Oxcarbazepine (16-23) Amitriptyline (24-25)	Gradient from 23 % B to 33 % B for 10 min, hold 100 % B for 1 min, equilibration at 23 % B for 4 min Solvent B: acetonitrile	This work
18	Oxcarbazepine	Gradient from 28 % B to 38 % B for 10 min, hold 100 % B for 1 min, equilibration at 38 % B for 4 min Solvent B: acetonitrile	This work

Supplementary Results

Gene expression in *E. coli* C43 (DE3)

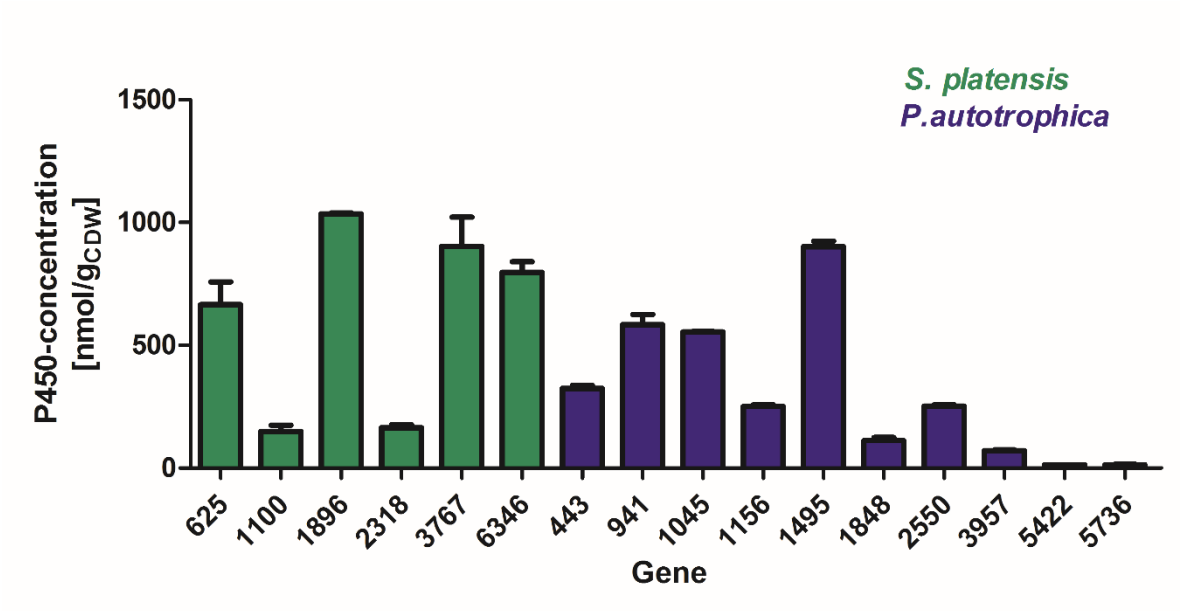


Figure S1: P450 concentrations achieved in *E. coli* C43 (DE3). Genes from *S. platensis* are shown in green; genes from *P. autotrophica* in blue. Measurements represent mean and standard deviation from two independent cultivations. Data were plotted with GraphPad Prism

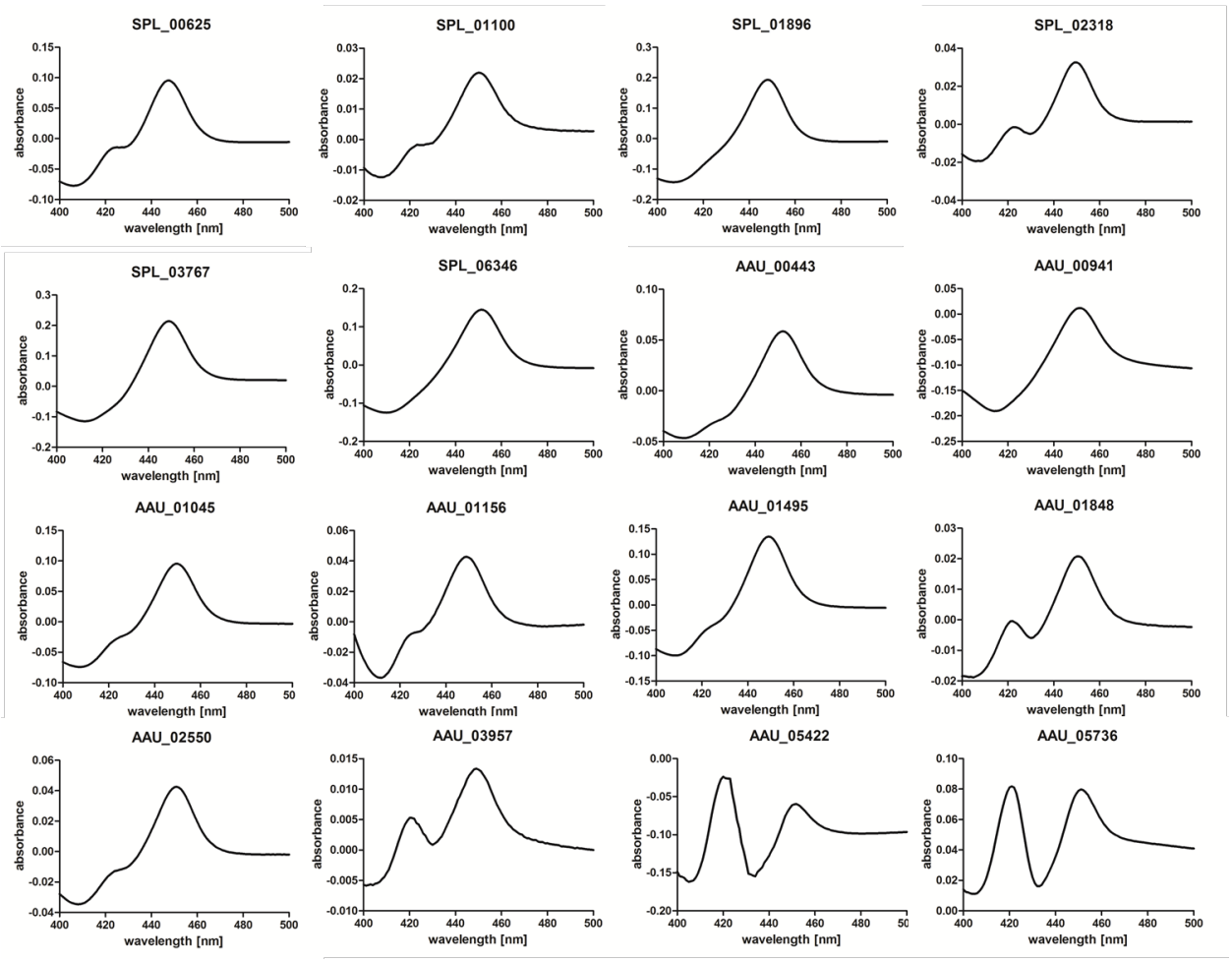


Figure S2: CO-difference spectra of the investigated P450s from *S. platensis* (SPL) and *P. autotrophica* (AAU). Data were plotted with GraphPad Prism

Identification of suitable redox partners for every reaction set up

For identification of the optimal redox partner pair for every model substrate class, purified P450s were tested for the compounds **1**, **7** and **16** with the redox partner proteins flavodoxin from *B. subtilis* (YkuN) and flavodoxin reductase (Fdr) from *E. coli* or putidaredoxin (Pdx) and putidaredoxin reductase (Pdr) from *P. putida.* The proteins were either highly enriched or purified up to 97 % purity as it was deduced from SDS-PAGE analysis (**Figure S4**). The redox partner/P450 combination, which yielded the highest conversion, was used for the systematic substrate screening described in the main paper.

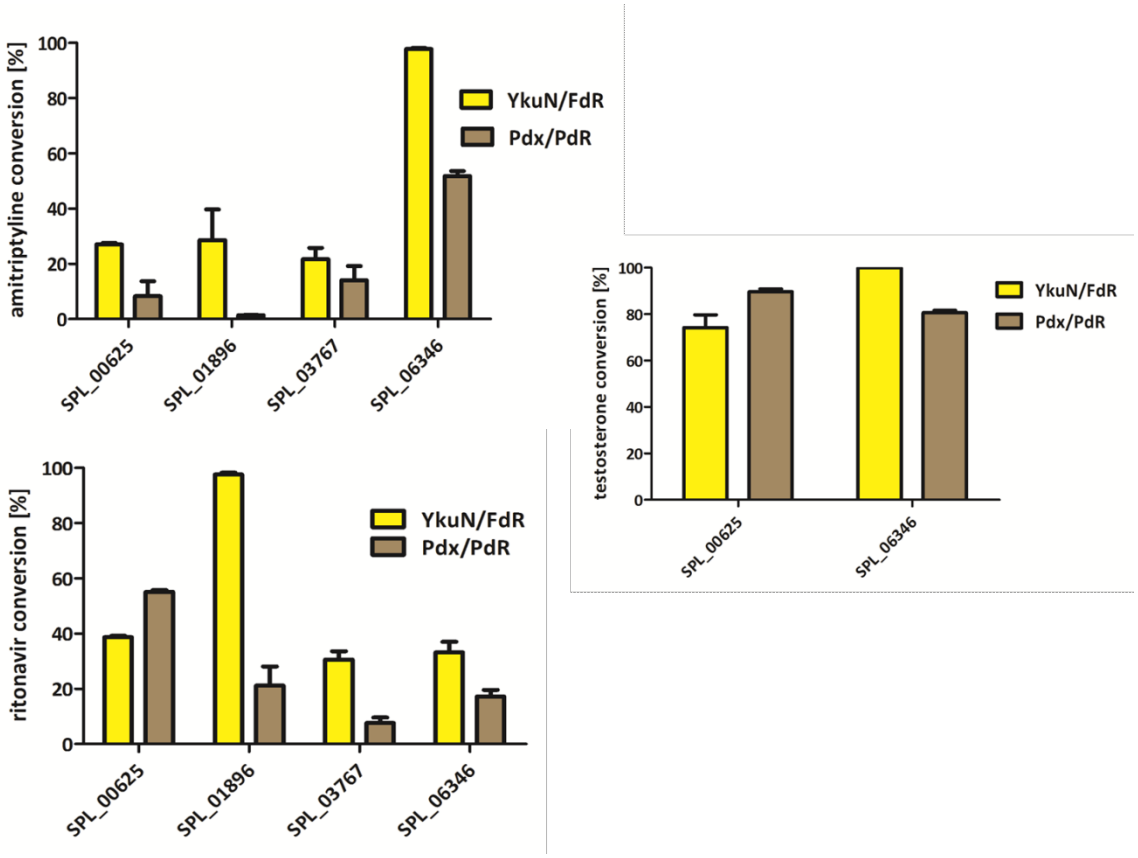


Figure S3: Comparison of conversions [%] with P450s from *S. platensis* with different redox partners for the model substrates ritonavir, amitriptyline and testosterone. Reaction conditions: Conversion of 0.2 mM substrate (in 2 % co-solvent) with 2 µM purified CYP, 10 µM YkuN or Pdx, 2 µM Fdr or Pdr (ratio 1:10:1) and 120 U catalase. Cofactor regeneration of $NAD(P)^+$ by 1 U GDH from *B. megaterium* with 20 mM glucose. The conversion was performed for 4 h (amitriptyline) or 20 h (testosterone, ritonavir) in 200 µL 50 mM phosphate buffer, pH 7.5 in 2 mL reaction tubes with open lids at 25 °C and 300 rpm in Eppendorf shakers. Due to high activity of SPL_00625 against testosterone, 1 mM substrate was used. Conversions are based on the amount of the ratio between the substrate and an internal standard. Measurements are means of technical duplicates. Data were plotted with GraphPad Prism

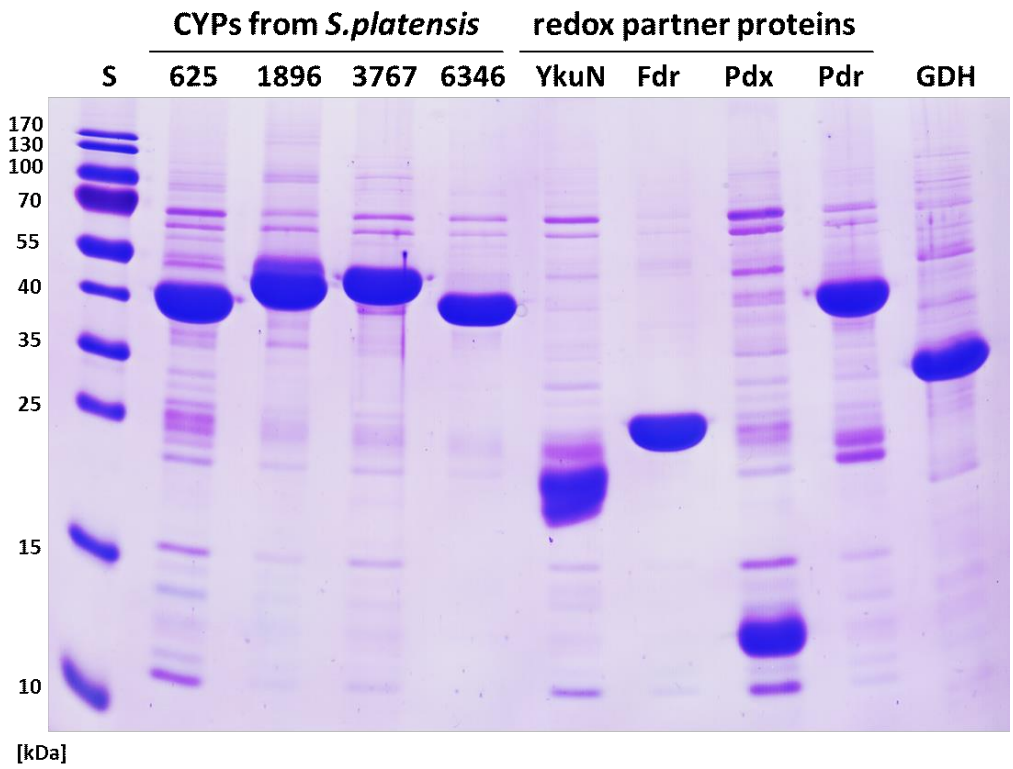


Figure S4: SDS-PAGE analysis of all enzymes used for the redox partner identification experiments. The purity of the target enzymes was between 45 % (Pdx) and 97 % (Fdr) as it was deduced from ImageJ (Rasband WS) analysis.

LC/MS-data

Anti-HIV agents

Table S8: Product analysis of ritonavir (1). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Ritonavir (1)	19.0	720.9	0		
main hydroxylation (1a)	15.6	736.9	+16	hydroxy ritonavir	commercial standard
demethylation (1b)	17.9	706.9	-14	N-demethylation of methyl urea nitrogen	LC-MS-MS data published in Worsch et al.[7]
dealkylation (1c)	13.6	581.7	-139	N-dealkylation on carbamid group	LC-MS-MS data published in Worsch et al.[7]
dealkylation (1d)	14.5	579.8	-141	N-dealkylation on carbamid group	
double oxidation (1 e)	14.0	752	+32	double oxidation products	LC-MS-MS data published in Worsch et al.[7]
double oxidation (1 f)	14.1		+32		
single oxidation and demethylation (1 g)	14.7	722.9	+2	demethylation product of 1a	LC-MS-MS data published in Worsch et al.[7]
other oxidation products (unidentified)	**Retention time ($[M+H]^+$; Δm):**10.9 min (m/z 597; **-125**); 16.5 min (m/z :722; **±0**); 16.9 min (m/z 736; **+16**); 17.3 min (m/z 718;**-4**)				

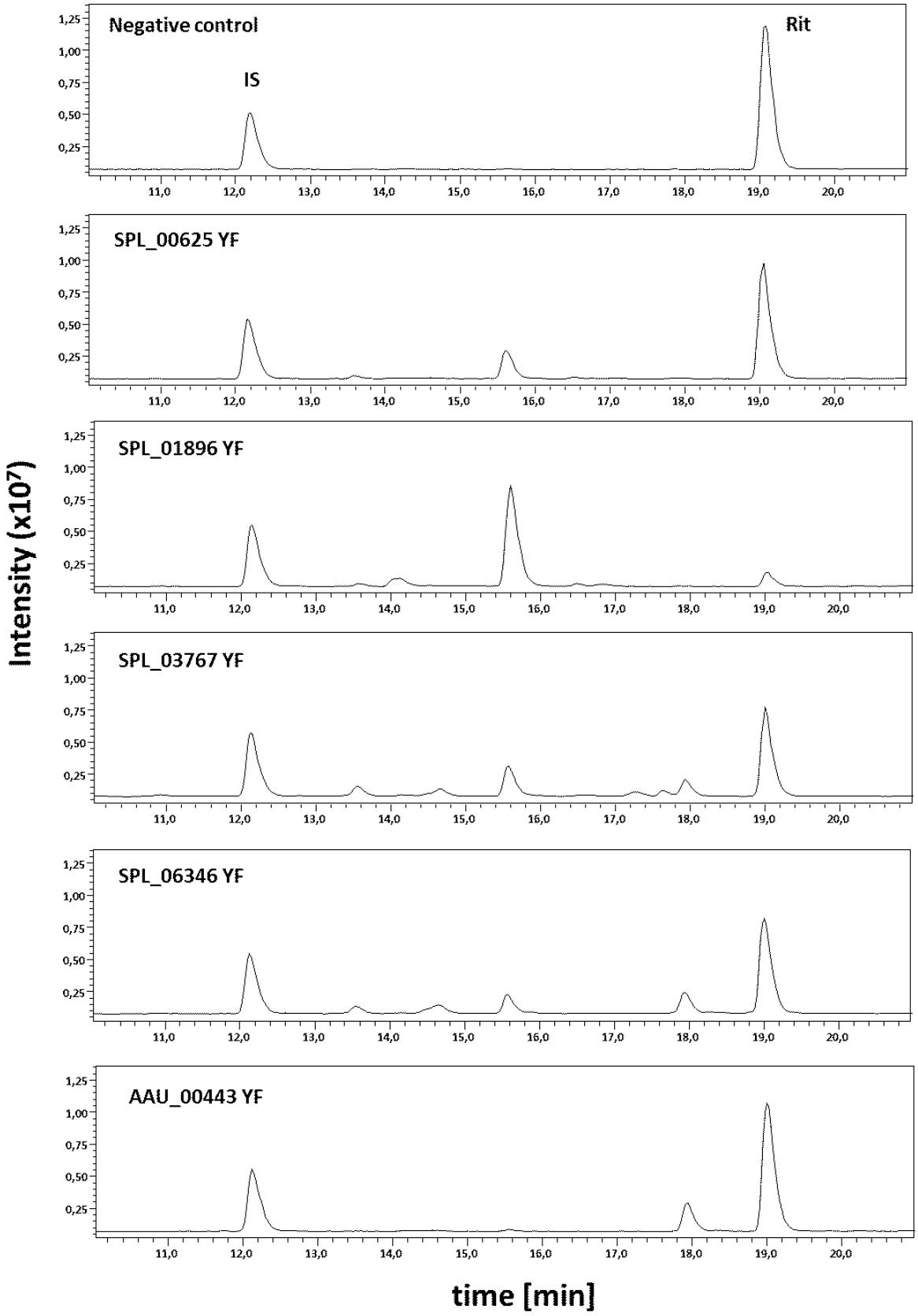


Figure S5: LC/MS-chromatograms for the conversion of ritonavir (1) in the initial screening. Rit: Ritonavir, IS: internal standard darunavir.

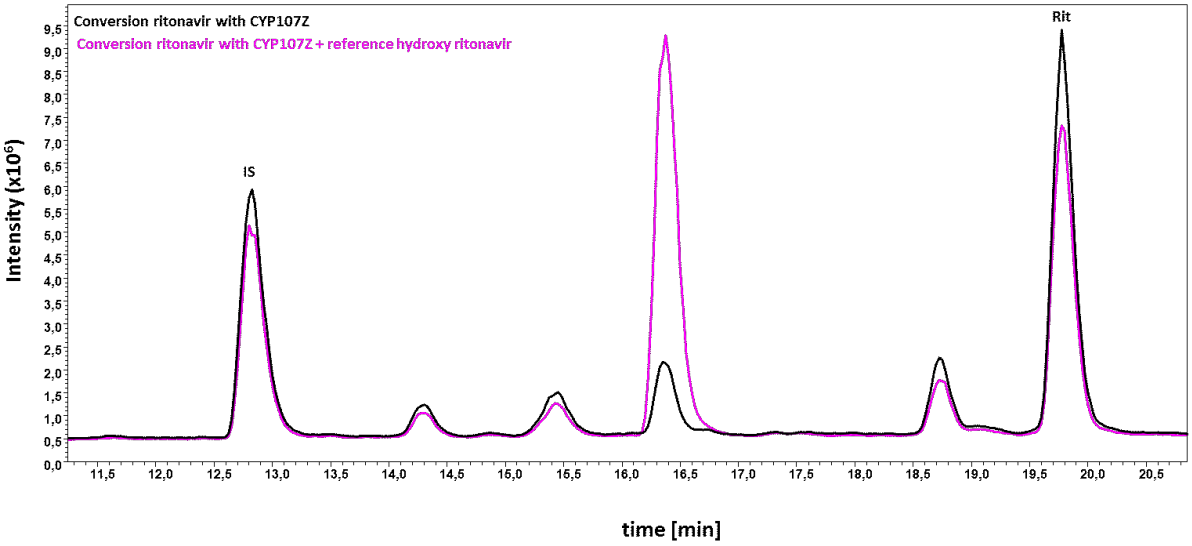


Figure S6: Identification of hydroxy ritonavir as a product via spiking of an authentic reference compound. Rit: ritonavir IS: internal standard darunavir. Black: conversion of ritonavir with CYP107Z, pink: conversion of ritonavir with CYP107Z spiked with hydroxy ritonavir.

Table S9: Product analysis of amprenavir (2). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Amprenavir (2)	15.7	505.6	0		
dealkylation (2a)	10.9	391.5	-114	4-amino-N-((2*R*,3*S*)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide	commercial standard
single oxidation (2b)	12.5	521.6	+16	proposed: hydroxylation	
single oxidation (2c	12.2	521.6	+16	proposed: hydroxylation	
single oxidation (2d)	13.0	521.6	+16	proposed: hydroxylation	
single oxidation (2e)	15.2	521.6	+16	proposed: hydroxylation	
single oxidation (2f)	15.3	521.6	+16	proposed: hydroxylation	
other oxidation products (unidentified)	**Retention time ($[M+H]^+$; Δm):** 11.2 min (m/z 521; **+14**); 13.0 min (m/z :505; **-2**); 13.7 min (m/z 521; **+14**); 13.8 min (m/z 525;**+18**), 14.0 min (m/z 505;**-2**), 14.2 min (m/z 505;**-2**),				

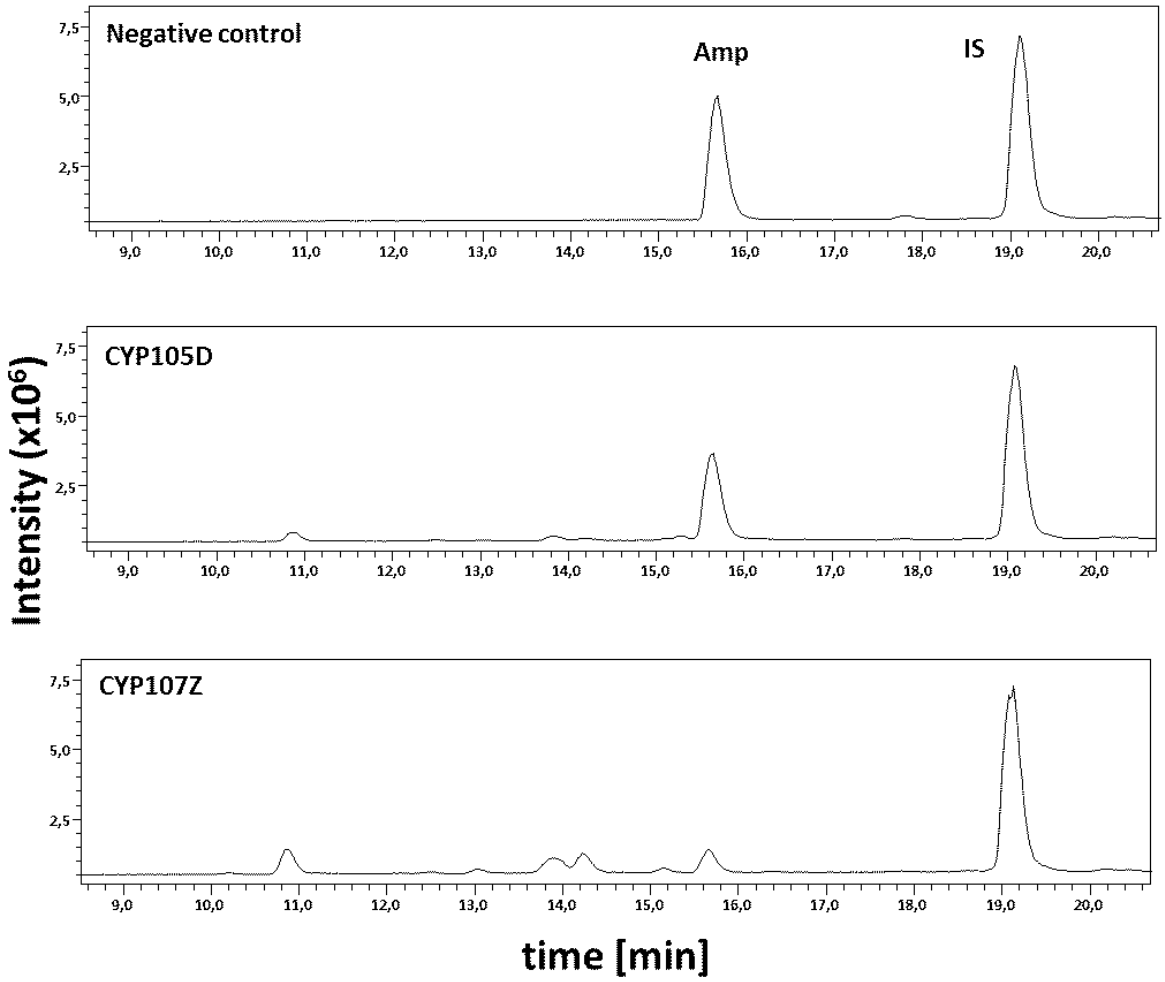


Figure S7: LC/MS-chromatograms for the conversion of amprenavir (2). Amp: Amprenavir, IS: internal standard rupintrivir.

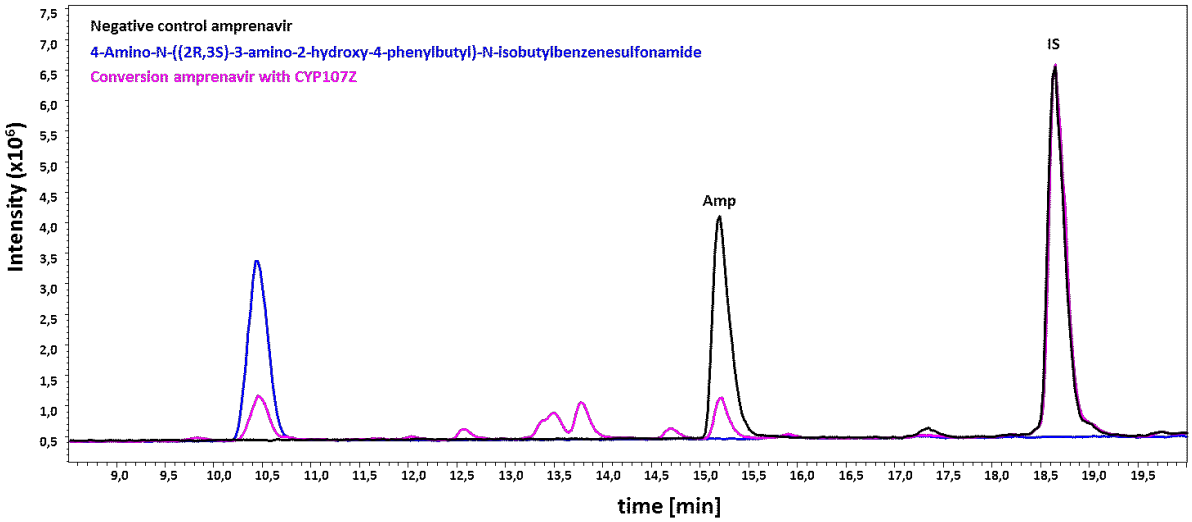


Figure S8: Comparison of the commercial product standard 4-Amino-N-((2R,3S)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide with the conversion of amprenavir (2). Amp: amprenavir, IS: internal standard rupintrivir. Black: control reaction without P450, pink: conversion of amprenavir with CYP107Z, pink: 4-Amino-N-((2R,3S)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide.

Table S10: Product analysis of darunavir (3). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Darunavir (3)	15.8	547.7	0		
dealkylation (2a)	10.9	391.5	-156	4-Amino-N-((2*R*,3*S*)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide	commercial standard
single oxidation (3b)	12.5	563.7	+16	proposed: hydroxylation	
single oxidation (3c)	13.3	563.7	+16	proposed: hydroxylation	
single oxidation (3d)	14.1	563.7	+16	proposed: hydroxylation	
single oxidation (3e)	14.5	563.7	+16	proposed: hydroxylation	
single oxidation (3f)	15.4	563.7	+16	proposed: hydroxylation	
double oxidation (3g)	8.1	407.5	-140	proposed: hydroxylation of **2a**	
double oxidation (3h)	10.2	407.5	-140	proposed: hydroxylation of **2a**	

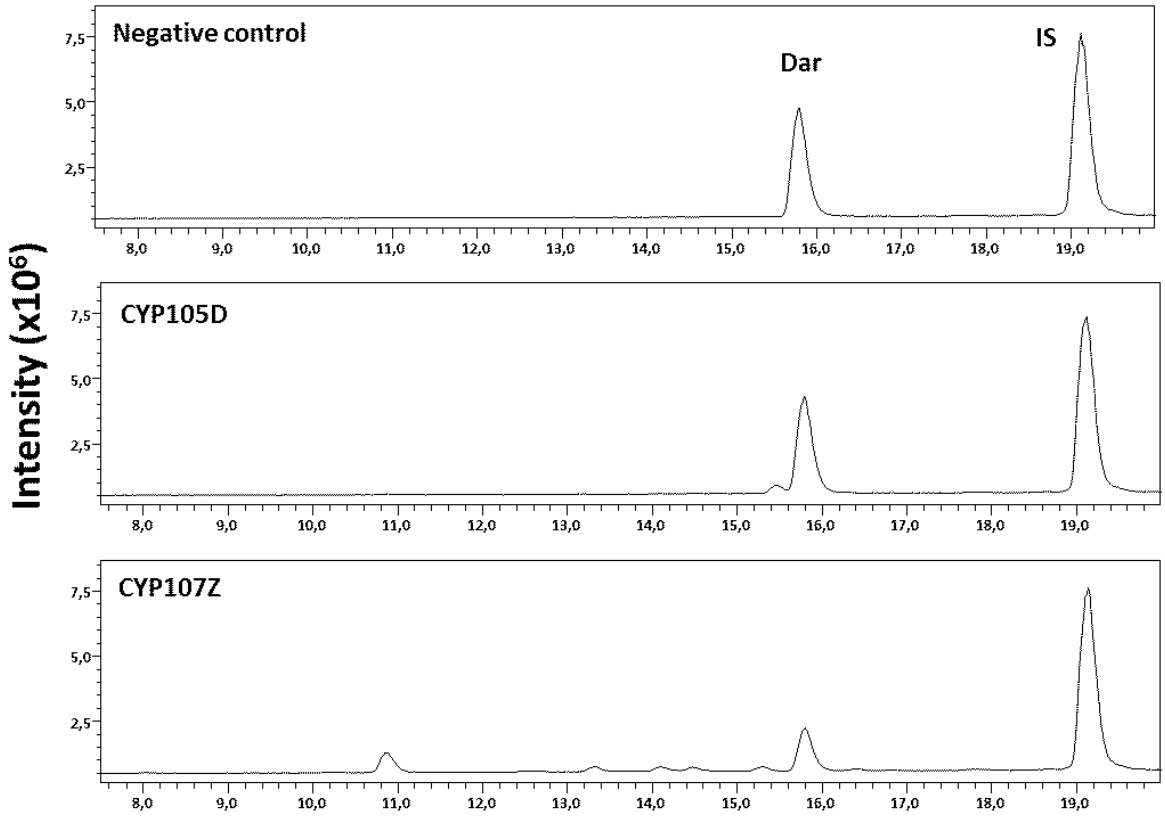


Figure S9: LC/MS-chromatograms for the conversion of darunavir (3). Dar: Darunavir, IS: internal standard rupintrivir.

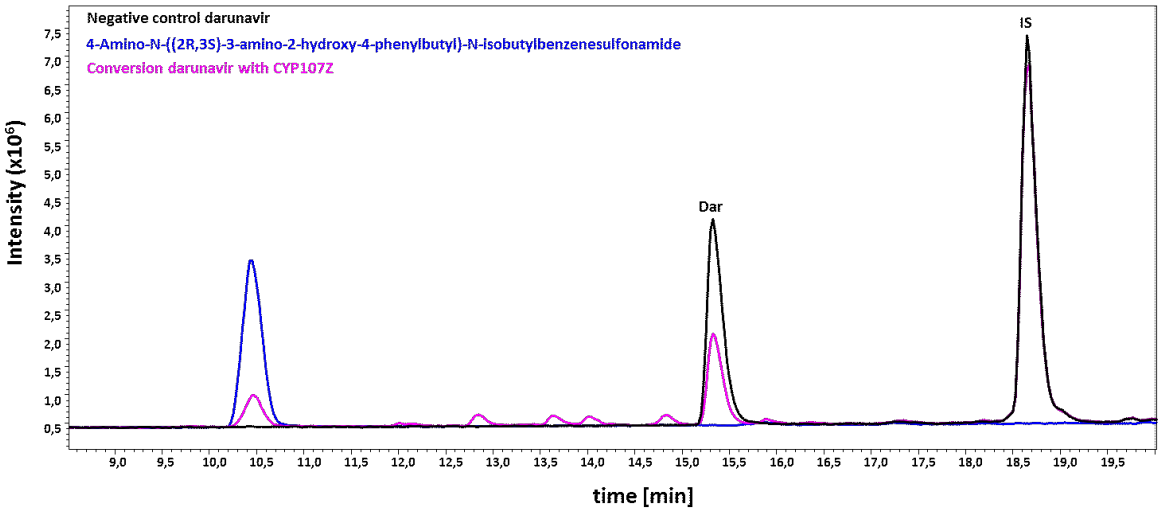


Figure S10 Comparison of the commercial product standard 4-Amino-N-((2R,3S)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide with the conversion of darunavir (3). Dar: darunavir, IS: internal standard rupintrivir. Black: control reaction without P450, pink: conversion of darunavir with CYP107Z, pink: 4-Amino-N-((2R,3S)-3-amino-2-hydroxy-4-phenylbutyl)-N-isobutylbenzenesulfonamide.

Table S11: Product analysis of atazanavir (4). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Atazanavir (4)	18.2	704.9	0	
single oxidation (4a)	13.2	720.9	+16	proposed: hydroxylation
single oxidation (4b)	13.7	720.9	+16	proposed: hydroxylation
single oxidation (4c)	16.3	720.9	+16	proposed: hydroxylation
dealkylation (4d)	9.0	646.9	-58	proposed: N-dealkylation on carbamid group
dealkylation (4e)	9.8	646.9	-58	proposed: N-dealkylation on carbamid group
double oxidation (4f)	14.5	718.9	+14	proposed: hydroxylation and follow up oxidation of a hydroxy group
double oxidation (4g)	15.8	718.9	+14	proposed: hydroxylation and follow up oxidation of a hydroxy group
double oxidation (4h)	16.9	718.9	+14	proposed: hydroxylation and follow up oxidation of a hydroxy group
double oxidation (4i)	19.3	718.9	+14	proposed: hydroxylation and follow up oxidation of a hydroxy group
double oxidation (4j)	8.9	736.9	+32	proposed: double hydroxylation
double oxidation (4k)	9.3	736.9	+32	proposed: double hydroxylation
double oxidation (4l)	10.4	736.9	+32	proposed: double hydroxylation
double oxidation (4m)	11.8	736.9	+32	proposed: double hydroxylation
double oxidation (4n)	12.4	736.9	+32	proposed: double hydroxylation
double oxidation (4o)	13.6	736.9	+32	proposed: double hydroxylation
double oxidation (4p)	14.9	736.9	+32	proposed: double hydroxylation
double oxidation (4q)	16.5	736.9	+32	proposed: double hydroxylation

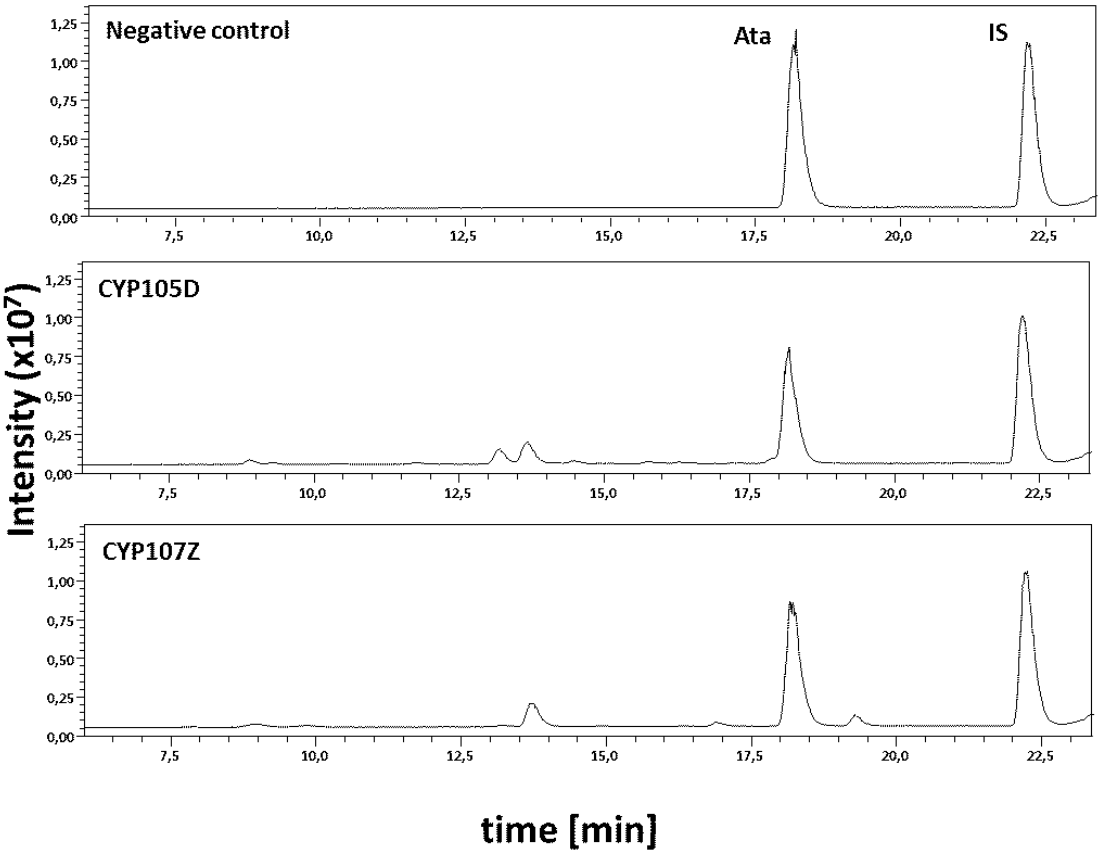


Figure S11: LC/MS-chromatograms for the conversion of atazanavir (4). Ata: Atazanavir, IS: internal standard ritonavir.

Table S12: Product analysis of saquinavir (5). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Saquinavir (5)	19.0	670.8	0	
single oxidation (5a)	12.6	686.8	+16	proposed: hydroxylation
single oxidation (5b)	13.5	686.8	+16	proposed: hydroxylation
single oxidation (5c)	14.6	686.8	+16	proposed: hydroxylation
single oxidation (5d)	15.4	686.8	+16	proposed: hydroxylation
single oxidation (5e)	16.9	686.8	+16	proposed: hydroxylation
single oxidation (5f)	18.6	686.8	+16	proposed: hydroxylation
double oxidation (5g)	5.9	702.8	+32	proposed: double hydroxylation
double oxidation (5h)	6.9	702.8	+32	proposed: double hydroxylation
double oxidation (5i)	8.0	702.8	+32	proposed: double hydroxylation
double oxidation (5j)	8.8	702.8	+32	proposed: double hydroxylation
double oxidation (5k)	9.7	702.8	+32	proposed: double hydroxylation
double oxidation (5l)	11.2	702.8	+32	proposed: double hydroxylation
double oxidation (5m)	13.2	702.8	+32	proposed: double hydroxylation
double oxidation (5n)	14.1	702.8	+32	proposed: double hydroxylation
double oxidation (5o)	16.3	702.8	+32	proposed: double hydroxylation

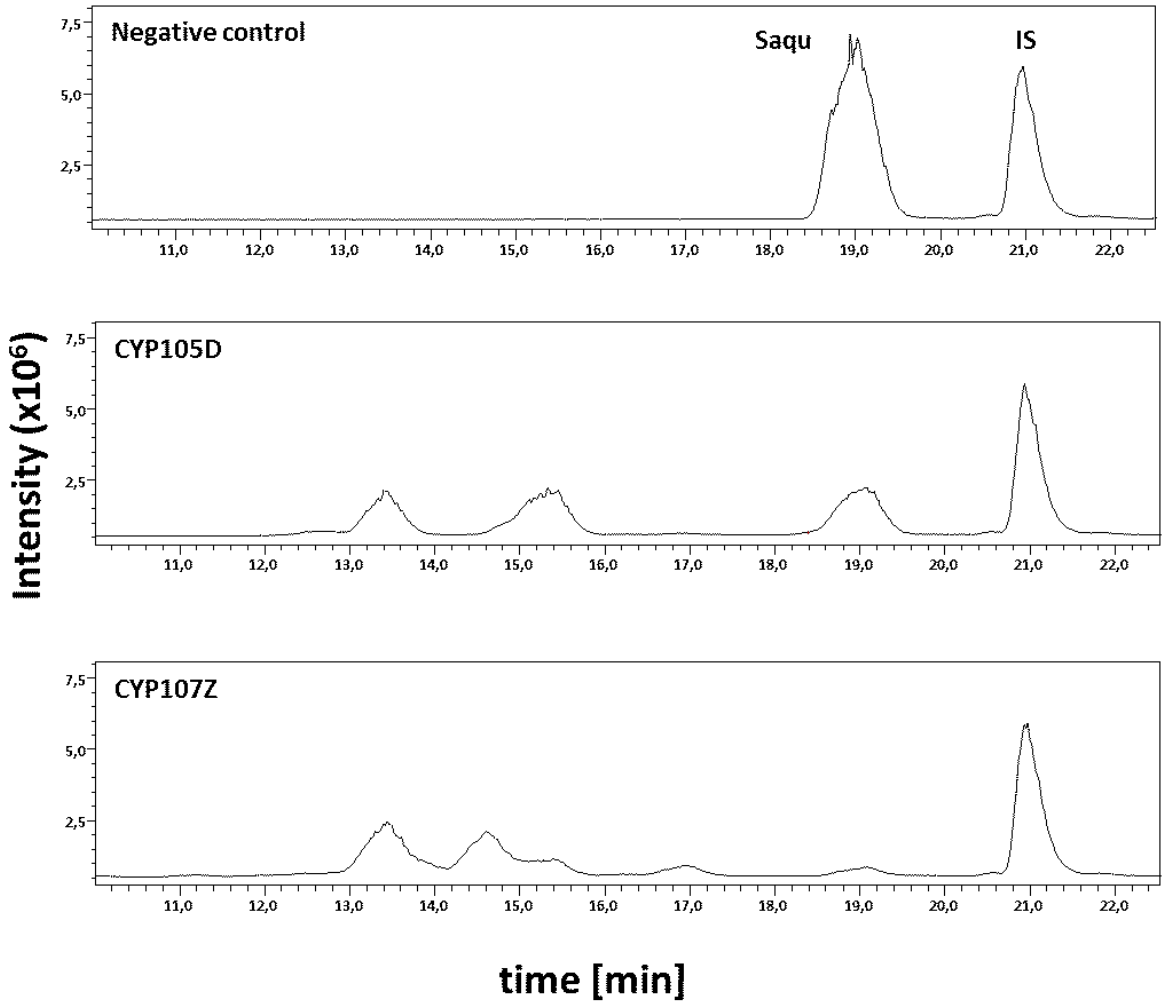


Figure S12: LC/MS-chromatograms for the conversion of saquinavir (5). Saqu: Saquinavir, IS: internal standard rupintrivir.

Table S13: Product analysis of indinavir (6). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Indinavir (6)	9.5	613.8	0	
single oxidation (6a)	4.5	629.8	+16	proposed: hydroxylation
single oxidation (6b)	5.4	629.8	+16	proposed: hydroxylation
single oxidation (6c)	6.5	629.8	+16	proposed: hydroxylation
single oxidation (6d)	10.5	629.8	+16	proposed: hydroxylation
single oxidation (6e)	12.1	629.8	+16	proposed: hydroxylation
dealkylation (6f)	10.2	522.7	-91	proposed: N-depyridomethylation

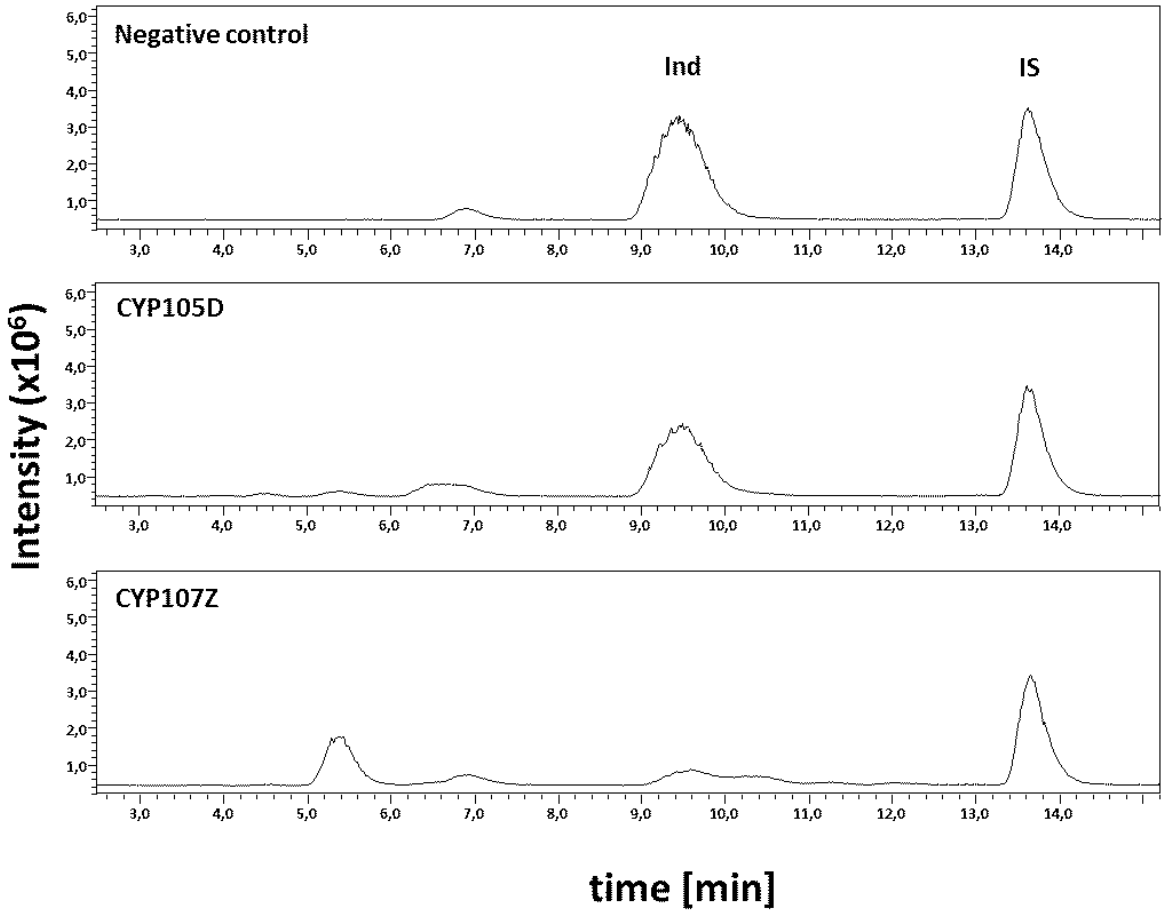


Figure S13: LC/MS-chromatograms for the conversion of indinavir (6). Ind: Indinavir, IS: internal standard amprenavir.

Steroids and glucocorticoids

Table S14: Product analysis of testosterone (7). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product	reference
Testosterone (7)	12.2	288.4	0		
6β-Hydroxytestosterone (7a)	10.0	304.4	+16	hydroxylation at C-6 in β-position	commercial standard
2β-Hydroxytestosterone (7b)	11.3	304.4	+16	hydroxylation at C-2 in β-position, main product of SPL_00625	commercial standard
16β-Hydroxytestosterone (7c)	10.9	304.4	+16	hydroxylation at C-16 in β-position	commercial standard
4-Androstene-3,17-dione (7d)	11.8	286.4	-2	oxidation to ketone at C-17, main product of AAU_01848	commercial standard
Hydroxylation product (7e)	9.6	304.4	+16	hydroxylation at unknown position, main product of SPL_06346	
Hydroxylation product (7f)	9.8	304.4	+16	hydroxylation at unknown position	
Hydroxylation product (7g)	10.3	304.4	+16	hydroxylation at unknown position	
Hydroxylation product (7h)	10.8	304.4	+16	hydroxylation at unknown position	
Hydroxylation product (7i)	11.0	304.4	+16	hydroxylation at unknown position	
Oxidation product (7j)	11.7	286.4	-2	proposed: oxidation to ketone	
Double oxidation product (7k)	7.3	320.4	+32	proposed: double hydroxylation	
Double oxidation product (7l)	8.2	320.4	+32	proposed: double hydroxylation	
Double oxidation product (7m)	9.3	320.4	+32	proposed: double hydroxylation	
Double oxidation product (7n)	9.8	320.4	+32	proposed: double hydroxylation	
Double oxidation product (7m)	10.5	320.4	+32	proposed: double hydroxylation	
Other oxidation products (unidentified)	**Retention time ($[M+H]^+$; Δm):** 8.0 min (m/z 375;+**87**); 9.9 min (m/z=303;**+15**) 10.1 min (m/z=303;**+15**);10.2 min (m/z =303; **+15**)				

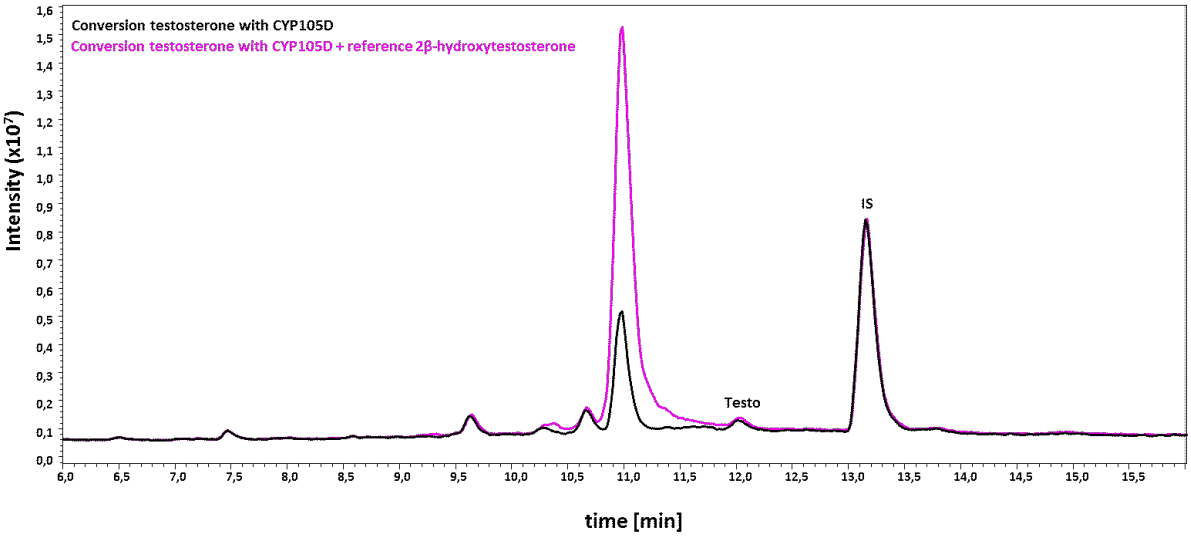


Figure S14: Identification of 2β-hydroxytestosterone as major product of testosterone oxidation by CYP105D via spiking of an authentic reference compound. Testo: Testosterone, IS: internal standard progesterone. Black: conversion of testosterone with CYP105D, pink: conversion of testosterone with CYP105D spiked with 2β-hydroxytestosterone.

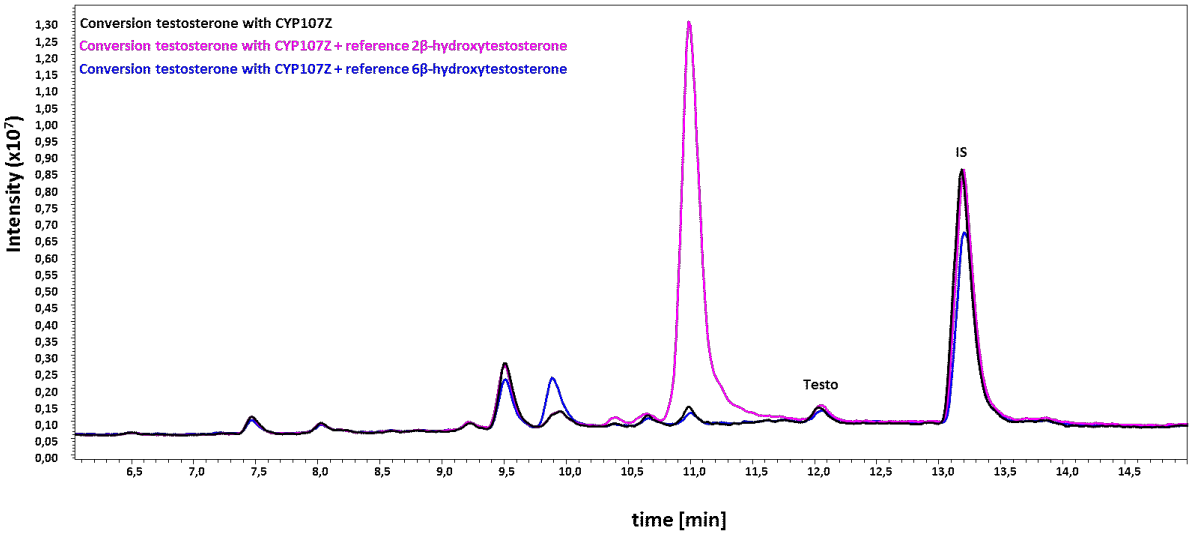


Figure S15: Identification of 2β-hydroxytestosterone and 6β-hydroxytestosterone as products of testosterone oxidation by CYP107Z via spiking of an authentic reference compound. Testo: Testosterone, IS: internal standard progesterone. Black: conversion of testosterone with CYP107Z, pink: conversion of testosterone with CYP107Z spiked with 2β-hydroxytestosterone, blue: conversion of testosterone with CYP107Z spiked with 6β-hydroxytestosterone.

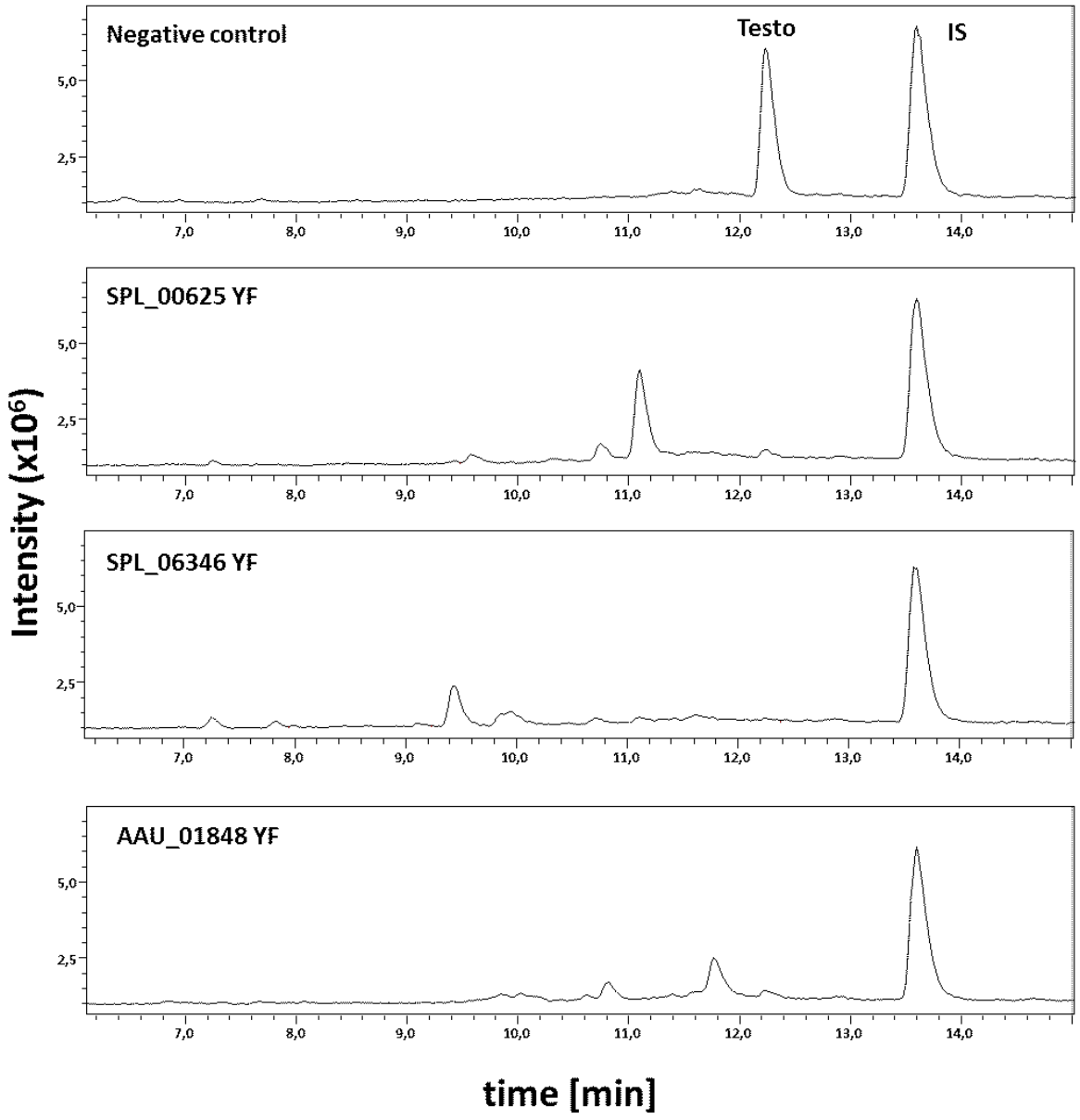


Figure S16: LC/MS-chromatograms for the conversion of testosterone (7) in the initial screening. Testo: Testosterone, IS: internal standard progesterone.

