HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Development of versatile biocatalysts based on P450 monooxygenases from the CYP154 family

Inaugural-Dissertation

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

vorgelegt von

Clemens-Jeremias von Bühler

aus Waiblingen

Düsseldorf, Mai 2014

Aus dem Institut für Biochemie
der Heinrich-Heine-Universität Düsseldorf

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

Referent:	Prof. Dr. Vlada B. Urlacher
Koreferent:	Apl. Prof. Dr. Martina Pohl
Tag der mündlichen Prüfung:	23.6.2014

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.

Düsseldorf, den 14. Mai 2014 Clemens-Jeremias von Bühler

Danksagungen

Frau Prof. Dr. Vlada Urlacher danke ich für die hervorragenden Arbeitsbedingungen am Institut und die Überlassung des interessanten Themas. Darüberhinaus möchte ich mich für die umfassende Betreuung meiner Arbeit, sowie alle hilfreichen Diskussionen und Hinweise und die Erstbegutachtung meiner Arbeit bedanken.

Frau Apl. Prof. Dr. Martina Pohl danke ich für die freundliche Übernahme des Koreferats.

Mein besonderer Dank gilt den Studenten, die zu dieser Arbeit beigetragen haben: Priska Le-Huu, Bodo Schmitz und Joshua Fehl.

Desweitern bedanke ich mich bei:

Dr. Astrid Höppner, Crystal and X-Ray Facility, HHU Düsseldorf für ihre Mitarbeit an der Kristallisation von CYP154E1 und ihren unermüdlichen Einsatz bei dem Versuch die Kristallstruktur zu lösen.

Dr. André Abts, Institut für Biochemie, HHU Düsseldorf für seine Hilfe bei den MALS-Messungen von CYP154E1.

Dr. Sabine Metzger, BMFZ, HHU Düsseldorf für ihre Hilfe bei den ESI-TOF-MS-Messungen von CYP154E1.

Dr. Helmut Keck und Dr. Peter Tommes für ihre Hilfe bei der Messung und Interpretation von HR-MS Spektren.

dem JGI für die Bereitstellung von genomischer DNA von *T. fusca.*

Dr. Matthias Gunne für das Korrekturlesen meiner Arbeit.

Allen weiteren Mitarbeitern des Instituts für Biochemie danke ich für die angenehme Arbeitsatmosphäre und fruchtbare Zusammenarbeit.

Meiner Frau Barbara und meinen Eltern bin ich ganz besonders für die Unterstützung während meiner Studienzeit, der Promotion und des damit verbundenen Umzugs nach Düsseldorf dankbar.

Contents

Zusammenfassung

Cytochrom P450 Monooxygenasen sind NAD(P)H-abhängige Hämoproteine, welche die reduktive Spaltung von molekularem Sauerstoff katalysieren. Dabei wird ein Sauerstoffatom zu Wasser reduziert und das andere auf das Substrat übertragen. Die dafür nötigen Reduktionsäquivalente werden von sogenannten Redoxpartnerproteinen auf die Monooxygenase übertragen. An der biokatalytischen Anwendung von P450 Monooxygenasen besteht ein ungebrochenes Interesse, da sie in der Lage sind selbst inerte, nicht-aktivierte Kohlenwasserstoff-Verbindungen zu oxidieren. Derzeit sind über 21000 P450 Gene klassifiziert. Um diesen Sequenzraum als Ressource für neue enzymatische Aktivitäten zu erschließen, sollten in dieser Arbeit neue P450 Enzyme charakterisiert werden. Die bakterielle CYP154-Familie wurde als potentielle Quelle für Biokatalysatoren mit breitem Substratspektrum und hoher Produktselektivität ausgewählt. Das biokatalytische Potential von zwei verschiedenen Familienmitgliedern, CYP154E1 aus *Thermobifida fusca* YX und CYP154A8 aus *Nocardia farcinica* IFM 10152, sollte ermittelt werden.

Um die Aktivität von P450 Enzymen rekonstituieren zu können, mussten zuerst geeignete Redoxpartnerproteine gefunden werden. Die Untersuchung von autologen Redoxpartnern führte zwar nicht zum Erfolg, die Aktivität beider P450 Enzyme konnte jedoch mit dem heterologen Putidaredoxin (Pdx) und der Putidaredoxinreduktase (PdR) aus *Pseudomonas putida* und noch effizienter mit dem Flavodoxin YkuN aus *Bacillus subtilis* und der *Escherichia coli* Flavodoxinreduktase (FdR) rekonstituiert werden.

Um das Substratspektrum und die Produktselektivität der beiden Monooxygenasen charakterisieren zu können, wurde eine neue Durchmusterungsmethode, genannt “Cluster-Screening”, entwickelt. Eine Bibliothek aus 51 organischen Molekülen, sortiert nach molekularer Masse und strukturellen Eigenschaften, wurde dafür zusammengestellt. Alle Verbindungen wurden *in vitro* mit beiden Monooxygenasen getestet und die gebildeten Produkte nach chromatographischer Trennung massenspektrometrisch detektiert. Diese Herangehensweise erlaubt die detaillierte Aufklärung der Regio- und Chemoselektivität der Enzyme. 30 Verbindungen der getesteten Bibliothek wurden von CYP154E1 und 23 von CYP154A8 umgesetzt. Darunter waren hauptsächlich lineare stäbchenförmige Moleküle, wie zum Beispiel Fettsäuren, Alkanole und azyklische Terpenoide. CYP154E1 zeigte generell einen höheren

Umsatz. CYP154A8 war im Allgemeinen weniger aktiv, jedoch meist deutlich selektiver und bildete weniger verschiedene Produkte als CYP154E1. Ein Produkt, das von beiden Enzymen gebildet wurde, war 25-Hydroxy-Grundmanns Keton (25-OH-GK), ein Synthesebaustein von Vitamin D_3-Derivaten.
Das ermittelte Substratspektrum wurde verwendet um eine neue Substratklasse, *n*-Alkane, für CYP154A8 vorherzusagen. Das Enzym akzeptierte tatsächlich *n*-Alkane einer Kettenlänge von C_7 bis C_{10}. In einem Zwei-Phasen-System konnte, unter Ausnutzung des *in situ* Extraktionseffekts der organischen Substratphase, eine Regioselektivität für die C2-Position von bis zu 90 % beobachtet werden. Ebenso wie die Regioselektivität war auch die Stereoselektivität kettenlängenabhängig. 2-(*S*)-Alkanole wurden daher mit 63-91 % *ee* gebildet. Die bestimmten TTN-Werte (total turnover number) von bis zu 4400 bestätigten, dass CYP154A8 mit den leistungsfähigsten und etablierten CYP102A1-Systemen konkurrieren kann. Im Gegensatz zu diesen hochgradig evolvierten Muteinen besitzt CYP154A8 als Wildtypenzym bereits eine höhere Regio- und Stereoselektivität als besagte CYP102A1-Varianten.
CYP154A8 und CYP154E1 hydroxylierten Grundmanns Keton (GK) absolut regioselektiv an der Position 25, CYP154E1 war dabei jedoch deutlich aktiver (100 % vs. 53 % Umsatz). 25-OH-GK ist eine wichtige Zwischenstufe für die Herstellung von 25-OH-Vitamin D_3-Derivaten. Auf Basis von CYP154E1 sollte eine biokatalytische Alternative zur ineffizienten, chemischen Hydroxylierung von GK entwickelt werden. *E. coli*-Zellen, die CYP154E1/Pdx/PdR exprimierten, wurden verwendet um Grundmann's Keton umzusetzen. Der Einfluss von Substratkonzentration und -löslichkeit, Produktinhibierung, Sauerstoffeintrag und das alternative Redoxpartnersystem YkuN/FdR wurde untersucht. Unter optimierten Bedingungen war die Herstellung von 1,1 mM (300 mg L^{-1}) 25-OH-GK möglich. Im Vergleich zur Fermentation von *Amycolata autotrophica* zur Produktion von 25-OH-Vitamin D_3 war die Endkonzentration von 25-OH-GK mindestens eine Größenordnung höher und wurde in nur einem Fünftel der Zeit erreicht. Die entwickelten Anwendungsbeispiele von Enzymen der CYP154-Familie werden in Zukunft hoffentlich den breiteren Einsatz dieser Enzyme in der Biokatalyse ermöglichen.
Um die beobachteten Selektivitäten von CYP154E1 rationalisieren zu können, wurde der Versuch unternommen das Enzym zu kristallisieren. Obwohl die erhaltenen Röntgenbeugungsdaten eine Lösung der Kristallstruktur noch nicht erlaubten, konnte mit der etablierten Kristallisation von CYP154E1 ein erster Schritt zur Bestimmung der 3D-Struktur gemacht werden.

Abstract

Cytochrome P450 monooxygenases (P450s) are NAD(P)H-dependent hemoproteins which catalyze the reductive scission of molecular oxygen leading to one molecule of water and the second oxygen atom being incorporated into a substrate molecule. For this process P450s require reducing equivalents which are transferred by so-called redox partner proteins. The interest in the biocatalytic application of P450s is high because they are able to transfer an oxygen atom on non-activated carbon atoms. To date more than 21000 P450 genes have been classified. To exploit this sequence space as a resource of new oxidation activities, in this study novel P450s should be characterized. The bacterial CYP154 family was selected as potential source of novel P450s with a broad substrate spectrum and a high product selectivity. The biocatalytic potential of two family members, CYP154E1 from *Thermobifida fusca* YX and CYP154A8 from *Nocardia farcinica* IFM 10152, was investigated.

To reconstitute P450 activity, suitable redox partners had to be identified first. The investigation of autologous redox partners from *T. fusca* was not successful but catalytic activity of both enzymes could be reconstituted with the heterologous putidaredoxin (Pdx) and putidaredoxin reductase (PdR) from *Pseudomonas putida*, and even more efficient with the flavodoxin YkuN from *Bacillus subtilis* in combination with *Escherichia coli* flavodoxin reductase (FdR).

In order to characterize the substrate spectra and the product selectivities of the reconstituted P450s a method called "cluster screening" was developed. A library of 51 organic molecules clustered into nine groups according to their chemical properties and size was established for screening. Subsequently, the products were separated by chromatography and detected by mass spectrometry. This approach allowed the detailed elucidation of enzyme chemo- and regioselectivity. From the library, 30 compounds were tested positive for CYP154E1 and 23 for CYP154A8. Mainly linear rod shaped molecules have been converted by both enzymes, e.g. fatty acids, alcohols, and acyclic terpenoids. CYP154E1 showed generally a higher degree of conversion whereas CYP154A8 displayed higher product selectivity.

The determined substrate spectrum was used to predict a novel substrate class of CYP154A8, namely *n*-alkanes. The enzyme accepted these substrates with a chain length from C_7 to C_{10}. Making use of the *in situ* product removal effect of a biphasic reaction system, maximal regioselectivity for the C2 position of up to 90 % was observed. The stereoselectivity of

CYP154A8 was chain length dependent and thus, 2-(*S*)-alkanols were formed with 63-91 % *ee*. The measured total turnover numbers of up to 4400 confirmed that CYP154A8 can compete with the best performing and well-established CYP102A1-based systems. But unlike those evolved CYP102A1 systems, CYP154A8 wild type shows already higher regio- and stereoselectivity.

Both enzymes catalyzed the hydroxylation of Grundmann's ketone at position 25. CYP154E1 and CYP154A8 showed absolute selectivity but conversion was higher with CYP154E1 (100 % vs. 53 %). 25-Hydroxy-Grundmann's ketone is an important synthetic intermediate for preparation of 25-hydroxyvitamin D_3 analogs. Based on CYP154E1 a biocatalytic alternative to the inefficient chemical synthesis of 25-OH-GK was investigated. *E. coli* cells expressing CYP154E1/Pdx/PdR were used to convert Grundmann's ketone. The influence of various process parameters like substrate concentration and solubility, product inhibition, oxygen supply, and redox partner systems were investigated. By combining all optimized parameters 1.1 mM (300 mg L^{-1}) 25-hydroxy-Grundmann's ketone were formed by the CYP154E1/Pdx/PdR system in 24 h. Compared to the fermentation of e.g. *Amycolata autotrophica* for the production of 25-hydroxyvitamin D_3 the resulting concentration of 25-hydroxy-Grundmann's ketone was at least one order of magnitude higher and was reached five times faster. The developed synthetic applications of CYP154-family enzymes and their detailed characterization will hopefully allow the broader application of these enzymes in biocatalysis in the future.

In order to rationalize the observed activities and selectivities of CYP154E1 the enzyme was crystallized. Although the obtained X-ray diffraction data have not yet allowed the solution of the crystal structure, with the established crystallization of CYP154E1 a first step to determine its three-dimensional structure was made.

1 Introduction

1.1 Biotechnology and biocatalysis

The application of enzymes and microbes in all day life dates back to ancient time, when yeasts were used to ferment food and beverages. Until today microorganisms are used for the production of tofu, cheese or koji as well as wine and beer. Only in the beginning of the last century scientists realized that they could make use of components of living cells for chemical transformations, a first step away from fermentation processes in the direction of enzyme catalysis. This transition is called "the first wave of biocatalysis".[1] In 1908 Rosenthaler for the first time used a crude enzyme extract of almonds containing a hydroxynitrile lyases for the stereoselective transformation of benzaldehyde into *(R)*-mandelonitrile.[2] Examples of recent date are the production of inverted sugar syrup with invertase and the semisynthetic production of β-lactam antibiotics with the help of an acylase.[3]

"The second wave of biocatalysis" (1980s and 1990s) is characterized by a deeper understanding of protein's three-dimensional structures. This in turn enabled structure-based engineering technologies for broadening the substrate range of enzymes what made non-natural biotransformations accessible. This was a prerequisite for using enzymes in the production of pharmaceuticals and fine chemicals. One example is the production of wax esters for cosmetics with the help of lipases.[4]

"The third wave of biocatalysis" started in the late 1990s when molecular biology methods were used to optimize enzymes for biocatalysis. This development was enabled by advances in gene sequencing, gene synthesis and bioinformatics methods. The so-called directed evolution, an iterative process of random mutagenesis and subsequent high-through-put screening for beneficial mutations, allowed the design of utterly new biocatalysts.[5]

The driving forces of the increased implementation of biotechnological processes are of economical, regulatory and scientific nature. Increasing costs for energy, waste disposal and raw materials motivate chemical companies to invest in the development of "greener" processes. Additionally, the increasingly strict environmental regulations force companies to

avoid processes involving harmful agents or producing high amounts of waste, byproducts, and emissions. Finally, the well-known and tragic example of the drug thalidomide, which was administered as a racemate and whose (*S*)-enantiomer had a strong teratogenic effect and caused malformation of thousands of newborns, led to a change in regulations by the US Food and Drug Administration (FDA). The development of novel pharmaceuticals as racemates was strongly disfavored which gave rise to a need of pure chiral pharmaceuticals.[6]

Enzymes possess a variety of unique features that make them promising candidates to achieve these goals in process development:[7,8] The reaction rates of enzyme catalyzed processes are typically 10^8–10^{10} times faster than uncatalyzed reactions.[9] This is much higher than what can be achieved with chemical catalysts and leads to lower amounts of catalyst needed to perform a bioconversion.

Enzymes work under mild conditions. Generally, the pH optimum of enzymes is between pH 5 and 8. Moreover, most enzymes originate from organisms growing between 20 and 40 °C and thus their optimal reaction temperatures are correspondingly mild. Compared to chemical catalysts this is rather low and helps to reduce energy consumption and avoids side reactions like isomerization and decomposition of the substrates. Most enzymes work in aqueous or buffered solutions which greatly reduces the amount of environmentally hazardous organic solvents needed in a biocatalytic process.

In general, enzymes catalyze a broad spectrum of reactions and generally display a high degree of chemo-, regio-, and stereoselectivity which makes them promising candidates for selective production of optically pure compounds. The reduced production of byproducts increases atom economy which is beneficial for the development of "greener" processes. Furthermore, the fact that enzymes are generally compatible with each other allows the design of artificial multi-enzyme system forming whole pathways *in vivo* or cascade reactions *in vitro*. This is especially interesting for the multi-step production of complex natural products by a combination of specialized enzymes as well as for establishing bio-based feedstocks for a sustainable chemical industry.

An analysis of enzymes used in established biotechnological processes gives first hints about the problems that are associated with the application of biocatalysts (Fig. 1).[10] The most often used enzymes in biocatalysis are hydrolases (EC 3.X). These enzymes are often commercially available, do not rely on cofactors, and generally demonstrate high robustness and stability.

Low enzyme stability and the need for expensive cofactors in stoichiometric amounts often hinder the application of oxidoreductases and especially that of P450s (Fig. 1). For details how these problems can be solved see section 1.2.8.

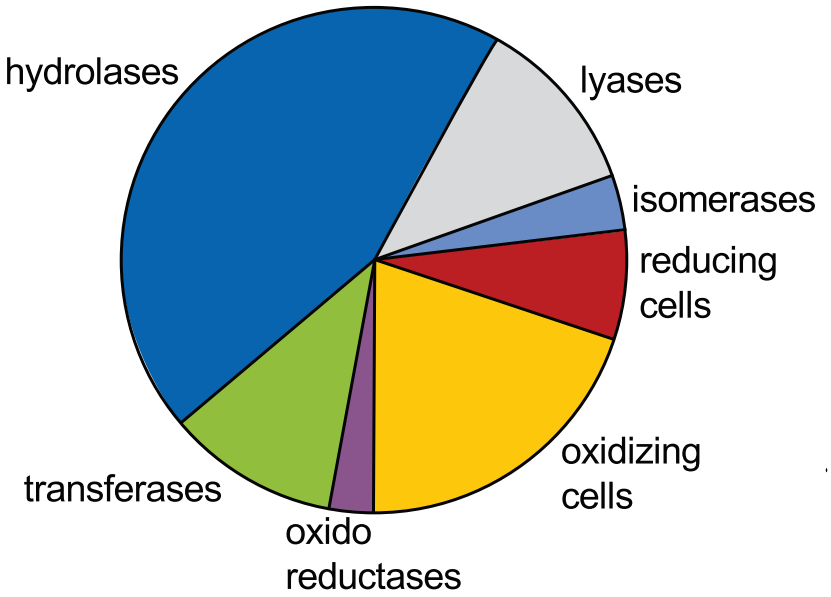


Figure 1: Enzyme types used in industrial biotransformations (based on 134 processes). Data taken from Straathof *et al.*[10]

Moreover, nature only provides one enantiomer of an enzyme. If it doesn't exhibit a desired stereoselectivity it is not possible to use the other stereoisomer of the catalyst. Hence, engineering of the catalyst is the only strategy to invert enzyme selectivity. This is often an uncertain and laborious task whose success is not guaranteed. Although enzymes in general display high substrate promiscuity, an enzyme catalyzing a specific transformation of a desired substrate does not necessarily exist. The same holds true for catalytic promiscuity which is high for enzymes in general, but for many chemical reactions appropriate enzymes are not known. Hence, there is a constant demand for the development of novel and advanced biocatalysts.

1.2 Cytochrome P450 monooxygenases

Cytochrome P450 monooxygenases (P450s) are hemoproteins belonging to the class of oxidoreductases (EC 1.14.14.1). They catalyze the reductive scission of molecular oxygen leading to one molecule of water while the second oxygen atom is incorporated into a substrate molecule. For this process P450s require the consecutive delivery of two reducing equivalents originating from nicotinamide cofactors. Most P450s are not able to accept reducing equivalents directly from NAD(P)H. Separate redox partner proteins are needed to transfer electrons to the heme iron. The name P450 was coined because of their characteristic

absorbance at 450 nm in ferrous CO-bound form that led to their discovery. P450s are ubiquitous enzymes found in all kingdoms of life. "The Cytochrome P450 homepage" (last update August 2013) lists more than 21000 classified P450 genes.[11] Even more P450 sequences are listed in databases but have not been classified yet.[12] The oxidative potential of P450s is utilized by their hosts for various purposes and ranges from assimilation of carbon sources to the formation of secondary metabolites or the detoxification of xenobiotics.

1.2.1 Occurrence and natural function of P450s

For the first time P450s have been observed in mammalian liver microsomes. The recently completed human genome project confirmed the existence of 57 putatively functional genes and 58 pseudogenes coding for P450s in the human genome. Among others they play a crucial role in the detoxification of xenobiotics, especially in phase I of the drug metabolism. During phase I of drug metabolism lipophilic substances become more water soluble upon enzymatic modification mediated, amongst others, by P450 monooxygenases. Subsequently, they become accessible to phase II enzymes responsible for the conjugation of the drug molecules to charged residues like glutathione (glutathione *S*-transferase) or acetyl moieties (*N*-acetyl transferase). In the human liver CYP3A4 is the major P450 which takes part in the oxidation of many drug molecules, e.g. erythromycin, midazolam, tamoxifen, amitriptyline and many others. CYP3A4 is induced by more than 50 % of all clinically used drugs.[13]
Furthermore many human P450s are involved in the biosynthesis and breakdown of steroid hormones.[14] The lack of such an enzyme due to genetic defects can have drastic effects. For instance, newborns lacking the cholesterol metabolizing CYP11A1 cannot survive without hormone replacement therapy.[15]
With 7446 gene sequences the number of classified P450 enzymes in plants exceeds those from any other kingdom of life. This diversity of P450 monooxygenases in plants demonstrates their importance for survival of these organisms. The genome of the well-studied organism *Arabidopsis thaliana* alone contains 273 CYP genes. P450s catalyze important steps in the biosynthesis of many plant secondary metabolites which are needed for plant defense, their communication with the environment, and also for the construction of physical barriers.[16]
While historically P450s were first discovered in liver microsomes of mammals, prokaryotic P450s and especially those of bacterial origin have contributed to a deeper understanding of the structure and the catalytic mechanism of P450s. CYP101A1 ($P450_{cam}$) from *Pseudomonas putida* was the first P450 which could be produced in amounts high enough to be

purified and crystallized. CYP101A1 mediates the hydroxylation of D-camphor to 5-exo-hydroxycamphor. Hence, *P. putida* is able to use camphor as a carbon source.[17] Apart from CYP101A1 there are also other P450s which have been shown to be involved in the utilization of organic molecules as carbon source, e.g. $P450_{terp}$ (α-terpineol) and $P450_{lin}$ (linalool).[18,19] A further important bacterial P450 involved in carbon source assimilation is CYP102A1 ($P450_{BM-3}$) from *Bacillus megaterium* that catalyzes the subterminal hydroxylation of fatty acids.[20]

1.2.2 Nomenclature of P450s

"Cytochrome" is a collective term for proteins containing a prosthetic heme group. Based on their spectroscopic properties which depend on the heme-type and the additional axial ligand of the heme iron center these enzymes can be classified in several groups. Beyond that, their secondary and tertiary structure are markedly different (see section 1.2.3).
Cytochrome P450s (CYPs) contain heme b with an axial thiolate ligand which is responsible for their characteristic absorbance at 450 nm. Besides their trivial names like $P450_{BM-3}$ or $P450_{cam}$ a nomenclature based on amino acid similarity of P450 proteins has been established. Protein sequences with a sequence identity > 40 % are classified to the same P450 family. These families are then divided into subfamilies for all proteins exhibiting > 55 % sequence identity. The identification code of P450s, e.g. CYP102A1 for $P450_{BM-3}$, is composed of the abbreviation CYP for cytochrome P450 followed by an Arabic numeral (102) denoting the sequence family and a letter coding the subfamily (A). The final Arabic numeral is sequentially increased with every new protein in a subfamily. Families are clustered according to the origin of the genes. CYP families 1-9 originate from mammals (including humans), families 10-69, 501-699 as well as 5001 and higher originate from animals, fungi, and other lower eukaryotes. P450s from plants are clustered in families 71-99 and 701-772. CYP families 101-299 and 1001-1047 contain the members of bacterial origin.

1.2.3 The P450 fold

Sequence identity within a P450 family is more than 40 % but this value can drop below 20 % when proteins from two different families are compared.[21,22] Nevertheless, all P450 enzyme have a common structural core and a conserved so-called P450 fold (Fig. 2). This consists of a total of 13 α-helices and 4 β-sheets. All α-helices are named with uppercase letters in the sequence of their occurrence in the polypeptide chain from N- to C-terminus. The β-sheets

are numbered in a similar way with Arabic numerals followed by a second Arabic numeral after a dash, e.g. 1-2, denoting the strand of a certain β-sheet. In the P450 fold two domains can be distinguished. The α-domain consists of the large β-sheet β1 and the smaller β2 as well as the helices αA, αB and αK'. Domain β forms the major part of the protein's structure and consists of mainly α-helices and the two smaller β-sheets 3 and 4. The presence of a fifth β-sheet is variable. In domain β also the heme coordinating core consisting of the six α helices αD, αE, αI, αL, αJ and αK is located. Among the so far known sequences there seems to exist only three absolutely conserved amino acids. Firstly, these are the glutamic acid and the arginine which form the ExxR. It is located in helix αK on the proximal side of the heme plane and is responsible for the stabilization of the core structure. The third absolutely conserved residue is the heme-complexing cysteine that serves as the axial, fifth ligand. This cysteine is located in a β-bulge called the Cys-pocket which envelops the thiolate ligand in a hydrophobic environment. It is formed by three highly conserved amino acids, two glycine and one phenylalanine residue. In theoretical calculations it could be shown that the redox potential of the heme is dependent on the hydrophobicity of its surrounding.[23] Thus, besides shielding the iron-sulfur bond from reduction, these residues could be responsible for tuning the redox potential of the P450.
A further structural element found in all P450s is the so-called meander loop between αK and the Cys-pocket. Even though it does not have a defined secondary structure a salt-bridge interaction of an ERR motive results in a highly conserved three-dimensional structure. Supposedly, it contributes to the stabilization of the heme group.

Due to the very high diversity of the accepted substrates between P450 enzymes, apart from all conserved regions there have to be also variable regions which account for these differences in selectivity. These have been identified by comparison between the CYP2 family and CYP101A1 and are so-called substrate recognition sites (SRS).[24] SRS-1 is located in the highly variable loop between αB and αC (BC-loop) whereas SRS-2 is part of the αF helix. SRS-3 and SRS-4 belong to helix αG and αI, respectively. SRS-5 lies in the β1-4 and SRS-6 in the β4-1 strand. While SRS-1 to SRS-3 are forming part of the substrate entrance channel, SRS-5 and SRS-6 protrude into the active site near the heme. In a study of 6300 sequences especially SRS-5 and one adjacent residue of the BC-loop could be identified to be an essential determinant of regio- and stereoselectivity.[25]
A further structural difference of P450s is the site for interaction with the redox partner proteins. The reductase interaction sites (RIS) of several CYPs were analyzed and differences in the sites αJ/J' (RIS1) and the insertion between meander loop and Cys-pocket (RIS2) were identified for different P450 classes.[26] The authors were able to show that cytochrome P450

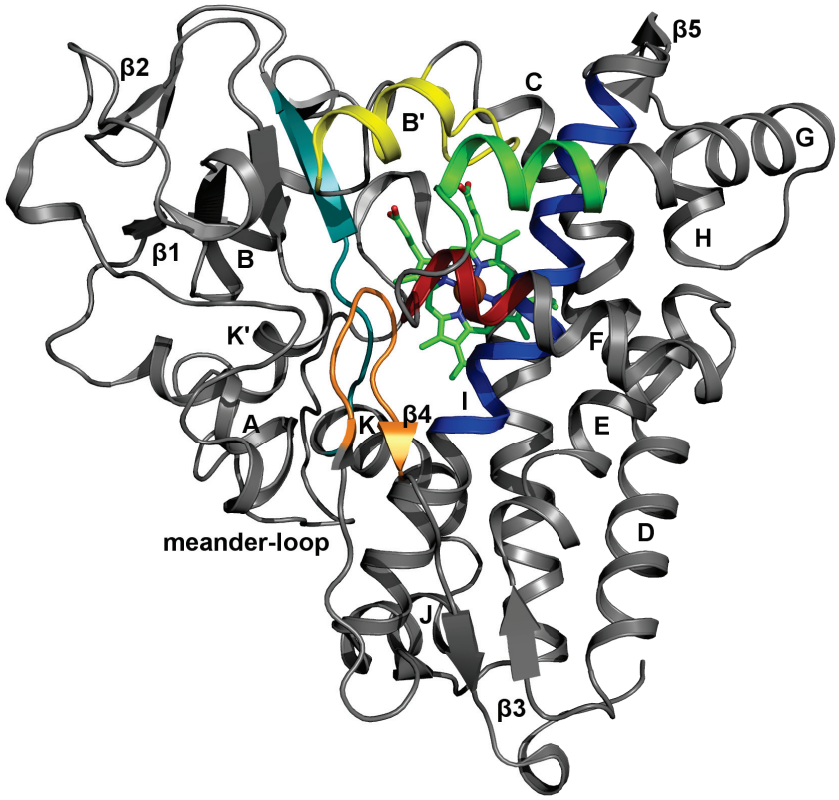


Figure 2: Representation of the crystal structure of CYP101A1 ($P450_{cam}$, PDB-ID: 2ZWT, chain A) with the α-helices and β-sheets labeled. Colored residues in cartoon representation belong to SRS1 (yellow), SRS2 (red), SRS3 (green), SRS4 (blue), SRS5 (turquoise), or SRS6 (orange). The prosthetic heme group is depicted in stick representation.

reductase-dependent enzymes (for details see section 1.2.4) possess a long RIS1 compared to class I and III P450 whose RIS1 is short or non-existent. The same correlation was observed for RIS2 which was short in class I and III enzymes (3-5 residues) and markedly longer (11-17 residues) for CPR-dependent enzymes.

1.2.4 Redox partner systems

A prerequisite for the catalytic reaction of CYPs are two single electron reduction steps. The necessary reducing equivalents have to be supplied via redox partner proteins and originate in the majority of all cases from the cofactors nicotinamide adenine dinucleotide phosphate (NADPH) or nicotinamide adenine dinucleotide (NADH). During evolution several different redox partner systems have evolved. According to the redox partners ten different P450 classes can be distinguished (Fig. 3), (Tab. 1).[27]

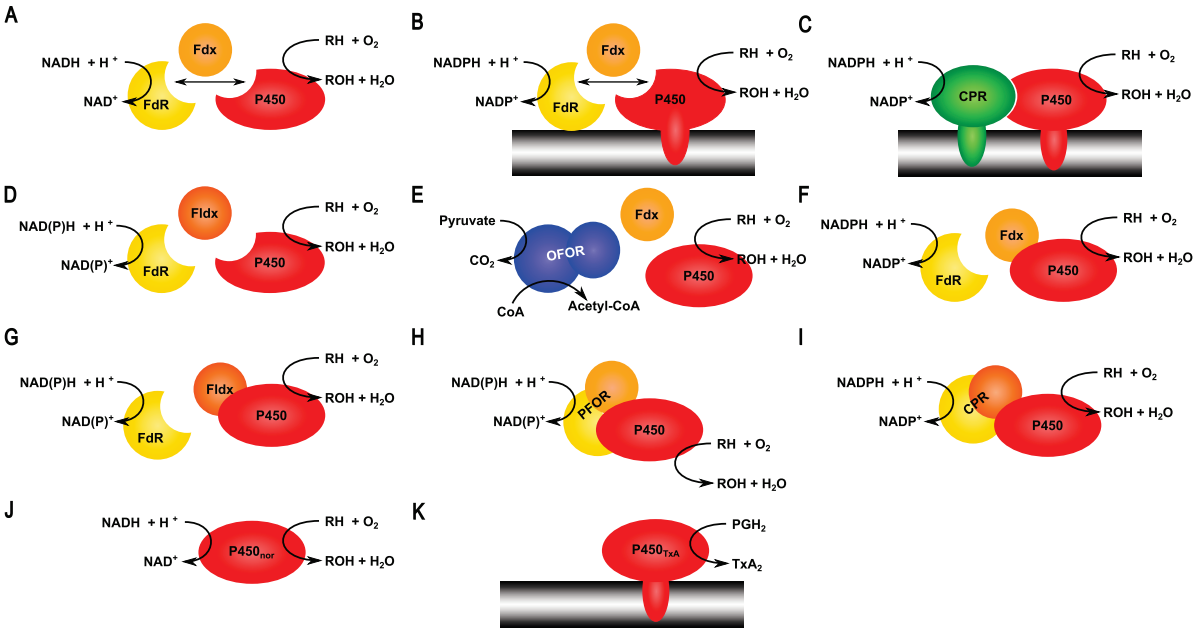


Figure 3: Schematic organization of different cytochrome P450 systems. **(A)** class I, bacterial system; **(B)** class I, mitochondrial system; **(C)** class II microsomal system; **(D)** class III, bacterial system; example P450cin; **(E)** class IV, bacterial thermophilic system; **(F)** class V, bacterial [Fdx]–[P450] fusion system; **(G)** class VI, bacterial [Fldx]–[P450] fusion system; **(H)** class VII, bacterial [PFOR]–[P450] fusion system; **(I)** class VIII, bacterial [CPR]–[P450] fusion system; **(J)** class IX, soluble eukaryotic P450nor; **(K)** independent eukaryotic system, example P450TxA.
Fdx, ferredoxin (iron-sulfur-cluster); FdR, ferredoxin reductase (FAD); CPR, cytochrome P450 reductase (FAD, FMN); Fldx, flavodoxin (FMN); OFOR, 2-oxoacid:ferredoxin oxidoreductase (thiamine pyrophosphate, [4Fe–4S]-cluster); PFOR, phthalate-family oxygenase reductase (FMN, [2Fe-2S]-cluster); prostaglandin H_2, PGH_2; TxA_2, thromboxane A2. Adapted from Hannemann *et al.* and reprinted with permission from Elsevier.[27]

Class I systems (Fig. 3 A and B) were among the first ones to be discovered in the adrenal mitochondria. Here a P450 is supplied with electrons transferred from a flavin adenine dinucleotide (FAD)-containing NADPH-reductase by a small iron-sulfur protein, a so-called ferredoxin. Both, P450 and reductase, are membrane bound or associated whereas the ferredoxin is soluble in the mitochondrial matrix of mammals. In bacteria class I systems exist as well, even though they are phylogenetically not related to the mammalian ones. Moreover, in bacterial class I systems all components are soluble in the cytosol. The prototype of a bacterial class I system is CYP101A1 with the NADH-dependent putidaredoxin reductase (PdR) and putidaredoxin (Pdx) which is a [2Fe-2S]-type ferredoxin. Both Pdx and PdR can easily be expressed in *E. coli*. It has been shown that they interact with a broad spectrum of P450s and therefore are often used to support a P450 if the physiological redox partners of a CYP are not known.[28,29]

Class II (Fig. 3 C) systems are most commonly found in the endoplasmic reticulum (ER) of eukaryotic organisms. The system consists of a P450 and a NADPH-dependent cytochrome P450 reductase (CPR). In mammals both are integral membrane proteins. The CPR has evolved from two ancestral genes. The N-terminal domain is homologous to a bacterial flavin mononucleotide (FMN)-containing flavodoxin and the C-terminal part resembles $NADP^+$-reductases and NADH-cytochrome *b5* reductases.[30,31] In two further variations of class II systems the later one as well as cytochrome *b5* itself can take part in the electron transfer chain of microsomal P450 systems. The only prokaryotic system known in class II is CYP105A3 ($P450_{sca}$) from *Streptomyces carbophilus* which catalyzes the oxidation of mevastatin to pravastatin.[32]

Class III P450 systems (Fig. 3 D) also consist of three components.But unlike class I systems the redox partner protein relaying the reducing equivalents to the P450 is a FMN-containing flavodoxin which is reduced by an FAD-containing flavodoxin reductase. Concerning the cofactors, class III (FAD ▸ FMN) resembles more the microsomal CPRs (FAD ▸ FMN) than the other bacterial class I system (FAD ▸ FeS). Besides the prototypical CYP176A1 ($P450_{cin}$) with its flavodoxin cindoxin and a corresponding reductase there is also evidence that CYP107H (P450BioI) from *Bacillus subtilis* belongs to class III. It could be shown that it interacts with the flavodoxins YkuN and YkuP from the same organism. Since the physiological reductase of these flavodoxins has not been identified yet, the *E. coli* flavodoxin reductase FdR has been used as a substitution.[33] Despite its interaction with YkuN and YkuP the definite classification of $P450_{BioI}$ is difficult since reduction by a *B. subtilis* ferredoxin has been observed as well.[34]

The well-studied P450 CYP102A1 ($P450_{BM-3}$) is a member of class VIII (Fig. 3 I). Since the P450 domain and a CPR-like domain are fused together within the same polypeptide chain this P450 does not rely on other redox partner and thus is termed catalytically self-sufficient. CYP102A1-catalyzed oxidation of fatty acids occurs with turnover frequencies of >17000 min^{-1}. It belongs to the fastest P450s known to date.[35] Monooxygenases of this type have been identified in many other prokaryotes but are also present in lower eukaryotes, e.g. $P450_{foxy}$ (CYP505) of the fungus *Fusarium oxysporum.*[36]

In the literature several different topologies of self-sufficient P450s have been reported: Class VIII systems consists of a P450 fused with reductases homologous to CPRs, e.g. CYP102A1. Electron flow within these enzymes involves two flavin cofactors. In class VII enzymes, e.g. CYP116B1 from *Rhodococcus sp.*, the reductase domain is homologous to phthalate dioxygenase reductase. This consists of a ferredoxin-like ([2Fe-2S]-cluster) and an FMN-containing domain which are fused to the P450 domain and supply electrons abstracted from the cofac-

Table 1: P450 systems classified depending on the topology of the protein components involved in the electron transfer to the P450 enzyme. Adapted from Hannemann *et al.* and reprinted with permission from Elsevier. [27]

Class/Source	Electron Transport Chain	Localization / Remarks
Class I		
Bacteria	NAD(P)H ▸ [FdR]▸[Fdx]▸[P450]	Cytosolic, soluble
Mitochondria	NADPH▸[FdR]▸[Fdx]▸[P450]	P450: inner mitochondrial membrane, FdR: membrane associated, Fdx: soluble
Class II		
Bacteria	NADH▸[CPR]▸[P450]	Cytosolic, soluble
Microsomes A	NADPH▸[CPR]▸[P450]	Membrane anchored, ER
Microsomes B	NADPH▸[CPR]▸[cytb5]▸[P450]	Membrane anchored, ER
Microsomes C	NADH▸[cytb5Red]▸[cytb5]▸[P450]	Membrane anchored, ER
Class III		
Bacteria	NAD(P)H▸[FdR]▸[Fldx]▸[P450]	Cytosolic, soluble
Class IV		
Bacteria	Pyruvate, CoA▸[OFOR]▸[Fdx]▸[P450]	Cytosolic, soluble
Class V		
Bacteria	NADH▸[FdR]▸[Fdx–P450]	Cytosolic, soluble
Class VI		
Bacteria	NAD(P)H▸[FdR]▸[Fldx–P450]	Cytosolic, soluble
Class VII		
Bacteria	NADH▸[PFOR–P450]	Cytosolic, soluble
Class VIII		
Bacteria, fungi	NADPH▸[CPR–P450]	Cytosolic, soluble
Class IX		
Fungi	NADH▸[P450]	Cytosolic, soluble
Class X		
Plants and mammals	[P450]	Cofactor independent, membrane bound, ER

Fdx, ferredoxin (iron–sulfur-cluster); FdR, ferredoxin reductase (FAD); CPR, cytochrome P450 reductase (FAD, FMN); Fldx, flavodoxin (FMN); OFOR, 2-oxoacid:ferredoxin oxidoreductase (thiamine pyrophosphate, [4Fe–4S] cluster); PFOR, phthalate-family oxygenase reductase (FMN, [2Fe-2S] cluster); cytb5, cytochrome b5; cytb5Red, cytochrome b5 reductase; ER, endoplasmic reticulum

tor.

1.2.5 P450 catalytic mechanism

The catalytic cycle of P450s was controversially discussed over the last two decades. Especially the mechanism of oxygen activation and the exact nature of the catalytically active species were difficult to elucidate due to the very unstable intermediates involved. Only in the year 2010 by the work of Green *et al.* the last and most important intermediate in the cycle, Compound I, could be characterized.[37]

As depicted in Fig. 4 the resting state of the enzyme is a ferric species with a sixth axial water ligand **(1)**. Upon binding of a substrate molecule the hydrogen bonding network near the heme iron is destabilized and the heme iron's water ligand is displaced. As a consequence, the heme iron is left in a five-coordinated, out-of-plane, high-spin coordination **(2)**. The spin-state shift of the heme iron is accompanied by a shift in its optical absorption spectrum from 415-417 nm for the low-spin (LS) state to 390-394 nm for the high-spin (HS) state.[38] This effect can be used to record binding spectra. In a typical type I spectrum a minimum at 420 nm and a maximum at 390 nm is observed when substrate free P450 is titrated with a real substrate.[39,40] The difference between LS and HS absorbance observed at different substrate concentrations is used to calculate binding constants (K_D), a measure for the affinity of a P450 towards a certain substrate.[41] Reverse type I spectra are observed when the initial preparation of P450 contains a significant amount of HS state species which is eliminated or reduced by the binding of the test compound. In case of the formation of six-coordinated complexes with strong ligands or inhibitors type II spectra are observed with an absorbance maximum at 425-440 nm and a minimum at 415 nm or lower. A further consequence of the LS-HS-transition is the shift of the midpoint potential of the heme iron to more positive values which enables the first reduction step. A second factor influencing the redox potential of the Fe^{II}/Fe^{III}-couple are conformational changes upon binding of the redox partner as observed for Pdx and CYP101A1 as well as adrenodoxin (Adx) and CYP11A1.[42]

Due to the higher redox potential of the substrate-bound P450 a single reduction of the heme iron center leading to a ferrous heme complex with a characteristic absorbance at 408-410 nm **(3)** becomes possible. The ferric species is a prerequisite for the binding of dioxygen **(4)**. This complex can be described either as ferrous dioxygen, because it is EPR silent like Fe^{2+}-OO-complexes, or as a ferric superoxide which is supported by Mössbauer experiments.[43] Recent studies support the view of a mixed character of the iron center with only partial electron density transfer between iron and oxygen.[44] The dioxygen bound complex **(4)** is subse-

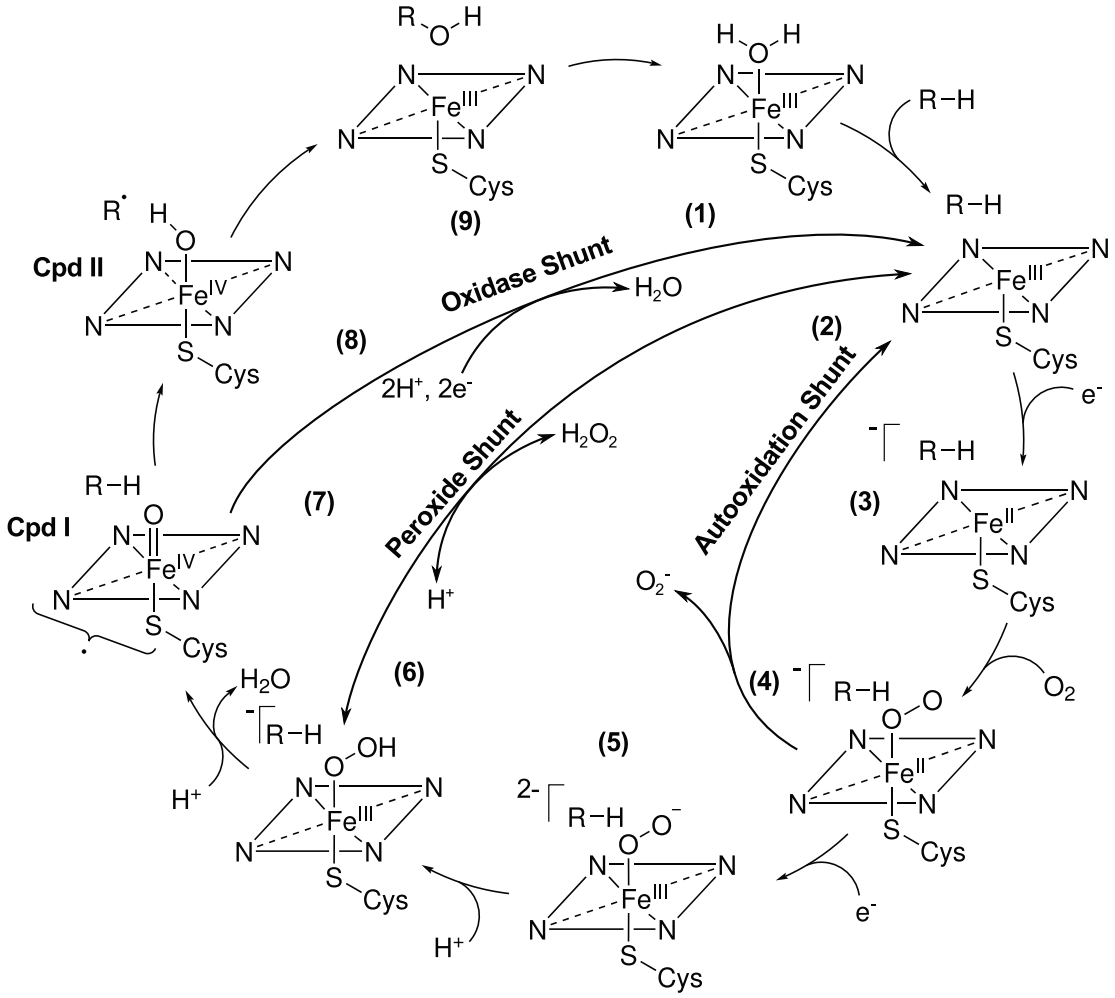


Figure 4: Catalytic cycle of P450 monooxygenases: **(1)** Ferric resting state (low-spin). **(2)** The axial water ligand is replaced by a substrate (R-H) inducing structural changes that lead to a ferric high-spin species **(3)**. The subsequent one electron reduction is the required for the binding of molecular oxygen **(4)**. A second single electron reduction leads to a ferric peroxo anion **(5)**, which then is protonated yielding a hydroperoxo species **(6)**. The following second protonation and water elimination form the catalytically active compound I **(Cpd I**, 7**)**. It abstracts a hydrogen atom from the substrate leading to a substrate radical and compound II **(Cpd II**, **8)**. The substrate radical is rebound to the hydroxyl group leaving the active site as hydroxylated product **(9)**. The heme returns to its ferric resting state **(1)**.

quently further reduced and a ferrous peroxo complex is formed **(5)** which forms by proton abstraction the hydroperoxo complex **(6)**. The ferric hydroperoxo complex **(6)** then undergoes water elimination and thereby the O-O bond is cleaved. The resulting complex **(7)** is the catalytically active species termed compound I. It is best described as an oxo-(Fe^{IV}) species electronically coupled to a ligand-based radical.[37] Its high reactivity enables compound I to oxidize even notoriously inert C-H bonds. Detailed characterization of this intermedi-

ate became possible after its synthesis by treating CYP119 with *m*-chloroperbenzoic acid.[37] Furthermore, Green *et al.* observed a kinetic isotope effect (KIE) supporting the rebound mechanism of compound I **(7-8)**. It was postulated previously that compound I abstracts a hydrogen radical from the substrate forming an Fe^{IV} hydroxide, compound II, **(8)** which is rapidly rebound to the substrate radical to yield hydroxylated product **(9)** and the P450 in its ferric resting state **(1)**.[45]
Besides the productive catalytic cycle there are several bypasses, so-called shunt reactions, which uncouple cofactor consumption from product formation. Instead, reactive oxygen species (ROX) are formed which in general have a destabilizing effect on proteins and the heme cofactor since they are potent oxidants themselves. On a physiological level, especially in eukaryotes, formation of ROX by P450s contributes at least partially, to the toxicity of xenobiotic substances.[44,46]

1.2.6 Catalyzed reactions

The high reactivity of the compound I intermediate enables P450s to oxidize even non-activated hydrocarbons resulting in the hydroxylation of those organic molecules. Carbon atom hydroxylation is the classical P450 reaction (Fig. 5 A, C, E). However, transformations mediated by P450s are not limited to this single type of reaction but comprise epoxidation of carbon-carbon double bonds (Fig. 5 B) as well as aromatic epoxidation (Fig. 5 D), *O*- and *N*-dealkylation (Fig. 5 F, H), dehalogenation (Fig. 5 I), formation of C-C-bonds in phenols (Fig. 5 J), and heteroatom oxidation (Fig. 5 K, G).[47,48] Most of the aforementioned reaction types can be rationalized as interaction of compound I with the substrate. Unusual P450 reactions often occur if the initial oxidation product is instable and undergoes a rearrangement reaction or if a second P450-catalyzed reaction takes place immediately.[49]
Examples for the hydroxylation activity of P450s both in eukaryotic and prokaryotic hosts are the oxidation of fatty acids by CYP4 family enzymes and by CYP102.[50,51] Bacteria exploit hydroxylation activity to access additional carbon sources which would be otherwise not accessible for their metabolism. This was firstly described for *Pseudomonas putida* which grows on (+)-camphor as sole carbon source.[17]

All processes described so far only deal with degradation processes in which P450s are involved. However, P450 also take part in many biosynthetic pathways which is the reason for the multitude of P450 genes in plant genomes. The biosynthesis of the promising anti malaria drug artemisinin in the plant *Artemisia annua* is a prominent example for P450s

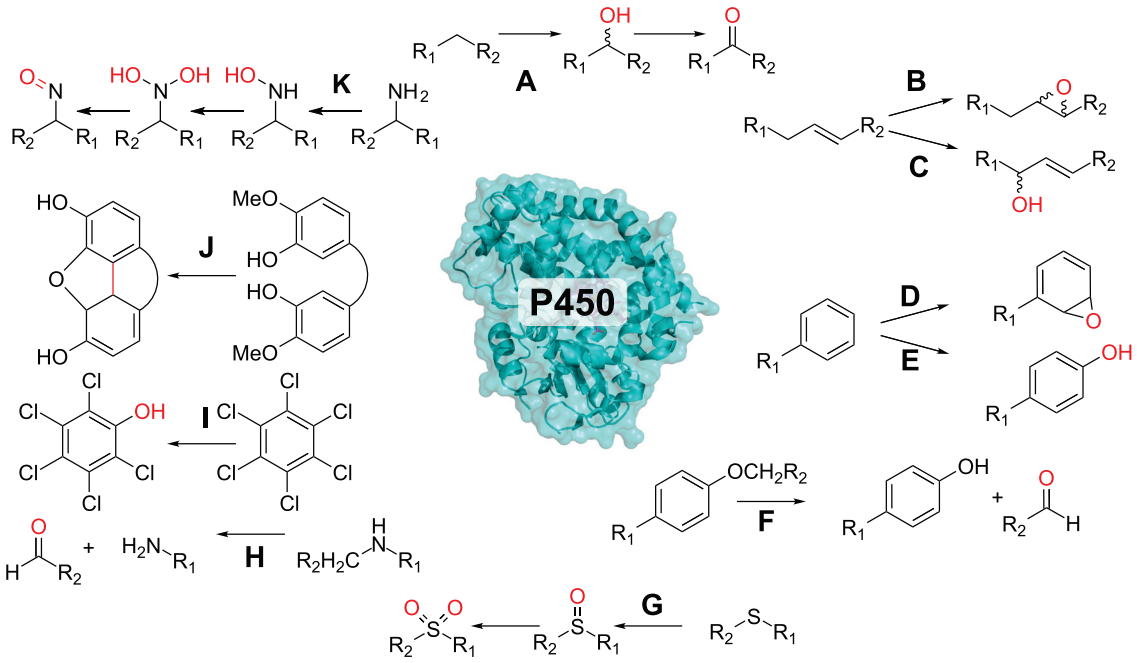


Figure 5: Selection of the most common reactions catalyzed by P450 monooxygenases. [47,48,52]

involved in secondary plant metabolism.[53–55] CYP71AV from *A. annua* catalyzes three consecutive oxidation steps (Fig. 6 A: I, II and IV) transforming an isopropenyl side chain into an α,β-unsaturated carboxylic acid, a prerequisite for the ring closure of the lactone present in the final compound artemisinin.

