

**Identification and characterization of new proteins
from *Bacillus megaterium* for biotechnological and
pharmaceutical applications**

Kumulative Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften (Dr. rer. nat.)

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

Von

Mohammed Milhim

Saarbrücken

2017



Tag des Kolloquiums : **28.06.2017**

Dekan : **Prof. Dr. Guido Kickelbick**

Berichterstatter : **Prof. Dr. Rita Bernhardt**

: **Prof. Dr. Gert-Wieland Kohring**

Vorsitz : **Prof. Dr. Uli Müller**

Akad. Mitarbeiter : **Dr.-Ing. Michael Kohlstedt**

"Imagination is more important than knowledge.

Knowledge is limited, whereas imagination

embraces the entire world, stimulating progress,

giving birth to evolution"

Albert Einstein (14 March 1879 – 18 April 1955)



ACKNOWLEDGMENT

I wish to thank all those who helped me. Without them, I could not have completed this project.

I would like to express my thanks to **Prof. Dr. Rita Bernhardt** for giving me the opportunity to make my thesis in her group. I especially would like to thank her for the continuous support and the intensive help in solving the scientific problems also for her patience, motivation, and immense knowledge. What I have learned from her will benefit me for life.

I would like to express my thanks to **Prof. Dr. Gert-Wieland Kohring** for kindly taking over the co-correction and being the second supervisor.

Sincere thanks are also to **Dr. Frank Hannemann** for all support, understanding and excellent advices since the very beginning of my PhD study, which helped to overcome all difficulties during my work.

Special thank for **Adrian Gerber, Philip Hartz, Ammar Abdulmughni, and Dr. Jens Neunzig** for the stimulating scientific discussions, teaching me many Laboratory techniques and for their kind help to finish some experiments and also for the very interesting talks in all fields.

I would also like to express my gratitude towards **Maximilian Ehrhardt** and **Ömer Kurt** for the very interesting dialogues in the different non-P450 fields.

I would like to thank all my colleagues in the group for the friendly and supportive atmospheres, **Dr. Azzam Mosa, Tanja Sagadin, Dr. Simone Brixius-Anderko, Benjamin Stenger, Lisa König, Natalia Putkaradze, Dr. Fredy Kern, Dr. Martin Litzenburger, Dr. Yogan Khatri, Dr. Lina Schiffer, and Dr. Alexander Schifrin** which made the work in the lab so enjoyable. I greatly look forward to having all of them as colleagues in the years ahead.

I also like to say thank for **Birgit Heider-Lips** and **Antje Eiden-Plach** for sharing me their excellent advices and experiences as well as for their efforts in facilitating the work in the lab. For her help with all administrative burdens, I would also like to express my thanks to **Gabi Schon**.

Special thanks to my friend **Musaab Al-Tuwaijari** for providing me with unfailing support and continuous encouragement.

Last but not the least; I would like to thank my family: my parents (Majid and Lamia), my brothers (Abood and Yahia) and sisters (Hiba and Saja) for supporting me spiritually throughout writing this thesis.

SCIENTIFIC CONTRIBUTIONS

This work is based on 4 original publications, which are reproduced in chapter 2 with permission of Journal of Biotechnology (2.1 and 2.2 Milhim et al., 2016a, 2016b), Metabolic Engineering (2.2 Gerber et al., 2016), and Steroid Biochemistry and Molecular Biology (2.4 Neunzig et al., 2017).

2.1: **Milhim, M.**, Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016a. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. J. Biotechnol. 231, 83–94.

The author planned and carried out all the presented experiments and wrote the manuscript.

2.2: **Milhim, M.**, Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016b. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. J. Biotechnol. 240, 68–75.

The author planned and carried out all the presented experiments and wrote the manuscript.

2.3: Gerber, A., **Milhim, M.**, Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. Metab. Eng. 36, 19–27.

The author participated in the purification of steroid products and contributed in the cloning experiments, in addition to the participation in the interpretation and discussion of the results.

2.4: Neunzig, J., **Milhim, M.**, Schiffer, L., Khatri, Y., Zapp, J., Sánchez-Guijo, A., Hartmann, M.F., Wudy, S.A., Bernhardt, R., 2017. The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1. *J. Steroid Biochem. Mol. Biol.* 167, 182–191.

The author established the CYP21A2-dependent whole-cell system. In addition to conducting *in vitro* experiments with CYP21A2.

CONTENTS

Acknowledgment	I
Scientific contributions	III
Contents	V
Abstract	1
Zusammenfassung	2
1 Introduction	3
1.1 Biotechnological biotransformation	3
1.2 Cytochromes P450	4
1.2.1 History and Nomenclature	4
1.2.2 Structure	6
1.2.3 Classification, reactions and catalytic cycle	8
1.2.4 Functions and examples	10
1.2.5 Application of P450 systems in biocatalysis	14
1.3 <i>Bacillus megaterium</i>	16
1.3.1 General	16
1.3.2 <i>Bacillus megaterium</i> P450 systems	16
1.4 Aim of the project	18
2 Publication of the results	20
2.1 Milhim et al., 2016a	21
A Novel NADPH-dependent flavoprotein reductase from <i>Bacillus megaterium</i> acts as an efficient cytochrome P450 reductase	
2.2 Milhim et al., 2016b	37
Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from <i>Bacillus megaterium</i> with a catalytic activity towards mevastatin	
2.3 Gerber et al., 2016	51
Genetic engineering of <i>Bacillus megaterium</i> for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one	
2.4 Neunzig et al., 2017	71
The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1	

3	Discussion.....	89
3.1	Characterization of a novel NADPH-dependent diflavin reductase.....	89
3.2	CYP107DY1 a new plasmid-encoded cytochrome P450 from <i>Bacillus megaterium</i>	94
3.3	<i>In vivo</i> whole cell pravastatin production.....	95
3.4	A new short-chain dehydrogenase/reductase from <i>Bacillus megaterium</i> with 17 β -HSD activity	96
3.5	BMD_2094 can convert the sesquiterpene nootkatol to (+)-nootkatone.....	98
4	Outlook.....	103
5	Appendix	107
5.1	Supplemental methods	107
5.1.1	Cloning of the gene encoding BMD_2094.....	107
5.1.2	Cultivation conditions of the <i>B. megaterium</i> and <i>in vivo</i> substrate conversion	107
5.1.3	Metabolites extraction and GC-MS analysis.....	108
5.2	Sequences.....	108
5.2.1	BmCPR (BMD_3122).....	108
5.2.2	CYP107DY1 (BMQ_pBM50008)	109
5.2.3	BMD_2094.....	109
5.3	Abbreviations	110
6	References	111

ABSTRACT

The oxyfunctionalization of substrates (such as steroids and drugs) in a stereo- and regioselective manner is of great interest for the biotechnological and pharmaceutical industry. This study presents the identification and characterization of three enzymes from the genome of the Gram-positive bacterium *Bacillus megaterium*. The first enzyme is a novel NADPH-dependent diflavin reductase (BmCPR for *Bacillus megaterium* cytochrome P450 reductase), which was found to be able to transfer electrons efficiently to the microsomal P450s as well as to different ferredoxins. The second enzyme is a cytochrome P450 (CYP107DY1), which is the first plasmid-encoded cytochrome P450 identified so far in the *Bacillus* species. The recombinant CYP107DY1 exhibiting a characteristic P450 absolute and reduced CO-bound difference spectrum was found to possess activity towards mevastatin to produce pravastatin. In addition, this work presents also the identification of a novel short-chain dehydrogenase/reductase (SDR) with 17 β -hydroxy steroid dehydrogenase (17 β -HSD) activity. The new dehydrogenase was found to possess activity towards the sesquiterpene (trans)-nootkatol converting it to the industrial valuable (+)-nootkatone. Moreover and in view of the biotechnological applications, the new enzymes could be utilized in the establishment of whole-cell based biocatalyst systems for the biotransformation of biotechnological relevant substances.

ZUSAMMENFASSUNG

Die stereo- und regioselektive Oxyfunktionalisierung von Substraten, wie etwa Steroiden und Medikamenten, ist von großem Interesse für die biotechnologische und pharmazeutische Industrie. In der vorliegenden Arbeit wurden drei verschiedene Enzyme des Gram-positiven Bakteriums *Bacillus megaterium* identifiziert und charakterisiert. Bei dem ersten Enzym handelt es sich um eine neuartige NADPH abhängige Diflavin Reduktase (BmCPR für *Bacillus megaterium* Cytochrom P450 Reduktase), die in der Lage ist, Elektronen effektiv auf mikrosomale Cytochrome P450, ebenso wie auf verschiedene Ferredoxine zu übertragen. Bei dem zweiten Enzym handelt es sich um ein Cytochrom P450 (CYP107DY1), welches zur Biotransformation von Mevastatin zu Pravastatin verwendet werden kann. Das rekombinant exprimierte CYP107DY1 zeigt das für Cytochrome P450 charakteristische Absorptionsspektrum im reduzierten und CO gebundenen Zustand. Soweit bekannt, ist dieses Enzym das erste Plasmid kodierte Cytochrom P450 der Gattung *Bacillus*. Weiterhin umfasst die vorliegende Arbeit die Identifizierung eines neuen Vertreters aus der Superfamilie der Short-chain Dehydrogenasen/Reduktasen (SDR) mit Aktivität von 17 β -Hydroxysteroid Dehydrogenasen (17 β -HSD). Die Aktivität dieser Dehydrogenase erlaubt zudem die Produktion des industriell bedeutenden Sesquiterpens (+)-Nootkaton aus (trans)-Nootkatol. Hinsichtlich einer biotechnologischen Anwendung besitzen alle in dieser Arbeit charakterisierten Enzyme das Potenzial für die Etablierung von Ganzzell basierten Biokatalysatoren für die Biotransformation biotechnologisch relevanter Substanzen.

1 INTRODUCTION

1.1 BIOTECHNOLOGICAL BIOTRANSFORMATION

Biotransformations (bio-conversions or microbial transformations) broadly refer to the processes, in which microorganisms or enzymes convert organic compounds into structurally related products. The amazing world of biotechnology was started by humans more than thousands of years ago by producing alcohol and vinegar by fermentation without even knowing about the presence of microorganisms (Solieri and Giudici, 2009). Later, living cells, such as yeast, bacteria, filamentous fungi and plants were deliberately used to improve the stability and taste of food (Frazzetto, 2003). A wide variety of biological catalysts can be used for biotransformation reactions. These include growing cells, resting cells, killed cells, immobilized cells, cell-free extracts, enzymes and immobilized enzymes (Zhang et al., 2016). The significance of bioconversion reactions becomes obvious when the production of a particular compound is either difficult or costly by chemical methods. Further, biotransformation have the advantage of converting substrates under relatively mild reaction conditions (neutral pH, room temperature, less activation energy) in comparison to conventional harsh chemical reactions (Frazzetto, 2003). Enzymes are stereo- and regionselective, and the recreation of this selectivity is often a major challenge for synthetic and pharmaceutical chemistry using chemocatalysis (Luo et al., 2015). Therefore, enzymes are commercially important and often provide a more environmental friendly way for the synthesis of chemical compounds. Many types of chemical reactions occur in biotransformation. These include oxidation, reduction, hydrolysis, condensation, isomerization, formation of new C—C bonds, synthesis of chiral compounds and reversal of hydrolytic reactions (Pervaiz et al., 2013).

One particularly valuable class of enzymes are the oxygenase, which catalyze the introduction of oxygen into non-functionalized carbon skeletons (Danielson, 2003). Oxygenations using conventional abiotic chemistry are energy consuming, often require either expensive catalysts and/or toxic reagents and are often neither regio- nor stereoselective. Hence there are substantial research efforts into the discovery and application of enzymes that catalyze oxygenation reactions, which have included

peroxidases, flavin dependent monooxygenases, and heme containing enzymes known as cytochromes P450 (Leak et al., 2009).

1.2 CYTOCHROMES P450

1.2.1 HISTORY AND NOMENCLATURE

Cytochromes P450 (P450, also referred as CYP) play a very central role in the drug metabolism, steroidogenesis and the metabolism of fatty acids (Danielson, 2003). Although the ability of the mammalian tissue to oxidize xenobiotics was widely recognized as early as in the 1950s (Axelrod, 1955; Brodie et al., 1955), the specific enzymes responsible for this catalysis were unknown. The first indication came in the year 1958 from the spectroscopic studies of the microsomal heme proteins from rat and pig livers (Garfinkel, 1958; Klingenberg, 1958), which showed the existence of a membrane-bound reduced red pigment exhibiting a Soret peak at 450 nm for the carbon monoxide-complexed form (**Figure 1.1**).

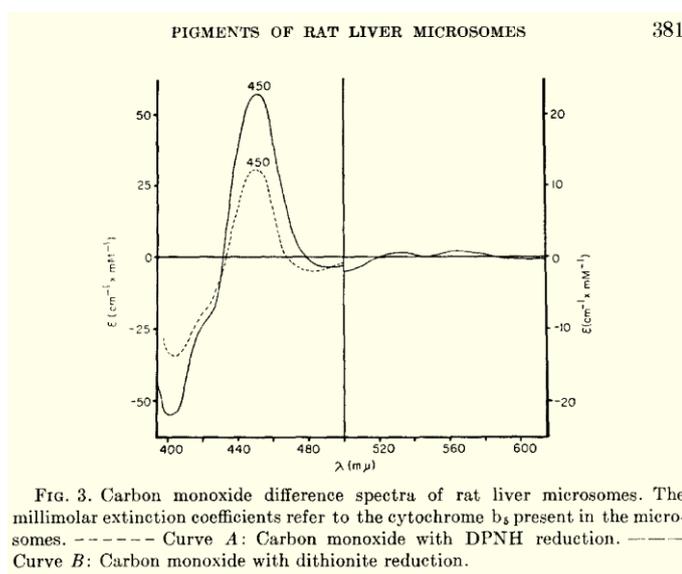


Figure 1.1. Original figure of the carbon monoxide difference spectra present in rat liver microsomes (Klingenberg, 1958).

These pigments were then identified as P450 hemoprotein (Omura and Sato, 1962). In 1963, Estabrook, Cooper and Rosenthal discovered their enzymatic activity in the metabolism of steroids (Estabrook et al., 1963). Further studies by Omura in 1964 and 1966 showed the localization of the P450s in the mitochondria of the adrenal cortex in addition to the discovery of the electron transfer

proteins adrenodoxin reductase and adrenodoxin (Omura et al., 1966; Omura and Sato, 1964). The first bacterial P450 was found in *Rhizobium* bacteroids, which was in contrast to its mammalian counterpart, a soluble P450 (Appleby, 1967). The cytochrome P450 reductase (CPR or POR) was first described in 1950 as cytochrome c reductase (Horecker, 1950). About 20 years later, Lu and Coon provided the evidence for the requirement of the POR in the microsomal P450s reactions (Lu et al., 1969). **Table 1** briefly lists some important milestones in the history of P450s discovery.

Table 1. Some selected milestones of cytochrome P450.

Year	description	Reference
1950	POR was described as cytochrome c reductase	(Horecker, 1950)
1958	Discovery of red pigments out of rat and pig livers	(Garfinkel, 1958; Klingenberg, 1958)
1962	The name cytochrome P450 was introduced by Omura and Sato	(Omura and Sato, 1962)
1963	The discovery of the involvement of cytochrome P450s in steroids metabolism	(Estabrook et al., 1963)
1964	Publication of the detailed spectroscopic characteristics as well as the location of the cytochrome P450s in mitochondria of adrenal cortex	(Omura and Sato, 1964)
1966	The discovery of the electron transfer proteins adrenodoxin reductase (AdR) and adrenodoxin (Adx) as well as the reconstitution of a NADPH-dependent steroid 11 β -hydroxylase activity of a cytochrome P450	(Omura et al., 1966)
1967	The discovery of the first bacterial soluble cytochrome P450 from <i>Rhizobium</i>	(Appleby, 1967)
1969	The requirement of POR in the cytochrome P450 reactions	(Lu et al., 1969)
1974	Identification of the fusion cytochrome P450 BM3 from <i>Bacillus megaterium</i>	(Miura and Fulco, 1974)
1987	The introduction of a systematic nomenclature for cytochrome P450s classification	(Nebert et al., 1987)
1987	The first high-resolution crystal structure of cytochrome P450cam	(Poulos et al., 1987)
1996	Description of more than 20 different reactions of the cytochrome P450s	(Sono et al., 1996)

Since that time, P450s have been discovered in all domains of life and over 35,000 different P450s (<http://drnelson.uthsc.edu/CytochromeP450.html>) were already identified (Nelson, 2009). They were found in mammals, bacteria (except *Escherichia coli*), archaea, fungi, plants and viruses (**Figure 1.2**) (Lamb and Waterman, 2013).

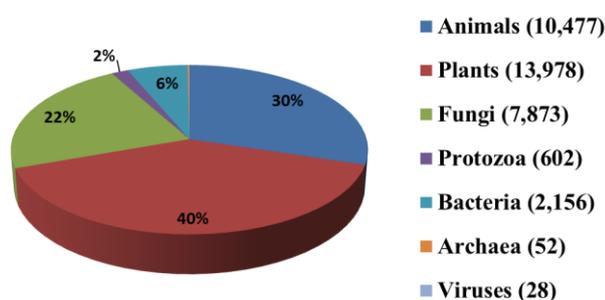


Figure 1.2. Schematic distribution of cytochromes P450 among the different kingdoms [Stand: April, 2016].

For the classification of P450s, a systematic nomenclature was suggested in 1987 (Nebert et al., 1987). As depicted in **Figure 1.3** the abbreviation CYP stands for cytochrome P450 while the first number gives the membership of the family (over 40% identity of the amino acid sequence). The first letter indicates the subfamily (over 55% identity of the amino acid sequence) and the last number represents the specific isoenzyme. The following figure shows CYP106A1 from *Bacillus megaterium* DSM319 as an example of P450s nomenclature.

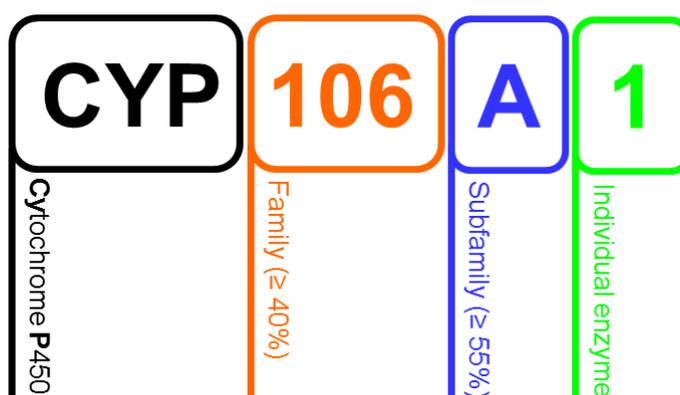


Figure 1.3. Nomenclature of cytochromes P450 as proposed by Nebert et al.(1987), exemplified by CYP106A1 from *B. megaterium* DSM319.

1.2.2 STRUCTURE

The sequence identity among P450 enzyme families is considered extremely low and it might not exceed 20%. However, the three dimensional organization of the resolved P450 crystal structures shows a remarkable conservation in the secondary structural folding and topology, consisting of two defined regions (**Figure 1.4**), the α -helices-rich region (or α -domain) comprising 12 main α -helices designated A-L and the β -sheets-rich region (or β -domain) (Werck-Reichhart and Feyereisen, 2000).

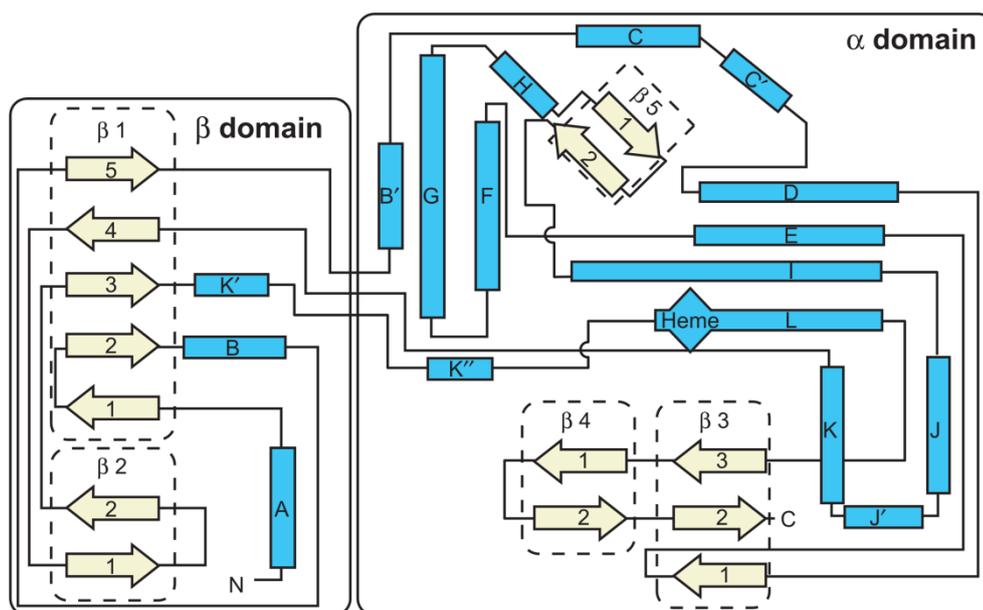


Figure 1.4. Topographic map showing the secondary structural elements arrangement of the P450s. The α -helices are represented in blue boxes while β -strands are in cream arrows. [Figure was adapted from Werck-Reichhart and Feyereisen, 2000]

The highest structural conservation is found in the core of the protein around the heme center and reflects a common mechanism of electron and proton transfer and oxygen activation (Sirim et al., 2010). These regions comprise: first, the heme-binding motif (F-x-x-G-x-x-C-x-G, x=any amino acid). The prosthetic heme group is embedded in the active site and surrounded by the I-helix from the proximal side and the L-helix from the distal side. The conserved cysteine residue is located in the loop region preceding the L-helix. The absolutely conserved cysteine coordinates as fifth ligand with the heme iron. The second highly conserved motif is (A/G-G-x-E/D-T-T/S), which is found in the I-helix and regarded as an oxygen-binding motif with the conserved glycine pointing at the center of the heme and the conserved threonine pointing to the oxygen-binding site. In addition, the presence of the (E-x-x-R) motif in the K-helix is responsible for the formation of a salt bridge, which is crucial for the heme center stabilization (Danielson, 2003; Lamb and Waterman, 2013).

The most variable regions are associated with the N-terminal and the substrate recognition sites; the later regions are located near the substrate access channel and catalytic site and are designated as substrate recognition sites (SRSs) (Gotoh, 1992). Six SRSs have been identified (**Figure 1.5**): SRS1 lies between helices B and C, SRS2 and SRS3 lie between the F and G helices, which form the upper

side of the active site, SRS4 includes the central part of the I-helix, while SRS5 and SRS6 are located at the N-terminus of β -sheet 1-4 and the turn at the end of β -sheet 4-1, respectively (Nair et al., 2016).

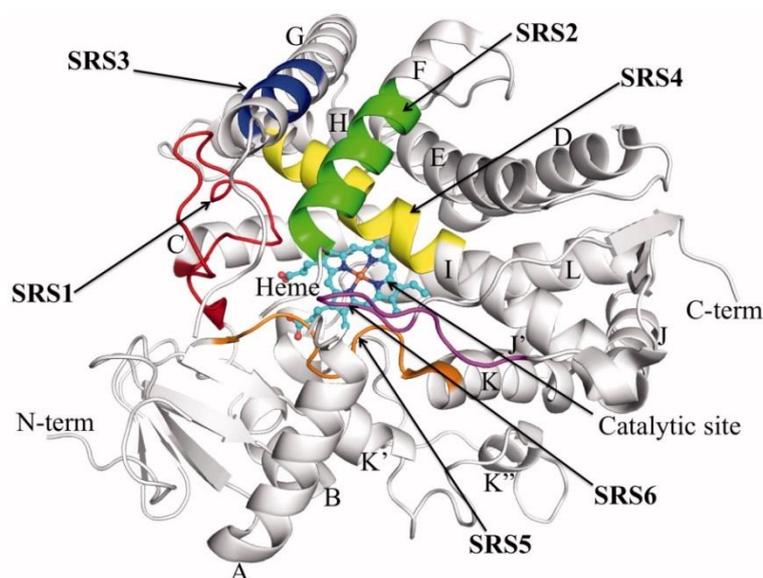


Figure 1.5. Substrate recognition sites (SRSs) in P450s exemplified with CYP2C9.

[Figure was adapted from Nair et al., 2016]

1.2.3 CLASSIFICATION, REACTIONS AND CATALYTIC CYCLE

P450s are heme-thiolate monooxygenases, which catalyze diverse reactions in a regio- and stereoselective manner and are involved in - for example but not limited to - the biosynthesis of hormones, signal molecules, defense-related chemicals and secondary metabolites, in addition to their central role in the metabolism of endogenous (steroids and fatty acids) and exogenous (drugs and toxins) substances (Bernhardt, 2006).

P450s require auxiliary redox partners for the activation of the molecular oxygen. These redox partners transfer two electrons in two single steps from the cofactor NAD(P)H to the heme of the P450. Depending on the components of the electron transfer chains, P450s are classified into different classes, of which the most researched ones are: class I, which comprises bacterial and mitochondrial P450s, and class II, which contains the microsomal P450s. Mitochondrial P450s are associated with the inner mitochondrial membrane and rely on a likewise membrane associated NADPH-oxidizing ferredoxin reductase for electron supply. The electrons are shuttled to the P450 via a soluble

ferredoxin. In bacterial class I systems, all three components are soluble. Microsomal P450s are embedded into the membrane of the endoplasmic reticulum and receive electrons through a cytochrome P450 reductase (CPR). This diflavin reductase from the FNR-like superfamily contains a flavin adenine dinucleotide (FAD) and a flavin mononucleotide (FMN) as prosthetic groups, allowing both the oxidation of NADPH and the shuttling of electrons to the P450 (Hannemann et al., 2007).

The function of P450s is to activate molecular oxygen to yield a reactive species that can attack relatively inert chemical sites in order to introduce hydroxyl group (**Figure 1.6**) into unreactive structures such as hydrocarbon chains and aromatic rings (Makris et al., 2002). P450s catalyze a very wide range of reactions including, epoxidations, deaminations, desulfurations, dehalogenations, N-, S-, and O-dealkylations, N-oxidations, peroxidations and sulfoxidations with the hydroxylations considered as the most important reactions (Bernhardt and Urlacher, 2014).

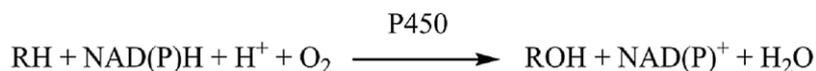


Figure 1.6. General reaction catalyzed by P450s.

During the catalytic cycle (**Figure 1.7**) the P450 enzyme binds the substrate and oxygen molecules, accepts two electrons and two protons, splits the molecular oxygen, inserts a single atom of oxygen into the cognate substrate, and finally releases a molecule of water and the oxygenated substrate as reaction products. In more detail, the binding of a substrate to a P450 causes the displacement of the water molecule from the active center and lowers the redox potential by approximately 100 mV, which makes the transfer of an electron favorable from its redox partner (Denisov et al., 2005). This is accompanied by a change in the spin state of the heme iron at the active site from low to high spin. The next step in the cycle is the reduction of the Fe^{+3} ion by an electron transferred from NAD(P)H via an electron transfer chain. An O_2 molecule binds rapidly to the Fe^{+2} ion forming the hexacoordinate low spin ferrous- O_2 adduct ($\text{Fe}^{+2}\text{-O}_2$). There is evidence suggesting that this complex then undergoes a slow conversion to a more stable complex ($\text{Fe}^{+3}\text{-O}_2^{-1}$). Reduction by a second electron is thought to produce a low spin ferric peroxyanion ($\text{Fe}^{+3}\text{-O}_2^{-2}$), followed by uptake of a proton leading to the formation of the ferric hydroperoxo compound (compound 0, $\text{Fe}^{+3}\text{-OH}_2^{-1}$). This intermediate can

release a water molecule after the acceptance of a further proton forming an iron-oxo state (compound I, $\text{Fe}^{4+}\text{-O}$). Compound I is highly reactive and can, therefore, interact with the substrate to form the oxygenated product. The newly formed product is then released and the cycle can restart (Denisov et al., 2005). Besides the “typical” cycle, there are some uncoupling pathways like the autooxidation shunt, peroxide shunt or oxidase shunt (Figure 1.7 dashed arrows). These shunts decrease the catalytic efficiency by the unproductive consumption of NAD(P)H (Meunier et al., 2004).

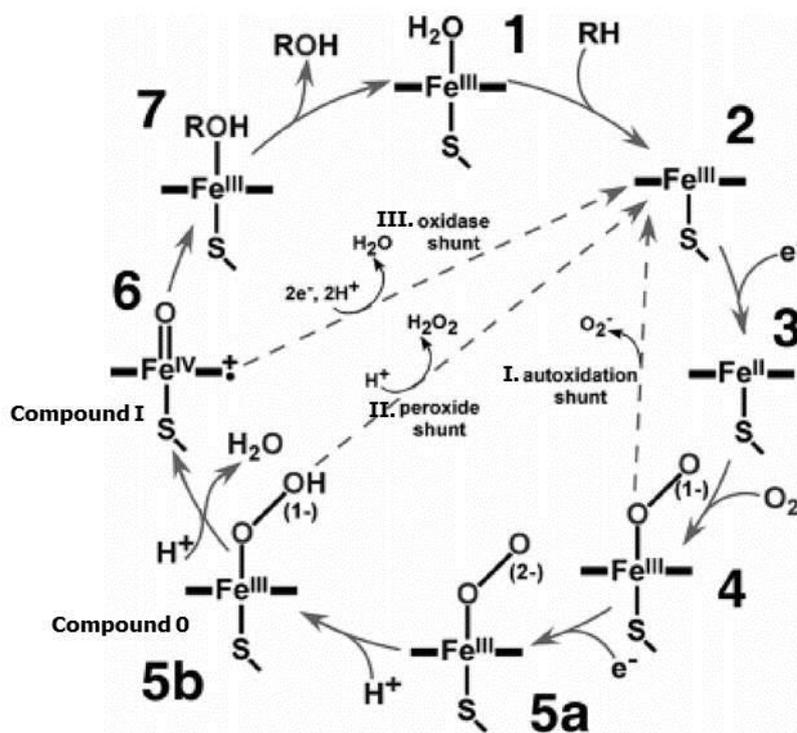


Figure 1.7. Catalytic cycle of cytochromes P450. Schematic representation of events leading to substrate hydroxylation at the heme iron. Dashed arrows indicate the autooxidation shunt (I), peroxide shunt (II) and oxidase shunt (III). [The cycle was adapted from Denisov et al., 2005]

1.2.4 FUNCTIONS AND EXAMPLES

From bacteria to human the broad spectrum of functions of P450s is very impressive. Bacterial P450s participate in the biosynthesis of secondary metabolites (e.g. antibiotics, antifungal, antitumor) and catalyze key reactions required for environmental bioremediation through the degradation of a variety of hydrocarbons (Lewis and Wiseman, 2005). The camphor-hydroxylating enzyme cytochrome P450_{cam} (CYP101A1) from the soil bacterium *Pseudomonas putida* (Gunsalus and Wagner, 1978) has

been studied in great detail. P450_{cam} was the first P450 enzyme to be structurally characterized (Poulos et al., 1987), and the role of the active site residues in substrate recognition has been examined by mutagenesis studies (Harford-Cross et al., 2000). This enzyme catalyzes the regio- and stereospecific hydroxylation of camphor to 5-*exo*-hydroxycamphor. The two large CYP105 and CYP107 families were found to participate in the degradation and biotransformation of a broad spectrum of xenobiotics as well as in secondary metabolites biosynthesis, and, therefore, are supposed to be of special interest for various biotechnological applications; for example CYP105P1 and CYP105H1 from *Streptomyces* sp. participate in the biosynthesis of the antifungals filipin and nystatin (Kelly et al., 2003; Xu et al., 2010), respectively, CYP107BR1 from *Pseudomonas autotrophica* participates in the activation of vitamin D₃ (Sakaki et al., 2011), CYP107E from *Micromonospora griseorubida* in the mycinamicin biosynthesis (Inouye et al., 1994), P450terf (CYP107L) from *Streptomyces platensis* in the hydroxylation of terfenadine (Lombard et al., 2011) and CYP107A1 from *Saccharopolyspora erythrea* in the synthesis of erythromycin.

Fungal P450s are required for the biosynthesis of various primary and secondary metabolites. Notable examples of the fungal primary metabolism are the synthesis of ergosterol, meiotic spore-wall biogenesis, and *n*-alkane hydroxylation, whereas fungal secondary metabolism deals with the biosynthesis of hormones and mycotoxins (gibberellin, aflatoxin and sterigmatocytin) (Crešnar and Petrič, 2011). Fungal P450s are also capable of detoxifying and degrading various xenobiotic compounds encountered in their environments, such as polycyclic aromatic hydrocarbons (PAHs), phenolic compounds, and other toxic environmental pollutants (Chen et al., 2014). A well-known and studied example is CYP51 (also called lanosterol 14 α -demethylase), which is involved in the conversion of lanosterol to 4,4-dimethylcholesta-8(9),14,24-trien-3 β -ol, a vital intermediate in the pathway leading to ergosterol production in fungi. Another very important role of fungal P450s in the cycling of nutrients includes CYP55, a fungal denitrification enzyme contributing to the nitrogen cycle (Durairaj et al., 2016).

In plants, P450s are involved in a wide range of biochemical pathways leading to the production of primary and secondary metabolites such as phenylpropanoids, alkaloids, terpenoids, lipids, cyanogenic

glycosides, and glucosinolates, as well as plant hormones (Mizutani, 2012). CYP51, CYP710, and CYP97 families catalyze different essential metabolic reactions, which are conserved in all plants (Mizutani and Ohta, 2010). *CYP51G* and *CYP710A* encode obtusifoliol 14 α -demethylase and sterol 22-desaturase, respectively, and are involved in sterol biosynthesis (Bak et al., 1997; Morikawa et al., 2006). CYP97A3 and CYP97C1 catalyze hydroxylation of the β - and ϵ -rings of carotenoids, respectively, in the biosynthetic pathway of xanthophylls, which play key roles in light harvesting and photoprotection (Kim and DellaPenna, 2006; Tian et al., 2004). CYP73A, CYP98A and CYP84A are involved in the hydroxylation of the aromatic ring of cinnamates, a core reactions conserved in all land plants, which provides vast arrays of phenolic compounds that function as structural components (e.g., lignin and suberin), UV protectants (flavonoids), antioxidants (polyphenols), antimicrobials (coumarins, lignans, isoflavonoids), and flavors (benzenoids, phenylpropenes) (Humphreys et al., 1999; Mizutani et al., 1993; Schoch et al., 2001). These chemical defenses are key strategies required for land plants to survive and expand their habitats. The key enzymes in the production of many plant hormones are also P450s. The biosynthesis of gibberellins (GAs), the diterpenoid plant hormones that regulate growth and influence various developmental processes (including stem elongation, germination, dormancy, flowering, sex expression, enzyme induction, and leaf and fruit senescence), require the activity of different P450s like CYP701A, CYP88A and CYP714D1 (Helliwell et al., 2001, 1998; Zhu et al., 2006).

P450s in arthropods are key players in metabolic resistance to natural plant chemicals and synthetic pesticides and are an integral component of growth, development and reproduction through their involvement in the processing of such endogenous substrates as pheromones and ecdysteroids (Feyereisen, 2011).

Humans have 57 sequenced P450 genes, which are essential contributors in the maintenance of general human health, particularly as they are related to the metabolism of pharmaceuticals, steroidogenesis, the metabolism of fatty acids and the conversion of procarcinogens and promutagens to deleterious genotoxic compounds. These genes are organized into 18 families and 43 subfamilies (Danielson, 2003). The largest number of human P450s, and most significant in terms of drug

metabolism, are those that are members of the CYP1, CYP2 and CYP3 families. CYP2 is the largest family of P450s in humans comprising approximately one third of human P450s sequences. This family of microsomal P450s is classified into 13 subfamilies that consist of 16 functional genes (CYP2A6, CYP2A7, CYP2A13, CYP2B6, CYP2C8, CYP2C9, CYP2C18, CYP2C19, CYP2D6, CYP2E1, CYP2F1, CYP2J2, CYP2R1, CYP2S1, CYP2U1, and CYP2W1) and 13 confirmed pseudogenes. The number of drugs metabolized primarily by CYP2D6 is very large including ~15–25% of all clinically used drugs from virtually all therapeutic classes, like antiarrhythmics (e.g. propafenone, mexiletine, flecainide), tricyclic and second generation antidepressants (e.g. amitriptyline, paroxetine, venlafaxin), antipsychotics (aripiprazole, risperidone), β -blockers (bufuralol, metoprolol), as well as anti-cancer drugs, in particular the selective estrogen receptor modifier (SERM) tamoxifen, several opioid analgesics including codeine and tramadol, and many others. The CYP3 family in humans consists of a single subfamily (CYP3A), within which there are four functional genes (CYP3A4, CYP3A5, CYP3A7, CYP3A43) and two pseudogenes. The CYP3A members overlap in their substrate specificities but show distinct patterns of tissue expression. The CYP3A subfamily enzymes play a major role in the metabolism of ~30% of clinically used drugs from almost all therapeutic categories. The active site of CYP3A4 is large and flexible and can accommodate and metabolize many preferentially lipophilic compounds with comparatively large structures (Zanger and Schwab, 2013). Typical large substrates are immunosuppressants like cyclosporin A and tacrolimus, macrolide antibiotics like erythromycin, and anticancer drugs including taxol, but smaller molecules are also accepted including ifosfamide, tamoxifen, benzodiazepines, several statins, antidepressants, opioids and many more (Nebert et al., 2013; Zhou et al., 2009).

The biosynthesis of steroid hormones (mineralocorticoids, glucocorticoids and sex hormones) starting from cholesterol requires the coordinated action of CYP11A1, CYP11B1/2, CYP17A1, CYP21A2 and the aromatase CYP19A1 in the adrenal cortex and gonads. CYP11A1 (cholesterol side-chain cleavage enzyme) catalyzes the initial and enzymatically rate-limiting step of steroid hormone biosynthesis in all steroidogenic tissues, which is the side-chain cleavage of cholesterol yielding pregnenolone, the common precursor of all steroid hormones. CYP17A1 (steroid 17 α -hydroxylase/17,20-lyase) possesses a steroid 17 α -hydroxylase and 17,20 carbon-carbon lyase activity and, thus, an important

branch point for the biosynthesis of glucocorticoids and sex hormones. It synthesizes the glucocorticoid precursor 17 α -hydroxyprogesterone in the adrenal zona fasciculata and testosterone, estrogen precursors DHEA and androstenedione from progesterone and pregnenolone in the adrenal zona reticularis and other steroidogenic tissues. CYP17A1 carbon-carbon lyase activity thereby depends on the presence of the allosteric effector cytochrome b₅, which is absent in the adrenal zona fasciculata. Following androgen biosynthesis by CYP17A1, CYP19A1 (aromatase) synthesizes estrogens from the androgen precursors in both, men and women. CYP21A2 (steroid 21-hydroxylase) produces the precursors for mineralo- and glucocorticoid biosynthesis in the adrenal by catalyzing the 21-hydroxylation of progesterone and 17-OH-progesterone yielding DOC and 11-deoxycortisol, respectively. The final steps of gluco- and mineralocorticoid biosynthesis are subsequently catalyzed by the 2 CYP11B subfamily members, CYP11B1 (steroid 11 β -hydroxylase) and CYP11B2 (aldosterone synthase) (Bernhardt and Waterman, 2007).

1.2.5 APPLICATION OF P450 SYSTEMS IN BIOCATALYSIS

In the past decades where “synthetic biology” is striving to replace or complement “synthetic chemistry” for production of high-value chemicals, the promiscuity and the versatility of P450s played a crucial role. According to the European Federation of Biotechnology, “Biotechnology” deals with the integrated use of biochemistry, microbiology, and engineering sciences in order to achieve technological (industrial) application of the capabilities of micro-organisms, cultured tissue cells, and parts thereof. Based on this, there is very promising progress for the commercial application of P450s enzymes.

Several P450s and engineered variants have been utilized in synthetic biology for valuable chemicals production by fermentation in microbial cells. For example CYP107BR1 (Vdh) from *Pseudonocardia autotrophica*, CYP105A1 from *Streptomyces griseolus* and CYP109E1 from *Bacillus megaterium* have been cloned and expressed for the production of 25-dihydroxyvitamin D₃ and 1 α ,25-dihydroxyvitamin D₃, the biologically active forms of vitamin D₃ (Abdulmughni et al., 2016; Sakaki et al., 2011). The production of cortisol from 11-deoxycortisol was achieved by using recombinant *Schizosaccharomyces pombe* and *Escherichia coli* strains expressing human CYP11B1 (Hakki et al.,

2008; Schiffer et al., 2015). Nootkatol and (+)-nootkatone were produced from valencene in *Saccharomyces cerevisiae* expressing the plant CYP71D51v2 (Gavira et al., 2013) and in *E. coli* expressing CYP109B1 (Girhard et al., 2010).

In most cases the implementation of the P450s in the industrial biotransformation processes requires further optimization steps including strain engineering, enzyme evolution and development of high cell density fermentation protocols as well as an effective downstream processing, which lead mostly to a very significant improvement of the biocatalytic system productivity. CYP105AS1 from *Amycolatopsis orientalis* was used for the production of the cholesterol lowering drug pravastatin in *Penicillium chrysogenum*. First, the integration of the complete compactin (ML-236B) gene cluster into the genome of the *P. chrysogenum* and the deletion of the esterase activity improved the compactin yield. For a stereoselective hydroxylation of compactin, CYP105AS1 was subjected to error-prone PCR and screening in order to achieve the desired pravastatin hydroxylase activity. Finally, to assess industrial performance of the *P. chrysogenum* strain, 10-L fed-batch fermentation was carried out leading to the production of more than 6 g/L pravastatin (McLean et al., 2015).

Another ambitious example is the production of the anti-malarial drug precursor artemisinic acid using the yeast *S. cerevisiae*, where the production could be improved drastically from few mg/L to about 25 g/L. The development of the process included the engineering of *S. cerevisiae* to increase the production of farnesyl pyrophosphate (FPP), which was subsequently converted to the amorphadiene by cloning of the plant amorphadiene synthase from *Artemisia annua*. By introducing CYP71AV1 from *A. annua*, amorphadiene was oxyfunctionalized via three intermediates yielding about 100 mg/L artemisinic acid (Ro et al., 2006). Further improvement of the system was achieved by overexpressing every enzyme of the mevalonate pathway (Westfall et al., 2012), co-expression of the cytochrome b5 gene and increasing the coupling efficiency between the CYP71AV1 and the CPR (Paddon et al., 2013). All these efforts and developments finally led to artemisinic acid yield of about 25 g/L in fermentation experiments.

1.3 BACILLUS MEGATERIUM

1.3.1 GENERAL

Bacillus megaterium is a nonpathogenic Gram-positive rod-shaped bacterium, which was described first in 1884 (De Bary, 1884). The complete genome sequence was published in 2011 (Eppinger et al., 2011). It is found in diverse ecological habitats and is able to grow on a wide variety of carbon sources, which allows for its cultivation on low-cost simple media. *B. megaterium* has gained a growing interest for the economic production of many enzymes and substances. It is one of the biggest known bacteria with a vegetative cell size of 4 μm in length and 1.5 μm in diameter, which is about 100 times larger than *E. coli* (Bunk et al., 2010). *B. megaterium* does not produce endotoxins associated with the outer membrane, nor does it have alkaline proteases, which improves the recovery of recombinant proteins (Vary, 1994). In the last decades enzymes from *B. megaterium* have been used industrially for the production of several substances and enzymes such as vitamin B₁₂ (Raux et al., 1998), oxetanocin, a viral inhibitor of HIV, hepatitis B virus and herpes virus (Morita et al., 1999; Shiota et al., 1996; Tseng et al., 1991), α - and β -amylase (Hebeda et al., 1988; Takasaki, 1989; Vihinen and Mantsiila, 1989), pencillin amidase (Martín et al., 1995; Suga et al., 1990), glucose dehydrogenase (Nagao et al., 1992) and HIV antigen (Shivakumar et al., 1998).

1.3.2 BACILLUS MEGATERIUM P450 SYSTEMS

Some of the most interesting proteins expressed naturally by *B. megaterium* are the P450s and related proteins (reductases and ferredoxins), which have been shown to be very important in many biotechnological and pharmaceutical applications (Abdulmughni et al., 2016; Brill et al., 2013; Kiss et al., 2015).

In general, *B. megaterium* encodes for several P450s. Among them, the self-sufficient CYP102A1 (also known as BM3), which is the most investigated bacterial P450 so far. BM3 is naturally a fatty acid hydroxylase (Miura and Fulco, 1974). It has been extensively studied over a period of almost forty years. The enzyme has been redesigned to catalyze the oxidation of non-natural substrates such as diverse as pharmaceuticals, terpenes and gaseous alkanes using a variety of engineering strategies

(Whitehouse et al., 2012). In addition, the biotechnologically valuable CYP106 family, CYP106A1 from *B. megaterium* strain DSM319 (Brill et al., 2013) and CYP106A2 from *B. megaterium* strain ATCC 13368 (Berg et al., 1979, 1976), was characterized to be associated with the biotransformation of a diverse array of substrates such as steroids and terpenoic substances (Brill et al., 2013; Schmitz et al., 2012). Furthermore, CYP109E1 was recently identified as vitamin D₃ hydroxylase from *B. megaterium strain* DSM319 (Abdulmughni et al., 2016; Brill et al., 2013; Jóźwik et al., 2016).

1.4 AIM OF THE PROJECT

The publication of the complete genome sequence of the *B. megaterium* strains DSM319 and QM B1551 in 2011 (Eppinger et al., 2011) enabled the characterization and identification of new enzymes. In addition, preliminary studies by our group showed the ability of *B. megaterium* to convert steroidal and non-steroidal substrates (Bleif et al., 2012; Brill et al., 2013; Gerber et al., 2015). Therefore, the scope of this work was the identification and characterization of new oxidoreductases (P450s, P450 reductases and dehydrogenases) from *B. megaterium*, in addition to the elucidation of their potential for biotechnological as well as pharmaceutical applications.

The first part of this thesis describes a di-flavin electron transfer system: a natural, soluble flavoprotein reductase from *B. megaterium* DSM319. The properties of the characterized reductase are analyzed at bioinformatic and biochemical levels. Its ability to support the electron transfer from NADPH either directly or indirectly using different ferredoxins to different classes of P450s should be studied. For this purpose the class I bacterial P450 (CYP106A1) and the eukaryotic, microsomal class II P450 (CYP21A2) will be used. In addition, the characterized reductase should be used in an *E. coli* whole-cell based biocatalyst system with both P450s classes.

The numbers of the newly identified P450s increased drastically over the past few years (Nelson, 2009), but there is nevertheless a still growing demand to exploit novel P450s as valuable biocatalysts in the industrial field. In the second part, this work reports the identification and characterization of a new plasmid-encoded P450 from the *B. megaterium* QM B1551. The new P450 will be first analyzed at the bioinformatic level, which is important for its classification. For biochemical investigations it has to be cloned and heterologously expressed in *E. coli*. Moreover, in view of the biotechnological application and since an effective electron transfer is essential in terms of cytochrome P450 activity, the search for a suitable redox partner/s will be crucial for the construction of *in vitro* as well as *in vivo* systems based on the new P450. This will then offer the possibility for the screening of new substrates, which could open up further biotechnological application of this enzyme.

The third part deals with the characterization of a short-chain dehydrogenase/reductase (SDR) with a 17 β -hydroxysteroid dehydrogenase activity. Moreover, the activity of the new dehydrogenase should also be investigated *in vivo* with biotechnologically valuable sesquiterpenes such as nootkatol for the production (+)-nootkatone.

2 PUBLICATION OF THE RESULTS

The results produced during the present work are published in the articles listed below:

2.1: **Milhim, M.**, Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016a. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. *J. Biotechnol.* 231, 83–94.

2.2: **Milhim, M.**, Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016b. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *J. Biotechnol.* 240, 68–75.

2.3: Gerber, A., **Milhim, M.**, Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. *Metab. Eng.* 36, 19–27.

2.4: Neunzig, J., **Milhim, M.**, Schiffer, L., Khatri, Y., Zapp, J., Sánchez-Guijo, A., Hartmann, M.F., Wudy, S.A., Bernhardt, R., 2017. The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1. *J. Steroid Biochem. Mol. Biol.* 167, 182–191.

2.1 Milhim et al., 2016a

A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase

Mohammed Milhim, Adrian Gerber, Jens Neunzig, Frank Hannemann and Rita Bernhardt

Journal of Biotechnology, 2016 Aug. 231: 83–94.

DOI: [10.1016/j.jbiotec.2016.05.035](https://doi.org/10.1016/j.jbiotec.2016.05.035)

Reprinted with permission of the Journal of Biotechnology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Biotechnology

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

A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase



Mohammed Milhim, Adrian Gerber, Jens Neunzig, Frank Hannemann, Rita Bernhardt*

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 9 February 2016
Received in revised form 20 May 2016
Accepted 25 May 2016
Available online 26 May 2016

Keywords:

NADPH-dependent cytochrome P450
oxidoreductase
Bacillus megaterium
CYP106A1
CYP21A2
Microsomal P450
Diflavin reductase

ABSTRACT

Cytochromes P450 (P450s) require electron transfer partners to catalyze substrate conversions. With regard to biotechnological approaches, the elucidation of novel electron transfer proteins is of special interest, as they can influence the enzymatic activity and specificity of the P450s. In the current work we present the identification and characterization of a novel soluble NADPH-dependent diflavin reductase from *Bacillus megaterium* with activity towards a bacterial (CYP106A1) and a microsomal (CYP21A2) P450 and, therefore, we referred to it as *B. megaterium* cytochrome P450 reductase (BmCPR). Sequence analysis of the protein revealed besides the conserved FMN-, FAD- and NADPH-binding motifs, the presence of negatively charged cluster, which is thought to represent the interaction domain with P450s and/or cytochrome c. BmCPR was expressed and purified to homogeneity in *Escherichia coli*. The purified BmCPR exhibited a characteristic diflavin reductase spectrum, and showed a cytochrome c reducing activity. Furthermore, in an *in vitro* reconstituted system, the BmCPR was able to support the hydroxylation of testosterone and progesterone with CYP106A1 and CYP21A2, respectively. Moreover, in view of the biotechnological application, the BmCPR is very promising, as it could be successfully utilized to establish CYP106A1- and CYP21A2-based whole-cell biotransformation systems, which yielded 0.3 g/L hydroxy-testosterone products within 8 h and 0.16 g/L 21-hydroxyprogesterone within 6 h, respectively. In conclusion, the BmCPR reported herein owns a great potential for further applications and studies and should be taken into consideration for bacterial and/or microsomal CYP-dependent bioconversions.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Bacillus megaterium is a nonpathogenic Gram-positive rod-shaped bacterium, which was described first in 1884. The complete genome sequence was published in 2011 (Eppinger et al., 2011). It is found in diverse ecological habitats and is able to grow on a wide variety of carbon sources, which allows its cultivation on low-cost simple media. It is one of the biggest known bacteria with a vegetative cell size of 4 μm in length and 1.5 μm in diameter, which is about 100 times larger than *Escherichia coli* (Bunk et al., 2010). *B. megaterium* does not produce endotoxins associated with the outer membrane, nor does it have alkaline proteases, which improves the recovery of recombinant proteins (Vary, 1994). In the last decades *B. megaterium* has been used industrially for the production of several substances and enzymes such as vitamin B₁₂ (Raux et al., 1998), oxetanocin, a viral inhibitor of HIV, hepatitis B virus and herpes virus (Morita et al., 1999), α - and β -amylase (Hebeda et al., 1988;

Takasaki, 1989), penicillin amidase (Martín et al., 1995), glucose dehydrogenase (Nagao et al., 1992) and HIV antigen (Shivakumar et al., 1998). Some of the most interesting proteins expressed naturally by *B. megaterium* are the cytochromes P450, which have been shown to be very important in many biotechnological and pharmaceutical applications (Brill et al., 2013; Kiss et al., 2015).

Cytochromes P450 (P450s) are heme-thiolate monooxygenases, which catalyze diverse reactions in a regio- and stereoselective manner and are involved in – for example but not limited to – the biosynthesis of hormones, signal molecules, defense-related chemicals and secondary metabolites in addition to their central role in the metabolism of endogenous (steroids and fatty acids) and exogenous (drugs and toxins) substances (Bernhardt, 2006). P450s catalyze a wide range of reactions including hydroxylations, epoxidations, deaminations, desulfurations, dehalogenations, N-, S-, and O-dealkylations, N-oxidations, peroxidations and sulfoxidations (Bernhardt and Urlacher, 2014). They are found in all kingdoms of life including mammals, plants, insects, fungi, bacteria and also in viruses (Nelson et al., 1996). P450s require auxiliary redox partners for the activation of the molecular oxygen. These redox partners transfer two electrons in two single steps from the cofactor NAD(P)H to the heme of the P450. Depending on com-

* Corresponding author at: Institute of Biochemistry, Campus B2.2, Saarland University, D-66123 Saarbrücken, Germany.
E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

of cytochrome *c* was monitored at 550 nm at 25 °C (the extinction coefficient of the reduced cytochrome *c* is 21 mM⁻¹ cm⁻¹) (Guengerich et al., 2009).

2.5. In vitro conversion and kinetic analysis

The *in vitro* conversion of the substrates was carried out with a reconstituted system in a final volume of 250 µl at 37 °C in conversion buffer (50 mM HEPES, pH 7.4, 20% glycerol, 100 µM DLPC). For the *in vitro* conversion of testosterone with CYP106A1 the reconstituted system contained 0.5 µM CYP106A1, 1 µM reductase (BmCPR, bovine AdR or Arh1), and 10 µM ferredoxin (Adx₄₋₁₀₈, Adx_{wt} or Fdx2). In case of CYP21A2, the reconstituted system contained 0.5 µM CYP21A2, 1 µM BmCPR or CPR. The NADPH regeneration system was 1 mM MgCl₂, 5 mM glucose-6-phosphate, 1 U glucose-6-phosphate dehydrogenase. Substrate was added to the needed final concentration. The reaction was started by adding NADPH (200 µM) and stopped by the addition of 250 µl of ethyl acetate, mixed vigorously, and extracted twice. After evaporating the combined organic phases to dryness, the residues were dissolved in the high performance liquid chromatography (HPLC) mobile phase (10% ACN) and subjected to HPLC analysis.

2.6. High performance liquid chromatography (HPLC) analysis

HPLC analysis was performed using a Jasco system (Pu-980 HPLC pump, an AS-950 sampler, a UV-975 UV/visible detector, and a LG-980-02 gradient unit; Jasco, Gross-Umstadt, Germany). A reversed-phase ec MN Nucleodur C18 (3 µm, 4.0 × 125 mm) column (Macherey-Nagel, Bethlehem, PA, USA) was used for all experiments at an oven temperature of 40 °C. Progesterone, testosterone and their metabolites were eluted from the column using a gradient method, starting with a mobile phase ratio of 10% ACN and increasing it to 100% ACN over 30 min. The detection wavelength was 240 nm.

2.7. Determination of kinetic parameters

K_M and k_{cat} values were determined by plotting the substrate conversion velocities versus the corresponding substrate concentrations using Michaelis–Menten kinetics and utilizing the program OriginPro 9.0G (OriginLab, Northampton, USA).

2.8. Activity measurement as a function of temperature

BmCPR or bovine CPR were incubated at the indicated temperatures (37 °C–75 °C) in a thermomixer for 10 min, followed by cooling on ice for 5 min. Then the protein solution was used to measure the ability of the BmCPR or bovine CPR to transfer electrons to the bovine CYP21A2 by assessing the conversion of progesterone to DOC (21-hydroxyprogesterone) as described previously.

2.9. Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectra were recorded at 30 °C using a JASCO J-715 spectropolarimeter over the wavelength range 190–260 nm at a protein concentration of 2 µM (0.135 mg/ml), dissolved in 10 mM potassium phosphate buffer pH 7.4 with the following parameters: path-length of 0.1 cm, data pitch of 0.1 nm, band width of 5 nm, accumulation 3 times. Spectra were recorded in triplicate and averaged.

Thermal denaturation measurement was performed in a 0.1 cm path-length thermostated quartz cell at a protein concentration of 2 µM (0.135 mg/ml) using a Jasco J-715 CD spectrophotometer. A heating rate of 1 °C/min was applied, and the CD signal at 209.5 nm

was monitored. The thermal denaturation data were fit to a derivation of the Hill fit equation (variable slope) using OriginPro 9.0 G to obtain the midpoint of denaturation (the thermal melting point).

2.10. Construction of vectors for E. coli whole-cell biotransformation

For the establishment of a CYP106A1-dependent whole-cell system, a tricistronic pET17b-based vector, expressing the CYP106A1-BmCPR-Fdx2 genes, was constructed. CYP106A1 coding sequence was amplified and cloned *via* the restriction sites *NdeI/HindIII*. The resulted vector (pET17b.CYP106A1) served then as a backbone for the cloning of BmCPR, which was amplified *via* PCR and cloned with the restriction sites *BamHI/NotI*. Fdx2 coding region was PCR amplified and cloned downstream CYP106A1-BmCPR using the restriction sites *NotI/XhoI*.

For the CYP21A2 whole-cell system, a bicistronic pET17b-based vector was constructed. The cDNAs of bovine CYP21A2 (modified as mentioned previously) (Arase et al., 2006) and the BmCPR coding gene were amplified and cloned in the pET17b vector *via* the restriction sites *NdeI/BamHI* and *BamHI/NotI*, respectively. All resulting vectors were verified by sequencing.

2.11. In vivo whole-cell biotransformation

E. coli C43 (DE3) cells were co-transformed with the suitable expression vector based on pET17b, coding for either CYP106A1-BmCPR-Fdx2 or bovine CYP21A2-BmCPR. In case of CYP21A2 expression, the cells were co-transformed with the chaperone GroEL/GroES-encoding plasmid pGro12, which has a kanamycin resistance gene (Brixius-Anderko et al., 2015; Nishihara et al., 1998). Transformed cells were grown overnight in 50 ml Terrific broth (TB) medium supplemented with 100 µg/ml ampicillin and 50 µg/ml kanamycin at 37 °C and shaking at 180 rpm. For the expression of proteins, 50 ml TB medium containing 100 µg/ml ampicillin and 50 µg/ml kanamycin were inoculated (1:100) with the transformed cells and cultivated at 37 °C with rotary shaking at 140 rpm. Protein expression was induced at OD₆₀₀ = 0.6–0.8 with 1 mM IPTG, 4 mg/ml arabinose (for induction of GroES/GroEL expression) and 1 mM δ-ALA as heme precursor. The temperature was then reduced to 30 °C. After incubation for 24 h, cells were harvested by centrifugation (4000g) for 20 min at 4 °C and washed once with 1 vol of conversion buffer (50 mM potassium phosphate buffer (pH 7.4) supplemented with 2% glycerol). After a second centrifugation, the cell pellets were resuspended in conversion buffer to an end cell suspension concentration of 60g wet cell weight (wcv)/L buffer. The substrate was added to a final concentration of 500 µM or 1 mM and the culture was incubated for the indicated time at 30 °C and 140 rpm. Substrate was extracted twice with the same volume of ethyl acetate and the organic phase was evaporated using a rotary evaporator. After that, the residues were dissolved in the high performance liquid chromatography (HPLC) mobile phase (10% ACN) and subjected to HPLC analysis.