Table S15: Product analysis of progesterone (8). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Progesterone (8)	13.8	314.5	0	
Hydroxylation product (8a)	11.0	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8b)	11.2	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8c)	11.3	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8d)	11.7	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8e)	11.8	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8f)	12.0	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8g)	12.1	330.5	+16	hydroxylation at unknown position
Hydroxylation product (8h)	12.2	330.5	+16	hydroxylation at unknown position, main product of SPL_00625
Hydroxylation product (8i)	12.3	330.5	+16	hydroxylation at unknown position, main product of SPL_06346
Hydroxylation product (8j)	12.6	330.5	+16	hydroxylation at unknown position
Double oxidation product (8k)	8.8	346.5	+32	proposed: double hydroxylation
Double oxidation product (8l)	9.0	346.5	+32	proposed: double hydroxylation
Double oxidation product (8m)	9.4	346.5	+32	proposed: double hydroxylation
Double oxidation product (8n)	9.7	346.5	+32	proposed: double hydroxylation
Double oxidation product (8o)	10.0	346.5	+32	proposed: double hydroxylation
Double oxidation product (8p)	10.1	346.5	+32	proposed: double hydroxylation
Double oxidation product (8q)	10.5	346.5	+32	proposed: double hydroxylation
Double oxidation product (8r)	10.6	346.5	+32	proposed: double hydroxylation
Double oxidation product (8s)	11.0	346.5	+32	proposed: double hydroxylation
Double oxidation product (8t)	11.1	346.5	+32	proposed: double hydroxylation
Double oxidation product (8u)	11.5	346.5	+32	proposed: double hydroxylation

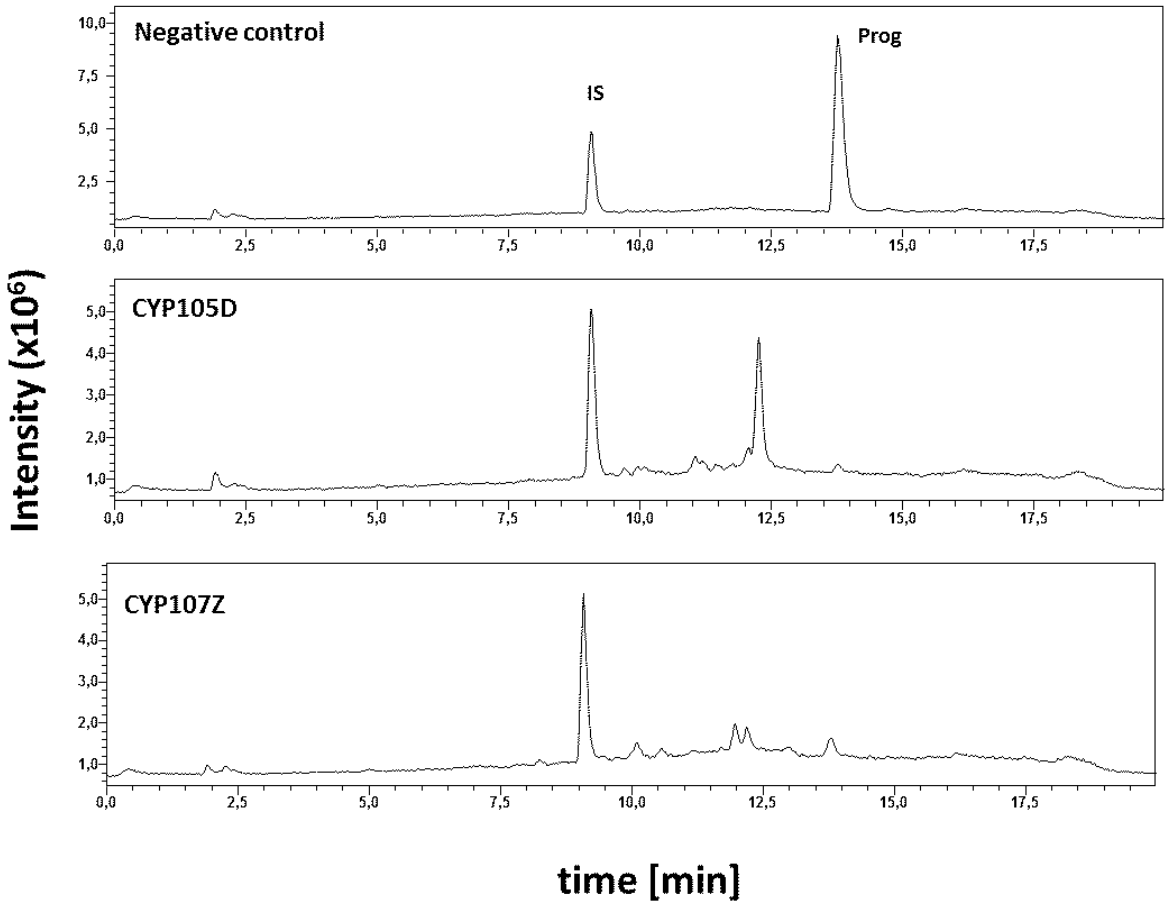


Figure S17: LC/MS-chromatograms for the conversion of progesterone (8). Prog: Progesterone, IS: internal standard triamcinolone.

Table S16: Product analysis of medrysone (9). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Medyrsone (9)	13.0	344.5	0	
Hydroxylation product (9a)	9.3	360.5	+16	hydroxylation at unknown position
Hydroxylation product (9b)	9.6	360.5	+16	hydroxylation at unknown position
Hydroxylation product (9c)	10.5	360.5	+16	hydroxylation at unknown position
Hydroxylation product (9d)	11.9	360.5	+16	hydroxylation at unknown position
Hydroxylation product (9e)	12.2	360.5	+16	hydroxylation at unknown position, main product of SPL_00625 and SPL_06346
Oxidation product (9f)	12.8	342.5	-2	proposed: oxidation to ketone
Double oxidation product (9g)	9.3	376.5	+32	proposed: double hydroxylation
Other oxidation products (unidentified)	**Retention time ($[M+H]^+$; Δm):** 10.5 min (m/z 360;+**15**); 11.0 min (m/z 344/362;-**2 or + 16**); 11.3 min (m/z 444, +**98**)			

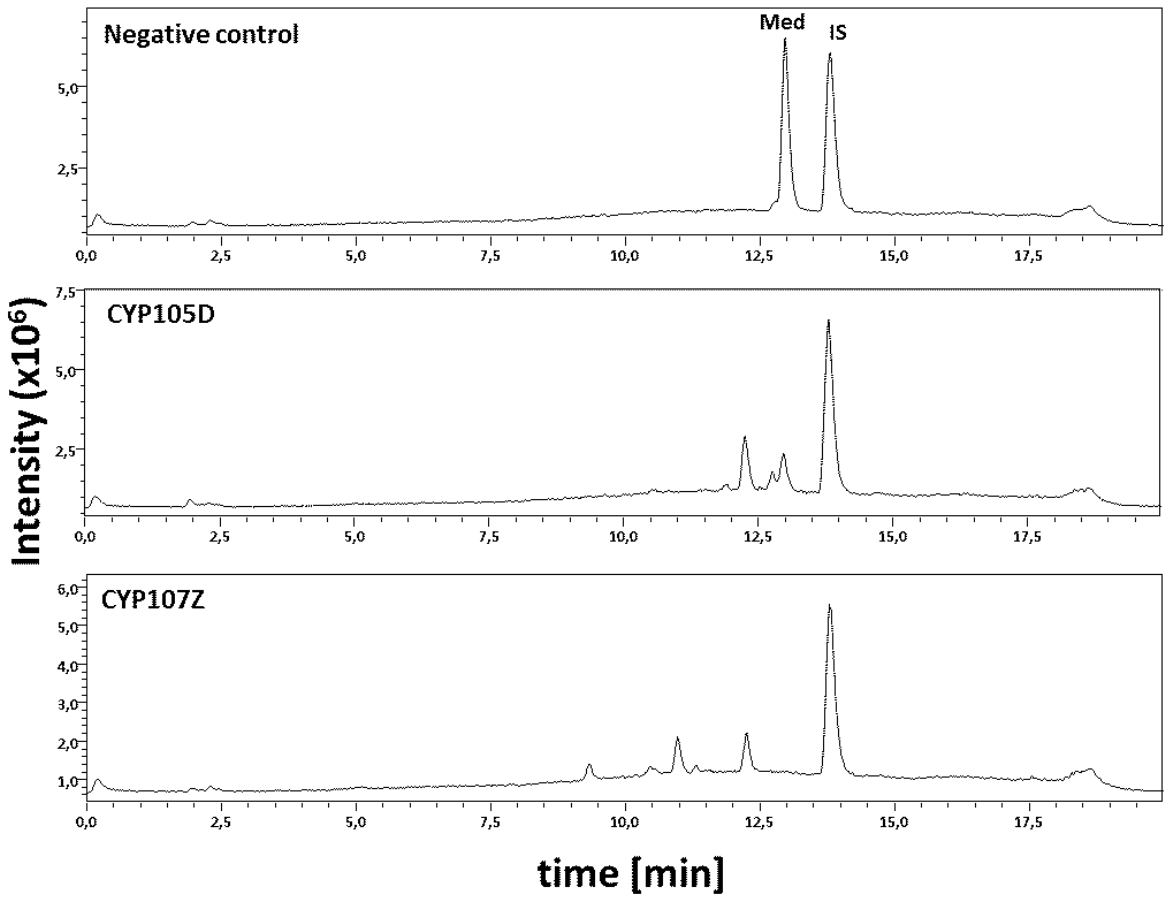


Figure S18: LC/MS-chromatograms for the conversion of medrysone (9). Med: Medrysone, IS: internal standard progesterone

Table S17: Product analysis of hydrocortisone (10). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Hydrocortisone (10)	10.5	362.5	0	
Hydroxylation product (10a)	9.9	378.5	+16	proposed: hydroxylation, main product of SPL_06346
Hydroxylation product (10b)	6.3	378.5	+16	proposed: hydroxylation

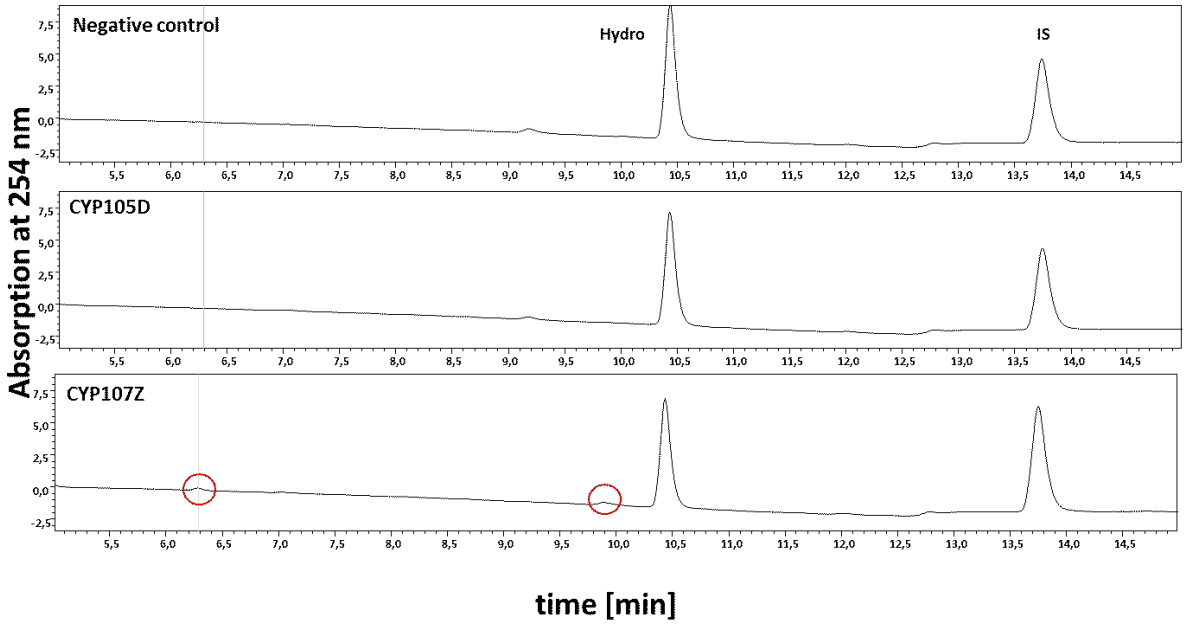


Figure S19: PDA-chromatograms for the conversion of hydrocortisone (10). Hydro: Hydrocortisone, IS: internal standard progesterone; proposed: products are highlighted with a red circle.

Table S18: Product analysis of prednisolone (11). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Prednisolone (11)	10.5	360.4	0	
Hydroxylation product (11a)	5.9	376.4	+16	proposed: hydroxylation, main product of SPL_06346
unknown compound (11b)	5.5	unknown		unknown product only visible in PDA
unknown compound (11c)	10.0	unknown		unknown product only visible in PDA

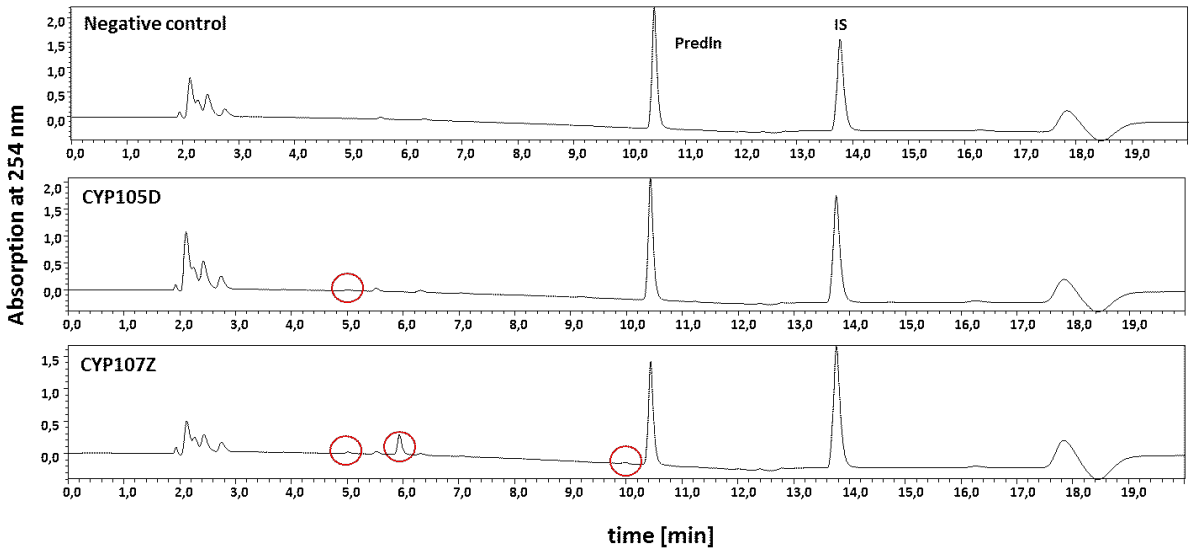


Figure S20: PDA-chromatograms for the conversion of prednisolone (11). Predn: Prednisolone, IS: internal standard progesterone; proposed products are highlighted with a red circle.

Table S19: Product analysis of prednisone (12). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Prednisone (12)	9.9	358.4	0	
Hydroxylation product (12a)	6.9	374.4	+16	proposed: hydroxylation, main product of SPL_06346

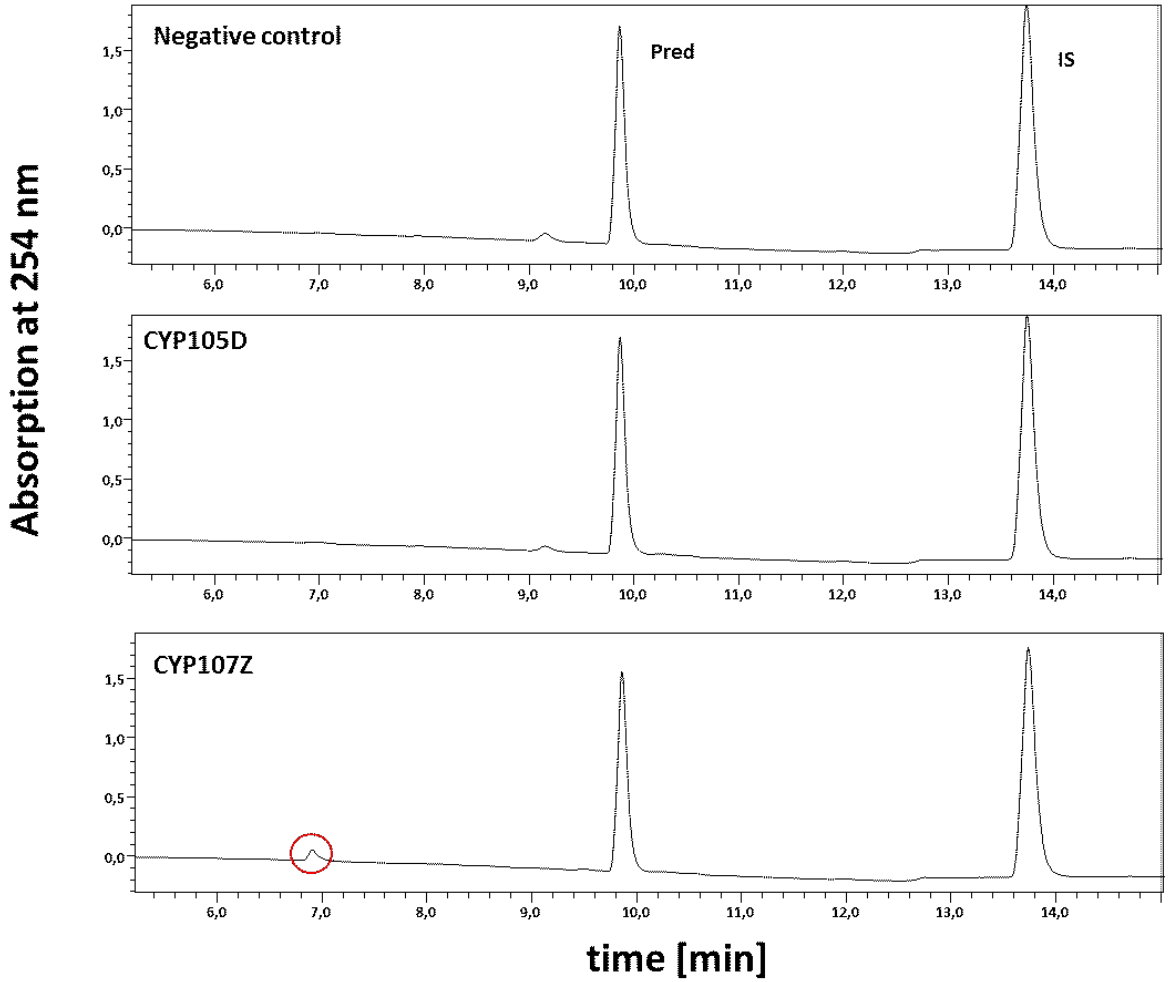


Figure S21: PDA-chromatograms for the conversion of prednisone (12). Pred: Prednisone, IS: internal standard progesterone; proposed: products are highlighted with a red circle.

Table S20: Product analysis of 6α-methylprednisolone (13). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
6α-Methylprednisolone (13)	11.3	374.5	0	
Hydroxylation product (13a)	6.7	390.5	+16	proposed: hydroxylation, main product of SPL_06346

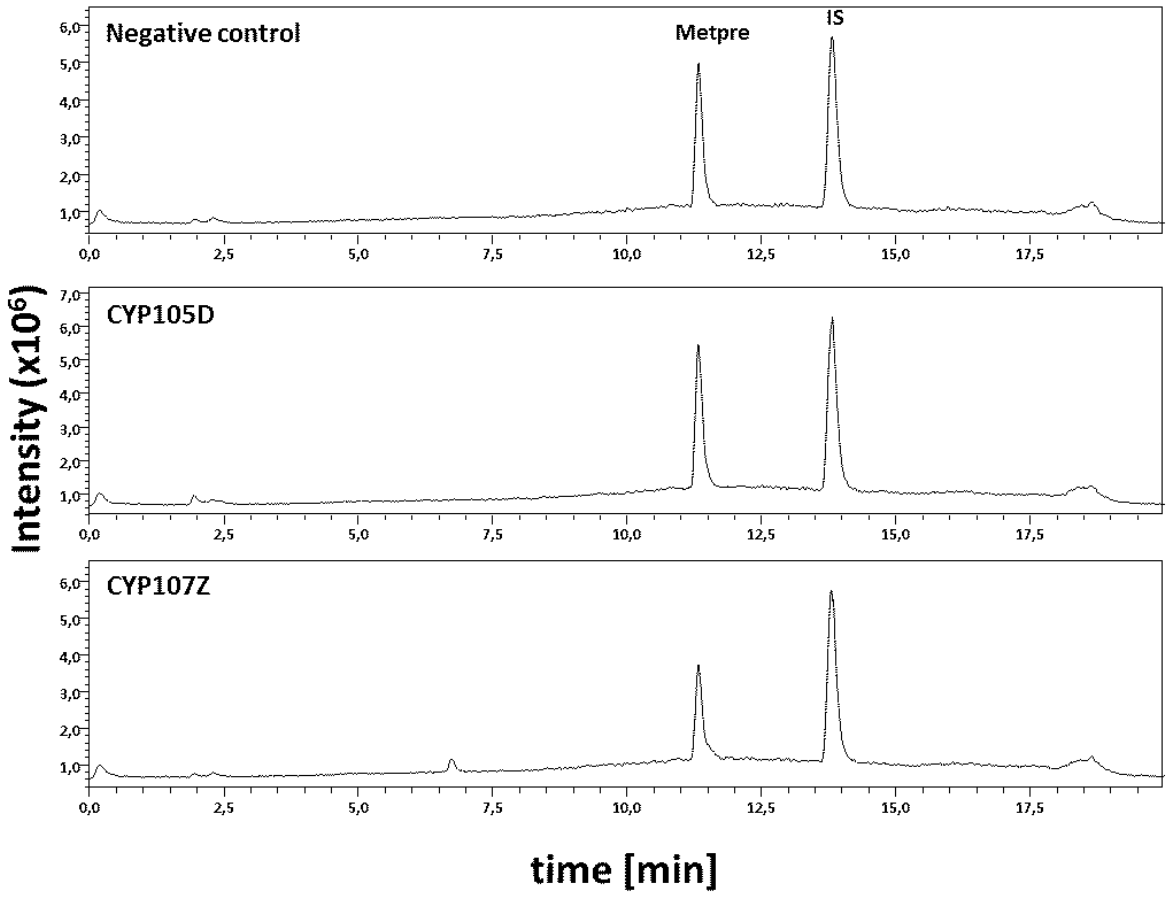


Figure S22: LC/MS-chromatograms for the conversion of 6α-methylprednisolone (13). Metpre: 6α-Methylprednisolone, IS: internal standard progesterone.

Table S21: Product analysis of dexamethasone (14). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product	reference
Dexanethasone (14)	11.2	392.4	0		
6β-hydroxydexamethasone (14a)	7.4	408.5	+16	hydroxylation at C-6 in β-position, main product of SPL_06346	commercial standard
unknown compound (14b)	5.0	unknown		unknown product only visible in PDA	

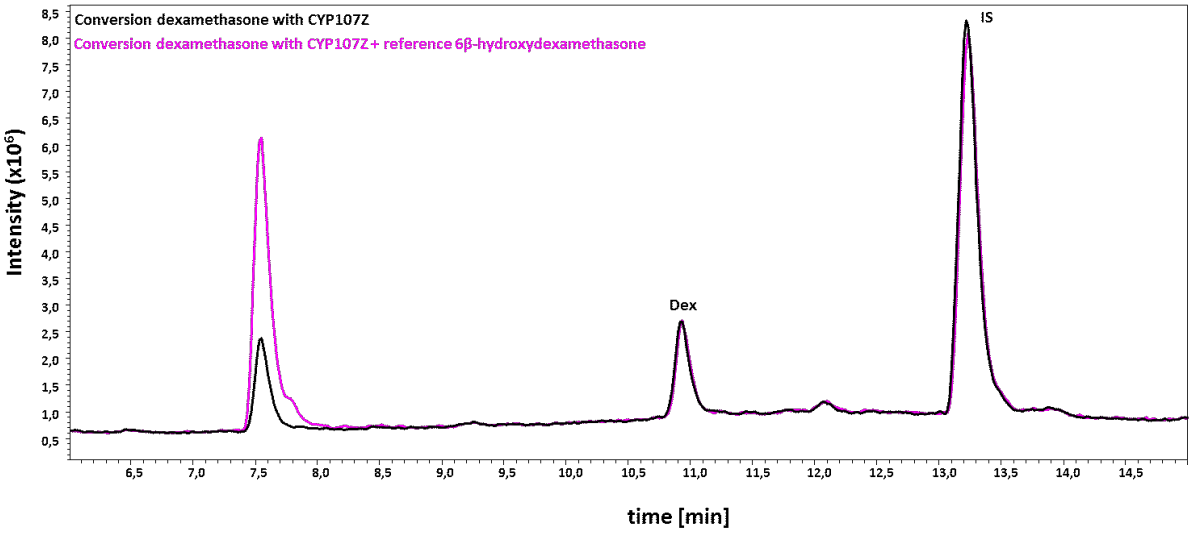


Figure S23: Identification of 6β-hydroxydexamethasone as major product of dexamethasone oxidation by CYP107Z via spiking with an authentic reference compound. Dex: Dexamethasone, IS: internal standard progesterone. Black: conversion of dexamethasone with CYP107Z, pink: conversion of dexamethasone with CYP107Z spiked with 6β-hydroxydexamethasone.

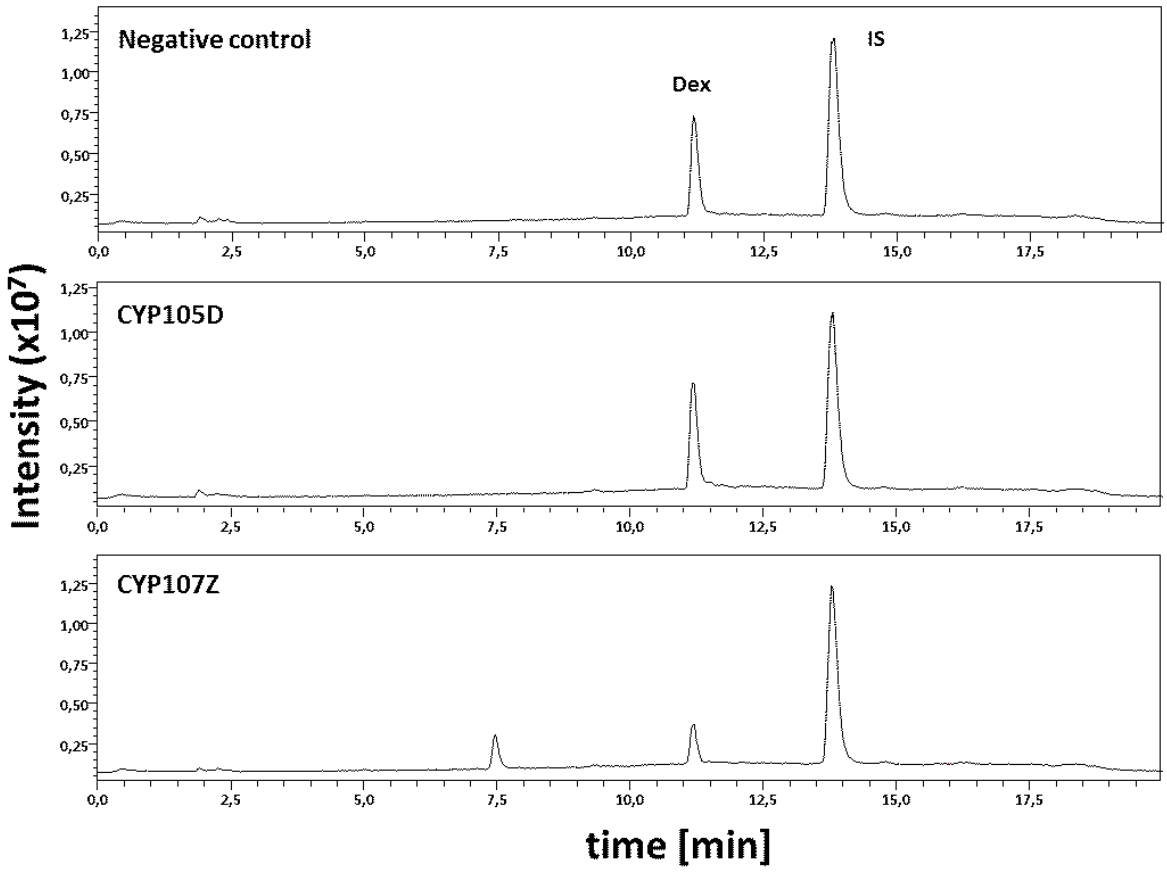


Figure S24: LC/MS-chromatograms for the conversion of dexamethasone (14). Dex: Dexamethasone, IS: internal standard progesterone.

Table S22: Product analysis of betamethasone (15). For unidentified products only the retention time and main m/z-value in LC/MS analysis is stated where possible.

Compound	Retention time [min]	Molecular weight[g/mol]	Δm	reaction/product
Betamethasone (15)	11.2	392.4	0	
Hydroxylation product (15a)	7.3	408.5	+16	proposed: hydroxylation, main product of SPL_06346
unknown compound (15b)	5.0	unknown		unknown product only visible in PDA

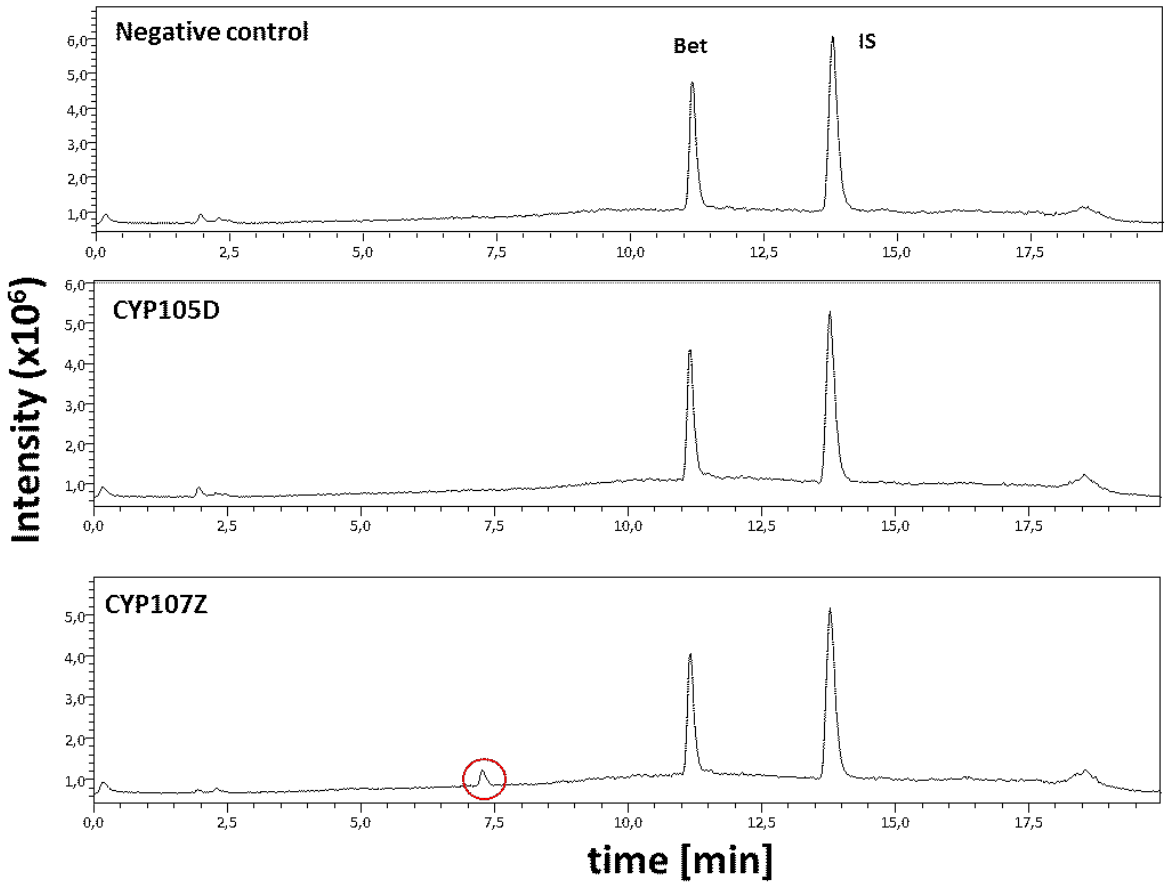


Figure S25: LC/MS-chromatograms for the conversion of betamethasone (15). Bet: Betamethasone, IS: internal standard progesterone.

Tricyclic antidepressants

Table S23: Product analysis of amitriptyline (16).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Amitriptyline (16)	18.3	277.4	0		
Nortriptyline (21)	17.4	263.4	-14	N-demethylation	commercial standard
(±)-*(E)*-10-Hydroxyamitriptyline (16a)	6.0	293.4	+16	hydroxylation at C-10	commercial standard
Desmethylnortriptyline (16b)	16.2	249.4	-28	2 x N-demethylation	commercial standard
Amitriptyline N-oxid (16c)	20.8	293.4	+16	N-oxidation at trialkyl moiety	commercial standard
single oxidation product (16d)	8.2	293.4	+16	proposed: hydroxylation	/
single oxidation product (16e)	12.9	293.4	+16		/
double oxidation product (16f)	5.7	279.4	+2	proposed: N-demethylation of **16a**	/
double oxidation product (16g)	7.7	279.4	+2	proposed: N-demethylation of **16d**	/
double oxidation product (16h)	21.1	279.4	+2	proposed: N-demethylation of **16c**	/

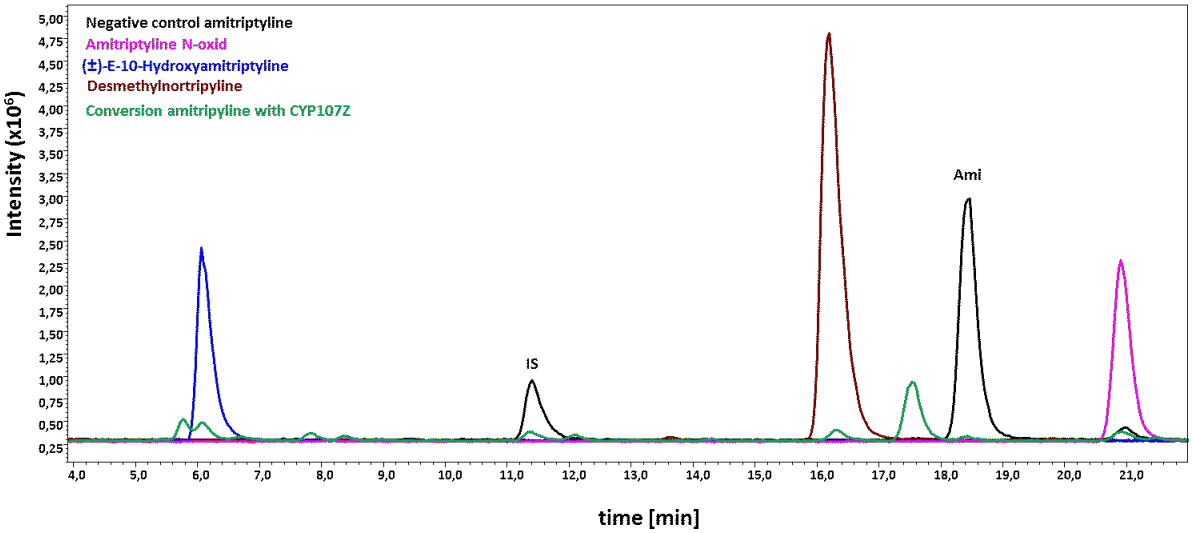


Figure S26: Comparison of commercial product standards with the conversion of amitriptyline (16). Ami: Amitriptyline, IS: internal standard oxcarbazepine Black: control reaction without P450, pink: amitriptyline N-oxid, blue: (±)-*E*-10-hydroxyamitriptyline, brown: desmethylnortriptyline, green: conversion of amitriptyline with CYP107Z.

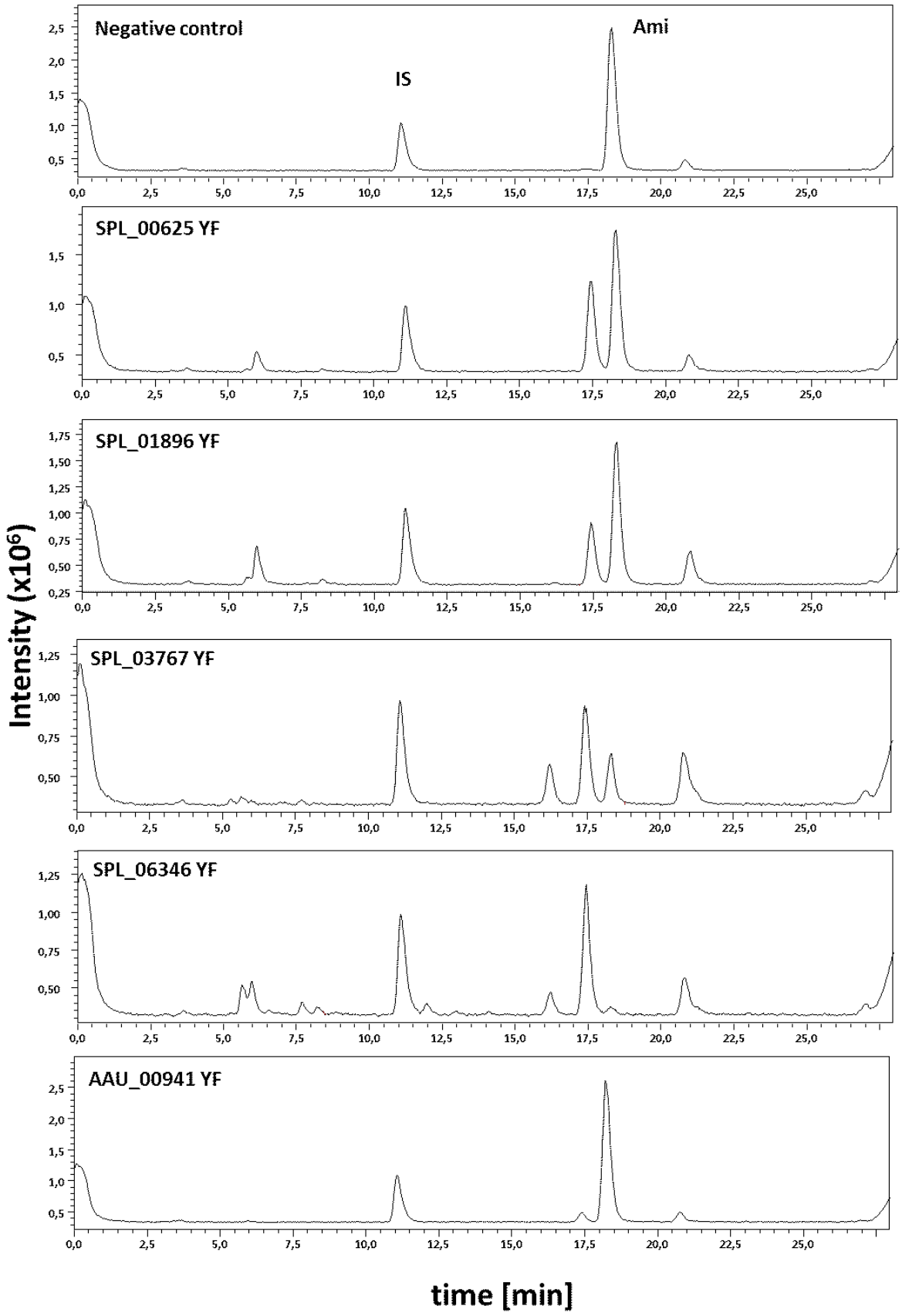


Figure S27: LC/MS-chromatograms for the conversion of amitriptyline (16) in the initial screening. Ami: Amitriptyline, IS: internal standard oxcarbazepine.

Table S24: Product analysis of imipramine (17).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Imipramine (17)	16.5	280.4	0	
demethylation product (17a)	15.7	266.4	-14	proposed: N-demethylation
single oxidation product (17b)	6.0	296.4	+16	proposed: aliphatic hydroxylation
didemethylation product (17c)	14.4	252.4	-28	proposed: 2 x N-demethylation
single oxidation product (17d)	8.1	296.4	+16	proposed: hydroxylation
double oxidation product (17e)	5.8	282.4	+2	proposed: N-demethylation of **17b**
double oxidation product (17f)	7.7	282.4	+2	proposed: N-demethylation of **17d**
double oxidation product (17g)	19.0	282.4	+2	proposed: N-demethylation of **N-oxid**

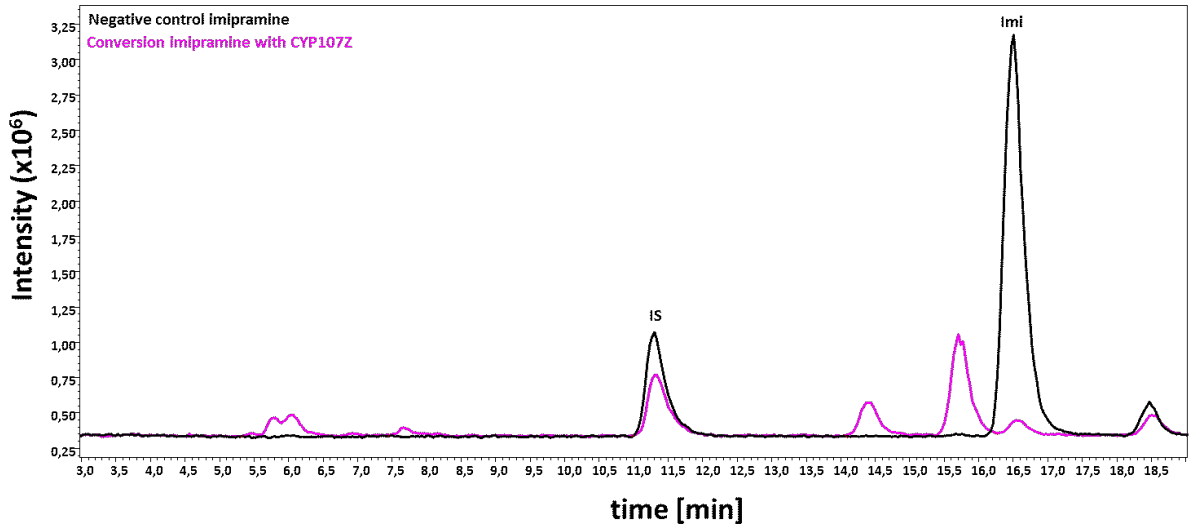


Figure S28: LC/MS-chromatograms for the conversion of imipramine (17). Imi: Imipramine, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of imipramine with CYP107Z.

Table S25: Product analysis of clomipramine (18).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Clomipramine (18)	14.7	314.9	0	
demethylation product (18a)	13.9	300.9	-14	proposed: N-demethylation
single oxidation product (18b)	6.3	330.9	+16	proposed: aliphatic hydroxylation
single oxidation product (18c)	6.5	330.9	+16	proposed: aliphatic hydroxylation
N-oxid (18d)	16.1	330.9	+16	proposed: N-oxid formation
didemethylation product (18e)	12.8	286.9	-28	proposed: 2 x N-demethylation
single oxidation product (18f)	7.0	330.9	+16	proposed: hydroxylation
single oxidation product (18g)	8.2	330.9	+16	proposed: hydroxylation
double oxidation product (18h)	6.0	316.9	+2	proposed: N-demethylation of **18b**
double oxidation product (18i)	6.3	316.9	+2	proposed: N-demethylation of **18c**
double oxidation product (18j)	6.6	316.9	+2	proposed: N-demethylation of **18f**
double oxidation product (18k)	7.8	316.9	+2	proposed: N-demethylation of **18g**

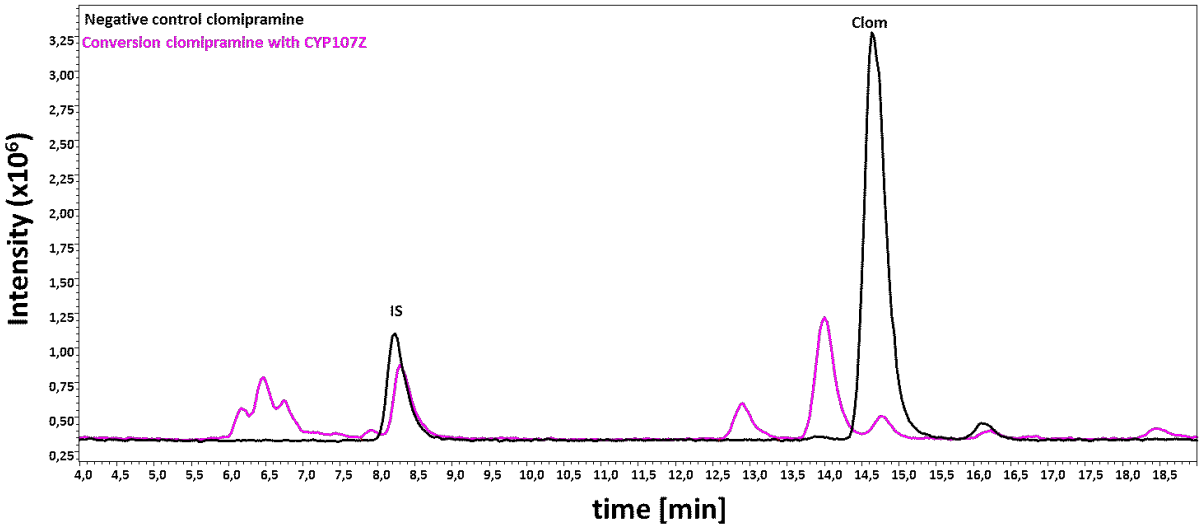


Figure S29: LC/MS-chromatograms for the conversion of clomipramine (18). Clom: Clomipramine, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of clomipramine with CYP107Z.

Table S26: Product analysis of chlorpromazine (19).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Chlorpromazine (19)	20.9	318.9	0		
demethylation product (19a)	20.0	304.9	-14	proposed: N-demethylation	/
Chlorpromazine sulfoxide (19b)	5.8	334.9	+16	sulfoxidation	commercial standard
didemethylation product (19c)	18.6	290.9	-28	proposed: 2 x N-demethylation	/
double oxidation product (19d)	5.5	320.9	+2	proposed: N-demethylation of **19b**	/
double oxidation product (19e)	6.8	350.9	+32	proposed: 2x single oxidation	/
Unknown oxidation (19f)	15.0	307.9	-11	/	/

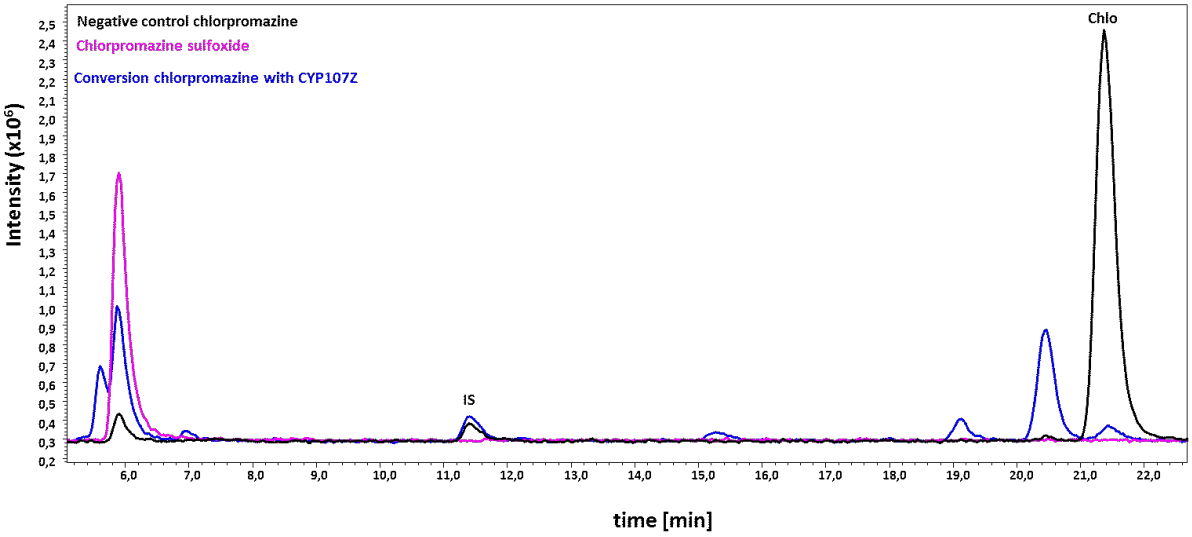


Figure S30: LC/MS-chromatograms for the conversion of chlorpromazine (19). Chlo: Chlorpromazine, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: chlorpromazine sulfoxide, blue: conversion of chlorpromazine with CYP107Z.

Table S27: Product analysis of cyclobenzaprine (20).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Cyclobenzaprine (20)	16.7	275.4	0	
demethylation product (20a)	15.9	261.4	-14	proposed: N-demethylation
didemethylation product (20b)	14.7	247.4	-28	proposed: 2 x N-demethylation
N-oxid (20c)	19.2	292.6	+16	proposed: N-oxid formation
Unknown oxidation (20d)	10.9	277.4	+2	/
Unknown oxidation (20e)	11.2	277.4	+2	/

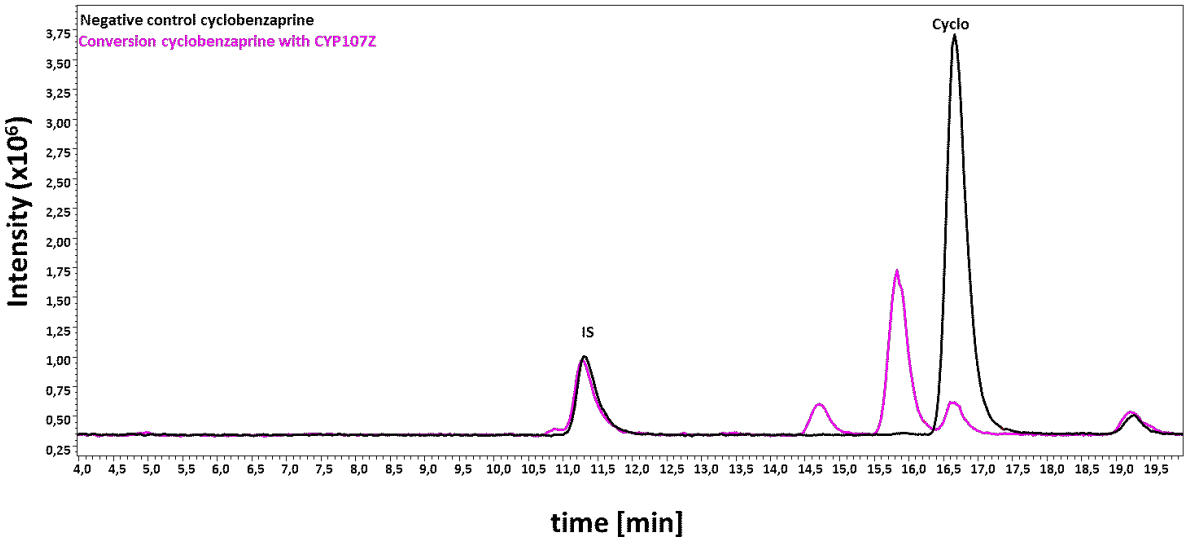


Figure S31: LC/MS-chromatograms for the conversion of cyclobenzaprine (20). Cyclo: Cyclobenzaprine, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of cyclobenzaprine with CYP107Z.

Table S28: Product analysis of nortriptyline (21).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Nortriptyline (21)	17.4	263.4	0		
desmethylnortriptyline (21a or 16b)	16.2	249.4	-14	N-demethylation	commercial standard
single oxidation product (21b or 16f)	5.7	279.4	+16	proposed: hydroxylation at C10	/
N-oxid (21c or 16h)	21.1	279.4	+16	proposed: N-oxid formation	/
single oxidation product (21d)	7.7	279.4	+16	proposed: hydroxylation	
single oxidation product (21e)	12.0	279.4	+16		/
double oxidation product (21f)	5.4	265.4	+2	proposed: N-demethylation of **21b**	/
double oxidation product (21g)	7.1	265.4	+2	proposed: N-demethylation of **21d**	

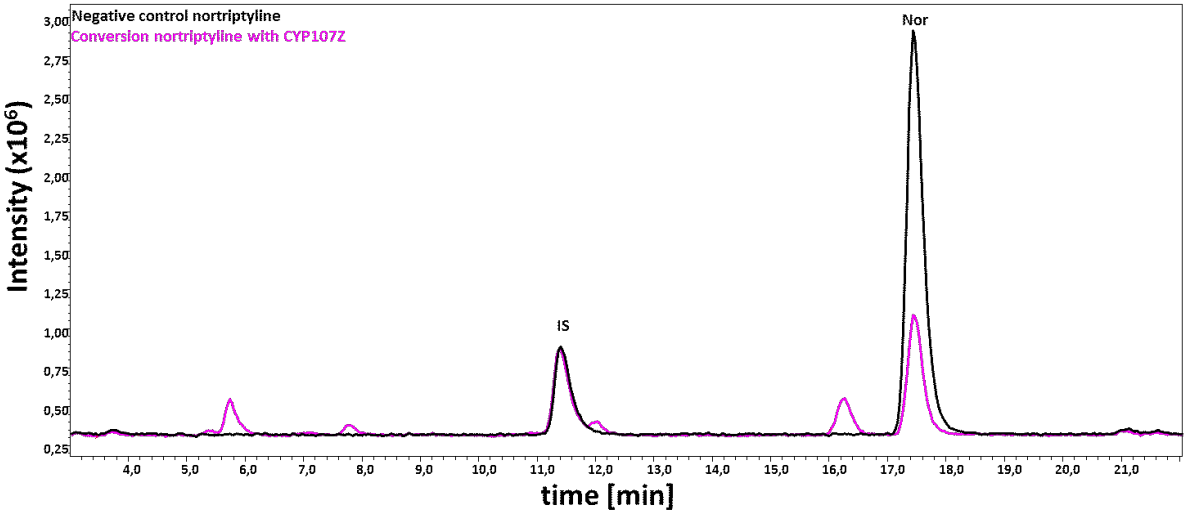


Figure S32: LC/MS-chromatograms for the conversion of nortriptyline (21). Nor: Nortriptyline, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of nortriptyline with CYP107Z.

Table S29: Product analysis of protriptyline (22).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Protriptyline (22)	16.1	263.4	0	
demethylation product (22a)	15.0	249.4	-14	proposed: N-demethylation
single oxidation product (22b)	10.2	279.4	+16	proposed: hydroxylation
N-oxid (22c)	19.2	279.4	+16	proposed: N-oxid formation

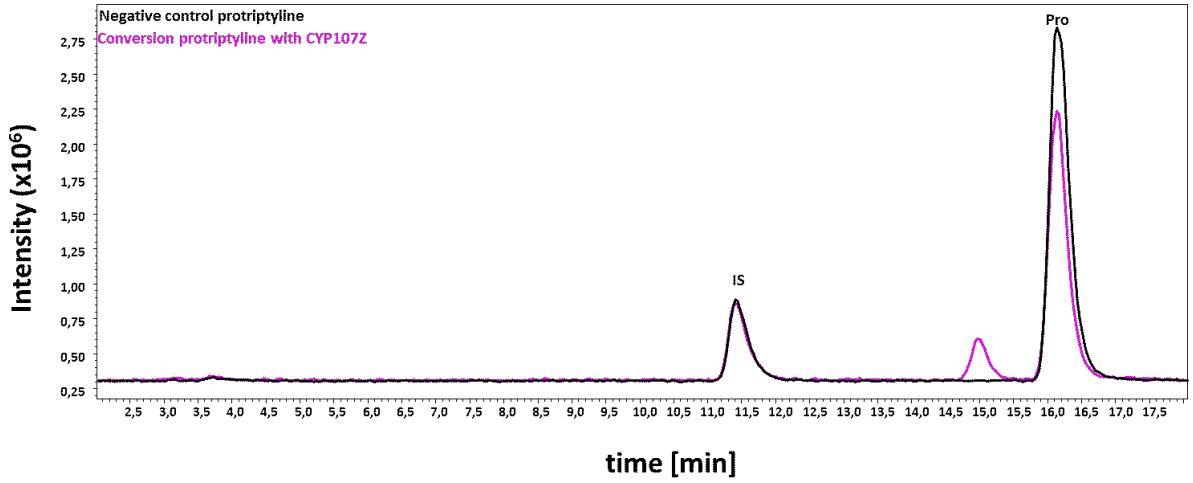


Figure S33: LC/MS-chromatograms for the conversion of protriptyline (22). Pro: Protriptyline, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of protriptyline with CYP107Z.

Table S30: Product analysis of carbamazepine (23).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Carbamazepine (23)	16.1	236.3	0		

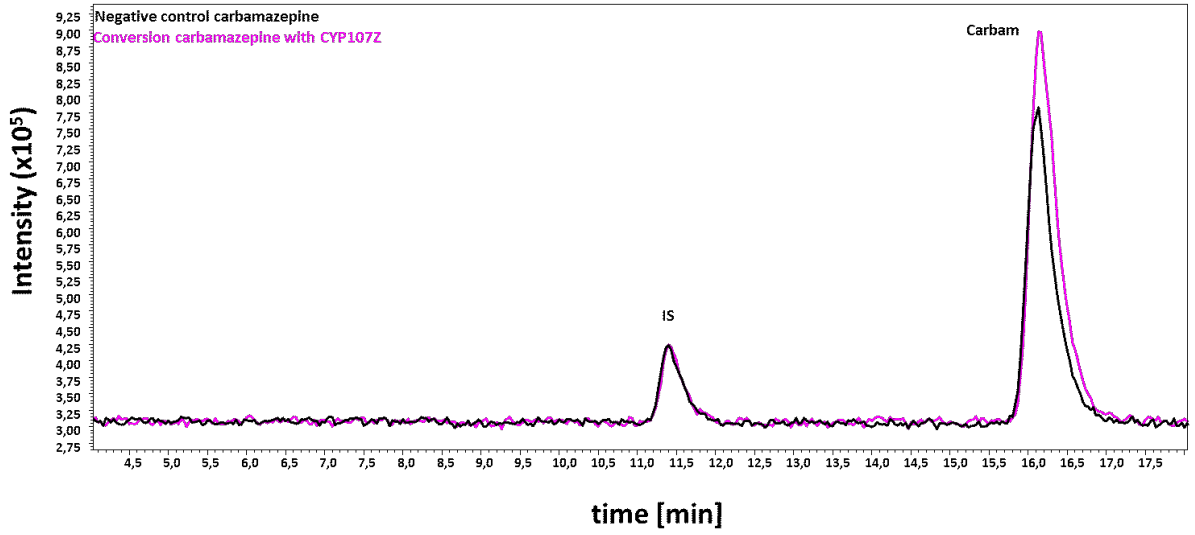


Figure S34: LC/MS-chromatograms for the conversion of carbamazepine (23). Carbam: Carbamazepine, IS: internal standard oxcarbazepine. Black: control reaction without P450, pink: conversion of carbamazepine with CYP107Z

Table S31: Product analysis of oxcarbazepine (24).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product	reference
Oxcarbazepine (24)	11.4	252.3	0		

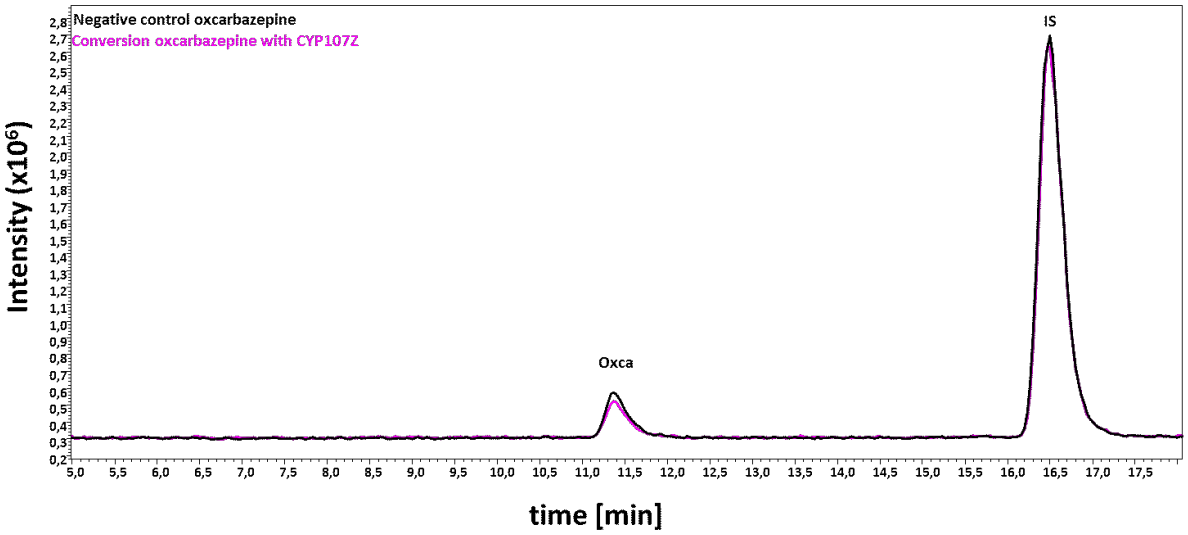


Figure S35: LC/MS-chromatograms for the conversion of oxcarbazepine (24). Oxcarba: Oxcarbazepine, IS: internal standard amitriptyline. Black: control reaction without P450, pink: conversion of oxcarbazepine with CYP107Z.

Table S32: Product analysis of opipramol (25).