The bacterium *Amycolatopsis orientalis* is the producer of the glycopeptide antibiotic vancomycin (Fig. 6 B) which has long been a drug of last resort to treat infections caused by multi-drug resistant bacteria. In the biosynthesis of vancomycin $P450_{OxyA/B}$ and $P450_{OxyC}$ are responsible for a rather uncommon reaction type: oxidative phenol coupling. During its biosynthesis P450s are responsible for the formation of three new bonds, one aromatic C-C-bond and two ether bridgeheads connecting the three phenol moieties.[56] The resulting rigid conformation confers much to the bioactivity of vancomycin which originates from the inhibition of proper cell wall synthesis in gram-positive bacteria.

P450-catalyzed *O*-, and *N*-dealkylation are commonly found during drug metabolism e.g. of dextromethorphan by CYP2D6 and CYP3A4.[13] Via hydroxylation on the alkyl terminus of a methyl ether an unstable hemiacetal is formed which subsequently releases formaldehyde and the demethylated alcohol.

Heteroatom oxidation like the *N*-oxidation in the synthesis of Nocardicin A in *Nocardia uniformis* can be explained with a double hydroxylation of a primary amine to a dihydroxyl-

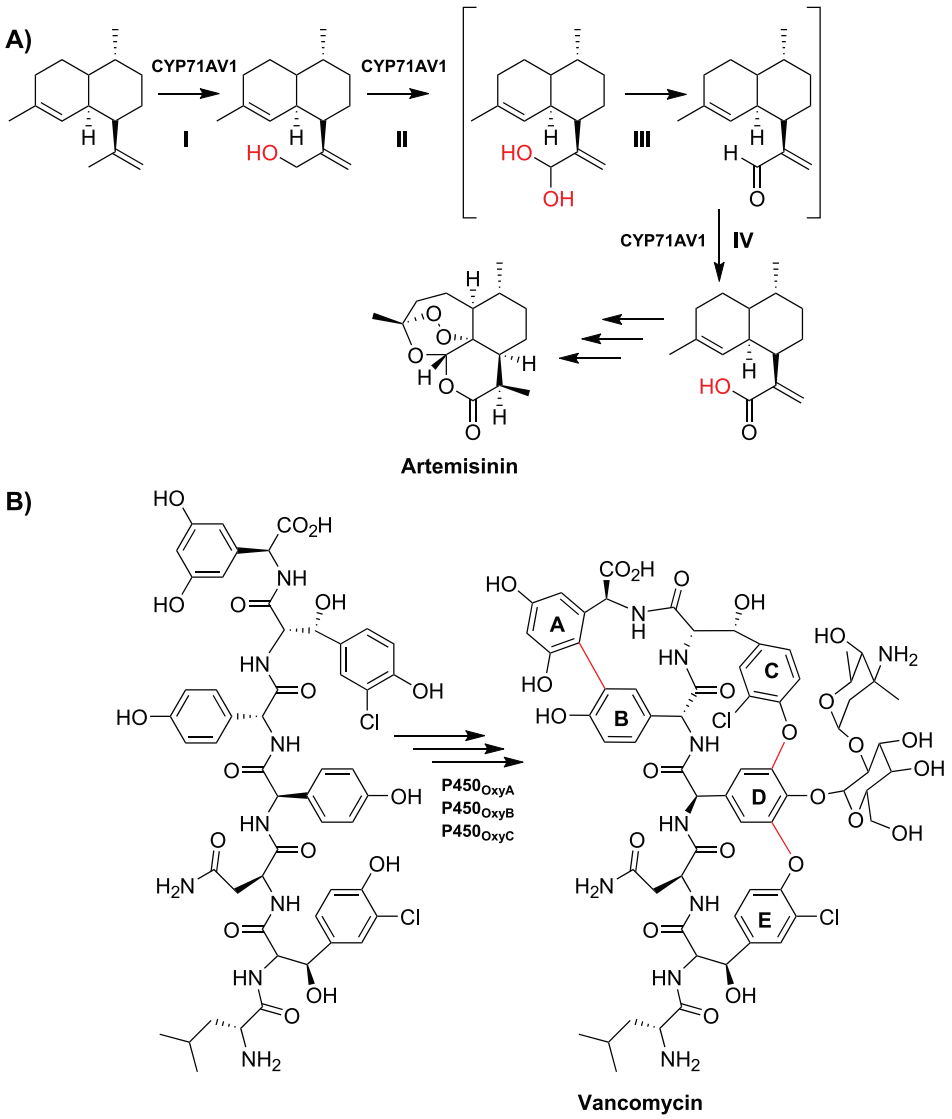


Figure 6: A) Biosynthesis of Artemisinin: CYP71AV1 is responsible for the first, allylic hydroxylation of amorpha-4,11-diene (I), as well as the second geminal hydroxylation (II) on the isopropenyl side chain. After rearrangement to the ketone (III) CYP71AV1 catalyzes a further oxidation step (IV) to artemisinic acid. In several further steps including ring closure and insertion of three further oxygen atoms, the final product artemisinin is formed.
B) Biosynthetic pathway of Vancomycin: $P450_{OxyB}$ closes the ether linkage C-O-D, $P450_{OxyA}$ forms the D-O-E bridgehead, and finally $P450_{OxyB}$ links ring A and B. The mechanism of these uncommon P450 reactions has not been investigated yet.

amine by $P450_{NocL}$ which subsequently forms the oxime after dehydration (Fig. 5 K) and tautomerization of the nitroso compound.[57]

Epoxidation is a common P450 reaction. Though it has been observed on aromatic substrates via a benzyne intermediate, olefinic epoxidation by P450s is more common. This reaction

type is also naturally occurring, e.g. in the biosynthesis of mycinamicin by $P450_{MycG}$ in *Microspora griseorubidia*.[58]

1.2.7 Synthetic applications of P450s

P450s have been described in several reviews as potentially useful biocatalysts.[5,59] A large number of accepted substrates and the high chemo-, regio- and stereoselectivity of the P450-catalyzed reactions make these enzymes attractive targets for the development of novel biotransformations. So far, the number of implemented large scale processes is still limited due to several limitations inherent to P450s and enzymatic processes in general, see also section 1.2.8.[60,61]

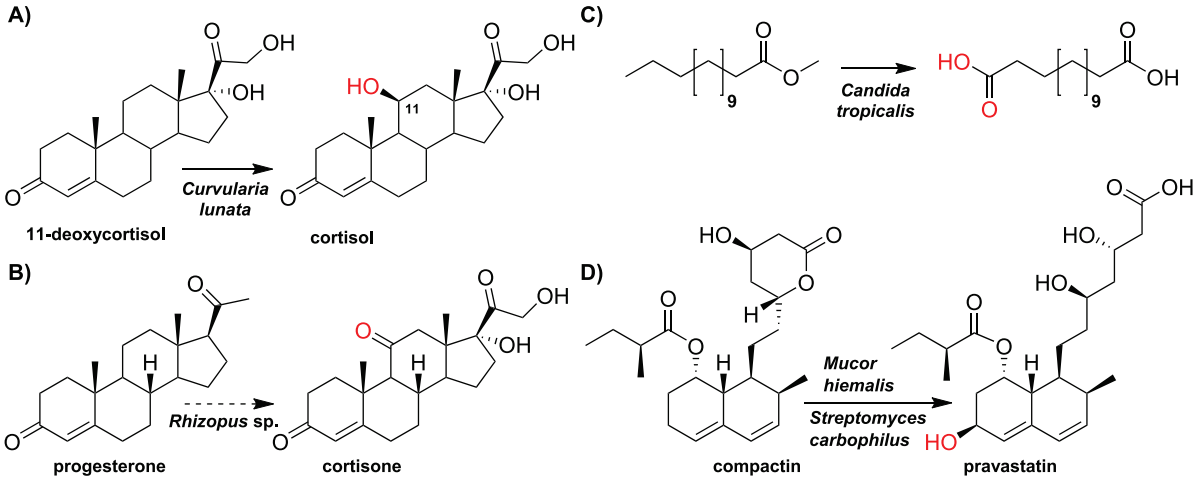


Figure 7: Implemented large scale industrial processes utilizing P450 monooxygenases: 11-β-hydroxylation activity of **A)** *Curvularia lunata* and **B)** *Rhizopus* sp. has been employed for steroid synthesis. **C)** Production of dicarboxylic acids is catalyzed by the yeast *Candida tropicalis*. **D)** A process for the production of pravastatin from compactin can be catalyzed either by *Mucor hiemalis* or *Streptomyces carbophilus* cells.

Two of the well-established processes make use of the 11-β-hydroxylation of steroids by microbial P450s (Fig. 7 A and B). Former Schering AG, now Bayer Health-Care Pharmaceuticals, established a process (100 t a^{-1}) for the production of cortisol from 11-deoxycortisol with the help of *Curvularia lunata* wild type cells expressing $P450_{lun}$.[62,63] A similar route was implemented by Pharmacia & Upjohn, now Pfizer Inc., for the synthesis of cortisone from progesterone utilizing *Rhizopus* sp.[64,65]

The production of dicarboxylic acids was demonstrated in > 100 m^3 scale with the help of *Candida tropicalis* harboring a number of P450s and several alcohol as well as aldehyde dehydrogenases (Fig. 7 C). The process was optimized in laboratory scale and a final product

concentration of 154 g L^{-1} was achieved with H_2O_2 as liquid oxygen source.[8]
A further process making use of a P450 as biocatalyst is the production of the cholesterol-lowering drug pravastatin (Fig. 7 D) from the precursor compactin. Two variants of the process using different cells have been described: Either *Mucor hiemalis* or *Streptomyces carbophilus* (CYP105A3) cells catalyze this reaction and are used by Daiichi Sankyo and Bristol-Myers Squibb, respectively.[66–68]

1.2.8 Limitations of P450 application

The limited number of established P450 processes is a first indication of the challenges these enzymes come with.[69] These drawbacks can be divided in three different groups: firstly, limitations related to substrates and cosubstrates; secondly, enzyme stability against various reaction parameters and, finally, all problems associated with the activity and selectivity of P450s.

1.2.8.1 Substrate and cosubstrates

Since P450s catalyze the incorporation of an oxygen atom into mostly hydrophobic substrates, their low solubility in the aqueous phase often can limit the productivity of P450-based processes. Substrate solubility can be increased by adding water-miscible organic co-solvents to the aqueous phase. A second approach is the application of biphasic reaction systems with water immiscible solvents which dissolve the substrate and ensure its constant supply into the aqueous phase.[70] A further advantage of a biphasic reaction system is the constant extraction of the formed products into the organic phase. Thereby, the products are protected from further oxidation by the P450 which has been demonstrated, e.g. for CYP109B1-catalyzed valencene oxidation.[71] Additionally, the substrate can be solubilized by cyclodextrins which intercalate apolar substrates into their hydrophobic core and due to their hydrophilic outer side are well soluble in the aqueous phase. Kühnel *et al.* demonstrated that the use of cyclodextrins significantly increased the substrate conversion of dodecanoic acid by CYP102A1.[72]
Since P450s depend on reducing equivalents in form of NAD(P)H in stoichiometric amounts, the high price of the nicotinamide cofactors precludes their use in large quantities and disfavors the P450 application *in vitro*. Hence, all previously described industrial processes utilize

whole cell biocatalysts. By employing the metabolism of the cell as cofactor regeneration system, cofactor costs can be greatly reduced.
An additional approach to overcome the need for stoichiometric amounts of cofactor involves the application of an additional enzyme for the cofactor regeneration. As cofactor regenerating enzymes, dehydrogenases are usually employed. These enzymes use a sacrificial substrate, e.g. glucose, glucose-6-phosphate, or formate in order to reduce the oxidized nicotinamide cofactors.[73]
Moreover, the catalytic cycle of P450s includes several branch points which lead to spontaneous autooxidative decay of reduced P450 without the formation of product. Hence, the the NAD(P)H consumption can be highly uncoupled from product formation, particularly, if non-natural substrates are used.[74] Uncoupling thereby further increases the necessary amounts of cofactor and simultaneously decreases enzyme stability due to formation of ROX (see section 1.2.8.2).

1.2.8.2 Stability

Although considerable progress has been made in tailoring P450 enzymes so that they can convert different substrates, their poor stability often remains a significant obstacle for applications in the synthesis of chemicals. Hydrogen peroxide and other reactive oxygen species produced in uncoupling reactions may lead to instability and degradation of the heme cofactor and the apoprotein. Low biocatalyst stability in this case represents an additional limiting factor for P450 biocatalysis. In order to reduce the negative effect of hydrogen peroxide on P450 stability catalase is usually added to the reaction mixture. So far, the source of the uncoupling problem itself has mainly been addressed by protein engineering of the P450.[75,76]
As mentioned in the previous section, organic solvents are usually used to increase substrate solubility in water and thereby their accessibility for P450 enzymes. However, it is generally recognized that organic solvents have a negative effect on enzymes, either by abstracting water molecules and thus degrading proteins as shown for water-miscible solvents, or by acting as inhibitors as shown for apolar solvents.
Solvent tolerance has been among others investigated for CYP1, CYP2 and CYP3-family enzymes in several studies with commonly used solvents.[77] The authors found that even small amounts of organic solvents can inactivate those P450s. To increase tolerance towards commonly used organic solvents, directed evolution has been successfully employed.[78,79]
A second approach to shield P450s from solvent is not based on protein engineering but on

reaction engineering. Possible methods are summarized in recent reviews and include, e.g. employment of P450s in reversed micells by using surfactants, or colyophilization of P450s with trehalose and subsequent suspension in *n*-decane.[79–81] On the other hand, the use of whole-cell biocatalysts providing accommodating environment for enzymes regarding ionic strength, pH and solvent contact can also be considered as a strategy for enzyme stabilization.

Under biocatalytic process conditions it can be desirable to use elevated temperature for several reasons, like influencing the reaction velocity and the reaction equilibrium but also the viscosity of the reaction solution and the solubility of reactants. Unfortunately, cytochrome P450s are generally temperature sensitive and the loss of the prosthetic heme group may occur spontaneously. Thus, higher thermal stability of P450s would allow novel applications of P450s in industrial processes. So far, only three thermophilic P450s have been described.[82] Among them CYP119 from *Sulfolobus solfataricus* was the first such enzyme whose crystal structure has been solved. The analysis revealed that it contains two unusual networks of aromatic amino acids and four networks of salt bridges which supposedly contribute to its thermostability.[83] In several studies it has been tried to confer thermostability to P450s of non-thermophilic origin by applying protein engineering methods.[84] Here especially rational design methods recombining sequence motives or whole domains from more thermostable enzymes have been proven successful.[85,86]

1.2.8.3 Activity and selectivity

To assess the activity of a P450 towards a certain substrate a working redox chain is required to supply reducing equivalents to the heme iron. Most eukaryotic P450s are supplied with electrons by only one CPR present in the organism, which makes the identification of a suitable redox partner quite straightforward. In the case of bacterial P450 cells contain often several P450 genes and a number of putative redox partner proteins. The identification of autologous redox partners is thus rather difficult and time-consuming. If their identification is not possible, heterologous redox partners originating from other organisms can be applied instead. Several well-established redox chains have been reported to be efficient with a number of "foreign" P450s.[28,29] In such cases if the protein-protein interaction is not optimal this leads to low P450 activities. In attempt to improve this interaction and to simplify the multi-component P450 system, a general strategy to construct "artificial" fusions has been introduced and successfully applied for several P450. This has been demonstrated with the

reductase domain of CYP102A1 as well as with CYP116B2.[87–89]

Drug-metabolizing hepatic P450s accept a broad range of substrates but show low activities and are difficult to express in recombinant hosts. In contrast, bacterial P450s are usually more stable, easy to express but exhibit a narrow substrate spectrum. Therefore, much work has been focused on broadening the substrate spectrum of the easy to handle bacterial P450s to mimic the more general catalytic properties of mammalian P450s.[69] An additional approach to circumvent the problem of a narrow substrate spectrum is the exploitation of the vast potential of all the P450 genes listed in databases instead of engineering a known P450 for a non-physiological reaction. To make use of these naturally occurring activities towards a specific substrate a lot of effort has been invested in the development of advanced screening methods, which is discussed in detail in the next section 1.2.9.

1.2.9 Screening of P450s

Traditional biocatalyst development is based on a target reaction for which a suitable activity has to be identified. The crucial point for the identification of a novel activity is the design of a robust screening method which can be subsequently applied to screen libraries of microbial strains *in vivo* or enzymes *in vitro*.[90] In both cases factors like membrane permeability of the target substance (*in vivo*), toxicity (*in vivo*), solubility and inhibition/inactivation have to be considered. In case of P450 enzymes besides the common chromatographic methods like liquid chromatography (LC) and gas chromatography (GC), applicable to both *in vivo* and *in vitro* screening, a variety of other methods, mostly suitable for *in vitro* screening, have been described.[91,92] A first spectroscopic *in vitro* method exploits the characteristic absorbance at 340 nm to detect the cofactor consumption of NAD(P)H-dependent enzymes. Since, contrary to many other enzymes, the NADPH consumption of P450s can be highly uncoupled from product formation, NADPH consumption assays often serve as a first rapid pre-screening.

Secondly, the spectroscopic features of P450s can be exploited for *in vitro* screening. Upon the addition of a potential substrate the absorbance spectrum is checked for a typical type I spectrum. This method has been successfully employed but comes with all the known limitation like compounds which induce a type I spectrum but are not converted or vice versa.[93–96]

Thirdly, several surrogate substrates have been described which result in a product molecule which can be detected by UV/VIS absorbance or fluorescence spectroscopy. Upon hydroxylation of a specific site they either release a chromogenic group, e.g. *p*-nitrophenolate, or

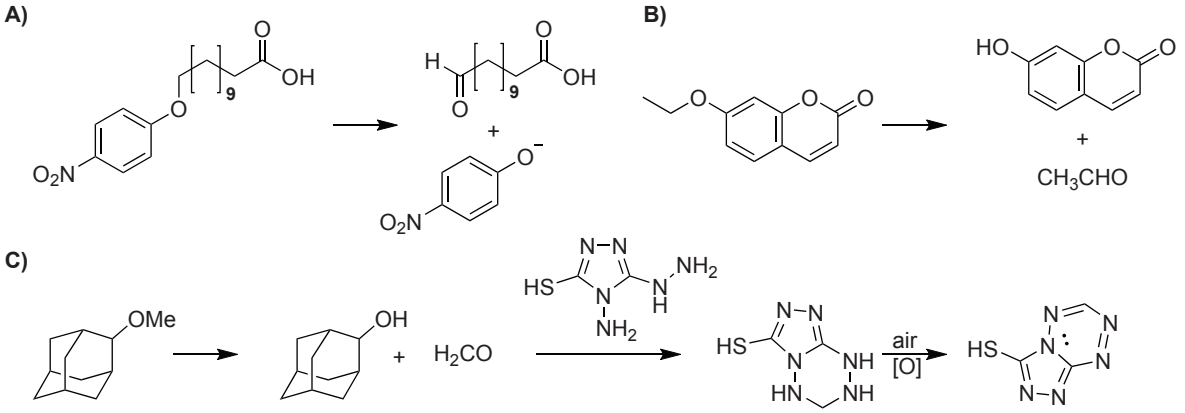


Figure 8: Selected screening methods for detection of P450 activity: **A)** The surrogate substrate *p*-nitrophenoxycarboxylic acid for detection of fatty acid hydroxylation via the absorbance of released *p*-nitrophenolate. **B)** Dealkylation of 7-ethoxycoumarin detected by fluorescence, **C)** Formaldehyde is released from the surrogate substrate methoxy adamantane and reacts with the dye purpald. After oxidation this can be detected by its purple color.

they are converted to a fluorophore, e.g. 7-ethoxycoumarin, or a secondary reaction with a dye is triggered, e.g. purpald (Fig. 8).[97–99]

A further approach uses mass spectroscopy to directly detect products which were formed by a screened P450. The complexity of the screening can be substantially reduced if several substrates are tested simultaneously. High resolution mass spectrometry (HR-MS) has been successfully employed to test eight P450s from *Bacillus subtilis* expressed in *E. coli* against a collection of 30 compounds. By using HR-MS it became possible to test all 30 compounds of the library simultaneously with each P450.[29] A limitation of this method, that has not been addressed by the authors, is the possible occurrence of false negative results if one of the tested compounds would act as an inhibitor and thereby mask activities towards the others. Furthermore this method does not reveal information about the product selectivity of the enzyme.

A last approach worth mentioning in this context is the fingerprinting method published by Fasan and coworkers.[100] As mentioned above (see section 1.2.6) five special surrogate substrates releasing formaldehyde are used as probes to detect site specific hydroxylation and thereby mapping the active site of a P450. The resulting fingerprints are used to predict activity of the P450 against other substances related to the probes. The method has been applied to a mutant library of CYP102A1 to characterize variants with respect to different regio- and stereoselectivity towards the substrate groups represented by the probes. Only recently the method has been extended to predict activity of the mutant library towards a substrate not represented by the probes.[101]

1.2.10 The CYP154 family

Literature data shows that the characterized P450s from the CYP154 family, in contrast to other families, accept substrates which are very diverse with respect to their molecular weight and structure as well as their chemical and biological functionalities. CYP154H1 mostly oxidizes small aromatic substrates like ethylbenzene, styrene, methyl phenyl sulfide, and indoles. The substrates of CYP154C1, C3, and C5 are larger and include steroids and lactone antibiotics with obviously very different chemical structures. For CYP154A8 two substrates were reported, 7-ethoxycoumarin and the isoflavone formononetin, which was *O*-demethylated and subsequently hydroxylated on an aromatic ring (Fig. 9). The only reported substrate of CYP154E1 which was identified previously in the group of Prof. Urlacher was the terpenoid geraniol which was converted to 8-hydroxy geraniol.[102]

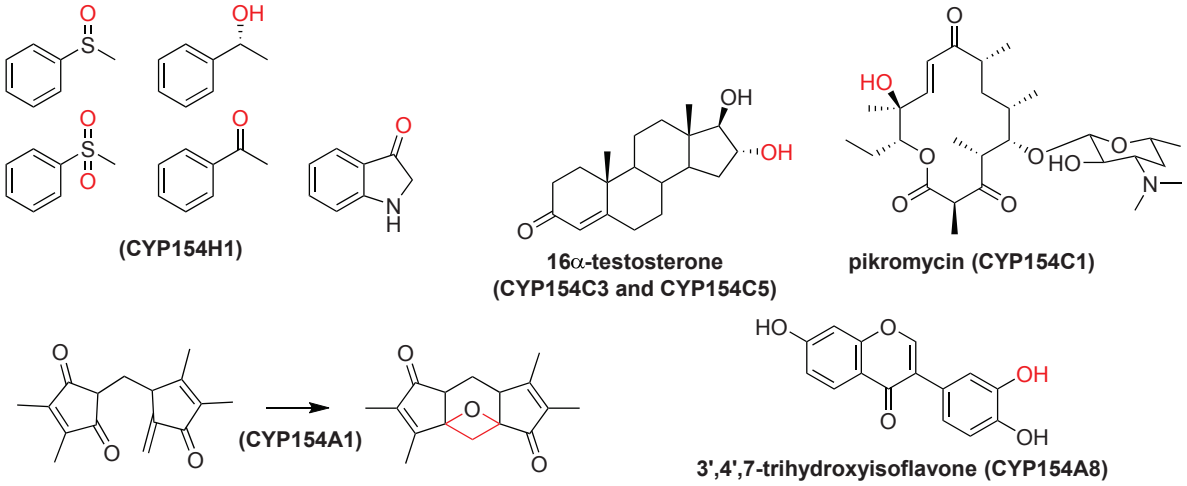


Figure 9: Overview of the reaction products of the so far characterized CYP154 enzymes: CYP154A1 and C1 from *Streptomyces coelicolor* A3(2), CYP154A8 and C5 from *Nocardia farcinica* IFM10152, CYP154C3 from *Streptomyces griseus*, CYP154E1 and CYP154H1 from *Thermobifida fusca* YX. The inserted functionalities or bonds are colored in red, respectively.

The CYP154 family consists of 35 sequences and is subdivided in subfamilies A-S (The Cytochrome P450 homepage; May 2014).[11,12] At least seven family members (CYP154A1, A8, C1, C3, C5, E1, H1) have been successfully expressed in *E. coli* and characterized.[28,98,102–105] Furthermore the crystal structures of CYP154A1 (protein data base (PDB)-ID: 1ODO) and C1 (PDB-ID: 1GWI) have been solved and published.[106,107] All P450s from the CYP154 family originate from Actinomycetes, Gram-positive bacteria, displaying a high GC-content of their genome. CYP154A1 and CYP154C1 whose physiological function has been elucidated, catalyze reactions which are part of the secondary metabolism of *Streptomyces coelicolor* A3(2).

CYP154C1 takes part in the biosynthesis of pikromycin (Fig. 9) whereas the cyclization reaction catalyzed by CYP154A1 is rather uncommon since it is redox neutral and hence takes place without the need of redox partners.
The general problem associated with the characterization of novel P450s, the need of suitable redox partners, has been solved for CYP154 enzymes using different strategies. The first approach which was chosen for CYP154C1, CYP154H1, CYP154E1, CYP154A8 was to use heterologous redox partner proteins. These were Pdx/PdR for C5, H1, A8, and E1 as well as ferredoxin and ferredoxin reductase from spinach for CYP154C1. A second approach which was successful with CYP154C3 was the construction of an artificial fusion protein with the reductase domain of CYP116B2.[104] The third approach was employed for the CYP154 members originating from *Streptomyces coelicolor*. A selection of flavin-containing reductases and ferredoxins from this strain was cloned and expressed in *E. coli* and tested with the P450s. The successful reduction of the CYPs was assayed by CO-binding studies *in vitro* after separate reconstitution of the P450s with the reductases and the ferredoxins.[103,108] In summary, so far most P450s in this family have been reconstituted with iron-sulfur cluster-containing redox partners. Whether this is a general property of CYP154 enzymes or whether this is due to the decision of the respective authors cannot be deduced from the literature so far.

1.3 Aim of the work

The interest in P450s for their application in biocatalytic processes is high, but the choice of suitable enzymes is difficult. On the one hand microsomal mammalian P450s have a broad substrate range since they often are responsible for detoxification of xenobiotics. On the other hand their low product selectivity, membrane association, and low expression levels in prokaryotes hinders their application.
Bacterial P450s in contrast, mostly are soluble enzymes and frequently show high expression levels in *E. coli*. A limitation for their application is however their narrower substrate spectrum which is related to their natural function in biosynthetic pathways or in the oxidation of defined carbon sources.
The long-term goal of this project was (i) the identification of a P450 family possessing versatile catalytic properties without the mentioned limitations and (ii) the development of biocatalytic processes based on enzymes from this family. The CYP154 family was identified based on literature data as potential source of novel P450 biocatalysts with suitable properties. Members of this family seem to combine the broad substrate spectrum of hepatic

CYPs, with the advantages of their bacterial counterparts, namely high product selectivity and efficient expression in prokaryotic hosts. Two members of this family, CYP154E1 and CYP154A8, have been selected to investigate whether their high phylogenetic diversity and distant subfamily relationship is reflected in their properties like activity and selectivity.
In this context the following aims have been set for this study:

1. In order to create effective P450 biocatalysts efficient redox partner systems are required. For this purpose, several redox partners for both CYP154 enzymes including autologous and heterologous ones should be tested and evaluated for P450 activity.
2. Individual substrate spectra of CYP154E1 and CYP154A8 should be determined. For this purpose, a substrate screening assay should be developed which on the one hand covers a broad space of chemical compounds and on the other hand does not detect false positives or negatives. Furthermore, the determined substrate spectra ideally should reveal systematic trends in the substrate spectra of CYP154E1 and CYP154A8.
3. In the next step, biocatalytic potential of CYP154 enzymes should be evaluated based on a few selected oxidation reactions leading to the formation of important fine chemicals or building blocks.
4. Finally, structural features of CYP154E1 should be identified which help to rationalize the observed substrate spectrum and selectivities. For this purpose, CYP154E1 should be crystallized and its three-dimensional structure should be solved.

2 Identification of appropriate redox partners for CYP154E1 and CYP154A8

2.1 Introduction

Most P450s rely on electrons supplied by redox partner proteins for catalytic activity. Thus, an essential prerequisite for screening is the reconstitution of P450 activity with appropriate redox partners. For the reconstitution of P450 activity two major strategies have been described in the literature. These are firstly the identification of autologous redox partners from the same host organism as the P450 and secondly, the use of heterologous redox partner which are able to reduce a P450 despite their origin from another organism.

1. Bacterial P450s typically rely on ferredoxins which accept electrons from specific flavin-containing reductases (class I). Ferredoxins are ubiquitous electron transport proteins in living cells.[109] In some species cell growth at low iron concentrations lets flavodoxins take over the electron transport role of ferredoxins. Above this flavodoxins are essential and constitutive proteins.[110] Since they take part in various redox processes microorganism generally can contain several different flavodoxins and ferredoxins which often interact with specific reductases only. As a consequence, the identification of autologous redox partners for a P450 out of this diversity of putative candidates is difficult and requires to test all combinations of P450, ferredoxins, flavodoxins, and the corresponding reductases, respectively. This approach has been successfully applied to P450s from *Streptomyces coelicolor* A3(2).[108] The advantage of this procedure is the knowledge that the host organism definitely contains a functional redox chain. Nevertheless, it is a very labor and time-consuming strategy. Moreover, it is not guaranteed that the redox partner proteins can be expressed in a heterologous expression system like *E. coli*.

2. The best established approach is to test versatile, available, heterologous redox partners from class I and III originating from other P450 redox chains in order to verify their ability to transfer electrons to a "foreign" P450. It has been demonstrated for P450s of various origin that this is possible e.g. with the adrenal Adx/AdR,[96] bacterial Pdx/PdR,[111] spinach ferredoxin and ferredoxin reductase,[107] and the *E. coli* Fldx/FdR system.[112]

   A variation of this approach is the construction of artificial fusion proteins consisting of a P450 monooxygenase artificially fused to the reductase domain of the naturally self-sufficient flavocytochromes of class VII or VIII.[87] These and other reductases have been successfully employed to obtain catalytically active P450 systems.[113–117] For an effective electron transport it is essential that the fused reductase domain interacts efficiently with the P450 domain. It has been demonstrated that a major determinant of this interaction is the linker length between the P450 and the reductase.[118]

For CYP154E1 and CYP154A8, different strategies have been chosen to reconstitute catalytically active P450 systems. In case of CYP154E1 autologous redox partners should be identified. In the genome of *T. fusca* no putative autologous redox partners genes could be found in close proximity to the sequence coding for CYP154E1, hence a broader search should be performed. Above this, for both P450s the compatibility with a set of heterologous redox partner proteins should be evaluated.
Because it is difficult to observe both reduction steps of the P450 performed by a redox partner protein spectroscopically the conversion of known substrates should be used as measure to quantify the electron transport. In general, problems can arise when no substrates are known but this was not the case for CYP154E1 and CYP154A8 since substrates for both enzymes have already been published.[98,102]

2.2 Experimental procedures

2.2.1 Materials and strains

All reagents were obtained from commercial sources and were of analytical grade or higher. Enzymes for molecular cloning were purchased from Fermentas (St. Leon-Roth, Germany) and New England Biolabs (Ipswich, USA). Formate dehydrogenase (FDH) for NADH regeneration was kindly provided by Prof. Tishkov (Department of Chemical Enzymol-

ogy, Lomonosov Moscow State University, Moscow, Russia).[119,120] The *Escherichia coli* strain DH5α was obtained from Clontech (Heidelberg, Germany). *E. coli* strains BL21(DE3), BL21(DE3)pLysS, NovaBlue(DE3), Rosetta(DE3), Origami 2(DE3) were obtained from Novagen (San Diego, USA). BL21-CodonPlus (DE3)-RP was commercially available from Stratagene (La Jolla, USA). Plasmid vectors pET-22b(+), pET-28b(+) and pCOLADuet-1 were purchased from Novagen (San Diego, USA). The constructs for expression of Pdx, PdR, YkuN, FdR, CYP154E1, and CYP154A8 were used as published.[96,102,121] Genomic DNA of *Thermobifida fusca* YX was obtained from the United States Department of Energy, Joint Genome Institute (Walnut Creek, USA).

2.2.2 Molecular cloning of redox partner proteins and CYP154E1

The *fdx1* and *fdx2* genes (Tab. 3) were amplified from genomic DNA by polymerase chain reaction (PCR) with Phusion DNA polymerase according to the manufacturer's protocol. For *fdx1* primers CVB_Tfu1226-F-NdeI / CVB_Tfu1226-R-EcoRI and for *fdx2* primers CVB_Tfu1275-F-NdeI / CVB_Tfu1275-R-EcoRI were used. All primers are listed in Tab. 2. After purification, the PCR products were separately cloned into pET-22b(+) exploiting the introduced restriction endonuclease sites (Tab. 2). The sequences of the resulting plasmids pET22-fdx1 and pET22-fdx2 were verified by a commercial DNA sequencing service (GATC, Konstanz, Germany). Plasmid maps of all constructed vectors can be found in section 9.2.
With the same protocol the reductase genes *fdr1* (CVB_Tfu_1227-F-NdeI / CVB_Tfu1227-R-EcoRI) and *fdr2* (CVB_Tfu1273-F-NdeI / CVB_Tfu1273-R-EcoRI) were cloned in pET-22b(+) leading to the expression plasmids pET22-fdr1 and pET22-fdr2.

For the coexpression of ferredoxins with the corresponding reductase each ferredoxin gene was inserted in the multiple cloning site (MCS) 1 of pCOLADuet-1 and the corresponding reductase gene was inserted in MCS 2. The following primers were used for amplification of the genes: *fdx1* (CVB1226EcoRI-F-D / CVB1226Hind3-R-D), *fdr1* (CVB_Tfu_1227-F-NdeI/CVB1227XhoI-R-D), *fdx2* (CVB1275EcoRI-F-D / CVB1275Hind3-R-D) and *fdr2* (CVB_Tfu1273-F-NdeI / CVB1273XhoI-R-D). The earlier described pET-22b(+)-based constructs pET-fdx1, pET-fdx2, pET22-fdr1 and pET22-fdr2 served as PCR templates. The resulting expression constructs pFdxFdr-1 and pFdxFdr-2, respectively, code for N-terminally His_6-tagged ferredoxins.
For coexpression with the pCOLADuet-1-based expression plasmids, CYP154E1 was subcloned from pET28-CYP154E1 into pET-22b(+) using the *Nde*I/*Eco*RI restriction sites.

Table 2: Oligonucleotides used for polymerase chain reaction (PCR). Restriction sites in the primer sequence are printed in lower case.

Designation	Sequence (5' → 3')	Restriction site
CVB_Tfu1226-F-NdeI	GAGTACT catatg GACATCCACGTCGACTACG	*Nde*I
CVB_Tfu1226-R-EcoRI	CC gaattc TCATTGTGCCCTCAGAGCGC	*Eco*RI
CVB_Tfu1275-F-NdeI	TATAATT catatg CGCGTCCGAGTCGACGGAAAG	*Nde*I
CVB_Tfu1275-R-EcoRI	CG gaattc TCACTCGACGGTGATGGCC	*Eco*RI
CVB1226EcoRI-F-D	CG gaattc G ATGGACATCCACGTCGACTAC	*Eco*RI
CVB1226Hind3-R-D	CCC aagctt TCATTGTGCCCTCAGAGCG	*Hind*III
CVB_Tfu_1227-F-NdeI	TATAATT catatg ACCGCGCCCGCCCGC	*Nde*I
CVB_Tfu1227-R-EcoRI	CC gaattc CTACGCACGTGTTCCTCCTG	*Eco*RI
CVB1227XhoI-R-D	TAT ctcgag CTACGCACGTGTTCCTCCTGTCTG	*Xho*I
CVB1275EcoRI-F-D	GC gaattc C ATGCGCGTCCGAGTCGAC	*Eco*RI
CVB1275Hind3-R-D	CCC aagctt TCACTCGACGGTGATGGCC	*Hind*III
CVB_Tfu1273-F-NdeI	GAGTACT catatg GAACGGATCGTCATTGTGGG	*Nde*I
CVB_Tfu1273-R-EcoRI	CC gaattc TCAACGCCCGCGAGGTG	*Eco*RI
CVB1273XhoI-R-D	TAT ctcgag TCAACGCCCGCGAGGTG	*Xho*I
CVB_154E1t-NdeI-F	GGATTAC catatg GCTTTTCCCGGGAACC	*Nde*I
ARO_154E1-EcoRI-R	CAGT gaattc TCAGGGTTTCGGGCG	*Eco*RI

2.2.3 Heterologous protein expression and purification

Chemically competent *E. coli* expression strains were transformed with the respective expression plasmids. From a single colony an overnight culture in 5 ml LB-medium was inoculated. The overnight cultures were used as inoculum for the expression cultures in shaking flasks (TB medium, 1:100). Cells were grown to an OD_{600} of 0.6-0.8 and gene expression was induced by the addition of 100 µM IPTG, unless otherwise stated. Supplementation of the medium with 100 µM $FeSO_4$ and 500 µM 5-aminolevulinic acid (5-ALA) ensured high availability of the heme cofactor for P450 expression. For the expression of ferredoxins only 100 µM $FeSO_4$ was added. Expression was performed at 25 °C and 140 rpm unless otherwise stated. After expression, generally 16 h, cells were harvested by centrifugation at 17700 g for 15 min.

CYP154E1 was expressed using BL21(DE3) cells harboring pET28-CYP154E1 as published previously.[122] Heterologous redox partners were expressed under the same conditions as CYP154E1. After harvesting, cells were resuspended in potassium phosphate buffer (50 mM, pH 7.5, 100 mM phenylmethanesulfonyl fluoride). Cell disruption was performed by ultrasound with a Branson Sonifier S-250A (6 x 45 s; 45 % duty cycle; output control 4.5) on ice. The P450 and the heterologous redox partners were purified by nickel affinity chromatography on two consecutive HisTrap FF crude 5 mL columns (GE Healthcare, Freiburg Germany) and subsequently desalted on a HiPrep 26/10 Desalting column (GE Healthcare, Freiburg Germany). Protein quantification of the heterologous redox partners was done by UV/VIS spectroscopy with published extinction coefficients: YkuN, ε_{461}=10.0 mM^{-1} cm^{-1}; FdR, ε_{456}=7.1 mM^{-1} cm^{-1}; Pdx, ε_{415}=11.1 mM^{-1} cm^{-1} and ε_{455}=10.4 mM^{-1} cm^{-1}; PdR, ε_{378}=9.7 mM^{-1} cm^{-1}, ε_{454}=10.0 mM^{-1} cm^{-1} and ε_{480}=8.5 mM^{-1} cm^{-1}.[96]
Glucose dehydrogenase (GDH) for cofactor regeneration was expressed and purified as described elsewhere.[122] Expression analysis was performed by Glycine-SDS-PAGE (Laemmli) with Coomassie staining.[123] Due to their low molecular weight for the analysis of ferredoxin expression Tricine-SDS-PAGE was used. Polyacrylamide concentrations of 4 % for the stacking gel, 10 % for the spacer gel and 16 % for the separating gel were used.[124] Tricine-SDS-polyacrylamide gels were either stained by Coomassie brilliant blue R250 or silver staining.