3. Results

3.1. Identification of the reductase ORF from *Bacillus megaterium* DSM319

The usefulness of *B. megaterium* based whole-cell systems for the expression of several P450s has been shown before in our group (Brill et al., 2013; Gerber et al., 2015). It turned out that *B. megaterium* based whole cell system/s expressing microsomal cytochrome P450s without co-expression of redox partners showed activity towards the corresponding substrate/s (unpublished data). Searching the *B. megaterium* strain DSM319 for a

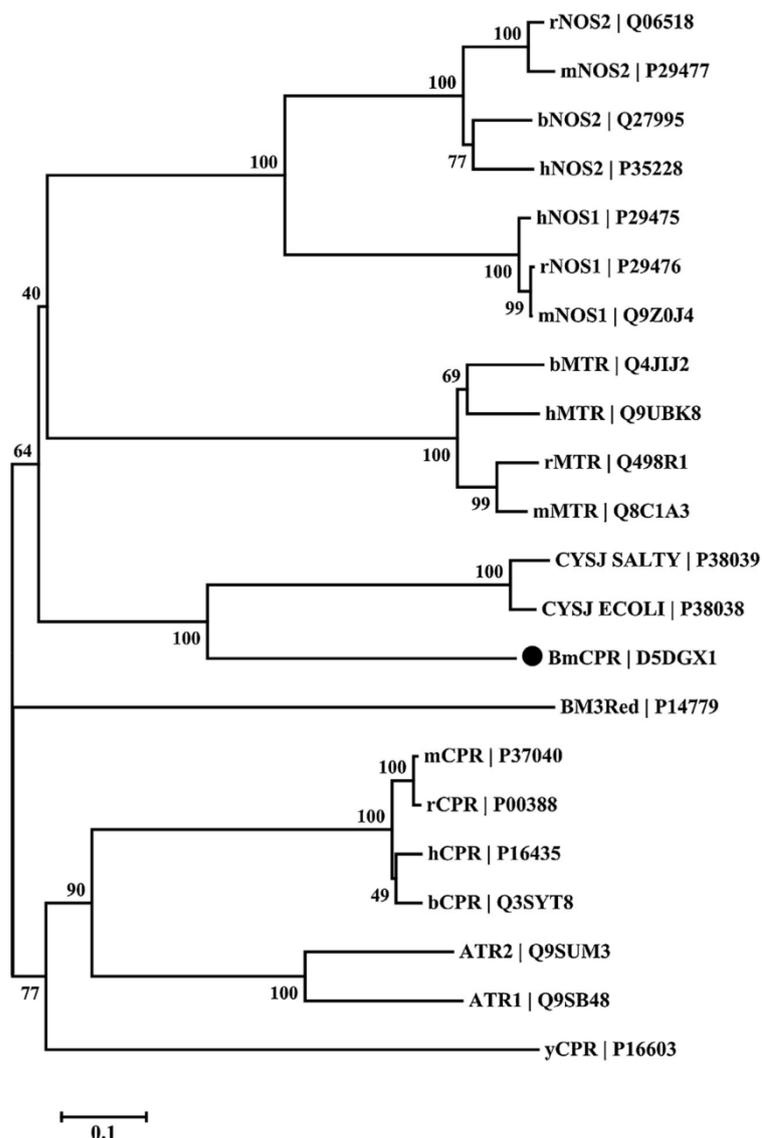


Fig. 1. Phylogenetic tree of di-flavin reductases displayed in Table 1. The alignment was done with 10-gap setting and 0.1-gap extension, with slow alignment input in ClustalW2 server. The tree was constructed by neighbor joining algorithm with bootstrap analysis of 1000 replicates. Bootstrap values are shown at the branch nodes. The scale bar represents 0.1 amino acid substitution per amino acid. The BmCPR is indicated with the closed circle (●). The number next to the gene name represents the UniProtKB accession number. For gene description refer to Table 1.

potential redox partner of P450s revealed an ORF (BMD_3122) [Uniprot acc. No. D5DGX1] encoding a suspected diflavin protein, which we referred to it through this publication *B. megaterium* cytochrome P450 reductase (BmCPR). Analyses of the cDNA showed that it encodes a 602 amino acid polypeptide starting with the rare start codon TTG (Supplementary material Fig. S1). The calculated molecular mass of the deduced amino acid sequence is 67.35 kDa, with a pI of 5.17. Prediction of transmembrane helices in BmCPR using the TMHMM Server v. 2.0 showed that the protein does not contain a transmembrane domain (Möller et al., 2001).

3.2. Sequence analysis

Alignment of the deduced amino acid sequence of BmCPR with some of the known diflavin CYPOR-like reductases of taxo-

nically diverse species (Table 1), shows a respective sequence identity and sequence similarity of 27%–30% and 44%–48% with the NADPH-cytochrome P450 reductases, 44% and 63% with the sulfite reductase [NADPH] flavoprotein alpha-component, 30%–33% and 45%–49% with the methionine synthase reductases, and 29%–32% and 44%–48% with the nitric oxide synthases. Amino acid sequences of the reductases displayed in Table 1 were used for the construction of a neighbor-joining phylogenetic tree, which shows that BmCPR is clustered with the CysJ, the alpha-component (or the flavoprotein component) of the sulfite reductases (SiR) (Fig. 1). This is in agreement with the analysis using the NCBI conserved domain database CDD server (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=271629>), which showed that the reductase domain of the BmCPR is related to the alpha subunit of the *E. coli* SiR subfamily, which belongs to the cytochrome P450 oxidoreductase-

BmCPR	-----MQLKVVNSPFNQEQADLLNRLLPRTLTEAQKMWLSGYLTAQAQSTSAEGTDPDVS-----TAAPAQTKQ	61
bCPR	MADSNMDAGTTTSEMVAEEVSLFSTTDVILFSLIVG---VMTYWFLFRKKK--EEVPEFTKIQTTSVSKDRSFVEKMKK	75
ATR1	-MTSALYASDLFKQLKSLMGTDLSLDDVVLVIATTSLALVAGFVLLWKKTTADRSGELKPLMIPKSLMAKDEDDLDLG	79
BM3Red	-----KKAEN	5
yCPR	-----MPFGIDNTDFTVLACLVLAVLLYVKNRSIKELLMSSDDGDITAVSSG-----NRDIAQVVT	55
	Phosphate Moiety	FMN Ring
BmCPR	TVSKDVTILYGSQTGNAQGLAENTGKTLEAKGFNVT--VSSMNDFKPN-----NLKKLENLIVVSTHGEPEPDNALS	133
bCPR	TGRN-IIVFYGSQTGTAEEFANRLSKDAHRYGMRG--MAADPEEYDLADLSSLPEIEKA-LAIFCMATYGEDPTDNAQD	151
ATR1	SGKTRVSIFFGTQTGTAEGFAKALSEEIKARYEKAQVAVKIDLDYAADDDQYEEKLLKKTLLAFFCVATYGDGEPDINAAR	159
BM3Red	AHNTPLLVLYGSNMGTAEGTARDLADIAMSKGFAP--QVATLDSHAG-----NLPREGAVLIVTASYN-GHPPDNAKQ	75
yCPR	ENNKNYLVLYASQTGTAEDYAKKFSKELVAKFNLN-VMCADVENYDFESLN----DVPVIVSIFISTYGEEDFPDGAVN	129
	FMN Ring	P450
BmCPR	FHEFLHGR--APKLENFRFVLSLGDSSYEF-FCQTGKEFDVRLAELGGERLYPR--VDCDLD-FEEPANKWLKGVIDG	207
bCPR	FYDWLQET---DVDLSGVKYAVFALGNKTYEH-FNAMGKYVDKRLQELGAQRIFDLGLGDDDN-LEEDFITWREQWPA	226
ATR1	FYKWFTEENERDIKLQQLAYGVFALGNRYEH-FNKIGIVLDELCKKAKRLIEVGLGDDQDQ--IEDDFNAWKESLWSE	237
BM3Red	FVDWLDQAS--ADEVKGVRYSVFGCGDKNWATYQKVPFIDETLAAKGAENIADRGEADASDD-FEGTYEWEHREHMWSD	152
yCPR	FEDFICNAE--AGALSNLRYNMFGLGNSTYEF-FNGAAKKAEKHLSAAGAIRLGLKGEADDGAGTTDEYMAWKDSILEV	206
BmCPR	LSEAKG-----HSASAAVPAEAPAGT-----SPYSRTNPFKAEVLENLNLNNGRGSNK	254
bCPR	VCEHFVGEATGEESIRQYELMVHTDMDMAKVYTGEMGRKLSYENQK-----PPFDAKNPFLAVVTTNRKLNQG-TER	298
ATR1	LDKLLKDEDKSVATPYTAVIPEYRVVTHDPRFTTQKSMESNVANGN-----TTIDIHPCRVDDVAVQKELHTHESDR	310
BM3Red	VAAYFNDIENSEDNK-----STLSLQFVDSADM-----PLAKMHGAFSTNVVASKLQQPGSAR	208
yCPR	LKDELHLDEQEAFTSQFYTVLNEITDSMSLGEPSAHYLPESHQLNRRNADGIQLGPFDLSPQYIAPIVKSRLEFSS-NDR	285
	FAD Ring	
BmCPR	ETRHLLESLGSGLTYPGDSLGIYPENDPELVLDLLLEFKWDASESVTVNKEGE-----TRPLR	314
bCPR	HLMHLELDISDSKIRYESGDHVAVYPANDSALVNQLGELGADLDIIMSLNLDL-----ESNKKHFPFCPTSYRTALTY	373
ATR1	SCIHLEFDISRTGITTYETGDHVGVAENHVEIVEEAGKLLGHSLLDVFVSIHADKEDGSPLESVAPPPFPCTGLTGLAR	390
BM3Red	STRHLEIELP-KEASYQEGDHLGVI PRNYEGIVNRVTRFGLDASQQIRLEAEEE-----KLAHLPLAKTVSVEELLQ	280
yCPR	NCIHSEFDLSGSNIKYSTGDHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDP-----TVKVFPTPTTIGAAIKH	358
BmCPR	EALISNFEITVLTKPLLKQAAELTGNKIKLALVEN-----REELKAYTQGRDVIDLVRDFGFWNSAQ-EFVAAILRKM	387
bCPR	YLDITNPPRTNVLYELAQYASEPTEHEQLRKMSSGEGKELYLRWVLEARRHILAILQDYPRLRPID-HLCELLPRLQ	452
ATR1	YADLLNPPRKSALVALAAYATEPSEAEKLLHLSF--DGKDEYSQWIVASQRSLEVMFAFPKPLGWFVFAAIAAPRLQ	468
BM3Red	YVELQDPVTRTQLRAMAAKTVCPPHKVELEALEK-----QAYKEQVLAKRLTMLELLEKYPACEMKFS-EFIALLP	354
yCPR	YLEITGPVSRQLFSSLIQFAPNADVKEKLTLLSKDK-DQFAVEITSKYFNADALKYLSDGAKWDTPVMQFLVESVPQMT	437
	FAD Ring	↓ ↓ ↓
BmCPR	ARLYSIASSLSANPDEVHVTIGAVRY--EAHGRE--RKGVCVLCSERLQ--PGDTIP-----V	440
bCPR	ARYYSIASSKVPNSVHICAVAVEY--ETKTGR-INKGVATSWLRRAKEPAGENGGRAL-----V	509
ATR1	PRYYSISSPRLAPSRVHVTLSALVYG--PTPTGR-IHKGVCSTWMMKNAVPAEKSHCSG-----A	525
BM3Red	PRYYSISSSPRVDEKQASITVSVVSG--EAWSGYGEYKGIASNYLAELQ--EGDTIT-----C	408
yCPR	PRYYSISSSSLSEKQTVHVTISIVENFPNPELDPAPPVGVTTNLLRNQLAQNNVNIATNLPHVYDLNGPRKLFANYKL	517
	NADPH	
BmCPR	YLQSNK-NFKLPQDQETPIIMVGPGTGVAPFRSFMQER---EETGA-----KGSWMPFGDQHFVTDFLYQTEWQKWL	509
bCPR	PMYVRKSQLRPPKATTPVIMVGPGTGVAPFFIGFIQERAWLRQQGKE-----VGETLLYGCRRSDEYLYREELAGFH	583
ATR1	PIFIRASNFKLPSNPSTPIVMVGPGTGLAPFRGFLQERMALKEDGEE-----LGSSLLFFGCRNRQMDFIYEDELNNFV	599
BM3Red	FISTPQSEFTLPKDPETPLIMVGPGTGVAPFRGFVQARKQLKEQGQS-----LGEAHLFYGCRSPHEDYLYQEELENAQ	482
yCPR	PVHVRRSNFRLPSNPSTPVIMIGPGTGVAPFRGFIIRERVAFLSESQKKGNNVSLGKHLIFLYGSRN-TDDFLYQDEWPEYA	596
	NADPH	
BmCPR	K--DGVLTKMDVAFSRDTEEKVYVQNRMLEHSEKLFQWLEEGA-SFYVCGDKTNMARDVHNTLVEIIEETEGKMSREQAEG	586
bCPR	K--DGALQNLNVAFSREQPQKVVYQHLLKKDKHEHLWKLIEGGAHIYVCGDARNMARDVQNTFYDVAEQGAMEHAQAVD	661
ATR1	D--QGVISELIMAFSREGAQKEYVQHKKMKEKAAQVVDLKEEG-YLYVCGDAKGMARDVHRTLHTIVQEQEGVSSSEAEA	676
BM3Red	S--EGITLHTAFSRMPNQPKTYVQHVMEQDGKLLIELLDQCA-HFYICGDSQMAPAVEATLMKSYADVHQVSEADARL	559
yCPR	KKLDGSEFMVVAHSRLPNTKVVYQDKLDYEDQVFEMINNGA-FIYVCGDAKGMAGVSTALVGLSRGKSIITTEATE	675
BmCPR	YLAEMKKQKRYQRDVY--	602
bCPR	YVKKLMTKGRYSLDVS-	678
ATR1	IVKKLQTEGRYLRDVW--	692
BM3Red	WLQOLEEKGRYAKDVWAG	577
yCPR	LIKMLKTSGRYQEDVW--	691

Fig. 2. ClustalW2 multiple sequence amino acid alignment of BmCPR and other CPRs. BmCPR (*B. megaterium* CPR), bCPR (bovine NADPH-cytochrome P450 reductase), ATR1 (NADPH-cytochrome P450 reductase 1 from *Arabidopsis thaliana*), BM3Red (Reductase domain of the bifunctional CYP102A2/BM3 from *B. megaterium*, amino acids 473–1049) and yCPR (yeast NADPH-cytochrome P450 reductase from *S. cerevisiae*). The FMN-, cytochrome P450-, FAD- and NADPH binding domains are labeled and highlighted gray. Arrows indicate the arginine residues proposed to be responsible for the interaction with the ferredoxin.

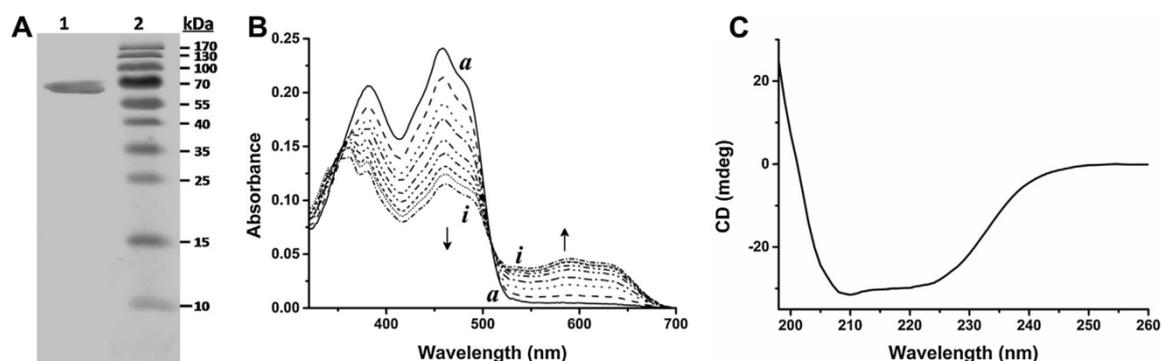


Fig. 3. Purification and spectral properties of the BmCPR. (A) 10% SDS-PAGE of the purified BmCPR. In lane 1 is a sample of the purified protein and lane 2 contains the bands of the prestained molecular weight ladder. (B) UV-vis spectrum of the purified BmCPR (spectrum a). Spectra (a–i) show the titration of the air-stable semiquinone BmCPR with NADPH under aerobic conditions. BmCPR (10 μ M) in 50 mM KPP (pH 7.4) was reduced with different NADPH concentrations ranging from 0 μ M (spectrum a) to 20 μ M (spectrum i) [added NADPH: 0, 2, 4, 6, 8, 10, 12, 16, 20 μ M]. Arrow shows the direction of the peak upon increasing NADPH concentrations. (C) Far-UV circular dichroism (CD) spectrum of the air-stable semiquinone BmCPR (2 μ M) in 10 mM potassium phosphate buffer pH 7.4. The CD spectrum was recorded in the region between 190 and 260 nm using the following parameters: path length 1 mm; time constant 2 s; band pass 5 nm; number of scans 3.

Table 1

Protein sequence identities and similarities values of the BMD_3122 with different reductases.

Closest match with different reductases					
Protein Abb.	Protein Name	No. Of aa	Species	Identity%/similarity% ^a	UniProtKB accession numbers ^b
BMD_3122	<i>Bacillus megaterium</i> CPR (BmCPR)	602	<i>B. megaterium</i> DSM319	100/100	D5DGX1
bCPR	NADPH-cytochrome P450 reductase	678	<i>Bos taurus</i>	30/48	Q35YT8
hCPR	NADPH-cytochrome P450 reductase	677	<i>Homo sapiens</i>	31/47	P16435
rCPR	NADPH-cytochrome P450 reductase	678	<i>Rattus norvegicus</i>	30/47	P00388
mCPR	NADPH-cytochrome P450 reductase	678	<i>Mus musculus</i>	30/47	P37040
BM3Red	Bifunctional P450/NADPH-P450 reductase domain	577	<i>B. megaterium</i> DSM319	30/51	P14779
ATR1	NADPH-cytochrome P450 reductase 1	692	<i>Arabidopsis thaliana</i>	30/48	Q9SB48
ATR2	NADPH-cytochrome P450 reductase 2	711	<i>Arabidopsis thaliana</i>	32/49	Q9SUM3
yCPR	NADPH-cytochrome P450 reductase	691	<i>S. cerevisiae</i> (strain ATCC 204508/S288c)	27/44	P16603
CYSJECOLI	Sulfite reductase [NADPH] flavoprotein alpha-component	599	<i>E. coli</i> (strain K12)	44/63	P38038
CYSJSALTY	Sulfite reductase [NADPH] flavoprotein alpha-component	599	<i>S. typhimurium</i> (strain LT2)	44/63	P38039
bMTR	Methionine synthase reductase	695	<i>Bos taurus</i>	33/49	Q4JJJ2
hMTR	Methionine synthase reductase	725	<i>Homo sapiens</i>	30/46	Q9UBK8
rMTR	Methionine synthase reductase	700	<i>Rattus norvegicus</i>	30/45	Q498R1
mMTR	Methionine synthase reductase	696	<i>Mus musculus</i>	32/47	Q8C1A3
hNOS1	Nitric oxide synthase, brain	1434	<i>Homo sapiens</i>	29/44	P29475
rNOS1	Nitric oxide synthase, brain	1429	<i>Rattus norvegicus</i>	29/44	P29476
mNOS1	Nitric oxide synthase, brain	1429	<i>Mus musculus</i>	29/44	Q9Z0J4
bNOS2	Nitric oxide synthase, inducible	1156	<i>Bos taurus</i>	30/48	Q27995
hNOS2	Nitric oxide synthase, inducible	1153	<i>Homo sapiens</i>	31/48	P35228
rNOS2	Nitric oxide synthase, inducible	1147	<i>Rattus norvegicus</i>	32/48	Q06518
mNOS2	Nitric oxide synthase, inducible	1144	<i>Mus musculus</i>	31/47	P29477

^a Based on NCBI BLAST (<http://www.blast.ncbi.nlm.nih.gov/>).

^b <http://www.uniprot.org/>.

like (CYPOR-like) sub-family having FAD and FMN as prosthetic groups and utilizing NADPH as electron donor. CYPOR-like reductases (including cytochrome P450 reductase, the reductase domain of the nitric oxide synthase, methionine synthase reductase and the reductase domain of the bifunctional cytochrome P450 BM-3) are ferredoxin reductase (FNR)-like proteins with an additional N-terminal FMN domain.

In order to analyze the similarity to its CYPOR homologs, the amino acid sequence of the BmCPR was aligned with NADPH-cytochrome P450 reductases of different origin: bovine CPR (bCPR), CPR from *Arabidopsis thaliana* (ATR1), the reductase domain of the bifunctional cytochrome P450 BM-3 from *B. megaterium* DSM319

and CPR from *S. cerevisiae* (yCPR) as shown in Fig. 2. Based on the distinct functional domains suggested for the FMN- and FAD-containing enzymes (Rana et al., 2013; Wang et al., 1997), BmCPR was found to contain all conserved [FMN/FAD/NADPH] domains, including the P450 interaction domain. FMN is located at the N-terminus (designated as phosphate moiety and FMN ring), followed by a flexible linker region, which separates the FAD binding domain from the FMN domain. The NADPH binding domain is present near the C-terminal end. An important feature in BmCPR is the presence of acidic residues, which are proposed to be important for the interaction with the microsomal P450s as well as cytochrome c (Fig. 2).

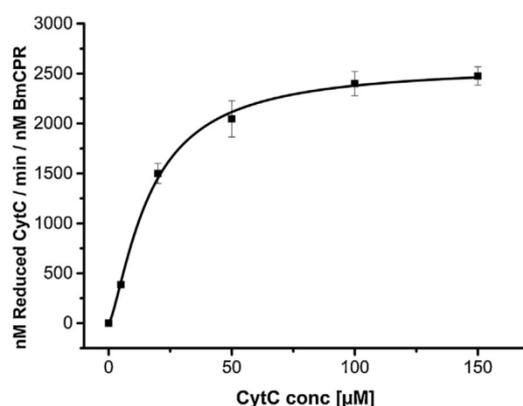


Fig. 4. Determination of the kinetic parameters of the reduction of cytochrome *c* (CytC) by BmCPR. The data is represented as mean \pm SD of three separate measurements. Data was fitted using originPro 9.0.G software.

3.3. Heterologous expression and purification

The DNA fragment encoding the full length BmCPR was PCR amplified from the genomic DNA of *B. megaterium* strain DSM319 and cloned into the expression vector pET17b. For purification with IMAC, a His-tag was introduced at the C-terminus. *E. coli* C43 (DE3) was transformed with the resulting vector. The full length BmCPR could be purified in soluble form as shown in the SDS-PAGE (Fig. 3A). The absorption spectrum of the purified BmCPR was characteristic for oxidized flavoproteins, with absorption maxima at 380 nm and 456 nm, a shoulder at 485 nm and a broad peak at 585 nm (Fig. 3B spectrum a), which represents the air-stable semiquinone form (Vermilion and Coon, 1978). This spectral property of the BmCPR suggested that the encoded protein is functionally active after purification. Titration of the oxidized enzyme with NADPH under aerobic conditions lead to a decrease of the absorbance at 456 nm and 380 nm, and an increase at 585 nm, with isosbestic points at 340 nm and 508 nm (Fig. 3B, spectra a–i). The CD spectrum was measured in the far-UV region for the air-stable semiquinone BmCPR, which showed a negative dichroic double band with minima at 208 and 223 nm (Fig. 3C), an indication of an α -helical protein (Greenfield, 2006), which is typical for the di-flavin reductases (French and Coon, 1979).

3.4. Enzymatic activity of BmCPR

The cytochrome *c* test is simple and widely used for the investigation of the electron transfer ability of reductases and, thus, it was applied to measure the reduction activity of the BmCPR. Different concentrations of cytochrome *c* were mixed with 50 pmol BmCPR in buffer followed by starting the reaction with NADPH. The enzyme was found to obey Michaelis–Menten kinetics (Fig. 4) and exhibited a K_M of $16.7 \pm 1.2 \mu\text{M}$ and a k_{cat} of $2582 \pm 72 \text{ min}^{-1}$.

After having established that BmCPR can reduce cytochrome *c*, the ability of the BmCPR to support electron transfer to two different classes of P450s was investigated. A class I P450, represented by the bacterial CYP106A1 and a class II P450s, represented by the microsomal bovine CYP21A2, were used.

3.5. In vitro reconstitution of CYP106A1 activity

To investigate the ability of the BmCPR to support electron transfer to the heme group of the P450s, *in vitro* assays were first carried out with the bacterial CYP106A1 from *B. megaterium* DSM319,

Table 2

In vitro conversion of testosterone (200 μM) with CYP106A1 and different redox systems.

System ^a	Conversion (%) ^b
CYP106A1-BmCPR	0
CYP106A1-BmCPR-Fdx2	95.5 \pm 0.9
CYP106A1-BmCPR-Adx ₄₋₁₀₈	20.4 \pm 0.1
CYP106A1-BmCPR-Adx _{wt}	3.9 \pm 0.1
CYP106A1-AdR-Fdx2	15.0 \pm 0.4
CYP106A1-AdR-Adx ₄₋₁₀₈	42.2 \pm 3.1
CYP106A1-AdR-Adx _{wt}	7.3 \pm 0.1
CYP106A1-Arh1-Fdx2	89.4 \pm 2.5
CYP106A1-Arh1-Adx ₄₋₁₀₈	35.8 \pm 1.0
CYP106A1-Arh1-Adx _{wt}	6.8 \pm 0.2

Fdx2: Ferredoxin from *B. megaterium* DSM319 (Brill et al., 2013).

^a CYP:Reductase:Ferredoxin ratio (1:3:20).

^b Conversion ratio within 1 h.

which was recently purified and characterized by our group (Brill et al., 2013).

Natural (or physiological) substrate for CYP106A1 is still unknown; therefore, testosterone was used as test substrate. Hence, testosterone has previously been identified as a substrate for CYP106A1 resulting in 3 products (15 β -, 6 β -, and an unknown product) (Kiss et al., 2015; Lee et al., 2015). In addition, our unpublished data [supported by data published by Lee et al. (2015)] showed that among all the characterized substrates for CYP106A1, testosterone was found to have the best affinity. The *in vitro* conversion of testosterone by CYP106A1 was performed using the BmCPR, the bovine AdR or the Arh1 (adrenodoxin reductase homologue 1 from *S. pombe*) reductases in combination with different ferredoxins; bovine Adx₄₋₁₀₈, Adx_{wt} or the previously identified ferredoxin Fdx2 form *B. megaterium* DSM319 (Brill et al., 2013). BmCPR was found to accept electrons from NADPH but not from NADH (data not shown). The results presented in Table 2 showed that the BmCPR, in combination with all tested ferredoxins, was able to support the activity of CYP106A1 *in vitro*, with an obvious higher efficiency using Fdx2. The conversion ratio of the 200 μM testosterone within 1 h was \sim 96% using the BmCPR and Fdx2 as redox partners, which is higher than the previously reported CYP106A1-AdR-Adx₄₋₁₀₈ (Kiss et al., 2015) and CYP106A1-Arh1-Fdx2 (Brill et al., 2013) systems with a conversion ratio of \sim 42% and \sim 89%, respectively.

3.6. BmCPR can replace microsomal P450 reductase

As mentioned previously, the BmCPR was found to contain the distinct functional domains present in FMN- and FAD- containing reductases, in addition to a suggested P450-interaction domain. These characteristics made it a good candidate for electron transfer from NADPH to the heme center of microsomal P450s. We demonstrated that the BmCPR was able to support the activity of bovine CYP21A2 as indicated by the conversion of progesterone to DOC using NADPH as electron donor. The conversion of progesterone to DOC by bovine CYP21A2 using BmCPR as redox partner followed Michaelis–Menten kinetics with a K_M of $1.82 \pm 0.12 \mu\text{M}$ and a k_{cat} of $1.27 \pm 0.02 \text{ min}^{-1}$ resulting in a k_{cat}/K_M of $697.80 \text{ mM}^{-1} \text{ min}^{-1}$. Interestingly, the catalytic efficiency of CYP21A2 in the presence of BmCPR is higher than that when using the bovine CPR as redox

Table 3

Kinetic parameters of the *in vitro* conversion of progesterone with bovine CYP21A2 and BmCPR or the bovine CPR as redox partner.

Reductase	K_M [μM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$\text{mM}^{-1} \text{ min}^{-1}$]	R^2
BmCPR	1.82 \pm 0.12	1.27 \pm 0.02	697.80	0.99
bCPR	2.24 \pm 0.13	0.91 \pm 0.01	406.25	0.99

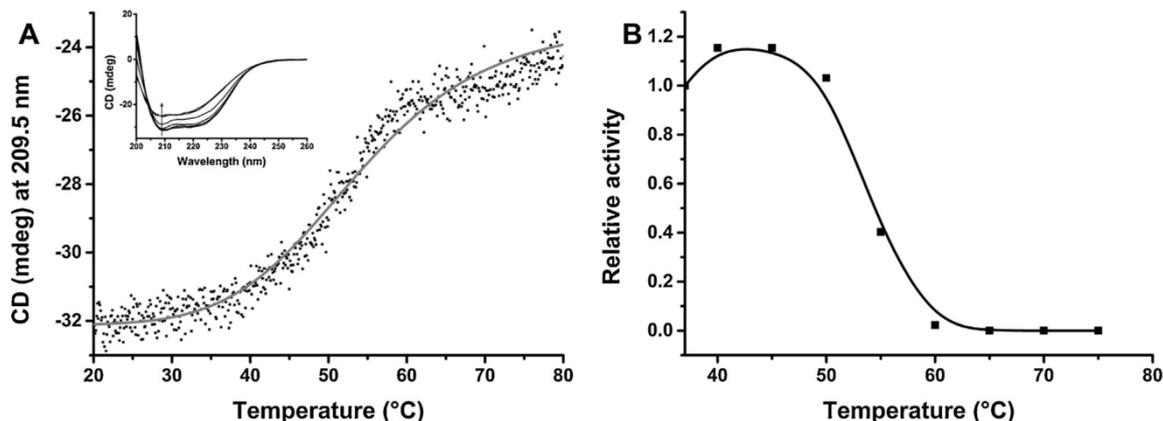


Fig. 5. The effect of temperature on the BmCPR. (A) Circular Dichroism (CD) thermal denaturation curve. The ellipticity was measured at 209.5 nm as a function of temperature in the range between 20 and 80 °C using a temperature slope of 1 °C/min, data pitch 0.1 °C. The data points were plotted and fitted (gray solid line) with OriginPro 9.0G program. For the measurement, the concentration of the BmCPR was 2 μM (0.135 mg/ml) resuspended in 10 mM potassium phosphate buffer pH 7.4. The inset shows the CD spectra of the BmCPR in the far-UV region at the indicated temperatures (Arrow show the direction of the peak by increasing the temperature at 209.5 nm). The CD spectra were recorded between 200 and 260 nm every 10 °C. (B) Thermal stability of the reductase. BmCPR was incubated at the indicated temperatures (37 °C to 75 °C) in a thermocycler for 10 min, followed by cooling on ice for 5 min. Then the protein solution was used to measure the ability of the BmCPR to transfer electrons to the bovine CYP21A2 by assessing the conversion of progesterone to DOC. The product formation was determined by HPLC. Relative activity on the Y-axis represents the normalized activity at the specific temperature against the activity at 37 °C.

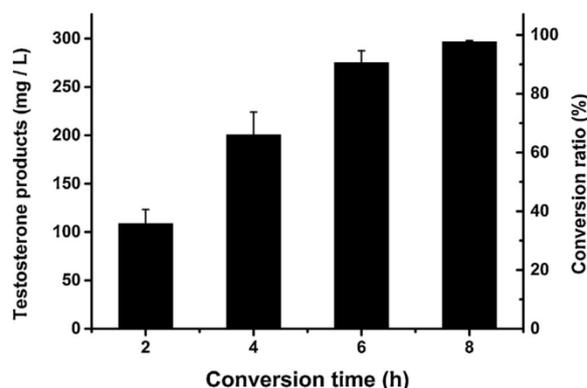


Fig. 6. Time dependent *in vivo* conversion of 1 mM testosterone by a CYP106A1 based whole-cell system with BmCPR and Fdx2 as redox partners. Whole-cell bio-transformation was performed using C43 (DE3) cells carrying the pET17b vector expressing CYP106A1, BmCPR and Fdx2. Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4+2% glycerol) and samples were taken at the indicated times and analyzed via HPLC. The data is represented as mean ± SD of three separate measurements.

partner, which has a k_{cat}/K_M of 406.25 mM⁻¹ min⁻¹ (Table 3) (Supplementary material Fig. S2).

3.7. Thermal stability of the BmCPR

Heating a protein solution gradually above a critical temperature causes a transition from the native state to the denatured state. The temperature at the transition midpoint, where the concentration ratio of native and denatured states is 1, is known as the melting temperature (T_m). Changes in the secondary structure of a protein can be monitored by circular dichroism (CD) spectroscopy in the far-UV region (190–260 nm) (Fig. 5A inset).

The thermal stability of the BmCPR was assessed by recording the increase in the CD signal at a single wavelength (209.5 nm) over the temperature range from 20 to 80 °C (Fig. 5A). The changes in the CD signal at 209.5 nm were recorded every 0.1 °C. Plotting the CD values at this wavelength as a function of temperature gave a

distinctive sigmoidal shaped curve with a midpoint of transition (the melting temperature) of 53.8 ± 0.2 °C. The denaturation of the BmCPR was found to be irreversible, which was observed from the turbidity of the solution, an indication of protein aggregation and precipitation. In addition, cooling the sample down did not cause BmCPR refolding.

In addition, we measured the temperature at which the reductase loses 50% of its activity, and we referred to it as (T_{50}). For this purpose, the reductase was incubated at the indicated temperatures for 10 min and cooled down on ice. The activity of the enzyme was estimated by its ability to support electron transfer to the bovine CYP21A2 for the conversion of progesterone to DOC. The T_{50} value was found to be 54.1 ± 0.4 °C (Fig. 5B) for BmCPR, which correlates with the melting temperature. To the best of our knowledge, there are very few data available concerning the thermal stability or the melting point of the different CPRs. For example the rat CPR (rCPR) has a thermal stability T_{50} of 46.1 °C, whereas, the CPR from hot pepper *Capsicum annuum* (CaCPR) showed a T_{50} of 56.1 °C (Lee et al., 2014).

3.8. Establishment of whole-cell systems using BmCPR as redox partner

After demonstrating the *in vitro* activity of the BmCPR, we investigated the efficiency of applying the BmCPR in *E. coli* whole-cell systems with CYP106A1 or CYP21A2.

CYP106A1 has a high biotechnological and industrial potential, as it has been found to catalyze the hydroxylation of a set of steroids as well as non-steroidal substrates such as the triterpene 11-keto-β-boswellic acid (Brill et al., 2013; Kiss et al., 2015). For the establishment of a CYP106A1-dependent whole-cell system, a tricistronic pET17b-based vector, expressing CYP106A1, BmCPR and Fdx2, was constructed. The whole-cell biotransformation was performed with resting cells in potassium phosphate buffer instead of the terrific broth complex medium, since indole, which results from tryptophan metabolism by *E. coli*, may have an inhibitory effect on the activity of P450s (Ringle et al., 2013).

Testosterone was found to be efficiently converted by the established system, verifying the *in vivo* functionality and efficiency of the BmCPR (Fig. 6). With a substrate concentration of 1 mM, a

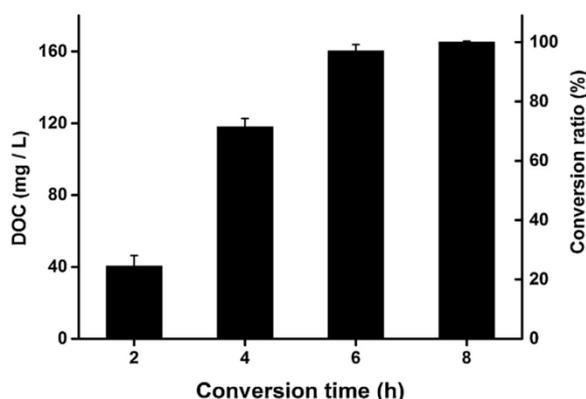


Fig. 7. Time dependent conversion of 0.5 mM progesterone by a bovine CYP21A2-dependent whole-cell system with BmCPR as redox partner. Whole-cell biotransformation was performed using C43 (DE3) cells carrying pET17b vector expressing CYP21A2 and BmCPR. Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4+2% glycerol) and samples were taken at the indicated time and analyzed via HPLC. The data is represented as mean \pm SD of three separate measurements. DOC: 21-hydroxyprogesterone.

conversion ratio of \sim 98% was achieved after 8 h, yielding \sim 0.3 g/L product.

Previously, efforts were performed to establish an *E. coli* based whole-cell system using CYP21A2. Naturally, the electron transfer to CYP21A2 is mediated by the membrane bound microsomal CPR, which is not easily expressed in *E. coli*. Therefore, for efficient supplementation of the CYP21A2 with electrons, CPR was replaced by soluble redox systems originating from the fission yeast *S. pombe* and consisting of Arh1 and the ferredoxin domain of electron transfer protein 1 (Etp1^{fd}) (Brixius-Anderko et al., 2015).

The high *in vitro* activity and the solubility of the BmCPR make it a very promising candidate for the establishment of an *in vivo* whole-cell system with CYP21A2. For this purpose, a bicistronic pET17b-based vector, expressing the CYP21A2 and BmCPR, was constructed. As illustrated in Fig. 7, using BmCPR as redox partner, 500 μ M progesterone could be completely converted after \sim 6 h with a conversion ratio of about 99%, yielding 0.16 g/L product.

4. Discussion

A soluble NADPH-dependent diflavin reductase from *B. megaterium* DSM319 was identified, cloned and expressed in *E. coli*. Sequence analysis showed that this reductase contains the FAD and FMN binding motifs as well as NADPH and cytochrome P450 interaction domains. This reductase exhibits high similarity to CysJ, the SiR alpha-component, implying its natural activity in *B. megaterium*. Additionally, in the *B. megaterium* genome the gene of the characterized reductase is located directly downstream of an open reading frame, which encodes the beta-component of the SiR. The SiR is known to be a hemoflavoprotein complex, composed of two subunits, the alpha (flavoprotein) component and the beta (hemoprotein) component, which is responsible for the reduction of sulfite to sulfide, an important step in the sulfate assimilation pathway in bacteria (Christner et al., 1981; Guillooard et al., 2002). The SiR alpha-component was first characterized in *S. typhimurium* and *E. coli*, and it was demonstrated that it shares spectroscopic and catalytic properties with the NADPH-cytochrome P450 reductase (Ostrowski et al., 1989). Evolutionary, there are indications that diflavin reductases evolved as a result of a fusion of two proteins, the bacterial flavodoxin (FMN-domain) and the ferredoxin-NADP⁺ reductase FNR (FAD-domain) (Porter and Kasper, 1986). Cytochrome P450 reductase (CPR) was the first

enzyme of this family to be isolated (Horecker, 1950), followed by several other dual flavin enzymes, sulfite reductase in bacteria (Christner et al., 1981; Ostrowski et al., 1989) as well as three proteins identified in human: nitric oxide synthase (Stuehr, 1997), methionine synthase reductase (Olteanu and Banerjee, 2001) and a cytoplasmic protein NR1 with yet unknown function that is expressed in cancer cells (Paine et al., 2000). This family includes also the diflavin domain of self-sufficient cytochromes P450 such as CYP102A1 (BM-3) from *B. megaterium*.

The characterized BmCPR reductase showed reducing activity towards cytochrome *c* with a catalytic activity of 2582 min⁻¹, which is slightly lower than that of the CPR from *Candida albicans* (2830 min⁻¹) (Park et al., 2010) and of rat CPR (3000 min⁻¹) (Guengerich et al., 2009) and much lower than that of the truncated yeast CPR comprising amino acids 44–691 (8940 min⁻¹) (Aigrain et al., 2011). The determined cytochrome *c* reducing activity was, however, higher than that of the truncated human CPR comprising amino acids 61–677 (1800 min⁻¹) (Aigrain et al., 2011) and of the newly characterized CPR from *Candida apicola* (1915 min⁻¹) (Girhard et al., 2012). Cytochrome *c* is not a physiological redox partner for CPRs. Nevertheless, there is evidence that the binding sites for cytochrome P450 and cytochrome *c* on the acidic FMN domain of the reductase are either identical or overlapping (Nisimoto, 1986; Shen and Kasper, 1995).

B. megaterium DSM319 contains different P450s. One of them (CYP106A1) has been recently characterized (Brill et al., 2013). However, since no homologous redox system could be identified so far for CYP106A1, the activity of CYP106A1 was supported formerly with heterologous redox partners such as the bovine AdR/Adx system (Kiss et al., 2015), a mixed system consisting of the yeast Arh1 and *B. megaterium* Fdx2 (Brill et al., 2013), or the spinach system FdR/Fdx (Lee et al., 2015). Since class I bacterial P450s are not able to accept electrons directly from ferredoxin reductases or diflavin reductases, the requirement of ferredoxin auxiliary protein in class I P450s system for electron transfer is crucial. We demonstrate here that the BmCPR/Fdx2 system was able to support the activity of the bacterial CYP106A1, and the activity was found to be higher than with the previously used redox partners. In class I P450 systems, FAD of the ferredoxin-reductase serves as an electron acceptor from NAD(P)H, which transfers the reducing equivalents to the iron-sulfur center of ferredoxin, which in turn reduces the P450s. Salt bridges have been identified previously to play an important role in the interactions between ferredoxin reductase and ferredoxin in the P450 redox chains (Lambeth et al., 1979). Crystallized putidaredoxin reductase-putidaredoxin (PdR/Pdx) complex from *Pseudomonas putida* showed that the interaction is mediated by basic amino acids on the surface of the PdR (R65, R310, K339, K387 and K409) (Sevrioukova and Poulos, 2010). In addition, site directed mutagenesis (Brandt and Vickery, 1993) and solved crystal structure of the bovine AdR/Adx complex (Müller et al., 2001) have provided a detailed insight into the electrostatic binding mechanism between these proteins. The interaction sites were found to be composed of the acidic amino acids of the Adx and a basic cleft region of the AdR (R211, R240 and R244). These residues are conserved in other ferredoxin-reductases such as the Arh1 from *S. pombe* as well as *S. cerevisiae* (Ewen et al., 2008). Considering the BmCPR FAD-domain depicted in Fig. 2, three arginines (R412, R418 and R420) and one lysine (K421) residue were identified (Fig. 2 arrows), which are likely to be involved in binding with the ferredoxins. The presence of such a basic region in the FAD-domain of the BmCPR suggests a similar interaction mechanism of the class I redox proteins and the *B. megaterium* system (BmCPR/Fdx2) during intermolecular electron transfer reaction to the CYP106A1. The geometrical localization of these basic amino acids on the surface of the BmCPR as well as their importance in the electron transfer

	Acidic cluster										
	Cluster 1					Cluster 2					
BmCPR	D	C	D	L	D	F	–	E	E	P	196
bCPR	D	D	D	G	N	L	–	E	E	D	215
ATR1	D	D	D	Q	S	I	–	E	D	D	226
yCPR	D	D	G	A	G	T	T	D	E	D	195
CYSJ_ECOLI	D	A	D	V	E	Y	–	Q	A	A	193

Fig. 8. Sequence comparison of the cytochrome P450 binding motif. ClustalW2 multiple sequence amino acid alignment of BmCPR and other diflavin reductases (see Table 1 for description) showing the acidic amino acid cluster involved in the interaction with the microsomal P450s and cytochrome c.

and/or formation of the BmCPR/Fdx2 complex is still to be elucidated.

Interactions between eukaryotic flavoproteins and bacterial P450s have been demonstrated in numerous examples, including the use of the bovine AdR together with Adx to transfer electrons to some bacterial P450s, namely CYP106A2 from *B. megaterium* (Hannemann et al., 2006), CYP109D1 and CYP264B1 from *S. celulosum* (Khatri et al., 2010; Ly et al., 2012), and CYP109B1 from *B. subtilis* (Girhard et al., 2010). On the other hand, examples of interactions between bacterial flavoproteins and eukaryotic P450s, such as the use of the reductase domain of the of the BM-3 form *B. megaterium* with CYP2E1 (Fairhead et al., 2005), CYP3A4 and CYP2B4 (Davydov et al., 2010), are rarely described in literature. So to the best of our knowledge, the BmCPR is the first soluble flavoprotein of bacterial origin, which can support the activity of eukaryotic class II microsomal P450s, without the need of an auxiliary ferredoxin. Interestingly, the characterized BmCPR was found to be 1.7 times more efficient than the bovine CPR in supporting the conversion of progesterone to DOC by the bovine CYP21A2 *in vitro*.

The exchangeability between the different diflavin reductases, which belong to the FNR-like superfamily, was first investigated by Barnes et al. (1991). In that study, the ability of the *E. coli* sulfite reductase to transfer electrons to P450s was excluded, since a soluble protein fraction from an *E. coli* strain lacking the SiR alpha-component of the SiR (CysJ) was still able to support electron transfer to bovine CYP17A1 (Barnes et al., 1991). In contrast, Zeghouf et al. (1998) could show that the SiR alpha-component of *E. coli* (CysJ) can support the electron transfer to CYP17A1 to hydroxylate pregnenolone *in vitro*, but the activity was 12–15 times less than that of the bovine CPR (Zeghouf et al., 1998). In the microsomal P450 systems, where the P450 and the reductase are both membrane associated, the electron transfer was reported to occur by a transient complex formation between the P450 and the reductase (Shen and Kasper, 1995). Similar to the interaction in the ferredoxin reductase/ferredoxin complex, it is known that electrostatic interactions are also responsible for the stabilization of the P450-reductase complex during the electron transfer from the FMN to the heme center (Bernhardt et al., 1988), where the acidic residues on the surface of the reductase interact with the basic residues on the surface of the P450. Sequence alignment of the BmCPR with different CPRs and CysJ from *E. coli* showed that the acidic residues cluster 1 and cluster 2 (¹⁸⁸DCDLDFEEP¹⁹⁶) in the P450 interaction domain are mostly conserved in the BmCPR (Fig. 8), while in the CysJ from *E. coli*, cluster 2 is composed of uncharged amino acids (¹⁹¹QAA¹⁹³).

Previously, several efforts were directed to improve the biotechnological applicability of P450s, however, P450s are dependent on an intermolecular electron transfer from NAD(P)H via one or two redox proteins. As a result, the search for an efficient electron supply from NAD(P)H is necessary for P450s to be applied in a biotechnological process.

The *in vitro* activity of the BmCPR with both CYP106A1 (class I) and CYP21A2 (class II) P450s was very encouraging to address the question whether the BmCPR is also applicable *in vivo*. Therefore, *E. coli* CYP106A1- and CYP21A2-based whole-cell biotransformation systems, which utilize BmCPR as redox partner, were established. The BmCPR containing systems were found to be functional and very efficient. The CYP106A1-BmCPR-Fdx2 system yielded ~0.9 g/L/d hydroxylated testosterone products (15β-, 6β-, and an unknown product) and the CYP21A2-BmCPR yielded ~0.64 g/L/d progesterone product (DOC). This clearly demonstrates the value of BmCPR for the development of efficient whole-cell systems using the bacterial or the mammalian cytochrome P450s for biotechnological applications.

Herein we demonstrated, to the best of our knowledge, for the first time the use of a diflavin reductase from prokaryotic origin to transfer electrons to a microsomal P450. The electron transfer from the CPR to the microsomal P450 needs the interaction between both enzymes, where the N-terminus of the CPR plays an important role (Pandey and Flück, 2013). Thus, the ability of a soluble diflavin reductase of bacterial origin (BmCPR) to support the activity of CYP21A2 more efficiently than its natural redox partner CPR seems to be very interesting: on one hand, for biotechnological application/s, as the expression of soluble proteins is much easier than that of membrane proteins, and, on the other hand, BmCPR could be a supportive tool in the detailed understanding and analyses of the interaction, which takes place between the reductase and the P450. From another point of view, the characterization and deeper understanding of such a reductase is of great importance for the re-evaluation and exploration of the phylogenetic relationships of NADPH cytochrome P450 reductases (CYPOR) between different taxa, which enable us to characterize non-conventional alternative redox partners.

In summary, we have cloned a diflavin reductase from *B. megaterium* DSM319 (BmCPR). The BmCPR showed *in vitro* and *in vivo* activity with class I and class II P450s, which make this reductase a very promising and advantageous candidate to provide a versatile redox partner for broad range of cytochromes P450 and of substrate classes.

Author contributions

M.M. carried out the experiments, analyzed and interpreted the data and drafted the manuscript. A.G. and J.N. participated in the establishment of the experiments. F.H. and R.B. designed the project, analyzed and interpreted the results, and assisted in drafting the manuscript.

Conflict of interest

The authors declare that they have no competing interest.

Acknowledgments

The authors would like to thank Birgit Heider-Lips for the purification of AdR and Adx, Dr. Elisa Brill for purification of Fdx2, Tanja Sagadin for purification of Arh1. Also we thank Dr. Azzam Mosa and Dr. Flora Marta Kiss for purification of the bovine CYP21A2 and CYP106A1, respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jbiotec.2016.05.035>.

References

- Aigrain, L., Pompon, D., Truan, G., 2011. Role of the interface between the FMN and FAD domains in the control of redox potential and electronic transfer of NADPH-cytochrome P450 reductase. *Biochem. J.* 435, 197–206.
- Arase, M., Waterman, M.R., Kagawa, N., 2006. Purification and characterization of bovine steroid 21-hydroxylase (P450c21) efficiently expressed in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 344, 400–405.
- Barnes, H.J., Arlotto, M.P., Waterman, M.R., 1991. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 5597–5601.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203.
- Bernhardt, R., Kraft, R., Otto, A., Ruckpaul, K., 1988. Electrostatic interactions between cytochrome P-450 LM2 and NADPH-cytochrome P-450 reductase. *Biomed. Biochim. Acta* 47, 581–592.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145.
- Brandt, M.E., Vickery, L.E., 1993. Charge pair interactions stabilizing ferredoxin-ferredoxin reductase complexes. Identification by complementary site-specific mutations. *J. Biol. Chem.* 268, 17126–17130.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2013. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1703–1717.
- Brixius-Anderko, S., Schiffer, L., Hannemann, F., Janocha, B., Bernhardt, R., 2015. A CYP21A2 based whole-cell system in *Escherichia coli* for the biotechnological production of premedrol. *Microb. Cell Fact.* 14, 135, <http://dx.doi.org/10.1186/s12934-015-0333-2>.
- Bunk, B., Schulz, A., Stammen, S., Münch, R., Warren, M.J., Rohde, M., Jahn, D., Biedendieck, R., 2010. A short story about a big magic bug. *Bioeng. Bugs* 1, 85–91.
- Bureik, M., Schiffer, B., Hiraoka, Y., Vogel, F., Bernhardt, R., 2002. Functional expression of human mitochondrial CYP11B2 in fission yeast and identification of a new internal electron transfer protein etp1. *Biochemistry (Mosc.)* 41, 2311–2321.
- Christner, J.A., Münck, E., Janick, P.A., Siegel, L.M., 1981. Mössbauer spectroscopic studies of *Escherichia coli* sulfite reductase: evidence for coupling between the siroheme and Fe4S4 cluster prosthetic groups. *J. Biol. Chem.* 256, 2098–2101.
- Davydov, D.R., Sineva, E.V., Sistla, S., Davydova, N.Y., Frank, D.J., Sligar, S.G., Halpert, J.R., 2010. Electron transfer in the complex of membrane-bound human cytochrome P450 3A4 with the flavin domain of P450BM-3: the effect of oligomerization of the heme protein and intermittent modulation of the spin equilibrium. *Biochim. Biophys. Acta* 1797, 378.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukanichuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Ewen, K.M., Schiffer, B., Uhlmann-Schiffer, H., Bernhardt, R., Hannemann, F., 2008. The endogenous adrenodoxin reductase-like flavoprotein arh1 supports heterologous cytochrome P450-dependent substrate conversions in *Schizosaccharomyces pombe*. *FEMS Yeast Res.* 8, 432–441.
- Fairhead, M., Giannini, S., Gillam, E.M.J., Gilardi, G., 2005. Functional characterisation of an engineered multidomain human P450 2E1 by molecular Lego. *J. Biol. Inorg. Chem.* 10, 842–853.
- French, J.S., Coon, M.J., 1979. Properties of NADPH-cytochrome P-450 reductase purified from rabbit liver microsomes. *Arch. Biochem. Biophys.* 195, 565–577.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Fact.* 14, 107, <http://dx.doi.org/10.1186/s12934-015-0300-y>.
- Girhard, M., Klaus, T., Khatri, Y., Bernhardt, R., Urlacher, V.B., 2010. Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* 87, 595–607.
- Girhard, M., Tieves, F., Weber, E., Smit, M.S., Urlacher, V.B., 2012. Cytochrome P450 reductase from *Candida apicola*: versatile redox partner for bacterial P450s. *Appl. Microbiol. Biotechnol.* 97, 1625–1635.
- Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* 1, 2876–2890.
- Guengerich, F.P., Martin, M.V., Sohl, C.D., Cheng, Q., 2009. Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nat. Protoc.* 4, 1245–1251.
- Guillouard, I., Auger, S., Hullo, M.-F., Chetouani, F., Danchin, A., Martin-Verstraete, I., 2002. Identification of *Bacillus subtilis* CysL, a regulator of the cysJI operon, which encodes sulfite reductase. *J. Bacteriol.* 184, 4681–4689.
- Hannemann, F., Virus, C., Bernhardt, R., 2006. Design of an *Escherichia coli* system for whole cell mediated steroid synthesis and molecular evolution of steroid hydroxylases. *J. Biotechnol.* 124, 172–181.
- Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., 2007. Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta* 1770, 330–344 (Subj. P450).
- Hebeda, R.E., Styrilund, C.R., Teague, W.M., 1988. Benefits of *Bacillus megaterium* amylase in dextrose production. *Starch*, 33–36.
- Hiwatashi, A., Ichikawa, Y., Maruya, N., Yamano, T., Aki, K., 1976. Properties of crystalline reduced nicotinamide adenine-dinucleotide phosphate-adrenodoxin reductase from bovine adrenocortical mitochondria. I. Physicochemical properties of holo-nadph-adrenodoxin and apo-nadph-adrenodoxin reductase and interaction between non-heme iron proteins and reductase. *Biochemistry (Mosc.)* 15, 3082–3090.
- Horecker, B.L., 1950. Triphosphopyridine nucleotide-cytochrome c reductase in liver. *J. Biol. Chem.* 183, 593–605.
- Khatri, Y., Girhard, M., Romankiewicz, A., Ringle, M., Hannemann, F., Urlacher, V.B., Hutter, M.C., Bernhardt, R., 2010. Regioselective hydroxylation of norisoprenoids by CYP109D1 from *Sorangium cellulosum* So ce56. *Appl. Microbiol. Biotechnol.* 88, 485–495.
- Kiss, F.M., Schmitz, D., Zapp, J., Dier, T.K.F., Volmer, D.A., Bernhardt, R., 2015. Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium*—identification of a novel 11-oxidase activity. *Appl. Microbiol. Biotechnol.* 99, 8495–8514.
- Lambeth, J.D., Seybert, D.W., Kamin, H., 1979. Ionic effects on adrenal steroidogenic electron transport. The role of adrenodoxin as an electron shuttle. *J. Biol. Chem.* 254, 7255–7264.
- Lee, G.-Y., Kim, H.M., Ma, S.H., Park, S.H., Joung, Y.H., Yun, C.-H., 2014. Heterologous expression and functional characterization of the NADPH-cytochrome P450 reductase from *Capsicum annuum*. *Plant Physiol. Biochem.* 82, 116–122.
- Lee, G.-Y., Kim, D.-H., Kim, D., Ahn, T., Yun, C.-H., 2015. Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Arch. Pharm. Res.* 38, 98–107.
- Ly, T.T.B., Khatri, Y., Zapp, J., Hutter, M.C., Bernhardt, R., 2012. CYP264B1 from *Sorangium cellulosum* So ce56: a fascinating norisoprenoid and sesquiterpene hydroxylase. *Appl. Microbiol. Biotechnol.* 95, 123–133.
- Martin, L., Prieto, M.A., Cortés, E., García, J., 1995. Cloning and sequencing of the *pac* gene encoding the penicillin G acylase of *Bacillus megaterium* ATCC 14945. *FEMS Microbiol. Lett.* 125, 287–292.
- Möller, S., Croning, M.D.R., Apweiler, R., 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646–653.
- Morita, M., Tomita, K., Ishizawa, M., Takagi, K., Kawamura, F., Takahashi, H., Morino, T., 1999. Cloning of oxetanocin a biosynthetic and resistance genes that reside on a plasmid of *Bacillus megaterium* Strain NK84-0128. *Biosci. Biotechnol. Biochem.* 63, 563–566.
- Müller, J.J., Lapko, A., Bourenkov, G., Ruckpaul, K., Heinemann, U., 2001. Adrenodoxin reductase-adrenodoxin complex structure suggests electron transfer path in steroid biosynthesis. *J. Biol. Chem.* 276, 2786–2789.
- Murataliev, M.B., Feyereisen, R., Walker, F.A., 2004. Electron transfer by diflavin reductases. *Biochem. Biophys. Acta* 1698, 1–26.
- Nagao, T., Mitamura, T., Wang, X.H., Negoro, S., Yomo, T., Urabe, I., Okada, H., 1992. Cloning, nucleotide sequences, and enzymatic properties of glucose dehydrogenase isozymes from *Bacillus megaterium* IAM1030. *J. Bacteriol.* 174, 5013–5020.
- Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C., Nebert, D.W., 1996. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6, 1–42.
- Neunzig, J., Sánchez-Guijo, A., Mosa, A., Hartmann, M.F., Geyer, J., Wudy, S.A., Bernhardt, R., 2014. A steroidogenic pathway for sulfonated steroids: the metabolism of pregnenolone sulfate. *J. Steroid Biochem. Mol. Biol.* 144, 324–333 (Part B).
- Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura, T., 1998. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694–1699.
- Nisimoto, Y., 1986. Localization of cytochrome c-binding domain on NADPH-cytochrome P-450 reductase. *J. Biol. Chem.* 261, 14232–14239.
- Olteanu, H., Banerjee, R., 2001. Human methionine synthase reductase, a soluble P-450 reductase-like dual flavoprotein, is sufficient for NADPH-dependent methionine synthase activation. *J. Biol. Chem.* 276, 35558–35563.
- Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. II. solubilization purification, and properties. *J. Biol. Chem.* 239, 2379–2385.