Compound	Retention time [min]	Molecular weight [g/mol]	Δm	reaction/product
Opipramol (25)	7.6	363.5	0	
Dealkylation product (25a)	5.6	319.4	-44	proposed N-dealkylation
Single oxidation product (25b)	4.4	379.5	+16	unknown single oxidation
Single oxidation product (25c)	9.0	379.5	+16	unknown single oxidation
Other oxidation products (unidentified)	**Retention time ($[M+H]^+$; Δm):**3.9 min (m/z 337; **-28**); 4.5 min (m/z 295; **-70**), 5.4 min (m/z 337; **-28**), 10.4 min (m/z 379; **+14**), 12.3 min (m/z 337; **-28**), 12.7 min (m/z 363; **±0**), 14.2 min (m/z 379; **+14**), 15.1 min (m/z 379; **+14**), 15.9 min (m/z 305; **-60**)			

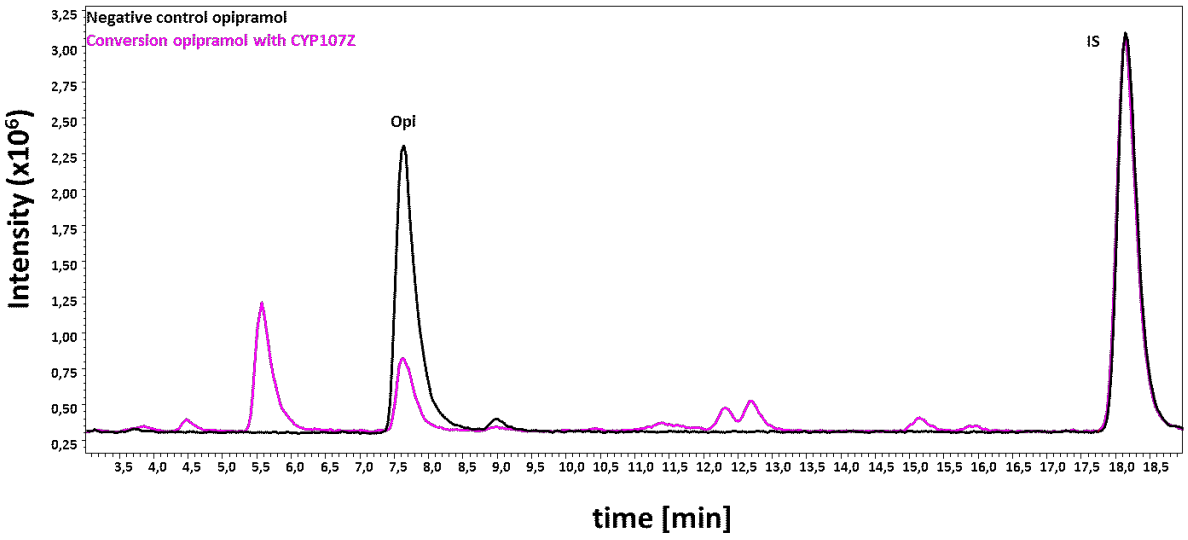
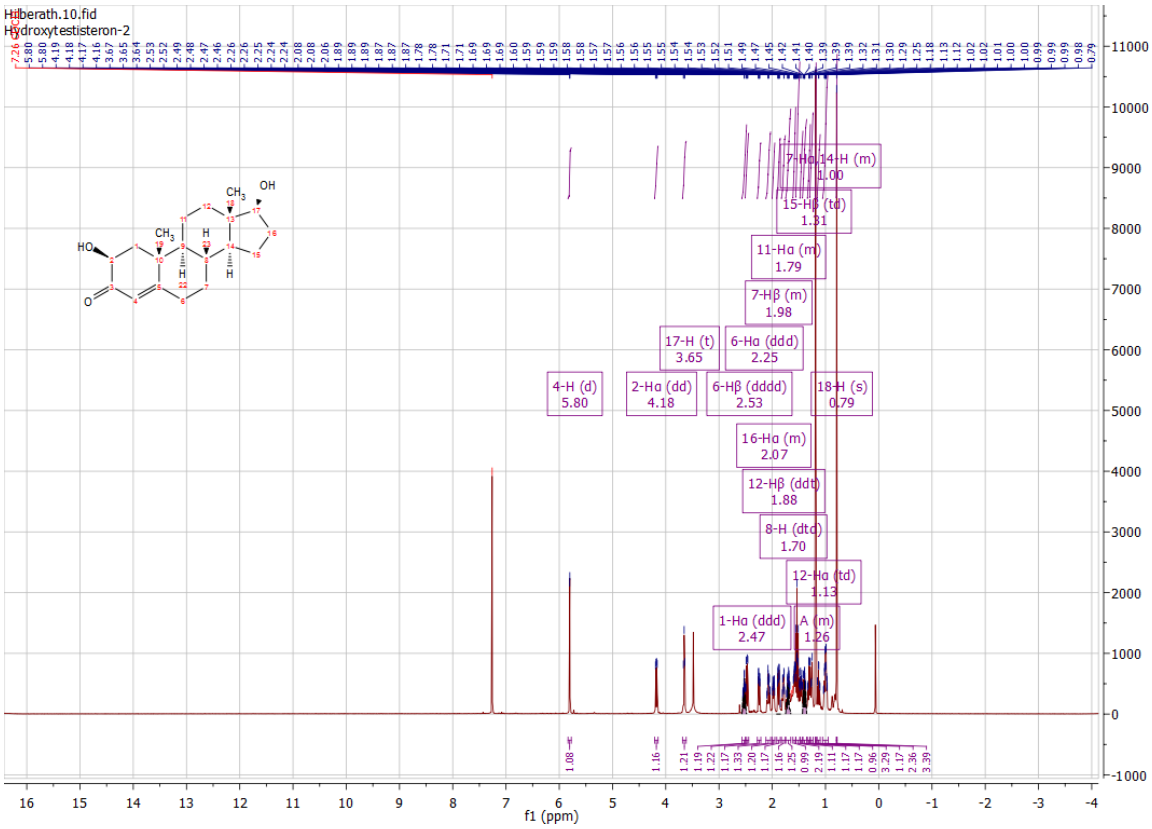


Figure S36: LC/MS-chromatograms for the conversion of opipramol (25). Opi: Opipramol, IS: internal standard amitriptyline. Black: control reaction without P450, pink: conversion of oxcarbazepine with CYP107Z.

NMR-data



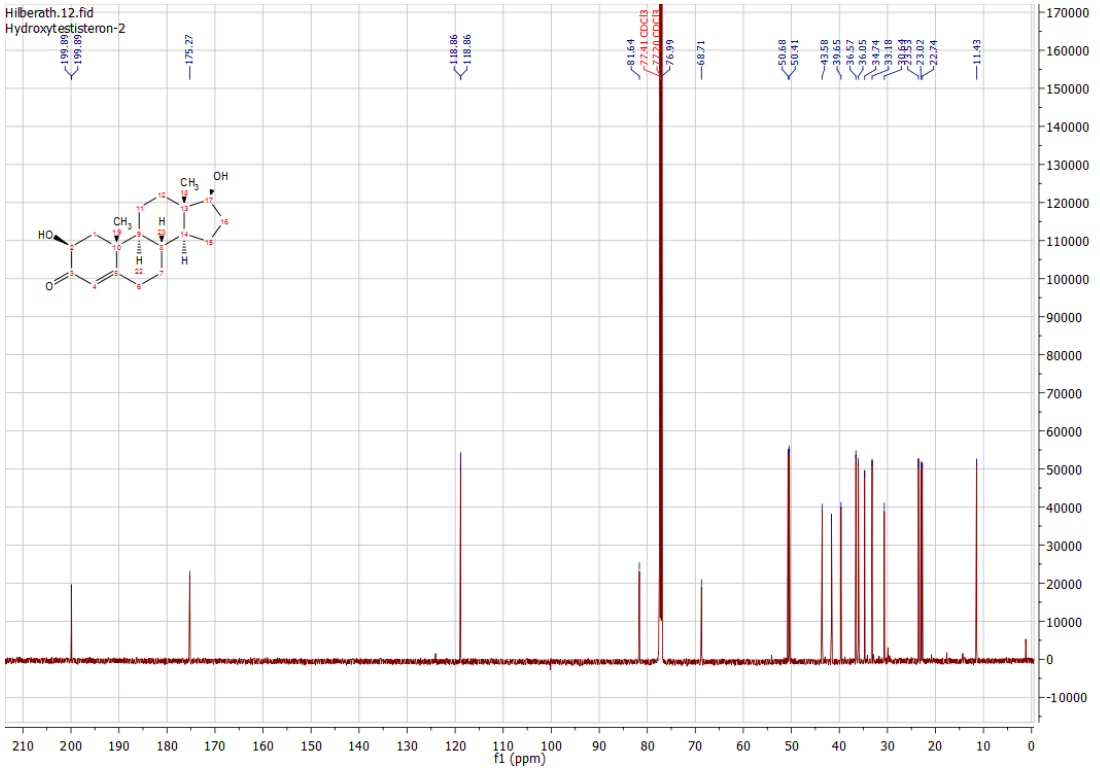


Figure S37: ^{1}**H-NMR (600 MHz) and** ^{13}**C-NMR (151 MHz) of 2β-hydroxytestosterone (7b) in** $CDCl_3$**.**

References

[1] M. Girhard, T. Klaus, Y. Khatri, R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2010**, *87*, 595-607.

[2] P. J. Bakkes, S. Biemann, A. Bokel, M. Eickholt, M. Girhard, V. B. Urlacher, *Sci. Rep.* **2015**, *5*, 12158.

[3] M. Girhard, K. Machida, M. Itoh, R. D. Schmid, A. Arisawa, V. B. Urlacher, *Microb. Cell Fact.* **2009**, *8*, 36.

[4] P. Le-Huu, T. Heidt, B. Claasen, S. Laschat, V. B. Urlacher, *ACS Catal.* **2015**, *5*, 1772-1780.

[5] a) D. Sirim, F. Wagner, A. Lisitsa, J. Pleiss, *BMC Biochem.* **2009**, *10*, 27; b) D. R. Nelson, *Hum. Genomics* **2009**, *4*, 59-65.

[6] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, *Mol. Syst. Biol.* **2011**, *7*, 539.

[7] A. Worsch, F. K. Eggimann, M. Girhard, C. J. von Bühler, F. Tieves, R. Czaja, A. Vogel, C. Grumaz, K. Sohn, S. Lütz, M. Kittelmann, V. B. Urlacher, *Biotechnol. Bioeng.* **2018**, *115*, 2156-2166.

[8] a) T. Nagao, T. Mitamura, X. H. Wang, S. Negoro, T. Yomo, I. Urabe, H. Okada, *J. Bacteriol.* **1992**, *174*, 5013-5020; b) S. Kranz-Finger, O. Mahmoud, E. Ricklefs, N. Ditz, P. J. Bakkes, V. B. Urlacher, *Biochim. Biophys. Acta, Proteins Proteomics* **2018**, *1866*, 2-10.

[9] M. M. Purdy, L. S. Koo, P. R. Ortiz de Montellano, J. P. Klinman, *Biochemistry* **2004**, *43*, 271-281.

[10] a) L. McIver, C. Leadbeater, D. J. Campopiano, R. L. Baxter, S. N. Daff, S. K. Chapman, A. W. Munro, *Eur. J. Biochem.* **1998**, *257*, 577-585; b) Z.-Q. Wang, R. J. Lawson, M. R. Buddha, C.-C. Wei, B. R. Crane, A. W. Munro, D. J. Stuehr, *J. Biol. Chem.* **2007**, *282*, 2196-2202.

[11] C. von Bühler, P. Le-Huu, V. B. Urlacher, *ChemBioChem* **2013**, *14*, 2189-2198.

2.3 The cytochrome P450 complement (CYPome) of *Streptomyces platensis*

Title: The cytochrome P450 complement (CYPome) of *Streptomyces platensis* DSM40041

Authors: Thomas Hilberath[+], Anne Worsch[+], Priska Le-Huu, Christian Grumaz, Kai Sohn, Marco Girhard, Vlada B. Urlacher*

[+]both authors contributed equally; *corresponding author

Manuscript in preparation

Own contribution: Conceptualization, design and conduction of all experiments with 18 P450s from *Streptomyces platensis*, analysis and interpretation of all data together with Anne Worsch, drafting of the manuscript together with Anne Worsch. Relative contribution: 40%.

The cytochrome P450 complement (CYPome) of *Streptomyces platensis* DSM40041

Thomas Hilberath[1+], Anne Worsch[1+], Priska Le-Huu[1], Christian Grumaz[2], Kai Sohn[2], Marco Girhard[1], Vlada B. Urlacher[1,*]

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

[2]Fraunhofer Institute for Interfacial Engineering and Biotechnology, Department of Molecular Biotechnology, 70569 Stuttgart, Germany

* Corresponding author: Vlada B. Urlacher. Email: vlada.urlacher@uni-duesseldorf.de

[+] Equal contribution

Abstract

The systematic analysis of the cytochrome P450 complement (CYPome) of *Streptomyces platensis* DSM 40041 is reported. *S. platensis* contains 39 annotated P450 genes which were classified in their CYP-families and further investigated regarding a putative role in the secondary metabolism. After evaluation of suitable conditions for heterologous expression in *E. coli*, the substrate spectra of P450 enzymes were determined and compared. Because *S. platensis* is used in the pharmaceutical industry for the production of drug metabolites, we evaluated all P450s regarding their activities against chemically diverse drugs. Five P450s from the CYP105 (OSY47991, OSY46529, OSY44745) and CYP107 (OSY40859, OSY37796) families showed activity and complementary selectivities against eight model substrates. Among those, OSY37796 (CYP107Z) is a promising candidate for drug metabolite syntheses as it converted all tested drugs with high activity.

1. Introduction

Streptomyces are Gram positive bacteria found in soil, compost or marine habitats.[85a] This genus is known for its inherent metabolic diversity and is an abundant source of biologically active secondary metabolites. Over 70% of all antibiotics and interesting non-antibiotic compounds are derived from *Streptomyces*, including macrolides like avermectin, non-ribosomal peptides like daptomycin, or terpenoids like pentalenolactone.[106] In addition, as soil bacteria *Streptomyces* play a key role in the decomposition of organic matter as they are capable to degrade structurally complex xenobiotics including pharmaceuticals, agro-chemicals, and environmental pollutants.[87, 104]

The biosynthesis of most secondary metabolites in *Streptomyces* involves cytochrome P450 monooxygenases (CYP or P450). These enzymes are heme b-thiolate enzymes that catalyze oxy-functionalization of a vast variety of chemically diverse compounds utilizing molecular oxygen. Along with diverse natural compounds ranging from terpenes and fatty acids to rather complex and sophisticated macrolides, non-ribosomal peptides, or alkaloids, P450 substrates also include non-natural drugs and xenobiotics.[19b, 88b]

P450-catalyzed reactions are diverse and include hydroxylations, dealkylations, epoxidation and many other “unusual” reactions such as bi-aryl coupling, nitration or dehydrogenation.[88a] The biocatalytic potential of P450s involved in biosynthetic pathways represents a research field of high interest. A prominent example is P450 TxtE from *Streptomyces scabies*, which catalyzes aromatic nitration of L-tryptophan. Arnold and co-workers investigated TxtE and homologous enzymes to make them accessible for the synthesis of valuable derivatives adjacent to the original substrate.[107] In addition, P450s from *Streptomyces* have become important target enzymes for the production of human-like drug metabolites. Drug metabolites are valuable fine chemicals required in the pharmaceutical industry for testing in Drug Metabolism and Pharmacokinetics studies (DMPK) and confirming identities of metabolite structures (MetID).

Recent studies of various research groups including our own work indicated that P450s from *actinomycetes* and especially *Streptomyces* can resemble activities of human P450s to form chemically diverse human drug metabolites as reviewed elsewhere.[19b, 104, 108] Since in many cases different P450s of the same strain differ in their chemo- and regioselectivities, they

could be used to furnish different human drug metabolites from the same lead structure.[11] Up to date, nearly all literature-described drug-metabolizing P450s from *Streptomyces* belong to only two P450 families, namely CYP105 and CYP107. These P450 families are the largest within the *Streptomyces* genus indicating functional diversity in substrate scope, activity and selectivity among the different isoforms.[88a, 109]

However, the identification and, even more important, biochemical characterization of P450s from *Streptomyces* with interesting properties remain the largest bottlenecks for further investigations although the number of sequenced genomes increases rapidly. The availability of genome sequences allows a global analysis of the entire cytochrome P450 complement (CYPome) of the respective strain.[110] Nevertheless, to date only the CYPomes of three *Streptomyces* strains, namely *S. avermitilis,*[111] *S. coelicolor* A3(2)[112] and *S. virginiae*[113] were examined, whereas comprehensive bioinformatics data of 48 other *Streptomyces* strains are available.[114]

This study provides a comprehensive analysis of the CYPome of *Streptomyces platensis* DSM 40041. This strain has been used by pharmaceutical companies for drug metabolites production.[57] In our previous studies, three P450s from this strain have been found to oxidize chemically diverse substrates.[11, 108a] Furthermore, *S. platensis* is also known as a producer of valuable secondary metabolites including pladienolides with high antitumor activities.[115] Based on the recently published genome sequence of *S. platensis* DSM 40041,[116] within this study P450 genes were first identified and analyzed by computational methods to assign them into their CYP-families according to sequence homology, and identify possible roles in secondary metabolism. All 39 identified P450 genes were then heterologously expressed in *E. coli* and their capability to produce metabolites from a panel of chemically diverse compounds mainly drugs was evaluated namely amitriptyline, amodiaquine, dextromethorphan, geraniol, ritonavir, testosterone, terfenadine, and thioridazine. A variety of reactions, such as N- dealkylation, sulfoxidation and hydroxylation was observed.

2. Material and methods

2.1 Bioinformatic analysis

Potential P450 genes were verified by comparison of the DNA sequences with P450 reference sequences stored in the online database CYPED (https://cyped.biocatnet.de/) as described earlier.[11, 117] P450s involved in secondary metabolism were revealed by the online tool 'antiSMASH'.[118] The classification of P450 sequences into their respective CYP families was performed according to a two-step protocol described previously.[11] Briefly, P450 sequences were first manually aligned with 'ClustalOmega' and then classified applying the reference sequences of David Nelson's 'Cytochrome P450 homepage'.[20]

2.2 Cloning, expression and enzyme preparation of P450s *from S. platensis*

The 39 identified P450 genes building the CYPome of *S. platensis* DSM 40041 were amplified from genomic DNA by PCR and cloned into pET expression vectors (Novagen/EMD Millipore, Germany). Details of the resulting plasmids are summarized in the Supplementary Information. *E. coli* BL21 (DE3) (Novagen, Germany) and C43 (DE3) (Lucigen, USA) were transformed with the expression vectors and gene expression was conducted in 50 mL buffered TB-medium supplemented with 30 µg/mL kanamycin. The cultures were incubated in 500 mL shaking flasks at 37 °C and 180 rpm until an OD_{600} of 1.2 - 1.8 was reached. 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.5 - 1 mM 5-aminolevulic acid (5-ALA) were added prior to lowering the temperature to 25°C. Cultures were cultivated overnight (~ 20 h) and harvested by centrifugation (4°C, 5251 x *g*, 20 min). In order to obtain crude cell extracts, cell pellets were first resuspended in 5 mL 50 mM potassium phosphate buffer (KPi), pH 7.5, containing 100 µM phenylmethylsulfonyl fluoride (PMSF). Cells were then disrupted by sonication using a Branson Sonfier 250 (3 cycles for 2 min, 40% amplitude, duty cycle 4). The insoluble protein fraction was removed by centrifugation (50.000 x *g*, 25 min, 4°C) and the crude cell extract was collected.

2.3 Purification and expression of redox partners Fdr and YkuN and GDH

Expression and purification of the flavodoxin reductase from *E. coli* (Fdr, EMBL-Bank no. L04757) and flavodoxin from *Bacillus subtilis* (YkuN, EMBL-Bank no. CAA10877) were done as described previously.[108a, 119] Glucose dehydrogenase from *Bacillus megaterium* (GDH, EMBL-Bank no. D10626) was used for cofactor regeneration and applied as crude cell extract as reported previously.[120]

2.4 Spectroscopic methods and enzyme assays

Spectrophotometric measurements were conducted using a double-beam photometer (Lambda 35, Perkin Elmer, Germany). P450 concentrations were estimated from their absorption maximum at 450 nm in CO-difference spectra using the extinction coefficient $\varepsilon_{450\text{-}490}$ = 91 mM^{-1} cm^{-1}.[14b] The activity of GDH was determined based on the reduction of $NAD(P)^+$ to NAD(P)H in presence of glucose in a continuous photometric assay following the increase in absorption at 340 nm (ε_{340} = 6.22 mM^{-1} cm^{-1}). Reaction mixtures were prepared with 100 mM glucose, 100 µM $NADP^+$ and 100 µL of suitable diluted enzyme solution in 50 mM KPi, pH 7.5 in a final reaction volume of 1 mL.

2.5 Reconstitution of P450 activities

Substrate screening was performed using *E. coli* crude cell extracts. Reaction mixtures (total 200 µL of 50 mM KPi, pH 7.5) contained 2 µM P450, 20 µM YkuN, 2 µM Fdr, 200 µM substrate dissolved in an organic solvent (DMSO, ethanol, or methanol; final concertation in the assay was 2% (v/v)), and 5 U ml^{-1} GDH and 20 mM glucose for cofactor regeneration. Reactions were initiated by addition of 200 µM NADPH and reaction mixtures were incubated at 25 °C using Thermomixer C (Eppendorf, Germany) with a shaking frequency of 600 min^{-1}. Depending on the substrate, an internal standard was added prior to extraction as described in detail in the Supplementary Information. For a better extraction of alkaline compounds, 300 µL 100 mM sodium carbonate buffer, pH 10 (pH 9.6 for amodiaquine) were added to the reaction and extraction was done with ethyl acetate. The organic phase was recovered, and analyzed directly by gas chromatography coupled to mass spectrometry (GC/MS) analysis, while the organic phase of samples was evaporated completely and the

residues were resuspended in methanol (LC/MS grade) prior to liquid chromatography coupled to mass spectrometry (LC/MS) analysis.

2.6 Product analysis

Conversions of all compounds (except geraniol) were analyzed by LC/MS on a Prominence/LCMS 2020 device (Shimadzu, Germany) equipped with either a Chromolith® Performance RP-18e column (100 x 4.6 mm, Merck, Germany) or a Chromolith® Performance RP-8e column (100 x 4.6 mm, Merck, Germany) for separation of the analytes. The column oven was kept at 30 °C. Solvent A was always ddH_2O with 0.1% formic acid, while solvent B was either methanol (MeOH) or acetonitrile (ACN). The injection volume was 1 µL and the flow varied between 0.5 and 1 mL min^{-1} depending on the respective substrate as described in detail in the Supplementary Information. All compounds were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI).

Samples containing geraniol were analyzed by GC/MS as described previously.[121]

3. Results and Discussion

3.1 Computational analysis

The genome sequencing of *S. platensis* DSM 40041 performed in one of our group revealed 39 P450 sequences identified via automated computational annotation and comparison to the online P450 database CYPED.[108a, 116] This number of P450 sequences is generally comparable or a bit higher than other *Streptomyces*' CYPomes, *e.g.* those from *S. avermitilis* (33 P450s) and *S. coelicolor*(A3)2 (18 P450s).[112] Considering the 7302 annotated (potential) protein coding sequences,[116] the genomic abundance of P450 genes in *S. platensis* accounts for 0.5%. The abundance of P450 genes of less than 1% is in line with previous results of other *Streptomyces*' CYPomes, with the highest amount reported in *S. clavuligerus* (1.1%).[114]

Table 2.3.1: The CYPome of *Streptomyces platensis* DSM 40041. Reference sequences were extracted from David Nelson's database.[20] Putative roles in biosynthetic clusters were assigned with "AntiSmash".[118]

Entry No.	GenBank number	Number of amino acids	Assigned CYP family according to D. Nelson nomenclature	Closest sequence identity	Sequence identity [%]	Biosynthetic cluster (AntiSmash)
1	OSY48498	1068	CYP102D	CYP102D1 from *S. avermitilis*	68	-
2	OSY47769	1062	CYP102G	CYP102G2 from *Saccharopolyspora erythraea* NRRL23338	66	-
3	OSY34726	412	CYP105B	CYP105B1 from *.S. griseolus*	77	-
4	OSY47991	395	CYP105D	CYP105D8 from *S. tubercidicus* strain I-1529	**85**	**-**
5	OSY40501	417	CYP105AA	CYP105AA1 from *S. tubercidicus* strain R-922	77	-
6	OSY46529	408	CYP105AA	CYP105AA5 from *S. lydicus*	88	-
7	OSY44745	428	CYP105AC	CYP105AC1 from *S. erythraea* NRRL23338	69	-
8	OSY47114	429	CYP105	new subfamily of CYP105	50	T1pks-NRPs
9	OSY46350	408	CYP105	new subfamily of CYP105	52	Terpene
10	OSY46937	396	CYP107E	CYP107E5 from *S. bingchengensis*	79	-
11	OSY44419	402	CYP107F	CYP107F1 from *S. griseus* D45916	72	Lantipeptide-T3pks
12	OSY48372	401	CYP107L	CYP107L10 from *S. griseus*	64	-
13	OSY40859	399	CYP107L	CYP107L11 from *S. griseus*	64	
14	OSY42650	439	CYP107U	CYP107U3 from *S. peucetius*	86	
15	OSY47824	407	CYP107X	CYP107X3 from *S. bingchengensis*	72	
16	OSY37796	430	CYP107Z9v2	CYP107Z9 from *S. tubercidicus* strain NRAA-7027	99	
17	OSY48350	389	CYP107	new subfamily of CYP107	40	
18	OSY47382	404	CYP107	new subfamily of CYP107	53	Thiopeptide-lantipeptide-NRPs
19	OSY46946	422	CYP107	new subfamily of CYP107	44	Terpene-bacteriocin
20	OSY48230	471	CYP107-like	/	37	
21	OSY48349	388	CYP107-like	/	35	
22	OSY47218	392	CYP113	new subfamily of CYP113	49	Lassopeptide-NRPs

Table 2.3.1: continued.

Entry No.	GenBank number	Number of amino acids	Assigned CYP family according to D. Nelson nomenclature	Closest sequence identity	Sequence identity [%]	Biosynthetic cluster (AntiSmash)
23	OSY41108	410	CYP125A	CYP125A16 from *S. bingchengensis*	83	
24	OSY46283	430	CYP125	new subfamily of CYP125	43	
25	OSY46264	403	CYP147B	CYP147B1 from *S. avermitilis*	63	
26	OSY42990	409	CYP147	new subfamily of CYP147	54	
27	OSY40215	381	CYP154A	CYP154A4 from *S. scabies* SCAB20211	57	
28	OSY48231	452	CYP157-like	/	39	
29	OSY40216	458	CYP157-like	/	34	
30	OSY34873	414	CYP159A	CYP159A1 from *S coelicolor*	68	
31	OSY47661	428	CYP165	new subfamily of CYP165	43	
32	OSY43763	466	CYP174-like	/	33	Terpene
33	OSY43764	467	CYP174-like	/	32	Terpene
34	OSY47468	527	CYP174-like	/	21	
35	OSY47408	472	CYP183	new subfamily of CYP183	41	Terpene
36	OSY46348	400	CYP183-like	/	35	Terpene
37	OSY45653	356	CYP183-like	/	21	
38	OSY47449	440	CYP208-like	/	34	NRPs-T1Pks
39	OSY40301	474	CYP208-like	/	38	

The first step of analysis was the classification of the P450 sequences into CYP-families according to the nomenclature of David Nelson according to which all P450 enzymes with an amino acid identity of at least 40% belong to the same family and with an amino acid identity of at least 55% to the same subfamily.[20] Based on this nomenclature, the *S. platensis*' P450s could be assigned to twelve CYP families.

A majority of 19 P450 sequences were classified as members of the CYP105 and CYP107 families, two families that are commonly found in *Streptomyces* (Table 2.3.1). Twelve of these 19 sequences share a sequence identity of ≥ 55% and belong to known P450 subfamilies. Among these twelve sequences are also three P450s (CYP107L, CYP105D and CYP107Z) that have been recently described by our group to oxidize a broad range of drugs.[11, 108a] Five of the 19 sequences could not be assigned to any known existing P450 subfamily within the CYP105 and CYP107 families indicating an affiliation to not yet

described subfamilies. The two remaining genes assigned to the CYP107 family are referred to as 'CYP107-like' genes meaning that the sequence identity is lower than the threshold of 40%. Since a more closely related family has not yet been described, we refer to these two genes as members of the CYP107 family.

Two genes belong to the CYP102 family; members of this family including the well-studied P450 BM3 (CYP102A1)[52] are of biocatalytic interest because of their generally high activity and facilitated handling due to the natural fusion of a diflavine reductase domain to the monooxygenase domain.[122] The remaining P450s belong to CYP113, CYP125, CYP147, CYP154, CYP159, CYP165 and CYP183 families, which are far less studied compared to CYP105, CYP107 and CYP102 families. Nevertheless, these enzymes are worthwhile for further investigations as oxidations of steroids,[123] fatty acids[121, 124] and alkanes[79a] have been described for some members of these P450 families.

Besides their classification into P450 families, a potential role of the P450s in secondary metabolism was assigned using 'antiSMASH'.[118] The online tool 'antibiotics and secondary metabolites analysis shell' (antiSMASH) is used to identify biosynthetic gene clusters in bacteria and fungi. By multiple sequence alignments of characterized signature proteins or protein domains and analyses of the cluster organization, antiSMASH provides information about the type of biosynthetic cluster and predicts possible chemical structures in case of polyketides and non-ribosomal peptides.[118, 125] P450s in *Streptomyces* are often involved in syntheses of secondary metabolites from different classes including terpenes, polyketides and peptides among others.[114] The enzymes involved in such biosynthesis pathways are often arranged in clusters within the genome of the respective strain. Indeed, twelve *cyp* genes were found within such putative enzyme clusters and therefore are potentially involved in secondary metabolite biosynthesis. Six of these P450 genes were proposed to be involved into four putative clusters for terpenes' biosynthesis. This suggestion is endorsed by the presence of other biosynthetic genes typically involved in terpene biosynthesis such as terpene cyclases and synthases (Figure 2.3.1). In two of the four assigned terpene clusters, two *cyp* genes are located in close proximity to each other and to the respective terpene cyclase probably suggesting subsequent oxidations by these P450s after cyclisation of the terpene substrate. The other six *cyp* genes of secondary metabolites' synthesis are arranged in either putative polyketide or peptide synthesis clusters.

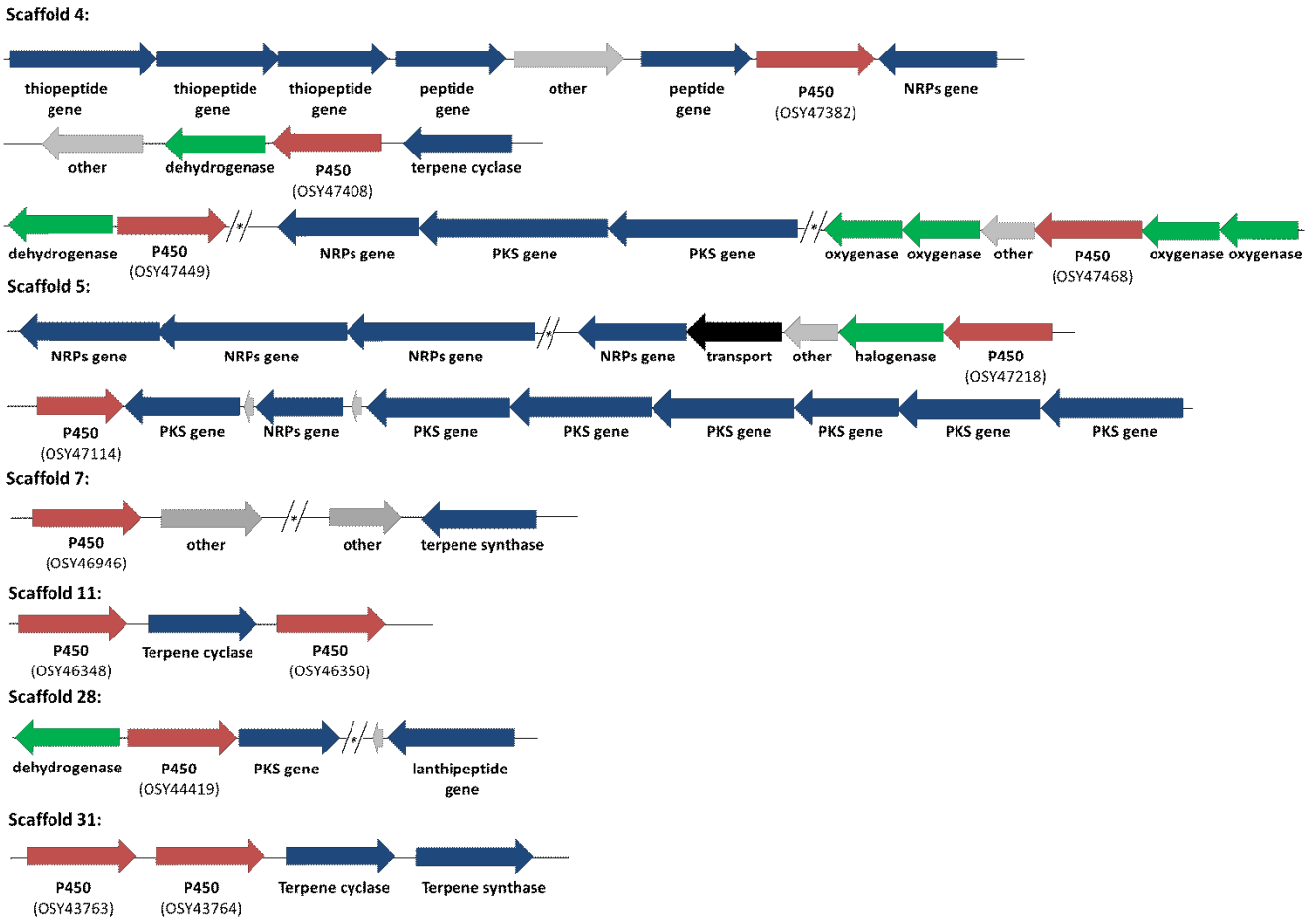


Figure 2.3.1: Schematic organization of putative biosynthetic clusters including a P450 in the genome of *Streptomyces platensis* DSM 40041. For simplification, data gained from 'AntiSmash' were reduced to the relevant biosynthetic genes around the P450 gene (further genes are indicated by an asterisk). The P450 genes (red) of the respective scaffolds are annotated with their GenBank number summarized in Table 2.3.1.

3.2 Heterologous expression of *cyp* genes in *E. coli*

The heterologous expression of all 39 identified P450s in *E. coli* under identical experimental conditions was investigated. To achieve high expression level of possibly all GC-rich *cyp* genes from *S. platensis* in *E. coli,* we used two *E. coli* strains: (a) *E. coli* BL21 (DE3) and (b) the 'Walker strain' *E. coli* C43 (DE3),[126] that is less sensitive to toxicity caused by high amounts of recombinant protein and thus often better suited for 'difficult-to-express' genes.[127] Evaluation of the expression levels of the individual P450s via CO-difference spectra (Figure 2.3.2) revealed that 25 of the 39 *cyp* genes from *S. platensis* could be expressed under the chosen conditions in *E. coli* with 15 genes expressed at high concentrations of up to ~ 1,600 nmol g_{CDW}^{-1}. The generally better suitability of C43 (DE3) over BL21 (DE3) for

heterologous expression of P450s from GC-rich organisms is in line with observations reported previously.[123c]

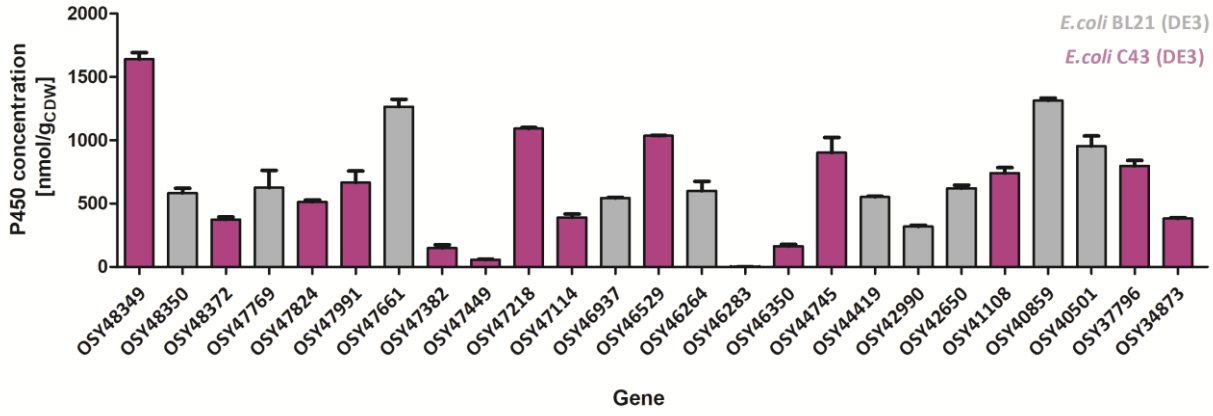


Figure 2.3.2: Expression of the *S. platensis* CYPome in *E. coli* BL21 (DE3) (grey) and C43 (DE3) (purple). Data represent the average of biological duplicates; P450 concentrations are given in nmol P450 per gram cell dry weight (g_{CDW}). Concentrations were determined according to the absorption difference in CO-difference spectra between 450 nm and 490 nm. Only results of gene expressions with a reproducible maximum absorbance at 450 nm are shown. The complete expression profile of all *p450* genes in strains BL21 (DE3) and C43 (DE3) is summarized in the Supplementary Material (Supplementary Figures 2.3.1 and 2.3.2).

P450 genes located in biosynthetic clusters (Figure 2.3.1) were much more difficult to express. Six out of these twelve P450s (OSY47408, OSY47468, OSY46946, OSY46348, OSY43763, OSY43764) could not be detected in soluble protein fractions in either *E. coli* strain meaning that they did not show a detectable maximum absorption at ~ 450 nm in CO-difference spectra. It is interesting to note that most of these P450s belong to CYP families, which have not been described in the literature to date and, as mentioned earlier, have a sequence identity of less than 40% to all known CYP families.

3.3 Activity screening of expressed P450s

All P450 enzymes that could be expressed in titers above 100 nmol g_{CDW}^{-1} were investigated for their ability to oxidize a panel of drug compounds. These compounds were chosen based on their chemical diversity allowing screening for most of the reaction types P450s are capable to catalyze. Except geraniol, all substrates are drugs differing in size and functional groups (Figure 2.3.3). Since a quantitative analysis was not always possible due to the difficult separation of metabolite mixtures, a semi-quantitative analysis was carried out in

order to get an overview of the activity of every individual P450. For this semi-quantitative analysis, conversion was divided into four categories: trace (1-10%), low (11-50%), moderate (51-90%) and high to full conversion (91-100%).

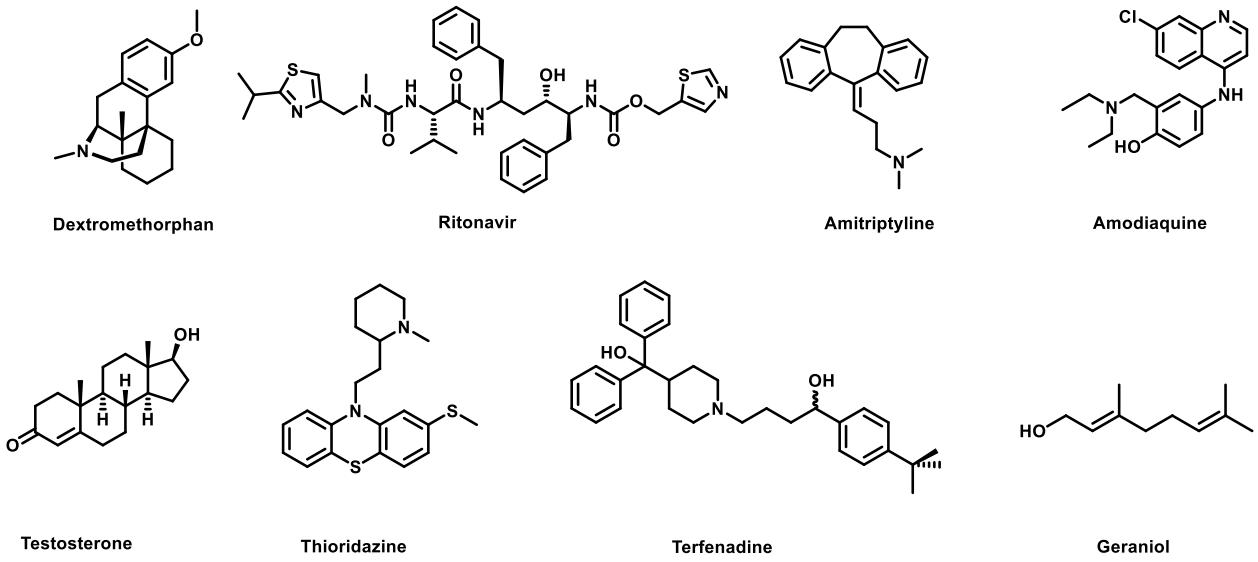


Figure 2.3.3: Chemical structure of the drug compounds applied in the P450 activity screening.

Out of the 23 P450s that were investigated, 16 showed activity against at least one of the substrates (Table 2.3.2). The best performance was observed for OSY37796 (CYP107Z); this enzyme was active against all eight substrates and led to the highest conversion values (high to full conversion for amitriptyline, terfenadine, testosterone and thioridazine; moderate conversion for dextromethorphan and geraniol; low conversion for amodiaquine and ritonavir).

The second best performance was observed for OSY44745 (CYP105AC). This P450 shows activity against all the tested compounds except testosterone. Compared to OSY37796, the conversion is higher for amodiaquine (moderate) and dextromethorphan (high), in the same range for terfenadine (high) and ritonavir (low) and less for amitriptyline (moderate), geraniol (trace) and thioridazine (low).

CYP105D (OSY47991) shows the third best performance with activity against all eight compounds. Compared to OSY37796, conversion is generally lower (amodiaquine,

dextromethorphan, geraniol: trace; amitriptyline, terfenadine, thioridazine: moderate), except for testosterone where it is in the same range (high to full) and for ritonavir with higher conversion (moderate).

OSY46529 (CYP105AA) and OSY40859 (CYP107L) show the same range of overall activity. However, OSY46529 shows activity against all eight compounds, whereas OSY40859 is only active against five of the substrates - namely amitriptyline (low), amodiaquine, ritonavir, thioridazine (moderate) and terfenadine (high to full). OSY47218 (CYP113) and OSY46937 (CYP107E) were capable of conversion of four (amitriptyline (moderate); amodiaquine (trace); ritonavir (low); thioridazine (high to full)) or five (amitriptyline, thioridazine (low); amodiaquine, terfenadine (trace); testosterone (high to full)) substrates, respectively.

The other P450s show generally lower activities with high to full conversion of geraniol by OSY34873 (CYP159A), moderate conversion of amitriptyline and thioridazine by OSY48372 (CYP107L); and only trace activity for amodiaquine and terfenadine conversion by OSY47669 (CYP102G). Further, amitriptyline, terfenadine and testosterone were converted by OSY47824 (CYP107X), amitriptyline and dextromethorphan by OSY47114 (CYP105), geraniol by OSY46350 (CYP107), terfenadine by OSY42990 (CYP147), amitriptyline and thioridazine by OSY42650 (CYP107U), and finally testosterone by OSY40501 (CYP105AA).

Generally, the most active P450s are members of the CYP105 and CYP107 families. This is not unexpected as these P450 families are known to have diverse substrate spectra and catalytic functions. It is likely that CYP105 and CYP107 enzymes with a broad substrate spectrum can adapt the size of their binding pockets to oxidize chemically diverse substrates.[128] However, due to great variation in binding modes between the individual CYP105 and CYP107 enzymes the diverse reactivity of a P450 remains to be explored individually.[36b, 109] Members of these two P450 families are predominantly found in *Actinomycetes* where they are involved in the degradation of xenobiotics and biosynthesis of secondary metabolites.[129] Because of these properties, P450s have become indispensable for the pharmaceutical industry not only with respect to the production of drug metabolites for toxicity studies during drug development (DMPK)[130] but also for their high potential to produce natural products and derivatives thereof.[88a]

Table 2.3.2: Semi-quantitative conversion of selected substrates with 23 P450s of the *Streptomyces platensis* complement. Conversion was divided into four categories: trace (1-10%), low (11-50%), moderate (51-90%) and high to full conversion (91-100%). The complete conversion profile for each compound by all P450s is summarized in the Supplementary Material (Supplementary Tables 2.3.4-11). Reaction conditions (0.2 mL, 25 °C, 600 min-^{1}): KPi buffer (50 mM, pH 7.5), co-solvent DMSO, ethanol or methanol (2% (v/v)), substrate (0.2 mM), NADPH (0.2 mM), GDH (5 U mL^{-1}), glucose (20 mM), P450 (2 µM), Fdr (2 µM), YkuN (20 µM). The conversion time was set to 4 h and 20 h. Data represent average values of at least two replicates.

Substrate	no conversion	trace conversion (1-10%)	low conversion (11-50%)	moderate conversion (51-90%)	high to full conversion (91-100%)
amitriptyline	OSY48349, OSY48350, OSY47769, OSY47661,OSY47382,OSY46264, OSY46350, OSY44419, OSY42990, OSY41108, OSY34873	OSY47824, OSY47114, OSY42650	OSY48372, OSY47991, OSY46937,OSY46529, OSY40859, OSY40501	OSY47218, OSY44745	OSY37796
amodiaquine	OSY48349, OSY48350, OSY48372,OSY47824, OSY47661,OSY47382,OSY47114, OSY46264, OSY46350, OSY44419, OSY42990, OSY42650, OSY41108, OSY40501,OSY34873	OSY47769, OSY47991, OSY47218, OSY46937	OSY46529, OSY37796	OSY44745, OSY40859	/
dextrometh-orphan	OSY48349, OSY48350, OSY48372, OSY47769, OSY47824, OSY47661,OSY47382,OSY47218, OSY46264, OSY46350, OSY44419, OSY42990, OSY41108, OSY40859, OSY40501,OSY34873	OSY47791, OSY47114, OSY46937	OSY46529	OSY37796	OSY44745
geraniol	OSY48349, OSY48350, OSY48372, OSY47769, OSY47824, OSY47661,OSY47382,OSY47218, OSY47114, OSY46937, OSY46264, OSY44419, OSY42990, OSY42650, OSY41108, OSY40859, OSY40501	OSY47991, OSY46529, OSY46350, OSY44745	/	OSY37796	OSY34873
ritonavir	OSY48349, OSY48350, OSY48372, OSY47769, OSY47824, OSY47661,OSY47382, OSY47114, OSY46937, OSY46264, OSY46350, OSY44419, OSY42990, OSY42650, OSY41108, OSY40501, OSY34873	/	OSY47218, OSY44745,OSY37796	OSY47991, OSY40859	OSY46529
terfenadine	OSY48349, OSY48350, OSY47824, OSY47661,OSY47382,OSY47218, OSY47114, OSY46937, OSY46264, OSY44419, OSY46350, OSY42650, OSY41108, OSY40859, OSY40501, OSY34873	OSY47769, OSY47824, OSY46529, OSY42990	OSY47991	/	OSY44745, OSY40859, OSY37796

Table 2.3.2: continued.

Substrate	no conversion	trace conversion	low conversion	moderate conversion	high to full conversion
testosterone	OSY48349, OSY48350, OSY48372, OSY47769, OSY47661,OSY47382,OSY47218, OSY47114, OSY46264, OSY46350, OSY44745, OSY44419, OSY42990, OSY42650, OSY41108, OSY40859, OSY34873	OSY47824, OSY46529, OSY40501	/	/	OSY47991,OSY 47114, OSY37796
thioridazine	OSY48349, OSY48350, OSY47769, OSY47824, OSY47661,OSY47382, OSY47218, OSY47114, OSY46264, OSY46350, OSY44419, OSY42990, OSY41108, OSY40501, OSY34873	OSY42650	OSY48372, OSY47991, OSY46937, OSY46529, OSY44745	OSY40859	OSY47218, OSY37796

It is interesting to note that out of ten expressed P450s classified as CYP107 members three exhibited high activity (OSY46937, OSY40859, OSY37796), three showed low activity (OSY48372, OSY47824, OSY42650) and four did not show any activity at all (OSY48349, OSY48350, OSY47382, OSY44419). The same holds true for CYP105 members, where three enzymes showed high activity (OSY47991, OSY46529, OSY44745) and three only very low activity (OSY47114, OSY46350, OSY40501).

The observation, that P450s of the same subfamilies show different activities, supports the general doctrine that P450 activity cannot be predicted from the amino acid sequence and their classification into subfamilies.[104] It should be noted, however, that besides the obvious reason that the chosen compounds might not be substrates for these P450s, the absence of conversion might also be explained by an unsuccessful activity reconstitution with the non-physiological redox partners, as potential physiological ones from *S. platensis* have not been investigated within this study.

It is interesting to note that one P450, capable to convert four compounds with high overall activity, belongs to the CYP113 family. Little is known about CYP113 except that some P450s of this family appear to be involved in the production of secondary metabolites, especially macrolactones in *Actinomycetes*, such as the 16-membered macrolide tylosin in *S. fradiae* (CYP113B1),[131] spiramycin in *S. ambofaciens,*[132] and midecamycin in *S. mycarofaciens*

(CYP113B3).[133] The best studied representative of this family is CYP113A1 (EryK), a C-12 hydroxylase from *S. erythraeus* involved in the biosynthesis of erythromycin A, a 14-membered macrolactone.[90a]

Among the P450s with minor activities, there is one member each belonging to the families CYP102 (OSY47769), CYP147 (OSY42990) and CYP159 (OSY34873). CYP102 members are naturally occurring self-sufficient P450s which are capable of fatty acid oxidation mostly at subterminal positions.[122] The CYP147 family belongs to one of seven P450 families that taken together contribute to 62% of all P450s.[114] Known members of the CYP147 family are CYP147F1 from *S. peucetius* and CYP147G1 from *Mycobacterium marinum*, both are fatty acid hydroxylases acting at ω- or ω-1 position.[124] The CYP159 family is quite common in *Streptomyces* species, but to the best of our knowledge, members of the CYP159 family have not been studied so far. However, the full conversion of geraniol by OSY34873 and the lack of any activity against the other seven compounds hint to a rather narrow substrate spectrum of this enzyme.

The three other P450s that did not show any activity against the tested compounds belong to the families CYP125 (OSY41108) and CYP165 (OSY47661). Besides the fact that the non-physiological redox partners might not have supported activity reconstitution it is likely that these P450s have a rather narrow substrate spectrum limiting their activity to one or few natural substrates. Members of the CYP125 family have been previously reported to be involved in the degradation of cholesterol which is either essential for the growth on cholesterol e.g. in *Rhodococcus jostii*, or is implicated in the pathogenicity of *Mycobacterium tuberculosis*.[123a, 123b] Since to the best of our knowledge no members of the CYP165 family have been studied so far, any assumption on their natural substrates is not possible.

3.4 Reaction types and reaction products of P450-catalyzed substrate conversions

In order to differentiate the P450s further, we investigated the active P450s regarding their selectivity towards the tested compounds and their product spectra. More specifically, we focused on the enzymes showing moderate to high activities.

Generally, hydroxylations and N-dealkylations were the two main reaction types observed on the model substrates catalyzed by P450s from *S. platensis* (Table 2.3.3). OSY47991

(CYP105D), OSY46529 (CYP105AA) and OSY40859 (CYP107L) were not only the most active P450s but oxidized their substrates also with high selectively via hydroxylation. This holds especially true for testosterone, ritonavir, and terfenadine. Apart from hydroxylations, other P450 reactions such as N-dealklyations (ritonavir,[134] terfenadine[95]) or alcohol oxidation (testosterone[100]) are also possible for these substrates.

The regio- and stereoselective introduction of hydroxy groups confer steroid drugs with different bioactive properties, such as anti-inflammatory, anti-cancer, immunosuppressive and diuretic just to name a few.[33] Within this study the steroid model substrate testosterone was oxidized by six P450s; three of which enabled complete conversion (Table 2.3.2). However, only CYP105D (OSY47991) could oxidize testosterone with high selectivity to 2β-hydroxy-testosterone as the main product, whereas the other enzymes yielded product mixtures. A number of microbial P450s have been reported to hydroxylate steroids at various positions; especially members of the families CYP106 (*e.g.* CYP106A1 at positions 6β, 7β, 9α, 11 and 15β),[135] CYP109 (*e.g.* CYP109E1 at position 16β),[136] CYP154 (*e.g.* CYP154C5 at position 16α),[102] and CYP260 (*e.g.* CYP260A1 at position 1α).[78] A recent study showed that CYP105D7 from *S. avermitilis* is also capable of catalyzing 2β-hydroxylation of testosterone, however with low activity (< 10% conversion).[98]

The HIV protease inhibitor ritonavir was also accepted by six enzymes as substrate (Table 2.3.2). Three P450s (CYP105D, CYP105AA, CYP107L) exhibited a high regioselectivity of up to 85% for hydroxylation of the *tert*-butyl moiety of ritonavir, resulting in the human metabolite hydroxy ritonavir.[11, 108a] From these, CYP105AA (OSY46529) resulted in nearly complete conversion maintaining a high selectivity.

Terfenadine can also be oxidized at its *tert*-butyl moiety. In three consecutive oxidation reactions, the human active metabolite fexofenadine can be obtained which differs from terfenadine in a carboxyl group at the *tert*-butyl moiety.[137] CYP107L (OSY40859) was capable of a full conversion with high activity and selectivity to fexofenadine as identified from co-elution with an authentic reference standard (Anne Worsch, unpublished data). Six P450s showing activity against terfenadine were not able to produce fexofenadine but produced mixtures of different dealkylation and single oxidation products, instead. The

production of fexofenadine from terfenadine by a CYP107L member is in line with a previous study on $P450_{terf}$ from a *S. platensis* strain NRRL 2364.[95]

The regioselective hydroxylation of drugs and other pharmaceutical molecules is an important key feature for a potential biocatalytic application of P450s.[15] The CYPome of *S. platensis* contains at least three P450s (CYP107L, CYP105AA, CYP105D) exhibiting both high activity and selectivity towards hydroxylation of drug compounds demonstrating the potential of this strain for application in the pharmaceutical industry for the production of drug metabolites.[57]

Table 2.3.3: Overview of selective oxidations by P450s from *S. platensis*. The P450 with the highest activity and selectivity is mentioned. Reaction conditions (0.2 mL, 25 °C, 600 min^{-1}): KPi buffer (50 mM, pH 7.5), co-solvent DMSO, ethanol or methanol (2% (v/v)), substrate (0.2 mM), NADPH (0.2 mM), GDH (5 U mL^{-1}), glucose (20 mM), P450 (2 µM), Fdr (2 µM), YkuN (20 µM). The conversion time was set to 4 h and 20 h. The product percentage was calculated based on the observed product peak areas in relation to the sum of product peaks areas. Details for every model substrate are provided in the Supplementary (Supplementary Table 2.3.4-11).

Substrate	Product	Reaction	P450	Activity	Product percentage [%]
Testosterone	2β-Hydroxytestosterone	Hydroxylation	OSY47991 (CYP105D)	full	70
Ritonavir	Hydroxy ritonavir	Hydroxylation	OSY46529 (CYP105AA)	full	85
Terfenadine	Fexofenadine	Triple oxidation	OSY40859 (CYP107L)	full	87
Amitriptyline	Nortriptyline	N-Demethylation	OSY47218 (CYP113)	moderate	87
Amodiaquine	Desethylamodiaquine	N-Deethylation	OSY44745 (CYP105AC)	moderate	89
Dextromethorphan	3-Methoxymorphinan	N-Demethylation	OSY44745 (CYP105AC)	full	44

Dealkylated metabolites can in principle be furnished from amodiaquine, amitriptyline, dextromethorphan, thioridazine and ritonavir. CYP105AC was identified as the most promising enzyme regarding its activity and selectivity for dealkylation reactions. Dextromethorphan can be converted either by N-demethylation to 3-methoxymorphinan or by O-demethylation to dextrorphan.[138] It was deduced from an authentic reference that the O-demethylated product was neither formed by CYP105AC nor by any of the other P450 from *S. platensis* showing activity towards dextromethorphan (Figure 2.3.S3). This observation suggests that N-dealkylation is favored over O-dealkylation by the dealkylating P450s from *S. platensis*.

Amodiaquine and amitriptyline contain one amine group which is either methylated (amitriptyline) or ethylated (amodiaquine) (Table 2.3.3). Eleven P450s were able to oxidize amitriptyline and eight P450s amodiaquine, albeit with low activities (Table 2.3.2). The most selective and active P450s with respect to dealkylation reactions were CYP113 (OSY47218) and CYP105AC, both showing high regioselectivity of about 90% towards N-dealkylation. As mentioned earlier, especially CYP105AC appears to prefer to catalyze N-dealkylation reactions making this P450 enzyme a worthwhile candidate for testing of other substrates with alkylated amine functionalities.

4. Conclusion

S. platensis DSM 40041 is an *Actinomycete* used by the pharmaceutical industry for the production of drug metabolites.[57] Within this study, the complete CYPome of *S. platensis* consisting of 39 *p450* genes was identified, and the capability of the individual P450s to oxy-functionalize a panel of eight chemically diverse compounds, mainly drugs, was analyzed. The most interesting P450s exhibiting high activities are members of either the CYP105 (OSY47991, OSY46529, OSY44745) or CYP107 (OSY40859, OSY37796) family. OSY37796 (CYP107Z) was the most active P450 and showed activity against all eight tested compounds confirming its involvement in xenobiotic degradation and making it a highly suitable candidate with respect to the production of drug metabolites.

Funding information

The research was funded by the Federal Ministry of Education and Research [grant numbers 031A223A and 031A223B] under the umbrella of the ERA-IB2 3rd call project “HyPerIn” [project number EIB.12.026].

Acknowledgements

We wish to thank Anna Mühlinghaus, Julia Rutsch, Davide Decembrino and Leonhard Kohleick for their contributions to enzyme expression and activity screening.

Author contributions

T.H. and A.W. planed, carried out or assisted in the experiments, analyzed the data, and drafted the manuscript. P.L-H. implemented the initial analytics and contributed to the initial design of experiments. C.G. and K.S. acquired funding and performed bioinformatic analyses in close cooperation with T.H. and A.W. M.G. and V.B.U. acquired funding, curated the data, and reviewed the manuscript draft.

Competing interests

The authors declare no competing financial interests.

2.3.1 Supporting information

Supplementary Material and methods

Supplementary Table 2.3.1: Synthetic oligonucleotides for gene amplification. Restriction sites are underlined.