2.2.4 Biotransformation of geraniol

E. coli cells coexpressing the autologous redox parter system together with CYP154E1 were harvested after 4 h. After cell lysis the P450 concentration in the soluble protein fraction was determined by CO-difference spectroscopy.[125] The ability of autologous redox partner systems to transfer electrons to CYP154E1 was investigated in biotransformation reactions of geraniol. The test substrate geraniol was shown to be accepted by CYP154E1 previously.[102]
In the reaction system 5 µM P450 originating from the coexpression was employed. Activity should be reconstituted by 200 µM NADPH and 200 µM NADH, cofactor regeneration with 5 U glucose-6-phosphate dehydrogenase, 2 mM glucose-6-phosphate, 300 U catalase, and 200 µM geraniol. After incubation in plastic reaction tubes overnight at 30 °C, the samples were analyzed via GC/MS using an established method.[102]
The test reaction of separately expressed heterologous redox partners with CYP154E1 and CYP154A8, respectively were performed in 500 µl scale at 25 °C for 5 h in plastic reaction tubes. Reactions for testing the Pdx/PdR system contained 1 µM P450, 5 µM PdR, 5 µM Pdx,

for cofactor regeneration 5 U FDH, 20 mM sodium formate, 200 µM NADH, 200 µM substrate, and 300 U catalase. Reactions testing the YkuN/FdR system contained 1 µM P450, 10 µM YkuN, 1 µM FdR, for cofactor regeneration 5 U GDH, 20 mM glucose, 200 µM NADH, 200 µM substrate, and 300 U catalase. The samples were analyzed via GC/MS using an established method.[102,122]

2.3 Results

2.3.1 Identification of autologous redox partners

In a work by Sirim *et al.* the reductase interaction sites (RIS) in 31 P450 monooxygenase structures, including CYP154A1 and CYP154C1, were examined.[26] The authors observed that P450s which receive electrons from CPRs possess larger RISs compared to those accepting electrons from ferredoxins and flavodoxins. The alignment of both CYP154A8 and CYP154E1 to the aforementioned CYP154A1 and CYP154C1 sequences revealed high sequence homology in the area of RIS1 and RIS2. Hence, the conclusion that both P450s depend on electrons supplied by class I or class III redox partner systems was drawn. The genome (GeneBank-ID: GI:72160406) of the soil bacterium *Thermobifida fusca* YX has been published in 2007 and is available in annotated form at the DOE Joint Genome Institute (JGI, http://genome.jgi-psf.org/thefu/thefu.home.html).[126] The annotations were screened for ferredoxins and flavodoxins as well as flavin-containing reductases in order to identify potential redox partners candidates of CYP154E1. The results are shown in Tab. 3.

The genome contains four ferredoxins and four ferredoxin reductases. According to their protein families database assignment all ferredoxins are of the [4Fe-4S]-type and belong to the Fer4 and the Fer4_19 families.[127] The ferredoxin reductases are annotated as NAD^+-dependent proteins except Tfu_1915 which is expected to accepts $NADP^+$ as cosubstrate.

Based on literature data which show that many of the P450s from the CYP154 family were successfully reconstituted with heterologous iron-sulfur-type redox partners, it was decided to clone *fdx1*/*fdr1* (redox partner system 1) and *fdx2*/*fdr2* (redox partner system 2). Since the two genes of each ferredoxin/ferredoxin reductase pair are located in close proximity to each other on the genome it is safe to assume that the respective reductase can transfer electrons to the corresponding ferredoxin.

Table 3: Overview of all genes coding for potential redox partners of CYP154E1.

Name	Locus ID	Protein ID	Annotation	GC-content (%)	Size (bp/kDa)
fdx1	Tfu_1226	YP_289287	Ferredoxin	66	195/7.1
fdx2	Tfu_1275	YP_289336	Ferredoxin	69	192/6.9
fdx3	Tfu_0491	YP_288552	Ferredoxin	62	321/11.7
fdx4	Tfu_1479	YP_289538	Ferredoxin	68	216/7.5
fdr1	Tfu_1227	YP_289288	Putative ferredoxin reductase	72	1245/44.1
fdr2	Tfu_1273	YP_289334	Ferredoxin-NAD^+ reductase	71	1191/42.5
fdr3	Tfu_1915	YP_289971	Ferredoxin-$NADP^+$ reductase.	66	1380/49.5
fdr4	Tfu_2348	YP_290404	Ferredoxin-NAD^+ reductase.	72	1203/42.7
fldx	Tfu_3023	YP_291079	Similar to Flavodoxin	68	513/18.7
fldr1	Tfu_0423	YP_288484	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	69	1185/43.1
fldr2	Tfu_1230	YP_289291	Putative NADH-flavin reductase	69	636/22.3

2.3.2 Expression of the reductases Fdr1 and Fdr2

To identify the optimal *E. coli* strain for recombinant expression of the selected reductases the strains BL21(DE3), BL21(DE3)pLysS, and BL21-CodonPlus (DE3)-RP were transformed with pET22-fdr1 and pET22-fdr2. Protein expression was induced with 500 µM IPTG and the cultures were grown at 30 ℃ and 180 rpm overnight.

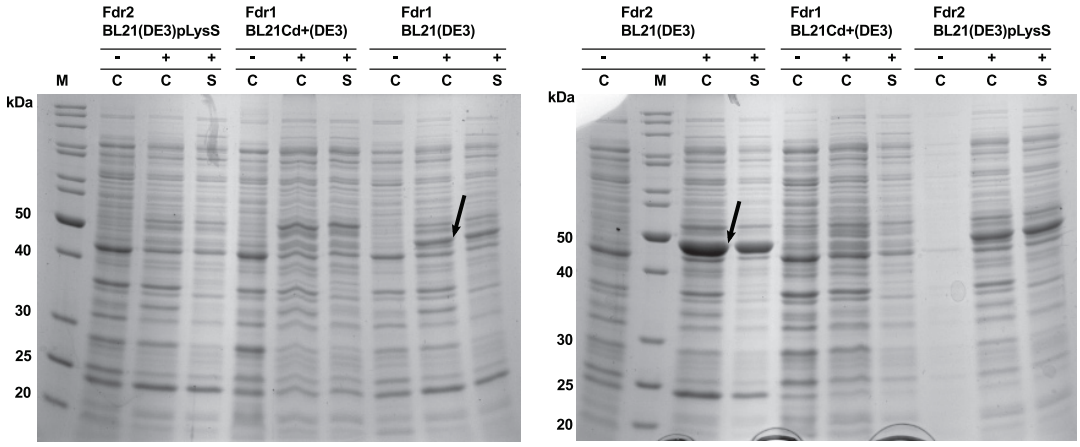


Figure 10: SDS-polyacrylamide gel of different *E. coli* strains expressing Fdr1 and Fdr2, respectively. Expression strains (**BL21Cd+(DE3)**: BL21-CodonPlus (DE3)-RP) and expressed genes are given on the top. **M**: protein marker, **C**: whole cell sample, **S**: soluble protein fraction after cells lysis and centrifugation, **-/+**: before/after induction with IPTG. Bands indicating the highest expression for Fdr1 and Fdr2, respectively, are marked with an arrow.

SDS-PAGE analysis of protein samples taken from different strains expressing the reductases revealed that *E. coli* BL21(DE3) yielded the highest protein concentration. This was indicated by bands at about 45 kDa for Fdr1 and Fdr2 on the SDS-polyacrylamide gel (Fig. 10). Thus, the expression conditions in the BL21(DE3) strain were further optimized regarding expression temperature (18-37 ℃). The SDS-polyacrylamide gels (Fig. 11 A) indicated that the optimal expression temperature for the given expression time was 18 ℃ for Fdr2 and 30-37 ℃ for Fdr1. In this experiment Fdr2 accounted for ~60 % of the total soluble *E. coli* protein.

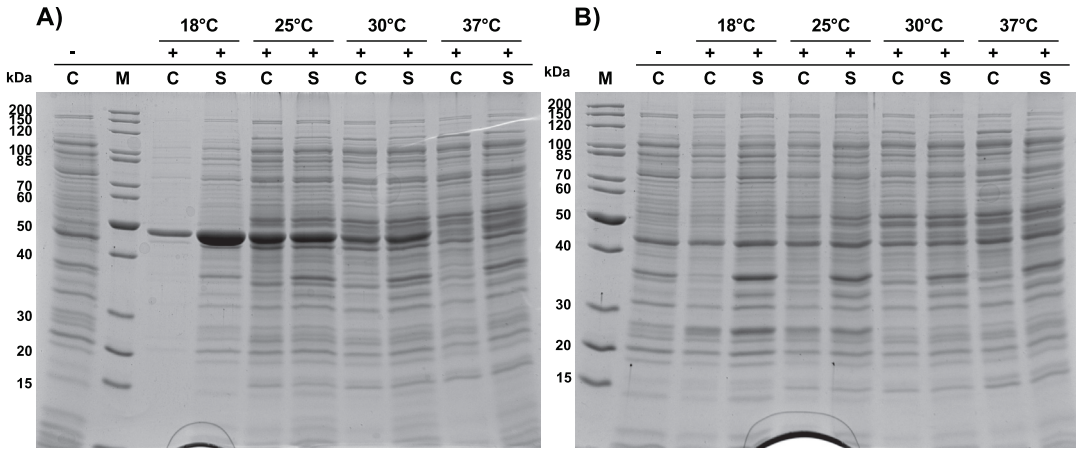


Figure 11: SDS-polyacrylamide gel of *E. coli* BL21(DE3) expressing (**A**) Fdr2 and (**B**) Fdr1 at different temperatures. **M**: protein marker, **C**: whole cell sample, **S**: soluble protein fraction after cells lysis and centrifugation, **-/+**: before/after induction with IPTG. Expression in BL21(DE3) was performed overnight at the temperatures indicated on the top.

2.3.3 Expression of the ferredoxins Fdx1 and Fdx2

To identify the optimal strain for expression of the ferredoxins, the strains *E. coli* strains BL21(DE3), BL21(DE3)pLysS, Rosetta(DE3), NovaBlue(DE3), Origami 2(DE3), BL21-CodonPlus (DE3)-RP were transformed with pET-fdx1 and pET-fdx2, respectively. Protein expression was induced with 500 µM IPTG and the cultures were grown at 25 °C and 180 rpm overnight. The SDS-PAGE analysis did not show any bands corresponding to the expected size of either Fdx1 (7.1 kDa) or Fdx2 (6.9 kDa). In order to evaluate the influence of the expression time on the protein yield, shorter intervals were tested. Samples of BL21(DE3) cells expressing each ferredoxin were taken for 6 h. The SDS-PAGE analysis revealed that the reduction of the expression time did not result in any detectable amounts of Fdx1 or Fdx2.

2.3.4 Coexpression of Fdx1/Fdr1 and Fdx2/Fdr2

In order to facilitate expression expression the coexpression of the ferredoxins with their corresponding reductase was investigated. Potentially, this should allow both proteins to interact during expression and, thus, might be beneficial for correct folding. BL21(DE3) cells harboring the pFdxFdr-1 or pFdxFdr-2 plasmids were induced with 100 µM or 500 µM IPTG. After expression at 25 °C or 30 °C overnight, protein samples were analyzed by conventional SDS-PAGE and simultaneously by Tricine-SDS-PAGE with silver staining. The detection by silver staining ensured that even very small amounts of the ferredoxins could be detected. As an example the SDS-polyacrylamide gel and the corresponding Tricine-SDS-polyacrylamide gel of the Fdx1/Fdr1 coexpression are depicted in Fig. 12. On the silver stained Tricine-SDS-polyacrylamide gel (Fig. 12 A) a band which correlates to the size of His_6-tagged Fdx1 (8.8 kD) could be identified after 1.5 h in whole cell samples and the soluble protein fraction. This band reached maximal intensity after 4 h of expression. A corresponding band on the Coomassie stained SDS-polyacrylamide gel (Fig. 12 B) was not visible. This suggested that either the expression level of Fdx1 was extremely low or that the band on the Tricine-SDS-polyacrylamide gel was an artifact not related to Fdx1.
Howsoever, since the autologous redox partners of CYP154E1 were investigated with the aim to design an efficient biocatalyst both redox partner system had to prove their ability to transfer electrons to CYP154E1 *in vitro*. For this purpose BL21(DE3) cells harboring pET22-CYP154E1 in combination with pFdxFdr-1 or pFdxFdr-2 were used to coexpress CYP154E1 with each redox partner system. Expression was performed for 4 h. The resulting soluble protein fraction was used to adjust the P450 concentration in test reactions to 5 µM. The

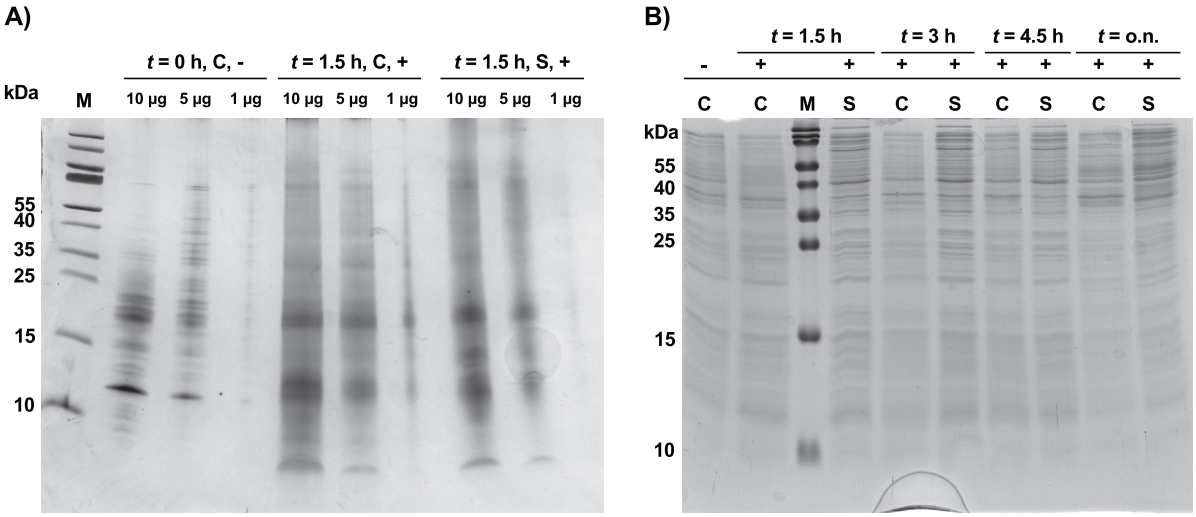


Figure 12: Silver stained Tricine-SDS-polyacrylamide gel **A)** and SDS-polyacrylamide gel **B)** of *E. coli* BL21(DE3) coexpressing Fdx1 and Fdr1. **M**: protein marker, **-/+**: before/after induction with IPTG, **C**: whole cell sample, **S**: soluble protein fraction after cells lysis and centrifugation. The expression time is given in the top line, **o.n.**: overnight. **A)** The total amount of protein loaded per lane is indicated in the second line. Expression conditions: 30 ℃, 100 µM IPTG.

test reactions with geraniol as substrate were incubated overnight at 30 ℃ and after extraction analyzed via GC/MS. The GC/MS-chromatograms are depicted in Fig. 13. Besides the substrate peak (1), citronellol (2), and indole (3) there were no further peaks indicating any P450 mediated reaction. Hydroxylated products were expected to elute at a retention time of ~12 min but have obviously not been formed. Instead citronellol and indole were present in the samples. Both were detected due to the *E. coli* cell extract used in the reaction. The

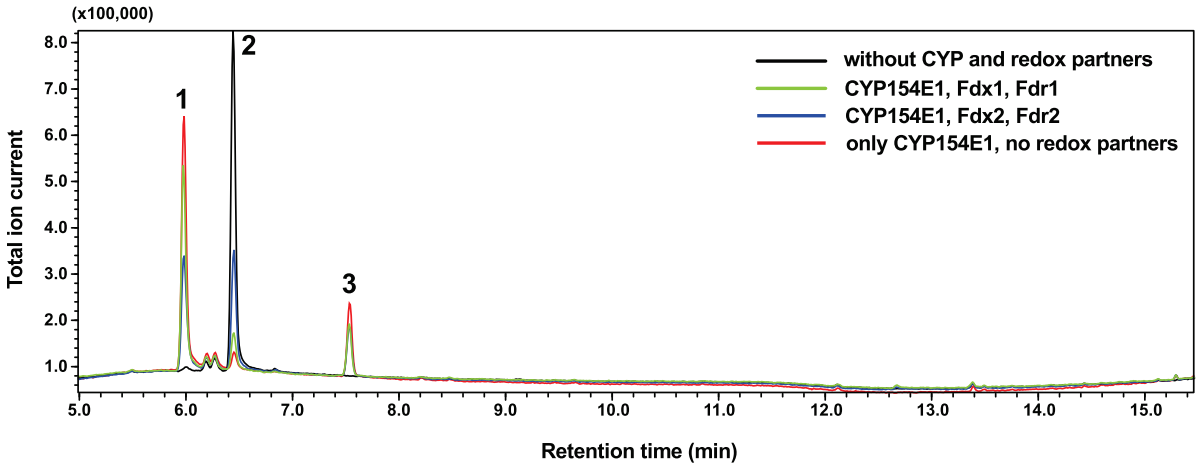


Figure 13: GC/MS chromatograms of geraniol conversions catalyzed by CYP154E1 and supported by the two *T. fusca* redox partner systems. Peak **1**: citronellol (reaction product of *E. coli* cell lysate with geraniol), **2**: substrate geraniol, **3**: indole (contamination from *E. coli* cell lysate).

expression level of CYP154E1 varied between the coexpression experiments. Thus, different amount of cell extract had to be used in the reaction to adjust the P450 concentration to 1 µM. *E. coli* cells contain an endogenous activity which is responsible for the formation of citronellol from geraniol. The different amounts of *E. coli* cells applied to the reaction also resulted in different amounts of this activity in the reaction. This explains the different peak areas of citronellol in the chromatograms (Fig. 13). Indole was a direct contamination from the *E. coli* cell extract.
At this stage the evaluation of autologous redox partner systems has been aborted. The expression levels obviously were too low to support CYP154E1-based catalysis and, hence, were of no practical use. Further options had to be considered and were investigated as described in the following section.

2.3.5 Investigation of heterologous redox partners

Since the identification of autologous redox partners for CYP154E1 failed, two well-established, heterologous redox partners systems available in the laboratory were tested. The Pdx/PdR system from *Pseudomonas putida* and the bacterial flavodoxin YkuN from *Bacillus subtilis* together with FdR from *E. coli*. The compatibility of the two heterologous redox partner systems with CYP154E1 and CYP154A8 was investigated in biotransformation reactions. The conversions of 200 µM Grundmann's ketone (CYP154E1) and 200 µM geraniol (CYP154A8) were calculated from GC/MS peak areas (Tab. 4).

Table 4: Evaluation of the interaction of CYP154E1 and CYP154A8 with heterologous redox partner systems by conversion of Grundmann's ketone (CYP154E1) and geraniol (CYP154A8).

	CYP154A8	**CYP154E1**
Redox Parter System	**Conversion (%)**	
YkuN/ FdR	100[(a)]	100[(a)]
Pdx/ PdR	49	9

[(a)] no remaining substrate detected

With YkuN/FdR complete conversion was observed for both P450s. The electron transport to CYP154E1 and CYP154A8 by Pdx/PdR was less efficient at the same cofactor concentration and resulted in conversions of 9 % and 49 %, respectively. YkuN/FdR as redox partner system led to a 2 and 11-fold increased conversion and thus was selected for the subsequent characterization of both P450s.[122]

2.4 Discussion

A prerequisite for the investigation of catalytic activities of P450 monooxygenases is a functional redox chain which can transfer electrons to the heme iron of a P450. Bacterial class I P450 systems consist of an iron-sulfur-containing ferredoxin and an FAD-containing reductase. Class III systems, also of bacterial origin, consist of a flavodoxin (FMN) and a corresponding FAD-containing reductase.[27]

The analysis of the *T. fusca* genome identified one flavodoxin (*fldx1*) gene and two potential flavodoxin reductases genes. Besides that, the genome contains four [4Fe-4S]-ferredoxin genes (*fdx1-4*) and four ferredoxin reductase genes (*fdr1-4*). According to the available literature the majority of active CYP154 systems has been reconstituted with ferredoxins. Two ferredoxins of *T. fusca* (Fdx1/Fdr1: redox partner system 1 and Fdx2/Fdr2: redox partner system 2) are located in direct vicinity to an NADH-dependent ferredoxin reductase gene. This might be an indication that they interact with each other and thus the probability to obtain functional redox partner systems is higher compared to the other candidates. Remarkably all four ferredoxins are of the [4Fe-4S]-type. Archetypical ferredoxins interacting with P450s, e.g. adrenodoxin and putidaredoxin, contains a [2Fe-2S]-cluster.[128] Ferredoxins containing [4Fe-4S]-clusters are reported to take part, among others, in reactions oxidizing H_2.[109] However, recently, the first [4Fe-4S]-ferredoxin from *B. subtilis* has been shown to interact with and transfer electrons to $P450_{BioI}$.[34] Similarly, a [3Fe-3S]-ferredoxin from *Mycobacterium tuberculosis* proved to be a reducing agent of the sterol demethylase (CYP51) from the same organism.[129]

All four redox partner genes have been cloned separately in pET-based expression vectors under the control of the T7 promoter. Utilizing these constructs expression conditions were optimized. Both reductases could be expressed in soluble form in *E. coli* BL21(DE3). Especially Fdr2 accounted for ~60 % of the total soluble protein in *E. coli*. Concerning the interaction with a P450, the respective ferredoxins are the more important part of the system. Expression of Fdx1 and Fdx2 in *E. coli* from the pET-22b(+)-based constructs could not be observed even after alteration of several expression parameters. To allow expression of the ferredoxins, plasmids for coexpression of both proteins of each redox partner system were constructed. Analysis of the expression revealed a protein band correlating to the size of Fdx1 in extremely low amounts and only detectable by silver staining.

To evaluate if the autologous redox partner systems support catalysis of CYP154E1, the monooxygenase was coexpressed with each system. The soluble protein fraction was then employed as biocatalyst in a test reaction with geraniol, a substrate of CYP154E1 known from

literature.[102] In case of redox partner systems 1 and 2 no product formation was observed. Whether this was due to an inability to transfer electrons onto CYP154E1 or due to their obviously low expression levels remained an open question. Difficulties with expression of ferredoxins have been reported in literature also in the case of *Sorangium cellulosum*.[130] Ewen *et al.* described the cloning and heterologous expression of eight ferredoxins in *E. coli*. Three of them, all of the [4Fe-4S]-type, could not be characterized due to insolubility, instability, or no expression at all.

The third and finally successful strategy was the reconstitution of both P450's catalytic activity with heterologous redox partners. In contrast to most other CYP154 monooxygenases described in literature, best conversions were not achieved with class I systems but with those belonging to class III. The combination of YkuN and FdR led to the highest conversions with both enzymes and enabled the substrate screening discussed in the following chapter.

3 Manuscript 1

Titel: Cluster Screening: An Effective Approach for Probing the Substrate Space of Uncharacterized Cytochrome P450s

Authors: Clemens von Bühler, Priska Le-Huu, and Vlada B. Urlacher

Published in: Chembiochem: a European journal of chemical biology, **2013**, *14*, 2189 (ISSN:1439-4227)

DOI: 10.1002/cbic.201300271

Contribution: Expression and purification of all proteins, spectroscopic characterization of both P450s, Screening and product analysis, all binding studies except cembrene, writing of the manuscript

Link: http://dx.doi.org/10.1002/cbic.201300271

CHEMBIOCHEM
FULL PAPERS

ChemPubSoc Europe

DOI: 10.1002/cbic.201300271

Cluster Screening: An Effective Approach for Probing the Substrate Space of Uncharacterized Cytochrome P450s

Clemens von Bühler, Priska Le-Huu, and Vlada B. Urlacher*[a]

Cytochrome P450 monooxygenases (P450s) are versatile enzymes with high potential for biocatalysis. The number of newly annotated P450 genes has been increasing constantly, and these thus represent a rich resource for new biocatalysts. However, the substrate scopes of newly identified P450s are often not known, and thus their exploitation is difficult. Herein we describe an approach, named "cluster screening", and its application for the primary characterization of two P450s: CYP154E1 and CYP154A8. A library comprising 51 compounds was designed and organized into nine groups according to their chemical properties. The activities of both P450s in vitro were maintained with suitable nonphysiological redox partners, and the cluster library was screened with these enzymes for product formation. From this library, 30 compounds tested positive for CYP154E1 and 23 were positive for CYP154A8. Cluster screening distinguishes subtle differences in activity and selectivity of enzymes as closely related as those of the same P450 family. For example, the alkaloid pergolide mesylate was converted by CYP154E1 (4%) but not by CYP154A8. A building block of vitamin D_3, Grundmann's ketone, was converted by both enzymes, although conversion was higher with CYP154E1 (100 vs 53%).

Introduction

Cytochrome P450 enzymes (CYPs, P450s) belong to a class of oxidoreductases (EC 1.14). They are ubiquitous enzymes found in all kingdoms of life.[2] The ability of P450s to catalyze selective oxidations of a vast variety of organic molecules makes them important candidates for synthetic applications. The rapidly growing number of newly identified P450 genes has attracted much attention as a rich resource for novel oxidation activities.[3] The P450 superfamily consists of more than 12000 classified enzymes and about 6000 genes that have been sequenced but not yet classified.[2] However, the pace of genome mining and gene annotation based on sequence similarity is much faster than the characterization of new P450s and elucidation of their functions. As a consequence, so far only a small number of these P450 genes have been expressed and characterized.

The characterization and exploitation of newly discovered P450s is limited by two major factors. Firstly, P450-catalyzed reactions depend on the consecutive delivery of two electrons to the heme iron by redox partner proteins. Most P450 systems consist of several proteins: P450 itself and one or two additional redox partners, such as ferredoxins/flavodoxins and flavin-containing reductases.[4] The identification of physiological redox partners of P450s is often challenging, because the encoding genes are rarely located in the same operon or in close proximity to the corresponding cytochrome P450 genes. Secondly, there is limited knowledge about the physiological functions of most P450s, thus there is no information about their potential substrates.

Several approaches have been described for identifying the substrates of P450s of unknown physiological function and/or substrate scope ("orphan P450s"). One of the approaches aims at the identification of (not necessarily physiological) substrates based on the screening for the characteristic spectral shift upon substrate binding (the so-called type I spectrum) against very large libraries of organic compounds.[5] However, some P450s reportedly do not exhibit a type I spectrum upon substrate binding.[6] Furthermore, induction of a type I spectrum through binding of a compound to a P450 is no guarantee that the bound compound is converted by the enzyme.[7] These limitations restrict the suitability of this method for identifying new catalytic activities of P450s.

Another approach, aimed at the identification of physiological substrates of P450s, employs metabolomics to detect compounds that are only formed in the host cell in the presence of a particular P450. An inherent limitation in this approach is that catalytic activity is detected only towards compounds occurring in the host cell.[8]

An interesting approach has been described recently for CYP102A1 from *Bacillus megaterium* (BM3) and ten of its mutants.[9] By employing five known terpenoid substrates of the enzyme as probes, the authors were able to predict activity for other structurally closely related compounds. Obviously this method requires knowledge about the already identified substrates to predict new activities of a P450. It is very well suited to distinguishing differences in stereo- and regioselectivity of mutants, rather than to provide a basis for initial substrate

[a] *Dipl.-Chem. C. von Bühler, M.Sc. P. Le-Huu, Prof. Dr. V. B. Urlacher*
Institute of Biochemistry, Heinrich-Heine University Düsseldorf
Universitätsstrasse 1, 40225 Düsseldorf (Germany)
E-mail: Vlada.Urlacher@uni-duesseldorf.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300271.

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

screening. Additionally, because of the colorimetric detection used in this method, it is only applicable for probes containing a methoxy group.

As we aspire to discover novel natural and non-natural activities of uncharacterized P450s, we have developed a semirational method for the identification of putative substrates. This method combines reconstitution of P450 activity with suitable redox partners, clustered substrate screening, and gas chromatography/mass spectrometry (GC/MS) or liquid chromatography/mass spectrometry (LC/MS) analysis of possible products. With this combined approach, data about product formation and distribution can be obtained. Hence, it is possible to elucidate the chemo- and regioselectivity of the investigated P450 in primary substrate screening.

In order to overcome the problem of unidentified physiological redox partners, we established a set of redox proteins originating from prokaryotes and eukaryotes; these can be applied to support the activity of uncharacterized P450s.[7,10]

Herein we describe the use of cluster screening for the investigation of two poorly characterized or completely uncharacterized enzymes from the CYP154 family: CYP154E1 from *Thermobifida fusca* YX and CYP154A8 from *Nocardia farcinica* IFM 10152.

CYP154E1 was cloned, heterologously expressed, and purified in our laboratory for the first time, and several nonphysiological redox partners were tested for electron transfer to this P450[10] with a substrate identified in the actual study. For the second enzyme, CYP154A8, O-demethylation and *ortho*-hydroxylation of the medium-sized isoflavonoid formononetin has already been reported.[11] All annotated CYP154 enzymes are from actinomycetes, and therefore might be interesting regarding formation of secondary metabolites.[12] The CYP154 family also attracted our attention because of the broad spectrum of characterized substrates that have been described for enzymes of this family. These substances belong to different chemical classes and have very different molecular weights and sizes: CYP154C5 from *N. farcinica* IFM 10152 was found to hydroxylate testosterone;[13] CYP154H1 from *T. fusca* YX catalyzes side-chain hydroxylation of small aromatic molecules like ethylstyrene, as well as S-oxidation of aromatic thioethers to their corresponding sulfoxides and sulfones; the substrates of CYP154C1 from *Streptomyces coelicolor* A3(2) are the 12- and 14-membered macrocyclic lactones YC-17 and narbomycin.[14] A metabolomics approach shed light on the rather uncommon cyclization reaction with the endogenous substrate catalyzed by CYP154A1 from *S. coelicolor* A3(2).[15] These observations suggest high substrate diversity for the CYP154 family. The question is, how similar (or different) are the catalytic properties of these enzymes? According to sequence similarity analysis, the two chosen candidates of the CYP154 family (41% identity) belong not only to different subfamilies but also to different branches of the phylogenetic tree (Figure 1). Hence, biocatalytic and structural diversity between the enzymes was expected to be high. Thus, by using the clustered library we aimed to find out whether this low sequence similarity is reflected in different substrate spectra of the enzymes.

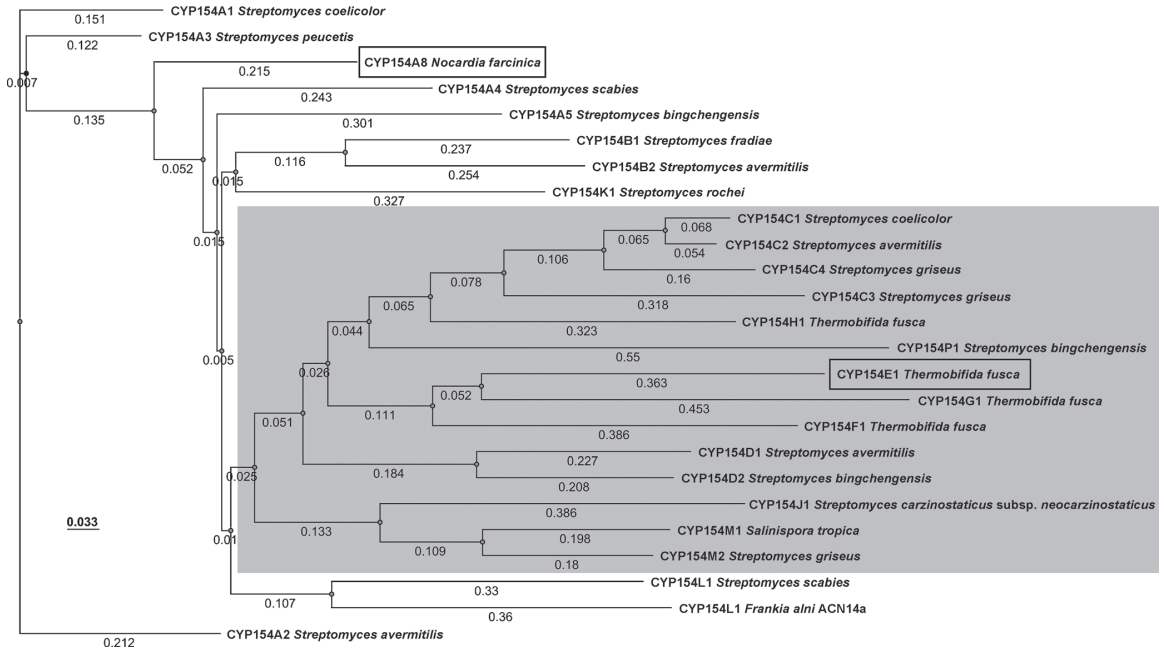


Figure 1. Phylogenetic tree of all available CYP154 family sequences (http://drnelson.uthsc.edu/bacterial.P450s.2011.htm). CYP154E1 from *T. fusca* and CYP154A8 from *N. farcinica* are marked with boxes. Alignment and tree building were performed with the Unipro UGENE bioinformatics suite.[1]

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

Results

Expression, purification, and spectral characterization of CYP154E1 and CYP154A8

Both P450 genes were cloned into the expression vector pET-28a(+) carrying an N-terminal His_6 tag. The enzymes were expressed in *Escherichia coli* BL21(DE3) in soluble form, and purified by immobilized metal (nickel) affinity chromatography (IMAC) and size-exclusion chromatography. The protein yields, calculated from CO-difference spectra, were 510 and 120 nmol L^{-1} for CYP154E1 and CYP154A8, respectively. For both proteins the characteristic Soret bands at 451 nm (CYP154E1) and 447 nm (CYP154A8) for the ferrous heme–CO complex were observed (Figure S1 in the Supporting Information). During CO-difference spectra measurements it became clear that CYP154A8 is not as stable as CYP154E1. Upon reduction, rapid formation of a shoulder peak at 420 nm was observed.

Identification of electron transfer partners

For reconstitution of activity, CYP154E1 and CYP154A8 require electron transfer proteins to supply the electrons. As the natural redox partners of both P450s have not been identified, we tested a set of nonphysiological redox partners known to interact with other P450s of bacterial origin, as well as hydrogen peroxide for direct oxidation of the heme iron by the "peroxide shunt". Among the tested redox partners were putidaredoxin and putidaredoxin reductase from *Pseudomonas putida*, cytochrome P450 reductase (CPR) from *Candida apicola* and the flavodoxin YkuN from *Bacillus subtilis* in combination with *E. coli* flavodoxin reductase (FdR). The highest conversion of test substrate **13** was observed with the YkuN/FdR combination; this was used for all subsequent experiments. This system proved to perform even better in our experiments than the published CYP154A8/Pdx/PdR system, under comparable conditions.[11] With hydrogen peroxide, no product formation was observed.

Library design for substrate cluster screening

In order to quickly assess the biocatalytic capabilities and compare the substrate spectra of CYP154E1 and CYP154A8, we compiled a library of organic molecules clustered into nine groups according to their chemical properties and size (Scheme 1).

Group I consisted of fatty acids of chain lengths C_7 to C_{12}, C_{14} (**1**–**7**), and C_{16} (**51**); these compounds are typical P450 substrates. In group II the position, type, and oxidation state of the hetero atom within the substrate was varied: alcohols **8**–**11** and the corresponding thioalcohol **12**. Group III comprised acyclic terpenoids (**13**–**21**) of varying chain length. The compounds in group III are of similar chain length to those of group II, and also contain oxygen atoms. The methyl side chains make these substrates more sterically demanding, and the carbon–carbon double bonds lead to a more rigid structure (less flexibility in the backbone). Furthermore, because of the double bonds, epoxidation is a second possible reaction pathway, in addition to hydroxylation of the activated allylic carbon atoms.

The compounds in group IV were highly methylated saturated fatty acids and alcohols. These do not possess the reactive allylic position (as for members of group III), but the methyl side chains lead to higher steric demand. The hetero function was limited to alcohols (**23**, **26**), carboxylic acids (**22**, **25**), and a methyl ester (**24**). With group V, we investigated a set of aliphatic (**29**) and aromatic (**27**–**28**) thioethers with significantly shorter chain lengths and chemical functionalities different from those in the other groups.

Group VI extended the substrate range to include bicyclic substances: both aliphatic (**31**, **34**) and N-hetero (**37**) and O-hetero (**30**–**33**, **35**, **36**, **38**) compounds. Substrate size was further increased with a tricyclic alizarin (**39**, group VII), quinizarin (**40**, group VII), and tetracyclic compounds (group VIII): 5α-cholestane-3β-ol (**41**), cholesterol (**42**), 5α-cholestane-3β-one (**43**), testosterone (**44**), deoxycholic acid (**45**), vitamin D_3 (**46**), 5α-cholestane (**47**), and pergolide mesylate (**48**). Finally, group IX comprised two macrocyclic compounds: cembrene (**49**) and erythromycin (**50**).

Substrate cluster screening

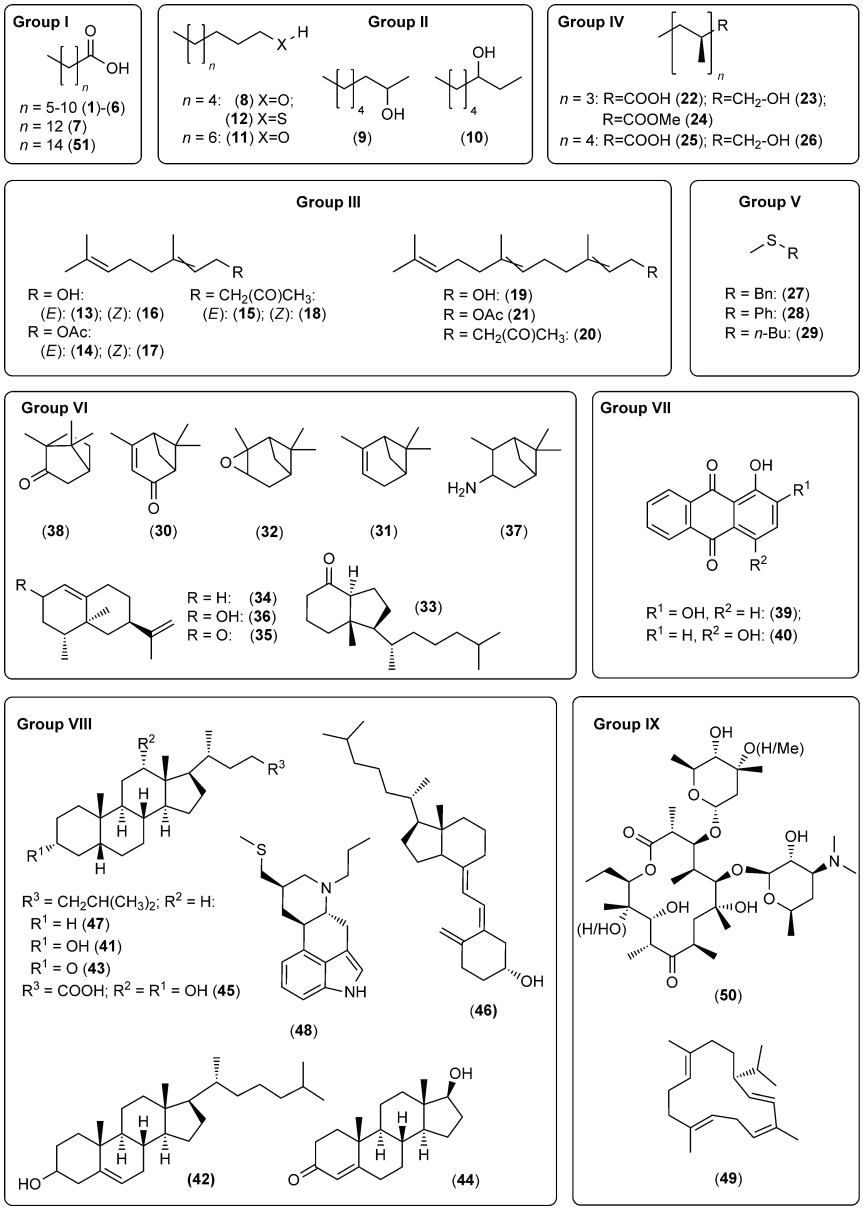
Substrate screening was performed independently for CYP154E1 and CYP154A8. After reconstitution of purified CYP154E1 with redox partners and the glucose dehydrogenase (GDH)-based cofactor regeneration system, substrates were screened in 500 μL reaction volume overnight, so that even products of substrates with very low conversion would be observable. An exception was made for group III, as after 24 h neither substrate nor product could be observed (possibly due to unspecific overoxidation); the negative control still contained the substrate (see Table 1). Hence, the reactions were only studied for 4 or 1 hour (farnesol).

Screening results with CYP154E1

In group I, the shortest fatty acid accepted by CYP154E1 was heptanoic acid (**1**, 29% conversion). For octanoic acid (**2**) conversion was 94%. For C_9–C_{14} fatty acids (**3**–**7**) quantitative conversion was achieved. Palmitic acid (**51**) was not accepted. The enzyme showed a clear preference for subterminal ω−1 hydroxylation in all cases, even though ω−2 hydroxy fatty acids were detected for all substrates other than **1** (Table 1). Product specificity ranged from exclusive formation of 6-hydroxy heptanoic acid from **1** to an ω−1/ω−2 ratio of 37:61 from tetradecanoic (myristic) acid (**7**).

In group II, octan-1-, -2-, and -3-ol (**8**, **9**, **10**) and the C_{10} homologue decan-1-ol (**11**) were quantitatively converted by CYP154E1. Product specificity clearly depended on chain length. The primary alcohols **8** and **11** led to only two products (ratios 50:50 and 10:90, respectively). In the case of the secondary alcohols **9** and **10**, the product pattern was more complex: **9** led to five products, of which the two main products

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org



Scheme 1. Compounds of the cluster library.

were formed with abundances of 36 and 32%. For **10**, six products were detected (main product, 56% of product peak area). Interestingly, for octane-1-thiol (**12**), which was oxidized to 68%, no monomeric oxidation product was observed: dimeric dioctyl disulfide was identified as the sole reaction product.

In group III, nerol (**16**) and geraniol (**13**) were converted quantitatively. Conversion of the derivatives geranyl-/neryl acetone (**15**/**18**), which were tested as a mixture of isomers, and farnesyl acetone (**22**) was lower (93 and 22%, respectively). Conversion of farnesol (**19**) reached 72% after 1 h. Product specificity was best for **16**: 84% 8-hydroxynerol was formed. The corresponding 8-hydroxy product was also formed from **13**, but it only appeared as a minor product (43%). Product specificity was lower for the conversion of **22** (seven products) and **15**/**18** (nine products). Conversion of acetic acid esters of geraniol (**14**) and nerol (**17**) only resulted in the formation of 8-hydroxygeraniol, after auto hydrolysis of the substrate in the aqueous reaction medium.

All compounds in group IV were recognized by CYP154E1 as substrates: 2,4,6-trimethyloctan-1-ol (**23**), 2,4,6,8-tetramethyl decanoic acid (**25**) and 2,4,6,8-tetramethyl decan-1-ol (**26**) were quantitatively converted; 2,4,6-trimethyloctanoic acid (**22**) and its methyl ester (**24**) were transformed with conversions of 90 and 96%, respectively. Product selectivity was highest for **22** (only two products, ratio 9:1). Product formation for all other compounds in this group was less specific: three products for **25**, and four for **24**, **23** (Figure 2A), and **26**.

All compounds of group V were identified as substrates and were oxidized to different extents (Table 1). For butyl methyl sulfide (**29**) the only product was the corresponding sulfoxide. The main product of phenyl methyl sulfide (**28**) was again the sulfoxide, but the corresponding sulfone was also detected (32%). Benzyl methyl sulfide (**27**) was transformed to the corresponding sulfoxide (12%), sulfone (19%), and a third main product (69%), which was identified by its mass spectrum to be dibenzyl sulfone.

Group VI had six hits: conversions for monoterpene verbenone (**30**) and α-pinene (**31**) were 18 and 79%, respectively; product specificity was relatively low (five and four products, respectively). Autoxidation could be excluded by comparison with negative controls (lacking P450). Grundmann's ketone (**33**), the only non-terpenoid in the group, was converted quantitatively with a single product being formed. By compari-

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

Table 1. Overview of the substrate cluster screening results: Conversion and product distribution (in brackets) were calculated from GC/MS and LC/MS peak areas. Names are provided for identified metabolites. Products of **9** and **23** are named according to Figure 2.

		Conversion [%] CYP154E1	CYP154A8	Product identification and distribution [%] (CYP154E1/CYP154A8)
1	heptanoic acid	29	0	6-hydroxy heptanoic acid (100/–)
2	octanoic acid	94	100	7-hydroxy octanoic acid (76/97), 6-hydroxy octanoic acid (24/3)
3	nonanoic acid	100	100	(9/–), 8-hydroxy nonanoic acid (81/95), 7-hydroxy nonanoic acid (10/5)
4	decanoic acid	100	100	(5/–), 9-hydroxy decanoic acid (91/78), 8-hydroxy decanoic acid (2/22), (2/–)
5	undecanoic acid	100	91	(12/–), 10-hydroxy undecanoic acid (66/69), 9-hydroxy undecanoic acid (17/31), (6/–)
6	dodecanoic acid	100	93	(6/–), 11-hydroxy dodecanoic acid (75/81), 10-hydroxy dodecanoic acid (3/19), (16/–)
7	tetradecanoic acid	100	81	13-hydroxy tetradecanoic acid (61/66), 12-tetradecanoic acid (37/34), (1/–), (1/–)
8	octan-1-ol	100	100	8-hydroxyoctan-2-one (50/36), octane-1,7-diol (50/64)
9	octan-2-ol	100	100	octane-2,7-dione (P3) (10/13), P2 (32/27), (15/–), P1 (36/60), (7/–)
10	octan-3-ol	100	100	(10/–), (9/–), (15/2), octane-2,3-diol (8/18), (56/80), (6/–)
11	decan-1-ol	100	100	(10/10), (90/85), (–/5)
12	octane-1-thiol	68	0	dioctyl disulfide (100/–)
13	geraniol	100[a]	100[a]	8-hydroxy geraniol (43/73), (57/13) (–/13)
15	geranyl acetone	93[a]	97	(6/–), (–/10), (8/–), (22/66), (22/–), (13/21), (–/3), (4/–), (5/–), (19/–), (2/–)
16	nerol	100[a]	100[a]	8-hydroxy nerol (84/96), (16/4)
18	neryl acetone	93[a]	67	(6/–), (–/10), (8/–), (22/66), (22/–), (13/21), (–/3), (4/–), (5/–), (19/–), (2/–)
19	farnesol	72[b]	18	(25/–), (10/–),(23/80),(43/20)
20	farnesyl acetone	22[a]	0	(12/–), (19/–), (35/–), (9/–), (10/–), (7/–), (9/–)
22	2,4,6-trimethyloctanoic acid	90	0	(89/–), (11/–)
23	2,4,6-trimethyloctan-1-ol	100	100	P1 (8/–), PI (–/4), P2/II (45/21), PIII (–/53), P3/IV (43/3), PV (–/19), P4 (4/–)
24	2,4,6-trimethyloctanoic acid methyl ester	96	75	(9/–), (83/86), (2/27) (10/–)
25	2,4,6,8-tetramethyl decanoic acid	100	77	(11/14), (83/86), (6/–)
26	2,4,6,8-tetramethyl decan-1-ol	100	95	(11/17), (52/72), (25/11), (12/–)
27	benzyl methyl sulfide	58	69	benzyl methyl sulfoxide (12/31), benzyl methyl sulfone (19/54), dibenzyl sulfone (69/15)
28	phenyl methyl sulfide	23	38	methyl phenyl sulfoxide (68/41), methyl phenyl sulfone (32/59)
29	butyl methyl sulfide	31	71	methyl butyl sulfoxide (100/21), methyl butyl sulfone (–/79)
30	verbenone	18	0	(28/–), (10/–), (9/–), (24/–), (29/–)
31	α-pinene	79	0	(15/–), (28/–), (33/–), (24/–)
33	Grundmann's ketone	100	53	25-hydroxy Gundmann's ketone (100/100)
35	nootkatone	100	0	13-hydroxy nootkantone (100/–)
36	nootkatol	79	0	(31/–), (69/–)
38	camphor	5	0	6-hydroxy camphor (100/–)
48	pergolide mesylate	4	0	(36/–), (64/–)

[a] Measurements were stopped after 4 h. [b] Measurements were stopped after 1 h.

son to an authentic standard this was identified as 25-hydroxy Grundmann's ketone. In contrast to the unsubstituted homologue valencene (**34**), the oxidized sesquiterpenes nootkatol (**36**) and nootkatone (**35**) were transformed with 79 and 100% conversion, respectively. The single oxidation product of **35** was identified as 13-hydroxy-nootkatone with an authentic standard.[16] The two oxidation products of **36** (ratio 31:69) did not match those from nootkatone (**35**) in terms of retention time and mass spectrum. Camphor (**38**) was also identified as a substrate for CYP154E1, although conversion was low (5%) and with a single reaction product. The EI mass spectrum did not match hydroxylation at the 5-position (as described for CYP101 enzymes) but displayed high similarity (match factor 791) to a published mass spectrum of 6-hydroxy camphor.[17]

The largest substance accepted by CYP154E1 as a substrate was pergolide mesylate (**48**, conversion 4%). Two single oxidation products were identified by the characteristic *m/z* 331 ion $[M+16+H]^+$ in their ESI-MS spectra. High resolution MS measurements confirmed the formula (*m/z* calcd for $C_{19}H_{27}N_2OS$: 331.1839 $[M+H]^+$, found: 331.1831). The MS/MS spectra of both peaks were identical, and initial water elimination ($\Delta m/z$ −18) and subsequent loss of a CH_2S-fragment ($\Delta m/z$ −46) is more consistent with hydroxylation than sulfoxidation (Figure S2). With other compounds from groups VII–IX, CYP154E1 did not demonstrate any product formation.