- Ostrowski, J., Barber, M.J., Rueger, D.C., Miller, B.E., Siegel, L.M., Kredich, N.M., 1989. Characterization of the flavoprotein moieties of NADPH-sulfite reductase from *Salmonella typhimurium* and *Escherichia coli*. Physicochemical and catalytic properties, amino acid sequence deduced from DNA sequence of *cysJ*, and comparison with NADPH-cytochrome P-450 reductase. *J. Biol. Chem.* 264, 15796–15808.
- Paine, M.J.I., Garner, A.P., Powell, D., Sibbald, J., Sales, M., Pratt, N., Smith, T., Tew, D.G., Wolf, C.R., 2000. Cloning and characterization of a novel human dual flavin reductase. *J. Biol. Chem.* 275, 1471–1478.
- Pandey, A.V., Flück, C.E., 2013. NADPH P450 oxidoreductase Structure, function, and pathology of diseases. *Pharmacol. Ther.* 138, 229–254.
- Park, H.-G., Lim, Y.-R., Eun, C.-Y., Han, S., Han, J.-S., Cho, K.S., Chun, Y.-J., Kim, D., 2010. *Candida albicans* NADPH-P450 reductase: expression, purification, and characterization of recombinant protein. *Biochem. Biophys. Res. Commun.* 396, 534–538.
- Porter, T.D., Kasper, C.B., 1986. NADPH-cytochrome P-450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. *Biochemistry (Mosc.)* 25, 1682–1687.
- Rana, S., Lattoo, S.K., Dhar, N., Razdan, S., Bhat, W.W., Dhar, R.S., Vishwakarma, R., 2013. NADPH-Cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from *Withania somnifera* (L.) dunal. *PLoS One* 8, <http://dx.doi.org/10.1371/journal.pone.0057068>.
- Raux, E., Lanois, A., Warren, M.J., Rambach, A., Thermes, C., 1998. Cobalamin (vitamin B12) biosynthesis: identification and characterization of a *Bacillus megaterium* *cobI* operon. *Biochem. J.* 335, 159–166.
- Ringle, M., Khatri, Y., Zapp, J., Hannemann, F., Bernhardt, R., 2013. Application of a new versatile electron transfer system for cytochrome P450-based *Escherichia coli* whole-cell bioconversions. *Appl. Microbiol. Biotechnol.* 97, 7741–7754.
- Sadeghi, S.J., Gilardi, G., 2013. Chimeric P450 enzymes: activity of artificial redox fusions driven by different reductases for biotechnological applications. *Biotechnol. Appl. Biochem.* 60, 102–110.
- Sagara, Y., Wada, A., Takata, Y., Waterman, M.R., Sekimizu, K., Horiuchi, T., 1993. Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. *Biol. Pharm. Bull.* 16, 627–630.
- Serizawa, N., Matsuoka, T., 1991. A two component-type cytochrome P-450 monooxygenase system in a prokaryote that catalyzes hydroxylation of ML-236 B to pravastatin, a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Biochim. Biophys. Acta: BBA* 1084, 35–40.
- Sevrioukova, I.F., Poulos, T.L., 2010. Arginines 65 and 310 in putidaredoxin reductase are critical for interaction with putidaredoxin. *Biochemistry (Mosc.)* 49, 5160–5166.
- Shen, A.L., Kasper, C.B., 1995. Role of acidic residues in the interaction of NADPH-cytochrome P450 oxidoreductase with cytochrome P450 and cytochrome c. *J. Biol. Chem.* 270, 27475–27480.
- Shivakumar, A.G., Katz, L., Cohen, L.B., Ginsburgh, C.L., Paul, S.L., Vanags, R.I., 1998. Expression of heterologous proteins in *Bacillus megaterium* utilizing sporulation promoters of *Bacillus subtilis*. US5726042 A. USA : Abbott Laboratories.
- Stuehr, D.J., 1997. Structure-function aspects in the nitric oxide synthases. *Annu. Rev. Pharmacol. Toxicol.* 37, 339–359.
- Takasaki, Y., 1989. Novel maltose-producing amylase from *Bacillus megaterium* G-2. *Agric. Biol. Chem.* 53, 341–347.
- Uhlmann, H., Beckert, V., Schwarz, D., Bernhardt, R., 1992. expression of bovine adrenodoxin in *Escherichia coli* and site-directed mutagenesis of (2-Fe-2s) cluster ligands. *Biochem. Biophys. Res. Commun.* 188, 1131–1138.
- Vary, P.S., 1994. Prime time for *Bacillus megaterium*. *Microbiology* 140, 1001–1013.
- Vermilion, J.L., Coon, M.J., 1978. Purified liver microsomal NADPH-cytochrome P-450 reductase: spectral characterization of oxidation-reduction states. *J. Biol. Chem.* 253, 2694–2704.
- Wang, M., Roberts, D.L., Paschke, R., Shea, T.M., Masters, B.S.S., Kim, J.-J.P., 1997. Three-dimensional structure of NADPH-cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U. S. A.* 94, 8411–8416.
- Zeghouf, M., Defaye, G., Fontecave, M., Coves, J., 1998. The flavoprotein component of the *Escherichia coli* sulfite reductase can act as a cytochrome P450c17 reductase. *Biochem. Biophys. Res. Commun.* 246, 602–605.

Supplementary material

A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase

Mohammed Milhim, Adrian Gerber, Jens Neunzig, Frank Hannemann and Rita Bernhardt^{*}

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

*Address correspondence to:

Prof. Dr. Rita Bernhardt, Institute of Biochemistry, Campus B 2.2, Saarland University, 66123 Saarbrücken (Germany)

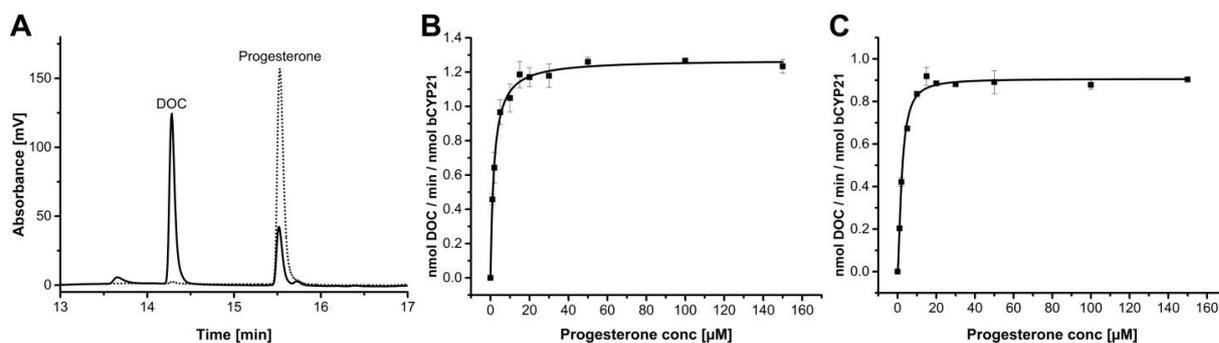
E-mail: ritabern@mx.uni-saarland.de

Phone: +49 681 302 4241

Fax: +49 681 302 4739

M Q L K V V N S P F N Q E Q A D L L N R L L P T L T E A Q K 90
 TTGCAACTTAAGGTAGTAAACAGCCCTTTTAAATCAAGAACAAGCAGATTGCTTAATCGCCTTCTGCCGACATTAACAGAAGCACAAAAA
 M W L S G Y L T A A Q S T S A E G T P D V S T A A P A Q T K 180
 ATGTGGTTGAGCGGTTATTTAACAGCAGCTCAATCTACGCTGCGGAAGGAACGCCAGAGCTTTCTACAGCAGCGCTGCTCAAAACGAAA
 Q T V S K D V T I L Y G S Q T G N A Q G L A E N T G K T L E 270
 CAGACAGTTTCAAAGACGTAACAATCTTTATGGATCACAAACAGGAAATGCTCAAGGACTTGTGAAAATACAGGCAAAACGCTTGAA
 A K G F N V T V S S M N D F K P N N L K K L E N L L I V V S 360
 GCAAAAGGCTTTAATGTAACGTATCTTCATGAATGATTTCAAACCAAATAATTTAAAGAAACTTGAAAATTTATTAATGTGCGTAAGT
 T H G E G E P P D N A L S F H E F L H G R R A P K L E N F R 450
 ACACATGGAGAAGGAGAGCCGCTGATAATGCGCTATCTTTCCATGAATTTCTTACGGCCGTCGAGCGCAAAACTTGAGAACTTCCGT
 F S V L S L G D S S Y E F F C Q T G K E F D V R L A E L G G 540
 TTCCTGTTTTATCGCTTGGAGACAGCTCATACGAATTTTCTGTCAAACAGGAAAAGAATTTGATGTGCGCTTAGCAGAACTTGGCGGT
 E R L Y P R V D C D L D F E E P A N K W L K G V I D G L S E 630
 GAAAGACTGTATCCCGCGCTTGACTGTGATTTAGATTTTGAAGAGCCCGCAAATAAATGGCTTAAAGGTGTTATTGACGGATTAAAGCGAA
 A K G H S A S A A V P A E A P A G T S P Y S R T N P F K A E 720
 CGGAAAGGACACAGCGCTTCGGCAGCTGTTCAGCGGAAGCTCTGCAGGAACTTACCATATTCAGAACAACCCCTTTTAAAGCAGAA
 V L E N L N L N G R G S N K E T R H L E L S L E G S G L T Y 810
 GTGCTTGAGAACTTAAACTTAAACGGCCGCGGATCAAATAAAGAAACGCGACTTTAGAACTATCTCTAGAAGGTTCAAGTTTGGACGTAT
 E P G D S L G I Y P E N D P E L V D L L L N E F K W D A S E 900
 GAACCAGGAGACAGTTTAGTATTTATCTGAAAATGACCCGGAGCTTGTGATCTTCTTAAACGAATCAAGTGGATGCAAGTAA
 S V T V N K E G E T R P L R E A L I S N F E I T V L T K P L 990
 AGTGTAACGGTTAATAAAGAAGGAGAAACGCGTCTCTTAGAGAAGCGCTAATCTCTAATTTTGAATTTACCGTTTAAACAAAGCCGCTT
 L K Q A A E L T G N D K L K A L V E N R E E L K A Y T Q G R 1080
 TTAAGCAAGCAGCTGAGCTTACTGGAAACGATAAGTTAAAAGCGCTTGTAGAAAATCGCGAGGAATTAAGAGCATATACACAAGCCGCT
 D V I D L V R D F G P W N V S A Q E F V A I L R K M P A R L 1170
 GATGTAATTGACTTAGTTCGTGACTTCGGTCCATGGAACGTATCGGCACAAGAGTTTGTAGCCATTTTACGCAAAATGCCAGCGCCCTT
 Y S I A S S L S A N P D E V H L T I G A V R Y E A H G R E R 1260
 TACTCGATTGCAAGCAGCTTATCAGCAAACCTGATGAAGTTTATCTAACAATCGGAGCGGTACGCTACGAAGCGCATGGACGCGAGCGT
 K G V C S V L C S E R L Q P G D T I P V Y L Q S N K N F K L 1350
 AAGGGTGTGTTTCAGTCCTTTGTTTCAAGCGTTTGCAGCCAGGCGATACGATTCTGTATACCTTCAAAGCAATAAAAACTTTAAGCTT
 P Q D Q E T P I I M V G P G T G V A P F R S F M Q E R E E T 1440
 CCTCAAGATCAGGAAACGCCAATTATATGGTAGGACCTGGTACAGGTGTGGCTCCGTTCCGCTCATTATGCAAGAGCGTGAAGAAACA
 G A K G K S W M F F G D Q H F V T D F L Y Q T E W Q K W L K 1530
 GGGCAAAAGGAAAGTCATGGATGTTCTTTGGAGATCAGCACCTTCGTAACAGACTTCTTTTACCAAAACAGAATGCAAAAGTGGTTAAAA
 D G V L T K M D V A F S R D T E E K V Y V Q N R M L E H S K 1620
 GATGGCGTACTGACAAAAATGGACGTGGCGTTTTTACGCGATACAGAAGAAAAGTATACGTACAAAACCGTATGCTCGAACATAGTAAA
 E L F Q W L E E G A S F Y V C G D K T N M A R D V H N T L V 1710
 GAATATTTCCAATGGTTAGAAGAAGCGCATCTTTTATGTGTGCGGAGATAAAACAAATATGGCACGCGACGTGCACAACACGCTAGTT
 E I I E T E G K M S R E Q A E G Y L A E M K K Q K R Y Q R D 1800
 GAAATTATCGAAACAGAAGGCAAGATGAGCCGTGAACAGCGGAAGGTTACCTTGCTGAAATGAAGAAAACAAAACGTTATCAGCGTGAT
 V Y
 GTATACTGA 1809

Supplementary Fig. S1. The open reading frame sequence of the BmCPR. The upper and the lower lines represent the deduced amino acid and nucleotide sequences, respectively. The one-letter code for each amino acid is aligned with the second nucleotide of each codon. The start codon TTG is underlined.



Supplementary Fig. S2. *In vitro* activity of the BmCPR. (A) *In vitro* conversion of 100 μ M progesterone to deoxycorticosterone (DOC) using bovine CYP21A2 and BmCPR at 37 $^{\circ}$ C for 1 h with NADPH (continuous chromatogram) or NADH (dashed chromatogram) as electron donor. (B) and (C) kinetic parameters for the conversion of progesterone catalyzed by CYP21A2 using BmCPR or bovine CPR as redox partner, respectively. The product formation was determined by HPLC and each value is represented as mean \pm SD of three separate measurements. Data was fitted using originPro 9.0G software.

2.2 Milhim et al., 2016b

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin

Mohammed Milhim, Natalia Putkaradze, Ammar Abdumughni, Fredy Kern, Philip Hartz and Rita Bernhardt

Journal of Biotechnology, 2016 Dec. 240: 68–75.

DOI: [10.1016/j.jbiotec.2016.11.002](https://doi.org/10.1016/j.jbiotec.2016.11.002)

Reprinted with permission of the Journal of Biotechnology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Biotechnology

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

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin



Mohammed Milhim, Natalia Putkaradze, Ammar Abdulmughni, Fredy Kern, Philip Hartz, Rita Bernhardt*

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 2 September 2016

Received in revised form 28 October 2016

Accepted 1 November 2016

Available online 2 November 2016

Keywords:

Bacillus megaterium

CYP107 family

Bacterial cytochrome P450

Mevastatin

Pravastatin

BmCPR

Fdx2

ABSTRACT

In the current work, we describe the identification and characterization of the first plasmid-encoded P450 (CYP107DY1) from a *Bacillus* species. The recombinant CYP107DY1 exhibits characteristic P450 absolute and reduced CO-bound difference spectra. Reconstitution with different redox systems revealed the autologous one, consisting of BmCPR and Fdx2, as the most effective one. Screening of a library of 18 pharmaceutically relevant compounds displayed activity towards mevastatin to produce pravastatin. Pravastatin is an important therapeutic drug to treat hypercholesterolemia, which was described to be produced by oxyfunctionalization of mevastatin (compactin) by members of CYP105 family. The hydroxylation at C6 of mevastatin was also suggested by docking this compound into a computer model created for CYP107DY1. Moreover, in view of the biotechnological application, CYP107DY1 as well as its redox partners (BmCPR and Fdx2) were successfully utilized to establish an *E. coli* based whole-cell system for an efficient biotransformation of mevastatin. The in vitro and in vivo application of the CYP07DY1 also offers the possibility for the screening of more substrates, which could open up further biotechnological usage of this enzyme.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Selective oxyfunctionalization of nonactivated carbon-hydrogen bonds represents a challenge in the industrial field. In general, the use of chemical methods has many disadvantages such as hazardous conditions, cost-efficiency and the lack of chemo-, stereo-, and regioselectivity. Therefore, in the last years many efforts have been carried out in the search for selective and efficient enzymatic systems that are able to incorporate oxygen into nonactivated carbon-hydrogen bonds. Cytochromes P450 (P450s) [E.C.1.14.-] are heme-iron containing enzymes that catalyze the monooxygenation of various nonactivated hydrocarbons with high regio-, stereo- and enantioselectivity including the biosynthesis of hormones, signal molecules, defense-related chemicals and secondary metabolites in addition to their central role in the metabolism of endogenous (steroids and fatty acids) and exogenous (drugs and toxins) substances (Bernhardt, 2006;

Bernhardt and Urlacher, 2014). They are found in all kingdoms of life including mammals, plants, insects, fungi, archaea, and bacteria (except *E. coli*) as well as in viruses. In eukaryotes, they are mostly integral membrane-bound proteins, whereas prokaryotic P450 systems are more likely soluble and located in the cytoplasm. P450s rely for their activities on redox partners. Based on the composition of the redox partner involved in the transfer of electrons, the P450 systems can be categorized into different classes (Hannemann et al., 2007) of which the most researched ones are the classes I and II. Class I contains the bacterial and the eukaryotic mitochondrial P450s, which obtain electrons from NADPH using two proteins, a flavin adenine dinucleotide (FAD)-containing ferredoxin reductase and an iron-sulfur containing ferredoxin. Class II comprises the eukaryotic microsomal P450s, which obtain electrons from NADPH via a FAD and FMN-containing P450 reductase.

The numbers of the newly identified P450s increased drastically over the past few years (Nelson, 2009), but there is nevertheless a still growing demand to exploit novel P450s as a valuable biocatalyst in the industrial field.

Bacillus megaterium is a nonpathogenic, aerobic, Gram-positive rod-shaped bacterium. Due to its high protein production capacity, plasmid stability and the ability to take up a variety of hydrophobic

* Corresponding author at: Institute of Biochemistry, Campus B2.2, Saarland University, D-66123 Saarbrücken, Germany.

E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

substrates, *B. megaterium* gained throughout the last decades a lot of interest in the industrial field for the production of biotechnologically relevant substances (Bunk et al., 2010). The publication of the complete genome sequence of the *B. megaterium* strains QM B1551 and DSM319 in 2011 (Eppinger et al., 2011) enabled the identification of new proteins. Among them are cytochromes P450 and a NADPH dependent diflavin reductase, which have been shown to be very important for biotechnological and pharmaceutical applications (Brill et al., 2013; Milhim et al., 2016). *B. megaterium* encodes for several P450s. The self-sufficient CYP102A1 (also known as BM3) is the most investigated bacterial P450 so far, which has been used and redesigned to catalyze the oxidation of a variety of biotechnologically interesting substances (Whitehouse et al., 2012). In addition, the biotechnologically valuable CYP106 family, CYP106A1 from *B. megaterium* strain DSM319 (Brill et al., 2013) and CYP106A2 from *B. megaterium* strain ATCC 13368 (Berg et al., 1976, 1979), was characterized to be associated with the biotransformation of a diverse array of substrates such as steroids and terpenoid substances (Brill et al., 2013; Schmitz et al., 2012). Furthermore, CYP109E1 was recently identified from *B. megaterium* strain DSM319 as steroid hydroxylase (Jóźwik et al., 2016).

In this study, we report the identification and characterization of a new plasmid-encoded P450 from the *B. megaterium* QM B1551. The bioinformatic analysis of the new P450 showed that it belongs to the CYP107 family. It was successfully cloned and expressed in *E. coli*. Screening of a potential substrates showed that CYP107DY1 possesses hydroxylation activity towards mevastatin.

2. Materials and methods

2.1. Strains, expression vectors, enzymes, and chemicals

E. coli TOP10 from Invitrogen (Karlsruhe, Germany) was used for cloning experiments. *E. coli* C43 (DE3) and the expression vector pET17b, both from Novagen (Darmstadt, Germany), were used for recombinant gene expression. Substrates were obtained from TCI (Eschborn, Germany). Pravastatin lactone was from Santa Cruz Biotechnology (Heidelberg, Germany). All other chemicals were purchased from Sigma–Aldrich (Schnelldorf, Germany).

2.2. Cloning of the gene encoding CYP107DY1

For protein purification purposes, the DNA fragment encoding the full length CYP107DY1 (Supplementary Fig. S1) was synthesized (Genart, Regensburg, Germany) and cloned into the expression vector pET17b with the *Nde*I/*Kpn*I restriction sites. For purification with IMAC, the 3' end of the gene was extended with a sequence coding for six histidines. Plasmid was verified by sequencing.

2.3. Heterologous gene expression and purification of CYP107DY1, reductases and ferredoxins

For heterologous gene expression, *E. coli* C43 (DE3) cells were co-transformed with the expression vector pET17b congaing the sequence for CYP107DY1 and the chaperone GroEL/GroES-encoding plasmid pGro12, which has a kanamycin resistance gene (Brixius-Anderko et al., 2015; Nishihara et al., 1998). Cultures were grown at 37 °C to an optical density of 0.6 in 200 ml TB medium containing the suitable antibiotics. The expression of the protein was induced by adding 1 mM IPTG, the synthesis of heme was enhanced by addition of 1 mM heme precursor δ -ALA. The cells were grown at 28 °C and 180 rpm for 24 h.

For purification, cell pellets were sonicated in 50 ml lysis buffer (50 mM potassium phosphate pH 7.4, 20% glycerol, 0.1 mM DTE, 500 mM sodium acetate, and 0.1 mM PMSF). After centrifugation

at 30,000g for 30 min at 4 °C, the supernatant was applied on a Ni–NTA agarose column equilibrated with lysis buffer. The column was washed with 100 ml equilibration buffer supplemented with 40 mM imidazole followed by 20 ml elution buffer supplemented with 200 mM imidazole (50 mM potassium phosphate pH 7.4, 20% glycerol, 0.1 mM DTE, and 0.1 mM PMSF). The eluted protein was dialyzed against elution buffer without imidazole, concentrated and stored at –80 °C.

The *B. megaterium* redox system BmCPR and Fdx2 were purified as reported previously (Brill et al., 2013; Milhim et al., 2016). The purification of the redox system from the fission yeast *Schizosaccharomyces pombe* Arh1 and Etp1^{fd} was carried out as described before (Bureik et al., 2002; Ewen et al., 2008). Recombinant bovine AdR and the Adx_{4–108} (truncated form of Adx comprising amino acids 4–108) were purified as mentioned elsewhere (Sagara et al., 1993; Uhlmann et al., 1992).

The concentration of recombinant P450 was estimated using the CO-difference spectral assay as described previously with $\epsilon_{450–490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ (Omura and Sato, 1964). The concentration of BmCPR was quantified by measuring the flavin absorbance at 456 nm with $\epsilon_{456} = 21 \text{ mM}^{-1} \text{ cm}^{-1}$ for the oxidized enzyme (Milhim et al., 2016). The concentrations of the AdR and Arh1 were measured using the extinction coefficient $\epsilon_{450} = 11.3 \text{ mM}^{-1} \text{ cm}^{-1}$ (Ewen et al., 2008; Hiwatashi et al., 1976). The concentrations of Fdx2 and Etp1^{fd} were measured using the extinction coefficient $\epsilon_{390} = 6.671 \text{ mM}^{-1} \text{ cm}^{-1}$ and $\epsilon_{414} = 9.8 \text{ mM}^{-1} \text{ cm}^{-1}$, respectively (Brill et al., 2013; Schiffler et al., 2004).

2.4. Investigation of electron transfer partners

The functional interaction of the electron transfer partners for a particular P450 can be determined by recording the NADPH reduced CO-complex peak at 450 nm when P450 was coupled with the different ferredoxins/ferredoxin reductases in the absence of substrate. For this, CYP107DY1 was mixed with ferredoxins (Fdx2, Etp1^{fd} or Adx_{4–108}) and ferredoxin reductases (BmCPR, Arh1 or AdR) with ratios of 1:40:5 μM [CYP107DY1:ferredoxin:ferredoxin reductase] in 50 mM HEPES buffer pH 7.4 and NADPH was added to a final concentration of 1 mM. The spectrum of NADPH-reduced samples was recorded after bubbling the sample with carbon monoxide (CO) gas. The reduction efficiency of the redox partners was then evaluated by comparing the peak at 450 nm of the CO-complexed CYP107DY1 reduced with the different redox systems and the peak at 450 nm of the CO-complexed CYP107DY1 reduced with sodium-dithionite.

2.5. In vitro conversion and HPLC analysis

The in vitro conversion of the substrates was carried out with a reconstituted system at 30 °C in conversion buffer (50 mM HEPES, pH 7.4, 20% glycerol). The reconstituted system contained 0.5 μM CYP107DY1, 2.5 μM BmCPR, 20 μM Fdx2, 1 mM MgCl_2 , 5 mM glucose-6-phosphate, 1 U glucose-6-phosphate dehydrogenase for NADPH regeneration and 100 μM substrate. The reaction was started by adding NADPH (200 μM) and stopped after 15 min by the addition of 1 vol ethyl acetate and extracted twice. The organic phase was evaporated under vacuum. Residuum was dissolved in 20% acetonitrile/water mixture and subjected to HPLC analysis. HPLC analysis was performed using a Jasco system. A reversed-phase ec MN Nucleodur C18 (4.0 × 125 mm) column (Macherey-Nagel) was used for all experiments at an oven temperature of 40 °C. Mevastatin and its metabolite pravastatin were eluted from the column using a gradient of acetonitrile from 20 to 100% in water over 20 min. The detection wavelength of mevastatin and its metabolite pravastatin was 236 nm.

2.6. Circular dichroism (CD) spectroscopy

Circular dichroism (CD) spectra were recorded at 30 °C using a JASCO J-715 spectropolarimeter over the wavelength range 190–260 nm for the far-UV region and 300–500 nm for the near-UV/Vis region at a protein concentration of 2 μ M (~0.1 mg/ml) and 20 (~1 mg/ml), respectively, dissolved in 10 mM potassium phosphate buffer pH 7.4 with the following parameters: path-length of 0.1 cm for the far-UV region and 1 cm for the near-UV/Vis region measurement, data pitch of 0.1 nm, band width of 5 nm, accumulation 3 times. Spectra were recorded in triplicate and averaged.

2.7. Homology modeling of CYP107DY1 and molecular docking with mevastatin

Using the homology modeling program Modeller 9.14 (University of California San Francisco, USA), a model of CYP107DY1 was calculated using CYP107RB1 (Vdh) (PDB accession code: 3A4G) as template. The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. No further structural refinement of the model was performed.

The three dimensional structure of the mevastatin molecule was obtained from PubChem (Kim et al., 2016). The docking simulations of the homology model of CYP107DY1 with the mevastatin molecule were carried out using Autodock 4.0 (Morris et al., 2009). The Windows version 1.5.6 of Autodock Tools was used to compute Kollman charges for the enzyme and Gasteiger-Marsili charges for the ligand (Sanner, 1999). 200 docking runs were carried out applying the Lamarckian genetic algorithm using default parameter settings.

2.8. In vivo whole-cell biotransformation

For the establishment of a CYP107DY1-dependent whole-cell system, a tricistronic pET17b-based vector, encoding the CYP107DY1-BmCPR-Fdx2 genes, was constructed. CYP107DY1 coding sequence was amplified and cloned via the restriction sites *NdeI/HindIII*. The resulting vector (pET17b.CYP107DY1) served then as a backbone for the cloning of BmCPR, which was amplified via PCR and cloned with the restriction sites *BamHI/NotI*. Fdx2 coding region was PCR amplified and cloned downstream CYP107DY1-BmCPR using the restriction sites *NotI/XhoI*. All resulting vectors were verified by sequencing.

E. coli C43 (DE3) cells were co-transformed with the suitable expression vector based on pET17b, coding for CYP107DY1-BmCPR-Fdx2, and the chaperone GroEL/GroES-coding plasmid pGro12, which has a kanamycin resistance gene (Brixius-Anderko et al., 2015; Nishihara et al., 1998). Transformed cells were grown overnight in 50 ml LB broth medium supplemented with 100 μ g/ml ampicillin and 50 μ g/ml kanamycin at 37 °C and shaking at 180 rpm. For the expression of proteins, 50 ml TB medium containing 100 μ g/ml ampicillin and 50 μ g/ml kanamycin were inoculated (1:100) with the transformed cells and cultivated at 37 °C with rotary shaking at 140 rpm. Protein expression was induced at OD₆₀₀ = 0.6–0.8 with 1 mM IPTG, 4 mg/ml arabinose (for induction of GroES/GroEL expression) and 1 mM δ -ALA as heme precursor. The temperature was then reduced to 28 °C. After incubation for 24 h, cells were harvested by centrifugation (4000g) for 20 min at 4 °C and washed once with 1 vol of conversion buffer (50 mM potassium phosphate buffer (pH 7.4) supplemented with 2% glycerol). After a second centrifugation, the cell pellets were resuspended in conversion buffer to an end cell-suspension concentration of 60g wet cell weight (wcw)/L buffer. The substrate mevastatin was added to a final concentration of 100 μ M and the culture was incubated for the indicated time at 30 °C and 140 rpm

for 20 h. To enable higher substrate conversion, EDTA (20 mM) or polymyxin B (32 μ g/ml) was added to increase permeability and substrate uptake of the *E. coli* cells (Janocha and Bernhardt, 2013; Kern et al., 2016). Substrate was extracted twice with the same volume of ethyl acetate and the organic phase was evaporated using a rotary evaporator. After that, the residues were dissolved in the high performance liquid chromatography (HPLC) mobile phase (20% ACN) and subjected to HPLC analysis using the same method mentioned in Section 2.5.

3. Results and discussion

3.1. Bioinformatic analysis

The publication of the complete genome sequence of the *B. megaterium* strains QM B1551 and DSM319 in 2011 (Eppinger et al., 2011) enabled the characterization of new proteins. Recently, we were able to identify and characterize different P450s and related proteins from the strain DSM319 (Brill et al., 2013; Gerber et al., 2015; Milhim et al., 2016). In contrast to the DSM319 strain, strain QM B1551 harbors seven indigenous plasmids (Eppinger et al., 2011). Sequence analysis of the indigenous plasmids showed that plasmid number 5 (pBM500) has an open reading frame (BMQ_pBM50008) containing 1233 base pairs, which encodes for a protein comprising 410 amino acids with a predicted molecular weight of about 46.742 kDa. The analysis of the BMQ_pBM50008 domains using the Pfam Database (Finn et al., 2016) for highly conserved motifs of the P450 family proved the presence of the heme-binding motif (F-x-x-G-x-x-x-C-x-G), the (A/G-G-x-E/D-T-T/S) motif in the I-helix and the (E-x-x-R) motif in the K-helix (Supplementary Fig. S1), suggesting its identification as a cytochrome P450.

Multiple sequence alignment showed that the protein sequence of BMQ_pBM50008 belongs to the CYP107 family. After submission of the protein sequence of BMQ_pBM50008 to the P450 nomenclature committee (Prof. Dr. David Nelson), it was found to match best to CYP107DA1 with 49% sequence identity and, therefore, was assigned as a new subfamily with the name CYP107DY1. CYP107 is the largest family among bacterial P450s comprising more than 2500 subfamily members (<https://cyped.biocatnet.de/sFam/107>). Besides the CYP105 family, members of the CYP107 family are shown to be the most studied bacterial P450s that participate in the degradation and biotransformation of a broad spectrum of xenobiotics as well as in secondary metabolite biosynthesis, and, therefore, are considered to be important for industrial biotechnology; for example CYP107BR1 from *Pseudomonas autotrophica* for the activation of Vitamin D₃ (Sakaki et al., 2011), CYP107E from *Micromonospora griseorubida* for mycinamicin biosynthesis (Inouye et al., 1994), and P450_{terf} (CYP107L) from *Streptomyces platensis* for the hydroxylation of terfenadine (Lombard et al., 2011).

CYP107DY1 is the first P450 found to be encoded on a bacillus sp. plasmid. There are only few examples of plasmid-encoded P450s described so far in the literature, such as the CYP107A2 (LkmF) and CYP107AP1 (LkmK) from *Streptomyces rochei*, which participate in the biosynthesis of the macrolide antibiotic lankomycin (Arakawa et al., 2006) and CYP102H1 from *Nocardia farcinica* that catalyzes the hydroxylation of linoleic acid (Chung et al., 2012).

The plasmid-encoded nature of the identified P450 is very interesting since the absence of CYP107 members in the genome of *B. megaterium* indicates that the presence of the corresponding gene of the CYP107DY1 on the plasmid may be due to horizontal gene transfer displacement rather than intragenic transfer from the chromosome to the plasmid. This suggests that the P450s can play an important role in adaptation and evolution in prokaryotes.

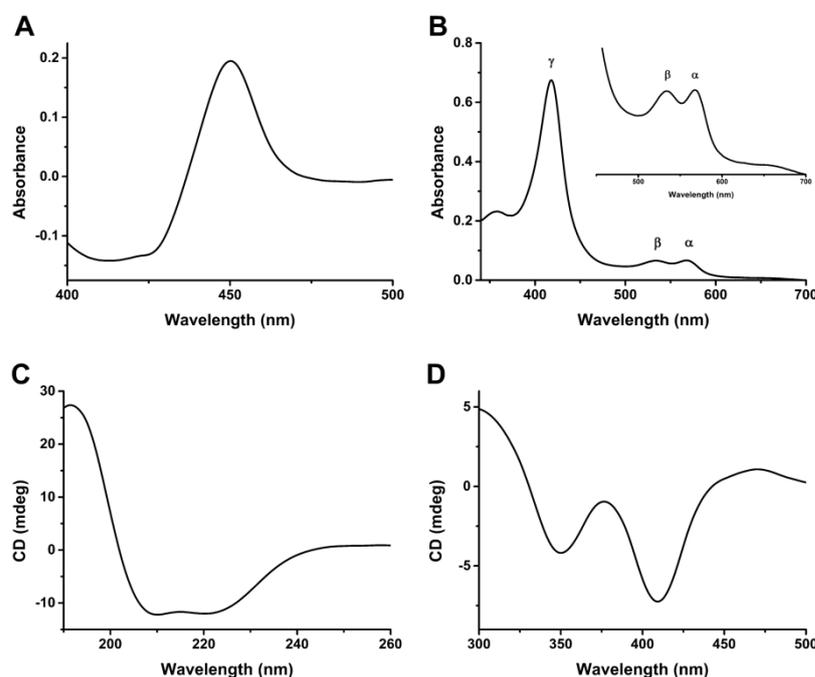


Fig. 1. Spectral characteristics of CYP107DY1. (A) The CO difference spectrum of CYP107DY1. (B) The UV-vis spectral characteristics of the purified CYP107DY1. The inset shows the magnification of the spectrum in the α and β band region. Circular dichroism (CD) spectra in the far-UV (C) and in the near-UV/visible (D) region at a CYP107DY1 concentration of 2 μ M and 20 μ M, respectively, resuspended in 10 mM potassium phosphate buffer pH 7.4. The CD spectrum was recorded using the following parameters: path length 1 mm (for far-UV) and 5 mm (near-UV/visible); time constant 2 s; band pass 5 nm; number of scans 3.

3.2. Expression, purification and spectrophotometric characterization

The DNA fragment encoding the full length CYP107DY1 was cloned into the expression vector pET17b and expressed in *E. coli* C43 (DE3). The expression levels were calculated by measuring reduced CO difference spectrum in the cell lysate (Omura and Sato, 1964). The carbon monoxide bound form gave a typical peak maximum at 450 nm (Fig. 1A). The full-length CYP107DY1 was purified in a soluble form with an expression level of 20 nmol/L. It was previously shown that the co-expression of some P450s with chaperones leads to an improvement of protein folding and thus increases of the expression level (Arase et al., 2006; Brixius-Anderko et al., 2015; Nishihara et al., 1998). The co-expression with the chaperones GroES and GroEL also improved the expression of CYP107DY1 more than 10 times, yielding 210 nmol/L.

Besides the CO-difference spectrum, the UV-vis absorption spectrum provides a simple technique for the characterization of P450 enzymes. The oxidized form of substrate free CYP107DY1 exhibited a major (γ) Soret peak at 417 and the smaller α and β bands at 567 and 535 nm, respectively (Fig. 1B), indicating a low spin state of the heme iron in the P450. In addition, the CD spectra of the oxidized CYP107DY1 were measured in the far-UV- and near UV-vis region. The far-UV CD spectrum showed a negative dichroic double band with minima at 208 and 222 nm (Fig. 1C), an indication of predominantly α -helical secondary structure (Poulos et al., 1986, 1987; Ravichandran et al., 1993). In the near UV-vis region, CYP107DY1 displayed two large negative signals at 350 nm and at 408 nm (Fig. 1D). These are in correspondence with the characteristic peaks for other bacterial P450s (Lepesheva et al., 2001; Munro et al., 1994).

Taken together, the spectrophotometric properties of the purified CYP107DY1 indicate that the enzyme is produced in the active form with proper heme incorporation.

3.3. Searching for a suitable redox partner

The availability of a redox partner/s is essential for studying the functionality of CYP107DY1. By searching the ORFs around the CYP107DY1 coding region as well as the ORFs on the seven indigenous plasmids of the QM B1551 strain, no redox partner (reductase and/or ferredoxin) was identified. Therefore, we tested herein different autologous and heterologous redox systems. As autologous electron transfer partners, the diflavin reductase BmCPR and the ferredoxin Fdx2 of *B. megaterium* DSM319 were selected. The BmCPR-Fdx2 system has been shown to support efficiently the activity of CYP106A1 (Milhim et al., 2016). As heterologous redox partners, the soluble *Schizosaccharomyces pombe* redox system adrenodoxin reductase homologue 1 (Arh1) and its ferredoxin (Etp1^{fd}) as well as the bovine adrenodoxin reductase (AdR) and adrenodoxin (Adx₄₋₁₀₈) were used. The redox systems Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ are described to transfer electrons to different classes of P450s (Brixius-Anderko et al., 2015; Ly et al., 2012).

Based on the measurement of the reduced CO-bound spectrum of CYP107DY1, the redox partners were tested and compared. Using the heterologous redox partners Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ very low peaks (<5%) compared with the dithionite reduced CO difference peak at 450 nm were recovered within 5 min. In contrast, the autologous redox partners BmCPR-Fdx2 were very efficient and able to recover ~99.5% of the peak (Fig. 2). Therefore, the BmCPR-Fdx2 redox system was selected for further investigations with CYP107DY1.

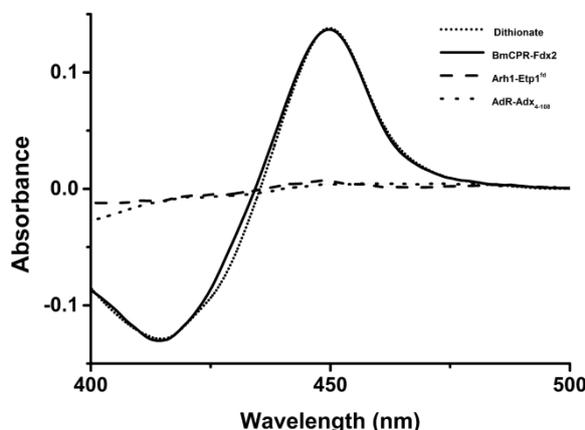


Fig. 2. Determination of CYP107DY1 reduction using autologous and heterologous electron transfer partners. The dithionite reduced CO-difference spectrum (short dot line) of CYP107DY1 was compared with the NADPH reduced BmCPR-Fdx2 (solid line), Arh1-Etp1^{fd} (dash line) and AdR-Adx₄₋₁₀₈ (dot line) CO-complexed spectrum. The NADPH (1 mM) reduced CO-difference spectra were recorded in a 200 μ l mixture of CYP/ferredoxin/reductase with a 1:40:5 molar ratio in 50 mM HEPES buffer pH 7.4 containing 20% glycerol (see Materials and methods).

3.4. In vitro conversion of mevastatin

The identification of suitable substrates and the characterization of the catalytic activity of a new enzyme are significant challenges. Therefore, at first a phylogenetic based approach was used in this study to predict a functional homolog of the new P450. The deduced amino acid sequence of CYP107DY1 was aligned with some related bacterial P450s with known function or substrate/s. The aligned sequences were then used for the construction of a neighbor-joining phylogenetic tree, which shows that CYP107DY1 is clustered with CYP267B1 from *Sorangium cellulosum* So ce56 (Fig. 3). CYP267B1 belongs to the CYP107 family (<http://drnelson.uthsc.edu/Bacteria.html>) and was previously characterized as a versatile drug metabolizer (Kern et al., 2016). Therefore, using an in vitro reconstituted assay we screened the activity of CYP107DY1 towards various pharmaceutical substrates (Supplementary Table 1). Among the 18 tested drugs, CYP107DY1 was found to metabolize mevastatin producing one product at a retention time of 9 min (Fig. 4B) compared with the negative control (Fig. 4A). The conversion ratio of 100 μ M mevastatin within 15 min was ~30%. The product retention time was identical to that of a pravastatin authentic standard (Fig. 4C). The results presented in Fig. 5 showed that CYP107DY1 activity can be supported by all tested redox systems, with an obvious higher efficiency using BmCPR-Fdx2. The conversion ratio of the 100 μ M mevastatin using Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈ as redox partners was ~5% and ~3%, respectively, compared with when using of the BmCPR-Fdx2.

The regioselective oxyfunctionalization of the precursor mevastatin is crucial for the production of pravastatin, the widely used therapeutic agent for hypercholesterolemia (Lamon-Fava, 2013). This was described previously using CYP105A3 (P450sca2) from *Streptomyces carbophilus* (Matsuoka et al., 1989) and mutant CYP105AS1 from *Amycolatopsis orientalis* (McLean et al., 2015). However, CYP107DY1 is the first reported P450 of the CYP107 family, which can highly selectively hydroxylate such statin drugs.

3.5. Homology modeling and docking of CYP107DY1 with mevastatin

For the prediction of the three dimensional structure of CYP107DY1, homology modeling was performed. In order to select

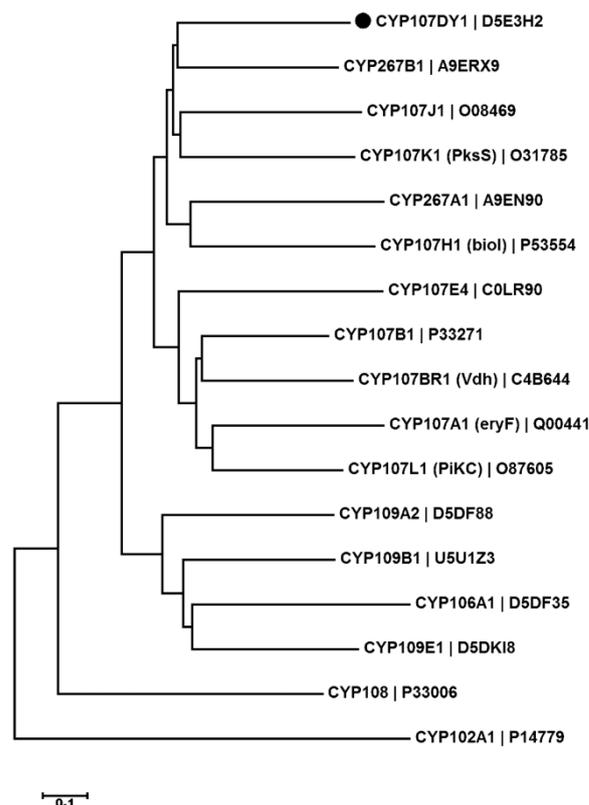


Fig. 3. Evolutionary relationships of CYP107DY1. The alignment was done with 10-gap setting and 0.1-gap extension, with slow alignment input in ClustalW2 server. The tree was constructed by neighbor joining algorithm with bootstrap analysis of 1000 replicates. The scale bar represents 0.1 amino acid substitution per amino acid. The CYP107DY1 is indicated with the closed circle (●). The number next to the gene name represents the UniProtKB accession number.

a suitable template, the amino acid sequence of the CYP107DY1 was submitted to the RCSB protein data bank BLAST server (<http://www.rcsb.org/pdb/search/advSearch.do?search=new>). The CYP107BR1 (Vdh) from *Pseudonocardia autotrophica* (UniProt acc. no.: C4B644) showed a sequence identity of 40% to CYP107DY1. The corresponding crystal structure of CYP107BR1 (Vdh) (PDB acc. code: 3A4G) was, therefore, chosen as a structural template. The CYP107DY1 homology model (Supplementary Fig. S2) has a typical three-dimensional P450 structure consisting of a C-terminus relatively rich in α -helices and an N-terminus relatively rich in β -sheets. Overall, it comprises 13 α -helices designated A, B, Bi and C-L and 12 β -sheets grouped into 5 regions. The prosthetic heme group is embedded in the active site and surrounded by the I-helix from the proximal side and the L-helix from the distal side. The conserved cysteine residue (Cys360) is located in the loop region preceding the L-helix. The key structural features of the CYP107DY1 suggest that its overall topology correlates with the general folding properties of P450s (Supplementary Fig. S2).

Three-dimensional structures of P450s are very helpful to understand the enzyme-substrate interaction. Therefore, mevastatin was docked into CYP107DY1. In the docking simulation the target protein was kept as rigid body. The flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box with a size of 52 $\text{Å} \times 52 \text{Å} \times 52 \text{Å}$ was fixed above the heme moiety to cover the whole active site of the enzyme. As illustrated in Fig. 6, the substrate molecule is located over the

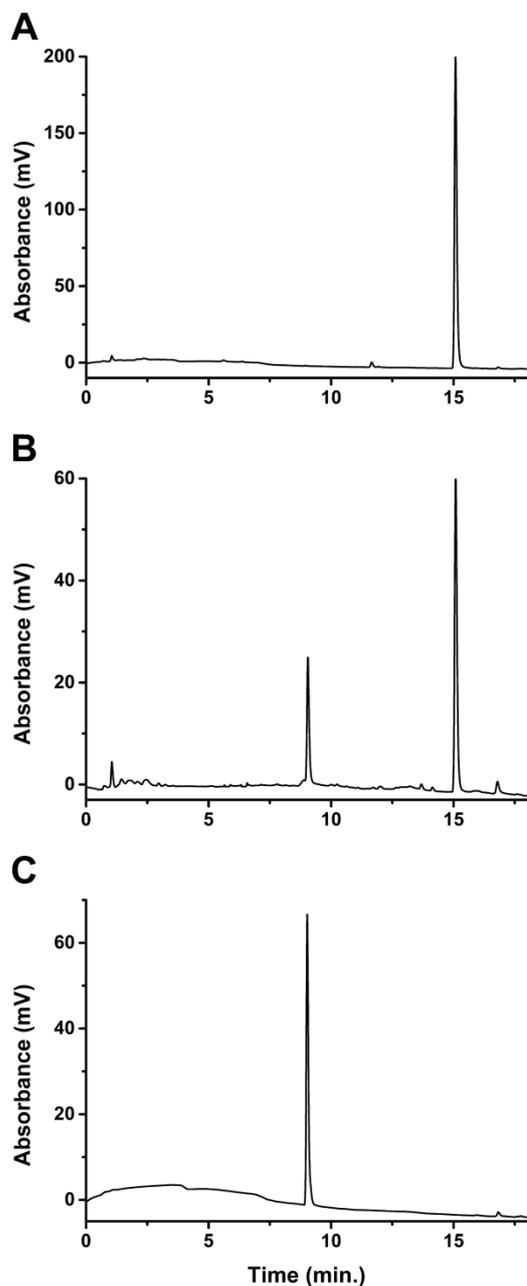


Fig. 4. HPLC analysis of the conversion of mevastatin with CYP107DY1. In vitro conversion of 100 μM mevastatin in the (A) absence of CYP107DY1 or (B) presence of CYP107DY1. The reaction was carried out in 50 mM HEPES buffer pH 7.4 with 20% glycerol. The reconstituted system contained 0.5 μM CYP107DY1, 2.5 μM BmCPR, 20 μM Fdx2 and the NADPH regeneration system. The conversion was carried out for 15 min at 30 °C. (C) Pravastatin authentic standard.

heme. In addition, the side chains of seven amino acid residues (including Leu99, Leu245, Thr249, Ala250, Glu253, Thr254 and Gly300) were found to define a cavity with mevastatin. The keto group of the 2-methylbutanoate side chain and the hydroxyl group on the lactone ring of the mevastatin molecule formed two hydrogen bonds with Thr249 and Glu253, respectively. These interactions enable the orientation of the molecule in a position that allows

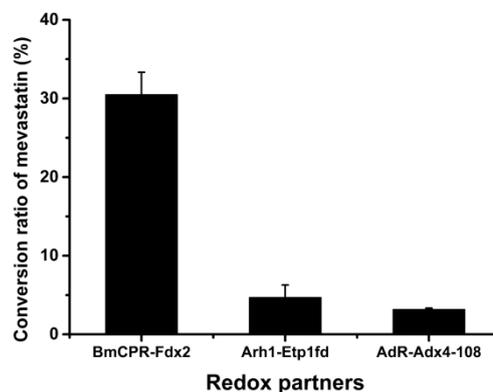


Fig. 5. Effect of the different redox partners on the in vitro conversion of mevastatin by CYP107DY1. In vitro conversion of 100 μM mevastatin by CYP107DY1 using the redox system BmCPR-Fdx2 of, Arh1-Etp1^{fd} and AdR-Adx₄₋₁₀₈. The reaction was carried out in 50 mM HEPES buffer pH 7.4 with 20% glycerol. The reconstituted system contained 0.5 μM CYP107DY1, 2.5 μM reductase (BmCPR, AdR or Arh1), and 20 μM ferredoxin (Fdx2, Etp1^{fd} or Adx₄₋₁₀₈) and the NADPH regeneration system. The reaction was carried out for 15 min at 30 °C. The data is represented as mean \pm SD of three separate measurements.

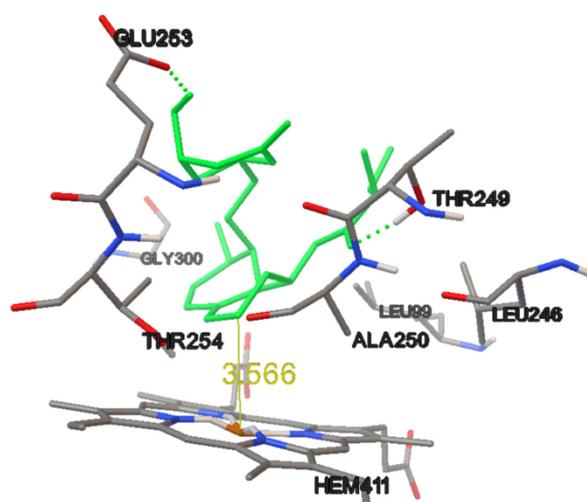
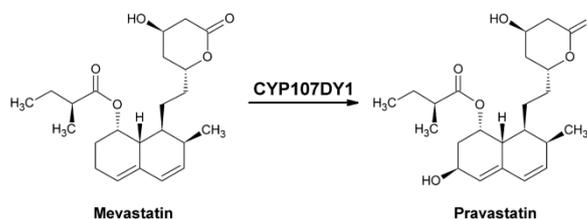


Fig. 6. Docking conformation of mevastatin. The ligand mevastatin is shown in green. The distance of the C-6 atom from the heme iron is shown in yellow in Å. Some of the amino acids that form the active site of the CYP107DY1 in the presence of mevastatin are shown and named. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Scheme 1. Conversion of mevastatin to pravastatin by CYP107DY1.

atom C-6 to face the heme iron (distance \sim 3.5 Å) (Fig. 6), favorable for a hydroxylation at this position to produce the pravastatin (Scheme 1), supporting our experimental data shown in Fig. 4.

Table 1
CYP107DY1-based ^(a) *Escherichia coli* Whole-cell biotransformation yield of mevastatin.

Additives ^(b)	Pravastatin yield (mg/L)
Without additives	13.2 ± 2.5
Polymyxin B	17.9 ± 2.1
EDTA	28.5 ± 3.1

^a *E. coli* C43 (DE3) cells were co-transformed with the expression vector based on pET17b, coding for CYP107DY1-BmCPR-Fdx2, and the chaperone GroEL/GroES-coding plasmid pGro12.

^b Polymyxin B and EDTA were added to a final concentration of 32 µg/ml and 20 mM, respectively.

3.6. Biotransformation of mevastatin using a whole-cell based system

The *in vitro* activity of CYP107DY1 towards mevastatin was very encouraging to address the question whether it is applicable for the *in vivo* production of pravastatin. Therefore, and as a proof of concept, a CYP107DY1 based *E. coli* whole-cell biotransformation system, utilizing BmCPR and Fdx2 as redox partners, was established. We chose *E. coli* as host, because, in contrast to *B. megaterium*, it has the advantage that no endogenous P450 can interfere with the desired reaction. For this purpose, a tricistronic pET17b-based vector, expressing CYP107DY1, BmCPR and Fdx2, was constructed. The whole-cell biotransformation was performed with resting cells in potassium phosphate buffer instead of the terrific broth complex medium, since indole, which results from the tryptophan metabolism by *E. coli*, may have an inhibitory effect on the activity of P450s (Brixius-Anderko et al., 2016; Ringle et al., 2013).

As shown in Table 1, the whole-cell biotransformation system yielded about 13.2 mg/L pravastatin. This amount is comparable to the previously reported WT CYP105A3 (P450sca2) *E. coli*-based whole-cell system, which yielded 12.8 mg/L pravastatin (Ba et al., 2013a, 2013b). It was shown previously that the hindered transport of the substrate into the *E. coli* cells is most likely responsible for the low biotransformation ratio (Janocha and Bernhardt, 2013). Therefore, we used the peptide antibiotic polymyxin B and the chelating agent EDTA, which are reported to exhibit permeabilization activity towards the outer membrane of the Gram-negative bacteria. Concerning this, *E. coli*-based whole cell systems were also designed for other P450s in our laboratory, in which optimal concentrations of polymyxin B (32 µg/ml) and EDTA (20 mM) were employed to achieve the highest conversion of substrates (Janocha and Bernhardt, 2013; Kern et al., 2016; Litzemberger et al., 2015). The use of polymyxin B lead to an 1.3 fold increase in the pravastatin production, yielding 17.9 mg/L, while the highest yield was achieved using EDTA with an 2.2 fold increase yielding 28.5 mg/L (Table 1), which suggests the potential of the constructed whole-cell system in the production of pravastatin.

Novel P450s are increasingly emerging because of the availability of many genome sequences. However, the multi-component nature of the P450s and other factors (Bernhardt and Urlacher, 2014) limit their biotechnological use. The characterization of CYP107DY1 and its successful employment in the *in vivo* pravastatin production is the first step in the industrialization of this interesting P450. The application of this novel pravastatin-producing P450 in an industrially optimized host as recently shown for the mutant CYP105AS1 (McLean et al., 2015) will certainly pave the way for a successful use in a biotechnological process.

4. Conclusion

Regio- and stereoselective oxidation of non-activated carbon atoms using P450s is of great interest for synthetic biology. The

robustness of the bacterial P450s is attractive for the industrial production of different valuable products (Bernhardt and Urlacher, 2014). However, despite the high potential and the recent development with respect to exploiting the P450s in biotechnological applications, there are various challenges limiting their use in industrial processes, for instance their dependency on electron transfer proteins. In the recent study, we have identified, cloned and expressed a new plasmid-encoded cytochrome P450 from *B. megaterium* QM B1551, which was categorized into a new subfamily (CYP107DY1). In addition, we were able to find a redox system (BmCPR-Fdx2) that can efficiently support the activity of the CYP107DY1, which enabled the identification of mevastatin as a substrate. The successful construction of *E. coli*-based whole cell system utilizing CYP107DY1 is of great importance not only for efficient production of pravastatin but also for future characterization of new substrates, functionalization of drugs and the production of new metabolites. Further improvements, either at the cellular level by expressing CYP107DY1 in other microorganisms (McLean et al., 2015), or at the molecular level by rational design to improve the activity of this P450 (Ba et al., 2013b) will contribute to the set-up of efficient industrial processes using this novel P450.

Author contributions

M.M. carried out all experiments, analyzed and interpreted the data and drafted the manuscript. N.P. and F.K. participated in the establishment of the experiments. A.A. participated in the substrate docking. P.H. and R.B. participated in the interpretation and discussion of experimental results and writing of the manuscript.

Conflict of interest

The authors declare that they have no competing interest.

Acknowledgments

The authors would like to thank Birgit Heider-Lips for the purification of AdR and Adx, Dr. Elisa Brill for cloning of Fdx2 and Tanja Sagadin for the purification of Arh1 and Etp1^{fd}. The authors also gratefully acknowledge Prof. Dr. David Nelson (The University of Tennessee Health Science Center) for helping in the CYP107DY1 classification.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jbiotec.2016.11.002>.

References

- Arakawa, K., Kodama, K., Tatsuno, S., Ide, S., Kinashi, H., 2006. Analysis of the loading and hydroxylation steps in lankamycin biosynthesis in *Streptomyces rochei*. *Antimicrob. Agents Chemother.* 50, 1946–1952.
- Arase, M., Waterman, M.R., Kagawa, N., 2006. Purification and characterization of bovine steroid 21-hydroxylase (P450c21) efficiently expressed in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 344, 400–405.
- Ba, L., Li, P., Zhang, H., Duan, Y., Lin, Z., 2013a. Engineering of a hybrid biotransformation system for cytochrome P450sca-2 in *Escherichia coli*. *Biotechnol. J.* 8, 785–793.
- Ba, L., Li, P., Zhang, H., Duan, Y., Lin, Z., 2013b. Semi-rational engineering of cytochrome P450sca-2 in a hybrid system for enhanced catalytic activity: insights into the important role of electron transfer. *Biotechnol. Bioeng.* 110, 2815–2825.
- Berg, A., Gustafsson, J.A., Ingelman-Sundberg, M., 1976. Characterization of a cytochrome P450-dependent steroid hydroxylase system present in *Bacillus megaterium*. *J. Biol. Chem.* 251, 2831–2838.
- Berg, A., Ingelman-Sundberg, M., Gustafsson, J.A., 1979. Isolation and characterization of cytochrome P450meg. *Acta Biol. Med. Ger.* 38, 333–344.

- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2013. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1703–1717.
- Brixius-Anderko, S., Schiffer, L., Hannemann, F., Janocha, B., Bernhardt, R., 2015. A CYP21A2 based whole-cell system in *Escherichia coli* for the biotechnological production of premedrol. *Microb. Cell Fact.* 14, <http://dx.doi.org/10.1186/s12934-015-0333-2>.
- Brixius-Anderko, S., Hannemann, F., Ringle, M., Khatri, Y., Bernhardt, R., 2016. An indole deficient *Escherichia coli* strain improves screening of cytochromes P450 for biotechnological applications. *Biotechnol. Appl. Biochem.*, <http://dx.doi.org/10.1002/bab.1488>.
- Bunk, B., Schulz, A., Stammes, S., Münch, R., Warren, M.J., Rohde, M., Jahn, D., Biedendieck, R., 2010. A short story about a big magic bug. *Bioeng. Bugs* 1, 85–91.
- Bureik, M., Schiffer, B., Hiraoka, Y., Vogel, F., Bernhardt, R., 2002. Functional expression of human mitochondrial CYP11B2 in fission yeast and identification of a new internal electron transfer protein. *etp1*. *Biochemistry (Mosc.)* 41, 2311–2321.
- Chung, Y.-H., Song, J.-W., Choi, K.-Y., Yoon, J.W., Yang, K.-M., Park, J.-B., 2012. Cloning, expression, and characterization of P450 monooxygenase CYP102H1 from *Nocardia farcinica*. *J. Korean Soc. Appl. Biol. Chem.* 55, 259–264.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Ewen, K.M., Schiffer, B., Uhlmann-Schiffer, H., Bernhardt, R., Hannemann, F., 2008. The endogenous adrenodoxin reductase-like flavoprotein arh1 supports heterologous cytochrome P450-dependent substrate conversions in *Schizosaccharomyces pombe*. *FEMS Yeast Res.* 8, 432–441.
- Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Fact.* 14, 107, <http://dx.doi.org/10.1186/s12934-015-0300-y>.
- Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., 2007. Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta BBA Gen. Subj.* 1770 (P450), 330–344.
- Hiwatashi, A., Ichikawa, Y., Maruya, N., Yamano, T., Aki, K., 1976. Properties of crystalline reduced nicotinamide adenine-dinucleotide phosphate-adrenodoxin reductase from bovine adrenocortical mitochondria. 1. Physicochemical properties of holo-NADPH-adrenodoxin and apo-NADPH-adrenodoxin reductase and interaction between non-heme iron proteins and reductase. *Biochemistry (Mosc.)* 15, 3082–3090.
- Inouye, M., Takada, Y., Muto, N., Beppu, T., Horinouchi, S., 1994. Characterization and expression of a P450-like mycinamicin biosynthesis gene using a novel Micromonospora-*Escherichia coli* shuttle cosmid vector. *Mol. Gen. Genet.* MGG 245, 456–464.
- Jóźwik, I.K., Kiss, F.M., Gricman, E., Abdumughni, A., Brill, E., Zapp, J., Pleiss, J., Bernhardt, R., Thunnissen, A.-M.W.H., 2016. Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J.*, <http://dx.doi.org/10.1111/febs.13911>.
- Janocha, S., Bernhardt, R., 2013. Design and characterization of an efficient CYP105A1-based whole-cell biocatalyst for the conversion of resin acid diterpenoids in permeabilized *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 97, 7639–7649.
- Kern, F., Khatri, Y., Litzenburger, M., Bernhardt, R., 2016. CYP267A1 and CYP267B1 from *Sorangium cellulosum* So ce56 are highly versatile drug metabolizers. *Drug Metab. Dispos.* 44, 495–504.
- Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H., 2016. PubChem substance and compound databases. *Nucleic Acids Res.* 44, D1202–1213.
- Lamon-Fava, S., 2013. Statins and lipid metabolism: an update. *Curr. Opin. Lipidol.* 24, 221–226.
- Lepesheva, G.I., Podust, L.M., Bellamine, A., Waterman, M.R., 2001. Folding requirements are different between sterol 14 α -demethylase (CYP51) from *Mycobacterium tuberculosis* and human or fungal orthologs. *J. Biol. Chem.* 276, 28413–28420.
- Litzenburger, M., Kern, F., Khatri, Y., Bernhardt, R., 2015. Conversions of tricyclic antidepressants and antipsychotics with selected P450s from *Sorangium cellulosum* So ce56. *Drug Metab. Dispos.* 43, 392–399.
- Lombard, M., Salard, I., Sari, M.-A., Mansuy, D., Buisson, D., 2011. A new cytochrome P450 belonging to the 107L subfamily is responsible for the efficient hydroxylation of the drug terfenadine by *Streptomyces platensis*. *Arch. Biochem. Biophys.* 508, 54–63.
- Ly, T.T.B., Khatri, Y., Zapp, J., Hutter, M.C., Bernhardt, R., 2012. CYP264B1 from *Sorangium cellulosum* So ce56: A fascinating norisoprenoid and sesquiterpene hydroxylase. *Appl. Microbiol. Biotechnol.* 95, 123–133.
- Matsuoka, T., Miyakoshi, S., Tanzawa, K., Nakahara, K., Hosobuchi, M., Serizawa, N., 1989. Purification and characterization of cytochrome P450sca from *Streptomyces carbophilus*. ML-236B (compactin) induces a cytochrome P450sca in *Streptomyces carbophilus* that hydroxylates ML-236B to pravastatin sodium (CS-514), a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl-coenzyme-A reductase. *Eur. J. Biochem. FEBS* 184, 707–713.
- McLean, K.J., Hans, M., Meijrink, B., van Scheppingen, W.B., Vollebregt, A., Tee, K.L., van der Laan, J.-M., Leys, D., Munro, A.W., van den Berg, M.A., 2015. Single-step fermentative production of the cholesterol-lowering drug pravastatin via reprogramming of *Penicillium chrysogenum*. *Proc. Natl. Acad. Sci. U. S. A.* 112, 2847–2852.
- Milhim, M., Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. *J. Biotechnol.* 231, 83–94.
- Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791.
- Munro, A.W., Lindsay, J.G., Coggins, J.R., Kelly, S.M., Price, N.C., 1994. Structural and enzymological analysis of the interaction of isolated domains of cytochrome P450 BM3. *FEBS Lett.* 343, 70–74.
- Nelson, D.R., 2009. The cytochrome P450 homepage. *Hum. Genomics* 4, 59–65.
- Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura, T., 1998. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694–1699.
- Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. II. Solubilization purification, and properties. *J. Biol. Chem.* 239, 2379–2385.
- Poulos, T.L., Finzel, B.C., Howard, A.J., 1986. Crystal structure of substrate-free *Pseudomonas putida* cytochrome P450. *Biochemistry (Mosc.)* 25, 5314–5322.
- Poulos, T.L., Finzel, B.C., Howard, A.J., 1987. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* 195, 687–700.
- Ravichandran, K.G., Boddupalli, S.S., Hasermann, C.A., Peterson, J.A., Deisenhofer, J., 1993. Crystal structure of hemoprotein domain of P450 BM3, a prototype for microsomal P450's. *Science* 261, 731–736.
- Ringle, M., Khatri, Y., Zapp, J., Hannemann, F., Bernhardt, R., 2013. Application of a new versatile electron transfer system for cytochrome P450-based *Escherichia coli* whole-cell bioconversions. *Appl. Microbiol. Biotechnol.* 97, 7741–7754.
- Sagara, Y., Wada, A., Takata, Y., Waterman, M.R., Sekimizu, K., Horiuchi, T., 1993. Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. *Biol. Pharm. Bull.* 16, 627–630.
- Sakaki, T., Sugimoto, H., Hayashi, K., Yasuda, K., Munetsuna, E., Kamakura, M., Ikushiro, S., Shiro, Y., 2011. Bioconversion of vitamin D to its active form by bacterial or mammalian cytochrome P450. *Biochim. Biophys. Acta BBA Proteins Proteomics* 1814, 249–256.
- Sanner, M.F., 1999. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* 17, 57–61.
- Schiffer, B., Bureik, M., Reinle, W., Müller, E.-C., Hannemann, F., Bernhardt, R., 2004. The adrenodoxin-like ferredoxin of *Schizosaccharomyces pombe* mitochondria. *J. Inorg. Biochem.* 98, 1229–1237.
- Schmitz, D., Zapp, J., Bernhardt, R., 2012. Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*. *FEBS J.* 279, 1663–1674.
- Uhlmann, H., Beckert, V., Schwarz, D., Bernhardt, R., 1992. Expression of bovine adrenodoxin in *Escherichia coli* and site-directed mutagenesis of (2-Fe-2s) cluster ligands. *Biochem. Biophys. Res. Commun.* 188, 1131–1138.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450 BM3 (CYP102A1): connecting the dots. *Chem. Soc. Rev.* 41, 1218–1260.

Supplementary material

Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin

**Mohammed Milhim, Natalia Putkaradze, Ammar Abdulmughni, Fredy Kern, Philip Hartz and
Rita Bernhardt***

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

*Address correspondence to:

Prof. Dr. Rita Bernhardt, Institute of Biochemistry, Campus B 2.2, Saarland University, 66123
Saarbrücken (Germany)

E-mail: ritabern@mx.uni-saarland.de

Phone: +49 681 302 4241

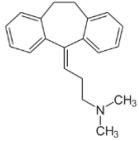
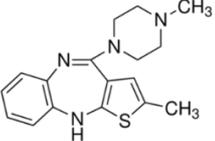
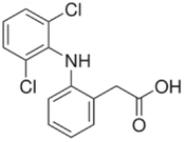
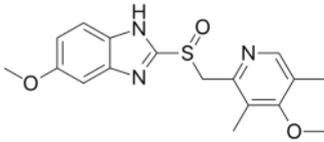
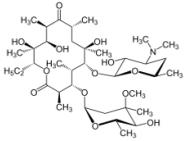
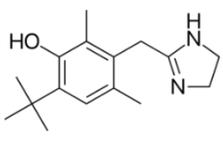
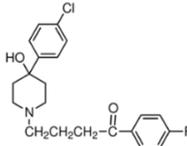
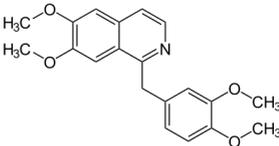
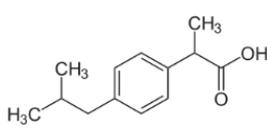
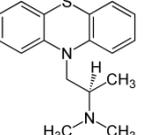
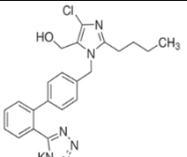
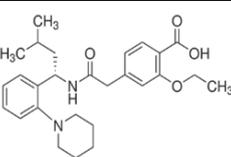
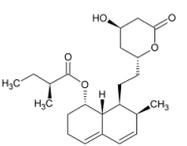
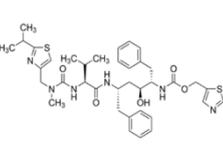
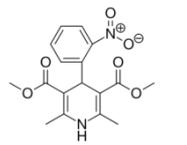
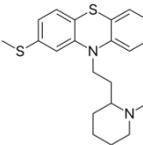
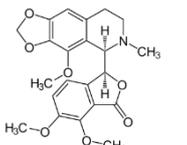
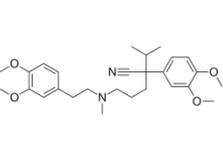
Fax: +49 681 302 4739

Supplementary Table 1. Protein sequence identity values of CYP107DY1 with different cytochrome P450s.

Closest match with different CYP107					
Protein Abb.	Natural function or substrate/s	No. of aa	Species	Identity% ^(a)	UniProtKB accession numbers ^(b)
CYP107DY1	---	410	<i>B. megaterium</i> QM B1551	100	D5E3H2
CYP267B1	Drug metabolism	405	<i>Sorangium cellulosum</i> So ce5	43.5	A9ERX9
CYP107B1	7-ethoxycoumarin	405	<i>Saccharopolyspora erythraea</i>	41.9	P33271
CYP107K1 (PksS)	Polyketide biosynthesis (bacillaene biosynthesis)	405	<i>Bacillus subtilis</i> (Strain 168)	41.8	O31785
CYP107H1 (bioI)	Biotin biosynthesis	395	<i>Bacillus subtilis</i> (Strain 168)	40.8	P53554
CYP107J1	Testosterone enanthate	410	<i>Bacillus subtilis</i> (Strain 168)	40.7	O08469
CYP107BR1 (Vdh)	Vitamin D3	403	<i>Pseudonocardia autotrophica</i>	40	C4B644
CYP107L1 (PIKC)	Pikromycin biosynthesis	416	<i>Streptomyces venezuelae</i>	39.2	O87605
CYP107A1 (eryF)	Erythromycin biosynthesis	404	<i>Saccharopolyspora erythraea</i>	36.2	Q00441
CYP107E4	Diclofenac	396	<i>Actinoplanes sp.ATCC 53771</i>	36	C0LR90
CYP109B1	Various substrates (e.g. (+)-valencene)	396	<i>Bacillus subtilis</i>	35.4	U5U1Z3
CYP267A1	Drugs metabolism	429	<i>Sorangium cellulosum</i> So ce5	35.3	A9EN90
CYP109A2	Steroids	403	<i>B. megaterium</i> DSM319	34.6	D5DF88
CYP109E1	Steroids	404	<i>B. megaterium</i> DSM319	32.7	D5DKI8
CYP106A1	Steroids and terpenoids	410	<i>B. megaterium</i> DSM319	29.1	D5DF35
CYP108	α -terpinol	428	<i>Pseudomonas sp.</i>	28	P33006
CYP102A1	Fatty acids	1,049	<i>B. megaterium</i> DSM319	18.4	P14779

(a) Based on Clustal Omega Multiple Sequence Alignment (MSA) (<http://www.ebi.ac.uk/Tools/msa/>)(b) <http://www.uniprot.org/>

Supplementary Table 2. List of tested substrates with CYP107DY1

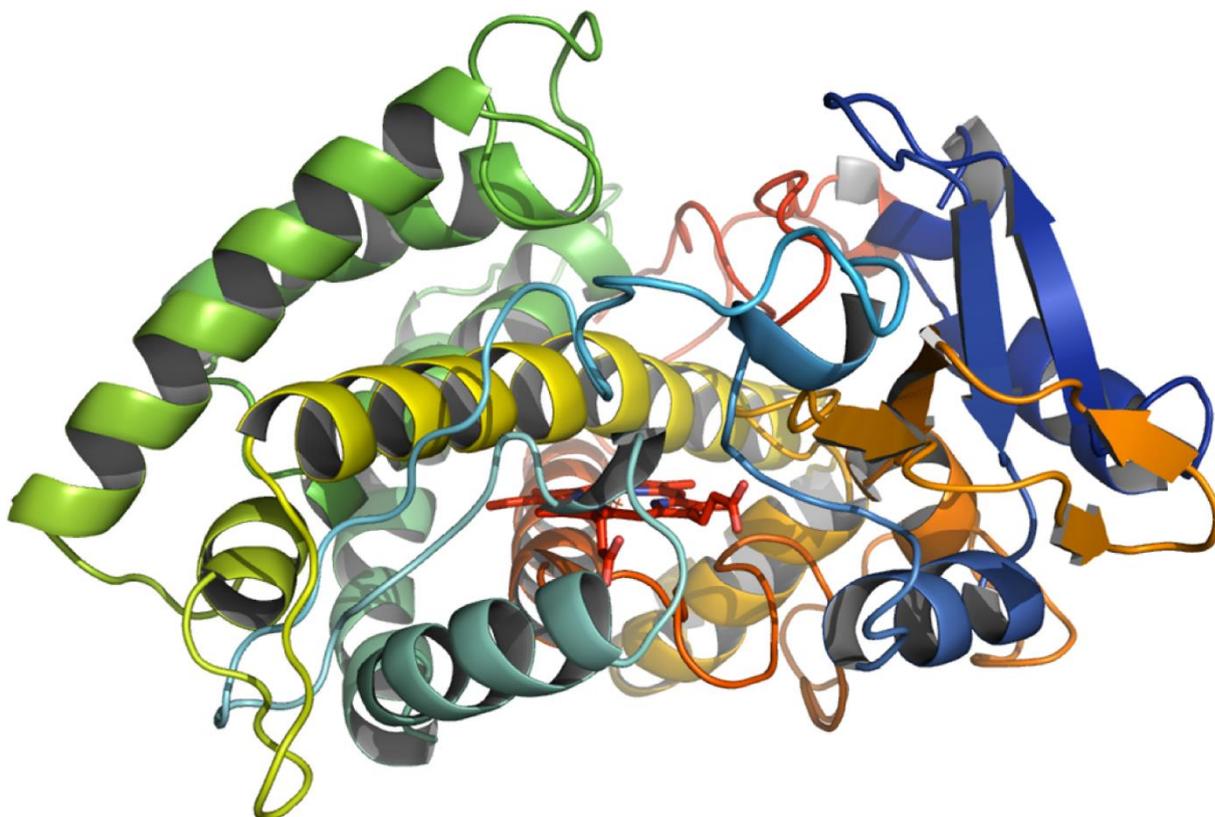
Name	Structure	Name	Structure
Amitriptyline		Olanzapine	
Diclofenac		Omeprazole	
Erythromycin		Oxymetazoline	
Haloperidol		Papaverine	
Ibuprofen		Promethazine	
Losartan		Repaglinide	
Mevastatin		Ritonavir	
Nifedipine		Thioridazine	
Noscopine		Verapamil	

```

1 ATGAAAAAGGTTACAGTTGATGATTTTAGCTCTCCAGAAAATATGCACGATGTCATCGGATTTTATAAAAACTCACTGAACATCAAGAACCTCTTATTCGTTG
1 M K K V T V D D F S S P E N M H D V I G F Y K K L T E H Q E P L I R L
106 GATGATTATTACGGTTGGGACCGGCATGGGTCGCATTACGTCATGACGATGTTGTTACGATACTAAAGAACCCCGCTTTTCTCAAAGATGTACGGAAAGTTCACA
36 D D Y Y G L G P A W V A L R H D D V V T I L K N P R F L K D V R K F T
211 CCATTGCAAGATAAAAAGGATTCTATAGATGATAGCACATCTGCAGCAAACTGTTTGAATGGATGATGAATATGCCGAATATGCTTACGGTCGATCCACCCGAT
71 P L Q D K K D S I D D S T S A S K L F E W M M N M P N M L T V D P P D
316 CACACTCGTTTGCAGGTTGGCCTCTAAAGCCTTTACGCCACGTATGATCGAGAATCTTCGACCTCGTATACAGCAGATTACCAATGAGCTATTGATTTCAGTA
106 H T R L R R L A S K A F T P R M I E N L R P R I Q Q I T N E L L D S V
421 GAAGGAAAAAGGAATATGGATCTTGTTCGGATTTTCTTTTCCTGCCCATTATTGTCATTTTCAGAGATGCTAGGGATTCCACCTTTAGATCAGAAAACGATTT
141 E G K R N M D L V A D F S F P L P I I V I S E M L G I P P L D Q K R F
526 CGCGACTGGACAGATAAACTCATCAAAGCAGCTATGGATCCTAGCCAAGGGGCTGTAGTTTATGGAAACACTCAAGGAGTTTATTGATTACATCAAAAAATGCTG
176 R D W T D K L I K A A M D P S Q G A V V M E T L K E F I D Y I K K M L
631 GTCGAAAAGCGCAACCATCCAGACGATGATGTGATGAGTGCTTTGTTGCAAGCACATGAGCAAGAAGATAAGTTGAGCGGAGAACGAGCTTCTTTCCACGATTTGG
211 V E K R N H P D D D V M S A L L Q A H E Q E D K L S E N E L L S T I W
736 CTACTCATTACAGCCGGACATGAGACGACGGCCCATCTAATCAGCAACGGCCTACTGGCGCTATTGAAGCATCCCGAACAAATGCGCCTGCTTCGGGATAATCCT
246 L L I T A G H E T T A H L I S N G V L A L L K H P E Q M R L L R D N P
841 TCTTTACTCCCTCTGCCGTTGAAGAGCTGCTACGCTATGCCGACCGGTCATGATTGGTGGCGTTTTGCGGGTGAAGATATCATCATGATGAAAAATGATT
281 S L L P S A V E E L L R Y A G P V M I G G R F A G E D I I M H G K M I
946 CCCAAAGGTGAAATGGTGTGTTCTCGCTGGTTGCCCAATATTGATTCACAGAAATTCCTTATCCTGAGGGATTGGATATTACACGCGAGGAGAATGAGCAT
316 P K G E M V L F S L V A A N I D S Q K F S Y P E G L D I T R E E N E H
1051 CTCACCTTCGAAAAGGTATCCATCATTGTTTGGGAGCGCCTTTGGCGCGCATGGAAGCACATATCGCTTTCGGCAGATTGCTTCAACGGTTTCTGATTACGA
351 L T F G K G I H H C L G A P L A R M E A H I A F G T L L Q R F P D L R
1156 TTGGCAATCGAATCGGAGCAACTGGTTTATAACAACAGCACATTGCGTTCCTTAAAGCTTGCCAGTTATTTCTAA
386 L A I E S E Q L V Y N N S T L R S L K S L P V I F *

```

Supplementary Fig. S1. The open reading frame sequence of the CYP107DY1. The upper and the lower lines represent the nucleotide and deduced amino acid sequences, respectively. The one-letter code for each amino acid is aligned with the first nucleotide of each codon. Several conserved motifs used for the identification of cytochrome P450s are underlined. I-helix (A/G-G-x-E/D-T-T/S), K-helix (E-x-x-R), and the heme pocket (F-x-x-G-x-x-x-C-x-G).



Supplementary Fig. S2. Homology model of CYP107DY1. A model of CYP107DY1 was calculated using CYP107RB1 (Vdh) (PDB accession code: 3A4G) as template. The coordinates of the heme-porphyrin atoms from the template structure were added subsequently to the obtained homology model. Program Modeller 9.14 (University of California San Francisco, USA) was used.

2.3 Gerber et al., 2016

Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one

Adrian Gerber, **Mohammed Milhim**, Philip Hartz, Josef Zapp and Rita Bernhardt

Metabolic Engineering, 2016 Jul. 36: 19–27.

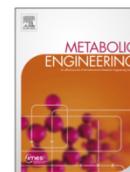
DOI: 10.1016/j.ymben.2016.02.010

Reprinted with permission of Metabolic Engineering. All rights reserved.



Contents lists available at ScienceDirect

Metabolic Engineering

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

Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one



Adrian Gerber^a, Mohammed Milhim^a, Philip Hartz^a, Josef Zapp^b, Rita Bernhardt^{a,*}

^a Saarland University, Institute of Biochemistry, Campus B2.2, 66123 Saarbrücken, Germany

^b Saarland University, Institute of Pharmaceutical Biology, Campus C2.3, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 19 October 2015
Received in revised form
1 February 2016
Accepted 23 February 2016
Available online 11 March 2016

Keywords:

Bacillus megaterium
Hydroxysteroid dehydrogenase
Steroid hormones
17 α ,20 β -Dihydroxy-4-pregnen-3-one
17 α ,20 β ,21 α -Trihydroxy-4-pregnen-3-one
MIH

ABSTRACT

17 α ,20 β -Dihydroxy-4-pregnen-3-one (17 α ,20 β DiOH-P) and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one (20 β OH-RSS) are the critical hormones required for oocyte maturation in fish. We utilized *B. megaterium*'s endogenous 20 β -hydroxysteroid dehydrogenase (20 β HSD) for the efficient production of both progestogens after genetically modifying the microorganism to reduce side-product formation. First, the gene encoding the autologous cytochrome P450 CYP106A1 was deleted, resulting in a strain devoid of any steroid hydroxylation activity. Cultivation of this strain in the presence of 17 α -hydroxyprogesterone (17 α OH-P) led to the formation of 17 α ,20 α -dihydroxy-4-pregnen-3-one (17 α ,20 α DiOH-P) as a major and 17 α ,20 β DiOH-P as a minor product. Four enzymes were identified as 20 α HSDs and their genes deleted to yield a strain with no 20 α HSD activity. The 3-oxoacyl-(acyl-carrier-protein) reductase FabG was found to exhibit 20 β HSD-activity and overexpressed to create a biocatalyst yielding 0.22 g/L 17 α ,20 β DiOH-P and 0.34 g/L 20 β OH-RSS after 8 h using shake-flask cultivation, thus obtaining products that are at least a thousand times more expensive than their substrates.

© 2016 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

1. Introduction

The oocyte maturation inducing hormones (MIHs) 17 α ,20 β DiOH-P and 20 β OH-RSS are the main progestogens in teleosts. In female fishes, MIH production in the ovarian follicle layers is stimulated by gonadotropin synthesized in the pituitary. MIH induces the formation of the maturation promoting factor (MPF) leading to further oocyte maturation (Nagahama and Yamashita, 2008). In addition, MIHs are produced in the testis of male and, among other functions, regulate spermiation and sperm motility (Scott et al., 2010). Both MIHs are produced from

cholesterol in a series of reaction steps involving the cytochromes P450 CYP11A1, CYP17, CYP21 as well as 3 β - and 20 β HSD.

The administration of MIHs to different species of aquacultured fish has been extensively investigated and shown to greatly increase spawning rates due to, for instance, overcoming ovulation disorder, improving sperm volume and concentration, attracting fish to spawning sites and induction of ovulation (Yamamoto et al., 2015; Martins Pinheiro et al., 2003; Hong et al., 2006; Haider and Rao, 1994; King and Young, 2001; Ohta et al., 1996; Miwa et al., 2001). Induced spawning is of particular interest for economically significant species of fish that do not reproduce spontaneously under captive cultivation conditions (Lee and Yang, 2002). However, the widespread use of both synthetic MIHs in aquacultures has been hindered by their high cost.

Steroid hormones are produced mainly through a combination of chemical and microbial conversion steps, since single-step stereospecific modifications of the unreactive steroid nucleus by chemical means alone are often not possible (Donova and Egorova, 2012). The chemical synthesis of 17 α ,20 β DiOH-P is tedious, requiring multiple reaction steps to prepare the steroid stereoselectively. One method to produce the progestogen includes the chlorination, dechlorination and reduction of RSS to produce the substance with an overall yield of 64 % (Ouedraogo et al., 2013). Another approach consists of reducing 17OH-P by NaBH₄ in

Abbreviations: AKR, aldo-keto reductase; 17OH-P, 17 α -hydroxyprogesterone; 17 α ,20 α DiOH-P, 17 α ,20 α -dihydroxy-4-pregnen-3-one; 17 α ,20 β DiOH-P, 17 α ,20 β -dihydroxy-4-pregnen-3-one; 20 β OH-RSS, 17 α ,20 α ,21 α -trihydroxy-4-pregnen-3-one; DOC, 11-deoxycorticosterone; FabG, 3-oxoacyl-(acyl-carrier-protein) reductase; HSD, hydroxysteroid dehydrogenase; LB, lysogeny broth; MIH, maturation inducing hormone; MPF, maturation promoting factor; NMR, nuclear magnetic resonance; ORF, open reading frame; PHB, poly(3-hydroxybutyrate); RP-HPLC, reverse phase high performance liquid chromatography; RSS, 11-deoxycortisol; SOE-PCR, splicing by overlapping polymerase chain reaction; TB, terrific broth; Upp, uracil phosphoribosyltransferase

* Corresponding author.

E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

<http://dx.doi.org/10.1016/j.ymben.2016.02.010>

1096-7176/© 2016 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

Table 1List of *B. megaterium* strains and plasmids used in this study.

Strains or plasmids	Description	Reference
Strains		
MS941	Mutant of DSM319, $\Delta nprM$ (extracellular protease)	Wittchen and Meinhardt (1995)
GHH1	Mutant of MS941, Δupp (selection marker)	Gerber et al. (2015)
GHH2	Mutant of GHH1, $\Delta cyp106A1$	This study
GHH5	Mutant of GHH2, ΔBMD_0912 (20 α HSD)	This study
GHH6	Mutant of GHH5, ΔBMD_1068 (20 α HSD)	This study
GHH7	Mutant of GHH6, ΔBMD_3715 (20 α HSD)	This study
GHH8	Mutant of GHH7, ΔBMD_1595 (20 α HSD)	This study
Plasmids		
pSMF2.1	Backbone vector for protein overexpression	Bleif et al. (2012)
pSMF2.1_0912	Expression of <i>BMD_0912</i> (20 α HSD)	This study
pSMF2.1_1595	Expression of <i>BMD_1595</i> (20 α HSD)	This study
pSMF2.1_1068	Expression of <i>BMD_1068</i> (20 α HSD)	This study
pSMF2.1_3715	Expression of <i>BMD_3715</i> (20 α HSD)	This study
pSMF2.1_FabG	Expression of FabG (20 β HSD)	This study
pUCTV2_Upp	Backbone vector for gene deletion	Gerber et al. (2015), based on Wittchen and Meinhardt (1995)
pUCTV2_Upp_ $\Delta 106A1$	Genomic deletion of <i>cyp106A1</i>	This study
pUCTV2_Upp_ $\Delta 0912$	Genomic deletion of <i>BMD_0912</i>	This study
pUCTV2_Upp_ $\Delta 1595$	Genomic deletion of <i>BMD_1595</i>	This study
pUCTV2_Upp_ $\Delta 1068$	Genomic deletion of <i>BMD_1068</i>	This study
pUCTV2_Upp_ $\Delta 3715$	Genomic deletion of <i>BMD_3715</i>	This study

methanol, resulting in a ~30 % yield of 17 α ,20 β DiOH-P (Kovganko et al., 2001). In microbial fermentations with strains of *B. megaterium*, *Yarrowia lipolytica*, *Saccharomyces cerevisiae* or *Bifidobacterium adolescentis* 17 α ,20 β DiOH-P appeared only as a minor side-product (Shkumatov et al., 2003; Winter et al., 1982; Koshcheenko et al., 1976). Genetically engineered microorganisms are often used to increase the accumulation of desired products by, for instance, deletion of genes to avoid by-product formation (Yao et al., 2014), augmenting the activity of endogenous enzymes by overexpression (Yao et al., 2013) or mutagenesis (Hunter et al., 2011), or increasing the tolerance of the organism towards metabolite and substrate stress (Nicolaou et al., 2010).

In this study, we present a genetically modified *B. megaterium* strain allowing the high yield production of both 17,20 β DiOH-P and 20 β OH-RSS from cheap steroid precursors. This Gram-positive, rod-shaped bacterium has gained considerable interest in recent years as a recombinant expression host due to its high protein production capacity, plasmid stability, ability to take up a variety of hydrophobic substrates as well as its large cell size, which allows detailed microscopic analyses (Vary et al., 2007). Like other *Bacillus* species, *B. megaterium* can be genetically engineered through homologous recombination with exogenous DNA, with or without applying a marker gene (Dong and Zhang, 2014). The classic procedure consists of protoplast transforming cells with the deletion vector, the integration of parts of the vector into the chromosome, screening the target locus for the mutation by PCR and finally curing the vector from the cells (Biedendieck et al., 2011). A more recent approach allows the one-step deletion of genes by using a transconjugation protocol with *Escherichia coli* as plasmid-donor cells (Richhardt et al., 2010). *B. megaterium* has been shown to be particularly suitable for the bioconversion of hydrophobic steroidal compounds, including cholesterol and analogs (Gerber et al., 2015), natural steroid hormones derived from cholesterol such as pregnenolone and dehydroepiandrosterone (Schmitz et al., 2014), synthetic steroid hormones such as prednisolone and dexamethasone (Kiss et al., 2015) and the secosteroid vitamin D3 (Ehrhardt et al., 2016). The genome of *B. megaterium* DSM319, the precursor of the strain used in this study, contains the gene for the cytochrome P450 CYP106A1, resulting in an endogenous steroid hydroxylation activity. Cytochromes P450 form a superfamily of heme-thiolate proteins, which are, in bacteria, mainly involved in the metabolism of xenobiotics and the

production of secondary metabolites (Bernhardt, 2006; Bernhardt and Urlacher, 2014). CYP106A1 is able to convert a wide variety of steroids, including testosterone, progesterone, 17-hydroxyprogesterone, RSS, DOC and cortisol (Kiss et al., 2015).

The aim of this study was to construct a biocatalyst for the production of the fish progestogens 17 α ,20 β DiOH-P and 20 β OH-RSS by utilizing an endogenous 20 β HSD and abolishing side-product formation from the likewise endogenous CYP106A1 and four 20 α HSDs through genetic engineering.

2. Materials and methods

2.1. Materials

20 β OH-RSS (4-pregnen-17 α ,20 β ,21-triol-3-one, catalog ID: Q4080-000), 20 β OH-cortisol (4-pregnen-11 β , 17, 20 β , 21-tetrol-3-one, Q3790-000) and 20 β OH-cortisone (4-pregnen-17, 20 β , 21-triol-3, 11-dione, Q3960-000) were purchased from Steraloids. 17OH-P (4-pregnen-17 α -ol-3,20-dione, H5752) RSS (4-Pregnene-17 α ,21-diol-3,20-dione, R0500), cortisol (4-pregnen-11 β ,17 α ,21-triol-3,20-dione, H4001), cortisone (pregnen-17 α ,21-diol-3,11,20-trione, C2755) and 17 α ,20 α DiOH-P (4-pregnen-17, 20 β -diol-3-one, P6285), xylose (95729), tetracycline (T7660), ethyl acetate (34858) and acetonitrile (34851) were from Sigma-Aldrich. LB broth (244610), yeast extract (212750) and tryptone (211705) were obtained from Becton Dickinson.

2.2. Molecular cloning and gene deletion

All plasmids were constructed by conventional cloning using restriction enzyme digestion and ligation. *E. coli* strain TOP10 (Invitrogen) was used for the propagation of plasmids. *B. megaterium* was transformed according to the PEG-mediated protoplast transformation method (Biedendieck et al., 2011). Table 1 lists all plasmids used in this study. Plasmid maps displaying backbone vectors and the restriction sites used for cloning are depicted in Supplemental Fig. 1. The xylose-inducible vector pSMF2.1 was used as a backbone for the overexpression of genes. ORFs of *BMD_0912*, *BMD_1595*, *BMD_1068*, *BMD_3715* and *FabG* were amplified from genomic DNA of *B. megaterium* MS941 and cloned into pSMF2.1. Genomic DNA was prepared using a genomic DNA

isolation kit (nexttec). For gene deletions, pUCTV2_Upp was used as a backbone vector (Gerber et al., 2015; Wittchen and Meinhardt, 1995). The knockout of *cyp106A1*, *BMD_0912*, *BMD_1595*, *BMD_1068* and *BMD_3715* was carried out as previously described (Gerber et al., 2015). In brief, flanking regions of the target genes were amplified from genomic DNA, fused via splicing by overlapping extension (SOE)-PCR and cloned into pSMF2.1_Upp (Table 1). *B. megaterium* was transformed with the resulting plasmids, grown colonies streaked out on minimal medium agar plates (Gerber et al., 2015) and incubated overnight at 42 °C. Colonies were then plated on minimal medium agar plates containing 1 μM 5-fluorouracil and again incubated overnight. Resulting colonies are plasmid-free and can be screened for target gene deletion by PCR. All primers used in this study are listed in Supplemental Table 1.

2.3. *B. megaterium* cultivation conditions and in vivo substrate conversions

Precultures of *B. megaterium* were started by inoculating LB-medium (25 g/L) with cells from a glycerol stock or colony and incubated overnight at 30 °C and 180 rpm shaking. TB-medium (24 g/L yeast extract, 12 g/L tryptone, 0.4% glycerol, 100 mM potassium phosphate buffer, pH 7.4) was used for the main cultures, which were inoculated 1:100 with the precultures. For cells harboring pSMF2.1 or pUCTV2 derivatives, tetracycline was added to the cultures to a final concentration of 10 μg/mL. Recombinant protein expression was induced at an optical density of ca. 0.4 at 600 nm by adding xylose dissolved in water to a final concentration of 0.5% (w/v). For the assessment of intrinsic CYP106A1, 20αHSD and 20βHSD activities, substrates were added in the early logarithmic growth phase (approximately 2 h after inoculation of the main culture), after being dissolved in ethanol (stock solutions: 14.9 mM progesterone, 15 mM 17OH-P, final ethanol concentration in the cultures: 2%). For *in vivo* conversions with over-expressed FabG, the substrate was added 20 h after protein induction, to achieve highest possible product yields (stock solutions: 25 mM 17OH-P and RSS, final ethanol concentration in the cultures: 4%).

2.4. Culture sample treatment and RP-HPLC analysis

1 mL culture samples were taken at different time points and extracted twice with equal volumes of ethylacetate. In case of the 20αHSD activity measurements, progesterone was added as an internal standard prior to that to account for extraction loss. After evaporation, the extracts were dissolved in 10% (v/v) acetonitrile in water. Gradient RP-HPLC analysis was carried out using a 125/4 Nucleodur C18 column (Macherey & Nagel) on a device manufactured by Jasco with detection at a wavelength of 240 nm. 10% (v/v) acetonitrile in water was used as mobile phase on channel A, pure acetonitrile on channel B. For all HPLC measurements the following method was used: 0–3.5 min: 100% A, 3.5–22 min: 100% A to 70% A, 22–25 min: 70% A to 0% A, 25–30 min: 0% A, 30–35 min: 100% A. The flow rate was 1 mL/min, column oven temperature was 40 °C. Purification of products for NMR analysis was performed with a VP 250/8 Nucleodur 100-5 C18 ec column, applying the above method with a flow rate of 4 mL/min. In order to quantify product yields, firstly absence of steroid degradation with strain GHH8 (Supplemental Fig. 2) and similarity of molar extinction coefficients of substrates and products (Supplemental Fig. 3) were verified. For the quantification, the product peak area was divided by the sum of product and substrate areas. The resulting quotient was multiplied with the applied substrate concentration. Molar concentrations of the products were then

converted to mass concentrations (mg/L) based on their molecular weight.

2.5. NMR characterization of the metabolites P3–P5

The NMR spectra were recorded in CDCl₃ and DMSO-d₆ (only P3) with a Bruker DRX 500 or a Bruker Avance 500 NMR spectrometer at 298 K. The chemical shifts were relative to CHCl₃ at δ 7.26 (¹H NMR) and CDCl₃ at δ 77.00 (¹³C NMR) or to DMSO at δ 2.50 (¹H NMR) respectively using the standard δ notation in parts per million. The 1D NMR (¹H and ¹³C NMR) and the 2D NMR spectra (gs-HH-COSY and gs-HSQCED) were recorded using the BRUKER pulse program library.

In contrast to S2, 17α-hydroxyprogesterone, the NMR data of its conversion products P3 and P4 lacked a carbonyl function at C-20. Therefore, resonances of an additional secondary hydroxyl group appeared in both spectra sets indicating that P3 (20α-OH) and P4 (20β-OH) were epimers concerning the hydroxyl group at C-20. Their spectra matched perfectly with those of commercially available reference compounds and were in agreement with data in literature (P3: (Wishart et al., 2013), P4: (Hunter and Carragher, 2003)).

P5 was found to be the 20β-hydroxy derivative of S3, 11-deoxycortisol. Its NMR data were in good accordance with those from literature (Hunter and Carragher, 2003) and its ¹H and ¹³C NMR spectra showed identical resonances when compared with those recorded with an authentic sample.

P3: 17α, 20α-Dihydroxy-4-pregnen-3-one (17α,20α,DiOH-P).

¹H NMR (DMSO-d₆, 500 MHz): δ 0.68 (s, 3xH-18), 0.82 (ddd, J = 12.4, 10.9 and 4.0 Hz, H-9), 0.96 (qd, J = 13.0 and 4.2, H-7a), 1.02 (d, J = 6.3 Hz, 3xH-21), 1.09 (m, H-15a), 1.14 (s, 3xH-19), 1.32 (qd, J = 13.0 and 4.0 Hz, H-11a), 1.41 (dt, J = 13.0 and 4.0 Hz, H-12a), 1.49 (m, H-11b), 1.52 (m, H-8), 1.57 (m, H-16a), 1.58 (m, H-15b), 1.60 (m, H-1a), 1.64 (m, H-12b), 1.70 (m, H-14), 1.78 (m, H-7b), 1.91 (t, J = 11.8 Hz, H-16b), 1.95 (ddd, J = 13.5, 5.1 and 3.5 Hz, H-1b), 2.14 (dt, J = 16.7 and 3.5 Hz, H-2a), 2.24 (ddd, J = 14.5, 3.7 and 2.2 Hz, H-6a), 2.39 (ddd, J = 16.7, 14.5 and 5.1 Hz, H-2b), 2.40 (m, H-6b), 3.52 (s, OH-17), 3.60 (quint, J = 6.3 Hz, H-20), 4.15 (d, J = 6.3 Hz, OH-20), 5.62 (br s, H-4). ¹³C NMR (CDCl₃, 500 MHz): δ 0.77 (s, 3xH-18), 0.97 (ddd, J = 12.4, 10.7 and 3.8 Hz, H-9), 1.10 (dddd, J = 13.7, 12.9, 11.7 and 4.5, H-7a), 1.19 (s, 3xH-19), 1.20 (d, J = 6.3 Hz, 3xH-21), 1.22 (m, H-15a), 1.42 (m, H-11a), 1.58 (m, 2H, H-8 and H-12a), 1.61 (m, H-11b), 1.70 (m, H-12b), 1.71 (m, H-1a), 1.74 (m, H-16a), 1.77 (m, H-14), 1.78 (m, H-15b), 1.86 (dddd, J = 12.9, 5.6, 3.6 and 2.5, H-7b), 2.02 (ddd, J = 13.5, 5.1 and 3.2 Hz, H-1b), 2.06 (m, H-16b), 2.28 (ddd, 14.7, 4.5 and 2.5 Hz, H-6a), 2.35 (dddd, J = 17.1, 5.0, 3.2 and 1.0 Hz, H-2a), 2.39 (m, H-6b), 2.41 (ddd, J = 17.1, 14.5 and 5.1 Hz, H-2b), 3.86 (q, J = 6.3 Hz, H-20), 5.73 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 14.06 (CH₃, C-18), 17.38 (CH₃, C-19), 18.52 (CH₃, C-21), 20.49 (CH₂, C-11), 23.34 (CH₂, C-15), 30.94 (CH₂, C-12), 31.95 (CH₂, C-7), 32.87 (CH₂, C-6), 33.93 (CH₂, C-2), 35.48 (CH, C-8), 35.68 (CH₂, C-1), 37.69 (CH₂, C-16), 38.52 (C, C-10), 45.64 (C, C-13), 50.54 (CH, C-14), 53.26 (CH, C-9), 72.28 (CH, C-20), 85.37 (C, C-17), 123.89 (CH, C-4), 171.22 (C, C-5), 199.53 (C, C-3).

P4: 17α, 20β-Dihydroxy-4-pregnen-3-one (17α,20β,DiOH-P).

¹H NMR (CDCl₃, 500 MHz): δ 0.85 (s, 3xH-18), 0.97 (ddd, J = 12.5, 10.7 and 4.0 Hz, H-9), 1.08 (dddd, J = 13.7, 12.9, 11.7 and 4.5, H-7a), 1.16 (m, H-15a), 1.18 (d, J = 6.3 Hz, 3xH-21), 1.19 (s, 3xH-19), 1.43 (m, H-11a), 1.46 (m, H-16a), 1.55 (m, H-8), 1.58 (m, H-11b), 1.59 (m, H-12a), 1.66 (m, H-16b), 1.70 (m, H-1a), 1.71 (m, H-14), 1.74 (m, 2H, H-12b and H-15b), 1.84 (dddd, J = 13.0, 5.5, 3.5 and 2.5, H-7b), 2.02 (ddd, J = 13.5, 5.1 and 3.1 Hz, H-1b), 2.27 (ddd, 14.7, 4.5 and 2.5 Hz, H-6a), 2.35 (dddd, J = 17.0, 4.8, 3.2 and 1.0 Hz, H-2a), 2.40 (m, H-6b), 2.41 (ddd, J = 17.0, 14.5 and 5.0 Hz, H-2b), 4.04 (q, J = 6.3 Hz, H-20), 5.72 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 15.09 (CH₃, C-18), 17.37 (CH₃, C-19), 18.73 (CH₃, C-21), 20.58 (CH₂, C-11),

23.73 (CH₂, C-15), 31.98 (CH₂, C-7), 32.08 (CH₂, C-12), 32.89 (CH₂, C-6), 33.78 (CH₂, C-16), 33.93 (CH₂, C-2), 35.64 (CH₂, C-1), 35.74 (CH, C-8), 38.56 (C, C-10), 47.13 (C, C-13), 49.60 (CH, C-14), 53.37 (CH, C-9), 70.32 (CH, C-20), 84.97 (C, C-17), 123.79 (CH, C-4), 171.55 (C, C-5), 199.72 (C, C-3).

P5: 17 α , 20 β , 21-Trihydroxy-4-pregnen-3-one (20 β OH-RSS).

¹H NMR (CDCl₃, 500 MHz): δ 0.85 (s, 3xH-18), 0.97 (ddd, J =12.5, 10.7 and 4.2 Hz, H-9), 1.09 (dddd, J =13.8, 12.9, 11.7 and 4.5, H-7a), 1.19 (s, 3xH-19), 1.23 (m, H-15a), 1.44 (m, H-11a), 1.58 (m, 3H, H-8, H-12a and H-16a), 1.62 (m, H-11b), 1.69 (m, H-14), 1.72 (m, H-1a), 1.78 (m, H-16b), 1.79 (m, H-15b), 1.86 (m, H-12b), 1.87 (m, H-7b), 2.04 (ddd, J =13.5, 5.1 and 3.1 Hz, H-1b), 2.28 (ddd, H =14.7, 4.5 and 2.5 Hz, H-6a), 2.37 (dddd, J =17.0, 4.8, 3.2 and 1.0 Hz, H-2a), 2.41 (m, H-6b), 2.44 (ddd, J =17.0, 14.5 and 5.0 Hz, H-2b), 3.77 (m, H-21a), 3.83 (m, 2H, H-20 and H-21b), 5.73 (br s, H-4). ¹³C NMR (CDCl₃, 125 MHz): δ 15.12 (CH₃, C-18), 17.37 (CH₃, C-19), 20.65 (CH₂, C-11), 23.84 (CH₂, C-15), 31.99 (CH₂, C-7), 31.74 (CH₂, C-12), 32.87 (CH₂, C-6), 33.22 (CH₂, C-16), 33.92 (CH₂, C-2), 35.65 (CH₂, C-1), 35.75 (CH, C-8), 38.56 (C, C-10), 47.41 (C, C-13), 49.04 (CH, C-14), 53.39 (CH, C-9), 64.69 (CH₂, C-21), 73.15 (CH, C-20), 85.18 (C, C-17), 123.82 (CH, C-4), 171.46 (C, C-5), 199.72 (C, C-3).

3. Results and discussion

3.1. Gene deletion of the innate steroid hydroxylase CYP106A1

First, in order to assess a potential HSD activity of *B. megaterium* strain GHH1 towards 17OH-P and RSS, the endogenous gene encoding the cytochrome P450 CYP106A1 was deleted, to construct a strain lacking any interfering steroid hydroxylase activity. The remaining innate cytochromes P450 are CYP102A1 (BM-3), CYP109E1 and CYP109A2. Wild-type BM-3 has not been reported to be involved in steroid metabolism. CYP109E1 and CYP109A2 did not show any activity towards common C21-steroids, as measured in our laboratory (unpublished data). *B. megaterium* strain GHH1 is a derivative of strain MS941 and differs only in the knockout of the *upp* gene, which was used as a counter-selection marker during the gene deletion experiments (Table 1 lists all strains used in this study) and still contains intact ORFs encoding the aforementioned cytochromes P450.

Cultivation of strain GHH1 in presence of the model substrate progesterone led to the formation of two products, with a substrate conversion rate of up to 20 % after 24 h. This is in agreement with results of Lee *et al.* (Lee *et al.*, 2015), who describe the formation of a mono- and a dihydroxyprogesterone by CYP106A1 from *B. megaterium* ATCC 14581 (98.8 % amino acid identity with CYP106A1 from GHH1) with progesterone as substrate, as determined by mass spectrometry. We deleted the gene for CYP106A1 by transforming *B. megaterium* with vector pUCTV2_Upp_Δ106A1 (plasmid maps of the backbone vectors and used restriction sites for cloning be found in Supplemental Fig. 1), containing flanking regions of the target gene fused by SOE-PCR, to subsequently undergo the knockout procedure, as described previously (Gerber *et al.*, 2015). The promoter of the CYP106A1 gene was kept intact to avoid negative effects on the transcription of genes downstream of the ORF. The genomic DNA of resulting colonies was screened for the deletion by PCR. A colony containing the deletion was identified and designated as GHH2.

Fig. 1A displays an agarose gel with genomic DNA of GHH1 (lane 1), the knockout vector (lane 2) and genomic DNA of GHH2 (lane 3) as template. The amplified DNA produced with DNA of GHH2 as template is truncated by approximately 1000 bp compared with GHH1 and exhibits the same size as the band produced with the knockout vector as a template indicating the deletion of the CYP106A1 gene. The knockout of CYP106A1 was further verified by carrying out *in vivo* whole cell conversions with progesterone as substrate followed by a RP-HPLC analysis. As evidenced in Fig. 1B, no products of CYP106A1 catalysis were formed with strain GHH2.

3.2. *In vivo* formation of 20 α - and 20 β -reduced products with 17OH-P as substrate

The next step was to investigate the formation of products of potential endogenous HSD activities with the newly obtained strain GHH2. All HSD reactions carried out in this study are summarized in Fig. 2.

The incubation of GHH2 with 17OH-P (S2) as a model substrate resulted in the formation of a major (P3) and a minor (P4) product, as assessed by RP-HPLC (Fig. 3). Product P3 was purified and a subsequent NMR analysis (see Section 2.5) revealed it to be

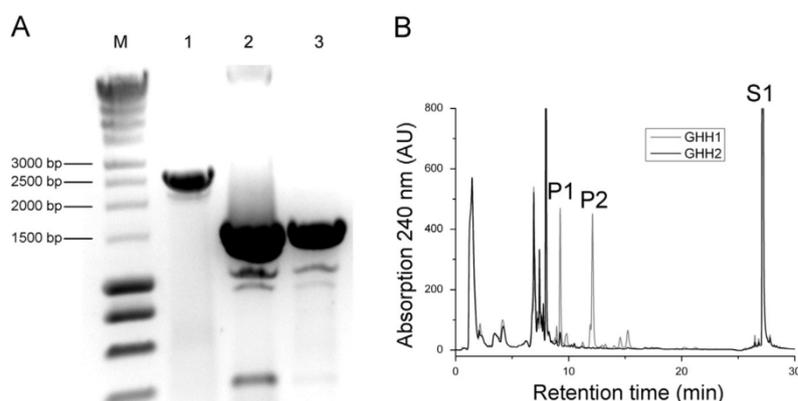


Fig. 1. Agarose gel electrophoresis and *in vivo* steroid conversion confirming CYP106A1 deletion. A: Screening for the knockout was performed by PCR, applying primers 106Afor and 106Brev which were used for the construction of the deletion construct by SOE-PCR (primer name and sequences are listed in Supplemental Table 1). A colony containing the knockout was identified and designated as strain GHH2. Agarose gel electrophoresis displaying amplified DNA with genomic DNA of the parent strain GHH1 (lane 1), knockout vector pUCTV2_Upp_Δ106A1 (lane 2) and genomic DNA of strain GHH2 (lane 3) as template (M: SmartLadder DNA marker (Eurogentec)). The major band in lane 3 is truncated by circa 1000 bp compared with the control DNA from strain GHH1 and corresponds to the band produced with the knockout vector as template (lane 2). B: RP-HPLC chromatogram showing the *in vivo* conversion of progesterone (S1, final concentration 300 μ M) with strains GHH1 (gray) and GHH2 (black) after 24 h. The two CYP106A1 products formed by strain GHH1, P1 and P2, are not present in cultures of GHH2.

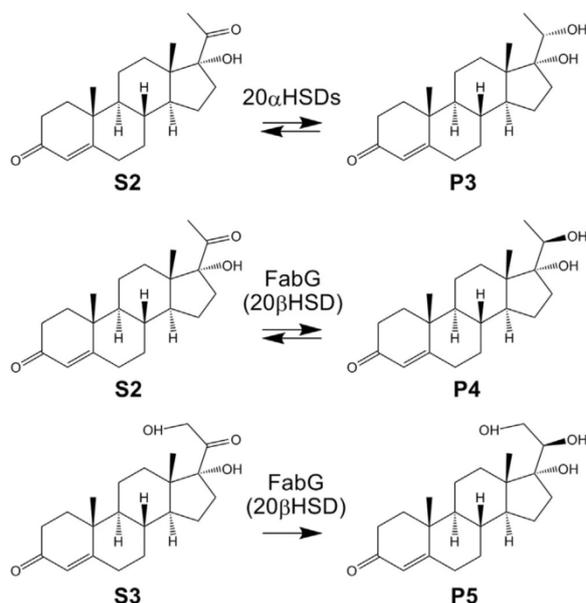


Fig. 2. Summary of the catalyzed HSD reactions described in this study. S2: 17 α -hydroxyprogesterone (17OH-P), P3: 17 α ,20 α -dihydroxy-4-pregnen-3-one (17 α ,20 α -DiOH-P), P4: 17 α ,20 β -dihydroxy-4-pregnen-3-one (17 α ,20 β -DiOH-P), S3: 11-deoxycortisol (RSS), P5: 20 β OH-RSS.

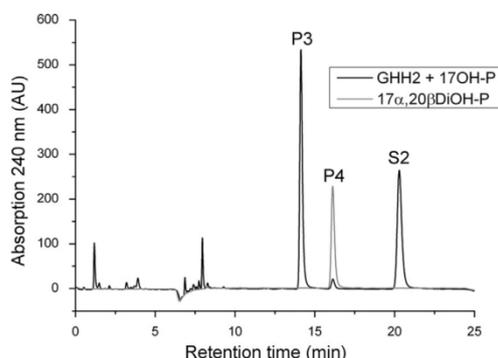


Fig. 3. Identification of 20 α - and 20 β -reduced products of 17OH-P. The HPLC-chromatogram shows the *in vivo* conversion of 17OH-P (S2, final concentration 300 μ M) with *B. megaterium* strain GHH2 (black) after 24 h. 17 α ,20 α -DiOH-P (P3) was formed as a major, 17 α ,20 β -DiOH-P (P4) as a minor product (gray: authentic 17 α ,20 β -DiOH-P standard). Additional peaks appearing during the first 10 min are also present in *B. megaterium* cultures incubated without the substrate 17OH-P (Supplemental Fig. 4).

17 α ,20 α -dihydroxy-4-pregnen-3-one (17 α ,20 α -dihydroxyprogesterone, 4-pregnene-17 α ,20 α -diol-3-one, 17 α ,20 α -DiOH-P), while the minor product P4 exhibited the same retention time as an authentic standard of the target product 17 α ,20 β -dihydroxy-4-pregnen-3-one (17 α ,20 β -dihydroxyprogesterone, 4-pregnene-17 α ,20 β -diol-3-one, 17 α ,20 β -DiOH-P) on the HPLC. The structure of P4 was further verified by NMR analysis (Section 2.5).

3.3. Identification and genomic deletion of 20 α HSDs

To remove the strain's unwanted 20 α HSD activity for the exclusive production of 20 β -reduced steroids, as a next step potential 20 α HSDs were identified from *B. megaterium*, their activities verified by overexpression experiments and their genes finally deleted.

Using MegaBac v9 (<http://megabac.tu-bs.de>), the genome of *B. megaterium* was scanned for potential HSDs applying amino acid sequences of human, bovine and rat 20 α HSDs as references. The open reading frames (ORF) of potential candidates were amplified from genomic DNA and cloned into vector pSMF2.1 under the control of a strong xylose-inducible promoter. *B. megaterium* GHH2 was transformed with the resulting plasmids. The 20 α HSD activity of the different recombinant strains towards the substrate 17OH-P was assessed by HPLC after carrying out *in vivo* whole-cell conversions. The following four ORFs were identified as putative 20 α HSD-encoding genes: BMD_0912 (annotated as oxidoreductase, aldo/keto reductase family), BMD_1068, BMD_1595 and BMD_3715 (all three annotated as a 2,5-diketo-D-gluconic acid reductase A). All four proteins belong to the NAD(P)(H) utilizing superfamily of aldo-keto reductases (AKRs) present in nearly all phyla. These mostly soluble oxidoreductases are involved in the reduction of aldehydes and ketones to primary and secondary alcohols during the phase I metabolism of xenobiotics and act on a broad range of substrates including lipids, prostaglandins, chemical carcinogens and steroids (Penning, 2015). Fig. 4A displays an exemplary HPLC chromatogram from the *in vivo* conversion of 17OH-P with strain GHH2 transformed with either pSMF2.1_0912 (expressing BMD_0912) or the empty backbone vector pSMF2.1 as a control after 20 hours. In comparison with the control strain, formation of 17 α ,20 α -DiOH-P is significantly increased while the substrate is almost completely depleted in the culture of the 20 α HSD overexpressing strain. The time course of the product formation for all 20 α HSD overexpressing strains compared with the control strain pictured in Fig. 4B demonstrates that all 20 α HSDs are able to carry out both the reduction and oxidation reaction. 17 α ,20 α -DiOH-P peak areas were normalized through multiplication by the quotients of the optical densities of the control strain, harboring the empty backbone vector pSMF2.1, and each 20 α HSD-overexpressing strain to account for negative or positive effects of the produced proteins on cell growth, which could influence 17 α ,20 α -DiOH-P production. The formation of 17 α ,20 α -DiOH-P is peaking during the first 20 h with the recombinant strains, then declines after being oxidized again, reaching a concentration comparable with the control strain. Oxidase or reductase activities of AKRs is determined by the prevailing ratio of NAD(P)H to NAD(P)⁺ in the cells (Rizner et al., 2003), which likely leads to the decrease in concentration of 17 α ,20 α -DiOH-P after 6 h in the cultures, due to NAD(P)H depletion. Each 20 α HSD was deleted stepwise in the same way as described for the knockout of CYP106A1 and the *in vivo* activity assessed after each deletion step. The final strain was designated as GHH8 (Table 1). Fig. 4C summarizes the allelic state of strain GHH8 in comparison with GHH2, as verified by PCR. All four 20 α HSDs have been truncated by approximately 800 bp. Whole-cell conversions with GHH2 and GHH8 were carried out, applying 17OH-P as substrate. After 72 h, no 17 α ,20 α -DiOH-P formation could be observed with the 20 α HSD-devoid strain (Fig. 4D). Lastly, strain GHH8 exhibited no difference in growth rate compared with the parental strain MS941, one of the most frequently applied and commercially available *B. megaterium* strains for biotechnological purposes (Fig. 4E), indicating that the combined gene deletions had no negative effects on the *B. megaterium* metabolism, at least under laboratory conditions.