Primer name	DNA-sequence (5'-3')	Restriction enzymes	usage
F-NdeI-SPL01126	GAAGTTCCATATGTGGCCGGGAACGACAAGGACCG	NdeI/XhoI	Amplification of OSY47408
RC-SPL1126-XhoI	GATCCTCGAGTTACCACCTGACGGGCAGTGC		
F-NdeI-SPL01167	GGAATTCCATATGCGTATCCCCGGCCCCGAGC	NdeI/XhoI	Amplification of OSY47449
RC-SPL1167-XhoI	GATCCTCGAGTCAGGCGCCAGAGAGGTCC		
F-NdeI-SPL01317	GAGCGGCCCATATGACTGTGCGCGACGAGATCG	NdeI/XhoI	Amplification of OSY47218
RC-SPL1317-XhoI	GATCCTCGAGTCAACGCCGGGCCGTCATCG		
F-NdeI-SPL02251	GAGATTCCATATGACCGTATCCGCAGCCGACGCC	NdeI/XhoI	Amplification of OSY46283
RC-SPL2251-XhoI	TAATATCTCGAGCTACCCCCCGCCGGGAACGTC		
F-NdeI-SPL02316	GAGATTCCATATGTTTGTTCTGAACTCGCCGGAGG	NdeI/XhoI	Amplification of OSY46348
RC-SPL2316-XhoI	TAATATCTCGAGTCACCGGGTTGGCCACGG		
F-NcoI-SPL02943	ATATCCATGGCTATGCACCGCGTCGGCGAG	NcoI/XhoI	Amplification of OSY45653
RC-SPL2943-XhoI	GGTCCTCGAGTCACGGACGAGCCATAGGAC		
F-NheI-SPL04452	GCCATATGGCTAGCTCCACTGTTTCTGCCCGATTAG	NheI/XhoI	Amplification of OSY43763
RC-SPL04452-XhoI	TAATATCTCGAGTCAGCGACGGTGCAGCCGGAG		
F-NdeI-SPL04453	GAGATTCCATATGCCCCTGACTCCGCGTACCCGTGC	NdeI/XhoI	Amplification of OSY43764
RC-SPL4453-XhoI	TAATATCTCGAGTCATCCGCGGGGCTCCGGCCGCAAG		
F-NdeI-SPL05184	TATATTCCATATGCCATGTCCCGCGCTGCCCGATG	NdeI/XhoI	Amplification of OSY41108
RC-SPL5184-XhoI	TACGGTCTCGAGTCAACCGTAGTGAACCCGGAGTTCC		
F-NdeI-SPL05473	GATGCGCCATATGTCCCTGACCGAATCCCTTCCC	NdeI/XhoI	Amplification of OSY40301
RC-SPL5473-XhoI	TATATTCTCGAGCTACTCCCCCGCCCCGGC		
F-NdeI-SPL05497	GCTAGACCATATGGCCCGGACCGAACTGCCC	NdeI/XhoI	Amplification of OSY40215
RC-SPL5497-XhoI	TGATCTCTCGAGTCACGCCGGAGTCAGGCGGAC		
F-NdeI-SPL07341	AGACATACATATGACCACTCTCCCGCAGCCGC	NdeI/XhoI	Amplification of OSY34873
RC-SPL7341-XhoI	TATATTCTCGAGTCAGCCCTCGCTCGCCCG		

Supplementary Table 2.3.2: Plasmids used in this study.

vector (internal number)	genes with EMBL bank number	vector properties	reference/source
pET28a-SPL_00107 (ANW01)	Gene „SPL_00107“ (EMBL-Bank: OSY48230)	SPL_00107 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET28a-SPL_00108 (ANW 02)	Gene „SPL_00108“ (EMBL-Bank: OSY48231)	SPL_00108 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_00226 (ANW 03)	Gene „SPL_00226“ (EMBL-Bank: OSY48349)	SPL_00226 (NdeI/SalI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_00227 (ANW 04)	Gene „SPL_00226“ (EMBL-Bank: OSY48350)	SPL_00227 (NdeI/SalI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_00249 (ANW 05)	Gene „SPL_00249“ (EMBL-Bank: OSY48372)	SPL_00249 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET24b_SPL_00375 (ANW 06)	Gene „SPL_00375“ (EMBL-Bank: OSY48498)	SPL_00375 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET24b_SPL_00400 (ANW 07)	Gene „SPL_00400“ (EMBL-Bank: OSY47769)	SPL_00400 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_00455 (ANW 08)	Gene „SPL_00455“ (EMBL-Bank: OSY47824)	SPL_00455 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_00625 (THI 09)	Gene „SPL_00625” (EMBL-Bank: OSY47991)	SPL_00625 (NdeI,XhoI) cloned in pET24b	Hilberath et al.[11]
pET28a-SPL_00916 (ANW 09)	Gene „SPL_00916“ (EMBL-Bank: OSY47661)	SPL_00916 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET28a-SPL_01100 (THI 10)	Gene „SPL_01100” (EMBL-Bank: OSY47382)	SPL_01100 (NcoI,XhoI) cloned in pET28a	Hilberath et al.[11]
pET24b-SPL_01126 (THI 03)	Gene „SPL_01126“ (EMBL-Bank: OSY47408)	SPL_01126 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_01167 (THI 04)	Gene „SPL_01167“ (EMBL-Bank: OSY47449)	SPL_01167 (NdeI,XhoI) cloned in pET24b	this work
pET28a-SPL_01186 (ANW 10)	Gene „SPL_01186“ (EMBL-Bank: OSY47468)	SPL_01186 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_01317 (THI 05)	Gene „SPL_01317“ (EMBL-Bank: OSY47218)	SPL_01317 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_01542 (ANW 11)	Gene „SPL_01542“ (EMBL-Bank: OSY47114)	SPL_01542 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET28a-SPL_01673 (ANW 12)	Gene „SPL_01673“ (EMBL-Bank: OSY46937)	SPL_01673 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET28a-SPL_01682 (ANW 13)	Gene „SPL_01682“ (EMBL-Bank: OSY46946)	SPL_01682 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_01896 (THI 11)	Gene „SPL_01896“ (EMBL-Bank: OSY46529)	SPL_01896 (NdeI,XhoI) cloned in pET24b	Hilberath et al.[11]
pET28a-SPL_02232 (ANW 14)	Gene „SPL_02232“ (EMBL-Bank: OSY46264)	SPL_02232 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_02251 (THI 17)	Gene „SPL_02251“ (EMBL-Bank: OSY46283)	SPL_02251 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_02316 (THI 06)	Gene „SPL_02316“ (EMBL-Bank: OSY46348)	SPL_02316 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_02318 (THI 07)	Gene „SPL_02318“ (EMBL-Bank: OSY46350)	SPL_02318 (NdeI,XhoI) cloned in pET24b	Hilberath et al.[11]
pET28a-SPL_02943 (THI 12)	Gene „SPL_02943“ (EMBL-Bank: OSY45653)	SPL_02943 (NcoI,XhoI) cloned in pET28a	this work

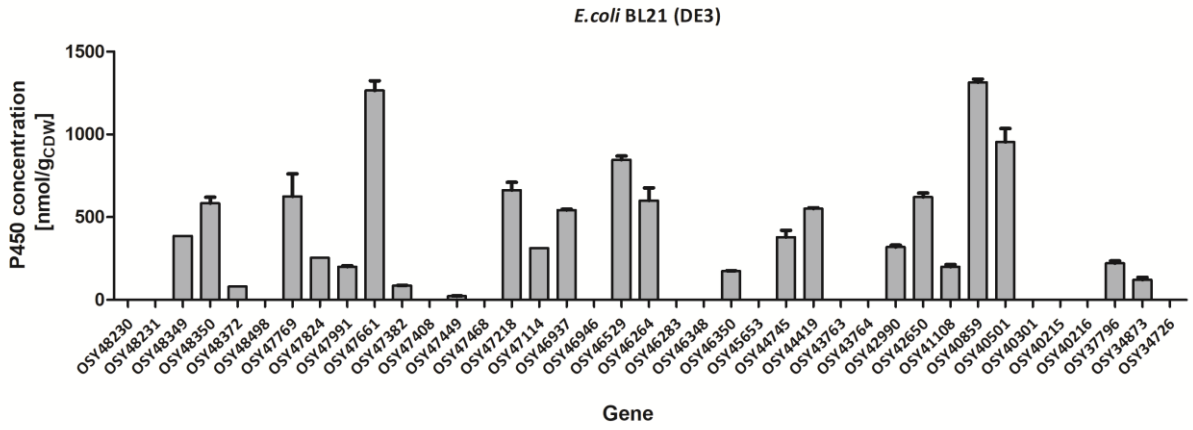
vector (internal number)	genes with EMBL bank number	vector properties	reference/source
pET24b-SPL_03767 (THI 13)	Gene „SPL_03767“ (EMBL-Bank: OSY44745)	SPL_03767 (NdeI,XhoI) cloned in pET24b	Hilberath et al.[11]
pET24b-SPL_04201 (ANW 15)	Gene „SPL_04201“ (EMBL-Bank: OSY44419)	SPL_04201 (NdeI/SalI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_04452 (THI 20)	Gene „SPL_04452“ (EMBL-Bank: OSY43763)	SPL_04452 (NheI,XhoI) cloned in pET24b	this work
pET24b-SPL_04453 (THI 08)	Gene „SPL_04453“ (EMBL-Bank: OSY43764)	SPL_04453 (NdeI,XhoI) cloned in pET24b	this work
pET28a-SPL_04587 (ANW 16)	Gene „SPL_04587“ (EMBL-Bank: OSY42990)	SPL_04587 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET28a-SPL_04666 (ANW 17)	Gene „SPL_04666“ (EMBL-Bank: OSY42650)	SPL_04666 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_05184 (THI 14)	Gene „SPL_05184“ (EMBL-Bank: OSY41108)	SPL_05184 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_05261 (ANW 18)	Gene „SPL_05261“ (EMBL-Bank: OSY40859)	SPL_05261 (NdeI/SalI) cloned in pET24b	Worsch et al.[108a]
pET28a-SPL_05402 (ANW 19)	Gene „SPL_05402“ (EMBL-Bank: OSY40501)	SPL_05402 (NcoI/XhoI) cloned in pET28a	Worsch et al.[108a]
pET24b-SPL_05473 (THI 18)	Gene „SPL_05473“ (EMBL-Bank: OSY40301)	SPL_05473 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_05497 (THI 15)	Gene „SPL_05497“ (EMBL-Bank: OSY40215)	SPL_05497 (NdeI,XhoI) cloned in pET24b	this work
pET24b-SPL_05498 (ANW 20)	Gene „SPL_05498“ (EMBL-Bank: OSY40216)	SPL_05498 (NdeI/EcoRI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_06346 (THI 16)	Gene „SPL_06346“ (EMBL-Bank: OSY37796)	SPL_06346 (NdeI,XhoI) cloned in pET24b	Hilberath et al.[11]
pET24b-SPL_07372 (ANW 21)	Gene „SPL_07372“ (EMBL-Bank: OSY34726)	SPL_07372 (NdeI/SalI) cloned in pET24b	Worsch et al.[108a]
pET24b-SPL_07341 (THI 19)	Gene „SPL_07341“ (EMBL-Bank: OSY34873)	SPL_07341 (NdeI,XhoI) cloned in pET24b	this work

LC and GC conditions

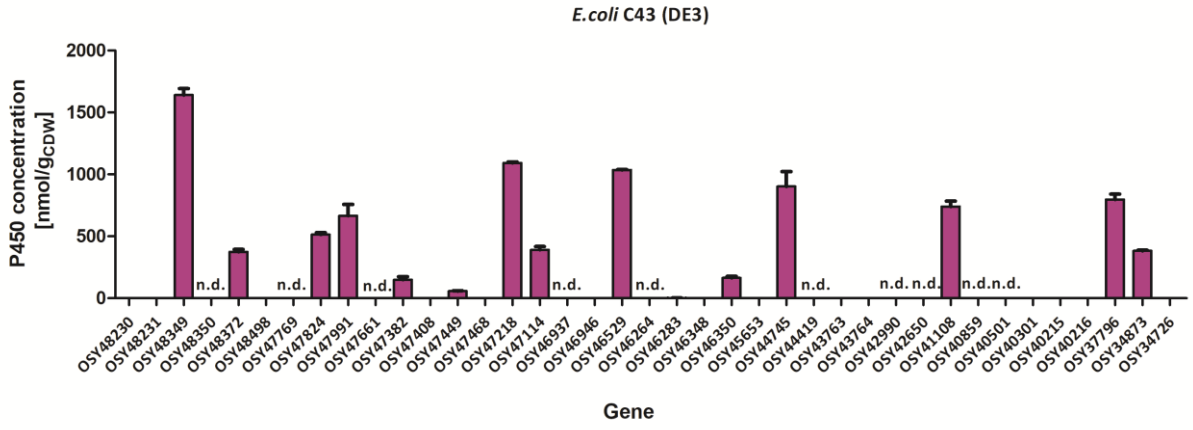
Supplementary Table 2.3.3: Chromatographic conditions for the LC/MS-or GC/MS-analysis.

Compound	Internal standard	Solvent program	Reference
Ritonavir	Darunavir	Gradient from 30 % B to 80 % B for 20 min, then gradient from 80 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 30 % B for 5 min Solvent A: 0.1% formic acid; solvent B: methanol	Worsch et al.[108a]
Testosterone	Progesterone	Gradient from 10 % B to 75 % B for 10 min, hold 75 % B for 5 min, equilibration at 10 % B for 5 min Solvent A: 0.1% formic acid; solvent B: methanol	von Bühler et al.[79a]
Amitriptyline	Oxcarbazepine	Gradient from 23 % B to 33 % B for 10 min, hold 100 % B for 1 min, equilibration at 23 % B for 4 min Solvent A: 0.1% formic acid; solvent B: acetonitrile	Hilberath et al.[11]
Amodiaquine	Quinidine	Gradient from 5 % B to 33 % B for 7 min, 33 % B to 100 % B for 1 min, hold 100 % B for 1 min, re-equilibration at 5 % B for 7 min Solvent A: 0.1% formic acid; solvent B: methanol	Worsch et al.[108a]
Terfenadine	Diphenhydramine	Gradient from 45 % B to 80 % B for 20 min, then gradient from 80 % B to 100 % B for 1 min, hold 100 % B for 1 min, equilibration at 45 % B for 5 min Solvent A: 0.1% formic acid; solvent B: methanol	This work
Thioridazine	Amitriptyline	Gradient from 45 % B to 80 % B for 20 min, 80 % B to 100 % B for 1 min, hold 100 % B for 1 min, re-equilibration at 45 % B for 5 min Solvent A: 0.1% formic acid; Solvent B: methanol	This work
Geraniol	/	starting column temperature at 120 °C, hold 120 °C for 3 min, and temperature increase to 300 °C at a rate of 30 °C min^{-1}, injection at 250 °C.	Rühlmann et al.[121]
Dextromethorphan	/	Gradient from 35 % B to 70 % B for 20 min, 70 % B to 100 % B for 1 min, hold 100 % B for 2 min, re-equilibration at 35 % B for 5 min Solvent A: 0.1% formic acid; Solvent B: methanol	This work

Supplementary Results

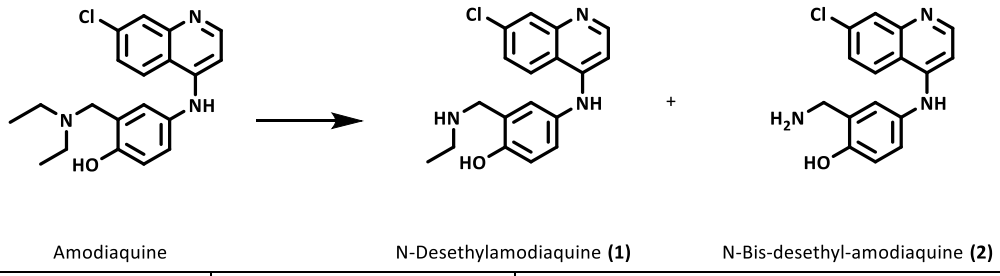


Supplementary Figure 2.3.1: P450 concentrations achieved in *E. coli* BL21 (DE3). Measurements were performed in biological duplicates in case a standard deviation is given. Data were plotted with GraphPad Prism.



Supplementary Figure 2.3.2: P450 concentrations achieved in *E. coli* C43 (DE3). Measurements were performed in biological duplicates in case a standard deviation is given. Data were plotted with GraphPad Prism. n.d.: not determined

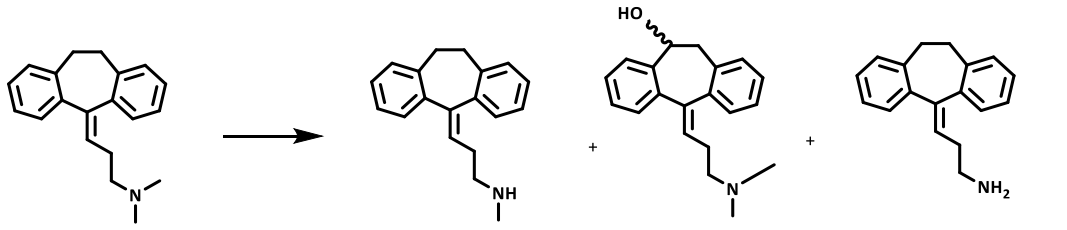
Supplementary Table 2.3.4: Results on product distribution and substrate conversion with amodiaquine. Retention times: Amodiaquine (6.4 min, 356 *m/z*), N-Desethylamodiaquine **1** (5.9 min, 328 *m/z*) and N-Bis-desethyl-amodiaquine **2** (5.4 min, 300 *m/z*). Conversion was calculated using the internal standard quindine (7.8 min, 326 *m/z*). Product identification was described previously.[108a]



Amodiaquine N-Desethylamodiaquine **(1)** N-Bis-desethyl-amodiaquine **(2)**

		Product distribution [%]	
P450	**conversion [%]**	**(1)**	**(2)**
OSY47769	≤ 10	100	0
OSY47991	≤ 10	100	0
OSY47218	≤ 10	100	0
OSY46937	≤ 10	100	0
OSY46529	13	100	0
OSY44745	64	89	11
OSY40859	76	72	28
OSY37796	25	100	0

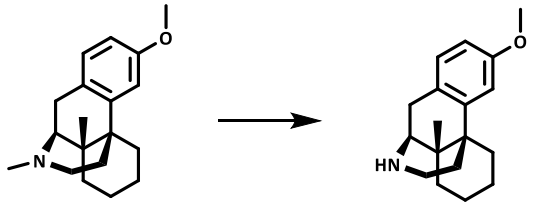
Supplementary Table 2.3.5: Results on product distribution and substrate conversion with amitriptyline. Retention times: Amitriptyline (18.3 min, 278 *m/z*), Nortriptyline **3** (17.4 min, 264 *m/z*), (±)-(*E*)-10-Hydroxyamitriptyline **4** (6.0 min, 294 *m/z*) and Desmethylnortripyline **5** (16.2 min, 250 *m/z*). Conversion was calculated using the internal standard oxcarbazepine (11.4 min, 253 *m/z*). Product identification was described previously.[11]



Amitriptyline Nortriptyline **(3)** (±)-(*E*)-10-Hydroxyamitriptyline **(4)** Desmethylnortriptyline **(5)**

		Product distribution [%]			
P450	**conversion [%]**	**(3)**	**(4)**	**(5)**	**other**
OSY48372	13	88	10	0	2
OSY47824	≤ 10	92	8	0	0
OSY47991	21	79	15	≤ 1	6
OSY47218	75	87	2	3	8
OSY47114	≤ 10	74	26	0	0
OSY46937	16	85	10	0	5
OSY46529	34	49	26	1	24
OSY44745	86	56	1	17	26
OSY42650	≤ 10	100	0	0	0
OSY40859	50	47	18	0	35
OSY37796	97	55	12	7	26

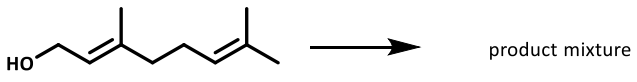
Supplementary Table 2.3.6: Results on product distribution and substrate conversion with dextromethorphan. Retention times: Dextromethorphan (10.5 min, 272 *m/z*) and 3-Methoxymorphinan **6** (11.4 min, 258 *m/z*). Conversion was determined from the integration of peak areas of the substrate and all products. Products were identified by co-elution with authentic reference (Supplementary Figure 2.3.3).



Dextromethorphan 3-Methoxymorphinan **(6)**

		Product distribution [%]	
P450	**conversion [%]**	**(6)**	**other**
OSY47991	≤ 10	76	24
OSY47114	≤ 10	100	0
OSY46937	≤ 10	73	27
OSY46529	39	93	7
OSY44745	93	44	56
OSY37796	78	77	23

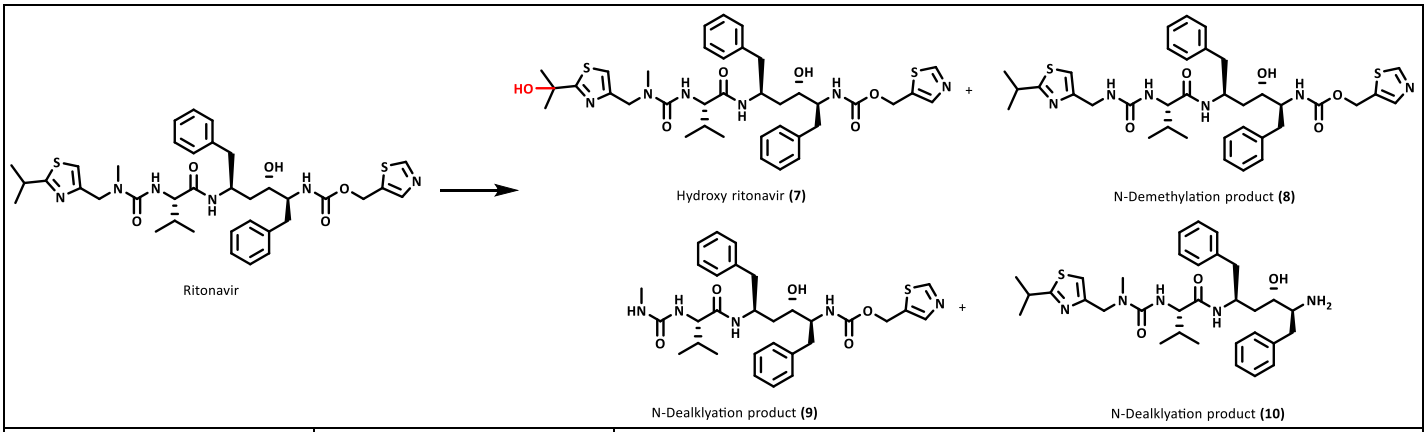
Supplementary Table 2.3.7: Results on substrate conversion with geraniol. Retention times: Geraniol (4.3 min). Conversion was estimated from the integration of peak areas of the substrate and putative products. It was not possible to calculate the product distribution due to the similar retention times of several products.



Geraniol

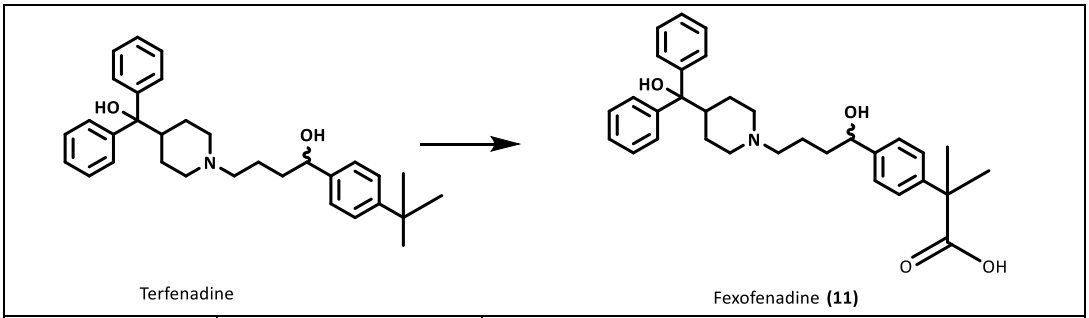
		Product distribution [%]
P450	**conversion [%]**	**other**
OSY47991	≤ 10	not determined
OSY46529	≤ 10	not determined
OSY46350	≤ 10	not determined
OSY44745	≤ 10	not determined
OSY37796	≈80	not determined
OSY34873	≈95	not determined

Supplementary Table 2.3.8: Results on product distribution and substrate conversion with ritonavir. Retention times: Ritonavir (19.0 min, 721 *m/z*), Hydroxy ritonavir **7** (15.6 min, 737 *m/z*), N-Demethylation product **8** (17.9 min, 707 *m/z*), N-Dealklyation product **9** (13.6 min, 583 *m/z*) and N-Dealklyation product **10** (14.5 min, 581 *m/z*). Conversion was calculated using the internal standard darunavir (12.1 min, 549 *m/z*). Product identification was described previously.[11] [108a]



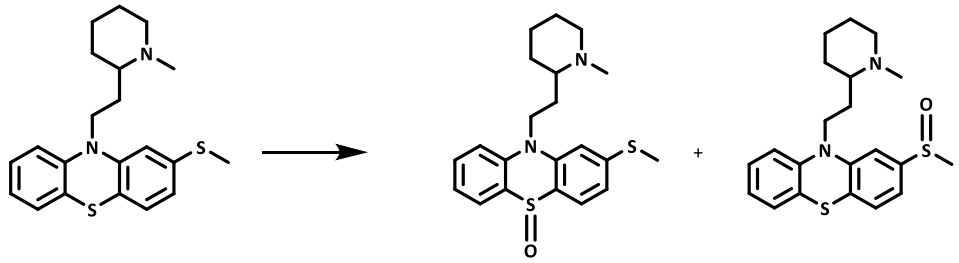
		Product distribution [%]				
P450	**conversion [%]**	**(7)**	**(8)**	**(9)**	**(10)**	**other**
OSY47991	68	86	0	8	0	6
OSY47218	14	39	22	6	21	12
OSY46529	92	84	0	3	0	13
OSY44745	46	36	19	14	5	26
OSY40859	86	84	0	1	≤ 1	14
OSY37796	41	30	32	14	13	11

Supplementary Table 2.3.9: Results on product distribution and substrate conversion with terfenadine. Retention times: Terfenadine (13.8 min, 472 *m/z*) and Fexofenadine **11** (8.5 min, 502 *m/z*). Conversion was calculated using the internal standard diphenhydramine (5.7 min, 256 *m/z*)..Products were identified by co-elution with authentic reference (unpublished data, Anne Worsch).



		Product distribution [%]	
P450	**conversion [%]**	**(11)**	**other**
OSY47769	≤ 10	0	100
OSY47991	24	0	100
OSY46529	≤ 10	0	100
OSY44745	95	0	100
OSY42990	≤ 10	0	100
OSY40859	97	87	13
OSY37796	92	0	100

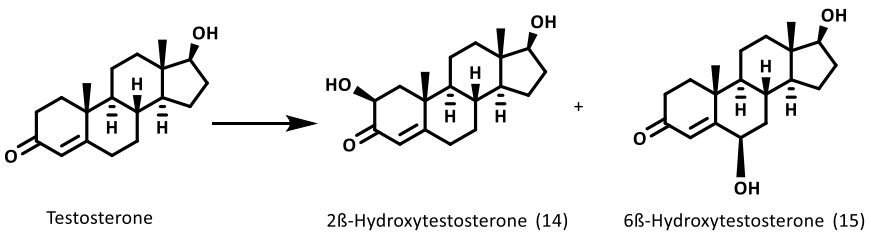
Supplementary Table 2.3.10: Results on product distribution and substrate conversion with thioridazine. Retention times: Thioridazine (20.3 min, 371 *m/z*), Thioridazine 5-sulfoxide **12** (8.5 min, 387 *m/z*) and Mesoridazine **13** (9.5 min, 387 *m/z*). Conversion was calculated using the internal standard amitriptyline (13.6 min, 278 *m/z*). Products were identified by co-elution with authentic reference (Supplementary Figure 2.3.4).



Thioridazine | Thioridazine 5-sulfoxide **(12)** | Mesoridazine **(13)**

		Product distribution [%]		
P450	**conversion [%]**	**(12)**	**(13)**	**other**
OSY48372	15	44	27	29
OSY47991	32	65	35	0
OSY47218	93	49	21	30
OSY46937	14	62	31	7
OSY46529	45	27	70	3
OSY44745	25	68	19	13
OSY42650	≤ 10	65	35	0
OSY40859	63	6	68	26
OSY37796	91	41	32	27

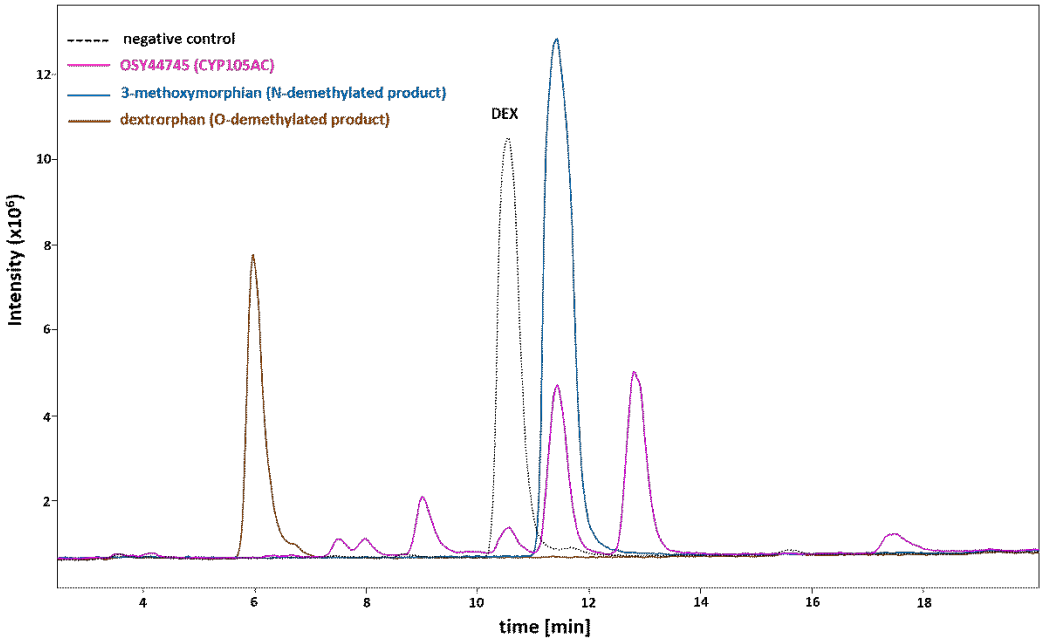
Supplementary Table 2.3.11: Results on product distribution and substrate conversion with testosterone. Retention times: Testosterone (12.2 min, 290 *m/z*), 2β-Hydroxytestosterone **14** (11.3 min, 306 *m/z*) and 6β-Hydroxytestosterone **15** (10 min, 306 *m/z*). Conversion was calculated using the internal standard progesterone (13.8 min, 316 *m/z*). Product identification was described previously.[11]



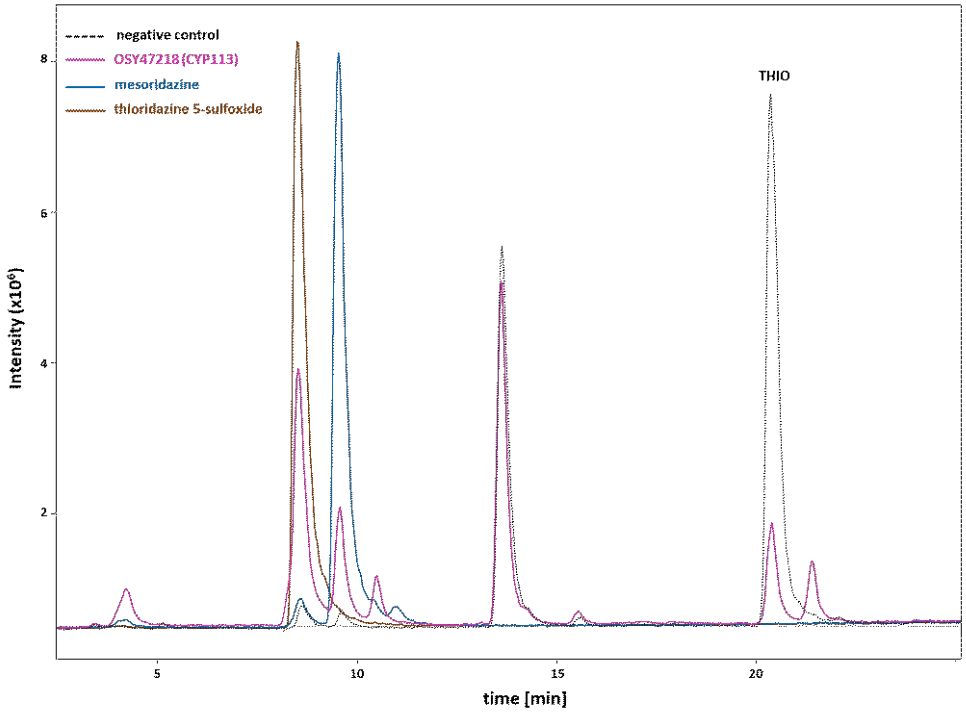
Testosterone | 2ß-Hydroxytestosterone (14) | 6ß-Hydroxytestosterone (15)

		Product distribution [%]		
P450	**conversion [%]**	**(14)**	**(15)**	**other**
OSY47824	≤ 10	0	0	100
OSY47991	100	70	0	30
OSY46937	97	0	0*	100
OSY46529	≤ 10	12	0	88
OSY40501	≤ 10	0	0	100
OSY37796	98	5	6	89

* Due to the similar retention times of several products, it could not be confirmed whether 6ß-hydroxytestosterone was formed by this P450.



Supplementary Figure 2.3.3: Comparison of the commercial product standards 3-methoxymorphian (blue) and dextorphan (brown) with the conversion of dextromethorphan (black). Dex: dextromethorphan, Black: control reaction without P450, pink: conversion of dextromethorphan with OSY44745 (CYP105AC).



Supplementary Figure 2.3.4: Comparison of the commercial product standards mesoridazine (blue) and thioridazine 5-sulfoxide (brown) with the conversion of thioridazine (pink). THIO: thioridazine, Black: control reaction without P450, pink: conversion of thioridazine with OSY47218 (CYP113).

2.4 Bile acid oxidation by CYP105D

Title: Insights into bile acid oxidation by a promiscuous bacterial cytochrome P450

Authors: **Thomas Hilberath**, Matthias Bureik, Vlada B. Urlacher*

* corresponding author

Manuscript in preparation

Own contribution: Conceptualization, design and conduction of all experiments, analysis and interpretation of all data, drafting of the manuscript. Relative contribution: 90%.

Insights into bile acid oxidation by a promiscuous bacterial cytochrome P450

Thomas Hilberath[1], Matthias Bureik[2], Vlada B. Urlacher[1]*

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

[2]School of Pharmaceutical Science and Technology, Health Science Platform, Tianjin University, Tianjin 300072, China

* Corresponding author: Vlada B. Urlacher. Email: vlada.urlacher@uni-duesseldorf.de

Abstract

The synthesis of hydroxylated bile acids can lead to the development of new drug targets to cure various metabolic disorders including cholestasis or gallstone disease. Chenodeoxycholic acid is a bile acid with the highest physiological relevance and thus a promising starting substrate to synthesize hydroxylated derivatives. In this study, we identified the cytochrome P450 CYP105D from *Streptomyces platensis* DSM 40041 as a general biocatalyst for the oxidation of structurally diverse bile acids. CYP105D oxidized 0.5 mM chenodeoxycholic acid completely to a main hydroxylated product with 55% selectivity. Testing of bile acids with different functional moieties at different positions of the tetracyclic hydrocarbon core revealed an influence of hydroxy- and keto groups on substrate binding, enzyme activity and product selectivity. Product analysis by NMR demonstrated the preference of CYP105D for the side chain hydroxylation of chenodeoxycholic acid, which has not been reported for any other P450 before.

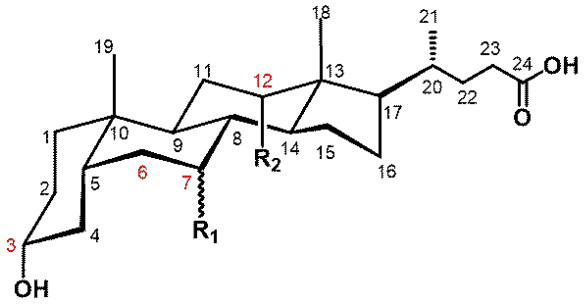
1. Introduction

Bile acids are tetracyclic steroid acids derived from cholesterol, which act as emulsifiers in the small intestine to dissolve hydrophobic substances such as lipids and fat-soluble vitamins.[139] They play a pivotal role as signaling molecules regulating members of the nuclear receptor family such as the farnesoid x receptor (FXR) or G-protein-coupled receptors (GPCRs) like TGR5.[140] They do not only modulate their own synthesis and enterohepatic recirculation, but also triglyceride, cholesterol, energy and glucose homeostasis.[141] Hence, a deficiency or misleading function in bile acid metabolism is associated with a number of diseases including nonalcoholic steatohepatitis,[142] primary biliary cirrhosis,[141, 143] and cholesterol gallstone disease.[144] In order to cure these diseases, bile acids have been exploited as drug targets. Five bile acids are mainly found in humans, namely chenodeoxycholic acid (CDCA), deoxycholic acid (DCA), cholic acid (CA), ursodeoxycholic acid (UDCA) and lithocholic acid (LCA) (Figure 2.4.1). In humans chenodeoxycholic acid is the most potent endogenous FXR ligand and together with cholic acid, it makes up to 80% of all bile acids.[145] Accordingly, derivatives of chenodeoxycholic acid are of great pharmaceutical interest. These derivatives include hydroxylated products such as muricholic acids, which possess an additional alcohol group at position C6. Muricholic acids and other oxy-functionalized bile acids are more hydrophilic making them interesting targets for preventing or treating cholesterol gallstones.[146] However, the synthesis of hydroxylated bile acids is a challenging task because tedious or potentially toxic, chemical or chemo-enzymatic reactions are required, and enzymes for hydroxylation remain unidentified.[147]

The hydroxylation of bile acids is naturally catalyzed by heme-containing cytochrome P450 monooxygenases. P450s are powerful biocatalysts due to their ability to oxidize a tremendous number of substrates in one step, often in a regio-and stereoselective manner under mild reaction conditions.[19b, 148] Bacterial P450s are especially interesting as they can be produced in high amounts in heterologous hosts and usually exhibit higher activities compared to their eukaryotic counterparts.[15] In a previous study we found that CYP105D from *Streptomyces platensis* DSM 40041 catalyzed the oxidation of a variety of chemically

diverse substrates. Among the accepted substrates, steroids like testosterone were oxidized with high regioselectivity leading to one main hydroxylated product.[11]

In the present study, we describe the identification and characterization of CYP105D from *S. platensis* regarding the oxidation of chenodeoxycholic acid and other bile acids. CYP105D could oxidize chenodeoxycholic acid *in vitro* mainly to one hydroxylated product. The binding constant K_D of CYP105D towards bile acids increased with a higher number of hydroxy- and keto groups. Activity and selectivity of CYP105D towards bile acids was strongly influenced by the position and molecular configuration of hydroxy and keto groups. The conditions for the *in vitro* reactions were optimized enabling complete conversion of 1 mM substrate concentration at a semi-preparative scale. CYP105D was found to have a unique selectivity in bile acid oxidation and thus is a promising bacterial P450 to form unusual oxy-functionalized bile acids.



	R_1	R_2
Chenodeoxycholic acid	OH_α	H
Cholic acid	OH_α	OH
Lithocholic acid	H	H
Deoxycholic acid	H	OH
Ursodeoxycholic acid	OH_β	H

Figure 2.4.1: Structure of the five common bile acids in humans. Common positions for hydroxy groups are highlighted in red.

2. Material and methods

2.1 Bacterial strains, enzymes and chemicals

E. coli strains BL21(DE3) and OverExpress C43(DE3) were obtained from Clontech and Lucigen. Catalase from *bovine* was purchased from Sigma Aldrich. Bile acid substrates were obtained from Sigma Aldrich, Acros Organics, Alfa Aesar and Serva. Muricholic acids used as reference metabolites were provided by Cayman Chemical.

2.2 Recombinant expression and purification of CYP105D, its redox partners and glucose dehydrogenase

All genes used in this study were cloned into the pET vectors allowing expression in *E. coli* (DE3) expression strains. Suitable conditions for expression of *cyp105D* from *Streptomyces platensis* DSM 40041 (EMBL-Bank: OSY47991),[11] flavodoxin reductase from *E. coli* (Fdr, EMBL-Bank: L04757), flavodoxin from *B. subtilis* (YkuN, EMBL-Bank: CAA10877),[108a, 119] and glucose dehydrogenase from *Bacillus megaterium* (GDH, EMBL-Bank: D10626)[120] are described elsewhere and were performed accordingly. All proteins except GDH contained an N-terminal His-tag and were purified via immobilized affinity chromatography using two HisTrap FF crude columns in series (5 mL, GE-Healthcare) following instructions described elsewhere.[11] GDH was purified via salt precipitation and heat treatment according to procedures described by Kranz Finger et al. and Nagao et al.[149]

2.3 Spectroscopic methods and enzyme assays

All photometric measurements were conducted using a double-beam photometer (Lambda 35, Perkin Elmer). Concentration of purified CYP105D was determined based on its absorption in CO-difference spectra using the published extinction coefficient ε_{450} = 91 mM^{-1} cm^{-1}.[14b] The concentration of YkuN and Fdr was obtained by using their extinction coefficients ε_{461}= 10 mM^{-1} $cm^{-}1$ and ε_{456}= 7.1 mM^{-1} cm^{-1}.[150] For determination of GDH activity, the reduction of $NADP^+$ to NADPH by GDH was tracked at 340 nm in a continuous photometric assay. Reaction mixtures contained 100 mM glucose, 100 µM $NADP^+$ and 100 µL of enzyme in 1 mL 50 mM potassium phosphate buffer (KPi), pH 7.5. The increase of absorption caused by NADPH formation was tracked for 120 s at 25 °C. The initial slope

(ΔA_{340}/min) between 20 - 80 s was linear and used to calculate the enzymatic activity [U/mL]: The activity of GDH is given in Units (U) where 1 U is defined as the amount of enzyme which is needed to convert 1 µmol substrate in 1 minute under assay conditions.

2.4 Substrate binding assays

Titration of a P450 with a substrate results in a continuous spectral shift of the soret band from 418 nm (low spin substrate free form) to 390 nm (high spin substrate bound form).[151] This spectral property was used to evaluate binding of bile acid substrates to CYP105D and to calculate the substrate dissociation constants (K_D). In a double-beam photometer (Lambda 35, Perkin Elmer) two tandem quartz cuvettes (Hellma) were placed. Each cuvette contained one chamber with 800 µL 1 µM purified CYP105D in 50 mM potassium phosphate buffer, pH 7.5 (chamber 1) and another chamber with buffer without enzyme (chamber 2). The bile acids were dissolved in ethanol and titrated in small volumes (0.5 – 3 µL) to the enzyme containing chamber of the 'sample cuvette' and in parallel to the buffer chamber of the "reference cuvette". After gentle mixing with a plastic spatula and incubation for 1 min at 25 °C, a spectrum between 330 and 500 nm was recorded. The addition of substrate was continued until no further increase in the absorption difference between 390 nm and 420 nm (Type I spectrum) occurred. K_D values were obtained by fitting the absorption difference versus the substrate concentration to the following hyperbolic function using Prism (GraphPad, version 5):

$$\Delta A = \frac{\Delta A_{max} \cdot [S]}{K_D + [S]}$$

ΔA = absorption difference between 390 nm and 420 nm

[S] = concentration of bile acids

ΔA_{max}= maximum absorption difference

2. 5 Analytical scale conversions

In vitro conversions were conducted with purified enzymes in 0.2 – 1 mL 50 mM potassium phosphate buffer, pH 7.5. For reconstitution of P450 activity, the non-physiological redox partners Fdr from *E. coli* together with the flavodoxin YkuN from *B. subtilis* were applied.

Reaction mixtures usually contained CYP105D (2 – 5 µM), YkuN (20 - 50 µM), Fdr (2 -5 µM), catalase from *bovine* (600 U/mL) and 0.2 – 1 mM substrate dissolved in 2 % (v/v) organic co-solvent. NADPH was added in stoichiometric amounts (200 µM) and regenerated with 5 U/mL purified GDH using 20 mM glucose as co-substrate. Samples were incubated at 25 °C in 2 mL reaction tubes with open lids in an Eppendorf thermomixer with a shaking speed of 300 - 600 min^{-1}. For extraction 200 µM of internal standard were added. Next, the sample was supplemented with 300 µL 0.1 M sodium acetate buffer, pH 4. After extraction with ethyl acetate, the organic phase was transferred into a new reaction tube and dried under reduced pressure. Samples were dissolved in methanol (LC/MS grade) and subjected to LC/MS-analysis.

2.7 Semi-preparative scale conversions and product isolation

For isolation of the oxidized products, conversions with chenodeoxycholic acid were performed similar to analytical scale conversions in 1 mL scale. Reactions were conducted in 48 mL final volume with 4 µM CYP105D, 40 µM YkuN, 4 µM Fdr and 1 mM substrate (19 mg). The reaction mixture was incubated at 25 °C and 600 min^{-1} in an Eppendorf thermomixer for 4 h. Next, samples for LC/MS analysis were taken, reactions pooled to 10 mL volume, and acidified with 15 mL 0.1 M sodium acetate buffer, pH 4. The reaction mixtures were extracted twice with ethyl acetate (2 x 20 mL), organic layers combined, dried over $MgSO_4$ and evaporated under reduced pressure. Isolation was conducted on a silica column via flash column chromatography starting with chloroform: methanol in a ratio of 9:1. During purification, the concentration of methanol was stepwise increased. Two metabolites (P1 and P2) could be isolated in sufficient concentration allowing first structural assignments via NMR-analysis. Metabolite P1 (3.3 mg, 17%) was isolated as white transparent solid in a LC/MS purity of ≥ 80%, while metabolite P2 was isolated with ≥ 60% purity (LC/MS) as white solid (4 mg, 21%).

2.9 Product analysis

Conversions were analyzed by liquid chromatography coupled to mass spectrometry (LC/MS) on a Prominence/LCMS 2020 device (Shimadzu). Analytes were separated on a Chromolith® Performance RP-18e column (100 x 4.6 mm, Merck) using water with 0.1% formic acid as

solvent A and methanol (MeOH) as solvent B. 1 µL of each sample was injected and the analytes were separated by a gradient with a flow rate of 1 mL/min at 30 °C. All bile acids were separated with the same program except dehydrocholic acid (Table 2.4.S1) Bile acids were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI) and detected in a range between 250-900 *m/z*.

NMR spectra (1D and 2D-spectra) were recorded on a Bruker Avance III-600 spectrometer (^{1}H-NMR: 600 mHz; ^{13}C-NMR: 150 mHz). The chemical shifts (δ) are stated in ppm and are related to the solvent methanol-d_4 (^{13}C: 49.00 ppm; ^{1}H: 3.31 ppm). NMR data were recorded for the substrate and products and compared to literature allowing signal assignments.[152]

3. Results

3.1 CYP105D as a candidate for bile acid oxidation

In order to identify a suitable candidate with activity towards bile acids, we focused on CYP105D and CYP107Z from *Streptomyces platensis* DSM 40041, because of their broad substrate spectrum and their complementary activities against steroidal compounds.[11] Both enzymes were tested *in vitro* against chenodeoxycholic acid as substrate using crude cell extracts. Since the physiological redox partners for both P450s remain unidentified, electrons from NADPH were provided by the NADPH-dependent flavodoxin reductase from *Escherichia coli* (Fdr) and the flavodoxin YkuN from *Bacillus subtilis*.[150, 153] Initially, the activity of both P450s was very low due to the appearance of an undesired by-product, which was also present in the negative control (data not shown). The expression host for all enzymes was *E. coli*, which is able to metabolize bile acids on its own due to activity of 7α-hydroxysteroid dehydrogenase (7α-HSDH).[154] 7α-HSDH from *E. coli* oxidizes bile acids at the 7α-hydroxy group to the corresponding ketone and produces undesired by-products in bile acid syntheses as reported earlier.[155] Thus, all enzymes including the P450 were purified avoiding bile acid oxidation by 7α-HSDH and tested again for activity against chenodeoxycholic acid (Figure 2.4.2). Both P450 enzymes oxidize chenodeoxycholic acid with high activity. CYP105D oxidized chenodeoxycholic acid almost completely (97% conversion)

and outperformed CYP107Z showing a slightly lower conversion of approximately 85%. Interestingly, both enzymes showed a similar product pattern. The main product (**P2**) was the same for both enzymes and accounted for more than half of the total product percentage. Because of its higher activity, CYP105D was further investigated.

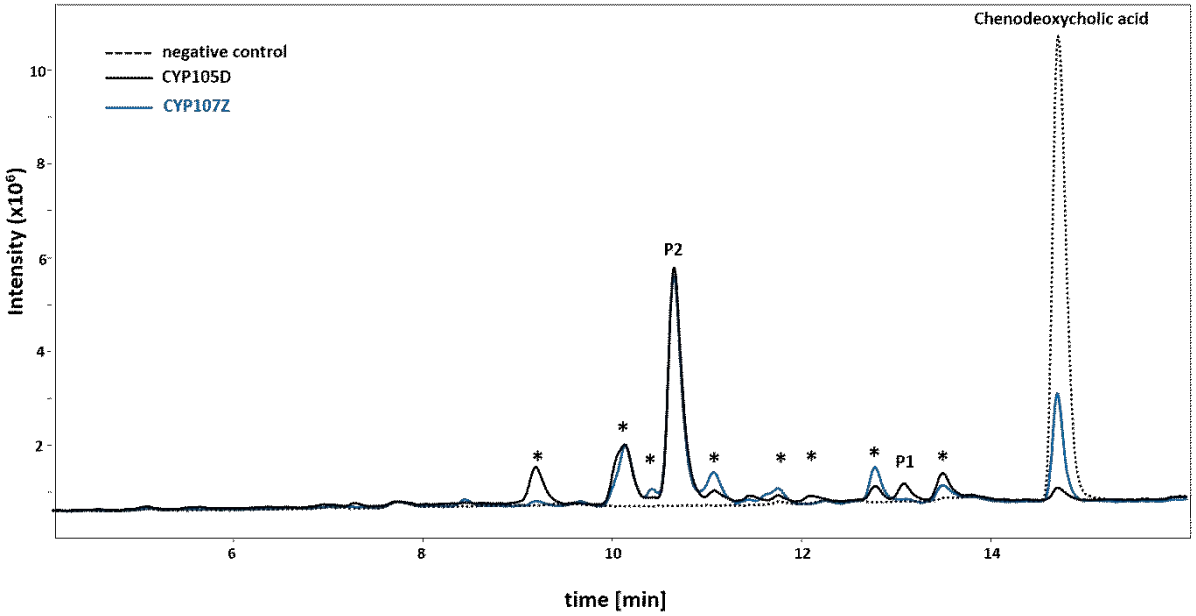


Figure 2.4.2: LC/MS-chromatograms for the conversion of chenodeoxycholic acid by CYP105D (black line) and CYP107Z (blue line). Products analyzed by NMR are numbered, while all other products are indicated with an (*). Reaction conditions: 0.2 mM chenodeoxycholic acid in 2 % (v/v) DMSO, 2 µM P450 together with 10 µM Fdr and 10 µM YkuN. NADPH (0.2 mM) was regenerated with 5 U/mL GDH and 20 mM glucose. Incubation at 25 °C and 300 min^{-1} shaking frequency for 4 h in 0.2 mL potassium phosphate buffer (50 mM KPi, pH 7.5). Conversion values were calculated from the ratio of the total product peak area relative to the sum of product and substrate peak areas.

3.2 Influence of the co-solvent and increasing substrate concentrations

Bile acids are amphipathic molecules that form micelles above their critical micelle concentration (cmc). Below their cmc of around 5-15 mM, bile acids are monomeric and thus almost insoluble in water.[147a, 156] Initially, chenodeoxycholic acid was dissolved in dimethyl sulfoxide (DMSO) and used for conversions at a concentration of 0.2 mM. To increase the activity of CYP105D against chenodeoxycholic acid, different co-solvents and increasing substrate concentrations were tested. Besides DMSO, acetonitrile (ACN), propan-2-ol (i-Prop) and ethanol (EtOH) were applied (Figure 2.4.3). The reaction time was set to four hours as this was the shortest time for maximum conversion (Supplementary Figure 2.4.1).

Using a substrate concentration of 0.5 mM, a conversion of 69 ± 2% was reached using DMSO as co-solvent. Due to the low solubility of chenodeoxycholic acid in ACN, only half of the substrate concentration could be used and thus the lowest conversion was achieved using this compound as co-solvent (50%). When i-Prop (87 ± 5% conversion) and EtOH (92 ± 9% conversion) were applied as co-solvent, higher conversions could be achieved in the same reaction time as compared to reactions with DMSO. Due to the higher P450 activity with EtOH, this co-solvent was used to increase the substrate concentration in aqueous solution (Figure 2.4.4). The doubling of the substrate concentration from 0.5 mM to 1 mM resulted in a more than 3-fold decrease of substrate conversion. The further increase of substrate concentration resulted in almost complete loss of P450 activity (conversion less than 1% at 2.5 mM substrate concentration). This suggests a potential inhibitory effect of chenodeoxycholic acid on CYP105D at higher substrate concentrations.

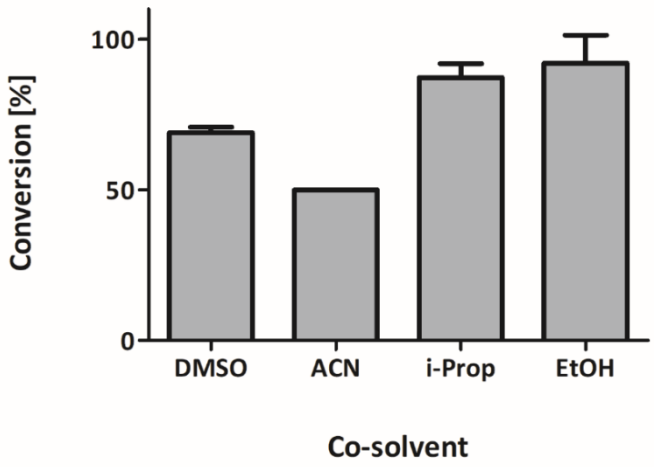


Figure 2.4.3: Conversion [%] of 0.5 mM chenodeoxycholic acid dissolved in 2 vol% (final concentration) of dimethyl sulfoxide (DMSO), acetonitrile (ACN), propan-2-ol (i-Prop) or ethanol (EtOH). Reaction conditions: 0.5 mM chenodeoxycholic acid in 2 % (v/v) co-solvent, 2 µM P450 together with 2 µM Fdr and 20 µM YkuN. NADPH (0.2 mM) was regenerated with 5 U/mL GDH and 20 mM glucose. The reaction was performed for 4 h in 200 µL 50 mM phosphate buffer (50 mM KPi, pH 7.5) in 2 mL reaction tubes with open lids at 25 °C and 300 min^{-1} in Eppendorf shakers. Due to the low solubility of chenodeoxycholic acid in ACN, only half of the substrate concentration than for the other solvents could be applied. Mean values and standard deviation are derived from at least two technical duplicates.

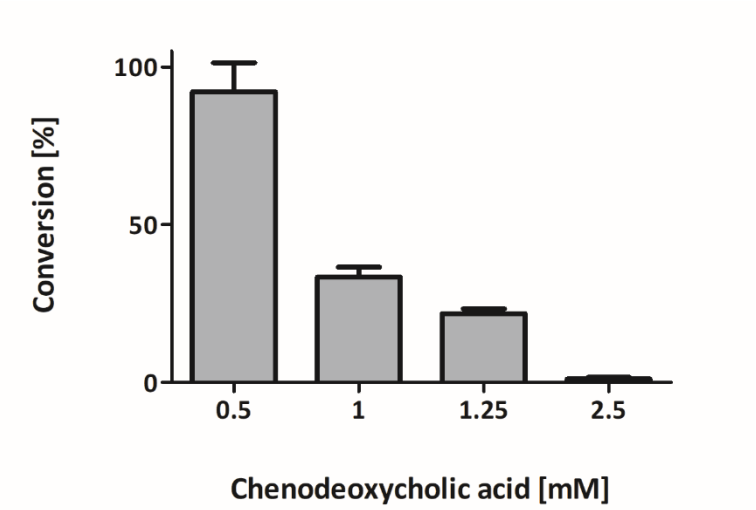


Figure 2.4.4: Conversion [%] of 0.5 – 2.5 mM chenodeoxycholic acid dissolved in 2 - 5 vol% EtOH (final concentration). Reaction conditions: 0.5 – 2.5 mM chenodeoxycholic acid in 2 – 5 % (v/v) EtOH, 2 µM P450 together with 2 µM Fdr and 20 µM YkuN. NADPH (0.2 mM) was regenerated with 5 U/mL GDH and 20 mM glucose. The reaction was performed for 4 h in 200 µL 50 mM phosphate buffer (50 mM KPi, pH 7.5) in 2 mL reaction tubes with open lids at 25 °C and 300 min^{-1} in Eppendorf shakers. Conversion was calculated from the ratio of the product peak areas relative to the sum of substrate and product peak areas. Mean values and standard deviation are derived from at least two technical duplicates.

3.3 Effect of functional moieties on substrate binding and enzyme activity

Bile acids share the same core structure, but differ in the position and number of their functional moieties (Figure 2.4.1). Overall six commercially available bile acids including chenodeoxycholic acid were tested regarding substrate binding and enzyme activity. The simplest bile acid is lithocholic acid which possesses only one hydroxy group at position 3. Ursodeoxycholic acid is the epimer of chenodeoxycholic acid and differs in the stereochemical configuration of the OH-group at the C7-Atom. Deoxycholic acid possesses like chenodeoxycholic acid two hydroxyl-groups, but one at the C12-atom instead of the C7- atom. Cholic acid contains three hydroxyl-groups at positions 3, 7 and 12. Its derivative dehydrocholic acid is the completely oxidized form of cholic acid possessing keto- instead of hydroxyl-groups at the same positions.

In order to unravel a correlation between substrate binding and structural variations of bile acids, the dissociation constant K_D was determined. The K_D-value was obtained by a hyperbolic fit of the difference in absorption between 390 and 420 nm against the substrate concentration (Figure 2.4.5). Generally, all bile acids induced a Type I binding spectrum, which is typically for most P450 substrates (Figure S2-S6). The highest affinity and

consequently the lowest K_D-value of 2.3 ± 0.2 µM was measured for lithocholic acid (Table 2.4.1). The K_D-value increased more than six-fold for bile acids with two hydroxy groups namely chenodeoxycholic acid (16.5 ± 0.4 µM), deoxycholic acid (14.8 ± 0.8 µM) and ursodeoxycholic acid (14.1 ± 1.2 µM). The different configuration of the hydroxy groups between these bile acids did not influence the K_D-value significantly. An even weaker binding affinity and consequently higher K_D value of 52.3 ± 1.4 µM was measured with cholic acid with three hydroxy groups. The lowest binding affinity (218 ± 13 µM) towards CYP105D was measured for dehydrocholic acid containing keto- instead of hydroxy groups. In summary it can be stated that binding affinity to the P450 was higher, the fewer hydroxy groups were present in the bile acid ligand. Ligands with keto groups like dehydrocholic acid were less tightly bound than ligands with hydroxy groups indicating that the binding of bile acids to CYP105D depended on both the number and type of functional groups.

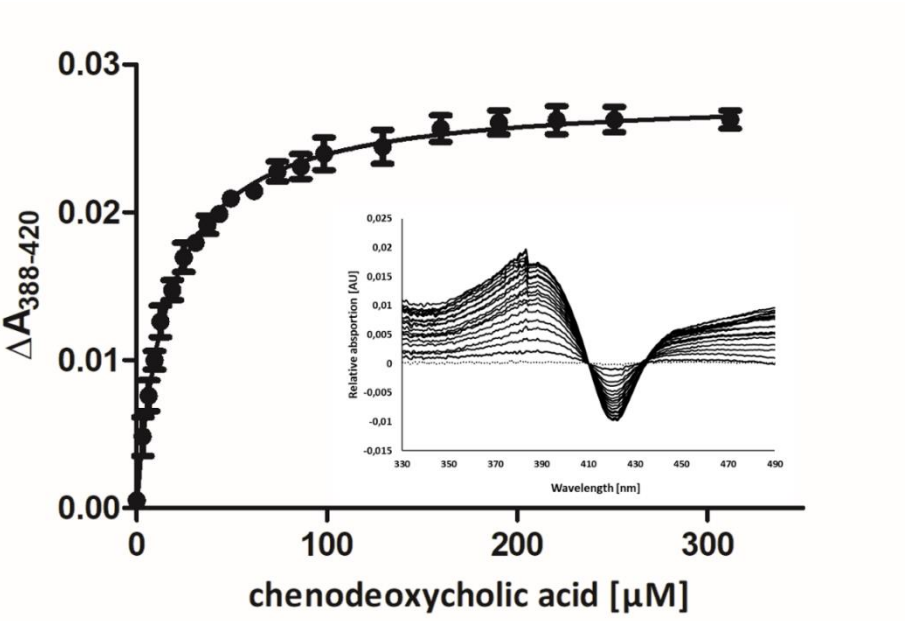


Figure 2.4.5: Type I binding of chenodeoxycholic acid to CYP105D. Peak and trough were observed at 388 and 420 nm, respectively. Measurements were performed in technical triplicates.

Complementary to the determination of the substrate dissociation constant K_D, the activity and selectivity of CYP105D against the selected bile acids was investigated (Supplementary Figures 2.4.7-11). The substrates were tested *in vitro* under optimized conditions identified before (chapter 2.4.3.3). With the exception of dehydrocholic acid all bile acids were accepted as substrates. The simplest substrate lithocholic acid carrying one hydroxy group

was oxidized with a conversion of 75 ± 6 %. Deoxycholic acid contains an additional hydroxy group at C12 and was the best substrate due to complete conversion. The constitutional isomers chenodeoxycholic acid and ursodeoxycholic acid miss the hydroxy group at the C12 position, but have a hydroxy group at the C7-position either α- or β-configurated. Both were oxidized less efficiently than deoxycholic acid (92% and 55% conversion). It is interesting that the activity between these two epimers differed by a factor of 1.5. The substrate with the lowest activity and also lowest selectivity was cholic acid (26 ± 2%). Every compound except cholic acid was converted by CYP105D to one major product accounting for more than 50% of the product ratio. Mass spectrometry indicated a monohydroxylated product due to a Δm of +16 *m/z*. Depending on the respective bile acid, up to 9 different site products were formed in small amounts. Overall, the activity and selectivity of CYP105D against bile acids was dependent on the configuration and position of hydroxy groups.

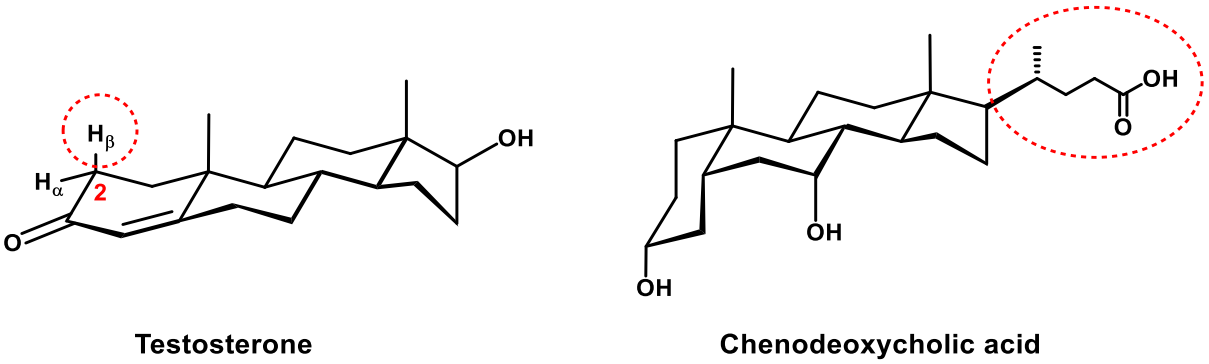
Table2.4.1: Summary of substrate binding, activities and product selectivities of CYP105D towards bile acids. No product was detected in conversions with dehydrocholic acid (n.p.d.). The percentage of the main product was calculated based on the relation of its peak area related to the sum of all product peaks areas. Conversion of 0.5 mM bile acid in 2 % vol% EtOH was performed as described in chapter 3.2.