Screening results with CYP154A8

From group I C_8–C_{14} fatty acids were accepted by CYP154 with relatvely high conversion (>80%, Table 1). Palmitic acid (**51**) was not converted. The preference of CYP154A8 to form ω−1 over ω−2 ranged from 97:3 (octanoic acid, **2**) to 34:66 (tetradecanoic (myristic) acid, **7**; Table 1).

For group II, substrate scope was comparable to that for CYP154E1. Alcohols octan-1-ol (**8**), octan-2-ol (**9**), octan-3-ol (**10**; Figure 2C–E) and decan-1-ol (**11**) were very good substrates (quantitative conversion). In the first step, **8** is converted into octane-1,7-diol, and then further oxidized to 8-hydroxyoctan-2-one (Figure 2C). Conversion of **9** resulted in three product peaks with 13, 27, and 60% relative abundance for

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

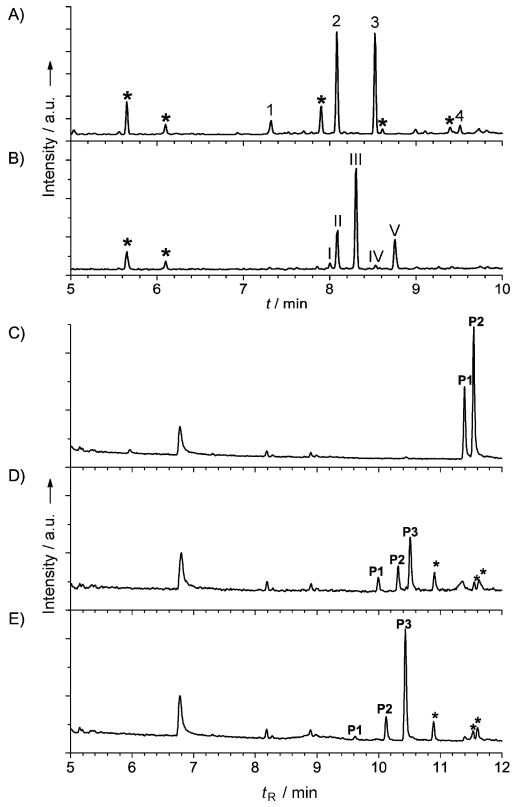


Figure 2. GC/MS chromatograms of the conversion of 2,4,6-trimethyloctan-1-ol with A) CYP154E1 (Arabic numerals), and B) CYP154A8 (Roman numerals), and conversion of C) octan-1-ol, D) octan-2-ol, E) octan-3-ol with CYP154A8. Background peaks are marked with an asterisk.

octan-2,7-dione, and unidentified products P2 and P1, respectively (Figure 2D). Database identification was only partly successful. As the retention times differed significantly from those of the products of **8**, a similar mechanism of secondary hydroxylation and overoxidation is possible, but not in combination with the carbon atom at positions C1 and C7. Compound **10** was equally well accepted and led to three products (ratio 2:18:80; Figure 2E). P2 was identified as octane-2,3-diol, but no database entry was found for P1 or P3, although (judging from retention times and mass spectra) they were not identical to the products of the conversion of **8** or **9**. The longer-chained **11** led to one main and two secondary products (10, 85, and 5%). Complete baseline separation was not possible, and hence only the last one could be identified (decan-1-al).

From group III only four substrates were accepted: geraniol (**13**) was converted to three products (73, 14, and 14% relative abundance). Baseline separation was not possible, thus only the major product could be identified (8-hydroxygeraniol). Diastereomer nerol (**16**) was converted into two products (ratio 4:96). The isomers geranyl- (**15**) and neryl acetone (**18**) were also recognized as substrates by CYP154A8, with a preference for **15** over **18** (97 and 67% conversion, respectively). The product distribution was less complex than for CYP154E1: four products (relative abundance 10:66:21:3). Transformation of farnesol (**19**) led to two products (ratio 1:5), with rather low conversion (18%).

Compounds in group IV proved to be substrates for CYP154A8 (as for CYP154E1). 2,4,6,-Trimethyloctanoic acid (**22**) was not accepted, but the corresponding alcohol 2,4,6-trimethyloctan-1-ol (**23**) was converted (100%) into five products (Figure 2B). Product 2/II was produced by both enzymes ($\omega-1$ hydroxylated compounds). The other products could not be identified. The methyl ester (**24**) of **22** was readily accepted and converted (75%) into two products (73 and 27% abundance). Though neither led to hits in the National Institute of Standards and Technology (NIST) database, the mass spectrum of the first indicated $\omega-1$ hydroxylation. The longer 2,4,6,8-tetramethyldecanoic acid (**25**) was converted (77%) into two products. The preference for $\omega-1$ hydroxylation here was reversed (14:86), compared to **24**. Conversion of the corresponding 2,4,6,8-tetramethyldecan-1-ol (**26**) was 95%, and resulted in the same product distribution (83:17) as for homologue **25**, but with reversed preference regarding $\omega-1$ hydroxylation.

The compounds in group V proved to be better substrates for CYP154A8 than for CYP154E1. Phenyl methyl sulfide (**28**) had the lowest conversion (38%), with the corresponding sulfoxide and sulfone in a ratio of 41:59. The benzyl derivative benzyl methyl sulfide (**27**) was converted (69%) into sulfoxide and sulfone products (15 and 54%, respectively), and into a third product, which was identified as dibenzyl sulfone (31%). The last compound in this group, the aliphatic butyl methyl sulfide (**29**), showed conversion of 71% and formation of the corresponding sulfoxide (21%) and sulfone (79%).

Of group VI compounds, only Grundmann's ketone (**33**) was identified as a substrate for CYP154A8. The conversion was 53%, with 25-hydroxy Grundmann's ketone as the single product. Group VII–IX compounds proved not to be substrates of CYP154A8.

Determination of substrate binding constants

For correlation of substrate binding and substrate conversion by the two CYP154 enzymes, we measured UV spectral changes upon titration with a few selected compounds from the library: geraniol (**13**), nerol (**16**), fatty acids C_9–C_{11} (**3**–**5**), and the macrocycle cembrene (**49**). Dissociation constant (K_d) values were calculated from the spectral data.

Both **13** and **16** showed perturbation of the heme spin state of both P450s (Figures S3 and S4), with a type I shift of the Soret band. For both CYP154 enzymes, affinity for these acyclic terpenes was high, whereas affinity for fatty acids (**3**–**5**) was much lower (Table 2). CYP154A8 demonstrated lower affinity than CYP154E1 with all three fatty acids.

Cembrene (**49**) did not induce a spin state shift in CYP154E1, so K_d could not be determined. In contrast, **49** induced a type I spectral change in CYP154A8 (K_d 920 μM). This suggests that CYP154A8 has a larger binding pocket than CYP154E1.

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

Table 2. Dissociation constants calculated from titration of test compounds with CYP154A8 and CYP154E1.

Compound	K_d [μM]	
	CYP154A8 ± SE	CYP154E1 ± SE
geraniol (**13**)	2.5 ± 0.2	6 ± 1
nerol (**16**)	17 ± 1	6 ± 0.5
nonanoic acid (**3**)	133 ± 6	38 ± 6
decanoic acid (**4**)	51 ± 9	38 ± 3
undecanoic acid (**5**)	96 ± 8	30 ± 3
cembrene (**49**)	919 ± 437	–

Discussion and Conclusion

The main goal of this study was to develop an effective primary substrate screening method for uncharacterized P450s. This method was verified by screening with CYP154E1 and CYP154A8.

The screening results obtained with the substances from group I showed that both CYP154E1 and CYP154A8 were able to oxidize C_8–C_{14} saturated fatty acids. Heptanoic acid (C_7) was accepted by CYP154E1 but not by CYP154A8. Palmitic acid (C_{16}) was not converted by either enzyme. Both monooxygenases predominantly hydroxylated fatty acids at the ω−1 position. As minor reaction products, ω−2 hydroxy fatty acids were detected for all substrates, except C_7, with CYP154E1.

The primary alcohols from group II were accepted by both enzymes. All octanols were quantitatively converted by both enzymes, but with different product selectivities (Table 1). CYP154A8 seems to be more regioselective in these reactions. Some products of these reactions were identified as ketones. Overoxidation by P450s with alcoholic substrates has been reported for different kinds of substrates[18] and has been extensively studied.[19]

Although both CYP154E1 and CYP154A8 demonstrated similar activities towards geraniol and nerol (group III), farnesol (**19**, C_{12}) was converted much better by CYP154E1. This was also the case for farnesyl acetone (**20**). Again, the trend was that CYP154A8 showed higher regioselectivity. Comparable activities of P450s with linear terpenoids have been found for monooxygenases from plants expressed in insect cells,[20] as well as for a chimera of CYP102A1 (BM3) with CYP4C7.[21] As CYP154E1 showed particularly high regioselectivity towards the 8-hydroxy products of geraniol and nerol, it might be an alternative to the BM3–CYP4C7 chimera.

For group IV compounds, the only difference in activity between the monooxygenases was for 2,4,6-trimethyloctanoic acid (**22**), which was accepted only by CYP154E1. Generally, CYP154A8 was again slightly more regioselective than CYP154E1. In comparison with linear fatty acids of the same carbon atom chain length, it is apparent that the product spectrum of CYP154E1 for the branched fatty acids is broader. To date, the only P450s known to oxidize these five branched fatty acid derivatives are BM3 and a triple mutant thereof.[22] Interestingly, these monooxygenases were highly regioselective in these reactions and produced only ω−1 hydroxylated products from the tested branched substrates. In contrast, CYP154E1 and especially CYP154A8 hydroxylated branched fatty acid derivatives with different regioselectivity, and thus provide access to other products. We recently showed that compounds and their hydroxylated products of this type can be used as synthones for the chemoenzymatic synthesis of the macrolactone borrelidin.[23] As the regioselectivity of CYP154E1 and CYP154A8 differs from that of BM3, synthetic routes to different derivatives can be realized.

Compounds in the above discussed groups are mainly linear rod-shaped molecules. The methyl sulfides in group V possess an aromatic (phenyl or benzyl) or an aliphatic (*n*-butyl) group. Conversion of these substrates was observed with both CYP154A8 and CYP154E1. As the aliphatic butyl methyl sulfide was also converted, the aromatic ring seems not to be required for substrate recognition, as could have been deduced from the only two published substrates of CYP154A8, 7-ethoxycoumarin and the isoflavonoid formononetin.[11] Enzymatic sulfoxidation of dialkyl sulfides is relatively rare: so far it has only been described for chloroperoxidase from *Caldariomyces fumago,* flavin-containing monooxygenase from *Methylophaga sp.* SK1, and for two Baeyer–Villiger monooxygenases.[24] For aryl alkyl sulfides **27**–**29**, formation of the respective sulfoxide and subsequent oxidation to the sulfone was observed for both P450s; this is consistent with the observation for CYP154H1.[25] Formation of dimeric reaction products has also been observed for CYP154H1. We detected the formation of dibenzyl sulfone from benzyl methyl sulfide (**27**) with CYP154E1, and at a much higher concentration with CYP154A8.

Among the compounds of group VI (increased steric demand), six were accepted by CYP154E1 but only Grundmann's ketone (**33**) was a substrate for both P450s. Grundmann's ketone is a product of ozonolysis of vitamin D_3; its selective 25-hydroxylation is part of a synthetic route to vitamin D_3 metabolites and derivatives. Chemical synthesis of 25-hydroxy Grundmann's ketone by using ruthenium catalysts has been reported.[26] The workup of the ruthenium-catalyzed reaction is laborious, and the remaining Grundmann's ketone (**33**) has to be recovered by column chromatography; the overall yield was reported as 49%.[26] CYP154E1 offers an alternative synthetic route, as it produces the desired 25-hydroxy product with absolute selectivity. To the best of our knowledge CYP154E1 and CYP154A8 are the only P450 monooxygenases known to accomplish this synthetic task.

The 6-hydroxylating activity of CYP154E1 with camphor has also been observed for CYP105D1 ($P450_{soy}$) from *Streptomyces griseus*.[27] CYP154A8 did not accept camphor as a substrate.

In order to determine whether large molecules such as macrolides or steroids can be accepted by CYP154E1 and CYP154A8, we tested compounds from groups VII, VIII, and IX. The only compound of group VIII that was identified as a substrate of CYP154E1 was pergolide mesylate (**48**). This substance, derived from ergot alkaloids, has been used in the treatment of Parkinson's disease and is known to be a potent inhibitor of human P450s especially CYP2D6 (IC_{50} 0.08 μM).[28] CYP3A4 is assumed to play an important role in the detoxification of pergolide, as it has been shown to catalyze the dealky-

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

lation of other ergot derivatives.[29] Another pathway in human pergolide metabolism results in the formation of the corresponding sulfoxide and sulfone. CYP154E1 forms different metabolites starting with pergolide.

Neither the tested anthraquinones nor the steroids were converted by the two CYP154 enzymes. Among the compounds with the highest steric demand in group IX were the nonpolar cyclic diterpene cembrene (**49**), which has a 14-membered ring, and the 14-membered macrolide antibiotic erythromycin (**50**). With its several oxygen functionalities and attached sugar side chains, **50** is more polar than **49**. Remarkably, a type I binding spectrum was observed for CYP154A8 with cembrene. The inability of this monooxygenase to convert cembrene shows that even though compounds of this size and molecular weight fit into its active site and induce a type I binding spectrum, product formation is not guaranteed; hence, a method based on detection of reaction products is needed to reliably identify substrates.

Unlike the recently published method of Furuya et al. for the identification of P450 substrates,[30] which is based on a small collection of generally known P450 substrates with no structural relationship, our approach is aimed at the systematic variation of the compounds used to probe the P450 activity. The systematic chemical relationship of our library compounds enabled us not only to identify new substrates for a P450 but, even more importantly, also to study the structural requirements which have to be satisfied for a compound to be converted. Remarkably, cluster screening distinguished subtle differences in activity and selectivity of enzymes as closely related as two in the same P450 family. Although this cluster library provides only limited coverage of the total chemical space, it was possible to identify useful biocatalysts for the following synthetic tasks: regioselective oxidation of Grundmann's ketone, sulfoxidation of a dialkyl sulfide, and production of a novel metabolite of the alkaloid pergolide mesylate.

The current cluster library can be extended or adapted by variation of the library compounds, depending on the preferred P450 activity. This renders the cluster screening approach suitable for identification of the substrate spectrum of any new P450 and hence might prove to be a universal tool in P450 enzymology and biocatalyst development.

Experimental Section

Strains, expression vectors, enzymes, and chemicals: *E. coli* strain DH5α (for molecular cloning), the expression strain *E. coli* BL21(DE3), and the expression vector pET-28a(+) were purchased from Novagen/EMD Millipore (Darmstadt, Germany). Phusion DNA-Polymerase, restriction endonucleases, and T4 DNA ligase were obtained from Fermentas (St. Leon-Rot, Germany) and New England Biolabs (Beverly, MA). Glucose dehydrogenase (GDH-2) from *B. megaterium*, which was used for NADPH regeneration, was cloned into pET-22b(+), expressed, and partially purified as described previously.[31]

All chemicals were purchased from commercial sources unless otherwise stated. The branched fatty acid substrates were available at our institute.[32] Grundmann's ketone and its 25-hydroxy derivative were kindly provided by Prof. Miguel Ferrero (Universidad de Oviedo, Oviedo, Spain). Prof. Rita Bernhardt, (Universität des Saarlandes, Saarbrücken, Germany) kindly provided 13-hydroxynootkatone. pET-11-derived plasmids containing genes encoding CYP154E1 (GenBank: AAZ57009) from *T. fusca* YX and CYP154A8 (GenBank: YP_118504) from *N. farcinica* IFM 10152 were kindly provided by Dr. Akira Arisawa (Mercian Cooperation, Iwata, Japan).

Molecular cloning: General molecular biology manipulations and microbiological experiments were carried out by standard methods.[33] The genes encoding CYP154E1 and CYP154A8 were amplified with high-fidelity Phusion DNA Polymerase according to the manufacture's protocol using specific primers, and, after subsequent enzymatic restriction digest, ligated into pET-28a(+) between NdeI and EcoRI restriction sites. The resulting DNA constructs were sequenced by automated DNA sequencing (GATC-Biotech, Konstanz, Germany). These carried genes for N-terminally His_6-tagged proteins, and were named pET-154E1 and pET-154A8, respectively. Vectors encoding N-terminally His_6-tagged YkuN and FdR are described elsewhere.[7]

Protein expression and purification: *E. coli* BL21(DE3) cells were used for protein expression. A single colony from an agar plate was inoculated into lysogeny broth preculture (5 mL) supplemented with kanamycin (35 mg mL^{-1}) and grown at 37 °C for 4 h. Terrific Broth main cultures (400 mL) were inoculated with preculture and grown under the same condition until OD_{600} = 0.6–0.8. Protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (100 μM), and the mixture was supplemented with $FeSO_4$ (125 μM) for iron-containing enzymes and 5′-aminolevulinic acid (80 μg mL^{-1}) for P450 expression, prior to incubation (16 h, 25 °C, 140 rpm). Cells were then harvested by centrifugation (17000 *g*, 15 min), resuspended in potassium phosphate buffer (50 mM, pH 7.5) containing phenylmethanesulfonyl fluoride (PMSF, 100 μM) and disrupted by sonication on ice. All enzymes except GDH were purified by nickel affinity chromatography on two consecutive HisTrap FF 5 mL columns (GE Healthcare), desalted with a HiPrep 26/10 Desalting column (GE Healthcare), and stored at −20 °C.

Enzyme assays and spectroscopic methods: The concentrations of CYP154A8 and CYP154E1 were determined based on the CO-difference spectra, as described previously ($\varepsilon_{450-490}$ = 91 mM^{-1} cm^{-1}).[34] Upon reduction with sodium dithionite (0.05 M), five spectra were recorded (400–500 nm). For spectroscopic concentration determination of YkuN (ε_{461} = 10.01 mM^{-1} cm^{-1})[35] and FdR (ε_{456} = 7.1 mM^{-1} cm^{-1})[36] published extinction coefficients were used.

Binding spectra from low-spin (LS) to high-spin (HS) transition of the heme iron upon substrate binding were recorded on a Lambda 35 dual-beam spectrophotometer (PerkinElmer) equipped with tandem cuvettes (Hellma, Müllheim, Germany). Spectra were acquired from 330 to 700 nm (480 nm min^{-1}, slit width 1 nm). In each cuvette one chamber was filled with P450 (1 μM) in potassium phosphate buffer (100 mM, pH 7.5); the other contained buffer. In the sample cuvette substrate was titrated into the chamber containing enzyme, and an equal amount of solvent used to dissolve the substrate was added to the buffer-containing chamber. In the reference cuvette substrate was added to the buffer filled chamber and the enzyme chamber was supplemented with an equal amount of solvent only. The difference in absorbance between the LS peak (417 nm) and the HS peak (391 nm) was plotted against substrate concentration, and a hyperbolic equation was fitted to a nonlinear Levenberg–Marquardt regression algorithm. The $NADP^+$ reduction activity of GDH was assayed photometrically at 340 nm in multiwell plates. Each well contained $NADP^+$ (100 μM), glucose (100 mM), and sample (2.5 μL) in potassium phosphate

CHEMBIOCHEM
FULL PAPERS
www.chembiochem.org

buffer (250 μL; 100 mM, pH 7.5). Initial reaction velocities were calculated by linear fit of the first 10 s.

Reconstitution of enzyme systems for substrate screening: All reactions for substrate screening were performed in potassium phosphate buffer (500 μL; 100 mM, pH 7.5) containing P450 (1 μM), YkuN (10 μM), and FdR (1 μM). Enzymatic cofactor regeneration for NADPH was achieved with GDH (5 U) and glucose (20 mM) as co-substrate. In order to remove hydrogen peroxide (which might be formed due to uncoupling), catalase (60 U) was added to the reaction mixture. Substrates (100 μM) were added from stock solutions (10 mM) in ethanol or DMSO. The reaction mixture was incubated (25 °C, overnight) on a heat block or incubator in standard 2 mL plastic reaction tubes. The reaction mixtures were extracted with ethyl acetate or diethyl ether by vortexing for 5 min, and centrifugation for phase separation. To facilitate extraction in the case of acidic substrates, the reaction mixture was acidified by addition of hydrochloric acid (17.5 μL, 6 N). For basic compounds, potassium hydroxide (50 μL, 1 N) was added. Experiments were performed in duplicate. For samples containing fatty acids and branched derivatives, the extracted organic phase was dried with anhydrous $MgCl_2$, evaporated to dryness, and the residue was dissolved in *N*,*O*-bis(trimethylsilyl)trifluoroacetamide containing trimethylchlorosilane (1%), transferred to a GC/MS vial, and heated (1 h, 60 °C) prior to analysis.

Product analysis: Analysis of substrate screening was carried out by GC/MS on a QP-2010 Plus spectrometer (Shimadzu, Tokyo, Japan) with a FS-Supreme-5 fused silica column (30 m × 0.25 mm × 0.25 μm; Chromatographie Service GmbH, Langerwehe, Germany) and helium as the carrier gas (injector temperature 250–300 °C, ion source 200 °C, interface 285–300 °C). Spectra were recorded in scan mode (*m/z* 40–350; GC methods in Table S1). Products were identified by comparison of acquired mass spectra with the NIST 08 database or authentic standards.

Compounds **41**–**46**, **48**, and **50** were detected by HPLC on a Prominence/LCMS 2020 system (Shimadzu, Tokyo, Japan). After extraction, the evaporated organic phase was resuspended in ethanol and injected into a Chromolith RP-18e 100-4.6 column (Merck, Darmstadt, Germany). The sample was eluted in water with formic acid (0.1%) as solvent A and methanol as solvent B.

High-resolution MS and MS/MS measurements were performed with a UHR-TOF maXis 4G spectrometer (Bruker Daltonik, Bremen, Germany). Mass calibration was performed immediately before sample measurement.

Acknowledgements

We wish to thank Professor Dr. Miguel Ferrero, Universidad de Oviedo (Spain) and Prof. Dr. Rita Bernhardt, Universität des Saarlandes, Saarbrücken (Germany) for kindly providing us with authentic standards. Furthermore the authors wish to thank Dr. Marco Girhard, Heinrich-Heine University Düsseldorf (Germany) for helpful ideas and fruitful discussion of the manuscript, as well as Drs. Helmut Keck and Peter Tommes, Heinrich-Heine University Düsseldorf for the help with the HR-MS measurements.

Keywords: biocatalysis · cluster screening · CYP154s · cytochromes P450 · enzyme catalysis

[1] K. Okonechnikov, O. Golosova, M. Fursov, *Bioinformatics* **2012**, *28*, 1166 – 1167.
[2] D. R. Nelson, *Biochim. Biophys. Acta Proteins Proteomics* **2011**, *1814*, 14 – 18.
[3] V. B. Urlacher, M. Girhard, *Trends Biotechnol.* **2012**, *30*, 26 – 36.
[4] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta Gen. Subj.* **2007**, *1770*, 330 – 344.
[5] a) M. Lisurek, B. Rupp, J. Wichard, M. Neuenschwander, J. P. von Kries, R. Frank, J. Rademann, R. Kühne, *Mol. Diversity* **2010**, *14*, 401 – 408; b) S. Bleif, F. Hannemann, M. Lisurek, J. P. von Kries, J. Zapp, M. Dietzen, I. Antes, R. Bernhardt, *ChemBioChem* **2011**, *12*, 576 – 582.
[6] a) B. Simgen, J. Contzen, R. Schwarzer, R. Bernhardt, C. Jung, *Biochem. Biophys. Res. Commun.* **2000**, *269*, 737 – 742; b) B. Sielaff, J. R. Andreesen, *FEBS J.* **2005**, *272*, 1148 – 1159; c) J. B. Schenkman, H. Remmer, R. W. Estabrook, *Mol. Pharmacol.* **1967**, *3*, 113 – 123.
[7] M. Girhard, T. Klaus, Y. Khatri, R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2010**, *87*, 595 – 607.
[8] a) F. P. Guengerich, Q. Cheng, *Pharmacol. Rev.* **2011**, *63*, 684 – 699; b) F. P. Guengerich, Z. Tang, S. G. Salamanca-Pinzón, Q. Cheng, *Mol. Interventions* **2010**, *10*, 153 – 163.
[9] K. Zhang, S. El Damaty, R. Fasan, *J. Am. Chem. Soc.* **2011**, *133*, 3242 – 3245.
[10] M. Girhard, F. Tieves, E. Weber, M. S. Smit, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1625 – 1635.
[11] K.-Y. Choi, H.-Y. Park, B.-G. Kim, *Enzyme Microb. Technol.* **2010**, *47*, 327 – 334.
[12] D. C. Lamb, F. P. Guengerich, S. L. Kelly, M. R. Waterman, *Expert Opin. Drug Metab. Toxicol.* **2006**, *2*, 27 – 40.
[13] H. Agematu, N. Matsumoto, Y. Fujii, H. Kabumoto, S. Doi, K. Machida, J. Ishikawa, A. Arisawa, *Biosci. Biotechnol. Biochem.* **2006**, *70*, 307 – 311.
[14] L. M. Podust, Y. Kim, M. Arase, B. A. Neely, B. J. Beck, H. Bach, D. H. Sherman, D. C. Lamb, S. L. Kelly, M. R. Waterman, *J. Biol. Chem.* **2003**, *278*, 12214 – 12221.
[15] Q. Cheng, D. C. Lamb, S. L. Kelly, L. Lei, F. P. Guengerich, *J. Am. Chem. Soc.* **2010**, *132*, 15173 – 15175.
[16] T. T. B. Ly, Y. Khatri, J. Zapp, M. C. Hutter, R. Bernhardt, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 123 – 133.
[17] J. A. Marco, J. F. Sanz-Cervera, V. García-Lliso, N. Batlle, *Phytochemistry* **1997**, *45*, 755 – 763.
[18] a) M. Girhard, K. Machida, M. Itoh, R. D. Schmid, A. Arisawa, V. B. Urlacher, *Microb. Cell Fact.* **2009**, *8*, 36; b) K. C. Cheng, J. B. Schenkman, *J. Biol. Chem.* **1983**, *258*, 11738 – 11744; c) M. Akhtar, M. R. Calder, D. L. Corina, J. N. Wright, *Biochem. J.* **1982**, *201*, 569 – 580.
[19] A. D. Vaz, M. J. Coon, *Biochemistry* **1994**, *33*, 6442 – 6449.
[20] P.-H. Sung, F.-C. Huang, Y.-Y. Do, P.-L. Huang, *J. Agric. Food Chem.* **2011**, *59*, 4637 – 4643.
[21] C.-K. J. Chen, R. E. Berry, T. K. Shokhireva, M. B. Murataliev, H. Zhang, F. A. Walker, *J. Biol. Inorg. Chem.* **2010**, *15*, 159 – 174.
[22] M. Budde, M. Morr, R. D. Schmid, V. B. Urlacher, *ChemBioChem* **2006**, *7*, 789 – 794.
[23] M. Theurer, Y. El Baz, K. Koschorreck, V. B. Urlacher, G. Rauhut, A. Baro, S. Laschat, *Eur. J. Org. Chem.* **2011**, 4241 – 4249.
[24] a) A. Rioz-Martínez, G. de Gonzalo, D. E. Torres Pazmiño, M. W. Fraaije, V. Gotor, *Eur. J. Org. Chem.* **2010**, 6409 – 6416; b) A. Rioz-Martínez, M. Kopacz, G. de Gonzalo, D. E. Torres Pazmiño, V. Gotor, M. W. Fraaije, *Org. Biomol. Chem.* **2011**, *9*, 1337; c) S. Colonna, N. Gaggero, C. Richelmi, P. Pasta, *Trends Biotechnol.* **1999**, *17*, 163 – 168.
[25] A. Schallmey, G. den Besten, I. G. P. Teune, R. F. Kembaren, D. B. Janssen, *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1475 – 1485.
[26] J. Kiegiel, P. M. Wovkulich, M. R. Uskoković, *Tetrahedron Lett.* **1991**, *32*, 6057 – 6060.
[27] F. S. Sariaslani, L. R. McGee, M. K. Trower, F. G. Kitson, *Biochem. Biophys. Res. Commun.* **1990**, *170*, 456 – 461.
[28] M. A. Wynalda, L. C. Wienkers, *Drug Metab. Dispos.* **1997**, *25*, 1211 – 1214.
[29] a) R. Rauschenbach, H. Gieschen, M. Husemann, B. Salomon, M. Hildebrand, *Eur. J. Pharmacol. Environ. Toxicol. Pharmacol. Sect.* **1995**, *293*, 183 – 190; b) S. E. Ball, G. Maurer, M. Zollinger, M. Ladona, A. E. Vickers, *Drug Metab. Dispos.* **1992**, *20*, 56 – 63.
[30] T. Furuya, T. Nishi, D. Shibata, H. Suzuki, D. Ohta, K. Kino, *Chem. Biol.* **2008**, *15*, 563 – 572.

[31] W. Hilt, G. Pfleiderer, P. Fortnagel, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **1991**, *1076*, 298 – 304.
[32] M. Morr, V. Wray, J. Fortkamp, R. D. Schmid, *Liebigs Ann. Chem.* **1992**, 433 – 439.
[33] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, **2001**.
[34] T. Omura, R. Sato, *J. Biol. Chem.* **1964**, *239*, 2370 – 2378.
[35] a) 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; b) R. J. Lawson, C. von Wachenfeldt, I. Haq, J. Perkins, A. W. Munro, *Biochemistry* **2004**, *43*, 12390 – 12409.
[36] 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.

Received: April 30, 2013
Published online on September 23, 2013

CHEMBIOCHEM

Supporting Information

Cluster Screening: An Effective Approach for Probing the Substrate Space of Uncharacterized Cytochrome P450s

Clemens von Bühler, Priska Le-Huu, and Vlada B. Urlacher*[a]

cbic_201300271_sm_miscellaneous_information.pdf

Supporting Information

Supporting Data

Figure S1: Absorbance spectra of **A**: purified CYP154A8 and **B**: purified CYP154E1.The spectra of the oxidized enzymes were recorded in the absence of substrates. The characteristic Soret band was detected at 451 nm for CYP154E1 and 447 nm for CYP154A8. Insets represent the CO difference spectra recorded after reduction with sodium dithionite and saturation with gaseous carbon monoxide.

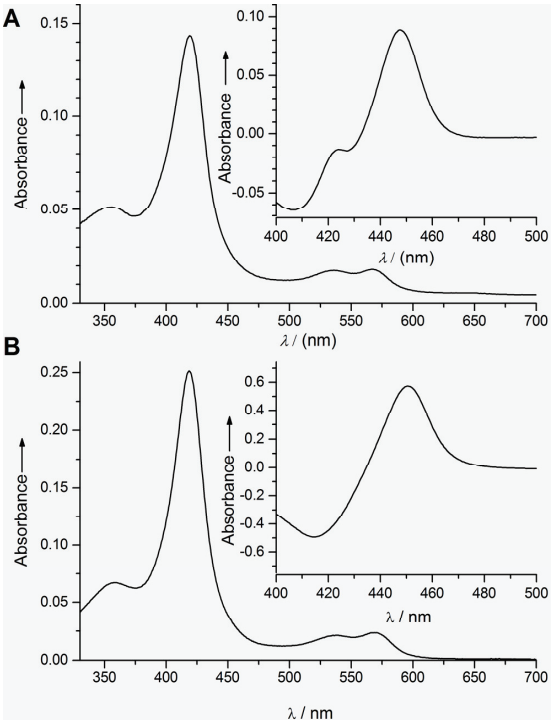


Figure S2: HR-MS product ion scan of the reaction products of **(48)**:

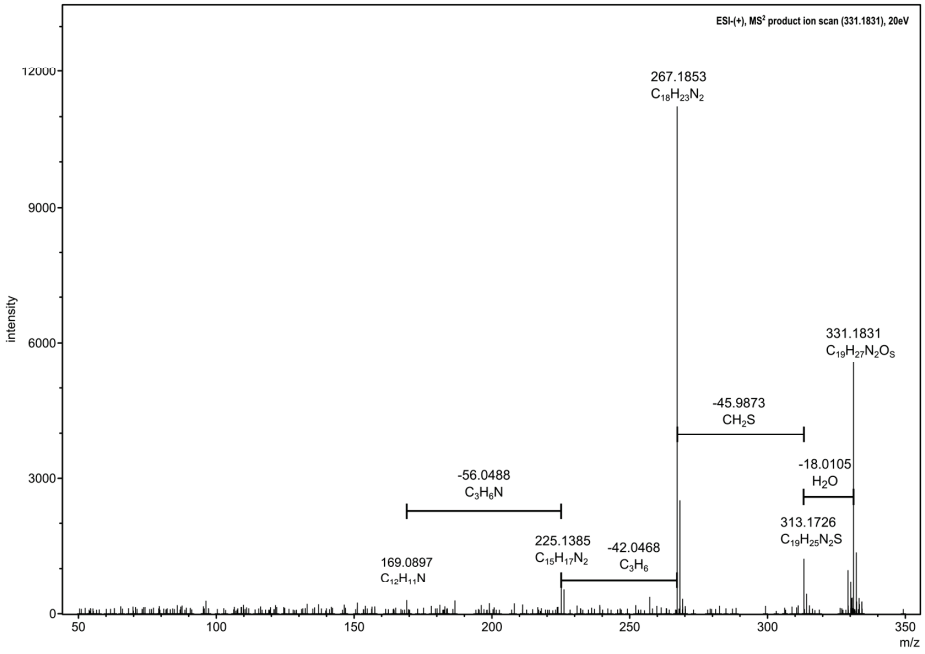


Figure S3 related to Table 2 Determination of binding constants of CYP154A8 **A)** and CYP154E1 **B)** with nerol **(16)**. Insets show the type I substrate binding spectrum. In the diagrams the absorbance difference (A_{390}-A_{420}) is plotted against substrate concentration.

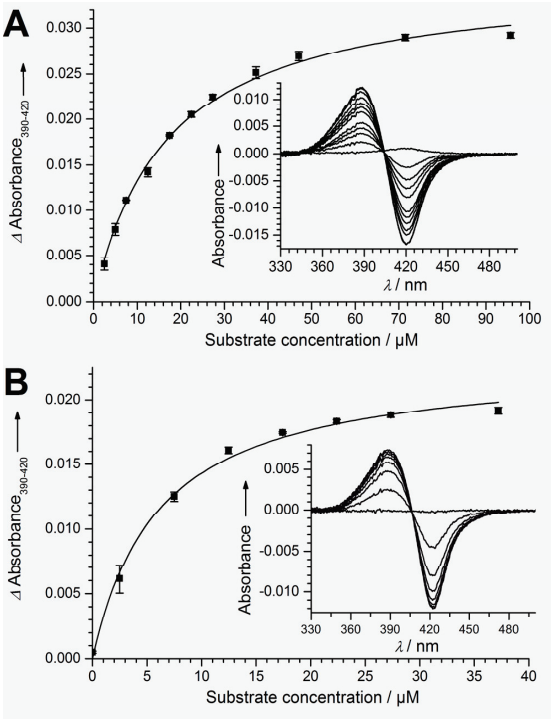
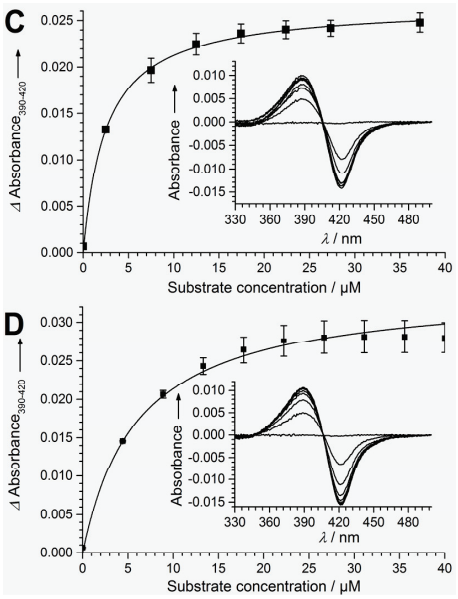


Figure S4: Determination of binding constants of CYP154A8 **C)** and CYP154E1 **D)** with geraniol (13). Insets show the type I substrate binding spectrum. In the diagrams the absorbance difference (A390-A420) is plotted against substrate concentration.



Supporting Experimental Procedures

Table S1: Chromatographic conditions for the GC/MS and LC/MS analysis of the reaction mixture.

Compounds	Oven program/solvent program	Other parameters
GC/MS **13–21**	start: 150 °C for 3 min ramp: 10 °C/min to 280 °C hold at 280 °C for 5 min	
GC/MS **22–26**	start: 100 °C for 1 min ramp: 15 °C/min to 300 °C hold: 300 °C for 2 min	
GC/MS **27–29** **8–12**	start: 40 °C ramp: 15 °C/min to 300 °C hold: 300 °C for 2 min [1]	
GC/MS **33**	start: 200 °C for 3 min ramp: 15 °C/min to 320 °C hold: 320 °C for 2 min	Injector and interface temperature of 300 °C

GC/MS **1–7, 51, 37,38**	start: 80 °C for 3 min ramp: 10 °C/min to 300 °C hold: 300 °C for 2 min	
GC/MS **34–36**	start: 130 °C for 3 min ramp: 10 °C/min to 260 °C hold: 260 °C for 1 min ramp: 15 °C/min to 320 °C hold: 320 °C for 2 min	
GC/MS **30-32, 38**	start: 90 °C for 5 min ramp: 15 °C/min to 320 °C hold: 320 °C for 5 min	
GC/MS **39-40**	start: 200 °C for 2 min ramp: 20 °C/min to 300 °C hold: 300 °C for 4 min ramp: 10 °C/min to 320 °C hold: 320 °C for 17 min	Injector temperature 300 °C
GC/MS **47**	start: 80 °C for 2 min ramp: 15 °C/min to 320 °C hold: 320 °C for 5 min	Injector and interface temperature of 300 °C
GC/MS **49**	start: 120 °C for 3 min ramp: 10 °C/min to 330 °C hold: 330 °C for 5 min	Injector and interface temperature of 300 °C
LC/MS: **41-46**	start: 50 %B ramp in 10 min to 100 % B hold: 100 % B for 7 min equilibration at 50 % for 5 min	A: water (0.1 % formic acid) B: methanol simultaneous: SIM(+)/SCAN(+/-) ESI/APCI dual ionization mode
LC/MS: **48**	start: 10 %B ramp in 20 min to 100 % B hold: 100 % B for 5 min equilibration at 10 % for 7 min	simultaneous: SIM(+)/SCAN(+/-) ESI/APCI dual ionization mode
LC/MS: **50**	start: 50 % B ramp in 15 min to 90 % B hold: 100 % B for 3 min equilibration at 50 % for 4 min	simultaneous: SIM(+)/SCAN(+/-) ESI/APCI dual ionization mode

Supplemental References

[1] A. Schallmey, G. den Besten, I. G. P. Teune, R. F. Kembaren, D. B. Janssen, Appl. Microbiol. Biotechnol 2011, 89, 1475–1485.

4 Manuscript 2

Titel: A novel P450-based biocatalyst for the selective production of chiral 2-alkanols

Authors: Clemens von Bühler and Vlada B. Urlacher

Published in: Chemical Communications (Cambridge, England), **2014**, 50, 4089 (ISSN:1359-7345)

DOI: 10.1039/c4cc00647j

Contribution: all experimental work, writing of the manuscript

Link: http://dx.doi.org/10.1039/c4cc00647j

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

A novel P450-based biocatalyst for the selective production of chiral 2-alkanols†

Clemens J. von Bühler and Vlada B. Urlacher*

Cite this: *Chem. Commun.*, 2014, 50, 4089

Received 24th January 2014,
Accepted 24th February 2014

DOI: 10.1039/c4cc00647j

www.rsc.org/chemcomm

A P450 monooxygenase from *Nocardia farcinica* (CYP154A8) catalyses the stereo- and regioselective hydroxylation of *n*-alkanes, still a challenging task in chemical catalysis. In a biphasic reaction system, the regioselectivity for the C2-position of C_7–C_9 alkanes was over 90%. The enzyme showed strict *S*-selectivity for all tested substrates, with enantiomeric excess (ee) of up to 91%.

Chiral linear alcohols are widely used as building blocks in the synthesis of pharmaceuticals, agrochemicals, pheromones and liquid crystals and therefore synthetic routes to these products are of high interest.[1] Besides the difficult and less selective chemical routes,[2] the enzymatic reduction of prochiral ketones catalysed by alcohol dehydrogenases is an established enzymatic process for the production of chiral alcohols.[3] However, a prerequisite for this enzymatic approach is the availability of functionalised substrates, namely ketones. Therefore, an enzymatic approach for the production of chiral secondary alcohols *via* direct hydroxylation of inert alkanes by a monooxygenase seems to be more beneficial. Alkanes are cheap starting compounds that are easily produced from mineral oil and hence are available in high amounts. Enzyme-catalysed selective hydroxylation of alkanes comes along with the typical advantages of biocatalytic processes: no requirement for toxic metal catalysts or high reaction temperatures and only limited use of potentially toxic additional organic solvents.

Several enzyme groups have been described that are capable of this type of reaction.[4] The number of oxygenases producing secondary and particularly 2-alcohols is however limited. The soluble methane monooxygenase (sMMO) and the particulate methane monooxygenase (pMMO) from *Methylococcus capsulatus* produce 2-alcohols from short chain alkanes but suffer from low expression.[5] Furthermore fungal peroxygenases accept small and medium chain alkanes as substrates but display moderate regioselectivity.[6]

Institute of Biochemistry, Heinrich-Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. E-mail: Vlada.Urlacher@uni-duesseldorf.de; Fax: +49 211 8113117; Tel: +49 211 8113866

† Electronic supplementary information (ESI) available: Materials and methods. GC/MS chromatograms, time dependence of octane conversion, and supplementation experiments. See DOI: 10.1039/c4cc00647j

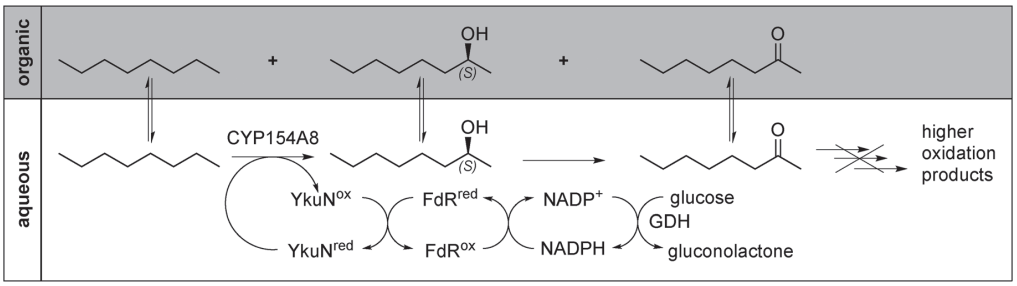
A very large and intensively studied group of enzymes catalysing hydroxylation reactions are the heme containing cytochrome P450 monooxygenases (P450s). Several wild type and engineered P450s are known to accept alkanes as substrates and especially CYP102A enzymes are known to produce 2-alcohols.[7]

Recently, we characterised two novel P450 enzymes belonging to the CYP154 family.[8] CYP154A8 from *Nocardia farcinica* IFM 10152 proved to be quite regioselective. Thus, this enzyme was chosen for a detailed investigation of its suitability as a biocatalyst for *n*-alkane oxidation.