3.4. Identification of *B. megaterium* 20 β HSD FabG and application for the production of 20 β -reduced steroids

After construction of strain GHH8 lacking any steroid hydroxylase and 20 α HSD activity, the responsible 20 β HSD was identified from *B. megaterium*'s genome and overexpressed to establish a whole-cell system for the production of 20 β -reduced steroids.

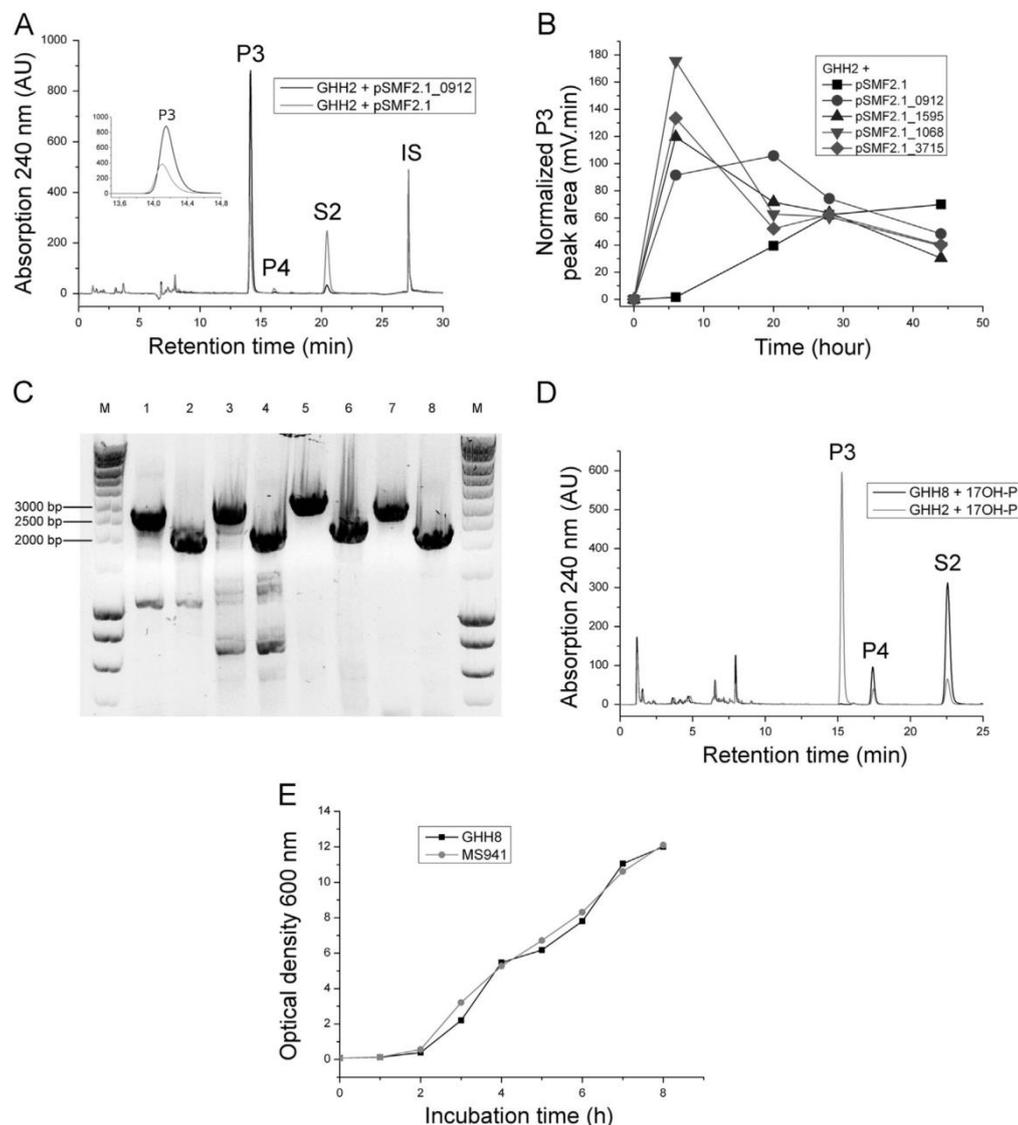


Fig. 4. Identification and genomic deletion of *B. megaterium* 20 α HSDs. A: Exemplary HPLC-chromatogram displaying the in vivo conversion of 17OH-P (S2) with strain GHH2 either transformed with pSMF2.1_0912 (black), encoding the 20 α HSD BMD_0912, or the empty vector pSMF2.1, as a control. 17 α ,20 α DiOH-P (P3) formation is significantly increased with the BMD_0912 overexpressing strain (inset: zoomed in P3 peak, IS: internal standard progesterone). B: Time course of 17 α ,20 α DiOH-P formation with the control strain and the 20 α HSD overexpressing strains. After 6 hours, the 20 α HSD overexpressing strains showed drastically increased product yields compared with the control strain. Each 20 α HSD catalyzed both the reduction and oxidation reaction, leading to a decline of 17 α ,20 α DiOH-P yield at later cultivation stages. Product areas were normalized to the optical density of the control strain, taking extraction loss into account by use of the internal standard. C: Agarose gel electrophoresis after PCR with genomic DNA of strain GHH2 and GHH8 applying primers that flank the locus of each specific 20 α HSD (0912Afor/0912Brev, 1595Afor/1595Brev, 1068Afor/1068Brev and 3715Afor/3715Brev). All four 20 α HSDs are truncated by ca. 800 bp in strain GHH8 (lanes 1, 3, 5, and 7: genomic DNA from GHH2 as template; lane 2: genomic DNA GHH5 (Δ BMD_0912); lane 4: genomic DNA GHH6 (Δ BMD_0912, Δ BMD_1068); lane 6: genomic DNA GHH7 (Δ BMD_0912, Δ BMD_1068, Δ BMD_3715) and lane 8: genomic DNA GHH8 (Δ BMD_0912, Δ BMD_1068, Δ BMD_3715, Δ BMD_1595); M: SmartLadder DNA marker). D: HPLC-chromatogram showing the conversion of 17OH-P (S2) with strain GHH2 (gray) and GHH8 (black) after 72 h. No 17 α ,20 α DiOH-P (P3) was produced with strain GHH8. E: Growth curve of strain GHH8 (black) and the parental strain MS941 (gray). In contrast to strain GHH2, MS941 still contains intact *cyp106A1* and *upp* ORFs. The combined gene deletions had no effect on growth of GHH8 under the applied cultivation conditions.

Similar genes to eukaryotic 20 β HSDs could not be found in the genome of *B. megaterium*. However, applying the amino acid sequence of a 3- α -(or 20- β)-HSD (Rv2002) from *Mycobacterium tuberculosis* as a reference (Yang et al., 2003), potential candidate genes could be obtained. These ORFs were again amplified from genomic *B. megaterium* DNA and cloned into pSMF2.1. The 20 α HSD activity-free strain GHH8 was transformed with the resulting plasmids and the in vivo activity towards 17OH-

P was measured by HPLC. FabG (BMD_4208; 3-oxoacyl-(acyl-carrier-protein) reductase) was identified as the responsible enzyme with 20 β HSD activity. FabG belongs to the family of short chain dehydrogenases involved in the type II biosynthesis of saturated and unsaturated fatty-acids, reducing the β -ketoacyl group to a β -hydroxy group during chain elongation. These enzymes exhibits a Rossmann-fold binding domain and can accept NADP(H) and/or NAD(H) as cofactor (Javidpour et al., 2014). Enzymes of that type

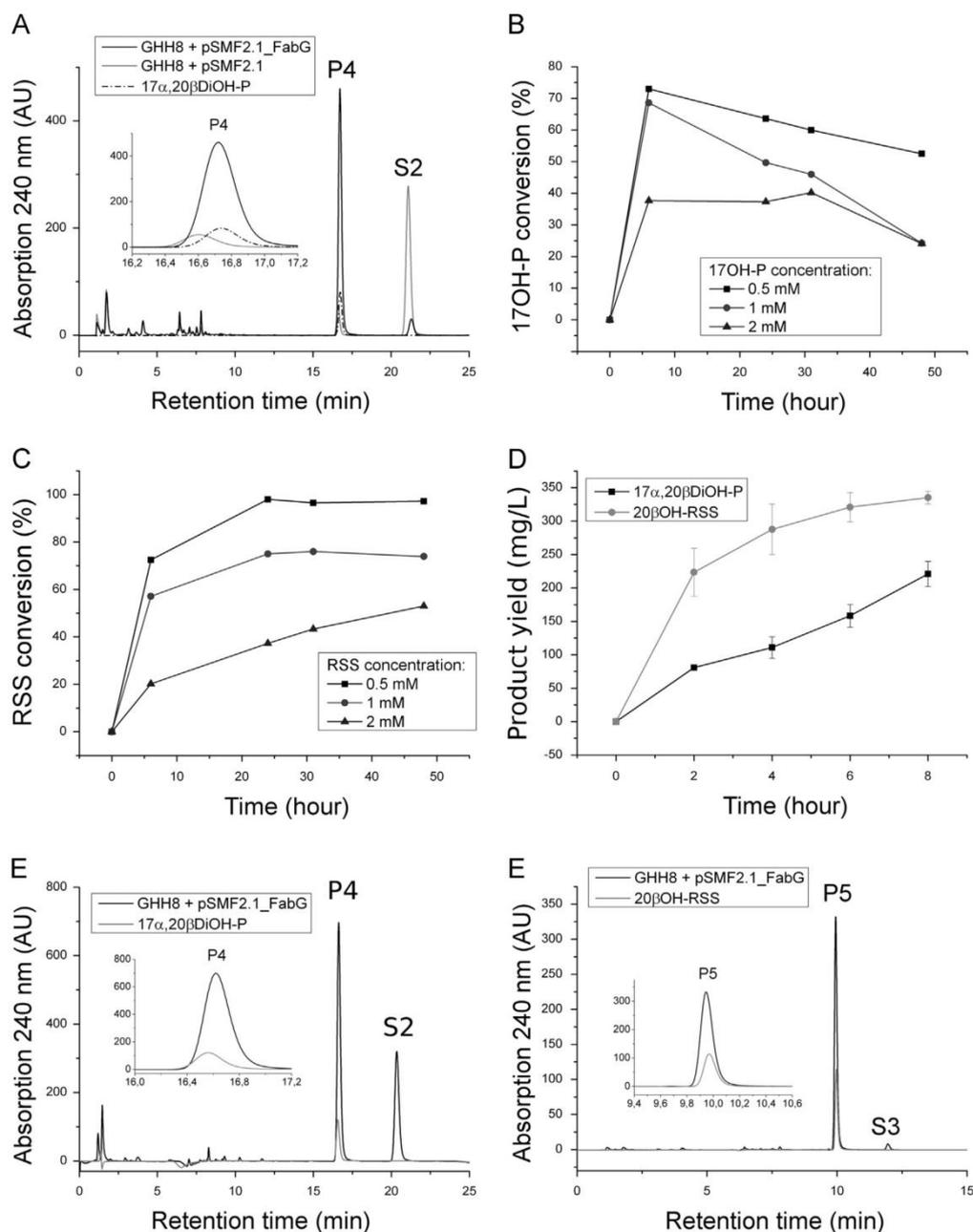


Fig. 5. Identification and overexpression of *B. megaterium* 20 β HSD FabG for 20 β -reduced steroid hormone production. A: HPLC-chromatogram showing the *in vivo* conversion of 17OH-P (S2, final concentration 300 μ M) to 17 α ,20 β DiOH-P (P4) with strain GHH8 transformed with pSMF2.1_FabG (black) or the empty pSMF2.1 vector (gray). The substrate was added simultaneously to protein induction. After 6 hours, product yield is strongly increased in the strain expressing FabG. B and C: Conversion of varying concentrations of 17OH-P and RSS after different time points (single experiments). Application of 1 mM of both substrates led to the highest absolute product yields within 10 h. D: Product yields at different time points using 17OH-P and RSS as substrates with a final concentration of 1 mM (triplicate biological experiments, error bars indicate standard deviation). Maximum yields of 17 α ,20 β DiOH-P and 20 β OH-RSS were achieved after 8 hours, with 220.9 and 335.25 mg/L, respectively. E: HPLC chromatogram showing the conversion of 17OH-P (S2, final concentration 1 mM) to 17 α ,20 β DiOH-P (P4) after 8 h. F: HPLC chromatogram showing the conversion of RSS (S3, final concentration 1 mM) to 20 β OH-RSS (P5) after 8 h.

are also involved in bacterial PHB synthesis by reducing acetyl-CoA and in the 20 β oxidation/reduction of steroids (Yang et al., 2003). After overexpression of FabG in GHH8, the 17 α ,20 β DiOH-P yield was drastically increased compared with strain GHH8 harboring the empty pSMF2.1 plasmid (Fig. 5A). In addition, RSS was found to be a substrate, being converted to

20 β OH-RSS (see Section 2.5 for NMR data). To the best of our knowledge, these steroids have not been described as substrates for any of the FabG isoform so far. Cortisol and cortisone were also identified as substrates of FabG (Supplemental Fig. 5, NMR data of the purified steroids in Supplemental Table 2), however a physiological function of the resulting products has not been

described in literature. Optimal substrate concentrations and conversion times were determined for 17OH-P and RSS (Fig. 5B and C). FabG could both reduce 17OH-P and oxidize 17 α ,20 β DiOH-P (Fig. 5B), RSS was reduced to 20 β OH-RSS with no reverse oxidation reaction of 20 β OH-RSS taking place (Fig. 5C). Whole-cell conversions with both substrates were carried out using a final concentration of 1 mM, taking samples within 10 hours (Fig. 5D). Maximum yields of 221 mg/L 17 α ,20 β DiOH-P (66.6 % relative substrate conversion) and 335 mg/L 20 β OH-RSS (96.4 %) were achieved after 8 hours. Formation rates during the first hour of incubation were as high as 40.4 mg/L/h 17 α ,20 β DiOH-P and 111.8 mg/L/h 20 β OH-RSS, as determined from Fig. 5D. Fig. 5E and F display HPLC-chromatograms for both conversions after 8 h, with the respective product standards.

4. Conclusions

17 α ,20 β DiOH-P and 20 β OH-RSS are important teleost steroid hormones and pheromones involved in oocyte maturation, spermiation, initiation of meiosis and increase of sperm motility. Their application in aquacultures has been shown to be beneficial for increasing fish spawning rates. However, both steroids are very expensive due to the lack of efficient chemical or microbial syntheses. We have established a *B. megaterium* based whole-cell system for the high-yield production of these progestogens without side-product formation. Both steroid hormones can be rapidly produced using cheap steroid precursors. Due to the high difference in price between substrates and products (compare vendor prices (Sigma-Aldrich, Steraloids) and Ouedraogo et al., 2013), carrying out the *in vivo* conversions using simple-shake flask cultivation already results in a profitable bioprocess, even when product loss from further purification steps can be expected, thus providing a good basis for large-scale fermentations.

The product of 20 α HSD activity, 17 α ,20 α -DiOH-P, occurs in higher concentrations in the human plasma than 17OH-P (Whitworth et al., 1983). It has been shown to have a hypertensive effect in sheep when injected concurrently with major corticosteroids (Butkus et al., 1982), but exhibits low affinities for classical mineralocorticoid or glucocorticoid receptors. In addition, it has been described as a competitive inhibitor of rat CYP17 lyase activity. Moreover, 17 α ,20 α DiOH-P appears to be involved in spermiation in several species of fish (Asahina et al., 1990; Tan et al., 1995). The lack of information on the specific physiological role of 17 α ,20 α DiOH-P makes it also an interesting target for a biotechnological production to study its function in more detail.

Overall, the biocatalyst is more efficient, environmentally friendly and less expensive than previously described chemical syntheses of both 20 β -reduced steroid hormones, which could open the way for the broad application of these progestogens in aquacultures.

Competing interests

The authors declare no competing interest.

Authors' contributions

A.G. designed research. A.G. performed all experiments, except for purification of 17 α ,20 α DiOH-P, carried out by P.H. M.M. provided plasmids and reagents. J.Z. performed NMR analyses. A.G. wrote the manuscript. M.M. and R.B. participated in the interpretation and discussion of experimental results and writing of the manuscript.

Acknowledgments

We would like to thank Friedhelm Meinhardt (Münster) for providing plasmid pUCTV2.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.ymben.2016.02.010>.

References

- Asahina, K., Barry, T.P., Aida, K., Fusetani, N., Hanyu, I., 1990. Biosynthesis of 17 α ,20 α -dihydroxy-4-pregnen-3-one from 17 α -hydroxyprogesterone by spermatozoa of the common carp, *Cyprinus carpio*. *J. Exp. Zool.* 255 (2), 244–249.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124 (1), 128–145.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98 (14), 6185–6203.
- Biedendieck, R., Borgmeier, C., Bunk, B., Stammen, S., Scherling, C., Meinhardt, F., et al., 2011. Systems biology of recombinant protein production using *Bacillus megaterium*. *Methods Enzymol.* 500, 165–195.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto-beta-boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93 (3), 1135–1146.
- Butkus, A., Congiu, M., Scoggins, B.A., Coghlan, J.P., 1982. The affinity of 17 alpha-hydroxyprogesterone and 17 alpha, 20 alpha-dihydroxyprogesterone for classical mineralocorticoid or glucocorticoid receptors. *Clin. Exp. Pharmacol. Physiol.* 9 (2), 157–163.
- Dong, H., Zhang, D., 2014. Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Factories* 13 (1), 1–11.
- Donova, M.V., Egorova, O.V., 2012. Microbial steroid transformations: current state and prospects. *Appl. Microbiol. Biotechnol.* 94 (6), 1423–1447.
- Ehrhardt, M., Gerber, A., Hannemann, F., Bernhardt, R., 2016. Expression of human CYP27A1 in *B. megaterium* for the efficient hydroxylation of cholesterol, vitamin D3 and 7-dehydrocholesterol. *J. Biotechnol.* 218, 34–40.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Factories.* 14 (1), 015–0300.
- Haider, S., Rao, N.V., 1994. Induced spawning of maturing Indian catfish, *Clarias batrachus* (L.), using low doses of steroid hormones and salmon gonadotropin. *Aquacult. Res.* 25 (4), 401–408.
- Hong, W.-S., Chen, S.-X., Zhang, Q.-Y., Zheng, W.-Y., 2006. Sex organ extracts and artificial hormonal compounds as sex pheromones to attract broodfish and to induce spawning of Chinese black sleeper (*Bostrichthys sinensis* Lacépède). *Aquacult. Res.* 37 (5), 529–534.
- Hunter, A.C., Carragher, N.E., 2003. Flexibility of the endogenous progesterone lactonisation pathway in *Aspergillus tamarii* KITA: transformation of a series of corticosteroid analogues. *J. Steroid Biochem. Mol. Biol.* 87 (4–5), 301–308.
- Hunter, D.J.B., Behrendorf, J.B.Y.H., Johnston, W.A., Hayes, P.Y., Huang, W., Bonn, B., et al., 2011. Facile production of minor metabolites for drug development using a CYP3A shuffled library. *Metab. Eng.* 13 (6), 682–693.
- Javidpour, P., Pereira, J.H., Goh, E.B., McAndrew, R.P., Ma, S.M., Friedland, G.D., et al., 2014. Biochemical and structural studies of NADH-dependent FabG used to increase the bacterial production of fatty acids under anaerobic conditions. *App. Environ. Microbiol.* 80 (2), 497–505.
- King, H.R., Young, G., 2001. Milt production by non-spermiating male Atlantic salmon (*Salmo salar*) after injection of a commercial gonadotropin releasing hormone analog preparation, 17 α -hydroxyprogesterone or 17 α ,20 β -dihydroxy-4-pregnen-3-one, alone or in combination. *Aquaculture* 193 (1–2), 179–195.
- Kiss, F.M., Khatri, Y., Zapp, J., Bernhardt, R., 2015. Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism. *FEBS Lett.* 589 (18), 2320–2326.
- Kiss, F.M., Schmitz, D., Zapp, J., Dier, T.K., Volmer, D.A., Bernhardt, R., 2015. Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* - identification of a novel 11-oxidase activity. *Appl. Microbiol. Biotechnol.* 24, 24.
- Koshcheenko, K.A., Krasnova, L.A., Garsiia-Rodriges, L.K., Surovtsev, V.I., Skriabin, G. K., 1976. 20alpha- and 20beta-reduction of steroids by immobilized enzyme preparations and *Bacillus megaterium* cells. *Prikl Biokhim Mikrobiol* 12 (3), 322–326.
- Kovganko, N.V., Kashkan, Z.N., Shkumatov, V.M., 2001. Simple Synthesis of 17 α ,20 β -Dihydroxypregn-4-en-3-one. *Chem. Nat. Compd.* 37 (1), 55–56.
- Lee, G.Y., Kim, D.H., Kim, D., Ahn, T., Yun, C.H., 2015. Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Arch. Pharm. Res.* 38 (1), 98–107.

- Lee, W.K., Yang, S.W., 2002. Relationship between ovarian development and serum levels of gonadal steroid hormones, and induction of oocyte maturation and ovulation in the cultured female Korean spotted sea bass *Lateolabrax maculatus* (Jeom-nong-eeo). *Aquaculture* 207 (1–2), 169–183.
- Martins Pinheiro, M.F., Guimarães de Souza, S.M., Gil Barcellos, L.J., 2003. Exposure to 17 α ,20 β -dihydroxy-4-pregnen-3-one changes seminal characteristics in Nile tilapia, *Oreochromis niloticus*. *Aquacult. Res.* 34 (12), 1047–1052.
- Miwa, T., Yoshizaki, G., Naka, H., Nakatani, M., Sakai, K., Kobayashi, M., et al., 2001. Ovarian steroid synthesis during oocyte maturation and ovulation in Japanese catfish (*Silurus asotus*). *Aquaculture* 198 (1–2), 179–191.
- Nagahama, Y., Yamashita, M., 2008. Regulation of oocyte maturation in fish. *Dev. Growth Differ.* 50 (1), 13.
- Nicolaou, S.A., Gaida, S.M., Papoutsakis, E.T., 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. *Metab. Eng.* 12 (4), 307–331.
- Ohta, H., Kagawa, H., Tanaka, H., Okuzawa, K., Hirose, K., 1996. Changes in fertilization and hatching rates with time after ovulation induced by 17, 20 β -dihydroxy-4-pregnen-3-one in the Japanese eel, *Anguilla japonica*. *Aquaculture* 139 (3–4), 291–301.
- Ouedraogo, Y.P., Huang, L., Torrente, M.P., Proni, G., Chadwick, E., Wehmschulte, R.J., et al., 2013. A direct stereoselective preparation of a fish pheromone and application of the zinc porphyrin tweezer chiroptical protocol in its stereochemical assignment. *Chirality* 25 (9), 575–581.
- Penning, T.M., 2015. The ald-keto reductases (AKRs): overview. *Chem. Biol. Interact.* 234, 236–246.
- Richhardt, J., Larsen, M., Meinhardt, F., 2010. An improved transconjugation protocol for *Bacillus megaterium* facilitating a direct genetic knockout. *Appl. Microbiol. Biotechnol.* 86.
- Rizner, T.L., Lin, H.K., Peehl, D.M., Steckelbroeck, S., Bauman, D.R., Penning, T.M., 2003. Human type 3 3 α -hydroxysteroid dehydrogenase (ald-keto reductase 1C2) and androgen metabolism in prostate cells. *Endocrinology* 144 (7), 2922–2932.
- Schmitz, D., Zapp, J., Bernhardt, R., 2014. Steroid conversion with CYP106A2 – production of pharmaceutically interesting DHEA metabolites. *Microb. Cell Factories* 13 (81), 1475–2859.
- Scott, A.P., Sumpter, J.P., Stacey, N., 2010. The role of the maturation-inducing steroid, 17,20 β -dihydroxypregn-4-en-3-one, in male fishes: a review. *J. Fish Biol.* 76 (1), 183–224.
- Shkumatov, V.M., Usova, E.V., Radyuk, V.G., Kashkan, Z.N., Kovganko, N.V., Juretzek, T., et al., 2003. Oxidation of 17 α ,20 β - and 17 α ,20 α -dihydroxypregn-4-en-3-ones, side products of progesterone biotransformation with recombinant microorganisms expressing cytochrome P-45017 α . *Russ. J. Bioorg. Chem.* 29 (6), 581–587.
- Tan, A.M.C., Lee, S.T.L., Kime, D.E., Chao, T.M., Lim, H.S., Chou, R., et al., 1995. 17 α , 20 α -dihydroxy-4-pregnen-3-one, not its 20 β isomer, is produced from 17 α -hydroxyprogesterone by spermatozoa of secondary male groupers (*Epinephelus tauvina*) derived from females implanted with 17 α -methyltestosterone. *J. Exp. Zool.* 271 (6), 462–465.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.D., et al., 2007. *Bacillus megaterium* – from simple soil bacterium to industrial protein production host. *Appl. Microbiol. Biotechnol.* 76 (5), 957–967.
- Whitworth, J.A., Butkus, A., Coghlan, J.P., Denton, D.A., Saines, D., Scoggins, B.A., 1983. Plasma 4-pregnen-17 alpha, 20 alpha-diol-3-one (17 alpha, 20 alpha-dihydroxyprogesterone) and 17 alpha-hydroxyprogesterone in man. *Acta Endocrinol.* 102 (2), 271–276.
- Winter, J., Cerone-McLernon, A., O'Rourke, S., Ponticorvo, L., Bokkenheuser, V.D., 1982. Formation of 20 β -dihydrosteroids by anaerobic bacteria. *J. Steroid Biochem.* 17 (6), 661–667.
- Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., et al., 2013. HMDB 3.0 – the human metabolome database in 2013. *Nucleic Acids Res.* 41, 17.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42 (6), 871–877.
- Yamamoto, Y., Yatabe, T., Higuchi, K., Takeuchi, Y., Yoshizaki, G., 2015. Improvement of ovulation induction by additive injection of 17,20 β -dihydroxy-4-pregnen-3-one after human chorionic gonadotropin administration in a pelagic egg spawning marine teleost, nibe croaker *Nibea mitsukurii* (Jordan & Snyder). *Aquacult. Res.* 46 (6), 1323–1331.
- Yang, J.K., Park, M.S., Waldo, G.S., Suh, S.W., 2003. Directed evolution approach to a structural genomics project: Rv2002 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* 100 (2), 455–460.
- Yao, K., Wang, F.-Q., Zhang, H.-C., Wei, D.-Z., 2013. Identification and engineering of cholesterol oxidases involved in the initial step of sterols catabolism in *Mycobacterium neoaurum*. *Metab. Eng.* 15, 75–87.
- Yao, K., Xu, L.-Q., Wang, F.-Q., Wei, D.-Z., 2014. Characterization and engineering of 3-ketosteroid- Δ^1 -dehydrogenase and 3-ketosteroid-9 α -hydroxylase in *Mycobacterium neoaurum* ATCC 25795 to produce 9 α -hydroxy-4-androstene-3,17-dione through the catabolism of sterols. *Metab. Eng.* 24, 181–191.

Supplementary material

Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one

Adrian Gerber, Mohammed Milhim, Philip Hartz, Josef Zapp and Rita Bernhardt*

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

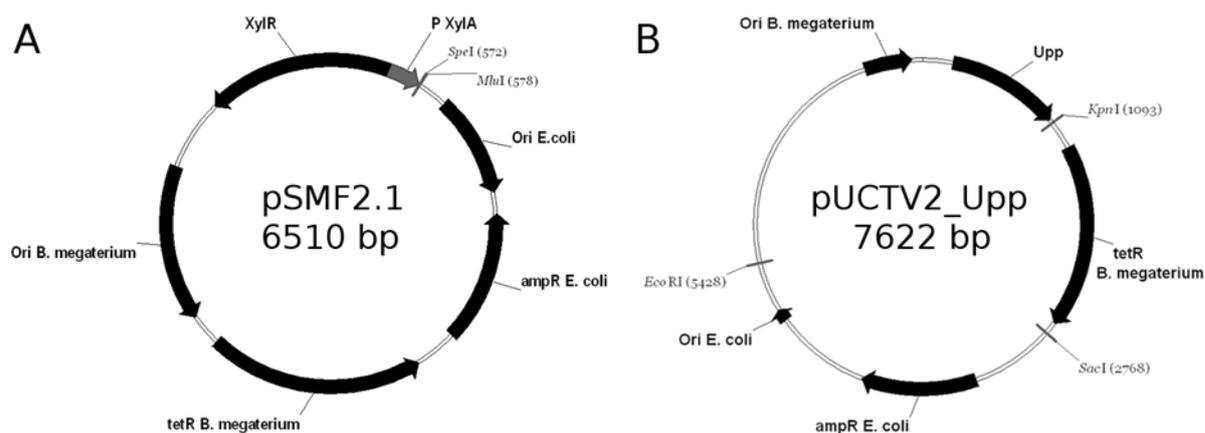
*Address correspondence to:

Prof. Dr. Rita Bernhardt, Institute of Biochemistry, Campus B 2.2, Saarland University, 66123 Saarbrücken (Germany)

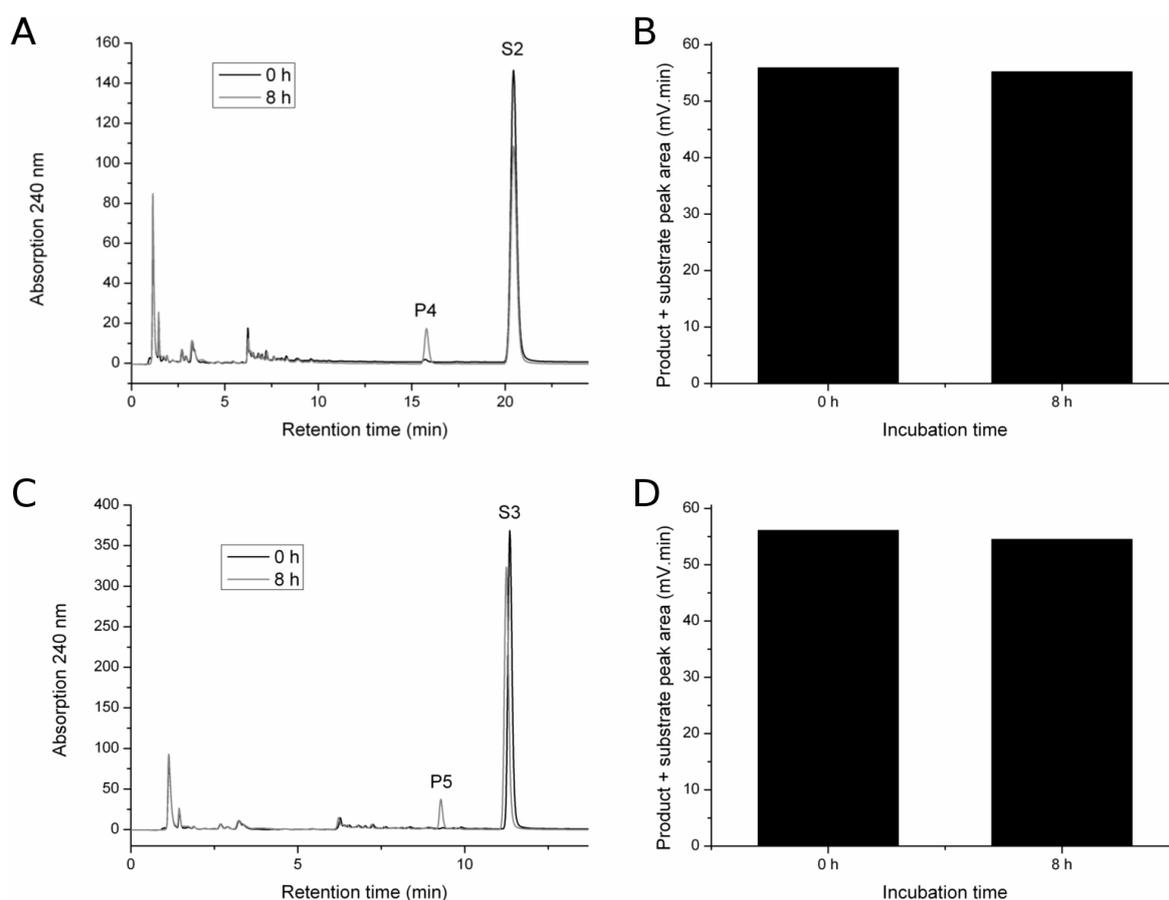
E-mail: ritabern@mx.uni-saarland.de

Phone: +49 681 302 4241

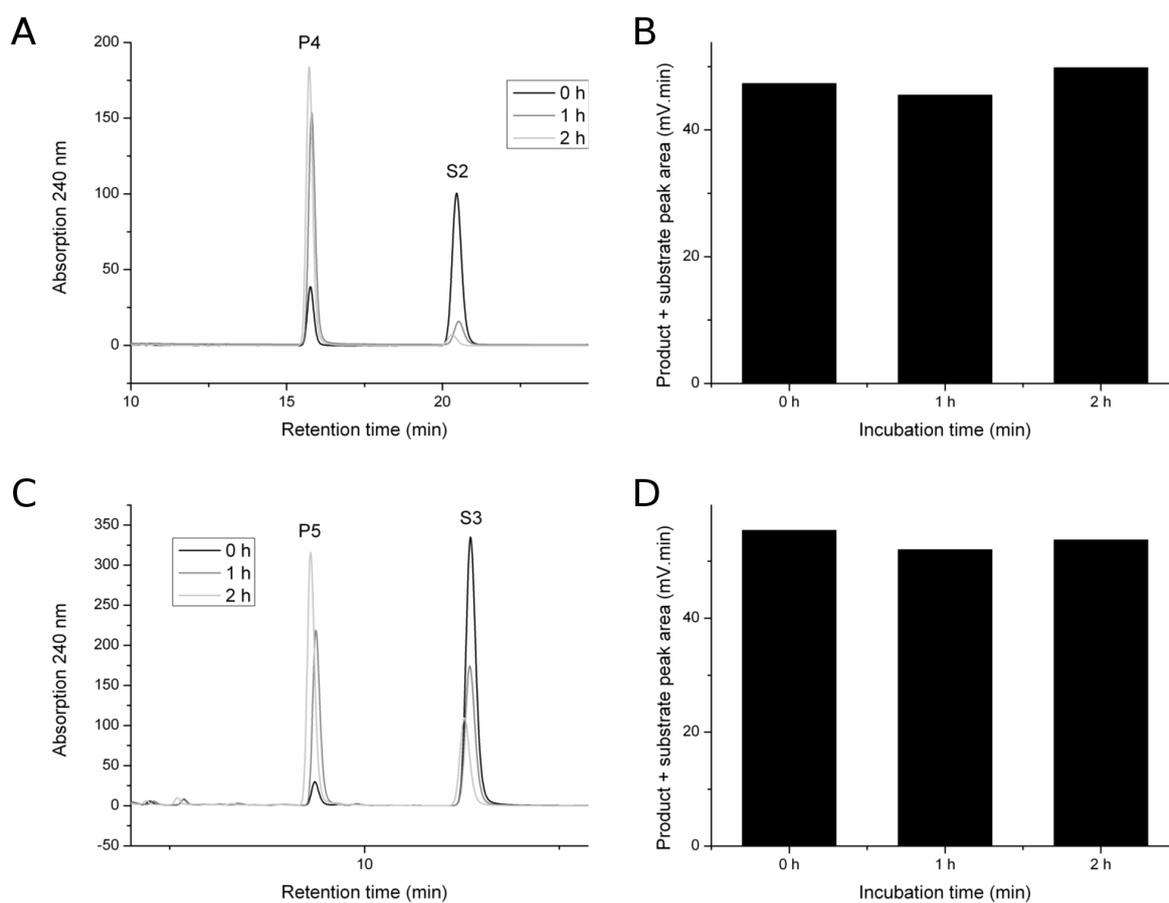
Fax: +49 681 302 4739



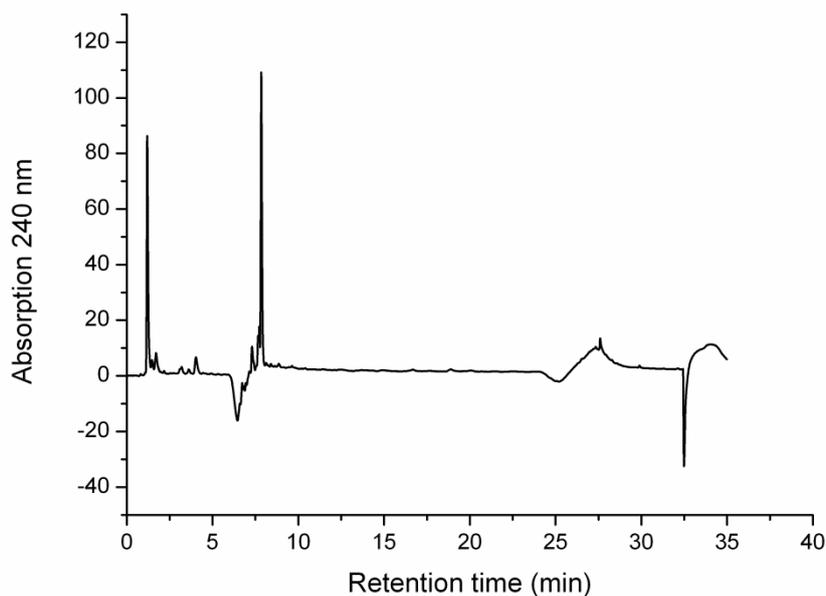
Supplemental Fig. 1. Map of backbone vectors. Plasmid pSMF2.1 was applied for overexpression experiments. SpeI and MluI restriction sites were used for cloning of BMD_0912, BMD_1595, BMD_1068, BMD_3715 and FabG resulting in plasmids pSMF2.1_0912, pSMF2.1_1595, pSMF2.1_1068, pSMF2.1_3715 and pSMF2.1_FabG, respectively (XylR: xylose repressor, P XylA: xylose isomerase promoter, Ori: origin of replication, ampR: ampicillin resistance gene, tetR: tetracycline resistance gene). Plasmid pUCTV2_Upp was applied for gene deletions. Fused flanking regions of *cyp106A1* and *BMD_3715* were cloned using the EcoRI restriction site, SacI site was used for flanking regions of *BMD_0912* and KpnI was used for the flanking regions of *BMD_1595* and *BMD_1068* resulting in plasmids pUCTV2_Upp_Δ106A1, pUCTV2_Upp_Δ3715, pUCTV2_Upp_Δ0912, pUCTV2_Upp_Δ1595 and pUCTV2_Upp_Δ1068, respectively (Upp: uracil phosphoribosyltransferase).



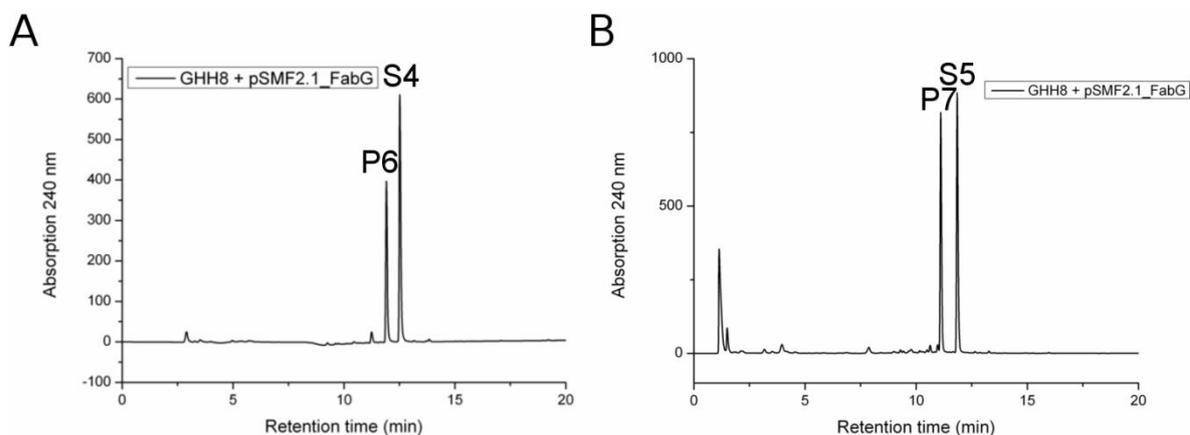
Supplemental Fig. 2. Absence of steroid degradation activity with *B. megaterium* strain GHH8. Main cultures were grown for 20 h, then substrates were added to a final concentration of 200 μ M. A and B: Addition of 17OH-P (S2). C and D: Addition of RSS (S3). No difference in the sum of substrate and product peak areas between 0 and 8 h could be observed for both substrates. 17 α ,20 α DiOH-P (P4) and 20 β OH-RSS (P5) are formed through the action of the chromosomally encoded FabG.



Supplemental Figure 3. Determination of the sum of product and substrate peak areas during 20 β HSD catalysis. A and B: Conversion of 200 μ M 17OH-P (S2) to 17 α ,20 β DiOH-P (P4). C and D: Conversion of 200 μ M RSS (S3) to 20 β OH-RSS (P5). During both reactions, the overall product and substrate peak areas remain constant, indicating that the substrates and their products exhibit similar molar extinction coefficients.



Supplemental Fig. 4. Cultivation of strain GHH8 medium with TB-medium in absence of a substrate after 24 hours. Peaks detected during the first 10 minutes are not related to conversions of the steroidal substrates.



Supplemental Fig. 5. *In-vivo* conversion of cortisol and cortisone by FabG. A: Conversion of cortisol (S, final concentration 300 μ M) to 20 β OH-cortisol (P) after 24 hours. The substrate was added simultaneously to protein induction. B: Conversion of cortisone (S, final concentration 300 μ M) to 20 β OH-cortisone (P) after 24 hours. The substrate was added simultaneously to protein induction.

Supplemental Table 1: List of primers used in this study.

Primer name	Sequence (5' -> 3')	Description
0912for	TATCAACTAGT TAAATCAAGGAGG TGAATGTACA <u>ATGG</u> AAAACCTTAC AGTCAACTAC	Forward primer for BMD_0912 amplification (SpeI site (<i>italic</i>), followed by ribosomal binding site (RBS, bold) and start codon of ORF (<u>underlined</u>))
0912rev	TATCAACGCGT <u>TTT</u> AAAATCAAAG TTATCTGGATC	Reverse primer for BMD_0912 amplification (MluI site, stop codon of ORF (<u>underlined</u>))
1595for	TATCAACTAGT TAAATCAAGGAGG TGAATGTACA <u>ATGA</u> ATATTGTTAC ATTAAACAA	Forward primer for BMD_1595 amplification (SpeI site, RBS, start codon)
1595rev	TATCAACGCGT <u>TTT</u> ATCGGACGTTT ATGTCACTTGGG	Reverse primer for BMD_1595 amplification (MluI site, stop codon)
1068for	TATCAACTAGT TAAATCAAGGAGG TGAATGTACA <u>ATG</u> AGTAATCATT GCAAGATACAGT	Forward primer for BMD_1068 amplification (SpeI site, RBS, start codon)
1068rev	TATCAACGCGT <u>TTT</u> AAAATCAAAA TTGTCCGGATC	Reverse primer for BMD_1068 amplification (MluI site, stop codon)
3715for	TATCAACGCGT <u>TTT</u> AAAATCAAAG TTGTCAAGGATCT	Forward primer for BMD_3715 (MluI site, stop codon)
3715rev	TATCAACTAGT TAAATCAAGGAGG TGAATGTACA <u>ATG</u> ATGAAAATTT ACAGGATACAG	Reverse primer for BMD_3715 (SpeI site, RBS, start codon)
Fabgfor	TATCAACGCGT <u>TTT</u> ACATCACCATT CCGCCGTCAACG	Forward primer for FabG amplification (MluI, stop codon)
Fabgrev	TATCAACTAGT TAAATCAAGGAGG TGAATGTACA <u>ATG</u> TTACAAGGGA AAGTTGCGGTTG	Reverse primer for FabG amplification (SpeI site, RBS, start codon)
106Afor	TATCAGAATTCTCTGTGATCATTC CCATTACTCGATTTTCT	Forward primer for amplification of flanking region downstream of CYP106A1 (EcoRI site)
106Arev	CATGTTAAACAAGTCTTGAGCGA CTACGAAGGCCTTTTCTCATATCG AACCATTTGAAG	Reverse primer for amplification of flanking region downstream of CYP106A1
106Bfor	CTTCAAATGGTTTCGATATGAGAAA AGGCCTTCGTAGTCGCTCAAGAC TTGTTTAAACATG	Forward primer for amplification of flanking region upstream of

106Brev	TATCAGAATTCGGTTAGCAA TATCACGTTTGATCTTAAGAATGA	CYP106A1 Reverse primer for amplification of flanking region upstream of CYP106A1 (EcoRI site)
0912Afor	TATCAGAGCTCTACGTCATATATT CTCTCTTACAGG	Forward primer for amplification of flanking region downstream of BMD_0912 (Sacl site)
0912Arev	CAAATACGTCTGCATTTTGGATAA TATAAACACCTAAACCTAACCAAG GC	Reverse primer for amplification of flanking region downstream of BMD_0912
0912Bfor	GCCTTGGTTAGGTTTAGGTGTTTA TATTATCCAAAATGCAGACGTATT TG	Forward primer for amplification of flanking region upstream of BMD_0912
0912Brev	TATCAGAGCTCGCTGTAGCTTTCT CAATTTCTTCTT	Reverse primer for amplification of flanking region upstream of BMD_0912 (Sacl site)
1595Afor	TATCAGGTACCGTTCGGGTATTG TTAATTTCGATTGCAC	Forward primer for amplification of flanking region downstream of BMD_1595 (KpnI site)
1595Arev	CGTTCATGTCACTTGGGTGCGGA TTTTTGACACCAGACCAACTGGTA A	Reverse primer for amplification of flanking region downstream of BMD_1595
1595Bfor	TTACCAGTTGGTCTGGTGTCAAAA ATCCGCACCCAAGTGACATGAAC G	Forward primer for amplification of flanking region upstream of BMD_1595
1595Brev	TATCAGGTACCCTACTCCTACAG CTGTCATTGCCTG	Reverse primer for amplification of flanking region upstream of BMD_1595 (KpnI site)
1068Afor	TATCAGGTACCCTGTCTGTAACG GCTGGAAACACA GC	Forward primer for amplification of flanking region downstream of BMD_1068 (KpnI site)
1068Arev	ATTTAATCCATCAATTTTGCTTACA TCTGATTACTCATTATAAAAAACC TCCTGCT	Reverse primer for amplification of flanking region downstream of BMD_1068
1068Bfor	AGCAGGAGGTTTTTTATAATGAGT AATCAGATGTAAGCAAATTGATG GATTAAT	Forward primer for amplification of flanking region upstream of BMD_1068
1068Brev	TATCAGGTACCAGAGTGAGTACA TTAGACTTGCTCTTT	Reverse primer for amplification of flanking

3715Afor	TATCAGAATTCTGTTGAGGCAAAC ATCTAATGA	region upstream of BMD_1068 (KpnI site) Forward primer for amplification of flanking region downstream of BMD_3715 (EcoRI site)
3715Arev	GAGTAAAGATGCCTGGCTTTGGC CGTGTAGGTCCAGATCCTGACAA CTT	Reverse primer for amplification of flanking region downstream of BMD_3715
3715Bfor	AAGTTGTCAGGATCTGGACCTAC ACGGCCAAAGCCAGGCATCTTTA CTC	Forward primer for amplification of flanking region upstream of BMD_3715
3715Brev	TATCAGAATTCATTTGAGAGGGTT GATTATTTATTTTCTTA	Reverse primer for amplification of flanking region upstream of BMD_3715 (EcoRI site)

Supplemental Table 2: NMR data for 20 β OH-cortisol (P6) and 20 β OH-cortisone (P7)

Product	P6: 11 β , 17 α , 20 β , 21-Tetrahydroxy-4-pregnen-3-one (20 β OH-Cortisol)
¹H NMR (CDCl₃, 500 MHz)	¹ H NMR (CDCl ₃ , 500 MHz): δ 1.08 (s, 3xH-18), 1.02 (dd, J=11.1 and 3.4 Hz, H-9), 1.12 (m, H-7a), 1.44 (s, 3xH-19), 1.30 (m, H-15a), 1.55 (m, H-16a), 4.41 (q, 3.3 Hz, H-11), 1.65 (m, H-14), 1.87 (m, H-1a), 1.81 (m, 2H, H-15b and H-16b), 1.86 (m, H-12a), 1.98 (m, H-7b), 1.92 (m, H-1b), 2.00 (m, H-8), 2.18 (dt, J= 13.4 and 4.7, H-12a), 2.24 (ddd, 14.5, 4.5 and 2.0 Hz, H-6a), 2.35 (dt, J=16.8 and 4.3 Hz, H-2a), 2.46 (m, H-6b), 2.48 (ddd, J=16.8, 13.9 and 5.0 Hz, H-2b), 3.76 (dd, J=11.0 and 3.4 Hz, H-21a), 3.80 (dd, J=11.0 and 5.2 Hz, H-21b), 3.83 (dd, J=5.2 and 3.4 Hz, H-20), 5.67 (br s, H-4). ¹³ C NMR (CDCl ₃ , 125 MHz): δ 17.83 (CH ₃ , C-18), 20.92 (CH ₃ , C-19), 23.82 (CH ₂ , C-15), 31.12 (CH ₂ , C-6), 31.55 (CH, C-8), 32.71 (CH ₂ , C-7), 33.07 (CH ₂ , C-16), 33.81 (CH ₂ , C-2), 34.94 (CH ₂ , C-12), 39.21 (C, C-10), 41.53 (CH ₂ , C-1), 46.64 (C, C-13), 50.53 (CH, C-14), 56.00 (CH, C-9), 64.68 (CH ₂ , C-21), 68.54 (CH, C-11), 72.88 (CH, C-20), 85.15 (C, C-17), 122.26 (CH, C-4), 172.55 (C, C-5), 199.78 (C, C-3).
Product	P7: 17 α , 20 β , 21-Trihydroxy-4-pregnen-3,11-dione (20 β OH-Cortisone)
¹H NMR (CDCl₃, 500 MHz)	δ 0.77 (s, 3xH-18), 1.30 (m, H-7a), 1.32 (m, H-15a), 1.39 (s, 3xH-19), 1.63 (td, J=14.0 and 4.4 Hz, H-1a), 1.73 (m, H-16a), 1.84 (m, H-16b), 1.90 (m, H-8), 1.92 (m, H-15b), 1.93 (m, H-9), 1.96 (m, H-7b), 2.28 (m, H-6a), 2.29 (m, H-2a), 2.33 (m, H-14), 2.40 (tdd, J= 14.5, 5.2 and 1.8 Hz, H-6b), 2.47 (ddd, J=17.0, 14.0 and 5.0 Hz, H-2b), 2.49 (d, J=12.5 Hz, H-12a), 2.66 (d, J=12.5 Hz, H-12b), 2.75 (ddd, J= 14.0 5.0 and 3.2 Hz, H-1b), 3.77 (m, 3H, H-20, H-21a and H-21b), 5.72 (br s, H-4). ¹³ C NMR

(CDCl₃, 125 MHz): δ 15.71 (CH₃, C-18), 17.19 (CH₃, C-19), 23.45 (CH₂, C-15), 32.31 (2xCH₂, C-6 and C-7), 33.33 (CH₂, C-16), 33.71 (CH₂, C-2), 34.68 (CH₂, C-1), 36.96 (CH, C-8), 38.15 (C, C-10), 48.35 (CH, C-14), 51.26 (CH₂, C-12), 51.58 (C, C-13), 62.43 (CH, C-9), 64.41 (CH₂, C-21), 72.57 (CH, C-20), 84.11 (C, C-17), 124.45 (CH, C-4), 169.13 (C, C-5), 200.02 (C, C-3), 210.92 (C, C-11).

2.4 Neunzig et al., 2017

The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1

Neunzig, J., **Milhim, M.**, Schiffer, L., Khatri, Y., Zapp, J., Sánchez-Guijo, A., Hartmann, M.F., Wudy, S.A., Bernhardt, R.

Journal of Steroid Biochemistry and Molecular Biology, 2017 Mar. 167: 182–191.

DOI: 10.1016/j.jsbmb.2017.01.002

Reprinted with permission of Journal of Steroid Biochemistry and Molecular Biology. All rights reserved.



Contents lists available at ScienceDirect

Journal of Steroid Biochemistry & Molecular Biology

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

The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1



J. Neunzig^a, M. Milhim^a, L. Schiffer^a, Y. Khatri^a, J. Zapp^b, A. Sánchez-Guijo^c, M.F. Hartmann^c, S.A. Wudy^c, R. Bernhardt^{a,*}

^a Department of Biochemistry, Faculty of Technical and Natural Sciences III, Saarland University, 66123 Saarbrücken, Germany

^b Institute of Pharmaceutical Biology, Faculty of Technical and Natural Sciences III, Saarland University, 66123 Saarbrücken, Germany

^c Steroid Research & Mass Spectrometry Unit, Division of Pediatric Endocrinology & Diabetology, Center of Child and Adolescent Medicine, Justus Liebig University, 35392 Giessen, Germany

ARTICLE INFO

Article history:

Received 30 September 2016

Received in revised form 2 January 2017

Accepted 3 January 2017

Available online 5 January 2017

Keywords:

Steroidogenesis
CYP21A2
CYP19A1
androstenedione
16(β)-OH-androstenedione
LC-MS/MS

ABSTRACT

The 21-hydroxylase (CYP21A2) is a steroidogenic enzyme crucial for the synthesis of mineralo- and glucocorticoids. It is described to convert progesterone as well as 17-OH-progesterone, through a hydroxylation at position C21, into 11-deoxycorticosterone (DOC) and 11-deoxycortisol (RSS), respectively. In this study we unraveled CYP21A2 to have a broader steroid substrate spectrum than assumed. Utilizing a reconstituted *in vitro* system, consisting of purified human CYP21A2 and human cytochrome P450 reductase (CPR) we demonstrated that CYP21A2 is capable to metabolize DOC, RSS, androstenedione (A4) and testosterone (T). In addition, the conversion of A4 rendered a product whose structure was elucidated through NMR spectroscopy, showing a hydroxylation at position C16- β . The androgenic properties of this steroid metabolite, 16(β)-OH-androstenedione (16bOHA4), were investigated and compared with A4. Both steroid metabolites were shown to be weak agonists for the human androgen receptor. Moreover, the interaction of 16bOHA4 with the aromatase (CYP19A1) was compared to that of A4, indicating that the C16 hydroxyl group does not influence the binding with CYP19A1. In contrast, the elucidation of the kinetic parameters showed an increased K_m and decreased k_{cat} value resulting in a 2-fold decreased catalytic efficiency compared to A4. These findings were in accordance with our docking studies, revealing a similar binding conformation and distance to the heme iron of both steroids. Furthermore, the product of 16bOHA4, presumably 16-hydroxy-estrone (16bOHE1), was investigated with regard to its estrogenic activity, which was negligible compared to estradiol and estrone. Finally, 16bOHA4 was found to be present in a patient with 11-hydroxylase deficiency and in a patient with an endocrine tumor. Taken together, this study provides novel information on the steroid hormone biosynthesis and presents a new method to detect further potential relevant novel steroid metabolites.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Steroid hormones are synthesized from cholesterol in a series of enzymatic reactions, involving six cytochromes P450 (CYPs), as

well as three hydroxysteroid-dehydrogenases (HSDs). The six cytochromes P450 belong to two different CYP classes: the mitochondrial CYPs (CYP11A1, CYP11B1 and CYP11B2) and the microsomal CYPs (CYP17A1, CYP21A2 and CYP19A1). The mitochondrial CYPs are arranged at the inner mitochondrial membrane. The electrons needed to catalyze their hydroxylation reactions are transferred from NADPH via a FAD containing adrenodoxin reductase (AdR) and an iron-sulphur containing adrenodoxin (Adx). In case of the microsomal CYPs, which are attached to the endoplasmic reticulum (ER) the electrons are delivered through a FAD and FMN containing NADPH-dependent oxidoreductase (CPR) [1,2].

Abbreviations: 16bOHA4, 16 β hydroxy-androstenedione; DOC, deoxycorticosterone; A4, androstenedione; E1, estrone; AdR, adrenodoxin reductase; E2, estradiol; Adx, adrenodoxin; ER, estrogen receptor; AR, androgen receptor; FAD, flavin adenine dinucleotide; CPR, cytochrome P450 reductase; FMN, flavin mononucleotide; CYP, cytochrome P450; p16bOHE1, putative 16 β hydroxy-estrone; DHEA, dehydroepiandrosterone; RSS, 11-deoxycortisol; DLPC, dilauroyl phosphatidylcholin; T, testosterone.

* Corresponding author.

E-mail address: ritabern@mx.uni-saarland.de (R. Bernhardt).

<http://dx.doi.org/10.1016/j.jsbmb.2017.01.002>

0960-0760/© 2017 Elsevier Ltd. All rights reserved.

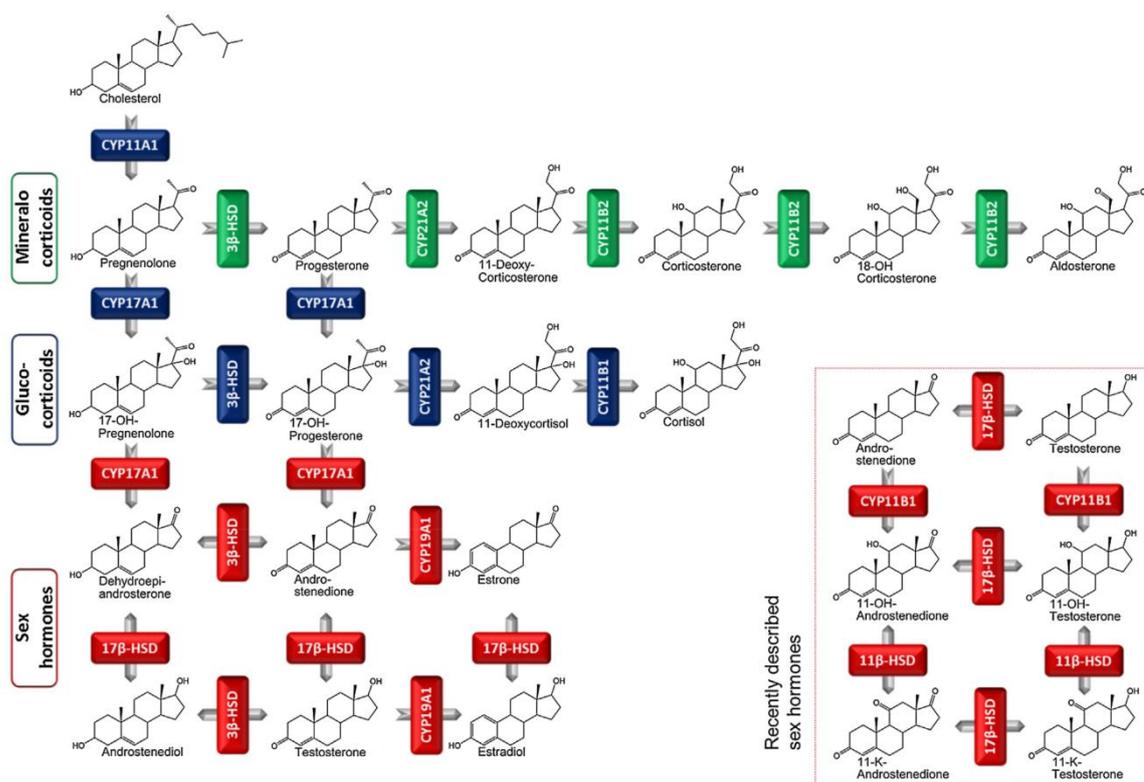


Fig. 1. Scheme of classical steroidogenesis including recently described sex hormones (inset) [6,7].