Substrate	**Functional group**	**K_D-value [µM]**	**Conversion [%]**	**Number of products**	**Percentage of the main product [%]**
Lithocholic acid	1 -OH group	2.3 ± 0.2	75 ± 6	10	55
Chenodeoxycholic acid	2 -OH groups	16.5 ± 0.4	92 ± 9	7	55
Ursodeoxycholic acid	2 -OH groups	14.1 ± 1.2	55 ± 9	6	50
Deoxycholic acid	2 -OH groups	14.8 ± 0.8	100	10	58
Cholic acid	3 -OH groups	52.3 ± 1.4	26 ± 2	4	44
Dehydrocholic acid	3 -C=O groups	218 ± 13	n. p. d.	0	0

3.4 Product analysis reveals an oxidation at the side chain

The oxidation of chenodeoxycholic acid by CYP105D was performed on semi-preparative scale in order to isolate products for structure elucidation. The substrate concentration was set to 1 mM. Since at this concentration the activity was three times lower (Figure 2.4.4), the doubled amount of P450 and redox partners was used (4 µM CYP, 40 µM YkuN, 4 µM Fdr). Under these conditions, complete conversion was achieved. CYP105D converted chenodeoxycholic acid to give the main product P2 (21% isolated yield) and the product P1 (17% isolated yield). Both products could only be obtained in moderate purities (60 – 80%) because of co-elution of other oxidation products with similar chemical properties. Nevertheless, both compounds were further analyzed by NMR spectroscopy and mass spectrometry.

NMR spectra of both products were compared to the reported ^{13}C and ^{1}H resonance assignments of the substrate chenodeoxycholic acid.[152] The addition of hydroxy groups at individual ^{1}H resonances results in a downfield shift of adjacent proton signals, while all other ^{1}H chemical shift changes are small. By comparing the ^{1}H and ^{13}C resonance signals of products and substrate in HSQC spectra, identical or very similar correlations for most atoms of the cyclic four-ring system could be seen (except position 11, Table 2.4.S2). Thus, hydroxylation or other oxidations by the P450 at these positions can be excluded. However, this means that CYP105D must oxidize chenodeoxycholic acid at the side chain. This observation is supported by the mass spectrometric analysis of product P1. Product P1 has two characteristic masses of $[M]^+$ = 327 *m/z* and 345 *m/z*. These masses are smaller compared to the main mass fragment of the substrate of 358 *m/z* [M- OH-OH] $^+$. A lower mass of a product can only be attributed to a modification of a functional group by the enzyme. Since the hydroxy groups of chenodeoxycholic acid could be excluded by NMR-analysis, a modification of the carboxyl group of the side chain is likely. Unfortunately, the exact modification and position of oxidation by CYP105D could not be determined.



Testosterone Chenodeoxycholic acid

Figure 2.4.6: Comparison of the selectivities of CYP105D towards testosterone (left) and chenodeoxycholic acid (right). The red circle indicates the preferred position of oxidation by the enzyme.

4. Discussion

Bile acids are attractive targets for the oxyfunctionalization by P450s because they are involved in the digestion process of the human body and consequently represent valuable targets for new pharmaceuticals.[139b, 141] In the present work, CYP105D was characterized for its ability to oxidize bile acids. As a model substrate, we focused mainly on chenodeoxycholic acid because it is one of the most abundant and one of the most important bile acids in the human body.[145]

CYP105D was previously shown to oxidize the steroid testosterone to 2ß-hydroxytestosterone with 70% regioselectivity after complete conversion. Thus, CYP105D is a good candidate for the oxidation of steroid derived compounds. Bile acids were accepted as substrates as well; however they were converted by this P450 monooxygenase with lower activities and also lower regioselectivies of maximum 58% (Table 2.4.1). In comparison to other steroids, bile acids like chenodeoxycholic acid exhibit two major structural differences (Figure 2.4.6). First, all bile acids have an additional five-carbon side chain ending with a carboxyl group. Second, bile acids have a different conformation of the first ring system because of the missing double bound and ketone group present in testosterone. Additionally, the individual steroids and bile acids feature a different number of functional groups. Previous investigations on CYP105D showed that steroids with more functional groups were much less efficiently oxidized compared to the substrates with less functional groups.[11] In the case of bile acids, we could show that already small structural differences at few positions of the core have an effect on substrate binding, selectivity and activity of

CYP105D (Table 2.4.1). Whereas the number and type of functional groups is decisive for a tight binding of the substrate, the configuration at one position can strongly influence the enzyme activity as demonstrated for cheno-and ursodeoxycholic acid. The influence of functional groups on the catalytic performance of a P450 was also investigated for other enzymes including CYP199A4 from *Rhodopseudomonas palustris* or P450 BM3.[157] Here, the availability of structural data and molecular docking helped to rationalize the observed results. This approach could also be useful in the present case to interpret the available data more precisely.

Chenodeoxycholic acid was not oxidized at the first ring like testosterone, but on the side chain indicated by the NMR analysis. To the best of our knowledge, the oxidation of the side chain of chenodeoxycholic acid by CYP105D or another bacterial P450 has not been reported so far. However, side-chain hydroxylated derivatives of chenodeoxycholic acid have been found in various vertebrates, including birds, fish, and marine mammals.[139a] The functional significance of side chain hydroxylation along with the enzymes catalyzing this reaction is not known yet. Thus, CYP105D might be a promising candidate for the synthesis of these commercially not available derivatives for several reasons. First, CYP105D is a bacterial enzyme and can be easily produced in heterologous hosts. Second, it has a broad substrate scope and oxidizes nearly all tested bile acids. This makes CYP105D a more promising enzyme candidate than the recently described CYP107D1 from *S. antibioticus*.[158] This P450 hydroxylates some bile acids such as lithocholic acid and deoxycholic acid but did not accept other bile acids including chenodeoxycholic acid.[158] However, in terms of selectivity CYP107D1 was superior above CYP105D because this enzyme hydroxylated bile acids only at the 6β-position.

In order to increase the selectivity of CYP105D towards bile acids, enzyme activity and selectivity can be optimized by site-directed mutagenesis. For this purpose, structural information such as a crystal structure or a reliable homology model has to be provided. Variants with improved selectivity could then be used to produce different hydroxylated bile acids. These bile acid derivatives could be tested for their biological activity and their potential use as novel drug targets.

In conclusion, we were able to identify CYP105D as biocatalyst for bile acid oxidation. The enzyme was characterized regarding substrate binding and *in vitro* oxidation of bile acids. The pharmacologically relevant chenodeoxycholic acid was oxidized at its side chain. This unusual reaction opens an interesting perspective for further engineering of CYP105D in diversified syntheses of oxyfunctionalized bile acids.

Author contributions

Thomas Hilberath: Conceptualization, Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing

Matthias Bureik: Investigation, Funding acquisition, Writing - review & editing

Vlada B. Urlacher: Supervision, Funding acquisition, Writing - review & editing

Acknowledgements

This work was generously supported by the Federal Ministry of Education and Research to Heinrich-Heine University Düsseldorf, grant number 031A223A. We gratefully acknowledge the German Academic Exchange Service (DAAD) for travel grants. We thank Alessandra Raffaele and Ansgar Bokel for proofreading as well as Sebastian Hölzel for technical assistance.

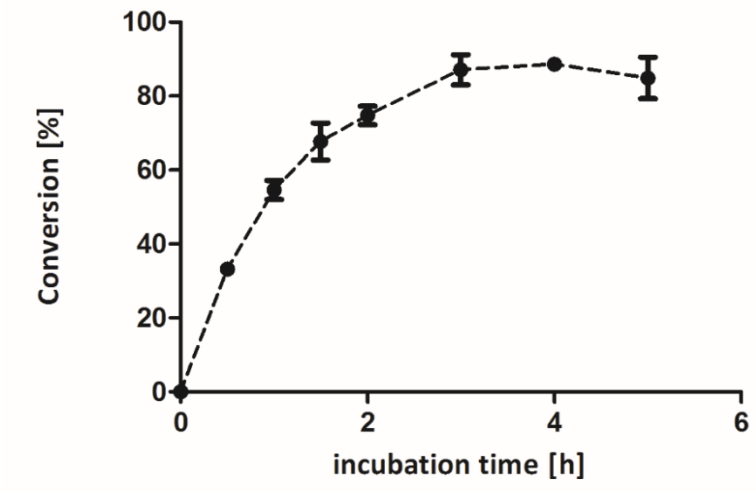
2.4.1 Supporting information

Chromatographic conditions for LC/MS analysis

Supplementary Table 2.4.1: Chromatographic conditions for the LC/MS-analysis of bile acids.

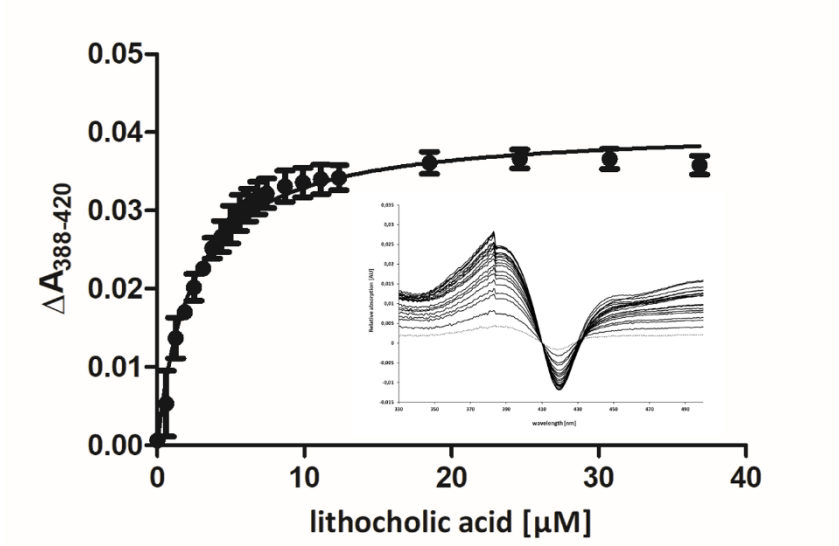
Bile acids		Dehydrocholic acid	
time [min]	methanol concentration [%]	time [min]	methanol concentration [%]
0.01	50	0.01	40
20	100	20	80
25	100	25	100
25.10	50	25.10	40
30	Stop	30	Stop

Time course of oxidation

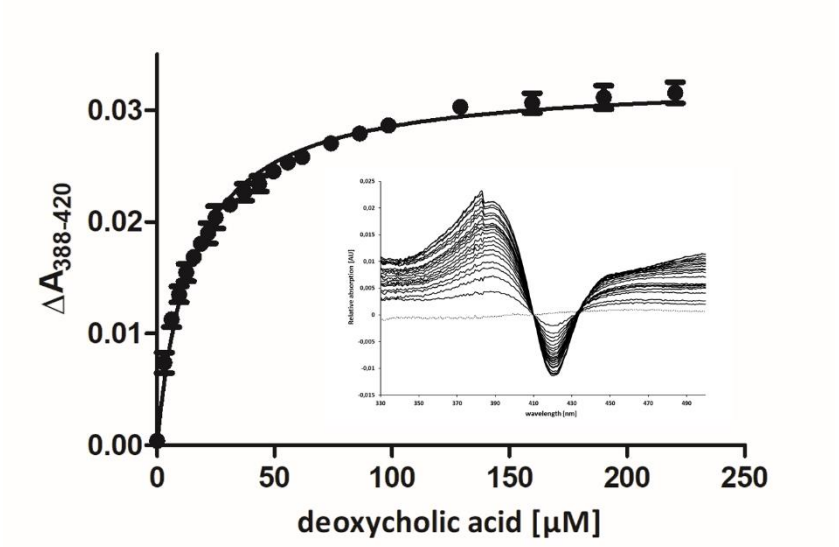


Supplementary Figure 2.4.1: Time course of the oxidation of chenodeoxycholic acid by CYP105D. Reaction conditions: 0.5 mM substrate (in 2 % EtOH final concentration), 2 µM CYP105D,20 µM YkuN, 2 µM Fdr. Reactions were performed in 0.5 mL scale with open lids at 600 min^{-1} shaking frequency in Eppendorf shakers. Conversion was calculated from the ratio of the product peak areas relative to the sum of substrate and product peak area. Measurements were performed in technical duplicates.

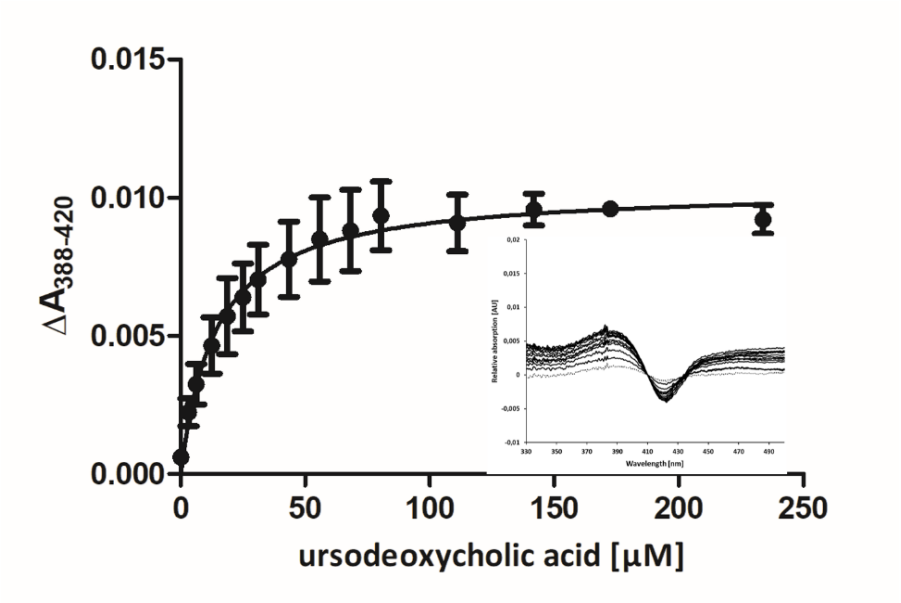
Substrate binding spectra



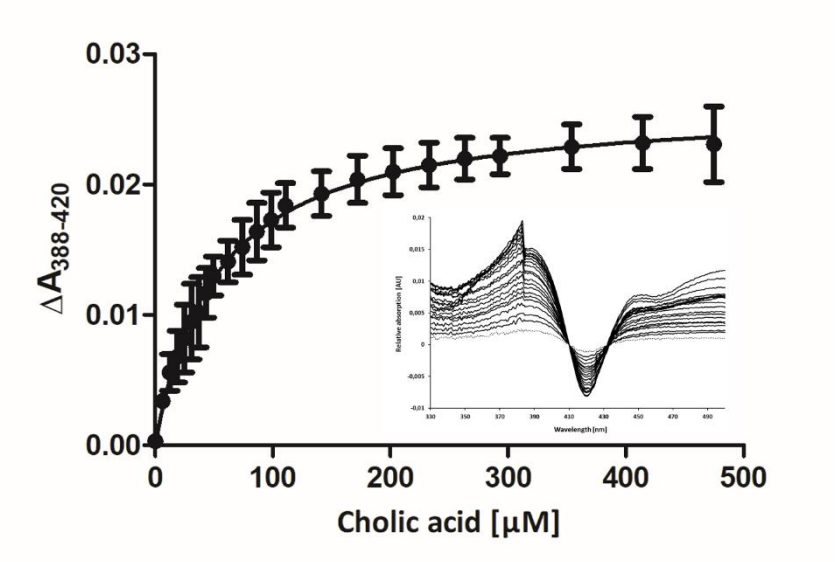
Supplementary Figure 2.4.2: Type I binding of lithocholic acid to CYP105D. Peak and trough were observed at 388 and 420 nm, respectively. Measurements were performed in technical triplicates.



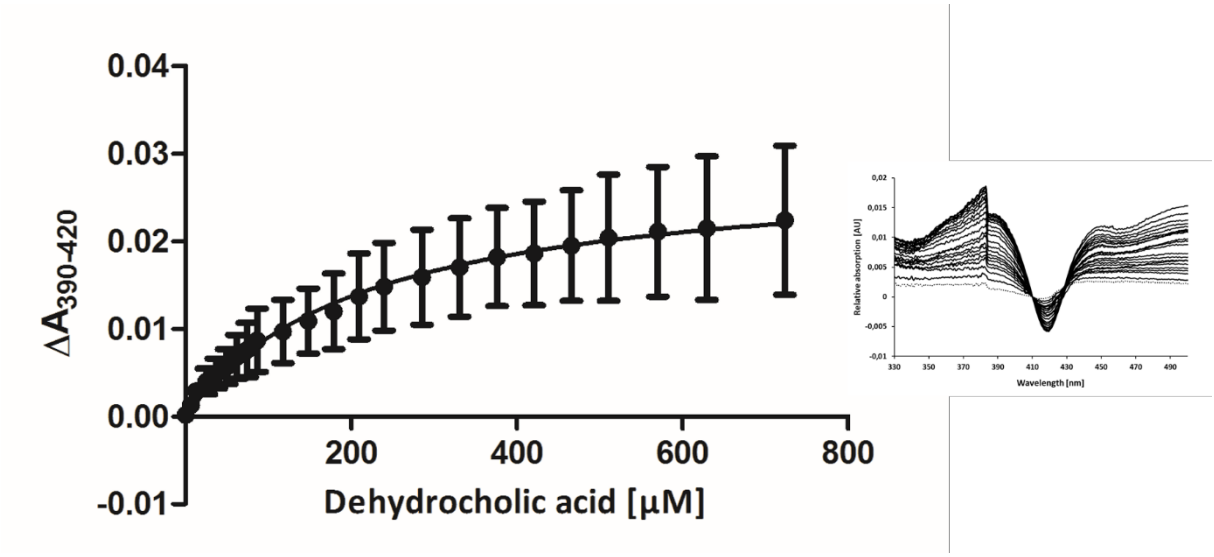
Supplementary Figure 2.4.3: Type I binding of deoxycholic acid to CYP105D. Peak and trough were observed at 388 and 420 nm, respectively. Measurements were performed in technical triplicates.



Supplementary Figure 2.4.4: Type I binding of ursodeoxycholic acid to CYP105D. Peak and trough were observed at 388 and 420 nm, respectively. Measurements were performed in technical triplicates.

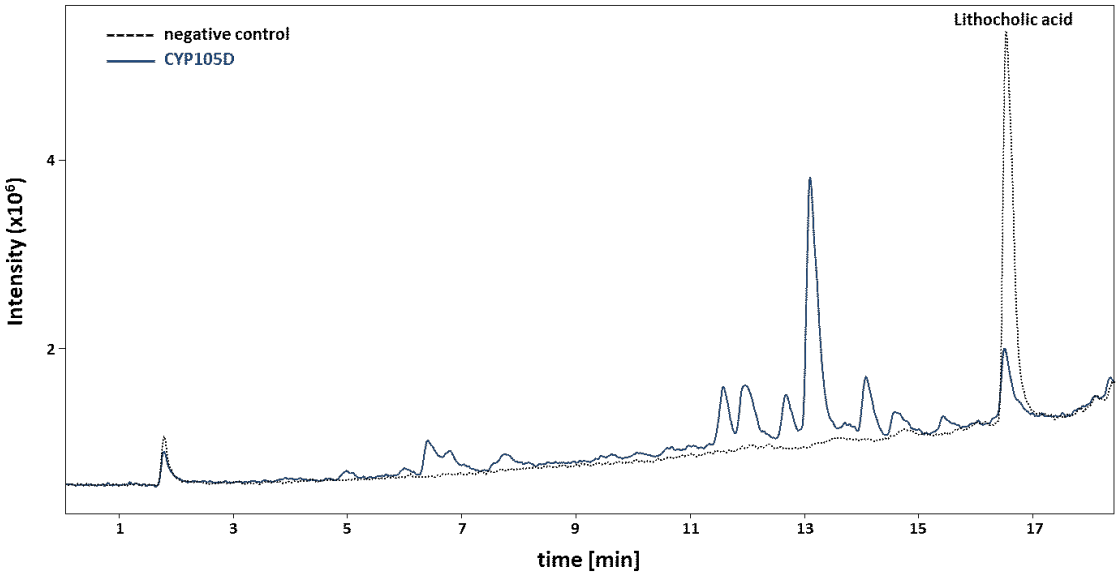


Supplementary Figure 2.4.5: Type I binding of cholic acid to CYP105D. Peak and trough were observed at 388 and 420 nm, respectively. Measurements were performed in technical triplicates.

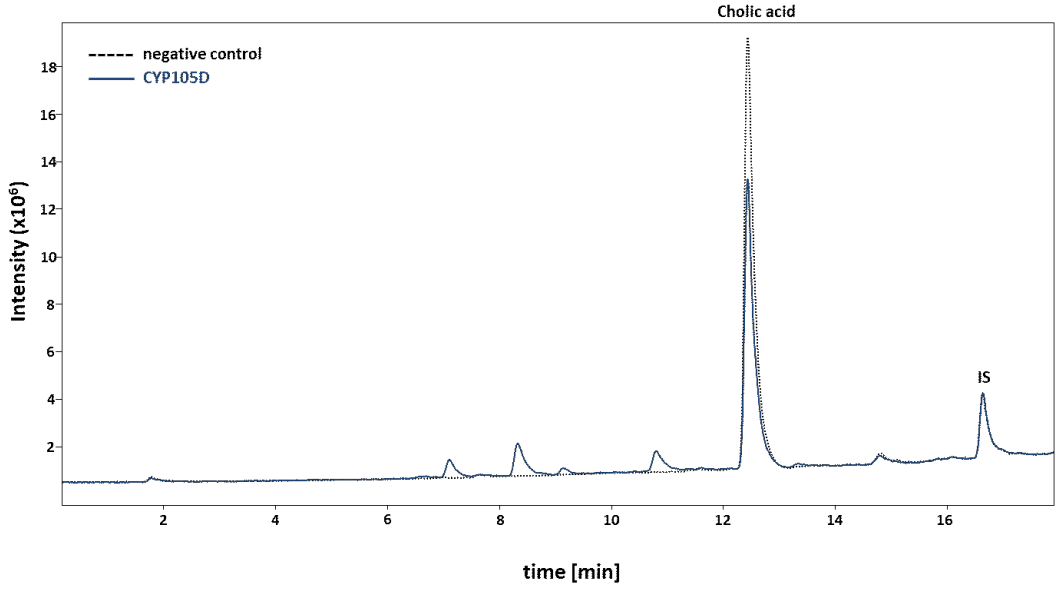


Supplementary Figure 2.4.6: Type I binding of dehydrocholic acid to CYP105D. Peak and trough were observed at 390 and 420 nm, respectively. Measurements were performed in technical triplicates.

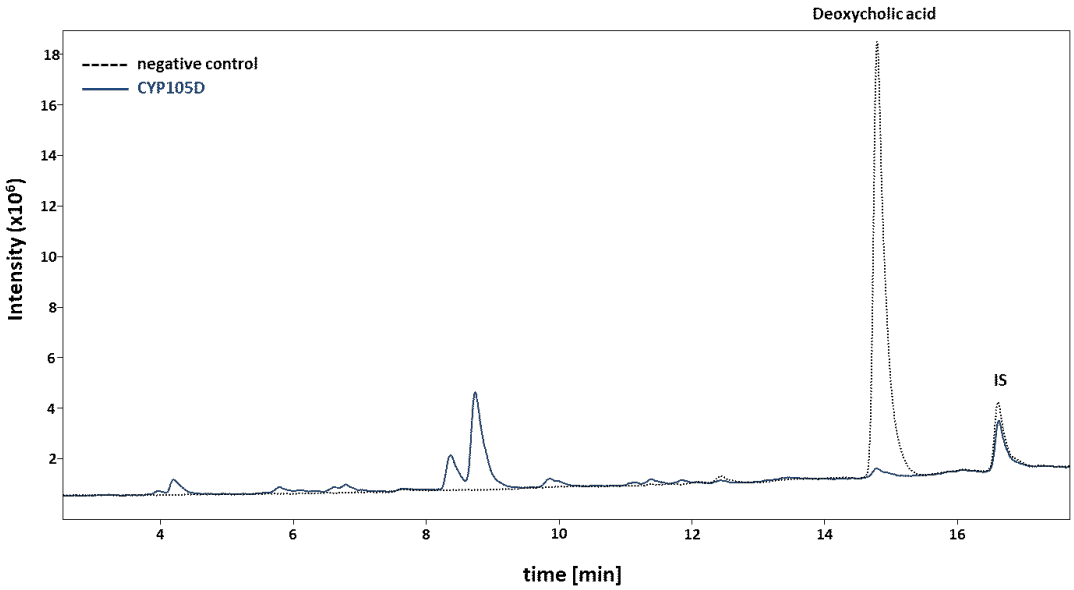
Conversion of bile acids



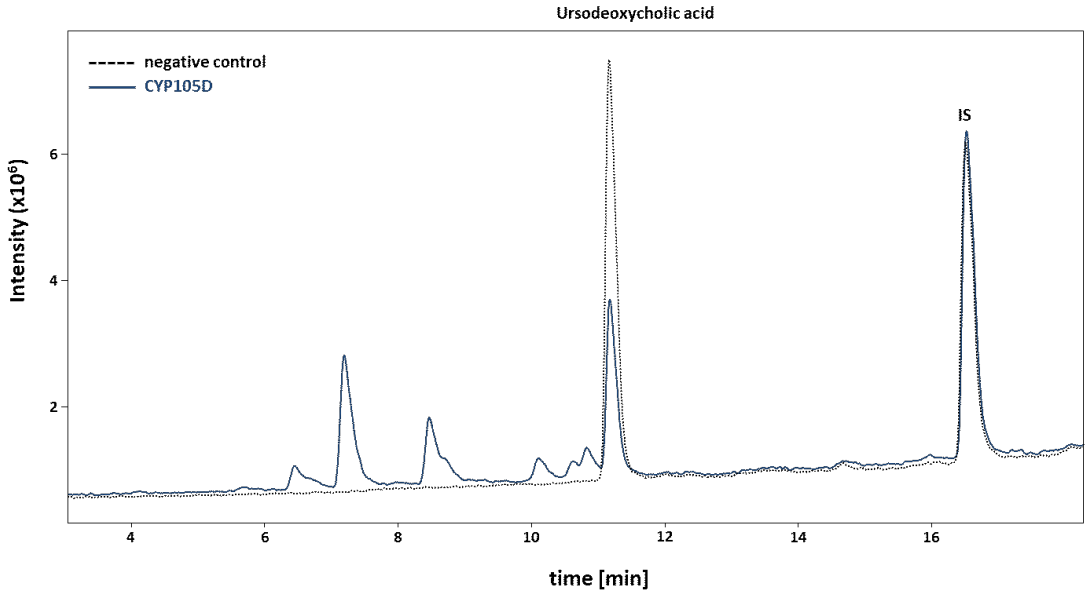
Supplementary Figure 2.4.7: LC/MS-chromatogram for the conversion of lithocholic acid by CYP105D (blue line) compared to the negative control without P450 (black line). The main product (retention time: 13.3 min) has a $[M]^-$ = 391 and an ΔM of = +16 to the substrate lithocholic acid ($[M]^-$ = 375).



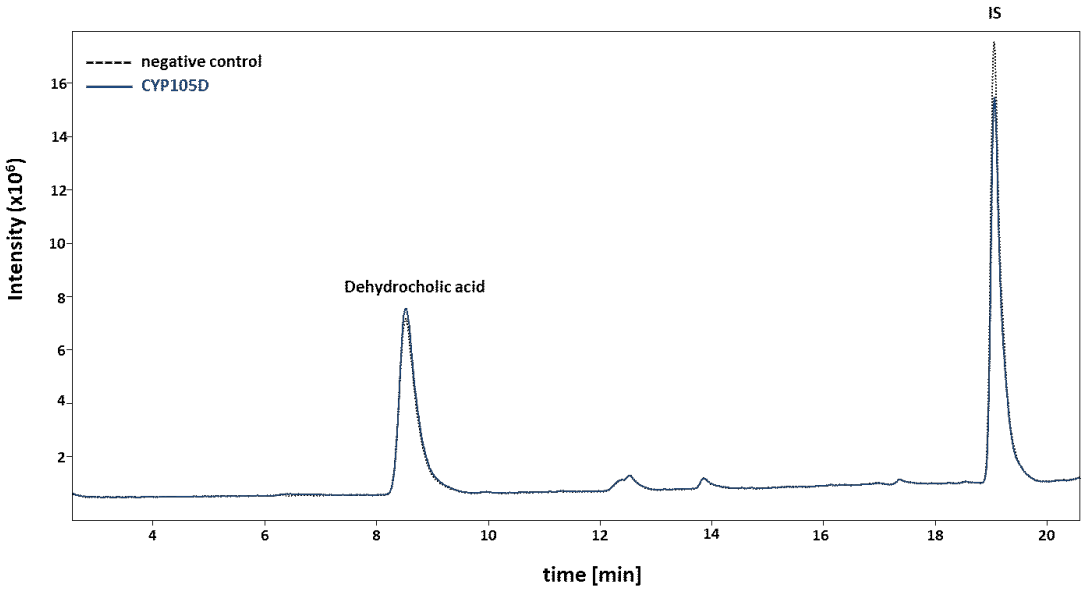
Supplementary Figure 2.4.8: LC/MS-chromatogram for the conversion of cholic acid by CYP105D (blue line) compared to the negative control without P450 (black line). The main product (retention time: 8.3 min) has a $[M]^-$ = 423 and an ΔM of = +16 to the substrate cholic acid ($[M]^-$ = 407). Internal standard (IS): lithocholic acid.



Supplementary Figure 2.4.9: LC/MS-chromatogram for the conversion of deoxycholic acid by CYP105D (blue line) compared to the negative control without P450 (black line). The main product (retention time: 8.8 min) has a $[M]^-$ = 407 and an ΔM of = +16 to the substrate deoxycholic acid ($[M]^-$ =391). Internal standard (IS): lithocholic acid.



Supplementary Figure 2.4.10: LC/MS-chromatogram for the conversion of ursodeoxycholic acid by CYP105D (blue line) compared to the negative control without P450 (black line). The main product (retention time: 7.3 min) has a $[M]^-$ = 407 and an ΔM of = +16 to the substrate ursodeoxycholic acid ($[M]^-$ =391). Internal standard (IS): lithocholic acid.

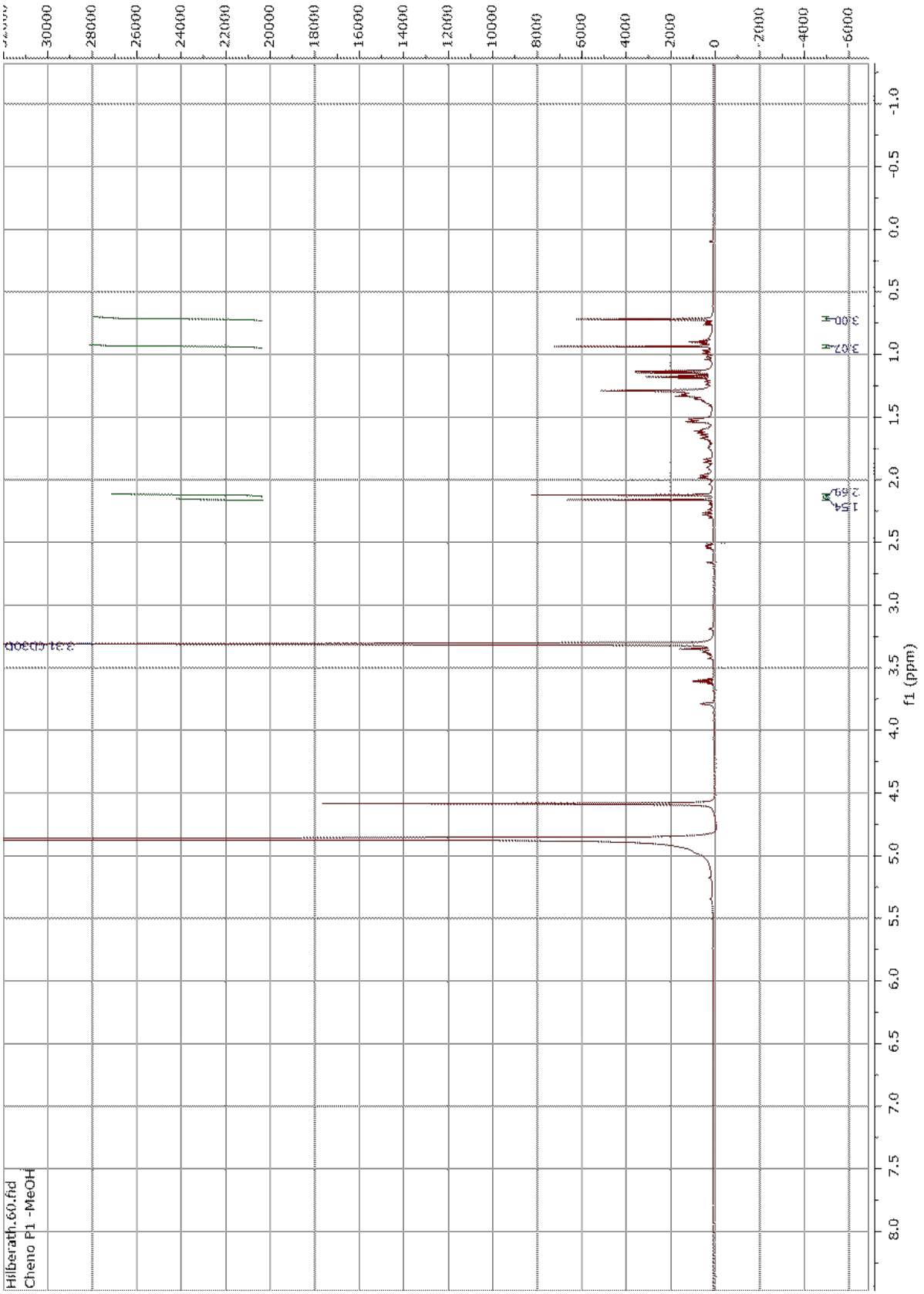


Supplementary Figure 2.4.11: LC/MS-chromatogram for the conversion of dehydrocholic acid by CYP105D (blue line) compared to the negative control without P450 (black line). Internal standard (IS): cholic acid.

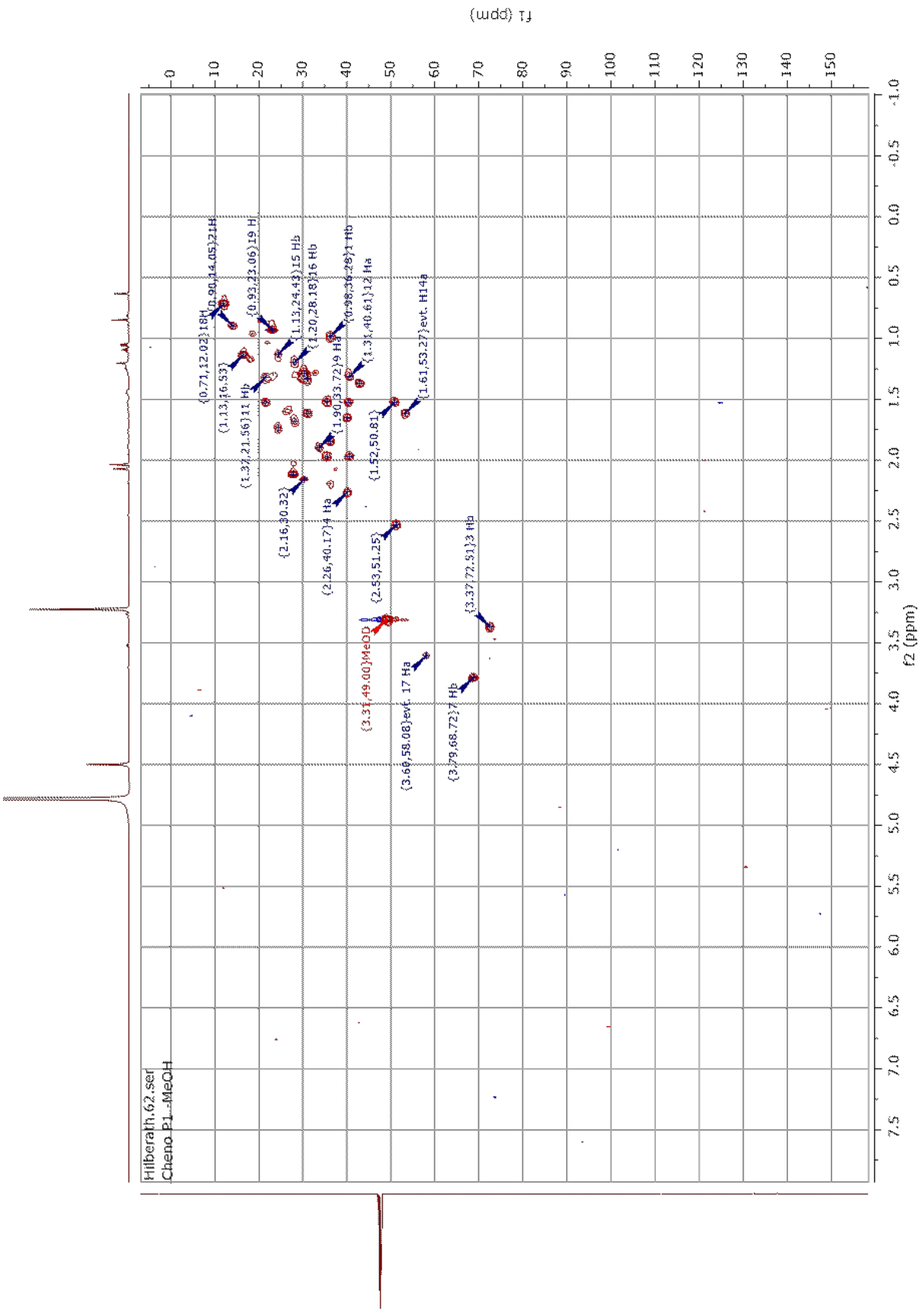
NMR-data

Supplementary Table 2.4.2: ^{13}C and ^{1}H resonance assignments for chenodeoxycholic acid and isolated products. The assignments are based on HSQC-spectra. Signals which are marked with an asterix (*) were assigned by HMBC spectra, whereas signals with a plus (+) were assigned by ^{1}H spectra. Pink colored lines indicate atoms which were not identified in HSQC spectra.

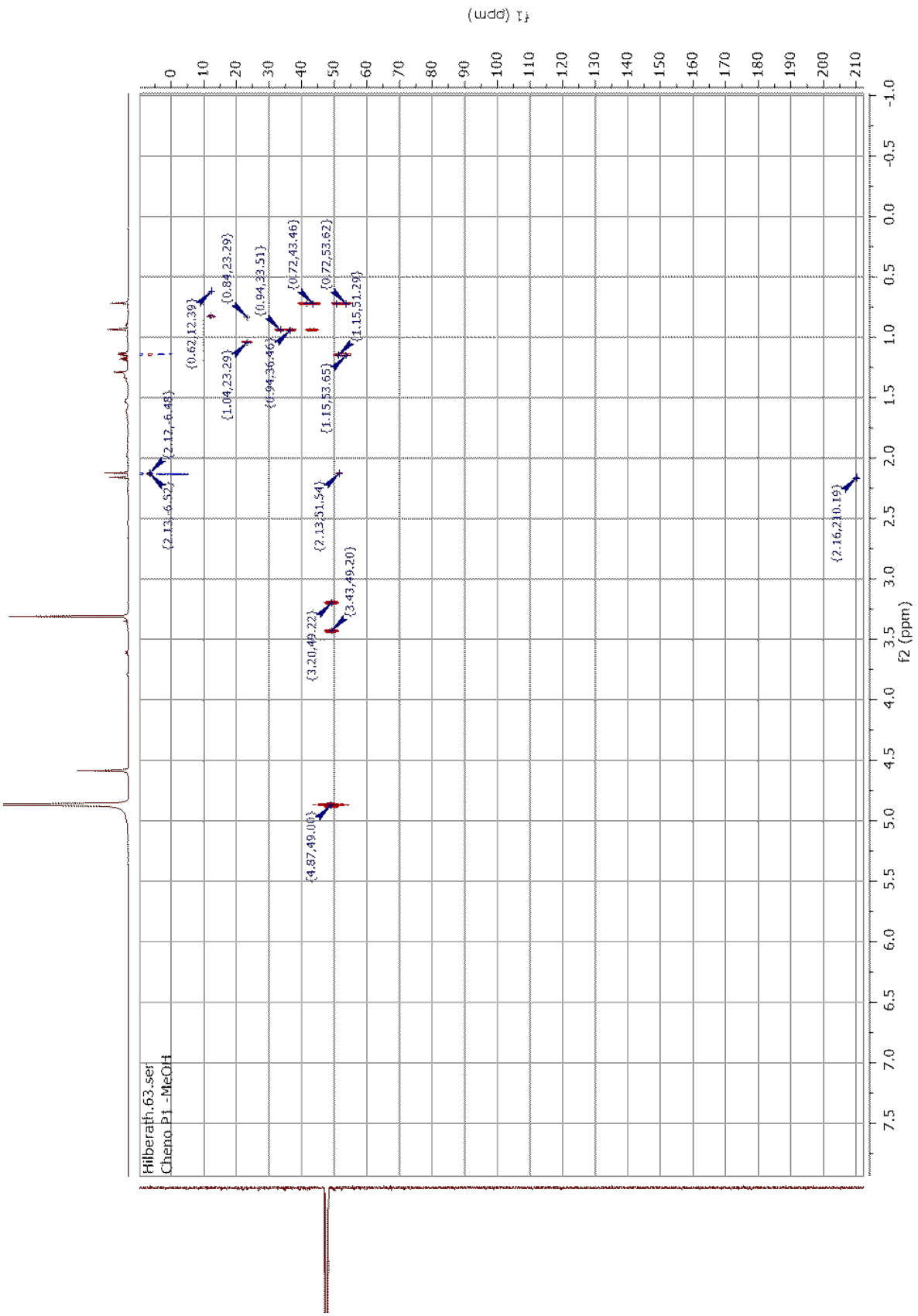
	Chenodesoxycholic acid			P1			P2		
Number	Carbon	Proton		Carbon	Proton		Carbon	Proton	
		α	β		α	β		α	β
1	36.2 ppm	1.86 ppm	0.98 ppm	36.2 ppm	1.85 ppm	0.98 ppm	36.2 ppm	1.86 ppm	0.98 ppm
2	31.1 ppm	1.34 ppm	1.61 ppm	31.1 ppm	1.34 ppm	1.61 ppm	30.9 ppm	1.34 ppm	1.62 ppm
3	72.5 ppm	-	3.37 ppm	72.5 ppm	-	3.37 ppm	72.4 ppm	-	3.37 ppm
4	40.1 ppm	2.26 ppm	1.66 ppm	40.1 ppm	2.27 ppm	1.65 ppm	39.9 ppm	2.28 ppm	1.65 ppm
5	42.9 ppm	-	1.37 ppm	42.9 ppm	-	1.37 ppm	42.9 ppm	-	1.37 ppm
6	35.5 ppm	1.54 ppm	1.98 ppm	35.5 ppm	1.51 ppm	1.98 ppm	35.3 ppm	1.51 ppm	1.99 ppm
7	68.8 ppm	-	3.80 ppm	68.7 ppm	-	3.79 ppm	68.7 ppm	-	3.79 ppm
8	40.4 ppm	-	1.50 ppm	40.4 ppm	-	1.52 ppm	40.3 ppm	-	1.50 ppm
9	33.7 ppm	1.87 ppm	-	33.7 ppm	1.90 ppm	-	33.6 ppm	1.88 ppm	-
10	no Signal HSQC			no Signal HSQC			no Signal HSQC		
11	21.5 ppm	1.49 ppm	1.33 ppm	21.6 ppm	1.52 ppm	1.32 ppm	21.6 ppm	1.51 ppm	
12	40.7 ppm	1.20 ppm	2.01 ppm	40.6 ppm	1.31 ppm	1.97 ppm	40.6 ppm	1.21 ppm	2.01 ppm
13	no Signal HSQC			no Signal HSQC			no Signal HSQC		
14	51.2 ppm	1.50 ppm	-	53.3 ppm	1.61 ppm	-	50.7 ppm	1.49 ppm	-
15	24.2 ppm	1.11 ppm	1.74 ppm	24.4 ppm	1.13 ppm	1.73 ppm	24.0 ppm	1.13 ppm	1.76 ppm
16	28.9 ppm	1.91 ppm	1.33 ppm	28.1 ppm	1.68 ppm	1.19 ppm	28.3 ppm	1.86 ppm	1.38 ppm
17	57.0 ppm	1.18 ppm	-	50.8 ppm*	1.52 ppm*	-	54.2 ppm	1.17 ppm	-
18	11.8 ppm	0.69 ppm		12.0 ppm	0.71 ppm		11.7 ppm	0.72 ppm	
19	23.1 ppm	0.93 ppm		23.1 ppm	0.93 ppm		23.1 ppm	0.93 ppm	
20	36.4 ppm	1.46 ppm							
21	18.5 ppm	0.96 ppm		14.1 ppm+	0.90 ppm+				
22	32.0 ppm	1.81 ppm	1.32 ppm						
23	31.7 ppm	2.33 ppm	2.21 ppm						
24	no Signal HSQC			no Signal HSQC			no Signal HSQC		
unassigned signals [ppm]:				(58.1/3.60), (51.3/2.53), (30.4/1.29), (30.3/2.16), (27.9/2.12), (18.1/1.17), (16.5/1.13)			(71.1/4.05), (41.9, 1.74), (37.3/ 2.22), (37.3/2.15), (30.3/1.29), (23.4/1.91), (12.7, 0.95)		



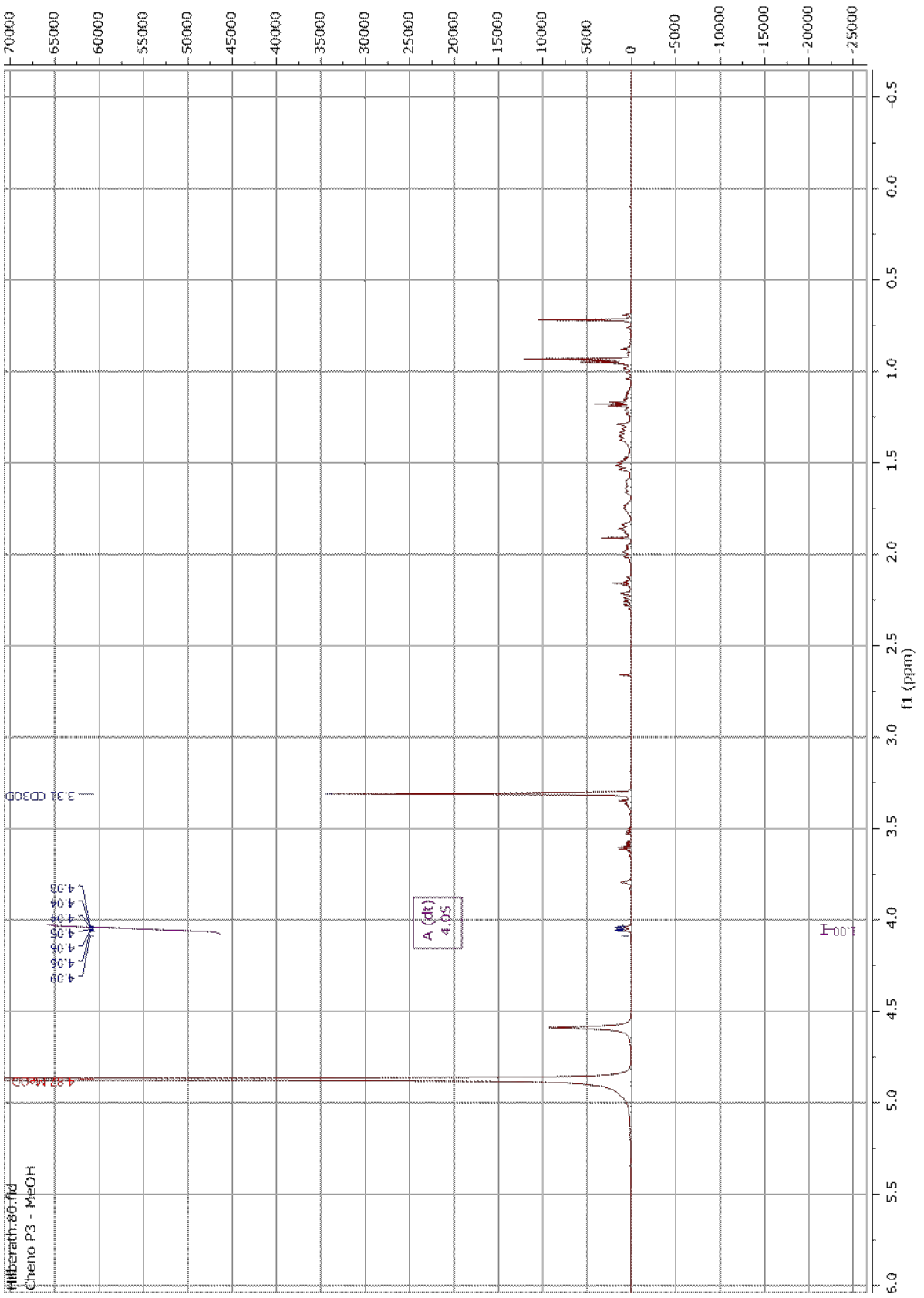
Supplementary Figure 2.4.12: ^{1}H NMR of product P1 in CD_3OD.



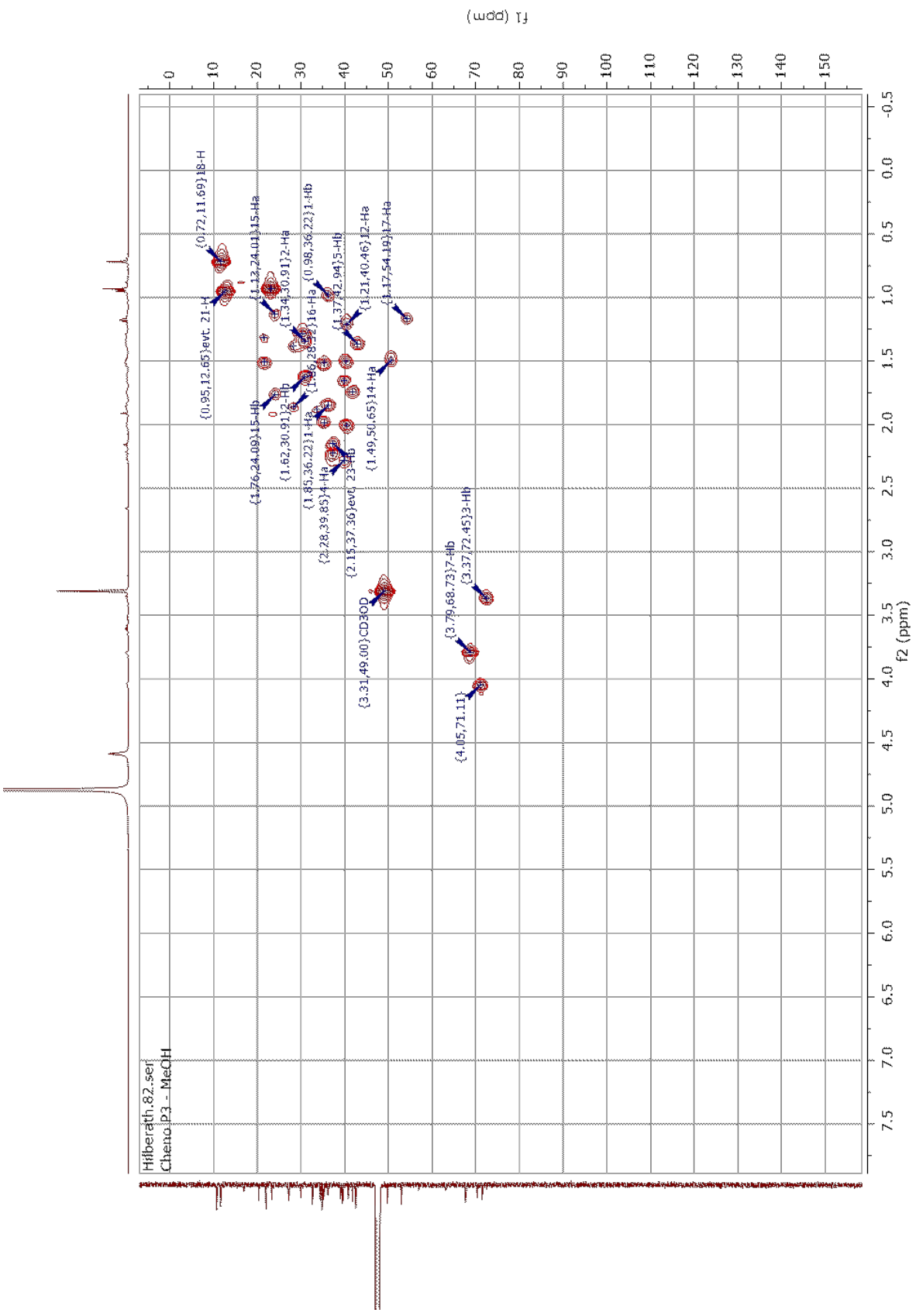
Supplementary Figure 2.4.13: $^{1}H,^{13}C$-HSQC spectrum of product P1 in CD_3OD.



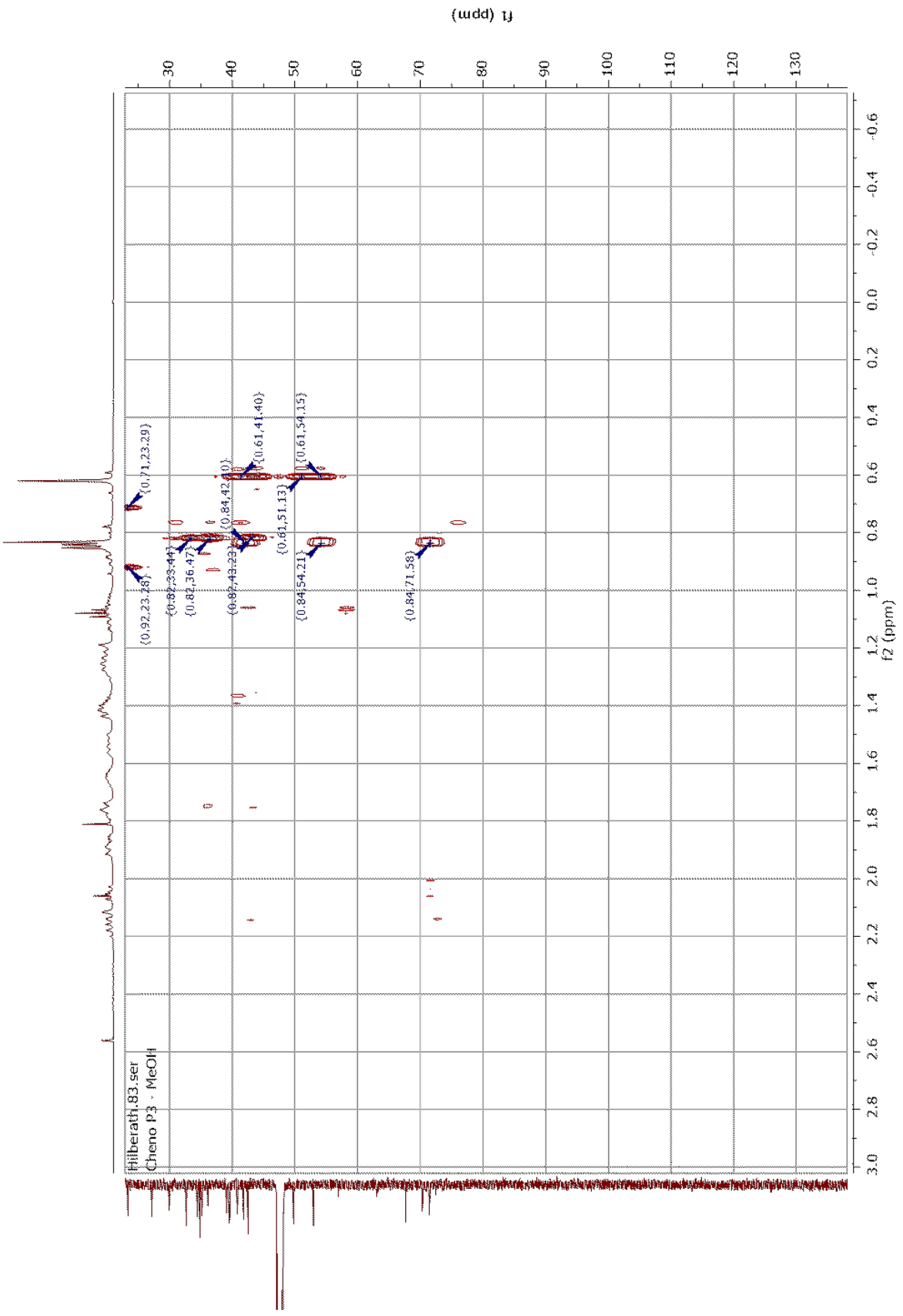
Supplementary Figure 2.4.14: $^{1}H,^{13}C$-HMBC spectrum of product P1 in CD_3OD.



Supplementary Figure 2.4.15: ^{1}**H NMR of product P2 in** CD_3OD**.**



Supplementary Figure 2.4.16: $^{1}H,^{13}C$-HSQC spectrum of product P2 in CD_3OD.



Supplementary Figure 2.4.17: $^{1}H,^{13}C$-HMBC spectrum of product P2 in CD_3OD.

2.5 Lyophilized P450 based whole-cell biocatalysts

Title: Evaluation of P450 activity in lyophilized recombinant *E. coli* cells

Authors: **Thomas Hilberath**, Alessandra Raffaele, Leonie M. Windeln, Vlada B. Urlacher*

* corresponding author

Manuscript in preparation

Own contribution: Conceptualization, design and conduction of most of the experiments, analysis and interpretation of all data, supervision of the master thesis of Leonie Windeln, drafting of the manuscript. Relative contribution: 80%.

Evaluation of P450 activity in lyophilized recombinant *E. coli* cells

Thomas Hilberath, Alessandra Raffaele, Leonie M. Windeln, Vlada B. Urlacher*

Institute of Biochemistry, Heinrich-Heine University Düsseldorf Universitätsstraße 1, D-40225 Düsseldorf, Germany

* Corresponding author: Vlada B. Urlacher. Email: vlada.urlacher@uni-duesseldorf.de

Abstract

Cytochromes P450 catalyze oxidation of chemically diverse compounds and thus offer great potential for biocatalysis. Due to the complexity of these enzymes, their dependency on nicotinamide cofactors and redox partner proteins, recombinant microbial whole cells appear most appropriate for effective P450-mediated reactions. However, some drawbacks exist that require individual solutions. Here, we compare lyophilized recombinant *E. coli* cells with different whole cell preparations and evaluate them for a P450-catalyzed oxidation. *E. coli* harboring CYP105D from *Streptomyces platensis* DSM 40041 was used as model system and testosterone as model substrate. Conversion was first enhanced by optimized handling of resting cells. Co-expression of the alcohol dehydrogenase from *Rhodococcus erythropolis* for cofactor regeneration was beneficial for biocatalysis with wet resting cells and crucial to achieve P450 activity with lyophilized recombinant *E. coli* cells. The use of recombinant lyophilized *E. coli* cells for P450 mediated oxidations is a promising starting point towards a broader application of these enzymes.

Keywords: cytochrome P450; whole-cell biotransformation; lyophilized cells; cofactor regeneration

Abbreviations: CYP: cytochrome P450, Pdr: putidaredoxin reductase from *Pseudomonas* putida, Pdx: putidaredoxin from *Pseudomonas putida,* Re-ADH: alcohol dehydrogenase from *Rhodococcus erythropolis*, NADH: Nicotinamide adenine dinucleotide

Cytochromes P450 (CYP or P450) are versatile heme-containing enzymes that catalyze oxidation reactions in the presence of molecular oxygen and NAD(P)H. Due to their ability to introduce one atom of molecular oxygen into a vast variety of organic molecules under mild reaction conditions with often high chemo- and regioselectivity, these enzymes have been recognized as attractive targets with high potential for biotechnological applications.[16a, 64, 159] Generally, whole-cell biocatalysis seems appealing because it allows to avoid cell lysis and enzyme isolation.[63] Enzymes are protected by the cell environment from the harmful influence of reaction components.[160] In case of NADH and NADPH dependent enzymes like P450s, these cofactors can be continuously regenerated via metabolism of the host cell, or optionally by the use of different heterologous cofactor-regenerating enzymes and co-substrates.[161] With regard to P450 enzymes, whole cell biocatalysis might be particularly interesting as electrons from NAD(P)H are transferred via one or two redox partner proteins to the catalytically active heme. Co-expression of the enzymes belonging to a P450 redox chain in one microbial cell appears more attractive than their separate expression and isolation.