CYP154A8 was expressed in *E. coli* in soluble form and purified as reported earlier.[8] A catalytically active P450 system was then reconstituted by the addition of the flavodoxin YkuN from *Bacillus subtilis* and *E. coli* flavodoxin reductase FdR.[8] Reactions were performed in 500 μl reaction volume under NADPH cofactor regeneration by glucose dehydrogenase. The following alkanes were tested as substrates: heptane (C_7), octane (C_8), nonane (C_9), decane (C_{10}) and dodecane (C_{12}). After extraction and derivatisation, the reaction mixtures were analysed *via* achiral and chiral GC coupled with MS. The stereoselectivity of the enzyme was chain length dependent. In comparison to authentic standards it could be determined that CYP154A8 showed strict *S*-selectivity for all tested alkanes: 84% ee for both 2-(*S*)-heptanol and 2-(*S*)-nonanol, 63% ee for 2-(*S*)-decanol and a maximum of 91% ee was found for 2-(*S*)-octanol (Fig. S1, ESI†). Dodecane was accepted by CYP154A8 but only trace amounts of products were formed and hence this alkane was not investigated further. In the case of decane oxidation, 27% 3-decanol was formed with 89% ee.

By monitoring the octane conversion over a period of 24 h it could be demonstrated that the ee value of the formed 2-(*S*)-octanol was not dependent on the degree of conversion, as is expected for an asymmetric synthesis (Fig. S2, ESI†).

So far the best-characterised P450s that catalyse this reaction with high stereoselectivity are the CYP102A1 mutants F87V/A328F,[9] 1-12G[10] and 139-3,[10,11] which all produce 2-octanol, but with significantly lower ee-values, *i.e.* 46% ee for 2-(*R*)-octanol, 39% ee for 2-(*R*)-octanol and 58% ee for 2-(*S*)-octanol, respectively. Thus CYP154A8 is the first P450 enzyme that is



Scheme 1 Hydroxylation of octane catalysed by CYP154A8 in a biphasic reaction system consisting of an organic substrate phase and an aqueous phase. The product scavenging effect of the organic phase prevents formation of higher oxidation products. YkuN (flavodoxin from *B. subtilis*), FdR (*E. coli* flavodoxin reductase), GDH (glucose dehydrogenase from *B. megaterium*).

shown to produce highly enantiomerically enriched 2-(*S*)-octanol from octane.

However, substrate conversion in aqueous solution resulted in the formation of higher oxidation products, diols, which are produced upon exposure of the formed 2-alcohol to the enzyme as described earlier.[8] In order to circumvent the formation of the undesired products, the reaction setup was modified: by employing the substrate as a second phase (20 vol% of the aqueous phase) a double benefit over the homogenous aqueous reaction system was achieved. First, the concentration of the substrate available to the enzyme was maximised, and second, the continuous extraction of formed products into the organic phase minimised or even prohibited the formation of higher oxidation products (Scheme 1). Partition coefficients for all 2-alcohols have been determined in the respective biphasic alkane–buffer-systems. Values ranged between $\log P_{C7}$ 1.0 ± 0.03 for 2-heptanol and $\log P_{C10}$ 1.42 ± 0.1 (Table S1 and Section S4.4, ESI†). This altered reaction setup yielded 2-alcohols as main reaction products with over 80% selectivity. Considering that some of the 2-alcohol molecules were obviously further oxidised by CYP154A8 to the corresponding ketones, an overall preference towards the C2-position of up to 90% can be calculated (Fig. 1A–C and Fig. S3, ESI†).

Similar to the stereoselectivity, the regioselectivity of CYP154A8 was chain length dependent and decreases for C_{10} down to 70% (Fig. 1D and Fig. S3, ESI†). For decane a significant amount of 3-decanol as well as its overoxidation product 3-decanone was detected. GC/MS data confirmed that 3-alkanones were not present for any other substrate but decane.

The benchmark for our reaction was the data published by Peters *et al.* (2003) for the CYP102A1 mutant 1-12G,[11] and data for the CYP102A1 mutant F87V/A328F designed in our group.[9] The mutant enzyme 1-12G produces 2-alcohols from alkanes ranging from C_7–C_{10} with a corresponding selectivity of 76, 82, 86, and 86%. However, in all cases alcohols with a hydroxy group at the positions C1, C3 and C4 were observed as by-products. Compared to 1-12G, CYP154A8 showed a higher or similar regioselectivity depending on the chain length of the substrate, but its product distribution was limited to the C2- and partially C3-position only. The mutant F87V/A328F is very selective for the C2-position of octane (92%).[9] Thus, the total regioselectivity of CYP154A8 for the C2-position was comparable to that of F87V/A328F.

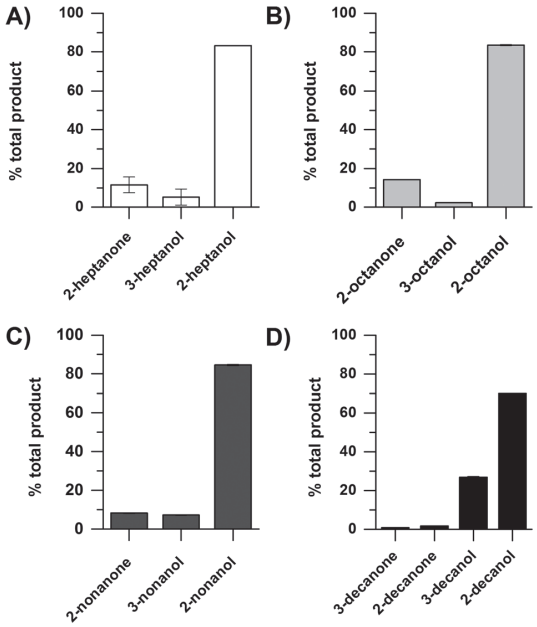


Fig. 1 Product distribution after hydroxylation of various alkanes: (A) heptane, (B) octane, (C) nonane, and (D) decane, determined from GC/MS peak areas and plotted as percentage of a specific product to the total amount of product (Fig. S3, ESI†). Where error bars cannot be seen, they are smaller than the linewidth.

In the next step challenges inherent to biphasic reaction systems such as interfacial surface area, low substrate solubility in the aqueous phase, and the choice of protein stabilising additives have been addressed. The optimised biphasic system included catalase to decompose H_2O_2, which might be formed by a shunt reaction in the catalytic cycle of P450s,[12] BSA for protein stabilisation,[13] ethanol as a cosolvent and the use of a rotator for mixing (Section S8 and Fig. S4, ESI†).

The following product titres were obtained after 24 h in the optimised biphasic reaction system: 3 mM 2-nonanol; 1.9 mM 2-heptanol; 2.2 mM 2-octanol; and 1.2 mM 2-decanol. Compared to values found for CYP102A1 mediated octane hydroxylation[9,11]

Table 1 Parameters for the 2-hydroxylation of heptane, octane, nonane and decane catalysed by CYP154A8

Sub.	Product formation rate[a,b]	Coupling[b] (%)	Conc.[c] (mM)	TTN[c,d]	ee-(*S*)[c] (%)
C_7	1.13 ± 0.08	7.3 ± 0.9	1.9	2800	84
C_8	4.6 ± 1.04	21.1 ± 4.2	2.2	3200	91
C_9	3.73 ± 0.44	15. 9 ± 1.8	3.0	4400	84
C_{10}	0.37 ± 0.05	4.0 ± 0.6	1.2	1700	63

[a] Given as nmol of 2-alcohol per nmol P450 per minute. [b] Determined using a 200 µl aqueous reaction system in multiwell plates with 1 mM substrate. [c] After 24 h in 500 µl of a biphasic reaction system. [d] Calculated as nmol of 2-alcohol per nmol P450.

the total turnover numbers (TTNs) of CYP154A8 were much higher than those reported for F87V/A328F and in the same order of magnitude as those of 1-12G (Table 1). Moreover, these results demonstrated that even though the catalytic system consisted of five enzymes, this complexity apparently does not have a negative effect on product formation.

In P450 systems at various stages of the electron transfer chain, electrons can be diverted from a productive reaction, leading to unproductive cofactor consumption, which reduces the catalytic performance of the system. This can take place between redox partners as well as through the earlier mentioned shunt reactions in the P450 catalytic cycle. To evaluate the performance of the CYP154A8 system we calculated the coupling between NADPH consumption and product formation (Table 1 and ESI†). The rather low values (Table 1) are however in accordance with those reported for other artificial P450 redox chains oxidising non-natural substrates.[14]

At a closer look these data seem to be contradictory: the total turnover number (TTN) and the final product concentration were highest for nonane but its product formation rate and coupling were lower than for octane.

Since we attributed this discrepancy to the stability of the involved enzymes we designed experiments to verify our hypothesis: octane conversion was supplemented after 5.5 h with either additional glucose, glucose dehydrogenase (both for cofactor regeneration), YkuN/FdR or YkuN/FdR/P450. After a total reaction time of 24 h each case was compared to a standard reaction setup without this extra addition. Neither the addition of glucose nor that of glucose regenerating GDH resulted in an elevated final product concentration. As opposed to this, the supplementation of the reaction with the redox partners or the P450 led to about 15% more product which could even be further maximised (45%) by the simultaneous addition of YkuN, FdR and CYP154A8 (Fig. S5, ESI†). These results indicated that as suggested the stability of the redox partner proteins and that of CYP154A8 turned out to have a major influence on the degree of conversion. In the C_7 to C_{12} alkane series the solubility decreases from 3.4 mg L^{-1} to 5.2 10^{-2} mg L^{-1}.[15] It is therefore reasonable to assume that enzyme denaturation is decreased with a longer alkane chain, due to reduced exposure. This in turn may explain the observed higher final concentrations of 2-nonanol as compared to 2-octanol (Table 1).

In conclusion, CYP154A8 from *N. farcinica* was identified as a novel regio- and stereoselective biocatalyst for the hydroxylation of inert *n*-alkanes. Depending on the substrate chain length, enzyme preference for oxidation at the C2-position was more than 90%. To the best of our knowledge, CYP154A8 is the only P450 monooxygenase where high regioselectivity for medium chain alkane hydroxylation coincides with a moderate to high enantiomeric excess of the formed 2-(*S*)-alcohols and high total turnover numbers. Furthermore, by means of reaction optimisation the total product amounts were significantly increased to levels that can compete with the best performing and well-established CYP102A1-based systems. But unlike those evolved CYP102A1 systems, CYP154A8 wild type shows already high selectivity, which might be even further increased by protein engineering methods. Compared to other enzymes like fungal peroxygenases both regio- and stereoselectivity of 2-alkanol production by CYP154A8 is significantly higher.[6] This study demonstrates the potential of CYP154A8 in alkane oxidation. Simultaneously, the described biocatalyst implies additional optimisations and the transfer to a whole cell system which are currently under investigation.

Notes and references

1 (*a*) K. Faber and M. C. Franssen, *Trends Biotechnol.*, 1993, **11**, 461; (*b*) N. Öhrner, C. Orrenius, A. Mattson, T. Norin and K. Hult, *Enzyme Microb. Technol.*, 1996, **19**, 328; (*c*) R. A. Sheldon, *J. Chem. Technol. Biotechnol.*, 1996, **67**, 1.
2 R. H. Crabtree, *J. Chem. Soc., Dalton Trans.*, 2001, 2437.
3 W. Kroutil, H. Mang, K. Edegger and K. Faber, *Curr. Opin. Chem. Biol.*, 2004, **8**, 120.
4 (*a*) Y. Ji, G. Mao, Y. Wang and M. Bartlam, *Front. Microbiol.*, 2013, **4**, 58; (*b*) J. B. Van Beilen and E. G. Funhoff, *Appl. Microbiol. Biotechnol.*, 2007, **74**, 13; (*c*) M. Bordeaux, A. Galarneau and J. Drone, *Angew. Chem., Int. Ed.*, 2012, **51**, 10712.
5 (*a*) S. J. Elliott, M. Zhu, L. Tso, H.-H. T. Nguyen, J. H.-K. Yip and S. I. Chan, *J. Am. Chem. Soc.*, 1997, **119**, 9949; (*b*) S. I. Chan, K. H.-C. Chen, S. S.-F. Yu, C.-L. Chen and S. S.-J. Kuo, *Biochemistry*, 2004, **43**, 4421.
6 S. Peter, M. Kinne, X. Wang, R. Ullrich, G. Kayser, J. T. Groves and M. Hofrichter, *FEBS J.*, 2011, **278**, 3667.
7 O. Lentz, V. Urlacher and R. D. Schmid, *J. Biotechnol.*, 2004, **108**, 41.
8 C. von Bühler, P. Le-Huu and V. B. Urlacher, *ChemBioChem*, 2013, **14**, 2189.
9 E. Weber, A. Seifert, M. Antonovici, C. Geinitz, J. Pleiss and V. B. Urlacher, *Chem. Commun.*, 2010, **46**, 944.
10 A. Glieder, E. T. Farinas and F. H. Arnold, *Nat. Biotechnol.*, 2002, **20**, 1135.
11 M. W. Peters, P. Meinhold, A. Glieder and F. H. Arnold, *J. Am. Chem. Soc.*, 2003, **125**, 13442.
12 P. R. Ortiz de Montellano, *Cytochrome P450. Structure, mechanism, and biochemistry*, Kluwer Academic/Plenum Publishers, New York, 3rd edn, 2005.
13 S. C. Maurer, K. Kühnel, L. A. Kaysser, S. Eiben, R. D. Schmid and V. B. Urlacher, *Adv. Synth. Catal.*, 2005, **347**, 1090.
14 (*a*) A. Schallmey, G. den Besten, I. G. P. Teune, R. F. Kembaren and D. B. Janssen, *Appl. Microbiol. Biotechnol.*, 2011, **89**, 1475; (*b*) M. Girhard, T. Klaus, Y. Khatri, R. Bernhardt and V. B. Urlacher, *Appl. Microbiol. Biotechnol.*, 2010, **87**, 595; (*c*) S. G. Bell, C. F. Harford-Cross and L. L. Wong, *Protein Eng.*, 2001, **14**, 797.
15 National Library of Medicine (US), Division of Specialized Information Services, *Hazardous Substances Data Bank*, available at: http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB.

Supplementary Information

Novel P450-based biocatalyst for the selective production of chiral 2-alkanols

Clemens J. von Bühler, Vlada B. Urlacher*

Institute of Biochemistry, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1*, 40225 Düsseldorf, Germany*, *E-Mail: Vlada.Urlacher@hhu.de*

Table of contents

Supplementary Materials and Methods

1. Chemicals and enzymes

The following chemicals were used: *n*-heptane (VWR), *n*-octane (Sigma-Aldrich), *n*-nonane (Alfa Aesar), *n*-decane (Alfa Aesar), 2-heptanol (TCI), 2-octanol (Sigma-Aldrich), 2-nonanol (Alfa Aesar) 2-decanol (TCI), and enantiomerically pure 2-(*S*)-heptanol (Alfa Aesar), 2-(*S*)-octanol (Merck), 2-(*S*)-nonanol (Sigma-Aldrich), 2-(*R*)-decanol (Sigma-Aldrich) and were of at least 97 % purity. Bovine serum albumin (BSA) and catalase from bovine liver was obtained from Sigma-Aldrich (Schnelldorf, Germany). NADPH and $NADP^+$ were commercially available from GERBU Biotechnik GmbH (Heidelberg, Germany).

2. Expression and quantification of the purified proteins

Enzymes were expressed and purified as described elsewhere.[1] Concentration of CYP154A8 was determined according to the method of Omura and Sato.[2] FdR and YkuN were quantified via UV/VIS spectroscopy using their published absorption coefficients.[3]

3. Biotransformations

3.1. Homogeneous aqueous reaction setup

The standard homogeneous reactions were performed in 500 ml reaction volume in 2 ml plastic reaction tubes. A master mix containing 0.675 mM CYP154A8, 6.75 mM YkuN, 0.675 mM FdR, 5 U GDH, 100 mM glucose, 10 mg/ml BSA, 600 U catalase and 200 mM $NADP^+$ was prepared and aliquoted to the reaction vessels containing *n*-alkane solution from a solution in ethanol to a final concentration of 4 mM. The reactions were performed at 25 °C for 24 h in a rotator (20 rpm). To ensure efficient cofactor regeneration the amount of glucose was set to 100 mM. Thus it is guaranteed that the cofactor regeneration is not rate limiting to the catalyst itself and additionally the low coupling of P450 with non-natural substrates does not preclude high product concentrations.

3.2. Biphasic reaction setup

The standard biphasic reactions were performed in 500 ml reaction volume in 2 ml plastic reaction tubes or 10 ml round bottoms flasks. A master mix containing 0.675 mM CYP154A8, 6.75 mM YkuN, 0.675 mM FdR, 5 U GDH, 100 mM glucose, 10 mg/ml BSA, 600 U catalase, 1 vol-% cosolvent (ethanol except otherwise stated) and 200 mM $NADP^+$ was prepared and aliquoted to the plastic tubes. The reaction was started by adding 100 µl of pure *n*-alkane (substrate) as second phase and was performed at 25 °C in either a rotator (20 rpm) or on a magnetic stirrer for 24 h.

3.3. Reaction setup for micro titer plate (MTP) measurements

For the determination of the substrate oxidation rate and coupling, reactions in 96 well plates were performed. A master mix containing 0.675 mM CYP154A8, 6.75 mM YkuN, 0.675 mM FdR, 10 mg/ml BSA, 600 U catalase, 1 vol-% ethanol and 1 mM substrate was prepared. Experiments in quadruplicates were done in 200 µl volume in a Tecan Infinite M200Pro UV/VIS MTP spectrophotometer. The measurements were started by injection of 20 µl of an NADPH solution. The concentration of the NADPH concentration was increased from 320 µM until the amount of product was high enough to be detected via GC/MS. By following the absorbance at 340 nm the concentration of NADPH in the reaction was monitored during the reaction and from the absorbance differences between the beginning and the end of the experiment the precise amount of NADPH was determined and used for subsequent calculations. After the end of the reaction 100 µl 1N HCl was injected to stop the reaction. After addition of 200 µM internal standard (ITSD) (s. section 4) each well was extracted separately with 500 µl toluene and subjected to GC/MS analysis as described in section 4.1.

4. Product analysis

Products were analysed on a GC/MS QP-2010 Plus instrument (Shimadzu, Tokyo, Japan) with helium as carrier gas. For all substrates except *n*-octane, 1-octanol was used as an internal standard. For *n*-octane the internal standard was 1-decanol. The reactions were stopped by the additions of 50 µl 6N HCl and subsequently the ITSD was added to a final concentration of 2 mM prior to extraction (2 × 500 µl toluene).

4.1. Achiral GC/MS

Toluene extracts were measured using an FS-Supreme-5 30 m × 0.25 mm × 0.25 µm (Chromatographie Service GmbH, Langerwehe, Germany) fused silica. The oven temperature was 80 °C for 6 min, ramped to 140 °C at 10 K s^{-1} and subsequently heated with 50 K s^{-1} to 300 °C and held isothermal for 5 min.

4.2. Chiral GC/MS

Toluene extracts were dried over anhydrous $MgSO_4$ and for derivatisation incubated in GC/MS vials for 2 h at 90 °C with acetic anhydride. After the derivatisation took place the samples were washed twice with 500 µl water to remove excess acetic acid. Chiral analysis was done using an FS-CYCLODEX beta-I/P (Chromatographie Service GmbH, Langerwehe, Germany) column. The GC was held at 100 °C for 17 min, ramped to 200 °C at 10 K s^{-1} and held isothermal for 2 min. By comparison of an authentic sample of the racemic 2-alcohols

with an authentic sample of a single enantiomer the identity of the formed products was verified.

4.3. Determination of coupling and product formation rate

Based on the via GC/MS determined amount of 2-alcohol the substrate oxidation rate was calculated as:

$$r^{ox.} = \frac{n^{Prod}}{n^{P450} * t}$$

Further, coupling was calculated as:

$$coupling = \left(\frac{n^{Prod.}}{n^{NADPH}}\right) * 100$$

using: oxidation rate, $r^{ox.}$; amount of product, n^{Prod}; amount of NADPH, n^{NADPH}; time until depletion of NADPH, t; and amount of P450, n^{P450}.

4.4. Determination of partition coefficients

In order to determine the distribution of the formed 2-alcohols between the aqueous and the organic phase in the reaction system, partition coefficients have been determined. 5 ml potassium phosphate buffer (100 mM, pH 7.5) with ethanol as cosolvent and 10 mg/ml BSA were overlaid with 1 ml of *n*-alkane containing the corresponding 2-alcohol. To obtain comparable data, concentrations of all 2-alcohols were set to 2 mM (referring to the aqueous phase). Phases were equilibrated for 45 min on a rotator at 100 rpm. Subsequently, further equilibration and phase separation took place for 30 min without agitation. Finally, the samples have been centrifuged for 15 min at 3500 g and 20 °C. 500 µl aliquots of the organic and the aqueous phase have been taken and ITSD was added. Samples from the organic phase were analysed directly by GC/MS. Aqueous samples were extracted with toluene and analysed by GC/MS. Values from the aqueous phase have been normalized using the ITSD. Peak areas of the 2-alcohol in both phases have been used to calculate log*P* values.

$$\log P = \frac{A^{aq.}_{norm.}}{A^{org.}}$$

Using $A^{aq.}_{norm.}$: normalized peak area of the 2-alcohol in the aqueous phase and $A^{org.}$: peak area of the 2-alcohol in the equilibrated organic phase. Measurements were performed in triplicates.

Supplementary Results

5. Partition Coefficients

Table S1: Partition coefficients of the 2-alkanols determined for the respective biphasic buffer/alkane-reaction systems

Solute/ org. solvents	2-heptanol / *n*-heptane	2-octanol/ *n*-octane	2-nonanol / *n*-nonane	2-decanol / *n*-decane
$\log P^{alkane/buffer}$	1.0 ± 0.03	1.36 ± 0.02	1.36 ± 0.03	1.42 ± 0.1

6. Gas chromatograms of 2-alkanols from chiral GC/MS analysis

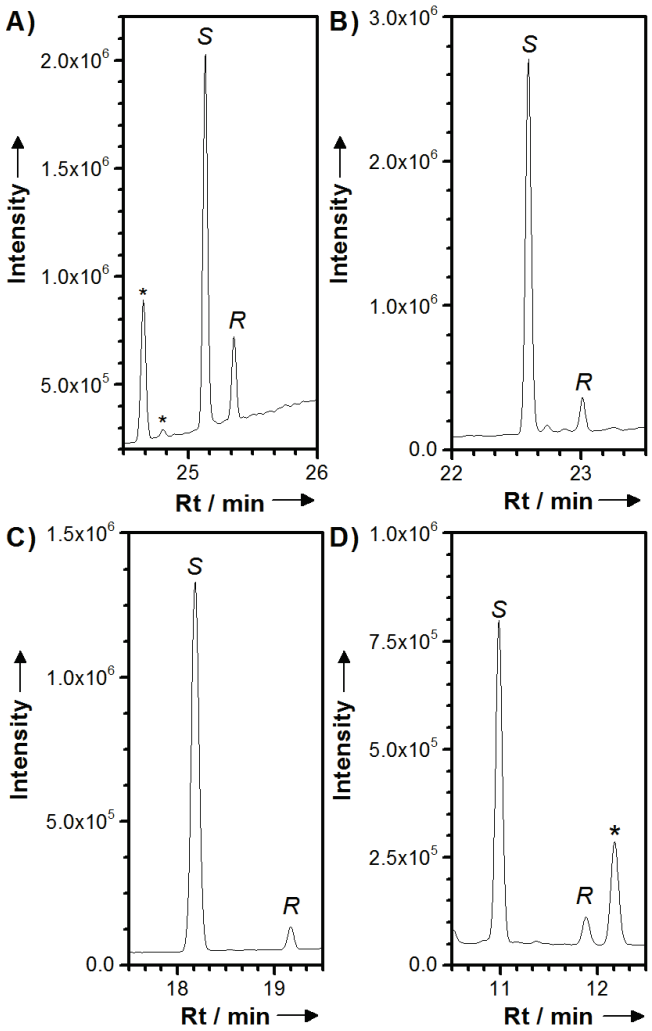


Figure S1: Chromatograms of the chiral GC/MS analysis of the 2-alkanols produced by CYP154A8 from: **A)** 2-decanol **B)** 2-nonanol **C)** 2-octanol **D)** 2-heptanol. Peaks are labelled with the stereodescriptors assigned with the help of authentic standards. An asterisk (*) denotes background peaks.

7. Time dependence of n-octane conversion

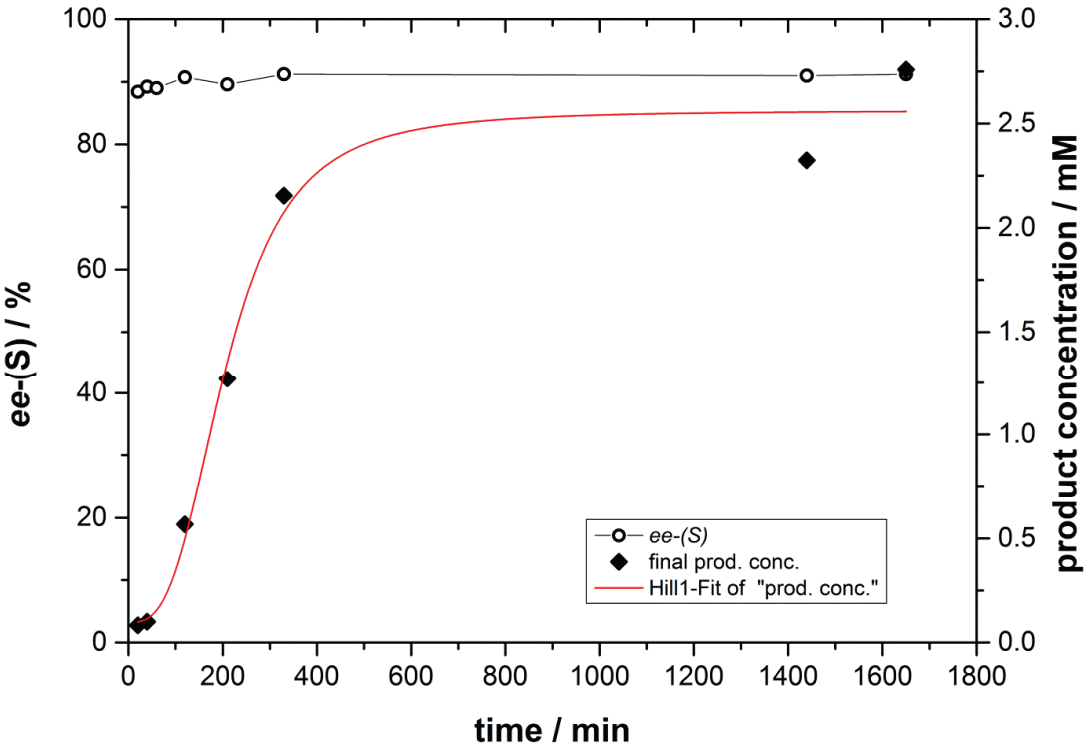


Figure S2: Monitoring of *n*-octane conversion and the *ee* value of 2-(*S*)-octanol over time. A common master mix (MM, composition as given in section 3.2) was prepared for all time points. 500 µl of the MM were aliquoted into separate 2 ml plastic reaction tubes. After the given time the reaction mixture was extracted to determine the amount of product and the enantiomeric excess via GC/MS.

8. Gas chromatograms of *n*-alkane conversion from achiral GC/MS analysis

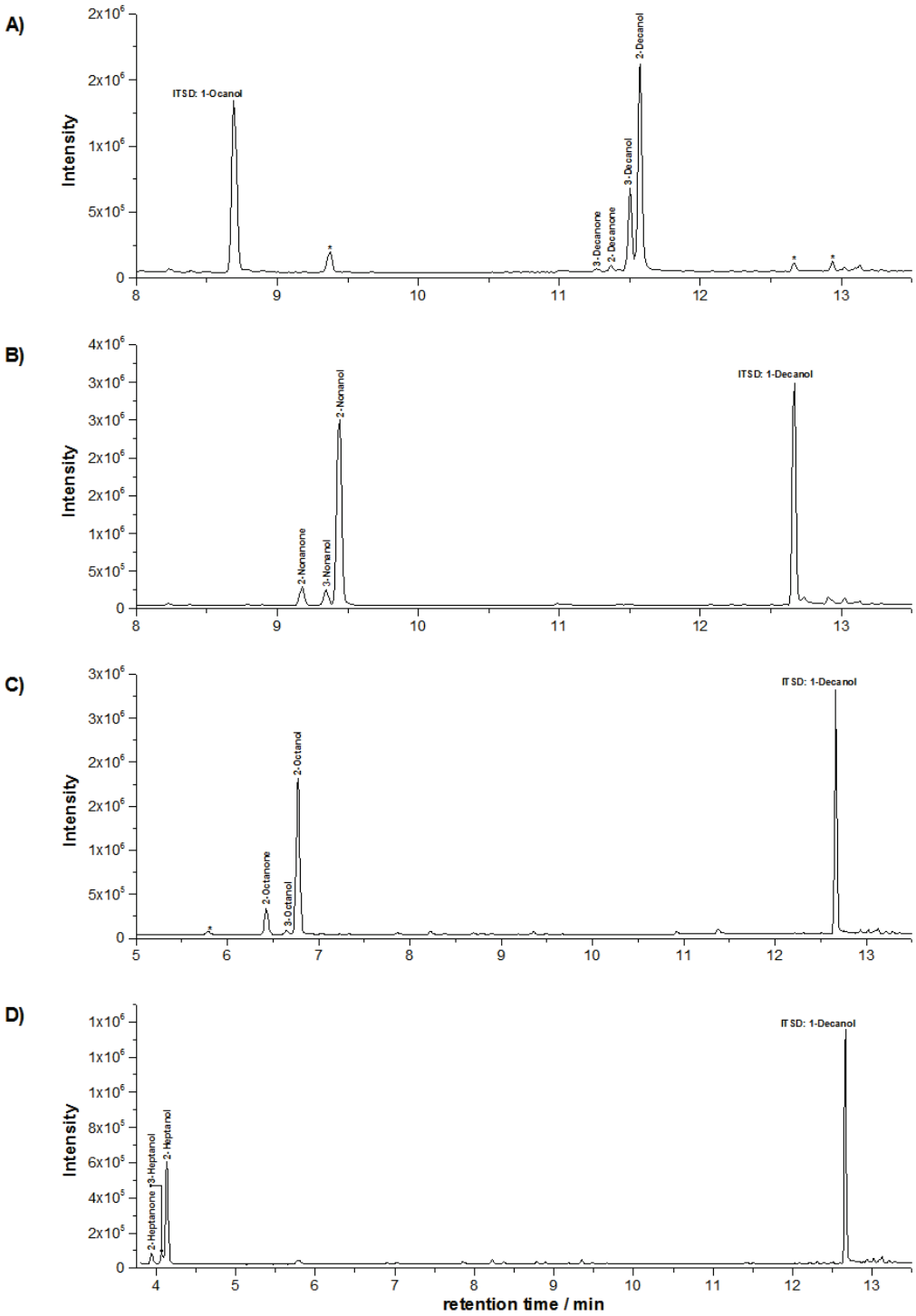


Figure S3 Chromatograms of the achiral GC/MS analysis of *n*-alkane conversions catalysed by CYP154A8: decane **A)**, nonane **B)**, octane **C)** and heptane **D)**. Peaks are labelled with the names of the corresponding products as well as internal standard (ITSD), and background peaks (*).

9. Optimization of the biphasic reaction system

In biphasic system several issues different to homogeneous reaction systems have to be addressed: Interfacial surface area is crucial for the mass transfer between aqueous and organic phase. This can be influenced by different mixing method.

With the standard biphasic reaction setup we tested: no mixing, shaking with a rotator and mixing with a magnetic stir bar in round bottom flasks. The results showed that the use of a rotator increased product formation by a factor of ~16 compared to the system without shaking (Figure S4 A). Mixing in a round bottom flask resulted however in lower amounts of products than in the system with a rotator. This indicated destabilisation and/or inactivation of the involved enzymes caused by shearing forces.

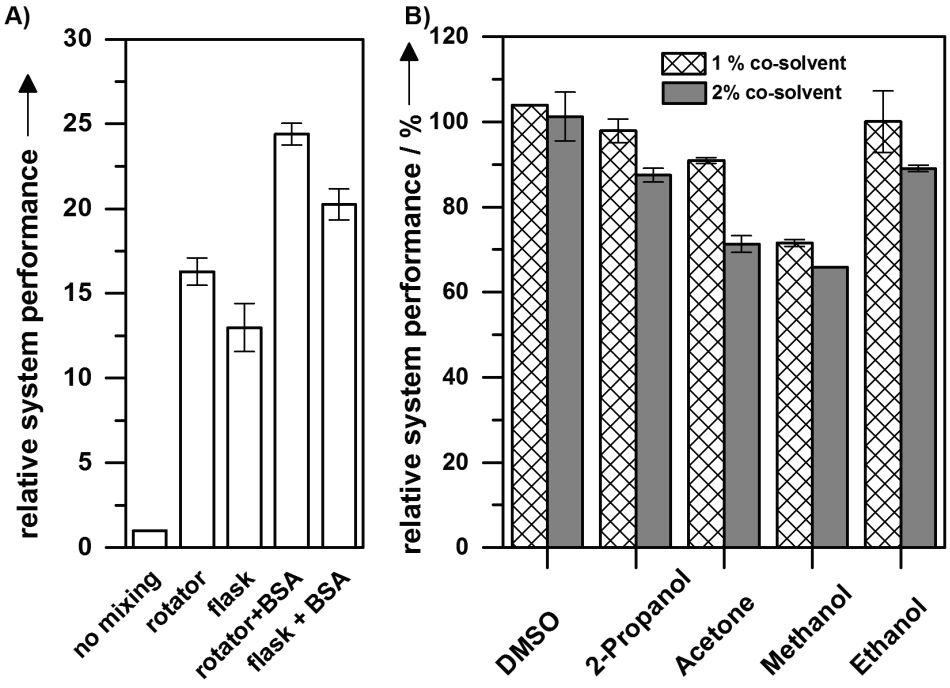


Figure S4. A) Octane oxidation catalysed by CYP154A8 in a biphasic system with different mixing techniques applied in order to increase the interfacial surface area as well as BSA to stabilise the enzymes. Values are normalized to "no mixing". **B**) Conversion of octane by CYP154A8 with different co-solvents tested in 1 vol-%and 2 vol-% concentration. Values are normalized to the value achieved with the default 1 vol-% ethanol. Where error bars cannot be seen, they are smaller than the line width.

Besides catalase, bovine serum albumin (BSA) was tested as well. In accordance with the data from Maurer et al. (2005) the protein stabilizing effect of BSA to the reaction mix led to an increase in product formation of about 50 %.[4] BSA may in addition act as sink for the substrate, because its natural function is to bind hydrophobic compounds.

Low substrate solubility is often a limiting factor for biocatalysis in biphasic systems, e.g. the solubility of octane in water is only 0.66 mg/l = 5.8 mM at 25 °C.[5] To increase the accessibility of alkanes for the enzyme in addition to ethanol which was used by default, a set of water miscible co-solvents in two different concentrations (1 vol-% and 2 vol-%) was tested: DMSO, 2-propanol, acetone and methanol. The results indicated that product formation was maximal with DMSO or ethanol (Figure S4 B), although the difference between 1 and 2 % co-solvent was minimal. Because DMSO was found to be oxidized by P450s to dimethyl sulfone we decided to use ethanol as co-solvent.

10. Results of the supplementation experiments

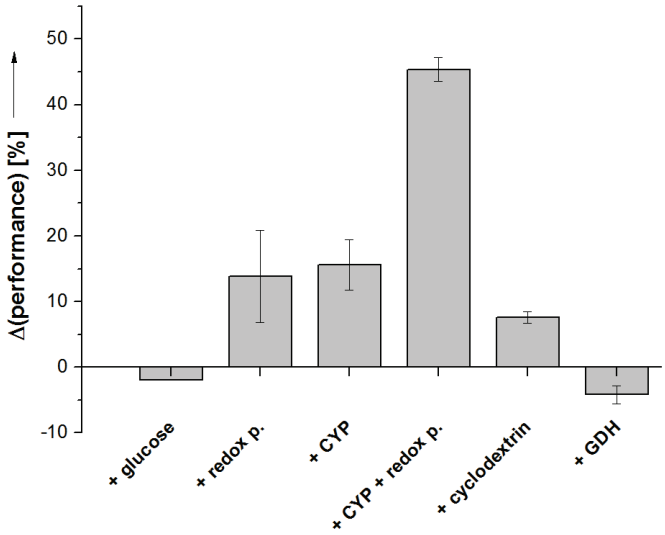


Figure S5: Results of the supplementation experiments: In order to prove the influence of enzyme stability on *n*-alkane conversion the following components were added to a standard biphasic reaction after 5,5 h of reaction: 1. glucose, 2. YkuN and FdR, 3. CYP154A8, 4. YkuN, FdR and CYP154A8, 5. methyl-β-cyclodextrin (to increase substrate/product solubility), 6. GDH. To evaluate the effect of the supplementation the samples were extracted after 24 h and product concentration analysed via GC/MS. Values are given as % relative to the system without supplementation.

SupplementalNotes and references

1 C. von Bühler, P. Le-Huu and V. B. Urlacher, *Chembiochem*, 2013, **14**, 2189–2198.

2 T. Omura and R. Sato, *J. Biol. Chem.*, 1964, **239**, 2370–2378.

3 Z.-Q. Wang, R. J. Lawson, M. R. Buddha, C.-C. Wei, B. R. Crane, A. W. Munro and D. J. Stuehr, *J. Biol. Chem.*, 2007, **282**, 2196–2202.

4 S. C. Maurer, K. Kühnel, L. A. Kaysser, S. Eiben, R. D. Schmid and V. B. Urlacher, *Adv. Synth. Catal.*, 2005, **347**, 1090–1098.

5 National Library of Medicine (US), Division of Specialized Information Services, *Hazardous Substances Data Bank,* available at: http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB.

5 Manuscript 3

Titel:	Whole-cell biotransformation with recombinant cytochrome P450 for the selective oxidation of Grundmann's ketone
Authors:	Alba Hernández-Martín, Clemens J. von Bühler, Florian Tieves, Susana Fernández, Miguel Ferrero and Vlada B. Urlacher*
Published in:	Bioorganic & Medicinal Chemistry, **2014**, *in press* (ISSN:0968-0896) http://www.sciencedirect.com/science/journal/09680896
DOI:	10.1016/j.bmc.2014.06.005
Contribution:	*In vitro* redox partner screening, whole cell reactions with the YkuN/FdR system, participation in the GC/MS method development, data analysis, writing of the manuscript
Notes:	Reprinted from Bioorganic & Medicinal Chemistry, with permission from Elsevier. License Number: 3415820913836
Link:	http://dx.doi.org/10.1016/j.bmc.2014.06.005

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Whole-cell biotransformation with recombinant cytochrome P450 for the selective oxidation of Grundmann's ketone

Alba Hernández-Martín [a,b,†], Clemens J. von Bühler [a,†], Florian Tieves [a], Susana Fernández [b], Miguel Ferrero [b], Vlada B. Urlacher [a,*]

[a] Institute of Biochemistry, Heinrich-Heine University Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany
[b] Departamento de Química Orgánica e Inorgánica, Instituto Universitario de Biotecnología de Asturias, Universidad de Oviedo, 33006 Oviedo (Asturias), Spain

ARTICLE INFO

Article history:
Received 17 April 2014
Revised 27 May 2014
Accepted 2 June 2014
Available online xxxx

Keywords:
Cytochrome P450
Vitamin D_3
Grundmann's ketone
Whole-cell biocatalysis
Hydroxylation

ABSTRACT

25-Hydroxy-Grundmann's ketone is a key building block in the chemical synthesis of vitamin D_3 and its derivatives through convergent routes. Generally, the chemical synthesis of this compound involves tedious procedures and results in a mixture of several products. Recently, the selective hydroxylation of Grundmann's ketone at position C25 by cytochrome P450 (CYP) 154E1 from *Thermobifida fusca* YX was described. In this study a recombinant whole-cell biocatalyst was developed and applied for hydroxylation of Grundmann's ketone. Biotransformation was performed by *Escherichia coli* cells expressing CYP154E1 along with two redox partner systems, Pdx/PdR and YkuN/FdR. The system comprising CYP154E1/Pdx/PdR showed the highest production of 25-hydroxy-Grundmann's ketone and resulted in 1.1 mM (300 mg L^{-1}) product concentration.

1. Introduction

The importance of vitamin D_3 **1** ((3β,5*Z*,7*E*)-9,10-secocholesta-5,7,10(19)-trien-3-ol, Fig. 1) in modulating calcium and phosphorus homeostasis has been well recognized for decades, but it is also of great interest because of its potential in the therapy of infectious diseases, cardiovascular diseases, autoimmune diseases and some forms of cancer.[1,2] Vitamin D_3 is synthesized in the skin from sun exposure upon ultraviolet irradiation of 7-dehydrocholesterol but can also be absorbed from the diet. Vitamin D_3 has no inherent hormonal activity itself, and has to be metabolized in vivo to be active.[3] In human body vitamin D_3 is first hydroxylated at position C25 mainly in the liver by mitochondrial CYP27A1[4] and microsomal CYP2R1[5] to yield 25-hydroxyvitamin D_3 (**2**, 25-OH-D_3), the major circulating form of vitamin D_3 in serum. The second hydroxylation catalysed by mitochondrial CYP27B1 mainly in the kidney[6] results in the biologically active form of vitamin D_3, 1α,25-dihydroxyvitamin D_3 (**3**, 1α,25-$(OH)_2$-D_3). This intermediate was found to possess anti-proliferative and anti-invasive activities in, for example, prostate cancer cells.[3,7]

The enzymes involved in the metabolism of vitamin D_3 are cytochrome P450 monooxygenases. The cytochromes P450 (P450s or CYPs) belong to a large superfamily of hemoproteins which catalyse a wide range of oxidation reactions.[8–10] They are widely distributed in nature and are responsible for regio- and stereoselective oxidation of a vast variety of organic compounds.[11] Most P450s are external monooxygenases: they require reducing equivalents for oxygen activation and substrate oxidation. Reducing equivalents are ultimately derived from the cell cofactors NADH or NADPH and are transferred by specific electron transfer proteins to the heme iron. In most cases one or two electron transfer proteins are involved in this process.[12]

Oxidation processes catalysed by P450s have attracted interest in the biotechnology sector, because they provide a selective route for the synthesis of high-value compounds.[13] The number of new P450 genes has been increasing constantly, and represents a high potential for the identification of novel biocatalysts.[11] Several examples of bioconversion processes of vitamin D_3 to its active form by bacterial or mammalian P450s have been reported.[14] For example, CYP105A1 from *Streptomyces griseolus* was found to produce 1α,25-$(OH)_2$-D_3 via two hydroxylations of vitamin D_3—first at position C25 and then at C1.[15] Site-directed mutagenesis at positions Arg73 and or Arg84 of this monooxygenase highly enhanced both 1-(*R*)- and 25-hydroxylation activities of CYP105A1.[16] Besides that, CYP107 from *Pseudonocardia autotrophica* was found to

* Corresponding author. Tel.: +49 211 81 13687.
E-mail address: vlada.urlacher@uni-duesseldorf.de (V.B. Urlacher).
† Both authors contributed equally.

http://dx.doi.org/10.1016/j.bmc.2014.06.005
0968-0896/

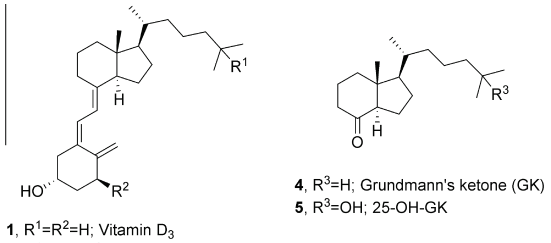


Figure 1. Vitamin D_3 related structures, Grundmann's ketone and 25-OH-Grundmann's ketone.

catalyse the same hydroxylations on vitamin D_3 leading to 1α,25-$(OH)_2$-D_3.[17]

There are several synthetic routes leading to 1α,25-$(OH)_2$-D_3 and its analogues.[18] The classical approach entails the photochemical ring opening of the 1-hydroxylated 7-dehydrocholesterol derivative to a previtamin, which is then thermolyzed to vitamin D. Total chemical synthesis of this compound starting from cholesterol can be completed in about 20 steps, and with low yields (approximately 1%).[19]

The therapeutic potential of 1α,25-$(OH)_2$-D_3 is limited by its tendency to induce hypercalcemia and hyperphosphatemia.[20] Therefore, during the last years, focus was laid on the synthesis of analogues of 1α,25-dihydroxy vitamin D_3 with dissociation of antiproliferative and differentiation activities from calcemic and phosphatemic effects.[21] Several methods involve the direct modification of vitamin D_3. Currently, the construction of the triene system by convergent coupling of an A-ring synthon with a CD ring/side chain fragment is mainly used. In this context, 25-hydroxy-Grundmann's ketone (**5**, 25-OH-GK) is a versatile precursor for the preparation of the CD ring/side chain part of 1α,25-$(OH)_2$-D_3 derivatives.[22] Several strategies for the synthesis of 25-OH-GK have been described in the literature. The most practical approach is the partial synthesis through degradation of vitamin D_3 to afford Grundmann's ketone **4** ((1*R*,3α*R*,7α*R*)-1-[(1*R*)-1,5-dimethylhexyl]octahydro-7α-methyl-4*H*-inden-4-one, Fig. 1)[23,24] followed by direct 25-hydroxylation; or by side chain modification of the Inhoffen-Lythgoe diol, obtained from degradation of vitamin D_2.[25–27] Although several methods have been described for the oxidation of the tertiary carbon at position 25, the one employing ruthenium tetroxide, developed in the Roche Research Center[28] and re-examined by DeLuca et al.,[29] is the most suitable for the synthesis of larger quantities. The reaction is carried out with catalytic amounts of $RuCl_3$ and excess $NaIO_4$ to generate 44–49% of 25-OH-GK with 10–12% of unreacted starting material. Oxidation reaction based on dioxiranes is reported in higher yields, but the low stability of these reagents and the tedious synthetic procedure limit their use.[30,31] Thus, a novel more sustainable method for the hydroxylation of Grundmann's ketone **4** is required.