In mammals, steroid hormones are crucial for life and reproduction. They are synthesized in steroidogenic tissues and organs such as the adrenal gland (Fig. 1), gonads, brain or the testis. Depending on their mode of action, steroid hormones are classified into mineralo- and glucocorticoids, sex hormones and neurosteroids. Through activation of the respective nuclear receptors, they participate in many different vital processes, such as the regulation of the salt and water homeostasis through the action of aldosterone, the main mineralocorticoid or the development of secondary sexual characteristics through the action of estrogens and androgens. Also, several steroid intermediates have been shown to possess biological relevance, such as DHEA or progesterone, which are involved in a variety of biological processes [3–5].

To date, the six CYPs participating in steroid hormone biosynthesis were reported to be successfully expressed in bacteria and subsequently purified, allowing their characterization and the investigation of the impact of diverse compounds on their activities on a molecular level [6–11]. Thereby researchers gained a deeper insight into the mechanistic mode of action of steroidogenic CYPs. In 2012, Hobler and colleagues [9] using recombinantly expressed and purified human CYP11B2 demonstrated its ability to hydroxylate the substrates corticosterone and deoxycorticosterone at position C19. Albeit the synthesis of 19OH-corticosterone has not yet been described so far in humans, the 19-hydroxylase activity towards deoxycorticosterone (DOC) was previously characterized by Kawamoto et al. [12] using COS-7 cells transfected with a plasmid encoding for human CYP11B2. These 19-hydroxylated products were proposed to be precursors of 19-normineralocorticoids such as 19-noraldosterone showing

hypertensinogenic activities and, therefore, a relation to primary aldosteronism. Moreover, CYP11B2 was shown to be able to hydroxylate deoxycorticosterone at position C18 yielding 18OH-deoxycorticosterone. It was concluded that aldosterone synthesis can alternatively be conducted via 18-hydroxylation of DOC, followed by 11β-hydroxylation and 18-oxidation of 18OH-DOC [9].

Very recently, a systematic analysis of the effect of steroid hormone intermediates on the action of CYP11A1 was performed in our laboratory. It was demonstrated that beside its side-chain cleavage activity, CYP11A1 possesses a 2β-, a 6β- and a 16β-hydroxylase activity towards different steroid hormone intermediates [13]. These novel enzymatic activities of CYP11A1 result in the formation of 2β- and 6βOH-deoxycorticosterone, 2β- and 6βOH-androstenedione, 6βOH-testosterone and 16βOH-DHEA. The hydroxylations at positions 6β and 16β of these steroids are already described, but mostly attributed to liver CYPs [14,15]. There are, however no reports available about the production 2βOH-deoxycorticosterone and 2βOH-androstenedione. Referring to the physiological role of 2βOH-deoxycorticosterone and 2βOH-androstenedione no data is available and thus, their function still has to be unravelled.

All of the above information indicates that the enzymatic properties of steroidogenic CYPs still have not been fully elucidated. The observation that CYP11A1, besides performing the side-chain cleavage of cholesterol, is also able to perform highly selective hydroxylations on steroid scaffolds opens up new paths for the understanding of the steroid hormone biosynthesis.

In this study we describe a novel activity of CYP21A2. This steroidogenic CYP hydroxylates progesterone and 17OH-progesterone at position C21, yielding DOC and deoxycortisol (RSS),

respectively, and, thus is essential for the formation of mineralo- and glucocorticoids. We demonstrated that besides the steroid intermediates depicted in Fig. 1, DOC, RSS, testosterone (T) and androstenedione (A4) serve as a substrate for CYP21A2. The conversion of A4 was characterized in detail, revealing a hydroxylation activity of CYP21A2 at position C16- β . Such a reaction performed by CYP21A2 was so far not described in the literature and, therefore, represents a completely novel reaction leading to the steroid hormone intermediate 16(β)-hydroxy-androstenedione (16bOHA4). It was demonstrated that 16OHA4 efficiently binds to the active site of CYP19A1 and is further metabolized to an aromatized product with negligible estrogenic activity. In order to understand binding and conversion in detail, docking studies were conducted using the crystal structure of human CYP19A1. Furthermore, the sulfonated 16OHA4 was found to be present in the urine of a patient with 11-hydroxylase deficiency as well as in a patient with an endocrine tumor, which suggests the physiological relevance of this novel steroid intermediate.

2. Materials

Steroids were purchased from Sigma–Aldrich (Taufkirchen, Germany) or from C/D/N Isotopes Inc. (Quebec, Canada). 1,2-Dilauroyl-sn-glycero-3-phosphocholine (DLPC), kanamycin sulfate, arabinose, magnesium chloride, sodium hydroxide, HPLC-grade acetonitrile were from Sigma–Aldrich (Taufkirchen, Germany). Yeast extract, “technical” was from Becton, Dickinson and Company (Heidelberg, Germany). Pepton, pancreatically digested and Na-acetate were from Merck (Darmstadt, Germany). LC–MS grade water and ammonium hydroxide were obtained from Fluka (Taufkirchen, Germany). Ethanol was obtained from Merck (Darmstadt, Germany). Ampicillin was from CarlRoth (Karlsruhe, Germany). NADPH was from Carbolution (Saarbrücken, Germany). Glucose-6-phosphate and glucose-6-phosphate dehydrogenase were purchased from Roche (Basel, Switzerland). Protino Ni-NTA was obtained from Macherey–Nagel (Düren, Germany). All other chemicals were of highest purity available.

3. Methods

3.1. Construction of recombinant expression plasmids in *E. coli*

The plasmid pET-17b was utilized to express human CYP21A2, CYP19A1 and CPR in *E. coli*. Each cDNA was cloned via the restriction sites NdeI and BamHI into the vector. The cDNA of human CYP21A2 was constructed according to Arase et al. with a replacement of the N-terminal hydrophobic anchor region with MAKKTSSKGGK from CYP2C3 and a C-terminal 6 \times histidine tag for protein purification [7]. CYP19A1 cDNA with a N-terminal replacement (MARQSFGRGKL, derived from CYP2C11) was kindly provided by Prof. N. Kagawa [16]. The CPR amino acid sequence is extended at the C-terminus by three glycines and six histidines and at the N-terminus has a lack of the first 27 amino acids [17]. For the *E. coli* based whole-cell system, the plasmid pET-17b was used to express CYP21A2 and CPR. The cDNAs of CYP21A2 and CPR were cloned via the restriction sites NdeI and BamHI (CYP21A2) and BamHI and NotI (CPR) into the vector.

3.2. Protein expression and purification

Human CYP21A2 was co-expressed with chaperones GroEL and GroES as described previously [7]. As host *E. coli* strain C43DE3 was used. The expression was performed in 2 l baffled flasks containing 400 ml TB medium supplemented with 100 mg/l ampicillin and 50 mg/l kanamycin. The expression culture was inoculated from an

overnight culture and grown at 37 °C and 120 rpm. Protein expression was induced at OD_{600nm} = 0.6 with 1 mM IPTG, 4 mg/ml arabinose, 1 mM δ -ALA and 50 mg/l ampicillin. Afterwards, temperature was decreased to 26 °C and the cells incubated at 95 rpm for 48 h. Cells were harvested and sonicated as described elsewhere [18]. The purification was performed as described for CYP11B1 [8]. Human CPR was expressed in *E. coli* strain C43DE3 using similar conditions as described for bovine CPR [11]. Cells were harvested and sonicated as described elsewhere [18]. CPR was purified via IMAC as described [8] with 40 mM imidazole in the washing buffer and 200 mM imidazole in the elution buffer. Human CYP19A1 was co-expressed with chaperones GroEL and GroES as described [19]. The expression was performed in 2.8 l Fernbach flasks containing 250 ml TB medium supplemented with 100 mg/l ampicillin and 50 mg/l kanamycin. The expression culture was inoculated from an overnight culture and grown at 37 °C and 220 rpm. Protein expression was induced at OD_{600nm} = 0.6 with 1 mM IPTG, 4 mg/ml arabinose, 1 mM δ -ALA and 50 mg/l ampicillin. Then, temperature was decreased to 30 °C and the cells incubated at 180 rpm for 24 h. Afterwards, human CYP19A1 was purified as described elsewhere [20].

3.3. UV-vis spectroscopy

CYP21A2 and CYP19A1 concentrations were calculated from a reduced carbon-monoxide difference spectrum according to Omura and Sato [21] with $\Delta\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$. Binding of A4 and 16bOHA4 to CYP19A1 was investigated using difference spectroscopy, which was carried out in tandem cuvettes according to Schenkman [22]. A4 and 16bOHA4 were dissolved in DMSO. The buffer utilized was composed of 50 mM HEPES (pH 7.4), 20% glycerol, and 100 mM dilauroyl phosphatidylcholine (DLPC). Difference spectra were recorded from 370 to 450 nm. To determine the dissociation constant (K_d), the values from four titrations were averaged and the resulting plots were fitted with hyperbolic regression.

3.4. Enzyme activity assay

In vitro substrate conversion assays were done as described [9] with slight modifications. The conversion buffer (50 mM HEPES, pH 7.4, 20% glycerol, 100 mM DLPC) contained 1 μ M CYP21A2 or CYP19A1, 1 μ M CPR, 1 mM MgCl₂, 5 mM glucose-6-phosphate, 4 U/ml glucose-6-phosphate dehydrogenase for NADPH regeneration and varying concentrations of A4 or 16bOHA4 as substrate. After starting the reaction with 1 mM NADPH at 37 °C (CYP21A2: 20 min, CYP19A1: 30s–1 min), the conversion was stopped by addition of one reaction volume of ethyl acetate. In the case of the CYP19A1-dependent conversion of A4 and 16bOHA4, testosterone (T) was added as internal standard. Extraction of steroids was performed twice with ethylacetate and the ethylacetate phase was evaporated. The steroids were suspended in 30% acetonitrile/water for subsequent HPLC analysis. Steroids were separated on a Jasco reversed phase HPLC system LC2100 using a 4 mm \times 125 mm Nucleodur C18 reversed phase column (Macherey–Nagel) with an acetonitrile/water gradient and a flow rate of 0.8 ml/min. Detection of the steroids was with a diode array spectrometer within 15–20 min at 40 °C.

In order to determine the kinetic constants, varying substrate concentrations in the range of 0–100 μ M were used to perform substrate conversion assays. Product formation was quantified by substrate consumption. The kinetic constants K_m and k_{cat} were determined by plotting the substrate conversion velocities versus the corresponding substrate concentrations and by applying Michaelis–Menten kinetics utilizing the program Origin 8.6

3.5. Androgen receptor (AR) and estrogen receptor (ER) transactivation assay

The cell culture based reporter assay was conducted utilizing the *Human AR or the Human ER Reporter Assay System (32-well Format Assays)* from Indigo Biosciences. The assay was carried out according to the manufacturer's protocol (Technical manual, Version 7.1.). For the agonist mode of the AR assay, A4 and 16OHA4 were applied individually with a concentration of 160 pM (EC₅₀ of 6FI-T). For the agonist mode of the ER assay E2, E1 and p16OHE1 were added individually with a concentration of 1.3 nM and, additionally, E1 and p16OHE1 were added with a concentration of 4 nM.

3.6. Whole-cell biotransformation and purification of 16bOHA4

The bicistronic expression vector pET17b_CYP21A2_CPR coding for CYP21A2 and cytochrome P450 reductase (CPR) was used in the present study for whole-cell biotransformations. *E. coli* C43(DE3) cells were co-transformed with pET17b_CYP21A2_CPR and the chaperone GroEL/GroES-encoding plasmid pGro12, possessing ampicillin and kanamycin resistance genes, respectively. Transformed cells were grown overnight in 50 ml Terrific broth (TB) medium supplemented with 100 mg/l ampicillin and 50 mg/l kanamycin at 37 °C and shaking at 180 rpm. For the preparative purification of 16bOHA4, 300 ml TB medium in 21 baffled flasks containing 100 mg/l ampicillin and 50 mg/l kanamycin were inoculated (1:100) with the transformed cells and cultivated at 37 °C with rotary shaking at 105 rpm. After the optical density at 600 nm reached to 0.8–1.0, the culture was induced with 1 mM (IPTG), 1 mM (d-ALA) and 4 mg/ml arabinose. The temperature and shaking speed were then reduced to 30 °C and 85 rpm, respectively. After incubation for 28 h, cells were harvested by centrifugation (4500g) for 20 min at 4 °C and washed once with 300 ml 50 mM potassium phosphate buffer (KPP) (pH 7.4) and 2% glycerol to remove indol, formed by the cells from tryptophane, since indol had been shown to have inhibitory effects towards CYPs [23,24]. After a second centrifugation, the cell pellets were suspended to a concentration of 60g wet cell pellet per liter KPP (pH 7.4) with 2% glycerol. Afterwards, 25 ml of the suspension were transferred to 300 ml baffled flask. A4 was added to a final concentration of 500 μM and the culture was incubated for another 24 h at 30 °C and 145 rpm. Steroids were extracted twice with the same volume of ethyl acetate and the organic phase was evaporated using a rotary evaporator. Steroids were separated by reversed phase HPLC on a C18ec column (Nucleodur 250/8) (Macherey-Nagel) with a mobile phase consisting of 20% acetonitrile in water. The fraction containing the product was collected, evaporated and subjected to subsequent nuclear magnetic resonance (NMR) spectroscopy analysis.

3.7. LC–MS analysis

LC–MS analyses were performed using a novel method for unconjugated [25]. The detection of 16bOHA4 (302.1 Da) was performed using the product ion scanning approach. For 16bOHA4, [M+H]⁺ has an *m/z* of 303.1 and, therefore, the first quadrupole was set at *m/z* 303.1 to exclude other ions to enter the MS. Subsequently, the ions were fragmented in the second quadrupole with argon and a collision energy of 40 eV, and then detected in the third quadrupole.

The purified standard 16bOHA4 appeared at a retention time of 1.48 min and generated two main fragments in Q3; *m/z* 97 and *m/z* 109 (see Supplementary information). The retention time and the signals 303.1→97 and 303.1→109 were used to identify the compound 16bOHA4 in biological samples after treatment with

steroid sulfatase. The detailed LC–MS/MS analysis of the patients tested are in the Supplementary information, S1).

3.8. NMR analysis

The NMR spectra were recorded in CDCl₃ at 298 K on a Bruker DRX 500 equipped with a 5 mm BBI probe head. The chemical shifts were relative to CHCl₃ at δ 7.24 (1H NMR) and CDCl₃ at δ 77.00 (13C NMR) respectively using the standard δ notation in parts per million. The 1D NMR (1H and 13C NMR, DEPT135) and the 2D NMR spectra (gs-HH-COSY, gs-NOESY, gs-HSQCED, and gs-HMBC) were recorded using the BRUKER pulse program library. All assignments were based on extensive NMR spectral evidence. NMR data is depicted in Fig. S2.

3.9. Docking of the crystal structure of human CYP19A1

The computer-simulated automated docking program AUTODOCK (version 4.00) [26] was applied for docking of the A4 and 16bOHA4 into the crystal structure of human CYP19A1 (PDB 3S79) after removing the associated ligands [27]. The Windows version 1.5.6 of the Autodock Tools was used to compute Kollman charges and Gasteiger-Marsili charges for the ligand free CYP19A1 and the steroids A4 and 16bOHA4, respectively [28]. Flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box (58 × 58 × 58 points with a grid spacing of 0.4 Å) was centered 5 Å above the heme-iron. For each of the ligands, 100 docking runs were carried out applying the Lamarckian genetic algorithm using default parameter settings.

4. Results and discussion

Although investigations on steroid biochemistry emerged in the 1930s and sophisticated analytical methods, such as GC–MS, arose in the 1960s, the field of steroid hormone research is still challenging and recent work reveals unexpected new findings. As discussed before, it has been shown that besides its side-chain cleavage activity, CYP11A1 is able to perform unexpected 2β-, 6β- and 16β-hydroxylations [13]. Further, Bloem et al. demonstrated the androgenic effect of 11-keto-dihydroxytestosterone, which was comparable to dihydrotestosterone, a full androgen receptor agonist [29]. This steroid metabolite is formed through an 11-hydroxylation of A4 by CYP11B1, followed by an oxidation of the OH-group catalysed by 11β-HSD.

The present work aimed to unravel further novel steroid hormone intermediates. For this purpose, it was examined whether CYP21A2 is capable to metabolize steroid metabolites others than progesterone and 17OH-progesterone. For this, the potential conversion of the steroid intermediates displayed in Fig. 1, have been investigated in an in vitro system consisting of purified human CYP21A2 and human CPR as well as a NADPH regenerating system. Four steroid metabolites displayed clear product peaks, as revealed by HPLC analysis (Fig. 2).

DOC and RSS, which are the products of the CYP21A2-dependent conversion of progesterone and 17OH-progesterone, were further metabolized by this enzyme. In the case of DOC, two products were formed with retention times of 14.5 and 16 min, which correspond to one more hydrophilic and one more hydrophobic steroid intermediate compared with the substrate DOC. In the case of RSS only one product was generated, which eluted at 13.5 min. The CYP21A2-dependent conversion of A4 revealed also one main product, which was monitored at 12.7 min. The conversion of T resulted in the formation of four different products, with retention times of 12.5, 13, 13.5 and 16 min. To get some information on these steroid metabolites we determined their mass using LC–MS (Table 1). The results show that these

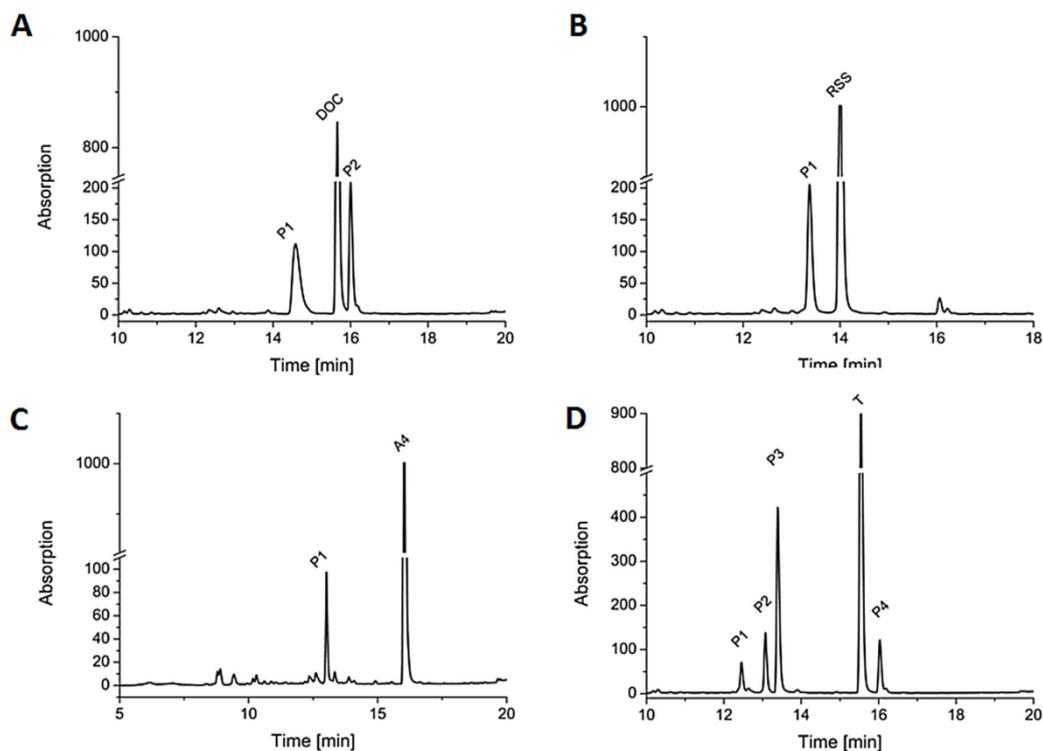


Fig. 2. HPLC chromatograms of CYP21A2-dependent conversions of A: DOC, showing two products at 14.2 min and 16 min. B: RSS with one product peak at 13.2 min. C: A4 showing one product at 13 min. D: T with four products at 12.4 min, 13.1 min, 13.5 min and 16 min.

metabolites were mainly hydroxylated or in the case of P2 of testosterone oxidated. Unfortunately, we could not get the mass of P2 and P4 of DOC and T, respectively. The mass obtained for these metabolites is 1242, which corresponds to a DLPC dimer (mass of 621). Therefore, an interference of the P2 and P4 metabolites and DLPC, which is necessary for the reaction, must be assumed. Nevertheless, since it was shown for P2 of T that CYP21A2 is able to catalyse an oxidation reaction leading to a more apolar metabolite, it can be proposed that a similar reaction could have taken place for the two not identified products

These unexpected findings demonstrate that CYP21A2 is more versatile than previously assumed and might play a role in the synthesis of, so far, unknown steroid metabolites. In order to elucidate the structure of these products, NMR analysis was performed. For this an *E. coli* based whole-cell system was established, expressing human CYP21A2 and human CPR to gain sufficient amounts of the novel metabolites. Preliminary

experiments showed a high product yield for A4, which was the reason to focus on this putative novel steroid intermediate in the following study.

One liter of the *E. coli* based whole-cell system was fed with 143 mg A4 (500 μ M), from which 66.5% were converted into the product (93 mg). The A4 product was then purified through a preparative HPLC, yielding ~22 mg of pure product (Fig. 3).

In contrast to A4 the NMR data of its conversion product showed resonances for an additional secondary hydroxyl group (δ_{H} 3.95 dd, $J = 9.5$ and 8.2 Hz; δ_{C} 75.21, CH₂). It could be located at C-16 by correlations of its proton resonance with those of H-15a (δ_{H} 1.54

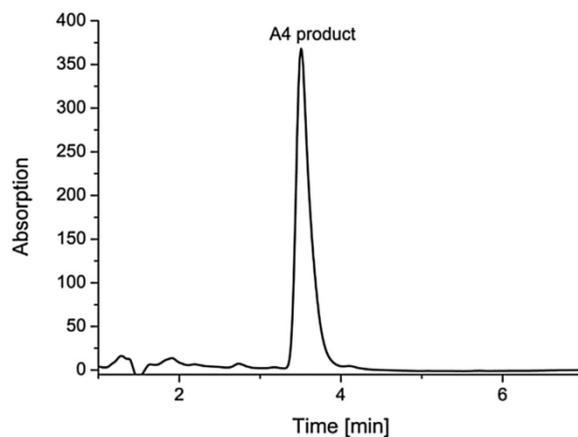


Fig. 3. HPLC-chromatogram of the purified A4 product eluting at 3.3 min (HPLC method for the CYP19A1-dependent conversion was used).

Table 1

Mass of the steroid metabolites, which are shown in Fig. 2, with probably catalyzed reaction of CYP21A2, determined by LC-MS.

Substrate		Masse	Possible reaction
DOC		331.40	Substrate
	P1	347.47	Hydroxylation (+16)
	P2	1242.42	Interference with DLPC
RSS		347.44	Substrate
	P1	363.31	Hydroxylation (+16)
Testosterone		289.41	Substrate
	P1	305.35	Hydroxylation (+16)
	P2	303.35	Oxidation of P1 (-2)
	P3	305.36	Hydroxylation (+16)
	P4	1242.55	Interference with DLPC

ddd, $J = 3.5, 12.0$ and 9.5 Hz) and H-15b ($\delta_{\text{H}} 2.39$ m) in the HHCOSEY and with carbonyl C-17 ($\delta_{\text{C}} 219.52$) in the HMBC. Unfortunately the stereochemistry at C-16 could not be solved by a NOESY spectrum, because H-16 showed only correlations to both H-15 protons. A literature search for both 16-hydroxy epimers revealed NMR data only for the 16 α -hydroxy form. Its proton data [30] resembled those of our compound, especially for H-1–H-11, but showed significant differences for H-14–H-16. 16 β -hydroxylation rarely occurs in microbial transformations of steroids [31]. That might be the reason that no NMR data could be found for the 16 β -hydroxy form. Instead of the dione ^{13}C NMR data for both 16-hydroxy epimers of 3 β ,16-hydroxy-androst-4-en-17-one where available [32]. Both data sets differed remarkably only in the chemical shifts of C-14 (16 α -OH form: 48.8 ppm; 16 β -OH form: 44.8 ppm) and C-16 (16 α -OH form: 71.3 ppm; 16 β -OH form: 75.3 ppm), indicating that our compound belonged to the 16 β -hydroxy form (C-14: 45.24 ppm; C-16: 75.21 ppm) (Fig. 4).

In humans there are several reports of the 16 α hydroxylation of A4 conducted by the liver CYPs CYP3A4 and CYP3A7 [33]. However, there are no literature reports about a 16 β hydroxylation of A4 in humans. Interestingly conversion studies of A4 using adrenal microsomes from guinea pigs, revealed the hydroxylation at position C16 β . The authors excluded the participation of liver CYPs and attributed this reaction to an adrenal CYP [34]. Here, we demonstrated the ability of the human CYP21A2 to perform this reaction, which suggests a potential role of this rare steroid intermediate in human. For this reason we decided to explore the action of 16 β OHA4 in more detail.

Kinetic experiments of A4 conversion catalyzed by CYP21A2 were not performed, since the conversion efficiency is too weak to obtain reasonable data. The solubility of A4 in aqueous solutions is around $200 \mu\text{M}$, which is not enough to reach the V_{max} value of this reaction.

Since A4 is a precursor of sex hormones, we were interested to know whether its product 16 β OHA4 possesses androgenic activity. Therefore a human androgen receptor assay was conducted (Fig. 5). The results demonstrate that 16 β OHA4 has a similar androgenic potential as A4, which has been described to weakly activate androgen receptors (AR) [35], indicating that the hydroxyl group at C16 β of 16 β OHA4 does thus not influence the interaction with androgen receptors. In contrast to the weak activation of the AR by A4 and 16 β OHA4, the reference substance 6 β I-testosterone strongly activated ARs, proving the functionality of the assay.

Next, we tested whether 16 β OHA4 is able to interact with CYP19A1, which physiologically converts A4, 16 α OHA4 and testosterone (T) into the estrogens estrone (E1), estriol (E3) and estradiol (E2), respectively.

Since the binding of substrates into the active site of CYPs often provokes a spectral change [22], we investigated in a first step the capability of 16 β OHA4 to induce a spin-shift of the CYP19A1 heme

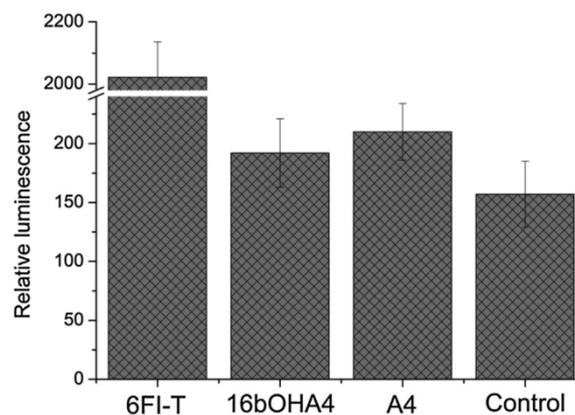


Fig. 5. Luminescence results of the androgen receptor assay. 6FI-T: 6 β I-testosterone as reference AR-agonist. 16 β OHA4: 16 β -OH-androstenedione. A4: androstenedione. Control: without steroid metabolite. Values are presented as mean-values of 4 independent measurements with standard deviation.

iron. This experiment revealed a tight binding of 16 β OHA4 to CYP19A1 with a K_{d} -value of $3.8 \pm 1.7 \mu\text{M}$, which is similar to the K_{d} -value $5.0 \pm 1.5 \mu\text{M}$ obtained for A4 (Fig. 6). The difference between these two values is not significant and therefore it is to assume that the hydroxyl group at position C16 β does not affect the binding to CYP19A1. In contrast, additional hydroxylations of A4 or T have been shown to dramatically affect binding properties of CYP19A1. For example, it was already shown that 4-hydroxy-androstenedione, which is used for the treatment of breast cancer, irreversibly binds to CYP19A1 [36,37].

In order to investigate whether 16 β OHA4 is metabolized by CYP19A1 we performed a reconstituted *in vitro* assay, consisting of recombinantly expressed and purified human CYP19A1 and human CPR. The results clearly show an absorbance change from 240 nm for 16 β OHA4 to 280 nm for the product, which is characteristic for an aromatization reaction (Fig. 7). Therefore, we propose the product to be 16 β -OH-estrone and refer to it in the following as putative 16 β -OH-estrone (p16 β OHE1).

The steady-state kinetic parameters for the 16 β OHA4 and A4 conversions are shown in Fig. 8 and Table 2. The conversions were found to follow Michaelis Menten kinetics. The K_{m} value is increased by a factor of 2 for 16 β OHA4, while the k_{cat} value is similar to that for A4 leading to a decrease in the catalytic efficiency of CYP19A1 with 16 β OHA4 by the factor of 2. From this it can be concluded that the hydroxyl group at position C16 β slightly hinders the optimal coordination of the substrate in the active site of CYP19A1. Nevertheless, the results demonstrate that this novel steroid metabolite is very efficiently converted by CYP19A1.

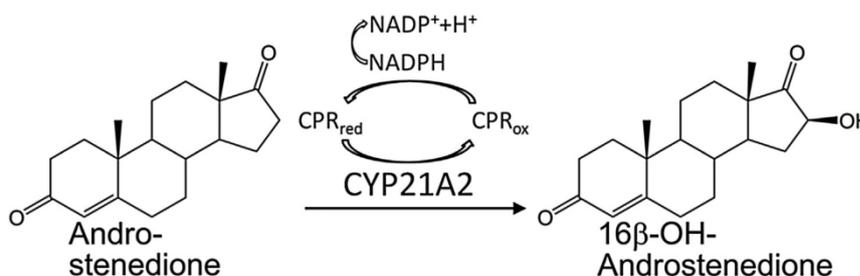


Fig. 4. A4 metabolism by CYP21A2 results in a hydroxylation at position C16 β .

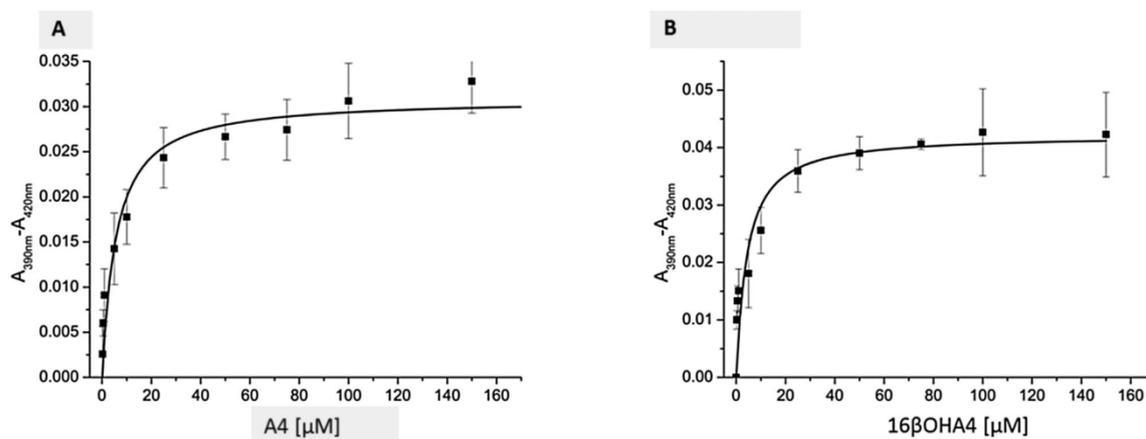


Fig. 6. Determination of substrate binding affinity to CYP19A1. The absorbance changes were plotted against the substrate concentration and fitted as described; $n = 3$.

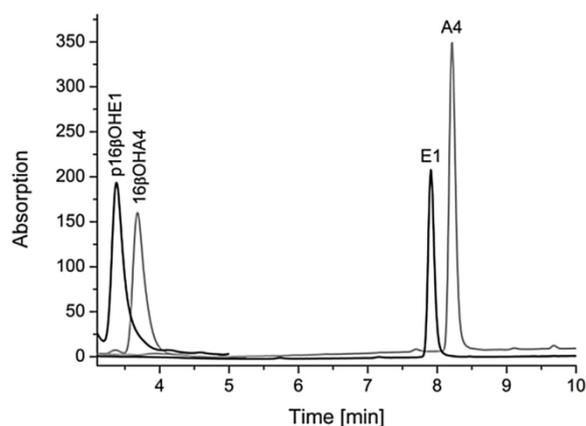


Fig. 7. HPLC-chromatogram of CYP19A1-dependent conversion of 16 β OHA4 and A4. The grey and black signals were recorded at a wavelength of 240 nm and 280 nm, respectively.

In order to get a deeper insight how 16 β OHA4 binds to CYP19A1, we performed docking studies using the crystal structure of human CYP19A1 (pdb 3S79). As shown in Fig. 9, the orientation of 16 β OHA4 and A4 in the active site of CYP19A1 is almost similar. In the case of 16 β OHA4, the substrate pocket is formed by the amino acid residues Trp224, Thr310, Ala306, Leu477, Leu372, Met374 and Val 370, whereas in the case of A4 the pocket is formed by Ala306, Trp224, Leu477, Met374, Arg115 and Leu133. Both steroid metabolites are fixed through a hydrogen bond between their C17 keto group and the residue Met374. In the case of 16 β OHA4 there is an additional hydrogen bond between the C16 hydroxyl group and the residue Leu372. The distance of the C19 atom of both substrates to the heme iron is slightly different. In the case of 16 β OHA4 this distance is 3.32 Å, whereas in the case of A4 the distance is 3.90 Å. The additional hydrogen bond, which is formed in the presence of 16 β OHA4 should lead to a significant increase of the affinity to CYP19A1. The binding of 16 β OHA4 that we observed is, however, not significantly different from the one of A4. At best we observe a tendency of a tighter binding, which is, however, not enough to be explained by the additional hydrogen bond. Hence, the K_d -values determined, question the formation of an additional hydrogen bond which is suggested by the docking model. To get

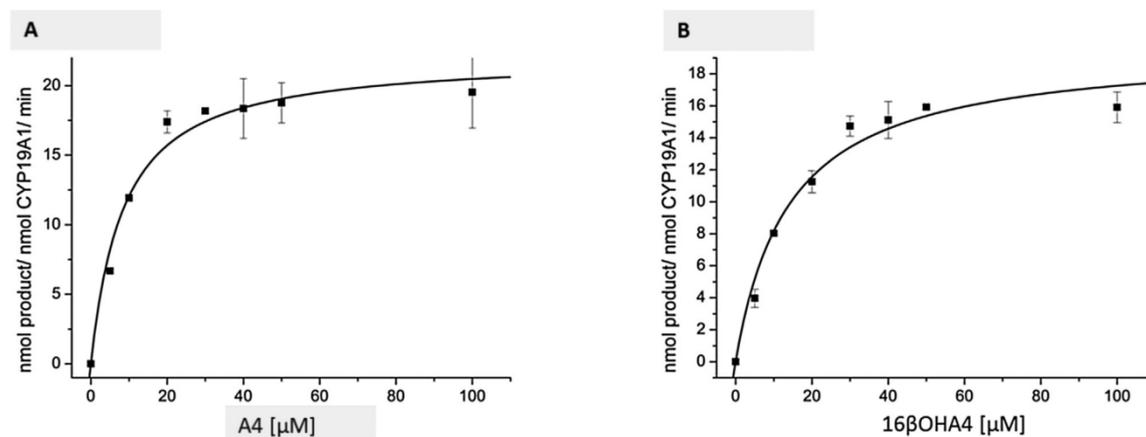


Fig. 8. Kinetics of A4 to E1 conversion catalyzed by CYP19A1 utilizing CPR as redox partner. B: Kinetic of 16 β OHA4 to p16 β OHE1 conversion catalyzed by CYP19A1 utilizing CPR as redox partner.

Table 2
Kinetic parameters of CYP19A1 metabolizing A4 and 16bOHA4, respectively.

	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{min}^{-1} \mu\text{M}^{-1}$)
A4	8 ± 1.5	22.1 ± 1.0	2.8
16bOHA4	14 ± 2	19.7 ± 1.2	1.4

further insight into the abovementioned discrepancy between experimental and theoretical data, a crystal of CYP19A1 with 16bOHA4 might be useful.

In order to investigate potential estrogenic properties of p16bOHE1, a human estrogen receptor assay was also performed. Since p16bOHE1 is not commercially available, we decided to generate this metabolite through a reconstituted *in vitro* reaction with CYP19A1 and CPR using 16bOHA4 as substrate. Likewise E1, E3 and E2 were produced. The positive control was conducted using E2 (in the following referred as E2 reference), which was provided by the manufacturer's assay, whereas the negative control was performed without the addition of any steroids (Fig. 10).

The luminescence measured in the presence of the E2 reference provided by the supplier was very similar to that obtained in the presence of E2, isolated from our *in vitro* system (8527 ± 494 and 8373 ± 758). This demonstrates the usefulness of the reconstituted *in vitro* system to obtain estrogens from androgens. The relative luminescence after addition of E1 was nearly 2.4 fold lower with a value of 3534 ± 380 . The lower estrogenic activity of E1 compared with E2 is in accordance with literature [38,39]. Interestingly, p16bOHE1 did not exhibit a significant estrogenic activity. The relative luminescence of p16bOHE1 was 458 ± 50 , whereas the negative control was 186 ± 12 . Even when adding 4 nM of p16bOHE1, the luminescence value increases only 1.5 fold to a value of 709 ± 72 , whereas the one in the presence of 4 nM E1 increased by a factor of 2 up to 7143 ± 613 .

In the abovementioned experiments it was possible to demonstrate that 16bOHA4 possesses a high affinity towards CYP19A1. Therefore it can be concluded that this novel steroid metabolite can act as a potent competitor of A4 and T. Moreover, it was shown that p16bOHE1, which is the resulting product of the CYP19A1-dependent conversion, does not exhibit significant estrogenic features. For this reason 16bOHA4 could be a promising candidate to test further, for the treatment of breast cancer caused by an estrogenic over-stimulation. However, further research in this field is necessary to explore the usefulness of this compound.

After identifying 16bOHA4 to be the resulting product of the CYP21A2-dependent A4 conversion, the question arose whether 16bOHA4 can be found in humans and thus, possesses physiological relevance. Hence, pure 16bOHA4 standard, which was obtained

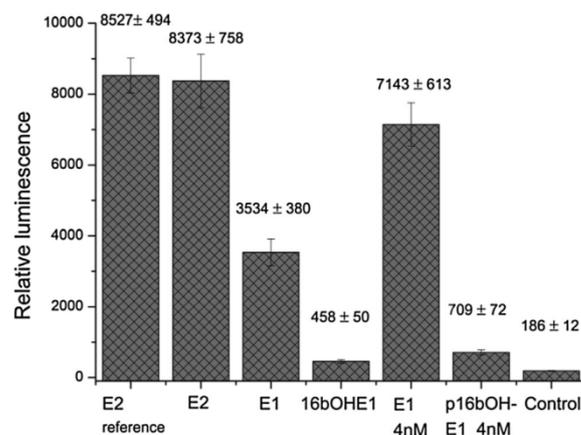


Fig. 10. Luminescence results of the estrogen receptor assay using 1.3 nM of each steroid. E2 reference: estradiol as reference ER-agonist. E2: estradiol obtained from T conversion. E1: estrone obtained from A4 conversion. p16bOHE1: putative 16-beta-OH-estrone obtained from 16bOHA4 conversion, E1 4 nM: 4 nM of estrone, 16bOHE1 4 nM: 4 nM of 16bOHE1. Control: without steroid metabolite. Values are presented as mean-values of 4 independent measurements.

after purification from the *E. coli* based whole-cell system, was included in the steroid profile analysis of humans. In a first step, the retention time and the fragmentation pattern of 16bOHA4 were analyzed. 16bOHA4 exhibits a retention time of 1.49 min, when employing the LC conditions described in the method section. The fragmentation shows two main fragments with m/z 107 and m/z 97 (Fig. S1).

Investigation of the steroid profile of a three years old healthy male control revealed no 16bOHA4. It has to be assumed that A4 has a significantly lower affinity towards CYP21A2 compared with its natural substrates progesterone and 17-OH-progesterone. This is suggested by the inability of A4 to induce a spin-shift of CYP21A2, and by the low conversion efficiency, as shown by the increased substrate conversion time (30 min), which was necessary to obtain the product. It seems that the concentration of A4 in human is not high enough to compete with Prog and 17OH-Prog for the active site of CYP21A2. For this reason we decided to explore the steroid profile of patients having an excess of androgens. We investigated the existence of 16bOHA4 in a patient with a defective 11-hydroxylase, in a patient with a defective 21-hydroxylase as well as a patient having an endocrine tumor overproducing androgens. As these patients have increased A4 concentrations, the chance to detect 16bOHA4 might be higher. The patient with a

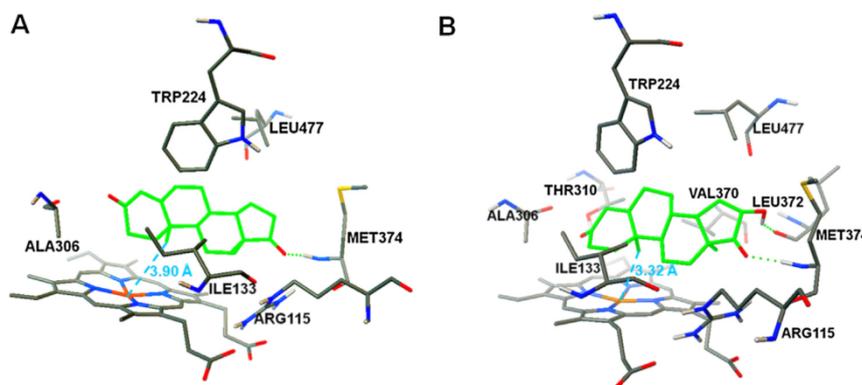


Fig. 9. Docking of A4 with human CYP19A1 (A) and of 16bOHA4 with human CYP19A1 (B). Amino acids forming the substrate pocket are labeled.

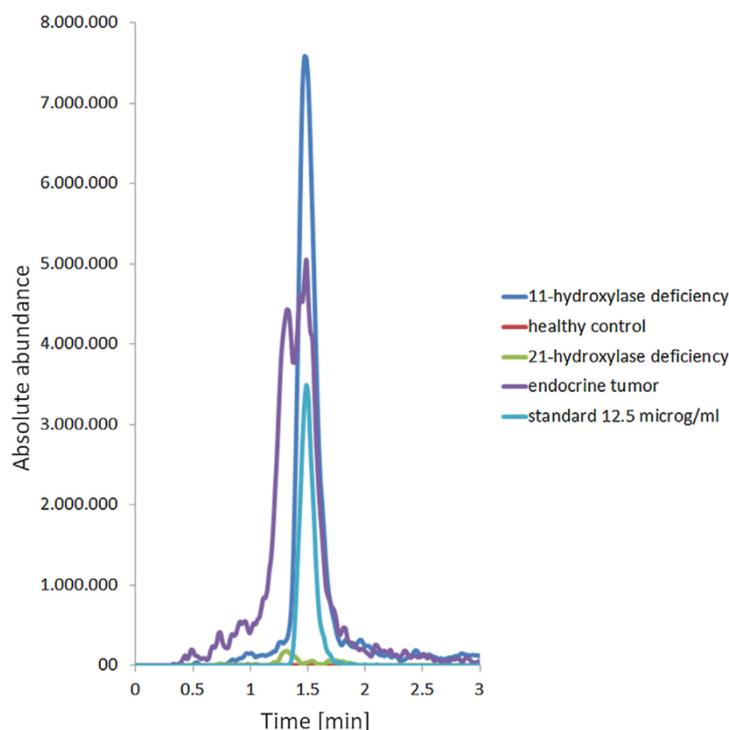


Fig. 11. Overlay of LC-chromatograms of: pure 16bOHA4 with a retention time at 1.49 min (light blue), the urine of a healthy child without any signal (orange), the urine of a child with 11-hydroxylase deficiency showing a signal at 1.49 min (blue), the urine of a child with 21-hydroxylase deficiency showing a no signal at 1.49 min (green), the urine of a child with an endocrine tumor showing signals at 1.32 min and 1.49 min (purple). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

defective 11-hydroxylase and the patient with an endocrine tumor exhibited a steroid metabolite with the retention time of 1.49 min, as the 16bOHA4 standard (Fig. 11). This metabolite was not found in the blood, but in the urine after sulfatase treatment. Additional analysis with a new LC-MS/MS method [40], could detect the sulfonated 16bOHA4 without sulfatase treatment as a whole molecule. These findings suggest that 16bOHA4 is very efficiently sulfonated in the phase II metabolism of the liver and designated for renal excretion. The fragmentation of this steroid, revealed the identical pattern as the one of the 16bOHA4 standard (Fig. S3), proving the existence of 16bOHA4 in humans with endocrine diseases. As expected, the urine of a patient with a 21-hydroxylase deficiency did not exhibit 16bOHA4, showing that CYP21A2 is crucial for the 16b-hydroxylation of A4 (Fig. 11). The presence of 16bOHA4 in humans suffering from 11-hydroxylase deficiency or from an endocrine tumor which was analysed by LC-MS/MS measurements was, furthermore, confirmed by GC-MS/MS analysis (data not shown).

The abovementioned findings indicate that 16bOHA4 can be formed in humans under excessive androgenic conditions. Presumably, the physiological concentrations of A4 are not sufficient to compete with the natural substrates of CYP21A2 and thus, A4 is not metabolized by this enzyme in healthy subjects under conditions of normal A4 levels. Therefore it should be considered whether 16bOHA4 could act as diagnostic marker, as it is present in certain hyperandrogenism-related diseases. One limitation of this study was the small number of patients and controls used to investigate the presence of 16bOHA4 and, therefore, these results are not representative for a specific disease and further research with representative cohorts is indispensable.

5. Summary and outlook

In this study the steroid metabolite 16bOHA4 was discovered employing an unconventional approach by selecting several steroid metabolites and testing their conversion by CYP21A2. The androgenic activity was low and comparable to the one of A4. Further, 16bOHA4 was shown to be present in patients with excessive androgen formation (11-hydroxylase deficiency, endocrine tumour). Therefore, we suggest that 16bOHA4 might possess a physiological meaning under certain endocrine circumstances, which still need to be elucidated. The fact that 16bOHA4 is absent in patients with 21-hydroxylase deficiency, which also results in increased androgen levels, shows that CYP21A2 is the exclusive human CYP capable to catalyse a hydroxylation at position C16 beta of A4. Keeping this in mind, it should be considered, whether 16bOHA4 might serve as a diagnostic marker.

Furthermore, 16bOHA4 was shown to be a substrate for CYP19A1, which is converted to an aromatized product, presumably to 16bOHE1. Due to an efficient *E. coli* based whole-cell system it was feasible to gain sufficient amounts of 16bOHA4 for the determination of kinetic parameters as well as binding studies, showing a tight binding and an effective conversion of 16bOHA4, similar to the natural substrates A4 and T. In contrast to E2 and E1, the estrogen receptor-assay experiments revealed that the estrogenic activity of p16bOHE1 was nearly absent. The facts that 16bOHA4 on the one hand, binds tightly to CYP19A1 and, on the other hand, the resulting product of the CYP19A1-dependent conversion displays a very low estrogenic activity, makes this steroid metabolite an promising compound acting as a CYP19A1 inhibitor. Furthermore, because of its single hydroxyl group,

16bOHA4 represents a valuable lead compound for the development of further steroidal CYP19A1 inhibitors, as the 16bOH group can be easily chemically modified.

Moreover, it was shown, that besides A4, more steroid intermediates can serve as a substrate for CYP21A2. These structures as well as their properties still remain to be elucidated and represent therefore an interesting starting point for further research in the field of steroid hormone biosynthesis.

Taken together, in this study we identified and characterized the steroid metabolite 16bOHA4, which was shown to be generated by CYP21A2 and further be metabolized by CYP19A1. The procedure presented here to reveal novel steroid metabolites is, furthermore, applicable to all steroidogenic enzymes and depicts therefore a new approach in steroid profiling.

Acknowledgements

The financial support of the Deutsche Forschungsgemeinschaft DFG FOR1369 is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsmb.2017.01.002>.

References

- [1] F. Hannemann, A. Bichet, K.M. Ewen, R. Bernhardt, Cytochrome P450 systems-biological variations of electron transport chains, *Biochim. Biophys. Acta* 1770 (2007) 330–344 (Epub 2006, Aug 2002).
- [2] R. Bernhardt, M.R. Waterman, Cytochrome P450 and steroid hormone biosynthesis, *The Ubiquitous Roles of Cytochrome P450 Proteins*, John Wiley & Sons, Ltd, 2007, 2017, pp. 361–396.
- [3] A. Pinto, B. Malacrida, J. Oieni, M.M. Serafini, A. Davin, V. Galbiati, E. Corsini, M. Racchi, DHEA modulates the effect of cortisol on RACK1 expression via interference with the splicing of the glucocorticoid receptor, *Br. J. Pharmacol.* 172 (2015) 2918–2927.
- [4] S. do Vale, L. Selinger, J.M. Martins, A.C. Gomes, M. Bicho, I. do Carmo, C. Escera, The relationship between dehydroepiandrosterone (DHEA), working memory and distraction—a behavioral and electrophysiological approach, *PLoS One* 9 (2014) e104869.
- [5] B. Mishra, J.Y. Park, K. Wilson, M. Jo, X-linked lymphocyte regulated gene 5c-like (Xlr5c-like) is a novel target of progesterone action in granulosa cells of periovulatory rat ovaries, *Mol. Cell. Endocrinol.* 412 (2015) 226–238.
- [6] S.T. Woods, J. Sadleir, T. Downs, T. Triantopoulos, M.J. Headlam, R.C. Tuckey, Expression of catalytically active human cytochrome p450sc in *Escherichia coli* and mutagenesis of isoleucine-462, *Arch. Biochem. Biophys.* 353 (1998) 109–115.
- [7] M. Arase, M.R. Waterman, N. Kagawa, Purification and characterization of bovine steroid 21-hydroxylase (P450c21) efficiently expressed in *Escherichia coli*, *Biochem. Biophys. Res. Commun.* 344 (2006) 400–405.
- [8] A. Zollner, N. Kagawa, M.R. Waterman, Y. Nonaka, K. Takio, Y. Shiro, F. Hannemann, R. Bernhardt, Purification and functional characterization of human 11beta hydroxylase expressed in *Escherichia coli*, *FEBS J.* 275 (2008) 799–810.
- [9] A. Hobler, N. Kagawa, M.C. Hutter, M.F. Hartmann, S.A. Wudy, F. Hannemann, R. Bernhardt, Human aldosterone synthase: recombinant expression in *E. coli* and purification enables a detailed biochemical analysis of the protein on the molecular level, *J. Steroid Biochem. Mol. Biol.* 132 (2012) 57–65.
- [10] N. Kagawa, Efficient expression of human aromatase (CYP19) in *E. coli*, *Methods Mol. Biol.* 705 (2011) 109–122.
- [11] J. Neunzig, A. Sanchez-Guijo, A. Mosa, M.F. Hartmann, J. Geyer, S.A. Wudy, R. Bernhardt, A steroidogenic pathway for sulfonated steroids: the metabolism of pregnenolone sulfate, *J. Steroid Biochem. Mol. Biol.* 144 (Pt B) (2014) 324–333.
- [12] T. Kawamoto, Y. Mitsuchi, K. Toda, K. Miyahara, Y. Yokoyama, K. Nakao, K. Hosoda, Y. Yamamoto, H. Imura, Y. Shizuta, Cloning of cDNA and genomic DNA for human cytochrome P-45011 beta, *FEBS Lett.* 269 (1990) 345–349.
- [13] A. Mosa, J. Neunzig, A. Gerber, J. Zapp, F. Hannemann, P. Pilak, R. Bernhardt, 2beta- and 16beta-hydroxylase activity of CYP11A1 and direct stimulatory effect of estrogens on pregnenolone formation, *J. Steroid Biochem. Mol. Biol.* 150 (2015) 1–10.
- [14] T. van der Hoeven, Testosterone oxidation by rat liver microsomes: effects of phenobarbital pretreatment and the detection of seven metabolites by HPLC, *Biochem. Biophys. Res. Commun.* 100 (1981) 1285–1291.
- [15] A.J. Draper, A. Madan, K. Smith, A. Parkinson, Development of a non-high pressure liquid chromatography assay to determine testosterone hydroxylase (CYP3A) activity in human liver microsomes, *Drug Metab. Dispos. Biol. Fate Chem.* 26 (1998) 299–304.
- [16] N. Kagawa, Efficient expression of human aromatase (CYP19) in *E. coli*, *Methods Mol. Biol.* 705 (2011) 109–122.
- [17] D. Sandee, W.L. Miller, High-yield expression of a catalytically active membrane-bound protein: human P450 oxidoreductase, *Endocrinology* 152 (2011) 2904–2908.
- [18] J. Neunzig, R. Bernhardt, Dehydroepiandrosterone sulfate (DHEAS) stimulates the first step in the biosynthesis of steroid hormones, *PLoS One* 9 (2014) e89727.
- [19] Y. Khatri, Y. Luthra, R. Duggal, S.G. Sligar, Kinetic solvent isotope effect in steady-state turnover by CYP19A1 suggests involvement of Compound 1 for both hydroxylation and aromatization steps, *FEBS Lett.* 588 (2014) 3117–3122.
- [20] S.L. Gantt, I.G. Denisov, Y.V. Grinkova, S.G. Sligar, The critical iron-oxygen intermediate in human aromatase, *Biochem. Biophys. Res. Commun.* 387 (2009) 169–173.
- [21] T. Omura, R. Sato, The carbon monoxide-binding pigment of liver microsomes. ii. solubilization, purification, and properties, *J. Biol. Chem.* 239 (1964) 2379–2385.
- [22] J.B. Schenkman, Studies on the nature of the type I and type II spectral changes in liver microsomes, *Biochemistry* 9 (1970) 2081–2091.
- [23] S. Brixius-Anderko, F. Hannemann, M. Ringle, Y. Khatri, R. Bernhardt, An indole deficient *Escherichia coli* strain improves screening of cytochromes P450 for biotechnological applications, *Biotechnol. Appl. Biochem.* (2016), doi:<http://dx.doi.org/10.1002/bab.1488>.
- [24] M. Ringle, Y. Khatri, J. Zapp, F. Hannemann, R. Bernhardt, Application of a new versatile electron transfer system for cytochrome P450-based *Escherichia coli* whole-cell bioconversions, *Appl. Microbiol. Biotechnol.* 97 (2013) 7741–7754.
- [25] A. Sanchez-Guijo, J. Neunzig, A. Gerber, V. Oji, M.F. Hartmann, H.C. Schuppe, H. Traupe, R. Bernhardt, S.A. Wudy, Role of steroid sulfatase in steroid homeostasis and characterization of the sulfated steroid pathway: evidence from steroid sulfatase deficiency, *Mol. Cell. Endocrinol.* 437 (2016) 142–153.
- [26] R. Huey, G.M. Morris, A.J. Olson, D.S. Goodsell, A semiempirical free energy force field with charge-based desolvation, *J. Comput. Chem.* 28 (2007) 1145–1152.
- [27] D. Ghosh, J. Lo, D. Morton, D. Valette, J. Xi, J. Griswold, S. Hubbell, C. Egbuta, W. Jiang, J. An, H.M. Davies, Novel aromatase inhibitors by structure-guided design, *J. Med. Chem.* 55 (2012) 8464–8476.
- [28] M.F. Sanner, Python: a programming language for software integration and development, *J. Mol. Graph. Model.* 17 (1999) 57–61.
- [29] L.M. Bloem, K.H. Storbek, L. Schloms, A.C. Swart, 11beta-hydroxyandrostenedione returns to the steroid arena: biosynthesis, metabolism and function, *Molecules* 18 (2013) 13228–13244.
- [30] D.N. Kirk, H.C. Toms, C. Douglas, K.A. White, K.E. Smith, S. Latif, R.W.P. Hubbard, A survey of the high-field 1H NMR spectra of the steroid hormones, their hydroxylated derivatives, and related compounds, *J. Chem. Soc. Perkin Trans. 2* (2) (1990) 1567–1594.
- [31] S. Z. Ge, Y. Mao, X. Li, F. Lu Liu, 16beta-hydroxylation of 4-androstene-3,17-dione by *Aspergillus niger*, *Sheng wu gong cheng xue bao=Chin. J. Biotechnol.* 30 (2014) 1481–1485.
- [32] D. Doller, E.G. Gros, Carbon-13 nuclear magnetic resonance spectral data of steroid vicinal ketols and related compounds, *Steroids* 56 (1991) 168–172.
- [33] J. Neunzig, M. Widjaja, C.A. Dragan, F.T. Peters, H.H. Maurer, M. Bureik, Engineering of human CYP3A enzymes by combination of activating polymorphic variants, *Appl. Biochem. Biotechnol.* 168 (2012) 785–796.
- [34] V.H. Black, J.C. Wittig, P. Cheung, Intraadrenal steroid metabolism in the guinea pig: guinea pig adrenal microsomes metabolize androstenedione in a manner distinct from liver microsomes, *Endocr. Res.* 21 (1995) 315–328.
- [35] F. Chen, K. Knecht, C. Leu, S.J. Rutledge, A. Scafanas, C. Gambone, R. Vogel, H. Zhang, V. Kasparcova, C. Bai, S. Harada, A. Schmidt, A. Reszka, L. Freedman, Partial agonist/antagonist properties of androstenedione and 4-androsten-3beta, 17beta-diol, *J. Steroid Biochem. Mol. Biol.* 91 (2004) 247–257.
- [36] M. Dowsett, R.C. Coombes, Second generation aromatase inhibitor—4-hydroxyandrostenedione, *Breast Cancer Res. Treat.* 30 (1994) 81–87.
- [37] W. Yue, A.M. Brodie, Mechanisms of the actions of aromatase inhibitors 4-hydroxyandrostenedione, fadrozole, and aminoglutethimide on aromatase in JEG-3 cell culture, *J. Steroid Biochem. Mol. Biol.* 63 (1997) 317–328.
- [38] M. Lippman, M.E. Monaco, G. Bolan, Effects of estrone, estradiol, and estril on hormone-responsive human breast cancer in long-term tissue culture, *Cancer Res.* 37 (1977) 1901–1907.
- [39] S. Sasson, A.C. Notides, Estriol and estrone interaction with the estrogen receptor. I. Temperature-induced modulation of the cooperative binding of [3H]estriol and [3H]estrone to the estrogen receptor, *J. Biol. Chem.* 258 (1983) 8113–8117.
- [40] A. Sanchez-Guijo, V. Oji, M.F. Hartmann, H. Traupe, S.A. Wudy, Simultaneous quantification of cholesterol sulfate, androgen sulfates, and progesterone sulfates in human serum by LC-MS/MS, *J. Lipid Res.* 56 (2015) 1843–1851.