Despite the apparent advantages of whole-cell systems for P450-catalyzed reactions, their application is often associated with challenges like substrate/product toxicity for the cell and limited substrate and product transfer across the cell membrane.[15, 64] Whereas substrate toxicity can be overcome by using more stable hosts[69a, 69b, 162], improved substrate uptake can be achieved by co-expression of transporters[70a], cell permeabilization[71] or other commonly used procedures like freezing and thawing[102, 163] In case of hydrophobic substrates of P450 enzymes, their low solubility in aqueous solution represents an additional drawback for biocatalysis. To increase substrate solubility organic solvents are added, which might negatively affect the biocatalysts. To this end, usage of lyophilized recombinant microbial cells carrying the target enzymes has been reported as an attractive alternative to both microbial whole cells and isolated enzyme because they allow working at high organic solvent concentrations and do not face the problem of substrate transport through the membrane.[164]

In this respect, it is important to explore the use of lyophilized recombinant *E. coli* cells for the P450-mediated biocatalysis and compare them with the better investigated whole-cell preparations. In this work we used as model system the recently characterized CYP105D

from *Streptomyces platensis* DSM 40041 that accepts a broad range of substrates including testosterone.[11] Oxyfunctionalized steroids like 2β-hydroxytestosterone are of high pharmaceutical interest as drug precursors and human drug metabolites.[165] Testosterone is a common steroid substrate often applied to evaluate the activity of P450s of prokaryotic and eukaryotic origin.[56h, 65a, 100, 166] As this compound has low solubility in water and a relatively large size which impair substrate uptake, different whole-cell preparations were compared to address the substrate accessibility issue. An *E. coli* C43 (DE3) whole-cell biocatalyst co-expressing CYP105D (GenBank accession no. **OSY47991**) with the NADH-dependent putidaredoxin reductase (Pdr, GenBank accession no. **BAA00413**) and putidaredoxin (Pdx, GenBank accession no. **BAA00414**) on two plasmids was constructed (details in the Supplementary material) and used for oxidation of 1 mM testosterone to 2β-hydroxytestosterone (Figure 2.5.1). Details on conditions for whole-cell biocatalysis and product quantification via LC/MS are provided in the Supplementary information.

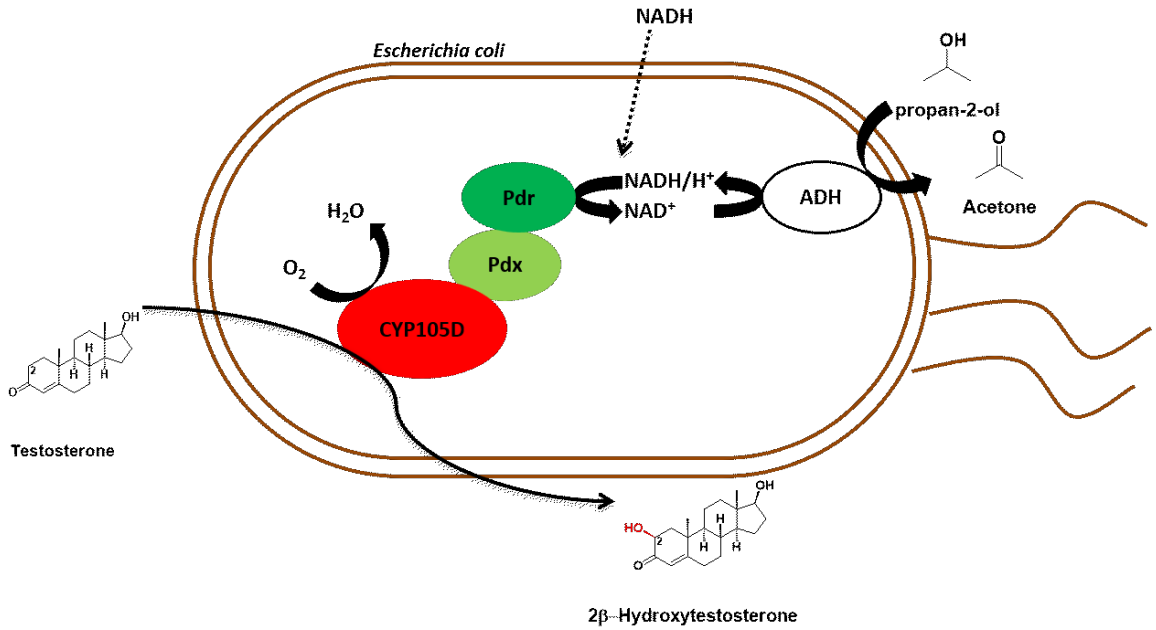


Figure 2.5.1: Schematic overview of the whole-biocatalyst for oxidation of testosterone based on CYP105D from *S. platensis*. Putidaredoxin reductase (Pdr) and putidaredoxin (Pdx) from *P. putida* are used as redox partners for CYP105D. ADH from *R. erythropolis* was implemented for cofactor regeneration using propan-2-ol as sacrificial substrate and solvent for testosterone.

E. coli cells were treated differently to identify the optimal cell preparation. Within this study the following cell preparations were investigated:

1. Resting cells obtained directly after cultivation and centrifugation without further treatment (‘non frozen’)

2. Resting cells obtained directly after cultivation and centrifugation and one additional sonication cycle after resuspension in PSE-buffer to a cell concentration (cww) of 100 mg/mL (‘sonified’)

3. Resting cells which were frozen at - 20°C as cell pellet (‘frozen cell pellet’)

4. Resting cells which were frozen at - 20°C as cell suspension in PSE buffer (‘frozen cell suspension’)

5. Lyophilized cells obtained using a Christ alpha 2-4 LSCplus (Martin Christ Gefriertrocknungsanlagen GmbH, Germany) (‘lyophilized cells’). For that purpose; cell pellets were thawed at room temperature, spread in a crystallization bowl and frozen at -20 °C without any additives. Lyophilization was conducted for at least one day at -80 °C under vacuum. Lyophilized cells were then transferred to a 50 mL reaction tube and stored at -20 °C.

Resting cells used immediately after cultivation and centrifugation (‘non frozen’) were least active as the conversion was not higher than 3% (Figure 2.5.2A). Freeze-thawing of *E. coli* cells has been reported to destabilize cell membrane by releasing some cell components, which make it more permeable.[167] Indeed, freezing the cells at -20 °C with subsequent thawing had a beneficial effect on activity of the whole-cell biocatalyst (Fig. 2.5.2A). However, it did matter in which manner the cells were frozen. If cells were first resuspended in buffer and then frozen at -20°C (‘frozen as cell suspension’), the activity was lower as compared to the procedure when the cell paste after centrifugation was frozen, thawed and resuspended in buffer just before the biotransformation (‘frozen as pellet’) (6% vs 31%). As previously reported, slow freeze-thawing mainly released components of the outer

membrane, whereas fast freeze-thawing caused a more drastic decay, also releasing cytoplasmic components.[167] In our experiments, individual cells resuspended in buffer can be frozen and thawed faster than cell paste. Additionally, ice crystals might also have an impact on the release of cell components. On this basis we hypothesize that cells resuspended in buffer lose cytoplasmic components after freeze-thawing and thus are less stable and active. The activity of the best performing resting cells frozen at -20°C ('frozen as pellet') did not show notable differences to cells which were additionally treated by sonication ('sonified') with which the conversion was about 30%. This suggests that both procedures similarly affect mass transfer over the membrane.

After identification of the most suitable whole-cell preparation ('frozen as pellet'), we aimed to increase conversion further by addition of cyclodextrins (Fig. 2.5.2B). Cyclodextrins are solubilizing agents that possess a hydrophilic outer surface and a hydrophobic cavity in which they can accommodate hydrophobic molecules in aqueous solution.[168] For whole-cell conversions of steroids in particular (2-hydroxypropyl)-β-cyclodextrin has been frequently used.[102, 169] In the present case, the addition of (2-hydroxypropyl)-β-cyclodextrin had a negative effect on conversion. In comparison to the whole-cell conversion without cyclodextrines, the equimolar addition of 1 mM (2-hydroxypropyl)-β-cyclodextrin already led to a 2-fold decrease of substrate conversion (17%). Increasing cyclodextrin concentrations caused a further decrease of conversion. It is assumed that the substrate probably got trapped by the cyclodextrin and thus is not accessible for the whole-cell biocatalyst any longer. Since no positive effect on substrate conversion was seen, (2-hydroxypropyl)-β-cyclodextrin was not added in the next experiments.

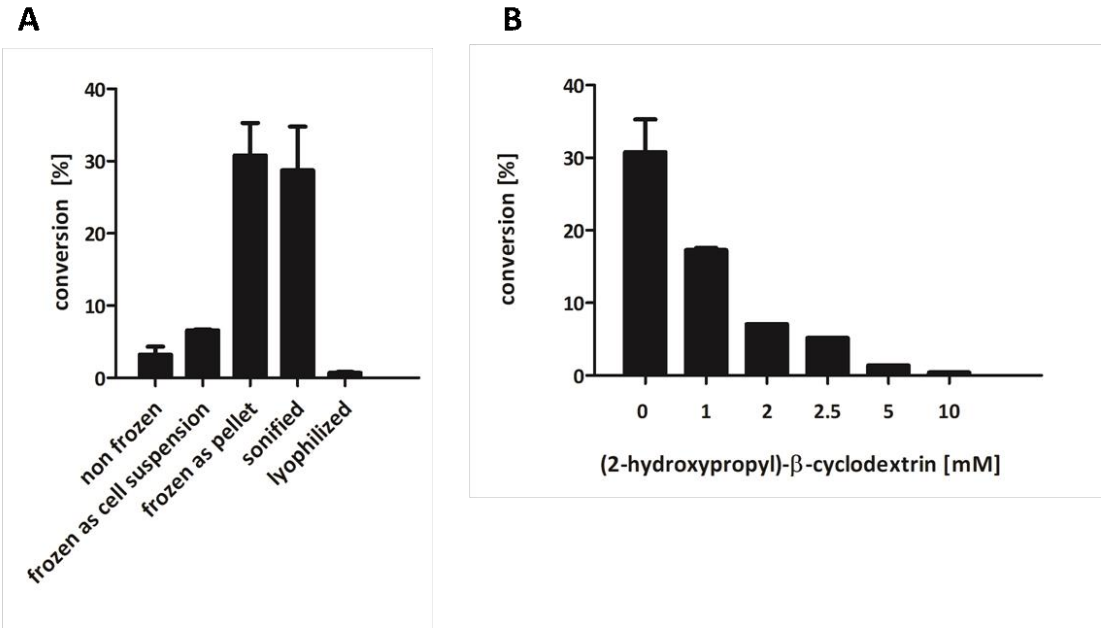


Figure 2.5.2: (A) Effect of different handling of resting cells of *E. coli* C43 (DE3) pET22b-*cyp105D* + pCOLADuet-*pdx-pdr* on testosterone conversion. (B) Conversion of testosterone by *E. coli* C43 (DE3) pET22b-*cyp105D* + pCOLADuet- *pdx-pdr* in presence of (2-hydroxypropyl)-β-cyclodextrin. Reaction conditions: 50 mg/mL wet cells in 0.5 mL Phosphate Sucrose EDTA (PSE)--buffer, pH 7.5 in 2 mL reaction tubes, 1 mM testosterone dissolved in 5% (v/v) propan-2-ol final concentration, 25 °C, 1100 min^{-1} shaking frequency. Cells were frozen at -20 °C for preparation of 'frozen cells.' (2-hydroxypropyl)-β-cyclodextrin was additionally supplemented to the best performing wet cell biocatalyst (`frozen as cell pellet') in concentrations of 1 – 10 mM. Experiments were performed in technical duplicates.

Activity of the lyophilized P450 whole-cell catalysts was less than 1% (Fig. 2.5.2A). We assumed that loss of activity in lyophilized cells was attributed to insufficient cofactor supplementation. As it would be advantageous to use lyophilized cells due to their easy handling, we further investigated if cofactor supply affected their catalytic performance in this case. To ensure cofactor regeneration in lyophilized cells, we additionally cloned the gene encoding for the alcohol dehydrogenase from *Rhodococcus erythropolis* DSM 43297 in the plasmid downstream of *pdx* and *pdr* (Supplementary Figure 2.5.1). The NAD^+-dependent alcohol dehydrogenase from *Rhodococcus erythropolis* DSM 43297 (Re-ADH, GenBank accession no. **CAF04319**)[170] catalyzes oxidation of the cheap sacrificial substrate propan-2-ol to acetone thereby reducing NAD^+ to NADH.[171] Hence, we used propan-2-ol as substrate of Re-ADH and simultaneously as co-solvent to dissolve testosterone. The P450 concentration in the cell was marginally affected by co-expression of an additional enzyme (278 ±11 nmol/g_{CDW} vs 268 ±2 nmol/g_{CDW}) as determined from CO-difference spectra.[14b] NADH production during propan-2-ol oxidation was evaluated by a photometric assay and was only detected with *E. coli* cells expressing Re-ADH (52 ± 0 U/g_{CDW}) and not with another

strain, which indicated that this ADH was successfully expressed (details in the Supplementary material). While the co-expression of Re-ADH had a beneficial effect on the activity of both, the best-performing resting wet cells (‘frozen as cell pellet’) and the lyophilized whole-cell biocatalyst, it was particularly advantageous for the latter one (Figure 2.5.3A). This effect indicates that targeted cofactor regeneration is crucial to support P450 activity in lyophilized cells.

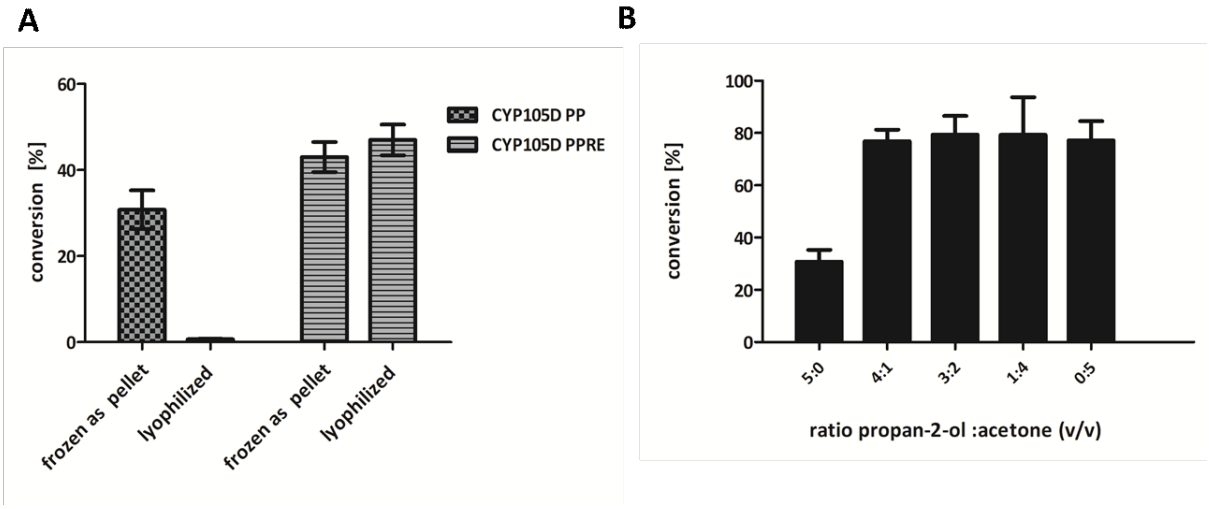


Figure 2.5.3: (A) Influence of cofactor regeneration by Re-ADH in *E. coli* C43 (DE3) pET22b-*cyp105D* + pCOLADuet- *pdx-pdr-adh* on testosterone conversion. (B) Effect of different ratios of propan-2-ol and acetone on testosterone conversion mediated by the wet whole-cell catalyst without ADH (CYP105D PP). The best performing wet cell biocatalyst (`frozen as cell pellet’) was investigated (see Figure 2.5.2A). Reaction conditions: 10 mg/mL lyophilized cells or 50 mg/mL wet cells in 0.5 mL Phosphate Sucrose EDTA (PSE)- buffer, pH 7.5 in 2 mL reaction tubes, 1 mM testosterone dissolved in 5% co-solvent (v/v) final concentration, 25 °C, 1100 min^{-1} shaking frequency. Experiments were performed in duplicates.

In order to further validate this hypothesis, we investigated the influence of external NADH on the activity of lyophilized cells since NADH might get lost or degraded during lyophilization.[65a] NADH was added to lyophilized cells in different concentrations and at different time points up to four times during biotransformation (Supplementary Figure 2.5.3). The activity of the lyophilized cells without Re-ADH could be increased only very slightly by NADH-addition (0.25 - 1 mM final concentration) independent of the time point and amount of added NADH (max conversion 3%). The combination of P450s with heterologous redox partners for non-physiological substrates often results in high uncoupling which leads to unproductive NADH consumption.[15] In the present case, the low conversion might reflect the uncoupling of the tested P450 system assuming that NADH cannot be regenerated by the metabolism in lyophilized *E. coli* cells. However, the

supplementation of 0.5 mM NADH after 4 h to the lyophilized cells where Re-ADH was present resulted in a 1.4-fold increase in activity towards testosterone. Testosterone conversion of 72% with lyophilized cells was similar or even slightly higher than that observed with resting cells (Table 2.5.1).

Additionally, the effect of the ADH by-product acetone was analyzed as acetone is a common organic solvent for chemical permeabilization of the cell membrane.[165, 172] Acetone is formed during NADH formation by Re-ADH and thus may contribute to change in solubility and cellular uptake of the substrate testosterone.[171] The oxidation of propan-2-ol to acetone is a reversible reaction, which leads to a thermodynamic equilibrium and consequently to different ratios of the two co-solvents over time.[173] We analyzed the effect of acetone on substrate conversion with wet cells and lyophilized cells, both containing the P450 system but no Re-ADH, by testing different ratios of the co-solvents propan-2-ol and acetone (Figure 2.5.3B). Increasing acetone concentrations had a positive effect on activity of the cells without Re-ADH and resulted in a 2.5-fold increase of conversion of maximum 79%. The increase in conversion catalyzed by the whole-cell catalyst with Re-ADH compared to the system without Re-ADH could be explained not only by the additional cofactor regeneration of ADH but also by the formation of acetone, which might have a positive effect on cell permeability. However, this effect was only observed for wet cells and not for lyophilized cells as conversion with increasing acetone concentrations was still less than 1% (data not shown). This supports the idea that targeted cofactor regeneration rather than improved substrate uptake is crucial to achieve P450 activity in lyophilized cells.

Table 2.5.1: Effect of external NADH addition on the activity of lyophilized P450 whole-cell catalysts. Reaction conditions: 10 mg/mL lyophilized cells in 0.5 mL Phosphate Sucrose EDTA (PSE)-buffer, pH 7.5, in 2 mL reaction tubes, 1 mM testosterone dissolved in 5% (v/v) propan-2-ol final concentration, 25 °C, 1100 rpm. 0.25 mM NADH was added up to four times at 0 h, 2 h, 4 h and 6 h incubation. For the cells co-expressing the *adh*, 0.5 mM NADH were added after 4 h. Experiments were performed in technical duplicates**.**

lyophilized *E.coli* C43 (DE3) harboring	testosterone conversion [%]	
	-NADH	+NADH
pET22b-*cyp105D* + pCOLADuet-*pdx-pdr*	≤ 1	3
pET22b-*cyp105D* + pCOLADuet-*pdx-pdr-adh*	51 ± 9	72 ± 5

Our results demonstrate that (i) handling procedure has a strong effect on the catalytic performance of recombinant P450-containing resting cells, and (ii) metabolism-independent regeneration of NADH is necessary to allow P450-catalysis with lyophilized cells. The use of these procedures illustrates interesting perspectives for convenient applications of cytochrome P450s for single- or multi-step reactions.

Acknowledgements

Financial support was kindly provided by the Federal Ministry of Education and Research [grant number 031A223A] under the umbrella of the ERA-IB2 3rd call project ‘HyPerIn’ [project number EIB.12.026].

2.5.1 Supporting information

Supplementary Material and methods

Supplementary Table 2.5.1: Synthetic oligonucleotides for cloning. Restriction sites are underlined.

Primer name	DNA-sequence (5'-3')	Restriction enzymes	usage
F-READH-NdeI	CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATA TACATATGAAGGCAATCCAGTACACGAGAATC	NdeI/XhoI	Amplification of *re-adh*
RC-READH-XhoI	CTGGCGTTCAAATTTCGCAGCAGCGGTTTCTTTACCA GACTCGAGTTACAGACCAGGGACCACAACCG		
fw-PP-pCOLADuet	CAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAAT AAGGAGATATACCATGGCTTCTAAAGTAGTGTATGT GTCACATG	NcoI/BamHI	Amplification of *camA* and *camB*
Rv-PP-pCOLADuet	CAAGCTTGTCGACCTGCAGGCGCGCCGAGCTCGAAT TCGGATCCTCAGGCACTACTCAGTTCAGCTTTG		

Supplementary Table 2.5.2: Plasmids used in this study.

vector (internal number)	genes with EMBL bank number	vector properties	reference/source
pET22b- *cyp105D* (THI 78)	*cyp105D* (EMBL-Bank: OSY47991)	SPL_00625 (NdeI,XhoI) cloned in pET22b	Hilberath et al.[11]
pET-28a(+)-RE-ADH (FTI 94)	*re-adh* (EMBL-Bank: CAF04319)	*re-adh* Y174F cloned in pET28a	Abokitse and Hummel[170]
pCOLADuet-PP (THI 87)	*camA* (EMBL-Bank: BAA00413) and *camB* (EMBL-Bank: BAA00414)	*camA* and *camB* (MCSI: NcoI, BamHI) cloned in pCOLADuet1	Hilberath et al.[11]
pCOLADuet-PP-RE (THI 103)	*camA* and *camB/re-adh*	*camA* and *camB* (MCSI: NcoI, BamHI); *re-adh* (MCSII: NdeI, XhoI) cloned in pCOLADuet1	this work

Construction of expression vectors and gene expression

The gene *cyp105D* from *S. platensis* was cloned with conventional methods in the expression vector pET22b between the recognition sites for the endonucleases NdeI and XhoI. Gibson assembly was used to clone the genes coding for alcohol dehydrogenase (*re-adh*), putidaredoxin reductase (*camA*) and putidaredoxin (*camB*) in the pCOLA-Duet vectors.[174] The genes encoding for CYP105D and redox partners were expressed from a two-plasmid system in *E. coli* C43 (DE3) similar as described previously.[108a] For gene expression, 100 mL TB-medium was inoculated with an overnight culture of the respective recombinant *E. coli*

strain to an OD_{600} of 0.05. The cultures were grown in 1 L flasks at 37 °C and 180 rpm for 2.5-3 h. At an OD_{600} of ≈1.0, 500 µM 5-aminolevulinic acid was added and expression of target genes was induced with 500 µM isopropyl β-d-1-thiogalactopyranoside (IPTG). All cultures were incubated at 20 °C and 140 rpm for 20 h after induction.

Preparation of recombinant *E. coli* cells

Different preparations of resting cells of *E. coli* C43 (DE3), carrying pET22b-*cyp105D* and pCOLADuet-PP, were investigated. After cultivation, the culture broth was split to several 50 mL falcon tubes and cells were harvested by centrifugation for at least 20 min at 5250 g and 4 °C. Cell pellets were then washed with 25 mL Phosphate Sucrose EDTA (PSE)-buffer (6.75 g/L KH_2PO_4, 85.5 g/L sucrose, 0.93 g/L EDTA-Na_2*2 H_2O, pH 7.5). The cell preperations were prepared as described in the main manuscript. Prior to the whole-cell biotransformation, cell preparations except for lyophilized cells were adjusted to a cell wet weight (cww) of 100 mg/L.

Preparation of crude cell extracts

Before cell disruption, cells were resuspended in 5 mL cold PSE-buffer supplemented with 0.1 mM phenylmethylsulfonyl fluoride (PMSF). The cell suspension was disrupted by sonication on ice (Branson Sonfier 250; 3 x 1.5 min, 40 % amplitude, duty cycle 4). Between the cycles the cell suspension was incubated for 2 min on ice. Cell debris was removed by centrifugation (40.000 *g*, 25 min and 4 °C). The soluble fraction (crude cell extract) was collected and directly used for determination of the P450-concentration and the ADH-activity. For cell dry weight (cdw) determination, 200 µL of the cell suspension were transferred to a dry 1.5 mL reaction tube. After centrifugation for 2 min at 13.500 *g* at room temperature, the supernatant was discarded and the cell pellets dried for 48 h at 60 °C before weighing. All measurements were performed in triplicate.

Whole-cell biocatalysis

Biotransformations were performed in 500 µL PSE-buffer containing cells in a final cell concentration of 50 mg/mL (cww) or 10 mg/mL (cdw), 1 x nutrient solution (6 mM glucose, 6 mM lactose and 12 mM citrate in PSE-buffer) and 1 mM testosterone (in 5 % (v/v) co-solvent final concentration). The tested co-solvents were propan-2-ol and acetone. The reaction was started by adding the substrate and transferring the reaction mixture to a 2 mL reaction tube. 2 mL reaction tubes with open lids were incubated at 25 °C up to 20 h at 1100 rpm in an Eppendorf shaker. At different time points 50-200 µL aliquots were taken for extraction with 1 mL ethyl acetate. 200 µM progesterone was added as internal standard for control of extraction. After phase separation the organic phase was transferred to a new reaction tube and concentrated under reduced pressure. The analytes were resolved in methanol for LC/MS analysis. Conversions were calculated from the sum of detected product peak areas relative to the substrate peak area either via PDA- or MS-analysis.

Determination of P450 concentration and ADH activity

Concentrations of P450 in crude cell extracts were calculated based on CO difference spectra using the extinction coefficient ε_{450} = 91 mM^{-1} cm^{-1} as published elsewhere.[14b] 2x 950 µL of protein sample, diluted in PSE-buffer if necessary, were filled into plastic cuvettes and placed in a double-beam photometer (Perkin Elmer). One of the samples was exposed to CO for a few seconds. Next, 50 µL of a 1 M sodium dithionite stock solution was added to reduce the heme-Fe(III) to Fe(II) and a difference spectrum between 400 and 500 nm was recorded. The measurements were continued until a constant absorption maximum was reached.

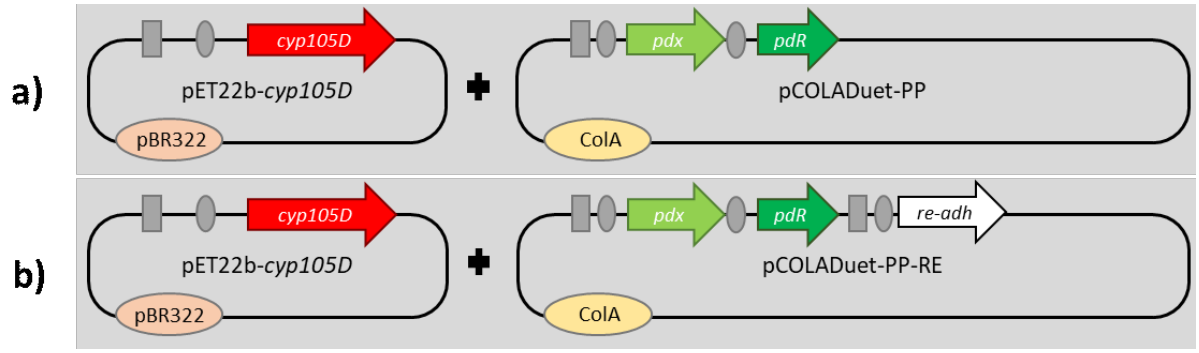
ADH-activity was measured in a continuous photometric assay monitoring NADH formation at 340 nm (ε_{340} = 6.22 mM^{-1} cm^{-1}) in presence of propan-2-ol as substrate. Reaction mixtures contained 50 mM Tris-HCl with 10 mM $MgCl_2$ (pH 8), 649 mM propan-2-ol (5 % (v/v)) and 50 µL of crude cell extract in appropriate dilution. After incubation for 2 min at 25 °C, the reaction was started by adding 0.5 mM NAD^+. The increase of absorption caused by NADH formation was tracked for 120 s at 25 °C in a double-beam photometer (Perkin Elmer). The initial slope (ΔA_{340}/min) between 20-80 s was linear and thus used to calculate the activity

[U/g_{CDW}]. 1 U is defined as the amount of enzyme which is needed to convert 1 µmol substrate in 1 minute under assay conditions. All measurements were done in duplicate.

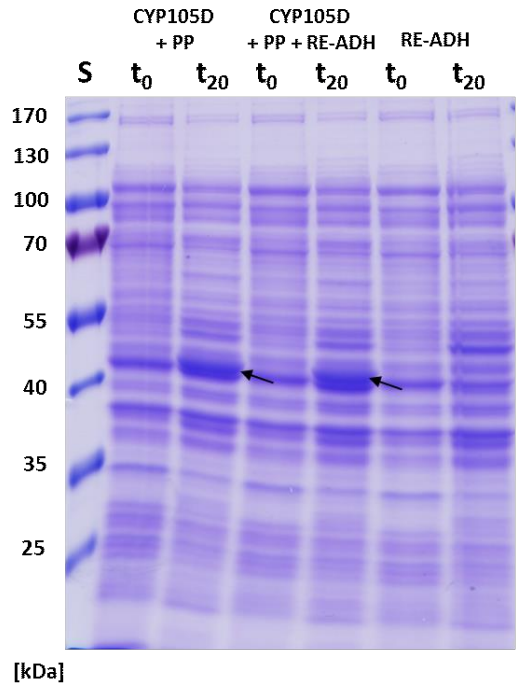
Product analysis

Product analysis was conducted by liquid chromatography coupled to mass spectrometry (LC/MS) on a Prominence/LCMS 2020 device (Shimadzu). Analytes were separated with a flow rate of 1 mL at 30 °C on a Chromolith® Performance RP-18e column (100 x 4.6 mm, Merck) using methanol as solvent B and ddH_2O with 0.1 % formic acid as solvent A. 1 µL of each sample was injected. The substances were ionized by electron spray ionization (ESI) and atmospheric pressure chemical ionization (APCI) in a dual ionization mode. Masses were detected in positive scan mode in a range between 100 – 500 *m/z*. Additionally, PDA chromatograms at 254 nm were recorded. The conditions for chromatographic separation were carried out as described previously.[11]

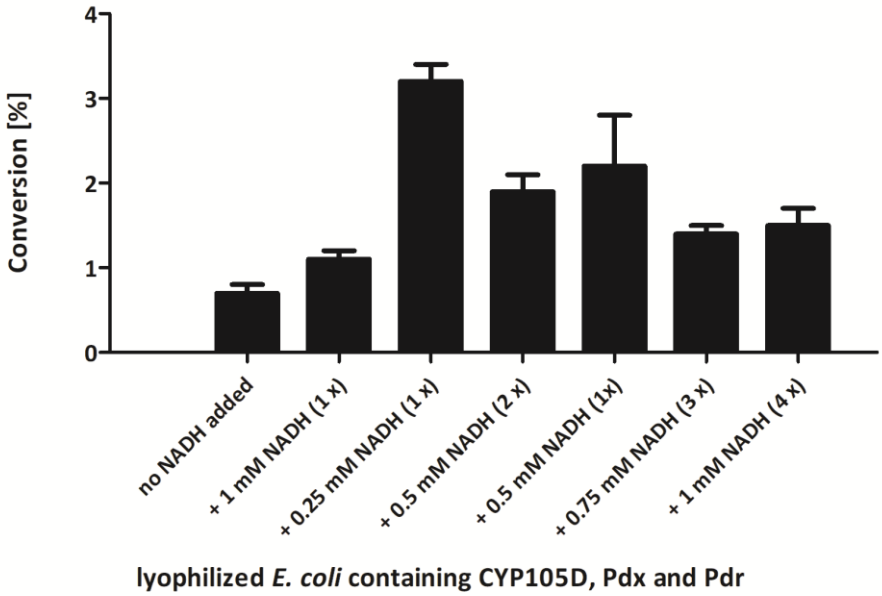
Supplementary Results



Supplementary Figure 2.5.1: Plasmid combinations used in this study. The P450 (*cyp105D*) is always encoded on the pET22b-vector. Redox partner genes (*pdx/pdr*) are integrated in MCSI of pCOLADuet. The *re-adh* (b) is integrated in the MCSII of the pCOLADuet-vector. Gray squares represent the T7-promoter; gray circles indicate the ribosome binding site.



Supplementary Figure 2.5.2 SDS-PAGE analysis of *E. coli* C43 (DE3) strains for whole-cell biocatalysis. The picture shows the comparison of whole-cell samples before induction (t_0) and 20 h after induction (t_{20}) as follows: *E. coli* C43 (DE3) pET22b-*cyp105D* + pCOLADuet-PP (lane 1 and 2), *E. coli* C43 (DE3) pET22b-*cyp105D* + pCOLADuet-PPRE (lane 3 and 4), *E. coli* C43 (DE3) pCOLADuet-Re-ADH (lane 5 and 6), PageRuler prestained protein ladder (lane S). CYP105D (43.9 kDa) is marked with a black arrow. Pdx (11.5 kDa), Pdr (45.8 kDa) and Re-ADH (36.2 kDa) are not seen.



Supplementary Figure 2.5.3: Effect of NADH addition on testosterone conversion mediated by the lyophilized whole-cell catalyst without ADH. NADH was added up to four times (number in brackets) every 2 h. Reaction conditions: 1 mM testosterone in 5 % (v/v) propan-2-ol final concentration, 25 °C 1100 rpm shaking frequency, 0.5 mL reaction volume in 2 mL tubes, 10 mg/mL lyophilized cells. 0.25 mM NADH was added up to four times at 0 h, 2 h, 4 h and 6 h incubation. The biotransformation was performed in technical duplicates.

3. General discussion

In the following, the most important results of the individual chapters are summarized, and discussed according to the literature.

3.1 Identification of P450s with target activity

The identification of a P450 capable of oxidizing a compound of interest is a challenging task. Hence, different screening methods have been described (see chapter 1.2.7). The chromatographic separation and mass spectrometric detection of the products, formed in the presence of a certain P450 enzyme, is particularly useful to directly analyze potential P450 substrates and product selectivity. However, it is decisive how the screening is designed in order to identify a P450 with the desired activity. If substrates of P450s are screened randomly, candidates with a broad substrate spectrum can be identified. However, it is not possible with randomized screening approaches to evaluate the influence of small changes in substrate structure on enzyme activity and selectivity. For this purpose, substrates with similar structure and different number and type of functional groups have to be tested. However, in this case, it can be not evaluated if the tested P450 has a broad or narrow substrate spectrum. Within this PhD project, a two-step screening was developed to identify bacterial P450s with activity towards drugs (2.2). In the first step, three model substrates allowed selection of the most promising P450s with activity towards the target substrates. In the second step, testing of structurally related molecules and derivatives of the model substrates provided information about the substrate scope and chemo -and regioselectivity of the selected P450 candidates from the first step (2.2). In this way, the advantages of both screening approaches described above could be combined.

The here developed two-step screening can be combined with other methods to obtain a higher throughput. For example, the selection of promising P450 enzymes in the first step of the two-step-screening can be facilitated by applying a ‘cocktail approach’. Cocktail approaches are less time-consuming screening methods because a mixture of potential P450 substrates is used for screening which can then be analyzed simultaneously via LC/MS or (U)HPLC-MS/MS.[175] However, this requires that the tested substrates can be easily

distinguished from each other and that there is no positive or negative effect on the conversion of the other substrates.[175] Additionally, different photometric methods can be implemented in the two-step screening. A bioluminescence assay could be particularly useful because it was used for screening of 17,143 compounds aiming at identification of novel substrates and inhibitors of human P450s. [176] In this assay, the tested P450 converts surrogate substrates to luciferin, which is then used by the auxiliary enzyme luciferase to generate a luminescent signal.[176] In the case a compound is an inhibitor or a substrate of the P450, the amount of light produced by luciferase gets reduced. This assumes, however, that no unwanted interaction between the auxiliary substrate and the substrate takes place and that a suitable surrogate substrate for the tested P450s is available. Overall, the developed two-step screening enlarges the methodologies of straight-forward approaches that facilitate the identification of a P450 with a target activity.

3.2 Substrate spectra of CYP105D and CYP107Z from *S. platensis*

The differences in substrate scopes of P450s are not easy to rationalize although a systematic nomenclature based on sequence alignments exists (1.2.2) which might provide first indications for substrates of uncharacterized members.[121, 177] In case of the CYP105 and CYP107 families the prediction of substrates is especially challenging. Many proteins have been assigned to these P450 families, but only a few have been characterized. Of the more than 41,000 P450 sequences, 2554 sequences are assigned to the CYP107 family and 1225 to the CYP105 family (https://cyped.biocatnet.de/ as of April 2020).[19a] Of these 3779 sequences, individual isoforms have often been characterized with respect to the oxidation of a particular substrate. In the following, CYP105D and CYP107Z are compared with characterized members of their phylogenetic families.

3.2.1 The CYP105 family

CYP105D from *S. platensis* was identified as promiscuous P450 which oxidizes a broad range of chemically diverse substrates (chapter 2.2, 2.3 and 2.4). Generally, the CYP105 family is characterized by broad substrate specificities and diverse catalytic functions.[129] Members of

this family are predominately involved in synthesis of rather big substrates like polyketides including amphotericin B (924.1 g/mol)[178] or tylosin (916.1 g/mol).[109, 179] The oxidation of such big macrolides by CYP105 enzymes are typically catalyzed with a high regioselectivity. This includes also the commercialized production of the cholesterol-lowering drug pravastatin, in which CYP105A3 selectively hydroxylates the substrate compactin at position 6β.[180]

CYP105A1 from *Streptomyces griseolus* is perhaps the best studied P450 of the CYP105 family with regard to activities against chemically diverse substrates. A comparison of the substrates scopes of CYP105D and CYP105A1 is given in Table 3.2.1. CYP105A1 was identified in studies towards the biotransformation of herbicides, such as chlorsulforon[181] or sulfometuron methyl.[181] In course of its potential use as biocatalyst, further substrates of CYP105A1 such as 7-ethoxycoumarin[182], ionones[183] and vitamin D forms[184] were found. More recently described examples extended the substrate scope to resin acid diterpenoids[185] and hypoglycemic drugs.[186] Overall, CYP105A1 accepts mainly small substrates, such as 7-ethoxycoumarin (190.2 g/mol) or α-and β-ionone (192.3 g/mol). The substrates with the highest molecular weight were glimepiride (490.7 g/mol) and glibenclamide (494.0 g/mol). The broad substrate promiscuity of CYP105A1 might be explained by the size of its binding pocket. The analysis of crystal structures of CYP105A1 in absence and in presence of the product 1α,25-dihydroxyvitamin D_3 (416,64 g/mol) revealed a large binding pocket that suggests accommodation of even bigger compounds than 1α,25-dihydroxyvitamin D_3 (416,64 g/mol).[96]

The substrate spectrum of CYP105D comprises chemically different compounds, most of which are terpenes and drugs like anti-HIV compounds, tricyclic antidepressants or synthetic glucocorticoids. Similar to CYP105A1, CYP105D can accept small substrates such as geraniol (154.3 g/mol), but especially large substrates such as ritonavir (720.9 g/mol) or atazanavir (704.9 g/mol). Although CYP105D differs in substrate scope from CYP105A1 it can be stated that this enzyme possesses a similar or even broader substrate scope. The binding pocket of CYP105D might be probably even bigger than those of CYP105A because substrates with a higher molecular weight were accepted. The idea is supported when looking at the regioselectivity of CYP105D. Because CYP105D oxidizes big substrates including ritonavir with a high regioselectivity of up to 86 %,[11] the substrates must get close to heme iron so

that positions in the vicinity of the heme iron can interact with the substrate and thus influence the orientation of substrate.[187] However, since P450s have large, diverse and flexible substrate binding cavities,[187] CYP105D needs to be further characterized structurally to verify these statements.

Table 3.2.1: Comparison of the substrate spectrum between CYP105A1 from *S. griseolus* and CYP105D from *S. platensis*. The substrates for CYP105D can be found in the chapters 2.2, 2.3 and 2.4. Selected substrates for CYP105D were identified in course of a master thesis.[188] Substrates marked with an asterix (*) were identfied in course of the 'HyPerIn' or the 'DAAD' project. These data are not or only partially included in this thesis.

CYP105A1 from *S. griseolus*		**CYP105D from *S. platensis***		
chlorsulforon[181]	sulfometuron methyl[189]	ritonavir	amprenavir	darunavir
chlorimuron ethyl[189]	W5822[190]	atazanavir	saquinavir	indinavir
7-ethoxycoumarin[182]	R7402[190]	testosterone	progesterone	medrysone
vitamin D_3[184]	vitamin D_2[184]	hydrocortisone	prednisolone	prednisone
25-hydroxyvitamin D_3[184]	1α-hydroxyvitamin D_2[184]	amitriptyline	imipramine	clomipramine
testosterone[100]	α-ionone[183]	cyclobenzaprine	nortriptyline	protriptyline
β-ionone[183]	glibenclamide[186]	opipramol	chenodeoxycholic acid	deoxycholic acid
glimepiride[186]	abietic acid[185]	lithocholic acid	ursodeoxycholic acid	cholic acid
dehydroabietic acid[185]	isopimaric acid[185]	diclofenac[188]	(*S*)-ketamine[188]	quinidine[188]
		(3a*R*)- (+)-sclareolide[188]*	abietic acid[188]	lupeol[188]
		thioridazine	terfenadine	geraniol
		dextromethorphan	amodiaquine	

3.2.2 The CYP107 family

CYP107Z from *S. platensis* has a broad substrate spectrum and shows high activity against drugs, outperforming other investigated members of the CYP107 family (chapter 2.2 and 2.3). Similar to the CYP105 family, members of the CYP107 family oxidize a broad range of chemically diverse substrates. Most substrates of characterized CYP107 enzymes have a high

molecular weight which includes physiological substrates like the macrolides pikromycin[89] and erythromycin,[90c] but also artificial substrates like avermectin[191]. Crystal structures of CYP107 enzymes including CYP107W1[192] and CYP107A1 (EryF)[193] revealed large binding pockets to accommodate their big substrates.

Most members of the CYP107 family were characterized against certain substrate classes. This includes especially the already mentioned macrolides. With regard to the oxidation of non-natural substrates including drugs like terfenadine[95]or vitamin D_3[97], only a few P450s were investigated. In this regard, CYP107Z might be an interesting candidate for biocatalytic applications because of its broad substrate spectrum (Table 3.2.2). The substrate spectrum of CYP107Z is largely complementary to that of CYP105D. However, substrates like valencene or dexamethasone have only been converted by CYP107Z. Additonally, CYP107Z showed a higher activity against most substrates compared to CYP105D (chapter 2.2 and 2.3). The large substrate binding pockets of already crystallised CYP107 enzymes suggest that such large binding pocket is also present in CYP107Z, which would somehow explain its broad substrate spectrum. If this is the case, protein engineering could be used to improve the selectivity of CYP107Z, which seems to be rather low at the moment compared to other CYP107 enzymes.

In general, it can be stated that CYP105D and CYP107Z are P450 enzymes with high potential for drug metabolite synthesis. Compared to other P450s from the CYP105 and CYP107 families, they have comparable or even broader substrate scopes. The activities and selectivities of CYP105D and CYP107Z are lower compared to evolved P450 enzymes like CYP102A1 that perform stereo-and regioselective oxidations of drugs. The activity and selectivity might be enhanced by protein engineering methods like random mutagenesis or rational protein design as well as reaction engineering.[17] These approaches were especially successful for CYP102A1. A great number of examples exist in literature, which describe the improvement of CYP102A1 by such means.[52, 56b, 56f, 194]

Table 3.2.2: Overview of the substrate spectrum of CYP107Z from *S. platensis*. The substrates for CYP107Z can be found in the chapters 2.2, 2.3 and 2.4. Selected substrates for CYP105D were identified in course of a master thesis.[188] Substrates marked with an asterix (*) were identfied in course of the 'HyPerIn' or the 'DAAD' project. These data are not or only partially included in this thesis.

Substrates accepted by CYP107Z from *S. platensis*				
ritonavir	amprenavir	darunavir	atazanavir	saquinavir
indinavir	testosterone	progesterone	medrysone	hydrocortisone
prednisolone	prednisone	amitriptyline	imipramine	clomipramine
cyclobenzaprine	nortriptyline	protriptyline	opipramol	thioridazine
chenodeoxycholic acid	deoxycholic acid*	lithocholic acid*	ursodeoxycholic acid*	cholic acid*
diclofenac[188]	(*S*)-ketamine[188]	(3a*R*)- (+)-sclareolide[188]*	abietic acid[188]	lupeol[188]
dextromethorphan	terfenadine	quinidine[188]	amodiaquine	geraniol
6α-methylprednisolone	valencene*	dexamethasone	betamethasone	farnesol[188]

3.3 P450 based whole-cell biocatalysis

Lyophilized cells are particular interesting for application, because they facilitate handling and enable easier storage compared to wet cells. On top of that, lyophilized cells can be applied in reaction systems with higher amount of organic co-solvent, allowing higher concentrations of hydrophobic substrates.[195] This is particularly interesting because substrates of P450s are typically hydrophobic. A CYP105D-based whole cell biocatalyst was developed with the aim of establishing a procedure that is based on the use of lyophilized *E. coli* cells (2.5). The hydroxylation of testosterone to 2β-hydroxytestosterone was chosen as model reaction. Initially, the activity of lyophilized cells (≤ 1% conversion) of the CYP105D-based whole cell system was very low compared to the activity of wet cells (31% conversion). The lower activity of lyophilized cells could be attributed to insufficient cofactor regeneration. When Re-ADH was co-expressed for cofactor regeneration, activities were comparable or even higher between lyophilized and wet cells. Under best conditions, a conversion of 72% of 1 mM substrate was achieved. This activity is in the same range which was observed with isolated enzymes.[11]

Although the limitations for implementation of lyophilized P450 whole cell systems could be identified, some questions still remain. Lyophilized whole cell catalysts with NAD(P)H-dependent enzymes such as alcohol dehydrogenases(ADH)[195] or imine reductases (IRED)[196] have already been described. Cofactor regeneration is enabled by an additional co-substrate such as cyclohexanol for ADHs and glucose for IREDs. Co-substrates like cyclohexanol can be directly used by ADHs for cofactor regeneration. Co-substrates like glucose must first be metabolized via the host metabolism (e.g. by glycolysis or the pentose phosphate pathway) in order to provide NAD(P)H for the reaction. This route was chosen for the lyophilized P450-based catalysts. However, in contrast to what is described in the literature, hardly any P450 activity was observed, when *re-adh* was not expressed (2.5). Therefore, it is questionable what the exact cause for the lack of P450 activity is, when no additional cofactor-regenerating enzyme is present. Possible reasons such as a lower activity of the *E.coli* metabolism for NADH cofactor regeneration or unfavorable reaction conditions must be investigated more closely in future studies.

3.4 Future perspectives

This work presents the identification of 28 P450 enzymes from *P. autotrophica* and *S. platensis*, which have been evaluated for their ability to oxidize drugs. Along with different aspects for identification of P450s with target activities (discussed in 3.1), CYP105D and CYP107Z were characterized for diverse oxidations of drug compounds (discussed in 3.2). Additionally, a procedure for the preparation of lyophilized P450 based whole-cell biocatalyst was developed (discussed in 3.3).

S. platensis proved to be an interesting source for new P450 enzymes. In this regard it would be worthwhile to investigate redox partner proteins from *S. platensis*. A recent review of *Li et al*. emphasized that 'not only the type, the amount, the combination, and the mode of action of bacterial redox partners affect the catalytic rate and product distribution but also the type and selectivity of P450 reactions'.[22] Accordingly, the characterization of additional redox partner proteins could provide new insights into the activity and selectivity of P450 from *S. platensis*.

The most promising drug-metabolizing P450s in this thesis were CYP105D and CYP107Z. Ten further representatives of the CYP105D family and 18 further representatives of the CYP107Z family are currently assigned in the databases. Although they have a rather 'high' sequence identity of 55% to each other, their substrate scope is quite heterogeneous. The already mentioned CYP105D7 from *S. grisoleus* accepts substrates such as 1-deoxypentalenic acid, diclofenac, naringenin, compactin, and testosterone.[98] CYP107Z enzymes were described for conversions of macrolides such as avermectin[191] or peptides such as actinomycin.[197] In this context, it might be interesting to study the substrate scope more intensively. This could eventually help to identify sequence-function relationships.

The structural diversity of the substrates described so far suggests that CYP105D and CYP107Z are also capable of converting a large number of other interesting substances. Current projects of the Institute of Biochemistry II deal with the oxidation of terpenoids and lignans. These are two classes of plant natural products with interesting biological properties. Interestingly, CYP105D and CYP107Z accept these compounds as substrates, albeit with low activities and insufficient selectivities (unpublished data A. Raffaele and D. Decembrino). As mentioned above, protein and reaction engineering can be applied to improve the activity and selectivity, thus opening the way for the establishment of CYP105D and CYP107Z in a biocatalytic process.

As pointed out in chapter 3.3, the use of lyophilized P450 based whole-cell biocatalysts offers advantages with regard to their application in reaction systems with a higher proportion of organic solvents. As P450 substrates are typically hydrophobic, the substrate concentration could be increased in such systems, which might lead to a higher P450 activity. However, P450 based processes in non-conventional media bring new challenges, such as altered oxygen solubility in organic solvents or the safe handling of increased concentration of organic solvents.[164]

4. References

[1] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185-194.

[2] R. A. Sheldon, *Green biocatalysis* **2016**, 1-15.

[3] K. Faber, *Biotransformations in organic chemistry, Vol. 4*, Springer, **1992**.

[4] a) U. T. Bornscheuer, *Philos.Trans. R. Soc. London Ser.A* **2018**, *376*; b) E. M. M. Abdelraheem, H. Busch, U. Hanefeld, F. Tonin, *React. Chem. Eng.* **2019**, *4*, 1878-1894.

[5] A. Fryszkowska, P. N. Devine, *Curr. Opin. Chem. Biol.* **2020**, *55*, 151-160.

[6] A. Schmid, J. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, *Nature* **2001**, *409*, 258-268.

[7] E. M. Gillam, M. A. Hayes, *Curr. Top. Med. Chem.* **2013**, *13*, 2254-2280.

[8] A. D. Roses, *Nat. Rev. Drug Discov.* **2008**, *7*, 807-817.

[9] D. J. Pollard, J. M. Woodley, *Trends Biotechnol.* **2007**, *25*, 66-73.

[10] S. D. Copley, *Curr. Opin. Struct. Biol.* **2017**, *47*, 167-175.

[11] T. Hilberath, L. M. Windeln, D. Decembrino, P. Le-Huu, F. L. Bilsing, V. B. Urlacher, *ChemCatChem* **2020**, *12*, 1710-1719.

[12] a) E. M. Gillam, *Clin. Exp. Pharmacol. Physiol.* **2005**, *32*, 147-152; b) M. Winkler, M. Geier, S. P. Hanlon, B. Nidetzky, A. Glieder, *Angew. Chem. Int. Ed.* **2018**, *57*, 13406-13423.

[13] J. A. Williams, R. Hyland, B. C. Jones, D. A. Smith, S. Hurst, T. C. Goosen, V. Peterkin, J. R. Koup, S. E. Ball, *Drug Metab. Dispos.* **2004**, *32*, 1201-1208.

[14] a) F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta-General Subjects* **2007**, *1770*, 330-344; b) T. Omura, R. Sato, *J. Biol. Chem.* **1964**, *239*, 2370-2378.

[15] R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 6185-6203.

[16] a) R. Bernhardt, *J. Biotechnol.* **2006**, *124*, 128-145; b) M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, *Chem. Rev.* **1996**, *96*, 2841-2888; c) B. Meunier, S. P. De Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947-3980; d) F. P. Guengerich, A. W. Munro, *J. Biol. Chem.* **2013**, *288*, 17065-17073.

[17] J. A. McIntosh, C. C. Farwell, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2014**, *19*, 126-134.

[18] D. C. Lamb, L. Lei, A. G. S. Warrilow, G. I. Lepesheva, J. G. L. Mullins, M. R. Waterman, S. L. Kelly, *J. Virol.* **2009**, *83*, 8266-8269.

[19] a) D. R. Nelson, *Biochim. Biophys. Acta-Proteins and Proteomics* **2018**, *1866*, 141-154; b) V. B. Urlacher, M. Girhard, *Trends Biotechnol.* **2019**.

[20] D. R. Nelson, *Hum. Genomics* **2009**, *4*, 59-65.

[21] D. R. Nelson, T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus, O. Gotoh, *DNA Cell Biol.* **1993**, *12*, 1-51.

[22] S. Li, L. Du, R. Bernhardt, *Trends Microbiol.* **2020**.

[23] V. B. Urlacher, M. Girhard, *Trends Biotechnol.* **2012**, *30*, 26-36.

[24] A. J. Warman, O. Roitel, R. Neeli, H. M. Girvan, H. E. Seward, S. A. Murray, K. J. McLean, M. G. Joyce, H. Toogood, R. A. Holt, Portland Press Ltd., **2005**.

[25] a) T. Kitazume, N. Takaya, N. Nakayama, H. Shoun, *J. Biol. Chem.* **2000**, *275*, 39734-39740; b) H. Ichinose, H. Wariishi, *Biochem. Biophys. Res. Commun.* **2013**, *438*, 289-294.

[26] a) C. M. Krest, E. L. Onderko, T. H. Yosca, J. C. Calixto, R. F. Karp, J. Livada, J. Rittle, M. T. Green, *J. Biol. Chem.* **2013**, *288*, 17074-17081; b) I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chem. Rev.* **2005**, *105*, 2253-2278.
[27] S. G. Sligar, *Biochemistry* **1976**, *15*, 5399-5406.
[28] J. Rittle, M. T. Green, *Science* **2010**, *330*, 933-937.
[29] J. P. T. Zaragoza, T. H. Yosca, M. A. Siegler, P. Moënne-Loccoz, M. T. Green, D. P. Goldberg, *J. Am. Chem. Soc.* **2017**, *139*, 13640-13643.
[30] P. R. O. de Montellano, in *Cytochrome P450*, Springer, **2015**, pp. 111-176.
[31] I. I. Karuzina, A. I. Archakov, *Free Radic. Biol. Med.* **1994**, *16*, 73-97.
[32] M. Bureik, R. Bernhardt, *Modern Biooxidation* **2007**, *155*, 176.
[33] M. V. Donova, O. V. Egorova, *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1423-1447.
[34] a) M. H. J. Zuidweg, W. F. Van der Waard, J. De Flines, *Biochim. Biophys. Acta* **1962**, *58*, 131-133; b) K. Petzoldt, K. Annen, H. Laurent, R. Wiechert*, Vol. US4353985 A*, **1982**.
[35] a) D. H. Peterson, H. C. Murray, S. H. Eppstein, L. M. Reineke, A. Weintraub, P. D. Meister, H. M. Leigh, *J. Am. Chem. Soc.* **1952**, *74*, 5933-5936; b) D. H. Peterson, H. C. Murray, *J. Am. Chem. Soc.* **1952**, *74*, 1871-1872.
[36] a) D. C. Lamb, M. R. Waterman, S. L. Kelly, F. P. Guengerich, *Curr. Opin. Biotechnol.* **2007**, *18*, 504-512; b) L. M. Podust, D. H. Sherman, *Nat. Prod. Rep.* **2012**, *29*, 1251-1266.
[37] D. K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong, J. D. Keasling, *Nature* **2006**, *440*, 940-943.
[38] M. Peplow, *Nature* **2016**, *530*, 389.
[39] M. Vermeir, S. Lachau-Durand, G. Mannens, F. Cuyckens, B. van Hoof, A. Raoof, *Drug Metab. Dispos.* **2009**, *37*, 809-820.
[40] W. Jacobsen, B. Kuhn, A. Soldner, G. Kirchner, K.-F. Sewing, P. A. Kollman, L. Z. Benet, U. Christians, *Drug Metab. Dispos.* **2000**, *28*, 1369-1378.
[41] a) R. J. Lantz, T. A. Gillespie, T. J. Rash, F. Kuo, M. Skinner, H. Y. Kuan, M. P. Knadler, *Drug Metab. Dispos.* **2003**, *31*, 1142-1150; b) G. V. Bakken, I. Rudberg, H. Christensen, E. Molden, H. Refsum, M. Hermann, *Drug Metab. Dispos.* **2009**, *37*, 254-258; c) K. M. Kirschbaum, M. J. Müller, J. Malevani, A. Mobascher, C. Burchardt, M. Piel, C. Hiemke, *World J. Biol. Psychiatry* **2008**, *9*, 212-218.
[42] K. O. Borg, E. Carlsson, K. J. Hoffmann, T. E. Jönsson, H. Thorin, B. Wallin, *Acta Pharmacol. Toxicol. (Copenh.)* **1975**, *36*, 125-135.
[43] a) D. C. Dahlin, G. T. Miwa, A. Y. Lu, S. D. Nelson, *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 1327-1331; b) B. Lalovic, B. Phillips, L. L. Risler, W. Howald, D. D. Shen, *Drug Metab. Dispos.* **2004**, *32*, 447-454; c) S. K. Paulson, J. Y. Zhang, A. P. Breau, J. D. Hribar, N. W. K. Liu, S. M. Jessen, Y. M. Lawal, J. N. Cogburn, C. J. Gresk, C. S. Markos, *Drug Metab. Dispos.* **2000**, *28*, 514-521.
[44] T. Andersson, J. O. Miners, M. E. Veronese, W. Tassaneeyakul, W. Tassaneeyakul, U. A. Meyer, D. J. Birkett, *Br. J. Clin. Pharmacol.* **1993**, *36*, 521-530.
[45] D. K. Walker, *Xenobiotica* **1999**, *29*, 297-310.
[46] H.-P. Gschwind, U. Pfaar, F. Waldmeier, M. Zollinger, C. Sayer, P. Zbinden, M. Hayes, R. Pokorny, M. Seiberling, M. Ben-Am, *Drug Metab. Dispos.* **2005**, *33*, 1503-1512.
[47] a) J. Knowles, G. Gromo, *Nat. Rev. Drug Discov.* **2003**, *2*, 63-69; b) E. Vitaku, B. R. Smith, D. T. Smith, J. T. Njardarson, **2013**.
[48] E. O'Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chem. Commun.* **2011**, *47*, 2490-2501.