Due to their chemical versatility and a wide substrate range, P450s appear as perfect candidates for achieving this goal. Recently, CYP154E1 from *Thermobifida fusca* YX and CYP154A8 from *Nocardia farcinica* were identified and characterized in our group, which selectively hydroxylate Grundmann's ketone (**4**, Fig. 1) at position C25.[32] To the best of our knowledge, these are the only reported P450s that enable this reaction. The established reaction was performed in 500 µL reaction volume using 100 µM substrate concentration. This biocatalytic system consisted of five isolated enzymes, namely, the catalytic P450, two electron transferring redox partner proteins, a glucose dehydrogenase to regenerate NADPH and catalase to detoxify reactive oxygen species which might be formed during the reaction. The complexity of this biocatalytic system as well as the need for the expensive cofactor NADPH can be overcome by the use of a whole-cell recombinant system instead of isolated enzymes. CYP154E1 demonstrated higher activity towards Grundmann's ketone **4**.[32]

The aim of this study was to investigate a whole-cell biocatalyst based on *Escherichia coli* expressing CYP154E1 together with corresponding redox partners in the reaction with this compound. As no physiological redox partners of CYP154E1 have been identified yet, two heterologous redox partner systems were tested. The first redox partner system included the putidaredoxin (Pdx) and the NADH-dependent putidaredoxin reductase (PdR) from *Pseudomonas putida*. The second redox partner system involved the flavodoxin YkuN from *Bacillus subtilis* and the NADPH dependent flavodoxin reductase (FdR) from *E. coli*, which were previously used in the in vitro approach. Each of the two redox partner systems were separately coexpressed with CYP154E1 and the resulting resting cells were tested for the hydroxylation of Grundmann's ketone **4**. The cofactor NAD(P)H needed for the reaction should be supplied via cell metabolism.

2. Results

2.1. First experiments with resting *E. coli* cells

E. coli BL21(DE3) was used as host for the coexpression of CYP154E1, Pdx and PdR. Biotransformations were carried out using resting cells in CV2 buffer containing glycerol. The P450 content verified from CO-difference spectra was in average 825 nmol L^{-1} expression medium. This corresponds to 4.7 mg g^{-1} cell wet weight (cww). For the first experiments (Table 1, entry 1) 70 g_{cww} - L^{-1} were used with a P450 content of 114 nmol g_{cww}^{-1}. Different initial concentrations of substrate **4** were tested and, due to the low solubility of GK **4** in water, also different amounts of cosolvent were used. Dimethyl sulfoxide (DMSO) was the cosolvent of choice. Ethanol and acetone were also tested but led to lower productivities (data not shown). First biotransformations were performed in 15 mL glass tubes containing 2 mL reaction volume each. Samples (1 mL) were taken after 5 and 24 h and analysed by GC/MS after extraction with ethyl acetate (Table 2).

As shown in Table 2, higher amounts of DMSO (entries 2, 4, 6 and 7) resulted in a decreased substrate conversion. Thus, the use of higher concentrations of the substrate **4** (entry 7) seemed to be counterproductive under these conditions due to solubility problems. Independently of the cosolvent concentration, the reaction rate was decreasing with increasing initial concentration of substrate **4** from 50 µM to 200 µM (entries 1–6). Nevertheless, when 2% DMSO was used, full conversion of 200 µM GK was achieved within 24 h (entries 1, 3 and 5). Thus, it could be concluded that the chosen whole-cell catalyst is suitable for biotransformation of Grundmann's ketone **4**. Substrate **4** concentrations of up to 200 µM seemed not to have a negative effect on conversion due to its possible toxicity for *E. coli* cells. Conditions in entry 5 were optimal for the chosen concentration of cells (70 g L^{-1}). It is also noteworthy that no side product formation was observed (Fig. 2). Negative controls in the absence of substrate **4** or with *E. coli* cells harbouring empty vector or with the cells expressing only Pdx and PdR but no CYP154E1, did not show any substrate conversion.

2.2. Product inhibition

The results presented in Table 2 demonstrate a decreased conversion when higher substrate concentrations were applied. In order to exclude product inhibition, biotransformations (Table 1,

Table 1
Overview of all biotransformation experiments performed in this study

Nr.	Substrate (μM)	Reaction vessel	Reaction volume (mL)	Amount of biocatalyst	Max. product (μM)	Reaction time (h)	Investigated parameter	Details
1	50–500	15 mL glass tube	2	70 g_{cww} L^{-1}	227	24	Cosolvent and product concentration	s. Table 2
2	200	15 mL glass tube	2	70 g_{cww} L^{-1}	198+$[P]_0$	24	Product inhibition ($[\mathbf{5}]_0$ = 50 μM, 100 μM)	s. Table 3
3	200	15 mL glass tube	2	70, 140, 210 g_{cww} L^{-1}	200	24	Amount of cells	s. Figure 3
4	200	100 mL flask	10	70, 140, 210 g_{cww} L^{-1}	200	24	Oxygen supply	s. Figure 3
5	1000–3000	100 mL flask	5, 10	140 g_{cww} L^{-1}	1050	24	Substrate concentration, reaction volume, redox partner Pdx/PdR	s. Figure 4
6	3000	100 mL flask	5, 10	140 g_{cww} L^{-1}	433/622	24/48	Reaction volume, redox partner YkuN/FdR	s. Figure 4
7	200	2 mL plastic tube	0.5	1 μM	200	5	Redox partners	in vitro, YkuN/FdR
8	200	2 mL plastic tube	0.5	1 μM	18	5	Redox partners	in vitro, Pdx/PdR

Table 2
Whole-cell biotransformation of Grundmann's ketone **4** into 25-OH-GK 5 by CYP154E1

Entry	**4** (μM)	DMSO[a] (%)	Conversion[b] (%)	
			5 h	24 h
1	50	2	37	100
2	50	5	23	77
3	100	2	26	100
4	100	5	15	61
5	200	2	17	100
6	200	5	10	45
7	500	5	9	45

[a] Percentage of cosolvent added to the CV2 biotransformation buffer.
[b] Qualitative data calculated from peak area integration of the GC/MS chromatogram.

Table 3
Product inhibition in the whole-cell biotransformation of Grundmann's ketone (**4**) to 25-OH-GK (**5**) by CYP154E1[a]

Entry	*t* (h)	Conversion[b] (%)		
		$[\mathbf{5}]_0$ = 0 μM[c]	$[\mathbf{5}]_0$ = 50 μM[c]	$[\mathbf{5}]_0$ = 100 μM[c]
1	1	13	10	11
2	2.5	35	42	44
3	5	67	75	75
4	8	82	91	87
5	24	80	99	99

[a] Reaction conditions: 70 g_{cww} L^{-1}, 200 μM Grundmann's ketone **4**, 2% DMSO, CV2 buffer, 180 rpm, 25 °C.
[b] Calculated by GC/MS after calibration using anthracene as ITSD.
[c] Initial concentration of 25-OH-GK (**5**) in the reaction mixture are subtracted for conversion calculation.

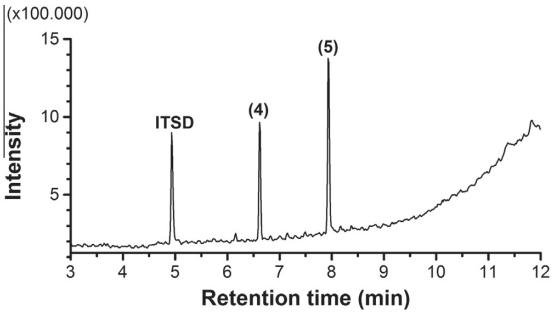


Figure 2. GC/MS chromatogram of a biotransformation of 200 μM GK in a 10 mL volume after 5 h. Besides the internal standard (ITSD) only peaks for substrate **4** and product **5** were observed.

entry 2) were run under the previously established conditions (Table 2, entry 5), but with addition of the product 25-OH-GK **5** at concentrations of 50 and 100 μM to the reaction mixture. GC/MS analysis of the samples showed that the degree of conversion was independent of the presence of product **5** (Table 3). The fact that under these conditions product **5** did not inhibit CYP154E1, supports our hypothesis that the decrease of substrate conversion with increasing initial substrate concentration is more related to the low solubility of the substrate 4 and its insufficient accessibility for the enzyme.

2.3. Effect of cell amounts and oxygen supply

With the aim to further improve productivity in the whole-cell system, the optimal amount of cells for the biotransformation should be determined. The experiments (Table 1, entry 3) were carried out using cell suspensions at concentrations of 70, 140 and 210 g_{cww} L^{-1} with a P450 content of 105 nmol g_{cww}^{-1}. Intuitively, a higher concentration of the biocatalyst in the reaction mixture should result in a faster reaction rate and a higher conversion. Surprisingly, the exact opposite was observed in our experiments: at higher cell concentrations lower substrate conversion at specific time points was observed. Thus, even if for 70 and 140 g_{cww} L^{-1} total conversion was achieved within the first 24 h, reaction rates with 140 g_{cww} L^{-1} were lower, and the reaction with 210 g_{cww} L^{-1} did not even reach full conversion (Fig. 3 A). The fact that conversion rates decreased with increasing cell concentrations indicates a limitation for the cells in performing biotransformation. Under the given conditions this limiting factor could be the availability of molecular oxygen. Higher amounts of cells increased oxygen consumption, either through their metabolism and/or because of a higher concentration of CYP154E1 consuming more oxygen as cosubstrate.

Higher oxygen supply in the reaction mixture can be achieved by increasing the air/liquid interfacial surface. With this aim in mind and in order to scale up the reaction, biotransformations (Table 1, entry 4) were performed in 100 mL shaking flasks (reaction volume 10 mL) instead of 15 mL glass tubes (Table 1, entry 3). In the larger reaction volume the influence of different cell concentrations (70, 140, 210 g_{cww} L^{-1}) was investigated as in the previous experiment in glass tubes. After 5 h reaction with 70 g_{cww} L^{-1} only about 52% conversion was achieved in glass tubes (Fig. 3 A), whereas in flasks conversion was 87% (Fig. 3 B). Full conversion was reached after 24 h in glass tubes, compared to 8 h in flasks. Higher reaction rates in flasks was attributed to a better oxygen supply which ensured full conversion with the highest cell amount of 210 g_{cww} L^{-1} which was not possible in glass tubes.

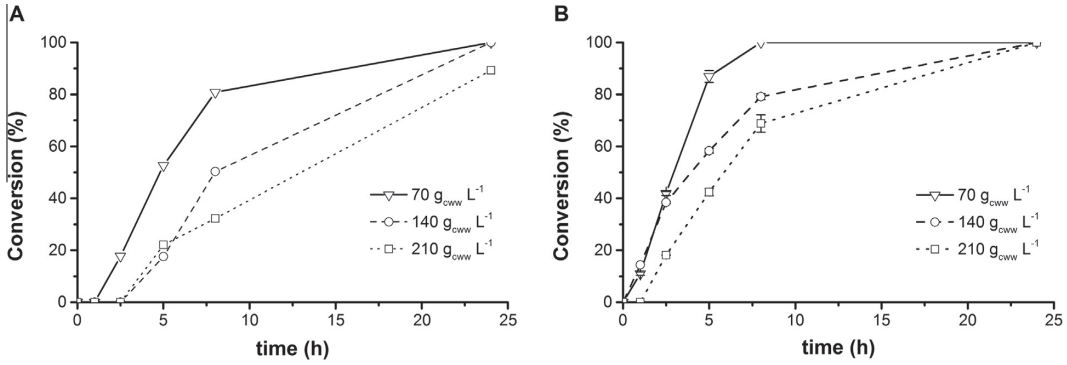


Figure 3. Effects of oxygen supply and cell concentration in the biotransformation of 200 μM GK: (A) glass tubes (15 mL), 2 mL reaction mixture (200 μM GK, 2% DMSO, CV2 buffer, 180 rpm, 25 °C), one tube per time point; (B) shaking flasks (100 mL), 10 mL reaction mixture (200 μM GK, 2% DMSO, CV2 buffer, 180 rpm, 25 °C), all time points from the same reaction mixture.

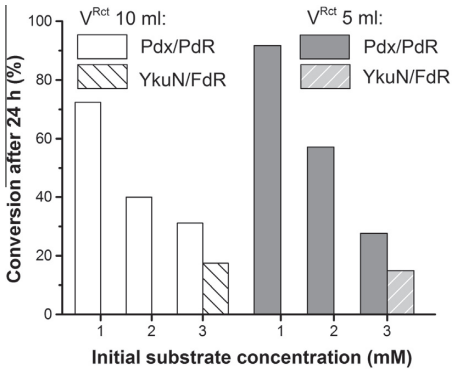


Figure 4. Scale-up and oxygen supply studies. Reaction conditions: 140 $g_{cww}{}^{-1}$, 2% DMSO, CV2 buffer, 180 rpm, 25 °C. White background: 10 mL reaction volume (V^{Rct}); grey background: 5 mL reaction volume (both 100 mL shaking flasks).

Further sets of experiments (Table 1, entry 5) were performed in 100 mL flasks with 140 g_{cww} L^{-1} at initial substrate concentrations of 1, 2, and 3 mM either in 10 mL (white bars, Fig. 4) or in 5 mL reaction volume (grey bars, Fig. 4). The decreased reaction volume from 10 ml to 5 ml increased the conversion of 1 and 2 mM GK **4** (Fig. 4, white vs grey bars). The maximal product concentration of 1.1 mM being produced from 2 mM GK **4** was achieved in 5 mL reaction volume after 24 h. Interestingly, when using 3 mM substrate concentration no increased conversion was observed in the smaller reaction volume, indicating that not only oxygen supply but other reaction parameters might be limiting.

2.4. Influence of redox partners

In order to evaluate the role of electron transfer partners, YkuN flavodoxin from *Bacillus subtilis* and the NADPH-dependent flavodoxin reductase (FdR) from *E. coli* were also tested as redox partners in this whole-cell process. In in vitro experiments YkuN/FdR combined with CYP154E1 led to a 10-fold increased conversion of GK **4** as compared to Pdx/PdR (Table 1, entry 7 vs 8). These results encouraged us to construct the tricistronic expression vector pCYP154E1-Ykun-FdR. Using cells harbouring this construct we repeated the biotransformation experiments in 10 mL and 5 mL reaction volume and 3 mM initial substrate concentration (Table 1, entry 6). Unfortunately, under the same expression conditions as for Pdx/PdR (Table 1, entry 5) the construct containing YkuN/FdR yielded only about half the P450 concentration (57 nmol g_{cww}^{-1} vs 114 nmol g_{cww}^{-1}) although the copy number and the promoter of both constructs were identical. Final product concentrations of 433 μM after 24 h and 620 μM after 48 h were observed with 140 g_{cww} L^{-1} in 10 ml reaction volume. As previously with Pdx/PdR (Table 1, entry 5) at 3 mM substrate concentration the use of reduced reaction volume did not increase substrate conversion. Although higher P450 concentrations were observed in *E. coli* with Pdx/PdR compared to the *E. coli* with YkuN/FdR, lower productivities of the latter system could not be explained by lower P450 expression only, since in vitro activities of CYP154E1/YkuN/FdR were at least 10 times higher. The total turnover values calculated per P450 in the whole-cell biocatalysts were up to 69% higher for the YkuN/FdR system compared to the Pdx/PdR system and thus reflect the same tendency as the results from the in vitro experiments. This was an indication that NADPH supply might be another limiting factor in this system, since *E. coli* flavodoxins reductase (FdR) is NADPH-dependent. Supplementation of the reaction buffer with 1% glucose (in order to accelerate NADPH regeneration) did not lead to the improved final product concentration, although the initial substrate conversion rate within the first 5 h was ~20% higher (data not shown).

3. Discussion

As mentioned in the introduction, vitamin D_3 and especially its biologically active metabolite 1α,25-$(OH)_2$-D_3, are involved in a multitude of biological processes and are important for the treatment of several diseases and disorders.[3,7] The therapeutic application of 1α,25-$(OH)_2$-D_3 is however hampered by its tendency to induce hypercalcemia and hyperphosphatemia.[20] 25-Hydroxy-Grundmann's ketone **5** is a key building block for the preparation of vitamin D_3 analogues with beneficial therapeutic properties through convergent strategies. As the established methods for the oxidation of Grundmann's ketone **4** at the physiologically important position C25 are accompanied by several problems, we decided to investigate an alternative biocatalytic approach. Interestingly, in our experiments CYP105A1 which accepts vitamin D_3 as substrate did not show any activity towards Grundmann's ketone **4**. Furthermore, we tested a library of 81 enzymes from the CYP102A family and their variants. These P450s are natural fusion proteins consisting of a heme-containing monooxygenase domain and an electron-delivering NADPH-dependent diflavin-reductase domain.[33] CYP102A enzymes, especially CYP102A1 from

Bacillus megaterium, are widely used as biocatalysts because they are stable, self-sufficient and their variants accept a broad range of substrates.[34] Activity of these variants against Grundmann's ketone **4** was screened based on NADPH-consumption and GC/MS analysis of the products. Although some of the tested variants accepted Grundmann's ketone **4** as substrate, they all showed low regioselectivity. Moreover, none of the oxidation products could be identified as the desired 25-hydroxylated product **5**.

CYP154E1 from *T. fusca* YX is to our knowledge the first P450 found to selectively oxidize substrate **4** at position C25 without any by-products being formed. Additionally, after several hours reaction time no degradation or additional products could be detected. The use of whole-cell biocatalysis is an established method to circumvent the problems associated with the complex multi-component nature of P450s as well as with their cofactor dependency. *E. coli* has often been used as host for heterologous P450 expression and for the development of whole-cell oxidation processes.[13] The challenge of low substrate solubility and substrate uptake could be solved (at least to some extent) by the use of DMSO. To achieve an effective electron transfer in the whole-cell system, two redox partner systems, Pdx/PdR and YkuN/FdR were tested. Both systems have been reported to support CYP154E1.[32,35] Whereas putidaredoxin reductase PdR is an NADH-dependent enzyme, the *E. coli* flavodoxin reductase FdR accepts mainly the phosphorylated nicotinamide cofactor NADPH. The intracellular concentration of NADH in *E. coli* has been reported to be at least 3-times higher than that of NADPH even when the cells grew on glucose.[36] The NADPH content and its regeneration is of paramount importance for whole-cell biocatalysis using enzymes depending on this cofactor. Thus, in previous studies several strategies to increase NADPH availability in *E. coli* strains were investigated.[36–38] Arnold and coworkers were able to improve the product-per-glucose yield of the P450-catalysed propane oxidation by 230% using a metabolic engineering approach. The authors identified three *E. coli* genes originating from fermentative pathways and the endogenous respiratory chain which were disrupted to redirect metabolic flux towards NADPH regeneration.[37] Thus, the limited supply of NADPH and the lower P450 expression levels are probably responsible for the lower product concentrations achieved by *E. coli* expressing CYP154E1/YkuN/FdR in comparison to the CYP154E1/Pdx/PdR system. Furthermore, we could demonstrate that for the whole-cell biocatalyst involving the NADH dependent PdR and Pdx, oxygen supply was a limiting factor.

By combining all optimized conditions, 1.1 mM 25-hydroxylated Grundmann's ketone **5** (300 mg L^{-1}) could be produced. Compared to the chemical synthesis using $RuCl_3$/$NaIO_4$ the established biotransformation clearly has the advantage of being absolutely regioselective. Moreover, a tedious reaction workup to remove a number of by-products is omitted.

Some examples of bioconversion processes of vitamin D_3 to its active intermediate have been described.[14] Omura and colleagues intensively investigated microbial strains capable of catalysing these reactions and identified several candidates for both hydroxylation reactions.[39,40] The highest activity was demonstrated by *Amycolata* sp. which produced 8.3 mg L^{-1} 25-(OH)-D_3 and 0.17 mg L^{-1} 1α,25-$(OH)_2$-D_3 in 120 h in 200 L fermentation.[40] Later CYP105A2 responsible for this two-steps biotransformation was identified and expressed in recombinant *Streptomyces lividans* cells.[41] This whole-cell biocatalyst produced 25(OH)-D_3 at concentration of 20 mg L^{-1}. Compared to this data the resulting concentrations of 25-OH-GK **5** achieved in this study were higher and reached faster, although one should take into account that the used reaction volume was much smaller and the concentration of vitamin D_3 in water much lower.

In conclusion, we were able to transfer the established in vitro system for the 25-hydroxylation of Grundmann's ketone **4** into an *E. coli* whole-cell catalyst expressing CYP154E1 and suitable redox partners. This system allows the selective production of 300 mg L^{-1} 25-OH-GK **5** within 24 h. Although several strategies are possible to further improve these results, our approach already now serves as proof of concept for the production of this important building block used in the chemical synthesis of 25-OH-D_3 and analogous.

4. Material and methods

4.1. Materials

Escherichia coli strain DH5α was obtained from Clontech (Heidelberg, Germany). *E. coli* strain BL21(DE3) and pET expression vectors were purchased from Novagen/EMD Millipore (Darmstadt, Germany). Pfu and Phusion DNA-Polymerase, restriction endonucleases and T4 DNA ligase were obtained from Fermentas (St. Leon-Rot, Germany) and New England Biolabs (Beverly, USA).

The pET-11-based tricistronic plasmid pCYP154E1-camAB which was used for coexpression of CYP154E1, Pdx and PdR, was kindly provided by Prof. Akira Arisawa (Mercian Corporation, Japan).

Grundmann's ketone **4** was prepared by ozonolysis of vitamin D_3 according to the published procedure[23,24] and its 25-hydroxy derivative was obtained via the ruthenium tetroxide method.[28,29] NMR analysis for confirmation of both structures were obtained using an AV300 spectrometer (Bruker Corporation, Billerica, USA).

All chemical reagents were purchased from commercial sources, were of analytical grade or higher and have been used without further purification.

4.2. Cloning of *CYP154E1*, YkuN and FdR

CYP154E1 was subcloned from the pET-11 based plasmid by using *Nde*I and *Eco*RI endonucleases.[35] The resulting DNA fragment was ligated into pET-22b(+) resulting in the plasmid pCYP154E1.

YkuN was amplified by PCR with *Pfu* DNA polymerase using the following primers: *YkuN*_fw1: CG ATA CCATGG CTA AAG CCT TGA TTA CAT ATG and *YkuN*_rv2 TAT AT CTCCTT CTT ATC TCA TGA AAC ATG GAT TTT TTC. In a separate PCR reaction *fdr* was amplified with the primers *fdr_fw* GAT AAG AAGGAG ATA TAA TGG CTG ATT GGG TAA CAG and *fdr_rv* GATT GGATCC TTA CCA GTA ATG CTC CGC. As template for PCR the published plasmids pET16b-YkuP(+) and pET11a-fpr were used, respectively.[42] After purification of the PCR products they were used in an overlap extension PCR. The resulting DNA fragment was cloned into the multiple cloning site 1 of pCOLADuet-1 using the introduced restriction sites *Nco*I and *Bam*HI. The resulting construct pCOLA-YkuN-FdR was subsequently used to amplify a bicistronic fragment containing one ribosome binding site upstream of each gene. The following primers were used for PCR: *YkuN_fw2* *GAA TTC* ACT AGT TTA TAG GAG GTC TTC CAT GGC TAA AGC CTT GAT TAC ATA TG and *FdR_rv2* CAC TGC TTC CGG TAG TCA ATA AAC CGG TAA GCT TGG GGT TTC TCG AG*G AGC TC*G GAT CCT TAC CAG TAA TGC TCC GCT G. The DNA fragment was cloned into pCYP154E1 with the help of the introduced restriction sites *Xho*I and *Eco*RI leading to the tricistronic expression plasmid pE1YF.

4.3. Protein expression and quantification

Pdx/PdR (putidaredoxin and putidaredoxin reductase from *Pseudomonas putida*), FdR (ferredoxin reductase from *Escherichia coli*) and GDH (glucose dehydrogenase from *Bacillus megaterium*) were available in purified form as described elsewhere.[32] Concentration of redox partner proteins were measured by UV/VIS spectroscopy

using their established extinction coefficients.[42] P450 concentration was determined by CO-difference spectroscopy using $\varepsilon_{450-490}$ - = 91 $mM^{-1}\,cm^{-1}$ after reduction with sodium dithionite and subsequent bubbling of gaseous CO as described elsewhere.[43]

4.4. In vitro experiments

Conditions for the in vitro reactions were as follows: 200 μM substrate (10 mM stock in DMSO), 200 μM NADPH, 5 U glucose dehydrogenase, 20 mM glucose and 300 U catalase to remove H_2O_2, which might be formed due to uncoupling. P450 concentration was 1 μM, whereas the redox partners were used in rations of 1:10:1 (P450:YkuN:FdR) and 1:5:5 (P450:Pdx:PdR). Reactions were adjusted to a final volume of 500 μL with potassium phosphate buffer (100 mM, pH 7.5) and were incubated at 25 °C in standard 2 mL plastic reaction tubes for 5 h. Each reaction was afterwards extracted with ethyl acetate by vortexing, and centrifugation for phase separation. Organic phases were directly transferred to a glass vial for analysis by GC/MS.

4.5. Whole-cell bioconversion

E. coli BL21(DE3) harbouring pCYP154E1-camAB for expression of CYP154E1/Pdx/PdR or pE1YF for expression of CYP154E1/YkuN/FdR were cultured overnight at 37 °C and 180 rpm in 5 mL LB medium (100 μg mL^{-1} ampicillin). The overnight cultures were used to inoculate main cultures of 400 mL LB medium in 2 L shaking flasks and were grown to an OD_{600} of 0.9 at 37 °C and 180 rpm. Protein expression was induced by addition of 100 μM IPTG. Furthermore the cells were supplemented with 64 μg mL^{-1} 5′-aminolevulinic acid and 500 μM $FeSO_4$. Expression was performed over night at 25 °C and 180 rpm. Cells were harvested by centrifugation at 8000*g* for 15 min at 4 °C. The cell pellet was resuspended in CV2 buffer (50 mM potassium phosphate, pH 7.5, 2% glycerol, 50 μg/mL ampicillin, 100 μM IPTG) resulting in a suspension containing 210 mg cell wet weight (cww) per mL. A sample from the cell suspension was taken to determine the P450 concentration via CO-difference spectroscopy. The remaining cells were used for biotransformations. Qualitative experiments were performed in 15 mL glass tubes. From those reactions, samples of 1 mL were taken and, after addition of NaCl for enhancing phase separation, were extracted twice with 500 μL ethyl acetate. All quantitative experiments were performed in 100 mL shaking flasks containing 5 or 10 mL reaction volume (CV2 buffer with 70–210 g_{cww} L cells, 2 % DMSO and 200 μM to 3 mM substrate) at 25 °C in an orbital shaker at 180 rpm. Duplicate samples of 500 μL were taken and, after addition of NaCl and 10 μL internal standard (ITSD, 5 mM anthracene in acetone), extracted twice with 250 μL ethyl acetate.

In order to calculate TTNs a 15 ml sample of the cell suspension used for biotransformation was taken. The cells were lysed by ultrasound with a Branson Sonifier S-250A (6 × 45 s; 45% duty cycle; output control 4.5) on ice. The P450 concentration has been determined photometrically using the CO-difference method as described elsewhere.[43] Total turnover numbers were calculated as amount of **5** (in nmol, determined by GC/MS) divided by the amount of P450 (in nmol) applied in the reaction.

4.6. Product analysis

The combined organic phases were analysed on a GC/MS QP-2010 Plus instrument (Shimadzu, Tokyo, Japan) equipped with a FS-Supreme-5 30 m × 0.25 mm × 0.25 μm (Chromatographie Service GmbH, Langerwehe, Germany) fused silica column and helium as carrier gas. The injector and interface temperature was set to 300 °C. The temperature gradient started at 200 °C isothermal for 3 min. Subsequently, the temperature was raised at 15 °C min^{-1} to 320 °C and held isothermal again for 2 min. The mass spectrometer (EI, 70 eV) was set to simultaneous recording of mass spectra in scan (*m*/*z* 40–350) and SIM mode. Peak areas in SIM were used for calibration.

Acknowledgements

A.H.-M., S.F. and M.F. wish to thank the Spanish Ministerio de Ciencia e Innovación (MICINN) for financial support (Project CTQ-2011-24237). A.H.-M. thanks the Spanish Ministerio de Educación for a predoctoral FPU fellowship. C.-J.v.B. and V.B.U. thank Heinrich-Heine University Düsseldorf, Germany for financial support. We wish to thank Dr. A. Arisawa, Mercian Cooperation for the kind donation of the plasmid pCYP154E1-camAB.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2014.06.005.

References and notes

1. Brown, A. J. In Ruth Wu-Wong, J., Ed.; Vitamin D and Your Body Volume Title: Why Does Vitamin D Matter?; Bentham Science Publishers, 2012; p 83.
2. Zhu, G.-D. In Ruth Wu-Wong, J., Ed.; Vitamin D and Your Body Volume Title: Why Does Vitamin D Matter?; Bentham Science Publishers, 2012; p 57.
3. Jones, G.; Strugnell, S. A.; DeLuca, H. F. *Physiol. Rev.* **1998**, *78*, 1193.
4. Sawada, N.; Sakaki, T.; Ohta, M.; Inouye, K. *Biochem. Biophys. Res. Commun.* **2000**, *273*, 977.
5. Strushkevich, N.; Usanov, S. A.; Plotnikov, A. N.; Jones, G.; Park, H. W. *J. Mol. Biol.* **2008**, *380*, 95.
6. Yamamoto, K.; Uchida, E.; Urushino, N.; Sakaki, T.; Kagawa, N.; Sawada, N.; Kamakura, M.; Kato, S.; Inouye, K.; Yamada, S. *J. Biol. Chem.* **2005**, *280*, 30511.
7. Prosser, D. E.; Jones, G. *Trends Biochem. Sci.* **2004**, *29*, 664.
8. Bernhardt, R. In *The Encyclopedia of Biological Chemistry*; Lennarz, W. J., Lane, M. D., Eds.; Academic Press: Waltham, MA, 2013; p 607.
9. Groves, J. T. *Nat. Chem.* **2014**, *6*, 89.
10. Ortiz de Montellano, P. R *Chem. Rev.* **2010**, *110*, 932.
11. Nelson, D. R. *Biochim. Biophys. Acta* **2011**, *1814*, 14.
12. Hannemann, F.; Bichet, A.; Ewen, K. M.; Bernhardt, R. *Biochim. Biophys. Acta* **2007**, *1770*, 330.
13. Martinez, C. A.; Rupashinghe, S. G. *Curr. Top. Med. Chem.* **2013**, *13*, 1470.
14. Sakaki, T.; Sugimoto, H.; Hayashi, K.; Yasuda, K.; Munetsuna, E.; Kamakura, M.; Ikushiro, S.; Shiro, Y. *Biochim. Biophys. Acta* **2011**, *1814*, 249.
15. Sawada, N.; Sakaki, T.; Yoneda, S.; Kusudo, T.; Shinkyo, R.; Ohta, M.; Inouye, K. *Biochem. Biophys. Res. Commun.* **2004**, *320*, 156.
16. Sugimoto, H.; Shinkyo, R.; Hayashi, K.; Yoneda, S.; Yamada, M.; Kamakura, M.; Ikushiro, S.; Shiro, Y.; Sakaki, T. *Biochemistry* **2008**, *47*, 4017.
17. Fujii, Y.; Kabumoto, H.; Nishimura, K.; Fujii, T.; Yanai, S.; Takeda, K.; Tamura, N.; Arisawa, A.; Tamura, T. *Biochem. Biophys. Res. Commun.* **2009**, *385*, 170.
18. Zhu, G. D.; Okamura, W. H. *Chem. Rev.* **1995**, *95*, 1877.
19. Kametani, T.; Furuyama, H. *Med. Res. Rev.* **1987**, *7*, 147.
20. van den Bemd, G. J.; Chang, G. T. *Curr. Drug Targets* **2002**, *3*, 85.
21. Iglesias-Gato, D.; Zheng, S.; Flanagan, J. N.; Jiang, L.; Kittaka, A.; Sakaki, T.; Yamamoto, K.; Itoh, T.; Lebrasseur, N. K.; Norstedt, G.; Chen, T. C. *J. Steroid Biochem. Mol. Biol.* **2011**, *127*, 269.
22. Windaus, A.; Grundmann, W. *Justus Liebigs Ann. Chem.* **1936**, *524*, 295.
23. Condran, P.; Hammond, M. L.; Mourino, A.; Okamura, W. H. *J. Am. Chem. Soc.* **1980**, *102*, 6259.
24. Inhoffen, H. H.; Quinkert, G.; Schütz, S.; Kampe, D.; Domagk, G. F. *Chem. Ber.* **1957**, *90*, 664.
25. Fall, Y.; Vitale, C.; Mouriño, A. *Tetrahedron Lett.* **2000**, *41*, 7337.
26. Leyes, G. A.; Okamura, W. H. *J. Am. Chem. Soc.* **1982**, *104*, 6099.
27. Sardina, F. J.; Mourino, A.; Castedo, L. *J. Org. Chem.* **1986**, *51*, 1264.
28. Kiegiel, J.; Wovkulich, P. M.; Uskoković, M. R. *Tetrahedron Lett.* **1991**, *32*, 6057.
29. Sicinski, R. R.; DeLuca, H. F. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 159.
30. Bovicelli, P.; Lupattelli, P.; Mincione, E.; Prencipe, T.; Curci, R. *J. Org. Chem.* **1992**, *57*, 5052.
31. Sikervar, V.; Fleet, J. C.; Fuchs, P. L. *J. Org. Chem.* **2012**, *77*, 5132.
32. von Bühler, C.; Le-Huu, P.; Urlacher, V. B. *ChemBioChem* **2013**, *14*, 2189.
33. Munro, A. W.; Girvan, H. M.; McLean, K. J. *Biochim. Biophys. Acta* **2007**, *1770*, 345.
34. Whitehouse, C. J. C.; Bell, S. G.; Wong, L.-L. *Chem. Soc. Rev.* **2012**, *41*, 1218.
35. Girhard, M.; Tieves, F.; Weber, E.; Smit, M. S.; Urlacher, V. B. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1625.
36. Chemler, J. A.; Fowler, Z. L.; McHugh, K. P.; Koffas, M. A. *Metab. Eng.* **2010**, *12*, 96.
37. Fasan, R.; Crook, N. C.; Peters, M. W.; Meinhold, P.; Buelter, T.; Landwehr, M.; Cirino, P. C.; Arnold, F. H. *Biotechnol. Bioeng.* **2011**, *108*, 500.

38. Schewe, H.; Kaup, B.-A.; Schrader, J. *Appl. Microbiol. Biotechnol.* **2008**, *78*, 55.
39. Sasaki, J.; Mikami, A.; Mizoue, K.; Omura, S. *Appl. Environ. Microbiol.* **1991**, *57*, 2841.
40. Sasaki, J.; Miyazaki, A.; Saito, M.; Adachi, T.; Mizoue, K.; Hanada, K.; Omura, S. *Appl. Microbiol. Biotechnol.* **1992**, *38*, 152.
41. Kawauchi, H.; Sasaki, J.; Adachi, T.; Hanada, K.; Beppu, T.; Horinouchi, S. *Biochim. Biophys. Acta* **1994**, *1219*, 179.
42. Girhard, M.; Klaus, T.; Khatri, Y.; Bernhardt, R.; Urlacher, V. B. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 595.
43. Omura, T.; Sato, R. *J. Biol. Chem.* **1964**, *239*, 2379.

Whole-cell biotransformation with recombinant cytochrome P450 for the selective oxidation of Grundmann's ketone

Alba Hernández-Martín,[a,b] Clemens J. von Bühler,[a] Florian Tieves,[a] Susana Fernández,[b] Miguel Ferrero[b] and Vlada B. Urlacher[a,*]

[a]*Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany*

[b]*Departamento de Química Orgánica e Inorgánica, Instituto Universitario de Biotecnología de Asturias, Universidad de Oviedo, 33006 Oviedo (Asturias), Spain*

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

SUPPLEMENTARY MATERIAL

1. Expression of CYP105A1, CYP102A1 and CYP102A2

Recombinant CYP105A1 was expressed under the following conditions: *E. coli* BL21(DE3) cells were transformed with pET-22b(+)-CYP105A1 and transformants were selected on LB-agar plates with ampicillin (100 $\mu g \cdot mL^{-1}$). 400 mL of TB medium supplemented with 100 $\mu g \cdot mL^{-1}$ of ampicillin were inoculated with 4 mL of an overnight culture (5 mL LB medium, 100 $\mu g \cdot mL^{-1}$ ampicillin, 37 ºC, 180 rpm) and grown at 37 ºC and 180 rpm, until the optical density at 600 nm (OD_{600}) reached approximately 0.7. Then 100 µM isopropyl β-D-1-thiogalactopyranoside (IPTG), 100 µM $FeSO_4$ and 40 $\mu g \cdot mL^{-1}$ 5'-aminolevulinic acid were added and the culture was grown for another 16 h at 25 ºC and 140 rpm. Cells were harvested by centrifugation at 8000 rpm (11305 *g*) for 20 min, the supernatant was discarded and the cell pellet was resuspended in 15 mL purification buffer (50 mM potassium phosphate, pH 7.5, 100 µM phenylmethanesulfolyl fluoride –PMSF-). Cells were lysed by sonification on ice (6 x 45 s, 1 min intermission), cell debris was removed by centrifugation (48000 *g*, 4 ºC, 1 h) and the soluble protein fraction was recovered (later on referred as cell lysate).

Variants of CYP102A1 from *Bacillus megaterium* and CYP102A2 from *Bacillus subtilis* were heterologously expressed in *E. coli* BL21(DE3) cells using pET-28a/b as vector. From each of the 81 mutants, a single colony from an agar plate was inoculated into 600 µL TB medium supplemented with 30 $\mu g \cdot mL^{-1}$ kanamycin and cultured overnight at 37 ºC and 600 rpm in 2 mL deep-well titer plates. 2.6% (v/v) of the overnight cultures were used to inoculate main

cultures of 600 µL TB medium (30 µg·mL^{-1} kanamycin) in a 2 mL deep-well plate and grown at 37 ºC and 600 rpm to an OD_{600} of 0.7. P450 expression was induced by addition of IPTG (1.75 mM). Cells were shaken at 600 rpm and at 30 ºC and cultured for 22 hours, followed by centrifugation of the plate at 4000 *g* and 4 ºC for 30 min for harvesting. Supernatants were then discarded and the plate was stored at -20 ºC. Cell pellets in the microtiter plate were washed 2 times by resuspending them in potassium phosphate buffer (600 µL, 50 mM, pH 7.5) followed by 15 min centrifugation at 4000 *g* and 4 ºC. For cell lysis they were resuspended in potassium phosphate buffer (250 µL, 50 mM, pH 7.5) containing lysozyme (1 mg·mL^{-1}), $MgCl_2$ (10 mM) and DNAse (0.05 U·mL^{-1}). After incubation at 4 ºC and 100 rpm for 3 hours, cells were treated with 3 freeze-thaw cycles (3 h at -80 ºC followed by 1.5 h at room temperature, cells being resuspended prior to re-freezing). Cell debris was removed by centrifugation (4000 *g*, 4 ºC, 1 h) and the soluble protein fraction was transferred into 1.5 mL microtubes.

For every P450 screened, expression levels were estimated using the CO-difference spectral assay as described in the literature using $\varepsilon_{450\text{-}490} = 91$ mM^{-1}·cm^{-1}.[1]

2. CYP102 activity screening

CYP102 family variants activity against Grundmann's ketone was checked by a NADPH consumption assay. The assay was performed by addition of NADPH (200 µM) to a cell lysate/substrate mixture containing cell lysate (1:4 diluted) and Grundmann's ketone (100 µM – from a stock solution in DMSO-) in a final volume of 200 µL potassium phosphate buffer (50 mM, pH 7.5). Absorbance at 340 nm was measured with a micro plate spectrophotometer (SPECTRAmax plus 384, Molecular Devices, USA) directly after addition of NADPH. Samples were compared with a blank reaction containing DMSO (4 µL) instead of substrate. Mutants showing a remarkable higher slope in the presence of the substrate in comparison with the curve for the blank reaction were supposed to be active against Grundmann's ketone and therefore selected for *in vitro* conversion experiments.

3. *In vitro* experiments

Conditions for the *in vitro* reactions with CYP105A1, CYP102A1 and CYP102A2 were as follows: 100 µM substrate (from 10 mM stock solution in DMSO), 200 µM NAD(P)H, 5 U of GDH (for cofactor regeneration), 20 mM glucose (co-substrate for glucose dehydrogenase) and 120 U catalase (for removing H_2O_2, which might be formed due to uncoupling). P450 concentration was variable in each case: CYP105A1 was added as cell lysate in a final concentration of 5 µM and in the CYP102 family screening, 120 µL of cell lysate were added in each case, with an average final concentration on the reaction of 1 µM. CYP105A1 was used in

different experiments with two different redox partners: Pdx:PdR (30 μM:15 μM) and Fdx:FdR (50 μM:5 μM). Reactions were set up to a final volume of 500 μL with potassium phosphate buffer (50 mM, pH 7.5). The reaction mixtures were incubated at 25 ºC in standard 2 mL closed plastic tubes overnight (except for BM-3 mutants, where conversion was measured after 3 hours). Each reaction was afterwards extracted with EtAcO (2x500 μL) by vortexing 5 min, and centrifugation for phase separation. Organic phases were directly transferred to a GC-MS 2 mL glass vial for qualitative determination of activity.

[1] Omura, T.; Sato, R. *J. Biol. Chem.* **1964,** *239,* 2379–2385

6 Crystallization of CYP154E1

6.1 Introduction

Elucidation of the structure-function relationship (SFR) of enzymes in general and of P450s in particular is a crucial step to understand their observed properties.[131] Hence, the SFR is an important key factor in many fields of work related to P450 monooxygenases. Among others it is a prerequisite to rationally modify and design novel biocatalysts.[79,132] The availability of X-ray diffraction as a standard tool for the determination of three-dimensional protein structures made it possible to solve P450 structures routinely. Thus, to date more than 600 P450 structures solved by X-ray diffraction are listen in the PDB.
In order to understand the differences and similarities in the activities and product selectivities of the two P450s CYP154A8 and CYP154E1 which were observed in the cluster screening[122] their crystal structures should be investigated. Furthermore, molecular determinants of the broad substrate spectrum of the CYP154 family should be identified.
CYP154E1 was chosen because in case of CYP154A8 the crystal structure of a close homologue, CYP154A1, already has been published.[106] A second reason for this choice was that, in contrast to CYP154A8, CYP154E1 is the sole member of its subfamily and no further structural knowledge about this enzyme is available. Thirdly, the origin from a thermophilic host potentially confers thermostability to CYP154E1. The elucidation of structural features which are responsible for thermostability were an additional motivation to focus our crystallization efforts on CYP154E1.

6.2 Experimental procedures

6.2.1 Materials and strains

All reagents were obtained from commercial sources and were of analytical grade or higher. Enzymes for molecular cloning were purchased from Fermentas (St. Leon-Roth, Germany) and New England Biolabs (Ipswich, USA). The *E. coli* strain DH5α was obtained from Clontech (Heidelberg, Germany). The *E. coli* strain BL21(DE3) and the plasmid vectors pET-22b(+) and pET-28b(+) were purchased from Novagen (San Diego, USA). The pET-28a(+)-based plasmid for expression of CYP154E1 was used as published previously.[102]

6.2.2 Molecular cloning of CYP154E1 variants

Variants of CYP154E1 were amplified by PCR with Phusion DNA polymerase from the pET-28a(+) plasmid coding for His_6-tagged CYP154E1 according to the manufacturer's protocol. For the variant CYP154E1t primers CVB_154E1t-NdeI-F / ARO_154E1-EcoRI-R were used. For CYP154E1-CHis primers CVB_154E1-CHis-F / CVB_154E1-CHis-R were employed. Details for all primers are listed in Tab. 5. After purification, the PCR products were separately cloned into pET-22b(+) (CYP154E1-CHis) or pET-28a(+) (CYP154E1t) by restriction and subsequent ligation using the introduced restriction endonuclease sites (Tab. 5). The sequence of the resulting plasmids was verified by a commercial DNA sequencing service (GATC, Konstanz, Germany). Plasmid maps of all constructed vectors can be found in section 9.2.

Table 5: Oligonucleotides used for molecular cloning of CYP154E1. Restriction sites in the primer sequence are printed in lower case.

Designation	Sequence (5' → 3')	Restriction site
CVB_154E1t-NdeI-F	GGATTAC catatg GCTTTTCCCGGGAACC	*Nde*I
ARO_154E1-EcoRI-R	CAGT gaattc TCAGGGTTTCGGGCG	*Eco*RI
CVB_154E1-CHis-F	GGATCTA catatg GGACAGTCCCGCCGAC	*Nde*I
CVB_154E1-CHis-R	TAT ctcgag GGGTTTCGGGCGCAAGAG	*Xho*I

6.2.3 Recombinant expression, purification and modification of CYP154E1

Expression and protein purification via nickel affinity chromatography was performed as described in section 2.2.3. To further enhance the purity of the CYP154E1 variants for crystallization, size exclusion chromatography (SEC) on a HiLoad Superdex 200 PG column (GE Healthcare, Freiburg, Germany) was performed. 50 mM Tris/HCl buffer pH 7.5 containing 150 mM NaCl and 2 % glycerol was used as isocratic eluent.
The endoprotease thrombin was used to cleave off the N-terminal His_6-tag from CYP154E1 (His-Tfus_2243). This was possible since pET-28a(+) contains a sequence coding for the thrombin recognition site (LVPR↓GS) between the His_6-tag and the *Nde*I restriction site. The commercial Thrombin Cleavage Capture Kit (EMD Chemicals Inc., an affiliate of Merck Bioscience, Darmstadt, Germany) was applied to purified protein and a cleavage protocol based on the manufacturer's suggestions was developed: 5 mg CYP154E1 (His-Tfus_2243), 500 µl 10X buffer, 2.5 U thrombin in a total volume of 5 ml ddH_2O. The reaction was performed at 20 °C in a heating block overnight. Subsequently, the cleaved His_6-tag was removed by nickel affinity chromatography. Size exclusion chromatography was applied to remove thrombin from the digested CYP154E1-Thr. To determine the success of the thrombin cleavage, samples taken after each step of the thrombin treatment were loaded on an SDS-polyacrylamide gel (12.5 %). After separation on an SDS-polyacrylamide gel the proteins were blotted on a nitrocellulose membrane in a semi-dry blotting apparatus (Biometra, Göttingen, Germany) according to the manufacturer's protocol. Proteins on the membrane containing a His_6-tag were specifically detected with the Anti-His_6-Peroxidase conjugate (Roche, Mannheim, Germany) converting "BM Blue POD substrate, precipitating" (Roche, Mannheim, Germany).