Supplementary material

The steroid metabolite 16(β)-OH-androstenedione generated by CYP21A2 serves as a substrate for CYP19A1

**Neunzig, J., Milhim, M., Schiffer, L., Khatri, Y., Zapp, J., Sánchez-Guijo, A., Hartmann, M.F.,
Wudy, S.A., Bernhardt, R.***

Institute of Biochemistry, Saarland University, 66123 Saarbrücken, Germany

*Address correspondence to:

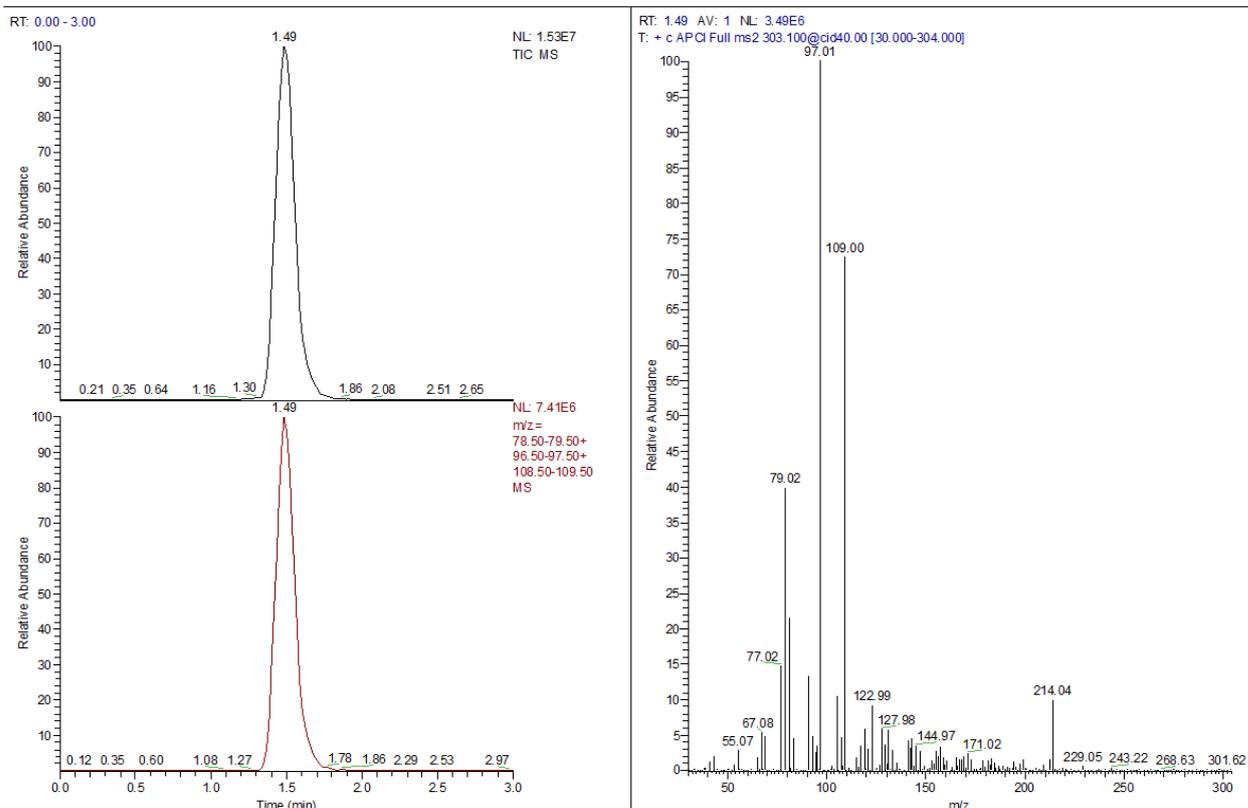
Prof. Dr. Rita Bernhardt, Institute of Biochemistry, Campus B 2.2, Saarland University, 66123
Saarbrücken (Germany)

E-mail: ritabern@mx.uni-saarland.de

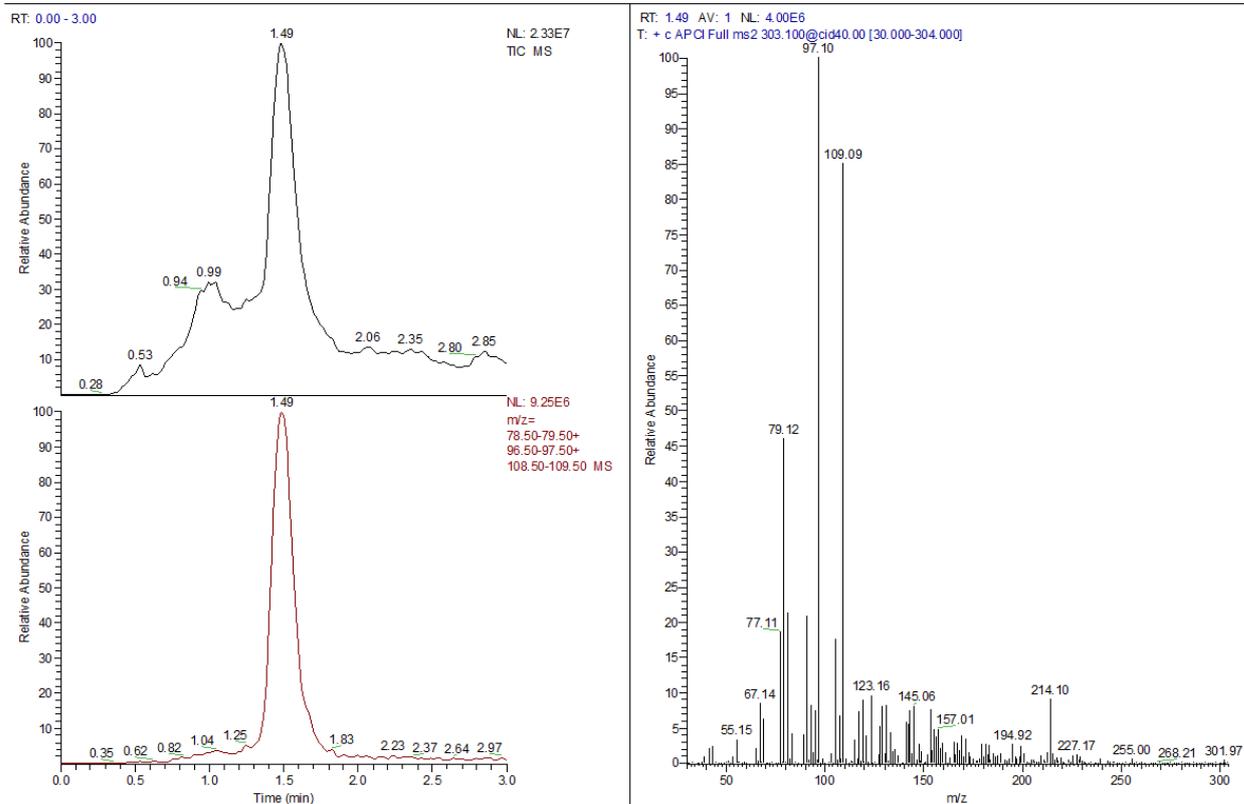
Phone: +49 681 302 4241

Fax: +49 681 302 4739

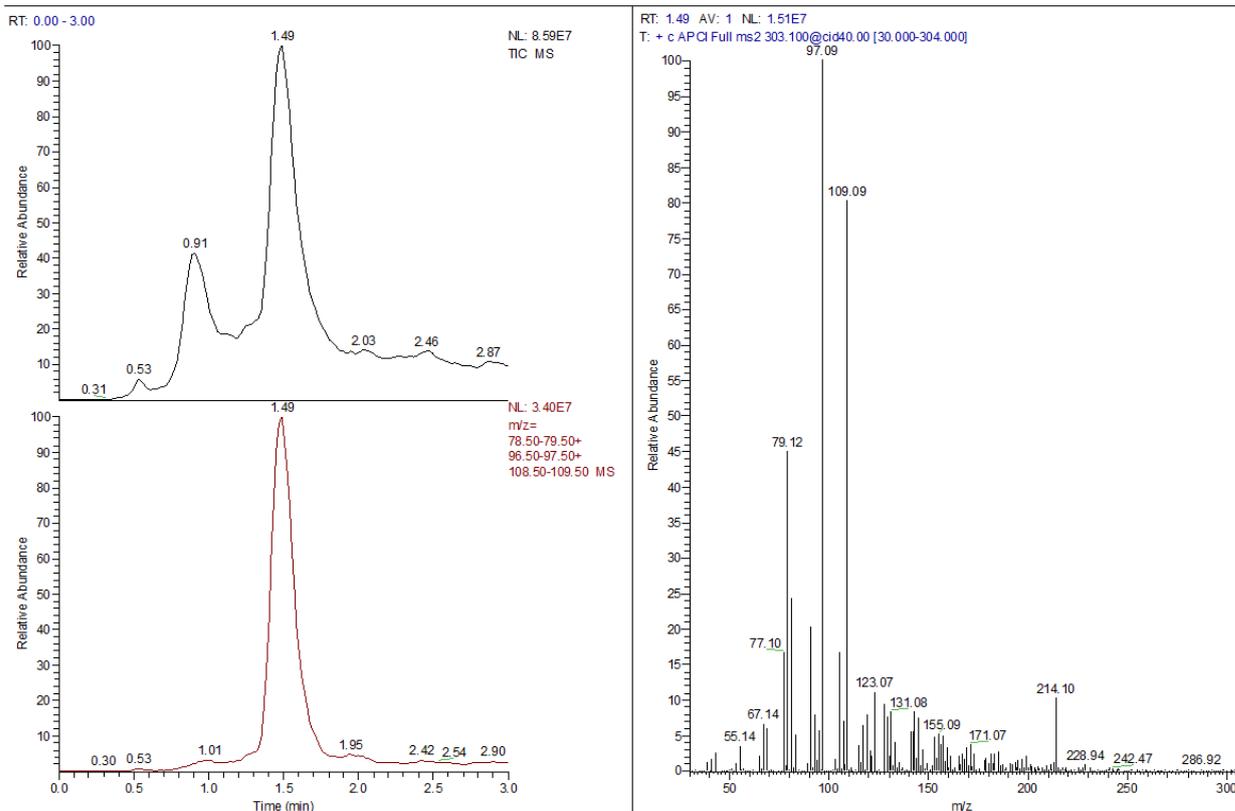
Figure S1: LC-MS/MS data- detection of 16OHA4 in human samples

12.5 µg/ml standard

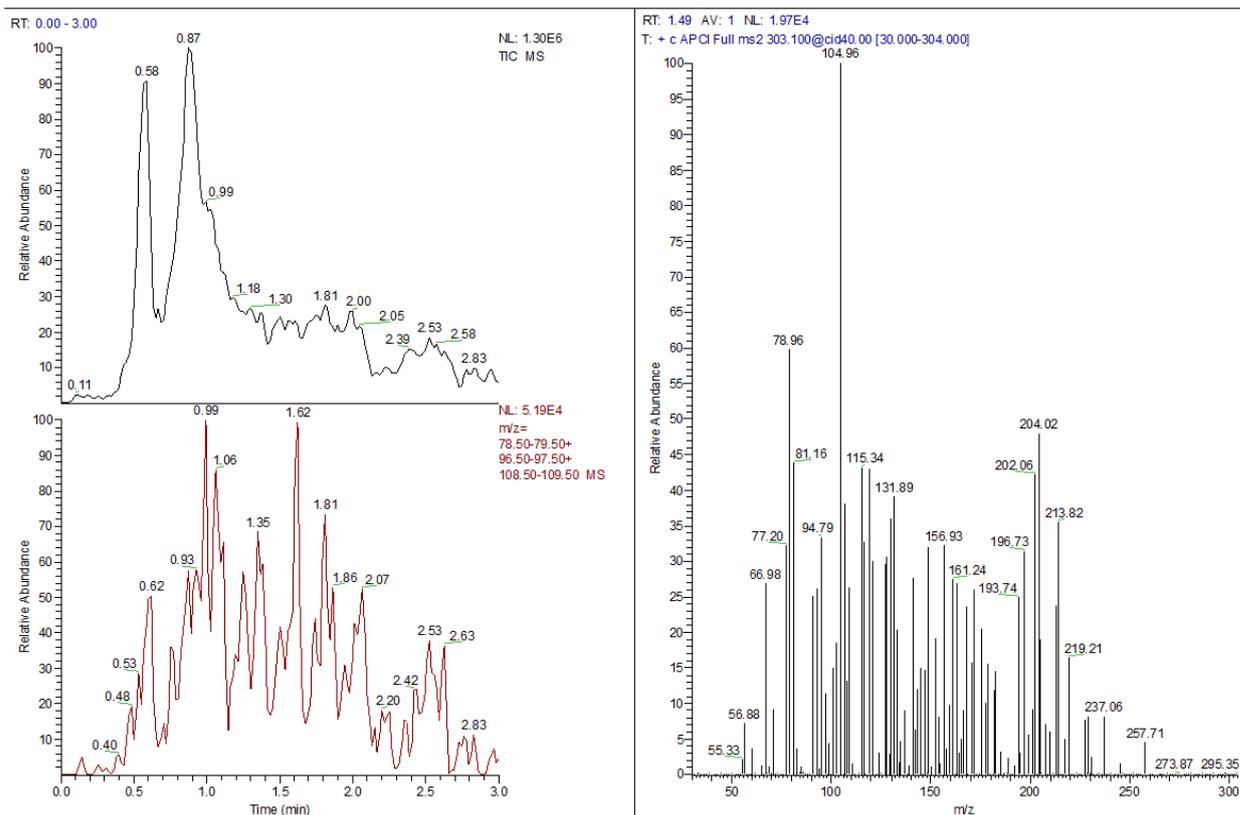
Urine from a male patient with 11-hydroxylase deficiency (2.4 years old, A4 concentration in plasma
12.5 ng/ml)



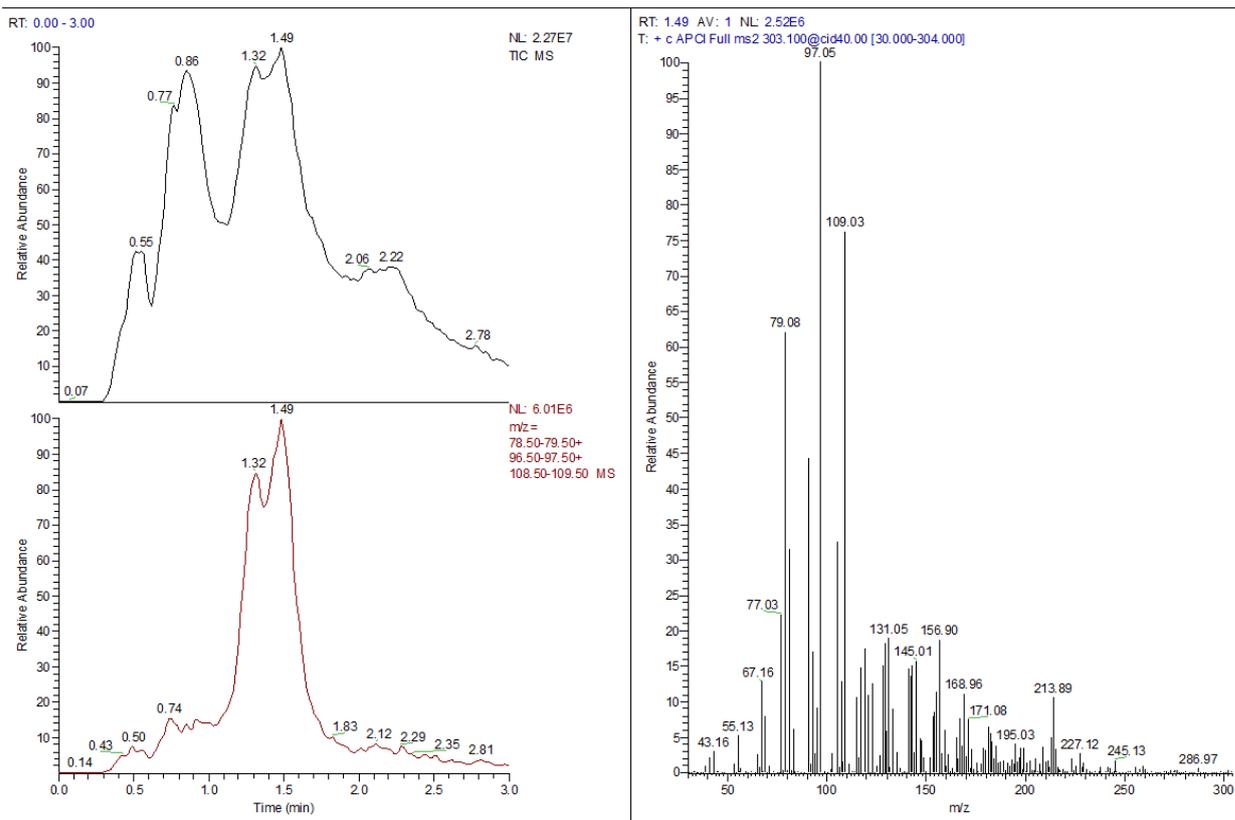
Urine from a male patient with 11-hydroxylase deficiency (3.4 years old, A4 concentration in plasma not measured)



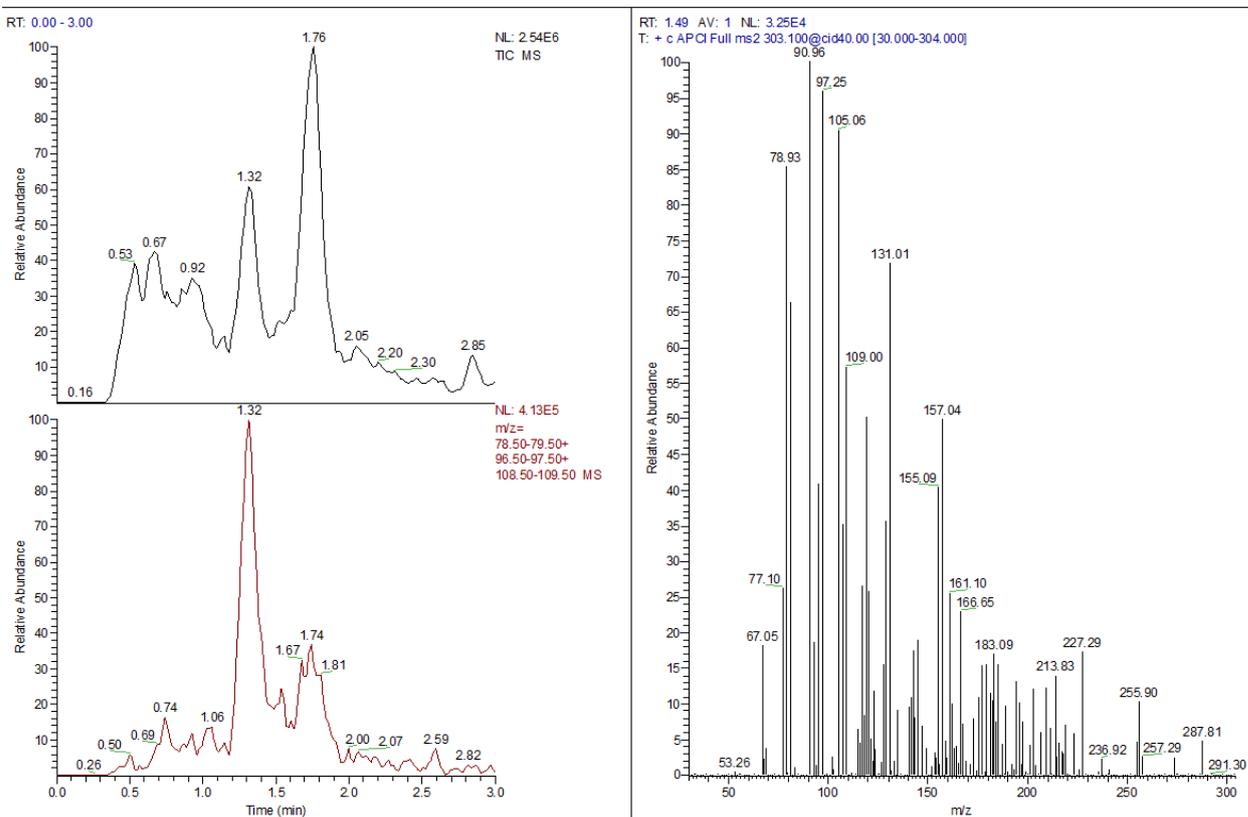
Control urine from a male healthy child 3.3 years old



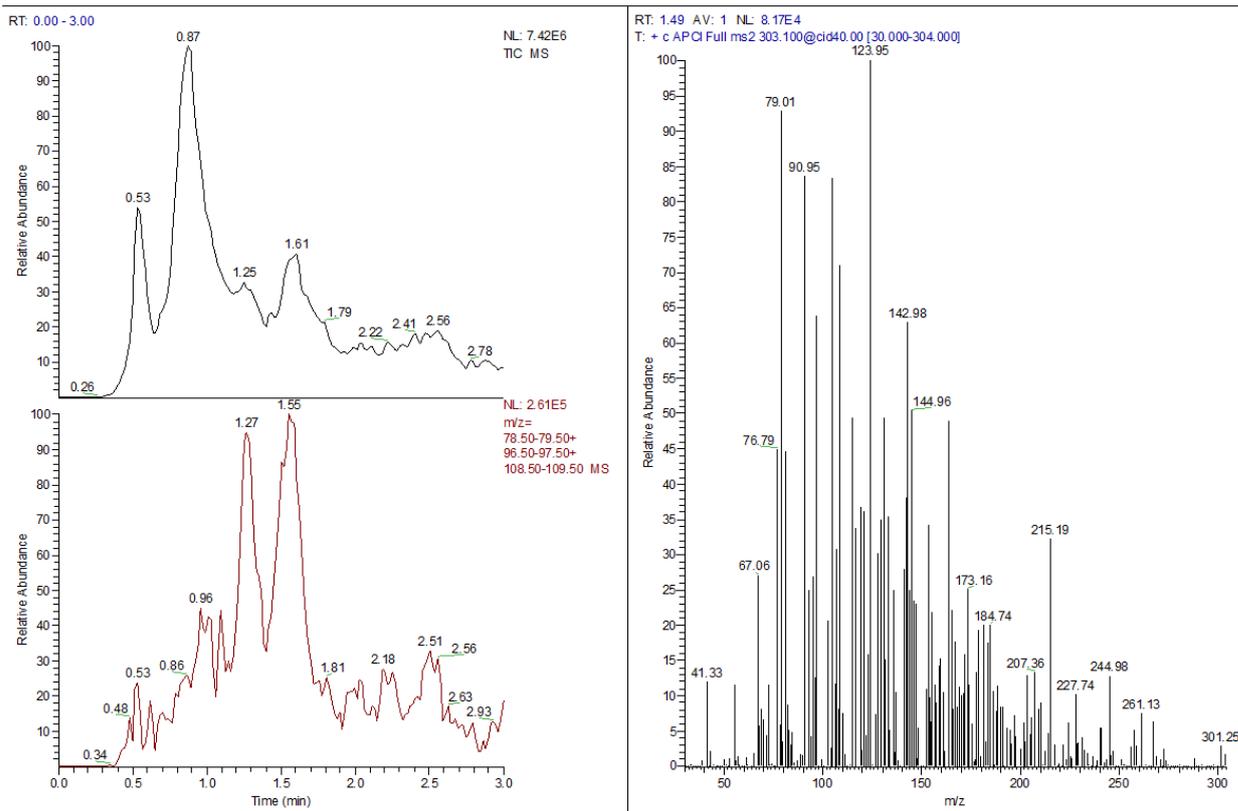
Urine from a female patient with an endocrine tumor (15.4 years old, A4 concentration not measured)



Urine from an untreated male patient with 21-hydroxylase deficiency (12.1 years old, A4 concentration not measured)



Control urine from a male healthy child 12.5 years old



S1: All urines after sulfatase treatment. Black chromatograms on the left panel show the product ion scan mode for m/z 303 at 40 eV collision energy for every sample. Red chromatograms are obtained as extracted chromatograms for the transition 303→79, 97, 109. On the right panel, the spectrum with the fragments obtained at a retention time of 1.49 min are depicted.

S2: NMR data of 16bOHA4

^1H NMR (CDCl_3 , 500 MHz): δ 0.98 (s, 3xH-18), 1.01 (ddd, $J=12.5$, 10.5 and 4.0 Hz, H-9), 1.11 (m, H-7a), 1.20 (s, 3xH-19), 1.24 (m, H-14), 1.33 (td, $J=13.0$ and 4.5 Hz, H-12a), 1.51 (tdd, $J=13.5$ 12.5 and 4.2 Hz, H-11a), 1.54 (ddd, $J=3.5$, 12.0 and 9.5 Hz, H-15a), 1.68 (m, H-11b), 1.70 (m, H-1a), 1.78 (dddd, $J=3 \times 11.0$ and 3.5 Hz, H-8), 1.91 (ddd, $J=13.0$, 4.2 and 2.8 Hz, H-12b), 1.96 (m, H-7b), 2.02 (ddd, $J=13.5$, 5.0 and 3.5 Hz, H-1b), 2.32 (m, H-6a), 2.34 (m, H-2a), 2.39 (m, 2H, H-2b and H-15b), 2.41 (m, H-6b), 3.95 (dd, $J=9.5$ and 8.2 Hz, H-16), 5.74 br s (H-4). ^{13}C NMR (CDCl_3 , 125 MHz): δ

14.84 (CH₃, C-18), 17.35 (CH₃, C-19), 20.11 (CH₂, C-11), 30.44 (CH₂, C-15), 30.99 (CH₂, C-7), 31.52 (CH₂, C-12), 32.46 (CH₂, C-6), 33.88 (CH₂, C-2), 34.24 (CH, C-8), 35.67 (CH₂, C-1), 38.69 (C, C-10), 45.24 (CH, C-14), 46.68 (C, C-13), 54.06 (CH, C-9), 75.21 (CH, C-16), 124.25 (CH, C-4), 169.88 (C, C-5), 199.17 (C, C-3), 219.52 (C, C-17).

3 DISCUSSION

The large-scale sequencing and the rapid completion of genomic sequences of many organisms led to far more gene products (Open Reading Frames) than functions, which can be referred to. Moreover, many of the 'known' functions may be uncertain as they are unsupported by experimental evidences, even in an organism as well studied as *E. coli*, in which there is experimental information for only 54% of the gene products (Guengerich et al., 2010).

Two approaches can be applied for the elucidation of the gene function. The classical biochemistry paradigm (or called forward strategy), which goes from an observation of a particular phenomenon or phenotype to the identification of the corresponding gene. While, the reverse genetic (or called backward strategy) goes from the gene sequence to the protein characterization and subsequently identification of the function.

The publication of the complete genome sequence of the *B. megaterium* strains QM B1551 and DSM319 in 2011 (Eppinger et al., 2011) and the immense development of bioinformatic and biochemical methods enabled the characterization of new proteins in this work for which the backward strategy (genome mining) has been applied.

3.1 CHARACTERIZATION OF A NOVEL NADPH-DEPENDENT DIFLAVIN REDUCTASE

As pointed out earlier, P450s rely for their activities on redox partners. The ability of the *B. megaterium* based whole cell system/s, expressing microsomal or bacterial P450s without co-expression of redox partners, to show activity towards the corresponding substrate/s suggested the presence of still unidentified redox partner/s (Bleif et al., 2012; Brill et al., 2013; Gerber et al., 2015).

Searching the *B. megaterium* genome for a homolog to the mammalian diflavin P450 reductase (CPR) revealed an open reading frame (BMD_3122) encoding an NADPH diflavin reductase, which was characterized and investigated at the bioinformatic and biochemical level in publication 2.1. The new

reductase was designated as *B. megaterium* cytochrome P450 reductase (BmCPR). The *BmCPR* gene encoding a protein of 602 amino acids was successfully cloned and expressed in *E. coli*. The purified soluble enzyme exhibited a characteristic diflavin reductase spectrum with a maximum peak at 380 nm, 456 nm and a broad peak at 585 nm. Sequence analysis showed that the (FMN-, FAD- and NADPH-) binding motifs are conserved, in addition to the presence of a cluster of negatively charged residues, which are thought to be responsible for the interaction with the cytochrome P450 as well as cytochrome c. BmCPR showed a cytochrome c reducing activity with a K_M of 16.7 μM and k_{cat} of 2,582 min^{-1} . Furthermore, in an *in vitro* reconstituted system, BmCPR was found to support the 21 α - and 17 α -hydroxylation of progesterone with the class II microsomal bovine CYP21A2 and CYP17A1, respectively, with a higher activity than that of the bovine NADPH-dependent cytochrome P450 reductase (CPR). In class I bacterial P450 systems, FAD of the ferredoxin-reductase serves as an electron acceptor from NAD(P)H, which transfers the reducing equivalents to the iron-sulfur center of ferredoxin, which in turn reduces the P450s (Hannemann et al., 2007). Our results showed that the BmCPR, in combination with different ferredoxins, was also able to support the activity of CYP106A1 *in vitro*, with the best results using the BmCPR-Fdx2 combination.

A phylogenetic analysis of the deduced BmCPR protein and comparison with other related diflavin reductases by the neighbor joining method showed that it belongs to the cytochrome P450_{like} reductase (CYPOR_{like}) sub-family having FAD and FMN as prosthetic groups and utilizing NADPH as electron donor. This sub-family comprises different members. Cytochrome P450 reductase (CPR) was the first enzyme of this family to be isolated (Horecker, 1950), followed by several other dual flavin enzymes, sulfite reductase in bacteria (Christner et al., 1981; Ostrowski et al., 1989) as well as three proteins identified in human: nitric oxide synthase NOS (Stuehr, 1997), methionine synthase reductase (Olteanu and Banerjee, 2001) and a cytoplasmic protein NR1 with yet unknown function that is expressed in cancer cells (Paine et al., 2000). This family includes also the diflavin domain of self-sufficient cytochromes P450 such as CYP102A1 (BM3) from *B. megaterium*. BmCPR is most related to the flavoprotein alpha-component of sulfite reductase (SiR) (44 % - 63%), an indication of the cross-functionality of the sub-family members. It is proposed that these family member enzymes

evolved from a common ancestor as a result of a fusion of two proteins, the bacterial flavodoxin (FMN-domain) and the ferredoxin-NADP⁺ reductase FNR (FAD-domain).

P450s are potential enzymes for biotechnological applications; nevertheless, the multi-component nature of P450s, in other words the absolute requirement of electron transfer partner/s for the P450 reaction is one of the most important factors, which hampered their powerful commercial utilization. Therefore, the co-expression of genes encoding the redox partners of the P450 is necessary when the host cell lacks such electron transfer proteins (*E. coli*) or the efficiency using an endogenous system is low.

BmCPR was employed in the establishment of whole cell biocatalyst systems with class I and II P450s as shown in publication 2.1 and 2.2 (Milhim et al., 2016a, 2016b). Therefore, the identification of this reductase provides a versatile redox partner for P450s with appealing biotechnological potential (Figure 4.1).

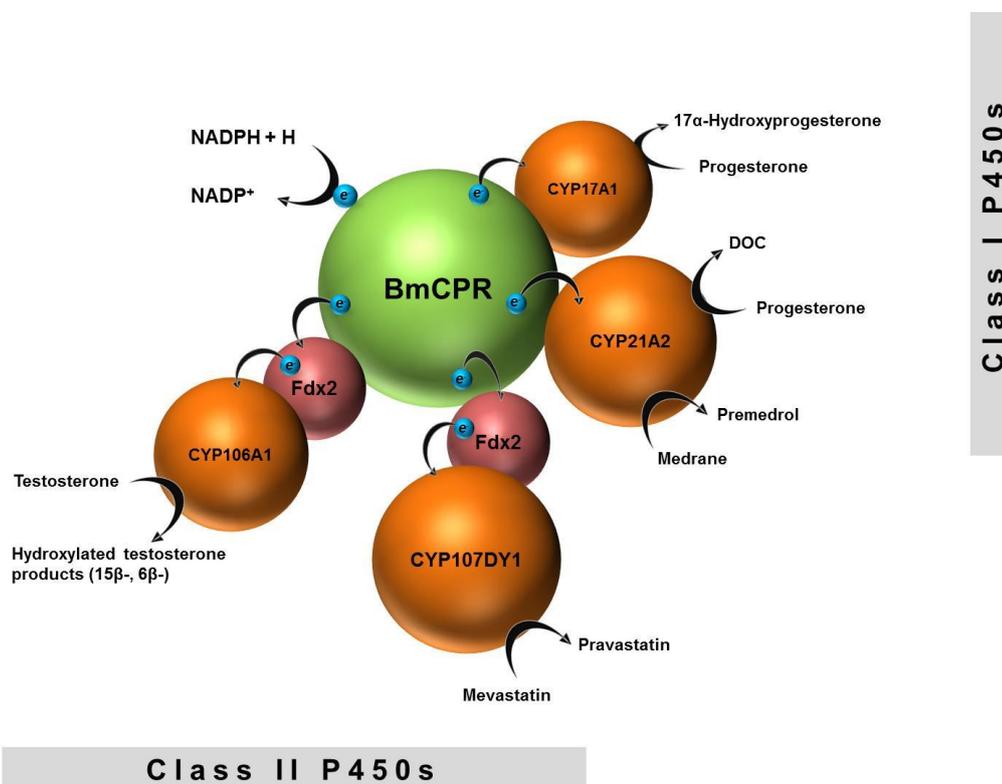


Figure 4.1. Schematic representation of the ability of BmCPR to support electrons transfer to class I (represented by CYP106A1 and CYP107DY1) and class II (represented by CYP21A2 and CYP17A1) P450s.

Some of the most potent anti-inflammatory and immunosuppressive agents are the synthetic glucocorticoids which are, therefore, considered to be very crucial in the pharmaceutical industry. Among them is medrol (methylprednisolone), which is widely used for the treatment of inflammatory and autoimmune diseases, such as: asthma (Fiel and Vincken, 2006); uveitis, rheumatoid arthritis (Saag, 2002); ulcerative colitis (Rosenberg et al., 1990); Crohn's disease (inflammatory bowel disease); Bell's palsy; multiple sclerosis (Thrower, 2009); cluster headaches; vasculitis; acute lymphoblastic leukemia (Lambrou et al., 2009); systemic lupus erythematosus, dermatomyositis and autoimmune hepatitis (Davis et al., 1978).

Medrol can be synthesized from the intermediate predmedrol. In our laboratories we successfully developed a CYP21A2-based whole-cell system in *E. coli* for the biotechnological 21-hydroxylation of the precursor medrane to produce predmedrol. It was demonstrated that the reaction efficiency is strongly dependent on the co-expressed redox partner with the best yield when the *S. pombe* redox system (Arh1/etp1^{fd}) was used (Brixius-Anderko et al., 2015).

The efficiency of the BmCPR to support CYP21A2 for the conversion of medrane to predmedrol was compared *in vivo* with Arh1/etp1^{fd}. The expression and conversion were carried out as described previously (Brixius-Anderko et al., 2015). **Figure 4.2** shows that the BmCPR was able to support the activity of CYP21A2 *in vivo*, with about 1.14 fold higher efficiency than the previously used Arh1/etp1^{fd} system.

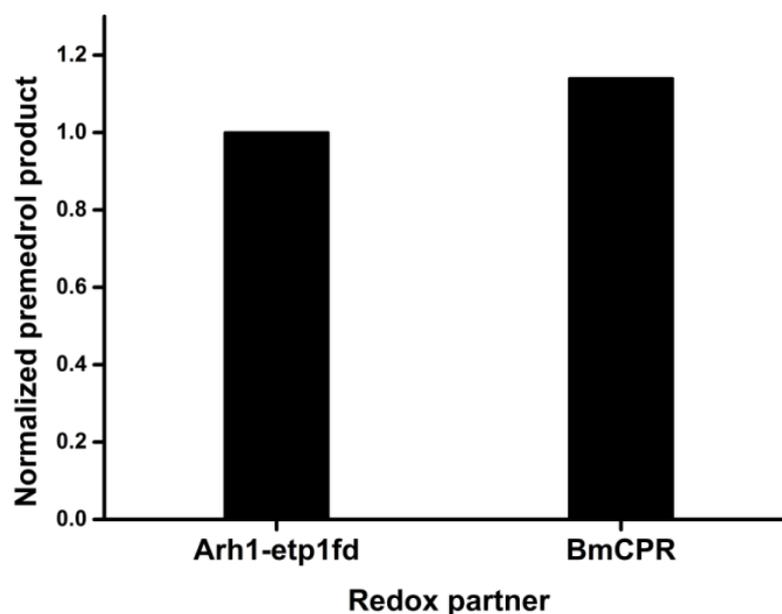


Figure 4.2. CYP21A2-dependent whole-cell conversion of medrane to premedrol using Arh1/etp1^{fd} or BmCPR as redox partner. Whole-cell biotransformation was performed using C43 (DE3) cells carrying pET17b vector expressing CYP21A2 and the redox partner (Arh1/etp1^{fd} or BmCPR). Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4 + 2% glycerol) and samples were taken after 24 hr conversion^(*). The data were normalized by setting Arh1/etp1^{fd} to 1.

^(*)This experiment was kindly carried out by Dr. Brixius-Anderko S.

The results pointed out the potential of the newly identified BmCPR in future biotechnological applications. In general, the use of BmCPR is advantageous since the expression of a soluble bacterial reductase is much easier than that of mammalian membrane anchored CPRs. Moreover, the thermal stability of the BmCPR is higher than that of its mammalian counterparts (publication 2.1).

Therefore, the characterization and deeper understanding of such a reductase is of great importance for the re-evaluation and exploration of the phylogenetic relationships of NADPH cytochrome P450 reductases between different taxa, which enables us to characterize non-conventional alternative redox partners.

3.2 CYP107DY1 A NEW PLASMID-ENCODED CYTOCHROME P450 FROM *BACILLUS MEGATERIUM*

B. megaterium strains QM B1551 harbors seven indigenous plasmids (Eppinger et al., 2011). About 11% of its cellular DNA was found to be carried on these plasmids, which emphasizes the importance and potential of the genes encoded by them. Sequence analysis of the indigenous plasmids showed that plasmid number 5 (pBM500) has an open reading frame (BMQ_pBM50008) containing 1,233 base pairs, which encodes for a protein comprising 410 amino acids. The amino acid sequence has a predicted isoelectric point pI of 5.81 and a molecular weight of about 46.742 kDa. Analysis of the protein domains using the Pfam Database (Finn et al., 2016) proved that BMQ_pBM50008 bears all highly conserved motifs of the P450 family. The new P450 was investigated in details in publication 2.2 (Milhim et al., 2016b). Considering the BMQ_pBM50008 amino acids sequence, three highly conserved regions for P450 characteristics were found. First, the heme-binding motif (F-x-x-G-x-x-x-C-x-G) was found at position 353 – 362 in the deduced amino acid sequence. This motif forms the heme binding loop, where the absolutely conserved cysteine coordinates as fifth ligand with the heme iron. The second highly conserved motif is (A/G-G-x-E/D-T-T/S), which is found in the I-helix and regarded as an oxygen-binding motif with the conserved glycine pointing at the center of the heme and the conserved threonine pointing to the oxygen-binding site. In addition, the presence of the (E-x-x-R) motif in the K-helix is responsible for the formation of a salt bridge, which is crucial for the heme center stabilization (Danielson, 2003).

Multiple sequence alignment and submission of the protein sequence of BMQ_pBM50008 to P450 nomenclature committee (Prof. Dr. David Nelson), turned out that it matches best to CYP107DA1 with 49% sequence identity and, therefore, was assigned as a new subfamily with the name CYP107DY1.

CYP107DY1 was successfully expressed in *E. coli* and purified. The recombinant CYP107DY1 exhibits a characteristic P450 absolute and reduced CO-bound difference spectrum. In addition, the CD spectra in the Far-UV and near UV-visible region were in correspondence with the characteristic

peaks for other bacterial P450s (Lepesheva et al., 2001; Munro et al., 1994), an indication that the purified enzyme is produced in the active form with proper heme incorporation.

Efforts were undertaken to search for suitable redox partner/s of CYP107DY1 since an effective electron transfer is essential in terms of cytochrome P450 activity. Based on the measurement of the reduced CO-bound spectrum of the CYP107DY1, the *Bacillus* redox system BmCPR-Fdx2 (publication 2.1) was found to be able to support the electron transfer from NADPH to CYP107DY1 and, therefore, it was used for substrates screening for CYP107DY1. Phylogenetically CYP107DY1 was found to be related to the CYP267B1 from *Sorangium cellulosum* So ce56, which was previously characterized as versatile drug metabolizer (Kern et al., 2016). Therefore, substrate screening using a drug library consisting of 18 drugs pointed out the ability of CYP107DY1 to convert mevastatin (compactin) to pravastatin as shown in publication 2.2.

Statin drugs inhibit the 3 β -hydroxy-3methylgluteryl CoA (HMG-CoA) reductase, the key enzyme involved in the rate limiting step of cholesterol biosynthesis. Statins reduce the low density lipoprotein (LDL) cholesterol level and thus are very effective in the treatment of hypercholesterolemia (Lamon-Fava, 2013). Among the most widely used statins is pravastatin, which is currently obtained by the regioselective hydroxylation of the natural product mevastatin (compactin). Mevastatin was first isolated from *Penicillium citrinum* (Endo et al., 1976) and *Penicillium brevicompactum* (Brown et al., 1976). An oxyfunctionalization of mevastatin to pravastatin was described using CYP105A3 (P450sca2) from *Streptomyces carbophilus* (Matsuoka et al., 1989) and CYP105AS1 from *Amycolatopsis orientalis* (McLean et al., 2015).

3.3 IN VIVO WHOLE CELL PRAVASTATIN PRODUCTION

After the successful *in vitro* conversion of mevastatin by CYP107DY1 into one main product pravastatin and due to the high pharmaceutical interest in the oxyfunctionalization of mevastatin, the biotechnological relevance of the newly identified P450 and reductase (publication 2.1 and 2.2) was assessed by construction of recombinant *E. coli*-based whole-cell system overexpressing CYP107DY1 together with the redox partners BmCPR-Fdx2.

The whole-cell biocatalyst system was successfully used for *in vivo* production of about 28.5 mg/L pravastatin (publication 2.2). This is still far away from the implementation of this system in an industrial biocatalyst process. However, the economic feasibility of the known examples of the industrially used P450s needed developmental research of more than 10 years (Bernhardt, 2006; Bernhardt and Urlacher, 2014; Girvan and Munro, 2016; Julsing et al., 2008). This started from gene characterization and functional identification, successful construction of an *in vivo* system, optimization of the system at the molecular (enzyme activity) and cellular level (use of industrially optimized and engineered host) before it was implemented in an industrial process. Therefore, the identification of CYP107DY1 as a highly regioselective mevastatin hydroxylase and the successful *in vivo* application is the first step for its future industrial use.

3.4 A NEW SHORT-CHAIN DEHYDROGENASE/REDUCTASE FROM *BACILLUS MEGATERIUM* WITH 17 β -HSD ACTIVITY

Biotransformation of testosterone with *B. megaterium* based whole cells harboring the pSMF2.1 empty vector resulted in the formation of the 17-oxo steroid form androst-4-ene-3,17-dione, an indication of the ability of *B. megaterium* to oxidize the 17 β -hydroxy- to 17-keto steroid. This is common in microorganisms and was firstly reported in 1937 by Mamoli and Vercelloni who described the transformation of androst-4-ene-3,17-dione to testosterone by *S. cerevisiae* (Liu et al., 2017). Later, this activity was also demonstrated in other microorganisms of different taxonomy such as bacteria (*Pseudomonas*, *Comamonas*, *Bacillus*, *Brevibacterium*, *Mycobacterium*, *Streptomyces*), fungi (*Aspergillus*, *Mucor*, *Penicillium*), yeasts (*Hansenula*, *Pichia*, *Saccharomyces*), algae (*Chlorella*, *Chlorococcum*) and protozoa (*Pentatrichomonas*, *Trichomonas*) (Donova et al., 2005).

To identify the enzyme, which possesses 17 β -hydroxysteroid dehydrogenase activity, the human testosterone 17-beta-dehydrogenase (17 β -HSD) (UniProt acc. Number: P37058) was blasted against the *B. megaterium* genome using MegaBac v9 (<http://megabac.tu-bs.de>). The best bit score was achieved with the open reading frame BMD_2094 with a sequence similarity of 39.4% with the human 17 β -HSD (ClustalW pairwise sequence alignment). BMD_2094 [Uniprot acc. No. D5DFE7] encodes a

259 amino acids polypeptide with a calculated molecular mass of the deduced amino acid sequence of 28.85 kDa and a pI of 9.40. The BMD_2094 protein belongs to the short-chain dehydrogenase/reductase (SDR) superfamily, which is found in all domains of life (Kallberg et al., 2010). SDRs are a functionally diverse family of NAD(P)(H)-dependent oxidoreductases and are known for a broad substrate spectrum ranging from alcohols, sugars, steroids and aromatic compounds to xenobiotics. The N-terminal region binds the coenzymes NAD(H) or NADP(H), while the C-terminal region constitutes the substrate binding part (Kallberg et al., 2002).

The activity of the candidate enzyme was verified by an overexpression approach. The 17β -HSD activity of the BMD_2094 towards testosterone substrate was assessed by HPLC after carrying out *in vivo* whole-cell conversions. As shown in **Figure 4.3A**, compared to the control strain (*B. megaterium* carrying the empty vector pSMF2.1), formation of androst-4-ene-3,17-dione is drastically increased in *B. megaterium* overexpressing the BMD_2094 with a conversion ratio of more than 88 %, yielding 0.6 g/L/day. This demonstrates the 17β -hydroxysteroids oxidation activity of BMD_2094. The ability of BMD_2094 to catalyze the reverse reaction (reduction of the oxygen at position 17) was assessed by using androst-4-ene-3,17-dione as a substrate. The reverse reaction is clearly weaker than the forward reaction with a conversion ratio of about 22 %, yielding 0.15 g/L/day (**Figure 4.3B**).

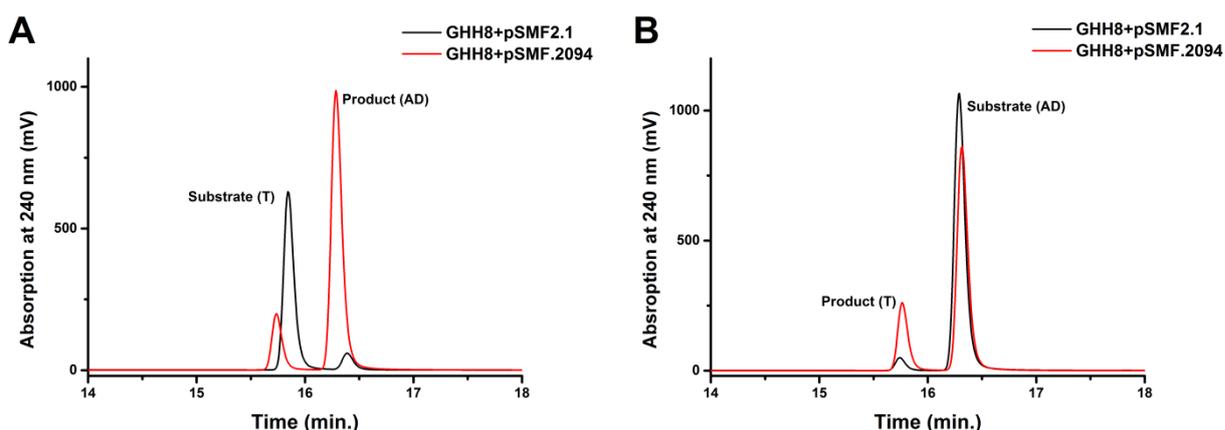


Figure 4.3. HPLC chromatograms of the *in vivo* conversion of (A) Testosterone [T] and (B) androst-4-ene-3,17-dione [AD]. *B. megaterium* GHH8 cells carrying empty vector (pSMF2.1) or overexpressing BMD_2094 (pSMF.2094) were induced with 5 mg/ml xylose at 30°C for 24 hours. The conversion was carried out using 0.5 mM substrate concentration for 5 h using resting cells (resuspended in 50 mM KPP buffer pH 7.4 + 2% glycerol) at 30°C. [GHH8: *B. megaterium* MS941 Δ UPP Δ CYP106A1 Δ 20aHSD (Gerber et al., 2016)].

Testosterone is included in the World Health Organization's list of essential medicines, which comprises the most important medications needed in a basic health system (World Health Organization, 2015). Testosterone is used as a medication for the treatment of males with androgen deficiency. This is known as hormone replacement therapy (HRT), which maintains serum testosterone levels in the normal range.

Currently, testosterone is chemically produced from androst-4-ene-3,17-dione (Ercoli and Ruggieri, 1953). Several attempts have been reported for the microbial biotransformation of sterol (such as cholesterol and β -sitosterol) to produce testosterone. Nevertheless, the presence of a final reduction step of the C17 was the main obstacle in a high testosterone yield. Although several microbial 17β -HSD enzymes have been cloned and characterized (Donova et al., 2005), just few of them were used to develop genetically engineered bacteria to improve the biotechnological production of testosterone (Fernández-Cabezón et al., 2016).

The 17β -HSD activity of BMD_2094 and its successful *in vivo* use offer new possibility for the development of biotechnological processes for the production of sex hormones. However, additional efforts are needed to examine the enzyme activity in more detail, which will enable the optimization of the production of testosterone from androst-4-ene-3,17-dione.

3.5 BMD_2094 CAN CONVERT THE SESQUITERPENE NOOTKATOL TO (+)-NOOTKATONE

Terpenes are a large group of naturally occurring hydrocarbons, which have valuable applications in the flavor and fragrance industry. The sesquiterpene (+)-nootkatone ($C_{15}H_{22}O$) has a characteristic grapefruit flavor and was first isolated from the cedar wood *Chamaecyparis nootkatensis*, but is also found in trace amounts in some citrus fruits such as grapefruit, mandarin and pummelo (Leonhardt and Berger, 2014). Therefore, for commercial applications (+)-nootkatone is synthesized from (+)-valencene, an abundant constituent of various citrus species (eg. Valencia orange) and thus, is a convenient and favorable bioresource. The conversion of (+)-valencene to (+)-nootkatone comprises two steps (**Figure 4.4**) proceeding via a regioselective hydroxylation of (+)-valencene at C-2

producing the (trans)-nootkatol, which is oxidized to (+)-nootkatone (Fraatz et al., 2009). Little is known about the enzymatic step/s involved in this biosynthesis. Therefore, many methods have been developed trying to synthesize (+)-nootkatone. A chemosynthesis method, which is based on the allylic oxidation of the precursor (+)-valencene using toxic hazardous oxidants such as chromate, manganite and *tert*-butyl hydroperoxide has been reported (Fraatz et al., 2009). More recently, whole-cell biotransformation methods (in bacteria, fungi or cell culture) employing different enzymes such as lipoxygenase, laccase or P450s enzymes have been also described (Leonhardt and Berger, 2014). While the first step can be catalyzed by P450s (Gavira et al., 2013; Girhard et al., 2009), an enzyme which catalyzes the second step is still required.

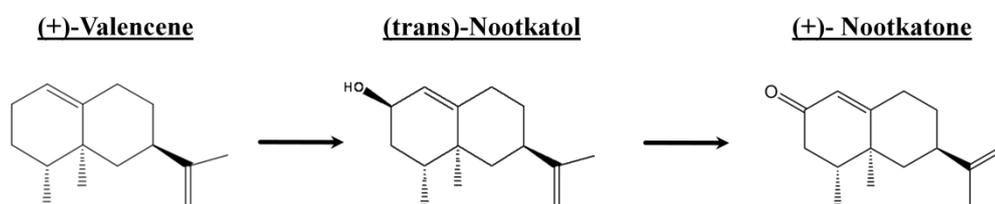


Figure 4.4. Oxidation of (+)-valencene via (trans)-nootkatol to (+)-nootkatone.

The investigation of various new potential substrates for the constructed *B. megaterium* based whole-cell systems by our group turned out the ability of *B. megaterium* cells to convert the sesquiterpene (trans)-nootkatol into the valuable ketone form (+)-nootkatone, although with low efficiency (**Figure 4.5A**). Information about the enzymatic step/s involved in the formation of (+)-nootkatone from (trans)-nootkatol is very scarce. In this study we aimed to identify an enzyme to enhance the final step in the formation of (+)-nootkatone, which is thought to be a dehydrogenation reaction. Therefore, we searched the *B. megaterium* strain DSM319 for a potential enzyme candidate/s of our cloned dehydrogenases collection shown in publication 2.3 (Gerber et al., 2016).

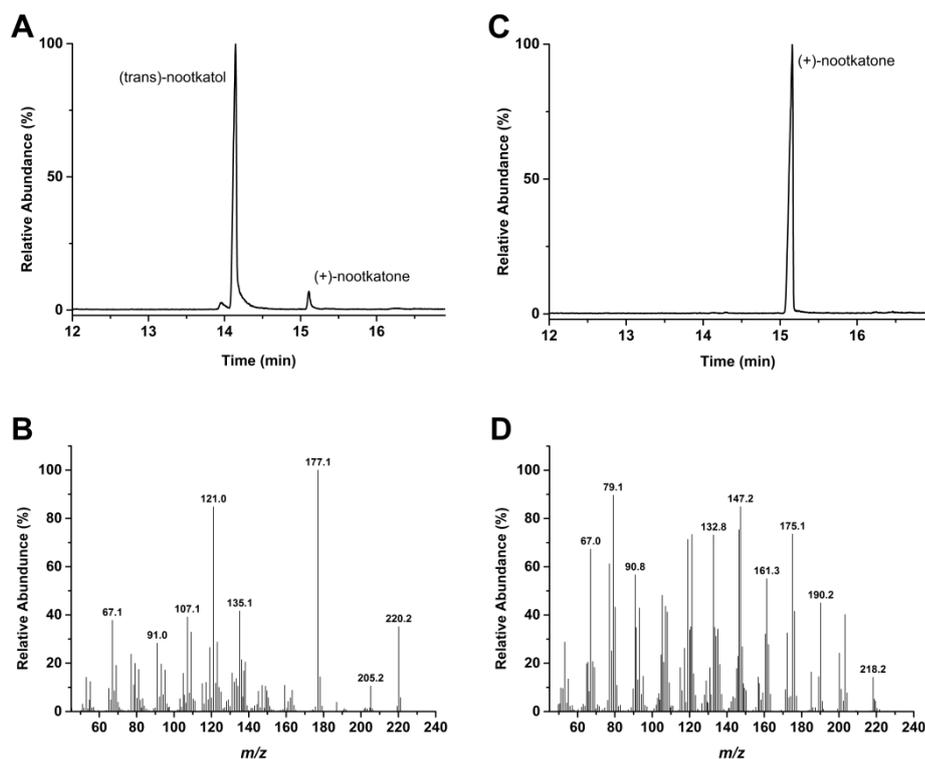


Figure 4.5. Product analysis and identification of the *in vivo* conversion of (trans)-nootkatol. Whole-cell biotransformation was performed using *B. megaterium* MS941 cells carrying the pSMF2.1 empty vector (A) or pSMF.2094 vector expressing BMD_2094 (C). (B) and (D) represent the product identification by mass spectrometry of (trans)-nootkatol and (+)-nootkatone, respectively. Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4 + 2% glycerol) and substrate concentration of 200 μ M. Samples were taken after 5 h conversion and analyzed via GC-MS.

The activity of the candidates enzymes towards (trans)-nootkatol was verified by an overexpression approach. To achieve this, we used the xylose inducible vector pSMF2.1 as a backbone for the overexpression of the genes (Gerber et al., 2016). Remarkably, none of the screened strains showed detectable activity except for the one transformed with the pSMF.2094 (expressing BMD_2094). Previously, many publications highlighted the participation of the dehydrogenases in the metabolic pathway of different terpenes namely the synthesis of zerumbone from 8-hydroxy- α -humulene from the ginger plant *Zingiber zerumbet* (Okamoto et al., 2011) and the germacrene-derived sesquiterpene lactones from the vegetable *Cichorium intybus* (de Kraker et al., 2001). As shown in **Figure 4.5B**, the *B. megaterium* strain expressing the BMD_2094 was able to convert (trans)-nootkatol (RT= 14.14 min) completely to (+)-nootkatone (RT= 15.16 min) compared to the control strain **Figure 4.5A**, where just minimal conversion was detected (~6%).

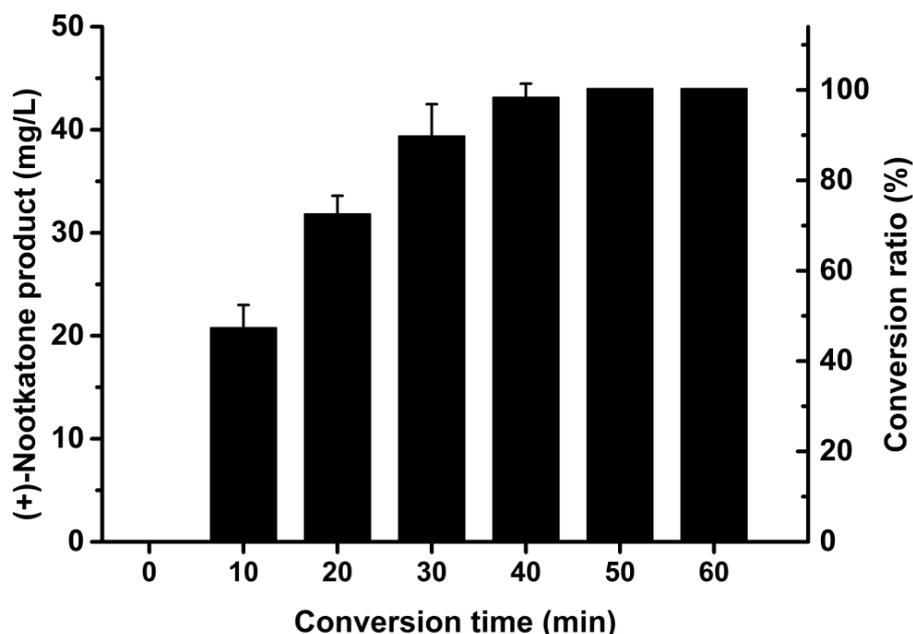


Figure 4.6. Time dependent *in vivo* conversion of 200 μ M (trans)-nootkatol by a BMD_2094 based whole-cell system. Whole-cell biotransformation was performed using *B. megaterium* MS941 cells transformed with pSMF.2094 vector (expressing BMD_2094). Conversion was carried out by using resting cells in potassium phosphate buffer (50 mM KPi pH 7.4 + 2% glycerol) and samples were taken at the indicated times and analyzed via GC-MS. The data is represented as mean \pm SD of three separate measurements.