[49] M. K. Julsing, S. Cornelissen, B. Bühler, A. Schmid, *Curr. Opin. Chem. Biol.* **2008**, *12*, 177-186.
[50] W. A. Duetz, J. B. Van Beilen, B. Witholt, *Curr. Opin. Biotechnol.* **2001**, *12*, 419-425.
[51] M. A. Noble, C. S. Miles, S. K. Chapman, D. A. Lysek, A. C. MacKay, G. A. Reid, R. P. Hanzlik, A. W. Munro, *Biochem. J.* **1999**, *339*, 371-379.
[52] C. J. Whitehouse, S. G. Bell, L.-L. Wong, *Chem. Soc. Rev.* **2012**, *41*, 1218-1260.
[53] E. M. J. Gillam, T. Baba, B.-R. Kim, S. Ohmori, F. P. Guengerich, *Arch. Biochem. Biophys.* **1993**, *305*, 123-131.
[54] M. C. Chang, R. A. Eachus, W. Trieu, D.-K. Ro, J. D. Keasling, *Nat. Chem. Biol.* **2007**, *3*, 274-277.
[55] H. Renault, J.-E. Bassard, B. Hamberger, D. Werck-Reichhart, *Curr. Opin. Plant Biol.* **2014**, *19*, 27-34.
[56] a) G. Di Nardo, G. Gilardi, *Int. J. Mol. Sci.* **2012**, *13*, 15901-15924; b) A. M. Sawayama, M. M. Y. Chen, P. Kulanthaivel, M. S. Kuo, H. Hemmerle, F. H. Arnold, *Chem. Eur. J.* **2009**, *15*, 11723-11729; c) H. Venkataraman, M. C. A. Verkade-Vreeker, L. Capoferri, D. P. Geerke, N. P. E. Vermeulen, J. N. M. Commandeur, *Biorg. Med. Chem.* **2014**, *22*, 5613-5620; d) C. F. Butler, C. Peet, K. J. McLean, M. T. Baynham, R. T. Blankley, K. Fisher, S. E. Rigby, D. Leys, M. W. Voice, A. W. Munro, *Biochem. J.* **2014**, *460*, 247-259; e) N. Beyer, J. K. Kulig, M. W. Fraaije, M. A. Hayes, D. B. Janssen, *ChemBioChem* **2018**, *19*, 326-337; f) X. Ren, J. A. Yorke, E. Taylor, T. Zhang, W. Zhou, L. L. Wong, *Chem. Eur. J.* **2015**, *21*, 15039-15047; g) C. G. Acevedo-Rocha, C. G. Gamble, R. Lonsdale, A. Li, N. Nett, S. Hoebenreich, J. B. Lingnau, C. Wirtz, C. Fares, H. Hinrichs, A. Deege, A. J. Mulholland, Y. Nov, D. Leys, K. J. McLean, A. W. Munro, M. T. Reetz, *ACS Catal.* **2018**, *8*, 3395-3410; h) S. Kille, F. E. Zilly, J. P. Acevedo, M. T. Reetz, *Nat. Chem.* **2011**, *3*, 738-743.
[57] L. M. Schmitz, J. Schäper, K. Rosenthal, S. Lütz, *ChemCatChem* **2019**, *11*, 5766-5777.
[58] a) N.-H. Tran, D. Nguyen, S. Dwaraknath, S. Mahadevan, G. Chavez, A. Nguyen, T. Dao, S. Mullen, T.-A. Nguyen, L. E. Cheruzel, *J. Am. Chem. Soc.* **2013**, *135*, 14484-14487; b) V. V. Shumyantseva, T. V. Bulko, A. I. Archakov, *J. Inorg. Biochem.* **2005**, *99*, 1051-1063; c) E. G. Hrycay, S. M. Bandiera, *Arch. Biochem. Biophys.* **2012**, *522*, 71-89.
[59] M. Sandroni, Y. Pellegrin, F. Odobel, *C. R. Chim.* **2016**, *19*, 79-93.
[60] H. Hirakawa, N. Kamiya, T. Tanaka, T. Nagamune, *Protein Eng. Des. Sel.* **2007**, *20*, 453-459.
[61] Y. Kitahama, M. Nakamura, Y. Yoshida, Y. Aoyama, *Biol. Pharm. Bull.* **2009**, *32*, 558-563.
[62] a) P. Hlavica, *Biotechnol. Adv.* **2009**, *27*, 103-121; b) P. J. Bakkes, J. L. Riehm, T. Sagadin, A. Rühlmann, P. Schubert, S. Biemann, M. Girhard, M. C. Hutter, R. Bernhardt, V. B. Urlacher, *Sci. Rep.* **2017**, *7*, 1-13; c) Y. Khatri, A. Schifrin, R. Bernhardt, *FEBS Lett.* **2017**, *591*, 1126-1140.
[63] J. Wachtmeister, D. Rother, *Curr. Opin. Biotechnol.* **2016**, *42*, 169-177.
[64] M. T. Lundemo, J. M. Woodley, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 2465-2483.
[65] a) D. Zehentgruber, F. Hannemann, S. Bleif, R. Bernhardt, S. Lütz, *ChemBioChem* **2010**, *11*, 713-721; b) H. Schewe, B.-A. Kaup, J. Schrader, *Appl. Microbiol. Biotechnol.* **2008**, *78*, 55-65.
[66] a) J. B. van Beilen, E. G. Funhoff, A. van Loon, A. Just, L. Kaysser, M. Bouza, R. Holtackers, M. Röthlisberger, Z. Li, B. Witholt, *Appl. Environ. Microbiol.* **2006**, *72*, 59-65; b) H. Schewe, D. Holtmann, J. Schrader, *Appl. Microbiol. Biotechnol.* **2009**, *83*,

849-857; c) S. Cornelissen, M. K. Julsing, J. Volmer, O. Riechert, A. Schmid, B. Bühler, *Biotechnol. Bioeng.* **2013**, *110*, 1282-1292.
[67] S. Cornelissen, S. Liu, A. T. Deshmukh, A. Schmid, B. Bühler, *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1359-1370.
[68] S. Schulz, PhD thesis, Heinrich-Heine-University Düsseldorf **2015**.
[69] a) J. Mi, D. Becher, P. Lubuta, S. Dany, K. Tusch, H. Schewe, M. Buchhaupt, J. Schrader, *Microb. Cell Fact.* **2014**, *13*, 170; b) F. Tieves, I. N. Erenburg, O. Mahmoud, V. B. Urlacher, *Biotechnol. Bioeng.* **2016**, *113*, 1845-1852; c) R. Karande, L. Debor, D. Salamanca, F. Bogdahn, K. H. Engesser, K. Buehler, A. Schmid, *Biotechnol. Bioeng.* **2016**, *113*, 52-61.
[70] a) M. K. Julsing, M. Schrewe, S. Cornelissen, I. Hermann, A. Schmid, B. Bühler, *Appl. Environ. Microbiol.* **2012**, *78*, 5724-5733; b) S. Schneider, M. G. Wubbolts, D. Sanglard, B. Witholt, *Appl. Environ. Microbiol.* **1998**, *64*, 3784-3790; c) A. J. Ruff, M. Arlt, M. van Ohlen, T. Kardashliev, M. Konarzycka-Bessler, M. Bocola, A. Dennig, V. B. Urlacher, U. Schwaneberg, *J. Mol. Catal. B: Enzym.* **2016**, *134*, 285-294.
[71] S. Janocha, R. Bernhardt, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7639-7649.
[72] a) I. Kaluzna, T. Schmitges, H. Straatman, D. van Tegelen, M. Müller, M. Schürmann, D. Mink, *Org. Process Res. Dev.* **2016**, *20*, 814-819; b) S. Pflug, S. M. Richter, V. B. Urlacher, *J. Biotechnol.* **2007**, *129*, 481-488; c) S. Liu, C. Li, X. Fang, Z. Cao, *Enzyme Microb. Technol.* **2004**, *34*, 73-77.
[73] A. Amberg, *In silico methods*, Springer, **2013**.
[74] D. C. Lamb, F. P. Guengerich, S. L. Kelly, M. R. Waterman, *Expert Opin. Drug Metab. Toxicol.* **2006**, *2*, 27-40.
[75] F. P. Guengerich, Z. Tang, Q. Cheng, S. G. Salamanca-Pinzón, *Biochim. Biophys. Acta-Proteins and Proteomics* **2011**, *1814*, 139-145.
[76] Q. Cheng, D. C. Lamb, S. L. Kelly, L. Lei, F. P. Guengerich, *J. Am. Chem. Soc.* **2010**, *132*, 15173-15175.
[77] S. Bleif, F. Hannemann, M. Lisurek, J. P. von Kries, J. Zapp, M. Dietzen, I. Antes, R. Bernhardt, *ChemBioChem* **2011**, *12*, 576-582.
[78] Y. Khatri, M. Ringle, M. Lisurek, J. P. von Kries, J. Zapp, R. Bernhardt, *ChemBioChem* **2016**, *17*, 90-101.
[79] a) C. von Bühler, P. Le-Huu, V. B. Urlacher, *ChemBioChem* **2013**, *14*, 2189-2198; b) B. Simgen, J. Contzen, R. Schwarzer, R. Bernhardt, C. Jung, *Biochem. Biophys. Res. Commun.* **2000**, *269*, 737-742; c) M. Girhard, T. Klaus, Y. Khatri, R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2010**, *87*, 595-607.
[80] A. Fulton, M. R. Hayes, U. Schwaneberg, J. Pietruszka, K.-E. Jaeger, in *Protein Eng.*, Springer, **2018**, pp. 209-231.
[81] U. Schwaneberg, C. Schmidt-Dannert, J. Schmitt, R. D. Schmid, *Anal. Biochem.* **1999**, *269*, 359-366.
[82] V. Ullrich, P. Weber, *Hoppe Seylers Z Physiol Chem.* **1972**, *353*, 1171-1177.
[83] L. K. Morlock, D. Böttcher, U. T. Bornscheuer, *Appl. Microbiol. Biotechnol.* **2018**, *102*, 985-994.
[84] A. C. Ward, N. Bora, *Curr. Opin. Microbiol.* **2006**, *9*, 279-286.
[85] a) C. V. Dilip, S. S. Mulaje, R. Y. Mohalkar, *Int. J. Pharm. Sci. Rev. Res.* **2013**, *4*, 1730; b) D. Tischler, W. J. H. Van Berkel, M. W. Fraaije, *Front. Microbiol.* **2019**, *10*, 800.
[86] a) M. Ventura, C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater, D. van Sinderen, *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 495-548; b) S. B. Zotchev, *J. Biotechnol.* **2012**, *158*, 168-175.

[87] M. Taylor, D. C. Lamb, R. Cannell, M. Dawson, S. L. Kelly, *Biochem. Biophys. Res. Commun.* **1999**, *263*, 838-842.
[88] a) J. D. Rudolf, C.-Y. Chang, M. Ma, B. Shen, *Nat. Prod. Rep.* **2017**, *34*, 1141-1172; b) A. Greule, J. E. Stok, J. J. De Voss, M. J. Cryle, *Nat. Prod. Rep.* **2018**, *35*, 757-791.
[89] Y. Xue, D. Wilson, L. Zhao, H.-w. Liu, D. H. Sherman, *Chem. Biol.* **1998**, *5*, 661-667.
[90] a) D. Stassi, S. Donadio, M. J. Staver, L. Katz, *J. Bacteriol.* **1993**, *175*, 182-189; b) M. C. Moncrieffe, M.-J. Fernandez, D. Spiteller, H. Matsumura, N. J. Gay, B. F. Luisi, P. F. Leadlay, *J. Mol. Biol.* **2012**, *415*, 92-101; c) J. F. Andersen, K. Tatsuta, H. Gunji, T. Ishiyama, C. R. Hutchinson, *Biochemistry* **1993**, *32*, 1905-1913.
[91] B. Zhao, X. Lin, L. Lei, D. C. Lamb, S. L. Kelly, M. R. Waterman, D. E. Cane, *J. Biol. Chem.* **2008**, *283*, 8183-8189.
[92] a) F. G. Healy, S. B. Krasnoff, M. Wach, D. M. Gibson, R. Loria, *J. Bacteriol.* **2002**, *184*, 2019-2029; b) S. M. Barry, J. A. Kers, E. G. Johnson, L. Song, P. R. Aston, B. Patel, S. B. Krasnoff, B. R. Crane, D. M. Gibson, R. Loria, *Nat. Chem. Biol.* **2012**, *8*, 814.
[93] a) A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 12289-12298; b) A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 11016-11017.
[94] H. Onaka, S. Asamizu, Y. Igarashi, R. Yoshida, T. Furumai, *Biosci. Biotechnol. Biochem.* **2005**, *69*, 1753-1759.
[95] M. Lombard, I. Salard, M.-A. Sari, D. Mansuy, D. Buisson, *Arch. Biochem. Biophys.* **2011**, *508*, 54-63.
[96] H. Sugimoto, R. Shinkyo, K. Hayashi, S. Yoneda, M. Yamada, M. Kamakura, S.-i. Ikushiro, Y. Shiro, T. Sakaki, *Biochemistry* **2008**, *47*, 4017-4027.
[97] Y. Fujii, H. Kabumoto, K. Nishimura, T. Fujii, S. Yanai, K. Takeda, N. Tamura, A. Arisawa, T. Tamura, *Biochem. Biophys. Res. Commun.* **2009**, *385*, 170-175.
[98] B. Ma, Q. Wang, H. Ikeda, C. Zhang, L.-H. Xu, *Appl. Environ. Microbiol.* **2019**, *85*.
[99] B. Dangi, C. W. Lee, K. H. Kim, S. H. Park, E. J. Yu, C. S. Jeong, H. Park, J. H. Lee, T. J. Oh, *FEBS J.* **2019**, *286*, 1683-1699.
[100] H. Agematu, N. Matsumoto, Y. Fujii, H. Kabumoto, S. Doi, K. Machida, J. Ishikawa, A. Arisawa, *Biosci. Biotechnol. Biochem.* **2006**, *70*, 307-311.
[101] T. Makino, Y. Katsuyama, T. Otomatsu, N. Misawa, Y. Ohnishi, *Appl. Environ. Microbiol.* **2014**, *80*, 1371-1379.
[102] P. Bracco, D. B. Janssen, A. Schallmey, *Microb. Cell Fact.* **2013**, *12*, 95.
[103] B. Dangi, K. H. Kim, S. H. Kang, T. J. Oh, *ChemBioChem* **2018**, *19*, 1066-1077.
[104] D. C. Lamb, M. R. Waterman, B. Zhao, *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 1279-1294.
[105] C. A. Martinez, S. G. Rupashinghe, *Curr. Top. Med. Chem.* **2013**, *13*, 1470-1490.
[106] a) M. Nett, H. Ikeda, B. S. Moore, *Nat. Prod. Rep.* **2009**, *26*, 1362-1384; b) K.-S. Hwang, H. U. Kim, P. Charusanti, B. Ø. Palsson, S. Y. Lee, *Biotechnol. Adv.* **2014**, *32*, 255-268.
[107] a) S. C. Dodani, J. K. B. Cahn, T. Heinisch, S. Brinkmann-Chen, J. A. McIntosh, F. H. Arnold, *ChemBioChem* **2014**, *15*, 2259-2267; b) S. C. Dodani, G. Kiss, J. K. B. Cahn, Y. Su, V. S. Pande, F. H. Arnold, *Nat. Chem.* **2016**, *8*, 419.
[108] a) A. Worsch, F. K. Eggimann, M. Girhard, C. J. von Bühler, F. Tieves, R. Czaja, A. Vogel, C. Grumaz, K. Sohn, S. Lütz, M. Kittelmann, V. B. Urlacher, *Biotechnol. Bioeng.* **2018**, *115*, 2156-2166; b) G. Di Nardo, G. Gilardi, *Trends Biochem. Sci.* **2020**.
[109] S. C. Moody, E. J. Loveridge, *J. Appl. Microbiol.* **2014**, *117*, 1549-1563.
[110] C. D. Murphy, *Biotechnol. Lett.* **2015**, *37*, 19-28.

[111] D. C. Lamb, H. Ikeda, D. R. Nelson, J. Ishikawa, T. Skaug, C. Jackson, S. Omura, M. R. Waterman, S. L. Kelly, *Biochem. Biophys. Res. Commun.* **2003**, *307*, 610-619.
[112] D. C. Lamb, T. Skaug, H.-L. Song, C. J. Jackson, L. M. Podust, M. R. Waterman, D. B. Kell, D. E. Kelly, S. L. Kelly, *J. Biol. Chem.* **2002**, *277*, 24000-24005.
[113] Z.-Z. Li, X.-F. Li, W. Yang, X. Dong, J. Yu, S.-L. Zhu, M. Li, L. Xie, W.-Y. Tong, *BMC Genomics* **2013**, *14*, 130.
[114] L. M. Senate, M. P. Tjatji, K. Pillay, W. Chen, N. M. Zondo, P. R. Syed, F. C. Mnguni, Z. E. Chiliza, H. D. Bamal, R. Karpoormath, *Sci. Rep.* **2019**, *9*, 1-12.
[115] K. Machida, A. Arisawa, S. Takeda, T. Tsuchida, Y. Aritoku, M. Yoshida, H. Ikeda, *Biosci. Biotechnol. Biochem.* **2008**, *72*, 2946-2952.
[116] C. Grumaz, Y. Vainshtein, P. Kirstahler, S. Luetz, M. Kittelmann, K. Schroer, F. K. Eggimann, R. Czaja, A. Vogel, T. Hilberath, A. Worsch, M. Girhard, V. B. Urlacher, M. Sandberg, K. Sohn, *Genome Announc.* **2017**, *5*, e00532-00517.
[117] D. Sirim, F. Wagner, A. Lisitsa, J. Pleiss, *BMC Biochem.* **2009**, *10*, 27.
[118] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81-W87.
[119] P. J. Bakkes, S. Biemann, A. Bokel, M. Eickholt, M. Girhard, V. B. Urlacher, *Sci. Rep.* **2015**, *5*, 12158.
[120] P. Le-Huu, T. Heidt, B. Claasen, S. Laschat, V. B. Urlacher, *ACS Catal.* **2015**, *5*, 1772-1780.
[121] A. Rühlmann, G. Groth, V. B. Urlacher, *ChemBioChem* **2018**, *19*, 478-485.
[122] S. Eiben, L. Kaysser, S. Maurer, K. Kühnel, V. B. Urlacher, R. D. Schmid, *J. Biotechnol.* **2006**, *124*, 662-669.
[123] a) J. K. Capyk, R. Kalscheuer, G. R. Stewart, J. Liu, H. Kwon, R. Zhao, S. Okamoto, W. R. Jacobs, L. D. Eltis, W. W. Mohn, *J. Biol. Chem.* **2009**, *284*, 35534-35542; b) K. Z. Rosłoniec, M. H. Wilbrink, J. K. Capyk, W. W. Mohn, M. Ostendorf, R. Van Der Geize, L. Dijkhuizen, L. D. Eltis, *Mol. Microbiol.* **2009**, *74*, 1031-1043; c) A. Schallmey, G. Den Besten, I. G. P. Teune, R. F. Kembaren, D. B. Janssen, *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1475-1485.
[124] a) S. A. Child, V. P. Rossi, S. G. Bell, *Biochim. Biophys. Acta-General Subjects* **2019**, *1863*, 408-417; b) S. Bhattarai, K. Liou, T.-J. Oh, *Arch. Biochem. Biophys.* **2013**, *539*, 63-69.
[125] M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, *Nucleic Acids Res.* **2011**, *39*, W339-W346.
[126] S. Wagner, M. M. Klepsch, S. Schlegel, A. Appel, R. Draheim, M. Tarry, M. Högbom, K. J. Van Wijk, D. J. Slotboom, J. O. Persson, *Proc. Natl. Acad. Sci. U S A.* **2008**, *105*, 14371-14376.
[127] B. Miroux, J. E. Walker, *J. Mol. Biol.* **1996**, *260*, 289-298.
[128] S. C. Gay, A. G. Roberts, J. R. Halpert, *Future Med. Chem.* **2010**, *2*, 1451-1468.
[129] K. J. McLean, D. Leys, A. W. Munro, in *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Ed.: P. R. Ortiz de Montellano), Springer International Publishing, Cham, **2015**, pp. 261-407.
[130] K. Schroer, M. Kittelmann, S. Lütz, *Biotechnol. Bioeng.* **2010**, *106*, 699-706.
[131] L. A. Merson-Davies, E. Cundiiffe, *Mol. Microbiol.* **1994**, *13*, 349-355.
[132] H.-C. Nguyen, E. Darbon, R. Thai, J.-L. Pernodet, S. Lautru, *Antimicrob. Agents Chemother.* **2013**, *57*, 3836-3842.
[133] L. Cong, W. Piepersberg, *Acta Biochim. Biophys. Sin. (Shanghai)* **2007**, *39*, 187-193.

[134] a) G. N. Kumar, A. D. Rodrigues, A. M. Buko, J. F. Denissen, *J. Pharmacol. Exp. Ther.* **1996**, *277*, 423-431; b) J. F. Denissen, B. A. Grabowski, M. K. Johnson, A. M. Buko, D. J. Kempf, S. B. Thomas, B. W. Surber, *Drug Metab. Dispos.* **1997**, *25*, 489-501.
[135] a) F. M. Kiss, D. Schmitz, J. Zapp, T. K. F. Dier, D. A. Volmer, R. Bernhardt, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 8495-8514; b) G.-Y. Lee, D.-H. Kim, D. Kim, T. Ahn, C.-H. Yun, *Arch. Pharm. Res.* **2015**, *38*, 98-107.
[136] I. K. Jóźwik, F. M. Kiss, Ł. Gricman, A. Abdulmughni, E. Brill, J. Zapp, J. Pleiss, R. Bernhardt, A. M. W. H. Thunnissen, *FEBS J.* **2016**, *283*, 4128-4148.
[137] C. M. Pratt, J. Mason, T. Russell, R. Reynolds, R. Ahlbrandt, *Am. J. Cardiol.* **1999**, *83*, 1451-1454.
[138] E. Jacqz-Aigrain, C. Funck-Brentano, T. Cresteil, *Pharmacogenetics* **1993**, *3*, 197-204.
[139] a) A. F. Hofmann, L. R. Hagey, *Cell. Mol. Life Sci.* **2008**, *65*, 2461-2483; b) D. W. Russell, *Annu. Rev. Biochem.* **2003**, *72*, 137-174.
[140] J. J. Eloranta, G. A. Kullak-Ublick, *Arch. Biochem. Biophys.* **2005**, *433*, 397-412.
[141] C. Thomas, R. Pellicciari, M. Pruzanski, J. Auwerx, K. Schoonjans, *Nat. Rev. Drug Discov.* **2008**, *7*, 678-693.
[142] N. Tanaka, T. Matsubara, K. W. Krausz, A. D. Patterson, F. J. Gonzalez, *Hepatology* **2012**, *56*, 118-129.
[143] U. Özcan, E. Yilmaz, L. Özcan, M. Furuhashi, E. Vaillancourt, R. O. Smith, C. Z. Görgün, G. S. Hotamisligil, *Science* **2006**, *313*, 1137-1140.
[144] F. Berr, E. Pratschke, S. Fischer, G. Paumgartner, *J. Clin. Invest.* **1992**, *90*, 859-868.
[145] M. Mikov, M. Đanić, N. Pavlović, B. Stanimirov, S. Goločorbin-Kon, K. Stankov, H. Al-Salami, *Eur. J. Drug Metab. Pharmacokinet.* **2018**, *43*, 269-280.
[146] D. Q. H. Wang, S. Tazuma, *J. Lipid Res.* **2002**, *43*, 1960-1968.
[147] a) F. Tonin, I. W. C. E. Arends, *Beilstein J. Org. Chem.* **2018**, *14*, 470-483; b) S. Takahashi, T. Fukami, Y. Masuo, C. N. Brocker, C. Xie, K. W. Krausz, C. R. Wolf, C. J. Henderson, F. J. Gonzalez, *J. Lipid Res.* **2016**, *57*, 2130-2137.
[148] N. D. Fessner, *ChemCatChem* **2019**, *11*, 2226-2242.
[149] a) S. Kranz-Finger, O. Mahmoud, E. Ricklefs, N. Ditz, P. J. Bakkes, V. B. Urlacher, *Biochim. Biophys. Acta, Proteins Proteomics* **2018**, *1866*, 2-10; b) T. Nagao, T. Mitamura, X. H. Wang, S. Negoro, T. Yomo, I. Urabe, H. Okada, *J. Bacteriol.* **1992**, *174*, 5013-5020.
[150] a) L. McIver, C. Leadbeater, D. J. Campopiano, R. L. Baxter, S. N. Daff, S. K. Chapman, A. W. Munro, *Eur. J. Biochem.* **1998**, *257*, 577-585; b) Z.-Q. Wang, R. J. Lawson, M. R. Buddha, C.-C. Wei, B. R. Crane, A. W. Munro, D. J. Stuehr, *J. Biol. Chem.* **2007**, *282*, 2196-2202.
[151] J. B. Schenkman, S. G. Sligar, D. L. Cinti, *Pharmacol. Ther.* **1981**, *12*, 43-71.
[152] D. V. Waterhous, S. Barnes, D. D. Muccio, *J. Lipid Res.* **1985**, *26*, 1068-1078.
[153] M. Girhard, K. Machida, M. Itoh, R. D. Schmid, A. Arisawa, V. B. Urlacher, *Microb. Cell Fact.* **2009**, *8*, 36.
[154] I. A. MacDonald, C. N. Williams, D. E. Mahony, *Biochim. Biophys. Acta-Enzymology* **1973**, *309*, 243-253.
[155] L. Liu, M. Braun, G. Gebhardt, D. Weuster-Botz, R. Gross, R. D. Schmid, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 633-639.
[156] T. Hisadome, T. Nakama, H. Itoh, T. Furusawa, *Gastroenterol. Jpn.* **1980**, *15*, 257-263.
[157] a) M. N. Podgorski, T. Coleman, R. R. Chao, J. J. De Voss, J. B. Bruning, S. G. Bell, *J. Inorg. Biochem.* **2020**, *203*, 110913; b) P. Le-Huu, D. Rekow, C. Krüger, A. Bokel, T.

Heidt, S. Schaubach, B. Claasen, S. Hölzel, W. Frey, S. Laschat, *Chem. Eur. J.* **2018**, *24*, 12010-12021.
[158] S. Grobe, A. Wszołek, H. Brundiek, M. Fekete, U. T. Bornscheuer, *Biotechnol. Lett.* **2020**, 1-6.
[159] a) H. M. Girvan, A. W. Munro, *Curr. Opin. Chem. Biol.* **2016**, *31*, 136-145; b) S. L. Kelly, D. E. Kelly, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20120476.
[160] a) M. Schrewe, M. K. Julsing, B. Buehler, A. Schmid, *Chem. Soc. Rev.* **2013**, *42*, 6346-6377; b) C. Willrodt, R. Karande, A. Schmid, M. K. Julsing, *Curr. Opin. Biotechnol.* **2015**, *35*, 52-62.
[161] S. P. Hanlon, T. Friedberg, C. R. Wolf, O. Ghisalba, M. Kittelmann, in *Modern Biooxidation (eds.: R. D. Schmid, V. B. Urlacher)*, Wiley-VCH, **2007**, pp. 233-252.
[162] R. Karande, D. Salamanca, A. Schmid, K. Buehler, *Biotechnol. Bioeng.* **2018**, *115*, 312-320.
[163] M. T. Lundemo, S. Notonier, G. Striedner, B. Hauer, J. M. Woodley, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 1197-1208.
[164] A. Jakoblinnert, D. Rother, *Green Chem.* **2014**, *16*, 3472-3482.
[165] F. M. Kiss, M. T. Lundemo, J. Zapp, J. M. Woodley, R. Bernhardt, *Microb. Cell Fact.* **2015**, *14*, 28.
[166] M. Geier, C. Schmid, A. Glieder, *Chim. Oggi* **2013**, *31*, 24-27.
[167] H. Souzu, *Biochim. Biophys. Acta* **1980**, *603*, 13-26.
[168] a) T. Loftsson, M. E. Brewster, *J. Pharm. Sci.* **1996**, *85*, 1017-1025; b) A. Rühlmann, D. Antovic, T. J. J. Müller, V. B. Urlacher, *Adv. Synth. Catal.* **2017**, *359*, 984-994.
[169] V. V. Fokina, A. V. Karpov, I. A. Sidorov, V. A. Andrjushina, A. Y. Arinbasarova, *Appl. Microbiol. Biotechnol.* **1997**, *47*, 645-649.
[170] K. Abokitse, W. Hummel, *Appl. Microbiol. Biotechnol.* **2003**, *62*, 380-386.
[171] W. Kroutil, H. Mang, K. Edegger, K. Faber, *Curr. Opin. Chem. Biol.* **2004**, *8*, 120-126.
[172] M. T. Lundemo, S. Notonier, G. Striedner, B. Hauer, J. M. Woodley, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 1197-1208.
[173] K. Schroer, E. Tacha, S. Lütz, *Org. Process Res. Dev.* **2007**, *11*, 836-841.
[174] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat. Methods* **2009**, *6*, 343-345.
[175] J. Reinen, G. Postma, C. Tump, T. Bloemberg, J. Engel, N. P. E. Vermeulen, J. N. M. Commandeur, M. Honing, *Anal. Bioanal. Chem.* **2016**, *408*, 1425-1443.
[176] H. Veith, N. Southall, R. Huang, T. James, D. Fayne, N. Artemenko, M. Shen, J. Inglese, C. P. Austin, D. G. Lloyd, D. S. Auld, *Nat. Biotechnol.* **2009**, *27*, 1050-1055.
[177] F. P. Guengerich, *Chem.-Biol. Interact.* **1997**, *106*, 161-182.
[178] P. Caffrey, S. Lynch, E. Flood, S. Finnan, M. Oliynyk, *Chem. Biol.* **2001**, *8*, 713-723.
[179] R. Fouces, E. Mellado, B. Díez, J. L. Barredo, *Microbiology* **1999**, *145*, 855-868.
[180] a) N. Serizawa, T. Matsuoka, *Biochim. Biophys. Acta, Lipids Lipid Metab.* **1991**, *1084*, 35-40; b) I. Watanabe, F. Nara, N. Serizawa, *Gene* **1995**, *163*, 81-85; c) T. Matsuoka, S. Miyakoshi, in *US-Patent 5179013*, **1993**.
[181] M. M. Joshi, H. M. Brown, J. A. Romesser, *Weed Sci.* **1985**, *33*, 888-893.
[182] H. A. Hussain, J. M. Ward, *Appl. Environ. Microbiol.* **2003**, *69*, 373-382.
[183] A. Çelik, S. L. Flitsch, N. J. Turner, *Org. Biomol. Chem.* **2005**, *3*, 2930-2934.
[184] N. Sawada, T. Sakaki, S. Yoneda, T. Kusudo, R. Shinkyo, M. Ohta, K. Inouye, *Biochem. Biophys. Res. Commun.* **2004**, *320*, 156-164.
[185] S. Janocha, J. Zapp, M. Hutter, M. Kleser, J. Bohlmann, R. Bernhardt, *ChemBioChem* **2013**, *14*, 467-473.

[186] M. Kleser, F. Hannemann, M. Hutter, J. Zapp, R. Bernhardt, *J. Biotechnol.* **2012**, *157*, 405-412.
[187] A. Seifert, M. Antonovici, B. Hauer, J. Pleiss, *ChemBioChem* **2011**, *12*, 1346-1351.
[188] D. Decembrino, Master thesis, Heinrich-Heine-University Düsseldorf and Universita'Degli Studi Di Torino. **2017**.
[189] J. A. Romesser, D. P. O'Keefe, *Biochem. Biophys. Res. Commun.* **1986**, *140*, 650-659.
[190] P. A. Harder, D. P. O'Keefe, J. A. Romesser, K. J. Leto, C. A. Omer, *Mol. Gen. Genet.* **1991**, *227*, 238-244.
[191] I. Molnár, V. Jungmann, J. Stege, A. Trefzer, J. P. Pachlatko, *Biochem. Soc. Trans.* **2006**, *34*, 1236-1240.
[192] S. Han, T.-V. Pham, J.-H. Kim, Y.-R. Lim, H.-G. Park, G.-S. Cha, C.-H. Yun, Y.-J. Chun, L.-W. Kang, D. Kim, *Arch. Biochem. Biophys.* **2015**, *575*, 1-7.
[193] J. R. Cupp-Vickery, T. L. Poulos, *Nat. Struct. Biol.* **1995**, *2*, 144-153.
[194] M. B. Buergler, A. Dennig, B. Nidetzky, *Biotechnol. Bioeng.* **2020**, *117*, 2377-2388.
[195] J. Wachtmeister, A. Jakoblinnert, J. Kulig, H. Offermann, D. Rother, *ChemCatChem* **2014**, *6*, 1051-1058.
[196] Z. Maugeri, D. Rother, *Adv. Synth. Catal.* **2016**, *358*, 2745-2750.
[197] S. Semsary, I. Crnovčić, R. Driller, J. Vater, B. Loll, U. Keller, *ChemBioChem* **2018**, *19*, 706-715.

5. Supplementary data

Amino acid	3 Letter code	1 Letter code	Amino acid	3 Letter code	1 Letter code
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

5.1 DNA sequences

All DNA sequences of the investigated P450s from *S. platensis* (SPL-number) and *P. autotrophica* (AAU-number) are listed below. The Gen bank number (OSY-number) and the P450 name (CYP) are given in the brackets.

SPL_00625 (OSY47991, CYP105D)

GTGACCGAAGCCATCCCCTACTTTCAGAACCGCACCTGTCCCTACCACCCGCCCGCCGCCTATCAGCCATTGCGCGAGGCC
GGCCCGCTGAGCCATGTCACGTTCTACGACGGCCGGAAGGTGTGGGCGGTCACCGGCCACCCCGAGGCACGGGCGCTGC
TGACCGACCAGCGGCTCTCCGCCGACCGGCAGAACCCGGCCTTCCCGGTCCCGTTCGAACGCTTCGCGGCCATCCGCCGG
GTCCGGACACCGCTGATCGGGGTCGACGACCCGGAGCACAACACCCAGCGCCGGATGCTGATCCCCAGCTTCAGCCTCAA
GCGGACCGCCGCACTGCGGCCGGAGATCCAGCGGATCGTCGACGGGCTGCTGGACCGGATGCTGGATCAGGGCCCGCCC
GCCGAGCTGGTCTCCGCGTTCGCGCTGCCGGTCCCGTCGATGGTGATCTGCGCACTGCTCGGCGTCCCGTACGCCGACCAC
GAGTTCTTCGAGGAGGAGTCCCGCCGTATCCTGCGCGGCCGGTCGGCCGAGGAGGCGGAGGACGCCCGGCTGAAGCTG
GAGGAGTACTTCACCGGGCTGATCGCCGCCAAGGAGAAGAAGCCGGGCGACGGGCTGCTGGACGAGCTGATCGAGGAC
CGGCTGCGGACCGGCGCGCTCACCCGCGACGAGCTGGTCCGGCTCGCCATGATCCTGCTGGTGGCCGGCCATGAGACCAC
CGCCAACATGATCTCGCTCGGCACCTTCACCCTGCTGGAGCACCCCGAGCAGCTGGCGCAGCTCAAGGCCGACGAGGGTC
TGATGCCGGCTGCCGTCGAGGAGCTGCTGCGGTTCCTGTCCATCGCGGACGGTCTGTTGCGGGTGGCGACGGAGGACAT
CGAGATCGGCGGTCAGGTGATCCGGGCCGACGACGCGGTCCTGTTTCCCGCCTCACTGATCAACCGGGACGAGGCCGCCT
ATCCGGCACCGGACGAGCTGGACCTCGGCCGTTCGGCCCGCCATCACGTGGCGTTCGGCTTCGGGATCCACCAGTGCCTG
GGGCAGAATCTGGCGCGCGCGGAGTTGGAGATCGCGCTGCGCTCACTGTTCAAAAGGATCCCGCAGCTGCGGCTCGCCG
TGCCGGCCGCCGAGATTCCGTTCAAGGCCGGCGACACCCTGCAAGGCATGATCGAACTGCCGCTGGCCTGGTAG

SPL_01100 (OSY47382, CYP107)

ATGGCCAGCACTGAAGAACCAGCAAGCCCGACCCCGCCTTCGTACCCTTTCCGCCAGGCCAGGGGAGTGGGCATCGACCC
GGAGTTCGCGCGCTTGCGGCACGATGCGCCGCTCGCCCGTGTCACCATGCCCTACGGCGGCCAGGCGTGGCTGGTGACGC
GGTACGAGGACGTGCGTACAGTCCTCGGCGACCGCAGATTCAGTCGGGCGGCCACGCTCGGGCGGGACGTACCGCGGCT
GGTCCCGCTCGTCCAGCAGGTGTCGAGCATTCTGACGCTTGACCCACCGGACCACACCCGTCACCGCAGGCTCGTTTCGCG
TGCATTCACCCCTCGCCGTATGGAGGAGCTACGTCCAAAGGTGGAGCAGTTCGTCGATGAGCTCATCGACGAACTGGTCG
GACACGGGCAGCCCGCTGACATGGTGGCTCACGTCACCCGACCGCTTCCTGTCATGGTGATCTGCGAGTTGCTGGGCGTT
CCCTTTGGCGAGCGGCACCTCTTTTACACATGGTCAGAAGCGATGCGGTCCACCGGCGCCGATGCGGTCGCCAAGATGAA
GCAGGCAGGACAGCTCCCCTGGGACTACCTGGCCGACAGGATCTCCGTCGAACGGGAGAATCCGTCCGACAACTTGCTCG
GAACGCTGGTGCGAGCGCGTGATGAAGAGGAGCGGCTCAGCGAGCAGGAACTGGTGAGCTTCGCGGTAACTCTCCTGTT
GGCCGGCCACGAGACGACCACCGACGAGCTCGGCAATTTCCTTTACACGCTCCTCGTGAACCCTGCACACCTTGAGCAACT
GCGCGCTCAGCCCGACATGTTGGAGTCGGCCATCGAAGAGTTGCTGCGCTTCGTGCCGATCGGAACCCTCTCCGGCTTCAC
CCGGATCGCCACTGAGGACGTCCGTCTGAGCGGTGGACTGGTACGAGCCGGCGACTCGGTCGTGGTGCAGGCGGATTCG
GCCAACCGCGACGAATCAGTCTTCGCGGACCCTGACGAGCTGGACTTCCAGCGGGAGCCCAATCGGCACCTGGCCTTCGG
CCACGGGCCGCATCACTGCCTGGGAGCGCAGCTCGCCAGGATCGAACTACGGGCCGCGATAGGCACCTTGCTGGCCCGCC
TGCCCGCCCTCGCGTTGGCGGTGTCGGCGCACGAAGTGTCCTGGAAGCTCGGGCGATCGGCGCTCGGGCCGCAGGCACT
GCCCGTCAGGTGGTAA

SPL_01126 (OSY47408, CYP183)

ATGTGGCCGGGAACGACAAGGACCGGCAGGAGCAGAAGCGGTGCACGGCTCACCGGCGGCCCGGGAGACACCGTCGCC
GATGCCCCCACCGGAGCCCCCGTCGCCCCCGGCCGCGTGCCGCTGGCCGGGCACGCCCCCTGGATCCACCGTGATGCGCT
GCGCTTCGTCCGGGAGCTGCGGCGCCACGGGCCGGTGACGAGGATCTACATCGGCCCACGGCAGGTACACACCGTCAAC
TCGCCCGAGATCGTCCGGGAGTTGCTGACCGTGCAGGCGCGCGCCTTCGACAAGGGCGCGATGTTCGACGCCCTGCGCG
AGCCGCTCGGCGACGGGCTGATCACCGCCGCCGGCGACCGCCACCTCCACCACCGCCGGCTGATGCAACCCGCCTTCCAC
CACGAACGGATCGCCGGCTACGCCCGCATCATGTCCGAAAGGTCGCTGGCACGGGCCGCTGTCTGGGAGCCCGGTACCAC
TCTCGACCTGGTACCGGAGATCCACCGCCTGACGCTCGACATCCTGCTGCGCACCCTCTTCGCGGATCCCCAGGACCCTGA
ACTGCGCACCGCCGTCCGGGACTGGCTGTCCGTCAAGTACCACGCGATGCGCCTTGCCCTCTCACCGCTGCACGCCTGGGC
GGAACGTCTCCCGCTGCCGGTCGGCTGGCGGCCGCCGCAGCCCGGCCCGCTGCGCCGGCTGGTGGCCGTCCAGCACCGC

ATCGTCGAGACCTACCGCGCCGATGGCCGGGACCGCGGCGATCTGCTCTCGATGCTCCTGGCCACGGGCGGCCCGGACG
GTTCCCTGACCGACGCCGAAGTGACCGACGAACTGATCACCCTGTTCCTGGCCGGTACCGGCACCGTCAGTGCGTCCCTGG
CCTGGGCGCTGTACGAGATCTCCCGCCGGCCCGACGTGCAGCGGCGGATCCACGACGAACTCGACACCGTCCTCGCCGGA
CGGCCTCCCGGCATCGAGGATCTTCCCGCACTGGTCCACACACGAAGGGTGCTGACCGAAGTGCTGCGGCTGCATCCACC
GTCGTGGCTGCTGATGCGCCGTGCGGTGCGGCCGGTCACCCTCGGCGGCGTGTCGCTCTCCCCCGGCGCCGAGGTGTTCT
TCAGCCCGTACGCCCTGCACCACGATCCGCAGCTGTACGAGAACCCGGACGACTTCGAGCCTGAGCGCTGGTCGCACGCC
GACGCCGCAAAGTCTTCGCGTCACACCTATCTGCCCTTCGGGGCCGGCAGCCGACTGTGCATCGGGGAAGACTTCGCCTG
GACCGAACTGATGCTGGCACTGGTCGCGTTCACCGCTCACCGGCGCCTGACACCTGCCGGCTCCACCCCCGTGCGGGCGC
TGGTCGGCACCGTCCTGCGCCCGGACCGGCTGCCGCTCACGGCGCATGCCCGCCCGTCCTGA

SPL_01167 (OSY47449, CYP208-like)

ATGCGTATCCCCGGCCCCGAGCACCAGGAGGACGGCGGGCTGAGCGCCGTTGCGGCCGCGGGTGGGCTGCACCGGTACC
AGTTACGCCTGCACGAGGCGTACGGGCCCGTCGTGCGCTTCCAACTGCCAGGCGCGGCAACGGCCGTGTCGGTCGCCGAT
CCCGTGCTTCTGGAGGCCACGGCGCACCTCGACACGCGGCCCGAGAAGCTGTTCGCGTTCCTGGACCCGCTGTGTGAGGC
AGGAAACCTGCAAGTGCTACCCGCCGACGAGCACACCCCCTGGCGCCGACTACTGCTCTCGGTGCTCGCCGGGCGGCCGG
CTCACGAGCGGCACTTCGCGCGCTTCACCGAGCTGGCGCGGGAGCTCGCGGACCGCTGGGCGGGGCAGACCGACGGCG
AACCCGTTGAGCTGCAACAGGAGTTGACCGCGCTGTCGCTCCGTATGATCTGCGGGTTCGCGCTCGGCGGCGCGGAAACC
GACACGGACCGTGTCGTCGCCGCGTTCGAGGAGGTGCTCACCGAGCACCTCGGGCGGCTCTACCAGGTGCCGCGGGCCG
ACCCGCGCTCGGCGGAGCGGGCCGAGGTGGCTCTTGCCTATCTGCGCGGGACGGTCGACCGCGTGGTGGCGGCGCACCG
CTCGGGCGGCCGCACCGACCGCAGCGACCTGATCGGGGCGCTCGTCGAGGCCGGTCAGAGCCCGGCGCGGGTCCGGGA
CACCGTGATGGTGACGCTGCTCGCCGCCCACCACACCACCGGCGTCGCCATCTCGTGGACGCTGCACGTGCTGGGCCGCC
ATCCCGACGCGGCCGATCGCGTCGCCGCCGAACTGGACCAGGTGCTGGGCGAGCGCGCCGCGCCCGATTACGGTGATCT
GCGGCGGCTCACATATTTGGACATGGTGCTCAAGGAGTCGATGCGCCTGTTCCCGCCCGGCCCGTACGGCGCACGCGAGA
CGACCGAGGACCTGGTTCTGGGCCCGTACGAGATCCCGGCCGGGACGACGGTGTTCTATCCGTTCTGGGCGGTGCACCTG
AACCCGGAACACTGGCCCGAGCCGGAGCGGTTCGACCCCGAACGGTTCACGCCGCAGGAGGTGGCCAAGCGGCCGAGG
CTCGCGTACGTCCCCTTCGGGATCGGGCCACGCAACTGCGAGGGCGCCGCTCTGGCCACGGTCGAGGCCCAGCTCGTCCT
GGCCGTCCTCCTCAAGCGCCTGCGCTTTCGGCCCGCGCCGGGGCATGAAGTGACGCCGGTGGAACGGTTCGTGCTCTGGG
CACGGGACGGGATCCGGATGATGGTGCGGCCACGGGACCTCTCTGGCGCCTGA

SPL_01317 (OSY47218, CYP113)

ATGACTGTGCGCGACGAGATCGCCGTCCCGGATGCCGACGGCAGCAACGCGTTGTTCGACTGGCTCGGCAGGATGCGGT
CCGAGCATCCCGTGTGGCAGGAGGGGGACGGGCCGTACCACGTCTTCCGCTACGAGGACGTGCAGCAGGTCATCTCCGA
CCCCAAGACGTTCTCCAACGACTCCAGCCGGGTCATGCCCCATCTGAAGCCGCTGACCGAGGGCAACATCAACTCGATGG
ACCCGCCGGACCACGGCAAGCTGCGCCGGCTGGTCAACCAGGCCTTCACCCCGAAGACGGTGGCCGGGCTCGAACCCCG
CATCGCGGCGGTGACGAACGAGCTGCTGGCCGCCGTCGACGCGGAGCGGTTCGACCTGGTGGACGTGCTGACGTATCCG
CTGCCGGTGATCGTCATCAGTGAGCTGCTGGGGGTGCCCGCCTCGGACCGTGATCTGTTCCGGGTGTGGGCGGACCGGTT
CATCGCGCTCGGCGACCAGCCGATCGCGCCGGAGGACTTCGTCGCCACGTTCCAGGCCGCGACCCGCGAGATGGACGAG
TATCTGCTGGGTCATTGCCGGCAGCGCCGGACGGACCCGAAGGACGACCTGATCAGCCGGCTGGCCACCGCCGAGATCG
ACGGGGAACGGCTCACCGACGACGAAGTCGTGAAATTCACCGGAATCCTCTTTCTGACCGGGCACCTCACGACGACACTG
CTGATCGGCAATGCCATGCAGTGTCTCGACAGCCATCCCGACGCCTATGCGGAACTCCGCGCCGACCGCTCACTGATCCCC
TCCGCGATTGAGGAAGTGCTGCGGTTCCGCTCGCCTTTCACCACGGTCAGCCGGGTCACCGTGACCGATGTCGAGGTGGG
CGGCCAGCTCATCCCGGCCGACCGGATGGTGACGCCCTGGGTGATTTCGGCGAACCATGACGAGGACCAGTTCCCGGAC
GCGCACCGCTTCGATATCCGGCGGACGCCCAACCGGCATGTCGCCTTCGGGCGCGGCGGACACTTCTGTGTGGGGGCGCC
GCTGGCCCGGCTGGAGGCGGAAGTGGCGCTCAATGCCCTGCTGGACACCTTCCGTGAACTGCGCGTCGACCACGGCCGG
GATGTCGCCTATCACGCCCGTGGAATGTATGGCGCGAAGAATCTGCCGATGACGGCCCGGCGTTGA

SPL_01896 (OSY46529, CYP105AA)

ATGAGTGAAATGGCCGCCATCCCGCACGGCCTCCCCATGGAGCGCGATGCGGGCCCCTTCGACCCGCCCCGCCAGATCAC
CCGGCTGCGCGACGCACGCCCGGTCAGTCCCCTGGTCTTCCCTGACGGTCATGAGGGCTGGCTCGTCACCGGCTACGACG
CGGTCCGCCAGCTCCTGGCCGACACCCGGTTCAGCTCCCGCCAGGACCTCGGCATCGTCCATGTGCCGTACGAGACCCCG
GGCATACCCGCCGCCACCGAACCGTCCCCGCAGGTGCCGGGCCTGTTCGTCGCCATGGACCCGCCGGACCACACCCGGCT
GCGGCGCAAGCTGACCGGCGCTTTCACCGTCAGACGCATGAAACAGCTCGAAGAGCACATCGTCGAGATCACCGAGCGG
CAACTGGACGAGATGGCGCGTCTGGCCCCGCCGGTCGACCTGGTCAAGGAGTTCGCGCTGCCGGTGCCCTCGCTGGTCAT
CTGCGAACTGCTCGGTGTCCCTTACGAGGACCGGGAGAACTTCCAGGTCAACTCCGCCAAGTTCCTGGTCAAGGACCAGC

CGCTGGACGAGAAGATGGCCGCGTACGGCGCGATGATGACGTACCTCGCCGAACTGGTCATGCGCAAACGGGCCGAACC
CGGCGAAGACATCCTGTCCGACCTGGCCCGCCATGACGACCTCACCATCGAGGAACTGACCGGCATCGCCTTCCTGCTGCT
GCTCGCGGGCCACGAGACCACCGCCAACATGCTGGCCCTGGGCACCTTCGCGCTCCTGGAGCACCCCGAGCAGCTGGCCG
AACTCCGCGCCGCCCCGGACCTGATCCCCGATGCCGTCGAGGAACTCATGCGTTACCTGTCCGTCGCCGACATCTTCTATC
GCTACGCCACGGAGGACATCGAACTCGGCGGGGAAACCATCGGCAAGGGCTCGACCGTCGTCGTGTCGCTGCTGGCCGC
CAACCACGACCCGCAGCGCTTCGACCACCCCGACACCTTGGACATCCACCGCAAGGCCCGCGGACACCTCTCCTTCGGCCA
CGGCATCCACCAGTGCCTGGGCCAGCAACTGGCCCGTATCGAGATGCGCGCCGGATTCGACGGACTGCTGCGCCGCTTCC
CGACCCTCCAGCTCGCCATCCCCGCCGACGAGGTGAAACTCAGGACCGACATGAACATCTACGGCGTCCACGAACTGCCG
GTCACCTGGACGGAAACGACCCGGTAG

SPL_02251 (OSY46283, CYP125)

ATGACCGTATCCGCAGCCGACGCCGCGGACCGGGAGGTGCCGGACGTCTTCGACCCGCGCGGCTACGCCGACGGGCTGC
CGCACGCGGCGTACCGGCTCTTGCGCGACCGGTATCCGGTGGCCTGGCAGGAGGAGCCCGAGGTGCTCGGCTGGCCGGC
CGGGCCGGGCTTCTGGGCGGTCACCCGCCACCAGGACGTCGTACGGGTCCTGAAGGACGCCCGGACCTTCTCCTCGCACC
TGGGCGCCACCCAGATCCGCGATCCCGACCCCGCCGATCTGCCGTTCATCCGCCGCATGATGCTCAATCAGGATCCCCCCG
ACCACGGGCGGCTGCGCCGTGCGGTCAGCCGCGCGTTCACCCCTCGCCGGATCGACCGCTTCGGCACCGTCATACGCGAG
CGGGCCCGTGCGCTGCTGACGGCGGCGGTCGGCACGGCGCGCGCCGGGGACGGGGTGGGTGACCTCGTCCCCGCGGTC
ACCGACGACTTCGCGCTGCTCAACCTCGCGGATCTGCTCGGGGTGCCGGCCGCTGAGCGCGGGCTGCTCCTGGACTGGAC
CCGGCGCGTCATCGGCTACCAGGACCCGGACGAGGCCGGCCCGCCCGCCTCCGGCCCGGACGGGCGGCCCGCCAACCCG
CGCTCCCCGGCGCTGCTCCAGGACATGTTCGGCTTCGCACGGGAACTGGCCGCCCACAAGCGCCGGTGTCCGGGCGACGA
TGTGATGACGGTGCTCGCCTCGGACCCGGAACTGGCCACGCCCGAGCTGGAGATGTTCTTCTTTCTGCTGACCATCGCGGG
CAATGACACCGTGCGCAGTGCGGCCCCCGGCGGACTGCTGGCGCTGGCCCGGCACCCGGACGCGTACCGCCGGCTGCGC
GCGGGCGCCGTGGGGACGGCGACGGCCGTCGAGGAACTGCTGCGCTGGCATCCGCCGGTGCTCAGCTTCCGCCGTACCG
CGGCATACGACACCGAACTGGCCGGCCGCACCCTTCACGCCGGCGACAAGGTGGTGGTCTTCCACGGTTCGGCGAATTAC
GACGAACGGGTCTTCGCCGACCCGCACCGCCTGGACCTGGCCCGCTCCGCCAATCCCCATGTCTCCTTCGGCGACGGGCCG
CATGTCTGCCTGGGCGCCCATCTCGCCCGGCTGCAACTGCGGGTGCTGTACGAGGAGACCTGCCGTCTGCTGCCCGCCCTG
GCGGTCGCCGGTCCGCCGCAGCGGCTGGTATCGAACTTCATTCACGGACTGAAGTCGCTGCCGCTGCGGCTGGTGGACGT
TCCCGGCGGGGGGTAG

SPL_02316 (OSY46348, CYP183-like)

ATGTTTGTTCTGAACTCGCCGGAGGCGGTCCACGACGTTCTCGTGACGCAGAGCAGAAAGTTCAGCAAAGGGTTGCTGTT
CGACATCGCGAGGCCATTTATCGGGAATGGCATAATCACGTCCGAACGAGAGTTTCACAGGCGCCAGCGTCGTGCGCTCC
AGCCTGCTTTCCATCGTGACAGTATCGCCGGATATGTCGGTACGATGATTGATGTCGCTGAGGAGCAGGTGTCGACGTGG
CGCCCGGGCGAGGTGATCGTCATGGATCAGGCGATGCGTCGGCTGATGACGGTCATGCTCGTAGCGACCTTGTTCAGCAC
CGAACGCCCGGACGGCGCCGAGGCTATGGCCGAGCTGGGGGATCGGGTCGCTGAGCACCTGACCACGGTGATGCGTGG
GGTCTTCGTGGGCACTGTCCTGCCCGCGCCGATCGCTTCTTTGCCTGCGCTCGGTCACCGCCGGTATCTGTCCGCGGCCGC
GGCCTTGCGCGACCTCGCGGACACAGCTGTCCGGCAGGCCCGGCAGGATCCCACCGGCCGCGGTGATCTGCTGTCGATCA
TGTTCACCGGCACAGCGGGGAATCCTCAGGGGATGAGTGACGTGGAAGCCCGCGATGAGTTGCTGTCCCTCCTCATGGCC
GGCGCTGAAACCACGTCAACCACCTTGTCCTGGGCTCTTTATGAGCTCGGTCGACGACCGGAGATCACAGATCGAATCAA
AGCCGAGTGCCGATCCCGACCTGTCGAGGTCATTCCCAGTGCAACGACCTTTCCCTTCACCGACCGGTTCCTCCGGGAGGT
TCTTCGATTACATCAGCCCAACTGGCTTCTTATGCGGCGCGCTTTGGAACCGGTGCGTGTCTCCGGAATCGAGCTTCCACCC
GGGGCCGAAATAATCTACAGCGCAGCGACGATGCATCGAGACCCGGCGTACTTCCCGGACCCCTTGCGATTCGACCCGGA
CCGATGGCTTGACCGTTCCCAAAAGGAACTTCCGCCGGGTGCCTACATCCCATTTGCAGTGGGCAACAGGAAGTGTATCG
GGGACTTCTTCTCAATGACCGAAATGCTGGTCGTCCTCCGCTCGATCGTATCGCGGTGGCGACTGCACCCGCTTGAGAAGC
ATCGGGTCCGGCCAGTGGTCCAAGCGCAGATCCGGCCGAATTCGCTGCCCATGCTCGTCTCACCGTGGCCAACCCGGTGA

SPL_02318 (OSY46350, CYP105)

ATGCCCGTTCGCCGCCCAGCGTCCTTCGTGACGTCCGAACGAGATCAAGACTGTCCATTCGACCCGTCACGGACTCTCGTT
CGCATGCGCGAAGAGGATGGTCTGCCGCGTCTGCGTGTATTTCACCCTGTGCGCGGTGAAATCGAAGCCATCGTTGTCAC
GAGATACGGTGATGTCCGGGCTGCCTTCGCGGAAGATCTCCTAGTGACCGGAAACGGCATCGATACAACGATGCCGCGG
ACCTTGTTCAACCAACCAGGTTTTCTACCCGCCTACAGCGGTGCAGAACATCAACGCCTCCGCCGCATGCTTACCGGCTGCT
TCACCGTGCGCCAGGTTGAGCAGCTGCGACCGGCGATCGAAGAAATGGTCTCGGCTCACCTCGACGCCATGGTGGCGCAC
GGGCCCGTGGTCGACCTGGTCGAGGCCTTCGCCCTCCCCATCCCCTCACTGGCGATCTGCGAACTGCTTGAAGTGCCGTAT
GAGACACGATCCATGTTTCAGCGCTGCTCTCAGGTGATCATGGACGGGACGGCCTGCGCCGACCAGTTGGTGGCCGCATC
GGCCGAGTTGAATGCAGCCATGCAAGAACTTGTCACCCGAAATCGTGCCGCGCCCGGGGACGGAGTGTTGGGAAGACTG

GTGCGGCAACACGGTGAGGACCTCACGGATGCTGAACTCCTCGGATTCGGGACCACTCTTCTCGTGGGGGGACACGAGA
CCACAGCGACCATGCTCGCCCTGAGCGTATTGGCGTTGTTGCGGAACCCAAACCAATTGGCTGCTCTCCGAGACGACACCT
CGGTCACGAACACAGCGGTCGAAGAACTTCTTCGGTATCTCGCCATTCCAGCGCCGCTTCTGAGGATGGCTGTGGAAGAC
GTCGAGATCTGCGGGAGCCGGATCGCGGCAGGTGAATACGTGCTGCTCAGCCCATTGACAGCGAACCGTGATCCCGACCT
CGTGCCCGAATGGCCTGACACGCTGGATGTCCATCGCCGTCCGGTAGCACAGATGTCGTTCGGTTTCGGTGTCCACCAGTG
CCTCGGACAACAGTTGGCAAGACTGGAGTTGAAGATCGCTCTGCCGGCGTTGCTCCGACGTTTTCCCACGCTCGAACTCGC
CGTCCCCGATACGGATCTCCGGTTCAGGGAAGGCTCGACGGTATATGGCGTCGAAGCGCTTCCAGTGTCCTGGCGACCGG
GCAACGCTGGAGGCCGAAAGTGA

SPL_02943 (OSY45653, CYP183-like)

ATGCCTATGCACCGCGTCGGCGAGAGCGGTCGGGTCATCGAGTTCGCGCCGAGGATCGACGAACTGTTAAGTCGTTATCG
TGGTGCGTCGCTTTTTCGTCTGGAGCCCGACACGATCGGGGTATCCGGCGCCGAGCTGATGGACCGCTTACTCCGCAGCA
GGCCGGCCAACGCGGAGGAGAGGCCCACCTTCAAACCCGTCCAGGGGCGGCTCGTTAGCCGTACGGATGCGTCCACCTTT
ATGCAGGCCGTGGGCTCGGACGTGCGGAACGCGCTGCAACGTCCCCTGGATGAGGCGGTCGACCTGACCGGCCGGTGGC
CGCATGTGGCGCATGTTTATCTGCGGGATCTCGTTTTCGGAAACGAACGTCTGCGATTCCGGGTTCTGGTCGACCGCAGGC
TGGAGTTGACTCCGAAACTGACCTGGTCGGCCGTCGCCTCCGGAGCCGCTTTGCTGGGCAGGCCCGGGCAGGACGTGCC
GCTGTCGAAACTTGCCGCTCTCGTGCTCGATTCCAAGACCTACGGTGACCGGCGCTATGCCATGTACCTGTATCGGCGGGT
GGCCGCACCGGTGTGCTTCACCGTATCCGCGCTGGTCACCAACGCCCTGTGGCTCGGGGCTCCGTTCAGCGATGACGTATC
GAACCGGAACATCCTCCTTGAGGCACTCCGGCTGCTTCCCCCCTCATGGAACATCTTGCGAATGGCTTCTCCGGAGTTCTCG
GACGTGGACGCACGGATCGGCCCGAAAGACGATGTCCTGTTGCTTCCCCTGCTGAGCCACCGCGATCCCGAGCTCTGGGA
CGAGCCCGACGCGTTTCGTCCCGACCGCTGGGAAGAGCTGGACGGAGACAACCACCCCGGGTATTTCCCCTTCGGCCACG
TAAACGAACGATGCTGGGGCCGGCATATGGTCCTGCCCCTGGCCGAACGGCTCCTCGACATCGTGCGCCGGGACGGACTG
GTGGTGAATCCGGGGCAGACCAGGGGGAAGGTCGAACTCGACGGTTTGCTCGAAGTGTCCGAAGTGCAGATGACTCGAC
CGATGGCTAGTCCTATGGCTCGTCCGTGA

SPL_03767 (OSY44745, CYP105AC)

ATGCAGCACGAACAGATCGCACCGTCTGACCTGTCCGCACCATCTGACCTGTCCGCACCCGCGACCGGGCAGACCGCTCAT
CAAGCGGCGGCGCAGCCCGAGCCGGTGACCCTGCCGCTGACCCGCGCCGCGGGCTGCCCCTTCGATCCGCCCGCCGAGCT
GGCGGAACTGCGCGAGCAGCAGCCGCTGCGCCCGATGCGATACCCCGACGGGCACGTCGGCTGGCTGGCCACCGGCCAC
GCCGTGACCCGGGCGATCGCCGCCGACCCCCGCTTCAGCTCCCGGTACGAGCTGCTGCGCCTCCCCCTGCCGGGCGGGCC
TCCCGGTCCGCTGCCCCCGGCACAGATCGGCGACCTCACCGGGATCGACGCACCGGAGCACACCCGCTACCGGCGGCTGC
TCGCGGGCAAGTTCACCGTCCGCCGGATGCGCCTGCTCACCGAGCGGGTCGAGCAGATCACCGCCGACCACCTCGACGCC
ATGGAGCGCCAGGGACCCGTGGTCGATCTGGTGCAGGCGTTCGCGCACCCGGTCCCGGCCCTGATGATCTGCGAACTGCT
CGGTGTGCCGCAGTCCGACCGCGCGGGATTCCAGGAGCACGCCGCGGTGCTGAGCAGTCCGGATGCCGGCCTGGAGGCG
CAGATGGCCGCCCTGACCGCGCTCACGGAATGCGTCGGCGAACTGGTGCCGGCCAAGCGCGCCCACCCCACCGATGACCT
GCTCAGCGATCTGACCACCAGCGAACTGACCGATGAAGAACTCGCCGGTATCGGCGGCTTCCTGCTGGCCGCCGGTCTCG
ACACCACCGCGAACATGCTCGGGCTCGGCACCTTCGCCCTGCTGAGCCACCCCGAGCAGGCCGCCGCCCTGCGCGCCGAT
CCCGGCCTCGCCGACCAGGCGGTGGAGGAACTGCTCCGCTATCTGAGCATCGCCCACACCGGCGTACGGGCGGCCCTGG
AGGACGTCGAGCTGGACGGCCAACTGATCAAGGCCGGCGACACCGTCACCGTGTCCGTGTCGGCCGCCAACCGCGACCC
GCTGAAGTTCCCCGACCCCGACACCCTCGATCTTCAGCGGAAGGCGACCGGGCATCTGGCCTTCGGCCACGGCGCCCACC
AGTGCCTGGGCCAGCAGCTGGCGCGGGTCGAGATGCGGGTCGCCTTCCCGGCGCTGTTCCGCCGCTTCCCGACGCTGCGG
CTGGCCGTCCCGCCCGAGGACGTGCCGCTGCGCGACGCCATGAACATCCACGGGGTGTACCAACTCCCCGTCACCTGGGA
CAAGGAGTAG