6.2.4 Characterization of CYP154E1 and CYP154A8

In order to determine the thermostability of CYP154E1 and CYP154A8, the proteins were incubated at different temperatures. The residual amount of correctly folded P450 after incubation was determined by CO-difference spectroscopy.[125]
In 100 mM potassium phosphate buffer pH 7.5 a master mix containing 6.88 µM CYP154A8 or CYP154E1 was prepared and samples of 125 µl volume were aliquoted to plastic PCR tubes. The samples were heated in a thermal cycler for 10 min to different temperatures between 30 °C and 85 °C. Thereafter the samples were stored on ice. The relative residual P450 concentration was calculated with respect to the P450 concentration in the master mix after

incubation at 20 °C.
Electrospray ionization time of flight mass spectrometry was performed at the "Biologisch-Medizinisches Forschungszentrum" (BMFZ), Heinrich-Heine-Universität Düsseldorf. Prior to the measurements a buffer exchange of the protein samples in 10 mM ammonium formate buffer, pH 7.5 was performed on a PD-10 Desalting column (GE-Healthcare, Freiburg, Germany). After diluting the protein 1:1 with methanol to a concentration of 1 mg ml^{-1} the samples were measured on an Applied Biosystems QStar XL ESI-TOF instrument.
Size exclusion chromatography (SEC) multi angle light scattering (MALS) experiments with purified CYP154E1 were performed at the "Institute of Biochemistry I", Heinrich-Heine Universität Düsseldorf. For the chromatography an ÄKTA purifier system (GE Healthcare, Freiburg, Germany) equipped with a Superdex 200 10/300 GL column (GE Healthcare, Freiburg, Germany) at a flow of 0.5 ml min^{-1} 50 mM Tris/HCl buffer pH 8 containing 150 mM NaCl was used. Refraction index and light scattering were measured on-line with an Optilab rEX and a miniDAWN TREOS detector (Wyatt Technology Europe, Dernbach, Germany). Data analysis was performed with the Astra software package.

6.2.5 Protein crystallization

Crystallization experiments were performed at room temperature using the sitting drop and the microbatch method. The appropriate protein concentration for the initial crystallization experiments was determined by ammonium sulfate precipitation (0.4-1.4 M in 50 mM HEPES buffer pH 7) in 24-well plates. The tested concentration range of CYP154E1 was 1.85-18.5 mg ml^{-1}.
The initial screening for crystallization conditions was conducted according to the sparse matrix sampling method first published in 1991.[133] Commercial kits listed in Tab. 6 were used to identify suitable crystallization conditions. For this purpose 90 µl of each solution contained in the screening kits (Tab. 6) were transferred into a reservoir well of a Crystal*EX*™ 96-well plate (Corning Inc., Corning, USA). In the protein well 500 nl of the CYP154E1 solution (10 mg ml^{-1}) were diluted with an equal amount of reservoir solution and subsequently equilibrated against the reservoir. Formation of protein crystals was regularly assessed.
Crystals for X-ray diffraction measurements were grown in Cryschem 24-well sitting drop plates (Hampton Research, Aliso Viejo, USA) with a drop size of 3 µl. The reservoir was filled with 400 µl of the precipitant solution identified with the screening kit. In the corresponding protein well 1.5 µl precipitant from the reservoir was used to dilute 1.5 µl of the protein solution (3.6-24 mg ml^{-1}) prior to the equilibration by vapor diffusion. For the identification of

beneficial additives enhancing crystal growth and diffraction behavior a collection of additives (additive screen I-III, Hampton Research, Aliso Viejo, USA) was screened. In a 24-well plate 0.3 µl of an additive was added to the respective protein well containing the aforementioned mixture of protein solution (14 mg ml^{-1}) and precipitant (100 mM Tris/HCl pH 8, 75 % MPD).

Microbatch reactions were set up according to the protocol first described by Chayen et al.[134] 72-well Terasaki plates (Greiner Bio One, Frickenhausen, Germany) were first filled with Paraffin oil and subsequently the wells were filled with a mixture of 1.5 µl protein solution (10-16 mg ml^{-1} CYP154E1) and 1.5 µl precipitant (100 mM Tris/HCl pH 7.5, 70-90 % MPD).

Crystal seeding experiments were performed in 24-well plates. Serial dilutions of crystal fragments were added to wells containing the same solution of precipitant and protein in which the crystal fragments were originally grown. Thereby the protein concentration was lowered to avoid any additional nucleation and to allow the protein only to accumulate on the added crystal fragments.

Table 6: List of all commercial crystallization screens used in this work. The upper part was used in initial screening. The lower set was used to screen for alternative crystallization conditions.

Manufacturer	**Screen**
Qiagen (Hilden, Germany)	JCSG Core I Suite, JCSG Core II Suite, JCSG Core III Suite, JCSG Core IV Suite, The MPD Suite, The Classics Suite, The Classics Suite II, The PEGs Suite, The PEGs II Suite, The ComPas Suite
Qiagen	The PACT Suite, MbClass Suite, MbClass II Suite, The Cryos Suite
Molecular Dimensions Ltd. (Newmarket, UK)	MemPlus, MemSys, MemGold, MIDAS

6.3 Results

6.4 Protein characterization

CYP154E1 was expressed in *E. coli* as N-terminally His_6-tagged protein (His-Tfus_2243) and purified to homogeneity (Fig. 14) as described in the "Experimental procedures" section (6.2.3).

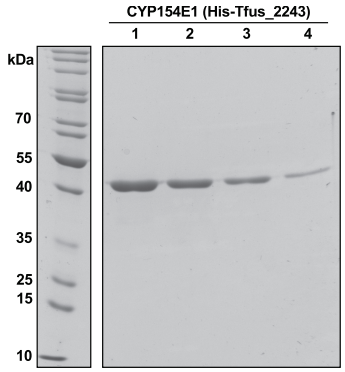


Figure 14: SDS-polyacrylamide gel of purified CYP154E1 (His-Tfus_2243) after nickel affinity chromatography and SEC. P450 amount per lane: 1 - 1 µg, 2 - 0.5 µg, 3 - 0.25 µg, 4 - 0.1 µg.

The expression in *E. coli* yielded up to 510 nmol purified P450 per L medium. CYP154E1 was subsequently characterized by multi angle light scattering (MALS) and electro spray ionization time of flight mass spectrometry (ESI-TOF-MS). Both measurements demonstrated that the enzymes molecular weight was in good accordance with the expected value of 47216.4 Da (Fig. 15). Furthermore the MALS measurements indicated that the protein in solution exists as monomers. The molecular weight determined by the ESI-TOF-MS measurements provided proof positive that CYP154E1 was a fully heme-loaded holoprotein.

CYP154E1 originates from *T. fusca* which is a moderate thermophilic soil bacterium. Hence, the thermostability of this enzyme was investigated and compared to the thermostability of CYP154A8 from the mesophilic bacterium *N. farcinica*. The characteristic absorbance at 450 nm for reduced and CO bound P450 is only detected for hexa-coordinated heme iron with an axial thiolate ligand. With increased temperature the enzyme becomes more and more flexible and finally unfolds. As a consequence, the heme group looses its axial ligand. This can be directly observed in the UV/VIS spectrum. Since the axial cysteinate ligand is also a requirement for the P450's catalytic function, absorbance at 450 nm serves as a direct

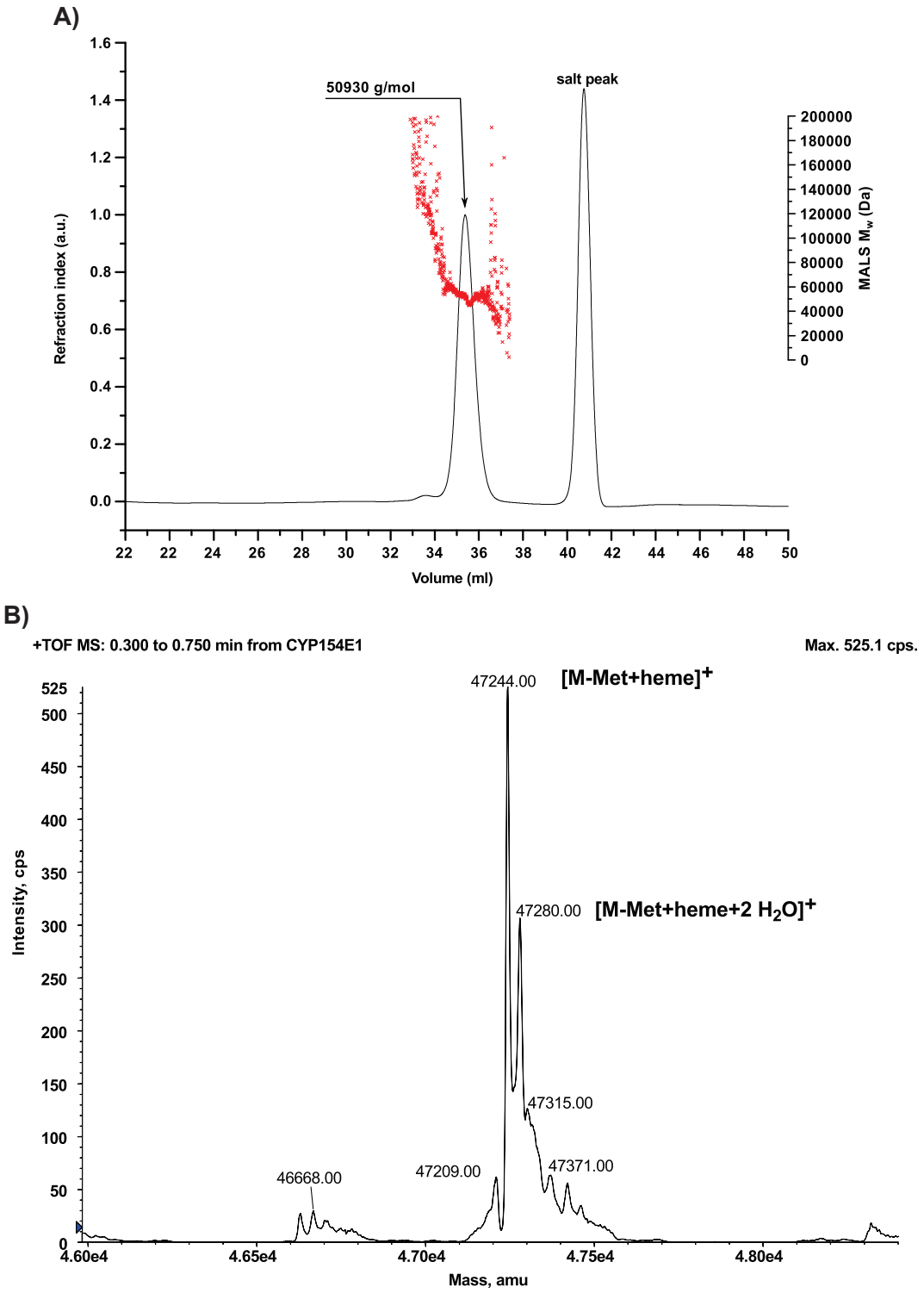


Figure 15: **A)** SEC-Chromatogram of CYP154E1 with detection of the refraction index and the MALS signal (red) to calculate the apparent molecular weight (M_w). **B)** Deconvoluted ESI-TOF-mass spectrum of CYP154E1.

measure for the residual amount of catalytically active P450. Thus, the residual P450 concentration after 10 min incubation of a solution of purified enzyme at elevated temperatures was measured via the CO-difference method and served as an estimation of its thermostability.

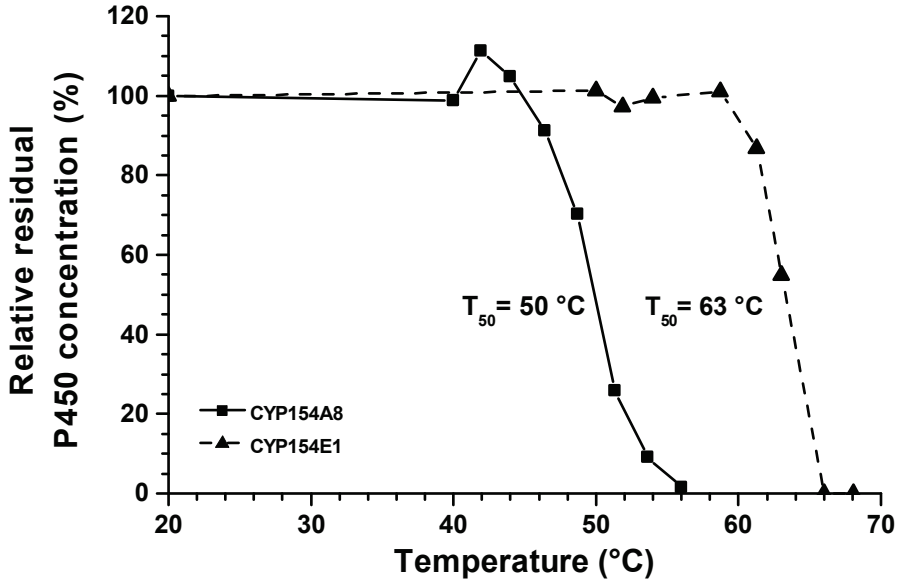


Figure 16: Denaturation curve obtained for CYP154E1 and CYP154A8. Measurements were performed in a thermo cycler. The relative residual concentration of the CO-bound P450 complex after 10 min at the indicated temperature was determined by CO-difference spectroscopy.

The concentration of correctly folded CYP154E1 did not decrease significantly up to 60 ℃. This was a clear indication that the P450 retained its stability and integrity under the chosen conditions. The denaturation curve showed a rapid decrease of correctly folded P450 between 60 ℃ and 70 ℃ (Fig. 16). At higher temperatures the correct coordination of the heme group was completely lost and CYP154E1 can be considered inactive. The T_{50}-value defined as the temperature at which the activity of the enzyme is reduced by half after incubation for 10 min,[135] was determined from to be 63 ℃. For CYP154A8 the same experiments have been performed. Incubation in the temperature range between 20 ℃ and 40 ℃ did not affect the enzyme. The maximal P450 concentration was observed after incubation at 42 ℃. Further increased temperatures resulted in a rapid loss of correctly folded P450. After incubation at 56 ℃ no residual CYP154A8 could be detected anymore. From this data the T_{50}-value of 50 ℃ was determined.

6.4.1 Initial crystal screening

The protein concentration of 10 mg ml^{-1} was found to be appropriate for the initial screening and was determined by ammonium sulfate precipitation. Subsequently the screening was performed with the first set of screens listed in Tab. 6. Out of the 960 different precipitants after 5 days only one showed formation of crystals (Fig. 17 A).

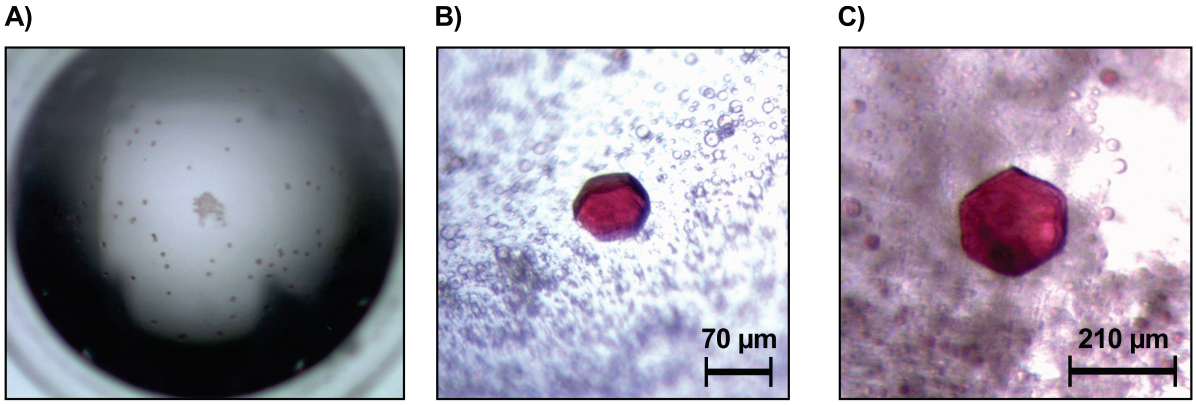


Figure 17: Crystals of CYP154E1 after **A)** initial screening, crystal size ~30 µm, **B)** upscaling and fine screening, **C)** addition of $Co(NH_3)_6Cl_3$.

The condition (The MPD Suite: F11) contained 100 mM Tris/HCl buffer pH 8, and 65 % 2-methyl-2,4-pentanediol (MPD). On a 24-well plate an upscaling and subsequent fine screening revealed that an MPD concentration of 70 % was beneficial for the larger sample volumes. In these experiments a crystal size of ~60–70 µm was achieved (Fig. 17 B). To further increase the crystal size a macro seeding experiment was performed but unfortunately did not lead to larger crystals. As a further way to increase the mass of the crystals an additive screening was performed in order to identify additives which are beneficial for crystal growth. This revealed that the addition of 10 mM $Co(NH_3)_6Cl_3$ allowed the crystals to grow to a diameter of up to ~200 µm (Fig. 17 C).

These optimized conditions (100 mM Tris/HCl buffer pH 8, and 70 % MPD and 10 mM $Co(NH_3)_6Cl_3$) were used to produce CYP154E1 crystals for the X-ray diffraction measurements. Additionally, the known P450 inhibitors 4-phenylimidazole (1 mM) and 2-methylimidazole (1 mM) as well as the substrates undecanoic acid (40 µM) and geraniol (100 µM) were cocrystallized. Ten crystals resulting from these different crystallization conditions were subjected to X-ray diffraction experiments (ESRF, Grenoble, France). Unfortunately, they all displayed very inhomogeneous diffraction behavior, exhibiting sharp reflections in one dimension and ill-defined ones in another spacial direction (Fig. 18). Due to this

diffraction behavior in this experiment no complete data set could be obtained.

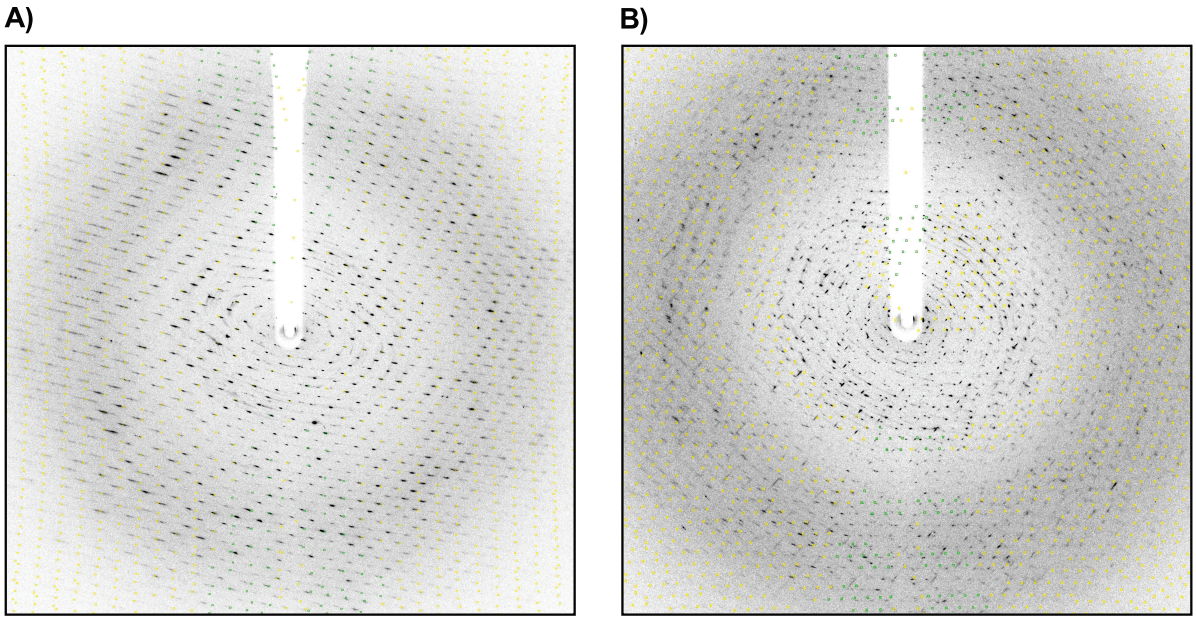


Figure 18: X-ray diffraction patterns of two CYP154E1 crystals cocrystallized with **A)** undecanoic acid and **B)** geraniol. Irregular formed spots can be identified on both images. Spots with a green mark can be assigned to a space group, those with a yellow mark cannot. (A and B: λ=0.8726 Å, $t^{Exp.}$=1 s, oscillation range: 1°)

6.4.2 Improving crystal quality

The irregular formed spots indicate that the degree of order in the crystal was not high enough for defined X-ray diffraction. The incomplete loading of the apoprotein with heme and the formation of dimers could be excluded due to the ESI-TOF-MS and MALS experiments (Fig. 15). Thus, it was hypothesized that the high degree of disorder might be due to the crystallization conditions.

The process of initial screening was repeated with the second part of the kits listed in Tab. 6. A second crystal forming precipitant could be identified. Similarly to the first identified crystallization condition it contained MPD. The major difference was the buffer substance 2-(Bis(2-hydroxyethyl)amino)acetic acid (Bicine) and the higher pH value (pH 9). The change of conditions resulted in crystals, which optically seemed to be of higher quality. The edges were sharper and crystal faces were flat without irregularities. X-ray diffraction of eight crystals of 40–50 µm diameter was measured on a synchrotron beam line (ESRF, Grenoble, France). The quality of the diffraction data was still low and again did not lead to complete data sets.

Since the changed crystallization conditions did not have a positive influence on crystal quality further possible reasons for crystal disorder should be identified. Literature research revealed that the amino terminus of CYP154A1 (PDB-ID: 1ODO) and CYP154C1 (PDB-ID: 1GWI) has not been completely resolved. Six and seven amino acid residues are not included in their crystal structures, respectively. This led to the second working hypothesis: A suitable N-terminus might be crucial for highly ordered protein crystals. In order to verify this hypothesis further CYP154E1 variants with alterations at the N-terminus were investigated. For their construction several factors had to be considered. In the course of this work the genome sequence of *T. fusca* including the annotation of open reading frames (ORF) was updated. This resulted in a new annotation of the CYP154E1-gene (Tfu_2976, AAZ57009) with an N-terminal truncation of 35 amino acids compared to the previously annotated sequence (Tfus_2243, ZP00058867). The differences are visualized in Fig. 19, line 1 and 3.

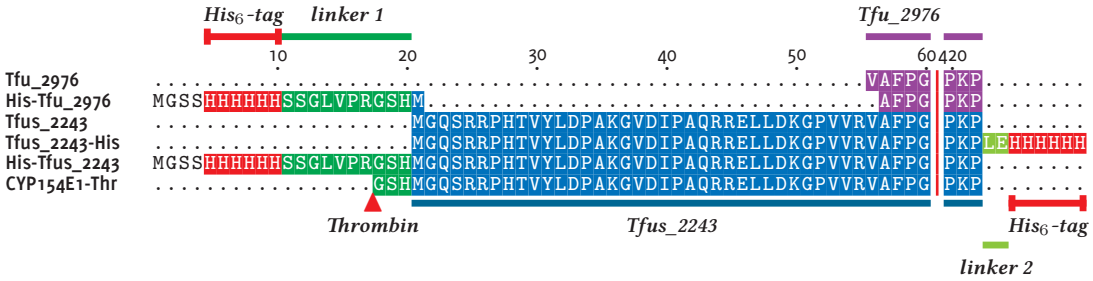


Figure 19: Alignment of all CYP154E1 variants used in this work. Tfus_2243, colored in blue, was the longer sequence originally deposited in GeneBank. In the updated version of the *T. fusca* genome the locus id was reassigned to Tfu_2976, colored in purple, and the open reading frame was N-terminally shortened by 35 amino acid residues. Sequence names are composed from locus id and the respective terminal modification.

The first tested variant of CYP154E1 was the N-terminally truncated and His_6-tagged version CYP154E1t (His-Tfu_2976, Fig 19, line 2). Samples taken during and after expression of the protein were analyzed by SDS-PAGE (Fig. 20). On the SDS-polyacrylamide gel the appearance of a band at ~40 kDa corresponds well with the expected molecular weight of CYP154E1t (42854 Da). Furthermore, the intensity of the band suggested that the protein concentration of CYP154E1t still increased after 2 h of expression. Nevertheless, after expression overnight the majority of the protein was found in the insoluble protein fraction. The residual soluble protein was assayed via CO-difference spectroscopy. No characteristic absorbance at 450 nm of the CO-bound P450 complex could be observed. This indicated that the heme coordination was incorrect and the protein inactive.

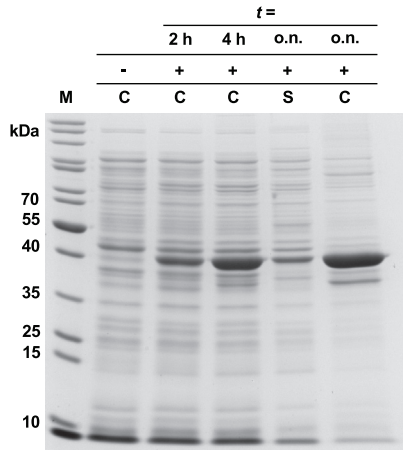


Figure 20: SDS-polyacrylamide gel of the expression of CYP154E1t (His-Tfu_2976). **M**: protein marker, **C**: whole cell sample, **S**: sample of the soluble protein fraction after cells lysis and centrifugation, **-/+**: before/after induction with IPTG, **o.n.**: overnight. Time points ***t*** of sampling are indicated in the top line.

Truncation of the N-terminus in Tfu_2976 obviously led to misfolded and inactive protein. It was concluded that the protein might be sensitive to N-terminal modification. Without sacrificing the convenience of affinity chromatographic purification the His_6-tag was moved to the 3'-end of the gene. Protein expression and purification was successful (Fig. 21) using the established two step protocol. CYP154E1-CHis exhibited a CO-difference spectrum similar to the N-terminal His_6-tagged variant (His-Tfus_2243, p. 50, Figure S1). The expression level of CYP154E1-CHis was 1150 nmol P450 per L medium. This value is about 125 % higher than that measured for N-terminally tagged CYP154E1 (His-Tfus_2243) under identical expression conditions (see p. 41).

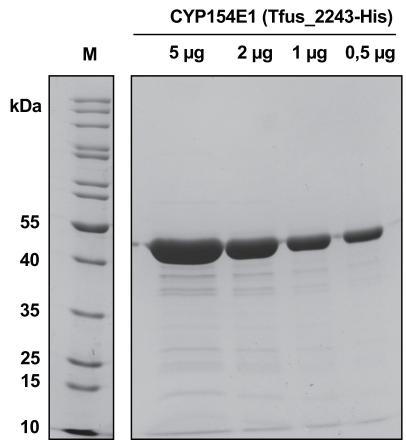


Figure 21: SDS-polyacrylamide gel of purified CYP154E1-CHis (Tfu_2243-His). **M**: protein marker. The amount of protein loaded per lane is indicated in the top line.

A last variant of CYP154E1 without any tags termed "CYP154E1-Thr" has been obtained by removing the His_6-tag from CYP154E1 via thrombin cleavage. This enzymatic restriction reduced the N-terminal modification to the tripeptide "GSH" originating from the thrombin recognition site and a single amino acid linker. The success of the cleavage reaction could be monitored by SDS-PAGE and western blotting with anti-His_6 immunodetection (Fig. 22). The calculated purification yield was 72 % based on the P450 concentration determined by CO-difference spectroscopy.

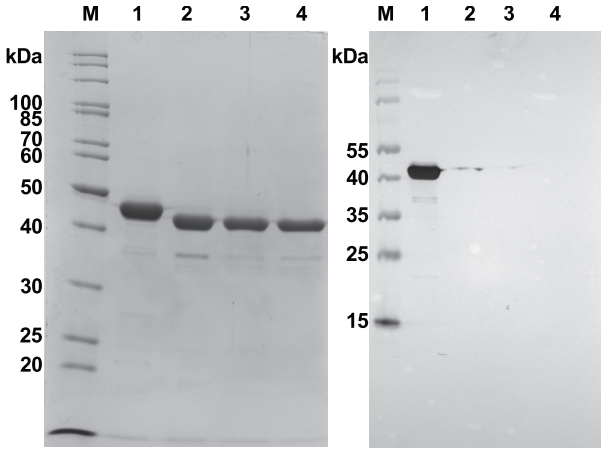


Figure 22: left: SDS-polyacrylamide gel and **right**: Western blot with Anti-His_6-immunodetection of NHis-CYP154E1 samples treated with thrombin. **M**: protein marker, **1**: before thrombin addition, **2**: after the cleavage reaction, **3**: after nickel affinity chromatography, **4**: after SEC.

Both variants, CYP154E1-CHis and CYP154E1-Thr, crystallized with the established conditions. Since the concentration of the CYP154E1-Thr solution was only 3.6 mg ml^{-1}, crystal growth was significantly slower, several weeks instead of days.

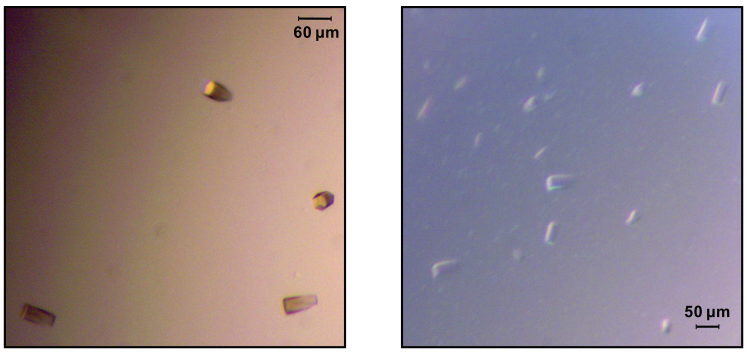


Figure 23: Crystals of CYP154E1-Thr crystallized with the precipitant 100 mM Tris/HCl at pH 8, 70 % MPD. The protein solution (1 µl, 3.6 mg ml^{-1}) was diluted with 1 µl precipitant (**left**) and 2 µl (**right**).

The formed crystals of CYP154E1-Thr exhibited a hexagonal prismatic geometry (Fig. 23)

compared to the tagged variants whose crystals were both hexagonal planar (Fig. 17). X-ray diffraction of both variants has been measured on a synchrotron beam line (DESY, Hamburg, Germany). Two full data sets with 4.0 Å and 2.8 Å resolution respectively could be obtained. The CYP154E1-Thr crystals were grown in 70 % MPD, 100 mM Tris/HCl at pH 8. The protein crystals demonstrated an apparent P6322 symmetry. With the resulting datasets the solution of the protein's structure was not possible since both crystals were perfect twins with a twinning ration α of 0.5. This prevented the solution by molecular replacement since the contribution to the intensity of each reflection could not be assigned.

6.5 Discussion

The substrate scope of two members of the CYP154 family has been successfully investigated with the developed cluster screening method (see chapter 3). In order to rationalize the observed properties, a deeper understanding of the molecular determinants which are responsible for the substrate spectrum as well as the regio-, chemo-, and stereoselectivity is essential. In the long term, this knowledge will eventually enable the tailoring of P450 biocatalysts for novel applications and perhaps even the creation of universal biocatalysts. Besides activity and reactivity further important properties like solvent- and thermostability can probably be rationalized based on the three-dimensional structure of an enzyme. Hence, an important aim of this study was to obtain crystals of CYP154 enzymes in order to be able to solve their crystal structure via X-ray diffraction.
CYP154E1 exhibited a T_{50} of 63 °C which was together with its interesting catalytic properties encouraging enough to select it as the candidate for protein crystallization. Crystallization experiments were performed in close cooperation with the "Crystal and X-Ray Facility", Heinrich-Heine-Universität Düsseldorf. A screening of possible precipitants with the sparse matrix sampling method revealed a crystal forming condition for N-terminally His_6-tagged CYP154E1: 100 mM Tris/HCl buffer pH 8 containing 65 % MPD. After upscaling and fine screening of the identified sitting drop crystallization condition, crystals with a diameter of ~60 µm could be obtained. Subsequent additive screening led to an improved precipitant containing additionally 10 mM $Co(NH_3)_6Cl_3$. These conditions allowed the hexagonal planar crystals to grow to a diameter of ~200 µm. For ten crystals X-ray diffraction has been measured but no complete data set could be obtained. To further improve the diffraction quality of the crystals additional precipitants were screened and a second crystal forming condition (100 mM Bicine buffer pH 9, 70 % MPD) could be identified. Unfortunately, this did not im-

prove the diffraction of the formed crystals.
To obtain high quality crystals a further variant of CYP154E1 with a His_6-tag located at the C-terminus instead of the N-terminus was constructed. Increased expression levels of CYP154E1-CHis in *E. coli* of about 125 % were the result of the changed position of the affinity tag.
A third variant of CYP154E1 was obtained by treating the original N-terminally His_6-tagged enzyme with the protease thrombin to cleave the affinity tag from the intact protein resulting in the version CYP154E1-Thr. Both new variants were subjected to crystallization experiments under the conditions identified earlier. Crystallization was observed in both cases and for CYP154E1-Thr even the geometry of the crystals changed to a hexagonal prismatic form.
X-ray diffraction measurements of the new crystals provided the first complete data sets. Nevertheless, the crystal quality did not suffice to solve the structure by molecular replacement since the crystals were twinned. Twinning is a known problem in protein crystallography and occurs when a crystal consist of more than one separate domain having different orientations.[136] It is distinguished from other crystal growth anomalies by the partial or complete coincidence of the lattices of each domain.[137] Therefore, the observed scattering intensities in X-ray diffraction are the weighted sum of the reflections of the sublattices. The ratio of the volume of the separate domains is called "twinning fraction" α. If α is different from 0.5 it can be possible to compensate the phenomenon during data analysis as demonstrated e.g. for CYP245A1 (PDB-ID: 3A1L).[138] If this does not succeed or α is equal to 0.5 (perfect twinning), then twinning has to be circumvented or at least reduced by a change in the crystallization procedures. From literature it is known, that the crystallization speed of the different domains is often slightly different. These differences can be exploited by lowering the crystallization temperature. Following this approach, the twinning of bacteriorhodopsin could be drastically reduced.[139] A further strategy frequently found in literature would be to screen for additional conditions leading to untwinned crystals.[140,141] Recent work on bacteriorhodopsin demonstrated that it is possible to separate two domains of a twinned crystal by slowly lowering the salt concentration of the mother liquor.[142] For CYP154E1-crystals this method would be applicable if the crystal only consists of two large domains. Finally, the variation of the affinity tag might be an option to obtain untwinned crystals as this has been reported for other proteins to have an influence on twinning.[143] In conclusion, an important step in the direction of solving the crystal structure of CYP154E1 has been achieved. However, further improvements of the diffraction properties of CYP154E1-crystals are required prior to solve their crystal structure.

7 Concluding discussion

The interest in P450 monooxygenases for biocatalysis is high since these enzymes are able to oxidize even non-activated carbon atoms. Ideally, a versatile P450-based biocatalyst for synthetic applications should accept a broad range of substrates, which are chemo-, regio-, and stereoselectively oxidized to a single product. Moreover, a biocatalyst should possess high activity and should be stable under process conditions. Finally, the expression of the enzyme in sufficient amounts should be possible in a simple prokaryotic system like *E. coli*.[69]
Although genome sequencing projects have rapidly increased the number of annotated P450 gene sequences in the recent years, the choice of enzymes is still difficult. Eukaryotic P450s, which are responsible for xenobiotic detoxification, have the desired broad substrate spectra, but their regioselectivity is often low. Furthermore, due to their eukaryotic origin, they often lead to low expression in *E. coli* and often are not stable.
In contrast, most bacterial P450s are soluble enzymes and frequently show high expression levels in *E. coli*. Often their limitation is the narrow substrate spectrum as they naturally often catalyze single steps in different biosynthetic pathways, only. CYP102A1 as one example for a bacterial P450 has been extensively engineered to extend its substrate spectrum.[81,144–147] Nevertheless, the regioselectivity of engineered CYP102A1 variants is still far from perfect.
Contrasting other P450 families, the substrate spectra of the characterized CYP154-family enzymes is not limited to one small class of chemical compounds. Whereas the CYP154 family in general accepts a broad substrate spectrum, the single enzymes often display high product selectivity. Two members of different subfamilies, CYP154E1 and CYP154A8, with 41 % sequence identity were chosen to characterize their substrate spectrum. Furthermore, their potential as biocatalyst in *in vitro* and *in vivo* systems was explored.
In the following, the achieved results of this work will be examined with respect to the aforementioned ideal properties of a P450-based biocatalyst. This will demonstrate the progress made towards the long-term goal: Development of a versatile P450-based biocatalyst and the elucidation of molecular determinants responsible for these properties.

7.1 Selectivity

"Selectivity" of P450 catalysis refers to several different aspects. First of all, it includes the substrate spectrum of a P450 enzyme which ideally should be very broad. Furthermore, it includes product selectivity, namely chemo-, regio-, and stereoselectivity of an enzyme in the catalyzed reaction. The selectivity of CYP154E1 and CYP154A8 had to be investigated. Established screening methods used to determine the substrate spectra of non-characterized P450s often do not reveal their product selectivity or, even more important, are not based on the detection of products at all. In this work an advanced screening approach, called cluster screening, was developed which enables the simultaneous collection of information on the accepted substrates and the formed products. As it relies on the chromatographic separation and mass spectrometric detection of the formed products it also elucidates the chemo- and regioselectivity of P450 enzymes. For the screening a library of 51 organic molecules was compiled. The compounds have been clustered into nine groups according to their chemical properties and size. The systematic chemical relationship of the compounds in the cluster library enables the identification of new substrates for the tested P450s. Even more importantly, also the structural requirements which have to be fulfilled by a chemical compound to be converted can be distinguished. By sacrificing the feasibility of high throughput in the chosen screening methodology, detailed information about the substrate space and product selectivity was obtained.

Phylogenetic analysis of the CYP154 family revealed that CYP154E1 and CYP154A8 belong to different branches of the phylogenetic tree. Furthermore, their sequence identity (41 %) is at the lower limit for P450s from the same family. Hence, differences in their selectivities were expected. Thirty compounds from the cluster library were identified as substrates for CYP154E1. The homologous CYP154A8 accepted 23 of the tested compounds. To a large extent the substrate spectra of both enzymes were overlapping in groups I-V (small, mostly rod-shaped molecules). Differences were observed for the substrates heptanoic acid, farnesyl acetone, octane-1-thiol, and 2,4,6-trimethyloctanoic acid which were only accepted by CYP154E1. In general, CYP154A8 displayed higher product selectivity but lower activity than CYP154E1.
In groups VI-IX makro- and polycyclic compounds with a higher steric demand have been tested. In these groups the substrate spectra of both enzymes differed even more than in the groups I-V. Six compounds from group VI (bicyclic compounds) were accepted by CYP154E1 but only one, Grundmann's ketone, was also converted by CYP154A8. An impor-

tant property of the aforementioned drug-metabolizing P450s is their large binding pocket. It enables them to accommodate molecules like bromoergocryptine (656 Da, CYP3A4),[148] erythromycin (734 Da, CYP3A4), or two molecules of ketoconazole (531 Da) in their active site.[149] Molecules with a correspondingly high molecular weight were clustered in groups VII to IX. The largest accepted compound for CYP154E1 was pergolide (314 Da). This compound, a derivative of ergot alkaloids, acts as dopamine receptor agonist and was until recently, when the risk of serious heart valve damage was reported, an approved drug to treat Parkinson's disease.[150] Several pathways for pergolide detoxification in humans are described. CYP3A4 is a major contributor to pergolide metabolism. Its dealkylation activity against other ergot alkaloid derivatives has also been demonstrated previously.[151,152] A second pathway observed in humans is the sulfoxidation of pergolide.[153] The products formed by CYP154E1 showed the incorporation of one oxygen atom into the alkaloid scaffold but the exact location could not be determined.
Grundmann's ketone (GK) was the only compound from groups VI-IX which was accepted by both P450s. In convergent routes to 25-hydroxyvitamin D_3 analogs, 25-hydroxy-Grundmann's ketone is an important intermediate. It is synthesized by ozonolysis of vitamin D_3 and, the 25-hydroxy functionality can be introduced by oxidation with a ruthenium catalyst.[154] After oxidation and a tedious workup with column chromatography an overall yield of 49 % is reported in the literature.[154] CYP154E1 offers an alternative synthetic route, since it shows absolute selectivity for the 25-hydroxy product without the formation of by-products. CYP154A8 accepts GK with similar selectivity and at the same time it was the largest molecule accepted by CYP154A8 in this screening. To the best of our knowledge these two P450s are the only CYPs to date which selectively convert GK to 25-hydroxy-Grundmann's ketone.
For the cyclic diterpene cembrene (272 Da), type I binding spectra could be recorded with CYP154A8 although no conversion was observed. From a mechanistic point of view it became obvious that binding of a substance to a P450 does not guarantee product formation. Furthermore, this observation emphasized the advantage of a product-based screening, like the cluster screening method used in this study, over mere screening for type I spectra when the reliable identification of substrates is the ultimate aim.

Based on the substrate profile obtained by cluster screening, *n*-alkanes were predicted to be a novel substrate class of CYP154A8. A detailed investigation of the selectivity of CYP154A8 towards this class of substances was performed. *n*-Alkanes with a chain length from C_7 to C_{10} were accepted by CYP154A8 and were converted to 2-alkanols and 3-alkanols. Already in the cluster screening some of these products have been identified as substrates of

CYP154A8. Thus, they were further oxidized what led to diols and subsequently hydroxy ketones. To prevent this overoxidation the substrate was employed as a second organic phase. This reaction design efficiently protected the mono alcohols from a second hydroxylation in the aqueous, enzyme-containing phase by constant, *in situ* liquid-liquid extraction. Additionally, the aqueous phase was saturated with the *n*-alkane and hence the availability of the substrate for the biocatalyst was maximal. As a consequence only 2- and 3-alkanols as well as 2- and 3-ketones were observed. Total regioselectivity for the C2-position was over 90 % for *n*-octane. With higher chain length the C2-selectivity decreased to ~70 % for *n*-decane (C_{10}). The stereoselectivity of the enzyme showed a similar trend. The maximal value of 91 % *ee* was reached for 2-(*S*)-octanol. For the longer-chained 2-(*S*)-decanol the enantiomeric excess decreased to 63 % *ee*. Simultaneously, 3-decanol was formed with 89 % *ee*. Other P450s which selectively hydroxylate *n*-octane to 2-octanol are the CYP102A1 variants 1-12G, 139-3 and F87V/A328F, all with notably lower enantiomeric excess: 46 % *ee* for 2-(*R*)-octanol, 39 % *ee* for 2-(*R*)-octanol, and 58 % *ee* for 2-(*S*)-octanol, respectively.[155–157] Until recently, when 2-(*S*)-octanol production with the $P450_{Pyr}$ variant SM1 (98 % *ee*) was published, CYP154A8 was the P450 reaching the highest enantioselectivity for octane 2-hydroxylation.[158] Nevertheless, it still is the most stereoselective P450 wild type enzyme for *n*-alkane hydroxylation known to date and thus further improvements by protein engineering methods seem possible.

7.2 Activity and catalytic performance

The activity and catalytic performance of a biocatalyst depend on many factors. First of all, an enzyme has to be expressed in a host organism either homo- or heterologously. *E. coli* is often the organism of choice due to its easy handling, low costs for culturing, fast growth, and established genetic manipulation methods. Both CYP154E1 and CYP154A8 could be heterologously overexpressed under standard conditions with up to 1.1 mmol L^{-1} culture (CYP154E1-CHis). A high expression level as one prerequisite for high performance was met by both enzymes.

Secondly, high activity of a P450 is always coupled to the availability of suitable redox partner proteins. In case of CYP154E1 two strategies have been explored. Cloning and expression of autologous redox partners from *T. fusca* failed because the ferredoxins were not expressed in high yields in *E. coli*. However, activity of CYP154E1 and CYP154A8 could be successfully

reconstituted with the heterologous putidaredoxin (Pdx) and its reductase putidaredoxin reductase (PdR) as well as the flavodoxin YkuN and *E. coli* flavodoxin reductase. Supported by YkuN and FdR the P450s reached 2 and 11-fold increased product conversion *in vitro* compared to Pdx/PdR.

Based on these results *E. coli* whole cell biocatalysts with YkuN/FdR and Pdx/PdR were designed for the CYP154E1-catalyzed hydroxylation of Grundmann's Ketone. In contrast to the *in vitro* system, in *E. coli* cells the productivity decreased when YkuN/FdR was used. Product concentrations reached only ~50 % of those achieved with Pdx/PdR. However, the total turnover numbers (TTNs) defined as amount of product formed by each molecule of P450, were still higher with YkuN/FdR than with Pdx/PdR. This phenomenon can be explained by the NADPH dependence of FdR. In *E. coli* cells the concentration of NADPH is reportedly lower than that of NADH.[159] Thereby, the more efficient YkuN/FdR system led to higher TTNs but after consumption of the cofactor this limitation prevented higher product concentrations.

A third, P450 specific factor which influences the performance of a biocatalysts is the coupling of cofactor consumption to product formation. The degree of uncoupling is governed by the rate of electron delivery to the heme iron which in turn is influenced by the redox potentials of the redox partners and by protein-protein interactions. If the second reduction of the heme center cannot take place due to missing reducing equivalents or occurs too slow reactive oxygen species are formed. Thus, redox partner compatibility, which determines the rate of electron delivery, is one of several factors influencing coupling. A further factor which determines the degree of coupling is the substrate itself. It is known that non-natural substrates promote uncoupling due to steric reasons. In this work coupling was determined for the CYP154A8-catalyzed *n*-alkane hydroxylation supported by YkuN/FdR. Maximal coupling was observed for *n*-octane. With 21 % it is far lower than the >90 % observed for CYP102A1 and its natural fatty acid substrates.[35] In contrast to some of its variants, CYP102A1 wild type exhibits low coupling for some non-natural substrates, e.g. only 3.2 % for 9-methylanthracene.[144] Coupling of the mentioned *n*-octane hydroxylating CYP102A1 variants 139-3 and 1-12G was reported to be 22 and 37 %, respectively. With respect to these values CYP154A8 (with 21 % coupling) was on a comparable level. It is noteworthy that this was possible despite the fact that all three components, CYP154A8, YkuN and FdR originate from three different organism.