Time dependent conversion demonstrates the efficiency of the BMD_2094 (**Figure 4.6**). A conversion ratio of about 100 % was achieved within 40 min yielding about 44 mg/L (+)-nootkatone. Most of the enzymatic based-methods for (+)-nootkatone production still have some limitations restricting their use in the industrial processes, including the substrate specificity, the need of auxiliary redox partners, the difficulties of expression of membrane enzymes in prokaryotic microorganism and the accumulation of the intermediate nootkatol as a result of unidentified enzyme, which can catalyze the second step efficiently (Fraatz et al., 2009). Most of these limitations have been overcome in many publications by expressing highly selective plant P450s in different yeast species or by the engineering of bacterial P450s to improve the formation of (trans)-nootkatol. Information about the enzymatic step leading from (trans)-nootkatol to (+)-nootkatone is very scarce. The first attempt for the use of cytochrome P450 in the production of (+)-nootkatone was described by Sowden et al (Sowden et al., 2005). P450cam from *Pseudomonas putida* and P450 BM3 from *B. megaterium* have been engineered for the oxidation of (+)-valencene. P450cam WT was not able to oxidize the (+)-valencene, therefore, a number of mutations were introduced. The P450cam mutants oxidize (+)-valencene with high

regioselectivity for C2 oxidation. The relative proportions of nootkatol and nootkatone vary from 86% (trans)-nootkatol and 4% nootkatone for the F87A/Y96F/L244A/V247L mutant, to 38% nootkatol and 47% nootkatone for F87V/Y96F/L244A. On the other hand the P450 BM3 WT showed activity towards (+)-valencene but with much lower regioselectivity (Sowden et al., 2005). CYP109B1 from *Bacillus subtilis* (Girhard et al., 2009), the CYP71AV8 from chicory (*Cichorium intybus*) (Cankar et al., 2011), and the tobacco CYP71D51v2 (Gavira et al., 2013) were also applied for the oxidation of (+)-valencene. Nevertheless, all these attempts resulted in multiple products, whereas (trans)-nootkatol was the main one. To overcome this problem, Wriessnegger et al (2014) identified and overexpressed an endogenous alcohol dehydrogenase (ADH) to enhance the final step of (+)-nootkatone formation in *Pichia pastoris*, which lead to >20- fold increase yielding 7 mg/L of the desired product. This amount was improved to 17 mg /L within 48 h by coexpressing truncated hydroxy-methylglutaryl-CoA reductase (tHmg1p) of *Saccharomyces cerevisiae* (Wriessnegger et al., 2014).

To the best of our knowledge, the BMD_2094 is the first identified dehydrogenase of bacterial origin, which can efficiently oxidize (trans)-nootkatol to (+)-nootkatone. Moreover, in view of the biotechnological application, the activity of BMD_2094 demonstrates its value in boosting the formation of (+)-nootkatone for the development of efficient whole-cell systems for the production of (+)-nootkatone. Nevertheless, further improvements are required, either at the cellular level by expressing BMD_2094 in other microorganisms, or at the methodological level by using bioreactor cultivation.

4 OUTLOOK

In conclusion, biocatalytic reactions involving incorporation of oxygen are potentially very useful in synthetic biology. Generally, reactions with molecular oxygen always involve some disadvantages associated with the formation of reactive oxygen species, which can cause damage of the reaction. Living cells have several ways to cope with ROS and thus the use of whole cell biocatalyst is very beneficial. In addition, whole cells contain enzymes that can be used for coenzyme regeneration, which is another critical issue in all biocatalytic redox reactions. Therefore, the biotransformation toolbox must be continuously expanded with new strains and novel enzymes and enzyme activities.

This work demonstrates the importance of three newly characterized enzymes and their impact in future biotechnological applications. The diflavin BmCPR enzyme was found to support the electron transfer *in vitro* and *in vivo* to class I and class II P450s (**Figure 4.1**), which make this reductase a very promising and advantageous candidate to provide a versatile redox partner for broad range of P450s and of substrate classes. This will facilitate the establishment of the whole cell systems, which utilize different classes of P450s. Significant progress has been achieved over the past decade in engineering of P450s and related proteins, and the technologies for this are quite well established nowadays. The tyrosine residue in the FMN binding motif (T-Y-G-E-G-D/E-P) of the P450 reductases had been shown previously to play an important role in the interaction with the isoalloxazine ring of the FMN (Wang et al., 1997). On the other hand, it is known that electrostatic interactions are responsible for the stabilization of the P450-reductase complex during the electron transfer from the FMN to the heme center (Bernhardt et al., 1988), where the acidic residues on the surface of the reductase interact with the basic residues on the surface of the P450. Sequence analysis of the BmCPR showed that the FMN as well as the P450 interaction motifs (**Figure 4.7**) can be further optimized in order to increase the overall activity of this reductase. Therefore, site directed mutagenesis in the FMN motif (H122Y) and in the P450 interaction motif (C189D / P196D) might be of great importance in improving the binding of the FMN co-factor and increasing the interaction with the class II P450s, respectively.

	Phosphate Moiety	FMN Ring	
BmCPR	TVSKDVTIILYGSQTGNAOGLAENTGKTLKAKGFNVT--VSSMNDFKPN-----NLKKLENLLIVVSTHGEGEPDNDALS		133
bCPR	TGRN-IIVFYGSQTGTAEFFANRLSKDAHRYGMRG--MAADPEEYDLADLSSLPEIEKA-LAIFCMATYGEGDPTDQAQD		151
ATR1	SGKTRVSIFFGTQTGTAEGFALSEEIKARYEKAARKVIDLDDYAADDQYEEKLKKETLAFFCVATYGDGEPDNDNAAR		159
yCPR	ENNKNYLVLYASQTGTAEYAKKFSKELVAKFNLN-VMCADVENYDFESLN-----DVPVIVSIFISTYGEGDFFPDGAVN		129
	FMN Ring	P450	
BmCPR	FHEFLHGRR--APKLENFRFSVLSLGDSSYEF-FCQTGKEFDVRLAELGGERLYPR--VDCDLD-FEEPANKWLKGVIDG		207
bCPR	FYDWLQET---DVDLSGVKYAVFALGNKTYEH-FNAMGKYVDKRLEQLGAQRIFDLGLGDDGN-LEEDFITWREQFWPA		226
ATR1	FYKWFTEENERDIKLQQLAYGVFALGNRQYEH-FNKIGIVLDEELCKKGAKRLIEVGLGDDQS-IEDDFNAWKESLWSE		237
yCPR	FEDFICNAE--AGALSNLRYNMFGLGNSTYEF-FNGAAKKAEKHLSAAGAIRLGKLGAEADDGAGTTDEEDYMAWKDSILEV		206

Figure 4.7. Part of the multiple sequence amino acids alignment of BmCPR and other cytochrome P450 reductases from bovine (bCPR), *Arabidopsis thaliana* (ATR1), and *S. cerevisiae* (yCPR). The residue H122 in the FMN binding motif is shown in green and the C189/C196 in the P450 interaction domain are shown in yellow.

CYP107DY1 was identified as a mevastatin hydroxylase. Its activity could be reconstituted *in vitro* and *in vivo*. The *in vivo* activity of CYP107DY1 demonstrates its value for the development of efficient whole-cell systems for the production of pravastatin. Nevertheless, further improvements are required, either at the cellular level by expressing CYP107DY1 in other microorganisms, or at the molecular level by improving the activity of this P450. The great potential offered by new protein engineering methods able to alter the properties of the enzyme, in addition to the availability of established high-throughput methods for substrate screening, opens a wide range of opportunities for biotechnological exploitation. Rational design of mutants by site-directed mutagenesis has been widely applied to different P450s such as P450cam (Harford-Cross et al., 2000; Loida and Sligar, 1993; Wong et al., 1997), P450 BM3 (Graham-Lorence et al., 1997; Oliver et al., 1997; Whitehouse et al., 2012; Yeom et al., 1995), CYP109E1 (Jóźwik et al., 2016) and others (Bernhardt and Urlacher, 2014; Janocha et al., 2015). An additional potential is offered by the application of random mutagenesis methods. These have already resulted in a range of successful examples of enzyme optimization (improvement in enzymatic activity, substrate specificity, protein folding and expression levels) with emerging applications to pharmaceuticals and vaccines (Arnold, 2009; Tsotsou et al., 2002). Therefore, the availability of *in vitro* reconstituted as well as functional *in vivo* systems motivates the identification of new compounds that are potential substrates of CYP107DY1. The screening can be carried out on substrates, which are analogue to mevastatin such as lovastatin and simvastatin or by systematic high throughput substrates screening.

This work presents also a proof of concept of the possibility of employing the 17 β -HSD activity of BMD_2094 for the production of pharmaceutical steroids. The reaction equilibrium of the BMD_2094

was found to be shifted towards oxidation (**Figure 4.8**), accordingly; the reaction and the system should be further optimized to shift it towards reduction in order to increase testosterone production efficiency. This could be achieved by changing the culture and conversion conditions. It was previously demonstrated that glucose utilization can promote a 17β -reduction and inhibit the reverse oxidation of the product, therefore, the control of the bacterial metabolic state will be very useful. Factors such as different carbon sources, pH and mode of substrate addition will be beneficial in increasing the NAD(P)H/NAD(P) ratio in the cell, and thus, will help shifting the reaction towards reduction. In addition, characterization of the BMD_2094 enzyme in more detail will be of great importance for future rational design of mutant with alterations in substrate specificity and coenzyme requirements.

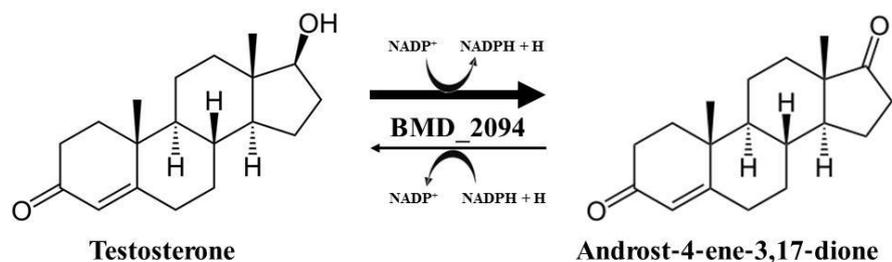


Figure 4.8. Oxidation of testosterone (forward reaction) and reduction of androst-4-ene-3,17-dione (reverse reaction) by BMD_2094.

Moreover, BMD_2094 was found to be very efficient in the conversion of the sesquiterpene (trans)-nootkatol to (+)-nootkatone. However, it will be worthful to coexpress a P450, which can catalyze the first step by converting valencene to (trans)-nootkatol (**Figure 4.4**). A suitable P450s are available such as the CYP71D51v2 from *Nicotiana tabacum* (Gavira et al., 2013) and the BM3 variant (BM3-AI) (Schulz et al., 2015).

The combination of different hosts (bacteria, yeast, and filamentous fungi), expression systems and plasmids theoretically allows many different possibilities to express and produce an enzyme. This variety of possibilities in gene cloning and enzyme production, potentially allows solutions to be developed, for example, for productivity increase and relief of inhibition phenomenon of substrates or products.

The future of genome mining for novel enzymes discovery seems bright and it is likely that the future will see an expansion of interest in this exciting field. Genome mining combined with biochemical and microbiological approaches can provide the stimuli needed for the construction of novel biocatalyst systems, which will drive the field of synthetic biology forward.

5 APPENDIX

5.1 SUPPLEMENTAL METHODS

5.1.1 CLONING OF THE GENE ENCODING BMD_2094

Genomic DNA was prepared using a genomic DNA isolation kit (nexttec). The DNA fragment encoding the full length BMD_2094 was PCR amplified from the genomic DNA of *B. megaterium* strain DSM319 using the forward primer (CGCttaattaa AAATCAAGGAGGTGAATGTACAATGTCACAGCATTATGCGC) and reverse primer (TTATCAactagtTTATTTAATTTCTTGTGTTCTGCGGACAACGG) [restriction sites are in small letters and the ribosomal binding site is underlined]. The resulting fragment was cloned in the xylose-inducible shuttle vector pSMF2.1 (Bleif et al., 2012) with the *PacI/SpeI* restriction sites, resulting in plasmid pSMF-2094. The plasmid was verified by sequencing.

5.1.2 CULTIVATION CONDITIONS OF THE *B. MEGATERIUM* AND IN VIVO SUBSTRATE CONVERSION

B. megaterium protoplasts preparation as well as the PEG-mediated transformation was carried out as described elsewhere (Biedendieck et al., 2011).

Transformed *B. megaterium* cells were grown overnight in 50 ml Luria-Bertani (LB) broth medium supplemented with 10 µg/ml tetracycline at 37 °C and shaking at 140 rpm. For the expression of proteins, 50 ml Terrific broth (TB) medium containing 10 µg/ml tetracycline were inoculated (1:100) with the transformed cells and cultivated in baffled erlenmeyer flask at 37 °C with rotary shaking at 140 rpm. Protein expression was induced at OD₆₀₀ of 0.4 – 0.6 with 5 mg/ml xylose. The temperature was then reduced to 30 °C. After incubation for 24 h, cells were harvested by centrifugation (4000 g) for 10 min at 4 °C and washed once with 1 volume with conversion buffer (50 mM potassium phosphate buffer (pH 7.4) supplemented with 2% glycerol). After a second centrifugation, the cell pellets were resuspended in conversion buffer to an end cell suspension concentration of 40 g wet cell

weight (wcw) / L buffer. The substrate was added and the culture was incubated for the indicated time at 30 °C and 140 rpm.

5.1.3 METABOLITES EXTRACTION AND GC-MS ANALYSIS

Culture sample/s (1 ml) were taken at different time points and extracted twice with equal volumes of ethylacetate followed by evaporating of the organic phase using a rotary evaporator. After that, the residues were dissolved in ethylacetate and analyzed with Gas chromatography mass spectrometry (GC-MS) system consisting of a DSQII quadrupole, a Focus GC column oven (Thermo scientific, Waltham, USA), and a DB-5 column (Agilent) with a length of 25 m, 0.32 mm ID, and 0.52 µm film thickness. The metabolites were analyzed in a m/z range of 20–350. The starting oven temperature was 50 °C for 1 min and then the temperature was ramped to 310 °C by 10 °C/min and held for 3 min with a flow rate of 1 ml/min. The EI-mass spectra were compared with the NIST mass spectral library (version 2.0).

5.2 SEQUENCES

5.2.1 BmCPR (BMD_3122)

1	TTGCAACTTA	AGGTTAGTAAA	CAGCCCTTTT	AATCAAGAAC	AAGCAGATTT	50
51	GCTTAATCGC	CTTCTGCCGA	CATTAACAGA	AGCACAAAAA	ATGTGGTTGA	100
101	GCGGTTATTT	AACAGCAGCT	CAATCTACGT	CTGCCGAAGG	AACGCCAGAC	150
151	GTTTCTACAG	CAGCGCCTGC	TCAAACGAAA	CAGACAGTTT	CAAAAAGACGT	200
201	AACAATCTTT	TATGGATCAC	AAACAGGAAA	TGCTCAAGGA	CTTGCTGAAA	250
251	ATACAGGCAA	AACGCTTGAA	GCAAAAAGGCT	TAAATGTAAC	TGTATCTTCT	300
301	ATGAATGATT	TCAAACCAAA	TAATTTAAAG	AAACTTGAAA	ATTTATTAAT	350
351	TGTCGTAAGT	ACACATGGAG	AAGGAGAGCC	GCCTGATAAT	GCGCTATCTT	400
401	TCCATGAATT	TCTTCACGGC	CGTCGAGCGC	CAAAACTTGA	GAACTTCCGT	450
451	TTCTCTGTTT	TATCGCTTGG	AGACAGCTCA	TACGAATTTT	TCTGTCAAAC	500
501	AGGAAAAGAA	TTTGATGTGC	GCTTAGCAGA	ACTTGGCGGT	GAAAGACTGT	550
551	ATCCGCGCGT	TGACTGTGAT	TTAGATTTTG	AAGAGCCCGC	AAATAAATGG	600
601	CTTAAAGGTG	TTATTGACGG	ATTAAGCGAA	GCGAAAGGAC	ACAGCGCTTC	650
651	GGCAGCTGTT	CCAGCGGAAG	CTCCTGCAGG	AACTTCACCA	TATTCAAGAA	700
701	CAAACCCTTT	TAAAGCAGAA	GTGCTTGAGA	ACTTAAACTT	AAACGGCCGC	750
751	GGATCAAATA	AAGAAACGCG	ACACTTAGAA	CTATCTCTAG	AAGGTTTCAGG	800
801	TTTGACGTAT	GAACCGAGG	ACAGTTTAGG	TATTTATCCT	GAAAATGACC	850
851	CGGAGCTTGT	TGATCTTCTT	CTTAACGAAT	TCAAGTGGGA	TGCAAGTGAA	900
901	AGTGTAACGG	TTAATAAAGA	AGGAGAAACG	CGTCCTCTTA	GAGAAGCGCT	950
951	AATCTCTAAT	TTTGAAATTA	CCGTTTTAAC	AAAGCCGCTT	TTAAAGCAAG	1000
1001	CAGCTGAGCT	TACTGGAAAC	GATAAGTTAA	AAGCGCTTGT	AGAAAATCGC	1050
1051	GAGGAATTA	AAGCATATAC	ACAAGGCCGT	GATGTAATTG	ACTTAGTTTCG	1100
1101	TGACTTCGGT	CCATGGAACG	TATCGGCACA	AGAGTTTGTA	GCCATTTTAC	1150
1151	GCAAAATGCC	AGCGCGCCTT	TACTCGATTG	CAAGCAGCTT	ATCAGCAAAC	1200
1201	CCTGATGAAG	TTCATCTAAC	AATCGGAGCG	GTACGCTACG	AAGCGCATGG	1250
1251	ACGCGAGCGT	AAGGGTGTTT	GTTTCAGTCT	TTGTTTCAGAA	CGTTTGCAGC	1300
1301	CAGGCGATAC	GATTCCTGTA	TACCTTCAAA	GCAATAAAAA	CTTTAAGCTT	1350
1351	CCTCAAGATC	AGGAAACGCC	AATTATTATG	GTAGGACCTG	GTACAGGTGT	1400
1401	GGCTCCGTTT	CGCTCATTTA	TGCAAGAGCG	TGAAGAAACA	GGGGCAAAAG	1450
1451	GAAAGTCATG	GATGTTCTTT	GGAGATCAGC	ACTTCGTAAC	AGACTTCCTT	1500
1501	TACCAAACAG	AATGGCAAAA	GTGGTTAAAA	GATGGCGTAC	TGACAAAAAT	1550
1551	GGACGTGGCG	TTTTCACGCG	ATACAGAAGA	AAAAGTATAC	GTACAAAACC	1600
1601	GTATGCTCGA	ACATAGTAAA	GAATTATTCC	AATGGTTAGA	AGAAGGCGCA	1650

1651	TCTTTTTATG	TGTGCGGAGA	TAAAACAAAT	ATGGCACGCG	ACGTGCACAA	1700
1701	CACGCTAGTT	GAAATTATCG	AAACAGAAGG	CAAGATGAGC	CGTGAACAGG	1750
1751	CGGAAGGTTA	CCTTGCTGAA	ATGAAGAAAC	AAAAACGTTA	TCAGCGTGAT	1800
1801	GTATACTGA					1809

5.2.2 CYP107DY1 (BMQ_pBM50008)

1	ATGAAAAAGG	TTACAGTTGA	TGATTTTAGC	TCTCCAGAAA	ATATGCACGA	50
51	TGTCATCGGA	TTTTATAAAA	AACTCACTGA	ACATCAAGAA	CCTCTTATTC	100
101	GTTTGGATGA	TTATTACGGG	TTGGGACCGG	CATGGGTCGC	ATTACGTCA	150
151	GACGATGTTG	TTACGATACT	AAAGAACCCC	CGTTTTCTCA	AAGATGTACG	200
201	GAAGTTCACA	CCATTGCAAG	ATAAAAAGGA	TTCTATAGAT	GATAGCACAT	250
251	CTGCGAGCAA	ACTGTTTGAA	TGGATGATGA	ATATGCCGAA	TATGCTTACG	300
301	GTCGATCCAC	CCGATCACAC	TCGTTGCGC	AGGTGCGCCT	CTAAAGCCTT	350
351	TACGCCACGT	ATGATCGAGA	ATCTTCGACC	TCGTATACAG	CAGATTACCA	400
401	ATGAGCTATT	GGATTGAGTA	GAAGGAAAAA	GGAATATGGA	TCTTGTTGCG	450
451	GATTTTTCTT	TTCTCTGCCC	CATTATTGTC	ATTTCAAGAG	TGCTAGGGAT	500
501	TCCACCTTTA	GATCAGAAAC	GATTTGCGCA	CTGGACAGAT	AAACTCATCA	550
551	AAGCAGCTAT	GGATCCTAGC	CAAGGGGCTG	TAGTTATGGA	AACACTCAAG	600
601	GAGTTTATTG	ATTACATCAA	AAAAATGCTG	GTCGAAAAGC	GCAACCATCC	650
651	AGACGATGAT	GTGATGAGTG	CTTTGTTGCA	AGCACATGAG	CAAGAAGATA	700
701	AGTTGAGCGA	GAACGAGCTT	CTTTCCACGA	TTTGGCTACT	CATTACAGCC	750
751	GGACATGAGA	CGACGGCCCA	TCTAATCAGC	AACGGCGTAC	TGGCGCTATT	800
801	GAAGCATCCC	GAACAATGTC	GCCTGCTTCG	GGATAATCCT	TCTTTACTCC	850
851	CCTCTGCCGT	TGAAGAGCTG	CTACGCTATG	CCGGACCGGT	CATGATTGGT	900
901	GGCCGTTTTG	CGGGTGAAGA	TATCATCATG	CATGGAAAAA	TGATTCCCAA	950
951	AGGTGAAATG	GTGCTGTTCT	CGCTGGTTGC	CGCCAATATT	GATTCACAGA	1000
1001	AATTCTCTTA	TCCTGAGGGA	TTGGATATTA	CACGCGAGGA	GAATGAGCAT	1050
1051	CTCACTTTCC	GAAAAGGTAT	CCATCATTGT	TTGGGAGCGC	CTTTGGCGCG	1100
1101	CATGGAAGCA	CATATCGCTT	TCGGCACATT	GCTTCAACGG	TTTCTTGATT	1150
1151	TACGATTGGC	AATCGAATCG	GAGCAACTGG	TTTATAACAA	CAGCACATTG	1200
1201	CGTTCTCTTA	AAAGCTTGCC	AGTTATTTTC	TAA		1233

5.2.3 BMD_2094

1	ATGTCACAGC	ATTATGCGCT	CATCACCGGT	GCGTCGGGAG	GAATAGGTAA	50
51	AGAGTTAGCT	TATCAATTCG	CAAAGGATGG	GCATCCTGTA	ATTTTAGTTG	100
101	CCAGAAGCGC	TGATAAACTA	GCGGCTATTG	GAGAGAATTT	AAAATCTACT	150
151	TATAATATTG	AAGTCATAAC	GGTTGTTAAG	GATTTATCAA	GAGAAGAAGA	200
201	GATACAGTCA	TTATATGAAG	AACTTAAAAA	TAAAAAATG	CACGTTGATT	250
251	ATTTAGTGAA	CAATGCTGGT	TTTGACTAT	ACGGCAAATT	TATTGAAACA	300
301	GCTTTAGATG	AAGAGTTAAA	CATGATTGAT	TTAAATATTC	GAGCATTAAC	350
351	TCATTTAACA	AAATTATTTT	TACCGGATAT	GCTCAAAAGA	AATCGTGGGA	400
401	AAATTTTAAA	CATTGCATCT	GTTGCTGCAT	TTATGCCAGG	TCCACTTATG	450
451	ACGGTATATT	ACGCAACGAA	AGCGTACGTA	TTATCTTTCA	CAGAGGCCCT	500
501	TGAAAATGAA	TTAAGAGGTA	CCAATGTAAC	GGTAAGTGCT	CTATGCCCCG	550
551	GTCCAACATA	AACGGGTTTT	AGCGACCGAG	CCCAGCTAAG	CAACTCCAAG	600
601	CTTTTTTAAA	GCGGTGCAAT	GGATGTAGAA	ACAGTAACGA	AAGTTGGCTA	650
651	TAAAAAATTT	ATGAAAGGGC	AGACGGTTAT	CGTACCAGGA	GTACAGTTTC	700
701	GTCTTGCTAC	GTTTATCCCG	CGCTTCGTGC	CAAGAAAGTG	GCTTACATCC	750
751	GTTGTCCGCA	GAACACAAGA	AATTAATAAA			780

5.3 ABBREVIATIONS

AdR	Adrenodoxin reductase
Adx	Adrenodoxin
Arh1	Adrenodoxin reductase homologue 1
<i>B. megaterium</i>	<i>Bacillus megaterium</i>
BmCPR	<i>Bacillus megaterium</i> cytochrome P450 reductase
CD	Circular dichroism
CO	Carbon monoxide
CPR/POR	Cytochrome P450 reductase / Cytochrome P450 oxidoreductase
CYP/P450	Cytochrome P450
CYP17A1	Cytochrome P450 17-hydroxylase/17,20 lyase
CYP21A2	Cytochrome P450 21-hydroxylase
DMSO	dimethyl sulfoxide
DOC	11-deoxycorticosterone
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetate
etp1 ^{fd}	Electron transfer protein 1
FAD	Flavine adenine dinucleotide
Fdx2	Ferredoxin 2 from <i>B. megaterium</i> DSM319
FMN	Flavine mononucleotide
GC	Gas chromatography
h	Hour
HPLC	High performance liquid chromatography
HSD	Hydroxy steroid dehydrogenase
KPP	Potassium phosphate buffer
L	Liter
M	Molar
mg	Milligram
µm	Micrometer
min	Minute
MS	Mass spectrometry
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
ORF	Open reading frame
SDR	Short-chain dehydrogenase/reductase
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
UV/Vis	Ultraviolet-visible spectroscope
wcw	wet cell weight

6 REFERENCES

- Abdulmughni, A., Jóźwik, I.K., Putkaradze, N., Brill, E., Zapp, J., Thunnissen, A.-M.W.H., Hannemann, F., Bernhardt, R., 2016. Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D3 hydroxylase. *J. Biotechnol.* 243, 38–47.
- Appleby, A.C., 1967. A soluble haemoprotein P450 from nitrogen-fixing *Rhizobium* bacteroids. *Biochim. Biophys. Acta* 147, 399–402.
- Arnold, F.H., 2009. How proteins adapt: lessons from directed evolution. *Cold Spring Harb. Symp. Quant. Biol.* 74, 41–46.
- Axelrod, J., 1955. The enzymatic demethylation of ephedrine. *J. Pharmacol. Exp. Ther.* 114, 430–438.
- Bak, S., Kahn, R.A., Olsen, C.E., Halkier, B.A., 1997. Cloning and expression in *Escherichia coli* of the obtusifoliol 14 α -demethylase of *Sorghum bicolor* (L.) Moench, a cytochrome P450 orthologous to the sterol 14 α -demethylases (CYP51) from fungi and mammals. *Plant J.* 11, 191–201.
- Berg, A., Gustafsson, J.A., Ingelman-Sundberg, M., 1976. Characterization of a cytochrome P450-dependent steroid hydroxylase system present in *Bacillus megaterium*. *J. Biol. Chem.* 251, 2831–2838.
- Berg, A., Ingelman-Sundberg, M., Gustafsson, J.A., 1979. Isolation and characterization of cytochrome P450meg. *Acta Biol. Med. Ger.* 38, 333–344.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145.
- Bernhardt, R., Kraft, R., Otto, A., Ruckpaul, K., 1988. Electrostatic interactions between cytochrome P450 LM2 and NADPH-cytochrome P450 reductase. *Biomed. Biochim. Acta* 47, 581–592.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203.
- Bernhardt, R., Waterman, M.R., 2007. Cytochrome P450 and steroid hormone biosynthesis, in: Sigel, A., Sigel, H., Sigel, R.K.O. (Eds.), *The Ubiquitous roles of cytochrome P450 proteins*. John Wiley & Sons, Ltd, pp. 361–396.
- Biedendieck, R., Borgmeier, C., Bunk, B., Stammen, S., Scherling, C., Meinhardt, F., Wittmann, C., Jahn, D., 2011. Systems biology of recombinant protein production using *Bacillus megaterium*, in: *Methods in Enzymology*. Elsevier, pp. 165–195.
- Bleif, S., Hannemann, F., Zapp, J., Hartmann, D., Jauch, J., Bernhardt, R., 2012. A new *Bacillus megaterium* whole-cell catalyst for the hydroxylation of the pentacyclic triterpene 11-keto- β -boswellic acid (KBA) based on a recombinant cytochrome P450 system. *Appl. Microbiol. Biotechnol.* 93, 1135–1146.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2013. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1703–1717.
- Brixius-Anderko, S., Schiffer, L., Hannemann, F., Janocha, B., Bernhardt, R., 2015. A CYP21A2 based whole-cell system in *Escherichia coli* for the biotechnological production of premedrol. *Microb. Cell Factories* 14:135. doi: 10.1186/s12934-015-0333-2
- Brodie, B.B., Axelrod, J., Cooper, J.R., Gaudette, L., Du, B.N.L., Mitoma, C., Udenfriend, S., 1955. Detoxication of drugs and other foreign compounds by liver microsomes. *Science* 121, 603–604.

- Brown, A.G., Smale, T.C., King, T.J., Hasenkamp, R., Thompson, R.H., 1976. Crystal and molecular structure of compactin, a new antifungal metabolite from *Penicillium brevicompactum*. *J. Chem. Soc. Perkin 1* 11, 1165–1170.
- Bunk, B., Schulz, A., Stammen, S., Münch, R., Warren, M.J., Rohde, M., Jahn, D., Biedendieck, R., 2010. A short story about a big magic bug. *Bioeng. Bugs* 1, 85–91.
- Cankar, K., van Houwelingen, A., Bosch, D., Sonke, T., Bouwmeester, H., Beekwilder, J., 2011. A chicory cytochrome P450 monooxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS Lett.* 585, 178–182.
- Chen, W., Lee, M.-K., Jefcoate, C., Kim, S.-C., Chen, F., Yu, J.-H., 2014. Fungal cytochrome P450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol. Evol.* 6, 1620–1634.
- Christner, J.A., Münck, E., Janick, P.A., Siegel, L.M., 1981. Mössbauer spectroscopic studies of *Escherichia coli* sulfite reductase. Evidence for coupling between the siroheme and Fe₄S₄ cluster prosthetic groups. *J. Biol. Chem.* 256, 2098–2101.
- Crešnar, B., Petrič, S., 2011. Cytochrome P450 enzymes in the fungal kingdom. *Biochim. Biophys. Acta* 1814, 29–0
- Danielson, P.B., 2003. The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Curr. Drug Metab.* 3, 561–97.
- Davis, M., Williams, R., Chakraborty, J., English, J., Marks, V., Ideo, G., Tempini, S., 1978. Prednisone or prednisolone for the treatment of chronic active hepatitis? A comparison of plasma availability. *Br. J. Clin. Pharmacol.* 5, 501–505.
- De Bary, A., 1884. *Vergleichende Morphologie und Biologie der Pilze, Mycetozoen, und Bacterien.* Wilhelm Engelmann, Leipzig, Germany.
- de Kraker, J.-W., Franssen, M.C.R., Dalm, M.C.F., de Groot, A., Bouwmeester, H.J., 2001. Biosynthesis of Germacrene A carboxylic acid in chicory roots. demonstration of a cytochrome P450 (+)-Germacrene A hydroxylase and NADP⁺-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiol.* 125, 1930–1940.
- Denisov, I.G., Makris, T.M., Sligar, S.G., Schlichting, I., 2005. Structure and chemistry of cytochrome P450. *Chem. Rev.* 105, 2253–2278.
- Donova, M.V., Egorova, O.V., Nikolayeva, V.M., 2005. Steroid 17 β -reduction by microorganisms. *Process Biochem.* 40, 2253–2262.
- Durairaj, P., Hur, J.-S., Yun, H., 2016. Versatile biocatalysis of fungal cytochrome P450 monooxygenases. *Microb. Cell Factories* 15(1):125. doi: 10.1186/s12934-016-0523-6.
- Endo, A., Kuroda, M., Tsujita, Y., 1976. ML-236A, ML-236B, and ML-236C, new inhibitors of cholesterologenesis produced by *Penicillium citrinium*. *J. Antibiot. (Tokyo)* 29, 1346–1348.
- Eppinger, M., Bunk, B., Johns, M.A., Edirisinghe, J.N., Kutumbaka, K.K., Koenig, S.S.K., Creasy, H.H., Rosovitz, M.J., Riley, D.R., Daugherty, S., Martin, M., Elbourne, L.D.H., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B., Ul-Qamar, R., Seibel, J., Bremer, E., Jahn, D., Ravel, J., Vary, P.S., 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J. Bacteriol.* 193, 4199–4213.
- Ercoli, A., Ruggieri, P. de, 1953. An improved method of preparing testosterone, dihydrotestosterone and some of their esters. *J. Am. Chem. Soc.* 75, 650–653.

- Estabrook, R.W., Cooper, D.Y., Rosenthal, O., 1963. The light reversible carbon monoxide inhibition of the steroid c21-hydroxylase system of the adrenal cortex. *Biochem. Z.* 338, 741–755.
- Fernández-Cabezón, L., Galán, B., García, J.L., 2016. Engineering *Mycobacterium smegmatis* for testosterone production. *Microb. Biotechnol.* 10, 151–161.
- Feyereisen, R., 2011. Arthropod CYPomes illustrate the tempo and mode in P450 evolution. *Biochim. Biophys. Acta* 1814, 19–28.
- Fiel, S.B., Vincken, W., 2006. Systemic corticosteroid therapy for acute asthma exacerbations. *J. Asthma* 43, 321–331.
- Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285.
- Fraatz, M.A., Berger, R.G., Zorn, H., 2009. Nootkatone—a biotechnological challenge. *Appl. Microbiol. Biotechnol.* 83, 35–41.
- Frazzetto, G., 2003. White biotechnology. *EMBO Rep.* 4, 835–837.
- Garfinkel, D., 1958. Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.* 77, 493–509.
- Gavira, C., Höfer, R., Lesot, A., Lambert, F., Zucca, J., Werck-Reichhart, D., 2013. Challenges and pitfalls of P450-dependent (+)-valencene bioconversion by *Saccharomyces cerevisiae*. *Metab. Eng.* 18, 25–35.
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Factories* 14, 107. doi: 10.1186/s12934-015-0300-y.
- Gerber, A., Milhim, M., Hartz, P., Zapp, J., Bernhardt, R., 2016. Genetic engineering of *Bacillus megaterium* for high-yield production of the major teleost progestogens 17 α ,20 β -di- and 17 α ,20 β ,21 α -trihydroxy-4-pregnen-3-one. *Metab. Eng.* 36, 19–27.
- Girhard, M., Klaus, T., Khatri, Y., Bernhardt, R., Urlacher, V.B., 2010. Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* 87, 595–607.
- Girhard, M., Machida, K., Itoh, M., Schmid, R.D., Arisawa, A., Urlacher, V.B., 2009. Regioselective biooxidation of (+)-valencene by recombinant *E. coli* expressing CYP109B1 from *Bacillus subtilis* in a two-liquid-phase system. *Microb. Cell Factories* 8, 36. doi: 10.1186/1475-2859-8-36.
- Girvan, H.M., Munro, A.W., 2016. Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Curr. Opin. Chem. Biol.* 31, 136–145.
- Gotoh, O., 1992. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* 267, 83–90.
- Graham-Lorence, S., Truan, G., Peterson, J.A., Falck, J.R., Wei, S., Helvig, C., Capdevila, J.H., 1997. An active Site substitution, F87V, converts cytochrome P450 BM3 into a regio- and stereoselective (14S,15R)-Arachidonic Acid Epoxygenase. *J. Biol. Chem.* 272, 1127–1135.
- Guengerich, F.P., Tang, Z., Salamanca-Pinzón, S.G., Cheng, Q., 2010. Characterizing proteins of unknown function: orphan cytochrome P450 enzymes as a paradigm. *Mol. Interv.* 10, 153–163.
- Gunsalus, I.C., Wagner, G.C., 1978. [17] Bacterial P450cam methylene monooxygenase components: cytochrome m, putidaredoxin, and putidaredoxin reductase, in: Packer, S.F. and L. (Ed.), *Methods in Enzymology, Biomembranes - Part C: Biological Oxidations*. Academic Press, pp. 166–188.

- Hakki, T., Zearo, S., Drăgan, C.-A., Bureik, M., Bernhardt, R., 2008. Coexpression of redox partners increases the hydrocortisone (cortisol) production efficiency in CYP11B1 expressing fission yeast *Schizosaccharomyces pombe*. *J. Biotechnol.* 133, 351–359.
- Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., 2007. Cytochrome P450 systems-biological variations of electron transport chains. *Biochim. Biophys. Acta BBA - Gen. Subj.*, P450 1770, 330–344.
- Harford-Cross, C.F., Carmichael, A.B., Allan, F.K., England, P.A., Rouch, D.A., Wong, L.-L., 2000. Protein engineering of cytochrome P450cam (CYP101) for the oxidation of polycyclic aromatic hydrocarbons. *Protein Eng.* 13, 121–128.
- Hebeda, R.E., Styrlund, C.R., Teague, W.M., 1988. Benefits of *Bacillus megaterium* amylase in dextrose production. *Starch* 33–36.
- Helliwell, C.A., Chandler, P.M., Poole, A., Dennis, E.S., Peacock, W.J., 2001. The CYP88A cytochrome P450, entkaurenic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2065–2070.
- Helliwell, C.A., Sheldon, C.C., Olive, M.R., Walker, A.R., Zeevaart, J.A., Peacock, W.J., Dennis, E.S., 1998. Cloning of the Arabidopsis ent-kaurene oxidase gene GA3. *Proc. Natl. Acad. Sci. U. S. A.* 95, 9019–9024.
- Horecker, B.L., 1950. Triphosphopyridine nucleotide-cytochrome c reductase in liver. *J. Biol. Chem.* 183, 593–605.
- Humphreys, J.M., Hemm, M.R., Chapple, C., 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10045–10050.
- Inouye, M., Takada, Y., Muto, N., Beppu, T., Horinouchi, S., 1994. Characterization and expression of a P450-like mycinamicin biosynthesis gene using a novel *Micromonospora-Escherichia coli* shuttle cosmid vector. *Mol. Gen. Genet.* MGG 245, 456–464.
- Janocha, S., Schmitz, D., Bernhardt, R., 2015. Terpene Hydroxylation with microbial cytochrome P450 monooxygenases, in: Schrader, J., Bohlmann, J. (Eds.), *Biotechnology of isoprenoids*. Springer International Publishing, Cham, pp. 215–250.
- Jóźwik, I.K., Kiss, F.M., Gricman, L., Abdulmughni, A., Brill, E., Zapp, J., Pleiss, J., Bernhardt, R., Thunnissen, A.-M.W.H., 2016. Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J.* 283, 4128–4148.
- Julsing, M.K., Cornelissen, S., Bühler, B., Schmid, A., 2008. Heme-iron oxygenases: powerful industrial biocatalysts? *Curr. Opin. Chem. Biol., Biocatalysis and biotransformation/Bioinorganic chemistry* 12, 177–186.
- Kallberg, Y., Oppermann, U., Jörnvall, H., Persson, B., 2002. Short-chain dehydrogenases/reductases (SDRs). *Eur. J. Biochem.* 269, 4409–4417.
- Kallberg, Y., Oppermann, U., Persson, B., 2010. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models: SDR classification using HMM. *FEBS J.* 277, 2375–2386.
- Kelly, S.L., Lamb, D.C., Jackson, C.J., Warrilow, A.G.S., Kelly, D.E., 2003. The biodiversity of microbial cytochromes P450, in: *Physiology, B.-A. in M.* (Ed.), . Academic Press, 131–186.
- Kern, F., Khatri, Y., Litzenburger, M., Bernhardt, R., 2016. CYP267A1 and CYP267B1 from *Sorangium cellulosum* So ce56 are Highly Versatile Drug Metabolizers. *Drug Metab. Dispos.* 44, 495–504.
- Kim, J., DellaPenna, D., 2006. Defining the primary route for lutein synthesis in plants: the role of Arabidopsis carotenoid β -ring hydroxylase CYP97A3. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3474–3479.

- Kiss, F.M., Schmitz, D., Zapp, J., Dier, T.K.F., Volmer, D.A., Bernhardt, R., 2015. Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* – identification of a novel 11-oxidase activity. *Appl. Microbiol. Biotechnol.* 99, 8495-514.
- Klingenberg, M., 1958. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* 75, 376–386.
- Lamb, D.C., Waterman, M.R., 2013. Unusual properties of the cytochrome P450 superfamily. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120434. doi: 10.1098/rstb.2012.0434.
- Lambrou, G.I., Vlahopoulos, S., Papanthanasidou, C., Papanikolaou, M., Karpusas, M., Zoumakis, E., Tzortzidou-Stathopoulou, F., 2009. Prednisolone exerts late mitogenic and biphasic effects on resistant acute lymphoblastic leukemia cells: relation to early gene expression. *Leuk. Res.* 33, 1684–1695.
- Lamon-Fava, S., 2013. Statins and lipid metabolism: an update. *Curr. Opin. Lipidol.* 24, 221–226.
- Leak, D.J., Sheldon, R.A., Woodley, J.M., Adlercreutz, P., 2009. Biocatalysts for selective introduction of oxygen. *Biocatal. Biotransformation* 27, 1–26.
- Leonhardt, R.-H., Berger, R.G., 2014. Nootkatone, in: Schrader, J., Bohlmann, J. (Eds.), *Biotechnology of isoprenoids*. Springer International Publishing, Cham, pp. 391–404.
- Lepesheva, G.I., Podust, L.M., Bellamine, A., Waterman, M.R., 2001. Folding requirements are different between sterol 14 α -Demethylase (CYP51) from *Mycobacterium tuberculosis* and human or fungal orthologs. *J. Biol. Chem.* 276, 28413–28420.
- Lewis, D.F.V., Wiseman, A., 2005. A selective review of bacterial forms of cytochrome P450 enzymes. *Enzyme Microb. Technol.* 36, 377–384.
- Liu, Y., Wang, Y., Chen, X., Wu, Q., Wang, M., Zhu, D., Ma, Y., 2017. Regio- and stereoselective reduction of 17-oxosteroids to 17 β -hydroxysteroids by a yeast strain *Zygowilliopsis* sp. WY7905. *Steroids* 118, 17–24.
- Loida, P.J., Sligar, S.G., 1993. Molecular recognition in cytochrome P450: Mechanism for the control of uncoupling reactions. *Biochemistry (Mosc.)* 32, 11530–11538.
- Lombard, M., Salard, I., Sari, M.-A., Mansuy, D., Buisson, D., 2011. A new cytochrome P450 belonging to the 107L subfamily is responsible for the efficient hydroxylation of the drug terfenadine by *Streptomyces platensis*. *Arch. Biochem. Biophys.* 508, 54–63.
- Lu, A.Y.H., Junk, K.W., Coon, M.J., 1969. Resolution of the cytochrome P450-containing ω -hydroxylation system of liver microsomes into three components. *J. Biol. Chem.* 244, 3714–3721.
- Luo, Y., Li, B.-Z., Liu, D., Zhang, L., Chen, Y., Jia, B., Zeng, B.-X., Zhao, H., Yuan, Y.-J., 2015. Engineered biosynthesis of natural products in heterologous hosts. *Chem Soc Rev* 44, 5265–5290.
- Makris, T.M., Davydov, R., Denisov, I.G., Hoffman, B.M., Sligar, S.G., 2002. Mechanistic enzymology of oxygen activation by the cytochromes P450. *Drug Metab. Rev.* 34, 691–708.
- Martín, L., Prieto, M.A., Cortés, E., García, J., 1995. Cloning and sequencing of the *pac* gene encoding the penicillin G acylase of *Bacillus megaterium* ATCC 14945. *FEMS Microbiol. Lett.* 125, 287–292.
- Matsuoka, T., Miyakoshi, S., Tanzawa, K., Nakahara, K., Hosobuchi, M., Serizawa, N., 1989. Purification and characterization of cytochrome P450_{sca} from *Streptomyces carbophilus*. ML-236B (compactin) induces a cytochrome P450_{sca} in *Streptomyces carbophilus* that hydroxylates ML-236B to pravastatin sodium (CS-514), a tissue-selective inhibitor of 3-hydroxy-3-methylglutaryl-coenzyme-A reductase. *Eur. J. Biochem. FEBS* 184, 707–713.
- McLean, K.J., Hans, M., Meijrink, B., van Scheppingen, W.B., Vollebregt, A., Tee, K.L., van der Laan, J.-M., Leys, D., Munro, A.W., van den Berg, M.A., 2015. Single-step fermentative production of the cholesterol-

- lowering drug pravastatin via reprogramming of *Penicillium chrysogenum*. Proc. Natl. Acad. Sci. U. S. A. 112, 2847–2852.
- Meunier, B., de Visser, S.P., Shaik, S., 2004. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. Chem. Rev. 104, 3947–3980.
- Milhim, M., Gerber, A., Neunzig, J., Hannemann, F., Bernhardt, R., 2016a. A Novel NADPH-dependent flavoprotein reductase from *Bacillus megaterium* acts as an efficient cytochrome P450 reductase. J. Biotechnol. 231, 83–94.
- Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016b. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. J. Biotechnol. 240, 68–75.
- Miura, Y., Fulco, A.J., 1974. (ω -2) Hydroxylation of fatty acids by a soluble system from *Bacillus megaterium*. J. Biol. Chem. 249, 1880–1888.
- Mizutani, M., 2012. Impacts of diversification of cytochrome P450 on plant metabolism. Biol. Pharm. Bull. 35, 824–832.
- Mizutani, M., Ohta, D., 2010. Diversification of P450 genes during land plant evolution. Annu. Rev. Plant Biol. 61, 291–315.
- Mizutani, M., Ward, E., DiMaio, J., Ohta, D., Ryals, J., Sato, R., 1993. Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. Biochem. Biophys. Res. Commun. 190, 875–880.
- Morikawa, T., Mizutani, M., Aoki, N., Watanabe, B., Saga, H., Saito, S., Oikawa, A., Suzuki, H., Sakurai, N., Shibata, D., Wadano, A., Sakata, K., Ohta, D., 2006. Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. Plant Cell 18, 1008–1022.
- Morita, M., Tomita, K., Ishizawa, M., Takagi, K., Kawamura, F., Takahashi, H., Morino, T., 1999. Cloning of oxetanocin A biosynthetic and resistance genes that reside on a plasmid of *Bacillus megaterium* strain NK84-0128. Biosci. Biotechnol. Biochem. 63, 563–566.
- Munro, A.W., Lindsay, J.G., Coggins, J.R., Kelly, S.M., Price, N.C., 1994. Structural and enzymological analysis of the interaction of isolated domains of cytochrome P450 BM3. FEBS Lett. 343, 70–74.
- Nagao, T., Mitamura, T., Wang, X.H., Negoro, S., Yomo, T., Urabe, I., Okada, H., 1992. Cloning, nucleotide sequences, and enzymatic properties of glucose dehydrogenase isozymes from *Bacillus megaterium* IAM1030. J. Bacteriol. 174, 5013–5020.
- Nair, P.C., McKinnon, R.A., Miners, J.O., 2016. Cytochrome P450 structure-function: insights from molecular dynamics simulations. Drug Metab. Rev. 48, 434–452.
- Nebert, D.W., Adesnik, M., Coon, M.J., Estabrook, R.W., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., Kemper, B., Levin, W., 1987. The P450 gene superfamily: recommended nomenclature. DNA Mary Ann Liebert Inc 6, 1–11.
- Nebert, D.W., Wikvall, K., Miller, W.L., 2013. Human cytochromes P450 in health and disease. Phil Trans R Soc B 368, 20120431. doi: 10.1098/rstb.2012.0431.
- Nelson, D.R., 2009. The cytochrome P450 homepage. Hum. Genomics 4, 59–65.
- Okamoto, S., Yu, F., Harada, H., Okajima, T., Hattan, J., Misawa, N., Utsumi, R., 2011. A short-chain dehydrogenase involved in terpene metabolism from Zingiber zerumbet. FEBS J. 278, 2892–2900.

- Oliver, C.F., Modi, S., Sutcliffe, M.J., Primrose, W.U., Lian, L.-Y., Roberts, G.C.K., 1997. A Single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regioselectivity of hydroxylation. *biochemistry (Mosc.)* 36, 1567–1572.
- Olteanu, H., Banerjee, R., 2001. Human methionine synthase reductase, a soluble P450 reductase-like dual flavoprotein, is sufficient for NADPH-dependent methionine synthase activation. *J. Biol. Chem.* 276, 35558–35563.
- Omura, T., Sanders, E., Estabrook, R.W., Cooper, D.Y., Rosenthal, O., 1966. Isolation from adrenal cortex of a nonheme iron protein and a flavoprotein functional as a reduced triphosphopyridine nucleotide-cytochrome P450 reductase. *Arch. Biochem. Biophys.* 117, 660–673.
- Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. II. solubilization, purification, and properties. *J. Biol. Chem.* 239, 2379–2385.
- Omura, T., Sato, R., 1962. A New cytochrome in liver microsomes. *J. Biol. Chem.* 237, 1375–1376.
- Ostrowski, J., Barber, M.J., Rueger, D.C., Miller, B.E., Siegel, L.M., Kredich, N.M., 1989. Characterization of the flavoprotein moieties of NADPH-sulfite reductase from *Salmonella typhimurium* and *Escherichia coli*. Physicochemical and catalytic properties, amino acid sequence deduced from DNA sequence of *cysJ*, and comparison with NADPH-cytochrome P450 reductase. *J. Biol. Chem.* 264, 15796–15808.
- Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S., Newman, J.D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532.
- Paine, M.J.I., Garner, A.P., Powell, D., Sibbald, J., Sales, M., Pratt, N., Smith, T., Tew, D.G., Wolf, C.R., 2000. Cloning and characterization of a novel human dual flavin reductase. *J. Biol. Chem.* 275, 1471–1478.
- Pervaiz, I., Ahmad, S., Madni, M.A., Ahmad, H., Khaliq, F.H., 2013. Microbial biotransformation: a tool for drug designing. *Appl. Biochem. Microbiol.* 49, 437–450.
- Poulos, T.L., Finzel, B.C., Howard, A.J., 1987. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* 195, 687–700.
- Raux, E., Lanois, A., Warren, M.J., Rambach, A., Thermes, C., 1998. Cobalamin (vitamin B12) biosynthesis: identification and characterization of a *Bacillus megaterium cobI* operon. *Biochem. J.* 335, 159–166.
- 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, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R., Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943.
- Rosenberg, W., Ireland, A., Jewell, D.P., 1990. High-dose methylprednisolone in the treatment of active ulcerative colitis: *J. Clin. Gastroenterol.* 12, 40–41.
- Saag, K.G., 2002. Glucocorticoid use in rheumatoid arthritis. *Curr. Rheumatol. Rep.* 4, 218–225.
- Sakaki, T., Sugimoto, H., Hayashi, K., Yasuda, K., Munetsuna, E., Kamakura, M., Ikushiro, S., Shiro, Y., 2011. Bioconversion of vitamin D to its active form by bacterial or mammalian cytochrome P450. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1814, 249–256.

- Schiffer, L., Anderko, S., Hobler, A., Hannemann, F., Kagawa, N., Bernhardt, R., 2015. A recombinant CYP11B1 dependent *Escherichia coli* biocatalyst for selective cortisol production and optimization towards a preparative scale. *Microb. Cell Factories* 14, 25. doi: 10.1186/s12934-015-0209-5.
- Schmitz, D., Zapp, J., Bernhardt, R., 2012. Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*. *FEBS J.* 279, 1663–1674.
- Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., Werck-Reichhart, D., 2001. CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.* 276, 36566–36574.
- Schulz, S., Girhard, M., Gaßmeyer, S.K., Jäger, V.D., Schwarze, D., Vogel, A., Urlacher, V.B., 2015. Selective enzymatic synthesis of the grapefruit flavor (+)-nootkatone. *ChemCatChem* 7, 601–604.
- Shiota, H., Nitta, K., Naito, T., Mimura, Y., Maruyama, T., 1996. Clinical evaluation of carbocyclic oxetanocin G eyedrops in the treatment of herpes simplex corneal ulcers. *Br. J. Ophthalmol.* 80, 413–415.
- Shivakumar, A.G., Katz, L., Cohen, L.B., Ginsburgh, C.L., Paul, S.L., Vanags, R.I., 1998. Expression of heterologous proteins in *Bacillus megaterium* utilizing sporulation promoters of *Bacillus subtilis*. US5726042 A (Patent).
- Sirim, D., Widmann, M., Wagner, F., Pleiss, J., 2010. Prediction and analysis of the modular structure of cytochrome P450 monooxygenases. *BMC Struct. Biol.* 10, 34. doi: 10.1186/1472-6807-10-34.
- Solieri, L., Giudici, P., 2009. *Vinegars of the world*. Springer, Milan; New York.
- Sono, M., Roach, M.P., Coulter, E.D., Dawson, J.H., 1996. Heme-containing oxygenases. *Chem. Rev.* 96, 2841–2888.
- Sowden, R.J., Yasmin, S., Rees, N.H., Bell, S.G., Wong, L.-L., 2005. Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450 BM3. *Org. Biomol. Chem.* 3, 57–64.
- Stuehr, D.J., 1997. Structure-function aspects in the nitric oxide synthases. *Annu. Rev. Pharmacol. Toxicol.* 37, 339–359.
- Suga, K.-I., Shiba, Y., Sorai, T., Shioya, S., Ishimura, F., 1990. Reaction kinetics and mechanism of immobilized penicillin acylase from *Bacillus megaterium*. *Ann. N. Y. Acad. Sci.* 613, 808–815.
- Takasaki, Y., 1989. Novel Maltose-producing Amylase from *Bacillus megaterium* G-2. *Agric. Biol. Chem.* 53, 341–347.
- Thrower, B.W., 2009. Relapse management in multiple sclerosis. *The Neurologist* 15, 1–5.
- Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M., DellaPenna, D., 2004. The *Arabidopsis* LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid ϵ -ring hydroxylation activity. *Proc. Natl. Acad. Sci.* 101, 402–407.
- Tseng, C.K.H., Marquez, V.E., Milne, G.W.A., Wysocki, R.J., Mitsuya, H., Shirasaki, T., Driscoll, J.S., 1991. A ring-enlarged oxetanocin A analog as an inhibitor HIV infectivity. *J. Med. Chem.* 34, 343–349.
- Tsotsou, G.E., Cass, A.E.G., Gilardi, G., 2002. High throughput assay for cytochrome P450 BM3 for screening libraries of substrates and combinatorial mutants. *Biosens. Bioelectron.* 17, 119–131.
- Vary, P.S., 1994. Prime time for *Bacillus megaterium*. *Microbiology* 140, 1001–1013. Vihinen, M., Mantsiila, P., 1989. Microbial amylolytic enzyme. *Crit. Rev. Biochem. Mol. Biol.* 24, 329–418.
- Wang, M., Roberts, D.L., Paschke, R., Shea, T.M., Masters, B.S.S., Kim, J.-J.P., 1997. Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci.* 94, 8411–8416.

- Werck-Reichhart, D., Feyereisen, R., 2000. Cytochromes P450: a success story. *Genome Biol.* 1, 3003.1-3003.9.
- Westfall, P.J., Pitera, D.J., Lenihan, J.R., Eng, D., Woolard, F.X., Regentin, R., Horning, T., Tsuruta, H., Melis, D.J., Owens, A., Fickes, S., Diola, D., Benjamin, K.R., Keasling, J.D., Leavell, M.D., McPhee, D.J., Renninger, N.S., Newman, J.D., Paddon, C.J., 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinin acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U. S. A.* 109, E111-118.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450(BM3) (CYP102A1): connecting the dots. *Chem. Soc. Rev.* 41, 1218–1260.
- Wong, L.-L., Westlake, A.C.G., Nickerson, D.P., 1997. Protein engineering of cytochrome P450cam, in: Hill, H.A.O., Sadler, P.J., Thomson, A.J. (Eds.), *metal sites in proteins and models, structure and bonding.* Springer Berlin Heidelberg, pp. 175–207.
- World Health Organization, 2015. 19th WHO Model List of Essential Medicines (April 2015). www.who.int/medicines/publications/essentialmedicines/en/.
- Wriessnegger, T., Augustin, P., Engleder, M., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H., Pichler, H., 2014. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. *Metab. Eng.* 24, 18–29.
- Xu, L.-H., Fushinobu, S., Takamatsu, S., Wakagi, T., Ikeda, H., Shoun, H., 2010. Regio- and stereospecificity of filipin hydroxylation sites revealed by crystal structures of cytochrome P450 105P1 and 105D6 from *Streptomyces avermitilis*. *J. Biol. Chem.* 285, 16844–16853.
- Yeom, H., Sligar, S.G., Li, H., Poulos, T.L., Fulco, A.J., 1995. The role of Thr268 in oxygen activation of cytochrome P450 BM3. *Biochemistry (Mosc.)* 34, 14733–14740.
- Zanger, U.M., Schwab, M., 2013. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* 138, 103–141.
- Zhang, M.M., Wang, Y., Ang, E.L., Zhao, H., 2016. Engineering microbial hosts for production of bacterial natural products. *Nat Prod Rep* 33, 963–987.
- Zhou, S.-F., Liu, J.-P., Chowbay, B., 2009. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab. Rev.* 41, 89–295.
- Zhu, Y., Nomura, T., Xu, Y., Zhang, Y., Peng, Y., Mao, B., Hanada, A., Zhou, H., Wang, R., Li, P., Zhu, X., Mander, L.N., Kamiya, Y., Yamaguchi, S., He, Z., 2006. Elongated uppermost internode encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* 18, 442–456.