SPL_04452 (OSY43763, CYP174-like)

GTGTCCACTGTTTCTGCCCGATTAGACTCGCCCGAAGGAAAAAACCCGCCACAGCCCGTGCCGCGACTGAGCCTGCGCGA
GAATCTCGTGGGGCTTGCGGAAATGCGGCGCGACCAACTGGCGTTCCTGGTGCGCTCGGTGCGGACCCATGGCGATATCT
TCCATATGCGCCTGGGTGGTCTGCCCACAGTGATGGTCAATCACCCCGACTATCTCCAGCACGTCCTGGTCGACAACCACC
AGAACTACGACAAGGACAACTTCCTCTATCGCGCCACCCGGACAGTGCTGCGCGAGGGGCTCATCGCGAACAAGGGCGG
CGAGCCATGGCGTCGGCACCGCCGGATCATGCAGCCCTCCTTCAGCCGTTCCAGCGTCGCCGACTTCACCACGAACATCAG
CGATCTGACCGGCGACCTGTTGCGCCAGTGGGAAGAGCCGGCCGGCCGAGGCGCGGTGGTCGAGGTGACGACCGAAAT
CGCCAATCTCGCGCTGCGGATCGTGCTGCGCGCGCTGTTCGGTGTCGACGCGGATGACCGCGGCCGGCGCTTCGAGCGG
GACTTCCTGGAAGTGAACGCCATCGCGGGGGATTTCTTCCGGTTCCCCTTCCCGCCGCTGTCCTGGTACTCGCCGTCCCGG
AACCGGCTCCGGCAGTTGATCCGGGAGATGGACGGCTTCATTGCGTATCTCATCGATGCGCGCCTGGCCCGGGAGAGCGA
GAAGCCCGATCTGTTCACGCTGCTGCTCAATGCGATGGACGAGGAGACCGGAACCGGGCTCACCCATGAGCAGTTGGCCA

ATGAGATCCTGGCCATGATCATTGCCGGATATGAAACCTCCAGCAATTCCATCTCCTGGATCTTCTACCAGCTCGCCGCGTA
CCCGGATATCCAGCGCCGGGTGCACGACGAGGTGGACAGCGTGCTGGACGGACGCGTACCCACTCTCGACGATCTGCCCC
GTCTGACCTACACCCGCATGGTGATCGACGAGACACTGCGCCTGTTCACCCCGGCGTGGCAGACCATGCGGCACGCCGTC
GACGACGATGTCATCGGCGGGTACCGGATTCCCCGGGGAACCGATGTCTATCTGAATCTCTTCACCTTCCACCGGCACCCC
GATTTCTGGCCGGACCCGGCCCGGTTCGATCCCGAGCGGTTCACCCCGGAAGAAATCGCCCGGCGGCCACGTCACGTCTA
TCAGCCGTTCGGAAGCGGGCCCCGCCACTGCCTCGGCAAGCATTTCGCGCTGACCGAACTGCACATCATCACGGCCATGCT
CGCCCAGGCCTTCCACATCACCCGGCCGGAGGGCCAGGCCCCCGTGGGTTTCGCGCCGCTGATCACCTTGCATCCCAAGG
GCGGCATACACCTCCGGCTGCACCGTCGCTGA

SPL_04453 (OSY43764, CYP174-like)

ATGCCCCTGACTCCGCGTACCCGTGCCGCGCGCCGGACGCCTTCGCTCAGAGAGATTCCCACGGTCACGGCACGGGACGG
TATCGGCGGAACGCTGGCCTTCAAACGCGACCAGCTCGCCTTTCTCGGCGACGGGCTGGCGCGGCACGGCGACATCTTCC
GGTTCCGCCCGCTCGGCCTGCCCATGGTCATGGTCAACCACCCGGACCACATCCGGCACATCCTGGTCGACAAGGGCGAG
CAGTACGACAAGGACGCCGCCATTTTCAAGGTGGTCCGTCCGGTCCTGCGCAACGGCCTGATCGCCAATGCCGATATGGC
ACTGTGGCGGCGCCAGCGCCGCATGATGGCCCCGCATTTCACCCCGCGCACGGTCAGCGCGTTCGCCCGCAACATGACGG
ACGAGACCGTGCAGATGCTGGAGCGCTGGGAATCCCGCACCGCCGCCGGTGACACCCTCGACATCACTGACGAAATAGGT
CAGTTGGCGCTCCGGATCGTGAACCGGTCACTCTTCAGTGCCGATGTGGGGGCCGCGGCCCAGGCTTTCGAGCGGGCCTT
CGGCGAGGCGAACAGTATCCTCGGCGCCTTCTTCCGCTTCCCGTTCCCTCCGCTCAGCGTCCCGACGCCCAGCCACCGGCG
GCTGCGCCGGGCGATCAACGGTATGGACGCCTTCGTCTCCGGCTTCATCAAGAAGCGGCTGAACGACGAGGTCGCCACG
GGCGAGGAAACCGATCTGCTGACTCTTCTGCTGCACTCCGTGGACGAGGAGGACGGCAAGGGAATGGACCTGGAGCAAC
TGCACCACGAGGTCCTGAACATCTGCATCGGCGCCTACGAAACGACCACCAACACCCTCTCGTGGGCGTTCTATCTGCTGG
CCCGGCATCCCGAGGTCGAGGAGCGGCTGCATGCCGAGGTCGACGAGGTGCTGGGCGGCCGGGTCCCGGCCTTCGAGG
ATCTGCCGAAGCTGGTGTACACCCGCATGGTCGTCGAAGAGACGCTGCGCATCTATTCCCCGGCGTATCAGTTCATGCGCC
GGGCGCGGGAAGAGGACGAGATCGACGGCTATCGGATGCCAGCCGGCACCAATGTCCTCATCAACAGCTATTTCCTGCAC
CGGCATCCGGACTTCTGGGAAGACCCCGAACGTTTCGTACCCGAGCGTTTCACCCCCGAGCAGGTCGCCCGGCGCCCCAA
GCACGTCTACATACCGTTCGGCAGCGGCCACCGCATCTGTATCGGCAAGCACTTCGCGCTGACCGAACTCACCCTCGTGCT
GGCCACCGTGGCCCGGCACCACCGGCTGGTCATGCCCGAAGGCGCCGCCGAGGTGCACCCGGAGGCCCTGATCACCCTG
CACCCCAAGGGCGGAGTGCACTTGCGGCCGGAGCCCCGCGGATGA

SPL_05184 (OSY41108, CYP125A)

ATGCCATGTCCCGCGCTGCCCGATGGGTTCGACTTCACCGACCCCGACGTCTACCAGTCCCGCGTTCCGCTCCCCGAGTTCG
CCCAGCTGCGGCGGACCGCGCCCGTGTGGTGGAACGCCCAGCCGCACGGCATCGCCGGCTTCGGTGACGACGGGTATTG
GGTCGTCACCCGTCACCAGGACGTCAAAGAGGTATCCACCAAGCCGGAGGTCTTCTCCGCGAACCTCAACACCTCGATCAT
CCGGTTCAACGCGGCCATGACCCGTGACCAGATAGACGTACAGAAACTCATCATGCTGAACATGGACCCGCCCGAGCACA
CCCGGGTCCGGCAGATCGTGCAGCGCGGCTTCACCCCGCGCTCCATACGGGCCCTGGAGGACGCGCTGCGCCACCGGGC
GGCGCAGATCGTCGCCGAGGCGCGCCGGAACGGCTCCGGGGATTTCGTCACCGATGTCGCCTGTGAACTCCCTCTCCAGG
CCATCGCGGAACTCATCGGCATCCCCCAGGACGACCGGGCGCGCATCTTCGACTGGTCGAACAAGATGGTCGCGTACGAC
GATCCCGAACTGGCCATCACCGAAGAGGTCGGCCTGAACGCGGCCACGGAGCTGATCTCGTACGCGATGAATCTCGCCGC
GGTGCGCAAGGAATGCCCGGCCAAGGACATCGTCAGCCAACTGGTGGCGGCGGAGCACGAAGGAAACCTCGGCTCCGAC
GAGTTCGGCTTCTTCGTGCTGCTGCTGGCGGTGGCCGGCAACGAGACCACGCGTAACGCGATCACGCACGGGATGCATGC
CTTCCTCACCCATCCCGACCAGTGGGAGCTCTACAAGCGCGAGCGCCCCGAAACGGCGGCCGAGGAGATCGTGCGGTGG
GCGACCCCGGTGGTCTCCTTCCAGCGCACCGCCACCCAGGACACCGAGCTGGGCGGCGCGAAGATCAAGAAGGGCCAGC
GGGTGGGGATCTTCTACTCCTCCGCCAACAACGACCCGGAGGTGTTCGACCGCCCCGAGGTCTTCGACATCACCCGGGAC
CCCAACCCCCATCTGGGATTCGGGGGCGGCGGCCCGCACTTCTGCCTCGGCAAGTCGCTCGCGGTCCTGGAGATCAACCT
GATCTTCGGCGCCATCGCGGACGCCATGCCCGACATGACTCTGGCGGGCGACCCGCGCCGGCTTCGCTCGGCCTGGCTCA
ACGGCGTCAAGGAACTCCGGGTTCACTACGGTTGA

SPL_05473 (OSY40301, CYP208-like)

ATGTCCCTGACCGAATCCCTTCCCCCTCAGCCCGGTTCCGCCCCGGAGCGCGCGATTCCGGGCCCGGCCGGCCTTCCCGTG
CTCGGCTCGATGCTCGACCTCCGGCGCGACTCGCTCAGCACGTTCTTACGGGCCCAGCGGGAGCACGGCGATGTGGTCCG
CCTGGAGGCCGGGCCGCCGGGCCTGCGCAGCGTCTTCCACGCGGTCTTCGCGCCCGAGGGCGTCCAGCAGATCCTCGCCT
CCCGGGCCGCCAACTTCCGCAAGGATCACCCGCTTTACGAGGAGGTCCGGCAGGCCTTCGGCAACGGCCTGCTGACCAGC
CAGGACGCCGACTACCTCAGACAGCGCCGGCTGGTGCAGCCGCTGTTCACCAAGCGCCGGGTCGACGGCTACGCCTCGGC
GGTGACCACCGAGGCCGCCACGCTGGCCGACCGCTGGCGGACCGCGGGCAGCGGCACCGTTGATCTGGTCCCCGAGATG
AACCGGCTGGCGCTGCGCACCGTCTCCCGGATCCTGTTCGGTCTGGACGCGGAGGCGGCGGTGGACGCCATCCACCGCTG

CGCCCCGGTCATCAACGCGTATGTCGTCCGGCGCGCCTACGTCCCCGTGAAGATCCCCCGCGAGTGGCCCACCCCCGGGA
ACCAGCGGGCGCGGGCCGCCACCGATGAGCTGAACGCGCTGTGCGACCGCATCGTGGCGGACCGGCGGGCCTCGCGCG
CCACCACCGGGGGCGCGGACGGCGCGGGCGAGGATCTGCTGTCCCTGCTCGCCGCTGCGGGCAATGACGAGGACGGCTC
GCTGGACACGACCGAGGTCCGCGAACAGGTCCTGATCTTCCTGCTGGCCGGGCACGAGACCACCGCGACCTCCCTCGCCT
TCACCCTCCACCTCCTCGCCCGGCACCCGGAGGAACAGACCCGGGTCCGCGACGAGCTCGACCGCGTCCTGGGCGACCGC
ACCCCCACGGCCGCCGACCTGGACCGACTGCCGTATCTGACACAGGCGCTCAAGGAATCGATGCGGCTGTATCCGGCGGC
GCCGGTGATCAGCCGCCGGGCCGTGGCGGCCGCCGAGATCGGCGGATGCCGCATCCCGGCCGGTGCCGATGTGGTCGTC
GCGCCCTGGGTCACCCACCGCGACCCGGAGCTGTGGCCGGACCCGGAGCGCTTTGACCCGCGGCGCTTCGACCCGGAGC
CGGAGGCGGCGCGGCACCGCTACGCCTGGTTCCCGTTCGGCGGCGGGCCGCGCGCCTGCATCGGACAGCACTTCTCGAT
GCTGGAGTCCGTGCTGGCGCTGGCCGTGCTGCTGCGCGGGTACGAACTCGACGCCGTGGACGAGGAGGTGCCGGTCTCG
GCCGGAATCACCCTCCAGGCGACCGGTCCGGCGCGGGTGCGTCTGCGGCCGCGCGCCGGGGCGGGGGAGTAG

SPL_05497 (OSY40215,CYP154A)

GTGGCCCGGACCGAACTGCCCGGCGGAATCGAGGCCTGGACGGTCACCAGCCACACCCTCCTCAAGCAGCTGCTGACCGA
CGACCGGGTCTCGAAGAACCCCCGGGACCACTGGCCGGAGTGGCAGCGCGAGGAGGTCCGCAGCAGCTGGCTCCAGCAC
TGGATCGGCGTCACCAACATGTTCACCGCGTACGGCGCCGATCACCGGCGGCTGCGCAAGCTGATCGCCCCGGCGTTCAC
GGCCCGCCGTACCGATGCCATCCGCCCCCGGGTGGAGAGCATCACCATGGCGCTGCTGGACGACCTCGCCGCGCACCCGG
CGGGCGAGCCCCTGGACCTGCGGGAGCGGTTCAACCACCCGCTGCCCATGCAGGTCATCTGCGAGCTGTTCGGCTTCCCC
GAGGGCGAGGCGCGCGCCGAACTCGCCCGCCTCATCACGGCCATCATGGACACCACGGCCACCCCGGAGCAGGCCGGTG
CCACCGGCGCGGCCGTCCATGCGCTGCTGTCCGACCTGGTCGCCGCCAAGCGCGAGCACCCGGCGGACGACCTCACCAGC
CTGCTGGTCACCGCGCGGGACGACGACGGCCGGCAGATGACCGAGAAGGAGCTGCTGGACACCCTGCTCCTGGTCATCG
GTGCCGGTCACGAGACCACCGTGGATCTCCTCGGCAACGCGGTGCACGCCCTGCTGACCCACCCCGAGCAGCTCGCGCTG
GTGCGGTCGGGCGAGGTCTCCTGGAACGACGTGATCGAGGAGACCCTCCGCTGGGCACCGAGCATCGCCGCGCTCCCCCT
CCGCTTCGCCGTGGAGGACATCGAGATCGCGGACGGCCCGACCATCCGCAAGGGCGAGGCGATCCTCCCGGGGTACGCG
GCGGCCGGCCGGGATGCCGCGTTCCATGGCGACACCGCCGCCGACTTCGACATCCGGCGCTCCCACCAGGAGCATCTGGC
GTTCGGCCACGGCGTCCACCACTGCCTCGGCGCCCCGCTGGCCCGTATGGAGGCCCGGATCGCCCTCCCGGCCCTCTTCGC
CCGCTTCCCGGACCTCCAACTGGCCACGGCGGCAGAGGAGCTGGAGGCCACGAGCGGCTTCATCTCGGGGGGTCTGCGG
AGCCTGCCGGTCCGCCTGACTCCGGCGTGA

SPL_06346 (OSY37796, CYP107Z)

ATGTCGGCATTATCCAACTCCCCGCTCGGCGCGCACGTCGGGAAACACCCTGGCGAGCCGAATGTGATGGACCCGGCGCT
GATCGCCGACCCGTTCGGCGGCTACGGCGCACTGCGTGAGCAAGGCCCGGTCGTACGGGGCCGGTTCATGGACGACTCG
CCCGTCTGGTTCGTGACGCGCTTCGAAGAGGTCCGCCAAGTCCTGCGCGATCAGCGGTTCGTGAACAACCCGGCCGCGCC
GTCCCTGGGGCGCTCGATCGACGAAAGCCCCGCGGTCAGACTTTTGGAAATGTTGGGGTTGCCCGACCATTTCCGGCCGT
ATCTGCTCGGGTCGATCCTCAACTACGACGCACCCGACCACACCCGGCTGCGCCGACTGGTCTCGCGCGCCTTCACGGCAC
GCAAGATCACCGACCTGCGGCCGCGGGTCGAGCAGATCACCGACGACCTGCTGACCCGGCTTCCCGAGCACGCCGAGGA
CGGTGTGGTCGACCTCATCCAGCACTTCGCCTACCCCCTGCCGATCACCGTGATCTGCGAACTGGTCGGCATCGCCGAAGC
GGATCGTCCGCAATGGCGGAAGTGGGGGGCCGACCTCGTCTCGCTGGAGCCGGGGCGGCTGAGCACCGCGTTCCCGGCG
ATGATCGAGCACATCCATGAGCTGATCCGCGAGCGGCGCGGGGCACTCACCGACGATCTGCTCAGCGAGCTGATCCGCAC
CCATGACGACGACGGCGGCCGGCTCAGCGACATCGAGATGGTCACCATGATCCTCACGATCGTCCTGGCAGGCCACGAGA
CCACCGCCCACCTCATAGGCAACGGCACGGCGGCACTGCTCACCCACCCCGACCAACTGCGCCTGCTCAAGGACGACCCG
GCGCTGCTGCCGCGCGCCGTCCACGAGCTGATGCGCTGGTGCGGGCCGGTGCACATGACCCAGCTGCGCTTCGCCTCCGA
GGACGTCGAGGTCGCCGGAACGCCGATCCGCAAGGGCGACGCCGTACAACTCATCCTGGTATCGGCGAACTTCGACCCCC
GCCACTACACCGACCCCGACCGTCTCGACCTGACCCGCCACCCCGCCGGCCACGCCGAGAACCACGTGGGGTTCGGCCAC
GGCATGCACTACTGCCTGGGCGCCACCCTCGCCAAACAGGAGGGCGAAGTCGCCTTCTCCCGGCTCTTCACGCACTACCCG
GAGCTGTCGCTGGGCGTCGCGGCGGAACAGCTGGCGCGGACGCAGGTGCCCGGCAGCTGGCGGCTGGACACCCTGCCG
CTGCGGCTGGGGTGA

SPL_07341 (OSY34873, CYP159A)

ATGACCACTCTCCCGCAGCCGCCCGACATCCTGTCACCGCACTTCGCGGCGGACCCCTACCCCGCGTACCGCGTTCTGCGG
GAGCACTACCCCGTCTTTCACGACCCGGGCACCGACAGTTTTCTGCTCTCGCGGTACGAGGACGTGTCCCGGGCCTTCCGC
GATCCGGTGTTCACCGGCGACAACTACGAATGGCAGCTCGAACCGGCGCACGGCGGCCGGACGCTGCCCCAGATGAGCG
GCCGTGAGCACGCCGTACGCCGGGCCCTGGTGGCCCCCGCCTTCCGGGGCCGGGAGCTGCGGGAGAAGTTTCTGCCGGT
CATCGAGCGCAATGCACGGGAGCTGATCGACACCTTCCGCGACGCGGCGGAGACCGATCTCGTGGCGCGGTTCGCGACC
CGCTTCCCCATCAACGTCATCGTCGACATGCTGGGGCTGGACCGGGCCGACCACGACCGCATCCACGACTGGTACGGCTC

CGTCGTCGGCTTCCTGGCCAATCTGGCCCAGGACCCCGAGATCGCGGAGGCCGGCCTGCGCGCCGGCCGGGAGCTCGCC
GACTACCTCCTCCCGGTCATCCACGAACGGCGGTCCGCCCCCGGCGACGATCTGCTCTCCATGCTCTGCACCGCCGAGGTC
GAGGGCACCAGGATGAACGACCAGGACATCACGGCGTTCGTCAGCCTGTTGCTCTCGGCCGGCGGCGAGACCACGGACA
AGGCCCTGGCCGCGCTGTTCCGCAATCTGCTGTGCCACCCGGACCAGTTGGCCGCCGTACGCGCCGACCGGTCGCTGATC
CCCGCGGCCTTCGCCGAAACGCTGCGCTTCACACCACCGGTGCAGATGATCATGCGTCAGACCGCCGCCGAAGTATCCCTC
AGCGGTGGCACCATCCCCGCGGGCGCCACCGTCACCTGTCTGATCGGCTCCGCGAACCGCGATGAGCGCCACTACACCGC
GCCGGACACCTTCCATGTCCTCCGCGACGATCTGACGCCCACCACCGCCTTCAGCGCCGCCGCCCAGCACGTCGCCTTCGG
CCTCGGCCGGCACTTCTGCGTGGGTGCGCTCCTGGCCAGGGCCGAGGTGGAGATCGCCGTCAATCAGCTGCTCGACGCGT
TCCCCGCGATGGACTTCGCGGACGGTACGCCGCCCACCGACGCGGGGGTTTTCACCCGCGGGCCGGAGCAGCTGCGGGT
GCGGCTGGTGCCGGCGGACGACCGGGCCCGGGCGAGCGAGGGCTGA

AAU_00443 (OSY43500, CYP107)

ATGACCGCACCGACCGGAGCGCCCGACCTGCTGGCACCCGACTTCGCAGCGGATCCGTACACCGGATACGCCCGGCTGCG
GGAGCAGGCTCCCGCGGTGACGGTGTCGCTGATGGGCGGCCCGCCGATGATGCTGGTGACCCGCTACGACGAGGTGCGG
ACGGTGCTCACCGATCCGCGTTTCGGCACCGATCCCGATCAGGCCGGTGCCGGGGTGAACGTCCGGGAAGCGATGCTGC
GCAAGGCCGGTGTGCCCGACGACGTGATCGGCTACCTGCTCCGCACGATCCTCACCTCCGACGGCGACGACCACACCCGG
CTGCGCAAGCTGGTGTCGCGTGCCTTCACCGTGCGGCGGGTGCAGTCGCTGGCACCCCGGGTCGAGCAGATCGCCGCCG
GCCTGCTCGACGACATCGCTGCGGCAGGCGCCGACGGCAGCCCGGTCGACCTGGTCGAGTCGTTCGGCTATCCACTTCCG
ATCGCGGTGATCTGCGAGCTGGTCGGGGTGCCCGAACCGGAGCGCGCGCAGTGGCACGGCTGGGGCAGGGTGCTGACC
ACGCTGGATCCCGAGCACACTCCCGGAGTGCTGCGGGCCGCCGTCGAGCACGTGCACGAGTTGATCGCGGCGCGTCGCG
CCGAACCCGCCGACGACCTGATCAGCGCGCTGATCGCGGCCCAGGAGGACGACGGCGACCGGCTCTCCGACCAGGAGAT
GGTCACGATGGTCTTCGCGCTGGTGATGGCCGGGCACGAGACCACCGCGCACCTGCTCTCGAACGGGGCGCTGGCCCTGC
TGACCCGTCCCGACCAGCTCGCGCTGCTGCGCGCCGAGCCCGACCGCTGGCCGGCCGCGGTGAACGAGCTGATGCGGGC
CCGCGGGCCGGTGCAGTTCACCCAGATCCGCTATCCGGTCACCGACGTCGAACTGGGCGGCGTCGCGATCCCCGCCGGGA
CCCCGGTGATCGCCGGTCTGCTGCCGGCGAACCACGATCCGCGGGCCTTCGACCGGCCGGACGACATCGACGTCCGGCG
GGACACCGGCCGTGGTGAGGGCCATCTCGGGTTCGGCCAGGGCGTGCACTACTGCCTGGGCGCAGCGCTCGCCCGGCAG
GAGGGCGACATCGGCCTCCGGATGCTGTTCGACCGGTTCCCGGGCATGCGGCTCGCGGTGCCCGCGCAGGAGCTCGCCT
GGCTGCCGCGGCCCGGGATGAGCCGGGTCGGGGAGCTGCCGGTGCTGCTGTCCTGA

AAU_00941 (OSY43336,CYP107)

ATGACCCCGACCGCCGCCGACGTCACCGGCGCCGACCTGCTCTCCCCCGAGGCGGTCGCCGATCCCTATCCGCTGCTCGCC
GCGATGCGCGAGCACGATCCCGTGCACTGGAGCGAGCGCTTCCGGTCGTGGTTCGTGACCCGTTTCGACGACGTCACCAC
CGCGCTGCGCGATCCCCGGTTCTCCTCCGACCGGATCACCCCCTACCGGAAGGCGAAGCTCGACGGCCCGGACGCCGATC
CGGCACTGCGGGCCGCGTTCGGTGTCCTCGGCGAGTGGATGGTGTTCAAGGATCCTCCCGACCACACCCGGTTGCGCCGG
TTGCTCAGCCGCAGCTTCACCCCGCGCGCGGTCGGCCGGATCGCCCCGCGGATCACCGAGCTCGTCGACGAGCTGCTCGA
CGCCGCACCGCCCGGTGAGCTCGATGTGGTCCGGGATCTCGCCTATCCGCTCACCGCGTCGGTGATCGCCGAGATGCTCG
GCGTGCCGCGGGCCGACCAGGAGCGGTTCAAGGACTGGTCCGACCGGATCACCGGCCTGGTGTTCGGCGGCCTGTCCGA
CTCCGACCGGCACCGCAGCGGCGCCGAGGGGATGGGTGAGCTCACCGGATACCTGACCGAGCTGGTCGCCCGCCGCGAG
CGGGAGCCGGCCGACGACCTGCTGACCGGGCTGATCACCGCCCGCGACGAGCAGGACGCCCTGACCCACGACGAGGTCA
TCGCGACCGGCGTGCTGCTGCTGTTCGCCGGCCACGAGACCACCACGAACCTGATCGGCTGCGGGGTGCGGGCGATGCTC
GCCCACCCCGATCAGGCCGCCCGGCTCGCCGCCGATCCCGGGCTGGTCAACGGGGCCGTGGAGGAGATGCTGCGCTTCG
ACGGCCCGGCCAAGACCGTGGTTCGGCTGATGGCCGAGGACGTCGAGCTGCGCGGCCGGGTGCTGCGCCGCGGCCAGC
GGGTCTTCCTCAGCCCGTCGTCGGCGAACCGGGATCCGGCGGCGTTCGACGATCCCGACACCTTCGACATCACCCGCAGG
CAGGGCCGCCAGCTCGGCTTCGGGGTCGGGATGCACTACTGCCTGGGCGCTCCACTTGCCCGGCTGGAGACGGCGATCG
CGGTCCCCCGGATCCTGGAGCGGCTGCCGGGTGCGGTCCCGGCGGGCGAACCGCACTGGGCGCCGGTGCTGCTCTCTCG
CGGCATGGAGCGCTACCCCCTGCGCACGGGCACCGGATGA

AAU_01045 (OSY42804, CYP107AL)

ATGGACGTCGACACCCGGGACGGCACGAGCACCACCTCGGGCTGCCCGCTGAGCTATCCCTTCAACCGGCCATCGGCGAT
CGAGCTGCCACCGGTGTACACCGACCTCCGGGCGGACACCCCGGTGGCCGAGGTGACACTGCCCAGCGGCGACCGCGGG
TACGTGGTGAGCCGCTACGACGACGCGAAGCTCGTCCTCGCCGATCCGCGCTTCAGCCGGGCGGCGATGGCCGCCGACG
GCGCACCGCGGCTGACCCCGGTCCCGATGCCGCCGAACAGCCTGTTCAGCACGGACCCGCCGGAGCACACCCGGCTCCGC
AAGCTCGTGACCGGCGAGTTCACCCGGCGCCGGATCACCGCGCTGGAGCCCCGCATCCAGGAGATGACCGACGAGCTGC
TCGACACGATGGCCGCGAACGGCGGCCCGGTCGATCTCAACCCGGCGCTGGCGTTCCCGCTGCCGGTCGGCGTGATCTGC
GAGCTGCTGGGCGTCCCGTTCGAGGACCGGGAGAAGTTCCGGGAGTGGTCGAACGCGATCGTCTCGCTCACCTCGCACAC

CCCCGAGCAGATGATCGAACAGCGGATGCAGATGGCGGGCTACCTCTACGAGCTCATCGGCCGGCGCGCCGACGAGCCG
GGTGAGGATCTGCTGAGCGCGCTCATCCAGGCGCACCAGGACCGGGGCGCCCTCGACGAGGGCGAGCTGATCGTGATGG
CGATGACCCTGCTCGTCGCGGGCCACGAGACGACGGTCAGCATGATCGGCGCCTGCGTGCTGACGCTGCTGCGGCACCCG
GAGATCCTCGCCGAGGTGGTCGAGAACCCGGGGCGGGTCGACGCCATGCTCGACGAGCTGCTGCGGGTCAACCCGATCG
GCGACGGCGGCCCGCTGCGGATCACGCTGGAGGACGTCGAGATCGCCGGCACCGTGATCCCGAAGGGCAGTGCGGTGCT
CGCCGCGGTCTGCTCGGCGAACCGGGACGAGAAGATCTTCGACCGGCCGGACGAGTTCGTCCCGGGCCGCGAGCACAAT
CCGCACCTCGCCTTCGGGCACGGCATCCACCACTGCCTCGGTGCGTCGCTGGCCCGGTCCGAGCTGCGGATCACCATCGA
GAGCCTGTTCCGGCGGTTCCCGACGCTGCGCCTGGCCGCACCGGTCGAGGAGCTGCGGATGAAGAGCGGGATGCTGGTG
CACGGCCTGGAGCGGCTGCCCGTCACCTGGTGA

AAU_01156 (OSY42914, CYP105)

ATGCAGCAATGCACACCGTGCAGCGATTCGGTAGGCTGGCGGGGAACCCGACGACGCCCAGGACAGGAGCCCGCCCTGA
CCGCCGACGAGAACACCACCGCCACCGACTCCCCCATCGCGCTGCGCAGCCTGGTCCCCGAGGACCTGCTCGAGCGCACG
CCGAGCTGCCCGTTCGATCCCGCGCCGGGGCTCGCGGCGCGGCGCGGCGCAGGCGCCGTTCAGCCGGTCGAGCTGCGCA
ACGGAGCCCGGGCGTTCCTCGTGACCGGCTTCGACGAGGCACGCGAGGTGCTGGCCGACTCCCGGTTCAGCGCCGACCG
GGTCCGCTACAAGGACGCGACGAAGCTCTCCGCACAGGAGGTCGAGCAGCTCGCCGCTGCCACCGAGCGCGGTGAGCCG
CTACGCCCGGCCGACGTGCAGCGCAACGACGGCATGTTCATCTTCATGGACCCGCCCGAGCACACCCGCATCCGCAGGCT
CCTGCAGGGCCAGTTCACGGTCCGCCGTATGCGCGCACTCGAGGAGCGCATGACCGAGATCGCCGCGGCGCACATCGAC
ACGATGCTCGCCGGTGGCCCCGGCGCCGACCTGGTGCCCGCCTATGCGCTGCCGCTGCCCTCGCAGATGATCTGTGAGCT
GCTCGGGGTCGACTATGCCGACCGGCAGGCGTTCCAGGACAACACCACGGTCGGGCTCAACGCCAACTCCACCGACGAG
GAACGCGGCCGGGCACTGGGCGAGCTCTACACGTTCCTGTCCGGGCTCGTCGCGCACAAGAAGGAGCACCCCGGCGACG
ATCTGCTGTCCGGGCTCATCCACGAGTCCGACCCGCCACTGCCGGACAACCAGCTGATCGACATCTCGCTGGTGCTGCTCG
GCGCCGGGCACGAGACCACGGCCAACATGCTCGCCCTCGGCGTGTTCGCGCTGCTCCAGCACCCCGAGCAGCTCGCCGCG
GTGACCACCGGGCCGGACGCGATCGACGGCGCCGTCGAGGAACTCCTGCGCTACCTGAGCATCATCCAGCTCGGGGTGTC
CCGGGTCGCGACCGAGACCGTCACGCTCGGCGGGGTCGACATCCCGGCGGGCTCCACCGTCATGGTGGCGGTGCCGGAG
ACGAACCGGGACGAGCGGCACCTCGACGCCGTCGACGAGCTCGACGTCCGCCGGGGGCGGGTGCCGCACCTCGCCTTCG
GGCACGGTGTCCACCAGTGCATCGGTTCCCAGCTGGCGCGCGTCGAGATGAAGGTCGGCTTCCGCGAGCTGTTCGCACGG
ATCCCCGGCCTACGCCTCGCCGTGGCAGCTGAGGATGTCCCGCTCCGCAACGACATGCTGATCTTCGGGGTGCACTCGTTG
CCGGTCACCTGGAACTGA

AAU_01495 (OSY42000, CYP107-like)

GTGTCCGTCGTACCCCCCGCCGGCTCCGCCGCCCCGCAGGCGCCCGGACCCGTGGAGTCACTGTTCCTGGGCCGGCTCGC
CGATCCGTACCCGGTCTACGCCGAGCTGCGTGAGACCGGCGACGGCGTCCACTGGTCCGATGTGCTGGGCGCCTACATCG
TCACCCGCTACGCCGACGTCCGGCAGATCGGGACCGACCCGAAGCTCTTCTCCAGCGACGTCTTCCACGACTCCGCGCCGA
GCACCCACGACCCGTCCGATCCCGAGCACCTGCGGTTCCTCGACGCCGCGTCGCGGCTGTTCATGTTCGCCGATCCGCCGG
TGCACACCCGGATCCGGTCGACCTTCCGGCACGTCTTCACGGTGCAGGCCACGCAGCGCTGGCGCCCGCTGGTCGAGCGG
GTCACCGCCGAGCTCATCGACCGCTACCCGCGCGGCGAGGAGTTCGACATCATGCCCGGCTTCGCCGCGGACGTGCCGGT
CGCCGTGATCGCCGCGATCCTCGGCGTGCCCGACGAGATGCGCCCGCGGTTCCGGGACTGGTCCTACGCCTACGCGTCCA
CCTTCGATCCGGTCGTCCAGGGCCCGCGCCGGGACCAGGCGATCGCCACCTCGCTGGAGCTGTTCGACTACCTCGGTGAG
CTGGTCGCGGCCCGCCGCGCCGAACCGCGCGACGACCTGATCAGCACCCTTGTCCGGACCGAGACGATCGACGGCGACA
CCCTCGGCGACATCGAGCTGCTCGCCCAGCTGGCCCTGCTGCTGGTCGCCGGCAACGAGACCACCACGACGCTGATCGGC
GCCGGCCTGACCCTGCTGTTCGACCATCCCGGTACCCGCCGGCGGATCGAGGCGGACATGTCGCTGCTGCCCGCGGCGAT
CGAGGAGATGCTGCGGATCGACCCTCCGCTGCACCTGGTGATCCGCAAGGCCACCGCCGACGTCGAGATCGGCGACCAC
ACGATCCCGGCCGGGAGCCTGCTGTTCCCCTGCCCGGCCGCGGCGAACCGGGACCCGCGCCGGTTCACCGAGCCGGACA
CCTTCCTGGTCCCCCGCGAGGACAACAAGCATCTGGCCTTCTACCACGGGGTGCACTTCTGCGTGGGCGCCCCGCTGGCCC
GGCTGGAGGGCCAGGTCGTGTTCGAGAACATCCTGCGGTCGTTCCCGGACGCGCGTCCCGGAGCCGTCCCGGCGGTGCG
CCGCACCACCAACTCCGTCGCCCGCGGCTGGGAGAGCCGGCCGGTGATCCTGTGA

AAU01848 (OSY41819, CYP107-like)

ATGATCACCGCACCGCAGCTGCCCTTCGACCGCCCGAACATCCTCGAGATCGCGCCGCTCTACGCGGTGCTCCGCAGGCAG
GCGCCGGTGGTGCCGGTGCGTACCCCGGTCGGTGATCCGGCCTGGCTGGTGACCCGGTACGCGGAGGCCCGTGAGCTGT
TCGGCGACTCCCGGCTCGGTCGCTCGCACCCGGAACCGGAGAACGCGGCGATGATCACCGAGGCGGCCATCCTGAACGG
GCCGAGCGGCAGCTACGAGACCGAGCAGGAGGAGCACGCCCGGCTGCGCGCGCTGCTGGTGCCGGCGTTCTCCGCGAA
GCGGATGCGGCTGCTCGGCGACCACGTCGGCGAGCTGGTCGACGGCTGCCTGGACGACCTGGTCGCGGCGCGCGACGCC
GACGGCGTCGCCGACCTGCACGCGCACCTGTCGTTCCCGCTGCCGGTGCTGGTGATCTGTGAGCTGCTGGGGGTGCCCTA

CTCCGATCGCGAGTACTTCAGCGGGCTGTCGGAGCGGGTGGCCGCGATCCGCTCCGGCGGGGACGCGGCGACCGCGATG
GCCGAGTTCCGCGCCTACACCGGCCGGCTCGCCGCCACCAAGCGGGCACAGCCCGGTGAGGACGTGATGACCGACCTGG
TCCGGGCGCAGGAGGCGGACCCGTCGTTCGGGGAGGAGGAGCTGACCCGGATCGCCGCCGGCCTGCTCTTCGCCGGGCA
CGAGACCACGGTCGGCCGGATCGATCTCGGCGTGTTGCTGCTGCTGACCGATCCGGCGCGGCGGGACGCCTTCGTCGCCG
ACGTCGAGGGCAGGCTGCACGGCACCGTCGAGGAGGTGCTGCGGCTGTCCGCGCCGGGCGGGCTGGGCCTGCTGCGCT
ACGCGCACGACGACATCGAGATCGGCGGGGTGACGATCCGGCGCGGTGACGCGGTGATGATCTCCACCGACTCGGCGAA
CCGGGGACACCGACGCCTTCGACGCCCCCGAGGAGTTCGACCCGACCCGCCGGCCGAACCCGCACGTCTCGTTCGGCTACG
GCGGATACTTCTGCATCGGTGCGAGCCTGGCCCGCACCGAGCTGTCGACGGTGTTCGCCCGGCTGTTCACCCGGCTGCCC
GGCCTGCGCCTTGCGGTCTCACCGGACGAGCTCGACGTCCGGTCCGACCAGATCACCGGGGGTGTGCGGGCGCTGCCGG
TCACCTGGTGA

AAU_02550 (OSY40792, CYP107-like)

GTGGCCAGTGTTGCCGCACCGCACGTCCCGATCGAGGATCCACAGTTCTATCTCGACGATCCGTGGCCGACGTTCGCCTGG
ATGCGCGAGAACGCGCCGTTCCACTACTACCCACCGCTCGACGCGTGCGTGCTCACCCGGCATGTGGACGTGAAGGAGGT
GGCGAGCAGGGCGACCGAGTTCGTGAGCTCGCGCGGGATCTTCCTCAACGACTGGAAGTACGCCGATCACGGCGGCGGC
GACCAGGACGAGACCCTGACCGACAGCTTCTTCCCCCGGGGTGGTGAGCAGGTCGGCACCACCGACCCGCCGCGGCACA
CCGAGCTCCGCCGGGTGATCGCGCCCGCCTTCGCGCCACGGGCACTGCGGCGGATGCAGGAGCGCCTCACCGTCGAGAT
CGAGACGATCCTCTCCGGCATCGTGCCCGGCGCGGTGACCGACTGGATGCCCTACGCCGGGCTGGTGCCGATCAAGGCTG
CGACCACGCTGCTCGGCCTGCCGGACGCCGACGTCGGCCGGGTGCAGTTCTGGAGCGACGAACTGGAGAAGCTCGGCGG
GGATCTCACCCTCGACGAGCTCCGGGCGGCGGCCGCGGAGTTCCACAGCCTGCAACAGTTCATCCTGGACAACGCCGAGG
CGAAGCGCCGCGATCCCGGCGGCGAGGACCTGCTCTCGGTGCTGCTGGAGGCGGAGCTGGACGACGACAAACTCTCCGA
GGCCAACGTCATCATGTTCGCGATGACGGTGCTGGCGGCCGGCAGCGACACCACCCGGGCGCTCCTCGCGGGACTCGTCC
ATCACCTCGCCCGGCACCCCGAGCAGTGGGAGATGCTGCGCGCGGACCGCTCGCTCGTGCCCCGGGCGATCGAGGAGAC
GCTGCGTTACGTCACCCCGGCCCGTGCCTTCGGCCGCACCGCCGTGGCCGACACGACCGTCAACGGCCGGGAGGTCGCGG
CGGGGCAGCGGCTCTACCTCATGTACATGGCGGCGAACCGCGACGAGACGGTGTTCCCGGATCCGGGCCGCTTCGACATC
ACCCGCCAGGAGAGCACCCAGCACCTGGCACTGAGCACGGGTGCCCATGTGTGTGCCGGCGCCGCACTCGTGCGGCTGC
AGGCGCCGCTGCTCGTCGAGGCGTTGCTCGACCGTTTCGCCGGGATCGAGGCGGCGGGCGAGGCCGTTCCGGTCCGGCA
CGTCATCCGGAACAGCTGGACCCGAATGCCGATGCGGTTCCGCACCTGA

AAU_03957 (OSY38754, CYP107AQ)

ATGCATGCGTCCGATTCCCCAGACCCCCGGTCCCGGCCCGCCCCGGCACTGTTCGACGCCACGGACCCGGGCTTCCTCGAC
GATCCCTACCCCGCCTACGCCGCACTGCGCGCCGCCGGCCCGGTGCACGAACACCCGGGGCTGGGCCTGCCGGTCGCGGT
CACCCACGCGGCCTGCTCCGCGGTGCTGCGCGACCGCGGTCTCGGCCGGATCTGGACCGATGCCCGCCCCGAGGCCGAGC
TGGCCGCGTTCAACCTCCTGCACCGCAACTCGCTGCTGGAGCGGGAGGGCGAACCGCACACCCGGCTACGACGCCTGGTG
GCCGCCGCGTTCGCCCGCGGGCACACCGAGCGGCTCGCCCCGCTGGTGCGCACCCGCGCCACCGCGCTCGTCGAGGACCT
CGTCACCAGAGTGCGCGACGGCGAGCCCGCCGACCTGATCCCGCTGGTCGCCGAGCCGCTGCCGGTCCAGGTGATCGCC
GACCTGCTGGGGGTGCCCGACGACCGGCGCGCCCCGCTGCGGGACTGGTCGGACGCGATCGTGCGGATGTACGAGCCCG
ATCCCGGGACGGACGGGCGGATCGCTGCCGAGCGGGCGTCCGCCGACTTCGTCGCGATGCTGCGGGACCTGGTCGCGCA
GCGCACCCGGCACGGCTCCGGCAGCGCCACGGACCGGCCCGGACCGGCGGCGGATCCGTCCGGACGACCGGGGGACCT
CATCGGCGATCTGCTCGCGGTGCGCGACTCCGGGGACCGGATCTCCGCCGACGAGCTCGTCGGCACCGCGGCCCTGCTCC
TGATGGCCGGGCACGAGGCGACCGTCAACGTCATCGGCAACGGGGTGCACGCGCTGCTGCGGCACCCGGACCAGTGGCG
ACGCCTGGTGGCCGATCCGGAGCTGGTCGGGACGGCCGTCGAGGAGCTGATCCGGTTCGACGCGCCACTGCAGCTGTTC
GAGCGGACCGCGGTCGTCGACACGACGATCGCCGGGCACCCGGTGCCCGCCGGCGGCCGGATCGGGACGCTGCTCGGC
GCCGCCGGTCGCGATCCCGCCGCGTTCGGCGACGACGCCGACCGGCTCGACGTCGGCCGGTCCCCGAACCCACACCTGGG
GTTCGGTGGCGGCGTGCACTACTGCCTCGGCGCCCCGCTGGCCCGGCTCGAGATCGCCGAGGTGCTGCGCGCGCTGCTGC
GCCTGCTCCCGGACGCGGTGCCGGACCGCGCCCCGGTCCGCCGGCCGCGTTTCGTCATGCGCGGCTGGTCGGAACTGAGA
CTGTCCGGCTCGGCCGAACGGCCCTCATTGCGTTGA

AAU_05422 (OSY36345, CYP107)

GTGGGCGGGGCGCACGCCGGCGGCCCGCTGCGGAACCGGGGGGAACCGACGATGACAGACACCACCGAACCGCTCGTG
CTCGACCAGGAGTTCTTCGACGACCCGGACCAGCTCTACCGCGAGCTGCGCGCCGAGCGGCCGGTCACCCGGGCGATCG
GCCCGAACGGCGTGGCCTTCTGGATGATCACCCGTTACGCCGACGCCCGCGCCGCGCTCAACGACCCGCGGCTGGCGAAG
GACGCCCGCAGGATCCCCGAGCTGCTGGCCCGTCAGGAGTCCGGGCCACCGGCCAGAGAGCTGGCCGAGTCGCTGGTCG
GGCACATGCTCAACGCCGACCCGCCTGACCACACCCGGCTGCGCAAGCTGGTCGGGCGGGCGTTCACGATGCGGGCGAT
CGGCAGGTTGCGACCCCGCATCGAGCAGATCGCCACCGAGCTCGCCGACGCGATGACGGCCTCCGGGCCCGAGGTCGAC

CTGCTCGACACGTTCGCGTTCCCGCTGCCCATGACGGTGATCTGCGAGATCCTCGGCGTCCCGCCGGACCGCCGCGAGGA
GTTCCGCACCTGGTCGAACACCCTTCTCTCCGCCGCCGGGGACTCCGAACGAGGCGCGGCCGCCGCCGCGATGGCGACCT
ACCTCTCCGAGCTCGTCGAGGACAAGGCCGCCCATCCCGCGGACGACATGCTCTCGGAGATCGTCCGCGCCTCCGAGGAC
GGCGACTCGCTGAGCCCCGGCGAGACGACGGCGATGGCGTTCCTGCTGCTCGTCGCCGGGCACGAGACGACCGTGAACC
TGATCGGCAACGGAGCGCTCGCGCTGCTGCGCGACCCGGAGCAGCTGGCCCTGCTGCGCGCCGACCCGGACCGGGTGCC
CGCCGCCGTCGAGGAGATCCTGCGCTACGACGGGCCGGTCAACCTGGCGACCTTCCGGTTCACCACCGAGCCGGTCGAGT
ACTCCGGCACGACGATCCCGGCCGACGCGTTCGTGCTCGTCTCGCTGCTCGGGGCGAACCGCGACCCGCAGCGCTGGCCG
GACGCGGACCGGTTCGACGTCGAGCGCGACCCGTCCGGCCACCTGGCCTTCGGCTTCGGCATCCACCACTGCGTCGGCGC
CCCGCTCGCCCGGCTGGAGGGGGAGATCGCCTTCCGCACCCTGCTGGCCCGGTTCCCGGATCTGCGGCTCGCCGGCGAAC
CCGGCCCGCACCGGATGAGCACGCTGATCCACGGCCGCACCCGGTTGCCGGTCAGATTGGATGCGCCGGCCGGGGCCGC
AACGGGACCGGCGGCCACGGGAGCATCCGCAACGGGACCGGCGGCCAGGTGA

AAU_05736 (OSY35899, CYP105AB)

ATGACGACCGTCGACGAGTTCCCGCAGTCCCGTACCTGCCCGTTCGCTCCGCCGCCGGCGTACGCGAGGATCCGCGAGGA
GGAGTCGGTCGCGCAGGTCCGGCTGCCCGACGGTGGGCGTGCGTGGGTGGTCAGCCGGCACGAGGACGTCCGCGCCGT
GCTCAACGACCGCCGGTTCAGCTCCGACCGCCGGCGACCGGACTTCCCGGAGCTGACGGCGGGCGCAGCCGCGGTCAAG
CGTCCCGACGAGGAGCCCACGCTGATCAGCATGGATGCGCCCGAACACCCGGTGGCCCGCCGAGCGGTGCTCGGCGAGT
TCACCGTGCGCCGCACCGAGGCGCTGCGGGCACGGGTCCAGGAGATCGTCGACGAGCGGATCGACGCGTTGCTGACCGG
GCCCCGGCCCGCCGACCTCGTCGAGGAACTGTCGCTGCCGATCCCGTCGCTCGTGATCTGCGAGCTGCTCGGCGTGCCCTA
CTCCGCGCACGCGTTCTTCCAGGAGAAGTCGACCGCGCTGATCGCCCGGGGGACGCCACCGCTGGAGCGAGCCGCCGCC
ATCGCAGCCCTGCGCGACTACCTGGACGAGTTGATCGCGGAGAAGGAGACCGACCCGCCGGACGATCTGCTCGGGCGGC
AGATCCACACCCTGCGCGAGCAGGGCTCCTACCGTCGGCCGGCCCTGGTGGGCATGGGAATCCTGCTGCTCGTGGCGGG
GCACGAGACGACGGCGAACATGATCTCGCTGTCGACGGTGGCGTTCCTGCGTGATCCCGAGCAGCTCGCGTTGATCCGCG
CCGATCCCACCAGGACCATCGCCGCGGTCGACGAGATGCTGCGGTACTTCACGATCGTGGACGCGGCGACGGCACGGCT
GTGCGTCGAGGACGCCGAGATCGGCGGGCAGCTGATCCGCGCGGGCGAGGGGGTGCTGGCGCTGACGTACTCGGCGAA
CCGCGATCCGCGGGCGTTCGCGGACCCCGACGAGCTGGACATCGAACGCGGCGCCCGCCACCACGTCGCTTTCGGGTTCG
GGCCACACCAGTGCCTGGGGCAGAACCTGGCCCGGATGGAGCTGCAGATCGTGTTCGACACCCTGTTCCGCCGCATCCCC
ACCCTCGCGCTCGCCGCCGACGTCGACGAGCTGCCCTTCAAGGACGACGCCGCCATCTTCGGCCTGCACCAGCTTCCGGTG
ACCTGGTGA

6. Acknowledgements

“I've battled demons that won't let me sleep
Called to the sea but she abandoned me
But I won't never give up, no”

Sia “Never give up” (2017)

With this thesis, I am not only finalizing my PhD, but also completing a chapter that truly had a great impact on my life. Looking back, I experienced a somehow “typical PhD” with all its ups and downs, which allowed me to grow scientifically and personally. I would like to dedicate the following chapter to all the people who have supported me along the way and have always believed in me, no matter what happened.

My special thanks go to my supervisor and first examiner Prof. **Vlada B. Urlacher**. Liebe **Vlada**, vielen Dank für dein Vertrauen, dass es mir ermöglichte mit großer Freiheit an spannenden und internationalen Projekten zu forschen. Meine Zeit in deiner Arbeitsgruppe war einzigartig, da sie nicht nur geprägt war von typischen Lehrtätigkeiten, sondern auch von außergewöhnlichen Ereignissen, wie der Organisation einer Konferenz und nicht zuletzt den Auswirkungen des Coronavirus. Gerade in diesen Momenten habe ich deine konstruktive und menschliche Art die Dinge anzugehen schätzen gelernt. Auch wenn wir in der einen oder anderen Diskussion nicht immer einer Meinung waren, so bin ich dir genau dafür dankbar, da dein Interesse und dein sachliches Feedback diese Arbeit, und somit auch mich persönlich, entscheidend vorangebracht haben. Ich danke dir für deinen Einsatz und werde die Zeit in deiner Gruppe stets in guter Erinnerung behalten!

I would also like to express my gratitude to my mentor and second examiner Prof. **Martina Pohl**. Liebe **Martina**, ich bin sehr froh, dass du dich nach meiner Masterarbeit in deiner Gruppe sofort als Mentorin meiner Doktorarbeit bereit erklärt hast. Gerade durch deine begeisterungsfähige und positive Art habe ich mich nach unseren Gesprächen immer besonders motiviert gefühlt und nochmals versucht, viele Dinge auf eine andere Weise zu sehen. Vielen Dank für Alles!

During my PhD I had the chance to collaborate with many people from all over the world. In this regard I want to thank all people from the '**HyPerIn**' project for the good work in advance and the great collaboration. I learned a lot and was truly lucky to be part of such a productive consortium. My special thanks goes also to Prof. **Matthias Bureik**, his group and all people I got to know during my research stays in Tianjin. I highly appreciate my time in China and learned a lot from you about studying and living in China (I am still dreaming of all the food we tasted).

I also want to thank all former and current members of the **Institute of Biochemistry II**. Thank you very much for the pleasant working atmosphere and the fruitful discussions about both scientific and non-scientific topics. I would like to thank the following people in particular:

My first and really deepest thank you goes to **Anne Worsch**. Liebe **Anne**, du warst nicht nur die Person, die mit mir zusammen am HyPerIn-Projekt gearbeitet hat, sondern auch diejenige, zu der ich immer bei jeglichem fachlichen Problem gehen konnte. Du bist eine wahre Powerfrau und ich habe meinen größten Respekt davor, wie du Familie und PhD so gut unter einen Hut bringen konntest. Danke für deinen Optimismus, die großartige Zusammenarbeit und Gespräche über alles- von „P450" bis „Gott und die Welt"!

Next I would like to thank the world best technician **Sebastian Hölzel**. Lieber **Sebastian**, wir haben fast zeitgleich angefangen und ich habe dich mit der Zeit immer mehr schätzen gelernt. Nicht nur, dass du der „gute Geist des Instituts" bist, sondern ich danke dir auch vor allem für deine lustige und unbeschwerte Art, die so manche schlechte Laune hat verfliegen lassen. Danke für unvergessliche Schalanderabende, Osteuropatrips, Fussballevents und vieles mehr!

This work would not have been possible without the help of five students namely **Judith Wamprecht**, **Jan Gebauer**, **Davide Decembrino**, **Leonie Windeln** and **Leonhard Kohleick**. I was privileged to supervise these five talented personalities who were not only scientifically good, but also identified with their projects in an outstanding way.

Liebe **Judith**, lieber **Jan**, auch wenn wir nur kurze Zeit zusammengearbeitet haben, so hat mir die Arbeit eures Praktikums doch wertvolle Erkenntnisse für die weitere Charakterisierung meiner P450 beschert. Vielen Dank für eure Hilfe und viel Erfolg bei euren künftigen (Forschungs-) Projekten!

Dear **Davide**, you came as an Erasmus student to our institute and stayed afterwards for your PhD. I could not expect that your project went that well and that you eventually became a good friend of mine. Thank you so much for your very accurate work which founded the basis for the first publication of my thesis. Grazie mille per i caffettini in combinazione con le conversazioni divertenti, le migliaia di birrette a "Rock am Ring", nell'Altstadt e in altri luoghi, ma soprattutto per la tua amicizia in ogni situazione!

Liebe **Leonie,** du hast in einer entscheidenden Phase meiner Arbeit dein Masterprojekt durchgeführt. Auch wenn im Laufe deines Projekts einige Stolpersteine auftauchten, so war ich doch immer wieder beeindruckt mit welcher Freude und großen Motivation du dich dem entgegengestellt hast. Gerade weil du immer daran geglaubt hast, konntest du entscheidend zu mindestens zwei meiner Projekte beitragen. Ich bedanke mich daher bei dir nicht nur für deine (zu Recht!) ausgezeichnete Masterarbeit, sondern auch für viele tolle Erlebnisse in Mosbach, bei Rock am Ring, diversen Biochemiefeiern oder bei anderen Gelegenheiten, von denen hoffentlich noch einige folgen werden!

Lieber **Leonhard**, du kamst gegen Ende meiner praktischen Arbeiten und hast mich bei den Umsetzungen für das „CYPome Projekt" unterstützt. Auch wenn manches auf Anhieb nicht so geklappt hat, so warst du doch stets sehr motiviert und hast sehr akkurat gearbeitet. Vielen Dank für deinen tollen Einsatz, wobei ich hoffe, dass du das ein oder andere von deinem Projekt für deine zukünftigen Arbeiten mitnehmen konntest!

I want to thank all other PhD students who shared with me the way during my time at the IBCII namely **Agathe Bronikowski**, **Anna Olbrich**, **Nikolas Ditz**, **Ansgar Bokel**, **Christopher Weber**, **Arsenij Kokorin**, **Alessandra Raffaele**, **Nina Jankowski** and **Stefan Wohlgemuth**. **Agathe**, danke für deine offene und herzliche Art! Die Miley Cyrus Version von „Jolene" wird immer einen besonderen Platz bei mir einnehmen! **Anna**, vielen Dank für zahlreiche Konzerte, Kaffeerunden, Sportevents, Kneipenabende und vieles mehr! Wir haben unsere Doktorarbeiten bald geschafft! **Niko**, danke für die genialen Rock am Ring Momente, Schalanderabende und einen unvergesslichen Trip nach Kolumbien. Natürlich „con leche"

und einer „Bandeja Paisa“, die uns nicht davon abhalten ließ den „Todesberg“ zu erklimmen. **Ansgar**, danke für die vielen hilfreichen Ratschläge rund um P450 und einzigartige Touren – egal ob im Schaukelstühlchen, in China, am Nürburgring oder sonst irgendwo. Gerade die Zeit in China mit Trips zur Chinesischen Mauer und Xi’an inklusive Paparazzi („it’s okay“) werde ich nie vergessen. **Chris**, vielen Dank für die vielen lustigen Diskussionen über die Maximierung des Mensaessens, Katzenvideos und vieles mehr. **Arsenij**, danke für deine herzliche und offene Sichtweise und besonders für die Hauspartys, natürlich mit echtem russischen Wodka. **Ale**, grazie per la tua cordialità e in particolare la tua aiuta con gli esperimenti per le ‘in vivo’ conversioni. Spero che vada in Sicilia presto per mangiare tutti il cibo bellissimo che hai raccontato. **Nina**, vielen Dank für deine ruhige und gewissenhafte Art und das Teilen von tollen Sport-und Urlaubsmomenten. **Stefan**, danke für deine Fürsorge im Labor und stets unterhaltsame Gespräche!

For practical advices, useful knowledge and interesting scientific discussions, I am grateful to Dr. **Sarah Kranz-Finger**, Dr. **Esther Ricklefs**, Dr. **Priska Le-Huu**, Dr. **Patrick Bakkes,** Dr. **Ansgar Rühlmann** and **Stefan Biemann**. Thank you so much also for pleasant coffee/tea breaks and nice barbecues! In this sense, I also want to acknowledge the work of my former supervisors of my Bachelor and Master projects – **Vlada**, **Osama**, **Dana**, **Marco**, **Alana** and **Martina**. Thanks for your effort and valuable advice! And also thanks to all those who took part in the proofreading of this thesis – **Alena**, **Alessandra**, **Ansgar**, **Florestan**, **Frauke**, **Leonie** and **Davide**.

I also want to acknowledge the people who made the ten years in Düsseldorf unforgettable. Ganz besonders danken möchte ich dabei **Jan Cox**. Vom ersten Semester bis zur Promotion haben wir im Studium, in der Fachschaft, in der GBM und sonst wo viel erlebt, sodass es den Rahmen sprengen würde, wenn ich das einzeln aufzählen würde. Danke für deine jahrelange Freundschaft und viele unterhaltsame Abende in der Altstadt oder sonst wo, bei dem wir natürlich immer nur ein Bier getrunken haben. ;)
Auch den vielen anderen (Wahl-)Düsseldorfern und anderen Nachwuchswissenschaftlern, die ich kennen und schätzen gelernt habe, möchte ich an dieser Stelle danke sagen. Besondere Grüße gehen dabei an alle Mitglieder von **Atlético Trichter** für geniale Festivals

(inklusive diverser Verurteilungen), **Gruppe 5** für eine schöne Studienzeit sowie der **juniorGBM Düsseldorf** für schöne Konferenzen und vieles mehr.

At the end, I want to acknowledge all the people from my home area in the Eifel. Mein ganz spezieller Dank geht dabei an **Niklas Spitzley**, **Michael Kaltenborn** und **Matthias Meyer** für unzählige Fifa Abende, bei dem ich zwar des Öfteren richtig vernichtet wurde, aber dafür immer Spaß hatte und alles andere vergessen konnte. Ferner danke ich euch sowie **Verena** und **Martin**, **Steffi** und **Jenny** sowie meinen Cousinen **Silke** und **Silvia** für viele schöne, unvergessliche Momente und eure Freundschaft über die ganzen Jahre. In diesem Sinne möchte ich auch meinen **Eltern**, **Peter** und **Irene**, und meinem Bruder **Achim** besonders danken. **Mama**, **Papa**, danke für all die Möglichkeiten, die ihr mir gegeben habt und vor allem für euer Verständnis in Zeiten, wo es mal nicht so lief. **Achim**, auch wenn wir uns manches Mal zusammen raffen mussten, so bin ich froh, dass ich immer auf dich zählen kann, egal wann und wo. Ich bin sehr froh, dass ihr drei immer für mich da seid. Ohne euch hätte ich das nicht geschafft!