A measure to assess the performance of a biocatalyst is the TTN. TTNs have been determined *in vitro* for the *n*-alkane oxidation mediated by CYP154A8 in biphasic reaction systems. The maximal TTN-value of 4400 was observed with *n*-nonane and the minimal value of 1700

with *n*-decane. On the one hand, this reflects the different specific activities of CYP154A8 towards the *n*-alkanes. On the other hand, the low solubility of the highly hydrophobic substrates have a major influence on TTNs. Additionally, the TTNs are influenced by the stability against the organic substrates and cosolvents, as discussed below.
In this work, TTNs have also been determined in the development of a CYP154E1-based whole cell biocatalyst for GK hydroxylation. The TTN for GK hydroxylation by CYP154E1/YkuN/FdR was about 30-fold lower compared to those observed for CYP154A8 mediated *n*-alkane hydroxylation. Moreover, different TTNs were observed depending on the used redox partner system. These differences were attributed to a more efficient electron transport by the YkuN/FdR system. It became apparent that besides its advantages the *E. coli* whole cell system also exhibits limitations. Especially, the intracellular NADPH concentration and the capacity of its regeneration by the *E. coli* metabolism had an influence on the performance of the whole-cell biocatalyst. Other limiting factors described in the literature, e.g. membrane permeability of the substrate, expression levels of the different enzymes co-expressed in the cell and the resulting undefined ratio of redox partners to P450 have not been investigated.[160] Prior to the implementation at a larger scale, these factors should be optimized in order to further increase the performance of the biocatalyst.

7.3 Stability

Enzyme stability is the major factor influencing the life time of a biocatalyst. As a consequence, the total stability of a biocatalytic system is dependent on the stability of each of the involved enzymes. Stability against various parameters including concentrations of substrate, product, and additional cosolvents as well as temperature and pressure can be distinguished. In this study all experiments were performed under ambient pressure, therefore only solvent tolerance and thermostability have been investigated.
The origin of CYP154E1 from a moderately thermophilic organism gave rise to investigate its thermostability. The T_{50} of CYP154E1 was 63 °C and thus 13 K higher than that of CYP154A8 which is of mesophilic origin. The value for CYP154E1 is in good accordance with the published apparent melting temperature of CYP154H1 (T_m = 68 °C) from the same host.[105] The thermostability of CYP154E1 lies between those of P450s from mesophilic bacteria like CYP102A1 (heme domain, T_m = 47 °C) and CYP101A1 (T_m = 61 °C) and those from hyperthermophiles like CYP175A1 (T_m = 88 °C) and CYP119A1 (T_m = 91 °C).[161,162] This enhanced stability renders CYP154E1 more attractive for biocatalytic applications where thermostable

enzymes are advantageous.

The solvent tolerance of the biocatalytic system became a major issue during the hydroxylation of *n*-alkanes with CYP154A8 in the described biphasic reaction system. The presence of an organic substrate phase maximized the concentration of the respective *n*-alkane in the aqueous phase. However, organic solvents in general are known to have a destabilizing effect on proteins and thus high concentrations can lead to enzyme inactivation. With CYP154A8/YkuN/FdR/GDH the initial rate *n*-octane hydroxylation was higher than that of *n*-nonane although the TTN achieved with *n*-nonane was ~38 % higher than that with *n*-octane. This behavior was attributed to a lower solubility of *n*-nonane and thereby to a reduced enzyme inactivation. Experiments in which the CYP154A8-catalyzed *n*-octane conversion was supplemented with different components of the reaction mixture after a certain reaction time, supported the hypothesis of enzyme inactivation by the substrate.

7.4 Summary and conclusion

On the way to novel, versatile, P450-based biocatalysts that combine broad substrate spectra with high product selectivity, enhanced stability, and high expression levels the CYP154 family was selected for investigation. Two family members, CYP154E1 from *T. fusca* and CYP154A8 from *N. farcinica* were characterized to determine their biocatalytic potential. Both P450s could be expressed in *E. coli* in high yields and several approaches to reconstitute P450 activity with different redox partners were explored. The strategy to use heterologous redox partners was successful and both monooxygenases displayed highest activity in presence of YkuN and FdR.

A new screening method based on product detection via GC/MS and LC/MS was developed and applied to CYP154E1 and CYP154A8. A library of chemically and structurally related compounds, which were clustered in nine different groups, allowed the identification of 30 and 23 novel substrates of CYP154E1 and CYP154A8, respectively. Moreover, subtle differences in activity and selectivity between the two homologous P450 monooxygenases were resolved using the cluster screening.

The substrate profile was then applied to predict a novel substrate class of CYP154A8, *n*-alkanes. A biphasic reaction system with an organic substrate phase was developed for this reaction with isolated enzymes. It could be shown that CYP154A8 accepts *n*-alkanes from C_7 to C_{10} with a total regioselectivity for the C2-position of up to 90 %. CYP154A8 is strictly *S*-selective and produces 2-(*S*)-octanol with 91 % *ee*. It is therefore so far the only

known wild type P450 monooxygenase that combines high stereoselectivity with high regioselectivity and high TTN leading to product concentrations of up to 3.0 mM after 24 h.
For the hydroxylation of Grundmann's Ketone a whole cell biocatalyst was developed which is able to produce 1.1 mM 25-hydroxy-Grundmann's ketone within 24 h. Investigations of the optimal redox partners *in vivo* demonstrated the superior productivity of the CYP154E1/Pdx/PdR system probably because of the higher availability of NADH over NAD(P)H in *E. coli* cells.
In an effort to elucidate the molecular determinants responsible for the selectivities and the stability of the investigated P450s, CYP154E1 was crystallized. This was a first important step to determine its three-dimensional structure, even though the obtained X-ray diffraction data did not allow the solution of the structure of CYP154E1.

In conclusion, it could be demonstrated that the CYP154 family enzymes are good candidates for the development of versatile P450-based oxidation catalysts. Exploiting their good product selectivity, synthetically useful biocatalysts for the hydroxylation of *n*-alkanes *in vitro* and of Grundmann's Ketone *in vivo* were developed. However, the two investigated P450s did not accept very large substrates like macrolide antibiotics or cembrene. To overcome this limitation and to tune selectivity even further, rational protein engineering can be applied. In an attempt to establish a structural basis for protein engineering of CYP154E1, the enzyme was crystallized. So far the resulting X-ray data did not allow the solution of the crystal structure due to twinning. Further efforts should be undertaken in order to solve the three-dimensional structure of CYP154E1, which will extend our knowledge about structure-function relationships in this P450 and would help to rationalize the observed results. Nevertheless, on the way to find molecular determinants responsible for the catalytic properties of CYP154E1 its crystallization was a step which might finally enable a wider application of CYP154-family enzymes in biocatalysis.

8 References

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

[2] Rosenthaler, L. *Biochem. Z.* **1908**, *14*, 238–253.

[3] Estell, D. A.; Graycar, T. P.; Wells, J. A. *J. Biol. Chem.* **1985**, *260*, 6518–6521.

[4] Hills, G. *Eur. J. Lipid Sci. Technol.* **2003**, *105*, 601–607.

[5] Meyer, H.-P.; Eichhorn, E.; Hanlon, S.; Lütz, S.; Schürmann, M.; Wohlgemuth, R.; Coppolecchia, R. *Catal. Sci. Technol.* **2012**, *3*, 29.

[6] Faber, K.; Franssen, M. C. *Trends Biotechnol.* **1993**, *11*, 461–470.

[7] Faber, K. *Biotransformations in Organic Chemistry*; Springer: Berlin and Heidelberg, 2011.

[8] Weiss, A. In *Modern biooxidation*; Schmid, R.; Urlacher, V. B., Eds.; Wiley-VCH: Weinheim, 2007; pp 193–209.

[9] Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938–945.

[10] Straathof, Adrie J. J.; Panke, S.; Schmid, A. *Curr. Opin. Biotechnol.* **2002**, *13*, 548–556.

[11] Nelson, D. R.; *The Cytochrome P450 homepage*. `http://drnelson.uthsc.edu/CytochromeP450.html`.

[12] Nelson, D. R. *Hum. Genomics* **2009**, *4*, 59–65.

[13] Zanger, U. M.; Schwab, M. *Pharmacol. Ther.* **2013**, *138*, 103–141.

[14] Nebert, D. W.; Russell, D. W. *Lancet* **2002**, *360*, 1155–1162.

[15] Pikuleva, I. A.; Waterman, M. R. *J. Biol. Chem.* **2013**, *288*, 17091–17098.

[16] Morant, M.; Bak, S.; Møller, B. L.; Werck-Reichhart, D. *Curr. Opin. Biotechnol.* **2003**, *14*, 151–162.

[17] Katagiri, M.; Ganguli, B. N.; Gunsalus, I. C. *J. Biol. Chem.* **1968**, *243*, 3543–3546.

[18] Peterson, J. A.; Lu, J. Y.; Geisselsoder, J.; Graham-Lorence, S.; Carmona, C.; Witney, F.; Lorence, M. C. *J. Biol. Chem.* **1992**, *267*, 14193–14203.

[19] Ropp, J. D.; Gunsalus, I. C.; Sligar, S. G. *J. Bacteriol.* **1993**, *175*, 6028–6037.

[20] Boddupalli, S. S.; Estabrook, R. W.; Peterson, J. A. *J. Biol. Chem.* **1990**, *265*, 4233–4239.

[21] Graham, S. E.; Peterson, J. A. *Arch. Biochem. Biophys.* **1999**, *369*, 24–29.

[22] Peterson, J. A.; Graham, S. E. *Structure* **1998**, *6*, 1079–1085.

[23] Kassner, R. J. *J. Am. Chem. Soc.* **1973**, *95*, 2674–2677.

[24] Gotoh, O. *J. Biol. Chem.* **1992**, *267*, 83–90.

[25] Seifert, A.; Pleiss, J. *Proteins* **2009**, *74*, 1028–1035.

[26] Sirim, D.; Widmann, M.; Wagner, F.; Pleiss, J. *BMC Struct. Biol.* **2010**, *10*, 34.

[27] Hannemann, F.; Bichet, A.; Ewen, K. M.; Bernhardt, R. *Biochim. Biophys. Acta* **2007**, *1770*, 330–344.

[28] Agematu, H.; Matsumoto, N.; Fujii, Y.; Kabumoto, H.; Doi, S.; Machida, K.; Ishikawa, J.; Arisawa, A. *Biosci. Biotechnol. Biochem.* **2006**, *70*, 307–311.

[29] Furuya, T.; Nishi, T.; Shibata, D.; Suzuki, H.; Ohta, D.; Kino, K. *Chem. Biol.* **2008**, *15*, 563–572.

[30] Porter, T. D.; Kasper, C. B. *Biochemistry* **1986**, *25*, 1682–1687.

[31] Smith, G. C.; Tew, D. G.; Wolf, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8710–8714.

[32] Serizawa, N.; Matsuoka, T. *Biochim. Biophys. Acta* **1991**, *1084*, 35–40.

[33] Lawson, R. J.; Wachenfeldt, C. v.; Haq, I.; Perkins, J.; Munro, A. W. *Biochemistry* **2004**, *43*, 12390–12409.

[34] Green, A. J.; Munro, A. W.; Cheesman, M. R.; Reid, G. A.; Wachenfeldt, C. v.; Chapman, S. K. *J. Inorg. Biochem.* **2003**, *93*, 92–99.

[35] Noble, M. A.; Miles, C. S.; Chapman, S. K.; Lysek, D. A.; MacKay, A. C.; Reid, G. A.; Hanzlik, R. P.; Munro, A. W. *Biochem. J.* **1999**, *339 (Pt 2)*, 371–379.

[36] Kitazume, T.; Takaya, N.; Nakayama, N.; Shoun, H. *J. Biol. Chem.* **2000**, *275*, 39734–39740.

[37] Rittle, J.; Green, M. T. *Science* **2010**, *330*, 933–937.

[38] Luthra, A.; Denisov, I. G.; Sligar, S. G. *Arch. Biochem. Biophys.* **2011**, *507*, 26–35.

[39] Schenkman, J. B.; Sligar, S. G.; Cinti, D. L. *Pharmacol. Ther.* **1981**, *12*, 43–71.

[40] Sligar, S. G. *Biochemistry* **1976**, *15*, 5399–5406.

[41] Truan, G.; Peterson, J. A. *Arch. Biochem. Biophys.* **1998**, *349*, 53–64.

[42] Schiffler, B.; Kiefer, M.; Wilken, A.; Hannemann, F.; Adolph, H. W.; Bernhardt, R. *J. Biol. Chem.* **2001**, *276*, 36225–36232.

[43] Sharrock, M.; Debrunner, P. G.; Schulz, C.; Lipscomb, J. D.; Marshall, V.; Gunsalus, I. C. *Biochim. Biophys. Acta* **1976**, *420*, 8–26.

[44] Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253–2277.

[45] Auclair, K.; Hu, Z.; Little, D. M.; Ortiz de Montellano, Paul R; Groves, J. T. *J. Am. Chem. Soc.* **2002**, *124*, 6020–6027.

[46] Serron, S. C.; Dwivedi, N.; Backes, W. L. *Toxicol. Appl. Pharmacol.* **2000**, *164*, 305–311.

[47] Cryle, M. J.; Stok, J. E.; Voss, J. J. d. *Aust. J. Chem.* **2003**, *56*, 749–762.

[48] Isin, E. M.; Guengerich, F. P. *Biochim. Biophys. Acta* **2007**, *1770*, 314–329.

[49] Guengerich, F. P.; Munro, A. W. *J. Biol. Chem.* **2013**, *288*, 17065–17073.

[50] Okita, R. T.; Okita, J. R. *Curr. Drug Metab.* **2001**, *2*, 265–281.

[51] Matson, R. S.; Stein, R. A.; Fulco, A. J. *Biochem. Biophys. Res. Commun.* **1980**, *97*, 955–961.

[52] Munro, A. W.; Girvan, H. M.; McLean, K. J. *Nat. Prod. Rep.* **2007**, *24*, 585.

[53] Bertea, C. M.; Freije, J. R.; van der Woude, H; Verstappen, F. W.; Perk, L.; Marquez, V.; Kraker, J.-W. d.; Posthumus, M. A.; Jansen, B. J.; Groot, A. d.; Franssen, M. C.; Bouwmeester, H. J. *Planta Med.* **2005**, *71*, 40–47.

[54] Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, Michelle C. Y.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D. *Nature* **2006**, *440*, 940–943.

[55] Shen, Q.; Chen, Y. F.; Wang, T.; Wu, S. Y.; Lu, X.; Zhang, L.; Zhang, F. Y.; Jiang, W. M.; Wang, G. F.; Tang, K. X. *Genet. Mol. Res.* **2012**, *11*, 3298–3309.

[56] Bischoff, D.; Pelzer, S.; Bister, B.; Nicholson, G. J.; Stockert, S.; Schirle, M.; Wohlleben, W.; Jung, G.; Süssmuth, R. D. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 4688–4691.

[57] Townsend, C. A.; Salituro, G. M. *J. Chem. Soc., Chem. Commun.* **1984**, 1631.

[58] Molnár, I.; et al. *Chem. Biol.* **2000**, *7*, 97–109.

[59] Wenda, S.; Illner, S.; Mell, A.; Kragl, U. *Green Chem.* **2011**, *13*, 3007.

[60] Schulz, S.; Girhard, M.; Urlacher, V. B. *ChemCatChem.* **2012**, *4*, 1889–1895.

[61] Sakaki, T. *Biol. Pharm. Bull.* **2012**, *35*, 844–849.

[62] Shull, G. M.; Kita, D. A.; Davisson, J. W.; *Oxygenation of steroids*; US-Patent: 2658023; 1953; PFIZER & CO Inc. (New York, US).

[63] Suzuki, K.; Sanga, K.-i.; Chikaoka, Y.; Itagaki, E. *Biochim. Biophys. Acta* **1993**, *1203*, 215–223.

[64] Peterson, D. H.; Murray, H. C.; Eppstein, S. H.; Reineke, L. M.; Weintraub, A.; Meister, P. D.; Leigh, H. M. *J. Am. Chem. Soc.* **1952**, *74*, 5933–5936.

[65] Hanson, F.R. and Maxon, W.D.; *Process for the microbiological oxygenation of progesterone*; US-Patent: 3201324; 1965; The Upjohn Company (Kalamazoo, US).

[66] Matsuoka, T.; Miyakoshi, S.; *Cytochrome P-450 enzymes*; US-Patent 5179013; 1993; Sankyo Company, Limited (Tokyo, JP).

[67] Watanabe, I.; Nara, F.; Serizawa, N. *Gene* **1995**, *163*, 81–85.

[68] Serizawa, N.; Nakagawa, K.; Hamano, K.; Tsujita, Y.; Terahara, A.; Kuwano, H. *J. Antibiot.* **1983**, *36*, 604–607.

[69] O'Reilly, E.; Köhler, V.; Flitsch, S. L.; Turner, N. J. *Chem. Commun.* **2011**, *47*, 2490–2501.

[70] Maurer, S. C.; Kühnel, K.; Kaysser, L. A.; Eiben, S.; Schmid, R. D.; Urlacher, V. B. *Adv. Synth. Catal.* **2005**, *347*, 1090–1098.

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

[72] Kühnel, K.; Maurer, S. C.; Galeyeva, Y.; Frey, W.; Laschat, S.; Urlacher, V. B. *Adv. Synth. Catal.* **2007**, *349*, 1451–1461.

[73] van der Donk, Wilfred A; Zhao, H. *Curr. Opin. Biotechnol.* **2003**, *14*, 421–426.

[74] Bernhardt, R. *J. Biotechnol.* **2006**, *124*, 128–145.

[75] Fasan, R.; Chen, M. M.; Crook, N. C.; Arnold, F. H. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 8414–8418.

[76] French, K. J.; Rock, D. A.; Rock, D. A.; Manchester, J. I.; Goldstein, B. M.; Jones, J. P. *Arch. Biochem. Biophys.* **2002**, *398*, 188–197.

[77] Easterbrook, J.; Lu, C.; Sakai, Y.; Li, A. P. *Drug Metab. Dispos.* **2001**, *29*, 141–144.

[78] Koschorreck K., von Bühler C.-J., Schulz S. and Urlacher V.B. In *Protein engineering handbook*; Lutz, S.; Bornscheuer, U. T., Eds.; Wiley-VCH: Weinheim, 2012; Vol. 3, pp 327–362.

[79] Gillam, Elizabeth M. J. *Chem. Res. Toxicol.* **2008**, *21*, 220–231.

[80] Fasan, R. *ACS Catal.* **2012**, *2*, 647–666.

[81] Jung, S. T.; Lauchli, R.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2011**, *22*, 809–817.

[82] Nishida, C. R.; Ortiz de Montellano, Paul R. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 437–445.

[83] Yano, J. K.; Koo, L. S.; Schuller, D. J.; Li, H.; Ortiz de Montellano, P R; Poulos, T. L. *J. Biol. Chem.* **2000**, *275*, 31086–31092.

[84] Salazar, O.; Cirino, P. C.; Arnold, F. H. *ChemBioChem* **2003**, *4*, 891–893.

[85] Li, Y.; Drummond, D. A.; Sawayama, A. M.; Snow, C. D.; Bloom, J. D.; Arnold, F. H. *Nat Biotechnol* **2007**, *25*, 1051–1056.

[86] Eiben, S.; Bartelmäs, H.; Urlacher, V. B. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 1055–1061.

[87] Munro, A. W.; Girvan, H. M.; McLean, K. J. *Biochim. Biophys. Acta* **2007**, *1770*, 345–359.

[88] Nodate, M.; Kubota, M.; Misawa, N. *Appl. Microbiol. Biotechnol.* **2006**, *71*, 455–462.

[89] Fairhead, M.; Giannini, S.; Gillam, Elizabeth M J; Gilardi, G. *J. Biol. Inorg. Chem.* **2005**, *10*, 842–853.

[90] Schmid, R. *Taschenatlas der Biotechnologie und Gentechnik*, 2nd ed.; Wiley-VCH: Weinheim, 2006.

[91] Li, W.; Josephs, J. L.; Skiles, G. L.; Humphreys, W. G. *Drug Metab. Dispos.* **2008**, *36*, 721–730.

[92] Rabe, K. S.; Gandubert, V. J.; Spengler, M.; Erkelenz, M.; Niemeyer, C. M. *Anal Bioanal Chem* **2008**, *392*, 1059–1073.

[93] Simgen, B.; Contzen, J.; Schwarzer, R.; Bernhardt, R.; Jung, C. *Biochem. Biophys. Res. Commun.* **2000**, *269*, 737–742.

[94] Sielaff, B.; Andreesen, J. R. *FEBS J.* **2005**, *272*, 1148–1159.

[95] Schenkman, J. B.; Remmer, H.; Estabrook, R. W. *Mol. Pharmacol.* **1967**, *3*, 113–123.

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

[97] Schwaneberg, U.; Schmidt-Dannert, C.; Schmitt, J.; Schmid, R. D. *Anal. Biochem.* **1999**, *269*, 359–366.

[98] Choi, K.-Y.; Park, H.-Y.; Kim, B.-G. *Enzyme Microb. Technol.* **2010**, *47*, 327–334.

[99] Hopps, H. B. *Aldrichimica Acta* **2000**, *33*, 28–30.

[100] Zhang, K.; El Damaty, S.; Fasan, R. *J. Am. Chem. Soc.* **2011**, *133*, 3242–3245.

[101] Zhang, K.; Shafer, B. M.; Demars, M. D.; Stern, H. A.; Fasan, R. *J. Am. Chem. Soc.* **2012**, *134*, 18695–18704.

[102] Girhard, M.; Tieves, F.; Weber, E.; Smit, M. S.; Urlacher, V. B. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1625–1635.

[103] Lamb, D. C.; Skaug, T.; Song, H.-L.; Jackson, C. J.; Podust, L. M.; Waterman, M. R.; Kell, D. B.; Kelly, D. E.; Kelly, S. L. *J. Biol. Chem.* **2002**, *277*, 24000–24005.

[104] Makino, T.; Katsuyama, Y.; Otomatsu, T.; Misawa, N.; Ohnishi, Y. *Appl. Environ. Microbiol.* **2013**, *80*, 1371–1379.

[105] Schallmey, A.; den Besten, G.; Teune, I. G. P.; Kembaren, R. F.; Janssen, D. B. *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1475–1485.

[106] Podust, L. M.; Bach, H.; Kim, Y.; Lamb, D. C.; Arase, M.; Sherman, D. H.; Kelly, S. L.; Waterman, M. R. *Protein Sci.* **2004**, *13*, 255–268.

[107] Podust, L. M.; Kim, Y.; Arase, M.; Neely, B. A.; Beck, B. J.; Bach, H.; Sherman, D. H.; Lamb, D. C.; Kelly, S. L.; Waterman, M. R. *J. Biol. Chem.* **2003**, *278*, 12214–12221.

[108] Lei, L.; Waterman, M. R.; Fulco, A. J.; Kelly, S. L.; Lamb, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 494–499.

[109] Bruschi, M.; Guerlesquin, F. *FEMS Microbiol. Rev.* **1988**, *4*, 155–175.

[110] Sancho, J. *Cell. Mol. Life Sci.* **2006**, *63*, 855–864.

[111] Bracco, P.; Janssen, D. B.; Schallmey, A. *Microb. Cell Fact.* **2013**, *12*, 95.

[112] Momoi, K.; Hofmann, U.; Schmid, R. D.; Urlacher, V. B. *Biochem. Biophys. Res. Commun.* **2006**, *339*, 331–336.

[113] Sabbadin, F.; Hyde, R.; Robin, A.; Hilgarth, E.-M.; Delenne, M.; Flitsch, S.; Turner, N.; Grogan, G.; Bruce, N. C. *ChemBioChem* **2010**, *11*, 987–994.

[114] Schückel, J.; Rylott, E. L.; Grogan, G.; Bruce, N. C. *ChemBioChem* **2012**, *13*, 2758–2763.

[115] Helvig, C.; Capdevila, J. H. *Biochemistry* **2000**, *39*, 5196–5205.

[116] Scheps, D.; Honda Malca, S.; Richter, S. M.; Marisch, K.; Nestl, B. M.; Hauer, B. *Microb. Biotechnol.* **2013**, *6*, 694–707.

[117] Murakami, H.; Yabusaki, Y.; Sakaki, T.; Shibata, M.; Ohkawa, H. *DNA* **1987**, *6*, 189–197.

[118] Robin, A.; Roberts, G. A.; Kisch, J.; Sabbadin, F.; Grogan, G.; Bruce, N. C.; Turner, N. J.; Flitsch, S. L. *Chem. Commun.* **2009**, 2478–2480.

[119] Seelbach, K.; Riebel, B.; Hummel, W.; Kula, M.-R.; Tishkov, V. I.; Egorov, A. M.; Wandrey, C.; Kragl, U. *Tetrahedron Lett.* **1996**, *37*, 1377–1380.

[120] Tishkov, V. I.; Galkin, A. G.; Fedorchuk, V. V.; Savitsky, P. A.; Rojkova, A. M.; Gieren, H.; Kula, M. R. *Biotechnol. Bioeng.* **1999**, *64*, 187–193.

[121] Priska Le-Huu; Master's thesis; Heinrich-Heine Universität Düsseldorf; Düsseldorf; November 2011.

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

[123] Laemmli, U. K. *Nature* **1970**, *227*, 680–685.

[124] Schägger, H. *Nat. Protoc.* **2006**, *1*, 16–22.

[125] Omura, T.; Sato, R. *J. Biol. Chem.* **1964**, *239*, 2370–2378.

[126] Lykidis, A.; Mavromatis, K.; Ivanova, N.; Anderson, I.; Land, M.; DiBartolo, G.; Martinez, M.; Lapidus, A.; Lucas, S.; Copeland, A.; Richardson, P.; Wilson, D. B.; Kyrpides, N. *J. Bacteriol.* **2007**, *189*, 2477–2486.

[127] Punta, M.; et al. *Nucleic Acids Res.* **2012**, *40*, D290–301.

[128] Ewen, K. M.; Kleser, M.; Bernhardt, R. *Biochim. Biophys. Acta* **2011**, *1814*, 111–125.

[129] McLean, K. J.; Warman, A. J.; Seward, H. E.; Marshall, K. R.; Girvan, H. M.; Cheesman, M. R.; Waterman, M. R.; Munro, A. W. *Biochemistry* **2006**, *45*, 8427–8443.

[130] Ewen, K. M.; Hannemann, F.; Khatri, Y.; Perlova, O.; Kappl, R.; Krug, D.; Huttermann, J.; Muller, R.; Bernhardt, R. *J. Biol. Chem.* **2009**, *284*, 28590–28598.

[131] Thomas C. Pochapsky, Sophia Kazanis, and Marina Dang; Pochapsky, T. C.; Kazanis, S.; Dang, M. *Antioxid. Redox Signal.* **2010**, *13*, 1273–1296.

[132] Kumar, S. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 115–131.

[133] Jancarik, J.; Kim, S. H. *J. Appl. Crystallogr.* **1991**, *24*, 409–411.

[134] Chayen, N. E.; Shaw Stewart, P. D.; Maeder, D. L.; Blow, D. M. *J. Appl. Crystallogr.* **1990**, *23*, 297–302.

[135] Polizzi, K. M.; Bommarius, A. S.; Broering, J. M.; Chaparro-Riggers, J. F. *Curr. Opin. Chem. Biol.* **2007**, *11*, 220–225.

[136] Yeates, T. O. *Meth. Enzymol.* **1997**, *276*, 344–358.

[137] Yeates, T. O.; Fam, B. C. *Structure* **1999**, *7*, R25–R29.

[138] Shiro, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11591–11596.

[139] Borshchevskiy, V.; Efremov, R.; Moiseeva, E.; Büldt, G.; Gordeliy, V. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 26–32.

[140] Redinbo, M. R.; Cascio, D.; Choukair, M. K.; Rice, D.; Merchant, S.; Yeates, T. O. *Biochemistry* **1993**, *32*, 10560–10567.

[141] Zhang, M.; Van Etten, Robert L.; Lawrence, C. M.; Stauffacher, C. V. *J. Mol. Biol.* **1994**, *238*, 281–283.

[142] Efremov, R.; Moukhametzianov, R.; Büldt, G.; Gordeliy, V. *Biophys. J.* **2004**, *87*, 3608–3613.

[143] Backmark, A.; Nyblom, M.; Törnroth-Horsefield, S.; Kosinska-Eriksson, U.; Nordén, K.; Fellert, M.; Kjellbom, P.; Johanson, U.; Hedfalk, K.; Lindkvist-Petersson, K.; Neutze, R.; Horsefield, R. *Acta Crystallogr D Biol Crystallogr* **2008**, *64*, 1183–1186.

[144] Li, Q. S.; Ogawa, J.; Schmid, R. D.; Shimizu, S. *Appl. Environ. Microbiol.* **2001**, *67*, 5735–5739.

[145] Appel, D.; Lutz-Wahl, S.; Fischer, P.; Schwaneberg, U.; Schmid, R. D. *J. Biotechnol.* **2001**, *88*, 167–171.

[146] Budde, M.; Morr, M.; Schmid, R. D.; Urlacher, V. B. *ChemBioChem* **2006**, *7*, 789–794.

[147] Meinhold, P.; Peters, M. W.; Chen, Michael M Y; Takahashi, K.; Arnold, F. H. *ChemBioChem* **2005**, *6*, 1765–1768.

[148] Sevrioukova, I. F.; Poulos, T. L. *J. Biol. Chem.* **2012**, *287*, 3510–3517.

[149] Farinas, E. T.; Alcalde, M.; Arnold, F. H. *Tetrahedron* **2004**, *60*, 525–528.

[150] Zanettini, R.; Antonini, A.; Gatto, G.; Gentile, R.; Tesei, S.; Pezzoli, G. *N. Engl. J. Med.* **2007**, *356*, 39–46.

[151] Rauschenbach, R.; Gieschen, H.; Husemann, M.; Salomon, B.; Hildebrand, M. *Eur. J. Pharmacol.* **1995**, *293*, 183–190.

[152] Ball, S. E.; Maurer, G.; Zollinger, M.; Ladona, M.; Vickers, A. E. *Drug Metab. Dispos.* **1992**, *20*, 56–63.

[153] Rubin, A.; Lemberger, L.; Dhahir, P. *Clin. Pharmacol. Ther.* **1981**, *30*, 258–265.

[154] Kiegiel, J.; Wovkulich, P. M.; Uskoković, M. R. *Tetrahedron Lett.* **1991**, *32*, 6057–6060.

[155] Peters, M. W.; Meinhold, P.; Glieder, A.; Arnold, F. H. *J. Am. Chem. Soc.* **2003**, *125*, 13442–13450.

[156] Glieder, A.; Farinas, E. T.; Arnold, F. H. *Nat. Biotech.* **2002**, *20*, 1135–1139.

[157] Weber, E.; Seifert, A.; Antonovici, M.; Geinitz, C.; Pleiss, J.; Urlacher, V. B. *Chem. Commun.* **2010**, *47*, 944.

[158] Yang, Y.; Liu, J.; Li, Z. *Angew. Chem.* **2014**, *126*, 3184–3188.

[159] Chemler, J. A.; Fowler, Z. L.; McHugh, K. P.; Koffas, M. A. *Metab. Eng.* **2010**, *12*, 96–104.

[160] Schrewe, M.; Julsing, M. K.; Bühler, B.; Schmid, A. *Chem. Soc. Rev.* **2013**, *42*, 6346.

[161] Yano, J. K.; Blasco, F.; Li, H.; Schmid, R. D.; Henne, A.; Poulos, T. L. *J. Biol. Chem.* **2003**, *278*, 608–616.

[162] McLean, M. A.; Maves, S. A.; Weiss, K. E.; Krepich, S.; Sligar, S. G. *Biochem. Biophys. Res. Commun.* **1998**, *252*, 166–172.

9 Appendix

9.1 Abbreviations

5-ALA	5-aminolevulinic acid
A. annua	*Artemisia annua*
AdR	adrenodoxin reductase
Adx	adrenodoxin
Amp	ampicillin
B. licheniformis	*Bacillus licheniformis*
Bicine	2-(bis(2-hydroxyethyl)amino)acetic acid
bp	base pairs
CO	carbon monoxide
CPR	cytochrome P450 reductase
CYP	cytochrome P450
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
ee	enantiomeric excess
EC	enzyme class
EPR	electron paramagnetic resonance
ER	endoplasmic reticulum
ESI-TOF-MS	electro spray ionization time of flight mass spectrometry
fer	*B. subtilis* ferredoxin
FAD	flavin adenine dinucleotide
FDA	U.S. Food and Drug Administration

FDH ..... formate dehydrogenase
FdR ..... *E. coli* flavodoxin reductase
FldR ..... flavodoxin reductase
Fldx ..... flavodoxin
FMN ..... flavin mononucleotide
GC ..... gas chromatography
GDH ..... glucose dehydrogenase
GK ..... Grundmann's ketone
HEPES ..... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR-MS ..... high resolution mass spectrometry
HS ..... high-spin
IMAC ..... immobilized metal ion affinity chromatography
IPTG isopropyl-β-D-1-thiogalactopyranoside
K_D ..... binding constant
KIE ..... kinetic isotope effect
LB ..... Luria-Bertani
LB_{Amp} ..... Luria-Bertani medium with ampicillin
LC ..... liquid chromatography
LS ..... low-spin
M_w ..... molecular weight
MALS ..... multi angle light scattering
MCS ..... multiple cloning site
MPD 2-methyl-2,4-pentanediol
MS ..... mass spectrometry
N. farcinica ..... *Nocardia farcinica*
NAD^+ ..... nicotinamide adenine dinucleotide, oxidized form
NADH ..... nicotinamide adenine dinucleotide, reduced form
$NADP^+$ ..... nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH ..... nicotinamide adenine dinucleotide phosphate, reduced form
OD ..... optical density
ORF ..... open reading frame

P450 ..... Cytochrome P450 monooxygenase
$P450_{BioI}$ ..... CYP107H1, from *Bacillus subtilis*
$P450_{BM-3}$ ..... CYP102A1, from *Bacillus megaterium*
$P450_{cam}$ ..... CYP101A1 from *Pseudomonas putida*
$P450_{cin}$ ..... CYP176A1, from *Citrobacter braakii*
$P450_{foxy}$ ..... CYP505, from *Fusarium oxysporum*
$P450_{lin}$ ..... CYP111A1, from *Pseudomonas incognita*
$P450_{lun}$ ..... CYP107, from *Curvularia lunata*
$P450_{MycG}$ ..... CYP107E1, from *Micromonospora griseorubida*
$P450_{OxyA}$ ..... CYP165A1, from *Amycolatopsis orientalis*
$P450_{OxyB}$ ..... CYP165B1, from *Amycolatopsis orientalis*
$P450_{OxyC}$ ..... CYP165C1, from *Amycolatopsis orientalis*
$P450_{pyr}$ ..... CYP153A7, from *Sphingomonas sp.*
$P450_{sca}$ ..... CYP105A3, from *Streptomyces carbophilus*
$P450_{terp}$ ..... CYP108A1, from *Pseudomonas* sp.
PCR ..... polymerase chain reaction
PDB ..... Protein Database, http://pdb.org/
PdR ..... putidaredoxin reductase
Pdx ..... putidaredoxin
PFAM ..... protein family database, http://pfam.xfam.org/
PFOR ..... phthalate-family oxygenase reductase
PMSF ..... phenylmethylsulfonyl fluorid
RIS ..... reductase interaction site
ROX ..... reactive oxygen species
rpm ..... rotations per minute
S. carbophilus ... *Streptomyces carbophilus*
SDS-PAGE ..... sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SEC ..... size exclusion chromatography
SFR ..... structure-function relationship
SRS ..... substrate recognition site
T_{50} ..... temperature which reduces enzyme activity by half

T_m ..... melting temperature

TB ..... terrific broth

Tricine ..... *N*-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine

Tris ..... tris(hydroxymethyl)aminomethane

TTN ..... total turnover number, defined as nmol product formed per nmol of P450 employed

U ..... enzyme unit

UV/VIS ..... ultraviolet-visible

YkuN ..... *B. subtilis* flavodoxin N

YkuP ..... *B. subtilis* flavodoxin P

9.2 Constructed plasmids

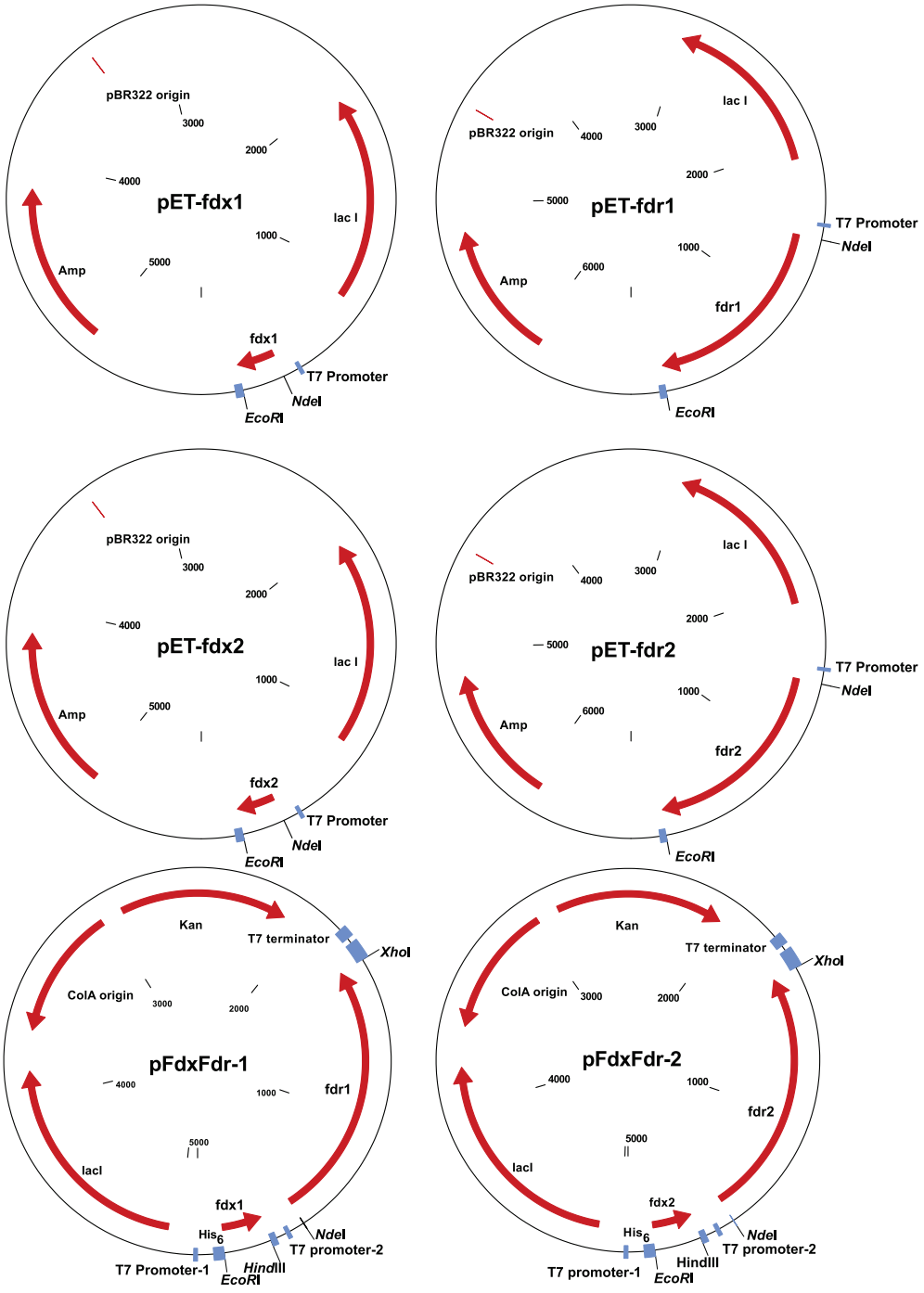


Figure 24: Plasmids pET-fdx1, pET-fdx2, pET-fdr1 and pET-fdr2 are based on pET-22b(+). The pCOLADuet-1-based plasmids pFdxFdr-1 and pFdxFdr-2 allow coexpression of the corresponding ferredoxin/ferredoxin reductase genes. **Kan**: kanamycin resistance gene, **Amp**: ampicillin resistance gene, **lacI**: *lacI* repressor gene, **His_6**: sequence coding for a hexa histidine tag.

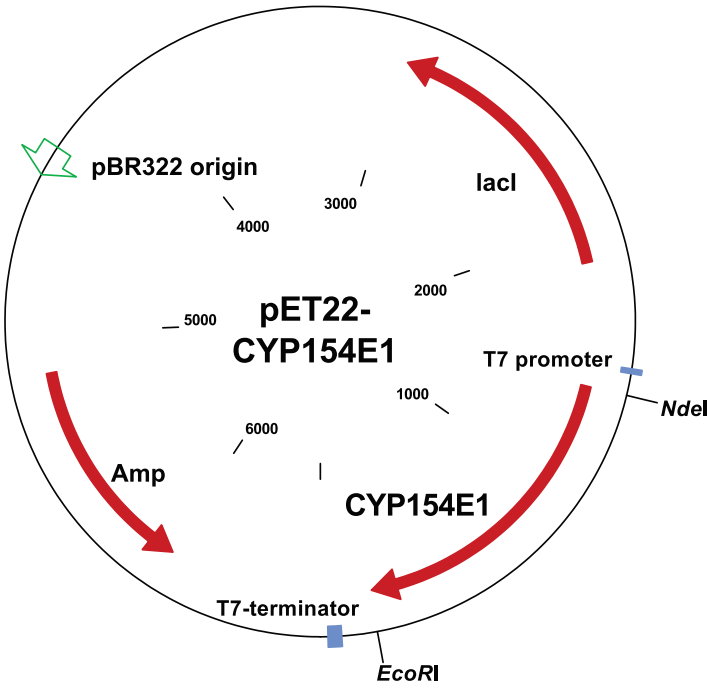


Figure 25: Plasmid map of the pET-22b(+)-based expression vector for CYP154E1. **Amp**: ampicillin resistance gene, **lacI**: *lacI:* repressor gene, **pBR322 origin**: origin of replication.

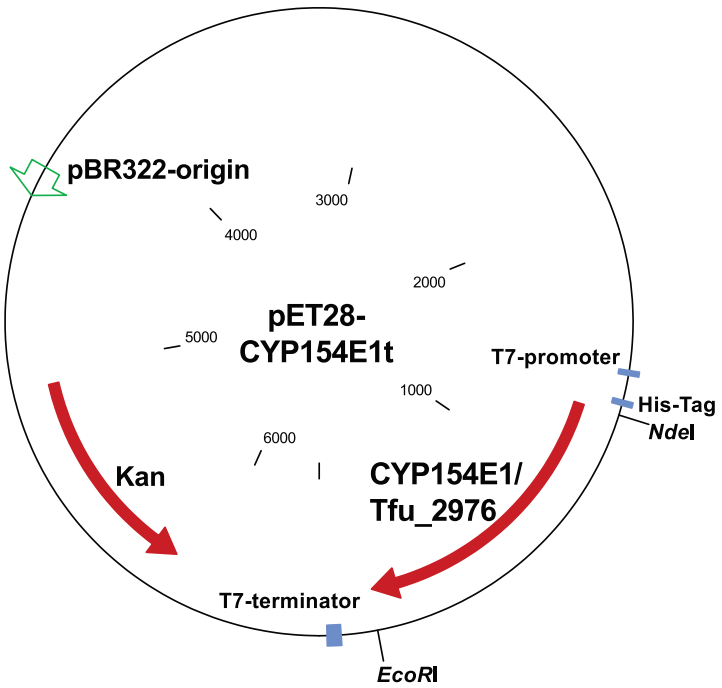


Figure 26: Plasmid map of pET28-CYP154E1t (Tfu_2976). **Kan**: kanamycin resistance gene, **pBR322-origin**: origin of replication.

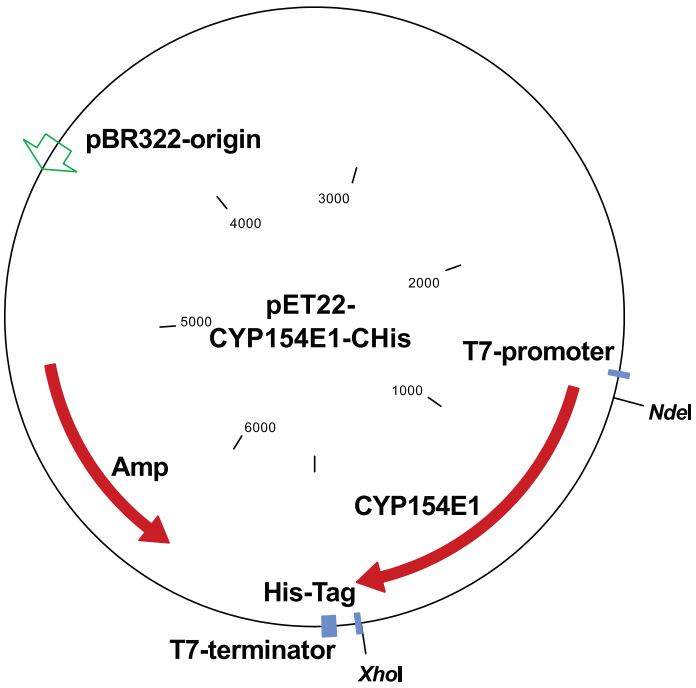


Figure 27: Plasmid map of pET22-CYP154E1-CHis. **Amp**: ampicillin resistance gene, **pBR322-origin**: origin of replication.